

1 VENKAT KONDA
2 6278 Grand Oak Way
3 San Jose, California 95135
4 Telephone: (408) 472-3273
5 Email: vkonda@gmail.com

6 Plaintiff *Pro se*

**Electronically Filed
by Superior Court of CA,
County of Santa Clara,
on 3/22/2021 11:37 PM
Reviewed By: F. Miller
Case #19CV345846
Envelope: 6087289**

7 **SUPERIOR COURT OF CALIFORNIA - COUNTY OF SANTA CLARA**
8 **UNLIMITED JURISDICTION**

9 **VENKAT KONDA, Ph.D., an individual,**

10 **Plaintiff,**

11 **v.**

12 **DEJAN MARKOVIC, Ph.D., an individual;**
13 **CHENG C. WANG, Ph.D., an individual;**
14 **FLEX LOGIX TECHNOLOGIES, INC., a**
15 **Delaware Corporation; THE REGENTS OF**
16 **THE UNIVERSITY OF CALIFORNIA;**
17 **GEOFFREY TATE, an individual; PIERRE**
18 **LAMOND, an individual; PETER HEBERT,**
19 **an individual; LESLIE M. LACKMAN, Ph.D.,**
20 **an individual; and DOES 1-20, inclusive,**

21 **Defendants.**

CASE NO. 19CV345846

**DECLARATION OF VIPIN CHAUDHARY,
Ph.D. IN SUPPORT OF PLAINTIFF'S
FOURTH AMENDED COMPLAINT**

**Department: 2
Before: Honorable Drew C. Takaichi**

**Date Complaint Filed: April 3, 2019
Trial Date: None**

22
23 I, Vipin Chaudhary, Ph.D., do hereby declare as follows:

24 1. I make the statements herein based on my personal knowledge and I could and would
25 competently testify thereto if called as a witness.

26 2. My current *curriculum vitae* is attached hereto as Exhibit K.

27 3. I earned a Bachelor Degree (Hons.) in Computer Science and Engineering from the
28 Indian Institute of Technology, Kharagpur, India in 1986, and the MS degree in Computer Science in

1 1989 and the Ph.D. degree in Electrical and Computer Engineering in 1992, both from The University
2 of Texas at Austin.

3 4. Currently I am the Endowed Kranzusch Professor and Inaugural Chair, Department of
4 Computer and Data Sciences, Case School of Engineering, Case Western Reserve University,
5 Cleveland, Ohio. Prior to this position, I was a SUNY Empire Innovation Professor between 2011 and
6 2020 and SUNY Empire Innovation Associate Professor between 2006 and 2011, Computer Science
7 and Engineering at University at Buffalo, The State University of New York. Between June 2016 and
8 June 2020, I was also a Program Director at the Office of Advanced Cyber Infrastructure, Directorate
9 for Computer and Information Science and Engineering, National Science Foundation, Alexandria,
10 Virginia. Prior to the University at Buffalo, I was an Associate Professor, Department of Computer
11 Science, and an Associate Professor, Department of Electrical and Computer Engineering at Wayne
12 State University, Detroit, Michigan between 1998 and 2006, and an Assistant Professor, Department of
13 Electrical and Computer Engineering, Wayne State University between 1992 and 1998.

14 5. I have been active in the field of integrated circuits and interconnection networks for
15 over 30 years, since my Ph.D. Dissertation, awarded in 1992, in the area of parallel and distributed
16 computing where interconnection networks is a major part of my dissertation.

17 6. I have received numerous awards for my work, including the 2019 National Science
18 Foundation Director's Superior Accomplishment Award for my contributions where as a Program
19 Director I co-led the National Strategic Computing Initiative from NSF for the United States and in the
20 working groups of the Quantum Leap Initiative, National Quantum Initiative, National Artificial
21 Intelligence Research Institutes, Cyber, and the I-Corps Program. The I-Corps program is now part of
22 "The American Innovation and Competitiveness Act" that enables commercialization of research and
23 venture startups. The U.S. National Strategic Computing Initiative incorporates many aspects of
24 interconnection networks to make large computer systems.

25 7. I was co-founder of several startups, including as a Senior Director of Advanced
26 Development at Cradle Technologies, Inc., where I was responsible for advanced programming tools
27 development for multi-processor chips where interconnection networks is a key component. In Scalable
28 Informatics, we designed and built some of the highest performance storage and analytics systems. As

1 the CEO of Computational Research Laboratories, we built the fourth largest supercomputer system in
2 the world with a unique interconnection network that enabled fast performance at a fraction of the cost.
3 This company was sold to Tata Consulting Services. Prior to this, I was the Chief Architect at Corio,
4 that is known as one of companies that really started the Software-as-a-Service revolution and had a
5 successful IPO in 2000.

6 8. I served as associate Guest Editor, Special Issue of IEICE Transactions on Information
7 and Systems on Hardware/Software Support for High Performance Scientific and Engineering
8 Computing, July 2004. I served as Conference or Symposium Chair in six of the relevant conferences
9 and as workshop chair in more than twenty workshops in the area of interconnection networks. I served
10 as program committee member in more than forty conferences. I taught and created numerous
11 undergraduate and graduate courses where integrated circuits and interconnection networks is an
12 integral part of the subject matter.

13 9. I have supervised numerous doctoral dissertations and Master's theses. I have
14 contributed to numerous book chapters and published numerous papers in refereed journal papers,
15 refereed conference papers and refereed workshops related to integrated circuits and interconnection
16 networks. I have given numerous invited talks at Academic Institutions, Industries, Research
17 Laboratories, conferences and workshops related to integrated circuits and interconnection networks.

18 10. I have known Venkat Konda, Ph.D. (hereinafter referred to as "Dr. Konda") since 1991.

19 11. Dr. Konda contacted me regarding the above-captioned lawsuit.

20 12. Dr. Konda requested that I review a few Konda Technologies (hereinafter referred to as
21 "Konda Tech") Documents listed below identifying Konda Tech's confidential information that is not
22 trade secret information. I agreed to review Konda Tech's documents listed below to identify Konda
23 Tech's confidential information that is non-trade-secret information, and provide the results of my
24 review in this declaration.

25 13. My understanding is that a trade secret is defined in California Civil Code Section
26 3426.1(d) as:

27 "Trade secret" means information, including a formula, pattern, compilation,
28 program, device, method, technique, or process, that:

- 1 (1) Derives independent economic value, actual or potential, from not being
2 generally known to the public or to other persons who can obtain economic value
3 from its disclosure or use; and
4 (2) Is the subject of efforts that are reasonable under the circumstances to
5 maintain its secrecy.”

6 In this declaration, I express my opinion with respect to the subjects covered in subparagraph
7 (1) of the above trade secret definition.

8 14. My understanding is confidential information that is not trade secret information means
9 “confidential information including knowledge about the FPGA industry including contemporary
10 industry analytics and trend analyses, business practices, disadvantages of the competition, advantages
11 vis-à-vis the competition, customer procurement, relationship building, and management, and business
12 successes.” This information can be compiled by sweat of the brow and is protectable by being
13 maintained confidential by the compiler of the information and not published.

14 15. Dr. Konda requested me to review the following documents: (1) Konda Tech’s
15 confidential Business Presentation to Defendant Markovic on October 7, 2009 in confidence (*See*,
16 Exhibit 13 in Fourth Amended Complaint “FAC”, hereinafter referred to as “Konda Tech 2009
17 Presentation”); (2) June 23, 2010 DARPA funding proposal (*See*, Exhibit 14 in FAC) and August 6,
18 2010 DARPA funding proposal (*See*, Exhibit 15 in FAC) (hereinafter collectively referred to as “Two
19 Confidential DARPA Proposals”); (3) Konda Tech WIPO WO 2008109756 A1 published on December
20 9, 2008 (*See*, Exhibit A attached hereto), Konda Tech WIPO WO 2008147926 A1 published on
21 December 4, 2008 (*See*, Exhibit C attached hereto), Konda Tech WIPO WO 2008147927 A1 published
22 on December 4, 2008 (*See*, Exhibit E attached hereto), and Konda Tech WIPO WO 2008147928 A1
23 published on December 4, 2008 (*See*, Exhibit G attached hereto) (hereinafter referred to as collectively
24 “2008 Konda Publications”); (4) Konda Tech US patent application US 2010/0135286 A1 publication
25 (*See*, Exhibit B attached hereto), Konda Tech US patent application US 2010/0172349 A1 publication
26 (*See*, Exhibit D attached hereto), Konda Tech US patent application US 2011/0044329 A1 publication
27 (*See*, Exhibit F attached hereto), and Konda Tech US patent application US 2011/0037498 A1
28 publication (*See*, Exhibit H attached hereto) (hereinafter referred to as collectively “pre-2010 Konda

1 Publications”); (5) Konda Tech WIPO WO 2011047368 A2 published on April 21, 2011 (hereinafter
2 referred to as “2011 Konda Publication”) (*See*, Exhibit I attached hereto); (6) Konda Tech US patent
3 application US 2012/0269190 A1 publication (hereinafter referred to as “2012 Konda Publication”)
4 (*See*, Exhibit J attached hereto); (7) Defendants Markovic and Wang’s paper presented in June 2011
5 (submitted in January 2011) at the 2011 VLSI Circuits Symposium titled “A 1.1 GOPS/mQ FPGA
6 Chip with Hierarchical Interconnect Fabric” unbeknownst to Dr. Konda and without his authorization,
7 (hereinafter referred to as “2011 VLSI Paper”) (*See*, Exhibit 20 in FAC); (8) Defendant Wang’s PhD
8 Dissertation presented in 2013 titled “Building Efficient, Reconfigurable Hardware using Hierarchical
9 Interconnects” unbeknownst to Dr. Konda and without his authorization (hereinafter referred to as
10 “Wang’s 2013 PhD Dissertation”) (*See*, Exhibit 22 in FAC); and (9) Defendants Markovic and Wang
11 presented paper titled “A Multi-Granularity FPGA with Hierarchical Interconnects for Efficient and
12 Flexible Mobile Computing” to the 2014 International Solid State Circuits Conference (hereinafter
13 referred to as “2014 ISSCC Paper”) (*See*, Exhibit 24 in FAC).

14 16. I was informed by Dr. Konda that the contents of the “Konda Tech 2009 Presentation,”
15 which is marked “confidential,” is confidential information even today.

16 17. I was asked to review if the Konda Tech 2009 Presentation and the Two Confidential
17 DARPA Proposals contain additional confidential information, particularly non-trade-secret
18 information, that is not disclosed in the 2008 Konda Publications. The objective was to determine the
19 Konda Tech confidential information, particularly non-trade-secret information, that was disclosed to
20 Markovic by Dr. Konda in confidence.

21 18. After my review, in my opinion, the Konda Tech 2009 Presentation and the Two
22 Confidential DARPA Proposals contain confidential information that is non-trade-secret information.

23 19. Accordingly in my opinion, the Konda Tech 2009 Presentation and the Two
24 Confidential DARPA Proposals contain confidential information that is non-trade-secret information
25 and that was not disclosed in the 2008 Konda Publications.

26 20. Dr. Konda also requested me to review if confidential information in the Konda Tech
27 2009 Presentation and the Two Confidential DARPA Proposals that is not disclosed in the 2008 Konda
28 Publications is contained in the 2011 VLSI Paper.

1 21. After my review, in my opinion, the 2011 VLSI Paper contains confidential information,
2 particularly non-trade-secret information, in the Konda Tech 2009 Presentation and the Two
3 Confidential DARPA Proposals that was not disclosed in the 2008 Konda Publications.

4 22. Accordingly in my opinion, at the time of submission, the 2011 VLSI Paper contains Dr.
5 Konda's confidential information including non-trade-secret information.

6 23. Dr. Konda further requested me to review if confidential non-trade-secret information in
7 the Konda Tech 2009 Presentation and the Two Confidential DARPA Proposals that is not disclosed in
8 the 2008 Konda Publications is contained in Wang's 2013 PhD Dissertation.

9 24. After my review, in my opinion, Wang's 2013 PhD Dissertation contains confidential
10 non-trade-secret information identified in the Konda Tech 2009 Presentation and the Two Confidential
11 DARPA Proposals that is not disclosed in the 2008 Konda Publications.

12 25. Accordingly in my opinion, at the time of submission, Wang's 2013 PhD Dissertation
13 contains Dr. Konda's confidential non-trade-secret information.

14 26. Additionally, Dr. Konda requested me to review if confidential information identified in
15 the Konda Tech 2009 Presentation and the Two Confidential DARPA Proposals that is not disclosed in
16 the 2008 Konda Publications.

17 27. After my review, in my opinion, the 2014 ISSCC Paper contains confidential non-trade-
18 secret information identified in the Konda Tech 2009 Presentation and the Two Confidential DARPA
19 Proposals that is not disclosed in the 2008 Konda Publications.

20 28. Accordingly in my opinion, at the time of submission, the 2014 ISSCC Paper contains
21 Dr. Konda's confidential non-trade-secret information.

22 29. Dr. Konda also requested me to review the texts and diagrams of the 2008 Konda
23 Publications and pre-2010 Konda Publications.

24 30. Based on my analysis, the 2011 VLSI Paper contains substantial portions of the texts
25 and related diagrams of the 2008 Konda Publications and pre-2010 Konda Publications without
26 attribution to Dr. Konda or Konda Tech.

27 31. Dr. Konda additionally requested me to review the texts and diagrams of the 2011
28 Konda Publication and 2012 Konda Publication.

EXHIBIT F



US 20110044329A1

(19) **United States**
 (12) **Patent Application Publication**
Konda

(10) **Pub. No.: US 2011/0044329 A1**
 (43) **Pub. Date: Feb. 24, 2011**

(54) **FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE NETWORKS**

(52) **U.S. Cl. 370/388**

(57) **ABSTRACT**

(76) **Inventor: Venkat Konda, San Jose, CA (US)**

Correspondence Address:
Konda Technologies, Inc
6278 GRAND OAK WAY
SAN JOSE, CA 95135 (US)

A generalized multi-link multi-stage network comprising $(2 \times \log_d N) - 1$ stages is operated in strictly nonblocking manner for unicast includes an input stage having N/d switches with each of them having d inlet links and $2 \times d$ outgoing links connecting to second stage switches, an output stage having N/d switches with each of them having d outlet links and $2 \times d$ incoming links connecting from switches in the penultimate stage. The network also has $(2 \times \log_d N) - 3$ middle stages with each middle stage having N/d switches, and each switch in the middle stage has $2 \times d$ incoming links connecting from the switches in its immediate preceding stage, and $2 \times d$ outgoing links connecting to the switches in its immediate succeeding stage. Also the same generalized multi-link multi-stage network is operated in rearrangeably nonblocking manner for arbitrary fan-out multicast and each multicast connection is set up by use of at most two outgoing links from the input stage switch.

(21) **Appl. No.: 12/601,274**

(22) **PCT Filed: May 22, 2008**

(86) **PCT No.: PCT/US08/64604**

§ 371 (c)(1),
 (2), (4) **Date: May 31, 2010**

Related U.S. Application Data

(60) Provisional application No. 60/940,389, filed on May 25, 2007, provisional application No. 60/940,391, filed on May 25, 2007, provisional application No. 60/940,392, filed on May 25, 2007.

Publication Classification

(51) **Int. Cl. H04L 12/50 (2006.01)**

A generalized multi-link multi-stage network comprising $(2 \times \log_d N) - 1$ stages is operated in strictly nonblocking manner for multicast includes an input stage having N/d switches with each of them having d inlet links and $3 \times d$ outgoing links connecting to second stage switches, an output stage having N/d switches with each of them having d outlet links and $3 \times d$ incoming links connecting from switches in the penultimate stage. The network also has $(2 \times \log_d N) - 3$ middle stages with each middle stage having N/d switches, and each switch in the middle stage has $3 \times d$ incoming links connecting from the switches in its immediate preceding stage, and $3 \times d$ outgoing links connecting to the switches in its immediate succeeding stage.

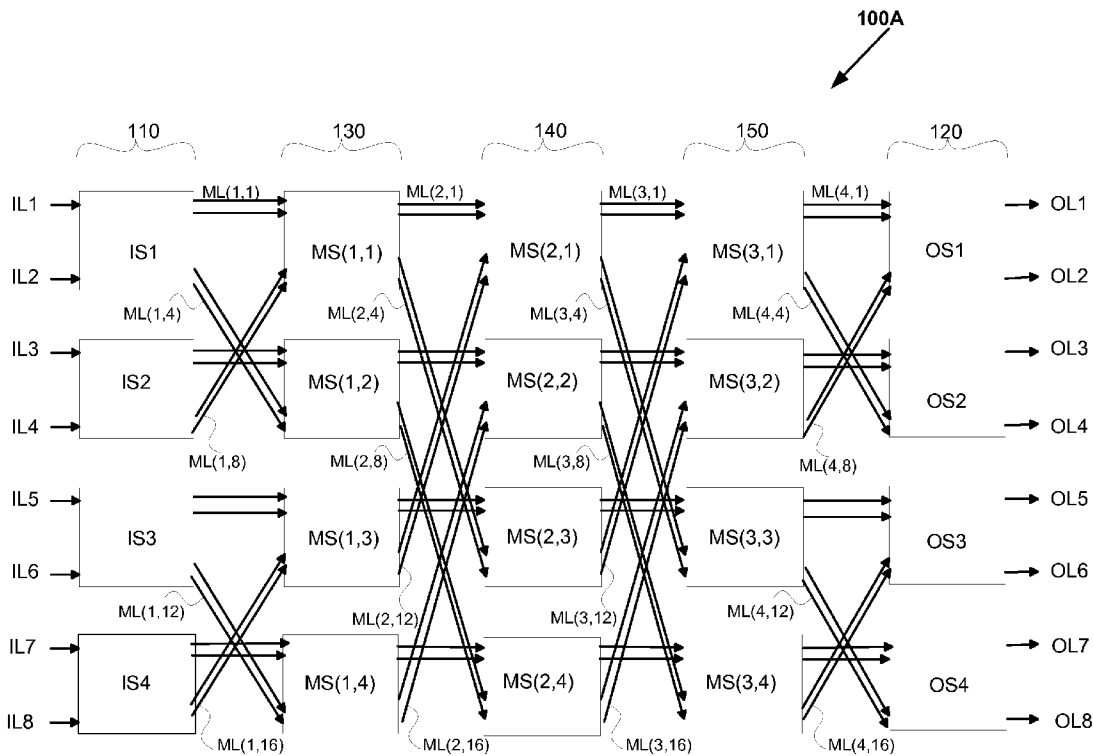


FIG. 1A

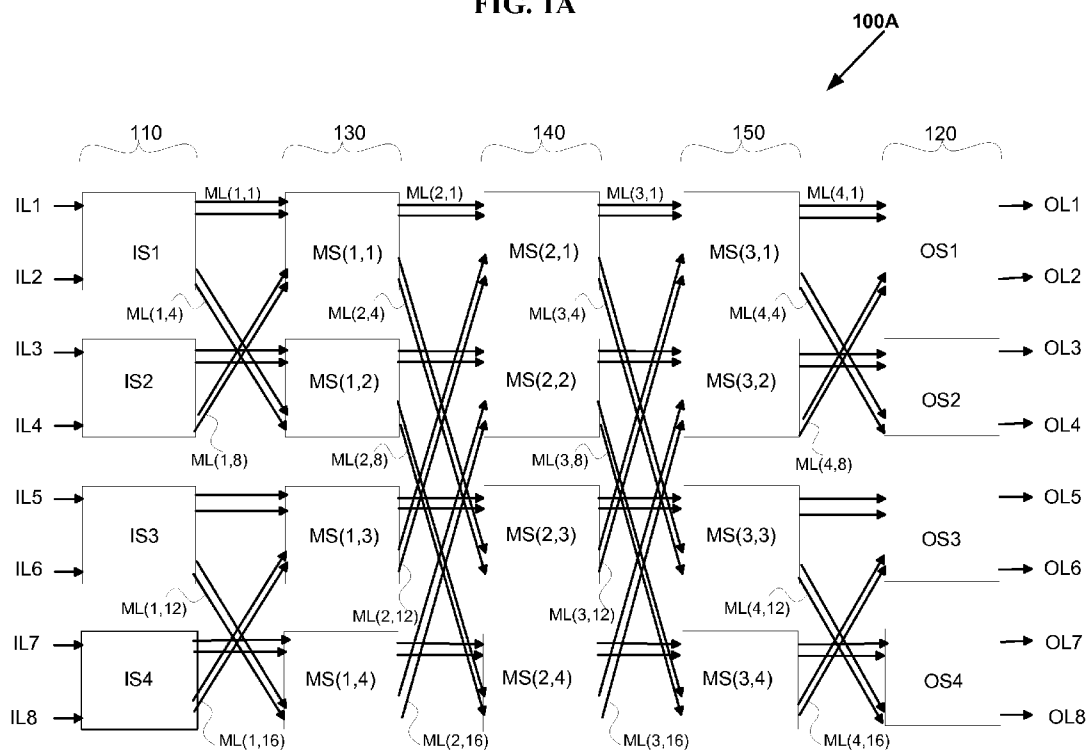


FIG. 1B

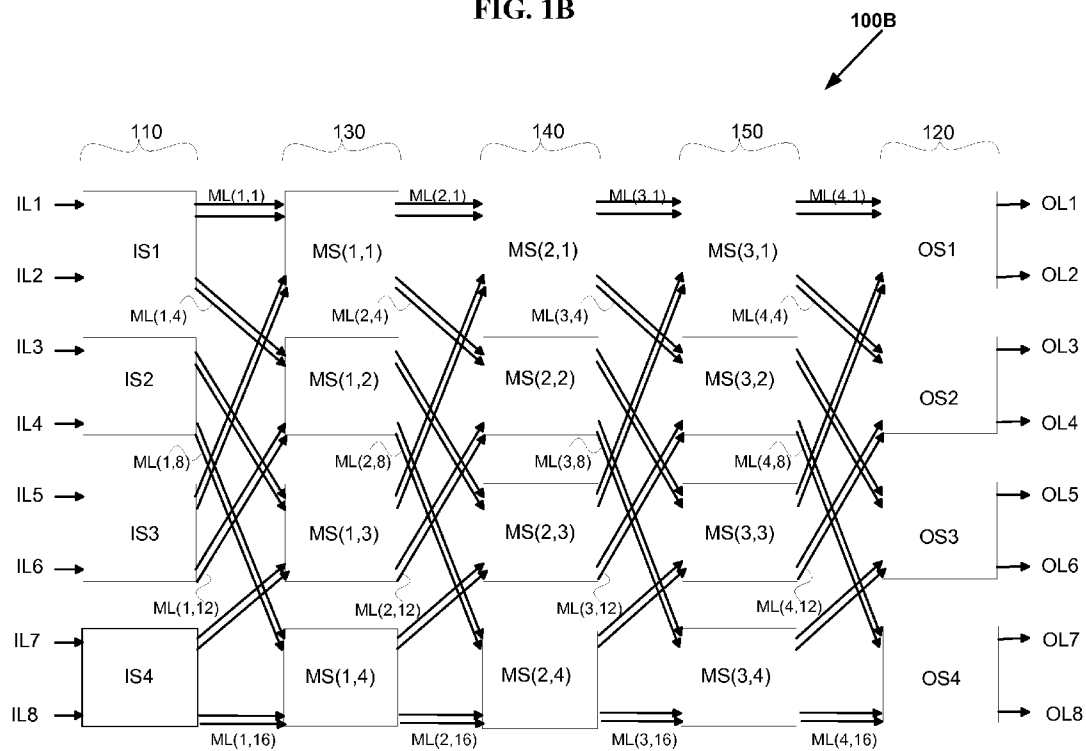


FIG. 1C

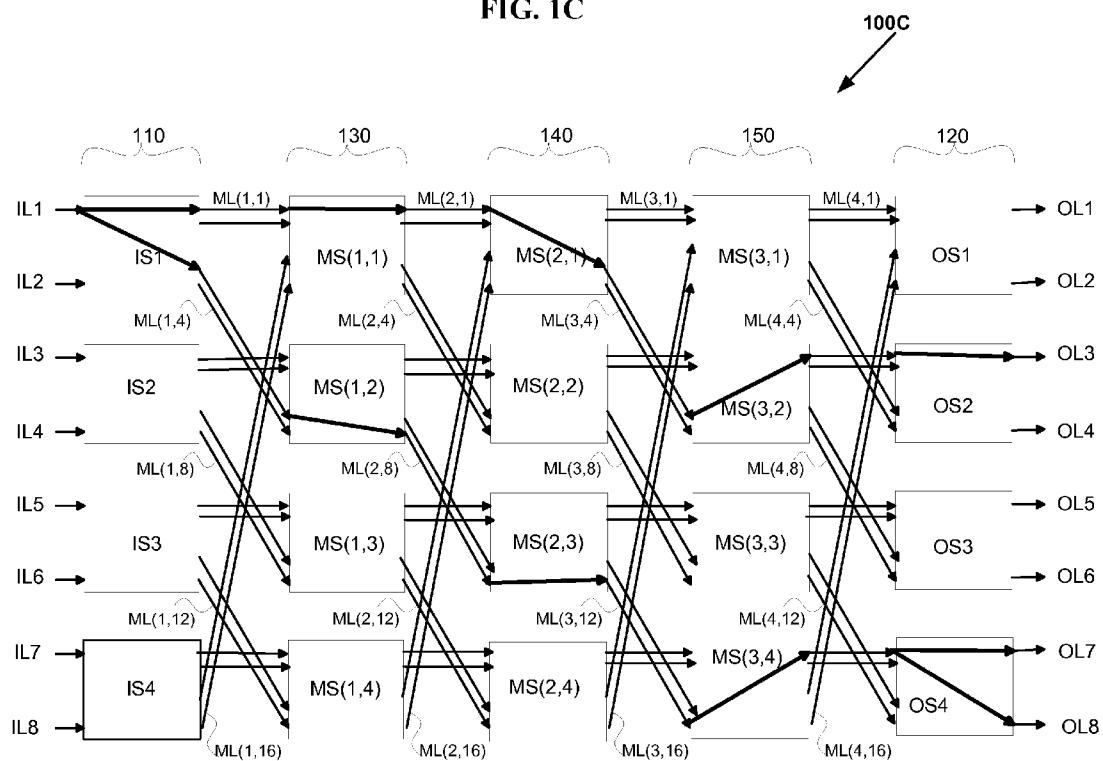


FIG. 1D

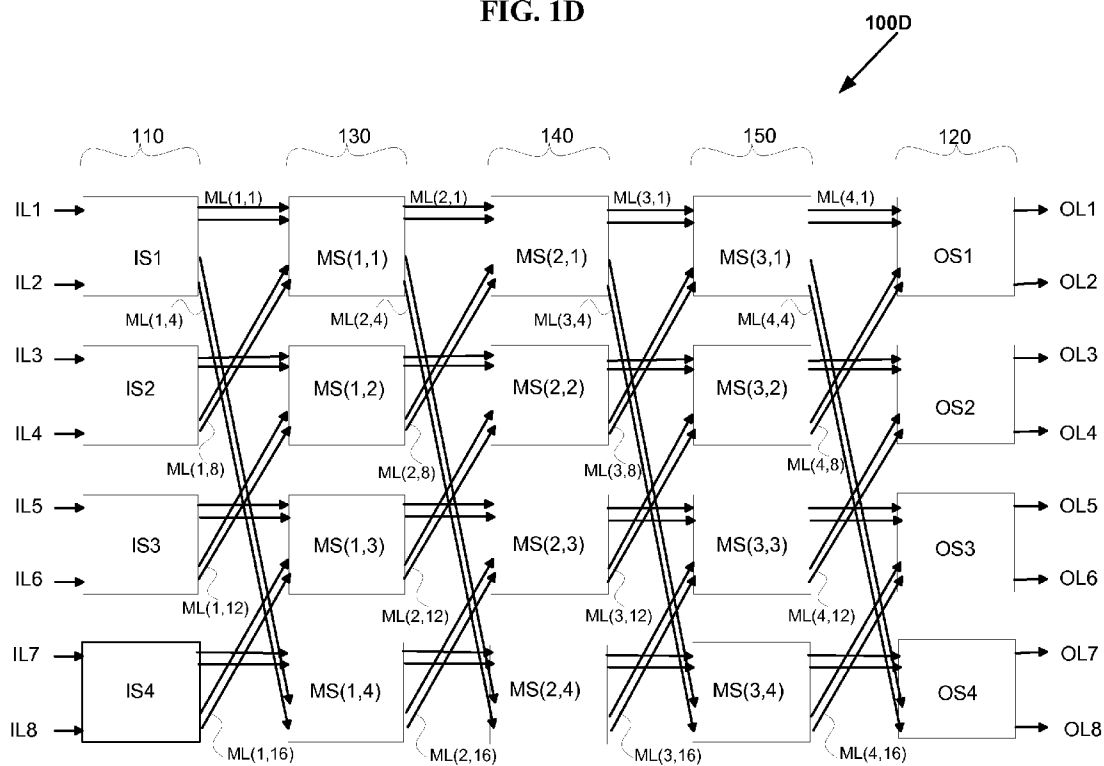


FIG. 1E

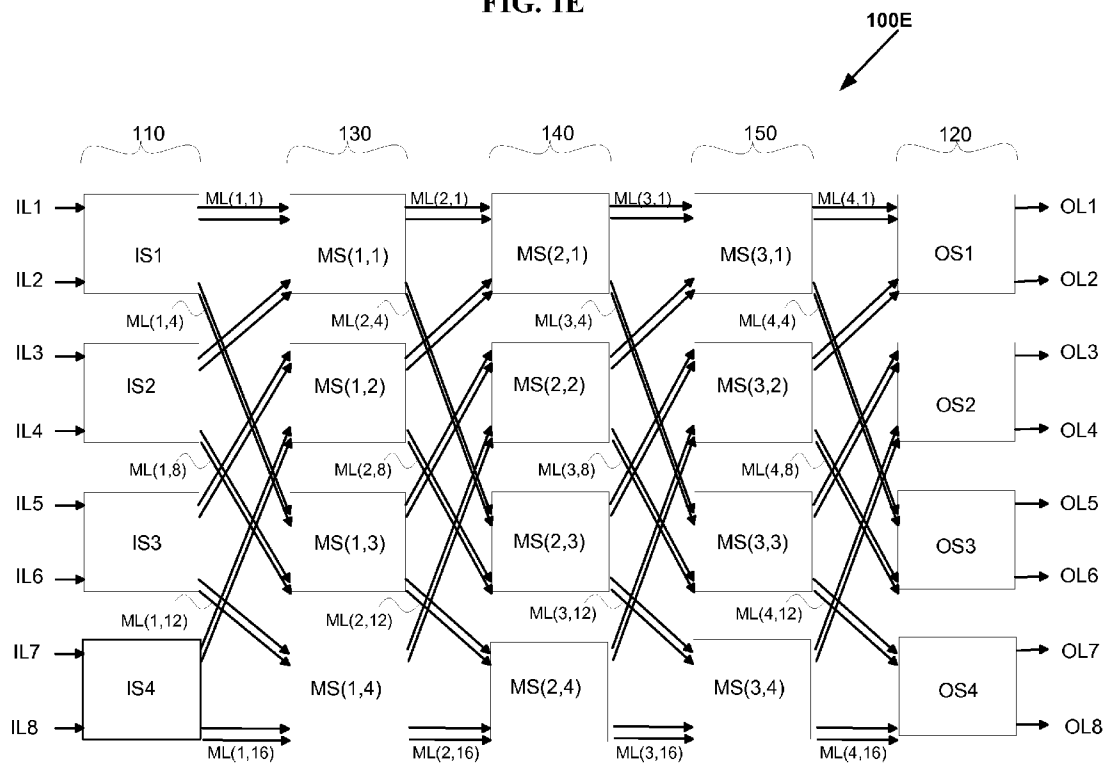


FIG. 1F

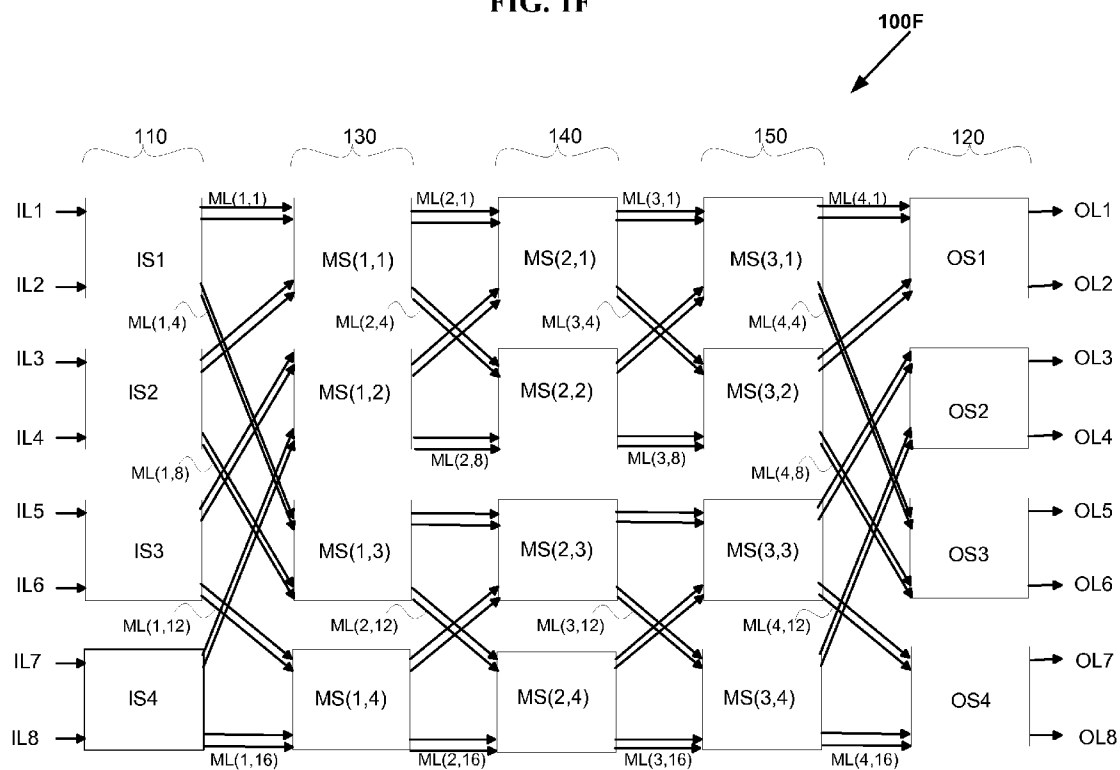


FIG. 1G

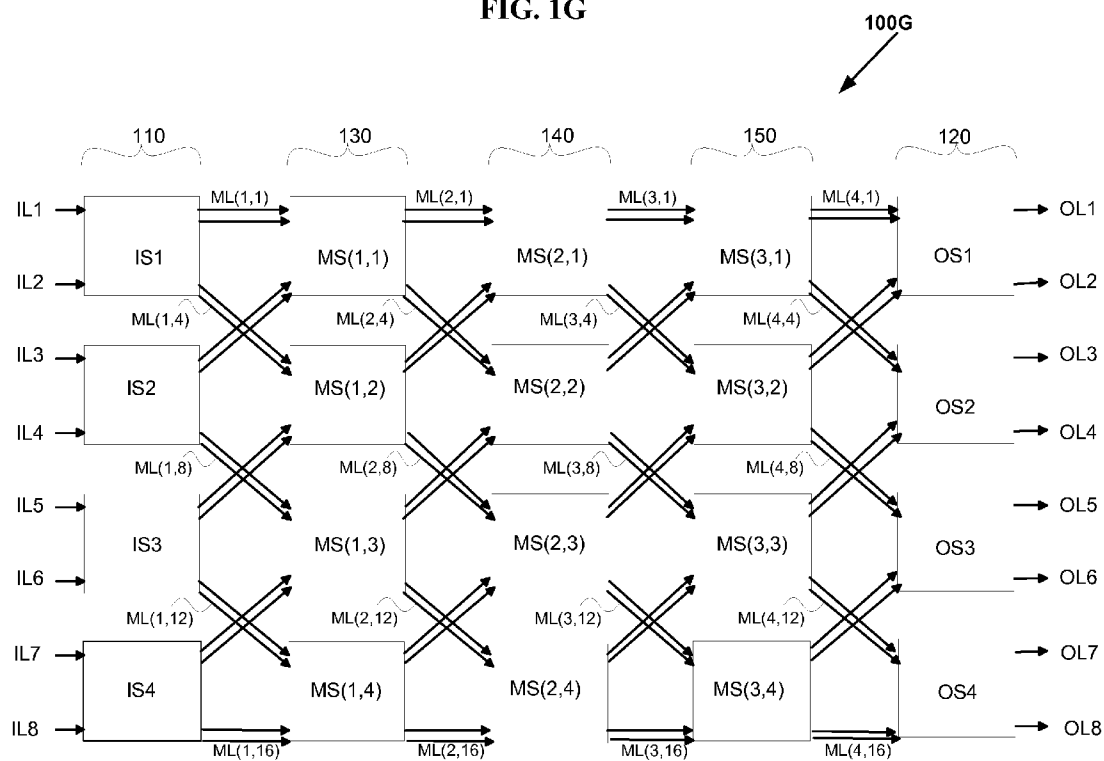


FIG. 1H

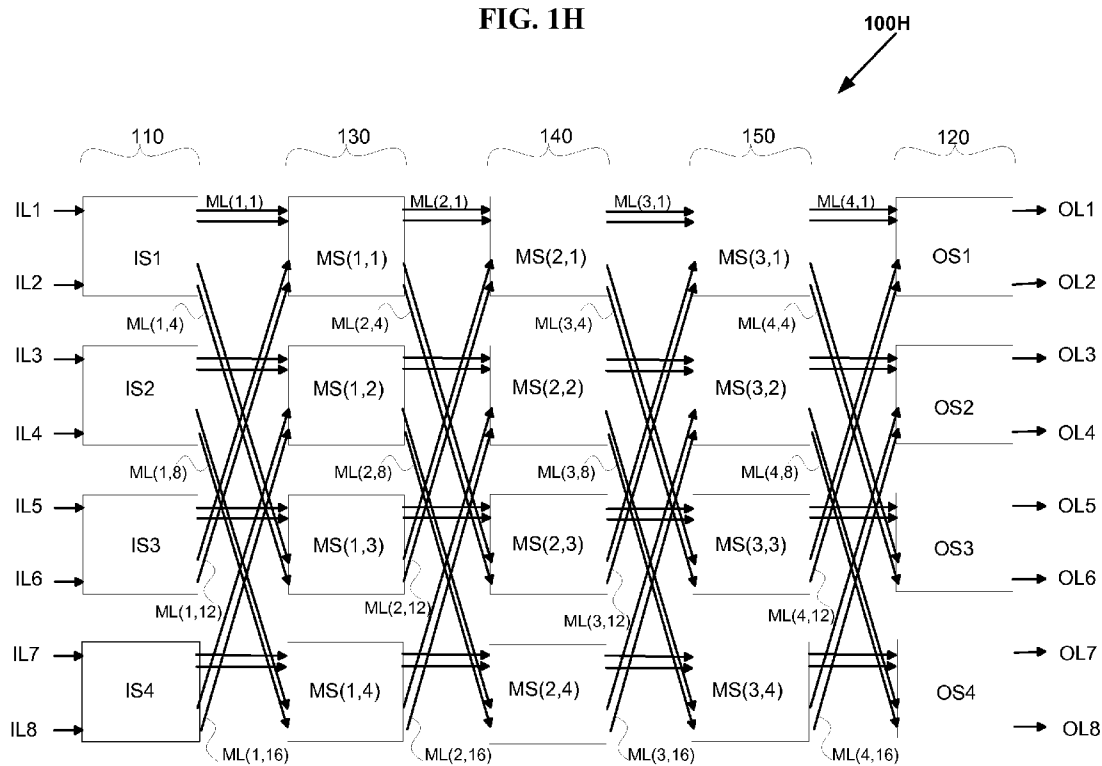


FIG. 11

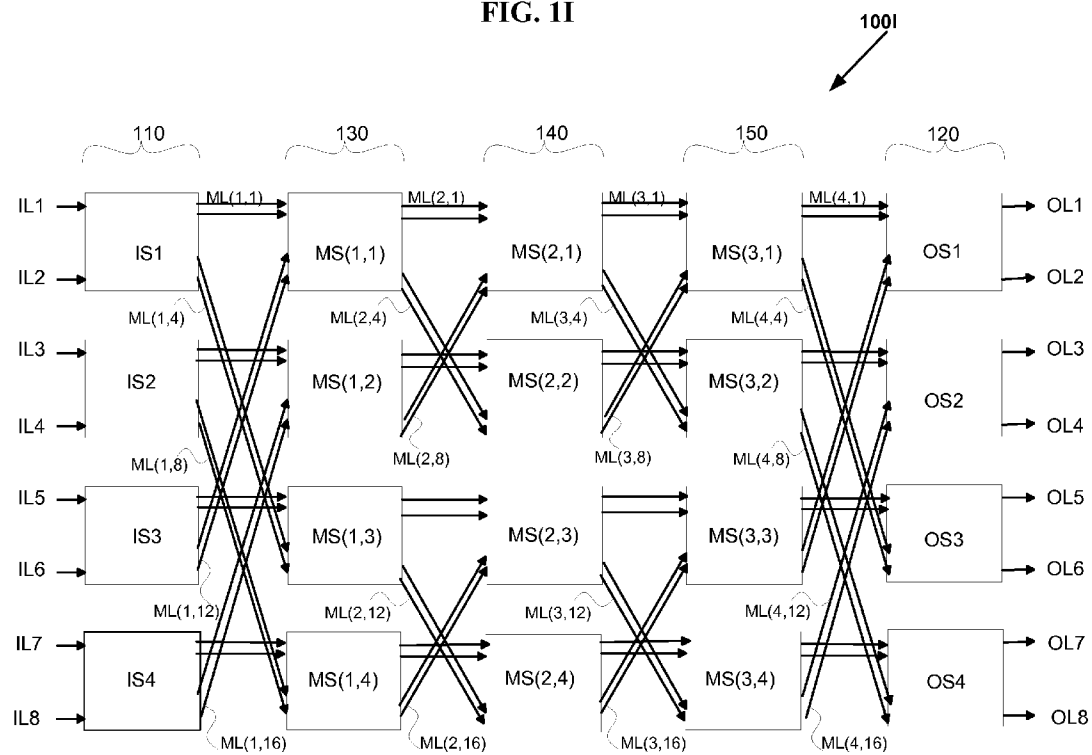


FIG. 1J

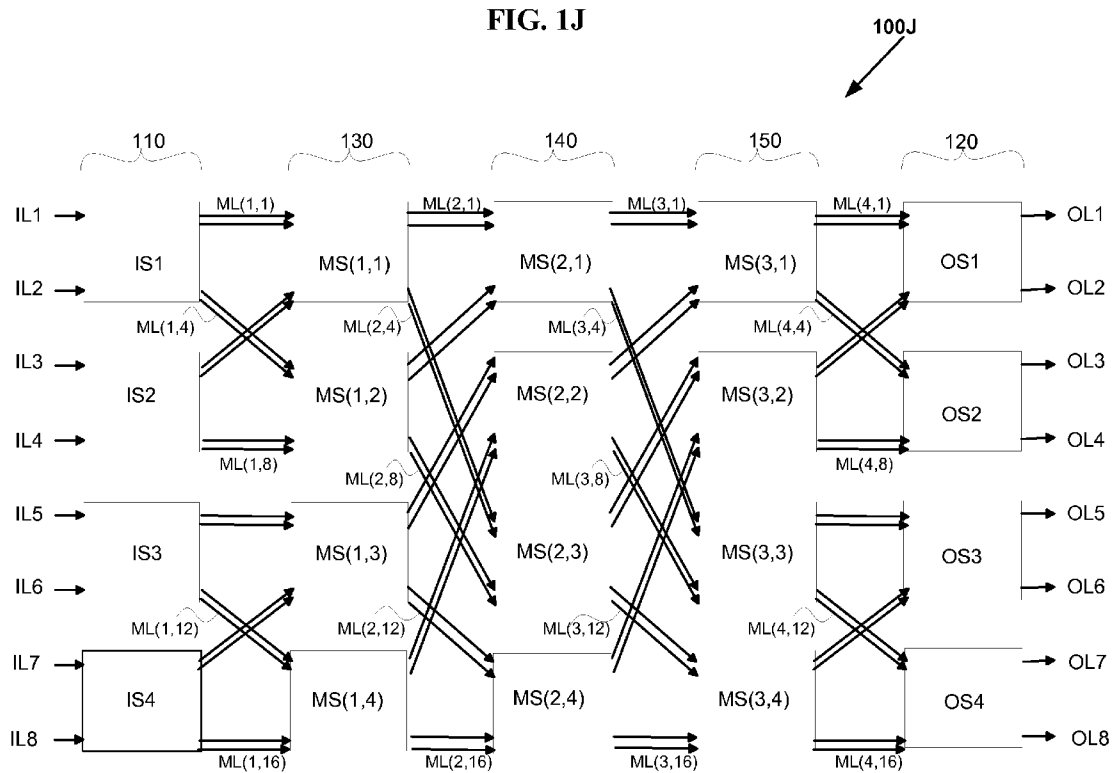


FIG. 1K

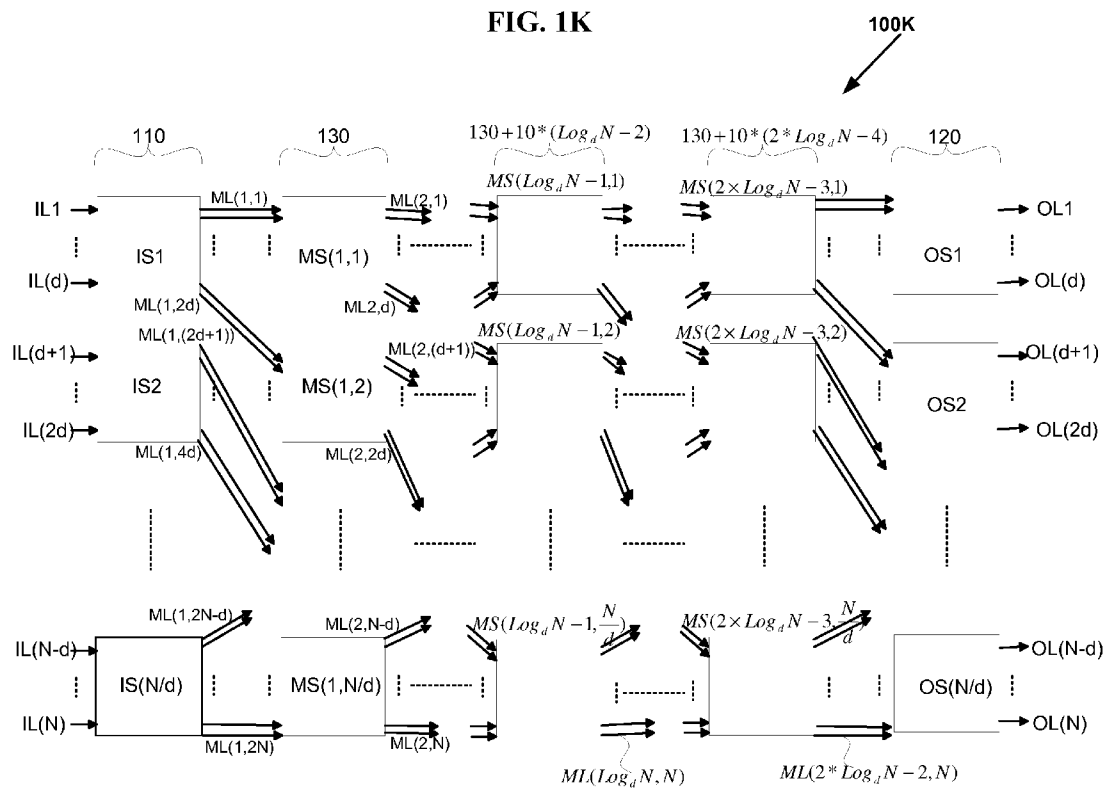


FIG. 1A1

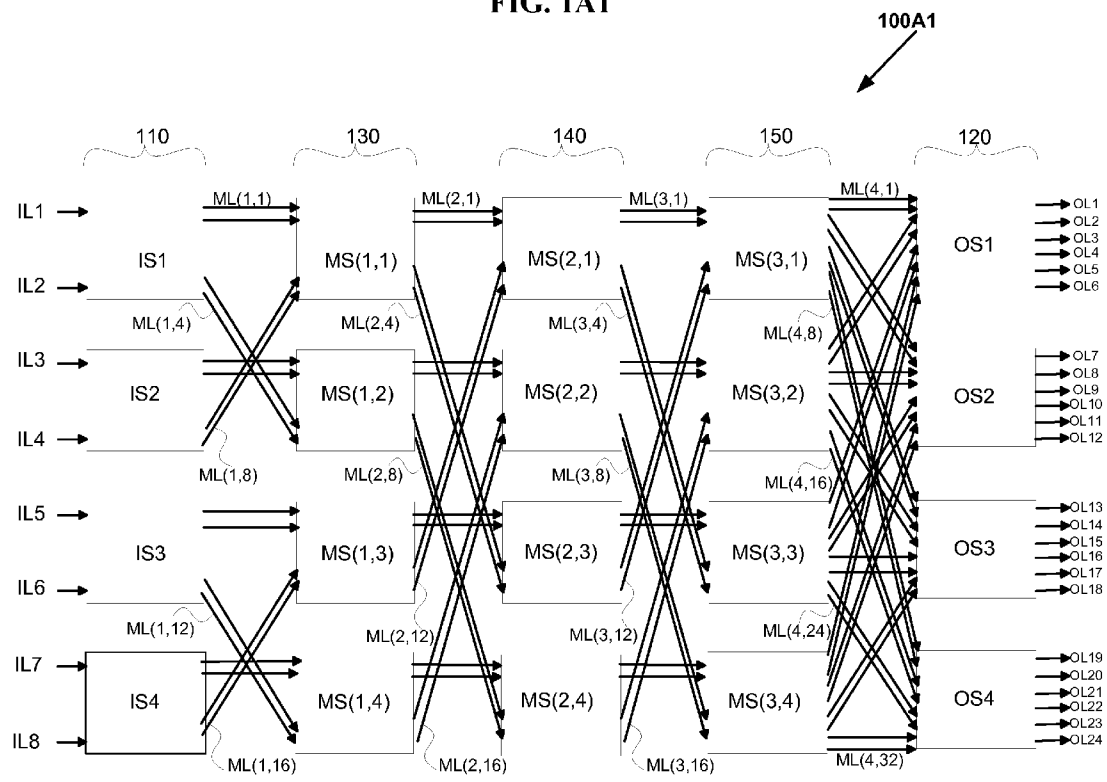


FIG. 1B1

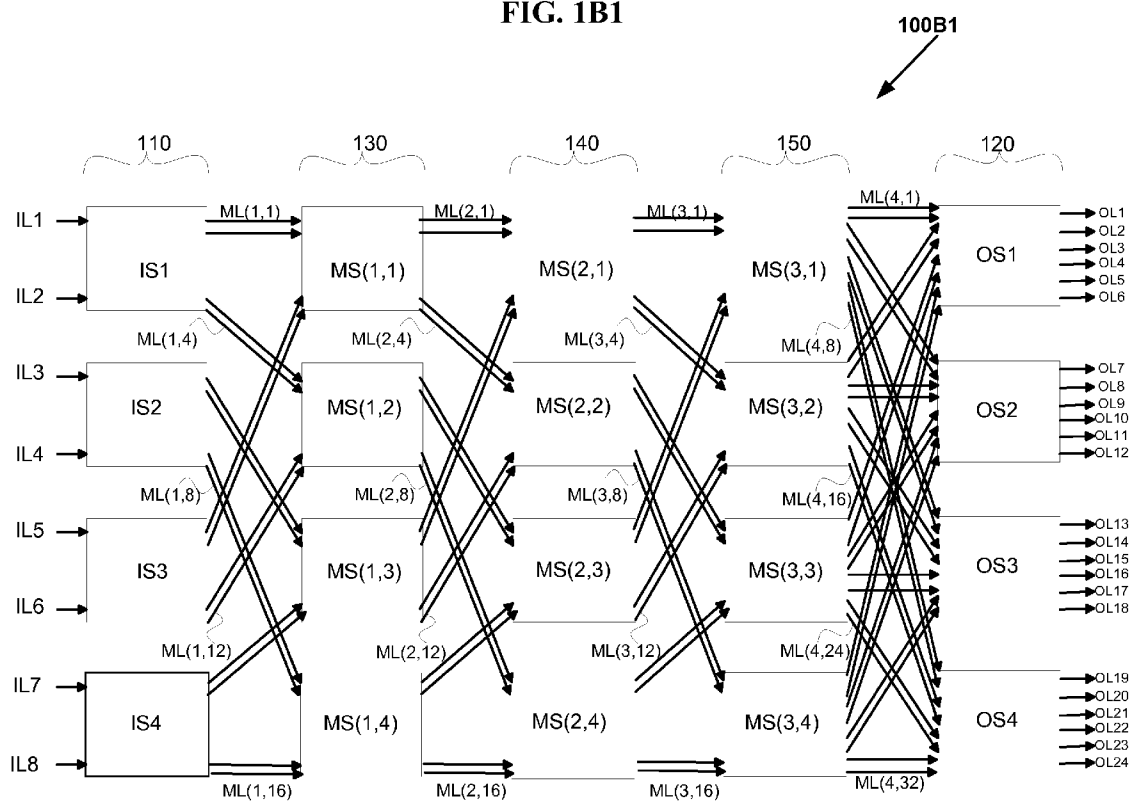


FIG. 1C1

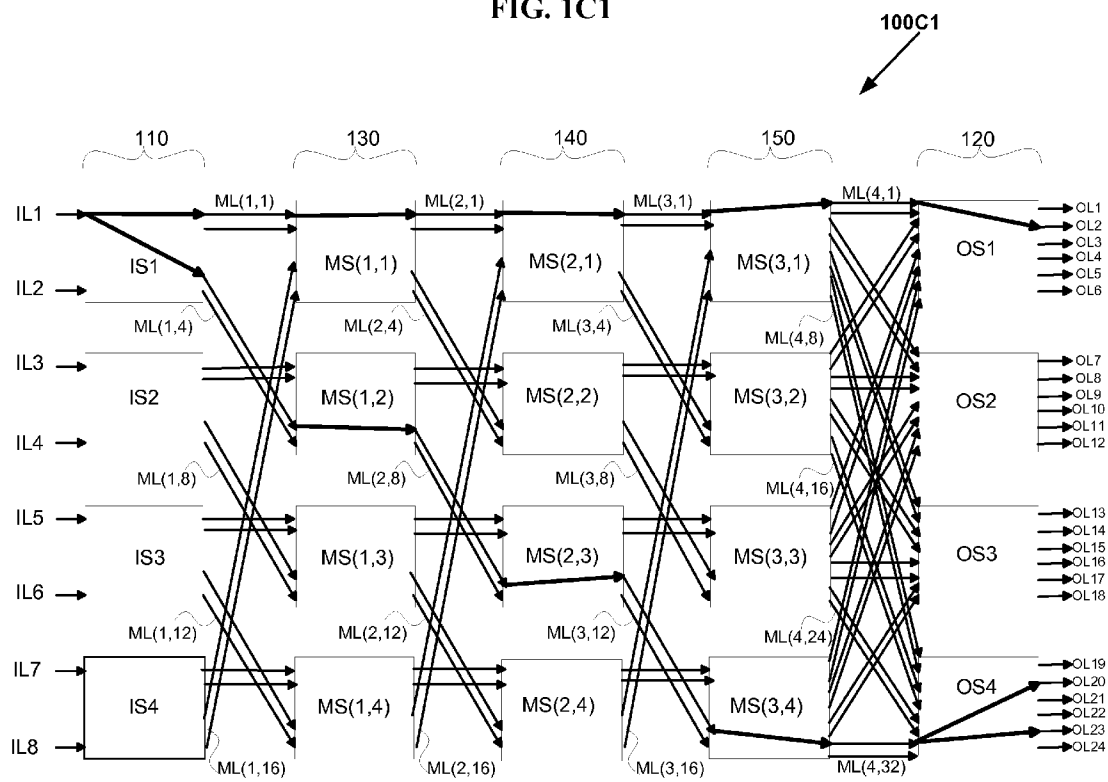


FIG. 1D1

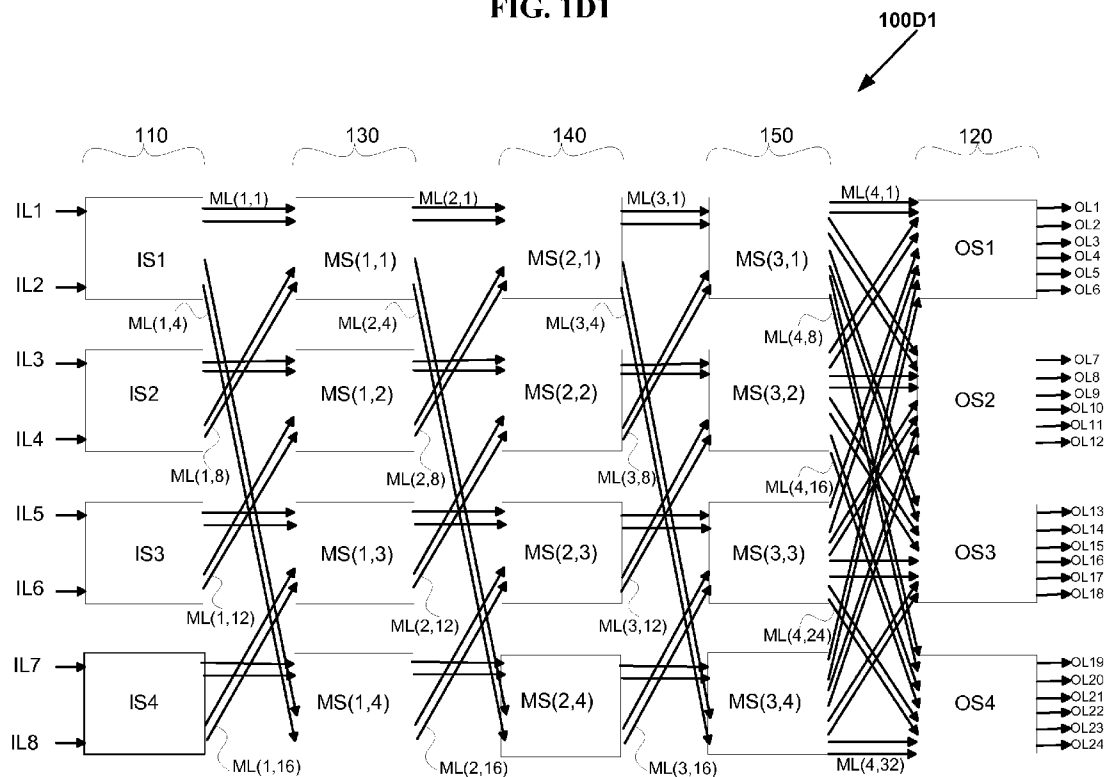


FIG. 1E1

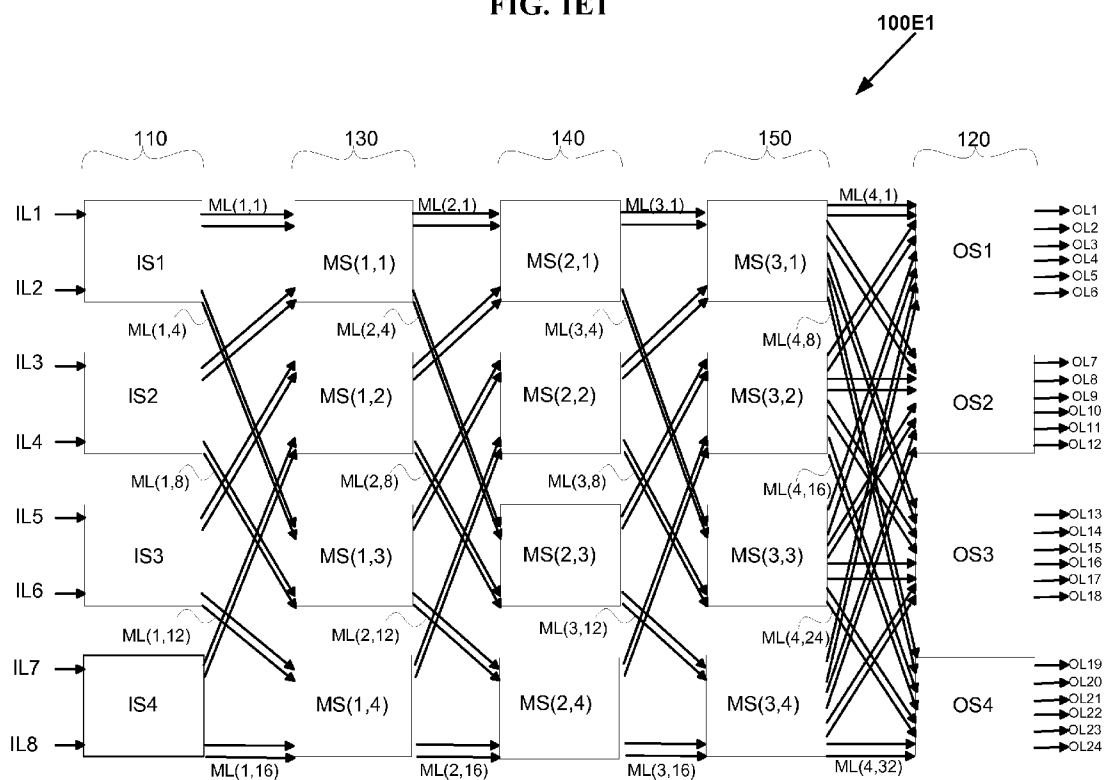


FIG. 1F1

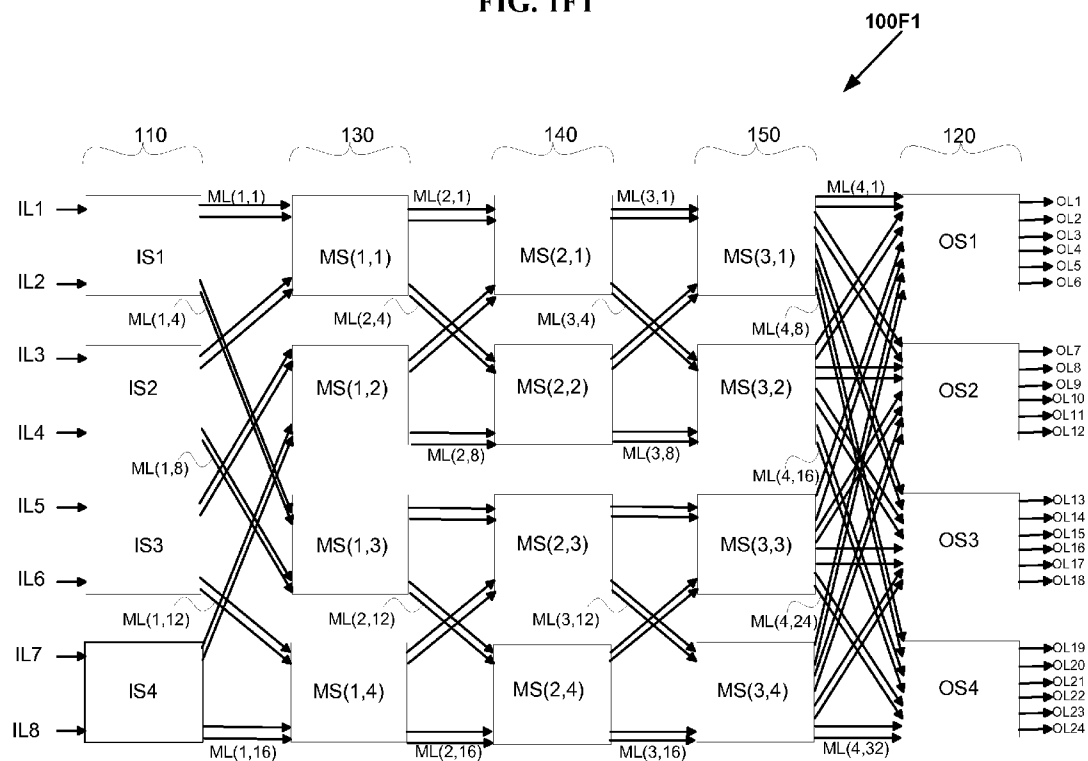


FIG. 1G1

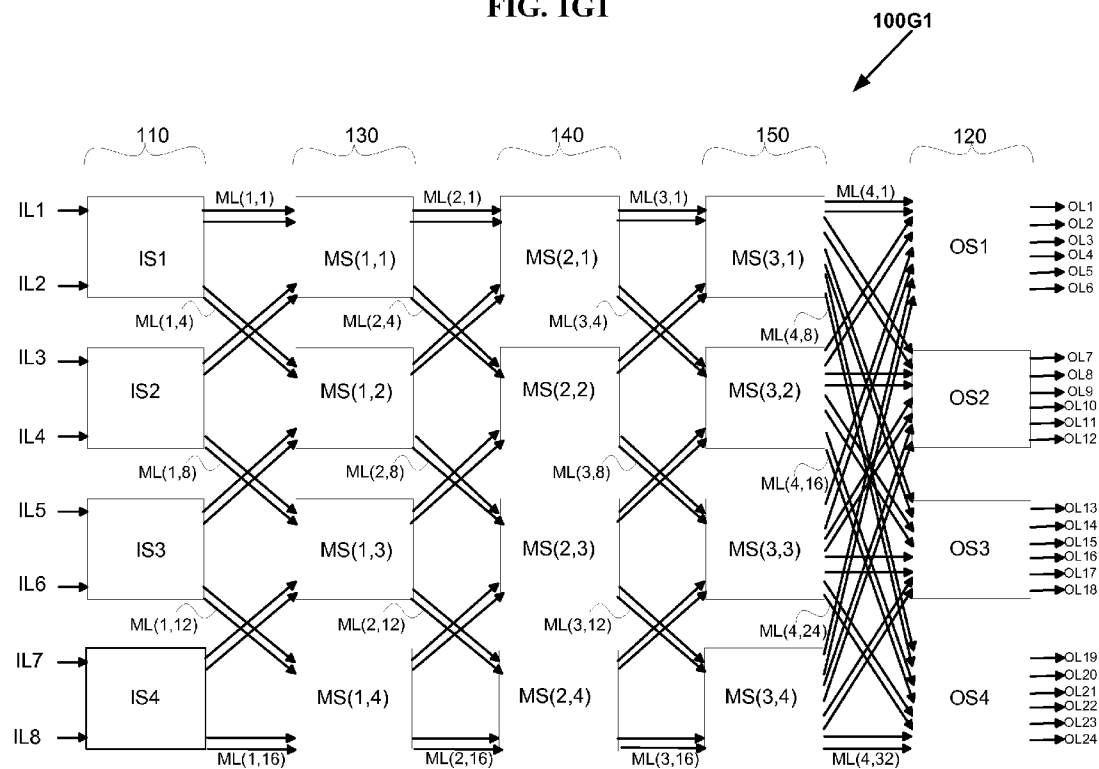


FIG. 1H1

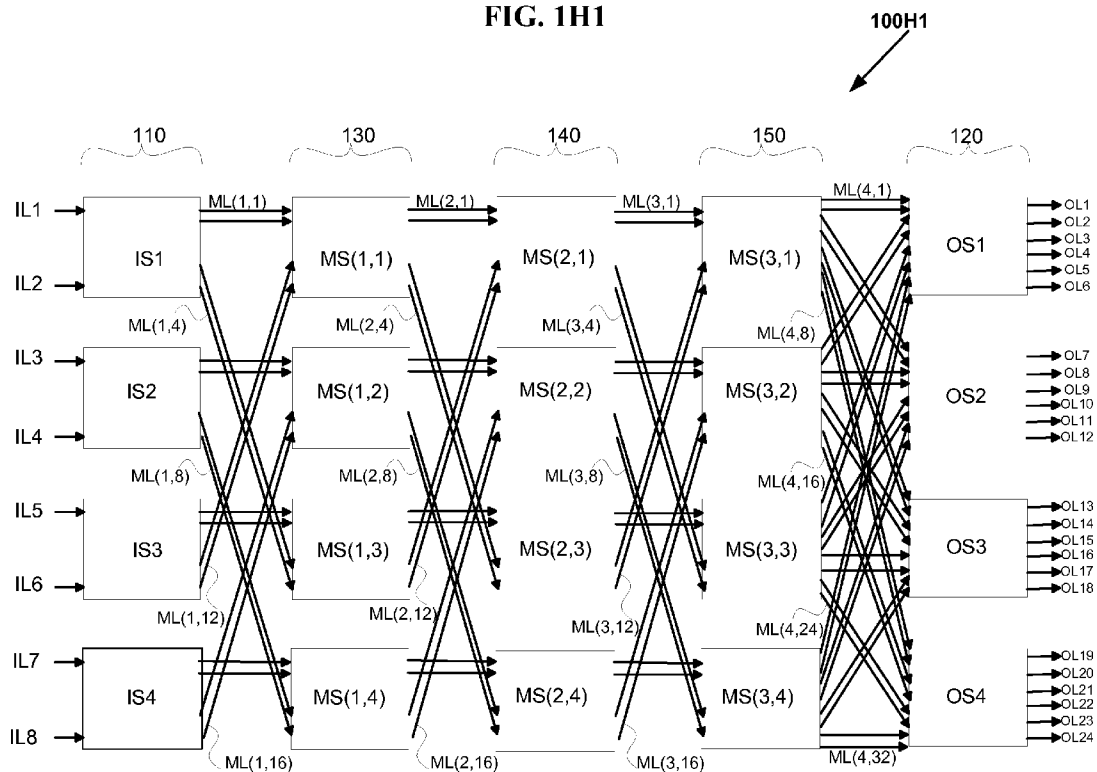


FIG. 111

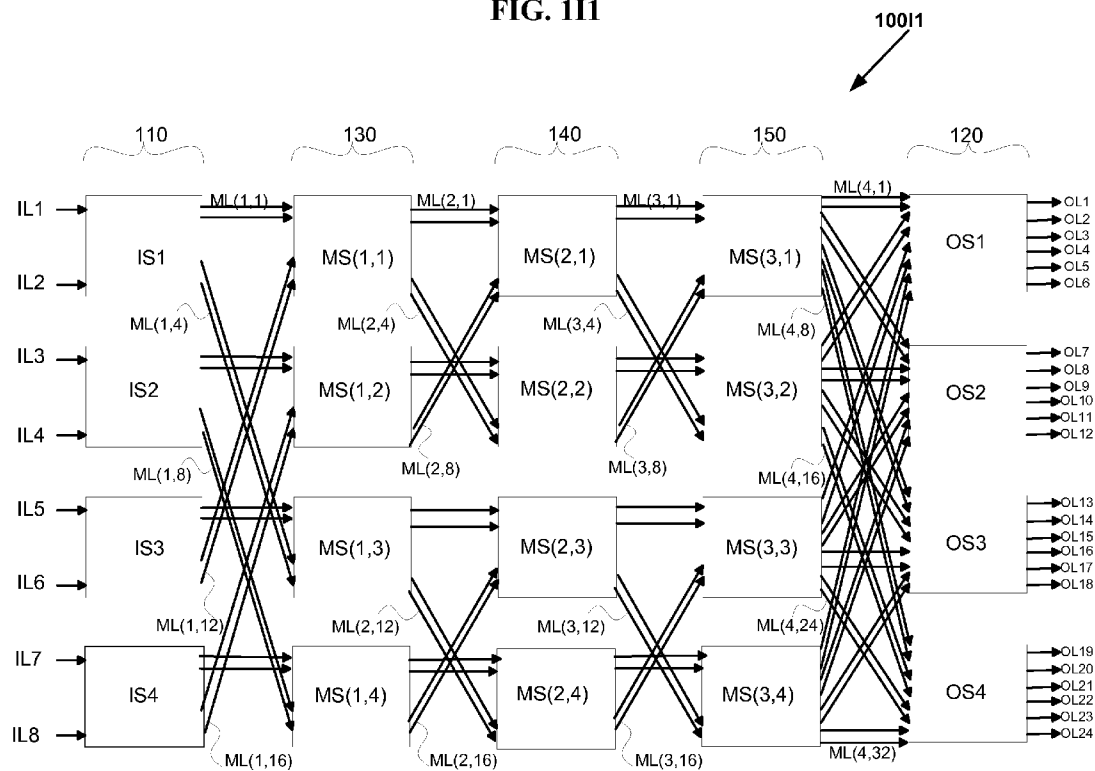


FIG. 1J1

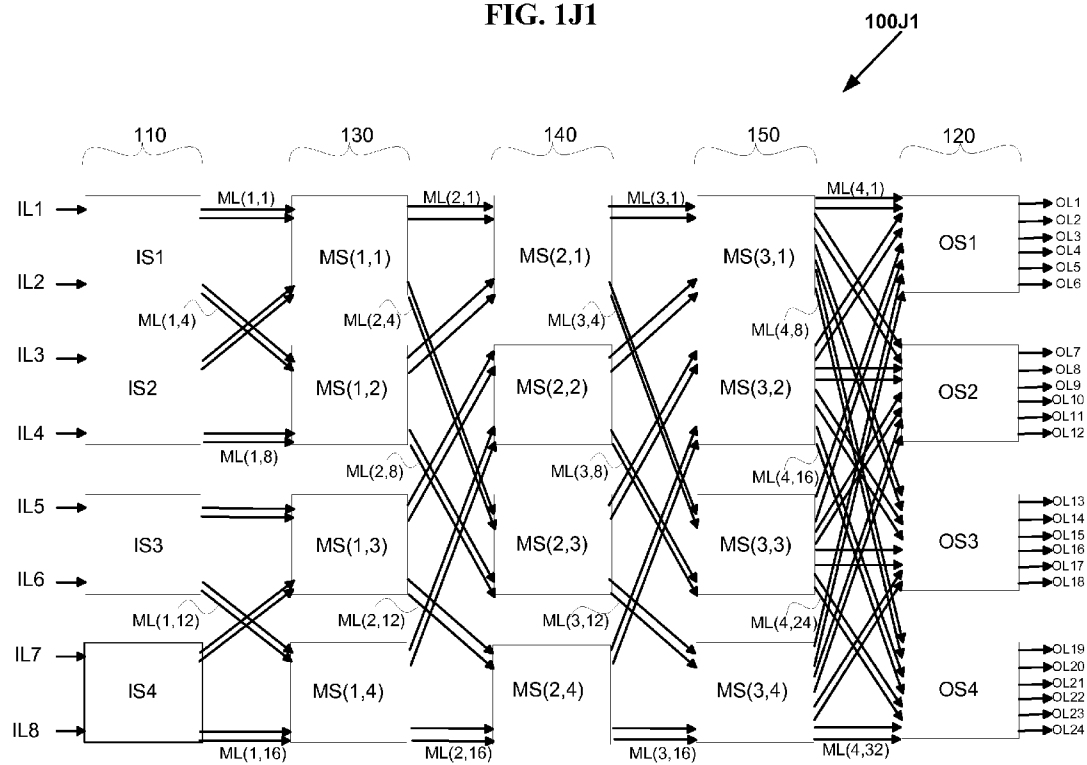


FIG. 1K1

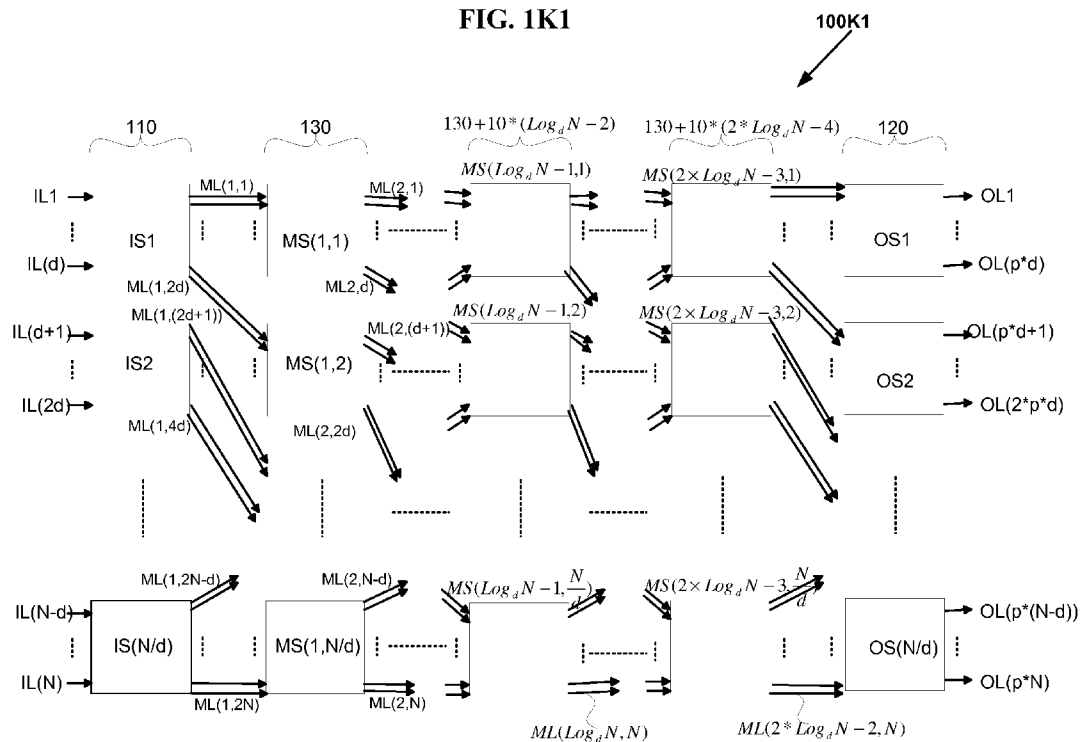


FIG. 1A2

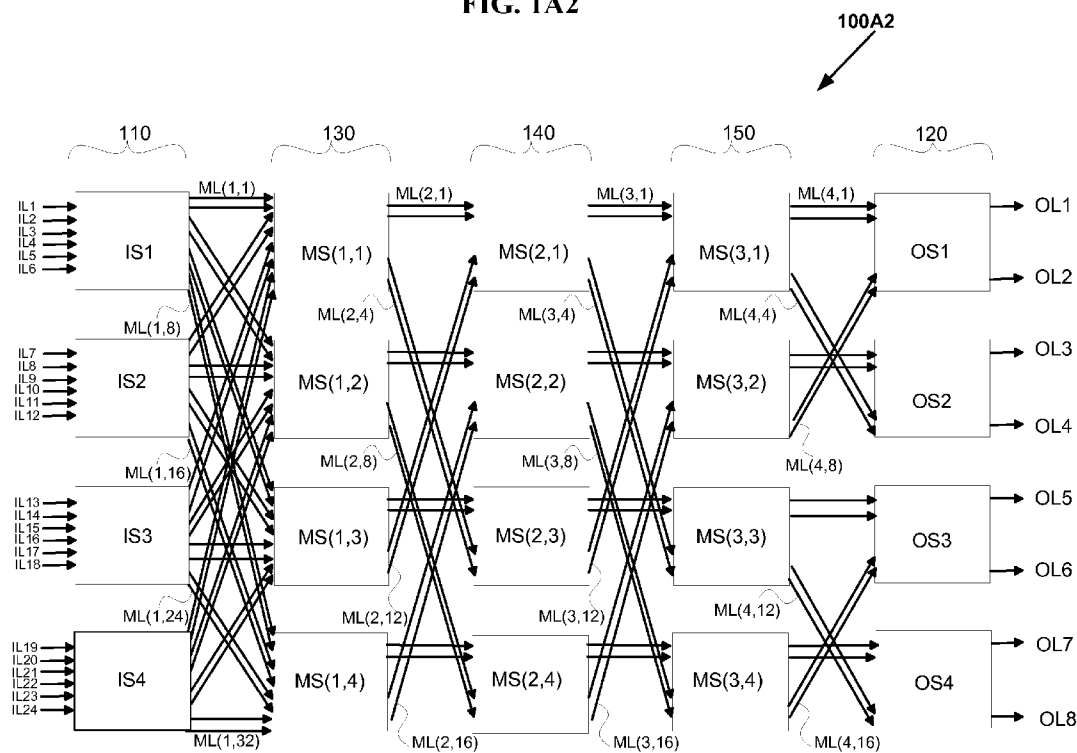


FIG. 1B2

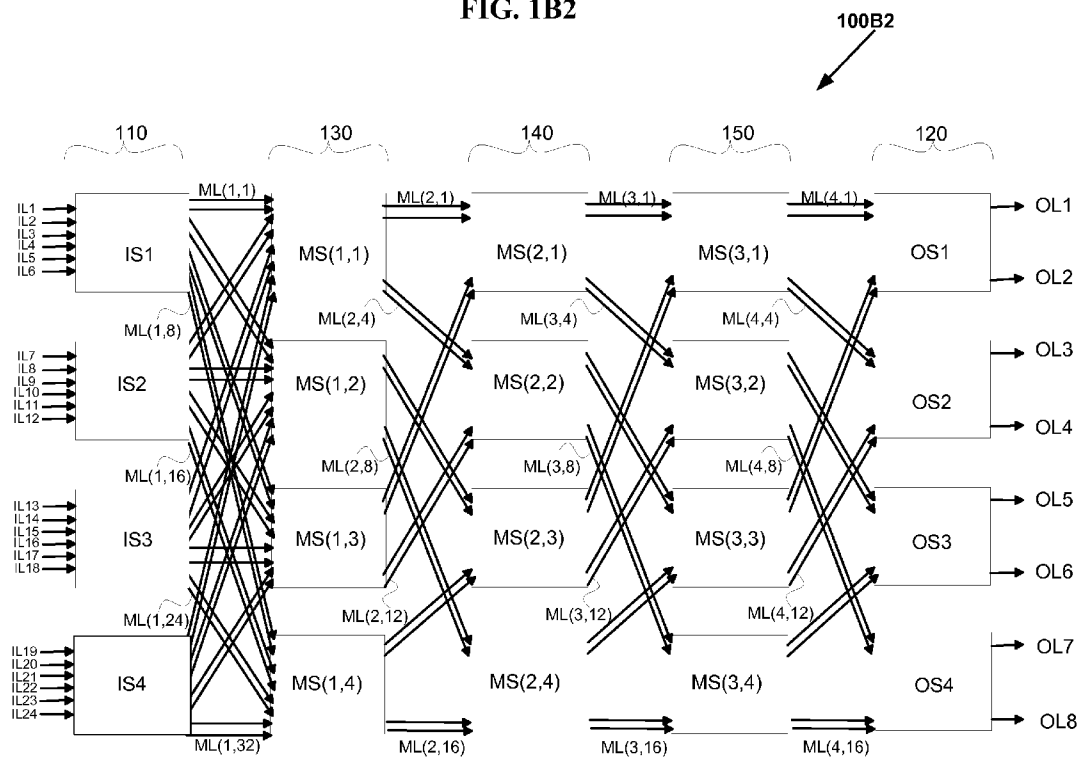


FIG. 1C2

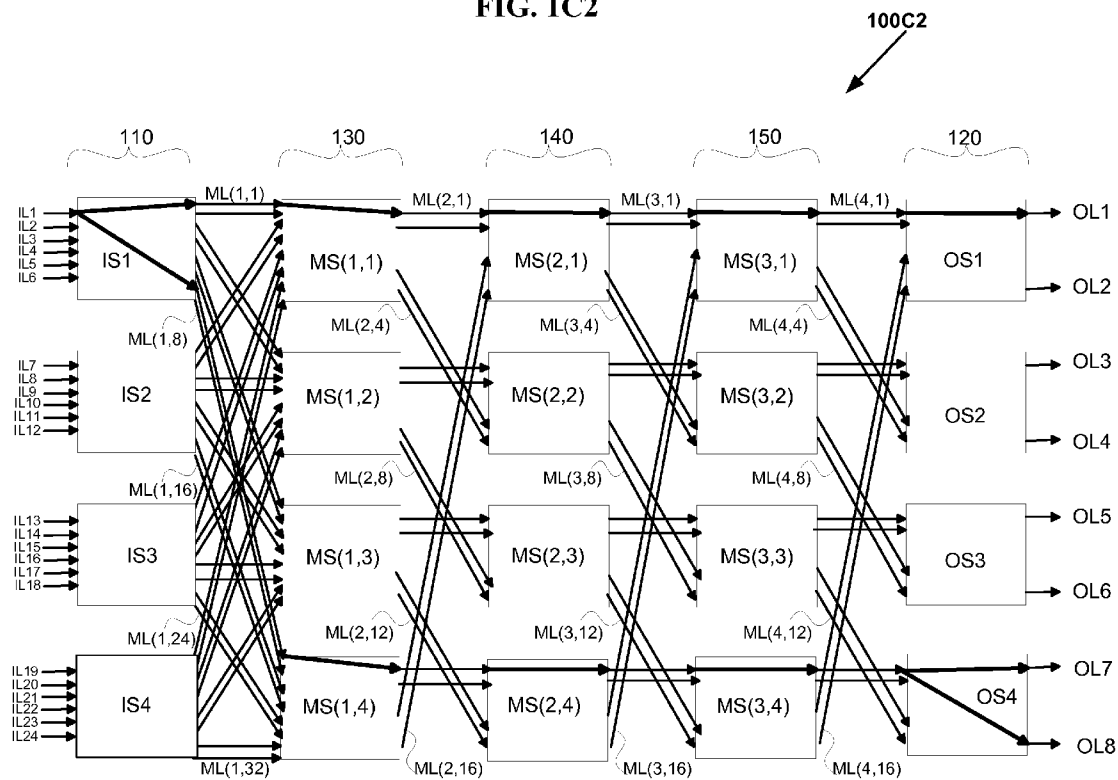


FIG. 1D2

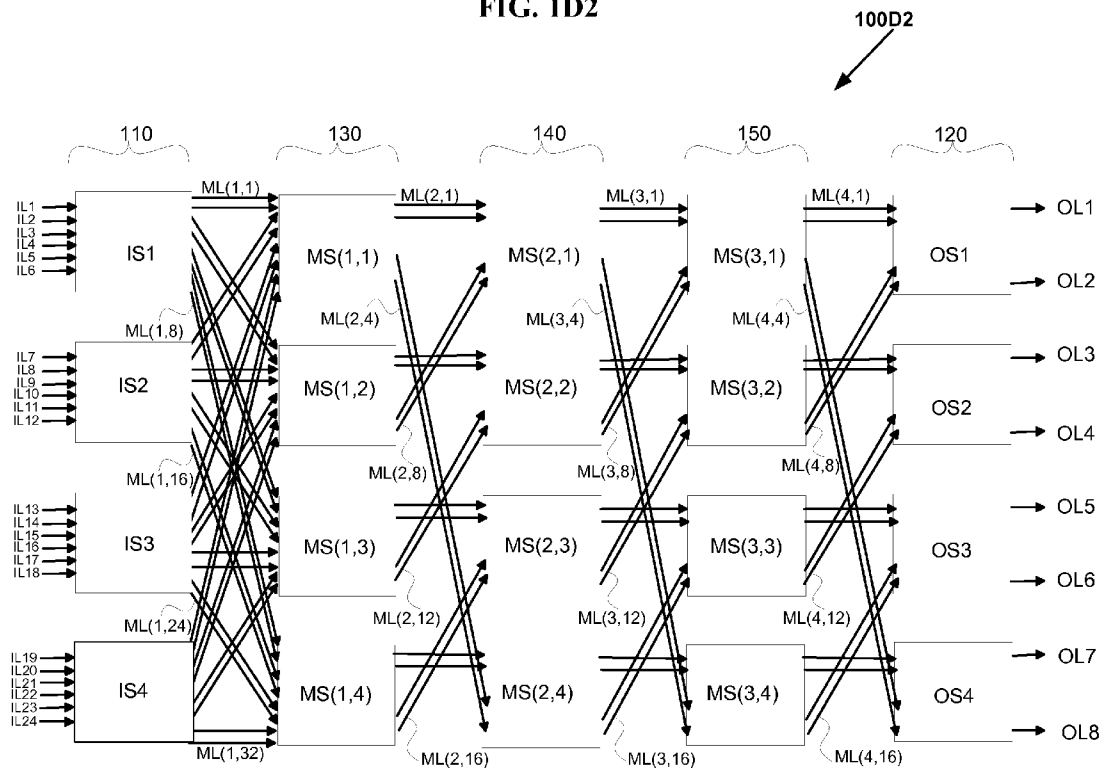


FIG. 1E2

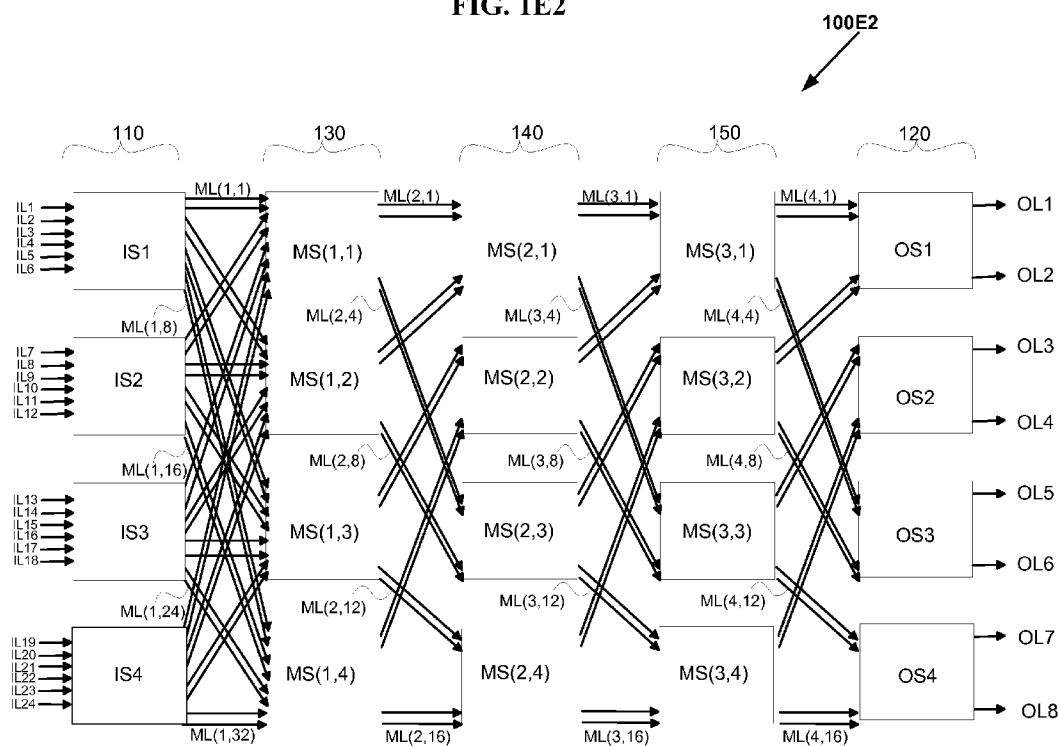


FIG. 1F2

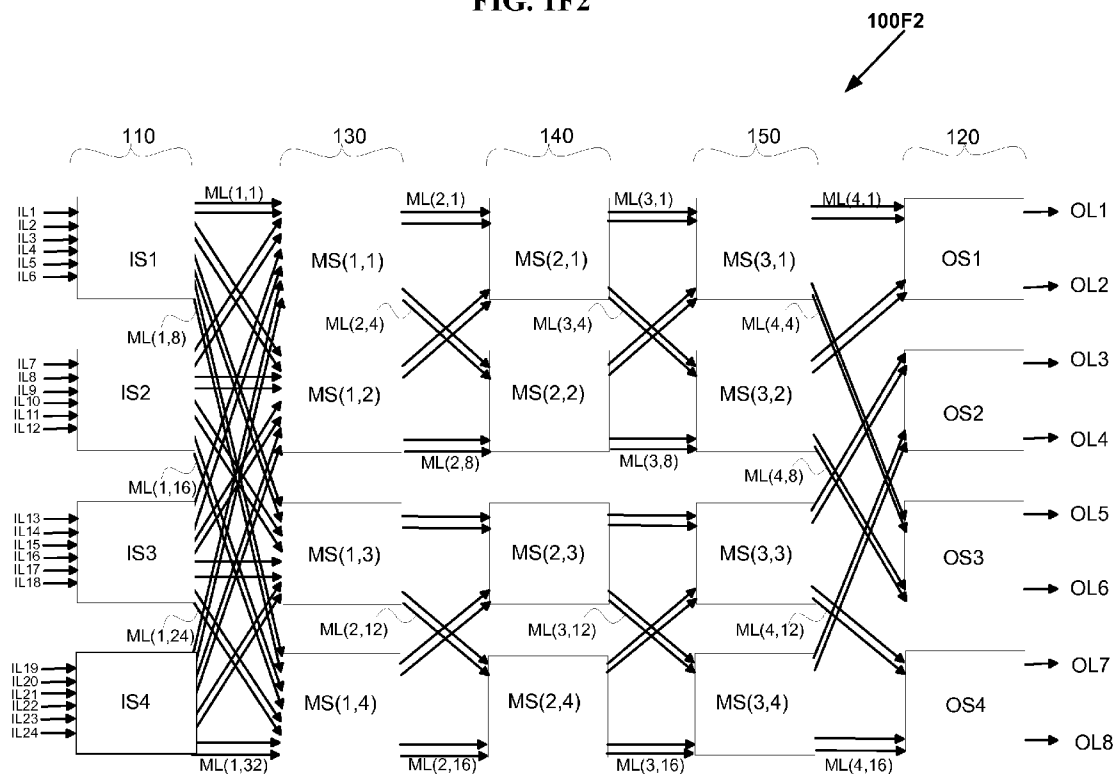


FIG. 1G2

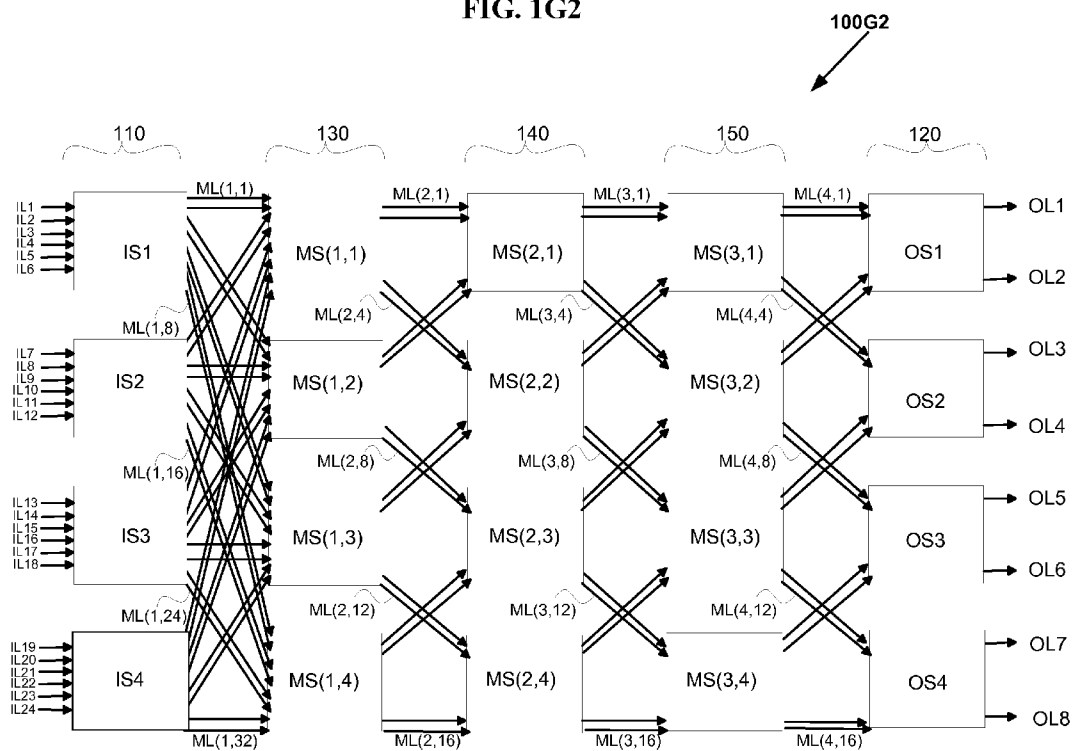


FIG. 1H2

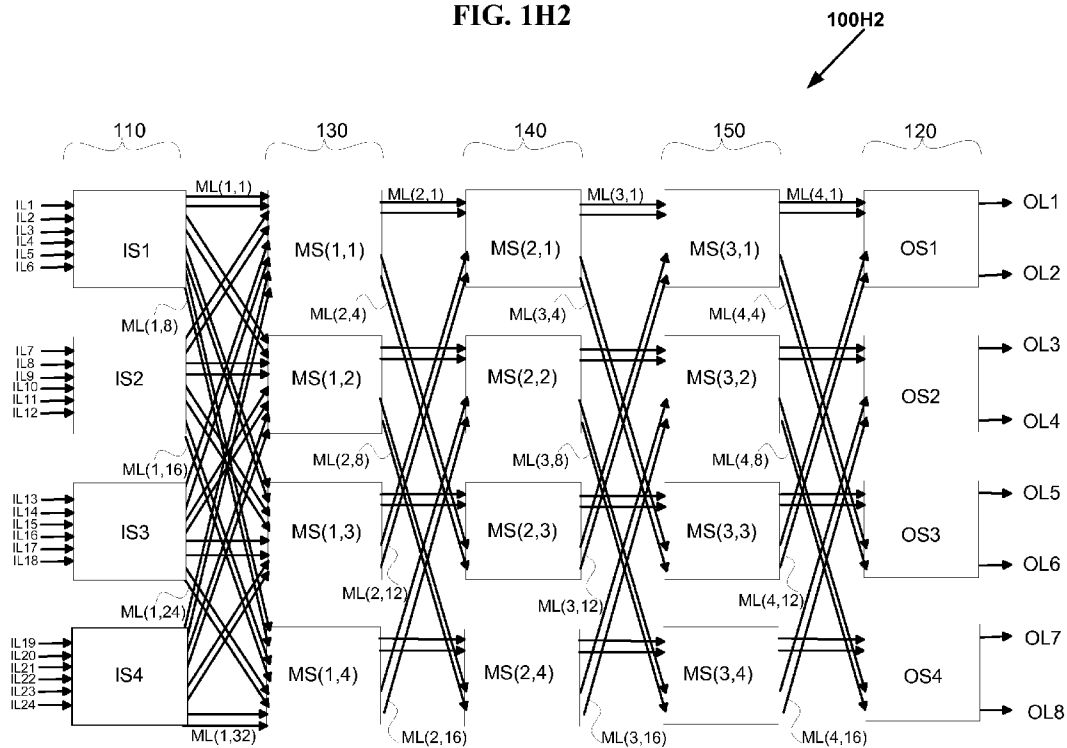
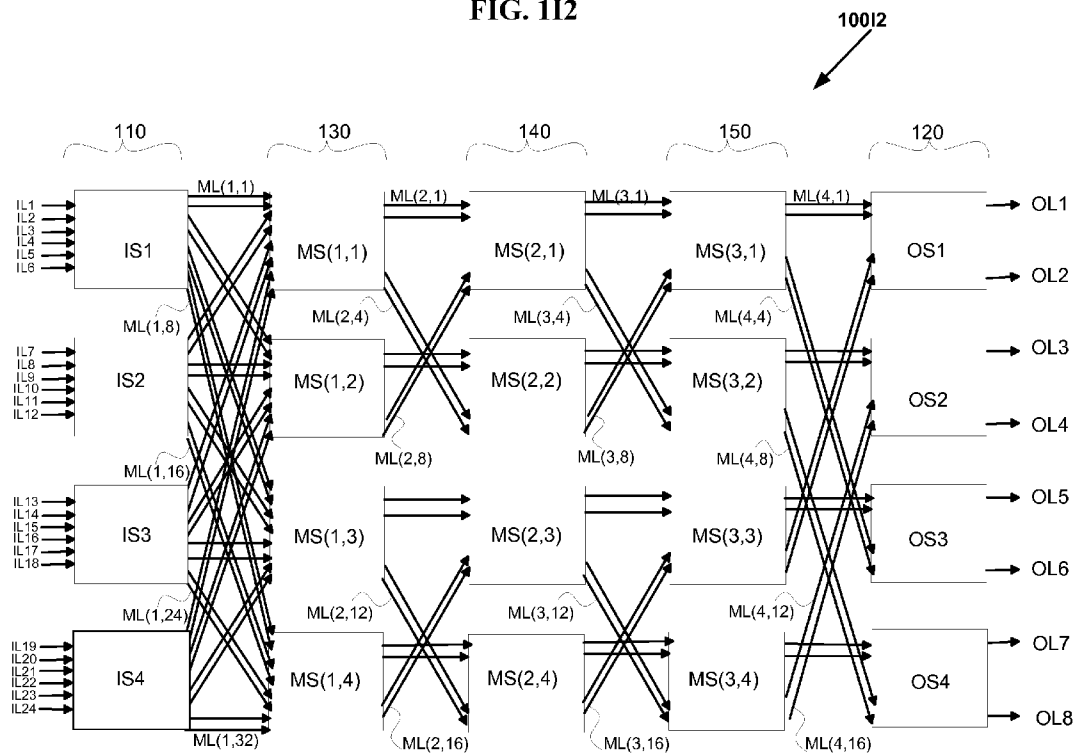


FIG. 112



10012

FIG. 1J2

100J2

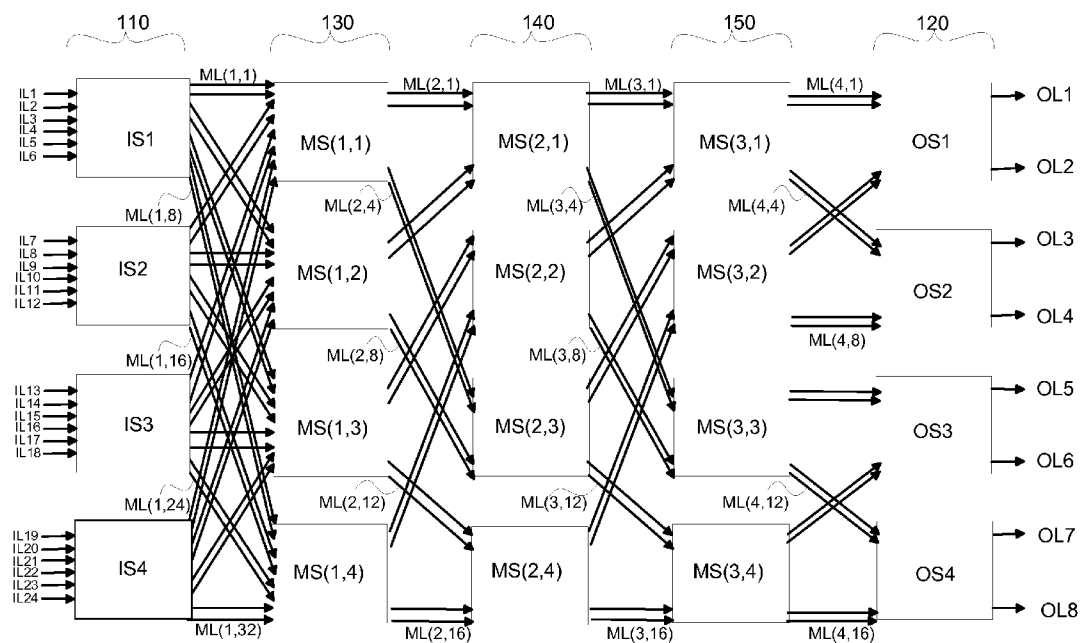


FIG. 1K2

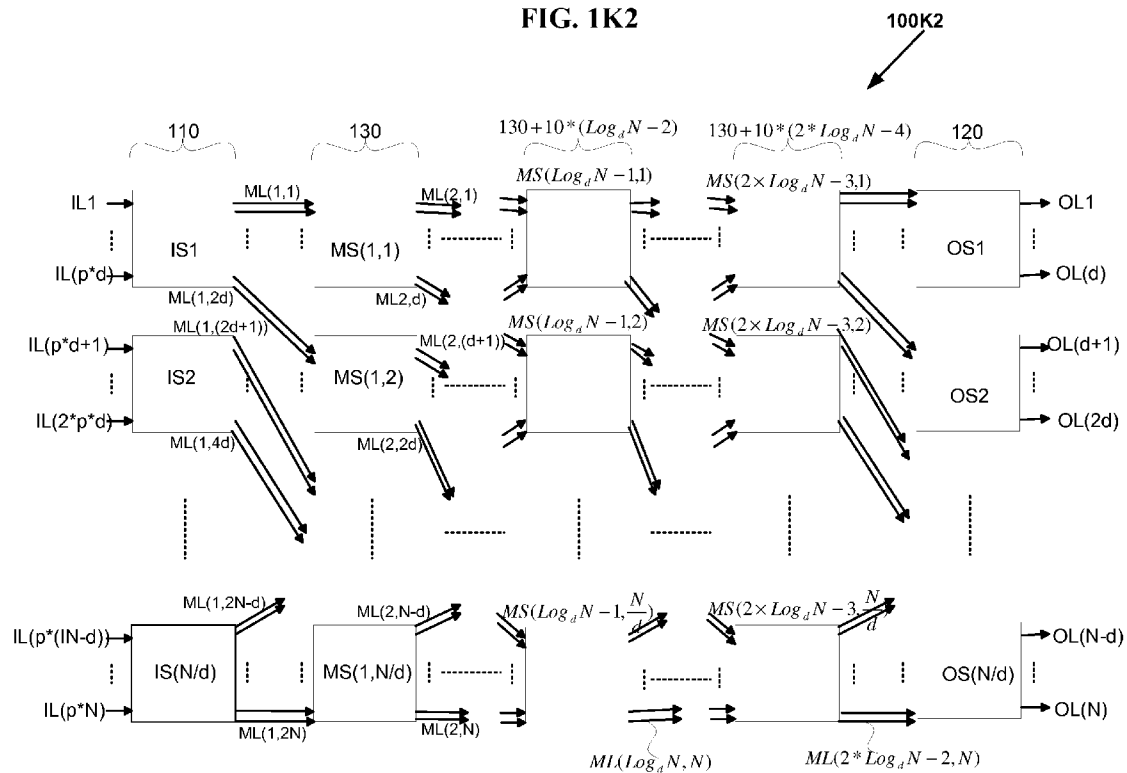
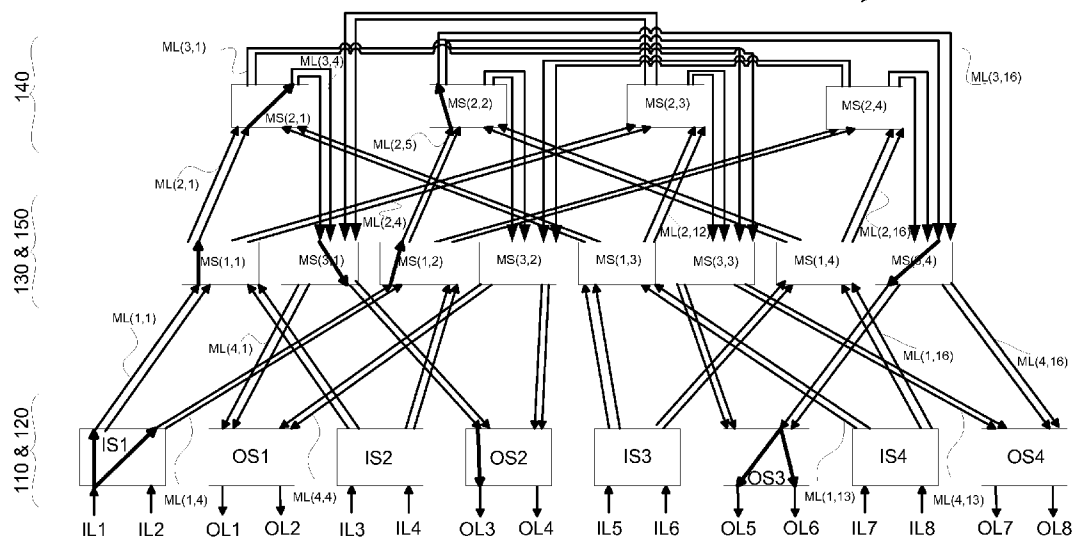


FIG. 2A

200A



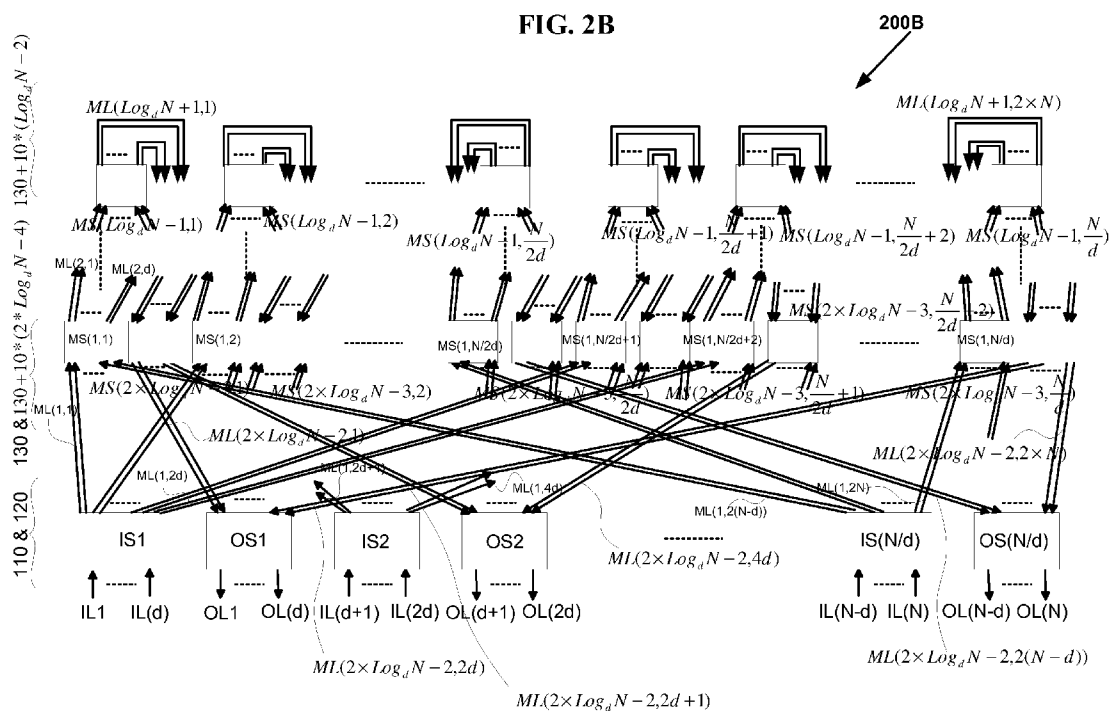
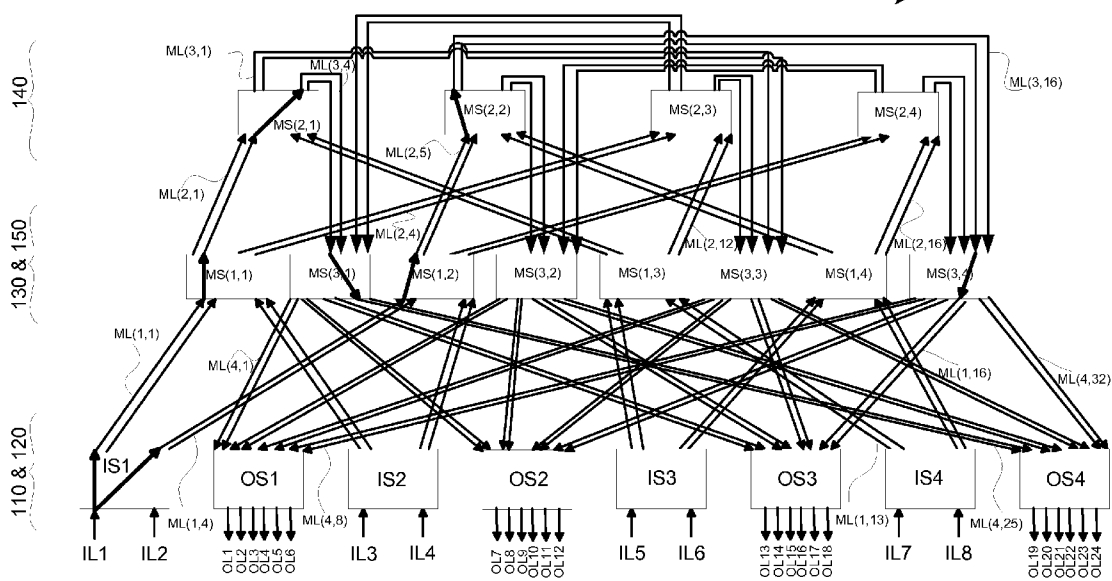


FIG. 2C

200C



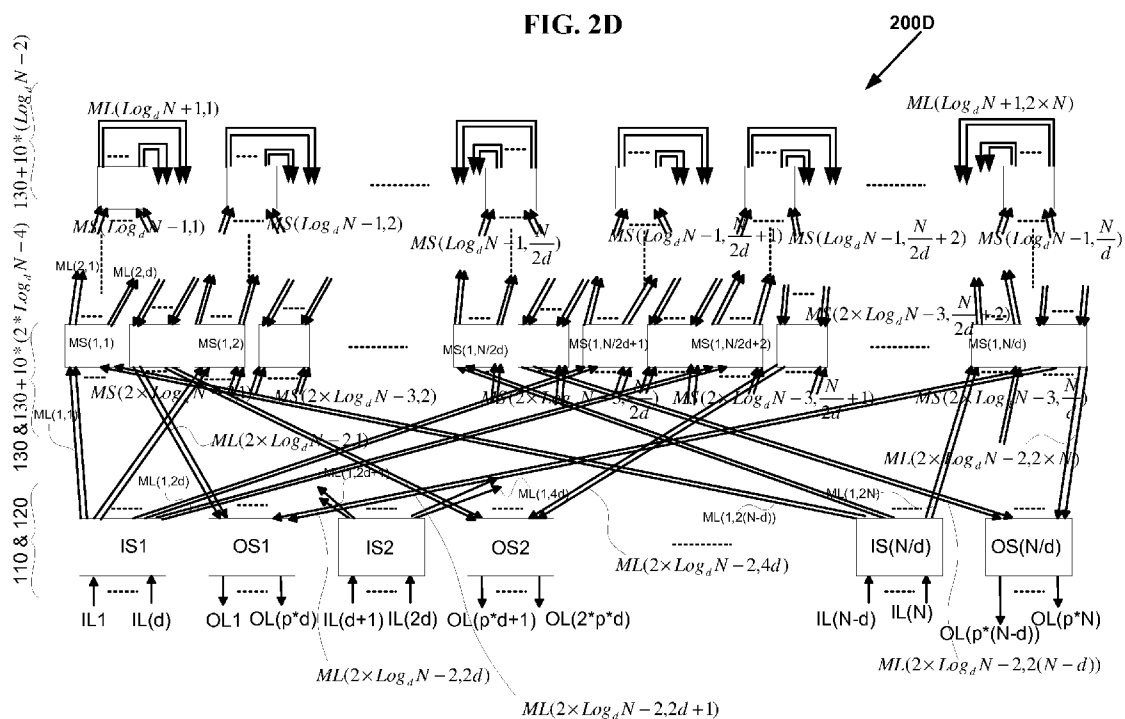
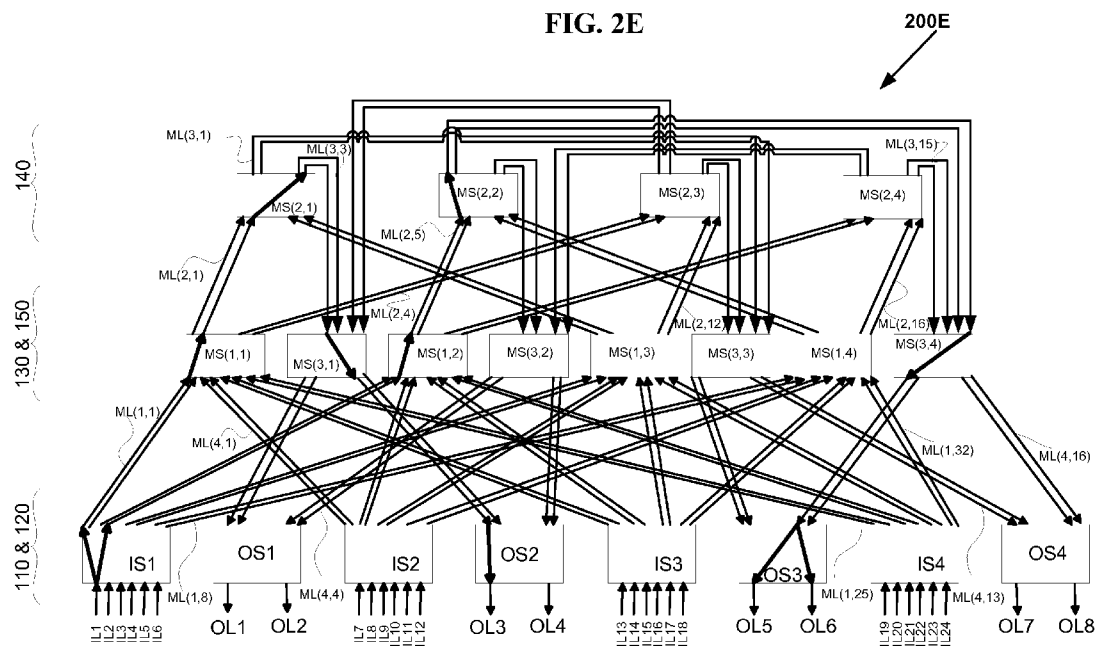


FIG. 2E



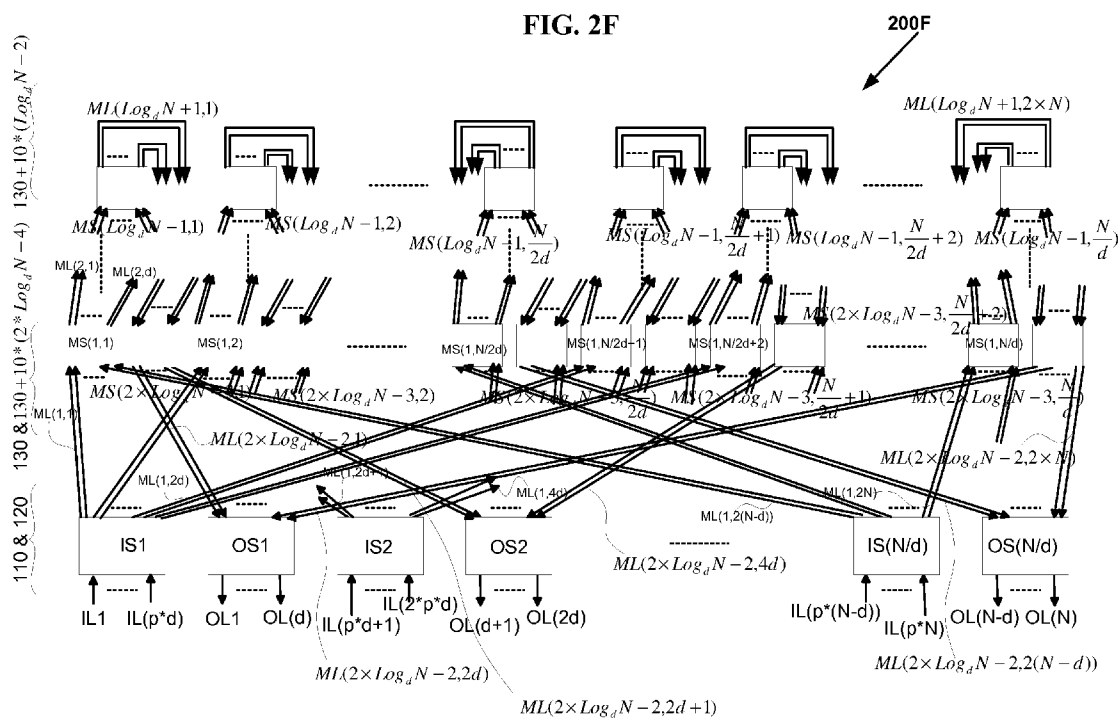


FIG. 3A

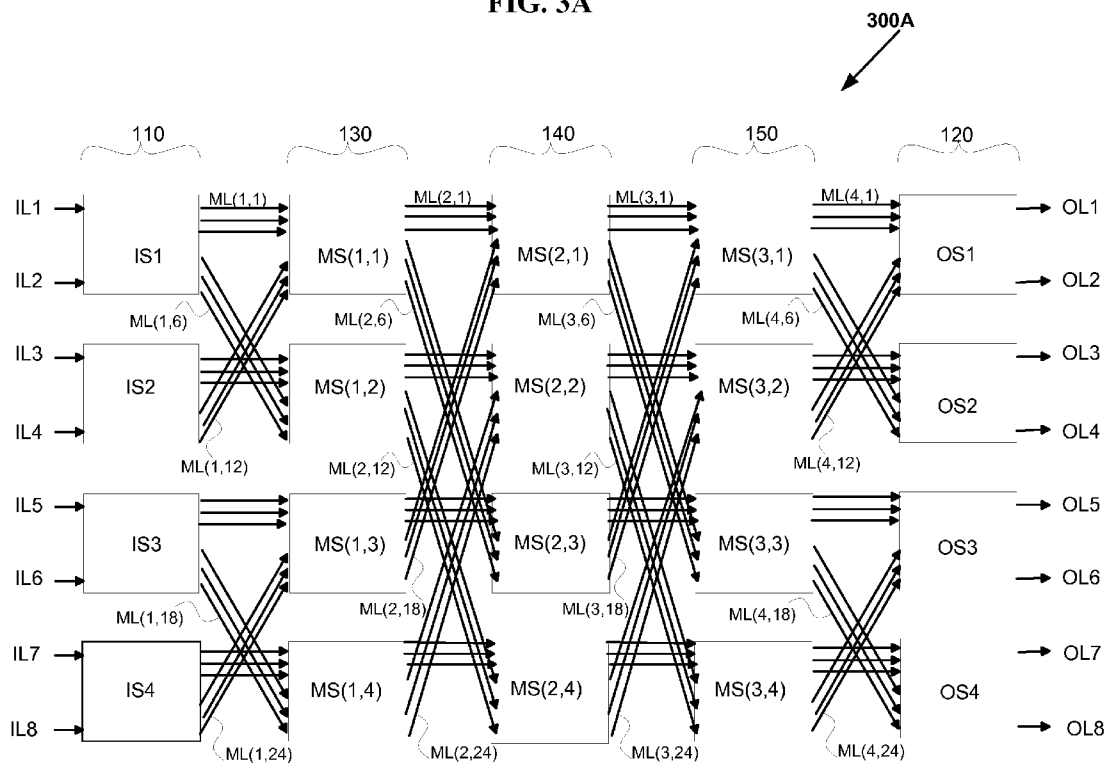


FIG. 3B

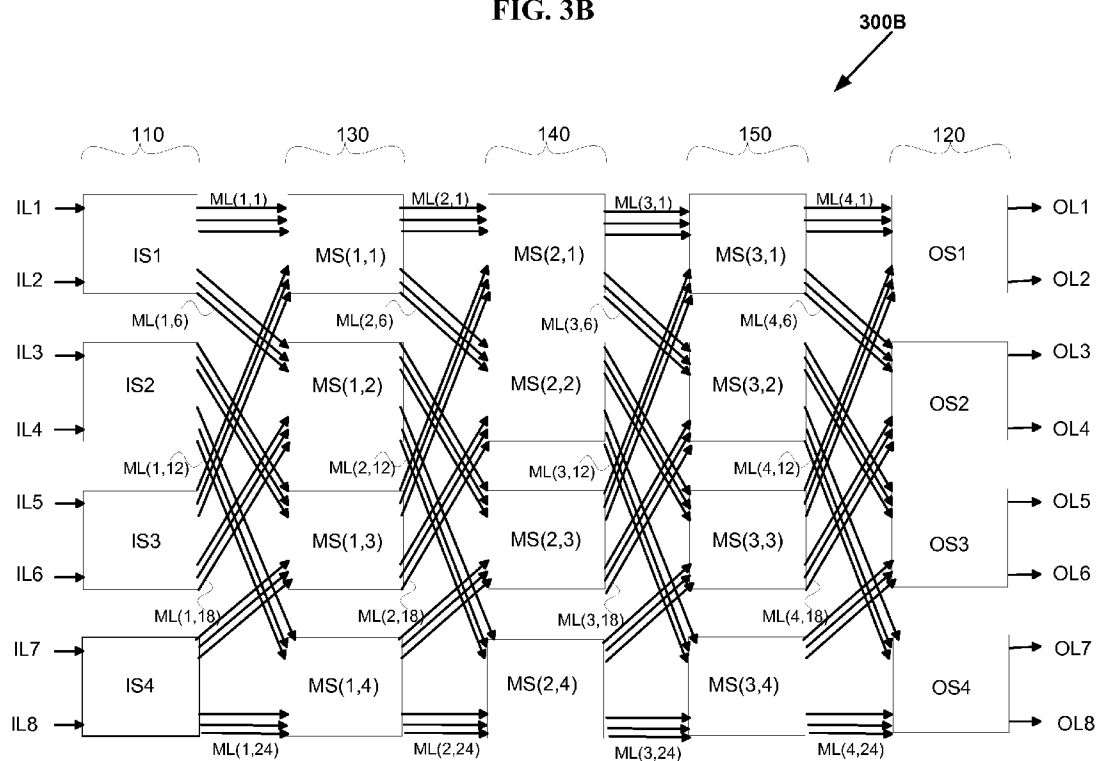


FIG. 3C

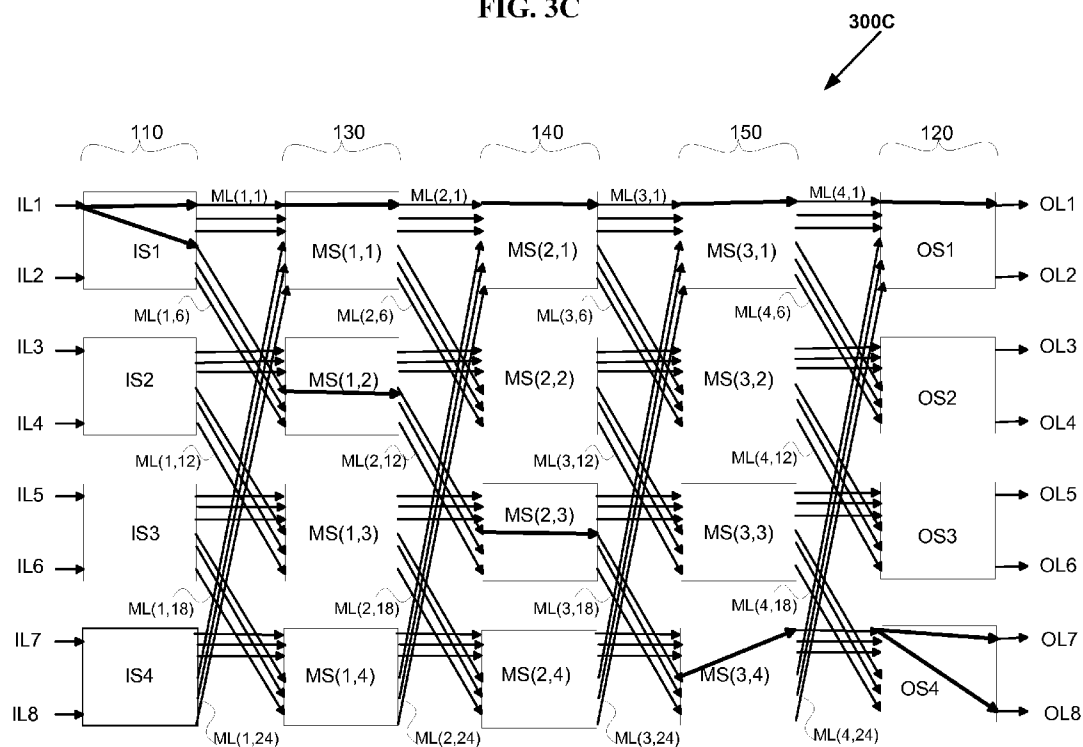


FIG. 3D

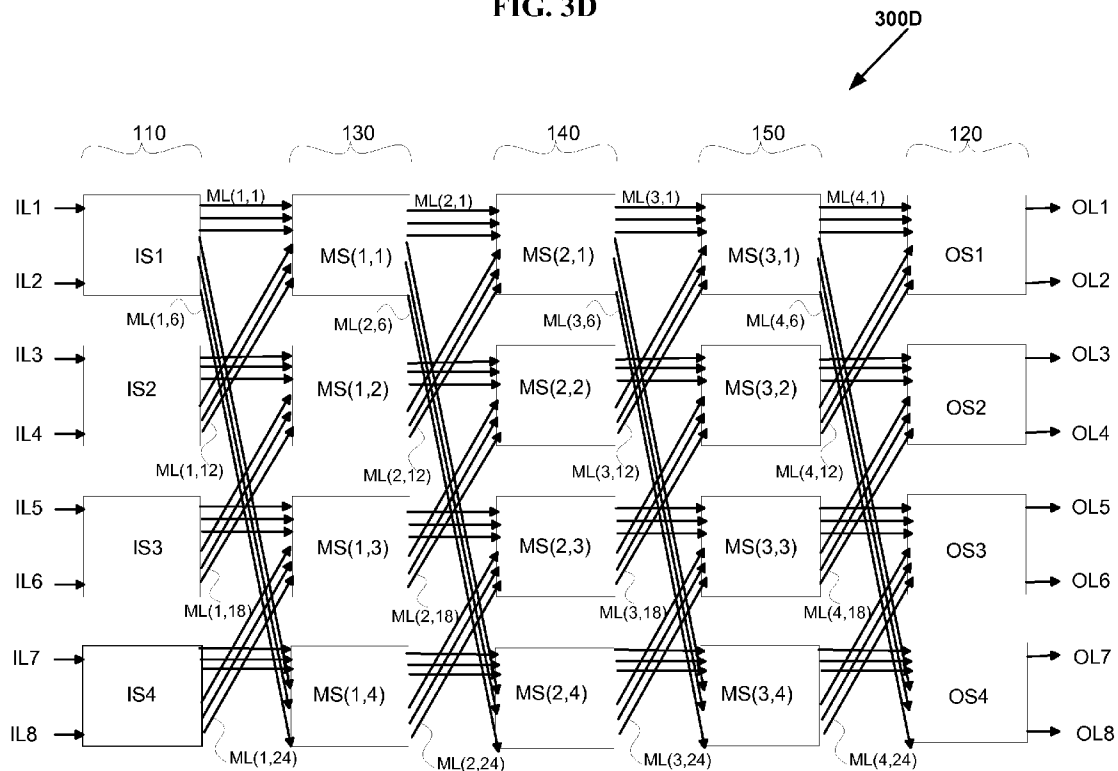


FIG. 3E

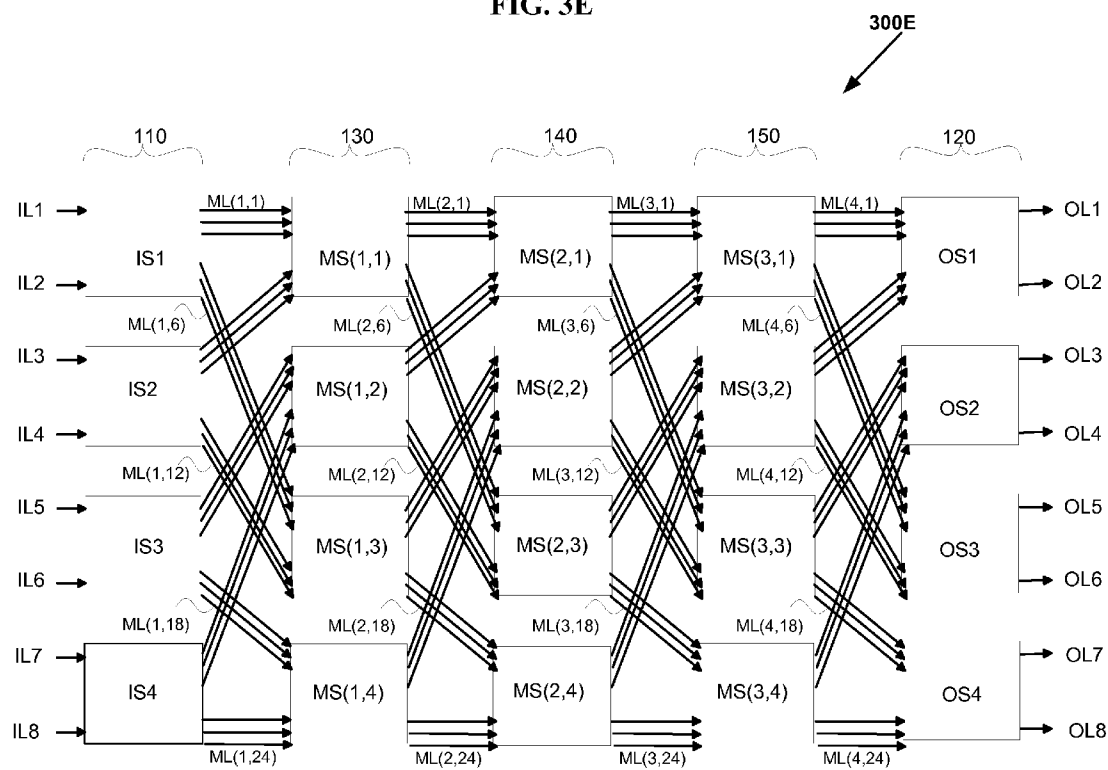


FIG. 3F

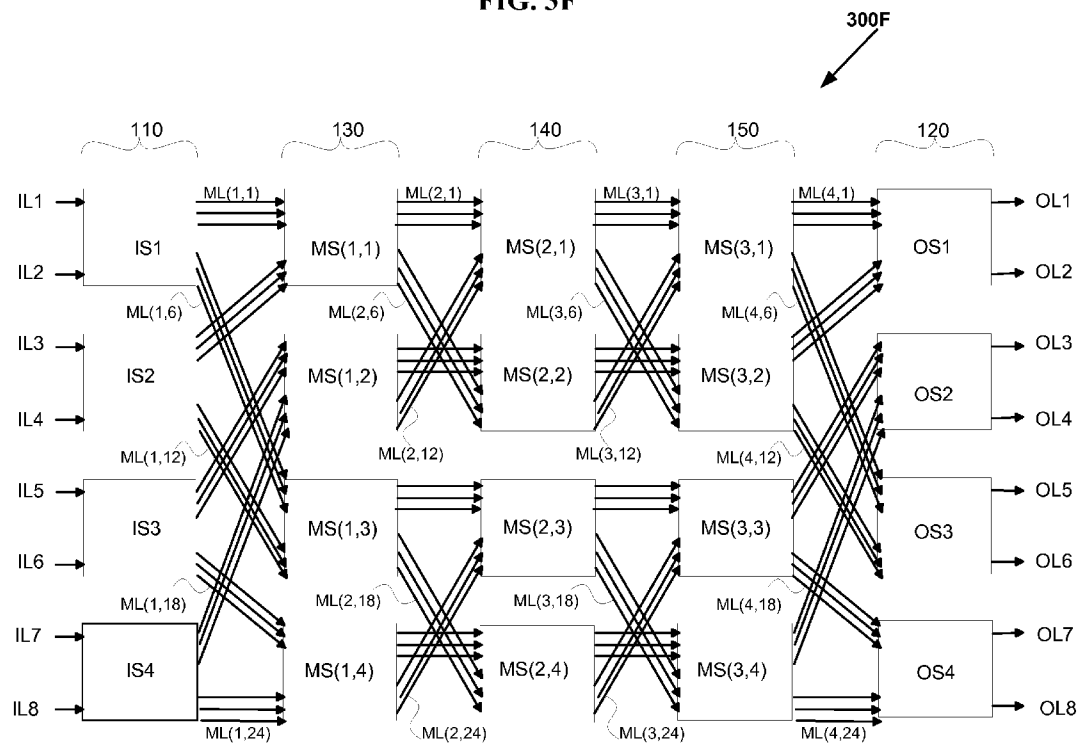


FIG. 3G

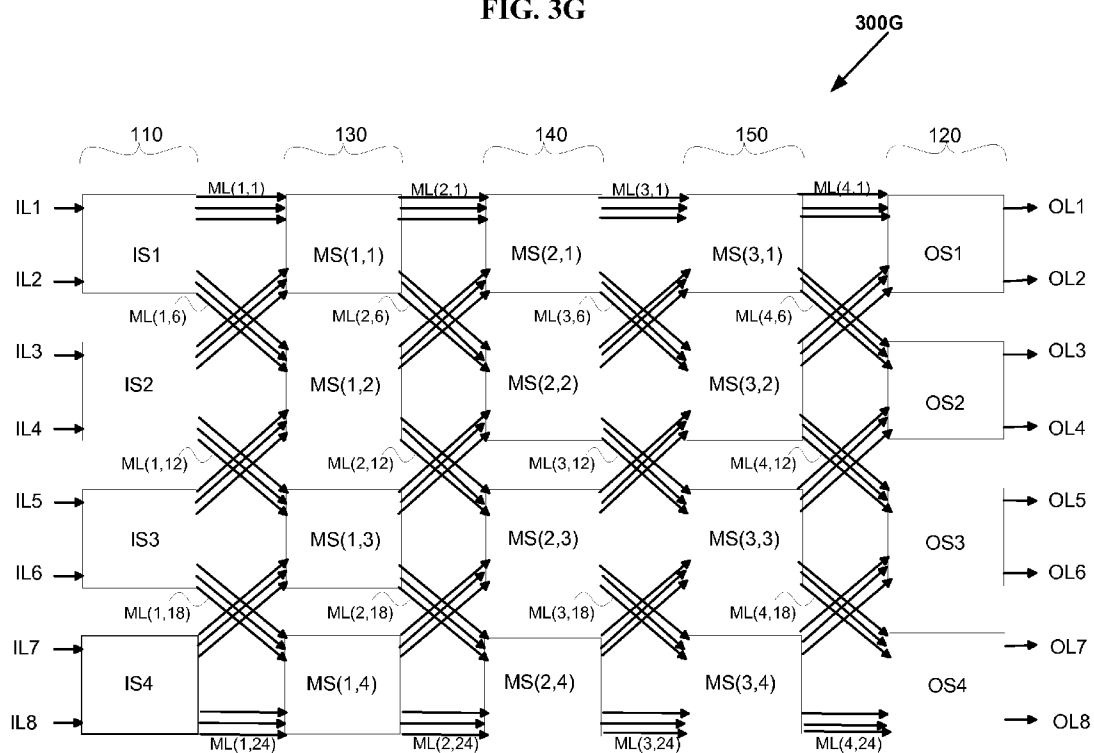


FIG. 3H

300H

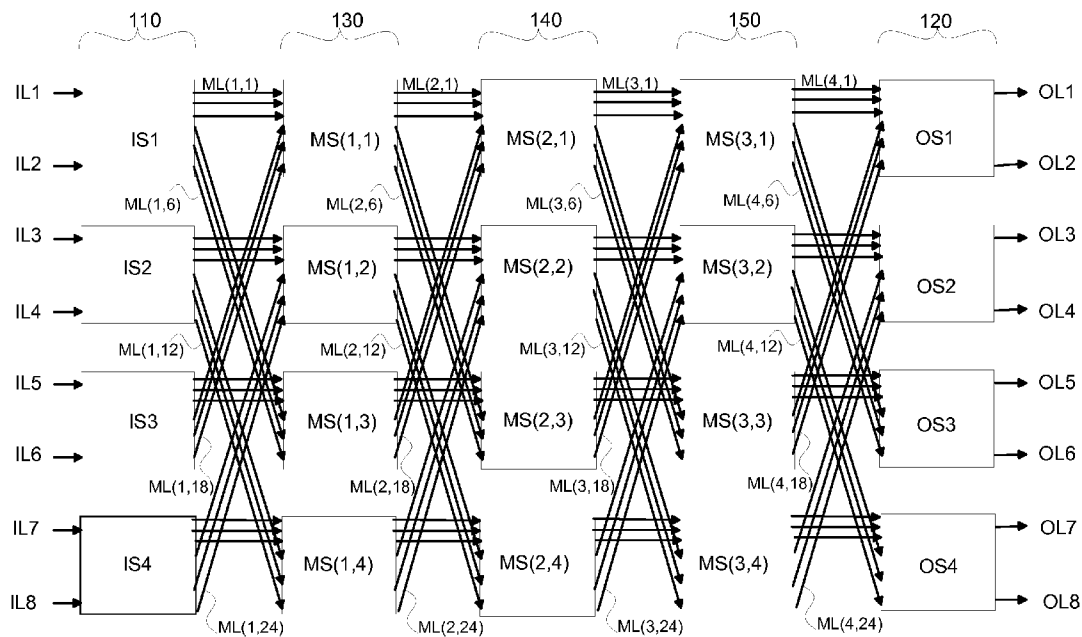


FIG. 3I

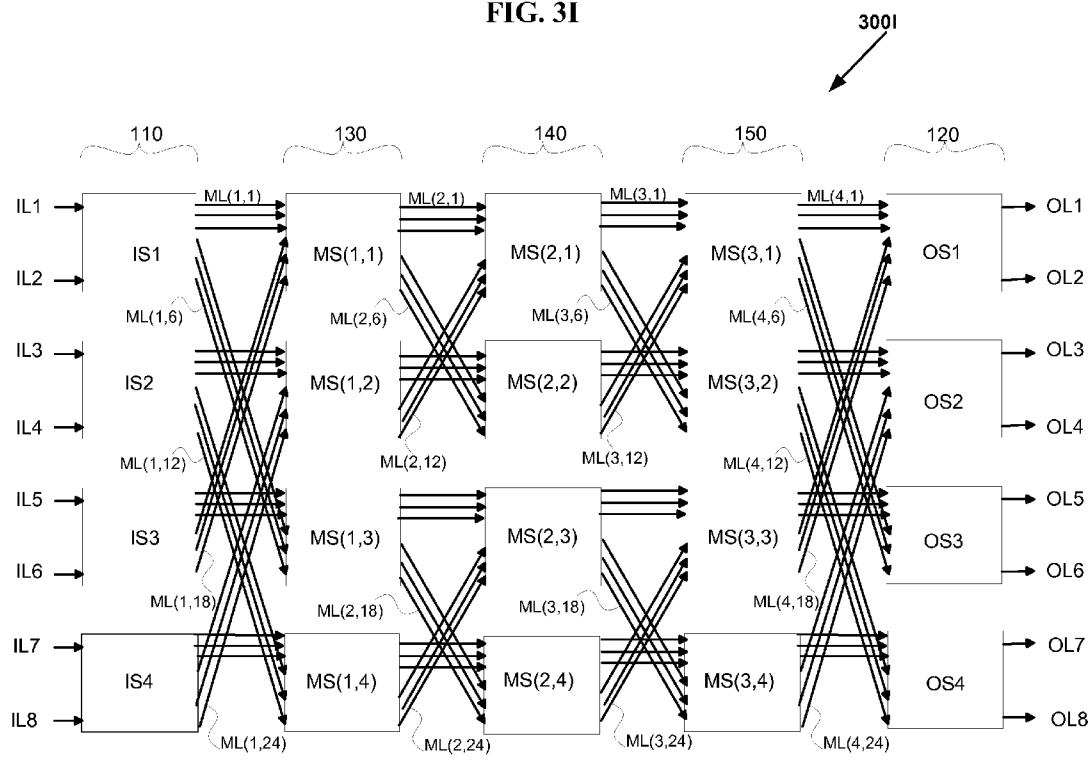


FIG. 3J

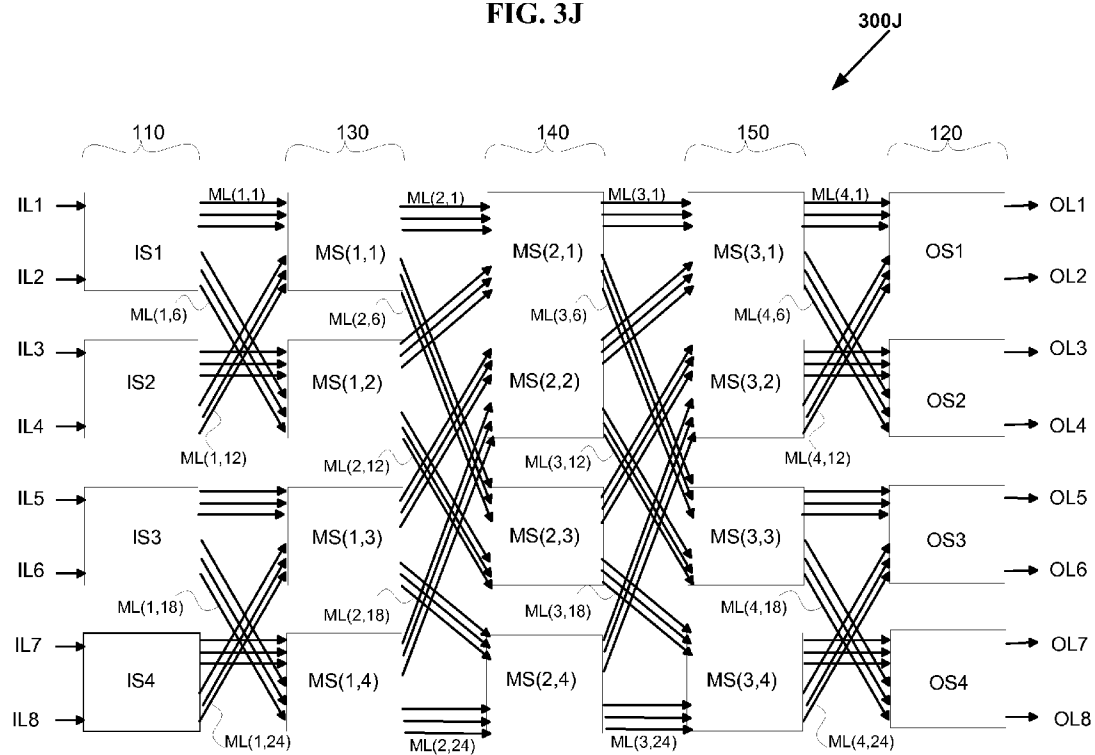


FIG. 3K

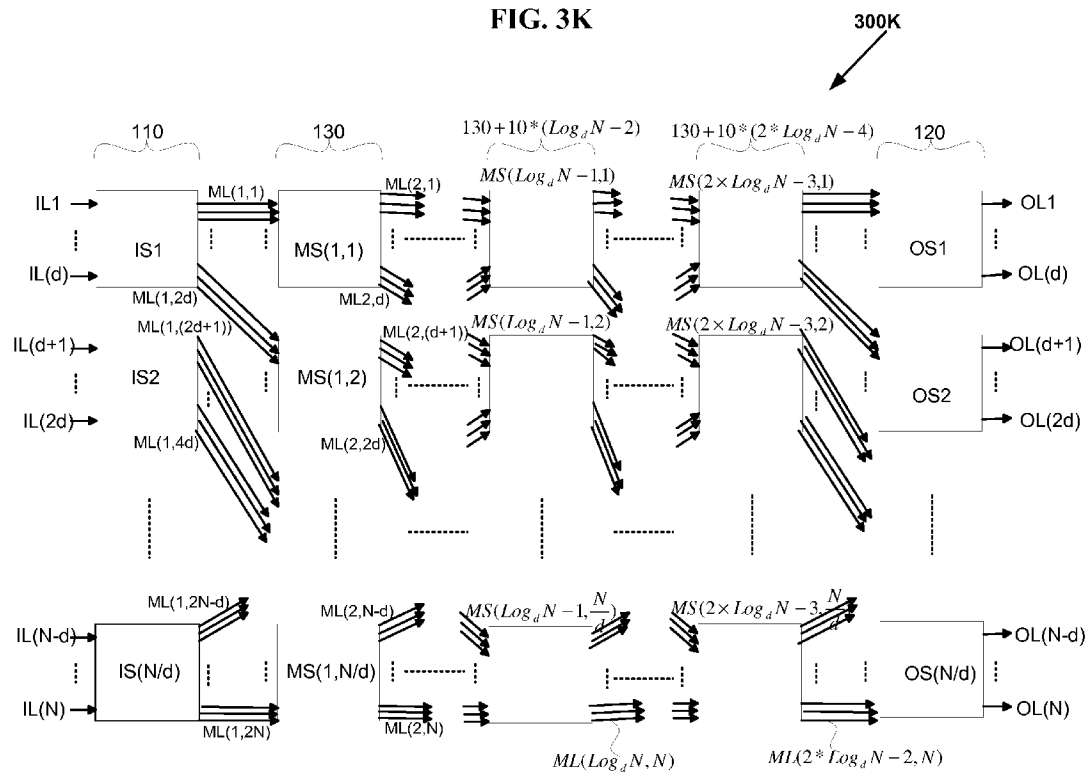


FIG. 3A1

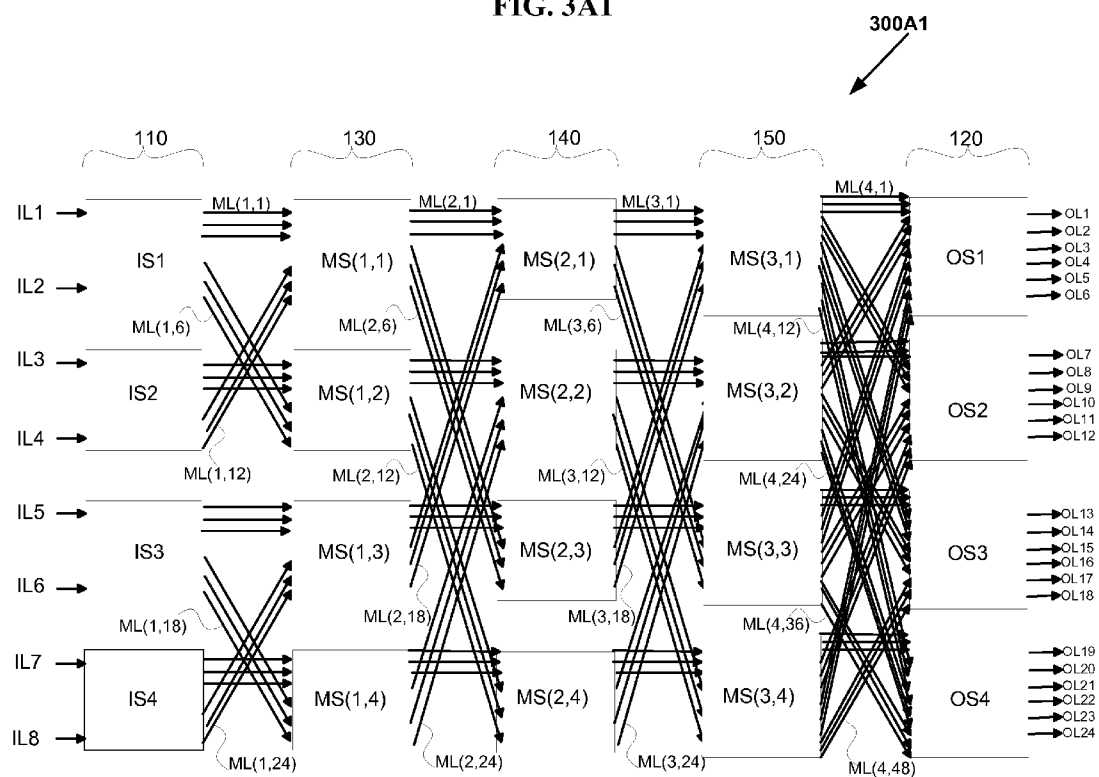


FIG. 3B1

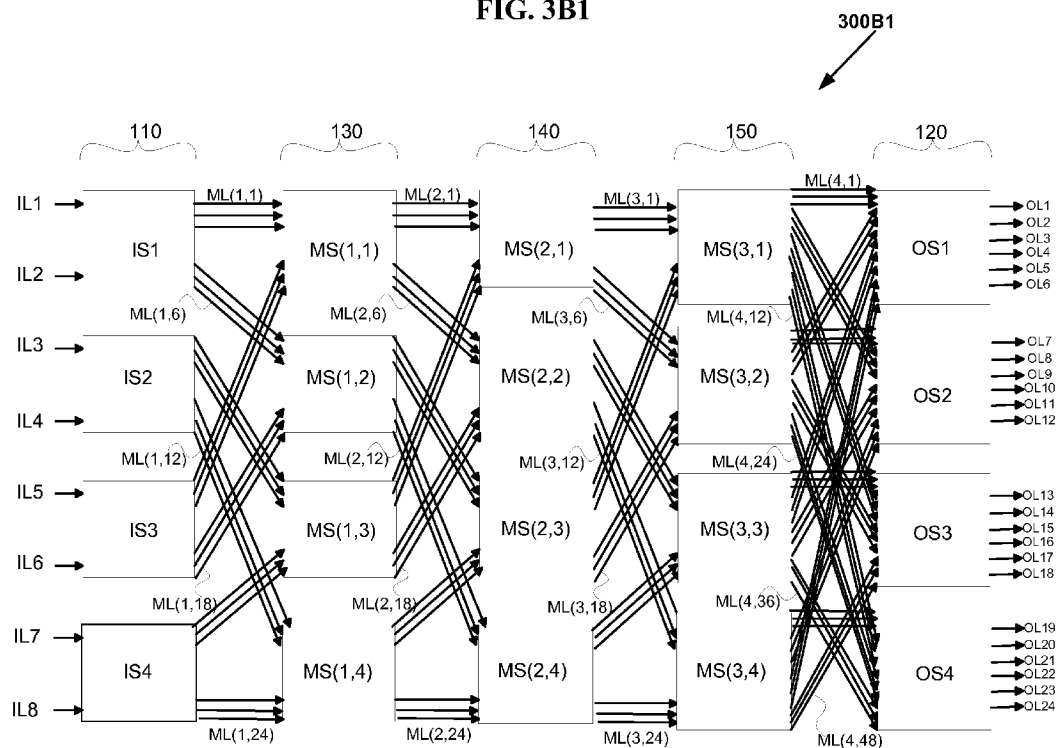


FIG. 3C1

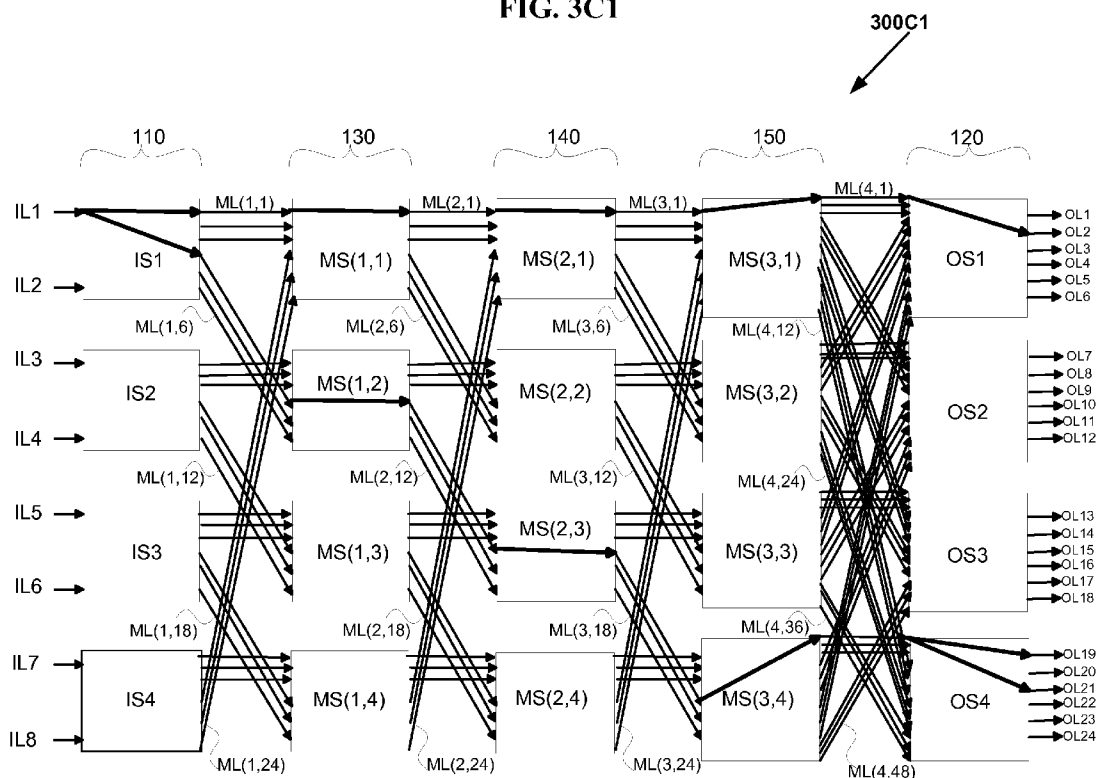


FIG. 3D1

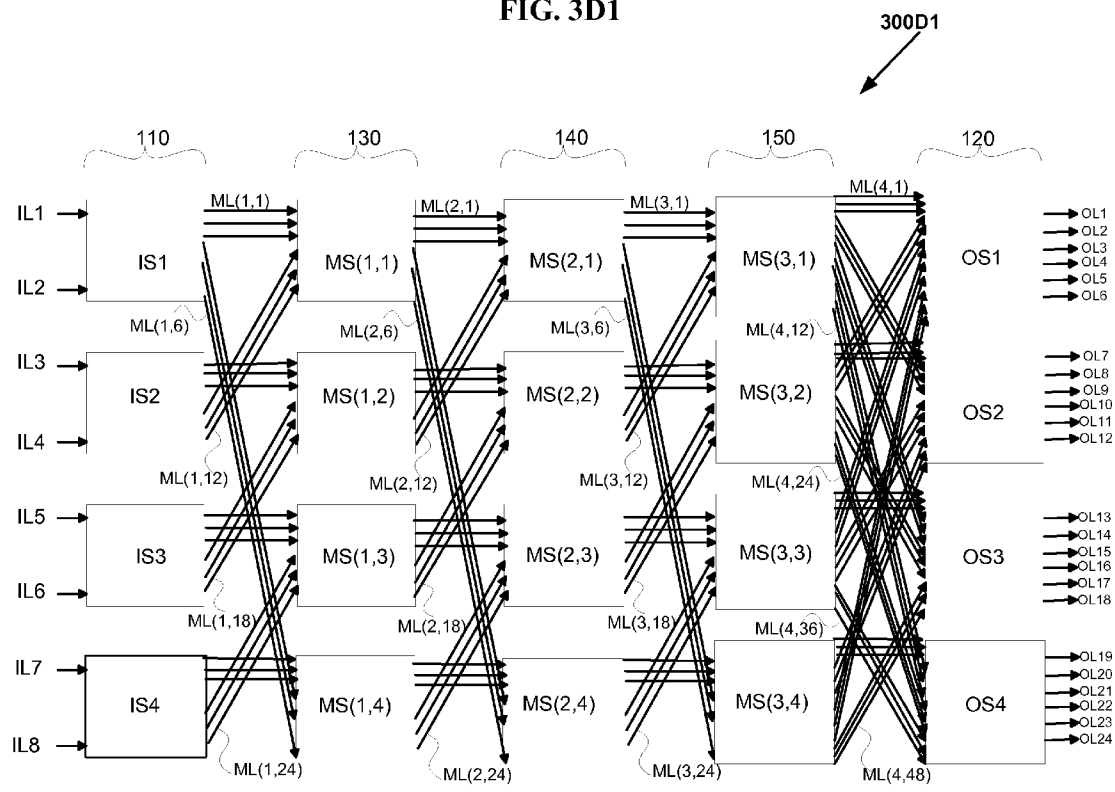


FIG. 3E1

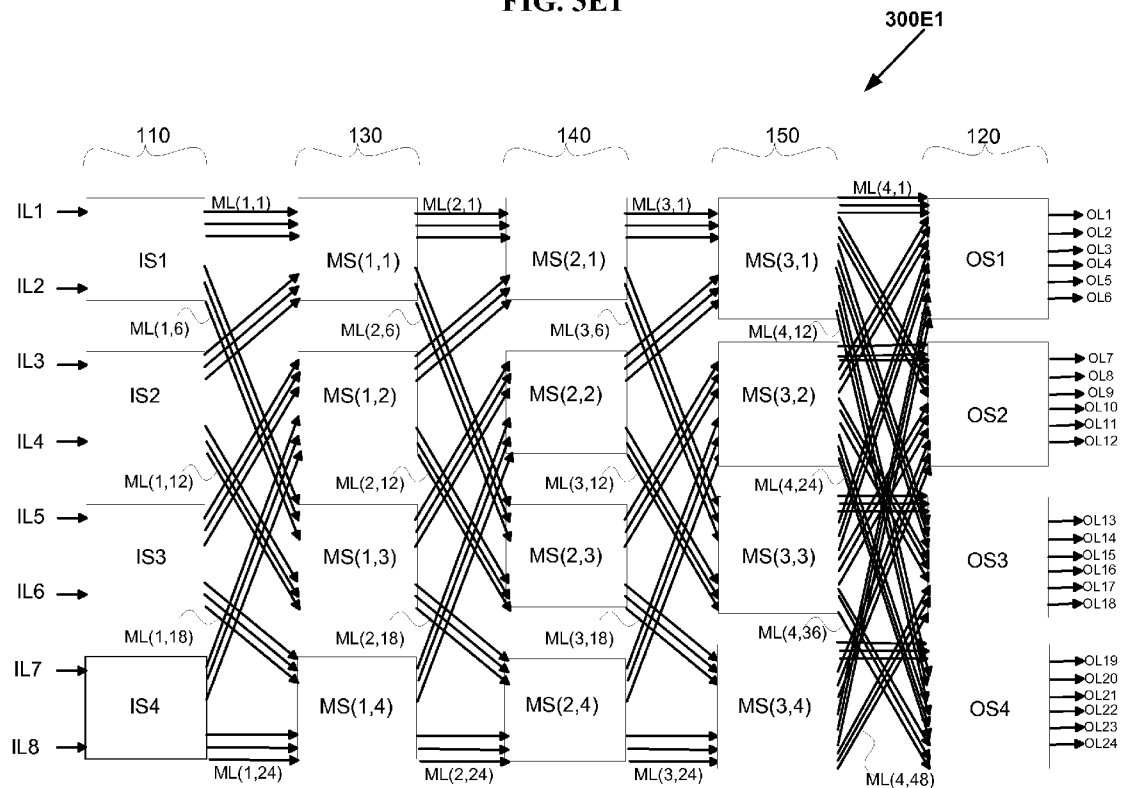


FIG. 3F1

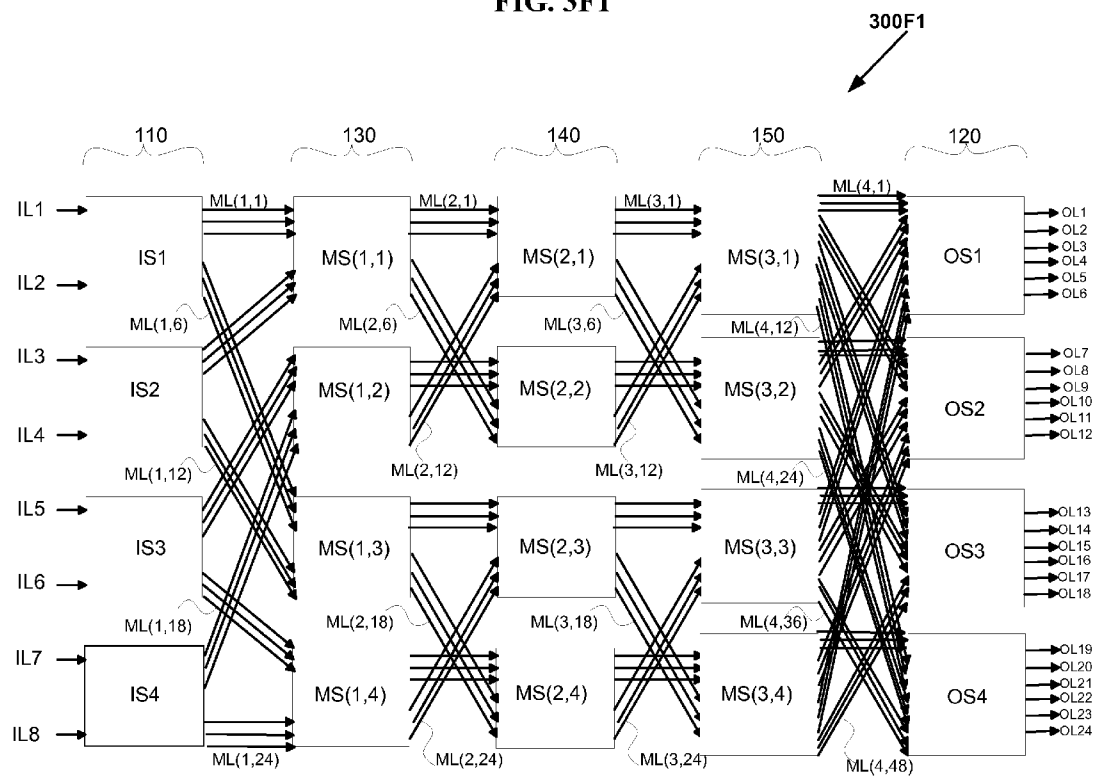


FIG. 3G1

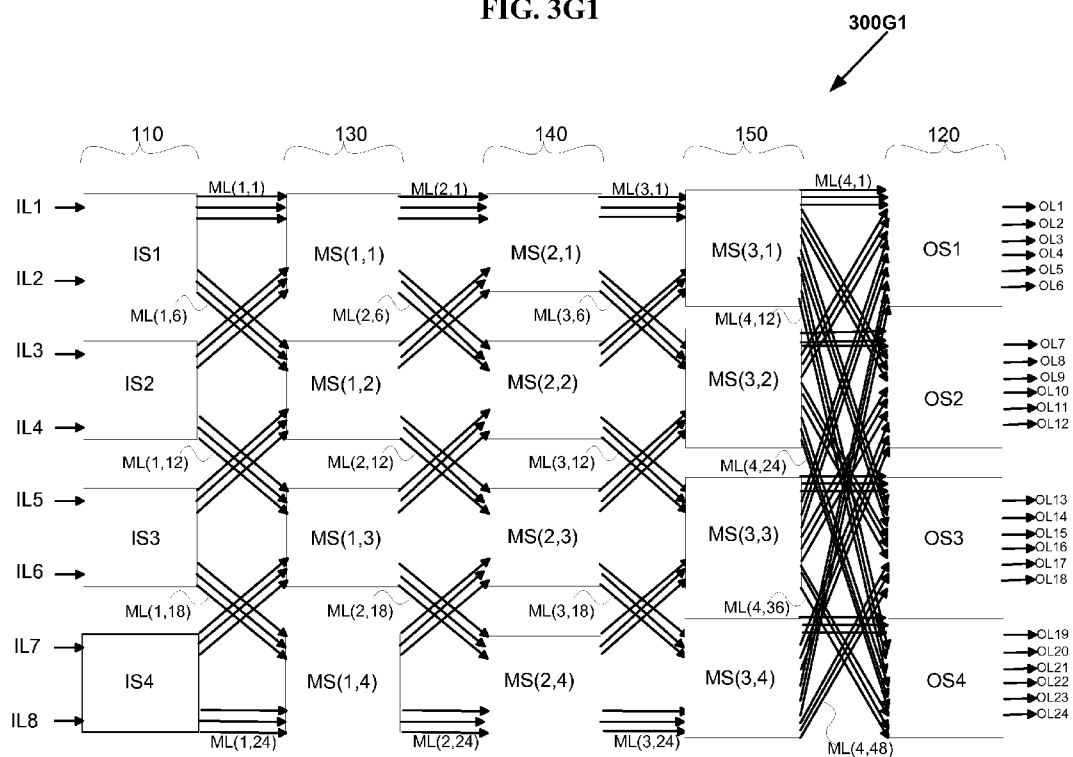


FIG. 3H1

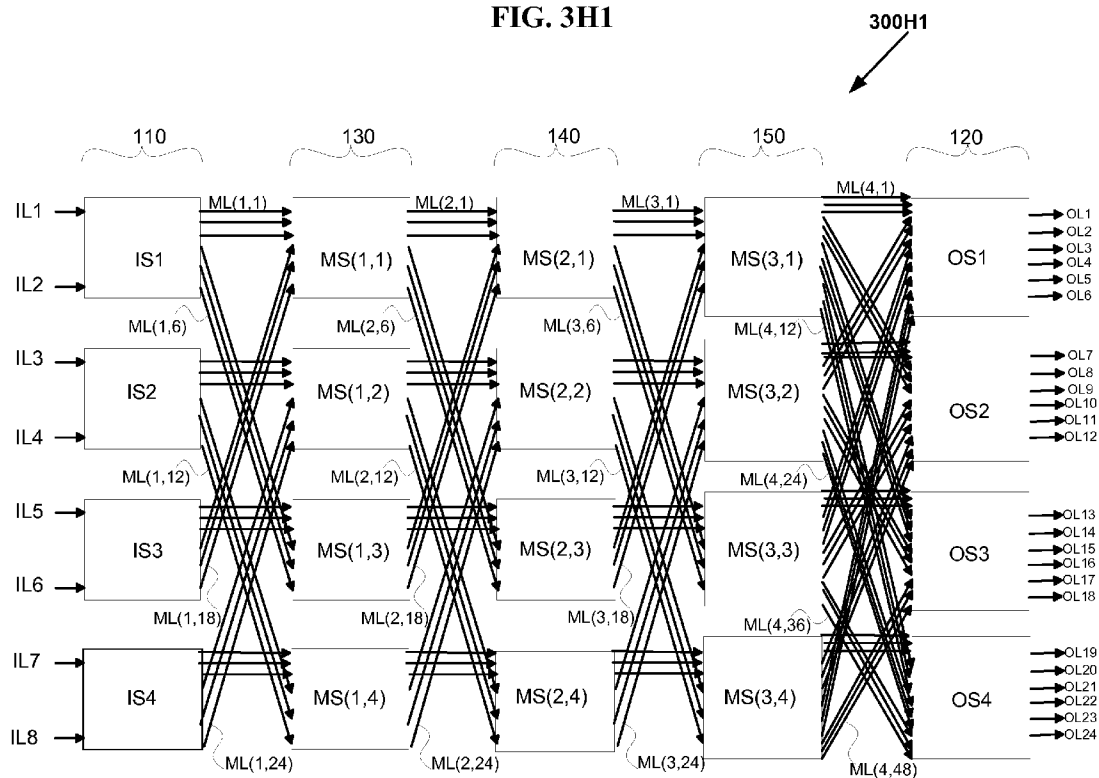


FIG. 3I1

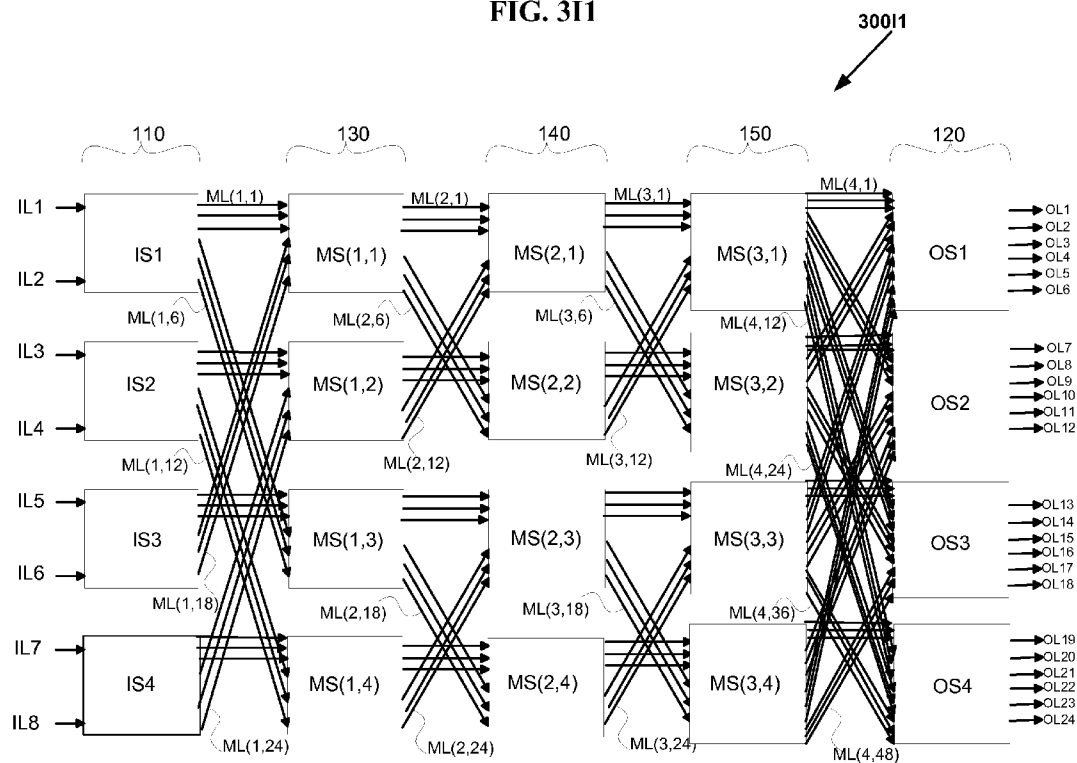


FIG. 3J1

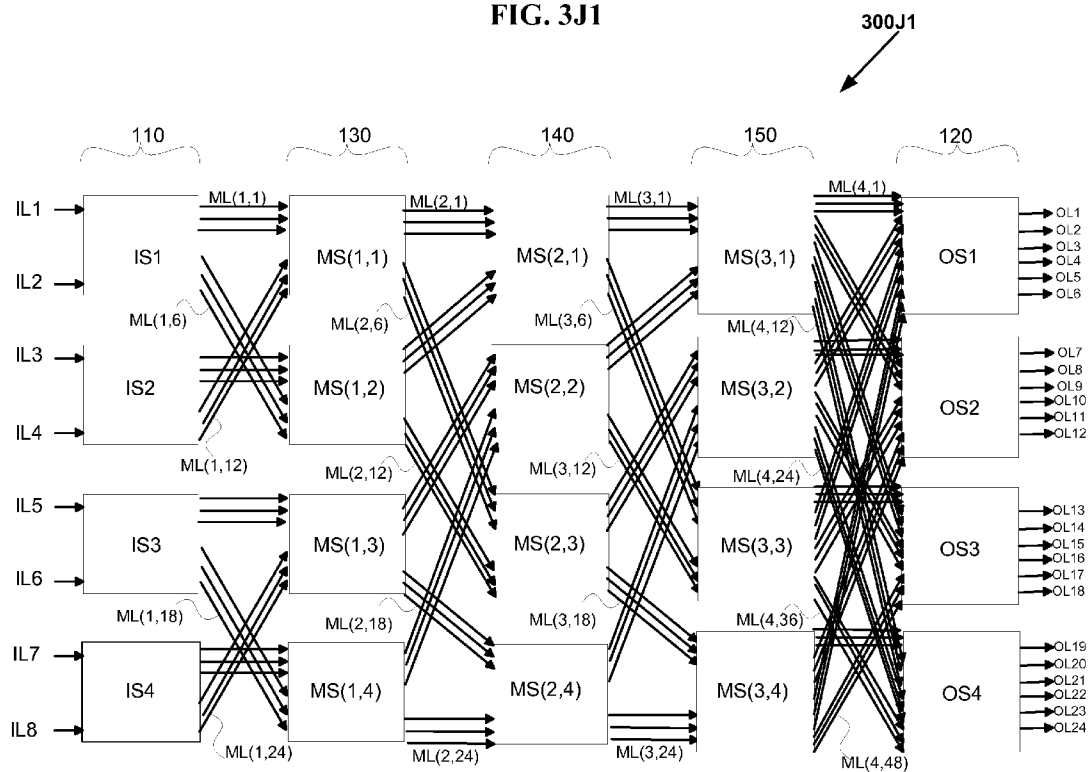


FIG. 3K1

300K1

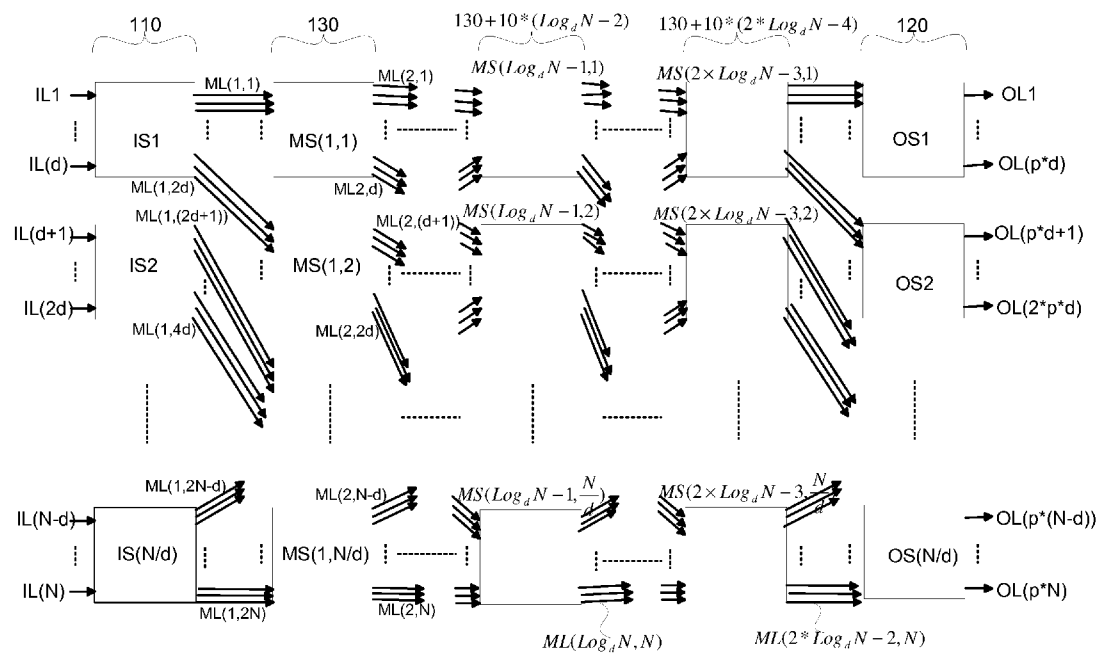


FIG. 3A2

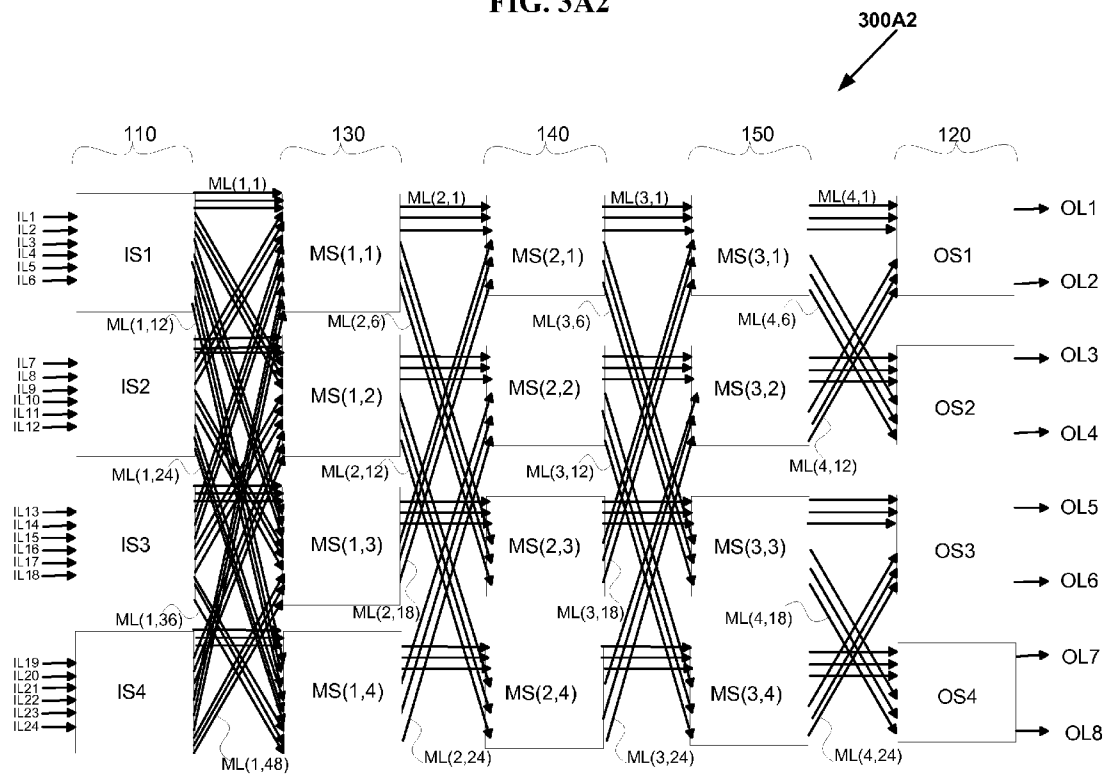


FIG. 3B2

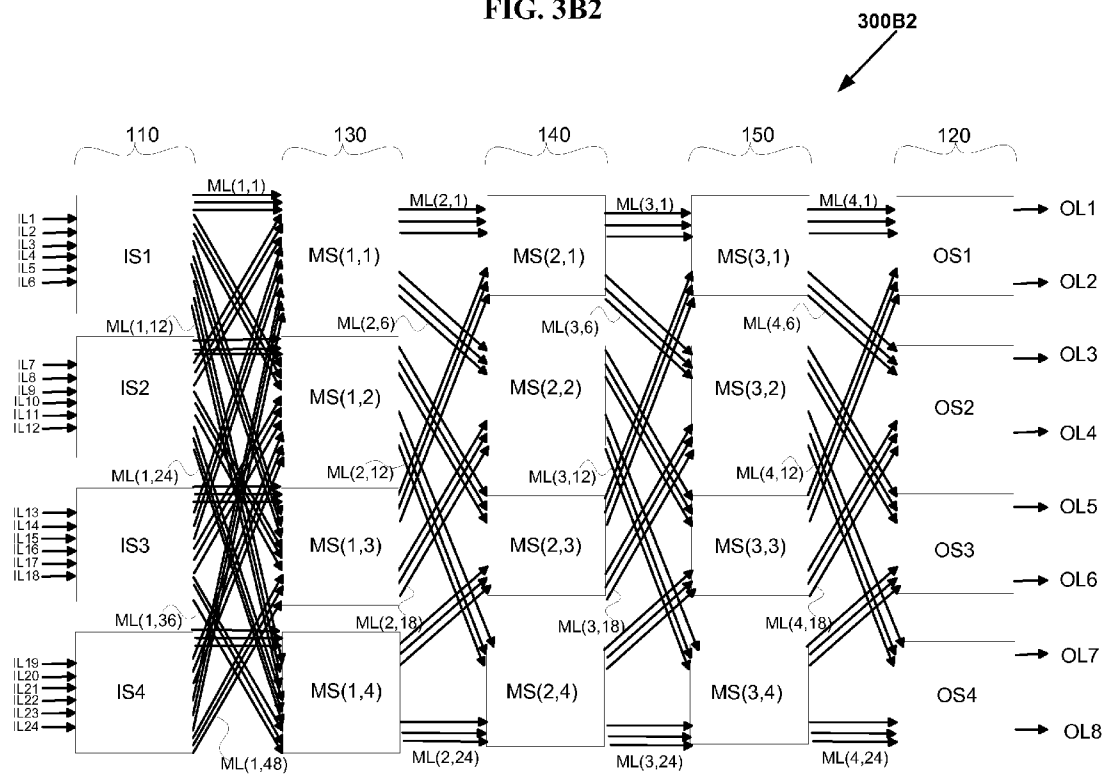


FIG. 3C2

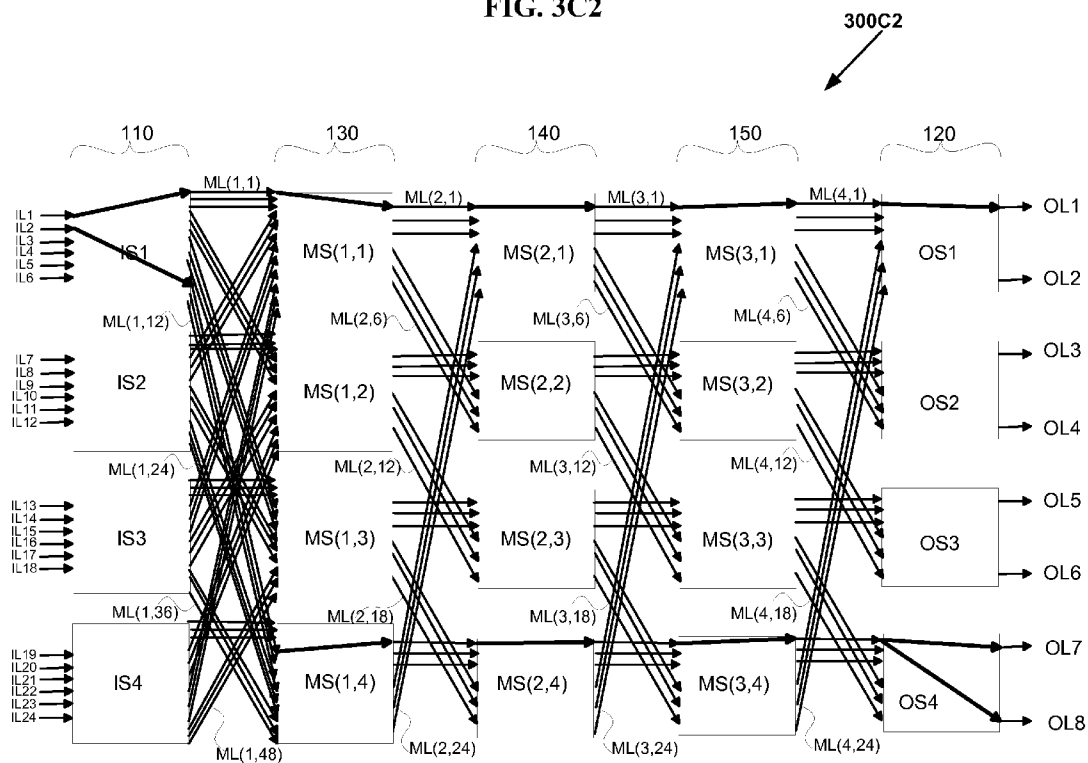


FIG. 3D2

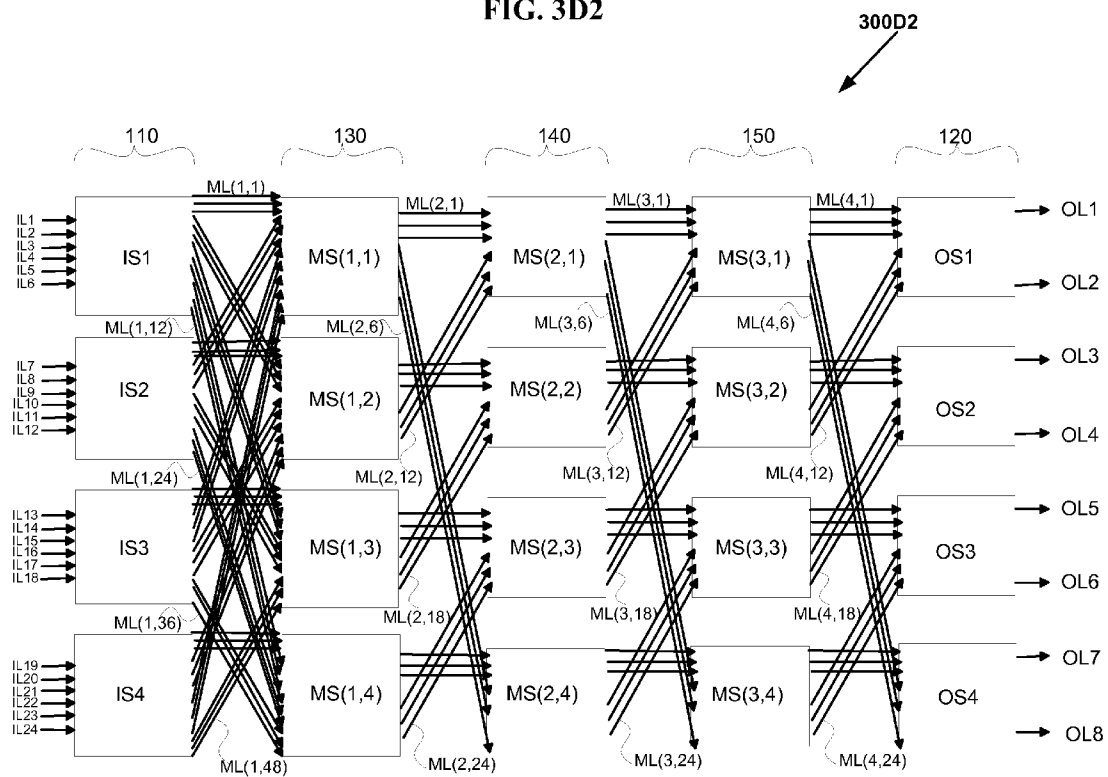


FIG. 3E2

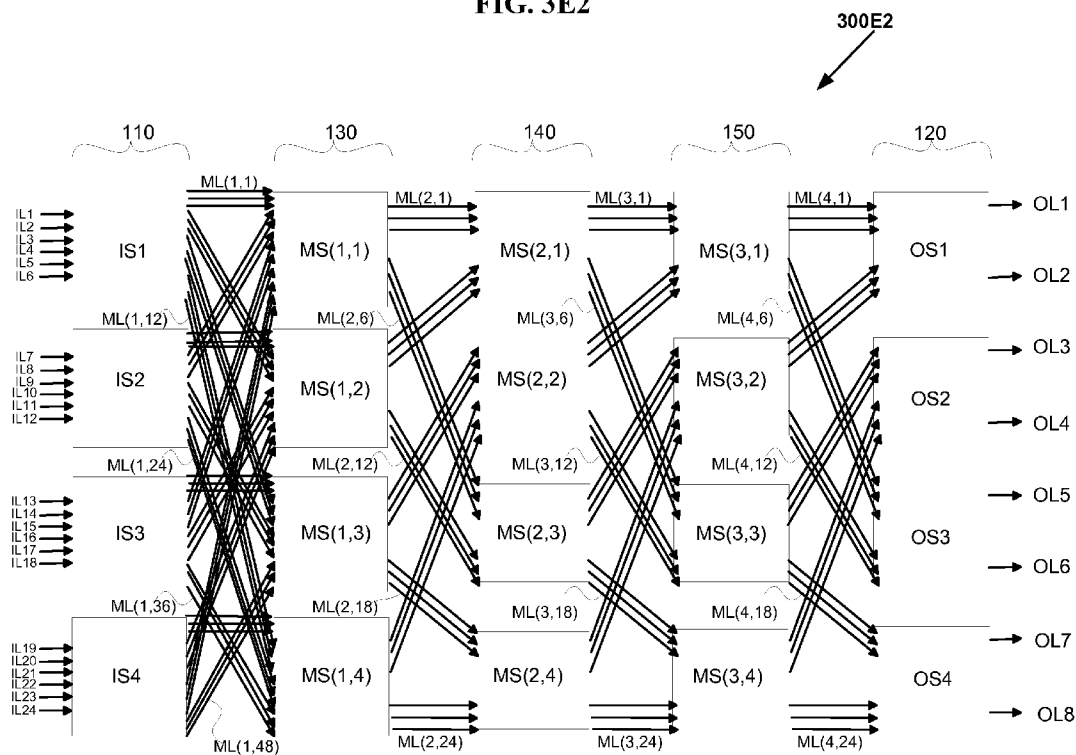


FIG. 3F2

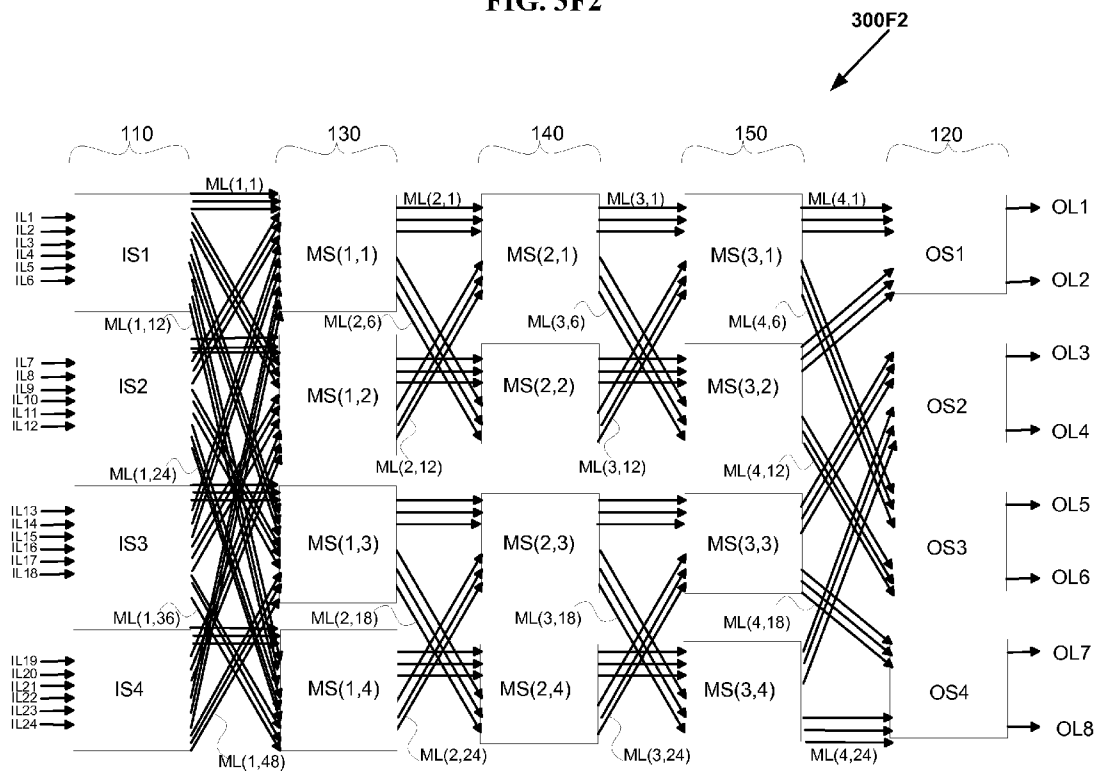


FIG. 3G2

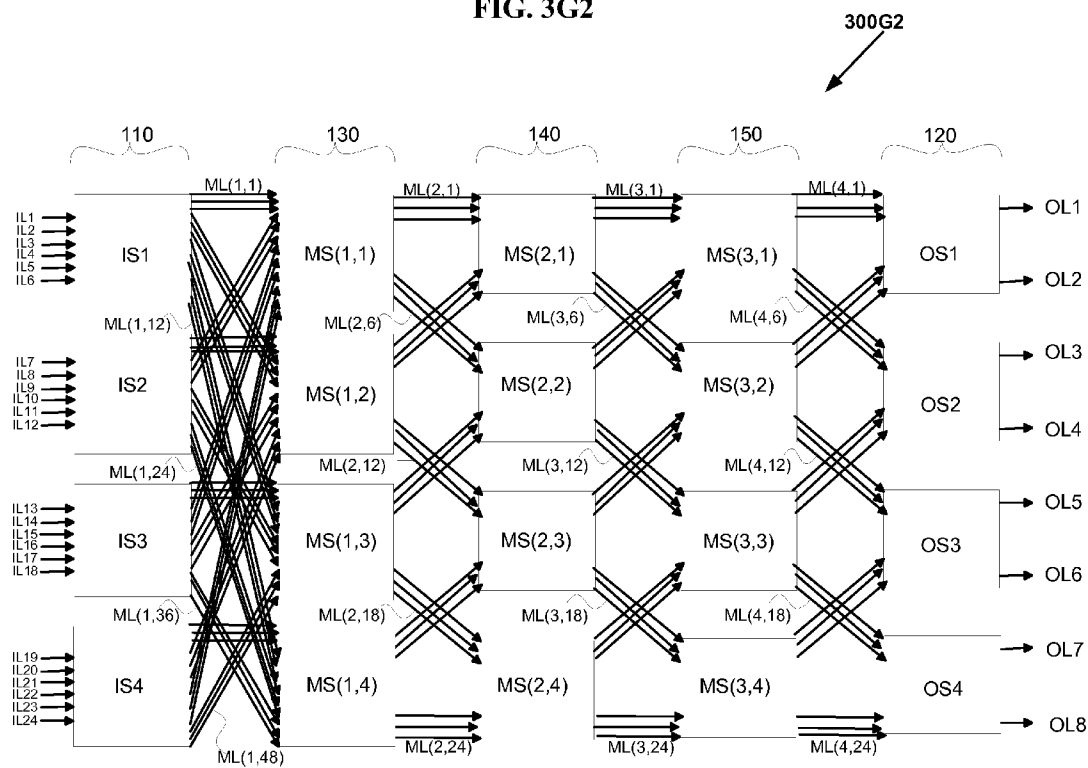


FIG. 3H2

300H2

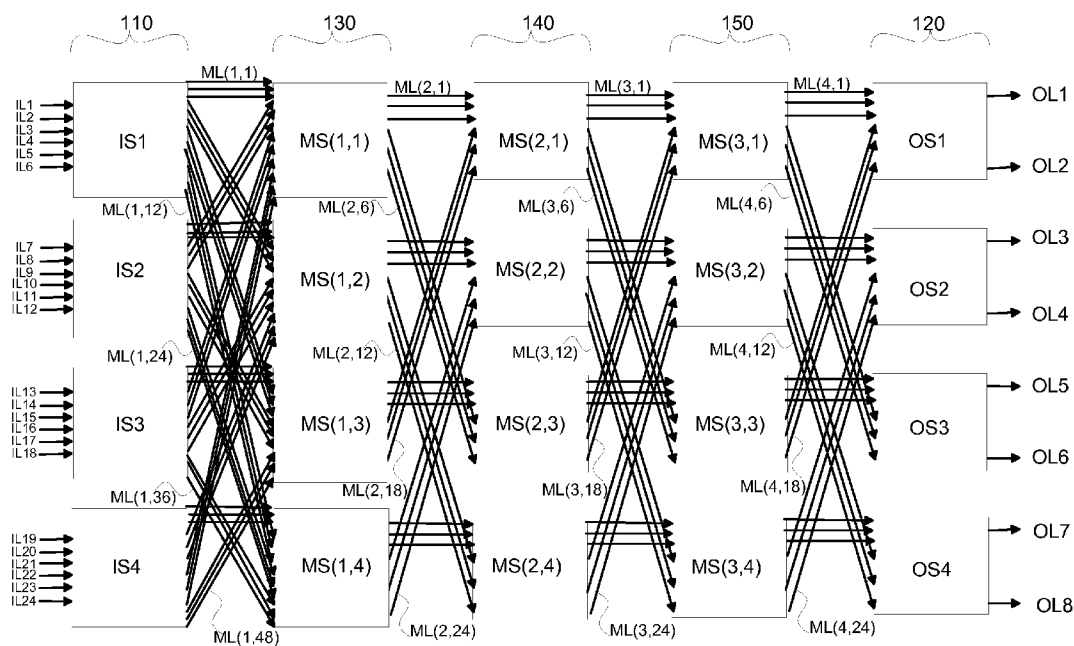


FIG. 3I2

30012

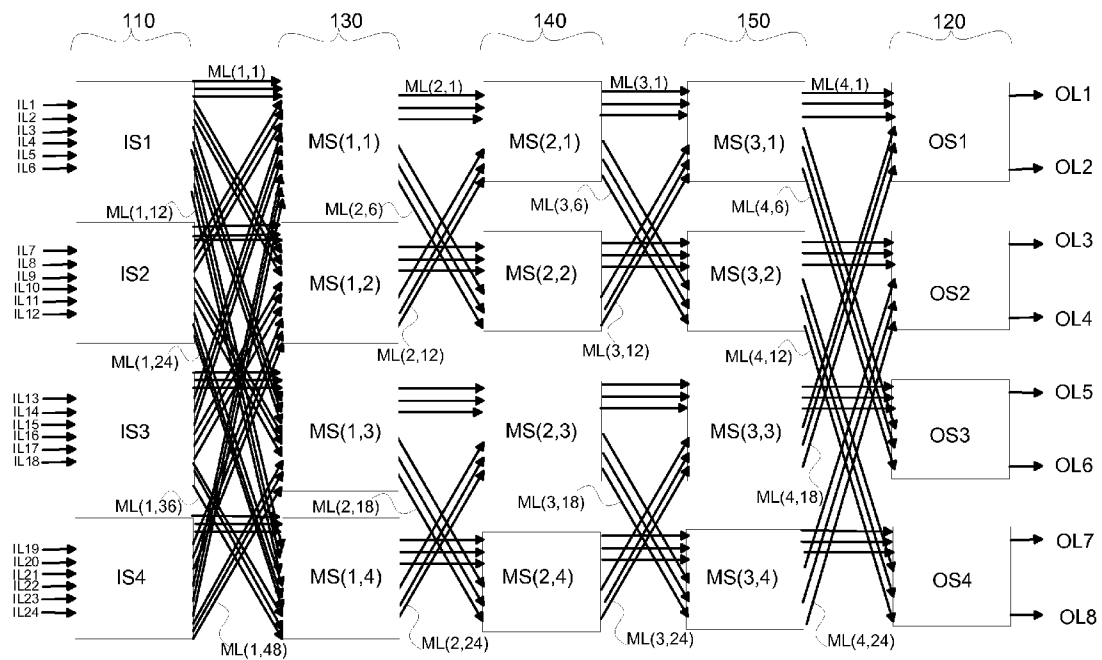


FIG. 3J2

300J2

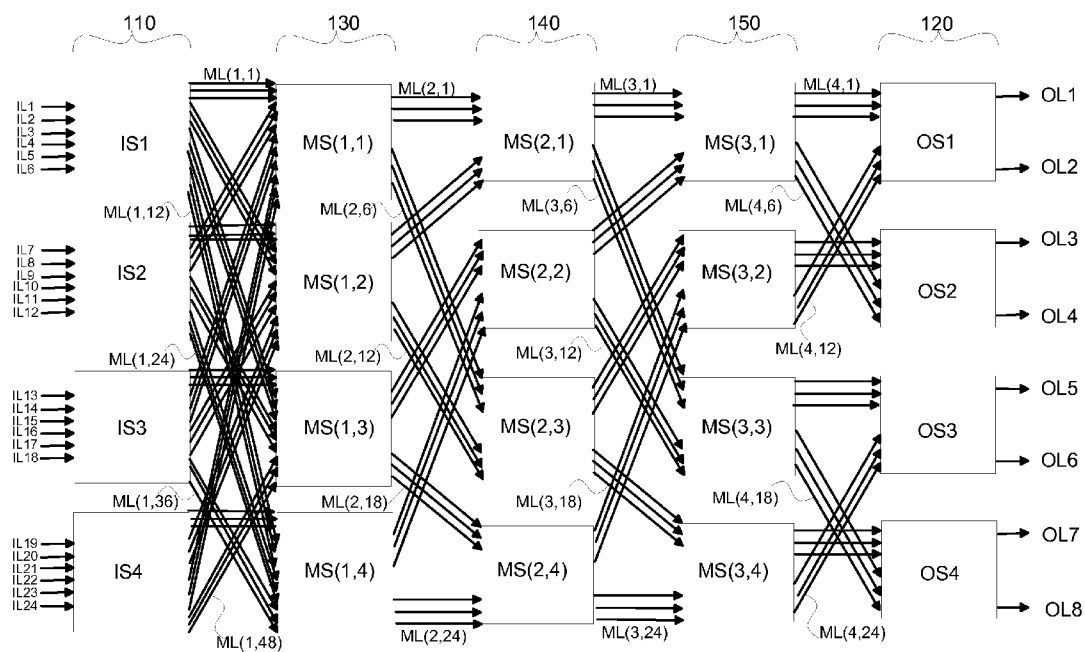


FIG. 3K2

300K2

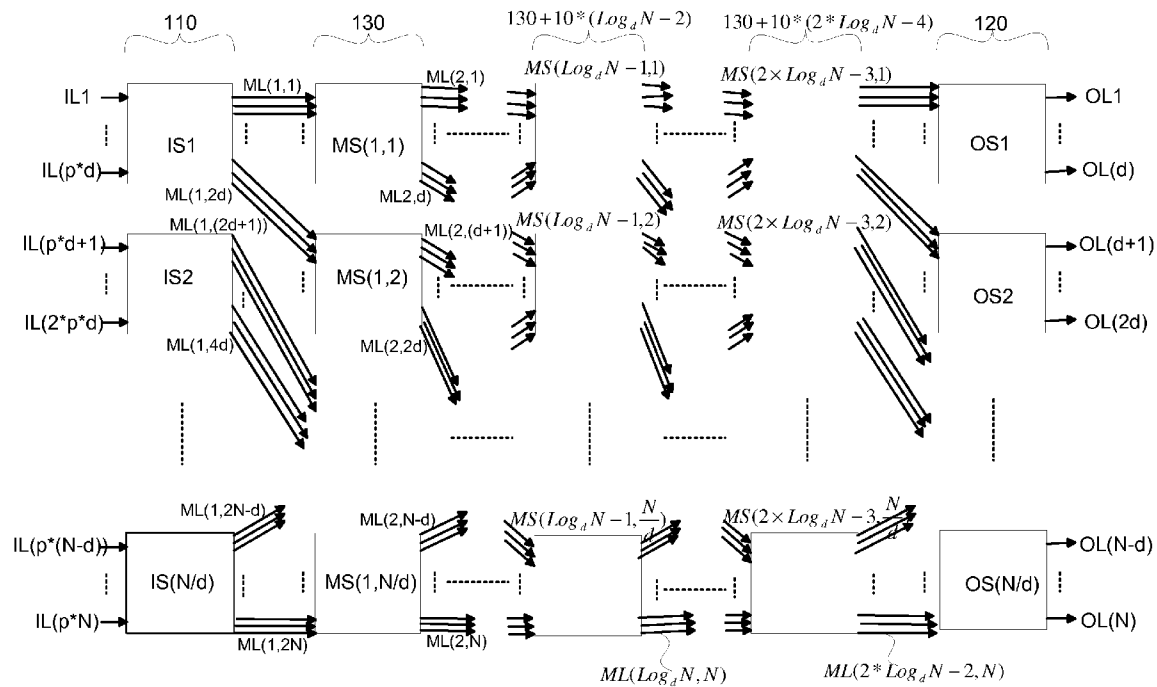


FIG. 4A

400A

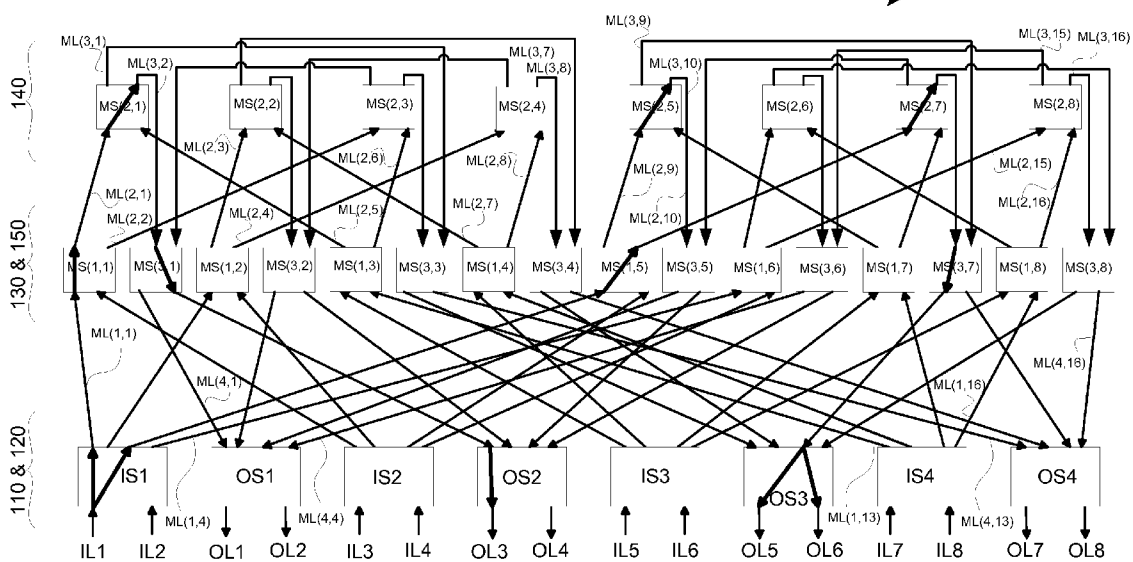


FIG. 4A1

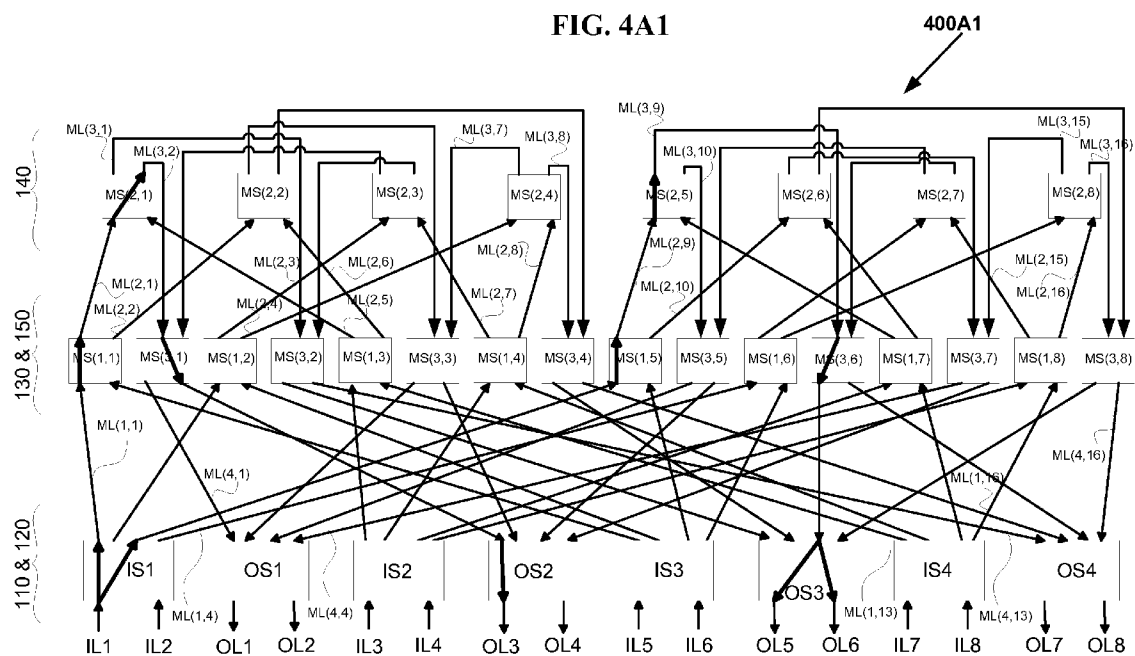


FIG. 4A2

400A2

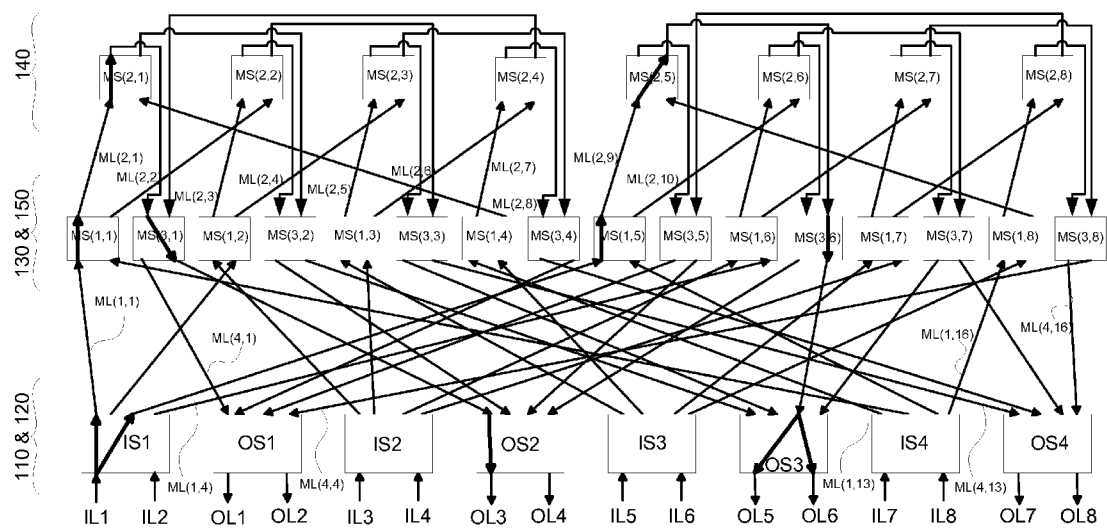


FIG. 4B

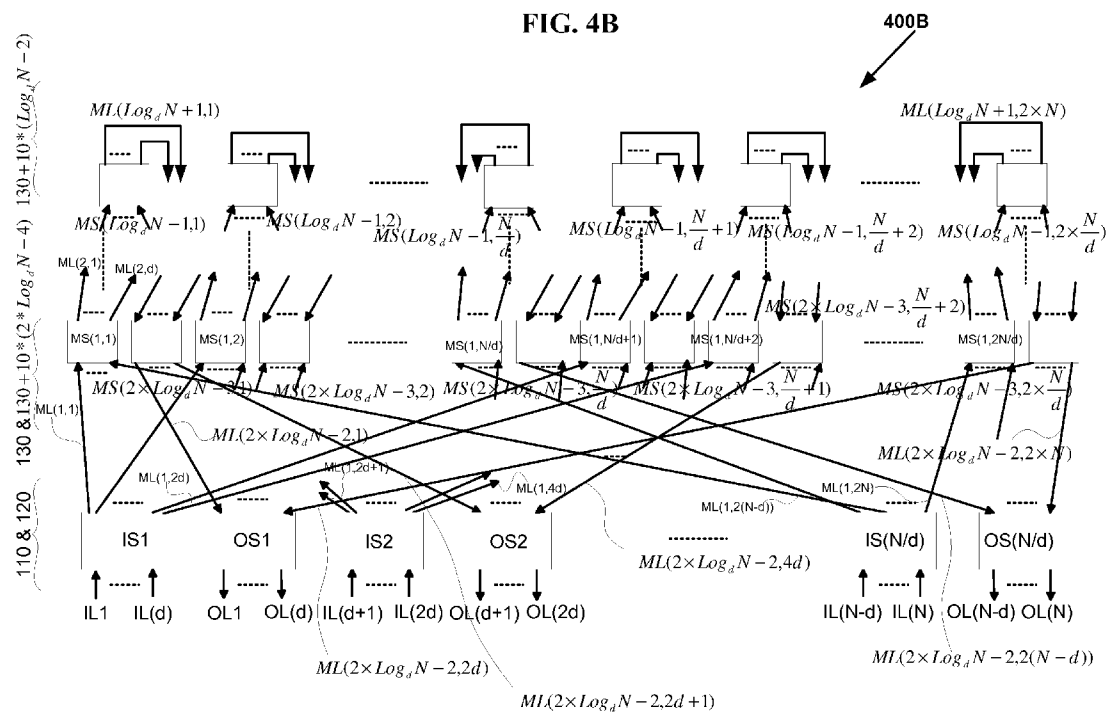


FIG. 4C

400C

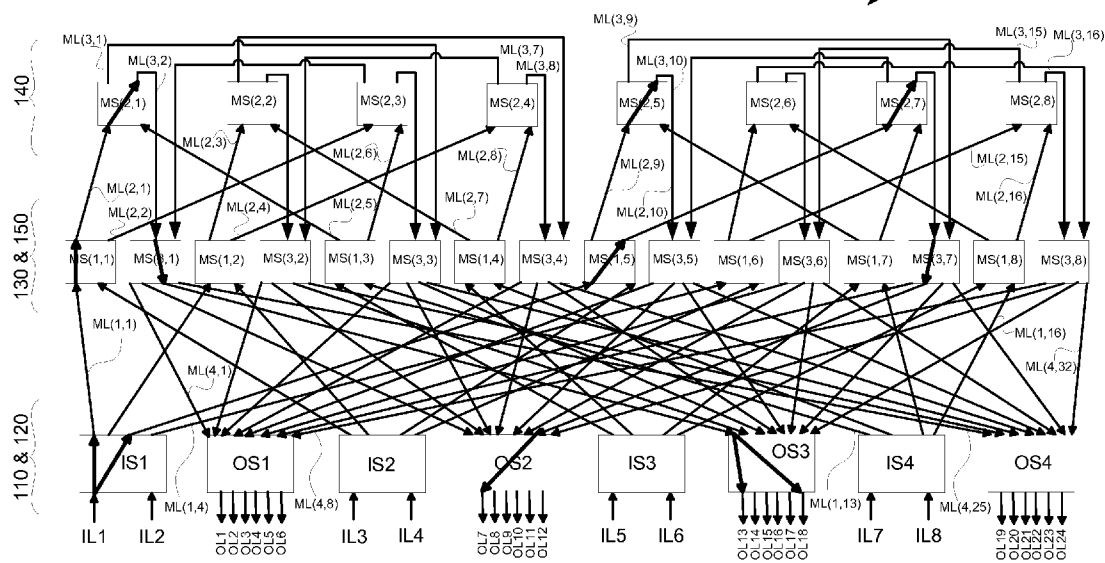


FIG. 4C1

400C1

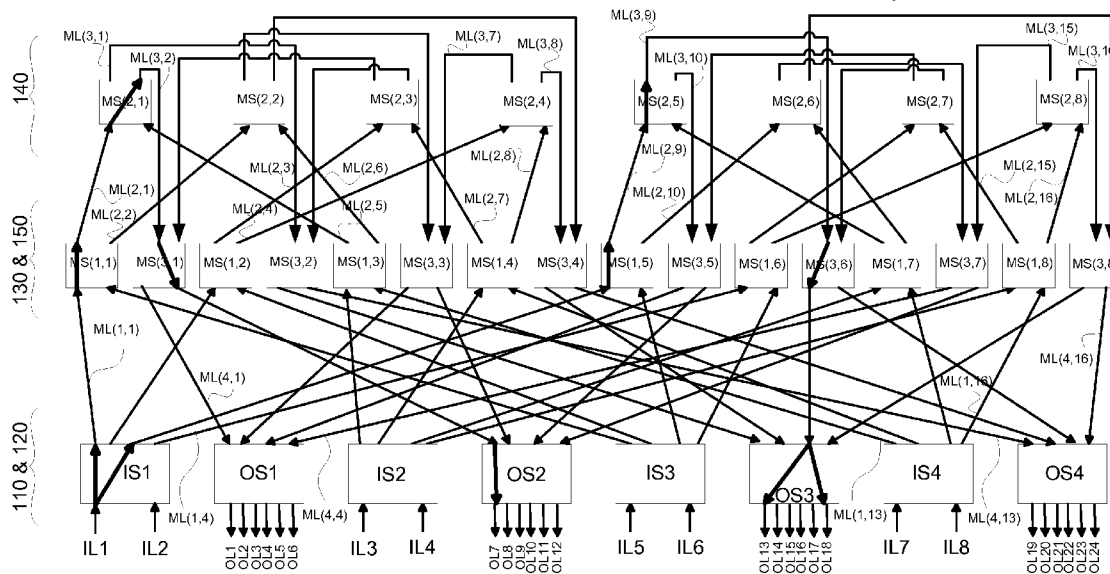


FIG. 4C2

400C2

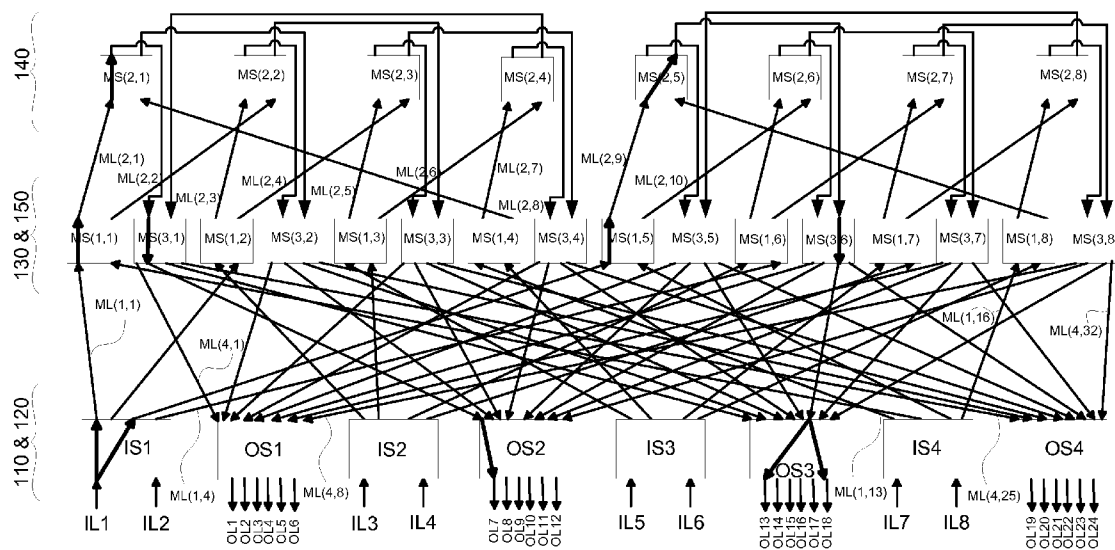


FIG. 4D

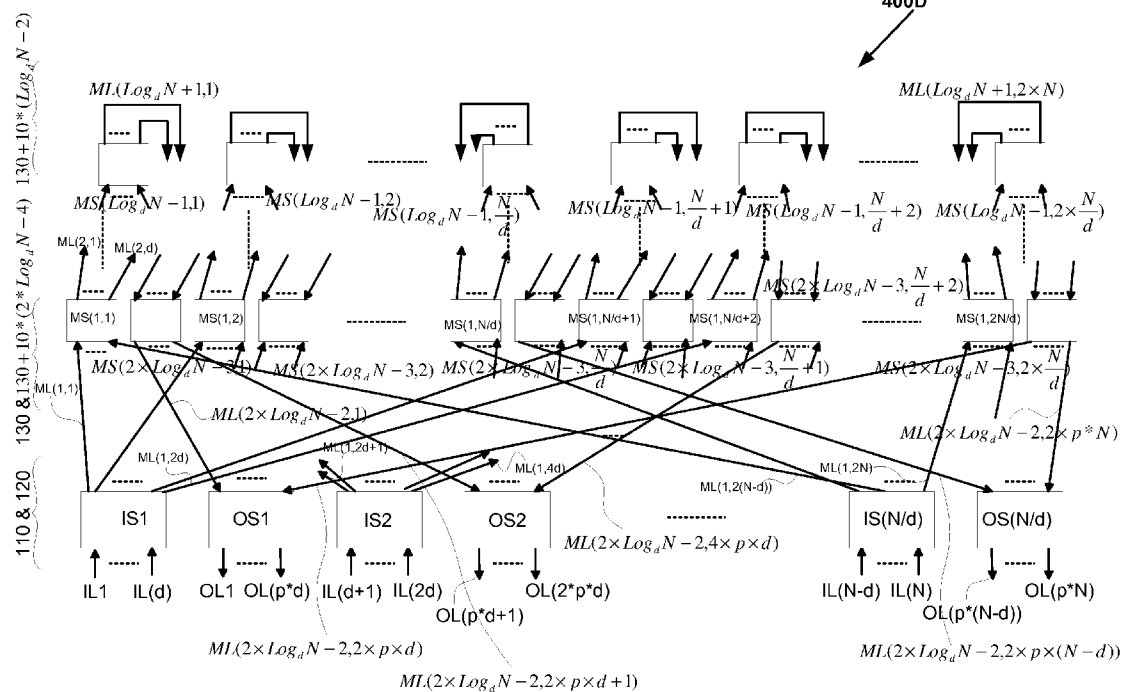


FIG. 4E

400E

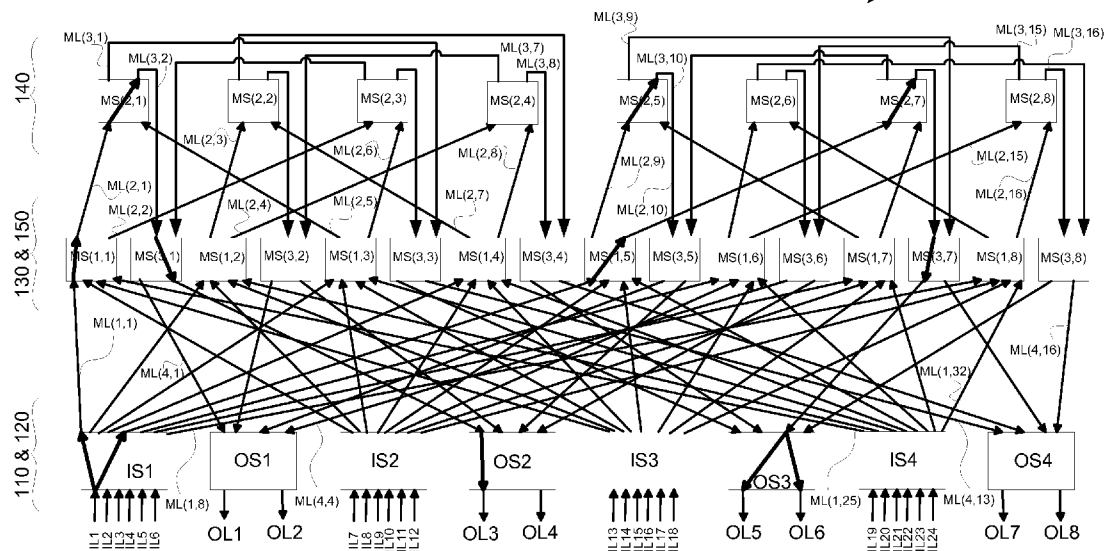


FIG. 4E1

400E1

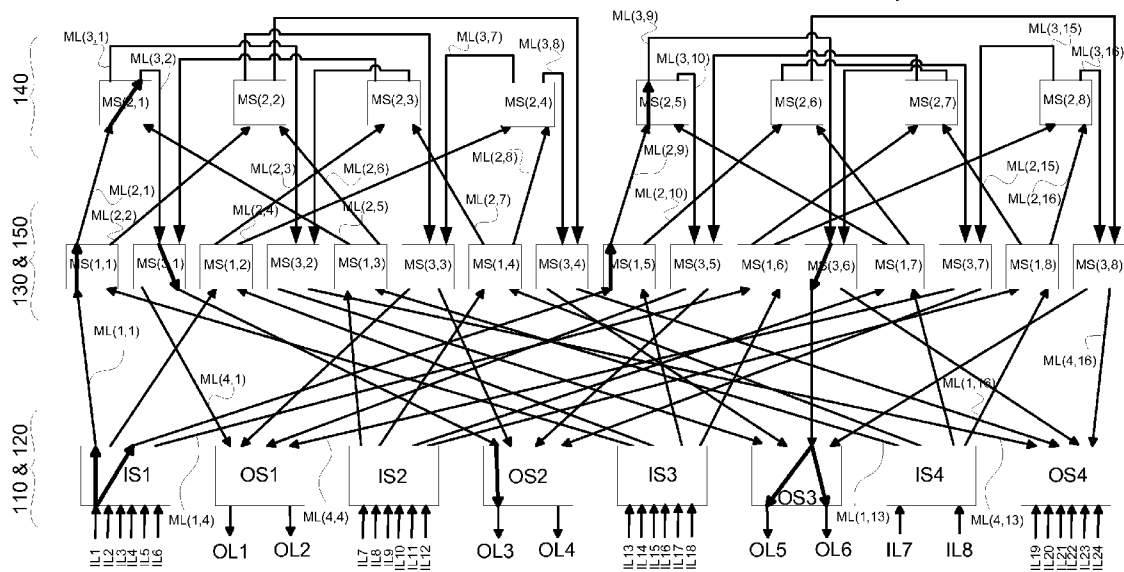


FIG. 4E2

400E2

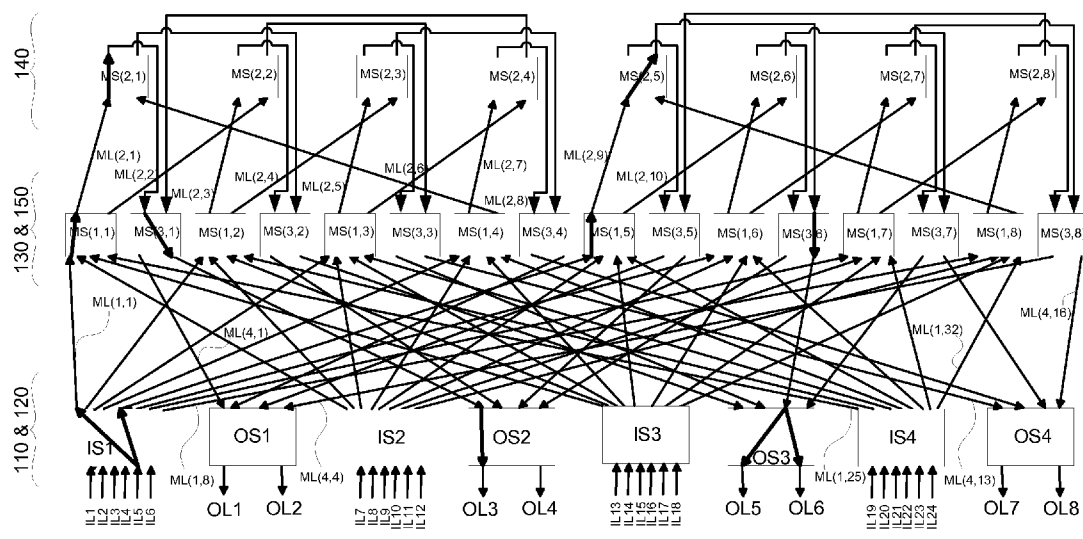


FIG. 4F

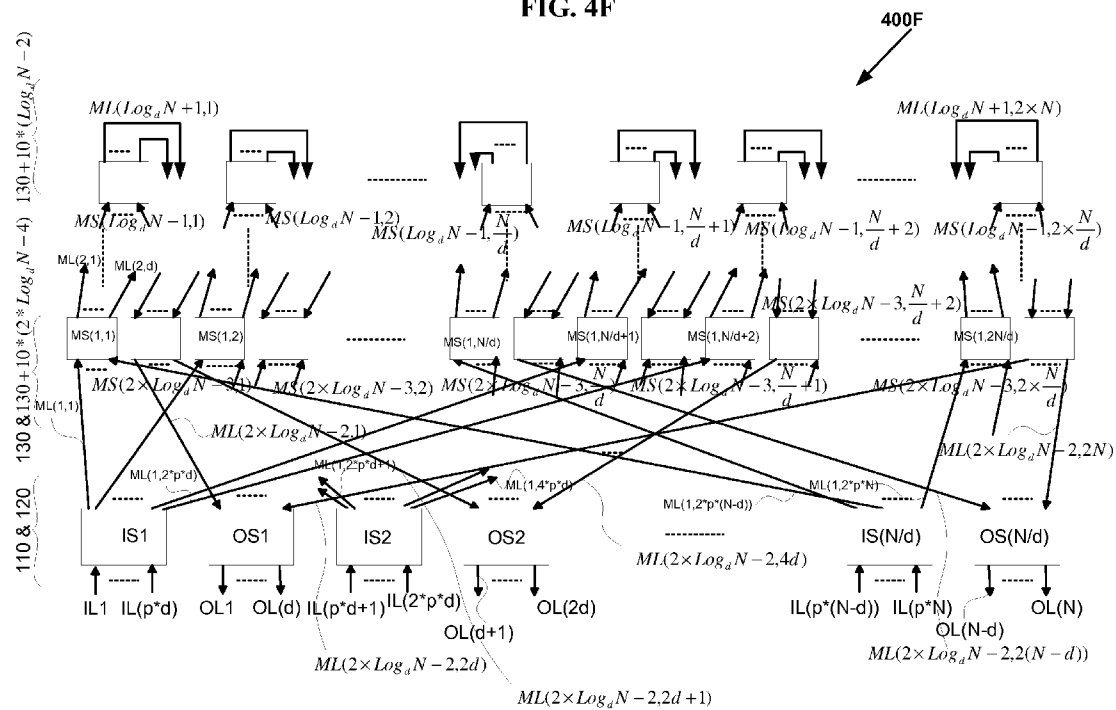
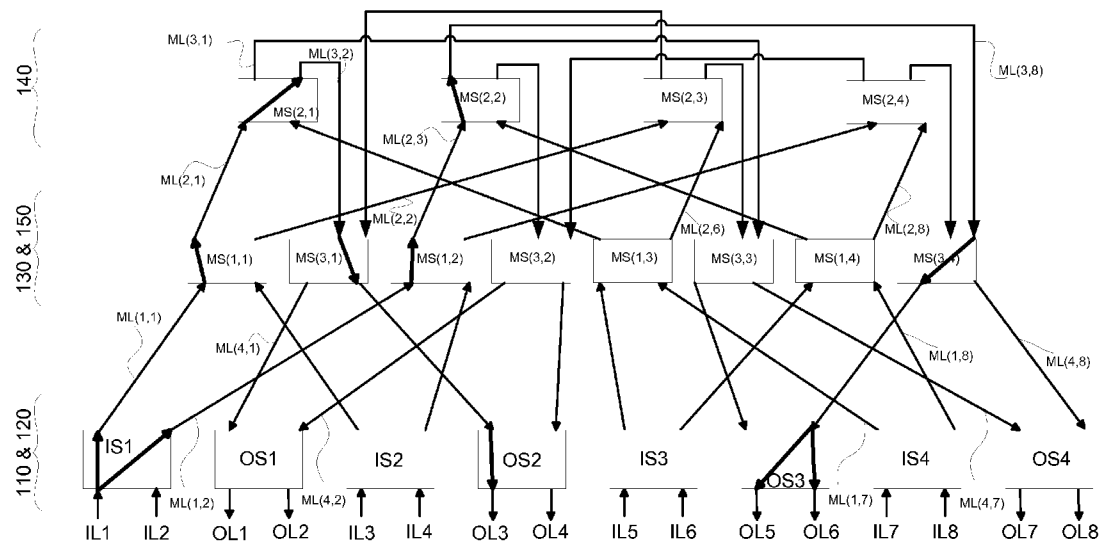


FIG. 5A

500A



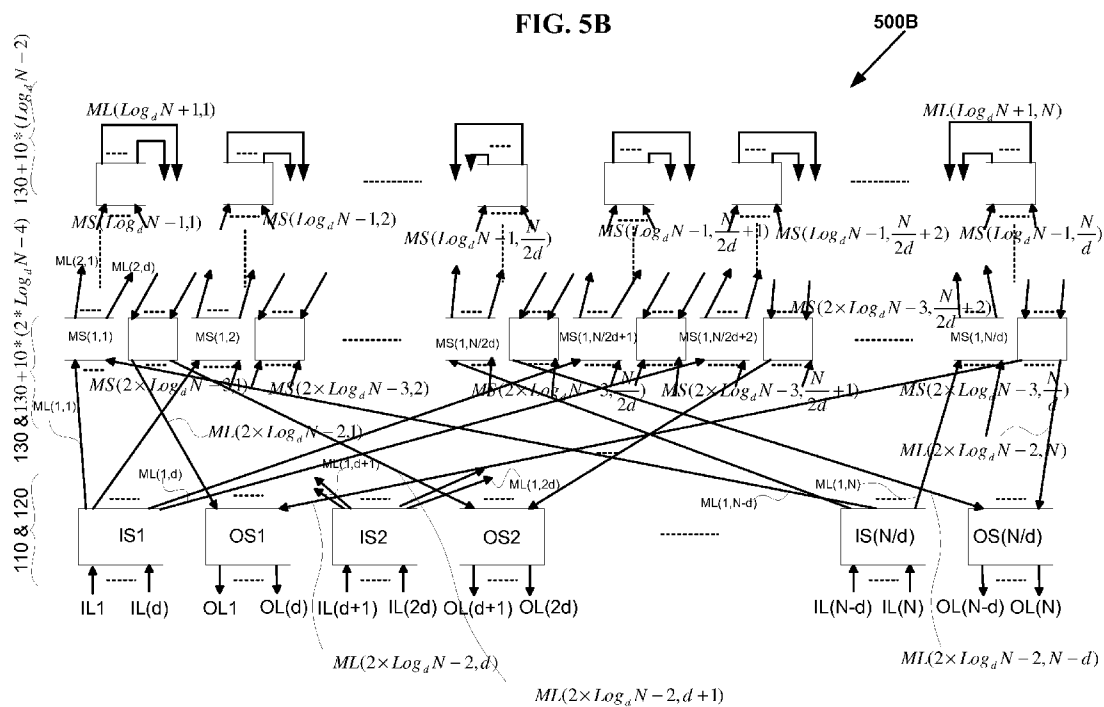


FIG. 5C

500C

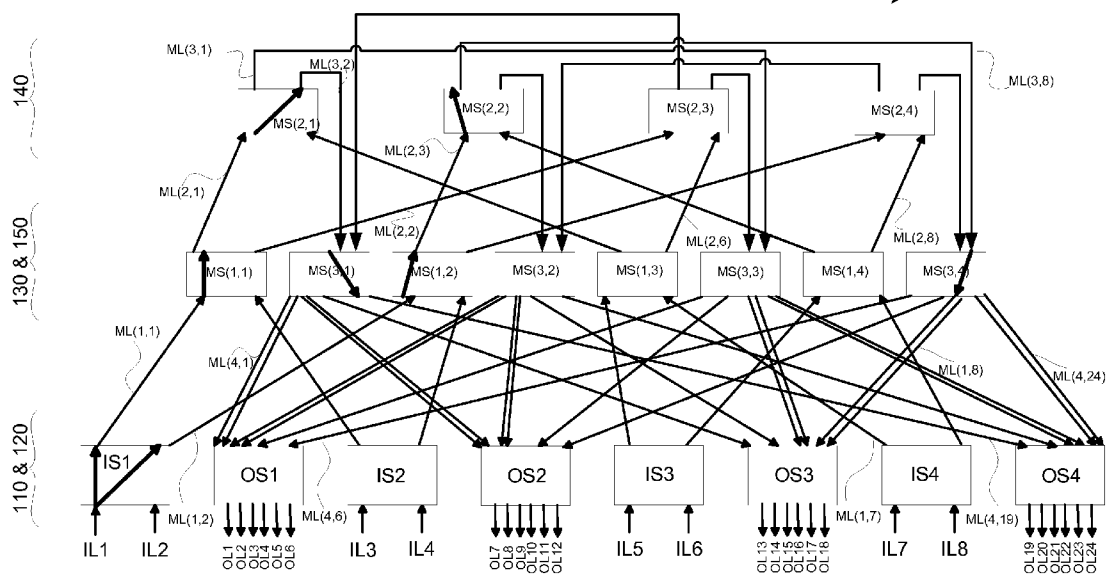


FIG. 5D

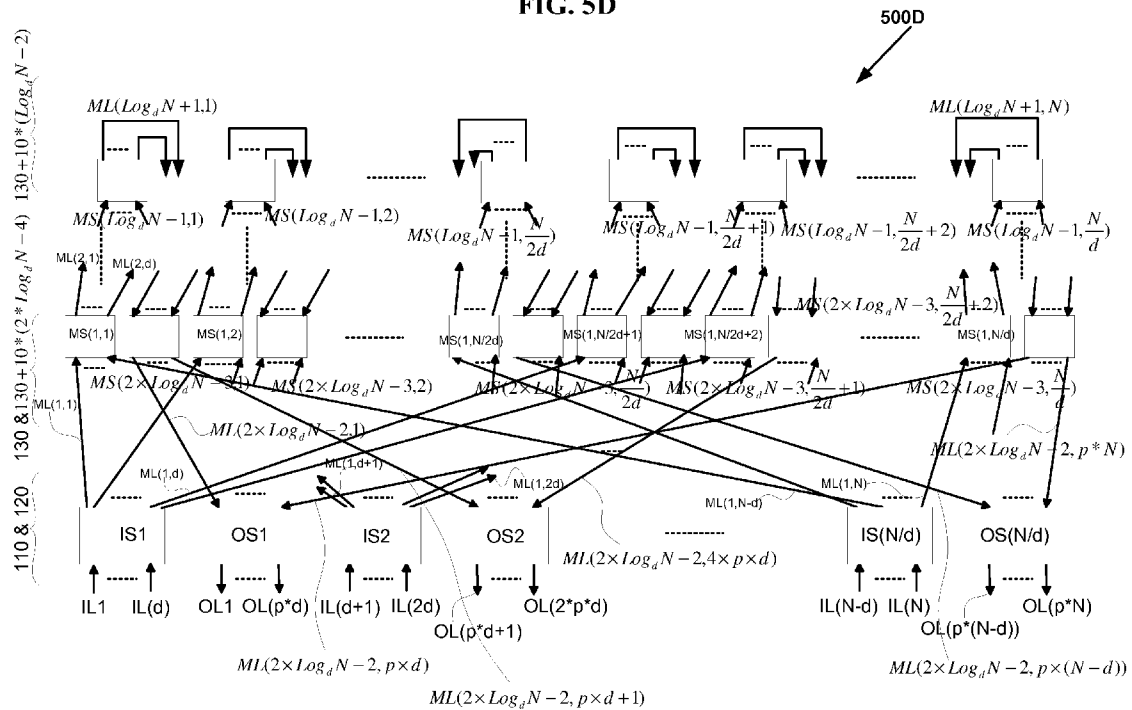


FIG. 5E

500E

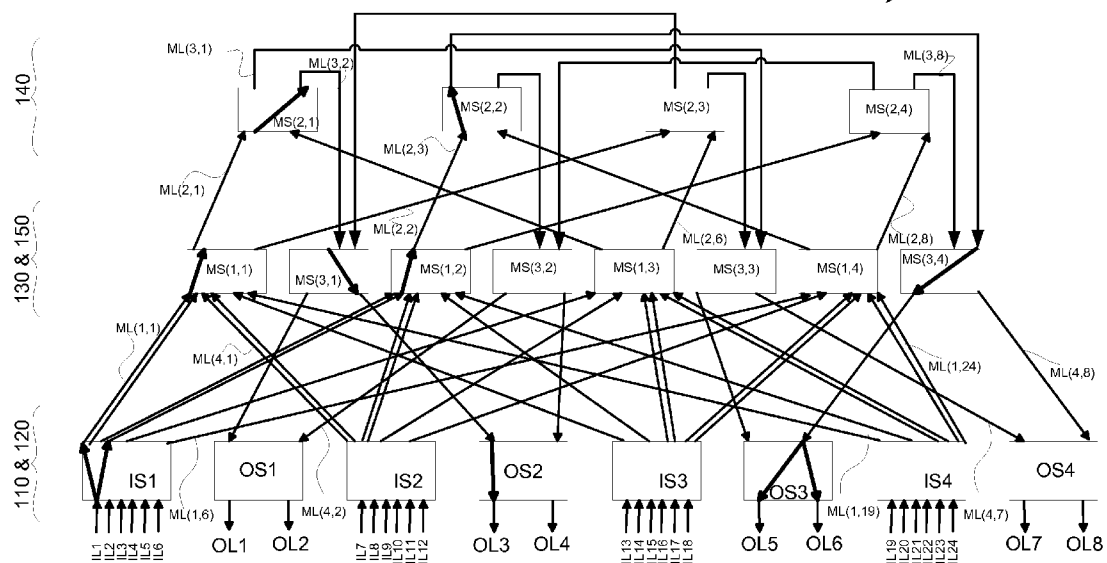


FIG. 5F

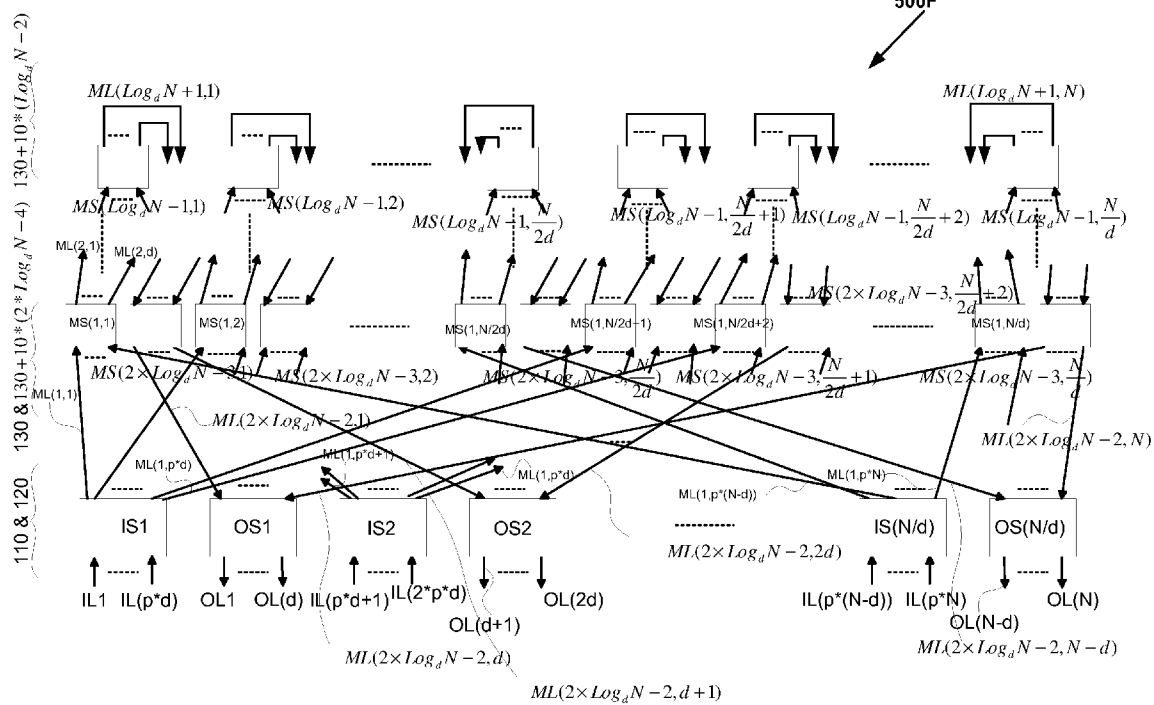


FIG. 6A

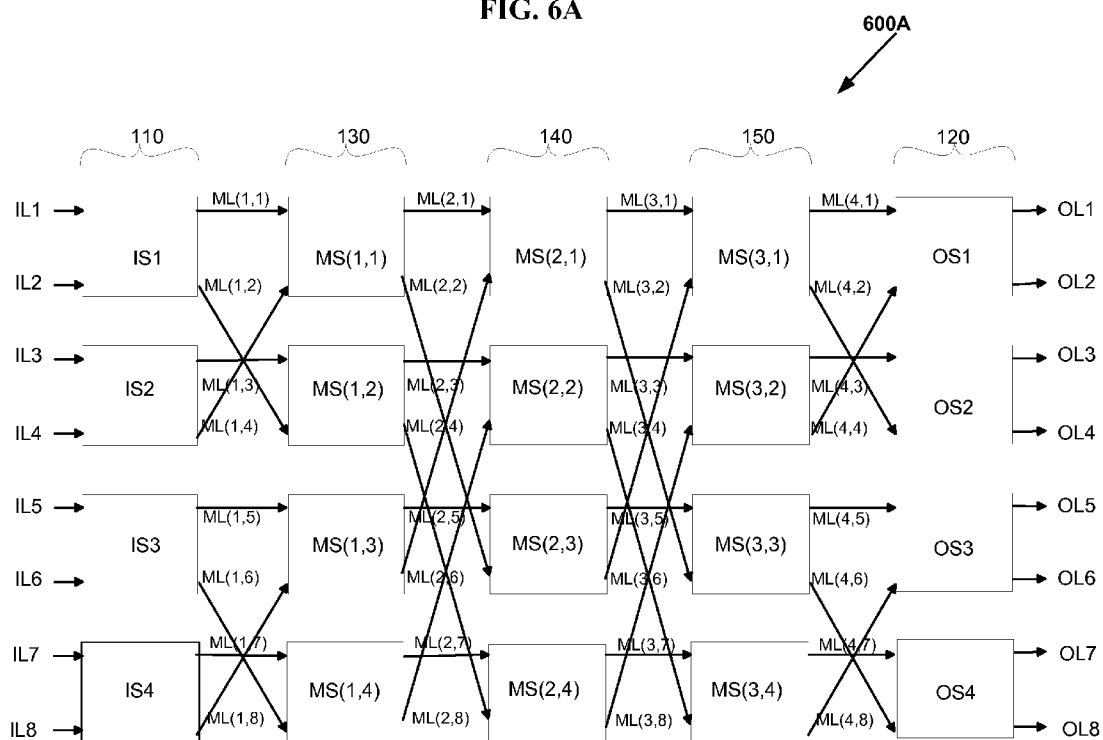
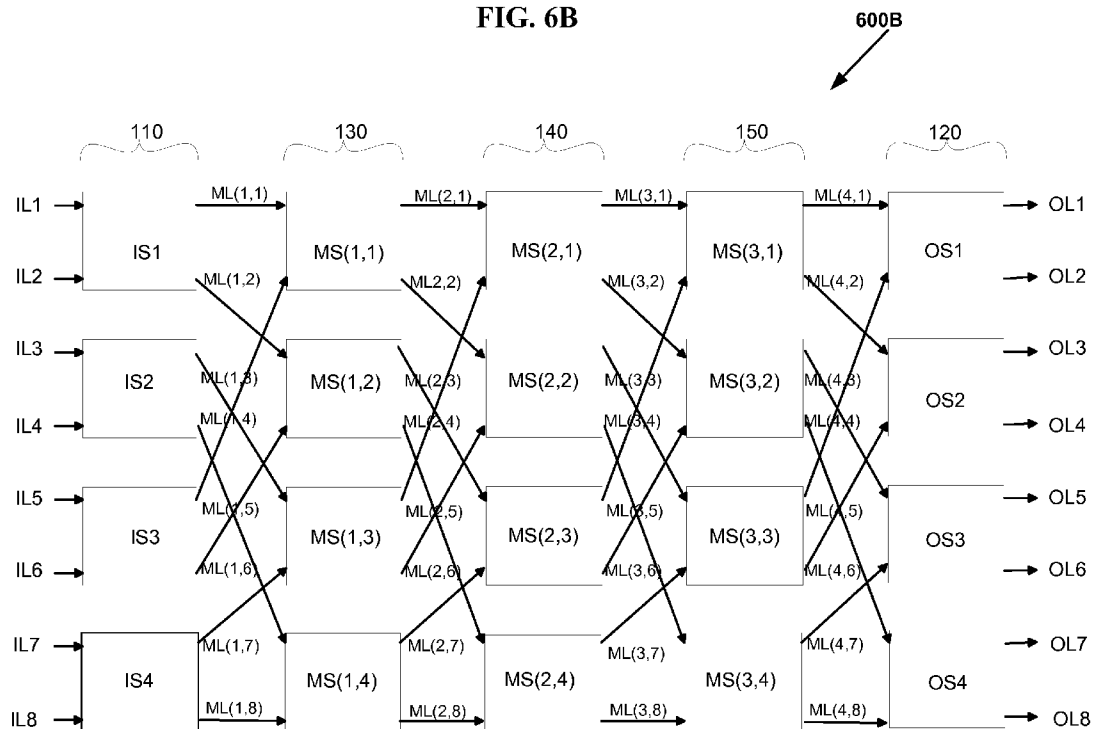


FIG. 6B



600B

FIG. 6C

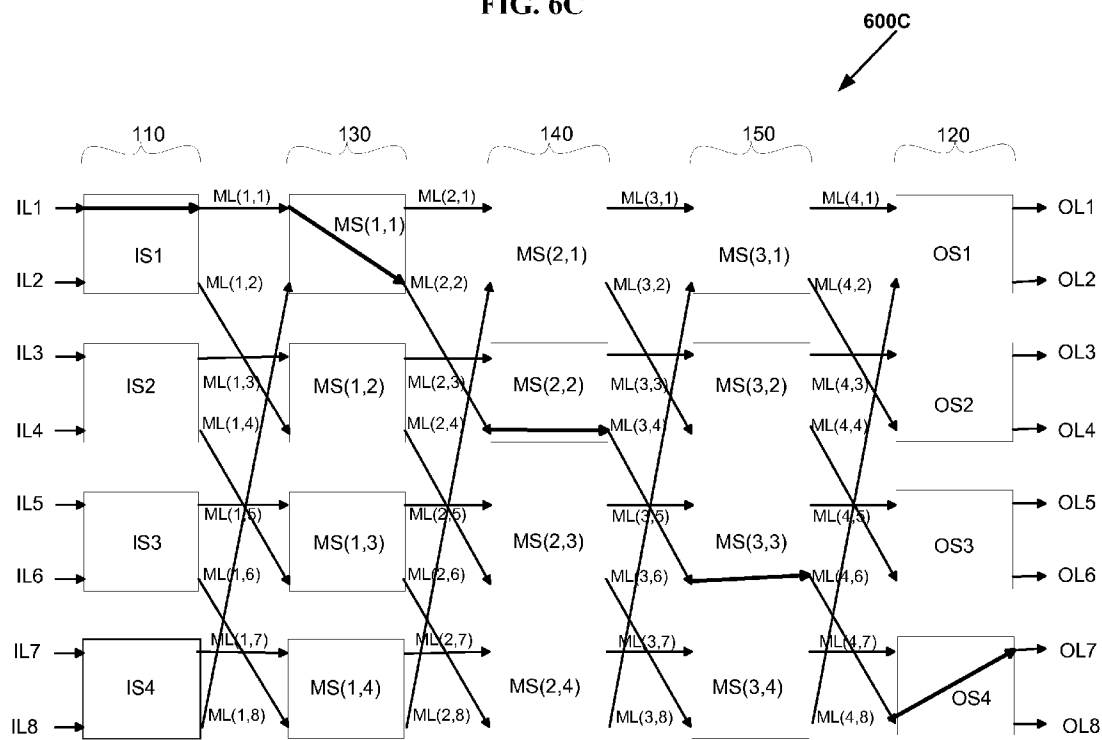


FIG. 6D

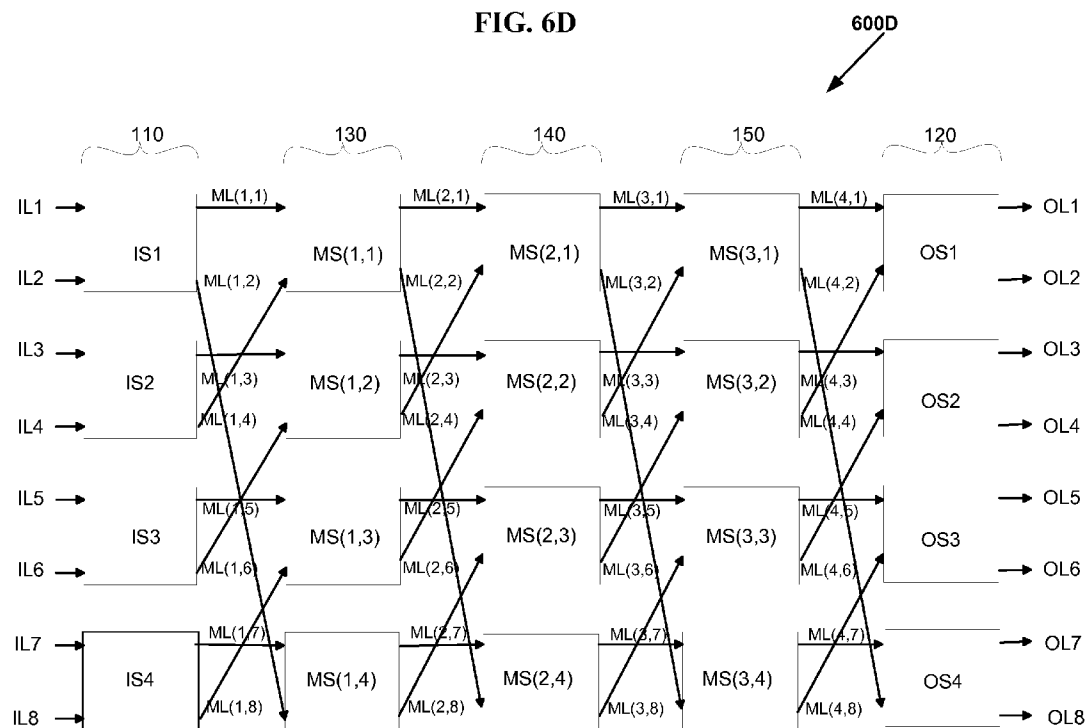


FIG. 6E

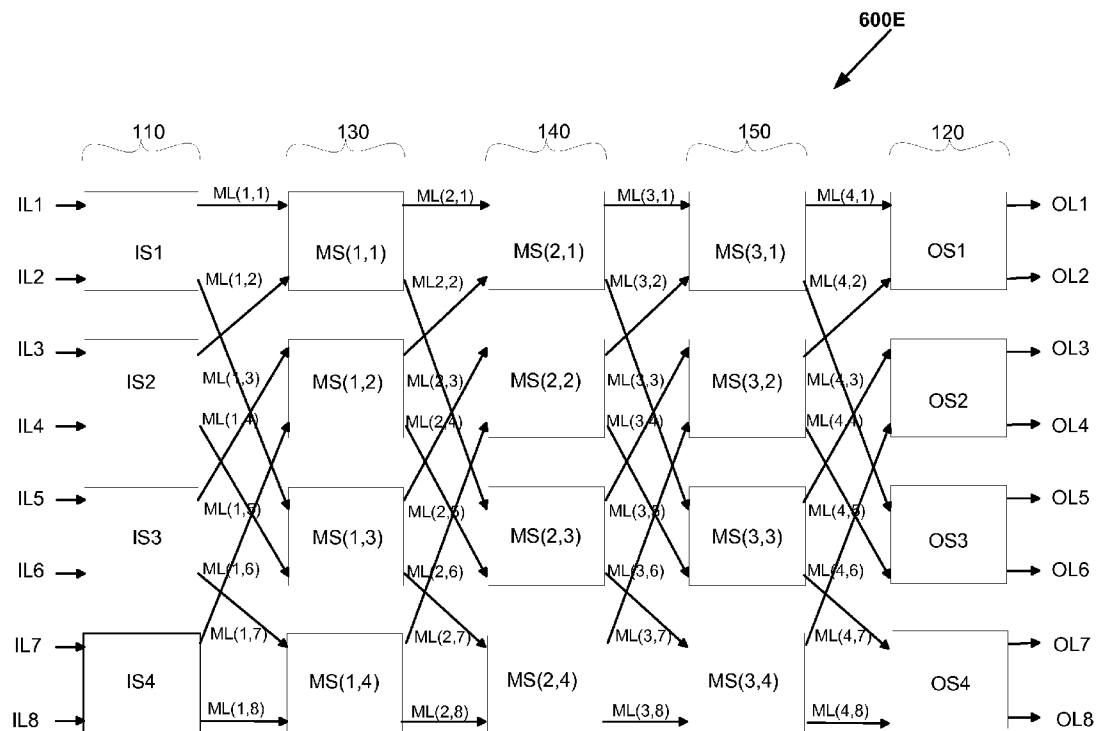


FIG. 6F

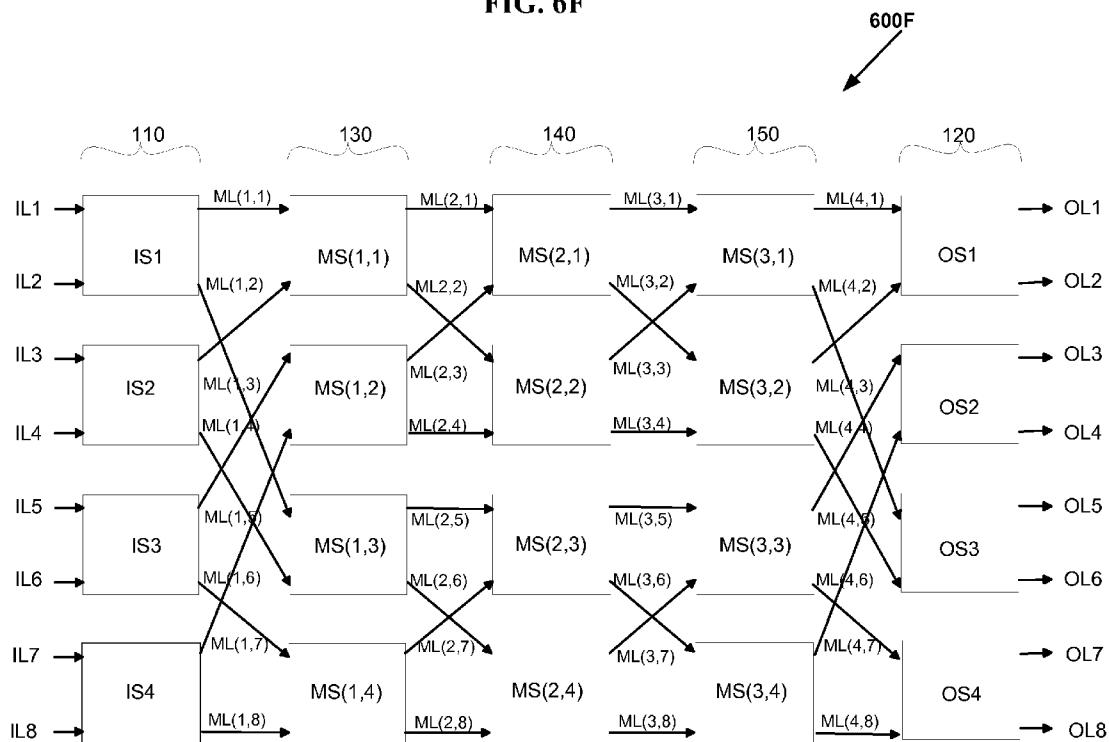


FIG. 6G

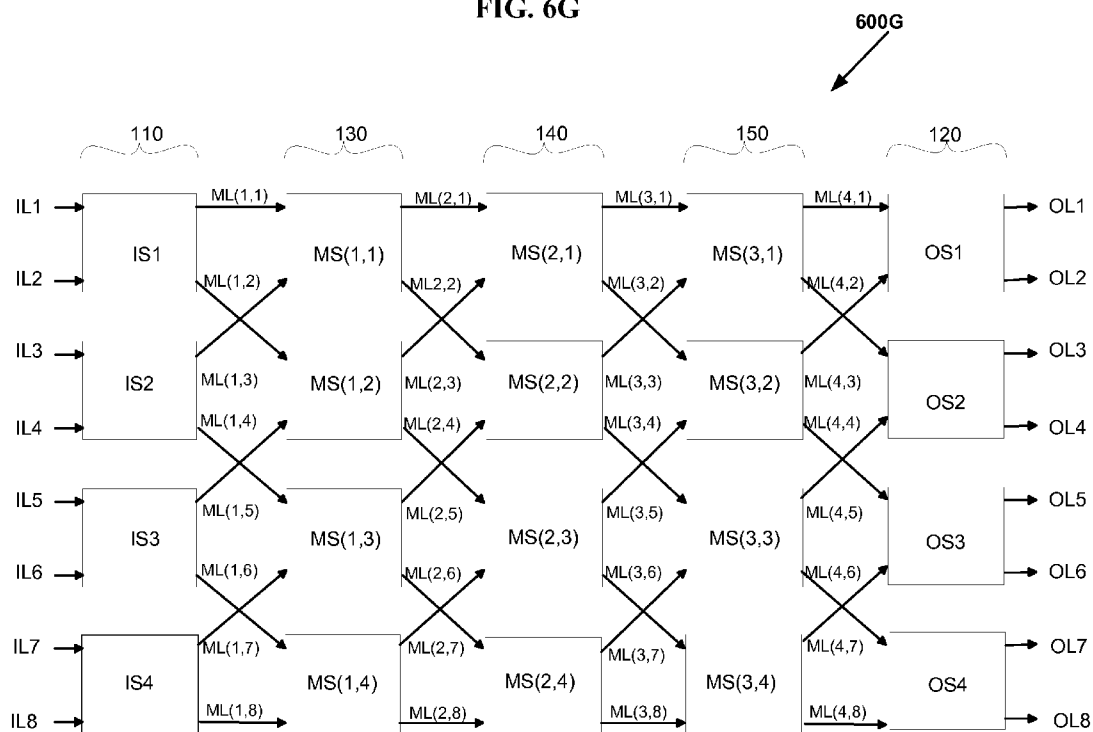


FIG. 6H

600H

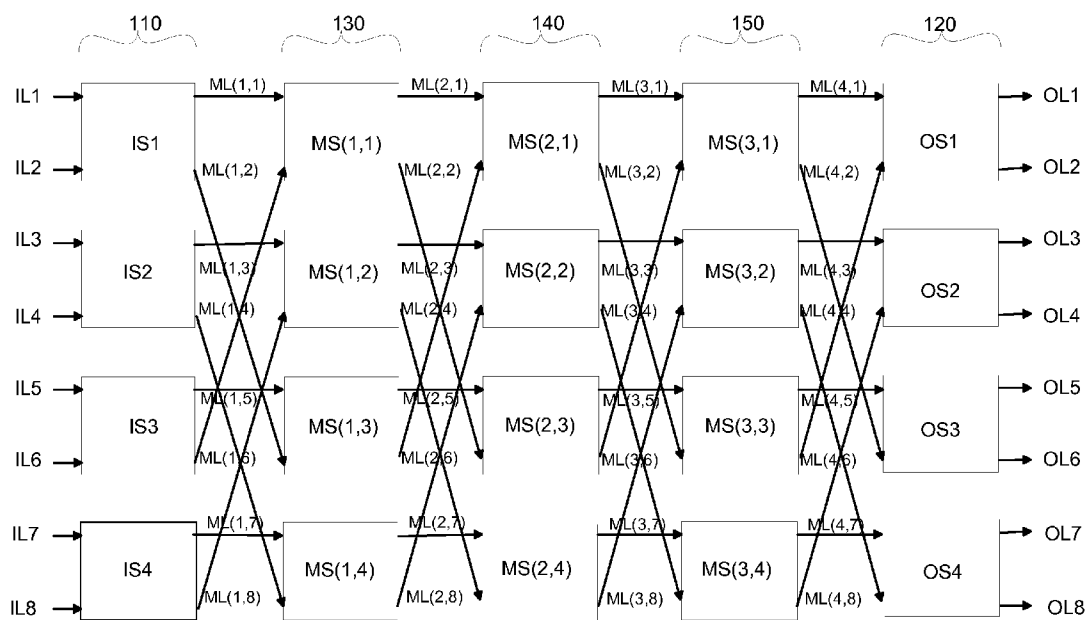


FIG. 6I

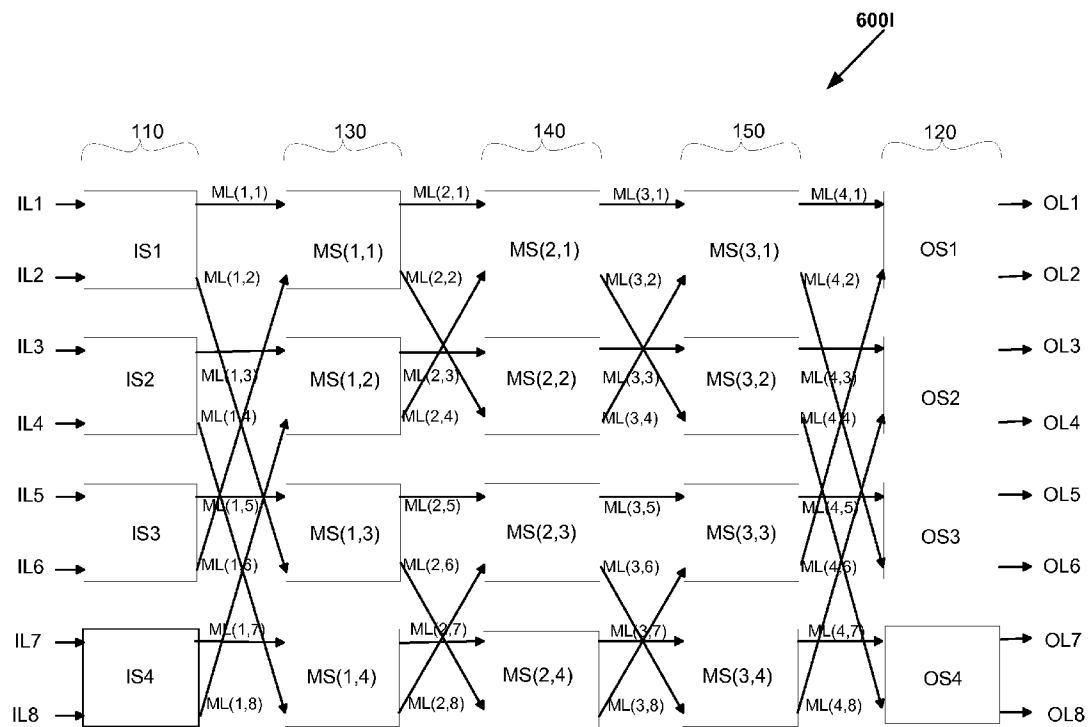


FIG. 6J

600J

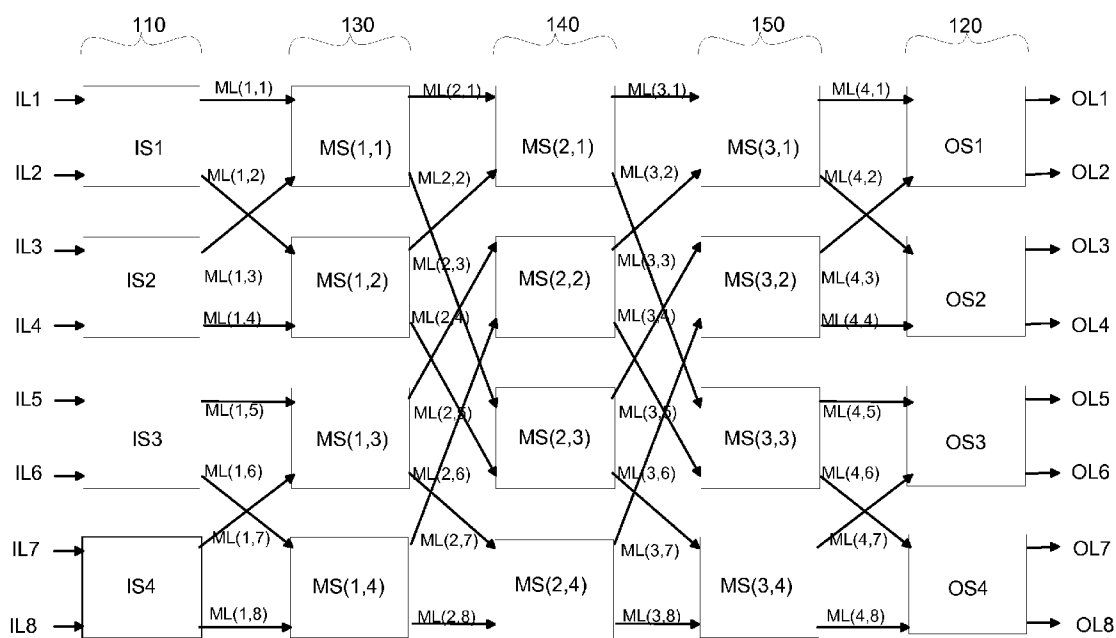


FIG. 6K

600K

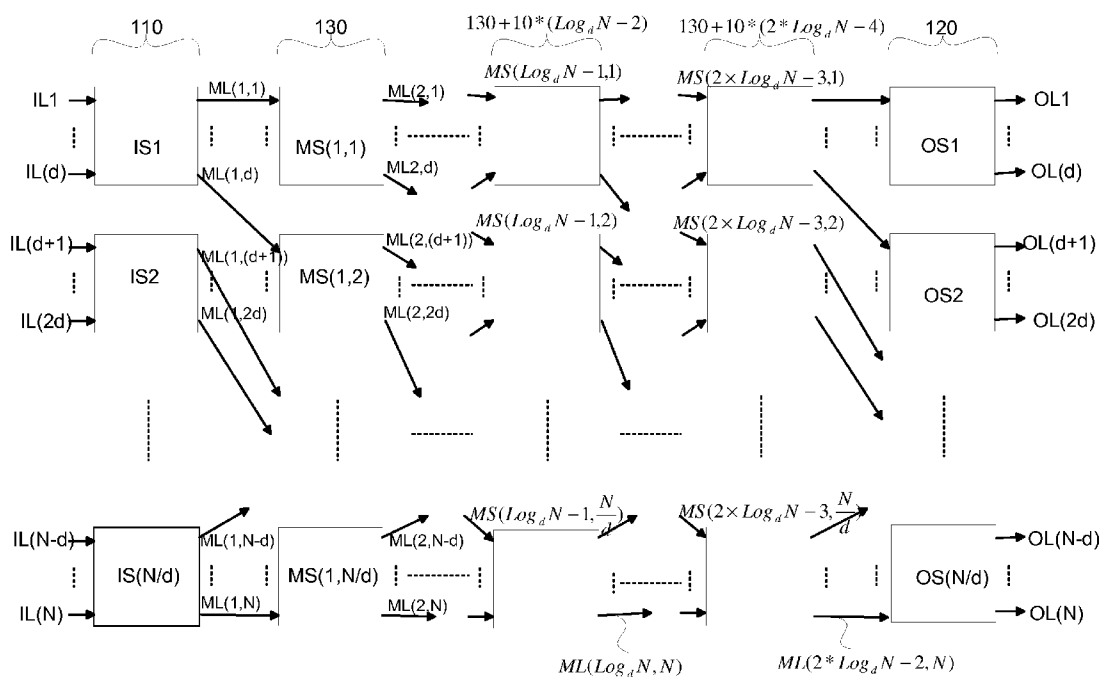


FIG. 6A1

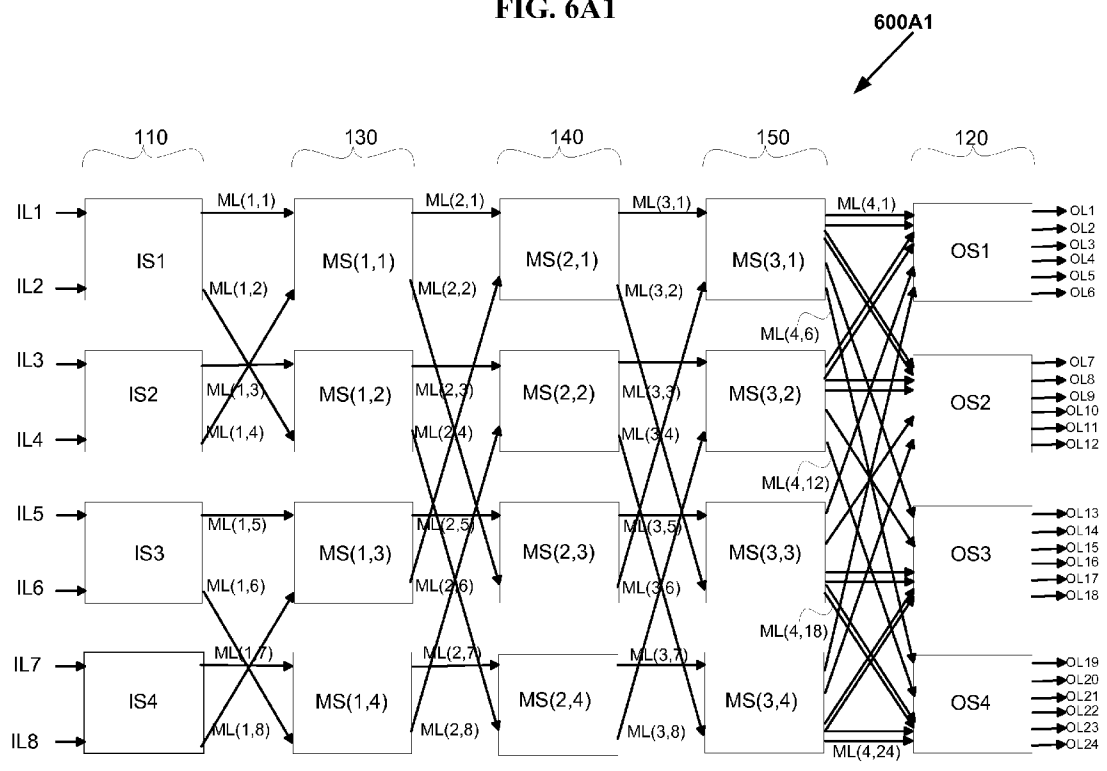


FIG. 6B1

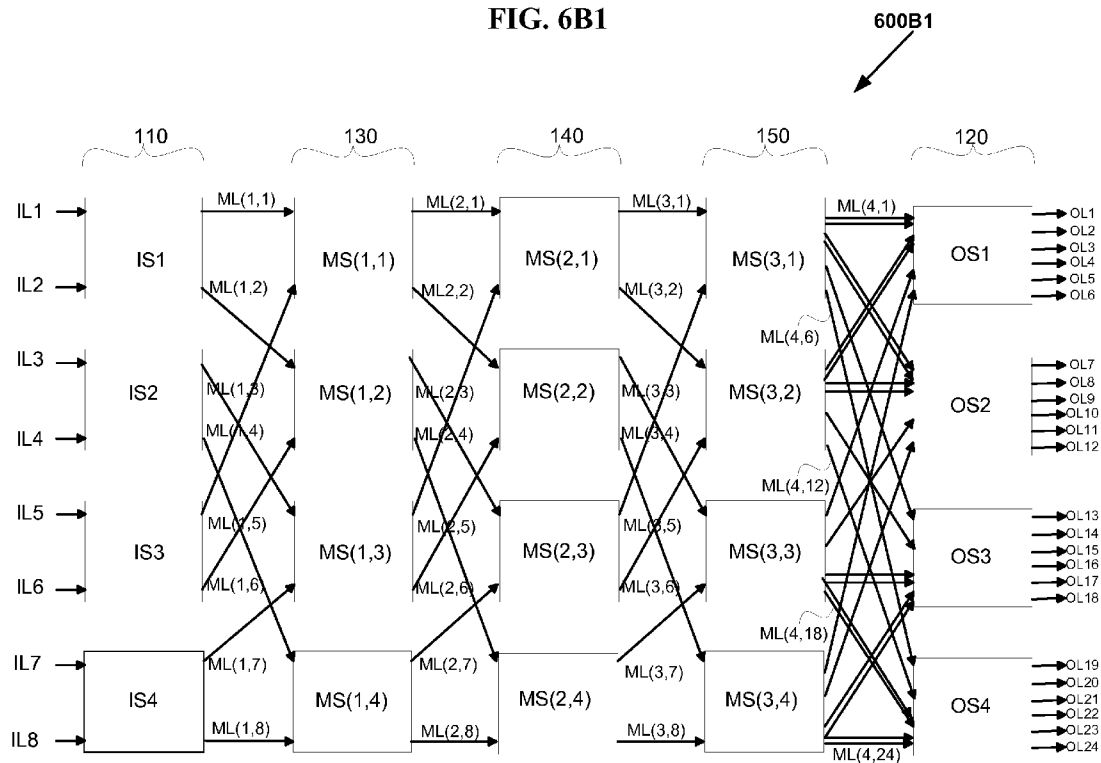


FIG. 6C1

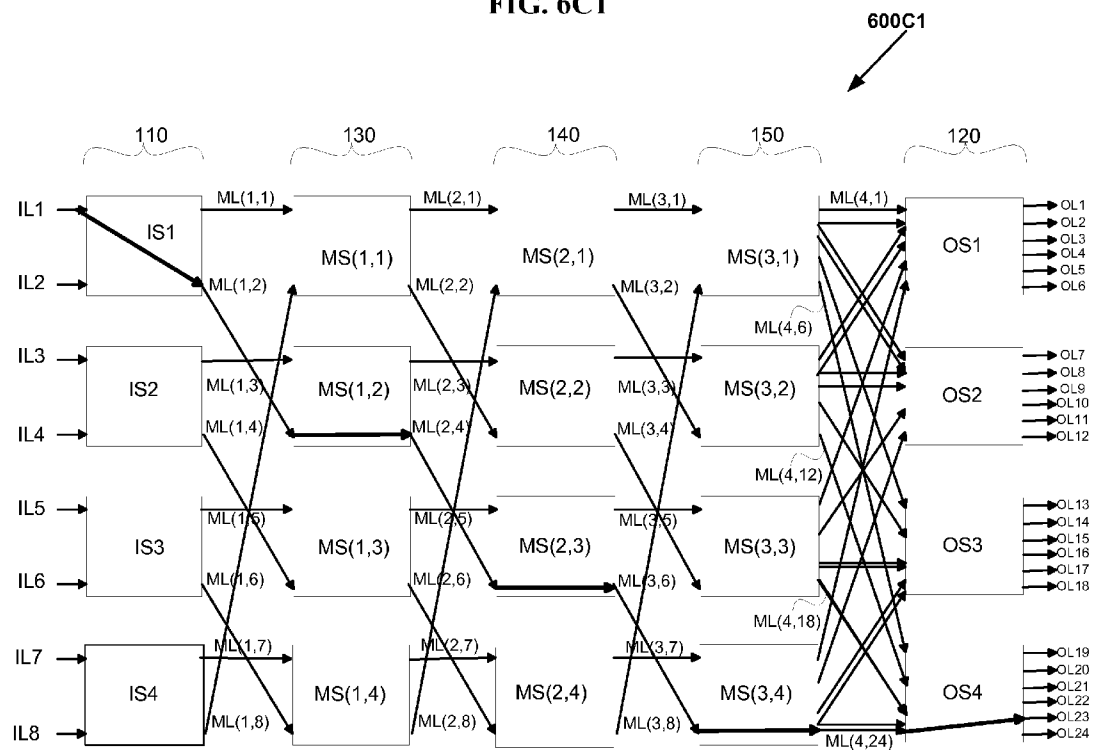


FIG. 6D1

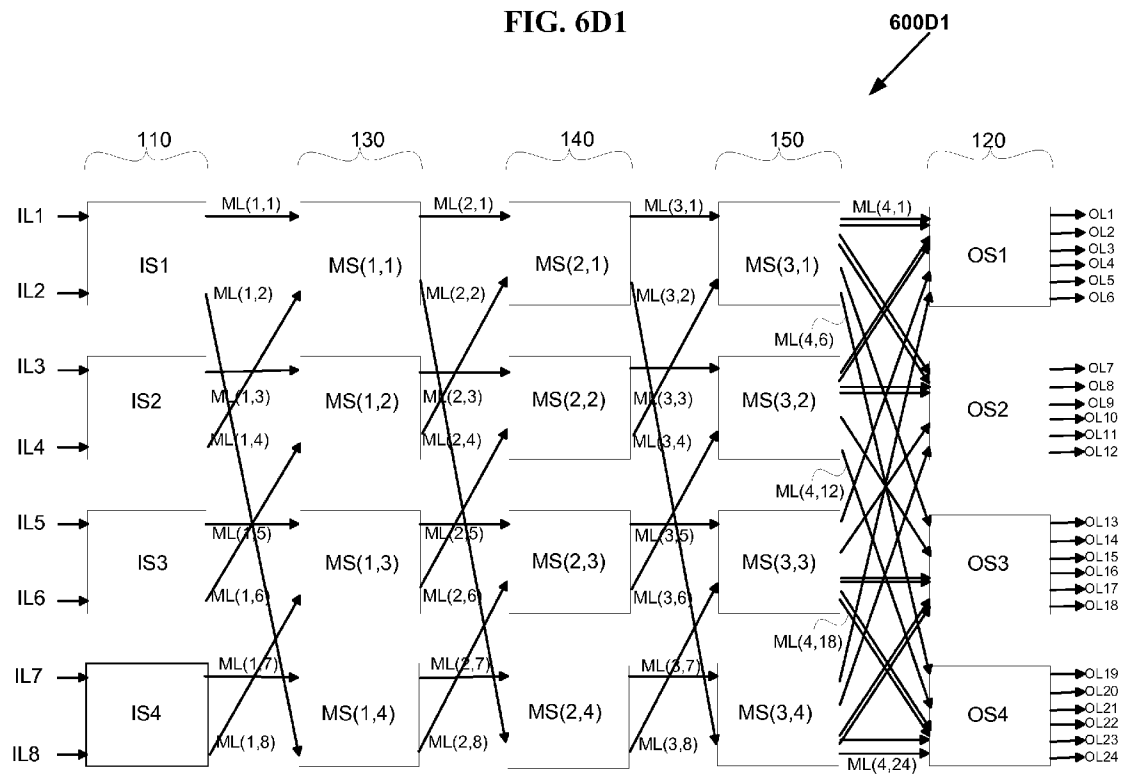


FIG. 6E1

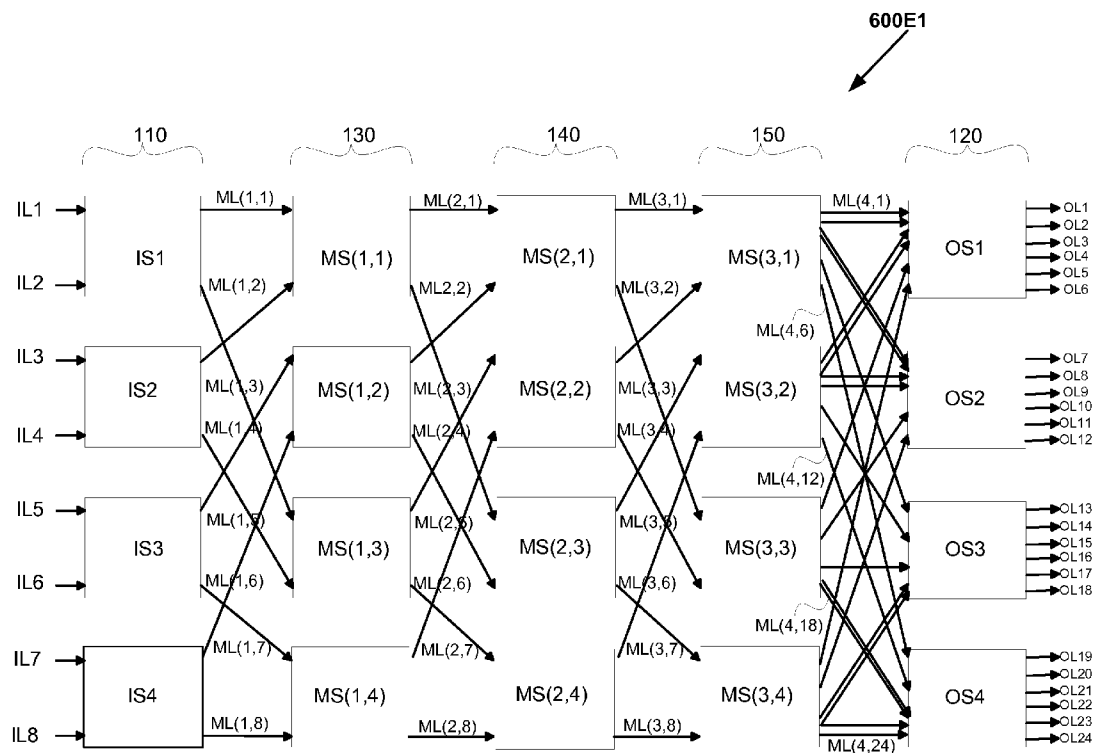


FIG. 6F1

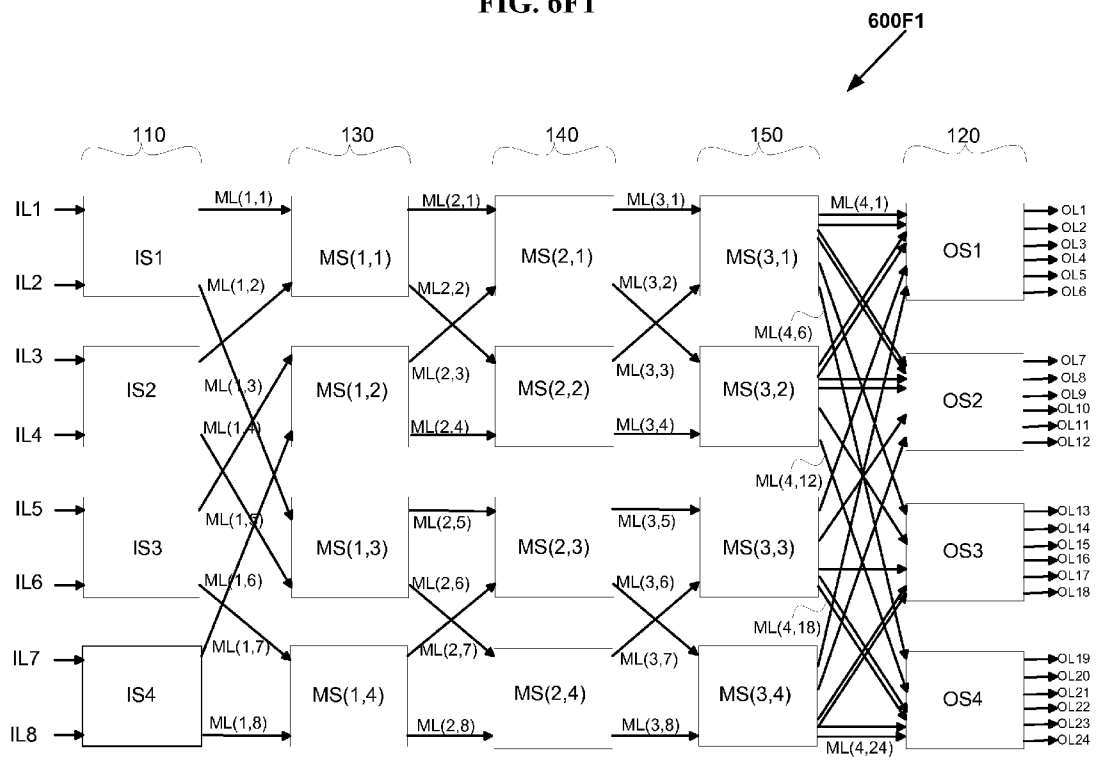


FIG. 6G1

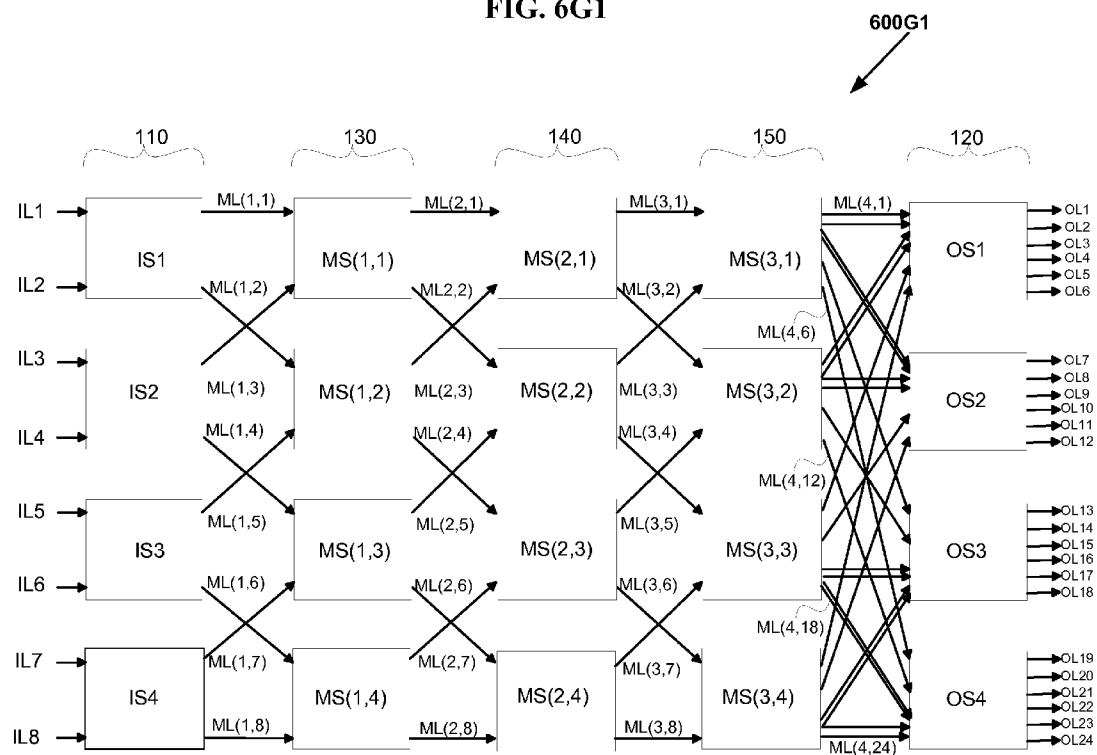


FIG. 6H1

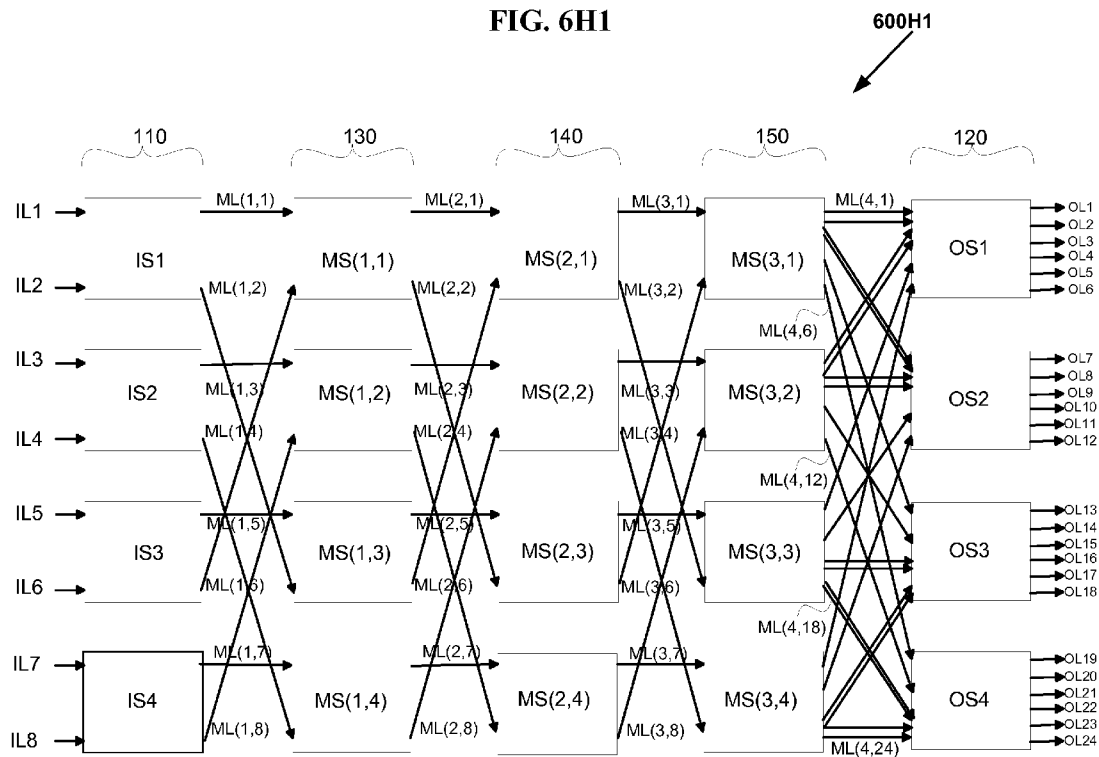


FIG. 6I1

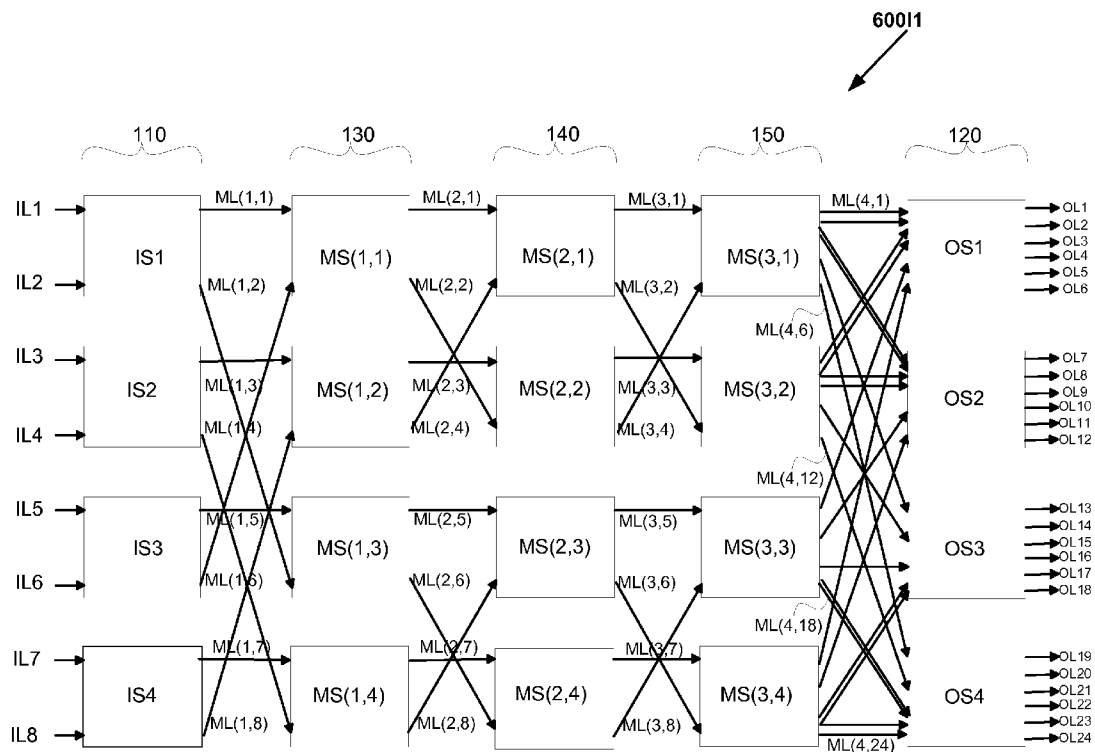


FIG. 6J1

600J1

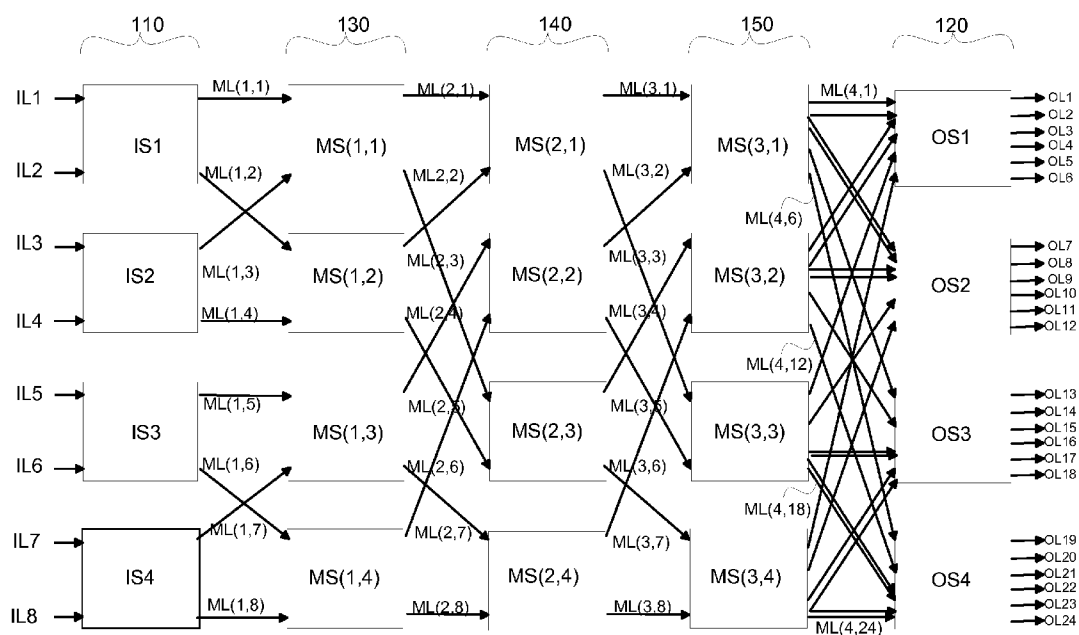


FIG. 6K1

600K1

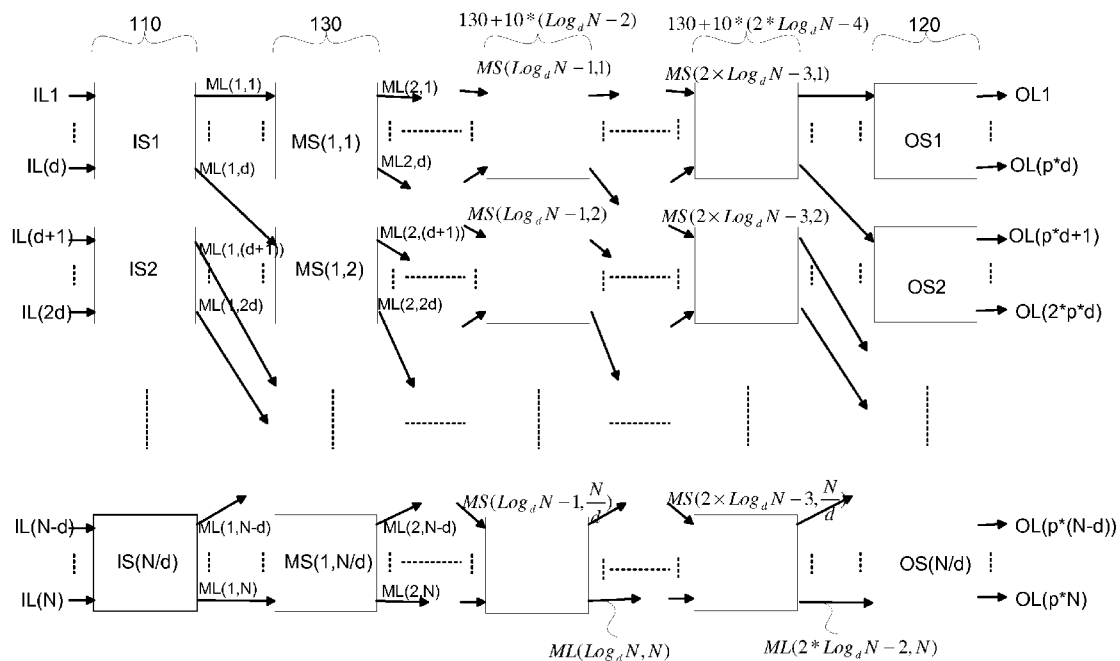


FIG. 6A2

600A2

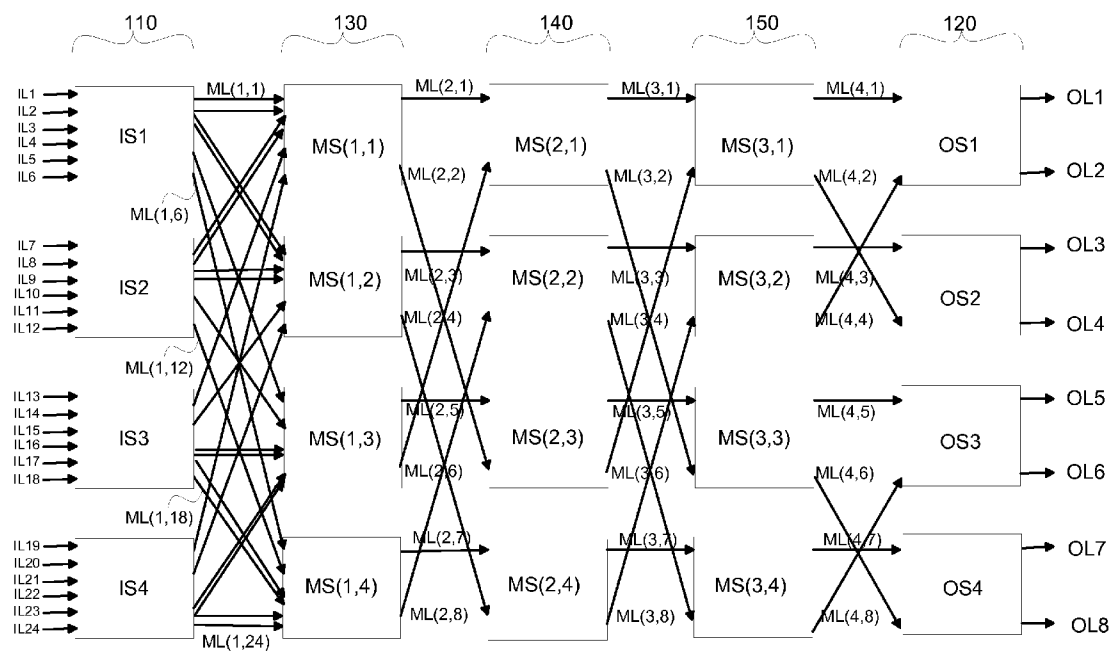


FIG. 6B2

600B2

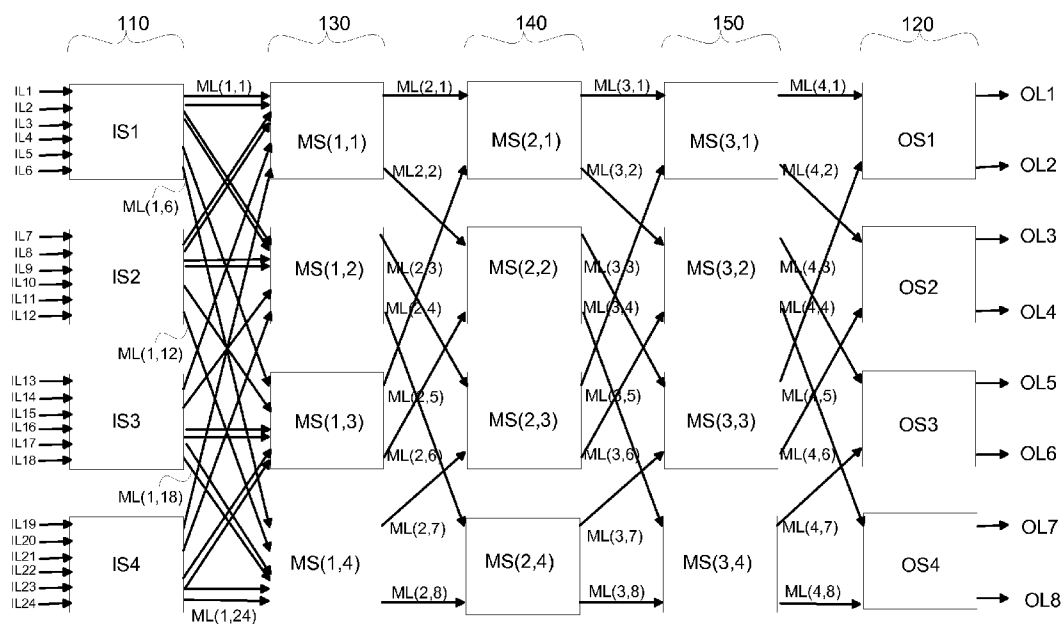


FIG. 6C2

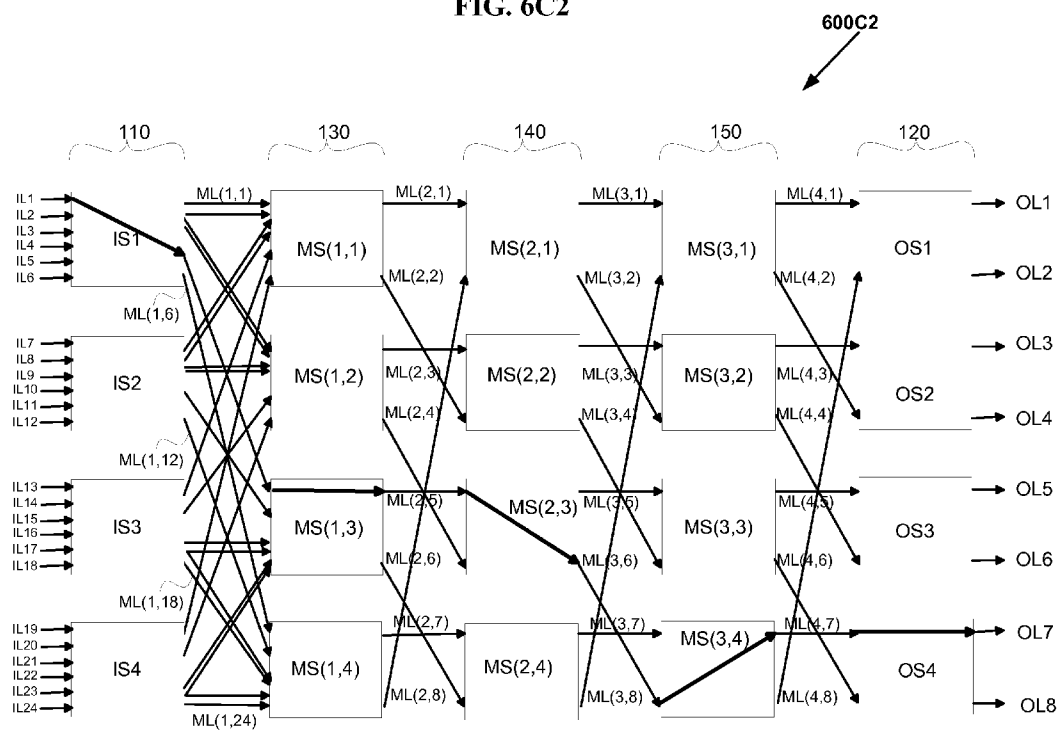


FIG. 6D2

600D2

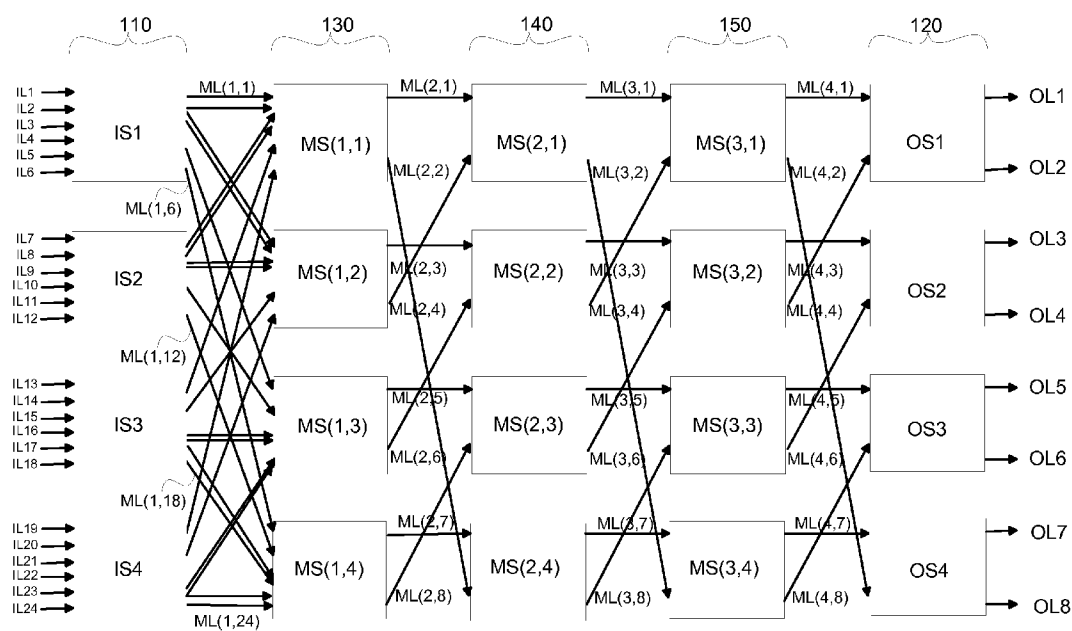


FIG. 6E2

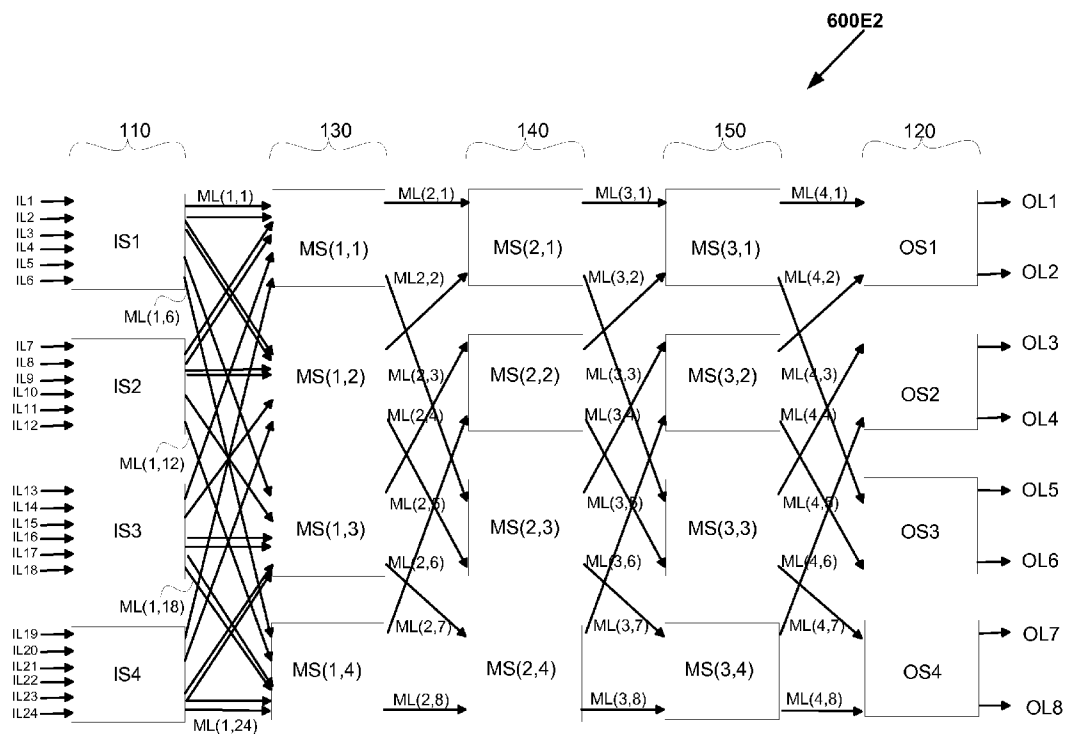


FIG. 6F2

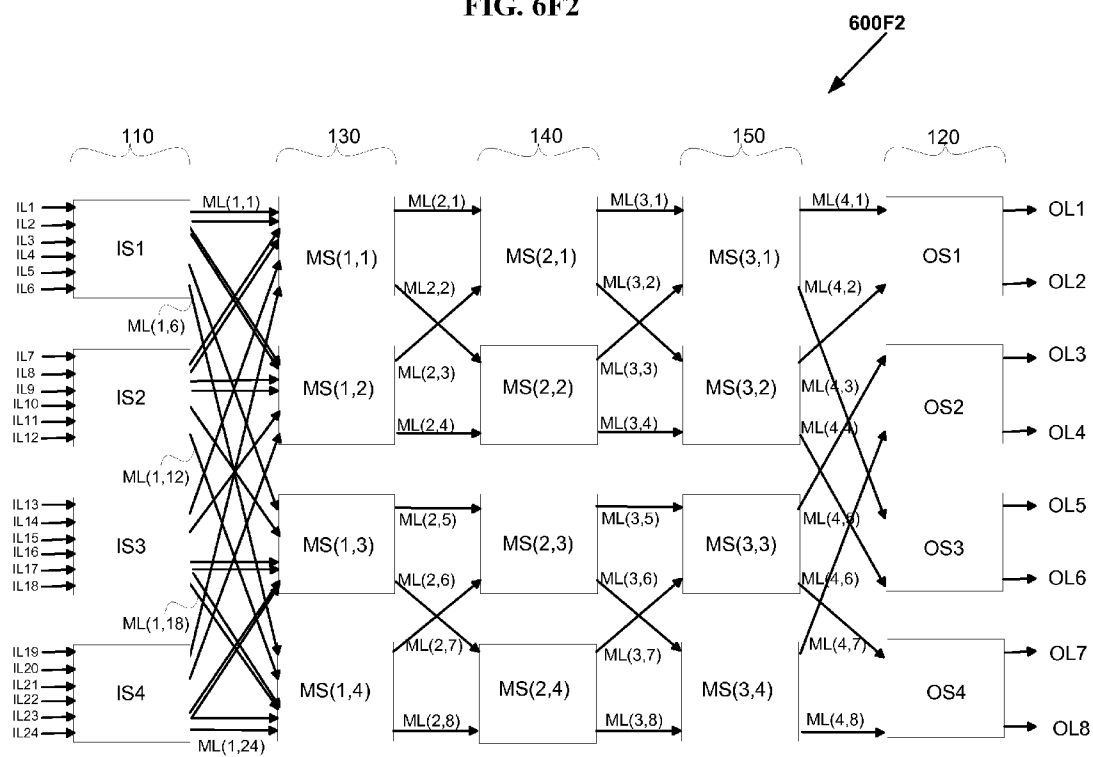


FIG. 6G2

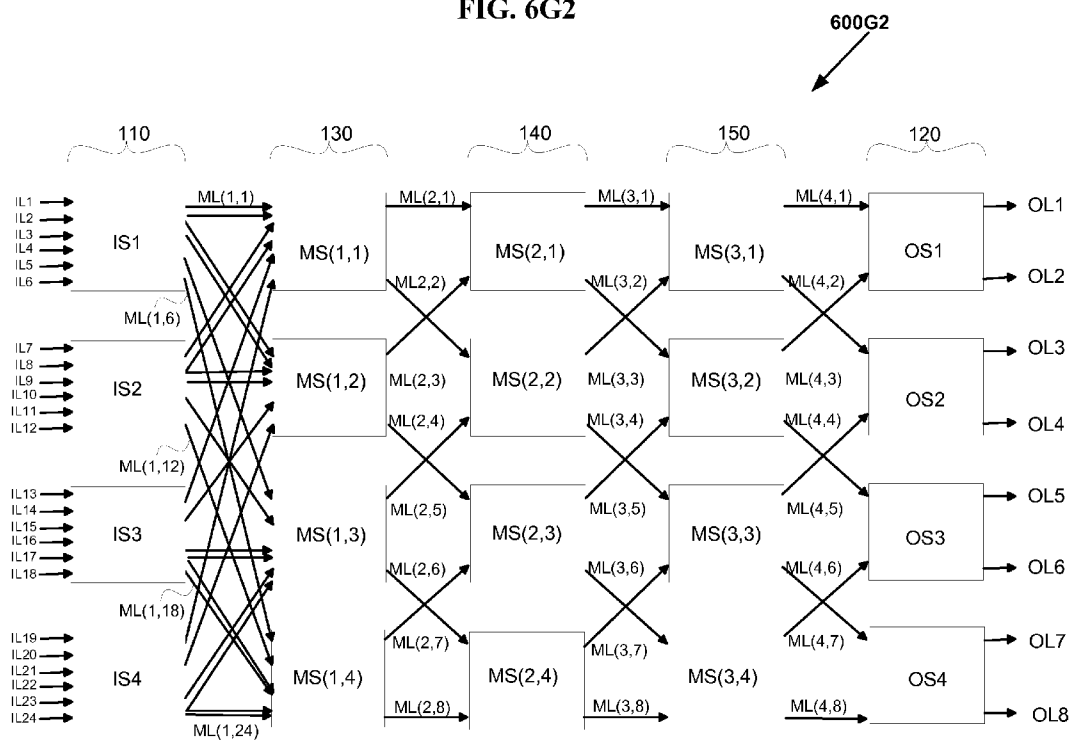


FIG. 6H2

600H2

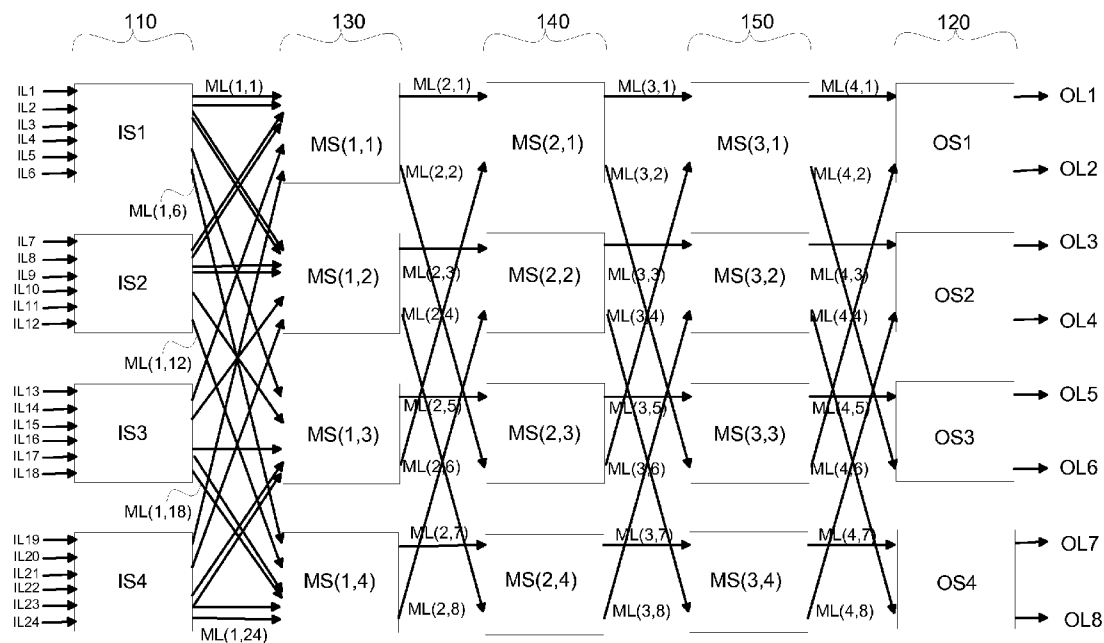


FIG. 6I2

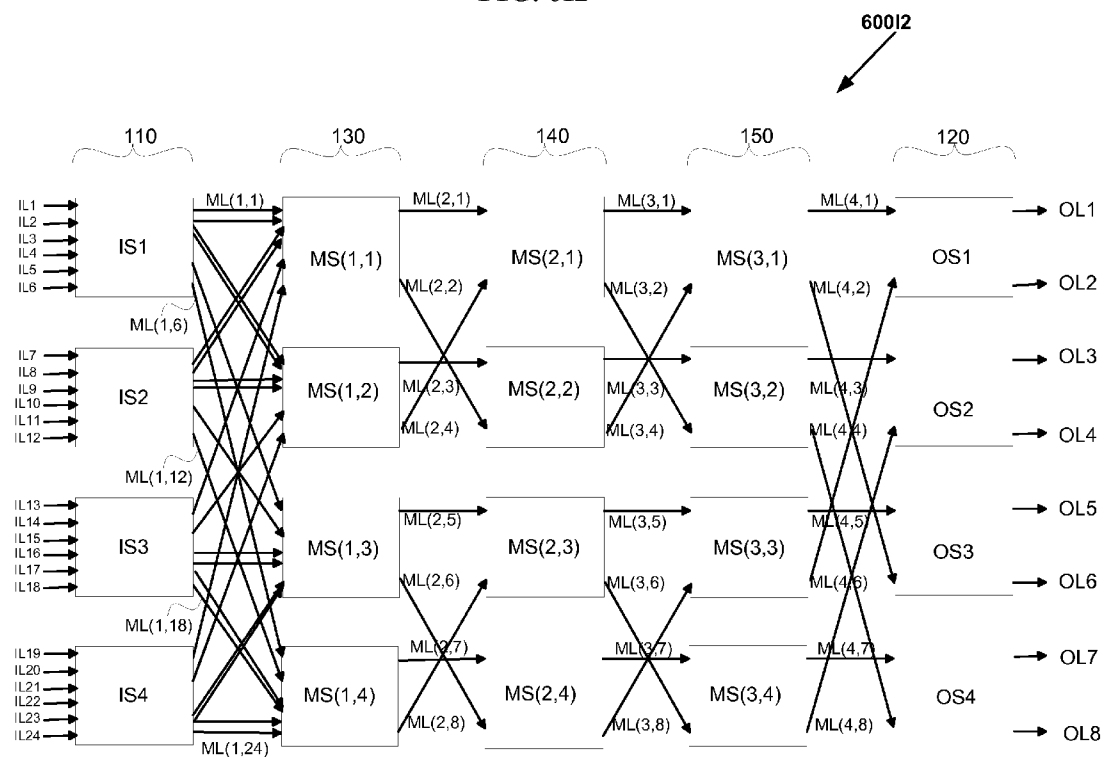


FIG. 6J2

600J2

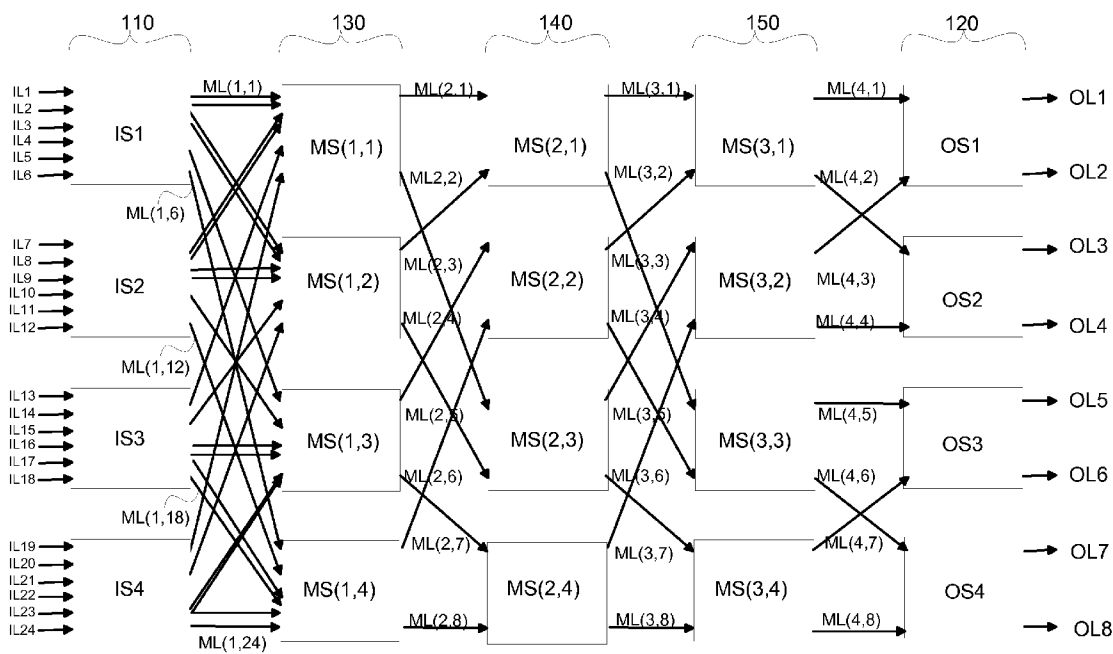
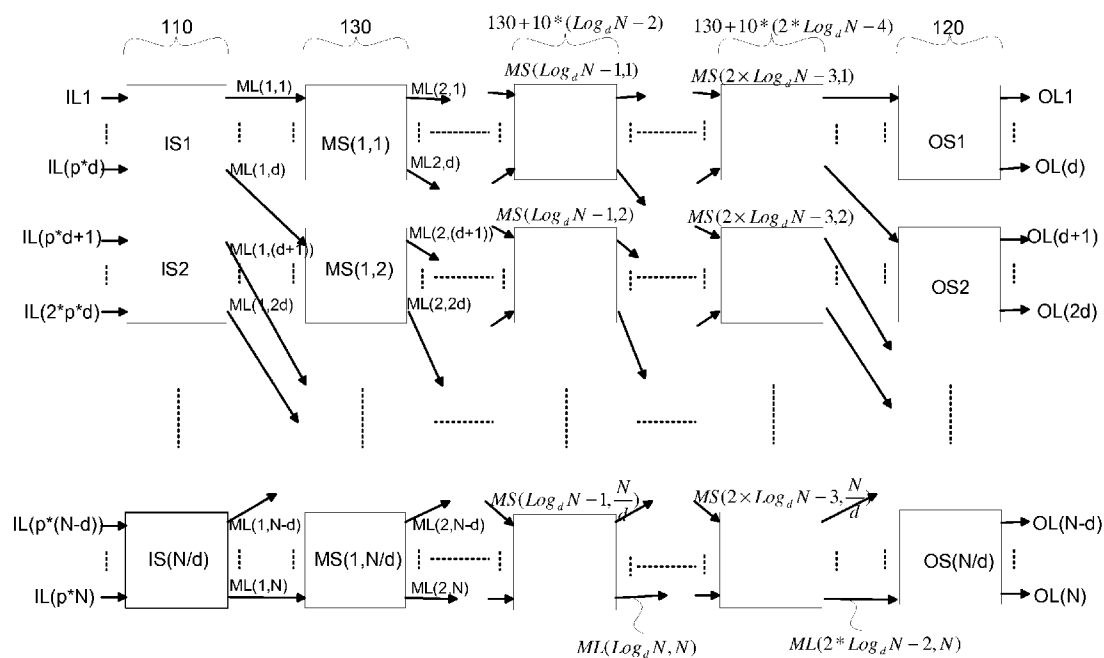


FIG. 6K2

600K2



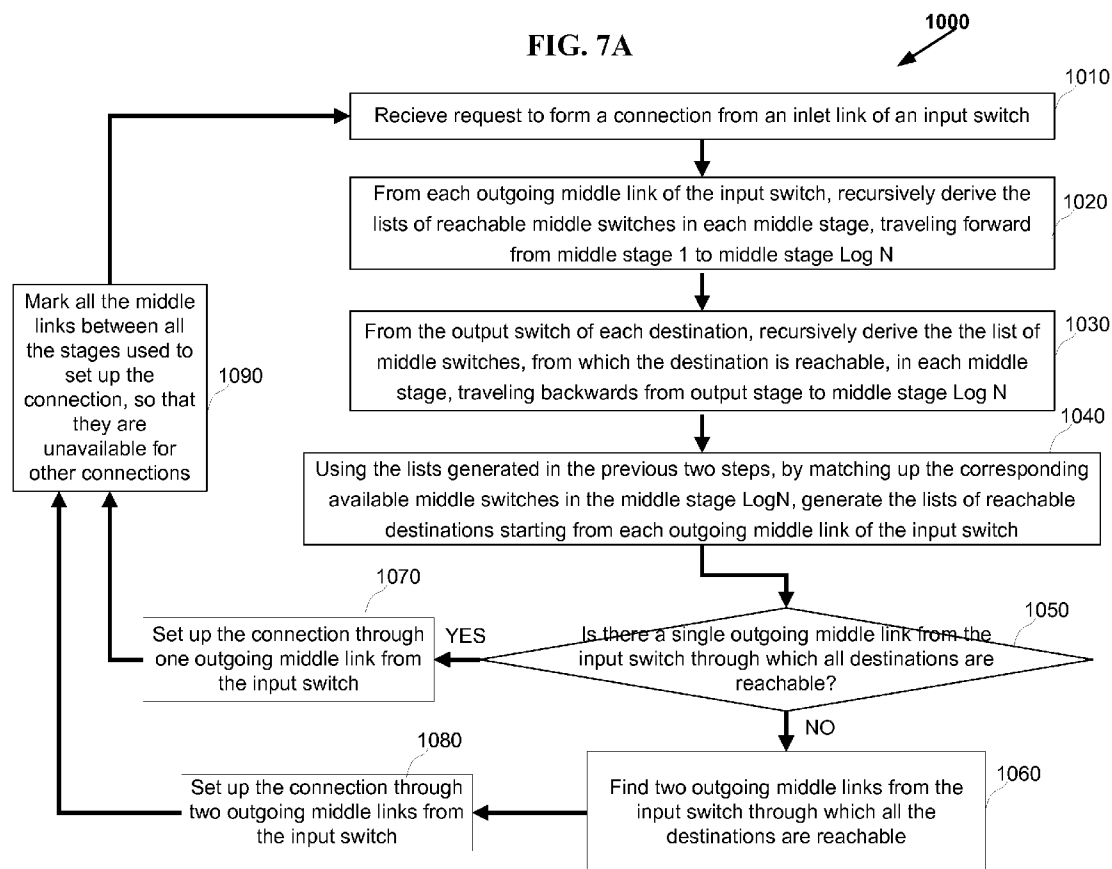
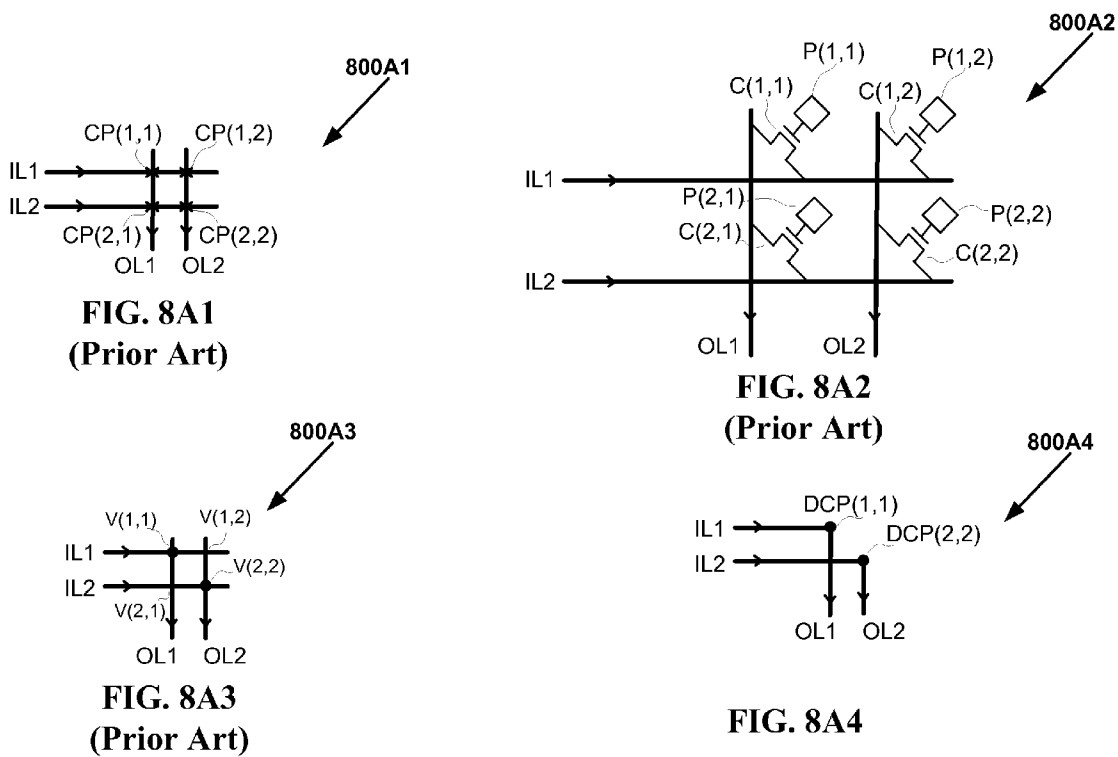


FIG. 8A



US 2011/0044329 A1

Feb. 24, 2011

1

**FULLY CONNECTED GENERALIZED
MULTI-LINK MULTI-STAGE NETWORKS****CROSS REFERENCE TO RELATED
APPLICATIONS**

[0001] This application is related to and claims priority of the PCT Application Serial No. PCT/US08/64604 entitled “FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 22, 2008, the U.S. Provisional Patent Application Ser. No. 60/940,389 entitled “FULLY CONNECTED GENERALIZED REARRANGEABLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, the U.S. Provisional Patent Application Ser. No. 60/940,391 entitled “FULLY CONNECTED GENERALIZED FOLDED MULTI-STAGE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007 and the U.S. Provisional Patent Application Ser. No. 60/940,392 entitled “FULLY CONNECTED GENERALIZED STRICTLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0002] This application is related to and incorporates by reference in its entirety the U.S. application Ser. No. 12/530,207 entitled “FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed Sep. 6, 2009, the PCT Application Serial No. PCT/US08/56064 entitled “FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed Mar. 6, 2008, the U.S. Provisional Patent Application Ser. No. 60/905,526 entitled “LARGE SCALE CROSSPOINT REDUCTION WITH NONBLOCKING UNICAST & MULTICAST IN ARBITRARILY LARGE MULTI-STAGE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed Mar. 6, 2007, and the U.S. Provisional Patent Application Ser. No. 60/940,383 entitled “FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0003] This application is related to and incorporates by reference in its entirety the US Patent Application Docket No. V-0038US entitled “FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS” by Venkat Konda assigned to the same assignee as the current application filed concurrently, the PCT Application Serial No. PCT/US08/64603 entitled “FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 22, 2008, the U.S. Provisional Patent Application Ser. No. 60/940,387 entitled “FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, and the U.S. Provisional Patent Application Ser. No. 60/940,390 entitled “FULLY CONNECTED GENERALIZED MULTI-LINK BUTTERFLY FAT TREE NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0004] This application is related to and incorporates by reference in its entirety the US Patent Application Docket No. V-0045US entitled “VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS” by Venkat Konda assigned to the same assignee as the current application filed concurrently, the PCT Application Serial No. PCT/US08/64605 entitled “VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 22, 2008, and the U.S. Provisional Patent Application Ser. No. 60/940,394 entitled “VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0005] This application is related to and incorporates by reference in its entirety the U.S. Provisional Patent Application Ser. No. 61/252,603 entitled “VLSI LAYOUTS OF FULLY CONNECTED NETWORKS WITH LOCALITY EXPLOITATION” by Venkat Konda assigned to the same assignee as the current application, filed Oct. 16, 2009.

[0006] This application is related to and incorporates by reference in its entirety the U.S. Provisional Patent Application Ser. No. 61/252,609 entitled “VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS” by Venkat Konda assigned to the same assignee as the current application, filed Oct. 16, 2009.

BACKGROUND OF INVENTION

[0007] Clos switching network, Benes switching network, and Cantor switching network are a network of switches configured as a multi-stage network so that fewer switching points are necessary to implement connections between its inlet links (also called “inputs”) and outlet links (also called “outputs”) than would be required by a single stage (e.g. crossbar) switch having the same number of inputs and outputs. Clos and Benes networks are very popularly used in digital crossconnects, switch fabrics and parallel computer systems. However Clos and Benes networks may block some of the connection requests.

[0008] There are generally three types of nonblocking networks: strictly nonblocking; wide sense nonblocking; and rearrangeably nonblocking (See V. E. Benes, “Mathematical Theory of Connecting Networks and Telephone Traffic” Academic Press, 1965 that is incorporated by reference, as background). In a rearrangeably nonblocking network, a connection path is guaranteed as a result of the networks ability to rearrange prior connections as new incoming calls are received. In strictly nonblocking network, for any connection request from an inlet link to some set of outlet links, it is always possible to provide a connection path through the network to satisfy the request without disturbing other existing connections, and if more than one such path is available, any path can be selected without being concerned about realization of future potential connection requests. In wide-sense nonblocking networks, it is also always possible to provide a connection path through the network to satisfy the request without disturbing other existing connections, but in this case the path used to satisfy the connection request must be carefully selected so as to maintain the nonblocking connecting capability for future potential connection requests.

[0009] Butterfly Networks, Banyan Networks, Batcher-Banyan Networks, Baseline Networks, Delta Networks, Omega Networks and Flip networks have been widely studied particularly for self routing packet switching applications.

US 2011/0044329 A1

Feb. 24, 2011

2

Also Benes Networks with radix of two have been widely studied and it is known that Benes Networks of radix two are shown to be built with back to back baseline networks which are rearrangeably nonblocking for unicast connections.

[0010] U.S. Pat. No. 5,451,936 entitled "Non-blocking Broadcast Network" granted to Yang et al. is incorporated by reference herein as background of the invention. This patent describes a number of well known nonblocking multi-stage switching network designs in the background section at column 1, line 22 to column 3, 59. An article by Y. Yang, and G. M., Masson entitled, "Non-blocking Broadcast Switching Networks" IEEE Transactions on Computers, Vol. 40, No. 9, September 1991 that is incorporated by reference as background indicates that if the number of switches in the middle stage, m , of a three-stage network satisfies the relation $m \geq \min((n-1)(x+r^{1/x}))$ where $1 \leq x \leq \min(n-1, r)$, the resulting network is nonblocking for multicast assignments. In the relation, r is the number of switches in the input stage, and n is the number of inlet links in each input switch.

[0011] U.S. Pat. No. 6,885,669 entitled "Rearrangeably Nonblocking Multicast Multi-stage Networks" by Konda showed that three-stage Clos network is rearrangeably nonblocking for arbitrary fan-out multicast connections when $m \geq 2 \times n$. And U.S. Pat. No. 6,868,084 entitled "Strictly Nonblocking Multicast Multi-stage Networks" by Konda showed that three-stage Clos network is strictly nonblocking for arbitrary fan-out multicast connections when $m \geq 3 \times n - 1$.

[0012] In general multi-stage networks for stages of more than three and radix of more than two are not well studied. An article by Charles Clos entitled "A Study of Non-Blocking Switching Networks" The Bell Systems Technical Journal, Volume XXXII, January 1953, No. 1, pp. 406-424 showed a way of constructing large multi-stage networks by recursive substitution with a crosspoint complexity of $d^2 \times N \times (\log_d N)^2$ for strictly nonblocking unicast network. Similarly U.S. Pat. No. 6,885,669 entitled "Rearrangeably Nonblocking Multicast Multi-stage Networks" by Konda showed a way of constructing large multi-stage networks by recursive substitution for rearrangeably nonblocking multicast network. An article by D. G. Cantor entitled "On Non-Blocking Switching Networks" 1: pp. 367-377, 1972 by John Wiley and Sons, Inc., showed a way of constructing large multi-stage networks with a crosspoint complexity of $d^2 \times N \times (\log_d N)^2$ for strictly nonblocking unicast, (by using $\log_d N$ number of Benes Networks for $d=2$) and without counting the crosspoints in multiplexers and demultiplexers. Jonathan Turner studied the cascaded Benes Networks with radices larger than two, for nonblocking multicast with 10 times the crosspoint complexity of that of nonblocking unicast for a network of size $N=256$.

[0013] The crosspoint complexity of all these networks is prohibitively large to implement the interconnect for multicast connections particularly in field programmable gate array (FPGA) devices, programmable logic devices (PLDs), field programmable interconnect Chips (FPICs), digital crossconnects, switch fabrics and parallel computer systems.

SUMMARY OF INVENTION

[0014] A generalized multi-link multi-stage network comprising $(2 \times \log_d N) - 1$ stages is operated in strictly nonblocking manner for unicast includes an input stage having N/d switches with each of them having d inlet links and $2 \times d$ outgoing links connecting to second stage switches, an output stage having N/d switches with each of them having d outlet

links and $2 \times d$ incoming links connecting from switches in the penultimate stage. The network also has $(2 \times \log_d N) - 3$ middle stages with each middle stage having N/d switches, and each switch in the middle stage has $2 \times d$ incoming links connecting from the switches in its immediate preceding stage, and $2 \times d$ outgoing links connecting to the switches in its immediate succeeding stage. Also the same generalized multi-link multi-stage network is operated in rearrangeably nonblocking manner for arbitrary fan-out multicast and each multicast connection is set up by use of at most two outgoing links from the input stage switch.

[0015] A generalized multi-link multi-stage network comprising $(2 \times \log_d N) - 1$ stages is operated in strictly nonblocking manner for multicast includes an input stage having N/d switches with each of them having d inlet links and $3 \times d$ outgoing links connecting to second stage switches, an output stage having N/d switches with each of them having d outlet links and $3 \times d$ incoming links connecting from switches in the penultimate stage. The network also has $(2 \times \log_d N) - 3$ middle stages with each middle stage having N/d switches, and each switch in the middle stage has $3 \times d$ incoming links connecting from the switches in its immediate preceding stage, and $3 \times d$ outgoing links connecting to the switches in its immediate succeeding stage.

BRIEF DESCRIPTION OF DRAWINGS

[0016] FIG. 1A is a diagram 100A of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having inverse Benes connection topology of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0017] FIG. 1B is a diagram 100B of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ (having a connection topology built using back-to-back Omega Networks) of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0018] FIG. 1C is a diagram 100C of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0019] FIG. 1D is a diagram 100D of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0020] FIG. 1E is a diagram 100E of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ (having a connection topology called flip network and also known as inverse shuffle exchange network) of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0021] FIG. 1F is a diagram 100F of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having

US 2011/0044329 A1

Feb. 24, 2011

3

Baseline connection topology of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0022] FIG. 1G is a diagram 100G of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0023] FIG. 1H is a diagram 100H of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0024] FIG. 1I is a diagram 100I of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ (having a connection topology built using back-to-back Banyan Networks or back-to-back Delta Networks or equivalently back-to-back Butterfly networks) of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0025] FIG. 1J is a diagram 100J of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0026] FIG. 1K is a diagram 100K of a general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages with $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0027] FIG. 1A1 is a diagram 100A1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having inverse Benes connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0028] FIG. 1B1 is a diagram 100B1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Omega Networks) of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0029] FIG. 1C1 is a diagram 100C1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0030] FIG. 1D1 is a diagram 100D1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0031] FIG. 1E1 is a diagram 100E1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology called flip network and also known as inverse shuffle exchange network) of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0032] FIG. 1F1 is a diagram 100F1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having Baseline connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0033] FIG. 1G1 is a diagram 100G1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0034] FIG. 1H1 is a diagram 100H1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0035] FIG. 1I1 is a diagram 100I1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Banyan Networks or back-to-back Delta Networks or equivalently back-to-back Butterfly networks) of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0036] FIG. 1J1 is a diagram 100J1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0037] FIG. 1K1 is a diagram 100K1 of a general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages with $N_1=p*N_2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0038] FIG. 1A2 is a diagram 100A2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having inverse Benes connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

US 2011/0044329 A1

Feb. 24, 2011

4

ably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0039] FIG. 1B2 is a diagram 100B2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Omega Networks) of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0040] FIG. 1C2 is a diagram 100C2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0041] FIG. 1D2 is a diagram 100D2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0042] FIG. 1E2 is a diagram 100E2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology called flip network and also known as inverse shuffle exchange network) of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0043] FIG. 1F2 is a diagram 100F2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having Baseline connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0044] FIG. 1G2 is a diagram 100G2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0045] FIG. 1H2 is a diagram 100H2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0046] FIG. 1I2 is a diagram 100I2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Banyan Networks or back-to-back Delta Networks or equivalently back-to-back Butterfly networks) of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0047] FIG. 1J2 is a diagram 100J2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$

having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0048] FIG. 1K2 is a diagram 100K2 of a general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages with $N_2=p*N_1$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0049] FIG. 2A is a diagram 200A of an exemplary symmetrical folded multi-link multi-stage network $V_{fold-link}(N, d, s)$ having inverse Benes connection topology of five stages with $N=8$, $d=2$ and $s=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0050] FIG. 2B is a diagram 200B of a general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, 2)$ with $(2 \times \log_d N) - 1$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0051] FIG. 2C is a diagram 200C of an exemplary asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, 2)$ having inverse Benes connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, and $d=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0052] FIG. 2D is a diagram 200D of a general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, 2)$ with $N_2=p*N_1$ and with $(2 \times \log_d N) - 1$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0053] FIG. 2E is a diagram 200E of an exemplary asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, 2)$ having inverse Benes connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, and $d=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0054] FIG. 2F is a diagram 200F of a general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, 2)$ with $N_1=p*N_2$ and with $(2 \times \log_d N) - 1$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0055] FIG. 3A is a diagram 300A of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having inverse Benes connection topology of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0056] FIG. 3B is a diagram 300B of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ (having a connection topology built using back-to-back Omega Networks) of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

US 2011/0044329 A1

Feb. 24, 2011

5

[0057] FIG. 3C is a diagram 300C of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0058] FIG. 3D is a diagram 300D of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0059] FIG. 3E is a diagram 300E of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ (having a connection topology called flip network and also known as inverse shuffle exchange network) of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0060] FIG. 3F is a diagram 300F of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having Baseline connection topology of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0061] FIG. 3G is a diagram 300G of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0062] FIG. 3H is a diagram 300H of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0063] FIG. 3I is a diagram 300I of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ (having a connection topology built using back-to-back Banyan Networks or back-to-back Delta Networks or equivalently back-to-back Butterfly networks) of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0064] FIG. 3J is a diagram 300J of an exemplary symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0065] FIG. 3K is a diagram 300K of a general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages with $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0066] FIG. 3A1 is a diagram 300A1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having inverse Benes connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0067] FIG. 3B1 is a diagram 300B1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Omega Networks) of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections.

[0068] FIG. 3C1 is a diagram 300C1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly

nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0069] FIG. 3D1 is a diagram 300D1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0070] FIG. 3E1 is a diagram 300E1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology called flip network and also known as inverse shuffle exchange network) of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0071] FIG. 3F1 is a diagram 300F1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having Baseline connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0072] FIG. 3G1 is a diagram 300G1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0073] FIG. 3H1 is a diagram 300H1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0074] FIG. 3I1 is a diagram 300I1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Banyan Networks or back-to-back Delta Networks or equivalently back-to-back Butterfly networks) of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0075] FIG. 3J1 is a diagram 300J1 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0076] FIG. 3K1 is a diagram 300K1 of a general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages with $N_1=p*N_2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0077] FIG. 3A2 is a diagram 300A2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ having inverse Benes connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0078] FIG. 3B2 is a diagram 300B2 of an exemplary asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Omega Networks) of five stages with $N_2=8$, $N_1=p*N_2=24$,

US 2011/0044329 A1

Feb. 24, 2011

6

where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0079] FIG. 3C2 is a diagram 300C2 of an exemplary asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0080] FIG. 3D2 is a diagram 300D2 of an exemplary asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0081] FIG. 3E2 is a diagram 300E2 of an exemplary asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ (having a connection topology called flip network and also known as inverse shuffle exchange network) of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0082] FIG. 3F2 is a diagram 300F2 of an exemplary asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ having Baseline connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0083] FIG. 3G2 is a diagram 300G2 of an exemplary asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0084] FIG. 3H2 is a diagram 300H2 of an exemplary asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0085] FIG. 3I2 is a diagram 300I2 of an exemplary asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Banyan Networks or back-to-back Delta Networks or equivalently back-to-back Butterfly networks) of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0086] FIG. 3J2 is a diagram 300J2 of an exemplary asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0087] FIG. 3K2 is a diagram 300K2 of a general asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages with $N_2=p*N_1$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0088] FIG. 4A is a diagram 400A of an exemplary symmetrical folded multi-stage network $V_{fold}(N, d, s)$ having inverse Benes connection topology of five stages with $N=8$, $d=2$ and $s=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrange-

ably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0089] FIG. 4A1 is a diagram 400A1 of an exemplary symmetrical folded multi-stage network $V_{fold}(N, d, 2)$ having Omega connection topology of five stages with $N=8$, $d=2$ and $s=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0090] FIG. 4A2 is a diagram 400A2 of an exemplary symmetrical folded multi-stage network $V_{fold}(N, d, 2)$ having nearest neighbor connection topology of five stages with $N=8$, $d=2$ and $s=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0091] FIG. 4B is a diagram 400B of a general symmetrical folded multi-stage network $V_{fold}(N, d, 2)$ with $(2 \times \log_d N) - 1$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0092] FIG. 4C is a diagram 400C of an exemplary asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 2)$ having inverse Benes connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, and $d=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0093] FIG. 4C1 is a diagram 400C1 of an exemplary asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 2)$ having Omega connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, and $d=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0094] FIG. 4C2 is a diagram 400C2 of an exemplary asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 2)$ having nearest neighbor connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, and $d=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0095] FIG. 4D is a diagram 400D of a general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 2)$ with $N_2=p*N_1$ and with $(2 \times \log_d N) - 1$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0096] FIG. 4E is a diagram 400E of an exemplary asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 2)$ having inverse Benes connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, and $d=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0097] FIG. 4E1 is a diagram 400E1 of an exemplary asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 2)$ having Omega connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, and $d=2$ with exemplary multicast connections, strictly nonblocking network for unicast con-

US 2011/0044329 A1

Feb. 24, 2011

7

nections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0098] FIG. 4E2 is a diagram 400E2 of an exemplary asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 2)$ having nearest neighbor connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, and $d=2$ with exemplary multicast connections, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0099] FIG. 4F is a diagram 400F of a general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 2)$ with $N_1=p*N_2$ and with $(2*\log_d N)-1$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0100] FIG. 5A is a diagram 500A of an exemplary symmetrical folded multi-stage network $V_{fold}(N, d, s)$ having inverse Benes connection topology of five stages with $N=8$, $d=2$ and $s=1$ with exemplary unicast connections rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0101] FIG. 5B is a diagram 500B of a general symmetrical folded multi-stage network $V_{fold}(N, d, 1)$ with $(2*\log_d N)-1$ stages rearrangeably nonblocking network for unicast connections in accordance with the invention.

[0102] FIG. 5C is a diagram 500C of an exemplary asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 1)$ having inverse Benes connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, and $d=2$ with exemplary unicast connections rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0103] FIG. 5D is a diagram 500D of a general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 1)$ with $N_2=p*N_1$ and with $(2*\log_d N)-1$ stages rearrangeably nonblocking network for unicast connections in accordance with the invention.

[0104] FIG. 5E is a diagram 500E of an exemplary asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 1)$ having inverse Benes connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, and $d=2$ with exemplary unicast connections rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0105] FIG. 5F is a diagram 500F of a general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, 1)$ with $N_1=p*N_2$ and with $(2*\log_d N)-1$ stages rearrangeably nonblocking network for unicast connections in accordance with the invention.

[0106] FIG. 6A is a diagram 600A of an exemplary symmetrical multi-stage network $V(N, d, s)$ having inverse Benes connection topology of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0107] FIG. 6B is a diagram 600B of an exemplary symmetrical multi-stage network $V(N, d, s)$ (having a connection topology built using back-to-back Omega Networks) of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0108] FIG. 6C is a diagram 600C of an exemplary symmetrical multi-stage network $V(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0109] FIG. 6D is a diagram 600D of an exemplary symmetrical multi-stage network $V(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0110] FIG. 6E is a diagram 600E of an exemplary symmetrical multi-stage network $V(N, d, s)$ (having a connection topology called flip network and also known as inverse shuffle exchange network) of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0111] FIG. 6F is a diagram 600F of an exemplary symmetrical multi-stage network $V(N, d, s)$ having Baseline connection topology of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0112] FIG. 6G is a diagram 600G of an exemplary symmetrical multi-stage network $V(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0113] FIG. 6H is a diagram 600H of an exemplary symmetrical multi-stage network $V(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0114] FIG. 6I is a diagram 600I of an exemplary symmetrical multi-stage network $V(N, d, s)$ (having a connection topology built using back-to-back Banyan Networks or back-to-back Delta Networks or equivalently back-to-back Butterfly networks) of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0115] FIG. 6J is a diagram 600J of an exemplary symmetrical multi-stage network $V(N, d, s)$ having an exemplary connection topology of five stages with $N=8$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0116] FIG. 6K is a diagram 600K of a general symmetrical multi-stage network $V(N, d, s)$ with $(2*\log_d N)-1$ stages with $s=1$, rearrangeably nonblocking network for unicast connections in accordance with the invention.

[0117] FIG. 6A1 is a diagram 600A1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having inverse Benes connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0118] FIG. 6B1 is a diagram 600B1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Omega Networks) of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0119] FIG. 6C1 is a diagram 600C1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0120] FIG. 6D1 is a diagram 600D1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0121] FIG. 6E1 is a diagram 600E1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ (having a con-

US 2011/0044329 A1

Feb. 24, 2011

8

nection topology called flip network and also known as inverse shuffle exchange network) of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections.

[0122] FIG. 6F1 is a diagram 600F1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having Baseline connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0123] FIG. 6G1 is a diagram 600G1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections, in accordance with the invention.

[0124] FIG. 6H1 is a diagram 600H1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections, in accordance with the invention.

[0125] FIG. 6I1 is a diagram 600I1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Banyan Networks or back-to-back Delta Networks or equivalently back-to-back Butterfly networks) of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0126] FIG. 6J1 is a diagram 600J1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_1=8$, $N_2=p*N_1=24$ where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections.

[0127] FIG. 6K1 is a diagram 600K1 of a general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages with $N_1=p*N_2$ and $s=1$, rearrangeably nonblocking network for unicast connections in accordance with the invention.

[0128] FIG. 6A2 is a diagram 600A2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having inverse Benes connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections, in accordance with the invention.

[0129] FIG. 6B2 is a diagram 600B2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Omega Networks) of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0130] FIG. 6C2 is a diagram 600C2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections, in accordance with the invention.

[0131] FIG. 6D2 is a diagram 600D2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections, in accordance with the invention.

[0132] FIG. 6E2 is a diagram 600E2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ (having a con-

nection topology called flip network and also known as inverse shuffle exchange network) of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections.

[0133] FIG. 6F2 is a diagram 600F2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having Baseline connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0134] FIG. 6G2 is a diagram 600G2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections, in accordance with the invention.

[0135] FIG. 6H2 is a diagram 600H2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections, in accordance with the invention.

[0136] FIG. 6I2 is a diagram 600I2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ (having a connection topology built using back-to-back Banyan Networks or back-to-back Delta Networks or equivalently back-to-back Butterfly networks) of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably nonblocking network for unicast connections.

[0137] FIG. 6J2 is a diagram 600J2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ having an exemplary connection topology of five stages with $N_2=8$, $N_1=p*N_2=24$, where $p=3$, $d=2$ and $s=1$, rearrangeably non-blocking network for unicast connections.

[0138] FIG. 6K2 is a diagram 600K2 of a general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages with $N_2=p*N_1$ and $s=1$, rearrangeably nonblocking network for unicast connections in accordance with the invention.

[0139] FIG. 7A is high-level flowchart of a scheduling method according to the invention, used to set up the multicast connections in all the networks disclosed in this invention.

[0140] FIG. 8A1 is a diagram 800A1 of an exemplary prior art implementation of a two by two switch; FIG. 8A2 is a diagram 800A2 for programmable integrated circuit prior art implementation of the diagram 800A1 of FIG. 8A1; FIG. 8A3 is a diagram 800A3 for one-time programmable integrated circuit prior art implementation of the diagram 800A1 of FIG. 8A1; FIG. 8A4 is a diagram 800A4 for integrated circuit placement and route implementation of the diagram 800A1 of FIG. 8A1.

DETAILED DESCRIPTION OF THE INVENTION

[0141] The present invention is concerned with the design and operation of large scale crosspoint reduction using arbitrarily large multi-link multi-stage switching networks for broadcast, unicast and multicast connections. Particularly multi-link multi-stage networks with stages more than three and radices greater than or equal to two offer large scale crosspoint reduction when configured with optimal links as disclosed in this invention.

[0142] When a transmitting device simultaneously sends information to more than one receiving device, the one-to-many connection required between the transmitting device

US 2011/0044329 A1

Feb. 24, 2011

9

and the receiving devices is called a multicast connection. A set of multicast connections is referred to as a multicast assignment. When a transmitting device sends information to one receiving device, the one-to-one connection required between the transmitting device and the receiving device is called unicast connection. When a transmitting device simultaneously sends information to all the available receiving devices, the one-to-all connection required between the transmitting device and the receiving devices is called a broadcast connection.

[0143] In general, a multicast connection is meant to be one-to-many connection, which includes unicast and broadcast connections. A multicast assignment in a switching network is nonblocking if any of the available inlet links can always be connected to any of the available outlet links.

[0144] In certain multi-link multi-stage networks, folded multi-link multi-stage networks, and folded multi-stage networks of the type described herein, any connection request of arbitrary fan-out, i.e. from an inlet link to an outlet link or to a set of outlet links of the network, can be satisfied without blocking if necessary by rearranging some of the previous connection requests. In certain other multi-link multi-stage networks of the type described herein, any connection request of arbitrary fan-out, i.e. from an inlet link to an outlet link or to a set of outlet links of the network, can be satisfied without blocking with never needing to rearrange any of the previous connection requests.

[0145] In certain multi-link multi-stage networks, folded multi-link multi-stage networks, and folded multi-stage networks of the type described herein, any connection request of unicast from an inlet link to an outlet link of the network, can be satisfied without blocking if necessary by rearranging some of the previous connection requests. In certain other multi-link multi-stage networks of the type described herein, any connection request of unicast from an inlet link to an outlet link of the network, can be satisfied without blocking with never needing to rearrange any of the previous connection requests.

[0146] Nonblocking configurations for other types of networks with numerous connection topologies and scheduling methods are disclosed as follows:

[0147] 1) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-stage networks $V(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the U.S. application Ser. No. 12/530,207 that is incorporated by reference above.

[0148] 2) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above.

[0149] 3) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above.

[0150] 4) VLSI layouts of generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link

multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ for $s=1, 2, 3$ or any number in general, are described in detail in the PCT Application Serial No. PCT/US08/64605 that is incorporated by reference above.

[0151] 5) VLSI layouts of numerous types of multi-stage networks with locality exploitation are described in U.S. Provisional Patent Application Ser. No. 61/252,603 that is incorporated by reference above.

[0152] 6) VLSI layouts of numerous types of multistage pyramid networks are described in U.S. Provisional Patent Application Ser. No. 61/252,609 that is incorporated by reference above.

RNB Multi-Link Multi-Stage Embodiments:

Symmetric RNB Embodiments:

[0153] Referring to FIG. 1A, in one embodiment, an exemplary symmetrical multi-link multi-stage network **100A** with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by four switches **IS1-IS4** and output stage **120** consists of four, four by two switches **OS1-OS4**. And all the middle stages namely middle stage **130** consists of four, four by four switches **MS(1,1)-MS(1,4)**, middle stage **140** consists of four, four by four switches **MS(2,1)-MS(2,4)**, and middle stage **150** consists of four, four by four switches **MS(3,1)-MS(3,4)**.

[0154] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0155] In one embodiment of this network each of the input switches **IS1-IS4** and output switches **OS1-OS4** are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch **IS1-IS4** can be denoted in general with the notation $d*2d$ and each output switch **OS1-OS4** can be denoted in general with the notation $2d*d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d*2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links **IL1-IL8**), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not neces-

US 2011/0044329 A1

Feb. 24, 2011

10

sary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0156] Each of the N/d input switches IS1-IS4 are connected to exactly 2xd switches in middle stage 130 through 2xd links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and also to middle switch MS(1,2) through the links ML(1,3) and ML(1,4)).

[0157] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through 2xd links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to exactly d switches in middle stage 140 through 2xd links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0158] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through 2xd links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,11) and ML(2,12) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through 2xd links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,3)).

[0159] Similarly each of the N/d middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through 2xd links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,11) and ML(3,12) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly d output switches in output stage 120 through 2xd links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from Middle switch MS(3,1), and the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1)).

[0160] Each of the N/d output switches OS1-OS4 are connected from exactly 2xd switches in middle stage 150 through 2xd links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2), and output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,7) and ML(4,8)).

[0161] Finally the connection topology of the network 100A shown in FIG. 1A is known to be back to back inverse Benes connection topology.

[0162] Referring to FIG. 1B, in another embodiment of network $V_{mlink}(N, d, s)$, an exemplary symmetrical multi-link multi-stage network 100B with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage 130

consists of four, four by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, four by four switches MS(3,1)-MS(3,4).

[0163] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0164] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d, where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d. The size of each input switch IS1-IS4 can be denoted in general with the notation $d^2 \times 2d$ and each output switch OS1-OS4 can be denoted in general with the notation $2d \times d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d \times 2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The symmetric multi-link multi-stage network of FIG. 1B is also the network of the type $V_{mlink}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0165] Each of the N/d input switches IS1-IS4 are connected to exactly 2xd switches in middle stage 130 through 2xd links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and also to middle switch MS(1,2) through the links ML(1,3) and ML(1,4)).

[0166] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through 2xd links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,9) and ML(1,10) are connected to the middle switch MS(1,1) from input switch IS3) and also are connected to exactly d switches in middle stage 140 through 2xd links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0167] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through 2xd links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,9) and ML(2,10) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through

US 2011/0044329 A1

Feb. 24, 2011

11

$2 \times d$ links (for example the links $ML(3,1)$ and $ML(3,2)$ are connected from middle switch $MS(2,1)$ to middle switch $MS(3,1)$, and the links $ML(3,3)$ and $ML(3,4)$ are connected from middle switch $MS(2,1)$ to middle switch $MS(3,2)$).

[0168] Similarly each of the N/d middle switches $MS(3,1)$ - $MS(3,4)$ in the middle stage 150 are connected from exactly d switches in middle stage 140 through $2 \times d$ links (for example the links $ML(3,1)$ and $ML(3,2)$ are connected to the middle switch $MS(3,1)$ from middle switch $MS(2,1)$, and the links $ML(3,9)$ and $ML(3,10)$ are connected to the middle switch $MS(3,1)$ from middle switch $MS(2,3)$) and also are connected to exactly d output switches in output stage 120 through $2 \times d$ links (for example the links $ML(4,1)$ and $ML(4,2)$ are connected to output switch $OS1$ from Middle switch $MS(3,1)$, and the links $ML(4,3)$ and $ML(4,4)$ are connected to output switch $OS2$ from middle switch $MS(3,1)$).

[0169] Each of the N/d output switches $OS1$ - $OS4$ are connected from exactly $2 \times d$ switches in middle stage 150 through $2 \times d$ links (for example output switch $OS1$ is connected from middle switch $MS(3,1)$ through the links $ML(4,1)$ and $ML(4,2)$, and output switch $OS1$ is also connected from middle switch $MS(3,3)$ through the links $ML(4,9)$ and $ML(4,10)$).

[0170] Finally the connection topology of the network 100B shown in FIG. 1B is known to be back to back Omega connection topology.

[0171] Referring to FIG. 1C, in another embodiment of network $V_{minik}(N, d, s)$, an exemplary symmetrical multi-link multi-stage network 100C with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches $IS1$ - $IS4$ and output stage 120 consists of four, four by two switches $OS1$ - $OS4$. And all the middle stages namely middle stage 130 consists of four, four by four switches $MS(1,1)$ - $MS(1,4)$, middle stage 140 consists of four, four by four switches $MS(2,1)$ - $MS(2,4)$, and middle stage 150 consists of four, four by four switches $MS(3,1)$ - $MS(3,4)$.

[0172] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0173] In one embodiment of this network each of the input switches $IS1$ - $IS4$ and output switches $OS1$ - $OS4$ are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch $IS1$ - $IS4$ can be denoted in general with the notation $d \times 2d$ and each output switch $OS1$ - $OS4$ can be denoted in general with the notation $2d \times d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d \times 2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The symmetric multi-link multi-stage network of

FIG. 1C is also the network of the type $V_{minik}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links $IL1$ - $IL8$), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links $IL1$ - $IL8$ as there are outlet links $OL1$ - $OL8$, in a symmetrical network they are the same.

[0174] Each of the N/d input switches $IS1$ - $IS4$ are connected to exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example input switch $IS1$ is connected to middle switch $MS(1,1)$ through the links $ML(1,1)$, $ML(1,2)$, and also to middle switch $MS(1,2)$ through the links $ML(1,3)$ and $ML(1,4)$).

[0175] Each of the N/d middle switches $MS(1,1)$ - $MS(1,4)$ in the middle stage 130 are connected from exactly d input switches through $2 \times d$ links (for example the links $ML(1,1)$ and $ML(1,2)$ are connected to the middle switch $MS(1,1)$ from input switch $IS1$, and the links $ML(1,15)$ and $ML(1,16)$ are connected to the middle switch $MS(1,1)$ from input switch $IS4$) and also are connected to exactly d switches in middle stage 140 through $2 \times d$ links (for example the links $ML(2,1)$ and $ML(2,2)$ are connected from middle switch $MS(1,1)$ to middle switch $MS(2,1)$, and the links $ML(2,3)$ and $ML(2,4)$ are connected from middle switch $MS(1,1)$ to middle switch $MS(2,2)$).

[0176] Similarly each of the N/d middle switches $MS(2,1)$ - $MS(2,4)$ in the middle stage 140 are connected from exactly d switches in middle stage 130 through $2 \times d$ links (for example the links $ML(2,1)$ and $ML(2,2)$ are connected to the middle switch $MS(2,1)$ from middle switch $MS(1,1)$, and the links $ML(2,15)$ and $ML(2,16)$ are connected to the middle switch $MS(2,1)$ from middle switch $MS(1,4)$) and also are connected to exactly d switches in middle stage 150 through $2 \times d$ links (for example the links $ML(3,1)$ and $ML(3,2)$ are connected from middle switch $MS(2,1)$ to middle switch $MS(3,1)$, and the links $ML(3,3)$ and $ML(3,4)$ are connected from middle switch $MS(2,1)$ to middle switch $MS(3,2)$).

[0177] Similarly each of the N/d middle switches $MS(3,1)$ - $MS(3,4)$ in the middle stage 150 are connected from exactly d switches in middle stage 140 through $2 \times d$ links (for example the links $ML(3,1)$ and $ML(3,2)$ are connected to the middle switch $MS(3,1)$ from middle switch $MS(2,1)$, and the links $ML(3,15)$ and $ML(3,16)$ are connected to the middle switch $MS(3,1)$ from middle switch $MS(2,4)$) and also are connected to exactly d output switches in output stage 120 through $2 \times d$ links (for example the links $ML(4,1)$ and $ML(4,2)$ are connected to output switch $OS1$ from middle switch $MS(3,1)$, and the links $ML(4,3)$ and $ML(4,4)$ are connected to output switch $OS2$ from middle switch $MS(3,1)$).

[0178] Each of the N/d output switches $OS1$ - $OS4$ are connected from exactly $2 \times d$ switches in middle stage 150 through $2 \times d$ links (for example output switch $OS1$ is connected from middle switch $MS(3,1)$ through the links $ML(4,1)$ and $ML(4,2)$, and output switch $OS1$ is also connected from middle switch $MS(3,4)$ through the links $ML(4,15)$ and $ML(4,16)$).

[0179] Finally the connection topology of the network 100C shown in FIG. 1C is hereinafter called nearest neighbor connection topology.

[0180] Similar to network 100A of FIG. 1A, 100B of FIG. 1B, and 100C of FIG. 1C, referring to FIG. 1D, FIG. 1E, FIG. 1F, FIG. 1G, FIG. 1H, FIG. 1I and FIG. 1J with exemplary symmetrical multi-link multi-stage networks 100D, 100E,

US 2011/0044329 A1

Feb. 24, 2011

12

100F, 100G, 100H, 100I, and 100J respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130, 140, and 150** is shown where input stage **110** consists of four, two by four switches IS1-IS4 and output stage **120** consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, four by four switches MS(1,1)-MS(1,4), middle stage **140** consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, four by four switches MS(3,1)-MS(3,4).

[0181] Such networks can also be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130, middle stage 140 and middle stage 150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130, middle stage 140 and middle stage 150**.

[0182] The networks **100D, 100E, 100F, 100G, 100H, 100I and 100J** of FIG. 1D, FIG. 1E, FIG. 1F, FIG. 1G, FIG. 1H, FIG. 1I, and FIG. 1J are also embodiments of symmetric multi-link multi-stage network can be represented with the notation $V_{link}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0183] Just like networks of **100A, 100B and 100C**, for all the networks **100D, 100E, 100F, 100G, 100H, 100I and 100J** of FIG. 1D, FIG. 1E, FIG. 1F, FIG. 1G, FIG. 1H, FIG. 1I, and FIG. 1J, each of the N/d input switches IS1-IS4 are connected to exactly $2 \times d$ switches in middle stage **130** through $2 \times d$ links.

[0184] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links.

[0185] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $2 \times d$ links and also are connected to exactly d switches in middle stage **150** through $2 \times d$ links.

[0186] Similarly each of the N/d middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through $2 \times d$ links and also are connected to exactly d output switches in output stage **120** through $2 \times d$ links.

[0187] Each of the N/d output switches OS1-OS4 are connected from exactly $2 \times d$ switches in middle stage **150** through $2 \times d$ links.

[0188] In all the ten embodiments of FIG. 1A to FIG. 1J the connection topology is different. That is the way the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16), and ML(4,1)-ML(4,16) are connected between the respective stages is different. Even though only ten embodi-

ments are illustrated, in general, the network $V_{mlink}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mlink}(N, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{mlink}(N, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N, d, s)$ can be built. The ten embodiments of FIG. 1A to FIG. 1J are only three examples of network $V_{mlink}(N, d, s)$.

[0189] In all the ten embodiments of FIG. 1A to FIG. 1J, each of the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16) and ML(4,1)-ML(4,16) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage **110** is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage **120** is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,4), MS(2,1)-MS(2,4), and MS(3,1)-MS(3,4) are referred to as middle switches or middle ports.

[0190] In the example illustrated in FIG. 1A (or in FIG. 1B to FIG. 1J), a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage **130** will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage **130** when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network **100A** (or **100B** to **100J**), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0191] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage **130** is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage **130**, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the rearrangeably nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Symmetric RNB Embodiments:

[0192] Network **100K** of FIG. 1K is an example of general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages. The general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention. (And in the example of FIG. 1K, $s=2$). The general symmetrical multi-link multi-

US 2011/0044329 A1

Feb. 24, 2011

13

stage network $V_{mlink}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of N/d input switches IS1-IS(N/d) (for example the links IL1-IL(d) to the input switch IS1) and $2 \times d$ outgoing links for each of N/d input switches IS1-IS(N/d) (for example the links ML(1,1)-ML(1,2 d) to the input switch IS1). There are d outlet links for each of N/d output switches OS1-OS(N/d) (for example the links OL1-OL(d) to the output switch OS1) and $2 \times d$ incoming links for each of N/d output switches OS1-OS(N/d) (for example ML($2 \times \log_d N - 2, 1$)-ML($2 \times \log_d N - 2, 2 \times d$) to the output switch OS1).

[0193] Each of the N/d input switches IS1-IS(N/d) are connected to exactly d switches in middle stage **130** through $2 \times d$ links.

[0194] Each of the N/d middle switches MS(1,1)-MS(1, N/d) in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links.

[0195] Similarly each of the N/d middle switches

$$MS(\log_d N - 1, 1) - MS(\log_d N - 1, \frac{N}{d})$$

in the middle stage **130** + $10^* (\log_d N - 2)$ are connected from exactly d switches in middle stage **130** + $10^* (\log_d N - 3)$ through $2 \times d$ links and also are connected to exactly d switches in middle stage **130** + $10^* (\log_d N - 1)$ through $2 \times d$ links.

[0196] Similarly each of the N/d middle switches MS($2 \times \log_d N - 3, 1$)-

$$MS(2 \times \log_d N - 3, \frac{N}{d})$$

in the middle stage **130** + $10^* (2 \times \log_d N - 4)$ are connected from exactly d switches in middle stage **130** + $10^* (2 \times \log_d N - 5)$ through $2 \times d$ links and also are connected to exactly d output switches in output stage **120** through $2 \times d$ links.

[0197] Each of the N/d output switches OS1-OS(N/d) are connected from exactly d switches in middle stage **130** + $10^* (2 \times \log_d N - 4)$ through $2 \times d$ links.

[0198] As described before, again the connection topology of a general $V_{mlink}(N, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink}(N, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{mlink}(N, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N, d, s)$ can be built. The embodiments of FIG. 1A to FIG. 1J are ten examples of network $V_{mlink}(N, d, s)$.

[0199] The general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ can be operated in rearrangeably non-blocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $S \geq 2$ according to the current invention.

[0200] Every switch in the multi-link multi-stage networks discussed herein has multicast capability. In a $V_{mlink}(N, d, s)$

network, if a network inlet link is to be connected to more than one outlet link on the same output switch, then it is only necessary for the corresponding input switch to have one path to that output switch. This follows because that path can be multicast within the output switch to as many outlet links as necessary. Multicast assignments can therefore be described in terms of connections between input switches and output switches. An existing connection or a new connection from an input switch to r' output switches is said to have fan-out r' . If all multicast assignments of a first type, wherein any inlet link of an input switch is to be connected in an output switch to at most one outlet link are realizable, then multicast assignments of a second type, wherein any inlet link of each input switch is to be connected to more than one outlet link in the same output switch, can also be realized. For this reason, the following discussion is limited to general multicast connections of the first type (with fan-out r' ,

$$1 \leq r' \leq \frac{N}{d}$$

although the same discussion is applicable to the second type.

[0201] To characterize a multicast assignment, for each inlet link

$$i \in \{1, 2, \dots, \frac{N}{d}\},$$

let $I_i = O$, where

$$O \subset \{1, 2, \dots, \frac{N}{d}\},$$

denote the subset of output switches to which inlet link i is to be connected in the multicast assignment. For example, the network of FIG. 1C shows an exemplary five-stage network, namely $V_{mlink}(8, 2, 2)$, with the following multicast assignment $I_1 = \{2, 4\}$ and all other $I_j = \emptyset$ for $j = [2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,2) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,2) only once into middle switches MS(2,1) and MS(2,3) respectively in middle stage **140**.

[0202] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,3) only once into middle switches MS(3,2) and MS(3,4) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,2) and MS(3,4) only once into output switches OS2 and OS4 in output stage **120**. Finally the connection I_1 fans out once in the output stage switch OS2 into outlet link OL3 and in the output stage switch OS4 twice into the outlet links OL7 and OL8. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage **130**.

Asymmetric RNB ($N_2 > N_1$) Embodiments:

[0203] Referring to FIG. 1A1, in one embodiment, an exemplary asymmetrical multi-link multi-stage network **100A1** with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call

US 2011/0044329 A1

Feb. 24, 2011

14

or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by four switches IS1-IS4 and output stage **120** consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, four by four switches MS(1,1)-MS(1,4), middle stage **140** consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, four by eight switches MS(3,1)-MS(3,4).

[0204] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size eight by six, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0205] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_1}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d * 2d$ and each output switch OS1-OS4 can be denoted in general with the notation $(d+d_2) * d_2$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $2d * 2d$. The size of each switch in the last middle stage can be denoted as $2d * (d+d_2)$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0206] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly d switches in middle stage **130** through $2 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and also to middle switch MS(1,2) through the links ML(1,3) and ML(1,4)).

[0207] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0208] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,11) and ML(2,12) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage **150** through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,3)).

[0209] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,11) and ML(3,12) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly

US 2011/0044329 A1

Feb. 24, 2011

15

$$\frac{d+d_2}{2}$$

output switches in output stage **120** through $d+d_2$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from Middle switch MS(3,1); the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1); the links ML(4,5) and ML(4,6) are connected to output switch OS3 from Middle switch MS(3,1); and the links ML(4,7) and ML(4,8) are connected to output switch OS4 from middle switch MS(3,1)).

[0210] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly

$$\frac{d+d_2}{2}$$

switches in middle stage **150** through $d+d_2$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,9) and ML(4,10); output switch OS1 is connected from middle switch MS(3,3) through the links ML(4,17) and ML(4,18); and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,25) and ML(4,26)).

[0211] Finally the connection topology of the network **100A1** shown in FIG. **1A1** is known to be back to back inverse Benes connection topology.

[0212] Referring to FIG. **1B1**, in one embodiment, an exemplary asymmetrical multi-link multi-stage network **100B1** with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by four switches IS1-IS4 and output stage **120** consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, four by four switches MS(1,1)-MS(1,4), middle stage **140** consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, four by eight switches MS(3,1)-MS(3,4).

[0213] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size eight by six, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0214] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_1}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d * 2d$ and each output switch OS1-OS4 can be denoted in general with the notation $(d+d_2) * d_2$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $2d * 2d$. The size of each switch in the last middle stage can be denoted as $2d * (d+d_2)$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{mink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0215] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly $2 \times d$ switches in middle stage **130** through $2 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and also to middle switch MS(1,2) through the links ML(1,3) and ML(1,4)).

[0216] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,9) and ML(1,10) are connected to the

US 2011/0044329 A1

Feb. 24, 2011

16

middle switch MS(1,1) from input switch IS3) and also are connected to exactly d switches in middle stage 140 through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0217] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,9) and ML(2,10) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0218] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,9) and ML(3,10) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly

$$\frac{d+d_2}{2}$$

output switches in output stage 120 through $d+d_2$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from middle switch MS(3,1); the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1); the links ML(4,5) and ML(4,6) are connected to output switch OS3 from Middle switch MS(3,1); and the links ML(4,7) and ML(4,8) are connected to output switch OS4 from middle switch MS(3,1)).

[0219] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly

$$\frac{d+d_2}{2}$$

switches in middle stage 150 through $d+d_2$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,9) and ML(4,10); output switch OS1 is connected from middle switch MS(3,3) through the links ML(4,17) and ML(4,18); and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,25) and ML(4,26)).

[0220] Finally the connection topology of the network 100B1 shown in FIG. 1B1 is known to be back to back Omega connection topology.

[0221] Referring to FIG. 1C1, in one embodiment, an exemplary asymmetrical multi-link multi-stage network 100C1 with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, four by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, four by eight switches MS(3,1)-MS(3,4).

[0222] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size eight by six, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0223] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p \times N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_1}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d \times 2d$ and each output switch OS1-OS4 can be denoted in general with the notation $(d+d_2) \times d_2$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

US 2011/0044329 A1

Feb. 24, 2011

17

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $2d \times 2d$. The size of each switch in the last middle stage can be denoted as $2d \times (d + d_2)$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{m-link}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0224] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and also to middle switch MS(1,2) through the links ML(1,3) and ML(1,4)).

[0225] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through $2 \times d$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,15) and ML(1,16) are connected to the middle switch MS(1,1) from input switch IS4) and also are connected to exactly d switches in middle stage 140 through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0226] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,15) and ML(2,16) are connected to the middle switch MS(2,1) from middle switch MS(1,4)) and also are connected to exactly d switches in middle stage 150 through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0227] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,15) and ML(3,16) are connected to the middle switch MS(3,1) from middle switch MS(2,4)) and also are connected to exactly

$$\frac{d + d_2}{2}$$

output switches in output stage 120 through $d + d_2$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from middle switch MS(3,1); the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1); the links ML(4,5) and ML(4,6) are connected to output switch OS3 from Middle switch MS(3,1); and the links ML(4,7) and ML(4,8) are connected to output switch OS4 from middle switch MS(3,1)).

[0228] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly

$$\frac{d + d_2}{2}$$

switches in middle stage 150 through $d + d_2$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,9) and ML(4,10); output switch OS1 is connected from middle switch MS(3,3) through the links ML(4,17) and ML(4,18); and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,25) and ML(4,26)).

[0229] Finally the connection topology of the network 100C1 shown in FIG. 1C1 is hereinafter called nearest neighbor connection topology.

[0230] Similar to network 100A1 of FIG. 1A1, 100B1 of FIG. 1B1, and 100C1 of FIG. 1C1, referring to FIG. 1D1, FIG. 1E1, FIG. 1F1, FIG. 1G1, FIG. 1H1, FIG. 1I1 and FIG. 1J1 with exemplary asymmetrical multi-link multi-stage networks 100D1, 100E1, 100F1, 100G1, 100H1, 100I1, and 100J1 respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, four by four switches MS(1,1)-MS(1,4),

US 2011/0044329 A1

Feb. 24, 2011

18

middle stage **140** consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, four by four switches MS(3,1)-MS(3,4).

[0231] Such networks can also be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0232] The networks **100D1**, **100E1**, **100F1**, **100G1**, **100H1**, **100I1** and **100J1** of FIG. 1D1, FIG. 1E1, FIG. 1F1, FIG. 1G1, FIG. 1H1, FIG. 1I1, and FIG. 1J1 are also embodiments of asymmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0233] Just like networks of **100A1**, **100B1** and **100C1**, for all the networks **100D1**, **100E1**, **100F1**, **100G1**, **100H1**, **100I1** and **100J1** of FIG. 1D1, FIG. 1E1, FIG. 1F1, FIG. 1G1, FIG. 1H1, FIG. 1I1, and FIG. 1J1, each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly d switches in middle stage **130** through $2 \times d$ links.

[0234] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links.

[0235] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $2 \times d$ links and also are connected to exactly d switches in middle stage **150** through $2 \times d$ links.

[0236] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through $2 \times d$ links and also are connected to exactly

$$\frac{d + d_2}{2}$$

output switches in output stage **120** through $d + d_2$ links.

[0237] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly

$$\frac{d + d_2}{2}$$

switches in middle stage **150** through $d + d_2$ links.

[0238] In all the ten embodiments of FIG. 1A1 to FIG. 1J1 the connection topology is different. That is the way the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16), and ML(4,1)-ML(4,16) are connected between the respective stages is different. Even though only ten embodiments are illustrated, in general, the network $V_{mlink}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mlink}(N_1, N_2, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{mlink}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N_1, N_2, d, s)$ can be built. The ten embodiments of FIG. 1A1 to FIG. 1J1 are only three examples of network $V_{mlink}(N_1, N_2, d, s)$.

[0239] In all the ten embodiments of FIG. 1A1 to FIG. 1J1, each of the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16) and ML(4,1)-ML(4,16) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage **110** is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage **120** is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,4), MS(2,1)-MS(2,4), and MS(3,1)-MS(3,4) are referred to as middle switches or middle ports.

[0240] In the example illustrated in FIG. 1A1 (or in FIG. 1B1 to FIG. 1J1), a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage **130** will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage **130** when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle

US 2011/0044329 A1

Feb. 24, 2011

19

switches permits the network **100A1** (or **100B1** to **100J1**), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0241] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage **130** is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage **130**, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the rearrangeably nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Asymmetric RNB ($N_2 > N_1$) Embodiments:

[0242] Network **100K1** of FIG. **1K1** is an example of general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_1) - 1$ stages where $N_2 > N_1$ and $N = p * N_1$, where $p > 1$. In network **100K1** of FIG. **1K1**, $N_1 = N$ and $N_2 = p * N$. The general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention. (And in the example of FIG. **1K1**, $s = 2$). The general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_1) - 1$ stages has d inlet links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links IL1-IL(d) to the input switch IS1) and $2 \times d$ outgoing links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links ML(1,1)-ML(1,2d) to the input switch IS1). There are d_2 (where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d)$$

outlet links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example the links OL1-OL($p \times d$) to the output switch OS1) and $d + d_2$ ($= d + p \times d$) incoming links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example ML($2 \times \log_d N_1 - 2, 1$)-ML($2 \times \log_d N_1 - 2, d + d_2$) to the output switch OS1).

[0243] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) are connected to exactly $2 \times d$ switches in middle stage **130** through $2 \times d$ links.

[0244] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1, N_1/d) in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links.

[0245] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(\log_d N_1 - 1, 1) - MS(\log_d N_1 - 1, \frac{N_1}{d})$$

in the middle stage **130** + $10 * (\log_d N_1 - 2)$ are connected from exactly d switches in middle stage **130** + $10 * (\log_d N_1 - 3)$ through $2 \times d$ links and also are connected to exactly d switches in middle stage **130** + $10 * (\log_d N_1 - 1)$ through $2 \times d$ links.

[0246] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(2 \times \log_d N_1 - 3, 1) - MS(2 \times \log_d N_1 - 3, \frac{N_1}{d})$$

in the middle stage **130** + $10 * (2 * \log_d N_1 - 4)$ are connected from exactly d switches in middle stage **130** + $10 * (2 * \log_d N_1 - 5)$ through $2 \times d$ links and also are connected to exactly

US 2011/0044329 A1

Feb. 24, 2011

20

$$\frac{d+d_2}{2}$$

output switches in output stage **120** through $d+d_2$ links.

[0247] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) are connected from exactly

$$\frac{d+d_2}{2}$$

switches in middle stage **130**+ $10*(2*\text{Log}_d N_1-4)$ through $d+d_2$ links

[0248] As described before, again the connection topology of a general $V_{mlink}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{mlink}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1A1 to FIG. 1J1 are ten examples of network $V_{mlink}(N_1, N_2, d, s)$ for $s=2$ and $N_2>N_1$.

[0249] The general symmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general symmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention.

[0250] For example, the network of FIG. 1C1 shows an exemplary five-stage network, namely $V_{mlink}(8, 24, 2, 2)$, with the following multicast assignment $I_1=\{1,4\}$ and all other $I_j=\phi$ for $j=[2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,2) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,2) only once into middle switches MS(2,1) and MS(2,3) respectively in middle stage **140**.

[0251] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,3) only once into middle switches MS(3,1) and MS(3,4) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,4) only once into output switches OS1 and OS4 in output stage **120**. Finally the connection I_1 fans out once in the output stage switch OS1 into outlet link OL2 and in the output stage switch OS4 twice into the outlet links OL20 and OL23. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage **130**.

Asymmetric RNB ($N_1>N_2$) Embodiments:

[0252] Referring to FIG. 1A2, in one embodiment, an exemplary asymmetrical multi-link multi-stage network

100A2 with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, six by eight switches IS1-IS4 and output stage **120** consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, eight by four switches MS(1,1)-MS(1,4), middle stage **140** consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, four by four switches MS(3,1)-MS(3,4).

[0253] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0254] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_1>N_2$ and $N_1=p*N_2$ where $p>1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_2}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d_1*(d+d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $(2*d*d)$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $2d*2d$. The size of each switch in the first middle stage can be denoted as $(d+d_1)*2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each input switch where $N_1>N_2$,

US 2011/0044329 A1

Feb. 24, 2011

21

and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0255] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly

$$\frac{d+d_1}{2}$$

switches in middle stage **130** through $d+d_1$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2); input switch IS1 is connected to middle switch MS(1,2) through the links ML(1,3) and ML(1,4); input switch IS1 is connected to middle switch MS(1,3) through the links ML(1,5), ML(1,6); and input switch IS1 is also connected to middle switch MS(1,4) through the links ML(1,7) and ML(1,8)).

[0256] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly

$$\frac{d+d_1}{2}$$

input switches through $d+d_1$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1; the links ML(1,9) and ML(1,10) are connected to the middle switch MS(1,1) from input switch IS2; the links ML(1,17) and ML(1,18) are connected to the middle switch MS(1,1) from input switch IS3; and the links ML(1,25) and ML(1,26) are connected to the middle switch MS(1,1) from input switch IS4) and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0257] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,11) and ML(2,12) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage **150** through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch

MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,3)).

[0258] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,11) and ML(3,12) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly d output switches in output stage **120** through $2 \times d$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from middle switch MS(3,1); and the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1)).

[0259] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage **150** through $2 \times d$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); and output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,7) and ML(4,8)).

[0260] Finally the connection topology of the network **100A2** shown in FIG. 1 A2 is known to be back to back inverse Benes connection topology.

[0261] Referring to FIG. 1B2, in one embodiment, an exemplary asymmetrical multi-link multi-stage network **100B2** with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, six by eight switches IS1-IS4 and output stage **120** consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, eight by four switches MS(1,1)-MS(1,4), middle stage **140** consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, four by four switches MS(3,1)-MS(3,4).

[0262] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0263] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar

US 2011/0044329 A1

Feb. 24, 2011

22

switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_2}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d_1 * (d + d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $(2 * d * d)$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $2d * 2d$. The size of each switch in the first middle stage can be denoted as $(d + d_1) * 2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{link}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0264] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly

$$\frac{d + d_1}{2}$$

switches in middle stage **130** through $d + d_1$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2); input switch IS1 is connected to middle switch MS(1,2) through the links ML(1,3) and ML(1,4); input switch IS1 is connected to middle switch MS(1,3) through the links ML(1,5), ML(1,6); and input switch IS1 is also connected to middle switch MS(1,4) through the links ML(1,7) and ML(1,8)).

[0265] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly

$$\frac{d + d_1}{2}$$

input switches through $d + d_1$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1; the links ML(1,9) and ML(1,10) are connected to the middle switch MS(1,1) from input switch IS2; the links ML(1,17) and ML(1,18) are connected to the middle switch MS(1,1) from input switch IS3; and the links ML(1,25) and ML(1,26) are connected to the middle switch MS(1,1) from input switch IS4) and also are connected to exactly d switches in middle stage **140** through $2 * d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0266] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $2 * d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,9) and ML(2,10) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage **150** through $2 * d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0267] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through $2 * d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,9) and ML(3,10) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly d output switches in output stage **120** through $2 * d$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from middle switch MS(3,1); and the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1)).

US 2011/0044329 A1

Feb. 24, 2011

23

[0268] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through $2 \times d$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); and output switch OS1 is also connected from middle switch MS(3,3) through the links ML(4,9) and ML(4,10)).

[0269] Finally the connection topology of the network 100B2 shown in FIG. 1B2 is known to be back to back Omega connection topology.

[0270] Referring to FIG. 1C2, in one embodiment, an exemplary asymmetrical multi-link multi-stage network 100C2 with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, six by eight switches IS1-IS4 and output stage 120 consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, eight by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, four by four switches MS(3,1)-MS(3,4).

[0271] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size six by eight, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size six by eight, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0272] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_2}{d}$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_2}{d}$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d_1 * (d + d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $(2 \times d * d)$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $2d * 2d$. The size of each switch in the first middle stage can be denoted as $(d + d_1) * 2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0273] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly

$$\frac{d + d_1}{2}$$

switches in middle stage 130 through $d + d_1$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2); input switch IS1 is connected to middle switch MS(1,2) through the links ML(1,3) and ML(1,4); input switch IS1 is connected to middle switch MS(1,3) through the links ML(1,5), ML(1,6); and input switch IS1 is also connected to middle switch MS(1,4) through the links ML(1,7) and ML(1,8)).

[0274] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly

$$\frac{d + d_1}{2}$$

input switches through $d + d_1$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1; the links ML(1,9) and ML(1,10) are connected to the middle switch MS(1,1) from input switch IS2; the links ML(1,17) and ML(1,18) are connected to the middle switch MS(1,1) from input switch IS3; and the links ML(1,25) and ML(1,26) are connected to the middle switch MS(1,1) from input switch IS4) and also are connected to exactly d switches in middle stage 140 through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and

US 2011/0044329 A1

Feb. 24, 2011

24

the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0275] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,15) and ML(2,16) are connected to the middle switch MS(2,1) from middle switch MS(1,4)) and also are connected to exactly d switches in middle stage 150 through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0276] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,15) and ML(3,16) are connected to the middle switch MS(3,1) from middle switch MS(2,4)) and also are connected to exactly d output switches in output stage 120 through $2 \times d$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from middle switch MS(3,1); and the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1)).

[0277] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through $2 \times d$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,15) and ML(4,16)).

[0278] Finally the connection topology of the network 100C2 shown in FIG. 1C2 is hereinafter called nearest neighbor connection topology.

[0279] Similar to network 100A2 of FIG. 1A2, 100B2 of FIG. 1B2, and 100C2 of FIG. 1C2, referring to FIG. 1D2, FIG. 1E2, FIG. 1F2, FIG. 1G2, FIG. 1H2, FIG. 1I2 and FIG. 1J2 with exemplary asymmetrical multi-link multi-stage networks 100D2, 100E2, 100F2, 100G2, 100H2, 100I2, and 100J2 respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where

input stage 110 consists of four, six by eight switches IS1-IS4 and output stage 120 consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, eight by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, four by four switches MS(3,1)-MS(3,4).

[0280] Such networks can also be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size six by eight, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0281] The networks 100D2, 100E2, 100F2, 100G2, 100H2, 100I2 and 100J2 of FIG. 1D2, FIG. 1E2, FIG. 1F2, FIG. 1G2, FIG. 1H2, FIG. 1I2, and FIG. 1J2 are also embodiments of asymmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0282] Just like networks of 100A2, 100B2 and 100C2, for all the networks 100D2, 100E2, 100F2, 100G2, 100H2, 100I2 and 100J2 of FIG. 1D2, FIG. 1E2, FIG. 1F2, FIG. 1G2, FIG. 1H2, FIG. 1I2, and FIG. 1J2, each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly

$$\frac{d + d_2}{2}$$

switches in middle stage 130 through $d + d_2$ links.

[0283] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly

$$\frac{d + d_2}{2}$$

input switches through $d + d_2$ links and also are connected to exactly d switches in middle stage 140 through $2 \times d$ links.

US 2011/0044329 A1

Feb. 24, 2011

25

[0284] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $2 \times d$ links and also are connected to exactly d switches in middle stage 150 through $2 \times d$ links.

[0285] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $2 \times d$ links and also are connected to exactly d output switches in output stage 120 through $2 \times d$ links.

[0286] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through $2 \times d$ links.

[0287] In all the ten embodiments of FIG. 1A2 to FIG. 1J2 the connection topology is different. That is the way the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16), and ML(4,1)-ML(4,16) are connected between the respective stages is different. Even though only ten embodiments are illustrated, in general, the network $V_{mlink}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mlink}(N_1, N_2, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{mlink}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N_1, N_2, d, s)$ can be built. The ten embodiments of FIG. 1A2 to FIG. 1J2 are only three examples of network $V_{mlink}(N_1, N_2, d, s)$.

[0288] In all the ten embodiments of FIG. 1A2 to FIG. 1J2, each of the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16) and ML(4,1)-ML(4,16) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage 110 is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage 120 is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,4), MS(2,1)-MS(2,4), and MS(3,1)-MS(3,4) are referred to as middle switches or middle ports.

[0289] In the example illustrated in FIG. 1A2 (or in FIG. 1B2 to FIG. 1J2), a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage 130 will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of

two is used. The specific middle switches that are chosen in middle stage 130 when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network 100A2 (or 100B2 to 100J2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0290] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage 130 is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage 130, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the rearrangeably nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Asymmetric RNB ($N_2 > N_1$) Embodiments:

[0291] Network 1001K2 of FIG. 1K2 is an example of general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_2) - 1$ stages where $N_1 > N_2$ and $N_1 = p \times N_2$ where $p > 1$. In network 100K2 of FIG. 1K2, $N_2 = N$ and $N_1 = p \times N$. The general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention. (And in the example of FIG. 1K2, $s = 2$). The general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_2) - 1$ stages has d_1 (where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d$$

inlet links for each of

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) (for example the links IL1-IL($p \times d$) to the input switch IS1) and $d + d_1$ ($= d + p \times d$) outgoing links for each of

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) (for example the links ML(1,1)-ML(1, ($d + p \times d$)) to the input switch IS1). There are d outlet links for each of

US 2011/0044329 A1

Feb. 24, 2011

26

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example the links OL1-OL(d) to the output switch OS1) and $2 \times d$ incoming links for each of

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example ML($2 \times \text{Log}_d N_2, 1$)-ML($2 \times \text{Log}_d N_2 - 2, 2 \times d$) to the output switch OS1).
[0292] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) are connected to exactly

$$\frac{d + d_1}{2}$$

switches in middle stage **130** through $d + d_1$ links.
[0293] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1, N_2/d) in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links.
[0294] Similarly each of the

$$\frac{N_2}{d}$$

middle switches

$$MS(\text{Log}_d N_2 - 1, 1) - MS\left(\text{Log}_d N_2 - 1, \frac{N_2}{d}\right)$$

in the middle stage $\mathbf{130} + 10^*(\text{Log}_d N_2 - 2)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10^*(\text{Log}_d N_2 - 3)$ through $2 \times d$ links and also are connected to exactly d switches in middle stage $\mathbf{130} + 10^*(\text{Log}_d N_2 - 1)$ through $2 \times d$ links.

[0295] Similarly each of the

$$\frac{N_2}{d}$$

middle switches

$$MS(2 \times \text{Log}_d N_2 - 3, 1) - MS\left(2 \times \text{Log}_d N_2 - 3, \frac{N_2}{d}\right)$$

in the middle stage $\mathbf{130} + 10^*(2 * \text{Log}_d N_2 - 4)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10^*(2 * \text{Log}_d N_2 - 5)$ through $2 \times d$ links and also are connected to exactly d output switches in output stage **120** through $2 \times d$ links.

[0296] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) are connected from exactly d switches in middle stage $\mathbf{130} + 10^*(2 * \text{Log}_d N_2 - 4)$ through $2 \times d$ links.

[0297] As described before, again the connection topology of a general $V_{mlink}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{mlink}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1A2 to FIG. 1J2 are ten examples of network $V_{mlink}(N_1, N_2, d, s)$ for $s=2$ and $N_2 > N_1$.

[0298] The general symmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0299] For example, the network of FIG. 1C2 shows an exemplary five-stage network, namely $V_{mlink}(8, 24, 2, 2)$, with the following multicast assignment $I_1 = \{1, 4\}$ and all other $I_j = \emptyset$ for $j = [2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,4) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,4) only once into middle switches MS(2,1) and MS(2,4) respectively in middle stage **140**.

[0300] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,4) only once into middle switches MS(3,1) and MS(3,4) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,4) only once into output switches OS1 and OS4 in output stage **120**. Finally the connection I_1 fans out once in the output stage switch OS1 into outlet link OL1 and in the output stage switch OS4 twice into the outlet links OL7 and OL8. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage **130**.

Symmetric Folded RNB Embodiments:

[0301] The folded multi-link multi-stage network $V_{fold,mlink}(N_1, N_2, d, s)$ disclosed, in the current invention, is topo-

US 2011/0044329 A1

Feb. 24, 2011

27

logically exactly the same as the multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$, disclosed in the current invention so far, excepting that in the illustrations folded network $V_{fold-mlink}(N_1, N_2, d, s)$ is shown as it is folded at middle stage $130+10*(\text{Log}_2 N_2-2)$. This is true for all the embodiments presented in the current invention.

[0302] Referring to FIG. 2A, in one embodiment, an exemplary symmetrical folded multi-link multi-stage network 200A with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, four by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, four by four switches MS(3,1)-MS(3,4).

[0303] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0304] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch IS1-IS4 can be denoted in general with the notation $d*2d$ and each output switch OS1-OS4 can be denoted in general with the notation $2d*d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d*2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric folded multi-link multi-stage network can be represented with the notation $V_{fold-mlink}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0305] Each of the N/d input switches IS1-IS4 are connected to exactly $2*d$ switches in middle stage 130 through $2*d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and also to middle switch MS(1,2) through the links ML(1,3) and ML(1,4)).

[0306] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through $2*d$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1)

from input switch IS1, and the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to exactly d switches in middle stage 140 through $2*d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0307] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $2*d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,11) and ML(2,12) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through $2*d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,3)).

[0308] Similarly each of the N/d middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $2*d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,11) and ML(3,12) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly d output switches in output stage 120 through $2*d$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from Middle switch MS(3,1), and the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1)).

[0309] Each of the N/d output switches OS1-OS4 are connected from exactly $2*d$ switches in middle stage 150 through $2*d$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2), and output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,7) and ML(4,8)).

[0310] Finally the connection topology of the network 200A shown in FIG. 2A is known to be back to back inverse Benes connection topology.

[0311] In other embodiments the connection topology may be different from the network 200A of FIG. 2A. That is the way the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16), and ML(4,1)-ML(4,16) are connected between the respective stages is different. Even though only one embodiment is illustrated, in general, the network $V_{fold-mlink}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold-mlink}(N, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{fold-mlink}(N, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{fold-mlink}(N, d, s)$ can be built. The embodiment of FIG. 2A is only one example of network $V_{fold-mlink}(N, d, s)$.

[0312] In the embodiment of FIG. 2A each of the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16) and ML(4,1)-ML(4,16) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage 110 is often referred to as the first stage. The output switches OS1-OS4 are also

US 2011/0044329 A1

Feb. 24, 2011

28

referred to as the network output ports. The output stage **120** is often referred to as the last stage. The middle stage switches **MS(1,1)**-**MS(1,4)** and **MS(2,1)**-**MS(2,4)** are referred to as middle switches or middle ports. The middle stage **130** is also referred to as root stage and middle stage switches **MS(2,1)**-**MS(2,4)** are referred to as root stage switches.

[0313] In the example illustrated in FIG. 2A, a fan-out of four is possible to satisfy a multicast connection request if input switch is **IS2**, but only two switches in middle stage **130** will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is **IS1**, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage **130** when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network **200A**, to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0314] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage **130** is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage **130**, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the rearrangeably nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Symmetric Folded RNB Embodiments:

[0315] Network **200B** of FIG. 2B is an example of general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages. The general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention. (And in the example of FIG. 2B, $s=2$). The general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of N/d input switches **IS1**-**IS(N/d)** (for example the links **IL1**-**IL(d)** to the input switch **IS1**) and $2 \times d$ outgoing links for each of N/d input switches **IS1**-**IS(N/d)** (for example the links **ML(1,1)**-**ML(1,2d)** to the input switch **IS1**). There are d outlet links for each of N/d output switches **OS1**-**OS(N/d)** (for example the links **OL1**-**OL(d)** to the output switch **OS1**) and $2 \times d$ incoming links for each of N/d output switches **OS1**-**OS(N/d)** (for example **ML(2 × Log_d N - 2, 1)**-**ML(2 × Log_d N - 2, 2 × d)** to the output switch **OS1**).

[0316] Each of the N/d input switches **IS1**-**IS(N/d)** are connected to exactly d switches in middle stage **130** through $2 \times d$ links.

[0317] Each of the N/d middle switches **MS(1,1)**-**MS(1,N/d)** in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links.

[0318] Similarly each of the N/d middle switches

$$MS(\log_d N - 1, 1) - MS\left(\log_d N - 1, \frac{N}{d}\right)$$

in the middle stage $130 + 10 * (\log_d N - 2)$ are connected from exactly d switches in middle stage $130 + 10 * (\log_d N - 3)$ through $2 \times d$ links and also are connected to exactly d switches in middle stage $130 + 10 * (\log_d N - 1)$ through $2 \times d$ links.

[0319] Similarly each of the N/d middle switches

$$MS(2 \times \log_d N - 3, 1) - MS\left(2 \times \log_d N - 3, \frac{N}{d}\right)$$

in the middle stage $130 + 10 * (2 * \log_d N - 4)$ are connected from exactly d switches in middle stage $130 + 10 * (2 * \log_d N - 5)$ through $2 \times d$ links and also are connected to exactly d output switches in output stage **120** through $2 \times d$ links.

[0320] Each of the N/d output switches **OS1**-**OS(N/d)** are connected from exactly d switches in middle stage $130 + 10 * (2 * \log_d N - 4)$ through $2 \times d$ links.

[0321] As described before, again the connection topology of a general $V_{fold-mlink}(N, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{fold-mlink}(N, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{fold-mlink}(N, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{fold-mlink}(N, d, s)$ can be built. The embodiment of FIG. 1A is one example of network $V_{fold-mlink}(N, d, s)$.

[0322] The general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0323] Every switch in the folded multi-link multi-stage networks discussed herein has multicast capability. In a $V_{fold-mlink}(N, d, s)$ network, if a network inlet link is to be connected to more than one outlet link on the same output switch, then it is only necessary for the corresponding input switch to have one path to that output switch. This follows because that path can be multicast within the output switch to as many outlet links as necessary. Multicast assignments can therefore be described in terms of connections between input switches and output switches. An existing connection or a new connection from an input switch to r' output switches is said to have fan-out r' . If all multicast assignments of a first type, wherein any inlet link of an input switch is to be connected in an output switch to at most one outlet link are realizable, then multicast assignments of a second type, wherein any inlet link of each input switch is to be connected to more than one outlet link in the same output switch, can also be realized. For this reason, the following discussion is limited to general multicast connections of the first type (with fan-out r' ,

US 2011/0044329 A1

Feb. 24, 2011

29

$$1 \leq r' \leq \frac{N}{d}$$

although the same discussion is applicable to the second type. [0324] To characterize a multicast assignment, for each inlet link

$$i \in \left\{ 1, 2, \dots, \frac{N}{d} \right\},$$

let $I_j = O$, where

$$O \subset \left\{ 1, 2, \dots, \frac{N}{d} \right\}.$$

denote the subset of output switches to which inlet link i is to be connected in the multicast assignment. For example, the network of FIG. 1C shows an exemplary five-stage network, namely $V_{m-link}(8, 2, 2)$, with the following multicast assignment $I_1 = \{2, 4\}$ and all other $I_j = \emptyset$ for $j = [2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,2) in middle stage 130, and fans out in middle switches MS(1,1) and MS(1,2) only once into middle switches MS(2,1) and MS(2,3) respectively in middle stage 140.

[0325] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,3) only once into middle switches MS(3,2) and MS(3,4) respectively in middle stage 150. The connection I_1 also fans out in middle switches MS(3,2) and MS(3,4) only once into output switches OS2 and OS4 in output stage 120. Finally the connection I_1 fans out once in the output stage switch OS2 into outlet link OL3 and in the output stage switch OS4 twice into the outlet links OL7 and OL8. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage 130.

Asymmetric Folded RNB ($N_2 > N_1$) Embodiments:

[0326] Referring to FIG. 2C, in one embodiment, an exemplary asymmetrical folded multi-link multi-stage network 200C with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, four by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, four by eight switches MS(3,1)-MS(3,4).

[0327] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in

rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size eight by six, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0328] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_1}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d * 2d$ and each output switch OS1-OS4 can be denoted in general with the notation $(d+d_2) * d_2$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $2d * 2d$. The size of each switch in the last middle stage can be denoted as $2d * (d+d_2)$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric folded multi-link multi-stage network can be represented with the notation $V_{fold-m-link}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0329] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly d switches in middle stage 130 through $2 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and also to middle switch MS(1,2) through the links ML(1,3) and ML(1,4)).

[0330] Each of the

$$\frac{N_1}{d}$$

US 2011/0044329 A1

Feb. 24, 2011

30

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0331] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,11) and ML(2,12) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage **150** through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,3)).

[0332] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,11) and ML(3,12) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly

$$\frac{d + d_2}{2}$$

output switches in output stage **120** through $d + d_1$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from Middle switch MS(3,1); the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1); the links ML(4,5) and ML(4,6) are connected to output switch OS3 from Middle switch MS(3,1); and the links ML(4,7) and ML(4,8) are connected to output switch OS4 from middle switch MS(3,1)).

[0333] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly

$$\frac{d + d_2}{2}$$

switches in middle stage **150** through $d + d_1$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,9) and ML(4,10); output switch OS1 is connected from middle switch MS(3,3) through the links ML(4,17) and ML(4,18); and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,25) and ML(4,26)).

[0334] Finally the connection topology of the network **200C** shown in FIG. 2C is known to be back to back inverse Benes connection topology.

[0335] In other embodiments the connection topology may be different from the network **200C** of FIG. 2C. That is the way the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16), and ML(4,1)-ML(4,16) are connected between the respective stages is different. Even though only one embodiment is illustrated, in general, the network $V_{fold-mlink}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold-mlink}(N, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{fold-mlink}(N, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{fold-mlink}(N, d, s)$ can be built. The embodiment of FIG. 2C is only one example of network $V_{fold-mlink}(N, d, s)$.

[0336] In the embodiment of FIG. 2C each of the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16) and ML(4,1)-ML(4,16) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage **110** is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage **120** is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,4) and MS(2,1)-MS(2,4) are referred to as middle switches or middle ports. The middle stage **130** is also referred to as root stage and middle stage switches MS(2,1)-MS(2,4) are referred to as root stage switches.

[0337] In the example illustrated in FIG. 2C, a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage **130** will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage **130** when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network **200C**, to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0338] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage

US 2011/0044329 A1

Feb. 24, 2011

31

130 is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage **130**, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the rearrangeably nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Asymmetric Folded RNB ($N_2 > N_1$) Embodiments:

[0339] Network **200D** of FIG. 2D is an example of general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_1) - 1$ stages where $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. In network **200D** of FIG. 2D, $N_1 = N$ and $N_2 = p * N$. The general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention. (And in the example of FIG. 2D, $s = 2$). The general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_1) - 1$ stages has d inlet links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links IL1-IL(d) to the input switch IS1) and $2 \times d$ outgoing links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links ML(1,1)-ML(1,2d) to the input switch IS1). There are d_2 (where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d$$

outlet links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example the links OL1-OL($p \times d$) to the output switch OS1) and $d + d_2$ ($= d + p \times d$) incoming links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example ML($2 \times \log_d N_1 - 2, 1$)-ML($2 \times \log_d N_1 - 2, d + d_2$) to the output switch OS1).

[0340] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) are connected to exactly $2 \times d$ switches in middle stage **130** through $2 \times d$ links.

[0341] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1, N_1/d) in the middle stage **130** are connected to exactly d input switches through $2 \times d$ links and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links.

[0342] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(\log_d N_1 - 1, 1) - MS(\log_d N_1 - 1, \frac{N_1}{d})$$

in the middle stage $\mathbf{130} + 10 * (\log_d N_1 - 2)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10 * (\log_d N_1 - 3)$ through $2 \times d$ links and also are connected to exactly d switches in middle stage $\mathbf{130} + 10 * (\log_d N_1 - 1)$ through $2 \times d$ links.

[0343] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(2 \times \log_d N_1 - 3, 1) - MS(2 \times \log_d N_1 - 3, \frac{N_1}{d})$$

in the middle stage $\mathbf{130} + 10 * (2 * \log_d N_1 - 4)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10 * (2 * \log_d N_1 - 5)$ through $2 \times d$ links and also are connected to exactly

$$\frac{d + d_2}{2}$$

output switches in output stage **120** through $d + d_2$ links.

US 2011/0044329 A1

Feb. 24, 2011

32

[0344] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) are connected from exactly

$$\frac{d+d_2}{2}$$

switches in middle stage **130**+ $10*(2*\text{Log}_d N_1-4)$ through $d+d_2$ links.

[0345] As described before, again the connection topology of a general $V_{fold-mlink}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{fold-mlink}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{fold-mlink}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{fold-mlink}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 1C is one example of network $V_{fold-mlink}(N_1, N_2, d, s)$ for $s=2$ and $N_2 > N_1$.

[0346] The general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0347] For example, the network of FIG. 2C shows an exemplary five-stage network, namely $V_{fold-mlink}(8, 24, 2, 2)$, with the following multicast assignment $I_1 = \{1, 4\}$ and all other $I_j = \emptyset$ for $j = [2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,2) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,2) only once into middle switches MS(2,1) and MS(2,3) respectively in middle stage **140**.

[0348] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,3) only once into middle switches MS(3,1) and MS(3,4) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,4) only once into output switches OS1 and OS4 in output stage **120**. Finally the connection I_1 fans out once in the output stage switch OS1 into outlet link OL2 and in the output stage switch OS4 twice into the outlet links OL20 and OL23. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage **130**.

Asymmetric Folded RNB ($N_1 > N_2$) Embodiments:

[0349] Referring to FIG. 2E, in one embodiment, an exemplary asymmetrical folded multi-link multi-stage network **200E** with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle

stages **130**, **140**, and **150** is shown where input stage **110** consists of four, six by eight switches IS1-IS4 and output stage **120** consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, eight by four switches MS(1,1)-MS(1,4), middle stage **140** consists of four, four by four switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, four by four switches MS(3,1)-MS(3,4).

[0350] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0351] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or N_2 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_2}{d},$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d_1 * (d+d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $(2*d*d)$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $2d*2d$. The size of each switch in the first middle stage can be denoted as $(d+d_1)*2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric folded multi-link multi-stage network can be represented with the notation $V_{fold-mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

US 2011/0044329 A1

Feb. 24, 2011

33

[0352] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly

$$\frac{d+d_1}{2}$$

switches in middle stage **130** through $d+d_1$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2); input switch IS1 is connected to middle switch MS(1,2) through the links ML(1,3) and ML(1,4); input switch IS1 is connected to middle switch MS(1,3) through the links ML(1,5), ML(1,6); and input switch IS1 is also connected to middle switch MS(1,4) through the links ML(1,7) and ML(1,8)).

[0353] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly

$$\frac{d+d_1}{2}$$

input switches through $d+d_1$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1; the links ML(1,9) and ML(1,10) are connected to the middle switch MS(1,1) from input switch IS2; the links ML(1,17) and ML(1,18) are connected to the middle switch MS(1,1) from input switch IS3; and the links ML(1,25) and ML(1,26) are connected to the middle switch MS(1,1) from input switch IS4) and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0354] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,11) and ML(2,12) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage **150** through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch

MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,3)).

[0355] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,11) and ML(3,12) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly d output switches in output stage **120** through $2 \times d$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from middle switch MS(3,1); and the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1)).

[0356] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage **150** through $2 \times d$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); and output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,7) and ML(4,8)).

[0357] Finally the connection topology of the network **200E** shown in FIG. 2E is known to be back to back inverse Benes connection topology.

[0358] In other embodiments the connection topology may be different from the network **200E** of FIG. 2E. That is the way the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16), and ML(4,1)-ML(4,16) are connected between the respective stages is different. Even though only one embodiment is illustrated, in general, the network $V_{fold-mlink}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold-mlink}(N, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{fold-mlink}(N, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{fold-mlink}(N, d, s)$ can be built. The embodiment of FIG. 2E is only one example of network $V_{fold-mlink}(N, d, s)$.

[0359] In the embodiment of FIG. 2E each of the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16) and ML(4,1)-ML(4,16) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage **110** is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage **120** is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,4) and MS(2,1)-MS(2,4) are referred to as middle switches or middle ports. The middle stage **130** is also

US 2011/0044329 A1

Feb. 24, 2011

34

referred to as root stage and middle stage switches MS(2,1)-MS(2,4) are referred to as root stage switches.

[0360] In the example illustrated in FIG. 2E, a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage **130** will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage **130** when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network **200E**, to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0361] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage **130**, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the rearrangeably nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Asymmetric Folded RNB ($N_2 > N_1$) Embodiments:

[0362] Network **200F** of FIG. 2F is an example of general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_2) - 1$ stages where $N_1 > N_2$ and $N_1 = p \times N_2$ where $p > 1$. In network **200F** of FIG. 2F, $N_2 = N$ and $N_1 = p \times N$. The general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention. (And in the example of FIG. 2F, $s = 2$). The general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_2) - 1$ stages has d , (where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d$$

inlet links for each of

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) (for example the links IL1-IL($p \times d$) to the input switch IS1) and $d+d$, ($=d+p \times d$) outgoing links for each of

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) (for example the links ML(1,1)-ML(1,($d+p \times d$)) to the input switch IS1). There are d outlet links for each of

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example the links OL1-OL(d) to the output switch OS1) and $2 \times d$ incoming links for each of

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example ML($2 \times \log_d N_2 - 2, 1$)-ML($2 \times \log_d N_2 - 2, 2 \times d$) to the output switch OS1).

[0363] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) are connected to exactly

$$\frac{d + d_1}{2}$$

switches in middle stage **130** through $d+d_1$ links.

[0364] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1, N_2/d) in the middle stage **130** are connected from exactly d input switches through $2 \times d$ links and also are connected to exactly d switches in middle stage **140** through $2 \times d$ links.

[0365] Similarly each of the

$$\frac{N_2}{d}$$

middle switches

$$MS(\log_d N_2 - 1, 1) - MS(\log_d N_2 - 1, \frac{N_2}{d})$$

in the middle stage **130** + $10 \times (\log_d N_2 - 2)$ are connected from exactly d switches in middle stage **130** + $10 \times (\log_d N_2 - 3)$

US 2011/0044329 A1

Feb. 24, 2011

35

through $2 \times d$ links and also are connected to exactly d switches in middle stage $130 + 10 * (\text{Log}_d N_2 - 1)$ through $2 \times d$ links.

[0366] Similarly each of the

$$\frac{N_2}{d}$$

middle switches

$$MS(2 \times \text{Log}_d N_2 - 3, 1) - MS\left(2 \times \text{Log}_d N_2 - 3, \frac{N_2}{d}\right)$$

in the middle stage $130 + 10 * (2 * \text{Log}_d N_2 - 4)$ are connected from exactly d switches in middle stage $130 + 10 * (2 * \text{Log}_d N_2 - 5)$ through $2 \times d$ links and also are connected to exactly d output switches in output stage **120** through $2 \times d$ links.

[0367] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) are connected from exactly d switches in middle stage $130 + 10 * (2 * \text{Log}_d N_2 - 4)$ through $2 \times d$ links.

[0368] As described before, again the connection topology of a general $V_{\text{fold-mlink}}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{\text{fold-mlink}}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{\text{fold-mlink}}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{\text{fold-mlink}}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 2F is one example of network $V_{\text{fold-mlink}}(N_1, N_2, d, s)$ for $s=2$ and $N_2 > N_1$.

[0369] The general symmetrical folded multi-link multi-stage network $V_{\text{fold-mlink}}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical folded multi-link multi-stage network $V_{\text{fold-mlink}}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0370] For example, the network of FIG. 2E shows an exemplary five-stage network, namely $V_{\text{fold-mlink}}(8, 24, 2, 2)$, with the following multicast assignment $I_1 = \{1, 4\}$ and all other $I_j = \emptyset$ for $j = [2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,4) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,4) only once into middle switches MS(2,1) and MS(2,4) respectively in middle stage **140**.

[0371] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,4) only once into middle switches MS(3,1) and MS(3,4) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,4) only once into output switches OS1 and OS4 in

output stage **120**. Finally the connection I_1 fans out once in the output stage switch OS1 into outlet link OL1 and in the output stage switch OS4 twice into the outlet links OL7 and OL8. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage **130**.

SNB Multi-Link Multi-Stage Embodiments:

Symmetric SNB Embodiments:

[0372] Referring to FIG. 3A, in one embodiment, an exemplary symmetrical multi-link multi-stage network **300A** with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by six switches IS1-IS4 and output stage **120** consists of four, six by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, six by six switches MS(1,1)-MS(1,4), middle stage **140** consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, six by six switches MS(3,1)-MS(3,4).

[0373] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by six, the switches in output stage **120** are of size six by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0374] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch IS1-IS4 can be denoted in general with the notation $d * 3d$ and each output switch OS1-OS4 can be denoted in general with the notation $3d * d$. Likewise, the size of each switch in any of the middle stages can be denoted as $3d * 3d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-link multi-stage network can be represented with the notation $V_{\text{mlink}}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0375] Each of the N/d input switches IS1-IS4 are connected to exactly $3 \times d$ switches in middle stage **130** through $3 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and ML(1,3); and also to middle switch MS(1,2) through the links ML(1,4), ML(1,5), and ML(1,6)).

[0376] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly d input switches through $3 \times d$ links (for example the links ML(1,1), ML(1,2), and ML(1,3) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,10), ML(1,11), and ML(1,12) are connected to the middle switch

US 2011/0044329 A1

Feb. 24, 2011

36

MS(1,1) from input switch IS2) and also are connected to exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,4), ML(2,5), and ML(2,6) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0377] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,16), ML(2,17), and ML(2,18) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,4), ML(3,5), and ML(3,6) are connected from middle switch MS(2,1) to middle switch MS(3,3)).

[0378] Similarly each of the N/d middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,16), ML(3,17), and ML(3,18) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly d output switches in output stage 120 through $3 \times d$ links (for example the links ML(4,1), ML(4,2), and ML(4,3) are connected to output switch OS1 from Middle switch MS(3,1), and the links ML(4,10), ML(4,11), and ML(4,12) are connected to output switch OS2 from middle switch MS(3,1)).

[0379] Each of the N/d output switches OS1-OS4 are connected from exactly $3 \times d$ switches in middle stage 150 through $3 \times d$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1), ML(4,2), and ML(4,3), and output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,10), ML(4,11) and ML(4,12)).

[0380] Finally the connection topology of the network 300A shown in FIG. 3A is known to be back to back inverse Benes connection topology.

[0381] Referring to FIG. 3B, in another embodiment of network $V_{mink}(N, d, s)$, an exemplary symmetrical multi-link multi-stage network 300B with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by six switches IS1-IS4 and output stage 120 consists of four, six by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, six by six switches MS(1,1)-MS(1,4), middle stage 140 consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, six by six switches MS(3,1)-MS(3,4).

[0382] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by six, the switches in output stage 120 are of size six by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0383] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch IS1-IS4 can be denoted in general with the notation $d^3 \times d$ and each output switch OS1-OS4 can be denoted in general with the notation $3d \times d$. Likewise, the size of each switch in any of the middle stages can be denoted as $3d \times 3d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The symmetric multi-link multi-stage network of FIG. 3B is also the network of the type $V_{mink}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0384] Each of the N/d input switches IS1-IS4 are connected to exactly $3 \times d$ switches in middle stage 130 through $3 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and ML(1,3); and also to middle switch MS(1,2) through the links ML(1,4), ML(1,5), and ML(1,6)).

[0385] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through $3 \times d$ links (for example the links ML(1,1), ML(1,2), and ML(1,3) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,13), ML(1,14), and ML(1,15) are connected to the middle switch MS(1,1) from input switch IS3) and also are connected to exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected from middle switch MS(1,1) to middle switch MS(2,1); and the links ML(2,4), ML(2,5), and ML(2,6) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0386] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,13), ML(2,14), and ML(2,15) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,4), ML(3,5) and ML(3,6) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0387] Similarly each of the N/d middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,13), ML(3,14), and ML(3,15) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly d output switches in output stage 120 through $3 \times d$ links (for example

US 2011/0044329 A1

Feb. 24, 2011

37

the links ML(4,1), ML(4,2), and ML(4,3) are connected to output switch OS1 from Middle switch MS(3,1), and the links ML(4,4), ML(4,5), and ML(4,6) are connected to output switch OS2 from middle switch MS(3,1)).

[0388] Each of the N/d output switches OS1-OS4 are connected from exactly 3xd switches in middle stage 150 through 3xd links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1), ML(4,2), and ML(4,3), and output switch OS1 is also connected from middle switch MS(3,3) through the links ML(4,13), ML(4,14), and ML(4,15)).

[0389] Finally the connection topology of the network 300B shown in FIG. 3B is known to be back to back Omega connection topology.

[0390] Referring to FIG. 3C, in another embodiment of network $V_{mlink}(N, d, s)$, an exemplary symmetrical multi-link multi-stage network 300C with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by six switches IS1-IS4 and output stage 120 consists of four, six by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, six by six switches MS(1,1)-MS(1,4), middle stage 140 consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, six by six switches MS(3,1)-MS(3,4).

[0391] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by six, the switches in output stage 120 are of size six by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0392] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d, where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d. The size of each input switch IS1-IS4 can be denoted in general with the notation $d*3d$ and each output switch OS1-OS4 can be denoted in general with the notation $3d*d$. Likewise, the size of each switch in any of the middle stages can be denoted as $3d*3d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The symmetric multi-link multi-stage network of FIG. 3C is also the network of the type $V_{mlink}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0393] Each of the N/d input switches IS1-IS4 are connected to exactly 3xd switches in middle stage 130 through 3xd links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and ML(1,3); and also to middle switch MS(1,2) through the links ML(1,4), ML(1,5), and ML(1,6)).

[0394] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through 3xd links (for example the links ML(1,1), ML(1,2), and ML(1,3) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,22), ML(1,23), and ML(1,24) are connected to the middle switch MS(1,1) from input switch IS4) and also are connected to exactly d switches in middle stage 140 through 3xd links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,4), ML(2,5), and ML(2,6) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0395] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through 3xd links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,22), ML(2,23), and ML(2,24) are connected to the middle switch MS(2,1) from middle switch MS(1,4)) and also are connected to exactly d switches in middle stage 150 through 3xd links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,4), ML(3,5), and ML(3,6) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0396] Similarly each of the N/d middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through 3xd links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,22), ML(3,23), and ML(3,24) are connected to the middle switch MS(3,1) from middle switch MS(2,4)) and also are connected to exactly d output switches in output stage 120 through 3xd links (for example the links ML(4,1), ML(4,2), and ML(4,3) are connected to output switch OS1 from middle switch MS(3,1), and the links ML(4,4), ML(4,5), and ML(4,6) are connected to output switch OS2 from middle switch MS(3,1)).

[0397] Each of the N/d output switches OS1-OS4 are connected from exactly 3xd switches in middle stage 150 through 3xd links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1), ML(4,2), and ML(4,3), and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,22), ML(4,23), and ML(4,24)).

[0398] Finally the connection topology of the network 300C shown in FIG. 3C is hereinafter called nearest neighbor connection topology.

[0399] Similar to network 300A of FIG. 3A, 300B of FIG. 3B, and 300C of FIG. 3C, referring to FIG. 3D, FIG. 3E, FIG. 3F, FIG. 3G, FIG. 3H, FIG. 3I and FIG. 3J with exemplary symmetrical multi-link multi-stage networks 300D, 300E, 300F, 300G, 300H, 300I, and 300J respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by six switches IS1-IS4 and output stage 120 consists of four, six by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, six by six switches MS(1,1)-MS(1,4), middle stage 140 consists of four, six by

US 2011/0044329 A1

Feb. 24, 2011

38

six switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, six by six switches MS(3,1)-MS(3,4).

[0400] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by six, the switches in output stage 120 are of size six by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0401] The networks 300D, 300E, 300F, 300G, 300H, 300I and 300J of FIG. 3D, FIG. 3E, FIG. 3F, FIG. 3G, FIG. 3H, FIG. 3I, and FIG. 3J are also embodiments of symmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0402] Just like networks of 300A, 300B and 300C, for all the networks 300D, 300E, 300F, 300G, 300H, 300I and 300J of FIG. 3D, FIG. 3E, FIG. 3F, FIG. 3G, FIG. 3H, FIG. 3I, and FIG. 3J, each of the N/d input switches IS1-IS4 are connected to exactly 3xd switches in middle stage 130 through 3xd links.

[0403] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through 3xd links and also are connected to exactly d switches in middle stage 140 through 3xd links.

[0404] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through 3xd links and also are connected to exactly d switches in middle stage 150 through 3xd links.

[0405] Similarly each of the N/d middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through 3xd links and also are connected to exactly d output switches in output stage 120 through 3xd links.

[0406] Each of the N/d output switches OS1-OS4 are connected from exactly 3xd switches in middle stage 150 through 3xd links.

[0407] In all the ten embodiments of FIG. 3A to FIG. 3J the connection topology is different. That is the way the links ML(1,1)-ML(1,24), ML(2,1)-ML(2,24), ML(3,1)-ML(3,24), and ML(4,1)-ML(4,24) are connected between the respective stages is different. Even though only ten embodiments are illustrated, in general, the network $V_{mlink}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mlink}(N, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{mlink}(N, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N, d, s)$ can be built. The ten embodiments of FIG. 3A to FIG. 3J are only three examples of network $V_{mlink}(N, d, s)$.

[0408] In all the ten embodiments of FIG. 3A to FIG. 3J, each of the links ML(1,1)-ML(1,24), ML(2,1)-ML(2,24), ML(3,1)-ML(3,24) and ML(4,1)-ML(4,24) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4

are also referred to as the network input ports. The input stage 110 is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage 120 is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,4), MS(2,1)-MS(2,4), and MS(3,1)-MS(3,4) are referred to as middle switches or middle ports.

[0409] In the example illustrated in FIG. 3A (or in FIG. 1B to FIG. 3J), a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage 130 will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage 130 when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network 300A (or 300B to 300J), to be operated in strictly nonblocking manner in accordance with the invention.

[0410] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage 130 is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage 130, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the strictly nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Symmetric SNB Embodiments:

[0411] Network 300K of FIG. 3K is an example of general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages. The general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ can be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention (and in the example of FIG. 3K, $s=3$). The general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of N/d input switches IS1-IS(N/d) (for example the links IL1-IL(d) to the input switch IS1) and 3xd outgoing links for each of N/d input switches IS1-IS(N/d) (for example the links ML(1,1)-ML(1,3d) to the input switch IS1). There are d outlet links for each of N/d output switches OS1-OS(N/d) (for example the links OL1-OL(d) to the output switch OS1) and 3xd incoming links for each of N/d output switches OS1-OS(N/d) (for example ML(2xLog_dN-2,1)-ML(2xLog_dN-2,3xd) to the output switch OS1).

[0412] Each of the N/d input switches IS1-IS(N/d) are connected to exactly d switches in middle stage 130 through 3xd links.

[0413] Each of the N/d middle switches MS(1,1)-MS(1,N/d) in the middle stage 130 are connected from exactly d input switches through 3xd links and also are connected to exactly d switches in middle stage 140 through 3xd links.

US 2011/0044329 A1

Feb. 24, 2011

39

[0414] Similarly each of the N/d middle switches

$$MS(\text{Log}_d N - 1, 1) - MS\left(\text{Log}_d N - 1, \frac{N}{d}\right)$$

in the middle stage $130+10*(\text{Log}_d N-2)$ are connected from exactly d switches in middle stage $130+10*(\text{Log}_d N-3)$ through $3 \times d$ links and also are connected to exactly d switches in middle stage $130+10*(\text{Log}_d N-1)$ through $3 \times d$ links.

[0415] Similarly each of the N/d middle switches

$$MS(2 \times \text{Log}_d N - 3, 1) - MS\left(2 \times \text{Log}_d N - 3, \frac{N}{d}\right)$$

in the middle stage $130+10*(2*\text{Log}_d N-4)$ are connected from exactly d switches in middle stage $130+10*(2*\text{Log}_d N-5)$ through $3 \times d$ links and also are connected to exactly d output switches in output stage **120** through $3 \times d$ links.

[0416] Each of the N/d output switches OS1-OS(N/d) are connected from exactly d switches in middle stage $130+10*(2*\text{Log}_d N-4)$ through $3 \times d$ links.

[0417] As described before, again the connection topology of a general $V_{mlink}(N, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink}(N, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{mlink}(N, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N, d, s)$ can be built. The embodiments of FIG. 3A to FIG. 3J are ten examples of network $V_{mlink}(N, d, s)$.

[0418] The general symmetrical multi-link multi-stage network $V_{mlink}(N, d, s)$ can be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention.

[0419] Every switch in the multi-link multi-stage networks discussed herein has multicast capability. In a $V_{mlink}(N, d, s)$ network, if a network inlet link is to be connected to more than one outlet link on the same output switch, then it is only necessary for the corresponding input switch to have one path to that output switch. This follows because that path can be multicast within the output switch to as many outlet links as necessary. Multicast assignments can therefore be described in terms of connections between input switches and output switches. An existing connection or a new connection from an input switch to r' output switches is said to have fan-out r' . If all multicast assignments of a first type, wherein any inlet link of an input switch is to be connected in an output switch to at most one outlet link are realizable, then multicast assignments of a second type, wherein any inlet link of each input switch is to be connected to more than one outlet link in the same output switch, can also be realized. For this reason, the following discussion is limited to general multicast connections of the first type (with fan-out r' ,

$$1 \leq r' \leq \frac{N}{d}$$

although the same discussion is applicable to the second type.

[0420] To characterize a multicast assignment, for each inlet link

$$i \in \left\{1, 2, \dots, \frac{N}{d}\right\},$$

let $I_i = O$, where

$$O \subset \left\{1, 2, \dots, \frac{N}{d}\right\},$$

denote the subset of output switches to which inlet link i is to be connected in the multicast assignment. For example, the network of FIG. 3C shows an exemplary five-stage network, namely $V_{mlink}(8, 2, 3)$, with the following multicast assignment $I_1 = \{1, 4\}$ and all other $I_j = \emptyset$ for $j = [2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,2) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,2) only once into middle switches MS(2,1) and MS(2,3) respectively in middle stage **140**.

[0421] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,3) only once into middle switches MS(3,1) and MS(3,4) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,4) only once into output switches OS1 and OS4 in output stage **120**. Finally the connection I_1 fans out once in the output stage switch OS1 into outlet link OL1 and in the output stage switch OS4 twice into the outlet links OL7 and OL8. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage **130**.

Asymmetric SNB ($N_2 > N_1$) Embodiments:

[0422] Referring to FIG. 3A1, in one embodiment, an exemplary asymmetrical multi-link multi-stage network **300A1** with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by six switches IS1-IS4 and output stage **120** consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, six by six switches MS(1,1)-MS(1,4), middle stage **140** consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, four by eight switches MS(3,1)-MS(3,4).

[0423] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by six, the switches in output stage **120** are of size eight by six, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

US 2011/0044329 A1

Feb. 24, 2011

40

[0424] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p \times N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_1}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d \times 3d$ and each output switch OS1-OS4 can be denoted in general with the notation $(2d + d_2) \times d_2$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $3d \times 3d$. The size of each switch in the last middle stage can be denoted as $3d \times (2d + d_2)$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{mink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0425] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly d switches in middle stage 130 through $3 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and ML(1,3); and also to middle switch MS(1,2) through the links ML(1,4), ML(1,5), and ML(1,6)).

[0426] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through $3 \times d$ links (for example the links ML(1,1), ML(1,2), and ML(1,3) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,13), ML(1,14), and ML(1,15) are

connected to the middle switch MS(1,1) from input switch IS3) and also are connected to exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected from middle switch MS(1,1) to middle switch MS(2,1); and the links ML(2,4), ML(2,5), and ML(2,6) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0427] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,13), ML(2,14), and ML(2,15) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,4), ML(3,5) and ML(3,6) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0428] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,13), ML(3,14), and ML(3,15) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly

$$\frac{2d + d_2}{3}$$

output switches in output stage 120 through $2d + d_1$ links

[0429] (For example the links ML(4,1), ML(4,2), and ML(4,3) are connected to output switch OS1 from Middle switch MS(3,1); the links ML(4,4), ML(4,5), and ML(4,6) are connected to output switch OS2 from middle switch MS(3,1); the links ML(4,7), ML(4,8), and ML(4,9) are connected to output switch OS3 from Middle switch MS(3,1); the links ML(4,10), ML(4,11), and ML(4,12) are connected to output switch OS2 from middle switch MS(3,1)).

[0430] Each of the

$$\frac{N_1}{d}$$

US 2011/0044329 A1

Feb. 24, 2011

41

output switches ON1-OS4 are connected from exactly

$$\frac{2d + d_2}{3}$$

switches in middle stage **150** through $2d+d_1$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1), ML(4,2), and ML(4,3); output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,16), ML(4,17), and ML(4,18); output switch OS1 is connected from middle switch MS(3,3) through the links ML(4,28), ML(4,29), and ML(4,30); and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,43), ML(4,44), and ML(4,45)).

[0431] Finally the connection topology of the network **300A1** shown in FIG. **3A1** is known to be back to back inverse Benes connection topology.

[0432] Referring to FIG. **3B1**, in one embodiment, an exemplary asymmetrical multi-link multi-stage network **300B1** with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by six switches IS1-IS4 and output stage **120** consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, six by six switches MS(1,1)-MS(1,4), middle stage **140** consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, four by eight switches MS(3,1)-MS(3,4).

[0433] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by six, the switches in output stage **120** are of size eight by six, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0434] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_1}{d}$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_1}{d}$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d * 3d$ and each output switch OS1-OS4 can be denoted in general with the notation $(2d+d_2) * d_2$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $3d * 3d$. The size of each switch in the last middle stage can be denoted as $2d * (2d+d_2)$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{link}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0435] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly $3 \times d$ switches in middle stage **130** through $3 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and ML(1,3); and also to middle switch MS(1,2) through the links ML(1,4), ML(1,5), and ML(1,6)).

[0436] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly d input switches through $3 \times d$ links (for example the links ML(1,1), ML(1,2), and ML(1,3) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,13), ML(1,14), and ML(1,15) are connected to the middle switch MS(1,1) from input switch IS3) and also are connected to exactly d switches in middle stage **140** through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected from middle switch MS(1,1) to middle switch MS(2,1); and the links ML(2,4), ML(2,5), and ML(2,6) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0437] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,13), ML(2,14), and ML(2,15) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage **150** through $3 \times d$ links (for

US 2011/0044329 A1

Feb. 24, 2011

42

example the links ML(3,1), ML(3,2), and ML(3,3) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,4), ML(3,5) and ML(3,6) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0438] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,13), ML(3,14), and ML(3,15) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly

$$\frac{2d + d_2}{3}$$

output switches in output stage 120 through $2d + d_2$ links

[0439] (For example the links ML(4,1), ML(4,2), and ML(4,3) are connected to output switch OS1 from Middle switch MS(3,1); the links ML(4,4), ML(4,5), and ML(4,6) are connected to output switch OS2 from middle switch MS(3,1); the links ML(4,7), ML(4,8), and ML(4,9) are connected to output switch OS3 from Middle switch MS(3,1); the links ML(4,10), ML(4,11), and ML(4,12) are connected to output switch OS2 from middle switch MS(3,1)).

[0440] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly

$$\frac{2d + d_2}{3}$$

switches in middle stage 150 through $2d + d_2$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1), ML(4,2), and ML(4,3); output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,16), ML(4,17), and ML(4,18); output switch OS1 is connected from middle switch MS(3,3) through the links ML(4,28), ML(4,29), and ML(4,30); and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,43), ML(4,44), and ML(4,45)).

[0441] Finally the connection topology of the network 300B1 shown in FIG. 3B1 is known to be back to back Omega connection topology.

[0442] Referring to FIG. 3C1, in one embodiment, an exemplary asymmetrical multi-link multi-stage network 300C1 with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic

blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by six switches IS1-IS4 and output stage 120 consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, six by six switches MS(1,1)-MS(1,4), middle stage 140 consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, four by eight switches MS(3,1)-MS(3,4).

[0443] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by six, the switches in output stage 120 are of size eight by six, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0444] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_1}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d * 3d$ and each output switch OS1-OS4 can be denoted in general with the notation $(2d + d_2) * d_2$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $3d * 3d$. The size of each switch in the last middle stage can be denoted as $2d * (2d + d_2)$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0445] Each of the

$$\frac{N_1}{d}$$

US 2011/0044329 A1

Feb. 24, 2011

43

input switches IS1-IS4 are connected to exactly $3 \times d$ switches in middle stage **130** through $3 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and ML(1,3); and also to middle switch MS(1,2) through the links ML(1,4), ML(1,5), and ML(1,6)).

[0446] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly d input switches through $3 \times d$ links (for example the links ML(1,1), ML(1,2), and ML(1,3) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,22), ML(1,23), and ML(1,24) are connected to the middle switch MS(1,1) from input switch IS4)) and also are connected to exactly d switches in middle stage **140** through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,4), ML(2,5), and ML(2,6) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0447] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,22), ML(2,23), and ML(2,24) are connected to the middle switch MS(2,1) from middle switch MS(1,4) and also are connected to exactly d switches in middle stage **150** through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,4), ML(3,5), and ML(3,6) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0448] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,22), ML(3,23), and ML(3,24) are connected to the middle switch MS(3,1) from middle switch MS(2,4) and also are connected to exactly

$$\frac{2d + d_2}{3}$$

output switches in output stage **120** through $2d + d_2$ links

[0449] (For example the links ML(4,1), ML(4,2), and ML(4,3) are connected to output switch OS1 from Middle switch MS(3,1); the links ML(4,4), ML(4,5), and ML(4,6) are connected to output switch OS2 from middle switch MS(3,1); the links ML(4,7), ML(4,8), and ML(4,9) are connected to output switch OS3 from Middle switch MS(3,1); the links ML(4,10), ML(4,11), and ML(4,12) are connected to output switch OS2 from middle switch MS(3,1)).

[0450] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly

$$\frac{2d + d_2}{3}$$

switches in middle stage **150** through $2d + d_2$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1), ML(4,2), and ML(4,3); output switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,16), ML(4,17), and ML(4,18); output switch OS1 is connected from middle switch MS(3,3) through the links ML(4,28), ML(4,29), and ML(4,30); and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,43), ML(4,44), and ML(4,45)).

[0451] Finally the connection topology of the network **300C1** shown in FIG. 3C1 is hereinafter called nearest neighbor connection topology.

[0452] Similar to network **300A1** of FIG. 3A1, **300B1** of FIG. 3B1, and **300C1** of FIG. 3C1, referring to FIG. 3D1, FIG. 3E1, FIG. 3F1, FIG. 3G1, FIG. 3H1, FIG. 3I1 and FIG. 3J1 with exemplary asymmetrical multi-link multi-stage networks **300D1**, **300E1**, **300F1**, **300G1**, **300H1**, **300I1**, and **300J1** respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by six switches IS1-IS4 and output stage **120** consists of four, six by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, six by six switches MS(1,1)-MS(1,4), middle stage **140** consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, six by six switches MS(3,1)-MS(3,4).

[0453] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by six, the switches in output stage **120** are of size six by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0454] The networks **300D1**, **300E1**, **300F1**, **300G1**, **300H1**, **300I1** and **300J1** of FIG. 3D1, FIG. 3E1, FIG. 3F1, FIG. 3G1, FIG. 3H1, FIG. 3I1, and FIG. 3J1 are also embodiments of asymmetric multi-link multi-stage network can be represented with the notation $V_{mink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links

US 2011/0044329 A1

Feb. 24, 2011

44

OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0455] Just like networks of 300A1, 300B1 and 300C1, for all the networks 300D1, 300E1, 300F1, 300G1, 300H1, 300I1 and 300J1 of FIG. 3D1, FIG. 3E1, FIG. 3F1, FIG. 3G1, FIG. 3H1, FIG. 3I1, and FIG. 3J1, each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly d switches in middle stage 130 through $3 \times d$ links.

[0456] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through $3 \times d$ links and also are connected to exactly d switches in middle stage 140 through $3 \times d$ links.

[0457] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $3 \times d$ links and also are connected to exactly d switches in middle stage 150 through $3 \times d$ links.

[0458] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $3 \times d$ links and also are connected to exactly

$$\frac{2d + d_2}{3}$$

output switches in output stage 120 through $2d + d$ links.

[0459] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly

$$\frac{2d + d_2}{3}$$

switches in middle stage 150 through $2d + d_2$ links

[0460] In all the ten embodiments of FIG. 3A1 to FIG. 3J1 the connection topology is different. That is the way the links ML(1,1)-ML(1,24), ML(2,1)-ML(2,24), ML(3,1)-ML(3,24), and ML(4,1)-ML(4,48) are connected between the respective stages is different. Even though only ten embodiments are illustrated, in general, the network $V_{mink}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mink}(N_1, N_2, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{mink}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{mink}(N_1, N_2, d, s)$ can be built. The ten embodiments of FIG. 3A1 to FIG. 3J1 are only three examples of network $V_{mink}(N, N_2, d, s)$.

[0461] In all the ten embodiments of FIG. 3A1 to FIG. 3J1, each of the links ML(1,1)-ML(1,24), ML(2,1)-ML(2,24), ML(3,1)-ML(3,24) and ML(4,1)-ML(4,48) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage 110 is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage 120 is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,4), MS(2,1)-MS(2,4), and MS(3,1)-MS(3,4) are referred to as middle switches or middle ports.

[0462] In the example illustrated in FIG. 3A1 (or in FIG. 3B1 to FIG. 3J1), a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage 130 will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage 130 when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network 300A1 (or 300B1 to 300J1), to be operated in strictly nonblocking manner in accordance with the invention.

[0463] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage 130 is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage 130, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the strictly nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Asymmetric SNB ($N_2 > N_1$) Embodiments:

[0464] Network 300K1 of FIG. 3K1 is an example of general asymmetrical multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_1) - 1$ stages where $N_2 > N_1$ and $N = p * N_1$ where $p > 1$. In network 300K1 of FIG. 3K1, $N_1 = N$ and $N_2 = p * N$. The general asymmetrical multi-link multi-

US 2011/0044329 A1

Feb. 24, 2011

45

stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for multicast when 3 according to the current invention (and in the example of FIG. 3K1, $s=3$). The general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_1) - 1$ stages has d inlet links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links IL1-IL(d) to the input switch IS1) and $3 \times d$ outgoing links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links ML(1,1)-ML(1,3d) to the input switch IS1). There are d_2 (where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d)$$

outlet links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example the links OL1-OL($p \times d$) to the output switch OS1) and $2d + d_2$ ($= 2d + p \times d$) incoming links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example ML($2 \times \log_d N_1 - 2, 1$)-ML($2 \times \log_d N_1 - 2, 2d + d_2$) to the output switch OS1).

[0465] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) are connected to exactly $3 \times d$ switches in middle stage **130** through $3 \times d$ links.

[0466] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1, N_1/d) in the middle stage **130** are connected from exactly d input switches through $3 \times d$ links and also are connected to exactly d switches in middle stage **140** through $3 \times d$ links.

[0467] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(\log_d N_1 - 1, 1) - MS(\log_d N_1 - 1, \frac{N_1}{d})$$

in the middle stage **130**+ $10^*(\log_d N_1 - 2)$ are connected from exactly d switches in middle stage **130**+ $10^*(\log_d N_1 - 3)$ through $3 \times d$ links and also are connected to exactly d switches in middle stage **130**+ $10^*(\log_d N_1 - 1)$ through $3 \times d$ links.

[0468] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(2 \times \log_d N_1 - 3, 1) - MS(2 \times \log_d N_1 - 3, \frac{N_1}{d})$$

in the middle stage **130**+ $10^*(2 \times \log_d N_1 - 4)$ are connected from exactly d switches in middle stage **130**+ $10^*(2 \times \log_d N_1 - 5)$ through $3 \times d$ links and also are connected to exactly

$$\frac{2d + d_2}{3}$$

output switches in output stage **120** through $2d + d_2$ links.

[0469] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) are connected from exactly

$$\frac{2d + d_2}{3}$$

switches in middle stage **130**+ $10^*(2 \times \log_d N_1 - 4)$ through $2d + d_2$ links.

[0470] As described before, again the connection topology of a general $V_{mlink}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{mlink}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N_1,$

US 2011/0044329 A1

Feb. 24, 2011

46

N_2, d, s) can be built. The embodiments of FIG. 3A1 to FIG. 3J1 are ten examples of network $V_{mlink}(N_1, N_2, d, s)$ for $s=3$ and $N_2 > N_1$.

[0471] The general symmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for multicast when 3 according to the current invention.

[0472] For example, the network of FIG. 3C1 shows an exemplary five-stage network, namely $V_{mlink}(8, 24, 2, 3)$, with the following multicast assignment $I_1 = \{1, 4\}$ and all other $I_j = \emptyset$ for $j = [2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,2) in middle stage 130, and fans out in middle switches MS(1,1) and MS(1,2) only once into middle switches MS(2,1) and MS(2,3) respectively in middle stage 140.

[0473] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,3) only once into middle switches MS(3,1) and MS(3,4) respectively in middle stage 150. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,4) only once into output switches OS1 and OS4 in output stage 120. Finally the connection I_1 fans out once in the output stage switch OS1 into outlet link OL2 and in the output stage switch OS4 twice into the outlet links OL19 and OL21. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage 130.

Asymmetric SNB ($N_1 > N_2$) Embodiments:

[0474] Referring to FIG. 3A2, in one embodiment, an exemplary asymmetrical multi-link multi-stage network 300A2 with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, six by eight switches IS1-IS4 and output stage 120 consists of four, six by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, eight by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, six by six switches MS(3,1)-MS(3,4).

[0475] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage 110 are of size six by eight, the switches in output stage 120 are of size six by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0476] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or and N_1 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p * N_2$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_2}{d},$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d_1 * (2d + d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $3d * d$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $3d * 3d$. The size of each switch in the first middle stage can be denoted as $(2d + d_1) * 3d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0477] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly

$$\frac{2d + d_1}{3}$$

switches in middle stage 130 through $2d + d_1$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and ML(1,3); input switch IS1 is also connected to middle switch MS(1,2) through the links ML(1,4), ML(1,5), and ML(1,6); input switch IS1 is connected to middle switch MS(1,3) through the links ML(1,7), ML(1,8), and ML(1,9); and input switch IS1 is also connected to middle switch MS(1,4) through the links ML(1,10), ML(1,11), and ML(1,12)).

[0478] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly

$$\frac{2d + d_1}{3}$$

US 2011/0044329 A1

Feb. 24, 2011

47

input switches through $2d+d_1$ links (for example middle switch MS(1,1) is connected from input switch IS1 through the links ML(1,1), ML(1,2), and ML(1,3); middle switch MS(1,1) is connected from input switch IS2 through the links ML(1,16), ML(1,17), and ML(1,18); middle switch MS(1,1) is connected from input switch IS3 through the links ML(1,28), ML(1,29), and ML(1,30); and middle switch MS(1,1) is connected from input switch IS4 through the links ML(1,43), ML(1,44), and ML(1,45)) and also are connected to exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,4), ML(2,5), and ML(2,6) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0479] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,16), ML(2,17), and ML(2,18) are connected to the middle switch MS(2,1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,4), ML(3,5), and ML(3,6) are connected from middle switch MS(2,1) to middle switch MS(3,3)).

[0480] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,16), ML(3,17), and ML(3,18) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly d output switches in output stage 120 through $3 \times d$ links (for example the links ML(4,1), ML(4,2), and ML(4,3) are connected to output switch OS1 from Middle switch MS(3,1), and the links ML(4,10), ML(4,11), and ML(4,12) are connected to output switch OS2 from middle switch MS(3,1)).

[0481] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through $3 \times d$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1), ML(4,2), and ML(4,3), and output

switch OS1 is also connected from middle switch MS(3,2) through the links ML(4,10), ML(4,11) and ML(4,12)).

[0482] Finally the connection topology of the network 300A2 shown in FIG. 3A2 is known to be back to back inverse Benes connection topology.

[0483] Referring to FIG. 3B2, in one embodiment, an exemplary asymmetrical multi-link multi-stage network 300B2 with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, six by eight switches IS1-IS4 and output stage 120 consists of four, six by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, eight by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, six by six switches MS(3,1)-MS(3,4).

[0484] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage 110 are of size six by eight, the switches in output stage 120 are of size six by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0485] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p \times N_2$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_2}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d_1 \times (2d+d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $3d \times d$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $3d \times 3d$. The size of each switch in the first middle stage can be denoted as $(2d+d_1) \times 3d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d

US 2011/0044329 A1

Feb. 24, 2011

48

represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0486] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly

$$\frac{2d + d_1}{3}$$

switches in middle stage 130 through $2d + d_1$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and ML(1,3); input switch IS1 is also connected to middle switch MS(1,2) through the links ML(1,4), ML(1,5), and ML(1,6); input switch IS1 is connected to middle switch MS(1,3) through the links ML(1,7), ML(1,8), and ML(1,9); and input switch IS1 is also connected to middle switch MS(1,4) through the links ML(1,10), ML(1,11), and ML(1,12)).

[0487] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly

$$\frac{2d + d_1}{3}$$

input switches through $2d + d_1$ links (for example middle switch MS(1,1) is connected from input switch IS1 through the links ML(1,1), ML(1,2), and ML(1,3); middle switch MS(1,1) is connected from input switch IS2 through the links ML(1,16), ML(1,17), and ML(1,18); middle switch MS(1,1) is connected from input switch IS3 through the links ML(1,28), ML(1,29), and ML(1,30); and middle switch MS(1,1) is connected from input switch IS4 through the links ML(1,43), ML(1,44), and ML(1,45)) and also are connected to exactly d switches in middle stage 140 through $2 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected from middle switch MS(1,1) to middle switch MS(2,1); and the links ML(2,4), ML(2,5), and ML(2,6) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0488] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,13), ML(2,14), and ML(2,15) are connected to the middle switch MS(2,

1) from middle switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,4), ML(3,5) and ML(3,6) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0489] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,13), ML(3,14), and ML(3,15) are connected to the middle switch MS(3,1) from middle switch MS(2,3)) and also are connected to exactly d output switches in output stage 120 through $3 \times d$ links (for example the links ML(4,1), ML(4,2), and ML(4,3) are connected to output switch OS1 from Middle switch MS(3,1), and the links ML(4,4), ML(4,5), and ML(4,6) are connected to output switch OS2 from middle switch MS(3,1)).

[0490] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through $3 \times d$ links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1), ML(4,2), and ML(4,3), and output switch OS1 is also connected from middle switch MS(3,3) through the links ML(4,13), ML(4,14), and ML(4,15)).

[0491] Finally the connection topology of the network 300B2 shown in FIG. 3B2 is known to be back to back Omega connection topology.

[0492] Referring to FIG. 3C2, in one embodiment, an exemplary asymmetrical multi-link multi-stage network 300C2 with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, six by eight switches IS1-IS4 and output stage 120 consists of four, six by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, eight by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, six by six switches MS(3,1)-MS(3,4).

[0493] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage 110 are of size six by eight, the switches in output stage 120 are of size six by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0494] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar

US 2011/0044329 A1

Feb. 24, 2011

49

switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p \cdot N_2$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_2}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d_1 \cdot (2d + d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $3d \cdot d$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $3d \cdot 3d$. The size of each switch in the first middle stage can be denoted as $(2d + d_1) \cdot 3d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-link multi-stage network can be represented with the notation $V_{link}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0495] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly

$$\frac{2d + d_1}{3}$$

switches in middle stage **130** through $2d + d_1$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and ML(1,3); input switch IS1 is also connected to middle switch MS(1,2) through the links ML(1,4), ML(1,5), and ML(1,6); input switch IS1 is connected to middle switch MS(1,3) through the links ML(1,7), ML(1,8), and ML(1,9); and input switch IS1 is also connected to middle switch MS(1,4) through the links ML(1,10), ML(1,11), and ML(1,12)).

[0496] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly

$$\frac{2d + d_1}{3}$$

input switches through $2d + d_1$ links (for example middle switch MS(1,1) is connected from input switch IS1 through the links ML(1,1), ML(1,2), and ML(1,3); middle switch MS(1,1) is connected from input switch IS2 through the links ML(1,16), ML(1,17), and ML(1,18); middle switch MS(1,1) is connected from input switch IS3 through the links ML(1,28), ML(1,29), and ML(1,30); and middle switch MS(1,1) is connected from input switch IS4 through the links ML(1,43), ML(1,44), and ML(1,45)) and also are connected to exactly d switches in middle stage **140** through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,4), ML(2,5), and ML(2,6) are connected from middle switch MS(1,1) to middle switch MS(2,2)).

[0497] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through $3 \times d$ links (for example the links ML(2,1), ML(2,2), and ML(2,3) are connected to the middle switch MS(2,1) from middle switch MS(1,1), and the links ML(2,22), ML(2,23), and ML(2,24) are connected to the middle switch MS(2,1) from middle switch MS(1,4)) and also are connected to exactly d switches in middle stage **150** through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,4), ML(3,5), and ML(3,6) are connected from middle switch MS(2,1) to middle switch MS(3,2)).

[0498] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through $3 \times d$ links (for example the links ML(3,1), ML(3,2), and ML(3,3) are connected to the middle switch MS(3,1) from middle switch MS(2,1), and the links ML(3,22), ML(3,23), and ML(3,24) are connected to the middle switch MS(3,1) from middle switch MS(2,4)) and also are connected to exactly d output switches in output stage **120** through $3 \times d$ links (for example the links ML(4,1), ML(4,2), and ML(4,3) are connected to output switch OS1 from middle switch

US 2011/0044329 A1

Feb. 24, 2011

50

MS(3,1), and the links ML(4,4), ML(4,5), and ML(4,6) are connected to output switch OS2 from middle switch MS(3,1)).

[0499] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through 3×d links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1), ML(4,2), and ML(4,3), and output switch OS1 is also connected from middle switch MS(3,4) through the links ML(4,22), ML(4,23), and ML(4,24)).

[0500] Finally the connection topology of the network 300C2 shown in FIG. 3C2 is hereinafter called nearest neighbor connection topology.

[0501] Similar to network 300A2 of FIG. 3A2, 300B2 of FIG. 3B2, and 300C2 of FIG. 3C2, referring to FIG. 3D2, FIG. 3E2, FIG. 3F2, FIG. 3G2, FIG. 3H2, FIG. 3I2 and FIG. 3J2 with exemplary asymmetrical multi-link multi-stage networks 300D2, 300E2, 300F2, 300G2, 300H2, 300I2, and 300J2 respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, six by eight switches IS1-IS4 and output stage 120 consists of four, six by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, eight by four switches MS(1,1)-MS(1,4), middle stage 140 consists of four, six by six switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, six by six switches MS(3,1)-MS(3,4).

[0502] Such a network can be operated in strictly non-blocking manner for multicast connections, because the switches in the input stage 110 are of size six by eight, the switches in output stage 120 are of size six by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0503] The networks 300D2, 300E2, 300F2, 300G2, 300H2, 300I2 and 300J2 of FIG. 3D2, FIG. 3E2, FIG. 3F2, FIG. 3G2, FIG. 3H2, FIG. 3I2, and FIG. 3J2 are also embodiments of asymmetric multi-link multi-stage network can be represented with the notation $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0504] Just like networks of 300A2, 300B2 and 300C2, for all the networks 300D2, 300E2, 300F2, 300G2, 300H2, 300I2 and 300J2 of FIG. 3D2, FIG. 3E2, FIG. 3F2, FIG. 3G2, FIG. 3H2, FIG. 3I2, and FIG. 3J2, each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly

$$\frac{2d + d_1}{3}$$

switches in middle stage 130 through 2d+d₁ links.

[0505] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly

$$\frac{2d + d_1}{3}$$

input switches through 2d+d₂ links and also are connected to exactly d switches in middle stage 140 through 3×d links.

[0506] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through 3×d links and also are connected to exactly d switches in middle stage 150 through 3×d links.

[0507] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through 3×d links and also are connected to exactly d output switches in output stage 120 through 3×d links.

[0508] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through 3×d links.

[0509] In all the ten embodiments of FIG. 3A2 to FIG. 3J2 the connection topology is different. That is the way the links ML(1,1)-ML(1,48), ML(2,1)-ML(2,24), ML(3,1)-ML(3,24), and ML(4,1)-ML(4,24) are connected between the respective stages is different. Even though only ten embodiments are illustrated, in general, the network $V_{mlink}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mlink}(N_1, N_2, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{mlink}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reach-

US 2011/0044329 A1

Feb. 24, 2011

51

able. Based on this property numerous embodiments of the network $V_{mlink}(N_1, N_2, d, s)$ can be built. The ten embodiments of FIG. 3A2 to FIG. 3J2 are only three examples of network $V_{mlink}(N_1, N_2, d, s)$.

[0510] In all the ten embodiments of FIG. 3A2 to FIG. 3J2, each of the links ML(1,1)-ML(1,48), ML(2,1)-ML(2,24), ML(3,1)-ML(3,24) and ML(4,1)-ML(4,24) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage 110 is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage 120 is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,4), MS(2,1)-MS(2,4), and MS(3,1)-MS(3,4) are referred to as middle switches or middle ports.

[0511] In the example illustrated in FIG. 3A2 (or in FIG. 3B2 to FIG. 3J2), a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage 130 will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage 130 when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network 300A2 (or 300B2 to 300J2), to be operated in strictly nonblocking manner in accordance with the invention.

[0512] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage 130 is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage 130, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the strictly nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Asymmetric SNB ($N_2 > N_1$) Embodiments:

[0513] Network 3001K2 of FIG. 3K2 is an example of general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_2) - 1$ stages where $N_1 > N_2$ and $N_1 = p \times N_2$ where $p > 1$. In network 300K2 of FIG. 3K2, $N_2 = N$ and $N_1 = p \times N$. The general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for multicast when $s=3$ according to the current invention (and in the example of FIG. 3K2, $s=3$). The general asymmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ with $(2 \times \log_d N_2) - 1$ stages has d , (where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d$$

inlet links for each of

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) (for example the links IL1-IL($p \times d$) to the input switch IS1) and $2d+d$, ($=2d+p \times d$) outgoing links for each of

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) (for example the links ML(1,1)-ML(1, ($d+p \times d$)) to the input switch IS1). There are d outlet links for each of

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example the links OL1-OL(d) to the output switch OS1) and $2 \times d$ incoming links for each of

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example ML($2 \times \log_d N_2 - 2, 1$)-ML($2 \times \log_d N_2 - 2, 3 \times d$) to the output switch OS1).

[0514] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) are connected to exactly

$$\frac{2d + d_1}{3}$$

switches in middle stage 130 through $2d+d_2$ links.

[0515] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1, N_2/d) in the middle stage 130 are connected from exactly d input switches through $3 \times d$ links and also are connected to exactly d switches in middle stage 140 through $3 \times d$ links.

[0516] Similarly each of the

$$\frac{N_2}{d}$$

US 2011/0044329 A1

Feb. 24, 2011

52

middle switches MS ($\text{Log}_d N_2 - 1, 1$)-

$$MS\left(\text{Log}_d N_2 - 1, \frac{N_2}{d}\right)$$

in the middle stage $130+10*(\text{Log}_d N_2 - 2)$ are connected from exactly d switches in middle stage $130+10*(\text{Log}_d N_2 - 3)$ through $3 \times d$ links and also are connected to exactly d switches in middle stage $130+10*(\text{Log}_d N_2 - 1)$ through $3 \times d$ links.

[0517] Similarly each of the

$$\frac{N_2}{d}$$

middle switches MS($2 \times \text{Log}_d N_2 - 3, 1$)-

$$MS\left(2 \times \text{Log}_d N_2 - 3, \frac{N_2}{d}\right)$$

in the middle stage $130+10*(2*\text{Log}_d N_2 - 4)$ are connected from exactly d switches in middle stage $130+10*(2*\text{Log}_d N_2 - 5)$ through $3 \times d$ links and also are connected to exactly d output switches in output stage **120** through $3 \times d$ links.

[0518] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) are connected from exactly d switches in middle stage $130+10*(2*\text{Log}_d N_2 - 4)$ through $2 \times d$ links.

[0519] As described before, again the connection topology of a general $V_{mlink}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{mlink}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{mlink}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 3A2 to FIG. 3J2 are ten examples of network $V_{mlink}(N_1, N_2, d, s)$ for $s=3$ and $N_2 > N_1$.

[0520] The general symmetrical multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention.

[0521] For example, the network of FIG. 3C2 shows an exemplary five-stage network, namely $V_{mlink}(8, 24, 2, 3)$, with the following multicast assignment $I_1 = \{1, 4\}$ and all other $I_j = \emptyset$ for $j = [2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,4) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,4) only once into middle switches MS(2,1) and MS(2,4) respectively in middle stage **140**.

[0522] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,4) only once into middle switches MS(3,1) and MS(3,4) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,4) only once into output switches OS1 and OS4 in output stage **120**. Finally the connection I_1 fans out once in the output stage switch OS1 into outlet link OL1 and in the output stage switch OS4 twice into the outlet links OL7 and OL8. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage **130**.

Folded Strictly Nonblocking Multi-Link Multi-Stage Networks:

[0523] The folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$, disclosed in the current invention, is topologically exactly the same as the multi-stage network $V_{mlink}(N_1, N_2, d, s)$, disclosed in U.S. Provisional Patent Application Ser. No. 60/940,392 that is incorporated by reference above, excepting that in the illustrations folded network $V_{fold-mlink}(N_1, N_2, d, s)$ is shown as it is folded at middle stage $130+10*(\text{Log}_d N_2 - 2)$.

[0524] The general symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention. Similarly the general asymmetrical folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention.

Folded Multi-Stage Network Embodiments:

Symmetric Folded RNB Embodiments:

[0525] Referring to FIG. 4A, in one embodiment, an exemplary symmetrical folded multi-stage network **400A** with five stages of thirty two switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by four switches IS1-IS4 and output stage **120** consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of eight, two by two switches MS(1,1)-MS(1,8), middle stage **140** consists of eight, two by two switches MS(2,1)-MS(2,8), and middle stage **150** consists of eight, two by two switches MS(3,1)-MS(3,8).

[0526] Such a network can be operated in strictly nonblocking manner for unicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are eight switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size two by four, the switches in output stage **120** are of size four by two, and there are eight switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0527] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

US 2011/0044329 A1

Feb. 24, 2011

53

N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by

$$2 \times \frac{N}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d \times 2d$ and each output switch OS1-OS4 can be denoted in general with the notation $2d \times d$. Likewise, the size of each switch in any of the middle stages can be denoted as $d \times d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric folded multi-stage network can be represented with the notation $V_{fold}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0528] Each of the N/d input switches IS1-IS4 are connected to exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,5) and MS(1,6) through the links ML(1,1), ML(1,2), ML(1,3) and ML(1,4) respectively).

[0529] Each of the

$$2 \times \frac{N}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,5) are connected to the middle switch MS(1,1) from input switch IS1 and IS2 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

[0530] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches MS(2,1)-MS(2,8) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,6) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,3) respectively).

[0531] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches MS(3,1)-MS(3,8) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,6) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage 120 through d links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switches MS(3,1)).

[0532] Each of the N/d output switches OS1-OS4 are connected from exactly $2 \times d$ switches in middle stage 150 through $2 \times d$ links (for example output switch OS1 is connected from middle switches MS(3,1), MS(3,2), MS(3,5) and MS(3,6) through the links ML(4,1), ML(4,3), ML(4,9) and ML(4,11) respectively).

[0533] Finally the connection topology of the network 400A shown in FIG. 4A is known to be back to back inverse Benes connection topology.

[0534] Referring to FIG. 4A1, in another embodiment of network $V_{fold}(N, d, s)$, an exemplary symmetrical folded multi-stage network 400A1 with five stages of thirty two switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of eight, two by two switches MS(1,1)-MS(1,8), middle stage 140 consists of eight, two by two switches MS(2,1)-MS(2,8), and middle stage 150 consists of eight, two by two switches MS(3,1)-MS(3,8).

[0535] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are eight switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are eight switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0536] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by

$$2 \times \frac{N}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d \times 2d$ and each output switch OS1-OS4 can be denoted in general with the notation $2d \times d$. Like-

US 2011/0044329 A1

Feb. 24, 2011

54

wise, the size of each switch in any of the middle stages can be denoted as $d*d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The symmetric folded multi-stage network of FIG. 4A1 is also the network of the type $V_{fold}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0537] Each of the N/d input switches IS1-IS4 are connected to exactly $2*d$ switches in middle stage 130 through $2*d$ links (for example input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,5) and MS(1,6) through the links ML(1,1), ML(1,2), ML(1,3) and ML(1,4) respectively).

[0538] Each of the

$$2 \times \frac{N}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,9) are connected to the middle switch MS(1,1) from input switch IS1 and IS3 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,2) respectively).

[0539] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches MS(2,1)-MS(2,8) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,5) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,2) respectively).

[0540] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches MS(3,1)-MS(3,8) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,5) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage 120 through d links (for example the links ML(4,1) and ML(4,2)

are connected to output switches OS1 and OS2 respectively from middle switches MS(3,1)).

[0541] Each of the N/d output switches OS1-OS4 are connected from exactly $2*d$ switches in middle stage 150 through $2*d$ links (for example output switch OS1 is connected from middle switches MS(3,1), MS(3,3), MS(3,5) and MS(3,7) through the links ML(4,1), ML(4,5), ML(4,9) and ML(4,13) respectively).

[0542] Finally the connection topology of the network 400A1 shown in FIG. 4A1 is known to be back to back Omega connection topology.

[0543] Referring to FIG. 4A2, in another embodiment of network $V_{fold}(N, d, s)$, an exemplary symmetrical folded multi-stage network 400A2 with five stages of thirty two switches for satisfying communication requests, such as setting up a telephone call or

[0544] a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of eight, two by two switches MS(1,1)-MS(1,8), middle stage 140 consists of eight, two by two switches MS(2,1)-MS(2,8), and middle stage 150 consists of eight, two by two switches MS(3,1)-MS(3,8).

[0545] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are eight switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size four by two, and there are eight switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0546] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by

$$2 \times \frac{N}{d}$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d*2d$ and each output switch OS1-OS4 can be denoted in general with the notation $2d*d$. Likewise, the size of each switch in any of the middle stages can be denoted as $d*d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The symmetric folded multi-stage network of FIG. 4A2 is also the network of the type $V_{fold}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary

US 2011/0044329 A1

Feb. 24, 2011

55

that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0547] Each of the N/d input switches IS1-IS4 are connected to exactly 2xd switches in middle stage 130 through 2xd links (for example input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,5) and MS(1,6) through the links ML(1,1), ML(1,2), ML(1,3) and ML(1,4) respectively).

[0548] Each of the

$$2 \times \frac{N}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,4) are connected to the middle switch MS(1,1) from input switch IS1 and IS4 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,2) respectively).

[0549] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches MS(2,1)-MS(2,8) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,8) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,4) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,2) respectively).

[0550] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches MS(3,1)-MS(3,8) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,8) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,4) respectively) and also are connected to exactly d output switches in output stage 120 through d links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switches MS(3,1)).

[0551] Each of the N/d output switches OS1-OS4 are connected from exactly 2xd switches in middle stage 150 through 2xd links (for example output switch OS1 is connected from middle switches MS(3,1), MS(3,4), MS(3,5) and MS(3,8) through the links ML(4,1), ML(4,2), ML(4,3) and ML(4,4) respectively).

[0552] Finally the connection topology of the network 400A2 shown in FIG. 4A2 is hereinafter called nearest neighbor connection topology.

[0553] In the three embodiments of FIG. 4A, FIG. 4A1 and FIG. 4A2 the connection topology is different. That is the way the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16), and ML(4,1)-ML(4,16) are connected between the respective stages is different. Even though only three embodiments are illustrated, in general, the network $V_{fold}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold}(N, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{fold}(N, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{fold}(N, d, s)$ can be built. The embodiments of FIG. 4A, FIG. 4A1, and FIG. 4A2 are only three examples of network $V_{fold}(N, d, s)$.

[0554] In the three embodiments of FIG. 4A, FIG. 4A1 and FIG. 4A2, each of the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16) and ML(4,1)-ML(4,16) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage 110 is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage 120 is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,8), MS(2,1)-MS(2,8), and MS(3,1)-MS(3,8) are referred to as middle switches or middle ports.

[0555] In the example illustrated in FIG. 4A (or in FIG. 1A1, or in FIG. 4A2), a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage 130 will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage 130 when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network 400A (or 400A1, or 400A2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0556] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage 130 is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage 130, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the rearrangeably nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Symmetric Folded RNB Embodiments:

[0557] Network 400B of FIG. 4B is an example of general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages. The general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the

US 2011/0044329 A1

Feb. 24, 2011

56

current invention. Also the general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can be operated in strictly non-blocking manner for unicast if $s \geq 2$ according to the current invention. (And in the example of FIG. 4B, $s=2$). The general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of N/d input switches IS1-IS(N/d) (for example the links IL1-IL(d) to the input switch IS1) and $2 \times d$ outgoing links for each of N/d input switches IS1-IS(N/d) (for example the links ML(1,1)-ML(1,2d) to the input switch IS1). There are d outlet links for each of N/d output switches OS1-OS(N/d) (for example the links OL1-OL(d) to the output switch OS1) and $2 \times d$ incoming links for each of N/d output switches OS1-OS(N/d) (for example ML($2 \times \log_d N - 2, 1$)-ML($2 \times \log_d N - 2, 2 \times d$) to the output switch OS1).

[0558] Each of the N/d input switches IS1-IS(N/d) are connected to exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example input switch IS1 is connected to middle switches MS(1,1)-MS(1, d) through the links ML(1,1)-ML(1, d) and to middle switches MS(1, $N/d+1$)-MS(1, $\{N/d\}+d$) through the links ML(1, $d+1$)-ML(1,2d) respectively.

[0559] Each of the

$$2 \times \frac{N}{d}$$

middle switches MS(1,1)-MS(1,2N/d) in the middle stage 130 are connected from exactly d input switches through d links and also are connected to exactly d switches in middle stage 140 through d links.

[0560] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches

$$MS(\log_d N - 1, 1) - MS(\log_d N - 1, 2 \times \frac{N}{d})$$

in the middle stage 130+10*($\log_d N - 2$) are connected from exactly d switches in middle stage 130+10*($\log_d N - 3$) through d links and also are connected to exactly d switches in middle stage 130+10*($\log_d N - 1$) through d links.

[0561] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches

$$MS(2 \times \log_d N - 3, 1) - MS(2 \times \log_d N - 3, 2 \times \frac{N}{d})$$

in the middle stage 130+10*($2 \times \log_d N - 4$) are connected from exactly d switches in middle stage 130+10*($2 \times \log_d N - 5$) through d links and also are connected to exactly d output switches in output stage 120 through d links.

[0562] Each of the N/d output switches OS1-OS(N/d) are connected from exactly $2 \times d$ switches in middle stage 130+10*($2 \times \log_d N - 4$) through $2 \times d$ links.

[0563] As described before, again the connection topology of a general $V_{fold}(N, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{fold}(N, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{fold}(N, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{fold}(N, d, s)$ can be built. The embodiments of FIG. 4A, FIG. 4A1, and FIG. 4A2 are three examples of network $V_{fold}(N, d, s)$.

[0564] The general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can be operated in rearrangeably non-blocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can be operated in strictly non-blocking manner for unicast if $s \geq 2$ according to the current invention.

[0565] Every switch in the folded multi-stage networks discussed herein has multicast capability. In a $V_{fold}(N, d, s)$ network, if a network inlet link is to be connected to more than one outlet link on the same output switch, then it is only necessary for the corresponding input switch to have one path to that output switch. This follows because that path can be multicast within the output switch to as many outlet links as necessary. Multicast assignments can therefore be described in terms of connections between input switches and output switches. An existing connection or a new connection from an input switch to r' output switches is said to have fan-out r' . If all multicast assignments of a first type, wherein any inlet link of an input switch is to be connected in an output switch to at most one outlet link are realizable, then multicast assignments of a second type, wherein any inlet link of each input switch is to be connected to more than one outlet link in the same output switch, can also be realized. For this reason, the following discussion is limited to general multicast connections of the first type (with fan-out r' ,

$$1 \leq r' \leq \frac{N}{d}$$

although the same discussion is applicable to the second type.

[0566] To characterize a multicast assignment, for each inlet link

$$i \in \{1, 2, \dots, \frac{N}{d}\},$$

let $I_i = O$, where

$$O \subset \{1, 2, \dots, \frac{N}{d}\},$$

denote the subset of output switches to which inlet link i is to be connected in the multicast assignment. For example, the

US 2011/0044329 A1

Feb. 24, 2011

57

network of FIG. 4A shows an exemplary five-stage network, namely $V_{fold}(8,2,2)$, with the following multicast assignment $I_1=\{2,3\}$ and all other $I_j=\emptyset$ for $j=[2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,5) in middle stage 130, and fans out in middle switches MS(1,1) and MS(1,5) only once into middle switches MS(2,1) and MS(2,5) respectively in middle stage 140.

[0567] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,5) only once into middle switches MS(3,1) and MS(3,7) respectively in middle stage 150. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,7) only once into output switches OS2 and OS3 in output stage 120. Finally the connection I_1 fans out once in the output stage switch OS2 into outlet link OL3 and in the output stage switch OS3 twice into the outlet links OL5 and OL6. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage 130.

Asymmetric Folded RNB ($N_2>N_1$) Embodiments:

[0568] Referring to FIG. 4C, in one embodiment, an exemplary asymmetrical folded multi-stage network 400C with five stages of thirty two switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage 130 consists of eight, two by two switches MS(1,1)-MS(1,8), middle stage 140 consists of eight, two by two switches MS(2,1)-MS(2,8), and middle stage 150 consists of eight, two by four switches MS(3,1)-MS(3,8).

[0569] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size eight by six, and there are eight switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size eight by six, and there are eight switches of size two by two in each of middle stage 130 and middle stage 140, and eight switches of size two by four in middle stage 150.

[0570] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2>N_1$ and $N_2=p*N_1$, where $p>1$. The number of middle switches in each middle stage is denoted by

$$2 \times \frac{N_1}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d*2d$ and each output switch OS1-OS4 can be denoted in general with the notation $(d+d_2)*d$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $d*d$. The size of each switch in the last middle stage can be denoted as

$$d * \frac{(d + d_2)}{2}.$$

A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric folded multi-stage network can be represented with the notation $V_{fold}(N_1, N_1, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2>N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0571] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,5) and MS(1,6) through the links ML(1,1), ML(1,2), ML(1,3) and ML(1,4) respectively).

[0572] Each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,5) are connected to the middle switch MS(1,1) from input switch IS1 and IS2 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

US 2011/0044329 A1

Feb. 24, 2011

58

[0573] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,8) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,6) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,3) respectively).

[0574] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,8) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,6) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly

$$\frac{d + d_2}{2}$$

output switches in output stage 120 through

$$\frac{d + d_2}{2}$$

links (for example the links ML(4,1), ML(4,2), ML(4,3) and ML(4,4) are connected to output switches OS1, OS2, OS3, and OS4 respectively from middle switches MS(3,1)).

[0575] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly d+d₂ switches in middle stage 150 through d+d₂ links (for example output switch OS1 is connected from middle switches MS(3,1), MS(3,2), MS(3,3), MS(3,4), MS(3,5), MS(3,6), MS(3,7), and MS(3,8) through the links ML(4,1), ML(4,5), ML(4,9), ML(4,13), ML(4,17), ML(4,21), ML(4,25) and ML(4,29) respectively).

[0576] Finally the connection topology of the network 400C shown in FIG. 4C is known to be back to back inverse Benes connection topology.

[0577] Referring to FIG. 4C1, in another embodiment of network V_{fold}(N₁, N₂, d, s), an exemplary asymmetrical folded multi-stage network 400C1 with five stages of thirty two switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage

110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage 130 consists of eight, two by two switches MS(1,1)-MS(1,8), middle stage 140 consists of eight, two by two switches MS(2,1)-MS(2,8), and middle stage 150 consists of eight, two by four switches MS(3,1)-MS(3,8).

[0578] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size eight by six, and there are eight switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size eight by six, and there are eight switches of size two by two in each of middle stage 130 and middle stage 140, and eight switches of size two by four in middle stage 150.

[0579] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N₁ is the total number of inlet links or and N₂ is the total number of outlet links and N₂>N₁ and N₂=p*N₁, where p>1. The number of middle switches in each middle stage is denoted by

$$2 \times \frac{N_1}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation d*2d and each output switch OS1-OS4 can be denoted in general with the notation (d+d₂)*d, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as d*d. The size of each switch in the last middle stage can be denoted as

$$d * \frac{(d + d_2)}{2}.$$

A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The asymmetric folded multi-stage network of FIG. 4C1 is also the network of the type V_{fold}(N₁, N₂, d, s), where N₁ represents the total number of inlet links of all input switches (for example the links IL1-IL8), N₂ represents the total number of outlet links of all output switches (for example the links OL1-OL24), d repre-

US 2011/0044329 A1

Feb. 24, 2011

59

sents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0580] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,5) and MS(1,6) through the links ML(1,1), ML(1,2), ML(1,3) and ML(1,4) respectively).

[0581] Each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,9) are connected to the middle switch MS(1,1) from input switch IS1 and IS3 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,2) respectively).

[0582] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,8) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,5) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,2) respectively).

[0583] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,8) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,5) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly

$$\frac{d + d_2}{2}$$

output switches in output stage 120 through

$$\frac{d + d_2}{2}$$

links (for example the links ML(4,1), ML(4,2), ML(4,3) and ML(4,4) are connected to output switches OS1, OS2, OS3, and OS4 respectively from middle switches MS(3,1)).

[0584] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly $d+d_1$ switches in middle stage 150 through $d+d_1$ links (for example output switch OS1 is connected from middle switches MS(3,1), MS(3,2), MS(3,3), MS(3,4), MS(3,5), MS(3,6), MS(3,7), and MS(3,8) through the links ML(4,1), ML(4,5), ML(4,9), ML(4,13), ML(4,17), ML(4,21), ML(4,25) and ML(4,29) respectively).

[0585] Finally the connection topology of the network 400C1 shown in FIG. 4C1 is known to be back to back Omega connection topology.

[0586] Referring to FIG. 4C2, in another embodiment of network $V_{fold}(N_1, N_2, d, s)$, an exemplary asymmetrical folded multi-stage network 400C2 with five stages of thirty two switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by four switches IS1-IS4 and output stage 120 consists of four, eight by six switches OS1-OS4. And all the middle stages namely middle stage 130 consists of eight, two by two switches MS(1,1)-MS(1,8), middle stage 140 consists of eight, two by two switches MS(2,1)-MS(2,8), and middle stage 150 consists of eight, two by four switches MS(3,1)-MS(3,8).

[0587] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size eight by six, and there are eight switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size two by four, the switches in output stage 120 are of size eight by six, and there are eight switches of size two by two in each of middle stage 130 and middle stage 140, and eight switches of size two by four in middle stage 150.

[0588] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_1}{d}$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

US 2011/0044329 A1

Feb. 24, 2011

60

$$2 \times \frac{N_1}{d}$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d \times 2d$ and each output switch OS1-OS4 can be denoted in general with the notation $(d+d_2) \times d$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $d \times d$. The size of each switch in the last middle stage can be denoted as

$$d * \frac{(d + d_2)}{2}.$$

A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The asymmetric folded multi-stage network of FIG. 4C2 is also the network of the type $V_{fold}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0589] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,5) and MS(1,6) through the links ML(1,1), ML(1,2), ML(1,3) and ML(1,4) respectively).

[0590] Each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,14) are connected to the middle switch MS(1,1) from input switch IS1 and IS4 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,2) respectively).

[0591] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,8) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,8) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,4) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,2) respectively).

[0592] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,8) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,8) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,4) respectively) and also are connected to exactly

$$\frac{d + d_2}{2}$$

output switches in output stage 120 through

$$\frac{d + d_2}{2}$$

links (for example the links ML(4,1), ML(4,2), ML(4,3) and ML(4,4) are connected to output switches OS1, OS2, OS3, and OS4 respectively from middle switches MS(3,1)).

[0593] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly $d+d_2$ switches in middle stage 150 through $d+d_1$ links (for example output switch OS1 is connected from middle switches MS(3,1), MS(3,2), MS(3,3), MS(3,4), MS(3,5), MS(3,6), MS(3,7), and MS(3,8) through the links ML(4,1), ML(4,5), ML(4,9), ML(4,13), ML(4,17), ML(4,21), ML(4,25) and ML(4,29) respectively).

[0594] Finally the connection topology of the network 400C2 shown in FIG. 4C2 is hereinafter called nearest neighbor connection topology.

[0595] In the three embodiments of FIG. 4C, FIG. 4C1 and FIG. 4C2 the connection topology is different. That is the way the links ML(1,1)-ML(1,16), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16), and ML(4,1)-ML(4,16) are connected between the respective stages is different. Even though only three embodiments are illustrated, in general, the network $V_{fold}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology.

US 2011/0044329 A1

Feb. 24, 2011

61

For example the connection topology of the network $V_{fold}(N_1, N_2, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{fold}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{fold}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 4C, FIG. 4C1, and FIG. 4C2 are only three examples of network $V_{fold}(N_1, N_2, d, s)$.

[0596] In the three embodiments of FIG. 4C, FIG. 4C1 and FIG. 4C2, each of the links ML(1,1)-ML(1,32), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16) and ML(4,1)-ML(4,16) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage 110 is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage 120 is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,8), MS(2,1)-MS(2,8), and MS(3,1)-MS(3,8) are referred to as middle switches or middle ports.

[0597] In the example illustrated in FIG. 4C (or in FIG. 1C1, or in FIG. 4C2), a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage 130 will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage 130 when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network 400C (or 400C1, or 400C2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0598] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage 130 is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage 130, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the rearrangeably nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Asymmetric Folded RNB ($N_2 > N_1$) Embodiments:

[0599] Network 400D of FIG. 4D is an example of general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ with $(2 \times \log_d N_1) - 1$ stages where $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. In network 400D of FIG. 4D, $N_1 = N$ and $N_2 = p * N$. The general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention. (And in the example of FIG. 4D, $s=2$). The

general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ with $(2 \times \log_d N_1) - 1$ stages has d inlet links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links IL1-IL(d) to the input switch IS1) and $2 \times d$ outgoing links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links ML(1,1)-ML(1,2d) to the input switch IS1). There are d_2 (where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d)$$

outlet links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example the links OL1-OL($p * d$) to the output switch OS1) and $d + d$, ($= d + p \times d$) incoming links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example ML($2 \times \log_d N_1 - 2, 1$)-ML($2 \times \log_d N_1 - 2, d + d_2$) to the output switch OS1).

[0600] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) are connected to exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example in one embodiment the input switch IS1 is connected to middle switches MS(1,1)-MS(1, d) through the links ML(1, 1)-ML(1, d) and to middle switches MS(1, $N_1/d + 1$)-MS(1, $\{N_1/d\} + d$) through the links ML(1, $d + 1$)-ML(1, $2d$) respectively.

[0601] Each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(1,1)-MS(1, $2N_1/d$) in the middle stage 130 are connected from exactly d input switches through d links and also are connected to exactly d switches in middle stage 140 through d links.

US 2011/0044329 A1

Feb. 24, 2011

62

[0602] Similarly each of the

$$MS(\text{Log}_d N_1 - 1, 1) - MS\left(\text{Log}_d N_1 - 1, 2 \times \frac{N_1}{d}\right)$$

middle switches

$$2 \times \frac{N_1}{d}$$

in the middle stage $130+10*(\text{Log}_d N_1-2)$ are connected from exactly d switches in middle stage $130+10*(\text{Log}_d N_1-3)$ through d links and also are connected to exactly d switches in middle stage $130+10*(\text{Log}_d N_1-1)$ through d links.

[0603] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches

$$MS(2 \times \text{Log}_d N_1 - 3, 1) - MS\left(2 \times \text{Log}_d N_1 - 3, 2 \times \frac{N_1}{d}\right)$$

[0604] in the middle stage $130+10*(2*\text{Log}_d N_1-4)$ are connected from exactly d switches in middle stage $130+10*(2*\text{Log}_d N_1-5)$ through d links and also are connected to exactly d output switches in output stage **120** through d links.

[0605] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) are connected from exactly $d+d_2$ switches in middle stage $130+10*(2*\text{Log}_d N_1-4)$ through $d+d_1$ links.

[0606] As described before, again the connection topology of a general $V_{fold}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{fold}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{fold}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{fold}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 4C, FIG. 4C1, and FIG. 4C2 are three examples of network $V_{fold}(N_1, N_2, d, s)$ for $s=2$ and $N_2 > N_1$.

[0607] The general symmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if 2 according to the current invention.

[0608] For example, the network of FIG. 4C shows an exemplary five-stage network, namely $V_{fold}(8, 24, 2, 2)$, with the following multicast assignment $I_1 = \{2, 3\}$ and all other $I_j = \phi$ for $j = [2-8]$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,5) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,5) only once into middle switches MS(2,1) and MS(2,5) respectively in middle stage **140**.

[0609] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,5) only once into middle switches MS(3,1) and MS(3,7) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,7) only once into output switches OS2 and OS3 in output stage **120**. Finally the connection I_1 fans out once in the output stage switch OS2 into outlet link OL7 and in the output stage switch OS3 twice into the outlet links OL13 and OL16. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage **130**.

Asymmetric Folded RNB ($N_1 > N_2$) Embodiments:

[0610] Referring to FIG. 4E, in one embodiment, an exemplary asymmetrical folded multi-stage network **400E** with five stages of thirty two switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, six by eight switches IS1-IS4 and output stage **120** consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of eight, four by two switches MS(1,1)-MS(1,8), middle stage **140** consists of eight, two by two switches MS(2,1)-MS(2,8), and middle stage **150** consists of eight, two by two switches MS(3,1)-MS(3,8).

[0611] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are eight switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are eight switches of size four by two in middle stage **130**, and eight switches of size two by two in middle stage **140** and middle stage **150**.

[0612] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p * N$, where $p > 1$. The number of middle switches in each middle stage is denoted by

US 2011/0044329 A1

Feb. 24, 2011

63

$$2 \times \frac{N_2}{d}$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d \times (d+d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $(2 \times d \times d)$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $d \times d$. The size of each switch in the first middle stage can be denoted as

$$\frac{(d + d_1)}{2} \times d.$$

A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric folded multi-stage network can be represented with the notation $V_{fold}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0613] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly $d+d_1$ switches in middle stage 130 through $d+d_1$ links (for example input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,3), MS(1,4), MS(1,5), MS(1,6), MS(1,7), and MS(1,8) through the links ML(1,1), ML(1,2), ML(1,3), ML(1,4), ML(1,5), ML(1,6), ML(1,7), and ML(1,8) respectively).

[0614] Each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 are connected from exactly

$$\frac{(d + d_1)}{2}$$

input switches through

$$\frac{(d + d_1)}{2}$$

links (for example the links ML(1,1), ML(1,9), ML(1,17) and ML(1,25) are connected to the middle switch MS(1,1) from input switch IS1, IS2, IS3, and IS4 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

[0615] Similarly each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,8) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,6) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,3) respectively).

[0616] Similarly each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,8) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,6) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage 120 through d links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switches MS(3,1)).

[0617] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly $2 \times d$ switches in middle stage 150 through $2 \times d$ links (for example output switch OS1 is connected from middle switches MS(3,1), MS(3,2), MS(3,5), and MS(3,6) through the links ML(4,1), ML(4,3), ML(4,9), and ML(4,11) respectively).

[0618] Finally the connection topology of the network 400E shown in FIG. 4E is known to be back to back inverse Benes connection topology.

[0619] Referring to FIG. 4E1, in another embodiment of network $V_{fold}(N_1, N_2, d, s)$, an exemplary asymmetrical folded multi-stage network 400E1 with five stages of thirty two switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150

US 2011/0044329 A1

Feb. 24, 2011

64

is shown where input stage **110** consists of four, six by eight switches IS1-IS4 and output stage **120** consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of eight, four by two switches MS(1,1)-MS(1,8), middle stage **140** consists of eight, two by two switches MS(2,1)-MS(2,8), and middle stage **150** consists of eight, two by two switches MS(3,1)-MS(3,8).

[0620] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are eight switches in each of middle stage **130**, middle stage **140** and middle stage **150**. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage **110** are of size six by eight, the switches in output stage **120** are of size four by two, and there are eight switches of size four by two in middle stage **130**, and eight switches of size two by two in middle stage **140** and middle stage **150**.

[0621] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p * N_2$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$2 \times \frac{N_2}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d * (d + d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $(2 * d * d)$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $d * d$. The size of each switch in the first middle stage can be denoted as

$$\frac{(d + d_1)}{2} * d.$$

A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The asymmetric folded multi-stage network of FIG. 4E1 is also the network of the type $V_{fold}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each input switch where $N_1 > N_2$, and s

is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0622] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly $d + d_1$ switches in middle stage **130** through $d + d_1$ links (for example input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,3), MS(1,4), MS(1,5), MS(1,6), MS(1,7), and MS(1,8) through the links ML(1,1), ML(1,2), ML(1,3), ML(1,4), ML(1,5), ML(1,6), ML(1,7), and ML(1,8) respectively).

[0623] Each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage **130** are connected from exactly

$$\frac{(d + d_1)}{2}$$

input switches through

$$\frac{(d + d_1)}{2}$$

links (for example the links ML(1,1), ML(1,9), ML(1,17) and ML(1,25) are connected to the middle switch MS(1,1) from input switch IS1, IS2, IS3, and IS4 respectively) and also are connected to exactly d switches in middle stage **140** through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,2) respectively).

[0624] Similarly each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,8) in the middle stage **140** are connected from exactly d switches in middle stage **130** through d links (for example the links ML(2,1) and ML(2,5) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage **150** through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,2) respectively).

[0625] Similarly each of the

$$2 \times \frac{N_2}{d}$$

US 2011/0044329 A1

Feb. 24, 2011

65

middle switches MS(3,1)-MS(3,8) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,5) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage 120 through d links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switches MS(3,1)).

[0626] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly $2 \times d$ switches in middle stage 150 through $2 \times d$ links (for example output switch OS1 is connected from middle switches MS(3,1), MS(3,3), MS(3,5), and MS(3,7) through the links ML(4,1), ML(4,5), ML(4,9), and ML(4,13) respectively).

[0627] Finally the connection topology of the network 400E1 shown in FIG. 4E1 is known to be back to back Omega connection topology.

[0628] Referring to FIG. 4E2, in another embodiment of network $V_{fold}(N_1, N_2, d, s)$, an exemplary asymmetrical folded multi-stage network 400E2 with five stages of thirty two switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, six by eight switches IS1-IS4 and output stage 120 consists of four, four by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of eight, four by two switches MS(1,1)-MS(1,8), middle stage 140 consists of eight, two by two switches MS(2,1)-MS(2,8), and middle stage 150 consists of eight, two by two switches MS(3,1)-MS(3,8).

[0629] Such a network can be operated in strictly non-blocking manner for unicast connections, because the switches in the input stage 110 are of size six by eight, the switches in output stage 120 are of size four by two, and there are eight switches in each of middle stage 130, middle stage 140 and middle stage 150. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the input stage 110 are of size six by eight, the switches in output stage 120 are of size four by two, and there are eight switches of size four by two in middle stage 130, and eight switches of size two by two in middle stage 140 and middle stage 150.

[0630] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. The number of middle switches in each middle stage is denoted by

$$2 \times \frac{N_2}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d * (d + d_1)$ and each output switch OS1-OS4 can be denoted in general with the notation $(2 \times d * d)$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $d * d$. The size of each switch in the first middle stage can be denoted as

$$\frac{(d + d_1)}{2} * d.$$

A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. The asymmetric folded multi-stage network of FIG. 4E1 is also the network of the type $V_{fold}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each input switch where $N_1 > N_2$, and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0631] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly $d + d_1$ switches in middle stage 130 through $d + d_1$ links (for example input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,3), MS(1,4), MS(1,5), MS(1,6), MS(1,7), and MS(1,8) through the links ML(1,1), ML(1,2), ML(1,3), ML(1,4), ML(1,5), ML(1,6), ML(1,7), and ML(1,8) respectively).

[0632] Each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 are connected from exactly

$$\frac{(d + d_1)}{2}$$

US 2011/0044329 A1

Feb. 24, 2011

66

input switches through

$$\frac{(d + d_1)}{2}$$

links (for example the links ML(1,1), ML(1,9), ML(1,17) and ML(1,25) are connected to the middle switch MS(1,1) from input switch IS1, IS2, IS3, and IS4 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,2) respectively).

[0633] Similarly each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,8) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,8) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,4) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,2) respectively).

[0634] Similarly each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,8) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,8) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,4) respectively) and also are connected to exactly d output switches in output stage 120 through d links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switches MS(3,1)).

[0635] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS4 are connected from exactly $2 \times d$ switches in middle stage 150 through $2 \times d$ links (for example output switch OS1 is connected from middle switches MS(3,1), MS(3,4), MS(3,5), and MS(3,8) through the links ML(4,1), ML(4,8), ML(4,9), and ML(4,16) respectively).

[0636] Finally the connection topology of the network 400E2 shown in FIG. 4E2 is hereinafter called nearest neighbor connection topology.

[0637] In the three embodiments of FIG. 4E, FIG. 4E1 and FIG. 4E2 the connection topology is different. That is the way the links ML(1,1)-ML(1,32), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16), and ML(4,1)-ML(4,16) are connected between the respective stages is different. Even though only three embodiments are illustrated, in general, the network $V_{fold}(N_1, N_2, d, s)$

s) can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold}(N_1, N_2, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{fold}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{fold}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 4E, FIG. 4E1, and FIG. 4E2 are only three examples of network $V_{fold}(N_1, N_2, d, s)$.

[0638] In the three embodiments of FIG. 4E, FIG. 4E1 and FIG. 4E2, each of the links ML(1,1)-ML(1,32), ML(2,1)-ML(2,16), ML(3,1)-ML(3,16) and ML(4,1)-ML(4,16) are either available for use by a new connection or not available if currently used by an existing connection. The input switches IS1-IS4 are also referred to as the network input ports. The input stage 110 is often referred to as the first stage. The output switches OS1-OS4 are also referred to as the network output ports. The output stage 120 is often referred to as the last stage. The middle stage switches MS(1,1)-MS(1,8), MS(2,1)-MS(2,8), and MS(3,1)-MS(3,8) are referred to as middle switches or middle ports.

[0639] In the example illustrated in FIG. 4E (or in FIG. 1E1, or in FIG. 4E2), a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two switches in middle stage 130 will be used. Similarly, although a fan-out of three is possible for a multicast connection request if the input switch is IS1, again only a fan-out of two is used. The specific middle switches that are chosen in middle stage 130 when selecting a fan-out of two is irrelevant so long as at most two middle switches are selected to ensure that the connection request is satisfied. In essence, limiting the fan-out from input switch to no more than two middle switches permits the network 400E (or 400E1, or 400E2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0640] The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, a fan-out of one is used, i.e. a single middle stage switch in middle stage 130 is used to satisfy the request. Moreover, although in the above-described embodiment a limit of two has been placed on the fan-out into the middle stage switches in middle stage 130, the limit can be greater depending on the number of middle stage switches in a network (while maintaining the rearrangeably nonblocking nature of operation of the network for multicast connections). However any arbitrary fan-out may be used within any of the middle stage switches and the output stage switches to satisfy the connection request.

Generalized Asymmetric Folded RNB ($N_1 > N_2$) Embodiments:

[0641] Network 400F of FIG. 4F is an example of general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ with $(2 \times \log_d N_2) - 1$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network 400D of FIG. 4F, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the cur-

US 2011/0044329 A1

Feb. 24, 2011

67

rent invention. (And in the example of FIG. 4F, $s=2$). The general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ with $(2 \times \log_d N_2) - 1$ stages has d_1 (where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d$$

inlet links for each of

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) (for example the links IL1-IL($p \times d$) to the input switch IS1) and $d+d$, ($=d+p \times d$) outgoing links for each of

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) (for example the links ML(1,1)-ML(1, $d+p \times d$) to the input switch IS1). There are d outlet links for each of

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example the links OL1-OL(d) to the output switch OS1) and $2 \times d$ incoming links for each of

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example ML($2 \times \log_d N_2 - 2, 1$)-ML($2 \times \log_d N_2 - 2, 2 \times d$) to the output switch OS1).

[0642] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) are connected to exactly $d+d_1$ switches in middle stage **130** through $d+d_1$ links (for example in one embodiment the input switch IS1 is connected to middle switches MS(1,1)-MS(1, $(d+d_1)/2$) through the links ML(1,1)-ML(1, $(d+d_1)/2$) and to middle switches MS(1, $N_1/d + 1$)-MS(1, $\{N_1/d\} + (d+d_1)/2$) through the links ML(1, $((d+d_1)/2 + 1)$ -ML(1, $(d+d_1)$) respectively.

[0643] Each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(1,1)-MS(1, $2 \times N_2/d$) in the middle stage **130** are connected from exactly d input switches through d links and also are connected to exactly d switches in middle stage **140** through d links.

[0644] Similarly each of the

$$2 \times \frac{N_2}{d}$$

middle switches

$$MS(\log_d N_2 - 1, 1) - MS(\log_d N_2 - 1, 2 \times \frac{N_2}{d})$$

in the middle stage **130** + $10 \times (\log_d N_2 - 2)$ are connected from exactly d switches in middle stage **130** + $10 \times (\log_d N_2 - 3)$ through d links and also are connected to exactly d switches in middle stage **130** + $10 \times (\log_d N_2 - 1)$ through d links.

[0645] Similarly each of the

$$2 \times \frac{N_2}{d}$$

middle switches

$$MS(2 \times \log_d N_2 - 3, 1) - MS(2 \times \log_d N_2 - 3, 2 \times \frac{N_2}{d})$$

in the middle stage **130** + $10 \times (2 \times \log_d N_2 - 4)$ are connected from exactly d switches in middle stage **130** + $10 \times (2 \times \log_d N_2 - 5)$ through d links and also are connected to exactly d output switches in output stage **120** through d links.

[0646] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) are connected from exactly $2 \times d$ switches in middle stage **130** + $10 \times (2 \times \log_d N_2 - 4)$ through $2 \times d$ links.

[0647] As described before, again the connection topology of a general $V_{fold}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{fold}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{fold}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{fold}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 4E, FIG. 4E1, and FIG. 4E2 are three examples of network $V_{fold}(N_1, N_2, d, s)$ for $s=2$ and $N_1 > N_2$.

[0648] The general symmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $S \geq 2$ according to the current invention.

US 2011/0044329 A1

Feb. 24, 2011

68

[0649] For example, the network of FIG. 4E shows an exemplary five-stage network, namely $V_{fold}(24,8,2,2)$, with the following multicast assignment $I_1=\{2,3\}$ and all other $I_j=\phi$ for $j=\{2-8\}$. It should be noted that the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and MS(1,5) in middle stage 130, and fans out in middle switches MS(1,1) and MS(1,5) only once into middle switches MS(2,1) and MS(2,5) respectively in middle stage 140.

[0650] The connection I_1 also fans out in middle switches MS(2,1) and MS(2,5) only once into middle switches MS(3,1) and MS(3,7) respectively in middle stage 150. The connection I_1 also fans out in middle switches MS(3,1) and MS(3,7) only once into output switches OS2 and OS3 in output stage 120. Finally the connection I_1 fans out once in the output stage switch OS2 into outlet link OL3 and in the output stage switch OS3 twice into the outlet links OL5 and OL6. In accordance with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage 130.

SNB Embodiments:

[0651] The folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ disclosed, in the current invention, is topologically exactly the same as the multi-stage network $V_{fold}(N_1, N_2, d, s)$, disclosed in U.S. Provisional Patent Application Ser. No. 60/940,391 that is incorporated by reference above, excepting that in the illustrations folded network $V_{fold}(N_1, N_2, d, s)$ is shown as it is folded at middle stage $130+10*(\text{Log}_d N_1-2)$.

[0652] The general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention. Similarly the general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can also be operated in strictly nonblocking manner for multicast when $S \geq 3$ according to the current invention.

Symmetric Folded RNB Unicast Embodiments:

[0653] Referring to FIG. 5A, an exemplary symmetrical folded multi-stage network 500A respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by two switches IS1-IS4 and output stage 120 consists of four, two by two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, two by two switches MS(1,1)-MS(1,4), middle stage 140 consists of four, two by two switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, two by two switches MS(3,1)-MS(3,4).

[0654] Such a network can be operated in rearrangeably nonblocking manner for unicast connections, because the switches in the input stage 110 are of size two by two, the switches in output stage 120 are of size two by two, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0655] The connection topology of the network 500A shown in FIG. 5A is known to be back to back inverse Benes connection topology. In other embodiments the connection topology is different. That is the way the links ML(1,1)-ML

(1,8), ML(2,1)-ML(2,8), ML(3,1)-ML(3,8), and ML(4,1)-ML(4,8) are connected between the respective stages is different.

[0656] Even though only one embodiment is illustrated, in general, the network $V_{fold}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold}(N, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{fold}(N, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{fold}(N, d, s)$ can be built. The embodiment of FIG. 5A is only one example of network $V_{fold}(N, d, s)$.

[0657] The network 500A of FIG. 5A is also rearrangeably nonblocking for unicast according to the current invention. In one embodiment of these networks each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch IS1-IS4 can be denoted in general with the notation $d*d$ and each output switch OS1-OS4 can be denoted in general with the notation $d*d$. Likewise, the size of each switch in any of the middle stages can be denoted as $d*d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric folded multi-stage network can be represented with the notation $V_{fold}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0658] In network 500A of FIG. 5A, each of the N/d input switches IS1-IS4 are connected to exactly d switches in middle stage 130 through d links (for example input switch IS1 is connected to middle switches MS(1,1) and MS(1,2) through the links ML(1,1) and ML(1,2) respectively).

[0659] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,4) are connected to the middle switch MS(1,1) from input switch IS1 and IS2 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

[0660] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,6) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,3) respectively).

US 2011/0044329 A1

Feb. 24, 2011

69

[0661] Similarly each of the N/d middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,6) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage 120 through d links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switch MS(3,1)).

[0662] Each of the N/d output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through d links (for example output switch OS1 is connected from middle switches MS(3,1) and MS(3,2) through the links ML(4,1) and ML(4,4) respectively).

Generalized Symmetric Folded RNB Unicast Embodiments:

[0663] Network 500B of FIG. 5B is an example of general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages. The general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention (and in the example of FIG. 5B, $s=1$). The general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of N/d input switches IS1-IS(N/d) (for example the links IL1-IL(d) to the input switch IS1) and d outgoing links for each of N/d input switches IS1-IS(N/d) (for example the links ML(1,1)-ML(1, d) to the input switch IS1). There are d outlet links for each of N/d output switches OS1-OS(N/d) (for example OL1-OL(d) to the output switch OS1) and d incoming links for each of N/d output switches OS1-OS(N/d) (for example ML(2xLog_dN-2,1)-ML(2xLog_dN-2, d) to the output switch OS1).

[0664] Each of the N/d input switches IS1-IS(N/d) are connected to exactly d switches in middle stage 130 through d links.

[0665] Each of the N/d middle switches MS(1,1)-MS(1,N/d) in the middle stage 130 are connected from exactly d input switches through d links and also are connected to exactly d switches in middle stage 140 through d links.

[0666] Similarly each of the N/d middle switches

$$MS(\text{Log}_d N - 1, 1) - MS\left(\text{Log}_d N - 1, \frac{N}{d}\right)$$

in the middle stage 130+10*(Log_d N-2) are connected from exactly d switches in middle stage 130+10*(Log_d N-3) through d links and also are connected to exactly d switches in middle stage 130+10*(Log_d N-1) through d links.

[0667] Similarly each of the N/d middle switches

$$MS(2 \times \text{Log}_d N - 3, 1) - MS\left(2 \times \text{Log}_d N - 3, \frac{N}{d}\right)$$

in the middle stage 130+10*(2*Log_d N-4) are connected from exactly d switches in middle stage 130+10*(2*Log_d N-5) through d links and also are connected to exactly d output switches in output stage 120 through d links.

[0668] Each of the N/d output switches OS1-OS(N/d) are connected from exactly d switches in middle stage 130+10*(2*Log_d N-4) through d links.

[0669] The general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=1$ according to the current invention.

Asymmetric Folded RNB ($N_2 > N_1$) Unicast Embodiments:

[0670] Referring to FIG. 5C, an exemplary symmetrical folded multi-stage network 500C respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, and 150 is shown where input stage 110 consists of four, two by two switches IS1-IS4 and output stage 120 consists of four, six by six switches OS1-OS4. And all the middle stages namely middle stage 130 consists of four, two by two switches MS(1,1)-MS(1,4), middle stage 140 consists of four, two by two switches MS(2,1)-MS(2,4), and middle stage 150 consists of four, two by six switches MS(3,1)-MS(3,4).

[0671] Such networks can be operated in rearrangeably nonblocking manner for unicast connections, because the switches in the input stage 110 are of size two by two, the switches in output stage 120 are of size six by six, and there are four switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0672] The connection topology of the network 500C shown in FIG. 5C is known to be back to back inverse Benes connection topology. The connection topology of the networks 500C is different in the other embodiments. That is the way the links ML(1,1)-ML(1,8), ML(2,1)-ML(2,8), ML(3,1)-ML(3,8), and ML(4,1)-ML(4,8) are connected between the respective stages is different.

[0673] Even though only one embodiment is illustrated, in general, the network $V_{fold}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold}(N_1, N_2, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{fold}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{fold}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 5C is only one example of network $V_{fold}(N_1, N_2, d, s)$.

[0674] The networks 500C of FIG. 5C is also rearrangeably nonblocking for unicast according to the current invention. In one embodiment of these networks each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

US 2011/0044329 A1

Feb. 24, 2011

70

$$\frac{N_1}{d}$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d \times d$ and each output switch OS1-OS4 can be denoted in general with the notation $d_2 \times d_2$, where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $d \times d$. The size of each switch in the last middle stage can be denoted as $d \times d_2$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric folded multi-stage network can be represented with the notation $V_{fold}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0675] In network 500C of FIG. 5C, each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly d switches in middle stage 130 through d links (for example input switch IS1 is connected to middle switches MS(1,1) and MS(1,2) through the links ML(1,1) and ML(1,2) respectively).

[0676] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,4) are connected to the middle switch MS(1,1) from input switch IS1 and IS2 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

[0677] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,6)

are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,3) respectively).

[0678] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,6) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage 120 through d_2 links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 from middle switch MS(3,1); the links ML(4,3) and ML(4,4) are connected to output switches OS2 from middle switch MS(3,1); the link ML(4,5) is connected to output switches OS3 from middle switch MS(3,1); and the link ML(4,6) is connected to output switches OS4 from middle switch MS(3,1)).

[0679] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through d_2 links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); output switch OS1 is connected from middle switch MS(3,2) through the links ML(4,7) and ML(4,8); output switch OS1 is connected from middle switch MS(3,3) through the link ML(4,13); and output switch OS1 is connected from middle switch MS(3,4) through the links ML(4,19)).

Generalized Asymmetric Folded RNB ($N_2 > N_1$) Unicast Embodiments:

[0680] Network 500D of FIG. 5D is an example of general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages where $N_2 > N_1$ and $N_2 = p \times N$, where $p > 1$. In network 500D of FIG. 5D, $N_1 = N$ and $N_2 = p \times N$. The general symmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention (and in the example of FIG. 5D, $s = 1$). The general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of

$$\frac{N_1}{d}$$

US 2011/0044329 A1

Feb. 24, 2011

71

input switches IS1-IS(N_1/d) (for example the links IL1-IL(d) to the input switch IS1) and d outgoing links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links ML(1,1)-ML(1, d) to the input switch IS1). There are d_2 (where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d$$

outlet links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example the links OL1-OL($p \times d$) to the output switch OS1) and d_2 ($=p \times d$) incoming links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example ML($2 \times \text{Log}_d N_1 - 2, 1$)-ML($2 \times \text{Log}_d N_1 - 2, d_2$) to the output switch OS1).
[0681] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) are connected to exactly d switches in middle stage **130** through d links.
[0682] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1, N_1/d) in the middle stage **130** are connected from exactly d input switches through d links and also are connected to exactly d switches in middle stage **140** through d links.
[0683] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(\text{Log}_d N_1 - 1, 1) - MS\left(\text{Log}_d N_1 - 1, \frac{N_1}{d}\right)$$

in the middle stage $\mathbf{130} + 10^* (\text{Log}_d N_1 - 2)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10^* (\text{Log}_d N_1 - 3)$

through d links and also are connected to exactly d switches in middle stage $\mathbf{130} + 10^* (\text{Log}_d N_1 - 1)$ through d links.

[0684] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(2 \times \text{Log}_d N_1 - 3, 1) - MS\left(2 \times \text{Log}_d N_1 - 3, \frac{N_1}{d}\right)$$

in the middle stage $\mathbf{130} + 10^* (2^* \text{Log}_d N_1 - 4)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10^* (2^* \text{Log}_d N_1 - 5)$ through d links and also are connected to exactly d output switches in output stage **120** through d_1 links.

[0685] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) are connected from exactly d switches in middle stage $\mathbf{130} + 10^* (2^* \text{Log}_d N_1 - 4)$ through d_2 links.

[0686] The general symmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when 1 according to the current invention.

Asymmetric Folded RNB ($N_1 > N_2$) Unicast Embodiments:

[0687] Referring to FIG. 5E, an exemplary symmetrical folded multi-stage network **500E** with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, six by six switches IS1-IS4 and output stage **120** consists of four, two by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, six by two switches MS(1,1)-MS(1,4), middle stage **140** consists of four, two by two switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, two by two switches MS(3,1)-MS(3,4).

[0688] Such a network can be operated in rearrangeably nonblocking manner for unicast connections, because the switches in the input stage **110** are of size six by six, the switches in output stage **120** are of size two by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0689] The connection topology of the network **500E** shown in FIG. 5E is known to be back to back inverse Benes connection topology. The connection topology of the networks **500E** is different in the other embodiments. That is the way the links ML(1,1)-ML(1,8), ML(2,1)-ML(2,8), ML(3,1)-ML(3,8), and ML(4,1)-ML(4,8) are connected between the respective stages is different.

[0690] Even though only one embodiment is illustrated, in general, the network $V_{fold}(N_1, N_2, d, s)$, comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold}(N_1, N_2, d, s)$ may be back

US 2011/0044329 A1

Feb. 24, 2011

72

to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{fold}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{fold}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 5E is only one example of network $V_{fold}(N_1, N_2, d, s)$.

[0691] The network 500E is rearrangeably nonblocking for unicast according to the current invention. In one embodiment of these networks each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_2}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d_1 * d_1$ and each output switch OS1-OS4 can be denoted in general with the notation $(d * d)$, where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $d * d$. The size of each switch in the first middle stage can be denoted as $d_1 * d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric folded multi-stage network can be represented with the notation $V_{fold}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each output switch where $N_1 > N_2$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0692] In network 500E of FIG. 5E, each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly d switches in middle stage 130 through d_1 links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1) and ML(1,2); input switch IS1 is connected to middle switch MS(1,2) through the links ML(1,3) and ML(1,4); input switch IS1 is connected to middle switch MS(1,3)

through the link ML(1,5); and input switch IS1 is connected to middle switch MS(1,4) through the links ML(1,6)).

[0693] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d , input switches through d links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1; the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2; the link ML(1,13) is connected to the middle switch MS(1,1) from input switch IS3; and the link ML(1,19) is connected to the middle switch MS(1,1) from input switch IS4), and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

[0694] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,6) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,3) respectively).

[0695] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,6) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage 120 through d_1 links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switch MS(3,1)).

[0696] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through d_1 links (for example output switch OS1 is connected from middle switches MS(3,1) and MS(3,2) through the links ML(4,1) and ML(4,4) respectively).

US 2011/0044329 A1

Feb. 24, 2011

73

Generalized Asymmetric Folded RNB ($N_1 > N_2$) Unicast Embodiments:

[0697] Network **500F** of FIG. **5F** is an example of general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network **500F** of FIG. **5F**, $N_2 = N$ and $N_1 = p * N$. The general symmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s=1$ according to the current invention (and in the example of FIG. **5F**, $s=1$). The general asymmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages has d_1 (where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d$$

inlet links for each of

$$\frac{N_2}{d}$$

input switches $IS1-IS(N_2/d)$ (for example the links $IL1-IL(p*d)$ to the input switch $IS1$) and $d_1 (=p*d)$ outgoing links for each of

$$\frac{N_2}{d}$$

input switches $IS1-IS(N_2/d)$ (for example the links $ML(1,1)-ML(1,(d+p*d))$ to the input switch $IS1$). There are d outlet links for each of

$$\frac{N_2}{d}$$

output switches $OS1-OS(N_2/d)$ (for example the links $OL1-OL(d)$ to the output switch $OS1$) and d incoming links for each of

$$\frac{N_2}{d}$$

output switches $OS1-OS(N_2/d)$ (for example $ML(2 \times \log_d N_2 - 2, 1) - ML(2 \times \log_d N_2 - 2, d)$ to the output switch $OS1$).

[0698] Each of the

$$\frac{N_2}{d}$$

input switches $IS1-IS(N_2/d)$ are connected to exactly d switches in middle stage **130** through d_1 links.

[0699] Each of the

$$\frac{N_2}{d}$$

middle switches $MS(1,1)-MS(1,N_2/d)$ in the middle stage **130** are connected from exactly d input switches through d_1 links and also are connected to exactly d switches in middle stage **140** through d links.

[0700] Similarly each of the

$$\frac{N_2}{d}$$

middle switches

$$MS(\log_d N_2 - 1, 1) - MS(\log_d N_2 - 1, \frac{N_2}{d})$$

in the middle stage $\mathbf{130} + 10 * (\log_d N_2 - 2)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10 * (\log_d N_2 - 3)$ through d links and also are connected to exactly d switches in middle stage $\mathbf{130} + 10 * (\log_d N_2 - 1)$ through d links.

[0701] Similarly each of the

$$\frac{N_2}{d}$$

middle switches

$$MS(2 \times \log_d N_2 - 3, 1) - MS(2 \times \log_d N_2 - 3, \frac{N_2}{d})$$

in the middle stage $\mathbf{130} + 10 * (2 * \log_d N_2 - 4)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10 * (2 * \log_d N_2 - 5)$ through d links and also are connected to exactly d output switches in output stage **120** through d links.

[0702] Each of the

$$\frac{N_2}{d}$$

output switches $OS1-OS(N_2/d)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10 * (2 * \log_d N_2 - 4)$ through d links.

[0703] The general symmetrical folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention.

Symmetric RNB Unicast Embodiments:

[0704] Referring to FIG. **6A**, FIG. **6B**, FIG. **6C**, FIG. **6D**, FIG. **6E**, FIG. **6F**, FIG. **6G**, FIG. **600H**, FIG. **600I** and FIG. **6J** with exemplary symmetrical multi-stage networks **600A**, **600B**, **600C**, **600D**, **600E**, **600F**, **600G**, **600H**, **600I**, and **600J** respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call

US 2011/0044329 A1

Feb. 24, 2011

74

or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by two switches IS1-IS4 and output stage **120** consists of four, two by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, two by two switches MS(1,1)-MS(1,4), middle stage **140** consists of four, two by two switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, two by two switches MS(3,1)-MS(3,4).

[0705] Such networks can be operated in rearrangeably nonblocking manner for unicast connections, because the switches in the input stage **110** are of size two by two, the switches in output stage **120** are of size two by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0706] In all the ten embodiments of FIG. 6A to FIG. 6J the connection topology is different. That is the way the links ML(1,1)-ML(1,8), ML(2,1)-ML(2,8), ML(3,1)-ML(3,8), and ML(4,1)-ML(4,8) are connected between the respective stages is different. For example, the connection topology of the network **600A** shown in FIG. 6A is known to be back to back inverse Benes connection topology; the connection topology of the network **600B** shown in FIG. 6B is known to be back to back Omega connection topology; and the connection topology of the network **600C** shown in FIG. 6C is hereinafter called nearest neighbor connection topology.

[0707] Even though only ten embodiments are illustrated, in general, the network $V(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V(N, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V(N, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V(N, d, s)$ can be built. The ten embodiments of FIG. 6A to FIG. 6J are only three examples of network $V(N, d, s)$.

[0708] The networks **600A-600J** of FIG. 6A-FIG. 6J are also rearrangeably nonblocking for unicast according to the current invention. In one embodiment of these networks each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch IS1-IS4 can be denoted in general with the notation $d*d$ and each output switch OS1-OS4 can be denoted in general with the notation $d*d$. Likewise, the size of each switch in any of the middle stages can be denoted as $d*d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-stage network can be represented with the notation $V(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL8), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch. Although it is not necessary that there be the same number of inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

[0709] In network **600A** of FIG. 6A, each of the N/d input switches IS1-IS4 are connected to exactly d switches in middle stage **130** through d links (for example input switch IS1 is connected to middle switches MS(1,1) and MS(1,2) through the links ML(1,1) and ML(1,2) respectively).

[0710] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage **130** are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,4) are connected to the middle switch MS(1,1) from input switch IS1 and IS2 respectively) and also are connected to exactly d switches in middle stage **140** through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

[0711] Similarly each of the N/d middle switches MS(2,1)-MS(2,4) in the middle stage **140** are connected from exactly d switches in middle stage **130** through d links (for example the links ML(2,1) and ML(2,6) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage **150** through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,3) respectively).

[0712] Similarly each of the N/d middle switches MS(3,1)-MS(3,4) in the middle stage **150** are connected from exactly d switches in middle stage **140** through d links (for example the links ML(3,1) and ML(3,6) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage **120** through d links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switch MS(3,1)).

[0713] Each of the N/d output switches OS1-OS4 are connected from exactly d switches in middle stage **150** through d links (for example output switch OS1 is connected from middle switches MS(3,1) and MS(3,2) through the links ML(4,1) and ML(4,4) respectively).

Generalized Symmetric RNB Unicast Embodiments:

[0714] Network **600K** of FIG. 6K is an example of general symmetrical multi-stage network $V(N, d, s)$ with $(2 \times \log_d N) - 1$ stages. The general symmetrical multi-stage network $V(N, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention (and in the example of FIG. 6K, $s=1$). The general symmetrical multi-stage network $V(N, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of N/d input switches IS1-IS (N/d) (for example the links IL1-IL(d) to the input switch IS1) and d outgoing links for each of N/d input switches IS1-IS (N/d) (for example the links ML(1,1)-ML(1, d) to the input switch IS1). There are d outlet links for each of N/d output switches OS1-OS(N/d) (for example OL1-OL(d) to the output switch OS1) and d incoming links for each of N/d output switches OS1-OS(N/d) (for example ML($2 \times \log_d N - 2, 1$)-ML($2 \times \log_d N - 2, d$) to the output switch OS1).

[0715] Each of the N/d input switches IS1-IS(N/d) are connected to exactly d switches in middle stage **130** through d links.

[0716] Each of the N/d middle switches MS(1,1)-MS(1, N/d) in the middle stage **130** are connected from exactly d input switches through d links and also are connected to exactly d switches in middle stage **140** through d links.

US 2011/0044329 A1

Feb. 24, 2011

75

[0717] Similarly each of the N/d middle switches

$$MS(\text{Log}_d N - 1, 1) - MS\left(\text{Log}_d N - 1, \frac{N}{d}\right)$$

in the middle stage $130+10^*(\text{Log}_d N-2)$ are connected from exactly d switches in middle stage $130+10^*(\text{Log}_d N-3)$ through d links and also are connected to exactly d switches in middle stage $130+10^*(\text{Log}_d N-1)$ through d links.

[0718] Similarly each of the N/d middle switches

$$MS(2 \times \text{Log}_d N - 3, 1) - MS\left(2 \times \text{Log}_d N - 3, \frac{N}{d}\right)$$

in the middle stage $130+10^*(2*\text{Log}_d N-4)$ are connected from exactly d switches in middle stage $130+10^*(2*\text{Log}_d N-5)$ through d links and also are connected to exactly d output switches in output stage **120** through d links.

[0719] Each of the N/d output switches OS1-OS(N/d) are connected from exactly d switches in middle stage $130+10^*(2*\text{Log}_d N-4)$ through d links.

[0720] The general symmetrical multi-stage network $V(N, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention.

Asymmetric RNB ($N_2 > N_1$) Unicast Embodiments:

[0721] Referring to FIG. 6A1, FIG. 6B1, FIG. 6C1, FIG. 6D1, FIG. 6E1, FIG. 6F1, FIG. 6G1, FIG. 600H1, FIG. 600H1 and FIG. 6J1 with exemplary symmetrical multi-stage networks **600A1**, **600B1**, **600C1**, **600D1**, **600E1**, **600F1**, **600G1**, **600H1**, **600I1**, and **600J1** respectively with five stages of twenty switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, two by two switches IS1-IS4 and output stage **120** consists of four, six by six switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, two by two switches MS(1,1)-MS(1,4), middle stage **140** consists of four, two by two switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, two by six switches MS(3,1)-MS(3,4).

[0722] Such networks can be operated in rearrangeably nonblocking manner for unicast connections, because the switches in the input stage **110** are of size two by two, the switches in output stage **120** are of size six by six, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0723] In all the ten embodiments of FIG. 6A1 to FIG. 6J1 the connection topology is different. That is the way the links ML(1,1)-ML(1,8), ML(2,1)-ML(2,8), ML(3,1)-ML(3,8), and ML(4,1)-ML(4,8) are connected between the respective stages is different. For example, the connection topology of the network **600A1** shown in FIG. 6A1 is known to be back to back inverse Benes connection topology; the connection topology of the network **600B1** shown in FIG. 6B1 is known to be back to back Omega connection topology; and the connection topology of the network **600C1** shown in FIG. 6C1 is hereinafter called nearest neighbor connection topology.

[0724] Even though only ten embodiments are illustrated, in general, the network $V(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V(N_1, N_2, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V(N_1, N_2, d, s)$ can be built. The ten embodiments of FIG. 6A1 to FIG. 6J1 are only three examples of network $V(N_1, N_2, d, s)$.

[0725] The networks **600A1-600J1** of FIG. 6A1-FIG. 6J1 are also rearrangeably nonblocking for unicast according to the current invention. In one embodiment of these networks each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_1}{d},$$

where N_1 is the total number of inlet links or and N_2 is the total number of outlet links and $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. The number of middle switches in each middle stage is denoted by

$$\frac{N_1}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation $d*d$ and each output switch OS1-OS4 can be denoted in general with the notation d_2*d_2 , where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d.$$

The size of each switch in any of the middle stages excepting the last middle stage can be denoted as $d*d$. The size of each switch in the last middle stage can be denoted as $d*d_2$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-stage network can be represented with the notation $V(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL8), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL24), d represents the inlet links of each input switch where $N_2 > N_1$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0726] In network **600A1** of FIG. 6A1, each of the

$$\frac{N_1}{d}$$

input switches IS1-IS4 are connected to exactly d switches in middle stage **130** through d links (for example input switch

US 2011/0044329 A1

Feb. 24, 2011

76

IS1 is connected to middle switches MS(1,1) and MS(1,2) through the links ML(1,1) and ML(1,2) respectively).

[0727] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,4) are connected to the middle switch MS(1,1) from input switch IS1 and IS2 respectively) and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

[0728] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,6) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,3) respectively).

[0729] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,6) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage 120 through d_2 links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 from middle switch MS(3,1); the links ML(4,3) and ML(4,4) are connected to output switches OS2 from middle switch MS(3,1); the link ML(4,5) is connected to output switches OS3 from middle switch MS(3,1); and the link ML(4,6) is connected to output switches OS4 from middle switch MS(3,1)).

[0730] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through d_2 links (for example output switch OS1 is connected from middle switch MS(3,1) through the links ML(4,1) and ML(4,2); output switch OS1 is connected from middle switch MS(3,2) through the links ML(4,7) and ML(4,8); output switch OS1 is connected from

middle switch MS(3,3) through the link ML(4,13); and output switch OS1 is connected from middle switch MS(3,4) through the links ML(4,19)).

Generalized Asymmetric RNB ($N_2 > N_1$) Unicast Embodiments:

[0731] Network 600K1 of FIG. 6K1 is an example of general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages where $N_2 > N_1$ and $N_2 = p * N$, where $p > 1$. In network 400K1 of FIG. 4K1, $N_1 = N$ and $N_2 = p * N$. The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention (and in the example of FIG. 6K1, $s = 1$). The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links IL1-IL(d) to the input switch IS1) and d outgoing links for each of

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) (for example the links ML(1,1)-ML(1, d) to the input switch IS1). There are d_2 (where

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d)$$

outlet links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example the links OL1-OL($p * d$) to the output switch OS1) and d_2 ($= p * d$) incoming links for each of

$$\frac{N_1}{d}$$

output switches OS1-OS(N_1/d) (for example ML($2 \times \log_d N_1 - 2, 1$)-ML($2 \times \log_d N_1 - 2, d_2$) to the output switch OS1).

[0732] Each of the

$$\frac{N_1}{d}$$

input switches IS1-IS(N_1/d) are connected to exactly d switches in middle stage 130 through d links.

US 2011/0044329 A1

Feb. 24, 2011

77

[0733] Each of the

$$\frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,N₁/d) in the middle stage **130** are connected from exactly d input switches through d links and also are connected to exactly d switches in middle stage **140** through d links.

[0734] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(\text{Log}_d N_1 - 1, 1) - MS\left(\text{Log}_d N_1 - 1, \frac{N_1}{d}\right)$$

in the middle stage **130**+10*(Log_d N₁-2) are connected from exactly d switches in middle stage **130**+10*(Log_d N₁-3) through d links and also are connected to exactly d switches in middle stage **130**+10*(Log_d N₁-1) through d links.

[0735] Similarly each of the

$$\frac{N_1}{d}$$

middle switches

$$MS(2 \times \text{Log}_d N_1 - 3, 1) - MS\left(2 \times \text{Log}_d N_1 - 3, \frac{N_1}{d}\right)$$

in the middle stage **130**+10*(2*Log_d N₁-4) are connected from exactly d switches in middle stage **130**+10*(2*Log_d N₁-5) through d links and also are connected to exactly d output switches in output stage **120** through d₁ links.

[0736] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS(N₁/d) are connected from exactly d switches in middle stage **130**+10*(2*Log_d N-4) through d₂ links.

[0737] The general symmetrical multi-stage network V(N₁, N₂, d, s) can be operated in rearrangeably nonblocking manner for multicast when s=1 according to the current invention.

Asymmetric RNB (N₁>N₂) Unicast Embodiments:

[0738] Referring to FIG. 6A2, FIG. 6B2, FIG. 6C2, FIG. 6D2, FIG. 6E2, FIG. 6F2, FIG. 6G2, FIG. 600H2, FIG. 600I2 and FIG. 6J2 with exemplary symmetrical multi-stage networks **600A2**, **600B2**, **600C2**, **600D2**, **600E2**, **600F2**, **600G2**, **600H2**, **600I2**, and **600J2** respectively with five stages of twenty switches for satisfying communication requests, such

as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage **110** and output stage **120** via middle stages **130**, **140**, and **150** is shown where input stage **110** consists of four, six by six switches IS1-IS4 and output stage **120** consists of four, two by two switches OS1-OS4. And all the middle stages namely middle stage **130** consists of four, six by two switches MS(1, 1)-MS(1,4), middle stage **140** consists of four, two by two switches MS(2,1)-MS(2,4), and middle stage **150** consists of four, two by two switches MS(3,1)-MS(3,4).

[0739] Such networks can be operated in rearrangeably nonblocking manner for unicast connections, because the switches in the input stage **110** are of size six by six, the switches in output stage **120** are of size two by two, and there are four switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0740] In all the ten embodiments of FIG. 6A2 to FIG. 6J2 the connection topology is different. That is the way the links ML(1,1)-ML(1,8), ML(2,1)-ML(2,8), ML(3,1)-ML(3,8), and ML(4,1)-ML(4,8) are connected between the respective stages is different. For example, the connection topology of the network **600A2** shown in FIG. 6A2 is known to be back to back inverse Benes connection topology; the connection topology of the network **600B2** shown in FIG. 6B2 is known to be back to back Omega connection topology; and the connection topology of the network **600C2** shown in FIG. 6C2 is hereinafter called nearest neighbor connection topology.

[0741] Even though only ten embodiments are illustrated, in general, the network V(N₁, N₂, d, s) can comprise any arbitrary type of connection topology. For example the connection topology of the network V(N₁, N₂, d, s) may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the V(N₁, N₂, d, s) network is, when no connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network V(N₁, N₂, d, s) can be built. The ten embodiments of FIG. 6A2 to FIG. 6J2 are only three examples of network V(N₁, N₂, d, s).

[0742] The networks **600A2**-**600J2** of FIG. 6A2-FIG. 6J2 are also rearrangeably nonblocking for unicast according to the current invention. In one embodiment of these networks each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage **110** and of output stage **120** can be denoted in general with the variable

$$\frac{N_2}{d},$$

where N₁ is the total number of inlet links or and N₂ is the total number of outlet links and N₁>N₂ and N₁=p*N₂ where p>1. The number of middle switches in each middle stage is denoted by

$$\frac{N_2}{d}.$$

The size of each input switch IS1-IS4 can be denoted in general with the notation d₁*d₁ and each output switch OS1-OS4 can be denoted in general with the notation (d*d), where

US 2011/0044329 A1

Feb. 24, 2011

78

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d.$$

The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $d \times d$. The size of each switch in the first middle stage can be denoted as $d_1 \times d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. An asymmetric multi-stage network can be represented with the notation $V(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all input switches (for example the links IL1-IL24), N_2 represents the total number of outlet links of all output switches (for example the links OL1-OL8), d represents the inlet links of each output switch where $N_1 > N_2$, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0743] In network 600A2 of FIG. 6A2, each of the

$$\frac{N_2}{d}$$

input switches IS1-IS4 are connected to exactly d switches in middle stage 130 through d_1 links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1) and ML(1,2); input switch IS1 is connected to middle switch MS(1,2) through the links ML(1,3) and ML(1,4); input switch IS1 is connected to middle switch MS(1,3) through the link ML(1,5); and input switch IS1 is connected to middle switch MS(1,4) through the links ML(1,6)).

[0744] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input switches through d links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1; the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2; the link ML(1,13) is connected to the middle switch MS(1,1) from input switch IS3; and the link ML(1,19) is connected to the middle switch MS(1,1) from input switch IS4), and also are connected to exactly d switches in middle stage 140 through d links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

[0745] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,4) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links (for example the links ML(2,1) and ML(2,6) are connected to the middle switch MS(2,1) from middle switches MS(1,1) and MS(1,3) respectively) and also are

connected to exactly d switches in middle stage 150 through d links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1) and MS(3,3) respectively).

[0746] Similarly each of the

$$\frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,4) in the middle stage 150 are connected from exactly d switches in middle stage 140 through d links (for example the links ML(3,1) and ML(3,6) are connected to the middle switch MS(3,1) from middle switches MS(2,1) and MS(2,3) respectively) and also are connected to exactly d output switches in output stage 120 through d_1 links (for example the links ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switch MS(3,1)).

[0747] Each of the

$$\frac{N_1}{d}$$

output switches OS1-OS4 are connected from exactly d switches in middle stage 150 through d_2 links (for example output switch OS1 is connected from middle switches MS(3,1) and MS(3,2) through the links ML(4,1) and ML(4,4) respectively).

Generalized Asymmetric RNB ($N_1 > N_2$) Unicast Embodiments:

[0748] Network 600K2 of FIG. 6K2 is an example of general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages where $N_1 > N_2$ and $N_1 = p \times N_2$ where $p > 1$. In network 400K2 of FIG. 4K2, $N_2 = N$ and $N_1 = p \times N$. The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention (and in the example of FIG. 6K2, $s = 1$). The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages has d_1 (where

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d$$

inlet links for each of

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) (for example the links IL1-IL($p \times d$) to the input switch IS1) and d , ($= p \times d$) outgoing links for each of

$$\frac{N_2}{d}$$

US 2011/0044329 A1

Feb. 24, 2011

79

input switches IS1-IS(N_2/d) (for example the links ML(1,1)-ML(1,($d+p*d$)) to the input switch IS1). There are d outlet links for each of

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example the links OL1-OL(d) to the output switch OS1) and d incoming links for each of

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) (for example ML($2 \times \text{Log}_d N_2 - 2, 1$)-ML($2 \times \text{Log}_d N_2 - 2, d$) to the output switch OS1).

[0749] Each of the

$$\frac{N_2}{d}$$

input switches IS1-IS(N_2/d) are connected to exactly d switches in middle stage **130** through d_1 links.

[0750] Each of the

$$\frac{N_2}{d}$$

middle switches MS(1,1)-MS(1, N_2/d) in the middle stage **130** are connected from exactly d input switches through d_1 links and also are connected to exactly d switches in middle stage **140** through d links.

[0751] Similarly each of the

$$\frac{N_2}{d}$$

middle switches

$$MS(\text{Log}_d N_2 - 1, 1) - MS\left(\text{Log}_d N_2 - 1, \frac{N_2}{d}\right)$$

[0752] in the middle stage $\mathbf{130} + 10^*(\text{Log}_d N_1 - 2)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10^*(\text{Log}_d N_1 - 3)$ through d links and also are connected to exactly d switches in middle stage $\mathbf{130} + 10^*(\text{Log}_d N_1 - 1)$ through d links.

[0753] Similarly each of the

$$\frac{N_2}{d}$$

middle switches

$$MS(2 \times \text{Log}_d N_2 - 3, 1) - MS\left(2 \times \text{Log}_d N_2 - 3, \frac{N_2}{d}\right)$$

in the middle stage $\mathbf{130} + 10^*(2 * \text{Log}_d N_1 - 4)$ are connected from exactly d switches in middle stage $\mathbf{130} + 10^*(2 * \text{Log}_d N_1 - 5)$ through d links and also are connected to exactly d output switches in output stage **120** through d links.

[0754] Each of the

$$\frac{N_2}{d}$$

output switches OS1-OS(N_2/d) are connected from exactly d switches in middle stage $\mathbf{130} + 10^*(2 * \text{Log}_d N_1 - 4)$ through d links.

[0755] The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention.

Scheduling Method Embodiments:

[0756] FIG. 7A shows a high-level flowchart of a scheduling method **1000**, in one embodiment executed to setup multicast and unicast connections in network **100A** of FIG. 1A (or any of the networks $V_{\text{mlink}}(N_1, N_2, d, s)$ and the networks $V(N_1, N_2, d, s)$ disclosed in this invention). According to this embodiment, a multicast connection request is received in act **1010**. Then the control goes to act **1020**.

[0757] In act **1020**, based on the inlet link and input switch of the multicast connection received in act **1010**, from each available outgoing middle link of the input switch of the multicast connection, by traveling forward from middle stage **130** to middle stage $\mathbf{130} + 10^*(\text{Log}_d N - 2)$, the lists of all reachable middle switches in each middle stage are derived recursively. That is, first, by following each available outgoing middle link of the input switch all the reachable middle switches in middle stage **130** are derived. Next, starting from the selected middle switches in middle stage **130** traveling through all of their available outgoing middle links to middle stage **140** all the available middle switches in middle stage **140** are derived. This process is repeated recursively until all the reachable middle switches, starting from the outgoing middle link of input switch, in middle stage $\mathbf{130} + 10^*(\text{Log}_d N - 2)$ are derived. This process is repeated for each available outgoing middle link from the input switch of the multicast connection and separate reachable lists are derived in each middle stage from middle stage **130** to middle stage $\mathbf{130} + 10^*(\text{Log}_d N - 2)$ for all the available outgoing middle links from the input switch. Then the control goes to act **1030**.

[0758] In act **1030**, based on the destinations of the multicast connection received in act **1010**, from the output switch of each destination, by traveling backward from output stage **120** to middle stage $\mathbf{130} + 10^*(\text{Log}_d N - 2)$, the lists of all middle switches in each middle stage from which each destination output switch (and hence the destination outlet links) is reachable, are derived recursively. That is, first, by following each available incoming middle link of the output switch of each destination link of the multicast connection, all the middle switches in middle stage $\mathbf{130} + 10^*(2 * \text{Log}_d N - 4)$ from which the output switch is reachable, are derived. Next, starting from the selected middle switches in middle stage **130**+

US 2011/0044329 A1

Feb. 24, 2011

80

$10^{*}(2^{*}\text{Log}_d N-4)$ traveling backward through all of their available incoming middle links from middle stage $130+10^{*}(2^{*}\text{Log}_d N-5)$ all the available middle switches in middle stage $130+10^{*}(2^{*}\text{Log}_d N-5)$ from which the output switch is reachable, are derived. This process is repeated recursively until all the middle switches in middle stage $130+10^{*}(\text{Log}_d N-2)$ from which the output switch is reachable, are derived. This process is repeated for each output switch of each destination link of the multicast connection and separate lists in each middle stage from middle stage $130+10^{*}(2^{*}\text{Log}_d N-4)$ to middle stage $130+10^{*}(\text{Log}_d N-2)$ for all the output switches of each destination link of the connection are derived. Then the control goes to act **1040**.

[0759] In act **1040**, using the lists generated in acts **1020** and **1030**, particularly list of middle switches derived in middle stage $130+10^{*}(\text{Log}_d N-2)$ corresponding to each outgoing link of the input switch of the multicast connection, and the list of middle switches derived in middle stage $130+10^{*}(\text{Log}_d N-2)$ corresponding to each output switch of the destination links, the list of all the reachable destination links from each outgoing link of the input switch are derived. Specifically if a middle switch in middle stage $130+10^{*}(\text{Log}_d N-2)$ is reachable from an outgoing link of the input switch, say “x”, and also from the same middle switch in middle stage $130+10^{*}(\text{Log}_d N-2)$ if the output switch of a destination link, say “y”, is reachable then using the outgoing link of the input switch x, destination link y is reachable. Accordingly, the list of all the reachable destination links from each outgoing link of the input switch is derived. The control then goes to act **1050**.

[0760] In act **1050**, among all the outgoing links of the input switch, it is checked if all the destinations are reachable using only one outgoing link of the input switch. If one outgoing link is available through which all the destinations of the multicast connection are reachable (i.e., act **1050** results in “yes”), the control goes to act **1070**. And in act **1070**, the multicast connection is setup by traversing from the selected only one outgoing middle link of the input switch in act **1050**, to all the destinations. Then the control transfers to act **1090**.

[0761] If act **1050** results “no”, that is one outgoing link is not available through which all the destinations of the multicast connection are reachable, then the control goes to act **1060**. In act **1060**, it is checked if all destination links of the multicast connection are reachable using two outgoing middle links from the input switch. According to the current invention, it is always possible to find at most two outgoing middle links from the input switch through which all the destinations of a multicast connection are reachable. So act **1060** always results in “yes”, and then the control transfers to act **1080**. In act **1080**, the multicast connection is setup by traversing from the selected only two outgoing middle links of the input switch in act **1060**, to all the destinations. Then the control transfers to act **1090**.

[0762] In act **1090**, all the middle links between any two stages of the network used to setup the connection in either act **1070** or act **1080** are marked unavailable so that these middle links will be made unavailable to other multicast connections. The control then returns to act **1010**, so that acts **1010**, **1020**, **1030**, **1040**, **1050**, **1060**, **1070**, **1080**, and **1090** are executed in a loop, for each connection request until the connections are set up.

[0763] In the example illustrated in FIG. 1A, four outgoing middle links are available to satisfy a multicast connection request if input switch is IS2, but only at most two outgoing

middle links of the input switch will be used in accordance with this method. Similarly, although three outgoing middle links is available for a multicast connection request if the input switch is IS1, again only at most two outgoing middle links is used. The specific outgoing middle links of the input switch that are chosen when selecting two outgoing middle links of the input switch is irrelevant to the method of FIG. 7A so long as at most two outgoing middle links of the input switch are selected to ensure that the connection request is satisfied, i.e. the destination switches identified by the connection request can be reached from the outgoing middle links of the input switch that are selected. In essence, limiting the outgoing middle links of the input switch to no more than two permits the network $V_{\text{mlink}}(N_1, N_2, d, s)$ and the network $V(N_1, N_2, d, s)$ to be operated in nonblocking manner in accordance with the invention.

[0764] According to the current invention, using the method **1040** of FIG. 7A, the network $V_{\text{mlink}}(N_1, N_2, d, s)$ and the networks $V(N_1, N_2, d, s)$ are operated in rearrangeably nonblocking for unicast connections when $s \geq 1$, are operated in strictly nonblocking for unicast connections when $s \geq 2$, and are operated in rearrangeably nonblocking for multicast connections when $s \geq 2$.

[0765] The connection request of the type described above in reference to method **1000** of FIG. 7A can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, only one outgoing middle link of the input switch is used to satisfy the request. Moreover, in method **1000** described above in reference to FIG. 7A any number of middle links may be used between any two stages excepting between the input stage and middle stage **130**, and also any arbitrary fan-out may be used within each output stage switch, to satisfy the connection request.

[0766] As noted above method **1000** of FIG. 7A can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks $V_{\text{mlink}}(N, d, s)$, $V_{\text{mlink}}(N_1, N_2, d, s)$, $V(N, d, s)$ and $V(N_1, N_2, d, s)$ disclosed in this invention.

Applications Embodiments:

[0767] All the embodiments disclosed in the current invention are useful in many varieties of applications. FIG. 8A1 illustrates the diagram of **800A1** which is a typical two by two switch with two inlet links namely IL1 and IL2, and two outlet links namely OL1 and OL2. The two by two switch also implements four crosspoints namely CP(1,1), CP(1,2), CP(2,1) and CP(2,2) as illustrated in FIG. 8A1. For example the diagram of **800A1** may be the implementation of middle switch MS(1,1) of the diagram **400A** of FIG. 4A where inlet link IL1 of diagram **800A1** corresponds to middle link ML(1,1) of diagram **400A**, inlet link IL2 of diagram **800A1** corresponds to middle link ML(1,5) of diagram **400A**, outlet link OL1 of diagram **800A1** corresponds to middle link ML(2,1) of diagram **400A**, outlet link OL2 of diagram **800A1** corresponds to middle link ML(2,2) of diagram **400A**.

1) Programmable Integrated Circuit Embodiments:

[0768] All the embodiments disclosed in the current invention are useful in programmable integrated circuit applications. FIG. 8A2 illustrates the detailed diagram **800A2** for the implementation of the diagram **800A1** in programmable integrated circuit embodiments. Each crosspoint is implemented

US 2011/0044329 A1

Feb. 24, 2011

81

by a transistor coupled between the corresponding inlet link and outlet link, and a programmable cell in programmable integrated circuit embodiments. Specifically crosspoint CP(1,1) is implemented by transistor C(1,1) coupled between inlet link IL1 and outlet link OL1, and programmable cell P(1,1); crosspoint CP(1,2) is implemented by transistor C(1,2) coupled between inlet link IL1 and outlet link OL2, and programmable cell P(1,2); crosspoint CP(2,1) is implemented by transistor C(2,1) coupled between inlet link IL2 and outlet link OL1, and programmable cell P(2,1); and crosspoint CP(2,2) is implemented by transistor C(2,2) coupled between inlet link IL2 and outlet link OL2, and programmable cell P(2,2).

[0769] If the programmable cell is programmed ON, the corresponding transistor couples the corresponding inlet link and outlet link. If the programmable cell is programmed OFF, the corresponding inlet link and outlet link are not connected. For example if the programmable cell P(1,1) is programmed ON, the corresponding transistor C(1,1) couples the corresponding inlet link IL1 and outlet link OL1. If the programmable cell P(1,1) is programmed OFF, the corresponding inlet link IL1 and outlet link OL1 are not connected. In volatile programmable integrated circuit embodiments the programmable cell may be an SRAM (Static Random Address Memory) cell. In non-volatile programmable integrated circuit embodiments the programmable cell may be a Flash memory cell. Also the programmable integrated circuit embodiments may implement field programmable logic arrays (FPGA) devices, or programmable Logic devices (PLD), or Application Specific Integrated Circuits (ASIC) embedded with programmable logic circuits or 3D-FPGAs.

2) One-Time Programmable Integrated Circuit Embodiments:

[0770] All the embodiments disclosed in the current invention are useful in one-time programmable integrated circuit applications. FIG. 8A3 illustrates the detailed diagram 800A3 for the implementation of the diagram 800A1 in one-time programmable integrated circuit embodiments. Each crosspoint is implemented by a via coupled between the corresponding inlet link and outlet link in one-time programmable integrated circuit embodiments. Specifically crosspoint CP(1,1) is implemented by via V(1,1) coupled between inlet link IL1 and outlet link OL1; crosspoint CP(1,2) is implemented by via V(1,2) coupled between inlet link IL1 and outlet link OL2; crosspoint CP(2,1) is implemented by via V(2,1) coupled between inlet link IL2 and outlet link OL1; and crosspoint CP(2,2) is implemented by via V(2,2) coupled between inlet link IL2 and outlet link OL2.

[0771] If the via is programmed ON, the corresponding inlet link and outlet link are permanently connected which is denoted by thick circle at the intersection of inlet link and outlet link. If the via is programmed OFF, the corresponding inlet link and outlet link are not connected which is denoted by the absence of thick circle at the intersection of inlet link and outlet link. For example in the diagram 800A3 the via V(1,1) is programmed ON, and the corresponding inlet link IL1 and outlet link OL1 are connected as denoted by thick circle at the intersection of inlet link IL1 and outlet link OL1; the via V(2,2) is programmed ON, and the corresponding inlet link IL2 and outlet link OL2 are connected as denoted by thick circle at the intersection of inlet link IL2 and outlet link OL2; the via V(1,2) is programmed OFF, and the corresponding inlet link IL1 and outlet link OL2 are not connected as

denoted by the absence of thick circle at the intersection of inlet link IL1 and outlet link OL2; the via V(2,1) is programmed OFF, and the corresponding inlet link IL2 and outlet link OL1 are not connected as denoted by the absence of thick circle at the intersection of inlet link IL2 and outlet link OL1. One-time programmable integrated circuit embodiments may be anti-fuse based programmable integrated circuit devices or mask programmable structured ASIC devices.

3) Integrated Circuit Placement and Route Embodiments:

[0772] All the embodiments disclosed in the current invention are useful in Integrated Circuit Placement and Route applications, for example in ASIC backend Placement and Route tools. FIG. 8A4 illustrates the detailed diagram 800A4 for the implementation of the diagram 800A1 in Integrated Circuit Placement and Route embodiments. In an integrated circuit since the connections are known a-priori, the switch and crosspoints are actually virtual. However the concept of virtual switch and virtual crosspoint using the embodiments disclosed in the current invention reduces the number of required wires, wire length needed to connect the inputs and outputs of different netlists and the time required by the tool for placement and route of netlists in the integrated circuit.

[0773] Each virtual crosspoint is used to either to hardwire or provide no connectivity between the corresponding inlet link and outlet link. Specifically crosspoint CP(1,1) is implemented by direct connect point DCP(1,1) to hardwire (i.e., to permanently connect) inlet link IL1 and outlet link OL1 which is denoted by the thick circle at the intersection of inlet link IL1 and outlet link OL1; crosspoint CP(2,2) is implemented by direct connect point DCP(2,2) to hardwire inlet link IL2 and outlet link OL2 which is denoted by the thick circle at the intersection of inlet link IL2 and outlet link OL2. The diagram 800A4 does not show direct connect point DCP(1,2) and direct connect point DCP(1,3) since they are not needed and in the hardware implementation they are eliminated. Alternatively inlet link IL1 needs to be connected to outlet link OL1 and inlet link IL1 does not need to be connected to outlet link OL2. Also inlet link IL2 needs to be connected to outlet link OL2 and inlet link IL2 does not need to be connected to outlet link OL1. Furthermore in the example of the diagram 800A4, there is no need to drive the signal of inlet link IL1 horizontally beyond outlet link OL1 and hence the inlet link IL1 is not even extended horizontally until the outlet link OL2. Also the absence of direct connect point DCP(2,1) illustrates there is no need to connect inlet link IL2 and outlet link OL1.

[0774] In summary in integrated circuit placement and route tools, the concept of virtual switches and virtual cross points is used during the implementation of the placement & routing algorithmically in software, however during the hardware implementation cross points in the cross state are implemented as hardwired connections between the corresponding inlet link and outlet link, and in the bar state are implemented as no connection between inlet link and outlet link

3) More Application Embodiments:

[0775] All the embodiments disclosed in the current invention are also useful in the design of SoC interconnects, Field programmable interconnect chips, parallel computer systems and in time-space-time switches.

US 2011/0044329 A1

Feb. 24, 2011

82

[0776] Numerous modifications and adaptations of the embodiments, implementations, and examples described herein will be apparent to the skilled artisan in view of the disclosure.

What is claimed is:

1. A network having a plurality of multicast connections, said network comprising:

N_1 inlet links and N_2 outlet links, and

when $N_2 > N_1$ and $N_2 = p * N_1$ where $p > 1$ then $N_1 = N$, $d_1 = d$, and

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d;$$

and

an input stage comprising

$$\frac{N_1}{d}$$

input switches, and each input switch comprising d inlet links and each said input switch further comprising $x \times d$ outgoing links connecting to switches in a second stage where $x > 0$; and an output stage comprising

$$\frac{N_1}{d}$$

output switches, and each output switch comprising d_2 outlet links and each said output switch further comprising

$$x \times \frac{(d + d_2)}{2}$$

incoming links connecting from switches in the penultimate stage; and

a plurality of y middle stages comprising N/d middle switches in each of said y middle stages wherein said second stage and said penultimate stage are one of said middle stages where $y > 3$, and

each middle switch in all said middle stages excepting said penultimate stage comprising $x \times d$ incoming links (hereinafter "incoming middle links") connecting from switches in its immediate preceding stage, and each middle switch further comprising $x \times d$ outgoing links (hereinafter "outgoing middle links") connecting to switches in its immediate succeeding stage; and

each middle switch in said penultimate stage comprising $x \times d$ incoming links connecting from switches in its immediate preceding stage, and each middle switch further comprising

$$x \times \frac{(d + d_2)}{2}$$

outgoing links connecting to switches in its immediate succeeding stage i.e., said output stage; or

when $N_1 > N_2$ and $N_1 = p * N_2$, where $p > 1$ then $N_2 = N$, $d_2 = d$ and

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d$$

and

an input stage comprising

$$\frac{N_2}{d}$$

input switches, and each input switch comprising d_1 inlet links and each input switch further comprising

$$x \times \frac{(d + d_1)}{2}$$

outgoing links connecting to switches in a second stage where $x > 0$; and

an output stage comprising

$$\frac{N_2}{d}$$

output switches, and each output switch comprising d outlet links and each output switch further comprising $x \times d$ incoming links connecting from switches in the penultimate stage; and

a plurality of y middle stages comprising N/d middle switches in each of said y middle stages wherein said second stage and said penultimate stage are one of said middle stages where $y > 3$, and

each middle switch in said second stage comprising

$$x \times \frac{(d + d_1)}{2}$$

incoming links connecting from switches in its immediate preceding stage i.e., said input stage, and each middle switch further comprising $x \times d$ outgoing links connecting to switches in its immediate succeeding stage; and

each middle switch in all said middle stages excepting said second stage comprising $x \times d$ incoming links (hereinafter "incoming middle links") connecting from switches in its immediate preceding stage, and each middle switch further comprising $x \times d$ outgoing links (hereinafter "outgoing middle links") connecting to switches in its immediate succeeding stage; and

wherein each multicast connection from an inlet link passes through at most two outgoing links in input switch, and said multicast connection further passes through a plurality of outgoing links in a plurality of switches in each said middle stage and in said output stage.

2. The network of claim 1, wherein all said incoming middle links and outgoing middle links are connected in any

US 2011/0044329 A1

Feb. 24, 2011

83

arbitrary topology such that when no connections are setup in said network, a connection from any said inlet link to any said outlet link can be setup.

3. The network of claim 2, wherein $y \cong (2 \times \log_d N_1) - 3$ when $N_2 > N_1$, and $y \cong (2 \times \log_d N_2) - 3$ when $N_1 > N_2$.

4. The network of claim 3, wherein $x \cong 1$, wherein said each multicast connection comprises only one destination link, and

said each multicast connection from an inlet link passes through only one outgoing link in input switch, and said multicast connection further passes through only one outgoing link in one of the switches in each said middle stage and in said output stage, and

further is always capable of setting up said multicast connection by changing the path, defined by passage of an existing multicast connection, thereby to change only one outgoing link of the input switch used by said existing multicast connection, and said network is hereinafter "rearrangeably nonblocking network for unicast".

5. The network of claim 3, wherein $x \cong 2$, wherein said each multicast connection comprises only one destination link, and

said each multicast connection from an inlet link passes through only one outgoing link in input switch, and said multicast connection further passes through only one outgoing link in one of the switches in each said middle stage and in said output stage, and

further is always capable of setting up said multicast connection by never changing path of an existing multicast connection, wherein said each multicast connection comprises only one destination link and the network is hereinafter "strictly nonblocking network for unicast".

6. The network of claim 3, wherein $x \cong 2$,

further is always capable of setting up said multicast connection by changing the path, defined by passage of an existing multicast connection, thereby to change one or two outgoing links of the input switch used by said existing multicast connection, and said network is hereinafter "rearrangeably nonblocking network".

7. The network of claim 3, wherein $x \cong 3$,

further is always capable of setting up said multicast connection by never changing path of an existing multicast connection, and the network is hereinafter "strictly nonblocking network".

8. The network of claim 1, further comprising a controller coupled to each of said input, output and middle stages to set up said multicast connection.

9. The network of claim 1, wherein said N_1 inlet links and N_2 outlet links are the same number of links, i.e., $N_1 = N_2 = N$, and $d_1 = d_2 = d$.

10. The network of claim 1, wherein said each input switch, said each output switch and said each middle switch is either fully populated or partially populated.

11. The network of claim 1,

wherein each of said input switches, or each of said output switches, or each of said middle switches further recursively comprise one or more networks.

12. A method for setting up one or more multicast connections in a network having N_1 inlet links and N_2 outlet links, and

when $N_2 > N_1$ and $N_2 = p * N_1$ where $p > 1$ then $N_1 = N$, $d_1 = d$, and

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d;$$

and having

an input stage having

$$\frac{N_1}{d}$$

input switches, and each input switch having d inlet links and each input switch further having $x \times d$ outgoing links connected to switches in a second stage where $x > 0$; and

an output stage having

$$\frac{N_1}{d}$$

output switches, and each output switch having d_2 outlet links and each output switch further having

$$x \times \frac{(d + d_2)}{2}$$

incoming links connected from switches in the penultimate stage; and

a plurality of y middle stages having N/d middle switches in each of said y middle stages wherein said second stage and said penultimate stage being one of said middle stages where $y > 3$, and

each middle switch in all said middle stages excepting said penultimate stage having $x \times d$ incoming links connected from switches in its immediate preceding stage, and each middle switch further having $x \times d$ outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said penultimate stage having $x \times d$ incoming links connected from switches in its immediate preceding stage, and each middle switch further having

$$x \times \frac{(d + d_2)}{2}$$

outgoing links connected to switches in its immediate succeeding stage; or

when $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$ then $N_2 = N$, $d_2 = d$ and

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d;$$

US 2011/0044329 A1

Feb. 24, 2011

84

and having

an input stage having

$$\frac{N_2}{d}$$

input switches, and each input switch having d_1 inlet links and each input switch further having

$$x \times \frac{(d + d_1)}{2}$$

outgoing links connected to switches in a second stage where $x > 0$; and

an output stage having

$$\frac{N_2}{d}$$

output switches, and each output switch having d outlet links and each output switch further having $x \times d$ incoming links connected from switches in the penultimate stage; anda plurality of y middle stages having N/d middle switches in each of said y middle stages wherein said second stage and said penultimate stage being one of said middle stages where $y > 3$, and

each middle switch in said second stage having

$$x \times \frac{(d + d_1)}{2}$$

incoming links connected from switches in its immediate preceding stage, and each middle switch further having $x \times d$ outgoing links connected to switches in its immediate succeeding stage; andeach middle switch in all said middle stages excepting said second stage having $x \times d$ incoming links connected from switches in its immediate preceding stage, and each middle switch further having $x \times d$ outgoing links connected to switches in its immediate succeeding stage; and said method comprising:

receiving a multicast connection at said input stage;

fanning out said multicast connection through at most two outgoing links in input switch and a plurality of outgoing links in a plurality of middle switches in each said middle stage to set up said multicast connection to a plurality of output switches among said

$$\frac{N_2}{d}$$

output switches, wherein said plurality of output switches are specified as destinations of said multicast connection, wherein said at most two outgoing links in input switch and said plurality of outgoing links in said plurality of middle switches in each said middle stage are available.

13. A method of claim 12 wherein said act of fanning out is performed without changing any existing connection to pass through another set of plurality of middle switches in each said middle stage.

14. A method of claim 12 wherein said act of fanning out is performed recursively.

15. A method of claim 12 wherein a connection exists through said network and passes through a plurality of middle switches in each said middle stage and said method further comprises:

if necessary, changing said connection to pass through another set of plurality of middle switches in each said middle stage, act hereinafter "rearranging connection".

16. A method of claim 12 wherein said acts of fanning out and rearranging are performed recursively.

17. A method for setting up one or more multicast connections in a network having N_1 inlet links and N_2 outlet links, and

when $N_2 > N_1$ and $N_2 = p \times N_1$, where $p > 1$ then $N_1 = N$, $d_1 = d$, and

$$d_2 = N_2 \times \frac{d}{N_1} = p \times d;$$

and having

an input stage having

$$\frac{N_1}{d}$$

input switches, and each input switch having d inlet links and each input switch further having $x \times d$ outgoing links connected to switches in a second stage where $x > 0$; and

an output stage having

$$\frac{N_1}{d}$$

output switches, and each output switch having d_2 outlet links and each output switch further having

$$x \times \frac{(d + d_2)}{2}$$

incoming links connected from switches in the penultimate stage; and

a plurality of y middle stages having N/d middle switches in each of said y middle stages wherein said second stage and said penultimate stage being one of said middle stages where $y > 3$, andeach middle switch in all said middle stages excepting said penultimate stage having $x \times d$ incoming links connected from switches in its immediate preceding stage, and each middle switch further having $x \times d$ outgoing links connected to switches in its immediate succeeding stage; and

US 2011/0044329 A1

Feb. 24, 2011

85

each middle switch in said penultimate stage having $x \times d$ incoming links connected from switches in its immediate preceding stage, and each middle switch further having

$$x \times \frac{(d + d_2)}{2}$$

outgoing links connected to switches in its immediate succeeding stage; or

when $N_1 > N_2$ and $N_1 = p \times N_2$ where $p > 1$ then $N_2 = N$, $d_2 = d$ and

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d;$$

and having

an input stage having

$$\frac{N_2}{d}$$

input switches, and each input switch having d_1 inlet links and each input switch further having

$$x \times \frac{(d + d_1)}{2}$$

outgoing links connected to switches in a second stage where $x > 0$; and

an output stage having

$$\frac{N_2}{d}$$

output switches, and each output switch having d outlet links and each output switch further having $x \times d$ incoming links connected from switches in the penultimate stage; and

a plurality of y middle stages having N/d middle switches in each of said y middle stages wherein said second stage and said penultimate stage being one of said middle stages where $y > 3$, and

each middle switch in said second stage having

$$x \times \frac{(d + d_1)}{2}$$

incoming links connected from switches in its immediate preceding stage, and each middle switch further having $x \times d$ outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in all said middle stages excepting said second stage having $x \times d$ incoming links connected from switches in its immediate preceding stage, and each middle switch further having $x \times d$ outgoing links connected to switches in its immediate succeeding stage; and said method comprising:

checking if a first outgoing link in input switch and a first plurality of outgoing links in plurality of middle switches in each said middle stage are available to at least a first subset of destination output switches of said multicast connection; and

checking if a second outgoing link in input switch and second plurality of outgoing links in plurality of middle switches in each said middle stage are available to a second subset of destination output switches of said multicast connection.

wherein each destination output switch of said multicast connection is one of said first subset of destination output switches and said second subset of destination output switches.

18. The method of claim 17 further comprising:

prior to said checkings, checking if all the destination output switches of said multicast connection are available through said first outgoing link in input switch and said first plurality of outgoing links in plurality of middle switches in each said middle stage

19. The method of claim 17 further comprising:

repeating said checkings of available second outgoing link in input switch and second plurality of outgoing links in plurality of middle switches in each said middle stage to a second subset of destination output switches of said multicast connection to each outgoing link in input switch other than said first and said second outgoing links in input switch.

wherein each destination output switch of said multicast connection is one of said first subset of destination output switches and said second subset of destination output switches.

20. The method of claim 17 further comprising:

repeating said checkings of available first outgoing link in input switch and first plurality of outgoing links in plurality of middle switches in each said middle stage to a first subset of destination output switches of said multicast connection to each outgoing link in input switch other than said first outgoing link in input switch.

21. The method of claim 17 further comprising:

setting up each of said multicast connection from its said input switch to its said output switches through not more than two outgoing links, selected by said checkings, by fanning out said multicast connection in its said input switch into not more than said two outgoing links.

22. The method of claim 17 wherein any of said acts of checking and setting up are performed recursively.

* * * * *

EXHIBIT G

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 December 2008 (04.12.2008)

PCT

(10) International Publication Number
WO 2008/147928 A1

(51) International Patent Classification:
H01L 25/00 (2006.01)

(21) International Application Number:
PCT/US2008/064605

(22) International Filing Date: 22 May 2008 (22.05.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/940,394 25 May 2007 (25.05.2007) US

(71) Applicant and

(72) Inventor: **KONDA, Venkat** [US/US]; 6278, Grand Oak Way, San Jose, CA 95135 (US).

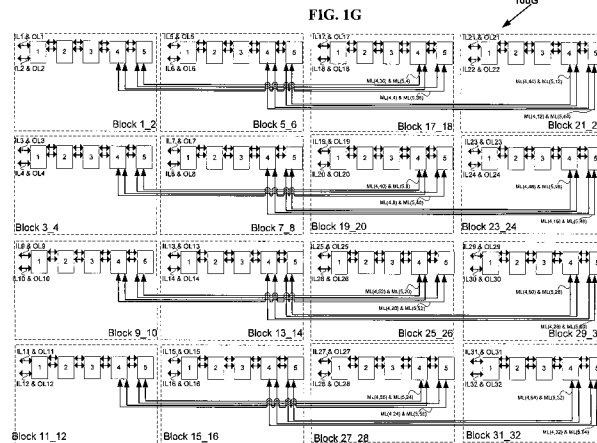
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

Published: — with international search report

(54) Title: VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS



(57) Abstract: In accordance with the invention, VLSI layouts of generalized multi-stage networks for broadcast, unicast and multicast connections are presented using only horizontal and vertical links. The VLSI layouts employ shuffle exchange links where outlet links of cross links from switches in a stage in one sub-integrated circuit block are connected to inlet links of switches in the succeeding stage in another sub-integrated circuit block so that said cross links are either vertical links or horizontal and vice versa. In one embodiment the sub-integrated circuit blocks are arranged in a hypercube arrangement in a two-dimensional plane. The VLSI layouts exploit the benefits of significantly lower cross points, lower signal latency, lower power and full connectivity with significantly fast compilation.



WO 2008/147928 A1

WO 2008/147928

PCT/US2008/064605

VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS**Venkat Konda**

5 CROSS REFERENCE TO RELATED APPLICATIONS

This application is Continuation In Part PCT Application to and incorporates by reference in its entirety the U.S. Provisional Patent Application Serial No. 60/940, 394 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25,
10 2007.

This application is related to and incorporates by reference in its entirety the PCT Application Serial No. PCT/US08/56064 entitled "FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed March 6, 2008, the U.S. Provisional Patent
15 Application Serial No. 60/905,526 entitled "LARGE SCALE CROSSPOINT REDUCTION WITH NONBLOCKING UNICAST & MULTICAST IN ARBITRARILY LARGE MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed March 6, 2007, and the U.S. Provisional Patent Application Serial No. 60/940, 383 entitled "FULLY CONNECTED
20 GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

This application is related to and incorporates by reference in its entirety the PCT Application Docket No. S-0038PCT entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same
25 assignee as the current application, filed concurrently, the U.S. Provisional Patent Application Serial No. 60/940, 387 entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same

WO 2008/147928

PCT/US2008/064605

assignee as the current application, filed May 25, 2007, and the U.S. Provisional Patent Application Serial No. 60/940, 390 entitled "FULLY CONNECTED GENERALIZED MULTI-LINK BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007

5 This application is related to and incorporates by reference in its entirety the PCT Application Docket No. S-0039PCT entitled "FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed concurrently, the U.S. Provisional Patent Application Serial No. 60/940, 389 entitled "FULLY CONNECTED GENERALIZED
10 REARRANGEABLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, the U.S. Provisional Patent Application Serial No. 60/940, 391 entitled "FULLY CONNECTED GENERALIZED FOLDED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007 and
15 the U.S. Provisional Patent Application Serial No. 60/940, 392 entitled "FULLY CONNECTED GENERALIZED STRICTLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

 This application is related to and incorporates by reference in its entirety the U.S.
20 Provisional Patent Application Serial No. 60/984, 724 entitled "VLSI LAYOUTS OF FULLY CONNECTED NETWORKS WITH LOCALITY EXPLOITATION" by Venkat Konda assigned to the same assignee as the current application, filed November 2, 2007.

 This application is related to and incorporates by reference in its entirety the U.S. Provisional Patent Application Serial No. 61/018, 494 entitled "VLSI LAYOUTS OF
25 FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed January 1, 2008.

WO 2008/147928

PCT/US2008/064605

BACKGROUND OF INVENTION

Multi-stage interconnection networks such as Benes networks and butterfly fat tree networks are widely useful in telecommunications, parallel and distributed computing. However VLSI layouts, known in the prior art, of these interconnection
5 networks in an integrated circuit are inefficient and complicated.

Other multi-stage interconnection networks including butterfly fat tree networks, Banyan networks, Batcher-Banyan networks, Baseline networks, Delta networks, Omega networks and Flip networks have been widely studied particularly for self routing packet switching applications. Also Benes Networks with radix of two have been widely studied
10 and it is known that Benes Networks of radix two are shown to be built with back to back baseline networks which are rearrangeably nonblocking for unicast connections.

The most commonly used VLSI layout in an integrated circuit is based on a two-dimensional grid model comprising only horizontal and vertical tracks. An intuitive interconnection network that utilizes two-dimensional grid model is 2D Mesh Network
15 and its variations such as segmented mesh networks. Hence routing networks used in VLSI layouts are typically 2D mesh networks and its variations. However Mesh Networks require large scale cross points typically with a growth rate of $O(N^2)$ where N is the number of computing elements, ports, or logic elements depending on the application.

20 Multi-stage interconnection with a growth rate of $O(N \times \log N)$ requires significantly small number of cross points. U.S. Patent 6,185,220 entitled "Grid Layouts of Switching and Sorting Networks" granted to Muthukrishnan et al. describes a VLSI layout using existing VLSI grid model for Benes and Butterfly networks. U.S. Patent 6,940,308 entitled "Interconnection Network for a Field Programmable Gate Array"
25 granted to Wong describes a VLSI layout where switches belonging to lower stage of Benes Network are layed out close to the logic cells and switches belonging to higher stages are layed out towards the center of the layout.

WO 2008/147928

PCT/US2008/064605

Due to the inefficient and in some cases impractical VLSI layout of Benes and butterfly fat tree networks on a semiconductor chip, today mesh networks and segmented mesh networks are widely used in the practical applications such as field programmable gate arrays (FPGAs), programmable logic devices (PLDs), and parallel computing interconnects. The prior art VLSI layouts of Benes and butterfly fat tree networks and VLSI layouts of mesh networks and segmented mesh networks require large area to implement the switches on the chip, large number of wires, longer wires, with increased power consumption, increased latency of the signals which effect the maximum clock speed of operation. Some networks may not even be implemented practically on a chip due to the lack of efficient layouts.

SUMMARY OF INVENTION

When large scale sub-integrated circuit blocks with inlet and outlet links are layed out in an integrated circuit device in a two-dimensional grid arrangement, (for example in an FPGA where the sub-integrated circuit blocks are Lookup Tables) the most intuitive routing network is a network that uses horizontal and vertical links only (the most often used such a network is one of the variations of a 2D Mesh network). A direct embedding of a generalized multi-stage network on to a 2D Mesh network is neither simple nor efficient.

In accordance with the invention, VLSI layouts of generalized multi-stage networks for broadcast, unicast and multicast connections are presented using only horizontal and vertical links. The VLSI layouts employ shuffle exchange links where outlet links of cross links from switches in a stage in one sub-integrated circuit block are connected to inlet links of switches in the succeeding stage in another sub-integrated circuit block so that said cross links are either vertical links or horizontal and vice versa. In one embodiment the sub-integrated circuit blocks are arranged in a hypercube arrangement in a two-dimensional plane. The VLSI layouts exploit the benefits of significantly lower cross points, lower signal latency, lower power and full connectivity with significantly fast compilation.

WO 2008/147928

PCT/US2008/064605

The VLSI layouts presented are applicable to generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general. The embodiments of VLSI layouts are useful in wide target applications such as FPGAs, CPLDs, pSoCs, ASIC placement and route tools, networking applications, parallel & distributed computing, and reconfigurable computing.

10

BRIEF DESCRIPTION OF DRAWINGS

FIG. 1A is a diagram 100A of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having inverse Benes connection topology of nine stages with $N = 32$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 1B is a diagram 100B of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network 100A shown in FIG. 1A, having inverse Benes connection topology of five stages with $N = 32$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 1C is a diagram 100C layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links belonging with in each block only.

WO 2008/147928

PCT/US2008/064605

FIG. 1D is a diagram 100D layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(1,i) for $i = [1, 64]$ and ML(8,i) for $i = [1, 64]$.

FIG. 1E is a diagram 100E layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(2,i) for $i = [1, 64]$ and ML(7,i) for $i = [1, 64]$.

FIG. 1F is a diagram 100F layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(3,i) for $i = [1, 64]$ and ML(6,i) for $i = [1, 64]$.

FIG. 1G is a diagram 100G layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(4,i) for $i = [1, 64]$ and ML(5,i) for $i = [1, 64]$.

FIG. 1H is a diagram 100H layout of a network $V_{fold-link}(N, d, s)$ where $N = 128$, $d = 2$, and $s = 2$, in one embodiment, illustrating the connection links belonging with in each block only.

FIG. 1I is a diagram 100I detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

FIG. 1J is a diagram 100J detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

FIG. 1K is a diagram 100K detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

WO 2008/147928

PCT/US2008/064605

FIG. 1K1 is a diagram 100M1 detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$ for $s = 1$.

FIG. 1L is a diagram 100L detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

FIG. 1L1 is a diagram 100L1 detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$ for $s = 1$.

FIG. 2A1 is a diagram 200A1 of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having inverse Benes connection topology of one stage with $N = 2$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2A2 is a diagram 200A2 of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network 200A1 shown in FIG. 2A1, having inverse Benes connection topology of one stage with $N = 2$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2A3 is a diagram 200A3 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2A2, in one embodiment, illustrating all the connection links.

FIG. 2B1 is a diagram 200B1 of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having inverse Benes connection topology of one stage with $N = 4$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2B2 is a diagram 200B2 of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network

WO 2008/147928

PCT/US2008/064605

200B1 shown in FIG. 2B1, having inverse Benes connection topology of one stage with $N = 4$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2B3 is a diagram 200B3 layout of the network

5 $V_{fold-link}(N, d, s)$ shown in FIG. 2B2, in one embodiment, illustrating the connection links belonging with in each block only. FIG. 2B4 is a diagram 200B4 layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2B2, in one embodiment, illustrating the connection links ML(1,i) for $i = [1, 8]$ and ML(2,i) for $i = [1, 8]$.

FIG. 2C11 is a diagram 200C11 of an exemplary symmetrical multi-link multi-

10 stage network $V_{fold-link}(N, d, s)$ having inverse Benes connection topology of one stage with $N = 8$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2C12 is a diagram 200C12 of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-link}(N, d, s)$ of the network

15 200C11 shown in FIG. 2C11, having inverse Benes connection topology of one stage with $N = 8$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2C21 is a diagram 200C21 layout of the network $V_{fold-link}(N, d, s)$ shown in

20 FIG. 2C12, in one embodiment, illustrating the connection links belonging with in each block only. FIG. 2C22 is a diagram 200C22 layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2C12, in one embodiment, illustrating the connection links ML(1,i) for $i = [1, 16]$ and ML(4,i) for $i = [1, 16]$. FIG. 2C23 is a diagram 200C23 layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2C12, in one embodiment, illustrating the connection

25 links ML(2,i) for $i = [1, 16]$ and ML(3,i) for $i = [1, 16]$.

FIG. 2D1 is a diagram 200D1 of an exemplary symmetrical multi-link multi-stage network $V_{fold-link}(N, d, s)$ having inverse Benes connection topology of one stage with $N = 16$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and

WO 2008/147928

PCT/US2008/064605

rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2D2 is a diagram 200D2 of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-link}(N, d, s)$ of the network 200D1 shown in FIG. 2D1, having
 5 inverse Benes connection topology of one stage with $N = 16$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2D3 is a diagram 200D3 layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2D2, in one embodiment, illustrating the connection links belonging with in each
 10 block only.

FIG. 2D4 is a diagram 200D4 layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2D2, in one embodiment, illustrating the connection links $ML(1,i)$ for $i = [1, 32]$ and $ML(6,i)$ for $i = [1,32]$.

FIG. 2D5 is a diagram 200D5 layout of the network $V_{fold-link}(N, d, s)$ shown in
 15 FIG. 2D2, in one embodiment, illustrating the connection links $ML(2,i)$ for $i = [1, 32]$ and $ML(5,i)$ for $i = [1,32]$.

FIG. 2D6 is a diagram 200D6 layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2D2, in one embodiment, illustrating the connection links $ML(3,i)$ for $i = [1, 32]$ and $ML(4,i)$ for $i = [1,32]$.

20 FIG. 3A is a diagram 300A of an exemplary symmetrical multi-link multi-stage network $V_{hcube}(N, d, s)$ having inverse Benes connection topology of nine stages with $N = 32$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

25 FIG. 3B is a diagram 300B of the equivalent symmetrical folded multi-link multi-stage network $V_{hcube}(N, d, s)$ of the network 300A shown in FIG. 3A, having inverse

WO 2008/147928

PCT/US2008/064605

Benes connection topology of five stages with $N = 32$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 3C is a diagram 300C layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links belonging with in each block only.

FIG. 3D is a diagram 100D layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links $ML(1,i)$ for $i = [1, 64]$ and $ML(8,i)$ for $i = [1,64]$.

FIG. 3E is a diagram 300E layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links $ML(2,i)$ for $i = [1, 64]$ and $ML(7,i)$ for $i = [1,64]$.

FIG. 3F is a diagram 300F layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links $ML(3,i)$ for $i = [1, 64]$ and $ML(6,i)$ for $i = [1,64]$.

FIG. 3G is a diagram 300G layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links $ML(4,i)$ for $i = [1, 64]$ and $ML(5,i)$ for $i = [1,64]$.

FIG. 3H is a diagram 300H layout of a network $V_{hcube}(N, d, s)$ where $N = 128$, $d = 2$, and $s = 2$, in one embodiment, illustrating the connection links belonging with in each block only.

FIG. 4A is a diagram 400A layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links belonging with in each block only.

WO 2008/147928

PCT/US2008/064605

FIG. 4B is a diagram 400B layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(1,i) for $i = [1, 64]$ and ML(8,i) for $i = [1, 64]$.

FIG. 4C is a diagram 400C layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 4C, in one embodiment, illustrating the connection links ML(2,i) for $i = [1, 64]$ and ML(7,i) for $i = [1, 64]$.

FIG. 4D is a diagram 400D layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 4D, in one embodiment, illustrating the connection links ML(3,i) for $i = [1, 64]$ and ML(6,i) for $i = [1, 64]$.

FIG. 4E is a diagram 400E layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 4E, in one embodiment, illustrating the connection links ML(4,i) for $i = [1, 64]$ and ML(5,i) for $i = [1, 64]$.

FIG. 4C1 is a diagram 400C1 layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links belonging with in each block only.

FIG. 5A1 is a diagram 500A1 of an exemplary prior art implementation of a two by two switch; FIG. 5A2 is a diagram 500A2 for programmable integrated circuit prior art implementation of the diagram 500A1 of FIG. 5A1; FIG. 5A3 is a diagram 500A3 for one-time programmable integrated circuit prior art implementation of the diagram 500A1 of FIG. 5A1; FIG. 5A4 is a diagram 500A4 for integrated circuit placement and route implementation of the diagram 500A1 of FIG. 5A1.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is concerned with the VLSI layouts of arbitrarily large switching networks for broadcast, unicast and multicast connections. Particularly switching networks considered in the current invention include: generalized multi-stage

WO 2008/147928

PCT/US2008/064605

- networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{cube}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general.

Efficient VLSI layout of networks on a semiconductor chip are very important and greatly influence many important design parameters such as the area taken up by the network on the chip, total number of wires, length of the wires, latency of the signals, capacitance and hence the maximum clock speed of operation. Some networks may not even be implemented practically on a chip due to the lack of efficient layouts. The different varieties of multi-stage networks described above have not been implemented previously on the semiconductor chips efficiently. For example in Field Programmable Gate Array (FPGA) designs, multi-stage networks described in the current invention have not been successfully implemented primarily due to the lack of efficient VLSI layouts. Current commercial FPGA products such as Xilinx Vertex, Altera's Stratix implement island-style architecture using mesh and segmented mesh routing interconnects using either full crossbars or sparse crossbars. These routing interconnects consume large silicon area for crosspoints, long wires, large signal propagation delay and hence consume lot of power.

The current invention discloses the VLSI layouts of numerous types of multi-stage networks which are very efficient. Moreover they can be embedded on to mesh and segmented mesh routing interconnects of current commercial FPGA products. The VLSI layouts disclosed in the current invention are applicable to including the numerous generalized multi-stage networks disclosed in the following patent applications, filed concurrently:

- 1) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-stage networks $V(N_1, N_2, d, s)$ with numerous connection

WO 2008/147928

PCT/US2008/064605

topologies and the scheduling methods are described in detail in the PCT Application Serial No. PCT/US08/56064 that is incorporated by reference above.

2) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$ with numerous
5 connection topologies and the scheduling methods are described in detail in U.S. Provisional Patent Application Serial No. 60/940, 387 that is incorporated by reference above.

3) Rearrangeably nonblocking for arbitrary fan-out multicast and unicast, and strictly nonblocking for unicast for generalized multi-link multi-stage networks
10 $V_{mlink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in U.S. Provisional Patent Application Serial No. 60/940, 389 that is incorporated by reference above.

4) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$ with
15 numerous connection topologies and the scheduling methods are described in detail in U.S. Provisional Patent Application Serial No. 60/940, 390 that is incorporated by reference above.

5) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$ with numerous
20 connection topologies and the scheduling methods are described in detail in U.S. Provisional Patent Application Serial No. 60/940, 391 that is incorporated by reference above.

6) Strictly nonblocking for arbitrary fan-out multicast for generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage
25 networks $V_{fold-mlink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling

WO 2008/147928

PCT/US2008/064605

methods are described in detail in U.S. Provisional Patent Application Serial No. 60/940, 392 that is incorporated by reference above.

7) VLSI layouts of numerous types of multi-stage networks with locality exploitation are described in U.S. Provisional Patent Application Serial No. 60/984, 724 that is incorporated by reference above.

8) VLSI layouts of numerous types of multistage pyramid networks are described in U.S. Provisional Patent Application Serial No. 61/018, 494 that is incorporated by reference above.

In addition the layouts of the current invention are also applicable to generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage pyramid networks $V_{fold-mlink-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{cube}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general.

Symmetric RNB generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$:

Referring to diagram 100A in FIG. 1A, in one embodiment, an exemplary generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages of one hundred and forty four switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, 150, 160, 170, 180 and 190 is shown where input stage 110 consists of sixteen, two by four switches IS1-IS16 and output stage 120 consists of sixteen, four by two switches OS1-OS16. And all the middle stages namely the middle

WO 2008/147928

PCT/US2008/064605

stage 130 consists of sixteen, four by four switches MS(1,1) - MS(1,16), middle stage 140 consists of sixteen, four by four switches MS(2,1) - MS(2,16), middle stage 150 consists of sixteen, four by four switches MS(3,1) - MS(3,16), middle stage 160 consists of sixteen, four by four switches MS(4,1) - MS(4,16), middle stage 170 consists of
 5 sixteen, four by four switches MS(5,1) - MS(5,16), middle stage 180 consists of sixteen, four by four switches MS(6,1) - MS(6,16), and middle stage 190 consists of sixteen, four by four switches MS(7,1) - MS(7,16).

As disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above, such a network can be operated in rearrangeably non-
 10 blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections.

In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable $\frac{N}{d}$, where N is the total
 15 number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by $\frac{N}{d}$. The size of each input switch IS1-IS4 can be denoted in general with the notation $d * 2d$ and each output switch OS1-OS4 can be denoted in general with the notation $2d * d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d * 2d$. A switch as used herein can be either a crossbar switch, or a
 20 network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-stage network can be represented with the notation $V_{link}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL32), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from
 25 each input switch to the inlet links of each input switch.

Each of the $\frac{N}{d}$ input switches IS1 - IS16 are connected to exactly d switches in middle stage 130 through two links each for a total of $2 \times d$ links (for example input

WO 2008/147928

PCT/US2008/064605

switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and also connected to middle switch MS(1,2) through the links ML(1,3) and ML(1,4). The middle links which connect switches in the same row in two successive middle stages are called hereinafter straight middle links; and the middle links which connect
 5 switches in different rows in two successive middle stages are called hereinafter cross middle links. For example, the middle links ML(1,1) and ML(1,2) connect input switch IS1 and middle switch MS(1,1), so middle links ML(1,1) and ML(1,2) are straight middle links; where as the middle links ML(1,3) and ML(1,4) connect input switch IS1 and middle switch MS(1,2), since input switch IS1 and middle switch MS(1,2) belong to two
 10 different rows in diagram 100A of FIG. 1A, middle links ML(1,3) and ML(1,4) are cross middle links.

Each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,16) in the middle stage 130 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1)
 15 from input switch IS1, and the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to exactly d switches in middle stage 140 through two links each for a total of $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1)
 20 to middle switch MS(2,3)).

Each of the $\frac{N}{d}$ middle switches MS(2,1) – MS(2,16) in the middle stage 140 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1)
 from input switch MS(1,1), and the links ML(1,11) and ML(1,12) are connected to the
 25 middle switch MS(2,1) from input switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through two links each for a total of $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,5)).

WO 2008/147928

PCT/US2008/064605

Each of the $\frac{N}{d}$ middle switches MS(3,1) – MS(3,16) in the middle stage 150 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from input switch MS(2,1), and the links ML(2,19) and ML(2,20) are connected to the middle switch MS(3,1) from input switch MS(2,5)) and also are connected to exactly d switches in middle stage 160 through two links each for a total of $2 \times d$ links (for example the links ML(4,1) and ML(4,2) are connected from middle switch MS(3,1) to middle switch MS(4,1), and the links ML(4,3) and ML(4,4) are connected from middle switch MS(3,1) to middle switch MS(4,9)).

10 Each of the $\frac{N}{d}$ middle switches MS(4,1) – MS(4,16) in the middle stage 160 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(4,1) and ML(4,2) are connected to the middle switch MS(4,1) from input switch MS(3,1), and the links ML(4,35) and ML(4,36) are connected to the middle switch MS(4,1) from input switch MS(3,9)) and also are connected to exactly d switches in middle stage 170 through two links each for a total of $2 \times d$ links (for example the links ML(5,1) and ML(5,2) are connected from middle switch MS(4,1) to middle switch MS(5,1), and the links ML(5,3) and ML(5,4) are connected from middle switch MS(4,1) to middle switch MS(5,9)).

20 Each of the $\frac{N}{d}$ middle switches MS(5,1) – MS(5,16) in the middle stage 170 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(5,1) and ML(5,2) are connected to the middle switch MS(5,1) from input switch MS(4,1), and the links ML(5,35) and ML(5,36) are connected to the middle switch MS(5,1) from input switch MS(4,9)) and also are connected to exactly d switches in middle stage 180 through two links each for a total of $2 \times d$ links (for example the links ML(6,1) and ML(6,2) are connected from middle switch MS(5,1) to middle switch MS(6,1), and the links ML(6,3) and ML(6,4) are connected from middle switch MS(5,1) to middle switch MS(6,5)).

WO 2008/147928

PCT/US2008/064605

Each of the $\frac{N}{d}$ middle switches MS(6,1) – MS(6,16) in the middle stage 180 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(6,1) and ML(6,2) are connected to the middle switch MS(6,1) from input switch MS(5,1), and the links ML(6,19) and ML(6,20) are connected to the middle switch MS(6,1) from input switch MS(5,5)) and also are connected to exactly d switches in middle stage 190 through two links each for a total of $2 \times d$ links (for example the links ML(7,1) and ML(7,2) are connected from middle switch MS(6,1) to middle switch MS(7,1), and the links ML(7,3) and ML(7,4) are connected from middle switch MS(6,1) to middle switch MS(7,3)).

10 Each of the $\frac{N}{d}$ middle switches MS(7,1) – MS(7,16) in the middle stage 190 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(7,1) and ML(7,2) are connected to the middle switch MS(7,1) from input switch MS(6,1), and the links ML(7,11) and ML(7,12) are connected to the middle switch MS(7,1) from input switch MS(6,3)) and also are connected to exactly d switches in middle stage 120 through two links each for a total of $2 \times d$ links (for example the links ML(8,1) and ML(8,2) are connected from middle switch MS(7,1) to middle switch MS(8,1), and the links ML(8,3) and ML(8,4) are connected from middle switch MS(7,1) to middle switch OS2).

Each of the $\frac{N}{d}$ middle switches OS1 – OS16 in the middle stage 120 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(8,1) and ML(8,2) are connected to the output switch OS1 from input switch MS(7,1), and the links ML(8,7) and ML(7,8) are connected to the output switch OS1 from input switch MS(7,2)).

Finally the connection topology of the network 100A shown in FIG. 1A is known to be back to back inverse Benes connection topology.

Referring to diagram 100B in FIG. 1B, is a folded version of the multi-link multi-stage network 100A shown in FIG. 1A. The network 100B in FIG. 1B shows input stage

WO 2008/147928

PCT/US2008/064605

110 and output stage 120 are placed together. That is input switch IS1 and output switch OS1 are placed together, input switch IS2 and output switch OS2 are placed together, and similarly input switch IS16 and output switch OS16 are placed together. All the right going middle links (hereinafter “forward connecting links”) {i.e., inlet links IL1 – IL32 and middle links ML(1,1) - ML(1,64)} correspond to input switches IS1 - IS16, and all the left going middle links (hereinafter “backward connecting links”) {i.e., middle links ML(8,1) - ML(8,64) and outlet links OL1-OL32} correspond to output switches OS1 - OS16.

Middle stage 130 and middle stage 190 are placed together. That is middle switches MS(1,1) and MS(7,1) are placed together, middle switches MS(1,2) and MS(7,2) are placed together, and similarly middle switches MS(1,16) and MS(7,16) are placed together. All the right going middle links {i.e., middle links ML(1,1) - ML(1,64) and middle links ML(2,1) – ML(2,64)} correspond to middle switches MS(1,1) – MS(1,16), and all the left going middle links {i.e., middle links ML(7,1) - ML(7,64) and middle links ML(8,1) and ML(8,64)} correspond to middle switches MS(7,1) – MS(7,16).

Middle stage 140 and middle stage 180 are placed together. That is middle switches MS(2,1) and MS(6,1) are placed together, middle switches MS(2,2) and MS(6,2) are placed together, and similarly middle switches MS(2,16) and MS(6,16) are placed together. All the right going middle links {i.e., middle links ML(2,1) - ML(2,64) and middle links ML(3,1) – ML(3,64)} correspond to middle switches MS(2,1) – MS(2,16), and all the left going middle links {i.e., middle links ML(6,1) - ML(6,64) and middle links ML(7,1) and ML(7,64)} correspond to middle switches MS(6,1) – MS(6,16).

Middle stage 150 and middle stage 170 are placed together. That is middle switches MS(3,1) and MS(5,1) are placed together, middle switches MS(3,2) and MS(5,2) are placed together, and similarly middle switches MS(3,16) and MS(5,16) are placed together. All the right going middle links {i.e., middle links ML(3,1) - ML(3,64) and middle links ML(4,1) – ML(4,64)} correspond to middle switches MS(3,1) – MS(3,16), and all the left going middle links {i.e., middle links ML(5,1) - ML(5,64) and

WO 2008/147928

PCT/US2008/064605

middle links ML(6,1) and ML(6,64)} correspond to middle switches MS(5,1) – MS(5,16).

Middle stage 160 is placed alone. All the right going middle links are the middle links ML(4,1) - ML(4,64) and all the left going middle links are middle links ML(5,1) -
5 ML(5,64).

In one embodiment, in the network 100B of FIG. 1B, the switches that are placed together are implemented as separate switches then the network 100B is the generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No.
10 60/940,389 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch and a four by two switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1) – ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is
15 – ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1 – OL2 being the outputs of the output switch OS1. Similarly in this embodiment of network 100B all the switches that are placed together in each middle stage are implemented as separate
20 switches.

Hypercube Topology layout schemes:

Referring to layout 100C of FIG. 1C, in one embodiment, there are sixteen blocks namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block
25 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, Block 23_24, Block 25_26, Block 27_28, Block 29_30, and Block 31_32. Each block implements all the switches in one row of the network 100B of FIG. 1B, one of the key aspects of the current invention. For example Block 1_2 implements the input switch IS1, output Switch OS1,

WO 2008/147928

PCT/US2008/064605

middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; Middle switch MS(4,1) is denoted by switch 5.

All the straight middle links are illustrated in layout 100C of FIG. 1C. For example in Block 1_2, inlet links IL1 – IL2, outlet links OL1 – OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 100C of FIG. 1C.

Even though it is not illustrated in layout 100C of FIG. 1C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit (hereinafter “sub-integrated circuit block”) depending on the applications in different embodiments. There are four quadrants in the layout 100C of FIG. 1C namely top-left, bottom-left, top-right and bottom-right quadrants. Top-left quadrant implements Block 1_2, Block 3_4, Block 5_6, and Block 7_8. Bottom-left quadrant implements Block 9_10, Block 11_12, Block 13_14, and Block 15_16. Top-right quadrant implements Block 17_18, Block 19_20, Block 21_22, and Block 23_24. Bottom-right quadrant implements Block 25_26, Block 27_28, Block 29_30, and Block 31_32. There are two halves in layout 100C of FIG. 1C namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants. Right-half consists of top-right and bottom-right quadrants.

Recursively in each quadrant there are four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements Block 1_2. Bottom-left sub-quadrant of top-left quadrant

WO 2008/147928

PCT/US2008/064605

implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 5_6. Finally bottom-right sub-quadrant of top-left quadrant implements Block 7_8. Similarly there are two sub-halves in each quadrant. For example in top-left quadrant there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left quadrant implements Block 1_2 and Block 3_4. Right-sub-half of top-left quadrant implements Block 5_6 and Block 7_8. Finally applicant notes that in each quadrant or half the blocks are arranged as a general binary hypercube. Recursively in larger multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 > 32$, the layout in this embodiment in accordance with the current invention, will be such that the super-quadrants will also be arranged in d-ary hypercube manner. (In the embodiment of the layout 100C of FIG. 1C, it is binary hypercube manner since $d = 2$, in the network $V_{fold-link}(N_1, N_2, d, s)$ 100B of FIG. 1B).

Layout 100D of FIG. 1D illustrates the inter-block links between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100D of FIG. 1D can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

Layout 100E of FIG. 1E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 3_4. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100E of FIG. 1E can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for

WO 2008/147928

PCT/US2008/064605

example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

5 Layout 100F of FIG. 1F illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated
10 in layout 100F of FIG. 1F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time
15 division multiplexed single track).

 Layout 100G of FIG. 1G illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of
20 Block 1_2 and switch 4 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100G of FIG. 1G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division
25 multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division multiplexed single track).

 The complete layout for the network 100B of FIG. 1B is given by combining the links in layout diagrams of 100C, 100D, 100E, 100F, and 100G. Applicant notes that in the layout 100C of FIG. 1C, the inter-block links between switch 1 and switch 2 of
30 corresponding blocks are vertical tracks as shown in layout 100D of FIG. 1D; the inter-

WO 2008/147928

PCT/US2008/064605

block links between switch 2 and switch 3 of corresponding blocks are horizontal tracks as shown in layout 100E of FIG. 1E; the inter-block links between switch 3 and switch 4 of corresponding blocks are vertical tracks as shown in layout 100F of FIG. 1F; and finally the inter-block links between switch 4 and switch 5 of corresponding blocks are horizontal tracks as shown in layout 100G of FIG. 1G. The pattern is alternate vertical tracks and horizontal tracks. It continues recursively for larger networks of $N > 32$ as will be illustrated later.

Some of the key aspects of the current invention are discussed. 1) All the switches in one row of the multi-stage network 100B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks or vertical tracks; 3) Since all the inter-block links are either horizontal or vertical tracks, all the inter-block links can be mapped on to island-style architectures in current commercial FPGA's; 4) The length of the longest wire is about half of the width (or length) of the complete layout (For example middle link ML(4,4) is about half the width of the complete layout).

In accordance with the current invention, the layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ the sub-quadrants, quadrants, and super-quadrants are arranged in d-ary hypercube manner and also the inter-blocks are accordingly connected in d-ary hypercube topology. Even though all the embodiments in the current invention are illustrated for $N_1 = N_2$, the embodiments can be extended for $N_1 \neq N_2$.

Referring to layout 100H of FIG. 1H, illustrates the extension of layout 100C for the network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 128$; $d = 2$; and $s = 2$. There are four super-quadrants in layout 100H namely top-left super-quadrant, bottom-left super-quadrant, top-right super-quadrant, bottom-right super-quadrant. Total number of blocks in the layout 100H is sixty four. Top-left super-quadrant implements the blocks from block 1_2 to block 31_32. Each block in all the super-quadrants has two more switches namely switch 6 and switch 7 in addition to the switches [1-5] illustrated in layout 100C of FIG. 1C. The inter-block link connection topology is the exactly the same between the

WO 2008/147928

PCT/US2008/064605

switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as it is shown in the layouts of FIG. 1D, FIG. 1E, FIG. 1F, and FIG. 1G respectively.

Bottom-left super-quadrant implements the blocks from block 33_34 to block 63_64. Top-right super-quadrant implements the blocks from block 65_66 to block 95_96. And bottom-right super-quadrant implements the blocks from block 97_98 to block 127_128. In all these three super-quadrants also, the inter-block link connection topology is the exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as that of the top-left super-quadrant.

10 Recursively in accordance with the current invention, the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-left super-quadrant and bottom-left super-quadrant. And similarly the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-right super-quadrant and bottom-right super-quadrant. The inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of top-left super-quadrant and top-right super-quadrant. And similarly the inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of bottom-left super-quadrant and bottom-right super-quadrant.

Referring to diagram 100I of FIG. 1I illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 100I illustrates both the intra-block and inter-block links connected to Block 1_2. The layout diagram 100I corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above.

WO 2008/147928

PCT/US2008/064605

That is the switches that are placed together in Block 1_2 as shown in FIG. 1I are namely input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switch
 5 MS(1,1) and middle switch MS(7,1) belonging to switch 2; middle switch MS(2,1) and middle switch MS(6,1) belonging to switch 3; middle switch MS(3,1) and middle switch MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1) – ML(1,4) being
 10 the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1) – ML(8,4) being the inputs of the output switch OS1 and outlet links OL1 – OL2 being the outputs of the output switch OS1.

Middle switch MS(1,1) is implemented as four by four switch with the middle links ML(1,1), ML(1,2), ML(1,7) and ML(1,8) being the inputs and middle links ML(2,1) – ML(2,4) being the outputs; and middle switch MS(7,1) is implemented as four by four
 15 switch with the middle links ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs and middle links ML(8,1) – ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 100I of FIG. 1I.

Now the VLSI layouts of generalized multi-link multi-stage network
 20 $V_{mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 < 32$; $d = 2$; $s = 2$ and its corresponding version of folded generalized multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 < 32$; $d = 2$; $s = 2$ are discussed. Referring to diagram 200A1 of FIG. 2A1 is generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 2$; $d = 2$. Diagram 200A2 of FIG. 2A2 illustrates the corresponding folded generalized multi-link multi-
 25 stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 2$; $d = 2$, version of the diagram 200A1 of FIG. 2A1. Layout 200A3 of FIG. 2A3 illustrates the VLSI layout of the network 200A2 of FIG. 2A2. There is only one block i.e., Block 1_2 comprising switch 1. Just like in the layout 100C of FIG. 1C, switch 1 consists of input switch IS1 and output switch OS1.

WO 2008/147928

PCT/US2008/064605

Referring to diagram 200B1 of FIG. 2B1 is generalized multi-link multi-stage network $V_{m-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 4$; $d = 2$; $s = 2$. Diagram 200B2 of FIG. 2B2 illustrates the corresponding folded generalized multi-link multi-stage network $V_{fold-m-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 4$; $d = 2$; $s = 2$, version of the diagram 200B1 of FIG. 2B1. Layout 200B3 of FIG. 2B3 illustrates the VLSI layout of the network 200B2 of FIG. 2B2. There are two blocks i.e., Block 1_2 and Block 3_4 each comprising switch 1 and switch 2. Switch 1 in each block consists of the corresponding input switch and output switch. For example switch 1 in Block 1_2 consists of input switch IS1 and output switch OS1. Similarly switch 2 in Block 1_2 consists of middle switch (1,1). Layout 200B4 of FIG. 2B4 illustrates the inter-block links of the VLSI layout diagram 200B3 of FIG. 2B3. For example middle links ML(1,4) and ML(2,8). It must be noted that all the inter-block links are vertical tracks in this layout. (Alternatively all the inter-blocks can also be implemented as horizontal tracks).

Referring to diagram 200C11 of FIG. 2C11 is generalized multi-link multi-stage network $V_{m-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 8$; $d = 2$; $s = 2$. Diagram 200C12 of FIG. 2C12 illustrates the corresponding folded generalized multi-link multi-stage network $V_{fold-m-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 8$; $d = 2$; $s = 2$, version of the diagram 200C11 of FIG. 2C11. Layout 200C21 of FIG. 2C21 illustrates the VLSI layout of the network 200C12 of FIG. 2C12. There are four blocks i.e., Block 1_2, Block 3_4, Block 5_6, and Block 7_8 each comprising switch 1, switch 2 and switch 3. For example switch 1 in Block 1_2 consists of input switch IS1 and output switch OS1; Switch 2 in Block 1_2 consists of MS(1,1) and MS(3,1). Switch 3 in Block 1_2 consists of MS(2,1).

Layout 200C22 of FIG. 2C22 illustrates the inter-block links between the switch 1 and switch 2 of the VLSI layout diagram 200C21 of FIG. 2C21. For example middle links ML(1,4) and ML(4,8) are connected between Block 1_2 and Block 3_4. It must be noted that all the inter-block links between switch 1 and switch 2 of all blocks are vertical tracks in this layout. Layout 200C23 of FIG. 2C23 illustrates the inter-block links between the switch 2 and switch 3 of the VLSI layout diagram 200C21 of FIG. 2C21. For example middle links ML(2,12) and ML(3,4) are connected between Block 1_2 and

WO 2008/147928

PCT/US2008/064605

Block 5_6. It must be noted that all the inter-block links between switch 2 and switch 3 of all blocks are horizontal tracks in this layout

Referring to diagram 200D1 of FIG. 2D1 is generalized multi-link multi-stage network $V_{m-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 16$; $d = 2$; $s = 2$. Diagram 200D2 of FIG. 2D2 illustrates the corresponding folded generalized multi-link multi-stage network $V_{fold-m-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 16$; $d = 2$; $s = 2$, version of the diagram 200D1 of FIG. 2D1. Layout 200D3 of FIG. 2D3 illustrates the VLSI layout of the network 200D2 of FIG. 2D2. There are eight blocks i.e., Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14 and Block 15_16 each comprising switch 1, switch 2, switch 3 and switch 4. For example switch 1 in Block 1_2 consists of input switch IS1 and output switch OS1; Switch 2 in Block 1_2 consists of MS(1,1) and MS(5,1). Switch 3 in Block 1_2 consists of MS(2,1) and MS(4,1), and switch 4 in Block 1_2 consists of MS(3,1).

Layout 200D4 of FIG. 2D4 illustrates the inter-block links between the switch 1 and switch 2 of the VLSI layout diagram 200D3 of FIG. 2D3. For example middle links ML(1,4) and ML(6,8) are connected between Block 1_2 and Block 3_4. It must be noted that all the inter-block links between switch 1 and switch 2 of all blocks are vertical tracks in this layout. Layout 200D5 of FIG. 2D5 illustrates the inter-block links between the switch 2 and switch 3 of the VLSI layout diagram 200D3 of FIG. 2D3. For example middle links ML(2,12) and ML(5,4) are connected between Block 1_2 and Block 5_6. It must be noted that all the inter-block links between switch 2 and switch 3 of all blocks are horizontal tracks in this layout. Layout 200D6 of FIG. 2D6 illustrates the inter-block links between the switch 3 and switch 4 of the VLSI layout diagram 200D3 of FIG. 2D3. For example middle links ML(3,4) and ML(4,20) are connected between Block 1_2 and Block 9_10. It must be noted that all the inter-block links between switch 3 and switch 4 of all blocks are vertical tracks in this layout.

Generalized Multi-link Butterfly Fat Tree Network Embodiment:

In another embodiment in the network 100B of FIG. 1B, the switches that are placed together are implemented as combined switch then the network 100B is the

WO 2008/147928

PCT/US2008/064605

generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,390 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch. For example the input switch IS1 and output switch OS1 are placed together; so input switch IS1 and output OS1 are implemented as a six by six switch with the inlet links IL1, IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and OL2 being the outputs of the combined switch IS1&OS1. Similarly in this embodiment of network 100B all the switches that are placed together are implemented as a combined switch.

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$.

Referring to diagram 100J of FIG. 1J illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 100J illustrates both the intra-block and inter-block links. The layout diagram 100J corresponds to the embodiment where the switches that are placed together are implemented as combined switch in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,390 that is incorporated by reference above.

WO 2008/147928

PCT/US2008/064605

That is the switches that are placed together in Block 1_2 as shown in FIG. 1J are namely the combined input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switch implemented is combined input and output switch IS1&OS1); middle switch
 5 MS(1,1) belonging to switch 2; middle switch MS(2,1) belonging to switch 3; middle switch MS(3,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links IL1, IL2 and ML(8,1) – ML(8,4) being the inputs and middle links ML(1,1) – ML(1,4), and outlet links OL1 – OL2 being the outputs.

10 Middle switch MS(1,1) is implemented as eight by eight switch with the middle links ML(1,1), ML(1,2), ML(1,7), ML(1,8), ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs and middle links ML(2,1) – ML(2,4) and middle links ML(8,1) – ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as eight by eight switches as illustrated in 100J of FIG. 1J. Applicant observes that in
 15 middle switch MS(1,1) any one of the right going middle links can be switched to any one of the left going middle links and hereinafter middle switch MS(1,1) provides U-turn links. In general, in the network $V_{mlink-bft}(N_1, N_2, d, s)$ each input switch, each output switch and each middle switch provides U-turn links.

In another embodiment, middle switch MS(1,1) (or the middle switches in any of
 20 the middle stage excepting the root middle stage) of Block 1_2 of

$V_{mlink-bft}(N_1, N_2, d, s)$ can be implemented as a four by eight switch and a four by four switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1J, the left going middle links namely ML(7,1),
 25 ML(7,2), ML(7,11), and ML(7,12) are never switched to the right going middle links ML(2,1), ML(2,2), ML(2,3), and ML(2,4). And hence to implement MS(1,1) two switches namely: 1) a four by eight switch with the middle links ML(1,1), ML(1,2), ML(1,7), and ML(1,8) as inputs and the middle links ML(2,1), ML(2,2), ML(2,3), ML(2,4), ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs and 2) a four by four

WO 2008/147928

PCT/US2008/064605

switch with the middle links ML(7,1), ML(7,2), ML(7,11), and ML(7,12) as inputs and the middle links ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs are sufficient without losing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

5 Generalized multi-stage network Embodiment:

In one embodiment, in the network 100B of FIG. 1B, the switches that are placed together are implemented as two separate switches in input stage 110 and output stage 120; and as four separate switches in all the middle stages, then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch and a four by two switch respectively. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,4), ML(8,7) and ML(8,8) being the inputs and outlet links OL1 – OL2 being the outputs.

The switches, corresponding to the middle stages that are placed together are implemented as four two by two switches. For example middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs;

WO 2008/147928

PCT/US2008/064605

Similarly in this embodiment of network 100B all the switches that are placed together are implemented as separate switches.

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized folded multi-stage network

- 5 $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$.

- Referring to diagram 100K of FIG. 1K illustrates a high-level implementation of
10 Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized folded multi-stage network

- $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 100K illustrates both the intra-block and inter-block links. The layout diagram 100K corresponds to the embodiment where the switches that are placed together are implemented as separate
15 switches in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that is incorporated by reference above.

- That is the switches that are placed together in Block 1_2 as shown in FIG. 1K are
20 namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches MS(1,1), MS(1,17), MS(7,1) and MS(7,17) belonging to switch 2; middle switches MS(2,1), MS(2,17), MS(6,1) and MS(6,17) belonging to switch 3; middle switches
25 MS(3,1), MS(3,17), MS(5,1) and MS(5,17) belonging to switch 4; And middle switches MS(4,1), and MS(4,17) belonging to switch 5.

Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and

WO 2008/147928

PCT/US2008/064605

middle links ML(1,1) – ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,4), ML(8,7) and ML(8,8) being the inputs and outlet links OL1 – OL2 being the outputs.

Middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as illustrated in 100K of FIG. 1K.

15 **Generalized multi-stage network Embodiment with S = 1:**

In one embodiment, in the network 100B of FIG. 1B (where it is implemented with $s = 1$), the switches that are placed together are implemented as two separate switches in input stage 110 and output stage 120; and as two separate switches in all the middle stages, then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by two switch and a two by two switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1 – OL2 being the outputs.

WO 2008/147928

PCT/US2008/064605

The switches, corresponding to the middle stages that are placed together are implemented as two, two by two switches. For example middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links
 5 ML(2,1) and ML(2,2) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs; Similarly in this embodiment of network 100B all the switches that are placed together are implemented as two separate switches.

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in
 10 FIG. 1G are also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$.

15 Referring to diagram 100K1 of FIG. 1K1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 100C of FIG. 1C when $s = 1$ which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ (All the double links are replaced by single links when $s = 1$). Block 1_2 in 100K1 illustrates both the intra-block and inter-
 20 block links. The layout diagram 100K1 corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 100B of FIG. 1B when $s = 1$. As noted before then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that
 25 is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 1K1 are namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches

WO 2008/147928

PCT/US2008/064605

MS(1,1) and MS(7,1) belonging to switch 2; middle switches MS(2,1) and MS(6,1) belonging to switch 3; middle switches MS(3,1) and MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is
 5 implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1 – OL2 being the outputs.

Middle switches MS(1,1) and MS(7,1) are placed together; so middle switch
 10 MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; And middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as
 15 illustrated in 100K1 of FIG. 1K1.

Generalized Butterfly Fat Tree Network Embodiment:

In another embodiment in the network 100B of FIG. 1B, the switches that are placed together are implemented as two combined switches then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and s
 20 $= 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch. For example the input switch IS1 and output switch OS1 are placed together; so input
 25 output switch IS1&OS1 are implemented as a six by six switch with the inlet links IL1, IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and OL2 being the outputs of the combined switch IS1&OS1.

WO 2008/147928

PCT/US2008/064605

The switches, corresponding to the middle stages that are placed together are implemented as two four by four switches. For example middle switches MS(1,1) and MS(1,17) are placed together; so middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; middle switch MS(1,17) is implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly in this embodiment of network 100B all the switches that are placed together are implemented as a two combined switches.

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$.

Referring to diagram 100L of FIG. 1L illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 100L illustrates both the intra-block and inter-block links. The layout diagram 100L corresponds to the embodiment where the switches that are placed together are implemented as two combined switches in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 1L are namely the combined input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switch implemented is combined input and output switch IS1&OS1); middle switch

WO 2008/147928

PCT/US2008/064605

MS(1,1) and MS(1,17) belonging to switch 2; middle switch MS(2,1) and MS(2,17) belonging to switch 3; middle switch MS(3,1) and MS(3,17) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

5 Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links IL1, IL2, ML(8,1) , ML(8,2), ML(8,7) and ML(8,8) being the inputs and middle links ML(1,1) – ML(1,4) and outlet links OL1 – OL2 being the outputs.

Middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; And middle switch MS(1,17) is
 10 implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two four by four switches as illustrated in 100L of FIG. 1L. Applicant observes that in middle switch MS(1,1) any one of the right going middle links can be switched to any one of the
 15 left going middle links and hereinafter middle switch MS(1,1) provides U-turn links. In general, in the network $V_{bfi}(N_1, N_2, d, s)$ each input switch, each output switch and each middle switch provides U-turn links.

In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of $V_{bfi}(N_1, N_2, d, s)$ can
 20 be implemented as a two by four switch and a two by two switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1L, the left going middle links namely ML(7,1) and ML(7,11) are never switched to the right going middle links ML(2,1) and ML(2,3). And hence to implement MS(1,1)
 25 two switches namely: 1) a two by four switch with the middle links ML(1,1) and ML(1,7) as inputs and the middle links ML(2,1), ML(2,3), ML(8,1), and ML(8,3) as outputs and 2) a two by two switch with the middle links ML(7,1) and ML(7,11) as inputs and the middle links ML(8,1) and ML(8,3) as outputs are sufficient without loosing any

WO 2008/147928

PCT/US2008/064605

connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

Generalized Butterfly Fat Tree Network Embodiment with $S = 1$:

5 In one embodiment, in the network 100B of FIG. 1B (where it is implemented with $s = 1$), the switches that are placed together are implemented as a combined switch in input stage 110 and output stage 120; and as a combined switch in all the middle stages, then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with five stages as disclosed in

10 U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a four by four switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input and output switch IS1&OS1 is implemented as four by four switch with the inlet links IL1, IL2, ML(8,1)

15 and ML(8,3) being the inputs and middle links ML(1,1) – ML(1,2) and outlet links OL1 – OL2 being the outputs

The switches, corresponding to the middle stages that are placed together are implemented as a four by four switch. For example middle switches MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,3), ML(7,1) and

20 ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs..

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with five stages. The layout 100C

25 in FIG. 1C can be recursively extended for any arbitrarily large generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$.

WO 2008/147928

PCT/US2008/064605

Referring to diagram 100L1 of FIG. 1L1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 100C of FIG. 1C when $s = 1$ which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ (All the double links are replaced by single links when $s = 1$). Block 1_2 in 100K1 illustrates both the intra-block and inter-block links. The layout diagram 100L1 corresponds to the embodiment where the switches that are placed together are implemented as a combined switch in the network 100B of FIG. 1B when $s = 1$. As noted before then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 1L1 are namely the input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switch MS(1,1) belonging to switch 2; middle switch MS(2,1) belonging to switch 3; middle switch MS(3,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Input and output switch IS1&OS1 are placed together; so input and output switch IS1&OS1 is implemented as four by four switch with the inlet links IL1, IL2, ML(8,1) and ML(8,3) being the inputs and middle links ML(1,1) – ML(1,2) and outlet links OL1 – OL2 being the outputs.

Middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,3), ML(7,1) and ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 100L1 of FIG. 1L1.

In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of

$V_{mink-bft}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two

WO 2008/147928

PCT/US2008/064605

switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1L1, the left going middle links namely ML(7,1) and ML(7,5) are never switched to the right going middle links ML(2,1) and ML(2,2).

5 And hence to implement MS(1,1) two switches namely: 1) a two by four switch with the middle links ML(1,1) and ML(1,3) as inputs and the middle links ML(2,1), ML(2,2), ML(8,1), and ML(8,2) as outputs and 2) a two by two switch with the middle links ML(7,1) and ML(7,5) as inputs and the middle links ML(8,1) and ML(8,2) as outputs are sufficient without loosing any connectivity of the embodiment of MS(1,1) being

10 implemented as an eight by eight switch as described before.)

Hypercube-like Topology layout schemes:

Referring to diagram 300A in FIG. 3A, in one embodiment, an exemplary generalized multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d =$

15 2; and $s = 2$ with nine stages of one hundred and forty four switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, 150, 170, 180 and 190 is shown where input stage 110 consists of sixteen, two by four switches IS1-IS16 and output stage 120 consists of

20 sixteen, four by two switches OS1-OS16.

As disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above, such a network can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections.

25 The diagram 300A in FIG. 3A is exactly the same as the diagram 100A in FIG. 1A excepting the connection links between middle stage 150 and middle stage 160 as well as between middle stage 160 and middle stage 170.

WO 2008/147928

PCT/US2008/064605

Each of the $\frac{N}{d}$ middle switches are connected to exactly d switches in middle stage 160 through two links each for a total of $2 \times d$ links (for example the links ML(4,1) and ML(4,2) are connected from middle switch MS(3,1) to middle switch MS(4,1), and the links ML(4,3) and ML(4,4) are connected from middle switch MS(3,1) to middle switch MS(4,15)).

Each of the $\frac{N}{d}$ middle switches MS(4,1) – MS(4,16) in the middle stage 160 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(4,1) and ML(4,2) are connected to the middle switch MS(4,1) from input switch MS(3,1), and the links ML(4,59) and ML(4,60) are connected to the middle switch MS(4,1) from input switch MS(3,15)) and also are connected to exactly d switches in middle stage 170 through two links each for a total of $2 \times d$ links (for example the links ML(5,1) and ML(5,2) are connected from middle switch MS(4,1) to middle switch MS(5,1), and the links ML(5,3) and ML(5,4) are connected from middle switch MS(4,1) to middle switch MS(5,15)).

Each of the $\frac{N}{d}$ middle switches MS(5,1) – MS(5,16) in the middle stage 170 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(5,1) and ML(5,2) are connected to the middle switch MS(5,1) from input switch MS(4,1), and the links ML(5,59) and ML(5,60) are connected to the middle switch MS(5,1) from input switch MS(4,15)).

Finally the connection topology of the network 100A shown in FIG. 1A is also basically back to back inverse Benes connection topology but with a slight variation. All the cross middle links from middle switches MS(3,1) - MS(3,8) connect to middle switches MS(4,9) – MS(4,16) and all the cross middle links from middle switches MS(3,9) - MS(3,16) connect to middle switches MS(4,1) – MS(4,8). Applicant makes a key observation that there are many combinations of connections possible using this property. The difference in the connection topology between diagram 100A of FIG. 1A and diagram 300A of FIG. 3A is that the connections formed by cross middle links

WO 2008/147928

PCT/US2008/064605

between middle stage 150 and middle stage 160 are made of two different combinations otherwise both the diagrams 100A and 300A implement back to back inverse Benes connection topology. Since these networks implement back to back inverse Benes topologies since there is difference in the connections of cross middle links between
5 middle stage 150 and middle stage 160, the same difference in the connections of cross middle links between 160 and middle stage 170 occurs.

Referring to diagram 300B in FIG. 3B, is a folded version of the multi-link multi-stage network 300A shown in FIG. 3A. The network 300B in FIG. 3B shows input stage 110 and output stage 120 are placed together. That is input switch IS1 and output switch
10 OS1 are placed together, input switch IS2 and output switch OS2 are placed together, and similarly input switch IS16 and output switch OS16 are placed together. All the right going middle links {i.e., inlet links IL1 – IL32 and middle links ML(1,1) - ML(1,64)} correspond to input switches IS1 - IS16, and all the left going middle links {i.e., middle links ML(7,1) - ML(7,64) and outlet links OL1-OL32} correspond to output switches
15 OS1 - OS16.

Just the same way there is difference in the connection topology between diagram 100A of FIG. 1A and diagram 300A of FIG. 3A in the way the connections are formed by cross middle links between middle stage 150 and middle stage 160 and also between
middle stage 160 and middle stage 170, the exact similar difference is there between the
20 diagram 100B of FIG. 1B and the diagram 300B of FIG. 3B, i.e., in the way the connections are formed by cross middle links between middle stage 150 and middle stage 160 and also between middle stage 160 and middle stage 170.

In one embodiment, in the network 300B of FIG. 3B, the switches that are placed together are implemented as separate switches then the network 300B is the generalized
25 folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch and a four by two switch. For example the switch input switch IS1 and output switch OS1
30 are placed together; so input switch IS1 is implemented as two by four switch with the

WO 2008/147928

PCT/US2008/064605

inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1) – ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1 – OL2 being the
5 outputs of the output switch OS1. Similarly in this embodiment of network 300B all the switches that are placed together are implemented as separate switches.

Referring to layout 300C of FIG. 3C, in one embodiment, there are sixteen blocks namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, Block 23_24, Block
10 25_26, Block 27_28, Block 29_30, and Block 31_32. Each block implements all the switches in one row of the network 300B of FIG. 3B, one of the key aspects of the current invention. For example Block 1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1).
15 For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; And middle switch MS(4,1) is denoted by switch 5.

All the straight middle links are illustrated in layout 300C of FIG. 3C. For example in Block 1_2, inlet links IL1 – IL2, outlet links OL1 – OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link
25 ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 300C of FIG. 3C.

Even though it is not illustrated in layout 300C of FIG. 3C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit or sub-integrated circuit block depending on the applications in different
30 embodiments. There are four quadrants in the layout 300C of FIG. 3C namely top-left,

WO 2008/147928

PCT/US2008/064605

bottom-left, top-right and bottom-right quadrants. Top-left quadrant implements Block 1_2, Block 3_4, Block 5_6, and Block 7_8. Bottom-left quadrant implements Block 9_10, Block 11_12, Block 13_14, and Block 15_16. Top-right quadrant implements Block 25_26, Block 27_28, Block 29_30, and Block 31_32. Bottom-right quadrant implements Block 17_18, Block 19_20, Block 21_22, and Block 23_24. There are two halves in layout 300C of FIG. 3C namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants. Right-half consists of top-right and bottom-right quadrants.

10 Recursively in each quadrant there are four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements Block 1_2. Bottom-left sub-quadrant of top-left quadrant implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 7_8. Finally bottom-right sub-quadrant of top-left quadrant implements Block 5_6.

15 Similarly there are two sub-halves in each quadrant. For example in top-left quadrant there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left quadrant implements Block 1_2 and Block 3_4. Right-sub-half of top-left quadrant implements Block 7_8 and Block 5_6. Recursively in larger multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 > 32$, the layout in this embodiment in accordance

20 with the current invention, will be such that the super-quadrants will also be arranged in a similar manner.

Layout 300D of FIG. 3D illustrates the inter-block links (in the layout 300C of FIG. 3C all the cross middle links are inter-block links) between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected

25 between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100D of FIG. 1D can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example

30 middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an

WO 2008/147928

PCT/US2008/064605

alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

Layout 300E of FIG. 3E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 3_4. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 300E of FIG. 3E can be implemented as diagonal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

Layout 300F of FIG. 3F illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 300F of FIG. 3F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time division multiplexed single track).

Layout 300G of FIG. 3G illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 3_4. Applicant notes that the inter-block links illustrated

WO 2008/147928

PCT/US2008/064605

in layout 300G of FIG. 3G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division
5 multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division multiplexed single track).

The complete layout for the network 300B of FIG. 3B is given by combining the links in layout diagrams of 300C, 300D, 300E, 300F, and 300G. Applicant notes that in the layout 300C of FIG. 3C, the inter-block links between switch 1 and switch 2 are
10 vertical tracks as shown in layout 300D of FIG. 3D; the inter-block links between switch 2 and switch 3 are horizontal tracks as shown in layout 300E of FIG. 3E; the inter-block links between switch 3 and switch 4 are vertical tracks as shown in layout 300F of FIG. 3F; and finally the inter-block links between switch 4 and switch 5 are horizontal tracks as shown in layout 300G of FIG. 3G. The pattern is either vertical tracks, horizontal
15 tracks or diagonal tracks. It continues recursively for larger networks of $N > 32$ as will be illustrated later.

Some of the key aspects of the current invention related to layout diagram 300C of IFG. 3C are noted. 1) All the switches in one row of the multi-stage network 300B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-
20 block links are either horizontal tracks, vertical tracks or diagonal tracks; 3) The length of the longest wire is about half of the width (or length) of the complete layout (For example middle link ML(4,4) is about half the width of the complete layout.);

The layout 300C in FIG. 3C can be recursively extended for any arbitrarily large generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$. Referring to
25 layout 300H of FIG. 3H, illustrates the extension of layout 300C for the network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 128$; $d = 2$; and $s = 2$. There are four super-quadrants in layout 300H namely top-left super-quadrant, bottom-left super-quadrant, top-right super-quadrant, bottom-right super-quadrant. Total number of blocks in the layout 300H is sixty four. Top-left super-quadrant implements the blocks from block 1_2

WO 2008/147928

PCT/US2008/064605

to block 31_32. Each block in all the super-quadrants has two more switches namely switch 6 and switch 7 in addition to the switches [1-5] illustrated in layout 300C of FIG. 3C. The inter-block link connection topology is the exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as it is shown in the layouts of FIG. 3D, FIG. 3E, FIG. 3F, and FIG. 3G respectively.

Bottom-left super-quadrant implements the blocks from block 33_34 to block 63_64. Top-right super-quadrant implements the blocks from block 65_66 to block 95_96. And bottom-right super-quadrant implements the blocks from block 97_98 to block 127_128. In all these three super-quadrants also, the inter-block link connection topology is the exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as that of the top-left super-quadrant.

Recursively in accordance with the current invention, the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-left super-quadrant and bottom-left super-quadrant. And similarly the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-right super-quadrant and bottom-right super-quadrant. The inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of top-left super-quadrant and top-right super-quadrant. And similarly the inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of bottom-left super-quadrant and bottom-right super-quadrant.

Ring Topology layout schemes:

Layout diagram 400C of FIG. 4C is another embodiment for the generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ diagram 100B in FIG. 1B.

Referring to layout 400C of FIG. 4C, there are sixteen blocks namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, Block 23_24, Block 25_26, Block 27_28,

WO 2008/147928

PCT/US2008/064605

Block 29_30, and Block 31_32. Each block implements all the switches in one row of the network 100B of FIG. 1B, one of the key aspects of the current invention. For example Block 1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; And middle switch MS(4,1) is denoted by switch 5.

All the straight middle links are illustrated in layout 400C of FIG. 4C. For example in Block 1_2, inlet links IL1 – IL2, outlet links OL1 – OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 400C of FIG. 4C.

Even though it is not illustrated in layout 400C of FIG. 4C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit or sub-integrated circuit block depending on the applications in different embodiments. The topology of the layout 400C in FIG. 4C is a ring. For each of the neighboring rows in diagram 100B of FIG. 1B the corresponding blocks are also physically neighbors in layout diagram 400C of FIG. 4C. In addition the topmost row is also logically considered as neighbor to the bottommost row. For example Block 1_2 (implementing the switches belonging to a row in diagram 100B of FIG. 1B) has Block 3_4 as neighbor since Block 3_4 implements the switches in its neighboring row. Similarly Block 1_2 also has Block 31_32 as neighbor since Block 1_2 implements topmost row of switches and Block 31_32 implements bottommost row of switches in diagram 100B of FIG. 1B. The ring layout scheme illustrated in 400C of FIG. 4C can be

WO 2008/147928

PCT/US2008/064605

generalized for a large multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 > 32$, in accordance with the current invention.

Layout 400B of FIG. 4B illustrates the inter-block links (in the layout 400A of FIG. 4A all the cross middle links are inter-block links) between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 400B of FIG. 4B are implemented as vertical tracks or horizontal tracks or diagonal tracks. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

Layout 400C of FIG. 4C illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 3_4. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 400C of FIG. 4C are implemented as vertical tracks or horizontal tracks or diagonal tracks. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

Layout 400D of FIG. 4D illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated

WO 2008/147928

PCT/US2008/064605

in layout 400D of FIG. 4D are implemented as vertical tracks or horizontal tracks or diagonal tracks. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time division multiplexed single track).

Layout 400E of FIG. 4E illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 400E of FIG. 4E are implemented as vertical tracks or horizontal tracks or diagonal tracks. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division multiplexed single track).

The complete layout for the network 100B of FIG. 1B is given by combining the links in layout diagrams of 400A, 400B, 400C, 400D, and 400E.

Some of the key aspects of the current invention related to layout diagram 400A of FIG. 4A are noted. 1) All the switches in one row of the multi-stage network 100B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks, vertical tracks or diagonal tracks; 3) Length of the different wires between the same two middle stages is not the same. However it gives an opportunity to implement the most connected circuits to place and route through the blocks which have shorter wires.

Layout diagram 400C1 of FIG. 4C1 is another embodiment for the generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ diagram 100B in FIG. 1B. Referring to layout 400C1 of FIG. 4C1, there are sixteen blocks namely Block 1_2, Block

WO 2008/147928

PCT/US2008/064605

3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, Block 23_24, Block 25_26, Block 27_28, Block 29_30, and Block 31_32. Each block implements all the switches in one row of the network 100B of FIG. 1B, one of the key aspects of the current invention. For example

5 Block 1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2;

10 Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; And middle switch MS(4,1) is denoted by switch 5.

All the straight middle links are illustrated in layout 400C1 of FIG. 4C1. For example in Block 1_2, inlet links IL1 – IL2, outlet links OL1 – OL2, middle link

15 ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 400C1 of FIG. 4C1.

20 Even though it is not illustrated in layout 400C1 of FIG. 4C1, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit or sub-integrated circuit block depending on the applications in different embodiments. The topology of the layout 400C1 in FIG. 4C1 is another embodiment of ring layout topology. For each of the neighboring rows in diagram 100B of FIG. 1B the

25 corresponding blocks are also physically neighbors in layout diagram 400C of FIG. 4C. In addition the topmost row is also logically considered as neighbor to the bottommost row. For example Block 1_2 (implementing the switches belonging to a row in diagram 100B of FIG. 1B) has Block 3_4 as neighbor since Block 3_4 implements the switches in its neighboring row. Similarly Block 1_2 also has Block 31_32 as neighbor since Block

30 1_2 implements topmost row of switches and Block 31_32 implements bottommost row

WO 2008/147928

PCT/US2008/064605

of switches in diagram 100B of FIG. 1B. The ring layout scheme illustrated in 400C of FIG. 4C can be generalized for a large multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 > 32$, in accordance with the current invention.

All the layout embodiments disclosed in the current invention are applicable to
 5 generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$
 10 for $s = 1, 2, 3$ or any number in general, and for both $N_1 = N_2 = N$ and $N_1 \neq N_2$, and d is any integer.

Conversely applicant makes another important observation that generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ are implemented with the layout topology being the hypercube topology shown in layout 100C of FIG. 1C with large scale cross point
 15 reduction as any one of the networks described in the current invention namely: generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree
 20 networks $V_{mlink-bft}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general, and for both $N_1 = N_2 = N$ and $N_1 \neq N_2$, and d is any integer.

Applications Embodiments:

All the embodiments disclosed in the current invention are useful in many
 25 varieties of applications. FIG. 5A1 illustrates the diagram of 500A1 which is a typical two by two switch with two inlet links namely IL1 and IL2, and two outlet links namely

WO 2008/147928

PCT/US2008/064605

OL1 and OL2. The two by two switch also implements four crosspoints namely CP(1,1), CP(1,2), CP(2,1) and CP(2,2) as illustrated in FIG. 5A1. For example the diagram of 500A1 may be the implementation of middle switch MS(1,1) of the diagram 100K of FIG. 1K where inlet link IL1 of diagram 500A1 corresponds to middle link ML(1,1) of diagram 100K, inlet link IL2 of diagram 500A1 corresponds to middle link ML(1,7) of diagram 100K, outlet link OL1 of diagram 500A1 corresponds to middle link ML(2,1) of diagram 100K, outlet link OL2 of diagram 500A1 corresponds to middle link ML(2,3) of diagram 100K.

10 1) Programmable Integrated Circuit Embodiments:

All the embodiments disclosed in the current invention are useful in programmable integrated circuit applications. FIG. 5A2 illustrates the detailed diagram 500A2 for the implementation of the diagram 500A1 in programmable integrated circuit embodiments. Each crosspoint is implemented by a transistor coupled between the corresponding inlet link and outlet link, and a programmable cell in programmable integrated circuit embodiments. Specifically crosspoint CP(1,1) is implemented by transistor C(1,1) coupled between inlet link IL1 and outlet link OL1, and programmable cell P(1,1); crosspoint CP(1,2) is implemented by transistor C(1,2) coupled between inlet link IL1 and outlet link OL2, and programmable cell P(1,2); crosspoint CP(2,1) is implemented by transistor C(2,1) coupled between inlet link IL2 and outlet link OL1, and programmable cell P(2,1); and crosspoint CP(2,2) is implemented by transistor C(2,2) coupled between inlet link IL2 and outlet link OL2, and programmable cell P(2,2).

If the programmable cell is programmed ON, the corresponding transistor couples the corresponding inlet link and outlet link. If the programmable cell is programmed OFF, the corresponding inlet link and outlet link are not connected. For example if the programmable cell P(1,1) is programmed ON, the corresponding transistor C(1,1) couples the corresponding inlet link IL1 and outlet link OL1. If the programmable cell P(1,1) is programmed OFF, the corresponding inlet link IL1 and outlet link OL1 are not connected. In volatile programmable integrated circuit embodiments the programmable

WO 2008/147928

PCT/US2008/064605

cell may be an SRAM (Static Random Address Memory) cell. In non-volatile programmable integrated circuit embodiments the programmable cell may be a Flash memory cell. Also the programmable integrated circuit embodiments may implement field programmable logic arrays (FPGA) devices, or programmable Logic devices (PLD),
5 or Application Specific Integrated Circuits (ASIC) embedded with programmable logic circuits or 3D-FPGAs.

FIG. 5A2 also illustrates a buffer B1 on inlet link IL2. The signals driven along inlet link IL2 are amplified by buffer B1. Buffer B1 can be inverting or non-inverting buffer. Buffers such as B1 are used to amplify the signal in links which are usually long.

10 2) One-time Programmable Integrated Circuit Embodiments:

All the embodiments disclosed in the current invention are useful in one-time programmable integrated circuit applications. FIG. 5A3 illustrates the detailed diagram 500A3 for the implementation of the diagram 500A1 in one-time programmable integrated circuit embodiments. Each crosspoint is implemented by a via coupled
15 between the corresponding inlet link and outlet link in one-time programmable integrated circuit embodiments. Specifically crosspoint CP(1,1) is implemented by via V(1,1) coupled between inlet link IL1 and outlet link OL1; crosspoint CP(1,2) is implemented by via V(1,2) coupled between inlet link IL1 and outlet link OL2; crosspoint CP(2,1) is implemented by via V(2,1) coupled between inlet link IL2 and outlet link OL1; and
20 crosspoint CP(2,2) is implemented by via V(2,2) coupled between inlet link IL2 and outlet link OL2.

If the via is programmed ON, the corresponding inlet link and outlet link are permanently connected which is denoted by thick circle at the intersection of inlet link and outlet link. If the via is programmed OFF, the corresponding inlet link and outlet link
25 are not connected which is denoted by the absence of thick circle at the intersection of inlet link and outlet link. For example in the diagram 500A3 the via V(1,1) is programmed ON, and the corresponding inlet link IL1 and outlet link OL1 are connected as denoted by thick circle at the intersection of inlet link IL1 and outlet link OL1; the via V(2,2) is programmed ON, and the corresponding inlet link IL2 and outlet link OL2 are

WO 2008/147928

PCT/US2008/064605

connected as denoted by thick circle at the intersection of inlet link IL2 and outlet link OL2; the via V(1,2) is programmed OFF, and the corresponding inlet link IL1 and outlet link OL2 are not connected as denoted by the absence of thick circle at the intersection of inlet link IL1 and outlet link OL2; the via V(2,1) is programmed OFF, and the

5 corresponding inlet link IL2 and outlet link OL1 are not connected as denoted by the absence of thick circle at the intersection of inlet link IL2 and outlet link OL1. One-time programmable integrated circuit embodiments may be anti-fuse based programmable integrated circuit devices or mask programmable structured ASIC devices.

3) Integrated Circuit Placement and Route Embodiments:

10 All the embodiments disclosed in the current invention are useful in Integrated Circuit Placement and Route applications, for example in ASIC backend Placement and Route tools. FIG. 5A4 illustrates the detailed diagram 500A4 for the implementation of the diagram 500A1 in Integrated Circuit Placement and Route embodiments. In an integrated circuit since the connections are known a-priori, the switch and crosspoints are

15 actually virtual. However the concept of virtual switch and virtual crosspoint using the embodiments disclosed in the current invention reduces the number of required wires, wire length needed to connect the inputs and outputs of different netlists and the time required by the tool for placement and route of netlists in the integrated circuit.

Each virtual crosspoint is used to either to hardwire or provide no connectivity

20 between the corresponding inlet link and outlet link. Specifically crosspoint CP(1,1) is implemented by direct connect point DCP(1,1) to hardwire (i.e., to permanently connect) inlet link IL1 and outlet link OL1 which is denoted by the thick circle at the intersection of inlet link IL1 and outlet link OL1; crosspoint CP(2,2) is implemented by direct connect point DCP(2,2) to hardwire inlet link IL2 and outlet link OL2 which is denoted

25 by the thick circle at the intersection of inlet link IL2 and outlet link OL2. The diagram 500A4 does not show direct connect point DCP(1,2) and direct connect point DCP(1,3) since they are not needed and in the hardware implementation they are eliminated. Alternatively inlet link IL1 needs to be connected to outlet link OL1 and inlet link IL1 does not need to be connected to outlet link OL2. Also inlet link IL2 needs to be

30 connected to outlet link OL2 and inlet link IL2 does not need to be connected to outlet

WO 2008/147928

PCT/US2008/064605

link OL1. Furthermore in the example of the diagram 500A4, there is no need to drive the signal of inlet link IL1 horizontally beyond outlet link OL1 and hence the inlet link IL1 is not even extended horizontally until the outlet link OL2. Also the absence of direct connect point DCP(2,1) illustrates there is no need to connect inlet link IL2 and outlet
5 link OL1.

In summary in integrated circuit placement and route tools, the concept of virtual switches and virtual cross points is used during the implementation of the placement & routing algorithmically in software, however during the hardware implementation cross points in the cross state are implemented as hardwired connections between the
10 corresponding inlet link and outlet link, and in the bar state are implemented as no connection between inlet link and outlet link.

3) More Application Embodiments:

All the embodiments disclosed in the current invention are also useful in the
15 design of SoC interconnects, Field programmable interconnect chips, parallel computer systems and in time-space-time switches.

Numerous modifications and adaptations of the embodiments, implementations, and examples described herein will be apparent to the skilled artisan in view of the disclosure.

WO 2008/147928

PCT/US2008/064605

CLAIMS

What is claimed is:

1. An integrated circuit device comprising a plurality of sub-integrated circuit blocks and a routing network, and
- 5 Said each plurality of sub-integrated circuit blocks comprising a plurality of inlet links and a plurality of outlet links; and
Said routing network interconnects any one of said outlet link of one of said sub-integrated circuit block to one or more said inlet links of one or more of said sub-integrated circuit blocks; and
- 10 Said routing network comprising of a plurality of stages y , starting from the lowest stage to the highest stage; and
Said routing network comprising a plurality of switches of size $d \times d$, where $d \geq 2$, in each said stage and each said switch of size $d \times d$ having d inlet links and d outlet links; and
- 15 Said each sub-integrated circuit block comprising a plurality of said switches corresponding to each said stage; and
Said each sub-integrated circuit block comprising a plurality of forward connecting links connecting from switches in lower stage to switches in the immediate succeeding higher stage, and also comprising a plurality of backward connecting links
- 20 connecting from switches in higher stage to switches in the immediate preceding lower stage; and
Said each sub-integrated circuit block comprising a plurality straight links in said forward connecting links from switches in lower stage to switches in the immediate succeeding higher stage and a plurality cross links in said forward connecting links from
- 25 switches in lower stage to switches in the immediate succeeding higher stage, and further comprising a plurality of straight links in said backward connecting links from switches

WO 2008/147928

PCT/US2008/064605

in higher stage to switches in the immediate preceding lower stage and a plurality of cross links in said backward connecting links from switches in higher stage to switches in the immediate preceding lower stage.

2. The integrated circuit device of claim 1, wherein said all straight links are
5 connecting from switches in each said sub-integrated circuit block are connecting to switches in the same said sub-integrated circuit block; and

said all cross links are connecting as either vertical or horizontal links between switches in two different said sub-integrated circuit blocks.

3. The integrated circuit device of claim 2, wherein said plurality of sub-
10 integrated circuit blocks arranged in a two-dimensional grid.

4. The integrated circuit device of claim 3, wherein said cross links in succeeding stages are connecting as alternative vertical and horizontal links between switches in said sub-integrated circuit blocks.

5. The integrated circuit device of claim 4, wherein said cross links from
15 switches in a stage in one of said sub-integrated circuit blocks are connecting to switches in the succeeding stage in another of said sub-integrated circuit blocks so that said cross links are either vertical links or horizontal and vice versa, and hereinafter such cross links are “shuffle exchange links”).

6. The integrated circuit device of claim 5, wherein said all horizontal shuffle
20 exchange links between switches in any two corresponding said succeeding stages are substantially of equal length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are substantially of equal length in the entire said integrated circuit device.

7. The integrated circuit device of claim 6, wherein the shortest horizontal
25 shuffle exchange links are connecting at the lowest stage and between switches in two

WO 2008/147928

PCT/US2008/064605

nearest neighboring said sub-integrated circuit blocks, and length of the horizontal shuffle exchange links is doubled in each succeeding stage; and the shortest vertical shuffle exchange links are connecting at the lowest stage and between switches in two nearest neighboring said sub-integrated circuit blocks, and length of the vertical shuffle exchange
5 links is doubled in each succeeding stage.

8. The integrated circuit device of claim 7, wherein $y \geq (\log_2 N)$ so that the length of the horizontal shuffle exchange links in the highest stage is equal to half the size of the horizontal size of said two dimensional grid of sub-integrated circuit blocks and the length of the vertical shuffle exchange links in the highest stage is equal to half the size of
10 the vertical size of said two dimensional grid of sub-integrated circuit blocks.

9. The integrated circuit device of claim 8, wherein $d = 2$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing
15 network is rearrangeably nonblocking for unicast Benes network with full bandwidth.

10. The integrated circuit device of claim 8, wherein $d = 2$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said
20 routing network is strictly nonblocking for unicast Benes network and rearrangeably nonblocking for arbitrary fan-out multicast Benes network with full bandwidth.

11. The integrated circuit device of claim 8, wherein $d = 2$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in
25 each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast Benes network with full bandwidth.

WO 2008/147928

PCT/US2008/064605

12. The integrated circuit device of claim 7, wherein $y \geq (\log_2 N)$ so that the length of the horizontal shuffle exchange links in the highest stage is equal to half the size of the horizontal size of said two dimensional grid of sub-integrated circuit blocks and the length of the vertical shuffle exchange links in the highest stage is equal to half the size of the vertical size of said two dimensional grid of sub-integrated circuit blocks, and
5 said each sub-integrated circuit block further comprising a plurality of U-turn links within switches in each of said stages in each of said sub-integrated circuit blocks.

13. The integrated circuit device of claim 12, wherein $d = 2$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said
10 forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast butterfly fat tree network with full bandwidth.

14. The integrated circuit device of claim 12, wherein $d = 2$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting
15 said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast butterfly fat tree network and rearrangeably nonblocking for arbitrary fan-out multicast butterfly fat tree network with
20 full bandwidth.

15. The integrated circuit device of claim 12, wherein $d = 2$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting
said forward connecting links and there are at least three switches in each said stage in
each said sub-integrated circuit block connecting said backward connecting links and said
25 routing network is strictly nonblocking for arbitrary fan-out multicast butterfly fat tree network with full bandwidth.

WO 2008/147928

PCT/US2008/064605

16. The integrated circuit device of claim 1, wherein said horizontal and vertical links are implemented on two or more metal layers.
17. The integrated circuit device of claim 1, wherein said switches comprising active and reprogrammable cross points and said each cross point is programmable by an SRAM cell or a Flash Cell.
18. The integrated circuit device of claim 1, wherein said sub-integrated circuit blocks are of equal die size.
19. The integrated circuit device of claim 16, wherein said sub-integrated circuit blocks are Lookup Tables (hereinafter "LUTs") and said integrated circuit device is a field programmable gate array (FPGA) device or field programmable gate array (FPGA) block embedded in another integrated circuit device.
20. The integrated circuit device of claim 16, wherein said sub-integrated circuit blocks are AND or OR gates and said integrated circuit device is a programmable logic device (PLD).
21. The integrated circuit device of claim 1, wherein said sub-integrated circuit blocks comprising any arbitrary hardware logic or memory circuits.
22. The integrated circuit device of claim 1, wherein said switches comprising active one-time programmable cross points and said integrated circuit device is a mask programmable gate array (MPGA) device or a structured ASIC device.
23. The integrated circuit device of claim 1, wherein said switches comprising passive cross points or just connection of two links or not and said integrated circuit device is a Application Specific Integrated Circuit (ASIC) device.
24. The integrated circuit device of claim 1, wherein said sub-integrated circuit blocks further recursively comprise one or more super-sub-integrated circuit blocks and a sub-routing network.

WO 2008/147928

PCT/US2008/064605

25. The integrated circuit device of claim 5, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \geq (\log_2 N)$.

5 26. The integrated circuit device of claim 25, wherein $d = 2$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-stage network with
10 full bandwidth.

27. The integrated circuit device of claim 25, wherein $d = 2$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said
15 routing network is strictly nonblocking for unicast generalized multi-stage network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-stage network with full bandwidth.

28. The integrated circuit device of claim 25, wherein $d = 2$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting
20 said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-stage network with full bandwidth.

29. The integrated circuit device of claim 5, wherein said all horizontal shuffle
25 exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \geq (\log_2 N)$, and

WO 2008/147928

PCT/US2008/064605

said each sub-integrated circuit block further comprising a plurality of U-turn links within switches in each of said stages in each of said sub-integrated circuit blocks.

30. The integrated circuit device of claim 29, wherein $d = 2$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized butterfly fat tree network with full bandwidth.

31. The integrated circuit device of claim 29, wherein $d = 2$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized butterfly fat tree Network and rearrangeably nonblocking for arbitrary fan-out multicast generalized butterfly fat tree network with full bandwidth.

32. The integrated circuit device of claim 29, wherein $d = 2$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized butterfly fat tree network with full bandwidth.

33. The integrated circuit device of claim 1, wherein said straight links connecting from switches in each said sub-integrated circuit block are connecting to switches in the same said sub-integrated circuit block; and said cross links are connecting as vertical or horizontal or diagonal links between two different said sub-integrated circuit blocks.

WO 2008/147928

PCT/US2008/064605

34. The integrated circuit device of claim 8, wherein $d = 4$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast multi-link Benes network with full bandwidth.

35. The integrated circuit device of claim 8, wherein $d = 4$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast multi-link Benes network and rearrangeably nonblocking for arbitrary fan-out multicast multi-link Benes network with full bandwidth.

36. The integrated circuit device of claim 8, wherein $d = 4$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast multi-link Benes network with full bandwidth.

37. The integrated circuit device of claim 12, wherein $d = 4$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast multi-link butterfly fat tree network with full bandwidth.

38. The integrated circuit device of claim 12, wherein $d = 4$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in

WO 2008/147928

PCT/US2008/064605

each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast multi-link butterfly fat tree network and rearrangeably nonblocking for arbitrary fan-out multicast multi-link butterfly fat tree network with full bandwidth.

5 39. The integrated circuit device of claim 12, wherein $d = 4$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast multi-link butterfly
10 fat tree network with full bandwidth.

40. The integrated circuit device of claim 5, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \geq (\log_2 N)$.

15 41. The integrated circuit device of claim 40, wherein $d = 4$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-link multi-stage
20 network with full bandwidth.

42. The integrated circuit device of claim 40, wherein $d = 4$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said
25 routing network is strictly nonblocking for unicast generalized multi-link multi-stage network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-link multi-stage network with full bandwidth.

WO 2008/147928

PCT/US2008/064605

43. The integrated circuit device of claim 40, wherein $d = 4$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-link multi-stage network with full bandwidth.

44. The integrated circuit device of claim 5, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \geq (\log_2 N)$, and said each sub-integrated circuit block further comprising a plurality of U-turn links within switches in each of said stages in each of said sub-integrated circuit blocks.

45. The integrated circuit device of claim 44, wherein $d = 4$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-link butterfly fat tree network with full bandwidth.

46. The integrated circuit device of claim 44, wherein $d = 4$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized multi-link butterfly fat tree Network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-link butterfly fat tree network with full bandwidth.

47. The integrated circuit device of claim 44, wherein $d = 4$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting

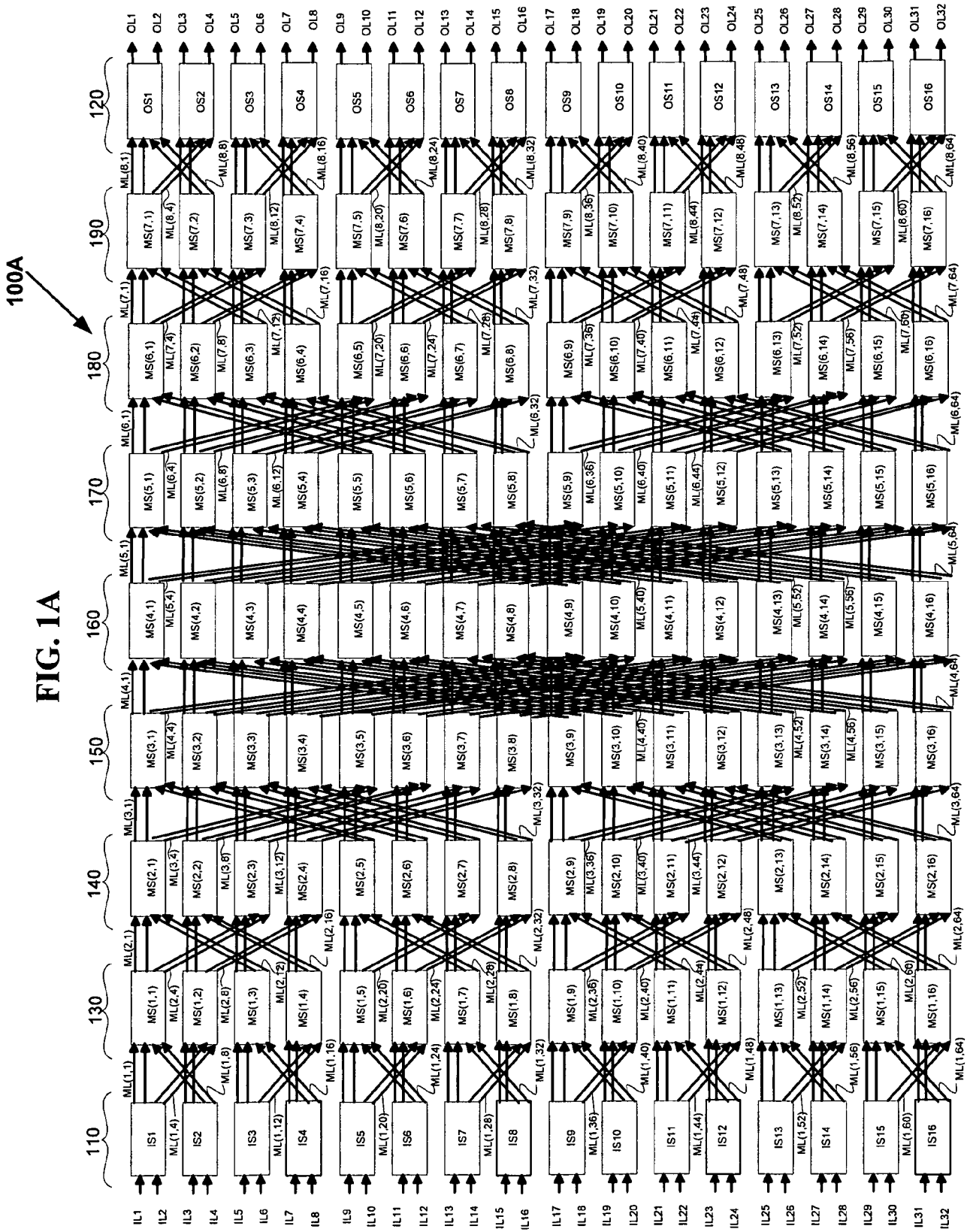
WO 2008/147928

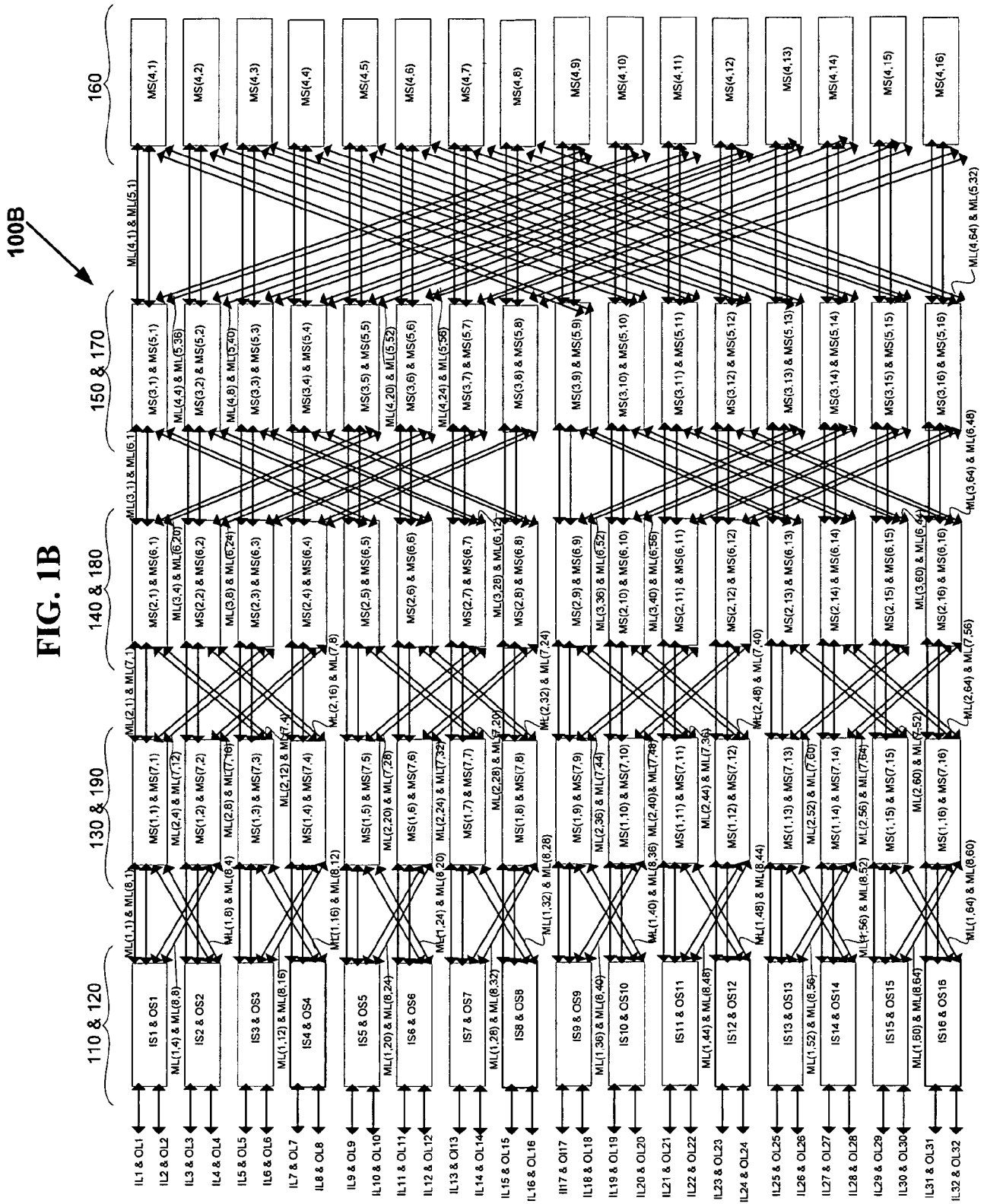
PCT/US2008/064605

said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-link butterfly fat tree network with full bandwidth.

5 48. The integrated circuit device of claim 1, wherein said plurality of forward connecting links use a plurality of buffers to amplify signals driven through them and said plurality of backward connecting links use a plurality of buffers to amplify signals driven through them; and said buffers can be inverting or non-inverting buffers.

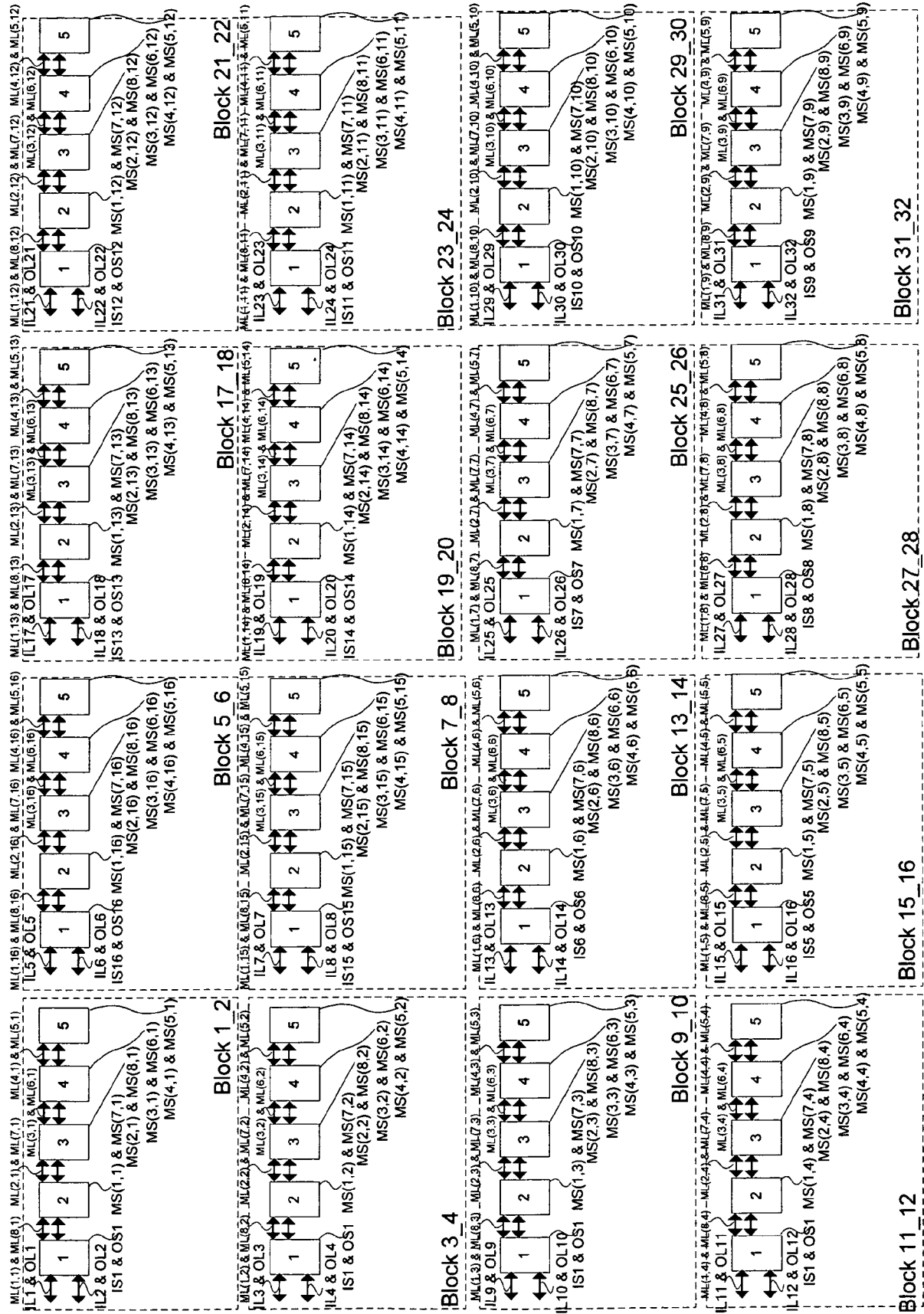
10 49. The integrated circuit device of claim 1, wherein said all switches of size $d \times d$ are either fully populated or partially populated.





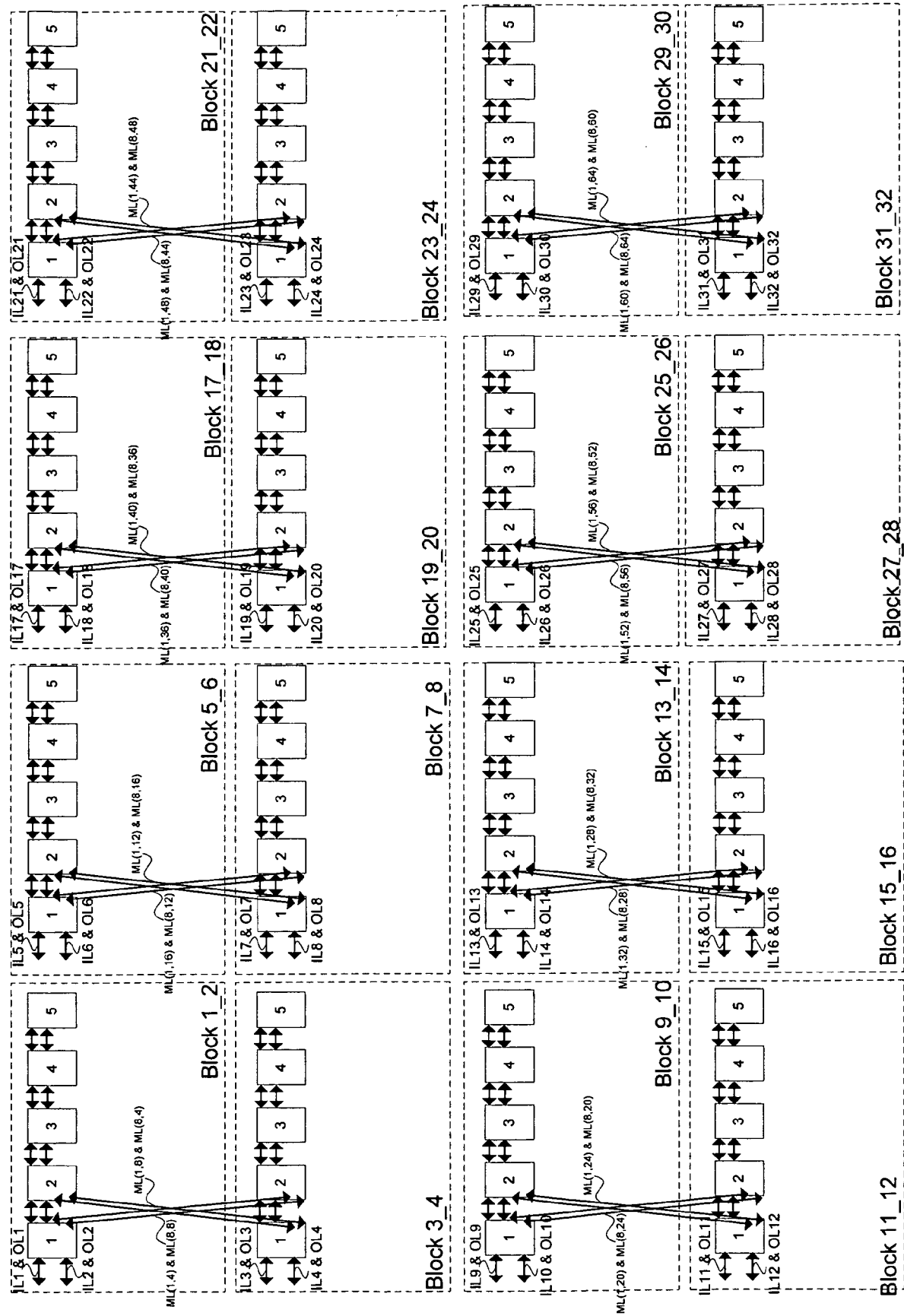
100C

FIG. 1C



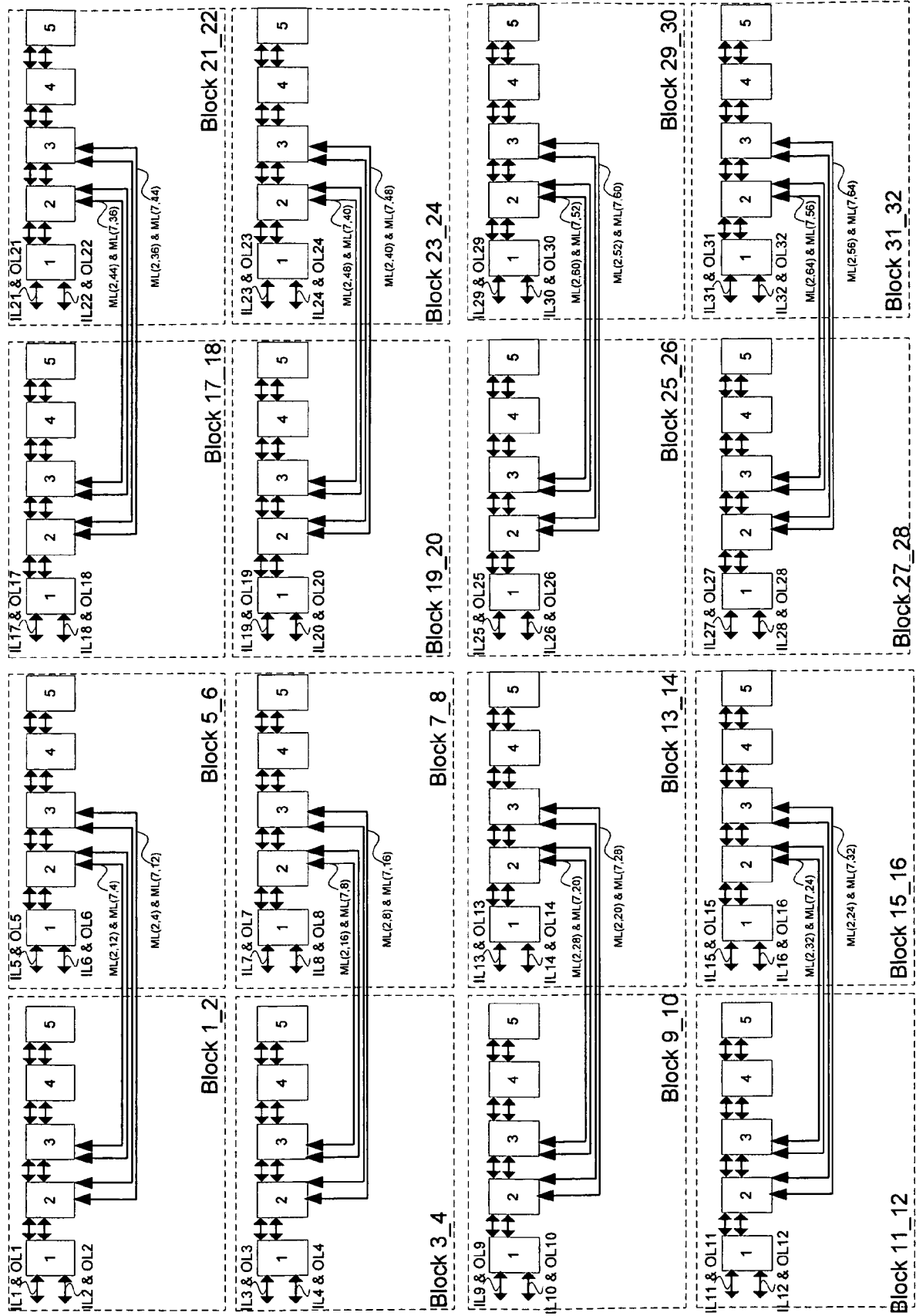
100D

FIG. 1D



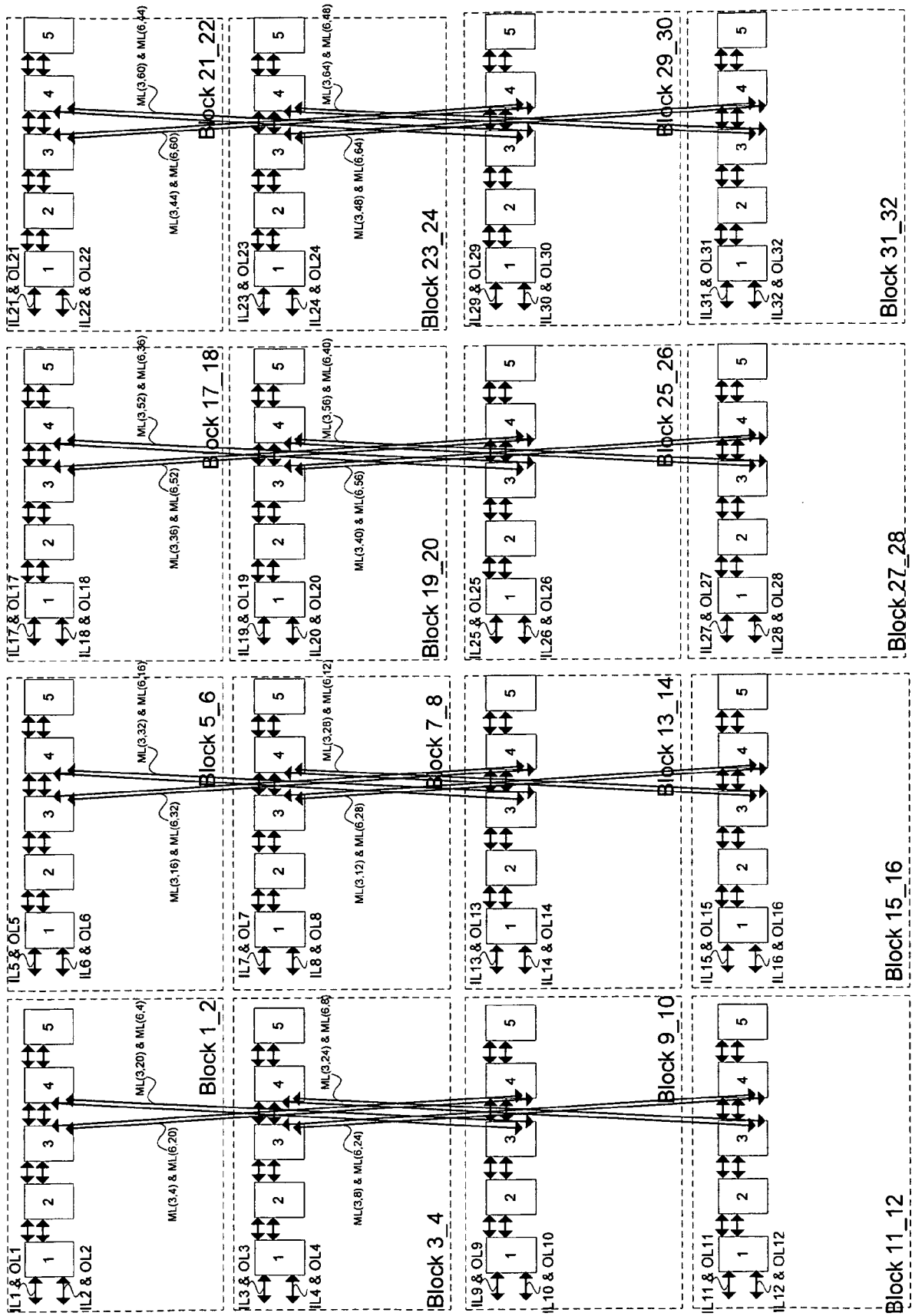
100E

FIG. 1E



100F

FIG. 1F



100G

FIG. 1G

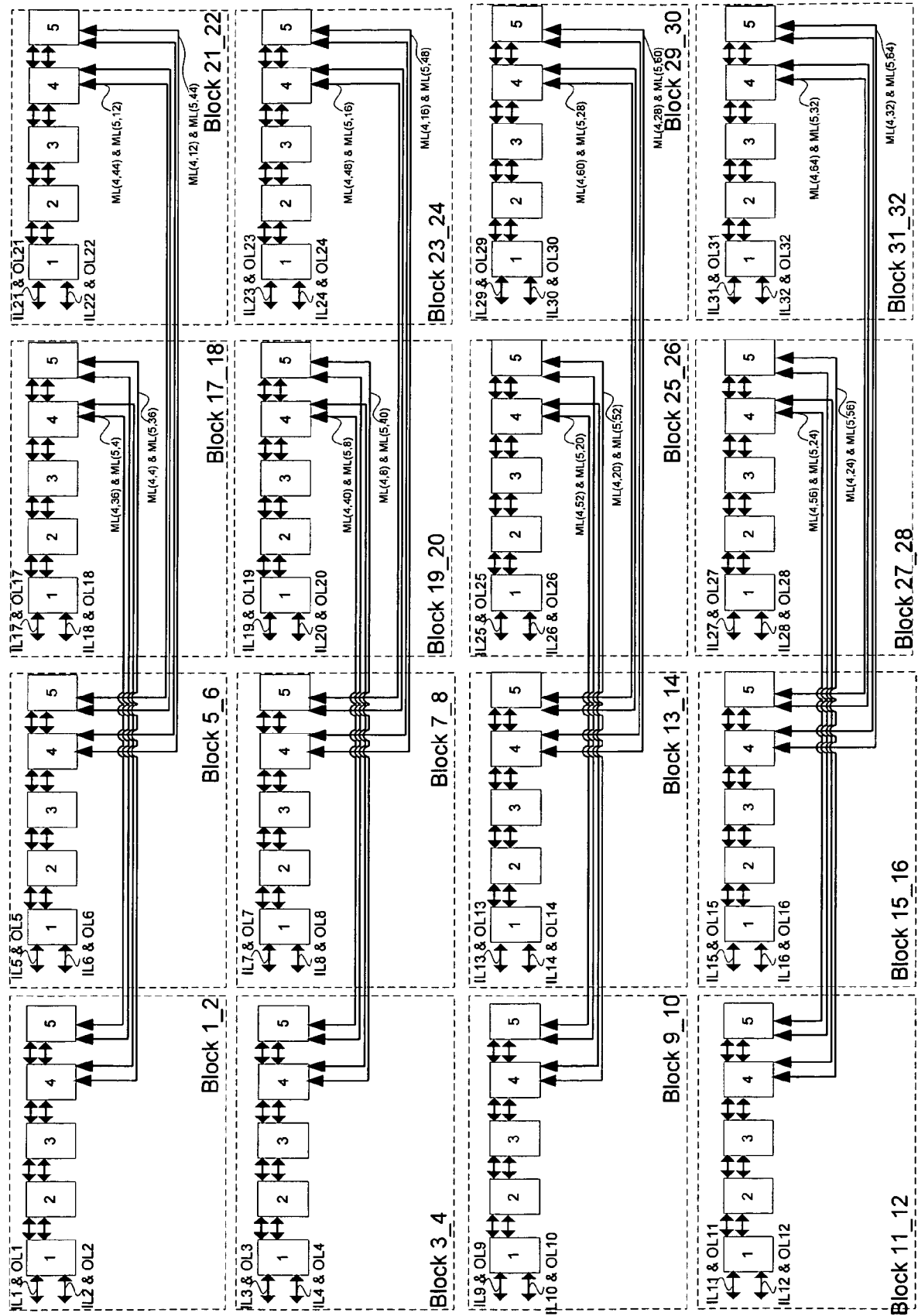
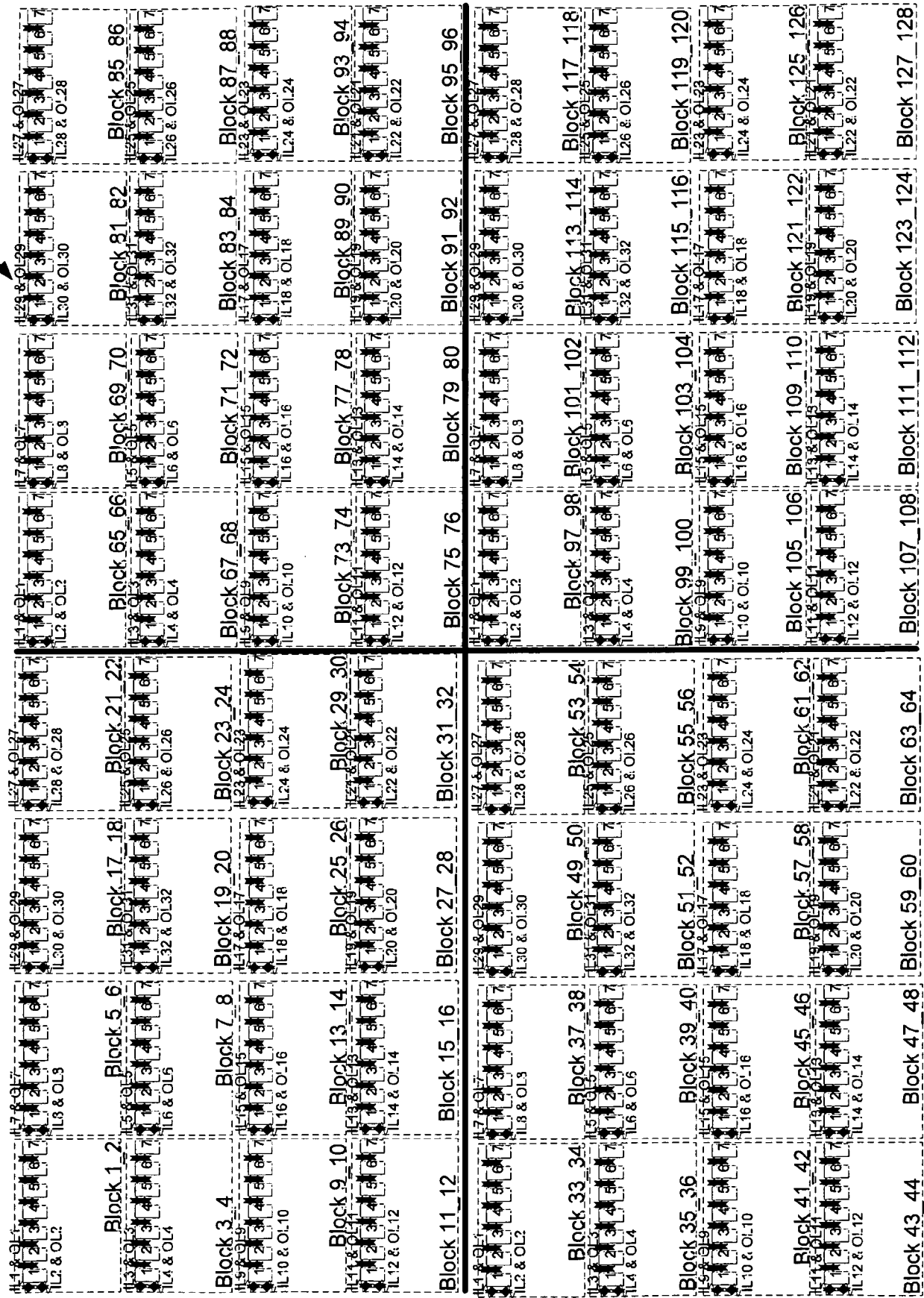
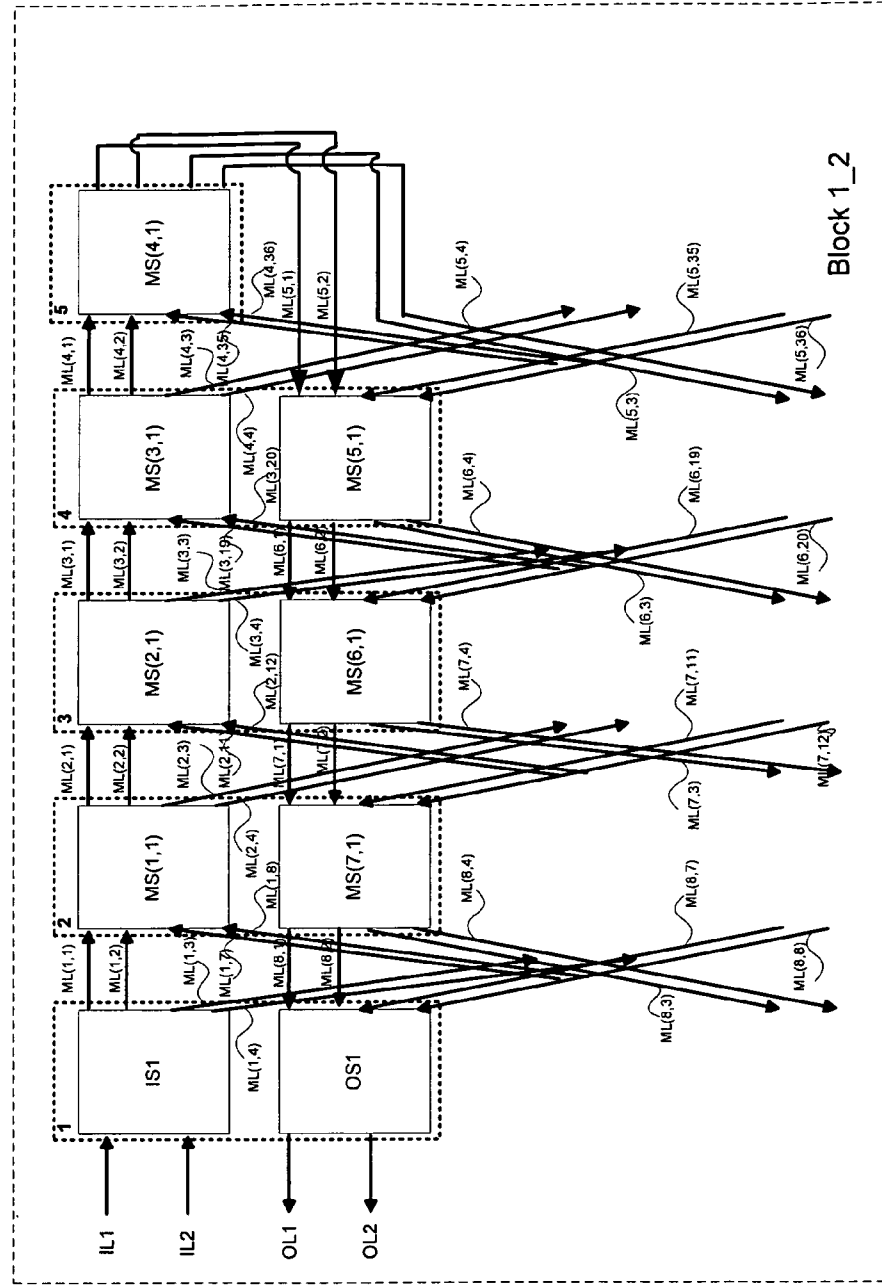


FIG. 1H
100H



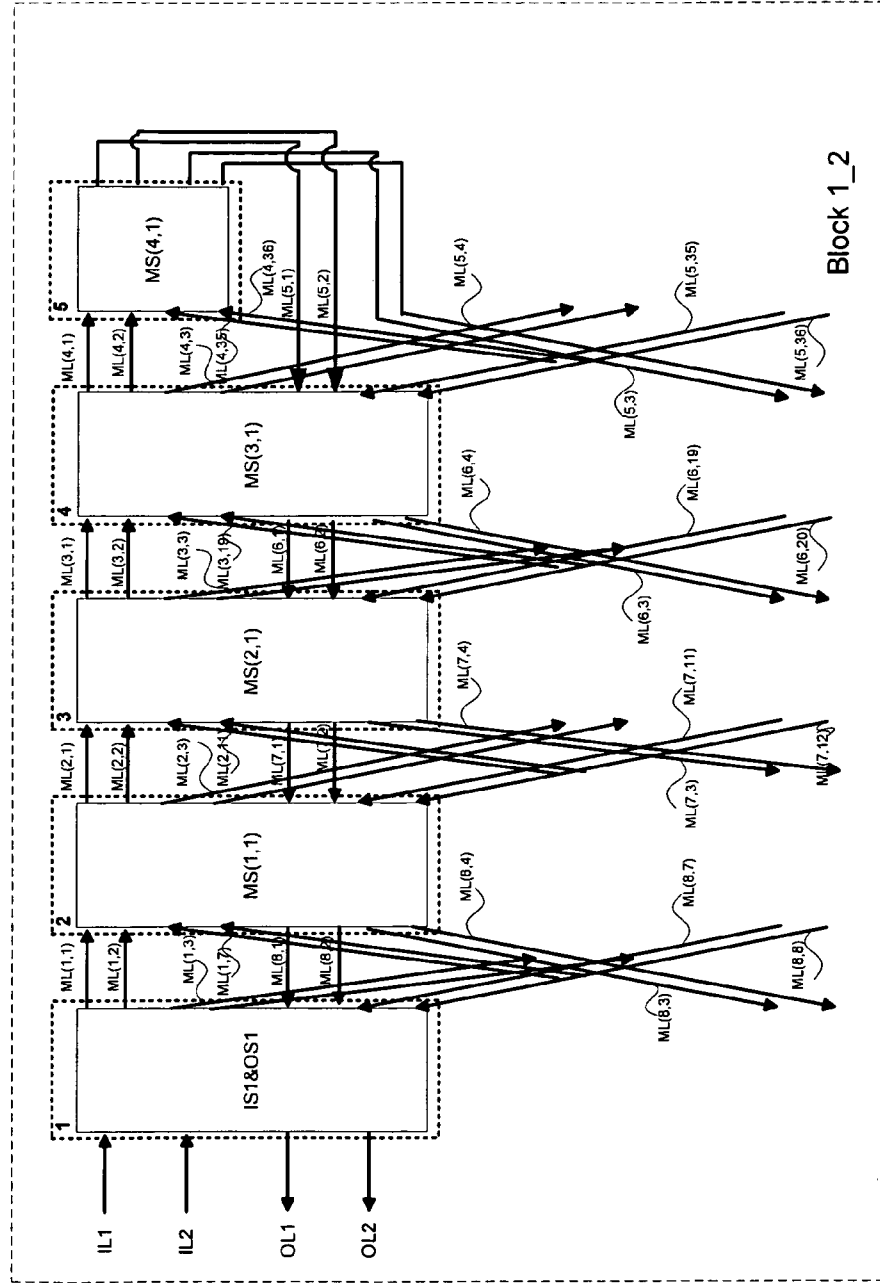
1001

FIG. 11



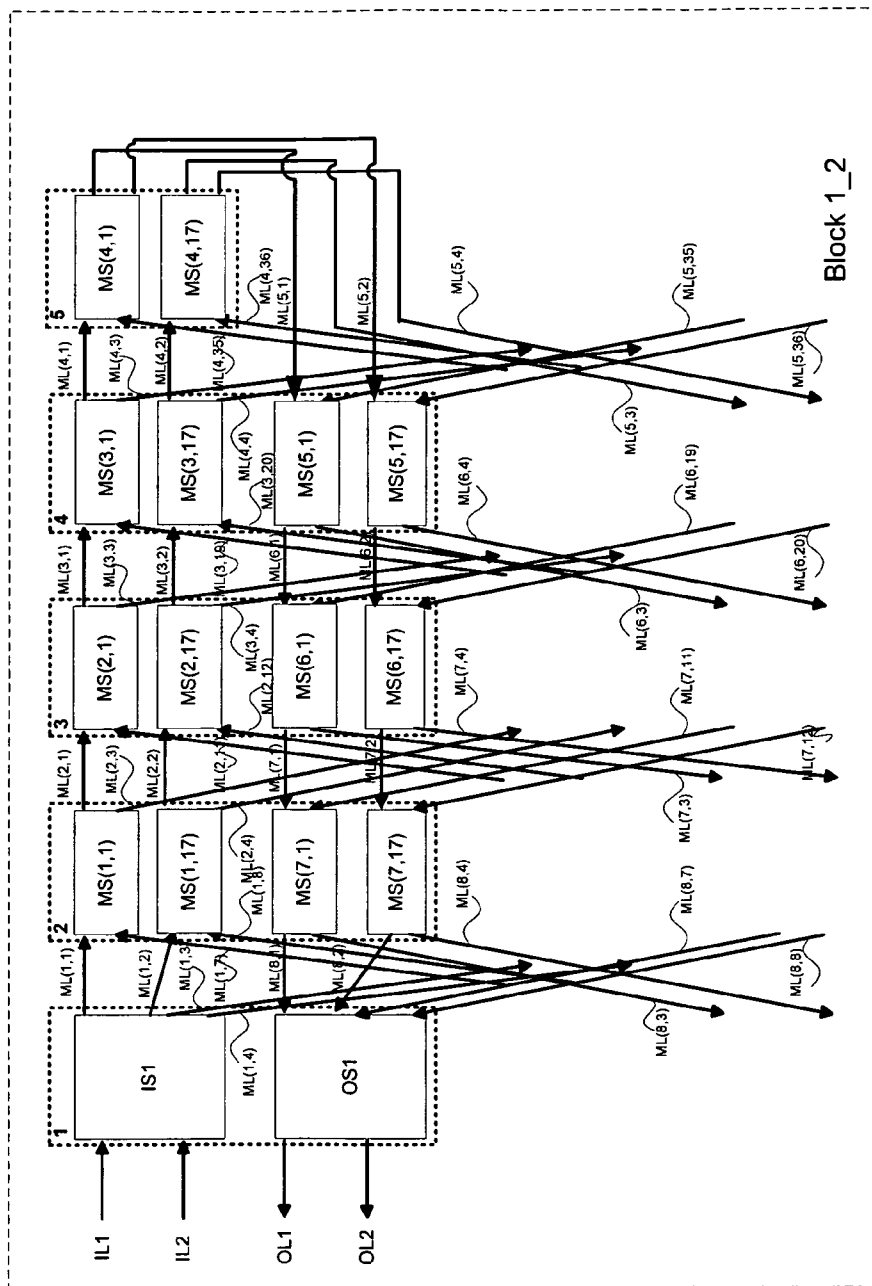
100J

FIG. 1J



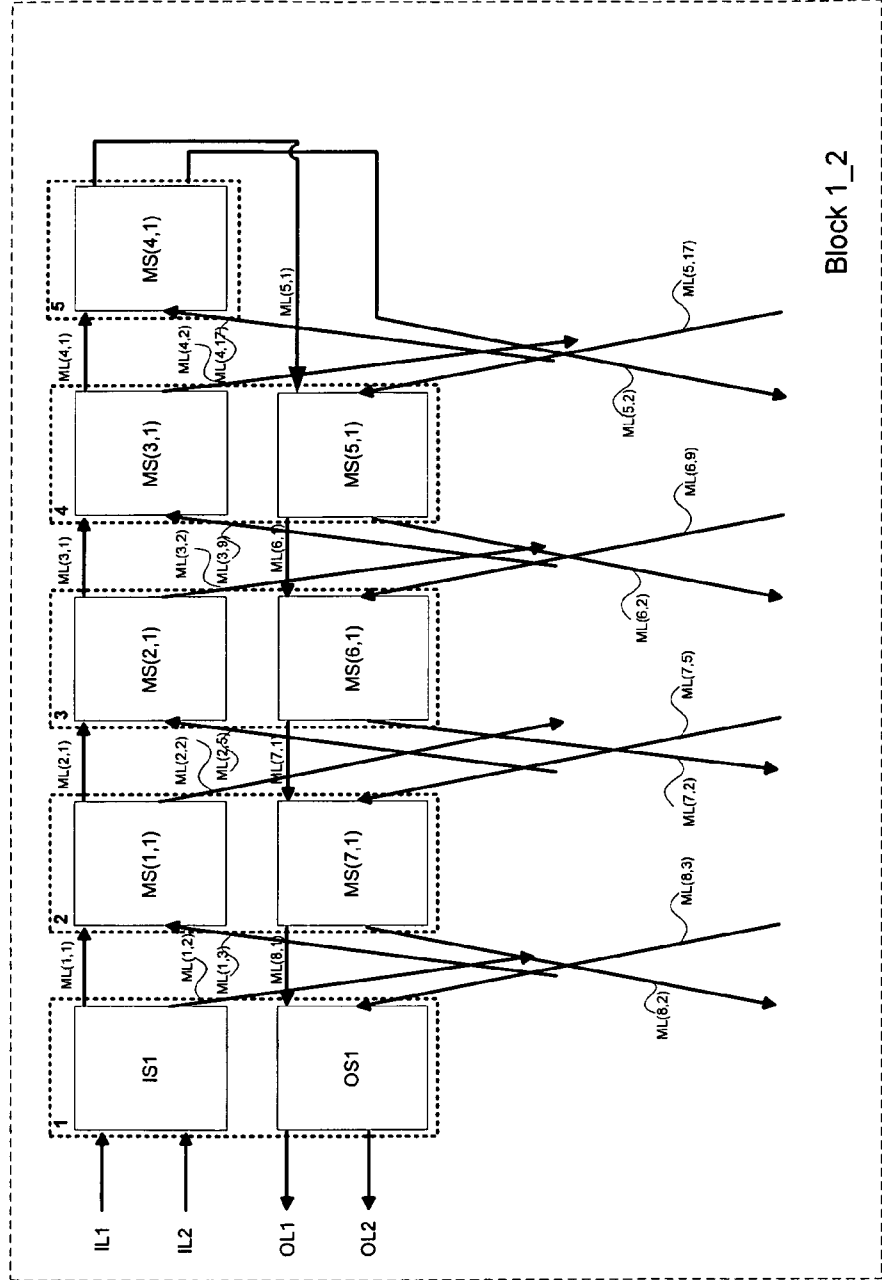
100K

FIG. 1K



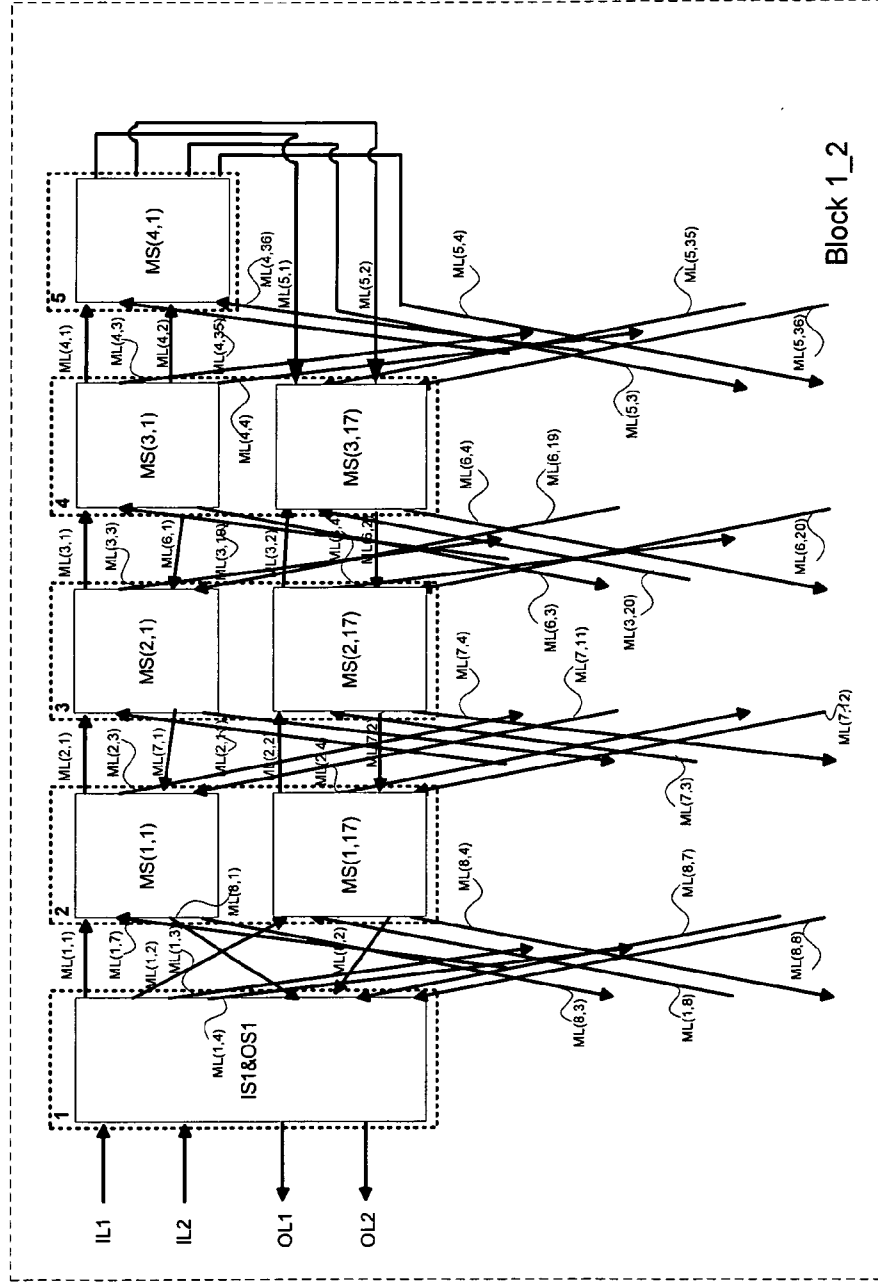
100K1

FIG. 1K1



100L

FIG. 1L



100L1

FIG. 1L1

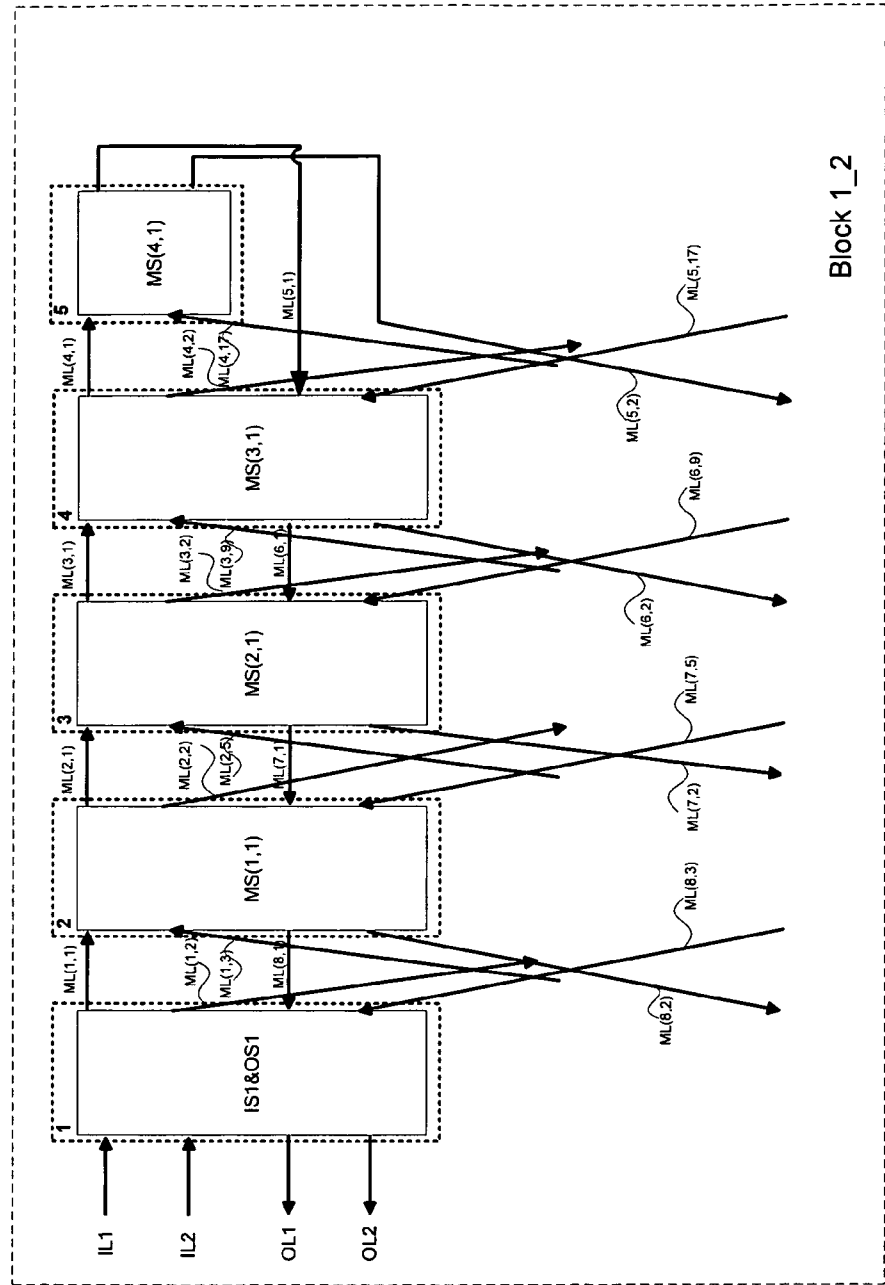


FIG. 2A1

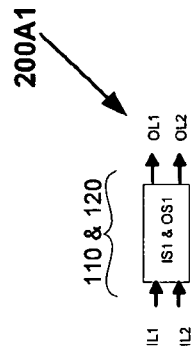


FIG. 2A2

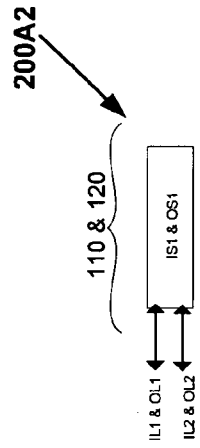
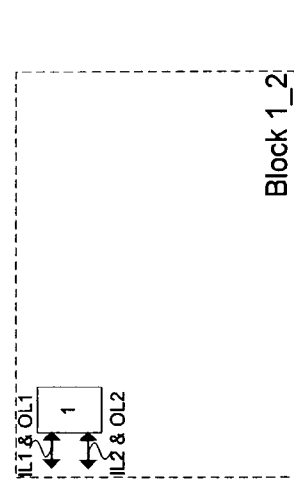
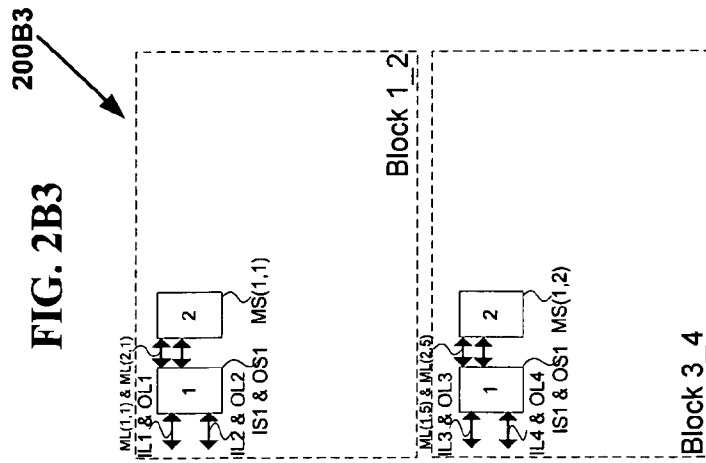
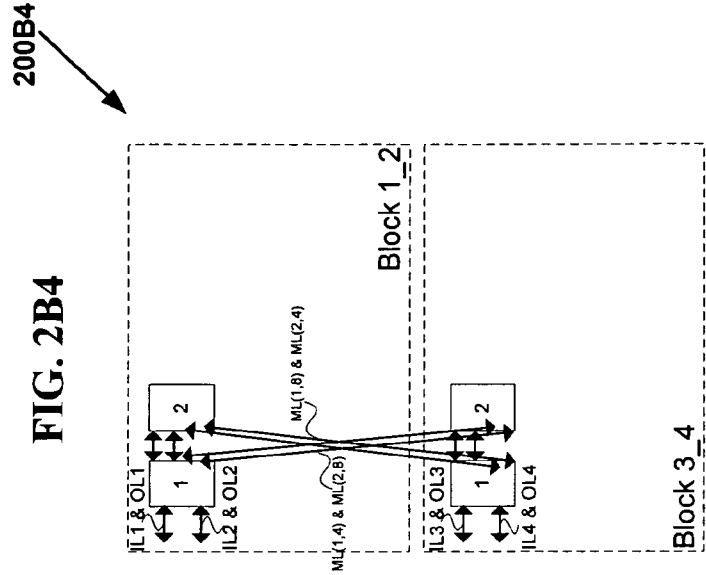
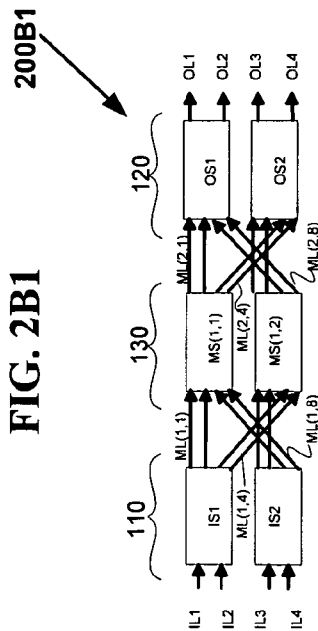
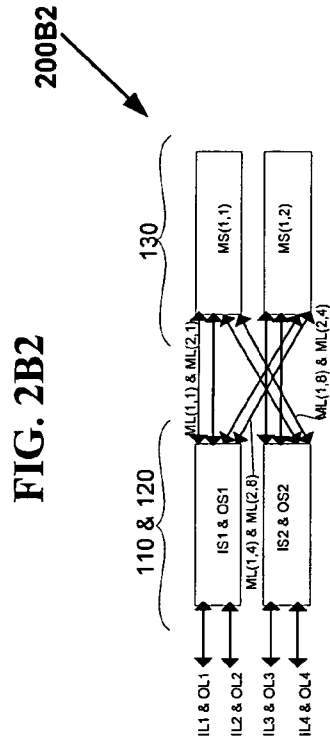
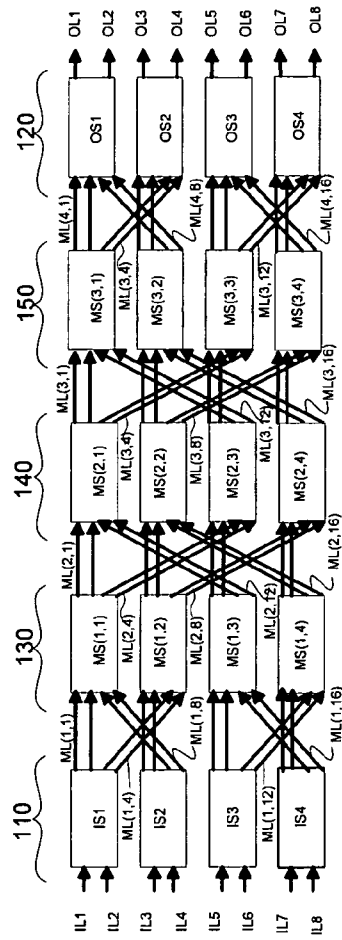


FIG. 2A3





200C11
FIG. 2C11



200C12
FIG. 2C12

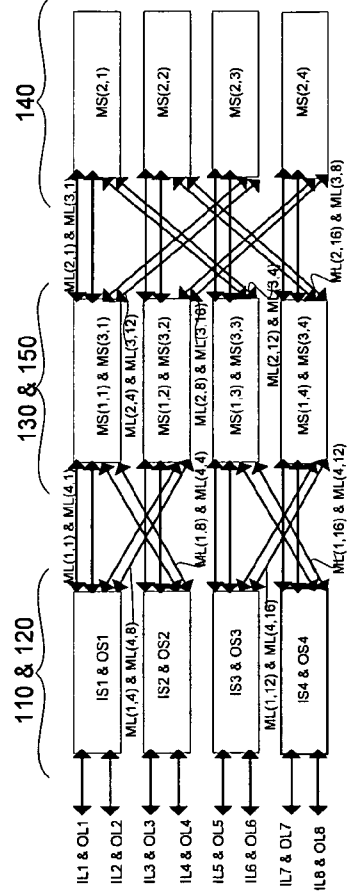


FIG. 2C21

200C21

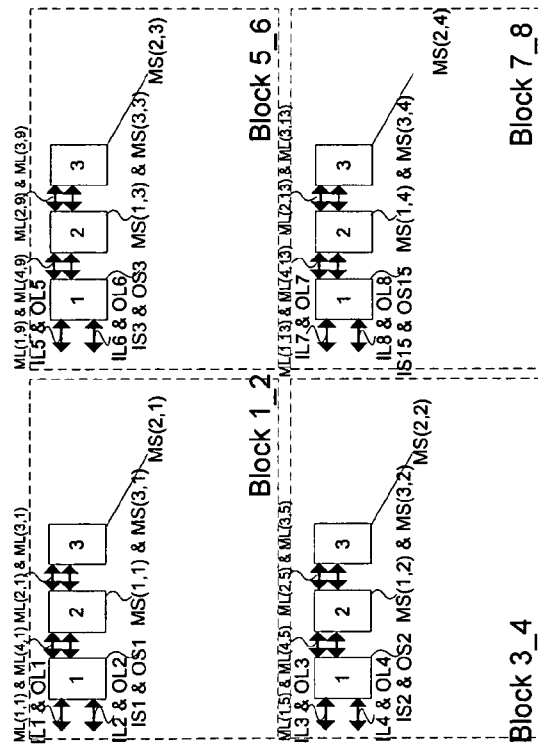


FIG. 2C22

200C22

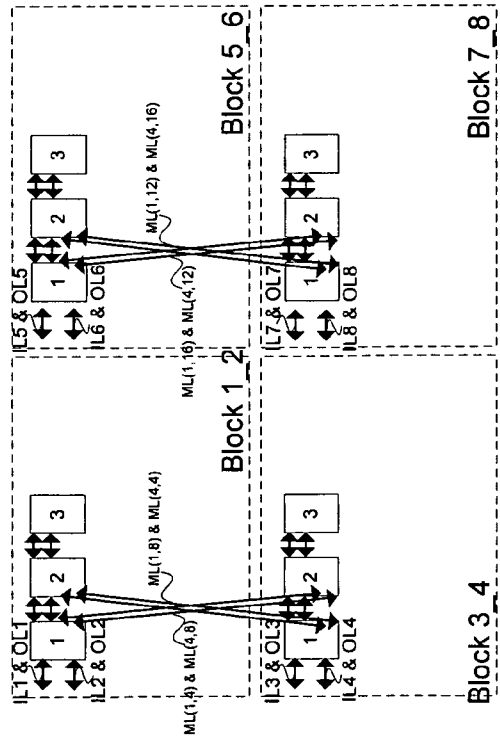


FIG. 2C23

200C23

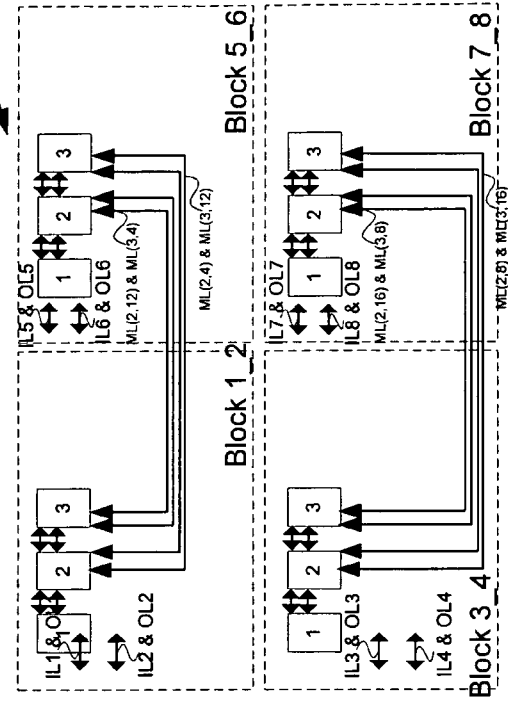


FIG. 2D1

200D1

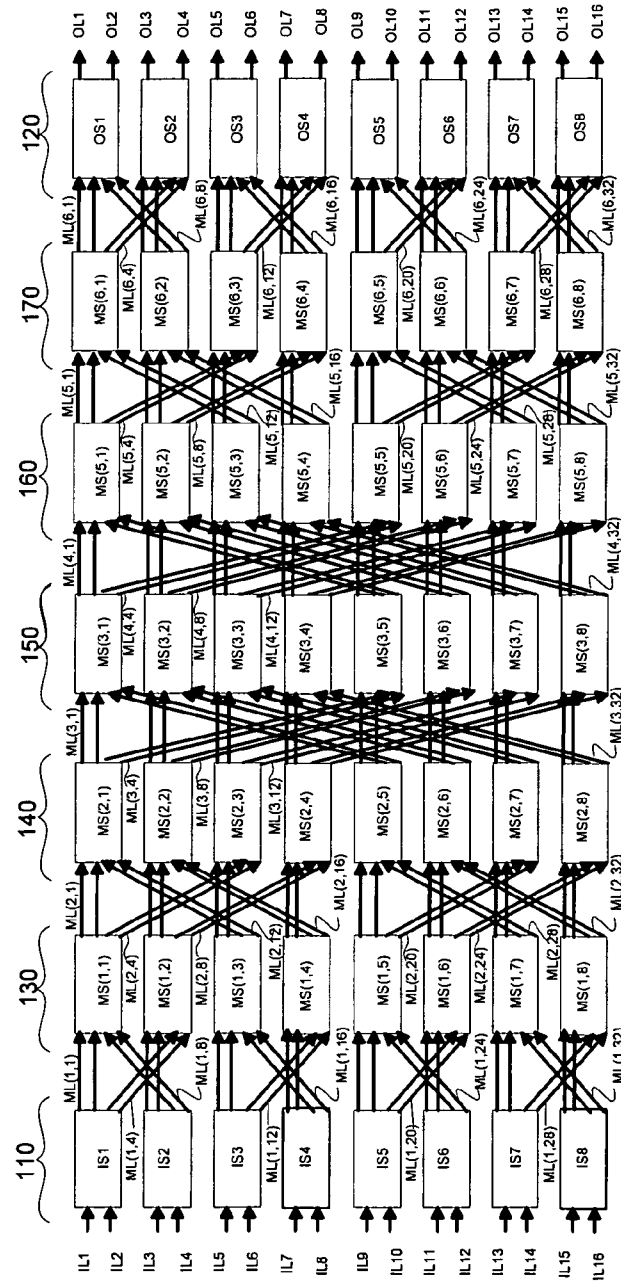
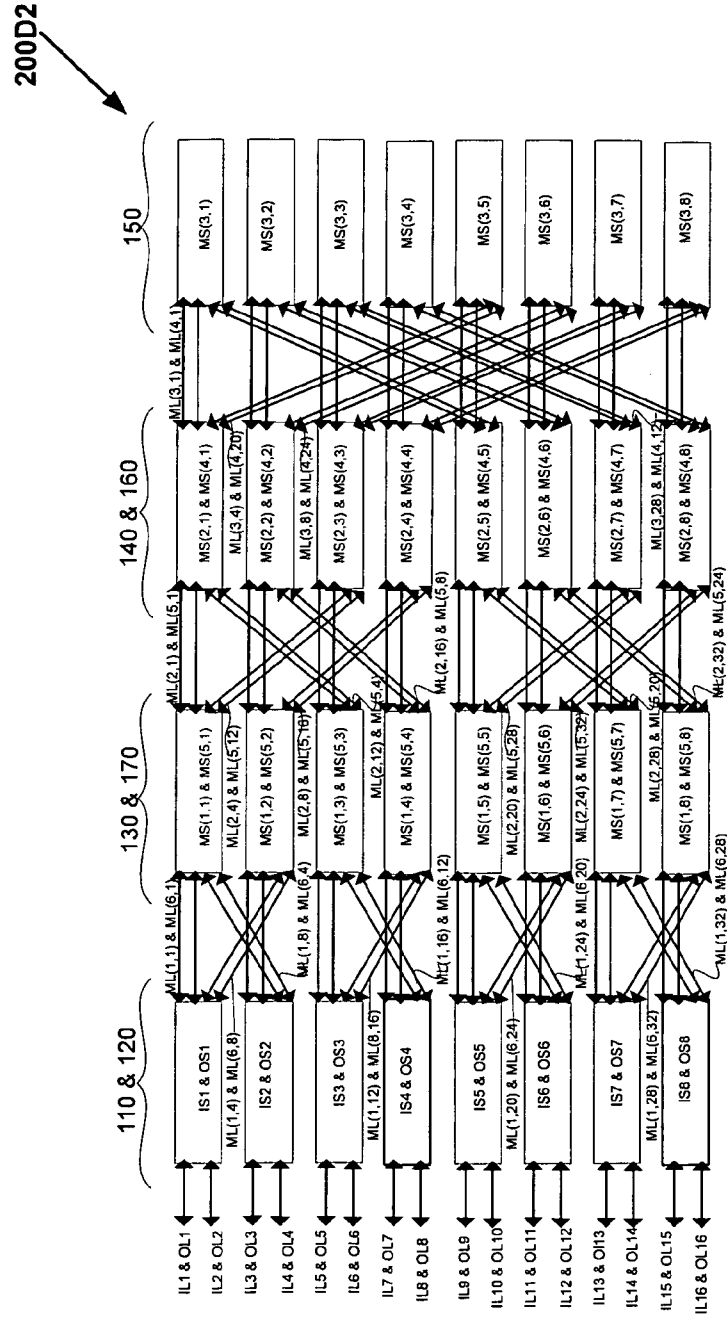
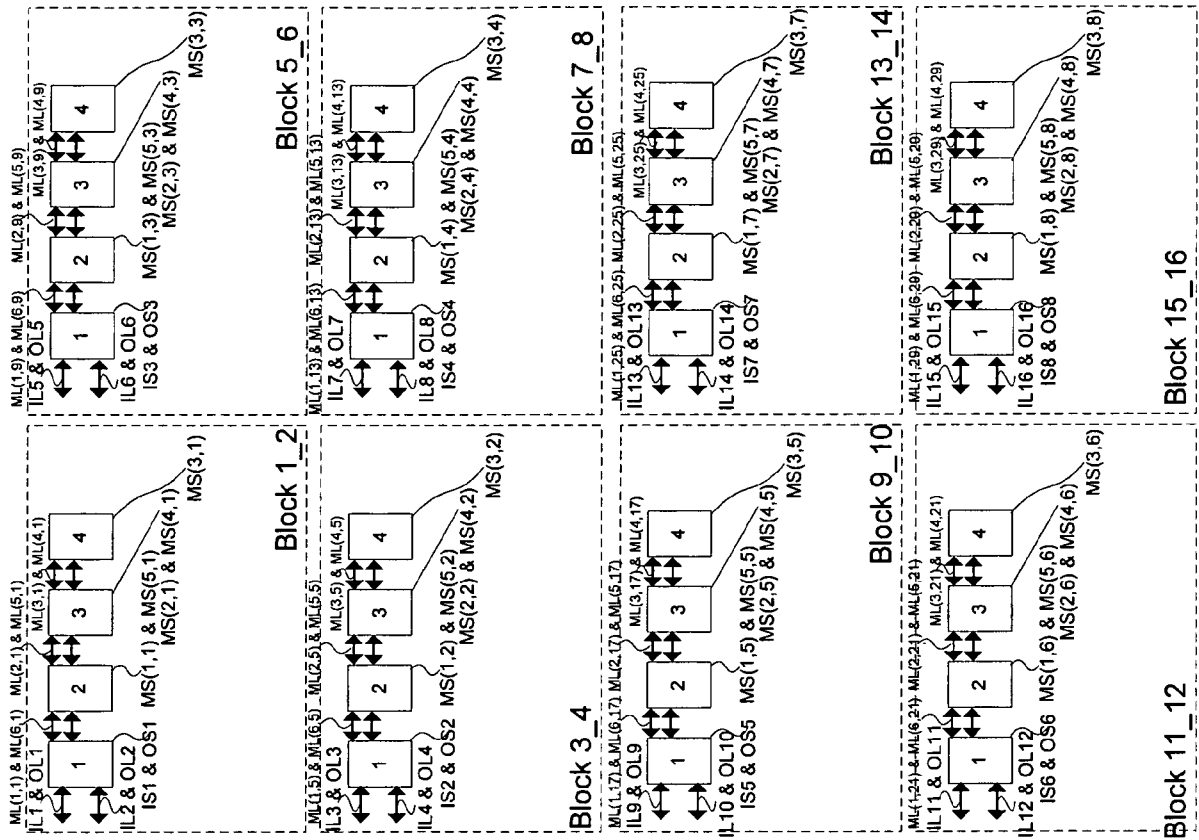


FIG. 2D2



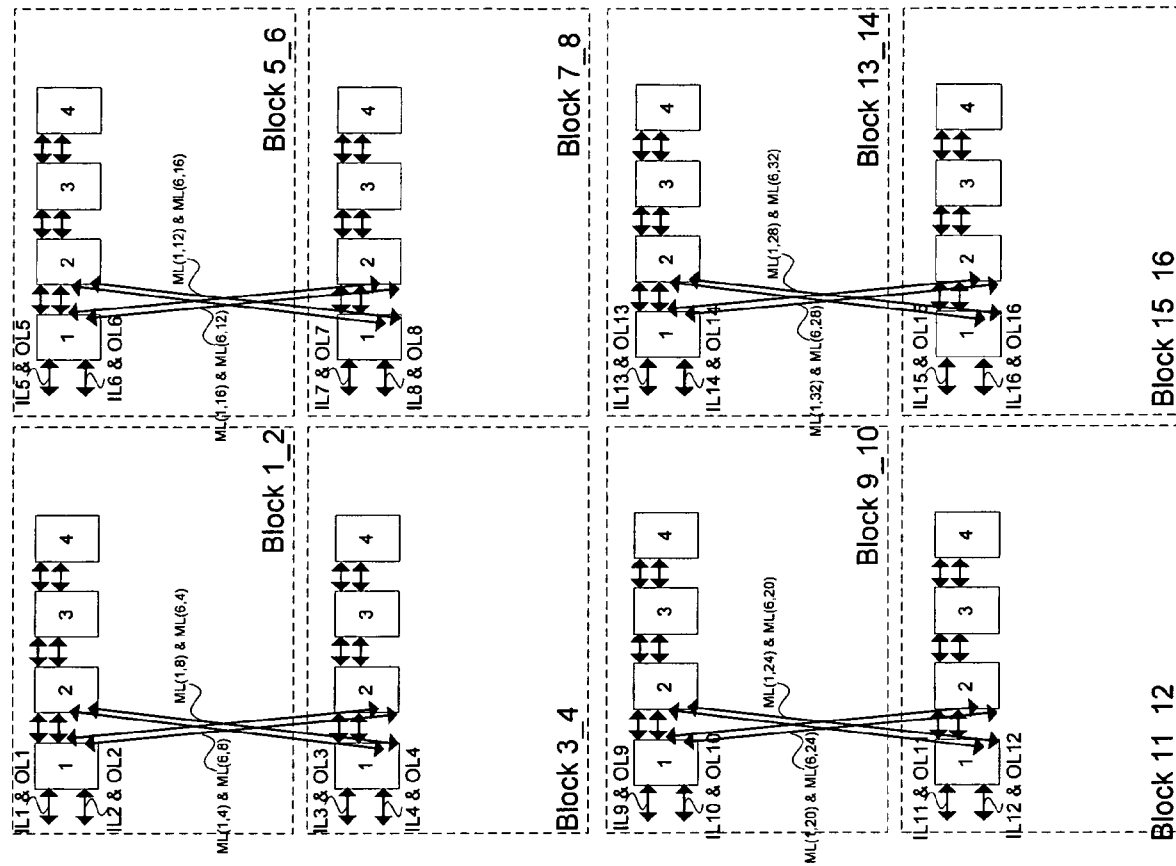
200D3

FIG. 2D3



200D4

FIG. 2D4



200D5

FIG. 2D5

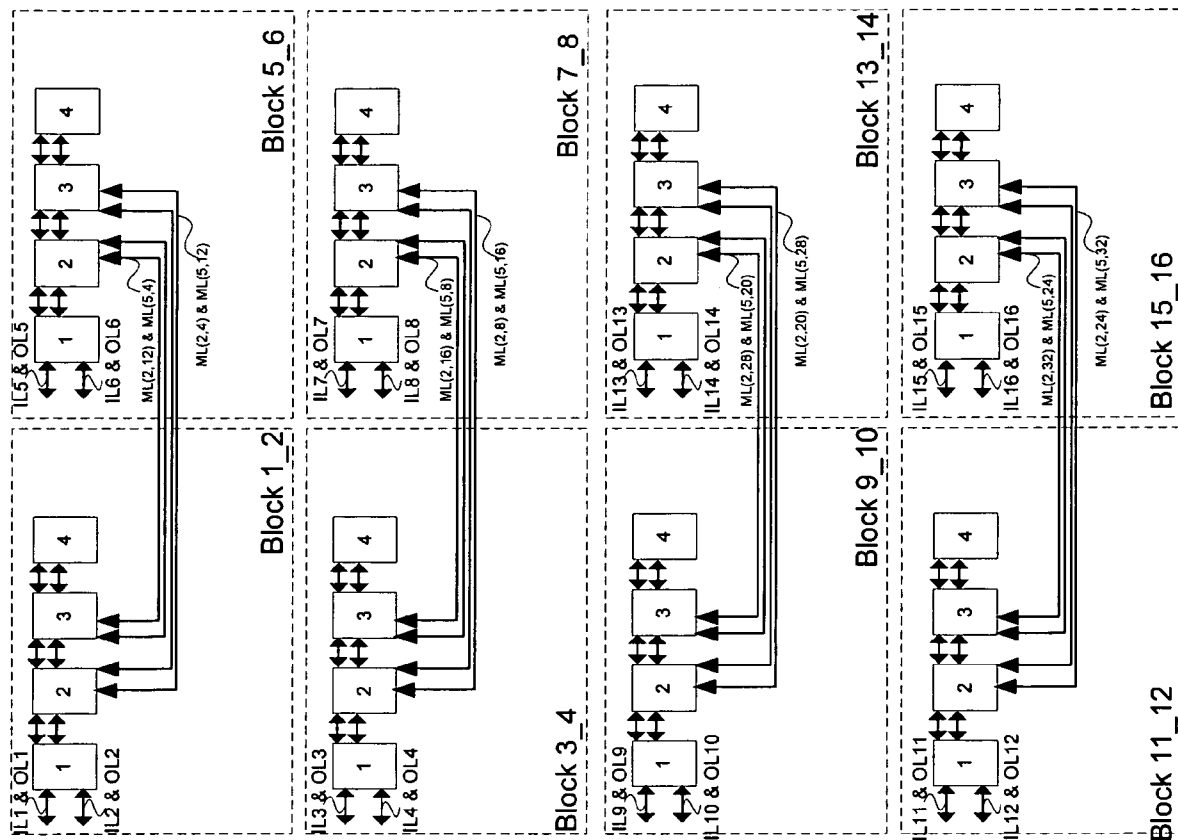
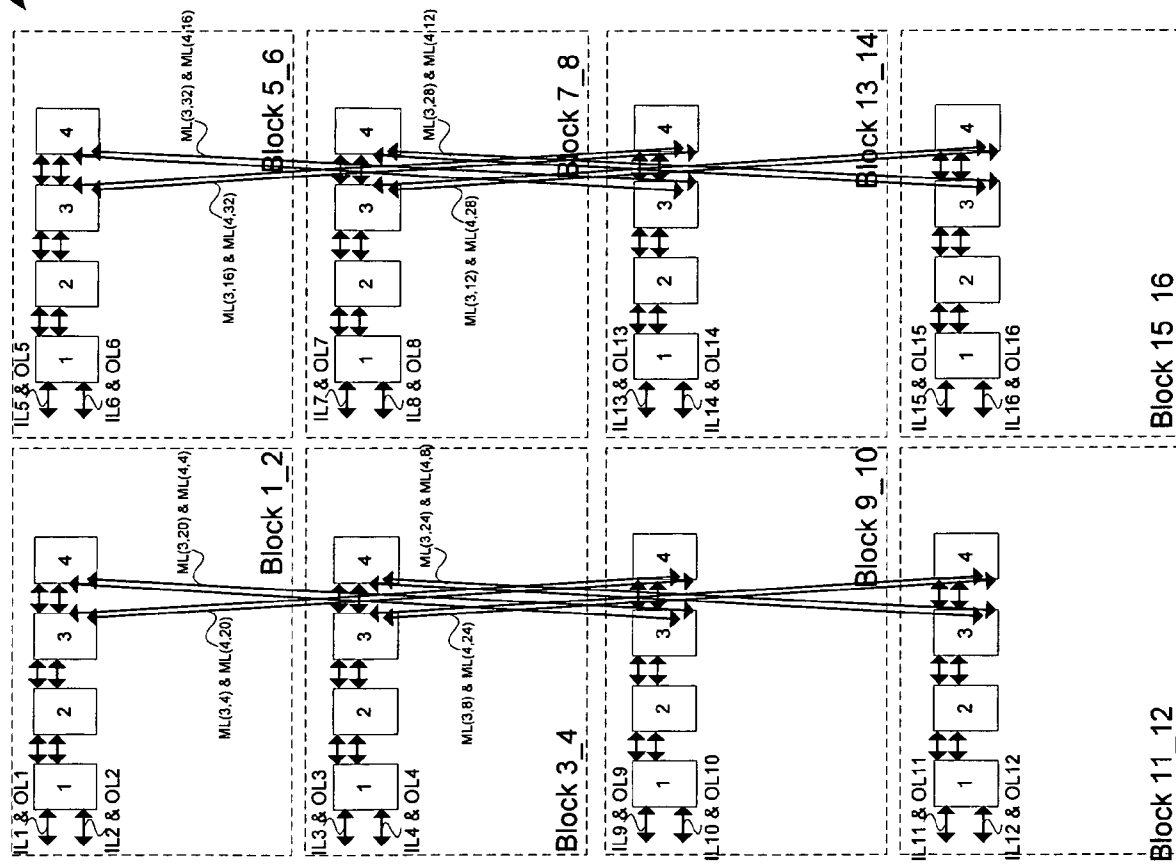


FIG. 2D6

200D6



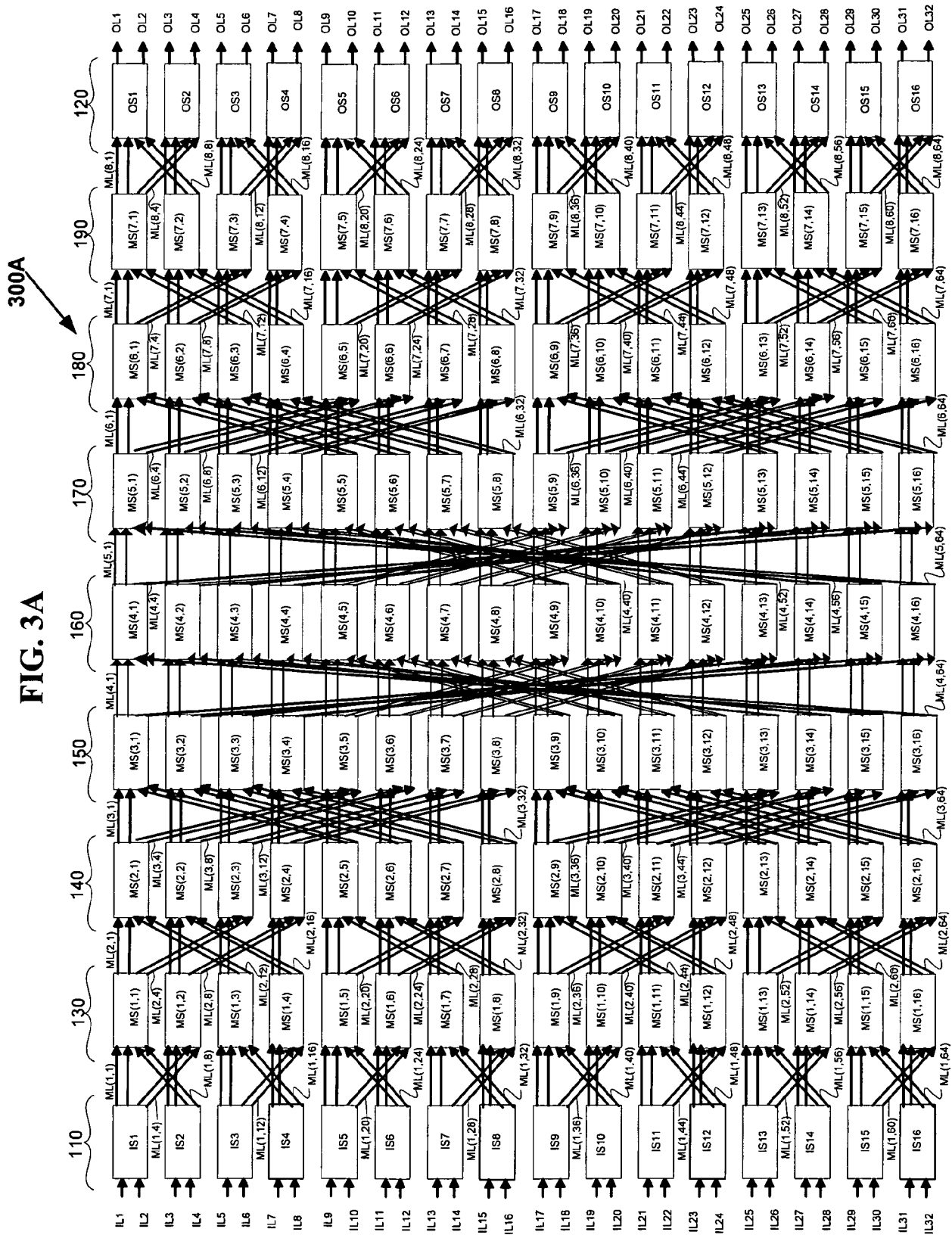
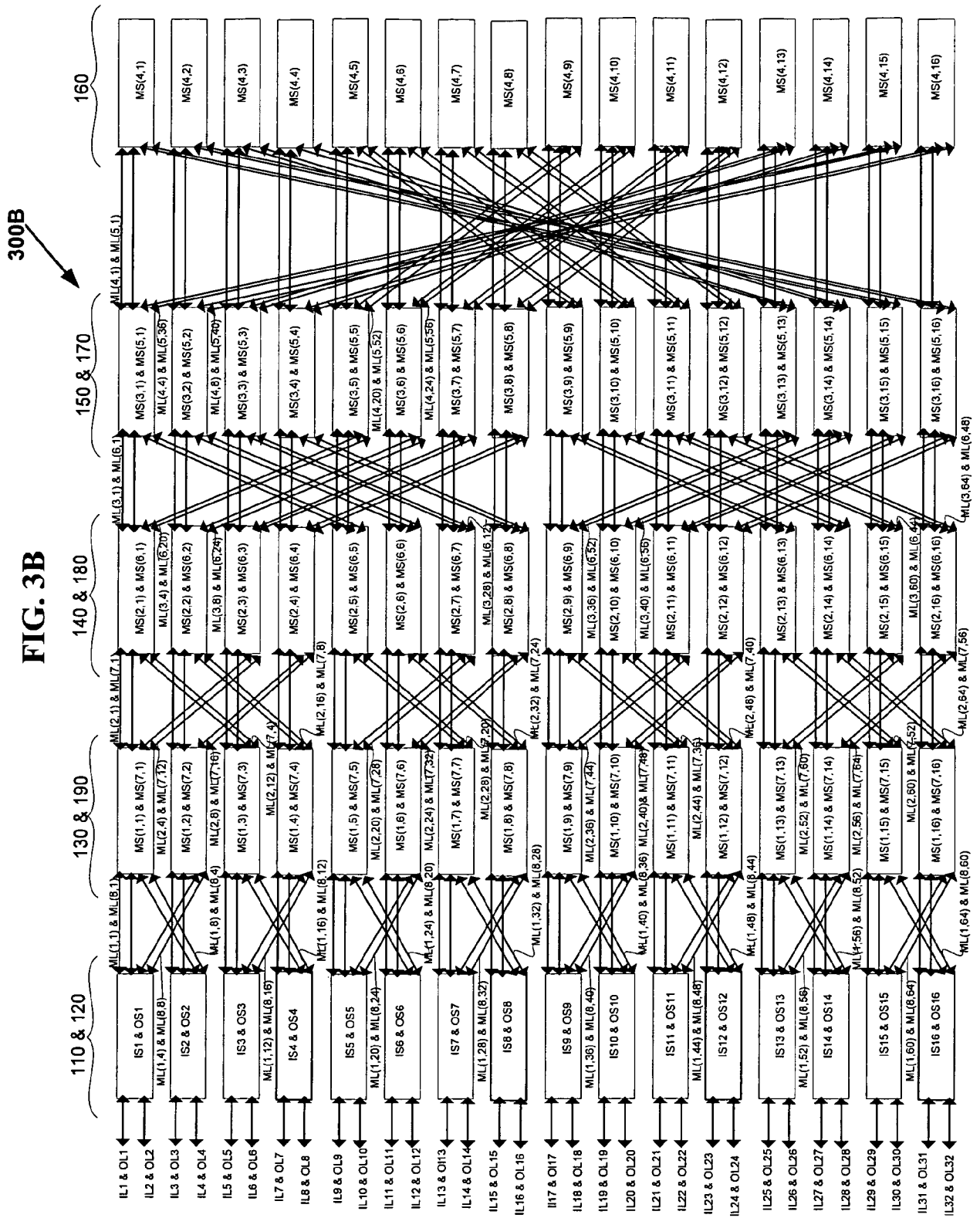
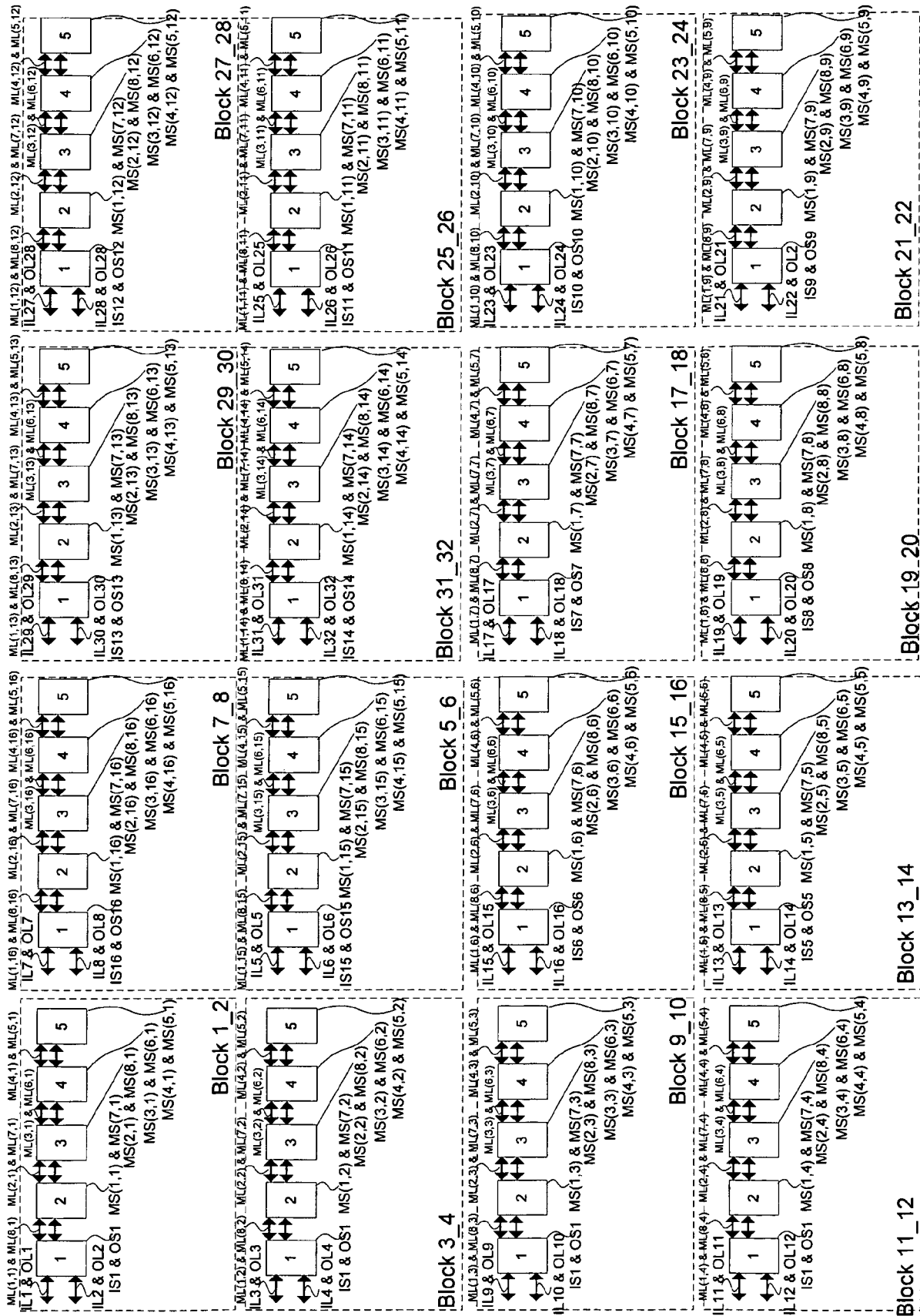


FIG. 3A



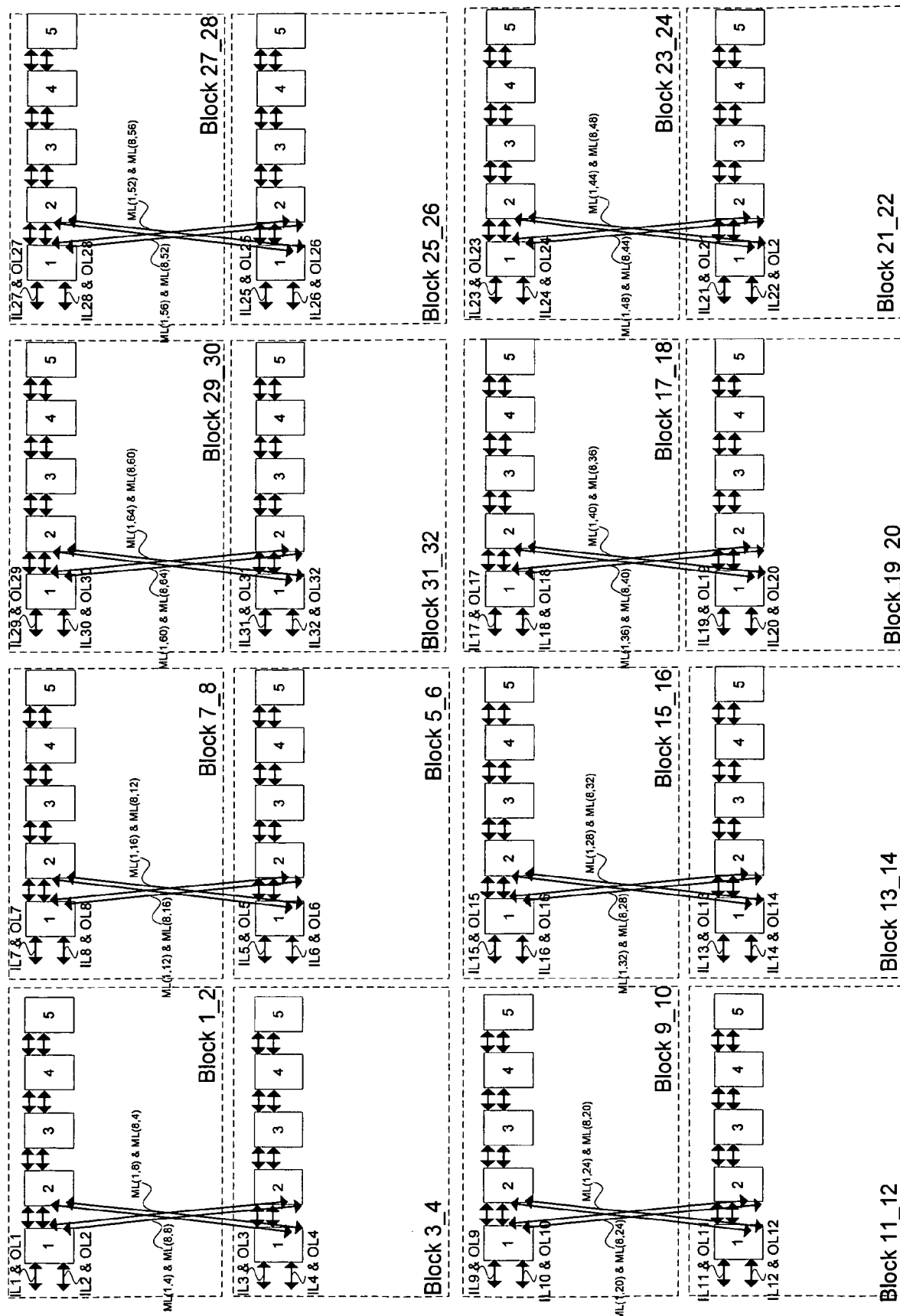
300C

FIG. 3C



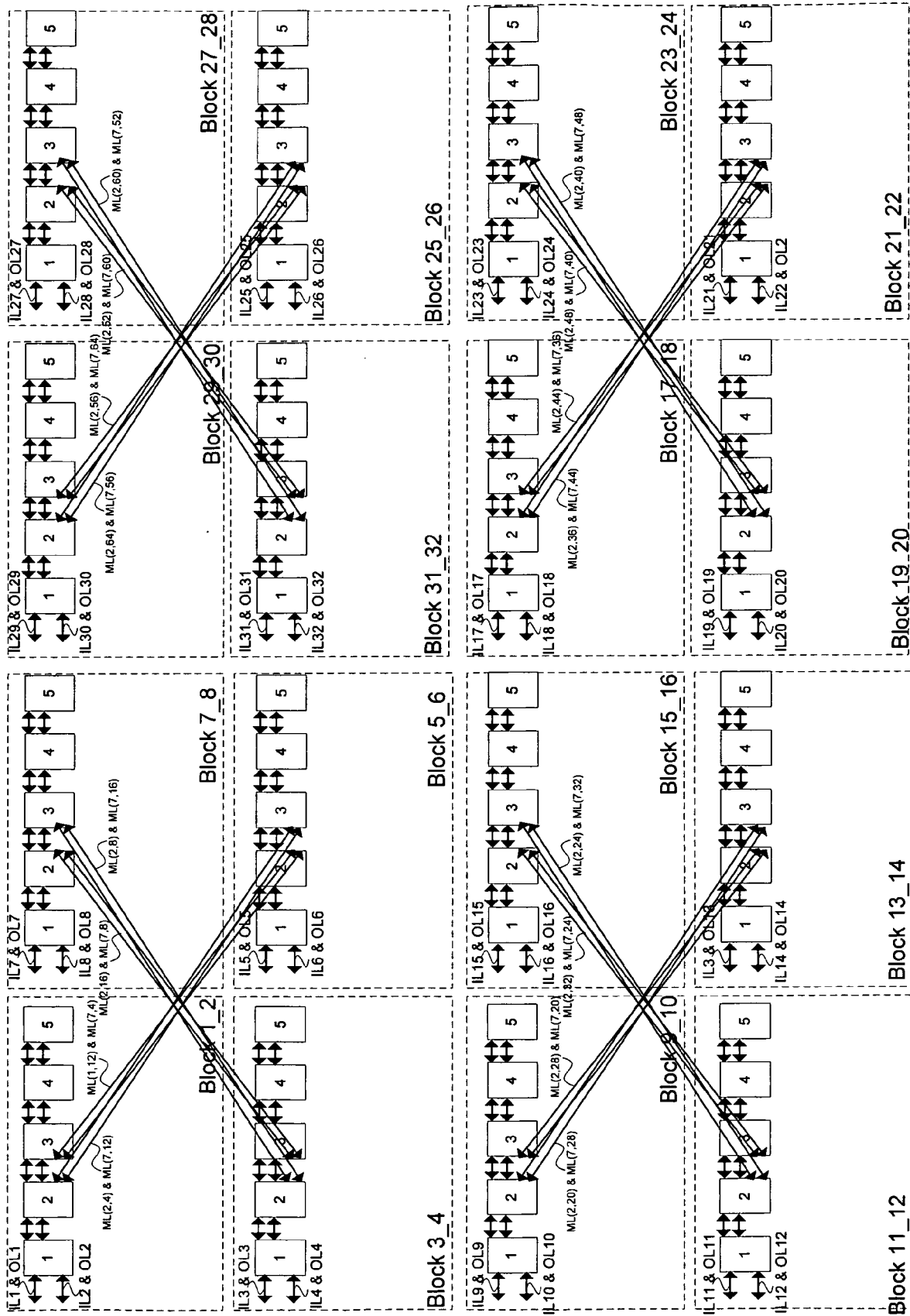
300D

FIG. 3D



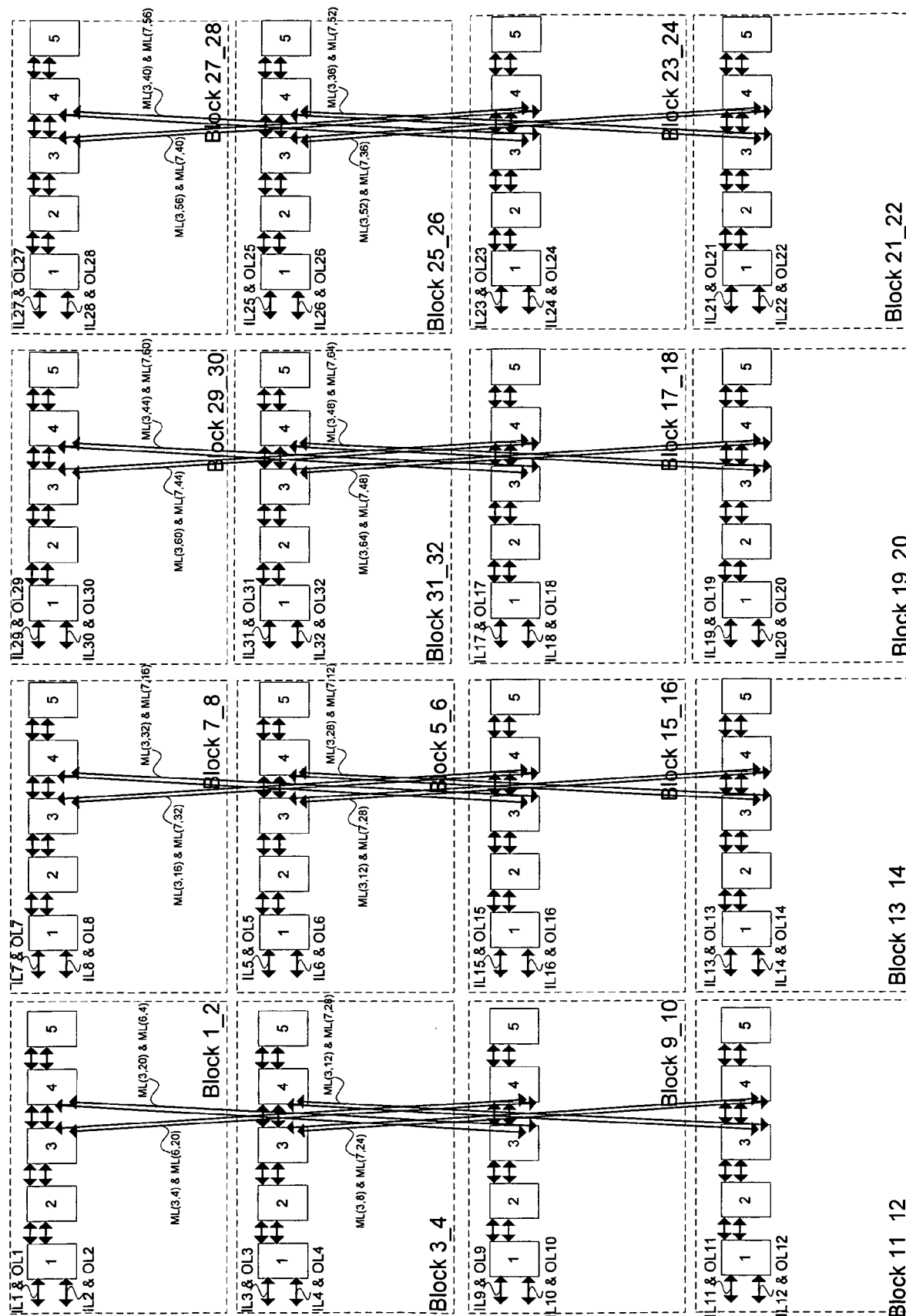
300E

FIG. 3E



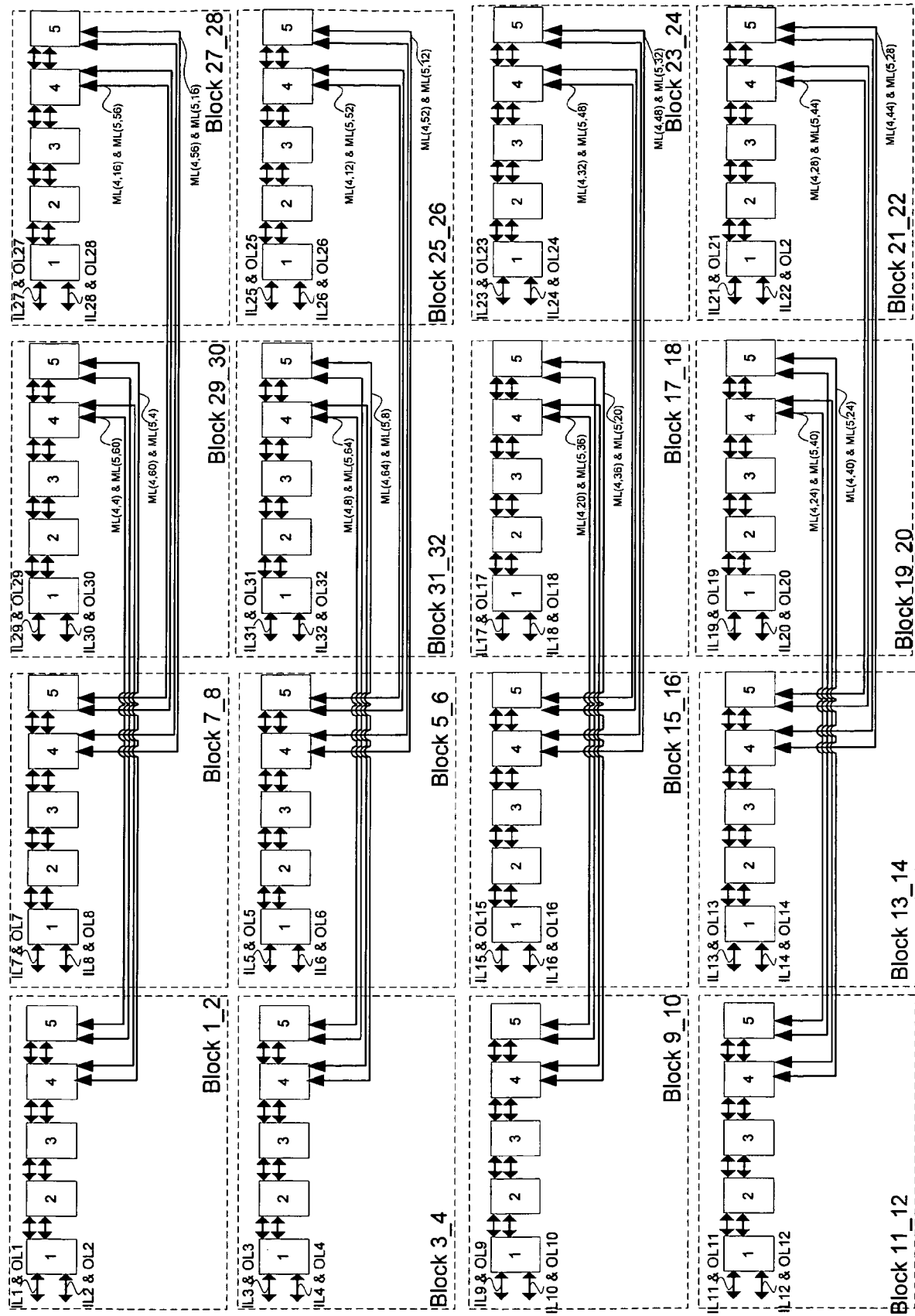
300F

FIG. 3F



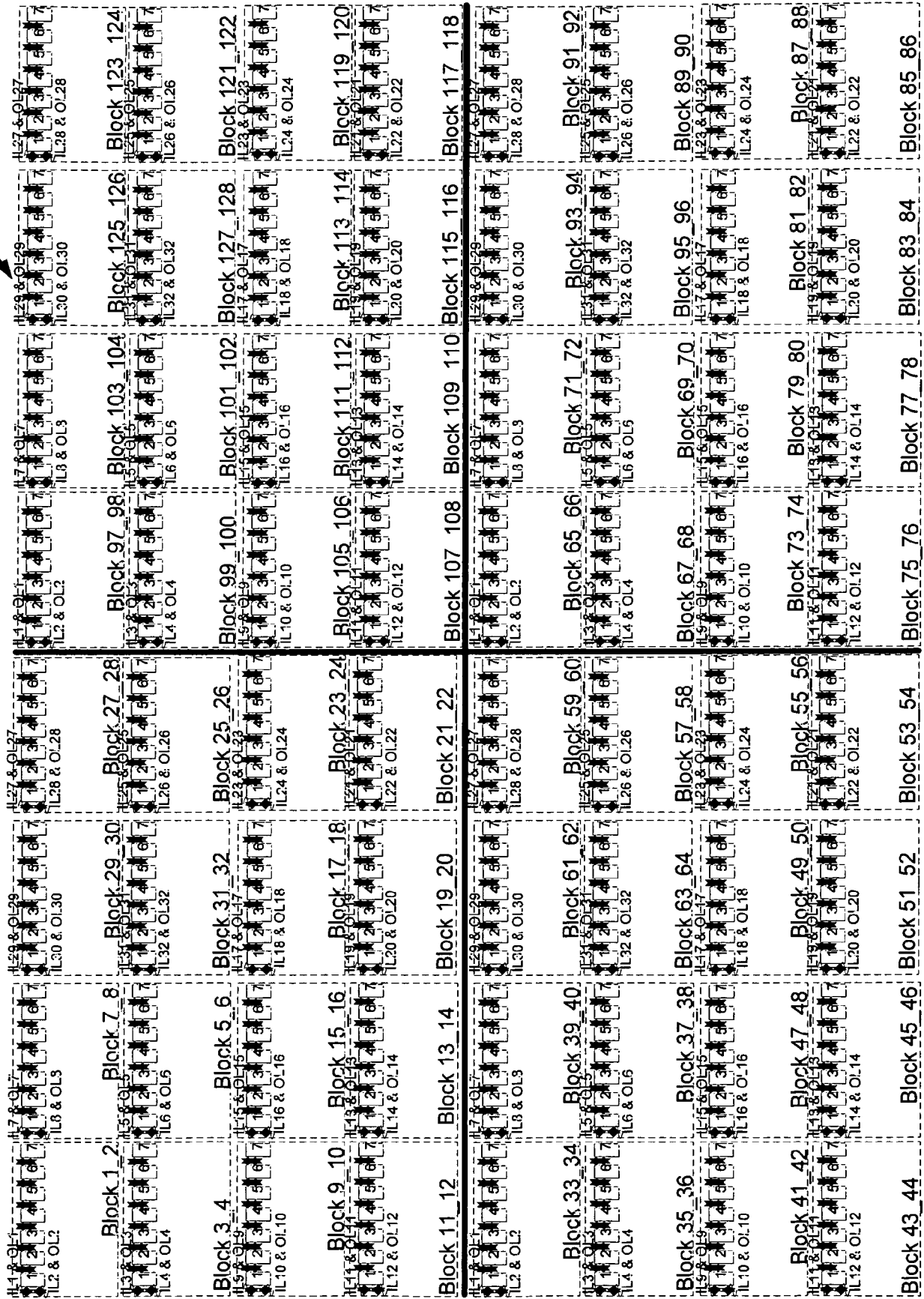
300G

FIG. 3G



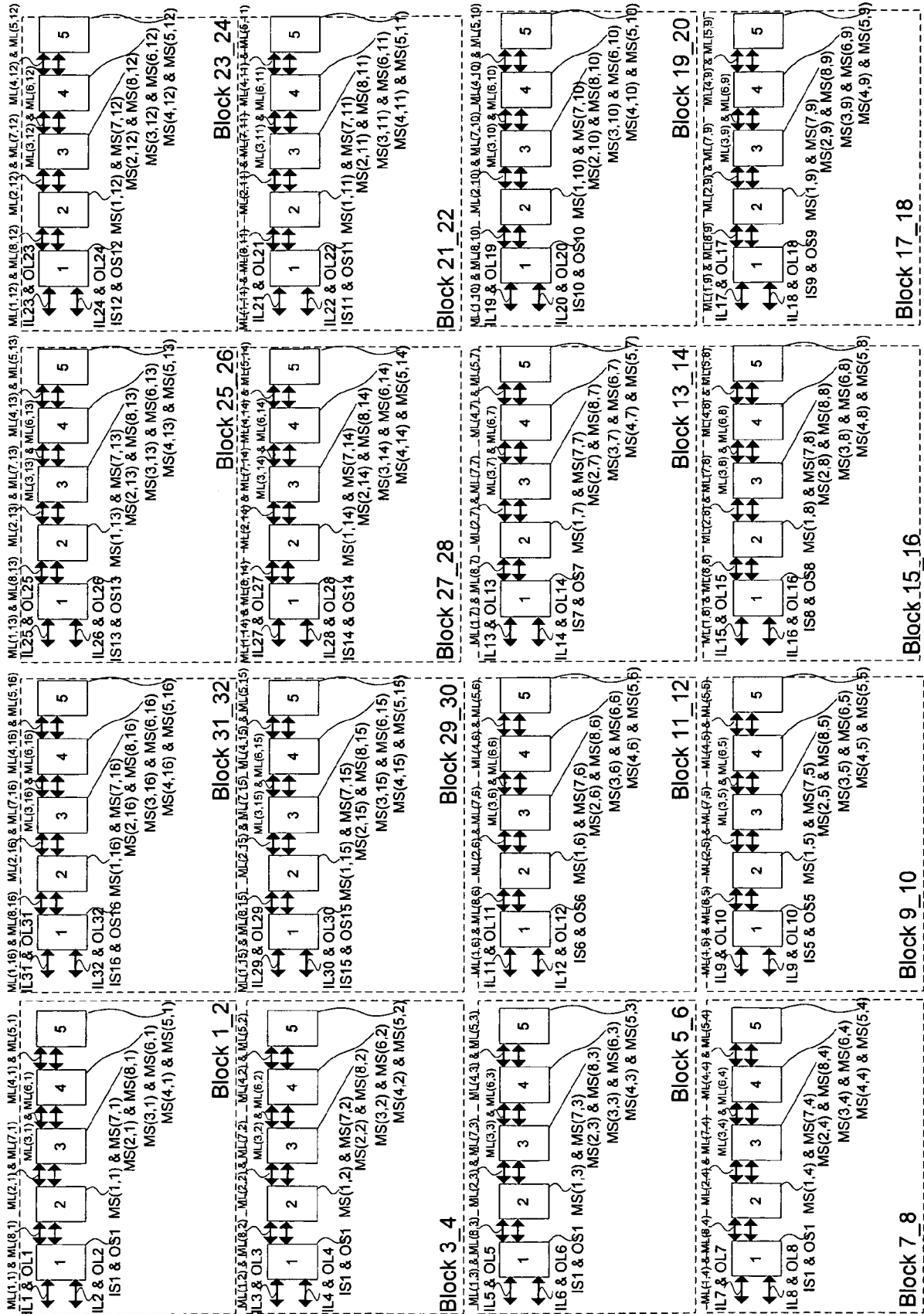
300H

FIG. 3H



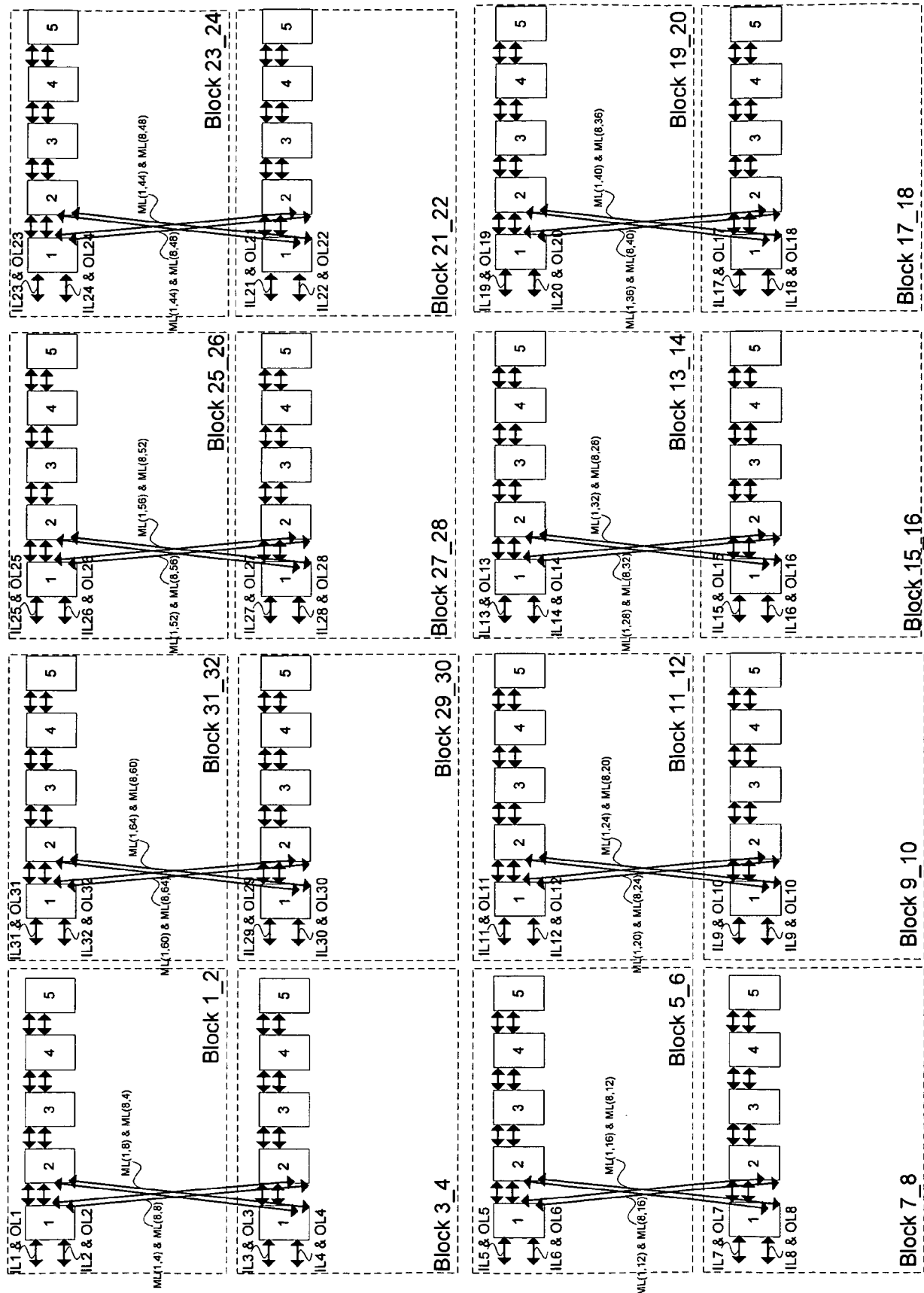
400A

FIG. 4A



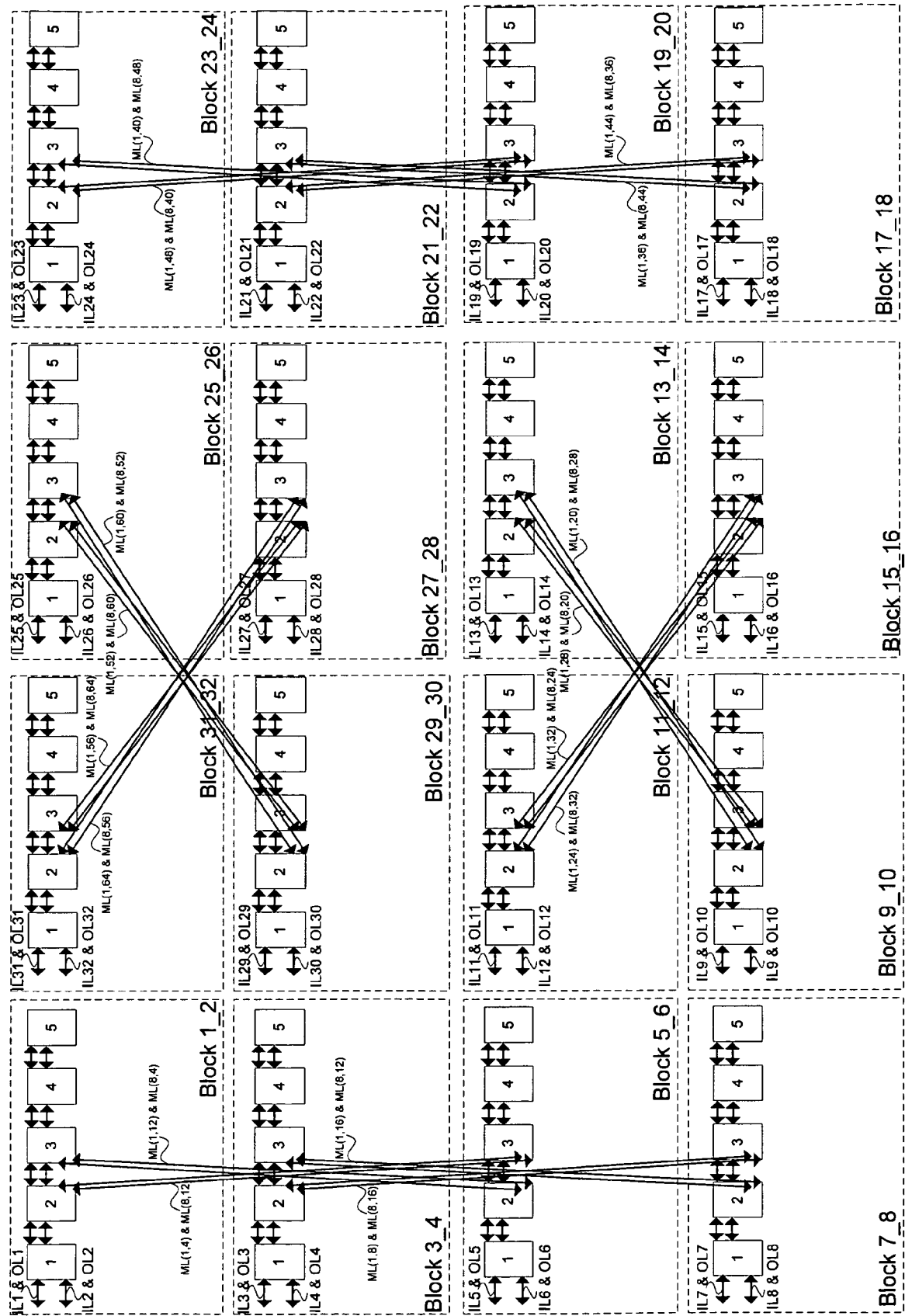
400B

FIG. 4B



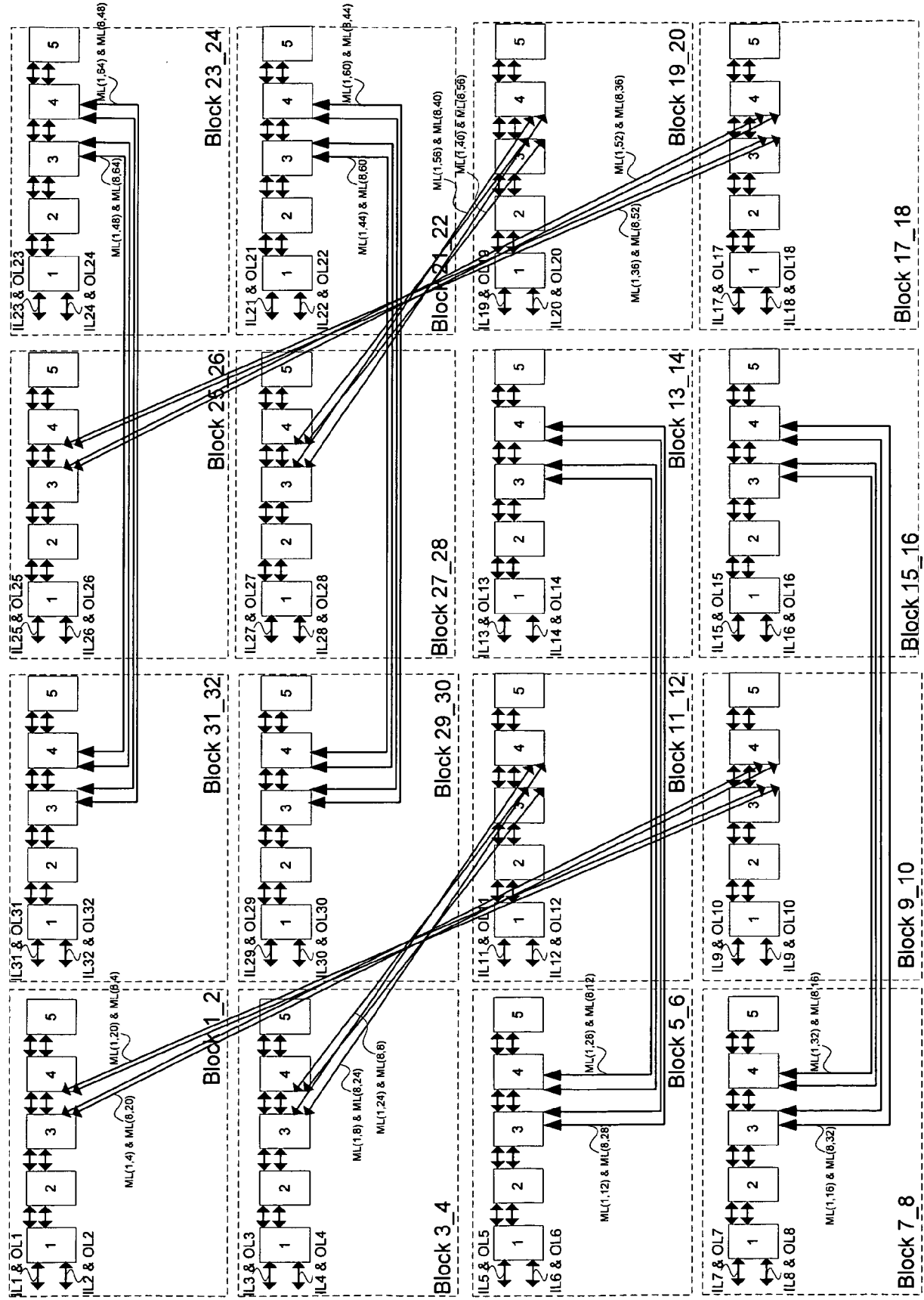
400C

FIG. 4C



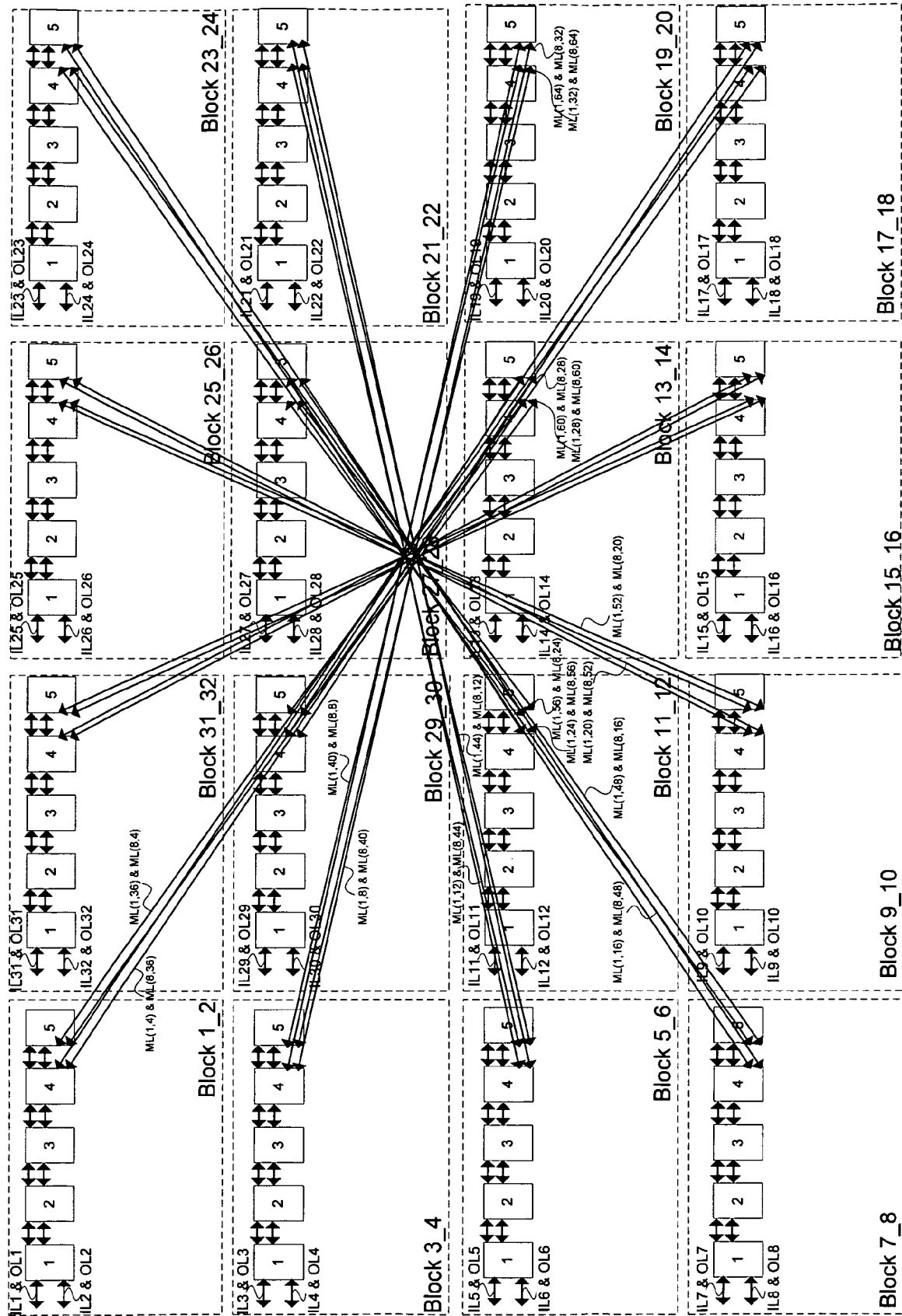
400D

FIG. 4D



400E

FIG. 4E



400C1

FIG. 4C1

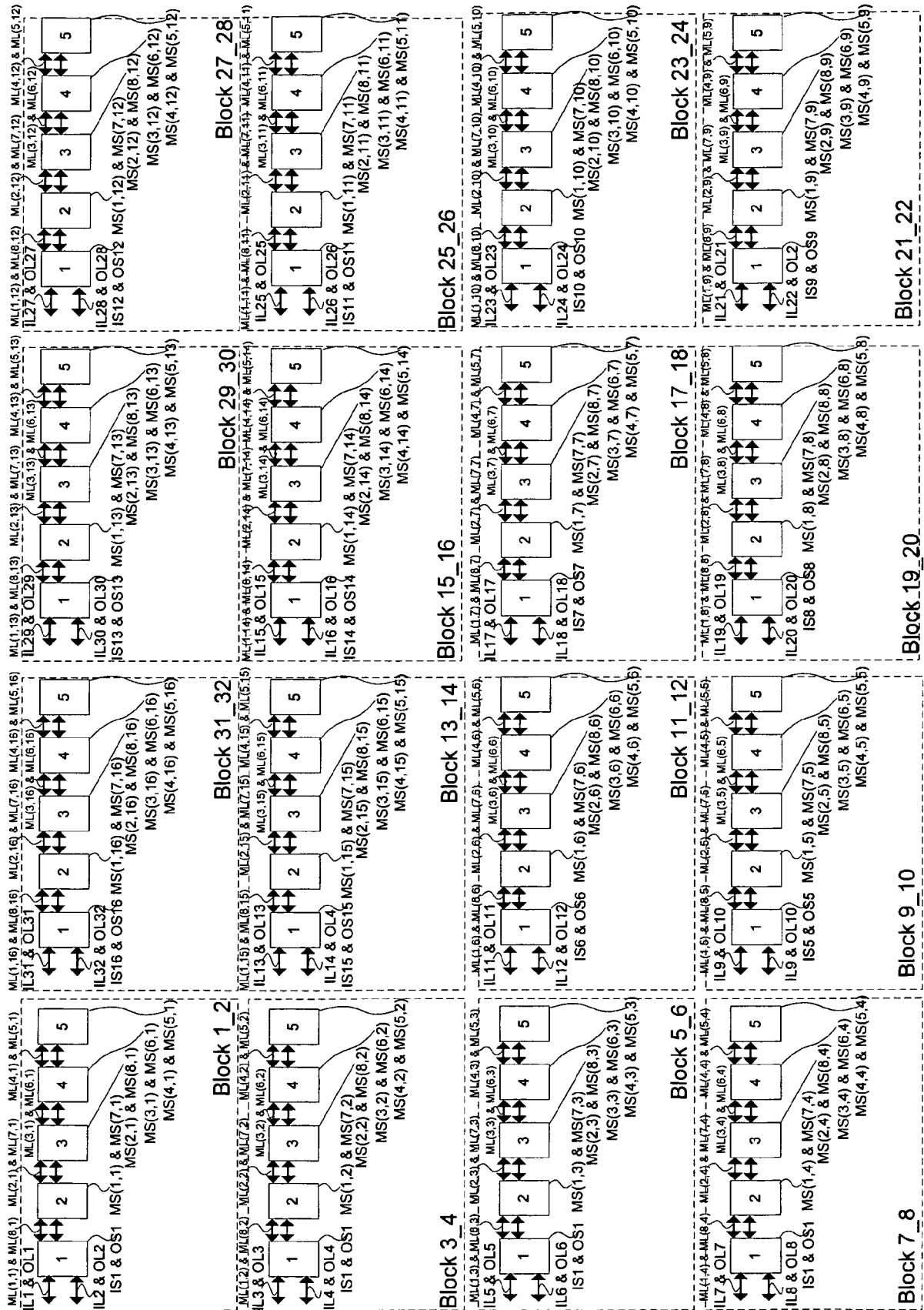


FIG. 5A

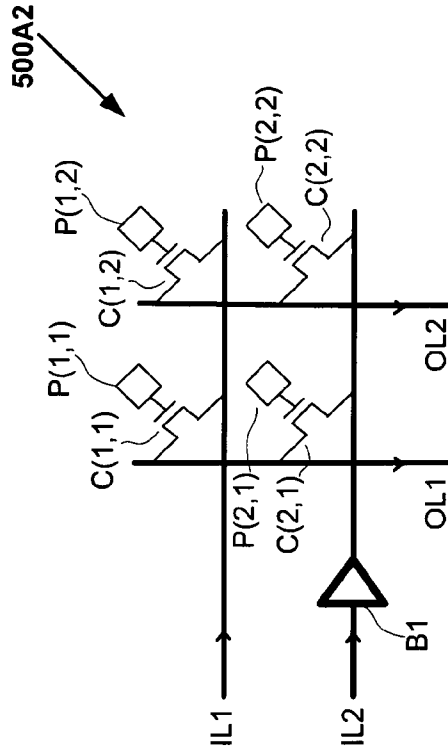


FIG. 5A2
(Prior Art)

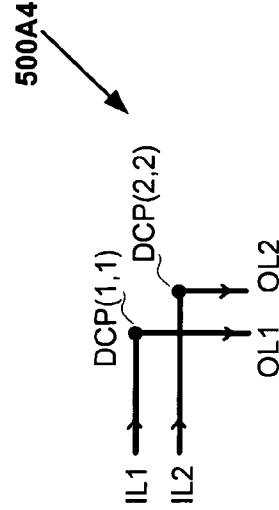


FIG. 5A4

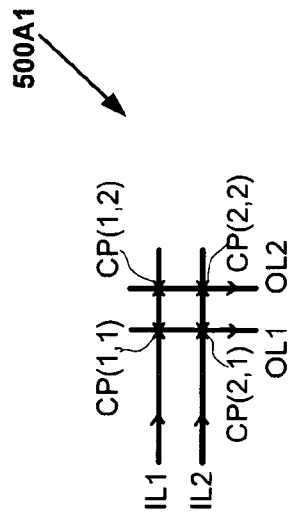


FIG. 5A1
(Prior Art)

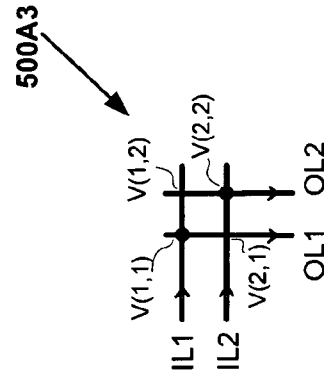


FIG. 5A3
(Prior Art)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2008/064605

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - H01L 25/00 (2008.04)

USPC - 326/41

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8) - H01L 25/00 (2008.04)

USPC - 326/41

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

MicroPatent

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6,940,308 B2 (WONG) 06 September 2005 (06.09.2005) entire document	1-8, 12, 16-19, 21-25, 29, 33, 40, 44, 48, 49
Y		9-11, 13-15, 20, 26-28, 30-32, 34-39, 41-43, 45-47
Y	US 7,154,887 B2 (WU et al) 26 December 2006 (26.12.2006) entire document	9-11, 13-15, 20, 26-28, 30-32, 34-39, 41-43, 45-47

 Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

25 August 2008

Date of mailing of the international search report

03 SEP 2008

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

EXHIBIT H

EXHIBIT H



US 20110037498A1

(19) **United States**

(12) **Patent Application Publication**
Konda

(10) **Pub. No.: US 2011/0037498 A1**

(43) **Pub. Date: Feb. 17, 2011**

(54) **VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS**

(52) **U.S. Cl. 326/41**

(57) **ABSTRACT**

(76) **Inventor: Venkat Konda, San Jose, CA (US)**

In accordance with the invention, VLSI layouts of generalized multi-stage networks for broadcast, unicast and multicast connections are presented using only horizontal and vertical links. The VLSI layouts employ shuffle exchange links where outlet links of cross links from switches in a stage in one sub-integrated circuit block are connected to inlet links of switches in the succeeding stage in another sub-integrated circuit block so that said cross links are either vertical links or horizontal and vice versa. In one embodiment the sub-integrated circuit blocks are arranged in a hypercube arrangement in a two-dimensional plane. The VLSI layouts exploit the benefits of significantly lower cross points, lower signal latency, lower power and full connectivity with significantly fast compilation.

Correspondence Address:
Konda Technologies, Inc
6278 GRAND OAK WAY
SAN JOSE, CA 95135 (US)

(21) **Appl. No.: 12/601,275**

(22) **PCT Filed: May 22, 2008**

(86) **PCT No.: PCT/US08/64605**

§ 371 (c)(1),
(2), (4) **Date: May 31, 2010**

Related U.S. Application Data

(60) **Provisional application No. 60/940,394, filed on May 25, 2007.**

Publication Classification

(51) **Int. Cl. H01L 25/00 (2006.01)**

The VLSI layouts presented are applicable to generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mink-bft}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ for $s=1, 2, 3$ or any number in general. The embodiments of VLSI layouts are useful in wide target applications such as FPGAs, CPLDs, pSoCs, ASIC placement and route tools, networking applications, parallel & distributed computing, and reconfigurable computing.

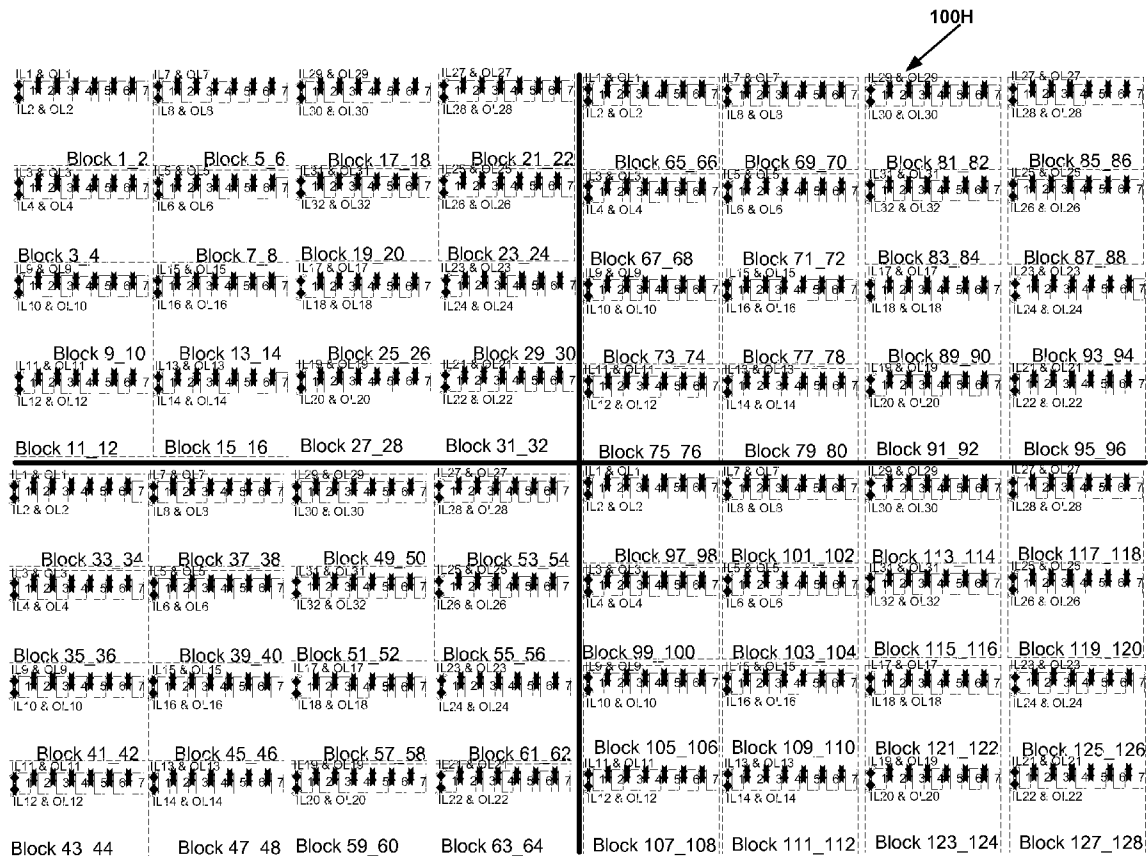


FIG. 1A

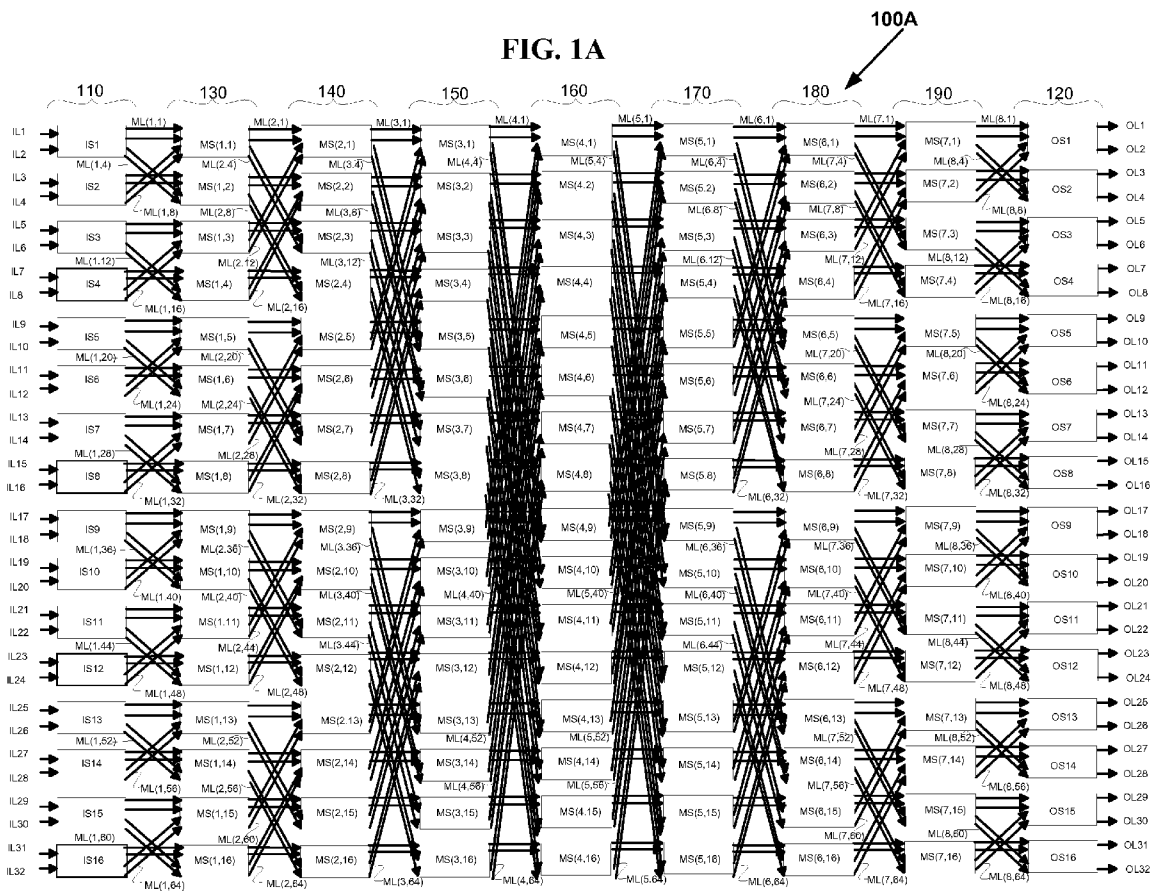
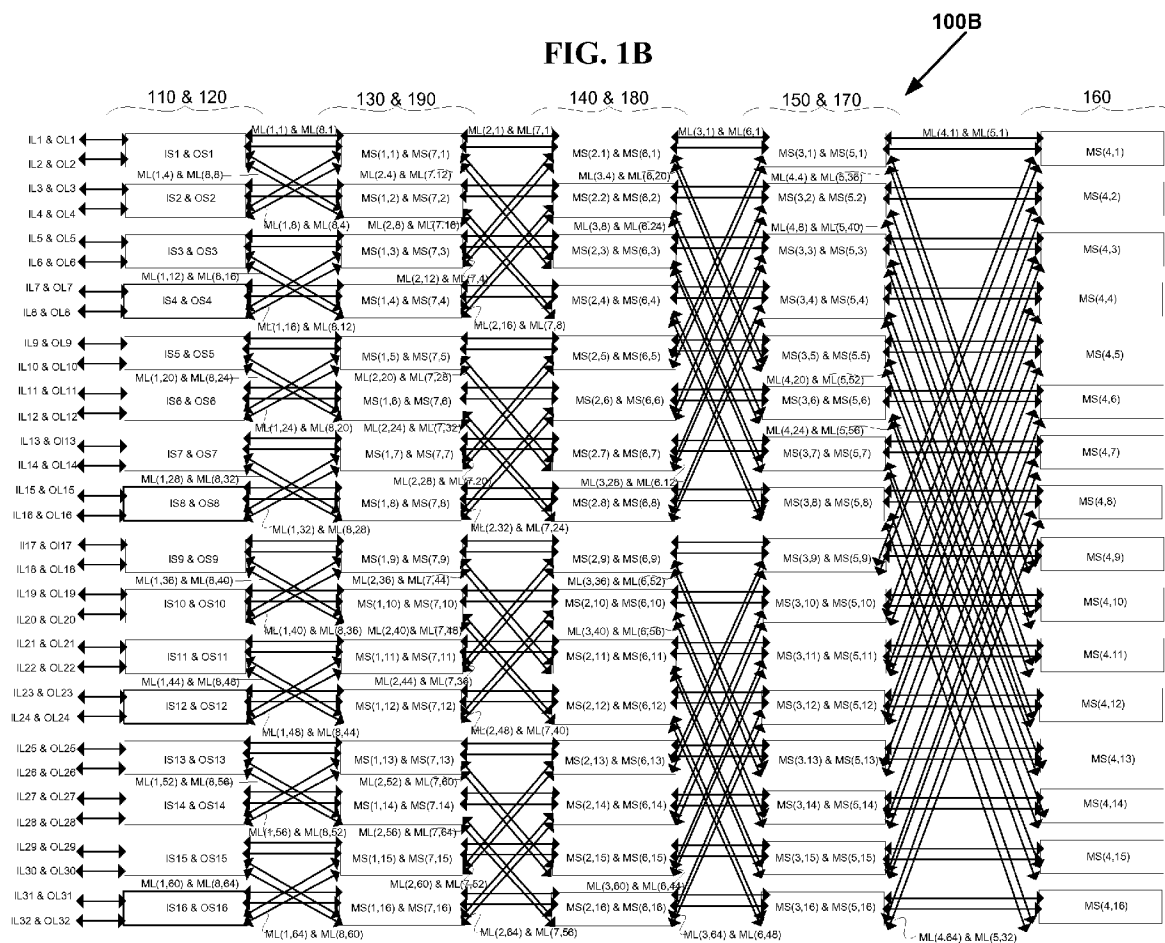


FIG. 1B



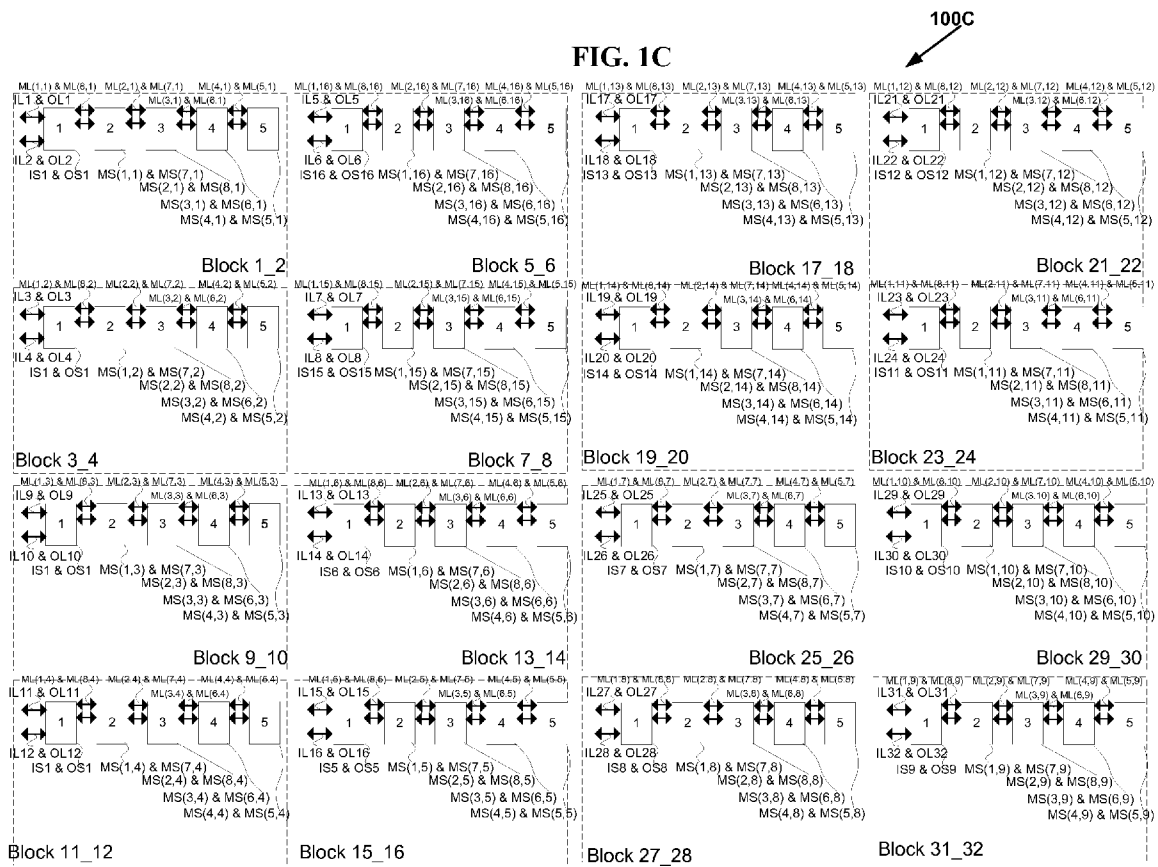


FIG. 1D

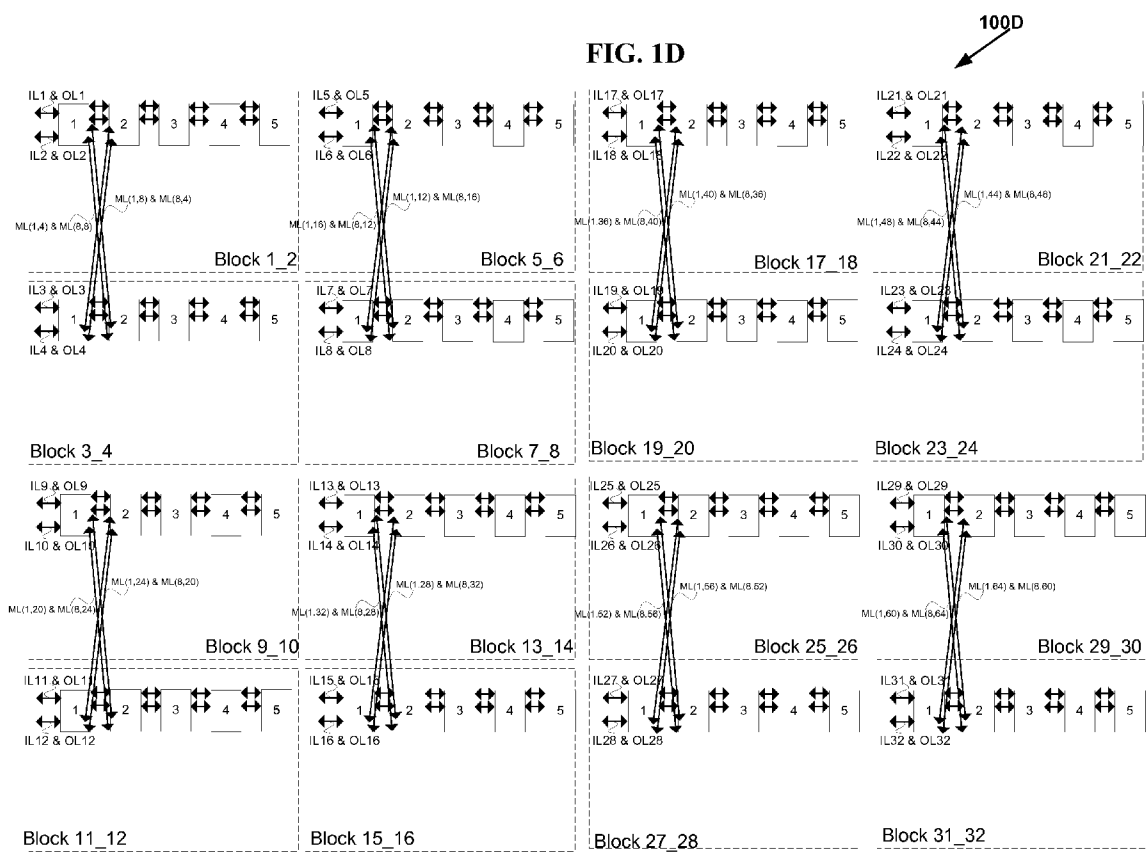
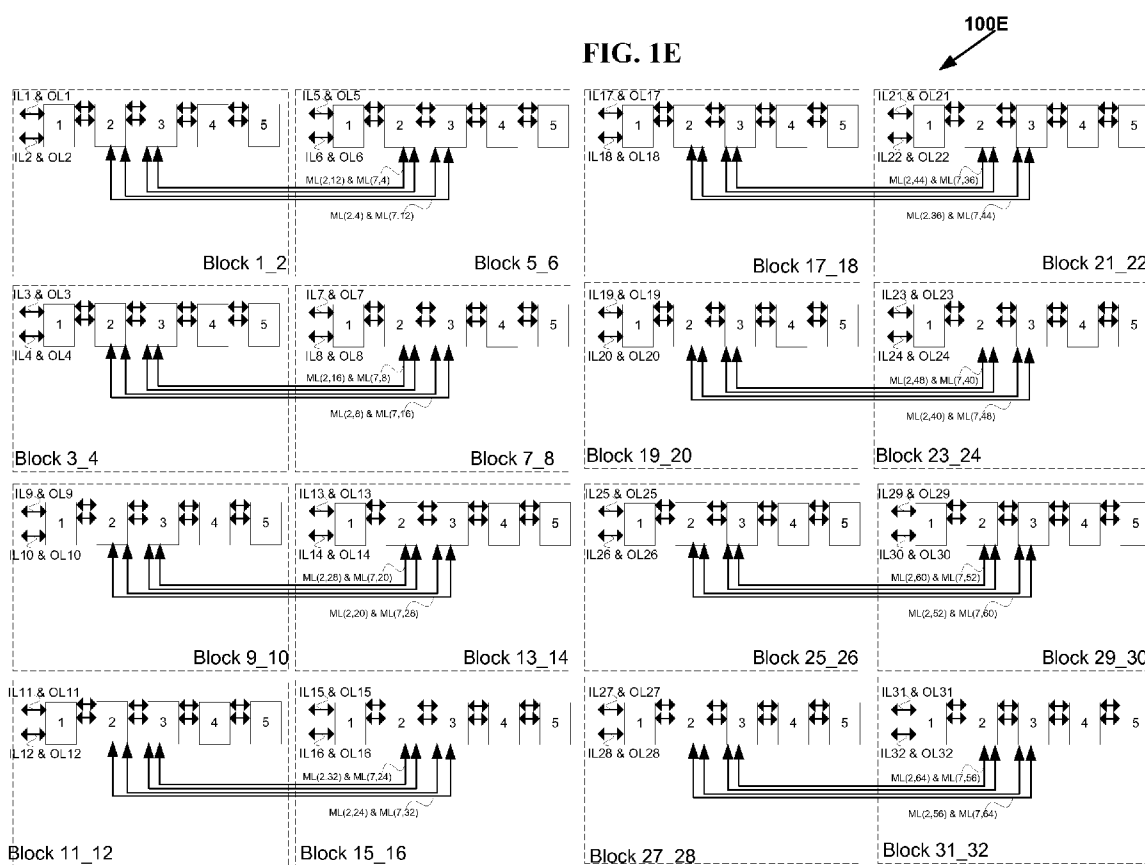
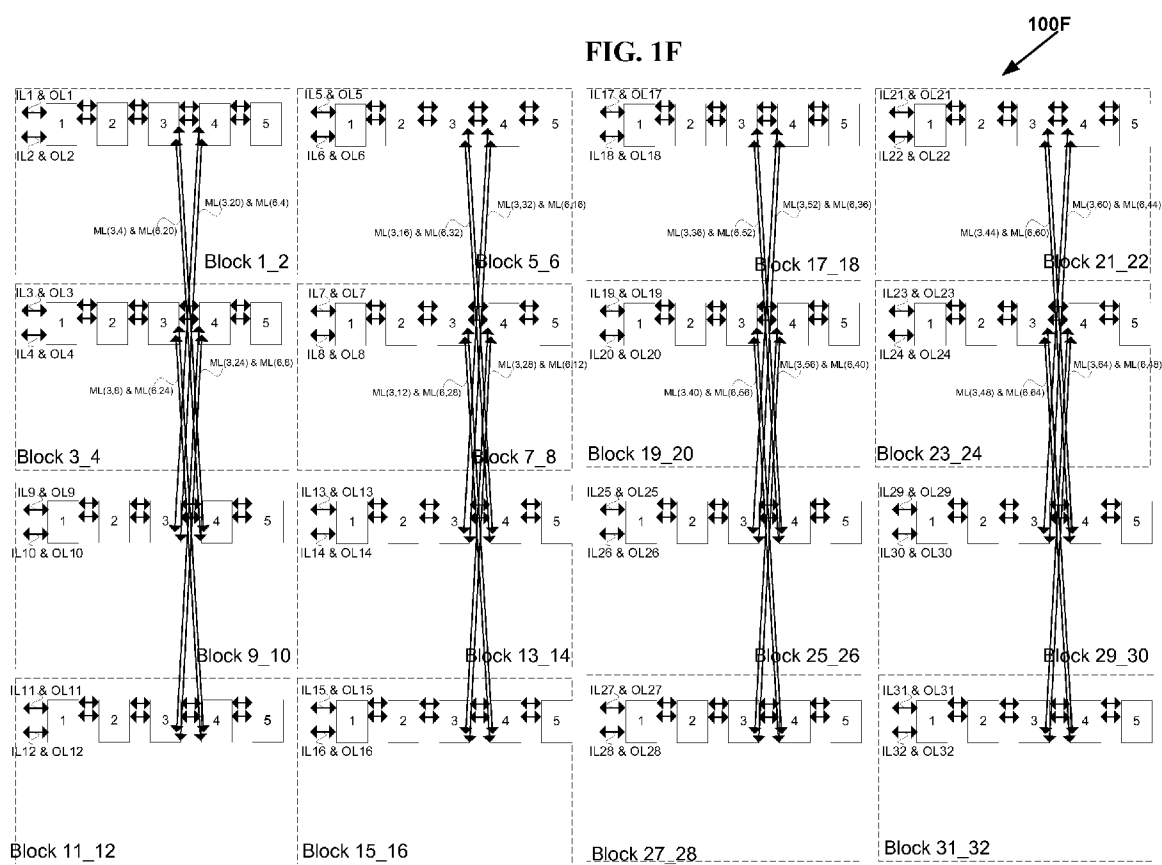
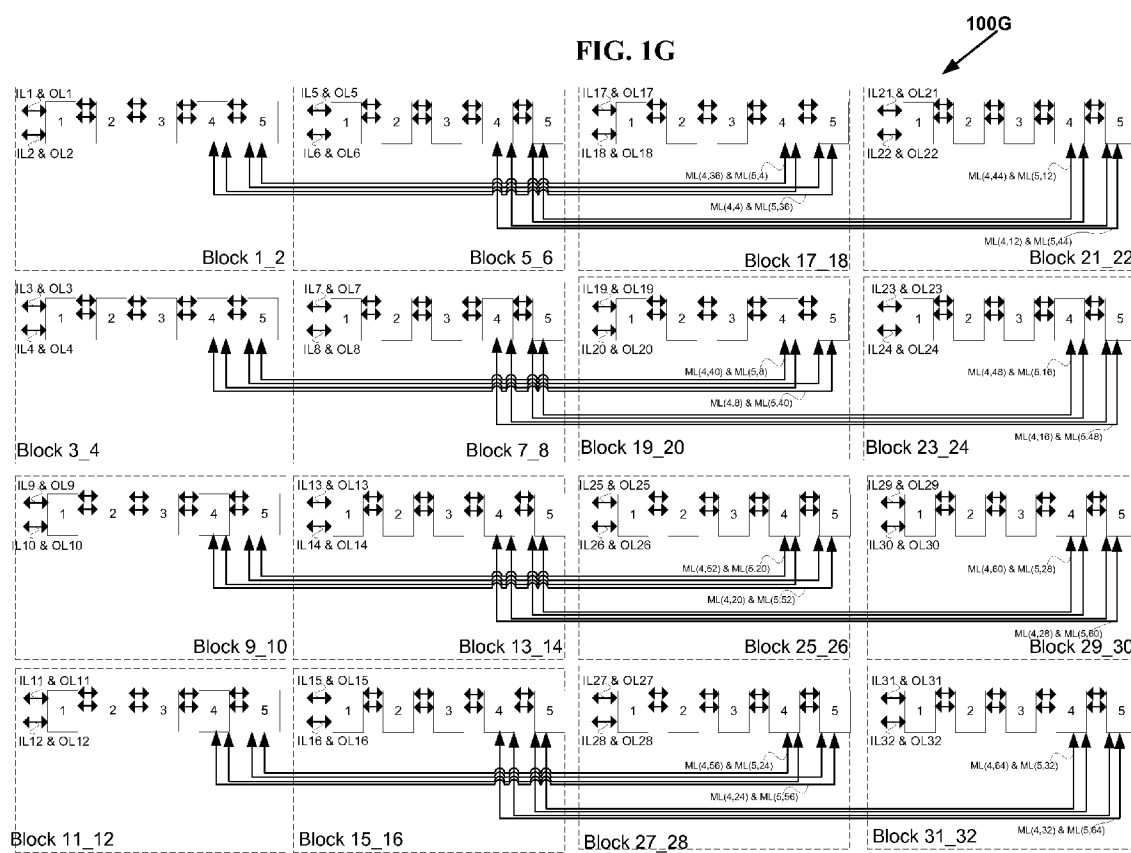


FIG. 1E







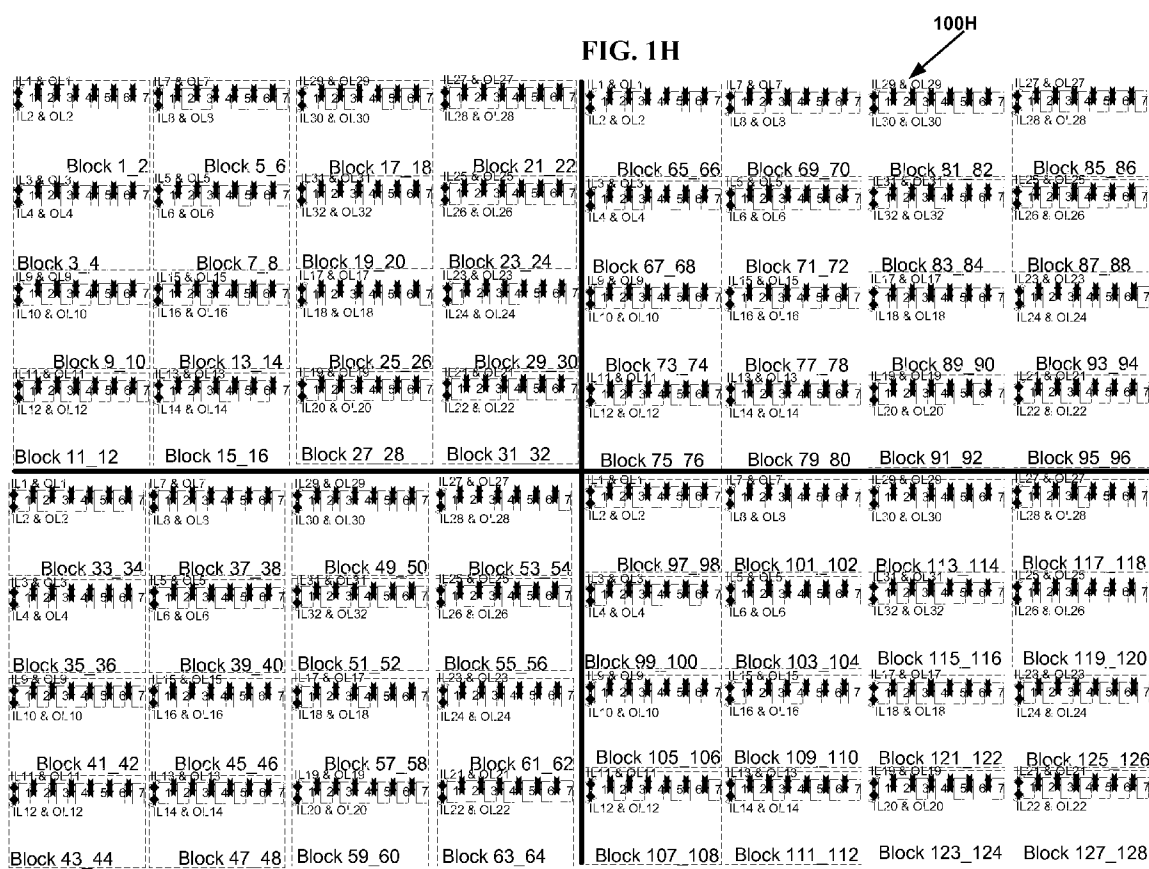


FIG. 11

1001

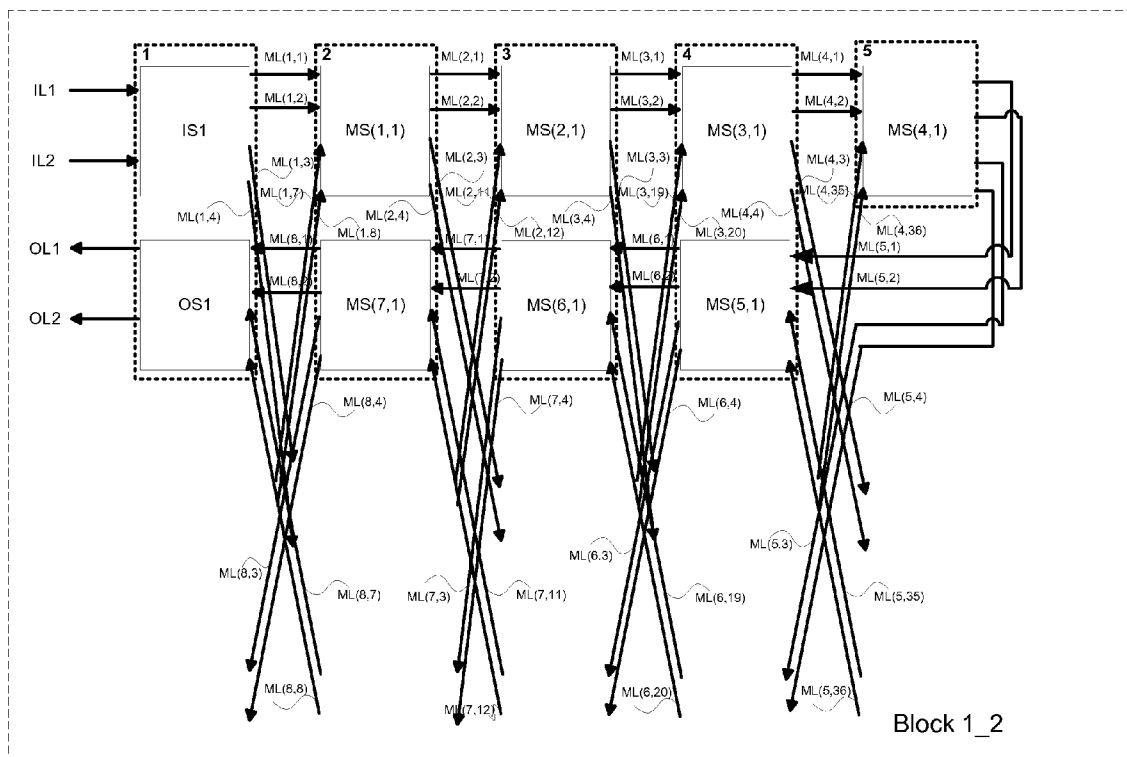


FIG. 1J

100J

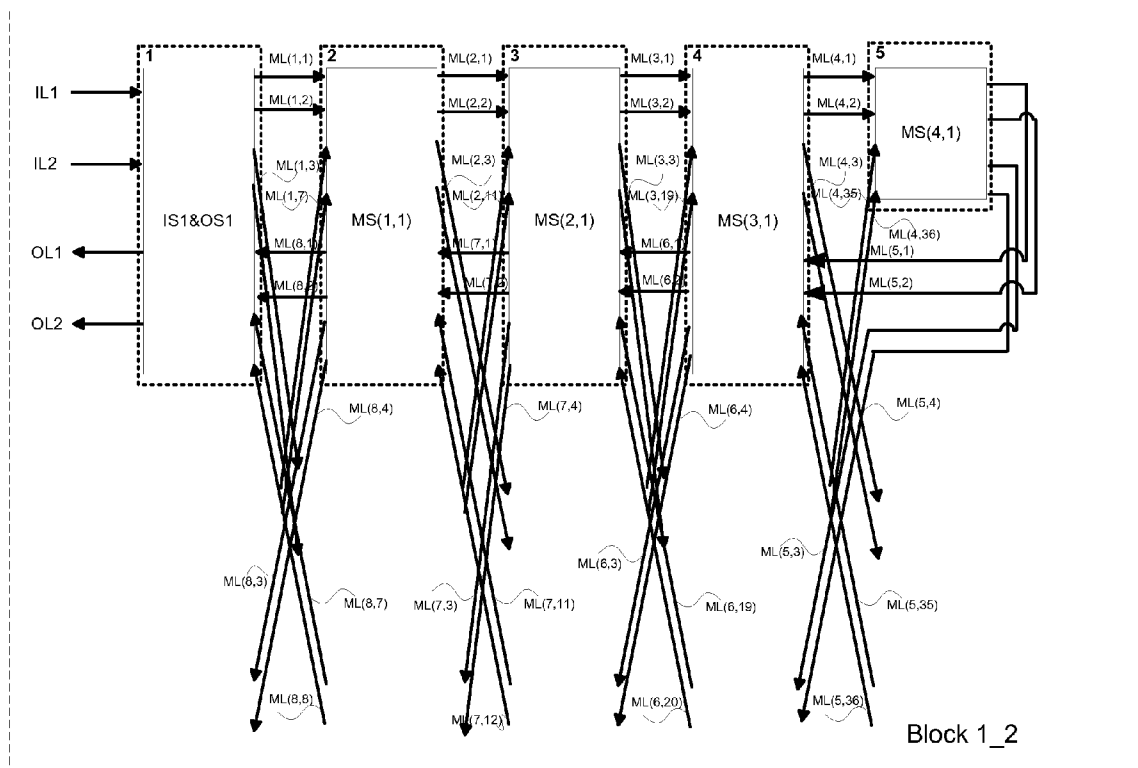


FIG. 1K

100K

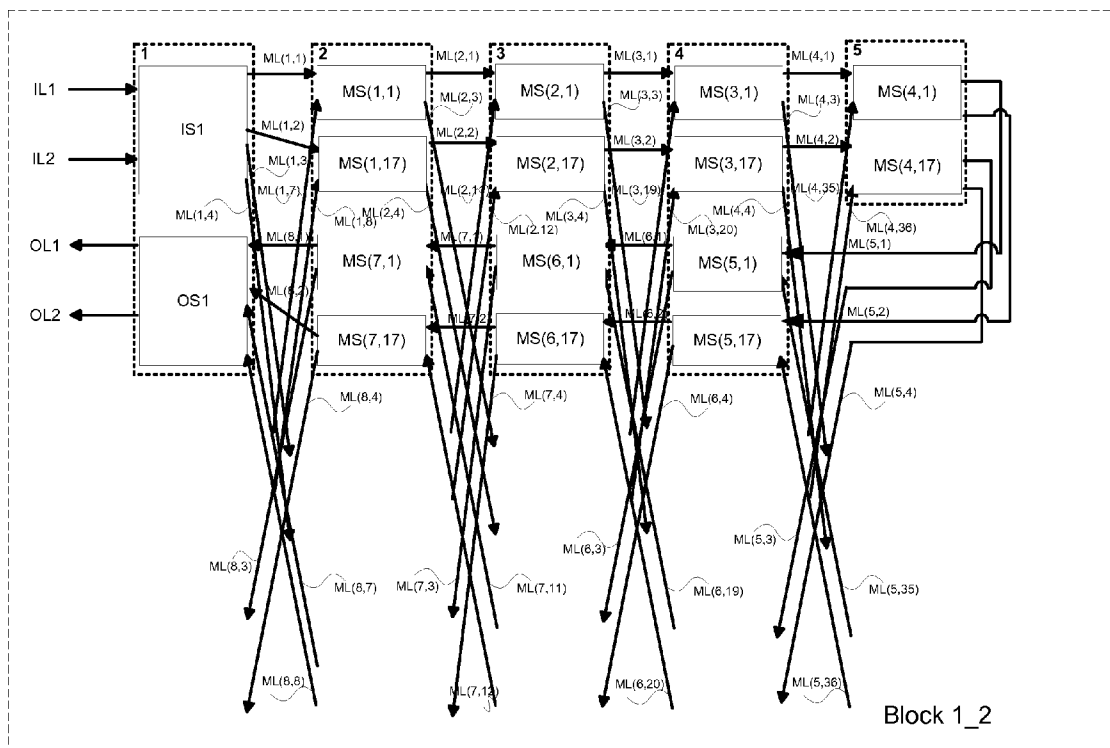


FIG. 1K1

100K1

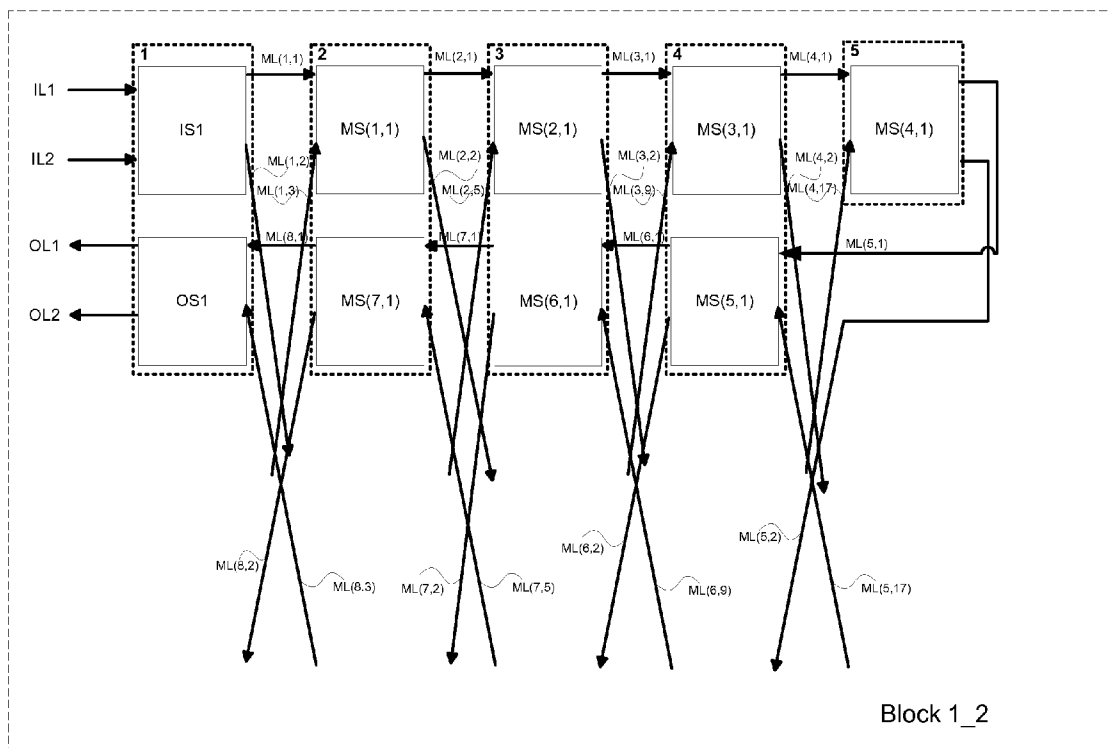


FIG. 1L

100L

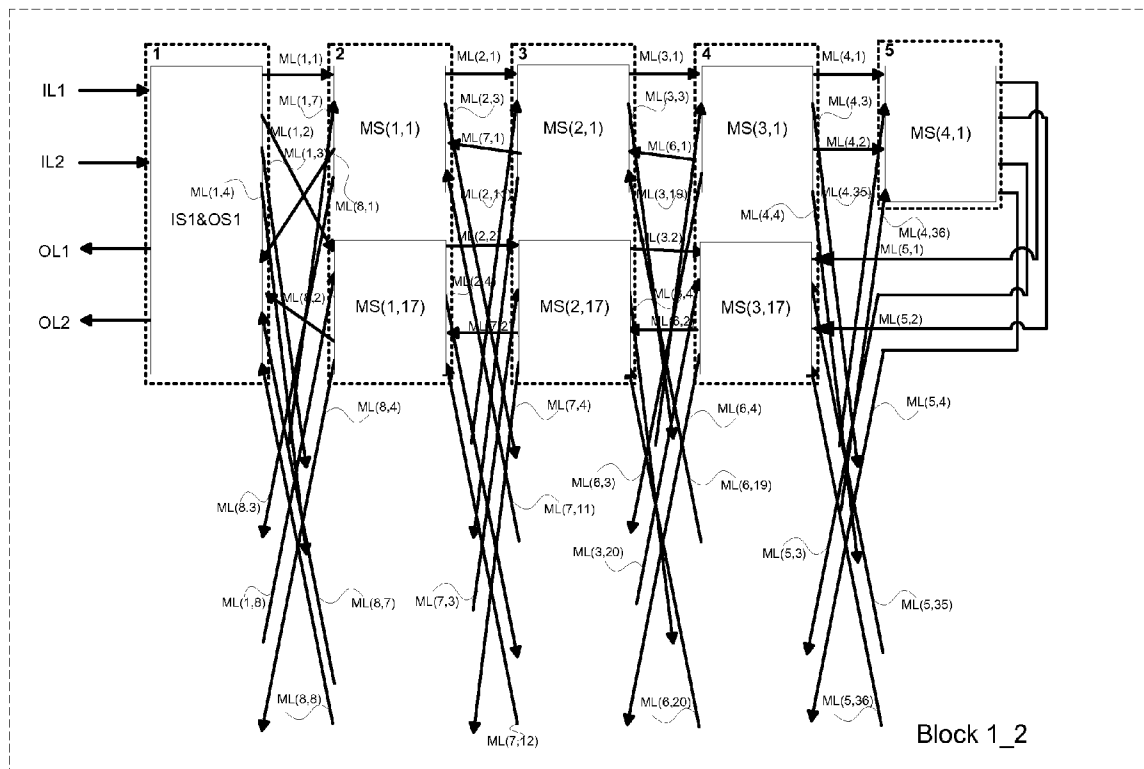


FIG. 1L1

100L1

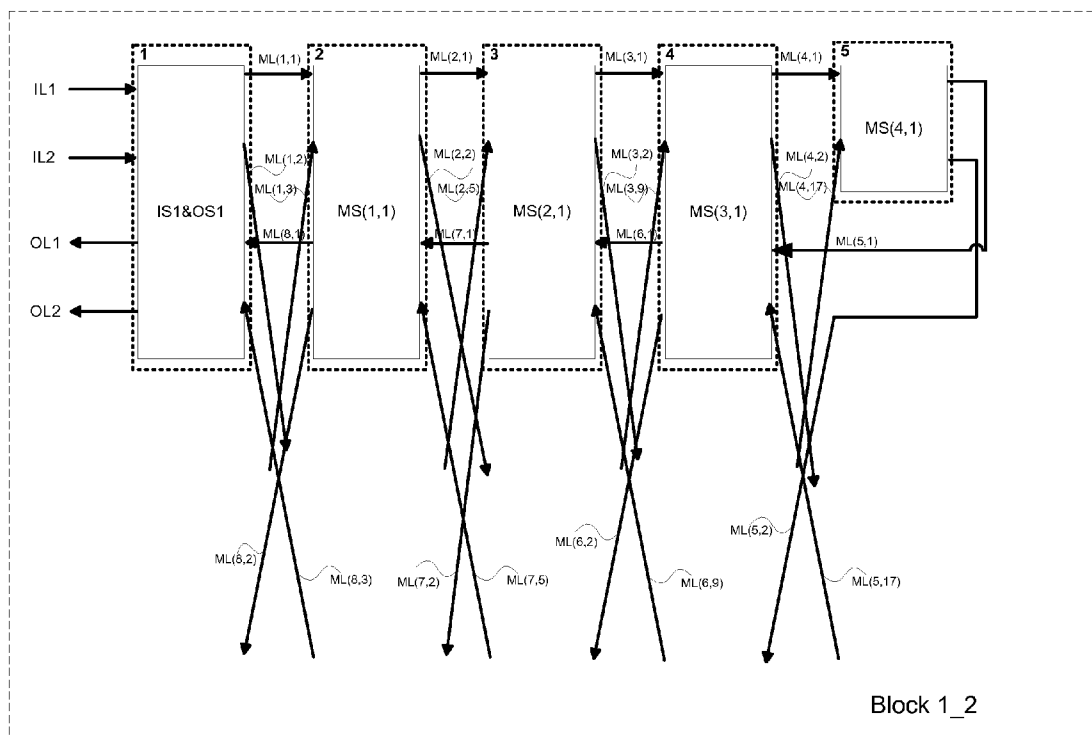


FIG. 2A1

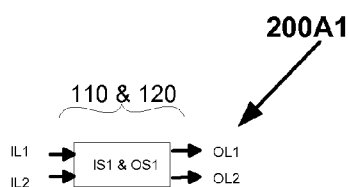


FIG. 2A2

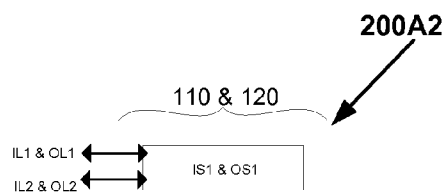


FIG. 2A3

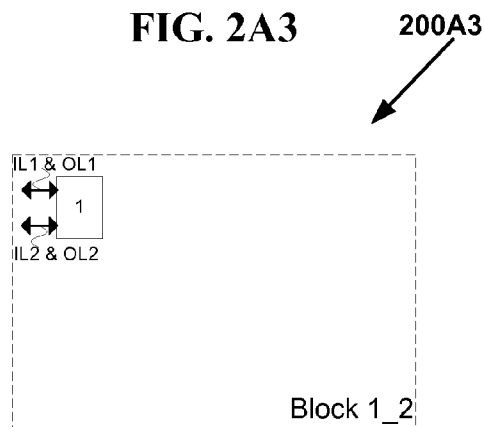


FIG. 2B1

200B1

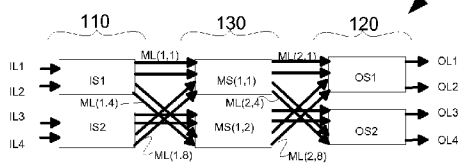


FIG. 2B2

200B2

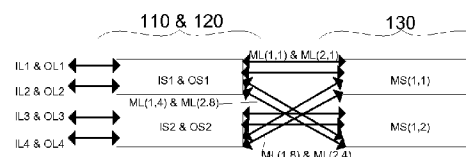


FIG. 2B3

200B3

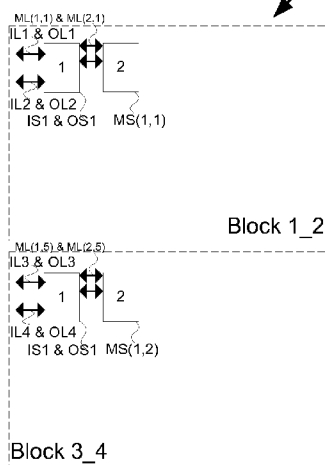


FIG. 2B4

200B4

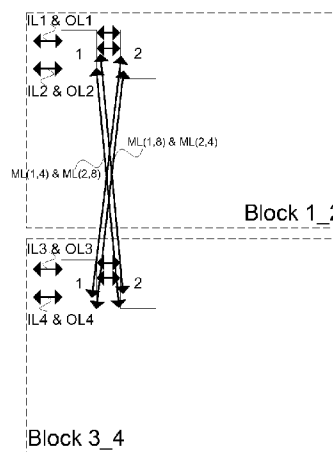


FIG. 2C11

200C11

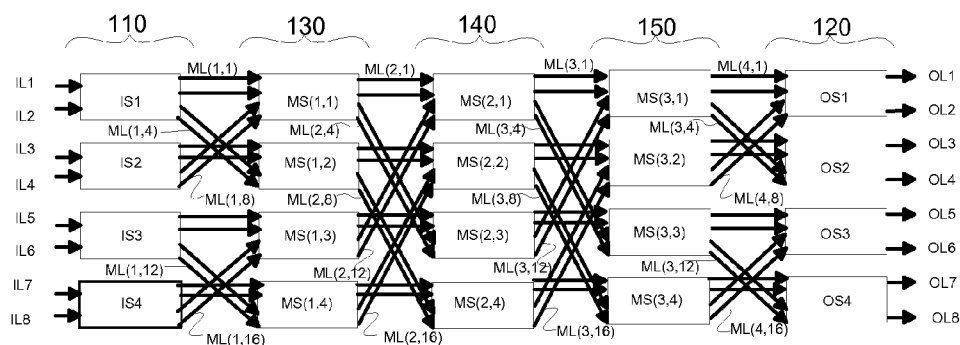


FIG. 2C12

200C12

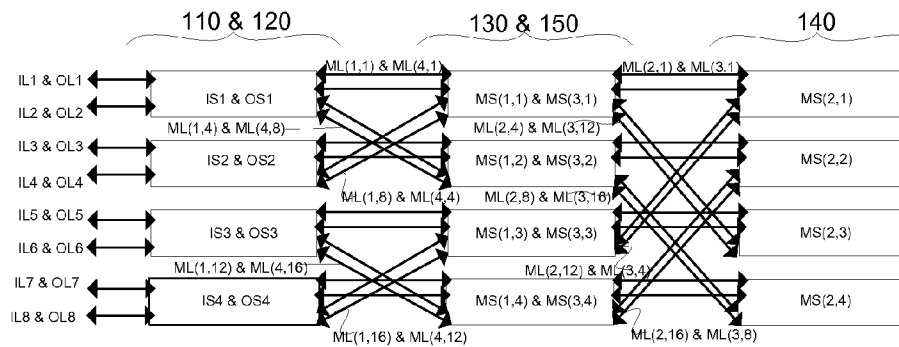


FIG. 2C21

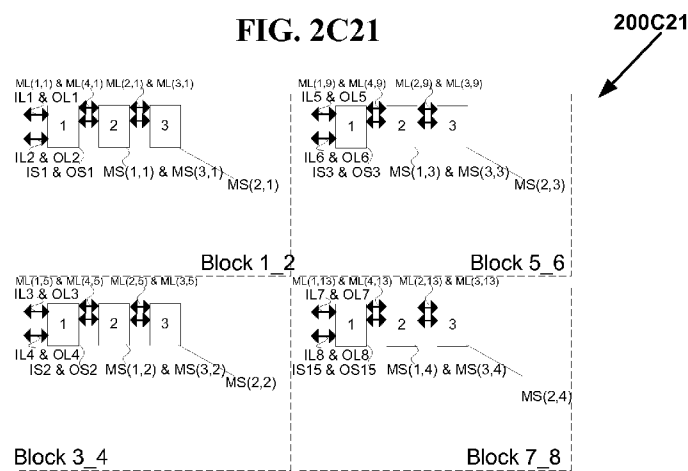


FIG. 2C22

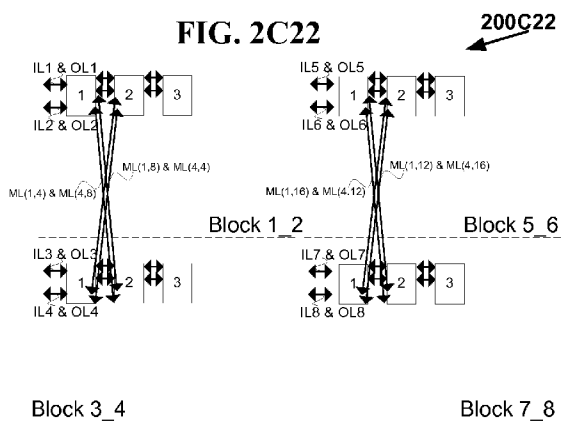


FIG. 2C23

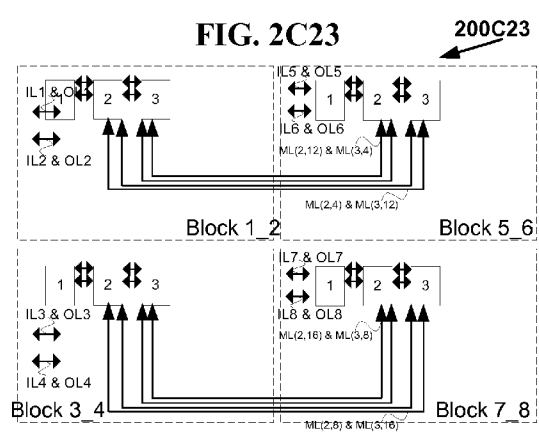


FIG. 2D1

200D1

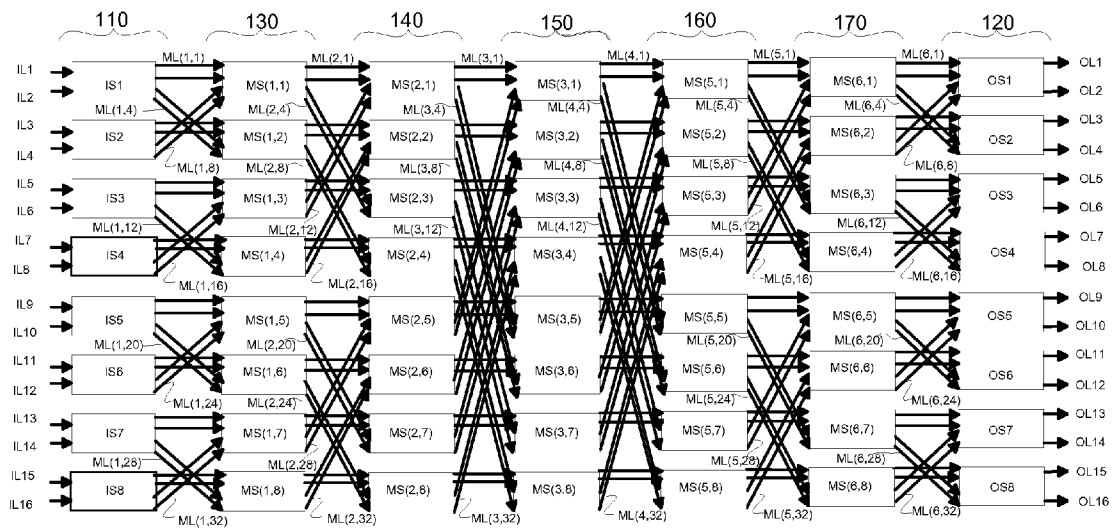


FIG. 2D2

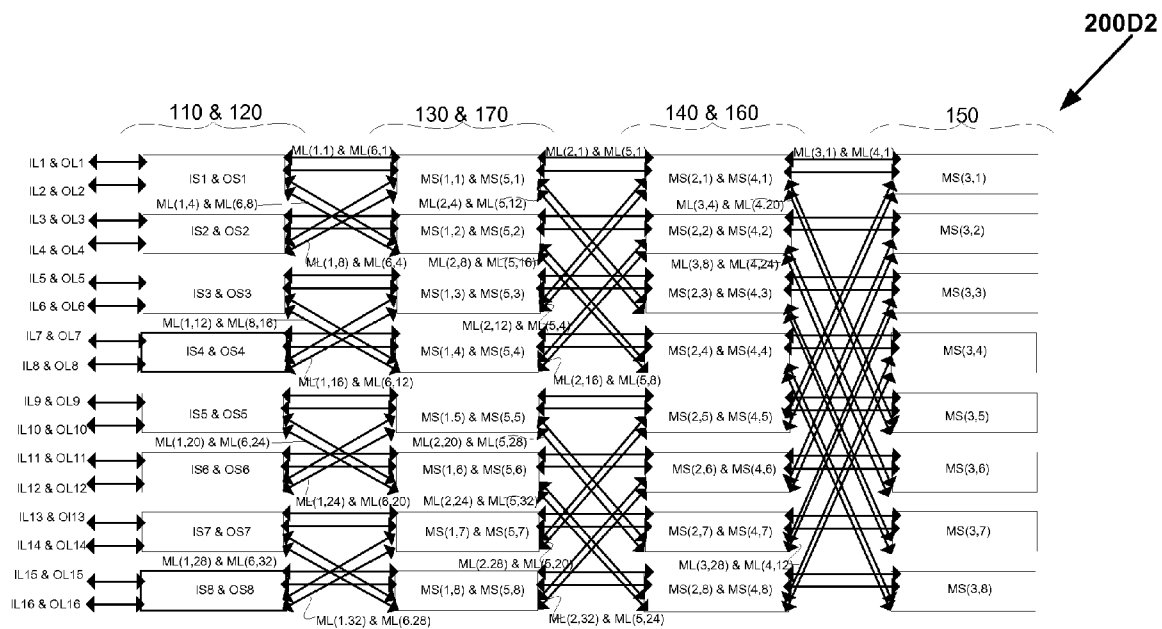
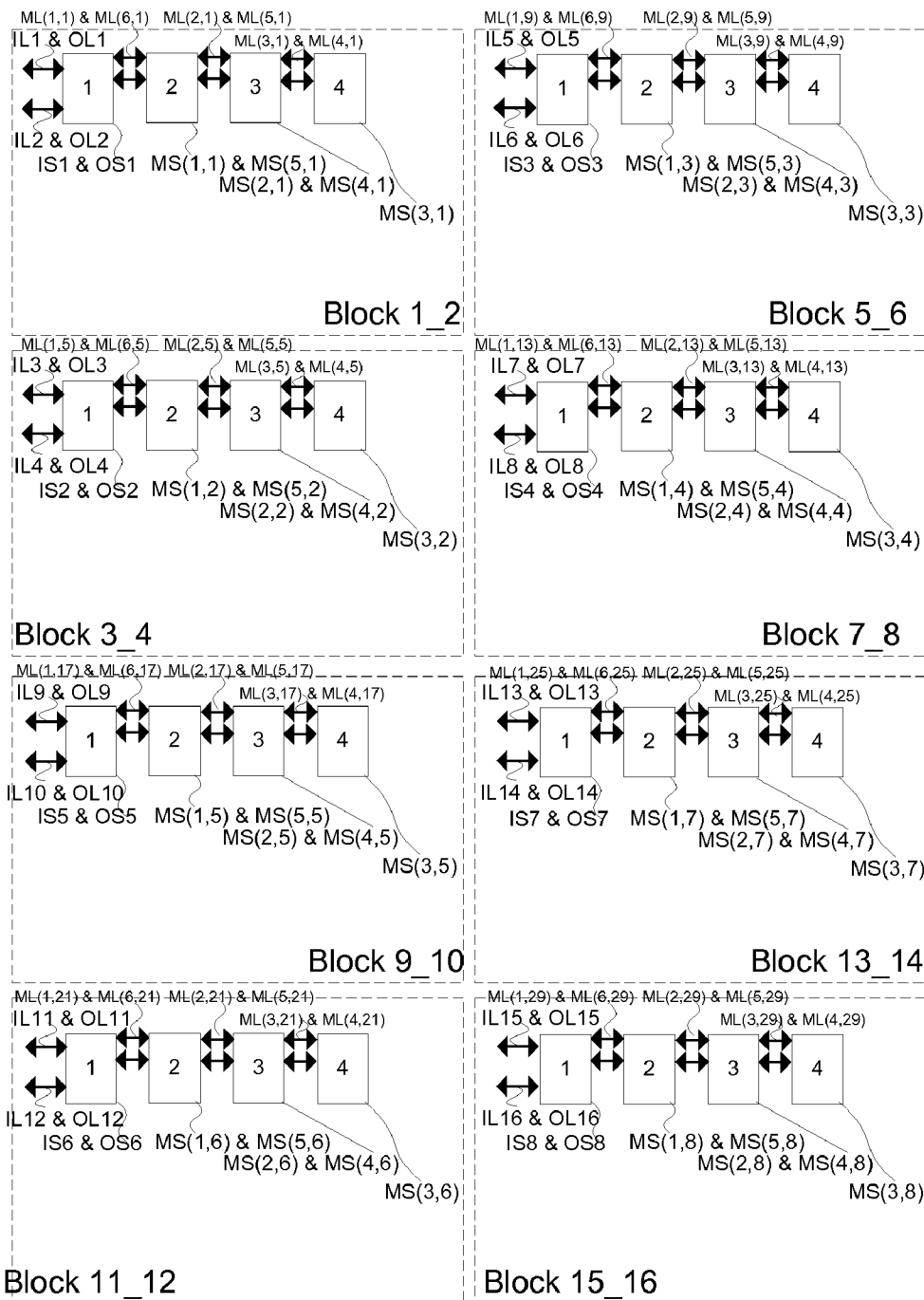


FIG. 2D3

200D3



200D4
↙

FIG. 2D4

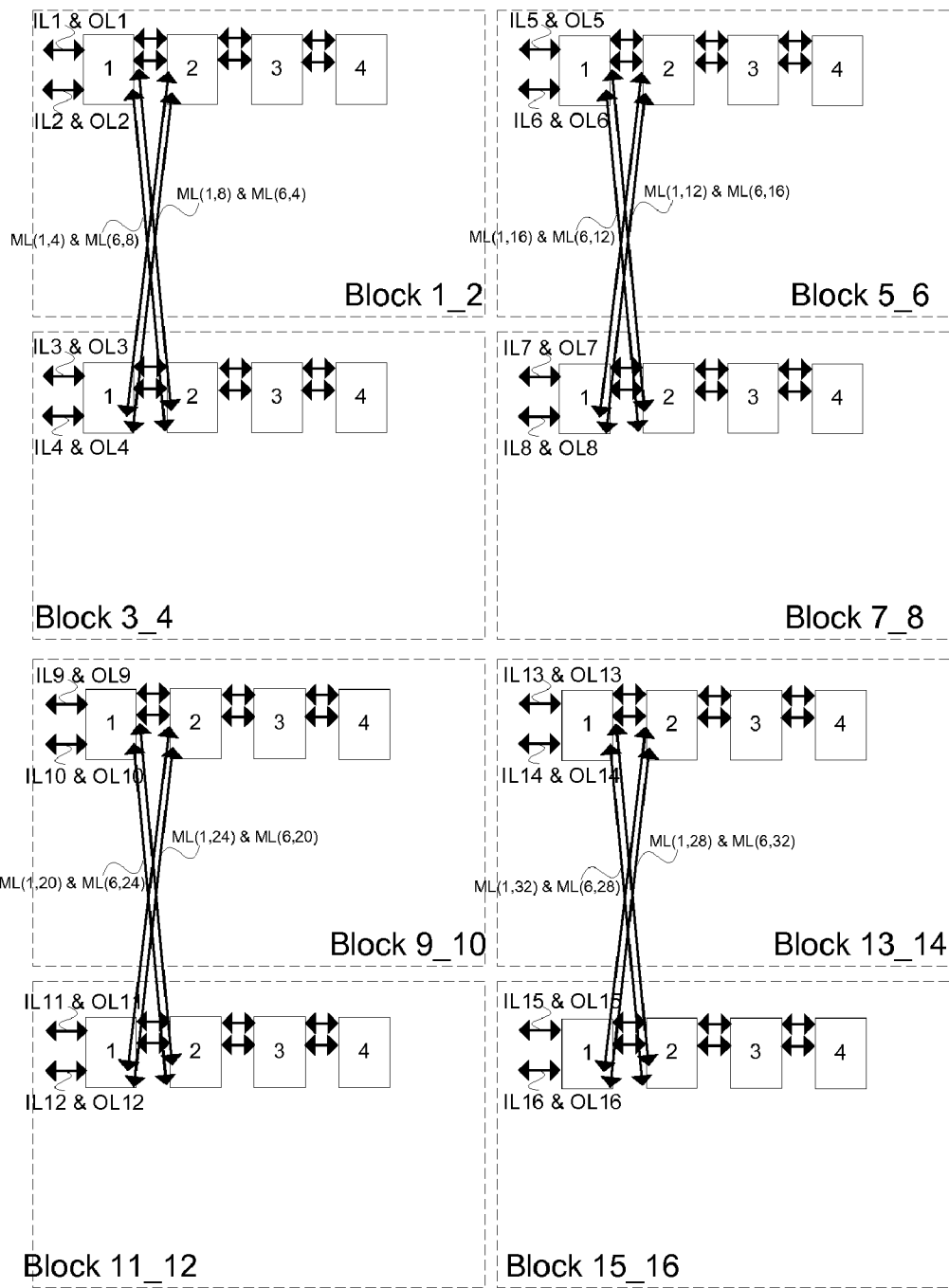


FIG. 2D5

200D5

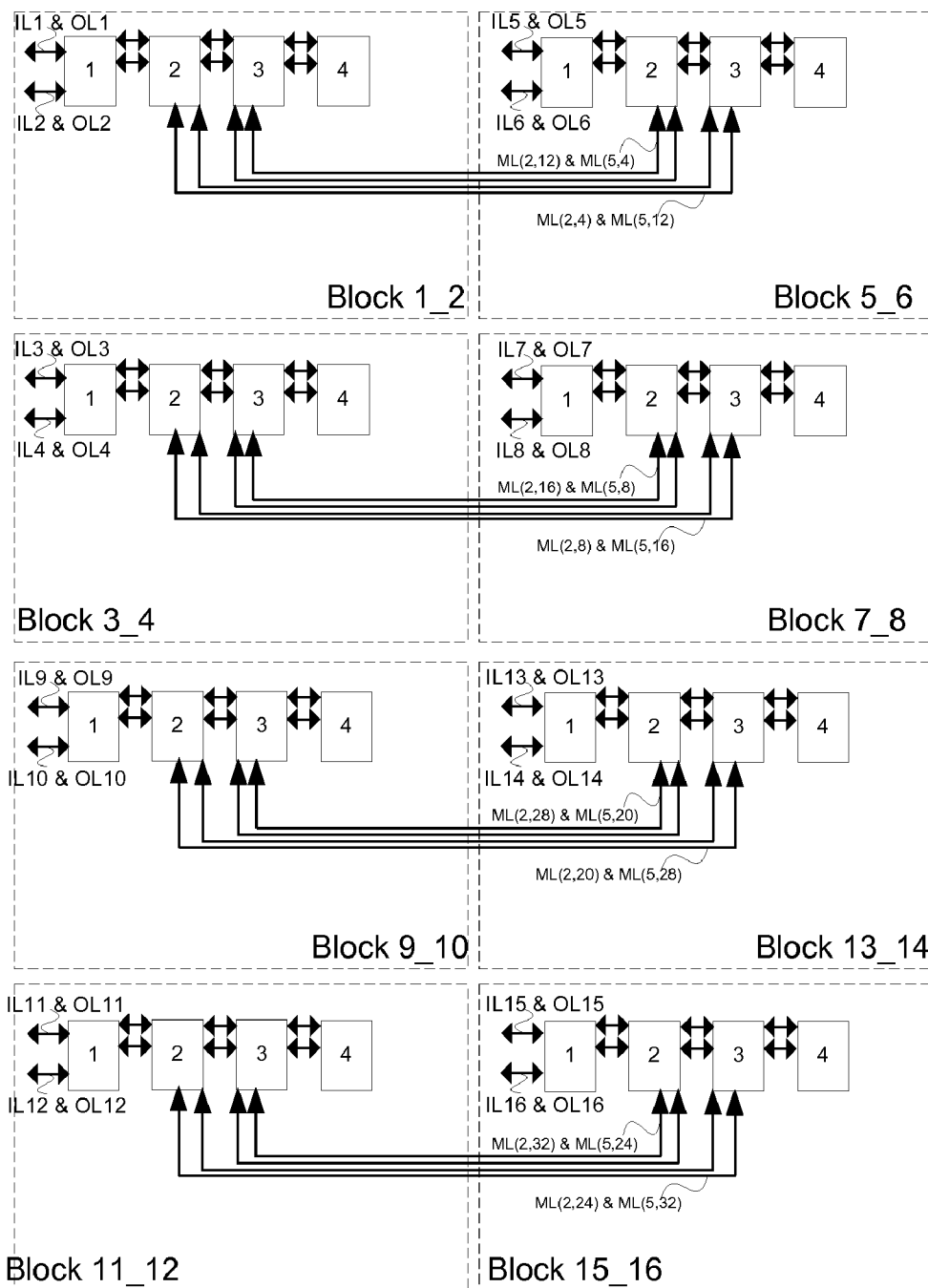


FIG. 2D6

200D6

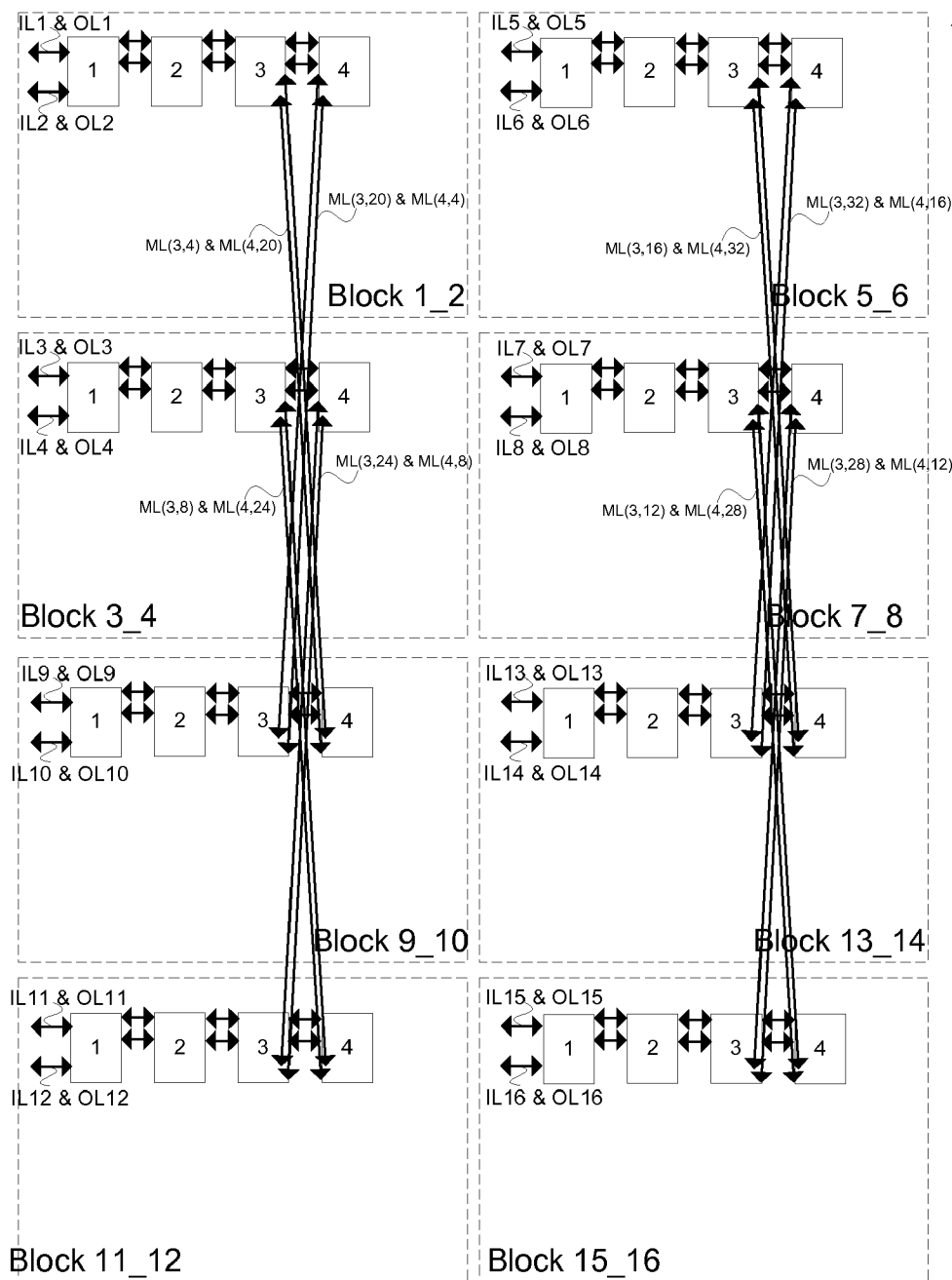


FIG. 3A

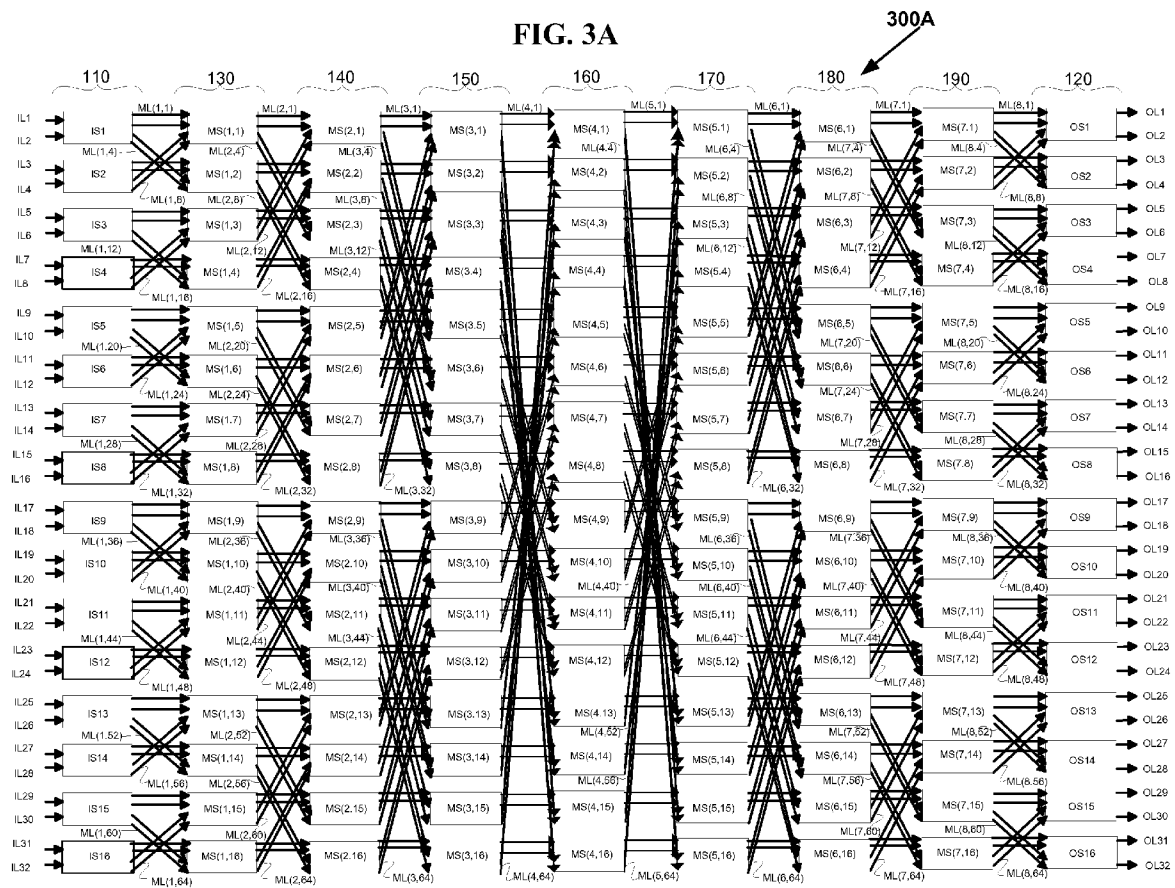
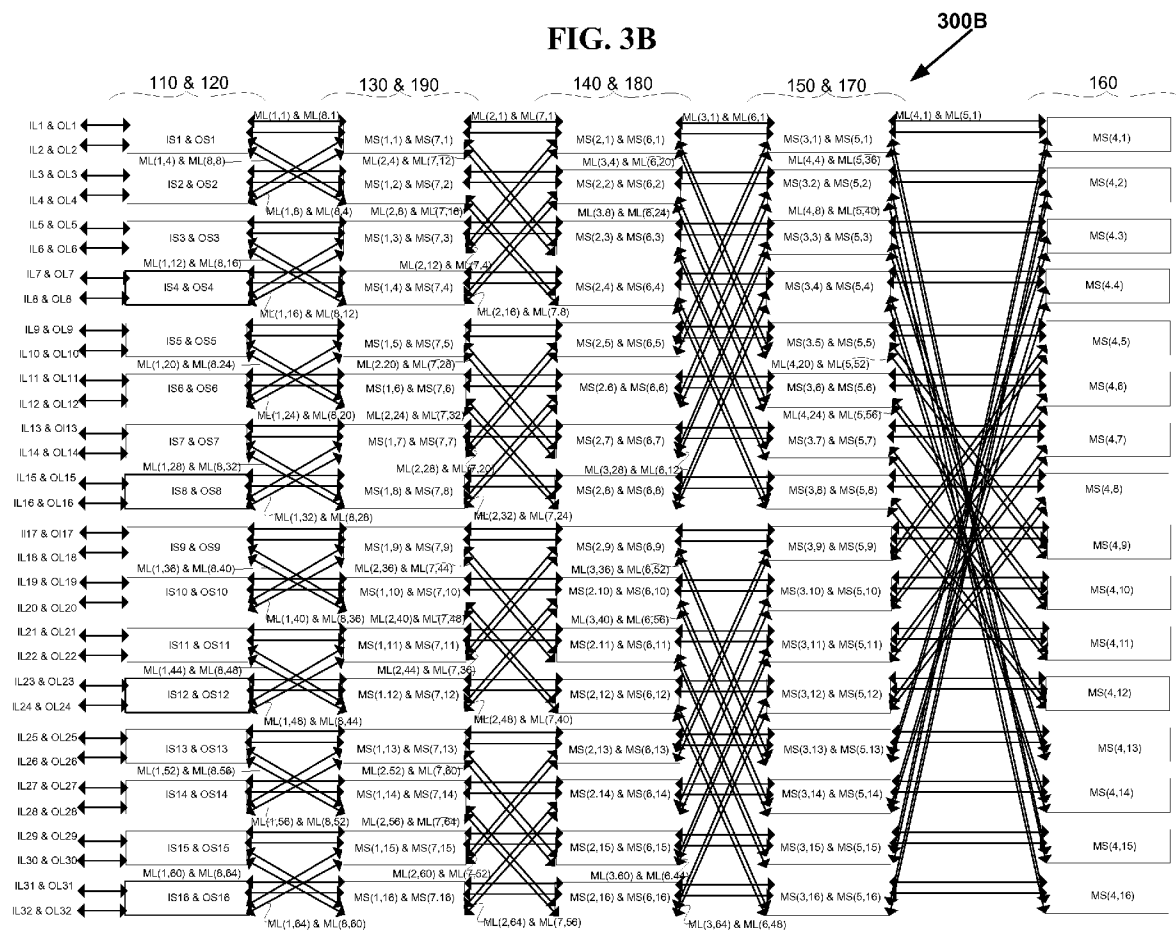
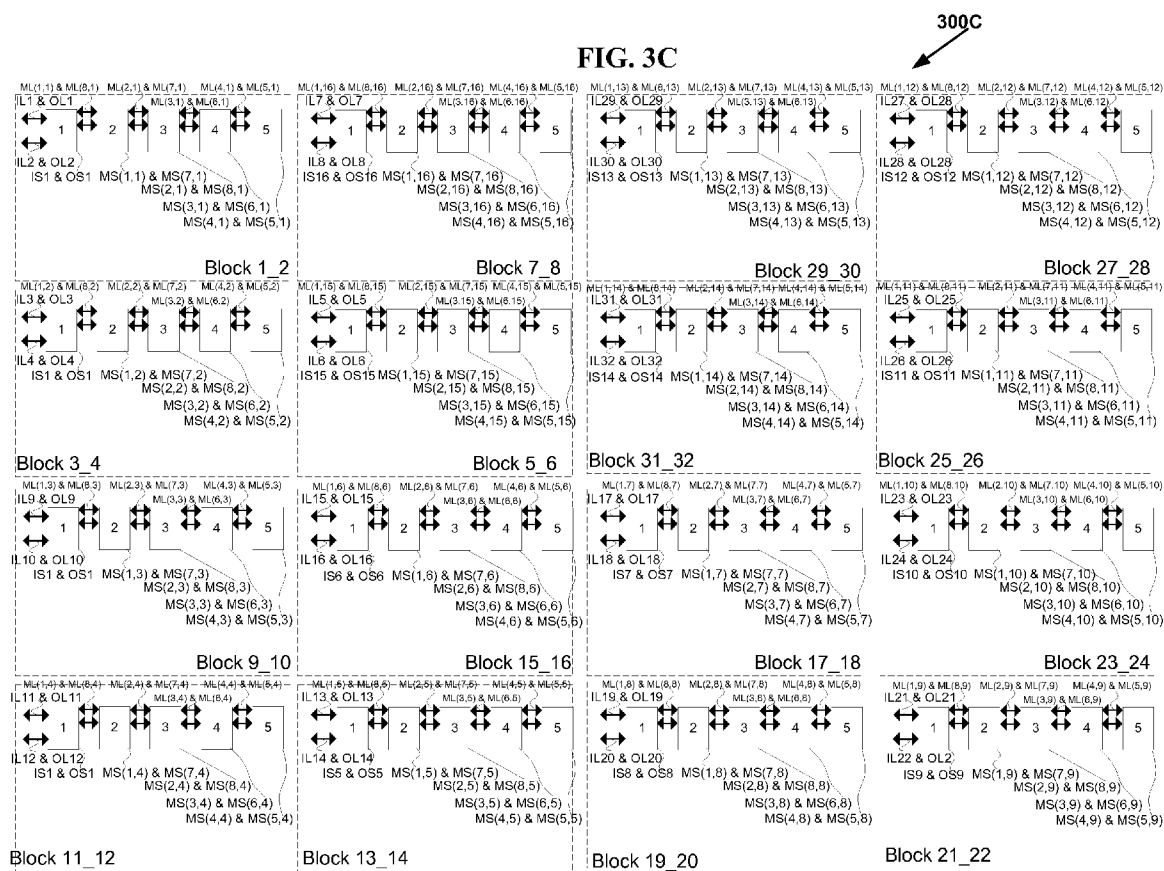
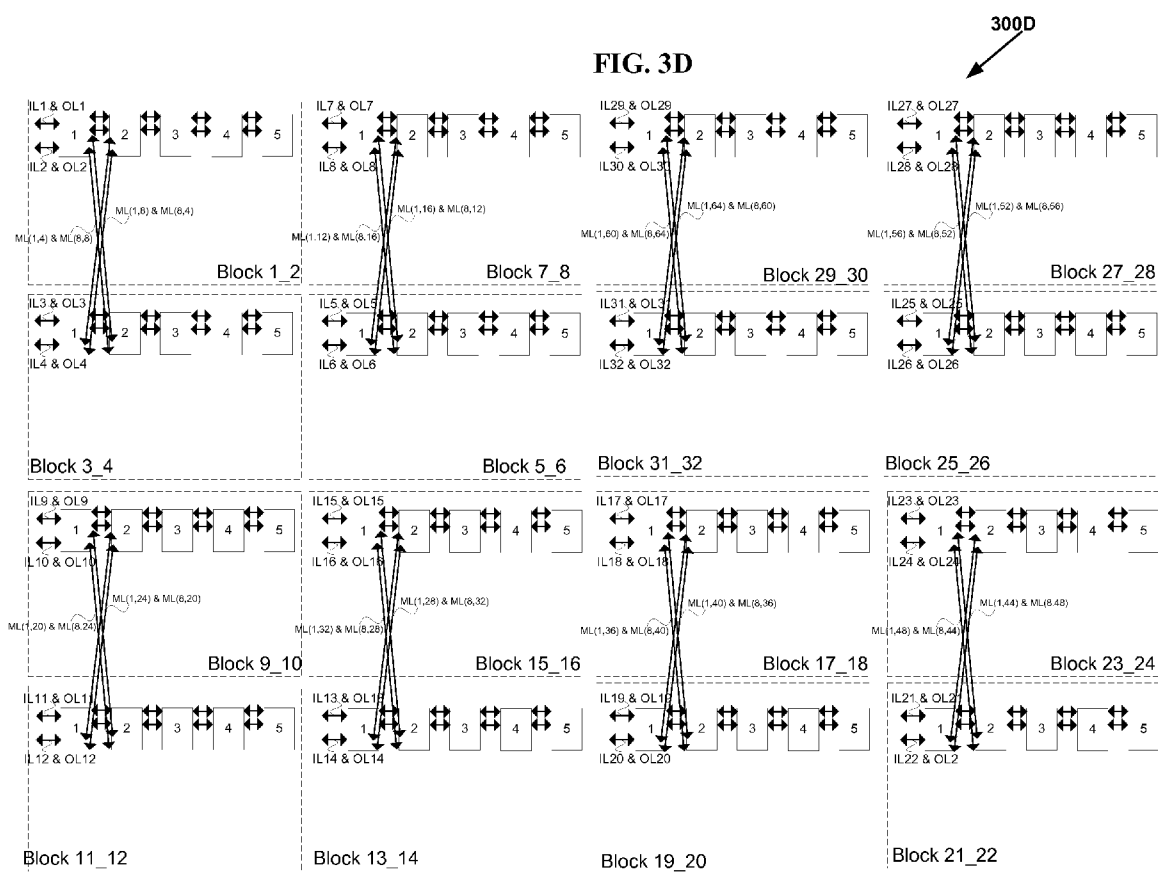


FIG. 3B







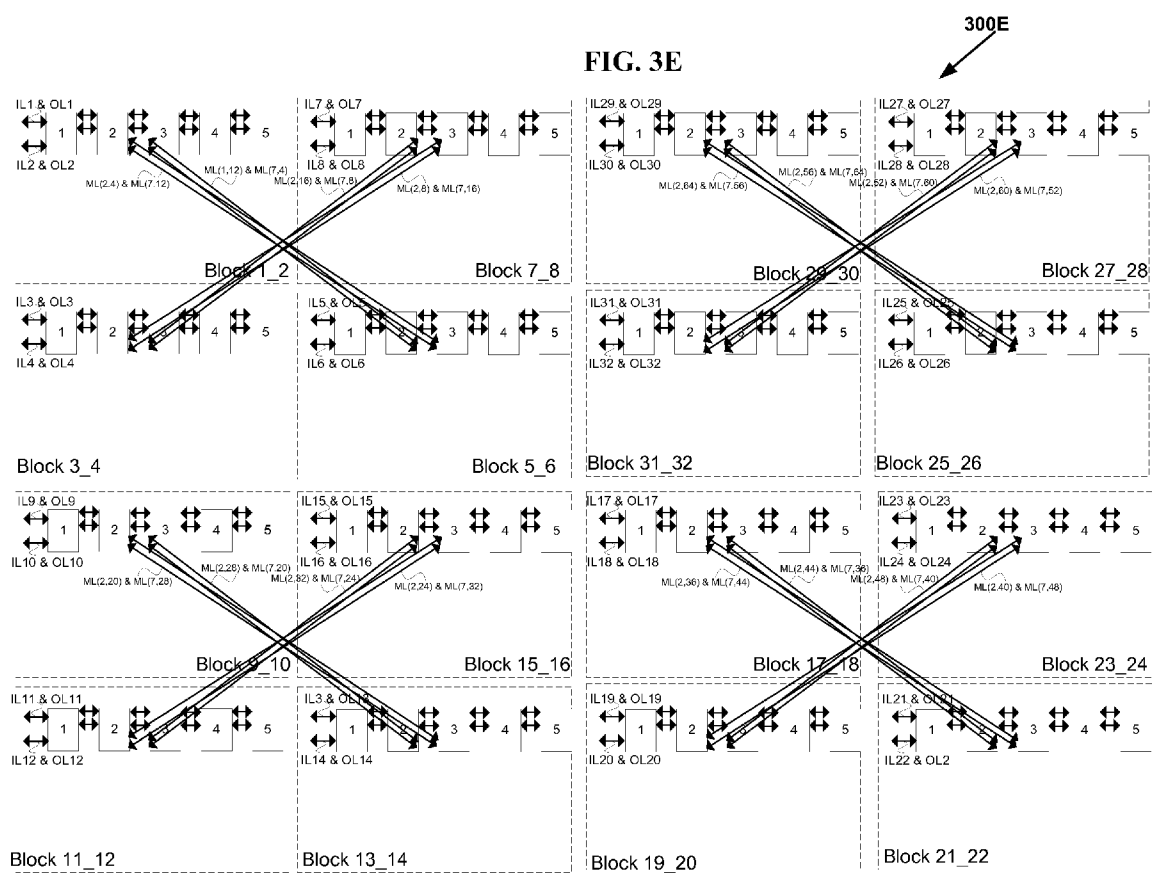


FIG. 3F

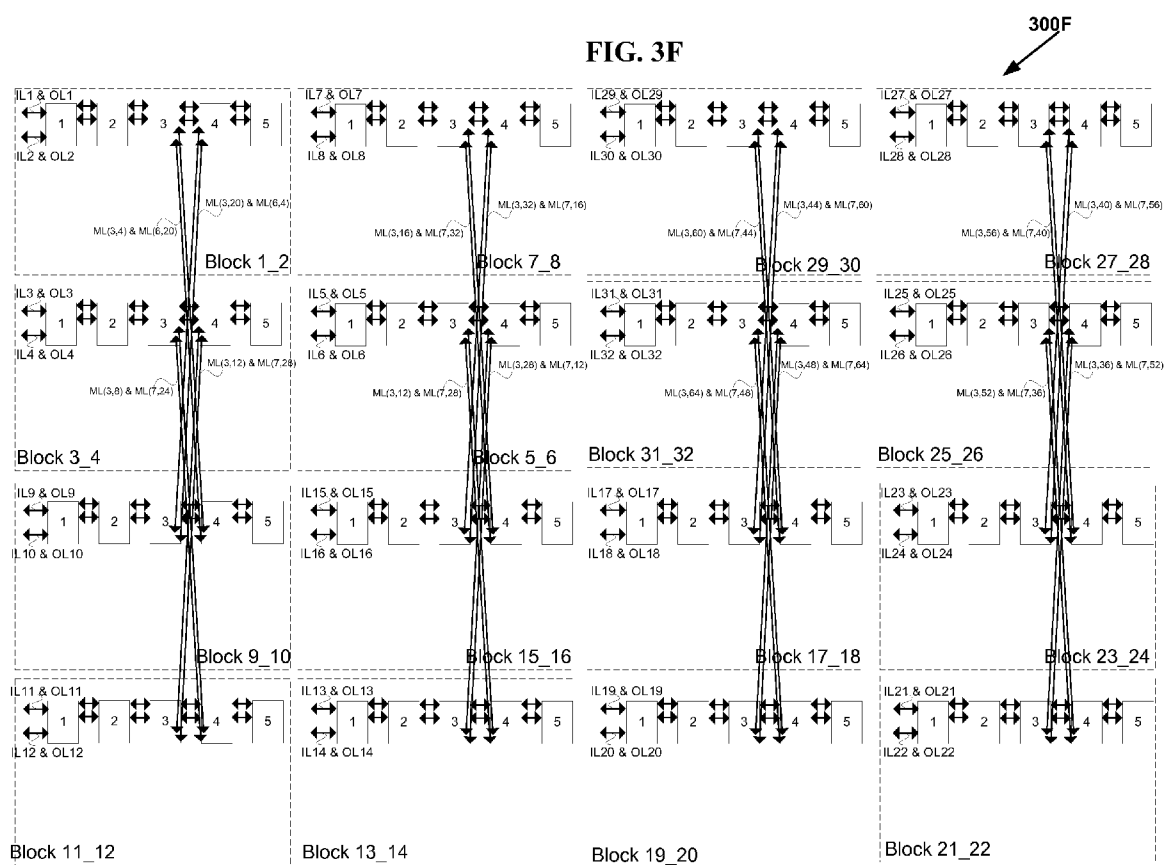


FIG. 3G

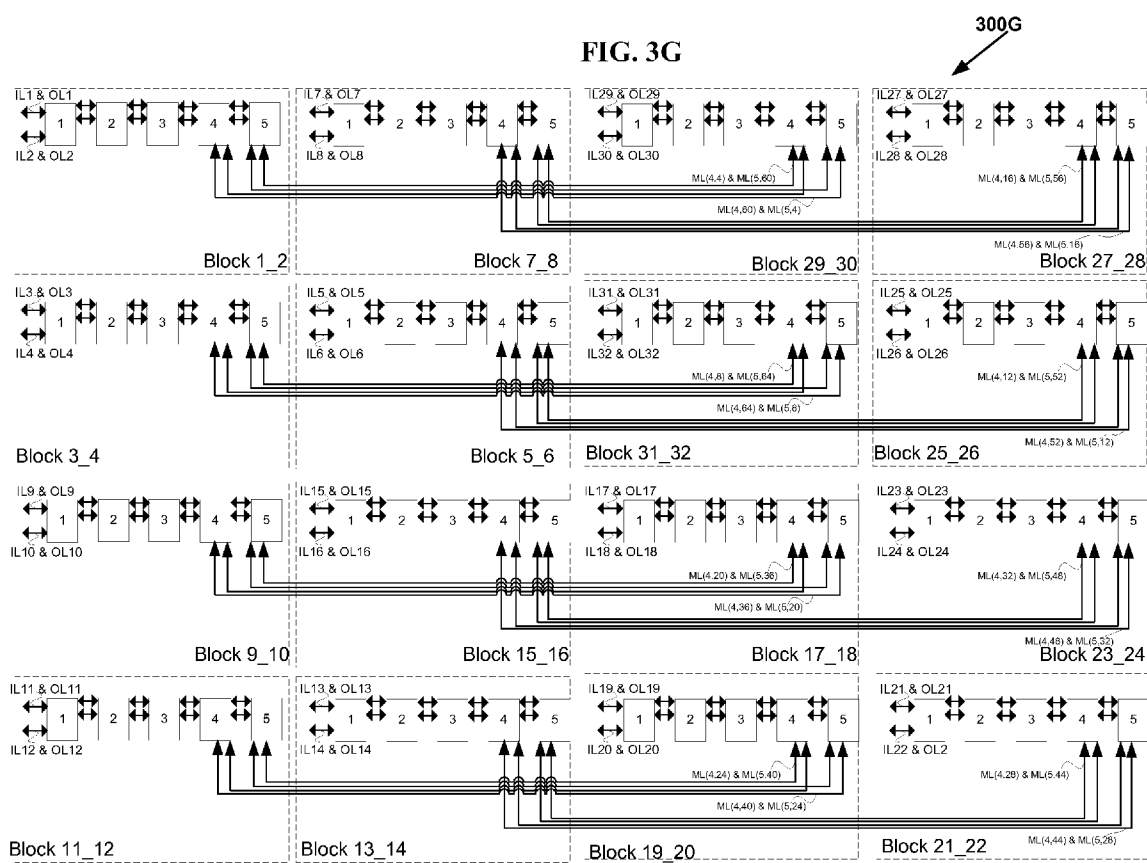


FIG. 3H

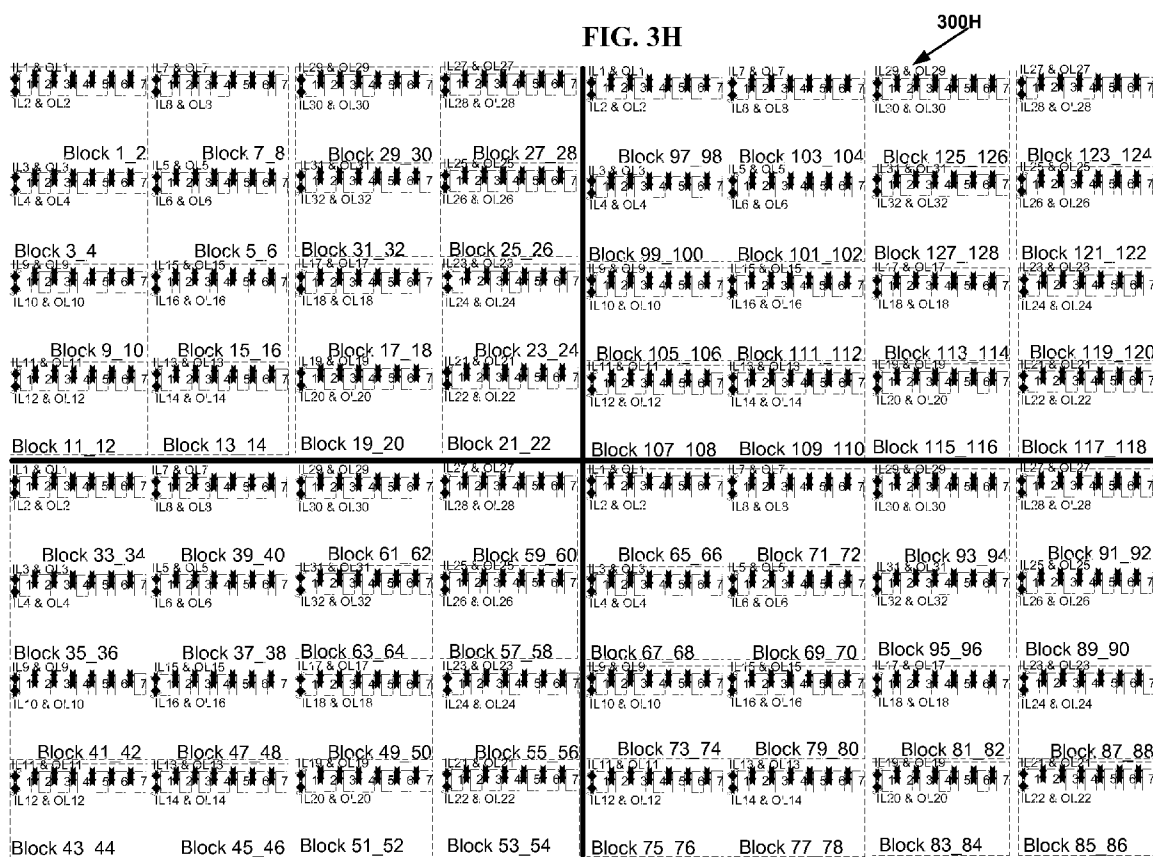


FIG. 4A

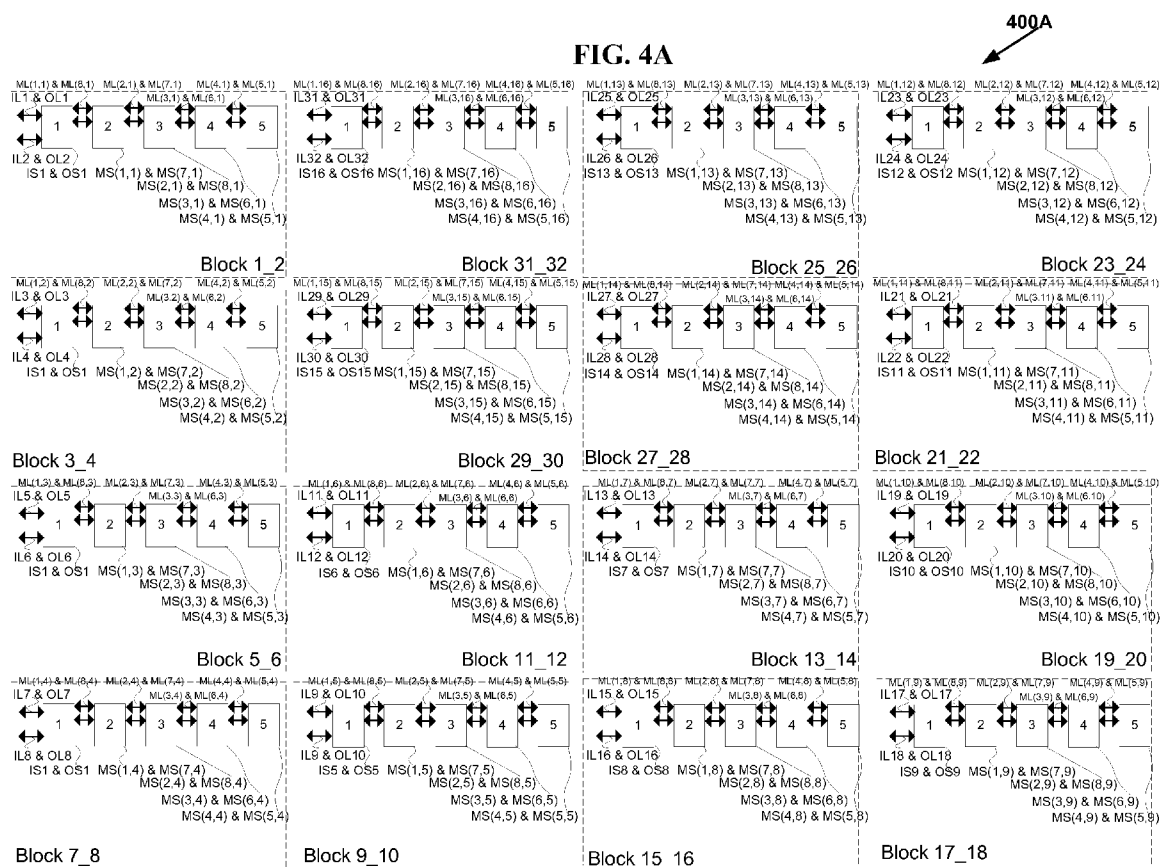


FIG. 4B

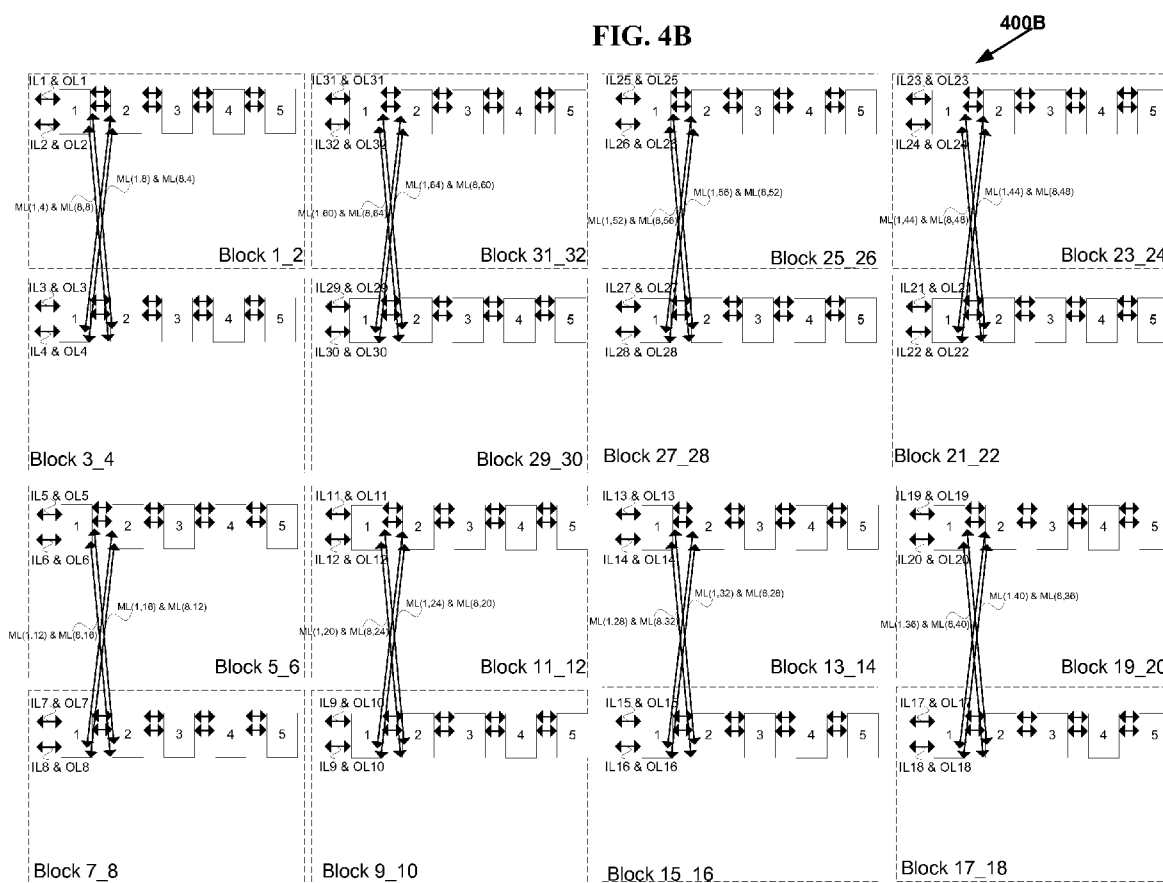


FIG. 4C

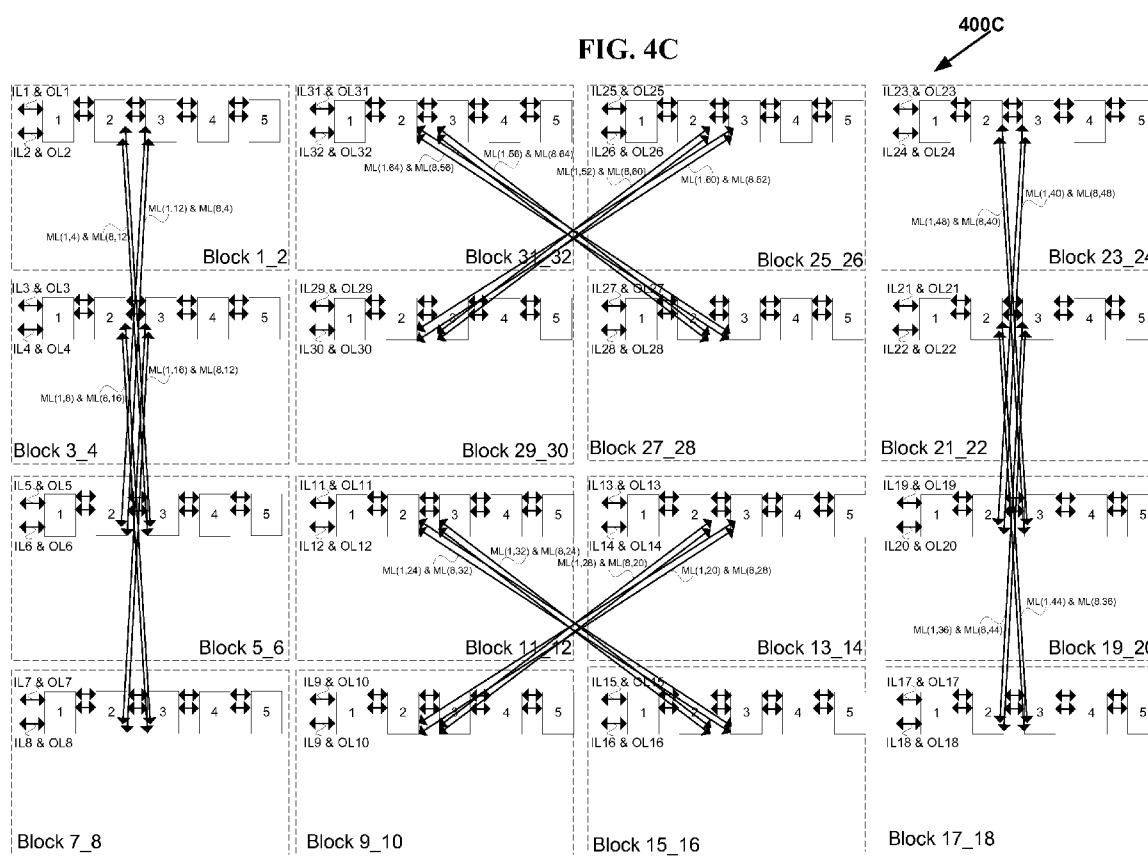
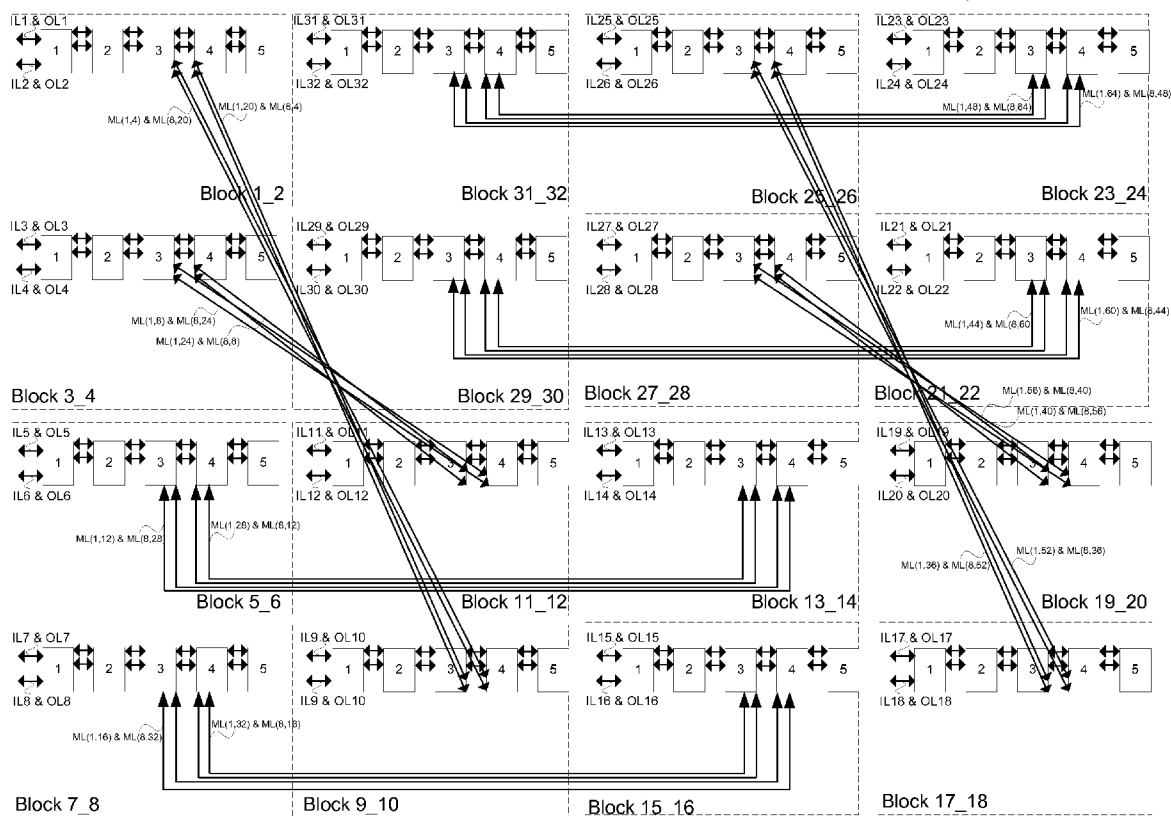


FIG. 4D

400D



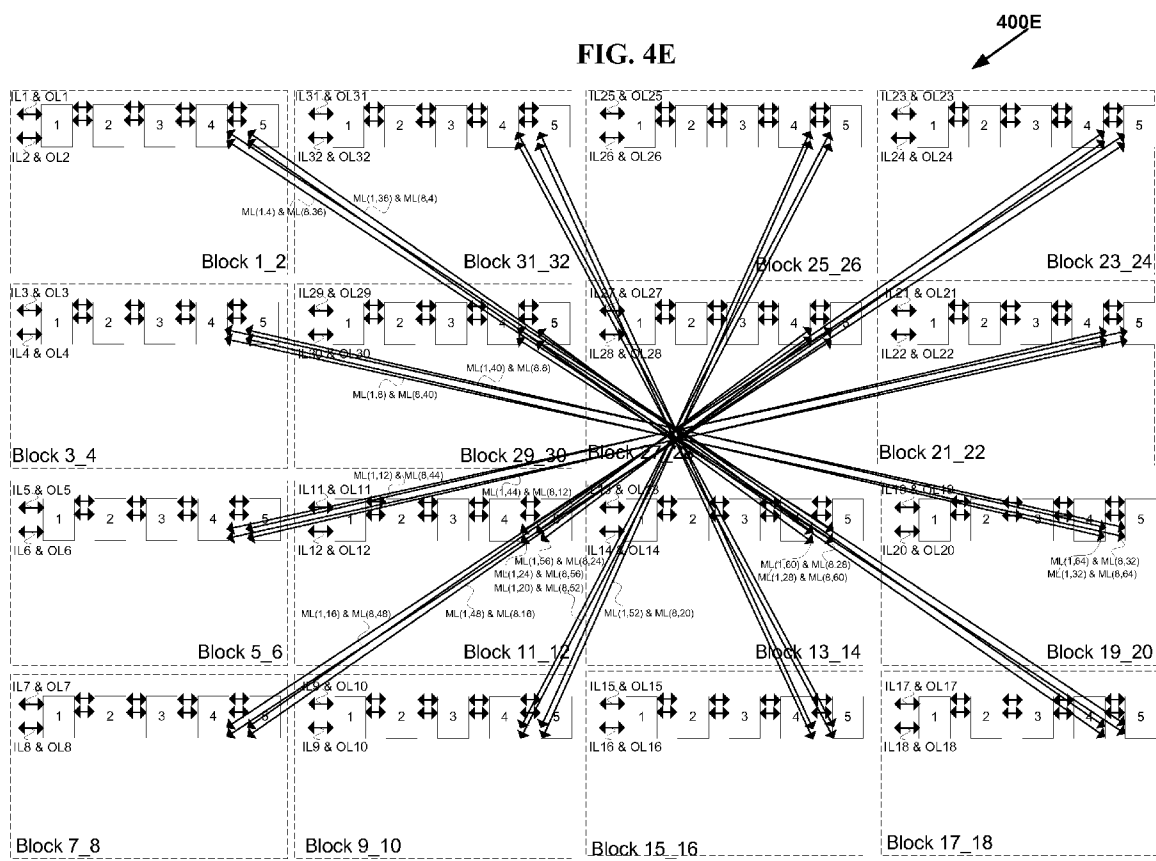
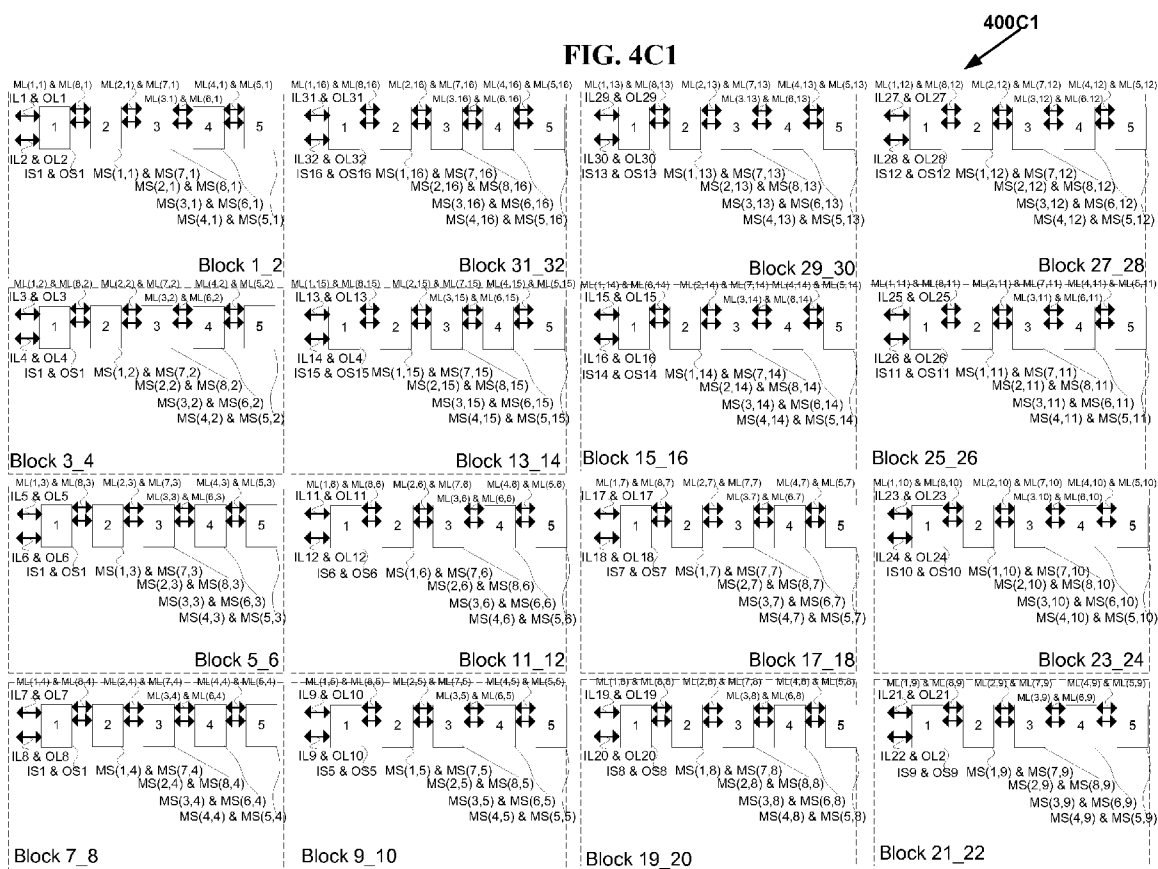
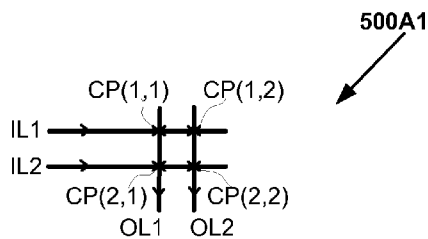


FIG. 4C1

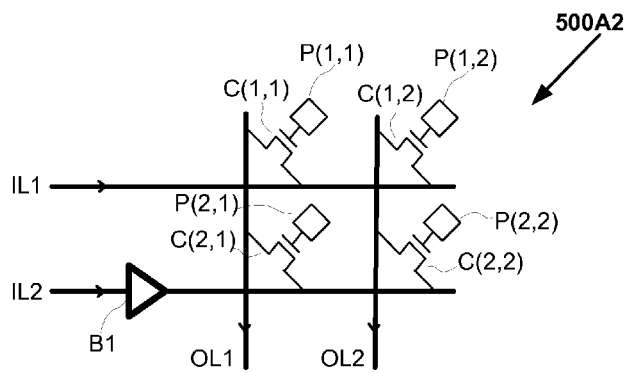


400C1

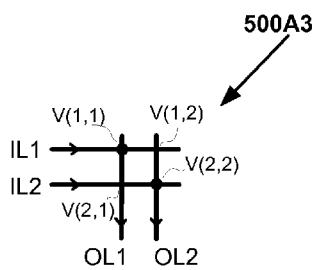
FIG. 5A



**FIG. 5A1
(Prior Art)**



**FIG. 5A2
(Prior Art)**



**FIG. 5A3
(Prior Art)**

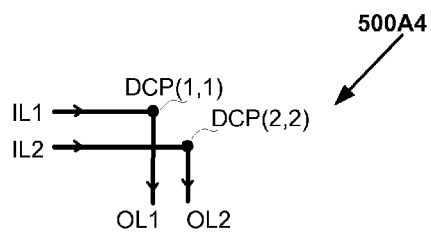


FIG. 5A4

US 2011/0037498 A1

Feb. 17, 2011

1

VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is related to and claims priority of the PCT Application Serial No. PCT/US08/64605 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 22, 2008, and the U.S. Provisional Patent Application Ser. No. 60/940,394 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0002] This application is related to and incorporates by reference in its entirety the U.S. application Ser. No. 12/530,207 entitled "FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Sep. 6, 2009, the PCT Application Serial No. PCT/US08/56064 entitled "FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Mar. 6, 2008, the U.S. Provisional Patent Application Ser. No. 60/905,526 entitled "LARGE SCALE CROSSPOINT REDUCTION WITH NONBLOCKING UNICAST & MULTICAST IN ARBITRARILY LARGE MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Mar. 6, 2007, and the U.S. Provisional Patent Application Ser. No. 60/940,383 entitled "FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0003] This application is related to and incorporates by reference in its entirety the US Patent Application Docket No. V-0038US entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application filed concurrently, the PCT Application Serial No. PCT/US08/64603 entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 22, 2008, the U.S. Provisional Patent Application Ser. No. 60/940,387 entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, and the U.S. Provisional Patent Application Ser. No. 60/940,390 entitled "FULLY CONNECTED GENERALIZED MULTI-LINK BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0004] This application is related to and incorporates by reference in its entirety the US Patent Application Docket No. V-0039US entitled "FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application filed concurrently, the PCT Application Serial No. PCT/US08/64604 entitled "FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 22, 2008, the U.S. Provisional Patent Application Ser. No. 60/940,389 entitled "FULLY CONNECTED GENERALIZED REARRANGE-

ABLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, the U.S. Provisional Patent Application Ser. No. 60/940,391 entitled "FULLY CONNECTED GENERALIZED FOLDED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007 and the U.S. Provisional Patent Application Ser. No. 60/940,392 entitled "FULLY CONNECTED GENERALIZED STRICTLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0005] This application is related to and incorporates by reference in its entirety the U.S. Provisional Patent Application Ser. No. 61/252,603 entitled "VLSI LAYOUTS OF FULLY CONNECTED NETWORKS WITH LOCALITY EXPLOITATION" by Venkat Konda assigned to the same assignee as the current application, filed Oct. 16, 2009.

[0006] This application is related to and incorporates by reference in its entirety the U.S. Provisional Patent Application Ser. No. 61/252,609 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Oct. 16, 2009.

BACKGROUND OF INVENTION

[0007] Multi-stage interconnection networks such as Benes networks and butterfly fat tree networks are widely useful in telecommunications, parallel and distributed computing. However VLSI layouts, known in the prior art, of these interconnection networks in an integrated circuit are inefficient and complicated.

[0008] Other multi-stage interconnection networks including butterfly fat tree networks, Banyan networks, Batcher-Banyan networks, Baseline networks, Delta networks, Omega networks and Flip networks have been widely studied particularly for self routing packet switching applications. Also Benes Networks with radix of two have been widely studied and it is known that Benes Networks of radix two are shown to be built with back to back baseline networks which are rearrangeably nonblocking for unicast connections.

[0009] The most commonly used VLSI layout in an integrated circuit is based on a two-dimensional grid model comprising only horizontal and vertical tracks. An intuitive interconnection network that utilizes two-dimensional grid model is 2D Mesh Network and its variations such as segmented mesh networks. Hence routing networks used in VLSI layouts are typically 2D mesh networks and its variations. However Mesh Networks require large scale cross points typically with a growth rate of $O(N^2)$ where N is the number of computing elements, ports, or logic elements depending on the application.

[0010] Multi-stage interconnection with a growth rate of $O(N \times \log N)$ requires significantly small number of cross points. U.S. Pat. No. 6,185,220 entitled "Grid Layouts of Switching and Sorting Networks" granted to Muthukrishnan et al. describes a VLSI layout using existing VLSI grid model for Benes and Butterfly networks. U.S. Pat. No. 6,940,308 entitled "Interconnection Network for a Field Programmable Gate Array" granted to Wong describes a VLSI layout where switches belonging to lower stage of Benes Network are layed out close to the logic cells and switches belonging to higher stages are layed out towards the center of the layout.

US 2011/0037498 A1

Feb. 17, 2011

2

[0011] Due to the inefficient and in some cases impractical VLSI layout of Benes and butterfly fat tree networks on a semiconductor chip, today mesh networks and segmented mesh networks are widely used in the practical applications such as field programmable gate arrays (FPGAs), programmable logic devices (PLDs), and parallel computing interconnects. The prior art VLSI layouts of Benes and butterfly fat tree networks and VLSI layouts of mesh networks and segmented mesh networks require large area to implement the switches on the chip, large number of wires, longer wires, with increased power consumption, increased latency of the signals which effect the maximum clock speed of operation. Some networks may not even be implemented practically on a chip due to the lack of efficient layouts.

SUMMARY OF INVENTION

[0012] When large scale sub-integrated circuit blocks with inlet and outlet links are layed out in an integrated circuit device in a two-dimensional grid arrangement, (for example in an FPGA where the sub-integrated circuit blocks are Lookup Tables) the most intuitive routing network is a network that uses horizontal and vertical links only (the most often used such a network is one of the variations of a 2D Mesh network). A direct embedding of a generalized multi-stage network on to a 2D Mesh network is neither simple nor efficient.

[0013] In accordance with the invention, VLSI layouts of generalized multi-stage networks for broadcast, unicast and multicast connections are presented using only horizontal and vertical links. The VLSI layouts employ shuffle exchange links where outlet links of cross links from switches in a stage in one sub-integrated circuit block are connected to inlet links of switches in the succeeding stage in another sub-integrated circuit block so that said cross links are either vertical links or horizontal and vice versa. In one embodiment the sub-integrated circuit blocks are arranged in a hypercube arrangement in a two-dimensional plane. The VLSI layouts exploit the benefits of significantly lower cross points, lower signal latency, lower power and full connectivity with significantly fast compilation.

[0014] The VLSI layouts presented are applicable to generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ for $s=1, 2, 3$ or any number in general. The embodiments of VLSI layouts are useful in wide target applications such as FPGAs, CPLDs, pSoCs, ASIC placement and route tools, networking applications, parallel & distributed computing, and reconfigurable computing.

BRIEF DESCRIPTION OF DRAWINGS

[0015] FIG. 1A is a diagram 100A of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having inverse Benes connection topology of nine stages with $N=32, d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0016] FIG. 1B is a diagram 100B of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network 100A shown in FIG. 1A, having inverse Benes connection topology of five stages with $N=32, d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0017] FIG. 1C is a diagram 100C layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links belonging with in each block only.

[0018] FIG. 1D is a diagram 100D layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(1,i) for $i=[1, 64]$ and ML(8,i) for $i=[1,64]$.

[0019] FIG. 1E is a diagram 100E layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(2,i) for $i=[1, 64]$ and ML(7,i) for $i=[1,64]$.

[0020] FIG. 1F is a diagram 100F layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(3,i) for $i=[1, 64]$ and ML(6,i) for $i=[1,64]$.

[0021] FIG. 1G is a diagram 100G layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(4,i) for $i=[1, 64]$ and ML(5,i) for $i=[1,64]$.

[0022] FIG. 1H is a diagram 100H layout of a network $V_{fold-mlink}(N, d, s)$ where $N=128, d=2$, and $s=2$, in one embodiment, illustrating the connection links belonging with in each block only.

[0023] FIG. 1I is a diagram 100I detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

[0024] FIG. 1J is a diagram 100J detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

[0025] FIG. 1K is a diagram 100K detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

[0026] FIG. 1K1 is a diagram 100M1 detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$ for $s=1$.

[0027] FIG. 1L is a diagram 100L detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

[0028] FIG. 1L1 is a diagram 100L1 detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$ for $s=1$.

[0029] FIG. 2A1 is a diagram 200A1 of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having inverse Benes connection topology of one stage with

US 2011/0037498 A1

Feb. 17, 2011

3

$N=2$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2A2 is a diagram 200A2 of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network 200A1 shown in FIG. 2A1, having inverse Benes connection topology of one stage with $N=2$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2A3 is a diagram 200A3 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2A2, in one embodiment, illustrating all the connection links.

[0030] FIG. 2B1 is a diagram 200B1 of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having inverse Benes connection topology of one stage with $N=4$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2B2 is a diagram 200B2 of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network 200B1 shown in FIG. 2B1, having inverse Benes connection topology of one stage with $N=4$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2B3 is a diagram 200B3 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2B2, in one embodiment, illustrating the connection links belonging with in each block only. FIG. 2B4 is a diagram 200B4 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2B2, in one embodiment, illustrating the connection links $ML(1,i)$ for $i=[1, 8]$ and $ML(2,i)$ for $i=[1,8]$.

[0031] FIG. 2C11 is a diagram 200C11 of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having inverse Benes connection topology of one stage with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention. FIG. 2C12 is a diagram 200C12 of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network 200C11 shown in FIG. 2C11, having inverse Benes connection topology of one stage with $N=8$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0032] FIG. 2C21 is a diagram 200C21 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2C12, in one embodiment, illustrating the connection links belonging with in each block only. FIG. 2C22 is a diagram 200C22 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2C12, in one embodiment, illustrating the connection links $ML(1,i)$ for $i=[1, 16]$ and $ML(4,i)$ for $i=[1,16]$. FIG. 2C23 is a diagram 200C23 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2C12, in one embodiment, illustrating the connection links $ML(2,i)$ for $i=[1, 16]$ and $ML(3,i)$ for $i=[1,16]$.

[0033] FIG. 2D1 is a diagram 200D1 of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having inverse Benes connection topology of one stage with $N=16$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0034] FIG. 2D2 is a diagram 200D2 of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network 200D1 shown in FIG. 2D1, having inverse Benes connection topology of one stage with $N=16$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0035] FIG. 2D3 is a diagram 200D3 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2D2, in one embodiment, illustrating the connection links belonging with in each block only.

[0036] FIG. 2D4 is a diagram 200D4 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2D2, in one embodiment, illustrating the connection links $ML(1,i)$ for $i=[1, 32]$ and $ML(6,i)$ for $i=[1,32]$.

[0037] FIG. 2D5 is a diagram 200D5 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2D2, in one embodiment, illustrating the connection links $ML(2,i)$ for $i=[1, 32]$ and $ML(5,i)$ for $i=[1,32]$.

[0038] FIG. 2D6 is a diagram 200D6 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2D2, in one embodiment, illustrating the connection links $ML(3,i)$ for $i=[1, 32]$ and $ML(4,i)$ for $i=[1,32]$.

[0039] FIG. 3A is a diagram 300A of an exemplary symmetrical multi-link multi-stage network $V_{hcube}(N, d, s)$ having inverse Benes connection topology of nine stages with $N=32$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0040] FIG. 3B is a diagram 300B of the equivalent symmetrical folded multi-link multi-stage network $V_{hcube}(N, d, s)$ of the network 300A shown in FIG. 3A, having inverse Benes connection topology of five stages with $N=32$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0041] FIG. 3C is a diagram 300C layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links belonging with in each block only.

[0042] FIG. 3D is a diagram 100D layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links $ML(1,i)$ for $i=[1, 64]$ and $ML(8,i)$ for $i=[1,64]$.

[0043] FIG. 3E is a diagram 300E layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links $ML(2,i)$ for $i=[1, 64]$ and $ML(7,i)$ for $i=[1,64]$.

[0044] FIG. 3F is a diagram 300F layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links $ML(3,i)$ for $i=[1, 64]$ and $ML(6,i)$ for $i=[1,64]$.

[0045] FIG. 3G is a diagram 300G layout of the network $V_{hcube}(N, d, s)$ shown in FIG. 3B, in one embodiment, illustrating the connection links $ML(4,i)$ for $i=[1, 64]$ and $ML(5,i)$ for $i=[1,64]$.

[0046] FIG. 3H is a diagram 300H layout of a network $V_{hcube}(N, d, s)$ where $N=128$, $d=2$, and $s=2$, in one embodiment, illustrating the connection links belonging with in each block only.

US 2011/0037498 A1

Feb. 17, 2011

4

[0047] FIG. 4A is a diagram 400A layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links belonging with in each block only.

[0048] FIG. 4B is a diagram 400B layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links $ML(1,i)$ for $i=[1, 64]$ and $ML(8,i)$ for $i=[1,64]$.

[0049] FIG. 4C is a diagram 400C layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 4C, in one embodiment, illustrating the connection links $ML(2,i)$ for $i=[1, 64]$ and $ML(7,i)$ for $i=[1,64]$.

[0050] FIG. 4D is a diagram 400D layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 4D, in one embodiment, illustrating the connection links $ML(3,i)$ for $i=[1, 64]$ and $ML(6,i)$ for $i=[1,64]$.

[0051] FIG. 4E is a diagram 400E layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 4E, in one embodiment, illustrating the connection links $ML(4,i)$ for $i=[1, 64]$ and $ML(5,i)$ for $i=[1,64]$.

[0052] FIG. 4C1 is a diagram 400C1 layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links belonging with in each block only.

[0053] FIG. 5A1 is a diagram 500A1 of an exemplary prior art implementation of a two by two switch; FIG. 5A2 is a diagram 500A2 for programmable integrated circuit prior art implementation of the diagram 500A1 of FIG. 5A1; FIG. 5A3 is a diagram 500A3 for one-time programmable integrated circuit prior art implementation of the diagram 500A1 of FIG. 5A1; FIG. 5A4 is a diagram 500A4 for integrated circuit placement and route implementation of the diagram 500A1 of FIG. 5A1.

DETAILED DESCRIPTION OF THE INVENTION

[0054] The present invention is concerned with the VLSI layouts of arbitrarily large switching networks for broadcast, unicast and multicast connections. Particularly switching networks considered in the current invention include: generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ for $s=1, 2, 3$ or any number in general.

[0055] Efficient VLSI layout of networks on a semiconductor chip are very important and greatly influence many important design parameters such as the area taken up by the network on the chip, total number of wires, length of the wires, latency of the signals, capacitance and hence the maximum clock speed of operation. Some networks may not even be implemented practically on a chip due to the lack of efficient layouts. The different varieties of multi-stage networks described above have not been implemented previously on the semiconductor chips efficiently. For example in Field Programmable Gate Array (FPGA) designs, multi-stage networks described in the current invention have not been successfully implemented primarily due to the lack of efficient VLSI layouts. Current commercial FPGA products such as Xilinx Vertex, Altera's Stratix implement island-style architecture using mesh and segmented mesh routing interconnects using either full crossbars or sparse crossbars. These

routing interconnects consume large silicon area for cross-points, long wires, large signal propagation delay and hence consume lot of power.

[0056] The current invention discloses the VLSI layouts of numerous types of multi-stage networks which are very efficient. Moreover they can be embedded on to mesh and segmented mesh routing interconnects of current commercial FPGA products. The VLSI layouts disclosed in the current invention are applicable to including the numerous generalized multi-stage networks disclosed in the following patent applications, filed concurrently:

[0057] 1) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-stage networks $V(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the PCT Application Serial No. PCT/US08/56064 that is incorporated by reference above.

[0058] 2) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above.

[0059] 3) Rearrangeably nonblocking for arbitrary fan-out multicast and unicast, and strictly nonblocking for unicast for generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above.

[0060] 4) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above.

[0061] 5) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above.

[0062] 6) Strictly nonblocking for arbitrary fan-out multicast for generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above.

[0063] 7) VLSI layouts of numerous types of multi-stage networks with locality exploitation are described in U.S. Provisional Patent Application Ser. No. 61/252,603 that is incorporated by reference above.

[0064] 8) VLSI layouts of numerous types of multistage pyramid networks are described in U.S. Provisional Patent Application Ser. No. 61/252,609 that is incorporated by reference above.

[0065] In addition the layouts of the current invention are also applicable to generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$, generalized

US 2011/0037498 A1

Feb. 17, 2011

5

folded multi-link multi-stage pyramid networks $V_{fold-mlink-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ for $s=1, 2, 3$ or any number in general.

Symmetric RNB Generalized Multi-Link Multi-Stage Network $V_{mlink}(N_1, N_2, d, s)$:

[0066] Referring to diagram 100A in FIG. 1A, in one embodiment, an exemplary generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages of one hundred and forty four switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, 150, 160, 170, 180 and 190 is shown where input stage 110 consists of sixteen, two by four switches IS1-IS16 and output stage 120 consists of sixteen, four by two switches OS1-OS16. And all the middle stages namely the middle stage 130 consists of sixteen, four by four switches MS(1,1)-MS(1,16), middle stage 140 consists of sixteen, four by four switches MS(2,1)-MS(2,16), middle stage 150 consists of sixteen, four by four switches MS(3,1)-MS(3,16), middle stage 160 consists of sixteen, four by four switches MS(4,1)-MS(4,16), middle stage 170 consists of sixteen, four by four switches MS(5,1)-MS(5,16), middle stage 180 consists of sixteen, four by four switches MS(6,1)-MS(6,16), and middle stage 190 consists of sixteen, four by four switches MS(7,1)-MS(7,16).

[0067] As disclosed in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above, such a network can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections.

[0068] In one embodiment of this network each of the input switches IS1-IS4 and output switches OS1-OS4 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch IS1-IS4 can be denoted in general with the notation $d*2d$ and each output switch OS1-OS4 can be denoted in general with the notation $2d*d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d*2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-stage network can be represented with the notation $V_{mlink}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL32), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0069] Each of the N/d input switches IS1-IS16 are connected to exactly d switches in middle stage 130 through two links each for a total of $2*d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the links ML(1,1), ML(1,2), and also connected to middle switch MS(1,2) through the links ML(1,3) and ML(1,4)). The middle links which connect switches in the same row in two successive middle stages are called hereinafter straight middle links; and the middle links which connect switches in

different rows in two successive middle stages are called hereinafter cross middle links. For example, the middle links ML(1,1) and ML(1,2) connect input switch IS1 and middle switch MS(1,1), so middle links ML(1,1) and ML(1,2) are straight middle links; whereas the middle links ML(1,3) and ML(1,4) connect input switch IS1 and middle switch MS(1,2), since input switch IS1 and middle switch MS(1,2) belong to two different rows in diagram 100A of FIG. 1A, middle links ML(1,3) and ML(1,4) are cross middle links

[0070] Each of the N/d middle switches MS(1,1)-MS(1,16) in the middle stage 130 are connected from exactly d input switches through two links each for a total of $2*d$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to exactly d switches in middle stage 140 through two links each for a total of $2*d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0071] Each of the N/d middle switches MS(2,1)-MS(2,16) in the middle stage 140 are connected from exactly d input switches through two links each for a total of $2*d$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from input switch MS(1,1), and the links ML(1,11) and ML(1,12) are connected to the middle switch MS(2,1) from input switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through two links each for a total of $2*d$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,5)).

[0072] Each of the N/d middle switches MS(3,1)-MS(3,16) in the middle stage 150 are connected from exactly d input switches through two links each for a total of $2*d$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from input switch MS(2,1), and the links ML(2,19) and ML(2,20) are connected to the middle switch MS(3,1) from input switch MS(2,5)) and also are connected to exactly d switches in middle stage 160 through two links each for a total of $2*d$ links (for example the links ML(4,1) and ML(4,2) are connected from middle switch MS(3,1) to middle switch MS(4,1), and the links ML(4,3) and ML(4,4) are connected from middle switch MS(3,1) to middle switch MS(4,9)).

[0073] Each of the N/d middle switches MS(4,1)-MS(4,16) in the middle stage 160 are connected from exactly d input switches through two links each for a total of $2*d$ links (for example the links ML(4,1) and ML(4,2) are connected to the middle switch MS(4,1) from input switch MS(3,1), and the links ML(4,35) and ML(4,36) are connected to the middle switch MS(4,1) from input switch MS(3,9)) and also are connected to exactly d switches in middle stage 170 through two links each for a total of $2*d$ links (for example the links ML(5,1) and ML(5,2) are connected from middle switch MS(4,1) to middle switch MS(5,1), and the links ML(5,3) and ML(5,4) are connected from middle switch MS(4,1) to middle switch MS(5,9)).

[0074] Each of the N/d middle switches MS(5,1)-MS(5,16) in the middle stage 170 are connected from exactly d input switches through two links each for a total of $2*d$ links (for

US 2011/0037498 A1

Feb. 17, 2011

6

example the links ML(5,1) and ML(5,2) are connected to the middle switch MS(5,1) from input switch MS(4,1), and the links ML(5,35) and ML(5,36) are connected to the middle switch MS(5,1) from input switch MS(4,9) and also are connected to exactly d switches in middle stage **180** through two links each for a total of $2 \times d$ links (for example the links ML(6,1) and ML(6,2) are connected from middle switch MS(5,1) to middle switch MS(6,1), and the links ML(6,3) and ML(6,4) are connected from middle switch MS(5,1) to middle switch MS(6,5)).

[0075] Each of the N/d middle switches MS(6,1)-MS(6,16) in the middle stage **180** are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(6,1) and ML(6,2) are connected to the middle switch MS(6,1) from input switch MS(5,1), and the links ML(6,19) and ML(6,20) are connected to the middle switch MS(6,1) from input switch MS(5,5) and also are connected to exactly d switches in middle stage **190** through two links each for a total of $2 \times d$ links (for example the links ML(7,1) and ML(7,2) are connected from middle switch MS(6,1) to middle switch MS(7,1), and the links ML(7,3) and ML(7,4) are connected from middle switch MS(6,1) to middle switch MS(7,3)).

[0076] Each of the N/d middle switches MS(7,1)-MS(7,16) in the middle stage **190** are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(7,1) and ML(7,2) are connected to the middle switch MS(7,1) from input switch MS(6,1), and the links ML(7,11) and ML(7,12) are connected to the middle switch MS(7,1) from input switch MS(6,3) and also are connected to exactly d switches in middle stage **120** through two links each for a total of $2 \times d$ links (for example the links ML(8,1) and ML(8,2) are connected from middle switch MS(7,1) to middle switch MS(8,1), and the links ML(8,3) and ML(8,4) are connected from middle switch MS(7,1) to middle switch OS2).

[0077] Each of the N/d middle switches OS1-OS16 in the middle stage **120** are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(8,1) and ML(8,2) are connected to the output switch OS1 from input switch MS(7,1), and the links ML(8,7) and ML(7,8) are connected to the output switch OS1 from input switch MS(7,2)).

[0078] Finally the connection topology of the network **100A** shown in FIG. 1A is known to be back to back inverse Benes connection topology.

[0079] Referring to diagram **100B** in FIG. 1B, is a folded version of the multi-link multi-stage network **100A** shown in FIG. 1A. The network **100B** in FIG. 1B shows input stage **110** and output stage **120** are placed together. That is input switch IS1 and output switch OS1 are placed together, input switch IS2 and output switch OS2 are placed together, and similarly input switch IS16 and output switch OS16 are placed together. All the right going middle links (hereinafter "forward connecting links") {i.e., inlet links IL1-IL32 and middle links ML(1,1)-ML(1,64)} correspond to input switches IS1-IS16, and all the left going middle links (hereinafter "backward connecting links") {i.e., middle links ML(8,1)-ML(8,64) and outlet links OL1-OL32} correspond to output switches OS1-OS16.

[0080] Middle stage **130** and middle stage **190** are placed together. That is middle switches MS(1,1) and MS(7,1) are placed together, middle switches MS(1,2) and MS(7,2) are placed together, and similarly middle switches MS(1,16) and

MS(7,16) are placed together. All the right going middle links {i.e., middle links ML(1,1)-ML(1,64) and middle links ML(2,1)-ML(2,64)} correspond to middle switches MS(1,1)-MS(1,16), and all the left going middle links {i.e., middle links ML(7,1)-ML(7,64) and middle links ML(8,1) and ML(8,64)} correspond to middle switches MS(7,1)-MS(7,16).

[0081] Middle stage **140** and middle stage **180** are placed together. That is middle switches MS(2,1) and MS(6,1) are placed together, middle switches MS(2,2) and MS(6,2) are placed together, and similarly middle switches MS(2,16) and MS(6,16) are placed together. All the right going middle links {i.e., middle links ML(2,1)-ML(2,64) and middle links ML(3,1)-ML(3,64)} correspond to middle switches MS(2,1)-MS(2,16), and all the left going middle links {i.e., middle links ML(6,1)-ML(6,64) and middle links ML(7,1) and ML(7,64)} correspond to middle switches MS(6,1)-MS(6,16).

[0082] Middle stage **150** and middle stage **170** are placed together. That is middle switches MS(3,1) and MS(5,1) are placed together, middle switches MS(3,2) and MS(5,2) are placed together, and similarly middle switches MS(3,16) and MS(5,16) are placed together. All the right going middle links {i.e., middle links ML(3,1)-ML(3,64) and middle links ML(4,1)-ML(4,64)} correspond to middle switches MS(3,1)-MS(3,16), and all the left going middle links {i.e., middle links ML(5,1)-ML(5,64) and middle links ML(6,1) and ML(6,64)} correspond to middle switches MS(5,1)-MS(5,16).

[0083] Middle stage **160** is placed alone. All the right going middle links are the middle links ML(4,1)-ML(4,64) and all the left going middle links are middle links ML(5,1)-ML(5,64).

[0084] In one embodiment, in the network **100B** of FIG. 1B, the switches that are placed together are implemented as separate switches then the network **100B** is the generalized folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above. That is the switches that are placed together in input stage **110** and output stage **120** are implemented as a two by four switch and a four by two switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1)-ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1-OL2 being the outputs of the output switch OS1. Similarly in this embodiment of network **100B** all the switches that are placed together in each middle stage are implemented as separate switches.

Hypercube Topology Layout Schemes:

[0085] Referring to layout **100C** of FIG. 1C, in one embodiment, there are sixteen blocks namely Block **1_2**, Block **3_4**, Block **5_6**, Block **7_8**, Block **9_10**, Block **11_12**, Block **13_14**, Block **15_16**, Block **17_18**, Block **19_20**, Block **21_22**, Block **23_24**, Block **25_26**, Block **27_28**, Block **29_30**, and Block **31_32**. Each block implements all the switches in one row of the network **100B** of FIG. 1B, one of the key aspects of the current invention. For example Block

US 2011/0037498 A1

Feb. 17, 2011

7

1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; Middle switch MS(4,1) is denoted by switch 5.

[0086] All the straight middle links are illustrated in layout 100C of FIG. 1C. For example in Block 1_2, inlet links IL1-IL2, outlet links OL1-OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 100C of FIG. 1C.

[0087] Even though it is not illustrated in layout 100C of FIG. 1C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit (hereinafter "sub-integrated circuit block") depending on the applications in different embodiments. There are four quadrants in the layout 100C of FIG. 1C namely top-left, bottom-left, top-right and bottom-right quadrants. Top-left quadrant implements Block 1_2, Block 3_4, Block 5_6, and Block 7_8. Bottom-left quadrant implements Block 9_10, Block 11_12, Block 13_14, and Block 15_16. Top-right quadrant implements Block 17_18, Block 19_20, Block 21_22, and Block 23_24. Bottom-right quadrant implements Block 25_26, Block 27_28, Block 29_30, and Block 31_32. There are two halves in layout 100C of FIG. 1C namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants. Right-half consists of top-right and bottom-right quadrants.

[0088] Recursively in each quadrant there are four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements Block 1_2. Bottom-left sub-quadrant of top-left quadrant implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 5_6. Finally bottom-right sub-quadrant of top-left quadrant implements Block 7_8. Similarly there are two sub-halves in each quadrant. For example in top-left quadrant there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left quadrant implements Block 1_2 and Block 3_4. Right-sub-half of top-left quadrant implements Block 5_6 and Block 7_8. Finally applicant notes that in each quadrant or half the blocks are arranged as a general binary hypercube. Recursively in larger multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 > 32$, the layout in this embodiment in accordance with the current invention, will be such that the super-quadrants will also be arranged in d-ary hypercube manner. (In the embodiment of the layout 100C of FIG. 1C, it is binary hypercube manner since $d=2$, in the network $V_{fold-link}(N_1, N_2, d, s)$ 100B of FIG. 1B).

[0089] Layout 100D of FIG. 1D illustrates the inter-block links between switches 1 and 2 of each block. For example

middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100D of FIG. 1D can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

[0090] Layout 100E of FIG. 1E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 3_4. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100E of FIG. 1E can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

[0091] Layout 100F of FIG. 1F illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100F of FIG. 1F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time division multiplexed single track).

[0092] Layout 100G of FIG. 1G illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100G of FIG. 1G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division multiplexed single track).

[0093] The complete layout for the network 100B of FIG. 1B is given by combining the links in layout diagrams of

US 2011/0037498 A1

Feb. 17, 2011

8

100C, **100D**, **100E**, **100F**, and **100G**. Applicant notes that in the layout **100C** of FIG. 1C, the inter-block links between switch **1** and switch **2** of corresponding blocks are vertical tracks as shown in layout **100D** of FIG. 1D; the inter-block links between switch **2** and switch **3** of corresponding blocks are horizontal tracks as shown in layout **100E** of FIG. 1E; the inter-block links between switch **3** and switch **4** of corresponding blocks are vertical tracks as shown in layout **100F** of FIG. 1F; and finally the inter-block links between switch **4** and switch **5** of corresponding blocks are horizontal tracks as shown in layout **100G** of FIG. 1G. The pattern is alternate vertical tracks and horizontal tracks. It continues recursively for larger networks of $N > 32$ as will be illustrated later.

[0094] Some of the key aspects of the current invention are discussed. 1) All the switches in one row of the multi-stage network **100B** are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks or vertical tracks; 3) Since all the inter-block links are either horizontal or vertical tracks, all the inter-block links can be mapped on to island-style architectures in current commercial FPGA's; 4) The length of the longest wire is about half of the width (or length) of the complete layout (For example middle link **ML(4,4)** is about half the width of the complete layout).

[0095] In accordance with the current invention, the layout **100C** in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ the sub-quadrants, quadrants, and super-quadrants are arranged in d-ary hypercube manner and also the inter-blocks are accordingly connected in d-ary hypercube topology. Even though all the embodiments in the current invention are illustrated for $N_1 = N_2$, the embodiments can be extended for $N_1 \neq N_2$. Referring to layout **100H** of FIG. 1H, illustrates the extension of layout **100C** for the network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 128$; $d = 2$; and $s = 2$. There are four super-quadrants in layout **100H** namely top-left super-quadrant, bottom-left super-quadrant, top-right super-quadrant, bottom-right super-quadrant. Total number of blocks in the layout **100H** is sixty four. Top-left super-quadrant implements the blocks from block **1_2** to block **31_32**. Each block in all the super-quadrants has two more switches namely switch **6** and switch **7** in addition to the switches **[1-5]** illustrated in layout **100C** of FIG. 1C. The inter-block link connection topology is the exactly the same between the switches **1** and **2**; switches **2** and **3**; switches **3** and **4**; switches **4** and **5** as it is shown in the layouts of FIG. 1D, FIG. 1E, FIG. 1F, and FIG. 1G respectively.

[0096] Bottom-left super-quadrant implements the blocks from block **33_34** to block **63_64**. Top-right super-quadrant implements the blocks from block **65_66** to block **95_96**. And bottom-right super-quadrant implements the blocks from block **97_98** to block **127_128**. In all these three super-quadrants also, the inter-block link connection topology is the exactly the same between the switches **1** and **2**; switches **2** and **3**; switches **3** and **4**; switches **4** and **5** as that of the top-left super-quadrant.

[0097] Recursively in accordance with the current invention, the inter-block links connecting the switch **5** and switch **6** will be vertical tracks between the corresponding switches of top-left super-quadrant and bottom-left super-quadrant. And similarly the inter-block links connecting the switch **5** and switch **6** will be vertical tracks between the corresponding switches of top-right super-quadrant and bottom-right super-quadrant. The inter-block links connecting the switch **6**

and switch **7** will be horizontal tracks between the corresponding switches of top-left super-quadrant and top-right super-quadrant. And similarly the inter-block links connecting the switch **6** and switch **7** will be horizontal tracks between the corresponding switches of bottom-left super-quadrant and bottom-right super-quadrant.

[0098] Referring to diagram **100I** of FIG. 1I illustrates a high-level implementation of Block **1_2** (Each of the other blocks have similar implementation) of the layout **100C** of FIG. 1C which represents a generalized folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block **1_2** in **100I** illustrates both the intra-block and inter-block links connected to Block **1_2**. The layout diagram **100I** corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network **100B** of FIG. 1B. As noted before then the network **100B** is the generalized folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above.

[0099] That is the switches that are placed together in Block **1_2** as shown in FIG. 1I are namely input switch **IS1** and output switch **OS1** belonging to switch **1**, illustrated by dotted lines, (as noted before switch **1** is for illustration purposes only, in practice the switches implemented are input switch **IS1** and output switch **OS1**); middle switch **MS(1,1)** and middle switch **MS(7,1)** belonging to switch **2**; middle switch **MS(2,1)** and middle switch **MS(6,1)** belonging to switch **3**; middle switch **MS(3,1)** and middle switch **MS(5,1)** belonging to switch **4**; And middle switch **MS(4,1)** belonging to switch **5**.

[0100] Input switch **IS1** is implemented as two by four switch with the inlet links **IL1** and **IL2** being the inputs of the input switch **IS1** and middle links **ML(1,1)**-**ML(1,4)** being the outputs of the input switch **IS1**; and output switch **OS1** is implemented as four by two switch with the middle links **ML(8,1)**-**ML(8,4)** being the inputs of the output switch **OS1** and outlet links **OL1**-**OL2** being the outputs of the output switch **OS1**.

[0101] Middle switch **MS(1,1)** is implemented as four by four switch with the middle links **ML(1,1)**, **ML(1,2)**, **ML(1,7)** and **ML(1,8)** being the inputs and middle links **ML(2,1)**-**ML(2,4)** being the outputs; and middle switch **MS(7,1)** is implemented as four by four switch with the middle links **ML(7,1)**, **ML(7,2)**, **ML(7,11)** and **ML(7,12)** being the inputs and middle links **ML(8,1)**-**ML(8,4)** being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in **100I** of FIG. 1I.

[0102] Now the VLSI layouts of generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 < 32$; $d = 2$; $s = 2$ and its corresponding version of folded generalized multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 < 32$; $d = 2$; $s = 2$ are discussed. Referring to diagram **200A1** of FIG. 2A1 is generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 2$; $d = 2$. Diagram **200A2** of FIG. 2A2 illustrates the corresponding folded generalized multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 2$; $d = 2$, version of the diagram **200A1** of FIG. 2A1. Layout **200A3** of FIG. 2A3 illustrates the VLSI layout of the network **200A2** of FIG. 2A2. There is only one block i.e., Block **1_2** comprising switch **1**. Just like in the layout **100C** of FIG. 1C, switch **1** consists of input switch **IS1** and output switch **OS1**.

US 2011/0037498 A1

Feb. 17, 2011

9

[0103] Referring to diagram **200B1** of FIG. **2B1** is generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1=N_2=4$; $d=2$; $s=2$. Diagram **200B2** of FIG. **2B2** illustrates the corresponding folded generalized multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2=4$; $d=2$; $s=2$, version of the diagram **200B1** of FIG. **2B1**. Layout **200B3** of FIG. **2B3** illustrates the VLSI layout of the network **200B2** of FIG. **2B2**. There are two blocks i.e., Block **1_2** and Block **3_4** each comprising switch **1** and switch **2**. Switch **1** in each block consists of the corresponding input switch and output switch. For example switch **1** in Block **1_2** consists of input switch **IS1** and output switch **OS1**. Similarly switch **2** in Block **1_2** consists of middle switch **(1,1)**. Layout **200B4** of FIG. **2B4** illustrates the inter-block links of the VLSI layout diagram **200B3** of FIG. **2B3**. For example middle links **ML(1,4)** and **ML(2,8)**. It must be noted that all the inter-block links are vertical tracks in this layout. (Alternatively all the inter-blocks can also be implemented as horizontal tracks).

[0104] Referring to diagram **200C11** of FIG. **2C11** is generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1=N_2=8$; $d=2$; $s=2$. Diagram **200C12** of FIG. **2C12** illustrates the corresponding folded generalized multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2=8$; $d=2$; $s=2$, version of the diagram **200C11** of FIG. **2C11**. Layout **200C21** of FIG. **2C21** illustrates the VLSI layout of the network **200C12** of FIG. **2C12**. There are four blocks i.e., Block **1_2**, Block **3_4**, Block **5_6**, and Block **7_8** each comprising switch **1**, switch **2** and switch **3**. For example switch **1** in Block **1_2** consists of input switch **IS1** and output switch **OS1**; Switch **2** in Block **1_2** consists of **MS(1,1)** and **MS(3,1)**. Switch **3** in Block **1_2** consists of **MS(2,1)**.

[0105] Layout **200C22** of FIG. **2C22** illustrates the inter-block links between the switch **1** and switch **2** of the VLSI layout diagram **200C21** of FIG. **2C21**. For example middle links **ML(1,4)** and **ML(4,8)** are connected between Block **1_2** and Block **3_4**. It must be noted that all the inter-block links between switch **1** and switch **2** of all blocks are vertical tracks in this layout. Layout **200C23** of FIG. **2C23** illustrates the inter-block links between the switch **2** and switch **3** of the VLSI layout diagram **200C21** of FIG. **2C21**. For example middle links **ML(2,12)** and **ML(3,4)** are connected between Block **1_2** and Block **5_6**. It must be noted that all the inter-block links between switch **2** and switch **3** of all blocks are horizontal tracks in this layout.

[0106] Referring to diagram **200D1** of FIG. **2D1** is generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1=N_2=16$; $d=2$; $s=2$. Diagram **200D2** of FIG. **2D2** illustrates the corresponding folded generalized multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2=16$; $d=2$; $s=2$, version of the diagram **200D1** of FIG. **2D1**. Layout **200D3** of FIG. **2D3** illustrates the VLSI layout of the network **200D2** of FIG. **2D2**. There are eight blocks i.e., Block **1_2**, Block **3_4**, Block **5_6**, Block **7_8**, Block **9_10**, Block **11_12**, Block **13_14** and Block **15_16** each comprising switch **1**, switch **2**, switch **3** and switch **4**. For example switch **1** in Block **1_2** consists of input switch **IS1** and output switch **OS1**; Switch **2** in Block **1_2** consists of **MS(1,1)** and **MS(5,1)**. Switch **3** in Block **1_2** consists of **MS(2,1)** and **MS(4,1)**, and switch **4** in Block **1_2** consists of **MS(3,1)**.

[0107] Layout **200D4** of FIG. **2D4** illustrates the inter-block links between the switch **1** and switch **2** of the VLSI layout diagram **200D3** of FIG. **2D3**. For example middle links **ML(1,4)** and **ML(6,8)** are connected between Block **1_2** and Block **3_4**. It must be noted that all the inter-block links

between switch **1** and switch **2** of all blocks are vertical tracks in this layout. Layout **200D5** of FIG. **2D5** illustrates the inter-block links between the switch **2** and switch **3** of the VLSI layout diagram **200D3** of FIG. **2D3**. For example middle links **ML(2,12)** and **ML(5,4)** are connected between Block **1_2** and Block **5_6**. It must be noted that all the inter-block links between switch **2** and switch **3** of all blocks are horizontal tracks in this layout. Layout **200D6** of FIG. **2D6** illustrates the inter-block links between the switch **3** and switch **4** of the VLSI layout diagram **200D3** of FIG. **2D3**. For example middle links **ML(3,4)** and **ML(4,20)** are connected between Block **1_2** and Block **9_10**. It must be noted that all the inter-block links between switch **3** and switch **4** of all blocks are vertical tracks in this layout.

Generalized Multi-link Butterfly Fat Tree Network Embodiment

[0108] In another embodiment in the network **100B** of FIG. **1B**, the switches that are placed together are implemented as combined switch then the network **100B** is the generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above. That is the switches that are placed together in input stage **110** and output stage **120** are implemented as a six by six switch. For example the input switch **IS1** and output switch **OS1** are placed together; so input switch **IS1** and output **OS1** are implemented as a six by six switch with the inlet links **IL1**, **IL2**, **ML(8,1)**, **ML(8,2)**, **ML(8,7)** and **ML(8,8)** being the inputs of the combined switch (denoted as **IS1&OS1**) and middle links **ML(1,1)**, **ML(1,2)**, **ML(1,3)**, **ML(1,4)**, **OL1** and **OL2** being the outputs of the combined switch **IS1&OS1**. Similarly in this embodiment of network **100B** all the switches that are placed together are implemented as a combined switch.

[0109] Layout diagrams **100C** in FIG. **1C**, **100D** in FIG. **1D**, **100E** in FIG. **1E**, **100F** in FIG. **1G** are also applicable to generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages. The layout **100C** in FIG. **1C** can be recursively extended for any arbitrarily large generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$. Accordingly layout **100H** of FIG. **1H** is also applicable to generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$.

[0110] Referring to diagram **100J** of FIG. **1J** illustrates a high-level implementation of Block **1_2** (Each of the other blocks have similar implementation) of the layout **100C** of FIG. **1C** which represents a generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block **1_2** in **100J** illustrates both the intra-block and inter-block links. The layout diagram **100J** corresponds to the embodiment where the switches that are placed together are implemented as combined switch in the network **100B** of FIG. **1B**. As noted before then the network **100B** is the generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above.

[0111] That is the switches that are placed together in Block **1_2** as shown in FIG. **1J** are namely the combined input and output switch **IS1&OS1** belonging to switch **1**, illustrated by dotted lines, (as noted before switch **1** is for illustration purposes only, in practice the switch implemented is combined input and output switch **IS1&OS1**); middle switch **MS(1,1)**

US 2011/0037498 A1

Feb. 17, 2011

10

belonging to switch 2; middle switch MS(2,1) belonging to switch 3; middle switch MS(3,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0112] Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links IL1, IL2 and ML(8,1)-ML(8,4) being the inputs and middle links ML(1,1)-ML(1,4), and outlet links OL1-OL2 being the outputs.

[0113] Middle switch MS(1,1) is implemented as eight by eight switch with the middle links ML(1,1), ML(1,2), ML(1,7), ML(1,8), ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs and middle links ML(2,1)-ML(2,4) and middle links ML(8,1)-ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as eight by eight switches as illustrated in 100J of FIG. 1J. Applicant observes that in middle switch MS(1,1) any one of the right going middle links can be switched to any one of the left going middle links and hereinafter middle switch MS(1,1) provides U-turn links. In general, in the network $V_{mink-bft}(N_1, N_2, d, s)$ each input switch, each output switch and each middle switch provides U-turn links.

[0114] In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of $V_{mink-bft}(N_1, N_2, d, s)$ can be implemented as a four by eight switch and a four by four switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1J, the left going middle links namely ML(7,1), ML(7,2), ML(7,11), and ML(7,12) are never switched to the right going middle links ML(2,1), ML(2,2), ML(2,3), and ML(2,4). And hence to implement MS(1,1) two switches namely: 1) a four by eight switch with the middle links ML(1,1), ML(1,2), ML(1,7), and ML(1,8) as inputs and the middle links ML(2,1), ML(2,2), ML(2,3), ML(2,4), ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs and 2) a four by four switch with the middle links ML(7,1), ML(7,2), ML(7,11), and ML(7,12) as inputs and the middle links ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs are sufficient without losing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

Generalized Multi-Stage Network Embodiment

[0115] In one embodiment, in the network 100B of FIG. 1B, the switches that are placed together are implemented as two separate switches in input stage 110 and output stage 120; and as four separate switches in all the middle stages, then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch and a four by two switch respectively. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,4), ML(8,7) and ML(8,8) being the inputs and outlet links OL1-OL2 being the outputs.

[0116] The switches, corresponding to the middle stages that are placed together are implemented as four two by two

switches. For example middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs; Similarly in this embodiment of network 100B all the switches that are placed together are implemented as separate switches.

[0117] Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$.

[0118] Referring to diagram 100K of FIG. 1K illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block 1_2 in 100K illustrates both the intra-block and inter-block links. The layout diagram 100K corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above.

[0119] That is the switches that are placed together in Block 1_2 as shown in FIG. 1K are namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches MS(1,1), MS(1,17), MS(7,1) and MS(7,17) belonging to switch 2; middle switches MS(2,1), MS(2,17), MS(6,1) and MS(6,17) belonging to switch 3; middle switches MS(3,1), MS(3,17), MS(5,1) and MS(5,17) belonging to switch 4; And middle switches MS(4,1), and MS(4,17) belonging to switch 5.

[0120] Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,4), ML(8,7) and ML(8,8) being the inputs and outlet links OL1-OL2 being the outputs.

[0121] Middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2)

US 2011/0037498 A1

Feb. 17, 2011

11

and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as illustrated in 100K of FIG. 1K.

Generalized Multi-Stage Network Embodiment with S=1

[0122] In one embodiment, in the network 100B of FIG. 1B (where it is implemented with $s=1$), the switches that are placed together are implemented as two separate switches in input stage 110 and output stage 120; and as two separate switches in all the middle stages, then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages as disclosed in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by two switch and a two by two switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1-OL2 being the outputs.

[0123] The switches, corresponding to the middle stages that are placed together are implemented as two, two by two switches. For example middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs; Similarly in this embodiment of network 100B all the switches that are placed together are implemented as two separate switches.

[0124] Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$.

[0125] Referring to diagram 100K1 of FIG. 1K1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 100C of FIG. 1C when $s=1$ which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ (All the double links are replaced by single links when $s=1$). Block 1_2 in 100K1 illustrates both the intra-block and inter-block links. The layout diagram 100K1 corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 100B of FIG. 1B when $s=1$. As noted before then the network 100B is the generalized folded multi-stage network

$V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages as disclosed in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above.

[0126] That is the switches that are placed together in Block 1_2 as shown in FIG. 1K1 are namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches MS(1,1) and MS(7,1) belonging to switch 2; middle switches MS(2,1) and MS(6,1) belonging to switch 3; middle switches MS(3,1) and MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0127] Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1-OL2 being the outputs.

[0128] Middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; And middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as illustrated in 100K1 of FIG. 1K1.

Generalized Butterfly Fat Tree Network Embodiment

[0129] In another embodiment in the network 100B of FIG. 1B, the switches that are placed together are implemented as two combined switches then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch. For example the input switch IS1 and output switch OS1 are placed together; so input output switch IS1&OS1 are implemented as a six by six switch with the inlet links IL1, IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and OL2 being the outputs of the combined switch IS1&OS1.

[0130] The switches, corresponding to the middle stages that are placed together are implemented as two four by four switches. For example middle switches MS(1,1) and MS(1,17) are placed together; so middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; middle switch MS(1,17) is implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly in this embodiment of network 100B all the switches that are placed together are implemented as two combined switches.

[0131] Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where

US 2011/0037498 A1

Feb. 17, 2011

12

$N_1=N_2=32$; $d=2$; and $s=2$ with five stages. The layout **100C** in FIG. **1C** can be recursively extended for any arbitrarily large generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$. Accordingly layout **100H** of FIG. **1H** is also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$.

[0132] Referring to diagram **100L** of FIG. **1L** illustrates a high-level implementation of Block **1_2** (Each of the other blocks have similar implementation) of the layout **100C** of FIG. **1C** which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block **1_2** in **100L** illustrates both the intra-block and inter-block links. The layout diagram **100L** corresponds to the embodiment where the switches that are placed together are implemented as two combined switches in the network **100B** of FIG. **1B**. As noted before then the network **100B** is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above.

[0133] That is the switches that are placed together in Block **1_2** as shown in FIG. **1L** are namely the combined input and output switch **IS1&OS1** belonging to switch **1**, illustrated by dotted lines, (as noted before switch **1** is for illustration purposes only, in practice the switch implemented is combined input and output switch **IS1&OS1**); middle switch **MS(1,1)** and **MS(1,17)** belonging to switch **2**; middle switch **MS(2,1)** and **MS(2,17)** belonging to switch **3**; middle switch **MS(3,1)** and **MS(3,17)** belonging to switch **4**; And middle switch **MS(4,1)** belonging to switch **5**.

[0134] Combined input and output switch **IS1&OS1** is implemented as six by six switch with the inlet links **IL1**, **IL2**, **ML(8,1)**, **ML(8,2)**, **ML(8,7)** and **ML(8,8)** being the inputs and middle links **ML(1,1)**-**ML(1,4)** and outlet links **OL1**-**OL2** being the outputs.

[0135] Middle switch **MS(1,1)** is implemented as four by four switch with middle links **ML(1,1)**, **ML(1,7)**, **ML(7,1)** and **ML(7,11)** being the inputs and middle links **ML(2,1)**, **ML(2,3)**, **ML(8,1)** and **ML(8,3)** being the outputs; And middle switch **MS(1,17)** is implemented as four by four switch with the middle links **ML(1,2)**, **ML(1,8)**, **ML(7,2)** and **ML(7,12)** being the inputs and middle links **ML(2,2)**, **ML(2,4)**, **ML(8,2)** and **ML(8,4)** being the outputs. Similarly all the other middle switches are also implemented as two four by four switches as illustrated in **100L** of FIG. **1L**. Applicant observes that in middle switch **MS(1,1)** any one of the right going middle links can be switched to any one of the left going middle links and hereinafter middle switch **MS(1,1)** provides U-turn links. In general, in the network $V_{bft}(N_1, N_2, d, s)$ each input switch, each output switch and each middle switch provides U-turn links.

[0136] In another embodiment, middle switch **MS(1,1)** (or the middle switches in any of the middle stage excepting the root middle stage) of Block **1_2** of $V_{bft}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch **MS(1,1)** of Block **1_2** as shown FIG. **1L**, the left going middle links namely **ML(7,1)** and **ML(7,11)** are never switched to the right going middle links **ML(2,1)** and **ML(2,3)**. And hence to implement **MS(1,1)** two switches namely: 1) a two by four switch with the middle links **ML(1,1)** and **ML(1,7)** as inputs and the middle links **ML(2,1)**, **ML(2,3)**, **ML(8,1)**, and **ML(8,3)** as outputs and 2) a two by two switch with the middle links

ML(7,1) and **ML(7,11)** as inputs and the middle links **ML(8,1)** and **ML(8,3)** as outputs are sufficient without losing any connectivity of the embodiment of **MS(1,1)** being implemented as an eight by eight switch as described before.)

Generalized Butterfly Fat Tree Network Embodiment with $S=1$

[0137] In one embodiment, in the network **100B** of FIG. **1B** (where it is implemented with $s=1$), the switches that are placed together are implemented as a combined switch in input stage **110** and output stage **120**; and as a combined switch in all the middle stages, then the network **100B** is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with five stages as disclosed in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above. That is the switches that are placed together in input stage **110** and output stage **120** are implemented as a four by four switch. For example the switch input switch **IS1** and output switch **OS1** are placed together; so input and output switch **IS1&OS1** is implemented as four by four switch with the inlet links **IL1**, **IL2**, **ML(8,1)** and **ML(8,3)** being the inputs and middle links **ML(1,1)**-**ML(1,2)** and outlet links **OL1**-**OL2** being the outputs

[0138] The switches, corresponding to the middle stages that are placed together are implemented as a four by four switch. For example middle switches **MS(1,1)** is implemented as four by four switch with middle links **ML(1,1)**, **ML(1,3)**, **ML(7,1)** and **ML(7,5)** being the inputs and middle links **ML(2,1)**, **ML(2,2)**, **ML(8,1)** and **ML(8,2)** being the outputs.

[0139] Layout diagrams **100C** in FIG. **1C**, **100D** in FIG. **1D**, **100E** in FIG. **1E**, **100F** in FIG. **1F** are also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with five stages. The layout **100C** in FIG. **1C** can be recursively extended for any arbitrarily large generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$. Accordingly layout **100H** of FIG. **1H** is also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$.

[0140] Referring to diagram **100L1** of FIG. **1L1** illustrates a high-level implementation of Block **1_2** (Each of the other blocks have similar implementation) for the layout **100C** of FIG. **1C** when $s=1$ which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ (All the double links are replaced by single links when $s=1$). Block **1_2** in **100L1** illustrates both the intra-block and inter-block links. The layout diagram **100L1** corresponds to the embodiment where the switches that are placed together are implemented as a combined switch in the network **100B** of FIG. **1B** when $s=1$. As noted before then the network **100B** is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages as disclosed in PCT Application Serial No. PCT/US08/64603 that is incorporated by reference above.

[0141] That is the switches that are placed together in Block **1_2** as shown in FIG. **1L1** are namely the input and output switch **IS1&OS1** belonging to switch **1**, illustrated by dotted lines, (as noted before switch **1** is for illustration purposes only, in practice the switches implemented are input switch **IS1** and output switch **OS1**); middle switch **MS(1,1)** belonging to switch **2**; middle switch **MS(2,1)** belonging to switch **3**; middle switch **MS(3,1)** belonging to switch **4**; And middle switch **MS(4,1)** belonging to switch **5**.

[0142] Input and output switch **IS1&OS1** are placed together; so input and output switch **IS1&OS1** is imple-

US 2011/0037498 A1

Feb. 17, 2011

13

mented as four by four switch with the inlet links IL1, IL2, ML(8,1) and ML(8,3) being the inputs and middle links ML(1,1)-ML(1,2) and outlet links OL1-OL2 being the outputs.

[0143] Middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,3), ML(7,1) and ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 100L1 of FIG. 1L1.

[0144] In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of $V_{m-link-bft}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1L1, the left going middle links namely ML(7,1) and ML(7,5) are never switched to the right going middle links ML(2,1) and ML(2,2). And hence to implement MS(1,1) two switches namely: 1) a two by four switch with the middle links ML(1,1) and ML(1,3) as inputs and the middle links ML(2,1), ML(2,2), ML(8,1), and ML(8,2) as outputs and 2) a two by two switch with the middle links ML(7,1) and ML(7,5) as inputs and the middle links ML(8,1) and ML(8,2) as outputs are sufficient without loosing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

Hypercube-Like Topology Layout Schemes:

[0145] Referring to diagram 300A in FIG. 3A, in one embodiment, an exemplary generalized multi-link multi-stage network $V_{m-link}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages of one hundred and forty four switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, 150, 170, 170, 180 and 190 is shown where input stage 110 consists of sixteen, two by four switches IS1-IS16 and output stage 120 consists of sixteen, four by two switches OS1-OS16.

[0146] As disclosed in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above, such a network can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections.

[0147] The diagram 300A in FIG. 3A is exactly the same as the diagram 100A in FIG. 1A excepting the connection links between middle stage 150 and middle stage 160 as well as between middle stage 160 and middle stage 170.

[0148] Each of the N/d middle switches are connected to exactly d switches in middle stage 160 through two links each for a total of $2 \times d$ links (for example the links ML(4,1) and ML(4,2) are connected from middle switch MS(3,1) to middle switch MS(4,1), and the links ML(4,3) and ML(4,4) are connected from middle switch MS(3,1) to middle switch MS(4,15)).

[0149] Each of the N/d middle switches MS(4,1)-MS(4,16) in the middle stage 160 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(4,1) and ML(4,2) are connected to the middle switch MS(4,1) from input switch MS(3,1), and the links ML(4,59) and ML(4,60) are connected to the middle

switch MS(4,1) from input switch MS(3,15)) and also are connected to exactly d switches in middle stage 170 through two links each for a total of $2 \times d$ links (for example the links ML(5,1) and ML(5,2) are connected from middle switch MS(4,1) to middle switch MS(5,1), and the links ML(5,3) and ML(5,4) are connected from middle switch MS(4,1) to middle switch MS(5,15)).

[0150] Each of the N/d middle switches MS(5,1)-MS(5,16) in the middle stage 170 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the links ML(5,1) and ML(5,2) are connected to the middle switch MS(5,1) from input switch MS(4,1), and the links ML(5,59) and ML(5,60) are connected to the middle switch MS(5,1) from input switch MS(4,15)).

[0151] Finally the connection topology of the network 100A shown in FIG. 1A is also basically back to back inverse Benes connection topology but with a slight variation. All the cross middle links from middle switches MS(3,1)-MS(3,8) connect to middle switches MS(4,9)-MS(4,16) and all the cross middle links from middle switches MS(3,9)-MS(3,16) connect to middle switches MS(4,1)-MS(4,8). Applicant makes a key observation that there are many combinations of connections possible using this property. The difference in the connection topology between diagram 100A of FIG. 1A and diagram 300A of FIG. 3A is that the connections formed by cross middle links between middle stage 150 and middle stage 160 are made of two different combinations otherwise both the diagrams 100A and 300A implement back to back inverse Benes connection topology. Since these networks implement back to back inverse Benes topologies since there is difference in the connections of cross middle links between middle stage 150 and middle stage 160, the same difference in the connections of cross middle links between 160 and middle stage 170 occurs.

[0152] Referring to diagram 300B in FIG. 3B, is a folded version of the multi-link multi-stage network 300A shown in FIG. 3A. The network 300B in FIG. 3B shows input stage 110 and output stage 120 are placed together. That is input switch IS1 and output switch OS1 are placed together, input switch IS2 and output switch OS2 are placed together, and similarly input switch IS16 and output switch OS16 are placed together. All the right going middle links {i.e., inlet links IL1-IL32 and middle links ML(1,1)-ML(1,64)} correspond to input switches IS1-IS16, and all the left going middle links {i.e., middle links ML(7,1)-ML(7,64) and outlet links OL1-OL32} correspond to output switches OS1-OS16.

[0153] Just the same way there is difference in the connection topology between diagram 100A of FIG. 1A and diagram 300A of FIG. 3A in the way the connections are formed by cross middle links between middle stage 150 and middle stage 160 and also between middle stage 160 and middle stage 170, the exact similar difference is there between the diagram 100B of FIG. 1B and the diagram 300B of FIG. 3B, i.e., in the way the connections are formed by cross middle links between middle stage 150 and middle stage 160 and also between middle stage 160 and middle stage 170.

[0154] In one embodiment, in the network 300B of FIG. 3B, the switches that are placed together are implemented as separate switches then the network 300B is the generalized folded multi-link multi-stage network $V_{fold-m-link}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in PCT Application Serial No. PCT/US08/64604 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are

US 2011/0037498 A1

Feb. 17, 2011

14

implemented as a two by four switch and a four by two switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1)-ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1-OL2 being the outputs of the output switch OS1. Similarly in this embodiment of network 300B all the switches that are placed together are implemented as separate switches.

[0155] Referring to layout 300C of FIG. 3C, in one embodiment, there are sixteen blocks namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, Block 23_24, Block 25_26, Block 27_28, Block 29_30, and Block 31_32. Each block implements all the switches in one row of the network 300B of FIG. 3B, one of the key aspects of the current invention. For example Block 1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; And middle switch MS(4,1) is denoted by switch 5.

[0156] All the straight middle links are illustrated in layout 300C of FIG. 3C. For example in Block 1_2, inlet links IL1-IL2, outlet links OL1-OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 300C of FIG. 3C.

[0157] Even though it is not illustrated in layout 300C of FIG. 3C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit or sub-integrated circuit block depending on the applications in different embodiments. There are four quadrants in the layout 300C of FIG. 3C namely top-left, bottom-left, top-right and bottom-right quadrants. Top-left quadrant implements Block 1_2, Block 3_4, Block 5_6, and Block 7_8. Bottom-left quadrant implements Block 9_10, Block 11_12, Block 13_14, and Block 15_16. Top-right quadrant implements Block 25_26, Block 27_28, Block 29_30, and Block 31_32. Bottom-right quadrant implements Block 17_18, Block 19_20, Block 21_22, and Block 23_24. There are two halves in layout 300C of FIG. 3C namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants. Right-half consists of top-right and bottom-right quadrants.

[0158] Recursively in each quadrant there are four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements

Block 1_2. Bottom-left sub-quadrant of top-left quadrant implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 7_8. Finally bottom-right sub-quadrant of top-left quadrant implements Block 5_6. Similarly there are two sub-halves in each quadrant. For example in top-left quadrant there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left quadrant implements Block 1_2 and Block 3_4. Right-sub-half of top-left quadrant implements Block 7_8 and Block 5_6. Recursively in larger multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2>32$, the layout in this embodiment in accordance with the current invention, will be such that the super-quadrants will also be arranged in a similar manner.

[0159] Layout 300D of FIG. 3D illustrates the inter-block links (in the layout 300C of FIG. 3C all the cross middle links are inter-block links) between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100D of FIG. 1D can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

[0160] Layout 300E of FIG. 3E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 3_4. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 300E of FIG. 3E can be implemented as diagonal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

[0161] Layout 300F of FIG. 3F illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 300F of FIG. 3F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time division multiplexed single track).

[0162] Layout 300G of FIG. 3G illustrates the inter-block links between switches 4 and 5 of each block. For example

US 2011/0037498 A1

Feb. 17, 2011

15

middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 300G of FIG. 3G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division multiplexed single track).

[0163] The complete layout for the network 300B of FIG. 3B is given by combining the links in layout diagrams of 300C, 300D, 300E, 300F, and 300G. Applicant notes that in the layout 300C of FIG. 3C, the inter-block links between switch 1 and switch 2 are vertical tracks as shown in layout 300D of FIG. 3D; the inter-block links between switch 2 and switch 3 are horizontal tracks as shown in layout 300E of FIG. 3E; the inter-block links between switch 3 and switch 4 are vertical tracks as shown in layout 300F of FIG. 3F; and finally the inter-block links between switch 4 and switch 5 are horizontal tracks as shown in layout 300G of FIG. 3G. The pattern is either vertical tracks, horizontal tracks or diagonal tracks. It continues recursively for larger networks of $N > 32$ as will be illustrated later.

[0164] Some of the key aspects of the current invention related to layout diagram 300C of FIG. 3C are noted. 1) All the switches in one row of the multi-stage network 300B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks, vertical tracks or diagonal tracks; 3) The length of the longest wire is about half of the width (or length) of the complete layout (For example middle link ML(4,4) is about half the width of the complete layout.);

[0165] The layout 300C in FIG. 3C can be recursively extended for any arbitrarily large generalized folded multi link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$. Referring to layout 300H of FIG. 3H, illustrates the extension of layout 300C for the network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 128$; $d = 2$; and $s = 2$. There are four super-quadrants in layout 300H namely top-left super-quadrant, bottom-left super-quadrant, top-right super-quadrant, bottom-right super-quadrant. Total number of blocks in the layout 300H is sixty four. Top-left super-quadrant implements the blocks from block 1_2 to block 31_32. Each block in all the super-quadrants has two more switches namely switch 6 and switch 7 in addition to the switches [1-5] illustrated in layout 300C of FIG. 3C. The inter-block link connection topology is the exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as it is shown in the layouts of FIG. 3D, FIG. 3E, FIG. 3F, and FIG. 3G respectively.

[0166] Bottom-left super-quadrant implements the blocks from block 33_34 to block 63_64. Top-right super-quadrant implements the blocks from block 65_66 to block 95_96. And bottom-right super-quadrant implements the blocks from block 97_98 to block 127_128. In all these three super-quadrants also, the inter-block link connection topology is the exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as that of the top-left super-quadrant.

[0167] Recursively in accordance with the current invention, the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-left super-quadrant and bottom-left super-quadrant. And similarly the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-right super-quadrant and bottom-right super-quadrant. The inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of top-left super-quadrant and top-right super-quadrant. And similarly the inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of bottom-left super-quadrant and bottom-right super-quadrant.

Ring Topology Layout Schemes:

[0168] Layout diagram 400C of FIG. 4C is another embodiment for the generalized folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ diagram 100B in FIG. 1B.

[0169] Referring to layout 400C of FIG. 4C, there are sixteen blocks namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, Block 23_24, Block 25_26, Block 27_28, Block 29_30, and Block 31_32. Each block implements all the switches in one row of the network 100B of FIG. 1B, one of the key aspects of the current invention. For example Block 1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; And middle switch MS(4,1) is denoted by switch 5.

[0170] All the straight middle links are illustrated in layout 400C of FIG. 4C. For example in Block 1_2, inlet links IL1-IL2, outlet links OL1-OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 400C of FIG. 4C.

[0171] Even though it is not illustrated in layout 400C of FIG. 4C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit or sub-integrated circuit block depending on the applications in different embodiments. The topology of the layout 400C in FIG. 4C is a ring. For each of the neighboring rows in diagram 100B of FIG. 1B the corresponding blocks are also physically neighbors in layout diagram 400C of FIG. 4C. In addition the topmost row is also logically considered as neighbor to the bottommost row. For example Block 1_2 (implementing the switches belonging to a row in diagram 100B of FIG. 1B) has Block 3_4 as neighbor since Block 3_4 implements the switches in its neighboring row. Similarly Block 1_2 also has Block 31_32 as neighbor since Block 1_2 implements topmost row of switches and Block 31_32 implements bottommost row of switches in diagram 100B of FIG.

US 2011/0037498 A1

Feb. 17, 2011

16

1B. The ring layout scheme illustrated in 400C of FIG. 4C can be generalized for a large multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2>32$, in accordance with the current invention.

[0172] Layout 400B of FIG. 4B illustrates the inter-block links (in the layout 400A of FIG. 4A all the cross middle links are inter-block links) between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 400B of FIG. 4B are implemented as vertical tracks or horizontal tracks or diagonal tracks. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

[0173] Layout 400C of FIG. 4C illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 3_4. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 400C of FIG. 4C are implemented as vertical tracks or horizontal tracks or diagonal tracks. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

[0174] Layout 400D of FIG. 4D illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 400D of FIG. 4D are implemented as vertical tracks or horizontal tracks or diagonal tracks. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time division multiplexed single track).

[0175] Layout 400E of FIG. 4E illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 400E of FIG. 4E are implemented as vertical tracks or horizontal tracks or diagonal tracks. Also in one embodiment inter-block links are implemented as two different tracks (for example middle

links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division multiplexed single track).

[0176] The complete layout for the network 100B of FIG. 1B is given by combining the links in layout diagrams of 400A, 400B, 400C, 400D, and 400E.

[0177] Some of the key aspects of the current invention related to layout diagram 400A of FIG. 4A are noted. 1) All the switches in one row of the multi-stage network 100B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks, vertical tracks or diagonal tracks; 3) Length of the different wires between the same two middle stages is not the same. However it gives an opportunity to implement the most connected circuits to place and route through the blocks which have shorter wires.

[0178] Layout diagram 400C1 of FIG. 4C1 is another embodiment for the generalized folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ diagram 100B in FIG. 1B. Referring to layout 400C1 of FIG. 4C1, there are sixteen blocks namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, Block 23_24, Block 25_26, Block 27_28, Block 29_30, and Block 31_32. Each block implements all the switches in one row of the network 100B of FIG. 1B, one of the key aspects of the current invention. For example Block 1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; And middle switch MS(4,1) is denoted by switch 5.

[0179] All the straight middle links are illustrated in layout 400C1 of FIG. 4C1. For example in Block 1_2, inlet links IL1-IL2, outlet links OL1-OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 400C1 of FIG. 4C1.

[0180] Even though it is not illustrated in layout 400C1 of FIG. 4C1, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit or sub-integrated circuit block depending on the applications in different embodiments. The topology of the layout 400C1 in FIG. 4C1 is another embodiment of ring layout topology. For each of the neighboring rows in diagram 100B of FIG. 1B the corresponding blocks are also physically neighbors in layout diagram 400C of FIG. 4C. In addition the topmost row is also logically considered as neighbor to the bottommost row. For example Block 1_2 (implementing the switches belonging to a row in diagram 100B of FIG. 1B) has Block 3_4 as neighbor since Block 3_4 implements the switches in its neighboring row. Similarly Block 1_2 also has

US 2011/0037498 A1

Feb. 17, 2011

17

Block **31_32** as neighbor since Block **1_2** implements top-most row of switches and Block **31_32** implements bottom-most row of switches in diagram **100B** of FIG. **1B**. The ring layout scheme illustrated in **400C** of FIG. **4C** can be generalized for a large multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2>32$, in accordance with the current invention.

[0181] All the layout embodiments disclosed in the current invention are applicable to generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ for $s=1, 2, 3$ or any number in general, and for both $N_1=N=N_2$ and $N_1 \neq N_2$, and d is any integer.

[0182] Conversely applicant makes another important observation that generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ are implemented with the layout topology being the hypercube topology shown in layout **100C** of FIG. **1C** with large scale cross point reduction as any one of the networks described in the current invention namely: generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$ for $s=1, 2, 3$ or any number in general, and for both $N_1=N_2=N$ and $N_1 \neq N_2$, and d is any integer.

Applications Embodiments

[0183] All the embodiments disclosed in the current invention are useful in many varieties of applications. FIG. **5A1** illustrates the diagram of **500A1** which is a typical two by two switch with two inlet links namely **IL1** and **IL2**, and two outlet links namely **OL1** and **OL2**. The two by two switch also implements four crosspoints namely **CP(1,1)**, **CP(1,2)**, **CP(2,1)** and **CP(2,2)** as illustrated in FIG. **5A1**. For example the diagram of **500A1** may the implementation of middle switch **MS(1,1)** of the diagram **100K** of FIG. **1K** where inlet link **IL1** of diagram **500A1** corresponds to middle link **ML(1,1)** of diagram **100K**, inlet link **IL2** of diagram **500A1** corresponds to middle link **ML(1,7)** of diagram **100K**, outlet link **OL1** of diagram **500A1** corresponds to middle link **ML(2,1)** of diagram **100K**, outlet link **OL2** of diagram **500A1** corresponds to middle link **ML(2,3)** of diagram **100K**.

1) Programmable Integrated Circuit Embodiments

[0184] All the embodiments disclosed in the current invention are useful in programmable integrated circuit applications. FIG. **5A2** illustrates the detailed diagram **500A2** for the implementation of the diagram **500A1** in programmable integrated circuit embodiments. Each crosspoint is implemented by a transistor coupled between the corresponding inlet link and outlet link, and a programmable cell in programmable integrated circuit embodiments. Specifically crosspoint **CP(1,1)** is implemented by transistor **C(1,1)** coupled between inlet link **IL1** and outlet link **OL1**, and programmable cell **P(1,1)**; crosspoint **CP(1,2)** is implemented by transistor **C(1,2)** coupled between inlet link **IL1** and outlet link **OL2**, and programmable cell **P(1,2)**; crosspoint **CP(2,1)** is imple-

mented by transistor **C(2,1)** coupled between inlet link **IL2** and outlet link **OL1**, and programmable cell **P(2,1)**; and crosspoint **CP(2,2)** is implemented by transistor **C(2,2)** coupled between inlet link **IL2** and outlet link **OL2**, and programmable cell **P(2,2)**.

[0185] If the programmable cell is programmed ON, the corresponding transistor couples the corresponding inlet link and outlet link. If the programmable cell is programmed OFF, the corresponding inlet link and outlet link are not connected. For example if the programmable cell **P(1,1)** is programmed ON, the corresponding transistor **C(1,1)** couples the corresponding inlet link **IL1** and outlet link **OL1**. If the programmable cell **P(1,1)** is programmed OFF, the corresponding inlet link **IL1** and outlet link **OL1** are not connected. In volatile programmable integrated circuit embodiments the programmable cell may be an SRAM (Static Random Access Memory) cell. In non-volatile programmable integrated circuit embodiments the programmable cell may be a Flash memory cell. Also the programmable integrated circuit embodiments may implement field programmable logic arrays (FPGA) devices, or programmable Logic devices (PLD), or Application Specific Integrated Circuits (ASIC) embedded with programmable logic circuits or 3D-FPGAs. **[0186]** FIG. **5A2** also illustrates a buffer **B1** on inlet link **IL2**. The signals driven along inlet link **IL2** are amplified by buffer **B1**. Buffer **B1** can be inverting or non-inverting buffer. Buffers such as **B1** are used to amplify the signal in links which are usually long.

2) One-Time Programmable Integrated Circuit Embodiments

[0187] All the embodiments disclosed in the current invention are useful in one-time programmable integrated circuit applications. FIG. **5A3** illustrates the detailed diagram **500A3** for the implementation of the diagram **500A1** in one-time programmable integrated circuit embodiments. Each crosspoint is implemented by a via coupled between the corresponding inlet link and outlet link in one-time programmable integrated circuit embodiments. Specifically crosspoint **CP(1,1)** is implemented by via **V(1,1)** coupled between inlet link **IL1** and outlet link **OL1**; crosspoint **CP(1,2)** is implemented by via **V(1,2)** coupled between inlet link **IL1** and outlet link **OL2**; crosspoint **CP(2,1)** is implemented by via **V(2,1)** coupled between inlet link **IL2** and outlet link **OL1**; and crosspoint **CP(2,2)** is implemented by via **V(2,2)** coupled between inlet link **IL2** and outlet link **OL2**.

[0188] If the via is programmed ON, the corresponding inlet link and outlet link are permanently connected which is denoted by thick circle at the intersection of inlet link and outlet link. If the via is programmed OFF, the corresponding inlet link and outlet link are not connected which is denoted by the absence of thick circle at the intersection of inlet link and outlet link. For example in the diagram **500A3** the via **V(1,1)** is programmed ON, and the corresponding inlet link **IL1** and outlet link **OL1** are connected as denoted by thick circle at the intersection of inlet link **IL1** and outlet link **OL1**; the via **V(2,2)** is programmed ON, and the corresponding inlet link **IL2** and outlet link **OL2** are connected as denoted by thick circle at the intersection of inlet link **IL2** and outlet link **OL2**; the via **V(1,2)** is programmed OFF, and the corresponding inlet link **IL1** and outlet link **OL2** are not connected as denoted by the absence of thick circle at the intersection of inlet link **IL1** and outlet link **OL2**; the via **V(2,1)** is programmed OFF, and the corresponding inlet link **IL2** and out-

US 2011/0037498 A1

Feb. 17, 2011

18

let link OL1 are not connected as denoted by the absence of thick circle at the intersection of inlet link IL2 and outlet link OL1. One-time programmable integrated circuit embodiments may be anti-fuse based programmable integrated circuit devices or mask programmable structured ASIC devices.

3) Integrated Circuit Placement and Route Embodiments

[0189] All the embodiments disclosed in the current invention are useful in Integrated Circuit Placement and Route applications, for example in ASIC backend Placement and Route tools. FIG. 5A4 illustrates the detailed diagram 500A4 for the implementation of the diagram 500A1 in Integrated Circuit Placement and Route embodiments. In an integrated circuit since the connections are known a-priori, the switch and crosspoints are actually virtual. However the concept of virtual switch and virtual crosspoint using the embodiments disclosed in the current invention reduces the number of required wires, wire length needed to connect the inputs and outputs of different netlists and the time required by the tool for placement and route of netlists in the integrated circuit.

[0190] Each virtual crosspoint is used to either to hardwire or provide no connectivity between the corresponding inlet link and outlet link. Specifically crosspoint CP(1,1) is implemented by direct connect point DCP(1,1) to hardwire (i.e., to permanently connect) inlet link IL1 and outlet link OL1 which is denoted by the thick circle at the intersection of inlet link IL1 and outlet link OL1; crosspoint CP(2,2) is implemented by direct connect point DCP(2,2) to hardwire inlet link IL2 and outlet link OL2 which is denoted by the thick circle at the intersection of inlet link IL2 and outlet link OL2. The diagram 500A4 does not show direct connect point DCP(1,2) and direct connect point DCP(1,3) since they are not needed and in the hardware implementation they are eliminated. Alternatively inlet link IL1 needs to be connected to outlet link OL1 and inlet link IL1 does not need to be connected to outlet link OL2. Also inlet link IL2 needs to be connected to outlet link OL2 and inlet link IL2 does not need to be connected to outlet link OL1. Furthermore in the example of the diagram 500A4, there is no need to drive the signal of inlet link IL1 horizontally beyond outlet link OL1 and hence the inlet link IL1 is not even extended horizontally until the outlet link OL2. Also the absence of direct connect point DCP(2,1) illustrates there is no need to connect inlet link IL2 and outlet link OL1.

[0191] In summary in integrated circuit placement and route tools, the concept of virtual switches and virtual cross points is used during the implementation of the placement & routing algorithmically in software, however during the hardware implementation cross points in the cross state are implemented as hardwired connections between the corresponding inlet link and outlet link, and in the bar state are implemented as no connection between inlet link and outlet link. 3) More Application Embodiments:

[0192] All the embodiments disclosed in the current invention are also useful in the design of SoC interconnects, Field programmable interconnect chips, parallel computer systems and in time-space-time switches.

[0193] Numerous modifications and adaptations of the embodiments, implementations, and examples described herein will be apparent to the skilled artisan in view of the disclosure.

What is claimed is:

1. An integrated circuit device comprising a plurality of sub-integrated circuit blocks and a routing network, and Said each plurality of sub-integrated circuit blocks comprising a plurality of inlet links and a plurality of outlet links; and Said routing network interconnects any one of said outlet link of one of said sub-integrated circuit block to one or more said inlet links of one or more of said sub-integrated circuit blocks; and Said routing network comprising of a plurality of stages y , starting from the lowest stage to the highest stage; and Said routing network comprising a plurality of switches of size $d \times d$, where $d \geq 2$, in each said stage and each said switch of size $d \times d$ having d inlet links and d outlet links; and Said each sub-integrated circuit block comprising a plurality of said switches corresponding to each said stage; and Said each sub-integrated circuit block comprising a plurality of forward connecting links connecting from switches in lower stage to switches in the immediate succeeding higher stage, and also comprising a plurality of backward connecting links connecting from switches in higher stage to switches in the immediate preceding lower stage; and Said each sub-integrated circuit block comprising a plurality straight links in said forward connecting links from switches in lower stage to switches in the immediate succeeding higher stage and a plurality cross links in said forward connecting links from switches in lower stage to switches in the immediate succeeding higher stage, and further comprising a plurality of straight links in said backward connecting links from switches in higher stage to switches in the immediate preceding lower stage and a plurality of cross links in said backward connecting links from switches in higher stage to switches in the immediate preceding lower stage.
2. The integrated circuit device of claim 1, wherein said all straight links are connecting from switches in each said sub-integrated circuit block are connecting to switches in the same said sub-integrated circuit block; and said all cross links are connecting as either vertical or horizontal links between switches in two different said sub-integrated circuit blocks.
3. The integrated circuit device of claim 2, wherein said plurality of sub-integrated circuit blocks arranged in a two-dimensional grid.
4. The integrated circuit device of claim 3, wherein said cross links in succeeding stages are connecting as alternative vertical and horizontal links between switches in said sub-integrated circuit blocks.
5. The integrated circuit device of claim 4, wherein said cross links from switches in a stage in one of said sub-integrated circuit blocks are connecting to switches in the succeeding stage in another of said sub-integrated circuit blocks so that said cross links are either vertical links or horizontal and vice versa, and hereinafter such cross links are "shuffle exchange links".
6. The integrated circuit device of claim 5, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are substantially of equal length and said vertical shuffle exchange links between

US 2011/0037498 A1

Feb. 17, 2011

19

switches in any two corresponding said succeeding stages are substantially of equal length in the entire said integrated circuit device.

7. The integrated circuit device of claim 6, wherein the shortest horizontal shuffle exchange links are connecting at the lowest stage and between switches in two nearest neighboring said sub-integrated circuit blocks, and length of the horizontal shuffle exchange links is doubled in each succeeding stage; and the shortest vertical shuffle exchange links are connecting at the lowest stage and between switches in two nearest neighboring said sub-integrated circuit blocks, and length of the vertical shuffle exchange links is doubled in each succeeding stage.

8. The integrated circuit device of claim 7, wherein $y \geq (\log_2 N)$ so that the length of the horizontal shuffle exchange links in the highest stage is equal to half the size of the horizontal size of said two dimensional grid of sub-integrated circuit blocks and the length of the vertical shuffle exchange links in the highest stage is equal to half the size of the vertical size of said two dimensional grid of sub-integrated circuit blocks.

9. The integrated circuit device of claim 8, wherein $d=2$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast Benes network with full bandwidth.

10. The integrated circuit device of claim 8, wherein $d=2$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast Benes network and rearrangeably nonblocking for arbitrary fan-out multicast Benes network with full bandwidth.

11. The integrated circuit device of claim 8, wherein $d=2$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast Benes network with full bandwidth.

12. The integrated circuit device of claim 7, wherein $y \geq (\log_2 N)$ so that the length of the horizontal shuffle exchange links in the highest stage is equal to half the size of the horizontal size of said two dimensional grid of sub-integrated circuit blocks and the length of the vertical shuffle exchange links in the highest stage is equal to half the size of the vertical size of said two dimensional grid of sub-integrated circuit blocks, and

said each sub-integrated circuit block further comprising a plurality of U-turn links within switches in each of said stages in each of said sub-integrated circuit blocks.

13. The integrated circuit device of claim 12, wherein $d=2$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast butterfly fat tree network with full bandwidth.

14. The integrated circuit device of claim 12, wherein $d=2$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast butterfly fat tree network and rearrangeably nonblocking for arbitrary fan-out multicast butterfly fat tree network with full bandwidth.

15. The integrated circuit device of claim 12, wherein $d=2$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast butterfly fat tree network with full bandwidth.

16. The integrated circuit device of claim 1, wherein said horizontal and vertical links are implemented on two or more metal layers.

17. The integrated circuit device of claim 1, wherein said switches comprising active and reprogrammable cross points and said each cross point is programmable by an SRAM cell or a Flash Cell.

18. The integrated circuit device of claim 1, wherein said sub-integrated circuit blocks are of equal die size.

19. The integrated circuit device of claim 16, wherein said sub-integrated circuit blocks are Lookup Tables (hereinafter "LUTs") and said integrated circuit device is a field programmable gate array (FPGA) device or field programmable gate array (FPGA) block embedded in another integrated circuit device.

20. The integrated circuit device of claim 16, wherein said sub-integrated circuit blocks are AND or OR gates and said integrated circuit device is a programmable logic device (PLD).

21. The integrated circuit device of claim 1, wherein said sub-integrated circuit blocks comprising any arbitrary hardware logic or memory circuits.

22. The integrated circuit device of claim 1, wherein said switches comprising active one-time programmable cross points and said integrated circuit device is a mask programmable gate array (MPGA) device or a structured ASIC device.

23. The integrated circuit device of claim 1, wherein said switches comprising passive cross points or just connection of two links or not and said integrated circuit device is a Application Specific Integrated Circuit (ASIC) device.

24. The integrated circuit device of claim 1, wherein said sub-integrated circuit blocks further recursively comprise one or more super-sub-integrated circuit blocks and a sub-routing network.

25. The integrated circuit device of claim 5, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \geq (\log_2 N)$.

26. The integrated circuit device of claim 25, wherein $d=2$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward

US 2011/0037498 A1

Feb. 17, 2011

20

connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-stage network with full bandwidth.

27. The integrated circuit device of claim 25, wherein $d=2$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized multi-stage network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-stage network with full bandwidth.

28. The integrated circuit device of claim 25, wherein $d=2$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-stage network with full bandwidth.

29. The integrated circuit device of claim 5, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \geq (\log_2 N)$, and

said each sub-integrated circuit block further comprising a plurality of U-turn links within switches in each of said stages in each of said sub-integrated circuit blocks.

30. The integrated circuit device of claim 29, wherein $d=2$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized butterfly fat tree network with full bandwidth.

31. The integrated circuit device of claim 29, wherein $d=2$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized butterfly fat tree Network and rearrangeably nonblocking for arbitrary fan-out multicast generalized butterfly fat tree network with full bandwidth.

32. The integrated circuit device of claim 29, wherein $d=2$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized butterfly fat tree network with full bandwidth.

33. The integrated circuit device of claim 1, wherein said straight links connecting from switches in each said sub-integrated circuit block are connecting to switches in the same said sub-integrated circuit block; and

said cross links are connecting as vertical or horizontal or diagonal links between two different said sub-integrated circuit blocks.

34. The integrated circuit device of claim 8, wherein $d=4$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast multi-link Benes network with full bandwidth.

35. The integrated circuit device of claim 8, wherein $d=4$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast multi-link Benes network and rearrangeably nonblocking for arbitrary fan-out multicast multi-link Benes network with full bandwidth.

36. The integrated circuit device of claim 8, wherein $d=4$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast multi-link Benes network with full bandwidth.

37. The integrated circuit device of claim 12, wherein $d=4$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast multi-link butterfly fat tree network with full bandwidth.

38. The integrated circuit device of claim 12, wherein $d=4$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast multi-link butterfly fat tree network and rearrangeably nonblocking for arbitrary fan-out multicast multi-link butterfly fat tree network with full bandwidth.

39. The integrated circuit device of claim 12, wherein $d=4$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast multi-link butterfly fat tree network with full bandwidth.

40. The integrated circuit device of claim 5, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \geq (\log_2 N)$.

41. The integrated circuit device of claim 40, wherein $d=4$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward

US 2011/0037498 A1

Feb. 17, 2011

21

connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-link multi-stage network with full bandwidth.

42. The integrated circuit device of claim **40**, wherein $d=4$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized multi-link multi-stage network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-link multi-stage network with full bandwidth.

43. The integrated circuit device of claim **40**, wherein $d=4$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-link multi-stage network with full bandwidth.

44. The integrated circuit device of claim **5**, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \cong (\log_2 N)$, and

said each sub-integrated circuit block further comprising a plurality of U-turn links within switches in each of said stages in each of said sub-integrated circuit blocks.

45. The integrated circuit device of claim **44**, wherein $d=4$ and there is only one switch in each said stage in each said sub-integrated circuit block connecting said forward connect-

ing links and there is only one switch in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-link butterfly fat tree network with full bandwidth.

46. The integrated circuit device of claim **44**, wherein $d=4$ and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least two switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized multi-link butterfly fat tree Network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-link butterfly fat tree network with full bandwidth.

47. The integrated circuit device of claim **44**, wherein $d=4$ and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said forward connecting links and there are at least three switches in each said stage in each said sub-integrated circuit block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-link butterfly fat tree network with full bandwidth.

48. The integrated circuit device of claim **1**, wherein said plurality of forward connecting links use a plurality of buffers to amplify signals driven through them and said plurality of backward connecting links use a plurality of buffers to amplify signals driven through them; and said buffers can be inverting or non-inverting buffers.

49. The integrated circuit device of claim **1**, wherein said wherein said all switches of size $d \times d$ are either fully populated or partially populated.

* * * * *

EXHIBIT I

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 April 2011 (21.04.2011)

PCT

(10) International Publication Number
WO 2011/047368 A2

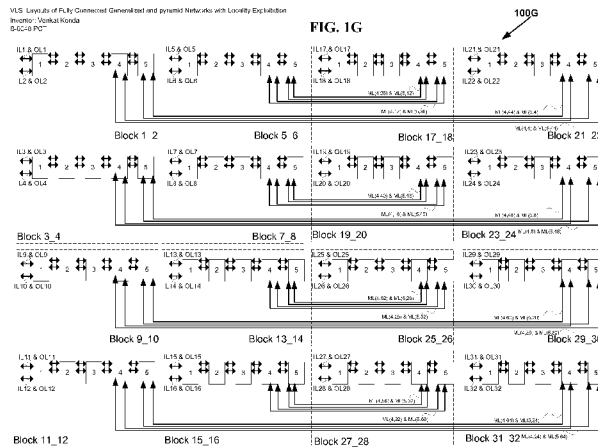
- (51) International Patent Classification:
G06F 17/50 (2006.01)
- (21) International Application Number:
PCT/US2010/052984
- (22) International Filing Date:
16 October 2010 (16.10.2010)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/252,603 16 October 2009 (16.10.2009) US
61/252,609 16 October 2009 (16.10.2009) US
- (72) Inventor; and
- (71) Applicant : **KONDA, Venkat** [US/US]; 6278, Grand Oak Way, San Jose, CA 95135 (US).
- (81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO,

DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) Title: VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS WITH LOCALITY EXPLOITATION



(57) Abstract: VLSI layouts of generalized multi-stage and pyramid networks for broadcast, unicast and multicast connections are presented using only horizontal and vertical links with spacial locality exploitation. The VLSI layouts employ shuffle exchange links where outlet links of cross links from switches in a stage in one sub-integrated circuit block are connected to inlet links of switches in the succeeding stage in another sub-integrated circuit block so that said cross links are either vertical links or horizontal and vice versa. Furthermore the shuffle exchange links are employed between different sub-integrated circuit blocks so that spacially nearer sub-integrated circuit blocks are connected with shorter links compared to the shuffle exchange links between spacially farther sub-integrated circuit blocks. In one embodiment the sub-integrated circuit blocks are arranged in a hypercube arrangement in a two-dimensional plane. The VLSI layouts exploit the benefits of significantly lower cross points, lower signal latency, lower power and full connectivity with significantly fast compilation. The VLSI layouts with spacial locality exploitation presented are applicable to generalized multi-stage and pyramid networks, generalized folded multi-stage and pyramid networks, generalized butterfly fat tree and pyramid networks, generalized multi-link multi-stage and pyramid networks, generalized folded multi-link multi-stage and pyramid networks,

[Continued on next page]



WO 2011/047368 A2

WO 2011/047368 A2 

generalized multi-link butterfly fat tree and pyramid networks, generalized hypercube networks, and generalized cube connected cycles networks for speedup of $s \geq 1$. The embodiments of VLSI layouts are useful in wide target applications such as FPGAs, CPLDs, pSoCs, ASIC placement and route tools, networking applications, parallel & distributed computing, and reconfigurable computing.

WO 2011/047368

PCT/US2010/052984

**VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID
NETWORKS WITH LOCALITY EXPLOITATION**

Venkat Konda

5 CROSS REFERENCE TO RELATED APPLICATIONS

This application is Continuation In Part PCT Application to and incorporates by reference in its entirety the U.S. Provisional Patent Application Serial No. 61/252, 603 entitled "VLSI LAYOUTS OF FULLY CONNECTED NETWORKS WITH LOCALITY EXPLOITATION" by Venkat Konda assigned to the same assignee as the current
10 application, filed October 16, 2009.

This application is Continuation In Part PCT Application to and incorporates by reference in its entirety the U.S. Provisional Patent Application Serial No. 61/252, 609 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS" by Venkat Konda assigned to the same assignee as the current
15 application, filed October 16, 2009.

This application is related to and incorporates by reference in its entirety the US Application Serial No. 12/530,207 entitled "FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed September 6, 2009, the U.S. Provisional Patent Application
20 Serial No. 60/905,526 entitled "LARGE SCALE CROSSPOINT REDUCTION WITH NONBLOCKING UNICAST & MULTICAST IN ARBITRARILY LARGE MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed March 6, 2007, and the U.S. Provisional Patent Application Serial No. 60/940, 383 entitled "FULLY CONNECTED GENERALIZED MULTI-STAGE
25 NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

WO 2011/047368

PCT/US2010/052984

This application is related to and incorporates by reference in its entirety the US Application Serial No. 12/601,273 entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed November 22, 2009, the U.S. Provisional Patent
5 Application Serial No. 60/940, 387 entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, and the U.S. Provisional Patent Application Serial No. 60/940, 390 entitled "FULLY CONNECTED GENERALIZED MULTI-LINK BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to
10 the same assignee as the current application, filed May 25, 2007

This application is related to and incorporates by reference in its entirety the US Application Serial No. 12/601,274 entitled "FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed November 22, 2009, the U.S. Provisional Patent
15 Application Serial No. 60/940, 389 entitled "FULLY CONNECTED GENERALIZED REARRANGEABLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, the U.S. Provisional Patent Application Serial No. 60/940, 391 entitled "FULLY CONNECTED GENERALIZED FOLDED MULTI-STAGE NETWORKS" by Venkat
20 Konda assigned to the same assignee as the current application, filed May 25, 2007 and the U.S. Provisional Patent Application Serial No. 60/940, 392 entitled "FULLY CONNECTED GENERALIZED STRICTLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

25 This application is related to and incorporates by reference in its entirety the US Application Serial No. 12/601,275 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed November 22, 2009, and the U.S. Provisional Patent Application Serial No. 60/940, 394 entitled "VLSI LAYOUTS OF FULLY

WO 2011/047368

PCT/US2010/052984

CONNECTED GENERALIZED NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

BACKGROUND OF INVENTION

5 Multi-stage interconnection networks such as Benes networks and butterfly fat tree networks are widely useful in telecommunications, parallel and distributed computing. However VLSI layouts, known in the prior art, of these interconnection networks in an integrated circuit are inefficient and complicated.

Other multi-stage interconnection networks including butterfly fat tree networks,
10 Banyan networks, Batcher-Banyan networks, Baseline networks, Delta networks, Omega networks and Flip networks have been widely studied particularly for self routing packet switching applications. Also Benes Networks with radix of two have been widely studied and it is known that Benes Networks of radix two are shown to be built with back to back baseline networks which are rearrangeably nonblocking for unicast connections.

15 The most commonly used VLSI layout in an integrated circuit is based on a two-dimensional grid model comprising only horizontal and vertical tracks. An intuitive interconnection network that utilizes two-dimensional grid model is 2D Mesh Network and its variations such as segmented mesh networks. Hence routing networks used in VLSI layouts are typically 2D mesh networks and its variations. However Mesh
20 Networks require large scale cross points typically with a growth rate of $O(N^2)$ where N is the number of computing elements, ports, or logic elements depending on the application.

Multi-stage interconnection network with a growth rate of $O(N \times \log N)$ requires significantly small number of cross points. U.S. Patent 6,185,220 entitled "Grid Layouts
25 of Switching and Sorting Networks" granted to Muthukrishnan et al. describes a VLSI layout using existing VLSI grid model for Benes and Butterfly networks. U.S. Patent 6,940,308 entitled "Interconnection Network for a Field Programmable Gate Array" granted to Wong describes a VLSI layout where switches belonging to lower stage of

WO 2011/047368

PCT/US2010/052984

Benes Network are laid out close to the logic cells and switches belonging to higher stages are laid out towards the center of the layout.

Due to the inefficient and in some cases impractical VLSI layout of Benes and butterfly fat tree networks on a semiconductor chip, today mesh networks and segmented mesh networks are widely used in the practical applications such as field programmable gate arrays (FPGAs), programmable logic devices (PLDs), and parallel computing interconnects. The prior art VLSI layouts of Benes and butterfly fat tree networks and VLSI layouts of mesh networks and segmented mesh networks require large area to implement the switches on the chip, large number of wires, longer wires, with increased power consumption, increased latency of the signals which effect the maximum clock speed of operation. Some networks may not even be implemented practically on a chip due to the lack of efficient layouts.

SUMMARY OF INVENTION

When large scale sub-integrated circuit blocks with inlet and outlet links are laid out in an integrated circuit device in a two-dimensional grid arrangement, (for example in an FPGA where the sub-integrated circuit blocks are Lookup Tables) the most intuitive routing network is a network that uses horizontal and vertical links only (the most often used such a network is one of the variations of a 2D Mesh network). A direct embedding of a generalized multi-stage network on to a 2D Mesh network is neither simple nor efficient.

In accordance with the invention, VLSI layouts of generalized multi-stage and pyramid networks for broadcast, unicast and multicast connections are presented using only horizontal and vertical links with spacial locality exploitation. The VLSI layouts employ shuffle exchange links where outlet links of cross links from switches in a stage in one sub-integrated circuit block are connected to inlet links of switches in the succeeding stage in another sub-integrated circuit block so that said cross links are either vertical links or horizontal and vice versa. Furthermore the shuffle exchange links are employed between different sub-integrated circuit blocks so that spacially nearer sub-

WO 2011/047368

PCT/US2010/052984

integrated circuit blocks are connected with shorter links compared to the shuffle exchange links between spacially farther sub-integrated circuit blocks. In one embodiment the sub-integrated circuit blocks are arranged in a hypercube arrangement in a two-dimensional plane. The VLSI layouts exploit the benefits of significantly lower cross points, lower signal latency, lower power and full connectivity with significantly fast compilation.

The VLSI layouts with spacial locality exploitation presented are applicable to generalized multi-stage and pyramid networks $V(N_1, N_2, d, s)$ & $V_p(N_1, N_2, d, s)$, generalized folded multi-stage and pyramid networks $V_{fold}(N_1, N_2, d, s)$ & $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat tree and butterfly fat pyramid networks $V_{bft}(N_1, N_2, d, s)$ & $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage and pyramid networks $V_{mlink}(N_1, N_2, d, s)$ & $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage and pyramid networks $V_{fold-mlink}(N_1, N_2, d, s)$ & $V_{fold-mlink-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree and butterfly fat pyramid networks $V_{mlink-bft}(N_1, N_2, d, s)$ & $V_{mlink-bfp}(N_1, N_2, d, s)$, generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$, and generalized cube connected cycles networks $V_{CCC}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general. The embodiments of VLSI layouts are useful in wide target applications such as FPGAs, CPLDs, pSoCs, ASIC placement and route tools, networking applications, parallel & distributed computing, and reconfigurable computing.

BRIEF DESCRIPTION OF DRAWINGS

FIG. 1A is a diagram 100A of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having a variation of inverse Benes connection topology of nine stages with $N = 32$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

WO 2011/047368

PCT/US2010/052984

FIG. 1B is a diagram 100B of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-link}(N, d, s)$ of the network 100A shown in FIG. 1A, having a variation of inverse Benes connection topology of five stages with $N = 32$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 1C is a diagram 100C layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links belonging with in each block only.

FIG. 1D is a diagram 100D layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(1,i) for $i = [1, 64]$ and ML(8,i) for $i = [1, 64]$.

FIG. 1E is a diagram 100E layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(2,i) for $i = [1, 64]$ and ML(7,i) for $i = [1, 64]$.

FIG. 1F is a diagram 100F layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(3,i) for $i = [1, 64]$ and ML(6,i) for $i = [1, 64]$.

FIG. 1G is a diagram 100G layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links ML(4,i) for $i = [1, 64]$ and ML(5,i) for $i = [1, 64]$.

FIG. 1H is a diagram 100H layout of a network $V_{fold-link}(N, d, s)$ where $N = 128$, $d = 2$, and $s = 2$, in one embodiment, illustrating the connection links belonging with in each block only.

WO 2011/047368

PCT/US2010/052984

FIG. 1I is a diagram 100I detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V_{mlink}(N, d, s)$ or $V_{fold-mlink}(N, d, s)$.

FIG. 1J is a diagram 100J detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V_{mlink-bft}(N, d, s)$.

FIG. 1K is a diagram 100K detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

FIG. 1K1 is a diagram 100M1 detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$ for $s = 1$.

FIG. 1L is a diagram 100L detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V_{bft}(N, d, s)$.

FIG. 1L1 is a diagram 100L1 detailed connections of BLOCK 1_2 in the network layout 100C in one embodiment, illustrating the connection links going in and coming out when the layout 100C is implementing $V_{bft}(N, d, s)$ for $s = 1$.

FIG. 2A is a diagram 200A of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having inverse Benes connection topology of nine stages with $N = 24$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2B is a diagram 200B of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network 200A shown in FIG. 2A, having inverse

WO 2011/047368

PCT/US2010/052984

Benes connection topology of five stages with $N = 24$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2C is a diagram 200C layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links belonging within each block only.

FIG. 2D is a diagram 200D layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links $ML(1,i)$ for $i = [1, 48]$ and $ML(8,i)$ for $i = [1,48]$.

FIG. 2E is a diagram 200E layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links $ML(2,i)$ for $i = [1, 32]$ and $ML(7,i)$ for $i = [1,32]$.

FIG. 2F is a diagram 200F layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links $ML(3,i)$ for $i = [1, 64]$ and $ML(6,i)$ for $i = [1,64]$.

FIG. 2G is a diagram 200G layout of the network $V_{fold-link}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links $ML(4,i)$ for $i = [1, 64]$ and $ML(5,i)$ for $i = [1,64]$.

FIG. 3A is a diagram 300A layout of the topmost row of the network $V_{fold-link}(N, d, s)$ with $N = 512$, $d = 2$ and $s=2$, in one embodiment, illustrating the provisioning of 2's BW.

FIG. 3B is a diagram 300B layout of the topmost row of the network $V_{fold-link}(N, d, s)$ with $N = 512$, $d = 2$ and $s=2$, in one embodiment, illustrating the provisioning of 4's BW.

WO 2011/047368

PCT/US2010/052984

FIG. 3C is a diagram 300C layout of the topmost row of the network $V_{fold-link}(N, d, s)$ with $N = 512$, $d = 2$ and $s=2$, in one embodiment, illustrating the provisioning of 8's BW with nearest neighbor connectivity first.

FIG. 3D is a diagram 300D layout of the topmost row of the network
5 $V_{fold-link}(N, d, s)$ with $N = 512$, $d = 2$ and $s=2$, in one embodiment, illustrating the provisioning of 8's BW with nearest neighbor connectivity recursively.

FIG. 4A is a diagram 400A layout of the topmost row of the network $V_{fold-link}(N, d, s)$ with $N = 512$, $d = 2$ and $s=2$, in one embodiment, illustrating the provisioning of 2's BW in first stage.

10 FIG. 4B is a diagram 400B layout of the topmost row of the network $V_{fold-link}(N, d, s)$ with $N = 512$, $d = 2$ and $s=2$, in one embodiment, illustrating the remaining nearest neighbor connectivity in the second stage by provisioning 4's BW, 8's BW etc.

FIG. 4C is a diagram 400C layout of the topmost row of the network
15 $V_{fold-link}(N, d, s)$ with $N = 512$, $d = 2$ and $s=2$, in one embodiment, illustrating the third stage, by provisioning 4's and 8's BW.

FIG. 5 is a diagram 500 layout of the topmost row of the network
20 $V_{fold-link}(N, d, s)$ with $N = 512$, $d = 2$ and $s = 2$, in one embodiment, illustrating the provisioning of 8's BW and 16's BW in Partial & Tapered Connectivity (Bandwidth) in a stage.

FIG. 6 is a diagram 600 layout of the topmost row of the network $V_{fold-link}(N, d, s)$ with $N = 2048$, $d = 2$ and $s = 2$, in one embodiment, illustrating the provisioning of 8's BW, 16's BW and 32's BW in Partial & Tapered Connectivity (Bandwidth) in a stage.

WO 2011/047368

PCT/US2010/052984

FIG. 7 is a diagram 700 layout of the topmost row of the network $V_{fold-link}(N, d, s)$ with $N = 2048$, $d = 2$ and $s = 2$, in one embodiment, illustrating the provisioning of 8's BW, 16's BW and 32's BW in Partial & Tapered Connectivity (Bandwidth) in a stage with equal length wires.

5 FIG. 8A is a diagram 800A of an exemplary symmetrical multi-link multi-stage pyramid network $V_{m-link-p}(N, d, s)$ having inverse Benes connection topology of nine stages with $N = 32$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

10 FIG. 8B is a diagram 800B of the equivalent symmetrical folded multi-link multi-stage pyramid network $V_{fold-m-link-p}(N, d, s)$ of the network 800A shown in FIG. 8A, having inverse Benes connection topology of five stages with $N = 32$, $d = 2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

15 FIG. 8C is a diagram 800C layout of the network $V_{fold-m-link-p}(N, d, s)$ shown in FIG. 8B, in one embodiment, illustrating the connection links belonging with in each block only.

FIG. 8D is a diagram 800D layout of the network $V_{fold-m-link-p}(N, d, s)$ shown in FIG. 8B, in one embodiment, illustrating the connection links ML(1,i) for $i = [1, 64]$ and
20 ML(8,i) for $i = [1,64]$.

FIG. 8E is a diagram 800E layout of the network $V_{fold-m-link-p}(N, d, s)$ shown in FIG. 8B, in one embodiment, illustrating the connection links ML(2,i) for $i = [1, 64]$ and ML(7,i) for $i = [1,64]$.

FIG. 8F is a diagram 800F layout of the network $V_{fold-m-link-p}(N, d, s)$ shown in
25 FIG. 8B, in one embodiment, illustrating the connection links ML(3,i) for $i = [1, 64]$ and ML(6,i) for $i = [1,64]$.

WO 2011/047368

PCT/US2010/052984

FIG. 8G is a diagram 800G layout of the network $V_{fold-link-p}(N, d, s)$ shown in FIG. 8B, in one embodiment, illustrating the connection links ML(4,i) for $i = [1, 64]$ and ML(5,i) for $i = [1, 64]$.

FIG. 8H is a diagram 800H layout of a network $V_{fold-link-p}(N, d, s)$ where $N = 5$ 128, $d = 2$, and $s = 2$, in one embodiment, illustrating the connection links belonging with in each block only.

FIG. 8I is a diagram 800I detailed connections of BLOCK 1_2 in the network layout 800C in one embodiment, illustrating the connection links going in and coming out when the layout 800C is implementing $V_{link-p}(N, d, s)$ or $V_{fold-link-p}(N, d, s)$.

10 FIG. 8J is a diagram 800J detailed connections of BLOCK 1_2 in the network layout 800C in one embodiment, illustrating the connection links going in and coming out when the layout 800C is implementing $V_{link-bfp}(N, d, s)$.

FIG. 8K is a diagram 800K detailed connections of BLOCK 1_2 in the network layout 800C in one embodiment, illustrating the connection links going in and coming
15 out when the layout 800C is implementing $V_p(N, d, s)$ or $V_{fold-p}(N, d, s)$.

FIG. 8K1 is a diagram 800M1 detailed connections of BLOCK 1_2 in the network layout 800C in one embodiment, illustrating the connection links going in and coming out when the layout 800C is implementing $V_p(N, d, s)$ or $V_{fold-p}(N, d, s)$ for $s = 1$.

FIG. 8L is a diagram 800L detailed connections of BLOCK 1_2 in the network
20 layout 800C in one embodiment, illustrating the connection links going in and coming out when the layout 800C is implementing $V_{bfp}(N, d, s)$.

FIG. 8L1 is a diagram 800L1 detailed connections of BLOCK 1_2 in the network layout 800C in one embodiment, illustrating the connection links going in and coming out when the layout 800C is implementing $V_{bfp}(N, d, s)$ for $s = 1$.

WO 2011/047368

PCT/US2010/052984

FIG. 9A is high-level flowchart of a scheduling method 900 according to the invention, used to set up the multicast connections in the generalized multi-stage pyramid network and the generalized multi-link multi-stage pyramid network disclosed in this invention.

5 FIG. 10A is high-level flowchart of a scheduling method 1000 according to the invention, used to set up the multicast connections in the generalized butterfly fat pyramid network and the generalized multi-link butterfly fat pyramid network disclosed in this invention.

FIG. 11A1 is a diagram 1100A1 of an exemplary prior art implementation of a two by two switch; FIG. 11A2 is a diagram 1100A2 for programmable integrated circuit prior art implementation of the diagram 1100A1 of FIG. 11A1; FIG. 11A3 is a diagram 1100A3 for one-time programmable integrated circuit prior art implementation of the diagram 1100A1 of FIG. 11A1; FIG. 11A4 is a diagram 1100A4 for integrated circuit placement and route implementation of the diagram 1100A1 of FIG. 11A1.

15

DETAILED DESCRIPTION OF THE INVENTION

The present invention is concerned with the VLSI layouts of arbitrarily large switching networks for broadcast, unicast and multicast connections. Particularly switching networks considered in the current invention include: generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, generalized hypercube networks $V_{cube}(N_1, N_2, d, s)$, and generalized cube connected cycles networks $V_{ccc}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general.

25

WO 2011/047368

PCT/US2010/052984

Efficient VLSI layout of networks on a semiconductor chip are very important and greatly influence many important design parameters such as the area taken up by the network on the chip, total number of wires, length of the wires, latency of the signals, capacitance and hence the maximum clock speed of operation. Some networks may not even be implemented practically on a chip due to the lack of efficient layouts. The different varieties of multi-stage networks described above have not been implemented previously on the semiconductor chips efficiently. For example in Field Programmable Gate Array (FPGA) designs, multi-stage networks described in the current invention have not been successfully implemented primarily due to the lack of efficient VLSI layouts.

Current commercial FPGA products such as Xilinx Vertex, Altera's Stratix implement island-style architecture using mesh and segmented mesh routing interconnects using either full crossbars or sparse crossbars. These routing interconnects consume large silicon area for crosspoints, long wires, large signal propagation delay and hence consume lot of power.

The current invention discloses the VLSI layouts of numerous types of multi-stage and pyramid networks which are very efficient and exploit spacial locality in the connectivity. Moreover they can be embedded on to mesh and segmented mesh routing interconnects of current commercial FPGA products. The VLSI layouts disclosed in the current invention are applicable to including the numerous generalized multi-stage networks disclosed in the following patent applications:

- 1) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-stage networks $V(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the US Application Serial No. 12/530,207 that is incorporated by reference above.
- 2) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the US Application Serial No. 12/601,273 that is incorporated by reference above.

WO 2011/047368

PCT/US2010/052984

3) Rearrangeably nonblocking for arbitrary fan-out multicast and unicast, and strictly nonblocking for unicast for generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the US Application Serial No. 12/601,274 that is incorporated by reference above.

4) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the US Application Serial No. 12/601,273 that is incorporated by reference above.

5) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the US Application Serial No. 12/601,274 that is incorporated by reference above.

6) Strictly nonblocking for arbitrary fan-out multicast and unicast for generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the US Application Serial No. 12/601,274 that is incorporated by reference above.

7) VLSI layouts of numerous types of multi-stage networks are described in the US Application Serial No. 12/601,275 entitled "VLSI LAYOUTS OF FULLY CONNECTED NETWORKS" that is incorporated by reference above.

In addition the layouts of the current invention are also applicable to generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks

WO 2011/047368

PCT/US2010/052984

$V_{m\text{link}-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage pyramid networks
 $V_{\text{fold}-m\text{link}-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks
 $V_{m\text{link}-bjp}(N_1, N_2, d, s)$, generalized hypercube networks $V_{\text{hcube}}(N_1, N_2, d, s)$ and
 generalized cube connected cycles networks $V_{\text{CCC}}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any
 5 number in general.

Symmetric RNB generalized multi-link multi-stage network $V_{m\text{link}}(N_1, N_2, d, s)$,

Connection Topology: Nearest Neighbor connectivity and with Full Bandwidth:

Referring to diagram 100A in FIG. 1A, in one embodiment, an exemplary
 10 generalized multi-link multi-stage network $V_{m\text{link}}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d =$
 2 ; and $s = 2$ with nine stages of one hundred and forty four switches for satisfying
 communication requests, such as setting up a telephone call or a data call, or a connection
 between configurable logic blocks, between an input stage 110 and output stage 120 via
 middle stages 130, 140, 150, 160, 170, 180 and 190 is shown where input stage 110
 15 consists of sixteen, two by four switches IS1-IS16 and output stage 120 consists of
 sixteen, four by two switches OS1-OS16. And all the middle stages namely the middle
 stage 130 consists of sixteen, four by four switches MS(1,1) - MS(1,16), middle stage
 140 consists of sixteen, four by four switches MS(2,1) - MS(2,16), middle stage 150
 consists of sixteen, four by four switches MS(3,1) - MS(3,16), middle stage 160 consists
 20 of sixteen, four by four switches MS(4,1) - MS(4,16), middle stage 170 consists of
 sixteen, four by four switches MS(5,1) - MS(5,16), middle stage 180 consists of sixteen,
 four by four switches MS(6,1) - MS(6,16), and middle stage 190 consists of sixteen, four
 by four switches MS(7,1) - MS(7,16).

As disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is
 25 incorporated by reference above, such a network can be operated in rearrangeably non-
 blocking manner for arbitrary fan-out multicast connections and also can be operated in
 strictly non-blocking manner for unicast connections.

WO 2011/047368

PCT/US2010/052984

In one embodiment of this network each of the input switches IS1-IS16 and output switches OS1-OS16 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable $\frac{N}{d}$, where N is the total number of inlet links or outlet links. The number of middle switches in each

5 middle stage is denoted by $\frac{N}{d}$. The size of each input switch IS1-IS16 can be denoted in general with the notation $d * 2d$ and each output switch OS1-OS16 can be denoted in general with the notation $2d * d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d * 2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a

10 network of switches. A symmetric multi-stage network can be represented with the notation $V_{m\text{link}}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL32), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

15 Each of the $\frac{N}{d}$ input switches IS1 – IS16 are connected to exactly d switches in middle stage 130 through two links each for a total of $2 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the middle links ML(1,1), ML(1,2), and also connected to middle switch MS(1,2) through the middle links ML(1,3) and ML(1,4)). The middle links which connect switches in the same row in two

20 successive middle stages are called hereinafter straight middle links; and the middle links which connect switches in different rows in two successive middle stages are called hereinafter cross middle links. For example, the middle links ML(1,1) and ML(1,2) connect input switch IS1 and middle switch MS(1,1), so middle links ML(1,1) and ML(1,2) are straight middle links; where as the middle links ML(1,3) and ML(1,4)

25 connect input switch IS1 and middle switch MS(1,2), since input switch IS1 and middle switch MS(1,2) belong to two different rows in diagram 100A of FIG. 1A, middle links ML(1,3) and ML(1,4) are cross middle links.

WO 2011/047368

PCT/US2010/052984

Each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,16) in the middle stage 130 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the middle links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the middle links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to exactly d switches in middle stage 140 through two links each for a total of $2 \times d$ links (for example the middle links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the middle links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

Each of the $\frac{N}{d}$ middle switches MS(2,1) – MS(2,16) in the middle stage 140 are connected from exactly d middle switches in middle stage 130 through two links each for a total of $2 \times d$ links (for example the middle links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from input switch MS(1,1), and the middle links ML(1,11) and ML(1,12) are connected to the middle switch MS(2,1) from input switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through two links each for a total of $2 \times d$ links (for example the middle links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the middle links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,6)).

Applicant notes that the topology of connections between middle switches MS(2,1) – MS(2,16) in the middle stage 140 and middle switches MS(3,1) – MS(3,16) in the middle stage 150 is not the typical inverse Benes topology but the connectivity of the generalized multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ 100A shown in FIG. 1A is effectively the same, or alternatively the network 100A shown in FIG. 1A is topologically equivalent to the network with inverse Benes network topology. However as will be described later in layouts of FIG. 1C – FIG. 1G, the length of the connection from a given inlet link to its destination outlet links may consist of different route resulting in different latency and different power dissipation for a given multicast or unicast assignment. As

WO 2011/047368

PCT/US2010/052984

will be described later in the layouts of FIG. 1C – FIG. 1G, the connection topology of middle links between middle stages 140 and 150 is in such a way that nearest neighbor blocks are connected directly and then the rest of the blocks are connected in inverse Benes topology.

5 Each of the $\frac{N}{d}$ middle switches MS(3,1) – MS(3,16) in the middle stage 150 are connected from exactly d middle switches in middle stage 140 through two links each for a total of $2 \times d$ links (for example the middle links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from input switch MS(2,1), and the middle links ML(2,23) and ML(2,24) are connected to the middle switch MS(3,1) from input switch
10 MS(2,6)) and also are connected to exactly d switches in middle stage 160 through two links each for a total of $2 \times d$ links (for example the middle links ML(4,1) and ML(4,2) are connected from middle switch MS(3,1) to middle switch MS(4,1), and the middle links ML(4,3) and ML(4,4) are connected from middle switch MS(3,1) to middle switch MS(4,11)).

15 Applicant notes that the topology of connections between middle switches MS(3,1) – MS(3,16) in the middle stage 150 and middle switches MS(4,1) – MS(4,16) in the middle stage 160 is not the typical inverse Benes topology but the connectivity of the generalized multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ 100A shown in FIG. 1A is effectively the same, or alternatively the network 100A shown in FIG. 1A is topologically
20 equivalent to the network with inverse Benes network topology. However as will be described later in layouts of FIG. 1C – FIG. 1G, the length of the connection from a given inlet link to its destination outlet links may consist of different route resulting in different latency and different power dissipation for a given multicast or unicast assignment. As will be described later in the layouts of FIG. 1C – FIG. 1G, the connection topology of
25 middle links between middle stages 150 and 160 is in such a way that nearest neighbor blocks are connected directly and then the rest of the blocks are connected in inverse Benes topology.

WO 2011/047368

PCT/US2010/052984

Each of the $\frac{N}{d}$ middle switches MS(4,1) – MS(4,16) in the middle stage 160 are connected from exactly d middle switches in middle stage 150 through two links each for a total of $2 \times d$ links (for example the middle links ML(4,1) and ML(4,2) are connected to the middle switch MS(4,1) from input switch MS(3,1), and the middle links ML(4,43) and ML(4,44) are connected to the middle switch MS(4,1) from input switch MS(3,11)) and also are connected to exactly d switches in middle stage 170 through two links each for a total of $2 \times d$ links (for example the middle links ML(5,1) and ML(5,2) are connected from middle switch MS(4,1) to middle switch MS(5,1), and the middle links ML(5,3) and ML(5,4) are connected from middle switch MS(4,1) to middle switch MS(5,11)).

Applicant notes that the topology of connections between middle switches MS(4,1) – MS(4,16) in the middle stage 160 and middle switches MS(5,1) – MS(5,16) in the middle stage 170 is not the typical inverse Benes topology but the connectivity of the generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ 100A shown in FIG. 1A is effectively the same or alternatively the network 100A shown in FIG. 1A is topologically equivalent to the network with inverse Benes network topology. However as will be described later in layouts of FIG. 1C – FIG. 1G, the length of the connection from a given inlet link to its destination outlet links may consist of different route resulting in different latency and different power dissipation for a given multicast or unicast assignment. As will be described later in the layouts of FIG. 1C – FIG. 1G, the connection topology of middle links between middle stages 160 and 170 is in such a way that nearest neighbor blocks are connected directly and then the rest of the blocks are connected in inverse Benes topology.

Each of the $\frac{N}{d}$ middle switches MS(5,1) – MS(5,16) in the middle stage 170 are connected from exactly d middle switches in middle stage 160 through two links each for a total of $2 \times d$ links (for example the middle links ML(5,1) and ML(5,2) are connected to the middle switch MS(5,1) from input switch MS(4,1), and the middle links ML(5,43) and ML(5,44) are connected to the middle switch MS(5,1) from input switch

WO 2011/047368

PCT/US2010/052984

MS(4,11)) and also are connected to exactly d switches in middle stage 180 through two links each for a total of $2 \times d$ links (for example the middle links ML(6,1) and ML(6,2) are connected from middle switch MS(5,1) to middle switch MS(6,1), and the middle links ML(6,3) and ML(6,4) are connected from middle switch MS(5,1) to middle switch
 5 MS(6,6)).

Applicant notes that the topology of connections between middle switches MS(5,1) – MS(5,16) in the middle stage 170 and middle switches MS(6,1) – MS(6,16) in the middle stage 180 is not the typical inverse Benes topology but the connectivity of the generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ 100A shown in FIG. 1A is
 10 effectively the same or alternatively the network 100A shown in FIG. 1A is topologically equivalent to the network with inverse Benes network topology. However as will be described later in layouts of FIG. 1C – FIG. 1G, the length of the connection from a given inlet link to its destination outlet links may consist of different route resulting in different latency and different power dissipation for a given multicast or unicast assignment. As
 15 will be described later in the layouts of FIG. 1C – FIG. 1G, the connection topology of middle links between middle stages 170 and 180 is in such a way that nearest neighbor blocks are connected directly and then the rest of the blocks are connected in inverse Benes topology.

Each of the $\frac{N}{d}$ middle switches MS(6,1) – MS(6,16) in the middle stage 180 are
 20 connected from exactly d middle switches in middle stage 170 through two links each for a total of $2 \times d$ links (for example the middle links ML(6,1) and ML(6,2) are connected to the middle switch MS(6,1) from input switch MS(5,1), and the middle links ML(6,23) and ML(6,24) are connected to the middle switch MS(6,1) from input switch MS(5,6)) and also are connected to exactly d switches in middle stage 190 through two
 25 links each for a total of $2 \times d$ links (for example the middle links ML(7,1) and ML(7,2) are connected from middle switch MS(6,1) to middle switch MS(7,1), and the middle links ML(7,3) and ML(7,4) are connected from middle switch MS(6,1) to middle switch MS(7,3)).

WO 2011/047368

PCT/US2010/052984

Each of the $\frac{N}{d}$ middle switches MS(7,1) – MS(7,16) in the middle stage 190 are connected from exactly d middle switches in middle stage 180 through two links each for a total of $2 \times d$ links (for example the middle links ML(7,1) and ML(7,2) are connected to the middle switch MS(7,1) from input switch MS(6,1), and the middle links ML(7,11) and ML(7,12) are connected to the middle switch MS(7,1) from input switch MS(6,3)) and also are connected to exactly d switches in middle stage 120 through two links each for a total of $2 \times d$ links (for example the middle links ML(8,1) and ML(8,2) are connected from middle switch MS(7,1) to middle switch MS(8,1), and the middle links ML(8,3) and ML(8,4) are connected from middle switch MS(7,1) to middle switch OS2).

Each of the $\frac{N}{d}$ middle switches OS1 – OS16 in the middle stage 120 are connected from exactly d middle switches in middle stage 190 through two links each for a total of $2 \times d$ links (for example the middle links ML(8,1) and ML(8,2) are connected to the output switch OS1 from input switch MS(7,1), and the middle links ML(8,7) and ML(8,8) are connected to the output switch OS1 from input switch MS(7,2)).

Finally the connection topology of the network 100A shown in FIG. 1A is logically similar to back to back inverse Benes connection topology with nearest neighbor connections between all the middle stages starting from middle stage 140 and middle stage 180.

Referring to diagram 100B in FIG. 1B, is a folded version of the multi-link multi-stage network 100A shown in FIG. 1A. The network 100B in FIG. 1B shows input stage 110 and output stage 120 are placed together. That is input switch IS1 and output switch OS1 are placed together, input switch IS2 and output switch OS2 are placed together, and similarly input switch IS16 and output switch OS16 are placed together. All the right going links {i.e., inlet links IL1 – IL32 and middle links ML(1,1) - ML(1,64)} correspond to input switches IS1 - IS16, and all the left going links {i.e., middle links ML(8,1) - ML(8,64) and outlet links OL1-OL32} correspond to output switches OS1 - OS16.

WO 2011/047368

PCT/US2010/052984

Middle stage 130 and middle stage 190 are placed together. That is middle switches MS(1,1) and MS(7,1) are placed together, middle switches MS(1,2) and MS(7,2) are placed together, and similarly middle switches MS(1,16) and MS(7,16) are placed together. All the right going middle links {i.e., middle links ML(1,1) - ML(1,64) and middle links ML(2,1) - ML(2,64)} correspond to middle switches MS(1,1) - MS(1,16), and all the left going middle links {i.e., middle links ML(7,1) - ML(7,64) and middle links ML(8,1) and ML(8,64)} correspond to middle switches MS(7,1) - MS(7,16).

Middle stage 140 and middle stage 180 are placed together. That is middle switches MS(2,1) and MS(6,1) are placed together, middle switches MS(2,2) and MS(6,2) are placed together, and similarly middle switches MS(2,16) and MS(6,16) are placed together. All the right going middle links {i.e., middle links ML(2,1) - ML(2,64) and middle links ML(3,1) - ML(3,64)} correspond to middle switches MS(2,1) - MS(2,16), and all the left going middle links {i.e., middle links ML(6,1) - ML(6,64) and middle links ML(7,1) and ML(7,64)} correspond to middle switches MS(6,1) - MS(6,16).

Middle stage 150 and middle stage 170 are placed together. That is middle switches MS(3,1) and MS(5,1) are placed together, middle switches MS(3,2) and MS(5,2) are placed together, and similarly middle switches MS(3,16) and MS(5,16) are placed together. All the right going middle links {i.e., middle links ML(3,1) - ML(3,64) and middle links ML(4,1) - ML(4,64)} correspond to middle switches MS(3,1) - MS(3,16), and all the left going middle links {i.e., middle links ML(5,1) - ML(5,64) and middle links ML(6,1) and ML(6,64)} correspond to middle switches MS(5,1) - MS(5,16).

Middle stage 160 is placed alone. All the right going middle links are the middle links ML(4,1) - ML(4,64) and all the left going middle links are middle links ML(5,1) - ML(5,64).

Just the same way as the connection topology of the network 100A shown in FIG. 1A, the connection topology of the network 100B shown in FIG. 1B is the folded version

WO 2011/047368

PCT/US2010/052984

and logically similar to back to back inverse Benes connection topology with nearest neighbor connections between all the middle stages starting from middle stage 140 and middle stage 180.

In one embodiment, in the network 100B of FIG. 1B, the switches that are placed
 5 together are implemented as separate switches then the network 100B is the generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch
 10 and a four by two switch respectively. For example the input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1) – ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and
 15 ML(8,8) being the inputs of the output switch OS1 and outlet links OL1 – OL2 being the outputs of the output switch OS1. Similarly in this embodiment of network 100B all the switches that are placed together in each middle stage are implemented as separate switches.

20 **Modified-Hypercube Topology layout scheme:**

Referring to layout 100C of FIG. 1C, in one embodiment, there are sixteen blocks namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, Block 23_24, Block 25_26, Block 27_28, Block 29_30, and Block 31_32. Each block implements all the
 25 switches in one row of the network 100B of FIG. 1B, one of the key aspects of the current invention. For example Block 1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are

WO 2011/047368

PCT/US2010/052984

denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; Middle switch MS(4,1) is denoted by switch 5.

5 All the straight middle links are illustrated in layout 100C of FIG. 1C. For example in Block 1_2, inlet links IL1 – IL2, outlet links OL1 – OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link
10 ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 100C of FIG. 1C.

Even though it is not illustrated in layout 100C of FIG. 1C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit depending on the applications in different embodiments. There are four
15 quadrants in the layout 100C of FIG. 1C namely top-left, bottom-left, top-right and bottom-right quadrants. Top-left quadrant implements Block 1_2, Block 3_4, Block 5_6, and Block 7_8. Bottom-left quadrant implements Block 9_10, Block 11_12, Block 13_14, and Block 15_16. Top-right quadrant implements Block 17_18, Block 19_20, Block 21_22, and Block 23_24. Bottom-right quadrant implements Block 25_26, Block
20 27_28, Block 29_30, and Block 31_32. There are two halves in layout 100C of FIG. 1C namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants. Right-half consists of top-right and bottom-right quadrants.

Recursively in each quadrant there are four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-
25 quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements Block 1_2. Bottom-left sub-quadrant of top-left quadrant implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 5_6. Finally bottom-right sub-quadrant of top-left quadrant implements Block 7_8. Similarly there are two sub-halves in each quadrant. For example in top-left quadrant
30 there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left

WO 2011/047368

PCT/US2010/052984

quadrant implements Block 1_2 and Block 3_4. Right-sub-half of top-left quadrant implements Block 5_6 and Block 7_8. Finally applicant notes that in each quadrant or half the blocks are arranged as a general binary hypercube. Recursively in larger multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 > 32$, the layout in this embodiment

5 in accordance with the current invention, will be such that the super-quadrants will also be arranged in d-ary hypercube manner. (In the embodiment of the layout 100C of FIG. 1C, it is binary hypercube manner since $d = 2$, in the network $V_{fold-link}(N_1, N_2, d, s)$ 100B of FIG. 1B).

Layout 100D of FIG. 1D illustrates the inter-block links between switches 1 and 2

10 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100D of FIG. 1D can be implemented as vertical tracks in one embodiment. Also

15 in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

20 The bandwidth provided between two physically adjacent blocks in the same column or same row, when a switch in the first block is connected to a switch in the second block through the corresponding inter-block links and also a second switch in the second block is connected to a second switch in the first block through the corresponding inter-block links, is hereinafter called 2's bandwidth or 2's BW. The bandwidth offered

25 between two diagonal blocks is also 2's BW when the corresponding row and columns provide 2's BW. For example the bandwidth provided between Block 1_2 and Block 3_4 of layout 100D of FIG. 1D is 2's BW because inter-block links between switch 1 of Block 1_2 and switch 2 of Block 3_4 are connected and also inter-block links between switch 2 of Block 1_2 and switch 1 of Block 3_4 are connected.

WO 2011/047368

PCT/US2010/052984

In general the bandwidth offered within a quadrant of the layout formed by two nearest neighboring blocks on each of the four sides is 2's BW. For example in layout 100C of FIG. 1C the bandwidth offered in top-left quadrant is 2's BW. Similarly the bandwidth offered within each of the other three quadrants bottom-left, top-right and bottom-right quadrants is 2' BW. Alternatively the bandwidth offered with in a square of blocks with the sides of the square consisting of two neighboring blocks is 2's BW. This definition can be generalized so that the bandwidth offered within a square of blocks with the sides consisting of "x" number of blocks, when $x = 2^y$ where y is an integer, is hereinafter x's BW. Hence the bandwidth offered between four neighboring quadrants is 4's BW. For example the bandwidth offered between top-left quadrant, bottom-left quadrant, top-right quadrant and bottom-right quadrant is 4's BW as will be described later. It must be noted that the 4's BW is the bandwidth offered between the four quadrants in a square of four quadrants and it is not the bandwidth offered with in each quadrant.

Layout 100E of FIG. 1E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 5_6. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 5_6. Applicant notes that the inter-block links illustrated in layout 100E of FIG. 1E can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

The bandwidth provided between Block 1_2 and Block 5_6 of layout 100E of FIG. 1E is 2's BW because inter-block links between switch 2 of Block 1_2 and switch 3 of Block 5_6 are connected and also inter-block links between switch 3 of Block 1_2 and switch 2 of Block 5_6 are connected. Similarly the bandwidth provided between Block 1_2 and Block 7_8 is also 2's BW since corresponding rows (formed by Block 1_2 and

WO 2011/047368

PCT/US2010/052984

Block 5_6; and by Block 3_4 and Block 7_8) and columns (formed by Block 1_2 and Block 3_4; and by Block 5_6 and Block 7_8) offer 2's BW. Similarly the bandwidth offered between Block 3_4 and Block 5_6 is 2's BW.

Layout 100F of FIG. 1F illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,23), and ML(6,24) are connected between switch 3 of Block 1_2 and switch 4 of Block 11_12. Similarly middle links ML(3,23), ML(3,24), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 11_12. Applicant notes that the inter-block links illustrated in layout 100F of FIG. 1F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,24) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,24) are implemented as a time division multiplexed single track).

Applicant notes that the topology of inter-block links between switches 3 and 4 of each block of layout 100F of FIG. 1F is not the typical inverse Benes Network topology. In layout 100F first the switches 3 and 4 of nearest neighbor blocks are connected and then the rest of the blocks are connected in inverse Benes Network topology. For example since Block 3_4 and Block 9_10 are nearest neighbors in the leftmost column of layout 100F the corresponding links from switches 3 and 4 are connected together first. Then the remaining blocks in each column are connected in inverse Benes topology. For example in layout 100F since the remaining block in the leftmost column of top-left quadrant is Block 1_2 and the remaining block in the leftmost column of bottom-left quadrant is Block 11_12 the inter-block links between their corresponding switches 3 and 4 are connected together. Similarly in all the columns, the inter-block links between switches 3 and 4 are connected.

The bandwidth offered in layout 100F of FIG. 1F is 4's BW, since the bandwidth offered with in a square of blocks with the sides of the square consisting of four neighboring blocks is 4's BW. It must be noted that the bandwidth offered between top-left quadrant and bottom-left quadrant is 4's BW. That is inter-block links of a switch in

WO 2011/047368

PCT/US2010/052984

each one of the blocks in top-left quadrant are connected to a switch in any one of the blocks in bottom-left quadrant and vice versa. Similarly the bandwidth offered between top-right quadrant and bottom-right quadrant is 4's BW. For example the bandwidth provided between Block 1_2 and Block 11_12 of layout 100F of FIG. 1F is 4's BW

5 because inter-block links between switch 3 of Block 1_2 and switch 4 of Block 11_12 are connected and also inter-block links between switch 4 of Block 1_2 and switch 3 of Block 11_12 are connected. Similarly the bandwidth provided between Block 3_4 and Block 9_10 of layout 100F of FIG. 1F is 4's BW, even though they are physically nearest neighbors. It must be noted that the 4's BW is the bandwidth offered between the four

10 quadrants in a square of four quadrants and it is not the bandwidth offered with in each quadrant.

Layout 100G of FIG. 1G illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,43), and ML(5,44) are connected between switch 4 of Block 1_2 and switch 5 of Block 21_22. Similarly middle

15 links ML(4,43), ML(4,44), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 21_22. Applicant notes that the inter-block links illustrated in layout 100G of FIG. 1G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,44) are implemented as two different

20 tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(4,4) and ML(5,44) are implemented as a time division multiplexed single track).

Applicant notes that the topology of inter-block links between switches 4 and 5 of each block of layout 100G of FIG. 1G is not the typical inverse Benes Network topology.

25 In layout 100G first the switches 4 and 5 of nearest neighbor blocks are connected and then the rest of the blocks are connected in inverse Benes Network topology. For example since Block 5_6 and Block 17_18 are nearest neighbors in the topmost row of layout 100G the corresponding links from switches 4 and 5 are connected together first. Then the remaining blocks in each row are connected in inverse Benes topology. For example

30 in layout 100G since the remaining block in the topmost row of top-left quadrant is Block

WO 2011/047368

PCT/US2010/052984

1_2 and the remaining block in the topmost row of top-right quadrant is Block 21_22 the inter-block links between their corresponding switches 4 and 5 are connected together. Similarly in all the rows, the inter-block links between switches 4 and 5 are connected.

The bandwidth offered in layout 100G of FIG. 1G is 4's BW, since the bandwidth offered with in a square of blocks with the sides of the square consisting of four neighboring blocks is 4's BW. It must be noted that the bandwidth offered between top-left quadrant and top-right quadrant is 4's BW. That is inter-block links of a switch in each one of the blocks in top-left quadrant are connected to a switch in any one of the blocks in top-right quadrant and vice versa. Similarly the bandwidth offered between bottom-left quadrant and bottom-right quadrant is 4's BW. For example the bandwidth provided between Block 1_2 and Block 21_22 of layout 100G of FIG. 1G is 4's BW because inter-block links between switch 4 of Block 1_2 and switch 5 of Block 21_22 are connected and also inter-block links between switch 5 of Block 1_2 and switch 4 of Block 21_22 are connected. Similarly the bandwidth provided between Block 5_6 and Block 17_18 of layout 100G of FIG. 1G is 4's BW, even though they are physically nearest neighbors. Just the same way 2's BW is provided between two diagonal blocks, the bandwidth offered between two diagonal quadrants is also 4's BW that is when the corresponding row and columns provide 4's BW.

The complete layout for the network 100B of FIG. 1B is given by combining the links in layout diagrams of 100C, 100D, 100E, 100F, and 100G. Applicant notes that in the layout 100C of FIG. 1C, the inter-block links between switch 1 and switch 2 of corresponding blocks are vertical tracks as shown in layout 100D of FIG. 1D; the inter-block links between switch 2 and switch 3 of corresponding blocks are horizontal tracks as shown in layout 100E of FIG. 1E; the inter-block links between switch 3 and switch 4 of corresponding blocks are vertical tracks as shown in layout 100F of FIG. 1F; and finally the inter-block links between switch 4 and switch 5 of corresponding blocks are horizontal tracks as shown in layout 100G of FIG. 1G. The pattern is alternate vertical tracks and horizontal tracks. It continues recursively for larger networks of $N > 32$ as will be illustrated later.

WO 2011/047368

PCT/US2010/052984

Some of the key aspects of the current invention are discussed. 1) All the switches in one row of the multi-stage network 100B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks or vertical tracks; 3) Since all the inter-block links are either horizontal or vertical tracks, all the inter-block links can be mapped on to island-style architectures in current commercial FPGA's; 4) The length of the wires in a given stage are not equal, for example the inter-block links between switches 3 and 4 of the nearest neighbor blocks Block 3_4 and Block 9_10 are smaller in length than the inter-block links between switches 3 and 4 of the blocks Block 1_2 and Block 11_12.

In accordance with the current invention, the layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ the sub-quadrants, quadrants, and super-quadrants are arranged in d-ary hypercube manner and also the inter-blocks are accordingly connected in d-ary hypercube topology. Even though all the embodiments in the current invention are illustrated for $N_1 = N_2$, the embodiments can be extended for $N_1 \neq N_2$.

Referring to layout 100H of FIG. 1H, illustrates the extension of layout 100C for the network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 128$; $d = 2$; and $s = 2$. There are four super-quadrants in layout 100H namely top-left super-quadrant, bottom-left super-quadrant, top-right super-quadrant, bottom-right super-quadrant. Total number of blocks in the layout 100H is sixty four. Top-left super-quadrant implements the blocks from block 1_2 to block 31_32. Each block in all the super-quadrants has two more switches namely switch 6 and switch 7 in addition to the switches [1-5] illustrated in layout 100C of FIG. 1C. The inter-block link connection topology is the exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as it is shown in the layouts of FIG. 1D, FIG. 1E, FIG. 1F, and FIG. 1G respectively.

Bottom-left super-quadrant implements the blocks from block 33_34 to block 63_64. Top-right super-quadrant implements the blocks from block 65_66 to block 95_96. And bottom-right super-quadrant implements the blocks from block 97_98 to block 127_128. In all these three super-quadrants also, the inter-block link connection

WO 2011/047368

PCT/US2010/052984

topology is exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as that of the top-left super-quadrant.

5 Recursively in accordance with the current invention, the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-left super-quadrant and bottom-left super-quadrant. And similarly the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-right super-quadrant and bottom-right super-quadrant. The inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of top-left super-quadrant and top-right super-quadrant. And
10 similarly the inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of bottom-left super-quadrant and bottom-right super-quadrant.

Just as described for layout 100F of FIG. 1F, Applicant notes that the connection topology of inter-block links between switches 5 and 6 of each block of layout 100H of
15 FIG. 1H is not the typical inverse Benes Network topology. In layout 100H first the switches 5 and 6 of nearest neighbor blocks are connected and then the rest of the blocks are connected in inverse Benes Network topology. For example since Block 11_12 and Block 33_34 are nearest neighbors in the leftmost column of layout 100H the corresponding inter-block links from switches 5 and 6 are connected together first. Then
20 the remaining blocks in the leftmost column are connected in inverse Benes topology. For example in layout 100H since the remaining blocks in the leftmost column of top-left super-quadrant are Block 1_2, Block 3_4, and Block 9_10 and the remaining blocks in the leftmost column of bottom-left super-quadrant are Block 35_36, Block 41_42 and Block 43_44 the inter-block links between their corresponding switches 5 and 6 are
25 connected together. In one embodiment the inter-block links of switches 5 and 6 corresponding to Block 1_2 and Block 35-36 are connected together; the inter-block links of switches 5 and 6 corresponding to Block 3_4 and Block 41_42 are connected together; and the inter-block links of switches 5 and 6 corresponding to Block 9_10 and Block 43_44 are connected together. (Similarly in another embodiment any one of the three
30 blocks in the leftmost column of top-left super-quadrant can be connected with any one of

WO 2011/047368

PCT/US2010/052984

the three blocks in the leftmost column of bottom-left super-quadrant of course as long as each block in leftmost column of top-left super-quadrant is connected to only one block in leftmost column of bottom-left super-quadrant and vice versa). Similarly in all the columns, the inter-block links between switches 5 and 6 are connected.

5 The bandwidth offered between top super-quadrants and bottom super-quadrants in layout 100H of FIG. 1H is 8's BW, since the bandwidth offered within a square of blocks with the sides of the square consisting of eight neighboring blocks is 8's BW. It must be noted that the bandwidth offered between top-left super-quadrant and bottom-left
10 super-quadrant is 8's BW. That is inter-block links of a switch in each one of the blocks in top-left super-quadrant are connected to a switch in any one of the blocks in bottom-left super-quadrant and vice versa. Similarly the bandwidth offered between top-right super-quadrant and bottom-right super-quadrant is 8's BW. For example in one embodiment the bandwidth provided between Block 1_2 and Block 35_36 of layout
15 100H of FIG. 1H is 8's BW because inter-block links between switch 5 of Block 1_2 and switch 6 of Block 35_36 are connected and also inter-block links between switch 5 of Block 1_2 and switch 6 of Block 35_36 are connected. Similarly the bandwidth provided between any one of the blocks in top-left super-quadrant and any one of the bottom-left super-quadrant of layout 100H of FIG. 1H is 8's BW. It must be noted that the 8's BW is the bandwidth offered between the four super-quadrants in a square of four super-
20 quadrants and it is neither the bandwidth offered between the four quadrants in one of the super-quadrants or within each quadrant.

Just as described for layout 100G of FIG. 1G, Applicant notes that the connection topology of inter-block links between switches 6 and 7 of each block of layout 100H of
25 FIG. 1H is not the typical inverse Benes Network topology. In layout 100H first the switches 6 and 7 of nearest neighbor blocks are connected and then the rest of the blocks are connected in inverse Benes Network topology. For example since Block 21_22 and Block 65_66 are nearest neighbors in the topmost row of layout 100H the corresponding inter-block links from switches 6 and 7 are connected together first. Then the remaining blocks in the topmost row are connected in inverse Benes topology. For example in
30 layout 100H since the remaining blocks in the topmost row of top-left super-quadrant are

WO 2011/047368

PCT/US2010/052984

Block 1_2, Block 5_6, and Block 17_18 and the remaining blocks in the topmost row of top-right super-quadrant are Block 69_70, Block 81_82 and Block 85_86 the inter-block links between their corresponding switches 6 and 7 are connected together. In one embodiment the inter-block links of switches 6 and 7 corresponding to Block 1_2 and Block 69-70 are connected together; the inter-block links of switches 6 and 7 corresponding to Block 5_6 and Block 81-82 are connected together; and the inter-block links of switches 6 and 7 corresponding to Block 17_18 and Block 85-86 are connected together. (Similarly in another embodiment any one of the three blocks in the topmost row of top-left super-quadrant can be connected with any one of the three blocks in the topmost row of top-right super-quadrant of course as long as each block in topmost row of top-right super-quadrant is connected to only one block in topmost row of top-right super-quadrant and vice versa). Similarly in all the rows, the inter-block links between switches 6 and 7 are connected.

The bandwidth offered between left super-quadrants and right super-quadrants in layout 100H of FIG. 1H is 8's BW, since the bandwidth offered with in a square of blocks with the sides of the square consisting of eight neighboring blocks is 8's BW. It must be noted that the bandwidth offered between top-left super-quadrant and top-right super-quadrant is 8's BW. That is inter-block links of a switch in each one of the blocks in top-left super-quadrant are connected to a switch in any one of the blocks in top-right super-quadrant and vice versa. Similarly the bandwidth offered between bottom-left super-quadrant and bottom-right super-quadrant is 8's BW. For example in one embodiment the bandwidth provided between Block 1_2 and Block 69_70 of layout 100H of FIG. 1H is 8's BW because inter-block links between switch 6 of Block 1_2 and switch 7 of Block 69_70 are connected and also inter-block links between switch 6 of Block 1_2 and switch 7 of Block 69_70 are connected. Similarly the bandwidth provided between any one of the blocks in top-left super-quadrant and any one of the blocks in top-right super-quadrant of layout 100H of FIG. 1H is 8's BW. Just the same way 2's BW is provided between two diagonal blocks, the bandwidth offered between two diagonal super-quadrants is 8's BW that is when the corresponding row and columns provide 8's BW.

WO 2011/047368

PCT/US2010/052984

Referring to diagram 100I of FIG. 1I illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized folded multi-link multi-stage network

$V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 100I illustrates

5 both the intra-block and inter-block links connected to Block 1_2. The layout diagram 100I corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized folded multi-link multi-stage network

$V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed

10 in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 1I are namely input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the

15 switches implemented are input switch IS1 and output switch OS1); middle switch MS(1,1) and middle switch MS(7,1) belonging to switch 2; middle switch MS(2,1) and middle switch MS(6,1) belonging to switch 3; middle switch MS(3,1) and middle switch MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1) – ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two

20 switch with the middle links ML(8,1), ML(8,2), ML(8,7), and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1 – OL2 being the outputs of the output switch OS1.

25 Middle switch MS(1,1) is implemented as four by four switch with the middle links ML(1,1), ML(1,2), ML(1,7) and ML(1,8) being the inputs and middle links ML(2,1) – ML(2,4) being the outputs; and middle switch MS(7,1) is implemented as four by four switch with the middle links ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs

WO 2011/047368

PCT/US2010/052984

and middle links ML(8,1) – ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 100I of FIG. 1I.

Generalized Multi-link Butterfly Fat Tree Network Embodiment:

In another embodiment in the network 100B of FIG. 1B, the switches that are placed together are implemented as combined switch then the network 100B is the generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,390 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch. For example the input switch IS1 and output switch OS1 are placed together; so input switch IS1 and output OS1 are implemented as a six by six switch with the inlet links IL1, IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and OL2 being the outputs of the combined switch IS1&OS1. Similarly in this embodiment of network 100B all the switches that are placed together are implemented as a combined switch.

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$.

Referring to diagram 100J of FIG. 1J illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 100J illustrates both the intra-block and inter-block links. The layout diagram 100J corresponds to the

WO 2011/047368

PCT/US2010/052984

embodiment where the switches that are placed together are implemented as combined switch in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application 5 Serial No. 60/940,390 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 1J are namely the combined input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switch implemented is combined input and output switch IS1&OS1); middle switch 10 MS(1,1) belonging to switch 2; middle switch MS(2,1) belonging to switch 3; middle switch MS(3,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links IL1, IL2 and ML(8,1), ML(8,2), ML(8,7), and ML(8,8) being the inputs and middle links ML(1,1) – ML(1,4), and outlet links OL1 – OL2 being the 15 outputs.

Middle switch MS(1,1) is implemented as eight by eight switch with the middle links ML(1,1), ML(1,2), ML(1,7), ML(1,8), ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs and middle links ML(2,1) – ML(2,4) and middle links ML(8,1) – ML(8,4) being the outputs. Similarly all the other middle switches are also implemented 20 as eight by eight switches as illustrated in 100J of FIG. 1J.

In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of $V_{mlink-bft}(N_1, N_2, d, s)$ can be implemented as a four by eight switch and a four by four switch to save cross points. This is because the left going middle links of these middle 25 switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1J, the left going middle links namely ML(7,1), ML(7,2), ML(7,11), and ML(7,12) are never switched to the right going middle links ML(2,1), ML(2,2), ML(2,3), and ML(2,4). And hence to implement MS(1,1) two switches namely: 1) a four by eight switch with the middle links ML(1,1), ML(1,2),

WO 2011/047368

PCT/US2010/052984

ML(1,7), and ML(1,8) as inputs and the middle links ML(2,1), ML(2,2), ML(2,3), ML(2,4), ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs and 2) a four by four switch with the middle links ML(7,1), ML(7,2), ML(7,11), and ML(7,12) as inputs and the middle links ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs are sufficient
 5 without losing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

Generalized multi-stage network Embodiment:

In one embodiment, in the network 100B of FIG. 1B, the switches that are placed together are implemented as two separate switches in input stage 110 and output stage
 10 120; and as four separate switches in all the middle stages, then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch
 15 and a four by two switch respectively. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs and
 20 outlet links OL1 – OL2 being the outputs.

The switches, corresponding to the middle stages that are placed together are implemented as four two by two switches. For example middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the
 25 inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle
 30 switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and

WO 2011/047368

PCT/US2010/052984

ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs; Similarly in this embodiment of network 100B all the switches that are placed together are implemented as separate switches.

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in
 5 FIG. 1G are also applicable to generalized folded multi-stage network
 $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages. The layout 100C
 in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-
 stage network $V_{fold}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable
 to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$.

10 Referring to diagram 100K of FIG. 1K illustrates a high-level implementation of
 Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of
 FIG. 1C which represents a generalized folded multi-stage network
 $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 100K illustrates
 both the intra-block and inter-block links. The layout diagram 100K corresponds to the
 15 embodiment where the switches that are placed together are implemented as separate
 switches in the network 100B of FIG. 1B. As noted before then the network 100B is the
 generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$;
 and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No.
 60/940,391 that is incorporated by reference above.

20 That is the switches that are placed together in Block 1_2 as shown in FIG. 1K are
 namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by
 dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the
 switches implemented are input switch IS1 and output switch OS1); middle switches
 MS(1,1), MS(1,17), MS(7,1) and MS(7,17) belonging to switch 2; middle switches
 25 MS(2,1), MS(2,17), MS(6,1) and MS(6,17) belonging to switch 3; middle switches
 MS(3,1), MS(3,17), MS(5,1) and MS(5,17) belonging to switch 4; And middle switches
 MS(4,1), and MS(4,17) belonging to switch 5.

WO 2011/047368

PCT/US2010/052984

Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and
 5 ML(8,8) being the inputs and outlet links OL1 – OL2 being the outputs.

Middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle
 10 links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2)
 15 and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as illustrated in 100K of FIG. 1K.

Generalized multi-stage network Embodiment with S = 1:

In one embodiment, in the network 100B of FIG. 1B (where it is implemented with $s = 1$), the switches that are placed together are implemented as two separate
 20 switches in input stage 110 and output stage 120; and as two separate switches in all the middle stages, then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and
 25 output stage 120 are implemented as two, two by two switches. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the
 30 inputs and outlet links OL1 – OL2 being the outputs.

WO 2011/047368

PCT/US2010/052984

The switches, corresponding to the middle stages that are placed together are implemented as two, two by two switches. For example middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links
 5 ML(2,1) and ML(2,2) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs; Similarly in this embodiment of network 100B all the switches that are placed together are implemented as two separate switches.

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in
 10 FIG. 1G are also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$.

15 Referring to diagram 100K1 of FIG. 1K1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 100C of FIG. 1C when $s = 1$ which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ (All the double links are replaced by single links when $s = 1$). Block 1_2 in 100K1 illustrates both the intra-block and inter-
 20 block links. The layout diagram 100K1 corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 100B of FIG. 1B when $s = 1$. As noted before then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that
 25 is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 1K1 are namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches

WO 2011/047368

PCT/US2010/052984

MS(1,1) and MS(7,1) belonging to switch 2; middle switches MS(2,1) and MS(6,1) belonging to switch 3; middle switches MS(3,1) and MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is
 5 implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1 – OL2 being the outputs.

Middle switches MS(1,1) and MS(7,1) are placed together; so middle switch
 10 MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; And middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as
 15 illustrated in 100K1 of FIG. 1K1.

Generalized Butterfly Fat Tree Network Embodiment:

In another embodiment in the network 100B of FIG. 1B, the switches that are placed together are implemented as two combined switches then the network 100B is the generalized butterfly fat tree network $V_{bf}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and s
 20 $= 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch. For example the input switch IS1 and output switch OS1 are placed together; so input output switch IS1&OS1 are implemented as a six by six switch with the inlet links IL1,
 25 IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and OL2 being the outputs of the combined switch IS1&OS1.

WO 2011/047368

PCT/US2010/052984

The switches, corresponding to the middle stages that are placed together are implemented as two four by four switches. For example middle switches MS(1,1) and MS(1,17) are placed together; so middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; middle switch MS(1,17) is implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly in this embodiment of network 100B all the switches that are placed together are implemented as a two combined switches.

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$.

Referring to diagram 100L of FIG. 1L illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 100L illustrates both the intra-block and inter-block links. The layout diagram 100L corresponds to the embodiment where the switches that are placed together are implemented as two combined switches in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 1L are namely the combined input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switch implemented is combined input and output switch IS1&OS1); middle switch

WO 2011/047368

PCT/US2010/052984

MS(1,1) and MS(1,17) belonging to switch 2; middle switch MS(2,1) and MS(2,17) belonging to switch 3; middle switch MS(3,1) and MS(3,17) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

5 Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links IL1, IL2, ML(8,1) , ML(8,2), ML(8,7) and ML(8,8) being the inputs and middle links ML(1,1) – ML(1,4) and outlet links OL1 – OL2 being the outputs.

Middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; And middle switch MS(1,17) is
 10 implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two four by four switches as illustrated in 100L of FIG. 1L.

In another embodiment, middle switch MS(1,1) (or the middle switches in any of
 15 the middle stage excepting the root middle stage) of Block 1_2 of $V_{\text{mlink-bjt}}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1L, the left going middle links namely ML(7,1) and
 20 ML(7,11) are never switched to the right going middle links ML(2,1) and ML(2,3). And hence to implement MS(1,1) two switches namely: 1) a two by four switch with the middle links ML(1,1) and ML(1,7) as inputs and the middle links ML(2,1), ML(2,3), ML(8,1), and ML(8,3) as outputs and 2) a two by two switch with the middle links ML(7,1) and ML(7,11) as inputs and the middle links ML(8,1) and ML(8,3) as outputs
 25 are sufficient without loosing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

WO 2011/047368

PCT/US2010/052984

Generalized Butterfly Fat Tree Network Embodiment with S = 1:

In one embodiment, in the network 100B of FIG. 1B (where it is implemented with $s = 1$), the switches that are placed together are implemented as a combined switch in input stage 110 and output stage 120; and as a combined switch in all the middle stages, then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a four by four switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input and output switch IS1&OS1 is implemented as four by four switch with the inlet links IL1, IL2, ML(8,1) and ML(8,3) being the inputs and middle links ML(1,1) – ML(1,2) and outlet links OL1 – OL2 being the outputs

The switches, corresponding to the middle stages that are placed together are implemented as a four by four switch. For example middle switches MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,3), ML(7,1) and ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs..

Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with five stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$.

Referring to diagram 100L1 of FIG. 1L1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 100C of FIG. 1C when $s = 1$ which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ (All the double links are replaced

WO 2011/047368

PCT/US2010/052984

by single links when $s = 1$). Block 1_2 in 100K1 illustrates both the intra-block and inter-block links. The layout diagram 100L1 corresponds to the embodiment where the switches that are placed together are implemented as a combined switch in the network 100B of FIG. 1B when $s = 1$. As noted before then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with
 5 butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 1L1 are namely the input and output switch IS1&OS1 belonging to switch 1, illustrated by
 10 dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switch MS(1,1) belonging to switch 2; middle switch MS(2,1) belonging to switch 3; middle switch MS(3,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Input and output switch IS1&OS1 are placed together; so input and output switch
 15 IS1&OS1 is implemented as four by four switch with the inlet links IL1, IL2, ML(8,1) and ML(8,3) being the inputs and middle links ML(1,1) – ML(1,2) and outlet links OL1 – OL2 being the outputs.

Middle switch MS(1,1) is implemented as four by four switch with middle links
 20 ML(1,1), ML(1,3), ML(7,1) and ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 100L1 of FIG. 1L1.

In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of
 $V_{mlink-bft}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two
 25 switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1L1, the left going middle links namely ML(7,1) and ML(7,5) are never switched to the right going middle links ML(2,1) and ML(2,2). And hence to implement MS(1,1) two switches namely: 1) a two by four switch with the

WO 2011/047368

PCT/US2010/052984

middle links ML(1,1) and ML(1,3) as inputs and the middle links ML(2,1), ML(2,2), ML(8,1), and ML(8,2) as outputs and 2) a two by two switch with the middle links ML(7,1) and ML(7,5) as inputs and the middle links ML(8,1) and ML(8,2) as outputs are sufficient without losing any connectivity of the embodiment of MS(1,1) being
 5 implemented as an eight by eight switch as described before.)

Symmetric RNB generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$,

Connection Topology with $N_1 \neq 2^x$ & $N_2 \neq 2^y$ where x and y are integers:

Referring to diagram 200A in FIG. 2A, in one embodiment, an exemplary
 10 generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 24$ and
 $2^4 < N = 24 < 2^5$; $d = 2$; and $s = 2$ with nine stages of ninety two switches for satisfying
 communication requests, such as setting up a telephone call or a data call, or a connection
 between configurable logic blocks, between an input stage 110 and output stage 120 via
 middle stages 130, 140, 150, 160, 170, 180 and 190 is shown where input stage 110
 15 consists of twelve, two by four switches IS1-IS12 and output stage 120 consists of
 twelve, four by two switches OS1-OS12. And the middle stages namely the middle stage
 130 consists of twelve, four by four switches MS(1,1) - MS(1,12), middle stage 140
 consists of eight, four by four switches MS(2,1) - MS(2,8), middle stage 180 consists of
 eight, four by four switches MS(6,1) - MS(6,8), and middle stage 190 consists of twelve,
 20 four by four switches MS(7,1) - MS(7,12); middle stage 150 consists of twelve, four by
 four switches MS(3,1) - MS(3,12), middle stage 160 consists of eight, four by four
 switches MS(4,1) - MS(4,2), MS(4,5) - MS(4,6), MS(4,9) - MS(4,12), middle stage 170
 consists of eight, four by four switches MS(5,1) - MS(5,2), MS(5,5) - MS(5,6), MS(5,9)
 - MS(5,12).

25 Such a generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where
 $N_1 \neq 2^x$ & $N_2 \neq 2^y$ where x and y are integers, can be operated in rearrangeably non-
 blocking manner for arbitrary fan-out multicast connections and also can be operated in
 strictly non-blocking manner for unicast connections, just the same way as when

WO 2011/047368

PCT/US2010/052984

$N_1 = 2^x$ & $N_2 = 2^y$ where x and y are integers, as disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above.

In one embodiment of this network each of the input switches IS1-IS12 and output switches OS1-OS12 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable $\frac{N}{d}$, where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by a maximum of $\frac{N}{d}$. The size of each input switch IS1-IS12 can be denoted in general with the notation $d * 2d$ and each output switch OS1-OS12 can be denoted in general with the notation $2d * d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d * 2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-stage network can be represented with the notation $V_{link}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL32), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

Each of the $\frac{N}{d}$ input switches IS1 – IS12 are connected to exactly d switches in middle stage 130 through two links each for a total of $2 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the middle links ML(1,1), ML(1,2), and also connected to middle switch MS(1,2) through the middle links ML(1,3) and ML(1,4)). Just the same way as defined before, the middle links which connect switches in the same row in two successive middle stages are called hereinafter straight middle links; and the middle links which connect switches in different rows in two successive middle stages are called hereinafter cross middle links. For example, the middle links ML(1,1) and ML(1,2) connect input switch IS1 and middle switch MS(1,1), so middle links ML(1,1) and ML(1,2) are straight middle links; where as the middle links ML(1,3) and ML(1,4) connect input switch IS1 and middle switch MS(1,2), since input

WO 2011/047368

PCT/US2010/052984

switch IS1 and middle switch MS(1,2) belong to two different rows in diagram 100A of FIG. 1A, middle links ML(1,3) and ML(1,4) are cross middle links.

Each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,12) in the middle stage 130 are connected from exactly d input switches through two links each for a total of $2 \times d$ links (for example the middle links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the middle links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2). Each of the middle switches MS(1,1) – MS(1,8) are connected to exactly d switches in middle stage 140 through two links each for a total of $2 \times d$ links (for example the middle links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the middle links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)); and each of the middle switches MS(1,9) – MS(1,12) are connected to exactly d switches in middle stage 150 through two links each for a total of $2 \times d$ links (for example the middle links ML(3,33) and ML(3,34) are connected from middle switch MS(1,9) to middle switch MS(3,9), and the middle links ML(3,35) and ML(3,36) are connected from middle switch MS(1,9) to middle switch MS(3,11)).

Each of the middle switches MS(2,1) – MS(2,8) in the middle stage 140 are connected from exactly d middle switches in middle stage 130 through two links each for a total of $2 \times d$ links (for example the middle links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from input switch MS(1,1), and the middle links ML(1,11) and ML(1,12) are connected to the middle switch MS(2,1) from input switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through two links each for a total of $2 \times d$ links (for example the middle links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the middle links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,5)).

Each of the $\frac{N}{d}$ middle switches MS(3,1) – MS(3,12) in the middle stage 150 are connected from exactly d middle switches in middle stage 140 through two links each

WO 2011/047368

PCT/US2010/052984

for a total of $2 \times d$ links (for example the middle links $ML(3,1)$ and $ML(3,2)$ are connected to the middle switch $MS(3,1)$ from input switch $MS(2,1)$, and the middle links $ML(2,19)$ and $ML(2,20)$ are connected to the middle switch $MS(3,1)$ from input switch $MS(2,5)$). Each of the middle switches $MS(3,1) - MS(3,2)$, $MS(3,5) - MS(3,6)$ and
5 $MS(3,9) - MS(3,12)$ are connected to exactly d switches in middle stage 160 through two links each for a total of $2 \times d$ links (for example the middle links $ML(4,1)$ and $ML(4,2)$ are connected from middle switch $MS(3,1)$ to middle switch $MS(4,1)$, and the middle links $ML(4,3)$ and $ML(4,4)$ are connected from middle switch $MS(3,1)$ to middle switch $MS(4,9)$); and each of the middle switches $MS(3,3) - MS(3,4)$ and $MS(3,7) -$
10 $MS(3,8)$ are connected to exactly d switches in middle stage 180 through two links each for a total of $2 \times d$ links (for example the middle links $ML(6,9)$ and $ML(6,10)$ are connected from middle switch $MS(3,3)$ to middle switch $MS(6,3)$, and the middle links $ML(6,11)$ and $ML(6,12)$ are connected from middle switch $MS(3,3)$ to middle switch $MS(6,7)$).

15 Each of the middle switches $MS(4,1) - MS(4,2)$, $MS(4,5) - MS(4,6)$ and $MS(4,9) - MS(4,12)$ in the middle stage 160 are connected from exactly d middle switches in middle stage 150 through two links each for a total of $2 \times d$ links (for example the middle links $ML(4,1)$ and $ML(4,2)$ are connected to the middle switch $MS(4,1)$ from input switch $MS(3,1)$, and the middle links $ML(4,35)$ and $ML(4,36)$ are connected to the
20 middle switch $MS(4,1)$ from input switch $MS(3,9)$) and also are connected to exactly d switches in middle stage 170 through two links each for a total of $2 \times d$ links (for example the middle links $ML(5,1)$ and $ML(5,2)$ are connected from middle switch $MS(4,1)$ to middle switch $MS(5,1)$, and the middle links $ML(5,3)$ and $ML(5,4)$ are connected from middle switch $MS(4,1)$ to middle switch $MS(5,9)$).

25 Each of the middle switches $MS(5,1) - MS(5,2)$, $MS(5,5) - MS(5,6)$ and $MS(5,9) - MS(5,12)$ in the middle stage 170 are connected from exactly d middle switches in middle stage 160 through two links each for a total of $2 \times d$ links (for example the middle links $ML(5,1)$ and $ML(5,2)$ are connected to the middle switch $MS(5,1)$ from input switch $MS(4,1)$, and the middle links $ML(5,35)$ and $ML(5,36)$ are connected to the
30 middle switch $MS(5,1)$ from input switch $MS(4,9)$). Each of the middle switches $MS(5,1)$

WO 2011/047368

PCT/US2010/052984

– MS(5,2), MS(5,5) – MS(5,6) are connected to exactly d switches in middle stage 180 through two links each for a total of $2 \times d$ links (for example the middle links ML(6,1) and ML(6,2) are connected from middle switch MS(5,1) to middle switch MS(6,1), and the middle links ML(6,3) and ML(6,4) are connected from middle switch MS(5,1) to middle switch MS(6,5)); and Each of the middle switches MS(5,9) – MS(5,12) are connected to exactly d switches in middle stage 190 through two links each for a total of $2 \times d$ links (for example the middle links ML(6,33) and ML(6,34) are connected from middle switch MS(5,9) to middle switch MS(7,9), and the middle links ML(6,35) and ML(6,36) are connected from middle switch MS(5,9) to middle switch MS(7,11)).

10 Each of the $\frac{N}{d}$ middle switches MS(6,1) – MS(6,8) in the middle stage 180 are connected from exactly d middle switches in middle stage 170 through two links each for a total of $2 \times d$ links (for example the middle links ML(6,1) and ML(6,2) are connected to the middle switch MS(6,1) from input switch MS(5,1), and the middle links ML(6,19) and ML(6,20) are connected to the middle switch MS(6,1) from input switch MS(5,5)) and also are connected to exactly d switches in middle stage 190 through two links each for a total of $2 \times d$ links (for example the middle links ML(7,1) and ML(7,2) are connected from middle switch MS(6,1) to middle switch MS(7,1), and the middle links ML(7,3) and ML(7,4) are connected from middle switch MS(6,1) to middle switch MS(7,3)).

20 Each of the $\frac{N}{d}$ middle switches MS(7,1) – MS(7,12) in the middle stage 190 are connected from exactly d middle switches in middle stage 180 through two links each for a total of $2 \times d$ links (for example the middle links ML(7,1) and ML(7,2) are connected to the middle switch MS(7,1) from input switch MS(6,1), and the middle links ML(7,11) and ML(7,12) are connected to the middle switch MS(7,1) from input switch MS(6,3)) and also are connected to exactly d switches in middle stage 120 through two links each for a total of $2 \times d$ links (for example the middle links ML(8,1) and ML(8,2) are connected from middle switch MS(7,1) to middle switch MS(8,1), and the middle links ML(8,3) and ML(8,4) are connected from middle switch MS(7,1) to middle switch OS2).

WO 2011/047368

PCT/US2010/052984

Each of the $\frac{N}{d}$ middle switches OS1 – OS12 in the middle stage 120 are connected from exactly d middle switches in middle stage 190 through two links each for a total of $2 \times d$ links (for example the middle links ML(8,1) and ML(8,2) are connected to the output switch OS1 from input switch MS(7,1), and the middle links
 5 ML(8,7) and ML(8,8) are connected to the output switch OS1 from input switch MS(7,2)).

Referring to diagram 200B in FIG. 2B, is a folded version of the multi-link multi-stage network 200A shown in FIG. 2A. The network 200B in FIG. 2B shows input stage 110 and output stage 120 are placed together. That is input switch IS1 and output switch
 10 OS1 are placed together, input switch IS2 and output switch OS2 are placed together, and similarly input switch IS12 and output switch OS12 are placed together. All the right going links {i.e., inlet links IL1 – IL24 and middle links ML(1,1) - ML(1,48)} correspond to input switches IS1 – IS12, and all the left going links {i.e., middle links ML(8,1) - ML(8,48) and outlet links OL1-OL24} correspond to output switches OS1 - OS12.

Middle stage 130 and middle stage 190 are placed together. That is middle switches MS(1,1) and MS(7,1) are placed together, middle switches MS(1,2) and MS(7,2) are placed together, and similarly middle switches MS(1,12) and MS(7,12) are placed together. All the right going middle links {i.e., middle links ML(1,1) - ML(1,48) and middle links ML(2,1) – ML(2,32) and the middle links ML(3,33) – ML(3,48)}
 20 correspond to middle switches MS(1,1) – MS(1,12), and all the left going middle links {i.e., middle links ML(7,1) - ML(7,32) and middle links ML(6,33) – ML(6,48) and middle links ML(8,1) and ML(8,48)} correspond to middle switches MS(7,1) – MS(7,12).

Middle stage 140 and middle stage 180 are placed together. That is middle switches MS(2,1) and MS(6,1) are placed together, middle switches MS(2,2) and MS(6,2) are placed together, and similarly middle switches MS(2,8) and MS(6,8) are placed together. All the right going middle links {i.e., middle links ML(2,1) - ML(2,48) and middle links ML(3,1) – ML(3,48)} correspond to middle switches MS(2,1) –

WO 2011/047368

PCT/US2010/052984

MS(2,8), and all the left going middle links {i.e., middle links ML(6,1) - ML(6,48) and middle links ML(7,1) and ML(7,48)} correspond to middle switches MS(6,1) – MS(6,8).

Middle stage 150 and middle stage 170 are placed together. That is middle switches MS(3,1) and MS(5,1) are placed together, middle switches MS(3,2) and MS(5,2) are placed together, and similarly middle switches MS(3,12) and MS(5,12) are placed together. All the right going middle links {i.e., middle links ML(3,1) - ML(3,48) and middle links ML(4,1) – ML(4,48)} correspond to middle switches MS(3,1) – MS(3,12, and all the left going middle links {i.e., middle links ML(5,1) - ML(5,48 and middle links ML(6,1) and ML(6,48)} correspond to middle switches MS(5,1) – MS(5,12).

Middle stage 160 is placed alone. All the right going middle links are the middle links ML(4,1) – ML(4,8), ML(4,17) – ML(4,24) and ML(4,33) - ML(4,48) and all the left going middle links are middle links ML(5,1) – ML(5,8), ML(5,17) – ML(5,24) and ML(5,33) - ML(5,48).

In one embodiment, in the network 200B of FIG. 2B, the switches that are placed together are implemented as separate switches then the network 200B is the generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 24$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch and a four by two switch. For example the input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1) – ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1 – OL2 being the outputs of the output switch OS1. Similarly in this embodiment of network 200B all the switches that are placed together in each middle stage are implemented as separate switches.

WO 2011/047368

PCT/US2010/052984

Modified-Hypercube Topology layout schemes:

Referring to layout 200C of FIG. 2C, in one embodiment, there are twelve blocks namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, and Block 23_24. Each block implements all the switches in one row of the network 200B of FIG. 2B, one of the key aspects of the current invention. For example Block 1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; Middle switch MS(4,1) is denoted by switch 5.

All the straight middle links are illustrated in layout 200C of FIG. 2C. For example in Block 1_2, inlet links IL1 – IL2, outlet links OL1 – OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 200C of FIG. 2C.

Even though it is not illustrated in layout 200C of FIG. 2C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit depending on the applications in different embodiments. There are a maximum of four quadrants in the layout 200C of FIG. 2C namely top-left, bottom-left, top-right and bottom-right quadrants. In each quadrant there are a maximum of four blocks. Top-left quadrant implements Block 1_2, Block 3_4, Block 5_6, and Block 7_8. Bottom-left quadrant implements Block 9_10, Block 11_12, Block 13_14, and Block 15_16. Top-right quadrant implements Block 17_18, Block 19_20. Bottom-right quadrant implements Block 21_22, and Block 23_24. There are two halves in layout 200C of FIG.

WO 2011/047368

PCT/US2010/052984

2C namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants. Right-half consists of top-right and bottom-right quadrants.

Recursively in each quadrant there are a maximum of four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements Block 1_2. Bottom-left sub-quadrant of top-left quadrant implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 5_6. Finally bottom-right sub-quadrant of top-left quadrant implements Block 7_8. Similarly there are a maximum of two sub-halves in each quadrant. For example in top-left quadrant there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left quadrant implements Block 1_2 and Block 3_4. Right-sub-half of top-left quadrant implements Block 5_6 and Block 7_8. Finally applicant notes that in each quadrant or half the blocks are arranged close to binary hypercube.

Layout 200D of FIG. 2D illustrates the inter-block links between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 200D of FIG. 2D can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track). As described before, the inter-link bandwidth provided between two physically adjacent blocks in the same column is hereinafter called 2's bandwidth or 2's BW. For example the inter-block links between switches 1 and 2 as illustrated in layout 200D of FIG. 2D is 2's BW.

Layout 200E of FIG. 2E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are

WO 2011/047368

PCT/US2010/052984

connected between switch 2 of Block 1_2 and switch 3 of Block 5_6. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 5_6. It must be noted that if there are an odd number of blocks in the rows of blocks then one of the blocks do not need inter-block links between switches 2 and 3, and also one of the switches for example switch 3 does not need to be implemented. For example in layout 200E there are three blocks in the topmost row namely Block 1_2, Block 5_6 and Block 17_18. In layout 200E there is no need to have inter-block links between switches 2 and 3 of Block 17_18 and hence there is no need to implement switch 3. Similarly in Block 19_20, Block 21_22 and Block 23_24 there is no need to provide inter-block links between switches 2 and 3 in those blocks. Also switch 3 is not implemented in those blocks.

Applicant notes that the inter-block links illustrated in layout 200E of FIG. 2E can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

In general the bandwidth offered within a quadrant or a partial quadrant of the layout formed by two nearest neighboring blocks is 2's BW. For example in layout 200C of FIG. 2C the bandwidth offered in top-right quadrant is 2's BW. Similarly the bandwidth offered within each of the other three quadrants top-left, bottom-left and bottom-right quadrants is 2' BW. Alternatively the bandwidth offered within a square or a partial square of blocks with the sides of the square consisting of two neighboring blocks is 2's BW. This definition can be generalized so that the bandwidth offered within a square of blocks with the sides consisting of "x" number of blocks, where $2^{y-1} \leq x \leq 2^y$ where "y" is an integer, is hereinafter x's BW.

Layout 200F of FIG. 2F illustrates the inter-block links between switches 3 and 4 of each block excepting that among the Block 17_18, Block 19_20, Block 21_22, and Block 23_24 the inter-block links are between the switches 2 and 4. For example middle

WO 2011/047368

PCT/US2010/052984

links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 200F of FIG. 2F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time division multiplexed single track). For example the inter-block links between switches 3 and 4 as illustrated in layout 200F of FIG. 2F is 4's BW.

Layout 200G of FIG. 2G illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 3_4. It must be noted that if the number of blocks in the rows of blocks is not a perfect multiple of four, then some of the blocks do not need inter-block links between switches 4 and 5, and also one of the switches for example switch 5 does not need to be implemented. For example in layout 200G there are three blocks in the topmost row namely Block 1_2, Block 5_6 and Block 17_18. In layout 200E there is no need to have inter-block links between switches 4 and 5 of Block 5_6 and hence there is no need to implement switch 5. Similarly in Block 7_8, Block 13_14 and Block 15_16 there is no need to provide inter-block links between switches 4 and 5 in those blocks. Also switch 5 is not implemented in those blocks.

Applicant notes that the inter-block links illustrated in layout 200G of FIG. 2G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division

WO 2011/047368

PCT/US2010/052984

multiplexed single track). The bandwidth offered between top-left quadrant, bottom-left quadrant, top-right partial quadrant and bottom-right partial quadrant is 4's BW in layout 200G of FIG. 2G.

The complete layout for the network 200B of FIG. 2B is given by combining the
5 links in layout diagrams of 200C, 200D, 200E, 200F, and 200G. Applicant notes that in the layout 200C of FIG. 2C, the inter-block links between switch 1 and switch 2 of corresponding blocks are vertical tracks as shown in layout 200D of FIG. 2D; the inter-block links between switch 2 and switch 3 of corresponding blocks are horizontal tracks as shown in layout 200E of FIG. 2E; the inter-block links between switch 3 and switch 4
10 of corresponding blocks are vertical tracks as shown in layout 200F of FIG. 2F; and finally the inter-block links between switch 4 and switch 5 of corresponding blocks are horizontal tracks as shown in layout 200G of FIG. 2G. The pattern is alternate vertical tracks and horizontal tracks.

Some of the key aspects of the current invention are discussed. 1) All the switches
15 in one row of the multi-stage network 200B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks or vertical tracks; 3) Since all the inter-block links are either horizontal or vertical tracks, all the inter-block links can be mapped on to island-style architectures in current commercial FPGAs; 4) The length of the longest wire is about half of the width (or length) of the
20 complete layout (For example middle link ML(4,4) is about half the width of the complete layout).

In accordance with the current invention, the layout 200C in FIG. 2C can be recursively extended for any arbitrarily large generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ the sub-quadrants, quadrants, and super-quadrants are
25 arranged in d-ary hypercube manner and also the inter-blocks are accordingly connected in d-ary hypercube topology. Even though all the embodiments in the current invention are illustrated for $N_1 = N_2$ when $N_1 = N_2 \neq 2^x$ where x is an integer, the embodiments can be extended for $N_1 \neq 2^x$ & $N_2 \neq 2^y$ where x and y are integers.

WO 2011/047368

PCT/US2010/052984

Just the same as was illustrated in diagram 100I of FIG. 1I for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$, a high-level

5 implementation of Block 1_2 of the layout 200C of FIG. 2C which represents a generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 24$; $d = 2$; and $s = 2$ is similar.

Just the same as was illustrated in diagram 100J of FIG. 1J for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of

10 the layout 100C of FIG. 1C which represents a generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$, a high-level implementation of Block 1_2 of the layout 200C of FIG. 2C which represents a generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 24$; $d = 2$; and $s = 2$ is similar.

Just the same as was illustrated in diagram 100K of FIG. 1K for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of

15 the layout 100C of FIG. 1C which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$, a high-level implementation of Block 1_2 of the layout 200C of FIG. 2C which represents a generalized folded multi-

20 stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 24$; $d = 2$; and $s = 2$ is similar.

Just the same as was illustrated in diagram 100K1 of FIG. 1K1 for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of

the layout 100C of FIG. 1C which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$, a high-level implementation of

25 Block 1_2 of the layout 200C of FIG. 2C which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1 = N_2 = 24$; $d = 2$; and $s = 1$ is similar.

WO 2011/047368

PCT/US2010/052984

Just the same as was illustrated in diagram 100L of FIG. 1L for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$, a high-level implementation of

5 Block 1_2 of the layout 200C of FIG. 2C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 24$; $d = 2$; and $s = 2$ is similar.

Just the same as was illustrated in diagram 100L1 of FIG. 1L1 for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized butterfly fat tree network

10 $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$, a high-level implementation of Block 1_2 of the layout 200C of FIG. 2C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1 = N_2 = 24$; $d = 2$; and $s = 1$ is similar.

Modified-Hypercube Topology with Nearest Neighbor connectivity first and the

15 **remaining with equal length wires, in every stage:**

Referring to layout 300A of FIG. 3A, 300B of FIG. 3B and 300C of FIG. 3C illustrate the topmost row of the extension of layout 100H for the network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 512$; $d = 2$; and $s = 2$. In one embodiment of the complete layout, not shown in FIGs. 3A-3C, there are four super-super-quadrants namely

20 top-left super-super-quadrant, bottom-left super-super-quadrant, top-right super-super-quadrant, and bottom-right super-super-quadrant. Total number of blocks in the complete layout is two hundred and fifty six. Top-left super-super-quadrant implements the blocks from block 1_2 to block 127_128. Bottom-left super-super-quadrant implements the blocks from block 129_130 to block 255_256. Top-right super-super-quadrant

25 implements the blocks from block 257_258 to block 319_320. Bottom-right super-super-quadrant implements the blocks from block 383_384 to block 511_512. Each block in all the super-super-quadrants has two more switches namely switch 8 and switch 9 in addition to the switches [1-7] described in layout 100H of FIG. 1H.

WO 2011/047368

PCT/US2010/052984

The embodiment of layout 300A of FIG. 3A illustrates the 2's BW provided in the top-most row of the complete layout namely between block 1_2 and block 5_6; between block 17_18 and block 21_22; between block 65_66 and block 69_90; between block 81_82 and block 85_86; between block 257_258 and block 261_262; between block 273_274 and block 275_276; between block 321_322 and block 325_326; and between block 337_338 and block 341_342. In one embodiment, the 2's BW provided between the respective blocks is through the inter-block links between corresponding switch 2 and switch 3 of the respective blocks.

The embodiment of layout 300B of FIG. 3B illustrates the 4's BW provided in the top-most row of the complete layout namely between block 1_2 and block 21_22; between block 5_6 and block 17_18; between block 65_66 and block 85_86; between block 69_70 and block 81_82; between block 257_258 and block 275_276; between block 261_262 and block 273_274; between block 321_322 and block 341_342; and between block 325_326 and block 337_338. In one embodiment, the 4's BW provided between the respective blocks is through the inter-block links between corresponding switch 4 and switch 5 of the respective blocks. In layout 300B, nearest neighbor blocks are connected together to provide 4's BW (for example the 4's BW provided between block 5_6 and block 17_18) and then the rest of the blocks are connected to provide the 4's BW (for example the 4's BW provided between block 1_2 and block 21_22).

The embodiment of layout 300C of FIG. 3C illustrates the 8's BW provided in the top-most row of the complete layout namely between block 1_2 and block 69_70; between block 5_6 and block 81_82; between block 17_18 and block 85_86; between block 21_22 and block 65_66; between block 257_258 and block 325_326; between block 261_262 and block 337_338; between block 273_274 and block 341_342; and between block 275_276 and block 321_322. In one embodiment, the 8's BW provided between the respective blocks is through the inter-block links between corresponding switch 6 and switch 7 of the respective blocks. In layout 300C, nearest neighbor blocks are connected together to provide 8's BW (for example the 8's BW provided between block 21_22 and block 65_66) and then the rest of the blocks are connected to provide the 8's BW (for example the 8's BW provided between block 1_2 and block 69_70).

WO 2011/047368

PCT/US2010/052984

Modified-Hypercube Topology with Recursive Nearest Neighbor connectivity, in every stage:

In another embodiment of the extension of layout 100H for the network

5 $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 512$; $d = 2$; and $s = 2$, the 2's BW and 4's BW are provided exactly the same as illustrated in FIG. 3A and FIG. 3B respectively; However 8's BW is offered as illustrated in layout 300D of FIG. 3D. The 8's BW is provided in the top-most row of the complete layout namely between block 21_22 and block 65_66; between block 17_18 and block 69_70; between block 5_6 and block
10 81_82; between block 1_2 and block 85_86; between block 275_276 and block 321_322; between block 273_274 and block 325_326; between block 261_262 and block 337_338; and between block 257_258 and block 341_342. In one embodiment, the 8's BW provided between the respective blocks is through the inter-block links between corresponding switch 6 and switch 7 of the respective blocks.

15 In layout 300D, nearest neighbor blocks are connected together to provide 8's BW recursively. Specifically first the 8's BW is provided between block 21_22 and block 65_66. Then the 8's BW is provided between the nearest neighbor blocks in the remaining blocks, i.e., between block 17_18 and block 69_70. Then the 8's BW is provided between the nearest neighbor blocks in the remaining blocks, i.e., between block
20 5_6 and block 81_82. Finally the 8's BW is provided between the nearest neighbor blocks in the remaining blocks, i.e., between block 1_2 and block 85_86. In the same manner, the 8's BW is provided in the remaining blocks between block 257_258 up to block 341_342.

25 **Modified-Hypercube Topology with the second stage implementing Nearest Neighbor connectivity:**

Referring to layout 400A of FIG. 4A, 400B of FIG. 4B and 400C of FIG. 4C illustrate the topmost row of the extension of layout 100H for the network

WO 2011/047368

PCT/US2010/052984

$V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 512$; $d = 2$; and $s = 2$. In another embodiment of the complete layout, not shown in FIGs. 4A-4C, there are four super-super-quadrants namely top-left super-super-quadrant, bottom-left super-super-quadrant, top-right super-super-quadrant, and bottom-right super-super-quadrant. Total number of blocks in the complete layout is two hundred fifty six. Top-left super-super-quadrant implements the blocks from block 1_2 to block 127_128. Bottom-left super-super-quadrant implements the blocks from block 129_130 to block 255_256. Top-right super-super-quadrant implements the blocks from block 257_258 to block 319_320. Bottom-right super-super-quadrant implements the blocks from block 383_384 to block 511_512. Each block in all the super-super-quadrants has two more switches namely switch 8 and switch 9 in addition to the switches [1-7] described in layout 100H of FIG. 1H.

In the embodiment of Layout 400A of FIG. 4A illustrates the 2's BW provided in the top-most row of the complete layout namely between block 1_2 and block 5_6; between block 17_18 and block 21_22; between block 65_66 and block 69_90; between block 81_82 and block 85_86; between block 257_258 and block 261_262; between block 273_274 and block 275_276; between block 321_322 and block 325_326; and between block 337_338 and block 341_342. In one embodiment, the 2's BW provided between the respective blocks is through the inter-block links between corresponding switch 2 and switch 3 of the respective blocks. Applicant notes that in layout 400A of FIG. 4A the first stage provides 2's BW between the blocks in the top-most row of the complete layout.

In the embodiment of Layout 400B of FIG. 4B illustrates the nearest neighbor connectivity between blocks of the top-most row of the complete layout to provide 4's BW, 8's BW, and 16's BW namely between block 5_6 and block 17_18 the bandwidth provided is 4's BW; between block 21_22 and block 65_66 the bandwidth provided is 8's BW; between block 69_70 and block 81_82 the bandwidth provided is 4's BW; between block 85_86 and block 257_258 the bandwidth provided is 16's BW; between block 261_262 and block 273_274 the bandwidth provided is 4's BW; between block 275_276 and block 321_322 the bandwidth provided is 8's BW; between block 325_326 and block 337_338 the bandwidth provided is 4's BW; and between block 1_2 and block 341_342

WO 2011/047368

PCT/US2010/052984

no bandwidth is provided. (Even though it is not illustrated, in another embodiment 16's BW can be provided between block 1_2 and block 342_342). In one embodiment, the BW provided between the respective blocks is through the inter-block links between corresponding switch 4 and switch 5 of the respective blocks. Applicant notes that in layout 400B of FIG. 4B the second stage provides the remaining nearest neighbor connectivity (i.e., after the first stage connectivity in layout 400A of FIG. 4A as illustrated provides nearest neighbor connectivity with 100% 2's BW) namely 50% of 4's BW, 25% of 8's BW and 12.5% of 16's BW, between the blocks in the top-most row of the complete layout.

The embodiment of layout 400C of FIG. 4C illustrates the 4's BW and 8's BW provided in the top-most row of the complete layout namely between block 1_2 and block 21_22 the bandwidth provided is 4's BW; between block 5_6 and block 69_70 the bandwidth provided is 8's BW; between block 17_18 and block 81_82 the bandwidth provided is 8's BW; between block 65_66 and block 85_86 the bandwidth provided is 4's BW; between block 257_258 and block 275_276 the bandwidth provided is 4's BW; between block 261_262 and block 325_326 the bandwidth provided is 8's BW; between block 273_274 and block 341_342 the bandwidth provided is 4's BW; between block 275_276 and block 337_338 the bandwidth provided is 8's BW; and between block 321_322 and block 341_342 the bandwidth provided is 4's BW. In one embodiment, the 4's BW and 8's BW provided between the respective blocks is through the inter-block links between corresponding switch 6 and switch 7 of the respective blocks. Applicant notes that in layout 400C of FIG. 4C the third stage provides 50% of 4's BW and 50% of 8's BW between the blocks in the top-most row of the complete layout.

The same process is repeated in the fourth stage by providing namely 25% of 8's BW and 87.5% of 16's BW is provided. This connectivity topology can be similarly extended to the network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 > 512$; $d = 2$; and $s = 2$.

WO 2011/047368

PCT/US2010/052984

Modified-Hypercube Topology with Partial & Tapered Connectivity (Bandwidth) in a stage, where $N_1 = N_2 = 512$:

Referring to layout 500 of FIG. 5 illustrates the topmost row of the extension of layout 100H for the network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 512$; $d = 2$; and $s =$

5 2. In another embodiment of the complete layout, not shown in FIG. 5, there are four super-super-quadrants namely top-left super-super-quadrant, bottom-left super-super-quadrant, top-right super-super-quadrant, and bottom-right super-super-quadrant. Total number of blocks in the complete layout is two hundred fifty six. Top-left super-super-quadrant implements the blocks from block 1_2 to block 127_128. Bottom-left super-

10 super-super-quadrant implements the blocks from block 129_130 to block 255_256. Top-right super-super-quadrant implements the blocks from block 257_258 to block 319_320. Bottom-right super-super-quadrant implements the blocks from block 383_384 to block 511_512. Each block in all the super-super-quadrants has two more switches namely switch 8 and switch 9 in addition to the switches [1-7] described in layout 100H of FIG.

15 1H.

The embodiment of layout 500 of FIG. 5 illustrates the 8's BW and 16's BW provided in the top-most row of the complete layout namely between block 21_22 and block 65_66 the bandwidth provided is 8's BW; between block 17_18 and block 69_70 the bandwidth provided is 8's BW; between block 85_86 and block 257_258 the

20 bandwidth provided is 16's BW; between block 81_82 and block 261_262 the bandwidth provided is 16's BW; between block 275_276 and block 321_322 the bandwidth provided is 8's BW; between block 273_274 and block 325_326 the bandwidth provided is 8's BW. In one embodiment, the 8's BW and 16's BW provided between the respective blocks is through the inter-block links between corresponding switch 6 and switch 7 of

25 the respective blocks. Applicant notes that in layout 500 of FIG. 5 the bandwidth provided between the blocks in the top-most row of the complete layout may be in anyone of the stages. Applicant observes that the 8's bandwidth provided in layout 500 of FIG. 5 is 50% of total 8's BW for full connectivity and 16's BW provided is 25% of the total 16's BW for full connectivity. In layout 500 of FIG. 5, the partial 8's BW and 16's

30 BW is provided in nearest neighbor connectivity manner recursively which makes the

WO 2011/047368

PCT/US2010/052984

wire lengths between different blocks to offer 8's BW is different and also makes the wire lengths between different blocks to offer 16's BW is different. Layout 500 of FIG. 5 illustrates an embodiment to provide partial bandwidth in a tapered manner, where it is not needed to provide the complete bandwidth in the higher stages.

5

Modified-Hypercube Topology with Partial & Tapered Connectivity (Bandwidth) in a stage, where $N_1 = N_2 = 2048$:

Referring to layout 600 of FIG. 6 illustrates the topmost row of the extension of layout 100H for the network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 2048$; $d = 2$; and $s =$
 10 2. In one embodiment of the complete layout, not shown in FIG. 6, there are four super-
 super-super-quadrants namely top-left super-super-super-quadrant, bottom-left super-
 super-super-quadrant, top-right super-super-super-quadrant, and bottom-right super-
 super-super-quadrant. Total number of blocks in the complete layout is one thousand and
 twenty four. Top-left super-super-quadrant implements the blocks from block 1_2 to
 15 block 511_512. Bottom-left super-super-quadrant implements the blocks from block
 513_514 to block 1023_1024. Top-right super-super-quadrant implements the blocks
 from block 1025_1026 to block 1535_1536. Bottom-right super-super-quadrant
 implements the blocks from block 1537_1538 to block 2047_2048. Each block in all the
 super-super-quadrants has four more switches namely switch 8, switch 9, switch 10 and
 20 switch 11 in addition to the switches [1-7] described in layout 100H of FIG. 1H.

In the embodiment of Layout 600 of FIG. 6 illustrates the 8's BW, 16's BW and
 32's BW provided in the top-most row of the complete layout namely between block
 21_22 and block 65_66 the bandwidth provided is 8's BW; between block 17_18 and
 block 69_70 the bandwidth provided is 8's BW; between block 85_86 and block 257_258
 25 the bandwidth provided is 16's BW; between block 81_82 and block 261_262 the
 bandwidth provided is 16's BW; between block 275_276 and block 321_322 the
 bandwidth provided is 8's BW; between block 273_274 and block 325_326 the
 bandwidth provided is 8's BW; between block 341_342 and block 1025_1026 the
 bandwidth provided is 32's BW; between block 337_338 and block 1029_1030 the

WO 2011/047368

PCT/US2010/052984

bandwidth provided is 32's BW; between block 1045_1046 and block 1089_1090 the bandwidth provided is 8's BW; between block 1041_1042 and block 1093_1094 the bandwidth provided is 8's BW; between block 1109_1110 and block 1281_1282 the bandwidth provided is 16's BW; between block 1105_1106 and block 1285_1286 the
5 bandwidth provided is 16's BW; between block 1299_1300 and block 1345_1346 the bandwidth provided is 8's BW; and between block 1297_1298 and block 1349_1350 the bandwidth provided is 8's BW.

In one embodiment, the 8's BW, 16's BW, and 32's BW provided between the respective blocks is through the inter-block links between corresponding switch 10 and
10 switch 11 of the respective blocks. Applicant notes that in layout 600 of FIG. 6 the bandwidth provided between the blocks in the top-most row of the complete layout may be in anyone of the stages. Applicant observes that the 8's bandwidth provided in layout 500 of FIG. 5 is 50% of total 8's BW for full connectivity, 16's BW provided is 25% of the total 16's BW for full connectivity and 32's BW provided is 12.5% of the total 32's
15 BW for full connectivity.

Applicant notes that in layout 600 of FIG. 6 the length of some of the wires providing bandwidth to 8's BW, 16's BW and 32's BW are of equal size, and the length of rest of the wires providing bandwidth to 8's BW, 16's BW and 32's BW are of equal size. In layout 600 of FIG. 6, the partial 8's BW, 16's BW and 32's BW is provided in
20 nearest neighbor connectivity manner recursively which makes the wire lengths between different blocks to offer 8's BW is different, also makes the wire lengths between different blocks to offer 16's BW is different and also makes the wire lengths between different blocks to offer 32's BW is different. Layout 600 of FIG. 6 illustrates an embodiment to provide partial bandwidth in a tapered manner, where it is not needed to
25 provide the complete bandwidth in the higher stages.

WO 2011/047368

PCT/US2010/052984

Modified-Hypercube Topology with Partial & Tapered Connectivity (Bandwidth) with equal length wires, in a stage:

Referring to layout 700 of FIG. 7 illustrates the topmost row of the extension of layout 100H for the network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 2048$; $d = 2$; and $s =$

5 2. In another embodiment of the complete layout, not shown in FIG. 7, there are four super-super-super-quadrants namely top-left super-super-super-quadrant, bottom-left super-super-super-quadrant, top-right super-super-super-quadrant, and bottom-right super-super-super-quadrant. Total number of blocks in the complete layout is one thousand and twenty four. Top-left super-super-quadrant implements the blocks from

10 block 1_2 to block 511_512. Bottom-left super-super-quadrant implements the blocks from block 513_514 to block 1023_1024. Top-right super-super-quadrant implements the blocks from block 1025_1026 to block 1535_1536. Bottom-right super-super-quadrant implements the blocks from block 1537_1538 to block 2047_2048. Each block in all the super-super-quadrants has four more switches namely switch 8, switch 9, switch 10 and

15 switch 11 in addition to the switches [1-7] described in layout 100H of FIG. 1H.

In the embodiment of Layout 700 of FIG. 7 illustrates the 8's BW, 16's BW and 32's BW provided in the top-most row of the complete layout namely between block 21_22 and block 69_70 the bandwidth provided is 8's BW; between block 17_18 and block 65_66 the bandwidth provided is 8's BW; between block 85_86 and block 261_262

20 the bandwidth provided is 16's BW; between block 81_82 and block 257_258 the bandwidth provided is 16's BW; between block 275_276 and block 325_326 the bandwidth provided is 8's BW; between block 273_274 and block 321_322 the bandwidth provided is 8's BW; between block 341_342 and block 1029_1030 the bandwidth provided is 32's BW; between block 337_338 and block 1025_1026 the

25 bandwidth provided is 32's BW; between block 1045_1046 and block 1093_1094 the bandwidth provided is 8's BW; between block 1041_1042 and block 1089_1090 the bandwidth provided is 8's BW; between block 1109_1110 and block 1285_1286 the bandwidth provided is 16's BW; between block 1105_1106 and block 1281_1282 the bandwidth provided is 16's BW; between block 1299_1300 and block 1349_1350 the

WO 2011/047368

PCT/US2010/052984

bandwidth provided is 8's BW; and between block 1297_1298 and block 1345_1346 the bandwidth provided is 8's BW.

In one embodiment, the 8's BW, 16's BW, and 32's BW provided between the respective blocks is through the inter-block links between corresponding switch 10 and switch 11 of the respective blocks. Applicant notes that in layout 700 of FIG. 7 the bandwidth provided between the blocks in the top-most row of the complete layout may be in anyone of the stages. Applicant observes that the 8's bandwidth provided in layout 500 of FIG. 5 is 50% of total 8's BW for full connectivity, 16's BW provided is 25% of the total 16's BW for full connectivity and 32's BW provided is 12.5% of the total 32's BW for full connectivity. Applicant notes that in layout 700 of FIG. 7 the length of the wires providing bandwidth to 8's BW, 16's BW and 32's BW are all of equal size. Layout 700 of FIG. 7 illustrates another embodiment to provide partial bandwidth in a tapered manner, where it is not needed to provide the complete bandwidth in the higher stages.

All the layout embodiments disclosed in the current invention are applicable to generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general, and for $N_1 = N_2 = N$.or $N_1 \neq N_2$, or $N_1 \neq 2^x$ & $N_2 \neq 2^y$ where x, y and d are integers.

Conversely applicant makes another important observation that generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ are implemented with the layout topology being the hypercube topology shown in layout 100C of FIG. 1C with large scale cross point reduction as any one of the networks described in the current invention namely: generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$,

WO 2011/047368

PCT/US2010/052984

generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general, and for $N_1 = N_2 = N$ or $N_1 \neq N_2$, or $N_1 \neq 2^x$ & $N_2 \neq 2^y$ where x, y and d are integers.

5

Symmetric RNB generalized multi-link multi-stage pyramid network

$V_{mlink-p}(N_1, N_2, d, s)$, **Connection Topology: Nearest Neighbor connectivity and with more than full Bandwidth:**

Referring to diagram 800A in FIG. 8A, in one embodiment, an exemplary
 10 generalized multi-link multi-stage pyramid $V_{mlink-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages of one hundred and forty four switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, 150, 160, 170, 180 and 190 is shown where input stage 110
 15 consists of sixteen switches with ten of two by four switches namely IS1, IS3, IS5, IS6, IS8, IS9, IS11, IS13, IS14, and IS16; and six of two by six switches namely IS2, IS4, IS7, IS10, IS12 and IS15.

The output stage 120 consists of sixteen switches with ten of four by two switches namely OS1, OS3, OS5, OS6, OS8, OS9, OS11, OS13, OS14, and OS16; and six of six
 20 by two switches namely OS2, OS4, OS7, OS10, OS12, and OS15.

The middle stage 130 consists of sixteen switches with four of four by four switches namely MS(1,1), MS(1,6), MS(1,11), and MS(1,16); four of six by four switches namely MS(1,2), MS(1,5), MS(1,12) and MS(1,15); four of four by six switches namely MS(1,3), MS(1,8), MS(1,9), and MS(1,14); and four of six by six switches
 25 namely MS(1,4), MS(1,7), MS(1,10), and MS(1,13).

The middle stage 190 consists of sixteen switches with four of four by four switches namely MS(7,1), MS(7,6), MS(7,11), and MS(7,16); four of four by six

WO 2011/047368

PCT/US2010/052984

switches namely MS(7,2), MS(7,5), MS(7,12) and MS(7,15); four of six by four switches namely MS(7,3), MS(7,8), MS(7,9), and MS(7,14); and four of six by six switches namely MS(7,4), MS(7,7), MS(7,10), and MS(7,13).

The middle stage 140 consists of sixteen switches with eight of four by four
 5 switches namely MS(2,1), MS(2,2), MS(2,5), MS(2,6), MS(2,11), MS(2,12), MS(2,15),
 and MS(2,16); and eight of six by four switches namely MS(2,3), MS(2,4), MS(2,7),
 MS(2,8), MS(2,9), MS(2,10), MS(2,13), and MS(2,14).

The middle stage 180 consists of sixteen switches with eight of four by four
 switches namely MS(6,1), MS(6,2), MS(6,5), MS(6,6), MS(6,11), MS(6,12), MS(6,15),
 10 and MS(6,16); and eight of four by six switches namely MS(6,3), MS(6,4), MS(6,7),
 MS(6,8), MS(6,9), MS(6,10), MS(6,13), and MS(6,14).

And all the remaining middle stages namely the middle stage 150 consists of
 sixteen, four by four switches MS(3,1) - MS(3,16), middle stage 160 consists of sixteen,
 four by four switches MS(4,1) - MS(4,16), and middle stage 170 consists of sixteen, four
 15 by four switches MS(5,1) - MS(5,16).

The multi-link multi-stage pyramid network $V_{mlink-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ shown in diagram 800A of FIG. 8A is built on top of the generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ by adding a few more links.

20 Since as disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above, a network $V_{mlink}(N_1, N_2, d, s)$ can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections, the network $V_{mlink-p}(N_1, N_2, d, s)$ can be operated in rearrangeably non-blocking manner for arbitrary
 25 fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections.

WO 2011/047368

PCT/US2010/052984

In one embodiment of this network each of the input switches IS1-IS16 and output switches OS1-OS16 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable $\frac{N}{d}$, where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by $\frac{N}{d}$. The size of each input switch IS1-IS16 can be denoted in general with the notation $d^+ * (2d)^+$ (hereinafter d^+ means d or more; or equivalently $\geq d$) and each output switch OS1-OS16 can be denoted in general with the notation $(2d)^+ * d^+$. Likewise, the size of each switch in any of the middle stages can be denoted as $(2d)^+ * (2d)^+$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-stage network can be represented with the notation $V_{m\text{link-p}}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL32), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

Each of the $\frac{N}{d}$ input switches IS1 – IS16 are connected to d^+ switches in middle stage 130 through two links each for a total of $(2 \times d)^+$ links (for example input switch IS2 is connected to middle switch MS(1,2) through the links ML(1,5), ML(1,6), and also connected to middle switch MS(1,1) through the links ML(1,7) and ML(1,8); In addition input switch IS2 is also connected to middle switch MS(1,5) through the links ML(1p,7) and ML(1p,8). The links ML(1,5), ML(1,6), ML(1,7) and ML(1,8) correspond to multistage network configuration and the links ML(1p,7) and ML(1p,8) correspond to the pyramid network configuration. Hereinafter all the pyramid links are denoted by ML(xp,y) where ‘x’ represents the stage the link belongs to and ‘y’ the link number in that stage.)

The middle links which connect switches in the same row in two successive middle stages are called hereinafter straight middle links; and the middle links which

WO 2011/047368

PCT/US2010/052984

connect switches in different rows in two successive middle stages are called hereinafter cross middle links. For example, the middle links ML(1,1) and ML(1,2) connect input switch IS1 and middle switch MS(1,1), so middle links ML(1,1) and ML(1,2) are straight middle links; where as the middle links ML(1,3) and ML(1,4) connect input switch IS1 and middle switch MS(1,2), since input switch IS1 and middle switch MS(1,2) belong to two different rows in diagram 800A of FIG. 8A, middle links ML(1,3) and ML(1,4) are cross middle links. It can be seen that pyramid links such as ML(1p,7) and ML(1p,8) are also cross middle links.

Each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,16) in the middle stage 130 are connected from d^+ input switches through two links each for a total of $(2 \times d)^+$ links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to d^+ switches in middle stage 140 through two links each for a total of $(2 \times d)^+$ links (for example the links ML(2,9) and ML(2,10) are connected from middle switch MS(1,3) to middle switch MS(2,3), and the links ML(2,11) and ML(2,12) are connected from middle switch MS(1,3) to middle switch MS(2,1) ; In addition middle switch MS(1,3) is also connected to middle switch MS(2,9) through the links ML(2p,11) and ML(2p,12). The links ML(2,9), ML(2,10), ML(2,11) and ML(2,12) correspond to multistage network configuration and the links ML(2p,11) and ML(2p,12) correspond to the pyramid network configuration.)

Each of the $\frac{N}{d}$ middle switches MS(2,1) – MS(2,16) in the middle stage 140 are connected from d^+ input switches through two links each for a total of $(2 \times d)^+$ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from input switch MS(1,1), and the links ML(1,11) and ML(1,12) are connected to the middle switch MS(2,1) from input switch MS(1,3)) and also are connected to d^+ switches in middle stage 150 through two links each for a total of $(2 \times d)^+$ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch

WO 2011/047368

PCT/US2010/052984

MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,6)).

Each of the $\frac{N}{d}$ middle switches MS(3,1) – MS(3,16) in the middle stage 150 are connected from d^+ input switches through two links each for a total of $(2 \times d)^+$ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from input switch MS(2,1), and the links ML(2,23) and ML(2,24) are connected to the middle switch MS(3,1) from input switch MS(2,6)) and also are connected to d^+ switches in middle stage 160 through two links each for a total of $(2 \times d)^+$ links (for example the links ML(4,1) and ML(4,2) are connected from middle switch MS(3,1) to middle switch MS(4,1), and the links ML(4,3) and ML(4,4) are connected from middle switch MS(3,1) to middle switch MS(4,11)).

Each of the $\frac{N}{d}$ middle switches MS(4,1) – MS(4,16) in the middle stage 160 are connected from d^+ input switches through two links each for a total of $(2 \times d)^+$ links (for example the links ML(4,1) and ML(4,2) are connected to the middle switch MS(4,1) from input switch MS(3,1), and the links ML(4,43) and ML(4,44) are connected to the middle switch MS(4,1) from input switch MS(3,11)) and also are connected to d^+ switches in middle stage 170 through two links each for a total of $(2 \times d)^+$ links (for example the links ML(5,1) and ML(5,2) are connected from middle switch MS(4,1) to middle switch MS(5,1), and the links ML(5,3) and ML(5,4) are connected from middle switch MS(4,1) to middle switch MS(5,11)).

Each of the $\frac{N}{d}$ middle switches MS(5,1) – MS(5,16) in the middle stage 170 are connected from d^+ input switches through two links each for a total of $(2 \times d)^+$ links (for example the links ML(5,1) and ML(5,2) are connected to the middle switch MS(5,1) from input switch MS(4,1), and the links ML(5,43) and ML(5,44) are connected to the middle switch MS(5,1) from input switch MS(4,11)) and also are connected to d^+ switches in middle stage 180 through two links each for a total of $(2 \times d)^+$ links (for example the

WO 2011/047368

PCT/US2010/052984

links ML(6,1) and ML(6,2) are connected from middle switch MS(5,1) to middle switch MS(6,1), and the links ML(6,3) and ML(6,4) are connected from middle switch MS(5,1) to middle switch MS(6,6)).

Each of the $\frac{N}{d}$ middle switches MS(6,1) – MS(6,16) in the middle stage 180 are
 5 connected from d^+ input switches through two links each for a total of $(2 \times d)^+$ links (for example the links ML(6,1) and ML(6,2) are connected to the middle switch MS(6,1) from input switch MS(5,1), and the links ML(6,23) and ML(6,24) are connected to the middle switch MS(6,1) from input switch MS(5,6)) and also are connected to d^+ switches in middle stage 190 through two links each for a total of $(2 \times d)^+$ links (for example the
 10 links ML(7,9) and ML(7,10) are connected from middle switch MS(6,3) to middle switch MS(7,3), and the links ML(7,11) and ML(7,12) are connected from middle switch MS(6,3) to middle switch MS(7,1); In addition middle switch MS(6,3) is also connected to middle switch MS(7,9) through the links ML(7p,11) and ML(7p,12). The links ML(7,9), ML(7,10), ML(7,11) and ML(7,12) correspond to multistage network
 15 configuration and the links ML(7p,11) and ML(7p,12) correspond to the pyramid network configuration.)

Each of the $\frac{N}{d}$ middle switches MS(7,1) – MS(7,16) in the middle stage 190 are
 connected from d^+ input switches through two links each for a total of $(2 \times d)^+$ links (for example the links ML(7,1) and ML(7,2) are connected to the middle switch MS(7,1) from
 20 input switch MS(6,1), and the links ML(7,11) and ML(7,12) are connected to the middle switch MS(7,1) from input switch MS(6,3)) and also are connected to d^+ switches in middle stage 120 through two links each for a total of $(2 \times d)^+$ links (for example middle switch MS(7,2) is connected to output switch OS2 through the links ML(8,5), ML(8,6), and also connected to middle switch OS1 through the links ML(8,7) and ML(8,8); In
 25 addition middle switch MS(7,2) is also connected to output switch OS5 through the links ML(8p,7) and ML(8p,8). The links ML(8,5), ML(8,6), ML(8,7) and ML(8,8) correspond to multistage network configuration and the links ML(8p,7) and ML(8p,8) correspond to the pyramid network configuration.)

WO 2011/047368

PCT/US2010/052984

Each of the $\frac{N}{d}$ middle switches OS1 – OS16 in the middle stage 120 are connected from d^+ input switches through two links each for a total of $(2 \times d)^+$ links (for example the links ML(8,1) and ML(8,2) are connected to the output switch OS1 from input switch MS(7,1), and the links ML(8,7) and ML(7,8) are connected to the output switch OS1 from input switch MS(7,2)).

Finally the connection topology of the network 800A shown in FIG. 8A is logically similar to back to back inverse Benes connection topology. In addition there are additional nearest neighbor links (i.e., pyramid links as described before) between the input stage 110 and middle stage 130; between middle stage 130 and middle stage 140; between middle stage 180 and middle stage 190; and middle stage 190 and output stage 120.

Applicant notes that in a multi-stage pyramid network with a fully connected multi-stage network configuration the pyramid links may not contribute for the connectivity however these links can be cleverly used to reduce the latency and power in an integrated circuit even though the number of cross points required are more to connect pyramid links than is required in a purely multi-stage network.

Applicant notes that in the generalized multi-link multi-stage pyramid network $V_{mlink-p}(N_1, N_2, d, s)$ the pyramid links are provided between any two successive stages as illustrated in the diagram 800A of FIG. 8A. The pyramid links in general are also provided between the switches in the same stage. The pyramid links are also provided between any two arbitrary stages.

Referring to diagram 800B in FIG. 8B, is a folded version of the multi-link multi-stage pyramid network 800A shown in FIG. 8A. The network 800B in FIG. 8B shows input stage 110 and output stage 120 are placed together. That is input switch IS1 and output switch OS1 are placed together, input switch IS2 and output switch OS2 are placed together, and similarly input switch IS16 and output switch OS16 are placed together. All the right going links {i.e., inlet links IL1 – IL32 and middle links ML(1,1) - ML(1,64)} correspond to input switches IS1 - IS16, and all the left going links {i.e.,

WO 2011/047368

PCT/US2010/052984

middle links ML(8,1) - ML(8,64) and outlet links OL1-OL32} correspond to output switches OS1 - OS16.

Middle stage 130 and middle stage 190 are placed together. That is middle switches MS(1,1) and MS(7,1) are placed together, middle switches MS(1,2) and
5 MS(7,2) are placed together, and similarly middle switches MS(1,16) and MS(7,16) are placed together. All the right going middle links {i.e., middle links ML(1,1) - ML(1,64) and middle links ML(2,1) - ML(2,64)} correspond to middle switches MS(1,1) - MS(1,16), and all the left going middle links {i.e., middle links ML(7,1) - ML(7,64) and middle links ML(8,1) and ML(8,64)} correspond to middle switches MS(7,1) -
10 MS(7,16).

Middle stage 140 and middle stage 180 are placed together. That is middle switches MS(2,1) and MS(6,1) are placed together, middle switches MS(2,2) and MS(6,2) are placed together, and similarly middle switches MS(2,16) and MS(6,16) are placed together. All the right going middle links {i.e., middle links ML(2,1) - ML(2,64)
15 and middle links ML(3,1) - ML(3,64)} correspond to middle switches MS(2,1) - MS(2,16), and all the left going middle links {i.e., middle links ML(6,1) - ML(6,64) and middle links ML(7,1) and ML(7,64)} correspond to middle switches MS(6,1) - MS(6,16).

Middle stage 150 and middle stage 170 are placed together. That is middle
20 switches MS(3,1) and MS(5,1) are placed together, middle switches MS(3,2) and MS(5,2) are placed together, and similarly middle switches MS(3,16) and MS(5,16) are placed together. All the right going middle links {i.e., middle links ML(3,1) - ML(3,64) and middle links ML(4,1) - ML(4,64)} correspond to middle switches MS(3,1) - MS(3,16), and all the left going middle links {i.e., middle links ML(5,1) - ML(5,64) and
25 middle links ML(6,1) and ML(6,64)} correspond to middle switches MS(5,1) - MS(5,16).

Middle stage 160 is placed alone. All the right going middle links are the middle links ML(4,1) - ML(4,64) and all the left going middle links are middle links ML(5,1) - ML(5,64).

WO 2011/047368

PCT/US2010/052984

Just the same way as the connection topology of the network 800A shown in FIG. 8A, the connection topology of the network 800B shown in FIG. 8B is the folded version and logically similar to back to back inverse Benes connection topology. In addition there are additional nearest neighbor links (i.e., pyramid links as described before) between the

5 input stage 110 and middle stage 130; between middle stage 130 and middle stage 140; between middle stage 180 and middle stage 190; and middle stage 190 and output stage 120.

The multi-link multi-stage pyramid network $V_{fold-link-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ shown in diagram 800B of FIG. 8B is built on top of the

10 generalized multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ by also adding a few more links.

Since as disclosed in U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above, a network $V_{fold-link}(N_1, N_2, d, s)$ can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and

15 also can be operated in strictly non-blocking manner for unicast connections, the network $V_{fold-link-p}(N_1, N_2, d, s)$ can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections.

In one embodiment, in the network 800B of FIG. 8B, the switches that are placed

20 together are implemented as separate switches then the network 800B is the generalized folded multi-link multi-stage pyramid network $V_{fold-link-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch and a four by two switch respectively. For example the input switch IS1 and output switch OS1 are

25 placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1) – ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1 – OL2 being the outputs

WO 2011/047368

PCT/US2010/052984

of the output switch OS1. Similarly in this embodiment of network 800B all the switches that are placed together in each middle stage are implemented as separate switches.

Modified-Hypercube Topology layout scheme:

Referring to layout 800C of FIG. 8C, in one embodiment, there are sixteen blocks
5 namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block
13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, Block 23_24, Block
25_26, Block 27_28, Block 29_30, and Block 31_32. Each block implements all the
switches in one row of the network 800B of FIG. 8B, one of the key aspects of the current
invention. For example Block 1_2 implements the input switch IS1, output Switch OS1,
10 middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch
MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1).
For the simplification of illustration, Input switch IS1 and output switch OS1 together are
denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are
denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are
15 denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are
denoted by switch 4; Middle switch MS(4,1) is denoted by switch 5.

All the straight middle links are illustrated in layout 800C of FIG. 8C. For
example in Block 1_2, inlet links IL1 – IL2, outlet links OL1 – OL2, middle link
ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link
20 ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link
ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link
ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are
illustrated in layout 800C of FIG. 8C.

Even though it is not illustrated in layout 800C of FIG. 8C, in each block, in
25 addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary
digital circuit depending on the applications in different embodiments. There are four
quadrants in the layout 800C of FIG. 8C namely top-left, bottom-left, top-right and
bottom-right quadrants. Top-left quadrant implements Block 1_2, Block 3_4, Block 5_6,
and Block 7_8. Bottom-left quadrant implements Block 9_10, Block 11_12, Block

WO 2011/047368

PCT/US2010/052984

- 13_14, and Block 15_16. Top-right quadrant implements Block 17_18, Block 19_20, Block 21_22, and Block 23_24. Bottom-right quadrant implements Block 25_26, Block 27_28, Block 29_30, and Block 31_32. There are two halves in layout 800C of FIG. 8C namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants.
- 5 Right-half consists of top-right and bottom-right quadrants.

Recursively in each quadrant there are four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements Block 1_2. Bottom-left sub-quadrant of top-left quadrant

10 implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 5_6. Finally bottom-right sub-quadrant of top-left quadrant implements Block 7_8. Similarly there are two sub-halves in each quadrant. For example in top-left quadrant there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left quadrant implements Block 1_2 and Block 3_4. Right-sub-half of top-left quadrant

15 implements Block 5_6 and Block 7_8. Finally applicant notes that in each quadrant or half the blocks are arranged as a general binary hypercube. Recursively in larger multi-stage network $V_{fold-link-p}(N_1, N_2, d, s)$ where $N_1 = N_2 > 32$, the layout in this embodiment in accordance with the current invention, will be such that the super-quadrants will also be arranged in d-ary hypercube manner. (In the embodiment of the

20 layout 800C of FIG. 8C, it is binary hypercube manner since $d = 2$, in the network $V_{fold-link-p}(N_1, N_2, d, s)$ 800B of FIG. 8B).

Layout 800D of FIG. 8D illustrates the inter-block links between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Middle links

25 ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Similarly pyramid middle links ML(1p,7), ML(1p,8), ML(8p,19), and ML(8p,20) are connected between switch 1 of Block 3_4 and switch 2 of Block 9_10. Similarly pyramid middle links ML(1p,19), ML(1p,20), ML(8p,7), and ML(8p,8) are connected between switch 2 of Block 3_4 and switch 1 of Block 9_10.

WO 2011/047368

PCT/US2010/052984

Applicant notes that the inter-block links illustrated in layout 800D of FIG. 8D can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment

5 inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

Layout 800E of FIG. 8E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are

10 connected between switch 2 of Block 1_2 and switch 3 of Block 3_4. Middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 3_4. Similarly pyramid middle links ML(2p,35), ML(2p,36), ML(7p,11), and ML(7p,12) are connected between switch 1 of Block 5_6 and switch 2 of Block 17_18. Similarly pyramid middle links ML(2p,11), ML(2p,12), ML(7p,35), and

15 ML(7p,36) are connected between switch 2 of Block 5_6 and switch 1 of Block 17_18.

Applicant notes that the inter-block links illustrated in layout 800E of FIG. 8E can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment

20 inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

Layout 800F of FIG. 8F illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are

25 connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 800F of FIG. 8F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example

30 middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an

WO 2011/047368

PCT/US2010/052984

alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time division multiplexed single track).

Layout 800G of FIG. 8G illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 800G of FIG. 8G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division multiplexed single track).

The complete layout for the network 800B of FIG. 8B is given by combining the links in layout diagrams of 800C, 800D, 800E, 800F, and 800G. Applicant notes that in the layout 800C of FIG. 8C, the inter-block links between switch 1 and switch 2 of corresponding blocks are vertical tracks as shown in layout 800D of FIG. 8D; the inter-block links between switch 2 and switch 3 of corresponding blocks are horizontal tracks as shown in layout 800E of FIG. 8E; the inter-block links between switch 3 and switch 4 of corresponding blocks are vertical tracks as shown in layout 800F of FIG. 8F; and finally the inter-block links between switch 4 and switch 5 of corresponding blocks are horizontal tracks as shown in layout 800G of FIG. 8G. The pattern is alternate vertical tracks and horizontal tracks. It continues recursively for larger networks of $N > 32$ as will be illustrated later.

Some of the key aspects of the current invention are discussed. 1) All the switches in one row of the multi-stage network 800B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks or vertical tracks; 3) Since all the inter-block links are either horizontal or vertical tracks, all the inter-block links can be mapped on to island-style architectures in current commercial

WO 2011/047368

PCT/US2010/052984

FPGA's; 4) The length of the longest wire is about half of the width (or length) of the complete layout (For example middle link ML(4,4) is about half the width of the complete layout).

In accordance with the current invention, the layout 800C in FIG. 8C can be
5 recursively extended for any arbitrarily large generalized folded multi-link multi-stage network $V_{fold-link-p}(N_1, N_2, d, s)$ the sub-quadrants, quadrants, and super-quadrants are arranged in d-ary hypercube manner and also the inter-blocks are accordingly connected in d-ary hypercube topology. Even though all the embodiments in the current invention are illustrated for $N_1 = N_2$, the embodiments can be extended for $N_1 \neq N_2$.

10 Referring to layout 800H of FIG. 8H, illustrates the extension of layout 800C for the network $V_{fold-link-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 128$; $d = 2$; and $s = 2$. There are four super-quadrants in layout 800H namely top-left super-quadrant, bottom-left super-quadrant, top-right super-quadrant, bottom-right super-quadrant. Total number of blocks in the layout 800H is sixty four. Top-left super-quadrant implements the blocks from
15 block 1_2 to block 31_32. Each block in all the super-quadrants has two more switches namely switch 6 and switch 7 in addition to the switches [1-5] illustrated in layout 800C of FIG. 8C. The inter-block link connection topology is the exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as it is shown in the layouts of FIG. 8D, FIG. 8E, FIG. 8F, and FIG. 8G respectively.

20 Bottom-left super-quadrant implements the blocks from block 33_34 to block 63_64. Top-right super-quadrant implements the blocks from block 65_66 to block 95_96. And bottom-right super-quadrant implements the blocks from block 97_98 to block 127_128. In all these three super-quadrants also, the inter-block link connection topology is exactly the same between the switches 1 and 2; switches 2 and 3; switches 3
25 and 4; switches 4 and 5 as that of the top-left super-quadrant.

Recursively in accordance with the current invention, the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-left super-quadrant and bottom-left super-quadrant. And similarly the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the

WO 2011/047368

PCT/US2010/052984

corresponding switches of top-right super-quadrant and bottom-right super-quadrant. The inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of top-left super-quadrant and top-right super-quadrant. And similarly the inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of bottom-left super-quadrant and bottom-right super-quadrant.

Referring to diagram 800I of FIG. 8I illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 800C of FIG. 8C which represents a generalized folded multi-link multi-stage network

10 $V_{fold-link-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 800I

illustrates both the intra-block and inter-block links connected to Block 1_2. The layout diagram 800I corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 800B of FIG. 8B. As noted before then the network 800B is the generalized folded multi-link multi-stage network

15 $V_{fold-link-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages.

That is the switches that are placed together in Block 1_2 as shown in FIG. 8I are namely input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switch MS(1,1) and middle switch MS(7,1) belonging to switch 2; middle switch MS(2,1) and middle switch MS(6,1) belonging to switch 3; middle switch MS(3,1) and middle switch MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1) – ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7), and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1 – OL2 being the outputs of the output switch OS1.

WO 2011/047368

PCT/US2010/052984

Middle switch MS(1,1) is implemented as four by four switch with the middle links ML(1,1), ML(1,2), ML(1,7) and ML(1,8) being the inputs and middle links ML(2,1) – ML(2,4) being the outputs; and middle switch MS(7,1) is implemented as four by four switch with the middle links ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs and middle links ML(8,1) – ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 800I of FIG. 8I.

Generalized Multi-link Butterfly Fat Pyramid Network Embodiment:

In another embodiment in the network 800B of FIG. 8B, the switches that are placed together are implemented as combined switch then the network 800B is the generalized multi-link butterfly fat pyramid network $V_{mlink-bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,390 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch. For example the input switch IS1 and output switch OS1 are placed together; so input switch IS1 and output OS1 are implemented as a six by six switch with the inlet links IL1, IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and OL2 being the outputs of the combined switch IS1&OS1. Similarly in this embodiment of network 800B all the switches that are placed together are implemented as a combined switch.

Layout diagrams 800C in FIG. 8C, 800D in FIG. 8D, 800E in FIG. 8E, 800F in FIG. 8G are also applicable to generalized multi-link butterfly fat pyramid network $V_{mlink-bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages. The layout 800C in FIG. 8C can be recursively extended for any arbitrarily large generalized multi-link butterfly fat pyramid network $V_{mlink-bfp}(N_1, N_2, d, s)$. Accordingly layout 800H of FIG. 8H is also applicable to generalized multi-link butterfly fat pyramid network $V_{mlink-bfp}(N_1, N_2, d, s)$.

WO 2011/047368

PCT/US2010/052984

Referring to diagram 800J of FIG. 8J illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 800C of FIG. 8C which represents a generalized multi-link butterfly fat pyramid network $V_{mlink-bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 800J illustrates

5 both the intra-block and inter-block links. The layout diagram 800J corresponds to the embodiment where the switches that are placed together are implemented as combined switch in the network 800B of FIG. 8B. As noted before then the network 800B is the generalized multi-link butterfly fat pyramid network $V_{mlink-bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent

10 Application Serial No. 60/940,390 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 8J are namely the combined input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switch implemented is combined input and output switch IS1&OS1); middle switch

15 MS(1,1) belonging to switch 2; middle switch MS(2,1) belonging to switch 3; middle switch MS(3,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links IL1, IL2 and ML(8,1), ML(8,2), ML(8,7), and ML(8,8) being the inputs and middle links ML(1,1) – ML(1,4), and outlet links OL1 – OL2 being the

20 outputs.

Middle switch MS(1,1) is implemented as eight by eight switch with the middle links ML(1,1), ML(1,2), ML(1,7), ML(1,8), ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs and middle links ML(2,1) – ML(2,4) and middle links ML(8,1) – ML(8,4) being the outputs. Similarly all the other middle switches are also implemented

25 as eight by eight switches as illustrated in 800J of FIG. 8J.

In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of

$V_{mlink-bfp}(N_1, N_2, d, s)$ can be implemented as a four by eight switch and a four by four

WO 2011/047368

PCT/US2010/052984

switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 8J, the left going middle links namely ML(7,1), ML(7,2), ML(7,11), and ML(7,12) are never switched to the right going middle links ML(2,1), ML(2,2), ML(2,3), and ML(2,4). And hence to implement MS(1,1) two switches namely: 1) a four by eight switch with the middle links ML(1,1), ML(1,2), ML(1,7), and ML(1,8) as inputs and the middle links ML(2,1), ML(2,2), ML(2,3), ML(2,4), ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs and 2) a four by four switch with the middle links ML(7,1), ML(7,2), ML(7,11), and ML(7,12) as inputs and the middle links ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs are sufficient without losing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

Generalized multi-stage pyramid network Embodiment:

In one embodiment, in the network 800B of FIG. 8B, the switches that are placed together are implemented as two separate switches in input stage 110 and output stage 120; and as four separate switches in all the middle stages, then the network 800B is the generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch and a four by two switch respectively. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs and outlet links OL1 – OL2 being the outputs.

The switches, corresponding to the middle stages that are placed together are implemented as four two by two switches. For example middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the

WO 2011/047368

PCT/US2010/052984

inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs; Similarly in this embodiment of network 800B all the switches that are placed together are implemented as separate switches.

Layout diagrams 800C in FIG. 8C, 800D in FIG. 8D, 800E in FIG. 8E, 800F in FIG. 8G are also applicable to generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages. The layout 800C in FIG. 8C can be recursively extended for any arbitrarily large generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$. Accordingly layout 800H of FIG. 8H is also applicable to generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$.

Referring to diagram 800K of FIG. 8K illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 800C of FIG. 8C which represents a generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 800K illustrates both the intra-block and inter-block links. The layout diagram 800K corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 800B of FIG. 8B. As noted before then the network 800B is the generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 8K are namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches

WO 2011/047368

PCT/US2010/052984

MS(1,1), MS(1,17), MS(7,1) and MS(7,17) belonging to switch 2; middle switches MS(2,1), MS(2,17), MS(6,1) and MS(6,17) belonging to switch 3; middle switches MS(3,1), MS(3,17), MS(5,1) and MS(5,17) belonging to switch 4; And middle switches MS(4,1), and MS(4,17) belonging to switch 5.

5 Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs and outlet links OL1 – OL2 being the outputs.

10 Middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being
 15 the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also
 20 implemented as two by two switches as illustrated in 800K of FIG. 8K.

Generalized multi-stage pyramid network Embodiment with S = 1:

In one embodiment, in the network 800B of FIG. 8B (where it is implemented with $s = 1$), the switches that are placed together are implemented as two separate switches in input stage 110 and output stage 120; and as two separate switches in all the
 25 middle stages, then the network 800B is the generalized folded multi-stage network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as two, two by two switches. For example the switch

WO 2011/047368

PCT/US2010/052984

input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1 – OL2 being the outputs.

The switches, corresponding to the middle stages that are placed together are implemented as two, two by two switches. For example middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs; Similarly in this embodiment of network 800B all the switches that are placed together are implemented as two separate switches.

Layout diagrams 800C in FIG. 8C, 800D in FIG. 8D, 800E in FIG. 8E, 800F in FIG. 8G are also applicable to generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages. The layout 800C in FIG. 8C can be recursively extended for any arbitrarily large generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$. Accordingly layout 800H of FIG. 8H is also applicable to generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$.

Referring to diagram 800K1 of FIG. 8K1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 800C of FIG. 8C when $s = 1$ which represents a generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ (All the double links are replaced by single links when $s = 1$). Block 1_2 in 800K1 illustrates both the intra-block and inter-block links. The layout diagram 800K1 corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 800B of FIG. 8B when $s = 1$. As noted before then the network 800B is the generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1 = N_2 =$

WO 2011/047368

PCT/US2010/052984

32; $d = 2$; and $s = 1$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,391 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 8K1 are namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated
 5 by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches MS(1,1) and MS(7,1) belonging to switch 2; middle switches MS(2,1) and MS(6,1) belonging to switch 3; middle switches MS(3,1) and MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

10 Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1) – ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1 – OL2 being the outputs.

15 Middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; And middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs.

20 Similarly all the other middle switches are also implemented as two by two switches as illustrated in 800K1 of FIG. 8K1.

Generalized Butterfly Fat Pyramid Network Embodiment:

In another embodiment in the network 800B of FIG. 8B, the switches that are placed together are implemented as two combined switches then the network 800B is the
 25 generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch.

WO 2011/047368

PCT/US2010/052984

For example the input switch IS1 and output switch OS1 are placed together; so input output switch IS1&OS1 are implemented as a six by six switch with the inlet links IL1, IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and
5 OL2 being the outputs of the combined switch IS1&OS1.

The switches, corresponding to the middle stages that are placed together are implemented as two four by four switches. For example middle switches MS(1,1) and MS(1,17) are placed together; so middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and
10 middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; middle switch MS(1,17) is implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly in this embodiment of network 800B all the switches that are placed together are implemented as a two combined switches.

Layout diagrams 800C in FIG. 8C, 800D in FIG. 8D, 800E in FIG. 8E, 800F in FIG. 8G are also applicable to generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages. The layout 800C in FIG. 8C can be recursively extended for any arbitrarily large generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$. Accordingly layout 800H of FIG. 8H is also
15 applicable to generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$.

Referring to diagram 800L of FIG. 8L illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 800C of FIG. 8C which represents a generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 800L illustrates both
25 the intra-block and inter-block links. The layout diagram 800L corresponds to the embodiment where the switches that are placed together are implemented as two combined switches in the network 800B of FIG. 8B. As noted before then the network 800B is the generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 =$

WO 2011/047368

PCT/US2010/052984

32; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 8L are namely the combined input and output switch IS1&OS1 belonging to switch 1, illustrated
 5 by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switch implemented is combined input and output switch IS1&OS1); middle switch MS(1,1) and MS(1,17) belonging to switch 2; middle switch MS(2,1) and MS(2,17) belonging to switch 3; middle switch MS(3,1) and MS(3,17) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

10 Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links IL1, IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs and middle links ML(1,1) – ML(1,4) and outlet links OL1 – OL2 being the outputs.

Middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1),
 15 ML(2,3), ML(8,1) and ML(8,3) being the outputs; And middle switch MS(1,17) is implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two four by four switches as illustrated in 800L of FIG. 8L.

20 In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of $V_{mink-bfp}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch
 25 MS(1,1) of Block 1_2 as shown FIG. 8L, the left going middle links namely ML(7,1) and ML(7,11) are never switched to the right going middle links ML(2,1) and ML(2,3). And hence to implement MS(1,1) two switches namely: 1) a two by four switch with the middle links ML(1,1) and ML(1,7) as inputs and the middle links ML(2,1), ML(2,3), ML(8,1), and ML(8,3) as outputs and 2) a two by two switch with the middle links

WO 2011/047368

PCT/US2010/052984

ML(7,1) and ML(7,11) as inputs and the middle links ML(8,1) and ML(8,3) as outputs are sufficient without losing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

5 Generalized Butterfly Fat Pyramid Network Embodiment with $S = 1$:

In one embodiment, in the network 800B of FIG. 8B (where it is implemented with $s = 1$), the switches that are placed together are implemented as a combined switch in input stage 110 and output stage 120; and as a combined switch in all the middle stages, then the network 800B is the generalized butterfly fat pyramid network

10 $V_{bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with five stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a four by four switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input and output switch
15 IS1&OS1 is implemented as four by four switch with the inlet links IL1, IL2, ML(8,1) and ML(8,3) being the inputs and middle links ML(1,1) – ML(1,2) and outlet links OL1 – OL2 being the outputs

The switches, corresponding to the middle stages that are placed together are implemented as a four by four switch. For example middle switches MS(1,1) is
20 implemented as four by four switch with middle links ML(1,1), ML(1,3), ML(7,1) and ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs..

Layout diagrams 800C in FIG. 8C, 800D in FIG. 8D, 800E in FIG. 8E, 800F in FIG. 8G are also applicable to generalized butterfly fat pyramid network
25 $V_{bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with five stages. The layout 800C in FIG. 8C can be recursively extended for any arbitrarily large generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$. Accordingly layout 800H of FIG. 8H is also applicable to generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$.

WO 2011/047368

PCT/US2010/052984

Referring to diagram 800L1 of FIG. 8L1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 800C of FIG. 8C when $s = 1$ which represents a generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ (All the double links are replaced by single links when $s = 1$). Block 1_2 in 800K1 illustrates both the intra-block and inter-block links. The layout diagram 800L1 corresponds to the embodiment where the switches that are placed together are implemented as a combined switch in the network 800B of FIG. 8B when $s = 1$. As noted before then the network 800B is the generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 1$ with nine stages as disclosed in U.S. Provisional Patent Application Serial No. 60/940,387 that is incorporated by reference above.

That is the switches that are placed together in Block 1_2 as shown in FIG. 8L1 are namely the input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switch MS(1,1) belonging to switch 2; middle switch MS(2,1) belonging to switch 3; middle switch MS(3,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

Input and output switch IS1&OS1 are placed together; so input and output switch IS1&OS1 is implemented as four by four switch with the inlet links IL1, IL2, ML(8,1) and ML(8,3) being the inputs and middle links ML(1,1) – ML(1,2) and outlet links OL1 – OL2 being the outputs.

Middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,3), ML(7,1) and ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 800L1 of FIG. 8L1.

In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of

$V_{mink-bfp}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two

WO 2011/047368

PCT/US2010/052984

switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 8L1, the left going middle links namely ML(7,1) and ML(7,5) are never switched to the right going middle links ML(2,1) and ML(2,2).

5 And hence to implement MS(1,1) two switches namely: 1) a two by four switch with the middle links ML(1,1) and ML(1,3) as inputs and the middle links ML(2,1), ML(2,2), ML(8,1), and ML(8,2) as outputs and 2) a two by two switch with the middle links ML(7,1) and ML(7,5) as inputs and the middle links ML(8,1) and ML(8,2) as outputs are sufficient without loosing any connectivity of the embodiment of MS(1,1) being

10 implemented as an eight by eight switch as described before.)

All the layout embodiments disclosed in the current invention are applicable to generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks

15 $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage pyramid networks $V_{fold-mlink-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{CCC}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general, and for both $N_1 = N_2 = N$ and $N_1 \neq N_2$, and d is any integer.

20 Conversely applicant makes another important observation that generalized cube connected cycles networks $V_{CCC}(N_1, N_2, d, s)$ are implemented with the layout topology being the hypercube topology shown in layout 200C of FIG. 2C with large scale cross point reduction as any one of the networks described in the current invention namely: generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-

25 stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage pyramid networks

WO 2011/047368

PCT/US2010/052984

$V_{fold-link-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks

$V_{mlink-bfp}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general, and for both

$N_1 = N_2 = N$ and $N_1 \neq N_2$, and d is any integer.

Applicant notes that in the generalized multi-stage pyramid networks

- 5 $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage pyramid networks $V_{fold-link-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$, and generalized hypercube networks
- 10 $V_{CCC}(N_1, N_2, d, s)$ the pyramid links are provided a) between the switches in any two successive stages, b) between the switches in the same stage, and c) between the switches any two arbitrary stages.

- In all the embodiments disclosed in the current invention, all the switches in some embodiments may be implemented as active switches consisting of cross points using
- 15 SRAM cells or Flash memory cells. Similarly in other embodiments the switches may be implemented as passive switches consisting of cross points using anti-fuse based vias or connections provided by metal layer programming as in structured ASICs. In another embodiment, the switches may be implemented as in 3D-FPGAs. In another embodiment where ASIC placement & routing, the switches are actually used to determine if two
- 20 wires are connected together or not; Alternatively they can be seen as switches during the implementation of the placement & routing however cross points in the cross state can be used as wire connections and in the bar state can be used as no connection of the wires.

25

WO 2011/047368

PCT/US2010/052984

Scheduling Method Embodiments for multi-stage pyramid networks and multi-link multi-stage pyramid networks:

FIG. 9A shows a high-level flowchart of a scheduling method 900, in one embodiment executed to setup multicast and unicast connections in the generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$ (for example the network 800A of FIG. 8A) or generalized folded multi-stage pyramid networks $V_{fold-mlink-p}(N_1, N_2, d, s)$ (for example the network 800B of FIG. 8B) or any of the generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$ disclosed in this invention. According to this embodiment, a multicast connection request is received in act 910. Then the control goes to act 920.

In act 920, based on the inlet link and input switch of the multicast connection received in act 910, from each available outgoing middle link of the input switch of the multicast connection, by traveling forward from middle stage 130 to middle stage $130 + 10 * (\log_d N - 2)$, the lists of all reachable middle switches in each middle stage are derived recursively. That is, first, by following each available outgoing middle link of the input switch all the reachable middle switches in middle stage 130 are derived. Next, starting from the selected middle switches in middle stage 130 traveling through all of their available outgoing middle links to middle stage 140 all the available middle switches in middle stage 140 are derived. This process is repeated recursively until all the reachable middle switches, starting from the outgoing middle link of input switch, in middle stage $130 + 10 * (\log_d N - 2)$ are derived. This process is repeated for each available outgoing middle link from the input switch of the multicast connection and separate reachable lists are derived in each middle stage from middle stage 130 to middle stage $130 + 10 * (\log_d N - 2)$ for all the available outgoing middle links from the input switch. Then the control goes to act 930.

In act 930, based on the destinations of the multicast connection received in act 910, from the output switch of each destination, by traveling backward from output stage

WO 2011/047368

PCT/US2010/052984

120 to middle stage $130+10*(\text{Log}_d N - 2)$, the lists of all middle switches in each middle stage from which each destination output switch (and hence the destination outlet links) is reachable, are derived recursively. That is, first, by following each available incoming middle link of the output switch of each destination link of the multicast

5 connection, all the middle switches in middle stage $130+10*(2*\text{Log}_d N - 4)$ from which the output switch is reachable, are derived. Next, starting from the selected middle switches in middle stage $130+10*(2*\text{Log}_d N - 4)$ traveling backward through all of their available incoming middle links from middle stage $130+10*(2*\text{Log}_d N - 5)$ all the available middle switches in middle stage $130+10*(2*\text{Log}_d N - 5)$ from which the

10 output switch is reachable, are derived. This process is repeated recursively until all the middle switches in middle stage $130+10*(\text{Log}_d N - 2)$ from which the output switch is reachable, are derived. This process is repeated for each output switch of each destination link of the multicast connection and separate lists in each middle stage from middle stage $130+10*(2*\text{Log}_d N - 4)$ to middle stage $130+10*(\text{Log}_d N - 2)$ for all the output

15 switches of each destination link of the connection are derived. Then the control goes to act 940.

In act 940, using the lists generated in acts 920 and 930, particularly list of middle switches derived in middle stage $130+10*(\text{Log}_d N - 2)$ corresponding to each outgoing link of the input switch of the multicast connection, and the list of middle switches

20 derived in middle stage $130+10*(\text{Log}_d N - 2)$ corresponding to each output switch of the destination links, the list of all the reachable destination links from each outgoing link of the input switch are derived. Specifically if a middle switch in middle stage $130+10*(\text{Log}_d N - 2)$ is reachable from an outgoing link of the input switch, say “x”, and also from the same middle switch in middle stage $130+10*(\text{Log}_d N - 2)$ if the output

25 switch of a destination link, say “y”, is reachable then using the outgoing link of the input switch x, destination link y is reachable. Accordingly, the list of all the reachable destination links from each outgoing link of the input switch is derived. The control then goes to act 950.

WO 2011/047368

PCT/US2010/052984

In act 950, among all the outgoing links of the input switch, it is checked if all the destinations are reachable using only one outgoing link of the input switch. If one outgoing link is available through which all the destinations of the multicast connection are reachable (i.e., act 950 results in “yes”), the control goes to act 970. And in act 970, the multicast connection is setup by traversing from the selected only one outgoing middle link of the input switch in act 950, to all the destinations. Then the control transfers to act 990.

If act 950 results “no”, that is one outgoing link is not available through which all the destinations of the multicast connection are reachable, then the control goes to act 960. In act 960, it is checked if all destination links of the multicast connection are reachable using two outgoing middle links from the input switch. According to the current invention, it is always possible to find at most two outgoing middle links from the input switch through which all the destinations of a multicast connection are reachable. So act 960 always results in “yes”, and then the control transfers to act 980. In act 980, the multicast connection is setup by traversing from the selected only two outgoing middle links of the input switch in act 960, to all the destinations. Then the control transfers to act 990.

In act 990, all the middle links between any two stages of the network used to setup the connection in either act 970 or act 980 are marked unavailable so that these middle links will be made unavailable to other multicast connections. The control then returns to act 910, so that acts 910, 920, 930, 940, 950, 960, 970, 980, and 990 are executed in a loop, for each connection request until the connections are set up.

In the example illustrated in FIG. 8A, four outgoing middle links are available to satisfy a multicast connection request if input switch is IS2, but only at most two outgoing middle links of the input switch will be used in accordance with this method. Similarly, although three outgoing middle links is available for a multicast connection request if the input switch is IS1, again only at most two outgoing middle links is used. The specific outgoing middle links of the input switch that are chosen when selecting two outgoing middle links of the input switch is irrelevant to the method of FIG. 9A so long as at most two outgoing middle links of the input switch are selected to ensure that the

WO 2011/047368

PCT/US2010/052984

connection request is satisfied, i.e. the destination switches identified by the connection request can be reached from the outgoing middle links of the input switch that are selected. In essence, limiting the outgoing middle links of the input switch to no more than two permits the network $V(N_1, N_2, d, s)$ to be operated in nonblocking manner in
5 accordance with the invention.

According to the current invention, using the method 940 of FIG. 9A, the network $V_p(N_1, N_2, d, s)$ or $V_{mlink-p}(N_1, N_2, d, s)$ is operated in rearrangeably nonblocking for unicast connections when $s \geq 1$, is operated in strictly nonblocking for unicast connections when $s \geq 2$, is operated in rearrangeably nonblocking for multicast
10 connections when $s \geq 2$, and is operated in strictly nonblocking for multicast connections when $s \geq 3$.

The connection request of the type described above in reference to method 900 of FIG. 9A can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request,
15 only one outgoing middle link of the input switch is used to satisfy the request. Moreover, in method 900 described above in reference to FIG. 9A any number of middle links may be used between any two stages excepting between the input stage and middle stage 130, and also any arbitrary fan-out may be used within each output stage switch, to satisfy the connection request.

As noted above method 900 of FIG. 9A can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks $V_p(N, d, s)$, $V_{mlink-p}(N, d, s)$, $V_p(N_1, N_2, d, s)$ or $V_{mlink-p}(N_1, N_2, d, s)$ disclosed in this
20 invention.

25

WO 2011/047368

PCT/US2010/052984

Scheduling Method Embodiments for butterfly fat pyramid networks and multi-link butterfly fat pyramid networks:

FIG. 10A shows a high-level flowchart of a scheduling method 1000, in one embodiment executed to setup multicast and unicast connections in the generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized folded butterfly fat pyramid networks $V_{fold-bfp}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$ or generalized folded multi-link butterfly fat pyramid networks $V_{fold-mlink-bfp}(N_1, N_2, d, s)$ disclosed in this invention. According to this embodiment, a multicast connection request is received in act 1010. Then the control goes to act 1020.

In act 1020, based on the inlet link and input switch of the multicast connection received in act 1010, from each available outgoing middle link of the input switch of the multicast connection, by traveling forward from middle stage 130 to middle stage $130 + 10 * (\text{Log}_d N - 2)$, the lists of all reachable middle switches in each middle stage are derived recursively. That is, first, by following each available outgoing middle link of the input switch all the reachable middle switches in middle stage 130 are derived. Next, starting from the selected middle switches in middle stage 130 traveling through all of their available outgoing middle links to middle stage 140 (reverse links from middle stage 130 to output stage 120 are ignored) all the available middle switches in middle stage 140 are derived. (In the traversal from any middle stage to the following middle stage only upward links are used and no reverse links or downward links are used. That is for example, while deriving the list of available middle switches in middle stage 140, the reverse links going from middle stage 130 to output stage 120 are ignored.) This process is repeated recursively until all the reachable middle switches, starting from the outgoing middle link of input switch, in middle stage $130 + 10 * (\text{Log}_d N - 2)$ are derived. This process is repeated for each available outgoing middle link from the input switch of the multicast connection and separate reachable lists are derived in each middle stage from middle stage 130 to middle stage $130 + 10 * (\text{Log}_d N - 2)$ for all the available outgoing middle links from the input switch. Then the control goes to act 1030.

WO 2011/047368

PCT/US2010/052984

In act 1030, based on the destinations of the multicast connection received in act 1010, from the output switch of each destination, by traveling backward from output stage 120 to middle stage $130+10*(\text{Log}_d N - 2)$, the lists of all middle switches in each middle stage from which each destination output switch (and hence the destination outlet links) is reachable, are derived recursively. That is, first, by following each available incoming middle link of the output switch of each destination link of the multicast connection, all the middle switches in middle stage 130 from which the output switch is reachable, are derived. Next, starting from the selected middle switches in middle stage 130 traveling backward through all of their available incoming middle links from middle stage 140 all the available middle switches in middle stage 140 (reverse links from middle stage 130 to input stage 120 are ignored) from which the output switch is reachable, are derived. (In the traversal from any middle stage to the following middle stage only upward links are used and no reverse links or downward links are used. That is for example, while deriving the list of available middle switches in middle stage 140, the reverse links coming to middle stage 130 from input stage 110 are ignored.) This process is repeated recursively until all the middle switches in middle stage $130+10*(\text{Log}_d N - 2)$ from which the output switch is reachable, are derived. This process is repeated for each output switch of each destination link of the multicast connection and separate lists in each middle stage from middle stage 130 to middle stage $130+10*(\text{Log}_d N - 2)$ for all the output switches of each destination link of the connection are derived. Then the control goes to act 1040.

In act 1040, using the lists generated in acts 1020 and 1030, particularly list of middle switches derived in middle stage $130+10*(\text{Log}_d N - 2)$ corresponding to each outgoing link of the input switch of the multicast connection, and the list of middle switches derived in middle stage $130+10*(\text{Log}_d N - 2)$ corresponding to each output switch of the destination links, the list of all the reachable destination links from each outgoing link of the input switch are derived. Specifically if a middle switch in middle stage $130+10*(\text{Log}_d N - 2)$ is reachable from an outgoing link of the input switch, say “x”, and also from the same middle switch in middle stage $130+10*(\text{Log}_d N - 2)$ if the output switch of a destination link, say “y”, is reachable then using the outgoing link of

WO 2011/047368

PCT/US2010/052984

the input switch x, destination link y is reachable. Accordingly, the list of all the reachable destination links from each outgoing link of the input switch is derived. The control then goes to act 1050.

5 In act 1050, among all the outgoing links of the input switch, it is checked if all the destinations are reachable using only one outgoing link of the input switch. If one outgoing link is available through which all the destinations of the multicast connection are reachable (i.e., act 1050 results in “yes”), the control goes to act 1070. And in act 1070, the multicast connection is setup by traversing from the selected only one outgoing middle link of the input switch in act 1050, to all the destinations. Also the nearest U-turn is taken while setting up the connection. That is at any middle stage if one of the middle switch in the lists derived in acts 1020 and 1030 are common then the connection is setup so that the U-turn is made to setup the connection from that middle switch for all the destination links reachable from that common middle switch. Then the control transfers to act 1090.

15 If act 1050 results “no”, that is one outgoing link is not available through which all the destinations of the multicast connection are reachable, then the control goes to act 1060. In act 1060, it is checked if all destination links of the multicast connection are reachable using two outgoing middle links from the input switch. According to the current invention, it is always possible to find at most two outgoing middle links from the input switch through which all the destinations of a multicast connection are reachable. So act 1060 always results in “yes”, and then the control transfers to act 1080. In act 1080, the multicast connection is setup by traversing from the selected only two outgoing middle links of the input switch in act 1060, to all the destinations. Also the nearest U-turn is taken while setting up the connection. That is at any middle stage if one of the middle switch in the lists derived in acts 1020 and 1030 are common then the connection is setup so that the U-turn is made to setup the connection from that middle switch for all the destination links reachable from that common middle switch. Then the control transfers to act 1090.

30 In act 1090, all the middle links between any two stages of the network used to setup the connection in either act 1070 or act 1080 are marked unavailable so that these

WO 2011/047368

PCT/US2010/052984

middle links will be made unavailable to other multicast connections. The control then returns to act 1010, so that acts 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, and 1090 are executed in a loop, for each connection request until the connections are set up.

According to the current invention, using the method 1040 of FIG. 10A, the
5 network $V_{bfp}(N_1, N_2, d, s)$ or $V_{mlink-bfp}(N_1, N_2, d, s)$ is operated in rearrangeably nonblocking for unicast connections when $s \geq 1$, is operated in strictly nonblocking for unicast connections when $s \geq 2$, is operated in rearrangeably nonblocking for multicast connections when $s \geq 2$, and is operated in strictly nonblocking for multicast connections when $s \geq 3$.

10 The connection request of the type described above in reference to method 1000 of FIG. 10A can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, only one outgoing middle link of the input switch is used to satisfy the request. Moreover, in method 1000 described above in reference to FIG. 10A any number of
15 middle links may be used between any two stages excepting between the input stage and middle stage 130, and also any arbitrary fan-out may be used within each output stage switch, to satisfy the connection request.

As noted above method 1000 of FIG. 10A can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks
20 $V_{bfp}(N, d, s)$, $V_{mlink-bfp}(N, d, s)$, $V_{bfp}(N_1, N_2, d, s)$ or $V_{mlink-bfp}(N_1, N_2, d, s)$ disclosed in this invention.

25

WO 2011/047368

PCT/US2010/052984

Applications Embodiments:

All the embodiments disclosed in the current invention are useful in many varieties of applications. FIG. 11A1 illustrates the diagram of 1100A1 which is a typical two by two switch with two inlet links namely IL1 and IL2, and two outlet links namely OL1 and OL2. The two by two switch also implements four crosspoints namely CP(1,1), CP(1,2), CP(2,1) and CP(2,2) as illustrated in FIG. 11A1. For example the diagram of 1100A1 may be the implementation of middle switch MS(1,1) of the diagram 100K of FIG. 1K where inlet link IL1 of diagram 1100A1 corresponds to middle link ML(1,1) of diagram 100K, inlet link IL2 of diagram 1100A1 corresponds to middle link ML(1,7) of diagram 100K, outlet link OL1 of diagram 1100A1 corresponds to middle link ML(2,1) of diagram 100K, outlet link OL2 of diagram 1100A1 corresponds to middle link ML(2,3) of diagram 100K.

1) Programmable Integrated Circuit Embodiments:

All the embodiments disclosed in the current invention are useful in programmable integrated circuit applications. FIG. 11A2 illustrates the detailed diagram 1100A2 for the implementation of the diagram 1100A1 in programmable integrated circuit embodiments. Each crosspoint is implemented by a transistor coupled between the corresponding inlet link and outlet link, and a programmable cell in programmable integrated circuit embodiments. Specifically crosspoint CP(1,1) is implemented by transistor C(1,1) coupled between inlet link IL1 and outlet link OL1, and programmable cell P(1,1); crosspoint CP(1,2) is implemented by transistor C(1,2) coupled between inlet link IL1 and outlet link OL2, and programmable cell P(1,2); crosspoint CP(2,1) is implemented by transistor C(2,1) coupled between inlet link IL2 and outlet link OL1, and programmable cell P(2,1); and crosspoint CP(2,2) is implemented by transistor C(2,2) coupled between inlet link IL2 and outlet link OL2, and programmable cell P(2,2).

If the programmable cell is programmed ON, the corresponding transistor couples the corresponding inlet link and outlet link. If the programmable cell is programmed OFF, the corresponding inlet link and outlet link are not connected. For example if the

WO 2011/047368

PCT/US2010/052984

programmable cell P(1,1) is programmed ON, the corresponding transistor C(1,1) couples the corresponding inlet link IL1 and outlet link OL1. If the programmable cell P(1,1) is programmed OFF, the corresponding inlet link IL1 and outlet link OL1 are not connected. In volatile programmable integrated circuit embodiments the programmable cell may be an SRAM (Static Random Address Memory) cell. In non-volatile programmable integrated circuit embodiments the programmable cell may be a Flash memory cell. Also the programmable integrated circuit embodiments may implement field programmable logic arrays (FPGA) devices, or programmable Logic devices (PLD), or Application Specific Integrated Circuits (ASIC) embedded with programmable logic circuits or 3D-FPGAs.

FIG. 11A2 also illustrates a buffer B1 on inlet link IL2. The signals driven along inlet link IL2 are amplified by buffer B1. Buffer B1 can be inverting or non-inverting buffer. Buffers such as B1 are used to amplify the signal in links which are usually long.

2) One-time Programmable Integrated Circuit Embodiments:

All the embodiments disclosed in the current invention are useful in one-time programmable integrated circuit applications. FIG. 11A3 illustrates the detailed diagram 1100A3 for the implementation of the diagram 1100A1 in one-time programmable integrated circuit embodiments. Each crosspoint is implemented by a via coupled between the corresponding inlet link and outlet link in one-time programmable integrated circuit embodiments. Specifically crosspoint CP(1,1) is implemented by via V(1,1) coupled between inlet link IL1 and outlet link OL1; crosspoint CP(1,2) is implemented by via V(1,2) coupled between inlet link IL1 and outlet link OL2; crosspoint CP(2,1) is implemented by via V(2,1) coupled between inlet link IL2 and outlet link OL1; and crosspoint CP(2,2) is implemented by via V(2,2) coupled between inlet link IL2 and outlet link OL2.

If the via is programmed ON, the corresponding inlet link and outlet link are permanently connected which is denoted by thick circle at the intersection of inlet link and outlet link. If the via is programmed OFF, the corresponding inlet link and outlet link are not connected which is denoted by the absence of thick circle at the intersection of

WO 2011/047368

PCT/US2010/052984

inlet link and outlet link. For example in the diagram 1100A3 the via V(1,1) is programmed ON, and the corresponding inlet link IL1 and outlet link OL1 are connected as denoted by thick circle at the intersection of inlet link IL1 and outlet link OL1; the via V(2,2) is programmed ON, and the corresponding inlet link IL2 and outlet link OL2 are connected as denoted by thick circle at the intersection of inlet link IL2 and outlet link OL2; the via V(1,2) is programmed OFF, and the corresponding inlet link IL1 and outlet link OL2 are not connected as denoted by the absence of thick circle at the intersection of inlet link IL1 and outlet link OL2; the via V(2,1) is programmed OFF, and the corresponding inlet link IL2 and outlet link OL1 are not connected as denoted by the absence of thick circle at the intersection of inlet link IL2 and outlet link OL1. One-time programmable integrated circuit embodiments may be anti-fuse based programmable integrated circuit devices or mask programmable structured ASIC devices.

3) Integrated Circuit Placement and Route Embodiments:

All the embodiments disclosed in the current invention are useful in Integrated Circuit Placement and Route applications, for example in ASIC backend Placement and Route tools. FIG. 11A4 illustrates the detailed diagram 1100A4 for the implementation of the diagram 1100A1 in Integrated Circuit Placement and Route embodiments. In an integrated circuit since the connections are known a-priori, the switch and crosspoints are actually virtual. However the concept of virtual switch and virtual crosspoint using the embodiments disclosed in the current invention reduces the number of required wires, wire length needed to connect the inputs and outputs of different netlists and the time required by the tool for placement and route of netlists in the integrated circuit.

Each virtual crosspoint is used to either to hardwire or provide no connectivity between the corresponding inlet link and outlet link. Specifically crosspoint CP(1,1) is implemented by direct connect point DCP(1,1) to hardwire (i.e., to permanently connect) inlet link IL1 and outlet link OL1 which is denoted by the thick circle at the intersection of inlet link IL1 and outlet link OL1; crosspoint CP(2,2) is implemented by direct connect point DCP(2,2) to hardwire inlet link IL2 and outlet link OL2 which is denoted by the thick circle at the intersection of inlet link IL2 and outlet link OL2. The diagram 1100A4 does not show direct connect point DCP(1,2) and direct connect point DCP(1,3)

WO 2011/047368

PCT/US2010/052984

since they are not needed and in the hardware implementation they are eliminated. Alternatively inlet link IL1 needs to be connected to outlet link OL1 and inlet link IL1 does not need to be connected to outlet link OL2. Also inlet link IL2 needs to be connected to outlet link OL2 and inlet link IL2 does not need to be connected to outlet link OL1. Furthermore in the example of the diagram 1100A4, there is no need to drive the signal of inlet link IL1 horizontally beyond outlet link OL1 and hence the inlet link IL1 is not even extended horizontally until the outlet link OL2. Also the absence of direct connect point DCP(2,1) illustrates there is no need to connect inlet link IL2 and outlet link OL1.

10 In summary in integrated circuit placement and route tools, the concept of virtual switches and virtual cross points is used during the implementation of the placement & routing algorithmically in software, however during the hardware implementation cross points in the cross state are implemented as hardwired connections between the corresponding inlet link and outlet link, and in the bar state are implemented as no
15 connection between inlet link and outlet link.

3) More Application Embodiments:

All the embodiments disclosed in the current invention are also useful in the design of SoC interconnects, Field programmable interconnect chips, parallel computer
20 systems and in time-space-time switches.

Numerous modifications and adaptations of the embodiments, implementations, and examples described herein will be apparent to the skilled artisan in view of the disclosure.

WO 2011/047368

PCT/US2010/052984

CLAIMS

What is claimed is:

1. A two-dimensional layout of hierarchical routing network comprising a total of $a \times b$ blocks with one side of said layout having the size of a blocks and the other side of
5 said layout having the size of b blocks where $a \geq 1$ and $b \geq 1$, and

Said routing network comprising a total of N_1 inlet links and a total of N_2 outlet links and y hierarchical stages wherein either

$N_2 = N_1 \times d_2$, $N_1 = (a \times b) \times d$, and said each block comprising d inlet links and $d \times d_2$ outlet links; or

- 10 $N_1 = N_2 \times d_1$, $N_2 = (a \times b) \times d$, and said each block comprising d outlet links and $d \times d_1$ inlet links, and

Said each stage comprising a switch of size $d \times d$, where $d \geq 2$ and each said switch of size $d \times d$ having d incoming links and d outgoing links; and

- 15 Said incoming links and outgoing links in each switch in said each stage of said each block comprising a plurality of forward connecting links connecting from switches in lower stage to switches in the immediate succeeding higher stage, and also comprising a plurality of backward connecting links connecting from switches in higher stage to switches in the immediate preceding lower stage; and

- 20 Said forward connecting links comprising a plurality of straight links connecting from a switch in a stage in a block to a switch in another stage in the same block and also comprising a plurality of cross links connecting from a switch in a stage in a block to a switch in another stage in a different block, and

WO 2011/047368

PCT/US2010/052984

Said backward connecting links comprising a plurality of straight links connecting from a switch in a stage in a block to a switch in another stage in the same block and also comprising a plurality of cross links connecting from a switch in a stage in a block to a switch in another stage in a different block.

- 5 2. The two-dimensional layout of claim 1, wherein said all cross links are connecting as either vertical or horizontal links between switches in two different said blocks.
3. The two-dimensional layout claim 2, wherein said cross links in succeeding stages are connecting as alternative vertical and horizontal links between switches in said blocks.
- 10 4. The two-dimensional layout of claim 3, wherein either said cross links from switches in a stage in one of said blocks are connecting to switches in the succeeding stage in another of said blocks so that said cross links are either vertical links or horizontal links and vice versa, and hereinafter such cross links are “shuffle exchange links”).
- 15 5. The two-dimensional layout of claim 4, wherein said all horizontal shuffle exchange links are connected, in each said stage, between two sets of neighboring blocks with each said set having neighboring blocks of size 2^x where $0 \leq x \leq y$ such that $x = 0$ for the lowest stage and $x = y$ for the highest stage and each block in one of the said sets is connected to at least one block in said second set excepting when the number of blocks
- 20 in one of the sets are less than the number of blocks in said second set, and
- said all vertical shuffle exchange links are connected, in each said stage, between two sets of neighboring blocks with each said set having neighboring blocks of size 2^x where $0 \leq x \leq y$ such that $x = 0$ for the lowest stage and $x = y$ for the highest stage and each block in one of the said sets is connected to at least one block in said second set
- 25 excepting when the number of blocks in one of the sets are less than the number of blocks in said second set.

WO 2011/047368

PCT/US2010/052984

6. The two-dimensional layout of claim 5, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are substantially of equal length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are substantially of equal length in the entire said two-dimensional layout, and

5 the shortest horizontal shuffle exchange links are connecting at the lowest stage and between switches in two nearest neighboring said blocks, and length of the horizontal shuffle exchange links is doubled in each succeeding stage; and the shortest vertical shuffle exchange links are connecting at the lowest stage and between switches in two

10 nearest neighboring said blocks, and length of the vertical shuffle exchange links is doubled in each succeeding stage.

7. The two-dimensional layout of claim 5, wherein said all horizontal shuffle exchange links between switches of a block one in each said two sets of blocks in any two corresponding said succeeding stages are connected so that the two nearest neighbors

15 are connected first and then the two nearest neighbors are connected in the remaining blocks which is repeated until the switches in all the blocks are connected, and said all vertical shuffle exchange links between switches of a block one in each said two sets of blocks in any two corresponding said succeeding stages are connected so that the two nearest neighbors are connected first and then the two nearest neighbors are connected in

20 the remaining blocks which is repeated until the switches in all the blocks are connected in the entire said two-dimensional layout.

8. The two-dimensional layout of claim 6, wherein $y \geq (\log_2(N_1))$ when $N_2 = N_1 \times d_2$, or $y \geq (\log_2(N_2))$ when $N_1 = N_2 \times d_2$ so that the length of the horizontal shuffle exchange links in the highest stage is equal to half the size of the horizontal size

25 of said two dimensional grid of blocks, and the length of the vertical shuffle exchange links in the highest stage is equal to half the size of the vertical size of said two dimensional grid of blocks.

9. The two-dimensional layout of claim 8, wherein $d = 2$ and there is only one switch in each said stage in each said block connecting said forward connecting links and

WO 2011/047368

PCT/US2010/052984

there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast Benes network with full bandwidth.

10. The two-dimensional layout of claim 8, wherein $d = 2$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast Benes network and rearrangeably nonblocking for arbitrary fan-out multicast Benes network with full bandwidth.

11. The two-dimensional layout of claim 8, wherein $d = 2$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast Benes network with full bandwidth.

12. The two-dimensional layout of claim 6, wherein $y \geq (\log_2(N_1))$ when $N_2 = N_1 \times d_2$, or $y \geq (\log_2(N_2))$ when $N_1 = N_2 \times d_2$ so that the length of the horizontal shuffle exchange links in the highest stage is equal to half the size of the horizontal size of said two dimensional grid of blocks and the length of the vertical shuffle exchange links in the highest stage is equal to half the size of the vertical size of said two dimensional grid of blocks, and

said each block further comprising a plurality of U-turn links within switches in each of said stages in each of said blocks.

13. The two-dimensional layout of claim 12, wherein $d = 2$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast butterfly fat tree network with full bandwidth.

WO 2011/047368

PCT/US2010/052984

14. The two-dimensional layout of claim 12, wherein $d = 2$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast butterfly fat tree network and rearrangeably nonblocking for arbitrary fan-out multicast butterfly fat tree network with full bandwidth.
15. The two-dimensional layout of claim 12, wherein $d = 2$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast butterfly fat tree network with full bandwidth.
16. The two-dimensional layout of claim 1, wherein said horizontal and vertical links are implemented on two or more metal layers.
17. The two-dimensional layout of claim 1, wherein said switches comprising active and reprogrammable cross points and said each cross point is programmable by an SRAM cell or a Flash Cell.
18. The two-dimensional layout of claim 1, wherein said blocks are of equal size.
19. The two-dimensional layout of claim 16, wherein said blocks further comprising Lookup Tables (hereinafter "LUTs") having outlet links connected to inlet links of said routing network and further having inlet links connected from outlet links of said routing network and said two-dimensional layout with said blocks of LUTs is a field programmable gate array (FPGA) integrated circuit device or field programmable gate array (FPGA) block embedded in another integrated circuit device.
20. The two-dimensional layout of claim 16, wherein said blocks further comprising AND or OR gates having outlet links connected to inlet links of said routing network and further having inlet links connected from outlet links of said routing network and said

WO 2011/047368

PCT/US2010/052984

two-dimensional layout with said blocks of AND or OR gates is a programmable logic device (PLD).

21. The two-dimensional layout of claim 1, wherein said blocks further comprising any arbitrary hardware logic or memory circuits.

5 22. The two-dimensional layout of claim 1, wherein said switches comprising active one-time programmable cross points.

23. The two-dimensional layout of claim 22, wherein said blocks further comprising Lookup Tables (hereinafter "LUTs") having outlet links connected to inlet links of said routing network and further having inlet links connected from outlet links of said routing
10 network and said two-dimensional layout with said blocks of LUTs is a mask programmable gate array (MPGA) device or a structured ASIC device.

24. The two-dimensional layout of claim 1, wherein said switches comprising passive cross points or just connection of two links or not and said blocks further comprising any arbitrary hardware logic or memory circuits having outlet links connected to inlet links of
15 said routing network and further having inlet links connected from outlet links of said routing network and said two-dimensional layout with said blocks of arbitrary hardware logic or memory circuits is an Application Specific Integrated Circuit (ASIC) device.

25. The two-dimensional layout of claim 1, wherein said blocks further recursively comprise one or more sub-blocks and a sub-routing network.

20 26. The two-dimensional layout of claim 4, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \geq (\log_2(N_1))$ when $N_2 = N_1 \times d_2$, or $y \geq (\log_2(N_2))$ when $N_1 = N_2 \times d_2$.

WO 2011/047368

PCT/US2010/052984

27. The two-dimensional layout of claim 26, wherein $d = 2$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast
5 generalized multi-stage network with full bandwidth.

28. The two-dimensional layout of claim 26, wherein $d = 2$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast
10 generalized multi-stage network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-stage network with full bandwidth.

29. The two-dimensional layout of claim 26, wherein $d = 2$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said
15 backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-stage network with full bandwidth.

30. The two-dimensional layout of claim 4, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two
20 corresponding said succeeding stages are of different length and $y \geq (\log_2(N_1))$ when $N_2 = N_1 \times d_2$, or $y \geq (\log_2(N_2))$ when $N_1 = N_2 \times d_2$, and
said each block further comprising a plurality of U-turn links within switches in each of said stages in each of said blocks.

31. The two-dimensional layout of claim 30, wherein $d = 2$ and there is only one
25 switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward

WO 2011/047368

PCT/US2010/052984

connecting links and said routing network is rearrangeably nonblocking for unicast generalized butterfly fat tree network with full bandwidth.

32. The two-dimensional layout of claim 30, wherein $d = 2$ and there are at least two switches in each said stage in each said block connecting said forward connecting links
5 and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized butterfly fat tree Network and rearrangeably nonblocking for arbitrary fan-out multicast generalized butterfly fat tree network with full bandwidth.

33. The two-dimensional layout of claim 30, wherein $d = 2$ and there are at least three
10 switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized butterfly fat tree network with full bandwidth.

34. The two-dimensional layout of claim 1, wherein said straight links connecting
15 from switches in each said block are connecting to switches in the same said block; and said cross links are connecting as vertical or horizontal or diagonal links between two different said blocks.

35. The two-dimensional layout of claim 8, wherein $d = 4$ and there is only one switch in each said stage in each said block connecting said forward connecting links and
20 there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast multi-link Benes network with full bandwidth.

36. The two-dimensional layout of claim 8, wherein $d = 4$ and there are at least two switches in each said stage in each said block connecting said forward connecting links
25 and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast

WO 2011/047368

PCT/US2010/052984

multi-link Benes network and rearrangeably nonblocking for arbitrary fan-out multicast multi-link Benes network with full bandwidth.

37. The two-dimensional layout of claim 8, wherein $d = 4$ and there are at least three switches in each said stage in each said block connecting said forward connecting links
5 and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast multi-link Benes network with full bandwidth.

38. The two-dimensional layout of claim 12, wherein $d = 4$ and there is only one switch in each said stage in each said block connecting said forward connecting links and
10 there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast multi-link butterfly fat tree network with full bandwidth.

39. The two-dimensional layout of claim 12, wherein $d = 4$ and there are at least two switches in each said stage in each said block connecting said forward connecting links
15 and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast multi-link butterfly fat tree network and rearrangeably nonblocking for arbitrary fan-out multicast multi-link butterfly fat tree network with full bandwidth.

40. The two-dimensional layout of claim 12, wherein $d = 4$ and there are at least three switches in each said stage in each said block connecting said forward connecting links
20 and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast multi-link butterfly fat tree network with full bandwidth.

41. The two-dimensional layout of claim 26, wherein $d = 4$ and there is only one switch in each said stage in each said block connecting said forward connecting links and
25 there is only one switch in each said stage in each said block connecting said backward

WO 2011/047368

PCT/US2010/052984

connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-link multi-stage network with full bandwidth.

42. The two-dimensional layout of claim 26, wherein $d = 4$ and there are at least two switches in each said stage in each said block connecting said forward connecting links
5 and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized multi-link multi-stage network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-link multi-stage network with full bandwidth.

43. The two-dimensional layout of claim 26, wherein $d = 4$ and there are at least three
10 switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-link multi-stage network with full bandwidth.

44. The two-dimensional layout of claim 30, wherein $d = 4$ and there is only one
15 switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-link butterfly fat tree network with full bandwidth.

45. The two-dimensional layout of claim 30, wherein $d = 4$ and there are at least two
20 switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized multi-link butterfly fat tree Network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-link butterfly fat tree network with full
25 bandwidth.

46. The two-dimensional layout of claim 30, wherein $d = 4$ and there are at least three switches in each said stage in each said block connecting said forward connecting links

WO 2011/047368

PCT/US2010/052984

and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-link butterfly fat tree network with full bandwidth.

47. The two-dimensional layout of claim 1, wherein said plurality of forward
5 connecting links use a plurality of buffers to amplify signals driven through them and said plurality of backward connecting links use a plurality of buffers to amplify signals driven through them; and said buffers are either inverting or non-inverting buffers.

48. The two-dimensional layout of claim 1, wherein said all switches of size $d \times d$ are either fully populated or partially populated.

WO 2011/047368

PCT/US2010/052984

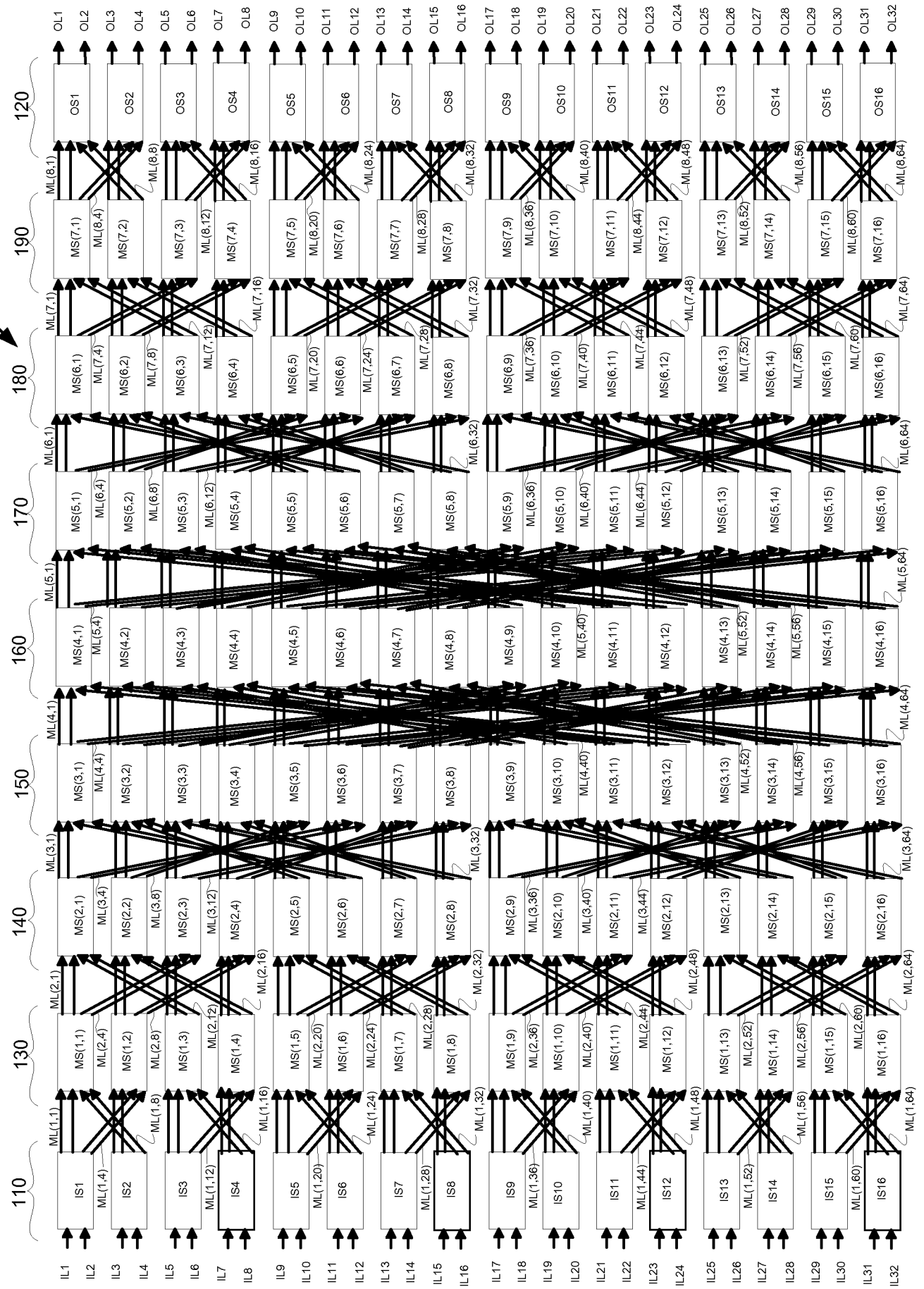


FIG. 1A

FIG. 1B

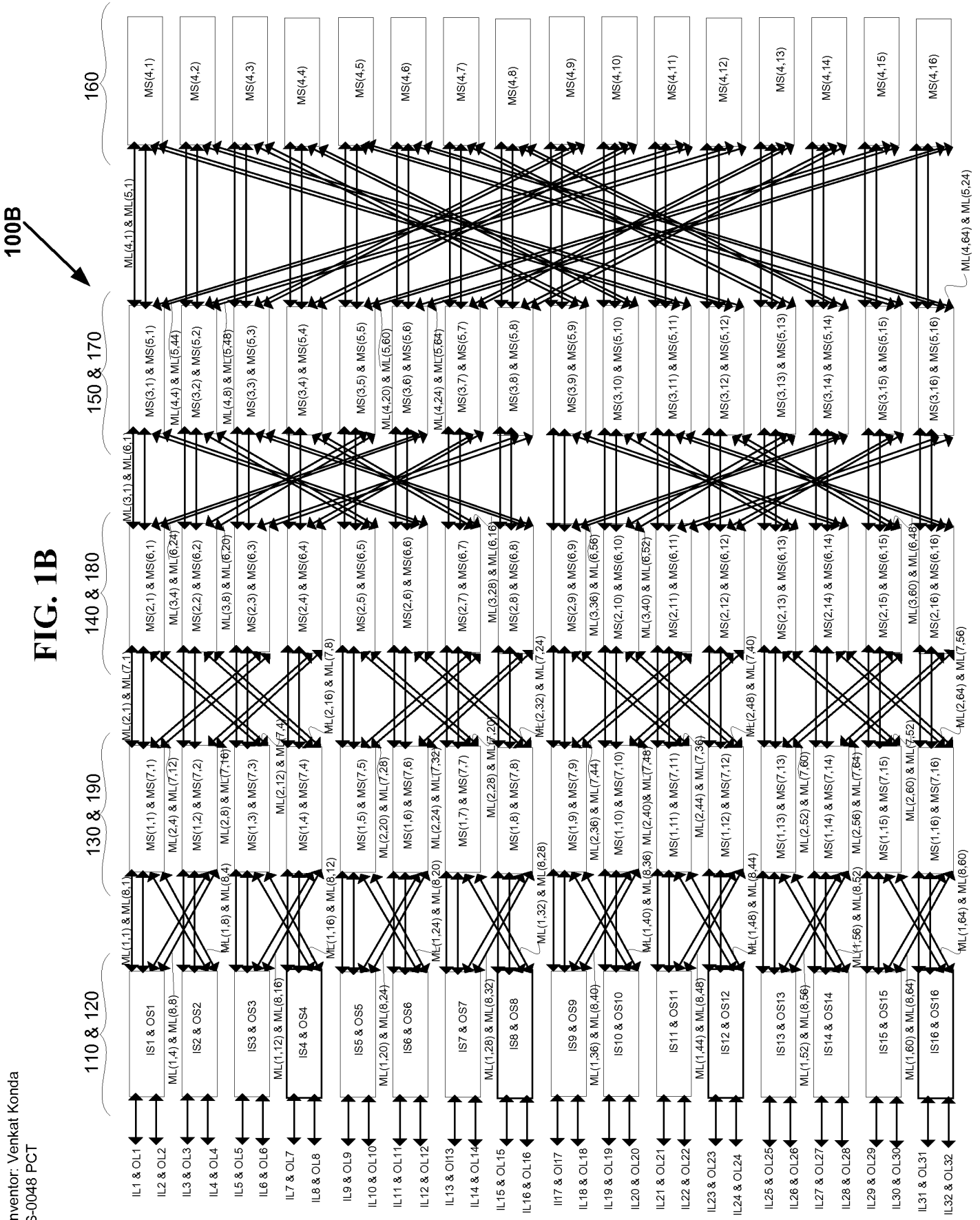
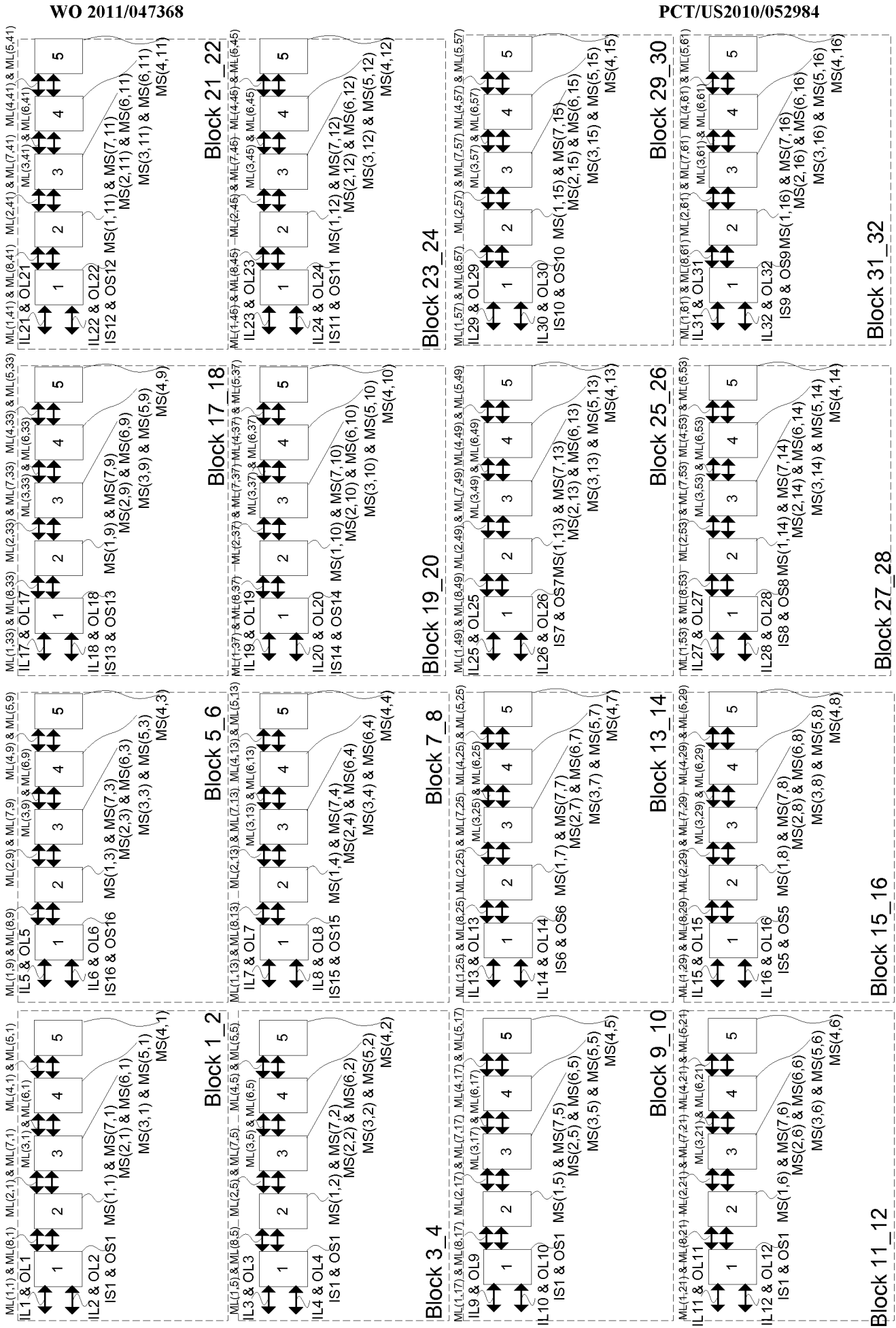


FIG. 1C



100C

100D

FIG. 1D

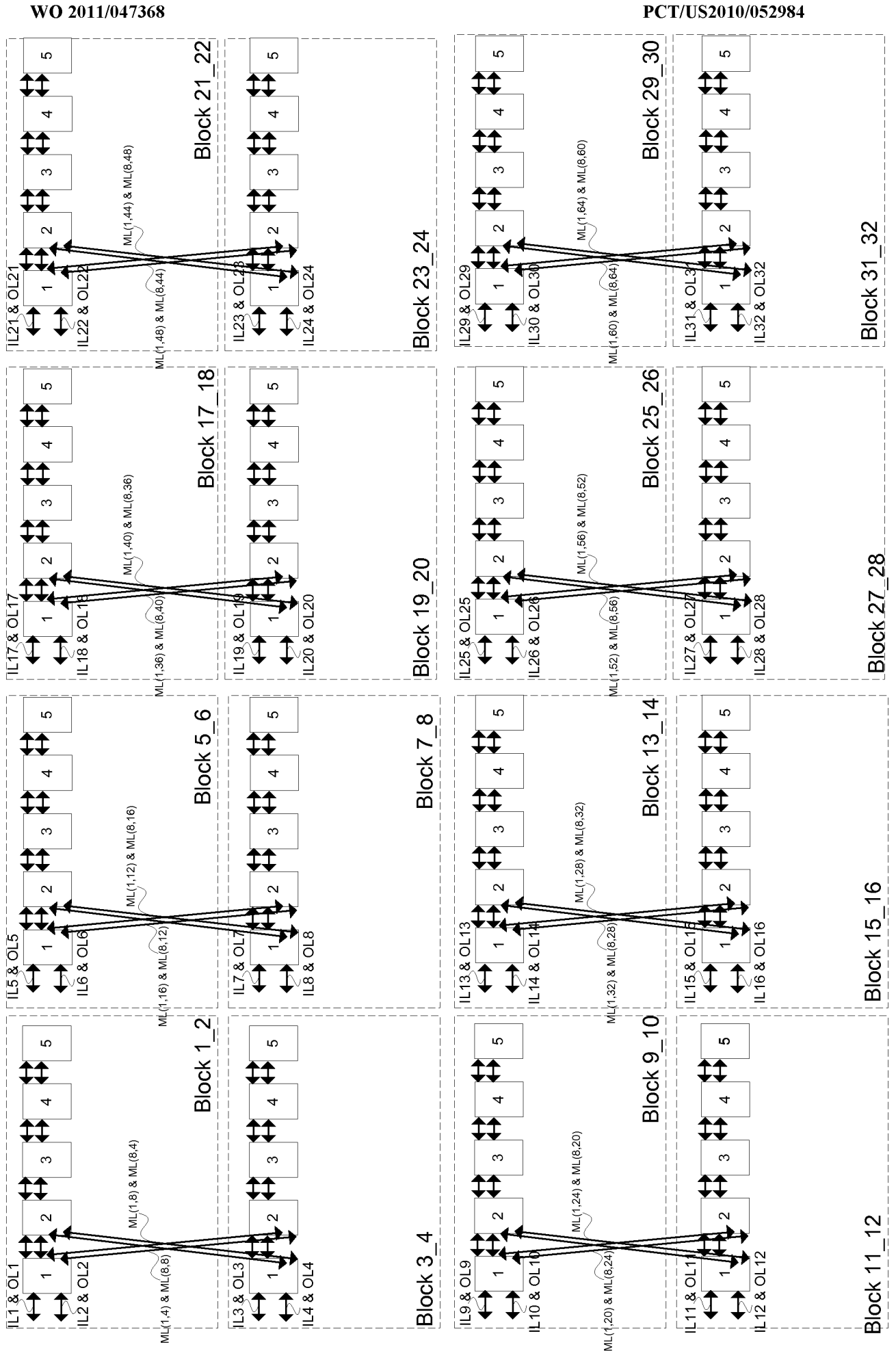
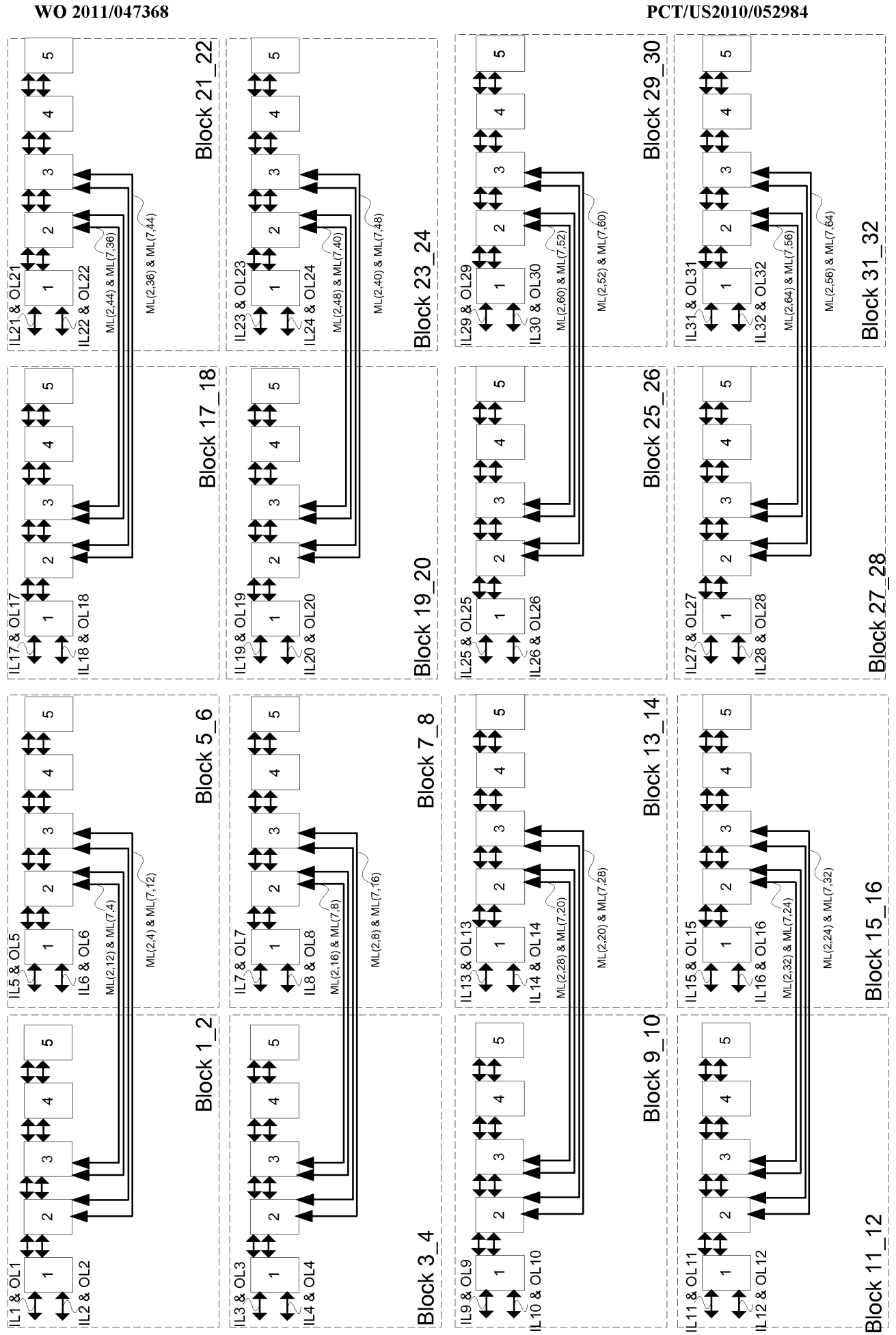


FIG. 1E



WO 2011/047368

PCT/US2010/052984

FIG. 1F

100F

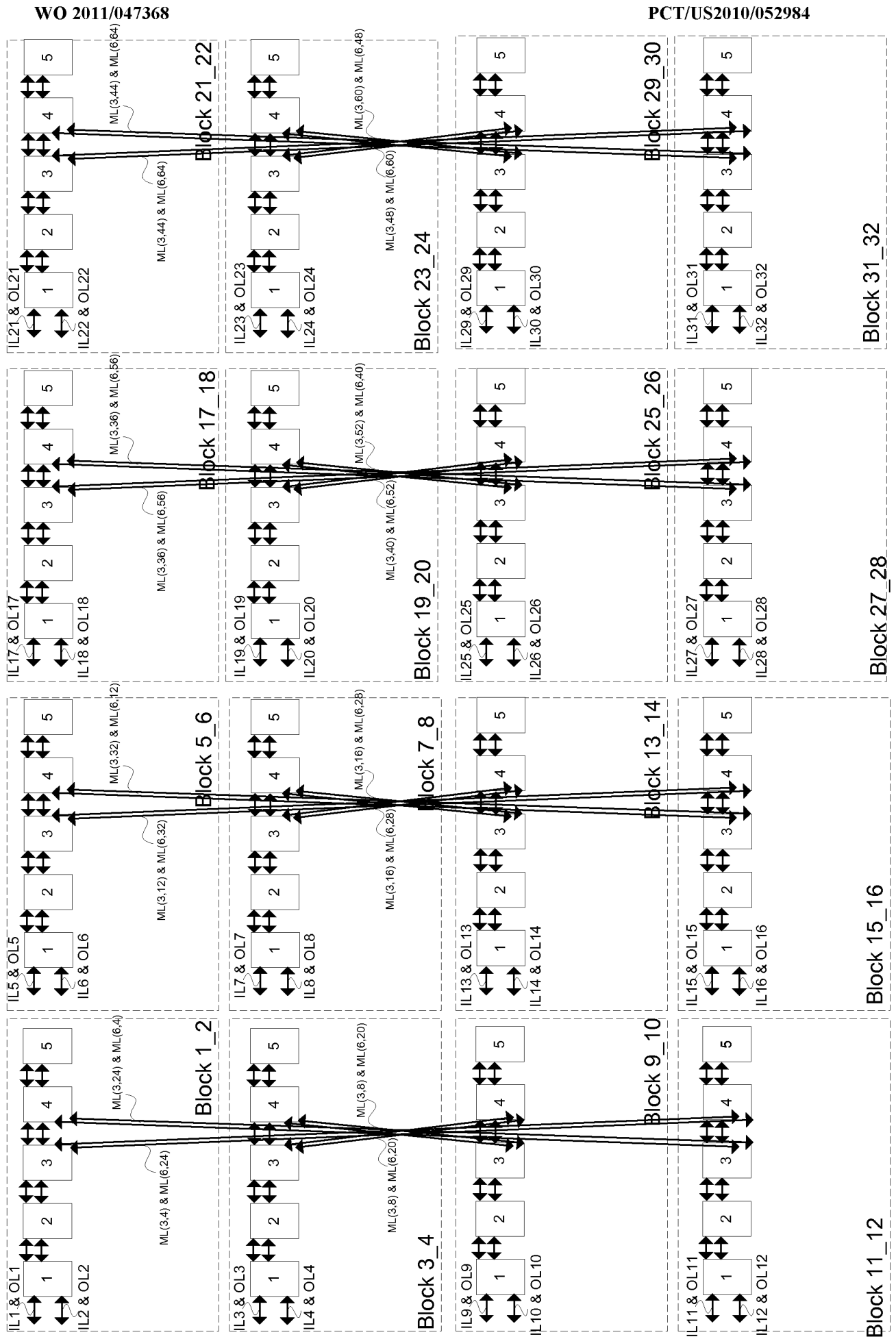
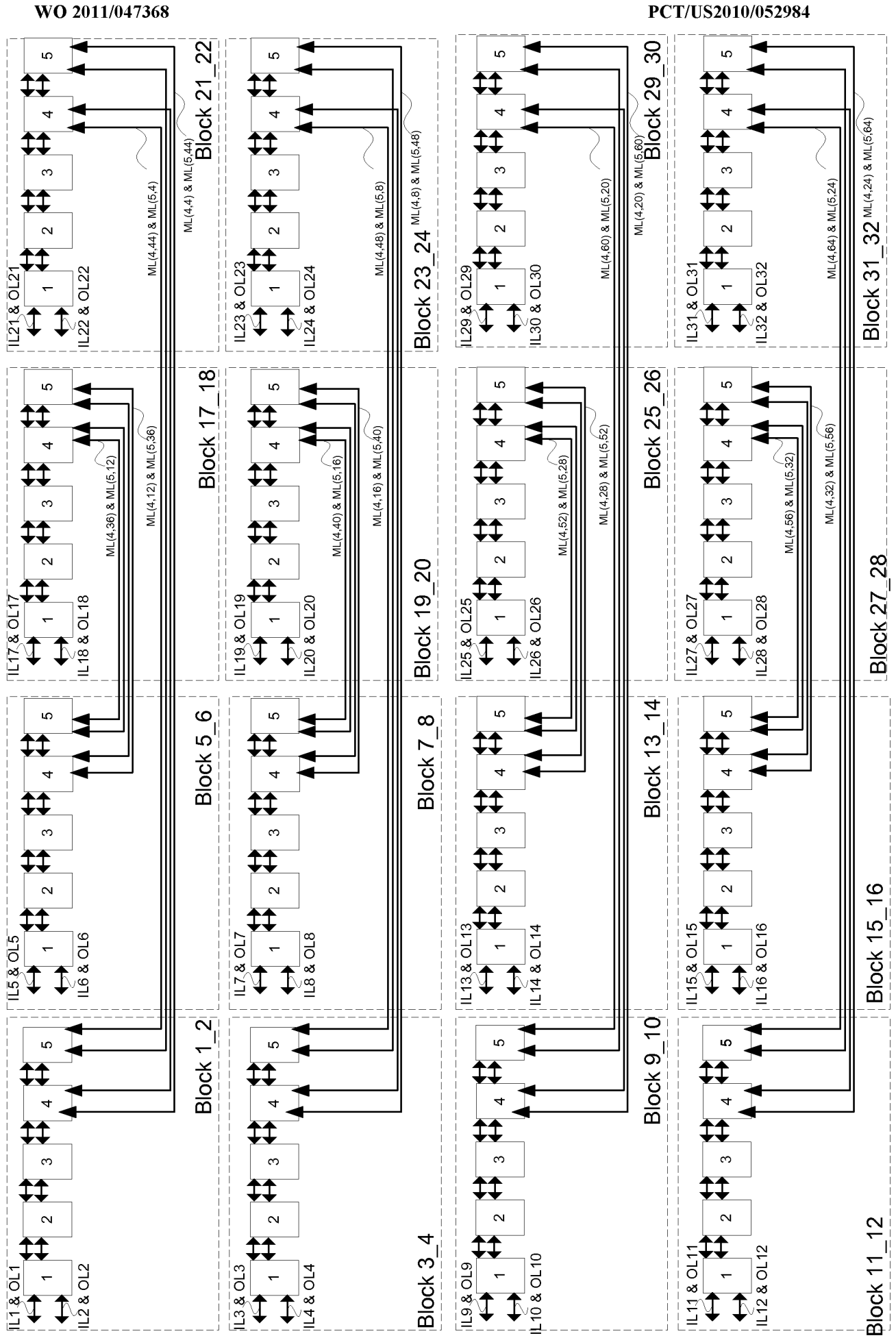


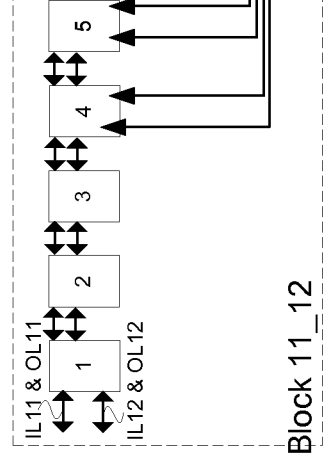
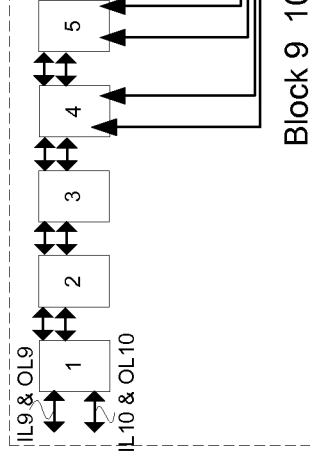
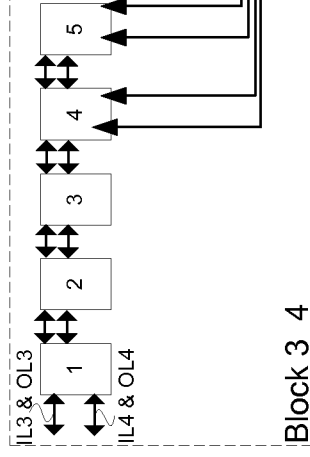
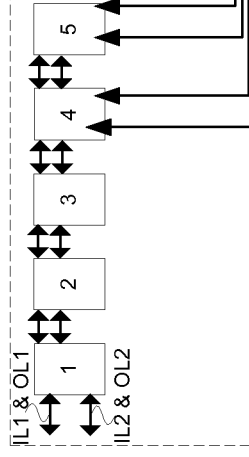
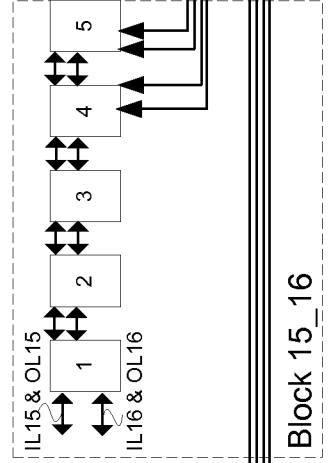
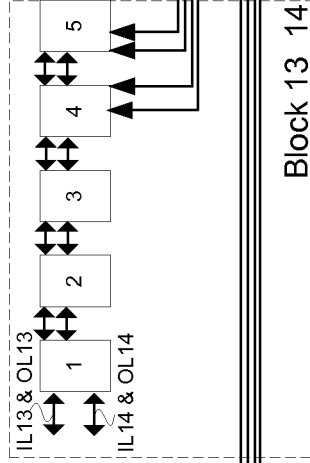
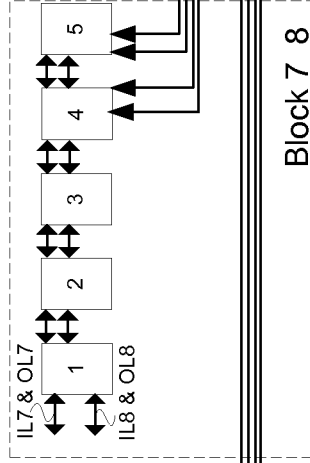
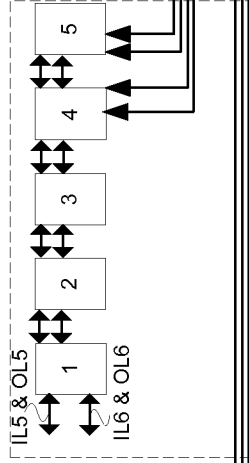
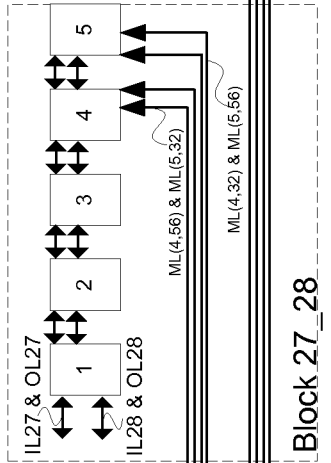
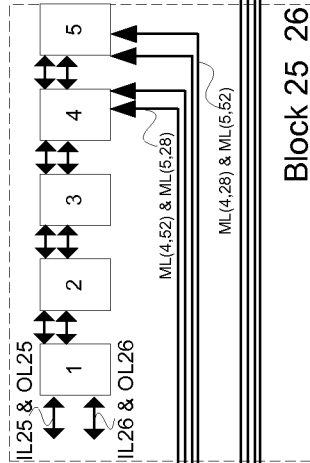
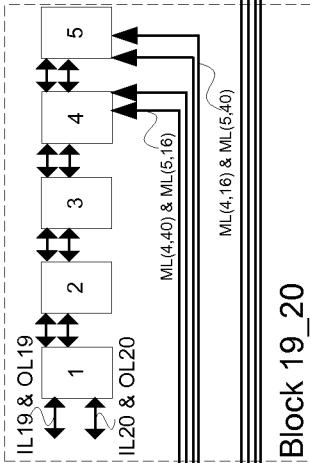
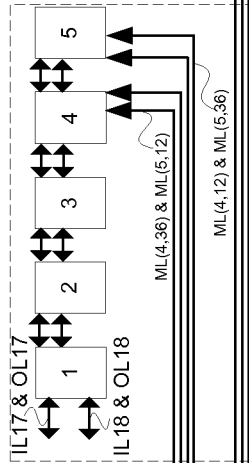
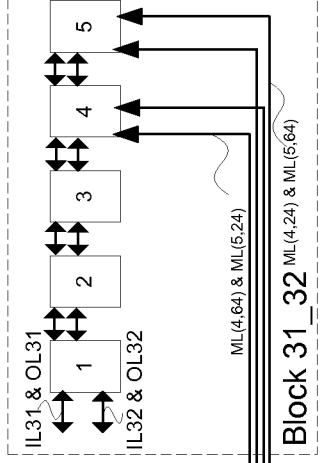
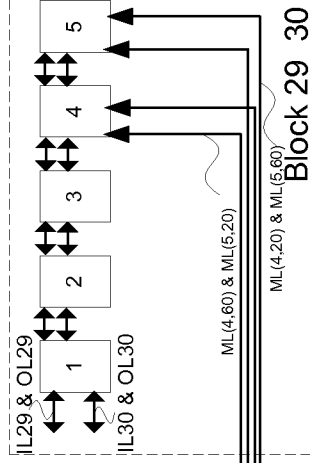
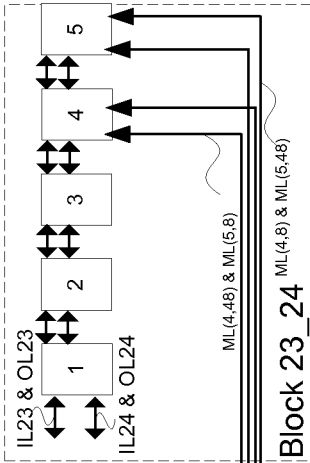
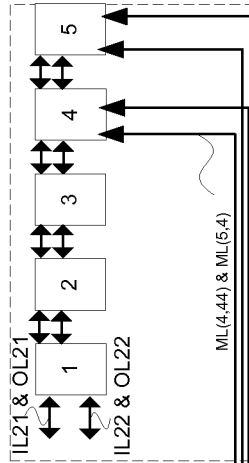
FIG. 1G

100G



WO 2011/047368

PCT/US2010/052984

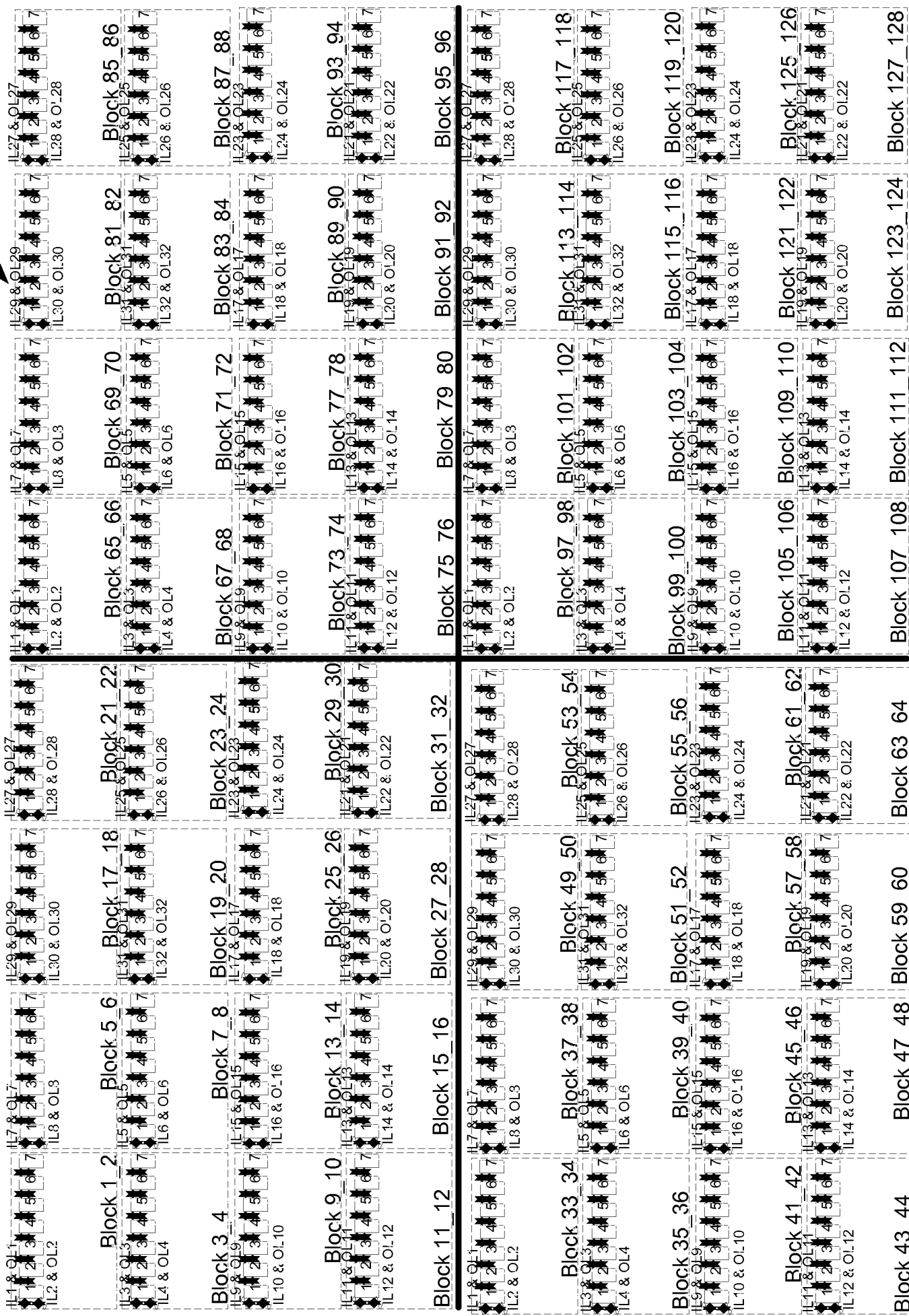


WO 2011/047368

PCT/US2010/052984

100H

FIG. 1H



1001

FIG. 11

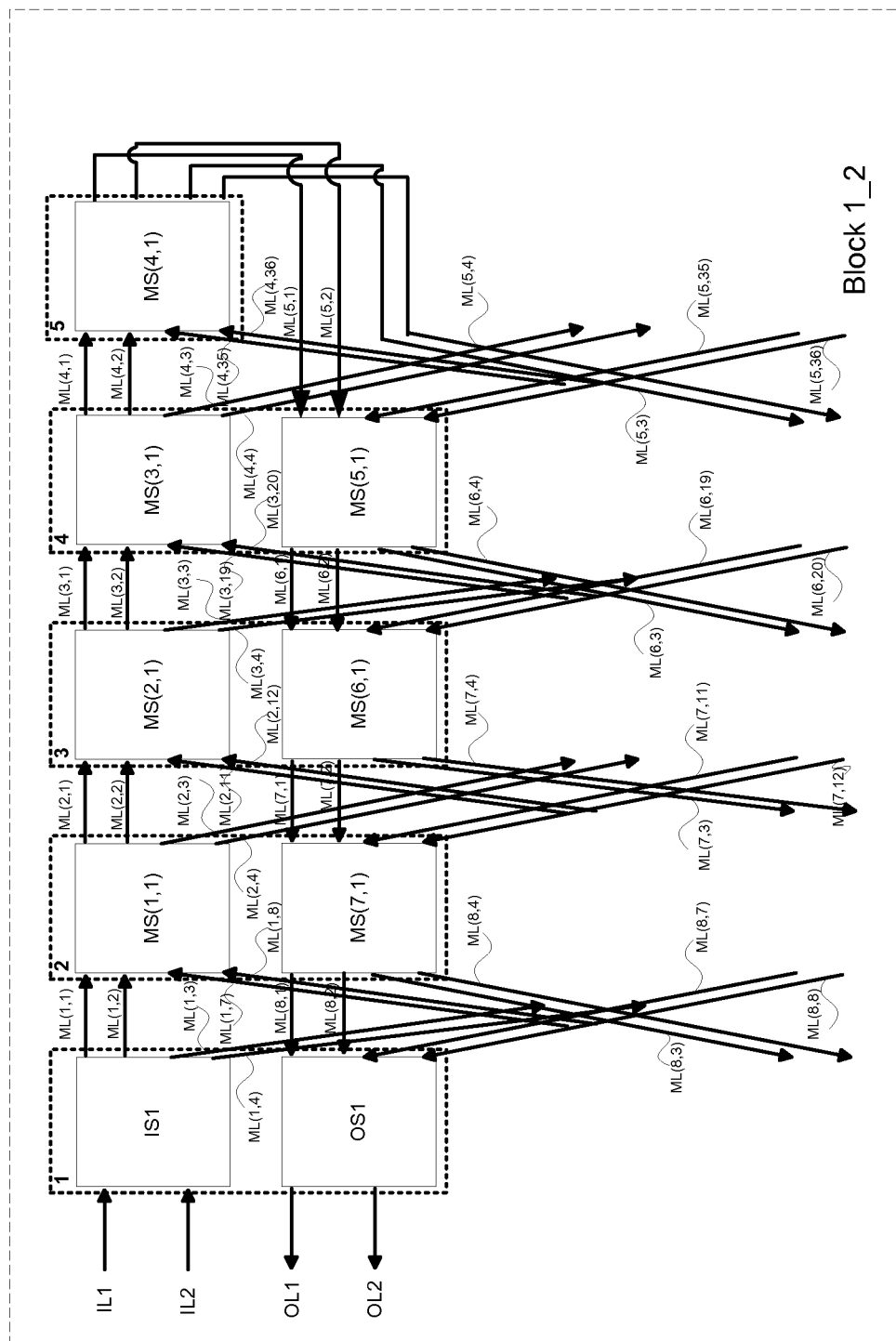
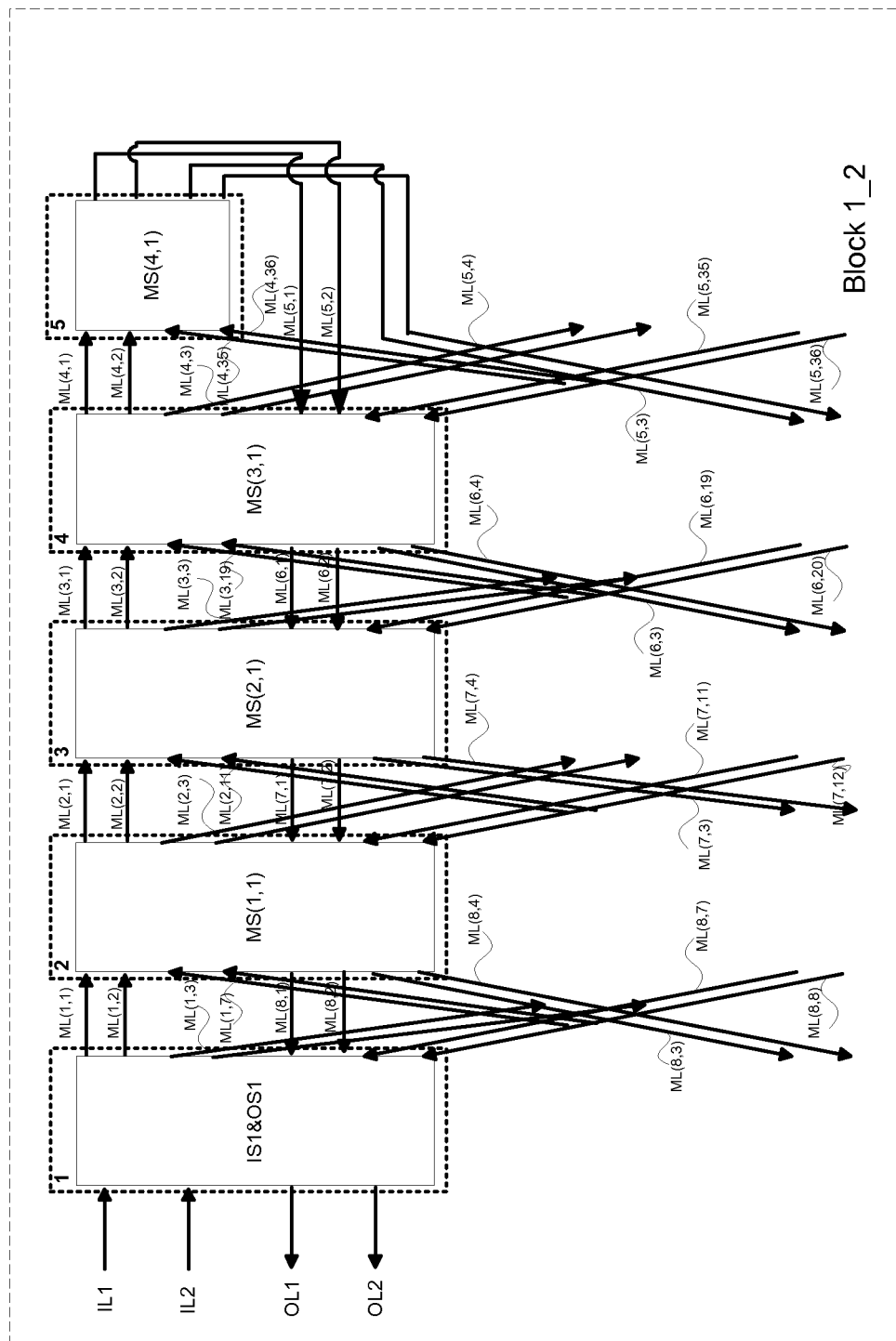


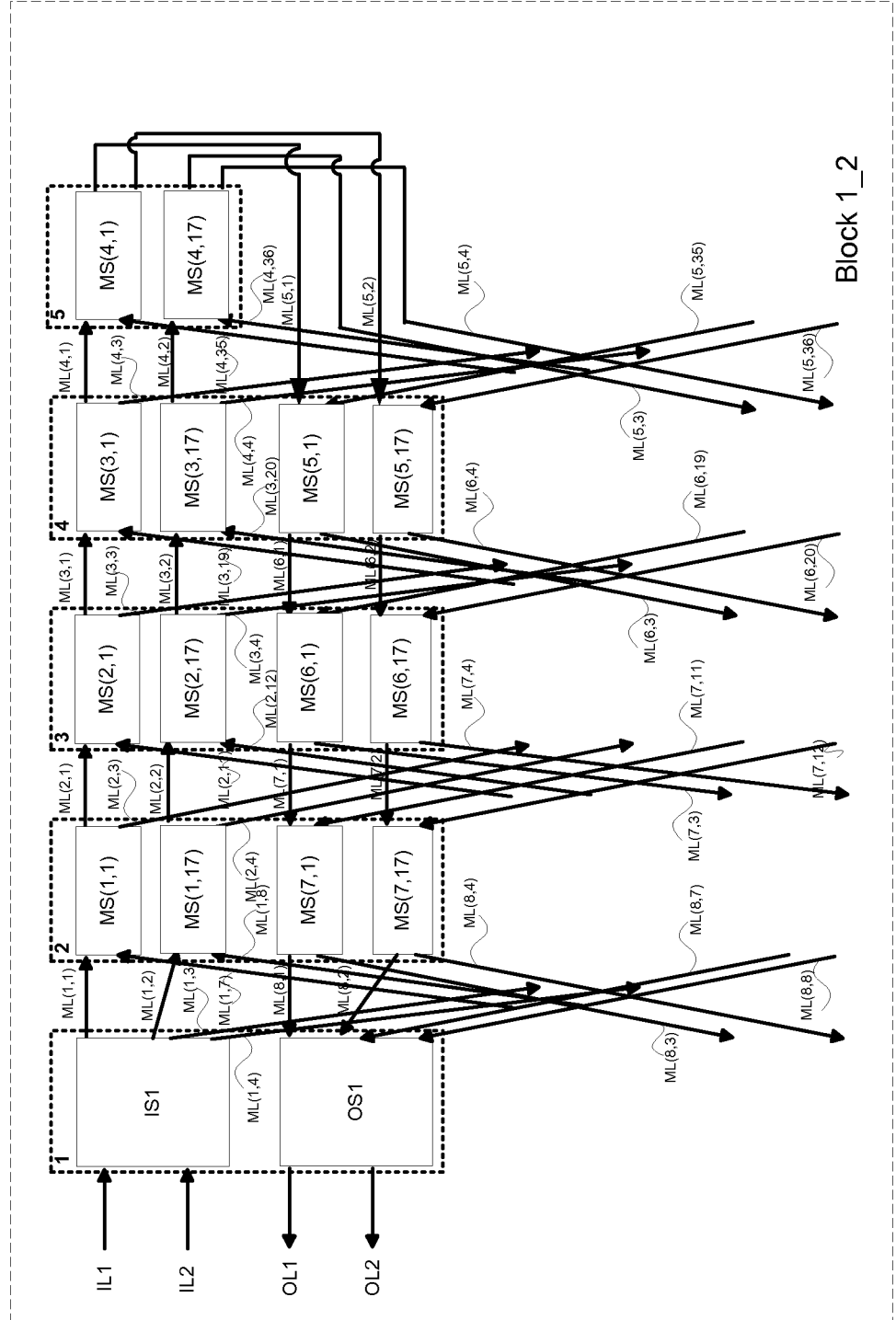
FIG. 1J

100J



100K

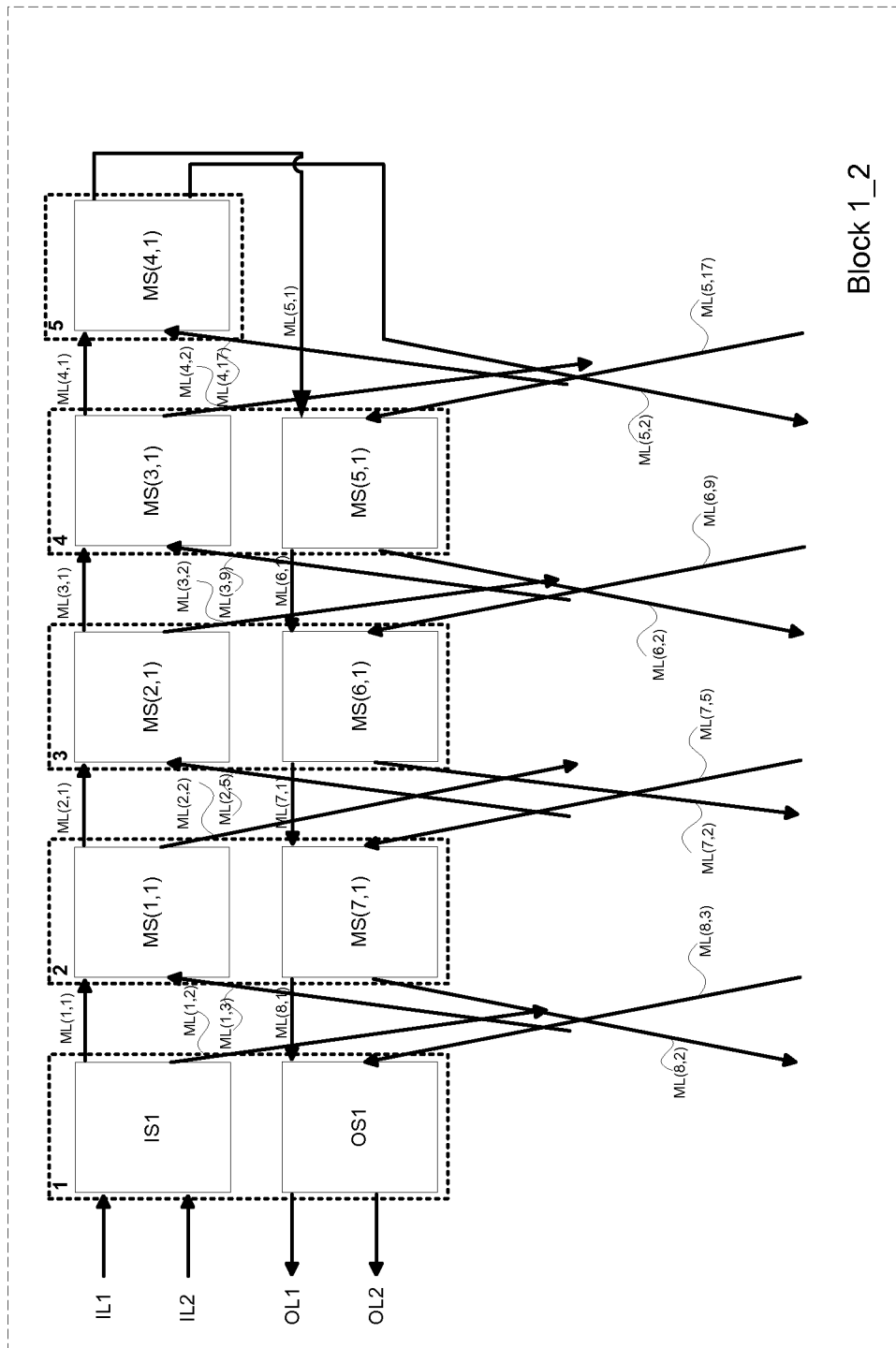
FIG. 1K



Block 1_2

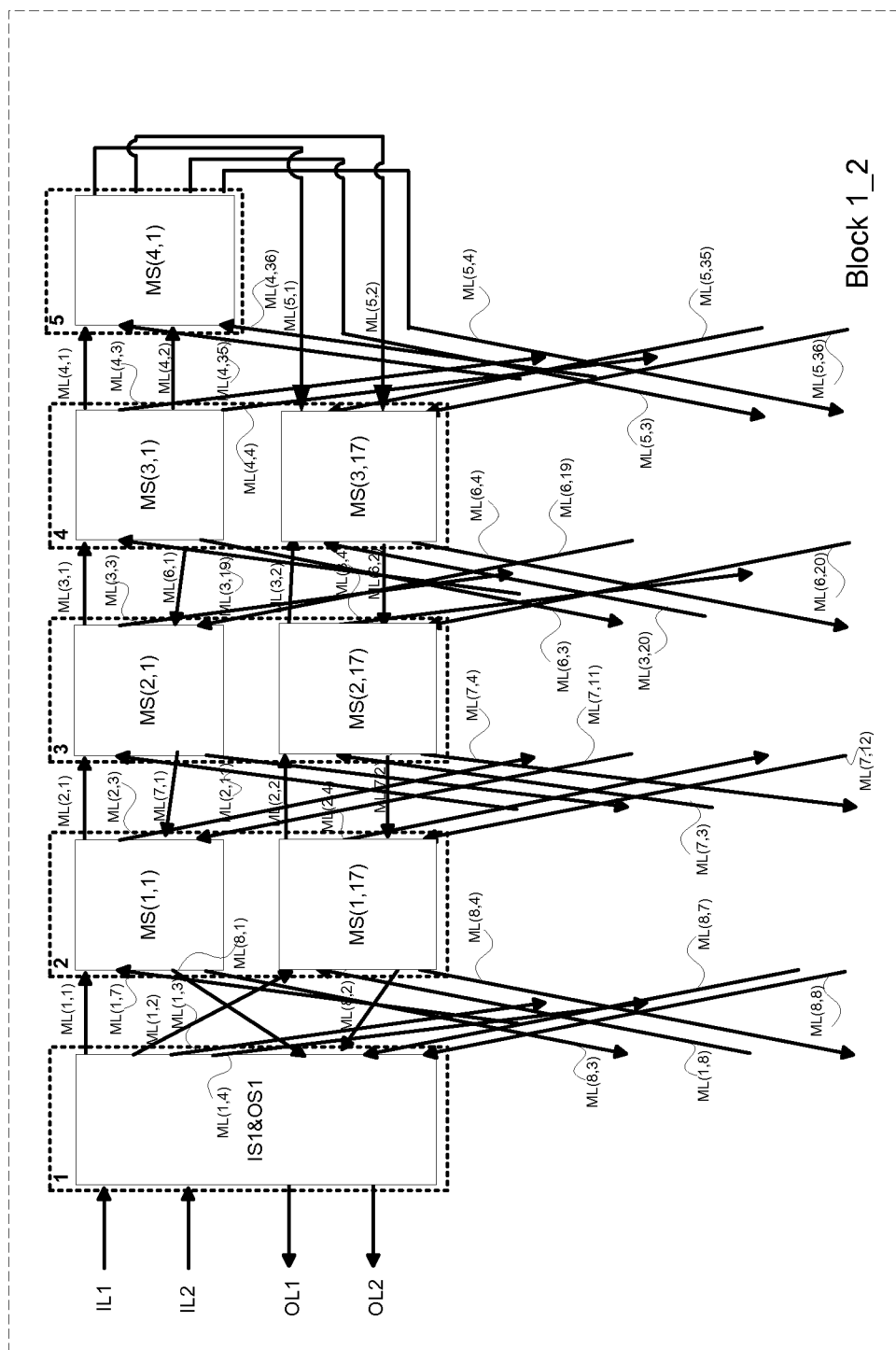
100K1

FIG. 1K1



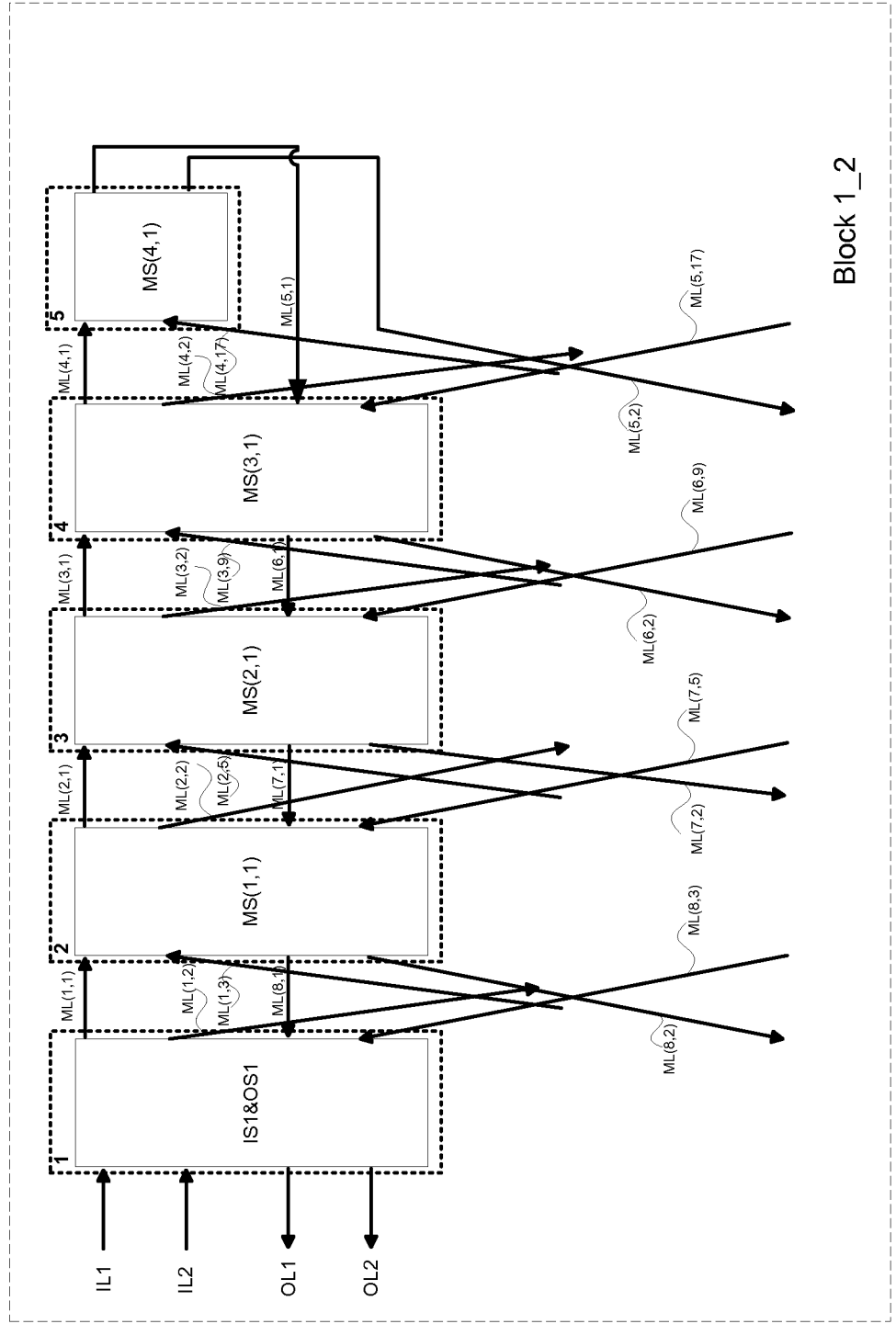
100L

FIG. 1L



100L1

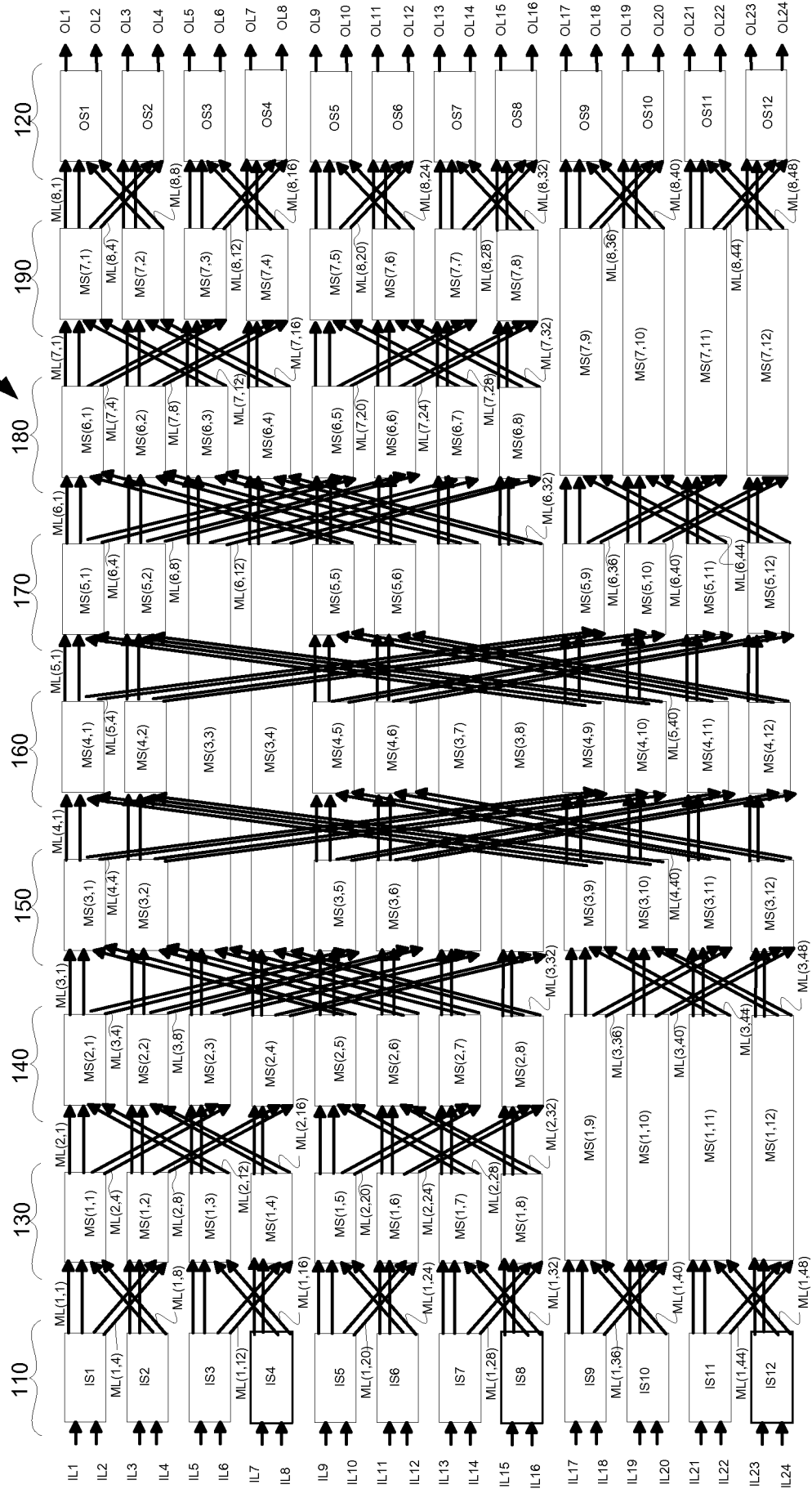
FIG. 1L1



VLSI Layouts of Fully Connected Generalized and pyramid Networks with Locality Exploitation
Inventor: Venkat Konda
S-0048 PCT

FIG. 2A

200A



VLSI Layouts of Fully Connected Generalized and pyramid Networks with Locality Exploitation
Inventor: Venkat Konda
S-0048 PCT

FIG. 2B

200B

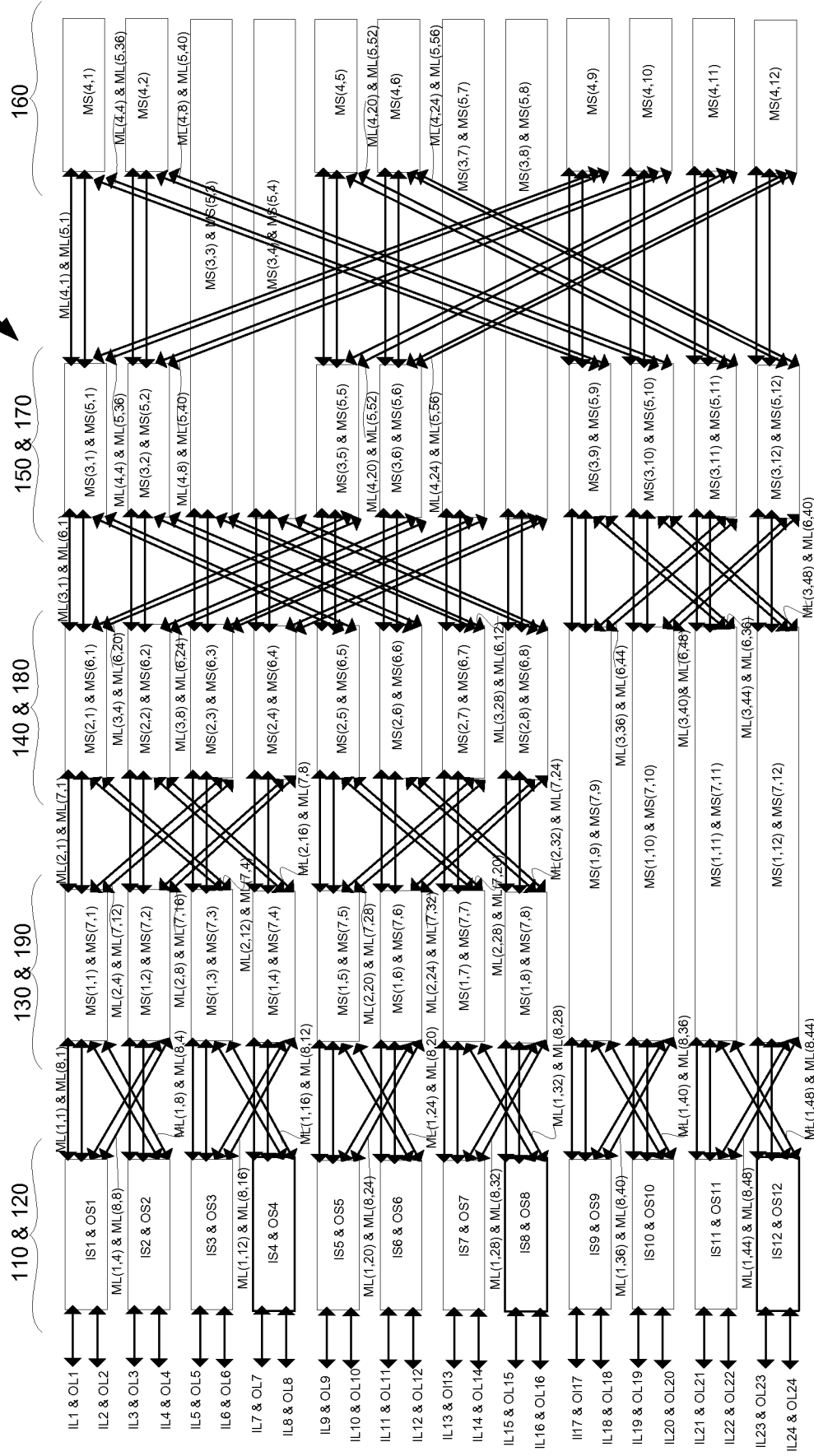




FIG. 2C

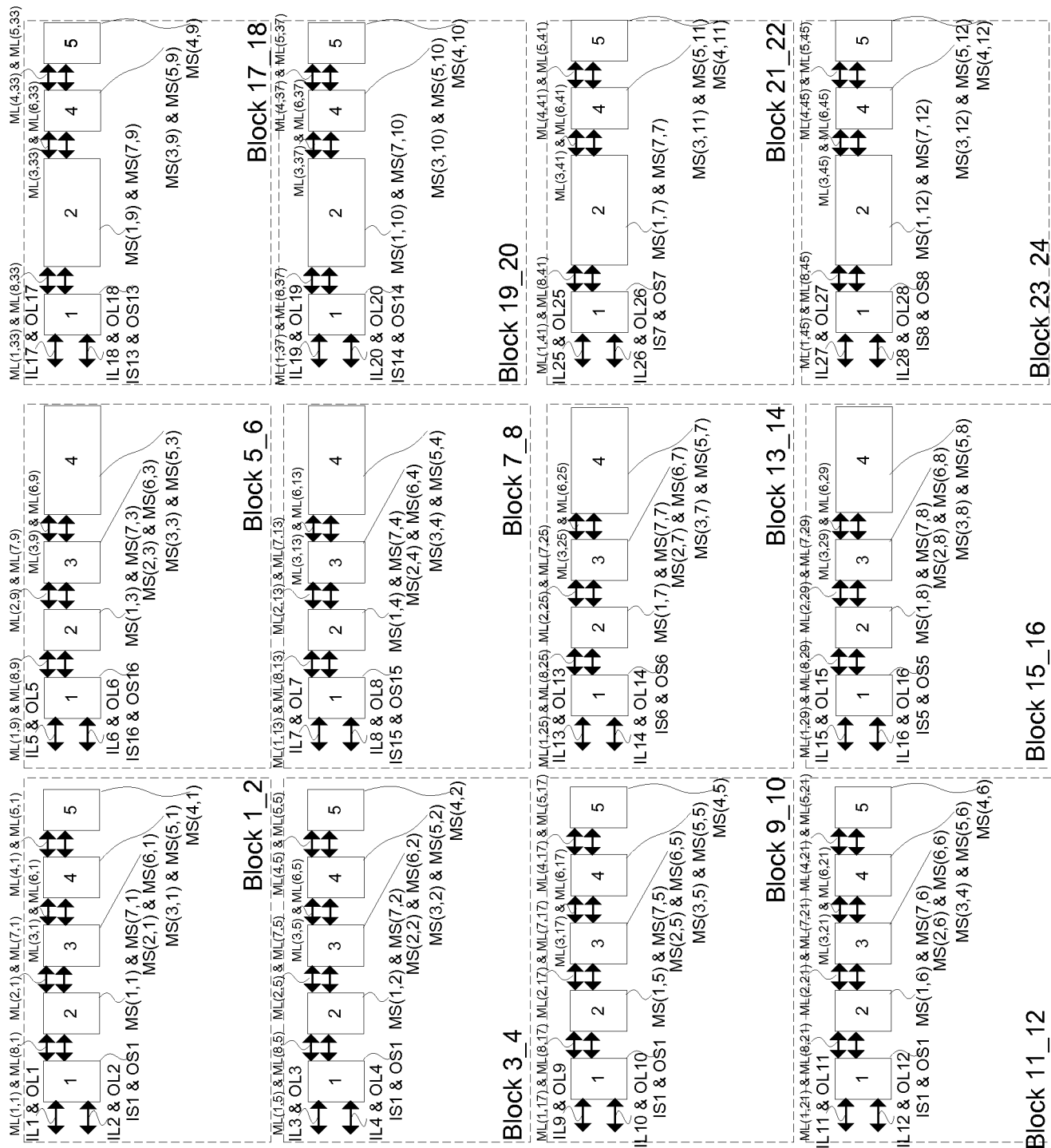




FIG. 2D

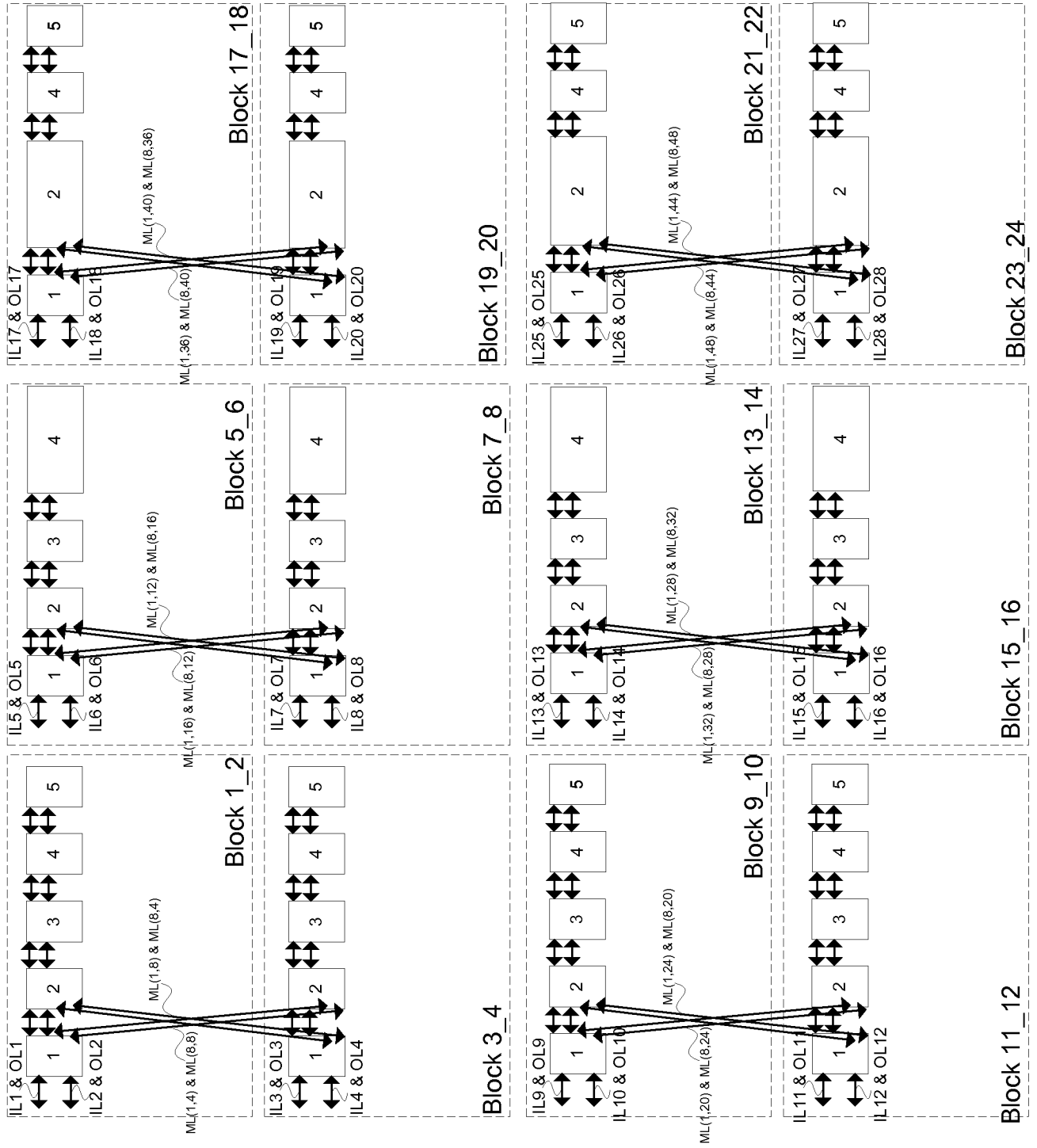
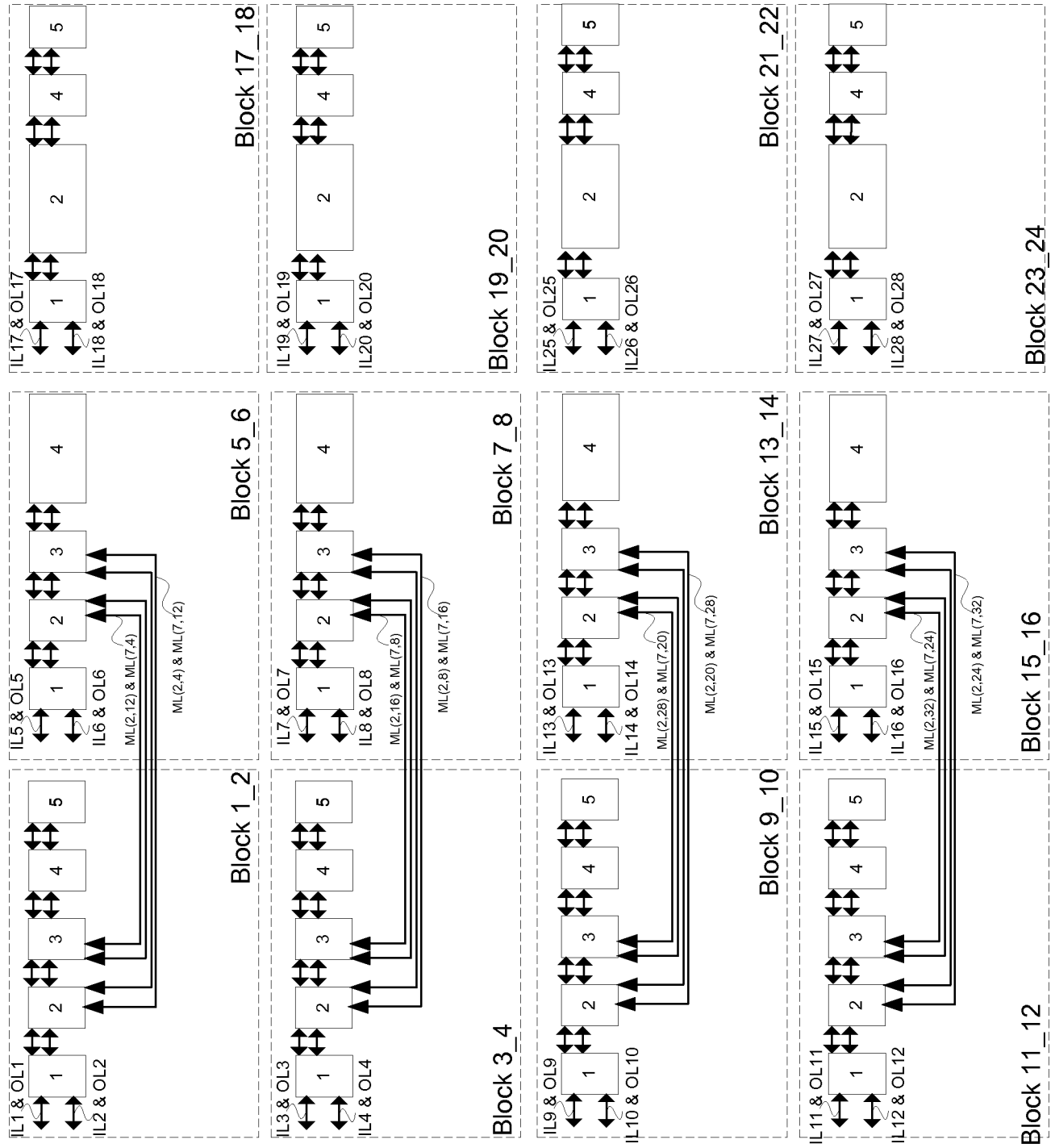


FIG. 2E



200F

FIG. 2F

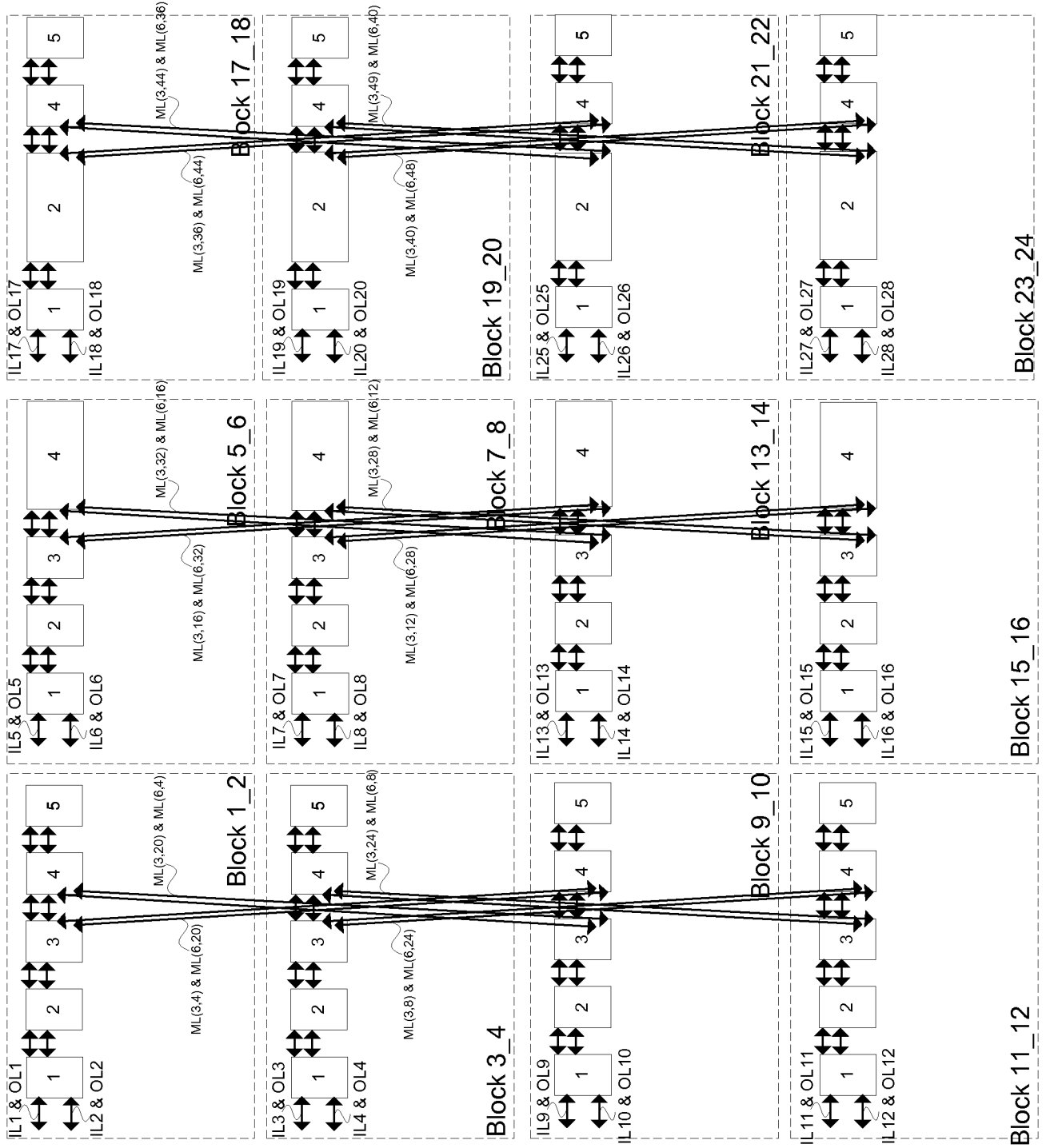
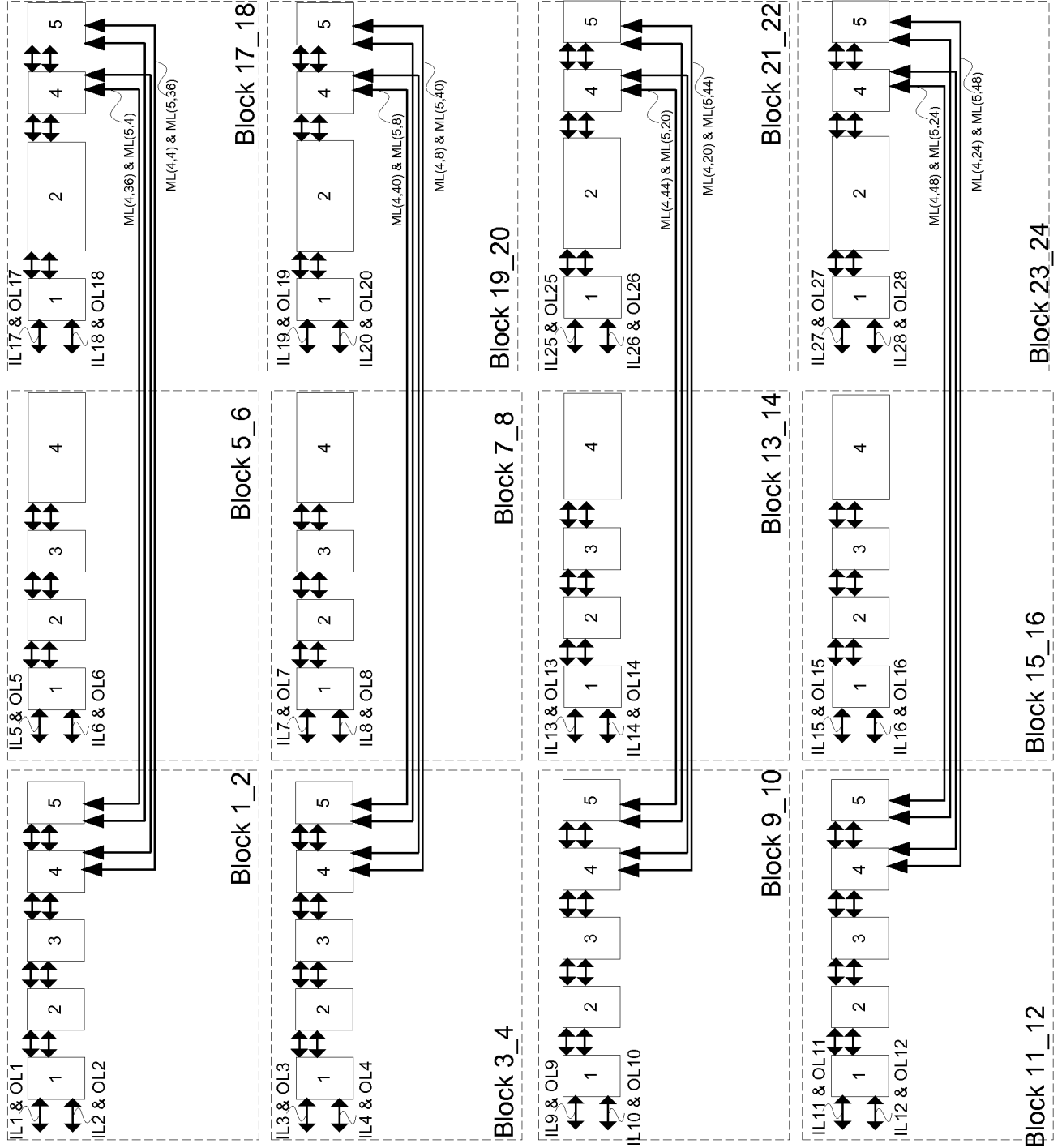




FIG. 2G



WO 2011/047368

PCT/US2010/052984

300A

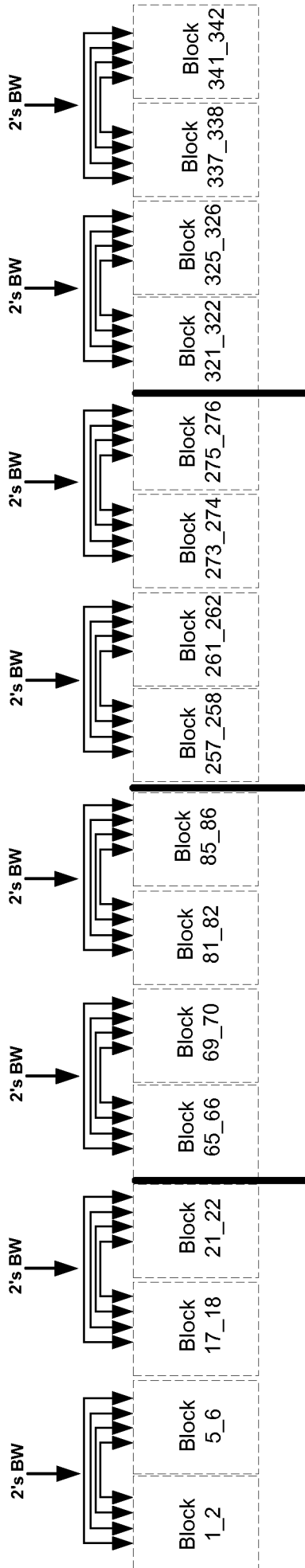


FIG. 3A

300B

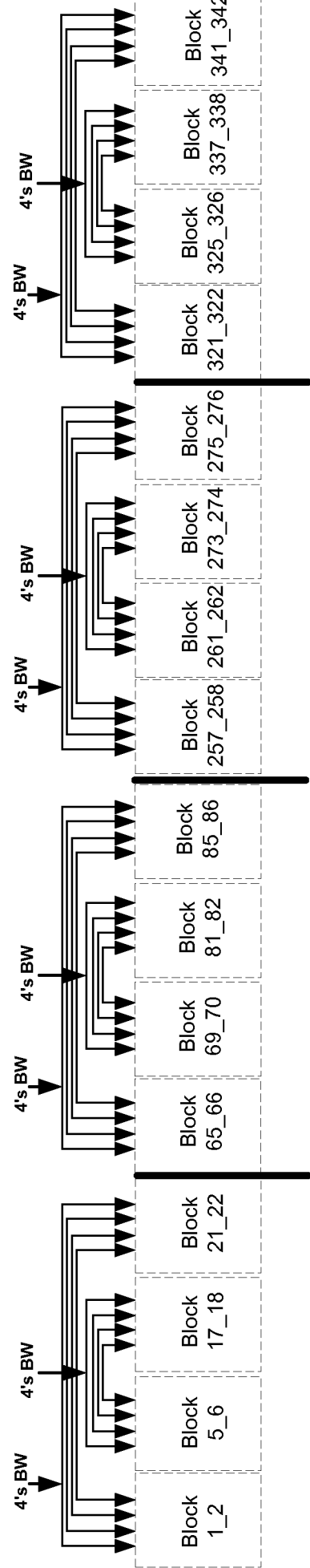


FIG. 3B

FIG. 3C

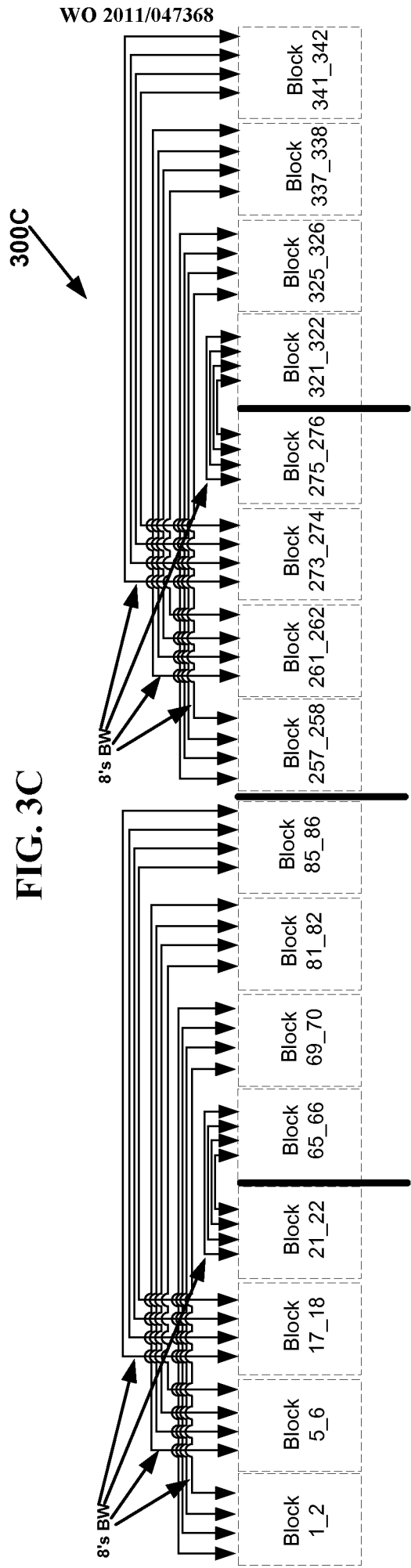
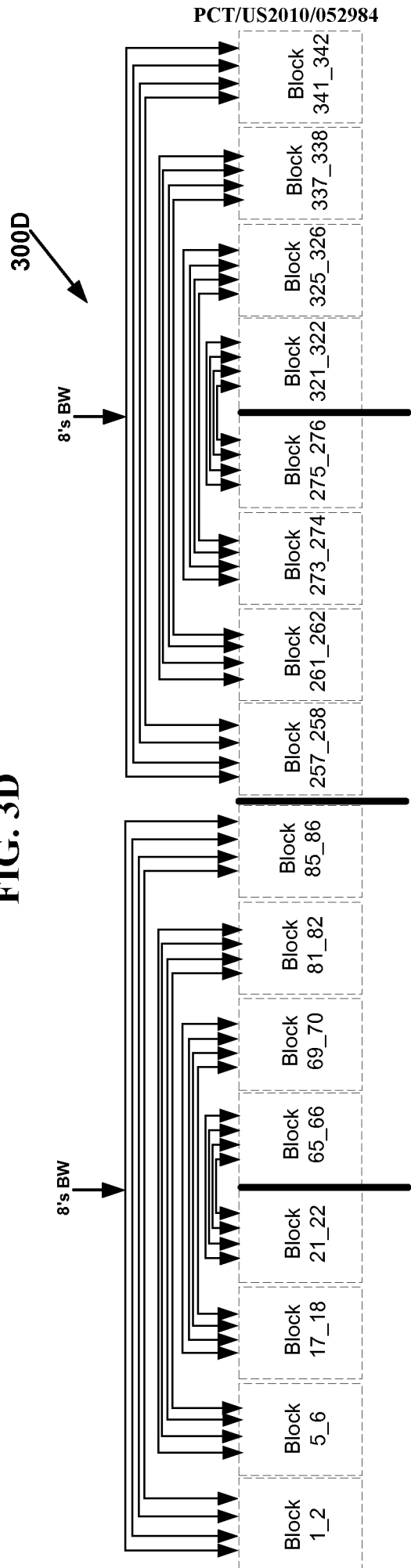


FIG. 3D

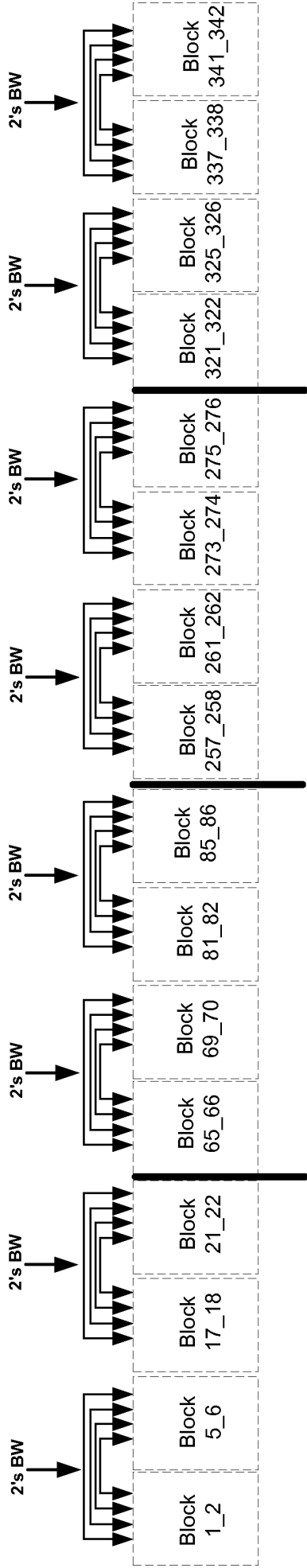


WO 2011/047368

PCT/US2010/052984

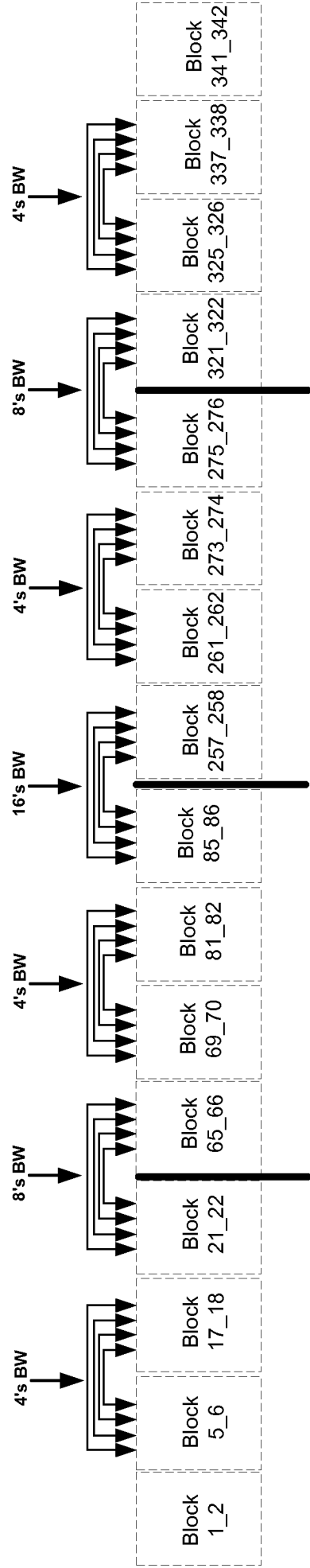
400A

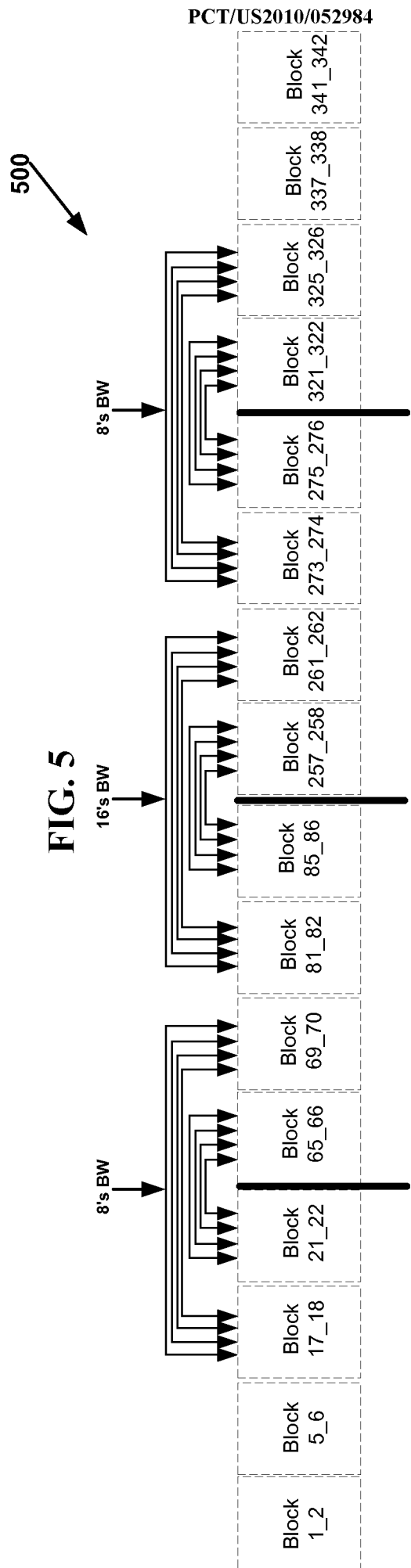
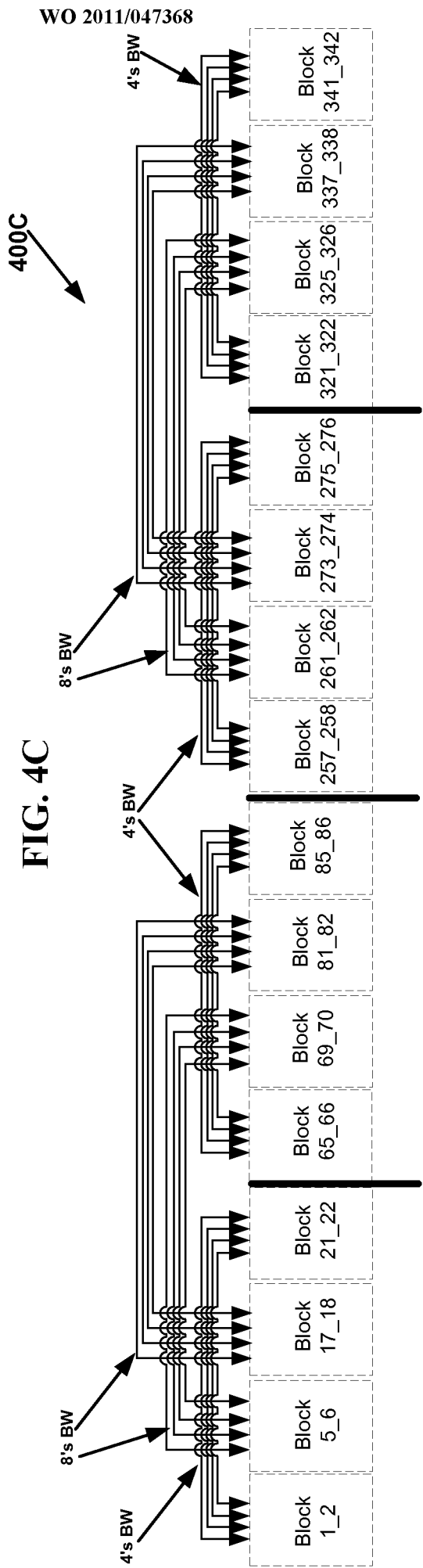
FIG. 4A



400B

FIG. 4B



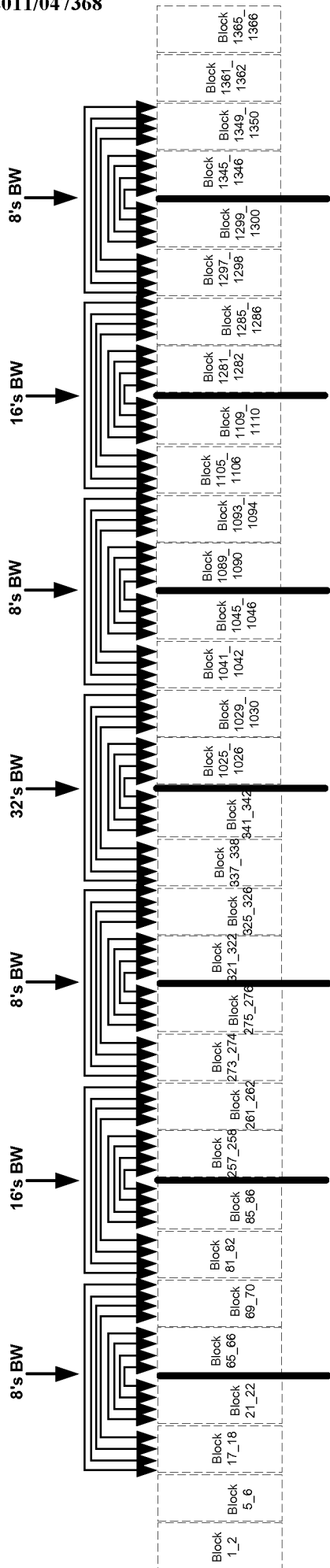


WO 2011/047368

PCT/US2010/052984

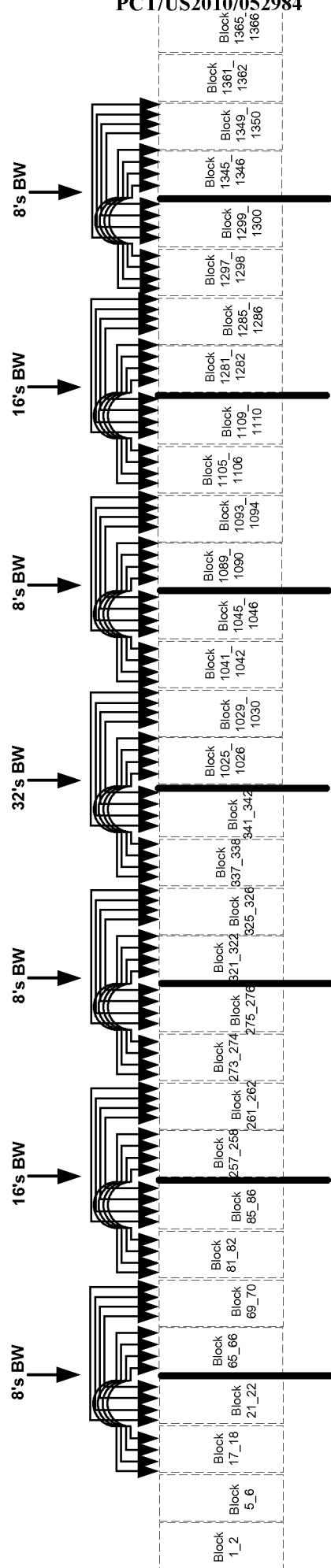
600

FIG. 6



700

FIG. 7



WO 2011/047368

PCT/US2010/052984

VLSI Layouts of Fully Connected Generalized and pyramid Networks with Locality Exploitation

Inventor: Venkat Konda

S-0048 PCT

800A

FIG. 8A

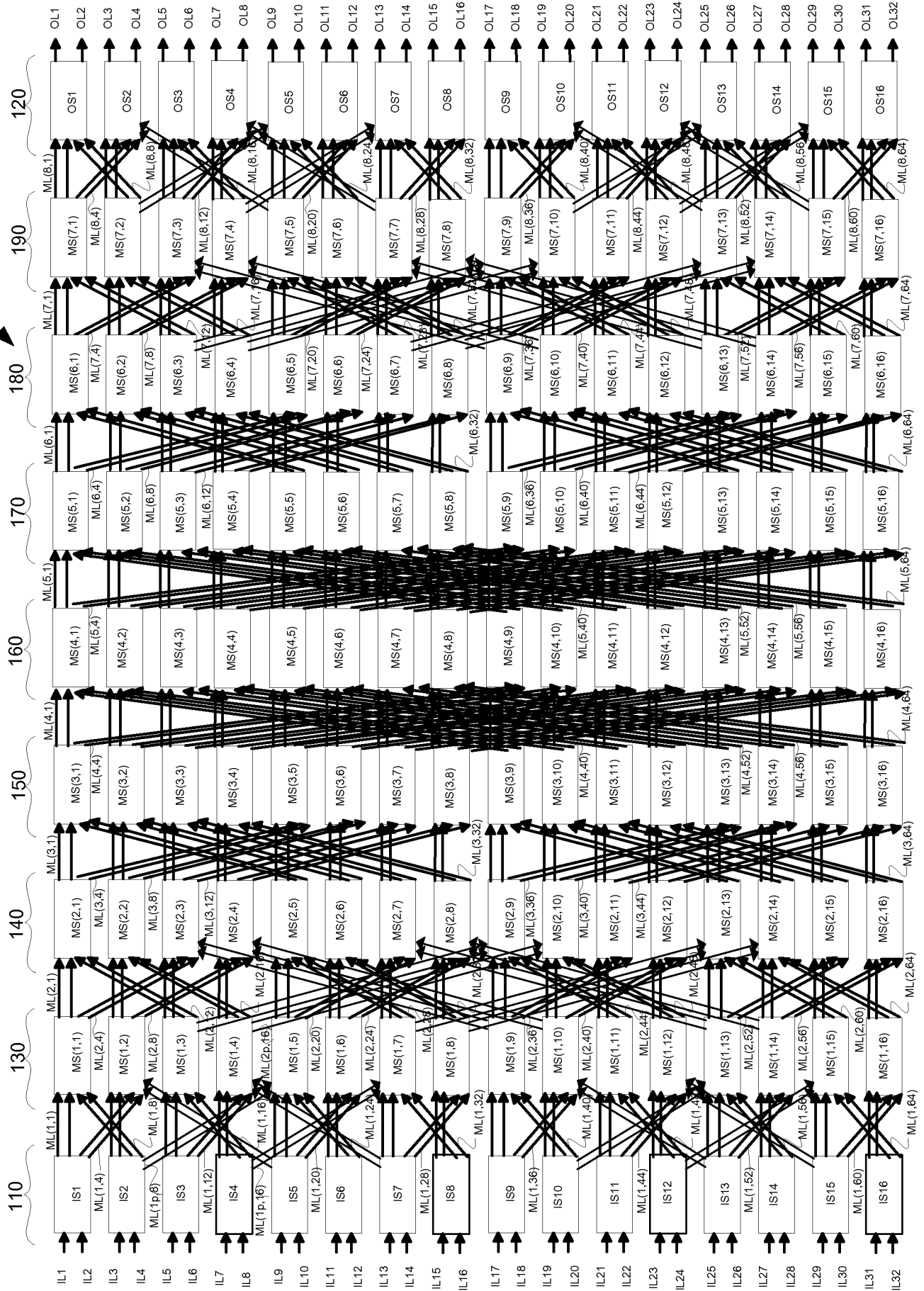


FIG. 8B

800B

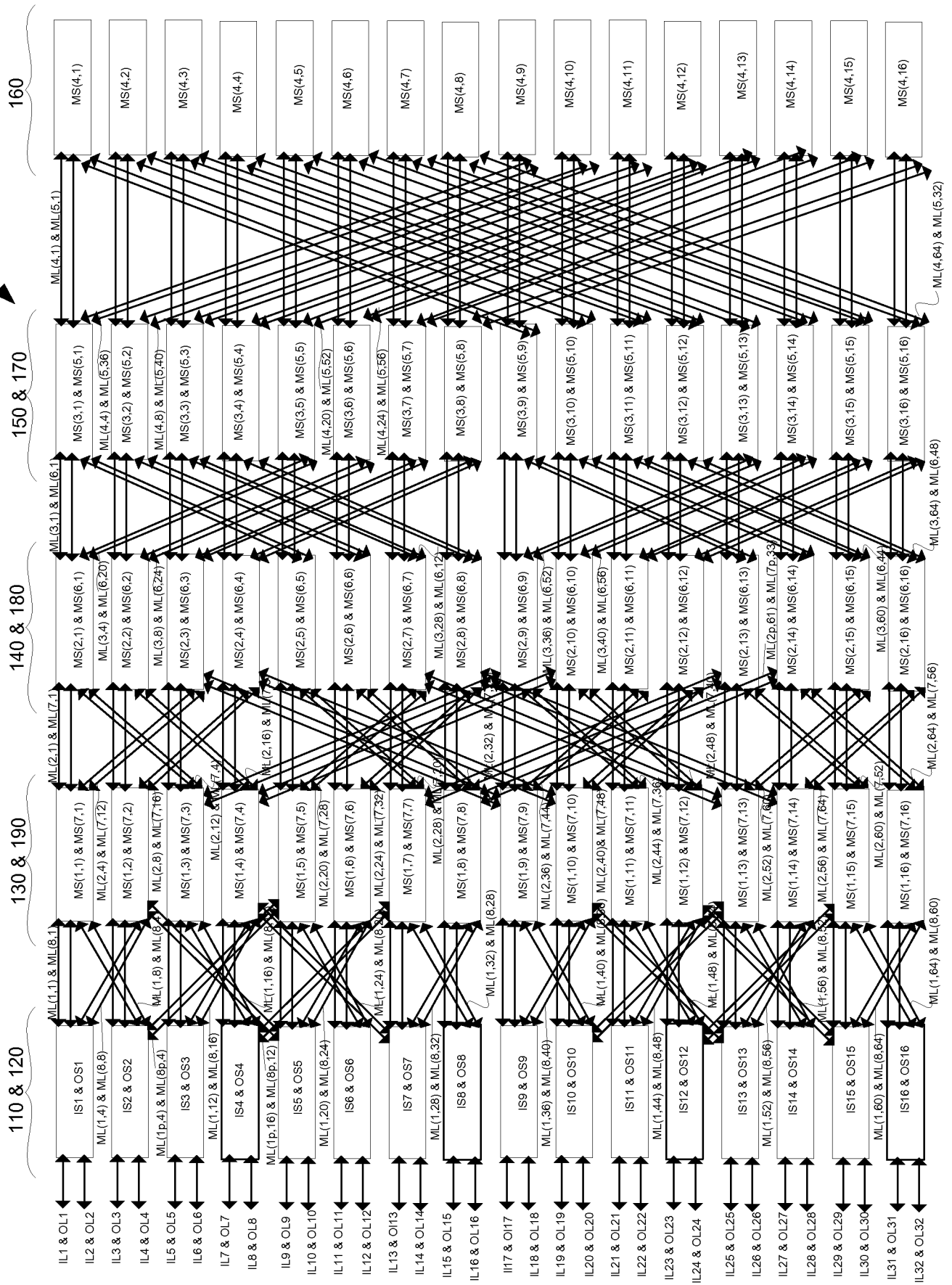
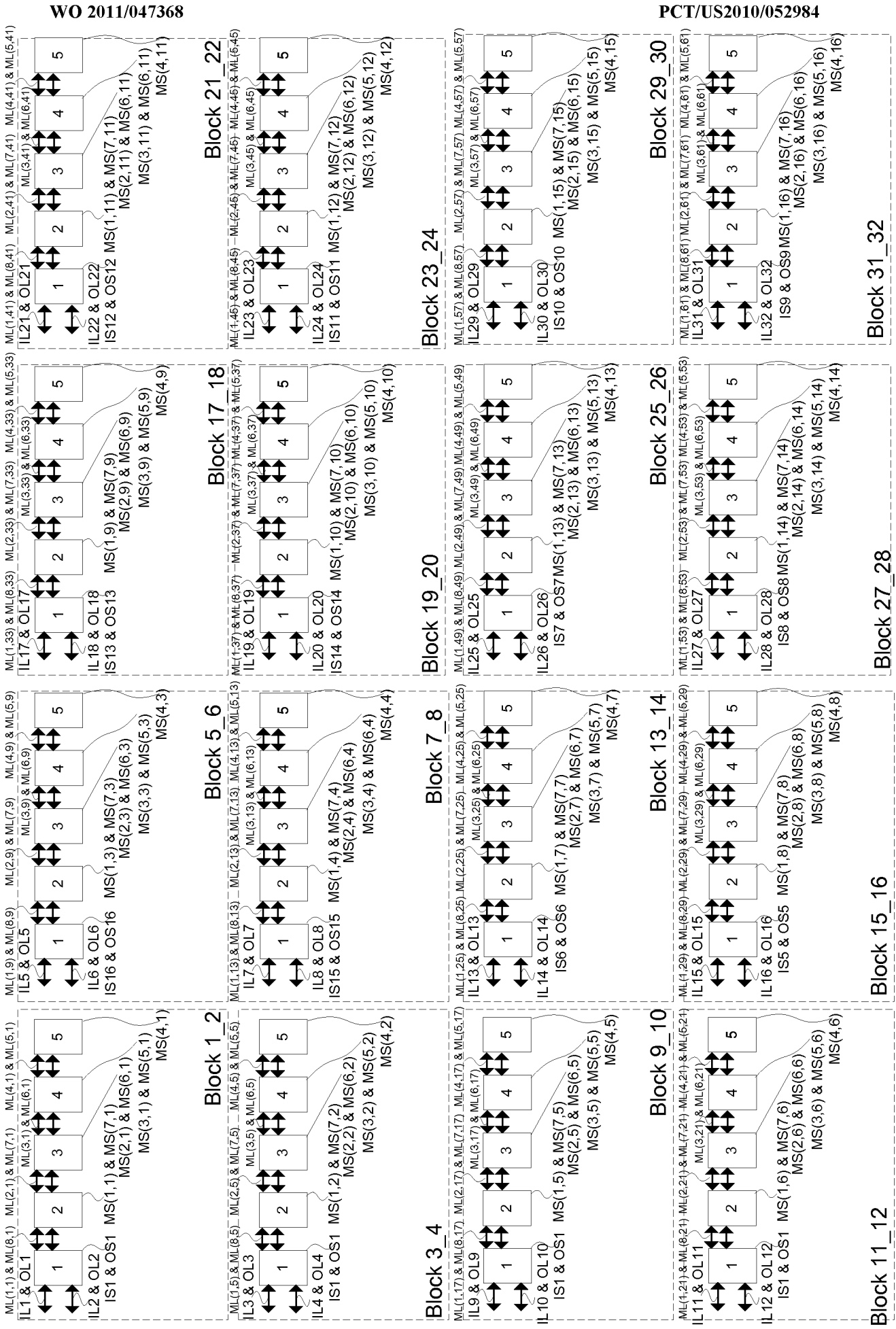


FIG. 8C

800C



WO 2011/047368

PCT/US2010/052984

800D

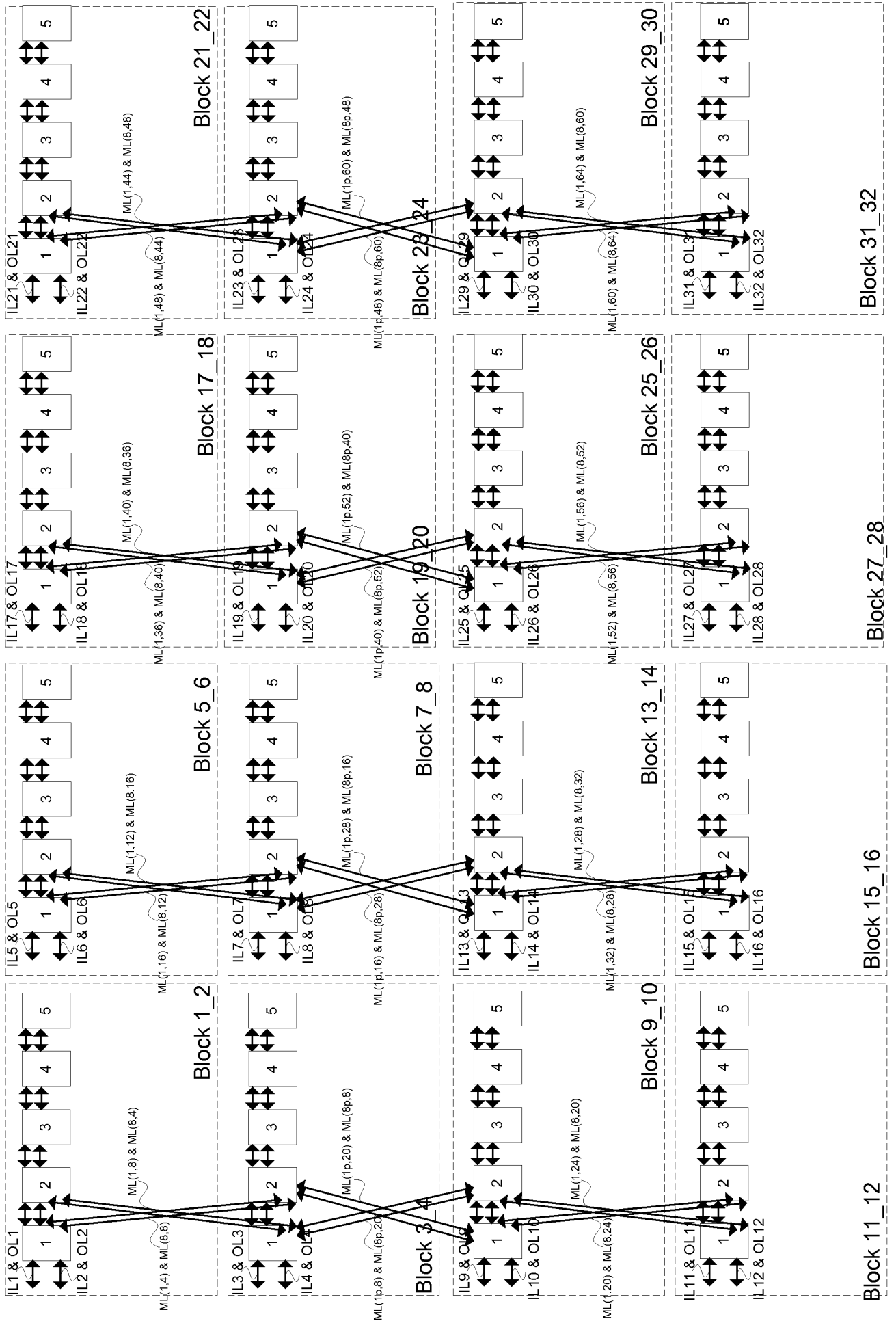


FIG. 8D

WO 2011/047368

PCT/US2010/052984

FIG. 8E

800E

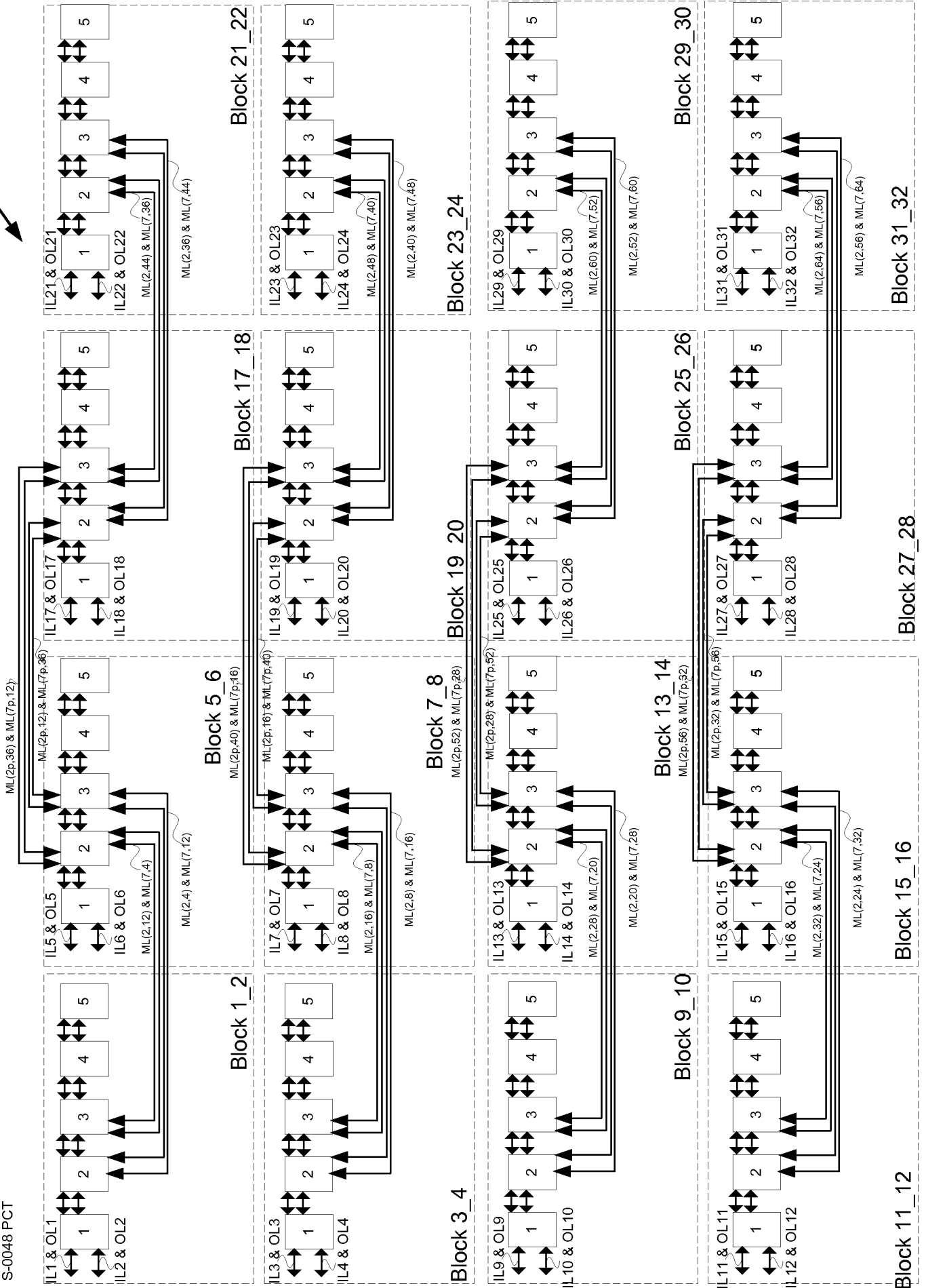


FIG. 8F

800F

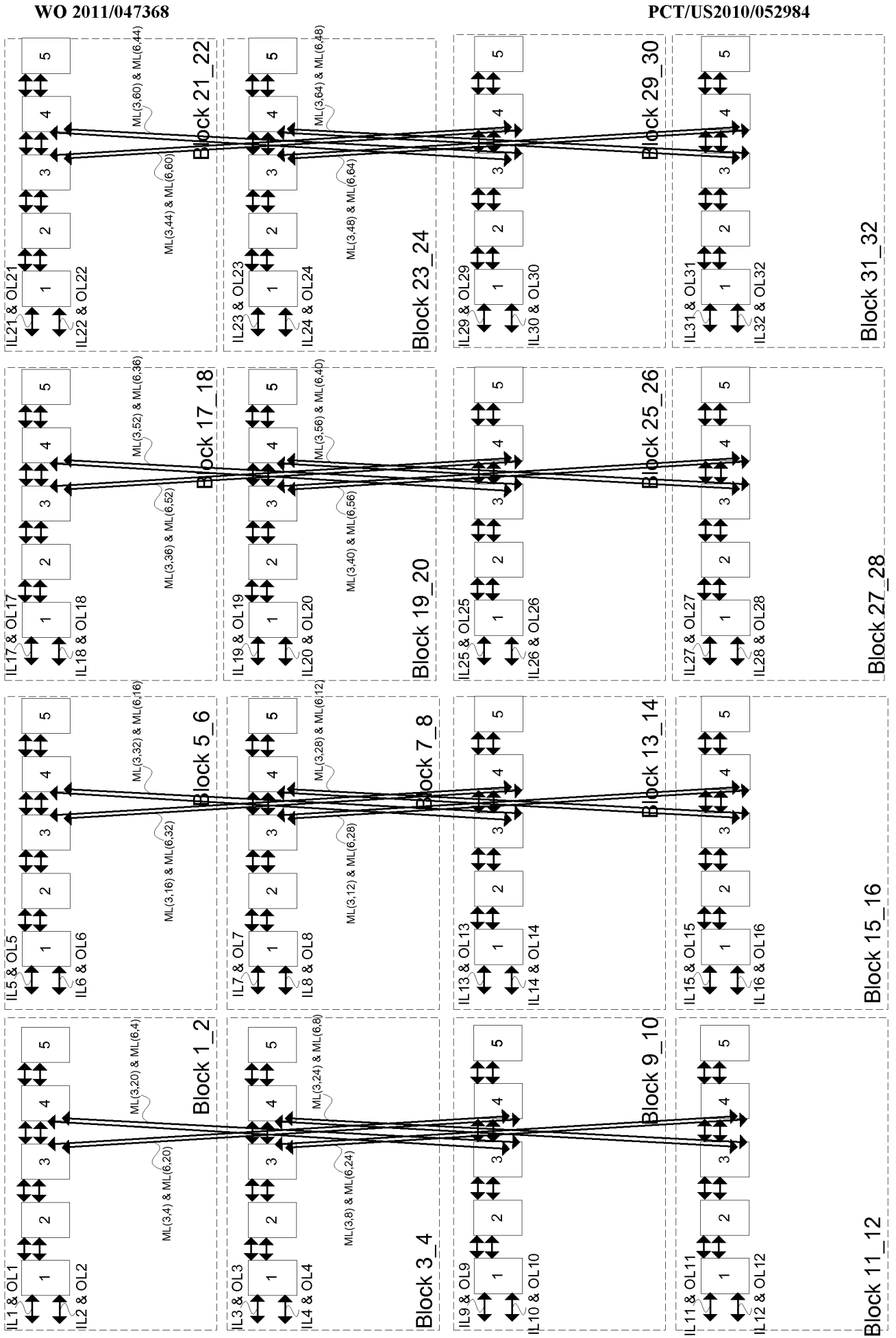
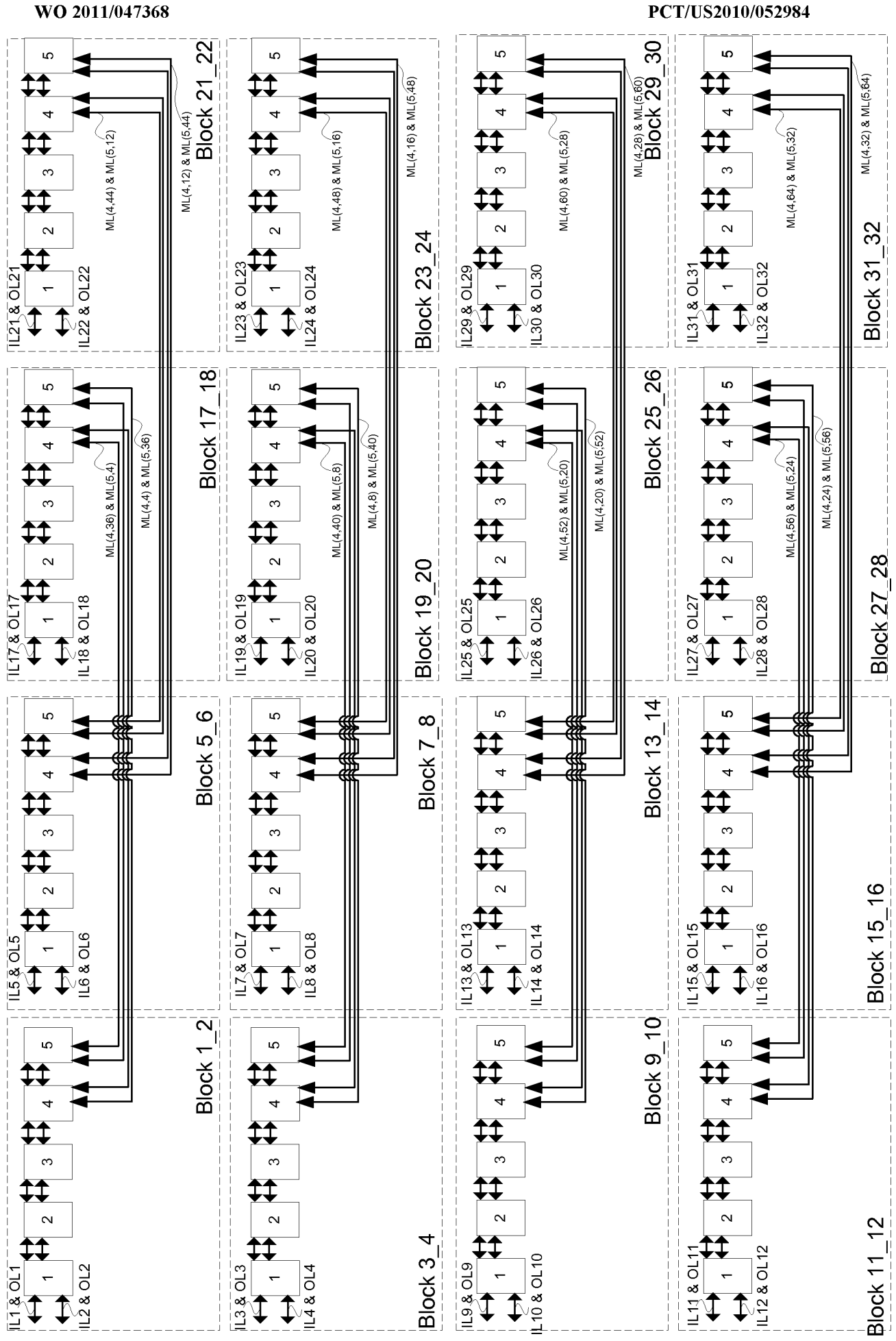


FIG. 8G

800G

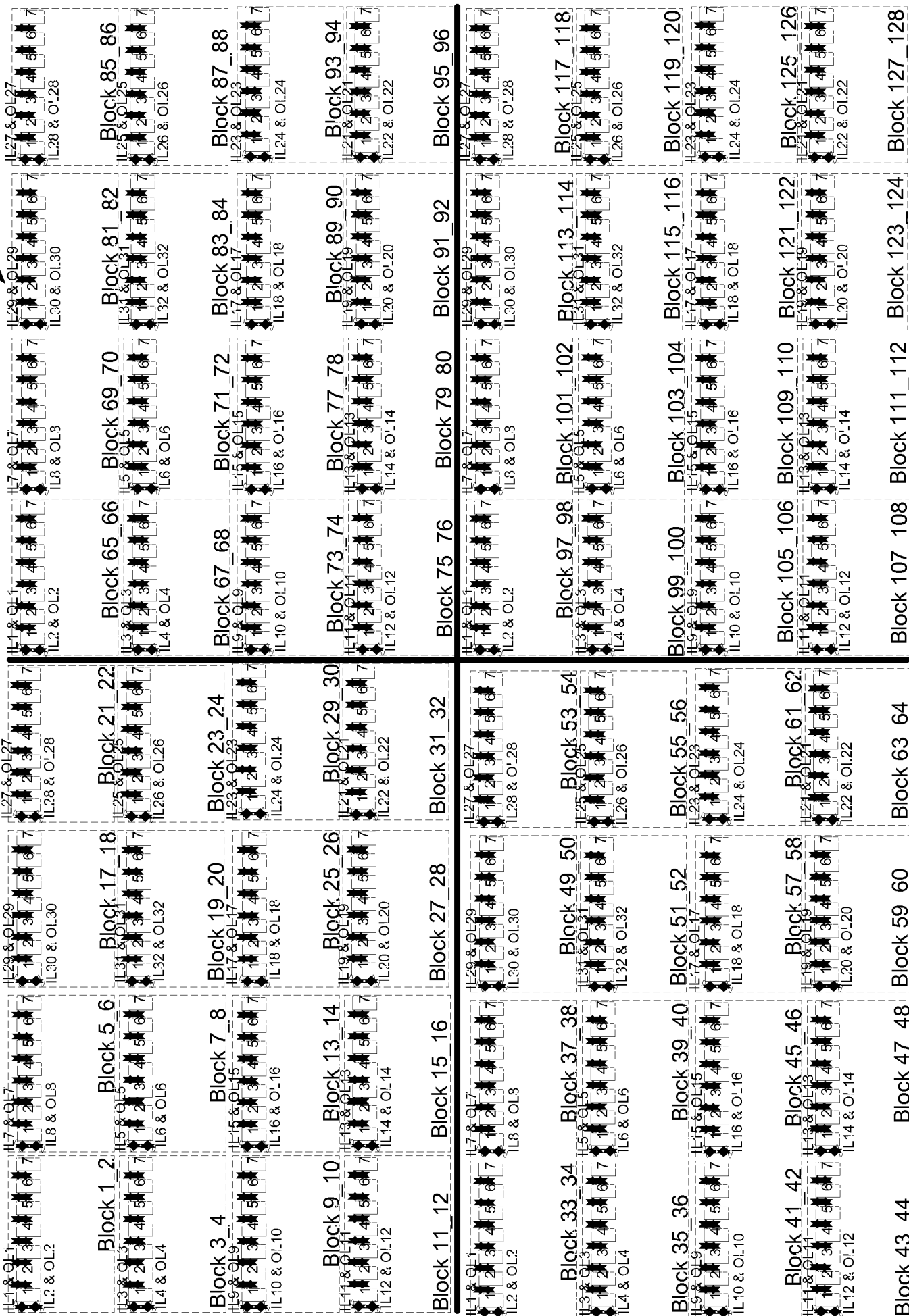


WO 2011/047368

PCT/US2010/052984

800H

FIG. 8H



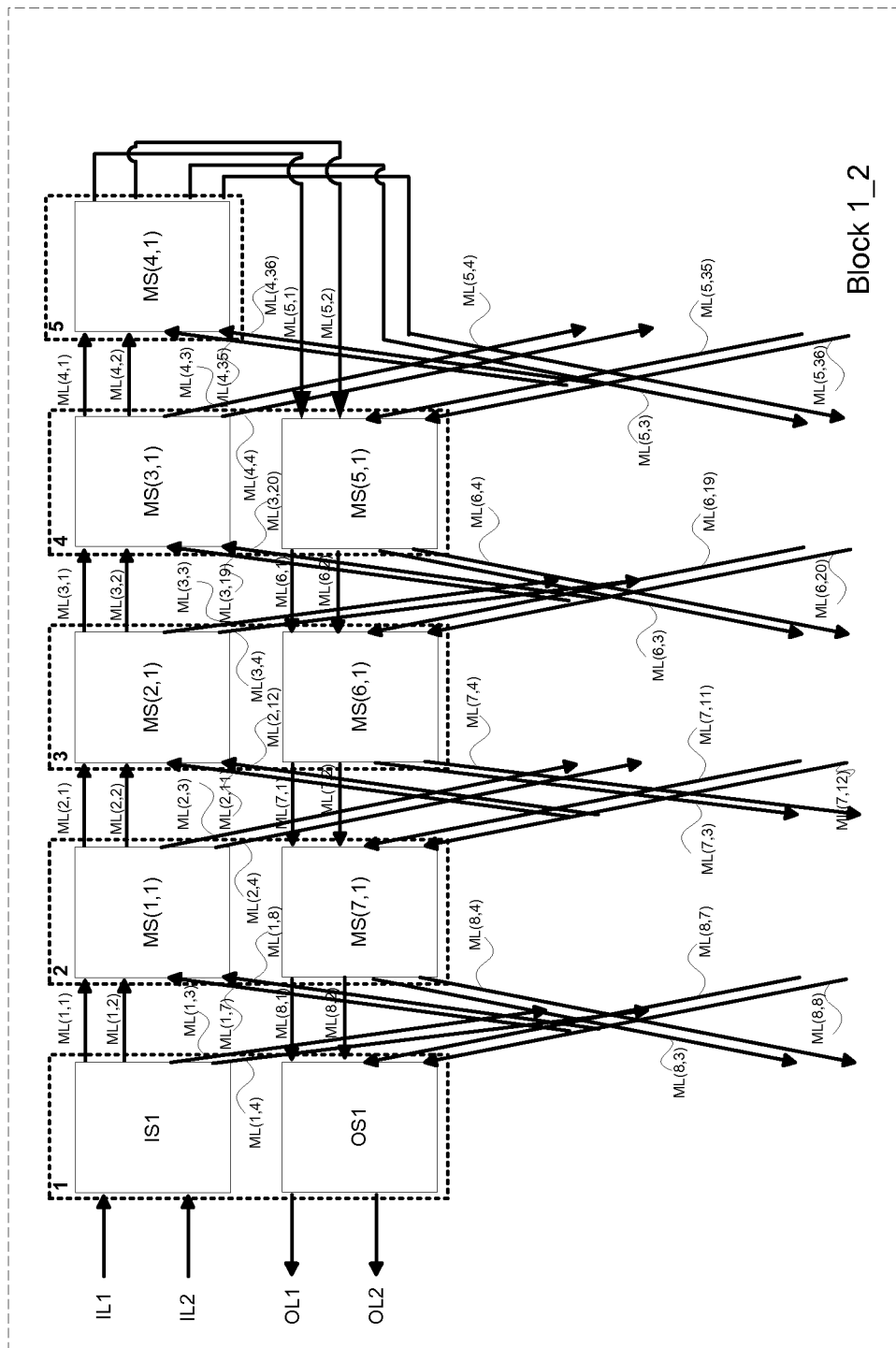
VLSI Layouts of Fully Connected Generalized and pyramid Networks with Locality Exploitation

Inventor: Venkat Konda

S-0048 PCT

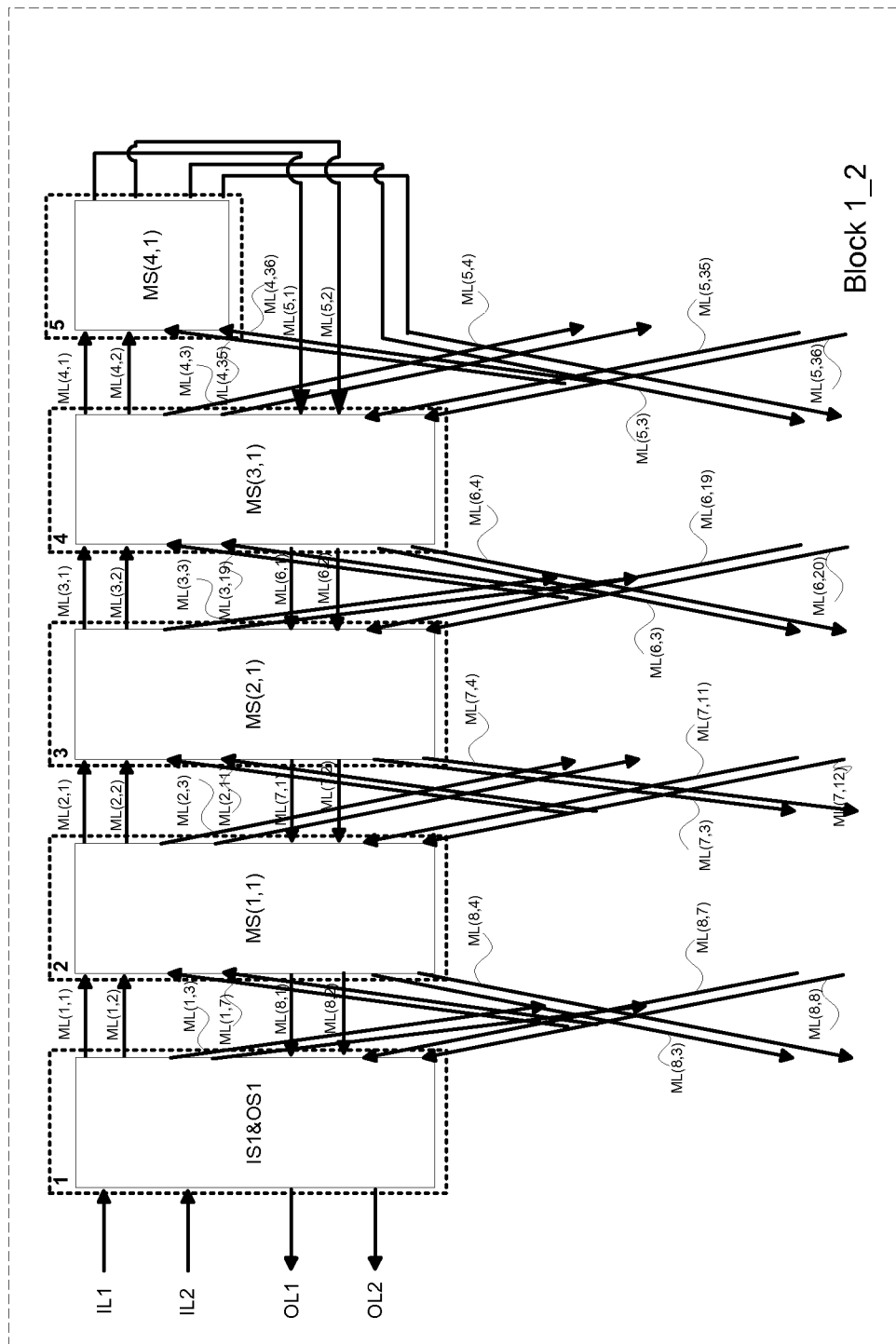
8001

FIG. 8I



800J

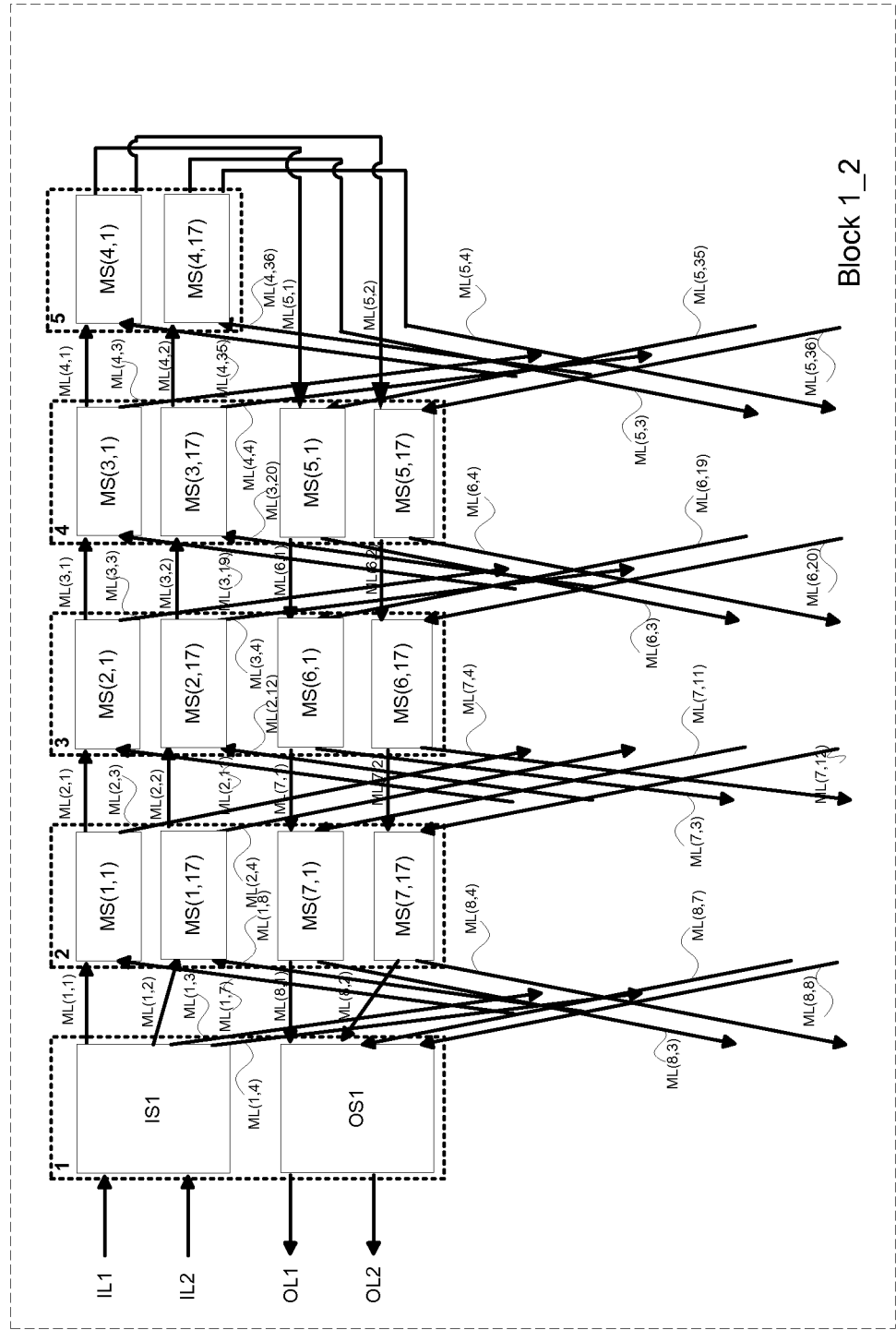
FIG. 8J



VLSI Layouts of Fully Connected Generalized and pyramid Networks with Locality Exploitation
Inventor: Venkat Konda
S-0048 PCT

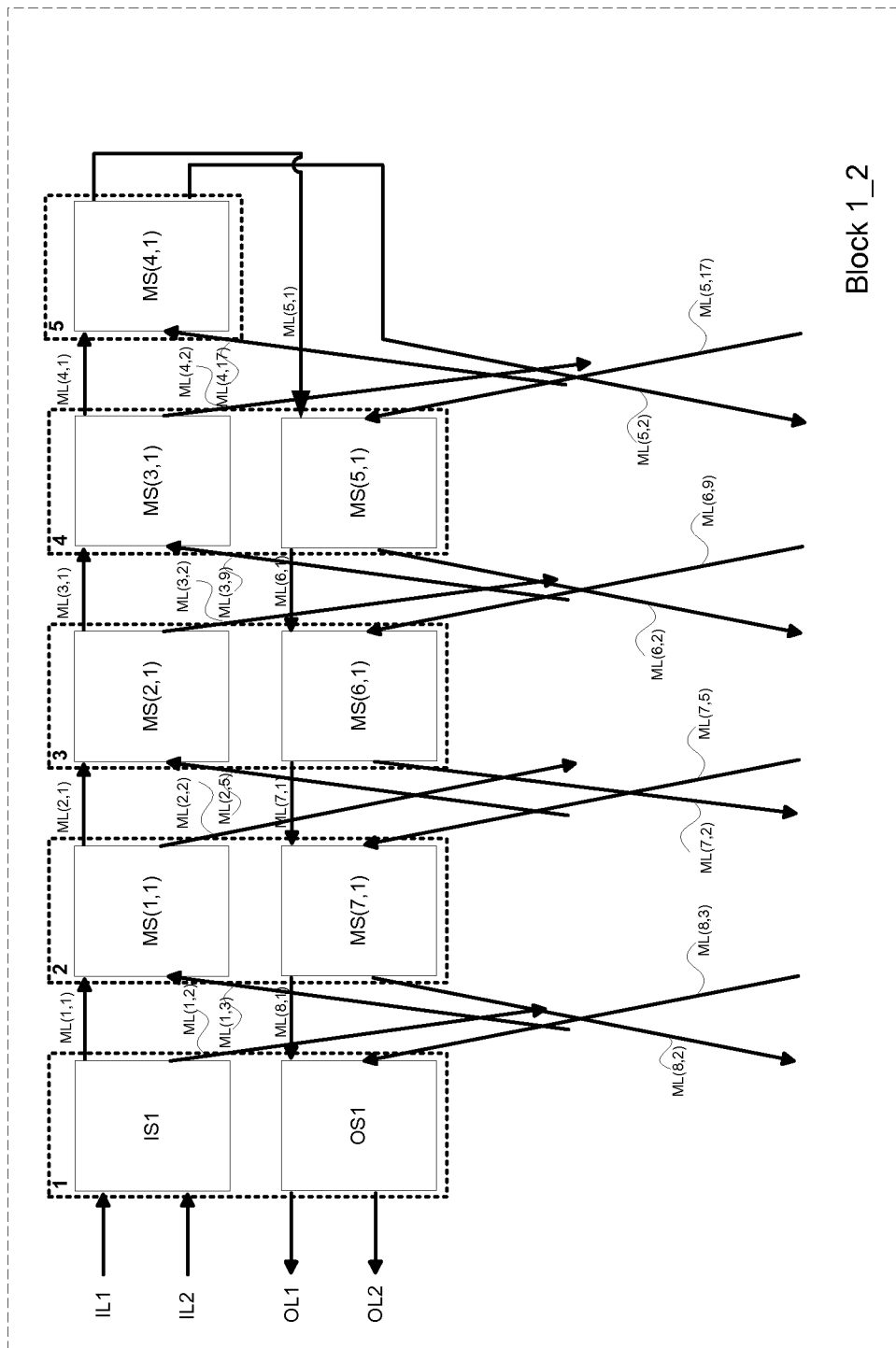
800K

FIG. 8K



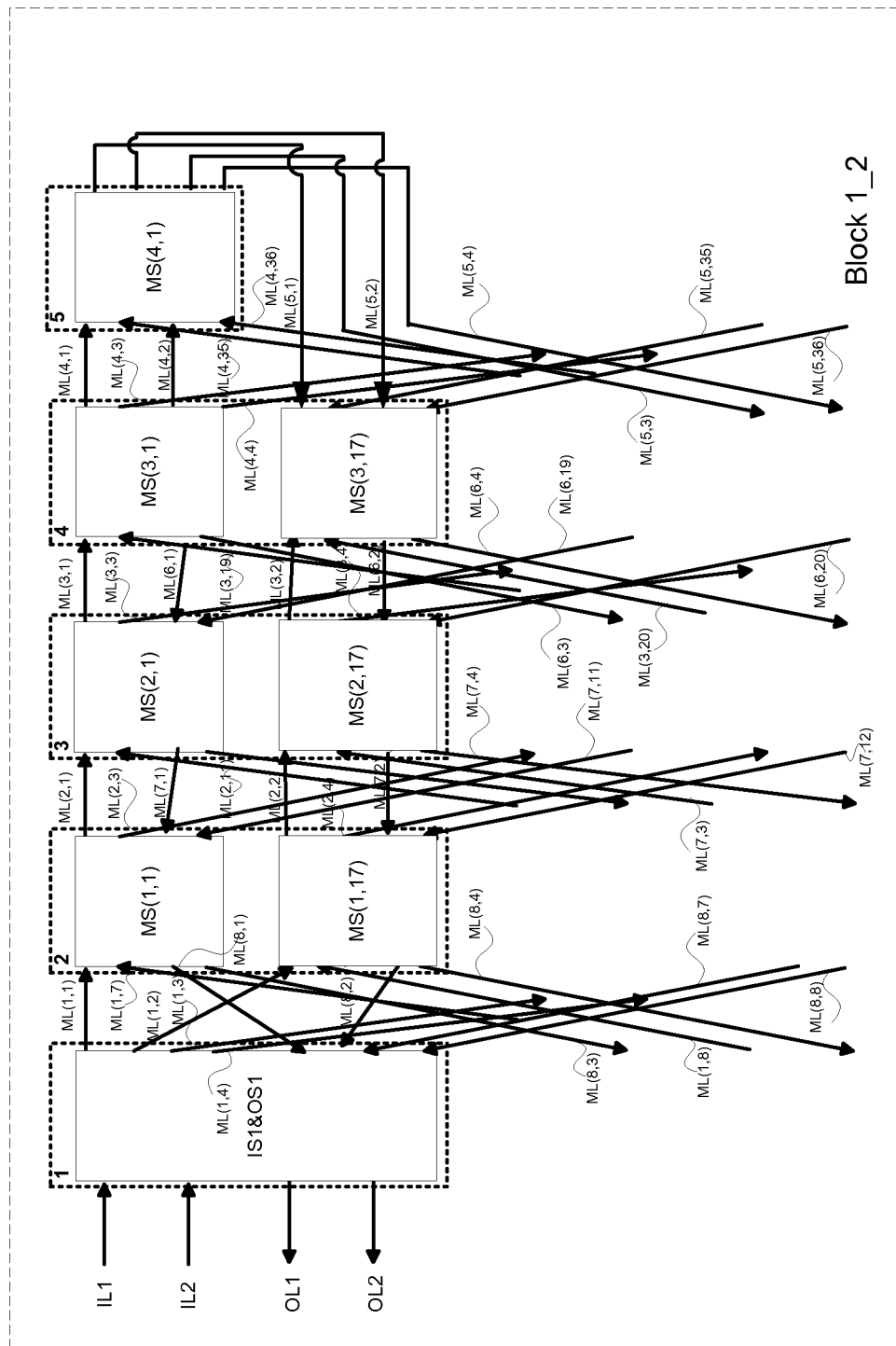
800K1

FIG. 8K1



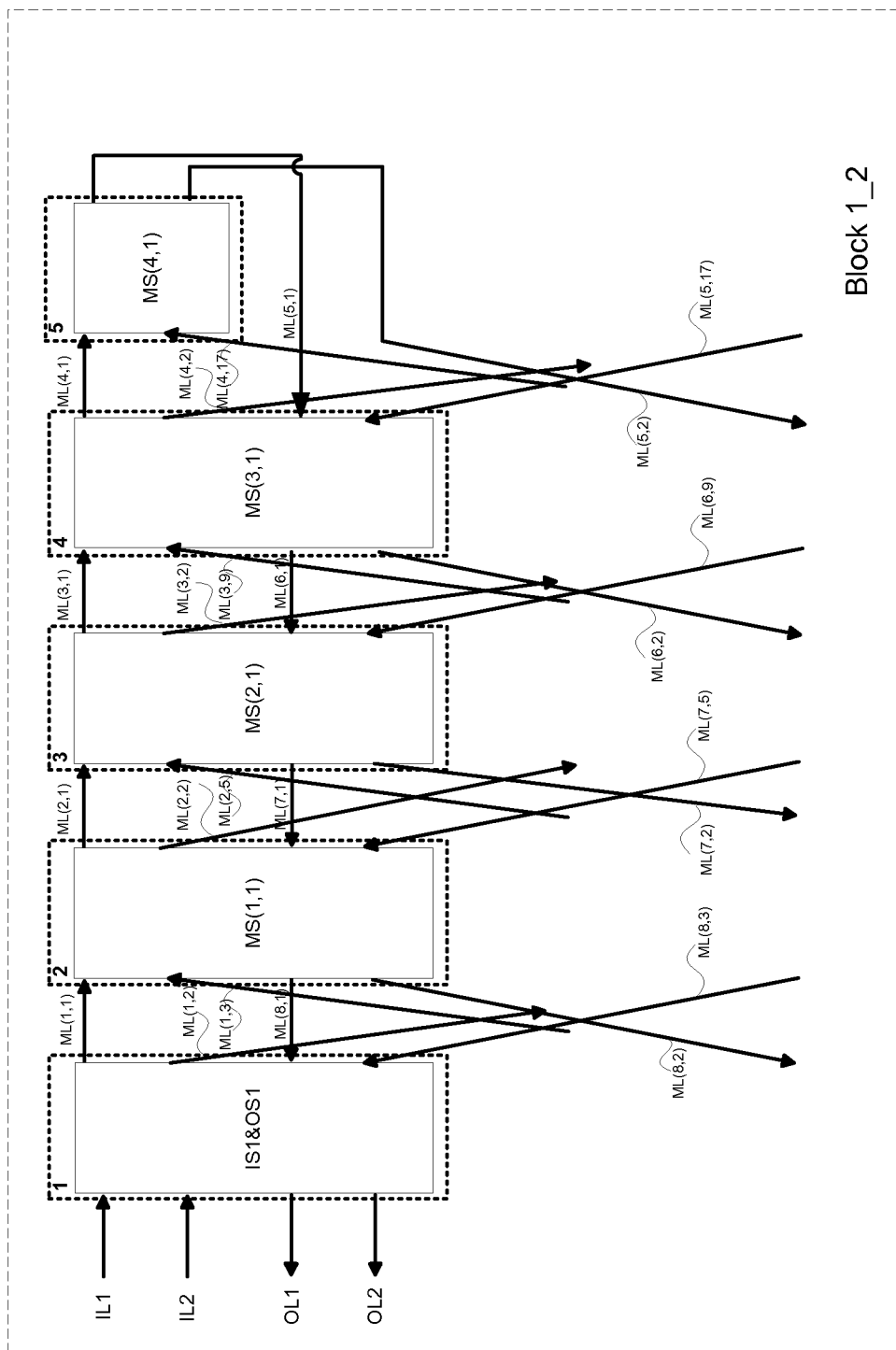
800L

FIG. 8L



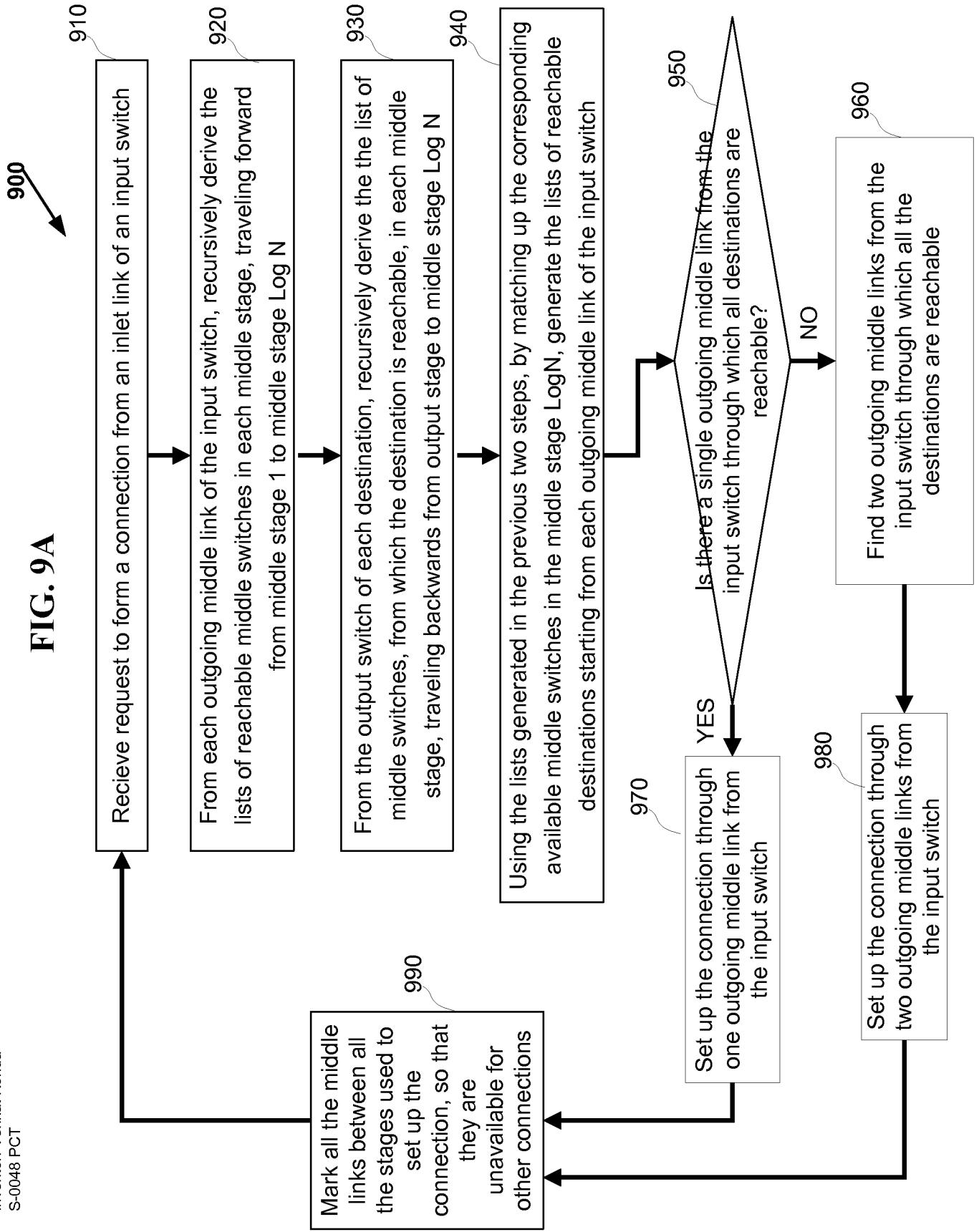
800L1

FIG. 8L1



Block 1_2

FIG. 9A



1000

FIG. 10A

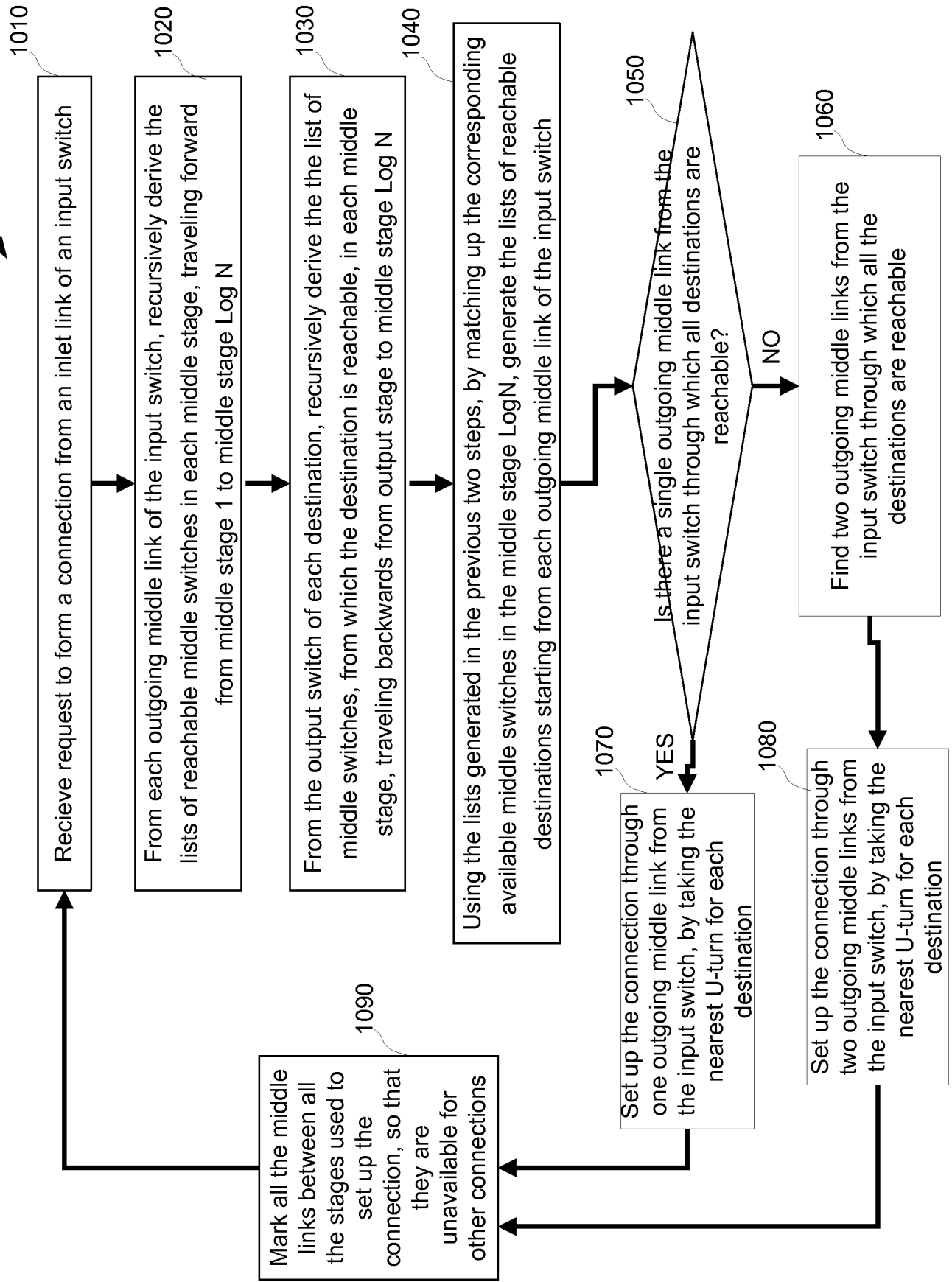
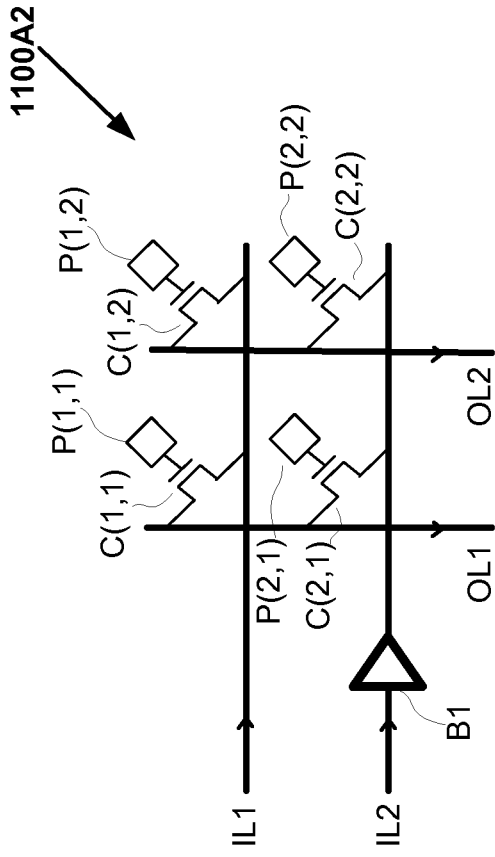
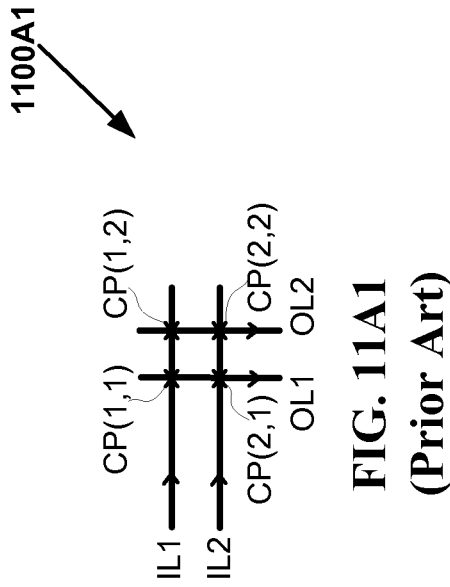


FIG. 11A



**FIG. 11A2
 (Prior Art)**



**FIG. 11A1
 (Prior Art)**

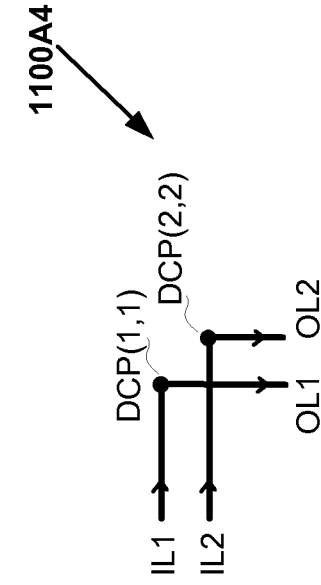
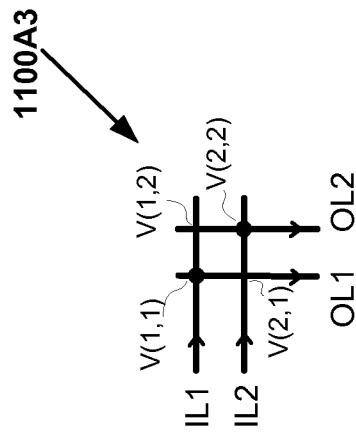


FIG. 11A4



**FIG. 11A3
 (Prior Art)**

EXHIBIT J



US 20120269190A1

(19) **United States**
(12) **Patent Application Publication**
Konda

(10) **Pub. No.: US 2012/0269190 A1**
(43) **Pub. Date: Oct. 25, 2012**

(54) **VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS WITH LOCALITY EXPLOITATION**

(75) Inventor: **Venkat Konda, San Jose, CA (US)**

(73) Assignee: **Konda Technologies Inc., San Jose, CA (US)**

(21) Appl. No.: **13/502,207**

(22) PCT Filed: **Oct. 16, 2010**

(86) PCT No.: **PCT/US10/52984**

§ 371 (c)(1),
(2), (4) Date: **Apr. 16, 2012**

Related U.S. Application Data

(60) Provisional application No. 61/252,603, filed on Oct. 16, 2009, provisional application No. 61/252,609, filed on Oct. 16, 2009.

Publication Classification

(51) **Int. Cl. H04L 12/50 (2006.01)**

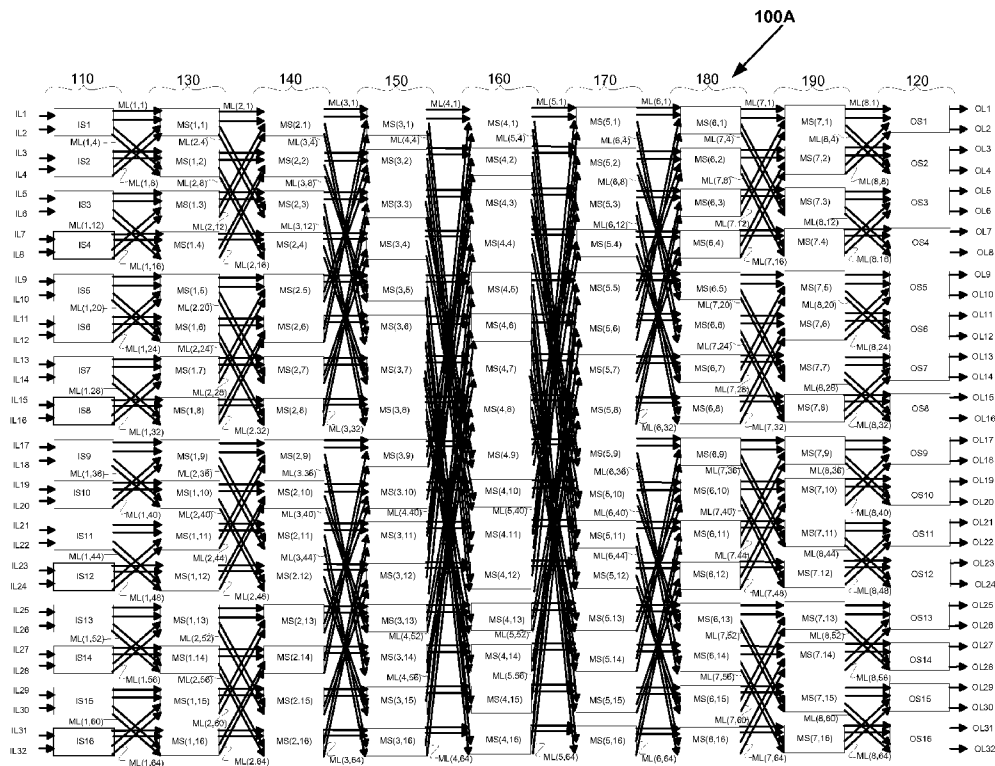
(52) **U.S. Cl. 370/388**

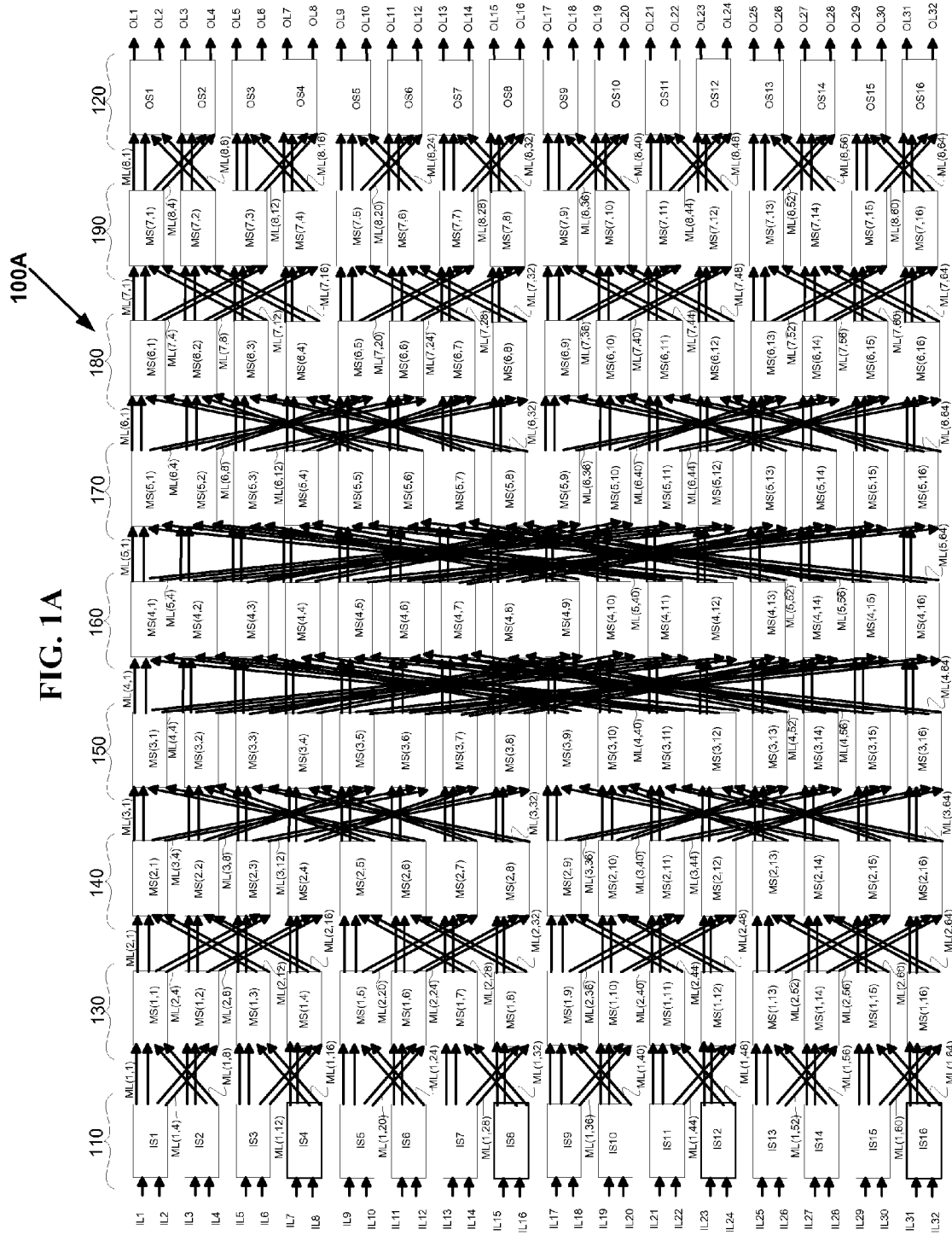
(57) **ABSTRACT**

VLSI layouts of generalized multi-stage and pyramid networks for broadcast, unicast and multicast connections are

presented using only horizontal and vertical links with spacial locality exploitation. The VLSI layouts employ shuffle exchange links where outlet links of cross links from switches in a stage in one sub-integrated circuit block are connected to inlet links of switches in the succeeding stage in another sub-integrated circuit block so that said cross links are either vertical links or horizontal and vice versa. Furthermore the shuffle exchange links are employed between different sub-integrated circuit blocks so that spatially nearer sub-integrated circuit blocks are connected with shorter links compared to the shuffle exchange links between spatially farther sub-integrated circuit blocks. In one embodiment the sub-integrated circuit blocks are arranged in a hypercube arrangement in a two-dimensional plane. The VLSI layouts exploit the benefits of significantly lower cross points, lower signal latency, lower power and full connectivity with significantly fast compilation.

The VLSI layouts with spacial locality exploitation presented are applicable to generalized multi-stage and pyramid networks, generalized folded multi-stage and pyramid networks, generalized butterfly fat tree and pyramid networks, generalized multi-link multi-stage and pyramid networks, generalized folded multi-link multi-stage and pyramid networks, generalized multi-link butterfly fat tree and pyramid networks, generalized hypercube networks, and generalized cube connected cycles networks for speedup of $s \geq 1$. The embodiments of VLSI layouts are useful in wide target applications such as FPGAs, CPLDs, pSoCs, ASIC placement and route tools, networking applications, parallel & distributed computing, and reconfigurable computing.





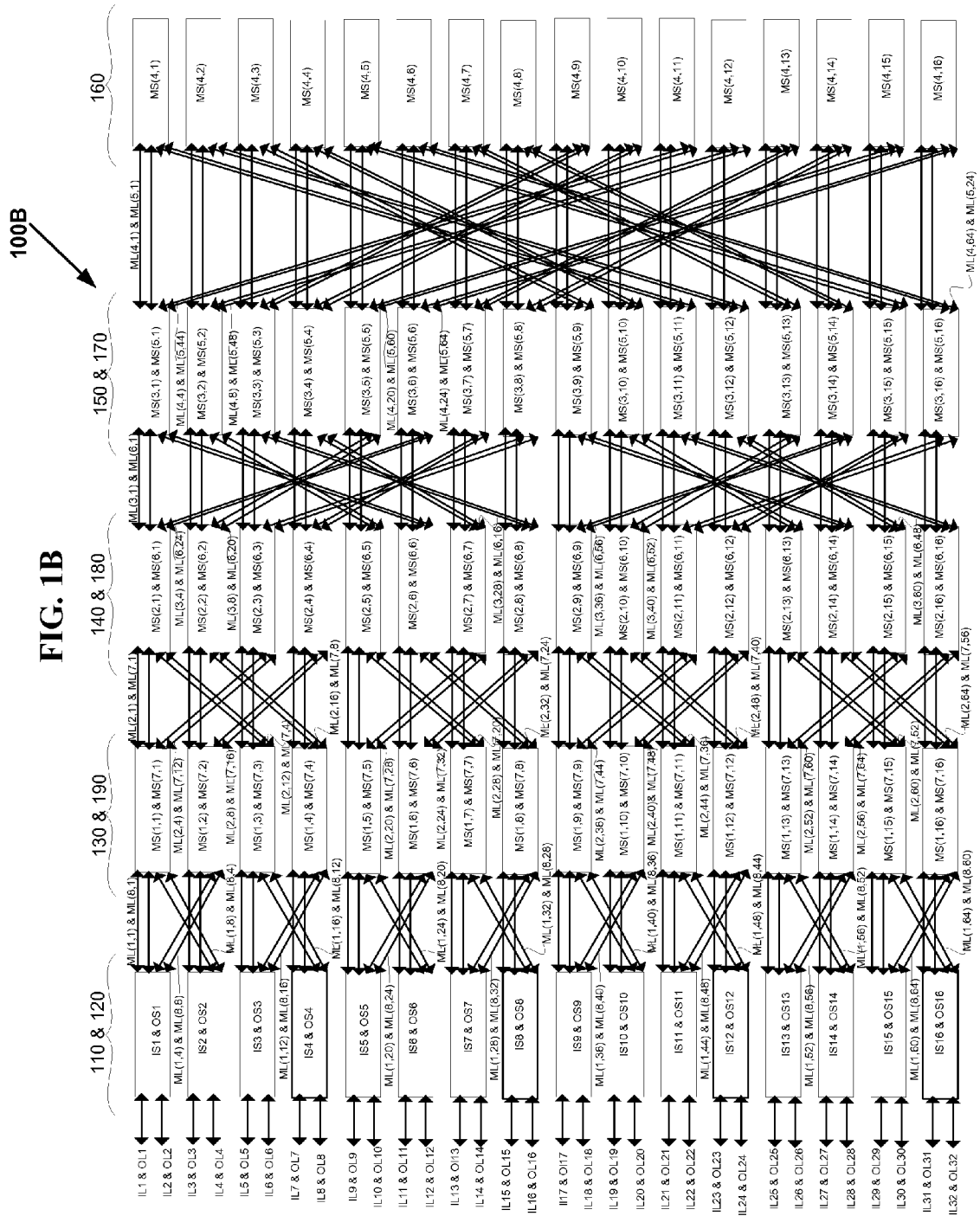
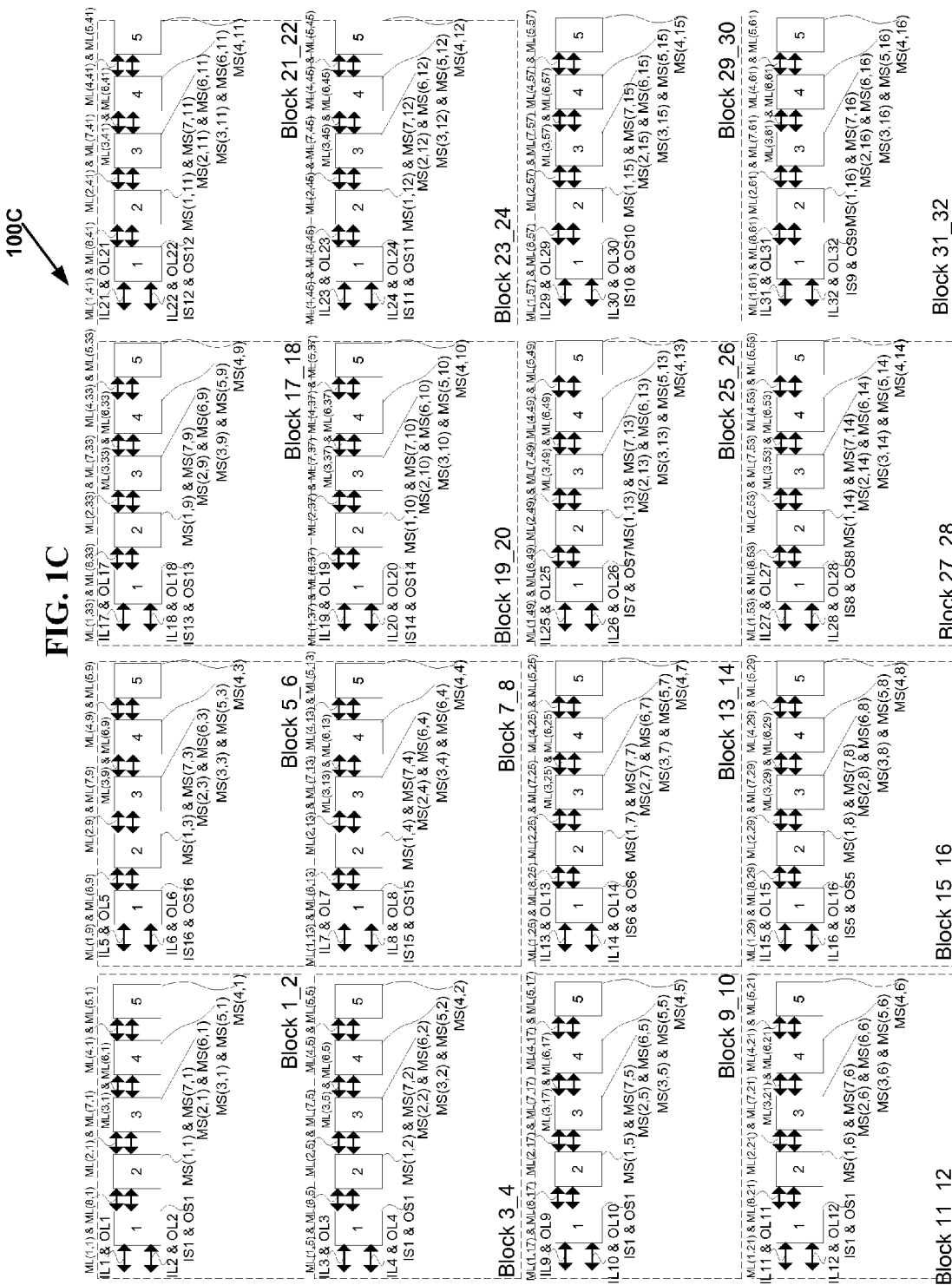
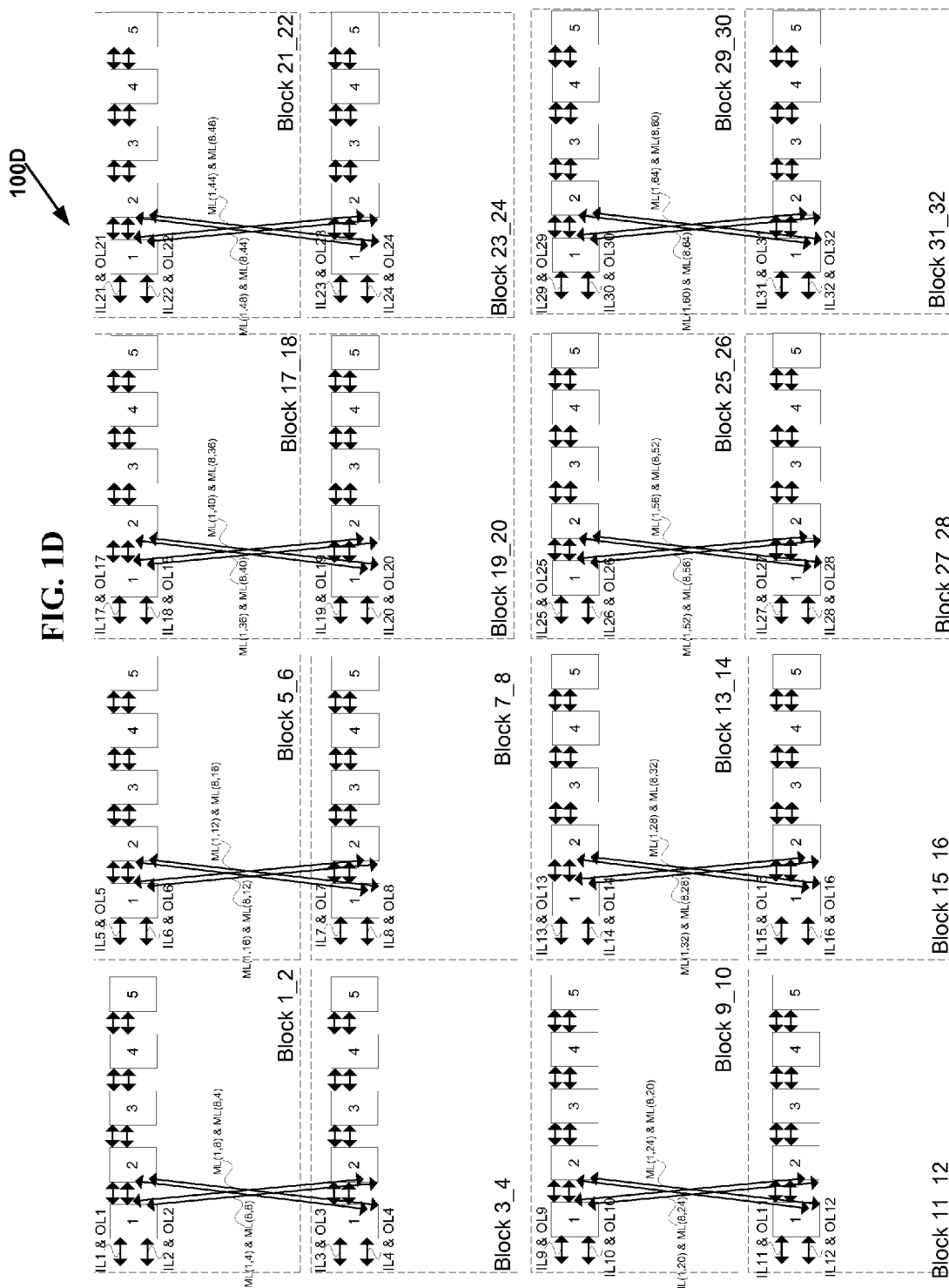


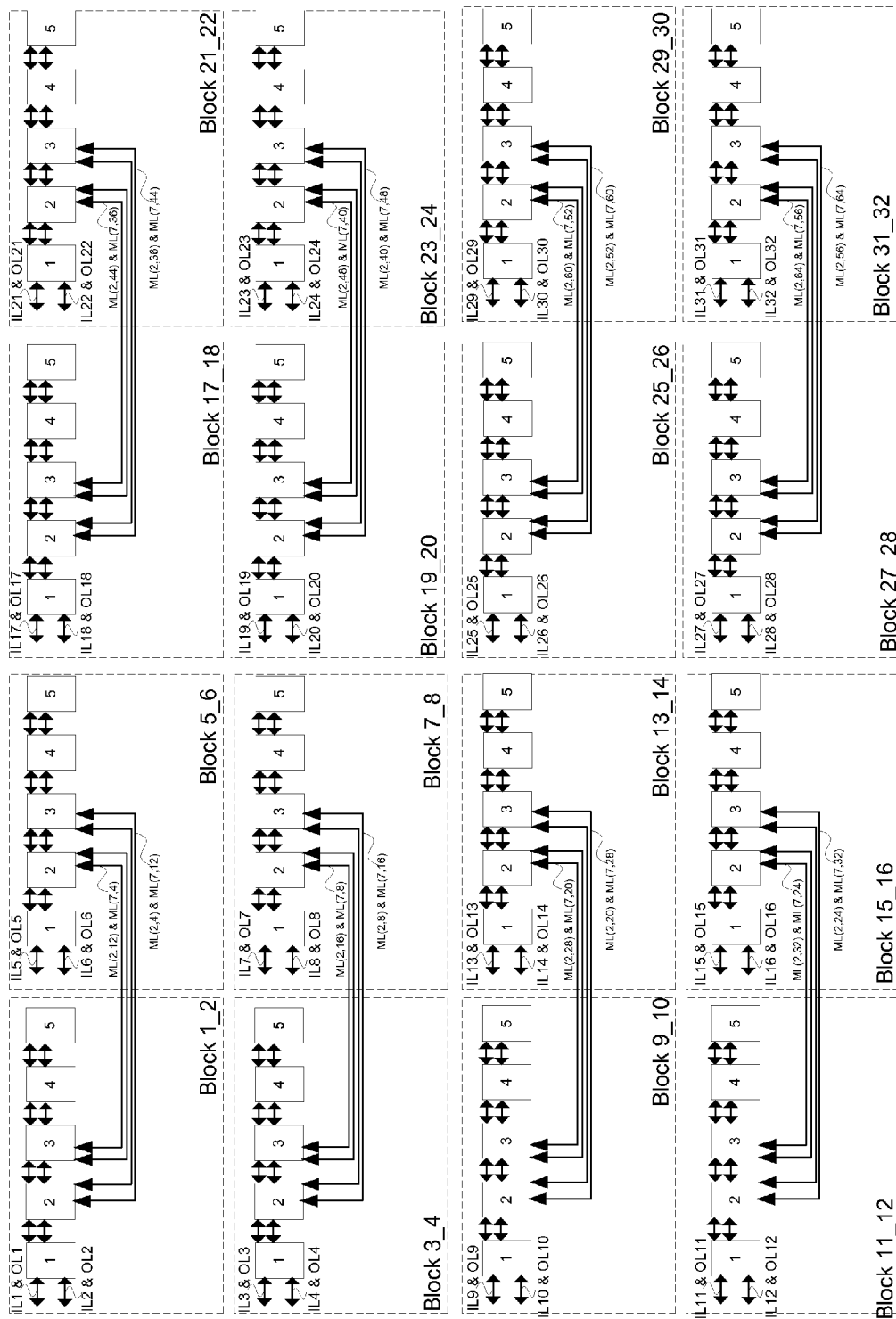
FIG. 1B

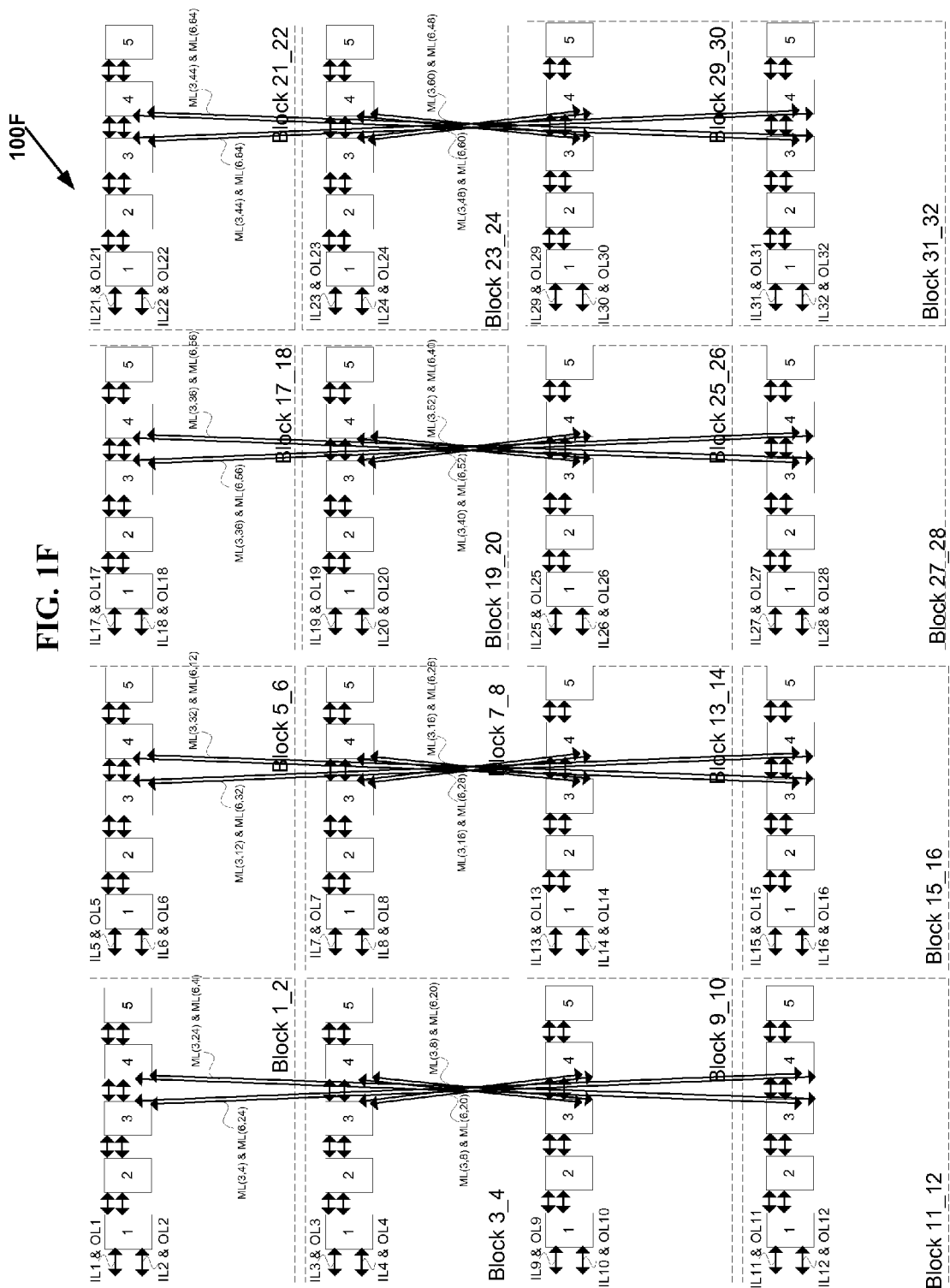


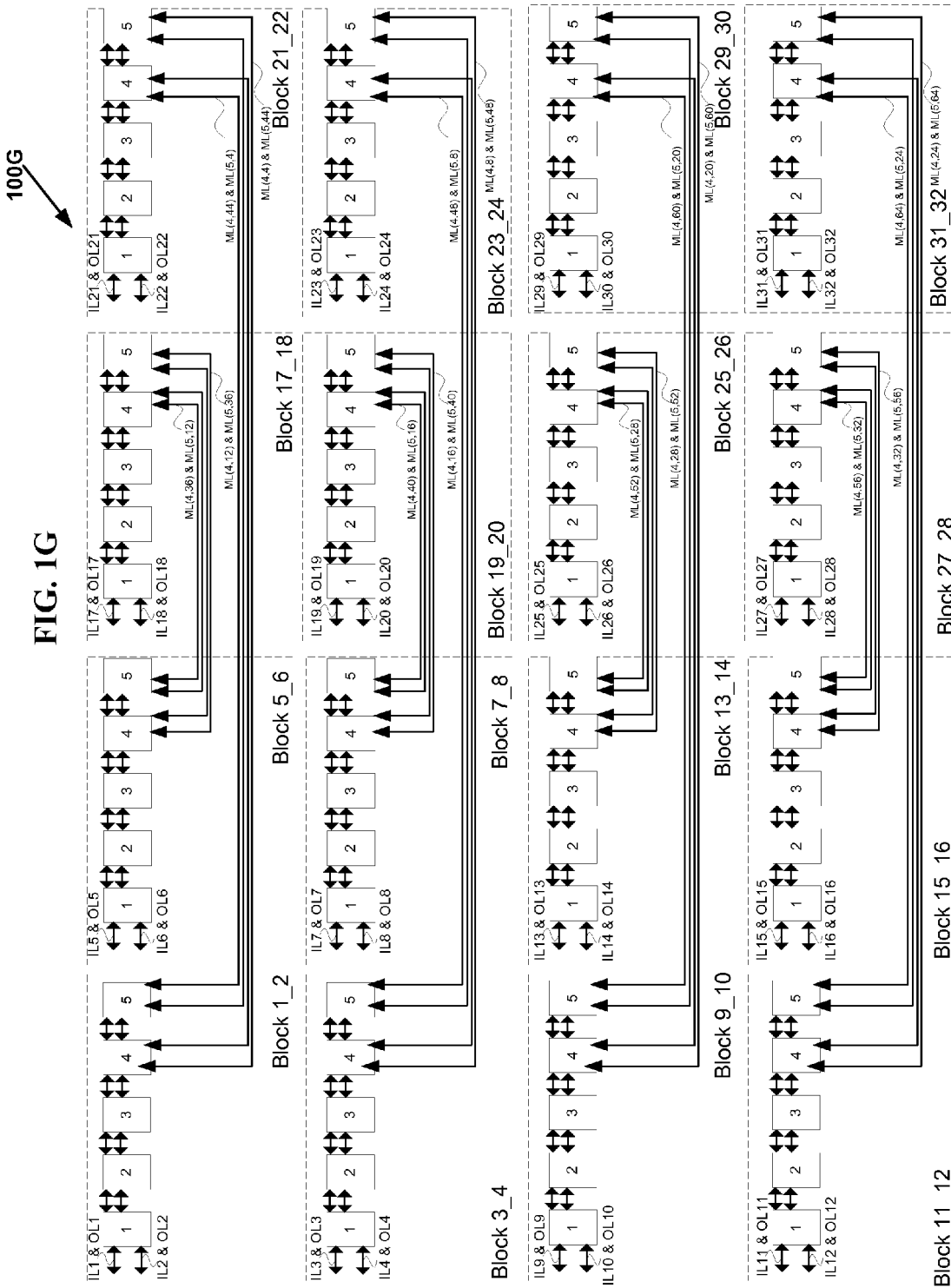


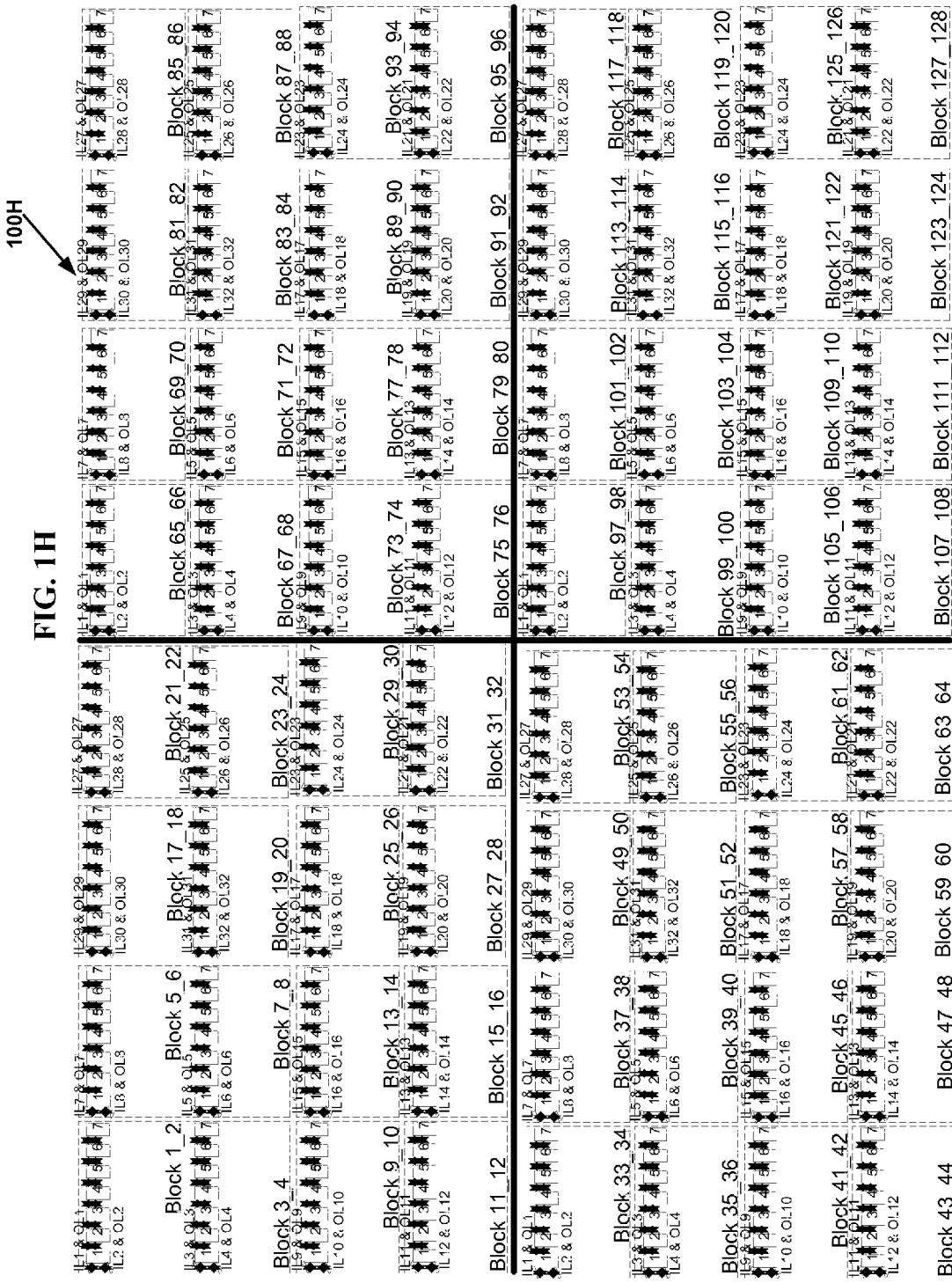
100E

FIG. 1E



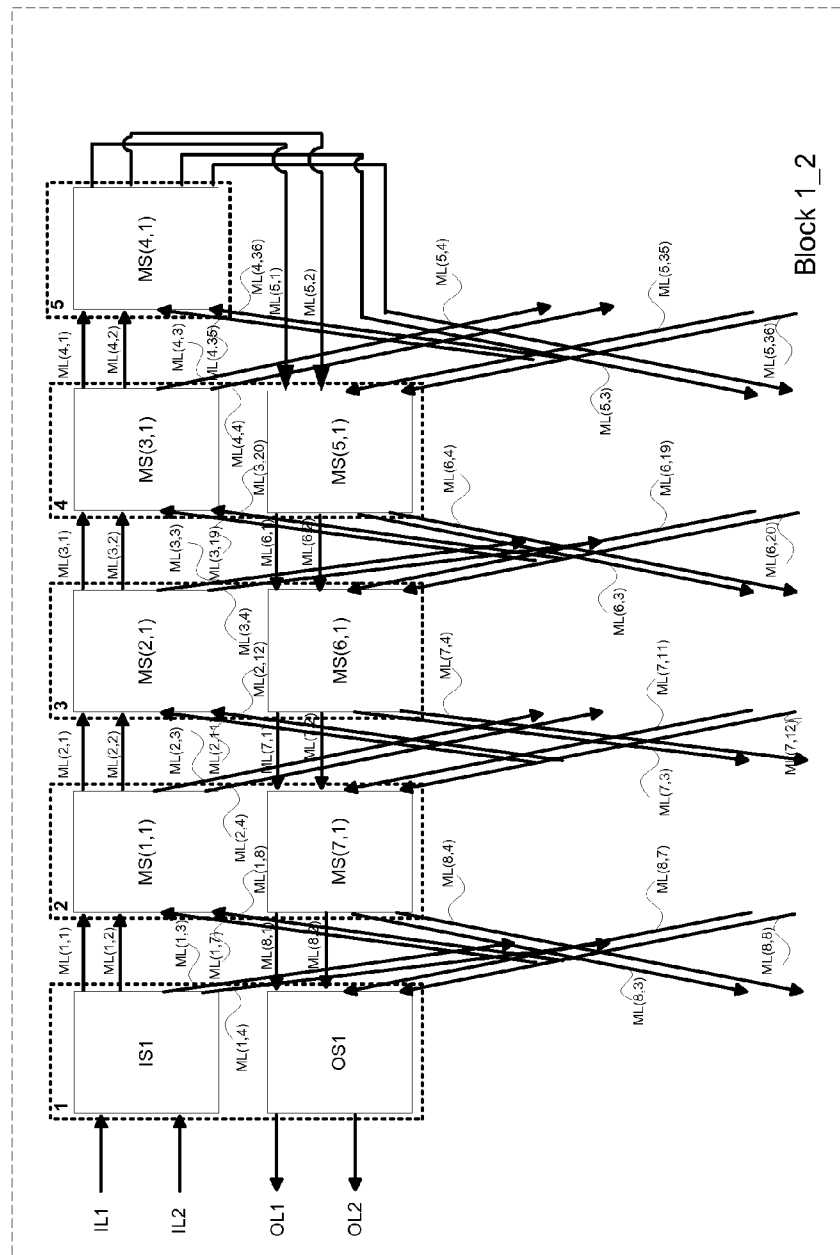






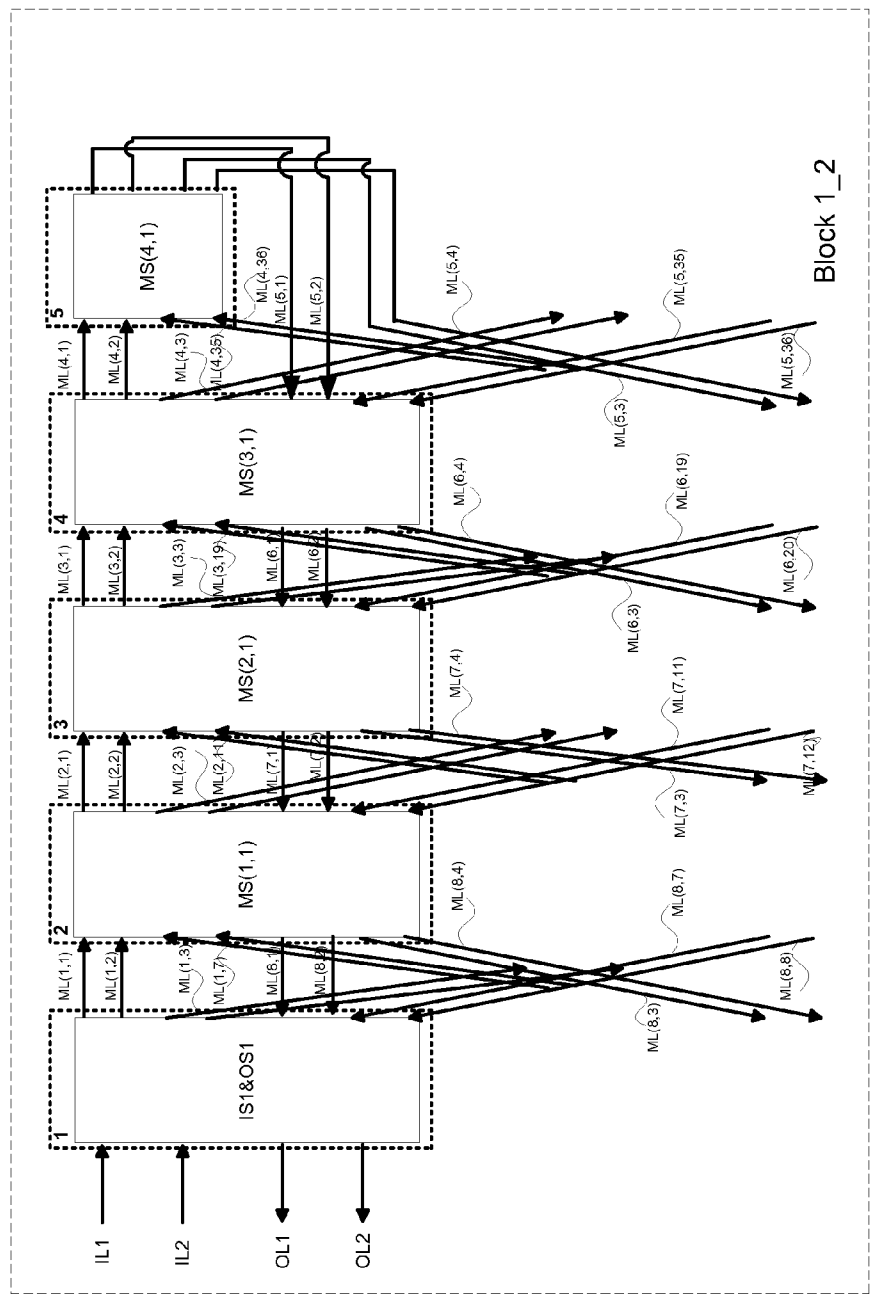
1001

FIG. 11



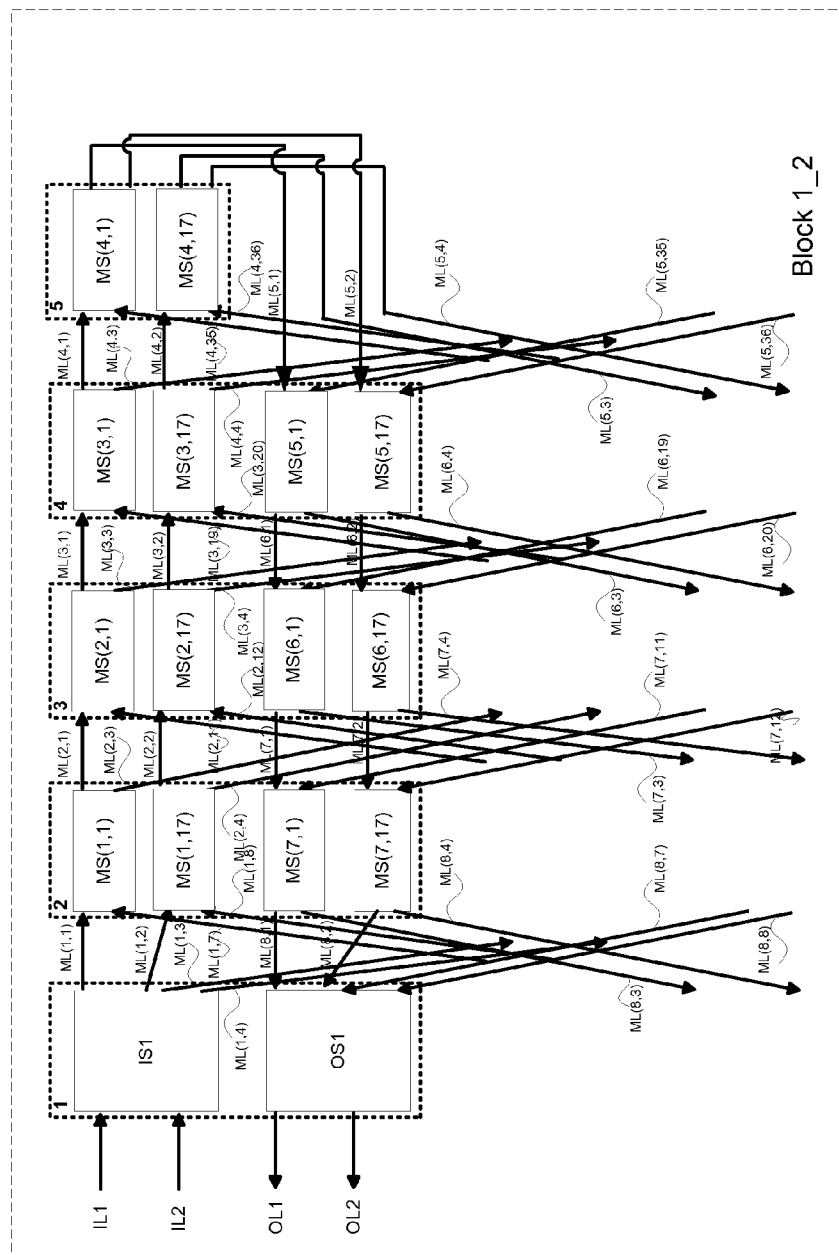
100J

FIG. 1J



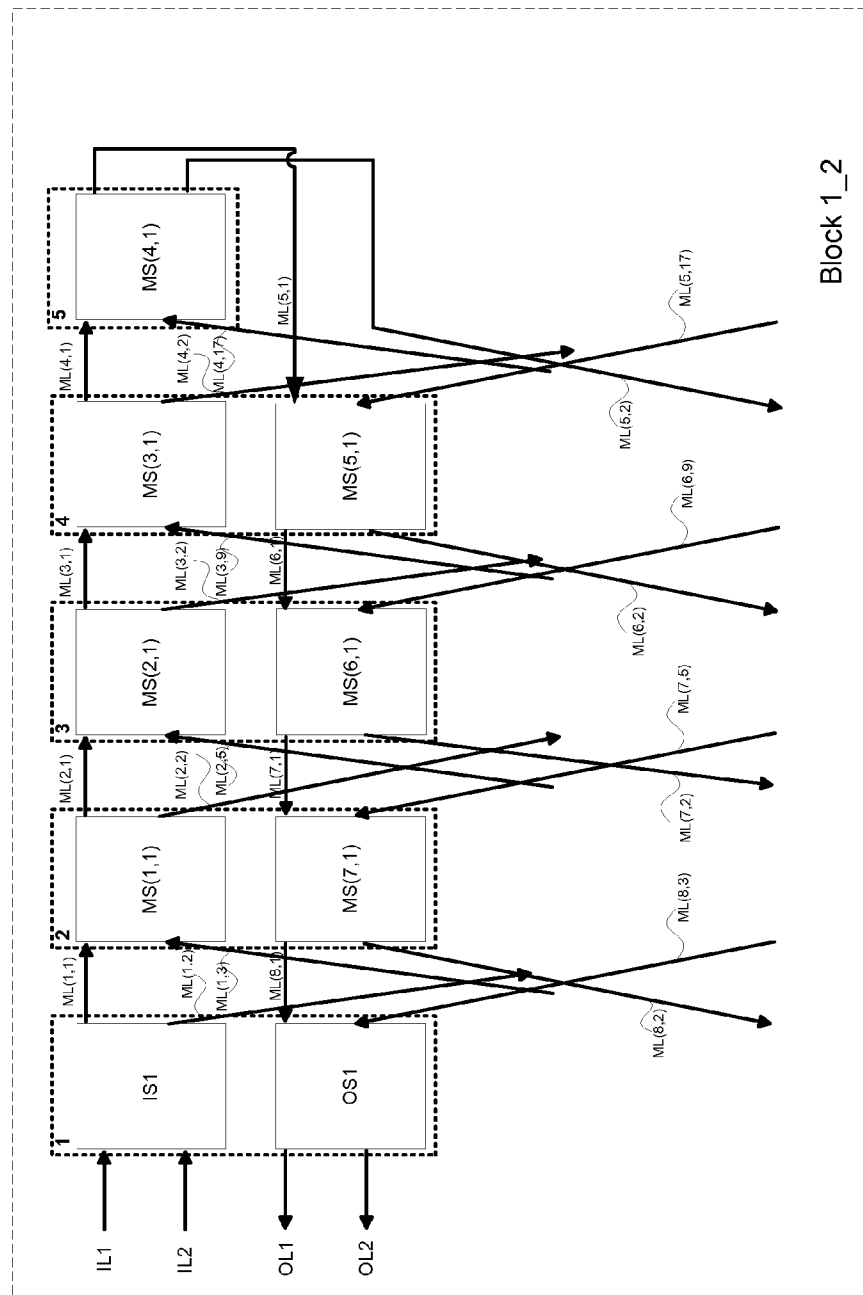
100K

FIG. 1K



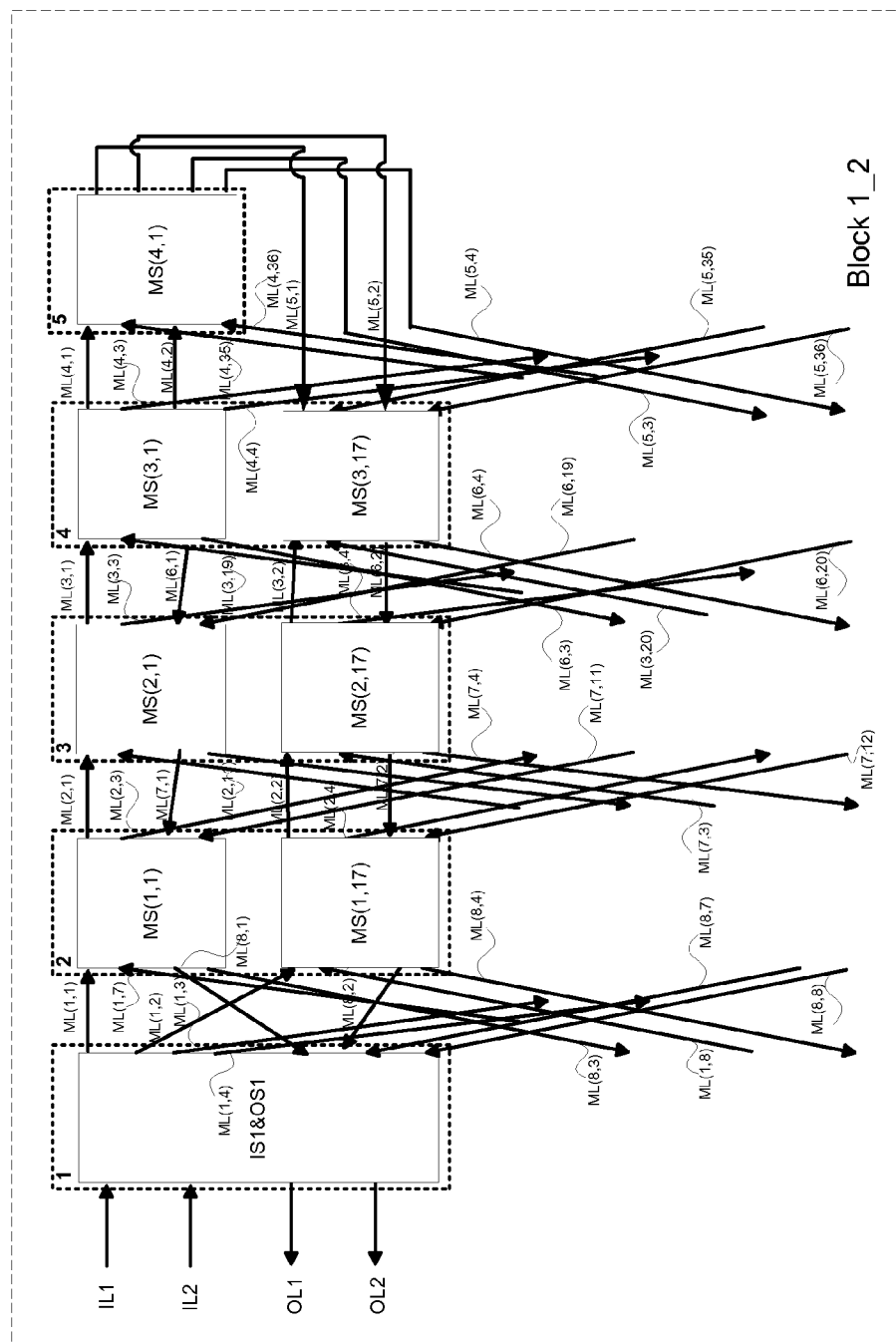
100K1

FIG. 1K1



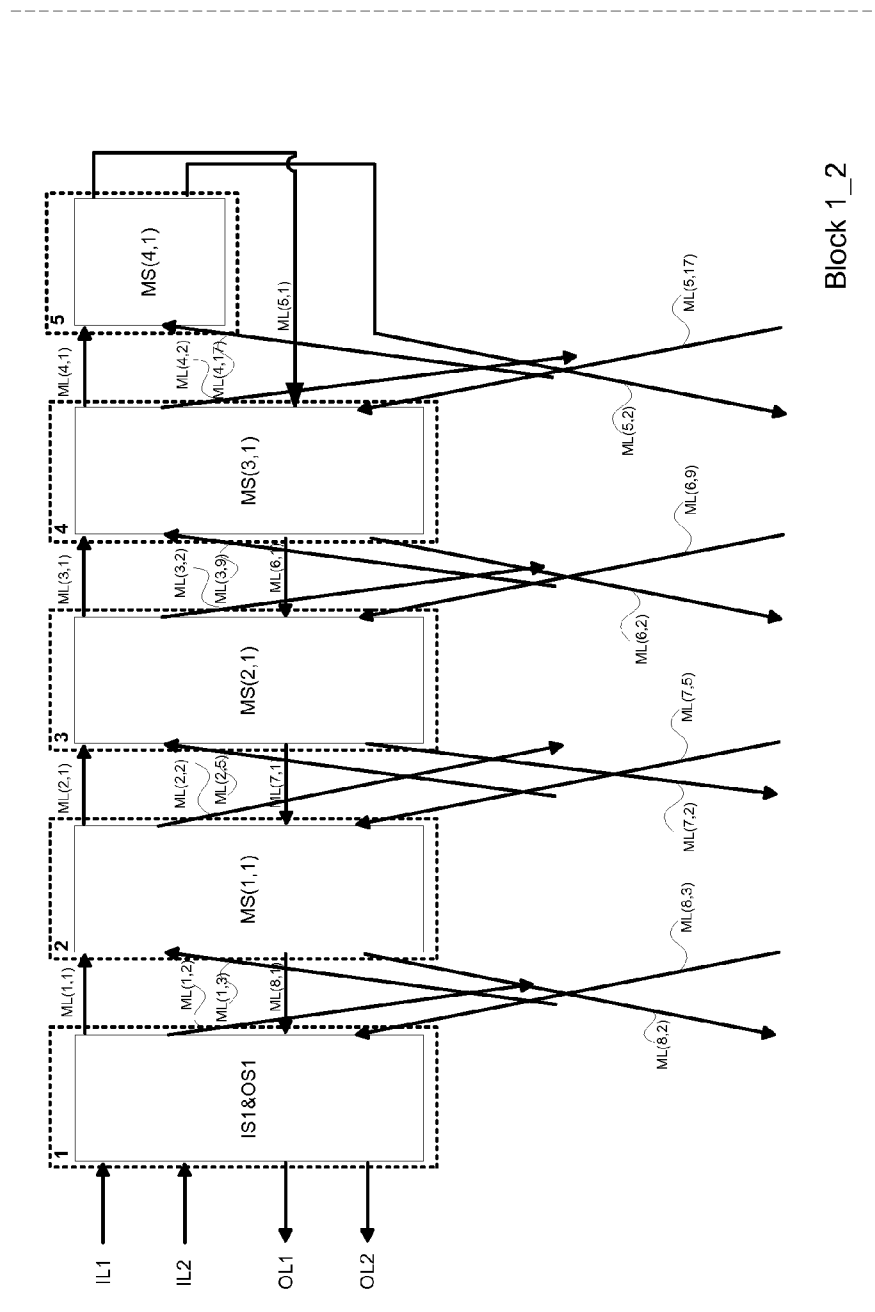
100L

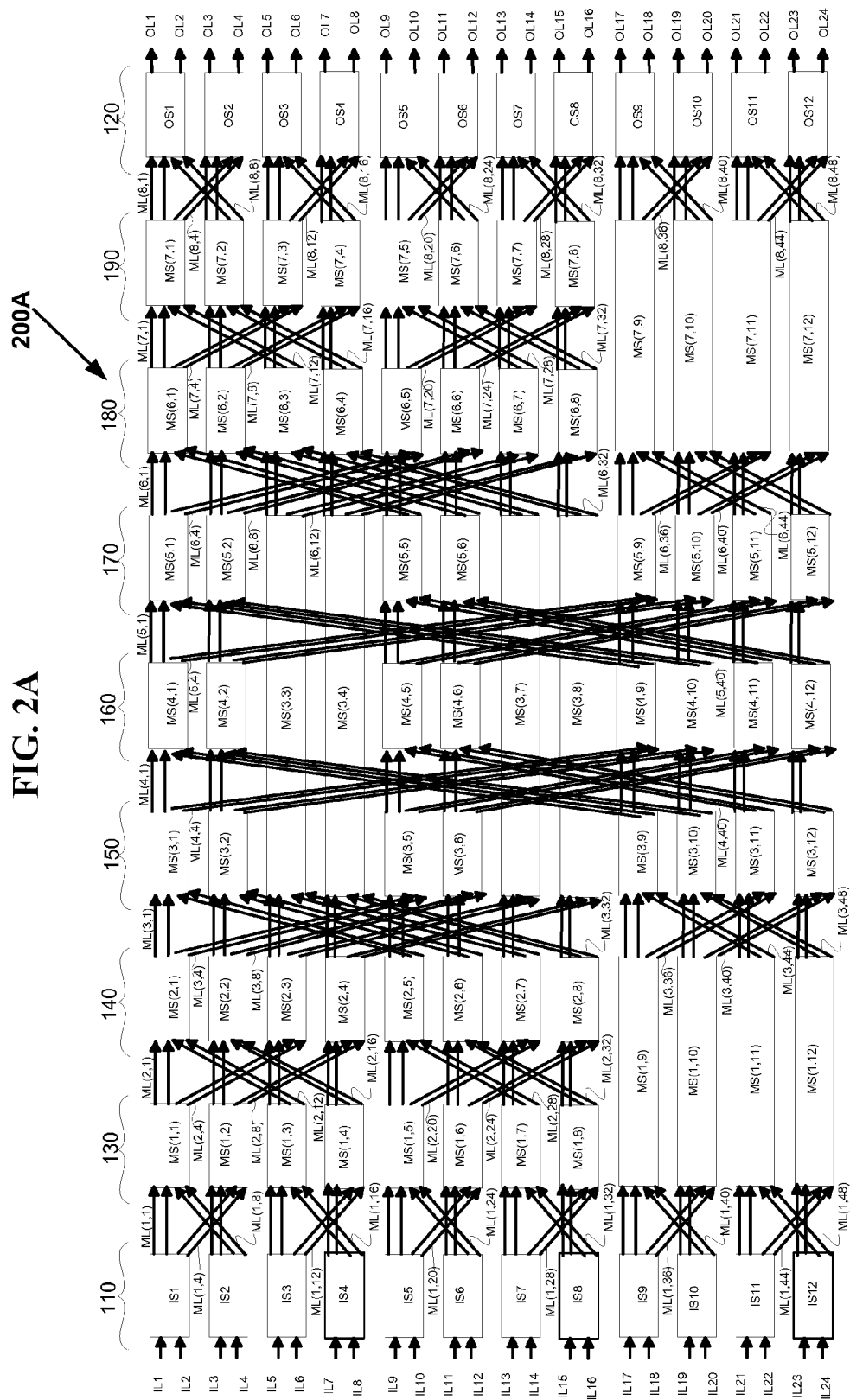
FIG. 1L



100L1

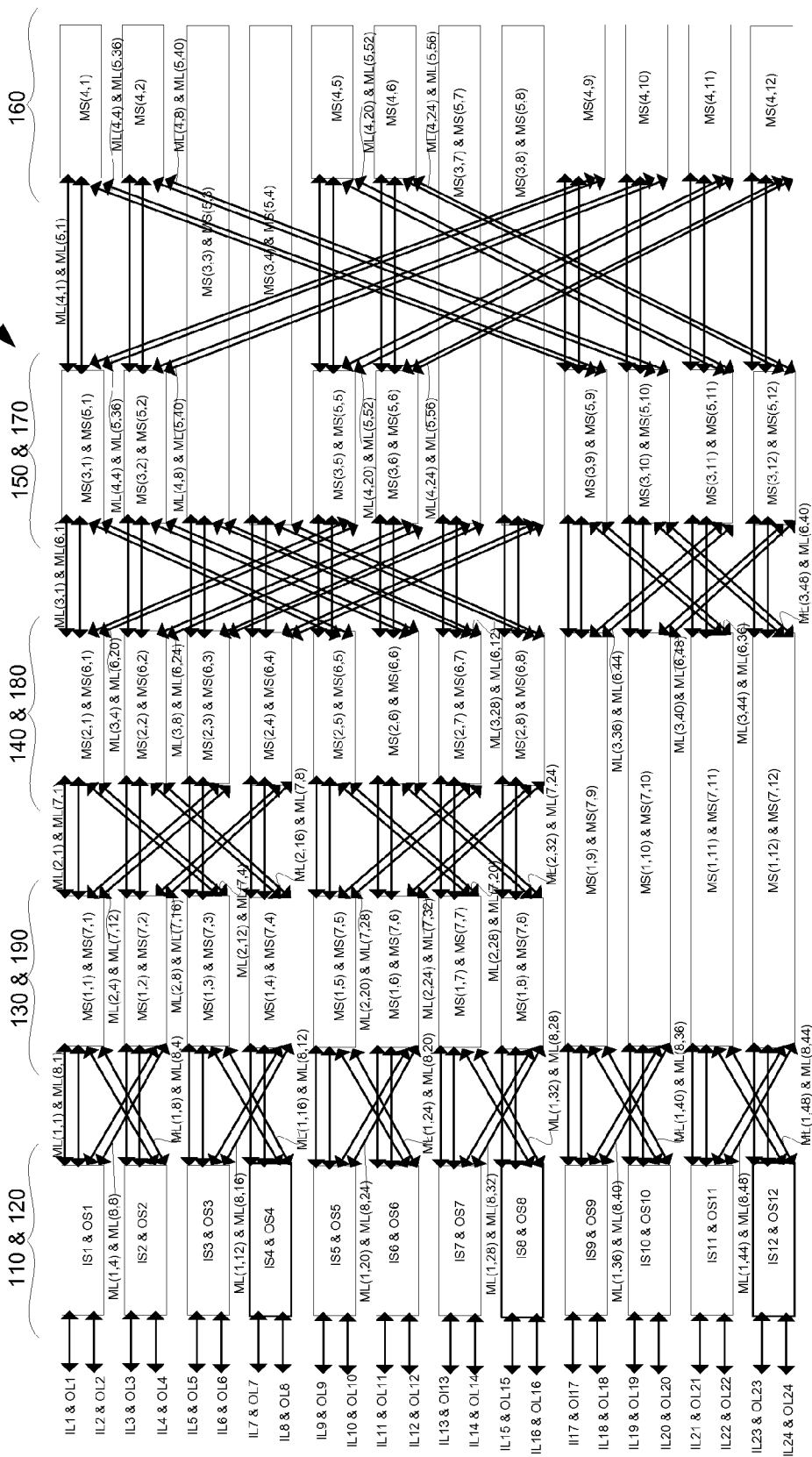
FIG. 1L1





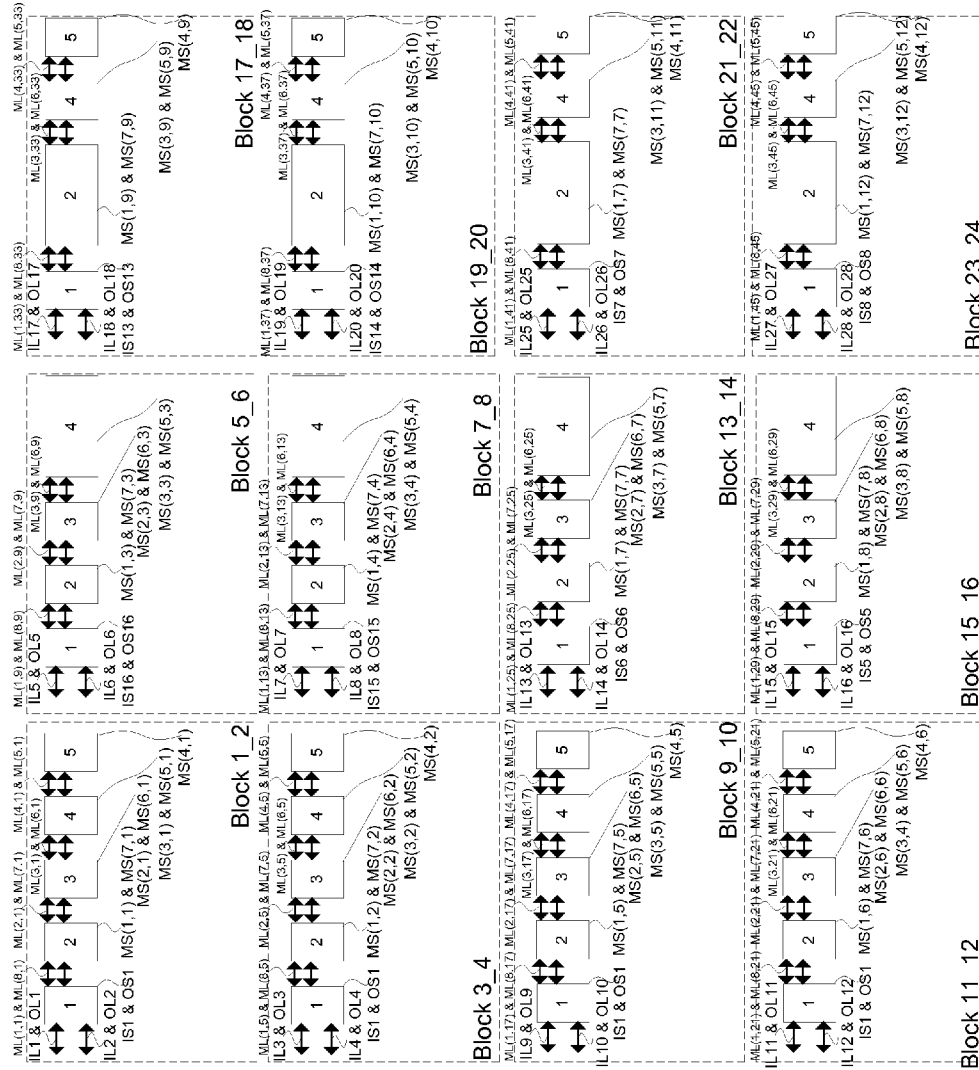
200B

FIG. 2B



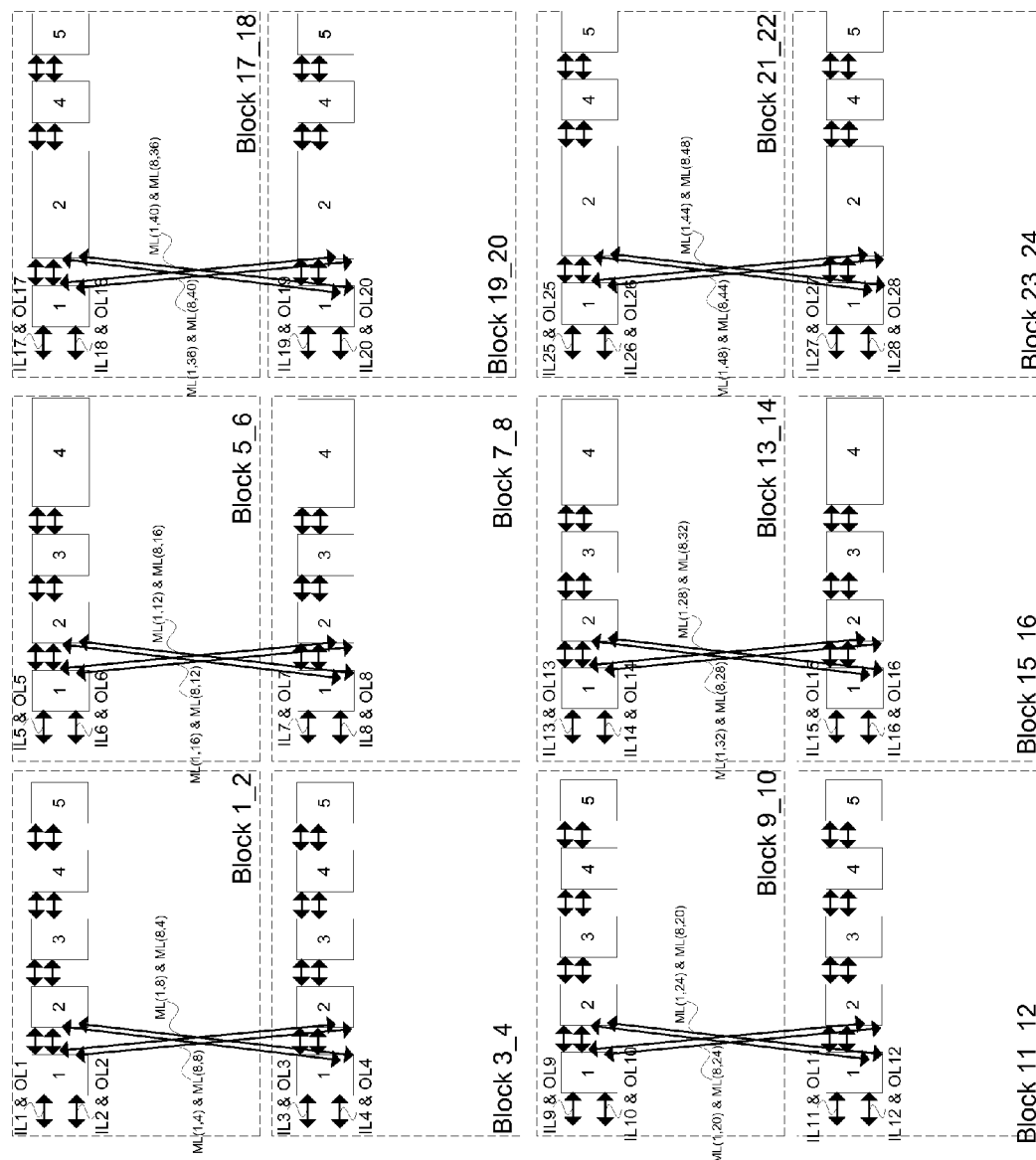
200C

FIG. 2C



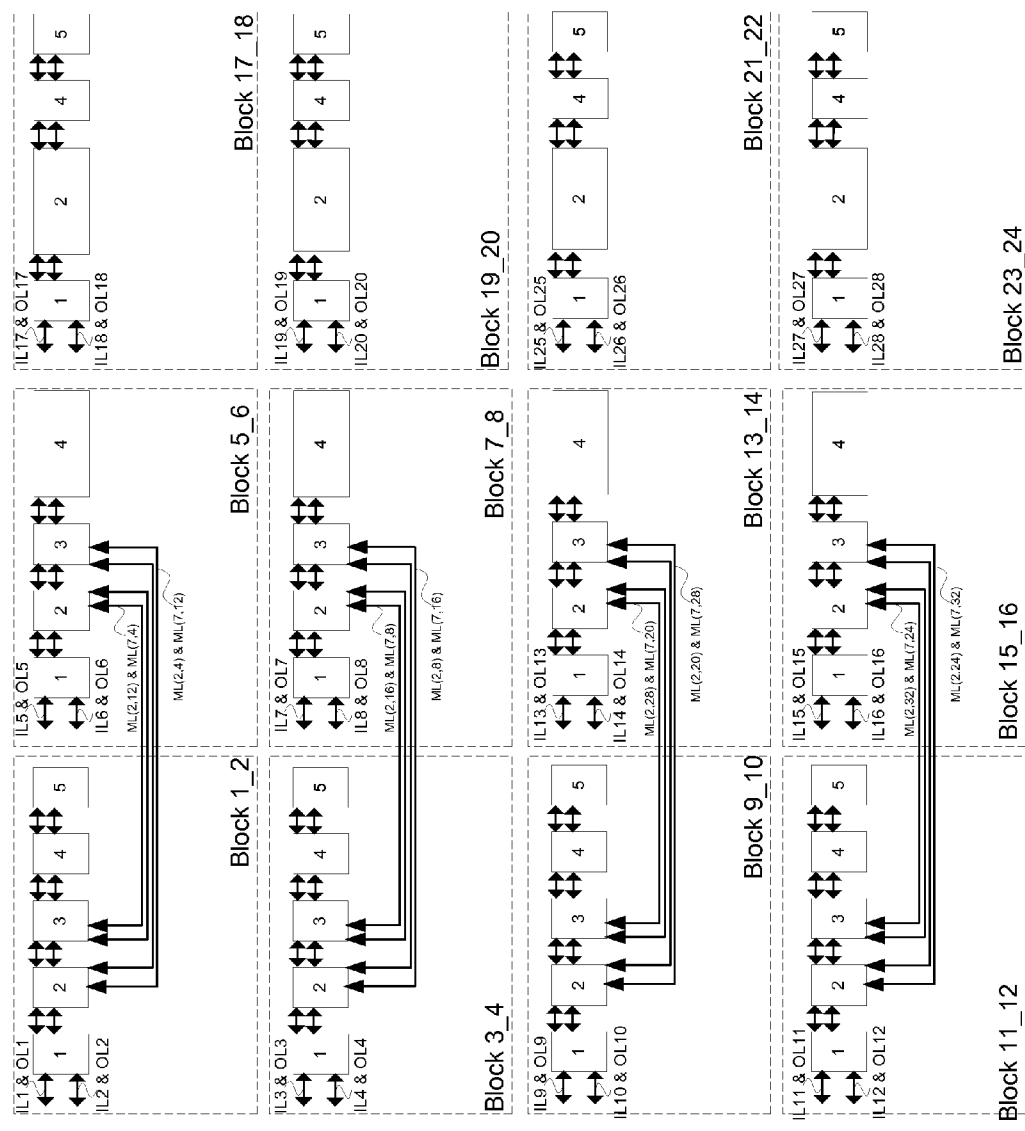
200D

FIG. 2D



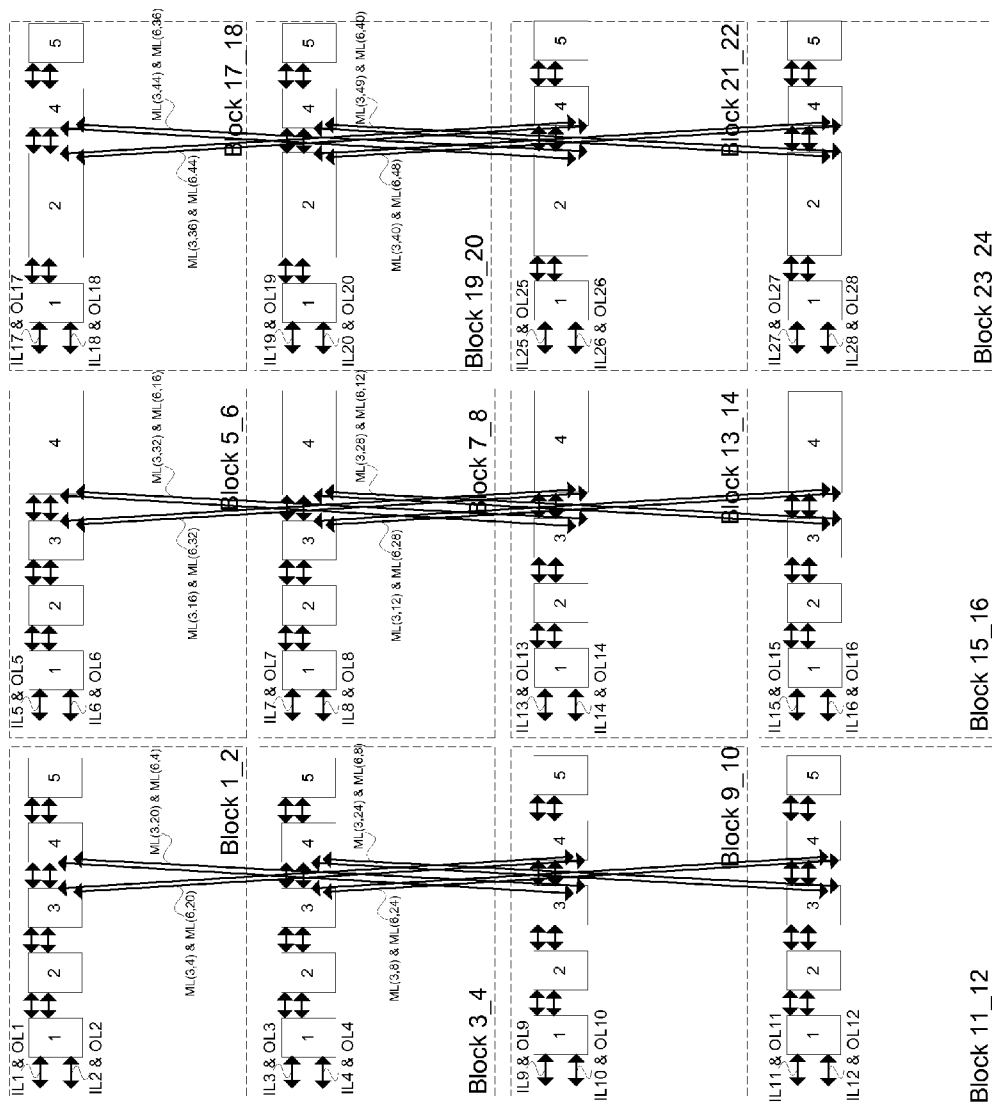
200E

FIG. 2E



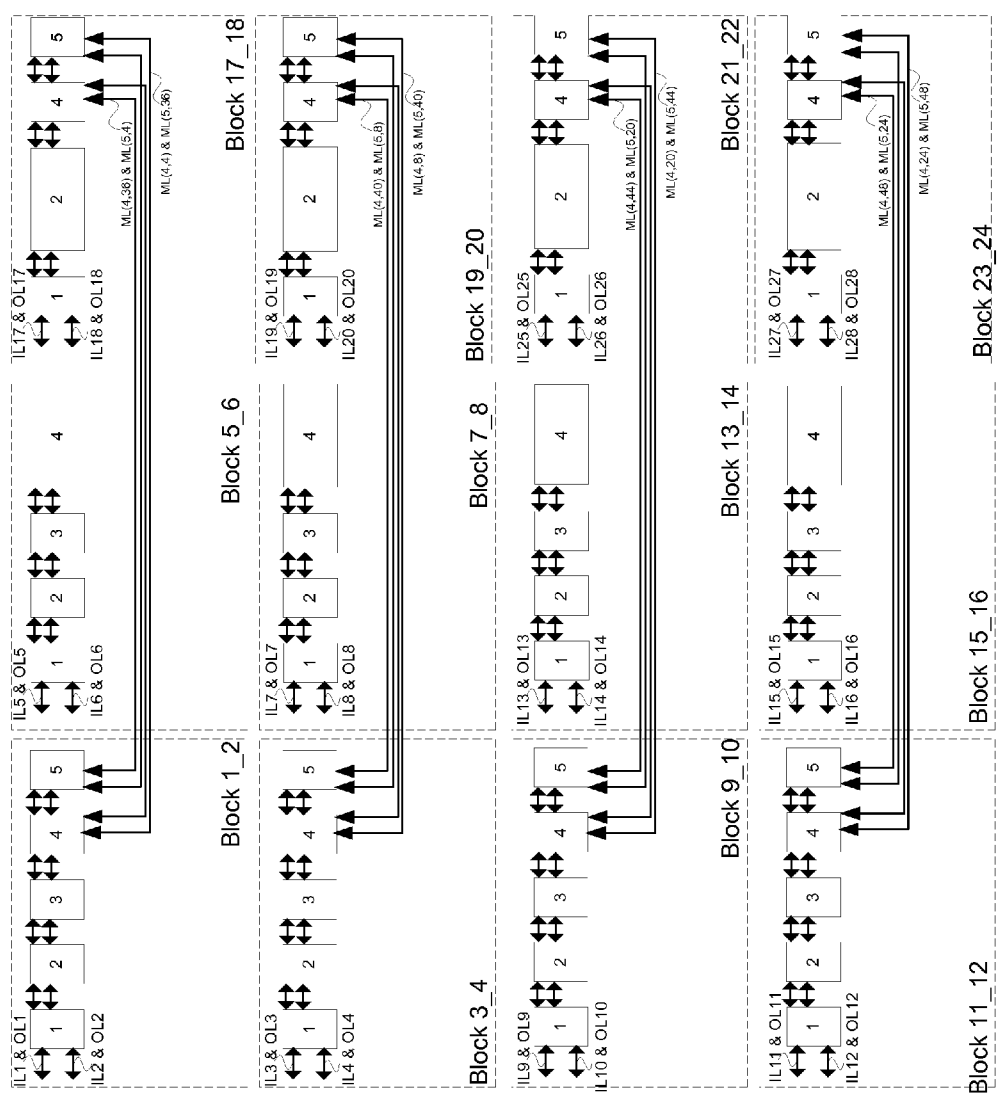
200F

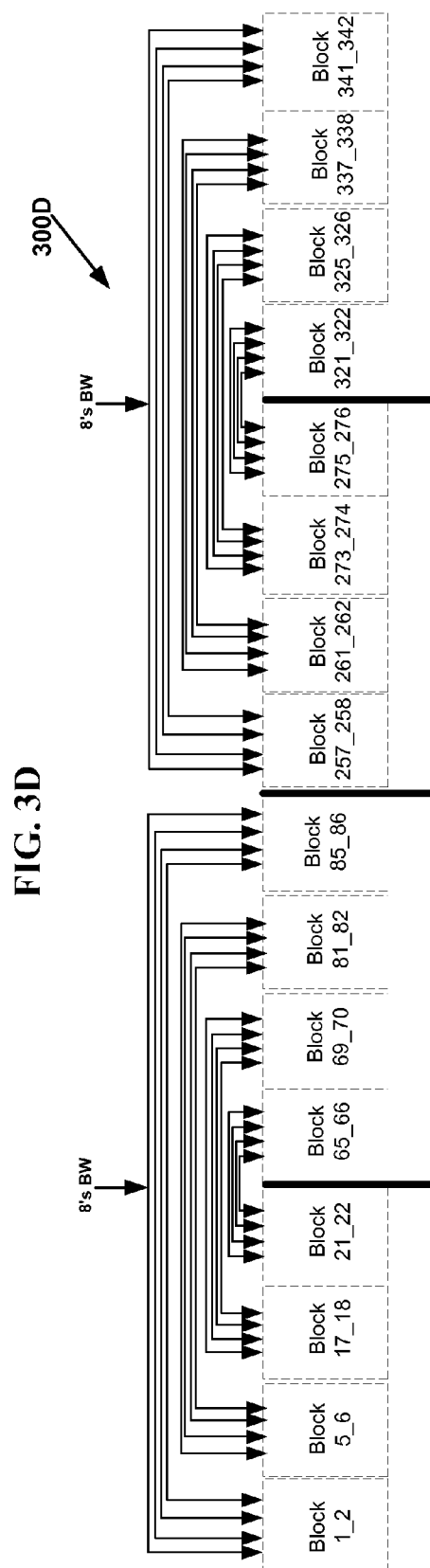
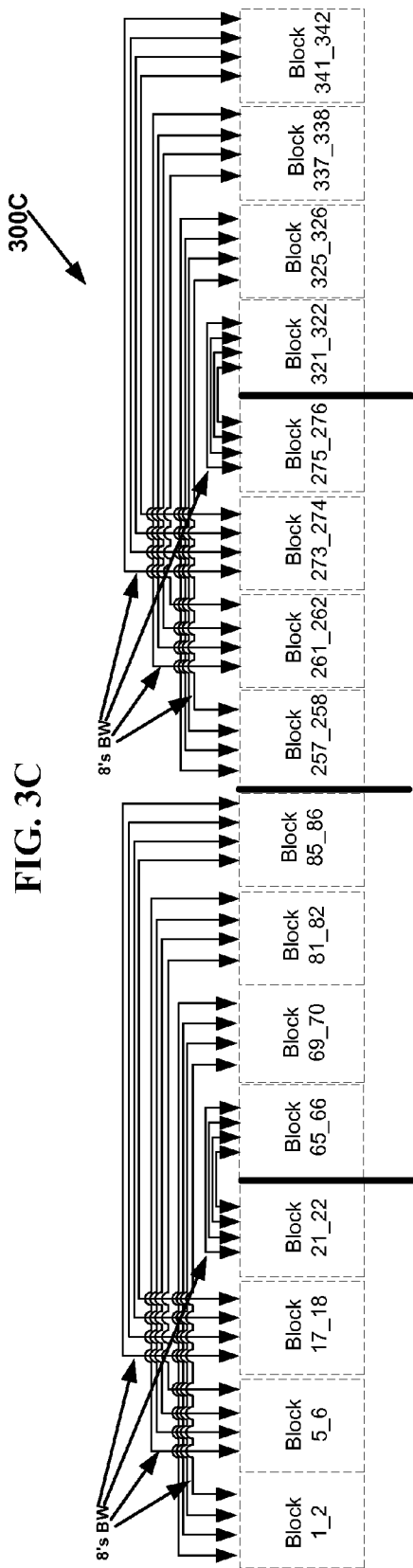
FIG. 2F

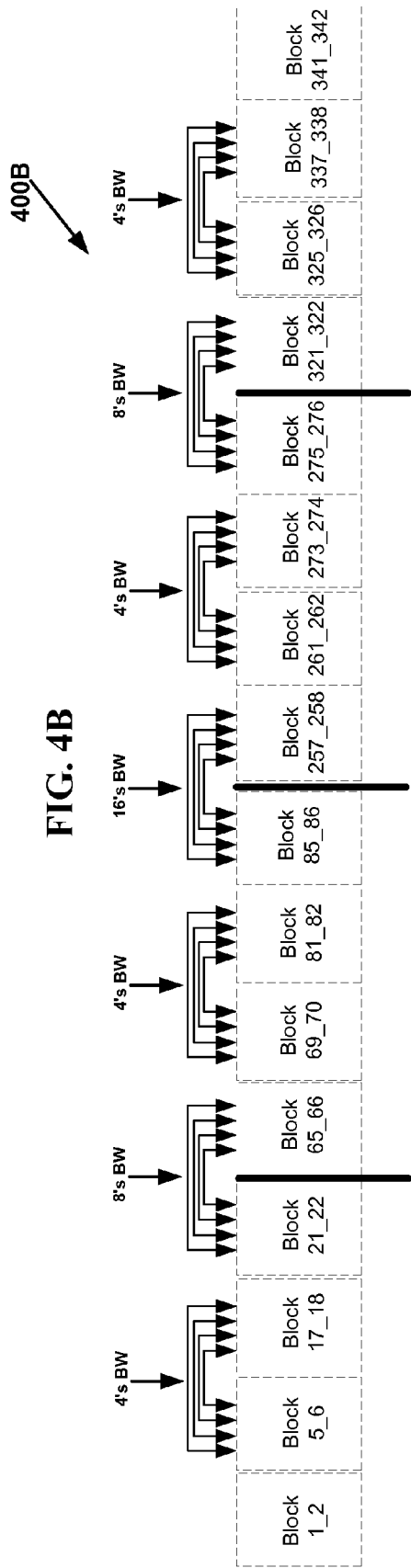
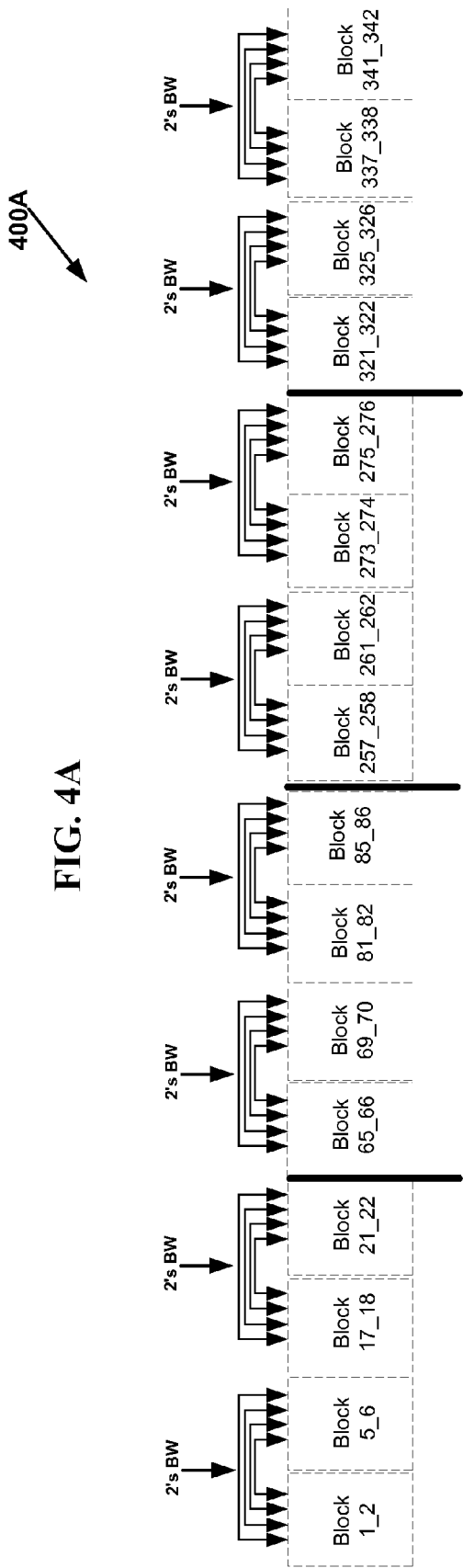


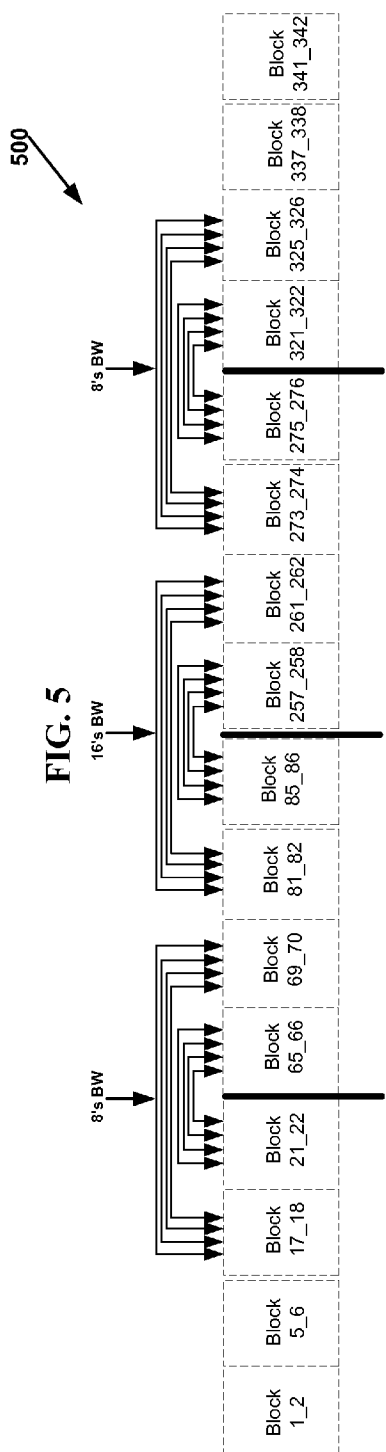
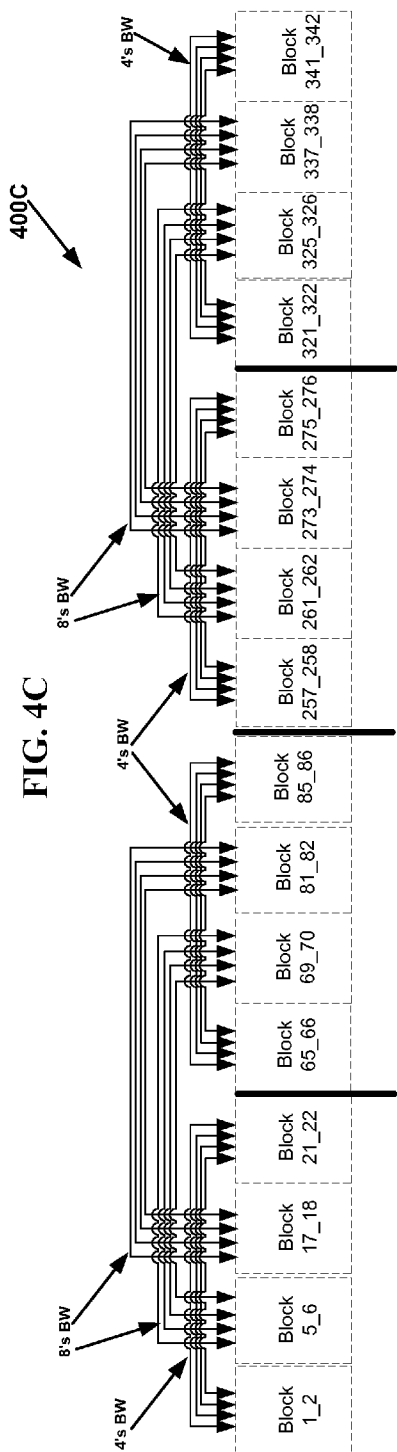
200G

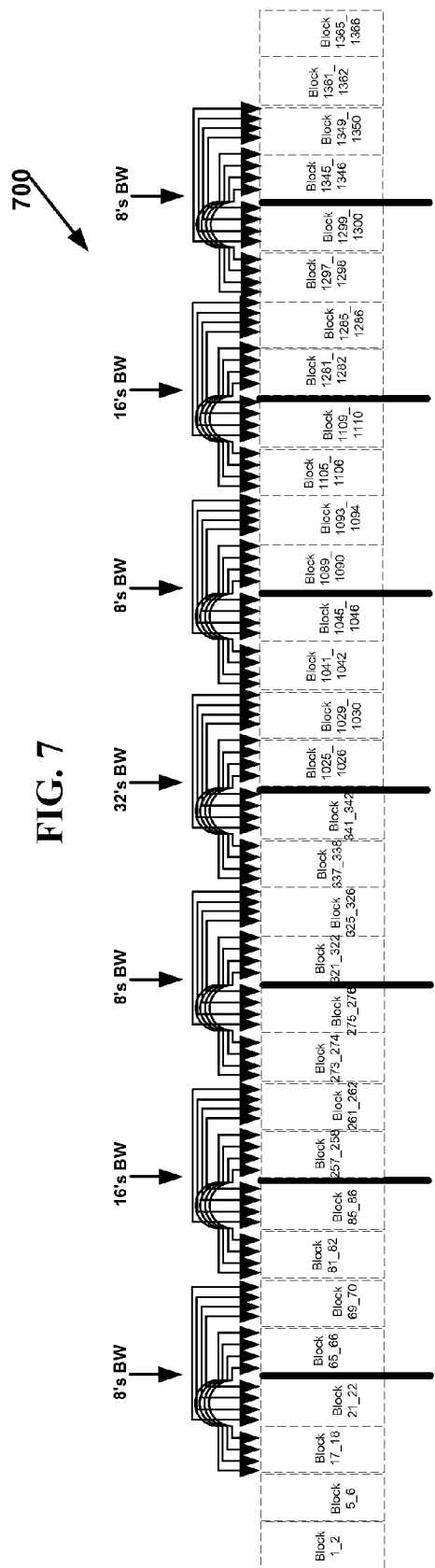
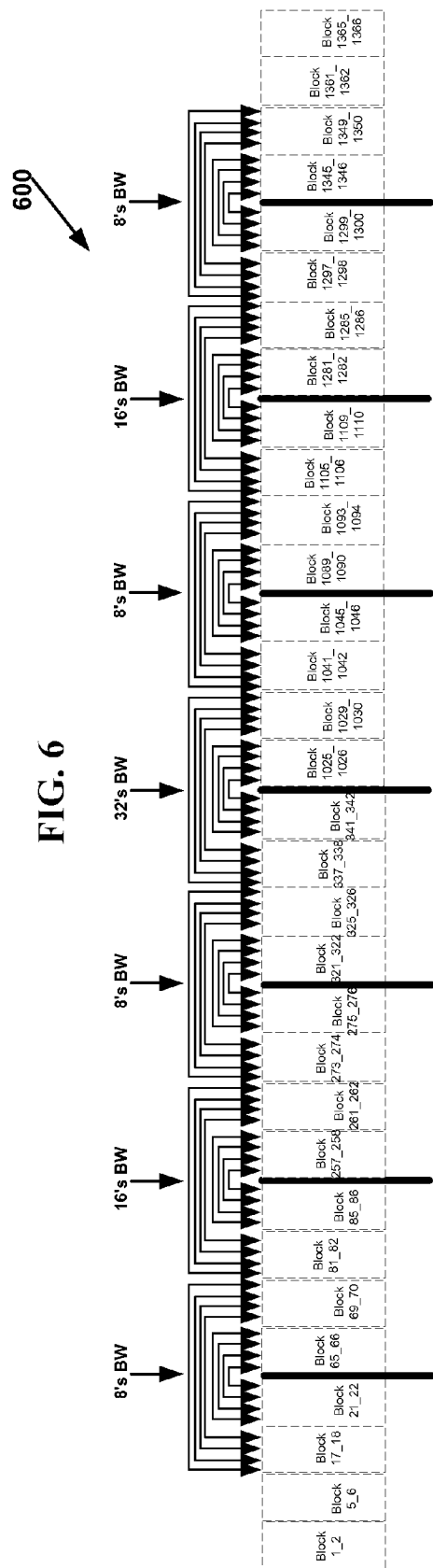
FIG. 2G











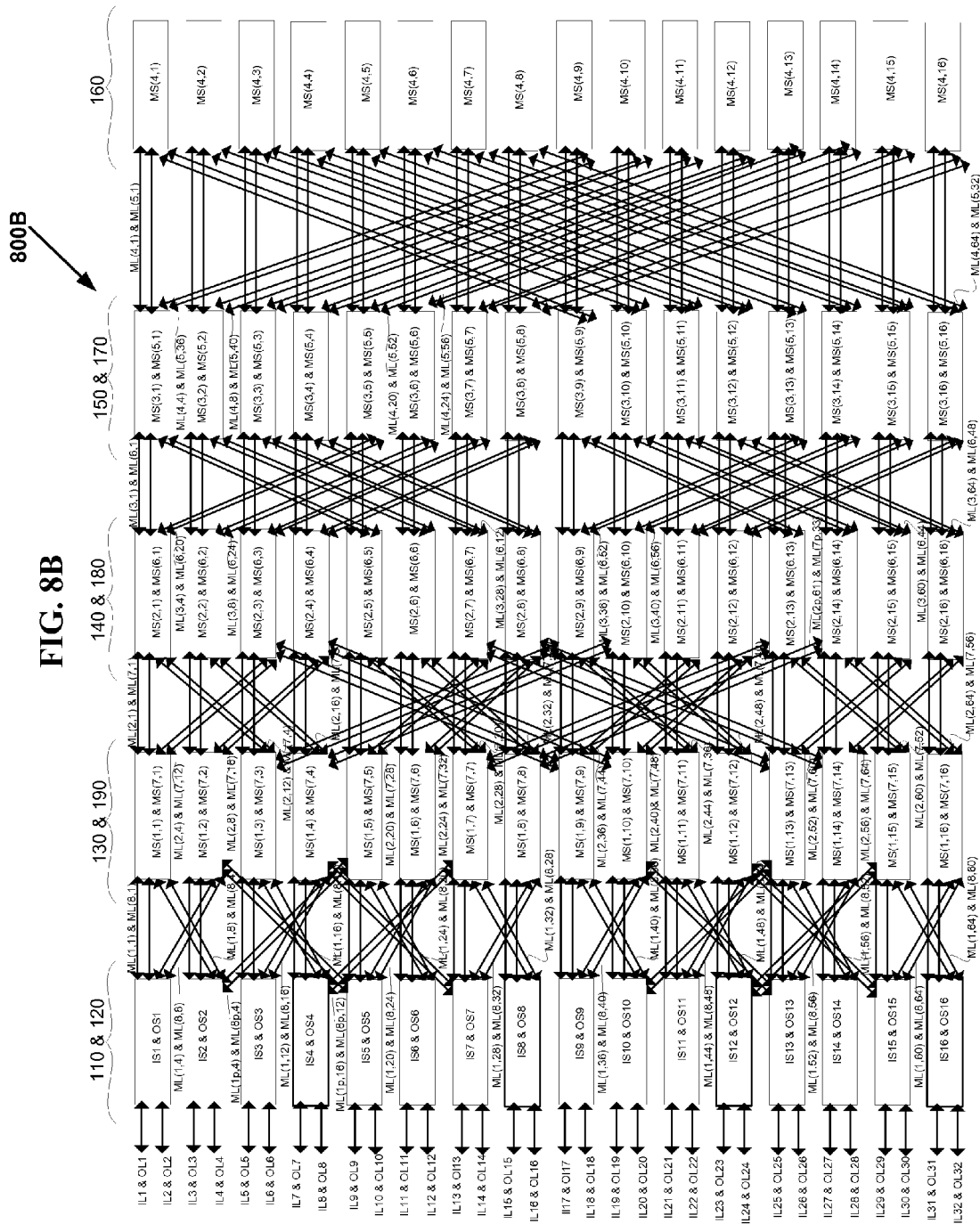
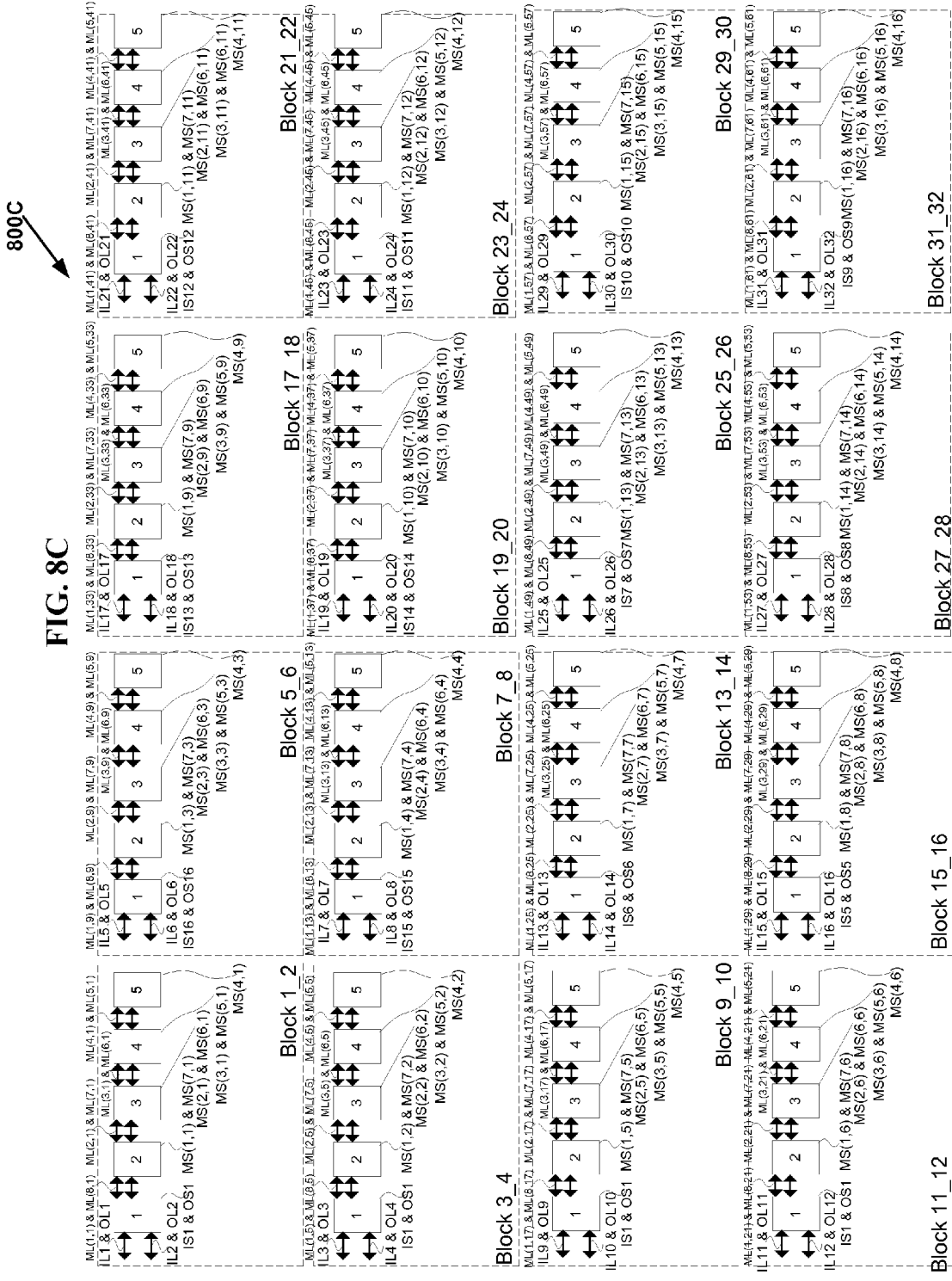


FIG. 8B



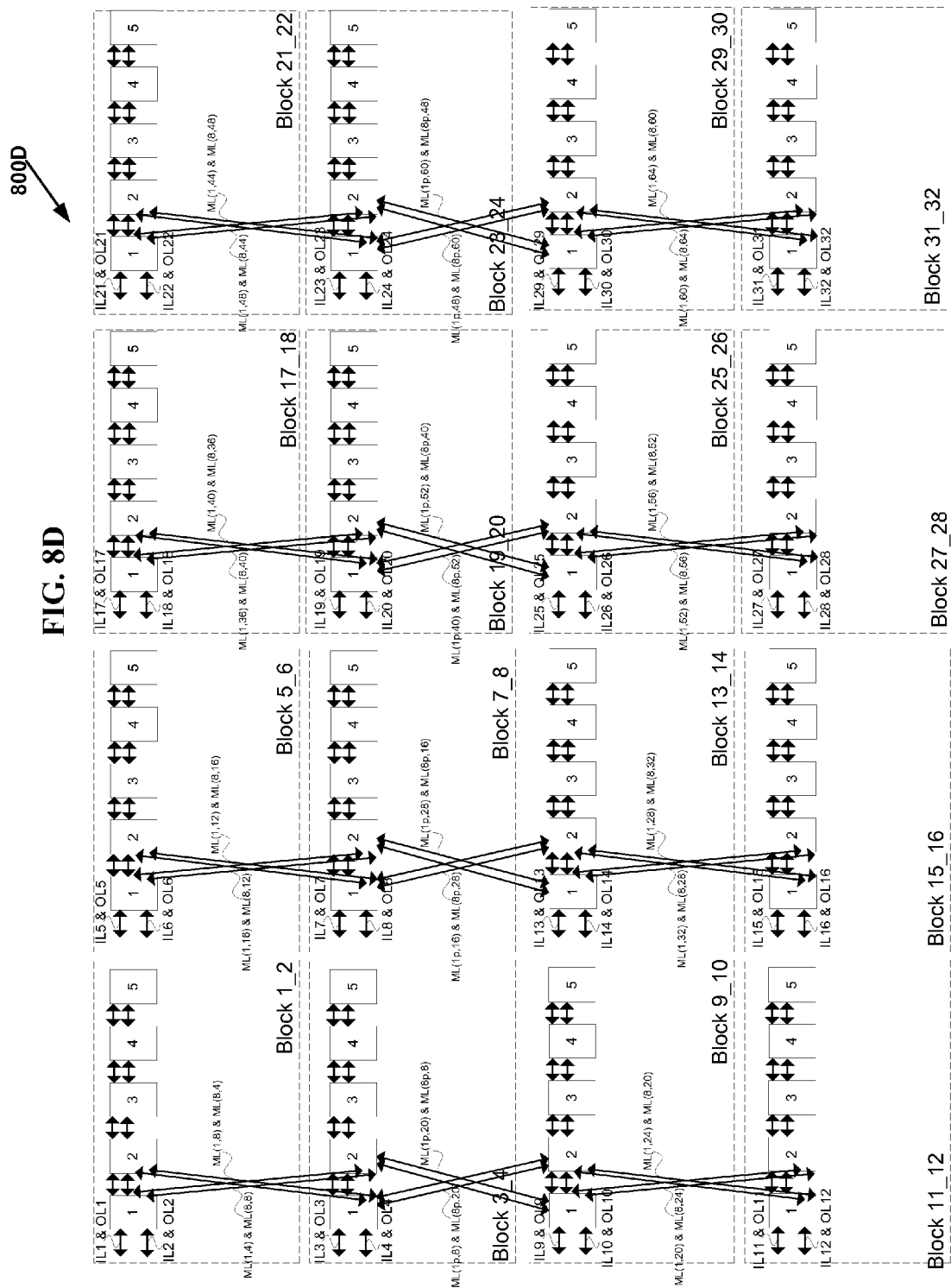
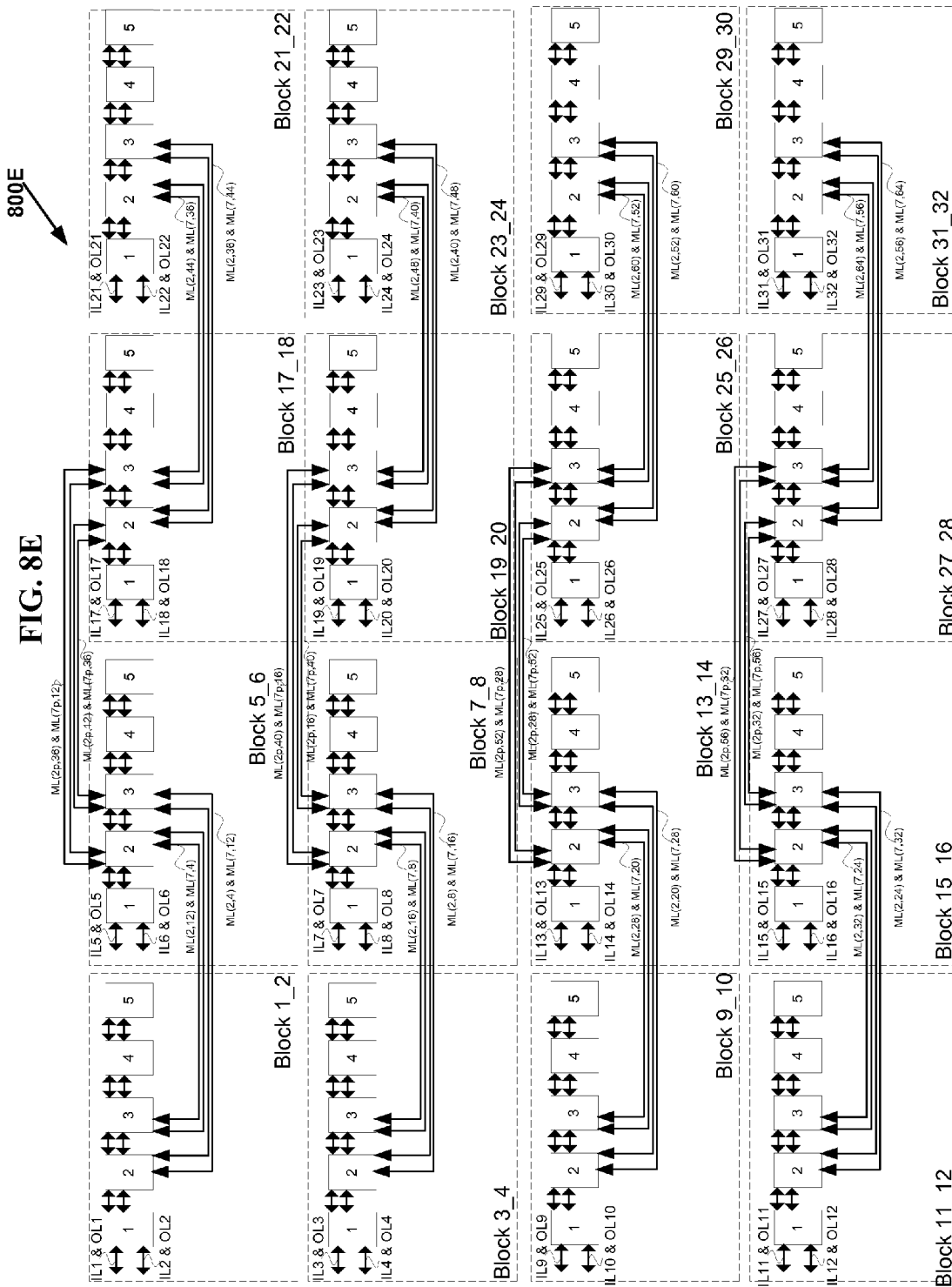
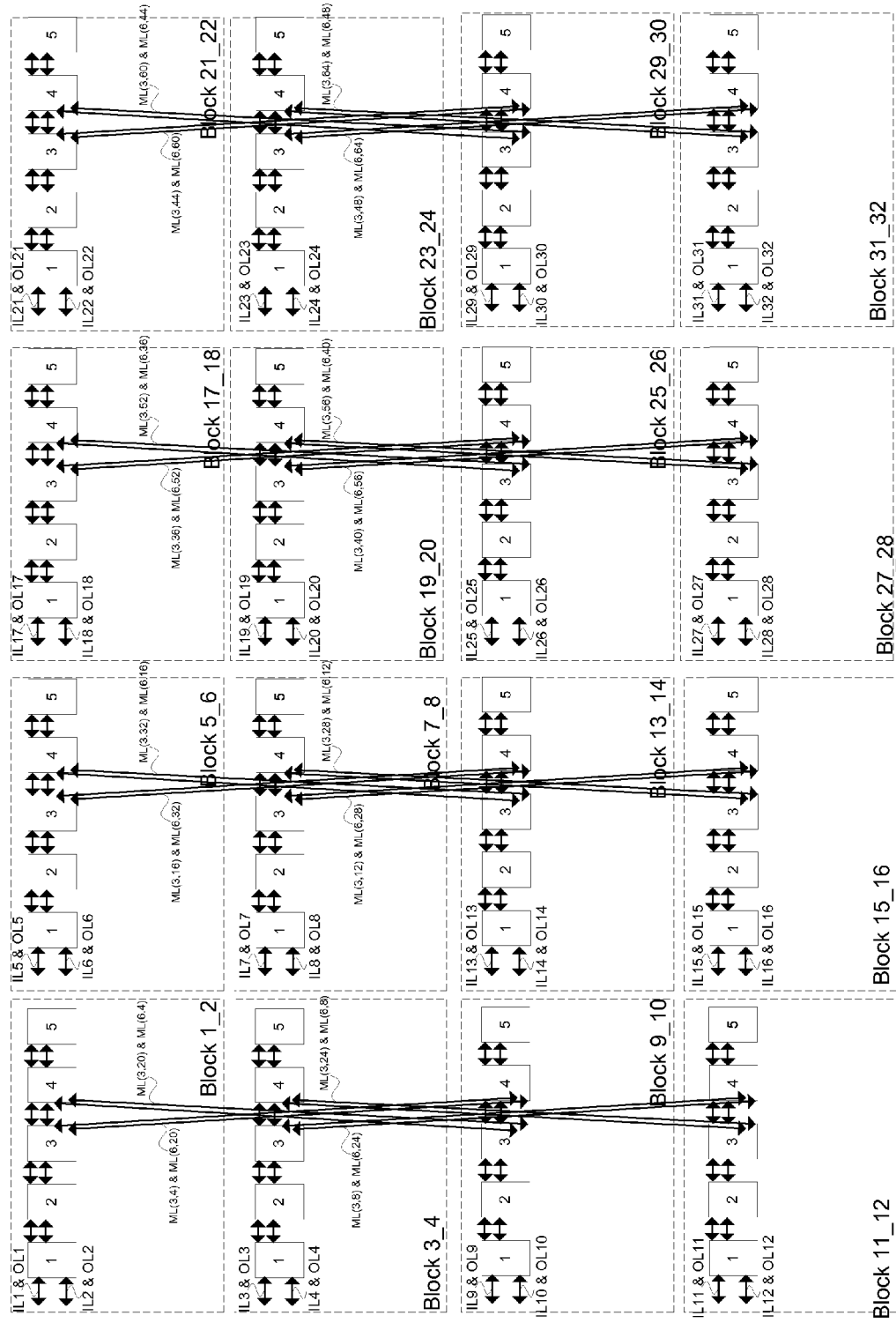


FIG. 8E



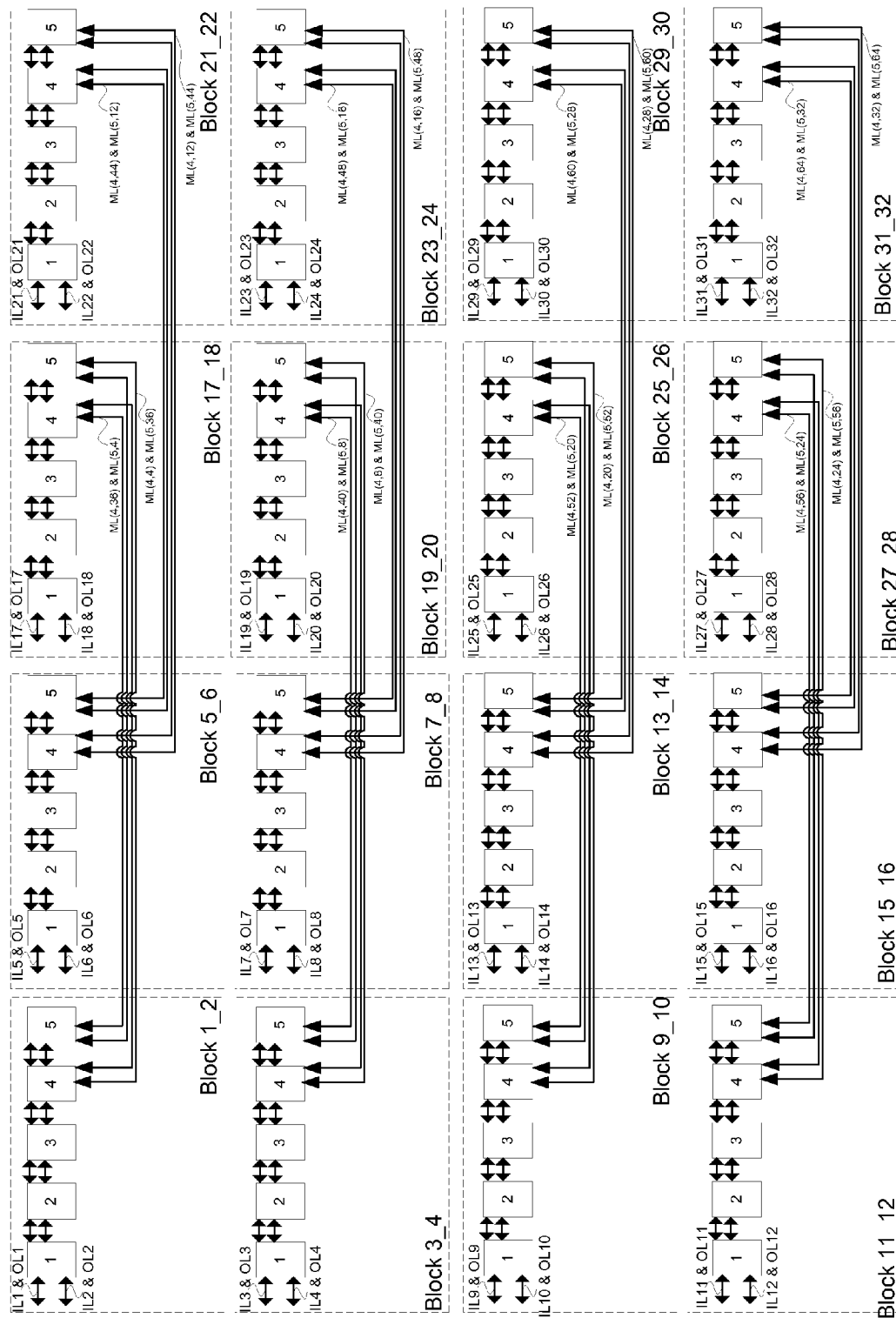
800F

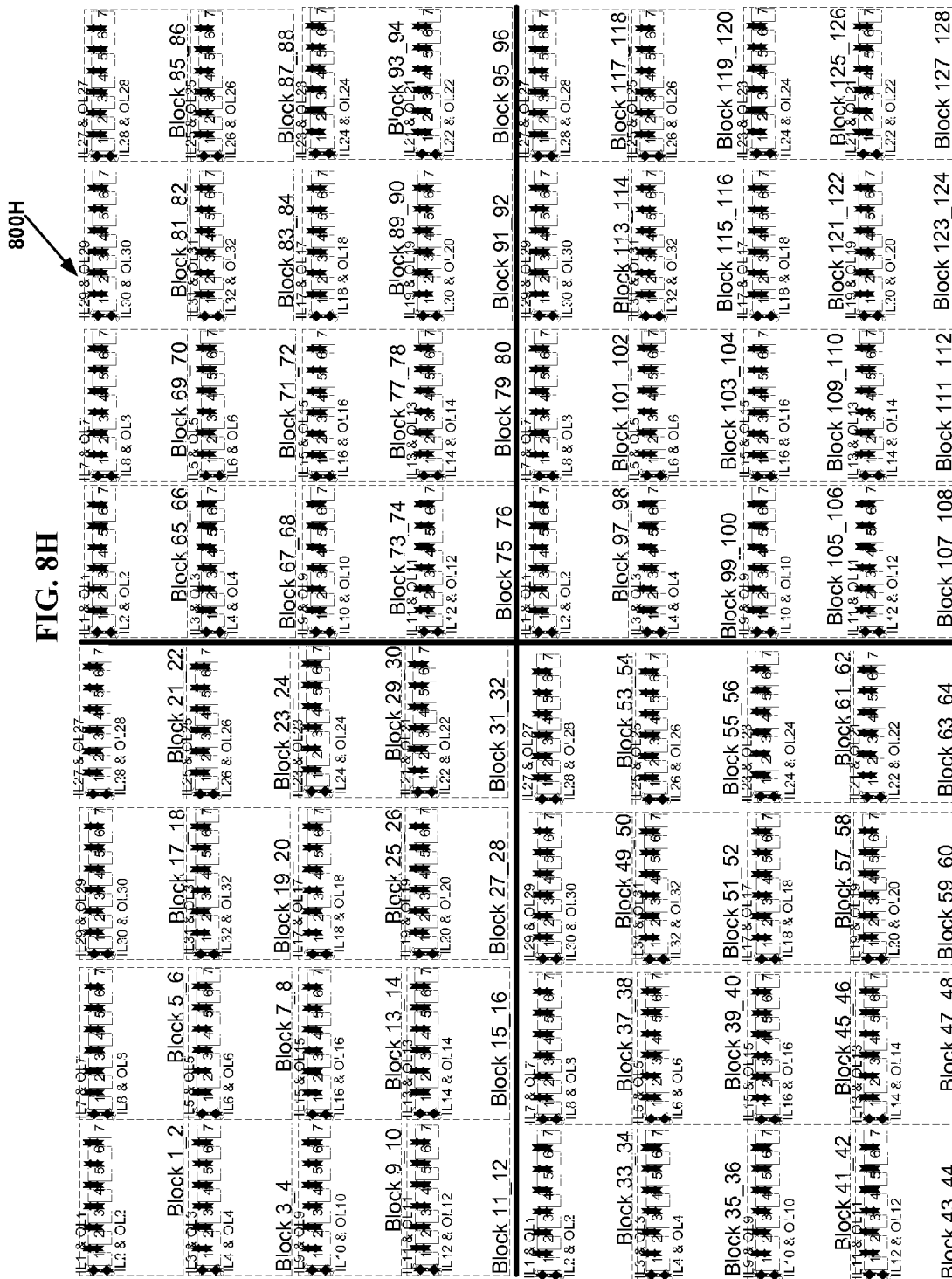
FIG. 8F



800G

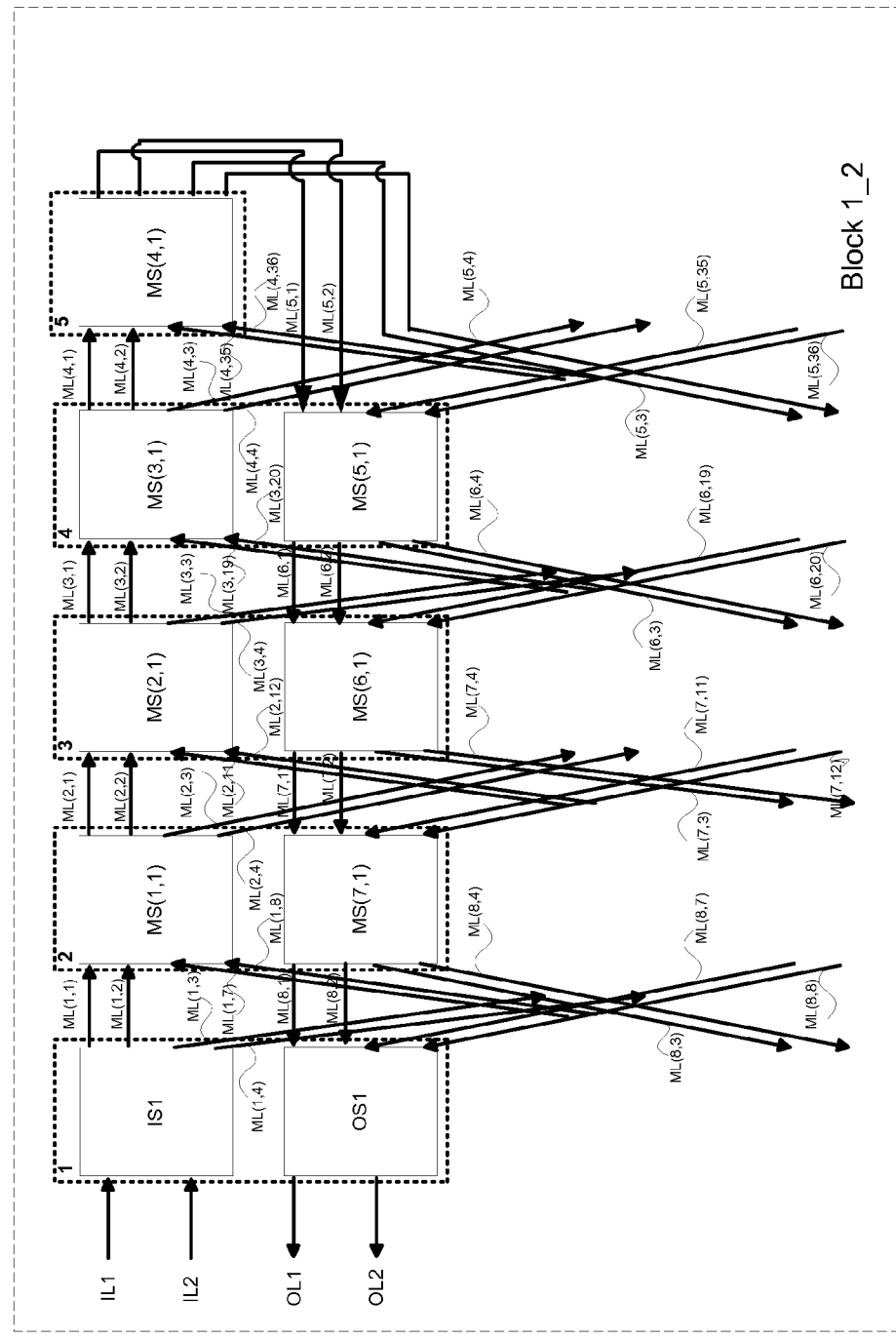
FIG. 8G





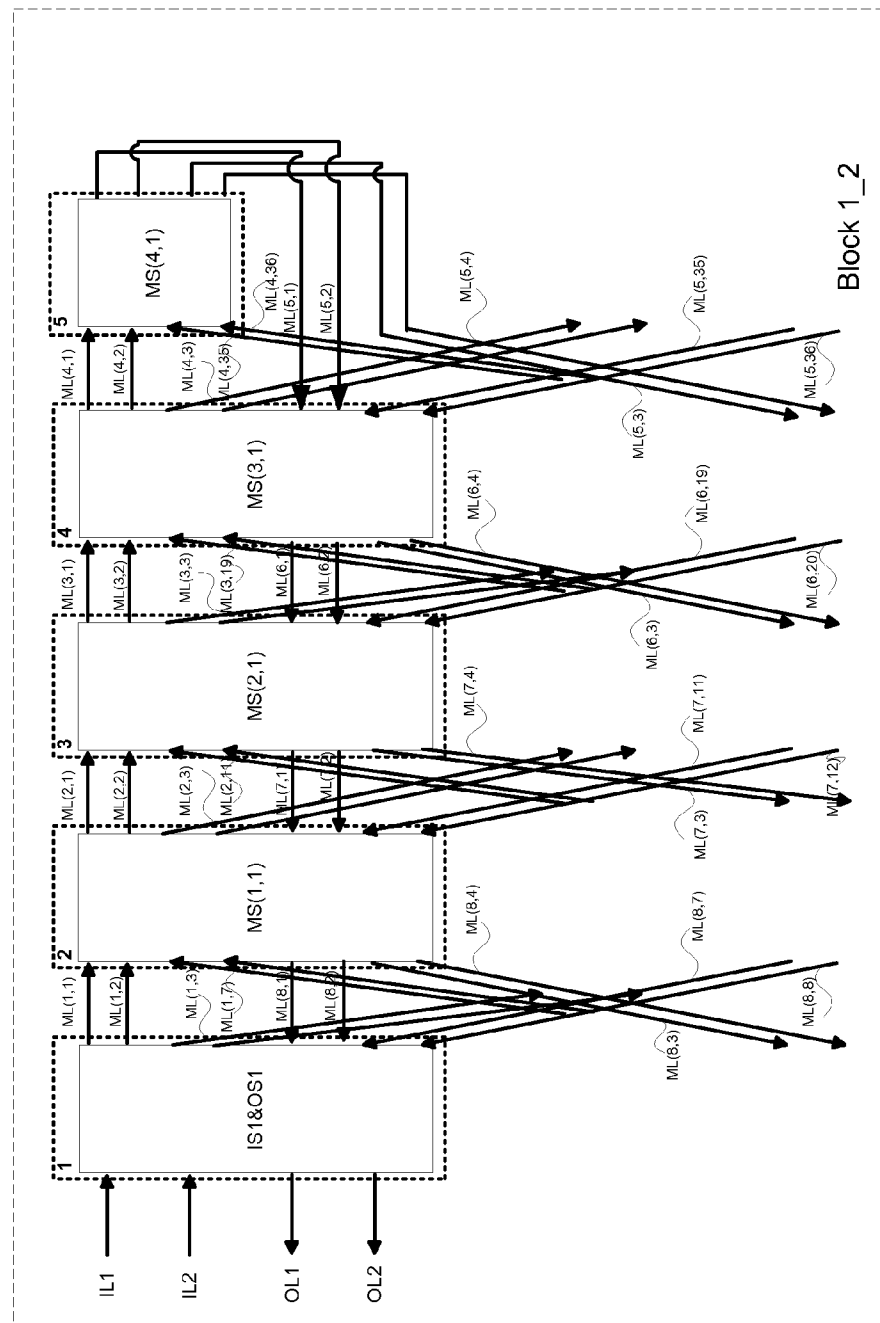
8001

FIG. 8I



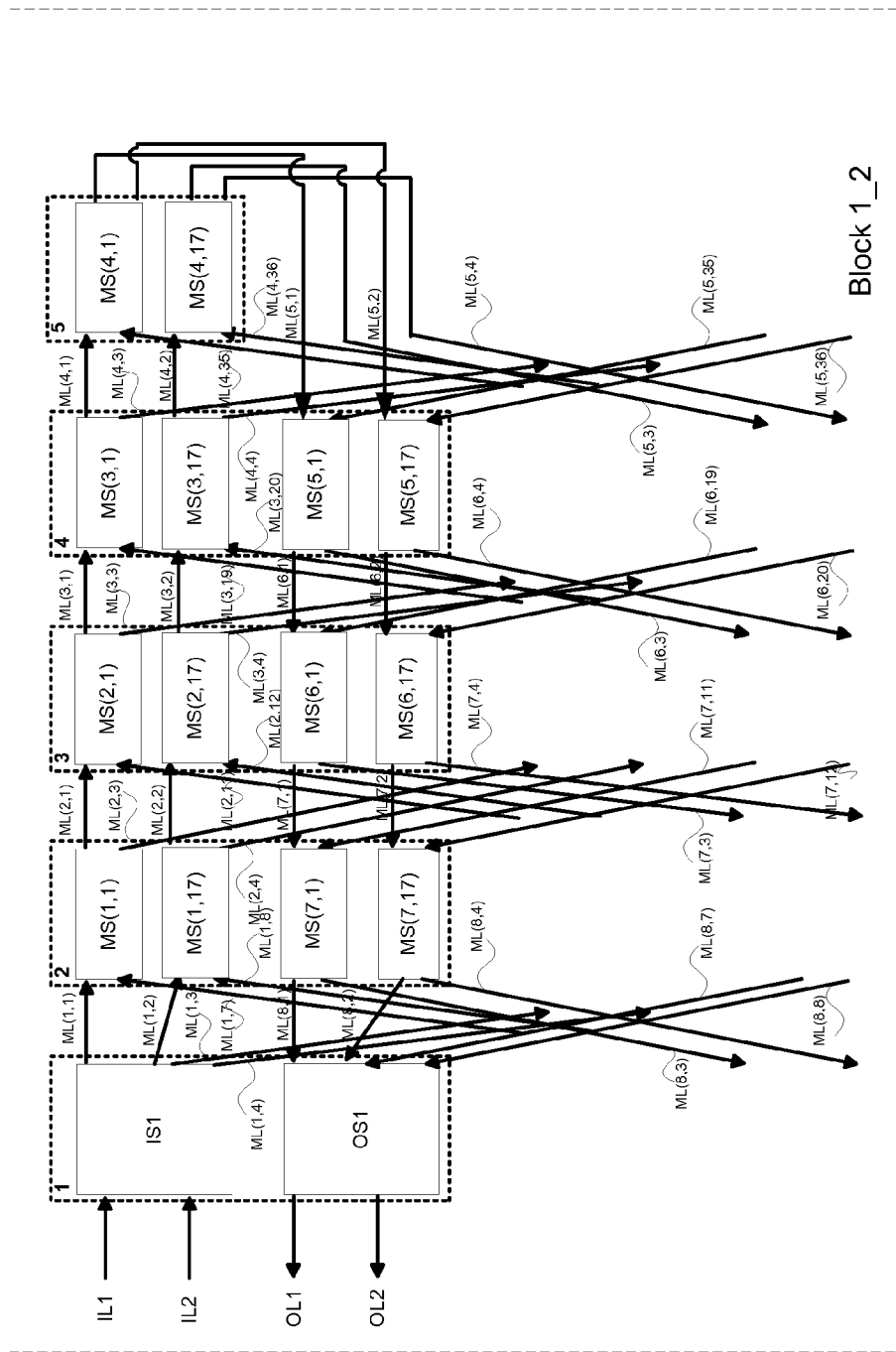
800J

FIG. 8J



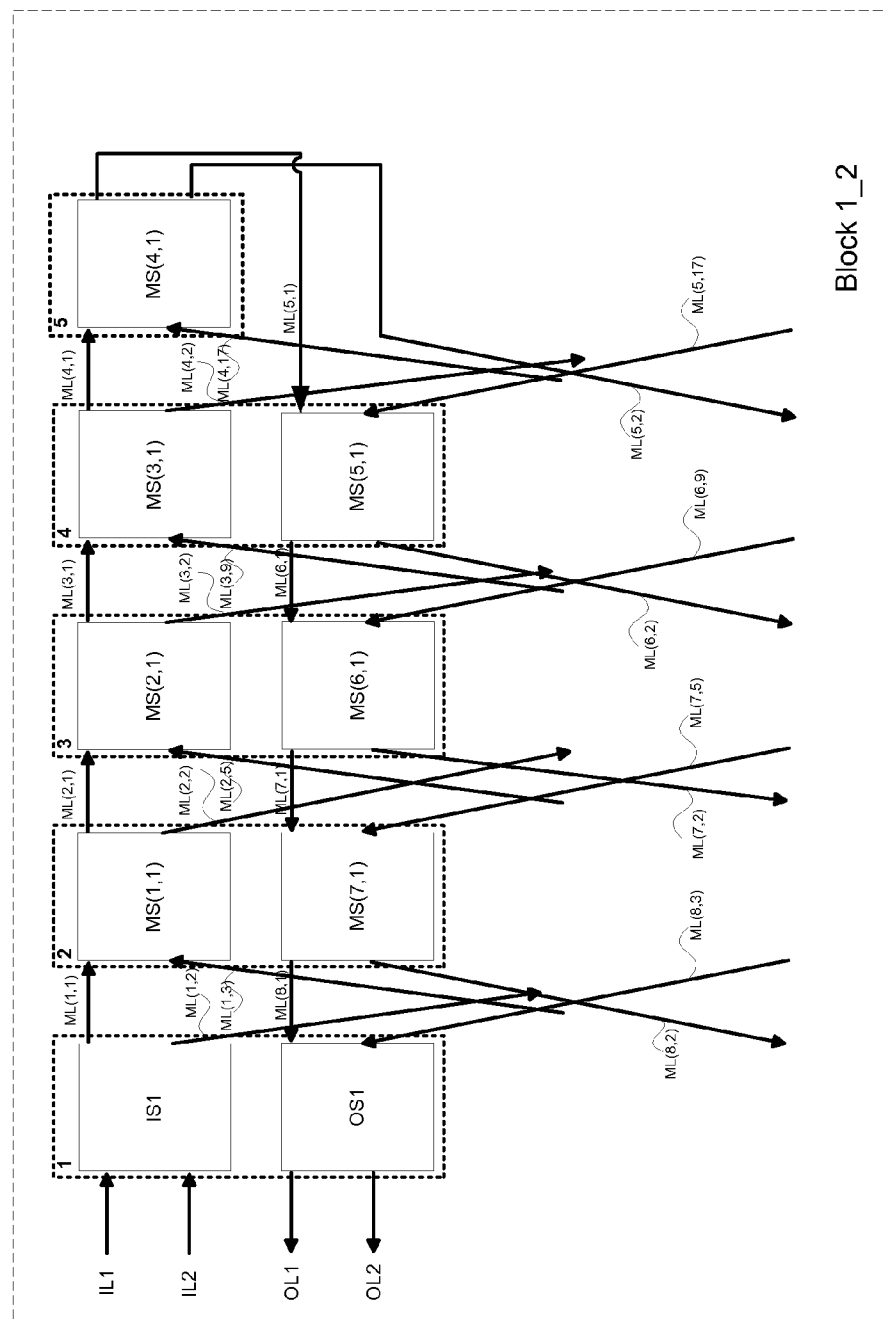
800K

FIG. 8K



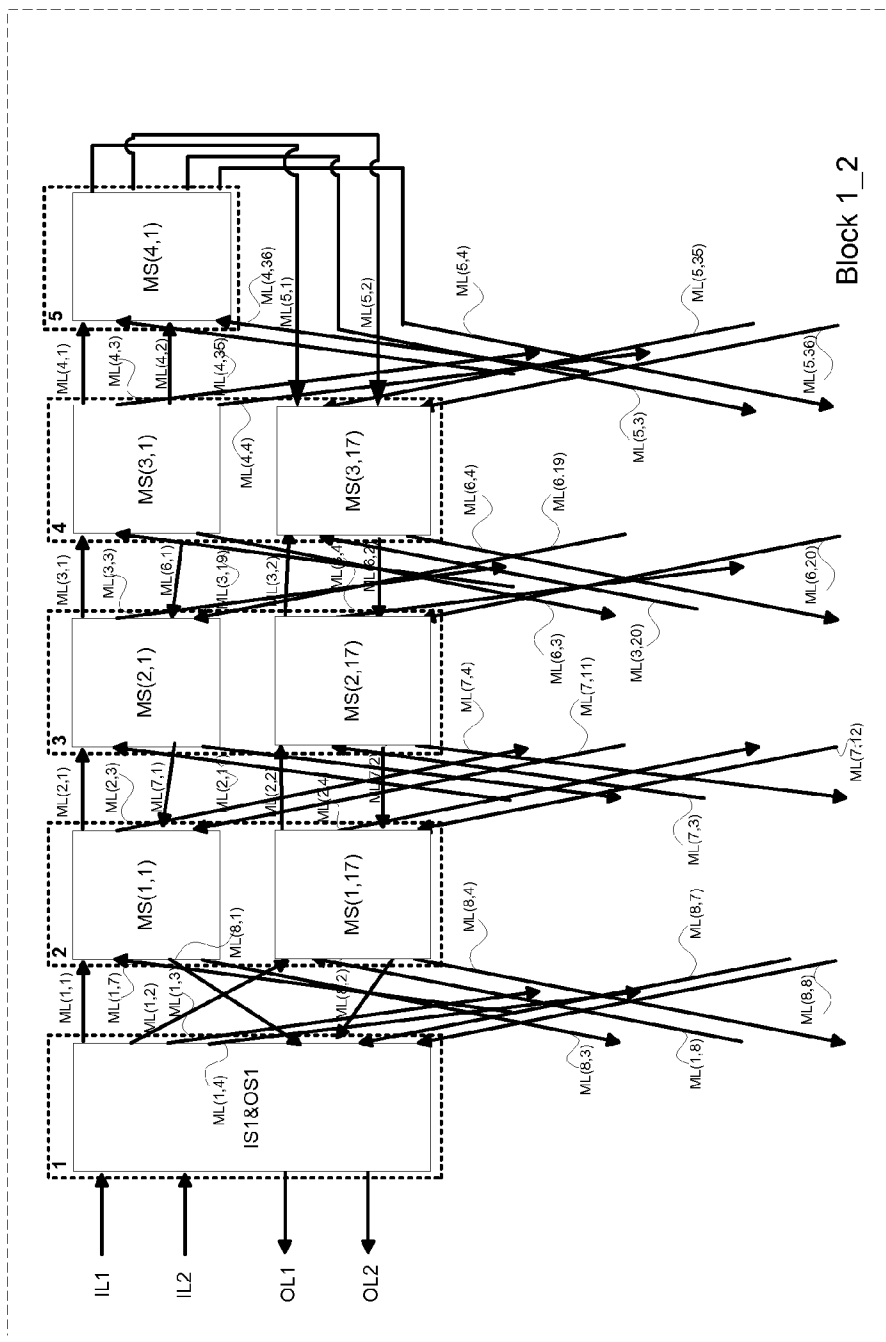
800K1

FIG. 8K1



800L

FIG. 8L



800L1

FIG. 8L1

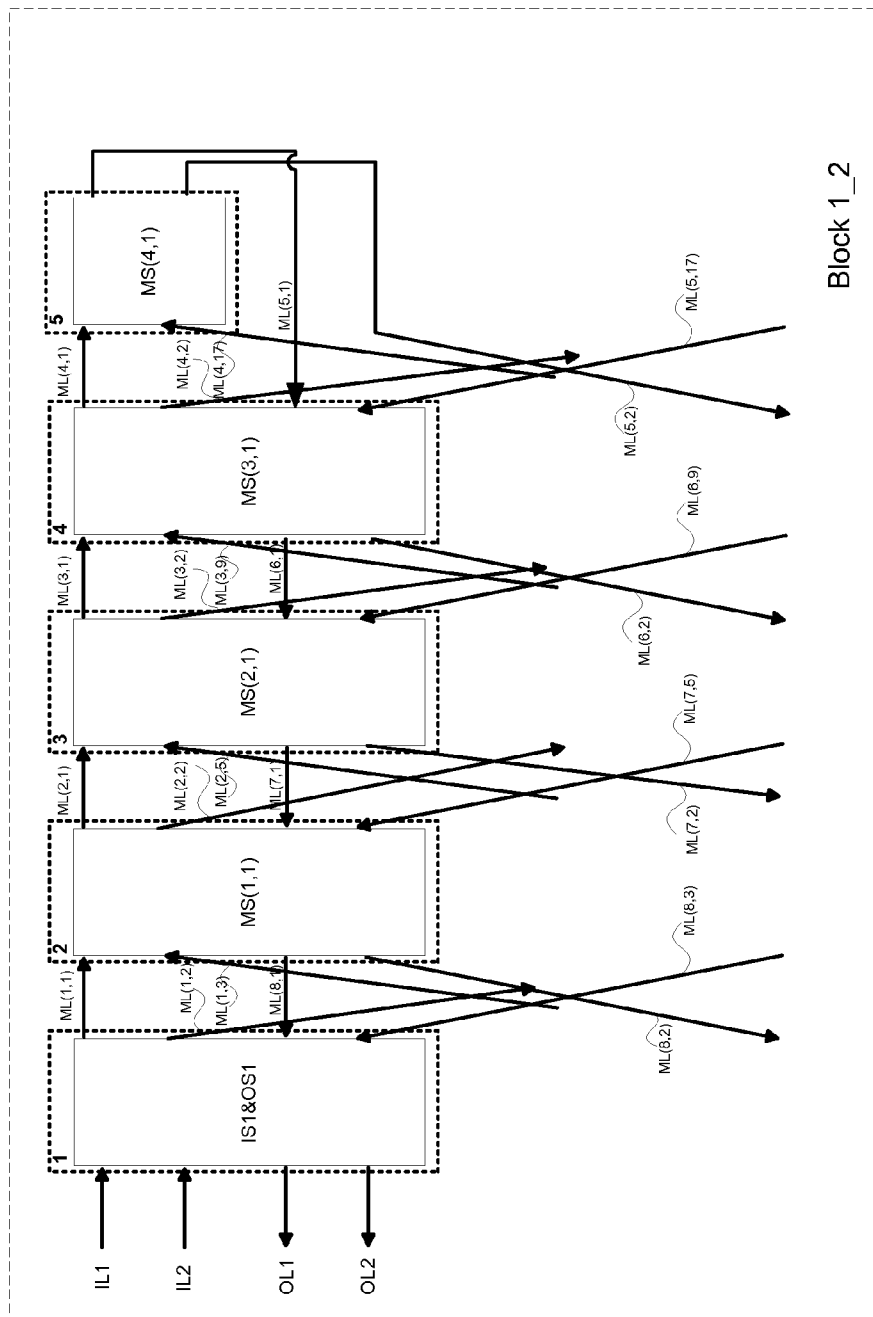


FIG. 9A

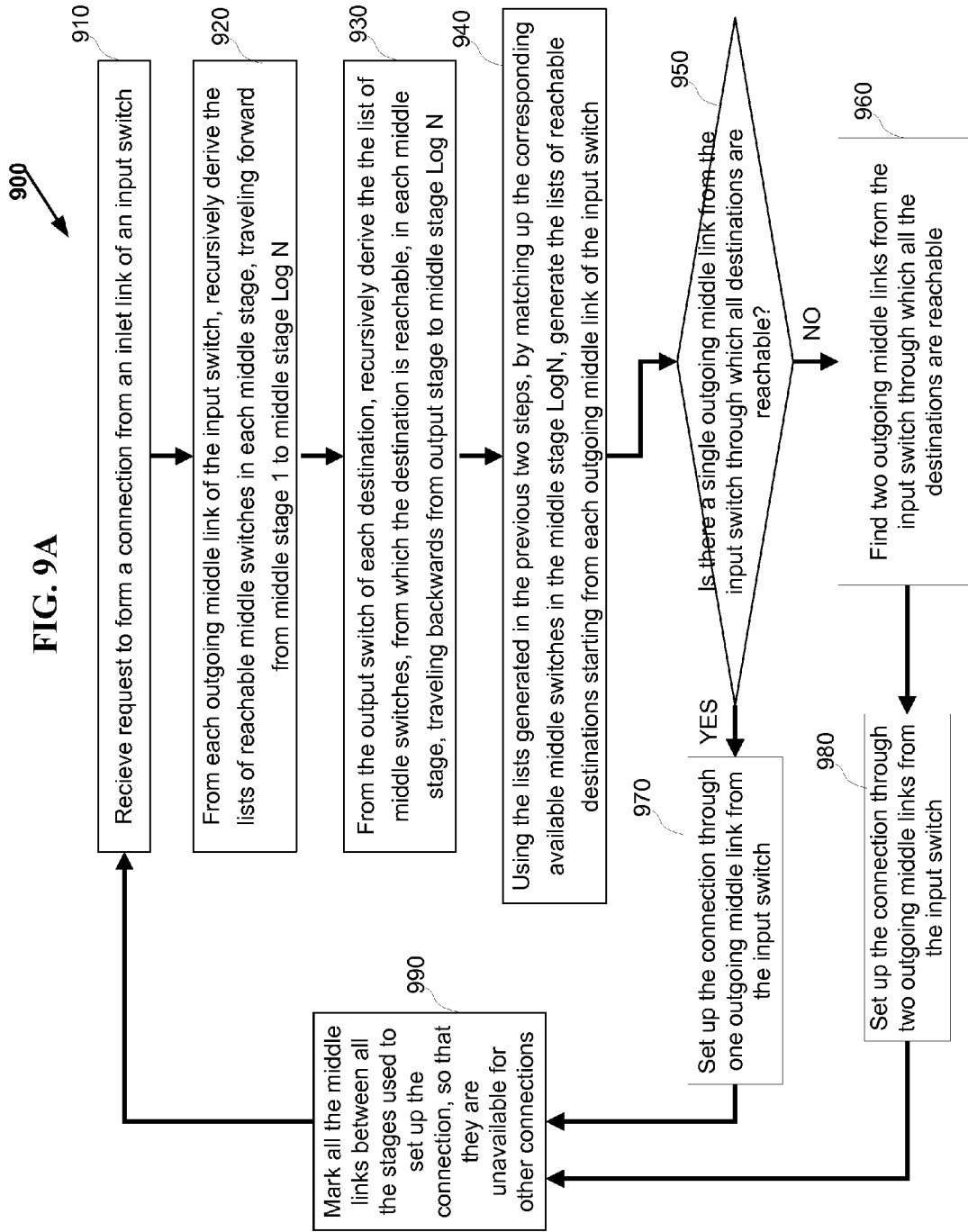


FIG. 10A

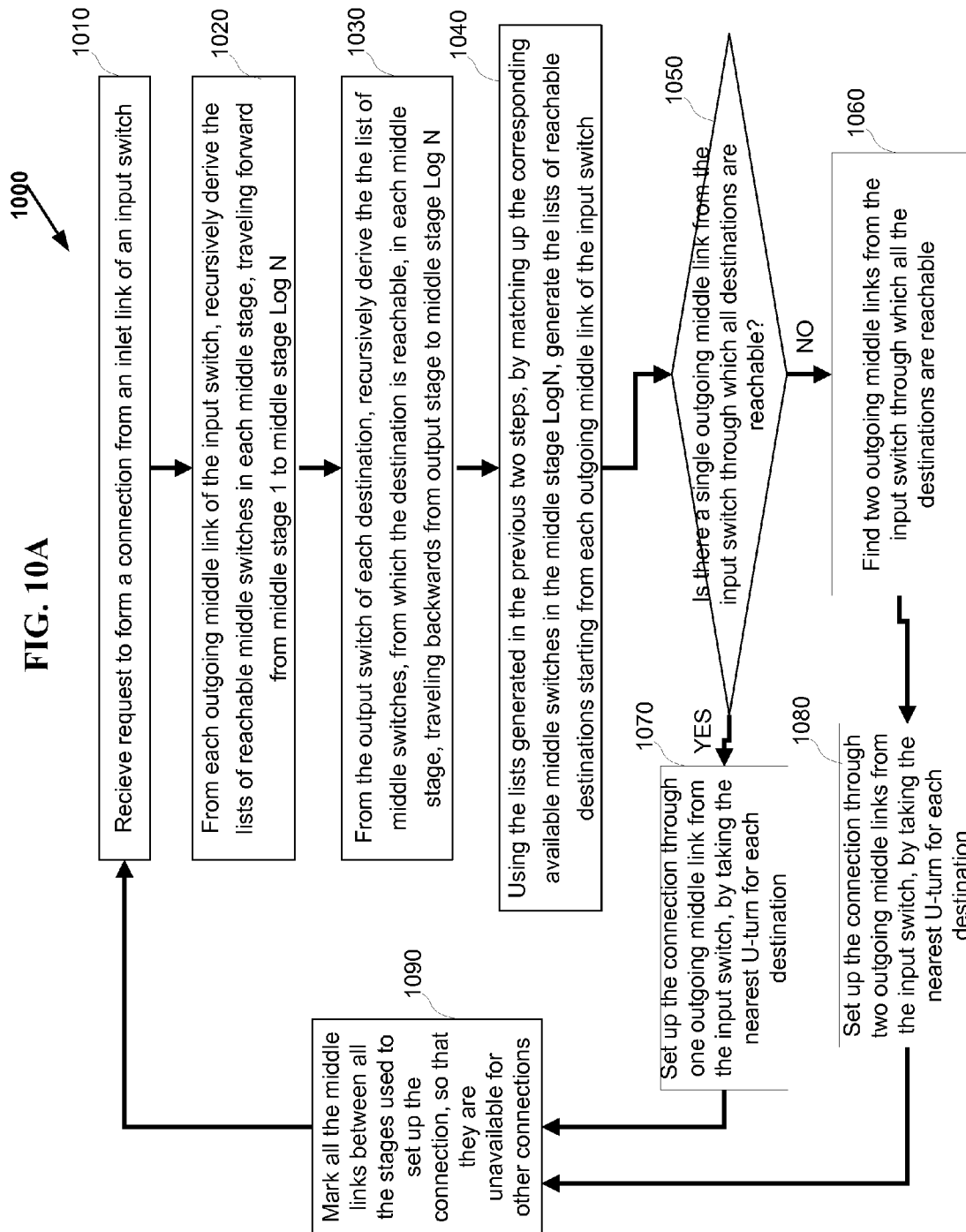


FIG. 11A

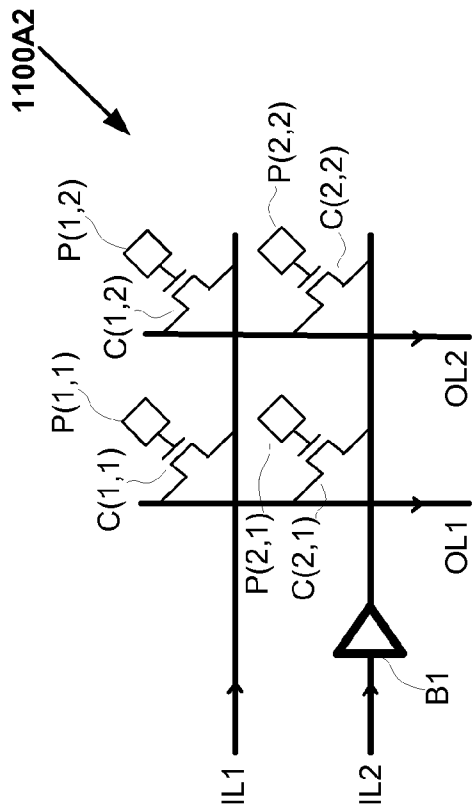


FIG. 11A2
(Prior Art)

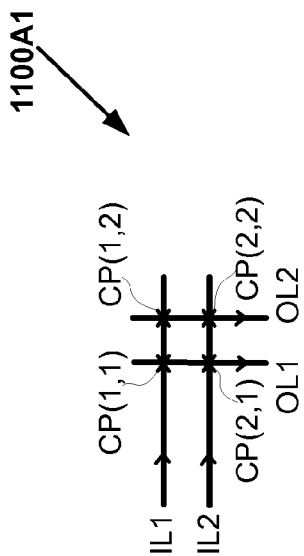


FIG. 11A1
(Prior Art)

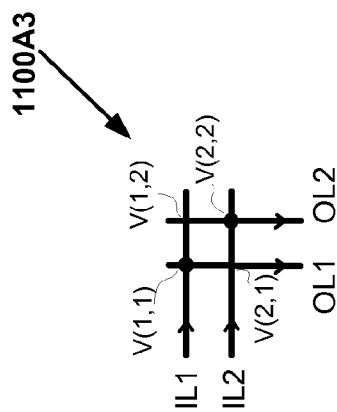


FIG. 11A3
(Prior Art)

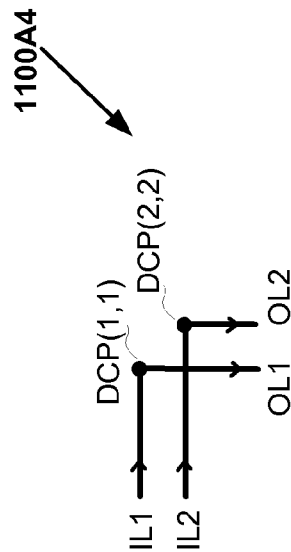


FIG. 11A4

US 2012/0269190 A1

Oct. 25, 2012

1

**VLSI LAYOUTS OF FULLY CONNECTED
GENERALIZED AND PYRAMID NETWORKS
WITH LOCALITY EXPLOITATION**

**CROSS REFERENCE TO RELATED
APPLICATIONS**

[0001] This application is Continuation In Part PCT Application to and incorporates by reference in its entirety the U.S. Provisional Patent Application Ser. No. 61/252,603 entitled "VLSI LAYOUTS OF FULLY CONNECTED NETWORKS WITH LOCALITY EXPLOITATION" by Venkat Konda assigned to the same assignee as the current application, filed Oct. 16, 2009.

[0002] This application is Continuation In Part PCT Application to and incorporates by reference in its entirety the U.S. Provisional Patent Application Ser. No. 61/252,609 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Oct. 16, 2009.

[0003] This application is related to and incorporates by reference in its entirety the U.S. application Ser. No. 12/530,207 entitled "FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Sep. 6, 2009, the U.S. Provisional Patent Application Ser. No. 60/905,526 entitled "LARGE SCALE CROSSPOINT REDUCTION WITH NONBLOCKING UNICAST & MULTICAST IN ARBITRARILY LARGE MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Mar. 6, 2007, and the U.S. Provisional Patent Application Ser. No. 60/940,383 entitled "FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0004] This application is related to and incorporates by reference in its entirety the U.S. application Ser. No. 12/601,273 entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Nov. 22, 2009, the U.S. Provisional Patent Application Ser. No. 60/940,387 entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, and the U.S. Provisional Patent Application Ser. No. 60/940,390 entitled "FULLY CONNECTED GENERALIZED MULTI-LINK BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0005] This application is related to and incorporates by reference in its entirety the U.S. application Ser. No. 12/601,274 entitled "FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Nov. 22, 2009, the U.S. Provisional Patent Application Ser. No. 60/940,389 entitled "FULLY CONNECTED GENERALIZED REARRANGEABLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, the U.S. Provisional Patent Application Ser. No. 60/940,391 entitled "FULLY CONNECTED GENERALIZED FOLDED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007 and the U.S. Provisional

Patent Application Ser. No. 60/940,392 entitled "FULLY CONNECTED GENERALIZED STRICTLY NON-BLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

[0006] This application is related to and incorporates by reference in its entirety the U.S. application Ser. No. 12/601,275 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Nov. 22, 2009, and the U.S. Provisional Patent Application Ser. No. 60/940,394 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

BACKGROUND OF INVENTION

[0007] Multi-stage interconnection networks such as Benes networks and butterfly fat tree networks are widely useful in telecommunications, parallel and distributed computing. However VLSI layouts, known in the prior art, of these interconnection networks in an integrated circuit are inefficient and complicated.

[0008] Other multi-stage interconnection networks including butterfly fat tree networks, Banyan networks, Batcher-Banyan networks, Baseline networks, Delta networks, Omega networks and Flip networks have been widely studied particularly for self routing packet switching applications. Also Benes Networks with radix of two have been widely studied and it is known that Benes Networks of radix two are shown to be built with back to back baseline networks which are rearrangeably nonblocking for unicast connections.

[0009] The most commonly used VLSI layout in an integrated circuit is based on a two-dimensional grid model comprising only horizontal and vertical tracks. An intuitive interconnection network that utilizes two-dimensional grid model is 2D Mesh Network and its variations such as segmented mesh networks. Hence routing networks used in VLSI layouts are typically 2D mesh networks and its variations. However Mesh Networks require large scale cross points typically with a growth rate of $O(N^2)$ where N is the number of computing elements, ports, or logic elements depending on the application.

[0010] Multi-stage interconnection network with a growth rate of $O(N \times \log N)$ requires significantly small number of cross points. U.S. Pat. No. 6,185,220 entitled "Grid Layouts of Switching and Sorting Networks" granted to Muthukrishnan et al. describes a VLSI layout using existing VLSI grid model for Benes and Butterfly networks. U.S. Pat. No. 6,940,308 entitled "Interconnection Network for a Field Programmable Gate Array" granted to Wong describes a VLSI layout where switches belonging to lower stage of Benes Network are layed out close to the logic cells and switches belonging to higher stages are layed out towards the center of the layout.

[0011] Due to the inefficient and in some cases impractical VLSI layout of Benes and butterfly fat tree networks on a semiconductor chip, today mesh networks and segmented mesh networks are widely used in the practical applications such as field programmable gate arrays (FPGAs), programmable logic devices (PLDs), and parallel computing interconnects. The prior art VLSI layouts of Benes and butterfly fat tree networks and VLSI layouts of mesh networks and segmented mesh networks require large area to implement the switches on the chip, large number of wires, longer wires,

US 2012/0269190 A1

Oct. 25, 2012

2

with increased power consumption, increased latency of the signals which effect the maximum clock speed of operation. Some networks may not even be implemented practically on a chip due to the lack of efficient layouts.

SUMMARY OF INVENTION

[0012] When large scale sub-integrated circuit blocks with inlet and outlet links are layed out in an integrated circuit device in a two-dimensional grid arrangement, (for example in an FPGA where the sub-integrated circuit blocks are Lookup Tables) the most intuitive routing network is a network that uses horizontal and vertical links only (the most often used such a network is one of the variations of a 2D Mesh network). A direct embedding of a generalized multi-stage network on to a 2D Mesh network is neither simple nor efficient.

[0013] In accordance with the invention, VLSI layouts of generalized multi-stage and pyramid networks for broadcast, unicast and multicast connections are presented using only horizontal and vertical links with spacial locality exploitation. The VLSI layouts employ shuffle exchange links where outlet links of cross links from switches in a stage in one sub-integrated circuit block are connected to inlet links of switches in the succeeding stage in another sub-integrated circuit block so that said cross links are either vertical links or horizontal and vice versa. Furthermore the shuffle exchange links are employed between different sub-integrated circuit blocks so that spacially nearer sub-integrated circuit blocks are connected with shorter links compared to the shuffle exchange links between spacially farther sub-integrated circuit blocks. In one embodiment the sub-integrated circuit blocks are arranged in a hypercube arrangement in a two-dimensional plane. The VLSI layouts exploit the benefits of significantly lower cross points, lower signal latency, lower power and full connectivity with significantly fast compilation.

[0014] The VLSI layouts with spacial locality exploitation presented are applicable to generalized multi-stage and pyramid networks $V(N_1, N_2, d, s)$ & $V_p(N_1, N_2, d, s)$, generalized folded multi-stage and pyramid networks $V_{fold}(N_1, N_2, d, s)$ & $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat tree and butterfly fat pyramid networks $V_{bft}(N_1, N_2, d, s)$ & $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage and pyramid networks $V_{mlink}(N_1, N_2, d, s)$ & $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage and pyramid networks $V_{fold-mlink}(N_1, N_2, d, s)$ & $V_{fold-mlink-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree and butterfly fat pyramid networks $V_{mlink-bft}(N_1, N_2, d, s)$ & $V_{mlink-bfp}(N_1, N_2, d, s)$, generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$, and generalized cube connected cycles networks $V_{CCC}(N_1, N_2, d, s)$ for $s=1,2,3$ or any number in general. The embodiments of VLSI layouts are useful in wide target applications such as FPGAs, CPLDs, pSoCs, ASIC placement and route tools, networking applications, parallel & distributed computing, and reconfigurable computing.

BRIEF DESCRIPTION OF DRAWINGS

[0015] FIG. 1A is a diagram **100A** of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ having a variation of inverse Benes connection topology of nine stages with $N=32$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblock-

ing network for arbitrary fan-out multicast connections, in accordance with the invention.

[0016] FIG. 1B is a diagram **100B** of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network **100A** shown in FIG. 1A, having a variation of inverse Benes connection topology of five stages with $N=32$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0017] FIG. 1C is a diagram **100C** layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links belonging with in each block only.

[0018] FIG. 1D is a diagram **100D** layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links $ML(e1,i)$ for $i=[1, 64]$ and $ML(8,i)$ for $i=[1, 64]$.

[0019] FIG. 1E is a diagram **100E** layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links $ML(2,i)$ for $i=[1, 64]$ and $ML(7,i)$ for $i=[1, 64]$.

[0020] FIG. 1F is a diagram **100F** layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links $ML(3,i)$ for $i=[1, 64]$ and $ML(6,i)$ for $i=[1, 64]$.

[0021] FIG. 1G is a diagram **100G** layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 1B, in one embodiment, illustrating the connection links $ML(4,i)$ for $i=[1, 64]$ and $ML(5,i)$ for $i=[1, 64]$.

[0022] FIG. 1H is a diagram **100H** layout of a network $V_{fold-mlink}(N, d, s)$ where $N=128$, $d=2$, and $s=2$, in one embodiment, illustrating the connection links belonging with in each block only.

[0023] FIG. 1I is a diagram **100I** detailed connections of BLOCK **1_2** in the network layout **100C** in one embodiment, illustrating the connection links going in and coming out when the layout **100C** is implementing $V_{mlink}(N, d, s)$ or $V_{fold-mlink}(N, d, s)$.

[0024] FIG. 1J is a diagram **100J** detailed connections of BLOCK **1_2** in the network layout **100C** in one embodiment, illustrating the connection links going in and coming out when the layout **100C** is implementing $V_{mlink-bft}(N, d, s)$.

[0025] FIG. 1K is a diagram **100K** detailed connections of BLOCK **1_2** in the network layout **100C** in one embodiment, illustrating the connection links going in and coming out when the layout **100C** is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$.

[0026] FIG. 1K1 is a diagram **100M1** detailed connections of BLOCK **1_2** in the network layout **100C** in one embodiment, illustrating the connection links going in and coming out when the layout **100C** is implementing $V(N, d, s)$ or $V_{fold}(N, d, s)$ for $s=1$.

[0027] FIG. 1L is a diagram **100L** detailed connections of BLOCK **1_2** in the network layout **100C** in one embodiment, illustrating the connection links going in and coming out when the layout **100C** is implementing $V_{bft}(N, d, s)$.

[0028] FIG. 1L1 is a diagram **100L1** detailed connections of BLOCK **1_2** in the network layout **100C** in one embodiment, illustrating the connection links going in and coming out when the layout **100C** is implementing $V_{bft}(N, d, s)$ for $s=1$.

[0029] FIG. 2A is a diagram **200A** of an exemplary symmetrical multi-link multi-stage network $V_{fold-mlink}(N, d, s)$

US 2012/0269190 A1

Oct. 25, 2012

3

having inverse Benes connection topology of nine stages with $N=24$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0030] FIG. 2B is a diagram 200B of the equivalent symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ of the network 200A shown in FIG. 2A, having inverse Benes connection topology of five stages with $N=24$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0031] FIG. 2C is a diagram 200C layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links belonging with in each block only.

[0032] FIG. 2D is a diagram 200D layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links $ML(1,i)$ for $i=[1, 48]$ and $ML(8,i)$ for $i=[1, 48]$.

[0033] FIG. 2E is a diagram 200E layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links $ML(2,i)$ for $i=[1, 32]$ and $ML(7,i)$ for $i=[1, 32]$.

[0034] FIG. 2F is a diagram 200F layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links $ML(3,i)$ for $i=[1, 64]$ and $ML(6,i)$ for $i=[1, 64]$.

[0035] FIG. 2G is a diagram 200G layout of the network $V_{fold-mlink}(N, d, s)$ shown in FIG. 2B, in one embodiment, illustrating the connection links $ML(4,i)$ for $i=[1, 64]$ and $ML(5,i)$ for $i=[1, 64]$.

[0036] FIG. 3A is a diagram 300A layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=512$, $d=2$ and $s=2$, in one embodiment, illustrating the provisioning of 2's BW.

[0037] FIG. 3B is a diagram 300B layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=512$, $d=2$ and $s=2$, in one embodiment, illustrating the provisioning of 4's BW.

[0038] FIG. 3C is a diagram 300C layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=512$, $d=2$ and $s=2$, in one embodiment, illustrating the provisioning of 8's BW with nearest neighbor connectivity first.

[0039] FIG. 3D is a diagram 300D layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=512$, $d=2$ and $s=2$, in one embodiment, illustrating the provisioning of 8's BW with nearest neighbor connectivity recursively.

[0040] FIG. 4A is a diagram 400A layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=512$, $d=2$ and $s=2$, in one embodiment, illustrating the provisioning of 2's BW in first stage.

[0041] FIG. 4B is a diagram 400B layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=512$, $d=2$ and $s=2$, in one embodiment, illustrating the remaining nearest neighbor connectivity in the second stage by provisioning 4's BW, 8's BW etc.

[0042] FIG. 4C is a diagram 400C layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=512$, $d=2$ and $s=2$, in one embodiment, illustrating the third stage, by provisioning 4's and 8's BW.

[0043] FIG. 5 is a diagram 500 layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=512$, $d=2$ and $s=2$, in

one embodiment, illustrating the provisioning of 8's BW and 16's BW in Partial & Tapered Connectivity (Bandwidth) in a stage.

[0044] FIG. 6 is a diagram 600 layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=2048$, $d=2$ and $s=2$, in one embodiment, illustrating the provisioning of 8's BW, 16's BW and 32's BW in Partial & Tapered Connectivity (Bandwidth) in a stage.

[0045] FIG. 7 is a diagram 700 layout of the topmost row of the network $V_{fold-mlink}(N, d, s)$ with $N=2048$, $d=2$ and $s=2$, in one embodiment, illustrating the provisioning of 8's BW, 16's BW and 32's BW in Partial & Tapered Connectivity (Bandwidth) in a stage with equal length wires.

[0046] FIG. 8A is a diagram 800A of an exemplary symmetrical multi-link multi-stage pyramid network $V_{mlink-p}(N, d, s)$ having inverse Benes connection topology of nine stages with $N=32$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0047] FIG. 8B is a diagram 800B of the equivalent symmetrical folded multi-link multi-stage pyramid network $V_{fold-mlink-p}(N, d, s)$ of the network 800A shown in FIG. 8A, having inverse Benes connection topology of five stages with $N=32$, $d=2$ and $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0048] FIG. 8C is a diagram 800C layout of the network $V_{fold-mlink-p}(N, d, s)$ shown in FIG. 8B, in one embodiment, illustrating the connection links belonging with in each block only.

[0049] FIG. 8D is a diagram 800D layout of the network $V_{fold-mlink-p}(N, d, s)$ shown in FIG. 8B, in one embodiment, illustrating the connection links $ML(1,i)$ for $i=[1, 64]$ and $ML(8,i)$ for $i=[1, 64]$.

[0050] FIG. 8E is a diagram 800E layout of the network $V_{fold-mlink-p}(N, d, s)$ shown in FIG. 8B, in one embodiment, illustrating the connection links $ML(2,i)$ for $i=[1, 64]$ and $ML(7,i)$ for $i=[1, 64]$.

[0051] FIG. 8F is a diagram 800F layout of the network $V_{fold-mlink-p}(N, d, s)$ shown in FIG. 8B, in one embodiment, illustrating the connection links $ML(3,i)$ for $i=[1, 64]$ and $ML(6,i)$ for $i=[1, 64]$.

[0052] FIG. 8G is a diagram 800G layout of the network $V_{fold-mlink-p}(N, d, s)$ shown in FIG. 8B, in one embodiment, illustrating the connection links $ML(4,i)$ for $i=[1, 64]$ and $ML(5,i)$ for $i=[1, 64]$.

[0053] FIG. 8H is a diagram 800H layout of a network $V_{fold-mlink-p}(N, d, s)$ where $N=128$, $d=2$, and $s=2$, in one embodiment, illustrating the connection links belonging with in each block only.

[0054] FIG. 8I is a diagram 800I detailed connections of BLOCK 1_2 in the network layout 800C in one embodiment, illustrating the connection links going in and coming out when the layout 800C is implementing $V_{mlink-p}(N, d, s)$ or $V_{fold-mlink-p}(N, d, s)$.

[0055] FIG. 8J is a diagram 800J detailed connections of BLOCK 1_2 in the network layout 800C in one embodiment, illustrating the connection links going in and coming out when the layout 800C is implementing $V_{mlink-bfp}(N, d, s)$.

[0056] FIG. 8K is a diagram 800K detailed connections of BLOCK 1_2 in the network layout 800C in one embodiment,

US 2012/0269190 A1

Oct. 25, 2012

4

illustrating the connection links going in and coming out when the layout **800C** is implementing $V_p(N, d, s)$ or $V_{fold-p}(N, d, s)$.

[0057] FIG. **8K1** is a diagram **800M1** detailed connections of BLOCK **1_2** in the network layout **800C** in one embodiment, illustrating the connection links going in and coming out when the layout **800C** is implementing $V_p(N, d, s)$ or $V_{fold-p}(N, d, s)$ for $s=1$.

[0058] FIG. **8L** is a diagram **800L** detailed connections of BLOCK **1_2** in the network layout **800C** in one embodiment, illustrating the connection links going in and coming out when the layout **800C** is implementing $V_{bfp}(N, d, s)$.

[0059] FIG. **8L1** is a diagram **800L1** detailed connections of BLOCK **1_2** in the network layout **800C** in one embodiment, illustrating the connection links going in and coming out when the layout **800C** is implementing $V_{bfp}(N, d, s)$ for $s=1$.

[0060] FIG. **9A** is high-level flowchart of a scheduling method **900** according to the invention, used to set up the multicast connections in the generalized multi-stage pyramid network and the generalized multi-link multi-stage pyramid network disclosed in this invention.

[0061] FIG. **10A** is high-level flowchart of a scheduling method **1000** according to the invention, used to set up the multicast connections in the generalized butterfly fat pyramid network and the generalized multi-link butterfly fat pyramid network disclosed in this invention.

[0062] FIG. **11A1** is a diagram **1100A1** of an exemplary prior art implementation of a two by two switch; FIG. **11A2** is a diagram **1100A2** for programmable integrated circuit prior art implementation of the diagram **1100A1** of FIG. **11A1**; FIG. **11A3** is a diagram **1100A3** for one-time programmable integrated circuit prior art implementation of the diagram **1100A1** of FIG. **11A1**; FIG. **11A4** is a diagram **1100A4** for integrated circuit placement and route implementation of the diagram **1100A1** of FIG. **11A1**.

DETAILED DESCRIPTION OF THE INVENTION

[0063] The present invention is concerned with the VLSI layouts of arbitrarily large switching networks for broadcast, unicast and multicast connections. Particularly switching networks considered in the current invention include: generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$, generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$, and generalized cube connected cycles networks $V_{ccc}(N_1, N_2, d, s)$ for $s=1,2,3$ or any number in general.

[0064] Efficient VLSI layout of networks on a semiconductor chip are very important and greatly influence many important design parameters such as the area taken up by the network on the chip, total number of wires, length of the wires, latency of the signals, capacitance and hence the maximum clock speed of operation. Some networks may not even be implemented practically on a chip due to the lack of efficient layouts. The different varieties of multi-stage networks described above have not been implemented previously on the semiconductor chips efficiently. For example in Field Programmable Gate Array (FPGA) designs, multi-stage networks described in the current invention have not been successfully implemented primarily due to the lack of efficient

VLSI layouts. Current commercial FPGA products such as Xilinx Vertex, Altera's Stratix implement island-style architecture using mesh and segmented mesh routing interconnects using either full crossbars or sparse crossbars. These routing interconnects consume large silicon area for cross-points, long wires, large signal propagation delay and hence consume lot of power.

[0065] The current invention discloses the VLSI layouts of numerous types of multi-stage and pyramid networks which are very efficient and exploit spacial locality in the connectivity. Moreover they can be embedded on to mesh and segmented mesh routing interconnects of current commercial FPGA products. The VLSI layouts disclosed in the current invention are applicable to including the numerous generalized multi-stage networks disclosed in the following patent applications:

[0066] 1) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-stage networks $V(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the U.S. application Ser. No. 12/530,207 that is incorporated by reference above.

[0067] 2) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the U.S. application Ser. No. 12/601,273 that is incorporated by reference above.

[0068] 3) Rearrangeably nonblocking for arbitrary fan-out multicast and unicast, and strictly nonblocking for unicast for generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the U.S. application Ser. No. 12/601,274 that is incorporated by reference above.

[0069] 4) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the U.S. application Ser. No. 12/601,273 that is incorporated by reference above.

[0070] 5) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the U.S. application Ser. No. 12/601,274 that is incorporated by reference above.

[0071] 6) Strictly nonblocking for arbitrary fan-out multicast and unicast for generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the U.S. application Ser. No. 12/601,274 that is incorporated by reference above.

[0072] 7) VLSI layouts of numerous types of multi-stage networks are described in the U.S. application Ser. No. 12/601,275 entitled "VLSI LAYOUTS OF FULLY CONNECTED NETWORKS" that is incorporated by reference above.

[0073] In addition the layouts of the current invention are also applicable to generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyra-

US 2012/0269190 A1

Oct. 25, 2012

5

mid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage pyramid networks $V_{fold-mlink-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$, generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ and generalized cube connected cycles networks $V_{ccc}(N_1, N_2, d, s)$ for $s=1,2,3$ or any number in general.

Symmetric RNB Generalized Multi-Link Multi-Stage Network $V_{mlink}(N_1, N_2, d, s)$, Connection Topology: Nearest Neighbor Connectivity and with Full Bandwidth:

[0074] Referring to diagram 100A in FIG. 1A, in one embodiment, an exemplary generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages of one hundred and forty four switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, 150, 160, 170, 180 and 190 is shown where input stage 110 consists of sixteen, two by four switches IS1-IS16 and output stage 120 consists of sixteen, four by two switches OS1-OS16. And all the middle stages namely the middle stage 130 consists of sixteen, four by four switches MS(1,1)-MS(1,16), middle stage 140 consists of sixteen, four by four switches MS(2,1)-MS(2,16), middle stage 150 consists of sixteen, four by four switches MS(3,1)-MS(3,16), middle stage 160 consists of sixteen, four by four switches MS(4,1)-MS(4,16), middle stage 170 consists of sixteen, four by four switches MS(5,1)-MS(5,16), middle stage 180 consists of sixteen, four by four switches MS(6,1)-MS(6,16), and middle stage 190 consists of sixteen, four by four switches MS(7,1)-MS(7,16).

[0075] As disclosed in U.S. Provisional Patent Application Ser. No. 60/940,389 that is incorporated by reference above, such a network can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections.

[0076] In one embodiment of this network each of the input switches IS1-IS16 and output switches OS1-OS16 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch IS1-IS16 can be denoted in general with the notation $d*2d$ and each output switch OS1-OS16 can be denoted in general with the notation $2d*d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d*2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-stage network can be represented with the notation $V_{mlink}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL32), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0077] Each of the N/d input switches IS1-IS16 are connected to exactly d switches in middle stage 130 through two links each for a total of $2xd$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the middle links ML(1,1), ML(1,2), and also connected to middle switch MS(1,2) through the middle links ML(1,3) and ML(1,

4)). The middle links which connect switches in the same row in two successive middle stages are called hereinafter straight middle links; and the middle links which connect switches in different rows in two successive middle stages are called hereinafter cross middle links. For example, the middle links ML(1,1) and ML(1,2) connect input switch IS1 and middle switch MS(1,1), so middle links ML(1,1) and ML(1,2) are straight middle links; whereas the middle links ML(1,3) and ML(1,4) connect input switch IS1 and middle switch MS(1,2), since input switch IS1 and middle switch MS(1,2) belong to two different rows in diagram 100A of FIG. 1A, middle links ML(1,3) and ML(1,4) are cross middle links.

[0078] Each of the N/d middle switches MS(1,1)-MS(1,16) in the middle stage 130 are connected from exactly d input switches through two links each for a total of $2xd$ links (for example the middle links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the middle links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to exactly d switches in middle stage 140 through two links each for a total of $2xd$ links (for example the middle links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the middle links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)).

[0079] Each of the N/d middle switches MS(2,1)-MS(2,16) in the middle stage 140 are connected from exactly d middle switches in middle stage 130 through two links each for a total of $2xd$ links (for example the middle links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from input switch MS(1,1), and the middle links ML(1,11) and ML(1,12) are connected to the middle switch MS(2,1) from input switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through two links each for a total of $2xd$ links (for example the middle links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the middle links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,6)).

[0080] Applicant notes that the topology of connections between middle switches MS(2,1)-MS(2,16) in the middle stage 140 and middle switches MS(3,1)-MS(3,16) in the middle stage 150 is not the typical inverse Benes topology but the connectivity of the generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ 100A shown in FIG. 1A is effectively the same, or alternatively the network 100A shown in FIG. 1A is topologically equivalent to the network with inverse Benes network topology. However as will be described later in layouts of FIG. 1C-FIG. 1G, the length of the connection from a given inlet link to its destination outlet links may consist of different route resulting in different latency and different power dissipation for a given multicast or unicast assignment. As will be described later in the layouts of FIG. 1C-FIG. 1G, the connection topology of middle links between middle stages 140 and 150 is in such a way that nearest neighbor blocks are connected directly and then the rest of the blocks are connected in inverse Benes topology.

[0081] Each of the N/d middle switches MS(3,1)-MS(3,16) in the middle stage 150 are connected from exactly d middle switches in middle stage 140 through two links each for a total of $2xd$ links (for example the middle links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from input switch MS(2,1), and the middle links ML(2,23) and ML(2,24) are connected to the middle switch MS(3,1) from

US 2012/0269190 A1

Oct. 25, 2012

6

input switch MS(2,6)) and also are connected to exactly d switches in middle stage **160** through two links each for a total of $2 \times d$ links (for example the middle links ML(4,1) and ML(4,2) are connected from middle switch MS(3,1) to middle switch MS(4,1), and the middle links ML(4,3) and ML(4,4) are connected from middle switch MS(3,1) to middle switch MS(4,11)).

[0082] Applicant notes that the topology of connections between middle switches MS(3,1)-MS(3,16) in the middle stage **150** and middle switches MS(4,1)-MS(4,16) in the middle stage **160** is not the typical inverse Benes topology but the connectivity of the generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ **100A** shown in FIG. 1A is effectively the same, or alternatively the network **100A** shown in FIG. 1A is topologically equivalent to the network with inverse Benes network topology. However as will be described later in layouts of FIG. 1C-FIG. 1G, the length of the connection from a given inlet link to its destination outlet links may consist of different route resulting in different latency and different power dissipation for a given multicast or unicast assignment. As will be described later in the layouts of FIG. 1C-FIG. 1G, the connection topology of middle links between middle stages **150** and **160** is in such a way that nearest neighbor blocks are connected directly and then the rest of the blocks are connected in inverse Benes topology.

[0083] Each of the N/d middle switches MS(4,1)-MS(4,16) in the middle stage **160** are connected from exactly d middle switches in middle stage **150** through two links each for a total of $2 \times d$ links (for example the middle links ML(4,1) and ML(4,2) are connected to the middle switch MS(4,1) from input switch MS(3,1), and the middle links ML(4,3) and ML(4,4) are connected to the middle switch MS(4,1) from input switch MS(3,11)) and also are connected to exactly d switches in middle stage **170** through two links each for a total of $2 \times d$ links (for example the middle links ML(5,1) and ML(5,2) are connected from middle switch MS(4,1) to middle switch MS(5,1), and the middle links ML(5,3) and ML(5,4) are connected from middle switch MS(4,1) to middle switch MS(5,11)).

[0084] Applicant notes that the topology of connections between middle switches MS(4,1)-MS(4,16) in the middle stage **160** and middle switches MS(5,1)-MS(5,16) in the middle stage **170** is not the typical inverse Benes topology but the connectivity of the generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ **100A** shown in FIG. 1A is effectively the same or alternatively the network **100A** shown in FIG. 1A is topologically equivalent to the network with inverse Benes network topology. However as will be described later in layouts of FIG. 1C-FIG. 1G, the length of the connection from a given inlet link to its destination outlet links may consist of different route resulting in different latency and different power dissipation for a given multicast or unicast assignment. As will be described later in the layouts of FIG. 1C-FIG. 1G, the connection topology of middle links between middle stages **160** and **170** is in such a way that nearest neighbor blocks are connected directly and then the rest of the blocks are connected in inverse Benes topology.

[0085] Each of the N/d middle switches MS(5,1)-MS(5,16) in the middle stage **170** are connected from exactly d middle switches in middle stage **160** through two links each for a total of $2 \times d$ links (for example the middle links ML(5,1) and ML(5,2) are connected to the middle switch MS(5,1) from input switch MS(4,1), and the middle links ML(5,3) and ML(5,4) are connected to the middle switch MS(5,1) from

input switch MS(4,11)) and also are connected to exactly d switches in middle stage **180** through two links each for a total of $2 \times d$ links (for example the middle links ML(6,1) and ML(6,2) are connected from middle switch MS(5,1) to middle switch MS(6,1), and the middle links ML(6,3) and ML(6,4) are connected from middle switch MS(5,1) to middle switch MS(6,6)).

[0086] Applicant notes that the topology of connections between middle switches MS(5,1)-MS(5,16) in the middle stage **170** and middle switches MS(6,1)-MS(6,16) in the middle stage **180** is not the typical inverse Benes topology but the connectivity of the generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ **100A** shown in FIG. 1A is effectively the same or alternatively the network **100A** shown in FIG. 1A is topologically equivalent to the network with inverse Benes network topology. However as will be described later in layouts of FIG. 1C-FIG. 1G, the length of the connection from a given inlet link to its destination outlet links may consist of different route resulting in different latency and different power dissipation for a given multicast or unicast assignment. As will be described later in the layouts of FIG. 1C-FIG. 1G, the connection topology of middle links between middle stages **170** and **180** is in such a way that nearest neighbor blocks are connected directly and then the rest of the blocks are connected in inverse Benes topology.

[0087] Each of the N/d middle switches MS(6,1)-MS(6,16) in the middle stage **180** are connected from exactly d middle switches in middle stage **170** through two links each for a total of $2 \times d$ links (for example the middle links ML(6,1) and ML(6,2) are connected to the middle switch MS(6,1) from input switch MS(5,1), and the middle links ML(6,23) and ML(6,24) are connected to the middle switch MS(6,1) from input switch MS(5,6)) and also are connected to exactly d switches in middle stage **190** through two links each for a total of $2 \times d$ links (for example the middle links ML(7,1) and ML(7,2) are connected from middle switch MS(6,1) to middle switch MS(7,1), and the middle links ML(7,3) and ML(7,4) are connected from middle switch MS(6,1) to middle switch MS(7,3)).

[0088] Each of the N/d middle switches MS(7,1)-MS(7,16) in the middle stage **190** are connected from exactly d middle switches in middle stage **180** through two links each for a total of $2 \times d$ links (for example the middle links ML(7,1) and ML(7,2) are connected to the middle switch MS(7,1) from input switch MS(6,1), and the middle links ML(7,11) and ML(7,12) are connected to the middle switch MS(7,1) from input switch MS(6,3)) and also are connected to exactly d switches in middle stage **120** through two links each for a total of $2 \times d$ links (for example the middle links ML(8,1) and ML(8,2) are connected from middle switch MS(7,1) to middle switch MS(8,1), and the middle links ML(8,3) and ML(8,4) are connected from middle switch MS(7,1) to middle switch OS2).

[0089] Each of the N/d middle switches OS1-OS16 in the middle stage **120** are connected from exactly d middle switches in middle stage **190** through two links each for a total of $2 \times d$ links (for example the middle links ML(8,1) and ML(8,2) are connected to the output switch OS1 from input switch MS(7,1), and the middle links ML(8,7) and ML(8,8) are connected to the output switch OS1 from input switch MS(7,2)).

[0090] Finally the connection topology of the network **100A** shown in FIG. 1A is logically similar to back to back inverse Benes connection topology with nearest neighbor

US 2012/0269190 A1

Oct. 25, 2012

7

connections between all the middle stages starting from middle stage **140** and middle stage **180**.

[0091] Referring to diagram **100B** in FIG. **1B**, is a folded version of the multi-link multi-stage network **100A** shown in FIG. **1A**. The network **100B** in FIG. **1B** shows input stage **110** and output stage **120** are placed together. That is input switch **IS1** and output switch **OS1** are placed together, input switch **IS2** and output switch **OS2** are placed together, and similarly input switch **IS16** and output switch **OS16** are placed together. All the right going links {i.e., inlet links **IL1-IL32** and middle links **ML(1,1)-ML(1, 64)**} correspond to input switches **IS1-IS16**, and all the left going links {i.e., middle links **ML(8,1)-ML(8,64)** and outlet links **OL1-OL32**} correspond to output switches **OS1-OS16**.

[0092] Middle stage **130** and middle stage **190** are placed together. That is middle switches **MS(1,1)** and **MS(7,1)** are placed together, middle switches **MS(1,2)** and **MS(7,2)** are placed together, and similarly middle switches **MS(1,16)** and **MS(7,16)** are placed together. All the right going middle links {i.e., middle links **ML(1,1)-ML(1, 64)** and middle links **ML(2,1)-ML(2,64)**} correspond to middle switches **MS(1, 1)-MS(1,16)**, and all the left going middle links {i.e., middle links **ML(7,1)-ML(7,64)** and middle links **ML(8,1)** and **ML(8,64)**} correspond to middle switches **MS(7,1)-MS(7, 16)**.

[0093] Middle stage **140** and middle stage **180** are placed together. That is middle switches **MS(2,1)** and **MS(6,1)** are placed together, middle switches **MS(2,2)** and **MS(6,2)** are placed together, and similarly middle switches **MS(2,16)** and **MS(6,16)** are placed together. All the right going middle links {i.e., middle links **ML(2,1)-ML(2,64)** and middle links **ML(3,1)-ML(3,64)**} correspond to middle switches **MS(2, 1)-MS(2,16)**, and all the left going middle links {i.e., middle links **ML(6,1)-ML(6,64)** and middle links **ML(7,1)** and **ML(7,64)**} correspond to middle switches **MS(6,1)-MS(6, 16)**.

[0094] Middle stage **150** and middle stage **170** are placed together. That is middle switches **MS(3,1)** and **MS(5,1)** are placed together, middle switches **MS(3,2)** and **MS(5,2)** are placed together, and similarly middle switches **MS(3,16)** and **MS(5,16)** are placed together. All the right going middle links {i.e., middle links **ML(3,1)-ML(3,64)** and middle links **ML(4,1)-ML(4,64)**} correspond to middle switches **MS(3, 1)-MS(3,16)**, and all the left going middle links {i.e., middle links **ML(5,1)-ML(5,64)** and middle links **ML(6,1)** and **ML(6,64)**} correspond to middle switches **MS(5,1)-MS(5, 16)**.

[0095] Middle stage **160** is placed alone. All the right going middle links are the middle links **ML(4,1)-ML(4,64)** and all the left going middle links are middle links **ML(5,1)-ML(5, 64)**.

[0096] Just the same way as the connection topology of the network **100A** shown in FIG. **1A**, the connection topology of the network **100B** shown in FIG. **1B** is the folded version and logically similar to back to back inverse Benes connection topology with nearest neighbor connections between all the middle stages starting from middle stage **140** and middle stage **180**.

[0097] In one embodiment, in the network **100B** of FIG. **1B**, the switches that are placed together are implemented as separate switches then the network **100B** is the generalized folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,389

that is incorporated by reference above. That is the switches that are placed together in input stage **110** and output stage **120** are implemented as a two by four switch and a four by two switch respectively. For example the input switch **IS1** and output switch **OS1** are placed together; so input switch **IS1** is implemented as two by four switch with the inlet links **IL1** and **IL2** being the inputs of the input switch **IS1** and middle links **ML(1,1)-ML(1,4)** being the outputs of the input switch **IS1**; and output switch **OS1** is implemented as four by two switch with the middle links **ML(8,1)**, **ML(8,2)**, **ML(8,7)** and **ML(8,8)** being the inputs of the output switch **OS1** and outlet links **OL1-OL2** being the outputs of the output switch **OS1**. Similarly in this embodiment of network **100B** all the switches that are placed together in each middle stage are implemented as separate switches.

Modified-Hypercube Topology Layout Scheme:

[0098] Referring to layout **100C** of FIG. **1C**, in one embodiment, there are sixteen blocks namely Block **1_2**, Block **3_4**, Block **5_6**, Block **7_8**, Block **9_10**, Block **11_12**, Block **13_14**, Block **15_16**, Block **17_18**, Block **19_20**, Block **21_22**, Block **23_24**, Block **25_26**, Block **27_28**, Block **29_30**, and Block **31_32**. Each block implements all the switches in one row of the network **100B** of FIG. **1B**, one of the key aspects of the current invention. For example Block **1_2** implements the input switch **IS1**, output Switch **OS1**, middle switch **MS(1,1)**, middle switch **MS(7,1)**, middle switch **MS(2,1)**, middle switch **MS(6,1)**, middle switch **MS(3,1)**, middle switch **MS(5,1)**, and middle switch **MS(4, 1)**. For the simplification of illustration, Input switch **IS1** and output switch **OS1** together are denoted as switch **1**; Middle switch **MS(1,1)** and middle switch **MS(7,1)** together are denoted by switch **2**; Middle switch **MS(2,1)** and middle switch **MS(6,1)** together are denoted by switch **3**; Middle switch **MS(3,1)** and middle switch **MS(5,1)** together are denoted by switch **4**; Middle switch **MS(4,1)** is denoted by switch **5**.

[0099] All the straight middle links are illustrated in layout **100C** of FIG. **1C**. For example in Block **1_2**, inlet links **IL1-IL2**, outlet links **OL1-OL2**, middle link **ML(1,1)**, middle link **ML(1,2)**, middle link **ML(8,1)**, middle link **ML(8,2)**, middle link **ML(2,1)**, middle link **ML(2,2)**, middle link **ML(7,1)**, middle link **ML(7,2)**, middle link **ML(3,1)**, middle link **ML(3,2)**, middle link **ML(6,1)**, middle link **ML(6,2)**, middle link **ML(4,1)**, middle link **ML(4,2)**, middle link **ML(5,1)** and middle link **ML(5,2)** are illustrated in layout **100C** of FIG. **1C**.

[0100] Even though it is not illustrated in layout **100C** of FIG. **1C**, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit depending on the applications in different embodiments. There are four quadrants in the layout **100C** of FIG. **1C** namely top-left, bottom-left, top-right and bottom-right quadrants. Top-left quadrant implements Block **1_2**, Block **3_4**, Block **5_6**, and Block **7_8**. Bottom-left quadrant implements Block **9_10**, Block **11_12**, Block **13_14**, and Block **15_16**. Top-right quadrant implements Block **17_18**, Block **19_20**, Block **21_22**, and Block **23_24**. Bottom-right quadrant implements Block **25_26**, Block **27_28**, Block **29_30**, and Block **31_32**. There are two halves in layout **100C** of FIG. **1C** namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants. Right-half consists of top-right and bottom-right quadrants.

US 2012/0269190 A1

Oct. 25, 2012

8

[0101] Recursively in each quadrant there are four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements Block 1_2. Bottom-left sub-quadrant of top-left quadrant implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 5_6. Finally bottom-right sub-quadrant of top-left quadrant implements Block 7_8. Similarly there are two sub-halves in each quadrant. For example in top-left quadrant there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left quadrant implements Block 1_2 and Block 3_4. Right-sub-half of top-left quadrant implements Block 5_6 and Block 7_8. Finally applicant notes that in each quadrant or half the blocks are arranged as a general binary hypercube. Recursively in larger multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2>32$, the layout in this embodiment in accordance with the current invention, will be such that the super-quadrants will also be arranged in d-ary hypercube manner. (In the embodiment of the layout 100C of FIG. 1C, it is binary hypercube manner since $d=2$, in the network $V_{fold-mlink}(N_1, N_2, d, s)$ 100B of FIG. 1B).

[0102] Layout 100D of FIG. 1D illustrates the inter-block links between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 100D of FIG. 1D can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

[0103] The bandwidth provided between two physically adjacent blocks in the same column or same row, when a switch in the first block is connected to a switch in the second block through the corresponding inter-block links and also a second switch in the second block is connected to a second switch in the first block through the corresponding inter-block links, is hereinafter called 2's bandwidth or 2's BW. The bandwidth offered between two diagonal blocks is also 2's BW when the corresponding row and columns provide 2's BW. For example the bandwidth provided between Block 1_2 and Block 3_4 of layout 100D of FIG. 1D is 2's BW because inter-block links between switch 1 of Block 1_2 and switch 2 of Block 3_4 are connected and also inter-block links between switch 2 of Block 1_2 and switch 1 of Block 3_4 are connected.

[0104] In general the bandwidth offered within a quadrant of the layout formed by two nearest neighboring blocks on each of the four sides is 2's BW. For example in layout 100C of FIG. 1C the bandwidth offered in top-left quadrant is 2's BW. Similarly the bandwidth offered within each of the other three quadrants bottom-left, top-right and bottom-right quadrants is 2' BW. Alternatively the bandwidth offered with in a square of blocks with the sides of the square consisting of two neighboring blocks is 2's BW. This definition can be generalized so that the bandwidth offered within a square of blocks

with the sides consisting of "x" number of blocks, when $x=2^y$ where y is an integer, is hereinafter x's BW. Hence the bandwidth offered between four neighboring quadrants is 4's BW. For example the bandwidth offered between top-left quadrant, bottom-left quadrant, top-right quadrant and bottom-right quadrant is 4's BW as will be described later. It must be noted that the 4's BW is the bandwidth offered between the four quadrants in a square of four quadrants and it is not the bandwidth offered with in each quadrant.

[0105] Layout 100E of FIG. 1E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 5_6. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 5_6. Applicant notes that the inter-block links illustrated in layout 100E of FIG. 1E can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

[0106] The bandwidth provided between Block 1_2 and Block 5_6 of layout 100E of FIG. 1E is 2's BW because inter-block links between switch 2 of Block 1_2 and switch 3 of Block 5_6 are connected and also inter-block links between switch 3 of Block 1_2 and switch 2 of Block 5_6 are connected. Similarly the bandwidth provided between Block 1_2 and Block 7_8 is also 2's BW since corresponding rows (formed by Block 1_2 and Block 5_6; and by Block 3_4 and Block 7_8) and columns (formed by Block 1_2 and Block 3_4; and by Block 5_6 and Block 7_8) offer 2's BW. Similarly the bandwidth offered between Block 3_4 and Block 5_6 is 2's BW.

[0107] Layout 100F of FIG. 1F illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,23), and ML(6,24) are connected between switch 3 of Block 1_2 and switch 4 of Block 11_12. Similarly middle links ML(3,23), ML(3,24), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 11_12. Applicant notes that the inter-block links illustrated in layout 100F of FIG. 1F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,24) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,24) are implemented as a time division multiplexed single track).

[0108] Applicant notes that the topology of inter-block links between switches 3 and 4 of each block of layout 100F of FIG. 1F is not the typical inverse Benes Network topology. In layout 100F first the switches 3 and 4 of nearest neighbor blocks are connected and then the rest of the blocks are connected in inverse Benes Network topology. For example since Block 3_4 and Block 9_10 are nearest neighbors in the leftmost column of layout 100F the corresponding links from switches 3 and 4 are connected together first. Then the remaining blocks in each column are connected in inverse Benes topology. For example in layout 100F since the remain-

US 2012/0269190 A1

Oct. 25, 2012

9

ing block in the leftmost column of top-left quadrant is Block 1_2 and the remaining block in the leftmost column of bottom-left quadrant is Block 11_12 the inter-block links between their corresponding switches 3 and 4 are connected together. Similarly in all the columns, the inter-block links between switches 3 and 4 are connected.

[0109] The bandwidth offered in layout 100F of FIG. 1F is 4's BW, since the bandwidth offered with in a square of blocks with the sides of the square consisting of four neighboring blocks is 4's BW. It must be noted that the bandwidth offered between top-left quadrant and bottom-left quadrant is 4's BW. That is inter-block links of a switch in each one of the blocks in top-left quadrant are connected to a switch in any one of the blocks in bottom-left quadrant and vice versa. Similarly the bandwidth offered between top-right quadrant and bottom-right quadrant is 4's BW. For example the bandwidth provided between Block 1_2 and Block 11_12 of layout 100F of FIG. 1F is 4's BW because inter-block links between switch 3 of Block 1_2 and switch 4 of Block 11_12 are connected and also inter-block links between switch 4 of Block 1_2 and switch 3 of Block 11_12 are connected. Similarly the bandwidth provided between Block 3_4 and Block 9_10 of layout 100F of FIG. 1F is 4's BW, even though they are physically nearest neighbors. It must be noted that the 4's BW is the bandwidth offered between the four quadrants in a square of four quadrants and it is not the bandwidth offered with in each quadrant.

[0110] Layout 100G of FIG. 1G illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,43), and ML(5,44) are connected between switch 4 of Block 1_2 and switch 5 of Block 21_22. Similarly middle links ML(4,43), ML(4,44), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 21_22. Applicant notes that the inter-block links illustrated in layout 100G of FIG. 1G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,44) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(4,4) and ML(5,44) are implemented as a time division multiplexed single track).

[0111] Applicant notes that the topology of inter-block links between switches 4 and 5 of each block of layout 100G of FIG. 1G is not the typical inverse Benes Network topology. In layout 100G first the switches 4 and 5 of nearest neighbor blocks are connected and then the rest of the blocks are connected in inverse Benes Network topology. For example since Block 5_6 and Block 17_18 are nearest neighbors in the topmost row of layout 100G the corresponding links from switches 4 and 5 are connected together first. Then the remaining blocks in each row are connected in inverse Benes topology. For example in layout 100G since the remaining block in the topmost row of top-left quadrant is Block 1_2 and the remaining block in the topmost row of top-right quadrant is Block 21_22 the inter-block links between their corresponding switches 4 and 5 are connected together. Similarly in all the rows, the inter-block links between switches 4 and 5 are connected.

[0112] The bandwidth offered in layout 100G of FIG. 1G is 4's BW, since the bandwidth offered with in a square of blocks with the sides of the square consisting of four neighboring blocks is 4's BW. It must be noted that the bandwidth offered

between top-left quadrant and top-right quadrant is 4's BW. That is inter-block links of a switch in each one of the blocks in top-left quadrant are connected to a switch in any one of the blocks in top-right quadrant and vice versa. Similarly the bandwidth offered between bottom-left quadrant and bottom-right quadrant is 4's BW. For example the bandwidth provided between Block 1_2 and Block 21_22 of layout 100G of FIG. 1G is 4's BW because inter-block links between switch 4 of Block 1_2 and switch 5 of Block 21_22 are connected and also inter-block links between switch 5 of Block 1_2 and switch 4 of Block 21_22 are connected. Similarly the bandwidth provided between Block 5_6 and Block 17_18 of layout 100G of FIG. 1G is 4's BW, even though they are physically nearest neighbors. Just the same way 2's BW is provided between two diagonal blocks, the bandwidth offered between two diagonal quadrants is also 4's BW that is when the corresponding row and columns provide 4's BW.

[0113] The complete layout for the network 100B of FIG. 1B is given by combining the links in layout diagrams of 100C, 100D, 100E, 100F, and 100G. Applicant notes that in the layout 100C of FIG. 1C, the inter-block links between switch 1 and switch 2 of corresponding blocks are vertical tracks as shown in layout 100D of FIG. 1D; the inter-block links between switch 2 and switch 3 of corresponding blocks are horizontal tracks as shown in layout 100E of FIG. 1E; the inter-block links between switch 3 and switch 4 of corresponding blocks are vertical tracks as shown in layout 100F of FIG. 1F; and finally the inter-block links between switch 4 and switch 5 of corresponding blocks are horizontal tracks as shown in layout 100G of FIG. 1G. The pattern is alternate vertical tracks and horizontal tracks. It continues recursively for larger networks of $N > 32$ as will be illustrated later.

[0114] Some of the key aspects of the current invention are discussed. 1) All the switches in one row of the multi-stage network 100B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks or vertical tracks; 3) Since all the inter-block links are either horizontal or vertical tracks, all the inter-block links can be mapped on to island-style architectures in current commercial FPGAs; 4) The length of the wires in a given stage are not equal, for example the inter-block links between switches 3 and 4 of the nearest neighbor blocks Block 3_4 and Block 9_10 are smaller in length than the inter-block links between switches 3 and 4 of the blocks Block 1_2 and Block 11_12.

[0115] In accordance with the current invention, the layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ the sub-quadrants, quadrants, and super-quadrants are arranged in d-ary hypercube manner and also the inter-blocks are accordingly connected in d-ary hypercube topology. Even though all the embodiments in the current invention are illustrated for $N_1 = N_2$, the embodiments can be extended for $N_1 \neq N_2$.

[0116] Referring to layout 100H of FIG. 1H, illustrates the extension of layout 100C for the network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1 = N_2 = 128$; $d = 2$; and $s = 2$. There are four super-quadrants in layout 100H namely top-left super-quadrant, bottom-left super-quadrant, top-right super-quadrant, bottom-right super-quadrant. Total number of blocks in the layout 100H is sixty four. Top-left super-quadrant implements the blocks from block 1_2 to block 31_32. Each block in all the super-quadrants has two more switches namely switch 6 and switch 7 in addition to the switches [1-5] illustrated in

US 2012/0269190 A1

Oct. 25, 2012

10

layout 100C of FIG. 1C. The inter-block link connection topology is the exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as it is shown in the layouts of FIG. 1D, FIG. 1E, FIG. 1F, and FIG. 1G respectively.

[0117] Bottom-left super-quadrant implements the blocks from block 33_34 to block 63_64. Top-right super-quadrant implements the blocks from block 65_66 to block 95_96. And bottom-right super-quadrant implements the blocks from block 97_98 to block 1_27_128. In all these three super-quadrants also, the inter-block link connection topology is exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as that of the top-left super-quadrant.

[0118] Recursively in accordance with the current invention, the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-left super-quadrant and bottom-left super-quadrant. And similarly the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-right super-quadrant and bottom-right super-quadrant. The inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of top-left super-quadrant and top-right super-quadrant. And similarly the inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of bottom-left super-quadrant and bottom-right super-quadrant.

[0119] Just as described for layout 100F of FIG. 1F, Applicant notes that the connection topology of inter-block links between switches 5 and 6 of each block of layout 100H of FIG. 1H is not the typical inverse Benes Network topology. In layout 100H first the switches 5 and 6 of nearest neighbor blocks are connected and then the rest of the blocks are connected in inverse Benes Network topology. For example since Block 11_12 and Block 33_34 are nearest neighbors in the leftmost column of layout 100H the corresponding inter-block links from switches 5 and 6 are connected together first. Then the remaining blocks in the leftmost column are connected in inverse Benes topology. For example in layout 100H since the remaining blocks in the leftmost column of top-left super-quadrant are Block 1_2, Block 3_4, and Block 9_10 and the remaining blocks in the leftmost column of bottom-left super-quadrant are Block 35_36, Block 41_42 and Block 43_44 the inter-block links between their corresponding switches 5 and 6 are connected together. In one embodiment the inter-block links of switches 5 and 6 corresponding to Block 1_2 and Block 35_36 are connected together; the inter-block links of switches 5 and 6 corresponding to Block 3_4 and Block 41_42 are connected together; and the inter-block links of switches 5 and 6 corresponding to Block 9_10 and Block 43_44 are connected together. (Similarly in another embodiment any one of the three blocks in the leftmost column of top-left super-quadrant can be connected with any one of the three blocks in the leftmost column of bottom-left super-quadrant of course as long as each block in leftmost column of top-left super-quadrant is connected to only one block in leftmost column of bottom-left super-quadrant and vice versa). Similarly in all the columns, the inter-block links between switches 5 and 6 are connected.

[0120] The bandwidth offered between top super-quadrants and bottom super-quadrants in layout 100H of FIG. 1H is 8's BW, since the bandwidth offered with in a square of blocks with the sides of the square consisting of eight neigh-

boring blocks is 8's BW. It must be noted that the bandwidth offered between top-left super-quadrant and bottom-left super-quadrant is 8's BW. That is inter-block links of a switch in each one of the blocks in top-left super-quadrant are connected to a switch in any one of the blocks in bottom-left super-quadrant and vice versa. Similarly the bandwidth offered between top-right super-quadrant and bottom-right super-quadrant is 8's BW. For example in one embodiment the bandwidth provided between Block 1_2 and Block 35_36 of layout 100H of FIG. 1H is 8's BW because inter-block links between switch 5 of Block 1_2 and switch 6 of Block 35_36 are connected and also inter-block links between switch 5 of Block 1_2 and switch 6 of Block 35_36 are connected. Similarly the bandwidth provided between any one of the blocks in top-left super-quadrant and any one of the bottom-left super-quadrant of layout 100H of FIG. 1H is 8's BW. It must be noted that the 8's BW is the bandwidth offered between the four super-quadrants in a square of four super-quadrants and it is neither the bandwidth offered between the four quadrants in one of the super-quadrants or with in each quadrant.

[0121] Just as described for layout 100G of FIG. 1G, Applicant notes that the connection topology of inter-block links between switches 6 and 7 of each block of layout 100H of FIG. 1H is not the typical inverse Benes Network topology. In layout 100H first the switches 6 and 7 of nearest neighbor blocks are connected and then the rest of the blocks are connected in inverse Benes Network topology. For example since Block 21_22 and Block 65_66 are nearest neighbors in the topmost row of layout 100H the corresponding inter-block links from switches 6 and 7 are connected together first. Then the remaining blocks in the topmost row are connected in inverse Benes topology. For example in layout 100H since the remaining blocks in the topmost row of top-left super-quadrant are Block 1_2, Block 5_6, and Block 17_18 and the remaining blocks in the topmost row of top-right super-quadrant are Block 69_70, Block 81_82 and Block 85_86 the inter-block links between their corresponding switches 6 and 7 are connected together. In one embodiment the inter-block links of switches 6 and 7 corresponding to Block 1_2 and Block 69_70 are connected together; the inter-block links of switches 6 and 7 corresponding to Block 5_6 and Block 81_82 are connected together; and the inter-block links of switches 6 and 7 corresponding to Block 17_18 and Block 85_86 are connected together. (Similarly in another embodiment any one of the three blocks in the topmost row of top-left super-quadrant can be connected with any one of the three blocks in the topmost row of top-right super-quadrant of course as long as each block in topmost row of top-right super-quadrant is connected to only one block in topmost row of top-right super-quadrant and vice versa). Similarly in all the rows, the inter-block links between switches 6 and 7 are connected.

[0122] The bandwidth offered between left super-quadrants and right super-quadrants in layout 100H of FIG. 1H is 8's BW, since the bandwidth offered with in a square of blocks with the sides of the square consisting of eight neighboring blocks is 8's BW. It must be noted that the bandwidth offered between top-left super-quadrant and top-right super-quadrant is 8's BW. That is inter-block links of a switch in each one of the blocks in top-left super-quadrant are connected to a switch in any one of the blocks in top-right super-quadrant and vice versa. Similarly the bandwidth offered between bottom-left super-quadrant and bottom-right super-quadrant is 8's BW. For example in one embodiment the bandwidth provided

US 2012/0269190 A1

Oct. 25, 2012

11

between Block 1_2 and Block 69_70 of layout 100H of FIG. 1H is 8's BW because inter-block links between switch 6 of Block 1_2 and switch 7 of Block 69_70 are connected and also inter-block links between switch 6 of Block 1_2 and switch 7 of Block 69_70 are connected. Similarly the bandwidth provided between any one of the blocks in top-left super-quadrant and any one of the blocks in top-right super-quadrant of layout 100H of FIG. 1H is 8's BW. Just the same way 2's BW is provided between two diagonal blocks, the bandwidth offered between two diagonal super-quadrants is 8's BW that is when the corresponding row and columns provide 8's BW.

[0123] Referring to diagram 100I of FIG. 1I illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block 1_2 in 100I illustrates both the intra-block and inter-block links connected to Block 1_2. The layout diagram 100I corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized folded multi-link multi-stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,389 that is incorporated by reference above.

[0124] That is the switches that are placed together in Block 1_2 as shown in FIG. 1I are namely input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switch MS(1,1) and middle switch MS(7,1) belonging to switch 2; middle switch MS(2,1) and middle switch MS(6,1) belonging to switch 3; middle switch MS(3,1) and middle switch MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0125] Input switch IS1 is implemented as two by four switch with the inlet links IL 1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1)-ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7), and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1-OL2 being the outputs of the output switch OS1.

[0126] Middle switch MS(1,1) is implemented as four by four switch with the middle links ML(1,1), ML(1,2), ML(1,7) and ML(1,8) being the inputs and middle links ML(2,1)-ML(2,4) being the outputs; and middle switch MS(7,1) is implemented as four by four switch with the middle links ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs and middle links ML(8,1)-ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 100I of FIG. 1I.

Generalized Multi-link Butterfly Fat Tree Network Embodiment:

[0127] In another embodiment in the network 100B of FIG. 1B, the switches that are placed together are implemented as combined switch then the network 100B is the generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,390

that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch. For example the input switch IS1 and output switch OS1 are placed together; so input switch IS1 and output OS1 are implemented as a six by six switch with the inlet links ILL IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and OL2 being the outputs of the combined switch IS1&OS1. Similarly in this embodiment of network 100B all the switches that are placed together are implemented as a combined switch.

[0128] Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$.

[0129] Referring to diagram 100J of FIG. 1J illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block 1_2 in 100J illustrates both the intra-block and inter-block links. The layout diagram 100J corresponds to the embodiment where the switches that are placed together are implemented as combined switch in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,390 that is incorporated by reference above.

[0130] That is the switches that are placed together in Block 1_2 as shown in FIG. 1J are namely the combined input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switch implemented is combined input and output switch IS1&OS1); middle switch MS(1,1) belonging to switch 2; middle switch MS(2,1) belonging to switch 3; middle switch MS(3,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0131] Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links ILL IL2 and ML(8,1), ML(8,2), ML(8,7), and ML(8,8) being the inputs and middle links ML(1,1)-ML(1,4), and outlet links OL1-OL2 being the outputs.

[0132] Middle switch MS(1,1) is implemented as eight by eight switch with the middle links ML(1,1), ML(1,2), ML(1,7), ML(1,8), ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs and middle links ML(2,1)-ML(2,4) and middle links ML(8,1)-ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as eight by eight switches as illustrated in 100J of FIG. 1J.

[0133] In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of $V_{mlink-bft}(N_1, N_2, d, s)$ can be implemented as a four by eight switch and a four by four switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1J, the left going middle

US 2012/0269190 A1

Oct. 25, 2012

12

links namely ML(7,1), ML(7,2), ML(7,11), and ML(7,12) are never switched to the right going middle links ML(2,1), ML(2,2), ML(2,3), and ML(2,4). And hence to implement MS(1,1) two switches namely: 1) a four by eight switch with the middle links ML(1,1), ML(1,2), ML(1,7), and ML(1,8) as inputs and the middle links ML(2,1), ML(2,2), ML(2,3), ML(2,4), ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs and 2) a four by four switch with the middle links ML(7,1), ML(7,2), ML(7,11), and ML(7,12) as inputs and the middle links ML(8,1), ML(8,2), ML(8,3), and ML(8,4) as outputs are sufficient without losing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

Generalized Multi-Stage Network Embodiment:

[0134] In one embodiment, in the network **100B** of FIG. 1B, the switches that are placed together are implemented as two separate switches in input stage **110** and output stage **120**; and as four separate switches in all the middle stages, then the network **100B** is the generalized folded multi stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage **110** and output stage **120** are implemented as a two by four switch and a four by two switch respectively. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs and outlet links OL1-OL2 being the outputs.

[0135] The switches, corresponding to the middle stages that are placed together are implemented as four two by two switches. For example middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs; Similarly in this embodiment of network **100B** all the switches that are placed together are implemented as separate switches.

[0136] Layout diagrams **100C** in FIG. 1C, **100D** in FIG. 1D, **100E** in FIG. 1E, **100F** in FIG. 1G are also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages. The layout **100C** in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$. Accordingly layout **100H** of FIG. 1H is also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$.

[0137] Referring to diagram **100K** of FIG. 1K illustrates a high-level implementation of Block **1_2** (Each of the other blocks have similar implementation) of the layout **100C** of

FIG. 1C which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block **1_2** in **100K** illustrates both the intra-block and inter-block links. The layout diagram **100K** corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network **100B** of FIG. 1B. As noted before then the network **100B** is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,391 that is incorporated by reference above.

[0138] That is the switches that are placed together in Block **1_2** as shown in FIG. 1K are namely the input switch IS1 and output switch OS1 belonging to switch **1**, illustrated by dotted lines, (as noted before switch **1** is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches MS(1,1), MS(1,17), MS(7,1) and MS(7,17) belonging to switch **2**; middle switches MS(2,1), MS(2,17), MS(6,1) and MS(6,17) belonging to switch **3**; middle switches MS(3,1), MS(3,17), MS(5,1) and MS(5,17) belonging to switch **4**; And middle switches MS(4,1), and MS(4,17) belonging to switch **5**.

[0139] Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs and outlet links OL1-OL2 being the outputs.

[0140] Middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as illustrated in **100K** of FIG. 1K.

Generalized Multi-Stage Network Embodiment with $S=1$:

[0141] In one embodiment, in the network **100B** of FIG. 1B (where it is implemented with $s=1$), the switches that are placed together are implemented as two separate switches in input stage **110** and output stage **120**; and as two separate switches in all the middle stages, then the network **100B** is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage **110** and output stage **120** are implemented as two, two by two switches. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1-OL2 being the outputs.

US 2012/0269190 A1

Oct. 25, 2012

13

[0142] The switches, corresponding to the middle stages that are placed together are implemented as two, two by two switches. For example middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs; Similarly in this embodiment of network 100B all the switches that are placed together are implemented as two separate switches.

[0143] Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$.

[0144] Referring to diagram 100K1 of FIG. 1K1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 100C of FIG. 1C when $s=1$ which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ (All the double links are replaced by single links when $s=1$). Block 1_2 in 100K1 illustrates both the intra-block and inter-block links. The layout diagram 100K1 corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 100B of FIG. 1B when $s=1$. As noted before then the network 100B is the generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,391 that is incorporated by reference above.

[0145] That is the switches that are placed together in Block 1_2 as shown in FIG. 1K1 are namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches MS(1,1) and MS(7,1) belonging to switch 2; middle switches MS(2,1) and MS(6,1) belonging to switch 3; middle switches MS(3,1) and MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0146] Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1-OL2 being the outputs.

[0147] Middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; And middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as illustrated in 100K1 of FIG. 1K1.

Generalized Butterfly Fat Tree Network Embodiment:

[0148] In another embodiment in the network 100B of FIG. 1B, the switches that are placed together are implemented as

two combined switches then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch. For example the input switch IS1 and output switch OS1 are placed together; so input output switch IS1&OS1 are implemented as a six by six switch with the inlet links ILL IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and OL2 being the outputs of the combined switch IS1&OS1.

[0149] The switches, corresponding to the middle stages that are placed together are implemented as two four by four switches. For example middle switches MS(1,1) and MS(1,17) are placed together; so middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; middle switch MS(1,17) is implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly in this embodiment of network 100B all the switches that are placed together are implemented as a two combined switches.

[0150] Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$.

[0151] Referring to diagram 100L of FIG. 1L illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block 1_2 in 100L illustrates both the intra-block and inter-block links. The layout diagram 100L corresponds to the embodiment where the switches that are placed together are implemented as two combined switches in the network 100B of FIG. 1B. As noted before then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,387 that is incorporated by reference above.

[0152] That is the switches that are placed together in Block 1_2 as shown in FIG. 1L are namely the combined input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switch implemented is combined input and output switch IS1&OS1); middle switch MS(1,1) and MS(1,17) belonging to switch 2; middle switch MS(2,1) and MS(2,17) belonging to switch 3; middle switch MS(3,1) and MS(3,17) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0153] Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links ILL IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs and middle links ML(1,1)-ML(1,4) and outlet links OL1-OL2 being the outputs.

US 2012/0269190 A1

Oct. 25, 2012

14

[0154] Middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; And middle switch MS(1,17) is implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two four by four switches as illustrated in 100L of FIG. 1L.

[0155] In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of $V_{m-link-bft}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1L, the left going middle links namely ML(7,1) and ML(7,11) are never switched to the right going middle links ML(2,1) and ML(2,3). And hence to implement MS(1,1) two switches namely: 1) a two by four switch with the middle links ML(1,1) and ML(1,7) as inputs and the middle links ML(2,1), ML(2,3), ML(8,1), and ML(8,3) as outputs and 2) a two by two switch with the middle links ML(7,1) and ML(7,11) as inputs and the middle links ML(8,1) and ML(8,3) as outputs are sufficient without loosing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

Generalized Butterfly Fat Tree Network Embodiment with S=1:

[0156] In one embodiment, in the network 100B of FIG. 1B (where it is implemented with $s=1$), the switches that are placed together are implemented as a combined switch in input stage 110 and output stage 120; and as a combined switch in all the middle stages, then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a four by four switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input and output switch IS1&OS1 is implemented as four by four switch with the inlet links ILL IL2, ML(8,1) and ML(8,3) being the inputs and middle links ML(1,1)-ML(1,2) and outlet links OL1-OL2 being the outputs

[0157] The switches, corresponding to the middle stages that are placed together are implemented as a four by four switch. For example middle switches MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,3), ML(7,1) and ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs.

[0158] Layout diagrams 100C in FIG. 1C, 100D in FIG. 1D, 100E in FIG. 1E, 100F in FIG. 1G are also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with five stages. The layout 100C in FIG. 1C can be recursively extended for any arbitrarily large generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$. Accordingly layout 100H of FIG. 1H is also applicable to generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$.

[0159] Referring to diagram 100L1 of FIG. 1L1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 100C of

FIG. 1C when $s=1$ which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ (All the double links are replaced by single links when $s=1$). Block 1_2 in 100K1 illustrates both the intra-block and inter-block links. The layout diagram 100L1 corresponds to the embodiment where the switches that are placed together are implemented as a combined switch in the network 100B of FIG. 1B when $s=1$. As noted before then the network 100B is the generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,387 that is incorporated by reference above.

[0160] That is the switches that are placed together in Block 1_2 as shown in FIG. 1L1 are namely the input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switch MS(1,1) belonging to switch 2; middle switch MS(2,1) belonging to switch 3; middle switch MS(3,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0161] Input and output switch IS1&OS1 are placed together; so input and output switch IS1&OS1 is implemented as four by four switch with the inlet links ILL IL2, ML(8,1) and ML(8,3) being the inputs and middle links ML(1,1)-ML(1,2) and outlet links OL1-OL2 being the outputs.

[0162] Middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,3), ML(7,1) and ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 100L1 of FIG. 1L1.

[0163] In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of $V_{m-link-bft}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 1L1, the left going middle links namely ML(7,1) and ML(7,5) are never switched to the right going middle links ML(2,1) and ML(2,2). And hence to implement MS(1,1) two switches namely: 1) a two by four switch with the middle links ML(1,1) and ML(1,3) as inputs and the middle links ML(2,1), ML(2,2), ML(8,1), and ML(8,2) as outputs and 2) a two by two switch with the middle links ML(7,1) and ML(7,5) as inputs and the middle links ML(8,1) and ML(8,2) as outputs are sufficient without loosing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

Symmetric RNB Generalized Multi-Link Multi-Stage Network $V_{m-link}(N_1, N_2, d, s)$, Connection Topology with $N_1 \neq 2^x$ & $N_2 \neq 2^y$ where x and y are integers:

[0164] Referring to diagram 200A in FIG. 2A, in one embodiment, an exemplary generalized multi-link multi-stage network $V_{m-link}(N_1, N_2, d, s)$ where $N_1=N_2=24$ and $2^4 < N = 24 < 2^5$; $d=2$; and $s=2$ with nine stages of ninety two switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, 150, 160, 170, 180 and 190 is shown where input stage 110 consists of

US 2012/0269190 A1

Oct. 25, 2012

15

twelve, two by four switches IS1-IS12 and output stage 120 consists of twelve, four by two switches OS1-OS12. And the middle stages namely the middle stage 130 consists of twelve, four by four switches MS(1,1)-MS(1,12), middle stage 140 consists of eight, four by four switches MS(2,1)-MS(2,8), middle stage 180 consists of eight, four by four switches MS(6,1)-MS(6,8), and middle stage 190 consists of twelve, four by four switches MS(7,1)-MS(7,12); middle stage 150 consists of twelve, four by four switches MS(3,1)-MS(3,12), middle stage 160 consists of eight, four by four switches MS(4,1)-MS(4,2), MS(4,5)-MS(4,6), MS(4,9)-MS(4,12), middle stage 170 consists of eight, four by four switches MS(5,1)-MS(5,2), MS(5,5)-MS(5,6), MS(5,9)-MS(5,12).

[0165] Such a generalized multi-link multi-stage network $V_{mink}(N_1, N_2, d, s)$ where $N_1=2^x$ & $N_2=2^y$ where x and y are integers, can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections, just the same way as when $N_1=2^x$ & $N_2=2^y$ where x and y are integers, as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,389 that is incorporated by reference above.

[0166] In one embodiment of this network each of the input switches IS1-IS12 and output switches OS1-OS12 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d, where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by a maximum of N/d. The size of each input switch IS1-IS12 can be denoted in general with the notation $d*2d$ and each output switch OS1-OS12 can be denoted in general with the notation $2d*d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d*2d$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-stage network can be represented with the notation $V_{mink}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL32), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0167] Each of the N/d input switches IS1-IS12 are connected to exactly d switches in middle stage 130 through two links each for a total of $2xd$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the middle links ML(1,1), ML(1,2), and also connected to middle switch MS(1,2) through the middle links ML(1,3) and ML(1,4)). Just the same way as defined before, the middle links which connect switches in the same row in two successive middle stages are called hereinafter straight middle links; and the middle links which connect switches in different rows in two successive middle stages are called hereinafter cross middle links. For example, the middle links ML(1,1) and ML(1,2) connect input switch IS1 and middle switch MS(1,1), so middle links ML(1,1) and ML(1,2) are straight middle links; where as the middle links ML(1,3) and ML(1,4) connect input switch IS1 and middle switch MS(1,2), since input switch IS1 and middle switch MS(1,2) belong to two different rows in diagram 100A of FIG. 1A, middle links ML(1,3) and ML(1,4) are cross middle links.

[0168] Each of the N/d middle switches MS(1,1)-MS(1,12) in the middle stage 130 are connected from exactly d input switches through two links each for a total of $2xd$ links (for

example the middle links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the middle links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2). Each of the middle switches MS(1,1)-MS(1,8) are connected to exactly d switches in middle stage 140 through two links each for a total of $2xd$ links (for example the middle links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to middle switch MS(2,1), and the middle links ML(2,3) and ML(2,4) are connected from middle switch MS(1,1) to middle switch MS(2,3)); and each of the middle switches MS(1,9)-MS(1,12) are connected to exactly d switches in middle stage 150 through two links each for a total of $2xd$ links (for example the middle links ML(3,33) and ML(3,34) are connected from middle switch MS(1,9) to middle switch MS(3,9), and the middle links ML(3,35) and ML(3,36) are connected from middle switch MS(1,9) to middle switch MS(3,11)).

[0169] Each of the middle switches MS(2,1)-MS(2,8) in the middle stage 140 are connected from exactly d middle switches in middle stage 130 through two links each for a total of $2xd$ links (for example the middle links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from input switch MS(1,1), and the middle links ML(1,11) and ML(1,12) are connected to the middle switch MS(2,1) from input switch MS(1,3)) and also are connected to exactly d switches in middle stage 150 through two links each for a total of $2xd$ links (for example the middle links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the middle links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,5)).

[0170] Each of the N/d middle switches MS(3,1)-MS(3,12) in the middle stage 150 are connected from exactly d middle switches in middle stage 140 through two links each for a total of $2xd$ links (for example the middle links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from input switch MS(2,1), and the middle links ML(2,19) and ML(2,20) are connected to the middle switch MS(3,1) from input switch MS(2,5)). Each of the middle switches MS(3,1)-MS(3,2), MS(3,5)-MS(3,6) and MS(3,9)-MS(3,12) are connected to exactly d switches in middle stage 160 through two links each for a total of $2xd$ links (for example the middle links ML(4,1) and ML(4,2) are connected from middle switch MS(3,1) to middle switch MS(4,1), and the middle links ML(4,3) and ML(4,4) are connected from middle switch MS(3,1) to middle switch MS(4,9)); and each of the middle switches MS(3,3)-MS(3,4) and MS(3,7)-MS(3,8) are connected to exactly d switches in middle stage 180 through two links each for a total of $2xd$ links (for example the middle links ML(6,9) and ML(6,10) are connected from middle switch MS(3,3) to middle switch MS(6,3), and the middle links ML(6,11) and ML(6,12) are connected from middle switch MS(3,3) to middle switch MS(6,7)).

[0171] Each of the middle switches MS(4,1)-MS(4,2), MS(4,5)-MS(4,6) and MS(4,9)-MS(4,12) in the middle stage 160 are connected from exactly d middle switches in middle stage 150 through two links each for a total of $2xd$ links (for example the middle links ML(4,1) and ML(4,2) are connected to the middle switch MS(4,1) from input switch MS(3,1), and the middle links ML(4,35) and ML(4,36) are connected to the middle switch MS(4,1) from input switch MS(3,9)) and also are connected to exactly d switches in middle stage 170 through two links each for a total of $2xd$ links (for

US 2012/0269190 A1

Oct. 25, 2012

16

example the middle links ML(5,1) and ML(5,2) are connected from middle switch MS(4,1) to middle switch MS(5,1), and the middle links ML(5,3) and ML(5,4) are connected from middle switch MS(4,1) to middle switch MS(5,9)).

[0172] Each of the middle switches MS(5,1)-MS(5,2), MS(5,5)-MS(5,6) and MS(5,9)-MS(5,12) in the middle stage 170 are connected from exactly d middle switches in middle stage 160 through two links each for a total of $2 \times d$ links (for example the middle links ML(5,1) and ML(5,2) are connected to the middle switch MS(5,1) from input switch MS(4,1), and the middle links ML(5,35) and ML(5,36) are connected to the middle switch MS(5,1) from input switch MS(4,9)). Each of the middle switches MS(5,1)-MS(5,2), MS(5,5)-MS(5,6) are connected to exactly d switches in middle stage 180 through two links each for a total of $2 \times d$ links (for example the middle links ML(6,1) and ML(6,2) are connected from middle switch MS(5,1) to middle switch MS(6,1), and the middle links ML(6,3) and ML(6,4) are connected from middle switch MS(5,1) to middle switch MS(6,5)); and Each of the middle switches MS(5,9)-MS(5,12) are connected to exactly d switches in middle stage 190 through two links each for a total of $2 \times d$ links (for example the middle links ML(6,33) and ML(6,34) are connected from middle switch MS(5,9) to middle switch MS(7,9), and the middle links ML(6,35) and ML(6,36) are connected from middle switch MS(5,9) to middle switch MS(7,11)).

[0173] Each of the N/d middle switches MS(6,1)-MS(6,8) in the middle stage 180 are connected from exactly d middle switches in middle stage 170 through two links each for a total of $2 \times d$ links (for example the middle links ML(6,1) and ML(6,2) are connected to the middle switch MS(6,1) from input switch MS(5,1), and the middle links ML(6,19) and ML(6,20) are connected to the middle switch MS(6,1) from input switch MS(5,5) and also are connected to exactly d switches in middle stage 190 through two links each for a total of $2 \times d$ links (for example the middle links ML(7,1) and ML(7,2) are connected from middle switch MS(6,1) to middle switch MS(7,1), and the middle links ML(7,3) and ML(7,4) are connected from middle switch MS(6,1) to middle switch MS(7,3)).

[0174] Each of the N/d middle switches MS(7,1)-MS(7,12) in the middle stage 190 are connected from exactly d middle switches in middle stage 180 through two links each for a total of $2 \times d$ links (for example the middle links ML(7,1) and ML(7,2) are connected to the middle switch MS(7,1) from input switch MS(6,1), and the middle links ML(7,11) and ML(7,12) are connected to the middle switch MS(7,1) from input switch MS(6,3)) and also are connected to exactly d switches in middle stage 120 through two links each for a total of $2 \times d$ links (for example the middle links ML(8,1) and ML(8,2) are connected from middle switch MS(7,1) to middle switch MS(8,1), and the middle links ML(8,3) and ML(8,4) are connected from middle switch MS(7,1) to middle switch OS2).

[0175] Each of the N/d middle switches OS1-OS12 in the middle stage 120 are connected from exactly d middle switches in middle stage 190 through two links each for a total of $2 \times d$ links (for example the middle links ML(8,1) and ML(8,2) are connected to the output switch OS1 from input switch MS(7,1), and the middle links ML(8,7) and ML(8,8) are connected to the output switch OS1 from input switch MS(7,2)).

[0176] Referring to diagram 200B in FIG. 2B, is a folded version of the multi-link multi-stage network 200A shown in

FIG. 2A. The network 200B in FIG. 2B shows input stage 110 and output stage 120 are placed together. That is input switch IS1 and output switch OS1 are placed together, input switch IS2 and output switch OS2 are placed together, and similarly input switch IS12 and output switch OS12 are placed together. All the right going links {i.e., inlet links IL1-IL24 and middle links ML(1,1)-ML(1,48)} correspond to input switches IS1-IS12, and all the left going links {i.e., middle links ML(8,1)-ML(8,48) and outlet links OL1-OL24} correspond to output switches OS1-OS12.

[0177] Middle stage 130 and middle stage 190 are placed together. That is middle switches MS(1,1) and MS(7,1) are placed together, middle switches MS(1,2) and MS(7,2) are placed together, and similarly middle switches MS(1,12) and MS(7,12) are placed together. All the right going middle links {i.e., middle links ML(1,1)-ML(1,48) and middle links ML(2,1)-ML(2,32) and the middle links ML(3,33)-ML(3,48)} correspond to middle switches MS(1,1)-MS(1,12), and all the left going middle links {i.e., middle links ML(7,1)-ML(7,32) and middle links ML(6,33)-ML(6,48) and middle links ML(8,1) and ML(8,48)} correspond to middle switches MS(7,1)-MS(7,12).

[0178] Middle stage 140 and middle stage 180 are placed together. That is middle switches MS(2,1) and MS(6,1) are placed together, middle switches MS(2,2) and MS(6,2) are placed together, and similarly middle switches MS(2,8) and MS(6,8) are placed together. All the right going middle links {i.e., middle links ML(2,1)-ML(2,48) and middle links ML(3,1)-ML(3,48)} correspond to middle switches MS(2,1)-MS(2,8), and all the left going middle links {i.e., middle links ML(6,1)-ML(6,48) and middle links ML(7,1) and ML(7,48)} correspond to middle switches MS(6,1)-MS(6,8).

[0179] Middle stage 150 and middle stage 170 are placed together. That is middle switches MS(3,1) and MS(5,1) are placed together, middle switches MS(3,2) and MS(5,2) are placed together, and similarly middle switches MS(3,12) and MS(5,12) are placed together. All the right going middle links {i.e., middle links ML(3,1)-ML(3,48) and middle links ML(4,1)-ML(4,48)} correspond to middle switches MS(3,1)-MS(3,12), and all the left going middle links {i.e., middle links ML(5,1)-ML(5,48) and middle links ML(6,1) and ML(6,48)} correspond to middle switches MS(5,1)-MS(5,12).

[0180] Middle stage 160 is placed alone. All the right going middle links are the middle links ML(4,1)-ML(4,8), ML(4,17)-ML(4,24) and ML(4,33)-ML(4,48) and all the left going middle links are middle links ML(5,1)-ML(5,8), ML(5,17)-ML(5,24) and ML(5,33)-ML(5,48).

[0181] In one embodiment, in the network 200B of FIG. 2B, the switches that are placed together are implemented as separate switches then the network 200B is the generalized folded multi link multi stage network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1 = N_2 = 24$; $d = 2$; and $s = 2$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,389 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a two by four switch and a four by two switch. For example the input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1)-ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1-OL2

US 2012/0269190 A1

Oct. 25, 2012

17

being the outputs of the output switch OS1. Similarly in this embodiment of network 200B all the switches that are placed together in each middle stage are implemented as separate switches.

Modified-Hypercube Topology Layout Schemes:

[0182] Referring to layout 200C of FIG. 2C, in one embodiment, there are twelve blocks namely Block 1_2, Block 3_4, Block 5_6, Block 7_8, Block 9_10, Block 11_12, Block 13_14, Block 15_16, Block 17_18, Block 19_20, Block 21_22, and Block 23_24. Each block implements all the switches in one row of the network 200B of FIG. 2B, one of the key aspects of the current invention. For example Block 1_2 implements the input switch IS1, output Switch OS1, middle switch MS(1,1), middle switch MS(7,1), middle switch MS(2,1), middle switch MS(6,1), middle switch MS(3,1), middle switch MS(5,1), and middle switch MS(4,1). For the simplification of illustration, Input switch IS1 and output switch OS1 together are denoted as switch 1; Middle switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; Middle switch MS(4,1) is denoted by switch 5.

[0183] All the straight middle links are illustrated in layout 200C of FIG. 2C. For example in Block 1_2, inlet links IL1-IL2, outlet links OL1-OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 200C of FIG. 2C.

[0184] Even though it is not illustrated in layout 200C of FIG. 2C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit depending on the applications in different embodiments. There are a maximum of four quadrants in the layout 200C of FIG. 2C namely top-left, bottom-left, top-right and bottom-right quadrants. In each quadrant there are a maximum of four blocks. Top-left quadrant implements Block 1_2, Block 3_4, Block 5_6, and Block 7_8. Bottom-left quadrant implements Block 9_10, Block 11_12, Block 13_14, and Block 15_16. Top-right quadrant implements Block 17_18, Block 19_20. Bottom-right quadrant implements Block 21_22, and Block 23_24. There are two halves in layout 200C of FIG. 2C namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants. Right-half consists of top-right and bottom-right quadrants.

[0185] Recursively in each quadrant there are a maximum of four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements Block 1_2. Bottom-left sub-quadrant of top-left quadrant implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 5_6. Finally bottom-right sub-quadrant of top-left quadrant implements Block 7_8. Similarly there are a maximum of two sub-halves in each quadrant. For example in top-left quadrant there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left quadrant implements Block 1_2 and Block

3_4. Right-sub-half of top-left quadrant implements Block 5_6 and Block 7_8. Finally applicant notes that in each quadrant or half the blocks are arranged close to binary hypercube.

[0186] Layout 200D of FIG. 2D illustrates the inter-block links between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Similarly middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 12 and switch 1 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 200D of FIG. 2D can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track). As described before, the inter-link bandwidth provided between two physically adjacent blocks in the same column is hereinafter called 2's bandwidth or 2's BW. For example the inter-block links between switches 1 and 2 as illustrated in layout 200D of FIG. 2D is 2's BW.

[0187] Layout 200E of FIG. 2E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 5_6. Similarly middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 1_2 and switch 2 of Block 5_6. It must be noted that if there are an odd number of blocks in the rows of blocks then one of the blocks do not need inter-block links between switches 2 and 3, and also one of the switches for example switch 3 does not need to be implemented. For example in layout 200E there are three blocks in the topmost row namely Block 1_2, Block 5_6 and Block 17_18. In layout 200E there is no need to have inter-block links between switches 2 and 3 of Block 17_18 and hence there is no need to implement switch 3. Similarly in Block 19_20, Block 21_22 and Block 23_24 there is no need to provide inter-block links between switches 2 and 3 in those blocks. Also switch 3 is not implemented in those blocks.

[0188] Applicant notes that the inter-block links illustrated in layout 200E of FIG. 2E can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

[0189] In general the bandwidth offered within a quadrant or a partial quadrant of the layout formed by two nearest neighboring blocks is 2's BW. For example in layout 200C of FIG. 2C the bandwidth offered in top-right quadrant is 2's BW. Similarly the bandwidth offered within each of the other three quadrants top-left, bottom-left and bottom-right quadrants is 2' BW. Alternatively the bandwidth offered within a square or a partial square of blocks with the sides of the square consisting of two neighboring blocks is 2's BW. This definition can be generalized so that the bandwidth offered within a

US 2012/0269190 A1

Oct. 25, 2012

18

square of blocks with the sides consisting of “x” number of blocks, where $2^{y-1} \leq x \leq 2^y$ where “y” is an integer, is hereinafter x’s BW.

[0190] Layout 200F of FIG. 2F illustrates the inter-block links between switches 3 and 4 of each block excepting that among the Block 17_18, Block 19_20, Block 21_22, and Block 23_24 the inter-block links are between the switches 2 and 4. For example middle links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 200F of FIG. 2F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time division multiplexed single track). For example the inter-block links between switches 3 and 4 as illustrated in layout 200F of FIG. 2F is 4’s BW.

[0191] Layout 200G of FIG. 2G illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 3_4. It must be noted that if the number of blocks in the rows of blocks is not a perfect multiple of four, then some of the blocks do not need inter-block links between switches 4 and 5, and also one of the switches for example switch 5 does not need to be implemented. For example in layout 200G there are three blocks in the topmost row namely Block 1_2, Block 5_6 and Block 17_18. In layout 200E there is no need to have inter-block links between switches 4 and 5 of Block 5_6 and hence there is no need to implement switch 5. Similarly in Block 7_8, Block 13_14 and Block 15_16 there is no need to provide inter-block links between switches 4 and 5 in those blocks. Also switch 5 is not implemented in those blocks.

[0192] Applicant notes that the inter-block links illustrated in layout 200G of FIG. 2G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division multiplexed single track). The bandwidth offered between top-left quadrant, bottom-left quadrant, top-right partial quadrant and bottom-right partial quadrant is 4’s BW in layout 200G of FIG. 2G.

[0193] The complete layout for the network 200B of FIG. 2B is given by combining the links in layout diagrams of 200C, 200D, 200E, 200F, and 200G. Applicant notes that in the layout 200C of FIG. 2C, the inter-block links between switch 1 and switch 2 of corresponding blocks are vertical tracks as shown in layout 200D of FIG. 2D; the inter-block links between switch 2 and switch 3 of corresponding blocks are horizontal tracks as shown in layout 200E of FIG. 2E; the inter-block links between switch 3 and switch 4 of corre-

sponding blocks are vertical tracks as shown in layout 200F of FIG. 2F; and finally the inter-block links between switch 4 and switch 5 of corresponding blocks are horizontal tracks as shown in layout 200G of FIG. 2G. The pattern is alternate vertical tracks and horizontal tracks.

[0194] Some of the key aspects of the current invention are discussed. 1) All the switches in one row of the multi-stage network 200B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks or vertical tracks; 3) Since all the inter-block links are either horizontal or vertical tracks, all the inter-block links can be mapped on to island-style architectures in current commercial FPGAs; 4) The length of the longest wire is about half of the width (or length) of the complete layout (For example middle link ML(4,4) is about half the width of the complete layout).

[0195] In accordance with the current invention, the layout 200C in FIG. 2C can be recursively extended for any arbitrarily large generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ the sub-quadrants, quadrants, and super-quadrants are arranged in d-ary hypercube manner and also the inter-blocks are accordingly connected in d-ary hypercube topology. Even though all the embodiments in the current invention are illustrated for $N_1=N_2$ when $N_1=N_2 \neq 2^x$ where x is an integer, the embodiments can be extended for $N_1 \neq 2^x$ & $N_2 \neq 2^y$ where x and y are integers.

[0196] Just the same as was illustrated in diagram 100I of FIG. 1I for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$, a high-level implementation of Block 1_2 of the layout 200C of FIG. 2C which represents a generalized folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ where $N_1=N_2=24$; $d=2$; and $s=2$ is similar.

[0197] Just the same as was illustrated in diagram 100J of FIG. 1J for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$, a high-level implementation of Block 1_2 of the layout 200C of FIG. 2C which represents a generalized multi-link butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ where $N_1=N_2=24$; $d=2$; and $s=2$ is similar.

[0198] Just the same as was illustrated in diagram 100K of FIG. 1K for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$, a high-level implementation of Block 1_2 of the layout 200C of FIG. 2C which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=24$; $d=2$; and $s=2$ is similar.

[0199] Just the same as was illustrated in diagram 100K1 of FIG. 1K1 for a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 100C of FIG. 1C which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$, a high-level implementation of Block 1_2 of the layout 200C of FIG. 2C which represents a generalized folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ where $N_1=N_2=24$; $d=2$; and $s=1$ is similar.

[0200] Just the same as was illustrated in diagram 100L of FIG. 1L for a high-level implementation of Block 1_2 (Each

US 2012/0269190 A1

Oct. 25, 2012

19

of the other blocks have similar implementation) of the layout **100C** of FIG. 1C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$, a high-level implementation of Block **1_2** of the layout **200C** of FIG. 2C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=24$; $d=2$; and $s=2$ is similar.

[0201] Just the same as was illustrated in diagram **100L1** of FIG. 1L1 for a high-level implementation of Block **1_2** (Each of the other blocks have similar implementation) of the layout **100C** of FIG. 1C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$, a high-level implementation of Block **1_2** of the layout **200C** of FIG. 2C which represents a generalized butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ where $N_1=N_2=24$; $d=2$; and $s=1$ is similar.

Modified-Hypercube Topology with Nearest Neighbor Connectivity First and the Remaining with Equal Length Wires, in Every Stage:

[0202] Referring to layout **300A** of FIG. 3A, **300B** of FIGS. 3B and **300C** of FIG. 3C illustrate the topmost row of the extension of layout **100H** for the network $V_{fold-mink}(N_1, N_2, d, s)$ where $N_1=N_2=512$; $d=2$; and $s=2$. In one embodiment of the complete layout, not shown in FIGS. 3A-3C, there are four super-super-quadrants namely top-left super-super-quadrant, bottom-left super-super-quadrant, top-right super-super-quadrant, and bottom-right super-super-quadrant. Total number of blocks in the complete layout is two hundred and fifty six. Top-left super-super-quadrant implements the blocks from block **1_2** to block **1_27_128**. Bottom-left super-super-quadrant implements the blocks from block **129_130** to block **255_256**. Top-right super-super-quadrant implements the blocks from block **257_258** to block **319_320**. Bottom-right super-super-quadrant implements the blocks from block **383_384** to block **511_512**. Each block in all the super-super-quadrants has two more switches namely switch **8** and switch **9** in addition to the switches [1-7] described in layout **100H** of FIG. 1H.

[0203] The embodiment of layout **300A** of FIG. 3A illustrates the 2's BW provided in the top-most row of the complete layout namely between block **1_2** and block **5_6**; between block **17_18** and block **21_22**; between block **65_66** and block **69_90**; between block **81_82** and block **85_86**; between block **257_258** and block **261_262**; between block **273_274** and block **275_276**; between block **321_322** and block **325_326**; and between block **337_338** and block **3_41_342**. In one embodiment, the 2's BW provided between the respective blocks is through the inter-block links between corresponding switch **2** and switch **3** of the respective blocks.

[0204] The embodiment of layout **300B** of FIG. 3B illustrates the 4's BW provided in the top-most row of the complete layout namely between block **1_2** and block **21_22**; between block **5_6** and block **17_18**; between block **65_66** and block **85_86**; between block **69_70** and block **81_82**; between block **257_258** and block **275_276**; between block **261_262** and block **273_274**; between block **321_322** and block **341_342**; and between block **325_326** and block **337_338**. In one embodiment, the 4's BW provided between the respective blocks is through the inter-block links between corresponding switch **4** and switch **5** of the respective blocks. In layout **300B**, nearest neighbor blocks are connected together to provide 4's BW (for example the 4's BW provided between block **5_6** and block **17_18**) and then the rest of the

blocks are connected to provide the 4's BW (for example the 4's BW provided between block **1_2** and block **21_22**).

[0205] The embodiment of layout **300C** of FIG. 3C illustrates the 8's BW provided in the top-most row of the complete layout namely between block **1_2** and block **69_70**; between block **5_6** and block **81_82**; between block **17_18** and block **85_86**; between block **21_22** and block **65_66**; between block **257_258** and block **325_326**; between block **261_262** and block **337_338**; between block **273_274** and block **341_342**; and between block **275_276** and block **321_322**. In one embodiment, the 8's BW provided between the respective blocks is through the inter-block links between corresponding switch **6** and switch **7** of the respective blocks. In layout **300C**, nearest neighbor blocks are connected together to provide 8's BW (for example the 8's BW provided between block **21_22** and block **65_66**) and then the rest of the blocks are connected to provide the 8's BW (for example the 8's BW provided between block **1_2** and block **69_70**). Modified-Hypercube Topology with Recursive Nearest Neighbor Connectivity, in Every Stage:

[0206] In another embodiment of the extension of layout **100H** for the network $V_{fold-mink}(N_1, N_2, d, s)$ where $N_1=N_2=512$; $d=2$; and $s=2$, the 2's BW and 4's BW are provided exactly the same as illustrated in FIG. 3A and FIG. 3B respectively; However 8's BW is offered as illustrated in layout **300D** of FIG. 3D. The 8's BW is provided in the top-most row of the complete layout namely between block **21_22** and block **65_66**; between block **17_18** and block **69_70**; between block **5_6** and block **81_82**; between block **1_2** and block **85_86**; between block **275_276** and block **321_322**; between block **273_274** and block **325_326**; between block **261_262** and block **337_338**; and between block **257_258** and block **341_342**. In one embodiment, the 8's BW provided between the respective blocks is through the inter-block links between corresponding switch **6** and switch **7** of the respective blocks.

[0207] In layout **300D**, nearest neighbor blocks are connected together to provide 8's BW recursively. Specifically first the 8's BW is provided between block **21_22** and block **65_66**. Then the 8's BW is provided between the nearest neighbor blocks in the remaining blocks, i.e., between block **17_18** and block **69_70**. Then the 8's BW is provided between the nearest neighbor blocks in the remaining blocks, i.e., between block **5_6** and block **81_82**. Finally the 8's BW is provided between the nearest neighbor blocks in the remaining blocks, i.e., between block **1_2** and block **85_86**. In the same manner, the 8's BW is provided in the remaining blocks between block **257_258** up to block **341_342**.

Modified-Hypercube Topology with the Second Stage Implementing Nearest Neighbor Connectivity:

[0208] Referring to layout **400A** of FIG. 4A, **400B** of FIGS. 4B and **400C** of FIG. 4C illustrate the topmost row of the extension of layout **100H** for the network $V_{fold-mink}(N_1, N_2, d, s)$ where $N_1=N_2=512$; $d=2$; and $s=2$. In another embodiment of the complete layout, not shown in FIGS. 4A-4C, there are four super-super-quadrants namely top-left super-super-quadrant, bottom-left super-super-quadrant, top-right super-super-quadrant, and bottom-right super-super-quadrant. Total number of blocks in the complete layout is two hundred fifty six. Top-left super-super-quadrant implements the blocks from block **1_2** to block **1_27_128**. Bottom-left super-super-quadrant implements the blocks from block **129_130** to block **255_256**. Top-right super-super-quadrant implements the blocks from block **257_258** to block **319_320**. Bottom-

US 2012/0269190 A1

Oct. 25, 2012

20

right super-super-quadrant implements the blocks from block 383_384 to block 511_512. Each block in all the super-super-quadrants has two more switches namely switch 8 and switch 9 in addition to the switches [1-7] described in layout 100H of FIG. 1H.

[0209] In the embodiment of Layout 400A of FIG. 4A illustrates the 2's BW provided in the top-most row of the complete layout namely between block 1_2 and block 5_6; between block 17_18 and block 21_22; between block 65_66 and block 69_90; between block 81_82 and block 85_86; between block 257_258 and block 261_262; between block 273_274 and block 275_276; between block 321_322 and block 325_326; and between block 337_338 and block 3_41_342. In one embodiment, the 2's BW provided between the respective blocks is through the inter-block links between corresponding switch 2 and switch 3 of the respective blocks. Applicant notes that in layout 400A of FIG. 4A the first stage provides 2's BW between the blocks in the top-most row of the complete layout.

[0210] In the embodiment of Layout 400B of FIG. 4B illustrates the nearest neighbor connectivity between blocks of the top-most row of the complete layout to provide 4's BW, 8's BW, and 16's BW namely between block 5_6 and block 17_18 the bandwidth provided is 4's BW; between block 21_22 and block 65_66 the bandwidth provided is 8's BW; between block 69_70 and block 81_82 the bandwidth provided is 4's BW; between block 85_86 and block 257_258 the bandwidth provided is 16's BW; between block 261_262 and block 273_274 the bandwidth provided is 4's BW; between block 275_276 and block 321_322 the bandwidth provided is 8's BW; between block 325_326 and block 337_338 the bandwidth provided is 4's BW; and between block 1_2 and block 341_342 no bandwidth is provided. (Even though it is not illustrated, in another embodiment 16's BW can be provided between block 1_2 and block 342_342). In one embodiment, the BW provided between the respective blocks is through the inter-block links between corresponding switch 4 and switch 5 of the respective blocks. Applicant notes that in layout 400B of FIG. 4B the second stage provides the remaining nearest neighbor connectivity (i.e., after the first stage connectivity in layout 400A of FIG. 4A as illustrated provides nearest neighbor connectivity with 100% 2's BW) namely 50% of 4's BW, 25% of 8's BW and 12.5% of 16's BW, between the blocks in the top-most row of the complete layout.

[0211] The embodiment of layout 400C of FIG. 4C illustrates the 4's BW and 8's BW provided in the top-most row of the complete layout namely between block 1_2 and block 21_22 the bandwidth provided is 4's BW; between block 5_6 and block 69_70 the bandwidth provided is 8's BW; between block 17_18 and block 81_82 the bandwidth provided is 8's BW; between block 65_66 and block 85_86 the bandwidth provided is 4's BW; between block 257_258 and block 275_276 the bandwidth provided is 4's BW; between block 261_262 and block 325_326 the bandwidth provided is 8's BW; between block 273_274 and block 341_342 the bandwidth provided is 4's BW; between block 275_276 and block 337_338 the bandwidth provided is 8's BW; and between block 321_322 and block 3_41_342 the bandwidth provided is 4's BW. In one embodiment, the 4's BW and 8's BW provided between the respective blocks is through the inter-block links between corresponding switch 6 and switch 7 of the respective blocks. Applicant notes that in layout 400C of FIG. 4C the

third stage provides 50% of 4's BW and 50% of 8's BW between the blocks in the top-most row of the complete layout.

[0212] The same process is repeated in the fourth stage by providing namely 25% of 8's BW and 87.5% of 16's BW is provided. This connectivity topology can be similarly extended to the network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2>512$; $d=2$; and $s=2$.

Modified-Hypercube Topology with Partial & Tapered Connectivity (Bandwidth) in a Stage, where $N_1=N_2=512$:

[0213] Referring to layout 500 of FIG. 5 illustrates the topmost row of the extension of layout 100H for the network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2=512$; $d=2$; and $s=2$. In another embodiment of the complete layout, not shown in FIG. 5, there are four super-super-quadrants namely top-left super-super-quadrant, bottom-left super-super-quadrant, top-right super-super-quadrant, and bottom-right super-super-quadrant. Total number of blocks in the complete layout is two hundred fifty six. Top-left super-super-quadrant implements the blocks from block 1_2 to block 1_27_128. Bottom-left super-super-quadrant implements the blocks from block 129_130 to block 255_256. Top-right super-super-quadrant implements the blocks from block 257_258 to block 319_320. Bottom-right super-super-quadrant implements the blocks from block 383_384 to block 511_512. Each block in all the super-super-quadrants has two more switches namely switch 8 and switch 9 in addition to the switches [1-7] described in layout 100H of FIG. 1H.

[0214] The embodiment of layout 500 of FIG. 5 illustrates the 8's BW and 16's BW provided in the top-most row of the complete layout namely between block 21_22 and block 65_66 the bandwidth provided is 8's BW; between block 17_18 and block 69_70 the bandwidth provided is 8's BW; between block 85_86 and block 257_258 the bandwidth provided is 16's BW; between block 81_82 and block 261_262 the bandwidth provided is 16's BW; between block 275_276 and block 321_322 the bandwidth provided is 8's BW; between block 273_274 and block 325_326 the bandwidth provided is 8's BW. In one embodiment, the 8's BW and 16's BW provided between the respective blocks is through the inter-block links between corresponding switch 6 and switch 7 of the respective blocks. Applicant notes that in layout 500 of FIG. 5 the bandwidth provided between the blocks in the top-most row of the complete layout may be in anyone of the stages. Applicant observes that the 8's bandwidth provided in layout 500 of FIG. 5 is 50% of total 8's BW for full connectivity and 16's BW provided is 25% of the total 16's BW for full connectivity. In layout 500 of FIG. 5, the partial 8's BW and 16's BW is provided in nearest neighbor connectivity manner recursively which makes the wire lengths between different blocks to offer 8's BW is different and also makes the wire lengths between different blocks to offer 16's BW is different. Layout 500 of FIG. 5 illustrates an embodiment to provide partial bandwidth in a tapered manner, where it is not needed to provide the complete bandwidth in the higher stages.

Modified-Hypercube Topology with Partial & Tapered Connectivity (Bandwidth) in a Stage, where $N_1=N_2=2048$:

[0215] Referring to layout 600 of FIG. 6 illustrates the topmost row of the extension of layout 100H for the network $V_{fold-mlink}(N_1, N_2, d, s)$ where $N_1=N_2=2048$; $d=2$; and $s=2$. In one embodiment of the complete layout, not shown in FIG. 6, there are four super-super-super-quadrants namely top-left super-super-super-quadrant, bottom-left super-super-super-

US 2012/0269190 A1

Oct. 25, 2012

21

quadrant, top-right super-super-super-quadrant, and bottom-right super-super-super-quadrant. Total number of blocks in the complete layout is one thousand and twenty four. Top-left super-super-quadrant implements the blocks from block 1_2 to block 511_512. Bottom-left super-super-quadrant implements the blocks from block 513_514 to block 1023_1024. Top-right super-super-quadrant implements the blocks from block 1025_1026 to block 1535_1536. Bottom-right super-super-quadrant implements the blocks from block 1537_1538 to block 2047_2048. Each block in all the super-super-quadrants has four more switches namely switch 8, switch 9, switch 10 and switch 11 in addition to the switches [1-7] described in layout 100H of FIG. 1H.

[0216] In the embodiment of Layout 600 of FIG. 6 illustrates the 8's BW, 16's BW and 32's BW provided in the top-most row of the complete layout namely between block 2122 and block 65_66 the bandwidth provided is 8's BW; between block 17_18 and block 69_70 the bandwidth provided is 8's BW; between block 85_86 and block 257_258 the bandwidth provided is 16's BW; between block 81_82 and block 261_262 the bandwidth provided is 16's BW; between block 275_276 and block 321_322 the bandwidth provided is 8's BW; between block 273_274 and block 325_326 the bandwidth provided is 8's BW; between block 3_41_342 and block 1025_1026 the bandwidth provided is 32's BW; between block 337_338 and block 1029_1030 the bandwidth provided is 32's BW; between block 1045_1046 and block 1089_1090 the bandwidth provided is 8's BW; between block 1041_1042 and block 1093_1094 the bandwidth provided is 8's BW; between block 1109_1110 and block 1_281_1282 the bandwidth provided is 16's BW; between block 1105_1106 and block 1_285_1286 the bandwidth provided is 16's BW; between block 1299_1300 and block 1345_1346 the bandwidth provided is 8's BW; and between block 1297_1298 and block 1349_1350 the bandwidth provided is 8's BW.

[0217] In one embodiment, the 8's BW, 16's BW, and 32's BW provided between the respective blocks is through the inter-block links between corresponding switch 10 and switch 11 of the respective blocks. Applicant notes that in layout 600 of FIG. 6 the bandwidth provided between the blocks in the top-most row of the complete layout may be in anyone of the stages. Applicant observes that the 8's bandwidth provided in layout 500 of FIG. 5 is 50% of total 8's BW for full connectivity, 16's BW provided is 25% of the total 16's BW for full connectivity and 32's BW provided is 12.5% of the total 32's BW for full connectivity.

[0218] Applicant notes that in layout 600 of FIG. 6 the length of some of the wires providing bandwidth to 8's BW, 16's BW and 32's BW are of equal size, and the length of rest of the wires providing bandwidth to 8's BW, 16's BW and 32's BW are of equal size. In layout 600 of FIG. 6, the partial 8's BW, 16's BW and 32's BW is provided in nearest neighbor connectivity manner recursively which makes the wire lengths between different blocks to offer 8's BW is different, also makes the wire lengths between different blocks to offer 16's BW is different and also makes the wire lengths between different blocks to offer 32's BW is different. Layout 600 of FIG. 6 illustrates an embodiment to provide partial bandwidth in a tapered manner, where it is not needed to provide the complete bandwidth in the higher stages.

Modified-Hypercube Topology with Partial & Tapered Connectivity (Bandwidth) with Equal Length Wires, in a Stage:

[0219] Referring to layout 700 of FIG. 7 illustrates the topmost row of the extension of layout 100H for the network $V_{fold-mink}(N_1, N_2, d, s)$ where $N_1=N_2=2048$; $d=2$; and $s=2$. In another embodiment of the complete layout, not shown in FIG. 7, there are four super-super-super-quadrants namely top-left super-super-super-quadrant, bottom-left super-super-super-quadrant, top-right super-super-super-quadrant, and bottom-right super-super-super-quadrant. Total number of blocks in the complete layout is one thousand and twenty four. Top-left super-super-quadrant implements the blocks from block 1_2 to block 511_512. Bottom-left super-super-quadrant implements the blocks from block 513_514 to block 1023_1024. Top-right super-super-quadrant implements the blocks from block 1025_1026 to block 1535_1536. Bottom-right super-super-quadrant implements the blocks from block 1537_1538 to block 2047_2048. Each block in all the super-super-quadrants has four more switches namely switch 8, switch 9, switch 10 and switch 11 in addition to the switches [1-7] described in layout 100H of FIG. 1H.

[0220] In the embodiment of Layout 700 of FIG. 7 illustrates the 8's BW, 16's BW and 32's BW provided in the top-most row of the complete layout namely between block 21_22 and block 69_70 the bandwidth provided is 8's BW; between block 17_18 and block 65_66 the bandwidth provided is 8's BW; between block 85_86 and block 261_262 the bandwidth provided is 16's BW; between block 81_82 and block 257_258 the bandwidth provided is 16's BW; between block 275_276 and block 325_326 the bandwidth provided is 8's BW; between block 273_274 and block 321_322 the bandwidth provided is 8's BW; between block 3_41_342 and block 1029_1030 the bandwidth provided is 32's BW; between block 337_338 and block 1025_1026 the bandwidth provided is 32's BW; between block 1045_1046 and block 1093_1094 the bandwidth provided is 8's BW; between block 1041_1042 and block 1089_1090 the bandwidth provided is 8's BW; between block 1109_1110 and block 1_285_1286 the bandwidth provided is 16's BW; between block 1105_1106 and block 1_281_1282 the bandwidth provided is 16's BW; between block 1299_1300 and block 1349_1350 the bandwidth provided is 8's BW; and between block 1297_1298 and block 1345_1346 the bandwidth provided is 8's BW.

[0221] In one embodiment, the 8's BW, 16's BW, and 32's BW provided between the respective blocks is through the inter-block links between corresponding switch 10 and switch 11 of the respective blocks. Applicant notes that in layout 700 of FIG. 7 the bandwidth provided between the blocks in the top-most row of the complete layout may be in anyone of the stages. Applicant observes that the 8's bandwidth provided in layout 500 of FIG. 5 is 50% of total 8's BW for full connectivity, 16's BW provided is 25% of the total 16's BW for full connectivity and 32's BW provided is 12.5% of the total 32's BW for full connectivity. Applicant notes that in layout 700 of FIG. 7 the length of the wires providing bandwidth to 8's BW, 16's BW and 32's BW are all of equal size. Layout 700 of FIG. 7 illustrates another embodiment to provide partial bandwidth in a tapered manner, where it is not needed to provide the complete bandwidth in the higher stages.

[0222] All the layout embodiments disclosed in the current invention are applicable to generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks V_{fold}

US 2012/0269190 A1

Oct. 25, 2012

22

(N_1, N_2, d, s) , generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$ and generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ for $s=1,2,3$ or any number in general, and for $N_1=N_2=N$, or $N_1 \neq N_2$, or $N_1 \neq 2^x$ & $N_2 \neq 2^y$ where x, y and d are integers.

[0223] Conversely applicant makes another important observation that generalized hypercube networks $V_{hcube}(N_1, N_2, d, s)$ are implemented with the layout topology being the hypercube topology shown in layout 100C of FIG. 1C with large scale cross point reduction as any one of the networks described in the current invention namely: generalized multi-stage networks $V(N_1, N_2, d, s)$, generalized folded multi-stage networks $V_{fold}(N_1, N_2, d, s)$, generalized butterfly fat tree networks $V_{bft}(N_1, N_2, d, s)$, generalized multi-link multi-stage networks $V_{mlink}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage networks $V_{fold-mlink}(N_1, N_2, d, s)$, generalized multi-link butterfly fat tree networks $V_{mlink-bft}(N_1, N_2, d, s)$ for $s=1,2,3$ or any number in general, and for $N_1=N_2=N$, or $N_1 \neq N_2$, or $N_1 \neq 2^x$ & $N_2 \neq 2^y$ where x, y and d are integers. Symmetric RNB Generalized Multi-Link Multi-Stage Pyramid Network $V_{mlink-p}(N_1, N_2, d, s)$, Connection Topology: Nearest Neighbor Connectivity and with More than Full Bandwidth:

[0224] Referring to diagram 800A in FIG. 8A, in one embodiment, an exemplary generalized multi-link multi-stage pyramid $V_{mlink-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages of one hundred and forty four switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, 140, 150, 160, 170, 180 and 190 is shown where input stage 110 consists of sixteen switches with ten of two by four switches namely IS1, IS3, IS5, IS6, IS8, IS9, IS11, IS13, IS14, and IS16; and six of two by six switches namely IS2, IS4, IS7, IS10, IS12 and IS15.

[0225] The output stage 120 consists of sixteen switches with ten of four by two switches namely OS1, OS3, OS5, OS6, OS8, OS9, OS11, OS13, OS14, and OS16; and six of six by two switches namely OS2, OS4, OS7, OS10, OS12, and OS15.

[0226] The middle stage 130 consists of sixteen switches with four of four by four switches namely MS(1,1), MS(1,6), MS(1,11), and MS(1,16); four of six by four switches namely MS(1,2), MS(1,5), MS(1,12) and MS(1,15); four of four by six switches namely MS(1,3), MS(1,8), MS(1,9), and MS(1,14); and four of six by six switches namely MS(1,4), MS(1,7), MS(1,10), and MS(1,13).

[0227] The middle stage 190 consists of sixteen switches with four of four by four switches namely MS(7,1), MS(7,6), MS(7,11), and MS(7,16); four of four by six switches namely MS(7,2), MS(7,5), MS(7,12) and MS(7,15); four of six by four switches namely MS(7,3), MS(7,8), MS(7,9), and MS(7,14); and four of six by six switches namely MS(7,4), MS(7,7), MS(7,10), and MS(7,13).

[0228] The middle stage 140 consists of sixteen switches with eight of four by four switches namely MS(2,1), MS(2,2), MS(2,5), MS(2,6), MS(2,11), MS(2,12), MS(2,15), and MS(2,16); and eight of six by four switches namely MS(2,3), MS(2,4), MS(2,7), MS(2,8), MS(2,9), MS(2,10), MS(2,13), and MS(2,14).

[0229] The middle stage 180 consists of sixteen switches with eight of four by four switches namely MS(6,1), MS(6,2), MS(6,5), MS(6,6), MS(6,11), MS(6,12), MS(6,15), and MS(6,16); and eight of four by six switches namely MS(6,3), MS(6,4), MS(6,7), MS(6,8), MS(6,9), MS(6,10), MS(6,13), and MS(6,14).

[0230] And all the remaining middle stages namely the middle stage 150 consists of sixteen, four by four switches MS(3,1)-MS(3,16), middle stage 160 consists of sixteen, four by four switches MS(4,1)-MS(4,16), and middle stage 170 consists of sixteen, four by four switches MS(5,1)-MS(5,16).

[0231] The multi-link multi-stage pyramid network $V_{mlink-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ shown in diagram 800A of FIG. 8A is built on top of the generalized multi-link multi-stage network $V_{mlink}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ by adding a few more links.

[0232] Since as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,389 that is incorporated by reference above, a network $V_{mlink}(N_1, N_2, d, s)$ can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections, the network $V_{mlink-p}(N_1, N_2, d, s)$ can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections.

[0233] In one embodiment of this network each of the input switches IS1-IS16 and output switches OS1-OS16 are crossbar switches. The number of switches of input stage 110 and of output stage 120 can be denoted in general with the variable N/d , where N is the total number of inlet links or outlet links. The number of middle switches in each middle stage is denoted by N/d . The size of each input switch IS1-IS16 can be denoted in general with the notation $d^+(2d)^+$ (hereinafter d^+ means d or more; or equivalently $\geq d$) and each output switch OS1-OS16 can be denoted in general with the notation $(2d)^+ * d^+$. Likewise, the size of each switch in any of the middle stages can be denoted as $(2d)^+ * (2d)^+$. A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a crossbar switch or a network of switches. A symmetric multi-stage network can be represented with the notation $V_{mlink-p}(N, d, s)$, where N represents the total number of inlet links of all input switches (for example the links IL1-IL32), d represents the inlet links of each input switch or outlet links of each output switch, and s is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0234] Each of the N/d input switches IS1-IS16 are connected to d^+ switches in middle stage 130 through two links each for a total of $(2 \times d)^+$ links (for example input switch IS2 is connected to middle switch MS(1,2) through the links ML(1,5), ML(1,6), and also connected to middle switch MS(1,1) through the links ML(1,7) and ML(1,8); In addition input switch IS2 is also connected to middle switch MS(1,5) through the links ML(1p,7) and ML(1p,8). The links ML(1,5), ML(1,6), ML(1,7) and ML(1,8) correspond to multistage network configuration and the links ML(1p,7) and ML(1p,8) correspond to the pyramid network configuration. Hereinafter all the pyramid links are denoted by ML(xp,y) where 'x' represents the stage the link belongs to and 'y' the link number in that stage.)

[0235] The middle links which connect switches in the same row in two successive middle stages are called hereinafter straight middle links; and the middle links which con-

US 2012/0269190 A1

Oct. 25, 2012

23

nect switches in different rows in two successive middle stages are called hereinafter cross middle links. For example, the middle links ML(1,1) and ML(1,2) connect input switch IS1 and middle switch MS(1,1), so middle links ML(1,1) and ML(1,2) are straight middle links; whereas the middle links ML(1,3) and ML(1,4) connect input switch IS1 and middle switch MS(1,2), since input switch IS1 and middle switch MS(1,2) belong to two different rows in diagram 800A of FIG. 8A, middle links ML(1,3) and ML(1,4) are cross middle links. It can be seen that pyramid links such as ML(1p,7) and ML(1p,8) are also cross middle links.

[0236] Each of the N/d middle switches MS(1,1)-MS(1,16) in the middle stage 130 are connected from d⁺ input switches through two links each for a total of (2×d)⁺0 links (for example the links ML(1,1) and ML(1,2) are connected to the middle switch MS(1,1) from input switch IS1, and the links ML(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from input switch IS2) and also are connected to d⁺ switches in middle stage 140 through two links each for a total of (2×d)⁺ links (for example the links ML(2,9) and ML(2,10) are connected from middle switch MS(1,3) to middle switch MS(2,3), and the links ML(2,11) and ML(2,12) are connected from middle switch MS(1,3) to middle switch MS(2,1); In addition middle switch MS(1,3) is also connected to middle switch MS(2,9) through the links ML(2p,11) and ML(2p,12). The links ML(2,9), ML(2,10), ML(2,11) and ML(2,12) correspond to multistage network configuration and the links ML(2p,11) and ML(2p,12) correspond to the pyramid network configuration.)

[0237] Each of the N/d middle switches MS(2,1)-MS(2,16) in the middle stage 140 are connected from d⁺ input switches through two links each for a total of (2×d)⁺ links (for example the links ML(2,1) and ML(2,2) are connected to the middle switch MS(2,1) from input switch MS(1,1), and the links ML(1,11) and ML(1,12) are connected to the middle switch MS(2,1) from input switch MS(1,3)) and also are connected to d⁺ switches in middle stage 150 through two links each for a total of (2×d)⁺ links (for example the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch MS(3,1), and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(3,6)).

[0238] Each of the N/d middle switches MS(3,1)-MS(3,16) in the middle stage 150 are connected from d⁺ input switches through two links each for a total of (2×d)⁺ links (for example the links ML(3,1) and ML(3,2) are connected to the middle switch MS(3,1) from input switch MS(2,1), and the links ML(2,23) and ML(2,24) are connected to the middle switch MS(3,1) from input switch MS(2,6)) and also are connected to d⁺ switches in middle stage 160 through two links each for a total of (2×d)⁺ links (for example the links ML(4,1) and ML(4,2) are connected from middle switch MS(3,1) to middle switch MS(4,1), and the links ML(4,3) and ML(4,4) are connected from middle switch MS(3,1) to middle switch MS(4,11)).

[0239] Each of the N/d middle switches MS(4,1)-MS(4,16) in the middle stage 160 are connected from d⁺ input switches through two links each for a total of (2×d)⁺ links (for example the links ML(4,1) and ML(4,2) are connected to the middle switch MS(4,1) from input switch MS(3,1), and the links ML(4,43) and ML(4,44) are connected to the middle switch MS(4,1) from input switch MS(3,11)) and also are connected to d⁺ switches in middle stage 170 through two links each for a total of (2×d)⁺ links (for example the links ML(5,1) and

ML(5,2) are connected from middle switch MS(4,1) to middle switch MS(5,1), and the links ML(5,3) and ML(5,4) are connected from middle switch MS(4,1) to middle switch MS(5,11)).

[0240] Each of the N/d middle switches MS(5,1)-MS(5,16) in the middle stage 170 are connected from d⁺ input switches through two links each for a total of (2×d)⁺ links (for example the links ML(5,1) and ML(5,2) are connected to the middle switch MS(5,1) from input switch MS(4,1), and the links ML(5,43) and ML(5,44) are connected to the middle switch MS(5,1) from input switch MS(4,11)) and also are connected to d⁺ switches in middle stage 180 through two links each for a total of (2×d)⁺ links (for example the links ML(6,1) and ML(6,2) are connected from middle switch MS(5,1) to middle switch MS(6,1), and the links ML(6,3) and ML(6,4) are connected from middle switch MS(5,1) to middle switch MS(6,6)).

[0241] Each of the N/d middle switches MS(6,1)-MS(6,16) in the middle stage 180 are connected from d⁺ input switches through two links each for a total of (2×d)⁺ links (for example the links ML(6,1) and ML(6,2) are connected to the middle switch MS(6,1) from input switch MS(5,1), and the links ML(6,23) and ML(6,24) are connected to the middle switch MS(6,1) from input switch MS(5,6)) and also are connected to d⁺ switches in middle stage 190 through two links each for a total of (2×d)⁺ links (for example the links ML(7,9) and ML(7,10) are connected from middle switch MS(6,3) to middle switch MS(7,3), and the links ML(7,11) and ML(7,12) are connected from middle switch MS(6,3) to middle switch MS(7,1); In addition middle switch MS(6,3) is also connected to middle switch MS(7,9) through the links ML(7p,11) and ML(7p,12). The links ML(7,9), ML(7,10), ML(7,11) and ML(7,12) correspond to multistage network configuration and the links ML(7p,11) and ML(7p,12) correspond to the pyramid network configuration.)

[0242] Each of the N/d middle switches MS(7,1)-MS(7,16) in the middle stage 190 are connected from d⁺ input switches through two links each for a total of (2×d)⁺ links (for example the links ML(7,1) and ML(7,2) are connected to the middle switch MS(7,1) from input switch MS(6,1), and the links ML(7,11) and ML(7,12) are connected to the middle switch MS(7,1) from input switch MS(6,3)) and also are connected to d⁺ switches in middle stage 120 through two links each for a total of (2×d)⁺ links (for example middle switch MS(7,2) is connected to output switch OS2 through the links ML(8,5), ML(8,6), and also connected to middle switch OS1 through the links ML(8,7) and ML(8,8); In addition middle switch MS(7,2) is also connected to output switch OS5 through the links ML(8p,7) and ML(8p,8). The links ML(8,5), ML(8,6), ML(8,7) and ML(8,8) correspond to multistage network configuration and the links ML(8p,7) and ML(8p,8) correspond to the pyramid network configuration.)

[0243] Each of the N/d middle switches OS1-OS16 in the middle stage 120 are connected from d⁺ input switches through two links each for a total of (2×d)⁺ links (for example the links ML(8,1) and ML(8,2) are connected to the output switch OS1 from input switch MS(7,1), and the links ML(8,7) and ML(7,8) are connected to the output switch OS1 from input switch MS(7,2)).

[0244] Finally the connection topology of the network 800A shown in FIG. 8A is logically similar to back to back inverse Benes connection topology. In addition there are additional nearest neighbor links (i.e., pyramid links as described before) between the input stage 110 and middle stage 130;

US 2012/0269190 A1

Oct. 25, 2012

24

between middle stage **130** and middle stage **140**; between middle stage **180** and middle stage **190**; and middle stage **190** and output stage **120**.

[0245] Applicant notes that in a multi-stage pyramid network with a fully connected multi-stage network configuration the pyramid links may not contribute for the connectivity however these links can be cleverly used to reduce the latency and power in an integrated circuit even though the number of cross points required are more to connect pyramid links than is required in a purely multi-stage network.

[0246] Applicant notes that in the generalized multi-link multi-stage pyramid network $V_{m-link-p}(N_1, N_2, d, s)$ the pyramid links are provided between any two successive stages as illustrated in the diagram **800A** of FIG. **8A**. The pyramid links in general are also provided between the switches in the same stage. The pyramid links are also provided between any two arbitrary stages.

[0247] Referring to diagram **800B** in FIG. **8B**, is a folded version of the multi-link multi-stage pyramid network **800A** shown in FIG. **8A**. The network **800B** in FIG. **8B** shows input stage **110** and output stage **120** are placed together. That is input switch **IS1** and output switch **OS1** are placed together, input switch **IS2** and output switch **OS2** are placed together, and similarly input switch **IS16** and output switch **OS16** are placed together. All the right going links {i.e., inlet links **IL1-IL32** and middle links **ML(1,1)-ML(1,64)**} correspond to input switches **IS1-IS16**, and all the left going links {i.e., middle links **ML(8,1)-ML(8,64)** and outlet links **OL1-OL32**} correspond to output switches **OS1-OS16**.

[0248] Middle stage **130** and middle stage **190** are placed together. That is middle switches **MS(1,1)** and **MS(7,1)** are placed together, middle switches **MS(1,2)** and **MS(7,2)** are placed together, and similarly middle switches **MS(1,16)** and **MS(7,16)** are placed together. All the right going middle links {i.e., middle links **ML(1,1)-ML(1,64)** and middle links **ML(2,1)-ML(2,64)**} correspond to middle switches **MS(1,1)-MS(1,16)**, and all the left going middle links {i.e., middle links **ML(7,1)-ML(7,64)** and middle links **ML(8,1)** and **ML(8,64)**} correspond to middle switches **MS(7,1)-MS(7,16)**.

[0249] Middle stage **140** and middle stage **180** are placed together. That is middle switches **MS(2,1)** and **MS(6,1)** are placed together, middle switches **MS(2,2)** and **MS(6,2)** are placed together, and similarly middle switches **MS(2,16)** and **MS(6,16)** are placed together. All the right going middle links {i.e., middle links **ML(2,1)-ML(2,64)** and middle links **ML(3,1)-ML(3,64)**} correspond to middle switches **MS(2,1)-MS(2,16)**, and all the left going middle links {i.e., middle links **ML(6,1)-ML(6,64)** and middle links **ML(7,1)** and **ML(7,64)**} correspond to middle switches **MS(6,1)-MS(6,16)**.

[0250] Middle stage **150** and middle stage **170** are placed together. That is middle switches **MS(3,1)** and **MS(5,1)** are placed together, middle switches **MS(3,2)** and **MS(5,2)** are placed together, and similarly middle switches **MS(3,16)** and **MS(5,16)** are placed together. All the right going middle links {i.e., middle links **ML(3,1)-ML(3,64)** and middle links **ML(4,1)-ML(4,64)**} correspond to middle switches **MS(3,1)-MS(3,16)**, and all the left going middle links {i.e., middle links **ML(5,1)-ML(5,64)** and middle links **ML(6,1)** and **ML(6,64)**} correspond to middle switches **MS(5,1)-MS(5,16)**.

[0251] Middle stage **160** is placed alone. All the right going middle links are the middle links **ML(4,1)-ML(4,64)** and all the left going middle links are middle links **ML(5,1)-ML(5,64)**.

[0252] Just the same way as the connection topology of the network **800A** shown in FIG. **8A**, the connection topology of the network **800B** shown in FIG. **8B** is the folded version and logically similar to back to back inverse Benes connection topology. In addition there are additional nearest neighbor links (i.e., pyramid links as described before) between the input stage **110** and middle stage **130**; between middle stage **130** and middle stage **140**; between middle stage **180** and middle stage **190**; and middle stage **190** and output stage **120**.

[0253] The multi-link multi-stage pyramid network $V_{fold-m-link-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ shown in diagram **800B** of FIG. **8B** is built on top of the generalized multi-link multi-stage network $V_{fold-m-link}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ by also adding a few more links.

[0254] Since as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,389 that is incorporated by reference above, a network $V_{fold-m-link}(N_1, N_2, d, s)$ can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections, the network $V_{fold-m-link-p}(N_1, N_2, d, s)$ can be operated in rearrangeably non-blocking manner for arbitrary fan-out multicast connections and also can be operated in strictly non-blocking manner for unicast connections.

[0255] In one embodiment, in the network **800B** of FIG. **8B**, the switches that are placed together are implemented as separate switches then the network **800B** is the generalized folded multi link multi stage pyramid network $V_{fold-m-link-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages. That is the switches that are placed together in input stage **110** and output stage **120** are implemented as a two by four switch and a four by two switch respectively. For example the input switch **IS1** and output switch **OS1** are placed together; so input switch **IS1** is implemented as two by four switch with the inlet links **IL1** and **IL2** being the inputs of the input switch **IS1** and middle links **ML(1,1)-ML(1,4)** being the outputs of the input switch **IS1**; and output switch **OS1** is implemented as four by two switch with the middle links **ML(8,1)**, **ML(8,2)**, **ML(8,7)** and **ML(8,8)** being the inputs of the output switch **OS1** and outlet links **OL1-OL2** being the outputs of the output switch **OS1**. Similarly in this embodiment of network **800B** all the switches that are placed together in each middle stage are implemented as separate switches.

Modified-Hypercube Topology Layout Scheme:

[0256] Referring to layout **800C** of FIG. **8C**, in one embodiment, there are sixteen blocks namely Block **1_2**, Block **3_4**, Block **5_6**, Block **7_8**, Block **9_10**, Block **11_12**, Block **13_14**, Block **15_16**, Block **17_18**, Block **19_20**, Block **21_22**, Block **23_24**, Block **25_26**, Block **27_28**, Block **29_30**, and Block **31_32**. Each block implements all the switches in one row of the network **800B** of FIG. **8B**, one of the key aspects of the current invention. For example Block **1_2** implements the input switch **IS1**, output Switch **OS1**, middle switch **MS(1,1)**, middle switch **MS(7,1)**, middle switch **MS(2,1)**, middle switch **MS(6,1)**, middle switch **MS(3,1)**, middle switch **MS(5,1)**, and middle switch **MS(4,1)**. For the simplification of illustration, Input switch **IS1** and output switch **OS1** together are denoted as switch **1**; Middle

US 2012/0269190 A1

Oct. 25, 2012

25

switch MS(1,1) and middle switch MS(7,1) together are denoted by switch 2; Middle switch MS(2,1) and middle switch MS(6,1) together are denoted by switch 3; Middle switch MS(3,1) and middle switch MS(5,1) together are denoted by switch 4; Middle switch MS(4,1) is denoted by switch 5.

[0257] All the straight middle links are illustrated in layout 800C of FIG. 8C. For example in Block 1_2, inlet links IL1-IL2, outlet links OL1-OL2, middle link ML(1,1), middle link ML(1,2), middle link ML(8,1), middle link ML(8,2), middle link ML(2,1), middle link ML(2,2), middle link ML(7,1), middle link ML(7,2), middle link ML(3,1), middle link ML(3,2), middle link ML(6,1), middle link ML(6,2), middle link ML(4,1), middle link ML(4,2), middle link ML(5,1) and middle link ML(5,2) are illustrated in layout 800C of FIG. 8C.

[0258] Even though it is not illustrated in layout 800C of FIG. 8C, in each block, in addition to the switches there may be Configurable Logic Blocks (CLB) or any arbitrary digital circuit depending on the applications in different embodiments. There are four quadrants in the layout 800C of FIG. 8C namely top-left, bottom-left, top-right and bottom-right quadrants. Top-left quadrant implements Block 1_2, Block 3_4, Block 5_6, and Block 7_8. Bottom-left quadrant implements Block 9_10, Block 11_12, Block 13_14, and Block 15_16. Top-right quadrant implements Block 17_18, Block 19_20, Block 21_22, and Block 23_24. Bottom-right quadrant implements Block 25_26, Block 27_28, Block 29_30, and Block 31_32. There are two halves in layout 800C of FIG. 8C namely left-half and right-half. Left-half consists of top-left and bottom-left quadrants. Right-half consists of top-right and bottom-right quadrants.

[0259] Recursively in each quadrant there are four sub-quadrants. For example in top-left quadrant there are four sub-quadrants namely top-left sub-quadrant, bottom-left sub-quadrant, top-right sub-quadrant and bottom-right sub-quadrant. Top-left sub-quadrant of top-left quadrant implements Block 1_2. Bottom-left sub-quadrant of top-left quadrant implements Block 3_4. Top-right sub-quadrant of top-left quadrant implements Block 5_6. Finally bottom-right sub-quadrant of top-left quadrant implements Block 7_8. Similarly there are two sub-halves in each quadrant. For example in top-left quadrant there are two sub-halves namely left-sub-half and right-sub-half. Left-sub-half of top-left quadrant implements Block 1_2 and Block 3_4. Right-sub-half of top-left quadrant implements Block 5_6 and Block 7_8. Finally applicant notes that in each quadrant or half the blocks are arranged as a general binary hypercube. Recursively in larger multi-stage network $V_{fold-mlink-p}(N_1, N_2, d, s)$ where $N_1=N_2>32$, the layout in this embodiment in accordance with the current invention, will be such that the super-quadrants will also be arranged in d-ary hypercube manner. (In the embodiment of the layout 800C of FIG. 8C, it is binary hypercube manner since $d=2$, in the network $V_{fold-mlink-p}(N_1, N_2, d, s)$ 800B of FIG. 8B).

[0260] Layout 800D of FIG. 8D illustrates the inter-block links between switches 1 and 2 of each block. For example middle links ML(1,3), ML(1,4), ML(8,7), and ML(8,8) are connected between switch 1 of Block 1_2 and switch 2 of Block 3_4. Middle links ML(1,7), ML(1,8), ML(8,3), and ML(8,4) are connected between switch 2 of Block 1_2 and switch 1 of Block 3_4. Similarly pyramid middle links ML(1p,7), ML(1p,8), ML(8p,19), and ML(8p,20) are connected between switch 1 of Block 3_4 and switch 2 of Block

9_10. Similarly pyramid middle links ML(1p,19), ML(1p,20), ML(8p,7), and ML(8p,8) are connected between switch 2 of Block 3_4 and switch 1 of Block 9_10.

[0261] Applicant notes that the inter-block links illustrated in layout 800D of FIG. 8D can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(1,4) and ML(8,8) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(1,4) and ML(8,8) are implemented as a time division multiplexed single track).

[0262] Layout 800E of FIG. 8E illustrates the inter-block links between switches 2 and 3 of each block. For example middle links ML(2,3), ML(2,4), ML(7,11), and ML(7,12) are connected between switch 2 of Block 1_2 and switch 3 of Block 3_4. Middle links ML(2,11), ML(2,12), ML(7,3), and ML(7,4) are connected between switch 3 of Block 12 and switch 2 of Block 3_4. Similarly pyramid middle links ML(2p,35), ML(2p,36), ML(7p,11), and ML(7p,12) are connected between switch 1 of Block 5_6 and switch 2 of Block 17_18. Similarly pyramid middle links ML(2p,11), ML(2p,12), ML(7p,35), and ML(7p,36) are connected between switch 2 of Block 5_6 and switch 1 of Block 17_18.

[0263] Applicant notes that the inter-block links illustrated in layout 800E of FIG. 8E can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(2,12) and ML(7,4) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(2,12) and ML(7,4) are implemented as a time division multiplexed single track).

[0264] Layout 800F of FIG. 8F illustrates the inter-block links between switches 3 and 4 of each block. For example middle links ML(3,3), ML(3,4), ML(6,19), and ML(6,20) are connected between switch 3 of Block 1_2 and switch 4 of Block 3_4. Similarly middle links ML(3,19), ML(3,20), ML(6,3), and ML(6,4) are connected between switch 4 of Block 1_2 and switch 3 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 800F of FIG. 8F can be implemented as vertical tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(3,4) and ML(6,20) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as a time division multiplexed single track (for example middle links ML(3,4) and ML(6,20) are implemented as a time division multiplexed single track).

[0265] Layout 800G of FIG. 8G illustrates the inter-block links between switches 4 and 5 of each block. For example middle links ML(4,3), ML(4,4), ML(5,35), and ML(5,36) are connected between switch 4 of Block 1_2 and switch 5 of Block 3_4. Similarly middle links ML(4,35), ML(4,36), ML(5,3), and ML(5,4) are connected between switch 5 of Block 1_2 and switch 4 of Block 3_4. Applicant notes that the inter-block links illustrated in layout 800G of FIG. 8G can be implemented as horizontal tracks in one embodiment. Also in one embodiment inter-block links are implemented as two different tracks (for example middle links ML(4,4) and ML(5,36) are implemented as two different tracks); or in an alternative embodiment inter-block links are implemented as

US 2012/0269190 A1

Oct. 25, 2012

26

a time division multiplexed single track (for example middle links ML(4,4) and ML(5,36) are implemented as a time division multiplexed single track).

[0266] The complete layout for the network 800B of FIG. 8B is given by combining the links in layout diagrams of 800C, 800D, 800E, 800F, and 800G. Applicant notes that in the layout 800C of FIG. 8C, the inter-block links between switch 1 and switch 2 of corresponding blocks are vertical tracks as shown in layout 800D of FIG. 8D; the inter-block links between switch 2 and switch 3 of corresponding blocks are horizontal tracks as shown in layout 800E of FIG. 8E; the inter-block links between switch 3 and switch 4 of corresponding blocks are vertical tracks as shown in layout 800F of FIG. 8F; and finally the inter-block links between switch 4 and switch 5 of corresponding blocks are horizontal tracks as shown in layout 800G of FIG. 8G. The pattern is alternate vertical tracks and horizontal tracks. It continues recursively for larger networks of $N > 32$ as will be illustrated later.

[0267] Some of the key aspects of the current invention are discussed. 1) All the switches in one row of the multi-stage network 800B are implemented in a single block. 2) The blocks are placed in such a way that all the inter-block links are either horizontal tracks or vertical tracks; 3) Since all the inter-block links are either horizontal or vertical tracks, all the inter-block links can be mapped on to island-style architectures in current commercial FPGA's; 4) The length of the longest wire is about half of the width (or length) of the complete layout (For example middle link ML(4,4) is about half the width of the complete layout).

[0268] In accordance with the current invention, the layout 800C in FIG. 8C can be recursively extended for any arbitrarily large generalized folded multi-link multi-stage network $V_{fold-mlink-p}(N_1, N_2, d, s)$ the sub-quadrants, quadrants, and super-quadrants are arranged in d-ary hypercube manner and also the inter-blocks are accordingly connected in d-ary hypercube topology. Even though all the embodiments in the current invention are illustrated for $N_1 = N_2$, the embodiments can be extended for $N_1 \neq N_2$.

[0269] Referring to layout 800H of FIG. 8H, illustrates the extension of layout 800C for the network $V_{fold-mlink-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 128$; $d = 2$; and $s = 2$. There are four super-quadrants in layout 800H namely top-left super-quadrant, bottom-left super-quadrant, top-right super-quadrant, bottom-right super-quadrant. Total number of blocks in the layout 800H is sixty four. Top-left super-quadrant implements the blocks from block 1_2 to block 31_32. Each block in all the super-quadrants has two more switches namely switch 6 and switch 7 in addition to the switches [1-5] illustrated in layout 800C of FIG. 8C. The inter-block link connection topology is the exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as it is shown in the layouts of FIG. 8D, FIG. 8E, FIG. 8F, and FIG. 8G respectively.

[0270] Bottom-left super-quadrant implements the blocks from block 33_34 to block 63_64. Top-right super-quadrant implements the blocks from block 65_66 to block 95_96. And bottom-right super-quadrant implements the blocks from block 97_98 to block 1_27_128. In all these three super-quadrants also, the inter-block link connection topology is exactly the same between the switches 1 and 2; switches 2 and 3; switches 3 and 4; switches 4 and 5 as that of the top-left super-quadrant.

[0271] Recursively in accordance with the current invention, the inter-block links connecting the switch 5 and switch

6 will be vertical tracks between the corresponding switches of top-left super-quadrant and bottom-left super-quadrant. And similarly the inter-block links connecting the switch 5 and switch 6 will be vertical tracks between the corresponding switches of top-right super-quadrant and bottom-right super-quadrant. The inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of top-left super-quadrant and top-right super-quadrant. And similarly the inter-block links connecting the switch 6 and switch 7 will be horizontal tracks between the corresponding switches of bottom-left super-quadrant and bottom-right super-quadrant.

[0272] Referring to diagram 800I of FIG. 8I illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 800C of FIG. 8C which represents a generalized folded multi-link multi-stage network $V_{fold-mlink-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$. Block 1_2 in 800I illustrates both the intra-block and inter-block links connected to Block 1_2. The layout diagram 800I corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 800B of FIG. 8B. As noted before then the network 800B is the generalized folded multi-link multi-stage network $V_{fold-mlink-p}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with nine stages.

[0273] That is the switches that are placed together in Block 1_2 as shown in FIG. 8I are namely input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switch MS(1,1) and middle switch MS(7,1) belonging to switch 2; middle switch MS(2,1) and middle switch MS(6,1) belonging to switch 3; middle switch MS(3,1) and middle switch MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0274] Input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs of the input switch IS1 and middle links ML(1,1)-ML(1,4) being the outputs of the input switch IS1; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7), and ML(8,8) being the inputs of the output switch OS1 and outlet links OL1-OL2 being the outputs of the output switch OS1.

[0275] Middle switch MS(1,1) is implemented as four by four switch with the middle links ML(1,1), ML(1,2), ML(1,7) and ML(1,8) being the inputs and middle links ML(2,1)-ML(2,4) being the outputs; and middle switch MS(7,1) is implemented as four by four switch with the middle links ML(7,1), ML(7,2), ML(7,11) and ML(7,12) being the inputs and middle links ML(8,1)-ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in 800I of FIG. 8I.

Generalized Multi-Link Butterfly Fat Pyramid Network Embodiment:

[0276] In another embodiment in the network 800B of FIG. 8B, the switches that are placed together are implemented as combined switch then the network 800B is the generalized multi-link butterfly fat pyramid network $V_{mink-bfp}(N_1, N_2, d, s)$ where $N_1 = N_2 = 32$; $d = 2$; and $s = 2$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,390 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and

US 2012/0269190 A1

Oct. 25, 2012

27

output stage **120** are implemented as a six by six switch. For example the input switch **IS1** and output switch **OS1** are placed together; so input switch **IS1** and output **OS1** are implemented as a six by six switch with the inlet links **IL1**, **IL2**, **ML(8,1)**, **ML(8,2)**, **ML(8,7)** and **ML(8,8)** being the inputs of the combined switch (denoted as **IS1&OS1**) and middle links **ML(1,1)**, **ML(1,2)**, **ML(1,3)**, **ML(1,4)**, **OL1** and **OL2** being the outputs of the combined switch **IS1&OS1**. Similarly in this embodiment of network **800B** all the switches that are placed together are implemented as a combined switch.

[0277] Layout diagrams **800C** in FIG. **8C**, **800D** in FIG. **8D**, **800E** in FIG. **8E**, **800F** in FIG. **8G** are also applicable to generalized multi-link butterfly fat pyramid network $V_{mink-bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages. The layout **800C** in FIG. **8C** can be recursively extended for any arbitrarily large generalized multi-link butterfly fat pyramid network $V_{mink-bfp}(N_1, N_2, d, s)$. Accordingly layout **800H** of FIG. **8H** is also applicable to generalized multi-link butterfly fat pyramid network $V_{mink-bfp}(N_1, N_2, d, s)$.

[0278] Referring to diagram **800J** of FIG. **8J** illustrates a high-level implementation of Block **1_2** (Each of the other blocks have similar implementation) of the layout **800C** of FIG. **8C** which represents a generalized multi-link butterfly fat pyramid network $V_{mink-bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block **1_2** in **800J** illustrates both the intra-block and inter-block links. The layout diagram **800J** corresponds to the embodiment where the switches that are placed together are implemented as combined switch in the network **800B** of FIG. **8B**. As noted before then the network **800B** is the generalized multi-link butterfly fat pyramid network $V_{mink-bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,390 that is incorporated by reference above.

[0279] That is the switches that are placed together in Block **1_2** as shown in FIG. **8J** are namely the combined input and output switch **IS1&OS1** belonging to switch **1**, illustrated by dotted lines, (as noted before switch **1** is for illustration purposes only, in practice the switch implemented is combined input and output switch **IS1&OS1**); middle switch **MS(1,1)** belonging to switch **2**; middle switch **MS(2,1)** belonging to switch **3**; middle switch **MS(3,1)** belonging to switch **4**; And middle switch **MS(4,1)** belonging to switch **5**.

[0280] Combined input and output switch **IS1&OS1** is implemented as six by six switch with the inlet links **IL1**, **IL2** and **ML(8,1)**, **ML(8,2)**, **ML(8,7)**, and **ML(8,8)** being the inputs and middle links **ML(1,1)**-**ML(1,4)**, and outlet links **OL1**-**OL2** being the outputs.

[0281] Middle switch **MS(1,1)** is implemented as eight by eight switch with the middle links **ML(1,1)**, **ML(1,2)**, **ML(1,7)**, **ML(1,8)**, **ML(7,1)**, **ML(7,2)**, **ML(7,11)** and **ML(7,12)** being the inputs and middle links **ML(2,1)**-**ML(2,4)** and middle links **ML(8,1)**-**ML(8,4)** being the outputs. Similarly all the other middle switches are also implemented as eight by eight switches as illustrated in **800J** of FIG. **8J**.

[0282] In another embodiment, middle switch **MS(1,1)** (or the middle switches in any of the middle stage excepting the root middle stage) of Block **1_2** of $V_{mink-bfp}(N_1, N_2, d, s)$ can be implemented as a four by eight switch and a four by four switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch

MS(1,1) of Block **1_2** as shown FIG. **8J**, the left going middle links namely **ML(7,1)**, **ML(7,2)**, **ML(7,11)**, and **ML(7,12)** are never switched to the right going middle links **ML(2,1)**, **ML(2,2)**, **ML(2,3)**, and **ML(2,4)**. And hence to implement **MS(1,1)** two switches namely: 1) a four by eight switch with the middle links **ML(1,1)**, **ML(1,2)**, **ML(1,7)**, and **ML(1,8)** as inputs and the middle links **ML(2,1)**, **ML(2,2)**, **ML(2,3)**, **ML(2,4)**, **ML(8,1)**, **ML(8,2)**, **ML(8,3)**, and **ML(8,4)** as outputs and 2) a four by four switch with the middle links **ML(7,1)**, **ML(7,2)**, **ML(7,11)**, and **ML(7,12)** as inputs and the middle links **ML(8,1)**, **ML(8,2)**, **ML(8,3)**, and **ML(8,4)** as outputs are sufficient without loosing any connectivity of the embodiment of **MS(1,1)** being implemented as an eight by eight switch as described before.)

Generalized Multi-Stage Pyramid Network Embodiment:

[0283] In one embodiment, in the network **800B** of FIG. **8B**, the switches that are placed together are implemented as two separate switches in input stage **110** and output stage **120**; and as four separate switches in all the middle stages, then the network **800B** is the generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage **110** and output stage **120** are implemented as a two by four switch and a four by two switch respectively. For example the switch input switch **IS1** and output switch **OS1** are placed together; so input switch **IS1** is implemented as two by four switch with the inlet links **IL1** and **IL2** being the inputs and middle links **ML(1,1)**-**ML(1,4)** being the outputs; and output switch **OS1** is implemented as four by two switch with the middle links **ML(8,1)**, **ML(8,2)**, **ML(8,7)** and **ML(8,8)** being the inputs and outlet links **OL1**-**OL2** being the outputs.

[0284] The switches, corresponding to the middle stages that are placed together are implemented as four two by two switches. For example middle switches **MS(1,1)**, **MS(1,17)**, **MS(7,1)**, and **MS(7,17)** are placed together; so middle switch **MS(1,1)** is implemented as two by two switch with middle links **ML(1,1)** and **ML(1,7)** being the inputs and middle links **ML(2,1)** and **ML(2,3)** being the outputs; middle switch **MS(1,17)** is implemented as two by two switch with the middle links **ML(1,2)** and **ML(1,8)** being the inputs and middle links **ML(2,2)** and **ML(2,4)** being the outputs; middle switch **MS(7,1)** is implemented as two by two switch with middle links **ML(7,1)** and **ML(7,11)** being the inputs and middle links **ML(8,1)** and **ML(8,3)** being the outputs; And middle switch **MS(7,17)** is implemented as two by two switch with the middle links **ML(7,2)** and **ML(7,12)** being the inputs and middle links **ML(8,2)** and **ML(8,4)** being the outputs; Similarly in this embodiment of network **800B** all the switches that are placed together are implemented as separate switches.

[0285] Layout diagrams **800C** in FIG. **8C**, **800D** in FIG. **8D**, **800E** in FIG. **8E**, **800F** in FIG. **8G** are also applicable to generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages. The layout **800C** in FIG. **8C** can be recursively extended for any arbitrarily large generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$. Accordingly layout **800H** of FIG. **8H** is also applicable to generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$.

US 2012/0269190 A1

Oct. 25, 2012

28

[0286] Referring to diagram 800K of FIG. 8K illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 800C of FIG. 8C which represents a generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block 1_2 in 800K illustrates both the intra-block and inter-block links. The layout diagram 800K corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 800B of FIG. 8B. As noted before then the network 800B is the generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,391 that is incorporated by reference above.

[0287] That is the switches that are placed together in Block 1_2 as shown in FIG. 8K are namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches MS(1,1), MS(1,17), MS(7,1) and MS(7,17) belonging to switch 2; middle switches MS(2,1), MS(2,17), MS(6,1) and MS(6,17) belonging to switch 3; middle switches MS(3,1), MS(3,17), MS(5,1) and MS(5,17) belonging to switch 4; And middle switches MS(4,1), and MS(4,17) belonging to switch 5.

[0288] Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by four switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,4) being the outputs; and output switch OS1 is implemented as four by two switch with the middle links ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs and outlet links OL1-OL2 being the outputs.

[0289] Middle switches MS(1,1), MS(1,17), MS(7,1), and MS(7,17) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,7) being the inputs and middle links ML(2,1) and ML(2,3) being the outputs; middle switch MS(1,17) is implemented as two by two switch with the middle links ML(1,2) and ML(1,8) being the inputs and middle links ML(2,2) and ML(2,4) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,11) being the inputs and middle links ML(8,1) and ML(8,3) being the outputs; And middle switch MS(7,17) is implemented as two by two switch with the middle links ML(7,2) and ML(7,12) being the inputs and middle links ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as illustrated in 800K of FIG. 8K.

Generalized Multi-Stage Pyramid Network Embodiment with $S=1$:

[0290] In one embodiment, in the network 800B of FIG. 8B (where it is implemented with $s=1$), the switches that are placed together are implemented as two separate switches in input stage 110 and output stage 120; and as two separate switches in all the middle stages, then the network 800B is the generalized folded multi-stage network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,391 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as two, two by two switches. For example the switch input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the

inputs and middle links ML(1,1)-ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1-OL2 being the outputs.

[0291] The switches, corresponding to the middle stages that are placed together are implemented as two, two by two switches. For example middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the inputs and middle links ML(8,1) and ML(8,2) being the outputs; Similarly in this embodiment of network 800B all the switches that are placed together are implemented as two separate switches.

[0292] Layout diagrams 800C in FIG. 8C, 800D in FIG. 8D, 800E in FIG. 8E, 800F in FIG. 8G are also applicable to generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages. The layout 800C in FIG. 8C can be recursively extended for any arbitrarily large generalized folded multi stage network $V_{fold-p}(N_1, N_2, d, s)$. Accordingly layout 800H of FIG. 8H is also applicable to generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$.

[0293] Referring to diagram 800K1 of FIG. 8K1 illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) for the layout 800C of FIG. 8C when $s=1$ which represents a generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ (All the double links are replaced by single links when $s=1$). Block 1_2 in 800K1 illustrates both the intra-block and inter-block links. The layout diagram 800K1 corresponds to the embodiment where the switches that are placed together are implemented as separate switches in the network 800B of FIG. 8B when $s=1$. As noted before then the network 800B is the generalized folded multi-stage pyramid network $V_{fold-p}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,391 that is incorporated by reference above.

[0294] That is the switches that are placed together in Block 1_2 as shown in FIG. 8K1 are namely the input switch IS1 and output switch OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switches implemented are input switch IS1 and output switch OS1); middle switches MS(1,1) and MS(7,1) belonging to switch 2; middle switches MS(2,1) and MS(6,1) belonging to switch 3; middle switches MS(3,1) and MS(5,1) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0295] Input switch IS1 and output switch OS1 are placed together; so input switch IS1 is implemented as two by two switch with the inlet links IL1 and IL2 being the inputs and middle links ML(1,1)-ML(1,2) being the outputs; and output switch OS1 is implemented as two by two switch with the middle links ML(8,1) and ML(8,3) being the inputs and outlet links OL1-OL2 being the outputs.

[0296] Middle switches MS(1,1) and MS(7,1) are placed together; so middle switch MS(1,1) is implemented as two by two switch with middle links ML(1,1) and ML(1,3) being the inputs and middle links ML(2,1) and ML(2,2) being the outputs; And middle switch MS(7,1) is implemented as two by two switch with middle links ML(7,1) and ML(7,5) being the

US 2012/0269190 A1

Oct. 25, 2012

29

inputs and middle links ML(8,1) and ML(8,2) being the outputs. Similarly all the other middle switches are also implemented as two by two switches as illustrated in 800K1 of FIG. 8K1.

Generalized Butterfly Fat Pyramid Network Embodiment:

[0297] In another embodiment in the network 800B of FIG. 8B, the switches that are placed together are implemented as two combined switches then the network 800B is the generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a six by six switch. For example the input switch IS1 and output switch OS1 are placed together; so input output switch IS1&OS1 are implemented as a six by six switch with the inlet links ILL IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs of the combined switch (denoted as IS1&OS1) and middle links ML(1,1), ML(1,2), ML(1,3), ML(1,4), OL1 and OL2 being the outputs of the combined switch IS1&OS1.

[0298] The switches, corresponding to the middle stages that are placed together are implemented as two four by four switches. For example middle switches MS(1,1) and MS(1,17) are placed together; so middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; middle switch MS(1,17) is implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly in this embodiment of network 800B all the switches that are placed together are implemented as a two combined switches.

[0299] Layout diagrams 800C in FIG. 8C, 800D in FIG. 8D, 800E in FIG. 8E, 800F in FIG. 8G are also applicable to generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages. The layout 800C in FIG. 8C can be recursively extended for any arbitrarily large generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$. Accordingly layout 800H of FIG. 8H is also applicable to generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$.

[0300] Referring to diagram 800L of FIG. 8L illustrates a high-level implementation of Block 1_2 (Each of the other blocks have similar implementation) of the layout 800C of FIG. 8C which represents a generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$. Block 1_2 in 800L illustrates both the intra-block and inter-block links. The layout diagram 800L corresponds to the embodiment where the switches that are placed together are implemented as two combined switches in the network 800B of FIG. 8B. As noted before then the network 800B is the generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=2$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,387 that is incorporated by reference above.

[0301] That is the switches that are placed together in Block 1_2 as shown in FIG. 8L are namely the combined input and output switch IS1&OS1 belonging to switch 1, illustrated by dotted lines, (as noted before switch 1 is for illustration purposes only, in practice the switch implemented is combined input and output switch IS1&OS1); middle switch MS(1,1)

and MS(1,17) belonging to switch 2; middle switch MS(2,1) and MS(2,17) belonging to switch 3; middle switch MS(3,1) and MS(3,17) belonging to switch 4; And middle switch MS(4,1) belonging to switch 5.

[0302] Combined input and output switch IS1&OS1 is implemented as six by six switch with the inlet links IL1, IL2, ML(8,1), ML(8,2), ML(8,7) and ML(8,8) being the inputs and middle links ML(1,1)-ML(1,4) and outlet links OL1-OL2 being the outputs.

[0303] Middle switch MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,7), ML(7,1) and ML(7,11) being the inputs and middle links ML(2,1), ML(2,3), ML(8,1) and ML(8,3) being the outputs; And middle switch MS(1,17) is implemented as four by four switch with the middle links ML(1,2), ML(1,8), ML(7,2) and ML(7,12) being the inputs and middle links ML(2,2), ML(2,4), ML(8,2) and ML(8,4) being the outputs. Similarly all the other middle switches are also implemented as two four by four switches as illustrated in 800L of FIG. 8L.

[0304] In another embodiment, middle switch MS(1,1) (or the middle switches in any of the middle stage excepting the root middle stage) of Block 1_2 of $V_{mlink-bfp}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch MS(1,1) of Block 1_2 as shown FIG. 8L, the left going middle links namely ML(7,1) and ML(7,11) are never switched to the right going middle links ML(2,1) and ML(2,3). And hence to implement MS(1,1) two switches namely: 1) a two by four switch with the middle links ML(1,1) and ML(1,7) as inputs and the middle links ML(2,1), ML(2,3), ML(8,1), and ML(8,3) as outputs and 2) a two by two switch with the middle links ML(7,1) and ML(7,11) as inputs and the middle links ML(8,1) and ML(8,3) as outputs are sufficient without losing any connectivity of the embodiment of MS(1,1) being implemented as an eight by eight switch as described before.)

Generalized Butterfly Fat Pyramid Network Embodiment with $S=1$:

[0305] In one embodiment, in the network 800B of FIG. 8B (where it is implemented with $s=1$), the switches that are placed together are implemented as a combined switch in input stage 110 and output stage 120; and as a combined switch in all the middle stages, then the network 800B is the generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with five stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,387 that is incorporated by reference above. That is the switches that are placed together in input stage 110 and output stage 120 are implemented as a four by four switch. For example the switch input switch IS1 and output switch OS1 are placed together; so input and output switch IS1&OS1 is implemented as four by four switch with the inlet links ILL IL2, ML(8,1) and ML(8,3) being the inputs and middle links ML(1,1)-ML(1,2) and outlet links OL1-OL2 being the outputs

[0306] The switches, corresponding to the middle stages that are placed together are implemented as a four by four switch. For example middle switches MS(1,1) is implemented as four by four switch with middle links ML(1,1), ML(1,3), ML(7,1) and ML(7,5) being the inputs and middle links ML(2,1), ML(2,2), ML(8,1) and ML(8,2) being the outputs.

US 2012/0269190 A1

Oct. 25, 2012

30

[0307] Layout diagrams **800C** in FIG. **8C**, **800D** in FIG. **8D**, **800E** in FIG. **8E**, **800F** in FIG. **8G** are also applicable to generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with five stages. The layout **800C** in FIG. **8C** can be recursively extended for any arbitrarily large generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$. Accordingly layout **800H** of FIG. **8H** is also applicable to generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$.

[0308] Referring to diagram **800L1** of FIG. **8L1** illustrates a high-level implementation of Block **1_2** (Each of the other blocks have similar implementation) for the layout **800C** of FIG. **8C** when $s=1$ which represents a generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ (All the double links are replaced by single links when $s=1$). Block **1_2** in **800K1** illustrates both the intra-block and inter-block links. The layout diagram **800L1** corresponds to the embodiment where the switches that are placed together are implemented as a combined switch in the network **800B** of FIG. **8B** when $s=1$. As noted before then the network **800B** is the generalized butterfly fat pyramid network $V_{bfp}(N_1, N_2, d, s)$ where $N_1=N_2=32$; $d=2$; and $s=1$ with nine stages as disclosed in U.S. Provisional Patent Application Ser. No. 60/940,387 that is incorporated by reference above.

[0309] That is the switches that are placed together in Block **1_2** as shown in FIG. **8L1** are namely the input and output switch **IS1&OS1** belonging to switch **1**, illustrated by dotted lines, (as noted before switch **1** is for illustration purposes only, in practice the switches implemented are input switch **IS1** and output switch **OS1**); middle switch **MS(1,1)** belonging to switch **2**; middle switch **MS(2,1)** belonging to switch **3**; middle switch **MS(3,1)** belonging to switch **4**; And middle switch **MS(4,1)** belonging to switch **5**.

[0310] Input and output switch **IS1&OS1** are placed together; so input and output switch **IS1&OS1** is implemented as four by four switch with the inlet links **ILL IL2**, **ML(8,1)** and **ML(8,3)** being the inputs and middle links **ML(1,1)-ML(1,2)** and outlet links **OL1-OL2** being the outputs.

[0311] Middle switch **MS(1,1)** is implemented as four by four switch with middle links **ML(1,1)**, **ML(1,3)**, **ML(7,1)** and **ML(7,5)** being the inputs and middle links **ML(2,1)**, **ML(2,2)**, **ML(8,1)** and **ML(8,2)** being the outputs. Similarly all the other middle switches are also implemented as four by four switches as illustrated in **800L1** of FIG. **8L1**.

[0312] In another embodiment, middle switch **MS(1,1)** (or the middle switches in any of the middle stage excepting the root middle stage) of Block **1_2** of $V_{mlink-bfp}(N_1, N_2, d, s)$ can be implemented as a two by four switch and a two by two switch to save cross points. This is because the left going middle links of these middle switches are never setup to the right going middle links. For example, in middle switch **MS(1,1)** of Block **1_2** as shown FIG. **8L1**, the left going middle links namely **ML(7,1)** and **ML(7,5)** are never switched to the right going middle links **ML(2,1)** and **ML(2,2)**. And hence to implement **MS(1,1)** two switches namely: 1) a two by four switch with the middle links **ML(1,1)** and **ML(1,3)** as inputs and the middle links **ML(2,1)**, **ML(2,2)**, **ML(8,1)**, and **ML(8,2)** as outputs and 2) a two by two switch with the middle links **ML(7,1)** and **ML(7,5)** as inputs and the middle links **ML(8,1)** and **ML(8,2)** as outputs are sufficient

without loosing any connectivity of the embodiment of **MS(1,1)** being implemented as an eight by eight switch as described before.)

[0313] All the layout embodiments disclosed in the current invention are applicable to generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage pyramid networks $V_{fold-mlink-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$, and generalized hypercube networks $V_{CCC}(N_1, N_2, d, s)$ for $s=1,2,3$ or any number in general, and for both $N_1=N_2=N$. and $N_1 \neq N_2$, and d is any integer.

[0314] Conversely applicant makes another important observation that generalized cube connected cycles networks $V_{CCC}(N_1, N_2, d, s)$ are implemented with the layout topology being the hypercube topology shown in layout **200C** of FIG. **2C** with large scale cross point reduction as any one of the networks described in the current invention namely: generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage pyramid networks $V_{fold-mlink-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$ for $s=1,2,3$ or any number in general, and for both $N_1=N_2=N$. and $N_1 \neq N_2$, and d is any integer.

[0315] Applicant notes that in the generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$, generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$, generalized folded multi-link multi-stage pyramid networks $V_{fold-mlink-p}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$ and generalized hypercube networks $V_{CCC}(N_1, N_2, d, s)$ the pyramid links are provided a) between the switches in any two successive stages, b) between the switches in the same stage, and c) between the switches any two arbitrary stages.

[0316] In all the embodiments disclosed in the current invention, all the switches in some embodiments may be implemented as active switches consisting of cross points using SRAM cells or Flash memory cells. Similarly in other embodiments the switches may be implemented as passive switches consisting of cross points using anti-fuse based vias or connections provided by metal layer programming as in structured ASICs. In another embodiment, the switches may be implemented as in 3D-FPGAs. In another embodiment where ASIC placement & routing, the switches are actually used to determine if two wires are connected together or not; Alternatively they can be seen as switches during the implementation of the placement & routing however cross points in the cross state can be used as wire connections and in the bar state can be used as no connection of the wires.

Scheduling Method Embodiments for Multi-Stage Pyramid Networks and Multi-Link Multi-Stage Pyramid Networks:

[0317] FIG. **9A** shows a high-level flowchart of a scheduling method **900**, in one embodiment executed to setup multicast and unicast connections in the generalized multi-link multi-stage pyramid networks $V_{mlink-p}(N_1, N_2, d, s)$ (for

US 2012/0269190 A1

Oct. 25, 2012

31

example the network **800A** of FIG. **8A**) or generalized folded multi-stage pyramid networks $V_{fold-mlink-p}(N_1, N_2, d, s)$ (for example the network **800B** of FIG. **8B**) or any of the generalized multi-stage pyramid networks $V_p(N_1, N_2, d, s)$, generalized folded multi-stage pyramid networks $V_{fold-p}(N_1, N_2, d, s)$ disclosed in this invention. According to this embodiment, a multicast connection request is received in act **910**. Then the control goes to act **920**.

[**0318**] In act **920**, based on the inlet link and input switch of the multicast connection received in act **910**, from each available outgoing middle link of the input switch of the multicast connection, by traveling forward from middle stage **130** to middle stage $130+10*(\text{Log}_d N-2)$, the lists of all reachable middle switches in each middle stage are derived recursively. That is, first, by following each available outgoing middle link of the input switch all the reachable middle switches in middle stage **130** are derived. Next, starting from the selected middle switches in middle stage **130** traveling through all of their available outgoing middle links to middle stage **140** all the available middle switches in middle stage **140** are derived. This process is repeated recursively until all the reachable middle switches, starting from the outgoing middle link of input switch, in middle stage $130+10*(\text{Log}_d N-2)$ are derived. This process is repeated for each available outgoing middle link from the input switch of the multicast connection and separate reachable lists are derived in each middle stage from middle stage **130** to middle stage $130+10*(\text{Log}_d N-2)$ for all the available outgoing middle links from the input switch. Then the control goes to act **930**.

[**0319**] In act **930**, based on the destinations of the multicast connection received in act **910**, from the output switch of each destination, by traveling backward from output stage **120** to middle stage $130+10*(\text{Log}_d N-2)$, the lists of all middle switches in each middle stage from which each destination output switch (and hence the destination outlet links) is reachable, are derived recursively. That is, first, by following each available incoming middle link of the output switch of each destination link of the multicast connection, all the middle switches in middle stage $130+10*(2*\text{Log}_d N-4)$ from which the output switch is reachable, are derived. Next, starting from the selected middle switches in middle stage $130+10*(2*\text{Log}_d N-4)$ traveling backward through all of their available incoming middle links from middle stage $130+10*(2*\text{Log}_d N-5)$ all the available middle switches in middle stage $130+10*(2*\text{Log}_d N-5)$ from which the output switch is reachable, are derived. This process is repeated recursively until all the middle switches in middle stage $130+10*(\text{Log}_d N-2)$ from which the output switch is reachable, are derived. This process is repeated for each output switch of each destination link of the multicast connection and separate lists in each middle stage from middle stage $130+10*(2*\text{Log}_d N-4)$ to middle stage $130+10*(\text{Log}_d N-2)$ for all the output switches of each destination link of the connection are derived. Then the control goes to act **940**.

[**0320**] In act **940**, using the lists generated in acts **920** and **930**, particularly list of middle switches derived in middle stage $130+10*(\text{Log}_d N-2)$ corresponding to each outgoing link of the input switch of the multicast connection, and the list of middle switches derived in middle stage $130+10*(\text{Log}_d N-2)$ corresponding to each output switch of the destination links, the list of all the reachable destination links from each outgoing link of the input switch are derived. Specifically if a middle switch in middle stage $130+10*(\text{Log}_d N-2)$ is reachable from an outgoing link of the input switch, say “x”,

and also from the same middle switch in middle stage $130+10*(\text{Log}_d N-2)$ if the output switch of a destination link, say “y”, is reachable then using the outgoing link of the input switch x, destination link y is reachable. Accordingly, the list of all the reachable destination links from each outgoing link of the input switch is derived. The control then goes to act **950**.

[**0321**] In act **950**, among all the outgoing links of the input switch, it is checked if all the destinations are reachable using only one outgoing link of the input switch. If one outgoing link is available through which all the destinations of the multicast connection are reachable (i.e., act **950** results in “yes”), the control goes to act **970**. And in act **970**, the multicast connection is setup by traversing from the selected only one outgoing middle link of the input switch in act **950**, to all the destinations. Then the control transfers to act **990**.

[**0322**] If act **950** results “no”, that is one outgoing link is not available through which all the destinations of the multicast connection are reachable, then the control goes to act **960**. In act **960**, it is checked if all destination links of the multicast connection are reachable using two outgoing middle links from the input switch. According to the current invention, it is always possible to find at most two outgoing middle links from the input switch through which all the destinations of a multicast connection are reachable. So act **960** always results in “yes”, and then the control transfers to act **980**. In act **980**, the multicast connection is setup by traversing from the selected only two outgoing middle links of the input switch in act **960**, to all the destinations. Then the control transfers to act **990**.

[**0323**] In act **990**, all the middle links between any two stages of the network used to setup the connection in either act **970** or act **980** are marked unavailable so that these middle links will be made unavailable to other multicast connections. The control then returns to act **910**, so that acts **910**, **920**, **930**, **940**, **950**, **960**, **970**, **980**, and **990** are executed in a loop, for each connection request until the connections are set up.

[**0324**] In the example illustrated in FIG. **8A**, four outgoing middle links are available to satisfy a multicast connection request if input switch is IS2, but only at most two outgoing middle links of the input switch will be used in accordance with this method. Similarly, although three outgoing middle links is available for a multicast connection request if the input switch is IS1, again only at most two outgoing middle links is used. The specific outgoing middle links of the input switch that are chosen when selecting two outgoing middle links of the input switch is irrelevant to the method of FIG. **9A** so long as at most two outgoing middle links of the input switch are selected to ensure that the connection request is satisfied, i.e. the destination switches identified by the connection request can be reached from the outgoing middle links of the input switch that are selected. In essence, limiting the outgoing middle links of the input switch to no more than two permits the network $V(N_1, N_2, d, s)$ to be operated in nonblocking manner in accordance with the invention.

[**0325**] According to the current invention, using the method **940** of FIG. **9A**, the network $V_p(N_1, N_2, d, s)$ or $V_{mink-p}(N_1, N_2, d, s)$ is operated in rearrangeably nonblocking for unicast connections when $s \geq 1$, is operated in strictly nonblocking for unicast connections when $s \geq 2$, is operated in rearrangeably nonblocking for multicast connections when $s \geq 2$, and is operated in strictly nonblocking for multicast connections when $s \geq 3$.

[0326] The connection request of the type described above in reference to method 900 of FIG. 9A can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, only one outgoing middle link of the input switch is used to satisfy the request. Moreover, in method 900 described above in reference to FIG. 9A any number of middle links may be used between any two stages excepting between the input stage and middle stage 130, and also any arbitrary fan-out may be used within each output stage switch, to satisfy the connection request.

[0327] As noted above method 900 of FIG. 9A can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks $V_p(N, d, s)$, $V_{mlink-p}(N, d, s)$, $V_p(N_1, N_2, d, s)$ or $V_{mlink-p}(N_1, N_2, d, s)$ disclosed in this invention.

Scheduling Method Embodiments for Butterfly Fat Pyramid Networks and Multi-Link Butterfly Fat Pyramid Networks:

[0328] FIG. 10A shows a high-level flowchart of a scheduling method 1000, in one embodiment executed to setup multicast and unicast connections in the generalized butterfly fat pyramid networks $V_{bfp}(N_1, N_2, d, s)$, generalized folded butterfly fat pyramid networks $V_{fold-bfp}(N_1, N_2, d, s)$, generalized multi-link butterfly fat pyramid networks $V_{mlink-bfp}(N_1, N_2, d, s)$ or generalized folded multi-link butterfly fat pyramid networks $V_{fold-mlink-bfp}(N_1, N_2, d, s)$ disclosed in this invention. According to this embodiment, a multicast connection request is received in act 1010. Then the control goes to act 1020.

[0329] In act 1020, based on the inlet link and input switch of the multicast connection received in act 1010, from each available outgoing middle link of the input switch of the multicast connection, by traveling forward from middle stage 130 to middle stage $130+10*(\text{Log}_d N-2)$, the lists of all reachable middle switches in each middle stage are derived recursively. That is, first, by following each available outgoing middle link of the input switch all the reachable middle switches in middle stage 130 are derived. Next, starting from the selected middle switches in middle stage 130 traveling through all of their available outgoing middle links to middle stage 140 (reverse links from middle stage 130 to output stage 120 are ignored) all the available middle switches in middle stage 140 are derived. (In the traversal from any middle stage to the following middle stage only upward links are used and no reverse links or downward links are used. That is for example, while deriving the list of available middle switches in middle stage 140, the reverse links going from middle stage 130 to output stage 120 are ignored.) This process is repeated recursively until all the reachable middle switches, starting from the outgoing middle link of input switch, in middle stage $130+10*(\text{Log}_d N-2)$ are derived. This process is repeated for each available outgoing middle link from the input switch of the multicast connection and separate reachable lists are derived in each middle stage from middle stage 130 to middle stage $130+10*(\text{Log}_d N-2)$ for all the available outgoing middle links from the input switch. Then the control goes to act 1030.

[0330] In act 1030, based on the destinations of the multicast connection received in act 1010, from the output switch of each destination, by traveling backward from output stage 120 to middle stage $130+10*(\text{Log}_d N-2)$, the lists of all middle switches in each middle stage from which each destination output switch (and hence the destination outlet links)

is reachable, are derived recursively. That is, first, by following each available incoming middle link of the output switch of each destination link of the multicast connection, all the middle switches in middle stage 130 from which the output switch is reachable, are derived. Next, starting from the selected middle switches in middle stage 130 traveling backward through all of their available incoming middle links from middle stage 140 all the available middle switches in middle stage 140 (reverse links from middle stage 130 to input stage 120 are ignored) from which the output switch is reachable, are derived. (In the traversal from any middle stage to the following middle stage only upward links are used and no reverse links or downward links are used. That is for example, while deriving the list of available middle switches in middle stage 140, the reverse links coming to middle stage 130 from input stage 110 are ignored.) This process is repeated recursively until all the middle switches in middle stage $130+10*(\text{Log}_d N-2)$ from which the output switch is reachable, are derived. This process is repeated for each output switch of each destination link of the multicast connection and separate lists in each middle stage from middle stage 130 to middle stage $130+10*(\text{Log}_d N-2)$ for all the output switches of each destination link of the connection are derived. Then the control goes to act 1040.

[0331] In act 1040, using the lists generated in acts 1020 and 1030, particularly list of middle switches derived in middle stage $130+10*(\text{Log}_d N-2)$ corresponding to each outgoing link of the input switch of the multicast connection, and the list of middle switches derived in middle stage $130+10*(\text{Log}_d N-2)$ corresponding to each output switch of the destination links, the list of all the reachable destination links from each outgoing link of the input switch are derived. Specifically if a middle switch in middle stage $130+10*(\text{Log}_d N-2)$ is reachable from an outgoing link of the input switch, say "x", and also from the same middle switch in middle stage $130+10*(\text{Log}_d N-2)$ if the output switch of a destination link, say "y", is reachable then using the outgoing link of the input switch x, destination link y is reachable. Accordingly, the list of all the reachable destination links from each outgoing link of the input switch is derived. The control then goes to act 1050.

[0332] In act 1050, among all the outgoing links of the input switch, it is checked if all the destinations are reachable using only one outgoing link of the input switch. If one outgoing link is available through which all the destinations of the multicast connection are reachable (i.e., act 1050 results in "yes"), the control goes to act 1070. And in act 1070, the multicast connection is setup by traversing from the selected only one outgoing middle link of the input switch in act 1050, to all the destinations. Also the nearest U-turn is taken while setting up the connection. That is at any middle stage if one of the middle switch in the lists derived in acts 1020 and 1030 are common then the connection is setup so that the U-turn is made to setup the connection from that middle switch for all the destination links reachable from that common middle switch. Then the control transfers to act 1090.

[0333] If act 1050 results "no", that is one outgoing link is not available through which all the destinations of the multicast connection are reachable, then the control goes to act 1060. In act 1060, it is checked if all destination links of the multicast connection are reachable using two outgoing middle links from the input switch. According to the current invention, it is always possible to find at most two outgoing middle links from the input switch through which all the

US 2012/0269190 A1

Oct. 25, 2012

33

destinations of a multicast connection are reachable. So act **1060** always results in “yes”, and then the control transfers to act **1080**. In act **1080**, the multicast connection is setup by traversing from the selected only two outgoing middle links of the input switch in act **1060**, to all the destinations. Also the nearest U-turn is taken while setting up the connection. That is at any middle stage if one of the middle switch in the lists derived in acts **1020** and **1030** are common then the connection is setup so that the U-turn is made to setup the connection from that middle switch for all the destination links reachable from that common middle switch. Then the control transfers to act **1090**.

[0334] In act **1090**, all the middle links between any two stages of the network used to setup the connection in either act **1070** or act **1080** are marked unavailable so that these middle links will be made unavailable to other multicast connections. The control then returns to act **1010**, so that acts **1010**, **1020**, **1030**, **1040**, **1050**, **1060**, **1070**, **1080**, and **1090** are executed in a loop, for each connection request until the connections are set up.

[0335] According to the current invention, using the method **1040** of FIG. **10A**, the network $V_{bfp}(N_1, N_2, d, s)$ or $V_{mlink-bfp}(N_1, N_2, d, s)$ is operated in rearrangeably nonblocking for unicast connections when $s \geq 1$, is operated in strictly nonblocking for unicast connections when $s \geq 2$, is operated in rearrangeably nonblocking for multicast connections when $s \geq 2$, and is operated in strictly nonblocking for multicast connections when $s \geq 3$.

[0336] The connection request of the type described above in reference to method **1000** of FIG. **10A** can be unicast connection request, a multicast connection request or a broadcast connection request, depending on the example. In case of a unicast connection request, only one outgoing middle link of the input switch is used to satisfy the request. Moreover, in method **1000** described above in reference to FIG. **10A** any number of middle links may be used between any two stages excepting between the input stage and middle stage **130**, and also any arbitrary fan-out may be used within each output stage switch, to satisfy the connection request.

[0337] As noted above method **1000** of FIG. **10A** can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks $V_{bfp}(N, d, s)$, $V_{mlink-bfp}(N, d, s)$, $V_{bfp}(N_1, N_2, d, s)$ or $V_{mlink-bfp}(N_1, N_2, d, s)$ disclosed in this invention.

Applications Embodiments:

[0338] All the embodiments disclosed in the current invention are useful in many varieties of applications. FIG. **11A1** illustrates the diagram of **1100A1** which is a typical two by two switch with two inlet links namely **IL1** and **IL2**, and two outlet links namely **OL1** and **OL2**. The two by two switch also implements four crosspoints namely **CP(1,1)**, **CP(1,2)**, **CP(2,1)** and **CP(2,2)** as illustrated in FIG. **11A1**. For example the diagram of **1100A1** may be the implementation of middle switch **MS(1,1)** of the diagram **100K** of FIG. **1K** where inlet link **IL1** of diagram **1100A1** corresponds to middle link **ML(1,1)** of diagram **100K**, inlet link **IL2** of diagram **1100A1** corresponds to middle link **ML(1,7)** of diagram **100K**, outlet link **OL1** of diagram **1100A1** corresponds to middle link **ML(2,1)** of diagram **100K**, outlet link **OL2** of diagram **1100A1** corresponds to middle link **ML(2,3)** of diagram **100K**.

1) Programmable Integrated Circuit Embodiments:

[0339] All the embodiments disclosed in the current invention are useful in programmable integrated circuit applica-

tions. FIG. **11A2** illustrates the detailed diagram **1100A2** for the implementation of the diagram **1100A1** in programmable integrated circuit embodiments. Each crosspoint is implemented by a transistor coupled between the corresponding inlet link and outlet link, and a programmable cell in programmable integrated circuit embodiments. Specifically crosspoint **CP(1,1)** is implemented by transistor **C(1,1)** coupled between inlet link **IL1** and outlet link **OL1**, and programmable cell **P(1,1)**; crosspoint **CP(1,2)** is implemented by transistor **C(1,2)** coupled between inlet link **IL1** and outlet link **OL2**, and programmable cell **P(1,2)**; crosspoint **CP(2,1)** is implemented by transistor **C(2,1)** coupled between inlet link **IL2** and outlet link **OL1**, and programmable cell **P(2,1)**; and crosspoint **CP(2,2)** is implemented by transistor **C(2,2)** coupled between inlet link **IL2** and outlet link **OL2**, and programmable cell **P(2,2)**.

[0340] If the programmable cell is programmed ON, the corresponding transistor couples the corresponding inlet link and outlet link. If the programmable cell is programmed OFF, the corresponding inlet link and outlet link are not connected. For example if the programmable cell **P(1,1)** is programmed ON, the corresponding transistor **C(1,1)** couples the corresponding inlet link **IL1** and outlet link **OL1**. If the programmable cell **P(1,1)** is programmed OFF, the corresponding inlet link **IL1** and outlet link **OL1** are not connected. In volatile programmable integrated circuit embodiments the programmable cell may be an SRAM (Static Random Access Memory) cell. In non-volatile programmable integrated circuit embodiments the programmable cell may be a Flash memory cell. Also the programmable integrated circuit embodiments may implement field programmable logic arrays (FPGA) devices, or programmable Logic devices (PLD), or Application Specific Integrated Circuits (ASIC) embedded with programmable logic circuits or 3D-FPGAs.

[0341] FIG. **11A2** also illustrates a buffer **B1** on inlet link **IL2**. The signals driven along inlet link **IL2** are amplified by buffer **B1**. Buffer **B1** can be inverting or non-inverting buffer. Buffers such as **B1** are used to amplify the signal in links which are usually long.

2) One-Time Programmable Integrated Circuit Embodiments:

[0342] All the embodiments disclosed in the current invention are useful in one-time programmable integrated circuit applications. FIG. **11A3** illustrates the detailed diagram **1100A3** for the implementation of the diagram **1100A1** in one-time programmable integrated circuit embodiments. Each crosspoint is implemented by a via coupled between the corresponding inlet link and outlet link in one-time programmable integrated circuit embodiments. Specifically crosspoint **CP(1,1)** is implemented by via **V(1,1)** coupled between inlet link **IL1** and outlet link **OL1**; crosspoint **CP(1,2)** is implemented by via **V(1,2)** coupled between inlet link **IL1** and outlet link **OL2**; crosspoint **CP(2,1)** is implemented by via **V(2,1)** coupled between inlet link **IL2** and outlet link **OL1**; and crosspoint **CP(2,2)** is implemented by via **V(2,2)** coupled between inlet link **IL2** and outlet link **OL2**.

[0343] If the via is programmed ON, the corresponding inlet link and outlet link are permanently connected which is denoted by thick circle at the intersection of inlet link and outlet link. If the via is programmed OFF, the corresponding inlet link and outlet link are not connected which is denoted by the absence of thick circle at the intersection of inlet link and outlet link. For example in the diagram **1100A3** the via

US 2012/0269190 A1

Oct. 25, 2012

34

V(1,1) is programmed ON, and the corresponding inlet link IL1 and outlet link OL1 are connected as denoted by thick circle at the intersection of inlet link IL1 and outlet link OL1; the via V(2,2) is programmed ON, and the corresponding inlet link IL2 and outlet link OL2 are connected as denoted by thick circle at the intersection of inlet link IL2 and outlet link OL2; the via V(1,2) is programmed OFF, and the corresponding inlet link IL1 and outlet link OL2 are not connected as denoted by the absence of thick circle at the intersection of inlet link IL1 and outlet link OL2; the via V(2,1) is programmed OFF, and the corresponding inlet link IL2 and outlet link OL1 are not connected as denoted by the absence of thick circle at the intersection of inlet link IL2 and outlet link OL1. One-time programmable integrated circuit embodiments may be anti-fuse based programmable integrated circuit devices or mask programmable structured ASIC devices.

3) Integrated Circuit Placement and Route Embodiments:

[0344] All the embodiments disclosed in the current invention are useful in Integrated Circuit Placement and Route applications, for example in ASIC backend Placement and Route tools. FIG. 11A4 illustrates the detailed diagram 1100A4 for the implementation of the diagram 1100A1 in Integrated Circuit Placement and Route embodiments. In an integrated circuit since the connections are known a-priori, the switch and crosspoints are actually virtual. However the concept of virtual switch and virtual crosspoint using the embodiments disclosed in the current invention reduces the number of required wires, wire length needed to connect the inputs and outputs of different netlists and the time required by the tool for placement and route of netlists in the integrated circuit.

[0345] Each virtual crosspoint is used to either to hardwire or provide no connectivity between the corresponding inlet link and outlet link. Specifically crosspoint CP(1,1) is implemented by direct connect point DCP(1,1) to hardwire (i.e., to permanently connect) inlet link IL1 and outlet link OL1 which is denoted by the thick circle at the intersection of inlet link IL1 and outlet link OL1; crosspoint CP(2,2) is implemented by direct connect point DCP(2,2) to hardwire inlet link IL2 and outlet link OL2 which is denoted by the thick circle at the intersection of inlet link IL2 and outlet link OL2. The diagram 1100A4 does not show direct connect point DCP(1,2) and direct connect point DCP(1,3) since they are not needed and in the hardware implementation they are eliminated. Alternatively inlet link IL1 needs to be connected to outlet link OL1 and inlet link IL1 does not need to be connected to outlet link OL2. Also inlet link IL2 needs to be connected to outlet link OL2 and inlet link IL2 does not need to be connected to outlet link OL1. Furthermore in the example of the diagram 1100A4, there is no need to drive the signal of inlet link IL1 horizontally beyond outlet link OL1 and hence the inlet link IL1 is not even extended horizontally until the outlet link OL2. Also the absence of direct connect point DCP(2,1) illustrates there is no need to connect inlet link IL2 and outlet link OL1.

[0346] In summary in integrated circuit placement and route tools, the concept of virtual switches and virtual cross points is used during the implementation of the placement & routing algorithmically in software, however during the hardware implementation cross points in the cross state are implemented as hardwired connections between the corresponding

inlet link and outlet link, and in the bar state are implemented as no connection between inlet link and outlet link.

3) More Application Embodiments:

[0347] All the embodiments disclosed in the current invention are also useful in the design of SoC interconnects, Field programmable interconnect chips, parallel computer systems and in time-space-time switches.

[0348] Numerous modifications and adaptations of the embodiments, implementations, and examples described herein will be apparent to the skilled artisan in view of the disclosure.

What is claimed is:

1. A two-dimensional layout of hierarchical routing network comprising a total of $a \times b$ blocks with one side of said layout having the size of a blocks and the other side of said layout having the size of b blocks where $a \geq 1$ and $b \geq 1$, and

Said routing network comprising a total of N_1 inlet links and a total of N_2 outlet links and y hierarchical stages wherein either

$N_2 = N_1 \times d_2$, $N_1 = (a \times b) \times d$, and said each block comprising d inlet links and $d \times d_2$ outlet links; or

$N_1 = N_2 \times d_1$, $N_2 = (a \times b) \times d$, and said each block comprising d outlet links and $d \times d_1$ inlet links, and

Said each stage comprising a switch of size $d \times d$, where $d \geq 2$ and each said switch of size $d \times d$ having d incoming links and d outgoing links; and

Said incoming links and outgoing links in each switch in said each stage of said each block comprising a plurality of forward connecting links connecting from switches in lower stage to switches in the immediate succeeding higher stage, and also comprising a plurality of backward connecting links connecting from switches in higher stage to switches in the immediate preceding lower stage; and

Said forward connecting links comprising a plurality of straight links connecting from a switch in a stage in a block to a switch in another stage in the same block and also comprising a plurality of cross links connecting from a switch in a stage in a block to a switch in another stage in a different block, and

Said backward connecting links comprising a plurality of straight links connecting from a switch in a stage in a block to a switch in another stage in the same block and also comprising a plurality of cross links connecting from a switch in a stage in a block to a switch in another stage in a different block.

2. The two-dimensional layout of claim 1, wherein said all cross links are connecting as either vertical or horizontal links between switches in two different said blocks.

3. The two-dimensional layout claim 2, wherein said cross links in succeeding stages are connecting as alternative vertical and horizontal links between switches in said blocks.

4. The two-dimensional layout of claim 3, wherein either said cross links from switches in a stage in one of said blocks are connecting to switches in the succeeding stage in another of said blocks so that said cross links are either vertical links or horizontal links and vice versa, and hereinafter such cross links are "shuffle exchange links").

5. The two-dimensional layout of claim 4, wherein said all horizontal shuffle exchange links are connected, in each said stage, between two sets of neighboring blocks with each said set having neighboring blocks of size 2^x where $0 \leq x \leq y$ such that $x=0$ for the lowest stage and $x=y$ for the highest stage and

US 2012/0269190 A1

Oct. 25, 2012

35

each block in one of the said sets is connected to at least one block in said second set excepting when the number of blocks in one of the sets are less than the number of blocks in said second set, and

said all vertical shuffle exchange links are connected, in each said stage, between two sets of neighboring blocks with each said set having neighboring blocks of size 2^x where $0 \leq x \leq y$ such that $x=0$ for the lowest stage and $x=y$ for the highest stage and each block in one of the said sets is connected to at least one block in said second set excepting when the number of blocks in one of the sets are less than the number of blocks in said second set.

6. The two-dimensional layout of claim 5, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are substantially of equal length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are substantially of equal length in the entire said two-dimensional layout, and

the shortest horizontal shuffle exchange links are connecting at the lowest stage and between switches in two nearest neighboring said blocks, and length of the horizontal shuffle exchange links is doubled in each succeeding stage; and the shortest vertical shuffle exchange links are connecting at the lowest stage and between switches in two nearest neighboring said blocks, and length of the vertical shuffle exchange links is doubled in each succeeding stage.

7. The two-dimensional layout of claim 5, wherein said all horizontal shuffle exchange links between switches of a block one in each said two sets of blocks in any two corresponding said succeeding stages are connected so that the two nearest neighbors are connected first and then the two nearest neighbors are connected in the remaining blocks which is repeated until the switches in all the blocks are connected, and said all vertical shuffle exchange links between switches of a block one in each said two sets of blocks in any two corresponding said succeeding stages are connected so that the two nearest neighbors are connected first and then the two nearest neighbors are connected in the remaining blocks which is repeated until the switches in all the blocks are connected in the entire said two-dimensional layout.

8. The two-dimensional layout of claim 6, wherein $y \geq (\log_2(N_1))$ when $N_2 = N_1 \times d_2$, or $y \geq (\log_2(N_2))$ when $N_1 = N_2 \times d_2$ so that the length of the horizontal shuffle exchange links in the highest stage is equal to half the size of the horizontal size of said two dimensional grid of blocks, and the length of the vertical shuffle exchange links in the highest stage is equal to half the size of the vertical size of said two dimensional grid of blocks.

9. The two-dimensional layout of claim 8, wherein $d=2$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast Benes network with full bandwidth.

10. The two-dimensional layout of claim 8, wherein $d=2$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast Benes network

and rearrangeably nonblocking for arbitrary fan-out multicast Benes network with full bandwidth.

11. The two-dimensional layout of claim 8, wherein $d=2$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast Benes network with full bandwidth.

12. The two-dimensional layout of claim 6, wherein $y \geq (\log_2(N_1))$ when $N_2 = N_1 \times d_2$, or $y \geq (\log_2(N_2))$ when $N_1 = N_2 \times d_2$ so that the length of the horizontal shuffle exchange links in the highest stage is equal to half the size of the horizontal size of said two dimensional grid of blocks and the length of the vertical shuffle exchange links in the highest stage is equal to half the size of the vertical size of said two dimensional grid of blocks, and

said each block further comprising a plurality of U-turn links within switches in each of said stages in each of said blocks.

13. The two-dimensional layout of claim 12, wherein $d=2$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast butterfly fat tree network with full bandwidth.

14. The two-dimensional layout of claim 12, wherein $d=2$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast butterfly fat tree network and rearrangeably nonblocking for arbitrary fan-out multicast butterfly fat tree network with full bandwidth.

15. The two-dimensional layout of claim 12, wherein $d=2$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast butterfly fat tree network with full bandwidth.

16. The two-dimensional layout of claim 1, wherein said horizontal and vertical links are implemented on two or more metal layers.

17. The two-dimensional layout of claim 1, wherein said switches comprising active and reprogrammable cross points and said each cross point is programmable by an SRAM cell or a Flash Cell.

18. The two-dimensional layout of claim 1, wherein said blocks are of equal size.

19. The two-dimensional layout of claim 16, wherein said blocks further comprising Lookup Tables (hereinafter "LUTs") having outlet links connected to inlet links of said routing network and further having inlet links connected from outlet links of said routing network and said two-dimensional layout with said blocks of LUTs is a field programmable gate array (FPGA) integrated circuit device or field programmable gate array (FPGA) block embedded in another integrated circuit device.

20. The two-dimensional layout of claim 16, wherein said blocks further comprising AND or OR gates having outlet links connected to inlet links of said routing network and further having inlet links connected from outlet links of said

US 2012/0269190 A1

Oct. 25, 2012

36

routing network and said two-dimensional layout with said blocks of AND or OR gates is a programmable logic device (PLD).

21. The two-dimensional layout of claim 1, wherein said blocks further comprising any arbitrary hardware logic or memory circuits.

22. The two-dimensional layout of claim 1, wherein said switches comprising active one-time programmable cross points.

23. The two-dimensional layout of claim 22, wherein said blocks further comprising Lookup Tables (hereinafter "LUTs") having outlet links connected to inlet links of said routing network and further having inlet links connected from outlet links of said routing network and said two-dimensional layout with said blocks of LUTs is a mask programmable gate array (MPGA) device or a structured ASIC device.

24. The two-dimensional layout of claim 1, wherein said switches comprising passive cross points or just connection of two links or not and said blocks further comprising any arbitrary hardware logic or memory circuits having outlet links connected to inlet links of said routing network and further having inlet links connected from outlet links of said routing network and said two-dimensional layout with said blocks of arbitrary hardware logic or memory circuits is an Application Specific Integrated Circuit (ASIC) device.

25. The two-dimensional layout of claim 1, wherein said blocks further recursively comprise one or more sub-blocks and a sub-routing network.

26. The two-dimensional layout of claim 4, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and $y \geq (\log_2(N_1))$ when $N_2 = N_1 \times d_2$, or $y \geq (\log_2(N_2))$ when $N_1 = N_2 \times d_2$.

27. The two-dimensional layout of claim 26, wherein $d=2$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-stage network with full bandwidth.

28. The two-dimensional layout of claim 26, wherein $d=2$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized multi-stage network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-stage network with full bandwidth.

29. The two-dimensional layout of claim 26, wherein $d=2$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-stage network with full bandwidth.

30. The two-dimensional layout of claim 4, wherein said all horizontal shuffle exchange links between switches in any two corresponding said succeeding stages are of different length and said vertical shuffle exchange links between switches in any two corresponding said succeeding stages are

of different length and $y \geq (\log_2(N_1))$ when $N_2 = N_1 \times d_2$, or $y \geq (\log_2(N_2))$ when $N_1 = N_2 \times d_2$, and

said each block further comprising a plurality of U-turn links within switches in each of said stages in each of said blocks.

31. The two-dimensional layout of claim 30, wherein $d=2$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized butterfly fat tree network with full bandwidth.

32. The two-dimensional layout of claim 30, wherein $d=2$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized butterfly fat tree Network and rearrangeably nonblocking for arbitrary fan-out multicast generalized butterfly fat tree network with full bandwidth.

33. The two-dimensional layout of claim 30, wherein $d=2$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized butterfly fat tree network with full bandwidth.

34. The two-dimensional layout of claim 1, wherein said straight links connecting from switches in each said block are connecting to switches in the same said block; and

said cross links are connecting as vertical or horizontal or diagonal links between two different said blocks.

35. The two-dimensional layout of claim 8, wherein $d=4$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast multi-link Benes network with full bandwidth.

36. The two-dimensional layout of claim 8, wherein $d=4$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast multi-link Benes network and rearrangeably nonblocking for arbitrary fan-out multicast multi-link Benes network with full bandwidth.

37. The two-dimensional layout of claim 8, wherein $d=4$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast multi-link Benes network with full bandwidth.

38. The two-dimensional layout of claim 12, wherein $d=4$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast multi-link butterfly fat tree network with full bandwidth.

39. The two-dimensional layout of claim 12, wherein $d=4$ and there are at least two switches in each said stage in each

US 2012/0269190 A1

Oct. 25, 2012

37

said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast multi-link butterfly fat tree network and rearrangeably nonblocking for arbitrary fan-out multicast multi-link butterfly fat tree network with full bandwidth.

40. The two-dimensional layout of claim **12**, wherein $d=4$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast multi-link butterfly fat tree network with full bandwidth.

41. The two-dimensional layout of claim **26**, wherein $d=4$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-link multi-stage network with full bandwidth.

42. The two-dimensional layout of claim **26**, wherein $d=4$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized multi-link multi-stage network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-link multi-stage network with full bandwidth.

43. The two-dimensional layout of claim **26**, wherein $d=4$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing

network is strictly nonblocking for arbitrary fan-out multicast generalized multi-link multi-stage network with full bandwidth.

44. The two-dimensional layout of claim **30**, wherein $d=4$ and there is only one switch in each said stage in each said block connecting said forward connecting links and there is only one switch in each said stage in each said block connecting said backward connecting links and said routing network is rearrangeably nonblocking for unicast generalized multi-link butterfly fat tree network with full bandwidth.

45. The two-dimensional layout of claim **30**, wherein $d=4$ and there are at least two switches in each said stage in each said block connecting said forward connecting links and there are at least two switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for unicast generalized multi-link butterfly fat tree network and rearrangeably nonblocking for arbitrary fan-out multicast generalized multi-link butterfly fat tree network with full bandwidth.

46. The two-dimensional layout of claim **30**, wherein $d=4$ and there are at least three switches in each said stage in each said block connecting said forward connecting links and there are at least three switches in each said stage in each said block connecting said backward connecting links and said routing network is strictly nonblocking for arbitrary fan-out multicast generalized multi-link butterfly fat tree network with full bandwidth.

47. The two-dimensional layout of claim **1**, wherein said plurality of forward connecting links use a plurality of buffers to amplify signals driven through them and said plurality of backward connecting links use a plurality of buffers to amplify signals driven through them; and said buffers are either inverting or non-inverting buffers.

48. The two-dimensional layout of claim **1**, wherein said all switches of size $d \times d$ are either fully populated or partially populated.

* * * * *

EXHIBIT K

*Curriculum Vitae***VIPIN CHAUDHARY**

Department of Computer and Data Sciences
Case Western Reserve University
Cleveland, Ohio 44106

Phone: (216) 368-2800
Email: vipin@case.edu

EDUCATION

- 1992 Ph.D., Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX
Dissertation: On Mapping Parallel Algorithms in a Distributed Computing Environment
Advisor: Dr. J. K. Aggarwal
- 1989 M.S., Computer Science, The University of Texas at Austin, Austin, TX. No Thesis.
- 1986 B.Tech (Honors), Computer Science and Engineering, Indian Institute of Technology,
Kharagpur, India
Thesis: Prolog Compiler
-

SUMMARY

A veteran of High Performance Computing (HPC), Dr. Chaudhary has been actively participating in the science, business, government, and technology innovation frontiers of HPC for almost three decades. His contributions range from heading research laboratories and holding executive management positions, to starting new technology ventures. He is currently the Kranzusch Professor and inaugural Chair of Department of Computer and Data Sciences at Case Western Reserve University.

Previously he was a Program Director in the Office of Advance Cyberinfrastructure at National Science Foundation where he co-led the National Strategic Computing Initiative from NSF for the United States and was in the working group of the Quantum Leap Initiative, National Quantum Initiative, National Artificial Intelligence Research Institutes, Cyber, and the I-Corps Program (where he was also a Program Director). I-Corps program is now part of "The American Innovation and Competitiveness Act" that enables commercialization of research and venture startups. He co-chaired the Networking and Information Technology Research and Technology Program's Middleware and Grid Interagency Coordination (MAGIC) Team for United States. He was also in the working group of the US Interagency Modeling and Analysis Group and a member of the Advanced Computing Roundtable of the Council on Competitiveness.

He was the Empire Innovation Professor of Computer Science and Engineering, the Director of the university's Data Intensive Computing Initiative and the co-founder of the Center for Computational and Data-Enabled Science and Engineering at University at Buffalo, State University of New York.

He cofounded Scalable Informatics, a leading provider of pragmatic, high performance software-defined storage and compute solutions to a wide range of markets, from financial and scientific computing to research and big data analytics. From 2010 to 2013, Dr. Chaudhary was the Chief

Executive Officer of Computational Research Laboratories (CRL) where he grew the company globally to be an HPC cloud and solutions leader before selling it to Tata Consulting Services. Prior to this, as Senior Director of Advanced Development at Cradle Technologies, Inc., he was responsible for advanced programming tools for multi-processor chips. He was also the Chief Architect at Corio Inc., which had a successful IPO in July, 2000.

Dr. Chaudhary was awarded the prestigious President of India Gold Medal in 1986 for securing the first rank amongst graduating students at the Indian Institute of Technology (IIT). He received the B.Tech. (Hons.) degree in Computer Science and Engineering from the Indian Institute of Technology, Kharagpur, in 1986 and a Ph.D. degree from The University of Texas at Austin in 1992.

RESEARCH INTERESTS

High Performance Computing and Applications to Science, Engineering, Biology, and Medicine; Big Data; Computer Assisted Diagnosis and Interventions; Medical Image Processing; Computer Architecture and Embedded Systems; Spectrum Management; Quantum Computing.

EMPLOYMENT HISTORY

08/2020 – Endowed Kranzusch Professor and Inaugural Chair, Department of Computer and Data Sciences, Case School of Engineering, Case Western Reserve University, Cleveland, Ohio.

06/2016 – 06/2020 Program Director, National Science Foundation, Directorate for Computer and Information Science and Engineering, Office of Advanced CyberInfrastructure

- Co-leading the National Strategic Computing Initiative from NSF for the United States
- Working group of the Quantum Leap Initiative and the National Quantum Initiative
- Working group of the National Artificial Intelligence Research Institutes Working Group
- Working group of the I-Corps Program (where he was also a Program Director). I-Corps program is now part of “The American Innovation and Competitiveness Act” that enables commercialization of research and venture startups.
- Co-chairing the Networking and Information Technology Research and Technology Program's Middleware and Grid Interagency Coordination (MAGIC) Team for United States.
- Working group of the US Interagency Modeling and Analysis Group

- Member of the Advanced Computing Roundtable of the Council on Competitiveness.
 - Started program with NASA - SWQU (Next Generation Software for Data Driven Models of Space Weather with Quantified Uncertainties)
 - Started multiple new programs (PPoSS, SWQU, CSSI, OAC Core, QCIS-FF, QII-TAQS, QLCI, National AI Research Institutes, Quantum Algorithm Challenge, Enabling Quantum Computing Platform Access) and managed many others (CDS&E, CDS&E-MSS, SPX, I-Corps, FMITF)
- 08/2011 – 08/2020 SUNY Empire Innovation Professor, Computer Science and Engineering, University at Buffalo, The State University of New York
- 05/2015 -- Founder, SpectrumFi, Inc.
- Licensing “Methods and systems for spectrum management” technology developed at UB.
- 02/2010 – 01/2013 CEO, Computational Research Labs (Tata CRL)
- Built the 4th largest supercomputer and the largest private supercomputer in the world
 - Provided complete High Performance Computing cloud and solutions
 - Grew the company worldwide
 - Sold to Tata Consulting Services
- 04/2009 – 05/2016 Founder, Diagnaid, Inc.
- Developed computer aided diagnosis for spine; raised small capital
- 09/2006 – 07/2011 SUNY Empire Innovation Associate Professor, Computer Science and Engineering, University at Buffalo, The State University of New York
- 09/2006 – 08/2011 Associate Professor (Adjunct), Department of Neurological Surgery, Wayne State University
- 09/2006 – 08/2007 Associate Professor (Adjunct), Department of Computer Science, Wayne State University
- 08/2003 – 03/2018 Co-Founder, Scalable Informatics, Inc
- Built the fastest storage and analytics computer systems targeting finance, pharmaceuticals, media, and scientific markets
 - Sold IP to private company
- 10/2005 – 05/2012 Founder and President, Micass L.L.C.
- Developed neurosurgery system; raised small capital

08/2005 – 07/2006	Associate Professor (Full Time Associate), Department of Biomedical Engineering, Wayne State University
09/2003 – 08/2006	Director, Institute for Scientific Computing, Wayne State University
09/2002 – 08/2006	Associate Professor, Institute for Manufacturing Research, Wayne State University
02/2001 – 01/2002	Senior Director, Cradle Technologies, Inc. <ul style="list-style-type: none"> • Responsible for architecture and programmability of the largest heterogeneous many-core media processor • Helped raise substantial funding
09/2000 – 08/2003	Founder and Associate Director, Institute for Scientific Computing, Wayne State University
01/2000 – 01/2001	Chief Architect, Corio, Inc. <ul style="list-style-type: none"> • Responsible for the data center architecture and delivery solution of the earliest cloud service • Successful IPO, July 2000 and later bought by IBM as their “On Demand” services.
09/1998 – 08/2006	Associate Professor (Full Time Associate before Nov 2002), Department of Computer Science, Wayne State University
08/1998 – 07/2007	Associate Professor (Adjunct since Nov 2002), Department of Electrical and Computer Engineering, Wayne State University
06/1992 – 07/1998	Assistant Professor, Department of Electrical and Computer Engineering, Wayne State University
01/1992 – 05/1992	Post-Doctoral Fellow (Jan-May), Computer and Vision Research Center, The University of Texas at Austin, Austin, TX
03/1987 – 12/1991	Graduate Research Assistant, Computer and Vision Research Center, The University of Texas at Austin, Austin, TX
08/1986 – 07/1988	Microelectronics and Computer Development (MCD) Fellow, Department of Computer Science, The University of Texas at Austin, Austin, TX

AWARDS

2019	National Science Foundation Director’s Superior Accomplishment Award
2018-	Honorary Professor, Amity University, Noida, India
2008	2007 Visionary Innovator, STOR, University at Buffalo, SUNY
1997	Excellence in Teaching Award, College of Engineering, Wayne State University

1996	Teaching Innovation Award, College of Engineering, Wayne State University
1993	National Science Foundation Research Initiation Award
1993	Wayne State University Faculty Research Award
1986 – 1988	Microelectronics and Computer Development (MCD) Fellow in Computer Sciences, The University of Texas at Austin
1986	<i>President of India Gold Medal</i> for first rank among all graduating students
1986	Institute Silver Medal for first rank among graduating Computer Science and Engineering students
1982 – 1986	General Proficiency Scholarship for academic accomplishments
1982	National Talent Search Scholarship, NCERT, India
1981	Certificate of Merit in the All India Annual Mathematics Talent Competition,
1980	National Scholarship and Certificate of Merit for outstanding performance in the All India Secondary School Examination, CBSE, India

TEAM AWARDS

2017	S. Liu, F. Shen, V. Chaudhary, and H. Liu, “ <i>MayoNLP at SemEval 2017 Task 10: Word Embedding Distance Pattern for Keyphrase Classification in Scientific Publications</i> ”, International Workshop on Semantic Evaluation (SemEval-2017), held with Annual Meeting of the Association for Computational Linguistics (ACL), August, 2017, Vancouver, Canada. (Top system in the challenge, tweet)
2013	“ <i>Scalable Informatics siFlash and Jackrabbit</i> ”, STAC-M3 Financial Industry Benchmarks (Best performance in 10 of 17 benchmarks. Continued to own these records for two years, only to be broken by Scalable Informatics hardware by customer.)
2005	“ <i>Cradle CT3600 MDSP</i> ”, EDN Innovation Award (Chip of the Year).
2003	“ <i>Cradle ECE3400/MPE3400</i> ”, Top 5 Microprocessor Report (MPR) Analysts’ Choice Award as Best Extreme Processor.

OTHER HONORS

2005	IEEE Computer Society, Certificate of Appreciation for Outstanding Service
1988	Who's Who among International students in American Universities.

PROFESSIONAL MEMBERSHIPS AND ACTIVITIES

Professional Memberships

- Senior Member, Institute of Electrical and Electronics Engineers (IEEE), IEEE Computer Society, IEEE Engineering in Medicine and Biology
- Member, American Association for the Advancement of Science (AAAS)
- Member, SPIE
- Member, American Medical Informatics Association (AMIA)
- Member, Association for Computing Machinery (ACM) – SIGPLAN, SIGARCH

- Member, USENIX (also University Representative at Wayne State University until 2006)

Editorial Services

- Associate Editor, IEEE Transactions on Cloud Computing, 2020 – present.
- Guest Editor, Computer Methods and Programs in Biomedicine, Elsevier, 2009.
- Editorial Board, International Journal of Smart Home, Science and Engineering Research Support Society, since 2006.
- Editorial Board, International Journal of Embedded Systems, Inderscience Publishers, 2003-2012.
- Guest Editor, Special Issue on High Performance Computing in Medicine and Biology, International Journal of Bioinformatics Research and Applications (IJBRA), 2006.
- Guest Editor, Special Issue on Media and Stream Processing, International Journal of Embedded Systems, 2006.
- Associate Guest Editor, Special issue of IEICE Transactions on Information and Systems on Hardware/Software Support for High Performance Scientific and Engineering Computing, July 2004.

Administrative Service

- Co-Founder, Computational and Data Enabled Science and Engineering Program, 2013-.
- Director, Data Intensive Discovery Initiative, 2009 –
- Vice Chair of Operations, South East Michigan IEEE Computer Society, 1997 -- 98.
- University Representative for USENIX, 1997 – 2006.

Conference/Symposium Chair

- Program Chair, 4th International Conference on Computing and Network Communications (CoCoNet'20), Chennai, India, October 14-17, 2020.
- General Chair, Fifth International Symposium on Signal Processing and Intelligent Recognition Systems, December 18-21, 2019, Trivandrum, India.
- Member of Advisory Committee, Fourth International Symposium on Signal Processing and Intelligent Recognition Systems, September 19-22, 2018, Bangalore, India.
- Member of Advisory Committee, International Symposium on Computational Science, December 12-15, 2015, Prasanthinilayam, India.
- Member of Steering Committee, *IEEE/ACM International Conference on High Performance Computing*, December 2015, Bangalore, India
- Member of Steering Committee, *IEEE/ACM International Conference on High Performance Computing*, December 2014, Goa, India
- Member of Steering Committee, *IEEE/ACM International Conference on High Performance Computing*, December 2013, Bangalore, India
- Member of Steering Committee, *IEEE/ACM International Conference on High Performance Computing*, December 2012, Pune, India
- Member of Steering Committee, *IEEE/ACM International Conference on High Performance Computing*, December 2011, Bangalore, India.

- Industry/User Symposium co-Chair, *IEEE/ACM International Conference on High Performance Computing*, December 2010, Goa, India.
- Program Vice-Chair, *IEEE International Conference on Advanced Information Networking and Applications* (AINA 2010), Perth, Australia, April 20-23, 2010.
- Program Vice-Chair, *IEEE International Conference on Advanced Information Networking and Applications* (AINA 2009), Bradford, UK, May 26-29, 2009.
- Track Chair for *Systems for Biological and Medical Applications, International Conference on Complex, Intelligent and Software Intensive Systems*, March 16-19, 2009, Fukuoka, Japan.
- Co-Founder and General Co-Chair, *IEEE International Symposium of Ubisafe Computing*, 2007, Niagara Falls, NY (<http://www.ubisafe.org/2007/>).
- Member of Steering Committee, *IEEE International Symposium of Ubisafe Computing*, 2007-.
- Co-Founder and General Co-Chair, *IEEE International Symposium on Bioinformatics and Life Science Computing*, 2007, Niagara Falls, NY (<http://www.laas.fr/BLSC07/>).
- Award Chair, *4th IEEE International Conference on Ubiquitous Intelligence and Computing*, Hong Kong, China, July 11-13, 2007.
- International Liaison Chair, *3rd IEEE International Conference on Ubiquitous Intelligence and Computing*, Wuhan and The Three Gorges, China, September 3-6, 2006.
- Program Vice-Chair, *2005 IFIP International Conference on Embedded and Ubiquitous Computing* (EUC 05), Nagasaki, Japan, 6-9 December 2005. (<http://euc05.euc-conference.org/>).

Workshop Chair

- Co-Organizer, *International Workshop on Healthcare Knowledge Discovery and Management (IHKDM)*, co-located with IEEE International Conference on Healthcare Informatics (ICHI 2017), Park City, Utah, August 23-26, 2017.
- Symposium co-Chair, *IEEE/ACM International Conference on High Performance Computing Industry, Research and User Symposium on Weather and Climate Modeling Challenges*, December 2011, Bangalore, India.
- Co-Chair, *2nd International Workshop on High-Performance Medical Image Computing for Image-Assisted Clinical Intervention and Decision Making* (held with MICCAI), Sep 24, 2010, Beijing, China.
- General Co-Chair, *IEEE International Workshop on Bioinformatics and Life Science Modeling and Computing*, Bradford, UK, May 26-29, 2009.
- Co-Chair, *International Workshop on High-Performance Medical Image Computing and Computer Aided Intervention* (held with MICCAI), Sep 10, 2008, New York, USA.
- General Co-Chair, *IEEE International Workshop on Bioinformatics and Life Science Modeling and Computing*, March 25-28, 2008, Okinawa, Japan. (<http://www.laas.fr/BLSMC08/>)
- Program Co-Chair, *International Workshop on Intelligent Informatics in Biology and Medicine* (IIBM 2008) will be held on March 4th – 7th, 2008, at Polytechnic University of Catalonia in Barcelona, Spain. (<http://www.cisis-conference.eu/>)

- Program Co-Chair, *2007 International Workshop on Embedded Parallel and Distributed Computing (EPDC-07)*, along with International Conference on Parallel Processing, XiAn China, September 2007 (<http://www.nwpu.edu.cn/epdc07/>).
- Co-Chair, *5th Workshop on Compile/Runtime Techniques for Parallel Computing*, along with 35th International Conference on Parallel Processing, August 14—18, Columbus, Ohio, 2006 (<http://www.pdcl.eng.wayne.edu/crtpc-2006>).
- Steering Committee Co-Chair and General Co-Chair, *2nd IEEE International Workshop on High Performance Computing in Medicine and Biology (HiPCoMB 2006)*, April 18-20, 2006, Vienna, Austria (<http://pdcl.wayne.edu/HiPCoMB-2006/>).
- 5th Annual US Army Vetronics Institute Winter Workshop Series, “*Workshop on Next Generation Embedded Processors: Combining RISC with Reconfigurable Fabric*”, January 11, 2006, Warren, MI.
- General Co-Chair, *7th Workshop on Media and Streaming Processors*, along with ACM MICRO, Barcelona, Spain, November 12-16, 2005 (<http://pcsostres.ac.upc.edu/micro38/>).
- General Co-Chair, *1st IEEE International Workshop on Parallel and Distributed Embedded System (PDES-05)*, July 20-22, 2005, Japan (<http://juliet.stfx.ca/~lyang/icpads05-pdes/>).
- Program Co-Chair, *1st IEEE International Workshop on High Performance Computing in Medicine and Biology (HiPCoMB 2005)*, July 20-22, 2005, Japan (<http://pdcl.wayne.edu/HiPCoMB-2005/>).
- Co-Chair, *4th Workshop on Compile/Runtime Techniques for Parallel Computing*, along with 34th International Conference on Parallel Processing, June 14—17, Oslo, Norway, 2005 (<http://www.pdcl.eng.wayne.edu/crtpc04>).
- 4th Annual US Army Vetronics Institute Winter Workshop Series, “*Workshop on Multiprocessor DSP: Reconfigurable, Real-Time, and High Performance System-on-Chip Solution*”, January 10, 2005, Warren, MI.
- Co-Chair, *6th Workshop on Media and Streaming Processors*, along with ACM MICRO, Portland, OR, December 8, 2004 (<http://www.pdcl.eng.wayne.edu/msp6>).
- Co-Chair, *3rd Workshop on Compile/Runtime Techniques for Parallel Computing*, along with 33rd International Conference on Parallel Processing, August 15—18, Montreal, Canada, 2004 (<http://www.pdcl.eng.wayne.edu/crtpc04>).
- *Organizer of Scientific and Bio-Computing Workshop*, sponsored by the Institute for Scientific Computing, Sep. 19, 2003. This workshop had four external distinguished speakers and a poster session. Over 130 faculty, students, and researchers participated in the workshop with 20 posters presented.
- 3rd Annual US Army Vetronics Institute Winter Workshop Series, “*Workshop on Software Scalable System on Chip: Reconfigurable, Real-Time, and High Performance System-on-Chip Solution*”, January 12, 2004, Warren, MI.
- Co-Chair, *5th Workshop on Media and Streaming Processors*, with ACM MICRO, San Diego, December 1, 2003 (<http://www.pdcl.eng.wayne.edu/msp5>).
- Co-Chair, *1st Workshop on Media Processing in Embedded Systems and SoCs*, along with ACM CASES, November 2003 (<http://www.crest.gatech.edu/conferences/cases2003/mases03.html>).
- Co-Chair, *2nd Workshop on Compile/Runtime Techniques for Parallel Computing*, along with International Conference on Parallel Processing, October 6—9, 2003 (<http://www.pdcl.eng.wayne.edu/crtpc03>).

- 2nd Annual US Army Vetronics Institute Winter Workshop Series, “*Workshop on Universal Micro Systems: Reconfigurable, Real-Time, and High Performance System-on-Chip Solution*”, December 4, 2002, Warren, MI.
- Co-Chair, *1st Workshop on Compile/Runtime Techniques for Parallel Computing*, along with International Conference on Parallel Processing, August 2002 (<http://www.pdcl.eng.wayne.edu/crtpc02>).
- Co-Chair, *4th Workshop on Media and Streaming Processors*, with ACM MICRO, November 2002 (<http://www.pdcl.eng.wayne.edu/msp4>).
- Co-Chair, *3rd Workshop on Media and Streaming Processors*, with ACM MICRO, December 2001 (<http://www.pdcl.eng.wayne.edu/msp01>).
- Chairman, International Workshop on “*Programming and Applications of Parallel/Distributed Systems*”, December 1997.

Program Committee Member

- The 9th IEEE Annual Ubiquitous Computing, Electronics, and Mobile Communication Conference (UEMCON), Columbia University, NY, USA, November 2018.
- The 7th IEEE Annual Ubiquitous Computing, Electronics, and Mobile Communication Conference (UEMCON), NY, USA, October 2016.
- IEEE/ACM International Conference on High Performance Computing, December 2015, Bangalore, India.
- IEEE/ACM International Conference on High Performance Computing, December 2014, Goa, India.
- 13th IEEE International Conference on High Performance Computing and Communications (HPCC-2011), September 2-4, 2011, Banff, Canada.
- 2nd Workshop on using Emerging Parallel Architectures (WEPA), Amsterdam, The Netherlands, May 31 - June 02, 2010, <http://www3.ntu.edu.sg/home/asbschmidt/>
- 2009 IEEE International Conference on Service-Oriented Computing and Applications (SOCA 2009), December 14-15, 2009, Taipei, Taiwan.
- Fourth International Workshop on Trustworthy Ubiquitous Computing (TwUC 2009), December 14-16, 2009, Kuala Lumpur, Malaysia.
- MICCAI-Grid Workshop: Medical imaging on GRID, HPC and GPU infrastructures: achievements and perspectives, London, UK, September 2-24, 2009.
- Workshop on High-End and Parallel Storage Systems (Co-held with PACT-2009 (Parallel Architectures and Compilation Techniques), Raleigh, North Carolina, USA, September 12-13, 2009
- 14th International Conference on Parallel and Distributed Systems (ICPADS 2008), December 8-10, 2008, Melbourne, Australia.
- Third International Workshop on Trustworthy Ubiquitous Computing (TwUC 2008), November 24-26, 2008 Linz, Austria.
- 21st International Conference on Parallel and Distributed Computing and Communication Systems, September 24-26, 2008, New Orleans, Louisiana, USA.
- 5th International Symposium on Embedded Computing, Oct 6-8, 2008, Beijing, China. (<http://conference.cs.cityu.edu.hk/sec08/sec.htm>)

- 20th International Conference on Parallel and Distributed Computing Systems, September 24-26, 2007, Las Vegas, USA.
- International Workshop on Intelligent Systems and Smart Home (WISH-07), Niagara Falls, Canada, August 2007.
- 5th International Symposium on Parallel and Distributed Processing and Applications (ISPA07), August 21-24, Niagara Falls, Canada, 2007 (<http://www.cs.umanitoba.ca/~ispa07/>).
- International Conference on Multimedia and Ubiquitous Engineering (MUE'07), April 2007, Seoul, Korea.
- 8th International Workshop on High Performance Parallel and Distributed Scientific and Engineering Computing (PDSEC), held with IEEE/ACM International Parallel and Distributed Processing Symposium, March 26—30, Long Beach, CA, 2007.
- 3rd International Conference on Distributed Computing & Internet Technology, 2006 (ICDCIT '06), Dec 20-23, 2006, Bhubaneshwar, India.
- 1st International Workshop on Trustworthy Ubiquitous Computing, Dec 4, 2006, Yogyakarta, Indonesia.
- 1st International Workshop on Smart Home, 2006 (IWSH '06), Nov 10-11, 2006, Cheju Island, Korea.
- 35th International Conference on Parallel Processing (ICPP), August 14-18, 2006, Columbus, Ohio, 2006.
- 8th International Workshop on High Performance Scientific and Engineering Computing with Applications (HPSECA), held with 35th International Conference on Parallel Processing, August 14—18, Columbus, Ohio, 2006.
- 3rd International Workshop on Embedded Computing, held with 35th International Conference on Parallel Processing, August 14—18, Columbus, Ohio, 2006.
- International Workshop on Embedded Software Optimization (ESO 2006), August 01-04, 2006, Seoul, Korea.
- 7th International Workshop on Parallel and Distributed Scientific and Engineering Computing (IPDPS-PDSEC06), April 25-29, 2006, Rhodes Island, Greece.
- IEEE 20th International Conference on Advanced Information Networking and Applications, April 18-20, 2006, Vienna, Austria.
- 3rd International Conference on Distributed Computing & Internet Technology, 2006 (ICDCIT '05), Dec 22-24, 2005, Bhubaneshwar, India.
- International Workshop on Ubiquitous Intelligence Smart Worlds (UISW2005) December 6-7, 2005, Nagasaki, Japan.
- 7th Workshop on High Performance Scientific and Engineering Computing with Applications (HPSECA), held with 34th International Conference on Parallel Processing, June 14—17, Oslo, Norway, 2005.
- 2nd International Workshop on Embedded Computing (EC-05), held with 34th International Conference on Parallel Processing, June 14—17, Oslo, Norway, 2005.
- 6th International Workshop on Parallel and Distributed Scientific and Engineering Computing (IPDPS-PDSEC05), April 4-8, 2005, Denver, Colorado, USA.
- International Workshop on Ubiquitous Smart Worlds (USW2005) March 28-30, 2005, Taipei, Taiwan.

- 2nd International Symposium on Parallel and Distributed Processing and Applications, December 13-15, 2004, Hong Kong, China.
- International Conference on Embedded and Ubiquitous Computing (EUC-04) August 26-28, 2004, Aizu, Japan.
- 6th Workshop on High Performance Scientific and Engineering Computing with Applications (HPSECA), held with 33rd International Conference on Parallel Processing, August 15—18, 2004, Montreal, Canada.
- 3rd Workshop on Compile/Runtime Techniques for Parallel Computing, held with 33rd International Conference on Parallel Processing, August 15—18, 2004, Montreal, Canada.
- 5th International Workshop on Parallel and Distributed Scientific and Engineering Computing (IPDPS-PDSEC04), April 26–30, Santa Fe, New Mexico, USA.
- 16th International Conference on Computer Applications in Industry and Engineering, Las Vegas, Nevada, November 11-13, 2003.
- 32nd International Conference on Parallel Processing, October 6-9, 2003, Kaohsiung, Taiwan.
- 5th Workshop on High Performance Scientific and Engineering Computing with Applications (HPSECA), held with 32nd International Conference on Parallel Processing, October 6-9, 2003, Kaohsiung, Taiwan.
- 2nd Workshop on Compile/Runtime Techniques for Parallel Computing, held with 32nd International Conference on Parallel Processing, October 6-9, 2003, Kaohsiung, Taiwan.
- 2nd Workshop on Hardware/Software Support for High Performance Scientific and Engineering Computing (SHPSEC-03), held with 12th IEEE Conference on parallel Architecture and Compilation Techniques, New Orleans, Louisiana, Sept. 27 – Oct. 1, 2003.
- 7th International Conference on Computer Science and Informatics, Cary, North Carolina, September 26-30, 2003.
- 4th Workshop on Parallel and Distributed Scientific and Engineering Computing with Applications (PDSECA), held with 17th International Parallel and Distributed Processing Symposium, April 22-26, 2003, Nice, France.
- 4th Workshop on Media and Streaming Processors, held with ACM Micro Conference, November 2002, Istanbul, Turkey.
- International Conference on Computer Applications in Industry and Engineering, November 2002.
- 3rd Workshop on Media and Streaming Processors, held with ACM Micro Conference, December 2002, Austin, Texas.
- International Conference on Computers and Their Applications, November 2001.
- International Conference on Computers and Their Applications, November 2000.
- International Conference on Computers and Their Applications, March 1999.
- International Conference on Advanced Computing, December 1998.
- International Conference on Computers and Their Applications, March 1998.
- International Conference on Distributed Processing and Networking, December 1997.
- International Conference on Parallel and Distributed Computing Systems, 1995.
- International Conference on Networks, 1996.

Referee

Grant Proposals:

- AAAS International Proposals, 2019
- Canada Innovation Fund, 2012
- DRDO, India – 2011
- National Science Foundation Reviewer – May 2007, Oct 2007, March 2008, June 2010, 2012, 13.
- National Science Foundation Panel (June 2010, March 2008, October 2005, April 2005, May 2002, August 2001, January 2001, January 2000, 1999)
- Research Grants Council, Hong Kong (March 2002, Feb 2004)
- Hong Kong Technology Cooperation Funding Scheme (TCFS) of the Innovation and Technology Support Program, May 2007
- Scientific Foundation, Ireland (October 2001)

External Evaluator for Tenure/Promotion Cases:

- Oakland University, MI – Computer Science and Engineering
- Cleveland State University, OH – Electrical and Computer Engineering

External Evaluator for Ph.D. Dissertation

- Indian Institute of Technology, Guwahati.

Recent Journals:

- IEEE Transaction on Medical Imaging
- IEEE Transactions on Knowledge and Data Engineering
- IEEE Transactions on Parallel and Distributed Systems
- IEEE Transactions on Computers
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- IEICE Transactions on Information and Systems
- Journal of Parallel and Distributed Computing
- The Computer Journal
- Journal of Supercomputing
- Journal of Systems and Software
- Journal of Parallel Algorithms and Architectures
- The Parallel Programming Journal
- ISCA Journal of Computers and Their Applications
- Parallel Computing
- International Journal of Bioinformatics Research and Applications
- Information Processing Letters
- Microprocessors and Microsystems

Conferences/Symposia/Workshops:

- Numerous conferences and workshops related to parallel computing, distributed computing, computer vision, computer architecture, embedded systems, bioinformatics, life science and medical computing, medical image processing, etc.

NATIONAL SERVICE

- Member of the Advanced Computing Advisory Committee (HPCAC), U.S. Council on Competitiveness, 2016 –
- Member of NSA-DOE HPC Technical Meeting, 2016 –
- Member of National Strategic Computing Initiative, Quantum Leap Initiatives, and I-Corps.
- NSF representative for the NIH Interagency Modeling and Analysis Group (IMAG) and Multi-Scale Modeling (MSM) Consortium
- Managing five programs at NSF

UNIVERSITY SERVICE

Chair

- Center for Computational Research Strategy Task Force, 2007-2008, University at Buffalo
- University Grid Operations Committee (2003 – 2006), Wayne State University

Committee Member

- CCR Faculty Advisory Committee, 2007 – 2010
- Faculty Senate Executive Committee, 2009-2010
- Faculty Senate, 2009-2011
- Future of the Informatics Program, 2007-2008, University at Buffalo
- Selective Salary Committee (2002, 2003), Wayne State University
- University High Performance Computing Advisory Group (2002 – 2006), Wayne State University
- University High Performance Computing Planning Committee (1995 – 1999), Wayne State University
- High Performance Imaging Task Force, Wayne State University and The Detroit Medical Center, (1996 – 99)

Ad hoc

- C&IT Committee for enhancing University computing facilities, Wayne State University
- Examiner, Doctoral Defense for Mr. Imran Ahmad, Department of Computer Science, Wayne State University

COLLEGE SERVICE

Committee Chair

- College Tenure Committee, 2014-15, University at Buffalo
- College Computer Advisory Committee (1995 – 99), Wayne State University

Committee Member

- Faculty Senate Member 2009-2011
- Tenure Committee (2009-2010; 2011-12)
- Biomedical Engineering Chair Search Committee, 2008
- New Engineering Building Core Team
- Biomedical Engineering Department Design Team, 2007, University at Buffalo
- ECE Chair Search (1993 – 96), Wayne State University
- Department Representative to the College of Engineering Executive Committee of the Faculty Assembly (1995 – 98), Wayne State University
- Secretary, College of Engineering Executive Committee of the Faculty Assembly, (1995 – 98), Wayne State University
- College of Engineering Computer Advisory Committee (1992 – 99), Wayne State University

Ad hoc:

- Faculty Task Group on Computers and Numerical Methods, Core Curricula for ABET Engineering Criteria 2000, Wayne State University

DEPARTMENTAL SERVICE*Committee Chair*

- Graduate Recruitment Committee (2002-03), Wayne State University

Committee Member

- Computer Science and Engineering, Steering Committee for Industrial Advisory Board (2009 – 2015) at University at Buffalo
- Computer Science and Engineering, Overseas MS-ES Program (2009 – 2010) at University at Buffalo
- Computer Science and Engineering, Course Load Policy Committee (2009 – 2010) at University at Buffalo
- Computer Science and Engineering, Facilities Committee (2009 – 2015) at University at Buffalo
- Computer Science and Engineering, Engineering Building Committee (2009 – 2010) at University at Buffalo
- Computer Science and Engineering, Faculty Search Committee (2007 – 10; 2011 – 15) at University at Buffalo
- Computer Science and Engineering, Overseas MS-ES Program (2008 – 2009) at University at Buffalo
- Computer Science and Engineering, Facilities Committee (2008 – 2009) at University at Buffalo

- Computer Science and Engineering, Engineering Building Committee (2008 – 2009) at University at Buffalo
 - Computer Science and Engineering, Facilities Committee (2007 – 2008) at University at Buffalo
 - Computer Science and Engineering, Engineering Building Committee (2007 – 2008) at University at Buffalo
 - Computer Science and Engineering, Graduate Admissions Committee (2006 – 2007) at University at Buffalo
 - Computer Science, Graduate Committee (2003 – 2006), Wayne State University
 - Computer Science Promotion and Tenure (2003 – 2006), Wayne State University
 - Computer Science Faculty Search (2001-02, 2005-06), Wayne State University
 - ECE Recruitment Committee (1996—97), Wayne State University
 - ECE Tenure, Promotion, and Salary Committee (1994 – 95), Wayne State University
 - ECE Faculty Search Committee (1993 -- 95, 96 – 98), Wayne State University
 - ECE Graduate Committee (1992 – 99), Wayne State University
 - ECE Math Committee (1993 – 98), Wayne State University
 - ECE Ph.D. Preliminary Exam Committee for Switching Theory, Computer Organization, and Math (1992 – 99), Wayne State University
-

COURSES TAUGHT

Undergraduate-level Courses Taught

- | | |
|----------|---|
| ECE 468 | Computer Organizations and Architecture, Wayne State University |
| CSC 4100 | Computer Architecture, Wayne State University |

Graduate-level Courses Taught

- | | |
|----------|--|
| CSE 726 | Seminar: Data Intensive and Cloud Computing and their Applications (Spring 2009 at University at Buffalo; Fall 2011) |
| CSE 702 | Seminar: Big Data (Spring 2013) |
| CSE 722 | Seminar: Big Data (Fall 2013, 2014) |
| CSE 703 | Seminar: Data Intensive High Performance Computing and Application (Spring 2008 at University at Buffalo) |
| CSE 603 | Parallel and Distributed Processing (Spring 2015) |
| CSE 587 | Data Intensive Computing (Spring 2015) |
| CSE 711 | Seminar: Data Intensive Supercomputing (Fall 2011) |
| CSE 633 | Parallel Algorithms (Spring 2015) |
| CSE 703 | Seminar: Trends in High Performance Computing (Spring 2007 at University at Buffalo) |
| CSE 487 | Data Intensive Computing (Spring 2015) |
| ECE 5610 | Introduction to Parallel and Distributed Systems, Wayne State University |
| CSC 6220 | Parallel Computing I, Wayne State University |
| ECE 5950 | Special Topics in Parallel Processing, Wayne State University |
| CSC 7100 | Advanced Computer Architecture, Wayne State University |
| ECE 7610 | Advanced Parallel and Distributed Systems, Wayne State University |

CSC 7220	Parallel Computing II, Wayne State University
ECE 7620	Real-Time Systems, Wayne State University
ECE 7660	Parallel Computer Architecture, Wayne State University
CSC 7991	Special Topics in Computer Science “Parallel Computer Architecture”, Wayne State University
ECE 7950	Parallelizing Compilers for High Performance Computing, Wayne State University
CSC 8260	Trends in HPC Architectures, Systems, and Software, Wayne State University

Courses Created

CSE 722	Seminar: Big Data (Fall 2013, 2014)
CSE 702	Seminar: Big Data (Spring 2013)
CSE 587/487	Data Intensive Computing (Spring 2015)
CSE 726	Seminar: Data Intensive and Cloud Computing and their Applications (Spring 2009 at University at Buffalo)
CSE 703	Seminar: Data Intensive High Performance Computing and Application (Spring 2008 at University at Buffalo)
CSE 711	Seminar: Data Intensive Supercomputing (Fall 2011)
CSE 603	Parallel and Distributed Processing
CSE 703	Seminar: Trends in High Performance Computing (Spring 2007 at University at Buffalo)
ECE 5610	Introduction to Parallel and Distributed Systems, Wayne State University
ECE 5950	Special Topics in Parallel Processing, Wayne State University
CSC 7100	Advanced Computer Architecture, Wayne State University
ECE 7610	Advanced Parallel and Distributed Systems, Wayne State University
CSC 7991	Special Topics in Computer Science “Parallel Computer Architecture”, Wayne State University
ECE 7950	Parallelizing Compilers for High Performance Computing, Wayne State University
CSC 8260	Trends in HPC Architectures, Systems, and Software, Wayne State University

Other Short Courses Created

Cray hands-on training workshop, , Wayne State University

<http://www.pdcl.eng.wayne.edu/training/CrayI90.html>

Faculty Development Workshop “Course development using the World Wide Web”, May 1996, Wayne State University

<http://www.ece.eng.wayne.edu/~pdcl/fdw/fdw.html>

RESEARCH SUPERVISION

Post-Doctoral Fellow/Research Associate Supervision

- Hanqiang Liu, 2017 - 2018, Computer Aided Diagnosis; currently Associate Professor of Computer Science at Shaanxi Normal University.

- Raja Alomari, 2010 – 2014, Computer Aided Diagnosis; currently at VMWare, Inc.
- John Paul Walters, 2007 – 2009, High Performance Computing; currently employed at ISI, University of Southern California.
- Suryaprakash Kompalli, 2006- 2008, Computer Aided Diagnosis; currently employed at HP Labs.
- Chengzhong Xu, 1994-98; Currently Professor, Department of Electrical and Computer Engineering, Wayne State University
- Guy Edjlali, 1997-99; Currently employed at Google; previously Assistant Professor, Department of Electrical and Computer Engineering, Wayne State University.
- Ryan Jin, 1997-99; MD, Wayne State University

Full Time Research Assistant Supervision

- Nathaniel Byrnes, 2011- 12, Data Intensive Computing
- Taruna Seth, 2009 – present, Virtual Surgeries and Data Intensive Computing.
- Chetan Bhole, 2006- 2007, Computer Assisted Diagnosis; currently Doctoral student at University of Rochester.
- Gulsheen Kaur, 2002-2005, Computer Assisted Surgery; currently employed with Compuware Inc, MI.
- Mohammad Alam, 2004- present, Computer Assisted Surgery, employed with Compuware Inc., MI.
- Raghu Venkatram, 2003-2004, Computer Assisted Surgery; Currently employed North Carolina State University, NC.
- Dingguo Chen, 2005 – 2006, Computer Assisted Surgery, employed with NC State.
- Jun Tan, 2005, Computer Assisted Surgery; employed with Detroit Medical Center.

Doctoral Student Supervision

Dissertations Supervised as Major Professor

- Jialin Ju, ``*Unique Sets Oriented Loop Parallelization*”, August 1998; Currently employed with IBM, San Jose, CA; previously with AT&T and Pacific Northwest National Laboratories.
- Sumit Roy, ``*Compilation issues for Distributed Shared Memory on Clusters of Symmetrical Multiprocessors*”, August 1998; Currently a Member of Technical Staff at Hewlett Packard Lab., Palo Alto, CA; Previously Assistant Professor (Research), Wayne State University.
- David A. Reimann, ``*Real-Time Cone Beam Tomography*”, November 1998; Co-Advisor: Dr. I. K. Sethi (Dept. of Computer Science at Wayne State University); Currently employed as Chair and Associate Professor, Dept. of Mathematics and Computer Science, Albion College.
- Hai Jiang, “*Process/Thread Migration and Checkpointing for Homogeneous and Heterogeneous Distributed Systems*”, November 2003; Currently employed as Assistant Professor in Computer Science at Arkansas State University; Was Visiting Assistant Professor, Dept. of Computer Science, Wayne State University.
- Xiandong Meng, “*A Heterogeneous Computing Platform for Biological Sequence Database Searches*”, April 2007, Employed with Texas A&M University.

- John Walters, “*Fault Tolerant Techniques for High Performance Computing and a Bioinformatics Application*,” May 2007, Employed with University at Buffalo, SUNY as Post-Doctoral Fellow.
- Julia Eizenkop, “*Computer and Analytical Modeling of Phenomena in Silicon Films under Excimer Laser Irradiation*,” (Co-Advisor: Ivan Avrutsky), July 2007; Startup company in Stealth mode.
- Raja Alomari, “*Computer Aided Diagnosis of Intervertebral Disc Pathology in Lumbar Spine*”, December 2009; Post-doctoral Fellow at University at Buffalo, SUNY.
- Nandan Garg, “*Dealing with Misbehavior in Distributed Systems: A game-theoretic approach*”, (Co-Advisor: Daniel Grosu), Wayne State University, May 2010.
- Jason Caravas, “*Evolution of the higher Diptera based on nuclear versus mitochondrial genes*” [Biology student] (Co-Advisor: Markus Friedrich), Wayne State University, June 2010.
- Jaehan Koh, “*LumbarDiagnostics: A Framework for Diagnosing Lumbar Spine Pathology*”, May 2012, Samsung Research.
- Subarna Ghosh, “*Algorithms for Automatic Localization, Segmentation and Diagnosis of Lumbar Pathology*”, August 2014, Self Employed.
- Nilesh Khambekar, “*Quantified Dynamic Spectrum Access Paradigm*”, August 18, 2015, SpectrumFi.
- Chao Feng, “*Techniques for computing High Order Virial Coefficients Using Mixed-Precision Approach for Hybrid Architectures*”, October 17, 2016, Oracle.
- Ruhan Sa, “*Algorithms for Automatic Object Detection and Landmark Detection in Medical Images*”, August 7, 2018, Siemens Research.
- Sijia Liu, “*Algorithms for Relation Extraction from Biomedical Texts*”, October 22, 2018, Mayo Clinic.

Current Doctoral Students as Supervisor

- Wei Yi, University at Buffalo: Computational Finance
- Taruna Seth, University at Buffalo: Computational Finance
- Shi Yan, University at Buffalo: Computer Aided Diagnosis
- Vinooth Kulkarni, University at Buffalo: Quantum Computing

Doctoral Dissertations Supervised as Committee Member

- Sharath Chandrasekhara, Chair: Steven Ko, Department of Computer Science and Engineering, University at Buffalo.
- Junfei Wang, Advisor: Dr. Sargur Srihari, Dept. of Computer Science and Engineering, University at Buffalo.
- Duygu Sarikaya, Chair: Jason Corso, Department of Computer Science and Engineering, University at Buffalo.
- Geoff Gross, Chair: Rakesh Nagi, Department of Industrial and Systems Engineering, University at Buffalo.
- Daekeun You, Chair: Venu Govindaraju, Department of Computer Science and Engineering, University at Buffalo.

- Huaigu Cao, Chair: Venu Govindaraju, Department of Computer Science and Engineering, University at Buffalo.
- Ifeoma Inwogu, Chair: Venu Govindaraju, Department of Computer Science and Engineering, University at Buffalo.
- Xin Liu, Chair: Chunming Qiao, Department of Computer Science and Engineering, University at Buffalo.
- Zhi Zhang, Chair: Venu Govindaraju, Department of Computer Science and Engineering, University at Buffalo.
- Yong Xi, Advisor: Dr. Loren Schwiebert, Department of Computer Science, Wayne State University
- Song Fu, Advisor: Dr. C. Z. Xu, Department of Electrical and Computer Engineering, Wayne State University
- Hanpei Lufei, Advisor: Dr. W. Shi, Department of Computer Science, Wayne State University
- Hrant Hratchian, Advisor: Dr. H. B. Schlegel, Department of Chemistry, Wayne State University
- Tianying Yan, Advisor: Dr. W. Hase, Department of Chemistry, Wayne State University
- Lipen Sun, Advisor: Dr. W. Hase, Department of Chemistry, Wayne State University
- Ayad Salhieh, Advisor: Dr. Loren Schwiebert, Department of Computer Science, Wayne State University
- Changli Jiao, Advisor: Dr. Loren Schwiebert, Department of Computer Science, Wayne State University
- Jie Sun, Advisor: Dr. W. Hase, Department of Chemistry, Wayne State University
- Tissa Samarantunga, Advisor: Dr. S. M. Mahmud, Dept. of ECE, Wayne State University
- Rajinder Kaushal, Advisor: Dr. J. Bedi, Dept. of ECE, Wayne State University
- Saravahan Agasaveeran, Advisor: Dr. N. Tsao, Department of Computer Science, Wayne State University
- Ansaf Alrabady, Advisor: Dr. S. M. Mahmud, Dept. of ECE, Wayne State University
- Shin Ping Wang, Advisor: Dr. P. Siy, Dept. of ECE, Wayne State University

Current Membership on Doctoral Committees

Masters Student Supervision

Masters Thesis Supervised as Major Professor

- Vikas Gautam, M.S., ``Allocation Strategies for k -ary n -cubes'', May 1994; Employed with IBM; previously with Sequent Computers, Inc., Schaumburg, Illinois.
- Swamy Punyamurtula, M.S., ``Parallelizing Loops with Non-Uniform Dependences'', December 1994; Employed with AMD (Advanced Micro Devices), Austin, Texas.
- Chegu Vinod, M.S., ``Parallel Hierarchical Radiosity Algorithms'', May 1995; Employed with Hewlett-Packard, Cupertino; previously with Convex Computers, Dallas; with IBM, Austin, Texas.

- Subburajan Ponnuswamy, M.S., ``*Directed and Undirected Cayley Interconnection Networks for Point-to-Point Communication and VLSI Implementation*'', December 1995; Started his own company; Previously employed with Malibu Networks; Honeywell Research; Sequent Computers, Inc., Beaverton; EDS, Troy, Michigan.
- Padmanabhan Menon, M.S., ``*Real-time MRI Imaging*'', December 1998; Employed with Hewlett-Packard, Cupertino.
- Niranjan Ghatge, M.S., ``*Optimizing Automatically Generated Programs for a Software Distributed Shared Memory System*'', December 2000; Employed with Intel, Phoenix, AZ.
- Manish Shah, M.S., ``*A Policy Based User Configurable Java Security Architecture*'', December 2000; Employed with Sun Microsystems, Mountain View, CA; previously with Pixo, Cupertino, CA; Corio, San Carlos, CA.
- Darshan Thaker, M.S., ``*DSim: Distributed Shared Memory Simulation and Performance Analysis*'', December 2002; currently Ph.D. student at UC Davis.
- Anil Nambiar, M.S., ``*Tools for Mapping Resource Constrained Applications on Chip-Multiprocessors*'', December 2004; Employed with Sharp Microelectronics, Camas, WA.
- Ganesh Yadav, M.S., ``*Software-configurable Mode Decision for Intra Prediction in H.264/AVC*'', May 2005; Employed with Stretch, Inc., Mountain View, CA.
- Mamatha Nanjundaiah, M.S., ``*Enhanced Class 0 RFID Anti-Collision Protocol*'', December 2005; Employed with Denlaw Inc., MI; Earlier with AW Technical Center, MI.
- Snehal Joshi, M.S., ``*Bluetooth Enabled Haptic User Interface for Handhelds*'', December 2006; Employed with Palm Inc., Sunnyvale, CA.
- Nilesh Khambekar, M.S., ``*Utilizing OFDM Guard Interval for Spectrum Sensing*'', December 2006; Employed with Juniper Networks, Sunnyvale, CA.
- Kshitij Gunjekar, M.S., ``*Misbehaviour Detection in Ad Hoc Wireless Networks*'', December 2006; Employed with Redback Networks (Now Ericsson), San Jose, CA.
- Vidyananth Balu, M.S., ``*Application based empirical evaluation of multi-core NVIDIA graphics and IBM Cell BE processors*'', July 2008; Employed with Netapp, San Jose, CA.
- Neville Mehta, M.S., ``*Content Based Sub-Image Retrieval in Pathology Images*'', August 2009; Employed with Bioimagene, Inc., Cupertino, CA.
- Ata E. Husain Bohra, M.S., ``*Energy Optimization Techniques for Virtualized Clouds: Enhanced Power Modeling and Energy Aware VM Placements*'', June 2010; Employed with Riverbed Networks, San Jose, CA.
- Vikas Ashok Patil, M.S., ``*Rack Aware Scheduling in HPC Data Centers*'', June 2011; Employed with Factset.

Masters Essays Supervised

- Krishna Raman, M.S., ``*Coterie Based Generalization of Distributed Mutual Exclusion Algorithms*'', May 1995; Currently Employed with UUNET, Washington, D.C.; previously with EDS, Troy, Michigan; Co-Advisor: Dr. S. P. Rana with Dept. of Computer Science, Wayne State University.
- Tyakal Ramachandrababu, M.S., ``*Data Redistribution Techniques for Distributed Memory Multicomputers*'', May 1995; Currently Employed with EDS, Troy, Michigan; Co-Advisor: Dr. S. P. Rana with Dept. of Computer Science, Wayne State University.

Current Masters Students with Thesis

- Leila Talebpour, M.S., Dept. of Computer Science, SUNY Buffalo.
- Muhammad Aamir Masood, M.S., Dept. of Computer Science, SUNY Buffalo.

Masters Students with Project Supervision

- Over 1000 students supervised

Masters Thesis Supervised as Committee Member

- Rohit Shivaswamy, Advisor: Abani Patra, Dept. of Mechanical and Aerospace Engineering, University at Buffalo, SUNY
- Dipankar Das, Advisor: Jason Corso, Dept. of Computer Science and Engineering, University at Buffalo, SUNY
- Duygu Sarikaya, Advisor: Jason Corso, Dept. of Computer Science and Engineering, University at Buffalo, SUNY
- Bhaskar Purkayastha, Advisor: Venu Govindaraju, Dept. of Computer Science and Engineering, University at Buffalo, SUNY
- Omar Mukhtar, Advisor: Venu Govindaraju, Dept. of Computer Science and Engineering, University at Buffalo, SUNY
- Hanpei Lufei, Advisor: Dr. W. Shi, Dept. of Computer Science, Wayne State University
- Anubhav Das, Advisor: : Dr. D. Grosu, Dept. of Computer Science, Wayne State University
- Umesh Kant, Advisor: Dr. D. Grosu, Dept. of Computer Science, Wayne State University
- Manish Kochar, Advisor: Dr. L. Schwiebert, Dept. of Computer Science, Wayne State University
- Sanjay Majumdar, Advisor: Dr. L. Schwiebert. Dept. of ECE, Wayne State University
- Jie Lu, Advisor: Dr. Y. Zhao, Dept. of ECE, Wayne State University
- Shyamsundar Chandrasekharan, Advisor: Dr. V. Rajlich, Department of Computer Science, Wayne State University

GRANT SUPPORT*Externally Funded Research*

- PI, National Science Foundation, 6/13/19 – 6/12/20
Project: Intergovernmental Personnel Act (IPA) Assignment
Amount Awarded: \$314,027
- PI, National Science Foundation, 6/13/18 – 6/12/19
Project: Intergovernmental Personnel Act (IPA) Assignment
Amount Awarded: \$284,529
- PI, National Science Foundation, 6/13/17 – 6/12/18

Project: Intergovernmental Personnel Act (IPA) Assignment
Amount Awarded: \$280,006

- PI, National Science Foundation, 6/13/16 – 6/12/17
Project: Intergovernmental Personnel Act (IPA) Assignment
Amount Awarded: \$267,318
- Co-PI, National Science Foundation, 5/1/14 – 11/30/14
Project: I-Corps: Integrated framework for risk assessment for catastrophic natural disasters
Amount Awarded: \$50,000 (PI: A. Patra)
- PI, National Science Foundation, 5/1/13 – 11/30/13
Project: I-Corps: Standardized MRI Interpretation for Low Back Pain Diagnosis
Amount Awarded: \$50,000
- PI, Netezza Gift
Project: Research in Data Intensive Computing, October 2010 –
Amount Awarded: \$96,000
- PI, Big Data Fast Gift
Project: Research in Data Intensive Computing, September 2010 –
Amount Awarded: \$96,000
- Co-PI, National Science Foundation Grant, 9/1/2010 – 8/31/2015
Project: CDI Type II: New cyber-enabled strategies to realize the promise of quantum chemistry as a far-reaching tool for engineering applications
Amount Awarded: \$ 1,426,482 (PI: D. Kofke, co-PI: T. Furlani)
- Co-PI, National Science Foundation Grant, 6/1/2010 – 5/31/2015
Project: Technology Audit and Insertion Service for TeraGrid
Amount Awarded: \$7,763,246 (PI: T. Furlani, co-PI: M. Green, M. Jones, and G. Laszweski)
- PI, National Science Foundation, 2/1/10 – 1/31/13
Project: MRI-R2: Acquisition of a Data Intensive Supercomputer for Innovative and Transformative Research in Science and Engineering
Amount Awarded: \$4,600,351 (co-PIs: T. Furlani, D. Kofke, A. Patra, and B. Pitman)
- PI, New York State Innovation Economy Matching Grant Program, 2/1/10 – 1/31/13
Project: MRI-R2: Acquisition of a Data Intensive Supercomputer for Innovative and Transformative Research in Science and Engineering
Amount Awarded: \$460,035
- PI, National Science Foundation Grant, 8/15//2009 – 8/14/2011
Project: Acquisition of BCI – A Biomedical Computing Infrastructure
Amount Awarded: \$ 688,315 (Co-PIs: J. Corso, T. Furlani, K. Hoffman, V. Krovi)

- Co-PI, National Science Foundation Grant, 9/1/2009 – 8/31/2012
Project: A Comprehensive Framework for Timely Introduction of Emerging Data-Intensive Computing to STEM Audiences
Amount Awarded: \$249,875 (PI: Bina Ramamurthy, co-PI: John Van Benschoten)
- PI, New York State Office of Science, Technology, and Academic Research, UB Center for Advanced Biomedical and Bioengineering Technology (UB CAT)
Project: Computer Aided Detection of Lumbar Spinal Pathology on MRI Exams
Amount Awarded: \$20,700
- PI, Diagnaid, Inc., 9/2008 – 5/2009
Project: Computer Aided Detection of Lumbar Spinal Pathology on MRI Exams
Amount Awarded: \$20,700
- PI, Medcotek, Inc. Gift, May 2007 --
Project: Technical Challenges in Radiology
Amount: \$6,125
- PI, New York State Office of Science, Technology, and Academic Research, 11/2006-10/2009
Project: High Performance Computing for Life Science and Medicine
Amount Awarded: \$700,000
- PI, New York State Office of Science, Technology, and Academic Research, UB Center for Advanced Biomedical and Bioengineering Technology (UB CAT), 9/2006 – 7/2007
Project: Infrastructure for radiographic visualization
Amount Awarded: \$10,000 (with \$5,000 matching from Medcotek Inc.)
- PI, Sun Microsystems, Inc. Academic Excellence Grant, 5/2005 – 12/2005
Project: Michigan Center for Advanced Lifescience Computing (Equipment Grant)
Amount Awarded: \$ 50,000
- PI, Stretch, Inc., 1/2005 – 5/2005
Project: Design and Development of an H.264 Encoder/Decoder for Stretch Processor
Amount Awarded: \$ 28,887
- PI, Cradle Technologies, Inc., 5/2004 – 8/2004
Project: MPEG4 and H.264 Optimization for Cradle C4
Amount Awarded: \$ 17,820
- PI, Michigan Life Science Corridor, 9/2003 – 8/2007
Project: Integration of Bioengineering and Biocomputing to Advance Michigan Computer-Assisted Surgery Research
Amount Awarded: \$ 3,377,560 (PI – 42%, Co-PIs: Wayne State University - Ming Dong, Sorin Draghici, Farshad Fotouhi, Albert King, Weisong Shi, King Yang, Murali

Guthikonda, Zhimin Zhang, Oakland Univeristy – Ishwar Sethi, University of Michigan – William Grosky)

- PI, National Science Foundation Grant, 6/2000 – 7/2007
Project: IGERT- Interdisciplinary Traineeship in High Performance Computing Applications
Amount Awarded: \$ 2,949,618 (Co-PIs: E. Goldfield – 33%, H. B. Schlegel – 33%)
- Co-PI, Air Force Office for Scientific Research, 9/2002 – 8/2005
PI, William Hase, Other Co-PI: Chuck Doubleday (Columbia University)
Project: Dynamics of O(3P) reactions with gaseous, liquid, and solid hydrocarbons
Amount Awarded: \$ 300,000
- PI, Sun Microsystems, Inc., 2001
Project: Research in High Performance Computing Applications (Equipment Grant)
Amount Awarded: \$ 1,000,000
- PI, National Science Foundation Grant, 9/2000 – 12/2003
Project: ITR/ACF+SSI - Opportunistic Parallel Computation
Amount Awarded: \$ 253,455 (PI: 80%, Co-PI: D. P. Agrawal (Univ of Cincinnati)– 20%)
- PI, National Science Foundation Grant, 9/1999 – 8/2002
Project: Acquisition of a Cluster of Symmetric Multiprocessors (Equipment Grant)
Amount Awarded: \$ 298,563 (Co-PIs (equal share): G. Edjlali, E. Goldfield, W. L. Hase, and H. B. Schlegel)
- PI, Sun Microsystems, Inc., 9/1999 – 8/2002
Project: Research in High Performance Computing Applications (Equipment Grant)
Amount Awarded: \$ 243,000
- Senior Personnel (Co-Investigator), National Science Foundation Grant, 1/1998 – 12/1999
PI: John Camp
Project: High Performance Connections to the Internet (Equipment Grant)
Amount Awarded: \$ 317,859
- PI, National Science Foundation Grant, 9/1997 – 8/2000
Project: REU Site for Parallel and Distributed Applications
Amount Awarded: \$ 224,100 (PI: 35%, Co-PIs: Loren Schwiebert and C. Z. Xu)
- PI, National Science Foundation Grant, 1/1998 – 12/1998
Project: High Performance Computing on an ATM-Connected Cluster of Symmetric Multiprocessors (Equipment Grant)
Amount Awarded: \$ 102,400 (PI: 49%, Co-PIs: Loren Schwiebert and C. Z. Xu)
- PI, Ford Motor Company, 1/1997 – 12/1999
Project: Automatic Parallelization Environment for Computational Fluid Dynamics
Amount Awarded: \$ 150,000 (Co-PI: Ming Chia Lai)

Note: (12.5% acceptance rate worldwide; pure research grant with no deliverables)

- PI, Ford Motor Company, 1/1997 – 12/1999
Project: Parallel Software Development for Interfaces and Adhesive Interactions
Amount Awarded: \$ 150,000 (Co-PI: W. L. Hase)
Note: (12.5% acceptance rate worldwide; pure research grant with no deliverables)
- PI, Cray Research Inc., 9/1996 – 8/1997
Project: Real-Time Optimization of 3-Dimensional Conformal Radiation Therapy Treatment Planning
Amount Awarded: \$ 30,000 (PI: 40%, Co-PIs: C. Xu, G. Ezzel, and C. Kota)
- PI, Army Research Laboratory, 10/1994 – 9/1995
Project: Automatic Program Parallelization
Amount Awarded: \$ 37,654
- PI, Ford Motor Company, 8/1995 – 7/1996
Project: Parallelization of Molecular Dynamics Programs
Amount Awarded: \$ 25,000
- PI, National Science Foundation Grant, 9/1993 – 2/1997
Project: Research Initiation Award - Adaptive Load Balancing for Heterogeneous Distributed Systems
Amount Awarded: \$ 89,999
- PI, General Motors Research, 6/1994 – 7/1994
Project: Algorithms for Texture Mapping onto Curved Surfaces
Amount Awarded: \$ 5,000
- PI, IBM, 8/1992 – 7/1993
Project: Tools for mapping algorithms and visualizing the performance of parallel and distributed computing environments (Equipment Grant)
Amount Awarded: \$ 53,200

Internal Competitive Funded Grants

- University at Buffalo, UB 2020 Interdisciplinary Research Development Fund, (IRDF), “Characterization of Vascular Flow”, 6/15/2009 – 6/30/2010, \$28,000, PI: Kenneth Hoffman, co-PIs: Vipin Chaudhary, Bruce Pitman, Jaehun Jung, Tom Furlani, Adnan Siddiqui, Matthew Jones.
- WSU Graduate Research Assistantship Award, 2 GRAs, 2004 (with Markus Friedrich)
- WSU Stipend Enhancement for NSF IGERT, \$620,000, 1999. [Cost sharing for NSF IGERT grant (PI)]

- WSU Equipment Match, “Acquisition of a Cluster of Symmetric Multiprocessors”, \$200,000, 1999. [Cost sharing for NSF IGERT grant (PI)]
- WSU Equipment Match for NSF MRI Award “Acquisition of a Cluster of Symmetric Multiprocessors”, \$200,000, 1999.
- Faculty Mentor for Dr. Guy Edjlali, \$2000, 1999; Dr. Chengzhong Xu, \$2,000, 1998; Dr. Loren Schwiebert, \$2,000, 1997.
- WSU Graduate Research Assistantship Award, \$17,000, 1998.
- WSU Equipment Match, “High Performance Computing on an ATM-Connected Cluster of Symmetric Multiprocessors”, \$100,000, 1998.
- WSU Supplemental Research Equipment Fund Grant, “Programming Environment APE for Automatic Parallelization of Programs, \$58,710, 3/97--6/97. (Co-investigators: L. Schwiebert and C. Xu)
- WSU Research Stimulation Fund Grant, “Software for the Cray J916”, \$12,000, 1996-97.
- WSU Faculty Research Award, \$7,000, 1993.

PUBLICATIONS (* against author’s name indicates he was my student; ** against author’s name indicates he was a student who worked with me on this project but has a different academic advisor, *** against author’s name indicates he was my post-doc)

Books

1. V. Chaudhary, J. P. Walters*, and H. Jiang*, *Computation Checkpointing and Migration*, Nova Science Publishers, New York 2010.
2. R. Alomari and V. Chaudhary, “*Automated Lumbar Spine Diagnosis*”, Lambert Academic Publishing, Germany, 2010.

Book Chapters

1. R. Alomari*, S. Ghosh*, J. Koh*, and V. Chaudhary, “*Vertebral Column Localization, Labeling, and Segmentation*”, in *Computational Methods and Clinical Applications for Spine Imaging*, Springer, pp. 193-229. *Spinal Imaging and Image Analysis, Lecture Notes in Computational Vision and Biomechanics Volume 18*, 2015.
2. T. Seth* and V. Chaudhary, “*Big Data in Finance*”, in *Big Data: Algorithms, Analytics, and Applications*, Chapman and Hall/CRC Big Data Series, CRC Press, 2014.
3. J. P. Walters*, V. Chaudhary, and B. Schmidt, “*Database Searching with Profile Hidden Markov Models on Reconfigurable and Many-Core Architectures*”, in *Bioinformatics: High Performance Parallel Computer Architectures*, Taylor and Francis/CRC Press, 2010.

4. J. P. Walters*, J. Landman, and V. Chaudhary, “*Optimized Cluster-Enabled HMMER Searchers*”, in *Grid Computing for Bioinformatics and Computational Biology*, Editors: El-Ghazali Talbi and Albert Zomaya, John Wiley and Sons, 2007, pp. 51-60.
5. J. P. Walters*, Z. Liang**, W. Shi, and V. Chaudhary, “*Wireless Sensor Network Security: A Survey*” in *Security in Distributed, Grid, and Pervasive Computing*, Editor: Yang Xiao, Auerbach Publications, CRC Press, 2007, ISBN: 0849379210 (to appear).
6. V. Chaudhary and H. Jiang*, “*Techniques for Migrating Computations on the Grid*”, in *Engineering the Grid: Status and Perspective*, Editors: Beniamino Di Martino, Jack Dongarra, Adolfo Hoisie, Hans Zima, and Laurence T. Yang, American Scientific Publishers, pp. 399 – 415, January 2006, ISBN: 1-58883-038-1.
7. V. Chaudhary, F. Liu*, X. Meng*, V. Matta*, A. Nambiar*, G. Yadav*, and L. T. Yang, “*Parallel Implementations of Local Sequence Alignment: Hardware and Software*”, in *Parallel Computing in Bioinformatics and Computational Biology*, Editor: Albert Zomaya, John Wiley and Sons, 2006.
8. H. Jiang*, V. Chaudhary, and J. P. Walters*, “*Data Conversion for Heterogeneous Migration/Checkpointing*”, in *High Performance Computing: Paradigm and Infrastructure*, Editors: Laurence T. Yang and Minyi Guo, John Wiley and Sons, N.J., pp. 241 – 260, 2006, ISBN: 13-978-0-471-65471-1.
9. F. Liu* and V. Chaudhary, “*OpenMP for Heterogeneous Chip Multiprocessors*”, in *High Performance Computing: Paradigm and Infrastructure*, Editors: Laurence T. Yang and Minyi Guo, John Wiley and Sons, N.J., pp. 117 – 134, 2006, ISBN: 13-978-0-471-65471-1.
10. V. Chaudhary, W. L. Hase, H. Jiang*, L. Sun**, and D. Thaker*, “*Comparing various parallelizing approaches for tribology simulations*,” in *Hardware/Software Support for Parallel and Distributed Scientific and Engineering Computing*, Editors: L. T. Yang and Y. Pan, Kluwer Academic Publishers, 2004, pp. 231—252.
11. G. Edjlali***, A. Acharya, and V. Chaudhary, “*History-Based Access Control for Mobile Code*”, in *Secure Internet Programming*, Eds. Jan Vitek and Christian Jensen, Springer-Verlag, 1999, pp. 413 — 432.
12. V. Chaudhary, K. Kumari, P. Arunachalam, and J. K. Aggarwal, “*Manipulation of Octrees and Quadrees on Multiprocessors*”, in *Parallel Image Analysis and Processing*, (Editors: K. Inoue, A. Nakamura, M. Nivat, A. Saoudi, and P. S. P. Wang), World Scientific Publishing Company, 1994, pp. 25 – 41. (Extended version of a journal paper)
13. V. Chaudhary and J. K. Aggarwal, “*Parallelism in computer vision -- a review*”, Published as a chapter in *Parallel Algorithms for Machine Intelligence and Computer Vision*, (Editors: V. Kumar, P. S. Gopalakrishnan, and L. N. Kanal), Springer-Verlag, 1990, pp. 271 – 309.
14. V. Chaudhary and J. K. Aggarwal, “*Parallelism in low-level computer vision*”, in *Data Analysis in Astronomy III*, Edited by V. Di Gesu, L. Scarsi, P. Crane, J. H. Friedman, S. Levialdi, and M.C. Macarone, Plenum Press, 1989, pp. 255—270.

Refereed Journal Papers

1. Sijia Liu*, F Shen, R Komandur Elayavilli, Y Wang, M, Rastegar-Mojarad, V Chaudhary, H Liu, “*Extracting Chemical Protein Relations using Attention-based Neural Networks*”, Database. 2018.

2. H. Liu***, F. Zhao, and V. Chaudhary, “*Pareto-based Interval Type-2 Fuzzy c-means with Multi-scale JND Color Histogram for Image Segmentation*”, Journal of Digital Signal Processing, Elsevier, 2018.
3. C. Feng*, A. J. Schultz, V. Chaudhary, and D. A. Kofke, “*Eighth to sixteenth virial coefficients of the Lennard-Jones model*”, in The Journal of Chemical Physics, 143, 044504, 2015 (9 pages), DOI: 10.1063/1.4927339
4. S. Ghosh* and V. Chaudhary, “*Supervised Methods for Detection and Segmentation of Tissues in Clinical Lumbar MRI*”, in Computerized Medical Imaging and Graphics, 2014, October, 38(7), pp. 639-649; DOI: 10.1016/j.compmedimag.2014.03.005
5. J. Koh*, V. Chaudhary, E. K. Jeon, and G. Dhillon, “*Automatic Spinal Canal Detection in Lumbar MR Images in the Sagittal View Using Dynamic Programming*”, in Computerized Medical Imaging and Graphics, 2014, October, 38(7), pp. 569-579.
6. T. Seth*, V. Chaudhary, C. Buyea, and L. Bone, “*A Haptic Enabled Virtual Reality Framework for Orthopedic Surgical Training and Interventions*”, International Journal of Computers and Their Applications (IJCA), Vol. 21, No. 4, 2014.
7. V. Patil* and V. Chaudhary, “*Rack Aware Scheduling in HPC Data Centers: An Energy Conservation Strategy*”, Cluster Computing, Vol. 16, Issue 3, Springer, 2013, pp. 559-573.
8. R. Shivaswamy**, A. Patra, and V. Chaudhary, “*Integrating Data and Compute Intensive Workflows for Uncertainty Quantification in Large Scale Simulation – Application to Model Based Hazard Analysis*”, International Journal of Computer Mathematics, Vol 91, Issue 4, 2014, pp. 730-747.
9. Hazem Hiary, Raja S Alomari*, and Vipin Chaudhary, Segmentation and Localization of Whole Slide Images Using Unsupervised Learning, Journal of IET Image Processing, July, 2013, 7(5), pp. 464-471.
10. S. Al-Helo, R. Alomari*, S. Ghosh*, V. Chaudhary, G. Dhillon, M. B. Al-Zoubi, H. Hiary, T. M. Hamtini, “*Compression Fracture Diagnosis in Lumbar: A Clinical CAD System*”, International Journal of Computer Assisted Radiology and Surgery, May 2013, Vol 8, Issue 3, pp. 461-469.
11. R. Alomari*, Hazem Hiary, Maha Saadah and V. Chaudhary, “*Automated Segmentation of Stromal Tissue in Histology Images using a Voting Bayesian Model*”, Journal of Signal Image and Video Processing, Nov 2012, DOI: 10.1007/s11760-012-0393-2
12. A. Schultz, N. S. Barlow, V. Chaudhary, and D. A. Kofke, “*Mayer Sampling Monte Carlo Calculation of Virial Coefficients on Graphics Processors*”, Molecular Physics: An International Journal at the Interface between Chemistry and Physics, Taylor and Francis, 2012, pp. 1-9.
13. J. Koh*, V. Chaudhary and G. Dhillon, “*Disc Herniation Diagnosis in MRI Using LumbarDiagnositics CAD Framework*”, International Journal of Computer Assisted Radiology and Surgery (IJCARS), DOI 10.1007/s11548-012-0674-9, 2012.
14. R. Alomari*, J. Corso, and V. Chaudhary, “*Towards a Clinical Lumbar CAD: Herniation Diagnosis*”, International Journal of Computer Assisted Radiology and Surgery, vol. 6, no1, pp. 119-126, 2011.
15. R. Alomari*, J. Corso, and V. Chaudhary, “*Labeling of Lumbar Discs using both Pixel- and Object-Level Features with a Two-Level Probabilistic Model*”, IEEE Transactions on Medical Imaging, 30(1):1-10, 2011.

16. T. Scofield, J. Delmerico*, V. Chaudhary, and G. Valente, “*XtremeData dbX: An FPGA-Based Data Warehouse Appliance*”, IEEE Computing in Science and Engineering, July/August 2010, pp. 66 – 73.
17. X. Meng* and V. Chaudhary, “*A High-Performance Heterogeneous Computing Platform for Biological Sequence Analysis*”, IEEE Transactions on Parallel and Distributed Systems, vol. 21, no. 9, pp. 1267-1280, 2010.
18. R. Alomari*, J.J. Corso, V. Chaudhary, and G. Dhillon, “*Computer-aided diagnosis of lumbar disc pathology from clinical lower spine*”, International Journal of Computer Assisted Radiology and Surgery, Vol. 5, no 3, pp. 287-293, May, 2010.
19. J. P. Walters* and V. Chaudhary, “*Replication-Based Fault-Tolerance for MPI Applications*”, IEEE Transactions on Parallel and Distributed Systems, Vol. 20, No. 7, pp. 997-1010, June 2009.
20. H. Lufei**, W. Shi, and V. Chaudhary, “*Adaptive Secure Access to Remoter Services in Mobile Environments*”, IEEE Transactions on Services Computing, January 2008, Vol. 1, No. 1, pp. 49-61.
21. J. P. Walters* and V. Chaudhary, “*A Fault-tolerant Strategy for Virtualized HPC Clusters*”, The Journal of Supercomputing, Vol. 50, No. 3, pp. 209-239, 2009.
22. X. Meng* and V. Chaudhary, “*Boosting Data Throughput for Sequence Database Similarity Searches on FPGAs using an Adaptive Buffering Scheme*”, Journal of Parallel Computing, Vol. 35, No. 1, January 2009, pp. 1-11.
23. J. Eizenkop*, I. Avrutsky, D. G. Georgiev, and V. Chaudhary, “*Single-pulse excimer laser nanostructuring of silicon: A heat transfer problem and surface morphology*,” Journal of Applied Physics (Vol.103, Issue 9), 2008.
24. J. Walters*, X. Meng*, V. Chaudhary, T. Oliver, L. Y. Yeow, B. Schmidt, D. Nathan, J. Landman, “*A Hardware/Software-Accelerated MPI-HMMER Solution*”, Special Issue on Computing Architectures and Acceleration for Bioinformatics Algorithms, The Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology (Elsevier), 48(3), pp. 223-238, 2007.
25. N. Garg*, D. Grosu, and V. Chaudhary, “*An Antisocial Strategy for Scheduling Mechanisms*”, IEEE Transactions on Systems, Man, and Cybernetics – Part A, Vol. 37, No. 6, November 2007, pp. 946-954.
26. J. Eizenkop, D.G. Georgiev, I. Avrutsky, G. Auner, V. Chaudhary, “*Investigation of the formation of nanostructures on silicon thin films after excimer laser irradiation*”, Journal of Physics: Conference Series 59 (2007), pp 458—461.
27. J. Eizenkop, I. Avrutsky, G. Auner, D.G. Georgiev, and V. Chaudhary, “*Single pulse excimer laser nanostructuring of thin silicon films: Nanosharp cones formation and a heat transfer problem*”, Journal of Applied Physics, 101, (2007), pp. 094301—094307.
28. J. Hu**, X. Jin**, J. B. Lee, L. Zhang**, V. Chaudhary, M. Guthikonda, K. H. Yang, and A. I. King, “*Intraoperative Brain Shift Prediction Using a 3D Inhomogeneous Patient-Specific Finite Element Model*”, Journal of Neurosurgery, Vol. 106, No. 1, January 2007, pp. 164—169.
29. X. Meng* and V. Chaudhary, “*Hybrid Parallelism for Sequence Alignment on Linux Clusters*”, in International Journal of Bioinformatics Research and Applications (IJBRA), Volume 2, Issue 4, pp. 430—441, 2006.
30. Gulsheen Kaur*, Jun Tan**, Mohammed Alam*, Vipin Chaudhary, Dingguo Chen**, Ming Dong, Hazem Eltahawy, Farshad Fotouhi, Christopher Gammage*, Jason Gong, William

- Grosky, Murali Guthikonda, Jingwen Hu**, Devkanak Jeyaraj**, Xin Jin**, Albert King, Joseph Landman, Jong Lee, Qing Hang Li, Hanping Lufei**, Michael Morse**, Jignesh Patel, Ishwar Sethi, Weisong Shi, King Yang, and Zhiming Zhang, “*CASMIL: A comprehensive software/toolkit for Image-guided Neurosurgeries*”, International Journal of Medical Robotics and Computer Assisted Surgery, John Wiley & Sons Ltd., Vol. 2, No. 2, June 2006, pp. 118—130.
31. D. Thaker* and V. Chaudhary, “*Simulation Tools to Study a Distributed Shared Memory for Clusters of Symmetric Multiprocessors*”, Future Generation Computer Systems The International Journal of Grid Computing: Theory, Methods, and Applications, Elsevier, Volume 22, Issues 1-2, January 2006, pp. 57-66.
 32. V. Chaudhary, W. L. Hase, H. Jiang*, L. Sun**, and D. Thaker*, “*Experiments with Parallelizing Tribology Simulations*,” Journal of Supercomputing – special issue on High Performance Scientific and Engineering Applications, Vol. 28, pp. 323—343, 2004.
 33. C. Xu*** and V. Chaudhary, “*Time Stamp Algorithms for Runtime Parallelization of DOACROSS Loops with Dynamic Dependences*”, IEEE Transactions on Parallel and Distributed Systems, May 2001, Vol. 12, No. 5, pp. 433 – 450.
 34. S. Roy*, R. Jin***, V. Chaudhary, and W. Hase, “*Parallel Molecular Dynamics Simulations of Adhesive Interactions of Alkanes/Hydroxylated α -Aluminum Oxide Interfaces*”, Computer Physics Communications, pp. 210-218, 128(2000), June 2000.
 35. S. Roy* and V. Chaudhary, “*Design Issues for a High-Performance Distributed Shared Memory on Symmetrical Multiprocessor Clusters*”, Cluster Computing: The Journal of Networks, Software Tools and Applications, pp. 177 - 186, 2 (1999) 3 1999.
 36. V. Chaudhary and J. K. Aggarwal, “*Parallel Image Component Labeling for Target Acquisition*”, in Journal of Optical Engineering – special issue on Target Acquisition, July 1998, Vol. 37, No. 7, pp. 2078—2090.
 37. J. Ju* and V. Chaudhary, “*Unique Sets Oriented Parallelization of Loops with Non-uniform Dependences*”, in The Computer Journal - Special issue on Automatic Loop Parallelization, Vol. 40, No. 6, 1997, pp. 322—339.
 38. S. Punyamurtula*, V. Chaudhary, J. Ju*, and S. Roy*, “*Compile Time Partitioning of Nested Loop Iteration Spaces with Non-Uniform Dependences*”, in Journal of Parallel Algorithms and Applications -- Special issue on Optimizing Compilers for Parallel Languages, Vol. 12, 1997, pp. 113 -- 141.
 39. V. Chaudhary, K. Kumari, P. Arunachalam, and J. K. Aggarwal, “*Manipulation of Octrees and Quadrees on Multiprocessors*”, in International Journal on Pattern Recognition and Artificial Intelligence, Vol. 8, No. 2, April '94, pp. 439—456.
 40. V. Chaudhary and J. K. Aggarwal, “*A generalized scheme for mapping parallel algorithms*”, in IEEE Transactions on Parallel and Distributed Systems, Mar '93, pp. 328 – 346.

Refereed Conference Papers

1. J. Spherhac, R. L. DeLeon, J. P. White, M. Jones, A. E. Bruno, R. Jones-Ivey, T. R. Furlani, J. E. Bard, and V. Chaudhary, “*Towards Performant Workflows, Monitoring and Measuring*”, International Conference on Computer Communication and Networks (ICCCN 2020), Aug 3-6, 2020, Honolulu, Hawaii, USA.

2. H. Liu***, A. Pi, and V. Chaudhary, “*Broad learning-based intervertebral discs localization and segmentation*”, in The Third International Symposium on Image Computing and Digital Medicine (ISICDM 2019), Xi'an, China, Aug. 24-26, 2019.
3. Y. Wei* and V. Chaudhary, “*The Representation of Wave Trend Direction in Stock Price Time Series via RNN*”, 15th International Conference on Machine Learning and Data Mining (MLDM), July 20-25, 2019, New York, USA.
4. Y. Wei* and V. Chaudhary, “*TST: An Effective Approach to Extract Trend Feature in Stock Time Series*”, 7th International Conference on Advances in Computing, Communications and Informatics, Bangalore, India, Sep 19-22, 2018.
5. T. Seth* and V. Chaudhary, “*Exploring Scalable Computing Architectures for Interaction Analysis*”, 27th International Conference on Computer Communications and Networks (ICCCN 2018), July 30- August 2, 2018, Hangzhou, China.
6. S. Chandrashekara*, M. R. Kumar*, M. Venkataramaiah*, and V. Chaudhary, “*Cider: A Case for Block Level Variable Redundancy on a Distributed Flash Array*”, 26th International Conference on Computer Communications and Networks (ICCCN 2017), July 31- August 3, 2017, Vancouver, Canada.
7. Ruhan Sa*, William Owens Jr, Raymond Wiegand, Mark Studin, Donald Capoferri, Kenneth Bahooora, Alexander Greaux, Robbrey Rattray, Adam Hutton, John Cintineo, Vipin Chaudhary, “*Intervertebral Disc Detection in X-Ray Images using Faster R-CNN*”, 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Jeju Island, South Korea, July 2017.
8. S. Liu*, L. Wang, D. Ihrke, V. Chaudhary, C. Tao, C. Weng, and H. Liu, “*Correlating Lab Test Results in Clinical Notes with Structured Lab Data: A Case Study in HbA1c and Glucose*”, Proceedings of 2017 AMIA (American Medical Informatics Association) Joint Summits on Clinical Research Informatics, March 27-30, 2017, San Francisco, USA.
9. R. Sa*, W. Owens, R. Wiegand, and V. Chaudhary, “*Towards an affordable Deep Learning system: automated intervertebral disc detection in x-ray images*”, SPIE Medical Imaging (ORAL presentation), February 11-16, 2017, Orlando, FL.
10. N. Khambekar*, V. Chaudhary, and C. Spooner, “*MUSE: A Methodology for Quantifying Spectrum Usage*”, IEEE GLOBECOM, 2016, December 26-28, Washington, DC, USA.
11. Dingcheng Li, Sijia Liu*, Majid Rastegar-Mojarad, Yanshan Wang, Xiaodi Li, Vipin Chaudhary, Terry Therneau, and Hongfang Liu, “*A Topic-modeling Based Framework for Drug-drug Interaction Classification from Biomedical Text*”, AMIA 2016 Annual Symposium, November 12-16, 2016, Chicago, IL, USA.
12. R. Sa*, W. Owens, R. Wiegand, and V. Chaudhary, “*Fast scale-invariant lateral lumbar vertebrae detection and segmentation in X-ray images*”, 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Orlando, FL, 17-20 August 2016.
13. S. Liu*, H. Liu, V. Chaudhary, and D. Li, “*An Infinite Mixture Model for Coreference Resolution in Clinical Notes*”, Proceedings of 2016 AMIA (American Medical Informatics Association) Joint Summits on Clinical Research Informatics, March 23-24, 2016, San Francisco, USA.
14. N. Khambekar*, C. Spooner, and V. Chaudhary, “*Characterizing of the Missed Spectrum – Access Opportunities under Dynamic Spectrum Sharing*”, The Seventh International Conference on COMMunication Systems and NETworkS (COMSNETS), January 5-9, 2016, Bangalore, India.

15. Y. Wei* and V. Chaudhary, “*The Influence of Sample Reconstruction on Stock Trend Prediction via NARX Neural Network*”, in IEEE 2015 International Conference on Machine Learning and Applications, Dec 9-12, 2015, Miami, USA (to appear).
16. N. Khambekar*, V. Chaudhary, and C. Spooner, “*Estimating the Use of Spectrum for Defining and Enforcing the Spectrum Access Rights*”, IEEE International Conference on Military Communication (MILCOM), 2015, October 26-28, Tampa, Florida, USA, pp. 250-257.
17. N. Khambekar*, C. Spooner, and V. Chaudhary, “*Quantified Discrete Spectrum Access (QDSA) Framework*”, 43rd Research Conference on Communications, Information and Internet Policy, Arlington, VA, USA, Sept 25-27, 2015.
18. Y. Wei* and V. Chaudhary, “*Fast Quantitative Analysis of Stock Trading Points in Dual Period of DMAC*”, IEEE International Conference on Big Data Computing Service and Applications, San Francisco, USA, Mar 30-Apr 2, 2015.
19. S. Liu*, R. Sa*, O. Maguire, H. Minderman, and V. Chaudhary, “*Spot counting on fluorescence in situ hybridization in suspension images using Gaussian mixture model*” SPIE Medical Imaging, Feb 21-26, 2015.
20. C. Feng*, A. Schultz, V. Chaudhary, and D. Kofke, “*Mixed-Precision Models for Calculation of High-Order Virial Coefficients on GPUs*”, in IEEE International Conference on High Performance Computing (HiPC), December 17-20, 2014.
21. N. Khambekar*, C. Spooner, and V. Chaudhary, “*On Improving Serviceability with Quantified Dynamic Spectrum Access*”, IEEE Dynamic Spectrum Access Networks (DySPAN), McLean, VA, USA, April 1-4, 2014.
22. D. Brahme, O. Bhardwaj, and V. Chaudhary, “*SymSig: A low latency interconnection topology for HPC clusters*”, Proceeding of International Conference on High Performance Computing, December 18-21, 2013, Bangalore, India.
23. S. Ghosh*, V. Chaudhary, and G. Dhillon, “*Exploring the utility of axial lumbar MRI for automatic diagnosis of intervertebral disc abnormalities*”, Proceedings of the SPIE Conference on Medical Imaging 2013 (Computer-Aided Diagnosis).
24. J. Koh*, V. Chaudhary, and G. Dhillon, “*An Automated Boundary Detection of the Spinal Canal Using Dynamic Programming*,” EMBC 2012 (Accepted), August 28-September 1, 2012.
25. R. Alomari*, S. Ghosh*, V. Chaudhary and O. Al-Kadi, “*Local binary patterns for stromal area removal in histology images*”, SPIE Medical Imaging, 831524 (February 23, 2012); doi:10.1117/12.911007.
26. Heba Z. Al-Lahham, R. Alomari*, V. Chaudhary, and H. Hiary, “*Automated proliferation rate estimation from Ki-67 histology images*”, Proc. SPIE 8315, Medical Imaging 2012: Computer-Aided Diagnosis, 83152A (February 23, 2012); doi:10.1117/12.911009
27. S. Ghosh* and V. Chaudhary. “*Feature Analysis for Automatic Classification of HEP-2 Florescence Patterns : Computer-Aided Diagnosis of Auto-Immune Diseases*”, 21st International Conference on Pattern Recognition, ICPR 2012, pp. 174-177, Nov 11-15, 2012.
28. S. Ghosh*, M. R. Malgireddy**, V. Chaudhary and G. Dhillon. “*A new approach to automatic disc localization in clinical lumbar MRI: Combining machine learning with heuristics*”, IEEE 9th International Symposium on Biomedical Imaging (ISBI), May 2-5, 2012, Barcelona, Spain.

29. Taruna Seth***, Vipin Chaudhary, Cathy Buyea and Lawrence Bone, “A *Virtual Interactive Navigation System for Orthopaedic Surgical Interventions*”, In Proc. of the 4th Int. Conf. on Applied Sciences in Biomedical and Communication Technologies, ISABEL'11, Spain.
30. Raja' S. Alomari*, Vipin Chaudhary and Gurmeet Dhillon, “*Computer Aided Diagnosis System for Lumbar spine*”, In Proc. of the 4th Int. Conf. on Applied Sciences in Biomedical and Communication Technologies, ISABEL'11, Spain.
31. S. Al-Helo, R. Alomari*, V. Chaudhary, M. B. Al-Zoubi, “*Segmentation of Lumbar Vertebrae from Clinical CT Using Active Shape Models and GVF-Snake*”. Proceedings of the 33rd Annual International IEEE EMBS'11, Sept 2011.
32. Subarna Ghosh*, Raja' S. Alomari*, Vipin Chaudhary and Gurmeet Dhillon, “*Composite Features for Automatic Diagnosis of Intervertebral Disc Herniation from Lumbar MRI*”, In the Proceedings of the 33rd Annual International IEEE EMBC , Sept 2011.
33. J. Koh*, P. Scott, V. Chaudhary and G. Dhillon, “*An Automated Segmentation Method of the Spinal Canal From Clinical MR Images Based on an Attention Model and an Active Contour Model*”, IEEE International Symposium on Biomedical Imaging (ISBI), March 30 - April 2, 2011.
34. S. Ghosh*, R. Alomari*, V. Chaudhary and G. Dhillon, “*Computer-Aided Diagnosis for Lumbar MRI using Heterogeneous Classifiers*”, IEEE International Symposium on Biomedical Imaging (ISBI), March 30 - April 2, 2011.
35. J. Koh*, R. Alomari*, V. Chaudhary, and G. Dhillon, “*Lumbar Spinal Stenosis CAD from Clinical MRM and MRI Based on Inter- and Intra-Context Features with a Two-Level Classifier*”, Proceedings of the SPIE Conference on Medical Imaging 2011 (Computer-Aided Diagnosis), Vol. 7963, 796304 (Podium Presentation).
36. S. Ghosh*, R. Alomari*, V. Chaudhary and G. Dhillon, “*Automatic lumbar vertebra segmentation from clinical CT for wedge compression fracture diagnosis*”, Proceedings of the SPIE Conference on Medical Imaging, 2011 (Computer Aided Diagnosis), Vol. 7963, 796303 (Podium Presentation).
37. J. Koh*, V. Chaudhary, and G. Dhillon, “*A fully automated method of associating axial slices with a disc based on labeling of multi-protocol lumbar MRI*”, IEEE International Conference on Image Processing, September 26-29, 2010, Hong Kong.
38. J. Koh*, T. Kim**, V. Chaudhary, and G. Dhillon, “*Automatic Segmentation of the Spinal Cord and the Dural Sac in Lumbar MR Images Using Gradient Vector Flow Field*”, 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'10), September 1-4, 2010, Buenos Aires, Argentina.
39. J. Koh*, V. Govindaraju, and V. Chaudhary, “*A Robust Iris Localization Method Using an Active Contour Model and Hough Transform*”, 20th International Conference on Pattern Recognition, August 23-26, 2010, Istanbul, Turkey.
40. R. Alomari*, V. Chaudhary, and G. Dhillon, “*Lumbar disc herniation CAD with GVF-snake model*”, Computer Assisted Radiology and Surgery – 24th International Congress and Exhibition, June 23-26, 2010, Geneva, Switzerland. (to appear).
41. R. Alomari*, J. J. Corso, V. Chaudhary, and G. Dhillon, “*Automatic diagnosis of lumbar disc herniation using shape and appearance features from MRI*”, SPIE Medical Imaging Conference, Feb 2010, San Diego, CA.
42. J. Koh*, V. Chaudhary, and G. Dhillon, “*Diagnosis of disc herniation based on classifiers and features generated from spine MR images*”, SPIE Medical Imaging Conference, Feb 2010, San Diego, CA.

43. J. Delmerico*, N. Byrnes, A. Bruno, M. Jones, S. Gallo, and V. Chaudhary, “*Comparing the performance of Clusters, Hadoop, and Active Disks on Microarray Correlation Computations*”, International Conference on High Performance Computing, 2009, pp.378-387.
44. N. Mehta*, R. Alomari*, and V. Chaudhary, “*Content Based Sub-Image Retrieval System for High Resolution Pathology Images using Salient Interest Points*”, Annual Conference of IEEE Engineering in Medicine and Biology Society, September 2009, Minneapolis, MN.
45. R. Alomari*, J.J. Corso, V. Chaudhary, and G. Dhillon, “*Desiccation Diagnosis in Lumbar Discs from Clinical MRI with a Probabilistic Model*”, IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2009), June 28 – July 1, 2009, Boston, pp. 546-549.
46. R. Alomari*, J.J. Corso, V. Chaudhary, and G. Dhillon, “*Abnormality detection in lumbar discs from clinical MR images with a probabilistic model*”, Computer Assisted Radiology and Surgery – 23rd International Congress and Exhibition, June 23-27, 2009, Berlin, Germany. (Podium Presentation).
47. J. P. Walters*, R. Darole*, and V. Chaudhary, “*Improving MPI-HMMER’s Scalability with Parallel I/O*”, IEEE/ACM International Parallel and Distributed Processing Symposium (IPDPS), May 25-29, 2009, Rome, Italy.
48. J. P. Walters*, V. Balu*, S. Kompalli***, and V. Chaudhary, “*Evaluating the use of GPUs in Liver Image Segmentation and HMMER Database Searches*”, IEEE/ACM International Parallel and Distributed Processing Symposium (IPDPS), May 25-29, 2009, Rome, Italy.
49. N. Khambekar*, C. Spooner, and V. Chaudhary, “*Listen-while-Talking: A Technique for Primary User Protection*”, IEEE Wireless Communications and Networking Conference, April 5-8, 2009, Budapest, Hungary.
50. R. Alomari*, R. Allen, B. Sabata, and V. Chaudhary, “*Localization of tissues in high resolution digital anatomic pathology images*”, SPIE Medical Imaging 2009, Feb 10-12, 2009, Florida (Podium Presentation).
51. C. Bhole*, S. Kompalli***, and V. Chaudhary, “*Context-sensitive labeling of spinal structures in MRI images*”, SPIE Medical Imaging 2009, Feb 10-12, 2009, Florida.
52. X. Meng and V. Chaudhary, “*Heterogeneous Computing for Biological Sequence Database Searches*”, 7th Asia-Pacific Bioinformatics Conference (APBC2009), Beijing, China, January 13-16, 2009.
53. N. Mehta*, S. Kompalli***, and V. Chaudhary, “*Web-based Architecture to Enable High-performance CAD tools and Multi-user Synchronization*”, eHealth 2008, London, UK.
54. J. Corso, R. Alomari*, and V. Chaudhary, “*Lumbar Disc Localization and Labeling with a Probabilistic Model on both Pixel and Object Features*”, 11th International Conference on Medical Image Computing and Computer Assisted Intervention, (MICCAI 2008), New York, September 6-10, 2008.
55. J. P. Walters**, V. Balu*, V. Chaudhary, D. Kofke, and A. Schultz, “*Accelerating Molecular Dynamics Simulations with GPUs*”, Proceedings of the 21st International Society for Computers and their Applications, Parallel and Distributed Computing and Communication Systems (ISCA-PDCCS’08), New Orleans, LA, 2008.
56. H. Lufei**, W. Shi, and V. Chaudhary, “*Adaptive Secure Access to Remoter Services*”, 2008 IEEE International Conference on Services Computing (SCC 2008), July 8-11, 2008, Hawaii, USA.

57. J. P. Walters*, V. Chaudhary, M. Cha, S. Guercio Jr, and S. Gallo, “A Comparison of Virtualization Technologies for HPC”, 22nd International Conference on Advanced Information Networking and Applications, Okinawa, Japan, March 2008.
58. S. Kompalli***, R. Alomari*, and V. Chaudhary, “Segmentation of the Liver from Abdominal CT Using Multi-featured Markov Random Fields”, Second International Conference on Complex, Intelligent and Software Intensive Systems, Barcelona, Spain, March 2008.
59. R. Alomari*, S. Kompalli***, S. Lau, and V. Chaudhary, “Design of a Benchmark Dataset, Similarity Metrics, and Tools for Liver Segmentation”, SPIE Medical Imaging Conference, February 2008, San Diego, CA.
60. I. Inwogu* and V. Chaudhary, “Enhancing Regional Lymph Nodes from Endoscopic Ultrasound Images”, SPIE Medical Imaging Conference, February 2008, San Diego, CA.
61. J. P. Walters* and V. Chaudhary, “A Scalable Asynchronous Replication based Strategy for Fault Tolerant MPI Applications,” IEEE International Conference on High Performance Computing, 2007, Goa, India, pp. 257-268.
62. R. Alomari*, S. Kompalli***, and Vipin Chaudhary, “Liver Segmentation from Abdominal CT images using GVF snake,” 2007 Western New York Image Processing Workshop, IEEE Signal Processing Society, September 2007, Rochester, USA.
63. J. P. Walters* and V. Chaudhary, “FT-OpenVZ: A Virtualized Approach to Fault-Tolerance in Distributed Systems,” International Conference on Parallel and Distributed Systems, Sep 24-26, Las Vegas, USA, 2007, pp. 85--90.
64. N. Khambekar*, L. Dong, and V. Chaudhary, “Utilizing OFDM Guard Interval for Spectrum Monitoring”, in Proceedings of IEEE Wireless Communications and Networking Conference, 2007, Hong Kong.
65. X. Meng* and V. Chaudhary, “An Adaptive Data Prefetching Scheme for Biosequence Database Search on Reconfigurable Platforms”, ACM Symposium on Applied Computing, Seoul, Korea, March 11—15, 2007, pp. 140—141.
66. J. P. Walters* and V. Chaudhary, “Application-Level Checkpointing Techniques for Parallel Programs”, Proceedings of the 3rd International Conference on Distributed Computing and Internet Technology, December 20-23, 2006, Bhubaneswar, India, pp. 221 – 234. (invited paper).
67. H. Liu**, H. Lufei**, W. Shi, and V. Chaudhary, “Towards Ubiquitous Access of Computer-Assisted Surgery Systems”, in Proceedings of 28th Annual International Conference of IEEE Engineering in Medicine and Biology Society (EMBS), August 30-September 3, 2006, New York City, USA.
68. J. Ma, Q. Zhao, V. Chaudhary, J. Cheng, L. T. Yang, R. Huang, and Q. Jin, “Ubisafe Computing: Vision and Challenges (I)”, in 3rd International Conference on Autonomic and Trusted Computing (ATC-06), September 3-6, 2006, Wuhan and Three Gorges, China, pp. 386-397 (Lecture Notes in Computer Science 4158, Springer).
69. M. Nanjundaiah* and V. Chaudhary, “A simulation study comparing the performance of two RFID protocols”, in 3rd International Conference on Ubiquitous Intelligence and Computing (UIC-06), September 3-6, 2006, Wuhan and Three Gorges, China, pp. 679—687.
70. J. Tan**, D. Chen**, V. Chaudhary, and I. Sethi, “A Template Based Technique For Automatic Detection Of Fiducial Markers In 3D Brain Images”, in CARS 2006 (Computer Assisted Radiology and Surgery, 20th International Congress and Exhibition), June 28 - July 1, 2006, Osaka, Japan, pp. 47—49.

71. H. Lufei**, W. Shi, and V. Chaudhary, “*M-CASEngine: a collaborative environment for computer-assisted surgery*”, in CARS 2006 (Computer Assisted Radiology and Surgery, 20th International Congress and Exhibition), June 28 - July 1, 2006, Osaka, Japan, pp. 447—449. (Podium Presentation)
72. D. Chen**, V. Chaudhary, and I. Sethi, “*Automatic fiducial localization in brain images*”, in CARS 2006 (Computer Assisted Radiology and Surgery, 20th International Congress and Exhibition), June 28 - July 1, 2006, Osaka, Japan, pp. 45—47.
73. D. Chen**, V. Chaudhary, and I. Sethi, “*3D digital brain atlas construction using Talairach & Tournoux 88*”, in CARS 2006 (Computer Assisted Radiology and Surgery, 20th International Congress and Exhibition), June 28 - July 1, 2006, Osaka, Japan, pp. 461.
74. J. P. Walters*, B. Qudah*, and V. Chaudhary, “*Accelerating the HMMER Sequence Analysis Suite using Conventional Processors*”, IEEE 20th International Conference on Advanced Information Networking and Applications, April 18-20, 2006, Vienna, Austria, pp. 289–294.
75. A. Nambiar* and V. Chaudhary, “*On Tools for Modeling High-Performance Embedded Systems*”, in Proc. of Intl. Conf. on Embedded and Ubiquitous Computing, Dec 2005, Nagasaki, Japan, pp. 360—370.
76. J. Hu**, X. Jin**, J. B. Lee, L. Zhang**, V. Chaudhary, K. H. Yang, and A. I. King, “*A 3D Patient-Specific Finite Element Model For Predicting Brain Shift During Neurosurgery*”, Proceedings of the 2005 BMES Annual Fall Meetings, Baltimore, MD, Sept. 28 – Oct. 1, 2005.
77. A. Nambiar* and V. Chaudhary, “*Mapping Resource Constrained Applications on Chip Multiprocessors*”, International Conference on Embedded Systems and Applications, Las Vegas, NV, June 2005, pp. 117-123.
78. Y. Ji*, H. Jiang*, and V. Chaudhary, “*Adaptation Strategies for Application-Level Computation Migration/Checkpointing*”, International Conference of Parallel and Distributed Processing Techniques and Applications, Las Vegas, NV, June 2005, pp. 1156-1162.
79. Y. Ji*, H. Jiang*, and V. Chaudhary, “*Adaptation Point Analysis for Computation Migration/Checkpointing*”, in ACM Symposium on Applied Computing, March 2005, Santa Fe, New Mexico, pp. 750-751.
80. V. Chaudhary and H. Jiang*, “*Migrating Processes/Threads in Grids*”, International Conference and Exposition on Communications and Computing, 28(3), IIT Kanpur, India, February 4-6, 2005.
81. G. Yadav*, R. K. Singh, V. Chaudhary, “*Software-only Multiple Variable Length Decoding for Real-Time Video on MDSP*”, Proc. of IEEE International Conference on Consumer Electronics, Jan 2005, Las Vegas.
82. G. Yadav*, R. K. Singh, V. Chaudhary, “*MAVD: MPEG-2 Audio Video Decode System on MDSP*”, in Proc. of IEEE Intl. Symposium on Consumer Electronics, pp. 19-24, Sept 2004, Reading, UK.
83. G. Yadav*, R. K. Singh, V. Chaudhary, “*On Implementation of MPEG-2 like Real-Time Parallel Media Applications on MDSP SoC Cradle Architecture*”, in Proc. of Intl. Conf. on Embedded and Ubiquitous Computing, pp. 281-290, Aug 2004, Aizu, Japan.
84. X. Meng* and V. Chaudhary, “*Bio-Sequence Analysis with Cradle’s 3SoC Software Scalable System on Chip*”, Proceedings of the 19th ACM Symposium on Applied Computing, Cyprus, March 2004, pp. 202—206.
85. H. Jiang* and V. Chaudhary, “*Process/Thread Migration and Checkpointing in Heterogeneous Distributed Systems*”, in Proceedings of the 37th Hawaii International

- Conference on System Sciences (HiCSS-37), IEEE Computer Society, Big Island, Hawaii, January 5-8, 2004.
86. H. Jiang* and V. Chaudhary, “*Thread Migration/Checkpointing for Type-Unsafe C Programs*”, Proceedings of ACM/IEEE International Conference on High Performance Computing (HiPC), Hyderabad, India, December 2003, pp. 469—479.
 87. H. Jiang*, V. Chaudhary, and J. P. Walters, “*Data Conversion for Process/Thread Migration and Checkpointing*”, Proceedings of the International Conference on Parallel Processing (ICPP), Kaohsiung, Taiwan, October 6-9, 2003, pp. 473 – 480.
 88. F. Liu* and V. Chaudhary, “*Extending OpenMP for Heterogeneous Chip Multiprocessors*”, Proceedings of the International Conference on Parallel Processing, Kaohsiung, Taiwan, October 2003, pp. 161 – 168.
 89. N. Ghate* and V. Chaudhary, “*Optimizing Automatically Generated Programs for a Software Distributed Shared Memory System*”, International Conference on Parallel and Distributed Computing and Systems, November 2002, pp. 735-744.
 90. H. Jiang* and V. Chaudhary, “*On Improving Thread Migration: Safety & Performance*”, ACM/IEEE International Conference on High Performance Computing, December 2002, pp. 474—484.
 91. M. Shah*, V. Chaudhary, and G. Edjlali***, “*Policy Based User Configurable Java Security Architecture*”, International Conference on Security and Management, June 2002, pp. 376—382.
 92. H. Jiang* and V. Chaudhary, “*Compile/Run time support for Thread Migration*”, ACM/IEEE International Parallel and Distributed Processing Symposium, April 2002, pp. 58--66.
 93. J. Ju* and V. Chaudhary, “*A Fission Technique Enabling Parallelization of Non-Perfectly Nested Loops*”, ACM/IEEE International Conference on High Performance Computing, Dec. 99, pp. 87—94.
 94. D. Reimann*, V. Chaudhary, and I. K. Sethi, “*Modeling Cone-Beam Tomographic Reconstruction Using LogSMP: An Extended LogP Model for Cluster of SMPs*”, ACM/IEEE International Conference on High Performance Computing, Dec. 99, pp. 77—83.
 95. V. Chaudhary, C. Xu***, S. Roy*, S. Jia*, G. Ezzell, and C. Kota, “*Parallelization of Radiation Therapy Treatment Planning (RTTP): A Case Study*”, International Conference on Parallel and Distributed Computing Systems, 1999, pp. 534 – 539.
 96. S. Roy*, V. Chaudhary, S. Jia*, and P. Menon*, “*Application Based Evaluation of Distributed Shared Memory Versus Message Passing*”, International Conference on Parallel and Distributed Computing Systems, 1999, pp. 15 – 20.
 97. V. Chaudhary, G. Edjlali***, S. Roy*, and D. Thaker*, “*Cost-Performance Evaluation of SMP Clusters*”, International Conference on Parallel and Distributed Processing Techniques and Applications, 1999, pp. 718 – 724.
 98. S. Roy* and V. Chaudhary, “*Evaluation of Cluster Interconnects for a Distributed Shared Memory*”, Proc. of the IEEE International Performance, Computing, and Communication Conference, 1999, pp. 1—7.
 99. S. Roy* and V. Chaudhary, “*Communication Requirements of Software Distributed Shared Memory Systems*”, Proc. of IEEE National Conference on Communications, 1999, pp. 409—416.
 100. S. Roy* and V. Chaudhary, “*Strings: A High-Performance Distributed Shared Memory for Symmetrical Multiprocessor Clusters*”, Proc. of IEEE Conf. on High Performance Distributed Computing, July 1998, pp. 90 – 97.

101. G. Edjlali***, A. Acharya, and V. Chaudhary, "History-Based Access Control for Mobile Code", ACM Conference on Computer and Communication Security, November 1998, pp. 38—48.
102. P. Menon*, V. Chaudhary, and J. Pipe, "Parallel Algorithms for Deblurring of MR Images", International Conference on Computers and Their Applications, March 1998, pp. 266—269.
103. V. Chaudhary and M. Ahluwalia*, "PEST: A system for evaluating parameters in a load balancing system", International Conference on Distributed Processing and Networking, December 1997, pp. 48-52.
104. V. Chaudhary, C. Xu***, J. Shi*, D. Hu*, G. Ezzell, and C. Kota, "Experiences on the parallelization of radiation therapy treatment planning", International Conference on Advanced Computing, December 1997, pp. 135-141.
105. C. Xu*** and V. Chaudhary, "Time-stamping algorithms for parallelization of loops at run-time", in Proceedings of the ACM/IEEE International Parallel Processing Symposium, April 1997, (8 pages).
106. S. Roy* and V. Chaudhary, "A New Metric for Processor Allocation Schemes in Multiprocessor Systems", in Proceedings of the IEEE International Performance, Computing, and Communications Conference, Feb. 1997, pp. 42--48.
107. S. Roy* and V. Chaudhary, "Parallelization of 3-D Range Image Segmentation on a SIMD Multiprocessor", in Proceedings of the International Conference on Computers and Their Applications, Mar. 1997, pp. 106--109.
108. D. Reimann*, M. J. Flynn, V. Chaudhary, and I. K. Sethi, "Parallel Computing Methods for X-Ray Cone Beam Tomography with Large Array Sizes", in Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference Record, November 1996, pp. 1710--1713.
109. D. Reimann*, V. Chaudhary, M. J. Flynn, I. K. Sethi, "Cone Beam Tomography using MPI on Heterogeneous Workstation Cluster", in Proceedings of the Second MPI Developers Conference, 1996, pp. 142--148.
110. J. Ju* and V. Chaudhary, "Unique Sets Based Partitioning of Nested Loops with Non-Uniform Dependences", in Proceedings of the International Conference on Parallel Processing, 1996, Vol III, pp. 45--52.
111. D. Reimann*, V. Chaudhary, M. J. Flynn, I. K. Sethi, "Parallel Implementation of Cone Beam Tomography", in Proceedings of the International Conference on Parallel Processing, 1996, Vol. II, pp. 170--173.
112. A. Alrabady**, S. M. Mahmud, and V. Chaudhary, "Placement of Resources in the Star Network", in Proceedings of the IEEE International Conference On Algorithms and Architectures for Parallel Processing, 1996.
113. V. Chaudhary, C. Z. Xu***, S. Roy*, J. Ju*, V. Sinha*, and L. Luo*, "Design and evaluation of an environment APE for Automatic parallelization of Programs", in Proceedings of the IEEE International Symposium on Parallel Architectures, Algorithms, and Networks, June 1996, 77--83.
114. G. Dommety*, V. Chaudhary, and B. Sabata, "Strategies for Processor allocation in k-ary n-cubes", in Proceedings of the International Conference on Parallel and Distributed Computing Systems, 1995, pp. 216--221.
115. C. Vinod* and V. Chaudhary, "A case study of parallel hierarchical radiosity algorithms on a DM-COMA architecture", in Proceedings of the International Conference on Parallel and Distributed Computing Systems, 1995, pp. 236--241.

116. S. P. Rana, K. Raman*, and V. Chaudhary, ``*Migrating controller based framework for mutual exclusion in distributed systems*'', in Proceedings of the IEEE International Performance, Computers, and Communications Conference, 1995, pp. 1--7.
117. R. Srinivasan*, V. Chaudhary, and S. M. Mahmud, ``*Contention sensitive fault-tolerant routing algorithms for hypercubes*'', in Proceedings of the IEEE International Symposium on Parallel Architectures, Algorithms, and Networks, 1994, pp. 197--204.
118. S. Punyamurtula* and V. Chaudhary, ``*Minimum dependence distance tiling of nested loops with non-uniform dependences*'', in Proceedings of the IEEE Symposium on Parallel and Distributed Processing, 1994, pp. 74--81.
119. S. Ponnuswamy* and V. Chaudhary, ``*Embedding Hamiltonians and cycles in rotator and cycle prefix digraphs*'' in Proceedings of the IEEE Symposium on Parallel and Distributed Processing, 1994, pp. 603--610.
120. T. Samaratunga**, R. Srinivasan*, V. Chaudhary, and S. M. Mahmud, ``*An optimal mapping algorithm for HIN-Based multiprocessor systems*'', in Proceedings of the International Conference on Parallel and Distributed Computing Systems, 1994, pp. 706--711.
121. S. Ponnuswamy* and V. Chaudhary, ``*Analysis of fault tolerance in Cayley digraphs using forbidden faulty sets*'' in Proceedings of the International Conference on Parallel and Distributed Computing and Systems, 1994, pp. 346--349.
122. S. Ponnuswamy* and V. Chaudhary, ``*A comparative study of star graphs and rotator graphs*'' in Proceedings of the International Conference on Parallel Processing, 1994, Vol. I, pp. 46-50.
123. V. Gautam* and V. Chaudhary, ``*Subcube allocation strategies in a k-ary n-cube*'', in Proceedings of the International Conference on Parallel and Distributed Computing Systems, pp. 141--146, 1993.
124. V. Chaudhary, ``*Mapping algorithms for permutation networks*'', in Proceedings of the IEEE Midwest Conference on Circuits and Systems, August 93, pp. 323--326.
125. S. Ponnuswamy* and V. Chaudhary, ``*Embedding meshes in rotator graphs*'', in Proceedings of the IEEE Midwest Conference on Circuits and Systems, August 93, pp. 5--8.
126. N. Ramesh* and V. Chaudhary, ``*Complexity analysis of range image segmentation on MasPar MP-1*'' in Proceedings of the IEEE Midwest Conference on Circuits and Systems, August 93, pp. 903--906.
127. V. Chaudhary, B. Sabata, and J. K. Aggarwal, ``*Mapping interconnection networks into VEDIC networks*'', in Proceedings of the IEEE/ACM International Parallel Processing Symposium, 1993, pp. 531--537.
128. V. Chaudhary, K. Kumari, P. Arunachalam, and J. K. Aggarwal, ``*Parallel manipulations of octrees and quadtrees*'', in Proceedings of Second International Conference on Parallel Image Analysis, Ube, Japan, Lecture Notes in Computer Science, pp. 69--86, Springer-Verlag, 1992.
129. V. Chaudhary, B. Sabata, and J. K. Aggarwal, ``*Deadlock-free multicast wormhole routing in VEDIC networks*'', poster at International Conference on Parallel Processing, 92.
130. V. Chaudhary and J. K. Aggarwal, ``*On the complexity of parallel image component labeling*'', in Proceedings of the International Conference on Parallel Processing, August 1991, Vol. III, pp. 183 -- 187.
131. V. Chaudhary, B. Sabata, and J. K. Aggarwal, ``*The VEDIC network for multicomputers*'' , in Proceedings of the International Conference on Parallel Processing, August 1991, Vol. I, pp. 686 -- 687.

132. V. Chaudhary and J. K. Aggarwal, "Generalized mapping of parallel algorithms onto parallel Architectures", in Proceedings of the International Conference on Parallel Processing, August 1990, Vol. II, pp. 137 -- 141.

Refereed Workshop Papers

1. S. Liu*, L. Wang, V. Chaudhary, and H. Liu, "Attention Neural Model for Temporal Relation Extraction", 2nd Clinical Natural Language Processing Workshop (with NAACL 2019), Minneapolis, USA, June 7, 2019.
2. S. Liu*, F. Shen, Y. Wang, M. Rastegar-Mojarad, R. K. Elayavilli, V. Chaudhary, and H. Liu, "Attention-based Neural Networks for Chemical Protein Relation Extraction", BioCreative VI Workshop Proceedings, October 2017, Washington DC.
3. S. Liu*, F. Shen, V. Chaudhary, and H. Liu, "MayoNLP at SemEval 2017 Task 10: Word Embedding Distance Pattern for Keyphrase Classification in Scientific Publications", International Workshop on Semantic Evaluation (SemEval-2017), held with Annual Meeting of the Association for Computational Linguistics (ACL), August, 2017, Vancouver, Canada. (Top system in participated evaluation scenario, [tweet](#))
4. R. Shivaswamy, A. Patra, and V. Chaudhary, "Large Data and Computation in a Hazard Map Workflow Using Hadoop and Netezza Architectures", International Workshop on Data-Intensive Scalable Computing Systems (DISCS-2013), in conjunction with 2013 ACM/IEEE Supercomputing Conference (SC'13), Nov 2013.
5. R. Alomari*, V. Chaudhary, J. Corso, and G. Dhillon, "Lumbar Spine Disc Herniation Diagnosis with a Joint Shape Model", MICCAI 2013 Workshop on "Computational Methods and Clinical Applications for Spine Imaging", Nagoya, Japan, 2013.
6. S. Ghosh*, M. Malgireddy, V. Chaudhary, G. Dhillon, "A supervised approach towards segmentation of clinical MRI for automatic lumbar diagnosis", MICCAI 2013 Workshop on "Computational Methods and Clinical Applications for Spine Imaging", Nagoya, Japan, 2013.
7. V. Patil*, V. Chaudhary, "Rack Aware Scheduling in HPC data centers", 7th Workshop on High-Performance Power Aware Computing (HPPAC), IEEE International Parallel and Distributed Processing Symposium (IPDPS), 2011, Anchorage, Alaska, USA.
8. Ata E. H. Bohra* and V. Chaudhary, "VMeter: Power Modelling for Virtualized Clouds", 6th High Performance and Power Aware Computing (HPPAC), IEEE International Parallel and Distributed Processing Symposium (IPDPS). April 19-23, Atlanta, Georgia, USA.
9. J. P. Walters*, B. Bantwal,* and V. Chaudhary, "Enabling Interactive Jobs in Virtualized Data Centers", First International Workshop on Cloud Computing and Its Applications, Chicago, IL, October 22-23, 2008.
10. Brent Rood, John Paul Walters*, Vipin Chaudhary, and Michael J. Lewis, "Failure Prediction and Scalable Checkpointing for Reliable Large-Scale Grid Computing", HPDC-16: The 16th IEEE International Symposium on High Performance Distributed Computing (Hot Topics Session), Monterey, California, June 27-29 2007.
11. J. P. Walters*, H. Jiang*, and V. Chaudhary, "An Adaptive Heterogeneous Software DSM", in 5th International Workshop on Compile and Runtime Techniques for Parallel Computing (CRTPC 2006) with 2006 International Conference on Parallel Processing, August 14-18, 2006, Columbus, Ohio, pp. 265 – 272.
12. C. Gammage* and V. Chaudhary, "On Optimization and Parallelization of Fuzzy Connected Segmentation for Medical Imaging", 2nd IEEE International Workshop on Workshop on High

- Performance Computing in Medicine and Biology (HiPCoMB 2006)*, held with IEEE 20th International Conference on Advanced Information Networking and Applications, April 18-20, 2006, Vienna, Austria, pp. 623 – 627.
13. V. Chaudhary, C. Gammage*, and J. Landman, “*Opportunities for Parallel/Distributed Computing in Surgery*”, 13th DPS Workshop, Okinawa, Japan, November 2005, pp. 240-244.
 14. X. Meng* and V. Chaudhary, “*Exploiting Multi-level Parallelism for Homology Search using General Purpose Processors*”, Proceedings of the 1st IEEE International Workshop on High Performance Computing for Medicine and Biology, with 11th International Conference on Parallel and Distributed Systems (ICPADS 2005), 20-22 July 2005 in Fukuoka, Japan, pp. 331-335.
 15. N. Garg*, D. Grosu, and V. Chaudhary, “*An Antisocial Strategy for Scheduling Mechanisms*”, Proc. of the 19th IEEE International Parallel and Distributed Processing Symposium, 7th Workshop on Advances in Parallel and Distributed Computational Models (APDCM'05), April 4-8, 2005, Denver, Colorado, USA.
 16. M. Nanjundaiah* and V. Chaudhary, “*Improvement to the Anticollision Protocol Specification for 900 MHz Class 0 Radio Frequency Identification Tag*”, in International Workshop on Ubiquitous Smart Worlds, held in conjunction with IEEE Conference on Advanced Information Networking and Applications, March 2005, Taipei, Taiwan, pp. 616-620.
 17. F. Liu* and V. Chaudhary, “*A Practical OpenMP Compiler for System on Chips*”, Proceedings of International Workshop on OpenMP Applications and Tools, June 2003, pp. 54 – 68.
 18. L. Zhang* and V. Chaudhary, “*On the performance of Bus Interconnection for SOCs*”, ACM 4th Workshop on Media and Stream Processors, Nov. 2002, pp. 1 – 9.
 19. H. Jiang* and V. Chaudhary, “*MigThread: Thread Migration in DSM Systems*”, Workshop on Compile/Runtime Techniques for Parallel Computing, held with International Conference on Parallel Processing, August 2002, pp. 581—588.
 20. V. Chaudhary, W. L. Hase, H. Jiang*, L. Sun**, and D. Thaker*, “*Experiments with Parallelizing a Tribology application*”, Workshop on High Performance Scientific and Engineering Computing with Applications, held with International Conference on Parallel Processing, August 2002, 344—351.
 21. V. Chaudhary, J. Ju*, L. Luo*, S. Roy*, V. Sinha*, C. Xu***, and V. Konda, “*Automatic parallelization of non-uniform dependences*”, in Proceedings of the First SUIF Compiler Workshop, 1996, pp. 148--152.

Abstracts with Podium Talks

1. J. J. Corso, J. A. Delmerico, P. David, R. Alomari, V. Chaudhary. “*Layered Models for Bridging from Low to High Level Vision*”, Frontiers in computer vision workshop, Massachusetts Institute of Technology, Aug, 2011. (poster)
2. V. Chaudhary, “*Heterogeneous Checkpointing of HPC Applications*”, in SIAM Conference on Mathematics for Industry: Challenges and Frontiers, Oct, 2005, Detroit, MI.
3. J. Eizenkop*, D.G. Georgiev, I. Avrutsky, G. Auner, and V. Chaudhary, “*Investigation of the formation of nanostructures on silicon thin films after excimer laser irradiation*”, 8th International Conference on Laser Ablation, September 11-16, 2005, Banff, Alberta, Canada, Printed in Journal of Physics, Conference Series.

PATENTS AND INTELLECTUAL PROPERTY DISCLOSURES

- 1 “*Methods and systems for spectrum management*”, N. Khambekar (University at Buffalo, SUNY), V. Chaudhary (University at Buffalo, SUNY), and C. Spooner (Northwestern Research Associates), **US Patent US15335429**, October 2016, (with additional application **US20170208476A1**), July 2017.
- 2 “*COCOA: command and control under dynamic spectrum sharing paradigm*”, N. Khambekar (University at Buffalo, SUNY) and V. Chaudhary (University at Buffalo, SUNY), Invention Disclosure: August 27, 2015.
- 3 “*HONEYCOMB: a real-time infrastructure for situational awareness and adaptation*”, N. Khambekar (University at Buffalo, SUNY) and V. Chaudhary (University at Buffalo, SUNY), Invention Disclosure: August 27, 2015.
- 4 “*COIN: A new model for spectrum commerce that brings in simplicity, precision, and efficiency in the spectrum trade*”, N. Khambekar (University at Buffalo, SUNY) and V. Chaudhary (University at Buffalo, SUNY), Invention Disclosure: August 27, 2015.
- 5 “*FLUX: An agile responsive spectrum management framework*”, N. Khambekar (University at Buffalo, SUNY) and V. Chaudhary (University at Buffalo, SUNY), Invention Disclosure: August 21, 2015.
- 6 “*SPHERE: A real time infrastructure for dynamic spectrum access, regulation, and management*”, N. Khambekar (University at Buffalo, SUNY), Chad Spooner (Northwestern Research Associates), and V. Chaudhary (University at Buffalo, SUNY), Invention Disclosure: August 21, 2015.
- 7 “*RAMP: Radio-environment Awareness Maps for understanding of the use of spectrum and performance of the spectrum functions*”, N. Khambekar (University at Buffalo, SUNY), V. Chaudhary (University at Buffalo, SUNY), and C. Spooner (Northwestern Research Associates), Invention Disclosure: August 21, 2015.
- 8 “*Spectrum-space Digitization*”, N. Khambekar (University at Buffalo, SUNY), V. Chaudhary (University at Buffalo, SUNY), and C. Spooner (Northwestern Research Associates), Invention Disclosure: August 21, 2015.
- 9 “*Quantified Dynamic Spectrum Access*”, N. Khambekar (University at Buffalo, SUNY), V. Chaudhary (University at Buffalo, SUNY), and C. Spooner (Northwestern Research Associates), Invention Disclosure: August 21, 2015.
- 10 “*MUSE: A Methodology for characterizing and quantifying the use of the spectrum in the space, time, and frequency by individual transceivers*”, N. Khambekar (University at Buffalo, SUNY), V. Chaudhary (University at Buffalo, SUNY), and C. Spooner (Northwestern Research Associates), Invention Disclosure: August 21, 2015.
- 11 “*System and method for fault-tolerant block data storage*”, S. Chandrasekhara (University at Buffalo, SUNY), M. R. Kumar (University at Buffalo, SUNY), and V. Chaudhary (University at Buffalo, SUNY), US Patent (and International) Application Number: **PCT/US2015/026267**, April 2015.
- 12 “*Throughput Enhancing Inband Sensing Technique for Cognitive Radio Devices Employing Dynamic Spectrum Access*”, N. Khambekar (University at Buffalo, SUNY), V. Chaudhary (University at Buffalo, SUNY), C. Spooner (Northwestern Research Associates), and L. Dong (Western Michigan University, Kalamazoo), Provisional Patent filed 2009.

- 13 “*Radiology Viewing System and Method*”, V. Chaudhary, S. Kompalli, C. Gammage, M. Yaqub, and M. Alam, Patent Pending, University at Buffalo, SUNY. Patent Licensed by Medcotek, Inc., 2008.
- 14 “*Heterogeneous Resource Management for High Performance Computational Cluster*”, V. Chaudhary, J. P. Walters and B. Bantwal, University at Buffalo, SUNY; provisional patent filed, 2008.
- 15 “*Context Sensitive Technique for labeling of Spinal and Vertebral Structure*”, V. Chaudhary, S. Kompalli and C. Bhole, University at Buffalo, SUNY; provisional patent filed, 2008. Licensing deal in progress.
- 16 “*Multilevel Automatic Ratio-based Medical Image Segmentation*”, Dingguo Chen, Vipin Chaudhary and Ishwar Sethi (Invention Disclosure to Wayne State University).
- 17 “*Haptic Interface for multi-resolution navigation, manipulation, and annotation of image data*”, Joseph Landman and Vipin Chaudhary (Invention Disclosure to Wayne State University).

SOFTWARE SYSTEMS DEVELOPED

1. GPU-HMMER: Graphics Processor implementation of HMMER that works for multiple GPUs and achieves over 100x speedup. It is being used at every major bioinformatics site. Three companies have created hardware specifically to run GPU-HMMER and are selling that. It is available at <http://www.mpihmmmer.org>
2. MPI-HMMER: HMMER is a freely distributable implementation of profile HMM software for protein sequence analysis. MPI-HMMER is a multiple-level optimization of the original HMMER 2.3.2 code that consists of two distinct optimizations: a portably tuned P7Viterbi function as well as an MPI implementation. The MPI implementation exhibits excellent speedups over the base PVM implementation available with HMMER. Further, a verification mode in both hmmpfam and hmmsearch is provided that ensures (at a cost of speed) results are returned in exactly the same order as the serial version. It is already downloaded by over 1000 institutions and groups worldwide. The next version will also incorporate working with specific FPGA platforms. <http://www.mpihmmmer.org>
3. FAST-FASTA: This is a multi-level parallel implementation of FASTA that also incorporates FPGA platforms. FASTA is used for local bio-sequence homology searches and is the most widely used implementation of Smith-Waterman algorithm. This implementation was made open source for public distribution.
4. CASMIL: Computer Assisted Surgery system being developed. Clinical study starting 9/06.
5. ADAM: A Distributed Adaptively-shared Memory System that adds adaptivity to software distributed shared memories.
6. MigThread: This is a thread migration tool that works for a heterogeneous environment and handles complex pointers.
7. DSMSim: This is a simulator for Strings DSM that helps in improving performance of DSMs.
8. APE: This is an automatic parallelization environment that takes as input Fortran/C programs and generates code that executes on a cluster of SMPs.
9. Pest: This is a tool for evaluating the parameters for load balancing in a network of computers. It is implemented in Java. This is an outcome of NSF-RIA.

10. Strings: This is a fully multi-threaded (kernel and user-level threads) Distributed Shared Memory available for SMP clusters. It is ported for Solaris, AIX, and Linux.
11. Nusuif: This is an extension of the Suif compiler system with capability for parallelizing perfectly and imperfectly nested non-uniform dependence loops in Fortran and C.
12. Deeds: This is a Java security tool that handles user specified policies for security.
13. Primari: This is a parallel MRI tool that uses a network of workstations and reconstructs a deblurred MRI. It is being used in Detroit Medical Center clinically.
14. Prttp: This is a parallel radiation therapy treatment planning system which will replace the current treatment planning system being used at the DMC Karmanos Cancer Institute.
15. Cbt: This is a software system for three dimensional cone beam tomography was developed on a network of workstations that is currently being clinically used at Henry Ford Hospital.

INVITED TALKS

Invited Talks at Academic Institutions, Industries, and Research Laboratories

1. Rowan University, Keynote Speech, “*Transforming Science by Advanced Cyberinfrastructure*”, NSF Career Workshop, April 14, 2020.
2. New Jersey Institute of Technology President’s Forum Keynote Speech, “*Four Decades of HPC: Architectures, Programming Environments, Systems and Applications*”, November 14, 2019.
3. Rutgers University, Department Seminar, Electrical and Computer Engineering, “*Four Decades of HPC: Architectures, Programming Environments, Systems and Applications*”, May 10, 2019.
4. University of Delaware, Department Seminar, Computer and Information Sciences, “*Four Decades of HPC: Architectures, Programming Environments, Systems and Applications*”, March 10, 2019.
5. University of Virginia at Charlottesville, Department of Computer Science Distinguished Speaker Series, “*Changing Landscape of Advanced Cyberinfrastructure in Science and Engineering*”, October 19, 2018.
6. University of California at Riverside, Department of Computer Science and Engineering Colloquium, “*Opportunities for Advanced Cyberinfrastructure in Science and Engineering*”, October 30, 2017.
7. University of California at Riverside, School of Business Colloquium, “*“NSF Innovation Corps (I-Corps™): Preparing Scientists and Engineers to Accelerate the Economic and Societal Benefits”*”, October 30, 2017.
8. Panelist, Confluence 2016 hosted by Zinnov, “*Riding the Interoperability Wave: Transitioning towards and Open-Source Paradigm*”, San Jose, CA, March 23, 2016.
9. Institute for Genomics and Multi-scale Biology, Mount Sinai School of Medicine, New York, “*Data Intensive Discovery Initiative*”, May 23, 2012.
10. Tata Institute for Fundamental Research, International Center for Theoretical Sciences, JNCASR Bangalore, “*Scientific Discovery through Intensive Exploration of Data- Data Intensive Computing Architecture and Discovery Initiative*” February 10, 2011.
11. Indian Institute of Technology, Kharagpur, Department of Computer Science and Engineering, “*Manycores and Data Intensive Computing*”, October 8, 2010.

12. Novartis, Inc., Cambridge, MA, "*Data Intensive Discovery Initiative – Challenges and Solutions*", September 18, 2009.
13. Pacific Northwest National Laboratory, Richland, WA, "*Data Intensive Computing*", August 14, 2009
14. Computational Research Laboratory, Pune, India, "*Research Directions in High Performance Computing*", August 25, 2009.
15. Netezza Inc., Marlborough, MA, "*Data Intensive Discovery Initiative- Challenges and Solutions*", August 18, 2009.
16. University at Buffalo, Information and Computing Technology (ICT) Day, "*Multicores, Clouds, and Data Intensive Computing: Issues and Opportunities*", May 1, 2009.
17. Howard Hughes Medical Institute (HHMI) Janelia Farm Research Campus, "*MPI-Hmmer: Status and Looking Ahead*", March 17, 2008.
18. Roswell Park Cancer Institute, Buffalo, "*Trends in High Performance Computing and Storage and its impact on Medicine and Biology*", January 9, 2008.
19. Persistent Systems Pvt. Ltd., Pune, India, "*Virtual Machines and Accelerators for High Performance Computing and Computer Assisted Diagnosis and Interventions*", Aug 10, 2007.
20. Bioimagine India Pvt. Ltd., Pune, India, "*Computer Assisted Diagnosis and Interventions (CADI)*", August 9, 2007.
21. Bioimagine Inc, Cupertino USA, "*Computer Assisted Diagnosis and Interventions (CADI)*", April, 2007.
22. Delhi University, India, "*High Performance Computing*", August 27, 2005.
23. Hosei University, Japan, "*CASMIL: A Comprehensive Tool for Image Guided Neurosurgeries*", July 17, 2005.
24. EPSCoR Centers Development Initiative (CDI), "*High Performance Computing Applications*, March 29, 2004, Alexandria, VA.
25. Oakland University, "*Thread/Process Migration in Heterogeneous Distributed Systems*", Nov 17, 2003.
26. Albion College, "*Software Scalable System on Chip Architectures*", Sep 18, 2003.
27. University of Toledo, "*Software Scalable System on Chip Architectures*", Sep 3, 2003.
28. University of California at Davis, "*Software Scalable System on Chip Architectures*", March 6, 2003.
29. Birla Institute for Technology, Ranchi, India, "*Distributed Shared Memories and Thread Migration*", February 15, 2003.
30. Indian Institute for Science, Bangalore, "*Universal Micro System*", July 2002.
31. University of Michigan at Ann Arbor, "*Universal Micro System: An architecture for Streaming*", Invited Talk, Department of EECS – The Advanced Architecture Group, August 2001.
32. Wayne State University, "*Automatic Program Parallelization*" Faculty Seminar, 1999.
33. University of Cincinnati, "*Automatic Program Parallelization*", Invited Lecture Series, 1999.
34. University of Toledo, "*Automatic Program Parallelization*", Invited Lecture Series, 1999
35. State University of New York at Buffalo, Department of Computer Science Colloquium, "*Automatic Parallelization of Programs*", September 26, 1997.
36. United States Naval Research Laboratory, Washington, D.C., "*Automatic Parallelization of Sequential High Performance Computing Programs*", September 12, 1997.
37. United States Army Research Laboratory, Aberdeen, MD, "*Automatic Parallelization of Sequential High Performance Computing Programs*", August 15, 1997.

38. TARDEC/Oakland University monthly technical seminar, U. S. Army Tank Commandment, Warren, MI, ``Automatic Parallelization of Sequential High Performance Computing Programs'', May 8, 1997.
39. Institute of Electrical and Electronics Engineers, South East Michigan Chapter, Detroit, ``Automatic Parallelization of Scientific and Engineering Programs'', Dec. 12, 1996.
40. United States Army Research Laboratory, Aberdeen, MD, ``Automatic Program Parallelization'', August 9, 1995.
41. The Ohio State University, Columbus, Ohio, Parallel and Distributed Computing Colloquium, Department of Computer and Information Sciences, ``Parallelizing loops with non-uniform dependences'', Nov. 4, 1994.
42. TARDEC/Oakland University monthly technical seminar, U.S. Army Tank Commandment, Warren, MI, ``Automatic parallel program development for HPC applications'', Sep. 8, 1994.
43. University of Virginia, Charlottesville, Department of EE/CS Seminar, ``On tiling iteration spaces of nested loops with irregular dependences'', May 1994.
44. United States Army Research Laboratory, Aberdeen, MD, ``Automatic parallel program development for HPC Applications'', May 1994.
45. Siemens Research, Princeton, New Jersey, ``Automatic parallel program development for HPC Applications'', May 1994.
46. MITL (Panasonic) Research Laboratory, Princeton, New Jersey, ``Automatic parallel program development for HPC Applications'', May 1994.
47. United States Naval Research Laboratory, Washington, D.C., ``A generalized scheme for mapping parallel algorithms'', March 1994.
48. Auburn University, Department of Electrical and Computer Engineering, ``Mapping parallel algorithm in a distributed computing environment'', Apr. 1992.
49. University of Alabama at Birmingham, Department of Computer Science, ``Mapping parallel algorithm in a distributed computing environment'', April 1992.
50. Wayne State University, Department of Electrical and Computer Engineering, ``Mapping parallel algorithm in a distributed computing environment'', March 1992.
51. International Business Machines, Austin, TX, ``Mapping parallel algorithm in a distributed computing environment'', October 1991.
52. AT&T Bell Laboratories, Murray Hill, NJ, ``Mapping parallel algorithm in a distributed computing environment and VEDIC networks'', February 1992.
53. Schlumberger, Laboratory of Computer Science, ``Mapping parallel algorithm in a distributed computing environment'', January 1992.
54. Computer Society of India, ``International standardization of message transfer protocols'', February 1986.

Invited Conference/Workshop Talks

1. Moderator and Panelist, ``Future of Parallel and Distributed Computing'', International Conference on Parallel and Distributed Processing Systems, May 18-20, 2020, New Orleans, USA.
2. Keynote Speaker, ``Changing Landscape of Supercomputing in Four Decades'', Fifth International Symposium on Signal Processing and Intelligent Recognition Systems, December 18-21, 2019, Trivandrum, India.

3. Panelist Speaker, “*Funding Opportunities for Data-Centric Engineering*”, International Workshop on Data-Centric Engineering, MIT, Cambridge, USA, December 9-12, 2019.
4. Invited Speaker, “*Future Challenges for eScience*”, International Conference on eScience, 2019, San Diego, CA, September 24-27, 2019.
5. Keynote Speaker, “*Changing Landscape of Advanced Cyberinfrastructure in Science and Engineering*”, International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS 2019), Shenzhen, China, May 9, 2019.
6. Panelist, “*Future Directions of Research Funding Programs*”, 2019 SIAM Conference on Computational Science and Engineering, February 25, 2019, Spokane, WA.
7. Keynote Speaker, “*Changing Landscape of Advanced Cyberinfrastructure in Science and Engineering*”, International Conference on Distributed Computing and Networking (ICDCN 2019), Bangalore, India, January 4-7, 2019.
8. Featured Speaker, “*Computer Aided Diagnosis of Lumbar Pathology*”, International Conference on Medical Imaging and Case Reports, Baltimore, MD, October 29-31, 2018
9. Keynote Speaker, “*Medical Image Diagnostics: A case study of spine pathology*”, Fourth International Symposium on Signal Processing and Intelligent Recognition Systems, September 19-22, 2018, Bangalore, India.
10. Keynote Speaker, “*Diagnosing Spine Pathology using Medical Image Analysis*”, 7th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO’2018), August 29-31, 2018, Noida, India.
11. Invited Speaker, *IIT Bay Area Leadership Conference*, June 9, 2018, Santa Clara Convention Center, Santa Clara, CA, USA.
12. Opening Remarks, “*NSF Perspective*”, Workshop on BisQue + Scalable Image Informatics, University of California at Santa Barbara, February 5, 2018.
13. Keynote Speaker, “*Software Infrastructure and Sustainability*”, Workshop on BisQue + Scalable Image Informatics, University of California at Santa Barbara, February 6, 2018.
14. Panelist, “*Software Infrastructure and Sustainability*”, Workshop on BisQue + Scalable Image Informatics, University of California at Santa Barbara, February 6, 2018.
15. Panelist, “*BOF: Exploring Opportunities for Big Data Programming Challenge at BigDF*”, International Workshop on Foundations of Big Data Computing (BigDF), held with HiPC 2017, Jaipur, India, December 18, 2017.
16. Keynote Speaker, “*Advanced Cyberinfrastructure in Science and Engineering*”, First Workshop on Software Challenges to Exascale Computing, Jaipur, India, December 17, 2017.
17. Panel Moderator, “*Global Preparedness for Exascale Computing*”, First Workshop on Software Challenges to Exascale Computing, Jaipur, India, December 17, 2017.
18. Keynote Speaker, IEEE Workshop on Big Data in Smart Grids (Co-located with IEEE Big Data Conference), “*Advanced CyberInfrastructure to Enable Big Data in Smart Grids*”, December 5, 2016, Washington DC, USA.
19. World Wide Web: Technology, Standards and Internationalization Conference, “*High Performance Data Intensive Computing Related Issues*”, May 6-9, 2010, New Delhi, India.
20. 9th Annual Radiation Oncology Conference, “*High Performance Computing – Playstations, GPUs, etc.*”, Niagara Falls, USA, September 6, 2008.
21. 3rd International Conference on Distributed Computing and Internet Technology, “*Application-Level Checkpointing Techniques for Parallel Programs – Invited Talk/Paper*”, December 20-23, 2006, Bhubaneswar, India.

22. Keynote talk at the 8th International Workshop on High Performance Scientific and Engineering Computing (HPSEC-06) held with the 35th International Conference on Parallel Processing, August 18, 2006, *Surgery requires HPC, Really!*, Columbus, Ohio, August 18, 2006.
 23. International Conference and Exposition on Communications and Computing, “*Migrating Processes/Threads in Grids – Invited Talk*”, IIT Kanpur, India, February 4-6, 2005.
 24. HPC Consortium at IEEE/ACM Supercomputing 2002, “*Distributed Shared Memories and Thread Migration*”, November 15, 2002.
 25. International Workshop on Programming and Applications of Parallel/Distributed Systems, “*Automatic Program Parallelization*”, Invited talk, December 7, 1997.
 26. Third International Workshop on Data Analysis in Astronomy, “*Parallel image processing: architectures and algorithms*”, Invited talk, June 1988.
-

Advisory Boards of Companies

1. Omegaband, Inc., 2001-2003
 2. Cradle Technologies, Inc., 2002-2004
 3. Opalsoft, Inc., 2001- 2010
 4. Medcotek, Inc., 2007-2008
 5. Bioimagene, Inc., 2006-2009
 6. Bombay Stock Exchange (BSE), 2011- 2016
-