

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 11:34 PM
Reviewed By: F. Miller
Case #19CV345846
Envelope: 6087286**

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

**EXHIBITS A - E IN 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**

EXHIBIT A

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
12 September 2008 (12.09.2008)

PCT

(10) International Publication Number
WO 2008/109756 A1

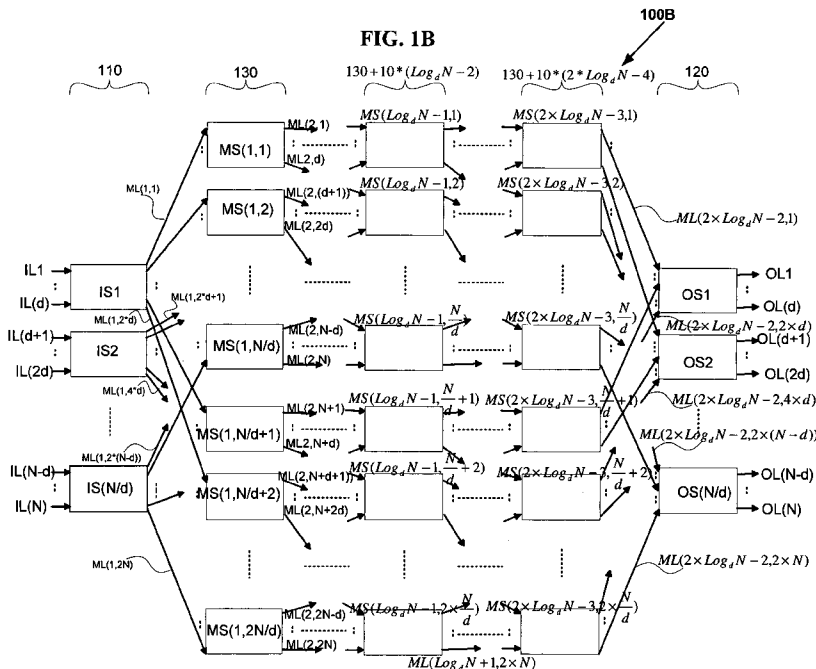
- (51) International Patent Classification:
H04Q 3/00 (2006.01)
- (21) International Application Number:
PCT/US2008/056064
- (22) International Filing Date: 6 March 2008 (06.03.2008)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/905,526 6 March 2007 (06.03.2007) US
60/940,383 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, — with international search report

(54) Title: FULLY CONNECTED GENERALIZED MULTI-STAGE NETWORKS



(57) Abstract: A multi-stage network comprising $(2x \log_d N) - 1$ 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 $2x d$ outgoing links connecting to second stage switches, an output stage having N/d switches with each of them having d outlet links and $2x d$ incoming links connecting from the switches in the penultimate stage. The network also has $(2x \log_d N) - 3$ middle stages with each middle stage having $2x N/d$ switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, and d outgoing links connecting to the switches in its immediate succeeding stage. Also the same 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. A multi-stage network comprising $(2x \log_d N) - 1$ 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 $3x d$ outgoing links connecting to second stage switches, an output stage having N/d switches with each of them having d outlet links and $3x d$ incoming links connecting from switches in the penultimate stage. The network also has $(2x \log_d N) - 3$ middle stages with each middle stage having $3x N/d$ switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, and d outgoing links connecting to the switches in its immediate succeeding stage.

WO 2008/109756 A1

WO 2008/109756

PCT/US2008/056064

1

FULLY CONNECTED GENERALIZED MULTI-STAGE 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. 60/905,526 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.

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, 383 entitled "FULLY CONNECTED 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 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 assignee as the current application, filed May 25, 2007.

This application is related to and incorporates by reference in its entirety the U.S. Provisional Patent 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.

WO 2008/109756

PCT/US2008/056064

2

This application is related to and incorporates by reference in its entirety 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 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.

10 This application is related to and incorporates by reference in its entirety 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.

15 This application is related 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, 2007.

20 This application is related to and incorporates by reference in its entirety the U.S. 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.

25 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 FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed January 1, 2008.

WO 2008/109756

PCT/US2008/056064

3

BACKGROUND OF INVENTION

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.

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 network's 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.

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. 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.

WO 2008/109756

PCT/US2008/056064

4

U.S. Patent 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.

U.S. Patent 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. Patent 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 $m \geq 3 \times n - 1$.

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, Jan. 1953, No.1, pp. 406-424 showed a way of constructing large multi-stage networks by recursive substitution with a crosspoint complexity of $d^2 \times N \times (\log_d N)^{2.58}$ for strictly nonblocking unicast network. Similarly U.S. Patent 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

WO 2008/109756

PCT/US2008/056064

5

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 $N=256$.

The crosspoint complexity of all these networks is prohibitively large to
5 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.

10 SUMMARY OF INVENTION

A multi-stage network comprising $(2 \times \log_d N) - 1$ stages is operated in strictly nonblocking manner for unicast includes an input stage having $\frac{N}{d}$ switches with each of them having d inlet links and $2 \times d$ outgoing links connecting to second stage switches, an output stage having $\frac{N}{d}$ switches with each of them having d outlet links and $2 \times d$
15 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 $\frac{2 \times N}{d}$ switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, and d outgoing links connecting to the switches in its immediate succeeding stage. Also the same multi-stage network is operated in
20 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.

A multi-stage network comprising $(2 \times \log_d N) - 1$ stages is operated in strictly nonblocking manner for multicast includes an input stage having $\frac{N}{d}$ switches with each of them having d inlet links and $3 \times d$ outgoing links connecting to second stage

switches, an output stage having $\frac{N}{d}$ 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 $\frac{3 \times N}{d}$ switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, and d outgoing links connecting to the switches in its immediate succeeding stage.

BRIEF DESCRIPTION OF DRAWINGS

FIG. 1A is a diagram 100A 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 $s = 2$ 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.

FIG. 1B is a diagram 100B of a general symmetrical multi-stage network $V(N, d, 2)$ 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.

FIG. 1C is a diagram 100C of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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.

FIG. 1D is a diagram 100D of a general asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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.

WO 2008/109756

PCT/US2008/056064

7

FIG. 1E is a diagram 100E of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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.

FIG. 1F is a diagram 100F of a general asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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.

FIG. 1A1 is a diagram 100A1 of an exemplary symmetrical multi-stage network $V(N, d, 2)$ 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.

FIG. 1C1 is a diagram 100C1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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.

FIG. 1E1 is a diagram 100E1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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 connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 1A2 is a diagram 100A2 of an exemplary symmetrical multi-stage network $V(N, d, 2)$ 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

WO 2008/109756

PCT/US2008/056064

8

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

FIG. 1C2 is a diagram 100C2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ having nearest neighbor connection topology of five stages with $N_1 = 8$,
5 $N_2 = p * N_1 = 24$ 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.

FIG. 1E2 is a diagram 100E2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ having nearest neighbor connection topology of five stages with $N_2 = 8$,
10 $N_1 = p * N_2 = 24$, 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.

FIG. 2A is a diagram 200A of an exemplary symmetrical multi-stage network $V(N, d, 3)$ having inverse Benes connection topology of five stages with $N = 8$, $d = 2$ and
15 $s=3$ with exemplary multicast connections strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2B1 & FIG. 2B2 is a diagram 200B of a general symmetrical multi-stage network $V(N, d, 3)$ with $(2 \times \log_d N) - 1$ stages strictly nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

FIG. 2C is a diagram 200C of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2D1 & FIG. 2D2 is a diagram 200D of a general asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ with $N_2 = p * N_1$ and with $(2 \times \log_d N) - 1$ stages strictly

WO 2008/109756

PCT/US2008/056064

9

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

FIG. 2E is a diagram 200E of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2F1 & FIG. 2F2 is a diagram 200F of a general asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ with $N_1 = p * N_2$ and with $(2 * \log_d N) - 1$ stages strictly nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

FIG. 2A1 is a diagram 200A1 of an exemplary symmetrical multi-stage network $V(N, d, 3)$ having Omega connection topology of five stages with $N = 8$, $d = 2$ and $s=3$ with exemplary multicast connections, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2C1 is a diagram 200C1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2E1 is a diagram 200E1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2A2 is a diagram 200A2 of an exemplary symmetrical multi-stage network $V(N, d, 3)$ having nearest neighbor connection topology of five stages with $N = 8$, $d = 2$

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

FIG. 2C2 is a diagram 200C2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ having nearest neighbor connection topology of five stages with $N_1 = 8$,
5 $N_2 = p * N_1 = 24$ where $p = 3$, and $d = 2$ with exemplary multicast connections, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 2E2 is a diagram 200E2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ having nearest neighbor connection topology of five stages with $N_2 = 8$,
10 $N_1 = p * N_2 = 24$, where $p = 3$, and $d = 2$ with exemplary multicast connections, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 3A 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
15 invention.

FIG. 4A1 is a diagram 400A1 of an exemplary prior art implementation of a two by two switch; FIG. 4A2 is a diagram 400A2 for programmable integrated circuit prior art implementation of the diagram 400A1 of FIG. 4A1; FIG. 4A3 is a diagram 400A3 for one-time programmable integrated circuit prior art implementation of the diagram 400A1
20 of FIG. 4A1; FIG. 4A4 is a diagram 400A4 for integrated circuit placement and route implementation of the diagram 400A1 of FIG. 4A1.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is concerned with the design and operation of large scale
25 crosspoint reduction using arbitrarily large multi-stage switching networks for broadcast, unicast and multicast connections including their generalized topologies. Particularly multi-stage networks with stages more than three and radices greater than or equal to two

WO 2008/109756

PCT/US2008/056064

11

offer large scale crosspoint reduction when configured with optimal links as disclosed in this invention.

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

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.

In certain 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-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.

In certain 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-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.

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

- 1) 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 U.S. Provisional Patent Application, Attorney Serial No. 60/940, 387 that is incorporated by reference above.
- 2) 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 U.S. Provisional Patent Application, Attorney Serial No. 60/940, 389 that is incorporated by reference above.
- 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 U.S. Provisional Patent Application, Attorney Serial No. 60/940, 390 that is incorporated by reference above.
- 4) 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 U.S. Provisional Patent Application, Attorney Serial No. 60/940, 391 that is incorporated by reference above.
- 5) 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 U.S. Provisional Patent Application, Attorney Serial No. 60/940,392 that is incorporated by reference above.

6) 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 U.S. Provisional Patent Application, Attorney Serial No. M-0045 US 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 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.

8) VLSI layouts of numerous types of multistage pyramid networks are described in U.S. Provisional Patent Application Serial No. 61/018,494 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed January 1, 2008.

Symmetric RNB Embodiments:

Referring to FIG. 1A, in one embodiment, an exemplary symmetrical multi-stage network 100A 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

WO 2008/109756

PCT/US2008/056064

14

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).

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.

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 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. 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.

WO 2008/109756

PCT/US2008/056064

15

Each of the $\frac{N}{d}$ 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).

5 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
10 switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

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
15 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).

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
20 (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)).

WO 2008/109756

PCT/US2008/056064

16

Each of the $\frac{N}{d}$ 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).

5 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 FIG. 1A1, in another embodiment of network $V(N, d, s)$, an exemplary symmetrical multi-stage network 100A1 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).

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.

25 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

WO 2008/109756

PCT/US2008/056064

17

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 multi-stage network of FIG. 1A1 is also the network of the type $V(N, d, s)$, 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.

Each of the $\frac{N}{d}$ 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).

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).

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)

WO 2008/109756

PCT/US2008/056064

18

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).

5 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
10 ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switches MS(3,1)).

Each of the $\frac{N}{d}$ 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(3,3), MS(3,5) and MS(3,7) through the
15 links ML(4,1), ML(4,5), ML(4,9) and ML(4,13) respectively).

Finally the connection topology of the network 100A1 shown in FIG. 1A1 is known to be back to back Omega connection topology.

Referring to FIG. 1A2, in another embodiment of network $V(N, d, s)$, an exemplary symmetrical multi-stage network 100A2 with five stages of thirty two
20 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
25 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).

WO 2008/109756

PCT/US2008/056064

19

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
 5 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.

In one embodiment of this network each of the input switches IS1-IS4 and output
 10 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 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
 15 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 multi-stage network of FIG. 1A2 is also the network of the type $V(N, d, s)$,
 20 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.

25 Each of the $\frac{N}{d}$ 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).

WO 2008/109756

PCT/US2008/056064

20

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(1,14) 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).

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).

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)).

Each of the $\frac{N}{d}$ 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(3,4), MS(3,5) and MS(3,8) through the links ML(4,1), ML(4,2), ML(4,3) and ML(4,4) respectively).

Finally the connection topology of the network 100A2 shown in FIG. 1A2 is hereinafter called nearest neighbor connection topology.

WO 2008/109756

PCT/US2008/056064

21

In the three embodiments of FIG. 1A, FIG. 1A1 and FIG. 1A2 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 $V(N, d, s)$ 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N, d, s)$ can be built. The embodiments of FIG. 1A, FIG. 1A1, and FIG. 1A2 are only three examples of network $V(N, d, s)$.

In the three embodiments of FIG. 1A, FIG. 1A1 and FIG. 1A2, 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.

In the example illustrated in FIG. 1A (or in FIG1A1, or in FIG. 1A2), 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 100A1,

or 100A2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

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:

Network 100B of FIG. 1B 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 multicast when $s = 2$ according to the current invention. Also the general symmetrical multi-stage network $V(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention (And in the example of FIG. 1B, $s = 2$). 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ output switches OS1-

WO 2008/109756

PCT/US2008/056064

23

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).

Each of the $\frac{N}{d}$ 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.

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.

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.

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 $130 + 10 * (2 * \log_d N - 4)$ 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.

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

WO 2008/109756

PCT/US2008/056064

24

As described before, again the connection topology of a general $V(N, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V(N, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations.

5 The applicant notes that the fundamental property of a valid connection topology of the general $V(N, d, s)$ network is, when no connections are setup in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N, d, s)$ can be built. The embodiments of FIG. 1A, FIG. 1A1, and FIG. 1A2 are three examples of network $V(N, d, s)$.

10 The general symmetrical multi-stage network $V(N, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 2$ according to the current invention. Also the general symmetrical multi-stage network $V(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention.

Every switch in the multi-stage networks discussed herein has multicast

15 capability. In a $V(N, d, s)$ 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

20 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

25 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.

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. 1A shows an exemplary five-stage network, namely $V(8,2,2)$, 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.

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 RNB ($N_2 > N_1$) Embodiments:

Referring to FIG. 1C, in one embodiment, an exemplary asymmetrical multi-stage network 100C 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).

WO 2008/109756

PCT/US2008/056064

26

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
 5 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.

10 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
 15 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
 20 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
 25 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.

WO 2008/109756

PCT/US2008/056064

27

- 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).
- 5 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
- 10 switch MS(1,1) to middle switch MS(2,1) and MS(2,3) respectively).

- 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
- 15 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).

- 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
- 20 (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)).

WO 2008/109756

PCT/US2008/056064

28

Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $d + d_2$ switches in middle stage 150 through $d + d_2$ 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),
 5 ML(4,17), ML(4,21), ML(4,25) and ML(4,29) respectively).

Finally the connection topology of the network 100C shown in FIG. 1C is known to be back to back inverse Benes connection topology.

Referring to FIG. 1C1, in another embodiment of network $V(N_1, N_2, d, s)$, an exemplary asymmetrical multi-stage network 100C1 with five stages of thirty two
 10 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
 15 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).

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
 20 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
 25 stage 130 and middle stage 140, and eight switches of size two by four in middle stage 150.

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

WO 2008/109756

29

PCT/US2008/056064

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. The asymmetric multi-stage network of FIG. 1C1 is also the network of the type $V(N_1, N_2, d, s)$, 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.

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).

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

WO 2008/109756

PCT/US2008/056064

30

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).

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).

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 $MS(3,1)$).

Each of the $\frac{N_1}{d}$ output switches $OS1 - OS4$ are connected from exactly $d+d_2$ switches in middle stage 150 through $d+d_2$ 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).

Finally the connection topology of the network 100C1 shown in FIG. 1C1 is known to be back to back Omega connection topology.

Referring to FIG. 1C2, in another embodiment of network $V(N_1, N_2, d, s)$, an exemplary asymmetrical multi-stage network 100C2 with five stages of thirty two

WO 2008/109756

PCT/US2008/056064

31

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).

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.

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. The asymmetric multi-stage network of FIG. 1C2 is also the network of the type $V(N_1, N_2, d, s)$, 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.

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).

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).

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).

WO 2008/109756

PCT/US2008/056064

33

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

5 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)).

Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $d+d_2$ switches in middle stage 150 through $d+d_2$ links (for example output switch OS1 is

10 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).

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

15 In the three embodiments of FIG. 1C, FIG. 1C1 and FIG. 1C2 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 $V(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For

20 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network

25 $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1C, FIG. 1C1, and FIG. 1C2 are only three examples of network $V(N_1, N_2, d, s)$.

WO 2008/109756

PCT/US2008/056064

34

In the three embodiments of FIG. 1C, FIG. 1C1 and FIG. 1C2, 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.

In the example illustrated in FIG. 1C (or in FIG1C1, or in FIG. 1C2), 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 100C (or 100C1, or 100C2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

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.

30

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

Network 100D of FIG. 1D is an example of general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ 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 100D of FIG. 1D, $N_1 = N$ and $N_2 = p * N$. The general

5 asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 2$ according to the current invention. Also the general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention. (And in the example of FIG. 1D, $s = 2$). The general asymmetrical multi-stage network

10 $V(N_1, N_2, d, s)$ 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

15 example the links OL1-OL($p*d$) 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).

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

20 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.

WO 2008/109756

36

PCT/US2008/056064

Each of the $2 \times \frac{N_1}{d}$ middle switches $MS(1,1) - MS(1,2 N_1/d)$ 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.

Similarly each of the $2 \times \frac{N_1}{d}$ middle switches $MS(\text{Log}_d N_1 - 1, 1) - MS(\text{Log}_d N_1 - 1, 2 \times \frac{N_1}{d})$ 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.

Similarly each of the $2 \times \frac{N_1}{d}$ middle switches $MS(2 * \text{Log}_d N_1 - 3, 1) - MS(2 * \text{Log}_d N_1 - 3, 2 \times \frac{N_1}{d})$ 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 $\frac{(d + d_2)}{2}$ output switches in output stage 120 through d links.

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 $d + d_2$ links.

As described before, again the connection topology of a general $V(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V(N_1, N_2, d, s)$ 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 $V(N_1, N_2, d, s)$ network is, when no connections are setup in the network, a connection from any inlet link to any outlet link can be setup. Based on this

WO 2008/109756

PCT/US2008/056064

37

property numerous embodiments of the network $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1C, FIG. 1C1, and FIG. 1C2 are three examples of network $V(N_1, N_2, d, s)$ for $s = 2$ and $N_2 > N_1$.

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

For example, the network of Fig. 1C shows an exemplary five-stage network, namely $V(8, 24, 2, 2)$, 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.

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 RNB ($N_1 > N_2$) Embodiments:

Referring to FIG. 1E, in one embodiment, an exemplary asymmetrical multi-stage network 100E 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

WO 2008/109756

PCT/US2008/056064

38

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).

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.

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 $2 * \frac{N_2}{d}$. 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 $d_1 = N_1 * \frac{d}{N_2} = p * d$. 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. 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

WO 2008/109756

PCT/US2008/056064

39

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.

5 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).

10 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).

15

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).

20

WO 2008/109756

PCT/US2008/056064

40

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
 5 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)).

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
 10 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).

Finally the connection topology of the network 100E shown in FIG. 1E is known to be back to back inverse Benes connection topology.

Referring to FIG. 1E1, in another embodiment of network $V(N_1, N_2, d, s)$, an
 15 exemplary asymmetrical multi-stage network 100E1 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
 20 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).

Such a network can be operated in strictly non-blocking manner for unicast
 25 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

WO 2008/109756

PCT/US2008/056064

41

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.

5 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

10 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

15 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 multi-stage network of FIG. 1E1 is also the network of the type $V(N_1, N_2, d, s)$, 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

20 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.

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

25 connected to middle switches MS(1,1), MS(1,2), MS(1,3), MS(1,4), MS(1,5), MS(1,6),

WO 2008/109756

PCT/US2008/056064

42

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).

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).

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).

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(3,5) 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)).

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

WO 2008/109756

PCT/US2008/056064

43

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).

Finally the connection topology of the network 100E1 shown in FIG. 1E1 is known to be back to back Omega connection topology.

5 Referring to FIG. 1E2, in another embodiment of network $V(N_1, N_2, d, s)$, an exemplary asymmetrical multi-stage network 100E2 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
10 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).

15 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
20 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.

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
25 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

WO 2008/109756

PCT/US2008/056064

44

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 multi-stage network of FIG. 1E1 is also the network of the type $V(N_1, N_2, d, s)$, 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 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.

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).

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).

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).

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)).

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).

Finally the connection topology of the network 100E2 shown in FIG. 1E2 is hereinafter called nearest neighbor connection topology.

In the three embodiments of FIG. 1E, FIG. 1E1 and FIG. 1E2 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

WO 2008/109756

PCT/US2008/056064

46

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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1E, FIG. 1E1, and FIG. 1E2 are only three examples of network $V(N_1, N_2, d, s)$.

In the three embodiments of FIG. 1E, FIG. 1E1 and FIG. 1E2, 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.

In the example illustrated in FIG. 1E (or in FIG. 1E1, or in FIG. 1E2), 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 100E (or 100E1, or 100E2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

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 ($N_1 > N_2$) Embodiments:

Network 100F of FIG. 1F is an example of general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ 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 100D of FIG. 1F, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 2$ according to the current invention. Also the general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention. (And in the example of FIG. 1F, $s = 2$). The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ 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($p*d$) 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,($d+p*d$)) 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).

WO 2008/109756

PCT/US2008/056064

48

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, {

5 N_1/d)+($d+d_1$)/2) through the links ML(1, (($d+d_1$)/2)+1) – ML(1, ($d+d_1$)) respectively.

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.

Similarly each of the $2 \times \frac{N_2}{d}$ middle switches $MS(\text{Log}_d N_2 - 1, 1)$ -

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

Similarly each of the $2 \times \frac{N_2}{d}$ middle switches $MS(2 \times \text{Log}_d N_2 - 3, 1)$ -

15 $MS(2 \times \text{Log}_d N_2 - 3, 2 \times \frac{N_2}{d})$ 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 d links and also are connected to exactly d output switches in output stage 120 through d links.

Each of the $\frac{N_2}{d}$ output switches OS1 – OS(N_2/d) are connected from exactly

20 $2 \times d$ switches in middle stage $130 + 10 * (2 * \text{Log}_d N_2 - 4)$ through $2 \times d$ links.

As described before, again the connection topology of a general $V(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V(N_1, N_2, d, s)$ 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 $V(N_1, N_2, d, s)$ network is, when no connections are setup in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1E, FIG. 1E1, and FIG. 1E2 are three examples of network $V(N_1, N_2, d, s)$ for $s = 2$ and $N_1 > N_2$.

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

For example, the network of Fig. 1E shows an exemplary five-stage network, namely $V(24, 8, 2, 2)$, 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.

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.

Symmetric SNB Embodiments:

Referring to FIG. 2C, FIG. 2C1, and FIG. 2C2, three exemplary symmetrical
 5 multi-stage networks 200C, 200C1, and 200C2 respectively with five stages of 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, 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
 10 two switches OS1-OS4. And all the middle stages namely middle stage 130 consists of twelve, two by two switches MS(1,1) - MS(1,12), middle stage 140 consists of twelve, two by two switches MS(2,1) - MS(2,12), and middle stage 150 consists of twelve, two by two switches MS(3,1) - MS(3,12).

Such a network can be operated in strictly nonblocking manner for multicast
 15 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 twelve switches in each of middle stage 130, middle stage 140 and middle stage 150.

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
 20 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 $3 \times \frac{N}{d}$. 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
 25 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

WO 2008/109756

PCT/US2008/056064

51

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
 5 inlet links IL1-IL8 as there are outlet links OL1-OL8, in a symmetrical network they are the same.

Each of the $\frac{N}{d}$ input switches IS1 – IS4 are connected to exactly $3 \times d$ switches
 in middle stage 130 through $3 \times d$ links (for example in FIG. 2A, input switch IS1 is
 connected to middle switches MS(1,1), MS(1,2), MS(1,5), MS(1,6), MS(1,9) and
 10 MS(1,10) through the links ML(1,1), ML(1,2), ML(1,3), ML(1,4), ML(1,5) and ML(1,6)
 respectively).

Each of the $3 \times \frac{N}{d}$ middle switches MS(1,1) – MS(1,12) in the middle stage 130
 are connected from exactly d input switches through d links (for example in FIG. 2A,
 the links ML(1,1) and ML(1,7) are connected to the middle switch MS(1,1) from input
 15 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).

Similarly each of the $3 \times \frac{N}{d}$ middle switches MS(2,1) – MS(2,12) in the middle
 stage 140 are connected from exactly d switches in middle stage 130 through d links
 20 (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).

Similarly each of the $3 \times \frac{N}{d}$ middle switches MS(3,1) – MS(3,12) in the middle
 25 stage 150 are connected from exactly d switches in middle stage 140 through d links

WO 2008/109756

52

PCT/US2008/056064

(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)).

Each of the $\frac{N}{d}$ 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 switches MS(3,1), MS(3,2), MS(3,5), MS(3,6), MS(3,9) and MS(3,10) through the links ML(4,1), ML(4,3), ML(4,9), ML(4,11), ML(4,17) and ML(4,19) respectively).

Finally the connection topology of the network 200A shown in FIG. 2A is known to be back to back inverse Benes connection topology; the connection topology of the network 200A1 shown in FIG. 2A1 is known to be back to back Omega connection topology; and the connection topology of the network 200A2 shown in FIG. 2A2 is hereinafter called nearest neighbor connection topology.

In the three embodiments of FIG. 2A, FIG. 2A1 and FIG. 2A2 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 three 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N, d, s)$ can be built. The embodiments of FIG. 2A, FIG. 2A1, and FIG. 2A2 are only three examples of network $V(N, d, s)$.

WO 2008/109756

PCT/US2008/056064

53

In the three embodiments of FIG. 2A, FIG. 2A1 and FIG. 2A2, 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(1,12), MS(2,1) – MS(2,12), and MS(3,1) – MS(3,12) are referred to as middle switches or middle ports.

In the example illustrated in FIG. 2A, FIG. 2A1, and 2A2, a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two middle 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 (or 200A1, or 200A2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

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.

30

Generalized SNB Embodiments:

Network 200B of FIG. 2B1 is an example of general symmetrical multi-stage network $V(N, d, s)$ with $(2 \times \log_d N) - 1$ stages. Network 200B of FIG. 2B1 contains three different copies of the network 200B2 in FIG. 2B2. The general symmetrical multi-stage network $V(N, d, s)$ can be operated in strictly nonblocking manner for multicast when $s = 3$ according to the current invention (and in the example of FIG. 2B1, $s = 3$). 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 $\frac{N}{d}$ input switches IS1-IS(N/d) (for example the links IL1-IL(d) to the input switch IS1) and $3 \times d$ outgoing links for each of $\frac{N}{d}$ 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 $\frac{N}{d}$ output switches OS1-OS(N/d) (for example the links OL1-OL(d) to the output switch OS1) and $3 \times d$ incoming links for each of $\frac{N}{d}$ output switches OS1-OS(N/d) (for example $ML(2 \times \log_d N - 2, 1) - ML(2 \times \log_d N - 2, 3 \times d)$ to the output switch OS1).

Each of the $\frac{N}{d}$ input switches IS1 - IS(N/d) are connected to exactly $3 \times d$ switches in middle stage 130 through $3 \times d$ links.

Each of the $3 \times \frac{N}{d}$ middle switches MS(1,1) - MS(1,3N/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.

Similarly each of the $3 \times \frac{N}{d}$ middle switches $MS(\log_d N - 1, 1) - MS(\log_d N - 1, 3 \times \frac{N}{d})$ in the middle stage $130 + 10 * (\log_d N - 2)$ are connected from

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

Similarly each of the $3 \times \frac{N}{d}$ middle switches $MS(2 \times \text{Log}_d N - 3, 1) -$

$MS(2 \times \text{Log}_d N - 3, 3 \times \frac{N}{d})$ in the middle stage $130+10*(2 * \text{Log}_d N - 4)$ are connected

5 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.

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

The general symmetrical multi-stage network $V(N, d, s)$ can be operated in
10 strictly nonblocking manner for multicast when $s = 3$ according to the current invention.

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. 2A shows an exemplary five-stage network, namely $V(8, 2, 3)$, with the following multicast
15 assignment $I_1 = \{1, 3\}$ 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(1,2) and MS(1,5) in middle stage 130, and fans out in middle switches MS(1,2) and MS(1,5) only once into middle switches MS(2,2) and MS(2,5) respectively in middle stage 140.

The connection I_1 also fans out in middle switches MS(2,2) and MS(2,5) only
20 once into middle switches MS(3,2) and MS(3,7) respectively in middle stage 150. The connection I_1 also fans out in middle switches MS(3,2) and MS(3,7) only once into output switches OS1 and OS3 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 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 switches in middle stage 130.

5 Asymmetric SNB ($N_2 > N_1$) Embodiments:

Referring to FIG. 2C, FIG. 2C1, and FIG. 2C2, three exemplary symmetrical multi-stage networks 200C, 200C1, and 200C2 respectively with five stages of 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, and 150 is shown where input stage 110 consists of four, two by six switches IS1-IS4 and output stage 120 consists of four, twelve by six switches OS1-OS4. And all the middle stages namely middle stage 130 consists of twelve, two by two switches MS(1,1) - MS(1,12), middle stage 140 consists of twelve, two by two switches MS(2,1) - MS(2,12), and middle stage 150 consists of twelve, two by four switches MS(3,1) - MS(3,12).

Such a network can be operated in strictly nonblocking 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 twelve by six, and there are twelve switches in each of middle stage 130, middle stage 140 and middle stage 150.

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 $3 \times \frac{N_1}{d}$. 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

WO 2008/109756

PCT/US2008/056064

57

notation $(2d + d_2) * d$, 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 $d * \frac{(2d + d_2)}{3}$. (Throughout the current invention, a fraction is rounded to the nearest higher integer). A switch as used

5 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

10 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.

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 in FIG. 2C, input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,5), MS(1,6), MS(1,9) and

15 MS(1,10) through the links ML(1,1), ML(1,2), ML(1,3), ML(1,4), ML(1,5) and ML(1,6) respectively).

Each of the $3 \times \frac{N_1}{d}$ middle switches MS(1,1) – MS(1,12) in the middle stage 130 are connected from exactly d input switches through d links (for example in FIG. 2C, the links ML(1,1) and ML(1,7) are connected to the middle switch MS(1,1) from input

20 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).

Similarly each of the $3 \times \frac{N_1}{d}$ middle switches MS(2,1) – MS(2,12) in the middle stage 140 are connected from exactly d switches in middle stage 130 through d links

25 (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).

5 Similarly each of the $3 \times \frac{N_1}{d}$ middle switches MS(3,1) – MS(3,12) 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 $\frac{(2d + d_2)}{3}$ output switches in output stage 120 through $\frac{(2d + d_2)}{3}$ links (for
10 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 switch MS(3,1)).

Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $2d + d_2$ switches in middle stage 150 through $2d + d_2$ 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),
15 MS(3,6), MS(3,7), MS(3,8), MS(3,9), MS(3,10), MS(3,11), and MS(3,12) through the links ML(4,1), ML(4,5), ML(4,9), ML(4,13), ML(4,17), ML(4,21), ML(4,25), ML(4,29), ML(4,33), ML(4,37), ML(4,41), and ML(4,45) respectively).

Finally the connection topology of the network 200C shown in FIG. 2C is known to be back to back inverse Benes connection topology; the connection topology of the
20 network 200C1 shown in FIG. 2C1 is known to be back to back Omega connection topology; and the connection topology of the network 200C2 shown in FIG. 2C2 is hereinafter called nearest neighbor connection topology.

In the three embodiments of FIG. 2C, FIG. 2C1 and FIG. 2C2 the connection topology is different. That is the way the links ML(1,1) - ML(1,24), ML(2,1) - ML(2,24),
25 ML(3,1) - ML(3,24), and ML(4,1) - ML(4,48) are connected between the respective stages is different. Even though only three embodiments are illustrated, in general, the

WO 2008/109756

PCT/US2008/056064

59

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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 2C, FIG. 2C1, and FIG. 2C2 are only three examples of network $V(N_1, N_2, d, s)$.

In the three embodiments of FIG. 2C, FIG. 2C1 and FIG. 2C2, 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(1,12), MS(2,1) – MS(2,12), and MS(3,1) – MS(3,12) are referred to as middle switches or middle ports.

In the example illustrated in FIG. 2C, FIG. 2C1, and 2C2, a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two middle 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 (or 200C1, or 200C2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

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:

Network 200D of FIG. 2D1 is an example of general symmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages where $N_2 > N_1$ and $N_2 = p * N_1$ where $p > 1$. In network 200D of FIG. 2D, $N_1 = N$ and $N_2 = p * N$. Network 200D of FIG. 2D1 contains three different copies of the network 200D2 in FIG. 2D2. The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for multicast when $s = 3$ according to the current invention (and in the example of FIG. 2D1, $s = 3$). 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 $3 \times d$ 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).

WO 2008/109756

PCT/US2008/056064

61

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.

Each of the $3 \times \frac{N_1}{d}$ middle switches MS(1,1) – MS(1, $3 \times \frac{N_1}{d}$) 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.

Similarly each of the $3 \times \frac{N_1}{d}$ middle switches $MS(\text{Log}_d N_1 - 1, 1)$ - $MS(\text{Log}_d N_1 - 1, 3 \times \frac{N_1}{d})$ 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.

Similarly each of the $3 \times \frac{N_1}{d}$ middle switches $MS(2 * \text{Log}_d N_1 - 3, 1)$ - $MS(2 * \text{Log}_d N_1 - 3, 3 \times \frac{N_1}{d})$ 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 $\frac{2d + d_2}{3}$ output switches in output stage 120 through $\frac{2d + d_2}{3}$ links.

Each of the $\frac{N_1}{d}$ output switches OS1 – OS(N_1/d) are connected from exactly $2d + d_2$ switches in middle stage $130 + 10 * (2 * \text{Log}_d N_1 - 4)$ through $2d + d_2$ links.

The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 3$ according to the current invention.

WO 2008/109756

PCT/US2008/056064

62

For example, the network of FIG. 2C shows an exemplary five-stage network, namely $V(8,2,3)$, 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(1,2) and MS(1,5) in middle stage 130, and fans out in middle switches MS(1,2) and MS(1,5) only once into middle switches MS(2,4) and MS(2,5) respectively in middle stage 140.

The connection I_1 also fans out in middle switches MS(2,4) and MS(2,5) only once into middle switches MS(3,4) and MS(3,7) respectively in middle stage 150. The connection I_1 also fans out in middle switches MS(3,4) 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 OS1 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 switches in middle stage 130.

15 **Asymmetric SNB ($N_1 > N_2$) Embodiments:**

Referring to FIG. 2E, FIG. 2E1, and FIG. 2E2, three exemplary symmetrical multi-stage networks 200E, 200E1, and 200E2 respectively with five stages of 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, and 150 is shown where input stage 110 consists of four, six by twelve 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 twelve, four by two switches MS(1,1) - MS(1,12), middle stage 140 consists of twelve, two by two switches MS(2,1) - MS(2,12), and middle stage 150 consists of twelve, two by two switches MS(3,1) - MS(3,12).

Such a network can be operated in strictly nonblocking manner for multicast connections, because the switches in the input stage 110 are of size six by twelve, the

WO 2008/109756

PCT/US2008/056064

63

switches in output stage 120 are of size six by two, and there are twelve switches in each of middle stage 130, middle stage 140 and middle stage 150.

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

5 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 $3 \times \frac{N_2}{d}$. The size of each input switch IS1-IS4 can be denoted in general with

10 the notation $d * (2d + d_1)$, where $d_1 = N_1 \times \frac{d}{N_2} = p \times d$ and each output switch OS1-OS4 can be denoted in general with the notation $3d * d$. 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 $\frac{(2d + d_1)}{3} * d$. (Throughout the current

invention, a fraction is rounded to the nearest higher integer). A switch as used herein can be either a crossbar switch, or a network of switches each of which in turn may be a

15 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 input switch where $N_1 > N_2$, and s is the ratio of

20 number of incoming links to each output switch to the outlet links of each output switch.

Each of the $\frac{N_2}{d}$ input switches IS1 – IS4 are connected to exactly $2d + d_1$ switches in middle stage 130 through $2d + d_1$ links (for example in FIG. 2E, 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), MS(1,8), MS(1,9), MS(1,10), MS(1,11) and MS(1,12)

WO 2008/109756

PCT/US2008/056064

64

through the links ML(1,1), ML(1,2), ML(1,3), ML(1,4), ML(1,5), ML(1,6), ML(1,7), ML(1,8), ML(1,9), ML(1,10), ML(1,11), and ML(1,12) respectively).

Each of the $3 \times \frac{N_2}{d}$ middle switches MS(1,1) – MS(1,12) in the middle stage 130 are connected from exactly $\frac{2d + d_1}{3}$ input switches through $\frac{2d + d_1}{3}$ links (for example in FIG. 2E, the links ML(1,1), ML(1,13), ML(1,25), and ML(1,37) 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).

Similarly each of the $3 \times \frac{N_2}{d}$ middle switches MS(2,1) – MS(2,12) 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).

Similarly each of the $3 \times \frac{N_2}{d}$ middle switches MS(3,1) – MS(3,12) 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), ML(4,2), ML(4,3), and ML(4,4) are connected to output switches OS1, OS2, OS3, and OS4 respectively from middle switch MS(3,1)).

Each of the $\frac{N_2}{d}$ output switches OS1 – OS4 are connected from exactly $3d$ switches in middle stage 150 through $3d$ 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), MS(3,8), MS(3,9), MS(3,10), MS(3,11), and MS(3,12) through the links ML(4,1), ML(4,5), ML(4,9), ML(4,13), ML(4,17), ML(4,21), ML(4,25), ML(4,29), ML(4,33), ML(4,37), ML(4,41), and ML(4,45) respectively).

5 Finally the connection topology of the network 200E shown in FIG. 2E is known to be back to back inverse Benes connection topology; the connection topology of the network 200E1 shown in FIG. 2E1 is known to be back to back Omega connection topology; and the connection topology of the network 200E2 shown in FIG. 2E2 is hereinafter called nearest neighbor connection topology.

10 In the three embodiments of FIG. 2E, FIG. 2E1 and FIG. 2E2 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 three embodiments are illustrated, in general, the network $V(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For
15 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network
20 $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 2E, FIG. 2E1, and FIG. 2E2 are only three examples of network $V(N_1, N_2, d, s)$.

In the three embodiments of FIG. 2E, FIG. 2E1 and FIG. 2E2, 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
25 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(1,12),

WO 2008/109756

PCT/US2008/056064

66

MS(2,1) – MS(2,12), and MS(3,1) – MS(3,12) are referred to as middle switches or middle ports.

In the example illustrated in FIG. 2E, FIG. 2E1, and 2E2, a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two middle 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 (or 200E1, or 200E2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

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:

Network 200F of FIG. 2F1 is an example of general symmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network 200F of FIG. 2F, $N_2 = N$ and $N_1 = p * N$. Network 200F of FIG. 2F1 contains three different copies of the network 200F2 in FIG. 2F2. The general

WO 2008/109756

67

PCT/US2008/056064

asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for multicast when $s = 3$ according to the current invention (and in the example of FIG. 2F1, $s = 3$). 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 $2d + d_1 (= 2d + p \times d)$ 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 $3 \times d$ 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)..

Each of the $\frac{N_2}{d}$ input switches IS1 - IS(N_2/d) are connected to exactly $2d + d_1$ switches in middle stage 130 through $2d + d_1$ links.

Each of the $3 \times \frac{N_2}{d}$ middle switches MS(1,1) - MS(1,3 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.

Similarly each of the $3 \times \frac{N_2}{d}$ middle switches MS($\log_d N_2 - 1, 1$) - MS($\log_d N_2 - 1, 3 \times \frac{N_2}{d}$) in the middle stage 130 + $10 * (\log_d N_2 - 2)$ are connected from exactly d switches in middle stage 130 + $10 * (\log_d N_2 - 3)$ through d links and also are connected to exactly d switches in middle stage 130 + $10 * (\log_d N_2 - 1)$ through d links.

WO 2008/109756

PCT/US2008/056064

68

Similarly each of the $3 \times \frac{N_2}{d}$ middle switches $MS(2 \times \text{Log}_d N_2 - 3, 1) - MS(2 \times \text{Log}_d N_2 - 3, 3 \times \frac{N_2}{d})$ 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 d links and also are connected to exactly d output switches in output stage 120 through d links.

Each of the $\frac{N_2}{d}$ output switches OS1 – OS(N_2/d) are connected from exactly $3 \times d$ switches in middle stage $130 + 10 * (2 * \text{Log}_d N_2 - 4)$ through $3 \times d$ links.

The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 3$ according to the current invention.

For example, the network of FIG. 2E shows an exemplary five-stage network, namely $V(8, 2, 3)$, with the following multicast assignment $I_1 = \{1, 3\}$ 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(1,2) and MS(1,5) in middle stage 130, and fans out in middle switches MS(1,2) and MS(1,5) only once into middle switches MS(2,4) and MS(2,5) respectively in middle stage 140.

The connection I_1 also fans out in middle switches MS(2,4) and MS(2,5) only once into middle switches MS(3,2) and MS(3,7) respectively in middle stage 150. The connection I_1 also fans out in middle switches MS(3,2) and MS(3,7) only once into output switches OS1 and OS3 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 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 switches in middle stage 130.

25

Applications Embodiments:

All the embodiments disclosed in the current invention are useful in many varieties of applications. FIG. 4A1 illustrates the diagram of 400A1 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. 4A1. For example the diagram of 400A1 may be the implementation of middle switch MS(1,1) of the diagram 100A of FIG. 1A where inlet link IL1 of diagram 400A1 corresponds to middle link ML(1,1) of diagram 100A, inlet link IL2 of diagram 400A1 corresponds to middle link ML(1,5) of diagram 100A, outlet link OL1 of diagram 400A1 corresponds to middle link ML(2,1) of diagram 100A, outlet link OL2 of diagram 400A1 corresponds to middle link ML(2,2) of diagram 100A.

1) Programmable Integrated Circuit Embodiments:

All the embodiments disclosed in the current invention are useful in programmable integrated circuit applications. FIG. 4A2 illustrates the detailed diagram 400A2 for the implementation of the diagram 400A1 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

WO 2008/109756

PCT/US2008/056064

70

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

5 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.

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. 4A3 illustrates the detailed diagram 400A3 for the implementation of the diagram 400A1 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 400A3 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

30 connected as denoted by thick circle at the intersection of inlet link IL2 and outlet link

WO 2008/109756

PCT/US2008/056064

71

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. 4A4 illustrates the detailed diagram 400A4 for the implementation of the diagram 400A1 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 400A4 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 400A4, there is no need to drive the

WO 2008/109756

PCT/US2008/056064

72

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.

5 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
10 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 systems and in time-space-time switches.

15 Scheduling Method Embodiments:

FIG. 3A 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 $V(N_1, N_2, d, s)$ disclosed in this invention). According to this embodiment, a multicast connection request is received in act 1010. Then the
20 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
25 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 out going middle links to middle stage 140 all the available middle

WO 2008/109756

PCT/US2008/056064

73

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 1030.

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 $130+10*(2*\text{Log}_d N - 4)$ 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 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

WO 2008/109756

PCT/US2008/056064

74

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.

10 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.

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.

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

WO 2008/109756

PCT/US2008/056064

75

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.

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. 3A 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.

According to the current invention, using the method 1000 of FIG. 3A, the network $V(N_1, N_2, d, s)$ 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 $s \geq 3$.

The connection request of the type described above in reference to method 1000 of FIG. 3A 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. 3A 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.

WO 2008/109756

PCT/US2008/056064

76

As noted above method 1000 of FIG. 3A can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks $V(N, d, s)$ and $V(N_1, N_2, d, s)$ disclosed in this invention.

Numerous modifications and adaptations of the embodiments, implementations,
5 and examples described herein will be apparent to the skilled artisan in view of the disclosure

WO 2008/109756

PCT/US2008/056064

77

CLAIMS

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
 - 5 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
 - 10 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 $x \times \frac{N}{d}$ middle switches in each of said y middle stages wherein said second stage and said penultimate stage are one of said
 - 15 middle stages where $y > 3$, and
 - each middle switch in all said middle stages excepting said penultimate stage comprising d incoming links (hereinafter "incoming middle links") connecting from switches in its immediate preceding stage, and each middle switch further comprising d outgoing links (hereinafter "outgoing middle links") connecting to switches in its
 - 20 immediate succeeding stage; and
 - each middle switch in said penultimate stage comprising d incoming links (hereinafter "incoming middle links") connecting from switches in its immediate preceding stage, and each middle switch further comprising $\frac{(d + d_2)}{2}$ outgoing links

WO 2008/109756

78

PCT/US2008/056064

(hereinafter “outgoing middle links”) connecting 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

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

5 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

10 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 $x \times \frac{N}{d}$ 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 $\frac{(d + d_1)}{2}$ incoming links

15 (hereinafter “incoming middle links”) connecting from switches in its immediate preceding stage, and each middle switch further comprising d outgoing links (hereinafter “outgoing middle links”) connecting to switches in its immediate succeeding stage; and

each middle switch in all said middle stages excepting said second stage comprising d incoming links (hereinafter “incoming middle links”) connecting from
20 switches in its immediate preceding stage, and each middle switch further comprising 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
25 plurality of outgoing links in a plurality switches in each said middle stage and in said output stage.

WO 2008/109756

PCT/US2008/056064

79

2. The network of claim 1, wherein all said incoming middle links and outgoing middle links are connected in any 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.
- 5 3. The network of claim 2, wherein $y \geq (2 \times \log_d N_1) - 3$ when $N_2 > N_1$, and $y \geq (2 \times \log_d N_2) - 3$ when $N_1 > N_2$.
4. The network of claim 3, wherein $x \geq 1$, wherein said each multicast connection comprises only one destination link, and
said each multicast connection from an inlet link passes through only one
10 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
15 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 $x \geq 2$, wherein said each multicast connection comprises only one destination link, and
said each multicast connection from an inlet link passes through only one
20 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
25 connection comprises only one destination link and the network is hereinafter "strictly nonblocking network for unicast".
6. The network of claim 3, wherein $x \geq 2$,

WO 2008/109756

PCT/US2008/056064

80

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".

5 7. The network of claim 3, wherein $x \geq 3$,

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
10 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 each of said input switches, or each of said output switches, or each of
15 said middle switches further recursively comprise one or more networks.

11. 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

20 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

WO 2008/109756

PCT/US2008/056064

81

- 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 $x \times \frac{N}{d}$ 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 d incoming links connected from switches in its immediate preceding stage, and each middle switch further having d outgoing links connected to switches in its immediate succeeding stage; and
- each middle switch in said penultimate stage having d incoming links connected from switches in its immediate preceding stage, and each middle switch further having $\frac{(d + d_2)}{2}$ 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 $d_1 = N_1 \times \frac{d}{N_2} = p \times d$; 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

WO 2008/109756

82

PCT/US2008/056064

- a plurality of y middle stages having $x \times \frac{N}{d}$ 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 $\frac{(d + d_1)}{2}$ incoming links
- 5 connected from switches in its immediate preceding stage, and each middle switch further having 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 d incoming links connected from switches in its immediate preceding stage, and each middle switch further having d outgoing links connected to switches in its immediate
- 10 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
- 15 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.
12. A method of claim 11 wherein said act of fanning out is performed without
- 20 changing any existing connection to pass through another set of plurality of middle switches in each said middle stage.
13. A method of claim 11 wherein said act of fanning out is performed recursively.
14. A method of claim 11 wherein a connection exists through said network and passes through a plurality of middle switches in each said middle stage and said method
- 25 further comprises:

WO 2008/109756

83

PCT/US2008/056064

if necessary, changing said connection to pass through another set of plurality of middle switches in each said middle stage, act hereinafter "rearranging connection".

15. A method of claim 11 wherein said acts of fanning out and rearranging are performed recursively.

5 16. 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

10 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

15 a plurality of y middle stages having $x \times \frac{N}{d}$ 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 d incoming links connected from switches in its immediate preceding stage, and
20 each middle switch further having d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said penultimate stage having d incoming links connected from switches in its immediate preceding stage, and each middle switch further

WO 2008/109756

84

PCT/US2008/056064

having $\frac{(d+d_2)}{2}$ 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

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

5 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

10 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 $x \times \frac{N}{d}$ 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 $\frac{(d+d_1)}{2}$ incoming links

15 connected from switches in its immediate preceding stage, and each middle switch further having 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 d incoming links connected from switches in its immediate preceding stage, and each middle switch further having d outgoing links connected to switches in its immediate

20 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

WO 2008/109756

PCT/US2008/056064

85

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
5 first subset of destination output switches and said second subset of destination output switches.

17. The method of claim 16 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
10 said first plurality of outgoing links in plurality of middle switches in each said middle stage

18. The method of claim 16 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
15 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
20 switches.

19. The method of claim 16 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
25 link in input switch other than said first outgoing link in input switch.

20. The method of claim 16 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

WO 2008/109756**PCT/US2008/056064****86**

fanning out said multicast connection in its said input switch into not more than said two outgoing links.

21. The method of claim 16 wherein any of said acts of checking and setting up are performed recursively.

5

FIG. 1A

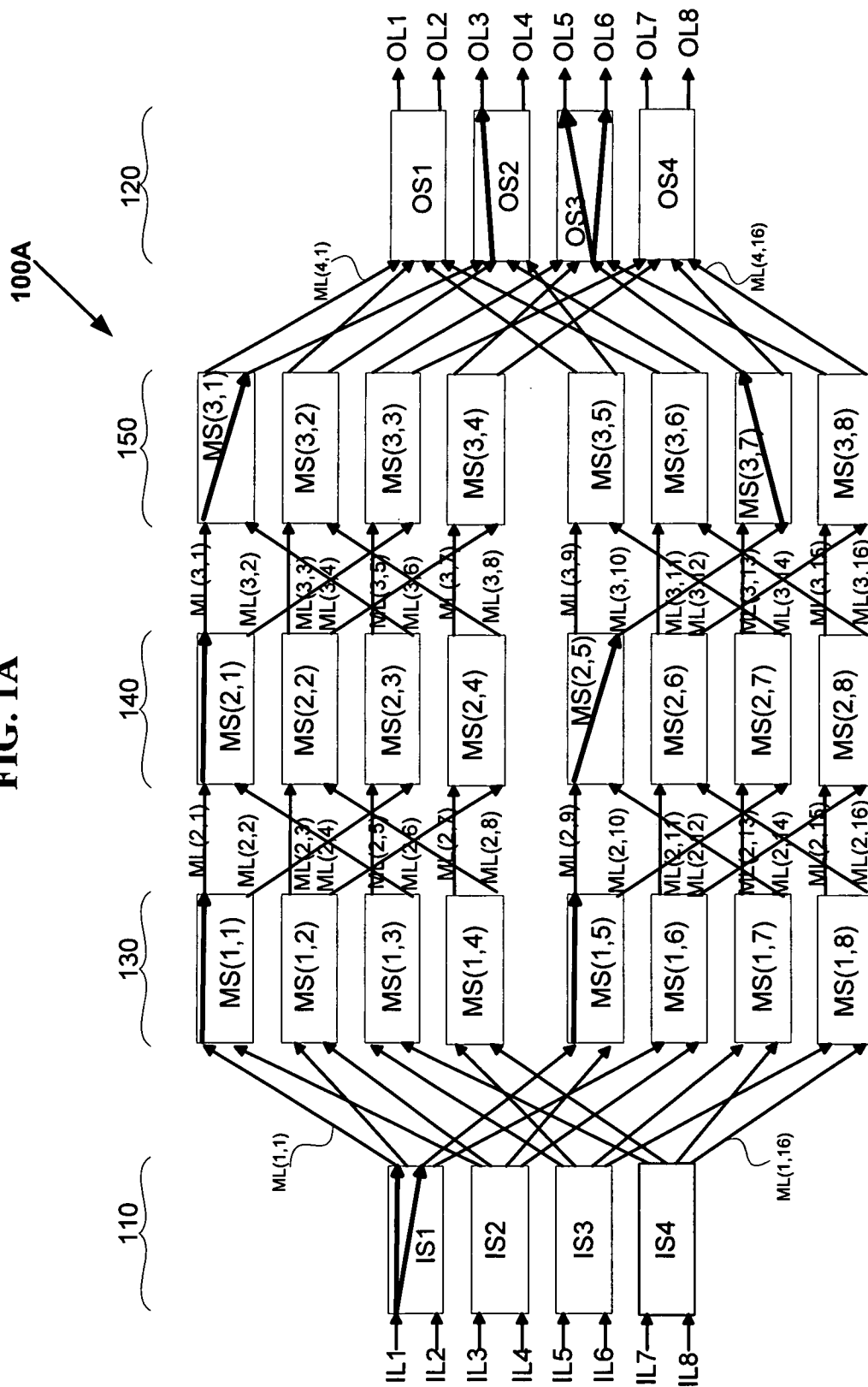


FIG. 1A1

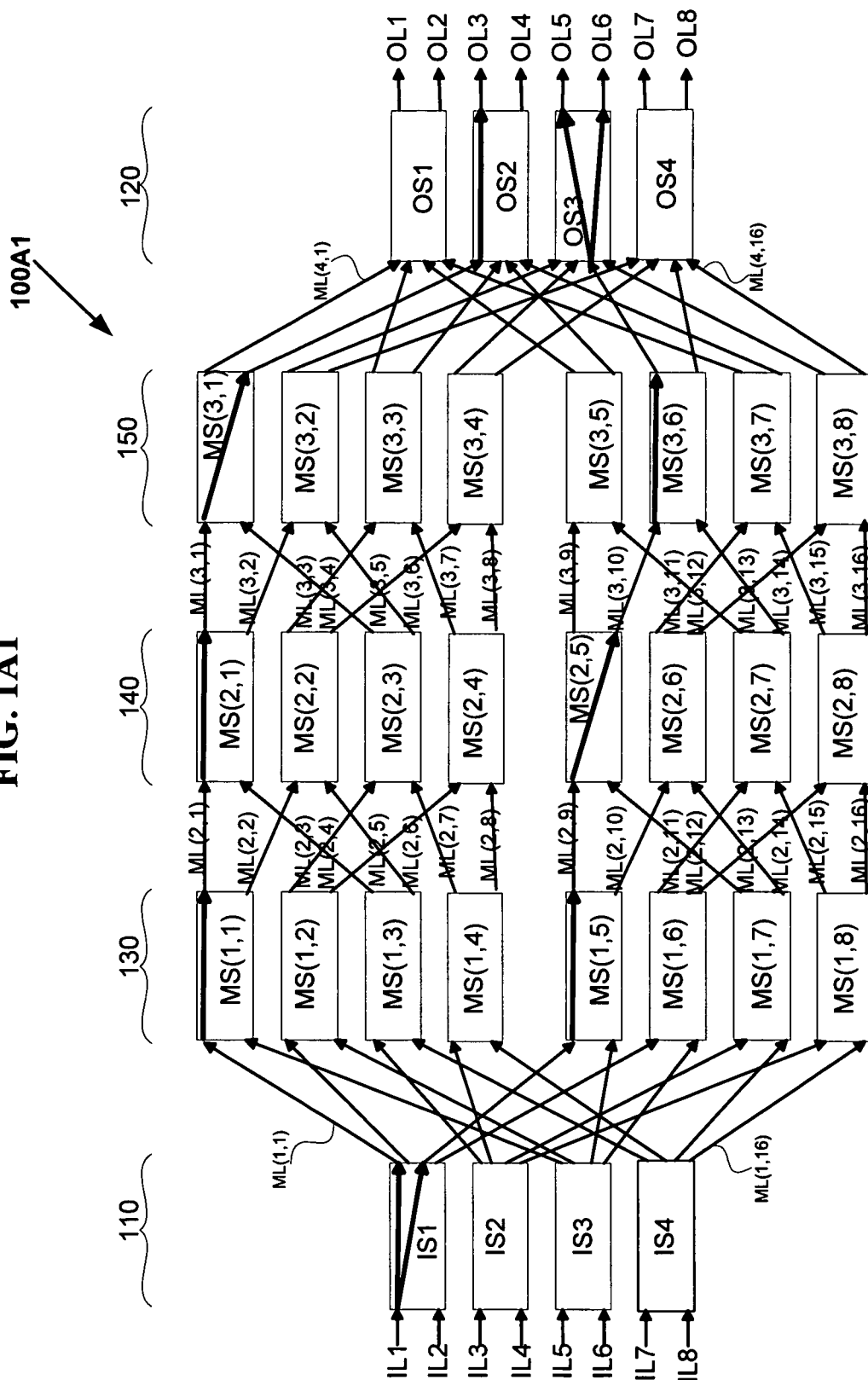
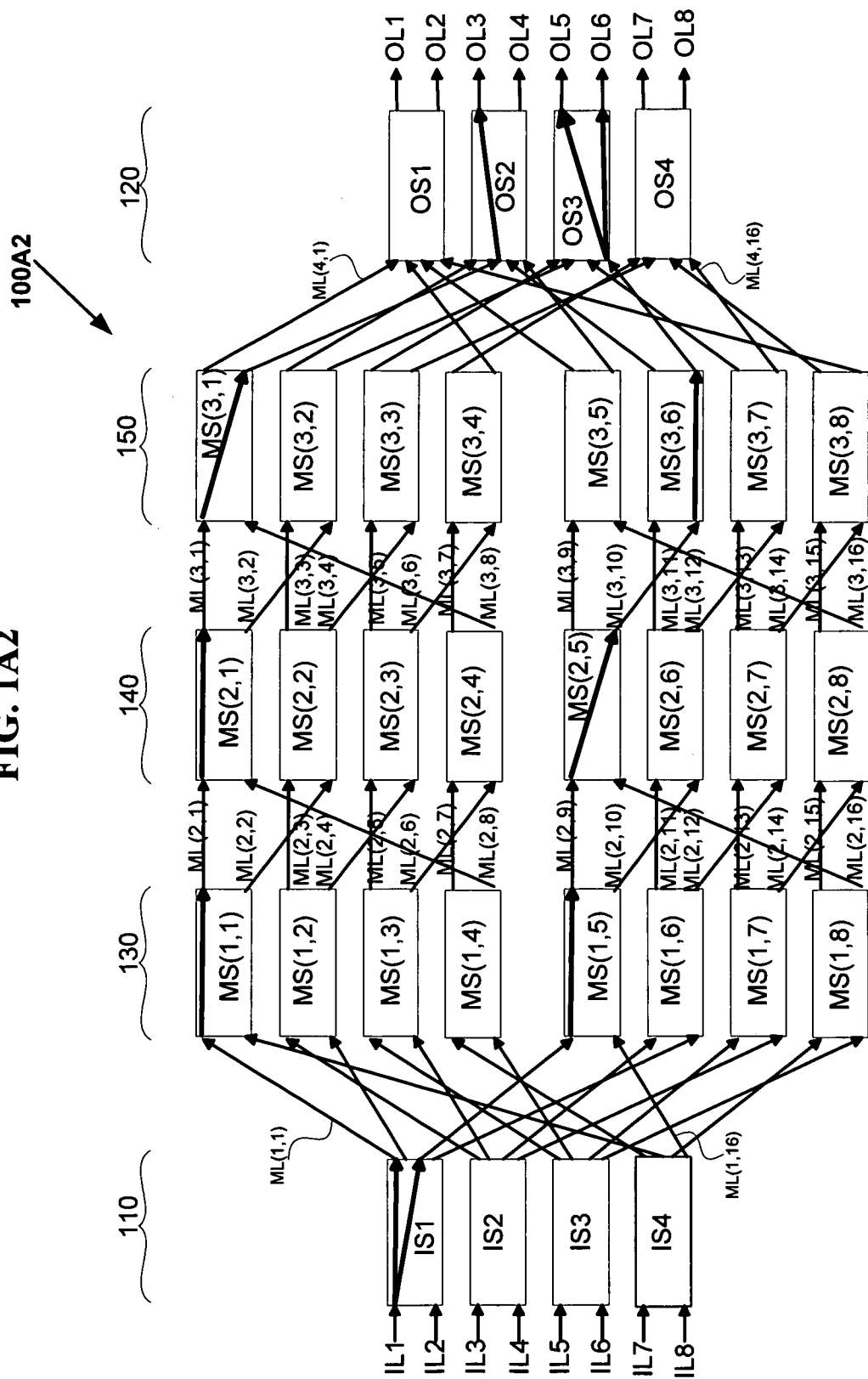


FIG. 1A2



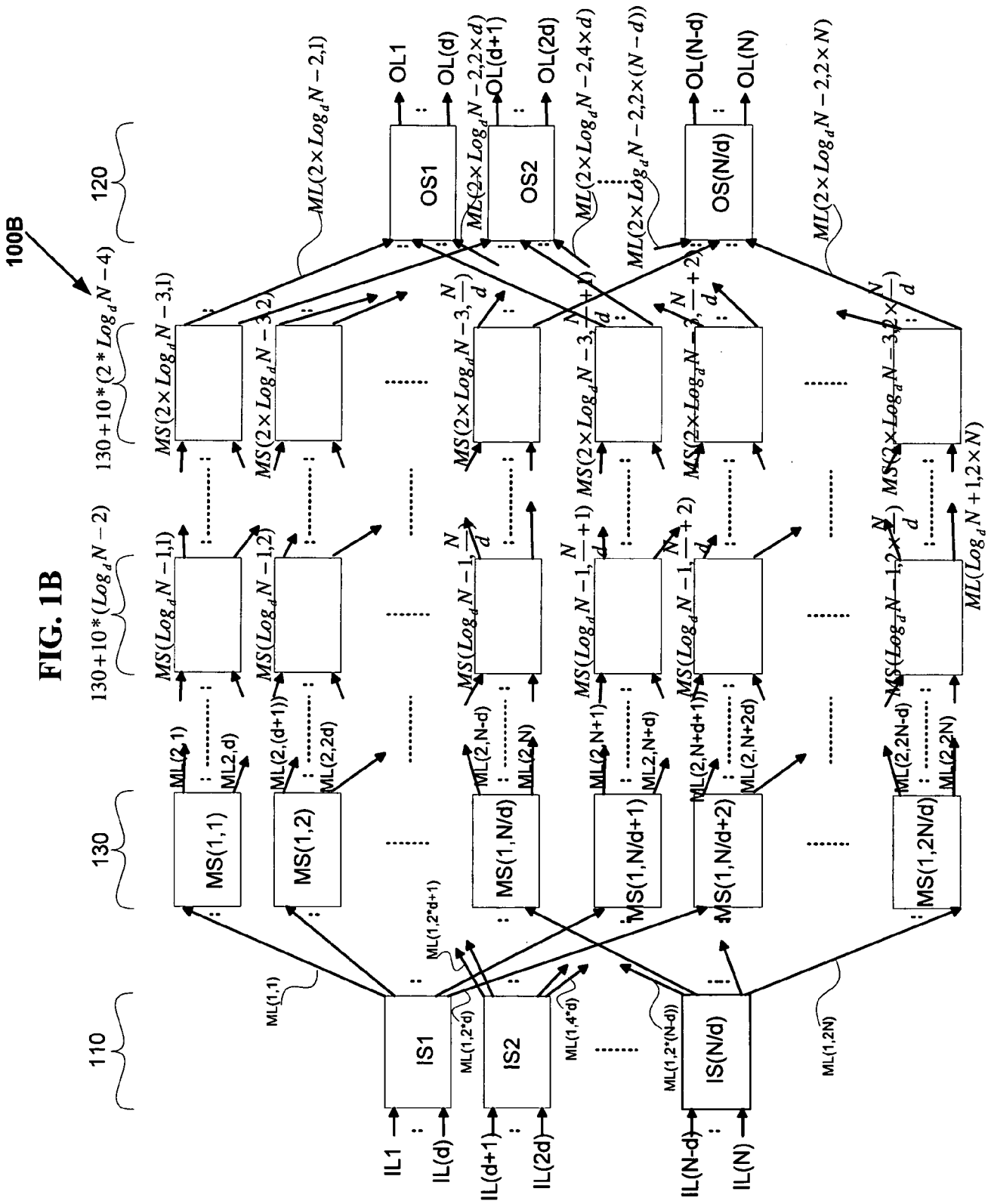


FIG. 1C

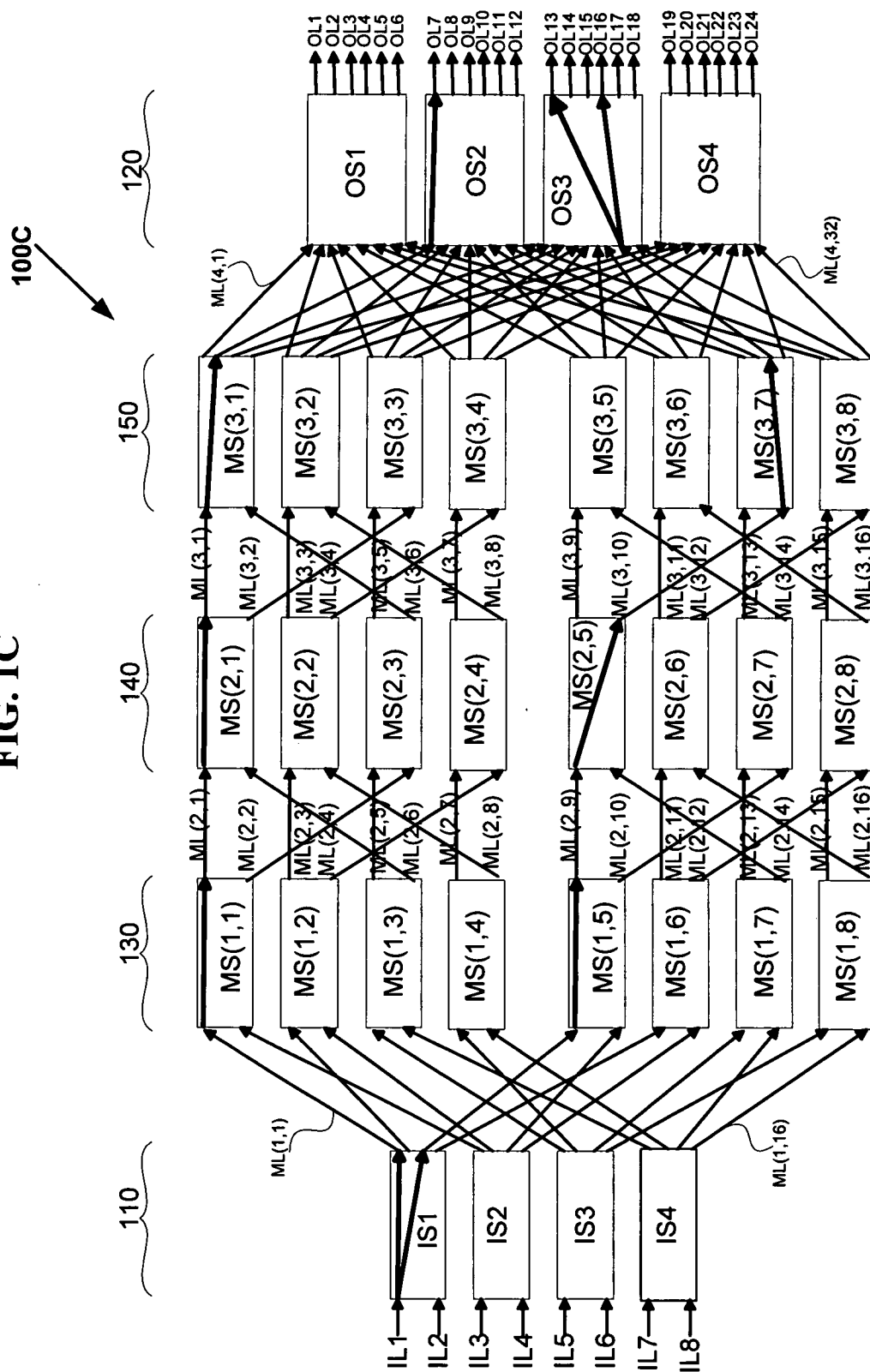


FIG. 1C1

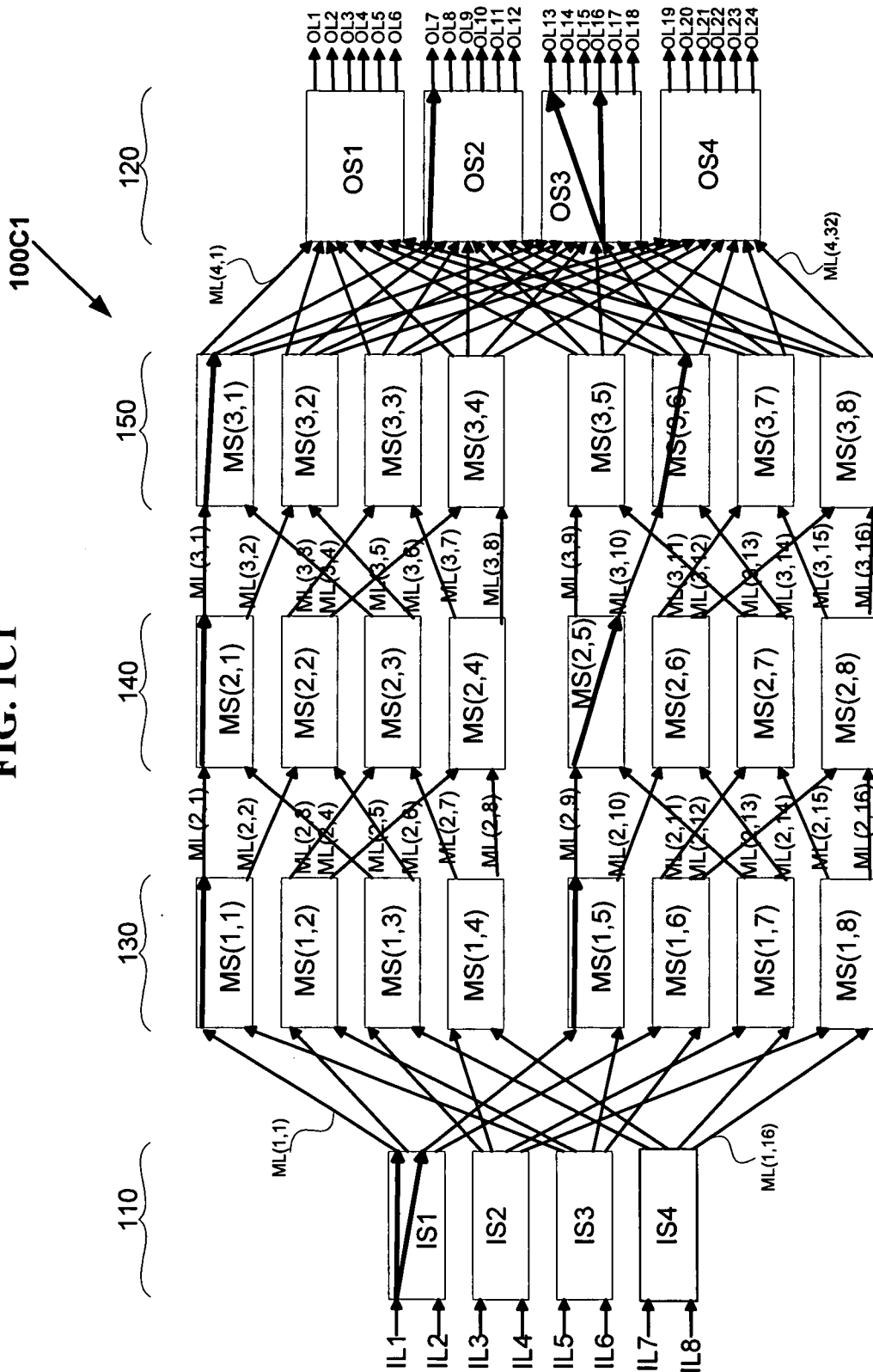
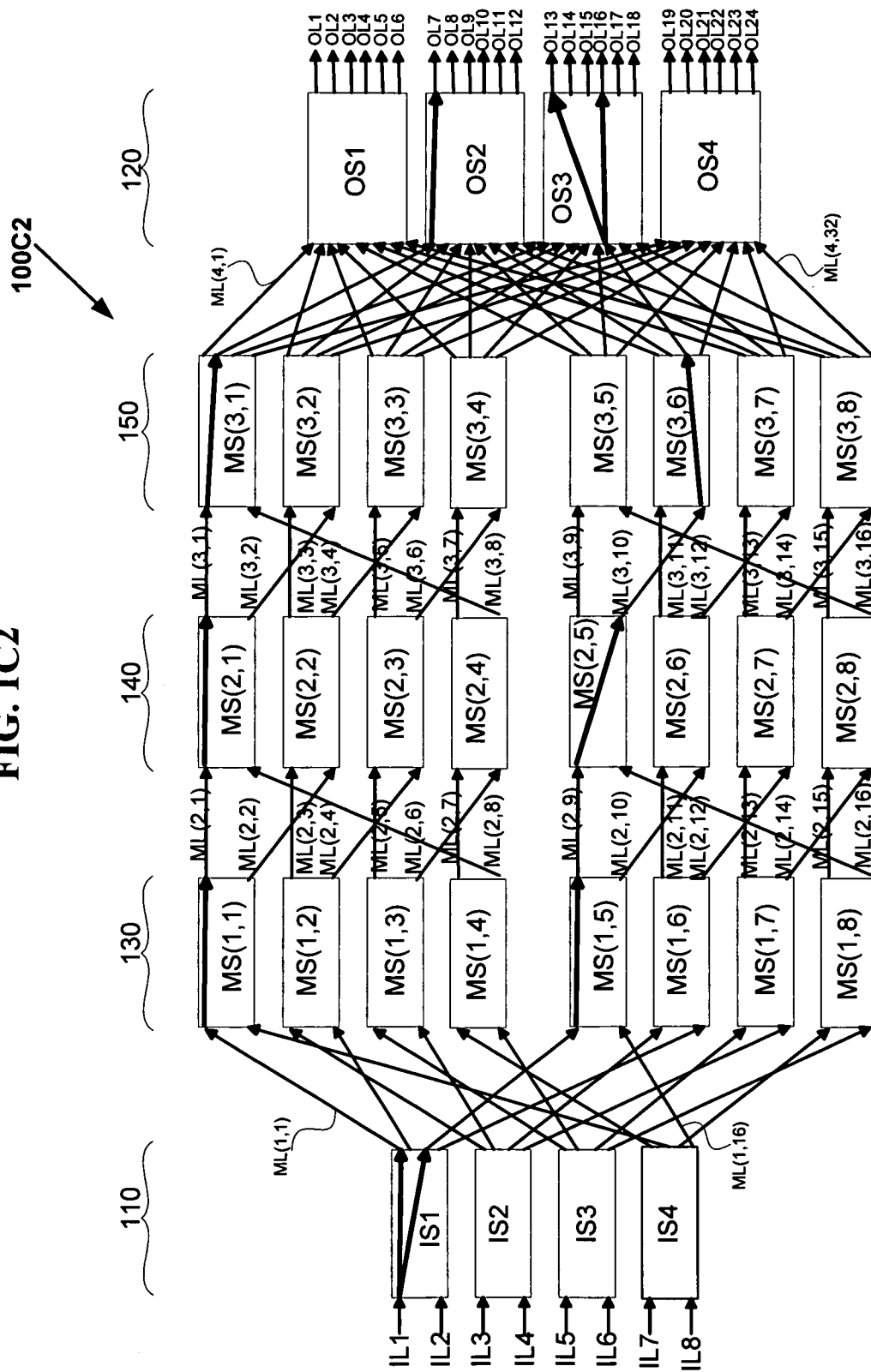


FIG. 1C2



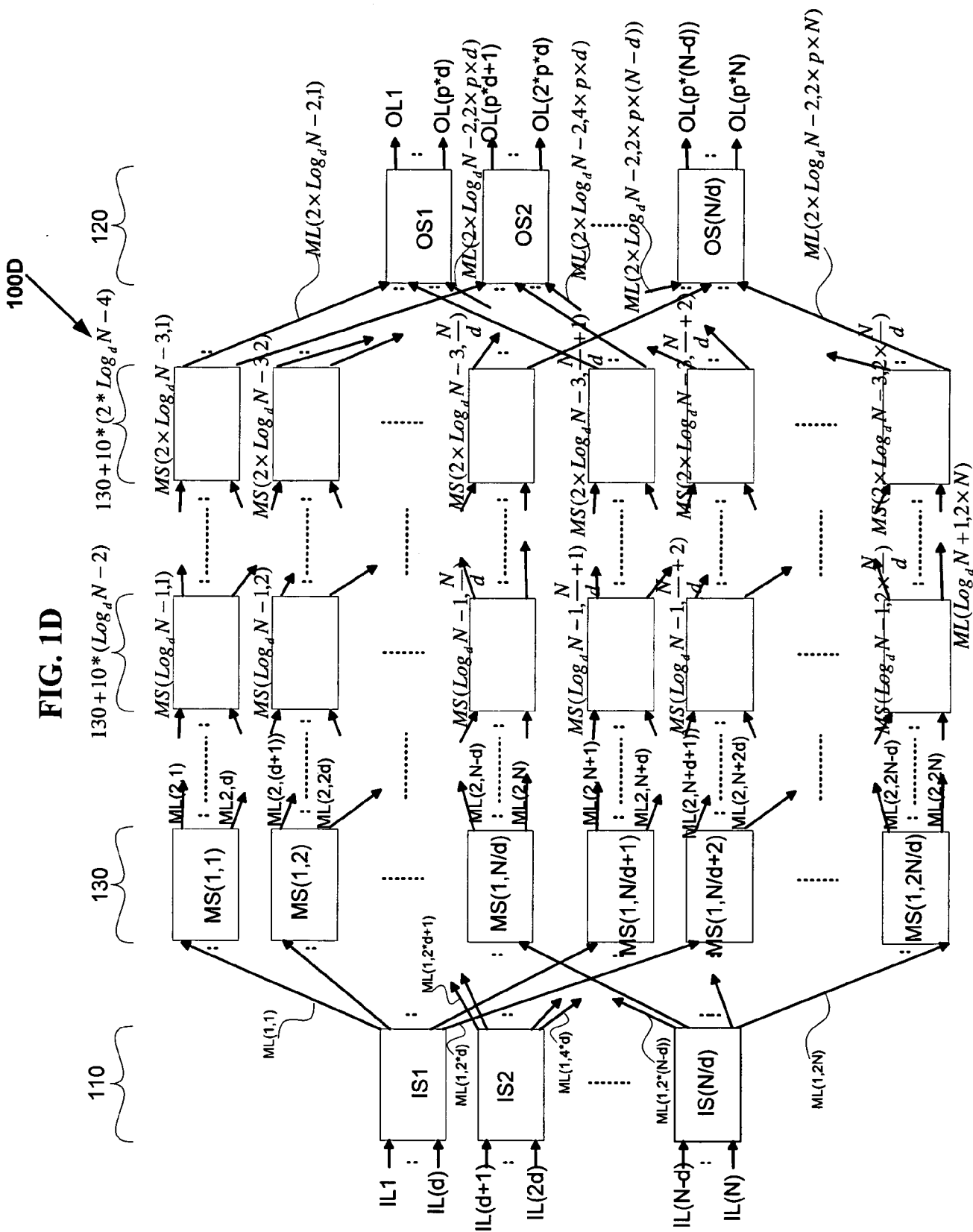


FIG. 1E

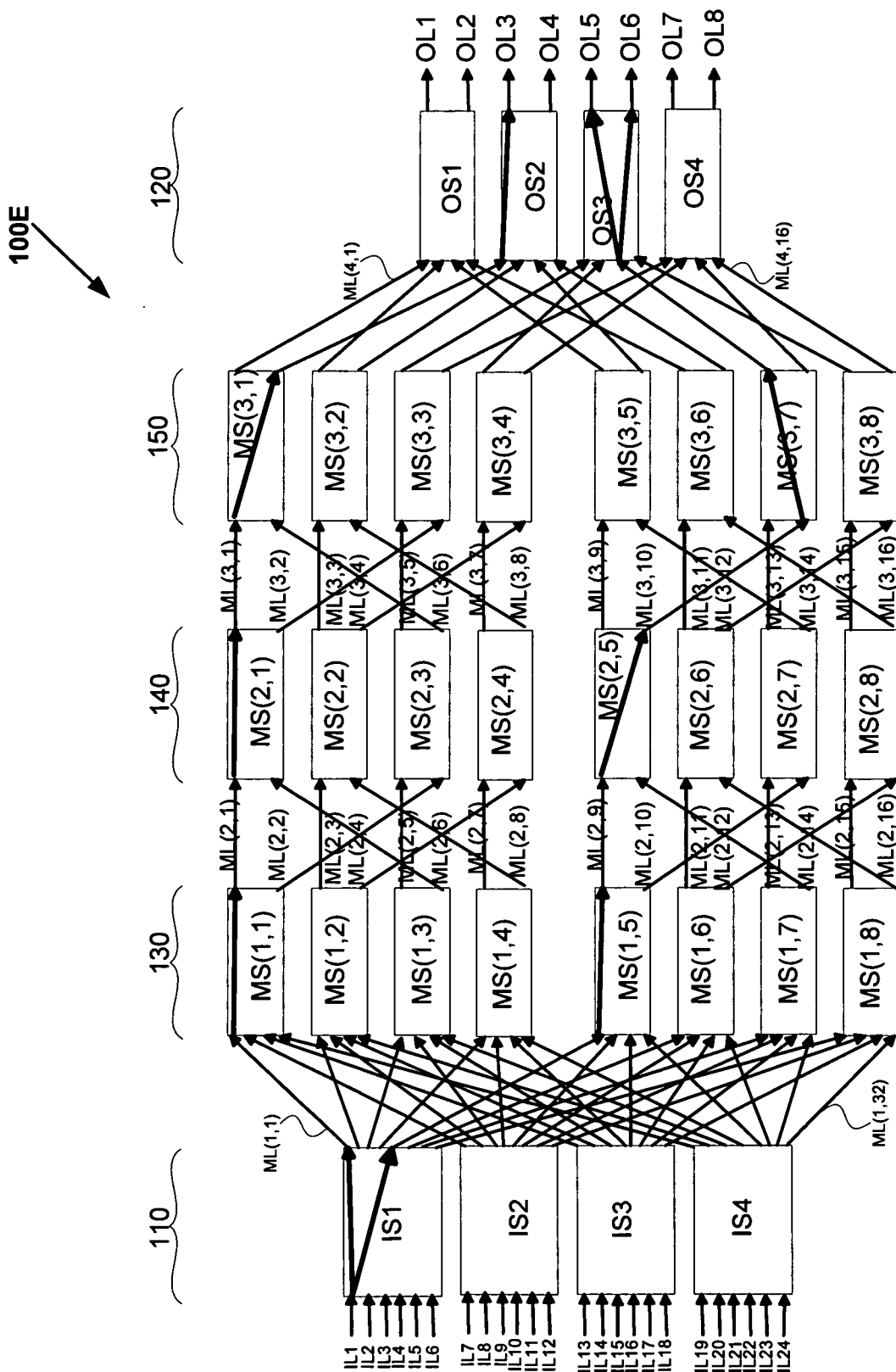


FIG. 1E1

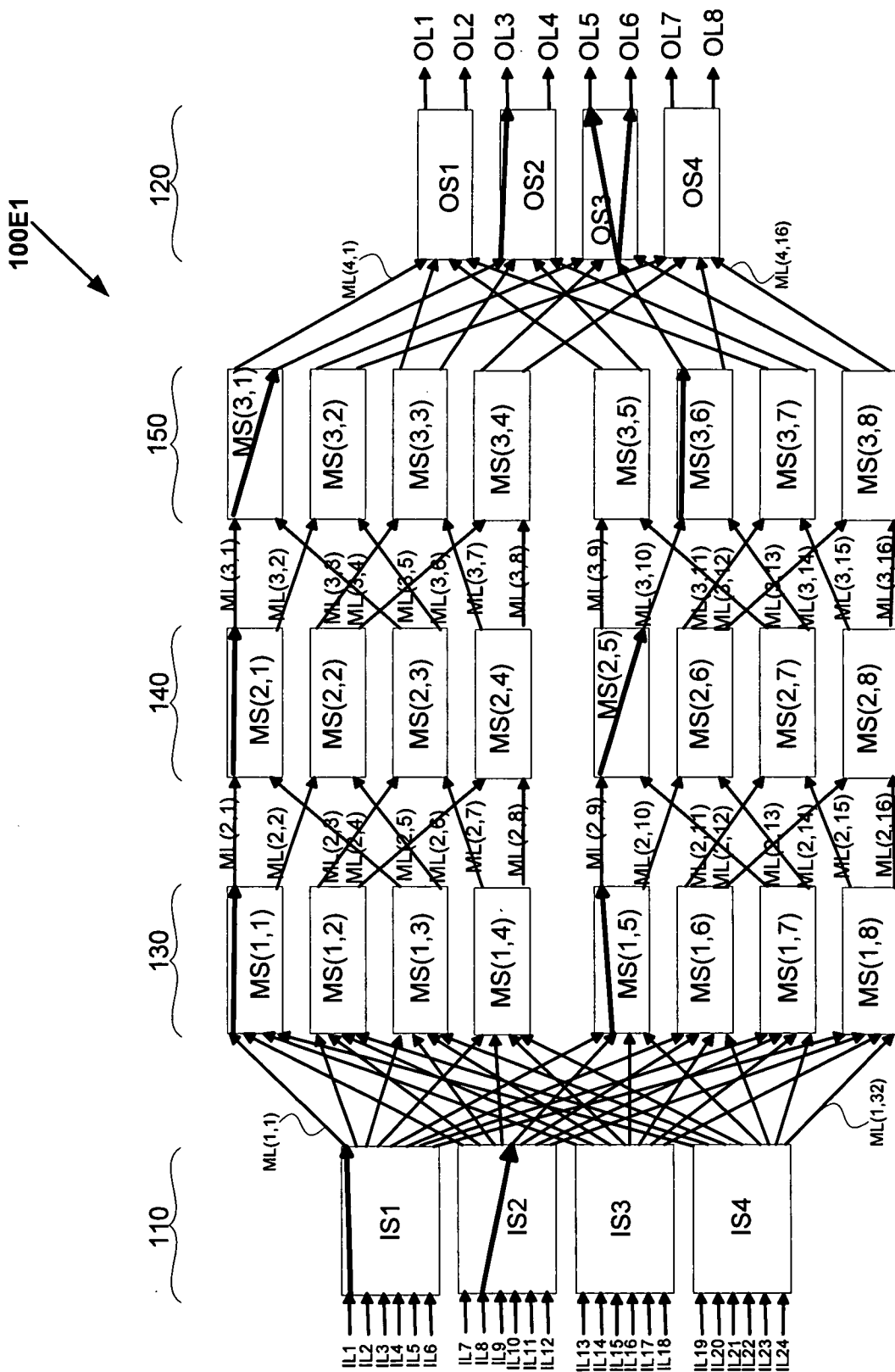


FIG. 1E2

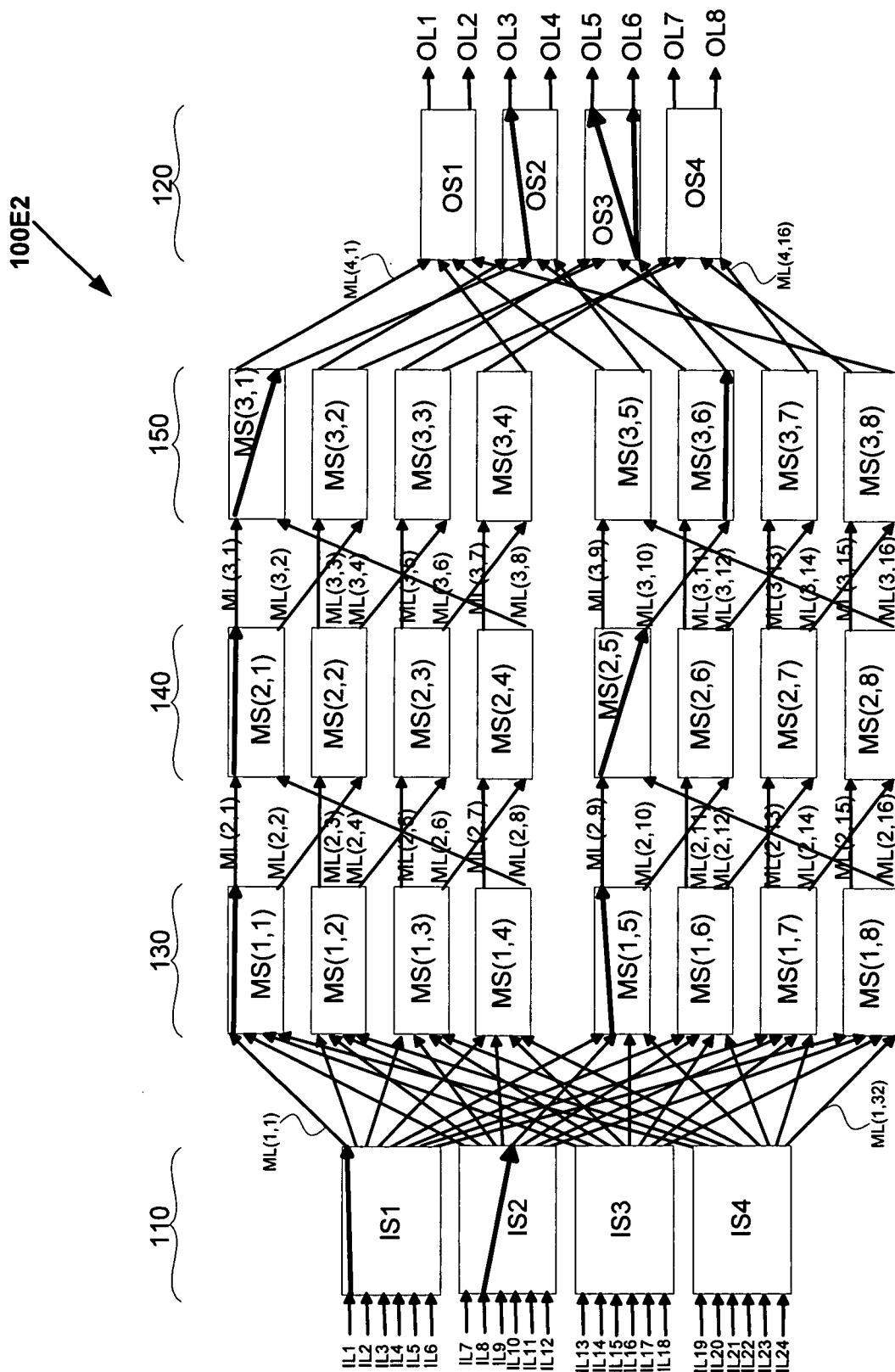
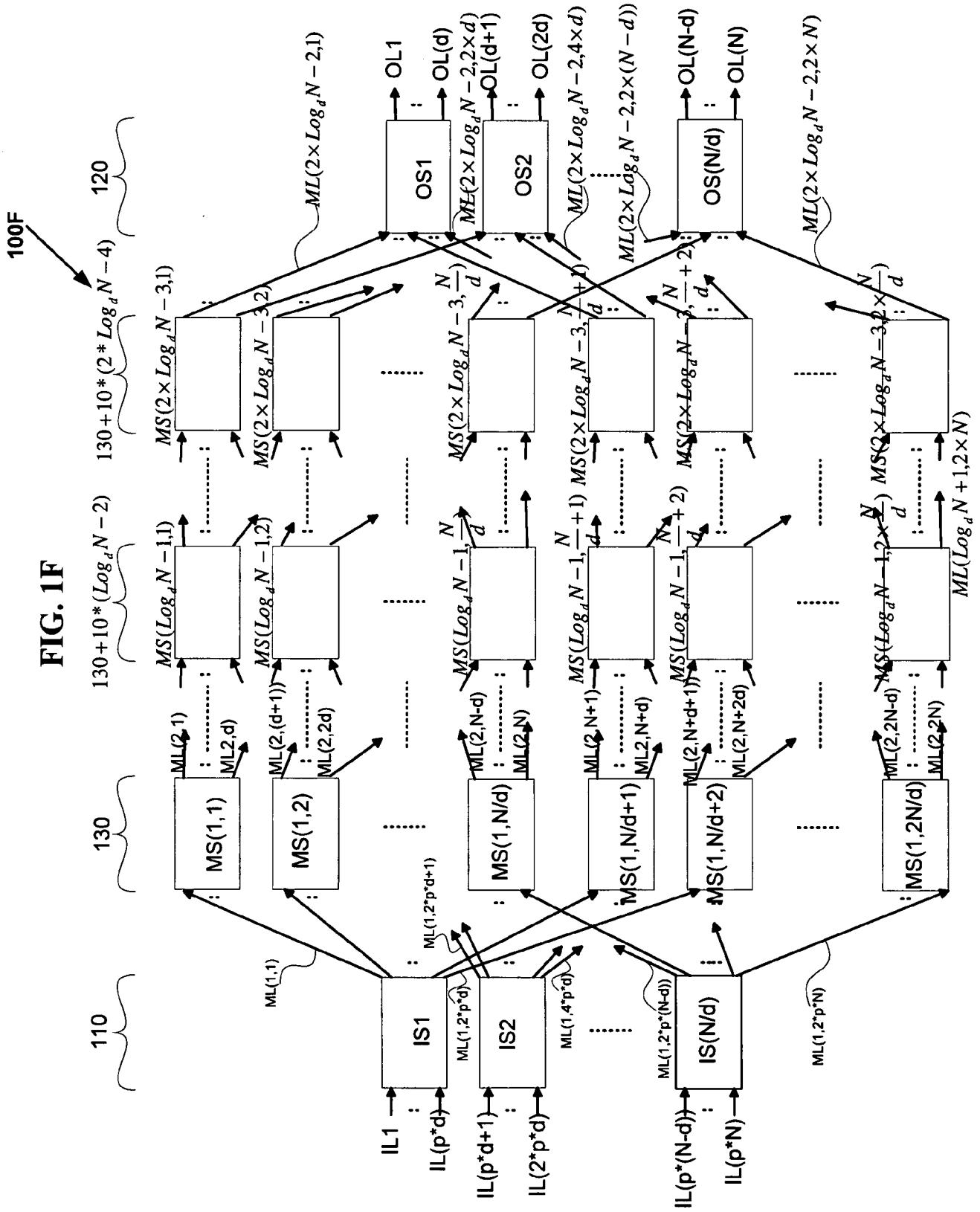


FIG. 1F



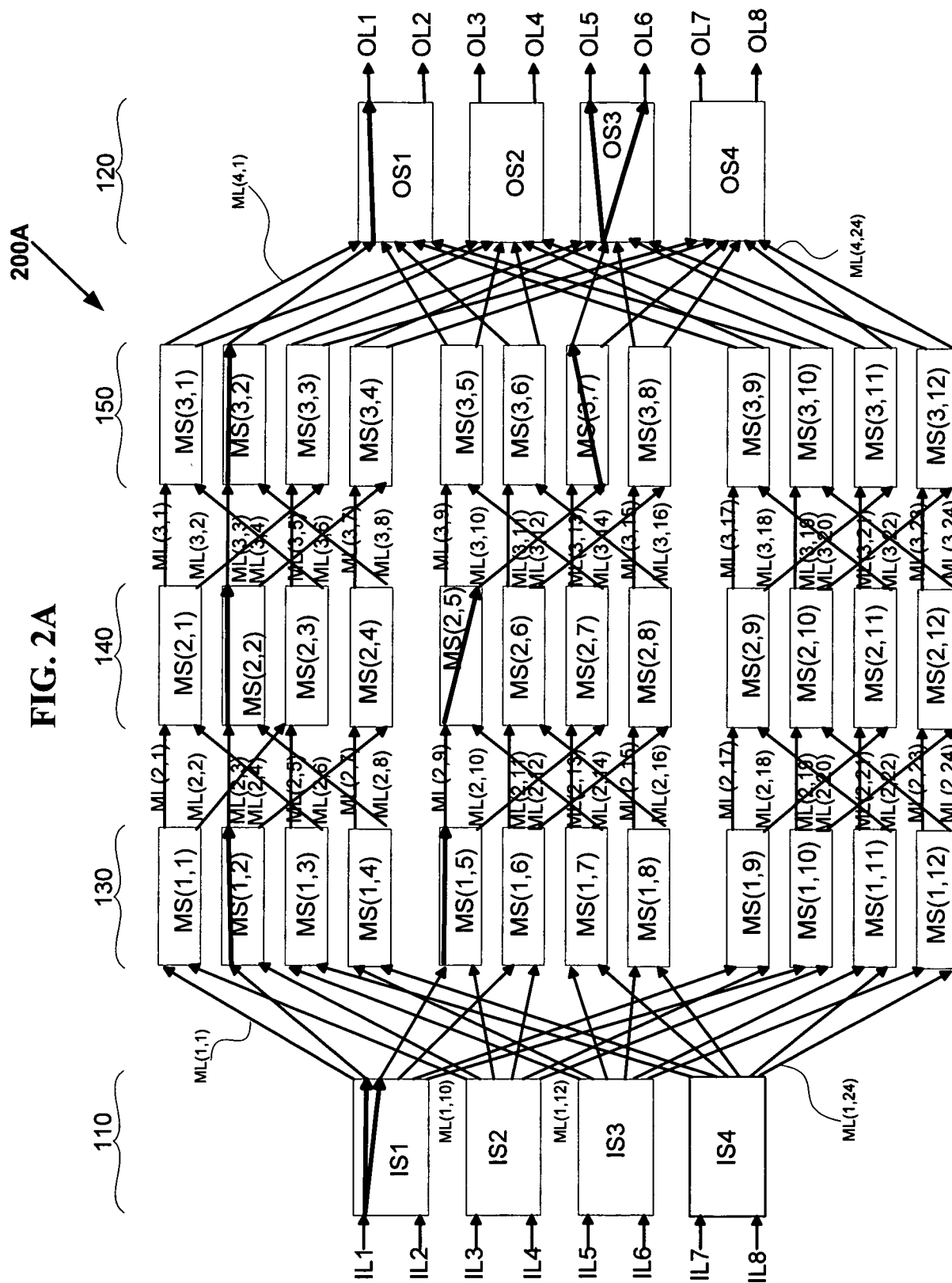
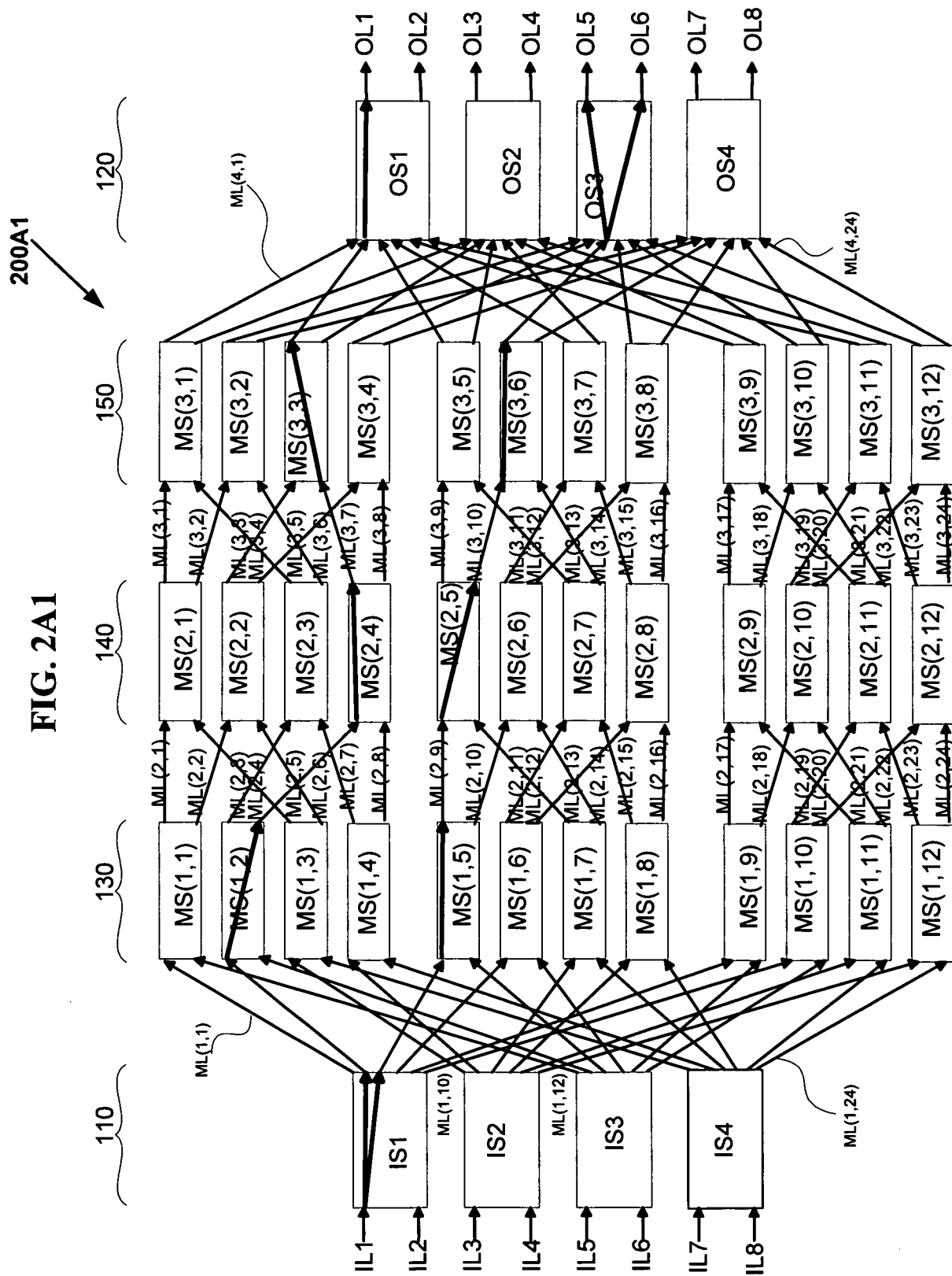


FIG. 2A

200A



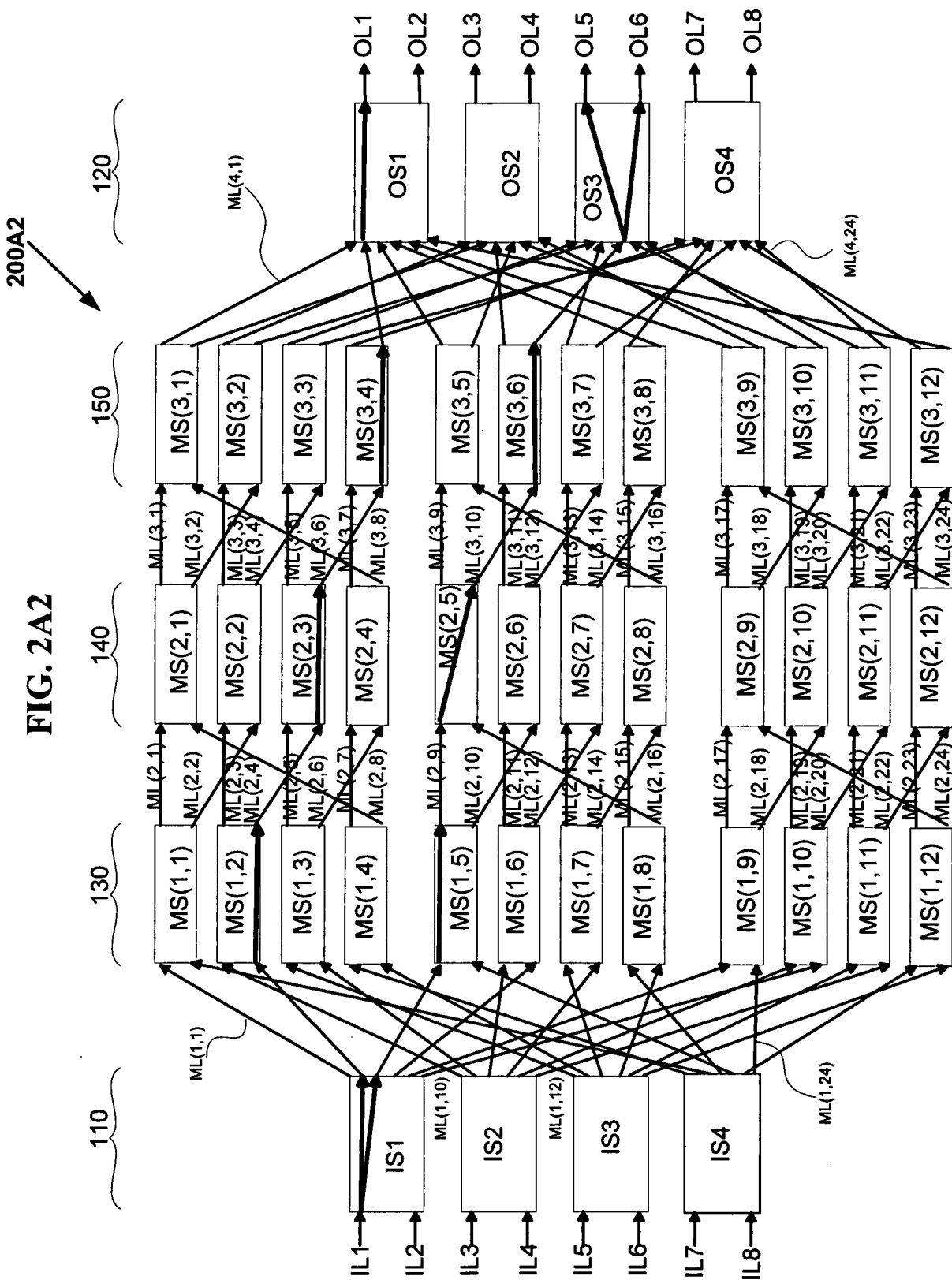
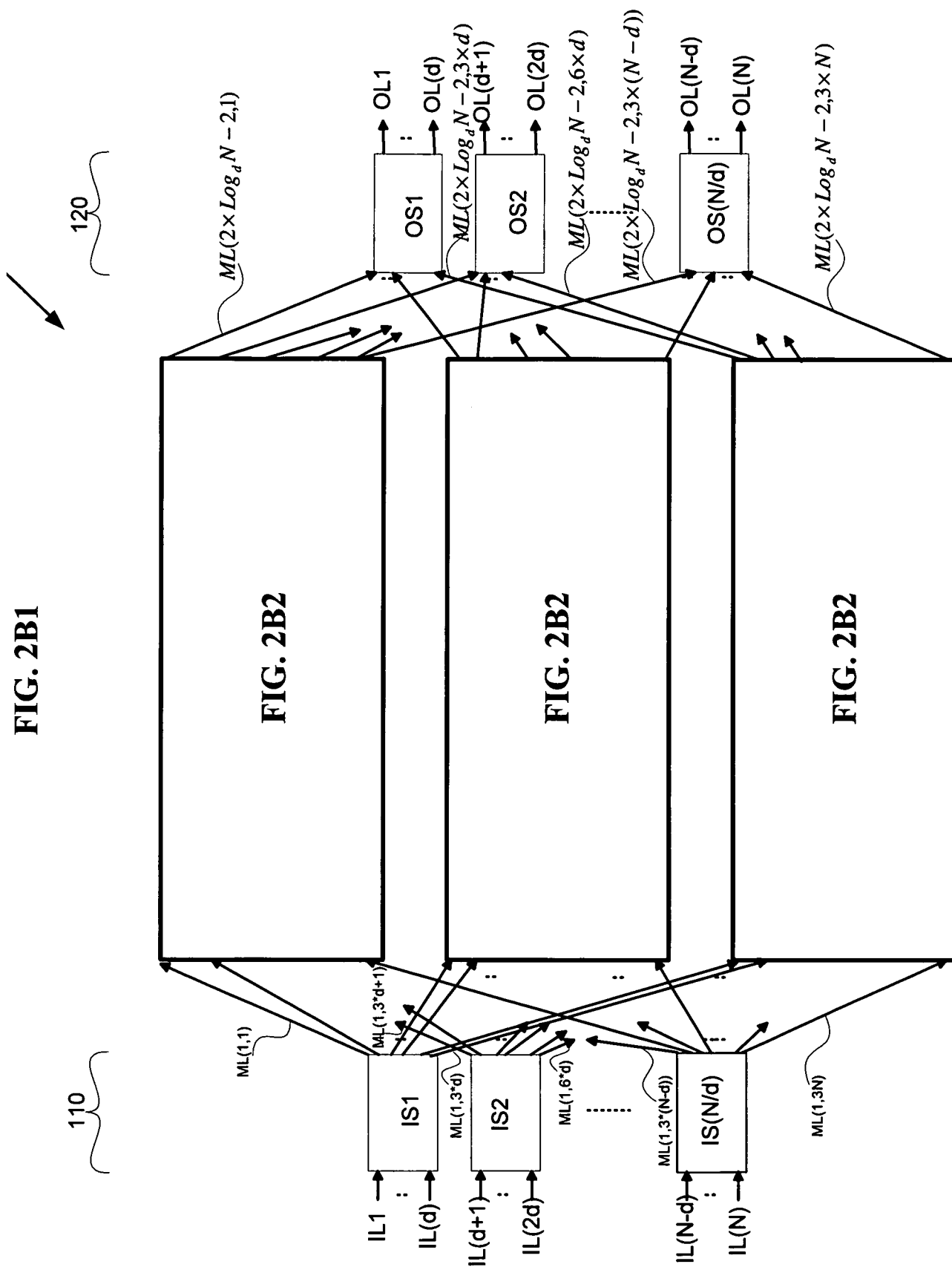
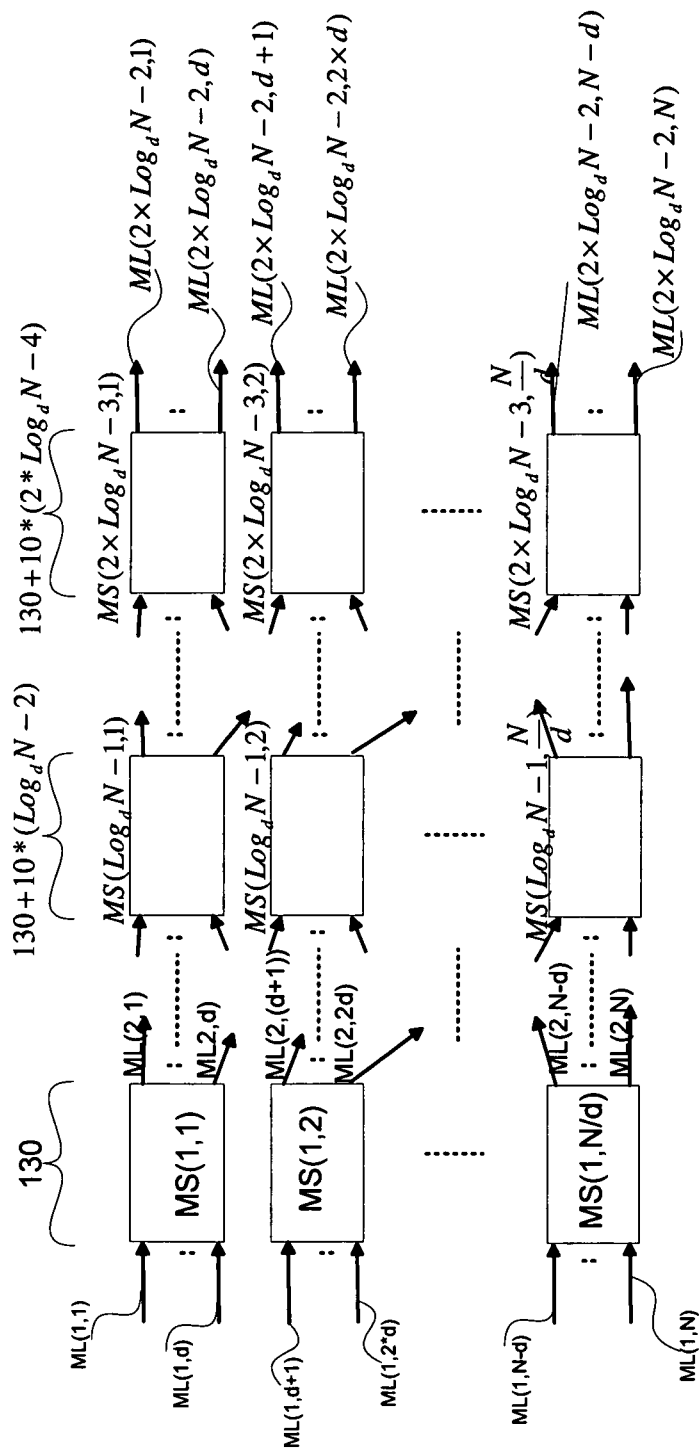


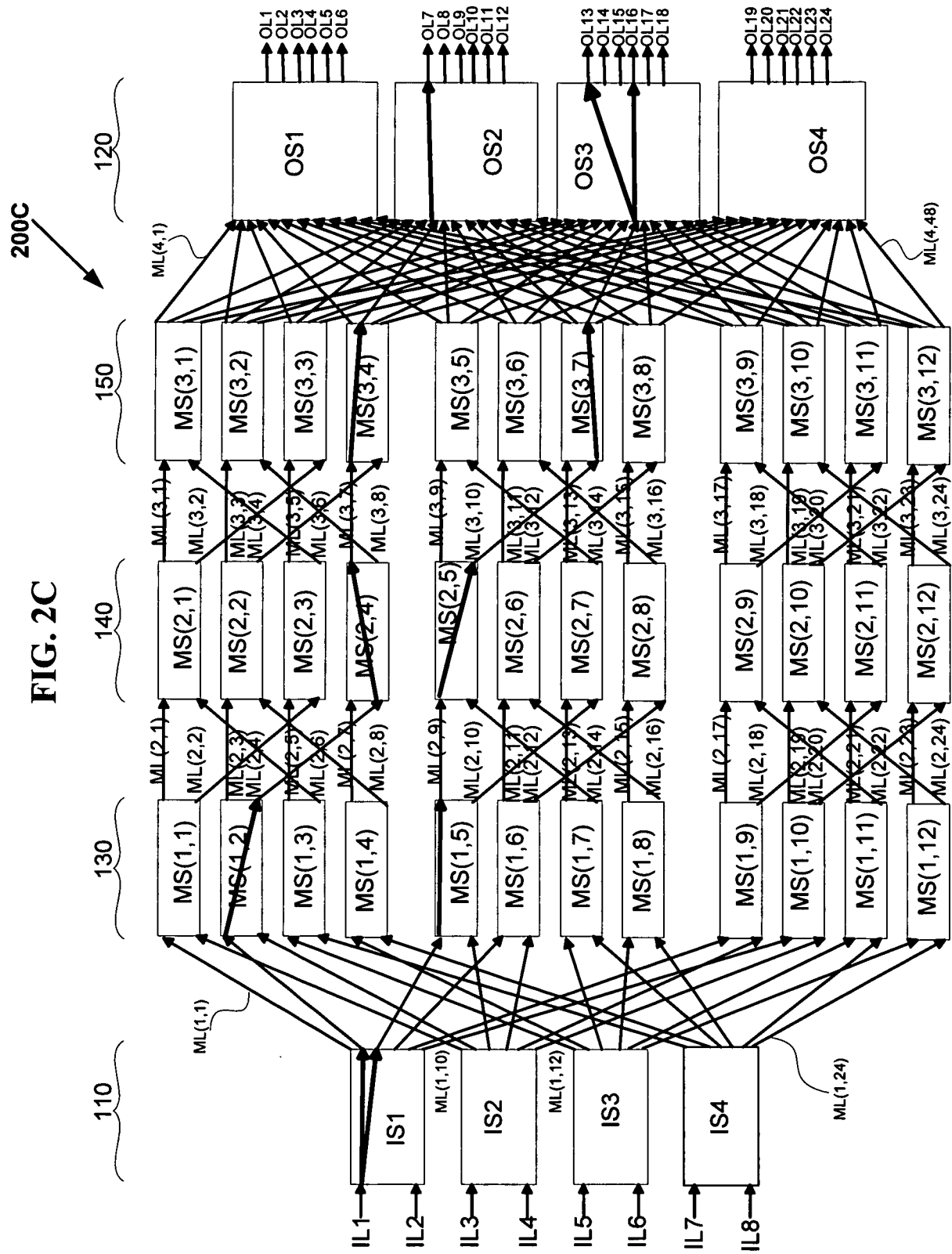
FIG. 2A2



200B2

FIG. 2B2





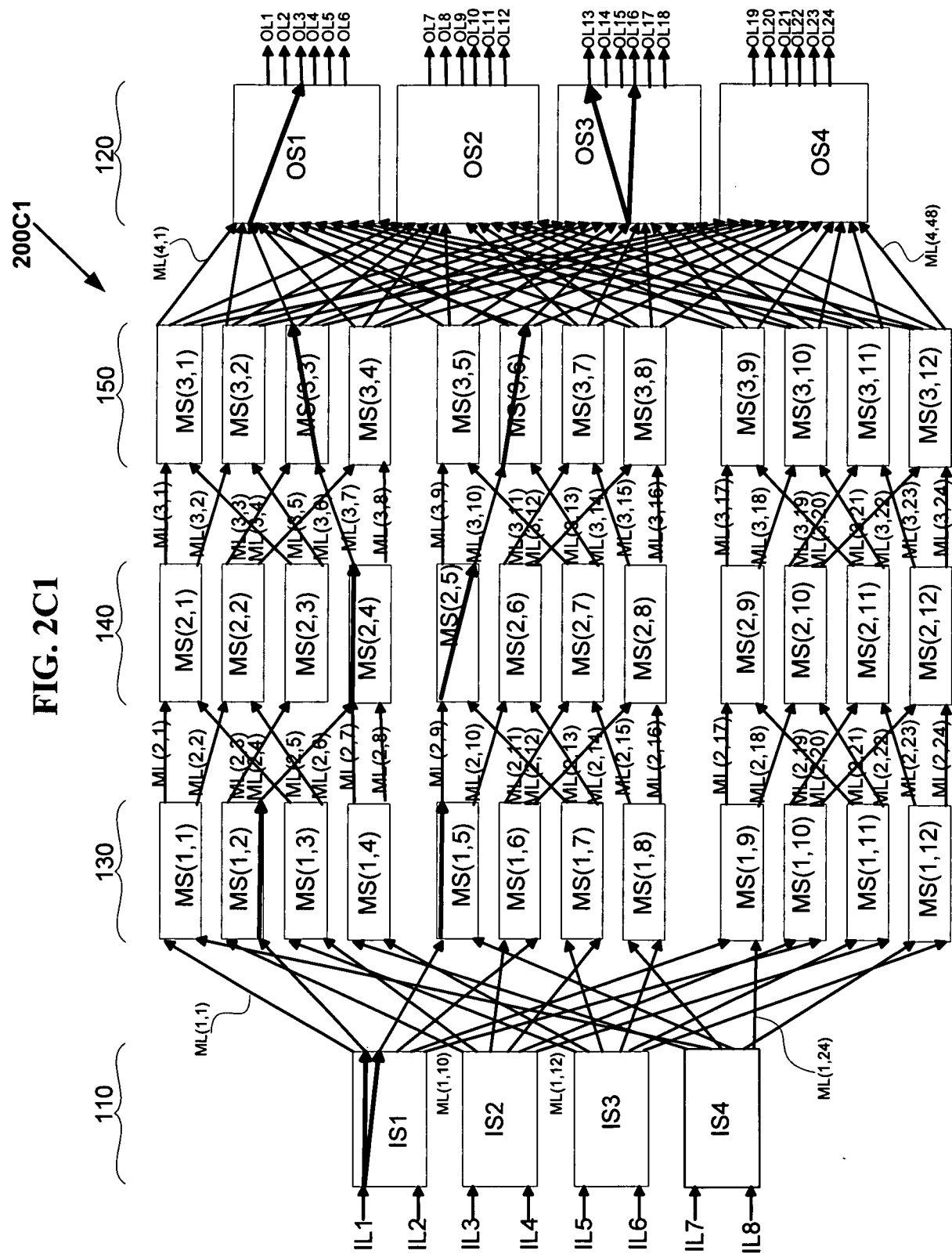


FIG. 2C1

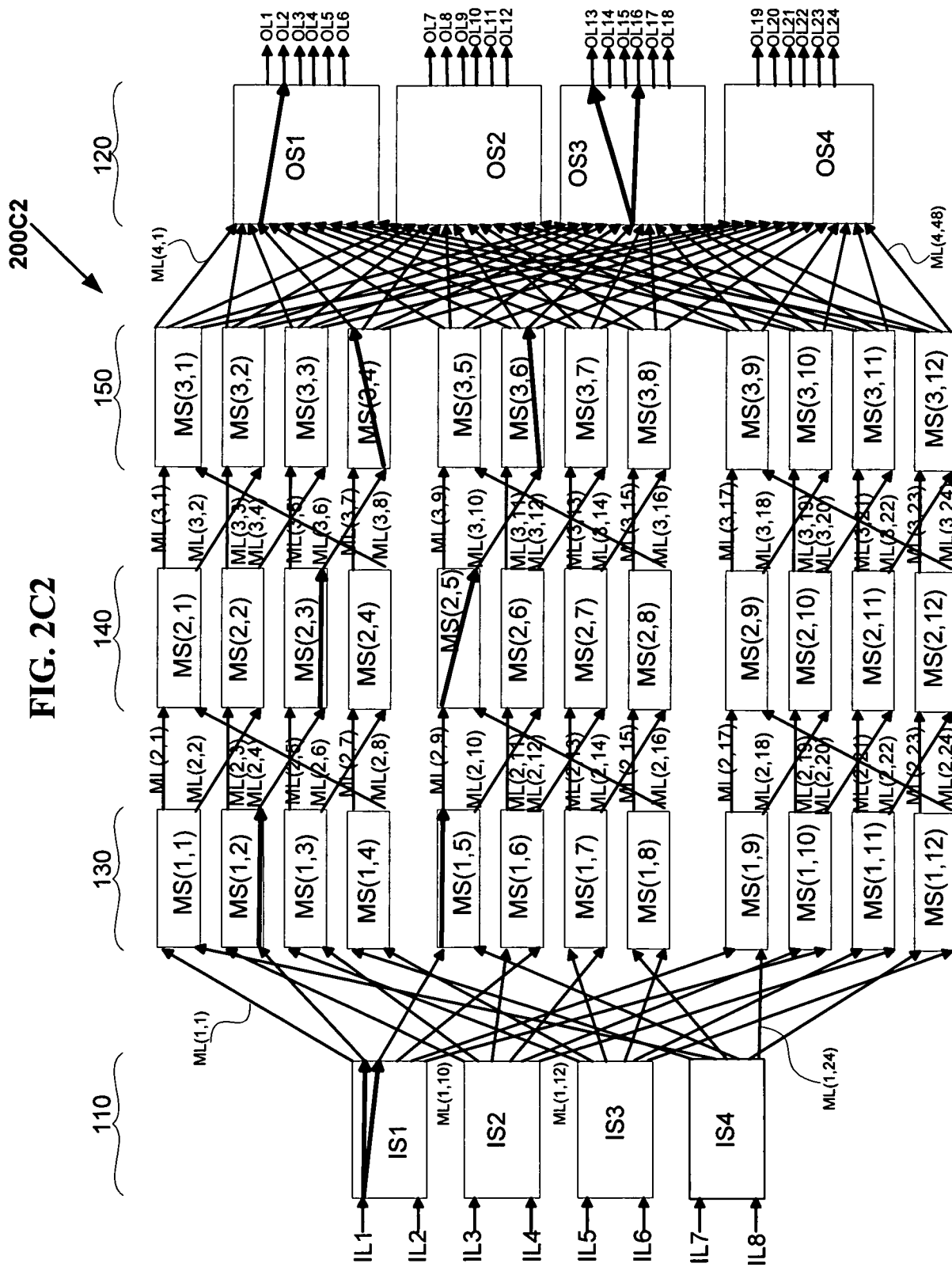


FIG. 2C2

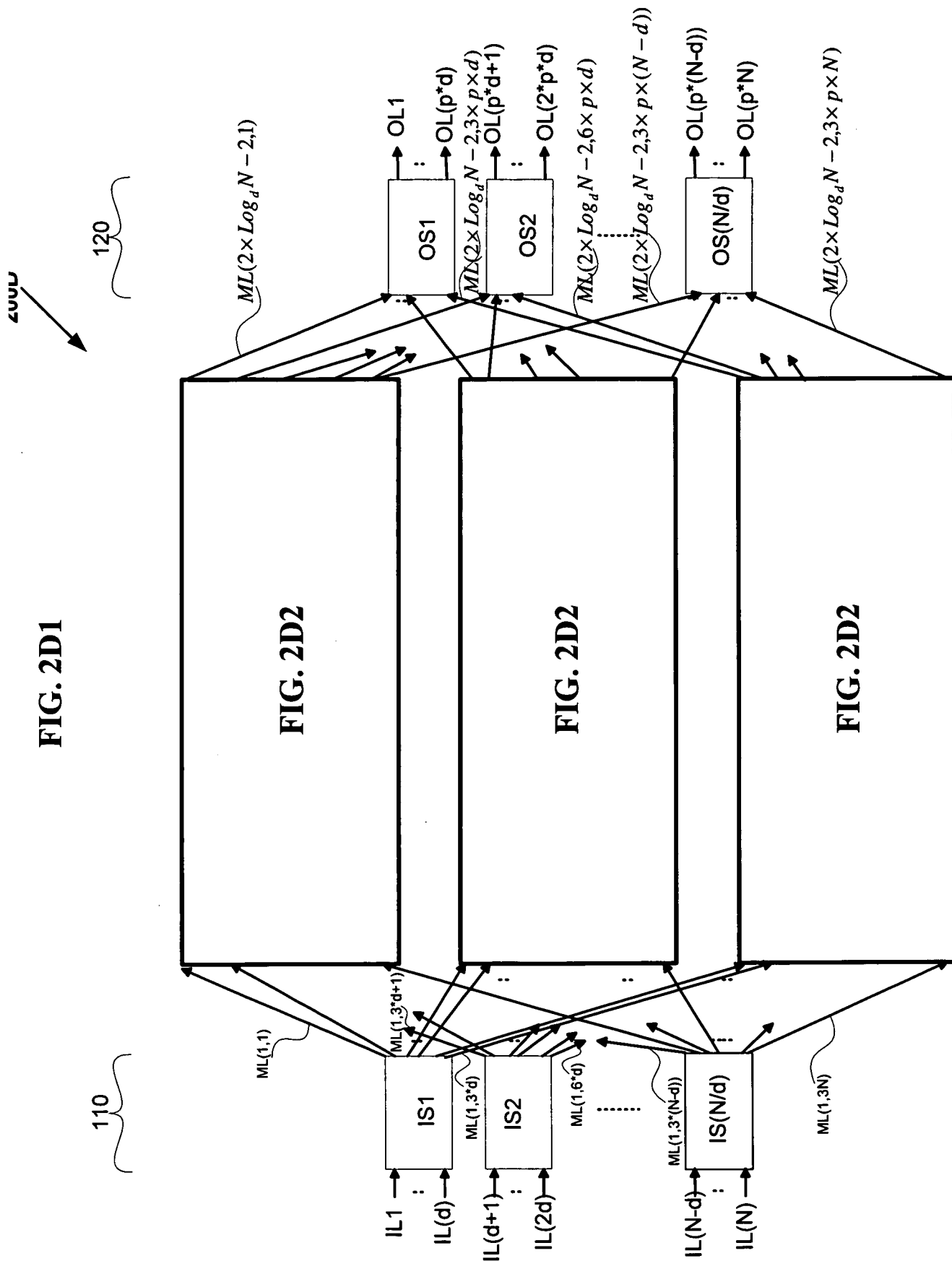
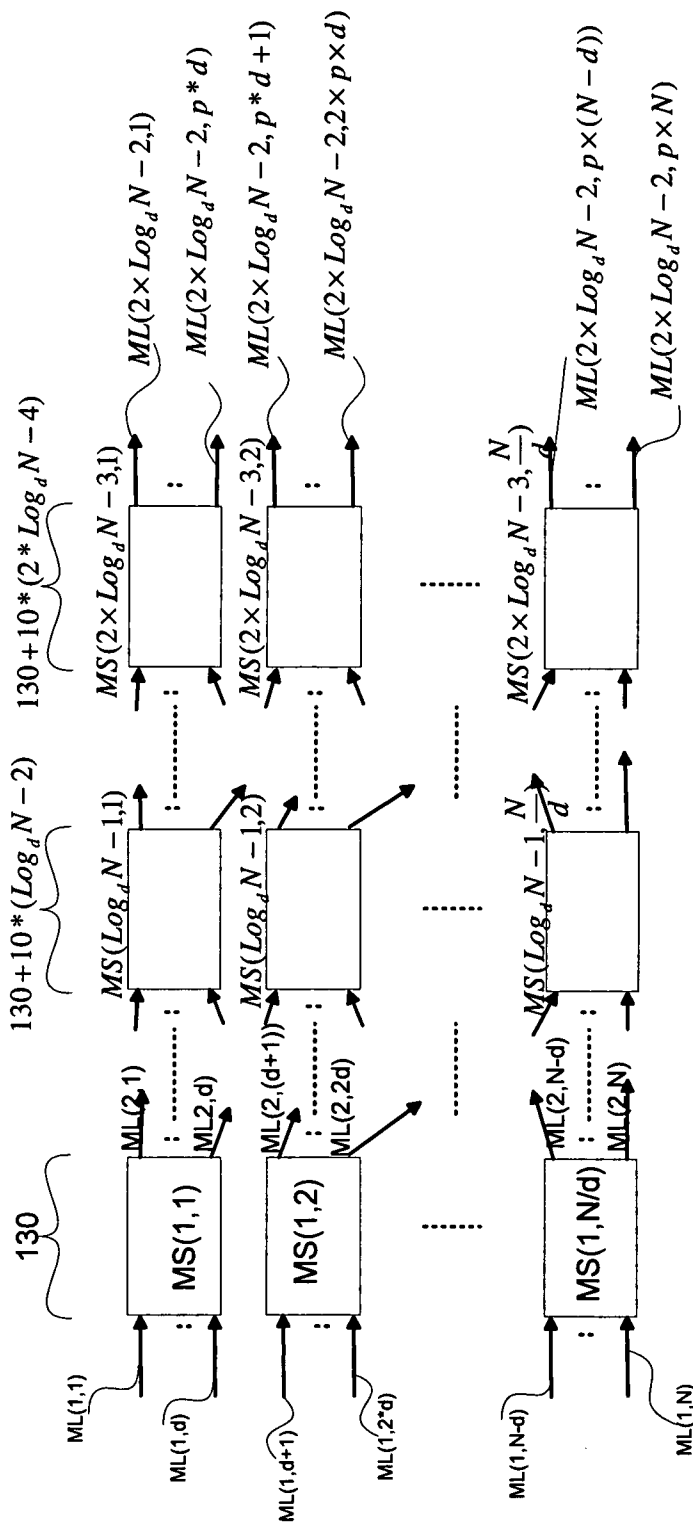


FIG. 2D2

200D2



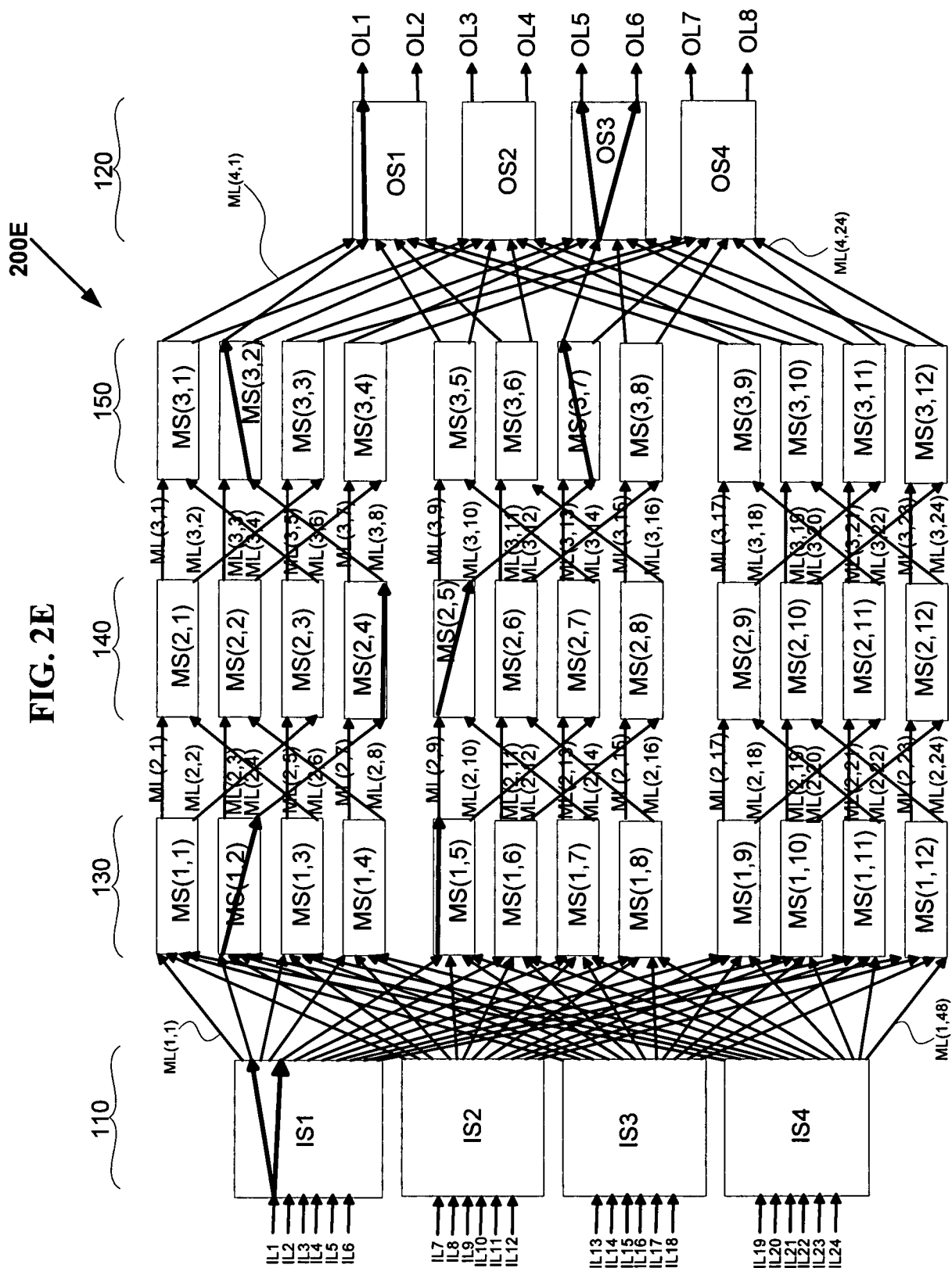


FIG. 2E

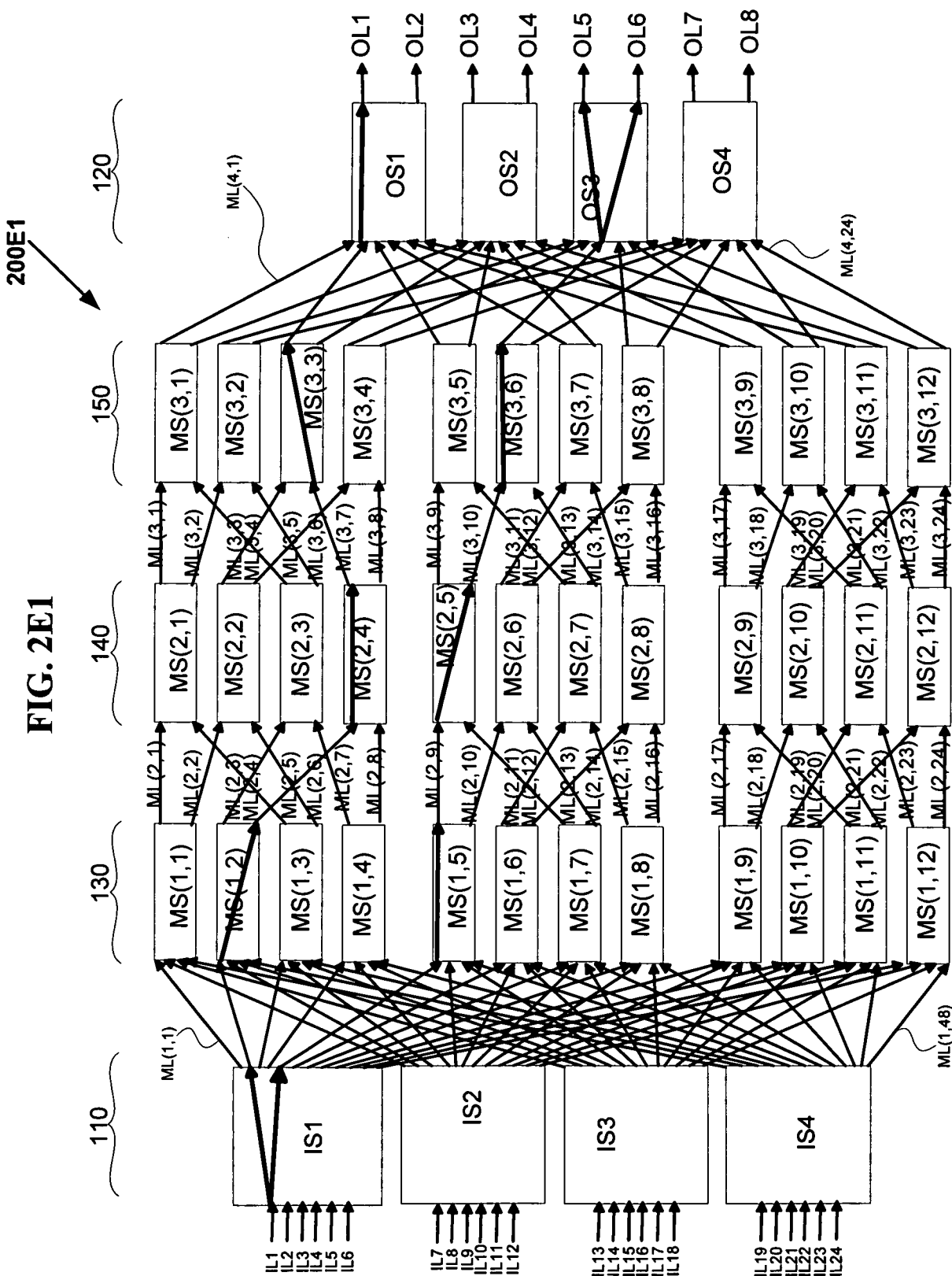
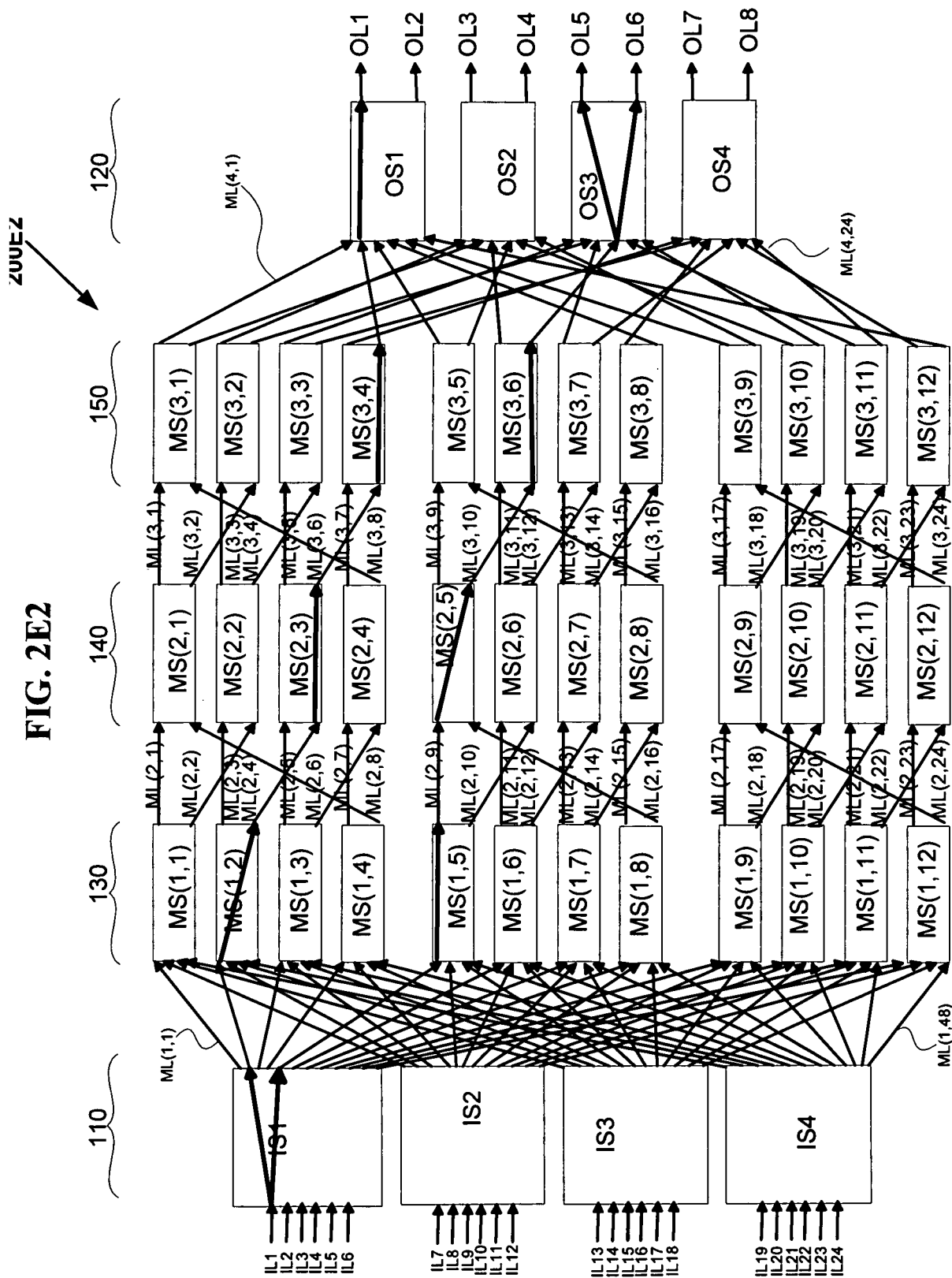


FIG. 2E1



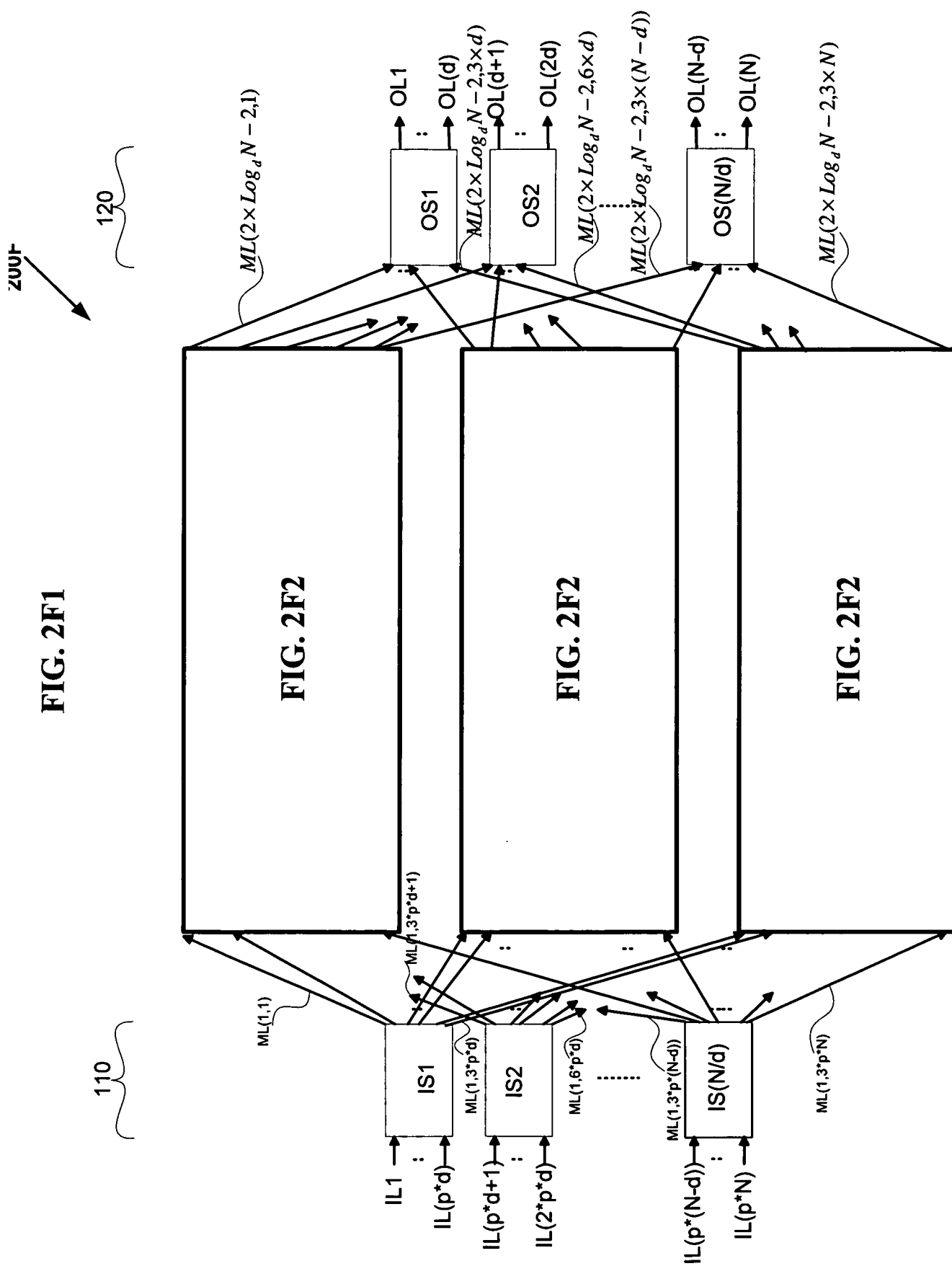
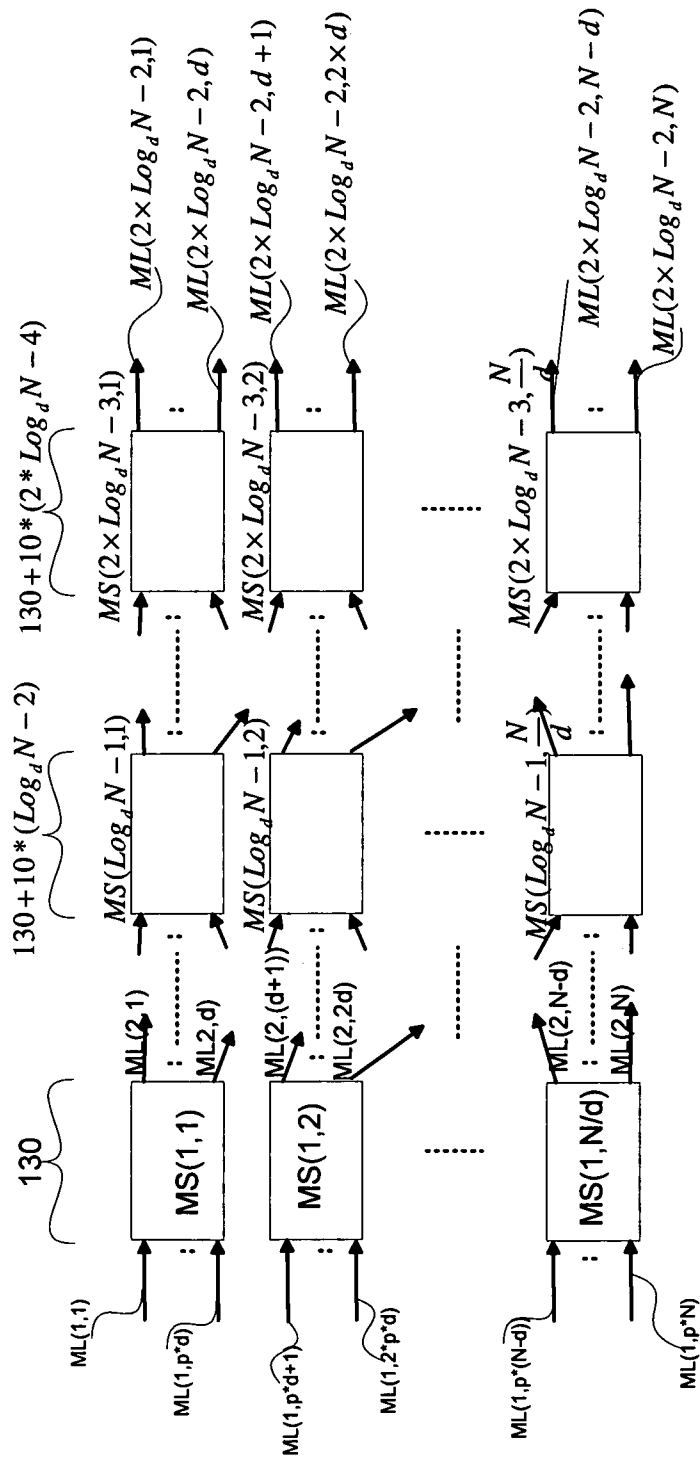


FIG. 2F2

200F2



WO 2008/109756

28/29

PCT/US2008/056064

FIG. 3A

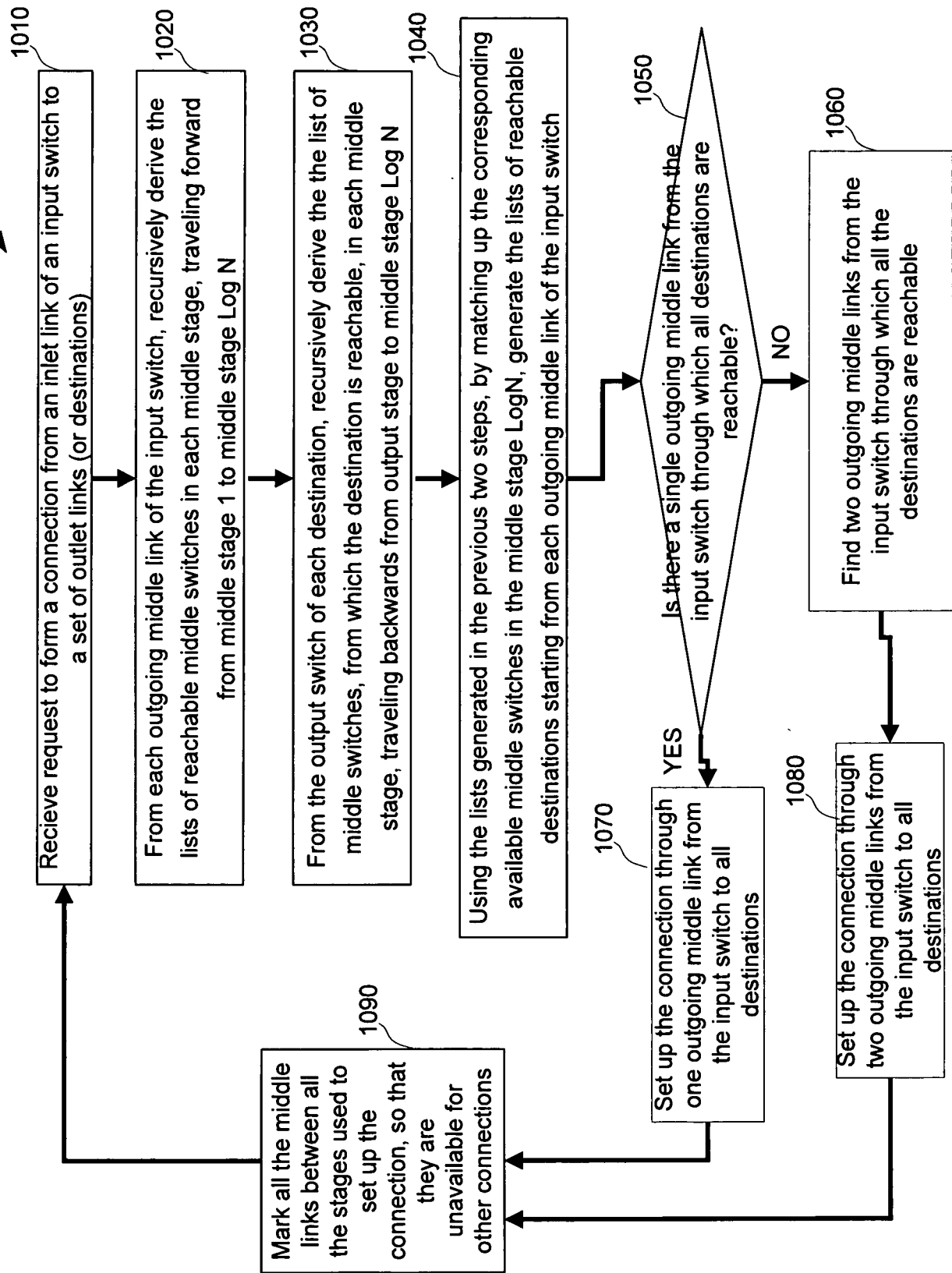


FIG. 4A

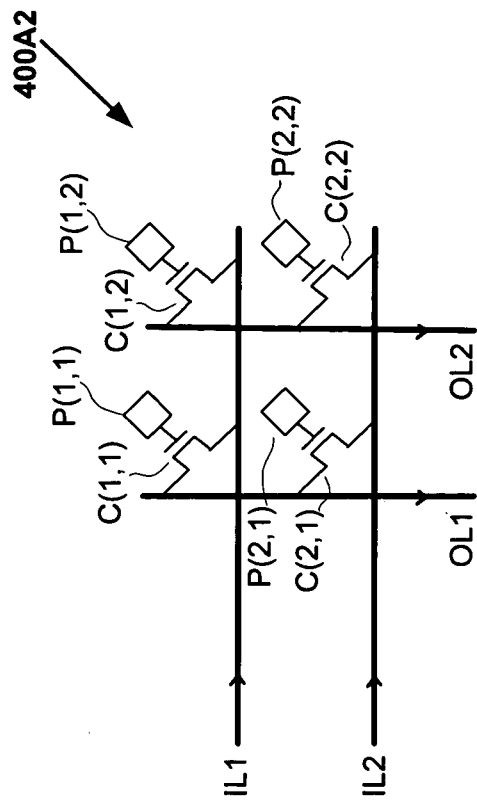


FIG. 4A2
(Prior Art)

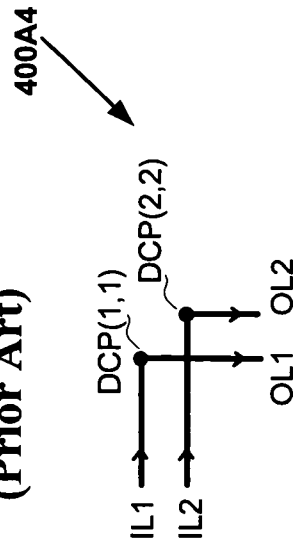


FIG. 4A4

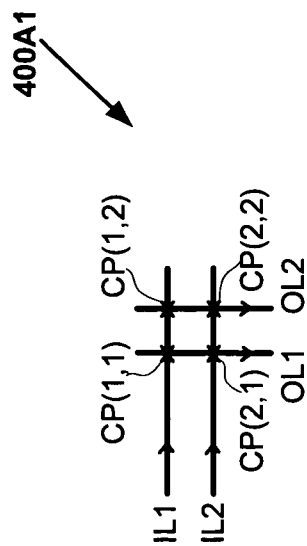


FIG. 4A1
(Prior Art)

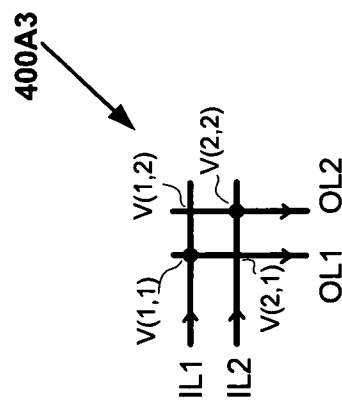


FIG. 4A3
(Prior Art)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 08/56064

A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - H04Q 3/00 (2008.04) USPC - 340/2.2 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) USPC: 340/2.2 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched US: 340/2.2, 370/390,427 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) PubWEST(PGPB,USPT,EPAB,JPAB); DialogPRO(Engineering); Google Scholar; Search Terms:multistage, network, multistage network, input, output, multicast, unicast, broadcast, switch, stage, incoming, outgoing, topology, nonblocking, rearrangeably, strictly nonblocking, fanning out, controller, recursively, path, checking, middle.		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- Y	US 5,541,914 A (Krishnamoorthy et al.) 30 July 1996 (30.07.1996), abstract, Fig. 64, entire document, especially col. 2, ln 1-7, col 4, ln 23, col 5, ln 62, col 6, ln 7, col 12, ln 3, col 16, ln 4-7, col. 17, ln 14-15, col 17, ln 64-67, col. 18, ln 10-14, col. 18, ln 14-15, ln 60-61, col 19, ln 5-6, ln 13-15, ln 15-20, col 28, ln 14-17, ln 46, col 29, ln 24-26, col 30, ln 58-62, col. 31, ln 11-16, col 36, ln 22, col. 40, ln 50, col. 42, line 45, col 45, ln 45-48, col. 59, ln 63-64, col 60, ln 29, col 63, ln 18-20, col. 65, ln 34-35, col. 66, ln 66-67, col 69, ln 32-38, col. 70, ln 33, col 76, ln 59, col 78, ln 15, and col 78, ln 20.	1-16 ----- 17-21
Y	US 5,666,360 A (Chen et al.) 09 September 1997 (09.09.1997), col 1, ln 25-26, col 2, ln 4-8, col 2, ln 10, ln 34, ln 43-45, col 6, ln 26, ln 47-50, col 7, ln 4-5, col 8, ln 9, col 11, ln 27, col 12, ln 2, ln 20, col 15, ln 9, ln 57, col 17, ln 54-59, ln 67, col 18, ln 1, ln 8-12, col 20, ln 14, and col 23. ln 4.	17-21
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>		
* 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 30 April 2008 (30.04.2008)		Date of mailing of the international search report 28 MAY 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: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

EXHIBIT A



US 20100135286A1

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

(10) **Pub. No.: US 2010/0135286 A1**
 (43) **Pub. Date: Jun. 3, 2010**

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

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

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

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

$$\frac{2 \times N}{d}$$

(21) Appl. No.: **12/530,207**

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, and d outgoing links connecting to the switches in its immediate succeeding stage. Also the same 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.

(22) PCT Filed: **Mar. 6, 2008**

(86) PCT No.: **PCT/US2008/056064**

§ 371 (c)(1),
 (2), (4) Date: **Sep. 6, 2009**

A 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

Related U.S. Application Data

(60) Provisional application No. 60/905,526, filed on Mar. 6, 2007, provisional application No. 60/940,383, filed on May 25, 2007.

Publication Classification

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

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

$$\frac{3 \times N}{d}$$

(57) **ABSTRACT**

A 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

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, and d outgoing links connecting to the switches in its immediate succeeding stage.

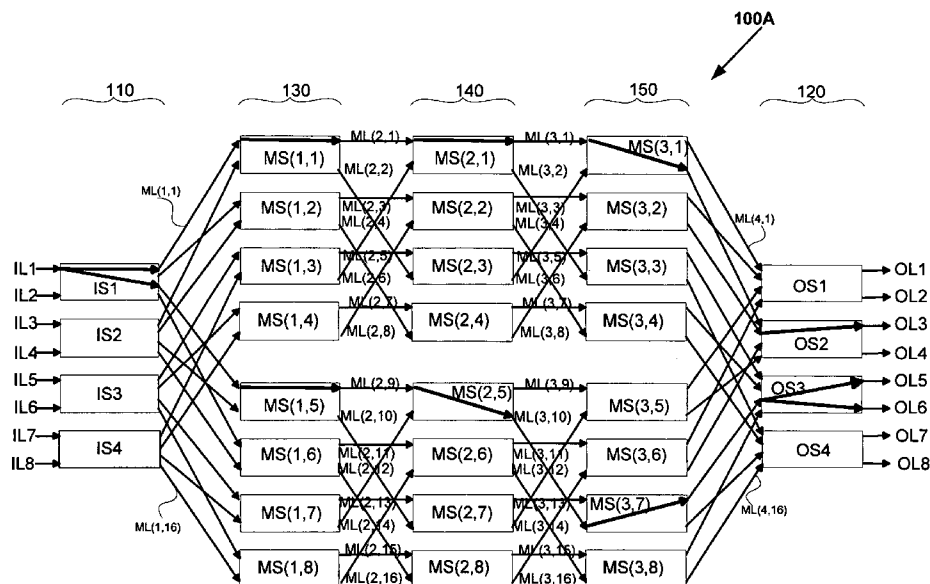


FIG. 1A

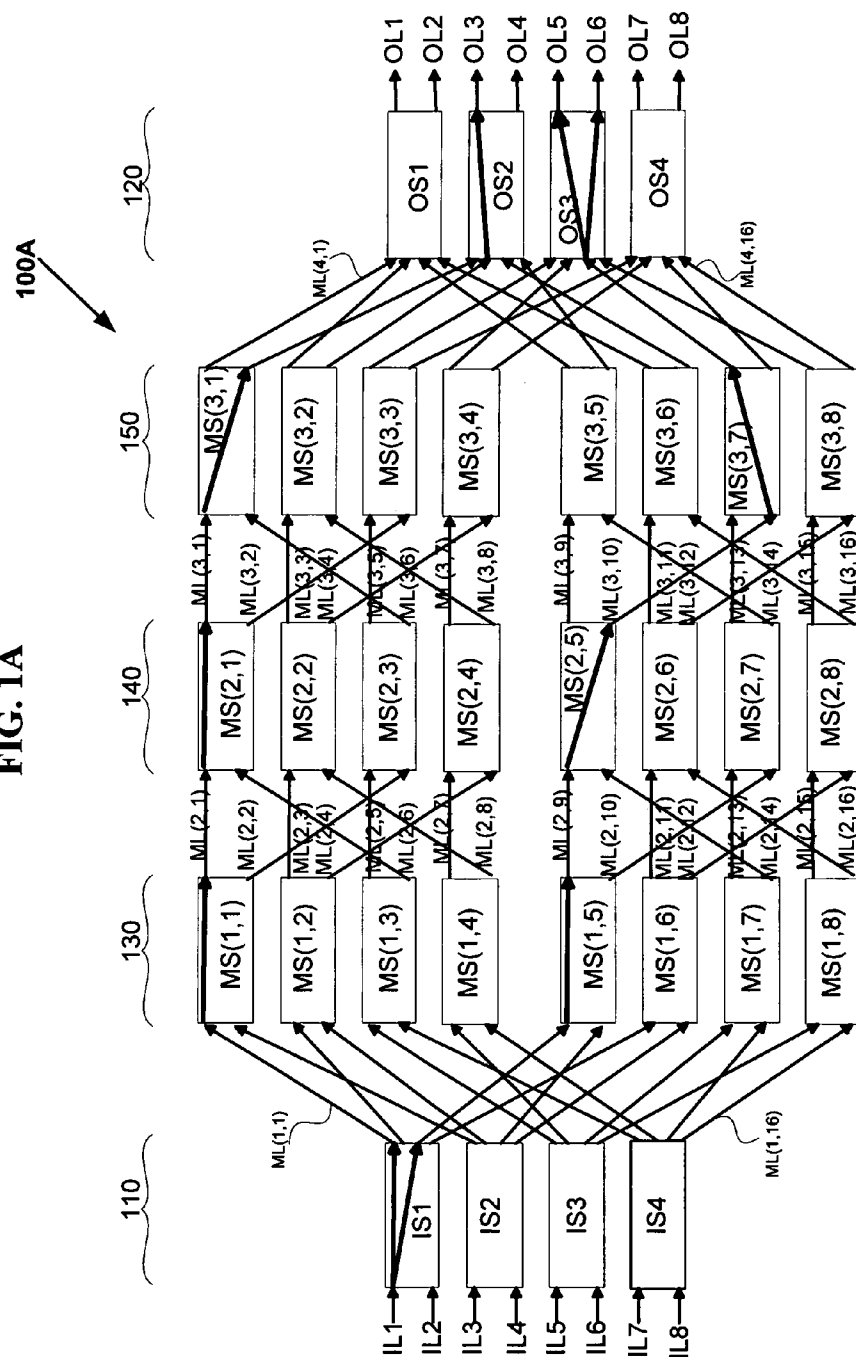


FIG. 1A1

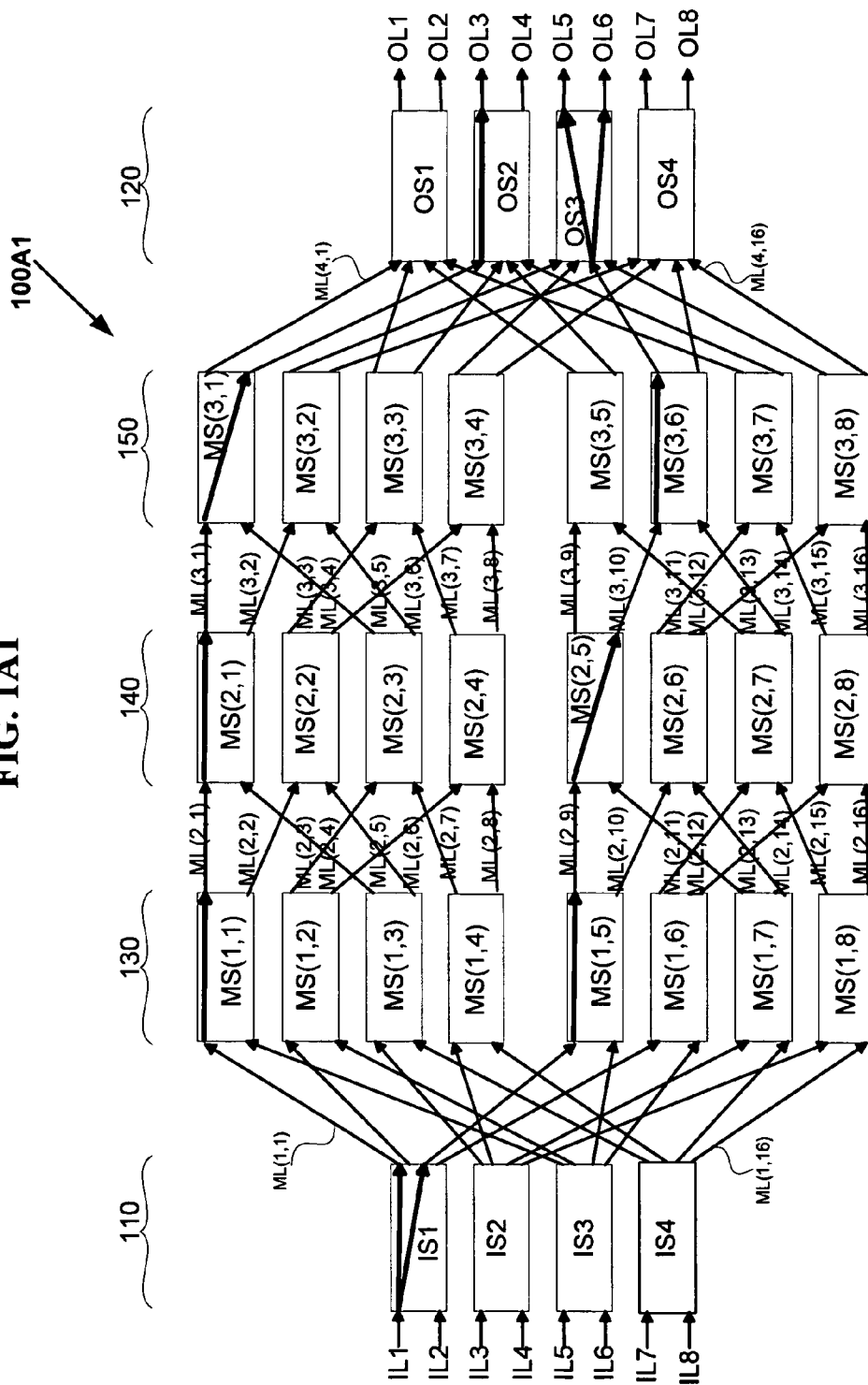
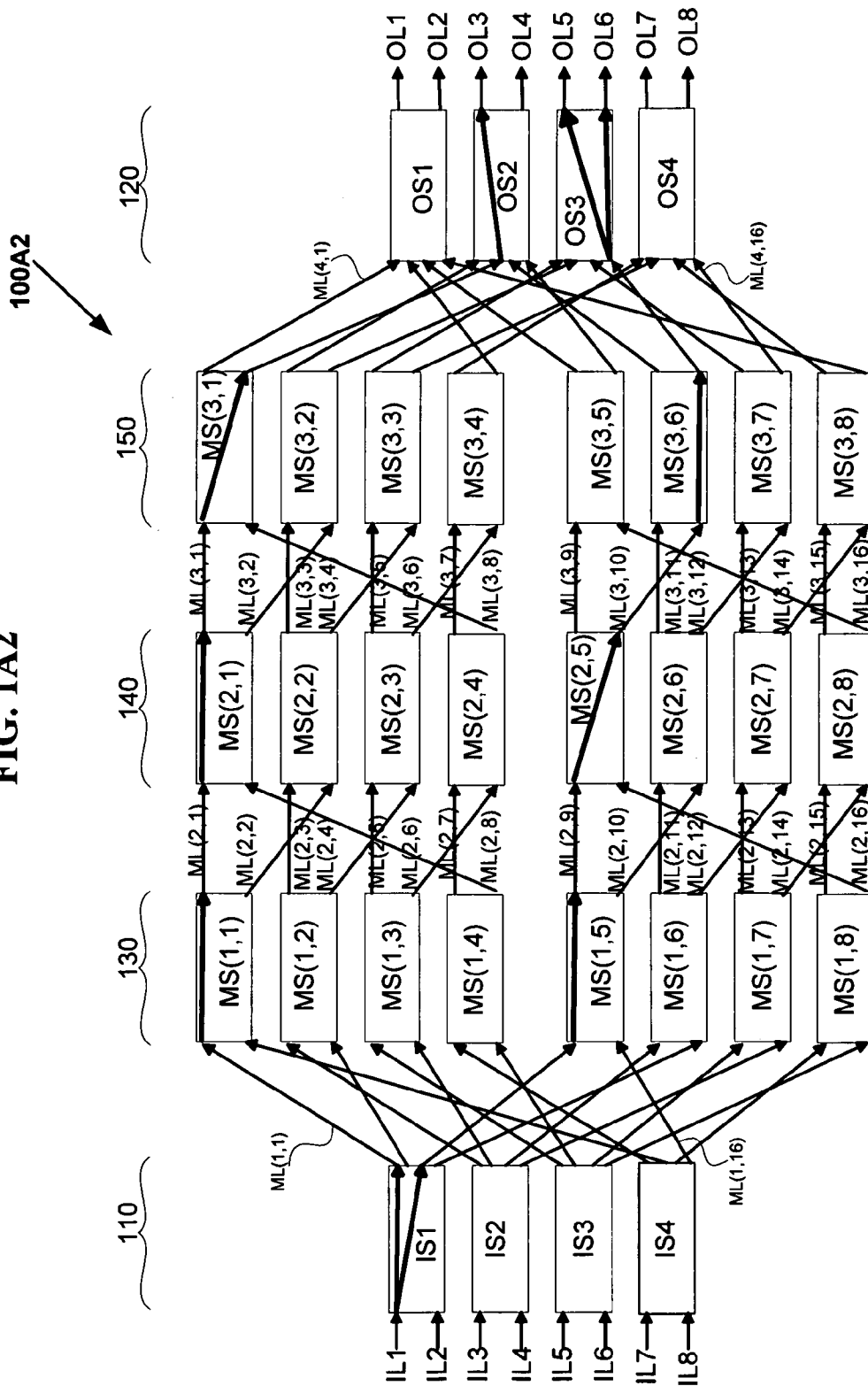


FIG. 1A2



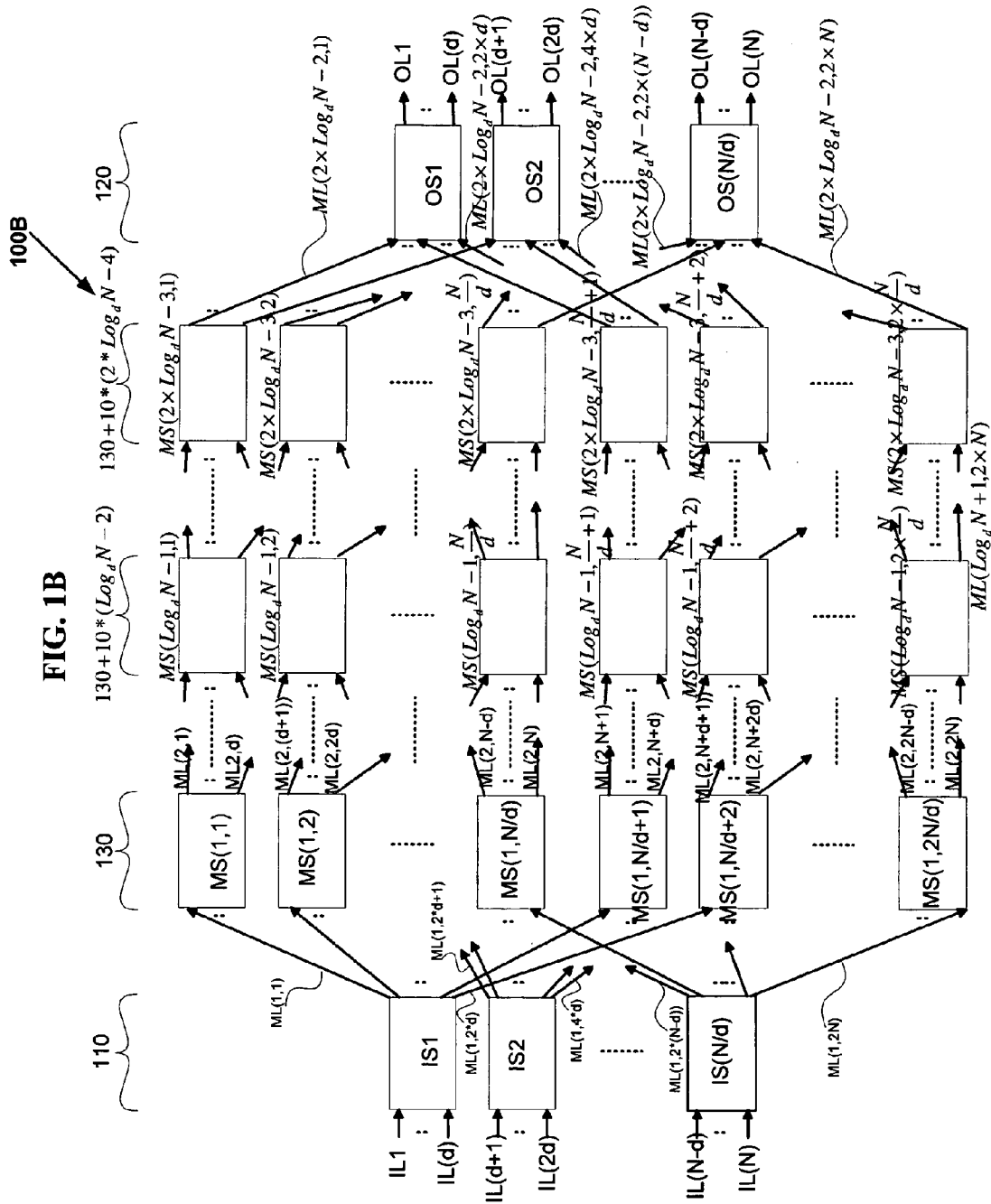


FIG. 1C

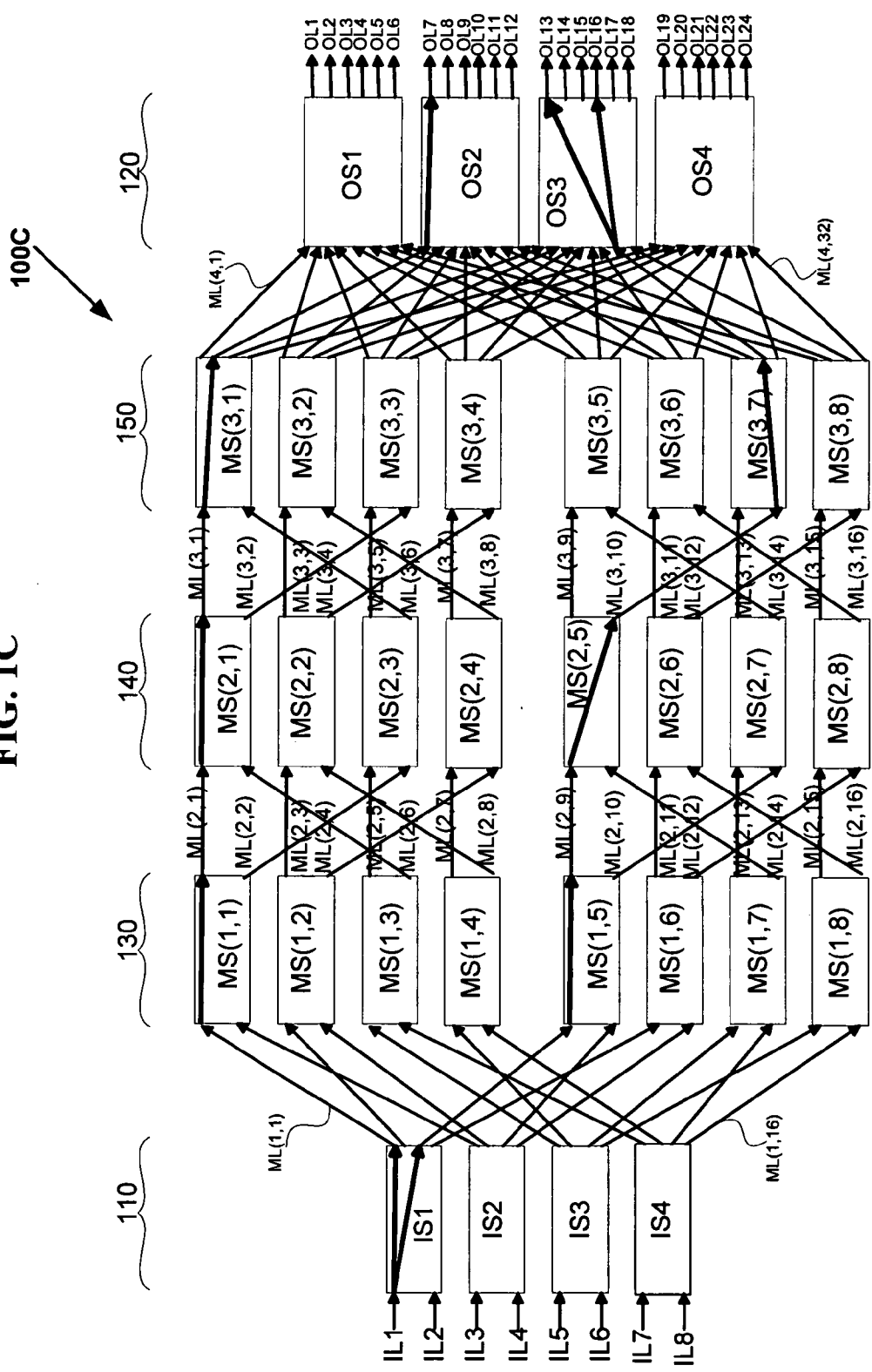


FIG. 1C1

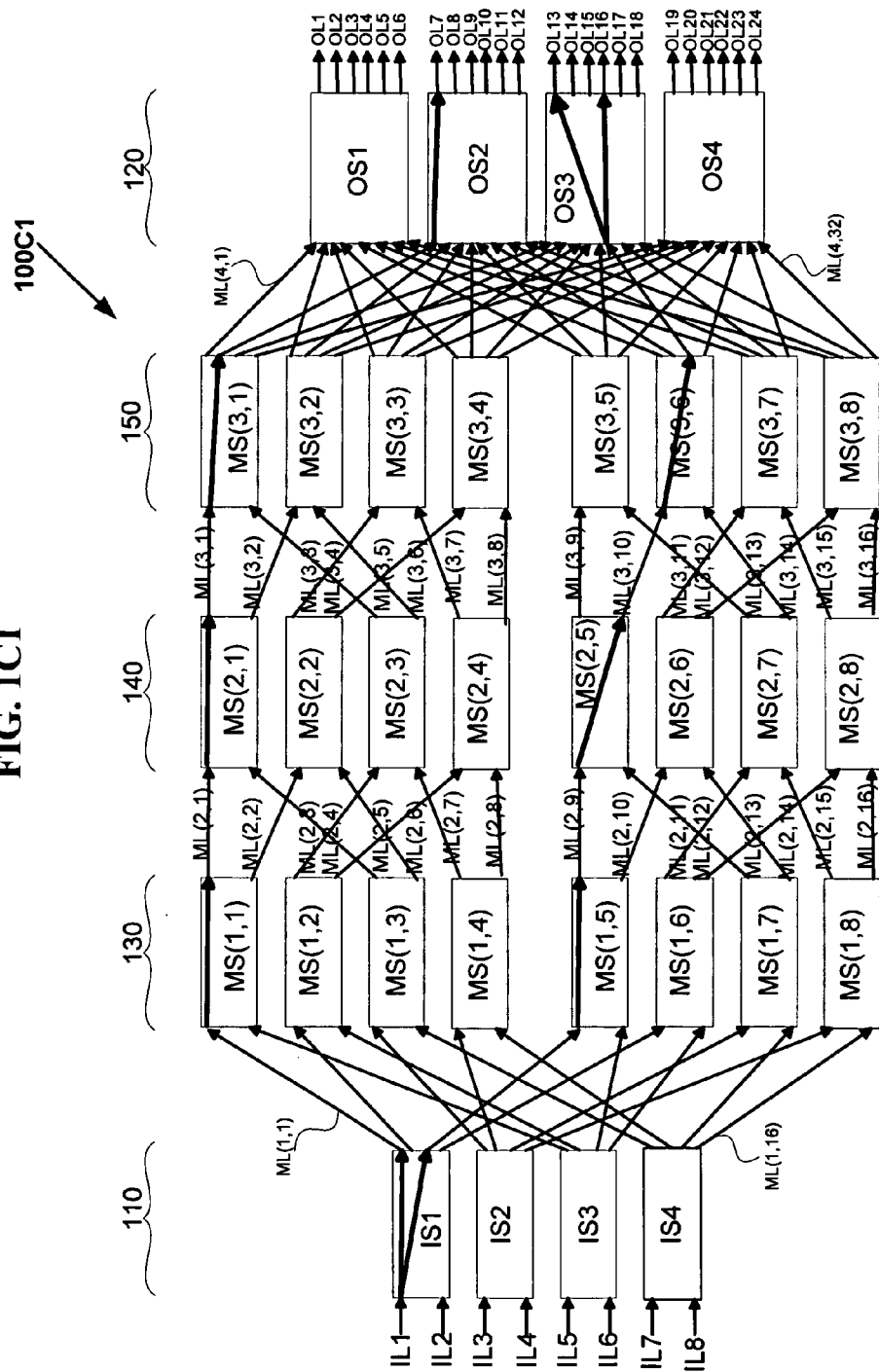
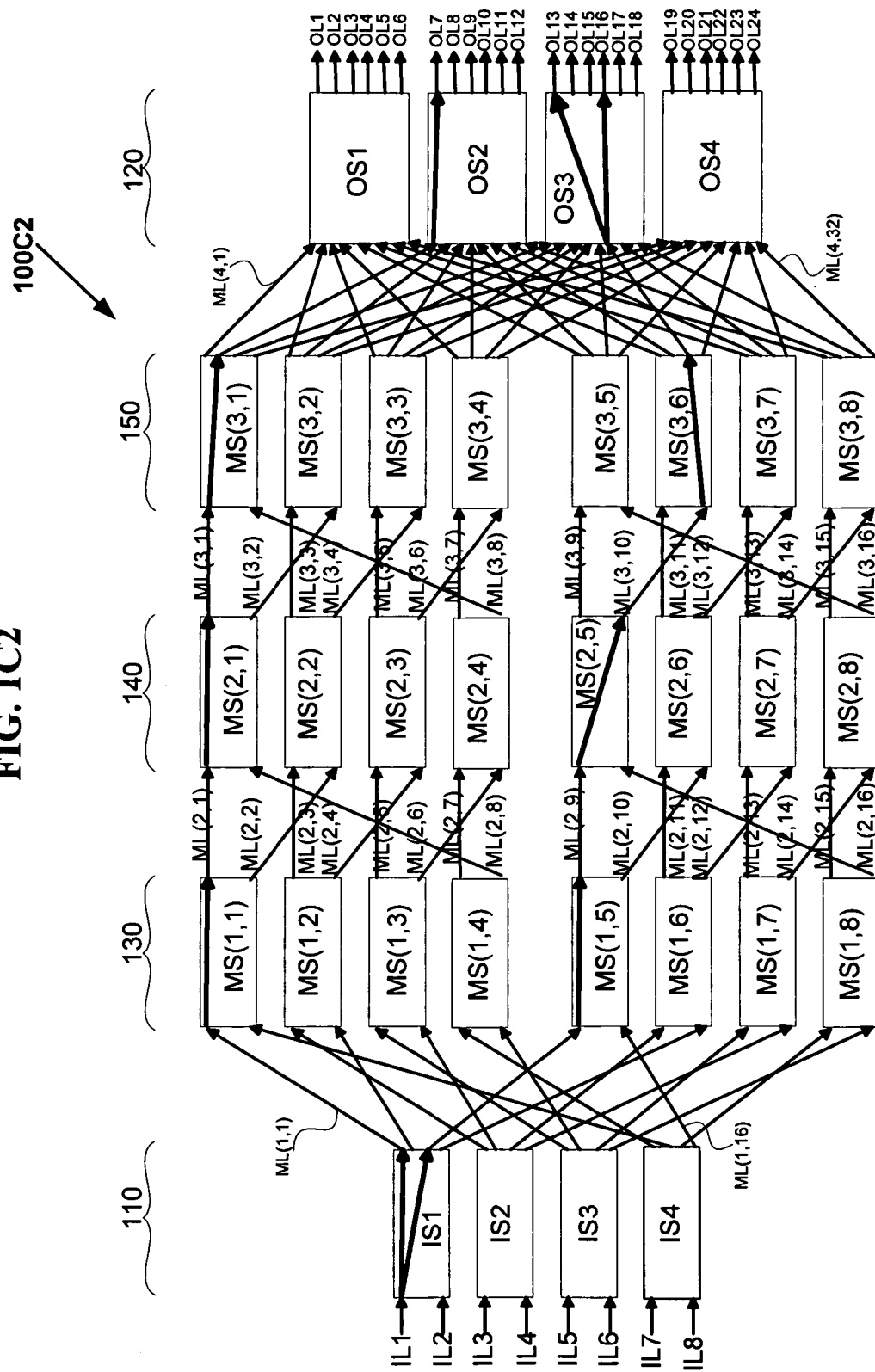


FIG. 1C2



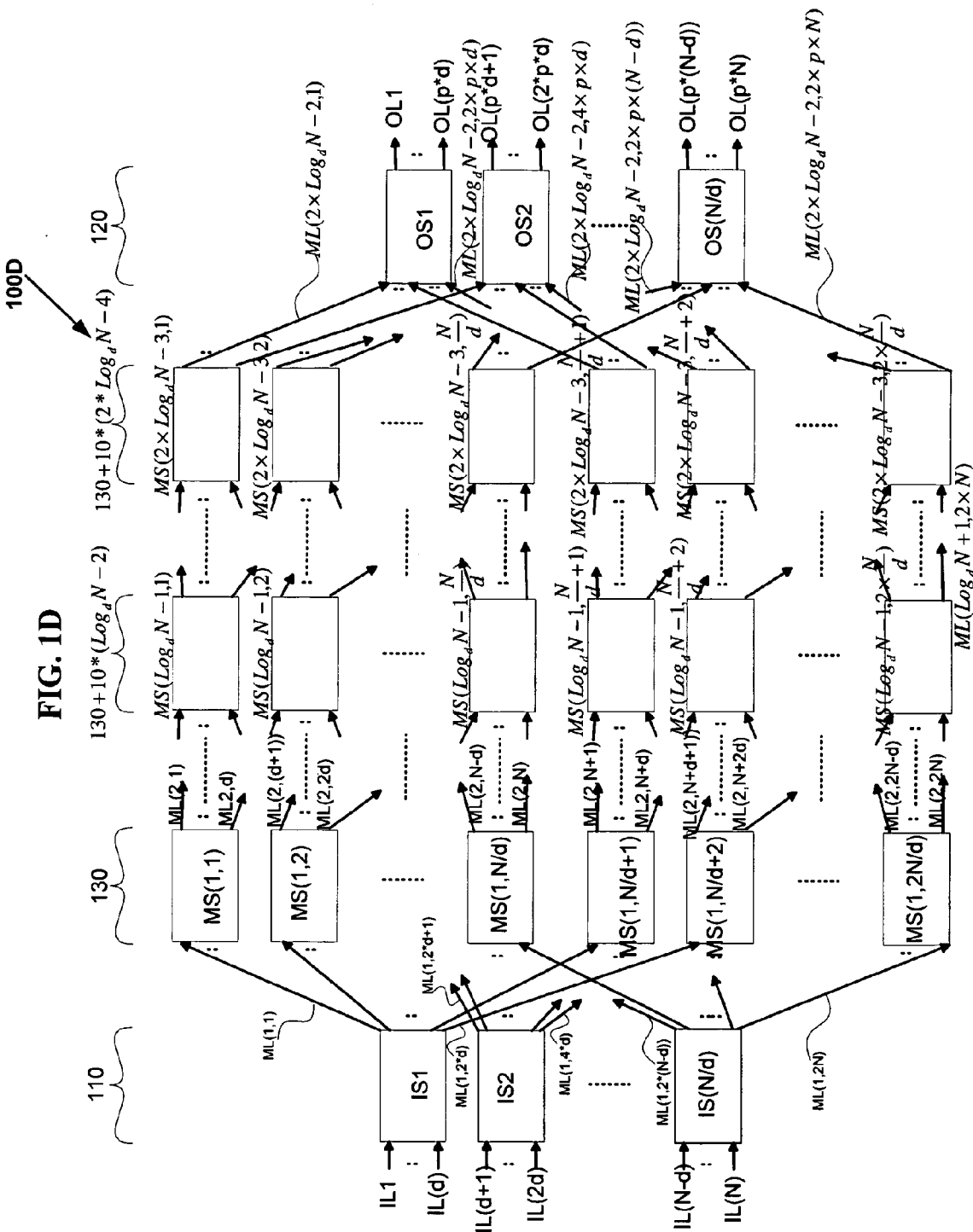


FIG. 1E

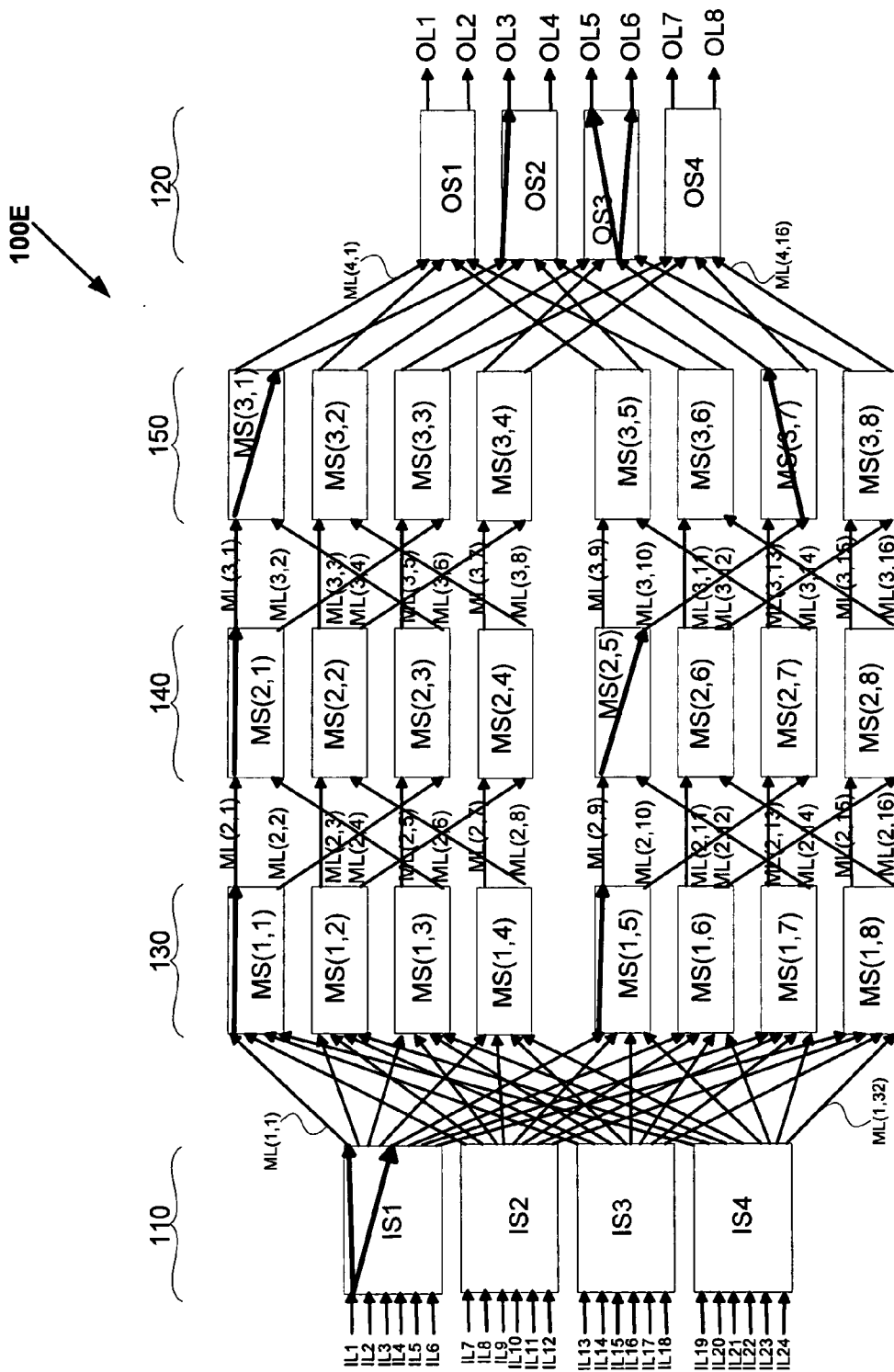


FIG. 1E1

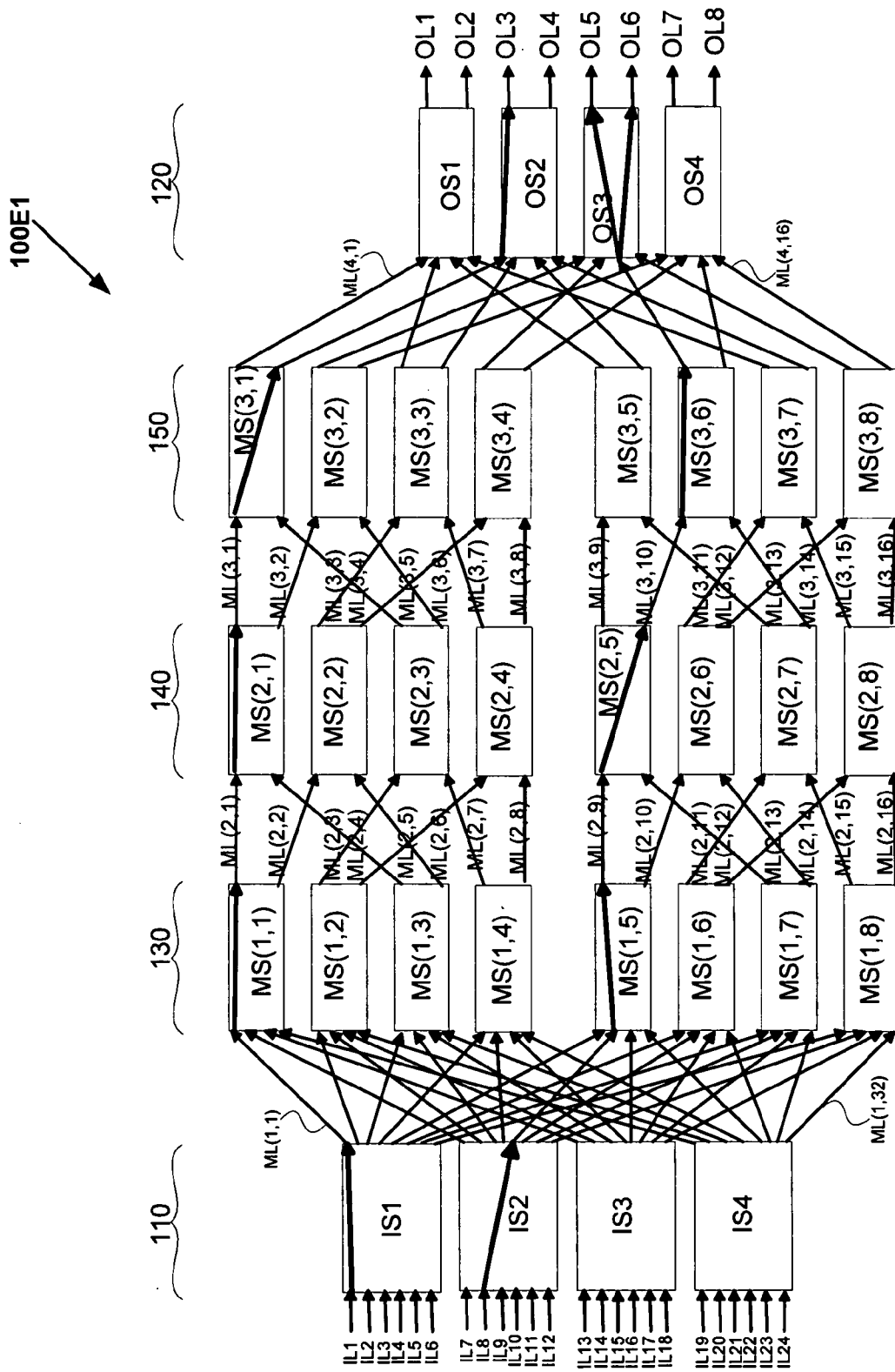
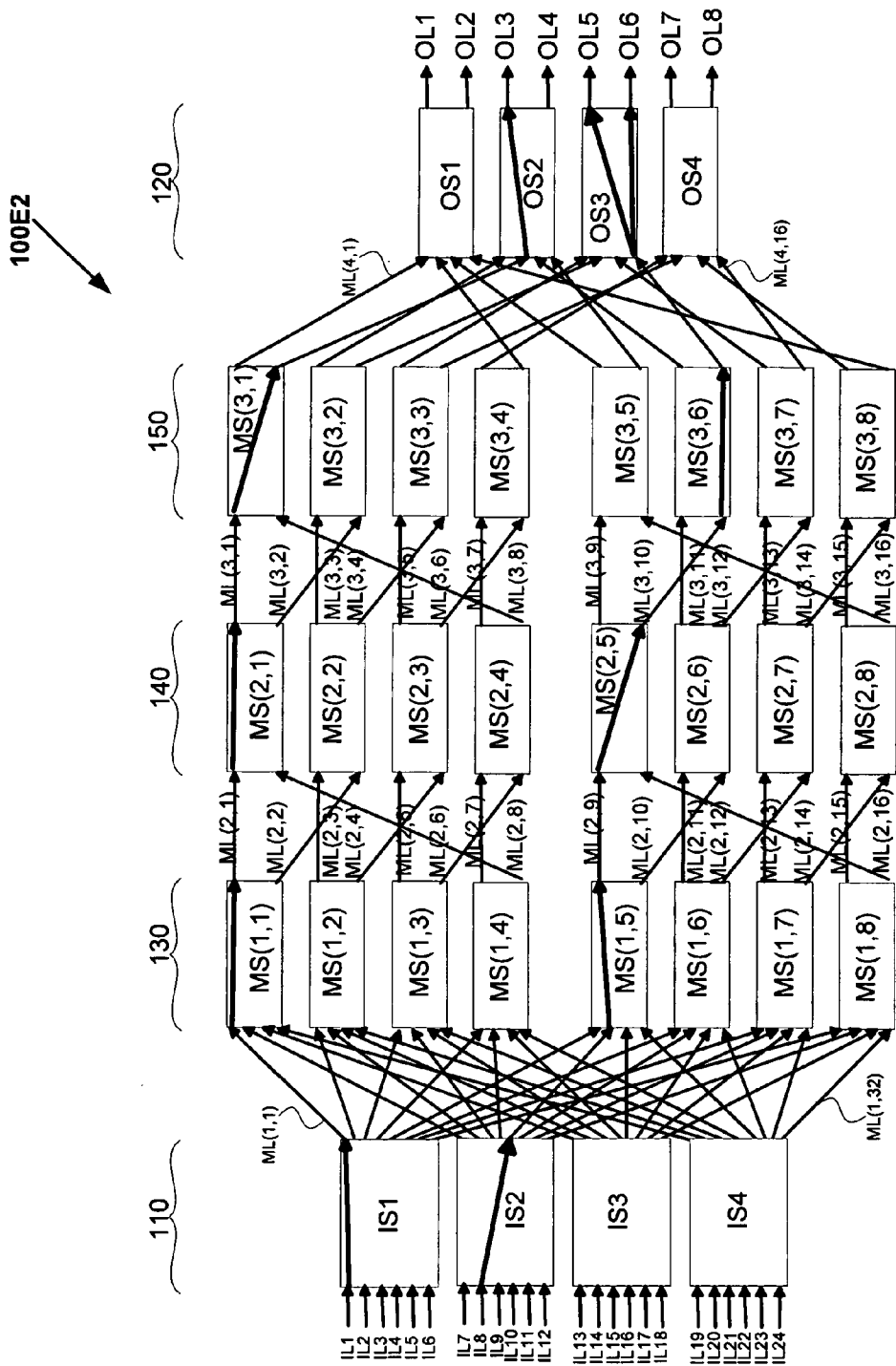


FIG. 1E2



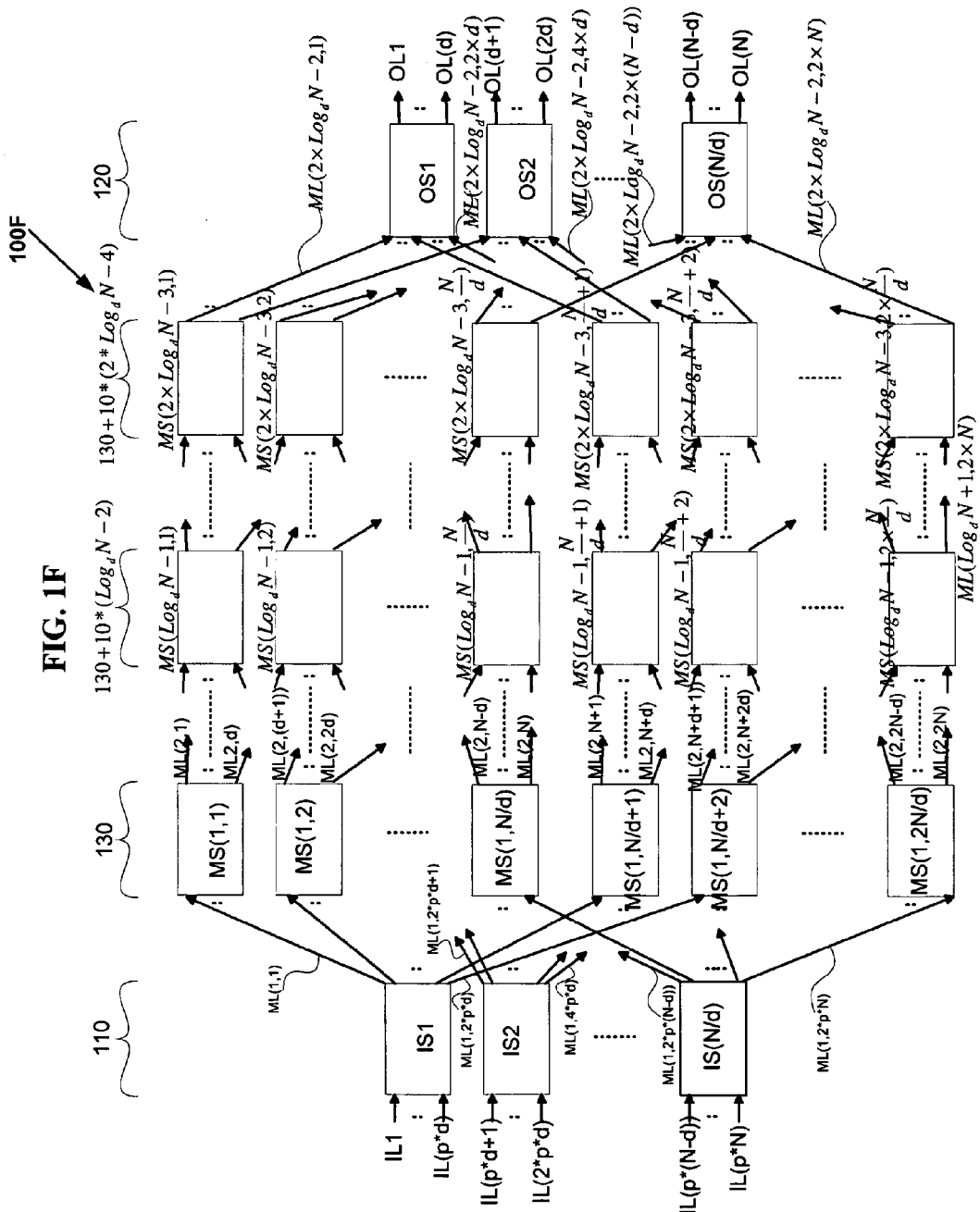


FIG. 1F

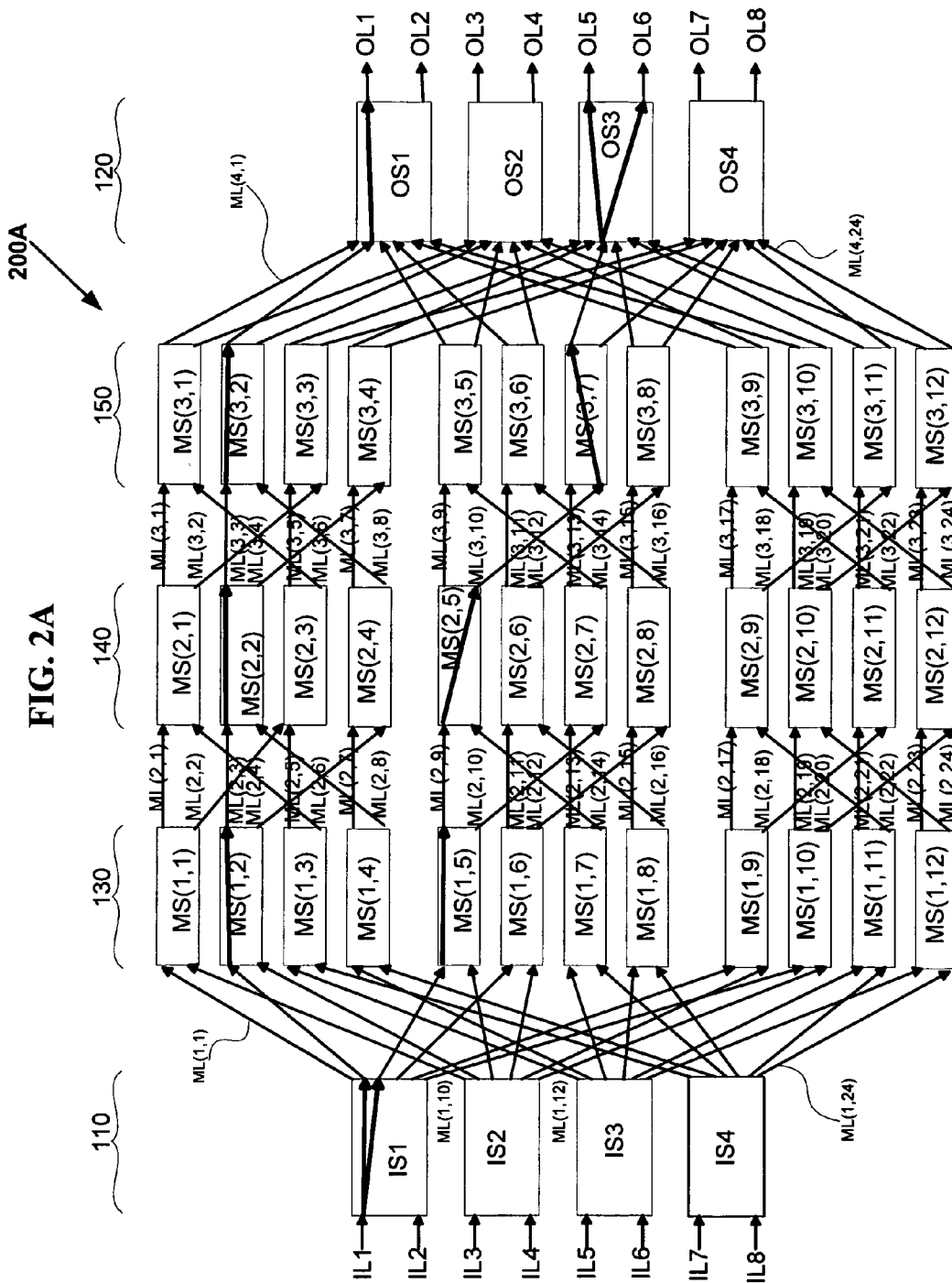


FIG. 2A

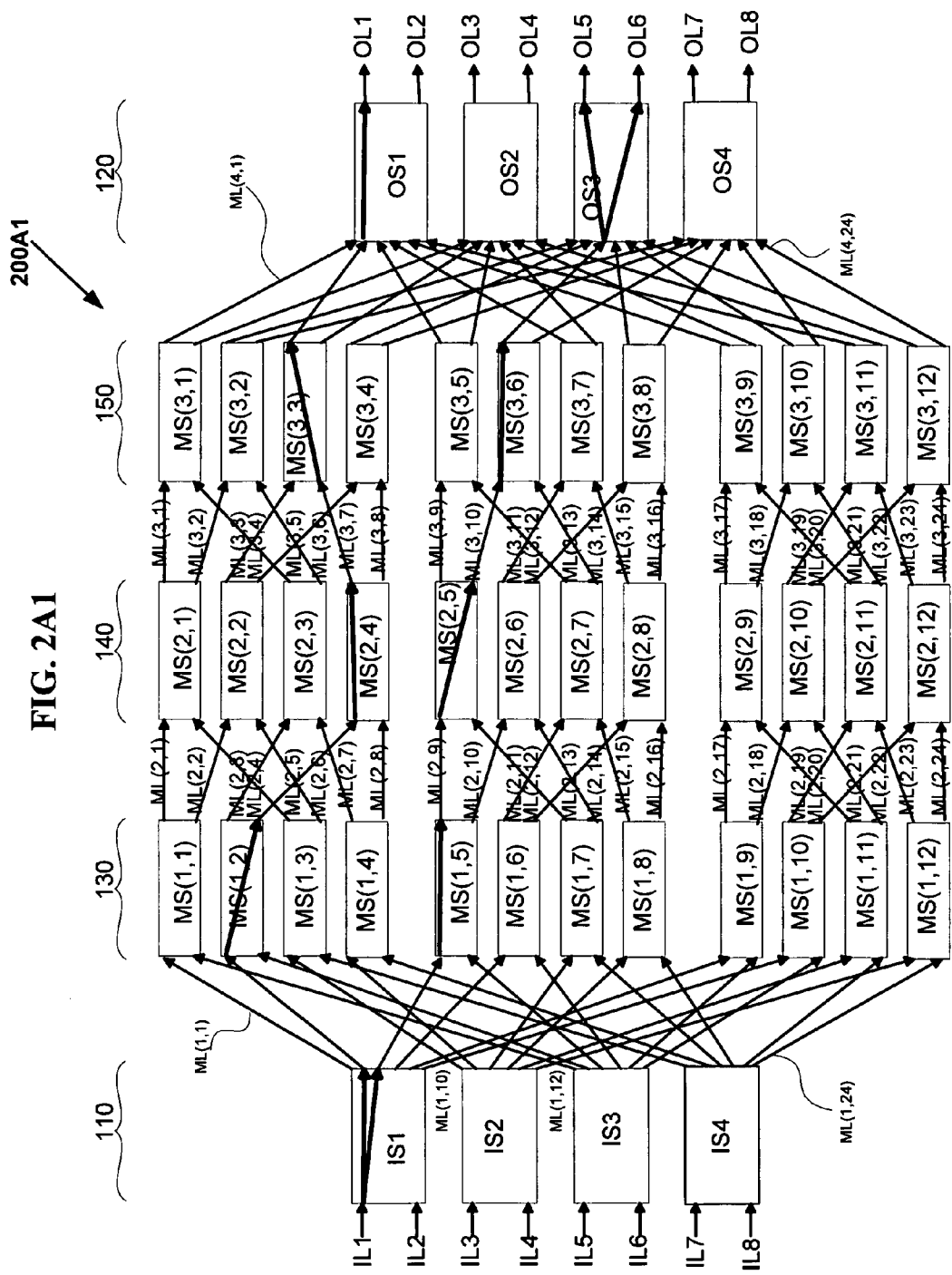


FIG. 2A1

200A1

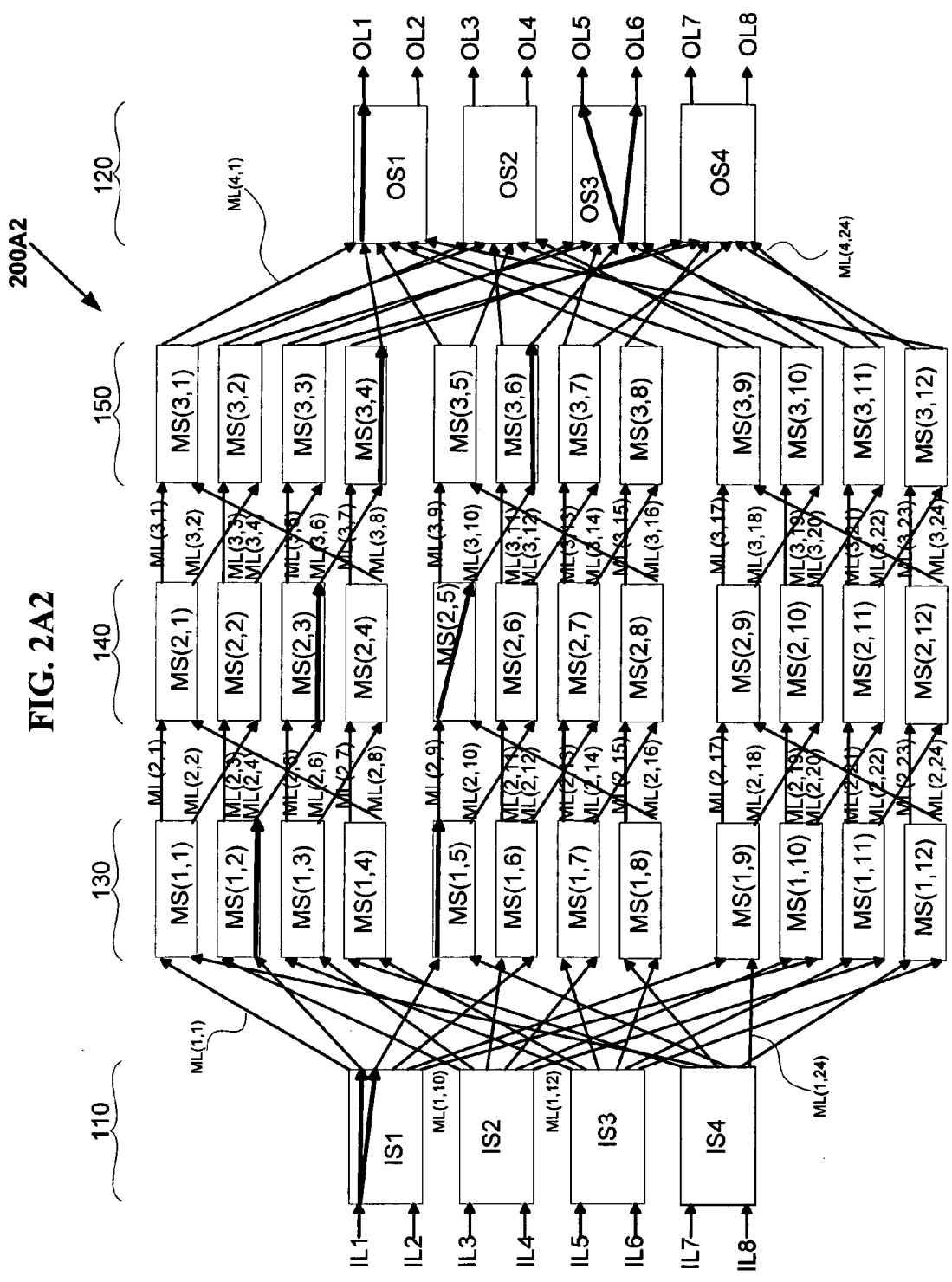


FIG. 2A2

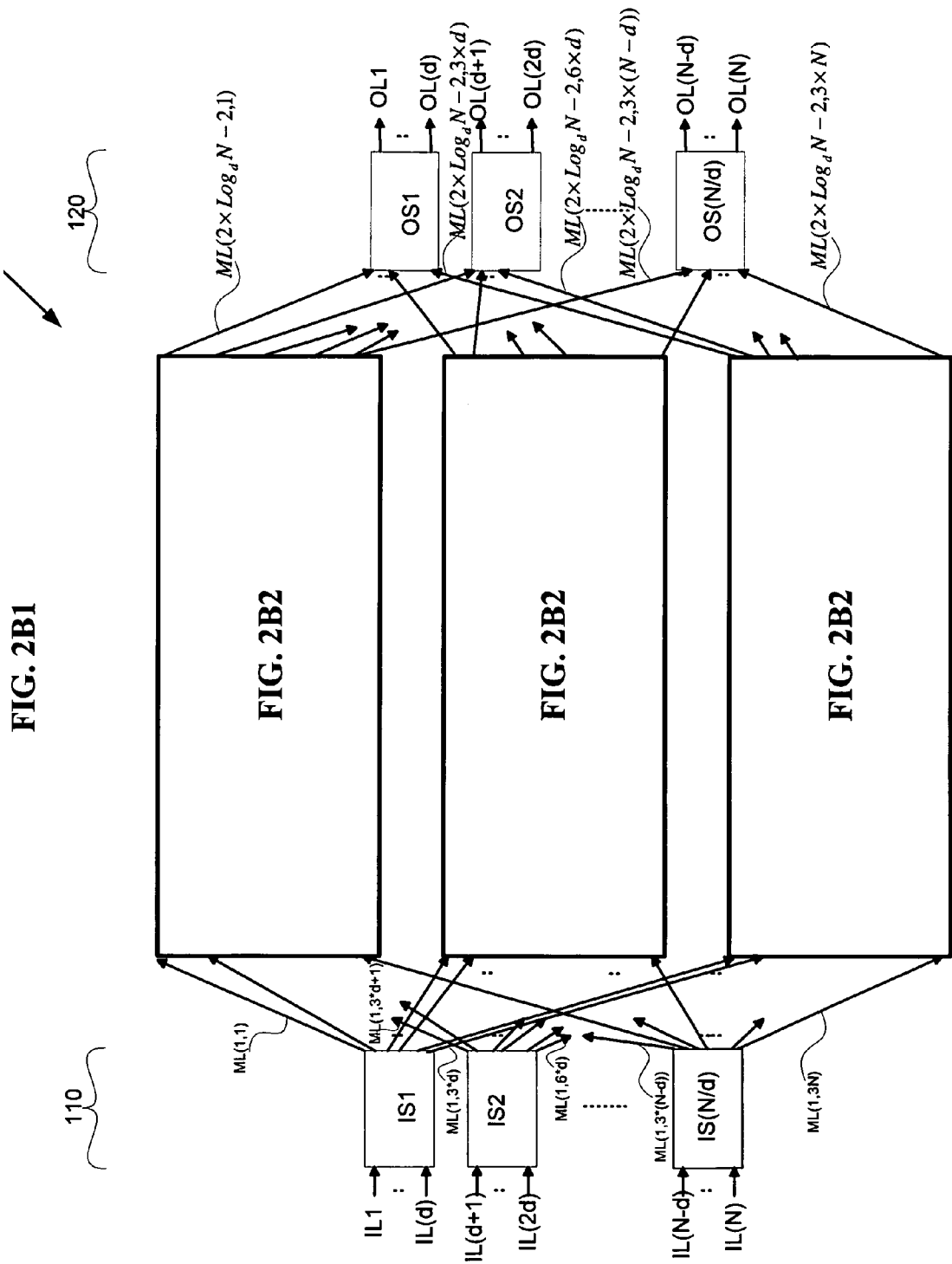
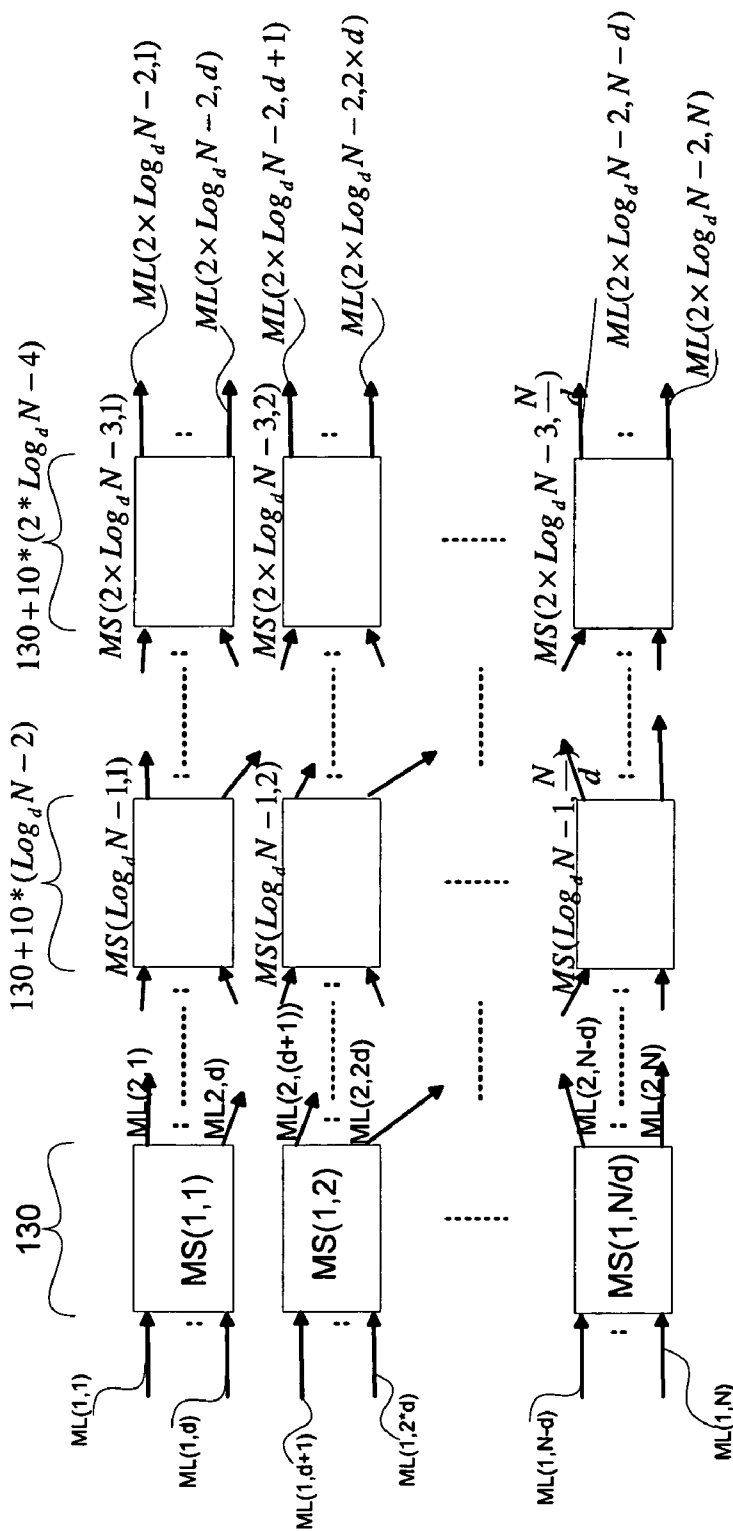
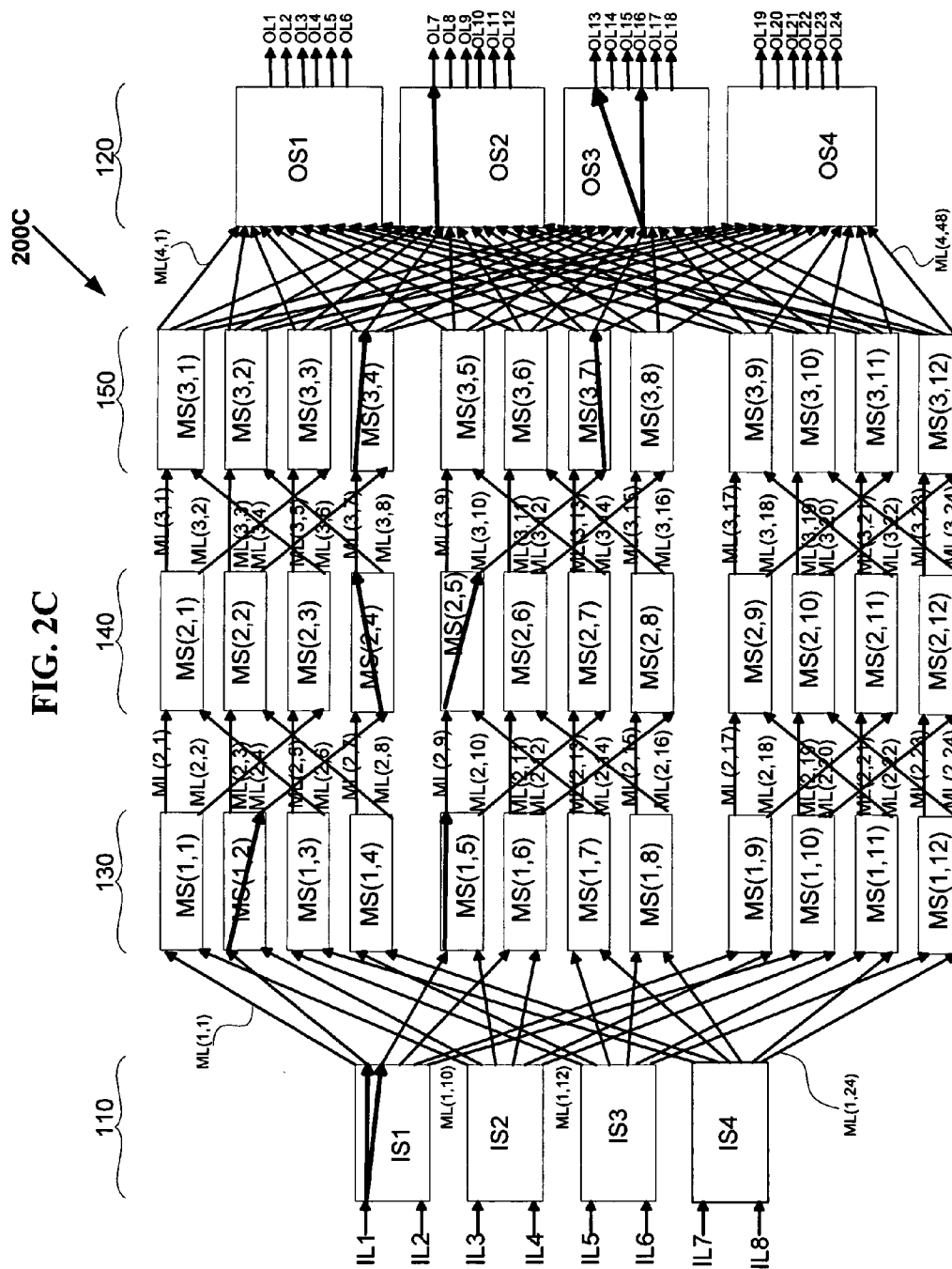


FIG. 2B2

200B2





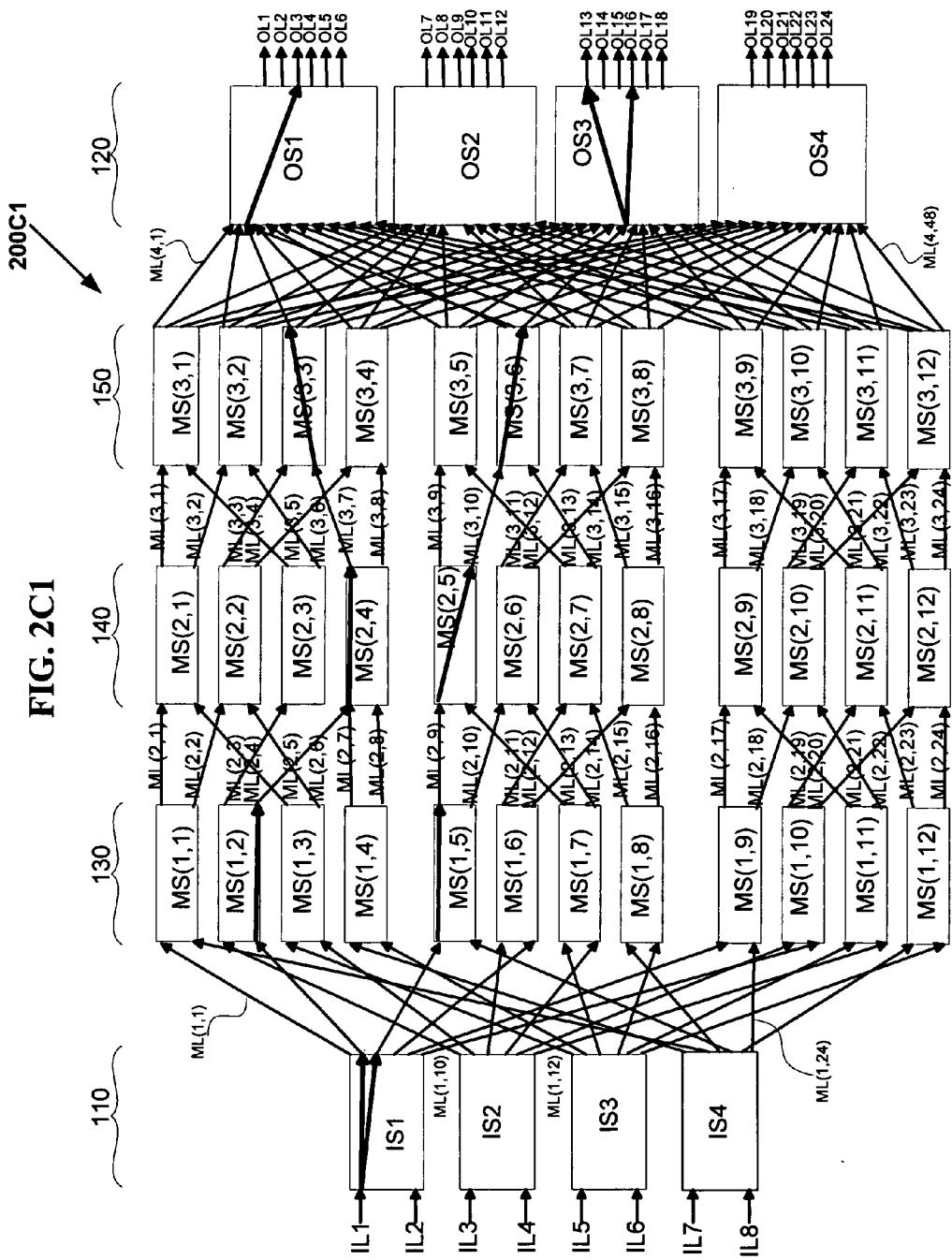


FIG. 2C1

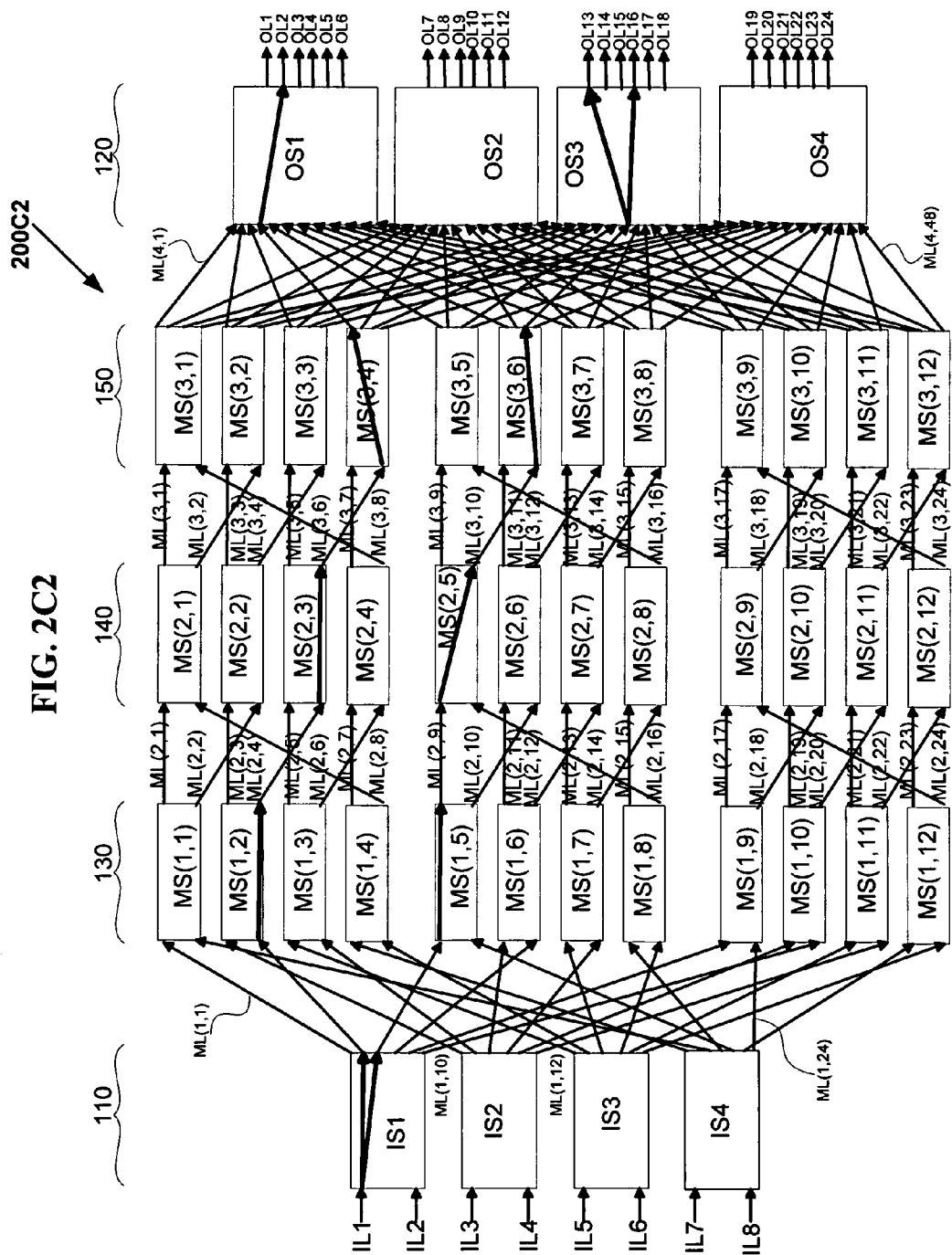


FIG. 2C2

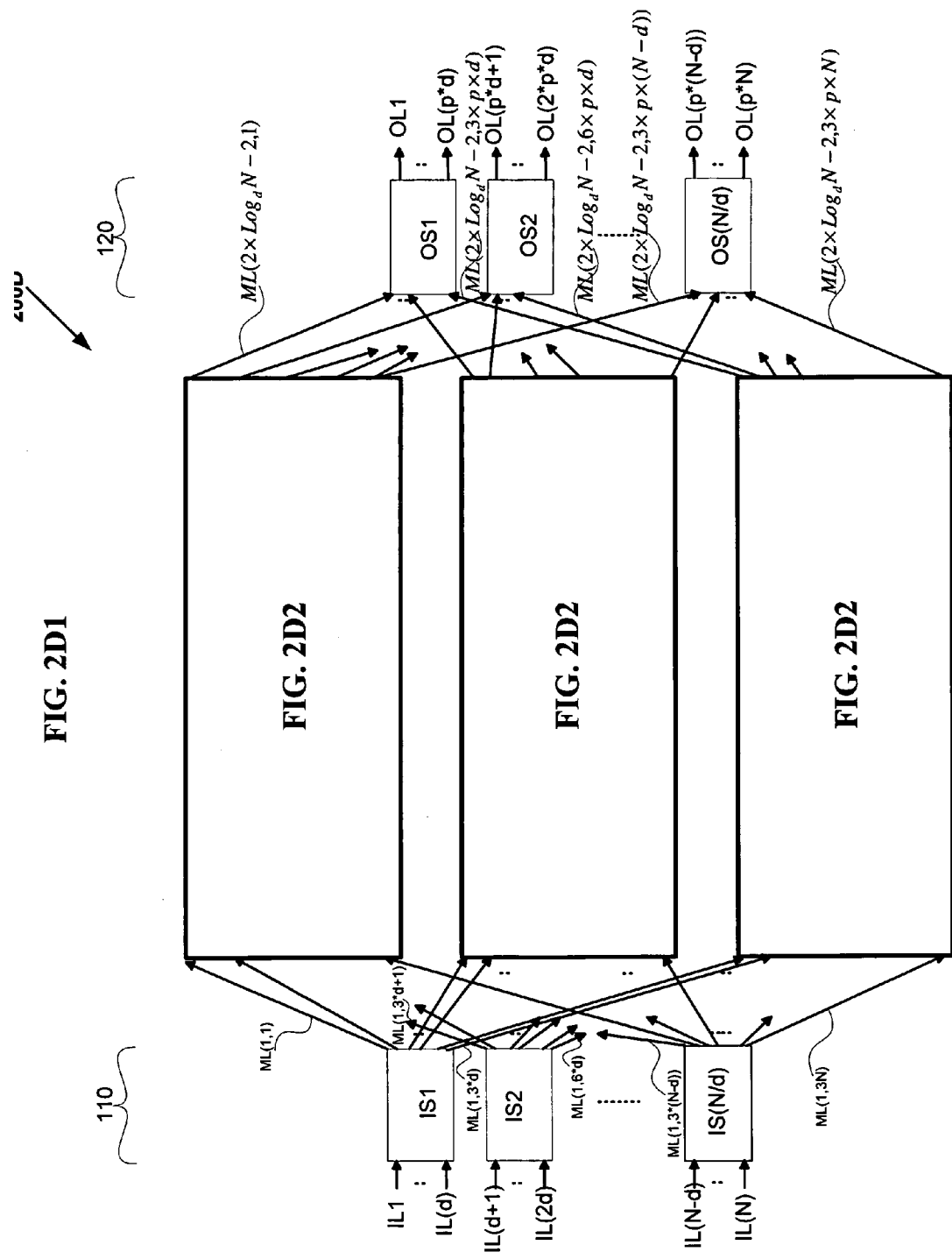
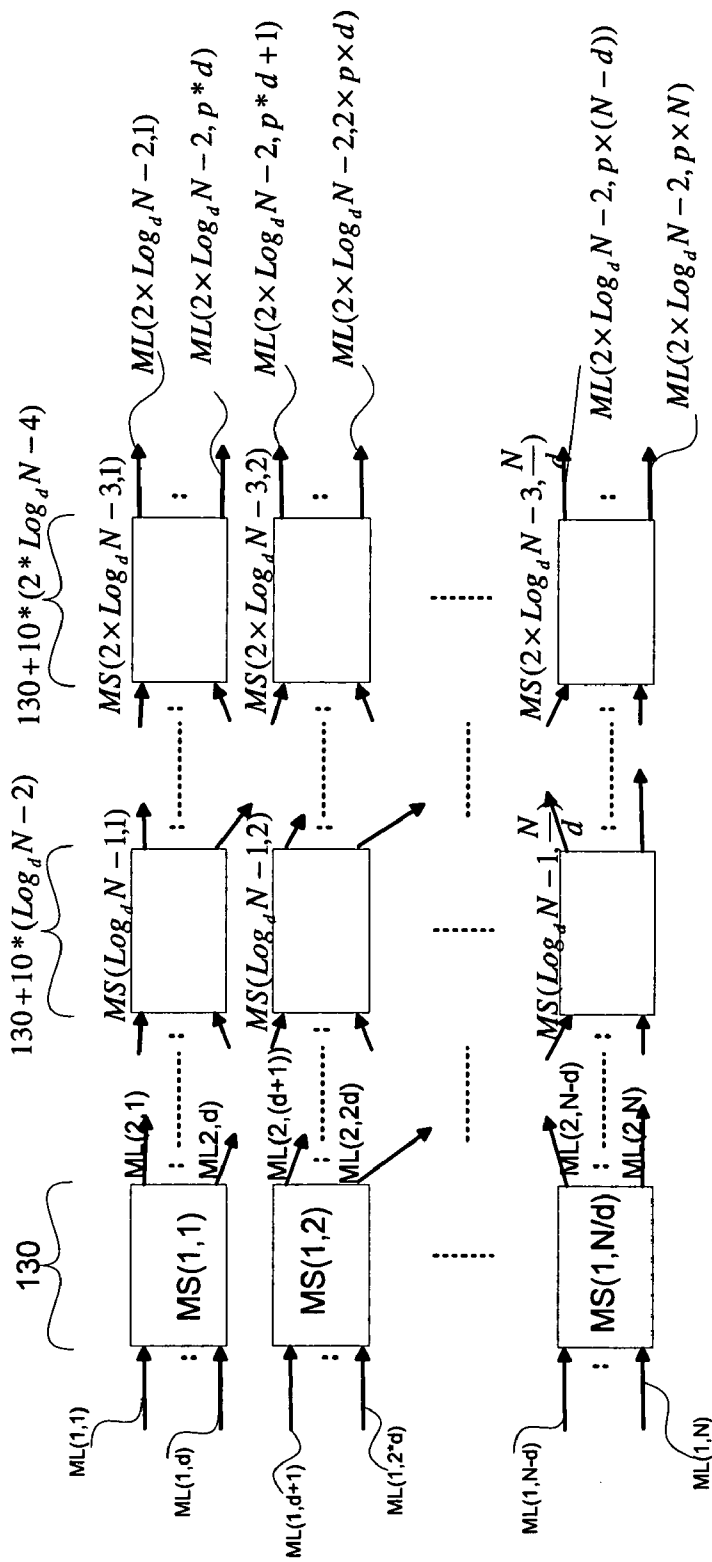


FIG. 2D2

200D2



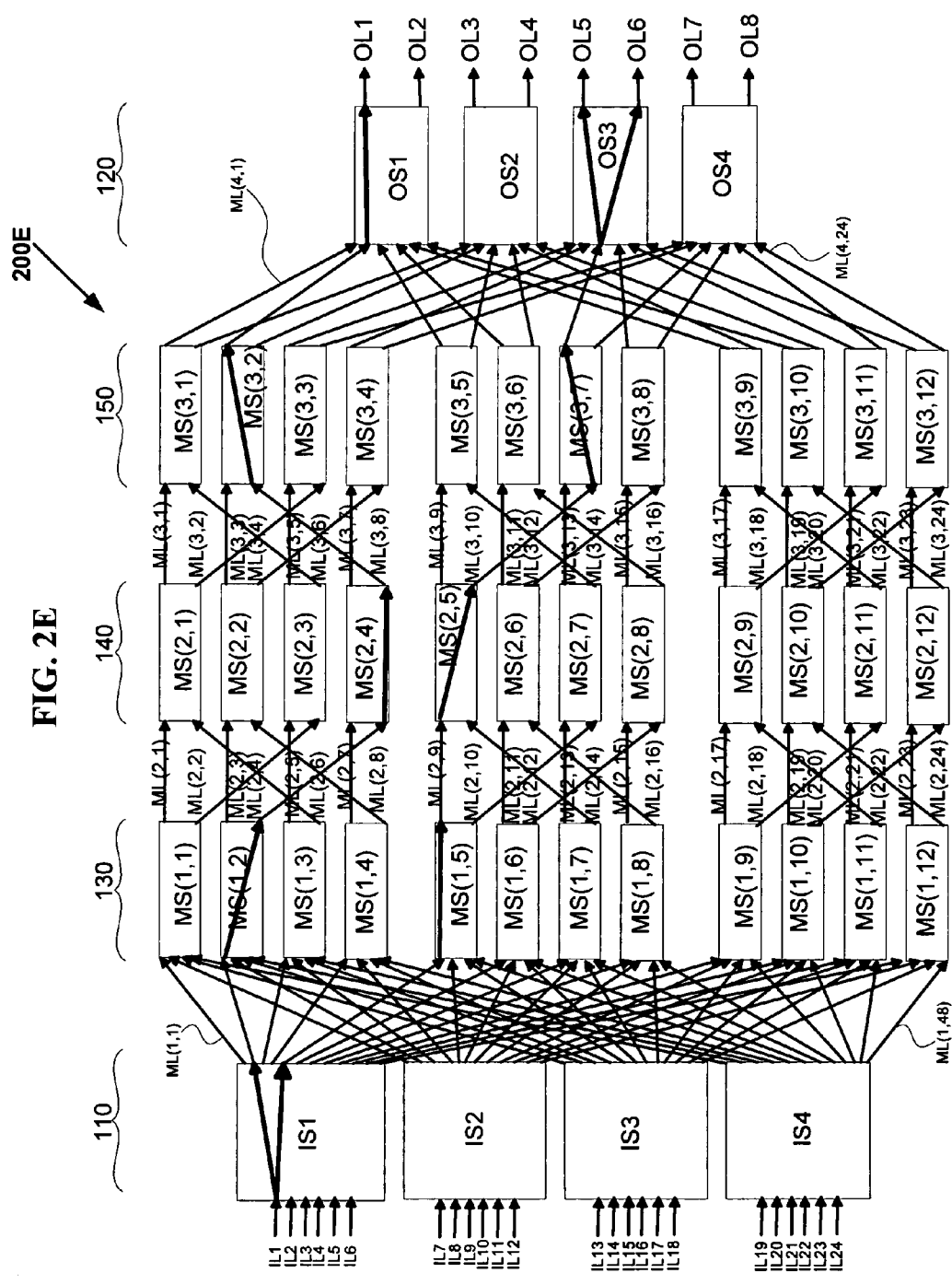


FIG. 2E

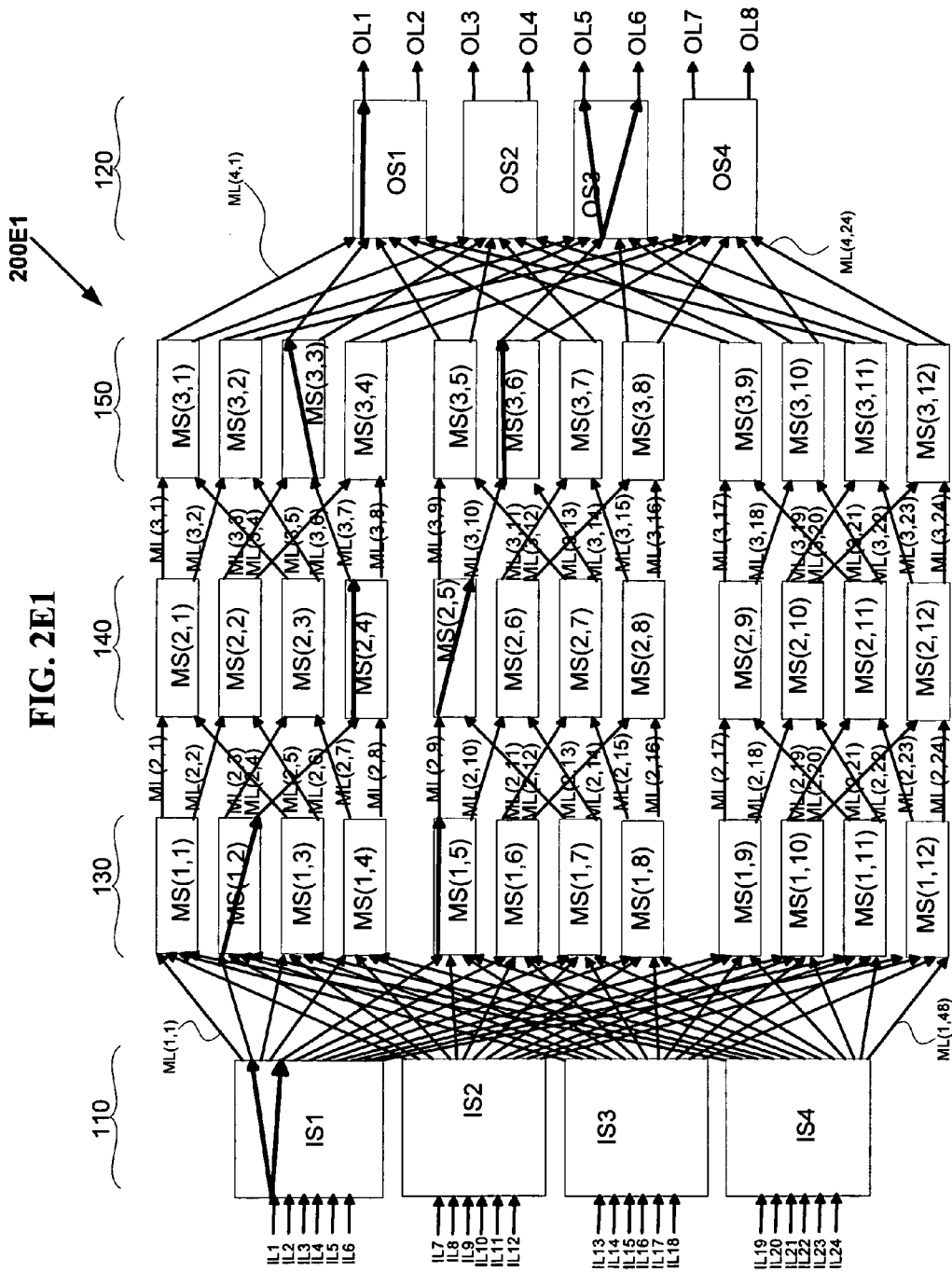


FIG. 2E1

200E1

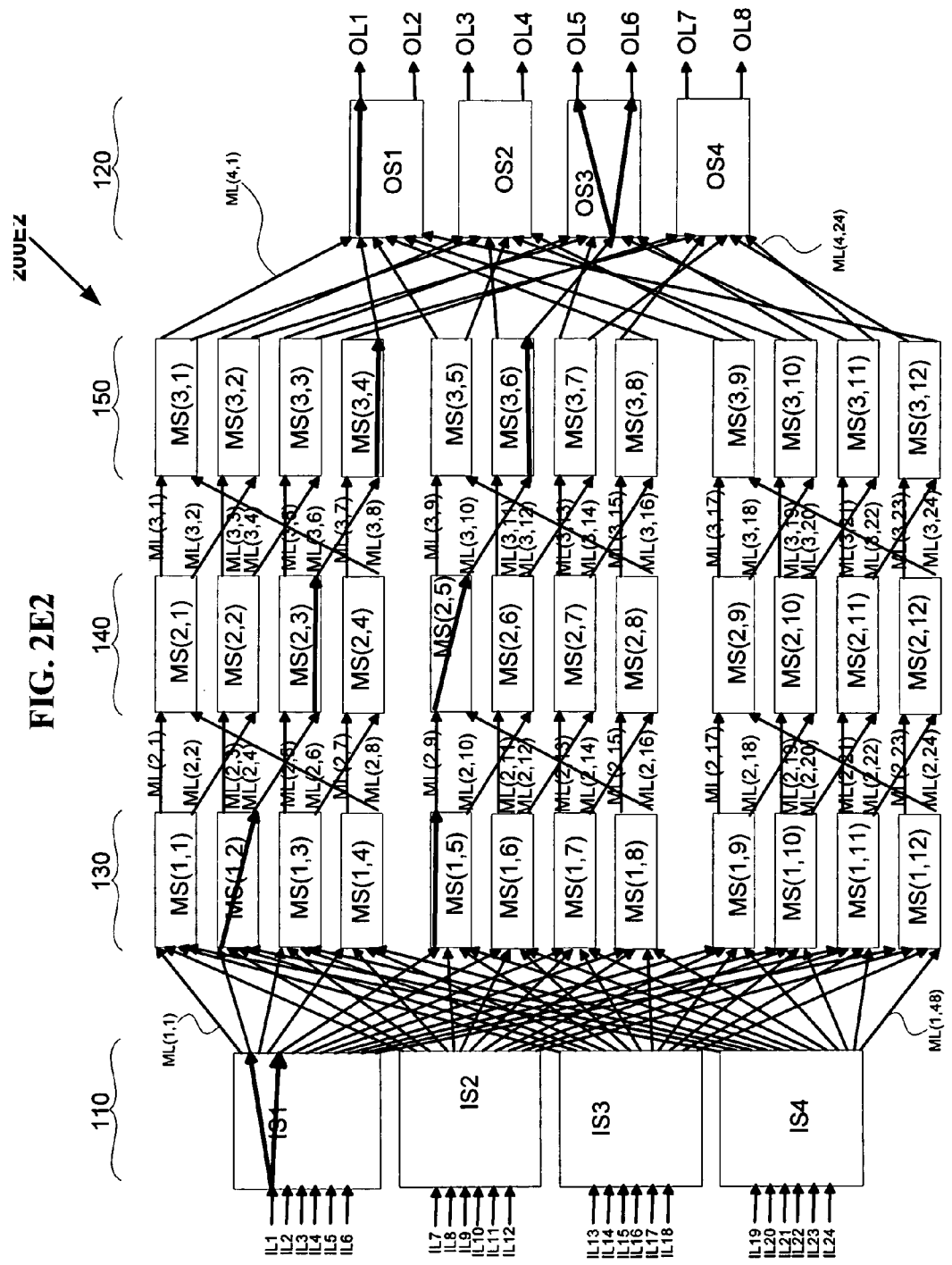


FIG. 2E2

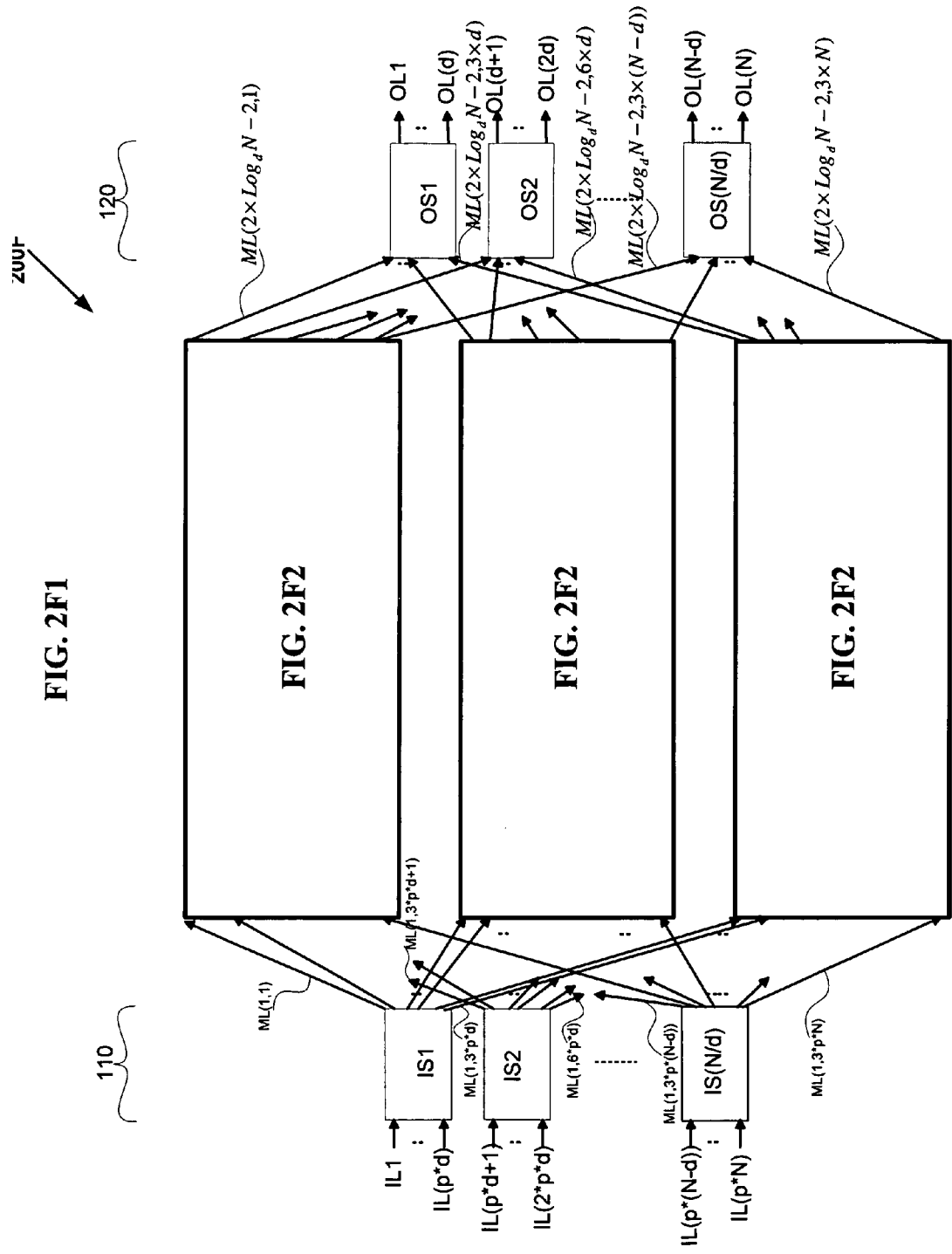
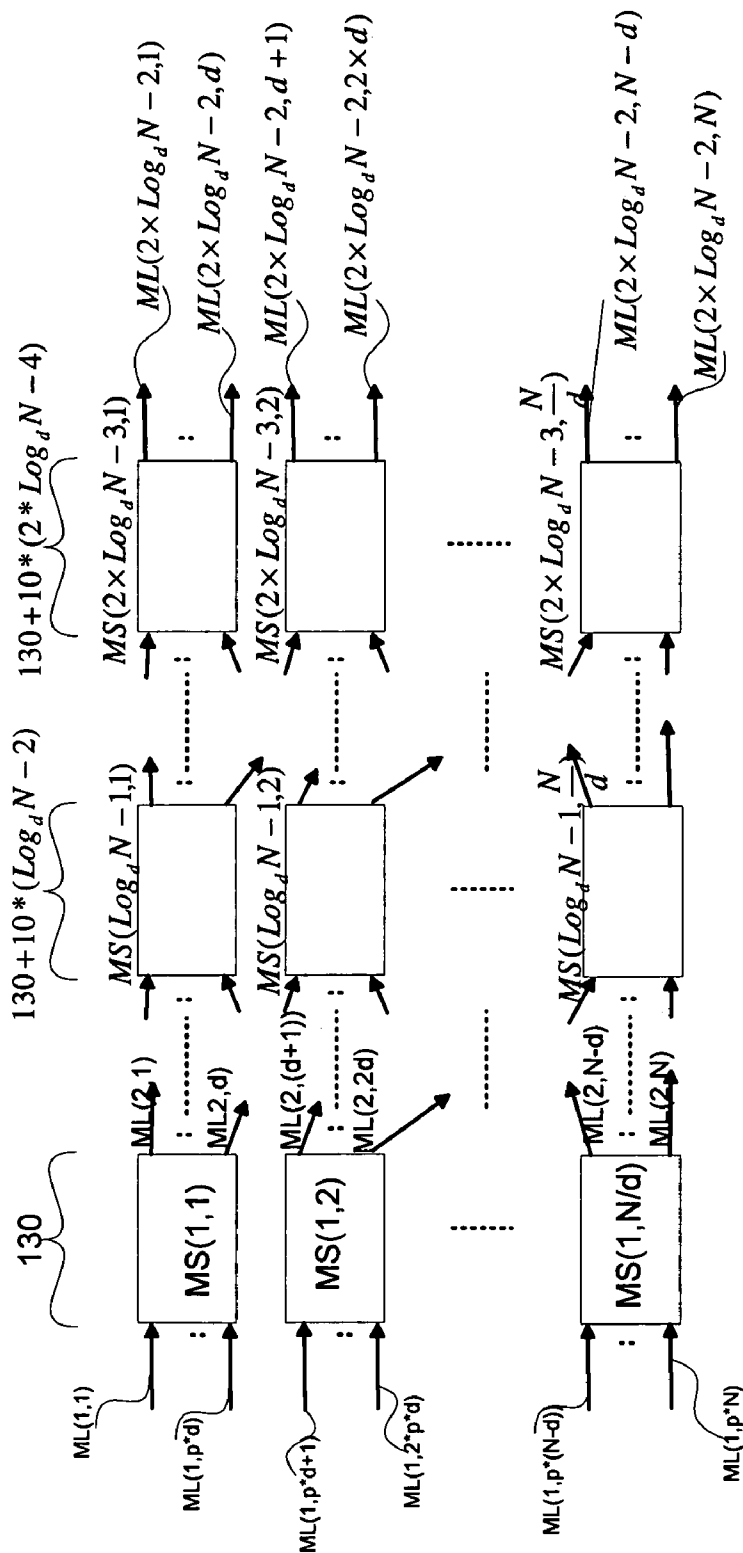


FIG. 2F1

FIG. 2F2

200F2



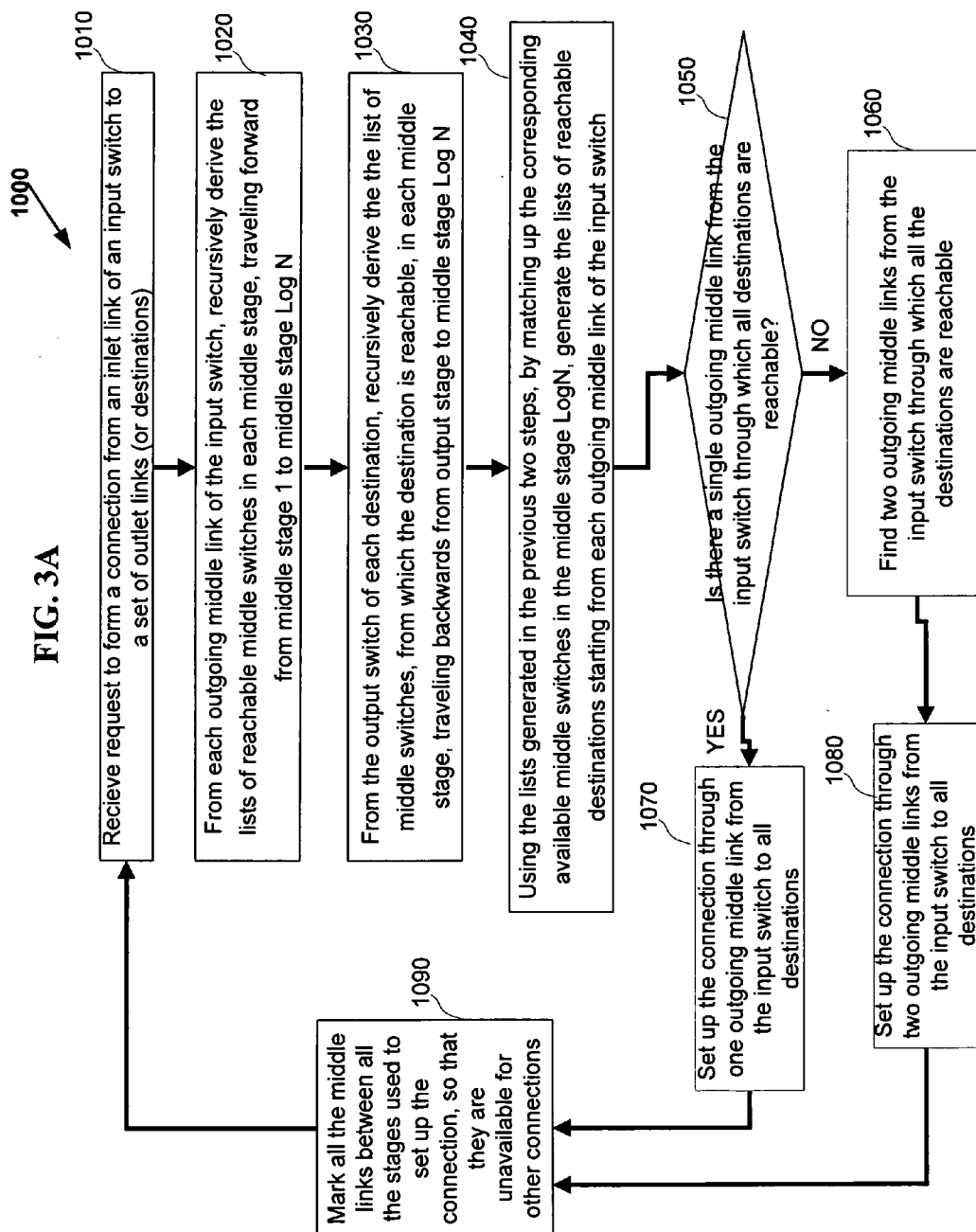


FIG. 4A

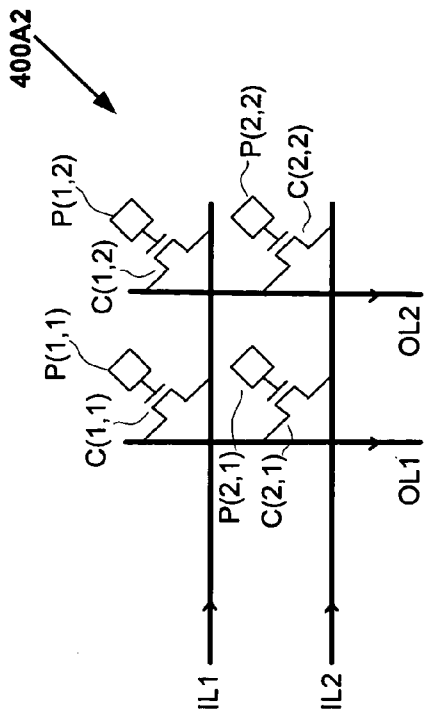


FIG. 4A2
(Prior Art)

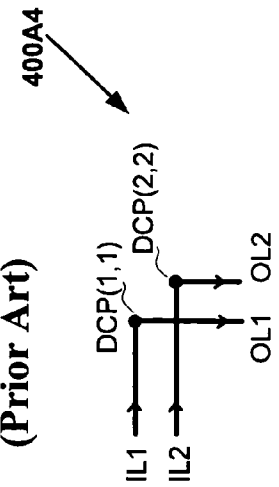


FIG. 4A4

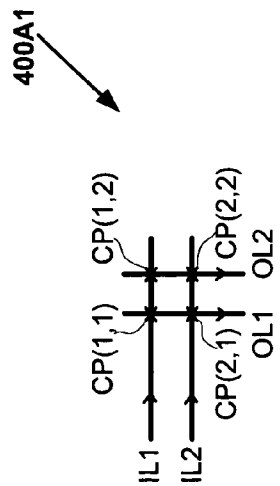


FIG. 4A1
(Prior Art)

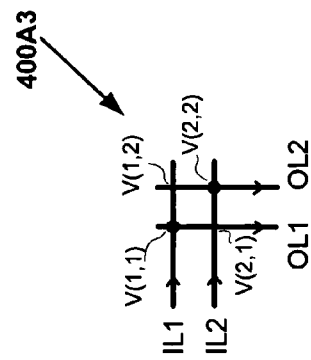


FIG. 4A3
(Prior Art)

US 2010/0135286 A1

Jun. 3, 2010

1

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

[0001] This application is related to and claims priority of PCT Application Serial No. PCT/U.S.08/56064 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.

[0002] This application is related to and incorporates by reference in its entirety the PCT Application Serial No. PCT/U.S.08/64603 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.

[0003] This application is related to and incorporates by reference in its entirety the PCT Application Serial No. PCT/U.S.08/64604 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.

[0004] This application is related to and incorporates by reference in its entirety the PCT Application Serial No. PCT/U.S.08/64605 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 PCT Application Serial No. PCT/U.S.08/82171 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS AND PYRAMID NETWORKS WITH LOCALITY EXPLOITATION" by Venkat Konda assigned to the same assignee as the current application, filed Nov. 2, 2008, the U.S. Provisional Patent Application Ser. 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 Nov. 2, 2007 and the U.S. Provisional Patent Application Ser. No. 61/018,494 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed Jan. 1, 2008.

BACKGROUND OF INVENTION

[0006] 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.

[0007] 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.

[0008] 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. 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] 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

US 2010/0135286 A1

Jun. 3, 2010

2

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.

[0010] 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 $m \leq 2 \times n$. 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 $m \geq 3 \times n - 1$.

[0011] 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 $N=256$.

[0012] 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

[0013] A 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

$$\frac{2 \times N}{d}$$

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, and d outgoing links connecting to the switches in its immediate succeeding stage. Also the same 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.

[0014] A 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

$$\frac{3 \times N}{d}$$

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, and d outgoing links connecting to the switches in its immediate succeeding stage.

BRIEF DESCRIPTION OF DRAWINGS

[0015] FIG. 1A is a diagram **100A** 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 $s=2$ 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.

[0016] FIG. 1B is a diagram **100B** of a general symmetrical multi-stage network $V(N, d, 2)$ 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.

[0017] FIG. 1C is a diagram **100C** of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ having inverse Benes connection topology of five stages with $N_1=8$, $N_2=p \times N_1=24$ 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.

[0018] FIG. 1D is a diagram **100D** of a general asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ with $N_2=p \times N_1$ 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.

[0019] FIG. 1E is a diagram **100E** of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ having inverse Benes connection topology of five stages with $N_2=8$, $N_1=p \times N_2=24$, where $p=3$, and $d=2$ with exemplary multicast

US 2010/0135286 A1

Jun. 3, 2010

3

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

[0020] FIG. 1F is a diagram 100F of a general asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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.

[0021] FIG. 1A1 is a diagram 100A1 of an exemplary symmetrical multi-stage network $V(N, d, 2)$ 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.

[0022] FIG. 1C1 is a diagram 100C1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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.

[0023] FIG. 1E1 is a diagram 100E1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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 connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0024] FIG. 1A2 is a diagram 100A2 of an exemplary symmetrical multi-stage network $V(N, d, 2)$ 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.

[0025] FIG. 1C2 is a diagram 100C2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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.

[0026] FIG. 1E2 is a diagram 100E2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 2)$ 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.

[0027] FIG. 2A is a diagram 200A of an exemplary symmetrical multi-stage network $V(N, d, 3)$ having inverse Benes connection topology of five stages with $N=8$, $d=2$ and $s=3$ with exemplary multicast connections strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0028] FIG. 2B1 & FIG. 2B2 is a diagram 200B of a general symmetrical multi-stage network $V(N, d, 3)$ with $(2 * \log_d$

$N) - 1$ stages strictly nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0029] FIG. 2C is a diagram 200C of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

[0030] FIG. 2D1 & FIG. 2D2 is a diagram 200D of a general asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ with $N_2=p*N_1$ and with $(2 * \log_d N) - 1$ stages strictly nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0031] FIG. 2E is a diagram 200E of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

[0032] FIG. 2F1 & FIG. 2F2 is a diagram 200F of a general asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ with $N_1=p*N_2$ and with $(2 * \log_d N) - 1$ stages strictly nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0033] FIG. 2A1 is a diagram 200A1 of an exemplary symmetrical multi-stage network $V(N, d, 3)$ having Omega connection topology of five stages with $N=8$, $d=2$ and $s=3$ with exemplary multicast connections, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0034] FIG. 2C1 is a diagram 200C1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

[0035] FIG. 2E1 is a diagram 200E1 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

[0036] FIG. 2A2 is a diagram 200A2 of an exemplary symmetrical multi-stage network $V(N, d, 3)$ having nearest neighbor connection topology of five stages with $N=8$, $d=2$ and $s=3$ with exemplary multicast connections, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

[0037] FIG. 2C2 is a diagram 200C2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

[0038] FIG. 2E2 is a diagram 200E2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, 3)$ 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 arbitrary fan-out multicast connections, in accordance with the invention.

[0039] FIG. 3A 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.

US 2010/0135286 A1

Jun. 3, 2010

4

[0040] FIG. 4A1 is a diagram 400A1 of an exemplary prior art implementation of a two by two switch; FIG. 4A2 is a diagram 400A2 for programmable integrated circuit prior art implementation of the diagram 400A1 of FIG. 4A1; FIG. 4A3 is a diagram 400A3 for one-time programmable integrated circuit prior art implementation of the diagram 400A1 of FIG. 4A1; FIG. 4A4 is a diagram 400A4 for integrated circuit placement and route implementation of the diagram 400A1 of FIG. 4A1.

DETAILED DESCRIPTION OF THE INVENTION

[0041] The present invention is concerned with the design and operation of large scale crosspoint reduction using arbitrarily large multi-stage switching networks for broadcast, unicast and multicast connections including their generalized topologies. Particularly 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.

[0042] When a transmitting device simultaneously sends information to more than one receiving device, the one-to-many connection required between the transmitting device 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.

[0043] 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.

[0044] In certain 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-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.

[0045] In certain 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-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.

[0046] Nonblocking configurations for other types of networks with numerous connection topologies and scheduling methods are disclosed as follows:

[0047] 1) 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 PCT Application Serial No. PCT/U.S.08/64603 that is incorporated by reference above.

[0048] 2) Rearrangeably nonblocking for arbitrary fan-out multicast and unicast, and strictly nonblocking for unicast for generalized multi-link multi-stage networks $V_{mink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the PCT Application Serial No. PCT/U.S.08/64604 that is incorporated by reference above.

[0049] 3) Strictly and rearrangeably nonblocking for arbitrary fan-out multicast and unicast for generalized multi-link butterfly fat tree networks $V_{mink-bft}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the PCT Application Serial No. PCT/U.S.08/64603 that is incorporated by reference above.

[0050] 4) 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 PCT Application Serial No. PCT/U.S.08/64604 that is incorporated by reference above.

[0051] 5) Strictly nonblocking for arbitrary fan-out multicast and unicast for generalized multi-link multi-stage networks $V_{mink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the PCT Application Serial No. PCT/U.S.08/64604 that is incorporated by reference above.

[0052] 6) VLSI layouts of numerous types of multi-stage networks are described in the PCT Application Serial No. PCT/U.S.08/64605 entitled "VLSI LAYOUTS OF FULLY CONNECTED NETWORKS" that is incorporated by reference above.

[0053] 7) VLSI layouts of numerous types of multi-stage networks with locality exploitation are described in PCT Application Serial No. PCT/U.S.08/82171 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS WITH LOCALITY EXPLOITATION" by Venkat Konda assigned to the same assignee as the current application, filed Nov. 2, 2008.

[0054] 8) VLSI layouts of numerous types of multistage pyramid networks are described in PCT Application Serial No. PCT/U.S.08/82171 entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS WITH LOCALITY EXPLOITATION" by Venkat Konda assigned to the same assignee as the current application, filed Nov. 2, 2008.

Symmetric RNB Embodiments

[0055] Referring to FIG. 1A, in one embodiment, an exemplary symmetrical multi-stage network 100A 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).

US 2010/0135286 A1

Jun. 3, 2010

5

[0056] 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**.

[0057] 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 $2d \times d$. 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 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.

[0058] 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).

[0059] 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).

[0060] 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).

[0061] 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)).

[0062] 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).

[0063] Finally the connection topology of the network **100A** shown in FIG. 1A is known to be back to back inverse Benes connection topology.

[0064] Referring to FIG. 1A1, in another embodiment of network $V(N,d,s)$, an exemplary symmetrical multi-stage network **100A1** 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).

[0065] 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**.

US 2010/0135286 A1

Jun. 3, 2010

6

[0066] 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 $2d \times d$. 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. The symmetric multi-stage network of FIG. 1A1 is also the network of the type $V(N,d,s)$, 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.

[0067] 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).

[0068] 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).

[0069] 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).

[0070] 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)).

[0071] 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(3,3), MS(3,5) and MS(3,7) through the links ML(4,1), ML(4,5), ML(4,9) and ML(4,13) respectively).

[0072] Finally the connection topology of the network **100A1** shown in FIG. 1A1 is known to be back to back Omega connection topology.

[0073] Referring to FIG. 1A2, in another embodiment of network $V(N,d,s)$, an exemplary symmetrical multi-stage network **100A2** 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).

[0074] 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**.

[0075] 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

$$2 \times \frac{N}{d}$$

US 2010/0135286 A1

Jun. 3, 2010

7

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. The symmetric multi-stage network of FIG. 1A2 is also the network of the type $V(N,d,s)$, 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.

[0076] 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).

[0077] 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).

[0078] 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).

[0079] 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)).

[0080] 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(3,4), MS(3,5) and MS(3,8) through the links ML(4,1), ML(4,2), ML(4,3) and ML(4,4) respectively).

[0081] Finally the connection topology of the network 100A2 shown in FIG. 1A2 is hereinafter called nearest neighbor connection topology.

[0082] In the three embodiments of FIG. 1A, FIG. 1A1 and FIG. 1A2 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 $V(N,d,s)$ 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N,d,s)$ can be built. The embodiments of FIG. 1A, FIG. 1A1, and FIG. 1A2 are only three examples of network $V(N,d,s)$.

[0083] In the three embodiments of FIG. 1A, FIG. 1A1 and FIG. 1A2, 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.

[0084] In the example illustrated in FIG. 1A (or in FIG. 1A1, or in FIG. 1A2), 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 100A1, or 100A2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0085] 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

US 2010/0135286 A1

Jun. 3, 2010

8

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

[0086] Network **100B** of FIG. **1B** 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 multicast when $s=2$ according to the current invention. Also the general symmetrical multi-stage network $V(N,d,s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention (And in the example of FIG. **1B**, $s=2$). 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 $2 \times d$ 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 $OS1$).

[0087] 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.

[0088] 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.

[0089] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches

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

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

[0090] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches

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

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

[0091] Each of the N/d output switches $OS1-OS(N/d)$ are connected from exactly $2 \times d$ switches in middle stage $\mathbf{130} + 10^*(2 * \log_d N - 4)$ through $2 \times d$ links.

[0092] As described before, again the connection topology of a general $V(N,d,s)$ may be any one of the connection topologies. For example the connection topology of the network $V(N,d,s)$ 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 $V(N,d,s)$ network is, when no connections are setup in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N,d,s)$ can be built. The embodiments of FIG. **1A**, FIG. **1A1**, and FIG. **1A2** are three examples of network $V(N,d,s)$.

[0093] The general symmetrical multi-stage network $V(N,d,s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general symmetrical multi-stage network $V(N,d,s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention.

[0094] Every switch in the multi-stage networks discussed herein has multicast capability. In a $V(N,d,s)$ 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.

US 2010/0135286 A1

Jun. 3, 2010

9

[0095] To characterize a multicast assignment, for each inlet link

$$i \in \left\{1, 2, \dots, \frac{N_1}{d}\right\},$$

let $I_i = O$, where

$$O \subset \left\{1, 2, \dots, \frac{N_1}{d}\right\},$$

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. 1A shows an exemplary five-stage network, namely $V(8,2,2)$, 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.

[0096] 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 RNB ($N_2 > N_1$) Embodiments

[0097] Referring to FIG. 1C, in one embodiment, an exemplary asymmetrical multi-stage network 100C 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).

[0098] 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.

[0099] 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

$$2 * \frac{N_1}{d}.$$

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

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

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 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.

[0100] Each of the

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

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).

[0101] Each of the

$$2 * \frac{N_1}{d}$$

US 2010/0135286 A1

Jun. 3, 2010

10

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).

[0102] 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).

[0103] 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)).

[0104] Each of the

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

output switches OS1-OS4 are connected from exactly $d+d_2$ switches in middle stage **150** through $d+d_2$ 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).

[0105] Finally the connection topology of the network **100C** shown in FIG. 1C is known to be back to back inverse Benes connection topology.

[0106] Referring to FIG. 1C1, in another embodiment of network $V(N_1, N_2, d, s)$, an exemplary asymmetrical multi-stage network **100C1** 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).

[0107] 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**.

[0108] 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

$$d * \frac{(d + d_2)}{2}.$$

US 2010/0135286 A1

Jun. 3, 2010

11

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 multi-stage network of FIG. 1C1 is also the network of the type $V(N_1, N_2, d, s)$, 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.

[0109] 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).

[0110] 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).

[0111] 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).

[0112] 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)).

[0113] Each of the

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

output switches OS1-OS4 are connected from exactly $d + d_2$ switches in middle stage 150 through $d + d_2$ 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).

[0114] Finally the connection topology of the network 100C1 shown in FIG. 1C1 is known to be back to back Omega connection topology.

[0115] Referring to FIG. 1C2, in another embodiment of network $V(N_1, N_2, d, s)$, an exemplary asymmetrical multi-stage network 100C2 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).

[0116] 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.

[0117] 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 2010/0135286 A1

Jun. 3, 2010

12

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

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. The asymmetric multi-stage network of FIG. 1C2 is also the network of the type $V(N_1, N_2, d, s)$, 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.

[0118] 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).

[0119] 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).

[0120] 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).

[0121] 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)).

[0122] Each of the

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

output switches OS1-OS4 are connected from exactly $d+d_1$ switches in middle stage 150 through $d+d_2$ 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).

[0123] Finally the connection topology of the network 100C2 shown in FIG. 1C2 is hereinafter called nearest neighbor connection topology.

US 2010/0135286 A1

Jun. 3, 2010

13

[0124] In the three embodiments of FIG. 1C, FIG. 1C1 and FIG. 1C2 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 $V(N_1, N_2, d, s)$ 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1C, FIG. 1C1, and FIG. 1C2 are only three examples of network $V(N_1, N_2, d, s)$.

[0125] In the three embodiments of FIG. 1C, FIG. 1C1 and FIG. 1C2, 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.

[0126] In the example illustrated in FIG. 1C (or in FIG. 1C1, or in FIG. 1C2), 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 100C (or 100C1, or 100C2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0127] 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

[0128] Network 100D of FIG. 1D is an example of general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ 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 100D of FIG. 1D, $N_1 = N$ and $N_2 = p * N$. The general

asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention. (And in the example of FIG. 1D, $s=2$). The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ 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 $d+d_2$ ($=d+p*d$) 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($2 \times \log_d N_1 - 2, d+d_2$) to the output switch OS1).

[0129] 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.

[0130] Each of the

$$2 \times \frac{N_1}{d}$$

US 2010/0135286 A1

Jun. 3, 2010

14

middle switches MS(1,1)-MS(1,2 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.

[0131] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches

$$MS(\text{Log}_d N_1 - 1, 1) - MS\left(\text{Log}_d N_1 - 1, 2 \times \frac{N_1}{d}\right)$$

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.

[0132] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches

$$MS(2 \times \text{Log}_d N_1 - 3, 1) - MS\left(2 \times \text{Log}_d N_1 - 3, 2 \times \frac{N_1}{d}\right)$$

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

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

output switches in output stage **120** through d links.

[0133] 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 $d + d_2$ links.

[0134] As described before, again the connection topology of a general $V(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V(N_1, N_2, d, s)$ 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 $V(N_1, N_2, d, s)$ network is, when no connections are setup in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N_1,$

$N_2, d, s)$ can be built. The embodiments of FIG. 1C, FIG. 1C1, and FIG. 1C2 are three examples of network $V(N_1, N_2, d, s)$ for $s=2$ and $N_2 > N_1$.

[0135] The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention.

[0136] For example, the network of FIG. 1C shows an exemplary five-stage network, namely $V(8, 24, 2, 2)$, 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**.

[0137] 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 RNB ($N_1 > N_2$) Embodiments

[0138] Referring to FIG. 1E, in one embodiment, an exemplary asymmetrical multi-stage network **100E** 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).

[0139] 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**.

[0140] 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 2010/0135286 A1

Jun. 3, 2010

15

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

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

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

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. 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 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.

[0141] 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).

[0142] 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).

[0143] 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).

[0144] 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)).

[0145] Each of the

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

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(3,2), MS(3,5), and MS(3,6) through the links ML(4,1), ML(4,3), ML(4,9), and ML(4,11) respectively).

US 2010/0135286 A1

Jun. 3, 2010

16

[0146] Finally the connection topology of the network 100E shown in FIG. 1E is known to be back to back inverse Benes connection topology.

[0147] Referring to FIG. 1E1, in another embodiment of network $V(N_1, N_2, d, s)$, an exemplary asymmetrical multi-stage network 100E1 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).

[0148] 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.

[0149] 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 multi-stage network of FIG. 1E1 is also the network of the type $V(N_1, N_2, d, s)$, 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 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.

[0150] 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).

[0151] 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).

[0152] Similarly each of the

$$2 \times \frac{N_2}{d}$$

US 2010/0135286 A1

Jun. 3, 2010

17

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).

[0153] 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(3,5) 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)).

[0154] 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).

[0155] Finally the connection topology of the network **100E1** shown in FIG. **1E1** is known to be back to back Omega connection topology.

[0156] Referring to FIG. **1E2**, in another embodiment of network $V(N_1, N_2, d, s)$, an exemplary asymmetrical multi-stage network **100E2** 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).

[0157] 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**.

[0158] 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 multi-stage network of FIG. **1E1** is also the network of the type $V(N_1, N_2, d, s)$, 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 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.

[0159] 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).

US 2010/0135286 A1

Jun. 3, 2010

18

[0160] 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).

[0161] 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).

[0162] 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)).

[0163] 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).

[0164] Finally the connection topology of the network **100E2** shown in FIG. 1E2 is hereinafter called nearest neighbor connection topology.**[0165]** In the three embodiments of FIG. 1E, FIG. 1E1 and FIG. 1E2 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 $V(N_1, N_2, d, s)$ 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1E, FIG. 1E1, and FIG. 1E2 are only three examples of network $V(N_1, N_2, d, s)$.**[0166]** In the three embodiments of FIG. 1E, FIG. 1E1 and FIG. 1E2, 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.**[0167]** In the example illustrated in FIG. 1E (or in FIG. 1E1, or in FIG. 1E2), 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 **100E** (or **100E1**, or **100E2**), to be operated in rearrangeably nonblocking manner in accordance with the invention.**[0168]** 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).

US 2010/0135286 A1

Jun. 3, 2010

19

[0169] 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 ($N_1 > N_2$) Embodiments

[0170] Network **100F** of FIG. **1F** is an example of general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ 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 **100D** of FIG. **1F**, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention. (And in the example of FIG. **1F**, $s=2$). The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ 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 $IL1-IL(p*d)$ to the input switch $IS1$) and $d+d_1 (=d+p*d)$ 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 $2 \times d$ 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 $OS1$).

[0171] 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.

[0172] Each of the

$$2 \times \frac{N_2}{d}$$

middle switches $MS(1,1)-MS(1, 2*N_2/d)$ 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.

[0173] 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** are connected from exactly d switches in middle stage **130** through d links and also are connected to exactly d switches in middle stage **130** through d links.

[0174] 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** are connected from exactly d switches in middle stage **130** through d links and also are connected to exactly d output switches in output stage **120** through d links.

[0175] 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** through $2 \times d$ links.

[0176] As described before, again the connection topology of a general $V(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V(N_1, N_2, d, s)$ may be back to back inverse Benes net-

US 2010/0135286 A1

Jun. 3, 2010

20

works, 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 $V(N_1, N_2, d, s)$ network is, when no connections are setup in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1E, FIG. 1E1, and FIG. 1E2 are three examples of network $V(N_1, N_2, d, s)$ for $s=2$ and $N_1 > N_2$.

[0177] The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention.

[0178] For example, the network of FIG. 1E shows an exemplary five-stage network, namely $V(24, 8, 2, 2)$, 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.

[0179] 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.

Symmetric SNB Embodiments

[0180] Referring to FIG. 2C, FIG. 2C1, and FIG. 2C2, three exemplary symmetrical multi-stage networks 200C, 200C1, and 200C2 respectively with five stages of 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, 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 twelve, two by two switches MS(1,1)-MS(1,12), middle stage 140 consists of twelve, two by two switches MS(2,1)-MS(2,12), and middle stage 150 consists of twelve, two by two switches MS(3,1)-MS(3,12).

[0181] Such a network can be operated in strictly nonblocking 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 twelve switches in each of middle stage 130, middle stage 140 and middle stage 150.

[0182] 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

$$3 \times \frac{N}{d}$$

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 $3d \times d$. 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 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.

[0183] 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 in FIG. 2A, input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,5), MS(1,6), MS(1,9) and MS(1,10) through the links ML(1,1), ML(1,2), ML(1,3), ML(1,4), ML(1,5) and ML(1,6) respectively).

[0184] Each of the

$$3 \times \frac{N}{d}$$

middle switches MS(1,1)-MS(1,12) in the middle stage 130 are connected from exactly d input switches through d links (for example in FIG. 2A, the links ML(1,1) and ML(1,7) 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).

[0185] Similarly each of the

$$3 \times \frac{N}{d}$$

middle switches MS(2,1)-MS(2,12) 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).

[0186] Similarly each of the

$$3 \times \frac{N}{d}$$

US 2010/0135286 A1

Jun. 3, 2010

21

middle switches MS(3,1)-MS(3,12) 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)).

[0187] 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 switches MS(3,1), MS(3,2), MS(3,5), MS(3,6), MS(3,9) and MS(3,10) through the links ML(4,1), ML(4,3), ML(4,9), ML(4,11), ML(4,17) and ML(4,19) respectively).

[0188] Finally the connection topology of the network **200A** shown in FIG. 2A is known to be back to back inverse Benes connection topology; the connection topology of the network **200A1** shown in FIG. 2A1 is known to be back to back Omega connection topology; and the connection topology of the network **200A2** shown in FIG. 2A2 is hereinafter called nearest neighbor connection topology.

[0189] In the three embodiments of FIG. 2A, FIG. 2A1 and FIG. 2A2 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 three 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N,d,s)$ can be built. The embodiments of FIG. 2A, FIG. 2A1, and FIG. 2A2 are only three examples of network $V(N,d,s)$.

[0190] In the three embodiments of FIG. 2A, FIG. 2A1 and FIG. 2A2, 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(1,12), MS(2,1)-MS(2,12), and MS(3,1)-MS(3,12) are referred to as middle switches or middle ports.

[0191] In the example illustrated in FIG. 2A, FIG. 2A1, and 2A2, a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two middle 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** (or **200A1**, or **200A2**), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0192] 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 SNB Embodiments

[0193] Network **200B** of FIG. 2B1 is an example of general symmetrical multi-stage network $V(N,d,s)$ with $(2 \times \log_d N) - 1$ stages. Network **200B** of FIG. 2B1 contains three different copies of the network **200B2** in FIG. 2B2. The general symmetrical multi-stage network $V(N,d,s)$ can be operated in strictly nonblocking manner for multicast when $s=3$ according to the current invention (and in the example of FIG. 2B1, $s=3$). 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 $3 \times d$ 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 $3 \times d$ incoming links for each of N/d output switches OS1-OS(N/d) (for example ML($2 \times \log_d N - 2, 1$)-ML($2 \times \log_d N - 2, 3 \times d$) to the output switch OS1).

[0194] Each of the N/d input switches IS1-IS(N/d) are connected to exactly $3 \times d$ switches in middle stage **130** through $3 \times d$ links.

[0195] Each of the

$$3 \times \frac{N}{d}$$

middle switches MS(1,1)-MS(1,3N/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.

[0196] Similarly each of the

$$3 \times \frac{N}{d}$$

middle switches

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

US 2010/0135286 A1

Jun. 3, 2010

22

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.

[0197] Similarly each of the

$$3 \times \frac{N}{d}$$

middle switches

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

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.

[0198] Each of the N/d output switches OS1-OS(N/d) are connected from exactly $3 \times d$ switches in middle stage $130+10*(2*\text{Log}_d N-4)$ through $3 \times d$ links.

[0199] The general symmetrical multi-stage network $V(N, d, s)$ can be operated in strictly nonblocking manner for multicast when $s=3$ according to the current invention.

[0200] 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. 2A shows an exemplary five-stage network, namely $V(8,2,3)$, with the following multicast assignment $I_1=\{1,3\}$ and all other $I_i=0$ 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(1,2) and MS(1,5) in middle stage **130**, and fans out in middle switches MS(1,2) and MS(1,5) only once into middle switches MS(2,2) and MS(2,5) respectively in middle stage **140**.

[0201] The connection I_1 also fans out in middle switches MS(2,2) and MS(2,5) only once into middle switches MS(3,2) and MS(3,7) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,2) and MS(3,7) only once into output switches OS1 and OS3 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 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 switches in middle stage **130**.

Asymmetric SNB ($N_2 > N_1$) Embodiments

[0202] Referring to FIG. 2C, FIG. 2C1, and FIG. 2C2, three exemplary symmetrical multi-stage networks **200C**, **200C1**,

and **200C2** respectively with five stages of 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**, and **150** is shown where input stage **110** consists of four, two by six switches IS1-IS4 and output stage **120** consists of four, twelve by six switches OS1-OS4. And all the middle stages namely middle stage **130** consists of twelve, two by two switches MS(1,1)-MS(1,12), middle stage **140** consists of twelve, two by two switches MS(2,1)-MS(2,12), and middle stage **150** consists of twelve, two by four switches MS(3,1)-MS(3,12).

[0203] Such a network can be operated in strictly nonblocking 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 twelve by six, and there are twelve switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0204] 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

$$3 \times \frac{N_1}{d}.$$

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 $(2d+d_2) * d$, 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

$$d * \frac{(2d + d_2)}{3}.$$

(Throughout the current invention, a fraction is rounded to the nearest higher integer). 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

US 2010/0135286 A1

Jun. 3, 2010

23

is the ratio of number of outgoing links from each input switch to the inlet links of each input switch.

[0205] 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 in FIG. 2C, input switch IS1 is connected to middle switches MS(1,1), MS(1,2), MS(1,5), MS(1,6), MS(1,9) and MS(1,10) through the links ML(1,1), ML(1,2), ML(1,3), ML(1,4), ML(1,5) and ML(1,6) respectively).

[0206] Each of the

$$3 \times \frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,12) in the middle stage **130** are connected from exactly d input switches through d links (for example in FIG. 2C, the links ML(1,1) and ML(1,7) 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).

[0207] Similarly each of the

$$3 \times \frac{N_1}{d}$$

middle switches MS(2,1)-MS(2,12) 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).

[0208] Similarly each of the

$$3 \times \frac{N_1}{d}$$

middle switches MS(3,1)-MS(3,12) 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

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

output switches in output stage **120** through

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

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 switch MS(3,1)).

[0209] Each of the

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

output switches OS1-OS4 are connected from exactly $2d + d_2$ switches in middle stage **150** through $2d + d_2$ 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), MS(3,8), MS(3,9), MS(3,10), MS(3,11), and MS(3,12) through the links ML(4,1), ML(4,5), ML(4,9), ML(4,13), ML(4,17), ML(4,21), ML(4,25), ML(4,29), ML(4,33), ML(4,37), ML(4,41), and ML(4,45) respectively).

[0210] Finally the connection topology of the network **200C** shown in FIG. 2C is known to be back to back inverse Benes connection topology; the connection topology of the network **200C1** shown in FIG. 2C1 is known to be back to back Omega connection topology; and the connection topology of the network **200C2** shown in FIG. 2C2 is hereinafter called nearest neighbor connection topology.

[0211] In the three embodiments of FIG. 2C, FIG. 2C1 and FIG. 2C2 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 three 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network $V(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 2C, FIG. 2C1, and FIG. 2C2 are only three examples of network $V(N_1, N_2, d, s)$.

[0212] In the three embodiments of FIG. 2C, FIG. 2C1 and FIG. 2C2, 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(1,12), MS(2,1)-MS(2,12), and MS(3,1)-MS(3,12) are referred to as middle switches or middle ports.

[0213] In the example illustrated in FIG. 2C, FIG. 2C1, and 2C2, a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two middle 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

US 2010/0135286 A1

Jun. 3, 2010

24

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** (or **200C1**, or **200C2**), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0214] 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

[0215] Network **200D** of FIG. **2D1** is an example of general symmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages where $N_2 > N_1$ and $N_2 = p \times N_1$ where $p > 1$. In network **200D** of FIG. **2D**, $N_1 = N$ and $N_2 = p \times N$. Network **200D** of FIG. **2D1** contains three different copies of the network **200D2** in FIG. **2D2**. The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for multicast when $s=3$ according to the current invention (and in the example of FIG. **2D1**, $s=3$). 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 $3 \times d$ 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 $OS1$).

[0216] 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.

[0217] Each of the

$$3 \times \frac{N_1}{d}$$

middle switches $MS(1,1)-MS(1,3 N_1/d)$ 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.

[0218] Similarly each of the

$$3 \times \frac{N_1}{d}$$

middle switches

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

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

[0219] Similarly each of the

$$3 \times \frac{N_1}{d}$$

middle switches

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

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 d links and also are connected to exactly

US 2010/0135286 A1

Jun. 3, 2010

25

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

output switches in output stage **120** through

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

links.

[0220] Each of the

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

output switches OS1-OS(N_1/d) are connected from exactly $2d+d_2$ switches in middle stage **130**+ $10*(2*\text{Log}_d N_1-4)$ through $2d+d_2$ links.

[0221] The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=3$ according to the current invention.

[0222] For example, the network of FIG. 2C shows an exemplary five-stage network, namely $V(8, 2, 3)$, 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(1,2) and MS(1,5) in middle stage **130**, and fans out in middle switches MS(1,2) and MS(1,5) only once into middle switches MS(2,4) and MS(2,5) respectively in middle stage **140**.

[0223] The connection I_1 also fans out in middle switches MS(2,4) and MS(2,5) only once into middle switches MS(3,4) and MS(3,7) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,4) 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 OS1 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 switches in middle stage **130**.

Asymmetric SNB ($N_1 > N_2$) Embodiments

[0224] Referring to FIG. 2E, FIG. 2E1, and FIG. 2E2, three exemplary symmetrical multi-stage networks **200E**, **200E1**, and **200E2** respectively with five stages of 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**, and **150** is shown where input stage **110** consists of four, six by twelve 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 twelve, four by two switches MS(1,1)-MS(1,12), middle stage **140** consists of twelve, two by two switches MS(2,1)-MS(2,12), and middle stage **150** consists of twelve, two by two switches MS(3,1)-MS(3,12).

[0225] Such a network can be operated in strictly nonblocking manner for multicast connections, because the switches in the input stage **110** are of size six by twelve, the switches in

output stage **120** are of size six by two, and there are twelve switches in each of middle stage **130**, middle stage **140** and middle stage **150**.

[0226] 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

$$3 * \frac{N_2}{d}$$

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

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

and each output switch OS1-OS4 can be denoted in general with the notation $3d*d$. 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

$$\frac{(2d + d_1)}{3} * d.$$

(Throughout the current invention, a fraction is rounded to the nearest higher integer). 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 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.

[0227] Each of the

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

input switches IS1-IS4 are connected to exactly $2d+d_1$ switches in middle stage **130** through $2d+d_1$ links (for example in FIG. 2E, 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), MS(1,8), MS(1,9), MS(1,10), MS(1,11) and MS(1,12) through the links ML(1,1), ML(1,2), ML(1,3),

US 2010/0135286 A1

Jun. 3, 2010

26

ML(1,4), ML(1,5), ML(1,6), ML(1,7), ML(1,8), ML(1,9), ML(1,10), ML(1,11), and ML(1,12) respectively).

[0228] Each of the

$$3 \times \frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,12) in the middle stage 130 are connected from exactly

$$\frac{2d + d_1}{3}$$

input switches through

$$\frac{2d + d_1}{3}$$

links (for example in FIG. 2E, the links ML(1,1), ML(1,13), ML(1,25), and ML(1,37) 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).

[0229] Similarly each of the

$$3 \times \frac{N_2}{d}$$

middle switches MS(2,1)-MS(2,12) 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).

[0230] Similarly each of the

$$3 \times \frac{N_2}{d}$$

middle switches MS(3,1)-MS(3,12) 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), ML(4,2), ML(4,3), and ML(4,4) are connected to output switches OS1, OS2, OS3, and OS4 respectively from middle switch MS(3,1)).

[0231] Each of the

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

output switches OS1-OS4 are connected from exactly 3d switches in middle stage 150 through 3d 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), MS(3,8), MS(3,9), MS(3,10), MS(3,11), and MS(3,12) through the links ML(4,1), ML(4,5), ML(4,9), ML(4,13), ML(4,17), ML(4,21), ML(4,25), ML(4,29), ML(4,33), ML(4,37), ML(4,41), and ML(4,45) respectively).

[0232] Finally the connection topology of the network 200E shown in FIG. 2E is known to be back to back inverse Benes connection topology; the connection topology of the network 200E1 shown in FIG. 2E1 is known to be back to back Omega connection topology; and the connection topology of the network 200E2 shown in FIG. 2E2 is hereinafter called nearest neighbor connection topology.

[0233] In the three embodiments of FIG. 2E, FIG. 2E1 and FIG. 2E2 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 three 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 in the network, a connection from any inlet link to any outlet link can be setup. Based on this property numerous embodiments of the network V(N₁,N₂,d,s) can be built. The embodiments of FIG. 2E, FIG. 2E1, and FIG. 2E2 are only three examples of network V(N₁,N₂,d,s).

[0234] In the three embodiments of FIG. 2E, FIG. 2E1 and FIG. 2E2, 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(1,12), MS(2,1)-MS(2,12), and MS(3,1)-MS(3,12) are referred to as middle switches or middle ports.

[0235] In the example illustrated in FIG. 2E, FIG. 2E1, and 2E2, a fan-out of four is possible to satisfy a multicast connection request if input switch is IS2, but only two middle 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 (or 200E1, or 200E2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

US 2010/0135286 A1

Jun. 3, 2010

27

[0236] 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

[0237] Network **200F** of FIG. **2F1** is an example of general symmetrical multi-stage network $V(N_1, N_2, d, s)$ with $(2 \times \log_d N) - 1$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network **200F** of FIG. **2F**, $N_2 = N$ and $N_1 = p * N$. Network **200F** of FIG. **2F1** contains three different copies of the network **200F2** in FIG. **2F2**. The general asymmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for multicast when $s=3$ according to the current invention (and in the example of FIG. **2F1**, $s=3$). 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 $IL1-IL(p*d)$ to the input switch $IS1$) and $2d+d_1 (=2d+p*d)$ 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 $3 \times d$ 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 $OS1$).

[0238] Each of the

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

input switches $IS1-IS(N_2/d)$ are connected to exactly $2d+d_1$ switches in middle stage **130** through $2d+d_1$ links.

[0239] Each of the

$$3 \times \frac{N_2}{d}$$

middle switches $MS(1,1)-MS(1,3 N_2/d)$ 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.

[0240] Similarly each of the

$$3 \times \frac{N_2}{d}$$

middle switches

$$MS(\log_d N_2 - 1, 1) - MS(\log_d N_2 - 1, 3 \times \frac{N_2}{d})$$

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.

[0241] Similarly each of the

$$3 \times \frac{N_2}{d}$$

middle switches

$$MS(2 \times \log_d N_2 - 3, 1) - MS(2 \times \log_d N_2 - 3, 3 \times \frac{N_2}{d})$$

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.

US 2010/0135286 A1

Jun. 3, 2010

28

[0242] Each of the

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

output switches OS1-OS(N_2/d) are connected from exactly $3 \times d$ switches in middle stage **130** + $10 \times (2 \times \text{Log}_d N_2 - 4)$ through $3 \times d$ links.

[0243] The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=3$ according to the current invention.

[0244] For example, the network of FIG. 2E shows an exemplary five-stage network, namely $V(8, 2, 3)$, with the following multicast assignment $I_1 = \{1, 3\}$ 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(1,2) and MS(1,5) in middle stage **130**, and fans out in middle switches MS(1,2) and MS(1,5) only once into middle switches MS(2,4) and MS(2,5) respectively in middle stage **140**.

[0245] The connection I_1 also fans out in middle switches MS(2,4) and MS(2,5) only once into middle switches MS(3,2) and MS(3,7) respectively in middle stage **150**. The connection I_1 also fans out in middle switches MS(3,2) and MS(3,7) only once into output switches OS1 and OS3 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 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 switches in middle stage **130**.

Applications Embodiments

[0246] All the embodiments disclosed in the current invention are useful in many varieties of applications. FIG. 4A1 illustrates the diagram of **400A1** 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. 4A1. For example the diagram of **400A1** may be the implementation of middle switch MS(1,1) of the diagram **100A** of FIG. 1A where inlet link IL1 of diagram **400A1** corresponds to middle link ML(1,1) of diagram **100A**, inlet link IL2 of diagram **400A1** corresponds to middle link ML(1,5) of diagram **100A**, outlet link OL1 of diagram **400A1** corresponds to middle link ML(2,1) of diagram **100A**, outlet link OL2 of diagram **400A1** corresponds to middle link ML(2,2) of diagram **100A**.

1) Programmable Integrated Circuit Embodiments

[0247] All the embodiments disclosed in the current invention are useful in programmable integrated circuit applications. FIG. 4A2 illustrates the detailed diagram **400A2** for the implementation of the diagram **400A1** 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).

[0248] 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.

2) One-time Programmable Integrated Circuit Embodiments

[0249] All the embodiments disclosed in the current invention are useful in one-time programmable integrated circuit applications. FIG. 4A3 illustrates the detailed diagram **400A3** for the implementation of the diagram **400A1** 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.

[0250] 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 **400A3** 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 embodi-

US 2010/0135286 A1

Jun. 3, 2010

29

ments may be anti-fuse based programmable integrated circuit devices or mask programmable structured ASIC devices.

3) Integrated Circuit Placement and Route Embodiments

[0251] 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. 4A4 illustrates the detailed diagram 400A4 for the implementation of the diagram 400A1 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.

[0252] 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 400A4 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 400A4, 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.

[0253] 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

[0254] 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.

Scheduling Method Embodiments

[0255] FIG. 3A 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 $V(N_1, N_2, d, s)$ disclosed in this inven-

tion). According to this embodiment, a multicast connection request is received in act 1010. Then the control goes to act 1020.

[0256] 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 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 1030.

[0257] 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 $130+10*(2*\text{Log}_d N-4)$ 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 1040.

[0258] 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

US 2010/0135286 A1

Jun. 3, 2010

30

of all the reachable destination links from each outgoing link of the input switch is derived. The control then goes to act **1050**.

[0259] 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**.

[0260] 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**.

[0261] 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.

[0262] 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. 3A 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 non-blocking manner in accordance with the invention.

[0263] According to the current invention, using the method **1000** of FIG. 3A, the network $V(N_1, N_2, d, s)$ 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 non-blocking for multicast connections when $s \geq 2$, and is operated in strictly nonblocking for multicast connections when $s \geq 3$.

[0264] The connection request of the type described above in reference to method **1000** of FIG. 3A 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. 3A 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.

[0265] As noted above method **1000** of FIG. 3A can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks $V(N, d, s)$ and $V(N_1, N_2, d, s)$ disclosed in this invention.

[0266] 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

$$x \times \frac{N}{d}$$

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 d incoming links (herein-

US 2010/0135286 A1

Jun. 3, 2010

31

after “incoming middle links”) connecting from switches in its immediate preceding stage, and each middle switch further comprising d outgoing links (hereinafter “outgoing middle links”) connecting to switches in its immediate succeeding stage; and

each middle switch in said penultimate stage comprising d incoming links (hereinafter “incoming middle links”) connecting from switches in its immediate preceding stage, and each middle switch further comprising

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

outgoing links (hereinafter “outgoing middle links”) connecting to switches in its immediate succeeding stage; or

when $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$ then $N_2 = N$, $d_1 = d$ 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

$$x \times \frac{N}{d}$$

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

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

incoming links (hereinafter “incoming middle links”) connecting from switches in its immediate preceding stage, and each middle switch further comprising d outgoing links (hereinafter “outgoing middle links”) connecting to switches in its immediate succeeding stage; and

each middle switch in all said middle stages excepting said second stage comprising d incoming links (hereinafter “incoming middle links”) connecting from switches in its immediate preceding stage, and each middle switch further comprising 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 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 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

US 2010/0135286 A1

Jun. 3, 2010

32

existing multicast connection, and said network is hereinafter “rearrangeably nonblocking network”.

7. The network of claim 3, wherein $x \geq 3$, 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 non-blocking 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 each of said input switches, or each of said output switches, or each of said middle switches further recursively comprise one or more networks.

11. 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

$$x \times \frac{N}{d}$$

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 d incoming links connected

from switches in its immediate preceding stage, and each middle switch further having d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said penultimate stage having d incoming links connected from switches in its immediate preceding stage, and each middle switch further having

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

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

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

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

$$x \times \frac{N}{d}$$

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

US 2010/0135286 A1

33

Jun. 3, 2010

each middle switch in said second stage having

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

incoming links connected from switches in its immediate preceding stage, and each middle switch further having 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 d incoming links connected from switches in its immediate preceding stage, and each middle switch further having 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.

12. The method of claim **11** 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.

13. The method of claim **11** wherein said act of fanning out is performed recursively.

14. The method of claim **11** 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".

15. The method of claim **11** wherein said acts of fanning out and rearranging are performed recursively.

16. 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

$$x \times \frac{N}{d}$$

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 d incoming links connected from switches in its immediate preceding stage, and each middle switch further having d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said penultimate stage having d incoming links connected from switches in its immediate preceding stage, and each middle switch further having

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

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

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

and having
an input stage having

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

US 2010/0135286 A1

34

Jun. 3, 2010

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

$$x \times \frac{N}{d}$$

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

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

incoming links connected from switches in its immediate preceding stage, and each middle switch further having 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 d incoming links connected from switches in its immediate preceding stage, and each middle switch further having 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.

17. The method of claim **16** 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

18. The method of claim **16** 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.

19. The method of claim **16** 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.

20. The method of claim **16** 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.

21. The method of claim **16** wherein any of said acts of checking and setting up are performed recursively.

* * * * *

EXHIBIT C

(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 2008/147926 A1

(51) International Patent Classification:
H03K 17/00 (2006.01)

(21) International Application Number:
PCT/US2008/064603

(22) International Filing Date: 22 May 2008 (22.05.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/940,387 25 May 2007 (25.05.2007) US
60/940,390 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).

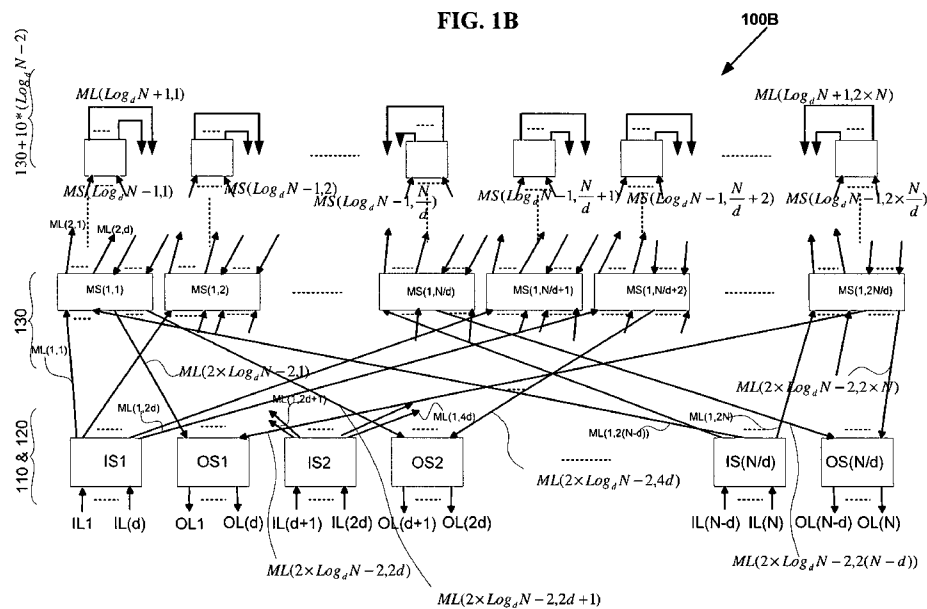
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, 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: FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS



(57) Abstract: A generalized butterfly fat tree network comprising $(\log_d N)$ stages is operated in strictly nonblocking manner for unicast, when $s > \text{or} = 2$, includes a leaf stage consisting of an input stage having N/d switches with each of them having d inlet links and $s \times d$ outgoing links connecting to its immediate succeeding stage switches, and an output stage having N/d switches with each of them having d outlet links and $s \times d$ incoming links connecting from switches in its immediate succeeding stage

WO 2008/147926 A1

WO 2008/147926

PCT/US2008/064603

FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE 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, 387 entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current application,
10 filed May 25, 2007.

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, 390 entitled "FULLY CONNECTED GENERALIZED MULTI-LINK BUTTERFLY FAT TREE NETWORKS" by Venkat Konda assigned to the same assignee as the current
15 application, filed May 25, 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
20 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
25 GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

WO 2008/147926

PCT/US2008/064603

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

5 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

10 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 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.

15 This application is related to and incorporates by reference in its entirety the PCT Application Docket No. S-0045PCT entitled "VLSI LAYOUTS OF FULLY

CONNECTED GENERALIZED NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed concurrently, and the U.S. Provisional Patent

Application Serial No. 60/940, 394 entitled "VLSI LAYOUTS OF FULLY

20 CONNECTED GENERALIZED 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. Provisional Patent Application Serial No. 60/984, 724 entitled "VLSI LAYOUTS OF

FULLY CONNECTED NETWORKS WITH LOCALITY EXPLOITATION" by Venkat

25 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

FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS" by Venkat

Konda assigned to the same assignee as the current application, filed January 1, 2008.

WO 2008/147926

PCT/US2008/064603

BACKGROUND OF INVENTION

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

5 "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.

10 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 network's ability to rearrange prior

15 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

20 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.

25 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. 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

30 nonblocking for unicast connections.

WO 2008/147926

PCT/US2008/064603

U.S. Patent 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.

U.S. Patent 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. Patent 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 $m \geq 3 \times n - 1$.

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, Jan. 1953, No.1, pp. 406-424 showed a way of constructing large multi-stage networks by recursive substitution with a crosspoint complexity of $d^2 \times N \times (\log_d N)^{2.58}$ for strictly nonblocking unicast network. Similarly U.S. Patent 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

WO 2008/147926

PCT/US2008/064603

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 $N=256$.

The crosspoint complexity of all these networks is prohibitively large to
5 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.

10 SUMMARY OF INVENTION

A generalized butterfly fat tree network comprising $(\log_d N)$ stages is operated in strictly nonblocking manner for unicast includes a leaf stage consisting of an input stage having $\frac{N}{d}$ switches with each of them having d inlet links and $2 \times d$ outgoing links connecting to its immediate succeeding stage switches, and an output stage having $\frac{N}{d}$
15 switches with each of them having d outlet links and $2 \times d$ incoming links connecting from switches in its immediate succeeding stage. The network also has $(\log_d N) - 1$ middle stages with each middle stage, excepting the root stage, having $\frac{2 \times N}{d}$ switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, d incoming links connecting from the switches in its
20 immediate succeeding stage, d outgoing links connecting to the switches in its immediate succeeding stage, d outgoing links connecting to the switches in its immediate preceding stage, and the root stage having $\frac{2 \times N}{d}$ switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage and d outgoing links connecting to the switches in its immediate
25 preceding stage. Also the same generalized butterfly fat tree network is operated in

WO 2008/147926

PCT/US2008/064603

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.

A generalized butterfly fat tree network comprising $(\log_d N)$ stages is operated in strictly nonblocking manner for multicast includes a leaf stage consisting of an input

5 stage having $\frac{N}{d}$ switches with each of them having d inlet links and $3 \times d$ outgoing links connecting to its immediate succeeding stage switches, an output stage having $\frac{N}{d}$ switches with each of them having d outlet links and $3 \times d$ incoming links connecting from switches in its immediate succeeding stage. The network also has $(\log_d N) - 1$ middle stages with each middle stage, excepting the root stage, having $\frac{3 \times N}{d}$ switches,

10 and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, d incoming links connecting from the switches in its immediate succeeding stage, d outgoing links connecting to the switches in its immediate succeeding stage, d outgoing links connecting to the switches in its immediate preceding stage, and the root stage having $\frac{3 \times N}{d}$ switches, and each switch in

15 the middle stage has d incoming links connecting from the switches in its immediate preceding stage and d outgoing links connecting to the switches in its immediate preceding stage.

BRIEF DESCRIPTION OF DRAWINGS

20 FIG. 1A is a diagram 100A of an exemplary Symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ having inverse Benes connection topology of three 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.

WO 2008/147926

PCT/US2008/064603

FIG. 1B is a diagram 100B of a general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ with $(\log_d N)$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

5 FIG. 1C is a diagram 100C of an exemplary Asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, 2)$ having inverse Benes connection topology of three 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.

10 FIG. 1D is a diagram 100D of a general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, 2)$ with $N_2 = p * N_1$ and with $(\log_d N)$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

15 FIG. 1E is a diagram 100E of an exemplary Asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, 2)$ having inverse Benes connection topology of three 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.

20 FIG. 1F is a diagram 100F of a general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, 2)$ with $N_1 = p * N_2$ and with $(\log_d N)$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

25 FIG. 2A is a diagram 200A of an exemplary Symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ having inverse Benes connection topology of three stages with $N = 8$, $d = 2$ and $s=1$ with exemplary unicast connections rearrangeably nonblocking network for unicast connections, in accordance with the invention.

WO 2008/147926

PCT/US2008/064603

FIG. 2B is a diagram 200B of a general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ with $(\log_d N)$ stages and $s=1$, rearrangeably nonblocking network for unicast connections in accordance with the invention.

FIG. 2C is a diagram 200C of an exemplary Asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, 1)$ having inverse Benes connection topology of three 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.

FIG. 2D is a diagram 200D of a general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, 1)$ with $N_2 = p * N_1$ and with $(\log_d N)$ stages rearrangeably nonblocking network for unicast connections in accordance with the invention.

FIG. 2E is a diagram 200E of an exemplary Asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, 1)$ having inverse Benes connection topology of three 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.

FIG. 2F is a diagram 200F of a general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, 1)$ with $N_1 = p * N_2$ and with $(\log_d N)$ stages rearrangeably nonblocking network for unicast connections in accordance with the invention.

FIG. 3A is a diagram 300A of an exemplary symmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N, d, s)$ 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.

FIG. 3B is a diagram 300B of a general symmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N, d, 2)$ with $(\log_d N)$ stages strictly nonblocking network for unicast

WO 2008/147926

PCT/US2008/064603

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

FIG. 3C is a diagram 300C of an exemplary asymmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, 2)$ 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.

FIG. 3D is a diagram 300D of a general asymmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, 2)$ with $N_2 = p * N_1$ and with $(\log_d N)$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

FIG. 3E is a diagram 300E of an exemplary asymmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, 2)$ 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.

FIG. 3F is a diagram 300F of a general asymmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, 2)$ with $N_1 = p * N_2$ and with $(\log_d N)$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

FIG. 4A 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.

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

WO 2008/147926

PCT/US2008/064603

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.

5

DETAILED DESCRIPTION OF THE INVENTION

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

When a transmitting device simultaneously sends information to more than one receiving device, the one-to-many connection required between the transmitting device
15 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
20 connection required between the transmitting device and the receiving devices is called a broadcast connection.

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
25 of the available outlet links.

In certain butterfly fat tree networks and multi-link butterfly fat tree 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

WO 2008/147926

PCT/US2008/064603

blocking if necessary by rearranging some of the previous connection requests. In certain other Butterfly fat tree 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.

In certain butterfly fat tree networks and multi-link butterfly fat tree 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 Butterfly fat tree 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.

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

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.

2) 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 U.S. Provisional Patent Application Serial No. 60/940,389 that is incorporated by reference above.

3) 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 U.S.

WO 2008/147926

PCT/US2008/064603

Provisional Patent Application Serial No. 60/940, 391 that is incorporated by reference above.

4) 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
5 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, 392 that is incorporated by reference above.

5) 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
10 $V_{bft}(N_1, N_2, d, s)$, 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, are described in detail in U.S. Provisional Patent Application Serial No. 60/940, 394 that is incorporated
15 by reference above.

6) 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.

7) VLSI layouts of numerous types of multistage pyramid networks are described
20 in U.S. Provisional Patent Application Serial No. 61/018, 494 that is incorporated by reference above.

WO 2008/147926

PCT/US2008/064603

BUTTERFLY FAT TREE EMBODIMENTS:

Symmetric RNB Embodiments:

Referring to FIG. 1A, in one embodiment, an exemplary symmetrical butterfly fat tree network 100A with three stages of twenty four switches for satisfying

5 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, and 140 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. Input stage 110 and output stage 120 together belong to leaf stage. And all the middle

10 stages excepting root stage namely middle stage 130 consists of eight, four by four switches MS(1,1) - MS(1,8), and root stage i.e., middle stage 140 consists of eight, two by two switches MS(2,1) - MS(2,8).

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

15 switches in output stage 120 are of size four by two, and there are eight switches in each of middle stage 130 and middle stage 140. 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 and middle stage 140.

20 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 number of inlet links or outlet links. Input stage 110 and output stage 120 together belong to leaf stage. The number of middle switches in each middle stage is denoted by

25 $2 \times \frac{N}{d}$. 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 $2d * 2d$ excepting that the size of each switch in middle stage 140 is denoted as $d * d$.

WO 2008/147926

PCT/US2008/064603

(In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $d * 2d$ and $d * d$ since the down coming middle links are never setup to the up going middle links. For example in network 100A of FIG. 1A, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

10 Middle stage 140 is called as root stage. 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 Butterfly fat tree network can be represented with the notation $V_{bft}(N, d, s)$, 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.

Each of the $\frac{N}{d}$ 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).

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 are also 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(1,1) from middle switches MS(2,1) and MS(2,3) respectively).

WO 2008/147926

PCT/US2008/064603

Each of the $2 \times \frac{N}{d}$ middle switches MS(1,1) – MS(1,8) in the middle stage 130 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) 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(1,1)).

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 130 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(1,1) and MS(1,3) respectively).

Each of the $\frac{N}{d}$ output switches OS1 – OS4 are connected from exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example output switch OS1 is connected from middle switches MS(1,1), MS(1,2), MS(1,5) and MS(1,6) through the links ML(4,1), ML(4,3), ML(4,9) and ML(4,11) respectively).

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

In the three embodiments of FIG. 1A, FIG. 1A1 and FIG. 1A2 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 $V_{bfi}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bfi}(N, d, s)$ may be back to back Benes

WO 2008/147926

PCT/US2008/064603

networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{bft}(N, d, s)$ 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 $V_{bft}(N, d, s)$ can be built.

5 The embodiments of FIG. 1A, FIG. 1A1, and FIG. 1A2 are only three examples of network $V_{bft}(N, d, s)$.

In the three embodiments of FIG. 1A, FIG. 1A1 and FIG. 1A2, 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
10 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) and MS(2,1) – MS(2,8) are referred to as middle switches or middle ports. The middle stage
15 130 is also referred to as root stage and middle stage switches MS(1,2) – MS(2,8) are referred to as root stage switches.

In the example illustrated in FIG. 1A (or in FIG1A1, or in FIG. 1A2), 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
20 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 100A1,
25 or 100A2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

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

WO 2008/147926

PCT/US2008/064603

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:

Network 100B of FIG. 1B is an example of general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ with $(\log_d N)$ stages. The general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 2$ according to the current invention. Also the general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention. (And in the example of FIG. 1B, $s = 2$). The general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ with $(\log_d N)$ stages has d inlet links for each of $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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).

Each of the $\frac{N}{d}$ 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

WO 2008/147926

PCT/US2008/064603

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.

Each of the $2 \times \frac{N}{d}$ middle switches $MS(1,1) - MS(1,2N/d)$ in the middle stage

5 130 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.

Similarly each of the $2 \times \frac{N}{d}$ middle switches $MS(1,1) - MS(1,2N/d)$ in the middle stage 130 are also connected from exactly d switches in middle stage 140 through d links and also are connected to exactly d output switches in output stage 120 through d links.

10

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.

15 Each of the $\frac{N}{d}$ output switches $OS1 - OS(N/d)$ are connected from exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links.

As described before, again the connection topology of a general $V_{bfi}(N, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bfi}(N, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more combinations.

20 The applicant notes that the fundamental property of a valid connection topology of the general $V_{bfi}(N, d, s)$ 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

WO 2008/147926

PCT/US2008/064603

network $V_{bft}(N, d, s)$ can be built. The embodiments of FIG. 1A, FIG. 1A1, and FIG. 1A2 are three examples of network $V_{bft}(N, d, s)$.

The general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

Every switch in the Butterfly fat tree networks discussed herein has multicast capability. In a $V_{bft}(N, d, s)$ 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.

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. 1A shows an exemplary three-stage network, namely $V_{bft}(8, 2, 2)$, with the following

WO 2008/147926

PCT/US2008/064603

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 output switch OS2 in output stage 120 and middle switch MS(2,7) in middle stage 140 respectively.

The connection I_1 also fans out in middle switch MS(2,7) 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 switch MS(1,7) only once into output switch 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 RNB ($N_2 > N_1$) Embodiments:

Referring to FIG. 1C, in one embodiment, an exemplary asymmetrical Butterfly fat tree network 100C with three stages of twenty 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 and 140 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. Input stage 110 and output stage 120 together belong to leaf stage. Middle stage 130 consists of eight, four by six switches MS(1,1) - MS(1,8) and middle stage 140 consists of eight, two by two switches MS(2,1) - MS(2,8).

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 and middle stage 140. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the

WO 2008/147926

PCT/US2008/064603

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 four by six in middle stage 130 and eight switches of size two by two in middle stage 140.

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 middle stage 130 can be denoted as $2d * (d + d_2)$. The size of each switch in the root stage (i.e., middle stage 140) can be denoted as $d * d$. The size of each switch in all the middle stages excepting middle stage 130 and root stage can be denoted as $2d * 2d$ (In network 100C of FIG. 1C, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $d * 2d$ and $d * d$ since the down coming middle links are never setup to the up going middle links. For example in network 100C of FIG. 1C, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

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 Butterfly fat tree network can be represented with the notation $V_{bft}(N_1, N_2, d, s)$, where

WO 2008/147926

PCT/US2008/064603

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.

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).

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 are also 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(1,1) from middle switches MS(2,1) and MS(2,3) respectively).

Similarly each of the $2 \times \frac{N_1}{d}$ middle switches MS(1,1) – MS(1,8) in the middle stage 130 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) 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 switch MS(1,1)).

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

WO 2008/147926

PCT/US2008/064603

(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 130 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(1,1) and
 5 MS(1,3) respectively).

Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $d + d_2$ switches in middle stage 130 through $d + d_2$ links (for example output switch OS1 is connected from 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(4,1), ML(4,5), ML(4,9), ML(4,13),
 10 ML(4,17), ML(4,21), ML(4,25) and ML(4,29) respectively).

Finally the connection topology of the network 100C shown in FIG. 1C is known to be back to back inverse Benes connection topology.

In the three embodiments of FIG. 1C, FIG. 1C1 and FIG. 1C2 the connection topology is different. That is the way the links ML(1,1) - ML(1,16), ML(2,1) - ML(2,16),
 15 ML(3,1) - ML(3,16), and ML(4,1) - ML(4,32) are connected between the respective stages is different. Even though only three embodiments are illustrated, in general, the network $V_{bfi}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bfi}(N_1, N_2, d, s)$ may be back to back Benes networks, Delta Networks and many more combinations. The applicant notes that
 20 the fundamental property of a valid connection topology of the $V_{bfi}(N_1, N_2, d, s)$ 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 $V_{bfi}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1C, FIG. 1C1, and FIG. 1C2 are only three examples of network $V_{bfi}(N_1, N_2, d, s)$.

25 In the three embodiments of FIG. 1C, FIG. 1C1 and FIG. 1C2, 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(4,32) are either available for use by a new connection or not available if currently

WO 2008/147926

PCT/US2008/064603

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) and
 5 MS(2,1) – MS(2,8) are referred to as middle switches or middle ports.

In the example illustrated in FIG. 1C (or in FIG1C1, or in FIG. 1C2), 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
 10 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 100C (or 100C1, or 100C2), to be operated in rearrangeably nonblocking manner in accordance with the
 15 invention.

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,
 20 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
 25 the output stage switches to satisfy the connection request.

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

Network 100D of FIG. 1D is an example of general asymmetrical Butterfly fat tree network $V_{bf}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_2 > N_1$ and $N_2 = p * N_1$ where $p > 1$. In network 100D of FIG. 1D, $N_1 = N$ and $N_2 = p * N$. The general

WO 2008/147926

PCT/US2008/064603

asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 2$ according to the current invention. Also the general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention. (And in the example of FIG. 1D, $s = 2$). The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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

10 $d_2 = N_2 \times \frac{d}{N_1} = p \times d$) 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).

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.

Each of the $2 \times \frac{N_1}{d}$ middle switches MS(1,1) - MS(1,2 N_1/d) in the middle stage 20 130 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.

WO 2008/147926

PCT/US2008/064603

Similarly each of the $2 \times \frac{N_1}{d}$ middle switches $MS(1,1) - MS(1,2 N_1/d)$ in the middle stage 130 are connected from exactly d switches in middle stage 140 through d links and also are connected to exactly d output switches in output stage 120 through $\frac{d + d_2}{2}$ links.

5 Similarly each of the $2 \times \frac{N_1}{d}$ middle switches $MS(\text{Log}_d N_1 - 1, 1) - MS(\text{Log}_d N_1 - 1, 2 \times \frac{N_1}{d})$ 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.

10 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 through $d + d_2$ links.

As described before, again the connection topology of a general $V_{bft}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back
 15 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiments of FIG.
 20 1C, FIG. 1C1, and FIG. 1C2 are three examples of network $V_{bft}(N_1, N_2, d, s)$ for $s = 2$ and $N_2 > N_1$.

The general symmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the

WO 2008/147926

PCT/US2008/064603

current invention. Also the general symmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

For example, the network of Fig. 1C shows an exemplary three-stage network, namely $V_{bft}(8, 24, 2, 2)$, 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 OS2 and OS3 respectively 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 OL18. 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.

15 **Asymmetric RNB ($N_1 > N_2$) Embodiments:**

Referring to FIG. 1E, in one embodiment, an exemplary asymmetrical Butterfly fat tree network 100E with three stages of twenty 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 and 140 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. Middle stage 130 consists of eight, six by four switches MS(1,1) - MS(1,8) and middle stage 140 consists of eight, two by two switches MS(2,1) - MS(2,8).

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 and middle stage 140. Such a network can be operated in

WO 2008/147926

PCT/US2008/064603

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 six by four in middle stage 130, and eight switches of size two by two in middle stage 140.

5 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

10 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 middle stage 130 can be denoted as $(d + d_1) * 2d$. The size of each switch in the root stage (i.e., middle stage 140) can be denoted as $d * d$. The size of each switch in all the middle

15 stages excepting middle stage 130 and root stage can be denoted as $2d * 2d$ (In network 100E of FIG. 1E, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $d * 2d$ and $d * d$ since the down coming middle links are never setup to the up going middle links. For example in network 100E of FIG. 1E, the down coming middle links

20 ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as

25 outputs).

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

WO 2008/147926

PCT/US2008/064603

Butterfly fat tree network can be represented with the notation $V_{bft}(N_1, N_2, d, s)$, 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
 5 $N_1 > N_2$, and s is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

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),
 10 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).

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
 15 MS(1,1) from input switch IS1, IS2, IS3, and IS4 respectively) and also 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(1,1) from middle switches MS(2,1) and MS(2,3) respectively).

Similarly each of the $2 \times \frac{N_2}{d}$ middle switches MS(1,1) – MS(1,8) in the middle
 20 stage 130 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), 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
 25 MS(1,1)).

WO 2008/147926

PCT/US2008/064603

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
 5 exactly d switches in middle stage 130 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(1,3) and MS(1,1) respectively).

Each of the $\frac{N_2}{d}$ output switches OS1 – OS4 are connected from exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example output switch OS1 is
 10 connected from middle switches MS(1,1), MS(1,2), MS(1,5), and MS(1,6) through the links ML(4,1), ML(4,3), ML(4,9), and ML(4,11) respectively).

Finally the connection topology of the network 100E shown in FIG. 1E is known to be back to back inverse Benes connection topology.

In the three embodiments of FIG. 1E, FIG. 1E1 and FIG. 1E2 the connection
 15 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 $V_{bft}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ may be back to back
 20 Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection topology of the $V_{bft}(N_1, N_2, d, s)$ 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
 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1E, FIG. 1E1, and FIG. 1E2 are
 25 only three examples of network $V_{bft}(N_1, N_2, d, s)$.

WO 2008/147926

PCT/US2008/064603

In the three embodiments of FIG. 1E, FIG. 1E1 and FIG. 1E2, 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
5 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) and MS(2,1) – MS(2,8) are referred to as middle switches or middle ports.

In the example illustrated in FIG. 1E (or in FIG1E1, or in FIG. 1E2), a fan-out of
10 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
15 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 100E (or 100E1, or 100E2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

The connection request of the type described above can be unicast connection
20 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
25 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.

WO 2008/147926

PCT/US2008/064603

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

Network 100F of FIG. 1F is an example of general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network 100F of FIG. 1F, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical

5 Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 2$ according to the current invention. Also the general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention. (And in the example of FIG. 1F, $s = 2$). The general asymmetrical Butterfly fat tree network

10 $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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 $d + d_1 (= d + p \times d)$ 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,($d+p*d$)) 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

15 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).

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

20 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.

WO 2008/147926

PCT/US2008/064603

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 from exactly d switches in middle stage 130 through d links.

Similarly 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 also are connected to exactly d switches in middle stage 140 through d links and also are connected to exactly d output switches in output stage 120 through d links.

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 $130 + 10 * (\log_d N_2 - 2)$ are connected from exactly d switches in middle stage $130 + 10 * (\log_d N_2 - 3)$ through d links and also are connected to exactly d switches in middle stage $130 + 10 * (\log_d N_2 - 1)$ through d links.

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 $130 + 10 * (2 * \log_d N_2 - 4)$ through $2 \times d$ links.

As described before, again the connection topology of a general $V_{bfi}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bfi}(N_1, N_2, d, s)$ 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 $V_{bfi}(N_1, N_2, d, s)$ 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 $V_{bfi}(N_1, N_2, d, s)$ can be built. The embodiments of FIG.

WO 2008/147926

PCT/US2008/064603

1E, FIG. 1E1, and FIG. 1E2 are three examples of network $V_{bft}(N_1, N_2, d, s)$ for $s = 2$ and $N_1 > N_2$.

The general symmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical Butterfly fat tree network
 5 $V_{bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

For example, the network of Fig. 1E shows an exemplary three-stage network, namely $V_{bft}(24, 8, 2, 2)$, with the following multicast assignment $I_1 = \{2, 3\}$ and all other
 10 $I_j = \phi$ 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 output switch OS2 in output stage 120 and middle switch and MS(2,7) in middle stage 140 respectively.

The connection I_1 also fans out in middle switch MS(2,7) only once into middle
 15 switch MS(1,7) in middle stage 130. The connection I_1 also fans out in middle switch MS(1,7) only once into output switch 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
 20 switches in middle stage 130.

Strictly Nonblocking Butterfly Fat Tree Networks:

The general symmetric Butterfly fat tree network $V_{bft}(N, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention. Similarly the general asymmetric Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$
 25 can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention.

WO 2008/147926

PCT/US2008/064603

Symmetric RNB Unicast Embodiments:

Referring to FIG. 2A, in one embodiment, an exemplary symmetrical Butterfly fat tree network 200A with three stages of sixteen switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection between

5 configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130, and 140 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. Input stage 110 and output stage 120 together belong to leaf stage. And all the middle stages

10 excepting root stage namely middle stage 130 consists of four, four by four switches MS(1,1) - MS(1,4), and root stage i.e., middle stage 140 consists of four, two by two switches MS(2,1) - MS(2,4).

Such a network can be operated in rearrangeably non-blocking 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

15 middle stage 130 and middle stage 140.

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 number of inlet links or outlet links. The number of middle switches in each middle stage

20 is denoted by $\frac{N}{d}$. 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 $2d * 2d$ excepting that the size of each switch in middle stage 140 is denoted as $d * d$. (In another embodiment, the size of each switch in any of the middle stages

25 other than the middle stage 140, can be implemented as $d * 2d$ and $d * d$ since the down coming middle links are never setup to the up going middle links. For example in network 200A of FIG. 2A, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1).

WO 2008/147926

PCT/US2008/064603

So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

5 Middle stage 140 is called as root stage. 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 Butterfly fat tree network can be represented with the notation $V_{bft}(N, d, s)$, 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
10 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.

Each of the $\frac{N}{d}$ input switches IS1 – IS4 are connected to exactly d switches in
15 middle stage 130 through $2 \times d$ links (for example input switch IS1 is connected to middle switch MS(1,1) through the link ML(1,1); and input switch IS1 is also connected to middle switch MS(1,2) through the link ML(1,2)).

Each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,4) in the middle stage 130 are
20 connected from exactly d input switches through d links (for example the link ML(1,1) is connected to the middle switch MS(1,1) from input switch IS1; and the link ML(1,3) is connected to the middle switch MS(1,1) from input switch IS2) and are also connected from exactly d switches in middle stage 140 through d links (for example the link ML(3,2) is connected to the middle switch MS(1,1) from middle switch MS(2,1) and also the link ML(3,5) is connected to the middle switch MS(1,1) from middle switch
25 MS(2,3)).

WO 2008/147926

PCT/US2008/064603

Each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,4) in the middle stage 130 are connected to exactly d switches in middle stage 140 through d links (for example the link ML(2,1) is connected from middle switch MS(1,1) to middle switch MS(2,1), and the link ML(2,2) is connected from middle switch MS(1,1) to middle switch MS(2,3)) and also are connected to exactly d output switches in output stage 120 through d links (for example the link ML(4,1) is connected to output switch OS1 from middle switch MS(1,1), and the link ML(4,2) is connected to output switch OS2 from middle switch MS(1,1)).

Similarly each of the $\frac{N}{d}$ 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 link ML(2,1) is connected to the middle switch MS(2,1) from middle switch MS(1,1), and the link ML(2,5) is 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 130 through d links (for example the link ML(3,1) is connected from middle switch MS(2,1) to middle switch MS(1,3); and the link ML(3,2) is connected from middle switch MS(2,1) to middle switch MS(1,1)).

Each of the $\frac{N}{d}$ output switches OS1 – OS4 are connected from exactly d switches in middle stage 130 through d links (for example output switch OS1 is connected from middle switch MS(1,1) through the link ML(4,1); and output switch OS1 is also connected from middle switch MS(1,2) through the link ML(4,2)).

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

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(1,8), 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. Even though only one embodiment is illustrated, in general, the network

WO 2008/147926

PCT/US2008/064603

$V_{bft}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bft}(N, d, s)$ 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 $V_{bft}(N, d, s)$ network is, when no

5 connections are setup from any input link all the output links should be reachable. Based on this property numerous embodiments of the network $V_{bft}(N, d, s)$ can be built. The embodiment of FIG. 2A is only one example of network $V_{bft}(N, d, s)$.

In the embodiment of FIG. 2A each of 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 either available for use by a

10 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

15 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.

Generalized Symmetric RNB Unicast Embodiments:

Network 200B of FIG. 2B is an example of general symmetrical Butterfly fat tree

20 network $V_{bft}(N, d, s)$ with $(\log_d N)$ stages. The general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s = 1$ according to the current invention (and in the example of FIG. 2B, $s = 1$). The general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ with $(\log_d N)$ stages has d inlet links for each of $\frac{N}{d}$ input switches IS1-IS(N/d) (for example the links IL1-IL(d) to

25 the input switch IS1) and d outgoing links for each of $\frac{N}{d}$ 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

WO 2008/147926

PCT/US2008/064603

links for each of $\frac{N}{d}$ output switches OS1-OS(N/d) (for example the links OL1-OL(d) to the output switch OS1) and d incoming links for each of $\frac{N}{d}$ output switches OS1-OS(N/d) (for example $ML(2 \times \text{Log}_d N - 2, 1) - ML(2 \times \text{Log}_d N - 2, d)$ to the output switch OS1).

5 Each of the $\frac{N}{d}$ input switches IS1 – IS(N/d) are connected to exactly d switches in middle stage 130 through d links.

Each of the $\frac{N}{d}$ 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.

10 Similarly each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,N/d) in the middle stage 130 are also connected from exactly d switches in middle stage 140 through d links and also are connected to exactly d output switches in output stage 120 through d links.

Similarly each of the $\frac{N}{d}$ middle switches $MS(\text{Log}_d N - 1, 1) - MS(\text{Log}_d N - 1, \frac{N}{d})$
 15 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.

Each of the $\frac{N}{d}$ output switches OS1 – OS(N/d) are connected from exactly d switches in middle stage 130 through d links.

20 As described before, again the connection topology of a general $V_{bfi}(N, d, s)$ may be any one of the connection topologies. For example the connection topology of the

WO 2008/147926

PCT/US2008/064603

network $V_{bft}(N, d, s)$ 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 $V_{bft}(N, d, s)$ network is, when no connections are setup from any input link if any

5 output link should be reachable. Based on this property numerous embodiments of the network $V_{bft}(N, d, s)$ can be built. The embodiment of FIG. 2A are one example of network $V_{bft}(N, d, s)$.

The general symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ is operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current

10 invention.

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

Referring to FIG. 2C, in one embodiment, an exemplary asymmetrical Butterfly fat tree network 200C with three stages of sixteen switches for satisfying communication

15 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 and 140 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. Middle stage 130 consists of four, four by eight switches MS(1,1) - MS(1,4) and middle stage 140

20 consists of four, two by two switches MS(2,1) - MS(2,4).

Such a network can be operated in rearrangeably non-blocking 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 of size four by eight in middle stage 130 and four switches of size two by two in middle stage 140.

25 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

WO 2008/147926

PCT/US2008/064603

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 middle stage 130 can be denoted as $2d * (d + d_2)$. The size of each switch in the root stage (i.e., middle stage 140) can be denoted as $d * d$. The size of each switch in all the middle stages excepting middle stage 130 and root stage can be denoted as $2d * 2d$ (In network 200C of FIG. 2C, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $d * 2d$ and $d * d$ since the down coming middle links are never setup to the up going middle links. For example in network 200C of FIG. 2C, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

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 Butterfly fat tree network can be represented with the notation $V_{bft}(N_1, N_2, d, s)$, 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.

WO 2008/147926

PCT/US2008/064603

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 switch MS(1,1) through the link ML(1,1), and input switch IS1 is also connected to MS(1,2) through the link ML(1,2)).

- 5 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 link ML(1,1) is connected to the middle switch MS(1,1) from input switch IS1 and the link ML(1,3) is connected to the middle switch MS(1,1) from input switch IS2) and are also connected from exactly d switches in middle stage 140 through d links (for example the link
10 ML(3,2) is connected to the middle switch MS(1,1) from middle switch MS(2,1), and the link ML(3,5) is connected to the middle switch MS(1,1) from middle switch MS(2,3)).

- Similarly each of the $\frac{N_1}{d}$ middle switches MS(1,1) – MS(1,4) in the middle stage 130 are connected to exactly d switches in middle stage 140 through d links (for example the link ML(2,1) is connected from middle switch MS(1,1) to middle switch
15 MS(2,1), and the link ML(2,2) is connected from middle switch MS(1,1) to middle switch MS(2,3)), and also are connected to exactly $\frac{d_2}{2}$ output switches in output stage 120 through d_2 links (for example the link ML(4,1) and ML(4,2) are connected from middle switch MS(1,1) to output switch OS1; the links ML(4,3) and ML(4,4) are connected from middle switch MS(1,1) to output switch OS2; the link ML(4,5) is
20 connected from middle switch MS(1,1) to output switch OS3; and the links ML(4,6) is connected from middle switch MS(1,1) to output switch OS4).

- 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 link ML(2,1) is connected to the middle switch MS(2,1) from middle switch
25 MS(1,1); and the link ML(2,5) is connected to the middle switch MS(2,1) from middle

WO 2008/147926

PCT/US2008/064603

switch MS(1,3)) and also are connected to exactly d switches in middle stage 130 through d links (for example the link ML(3,2) is connected from middle switch MS(2,1) to middle switch MS(1,1); and the link ML(3,1) is connected from middle switch MS(2,1) to middle switch MS(1,3)).

5 Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $\frac{d_2}{2}$ switches in middle stage 130 through d_2 links (for example output switch OS1 is connected from middle switch MS(1,1) through the links ML(4,1) and ML(4,2); output switch OS1 is connected from middle switch MS(1,2) through the links ML(4,7) and ML(4,8); output switch OS1 is connected from middle switch MS(1,3) through the link
10 ML(4,13); output switch OS1 is connected from middle switch MS(1,4) through the link ML(4,19)).

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

In other embodiments the connection topology may be different from the
15 embodiment of the network 200C of FIG. 2C. 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(4,24) are connected between the respective stages is different. Even though only one embodiment is illustrated, in general, the network $V_{bft}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network
20 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 2C,
25 are only one example of network $V_{bft}(N_1, N_2, d, s)$.

In the embodiment of FIG. 2C, each of 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(4,24) are either available for use by a

WO 2008/147926

PCT/US2008/064603

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(1,2) – MS(2,4) are referred to as root stage switches.

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

Network 200D of FIG. 2D is an example of general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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 Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s = 1$ according to the current invention (and in the example of FIG. 2D, $s = 1$). The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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 ($= p \times d$) incoming links for each of $\frac{N_1}{d}$ output switches OS1-OS(N_1/d) (for example $ML(2 \times \log_d N_1 - 2, 1)$ - $ML(2 \times \log_d N_1 - 2, d_2)$ to the output switch OS1).

WO 2008/147926

PCT/US2008/064603

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.

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.

Similarly 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 switches in middle stage 140 through d links and also are connected to exactly $\frac{d_2}{2}$ output switches in output stage 120 through d_2 links.

10 Similarly each of the $\frac{N_1}{d}$ middle switches $MS(\text{Log}_d N_1 - 1, 1)$ - $MS(\text{Log}_d N_1 - 1, \frac{N_1}{d})$ 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.

15 Each of the $\frac{N_1}{d}$ output switches OS1 – OS(N_1/d) are connected from exactly $\frac{d_2}{2}$ switches in middle stage 130 through d_2 links.

As described before, again the connection topology of a general $V_{bfr}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bfr}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back
 20 Omega networks, back to back Benes networks, Delta Networks and many more combinations. The applicant notes that the fundamental property of a valid connection

WO 2008/147926

PCT/US2008/064603

topology of the general $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 2C is one example of network $V_{bft}(N_1, N_2, d, s)$ for $s = 1$ and $N_2 > N_1$.

- 5 The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention.

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

Referring to FIG. 2E, in one embodiment, an exemplary asymmetrical Butterfly fat tree network 200E with three stages of sixteen 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 and 140 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. Middle stage 130 consists of four, eight by four switches MS(1,1) - MS(1,4) and middle stage 140 consists of four, two by two switches MS(2,1) - MS(2,4).

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 six, the switches in output stage 120 are of size two by two, and there are four switches in each of middle stage 130 and middle stage 140. Such a network can be operated in rearrangeably non-blocking 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 of size eight by four in middle stage 130, and four switches of size two by two in middle stage 140.

25 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

WO 2008/147926

PCT/US2008/064603

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 middle stage 130 can be denoted as $(d + d_1) * 2d$. The size of each switch in the root stage (i.e., middle stage 140) can be denoted as $d * d$. The size of each switch in all the middle stages excepting middle stage 130 and root stage can be denoted as $2d * 2d$ (In network 200E of FIG. 2E, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $d * 2d$ and $d * d$ since the down coming middle links are never setup to the up going middle links. For example in network 200E of FIG. 2E, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

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 Butterfly fat tree network can be represented with the notation $V_{bf}(N_1, N_2, d, s)$, 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.

WO 2008/147926

PCT/US2008/064603

Each of the $\frac{N_2}{d}$ input switches IS1 – IS4 are connected to exactly $\frac{d_1}{2}$ 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); input switch IS1 is connected to middle switch MS(1,4) through the link ML(1,6)).

Each of the $\frac{N_2}{d}$ middle switches MS(1,1) – MS(1,4) in the middle stage 130 are connected from exactly $\frac{d_1}{2}$ input switches through d_1 links (for example the links ML(1,1) and ML(1,2) are connected from input switch IS1 to middle switch MS(1,1); the links ML(1,7) and ML(1,8) are connected from input switch IS2 to middle switch MS(1,1); the link ML(1,13) is connected from input switch IS3 to middle switch MS(1,1); the link ML(1,19) is connected from input switch IS4 to middle switch MS(1,1)), and also are connected from exactly d switches in middle stage 140 through d links (for example the link ML(3,2) is connected to the middle switch MS(1,1) from middle switch MS(2,1); and the link ML(3,5) is connected to the middle switch MS(1,1) from middle switch MS(2,3)).

Similarly each of the $\frac{N_2}{d}$ middle switches MS(1,1) – MS(1,4) in the middle stage 130 are connected to exactly d switches in middle stage 140 through d links (for example the link ML(2,1) is connected from middle switch MS(1,1) to middle switch MS(2,1) and the link ML(2,2) is connected from middle switch MS(1,1) to middle switch MS(2,3)), and also are connected to exactly d output switches in output stage 120 through d links (for example the link ML(4,1) is connected to output switch OS1 from middle switch MS(1,1) and the link ML(4,2) is connected to output switch OS2 from middle switch MS(1,1)).

WO 2008/147926

PCT/US2008/064603

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 d links (for example the link ML(2,1) is connected to the middle switch MS(2,1) from middle switch MS(1,1) and the link ML(2,5) is 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 130 through d links (for example the link ML(3,2) is connected from middle switch MS(2,1) to middle switch MS(1,1) and the link ML(3,1) is connected from middle switch MS(2,1) to middle switch MS(1,3)).

Each of the $\frac{N_2}{d}$ output switches OS1 – OS4 are connected from exactly d switches in middle stage 130 through d links (for example output switch OS1 is connected from middle switch MS(1,1) through the link ML(4,1), and output switch OS1 is connected from middle switch MS(1,2) through the link ML(4,3)).

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

In other embodiments the connection topology may be different from the embodiment of the network 200E of FIG. 2E. That is the way the links ML(1,1) - ML(1,24), 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. Even though only one embodiment is illustrated, in general, the network $V_{bft}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 2E is only one example of network $V_{bft}(N_1, N_2, d, s)$.

WO 2008/147926

PCT/US2008/064603

In the embodiment of FIG. 2E, each of the links ML(1,1) – ML(1,24), ML(2,1) – ML(2,8), ML(3,1) – ML(3,8) and ML(4,1) – ML(4,8) 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.

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

Network 200F of FIG. 2F is an example of general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network 200F of FIG. 2F, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s = 1$ according to the current invention. (And in the example of FIG. 2F, $s = 1$). The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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 $d_1 (= p \times d)$ 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,($d+p*d$)) 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).

WO 2008/147926

PCT/US2008/064603

Each of the $\frac{N_2}{d}$ input switches IS1 – IS(N_2/d) are connected to exactly $\frac{d_1}{2}$

switches in middle stage 130 through d_1 links.

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 $\frac{d_1}{2}$ input switches through d_1 links and also are connected

5 from exactly d switches in middle stage 140 through d links.

Similarly each of the $\frac{N_2}{d}$ middle switches MS(1,1) – MS(1, $2N_2/d$) in the middle

stage 130 also are connected to exactly d switches in middle stage 140 through d links and also are connected to exactly d output switches in output stage 120 through d links.

Similarly each of the $\frac{N_2}{d}$ middle switches $MS(\text{Log}_d N_2 - 1, 1) -$

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

Each of the $\frac{N_2}{d}$ output switches OS1 – OS(N_2/d) are connected from exactly d

15 switches in middle stage $130 + 10 * (2 * \text{Log}_d N_2 - 4)$ through d links.

As described before, again the connection topology of a general $V_{bft}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ may be back to back inverse Benes networks, back to back Omega networks, back to back Benes networks, Delta Networks and many more

20 combinations. The applicant notes that the fundamental property of a valid connection topology of the general $V_{bft}(N_1, N_2, d, s)$ network is, when no connections are setup from

WO 2008/147926

PCT/US2008/064603

any input link if any output link should be reachable. Based on this property numerous embodiments of the network $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 2E is one example of network $V_{bft}(N_1, N_2, d, s)$ for $s = 1$ and $N_1 > N_2$.

The general symmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention.

MULTI-LINK BUTTERFLY FAT TREE EMBODIMENTS:

Symmetric RNB Embodiments:

Referring to FIG. 3A, in one embodiment, an exemplary symmetrical Multi-link Butterfly fat tree network 300A with three stages of sixteen 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, and 140 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. Input stage 110 and output stage 120 together belong to leaf stage. And all the middle stages excepting root stage namely middle stage 130 consists of four, eight by eight switches MS(1,1) - MS(1,4), and root stage i.e., middle stage 140 consists of four, four by four switches MS(2,1) - MS(2,4).

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 and middle stage 140. 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 and middle stage 140.

WO 2008/147926

PCT/US2008/064603

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 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 $4d * 4d$ excepting that the size of each switch in middle stage 140 is denoted as $2d * 2d$. (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $2d * 4d$ and $2d * 2d$ since the down coming middle links are never setup to the up going middle links. For example in network 300A of FIG. 3A, the down coming middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) are never setup to the up going middle links ML(2,1), ML(2,2), ML(2,3) and ML(2,4) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a four by eight switch with middle links ML(1,1), ML(1,2), ML(1,5) and ML(1,6) as inputs and middle links ML(2,1), ML(2,2), ML(2,3), ML(2,4), ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs; and a four by four switch with middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) as inputs and middle links ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs).

Middle stage 140 is called as root stage. 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 Butterfly fat tree network can be represented with the notation $V_{mlink-bft}(N, d, s)$, 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.

WO 2008/147926

PCT/US2008/064603

Each of the $\frac{N}{d}$ 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) and ML(1,2); and input switch IS1 is also connected to middle switch MS(1,2) through the links ML(1,3) and ML(1,4)).

5 Each of the $\frac{N}{d}$ 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(1,5) and ML(1,6) are connected to the middle switch MS(1,1) from input switch IS2) and are also connected from exactly d switches in middle stage 140
10 through $2 \times d$ links (for example the links ML(3,3) and ML(3,4) are connected to the middle switch MS(1,1) from middle switch MS(2,1) and also the links ML(3,9) and ML(3,10) are connected to the middle switch MS(1,1) from middle switch MS(2,3)).

Each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,4) in the middle stage 130 are connected to exactly d switches in middle stage 140 through $2 \times d$ links (for example
15 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)) 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(1,1), and the links ML(4,3) and
20 ML(4,4) are connected to output switch OS2 from middle switch MS(1,1)).

Similarly each of the $\frac{N}{d}$ 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(2,9) and ML(2,10) are connected to the middle
25 switch MS(2,1) from middle switch MS(1,3)), and also are connected to exactly d switches in middle stage 130 through $2 \times d$ links (for example the links ML(3,1) and

WO 2008/147926

PCT/US2008/064603

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

Each of the $\frac{N}{d}$ output switches OS1 – OS4 are connected from exactly d

- 5 switches in middle stage 130 through $2 \times d$ links (for example output switch OS1 is connected from middle switch MS(1,1) through the links ML(4,1), ML(4,2); and output switch OS1 is also connected from middle switch MS(1,2) through the links ML(4,5) and ML(4,6)).

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

In other embodiments the connection topology may be different from the network 300A of FIG. 3A. 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

15 network $V_{mlink-bft}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mlink-bft}(N, d, s)$ 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 $V_{mlink-bft}(N, d, s)$ network is, when no connections are setup from any input link all the output links should be

20 reachable. Based on this property numerous embodiments of the network $V_{mlink-bft}(N, d, s)$ can be built. The embodiment of FIG. 3A is only one example of network $V_{mlink-bft}(N, d, s)$.

In the embodiment of FIG. 3A 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

25 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

WO 2008/147926

PCT/US2008/064603

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.

5 In the example illustrated in FIG. 3A, 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
10 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, to be operated in rearrangeably nonblocking manner in accordance with the invention.

The connection request of the type described above can be unicast connection
15 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
20 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:

25 Network 300B of FIG. 3B is an example of general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N, d, s)$ with $(\log_d N)$ stages. The general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 2$ according to the current invention. Also the general symmetrical Multi-link Butterfly fat tree network

WO 2008/147926

PCT/US2008/064603

$V_{mlink-bft}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention. (And in the example of FIG. 3B, $s = 2$). The general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N, d, s)$ with $(\log_d N)$ stages has d inlet links for each of $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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).

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

Each of the $\frac{N}{d}$ 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.

Similarly each of the $\frac{N}{d}$ middle switches MS(1,1) - MS(1,N/d) in the middle stage 130 are also 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.

Similarly each of the $\frac{N}{d}$ 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

WO 2008/147926

PCT/US2008/064603

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

Each of the $\frac{N}{d}$ output switches OS1 – OS(N/d) are connected from exactly d switches in middle stage 130 through $2 \times d$ links.

5 As described before, again the connection topology of a general $V_{mlink-bft}(N, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink-bft}(N, d, s)$ 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
10 topology of the general $V_{mlink-bft}(N, d, s)$ 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 $V_{mlink-bft}(N, d, s)$ can be built. The embodiment of FIG. 3A are one example of network $V_{mlink-bft}(N, d, s)$.

The general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N, d, s)$
15 can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

Every switch in the Multi-link Butterfly fat tree networks discussed herein has
20 multicast capability. In a $V_{mlink-bft}(N, d, s)$ 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
25 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

WO 2008/147926

PCT/US2008/064603

- 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.
- 5 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.

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

- 10 is to be connected in the multicast assignment. For example, the network of FIG. 3A shows an exemplary three-stage network, namely $V_{\text{mlink-bft}}(8,2,2)$, 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(1,2) in middle stage 130, and fans out in middle switches MS(1,1) and MS(1,2) only once into output switch OS2 in output stage 120 and middle switch MS(2,2) in middle stage 140 respectively.

- The connection I_1 also fans out in middle switch MS(2,2) only once into middle switches MS(1,4) in middle stage 130. The connection I_1 also fans out in middle switch MS(1,4) only once into output switch 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.

25

WO 2008/147926

PCT/US2008/064603

Asymmetric RNB ($N_2 > N_1$) Embodiments:

Referring to FIG. 3C, in one embodiment, an exemplary asymmetrical Multi-link Butterfly fat tree network 300C with three stages of sixteen switches for satisfying communication requests, such as setting up a telephone call or a data call, or a connection
 5 between configurable logic blocks, between an input stage 110 and output stage 120 via middle stages 130 and 140 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. Middle stage 130 consists of four, eight by twelve switches MS(1,1) - MS(1,4) and middle stage 140 consists of four, four by four switches MS(2,1) - MS(2,4).

10 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 four switches in each of middle stage 130 and middle stage 140. Such a network can be operated in rearrangeably non-blocking manner for multicast connections, because the switches in the
 15 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 of size eight by twelve in middle stage 130 and four switches of size four by four in middle stage 140.

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
 20 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
 25 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 middle stage 130 can be denoted as $4d * 2(d + d_2)$. The size of each switch in the root stage (i.e., middle stage 140) can be denoted as $2d * 2d$. The size of each switch in all the

WO 2008/147926

PCT/US2008/064603

middle stages excepting middle stage 130 and root stage can be denoted as $4d * 4d$ (In network 300C of FIG. 3C, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $2d * 4d$ and $2d * 2d$ since the down coming middle links are never setup

5 to the up going middle links. For example in network 300C of FIG. 3C, the down coming middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) are never setup to the up going middle links ML(2,1), ML(2,2), ML(2,3) and ML(2,4) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a four by eight switch with middle links ML(1,1), ML(1,2), ML(1,5) and ML(1,6) as inputs and middle links ML(2,1), ML(2,2),

10 ML(2,3), ML(2,4), ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs; and a four by four switch with middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) as inputs and middle links ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs).

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

15 Multi-link Butterfly fat tree network can be represented with the notation $V_{mink-bft}(N_1, N_2, d, s)$, 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

20 each input switch to the inlet links of each input switch.

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) and ML(1,2), and input switch IS1 is also connected to MS(1,2) through the links ML(1,3) and ML(1,4)).

25 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

WO 2008/147926

PCT/US2008/064603

and the links ML(1,5) and ML(1,6) are connected to the middle switch MS(1,1) from input switch IS2) and are also connected from exactly d switches in middle stage 140 through $2 \times d$ links (for example the links ML(3,3) and ML(3,4) are connected to the middle switch MS(1,1) from middle switch MS(2,1), and the links ML(3,9) and
 5 ML(3,10) are connected to the middle switch MS(1,1) from middle switch MS(2,3)).

Similarly each of the $\frac{N_1}{d}$ middle switches MS(1,1) – MS(1,4) in the middle stage 130 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
 10 switch MS(1,1) to middle switch MS(2,3)), and also are connected to exactly $\frac{d_2}{2}$ output switches in output stage 120 through d_2 links (for example the links ML(4,1) and ML(4,2) are connected from middle switch MS(1,1) to output switch OS1; the links ML(4,3) and ML(4,4) are connected from middle switch MS(1,1) to output switch OS2; the links ML(4,4) and ML(4,6) are connected from middle switch MS(1,1) to output
 15 switch OS3; and the links ML(4,7) and ML(4,8) are connected from middle switch MS(1,1) to output switch OS4).

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
 20 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 130 through $2 \times d$ links (for example the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(1,1); and the links ML(3,1) and ML(3,2) are connected from middle switch MS(2,1) to middle switch
 25 MS(1,3)).

WO 2008/147926

PCT/US2008/064603

Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $\frac{d_2}{2}$

switches in middle stage 130 through d_2 links (for example output switch OS1 is connected from middle switch MS(1,1) through the links ML(4,1) and ML(4,2); output switch OS1 is connected from middle switch MS(1,2) through the links ML(4,9) and
 5 ML(4,10); output switch OS1 is connected from middle switch MS(1,3) through the links ML(4,17) and ML(4,18); output switch OS1 is connected from middle switch MS(1,4) through the links ML(4,25) and ML(4,26)).

Finally the connection topology of the network 300C shown in FIG. 3C is known to be back to back inverse Benes connection topology.

10 In other embodiments the connection topology may be different from the embodiment of the network 300C of FIG. 3C. 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(4,32) are connected between the respective stages is different. Even though only one embodiment is illustrated, in general, the network $V_{mlink-bft}(N_1, N_2, d, s)$ can comprise any arbitrary
 15 type of connection topology. For example the connection topology of the network $V_{mlink-bft}(N_1, N_2, d, s)$ 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 $V_{mlink-bft}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property
 20 numerous embodiments of the network $V_{mlink-bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 3C, are only one example of network $V_{mlink-bft}(N_1, N_2, d, s)$.

In the embodiment of FIG. 3C, 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(4,32) are either available for use by a new connection or not available if currently used by an existing connection. The input
 25 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

WO 2008/147926

PCT/US2008/064603

middle switches or middle ports. The middle stage 130 is also referred to as root stage and middle stage switches MS(1,2) – MS(2,4) are referred to as root stage switches.

In the example illustrated in FIG. 3C, 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 300C, to be operated in rearrangeably nonblocking manner in accordance with the invention.

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:

Network 300D of FIG. 3D is an example of general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_2 > N_1$ and $N_2 = p * N_1$ where $p > 1$. In network 300D of FIG. 3D, $N_1 = N$ and $N_2 = p * N$. The general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 2$ according to the current invention. Also the general asymmetrical Multi-link Butterfly fat tree network

WO 2008/147926

PCT/US2008/064603

$V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention. (And in the example of FIG. 3D, $s = 2$). The general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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).

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 $2 \times d$ links.

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.

Similarly 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 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.

WO 2008/147926

PCT/US2008/064603

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(\text{Log}_d N_1 - 1, 1)$ -
 $MS(\text{Log}_d N_1 - 1, \frac{N_1}{d})$ 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 $2 \times d$ links and also
 are connected to exactly d switches in middle stage $130 + 10 * (\text{Log}_d N_1 - 1)$ through
 5 $2 \times d$ links.

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 through $d + d_2$ links.

As described before, again the connection topology of a general
 $V_{m\text{link-bft}}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the
 10 connection topology of the network $V_{m\text{link-bft}}(N_1, N_2, d, s)$ 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 $V_{m\text{link-bft}}(N_1, N_2, d, s)$ network is,
 when no connections are setup from any input link if any output link should be reachable.
 15 Based on this property numerous embodiments of the network $V_{m\text{link-bft}}(N_1, N_2, d, s)$ can
 be built. The embodiment of FIG. 3C is one example of network $V_{m\text{link-bft}}(N_1, N_2, d, s)$
 for $s = 2$ and $N_2 > N_1$.

The general asymmetrical Multi-link Butterfly fat tree network
 $V_{m\text{link-bft}}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast
 20 when $s \geq 2$ according to the current invention. Also the general symmetrical Multi-link
 Butterfly fat tree network $V_{m\text{link-bft}}(N_1, N_2, d, s)$ can be operated in strictly nonblocking
 manner for unicast if $s \geq 2$ according to the current invention.

WO 2008/147926

PCT/US2008/064603

For example, the network of FIG. 3C shows an exemplary three-stage network, namely $V_{\text{mlink-bft}}(8,24,2,2)$, 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(1,2) in middle stage 130, and fans out in middle switches MS(1,1) and MS(1,2) only once into output switch OS2 in output stage 120 and middle switch MS(2,2) in middle stage 140.

The connection I_1 also fans out in middle switch MS(2,2) only once into middle switches MS(1,4) in middle stage 130. The connection I_1 also fans out in middle switch MS(1,4) only once into output switch 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 OL18. 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:

Referring to FIG. 3E, in one embodiment, an exemplary asymmetrical Multi-link Butterfly fat tree network 300E with three stages of sixteen 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 and 140 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. Middle stage 130 consists of four, twelve by eight switches MS(1,1) - MS(1,4) and middle stage 140 consists of four, four by four switches MS(2,1) - MS(2,4).

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 and middle stage 140. 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

WO 2008/147926

PCT/US2008/064603

by two, and there are four switches of size twelve by eight in middle stage 130, and four switches of size four by four in middle stage 140.

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

5 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

10 with the notation $(2d * d)$, where $d_1 = N_1 \times \frac{d}{N_2} = p \times d$. The size of each switch in middle stage 130 can be denoted as $2(d + d_1) * 4d$. The size of each switch in the root stage (i.e., middle stage 140) can be denoted as $2d * 2d$. The size of each switch in all the middle stages excepting middle stage 130 and root stage can be denoted as $4d * 4d$ (In network 300C of FIG. 3C, there is no such middle stage). (In another embodiment, the

15 size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $2d * 4d$ and $2d * 2d$ since the down coming middle links are never setup to the up going middle links. For example in network 300E of FIG. 3E, the down coming middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) are never setup to the up going middle links ML(2,1), ML(2,2), ML(2,3) and ML(2,4) for the middle switch MS(1,1). So

20 middle switch MS(1,1) can be implemented as a four by eight switch with middle links ML(1,1), ML(1,2), ML(1,5) and ML(1,6) as inputs and middle links ML(2,1), ML(2,2), ML(2,3), ML(2,4), ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs; and a four by four switch with middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) as inputs and middle links ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs).

25 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 Butterfly fat tree network can be represented with the notation

WO 2008/147926

PCT/US2008/064603

$V_{mink-bft}(N_1, N_2, d, s)$, 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.

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) 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 links ML(1,5) and ML(1,6); input switch IS1 is connected to middle switch MS(1,4) through the links ML(1,7) and ML(1,8)).

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

Similarly each of the $\frac{N_2}{d}$ middle switches MS(1,1) – MS(1,4) in the middle stage 130 are connected to exactly d switches in middle stage 140 through $2d$ links (for example the links ML(2,1) and ML(2,2) are connected from middle switch MS(1,1) to

WO 2008/147926

PCT/US2008/064603

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

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

15 Each of the $\frac{N_2}{d}$ output switches OS1 – OS4 are connected from exactly d switches in middle stage 130 through $2 \times d$ links (for example output switch OS1 is connected from middle switch MS(1,1) through the links ML(4,1) and ML(4,2), and output switch OS1 is connected from middle switch MS(1,2) through the links ML(4,5) and ML(4,6).

20 Finally the connection topology of the network 300E shown in FIG. 3E is known to be back to back inverse Benes connection topology.

In other embodiments the connection topology may be different from the embodiment of the network 300E of FIG. 3E. 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
25 connected between the respective stages is different. Even though only one embodiment is illustrated, in general, the network $V_{\text{mlink-bft}}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network

WO 2008/147926

PCT/US2008/064603

$V_{mlink-bft}(N_1, N_2, d, s)$ 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 $V_{mlink-bft}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link all the output links should be reachable. Based on this property
5 numerous embodiments of the network $V_{mlink-bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 3E is only one example of network $V_{mlink-bft}(N_1, N_2, d, s)$.

In the embodiment of FIG. 3E, 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
10 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.

In the example illustrated in FIG. 3E, a fan-out of four is possible to satisfy a
15 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
20 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 300E, to be operated in rearrangeably nonblocking manner in accordance with the invention.

The connection request of the type described above can be unicast connection
25 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

WO 2008/147926

PCT/US2008/064603

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.

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

Network 300F of FIG. 3F is an example of general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network 300F of FIG. 3F, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 2$ according to the current invention. Also the general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s = 2$ according to the current invention. (And in the example of FIG. 3F, $s = 2$). The general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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 $d + d_1 (= d + p \times d)$ 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, ($d+p*d$)) 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).

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.

WO 2008/147926

PCT/US2008/064603

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 $\frac{d+d_1}{2}$ input switches through $d+d_1$ links and also are connected from exactly d switches in middle stage 140 through $2 \times d$ links.

Similarly each of the $\frac{N_2}{d}$ middle switches MS(1,1) – MS(1, $2N_2/d$) in the middle 5 stage 130 also are connected to 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.

Similarly each of the $\frac{N_2}{d}$ middle switches MS($\log_d N_2 - 1, 1$) - 10 MS($\log_d N_2 - 1, \frac{N_2}{d}$) in the middle stage $130 + 10 * (\log_d N_2 - 2)$ are connected from exactly d switches in middle stage $130 + 10 * (\log_d N_2 - 3)$ through $2 \times d$ links and also are connected to exactly d switches in middle stage $130 + 10 * (\log_d N_2 - 1)$ through $2 \times d$ links.

Each of the $\frac{N_2}{d}$ output switches OS1 – OS(N_2/d) are connected from exactly d switches in middle stage $130 + 10 * (2 * \log_d N_2 - 4)$ through $2 \times d$ links.

15 As described before, again the connection topology of a general $V_{mlink-bft}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink-bft}(N_1, N_2, d, s)$ 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 20 property of a valid connection topology of the general $V_{mlink-bft}(N_1, N_2, d, s)$ 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 $V_{mlink-bft}(N_1, N_2, d, s)$ can

WO 2008/147926

PCT/US2008/064603

be built. The embodiments of FIG. 3E is one example of network $V_{mlink-bft}(N_1, N_2, d, s)$ for $s = 2$ and $N_1 > N_2$.

The general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast
 5 when $s \geq 2$ according to the current invention. Also the general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

For example, the network of FIG. 3E shows an exemplary three-stage network, namely $V_{mlink-bft}(24, 8, 2, 2)$, with the following multicast assignment $I_1 = \{2, 3\}$ and all
 10 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 output switch OS2 in output stage 120 and middle switch and MS(2,2) in middle stage 140 respectively.

The connection I_1 also fans out in middle switch MS(2,2) only once into middle
 15 switch MS(1,4) in middle stage 130. The connection I_1 also fans out in middle switch MS(1,4) only once into output switch 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
 20 switches in middle stage 130.

Strictly Nonblocking Multi-link Butterfly Fat Tree Networks:

The general symmetric multi-link Butterfly fat tree network $V_{mlink-bft}(N, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention. Similarly the general asymmetric multi-link Butterfly fat tree network
 25 $V_{mlink-bft}(N_1, N_2, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention.

WO 2008/147926

PCT/US2008/064603

Scheduling Method Embodiments:

FIG. 4A shows a high-level flowchart of a scheduling method 1000, in one embodiment executed to setup multicast and unicast connections in network 400A of FIG. 4A (or any of the networks $V_{bft}(N_1, N_2, d, s)$ and $V_{mlink-bft}(N_1, N_2, d, s)$ 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.

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

WO 2008/147926

PCT/US2008/064603

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 5 130 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 10 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 $130 + 10 * (\log_d N - 2)$ from which the output switch is reachable, are derived. This process is repeated for each output switch of each destination link of the multicast 15 connection and separate lists in each middle stage from middle stage 130 to middle stage $130 + 10 * (\log_d N - 2)$ 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 $130 + 10 * (\log_d N - 2)$ corresponding to each 20 outgoing link of the input switch of the multicast connection, and the list of middle switches derived in middle stage $130 + 10 * (\log_d N - 2)$ 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 $130 + 10 * (\log_d N - 2)$ is reachable from an outgoing link of the input switch, say 25 "x", and also from the same middle switch in middle stage $130 + 10 * (\log_d N - 2)$ 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.

WO 2008/147926

PCT/US2008/064603

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 5 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 10 destination links reachable from that common middle switch. Then the control transfers to act 1090.

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 15 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 20 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 25 transfers to act 1090.

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 30 1090 are executed in a loop, for each connection request until the connections are set up.

WO 2008/147926

PCT/US2008/064603

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. 4A 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 $V_{bft}(N_1, N_2, d, s)$ and the network $V_{mlink-bft}(N_1, N_2, d, s)$ to be operated in nonblocking manner in accordance with the invention.

According to the current invention, using the method 1040 of FIG. 4A, the network $V_{bft}(N_1, N_2, d, s)$ and the network $V_{mlink-bft}(N_1, N_2, d, s)$ are operated in rearrangeably nonblocking for unicast connections when $s \geq 1$, are operated in strictly nonblocking for unicast connections when $s \geq 2$, are operated in rearrangeably nonblocking for multicast connections when $s \geq 2$, and are operated in strictly nonblocking for multicast connections when $s \geq 3$.

The connection request of the type described above in reference to method 1000 of FIG. 4A 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. 4A 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 1000 of FIG. 4A can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks

WO 2008/147926

PCT/US2008/064603

$V_{bft}(N, d, s)$, $V_{bft}(N_1, N_2, d, s)$, $V_{mlink-bft}(N, d, s)$, and $V_{mlink-bft}(N_1, N_2, d, s)$ disclosed in this invention.

Applications Embodiments:

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 be the implementation of middle switch MS(2,1) of the diagram 100A of FIG. 1A where inlet link IL1 of diagram 500A1 corresponds to middle link ML(2,1) of diagram 100A, inlet link IL2 of diagram 500A1 corresponds to middle link ML(2,5) of diagram 100A, outlet link OL1 of diagram 500A1 corresponds to middle link ML(3,1) of diagram 100A, outlet link OL2 of diagram 500A1 corresponds to middle link ML(3,2) of diagram 100A.

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).

WO 2008/147926

PCT/US2008/064603

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:

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.

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

WO 2008/147926

PCT/US2008/064603

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. 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.

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.

WO 2008/147926

PCT/US2008/064603

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.

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:

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.

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 2008/147926

PCT/US2008/064603

CLAIMS

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
 - 5 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
 - a leaf stage comprising an input stage and an output stage; and said 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 its
 - 10 immediate succeeding stage where $x > 0$; and said 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 its
 - immediate succeeding stage; and
 - a plurality of y middle stages, excepting a root stage, comprising $x \times \frac{N}{d}$ middle
 - 15 switches in each of said y middle stages wherein one of said middle stages is the immediate succeeding stage to both said input stage and said output stage, where $y > 1$, and said root stage comprising $\frac{N}{d}$ middle switches; and
 - each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, comprising d incoming links
 - 20 (hereinafter "incoming middle links") connecting from switches in its immediate preceding stage and d incoming links connecting from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links

WO 2008/147926

PCT/US2008/064603

(hereinafter “outgoing middle links”) connecting to switches in its immediate succeeding stage and d outgoing links connecting to switches in its immediate succeeding stage; and

each middle switch in said succeeding stage to both said input stage and said output stage comprising d incoming links connecting from switches in said input stage and d incoming links connecting from switches in its immediate succeeding stage, and
 5 each middle switch further comprising $\frac{(d + d_2)}{2}$ outgoing links connecting to switches in said output stage and d outgoing links connecting to switches in its immediate succeeding stage; and

each middle switch in said root stage comprising d incoming links connecting from switches in its immediate preceding stage and each middle switch further
 10 comprising d outgoing links connecting to switches in its immediate preceding stage; or

when $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$ then $N_2 = N$, $d_2 = d$ and

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

a leaf stage comprising an input stage and an output stage; said input stage
 15 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 its

immediate succeeding stage where $x > 0$; and said output stage comprising $\frac{N_2}{d}$ output

switches, and each output switch comprising d outlet links and each output switch further
 20 comprising $x \times d$ incoming links connecting from switches in its immediate succeeding stage; and

a plurality of y middle stages, excepting a root stage, comprising $x \times \frac{N}{d}$ middle
 switches in each of said y middle stages wherein one of said middle stages is the
 immediate succeeding stage to both said input stage and said output stage, where $y > 1$,

and said root stage comprising $\frac{N}{d}$ middle switches; and

WO 2008/147926

PCT/US2008/064603

each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, comprising d incoming links (hereinafter “incoming middle links”) connecting from switches in its immediate preceding stage and d incoming links connecting from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links (hereinafter “outgoing middle links”) connecting to switches in its immediate succeeding stage and d outgoing links connecting to switches in its immediate preceding stage; and

each middle switch in said succeeding stage to both said input stage and said output stage comprising $\frac{(d + d_1)}{2}$ incoming links connecting from switches in said input stage and d incoming links connecting from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links connecting to switches in said output stage and d outgoing links connecting to switches in its immediate preceding stage; and

each middle switch in said root stage comprising d incoming links connecting from switches in its immediate preceding stage and each middle switch further comprising d outgoing links connecting to switches in its immediate preceding 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 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 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 $y \geq (\log_d N_1) - 1$ when $N_2 > N_1$, and $y \geq (\log_d N_2) - 1$ when $N_1 > N_2$.

WO 2008/147926

PCT/US2008/064603

4. The network of claim 3, wherein $x \geq 1$, 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
5 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
10 network is hereinafter “rearrangeably nonblocking network for unicast”.
5. The network of claim 3, wherein $x \geq 2$, 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
15 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
20 nonblocking network for unicast”.
6. The network of claim 3, wherein $x \geq 2$,
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
25 said network is hereinafter “rearrangeably nonblocking network”.
7. The network of claim 3, wherein $x \geq 3$,
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”.

WO 2008/147926

PCT/US2008/064603

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 $d_1 = d_2 = d$.
- 5 10. The network of claim 1, wherein said input switches, said output switches and said middle switches are not fully 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.
- 10 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
$$d_2 = N_2 \times \frac{d}{N_1} = p \times d ; \text{ and}$$
having a leaf stage comprising an input stage and an output stage; and said input
15 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 its immediate succeeding stage where $x > 0$; and said 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 its immediate succeeding stage;
20 and
a plurality of y middle stages, excepting a root stage, having $x \times \frac{N}{d}$ middle switches in each of said y middle stages wherein one of said middle stages is the

WO 2008/147926

PCT/US2008/064603

immediate succeeding stage to both said input stage and said output stage, where $y > 1$,
 and said root stage having $\frac{N}{d}$ middle switches, and

each middle switch in all said middle stages, excepting said root stage and said
 succeeding stage to both said input stage said output stage, having d incoming links
 5 connected from switches in its immediate preceding stage and d incoming links
 connected from switches in its immediate succeeding stage, and each middle switch
 further comprising d outgoing links connected to switches in its immediate succeeding
 stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said succeeding stage to both said input stage and said
 10 output stage having d incoming links connected from switches in said input stage and d
 incoming links connected from switches in its immediate succeeding stage, and each
 middle switch further having $\frac{(d + d_2)}{2}$ outgoing links connected to switches in said output
 stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said root stage having d incoming links connected from
 15 switches in its immediate preceding stage and each middle switch further having d
 outgoing links connected to switches in its immediate preceding stage; or

when $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$ then $N_2 = N$, $d_2 = d$ and

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d ; \text{ and having}$$

having a leaf stage having an input stage and an output stage; and said input stage
 20 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 its immediate

succeeding stage where $x > 0$; and said 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 its immediate succeeding stage; and

WO 2008/147926

PCT/US2008/064603

- a plurality of y middle stages, excepting a root stage, having $x \times \frac{N}{d}$ middle switches in each of said y middle stages wherein one of said middle stages is the immediate succeeding stage to both said input stage and said output stage, where $y > 1$, and said root stage having $\frac{N}{d}$ middle switches, and
- 5 each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, having d incoming links connected from switches in its immediate preceding stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links connected to switches in its immediate succeeding
- 10 stage and d outgoing links connected to switches in its immediate succeeding stage; and each middle switch in said succeeding stage to both said input stage and said output stage having $\frac{(d + d_1)}{2}$ incoming links connected from switches in said input stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further having d outgoing links connected to switches in said output
- 15 stage and d outgoing links connected to switches in its immediate succeeding stage; and each middle switch in said root stage having d incoming links connected from switches in its immediate preceding stage and each middle switch further having d outgoing links connected to switches in its immediate preceding stage; and said method comprising:
- 20 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
- 25 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.

WO 2008/147926

PCT/US2008/064603

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.
- 5 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".
- 10 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 $N_2 = p * N_1$ where $p > 1$ then $N_1 = N$, $d_1 = d$, and
15 $d_2 = N_2 \times \frac{d}{N_1} = p \times d$; and
having a leaf stage comprising an input stage and an output stage; and said 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 its immediate succeeding stage where $x > 0$; and said output stage having $\frac{N_1}{d}$ output switches, and
20 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 its immediate succeeding stage;
and

WO 2008/147926

PCT/US2008/064603

a plurality of y middle stages, excepting a root stage, having $x \times \frac{N}{d}$ middle switches in each of said y middle stages wherein one of said middle stages is the immediate succeeding stage to both said input stage and said output stage, where $y > 1$, and said root stage having $\frac{N}{d}$ middle switches, and

5 each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, having d incoming links connected from switches in its immediate preceding stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links connected to switches in its immediate succeeding

10 stage and d outgoing links connected to switches in its immediate succeeding stage; and each middle switch in said succeeding stage to both said input stage and said output stage having d incoming links connected from switches in said input stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further having $\frac{(d + d_2)}{2}$ outgoing links connected to switches in said output

15 stage and d outgoing links connected to switches in its immediate succeeding stage; and each middle switch in said root stage having d incoming links connected from switches in its immediate preceding stage and each middle switch further having d outgoing links connected to switches in its immediate preceding stage; or

when $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$ then $N_2 = N$, $d_2 = d$ and

20 $d_1 = N_1 \times \frac{d}{N_2} = p \times d$; and having

having a leaf stage having an input stage and an output stage; and said 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 its immediate succeeding stage where $x > 0$; and said output stage having $\frac{N_2}{d}$ output switches, and

WO 2008/147926

PCT/US2008/064603

each output switch having d outlet links and each output switch further having $x \times d$ incoming links connected from switches in its immediate succeeding stage; and

a plurality of y middle stages, excepting a root stage, having $x \times \frac{N}{d}$ middle switches in each of said y middle stages wherein one of said middle stages is the
 5 immediate succeeding stage to both said input stage and said output stage, where $y > 1$, and said root stage having $\frac{N}{d}$ middle switches, and

each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, having d incoming links connected from switches in its immediate preceding stage and d incoming links
 10 connected from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links connected to switches in its immediate succeeding stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said succeeding stage to both said input stage and said output stage having $\frac{(d + d_1)}{2}$ incoming links connected from switches in said input stage
 15 and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further having d outgoing links connected to switches in said output stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said root stage having d incoming links connected from switches in its immediate preceding stage and each middle switch further having d
 20 outgoing links connected to switches in its immediate preceding 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

25 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.

WO 2008/147926

PCT/US2008/064603

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 **Error! Reference source not found.** further comprising:
5 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 **Error! Reference source not found.** further comprising:
10 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.

15 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 **Error! Reference source not found.** further comprising:
20 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 **Error! Reference source not found.** further comprising:
25 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.

WO 2008/147926

PCT/US2008/064603

22. The method of claim **Error! Reference source not found.** wherein any of said acts of checking and setting up are performed recursively.

100A

FIG. 1A

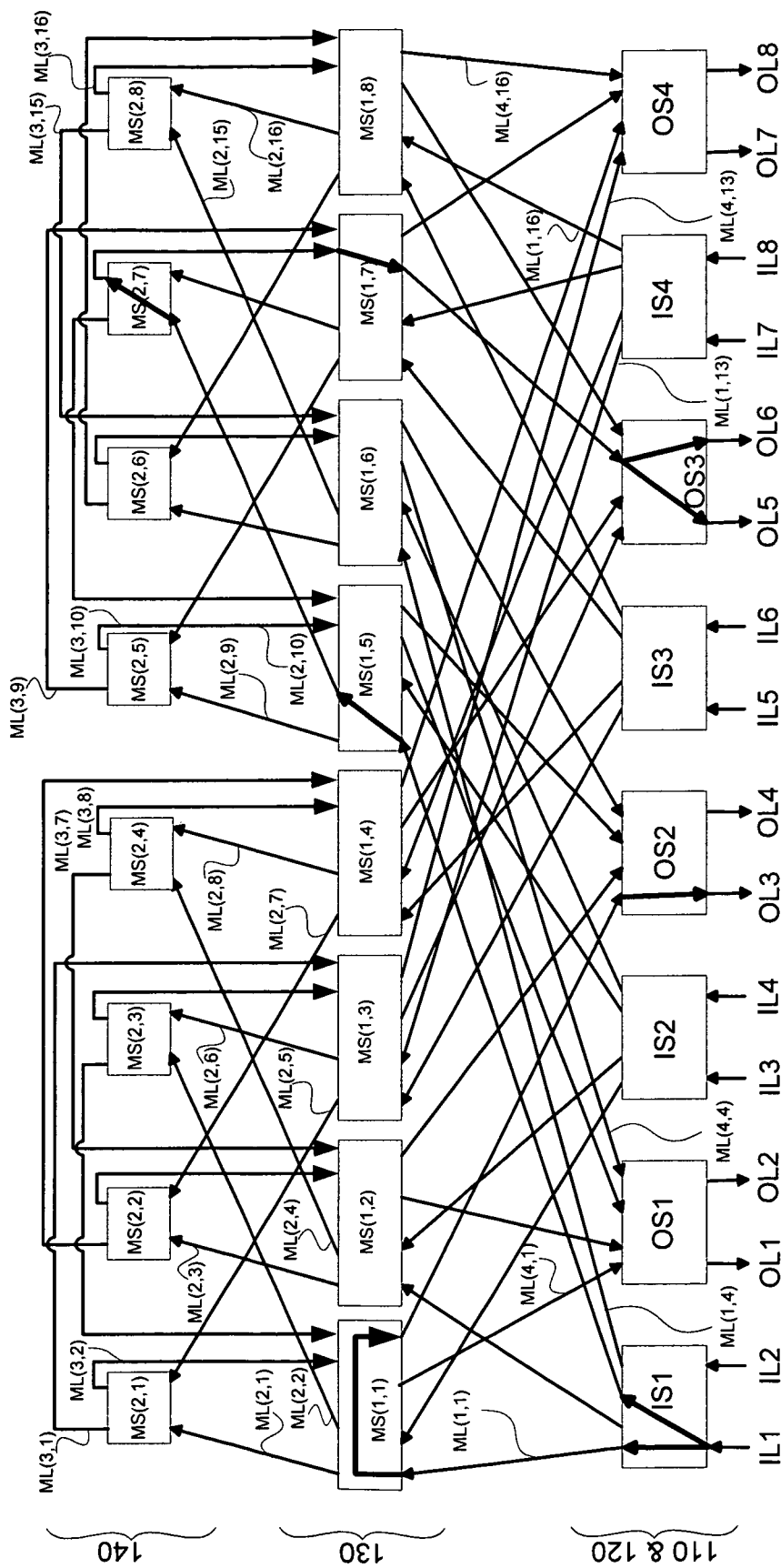
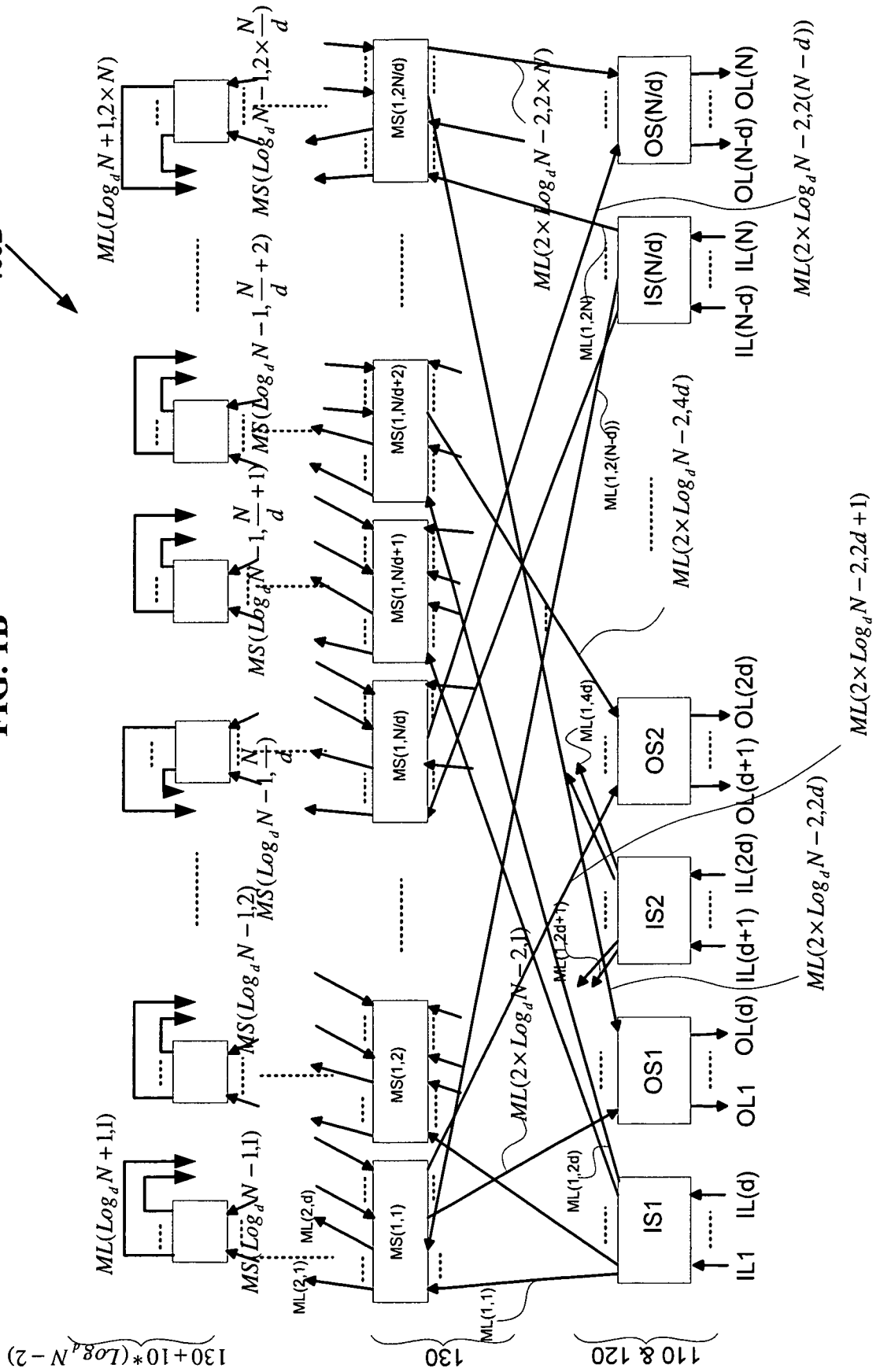


FIG. 1B

100B



100C

FIG. 1C

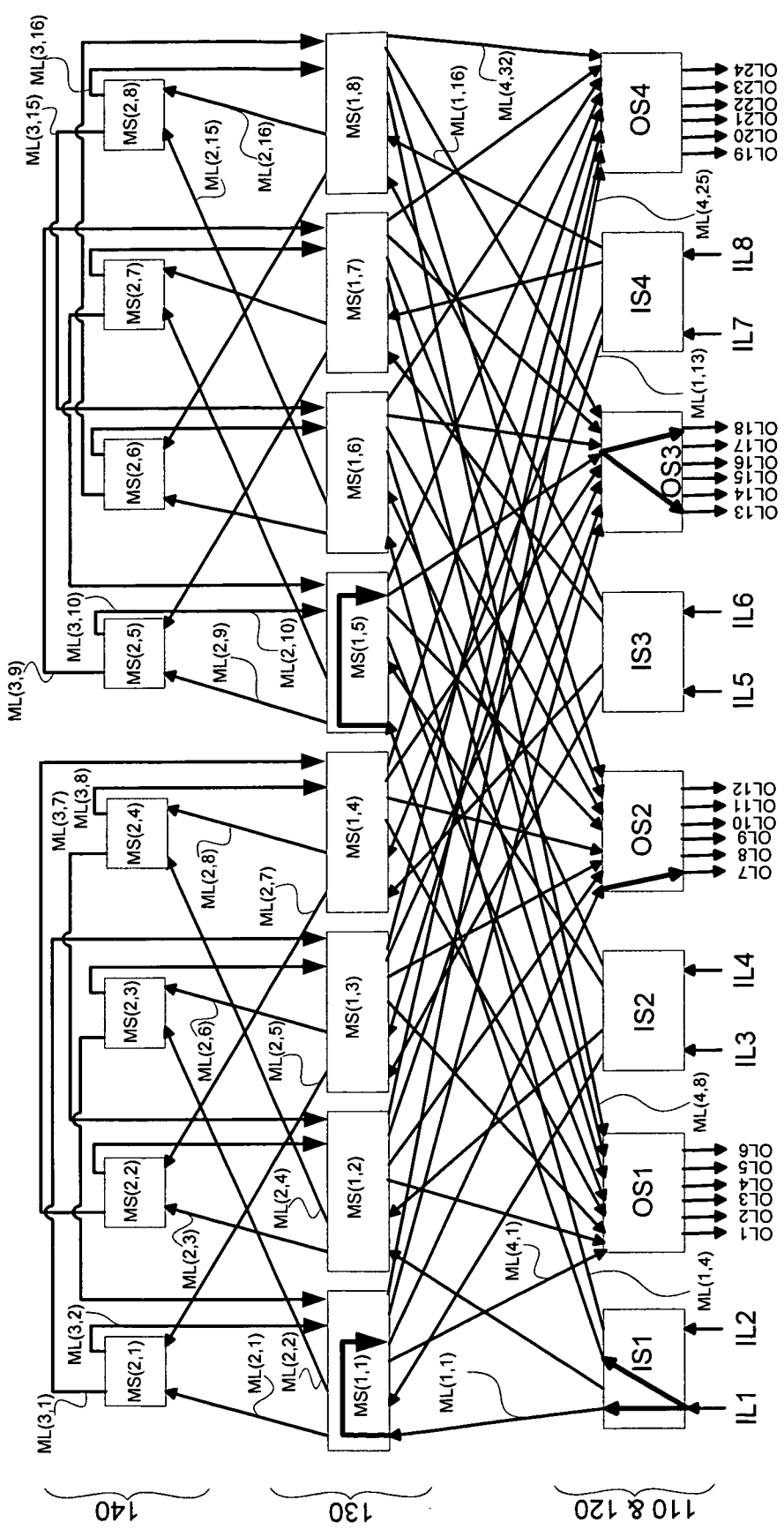
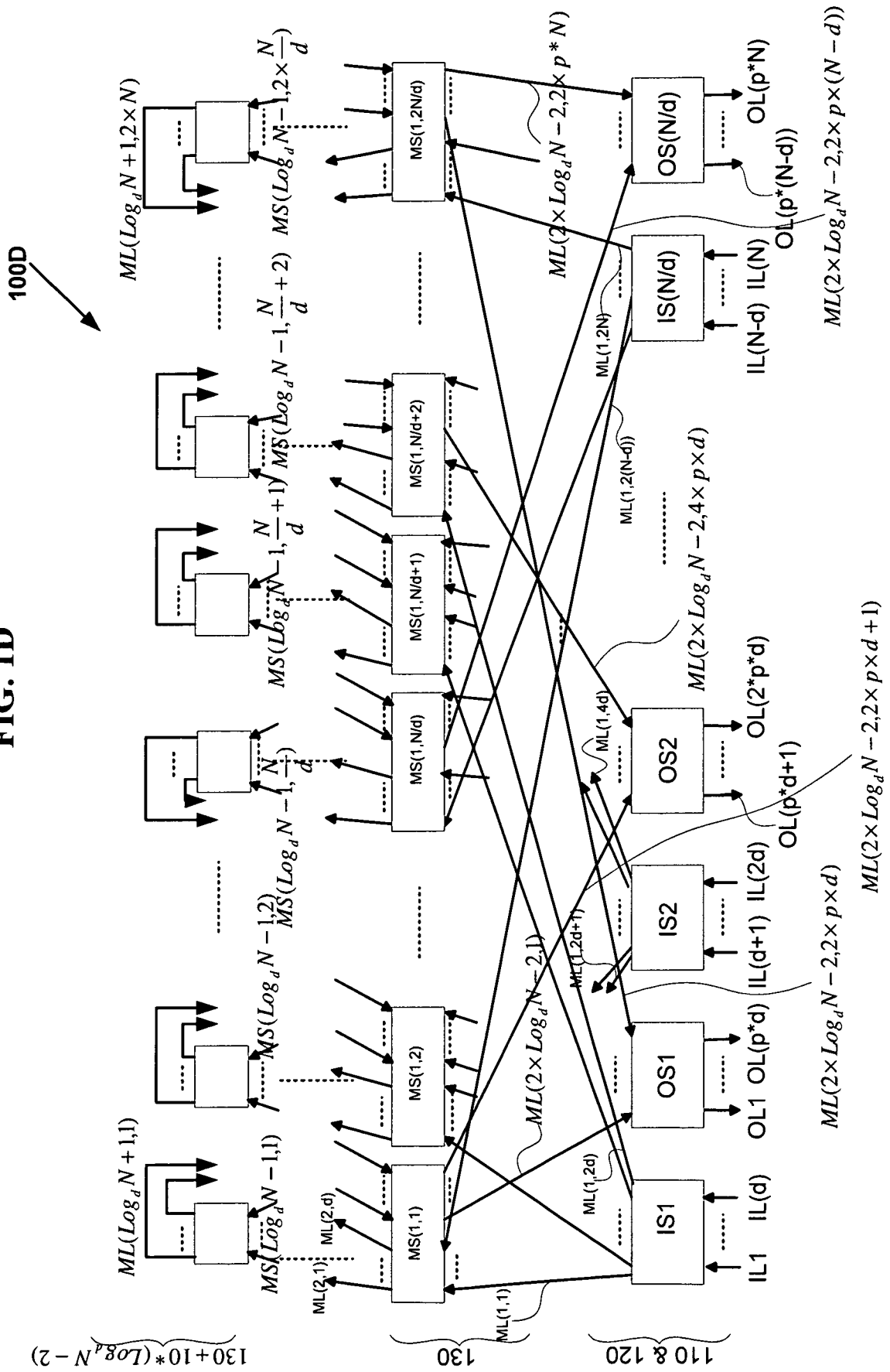


FIG. 1D



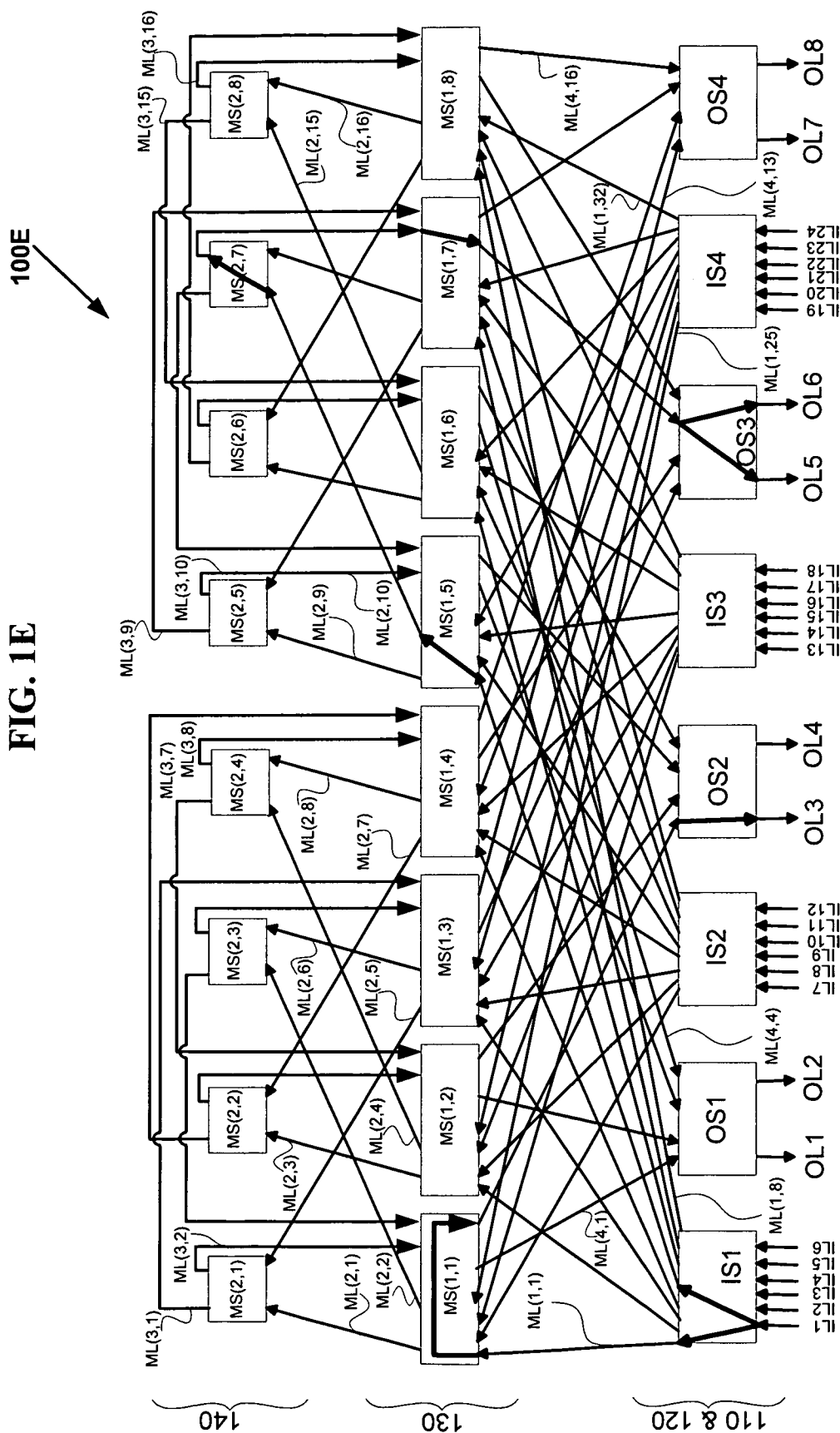


FIG. 1E

FIG. 1F

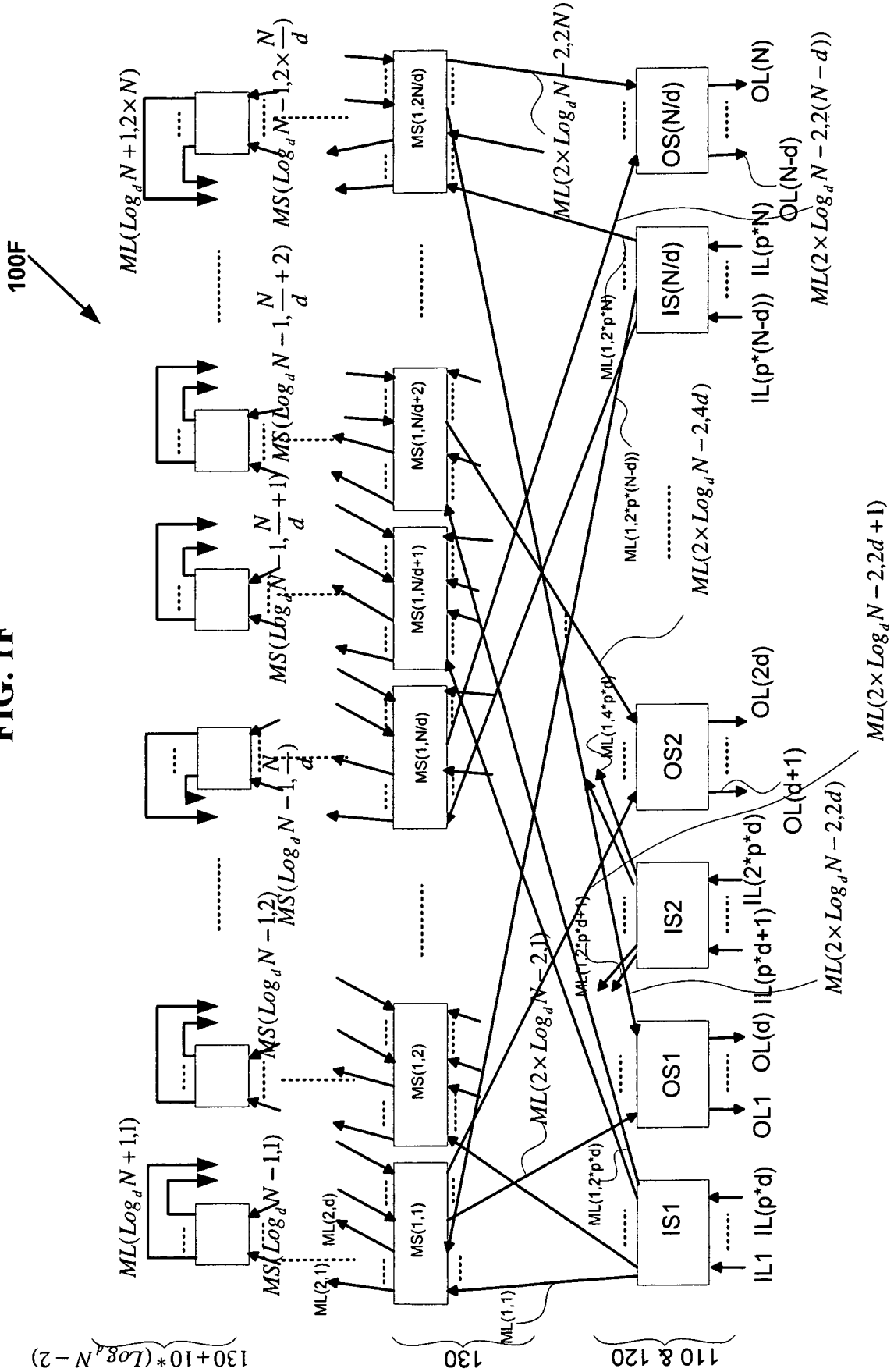


FIG. 2A

200A

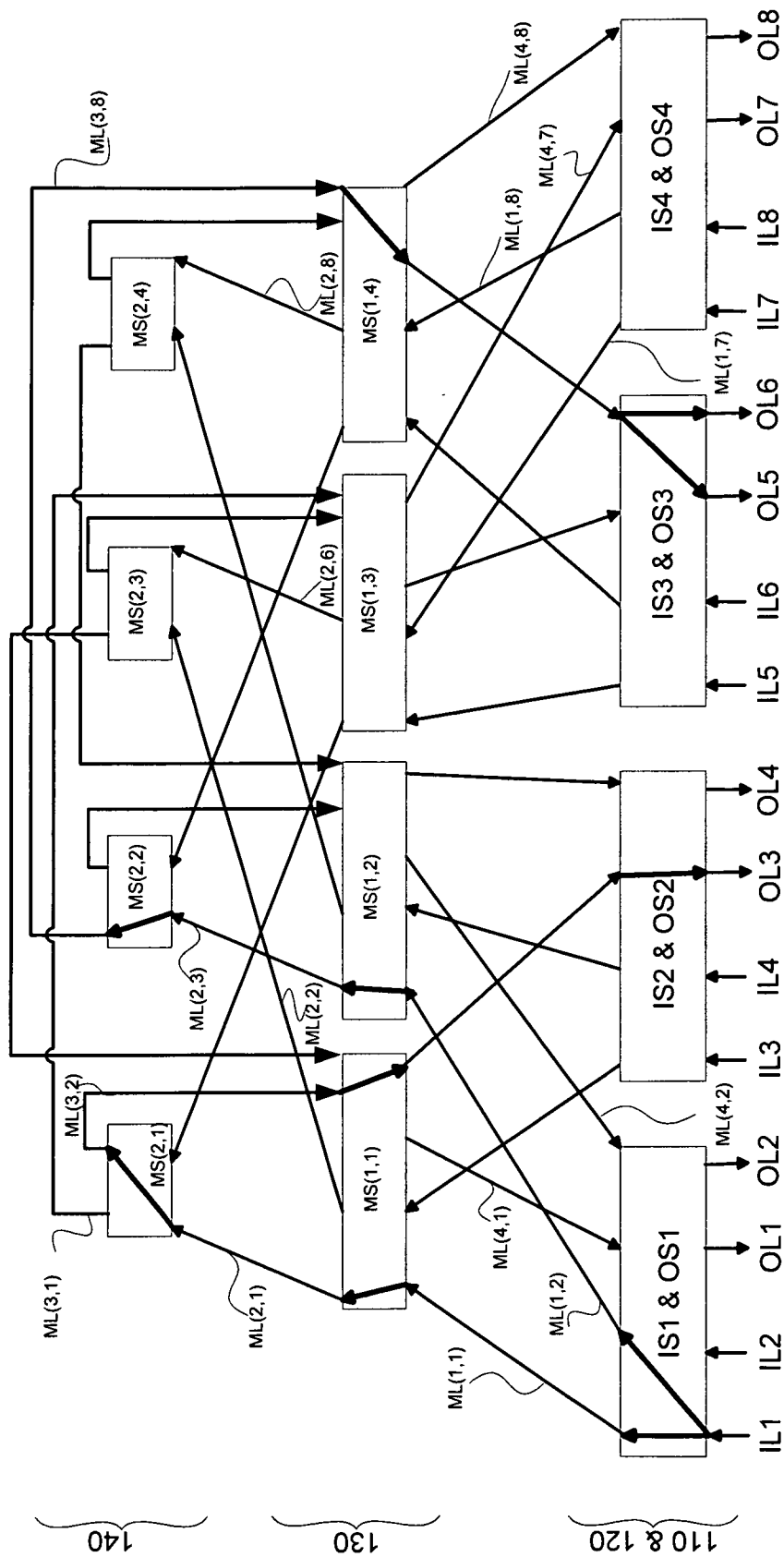


FIG. 2B

200B

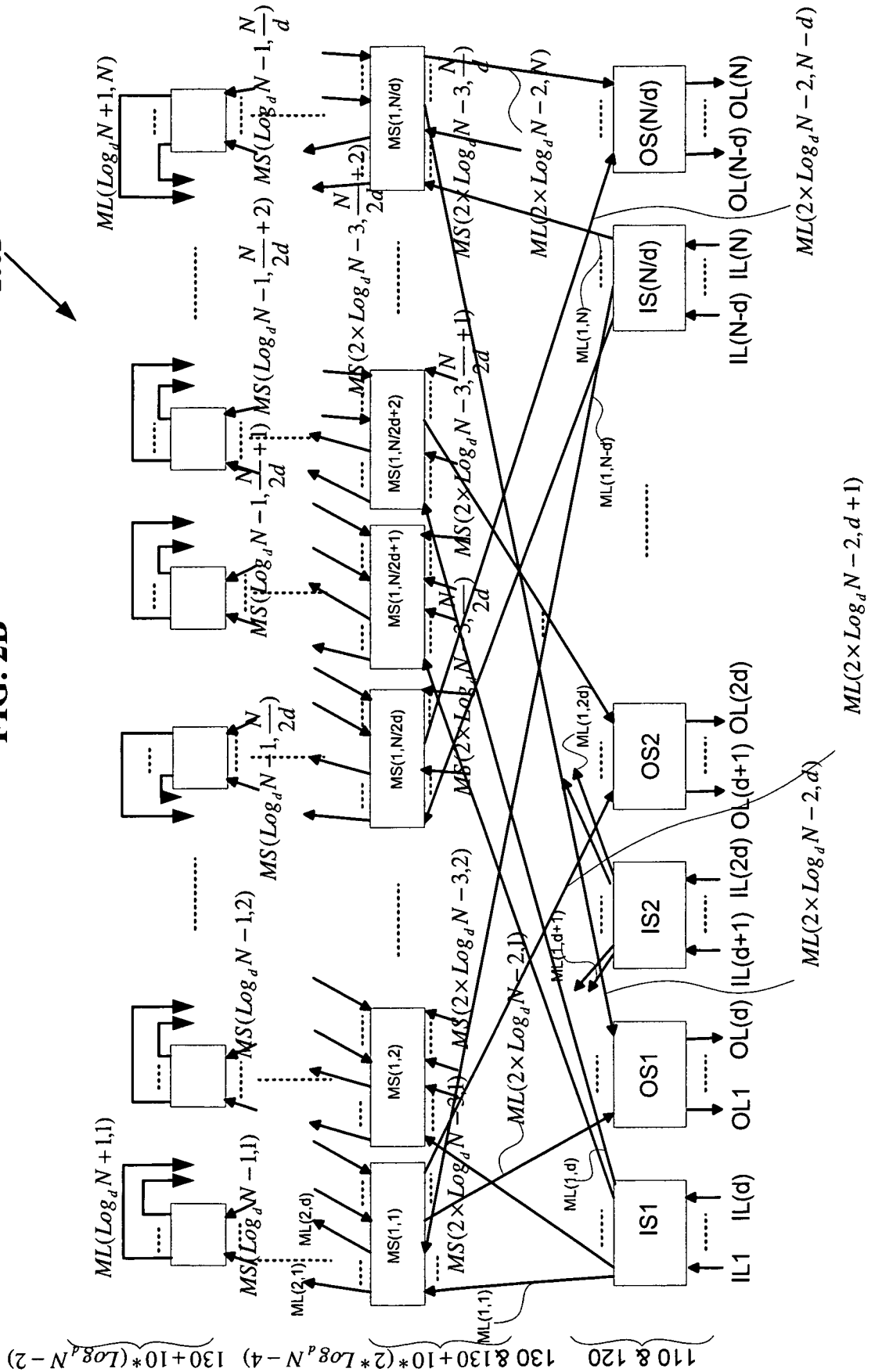


FIG. 2C

200C

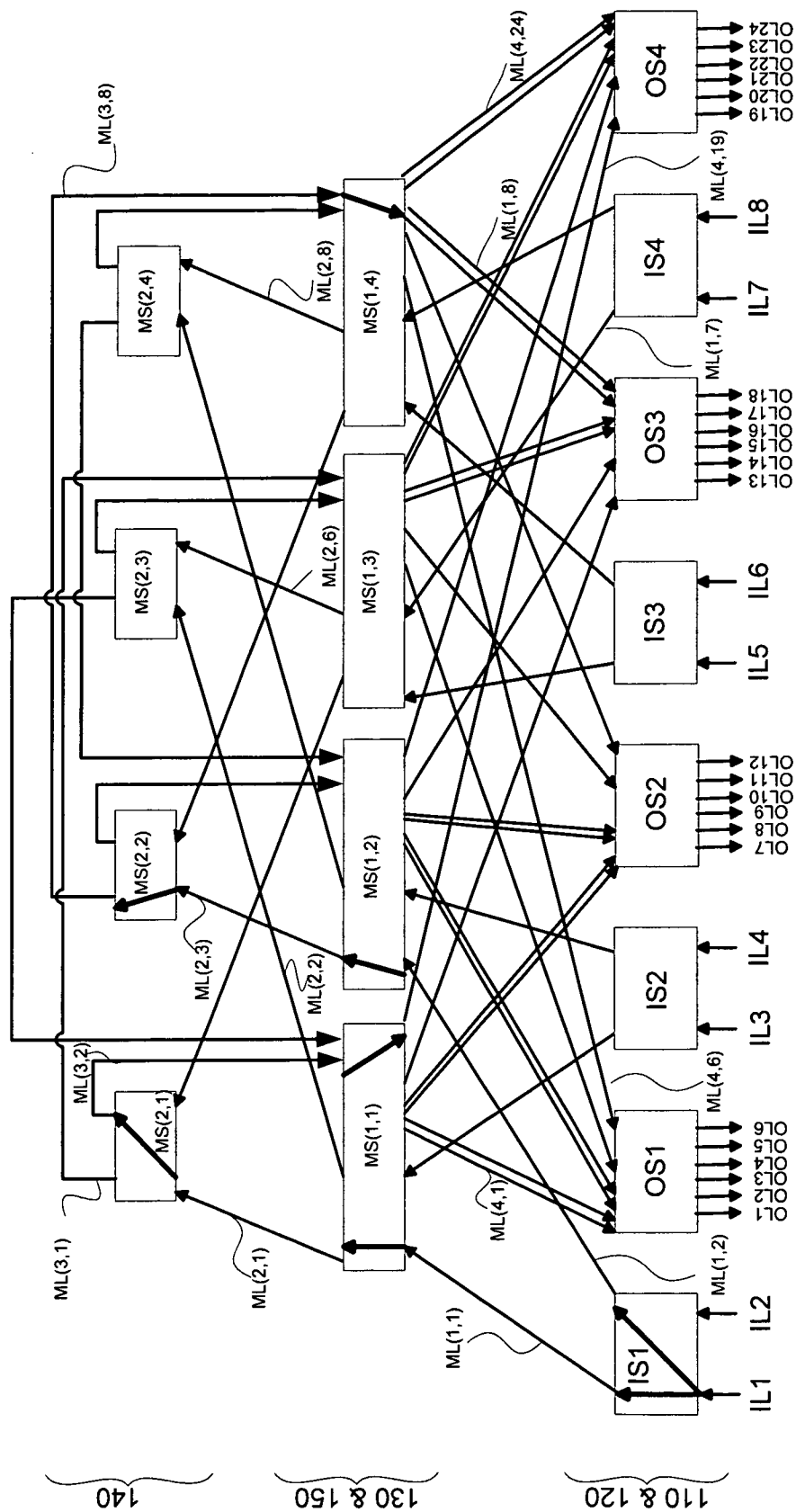


FIG. 2D

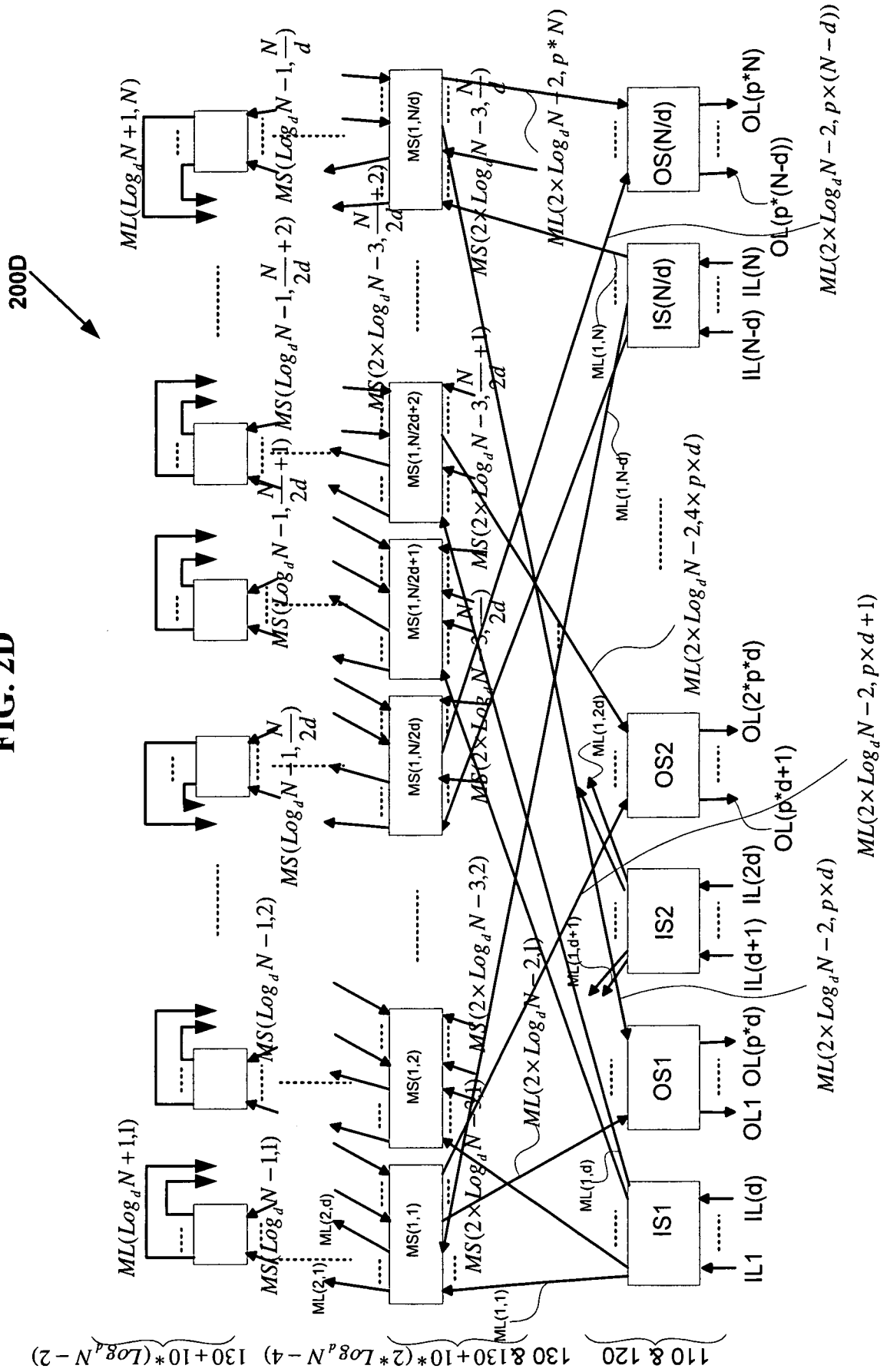


FIG. 2E

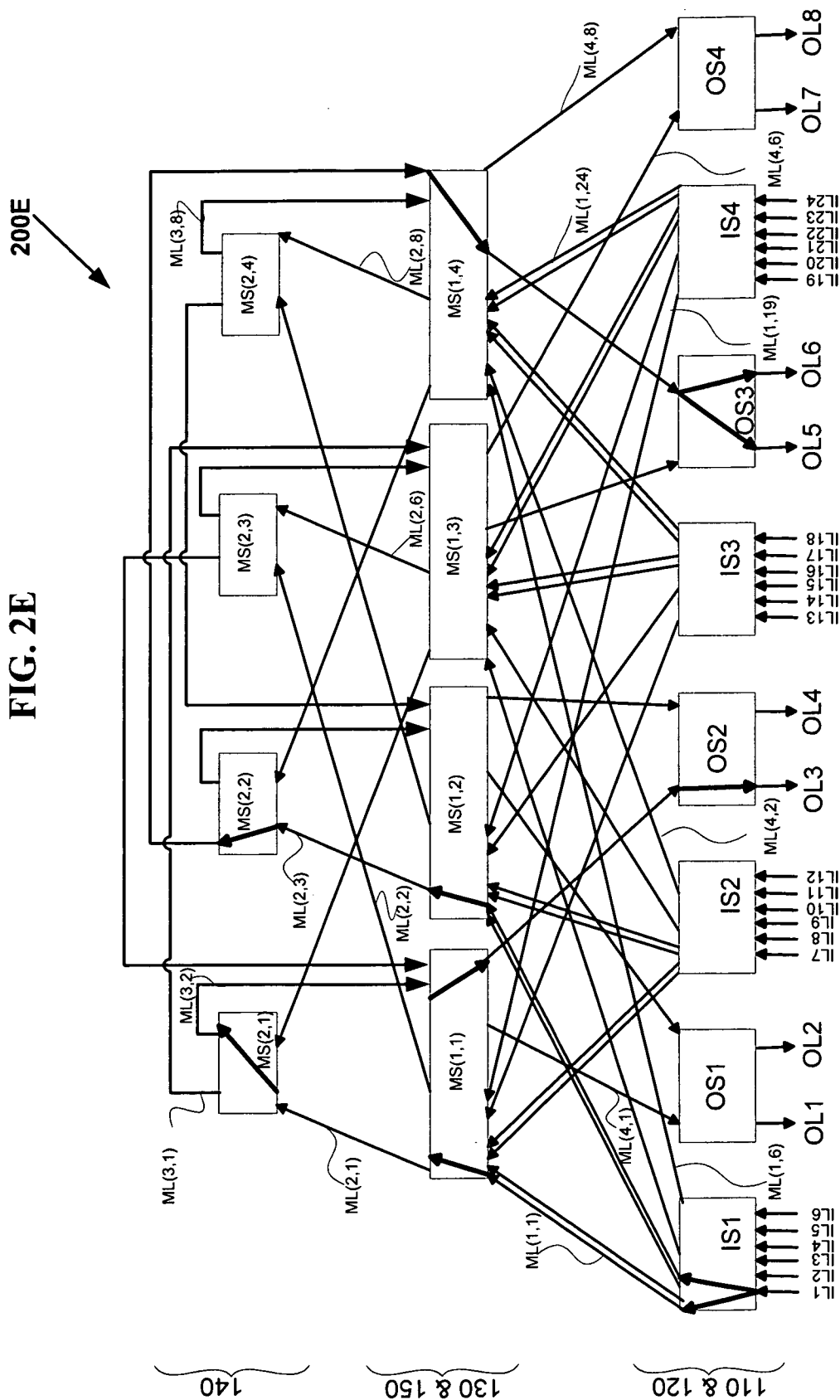
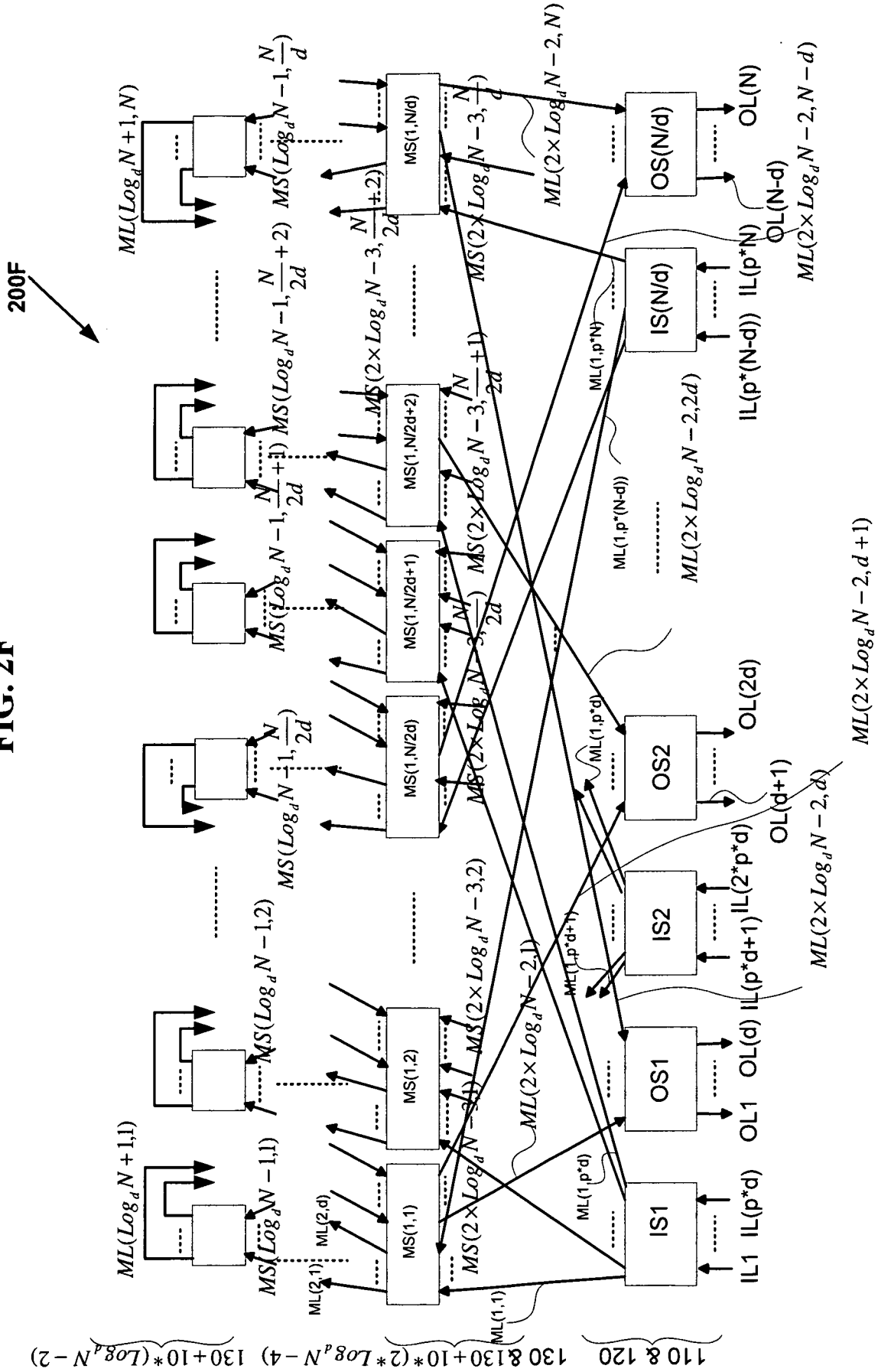


FIG. 2F



200F

300A

FIG. 3A

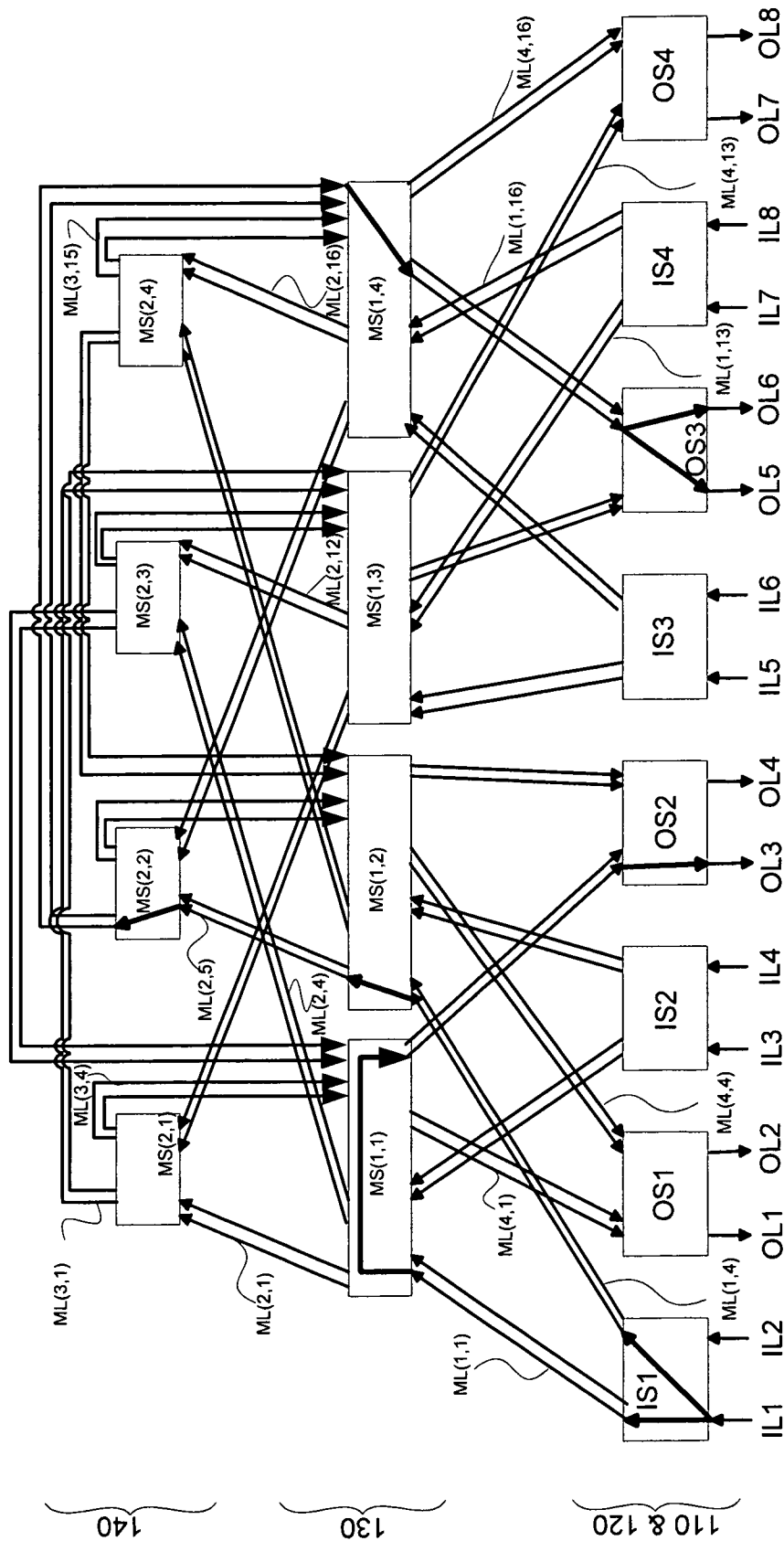


FIG. 3B

300B

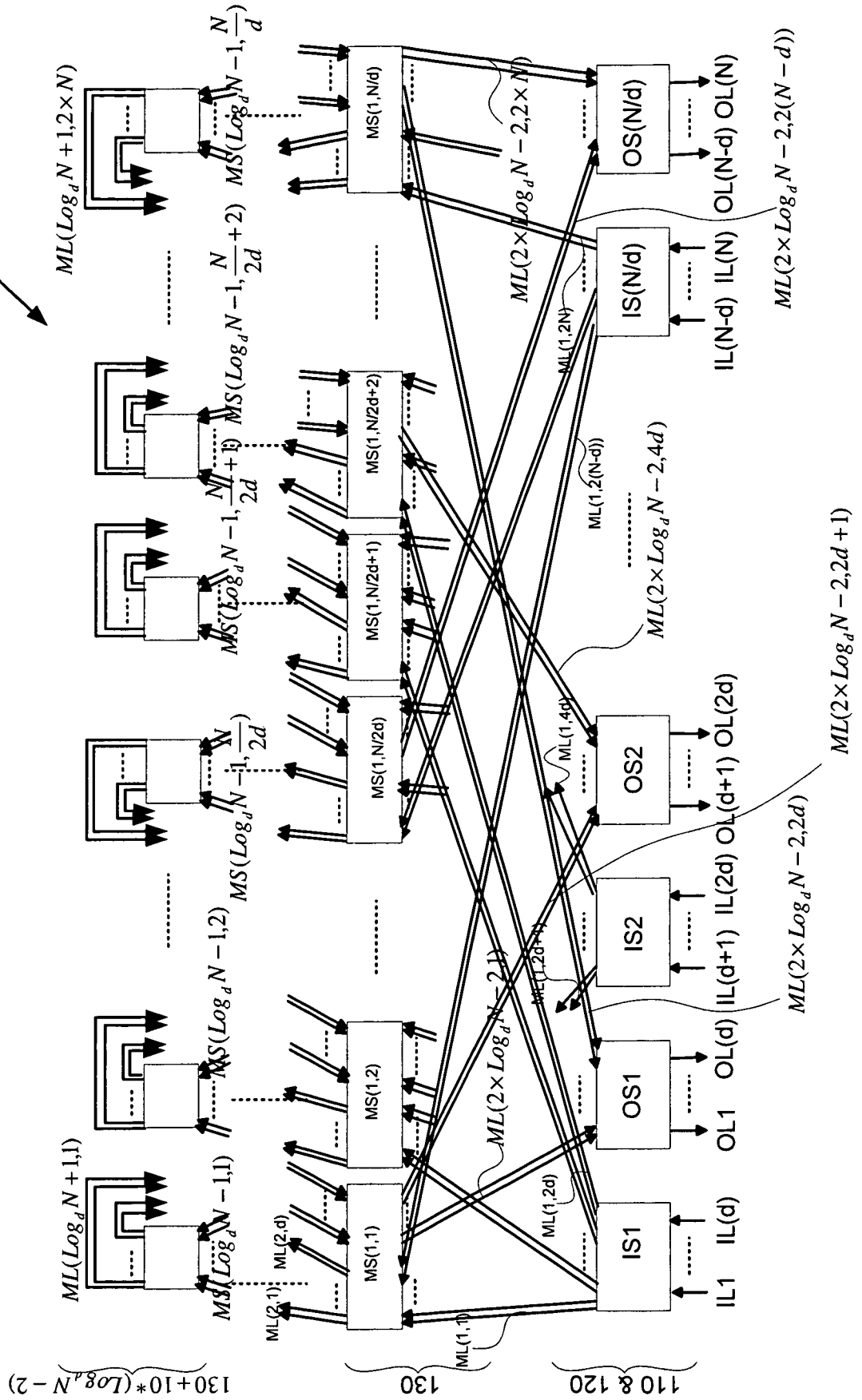


FIG. 3C

300C

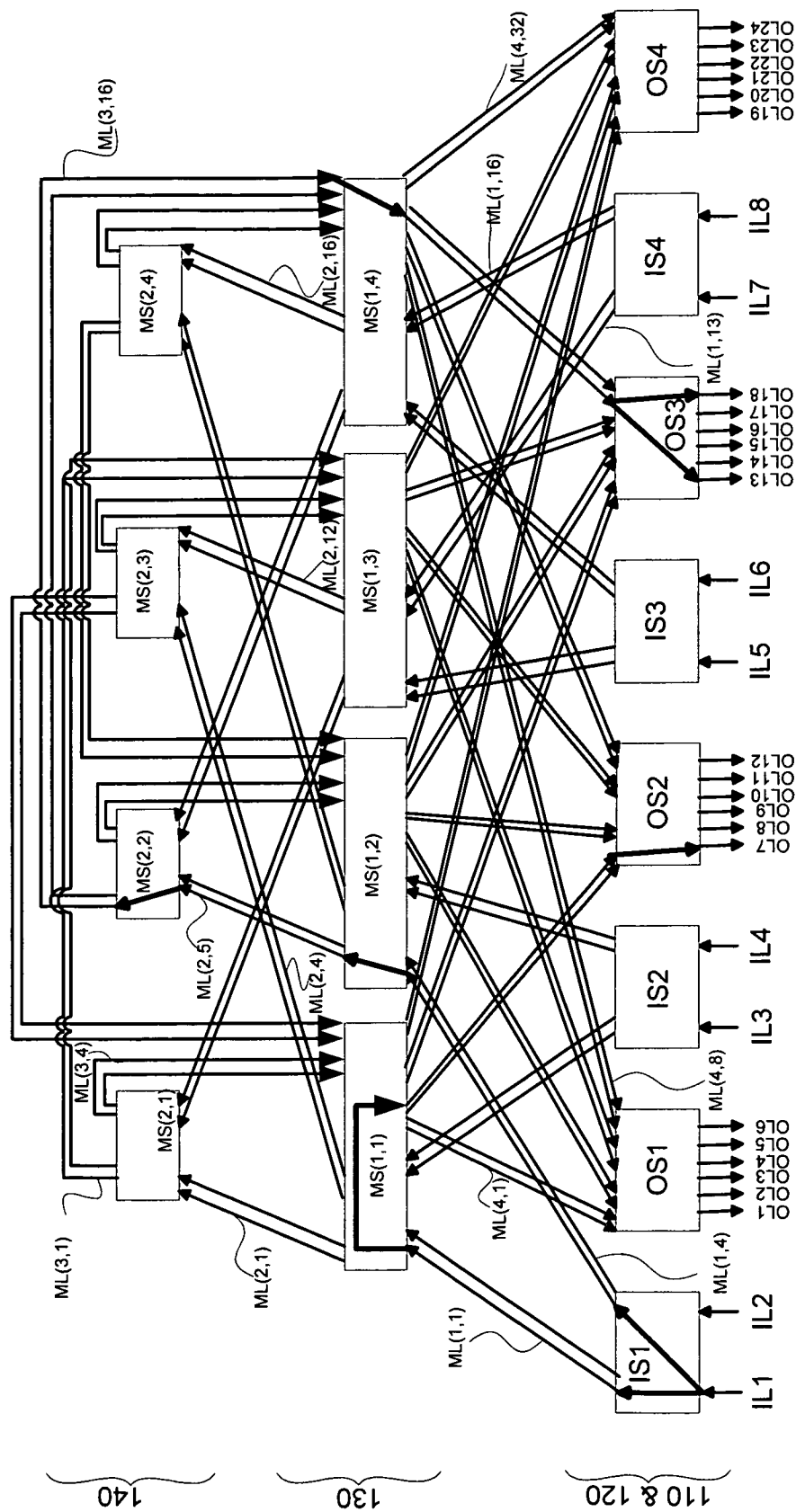


FIG. 3D

300D

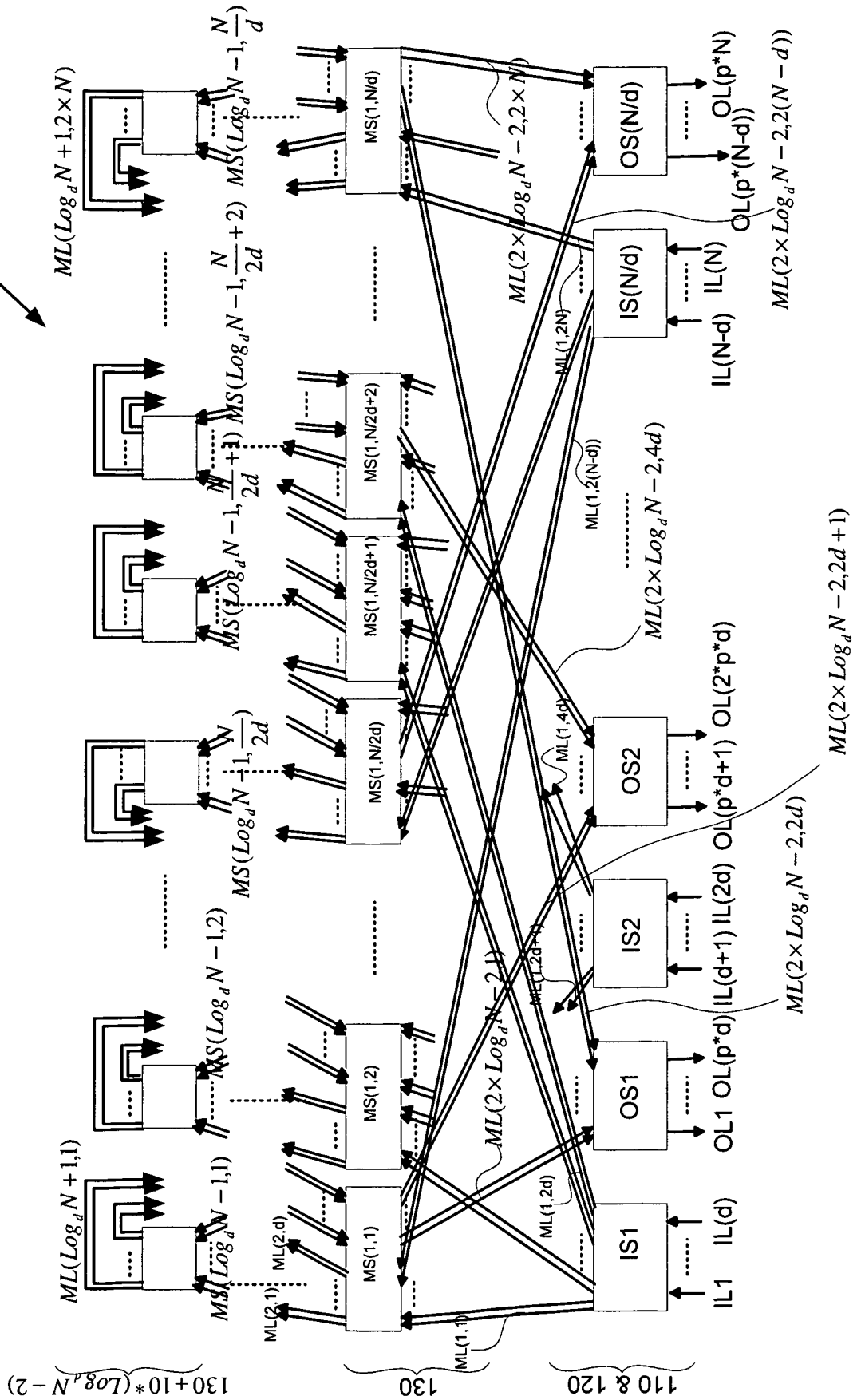


FIG. 3E

300E

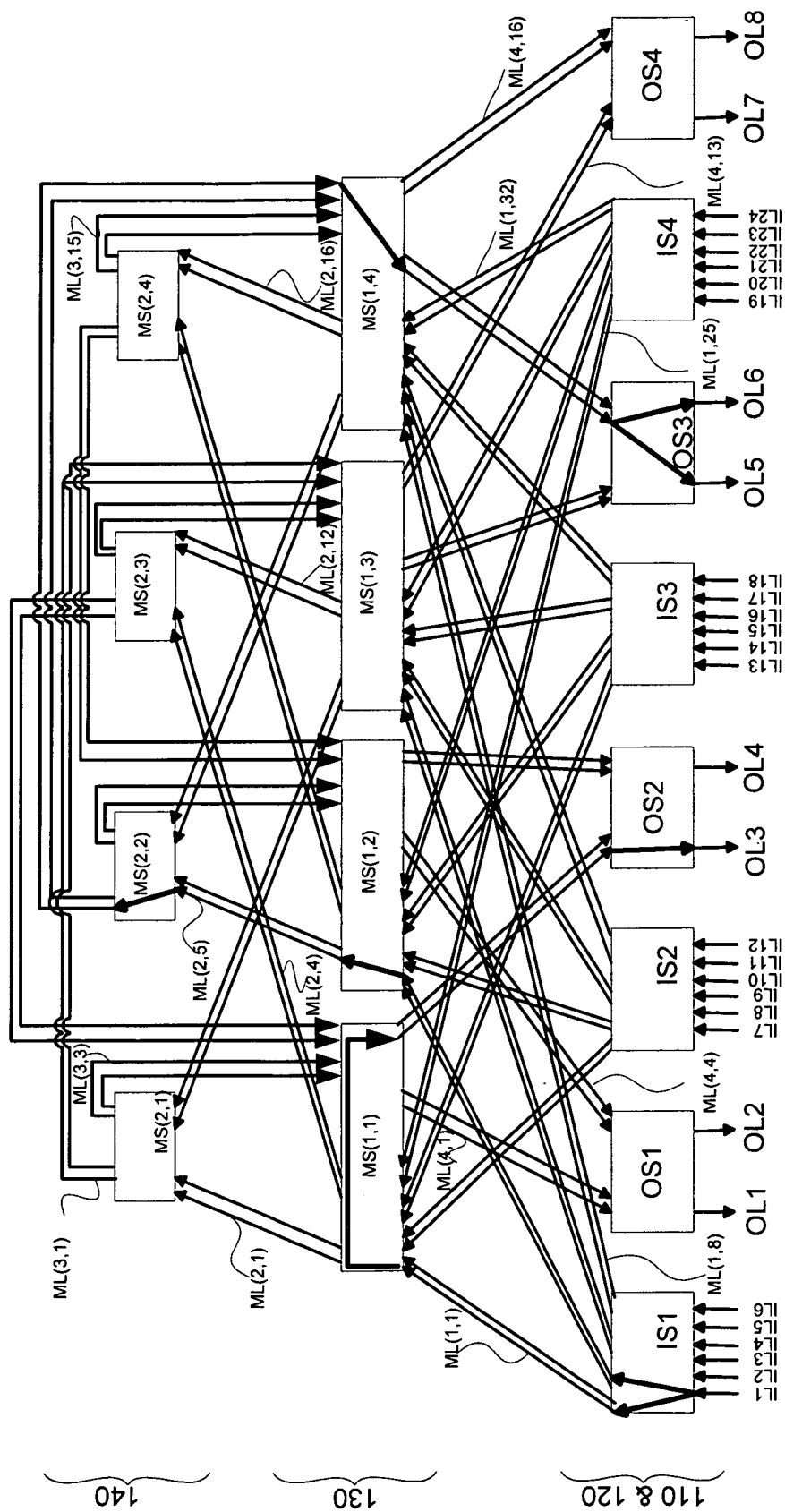


FIG. 3F

300F

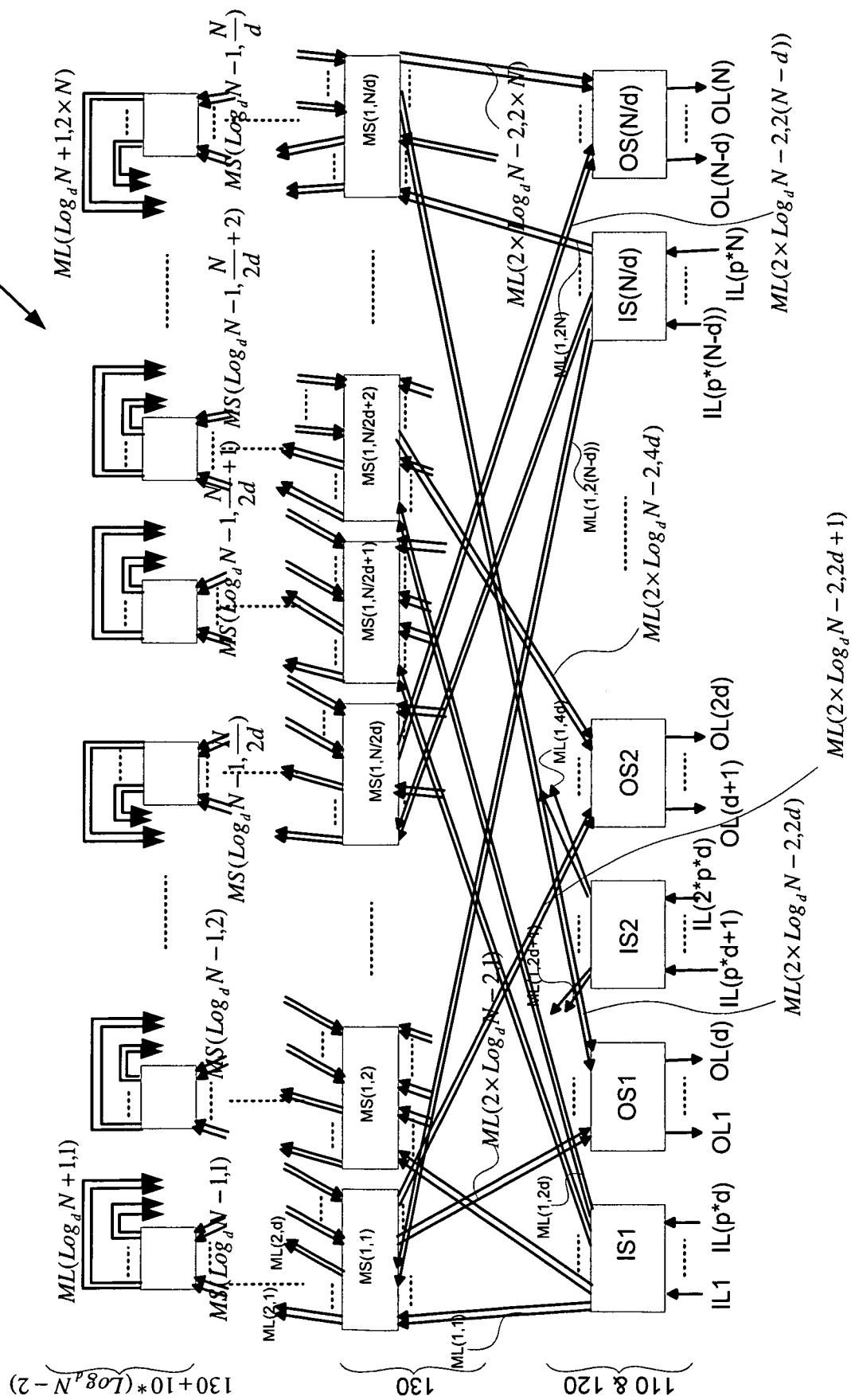


FIG. 4A

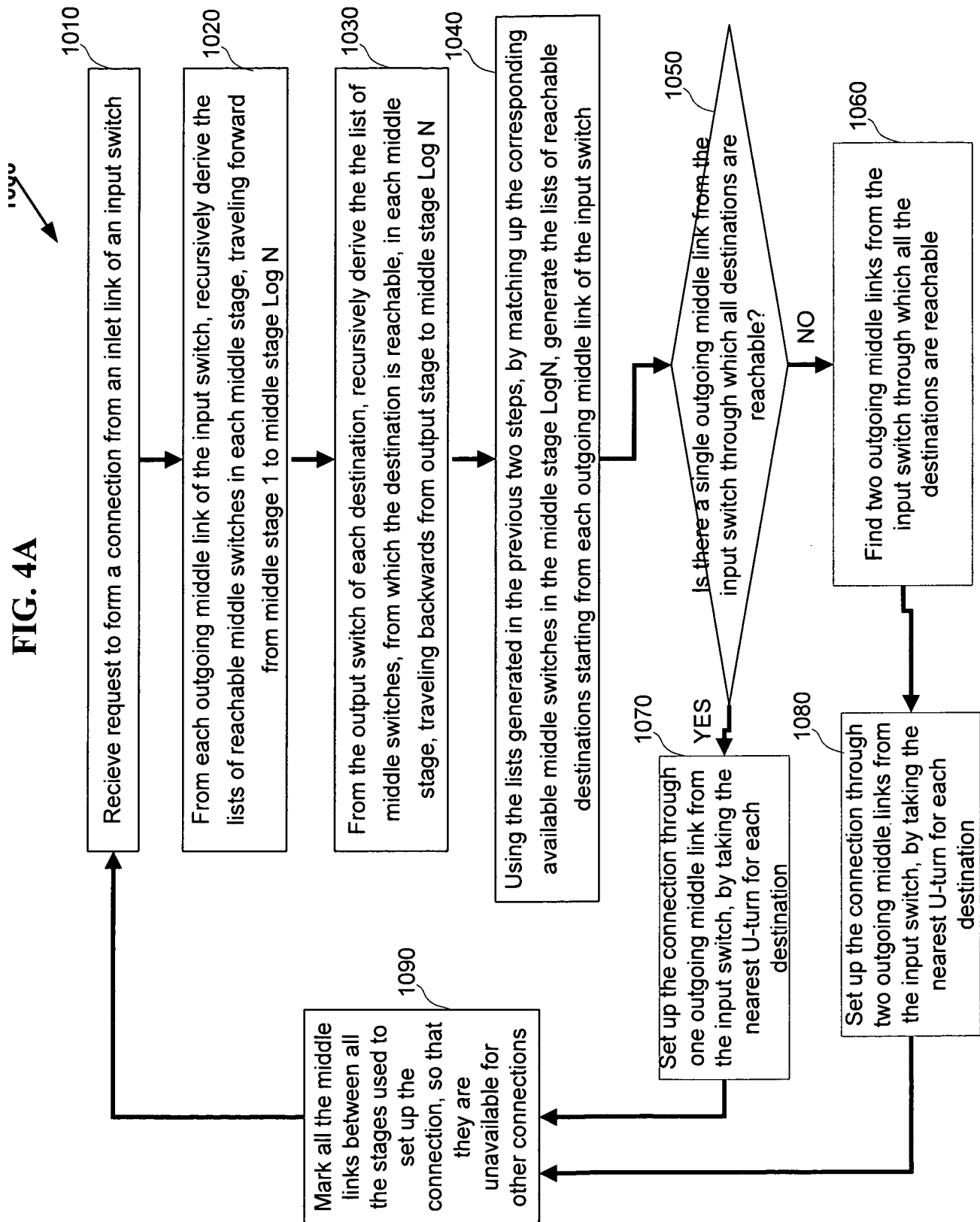


FIG. 5A

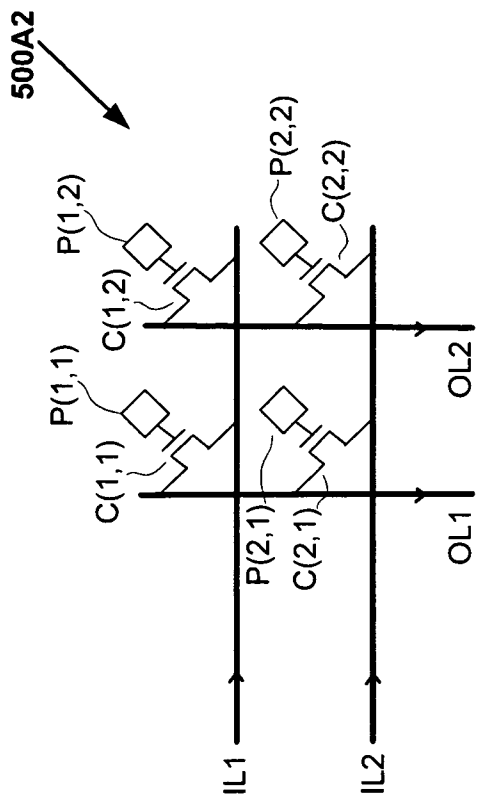


FIG. 5A2
(Prior Art)

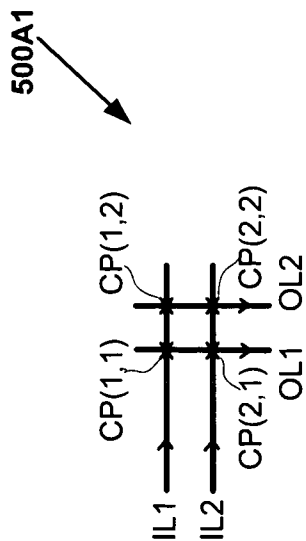


FIG. 5A1
(Prior Art)

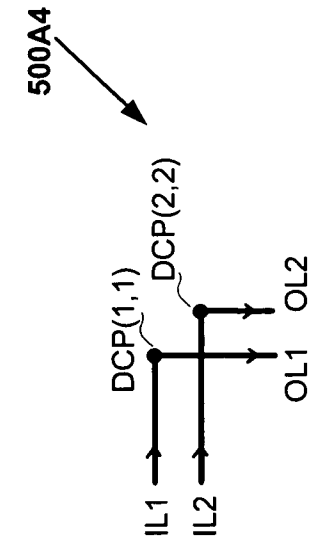


FIG. 5A4

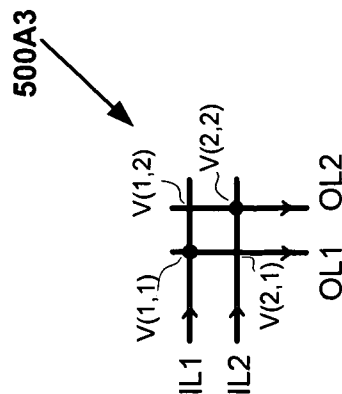


FIG. 5A3
(Prior Art)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2008/064603

A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - H03K 17/00 (2008.04) USPC - 340/2.2 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) - H03K 17/00; H04L 12/28 (2008.04) USPC - 340/2.2; 370/395.1 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, IP.com, DialogPro												
C. DOCUMENTS CONSIDERED TO BE RELEVANT												
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.										
A	US 6,868,084 B2 (KONDA) 15 March 2005 (15.03.2005) entire document	1-22										
A	US 2007/0053356 A1 (KONDA) 08 March 2007 (08.03.2007) entire document	1-22										
A	US 5,179,551 A (TURNER) 12 January 1993 (12.01.1993) entire document	1-22										
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>												
<p>* Special categories of cited documents:</p> <table border="0"> <tr> <td>"A" document defining the general state of the art which is not considered to be of particular relevance</td> <td>"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</td> </tr> <tr> <td>"E" earlier application or patent but published on or after the international filing date</td> <td>"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</td> </tr> <tr> <td>"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)</td> <td>"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</td> </tr> <tr> <td>"O" document referring to an oral disclosure, use, exhibition or other means</td> <td>"&" document member of the same patent family</td> </tr> <tr> <td>"P" document published prior to the international filing date but later than the priority date claimed</td> <td></td> </tr> </table>			"A" document defining the general state of the art which is not considered to be of particular relevance	"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	"E" earlier application or patent but published on or after the international filing date	"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	"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)	"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	"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	"P" document published prior to the international filing date but later than the priority date claimed	
"A" document defining the general state of the art which is not considered to be of particular relevance	"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											
"E" earlier application or patent but published on or after the international filing date	"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											
"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)	"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											
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family											
"P" document published prior to the international filing date but later than the priority date claimed												
Date of the actual completion of the international search 22 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 D



US 20100172349A1

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

(10) **Pub. No.: US 2010/0172349 A1**
 (43) **Pub. Date: Jul. 8, 2010**

(54) **FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS**

(76) Inventor: **Venkat Konda, (US)**

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

(21) Appl. No.: **12/601,273**

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

(86) PCT No.: **PCT/US08/64603**

§ 371 (c)(1),
 (2), (4) Date: **Nov. 22, 2009**

Related U.S. Application Data

(60) Provisional application No. 60/940,387, filed on May 25, 2007, provisional application No. 60/940,390, filed on May 25, 2007.

Publication Classification

(51) **Int. Cl.**
H04L 12/56 (2006.01)
 (52) **U.S. Cl.** **370/390**

(57) **ABSTRACT**

A generalized butterfly fat tree network comprising $(\log_d N)$ stages is operated in strictly nonblocking manner for unicast, when $s \geq 2$, includes a leaf stage consisting of an input stage having N/d switches with each of them having d inlet links

and $s \times d$ outgoing links connecting to its immediate succeeding stage switches, and an output stage having N/d switches with each of them having d outlet links and $s \times d$ incoming links connecting from switches in its immediate preceding stage. The network also has $(\log_d N) - 1$ middle stages with each middle stage, excepting the root stage, having

$$\frac{s \times N}{d}$$

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, d incoming links connecting from the switches in its immediate succeeding stage, d outgoing links connecting to the switches in its immediate succeeding stage, d outgoing links connecting to the switches in its immediate preceding stage, and the root stage having

$$\frac{s \times N}{d}$$

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage and d outgoing links connecting to the switches in its immediate preceding stage. Also the same generalized butterfly fat tree network, i.e. when $s \geq 2$, 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. Also the generalized butterfly fat tree network, when $s \geq 3$, is operated in strictly 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.

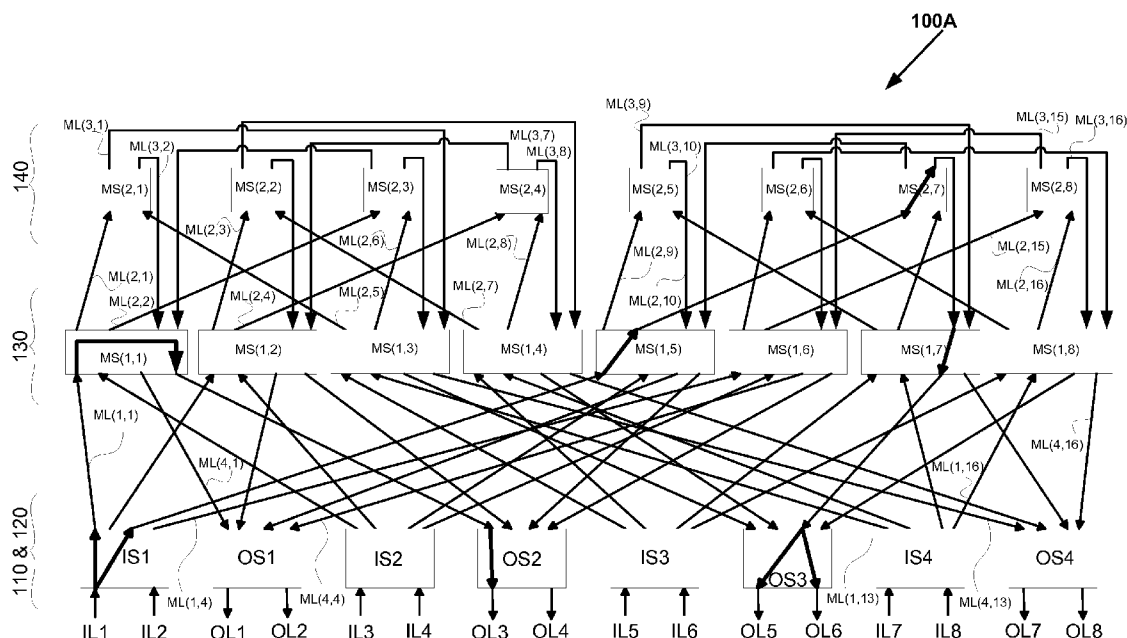


FIG. 1A

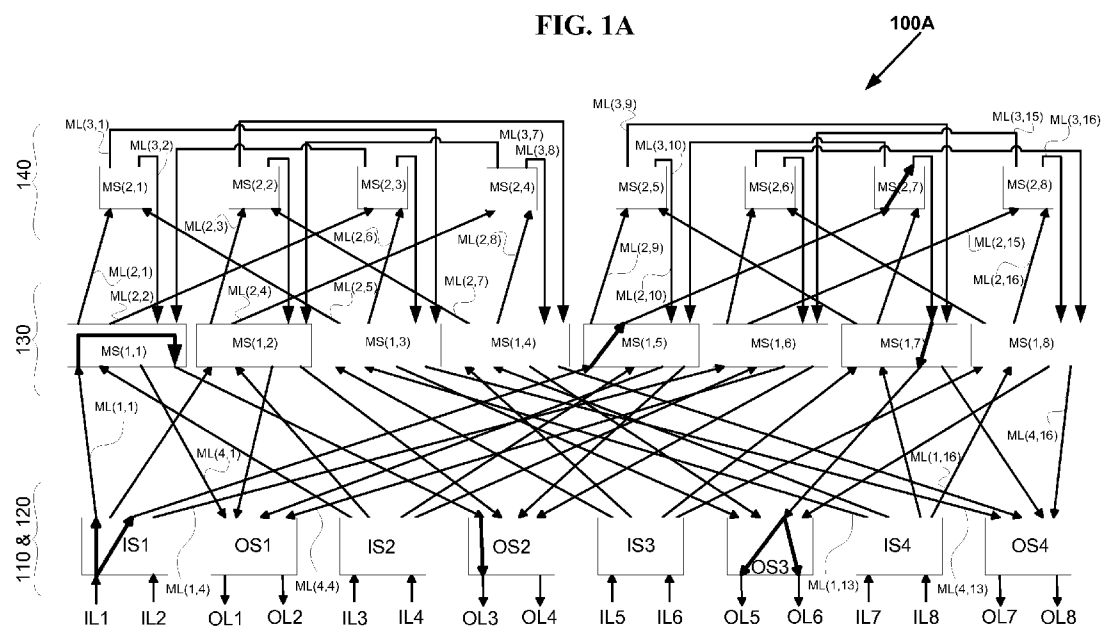


FIG. 1B

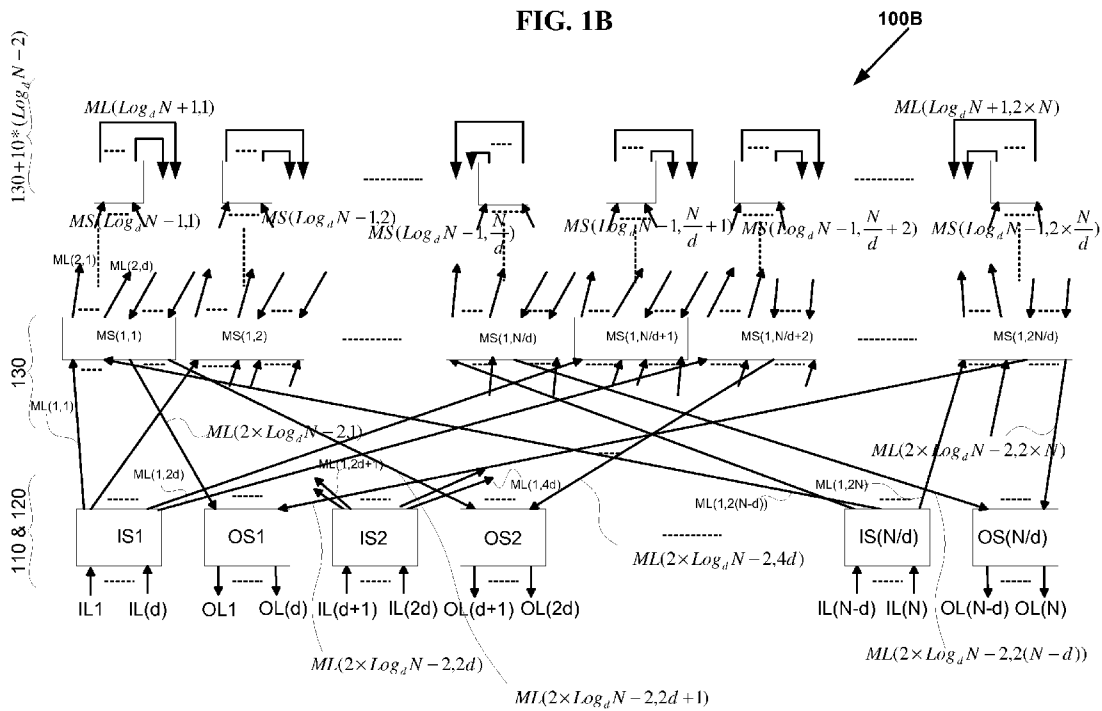


FIG. 1C

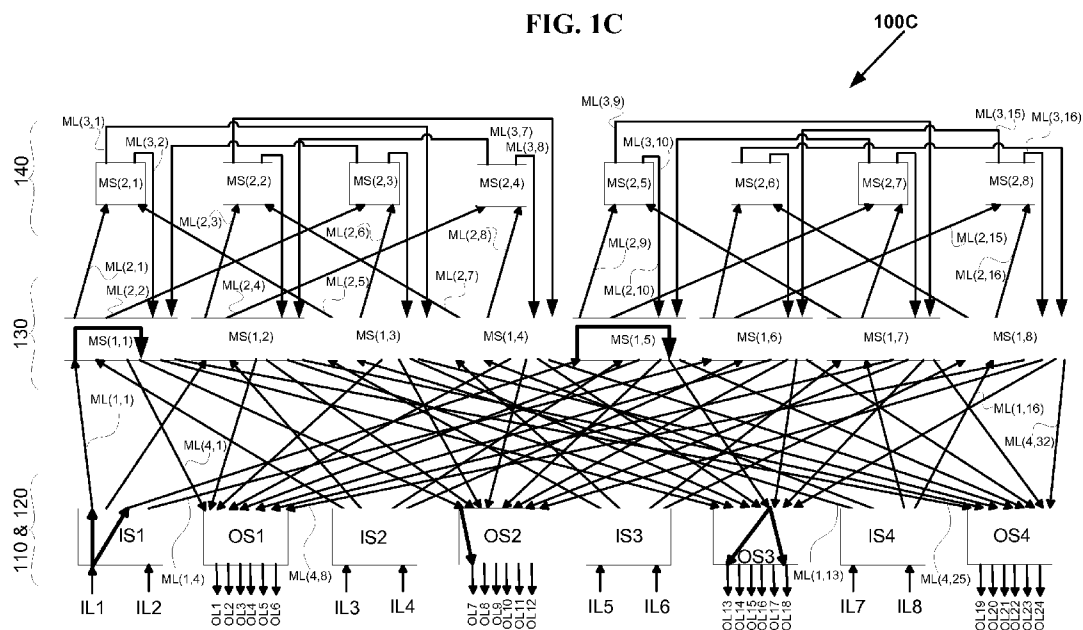


FIG. 1D

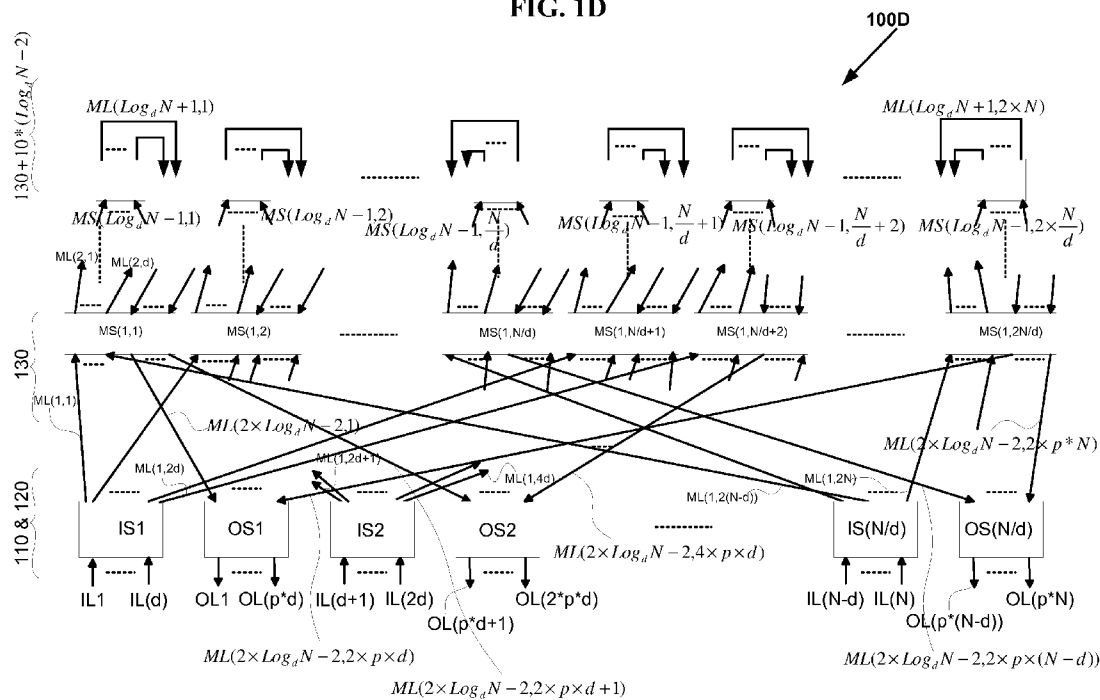


FIG. 1E

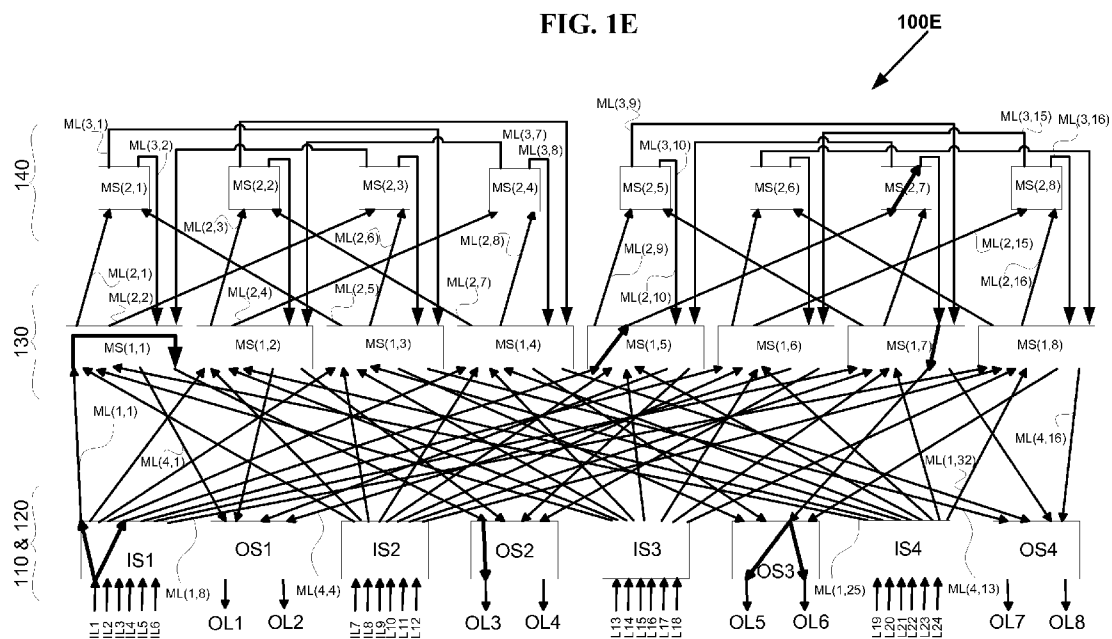


FIG. 1F

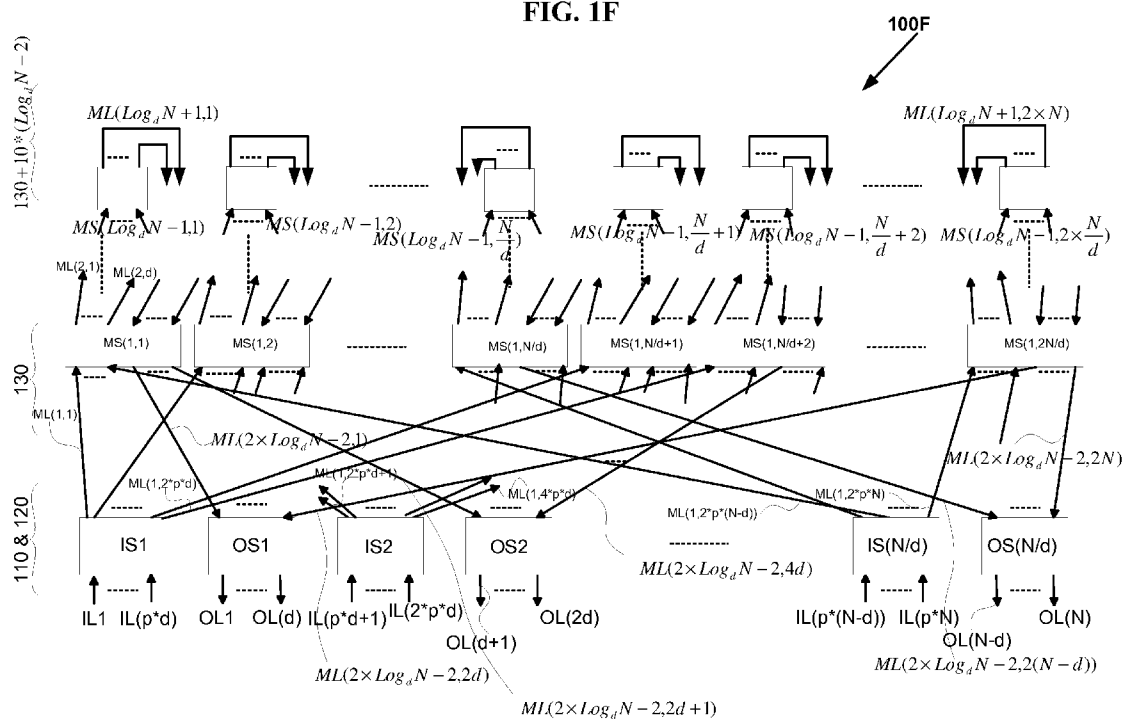
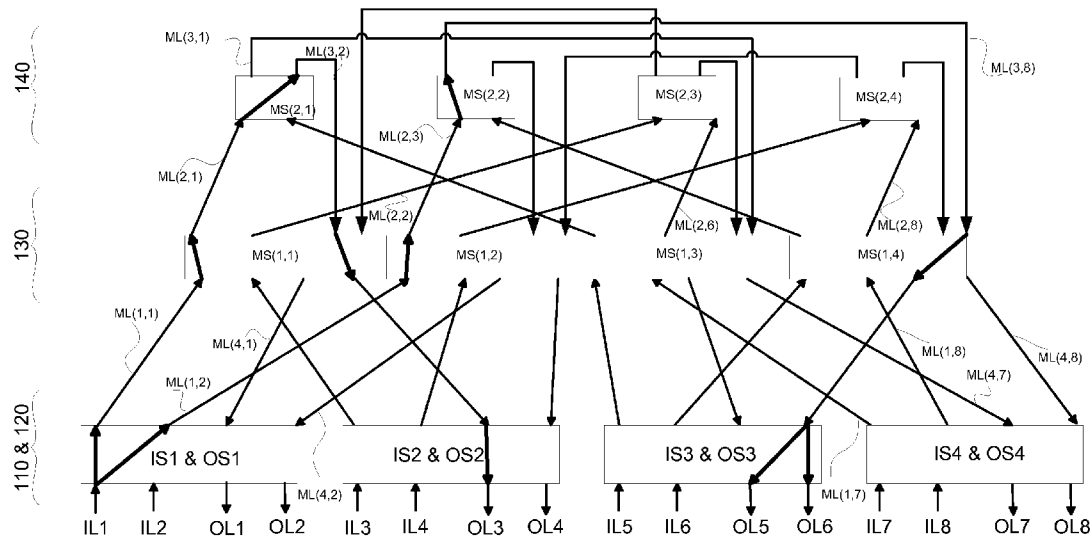


FIG. 2A

200A



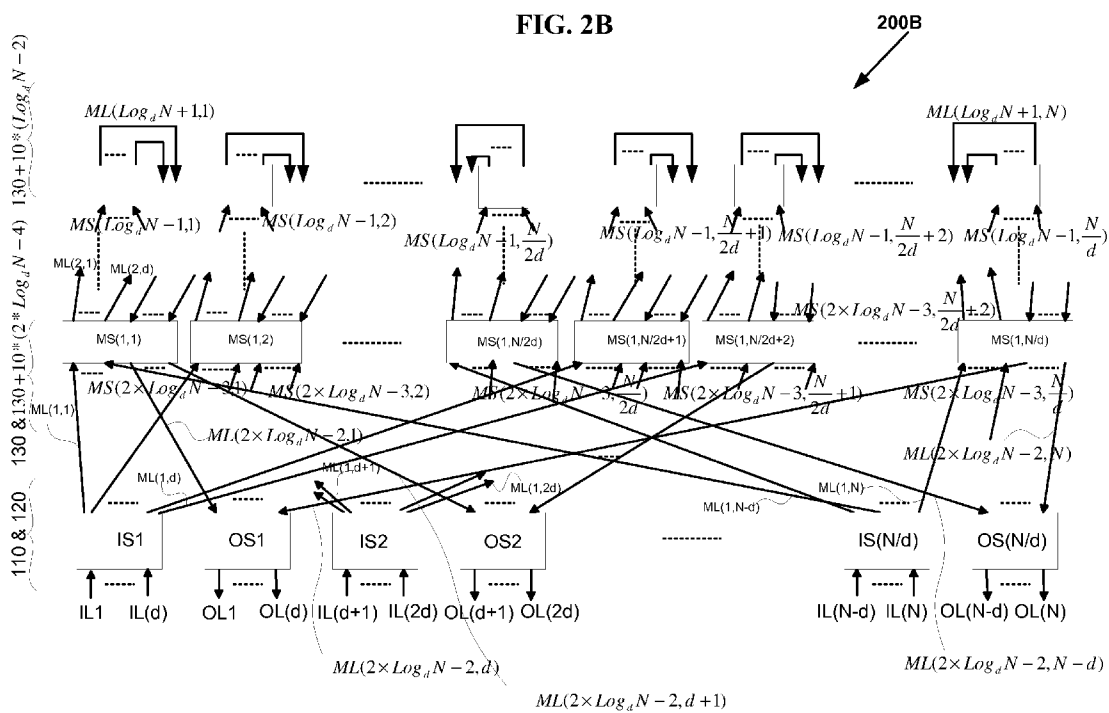


FIG. 2C

200C

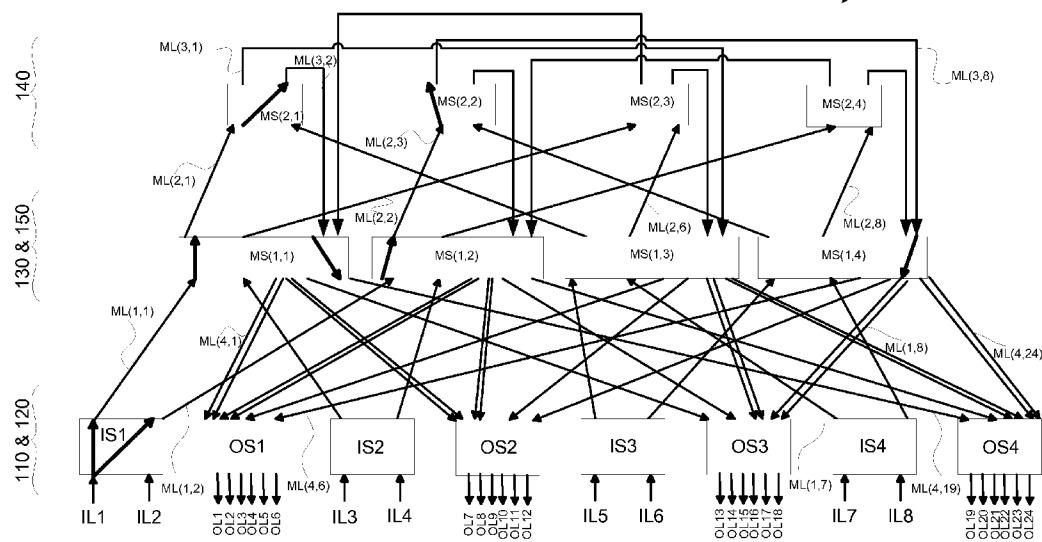


FIG. 2D

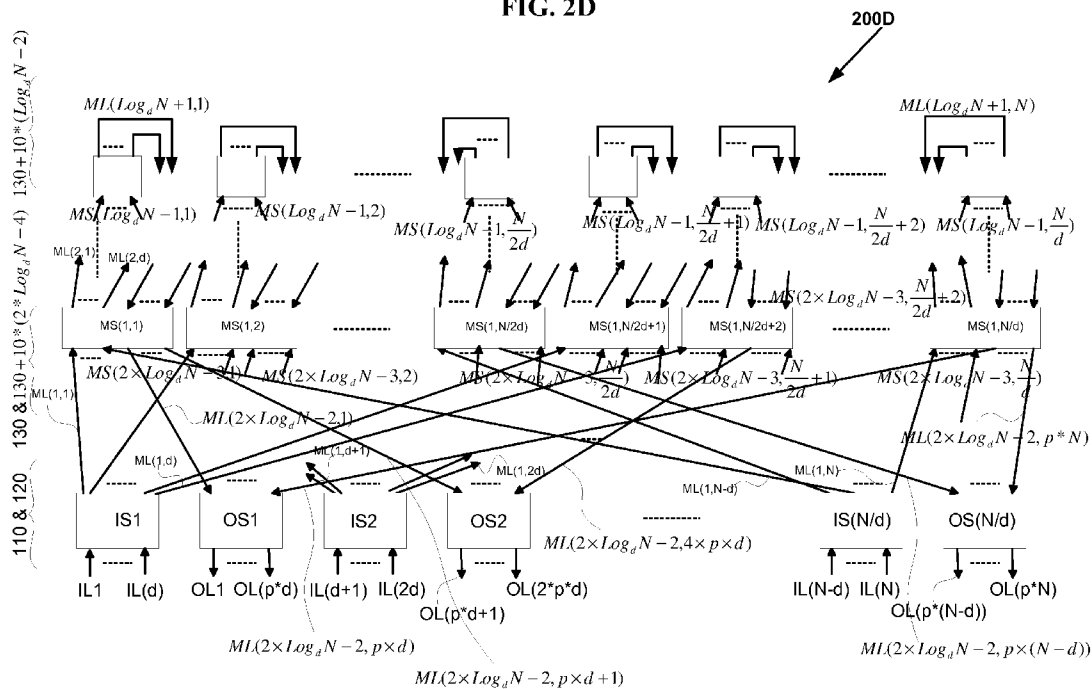


FIG. 2E

200E

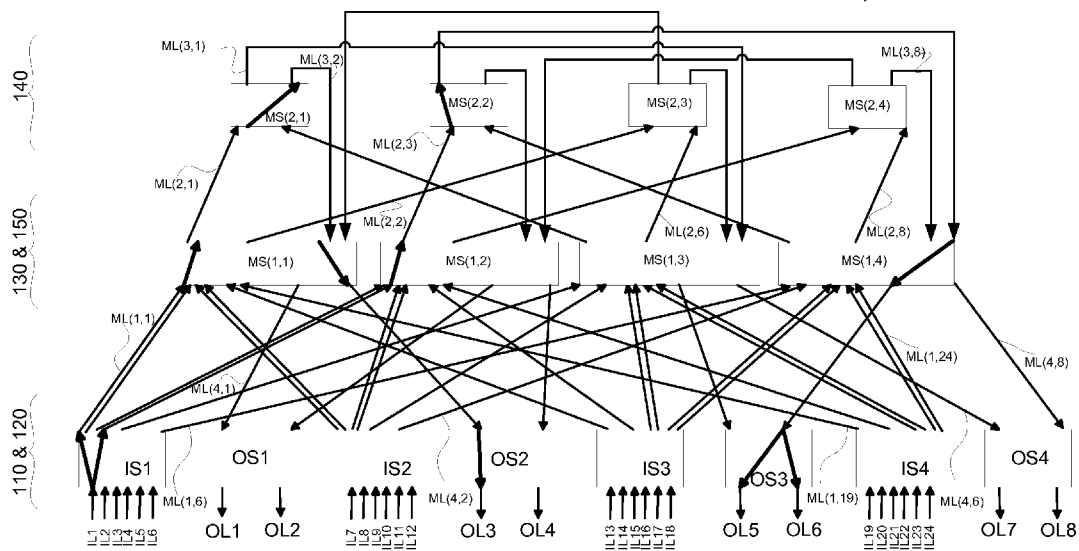


FIG. 2F

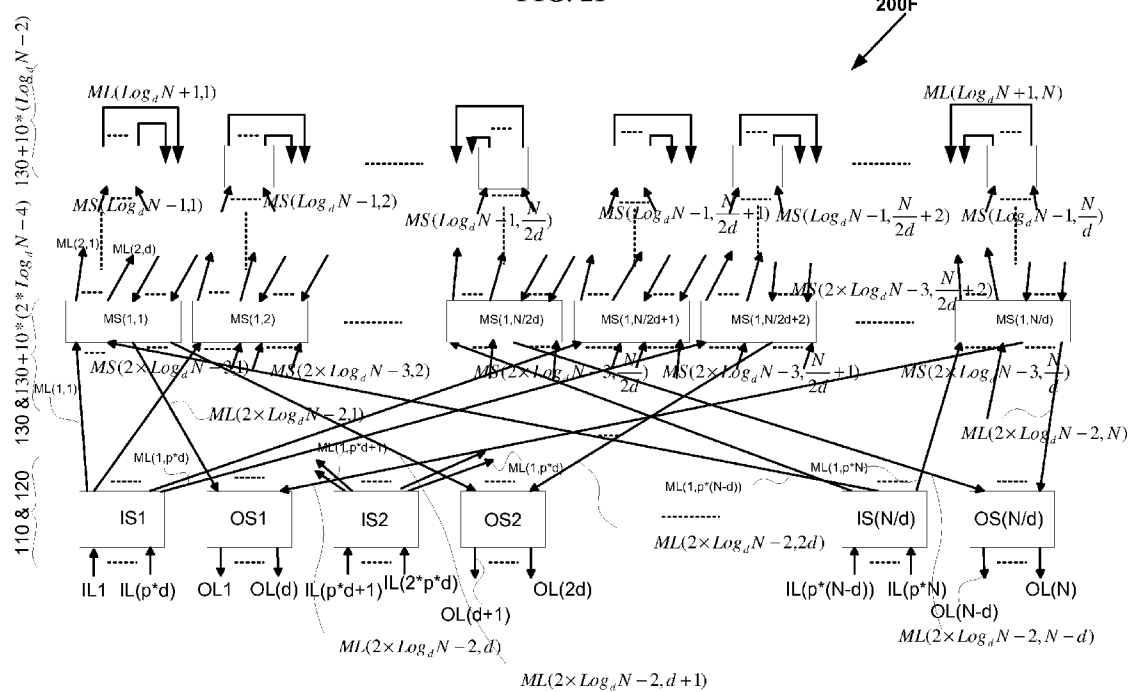


FIG. 3A

300A

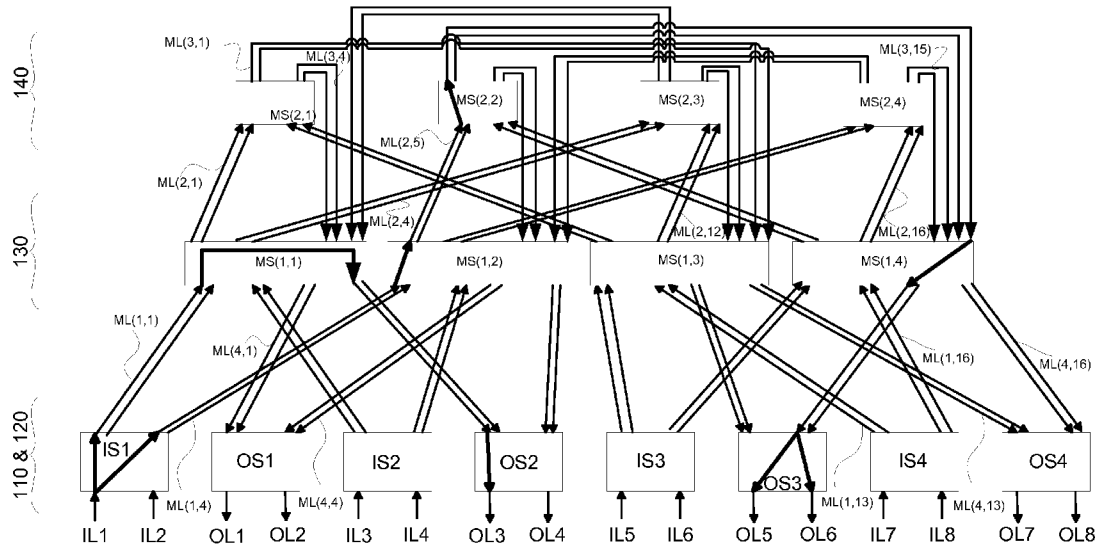


FIG. 3B

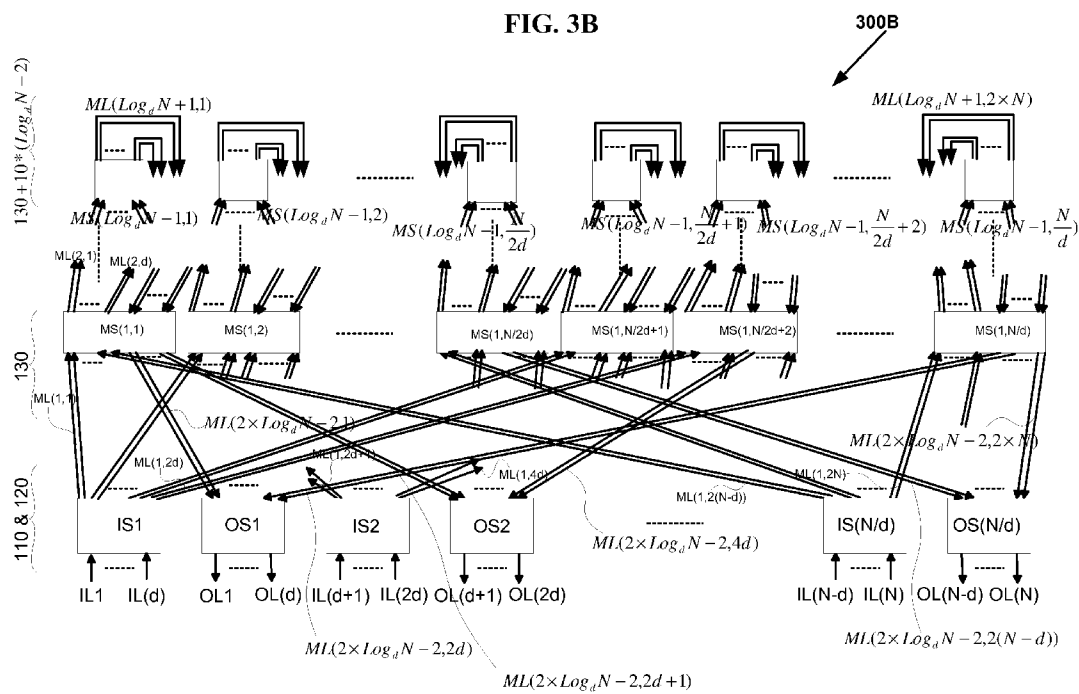


FIG. 3C

300C

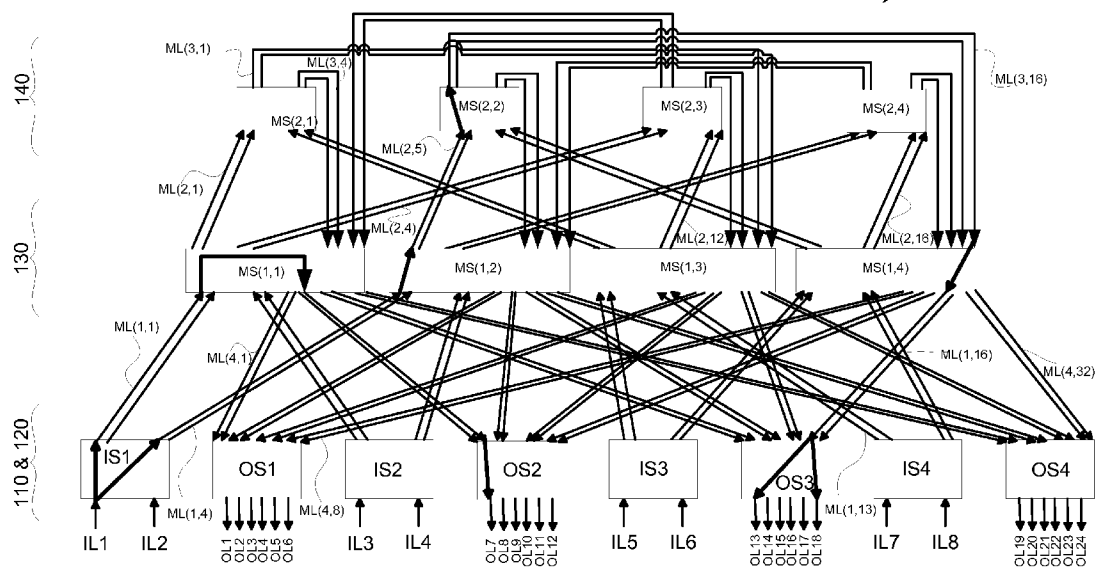


FIG. 3D

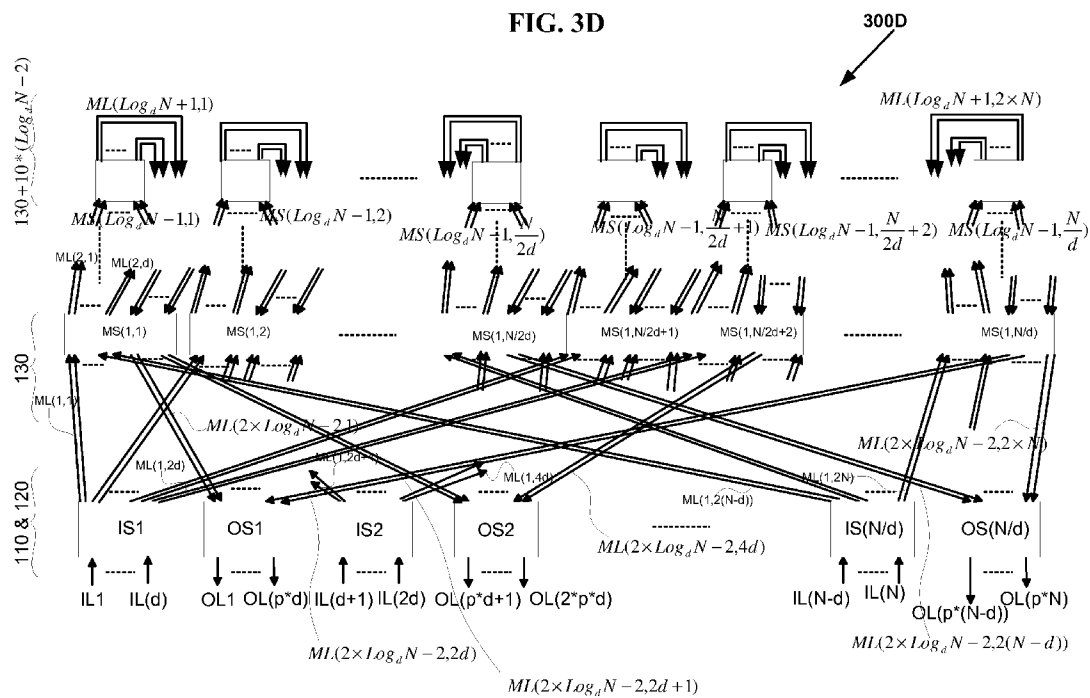


FIG. 3E

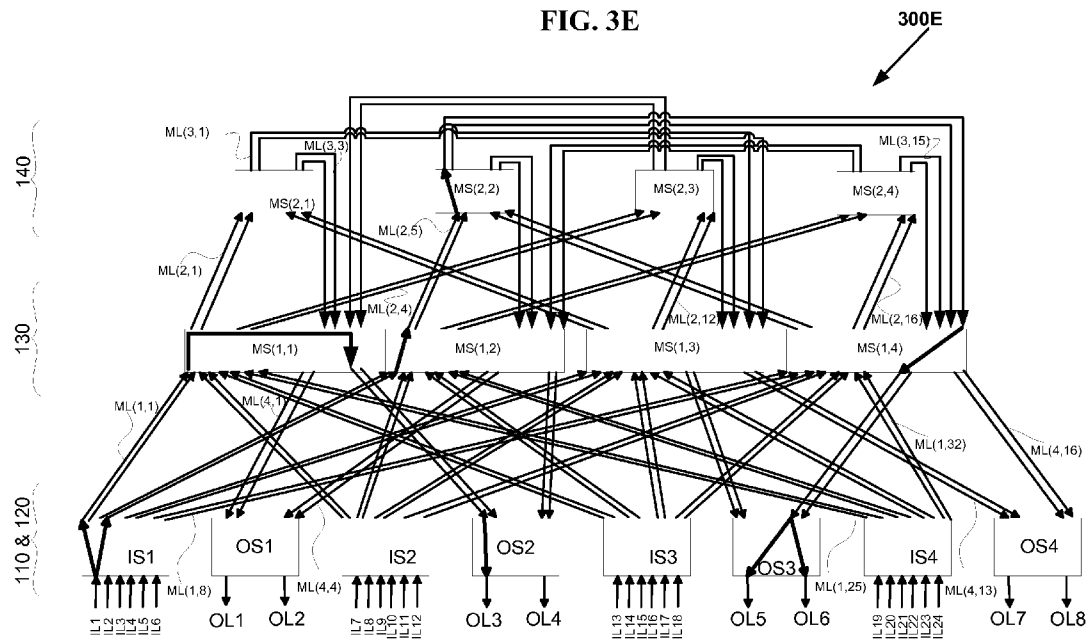


FIG. 3F

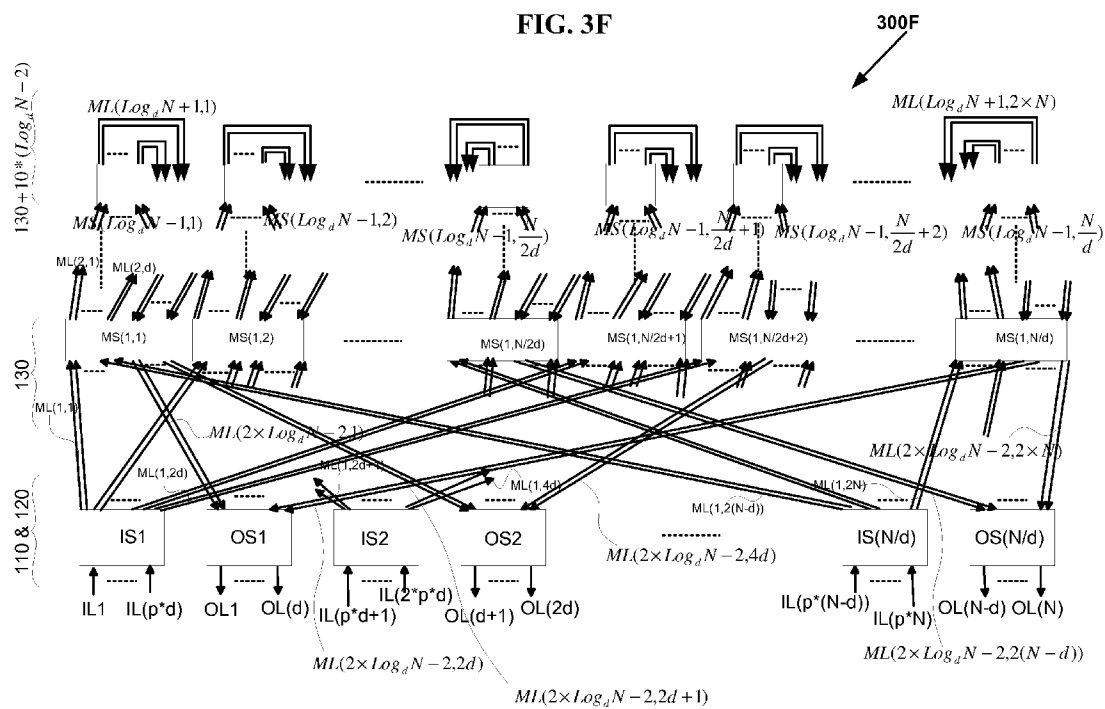


FIG. 4A

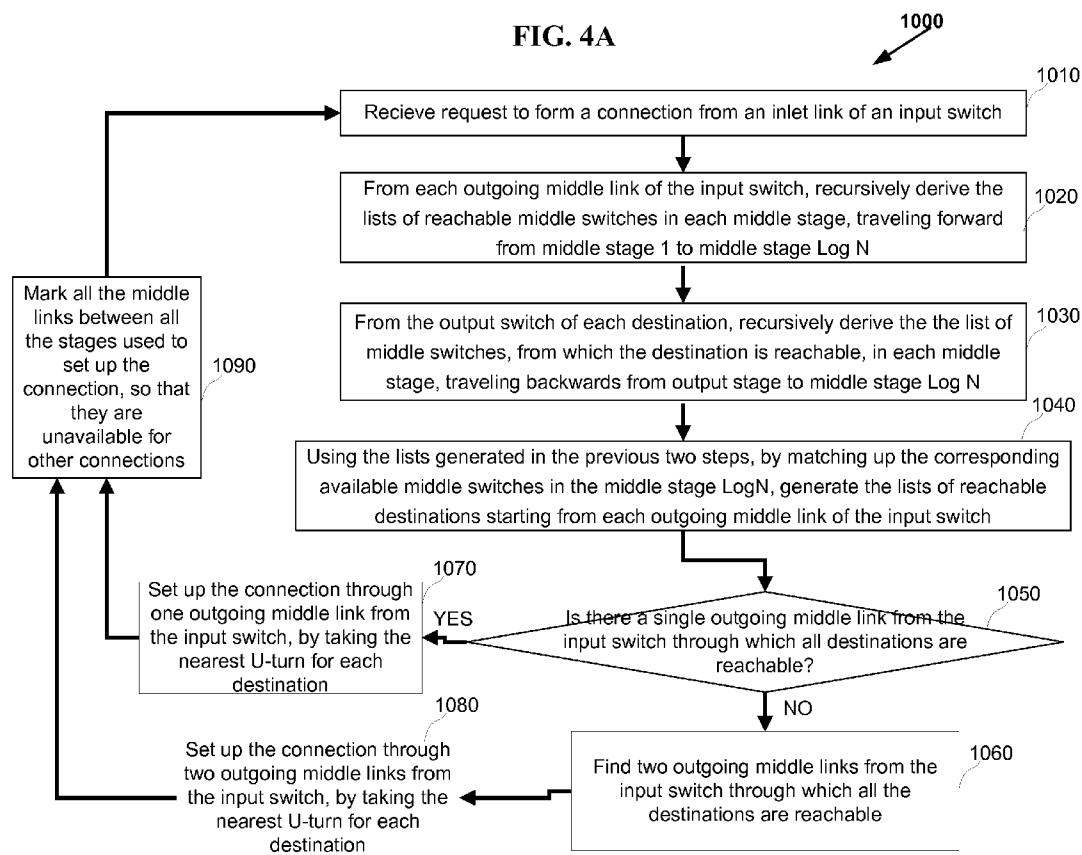
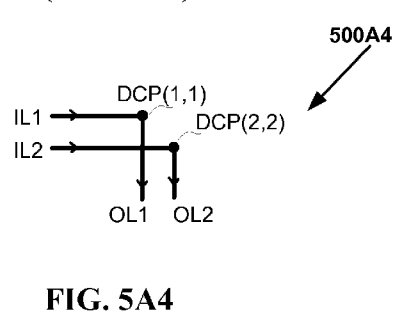
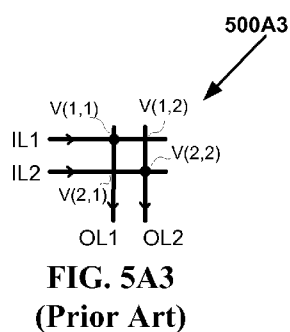
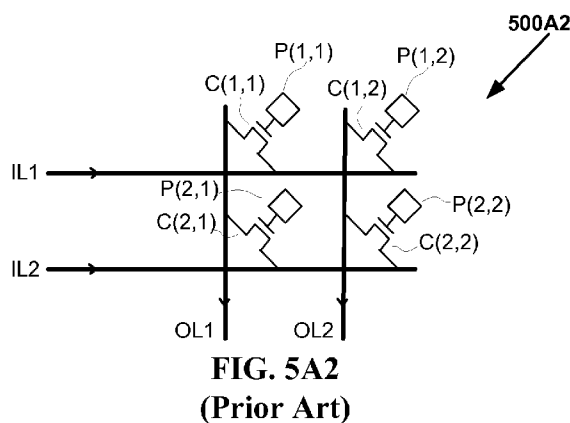
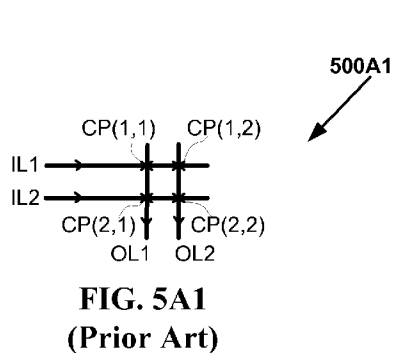


FIG. 5A



US 2010/0172349 A1

Jul. 8, 2010

1

FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is related to and claims priority of PCT Application Serial No. PCT/U.S.08/64603 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.

[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/U.S.08/56064 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-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/U.S.08/64604 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.

[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/U.S.08/64605 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 2010/0172349 A1

Jul. 8, 2010

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 $2 \leq x \leq \min(n-1, r)$, 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 butterfly fat tree network comprising $(\log_d N)$ stages is operated in strictly nonblocking manner for unicast includes a leaf stage consisting of an input stage having N/d switches with each of them having d inlet links and $2 \times d$ outgoing links connecting to its immediate succeeding stage switches, and 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 its immediate succeeding stage. The network also has $(\log_d N) - 1$ middle stages with each middle stage, excepting the root stage, having

$$\frac{2 \times N}{d}$$

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, d incoming links connecting from the switches in its immediate succeeding stage, d outgoing links connecting to the switches in its immediate succeeding stage, d outgoing links connecting to the switches in its immediate preceding stage, and the root stage having

$$\frac{2 \times N}{d}$$

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage and d outgoing links connecting to the switches in its immediate preceding stage. Also the same generalized butterfly fat tree 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 butterfly fat tree network comprising $(\log_d N)$ stages is operated in strictly nonblocking manner for multicast includes a leaf stage consisting of an input stage having N/d switches with each of them having d inlet links and $3 \times d$ outgoing links connecting to its immediate succeeding 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 its immediate succeeding stage. The network also has $(\log_d N) - 1$ middle stages with each middle stage, excepting the root stage, having

$$\frac{3 \times N}{d}$$

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage, d incoming links connecting from the switches in its immediate succeeding stage, d outgoing links connecting to the switches in its immediate succeeding stage, d outgoing links connecting to the switches in its immediate preceding stage, and the root stage having

$$\frac{3 \times N}{d}$$

switches, and each switch in the middle stage has d incoming links connecting from the switches in its immediate preceding stage and d outgoing links connecting to the switches in its immediate preceding stage.

BRIEF DESCRIPTION OF DRAWINGS

[0016] FIG. 1A is a diagram 100A of an exemplary Symmetrical Butterfly fat tree network $V_{bft}(N, d, s)$ having inverse

US 2010/0172349 A1

Jul. 8, 2010

3

Benes connection topology of three 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.

[0017] FIG. 1B is a diagram 100B of a general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ with $(\log_d N)$ stages 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 Asymmetrical Butterfly fat tree network $V_{bft}(N_1,N_2,d,2)$ having inverse Benes connection topology of three 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.

[0019] FIG. 1D is a diagram 100D of a general asymmetrical Butterfly fat tree network $V_{bft}(N_1,N_2,d,2)$ with $N_2=p*N_1$ and with $(\log_d N)$ stages 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 Asymmetrical Butterfly fat tree network $V_{bft}(N_1,N_2,d,2)$ having inverse Benes connection topology of three 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.

[0021] FIG. 1F is a diagram 100F of a general asymmetrical Butterfly fat tree network $V_{bft}(N_1,N_2,d,2)$ with $N_1=p*N_2$ and with $(\log_d N)$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0022] FIG. 2A is a diagram 200A of an exemplary Symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ having inverse Benes connection topology of three stages with $N=8$, $d=2$ and $s=1$ with exemplary unicast connections rearrangeably nonblocking network for unicast connections, in accordance with the invention.

[0023] FIG. 2B is a diagram 200B of a general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ with $(\log_d N)$ stages and $s=1$, rearrangeably nonblocking network for unicast connections in accordance with the invention.

[0024] FIG. 2C is a diagram 200C of an exemplary Asymmetrical Butterfly fat tree network $V_{bft}(N_1,N_2,d,1)$ having inverse Benes connection topology of three 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.

[0025] FIG. 2D is a diagram 200D of a general asymmetrical Butterfly fat tree network $V_{bft}(N_1,N_2,d,1)$ with $N_2=p*N_1$ and with $(\log_d N)$ stages rearrangeably nonblocking network for unicast connections in accordance with the invention.

[0026] FIG. 2E is a diagram 200E of an exemplary Asymmetrical Butterfly fat tree network $V_{bft}(N_1,N_2,d,1)$ having inverse Benes connection topology of three 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.

[0027] FIG. 2F is a diagram 200F of a general asymmetrical Butterfly fat tree network $V_{bft}(N_1,N_2,d,1)$ with $N_1=p*N_2$ and with $(\log_d N)$ stages rearrangeably nonblocking network for unicast connections in accordance with the invention.

[0028] FIG. 3A is a diagram 300A of an exemplary symmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N,d,s)$ 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.

[0029] FIG. 3B is a diagram 300B of a general symmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N,d,2)$ with $(\log_d N)$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0030] FIG. 3C is a diagram 300C of an exemplary asymmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N_1,N_2,d,2)$ 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.

[0031] FIG. 3D is a diagram 300D of a general asymmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N_1,N_2,d,2)$ with $N_2=p*N_1$ and with $(\log_d N)$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0032] FIG. 3E is a diagram 300E of an exemplary asymmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N_1,N_2,d,2)$ 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.

[0033] FIG. 3F is a diagram 300F of a general asymmetrical multi-link Butterfly fat tree network $V_{mlink-bft}(N_1,N_2,d,2)$ with $N_1=p*N_2$ and with $(\log_d N)$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

[0034] FIG. 4A 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.

[0035] 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

[0036] The present invention is concerned with the design and operation of large scale crosspoint reduction using arbitrarily large Butterfly fat tree networks and Multi-link Butterfly fat tree networks for broadcast, unicast and multicast

US 2010/0172349 A1

Jul. 8, 2010

4

connections. Particularly Butterfly fat tree networks and Multi-link Butterfly fat tree networks with stages more than or equal to three and radices greater than or equal to two offer large scale crosspoint reduction when configured with optimal links as disclosed in this invention.

[0037] When a transmitting device simultaneously sends information to more than one receiving device, the one-to-many connection required between the transmitting device 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.

[0038] 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.

[0039] In certain butterfly fat tree networks and multi-link butterfly fat tree 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 Butterfly fat tree 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.

[0040] In certain butterfly fat tree networks and multi-link butterfly fat tree 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 Butterfly fat tree 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.

[0041] Nonblocking configurations for other types of networks with numerous connection topologies and scheduling methods are disclosed as follows:

[0042] 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.

[0043] 2) Rearrangeably nonblocking for arbitrary fan-out multicast and unicast, and strictly nonblocking for unicast for generalized multi-link multi-stage networks $V_{mink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the PCT Application Serial No. PCT/U.S.08/64604 that is incorporated by reference above.

[0044] 3) 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 con-

nection topologies and the scheduling methods are described in detail in the PCT Application Serial No. PCT/U.S.08/64604 that is incorporated by reference above.

[0045] 4) Strictly nonblocking for arbitrary fan-out multicast for generalized multi-link multi-stage networks $V_{mink}(N_1, N_2, d, s)$ and generalized folded multi-link multi-stage networks $V_{fold-mink}(N_1, N_2, d, s)$ with numerous connection topologies and the scheduling methods are described in detail in the PCT Application Serial No. PCT/U.S.08/64604 that is incorporated by reference above.

[0046] 5) 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 $V_{mink}(N_1, N_2, d, s)$, 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, are described in detail in the PCT Application Serial No. PCT/U.S.08/64605 that is incorporated by reference above.

[0047] 6) 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.

[0048] 7) 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.

Butterfly Fat Tree Embodiments:

Symmetric RNB Embodiments:

[0049] Referring to FIG. 1A, in one embodiment, an exemplary symmetrical butterfly fat tree network **100A** with three stages of twenty 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**, and **140** 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. Input stage **110** and output stage **120** together belong to leaf stage. And all the middle stages excepting root stage namely middle stage **130** consists of eight, four by four switches MS(1,1)-MS(1,8), and root stage i.e., middle stage **140** consists of eight, two by two switches MS(2,1)-MS(2,8).

[0050] 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** and middle stage **140**. 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** and middle stage **140**.

[0051] 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

US 2010/0172349 A1

Jul. 8, 2010

5

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. Input stage **110** and output stage **120** together belong to leaf stage. 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 $2d*2d$ excepting that the size of each switch in middle stage **140** is denoted as $d*d$. (In another embodiment, the size of each switch in any of the middle stages other than the middle stage **140**, can be implemented as $d*2d$ and $d*d$ since the down coming middle links are never setup to the up going middle links For example in network **100A** of FIG. 1A, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

[0052] Middle stage **140** is called as root stage. 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 Butterfly fat tree network can be represented with the notation $V_{bft}(N,d,s)$, 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.

[0053] 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).

[0054] 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 are also 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(1,1) from middle switches MS(2,1) and MS(2,3) respectively).

[0055] Each of the

$$2 \times \frac{N}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage **130** **[0056]** 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) 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(1,1)).

[0057] 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 **130** 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(1,1) and MS(1,3) respectively).

[0058] Each of the N/d output switches OS1-OS4 are connected from exactly $2 \times d$ switches in middle stage **130** through $2 \times d$ links (for example output switch OS1 is connected from middle switches MS(1,1), MS(1,2), MS(1,5) and MS(1,6) through the links ML(4,1), ML(4,3), ML(4,9) and ML(4,11) respectively).

[0059] Finally the connection topology of the network **100A** shown in FIG. 1A is known to be back to back inverse Benes connection topology.

[0060] In the three embodiments of FIG. 1A, FIG. 1A1 and FIG. 1A2 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 $V_{bft}(N,d,s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$ can be built. The embodiments of FIG. 1A, FIG. 1A1, and FIG. 1A2 are only three examples of network $V_{bft}(N,d,s)$.

[0061] In the three embodiments of FIG. 1A, FIG. 1A1 and FIG. 1A2, 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) and

US 2010/0172349 A1

Jul. 8, 2010

6

MS(2,1)-MS(2,8) 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(1,2)-MS(2,8) are referred to as root stage switches.

[0062] In the example illustrated in FIG. 1A (or in FIG. 1A1, or in FIG. 1A2), 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 **100A1**, or **100A2**), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0063] 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:

[0064] Network **100B** of FIG. 1B is an example of general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ with $(\log_d N)$ stages. The general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention. (And in the example of FIG. 1B, $s=2$). The general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ with $(\log_d N)$ stages has d inlet links for each 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).

[0065] 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(1,2 d) respectively.

[0066] Each of the

$$2 \times \frac{N}{d}$$

middle switches MS(1,1)-MS(1,2 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.

[0067] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches MS(1,1)-MS(1,2 N/d) in the middle stage **130** are also connected from exactly d switches in middle stage **140** through d links and also are connected to exactly d output switches in output stage **120** through d links.

[0068] Similarly each of the

$$2 \times \frac{N}{d}$$

middle switches

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

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.

[0069] Each of the N/d output switches OS1-OS(N/d) are connected from exactly $2 \times d$ switches in middle stage **130** through $2 \times d$ links.

[0070] As described before, again the connection topology of a general $V_{bft}(N,d,s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$ can be built. The embodiments of FIG. 1A, FIG. 1A1, and FIG. 1A2 are three examples of network $V_{bft}(N,d,s)$.

[0071] The general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0072] Every switch in the Butterfly fat tree networks discussed herein has multicast capability. In a $V_{bft}(N,d,s)$ 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

US 2010/0172349 A1

Jul. 8, 2010

7

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 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.

[0073] To characterize a multicast assignment, for each inlet link

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

let $I_i=0$, 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. 1A shows an exemplary three-stage network, namely $V_{bft}(8,2,2)$, 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 output switch OS2 in output stage 120 and middle switch MS(2,7) in middle stage 140 respectively.

[0074] The connection I_1 also fans out in middle switch MS(2,7) 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 switch MS(1,7) only once into output switch 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 RNB ($N_2 > N_1$) Embodiments:

[0075] Referring to FIG. 1C, in one embodiment, an exemplary asymmetrical Butterfly fat tree network 100C with three stages of twenty 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 and 140 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. Input stage 110 and

output stage 120 together belong to leaf stage. Middle stage 130 consists of eight, four by six switches MS(1,1)-MS(1,8) and middle stage 140 consists of eight, two by two switches MS(2,1)-MS(2,8).

[0076] 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 and middle stage 140. 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 four by six in middle stage 130 and eight switches of size two by two in middle stage 140.

[0077] 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

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

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 middle stage 130 can be denoted as $2d * (d+d_2)$. The size of each switch in the root stage (i.e., middle stage 140) can be denoted as $d * d$. The size of each switch in all the middle stages excepting middle stage 130 and root stage can be denoted as $2d * 2d$ (In network 100C of FIG. 1C, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $d * 2d$ and $d * d$ since the down coming middle links are never setup to the up going middle links. For example in network 100C of FIG. 1C, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

[0078] 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

US 2010/0172349 A1

Jul. 8, 2010

8

Butterfly fat tree network can be represented with the notation $V_{bft}(N_1, N_2, d, s)$, 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.

[0079] 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).

[0080] 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 are also 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(1,1) from middle switches MS(2,1) and MS(2,3) respectively).

[0081] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 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) 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 switch MS(1,1)).

[0082] 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 130 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(1,1) and MS(1,3) respectively).

[0083] Each of the

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

output switches OS1-OS4 are connected from exactly $d+d_2$ switches in middle stage 130 through $d+d_2$ links (for example output switch OS1 is connected from 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(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).

[0084] Finally the connection topology of the network 100C shown in FIG. 1C is known to be back to back inverse Benes connection topology.

[0085] In the three embodiments of FIG. 1C, FIG. 1C1 and FIG. 1C2 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(4,32) are connected between the respective stages is different. Even though only three embodiments are illustrated, in general, the network $V_{bft}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1C, FIG. 1C1, and FIG. 1C2 are only three examples of network $V_{bft}(N_1, N_2, d, s)$.

[0086] In the three embodiments of FIG. 1C, FIG. 1C1 and FIG. 1C2, 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(4,32) 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) and MS(2,1)-MS(2,8) are referred to as middle switches or middle ports.

[0087] In the example illustrated in FIG. 1C (or in FIG1C1, or in FIG. 1C2), 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 connec-

US 2010/0172349 A1

Jul. 8, 2010

9

tion 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 100C (or 100C1, or 100C2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0088] 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:

[0089] Network 100D of FIG. 1D is an example of general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_2 > N_1$, and $N_2 = p * N_1$ where $p > 1$. In network 100D of FIG. 1D, $N_1 = N$ and $N_2 = p * N$. The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention. (And in the example of FIG. 1D, $s=2$). The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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).

[0090] 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.

[0091] Each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,2 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.

[0092] Similarly each of the

$$2 \times \frac{N_1}{d}$$

middle switches MS(1,1)-MS(1,2 N_1/d) in the middle stage 130 are connected from exactly d switches in middle stage 140 through d links and also are connected to exactly d output switches in output stage 120 through

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

links.

[0093] Similarly each of the

$$2 \times \frac{N_1}{d}$$

US 2010/0172349 A1

Jul. 8, 2010

10

middle switches

$$MS(\text{Log}_d N_1 - 1, 1) - MS(\text{Log}_d N_1 - 1, 2 \times \frac{N_1}{d})$$

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.

[0094] 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** through $d+d_2$ links.

[0095] As described before, again the connection topology of a general $V_{bft}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1C, FIG. 1C1, and FIG. 1C2 are three examples of network $V_{bft}(N_1, N_2, d, s)$ for $s=2$ and $N_2 > N_1$.

[0096] The general symmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0097] For example, the network of FIG. 1C shows an exemplary three-stage network, namely $V_{bft}(8, 24, 22)$, 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 OS2 and OS3 respectively in output stage **120**.

[0098] 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 OL18. 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:

[0099] Referring to FIG. 1E, in one embodiment, an exemplary asymmetrical Butterfly fat tree network **100E** with three stages of twenty 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** and **140** 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. Middle stage **130** consists of eight, six by four switches MS(1,1)-MS(1,8) and middle stage **140** consists of eight, two by two switches MS(2,1)-MS(2,8).

[0100] 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** and middle stage **140**. 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 six by four in middle stage **130**, and eight switches of size two by two in middle stage **140**.

[0101] 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

$$2 \times \frac{N_2}{d}.$$

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 middle stage **130** can be denoted as $(d+d_1)*2d$. The size of each switch in the root stage (i.e., middle stage **140**) can be denoted as $d*d$. The size of each switch in all the middle stages excepting middle stage **130** and root stage can be denoted as $2d*2d$ (In network **100E** of FIG. 1E, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage **140**, can be implemented as $d*2d$ and $d*d$ since the down coming middle links are never setup to the up going middle links. For example in network **100E** of FIG. 1E, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

[0102] 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

US 2010/0172349 A1

Jul. 8, 2010

11

Butterfly fat tree network can be represented with the notation $V_{bft}(N_1, N_2, d, s)$, 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 is the ratio of number of incoming links to each output switch to the outlet links of each output switch.

[0103] 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).

[0104] 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 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(1,1) from middle switches MS(2,1) and MS(2,3) respectively).

[0105] Similarly each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(1,1)-MS(1,8) in the middle stage 130 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), 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(1,1)).

[0106] 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 130 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(1,3) and MS(1,1) respectively).

[0107] Each of the

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

output switches OS1-OS4 are connected from exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links (for example output switch OS1 is connected from middle switches MS(1,1), MS(1,2), MS(1,5), and MS(1,6) through the links ML(4,1), ML(4,3), ML(4,9), and ML(4,11) respectively).

[0108] Finally the connection topology of the network 100E shown in FIG. 1E is known to be back to back inverse Benes connection topology.

[0109] In the three embodiments of FIG. 1E, FIG. 1E1 and FIG. 1E2 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 $V_{bft}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1E, FIG. 1E1, and FIG. 1E2 are only three examples of network $V_{bft}(N_1, N_2, d, s)$.

[0110] In the three embodiments of FIG. 1E, FIG. 1E1 and FIG. 1E2, 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) and MS(2,1)-MS(2,8) are referred to as middle switches or middle ports.

[0111] In the example illustrated in FIG. 1E (or in FIG. 1E1, or in FIG. 1E2), 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,

US 2010/0172349 A1

Jul. 8, 2010

12

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 100E (or 100E1, or 100E2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0112] 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 ($N_1 > N_2$) Embodiments:

[0113] Network 100F of FIG. 1F is an example of general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network 100F of FIG. 1F, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention. (And in the example of FIG. 1F, $s=2$). The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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 $d+d_1$ ($=d+p*d$) 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 $2 \times d$ 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).

[0114] 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.

[0115] Each of the

$$2 \times \frac{N_2}{d}$$

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

[0116] Similarly each of the

$$2 \times \frac{N_2}{d}$$

middle switches MS(1,1)-MS(1, $2 * N_2/d$) in the middle stage 130 also are connected to exactly d switches in middle stage 140 through d links and also are connected to exactly d output switches in output stage 120 through d links.

[0117] 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+ $10 * (\log_d N_2 - 2)$ are connected from exactly d switches in middle stage 130+ $10 * (\log_d N_2 - 3)$

US 2010/0172349 A1

Jul. 8, 2010

13

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

[0118] 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 $130+10*(2*\text{Log}_d N_2-4)$ through $2 \times d$ links.

[0119] As described before, again the connection topology of a general $V_{bft}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 1E, FIG. 1E1, and FIG. 1E2 are three examples of network $V_{bft}(N_1, N_2, d, s)$ for $s=2$ and $N_1 > N_2$.

[0120] The general symmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0121] For example, the network of FIG. 1E shows an exemplary three-stage network, namely $V_{bft}(24, 8, 2, 2)$, 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 output switch OS2 in output stage 120 and middle switch and MS(2,7) in middle stage 140 respectively.

[0122] The connection I_1 also fans out in middle switch MS(2,7) only once into middle switch MS(1,7) in middle stage 130. The connection I_1 also fans out in middle switch MS(1,7) only once into output switch 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.

Strictly Nonblocking Butterfly Fat Tree Networks:

[0123] The general symmetric Butterfly fat tree network $V_{bft}(N, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention. Similarly the general asymmetric Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention.

Symmetric RNB Unicast Embodiments:

[0124] Referring to FIG. 2A, in one embodiment, an exemplary symmetrical Butterfly fat tree network 200A with three

stages of sixteen 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, and 140 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. Input stage 110 and output stage 120 together belong to leaf stage. And all the middle stages excepting root stage namely middle stage 130 consists of four, four by four switches MS(1,1)-MS(1,4), and root stage i.e., middle stage 140 consists of four, two by two switches MS(2,1)-MS(2,4).

[0125] Such a network can be operated in rearrangeably non-blocking 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 and middle stage 140.

[0126] 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 $d \times d$ and each output switch OS1-OS4 can be denoted in general with the notation $d \times d$. Likewise, the size of each switch in any of the middle stages can be denoted as $2d \times 2d$ excepting that the size of each switch in middle stage 140 is denoted as $d \times d$. (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $d \times 2d$ and $d \times d$ since the down coming middle links are never setup to the up going middle links. For example in network 200A of FIG. 2A, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

[0127] Middle stage 140 is called as root stage. 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 Butterfly fat tree network can be represented with the notation $V_{bft}(N, d, s)$, 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.

[0128] Each of the N/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 link ML(1,1); and input switch IS1 is also connected to middle switch MS(1,2) through the link ML(1,2)).

[0129] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected from exactly d input

US 2010/0172349 A1

Jul. 8, 2010

14

switches through d links (for example the link $ML(1,1)$ is connected to the middle switch $MS(1,1)$ from input switch $IS1$; and the link $ML(1,3)$ is connected to the middle switch $MS(1,1)$ from input switch $IS2$) and are also connected from exactly d switches in middle stage **140** through d links (for example the link $ML(3,2)$ is connected to the middle switch $MS(1,1)$ from middle switch $MS(2,1)$ and also the link $ML(3,5)$ is connected to the middle switch $MS(1,1)$ from middle switch $MS(2,3)$).

[0130] Each of the N/d middle switches $MS(1,1)$ - $MS(1,4)$ in the middle stage **130** are connected to exactly d switches in middle stage **140** through d links (for example the link $ML(2,1)$ is connected from middle switch $MS(1,1)$ to middle switch $MS(2,1)$, and the link $ML(2,2)$ is connected from middle switch $MS(1,1)$ to middle switch $MS(2,3)$) and also are connected to exactly d output switches in output stage **120** through d links (for example the link $ML(4,1)$ is connected to output switch $OS1$ from middle switch $MS(1,1)$, and the link $ML(4,2)$ is connected to output switch $OS2$ from middle switch $MS(1,1)$).

[0131] 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 link $ML(2,1)$ is connected to the middle switch $MS(2,1)$ from middle switch $MS(1,1)$, and the link $ML(2,5)$ is 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 **130** through d links (for example the link $ML(3,1)$ is connected from middle switch $MS(2,1)$ to middle switch $MS(1,3)$; and the link $ML(3,2)$ is connected from middle switch $MS(2,1)$ to middle switch $MS(1,1)$).

[0132] Each of the N/d output switches $OS1$ - $OS4$ are connected from exactly d switches in middle stage **130** through d links (for example output switch $OS1$ is connected from middle switch $MS(1,1)$ through the link $ML(4,1)$; and output switch $OS1$ is also connected from middle switch $MS(1,2)$ through the link $ML(4,2)$).

[0133] Finally the connection topology of the network **200A** shown in FIG. 2A is known to be back to back inverse Benes connection topology.

[0134] In other embodiments the connection topology may be different from the network **200A** of FIG. 2A. That is the way the links $ML(1,1)$ - $ML(1,8)$, $ML(2,1)$ - $ML(2,8)$, $ML(3,1)$ - $ML(3,8)$, and $ML(4,1)$ - $ML(4,8)$ are connected between the respective stages is different. Even though only one embodiment is illustrated, in general, the network $V_{bft}(N,d,s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$ can be built. The embodiment of FIG. 2A is only one example of network $V_{bft}(N,d,s)$.

[0135] In the embodiment of FIG. 2A each of 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 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.

Generalized Symmetric RNB Unicast Embodiments:

[0136] Network **200B** of FIG. 2B is an example of general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ with $(\log_d N)$ stages. The general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ can be operated in rearrangeably nonblocking manner for unicast when $s=1$ according to the current invention (and in the example of FIG. 2B, $s=1$). The general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ with $(\log_d N)$ 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 the links $OL1$ - $OL(d)$ 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 $OS1$).

[0137] Each of the N/d input switches $IS1$ - $IS(N/d)$ are connected to exactly d switches in middle stage **130** through d links.

[0138] 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.

[0139] Similarly each of the N/d middle switches $MS(1,1)$ - $MS(1,N/d)$ in the middle stage **130** are also connected from exactly d switches in middle stage **140** through d links and also are connected to exactly d output switches in output stage **120** through d links.

[0140] 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 d links and also are connected to exactly d switches in middle stage **130**+ $10*(\log_d N - 1)$ through d links.

[0141] Each of the N/d output switches $OS1$ - $OS(N/d)$ are connected from exactly d switches in middle stage **130** through d links.

[0142] As described before, again the connection topology of a general $V_{bft}(N,d,s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$ can be built. The embodiment of FIG. 2A are one example of network $V_{bft}(N,d,s)$.

US 2010/0172349 A1

Jul. 8, 2010

15

[0143] The general symmetrical Butterfly fat tree network $V_{bft}(N,d,s)$ is operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention.

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

[0144] Referring to FIG. 2C, in one embodiment, an exemplary asymmetrical Butterfly fat tree network **200C** with three stages of sixteen 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** and **140** 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. Middle stage **130** consists of four, four by eight switches MS(1,1)-MS(1,4) and middle stage **140** consists of four, two by two switches MS(2,1)-MS(2,4).

[0145] Such a network can be operated in rearrangeably non-blocking 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 of size four by eight in middle stage **130** and four switches of size two by two in middle stage **140**.

[0146] 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 $d * d$ and each output switch OS1-OS4 can be denoted in general with the notation $d_2 * d_2$, where

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

[0147] The size of each switch in middle stage **130** can be denoted as $2d * (d + d_2)$. The size of each switch in the root stage (i.e., middle stage **140**) can be denoted as $d * d$. The size of each switch in all the middle stages excepting middle stage **130** and root stage can be denoted as $2d * 2d$ (In network **200C** of FIG. 2C, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage **140**, can be implemented as $d * 2d$ and $d * d$ since the down coming middle links are never setup to the up going middle links. For example in network **200C** of FIG. 2C, the down coming middle links ML(3,2) and ML(3,5) are never setup to the up going middle links ML(2,1) and ML(2,2) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a two by four switch

with middle links ML(1,1) and ML(1,3) as inputs and middle links ML(2,1), ML(2,2), ML(4,1) and ML(4,2) as outputs; and a two by two switch with middle links ML(3,2) and ML(3,5) as inputs and middle links ML(4,1) and ML(4,2) as outputs).

[0148] 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 Butterfly fat tree network can be represented with the notation $V_{bft}(N_1, N_2, d, s)$, 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.

[0149] 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 switch MS(1,1) through the link ML(1,1), and input switch IS1 is also connected to MS(1,2) through the link ML(1,2)).

[0150] 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 link ML(1,1) is connected to the middle switch MS(1,1) from input switch IS1 and the link ML(1,3) is connected to the middle switch MS(1,1) from input switch IS2) and are also connected from exactly d switches in middle stage **140** through d links (for example the link ML(3,2) is connected to the middle switch MS(1,1) from middle switch MS(2,1), and the link ML(3,5) is connected to the middle switch MS(1,1) from middle switch MS(2,3)).

[0151] Similarly each of the

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

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

$$\frac{d_2}{2}$$

output switches in output stage **120** through d_2 links (for example the link ML(4,1) and ML(4,2) are connected from middle switch MS(1,1) to output switch OS1; the links ML(4,

US 2010/0172349 A1

Jul. 8, 2010

16

3) and ML(4,4) are connected from middle switch MS(1,1) to output switch OS2; the link ML(4,5) is connected from middle switch MS(1,1) to output switch OS3; and the links ML(4,6) is connected from middle switch MS(1,1) to output switch OS4).

[0152] 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 link ML(2,1) is connected to the middle switch MS(2,1) from middle switch MS(1,1); and the link ML(2,5) is 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 130 through d links (for example the link ML(3,2) is connected from middle switch MS(2,1) to middle switch MS(1,1); and the link ML(3,1) is connected from middle switch MS(2,1) to middle switch MS(1,3)).

[0153] Each of the

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

output switches OS1-OS4 are connected from exactly

$$\frac{d_2}{2}$$

switches in middle stage 130 through d_2 links (for example output switch OS1 is connected from middle switch MS(1,1) through the links ML(4,1) and ML(4,2); output switch OS1 is connected from middle switch MS(1,2) through the links ML(4,7) and ML(4,8); output switch OS1 is connected from middle switch MS(1,3) through the link ML(4,13); output switch OS1 is connected from middle switch MS(1,4) through the link ML(4,19)).

[0154] Finally the connection topology of the network 200C shown in FIG. 2C is known to be back to back inverse Benes connection topology.

[0155] In other embodiments the connection topology may be different from the embodiment of the network 200C of FIG. 2C. 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(4,24) are connected between the respective stages is different. Even though only one embodiment is illustrated, in general, the network $V_{bft}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 2C, are only one example of network $V_{bft}(N_1, N_2, d, s)$.

[0156] In the embodiment of FIG. 2C, each of 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(4,24) 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) 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(1,2)-MS(2,4) are referred to as root stage switches.

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

[0157] Network 200D of FIG. 2D is an example of general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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 Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s=1$ according to the current invention (and in the example of FIG. 2D, $s=1$). The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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).

US 2010/0172349 A1

Jul. 8, 2010

17

[0158] 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.

[0159] 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.

[0160] Similarly 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 switches in middle stage **140** through d links and also are connected to exactly

$$\frac{d_2}{2}$$

output switches in output stage **120** through d_2 links.

[0161] Similarly each of the

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

[0162] middle switches

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

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

[0163] Each of the

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

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

$$\frac{d_2}{2}$$

switches in middle stage **130** through d_2 links.

[0164] As described before, again the connection topology of a general $V_{bft}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 2C is one example of network $V_{bft}(N_1, N_2, d, s)$ for $s=1$ and $N_2 > N_1$.

[0165] The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention.

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

[0166] Referring to FIG. 2E, in one embodiment, an exemplary asymmetrical Butterfly fat tree network **200E** with three stages of sixteen 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** and **140** 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. Middle stage **130** consists of four, eight by four switches MS(1,1)-MS(1,4) and middle stage **140** consists of four, two by two switches MS(2,1)-MS(2,4).

[0167] 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 six, the switches in output stage **120** are of size two by two, and there are four switches in each of middle stage **130** and middle stage **140**. Such a network can be operated in rearrangeably non-blocking 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 of size eight by four in middle stage **130**, and four switches of size two by two in middle stage **140**.

[0168] 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 $d_1 * d_1$ and each output switch OS1-OS4 can be denoted in general with the notation $(d * d)$, where

US 2010/0172349 A1

Jul. 8, 2010

18

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

The size of each switch in middle stage **130** can be denoted as $(d+d_1)*2d$. The size of each switch in the root stage (i.e., middle stage **140**) can be denoted as $d*d$. The size of each switch in all the middle stages excepting middle stage **130** and root stage can be denoted as $2d*2d$ (In network **200E** of FIG. **2E**, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage **140**, can be implemented as $d*2d$ and $d*d$ since the down coming middle links are never setup to the up going middle links. For example in network **200E** of FIG. **2E**, the down coming middle links $ML(3,2)$ and $ML(3,5)$ are never setup to the up going middle links $ML(2,1)$ and $ML(2,2)$ for the middle switch $MS(1,1)$. So middle switch $MS(1,1)$ can be implemented as a two by four switch with middle links $ML(1,1)$ and $ML(1,3)$ as inputs and middle links $ML(2,1)$, $ML(2,2)$, $ML(4,1)$ and $ML(4,2)$ as outputs; and a two by two switch with middle links $ML(3,2)$ and $ML(3,5)$ as inputs and middle links $ML(4,1)$ and $ML(4,2)$ as outputs).

[0169] 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 Butterfly fat tree network can be represented with the notation $V_{bft}(N_1, N_2, d, s)$, 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.

[0170] Each of the

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

input switches $IS1-IS4$ are connected to exactly

$$\frac{d_1}{2}$$

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 $ML(1,5)$; input switch $IS1$ is connected to middle switch $MS(1,4)$ through the link $ML(1,6)$).

[0171] Each of the

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

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

$$\frac{d_1}{2}$$

input switches through d_1 links (for example the links $ML(1,1)$ and $ML(1,2)$ are connected from input switch $IS1$ to middle switch $MS(1,1)$; the links $ML(1,7)$ and $ML(1,8)$ are connected from input switch $IS2$ to middle switch $MS(1,1)$; the link $ML(1,13)$ is connected from input switch $IS3$ to middle switch $MS(1,1)$; the link $ML(1,19)$ is connected from input switch $IS4$ to middle switch $MS(1,1)$), and also are connected from exactly d switches in middle stage **140** through d links (for example the link $ML(3,2)$ is connected to the middle switch $MS(1,1)$ from middle switch $MS(2,1)$; and the link $ML(3,5)$ is connected to the middle switch $MS(1,1)$ from middle switch $MS(2,3)$).

[0172] Similarly each of the

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

middle switches $MS(1,1)-MS(1,4)$ in the middle stage **130** are connected to exactly d switches in middle stage **140** through d links (for example the link $ML(2,1)$ is connected from middle switch $MS(1,1)$ to middle switch $MS(2,1)$ and the link $ML(2,2)$ is connected from middle switch $MS(1,1)$ to middle switch $MS(2,3)$), and also are connected to exactly d output switches in output stage **120** through d links (for example the link $ML(4,1)$ is connected to output switch $OS1$ from middle switch $MS(1,1)$ and the link $ML(4,2)$ is connected to output switch $OS2$ from middle switch $MS(1,1)$).

[0173] 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 d links (for example the link $ML(2,1)$ is connected to the middle switch $MS(2,1)$ from middle switch $MS(1,1)$ and the link $ML(2,5)$ is 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 **130** through d links (for example the link $ML(3,2)$ is connected from middle switch $MS(2,1)$ to middle switch $MS(1,1)$ and the link $ML(3,1)$ is connected from middle switch $MS(2,1)$ to middle switch $MS(1,3)$).

[0174] Each of the

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

output switches $OS1-OS4$ are connected from exactly d switches in middle stage **130** through d links (for example output switch $OS1$ is connected from middle switch $MS(1,1)$ through the link $ML(4,1)$, and output switch $OS1$ is connected from middle switch $MS(1,2)$ through the link $ML(4,3)$).

US 2010/0172349 A1

Jul. 8, 2010

19

[0175] Finally the connection topology of the network 200E shown in FIG. 2E is known to be back to back inverse Benes connection topology.

[0176] In other embodiments the connection topology may be different from the embodiment of the network 200E of FIG. 2E. That is the way the links ML(1,1)-ML(1,24), 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. Even though only one embodiment is illustrated, in general, the network $V_{bft}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 2E is only one example of network $V_{bft}(N_1, N_2, d, s)$.

[0177] In the embodiment of FIG. 2E, each of the links ML(1,1)-ML(1,24), ML(2,1)-ML(2,8), ML(3,1)-ML(3,8) and ML(4,1)-ML(4,8) 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.

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

[0178] Network 200F of FIG. 2F is an example of general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network 200F of FIG. 2F, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s=1$ according to the current invention. (And in the example of FIG. 2F, $s=1$). The general asymmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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).

[0179] Each of the

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

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

$$\frac{d_1}{2}$$

switches in middle stage 130 through d_1 links

[0180] 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

$$\frac{d_1}{2}$$

input switches through d_1 links and also are connected from exactly d switches in middle stage 140 through d links.

[0181] Similarly each of the

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

middle switches MS(1,1)-MS(1, $2N_2/d$) in the middle stage 130 also are connected to exactly d switches in middle stage 140 through d links and also are connected to exactly d output switches in output stage 120 through d links.

US 2010/0172349 A1

Jul. 8, 2010

20

[0182] 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 $130+10*(\text{Log}_d N_2-2)$ are connected from exactly d switches in middle stage $130+10*(\text{Log}_d N_2-3)$ through d links and also are connected to exactly d switches in middle stage $130+10*(\text{Log}_d N_2-1)$ through d links.

[0183] 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 d links.

[0184] As described before, again the connection topology of a general $V_{bft}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ 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 $V_{bft}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 2E is one example of network $V_{bft}(N_1, N_2, d, s)$ for $s=1$ and $N_1 > N_2$.

[0185] The general symmetrical Butterfly fat tree network $V_{bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for unicast when $s \geq 1$ according to the current invention.

Multi-Link Butterfly Fat Tree Embodiments:

Symmetric RNB Embodiments:

[0186] Referring to FIG. 3A, in one embodiment, an exemplary symmetrical Multi-link Butterfly fat tree network **300A** with three stages of sixteen 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**, and **140** 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**. Input stage **110** and output stage **120** together belong to leaf stage. And all the middle stages excepting root stage namely middle stage **130** consists of four, eight by eight switches **MS(1,1)-MS(1,4)**, and root stage i.e., middle stage **140** consists of four, four by four switches **MS(2,1)-MS(2,4)**.

[0187] 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** and middle stage **140**. 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** and middle stage **140**.

[0188] 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 $4d*4d$ excepting that the size of each switch in middle stage **140** is denoted as $2d*2d$. (In another embodiment, the size of each switch in any of the middle stages other than the middle stage **140**, can be implemented as $2d*4d$ and $2d*2d$ since the down coming middle links are never setup to the up going middle links. For example in network **300A** of FIG. 3A, the down coming middle links **ML(3,3)**, **ML(3,4)**, **ML(3,9)** and **ML(3,10)** are never setup to the up going middle links **ML(2,1)**, **ML(2,2)**, **ML(2,3)** and **ML(2,4)** for the middle switch **MS(1,1)**. So middle switch **MS(1,1)** can be implemented as a four by eight switch with middle links **ML(1,1)**, **ML(1,2)**, **ML(1,5)** and **ML(1,6)** as inputs and middle links **ML(2,1)**, **ML(2,2)**, **ML(2,3)**, **ML(2,4)**, **ML(4,1)**, **ML(4,2)**, **ML(4,3)**, and **ML(4,4)** as outputs; and a four by four switch with middle links **ML(3,3)**, **ML(3,4)**, **ML(3,9)** and **ML(3,10)** as inputs and middle links **ML(4,1)**, **ML(4,2)**, **ML(4,3)**, and **ML(4,4)** as outputs).

[0189] Middle stage **140** is called as root stage. 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 Butterfly fat tree network can be represented with the notation $V_{mlink-bft}(N, d, s)$, 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.

[0190] Each of the N/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)** and **ML(1,2)**; and input switch **IS1** is also connected to middle switch **MS(1,2)** through the links **ML(1,3)** and **ML(1,4)**).

[0191] 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(1,5)** and **ML(1,6)** are connected to the middle switch **MS(1,1)** from input switch **IS2**) and are also connected from exactly d switches in middle stage **140** through $2 \times d$ links (for example the links **ML(3,3)** and **ML(3,4)** are connected to the middle switch **MS(1,1)**

US 2010/0172349 A1

Jul. 8, 2010

21

from middle switch MS(2,1) and also the links ML(3,9) and ML(3,10) are connected to the middle switch MS(1,1) from middle switch MS(2,3)).

[0192] Each of the N/d middle switches MS(1,1)-MS(1,4) in the middle stage 130 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)) 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(1,1), and the links ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(1,1)).

[0193] 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 130 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(1,3); and the links ML(3,3) and ML(3,4) are connected from middle switch MS(2,1) to middle switch MS(1,1)).

[0194] Each of the N/d output switches OS1-OS4 are connected from exactly d switches in middle stage 130 through 2xd links (for example output switch OS1 is connected from middle switch MS(1,1) through the links ML(4,1), ML(4,2); and output switch OS1 is also connected from middle switch MS(1,2) through the links ML(4,5) and ML(4,6)).

[0195] Finally the connection topology of the network 300A shown in FIG. 3A is known to be back to back inverse Benes connection topology.

[0196] In other embodiments the connection topology may be different from the network 300A of FIG. 3A. 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 $V_{mlink-bft}(N,d,s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mlink-bft}(N,d,s)$ 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 $V_{mlink-bft}(N,d,s)$ 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 $V_{mlink-bft}(N,d,s)$ can be built. The embodiment of FIG. 3A is only one example of network $V_{mlink-bft}(N,d,s)$.

[0197] In the embodiment of FIG. 3A 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.

[0198] In the example illustrated in FIG. 3A, 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, to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0199] 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:

[0200] Network 300B of FIG. 3B is an example of general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N,d,s)$ with $(\log_d N)$ stages. The general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N,d,s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N,d,s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention. (And in the example of FIG. 3B, $s=2$). The general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N,d,s)$ with $(\log_d N)$ 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 2xd 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 2xd incoming links for each of N/d output switches OS1-OS(N/d) (for example ML(2xLog_dN-2,1)-ML(2xLog_dN-2,2xd) to the output switch OS1).

[0201] Each of the N/d input switches IS1-IS(N/d) are connected to exactly d switches in middle stage 130 through 2xd links.

[0202] 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 2xd links and also are connected to exactly d switches in middle stage 140 through 2xd links.

[0203] Similarly each of the N/d middle switches MS(1,1)-MS(1,N/d) in the middle stage 130 are also connected from exactly d switches in middle stage 140 through 2xd links and also are connected to exactly d output switches in output stage 120 through 2xd links.

US 2010/0172349 A1

Jul. 8, 2010

22

[0204] Similarly each of the N/d middle switches

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

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

[0205] Each of the N/d output switches OS1-OS(N/d) are connected from exactly d switches in middle stage **130** through $2 \times d$ links.

[0206] As described before, again the connection topology of a general $V_{mlink-bft}(N,d,s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink-bft}(N,d,s)$ 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 $V_{mlink-bft}(N,d,s)$ 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 $V_{mlink-bft}(N,d,s)$ can be built. The embodiment of FIG. 3A are one example of network $V_{mlink-bft}(N,d,s)$.

[0207] The general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N,d,s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N,d,s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0208] Every switch in the Multi-link Butterfly fat tree networks discussed herein has multicast capability. In a $V_{mlink-bft}(N,d,s)$ 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 type (with fan-out r' ,

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

although the same discussion is applicable to the second type.

[0209] To characterize a multicast assignment, for each inlet link

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

let $I_i=0$, 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. 3A shows an exemplary three-stage network, namely $V_{mlink-bft}(8,2,2)$, 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(1,2) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,2) only once into output switch OS2 in output stage **120** and middle switch MS(2,2) in middle stage **140** respectively.

[0210] The connection I_1 also fans out in middle switch MS(2,2) only once into middle switches MS(1,4) in middle stage **130**. The connection I_1 also fans out in middle switch MS(1,4) only once into output switch 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 RNB ($N_2 > N_1$) Embodiments:

[0211] Referring to FIG. 3C, in one embodiment, an exemplary asymmetrical Multi-link Butterfly fat tree network **300C** with three stages of sixteen 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** and **140** 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. Middle stage **130** consists of four, eight by twelve switches MS(1,1)-MS(1,4) and middle stage **140** consists of four, four by four switches MS(2,1)-MS(2,4).

[0212] 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 four switches in each of middle stage **130** and middle stage **140**. 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 of size eight by twelve in middle stage **130** and four switches of size four by four in middle stage **140**.

[0213] 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 2010/0172349 A1

Jul. 8, 2010

23

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

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 middle stage 130 can be denoted as $4d * 2(d+d_2)$. The size of each switch in the root stage (i.e., middle stage 140) can be denoted as $2d * 2d$. The size of each switch in all the middle stages excepting middle stage 130 and root stage can be denoted as $4d * 4d$ (In network 300C of FIG. 3C, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $2d * 4d$ and $2d * 2d$ since the down coming middle links are never setup to the up going middle links. For example in network 300C of FIG. 3C, the down coming middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) are never setup to the up going middle links ML(2,1), ML(2,2), ML(2,3) and ML(2,4) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a four by eight switch with middle links ML(1,1), ML(1,2), ML(1,5) and ML(1,6) as inputs and middle links ML(2,1), ML(2,2), ML(2,3), ML(2,4), ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs; and a four by four switch with middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) as inputs and middle links ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs).

[0214] 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 Butterfly fat tree network can be represented with the notation $V_{mlink-bft}(N_1, N_2, d, s)$, 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 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) and ML(1,2), and input switch IS1 is also connected to 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(1,5) and ML(1,6) are connected to the middle switch MS(1,1) from input switch IS2) and are also connected from exactly d switches in middle stage 140 through $2 \times d$ links (for example the links ML(3,3) and ML(3,4) are connected to the middle switch MS(1,1) from middle switch MS(2,1), and the links ML(3,9) and ML(3,10) are connected to the middle switch MS(1,1) from middle switch MS(2,3)).

[0217] Similarly each of the

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

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

$$\frac{d_2}{2}$$

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

[0218] 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 130 through $2 \times d$ links (for example the links ML(3,3) and ML(3,4) are connected from middle switch

US 2010/0172349 A1

Jul. 8, 2010

24

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

[0219] Each of the

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

output switches OS1-OS4 are connected from exactly

$$\frac{d_2}{2}$$

switches in middle stage 130 through d_2 links (for example output switch OS1 is connected from middle switch MS(1,1) through the links ML(4,1) and ML(4,2); output switch OS1 is connected from middle switch MS(1,2) through the links ML(4,9) and ML(4,10); output switch OS1 is connected from middle switch MS(1,3) through the links ML(4,17) and ML(4,18); output switch OS1 is connected from middle switch MS(1,4) through the links ML(4,25) and ML(4,26)).

[0220] Finally the connection topology of the network 300C shown in FIG. 3C is known to be back to back inverse Benes connection topology.

[0221] In other embodiments the connection topology may be different from the embodiment of the network 300C of FIG. 3C. 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(4,32) are connected between the respective stages is different. Even though only one embodiment is illustrated, in general, the network $V_{mlink-bft}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mlink-bft}(N_1, N_2, d, s)$ 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 $V_{mlink-bft}(N_1, N_2, d, s)$ 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 $V_{mlink-bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 3C, are only one example of network $V_{mlink-bft}(N_1, N_2, d, s)$.

[0222] In the embodiment of FIG. 3C, 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(4,32) 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(1,2)-MS(2,4) are referred to as root stage switches.

[0223] In the example illustrated in FIG. 3C, 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 300C, to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0224] 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:

[0225] Network 300D of FIG. 3D is an example of general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_2 > N_1$ and $N_2 = p * N_1$, where $p > 1$. In network 300D of FIG. 3D, $N_1 = N$ and $N_2 = p * N$. The general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention. (And in the example of FIG. 3D, $s=2$). The general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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

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

US 2010/0172349 A1

25

Jul. 8, 2010

output switches OS1-OS(N_1/d) (for example the links OL1-OL($p*d$) to the output switch OS1) and $d+d_2$ ($=d+p*d$) 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($2 \times \text{Log}_d N_1 - 2, d+d_2$) to the output switch OS1).

[0226] 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 $2 \times d$ links.

[0227] 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.

[0228] Similarly 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 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.

[0229] 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+ $10^*(\text{Log}_d N_1 - 2)$ are connected from exactly d switches in middle stage 130+ $10^*(\text{Log}_d N_1 - 3)$ through $2 \times d$ links and also are connected to exactly d switches in middle stage 130+ $10^*(\text{Log}_d N_1 - 1)$ through $2 \times d$ links.

[0230] 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 through $d+d_2$ links.

[0231] As described before, again the connection topology of a general $V_{mlink-bft}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{mlink-bft}(N_1, N_2, d, s)$ 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 $V_{mlink-bft}(N_1, N_2, d, s)$ 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 $V_{mlink-bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 3C is one example of network $V_{mlink-bft}(N_1, N_2, d, s)$ for $s=2$ and $N_2 > N_1$.

[0232] The general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0233] For example, the network of FIG. 3C shows an exemplary three-stage network, namely $V_{mlink-bft}(8, 24, 2, 2)$, 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(1,2) in middle stage 130, and fans out in middle switches MS(1,1) and MS(1,2) only once into output switch OS2 in output stage 120 and middle switch MS(2,2) in middle stage 140.

[0234] The connection I_1 also fans out in middle switch MS(2,2) only once into middle switches MS(1,4) in middle stage 130. The connection I_1 also fans out in middle switch MS(1,4) only once into output switch 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 OL18. 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:

[0235] Referring to FIG. 3E, in one embodiment, an exemplary asymmetrical Multi-link Butterfly fat tree network 300E with three stages of sixteen 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 and 140 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. Middle stage 130 consists of four, twelve by eight switches

US 2010/0172349 A1

Jul. 8, 2010

26

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

[0236] 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 and middle stage 140. 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 of size twelve by eight in middle stage 130, and four switches of size four by four in middle stage 140.

[0237] 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 $(2d * d)$, where

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

The size of each switch in middle stage 130 can be denoted as $2(d + d_1) * 4d$. The size of each switch in the root stage (i.e., middle stage 140) can be denoted as $2d * 2d$. The size of each switch in all the middle stages excepting middle stage 130 and root stage can be denoted as $4d * 4d$ (In network 300C of FIG. 3C, there is no such middle stage). (In another embodiment, the size of each switch in any of the middle stages other than the middle stage 140, can be implemented as $2d * 4d$ and $2d * 2d$ since the down coming middle links are never setup to the up going middle links. For example in network 300E of FIG. 3E, the down coming middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) are never setup to the up going middle links ML(2,1), ML(2,2), ML(2,3) and ML(2,4) for the middle switch MS(1,1). So middle switch MS(1,1) can be implemented as a four by eight switch with middle links ML(1,1), ML(1,2), ML(1,5) and ML(1,6) as inputs and middle links ML(2,1), ML(2,2), ML(2,3), ML(2,4), ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs; and a four by four switch with middle links ML(3,3), ML(3,4), ML(3,9) and ML(3,10) as inputs and middle links ML(4,1), ML(4,2), ML(4,3), and ML(4,4) as outputs).

[0238] 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 Butterfly fat tree network can be represented with the notation $V_{mlink-bft}(N_1, N_2, d, s)$, 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.

[0239] 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) 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 links ML(1,5) and ML(1,6); input switch IS1 is connected to middle switch MS(1,4) through the links ML(1,7) and ML(1,8)).

[0240] 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 from input switch IS1 to middle switch MS(1,1); the links ML(1,9) and ML(1,10) are connected from input switch IS2 to middle switch MS(1,1); the links ML(1,17) and ML(1,18) are connected from input switch IS3 to middle switch MS(1,1); the links ML(1,25) and ML(1,26) are connected from input switch IS4 to middle switch MS(1,1)), and also are connected to exactly d switches in middle stage 140 through $2d$ links (for example the links ML(3,3) and ML(3,4) are connected to the middle switch MS(1,1) from middle switch MS(2,1); and the links ML(3,9) and ML(3,10) are connected to the middle switch MS(1,1) from middle switch MS(2,3)).

[0241] Similarly each of the

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

middle switches MS(1,1)-MS(1,4) in the middle stage 130 are connected to exactly d switches in middle stage 140

US 2010/0172349 A1

Jul. 8, 2010

27

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

[0242] 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 $2d$ 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 **130** through $2d$ links (for example the links $ML(3,3)$ and $ML(3,4)$ are connected from middle switch $MS(2,1)$ to middle switch $MS(1,1)$ and the links $ML(3,1)$ and $ML(3,2)$ are connected from middle switch $MS(2,1)$ to middle switch $MS(1,3)$).

[0243] Each of the

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

output switches $OS1$ - $OS4$ are connected from exactly d switches in middle stage **130** through $2 \times d$ links (for example output switch $OS1$ is connected from middle switch $MS(1,1)$ through the links $ML(4,1)$ and $ML(4,2)$, and output switch $OS1$ is connected from middle switch $MS(1,2)$ through the links $ML(4,5)$ and $ML(4,6)$).

[0244] Finally the connection topology of the network **300E** shown in FIG. **3E** is known to be back to back inverse Benes connection topology.

[0245] In other embodiments the connection topology may be different from the embodiment of the network **300E** of FIG. **3E**. 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 one embodiment is illustrated, in general, the network $V_{mlink-bft}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{mlink-bft}(N_1, N_2, d, s)$ 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 $V_{mlink-bft}(N_1, N_2, d, s)$ 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 $V_{mlink-bft}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. **3E** is only one example of network $V_{mlink-bft}(N_1, N_2, d, s)$.

[0246] In the embodiment of FIG. **3E**, 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 $MS(1,1)$ - $MS(1,4)$ and $MS(2,1)$ - $MS(2,4)$ are referred to as middle switches or middle ports.

[0247] In the example illustrated in FIG. **3E**, 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 **300E**, to be operated in rearrangeably nonblocking manner in accordance with the invention.

[0248] 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 RNB ($N_1 > N_2$) Embodiments:

[0249] Network **300F** of FIG. **3F** is an example of general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ with $(\log_d N)$ stages where $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$. In network **300F** of FIG. **3F**, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s=2$ according to the current invention. Also the general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s=2$ according to the current invention. (And in the example of FIG. **3F**, $s=2$). The general asymmetrical Multi-link Butterfly fat tree network $V_{mlink-bft}(N_1, N_2, d, s)$ with $(\log_d N)$ 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 $d+d_1$ ($=d+p \times d$) outgoing links for each of

US 2010/0172349 A1

Jul. 8, 2010

28

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

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 $2*d$ incoming links for each of

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

output switches OS1-OS(N_2/d) (for example ML($2*\text{Log}_d N_2-2,1$)-ML($2*\text{Log}_d N_2-2,2*d$) to the output switch OS1).

[0250] 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.

[0251] 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

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

input switches through $d+d_1$ links and also are connected from exactly d switches in middle stage **140** through $2*d$ links.

[0252] Similarly each of the

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

middle switches MS(1,1)-MS(1, $2N_2/d$) in the middle stage **130** also are connected to exactly d switches in middle stage **140** through $2*d$ links and also are connected to exactly d output switches in output stage **120** through $2*d$ links.

[0253] Similarly each of the

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

middle switches

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

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

[0254] 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 $2*d$ links.

[0255] As described before, again the connection topology of a general $V_{\text{mlink-bft}}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{\text{mlink-bft}}(N_1, N_2, d, s)$ 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 $V_{\text{mlink-bft}}(N_1, N_2, d, s)$ 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 $V_{\text{mlink-bft}}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 3E is one example of network $V_{\text{mlink-bft}}(N_1, N_2, d, s)$ for $s=2$ and $N_1 > N_2$.

[0256] The general symmetrical Multi-link Butterfly fat tree network $V_{\text{mlink-bft}}(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s \geq 2$ according to the current invention. Also the general symmetrical Multi-link Butterfly fat tree network $V_{\text{mlink-bft}}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

[0257] For example, the network of FIG. 3E shows an exemplary three-stage network, namely $V_{\text{mlink-bft}}(24, 8, 2, 2)$, 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(1,2) in middle stage **130**, and fans out in middle switches MS(1,1) and MS(1,2) only once into output switch OS2 in output stage **120** and middle switch and MS(2, 2) in middle stage **140** respectively.

[0258] The connection I_1 also fans out in middle switch MS(2,2) only once into middle switch MS(1,4) in middle stage **130**. The connection I_1 also fans out in middle switch MS(1,4) only once into output switch 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

US 2010/0172349 A1

Jul. 8, 2010

29

with the invention, each connection can fan out in the input stage switch into at most two middle stage switches in middle stage **130**.

Strictly Nonblocking Multi-link Butterfly Fat Tree Networks:

[0259] The general symmetric multi-link Butterfly fat tree network $V_{mlink-bft}(N,d,s)$ can also be operated in strictly non-blocking manner for multicast when $s \geq 3$ according to the current invention. Similarly the general asymmetric multi-link Butterfly fat tree network $V_{mlink-bft}(N_1,N_2,d,s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention.

Scheduling Method Embodiments:

[0260] FIG. 4A shows a high-level flowchart of a scheduling method **1000**, in one embodiment executed to setup multicast and unicast connections in network **400A** of FIG. 4A (or any of the networks $V_{bft}(N_1,N_2,d,s)$ and $V_{mlink-bft}(N_1,N_2,d,s)$ disclosed in this invention). According to this embodiment, a multicast connection request is received in act **1010**. Then the control goes to act **1020**.

[0261] 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**.

[0262] 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**.

[0263] 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**.

[0264] 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**.

[0265] 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 connec-

US 2010/0172349 A1

Jul. 8, 2010

30

tion 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**.

[0266] 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.

[0267] 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. 4A 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 $V_{bft}(N_1, N_2, d, s)$ and the network $V_{mink-bft}(N_1, N_2, d, s)$ to be operated in nonblocking manner in accordance with the invention.

[0268] According to the current invention, using the method **1040** of FIG. 4A, the network $V_{bft}(N_1, N_2, d, s)$ and the network $V_{mink-bft}(N_1, N_2, d, s)$ are operated in rearrangeably nonblocking for unicast connections when $s \geq 1$, are operated in strictly nonblocking for unicast connections when $s \geq 2$, are operated in rearrangeably nonblocking for multicast connections when $s \geq 2$, and are operated in strictly nonblocking for multicast connections when $s \geq 3$.

[0269] The connection request of the type described above in reference to method **1000** of FIG. 4A 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. 4A 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.

[0270] As noted above method **1000** of FIG. 4A can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks $V_{bft}(N, d, s)$, $V_{bft}(N_1, N_2, d, s)$, $V_{mink-bft}(N, d, s)$, and $V_{mink-bft}(N_1, N_2, d, s)$ disclosed in this invention.

Applications Embodiments

[0271] 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 be the implementation of middle switch MS(2,1) of the diagram **100A** of FIG. 1A where inlet link IL1 of diagram **500A1** corresponds to middle link ML(2,1) of diagram **100A**, inlet link IL2 of diagram **500A1** corresponds to middle link ML(2,5) of diagram **100A**, outlet link OL1 of diagram **500A1** corresponds to middle link ML(3,1) of diagram **100A**, outlet link OL2 of diagram **500A1** corresponds to middle link ML(3,2) of diagram **100A**.

1) Programmable Integrated Circuit Embodiments:

[0272] 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).

[0273] 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.

2) One-time Programmable Integrated Circuit Embodiments:

[0274] 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.

US 2010/0172349 A1

Jul. 8, 2010

31

[0275] 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 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:

[0276] 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.

[0277] 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.

[0278] 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:

[0279] 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.

[0280] 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

a leaf stage comprising an input stage and an output stage; and said 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 its immediate succeeding stage where $x > 0$; and said 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 its immediate succeeding stage; and

a plurality of y middle stages, excepting a root stage, comprising

$$x \times \frac{N}{d}$$

middle switches in each of said y middle stages wherein one of said middle stages is the immediate succeeding stage to

US 2010/0172349 A1

Jul. 8, 2010

32

both said input stage and said output stage, where $y > 1$, and said root stage comprising N/d middle switches; and

each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, comprising d incoming links (hereinafter "incoming middle links") connecting from switches in its immediate preceding stage and d incoming links connecting from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links (hereinafter "outgoing middle links") connecting to switches in its immediate succeeding stage and d outgoing links connecting to switches in its immediate preceding stage; and

each middle switch in said succeeding stage to both said input stage and said output stage comprising d incoming links connecting from switches in said input stage and d incoming links connecting from switches in its immediate preceding stage, and each middle switch further comprising

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

outgoing links connecting to switches in said output stage and d outgoing links connecting to switches in its immediate succeeding stage; and

each middle switch in said root stage comprising d incoming links connecting from switches in its immediate preceding stage and each middle switch further comprising d outgoing links connecting to switches in its immediate preceding 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

a leaf stage comprising an input stage and an output stage; said 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 its immediate succeeding stage where $x > 0$; and said 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 its immediate succeeding stage; and

a plurality of y middle stages, excepting a root stage, comprising

$$x \times \frac{N}{d}$$

middle switches in each of said y middle stages wherein one of said middle stages is the immediate succeeding stage to both said input stage and said output stage, where $y > 1$, and said root stage comprising N/d middle switches; and

each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, comprising d incoming links (hereinafter "incoming middle links") connecting from switches in its immediate preceding stage and d incoming links connecting from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links (hereinafter "outgoing middle links") connecting to switches in its immediate succeeding stage and d outgoing links connecting to switches in its immediate preceding stage; and

each middle switch in said succeeding stage to both said input stage and said output stage comprising

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

incoming links connecting from switches in said input stage and d incoming links connecting from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links connecting to switches in said output stage and d outgoing links connecting to switches in its immediate succeeding stage; and

each middle switch in said root stage comprising d incoming links connecting from switches in its immediate preceding stage and each middle switch further comprising d outgoing links connecting to switches in its immediate preceding 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 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 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 $y \cong (\log_d N_1) - 1$ when $N_2 > N_1$, and $y \cong (\log_d N_2) - 1$ 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

US 2010/0172349 A1

Jul. 8, 2010

33

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 $x \geq 2$, 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 $x \geq 2$,

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 $x \geq 3$,

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 input switches, said output switches and said middle switches are not fully 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 a leaf stage comprising an input stage and an output stage; and said 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 its immediate succeeding stage where $x > 0$; and said 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 its immediate succeeding stage; and

a plurality of y middle stages, excepting a root stage, having

$$x \times \frac{N}{d}$$

middle switches in each of said y middle stages wherein one of said middle stages is the immediate succeeding stage to both said input stage and said output stage, where $y > 1$, and said root stage having N/d middle switches, and

each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, having d incoming links connected from switches in its immediate preceding stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links connected to switches in its immediate succeeding stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said succeeding stage to both said input stage and said output stage having d incoming links connected from switches in said input stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further having

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

outgoing links connected to switches in said output stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said root stage having d incoming links connected from switches in its immediate preceding stage and each middle switch further having d outgoing links connected to switches in its immediate preceding stage; or

US 2010/0172349 A1

Jul. 8, 2010

34

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 having

having a leaf stage having an input stage and an output stage; and said 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 its immediate succeeding stage where $x > 0$; and said 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 its immediate succeeding stage; and

a plurality of y middle stages, excepting a root stage, having

$$x \times \frac{N}{d}$$

middle switches in each of said y middle stages wherein one of said middle stages is the immediate succeeding stage to both said input stage and said output stage, where $y > 1$, and said root stage having N/d middle switches, and

each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, having d incoming links connected from switches in its immediate preceding stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links connected to switches in its immediate succeeding stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said succeeding stage to both said input stage and said output stage having

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

incoming links connected from switches in said input stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further having d outgoing links connected to switches in said output stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said root stage having d incoming links connected from switches in its immediate preceding stage and each middle switch further having d outgoing links connected to switches in its immediate preceding 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 $N_2 = p * N_1$ 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 a leaf stage comprising an input stage and an output stage; and said 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 con-

US 2010/0172349 A1

Jul. 8, 2010

35

nected to switches in its immediate succeeding stage where $x > 0$; and said 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 its immediate succeeding stage; and

a plurality of y middle stages, excepting a root stage, having

$$x \times \frac{N}{d}$$

middle switches in each of said y middle stages wherein one of said middle stages is the immediate succeeding stage to both said input stage and said output stage, where $y > 1$, and said root stage having N/d middle switches, and

each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, having d incoming links connected from switches in its immediate preceding stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links connected to switches in its immediate succeeding stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said succeeding stage to both said input stage and said output stage having d incoming links connected from switches in said input stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further having

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

outgoing links connected to switches in said output stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said root stage having d incoming links connected from switches in its immediate preceding stage and each middle switch further having d outgoing links connected to switches in its immediate preceding 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

having a leaf stage having an input stage and an output stage; and said 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 its immediate succeeding stage where $x > 0$; and said 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 its immediate succeeding stage; and

a plurality of y middle stages, excepting a root stage, having

$$x \times \frac{N}{d}$$

middle switches in each of said y middle stages wherein one of said middle stages is the immediate succeeding stage to both said input stage and said output stage, where $y > 1$, and said root stage having N/d middle switches, and

each middle switch in all said middle stages, excepting said root stage and said succeeding stage to both said input stage said output stage, having d incoming links connected from switches in its immediate preceding stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further comprising d outgoing links connected to switches in its immediate succeeding stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said succeeding stage to both said input stage and said output stage having

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

incoming links connected from switches in said input stage and d incoming links connected from switches in its immediate succeeding stage, and each middle switch further having d outgoing links connected to switches in said output stage and d outgoing links connected to switches in its immediate succeeding stage; and

each middle switch in said root stage having d incoming links connected from switches in its immediate preceding stage and each middle switch further having d out-

US 2010/0172349 A1

Jul. 8, 2010

36

going links connected to switches in its immediate preceding 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 E

(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 2008/147927 A1

(51) International Patent Classification:
H04L 12/50 (2006.01)

(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, 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.

(21) International Application Number:
PCT/US2008/064604

(22) International Filing Date: 22 May 2008 (22.05.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/940,389 25 May 2007 (25.05.2007) US
60/940,391 25 May 2007 (25.05.2007) US
60/940,392 25 May 2007 (25.05.2007) US

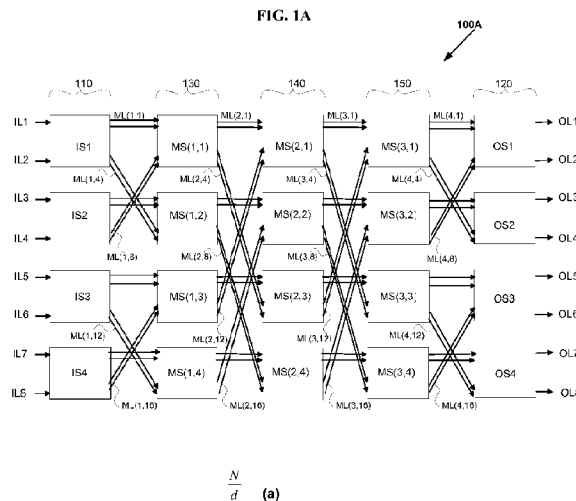
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, 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).

(71) Applicant and

(72) Inventor: **KONDA, Venkat** [US/US]; 6278 Grand Oak Way, San Jose, CA 95135 (US).

Published:
— with international search report

(54) Title: FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE NETWORKS



(57) Abstract: A generalized multi-link multi-stage network comprising $(2 \times \log_2 N) - 1$ 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 $(2 \times \log_2 N) - 3$ 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 multistage 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.

WO 2008/147927 A1

WO 2008/147927

PCT/US2008/064604

**FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE
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, 389 entitled "FULLY CONNECTED GENERALIZED REARRANGEABLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda
10 assigned to the same assignee as the current application, filed May 25, 2007.

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, 392 entitled "FULLY CONNECTED GENERALIZED STRICTLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same
15 assignee as the current application, filed concurrently.

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, 391 entitled "FULLY CONNECTED GENERALIZED FOLDED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application,
20 filed concurrently.

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
25 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

WO 2008/147927

PCT/US2008/064604

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 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-0038PCT entitled "FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE 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, 387 entitled "FULLY CONNECTED GENERALIZED
10 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 the same assignee as the current application, filed May 25, 2007.

15 This application is related to and incorporates by reference in its entirety the PCT Application Docket No. S-0045PCT entitled "VLSI LAYOUTS OF FULLY CONNECTED GENERALIZED NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed concurrently, and the U.S. Provisional Patent Application Serial No. 60/940, 394 entitled "VLSI LAYOUTS OF FULLY
20 CONNECTED GENERALIZED 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. Provisional Patent Application Serial No. 60/984, 724 entitled "VLSI LAYOUTS OF FULLY CONNECTED NETWORKS WITH LOCALITY EXPLOITATION" by Venkat
25 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 FULLY CONNECTED GENERALIZED AND PYRAMID NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed January 1, 2008.

WO 2008/147927

PCT/US2008/064604

BACKGROUND OF INVENTION

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

5 "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.

10 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 network's ability to rearrange prior

15 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

20 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.

25 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. 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

30 nonblocking for unicast connections.

WO 2008/147927

PCT/US2008/064604

U.S. Patent 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.

U.S. Patent 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. Patent 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 $m \geq 3 \times n - 1$.

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, Jan. 1953, No.1, pp. 406-424 showed a way of constructing large multi-stage networks by recursive substitution with a crosspoint complexity of $d^2 \times N \times (\log_d N)^{2.58}$ for strictly nonblocking unicast network. Similarly U.S. Patent 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

WO 2008/147927

PCT/US2008/064604

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 $N=256$.

The crosspoint complexity of all these networks is prohibitively large to
5 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.

10 SUMMARY OF INVENTION

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 $\frac{N}{d}$ switches with each of them having d inlet links and $2 \times d$ outgoing links connecting to second stage switches, an output stage having $\frac{N}{d}$ switches with each of them having d
15 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 $\frac{N}{d}$ 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-
20 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.

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 $\frac{N}{d}$
25 switches with each of them having d inlet links and $3 \times d$ outgoing links connecting to

WO 2008/147927

PCT/US2008/064604

second stage switches, an output stage having $\frac{N}{d}$ 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 $\frac{N}{d}$ 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

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.

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.

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.

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

WO 2008/147927

PCT/US2008/064604

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

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.

FIG. 1F is a diagram 100F 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 $s=2$, strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

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.

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.

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.

WO 2008/147927

PCT/US2008/064604

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.

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.

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.

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.

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.

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

WO 2008/147927

PCT/US2008/064604

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

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.

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.

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.

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.

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 $s = 2$,

WO 2008/147927

PCT/US2008/064604

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

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
 5 $N_1 = 8$, $N_2 = p * N_1 = 24$ 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.

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 $s = 2$,
 10 strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

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

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

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

WO 2008/147927

PCT/US2008/064604

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
5 multicast connections, in accordance with the invention.

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

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

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

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

WO 2008/147927

PCT/US2008/064604

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.

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.

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.

FIG. 2A is a diagram 200A of an exemplary symmetrical folded multi-link multi-stage network $V_{fold-mlink}(N, d, s)$ 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.

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.

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

WO 2008/147927

PCT/US2008/064604

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

FIG. 2D is a diagram 200D of a general asymmetrical folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, 2)$ 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.

FIG. 2E is a diagram 200E of an exemplary asymmetrical folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, 2)$ 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.

FIG. 2F is a diagram 200F of a general asymmetrical folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, 2)$ 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.

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.

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.

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 $N =$

WO 2008/147927

PCT/US2008/064604

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

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 $N =$
5 8, $d = 2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

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

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
15 accordance with the invention.

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.

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.
20

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
25

WO 2008/147927

PCT/US2008/064604

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.

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 $N =$
 5 8 , $d = 2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

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.

10 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.

15 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.

20 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.

25 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.

WO 2008/147927

PCT/US2008/064604

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
5 multicast connections, in accordance with the invention.

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.

10 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.

15 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.

20 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.

25 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.

WO 2008/147927

PCT/US2008/064604

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.

5 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.

FIG. 3B2 is a diagram 300B2 of an exemplary asymmetrical multi-link multi-
10 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 $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 3C2 is a diagram 300C2 of an exemplary asymmetrical multi-link multi-
15 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 $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 3D2 is a diagram 300D2 of an exemplary asymmetrical multi-link multi-
20 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 $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

FIG. 3E2 is a diagram 300E2 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 $N_2 = 8$, $N_1 = p * N_2$
25 = 24, where $p = 3$, $d = 2$ and $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

WO 2008/147927

PCT/US2008/064604

FIG. 3F2 is a diagram 300F2 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 $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

5 FIG. 3G2 is a diagram 300G2 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 $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

10 FIG. 3H2 is a diagram 300H2 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 $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

15 FIG. 3I2 is a diagram 300I2 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 $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

20 FIG. 3J2 is a diagram 300J2 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 $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

25 FIG. 3K2 is a diagram 300K2 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 $s=3$, strictly nonblocking network for arbitrary fan-out multicast connections, in accordance with the invention.

WO 2008/147927

PCT/US2008/064604

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

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.

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.

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.

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.

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

WO 2008/147927

PCT/US2008/064604

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

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.

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 $(2 * \log_d N) - 1$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

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.

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

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

WO 2008/147927

PCT/US2008/064604

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

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 $(2 \times \log_d N) - 1$ stages strictly nonblocking network for unicast connections and rearrangeably nonblocking network for arbitrary fan-out multicast connections in accordance with the invention.

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.

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

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.

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 $(2 \times \log_d N) - 1$ stages rearrangeably nonblocking network for unicast connections in accordance with the invention.

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.

WO 2008/147927

PCT/US2008/064604

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 $(2 \times \log_d N) - 1$ stages rearrangeably nonblocking network for unicast connections in accordance with the invention.

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.

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.

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.

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.

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.

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.

WO 2008/147927

PCT/US2008/064604

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.

5 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.

FIG. 6I is a diagram 600I of an exemplary symmetrical multi-stage network
10 $V(N, d, s)$ (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.

FIG. 6J is a diagram 600J of an exemplary symmetrical multi-stage network
15 $V(N, d, s)$ having an exemplary connection topology of five stages with $N = 8$, $d = 2$ and $s = 1$, rearrangeably nonblocking network for unicast connections.

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

20 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.

FIG. 6B1 is a diagram 600B1 of an exemplary asymmetrical multi-stage network
25 $V(N_1, N_2, d, s)$ (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.

WO 2008/147927

PCT/US2008/064604

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.

5 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.

FIG. 6E1 is a diagram 600E1 of an exemplary asymmetrical multi-stage network
10 $V(N_1, N_2, d, s)$ (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 $s = 1$, rearrangeably nonblocking network for unicast connections.

FIG. 6F1 is a diagram 600F1 of an exemplary asymmetrical multi-stage network
15 $V(N_1, N_2, d, s)$ 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.

FIG. 6G1 is a diagram 600G1 of an exemplary asymmetrical multi-stage network
20 $V(N_1, N_2, d, s)$ 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.

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 nonblocking network for unicast connections, in accordance with the invention.

25 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

WO 2008/147927

PCT/US2008/064604

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.

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 nonblocking network for unicast connections.

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.

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 nonblocking network for unicast connections, in accordance with the invention.

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.

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 nonblocking network for unicast connections, in accordance with the invention.

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 nonblocking network for unicast connections, in accordance with the invention.

WO 2008/147927

PCT/US2008/064604

FIG. 6E2 is a diagram 600E2 of an exemplary asymmetrical multi-stage network $V(N_1, N_2, d, s)$ (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 $s = 1$, rearrangeably nonblocking network for unicast connections.

5 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.

FIG. 6G2 is a diagram 600G2 of an exemplary asymmetrical multi-stage network
10 $V(N_1, N_2, d, s)$ 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 nonblocking network for unicast connections, in accordance with the invention.

FIG. 6H2 is a diagram 600H2 of an exemplary asymmetrical multi-stage network
15 $V(N_1, N_2, d, s)$ 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 nonblocking network for unicast connections, in accordance with the invention.

FIG. 6I2 is a diagram 600I2 of an exemplary asymmetrical multi-stage network
20 $V(N_1, N_2, d, s)$ (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.

FIG. 6J2 is a diagram 600J2 of an exemplary asymmetrical multi-stage network
25 $V(N_1, N_2, d, s)$ 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 nonblocking network for unicast connections.

WO 2008/147927

PCT/US2008/064604

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.

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.

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

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.

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

WO 2008/147927

PCT/US2008/064604

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.

5 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
10 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.

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
15 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.

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

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
25 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 connection topologies and the scheduling methods are described in detail in U.S.

WO 2008/147927

PCT/US2008/064604

Provisional Patent Application Serial No. 60/940, 387 that is incorporated by reference above.

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 U.S. Provisional Patent Application Serial No. 60/940, 390 that is incorporated by reference above.

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 $V_{cube}(N_1, N_2, d, s)$ for $s = 1, 2, 3$ or any number in general, are described in detail in U.S. Provisional Patent Application Serial No. 60/940, 394 that is incorporated by reference above.

5) 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.

6) 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.

RNB MULTI-LINK MULTI-STAGE EMBODIMENTS:

Symmetric RNB Embodiments:

Referring to FIG. 1A, in one embodiment, an exemplary symmetrical multi-link multi-stage network 100A with five stages of twenty switches for satisfying

WO 2008/147927

PCT/US2008/064604

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).

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.

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 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 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

WO 2008/147927

PCT/US2008/064604

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.

5 Each of the $\frac{N}{d}$ 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)).

 Each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,4) in the middle stage 130 are
 10 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
 15 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)).

 Similarly each of the $\frac{N}{d}$ 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
 20 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
 25 MS(3,3)).

WO 2008/147927

PCT/US2008/064604

Similarly each of the $\frac{N}{d}$ 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(3,11) 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)).

10 Each of the $\frac{N}{d}$ 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(3,2) through the links ML(4,7) and ML(4,8)).

15 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 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).

WO 2008/147927

PCT/US2008/064604

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
 5 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.

In one embodiment of this network each of the input switches IS1-IS4 and output
 10 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 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
 15 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
 20 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.

25 Each of the $\frac{N}{d}$ 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)).

WO 2008/147927

PCT/US2008/064604

Each of the $\frac{N}{d}$ 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(1,9) and ML(1,10) are connected to the middle switch MS(1,1) from
 5 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)).

Similarly each of the $\frac{N}{d}$ middle switches MS(2,1) – MS(2,4) in the middle stage
 10 140 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
 15 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)).

Similarly each of the $\frac{N}{d}$ 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
 20 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
 25 ML(4,3) and ML(4,4) are connected to output switch OS2 from middle switch MS(3,1)).

WO 2008/147927

PCT/US2008/064604

Each of the $\frac{N}{d}$ 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
 5 ML(4,9) and ML(4,10)).

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

Referring to FIG. 1C, in another embodiment of network $V_{mlink}(N, d, s)$, an exemplary symmetrical multi-link multi-stage network 100C with five stages of twenty
 10 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
 15 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).

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
 20 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
 25 stage 140 and middle stage 150.

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

WO 2008/147927

PCT/US2008/064604

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-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. The symmetric multi-link multi-stage network of FIG. 1C 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.

Each of the $\frac{N}{d}$ 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)).

Each of the $\frac{N}{d}$ 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)).

WO 2008/147927

PCT/US2008/064604

Similarly each of the $\frac{N}{d}$ 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)).

10 Similarly each of the $\frac{N}{d}$ 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)).

Each of the $\frac{N}{d}$ 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)).

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

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

WO 2008/147927

PCT/US2008/064604

exemplary symmetrical multi-link multi-stage networks 100D, 100E, 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).

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.

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 $V_{mlink}(N, d, s)$, 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.

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.

WO 2008/147927

PCT/US2008/064604

11, and FIG. 1J, each of the $\frac{N}{d}$ input switches IS1 – IS4 are connected to exactly $2 \times d$ switches in middle stage 130 through $2 \times d$ links.

Each of the $\frac{N}{d}$ 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.

Similarly each of the $\frac{N}{d}$ 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.

Similarly each of the $\frac{N}{d}$ 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.

Each of the $\frac{N}{d}$ output switches OS1 – OS4 are connected from exactly $2 \times d$ switches in middle stage 150 through $2 \times d$ links.

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 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

WO 2008/147927

PCT/US2008/064604

ten embodiments of FIG. 1A to FIG. 1J are only three examples of network
 $V_{mlink}(N, d, s)$.

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
5 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) –
10 MS(2,4), and MS(3,1) – MS(3,4) are referred to as middle switches or middle ports.

In the example illustrated in FIG. 1A (or in FIG1B 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
15 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
20 invention.

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,
25 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
30 the output stage switches to satisfy the connection request.

WO 2008/147927

PCT/US2008/064604

Generalized Symmetric RNB Embodiments:

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-stage network $V_{mlink}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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).

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

Each of the $\frac{N}{d}$ 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.

Similarly each of the $\frac{N}{d}$ 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

WO 2008/147927

PCT/US2008/064604

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

Similarly each of the $\frac{N}{d}$ middle switches $MS(2 \times \text{Log}_d N - 3, 1)$ -

$MS(2 \times \text{Log}_d N - 3, \frac{N}{d})$ in the middle stage $130 + 10 * (2 * \text{Log}_d N - 4)$ are connected from

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

Each of the $\frac{N}{d}$ output switches OS1 - OS(N/d) are connected from exactly d

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

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 15 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 $V_{mlink}(N, d, s)$.

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 20 $V_{mlink}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

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

WO 2008/147927

PCT/US2008/064604

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.

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. 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.

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

WO 2008/147927

PCT/US2008/064604

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:

Referring to FIG. 1A1, in one embodiment, an exemplary asymmetrical multi-link
 5 multi-stage network 100A1 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-
 10 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).

Such a network can be operated in strictly non-blocking manner for unicast
 15 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
 20 are of size eight by six, and there are four switches in each of middle stage 130, middle
 stage 140 and middle stage 150.

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
 25 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

WO 2008/147927

PCT/US2008/064604

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

5 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

10 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.

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

15 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)).

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,

20 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)).

WO 2008/147927

PCT/US2008/064604

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)).

10 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

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

20 switch OS4 from middle switch MS(3,1)).

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

25 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)).

WO 2008/147927

PCT/US2008/064604

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

Referring to FIG. 1B1, in one embodiment, an exemplary asymmetrical multi-link multi-stage network 100B1 with five stages of twenty switches for satisfying
 5 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
 10 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).

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
 15 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
 20 stage 140 and middle stage 150.

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
 25 $N_2 = p * N_1$ 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

WO 2008/147927

PCT/US2008/064604

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.

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)).

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 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)).

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

WO 2008/147927

PCT/US2008/064604

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)).

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)).

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)).

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

WO 2008/147927

PCT/US2008/064604

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).

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.

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

WO 2008/147927

PCT/US2008/064604

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.

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)).

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)).

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

WO 2008/147927

PCT/US2008/064604

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)).

5 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

10 $\frac{d + d_2}{2}$ 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

15 switch OS4 from middle switch MS(3,1)).

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

20 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)).

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

25 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.

WO 2008/147927

PCT/US2008/064604

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), 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).

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.

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 $V_{link}(N_1, N_2, d, s)$, 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.

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.

WO 2008/147927

PCT/US2008/064604

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.

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
 5 exactly d switches in middle stage 140 through $2 \times d$ links.

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.

Similarly each of the $\frac{N_1}{d}$ middle switches MS(3,1) – MS(3,4) in the middle stage
 10 150 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.

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.

15 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
 20 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

WO 2008/147927

PCT/US2008/064604

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. 1A1 to FIG. 1J1 are only three examples of network $V_{mink}(N_1, N_2, d, s)$.

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.

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

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).

WO 2008/147927

PCT/US2008/064604

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:

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 $N_2 = p * N_1$ 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($p*d$) 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).

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.

WO 2008/147927

PCT/US2008/064604

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.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(\text{Log}_d N_1 - 1, 1) -$

- 5 $MS(\text{Log}_d N_1 - 1, \frac{N_1}{d})$ 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 $2 \times d$ links and also are connected to exactly d switches in middle stage $130 + 10 * (\text{Log}_d N_1 - 1)$ through $2 \times d$ links.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(2 * \text{Log}_d N_1 - 3, 1) -$

- 10 $MS(2 * \text{Log}_d N_1 - 3, \frac{N_1}{d})$ 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 $2 \times d$ links and also are connected to exactly $\frac{d + d_2}{2}$ output switches in output stage 120 through $d + d_2$ links.

Each of the $\frac{N_1}{d}$ output switches $OS1 - OS(N_1/d)$ are connected from exactly

- 15 $\frac{d + d_2}{2}$ switches in middle stage $130 + 10 * (2 * \text{Log}_d N_1 - 4)$ through $d + d_2$ links.

- 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
- 20 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,

WO 2008/147927

PCT/US2008/064604

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 $N_2 > N_1$.

5 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.

10 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
15 MS(2,3) respectively in middle stage 140.

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
20 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:

25 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

WO 2008/147927

PCT/US2008/064604

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
 5 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).

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
 10 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
 15 stage 140 and middle stage 150.

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
 20 $N_1 = p * N_2$ 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 $d_1 = N_1 * \frac{d}{N_2} = p * d$. The size of each switch in any of the middle stages excepting the first middle stage can be denoted as $2d * 2d$. The
 25 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-

WO 2008/147927

PCT/US2008/064604

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-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.

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)).

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)).

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

WO 2008/147927

PCT/US2008/064604

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)).

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)).

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)).

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

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

WO 2008/147927

PCT/US2008/064604

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).

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.

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

WO 2008/147927

PCT/US2008/064604

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.

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)).

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(2,2)).

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

WO 2008/147927

PCT/US2008/064604

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)).

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(3,9) 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)).

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)).

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

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-

WO 2008/147927

PCT/US2008/064604

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).

5 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
10 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.

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
15 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
20 with the notation $(2 * d * d)$, where $d_1 = N_1 * \frac{d}{N_2} = p * d$. 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-
25 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-IL24), N_2 represents the total number of outlet links of all output switches (for example

WO 2008/147927

PCT/US2008/064604

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.

Each of the $\frac{N_2}{d}$ input switches IS1 – IS4 are connected to exactly $\frac{d + d_1}{2}$

5 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
10 ML(1,7) and ML(1,8)).

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)
15 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
20 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)).

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
25 switch MS(2,1) from middle switch MS(1,4)) and also are connected to exactly d

WO 2008/147927

PCT/US2008/064604

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)).

5 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

10 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)).

 Each of the $\frac{N_2}{d}$ output switches OS1 – OS4 are connected from exactly d

15 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)).

 Finally the connection topology of the network 100C2 shown in FIG. 1C2 is

20 hereinafter called nearest neighbor connection topology.

 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

25 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

WO 2008/147927

PCT/US2008/064604

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
5 MS(3,1) - MS(3,4).

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
10 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.

The networks 100D2, 100E2, 100F2, 100G2, 100H2, 100I2 and 100J2 of
15 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
20 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.

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
25 connected to exactly $\frac{d + d_2}{2}$ switches in middle stage 130 through $d + d_2$ links.

WO 2008/147927

PCT/US2008/064604

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.

Similarly each of the $\frac{N_2}{d}$ middle switches MS(2,1) – MS(2,4) in the middle stage 5 140 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.

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.

10 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.

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 15 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 20 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 $V_{mlink}(N_1, N_2, d, s)$.

WO 2008/147927

PCT/US2008/064604

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
5 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.

In the example illustrated in FIG. 1A2 (or in FIG. 1B2 to FIG. 1J2), a fan-out of
10 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
15 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.

The connection request of the type described above can be unicast connection
20 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
25 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.

WO 2008/147927

PCT/US2008/064604

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

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 $N_1 = p * N_2$ where $p > 1$. In network 100K2 of FIG. 1K2, $N_2 = N$ and $N_1 = p * N$. The

5 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

10 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($p*d$) 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,($d+p*d$)) to the input switch IS1). There are d outlet

15 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 $ML(2 \times \text{Log}_d N_2 - 2, 1) - ML(2 \times \text{Log}_d N_2 - 2, 2 \times d)$ to the output switch OS1).

Each of the $\frac{N_2}{d}$ input switches IS1 - IS(N_2/d) are connected to exactly $\frac{d + d_1}{2}$

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

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.

WO 2008/147927

PCT/US2008/064604

Similarly each of the $\frac{N_2}{d}$ middle switches $MS(\text{Log}_d N_2 - 1, 1)$ -
 $MS(\text{Log}_d N_2 - 1, \frac{N_2}{d})$ 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 $2 \times d$ links and also
 are connected to exactly d switches in middle stage $130 + 10 * (\text{Log}_d N_2 - 1)$ through
 5 $2 \times d$ links.

Similarly each of the $\frac{N_2}{d}$ middle switches $MS(2 * \text{Log}_d N_2 - 3, 1)$ -
 $MS(2 * \text{Log}_d N_2 - 3, \frac{N_2}{d})$ 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 $2 \times d$
 10 links.

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.

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
 15 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.
 20 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 $N_2 > N_1$.

WO 2008/147927

PCT/US2008/064604

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

5 $s \geq 2$ according to the current invention.

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 $I_j = \phi$ 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(1,4) in middle stage 130, and fans out

10 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.

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

15 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.

20 **Symmetric Folded RNB Embodiments:**

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-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

25 stage $130 + 10 * (\text{Log}_d N_2 - 2)$. This is true for all the embodiments presented in the current invention.

WO 2008/147927

PCT/US2008/064604

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).

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.

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 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 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

WO 2008/147927

PCT/US2008/064604

notation $V_{fold-link}(N, d, s)$, 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
 5 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.

Each of the $\frac{N}{d}$ 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
 10 MS(1,2) through the links ML(1,3) and ML(1,4)).

Each of the $\frac{N}{d}$ 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(1,7) and ML(1,8) are connected to the middle switch MS(1,1) from
 15 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)).

Similarly each of the $\frac{N}{d}$ middle switches MS(2,1) – MS(2,4) in the middle stage
 20 140 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
 25 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)).

WO 2008/147927

PCT/US2008/064604

Similarly each of the $\frac{N}{d}$ 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(3,11) 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)).

10 Each of the $\frac{N}{d}$ 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(3,2) through the links ML(4,7) and ML(4,8)).

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

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 $V_{fold-link}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold-link}(N, d, s)$ 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 $V_{fold-link}(N, d, s)$ 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

WO 2008/147927

PCT/US2008/064604

$V_{fold-link}(N, d, s)$ can be built. The embodiment of FIG. 2A is only one example of network $V_{fold-link}(N, d, s)$.

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 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.

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.

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).

WO 2008/147927

PCT/US2008/064604

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:

Network 200B of FIG. 2B is an example of general symmetrical folded multi-link multi-stage network $V_{fold-link}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages. The general symmetrical folded multi-link multi-stage network $V_{fold-link}(N, d, s)$ 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 $V_{fold-link}(N, d, s)$ 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 $V_{fold-link}(N, d, s)$ with $(2 \times \log_d N) - 1$ stages has d inlet links for each of $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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).

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

Each of the $\frac{N}{d}$ 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.

WO 2008/147927

PCT/US2008/064604

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

5 Similarly each of the $\frac{N}{d}$ middle switches $MS(2 \times \text{Log}_d N - 3, 1) - MS(2 \times \text{Log}_d N - 3, \frac{N}{d})$ 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 $2 \times d$ links and also are connected to exactly d output switches in output stage 120 through $2 \times d$ links.

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

As described before, again the connection topology of a general $V_{\text{fold-link}}(N, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{\text{fold-link}}(N, d, s)$ 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 $V_{\text{fold-link}}(N, d, s)$ 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 $V_{\text{fold-link}}(N, d, s)$ can be built. The embodiment of FIG. 1A is one example of network $V_{\text{fold-link}}(N, d, s)$.

20 The general symmetrical folded multi-link multi-stage network $V_{\text{fold-link}}(N, d, s)$ 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-

WO 2008/147927

PCT/US2008/064604

stage network $V_{fold-link}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

Every switch in the folded multi-link multi-stage networks discussed herein has multicast capability. In a $V_{fold-link}(N, d, s)$ network, if a network inlet link is to be
 5 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
 10 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.
 15 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.

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

20 is to be connected in the multicast assignment. For example, the network of Fig. 1C shows an exemplary five-stage network, namely $V_{link}(8, 2, 2)$, 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
 25 once into middle switches MS(2,1) and MS(2,3) respectively in middle stage 140.

WO 2008/147927

PCT/US2008/064604

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:

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).

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.

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

WO 2008/147927

PCT/US2008/064604

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 $V_{fold-link}(N_1, N_2, d, s)$, 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.

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)).

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

WO 2008/147927

PCT/US2008/064604

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)).

Similarly each of the $\frac{N_1}{d}$ middle switches MS(2,1) – MS(2,4) in the middle stage

- 5 140 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
10 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)).

Similarly each of the $\frac{N_1}{d}$ middle switches MS(3,1) – MS(3,4) in the middle stage

- 15 150 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 $d + d_2$ links (for example the links ML(4,1) and ML(4,2) are connected to output switch OS1 from Middle switch MS(3,1);
20 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)).

Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $\frac{d + d_2}{2}$

- 25 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

WO 2008/147927

PCT/US2008/064604

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)).

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

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
10 stages is different. Even though only one embodiment is illustrated, in general, the network $V_{fold-link}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold-link}(N, d, s)$ 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 $V_{fold-link}(N, d, s)$
15 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 $V_{fold-link}(N, d, s)$ can be built. The embodiment of FIG. 2C is only one example of network $V_{fold-link}(N, d, s)$.

In the embodiment of FIG. 2C each of the links ML(1,1) – ML(1,16), ML(2,1) –
20 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
25 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.

WO 2008/147927

PCT/US2008/064604

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.

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.

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

Network 200D of FIG. 2D is an example of general asymmetrical folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ 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 $V_{fold-link}(N_1, N_2, d, s)$ 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 $V_{fold-link}(N_1, N_2, d, s)$ 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

WO 2008/147927

PCT/US2008/064604

- network $V_{fold-link}(N_1, N_2, d, s)$ 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
- 5 $d_2 = N_2 \times \frac{d}{N_1} = p \times d$) 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).

- Each of the $\frac{N_1}{d}$ input switches IS1 - IS(N_1/d) are connected to exactly $2 \times d$
- 10 switches in middle stage 130 through $2 \times d$ links.

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.

- Similarly each of the $\frac{N_1}{d}$ middle switches MS($\log_d N_1 - 1, 1$) -
- 15 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.

WO 2008/147927

PCT/US2008/064604

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(2 \times \text{Log}_d N_1 - 3, 1)$ -
 $MS(2 \times \text{Log}_d N_1 - 3, \frac{N_1}{d})$ 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 $2 \times d$ links
 and also are connected to exactly $\frac{d + d_2}{2}$ output switches in output stage 120 through
 5 $d + d_2$ links.

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.

As described before, again the connection topology of a general
 $V_{\text{fold-link}}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the
 10 connection topology of the network $V_{\text{fold-link}}(N_1, N_2, d, s)$ 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 $V_{\text{fold-link}}(N_1, N_2, d, s)$ network is,
 when no connections are setup from any input link if any output link should be reachable.
 15 Based on this property numerous embodiments of the network $V_{\text{fold-link}}(N_1, N_2, d, s)$
 can be built. The embodiment of FIG. 1C is one example of network
 $V_{\text{fold-link}}(N_1, N_2, d, s)$ for $s = 2$ and $N_2 > N_1$.

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

WO 2008/147927

PCT/US2008/064604

For example, the network of Fig. 2C shows an exemplary five-stage network, namely $V_{fold-link}(8,24,2,2)$, 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.

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.

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

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).

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

WO 2008/147927

PCT/US2008/064604

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.

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 $d_1 = N_1 * \frac{d}{N_2} = p * d$. 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 $V_{fold-link}(N_1, N_2, d, s)$, 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.

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

WO 2008/147927

PCT/US2008/064604

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
 5 ML(1,7) and ML(1,8)).

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)
 10 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
 15 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)).

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
 25 MS(3,3)).

WO 2008/147927

PCT/US2008/064604

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)).

10 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)).

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

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 $V_{fold-link}(N, d, s)$ can comprise any arbitrary type of connection topology. For example the connection topology of the network $V_{fold-link}(N, d, s)$ 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 $V_{fold-link}(N, d, s)$ 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

WO 2008/147927

PCT/US2008/064604

$V_{fold-link}(N, d, s)$ can be built. The embodiment of FIG. 2E is only one example of network $V_{fold-link}(N, d, s)$.

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

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.

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

WO 2008/147927

PCT/US2008/064604

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:

Network 200F of FIG. 2F is an example of general asymmetrical folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ 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 200F of FIG. 2F, $N_2 = N$ and $N_1 = p * N$. The general asymmetrical folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ 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 $V_{fold-link}(N_1, N_2, d, s)$ 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

$V_{fold-link}(N_1, N_2, d, s)$ 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($p*d$) 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,($d+p*d$)) 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).

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.

WO 2008/147927

PCT/US2008/064604

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.

- Similarly each of the $\frac{N_2}{d}$ middle switches $MS(\text{Log}_d N_2 - 1, 1) - MS(\text{Log}_d N_2 - 1, \frac{N_2}{d})$ 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 $2 \times d$ 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.

- Similarly each of the $\frac{N_2}{d}$ middle switches $MS(2 * \text{Log}_d N_2 - 3, 1) - MS(2 * \text{Log}_d N_2 - 3, \frac{N_2}{d})$ 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.

- 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.

- As described before, again the connection topology of a general $V_{\text{fold-link}}(N_1, N_2, d, s)$ may be any one of the connection topologies. For example the connection topology of the network $V_{\text{fold-link}}(N_1, N_2, d, s)$ 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 $V_{\text{fold-link}}(N_1, N_2, d, s)$ network is, when no connections are setup from any input link if any output link should be reachable.

WO 2008/147927

PCT/US2008/064604

Based on this property numerous embodiments of the network $V_{fold-link}(N_1, N_2, d, s)$ can be built. The embodiment of FIG. 2F is one example of network $V_{fold-link}(N_1, N_2, d, s)$ for $s = 2$ and $N_2 > N_1$.

The general symmetrical folded multi-link multi-stage network

5 $V_{fold-link}(N_1, N_2, d, s)$ 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 $V_{fold-link}(N_1, N_2, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current invention.

For example, the network of Fig. 2E shows an exemplary five-stage network,

10 namely $V_{fold-link}(8, 24, 2, 2)$, 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.

15 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

20 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:

25 Referring to FIG. 3A, in one embodiment, an exemplary symmetrical multi-link multi-stage network 300A with five stages of twenty switches for satisfying

WO 2008/147927

PCT/US2008/064604

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.

5 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).

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

10 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.

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 $d * 3d$ 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

20 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 $V_{mink}(N, d, s)$, 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

25 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.

WO 2008/147927

PCT/US2008/064604

- Each of the $\frac{N}{d}$ 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)).
- 5 Each of the $\frac{N}{d}$ 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 MS(1,1) from input switch IS2) and also are connected to exactly d switches in
- 10 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)).

- Similarly each of the $\frac{N}{d}$ middle switches MS(2,1) – MS(2,4) in the middle stage
- 15 140 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
- 20 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)).

- Similarly each of the $\frac{N}{d}$ 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
- 25 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)

WO 2008/147927

PCT/US2008/064604

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
5 connected to output switch OS2 from middle switch MS(3,1)).

Each of the $\frac{N}{d}$ 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
10 the links ML(4,10), ML(4,11) and ML(4,12)).

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

Referring to FIG. 3B, in another embodiment of network $V_{mlink}(N, d, s)$, an exemplary symmetrical multi-link multi-stage network 300B with five stages of twenty
15 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
20 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).

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
25 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.

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

WO 2008/147927

PCT/US2008/064604

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-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. 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.

Each of the $\frac{N}{d}$ 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)).

Each of the $\frac{N}{d}$ 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)).

WO 2008/147927

PCT/US2008/064604

Similarly each of the $\frac{N}{d}$ 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)).

10 Similarly each of the $\frac{N}{d}$ 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 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)).

Each of the $\frac{N}{d}$ 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(3,3) through the links ML(4,13), ML(4,14), and ML(4,15)).

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

WO 2008/147927

PCT/US2008/064604

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).

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.

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 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 $d * 3d$ 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

WO 2008/147927

PCT/US2008/064604

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

Each of the $\frac{N}{d}$ 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)).

Each of the $\frac{N}{d}$ 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(1,22), 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)).

Similarly each of the $\frac{N}{d}$ 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(2,22), 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)).

Similarly each of the $\frac{N}{d}$ 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

WO 2008/147927

PCT/US2008/064604

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
 5 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)).

Each of the $\frac{N}{d}$ 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
 10 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)).

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

15 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
 20 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).
 25

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

WO 2008/147927

PCT/US2008/064604

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.

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.

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 $\frac{N}{d}$ input switches IS1 – IS4 are connected to exactly $3 \times d$ switches in middle stage 130 through $3 \times d$ links.

Each of the $\frac{N}{d}$ 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 and also are connected to exactly d switches in middle stage 140 through $3 \times d$ links.

Similarly each of the $\frac{N}{d}$ 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 and also are connected to exactly d switches in middle stage 150 through $3 \times d$ links.

Similarly each of the $\frac{N}{d}$ 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 and also are connected to exactly d output switches in output stage 120 through $3 \times d$ links.

WO 2008/147927

PCT/US2008/064604

Each of the $\frac{N}{d}$ output switches OS1 – OS4 are connected from exactly $3 \times d$ switches in middle stage 150 through $3 \times d$ links.

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 $V_{mlink}(N, d, s)$.

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.

In the example illustrated in FIG. 3A (or in FIG1B 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

WO 2008/147927

PCT/US2008/064604

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.

The connection request of the type described above can be unicast connection
 5 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
 10 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:

15 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
 20 $(2 \times \log_d N) - 1$ stages has d inlet links for each of $\frac{N}{d}$ input switches IS1-IS(N/d) (for example the links IL1-IL(d) to the input switch IS1) and $3 \times d$ outgoing links for each of $\frac{N}{d}$ 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 $\frac{N}{d}$ output switches OS1-OS(N/d) (for example the links OL1-OL(d) to the output switch OS1) and $3 \times d$ incoming links for
 25 each of $\frac{N}{d}$ output switches OS1-OS(N/d) (for example $ML(2 \times \log_d N - 2, 1) - ML(2 \times \log_d N - 2, 3 \times d)$ to the output switch OS1).

WO 2008/147927

PCT/US2008/064604

Each of the $\frac{N}{d}$ input switches IS1 – IS(N/d) are connected to exactly d switches in middle stage 130 through $3 \times d$ links.

Each of the $\frac{N}{d}$ middle switches MS(1,1) – MS(1,N/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.

Similarly each of the $\frac{N}{d}$ 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 $3 \times d$ links and also are connected to exactly d switches in middle stage $130 + 10 * (\log_d N - 1)$ through $3 \times d$ links.

Similarly each of the $\frac{N}{d}$ middle switches $MS(2 \times \log_d N - 3, 1) - MS(2 \times \log_d N - 3, \frac{N}{d})$ in the middle stage $130 + 10 * (2 * \log_d N - 4)$ are connected from exactly d switches in middle stage $130 + 10 * (2 * \log_d N - 5)$ through $3 \times d$ links and also are connected to exactly d output switches in output stage 120 through $3 \times d$ links.

Each of the $\frac{N}{d}$ output switches OS1 – OS(N/d) are connected from exactly d switches in middle stage $130 + 10 * (2 * \log_d N - 4)$ through $3 \times d$ links.

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

WO 2008/147927

PCT/US2008/064604

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 $V_{mlink}(N, d, s)$.

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
5 invention.

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
10 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
15 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.

20 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
25 the connection I_1 fans out in the first stage switch IS1 into middle switches MS(1,1) and

WO 2008/147927

PCT/US2008/064604

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.

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:

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).

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.

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

WO 2008/147927

PCT/US2008/064604

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 $3d * (2d + d_2)$. 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.

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)).

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

WO 2008/147927

PCT/US2008/064604

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)).

5 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
10 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)).

 Similarly each of the $\frac{N_1}{d}$ middle switches MS(3,1) – MS(3,4) in the middle stage
15 150 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
20 (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
25 middle switch MS(3,1)).

WO 2008/147927

PCT/US2008/064604

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)).

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

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).

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.

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

WO 2008/147927

PCT/US2008/064604

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.

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)).

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

WO 2008/147927

PCT/US2008/064604

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)).

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)).

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 (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)).

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

WO 2008/147927

PCT/US2008/064604

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),
 5 ML(4,44), and ML(4,45)).

Finally the connection topology of the network 300B1 shown in FIG. 3B1 is known to be back to back Omega connection topology.

Referring to FIG. 3C1, in one embodiment, an exemplary asymmetrical multi-link multi-stage network 300C1 with five stages of twenty switches for satisfying
 10 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
 15 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).

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
 20 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.

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
 25 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

WO 2008/147927

PCT/US2008/064604

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

5 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

10 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.

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)).

15

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)).

20

25

WO 2008/147927

PCT/US2008/064604

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)).

10 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 (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)).

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

WO 2008/147927

PCT/US2008/064604

OS1 is also connected from middle switch MS(3,4) through the links ML(4,43), ML(4,44), and ML(4,45)).

Finally the connection topology of the network 300C1 shown in FIG. 3C1 is hereinafter called nearest neighbor connection topology.

5 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
 10 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
 15 switches MS(2,1) - MS(2,4), and middle stage 150 consists of four, six by six switches MS(3,1) - MS(3,4).

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
 20 middle stage 130, middle stage 140 and middle stage 150.

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 $V_{mlink}(N_1, N_2, d, s)$, where N_1 represents the total number of inlet links of all
 25 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.

WO 2008/147927

PCT/US2008/064604

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.

- 5 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.

- 10 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.

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 $2d + d_2$ links.

- 15 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.

- 20 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 $V_{mlink}(N_1, N_2, d, s)$ 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

WO 2008/147927

PCT/US2008/064604

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. 3A1 to FIG. 3J1 are only
5 three examples of network $V_{mlink}(N_1, N_2, d, s)$.

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
10 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.

In the example illustrated in FIG. 3A1 (or in FIG. 3B1 to FIG. 3J1), a fan-out of
15 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
20 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.

The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on
25 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

WO 2008/147927

PCT/US2008/064604

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:

5 Network 300K1 of FIG. 3K1 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 $N_2 = p * N_1$ where $p > 1$. In network 300K1 of FIG. 3K1, $N_1 = N$ and $N_2 = p * N$. The general asymmetrical multi-link multi-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
 10 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
 15 $d_2 = N_2 \times \frac{d}{N_1} = p \times d$) 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).

Each of the $\frac{N_1}{d}$ input switches IS1 - IS(N_1/d) are connected to exactly $3 \times d$
 20 switches in middle stage 130 through $3 \times d$ links.

WO 2008/147927

PCT/US2008/064604

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.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(\log_d N_1 - 1, 1) -$

- 5 $MS(\log_d N_1 - 1, \frac{N_1}{d})$ 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.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(2 * \log_d N_1 - 3, 1) -$

- 10 $MS(2 * \log_d N_1 - 3, \frac{N_1}{d})$ 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 $3 \times d$ links and also are connected to exactly $\frac{2d + d_2}{3}$ output switches in output stage 120 through $2d + d_2$ links.

Each of the $\frac{N_1}{d}$ output switches $OS1 - OS(N_1/d)$ are connected from exactly

- 15 $\frac{2d + d_2}{3}$ switches in middle stage $130 + 10 * (2 * \log_d N_1 - 4)$ through $2d + d_2$ links.

- 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
- 20 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,

WO 2008/147927

PCT/US2008/064604

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. 3A1 to FIG. 3J1 are ten examples of network $V_{mlink}(N_1, N_2, d, s)$ for $s = 3$ and $N_2 > N_1$.

5 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.

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
 10 $I_j = \phi$ 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.

The connection I_1 also fans out in middle switches MS(2,1) and MS(2,3) only
 15 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
 20 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:

Referring to FIG. 3A2, in one embodiment, an exemplary asymmetrical multi-link multi-stage network 300A2 with five stages of twenty switches for satisfying
 25 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

WO 2008/147927

PCT/US2008/064604

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).

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.

10 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

15 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

20 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

25 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.

WO 2008/147927

PCT/US2008/064604

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)).

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(2,3)).

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

WO 2008/147927

PCT/US2008/064604

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)).

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)).

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)).

Finally the connection topology of the network 300A2 shown in FIG. 3A2 is known to be back to back inverse Benes connection topology.

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

WO 2008/147927

PCT/US2008/064604

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

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
5 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.

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
10 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 $d_1 * (2d + d_1)$ 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
15 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
20 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.

WO 2008/147927

PCT/US2008/064604

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)).

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)).

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

WO 2008/147927

PCT/US2008/064604

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)).

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)).

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)).

Finally the connection topology of the network 300B2 shown in FIG. 3B2 is known to be back to back Omega connection topology.

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

WO 2008/147927

PCT/US2008/064604

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

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
5 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.

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
10 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 $d_1 * (2d + d_1)$ 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
15 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
20 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.

WO 2008/147927

PCT/US2008/064604

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)).

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)).

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

WO 2008/147927

PCT/US2008/064604

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)).

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 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)).

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(3,4) through the links ML(4,22), ML(4,23), and ML(4,24)).

Finally the connection topology of the network 300C2 shown in FIG. 3C2 is hereinafter called nearest neighbor connection topology.

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,

WO 2008/147927

PCT/US2008/064604

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).

Such a network can be operated in strictly non-blocking manner for multicast
 5 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.

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
 10 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
 15 from each input switch to the inlet links of each input switch.

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 $2d + d_2$ links.

20 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 $2d + d_2$ links and also are connected to exactly d switches in middle stage 140 through $3 \times d$ links.

WO 2008/147927

PCT/US2008/064604

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 and also are connected to exactly d switches in middle stage 150 through $3 \times d$ links.

5 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 and also are connected to exactly d output switches in output stage 120 through $3 \times d$ links.

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.

In all the ten embodiments of FIG. 3A2 to FIG. 3J2 the connection topology is
 10 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
 15 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. 3A2 to FIG. 3J2 are only
 20 three examples of network $V_{mlink}(N_1, N_2, d, s)$.

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
 25 ports. The input stage 110 is often referred to as the first stage. The output switches

WO 2008/147927

PCT/US2008/064604

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.

In the example illustrated in FIG. 3A2 (or in FIG. 3B2 to FIG. 3J2), a fan-out of
 5 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
 10 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.

The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on
 15 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
 20 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:

Network 3001K2 of FIG. 3K2 is an example of general asymmetrical multi-link
 25 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 $N_1 = p * N_2$ where $p > 1$. In network 300K2 of FIG. 3K2, $N_2 = N$ and $N_1 = p * N$. 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

WO 2008/147927

PCT/US2008/064604

- 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 $2d + d_1 (= 2d + p \times d)$ 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).
- 10 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.
- 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.
- 15 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 $130 + 10 * (\log_d N_2 - 2)$ are connected from exactly d switches in middle stage $130 + 10 * (\log_d N_2 - 3)$ through $3 \times d$ links and also are connected to exactly d switches in middle stage $130 + 10 * (\log_d N_2 - 1)$ through $3 \times d$ links.

WO 2008/147927

PCT/US2008/064604

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 $130 + 10 * (2 * \log_d N_2 - 4)$ are connected
 from exactly d switches in middle stage $130 + 10 * (2 * \log_d N_2 - 5)$ through $3 \times d$ links
 and also are connected to exactly d output switches in output stage 120 through $3 \times d$
 5 links.

Each of the $\frac{N_2}{d}$ output switches OS1 – OS(N_2/d) are connected from exactly d
 switches in middle stage $130 + 10 * (2 * \log_d N_2 - 4)$ through $2 \times d$ links.

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
 10 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.
 15 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 $N_2 > N_1$.

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
 20 current invention.

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
 $I_j = \phi$ 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(1,4) in middle stage 130, and fans out

WO 2008/147927

PCT/US2008/064604

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.

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:

The folded multi-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$, disclosed in the current invention, is topologically exactly the same as the multi-stage network $V_{link}(N_1, N_2, d, s)$, disclosed in U.S. Provisional Patent Application Docket No. M-0037US that is incorporated by reference above, excepting that in the illustrations folded network $V_{fold-link}(N_1, N_2, d, s)$ is shown as it is folded at middle stage $130 + 10 * (\log_d N_2 - 2)$.

The general symmetrical folded multi-link multi-stage network $V_{fold-link}(N, d, s)$ 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-link multi-stage network $V_{fold-link}(N_1, N_2, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention.

25

WO 2008/147927

PCT/US2008/064604

FOLDED MULTI-STAGE NETWORK EMBODIMENTS:

Symmetric folded RNB Embodiments:

Referring to FIG. 4A, in one embodiment, an exemplary symmetrical folded multi-stage network 400A with five stages of thirty two switches for satisfying

5 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

10 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).

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

15 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

20 stage 140 and middle stage 150.

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 number of inlet links or outlet links. The number of middle switches in each middle stage

25 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

WO 2008/147927

PCT/US2008/064604

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.

Each of the $\frac{N}{d}$ 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).

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).

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).

WO 2008/147927

PCT/US2008/064604

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
 5 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)).

Each of the $\frac{N}{d}$ 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
 10 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).

Finally the connection topology of the network 400A shown in FIG. 4A is known to be back to back inverse Benes connection topology.

Referring to FIG. 4A1, in another embodiment of network $V_{fold}(N, d, s)$, an
 15 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
 20 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).

Such a network can be operated in strictly non-blocking manner for unicast
 25 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

WO 2008/147927

PCT/US2008/064604

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.

- 5 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 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
- 10 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. 4A1 is also the network of the type
- 15 $V_{fold}(N, d, s)$, 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
- 20 symmetrical network they are the same.

Each of the $\frac{N}{d}$ 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).

- 25 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

WO 2008/147927

PCT/US2008/064604

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).

5 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).

 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)).

 Each of the $\frac{N}{d}$ 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(3,3), MS(3,5) and MS(3,7) through the links ML(4,1), ML(4,5), ML(4,9) and ML(4,13) respectively).

 Finally the connection topology of the network 400A1 shown in FIG. 4A1 is known to be back to back Omega connection topology.

25 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

WO 2008/147927

PCT/US2008/064604

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).

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.

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 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

WO 2008/147927

PCT/US2008/064604

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.

5 Each of the $\frac{N}{d}$ 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).

10 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(1,14) 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).

15 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).

20

25 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

WO 2008/147927

PCT/US2008/064604

ML(4,1) and ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switches MS(3,1)).

Each of the $\frac{N}{d}$ 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
5 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).

Finally the connection topology of the network 400A2 shown in FIG. 4A2 is hereinafter called nearest neighbor connection topology.

In the three embodiments of FIG. 4A, FIG. 4A1 and FIG. 4A2 the connection
10 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
15 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
20 examples of network $V_{fold}(N, d, s)$.

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
25 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),

WO 2008/147927

PCT/US2008/064604

MS(2,1) – MS(2,8), and MS(3,1) – MS(3,8) are referred to as middle switches or middle ports.

In the example illustrated in FIG. 4A (or in FIG1A1, 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.

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:

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 current invention. Also the general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can be operated in strictly

WO 2008/147927

PCT/US2008/064604

nonblocking 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 $\frac{N}{d}$ input switches

IS1-IS(N/d) (for example the links IL1-IL(d) to the input switch IS1) and $2 \times d$ outgoing
 5 links for each of $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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).

10 Each of the $\frac{N}{d}$ 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.

15 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.

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

20 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.

WO 2008/147927

PCT/US2008/064604

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 $130 + 10 * (2 * \log_d N - 4)$ 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.

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

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
10 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
15 FIG. 4A2 are three examples of network $V_{fold}(N, d, s)$.

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 current invention. Also the general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the
20 current invention.

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

WO 2008/147927

PCT/US2008/064604

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.

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. 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 $I_j = \phi$ 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.

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.

WO 2008/147927

PCT/US2008/064604

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

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
 5 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
 10 MS(2,1) - MS(2,8), and middle stage 150 consists of eight, two by four switches MS(3,1) - MS(3,8).

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
 15 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
 20 150.

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
 25 $N_2 = p * N_1$ where $p > 1$. The number of middle switches in each middle stage is denoted by $2 * \frac{N_1}{d}$. 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

WO 2008/147927

PCT/US2008/064604

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 $V_{fold}(N_1, N_2, d, s)$, 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.

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).

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).

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

WO 2008/147927

PCT/US2008/064604

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).

Similarly each of the $2 \times \frac{N_1}{d}$ middle switches MS(3,1) – MS(3,8) in the middle
 5 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,
 10 OS2, OS3, and OS4 respectively from middle switches MS(3,1)).

Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $d+d_2$ switches in middle stage 150 through $d+d_2$ 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),
 15 ML(4,17), ML(4,21), ML(4,25) and ML(4,29) respectively).

Finally the connection topology of the network 400C shown in FIG. 4C is known to be back to back inverse Benes connection topology.

Referring to FIG. 4C1, in another embodiment of network $V_{fold}(N_1, N_2, d, s)$, an exemplary asymmetrical folded multi-stage network 400C1 with five stages of thirty two
 20 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
 25 eight, two by two switches MS(1,1) - MS(1,8), middle stage 140 consists of eight, two by

WO 2008/147927

PCT/US2008/064604

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).

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.

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 $2 * \frac{N_1}{d}$. 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 $d_2 = N_2 * \frac{d}{N_1} = p * d$. 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 $V_{fold}(N_1, N_2, d, s)$, 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

WO 2008/147927

PCT/US2008/064604

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.

Each of the $\frac{N_1}{d}$ input switches IS1 – IS4 are connected to exactly $2 \times d$ switches
 5 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).

Each of the $2 \times \frac{N_1}{d}$ middle switches MS(1,1) – MS(1,8) in the middle stage 130
 10 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).

Similarly each of the $2 \times \frac{N_1}{d}$ middle switches MS(2,1) – MS(2,8) in the middle
 15 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
 20 MS(3,2) respectively).

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
 25 exactly $\frac{d + d_2}{2}$ output switches in output stage 120 through $\frac{d + d_2}{2}$ links (for example

WO 2008/147927

PCT/US2008/064604

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)).

Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $d + d_2$

switches in middle stage 150 through $d + d_2$ links (for example output switch OS1 is
 5 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).

Finally the connection topology of the network 400C1 shown in FIG. 4C1 is known to be back to back Omega connection topology.

10 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
 15 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).

20 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
 25 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.

WO 2008/147927

PCT/US2008/064604

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

5 $N_2 = p * N_1$ where $p > 1$. The number of middle switches in each middle stage is denoted by $2 * \frac{N_1}{d}$. 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 $d_2 = N_2 * \frac{d}{N_1} = p * d$. 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

10 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.

Each of the $\frac{N_1}{d}$ input switches IS1 – IS4 are connected to exactly $2 * d$ switches

20 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).

Each of the $2 * \frac{N_1}{d}$ 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

WO 2008/147927

PCT/US2008/064604

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).

5 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).

10

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)).

15

Each of the $\frac{N_1}{d}$ output switches OS1 – OS4 are connected from exactly $d+d_2$ switches in middle stage 150 through $d+d_2$ 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).

20

Finally the connection topology of the network 400C2 shown in FIG. 4C2 is hereinafter called nearest neighbor connection topology.

25

WO 2008/147927

PCT/US2008/064604

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. 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 $V_{fold}(N_1, N_2, d, s)$.

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.

In the example illustrated in FIG. 4C (or in FIG1C1, 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,

WO 2008/147927

PCT/US2008/064604

or 400C2), to be operated in rearrangeably nonblocking manner in accordance with the invention.

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:

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

WO 2008/147927

PCT/US2008/064604

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).

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.

Each of the $2 \times \frac{N_1}{d}$ middle switches MS(1,1) – MS(1, $2 \times \frac{N_1}{d}$) 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.

Similarly each of the $2 \times \frac{N_1}{d}$ middle switches $MS(\log_d N_1 - 1, 1) - MS(\log_d N_1 - 1, 2 \times \frac{N_1}{d})$ 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 d links and also are connected to exactly d switches in middle stage $130 + 10 * (\log_d N_1 - 1)$ through d links.

Similarly each of the $2 \times \frac{N_1}{d}$ middle switches $MS(2 \times \log_d N_1 - 3, 1) - MS(2 \times \log_d N_1 - 3, 2 \times \frac{N_1}{d})$ 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 d links and also are connected to exactly d output switches in output stage 120 through d links.

WO 2008/147927

PCT/US2008/064604

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 + $10 * (2 * \log_d N_1 - 4)$ through $d + d_2$ links.

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 $N_2 > N_1$.

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.

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.

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

WO 2008/147927

PCT/US2008/064604

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:

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).

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.

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

WO 2008/147927

PCT/US2008/064604

$N_1 = p * N_2$ 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. 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.

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).

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)

WO 2008/147927

PCT/US2008/064604

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

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).

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)).

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).

Finally the connection topology of the network 400E shown in FIG. 4E is known to be back to back inverse Benes connection topology.

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

WO 2008/147927

PCT/US2008/064604

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).

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.

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

WO 2008/147927

PCT/US2008/064604

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.

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).

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).

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).

WO 2008/147927

PCT/US2008/064604

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(3,5) 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
 5 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)).

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
 10 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).

Finally the connection topology of the network 400E1 shown in FIG. 4E1 is known to be back to back Omega connection topology.

Referring to FIG. 4E2, in another embodiment of network $V_{fold}(N_1, N_2, d, s)$, an
 15 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
 20 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).

Such a network can be operated in strictly non-blocking manner for unicast
 25 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

WO 2008/147927

PCT/US2008/064604

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.

5 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

10 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

15 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.

20

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

25 connected to middle switches MS(1,1), MS(1,2), MS(1,3), MS(1,4), MS(1,5), MS(1,6),

WO 2008/147927

PCT/US2008/064604

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).

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).

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).

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)).

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

WO 2008/147927

PCT/US2008/064604

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).

Finally the connection topology of the network 400E2 shown in FIG. 4E2 is hereinafter called nearest neighbor connection topology.

5 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 $V_{fold}(N_1, N_2, d, s)$ can comprise any arbitrary type of connection topology. For
10 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
15 $V_{fold}(N_1, N_2, d, s)$ can be built. The embodiments of FIG. 4E, FIG. 4E1, and FIG. 4E2 are only three examples of network $V_{fold}(N_1, N_2, d, s)$.

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
20 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
25 ports.

In the example illustrated in FIG. 4E (or in FIG. 4E1, 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

WO 2008/147927

PCT/US2008/064604

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
 5 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.

The connection request of the type described above can be unicast connection request, a multicast connection request or a broadcast connection request, depending on
 10 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
 15 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:

Network 400F of FIG. 4F is an example of general asymmetrical folded multi-
 20 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 $V_{fold}(N_1, N_2, d, s)$
 25 can be operated in strictly nonblocking manner for unicast if $s \geq 2$ according to the current 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

WO 2008/147927

PCT/US2008/064604

example the links $IL1-IL(p*d)$ to the input switch IS1) and $d + d_1 (= d + p*d)$ 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

5 $2*d$ incoming links for each of $\frac{N_2}{d}$ output switches OS1-OS(N_2/d) (for example $ML(2*Log_d N_2 - 2,1) - ML(2*Log_d N_2 - 2,2*d)$ to the output switch OS1).

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

10 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.

Each of the $2*\frac{N_2}{d}$ middle switches $MS(1,1) - MS(1,2*N_2/d)$ 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.

15 Similarly each of the $2*\frac{N_2}{d}$ middle switches $MS(Log_d N_2 - 1,1) - MS(Log_d N_2 - 1,2*\frac{N_2}{d})$ in the middle stage $130 + 10*(Log_d N_2 - 2)$ are connected from exactly d switches in middle stage $130 + 10*(Log_d N_2 - 3)$ through d links and also are connected to exactly d switches in middle stage $130 + 10*(Log_d N_2 - 1)$ through d links.

WO 2008/147927

PCT/US2008/064604

Similarly each of the $2 \times \frac{N_2}{d}$ middle switches $MS(2 \times \text{Log}_d N_2 - 3, 1)$ -
 $MS(2 \times \text{Log}_d N_2 - 3, 2 \times \frac{N_2}{d})$ 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
 d links and also are connected to exactly d output switches in output stage 120 through
 5 d links.

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 $130 + 10 * (2 * \text{Log}_d N_2 - 4)$ through $2 \times d$ links.

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
 10 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.
 15 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 $N_1 > N_2$.

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
 20 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.

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

WO 2008/147927

PCT/US2008/064604

$I_j = \phi$ 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.

5 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
10 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:

The folded multi-stage network $V_{fold}(N_1, N_2, d, s)$ disclosed, in the current
15 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 Docket No. M-0037US 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 $130 + 10 * (\log_d N_2 - 2)$.

The general symmetrical folded multi-stage network $V_{fold}(N, d, s)$ can also be
20 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:

25 Referring to FIG. 5A, an exemplary symmetrical folded multi-stage network 500A respectively with five stages of twenty switches for satisfying communication

WO 2008/147927

PCT/US2008/064604

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.

5 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).

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

10 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.

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),

15 ML(3,1) - ML(3,8), and ML(4,1) - ML(4,8) are connected between the respective stages is different.

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,

20 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 $V_{fold}(N, d, s)$.

25 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

WO 2008/147927

PCT/US2008/064604

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-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.

In network 500A of FIG. 5A, each of the $\frac{N}{d}$ 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).

Each of the $\frac{N}{d}$ 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).

Similarly each of the $\frac{N}{d}$ 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

WO 2008/147927

PCT/US2008/064604

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
 5 MS(3,3) respectively).

Similarly each of the $\frac{N}{d}$ 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
 10 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)).

Each of the $\frac{N}{d}$ output switches OS1 – OS4 are connected from exactly d switches in middle stage 150 through d links (for example output switch OS1 is
 15 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:

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
 20 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 $\frac{N}{d}$ 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 $\frac{N}{d}$
 25 input switches IS1-IS(N/d) (for example the links ML(1,1) - ML(1,d) to the input switch

WO 2008/147927

PCT/US2008/064604

IS1). There are d outlet links for each of $\frac{N}{d}$ output switches OS1-OS(N/d) (for example OL1-OL(d) to the output switch OS1) and d incoming links for each of $\frac{N}{d}$ output switches OS1-OS(N/d) (for example $ML(2 \times \text{Log}_d N - 2, 1) - ML(2 \times \text{Log}_d N - 2, d)$ to the output switch OS1).

- 5 Each of the $\frac{N}{d}$ input switches IS1 – IS(N/d) are connected to exactly d switches in middle stage 130 through d links.

Each of the $\frac{N}{d}$ 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.

- 10 Similarly each of the $\frac{N}{d}$ middle switches $MS(\text{Log}_d N - 1, 1) - MS(\text{Log}_d N - 1, \frac{N}{d})$ 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.

- Similarly each of the $\frac{N}{d}$ middle switches $MS(2 \times \text{Log}_d N - 3, 1) - MS(2 \times \text{Log}_d N - 3, \frac{N}{d})$ 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.

Each of the $\frac{N}{d}$ 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.

WO 2008/147927

PCT/US2008/064604

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:

5 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
10 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).

Such networks can be operated in rearrangeably nonblocking manner for unicast
15 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.

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
20 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.

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
25 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

WO 2008/147927

PCT/US2008/064604

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 $V_{fold}(N_1, N_2, d, s)$.

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

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

WO 2008/147927

PCT/US2008/064604

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).

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).

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).

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)).

WO 2008/147927

PCT/US2008/064604

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:

10 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 $N_2 = p * N_1$ where $p > 1$. In network 500D of FIG. 5D, $N_1 = N$ and $N_2 = p * N$. 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 $s \geq 1$ according to the current

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

20 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 $ML(2 \times \log_d N_1 - 2, 1) - ML(2 \times \log_d N_1 - 2, d_2)$ to the output switch OS1).

WO 2008/147927

PCT/US2008/064604

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.

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.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(\text{Log}_d N_1 - 1, 1)$ - $MS(\text{Log}_d N_1 - 1, \frac{N_1}{d})$ 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.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(2 * \text{Log}_d N_1 - 3, 1)$ - $MS(2 * \text{Log}_d N_1 - 3, \frac{N_1}{d})$ 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_2 links.

Each of the $\frac{N_1}{d}$ output switches OS1 – OS(N_1/d) are connected from exactly d switches in middle stage $130 + 10 * (2 * \text{Log}_d N - 4)$ through d_2 links.

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 $s \geq 1$ according to the current invention.

20

WO 2008/147927

PCT/US2008/064604

Asymmetric folded RNB ($N_1 > N_2$) Unicast Embodiments:

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).

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.

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.

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 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 $V_{fold}(N_1, N_2, d, s)$.

WO 2008/147927

PCT/US2008/064604

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

5 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

10 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

15 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.

In network 500E of FIG. 5E, each of the $\frac{N_2}{d}$ input switches IS1 – IS4 are

20 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

25 the links ML(1,6)).

WO 2008/147927

PCT/US2008/064604

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_1 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).

10 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).

15 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 and OS2 respectively from middle switch MS(3,1)).

25 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

WO 2008/147927

PCT/US2008/064604

connected from middle switches MS(3,1) and MS(3,2) through the links ML(4,1) and ML(4,4) respectively).

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

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 $d_1 (= p \times d)$ 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,($d+p*d$)) 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).

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.

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.

WO 2008/147927

PCT/US2008/064604

Similarly each of the $\frac{N_2}{d}$ middle switches $MS(\text{Log}_d N_2 - 1, 1)$ -
 $MS(\text{Log}_d N_2 - 1, \frac{N_2}{d})$ 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 d links and also are
 connected to exactly d switches in middle stage $130 + 10 * (\text{Log}_d N_2 - 1)$ through d
 5 links.

Similarly each of the $\frac{N_2}{d}$ middle switches $MS(2 * \text{Log}_d N_2 - 3, 1)$ -
 $MS(2 * \text{Log}_d N_2 - 3, \frac{N_2}{d})$ 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 d links
 and also are connected to exactly d output switches in output stage 120 through d links.
 10 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 d links.

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.

15 Symmetric RNB Unicast Embodiments:

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
 20 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

WO 2008/147927

PCT/US2008/064604

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).

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.

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.

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 $V(N, d, s)$.

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 $\frac{N}{d}$, where N is the total number of inlet links or

WO 2008/147927

PCT/US2008/064604

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 $d * d$ 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.

In network 600A of FIG. 6A, each of the $\frac{N}{d}$ 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).

Each of the $\frac{N}{d}$ 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).

Similarly each of the $\frac{N}{d}$ 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

WO 2008/147927

PCT/US2008/064604

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).

Similarly each of the $\frac{N}{d}$ middle switches MS(3,1) – MS(3,4) in the middle stage
 5 150 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
 10 MS(3,1)).

Each of the $\frac{N}{d}$ 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).

15 **Generalized Symmetric RNB Unicast Embodiments:**

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, $s = 1$).
 20 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ 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 $\frac{N}{d}$ output switches OS1-OS(N/d) (for example OL1-OL(d) to the output

WO 2008/147927

PCT/US2008/064604

switch OS1) and d incoming links for each of $\frac{N}{d}$ 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).

Each of the $\frac{N}{d}$ input switches IS1 – IS(N/d) are connected to exactly d switches in middle stage 130 through d links.

- 5 Each of the $\frac{N}{d}$ 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.

Similarly each of the $\frac{N}{d}$ 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 d links and also are connected to exactly d switches in middle stage $130 + 10 * (\log_d N - 1)$ through d links.

Similarly each of the $\frac{N}{d}$ middle switches $MS(2 \times \log_d N - 3, 1) - MS(2 \times \log_d N - 3, \frac{N}{d})$ in the middle stage $130 + 10 * (2 * \log_d N - 4)$ 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.

Each of the $\frac{N}{d}$ 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.

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.

WO 2008/147927

PCT/US2008/064604

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

Referring to FIG. 6A1, FIG. 6B1, FIG. 6C1, FIG. 6D1, FIG. 6E1, FIG. 6F1, FIG. 6G1, FIG. 600H1, FIG. 600I1 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).

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.

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.

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

WO 2008/147927

PCT/US2008/064604

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 $V(N_1, N_2, d, s)$.

- 5 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
- 10 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
- $$d_2 = N_2 \times \frac{d}{N_1} = p \times d .$$
- 15 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
- 20 (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.

- 25 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

WO 2008/147927

PCT/US2008/064604

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).

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).

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).

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)).

WO 2008/147927

PCT/US2008/064604

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:

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 $N_2 = p * N_1$ 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 ($= p \times d$) 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($2 \times \log_d N_1 - 2, d_2$) to the output switch OS1).

WO 2008/147927

PCT/US2008/064604

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.

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.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(\text{Log}_d N_1 - 1, 1)$ - $MS(\text{Log}_d N_1 - 1, \frac{N_1}{d})$ 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.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(2 * \text{Log}_d N_1 - 3, 1)$ - $MS(2 * \text{Log}_d N_1 - 3, \frac{N_1}{d})$ 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_2 links.

Each of the $\frac{N_1}{d}$ output switches OS1 – OS(N_1/d) are connected from exactly d switches in middle stage $130 + 10 * (2 * \text{Log}_d N - 4)$ through d_2 links.

The general symmetrical multi-stage network $V(N_1, N_2, d, s)$ can be operated in rearrangeably nonblocking manner for multicast when $s = 1$ according to the current invention.

20

WO 2008/147927

PCT/US2008/064604

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

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).

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.

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.

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

WO 2008/147927

PCT/US2008/064604

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. 6A2 to FIG. 6J2 are only three examples of network $V(N_1, N_2, d, s)$.

- 5 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 N_1 is the total number of inlet links or
- 10 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
- $$d_1 = N_1 \times \frac{d}{N_2} = p \times d .$$
- 15 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 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
- 20 (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.

- 25 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

WO 2008/147927

PCT/US2008/064604

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)).

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_1 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).

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).

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

WO 2008/147927

PCT/US2008/064604

ML(4,2) are connected to output switches OS1 and OS2 respectively from middle switch MS(3,1)).

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:

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 $N_1 = p * N_2$ where $p > 1$. In network 400K2 of FIG. 4K2, $N_2 = N$ and $N_1 = p * N$. 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. 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($p*d$) to the input switch IS1) and $d_1 (= p \times d)$ 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,($d+p*d$)) 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).

WO 2008/147927

PCT/US2008/064604

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.

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.

Similarly each of the $\frac{N_2}{d}$ middle switches $MS(\text{Log}_d N_2 - 1, 1)$ - $MS(\text{Log}_d N_2 - 1, \frac{N_2}{d})$ 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 d links and also are connected to exactly d switches in middle stage $130 + 10 * (\text{Log}_d N_2 - 1)$ through d links.

Similarly each of the $\frac{N_2}{d}$ middle switches $MS(2 * \text{Log}_d N_2 - 3, 1)$ - $MS(2 * \text{Log}_d N_2 - 3, \frac{N_2}{d})$ 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 d links and also are connected to exactly d output switches in output stage 120 through d links.

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 d links.

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.

20

WO 2008/147927

PCT/US2008/064604

Scheduling Method Embodiments:

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 $V_{mink}(N_1, N_2, d, s)$ 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.

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 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 1030.

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 $130+10*(2*\text{Log}_d N - 4)$ from which the output switch is reachable, are derived. Next, starting from the selected middle

WO 2008/147927

PCT/US2008/064604

switches in middle stage $130+10*(2*Log_d N - 4)$ traveling backward through all of their available incoming middle links from middle stage $130+10*(2*Log_d N - 5)$ all the available middle switches in middle stage $130+10*(2*Log_d N - 5)$ from which the output switch is reachable, are derived. This process is repeated recursively until all the
5 middle switches in middle stage $130+10*(Log_d N - 2)$ 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 $130+10*(2*Log_d N - 4)$ to middle stage $130+10*(Log_d N - 2)$ for all the output switches of each destination link of the connection are derived. Then the control goes to
10 act 1040.

In act 1040, using the lists generated in acts 1020 and 1030, particularly list of middle switches derived in middle stage $130+10*(Log_d N - 2)$ corresponding to each outgoing link of the input switch of the multicast connection, and the list of middle switches derived in middle stage $130+10*(Log_d N - 2)$ corresponding to each output
15 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 $130+10*(Log_d N - 2)$ is reachable from an outgoing link of the input switch, say “x”, and also from the same middle switch in middle stage $130+10*(Log_d N - 2)$ if the output switch of a destination link, say “y”, is reachable then using the outgoing link of
20 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.

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
25 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.

WO 2008/147927

PCT/US2008/064604

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
5 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
10 transfers to act 1090.

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
15 1090 are executed in a loop, for each connection request until the connections are set up.

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
20 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
25 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 $V_{mlink}(N_1, N_2, d, s)$ and the network $V(N_1, N_2, d, s)$ to be operated in nonblocking manner in accordance with the invention.

WO 2008/147927

PCT/US2008/064604

According to the current invention, using the method 1040 of FIG. 7A, the network $V_{mlink}(N_1, N_2, d, s)$ 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 $s \geq 2$.

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.

As noted above method 1000 of FIG. 7A can be used to setup multicast connections, unicast connections, or broadcast connection of all the networks $V_{mlink}(N, d, s)$, $V_{mlink}(N_1, N_2, d, s)$, $V(N, d, s)$ and $V(N_1, N_2, d, s)$ disclosed in this invention.

Applications Embodiments:

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

WO 2008/147927

PCT/US2008/064604

diagram 400A, outlet link OL2 of diagram 800A1 corresponds to middle link ML(2,2) of diagram 400A.

1) Programmable Integrated Circuit Embodiments:

All the embodiments disclosed in the current invention are useful in
5 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 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
10 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)
15 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
20 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
25 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.

WO 2008/147927

PCT/US2008/064604

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. 8A3 illustrates the detailed diagram 800A3 for the implementation of the diagram 800A1 in one-time programmable
5 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
10 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
15 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
20 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
25 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.

WO 2008/147927

PCT/US2008/064604

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. 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.

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.

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

WO 2008/147927

PCT/US2008/064604

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:

5 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.

Numerous modifications and adaptations of the embodiments, implementations, and examples described herein will be apparent to the skilled artisan in view of the
10 disclosure.

WO 2008/147927

PCT/US2008/064604

CLAIMS

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
 - 5 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
 - 10 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 $\frac{N}{d}$ middle switches in each of said y middle stages wherein said second stage and said penultimate stage are one of said
 - 15 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
 - 20 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

WO 2008/147927

PCT/US2008/064604

when $N_1 > N_2$ and $N_1 = p * N_2$ where $p > 1$ then $N_2 = N$, $d_2 = d$ and

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

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

5 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 $\frac{N}{d}$ middle switches in each of said y

10 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

15 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

20 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 switches in each said middle stage and in said output stage.

WO 2008/147927

PCT/US2008/064604

2. The network of claim 1, wherein all said incoming middle links and outgoing middle links are connected in any 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.
- 5 3. The network of claim 2, wherein $y \geq (2 \times \log_d N_1) - 3$ when $N_2 > N_1$, and $y \geq (2 \times \log_d N_2) - 3$ when $N_1 > N_2$.
4. The network of claim 3, wherein $x \geq 1$, wherein said each multicast connection comprises only one destination link, and
said each multicast connection from an inlet link passes through only one
10 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
15 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 $x \geq 2$, wherein said each multicast connection comprises only one destination link, and
said each multicast connection from an inlet link passes through only one
20 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
25 connection comprises only one destination link and the network is hereinafter "strictly nonblocking network for unicast".
6. The network of claim 3, wherein $x \geq 2$,

WO 2008/147927

PCT/US2008/064604

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".

5 7. The network of claim 3, wherein $x \geq 3$,

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
10 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.

15 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

20 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

WO 2008/147927

PCT/US2008/064604

- 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 $\frac{N}{d}$ 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
- 5 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
- 10 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
- 15 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 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
- 20 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

WO 2008/147927

PCT/US2008/064604

- a plurality of y middle stages having $\frac{N}{d}$ 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
- 5 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
- 10 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
- 15 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
- 20 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
- 25 further comprises:

WO 2008/147927

PCT/US2008/064604

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.

5 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 $N_2 = p * N_1$ 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

10 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

15 a plurality of y middle stages having $\frac{N}{d}$ 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,

20 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

WO 2008/147927

PCT/US2008/064604

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

$$d_1 = N_1 \times \frac{d}{N_2} = p \times d ; \text{ and having}$$

5 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

10 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 $\frac{N}{d}$ 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

15 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

20 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

WO 2008/147927

PCT/US2008/064604

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
5 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
10 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
15 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
20 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
25 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

WO 2008/147927**PCT/US2008/064604**

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.

5

FIG. 1A

100A

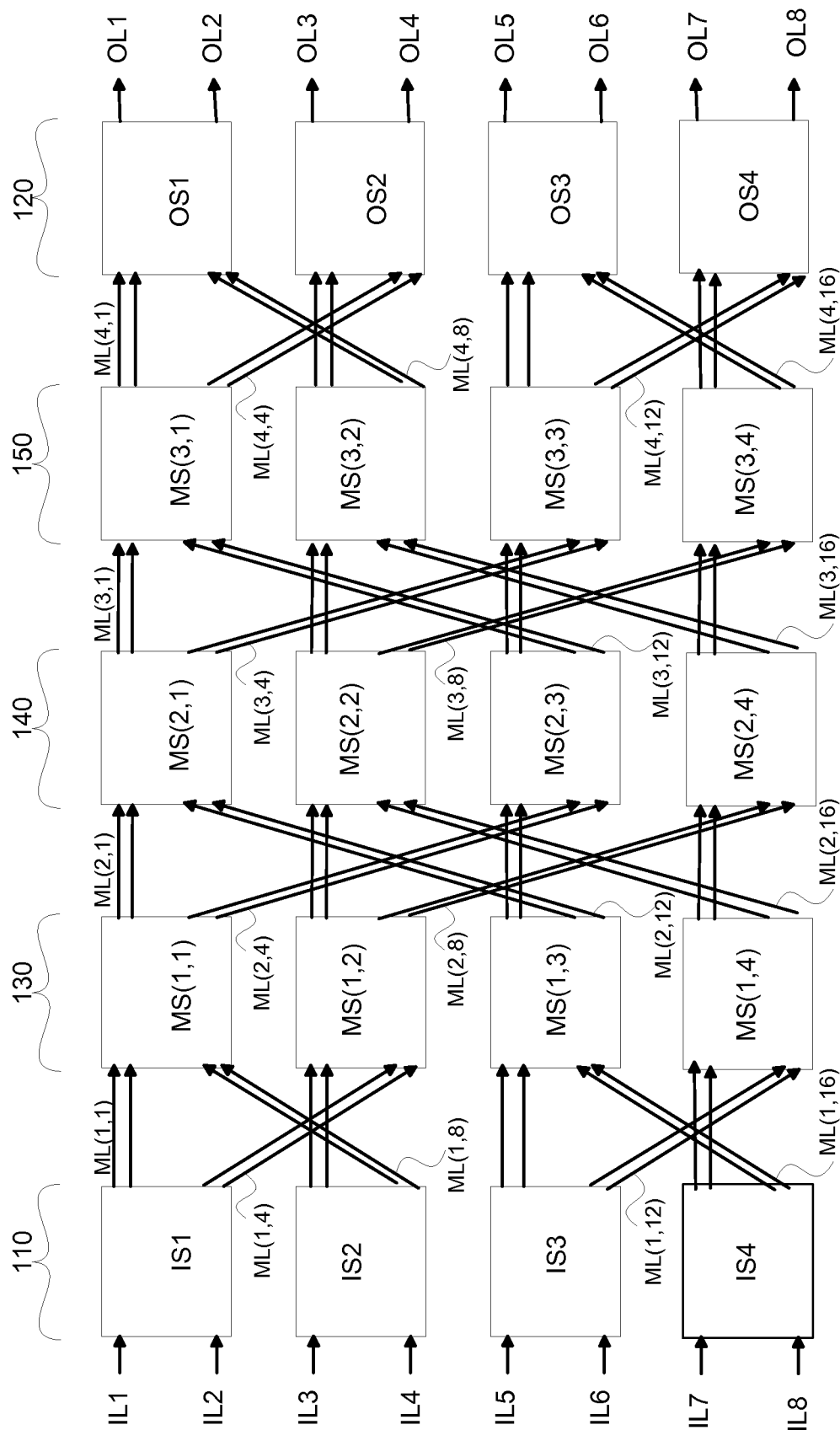


FIG. 1B

100B

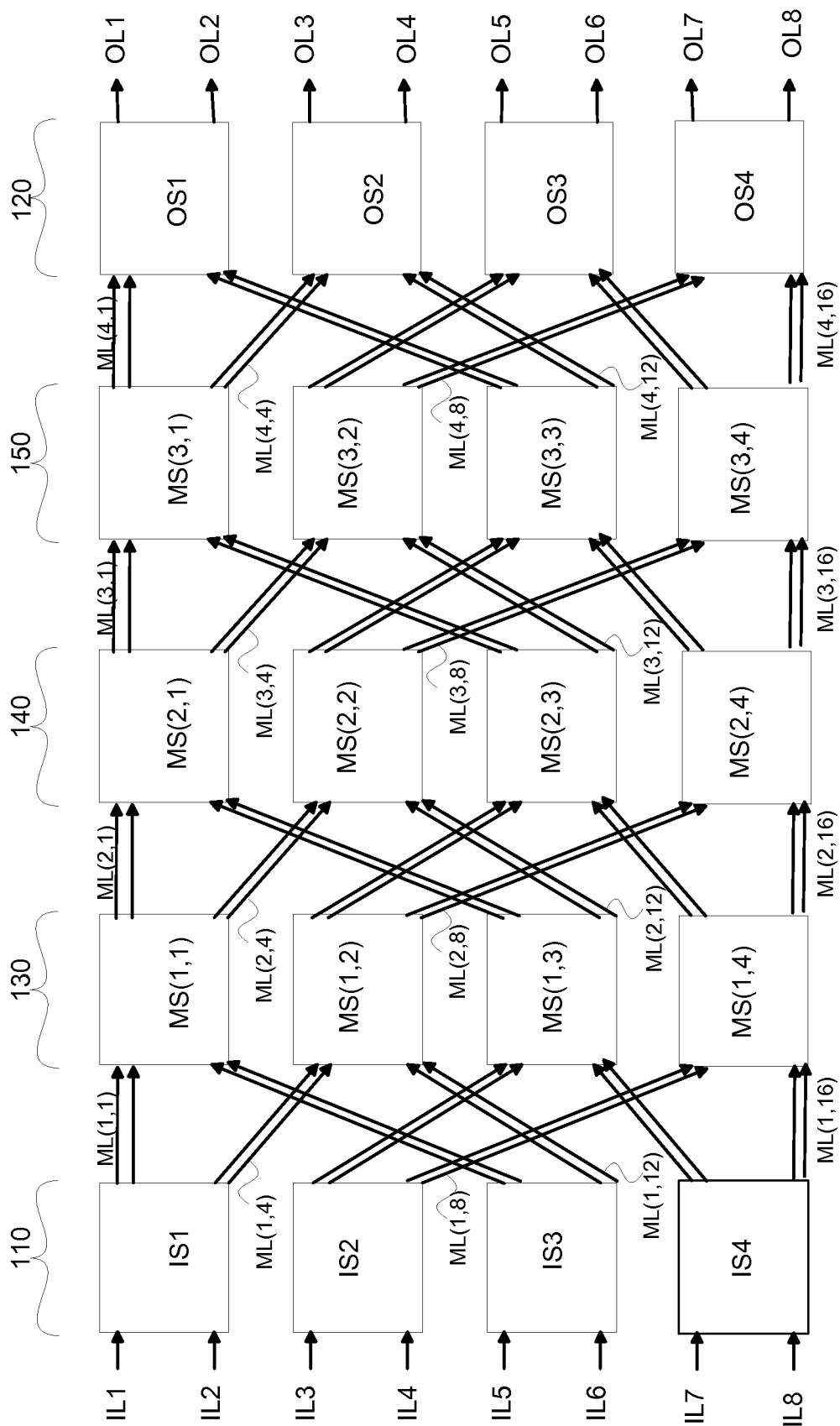


FIG. 1C

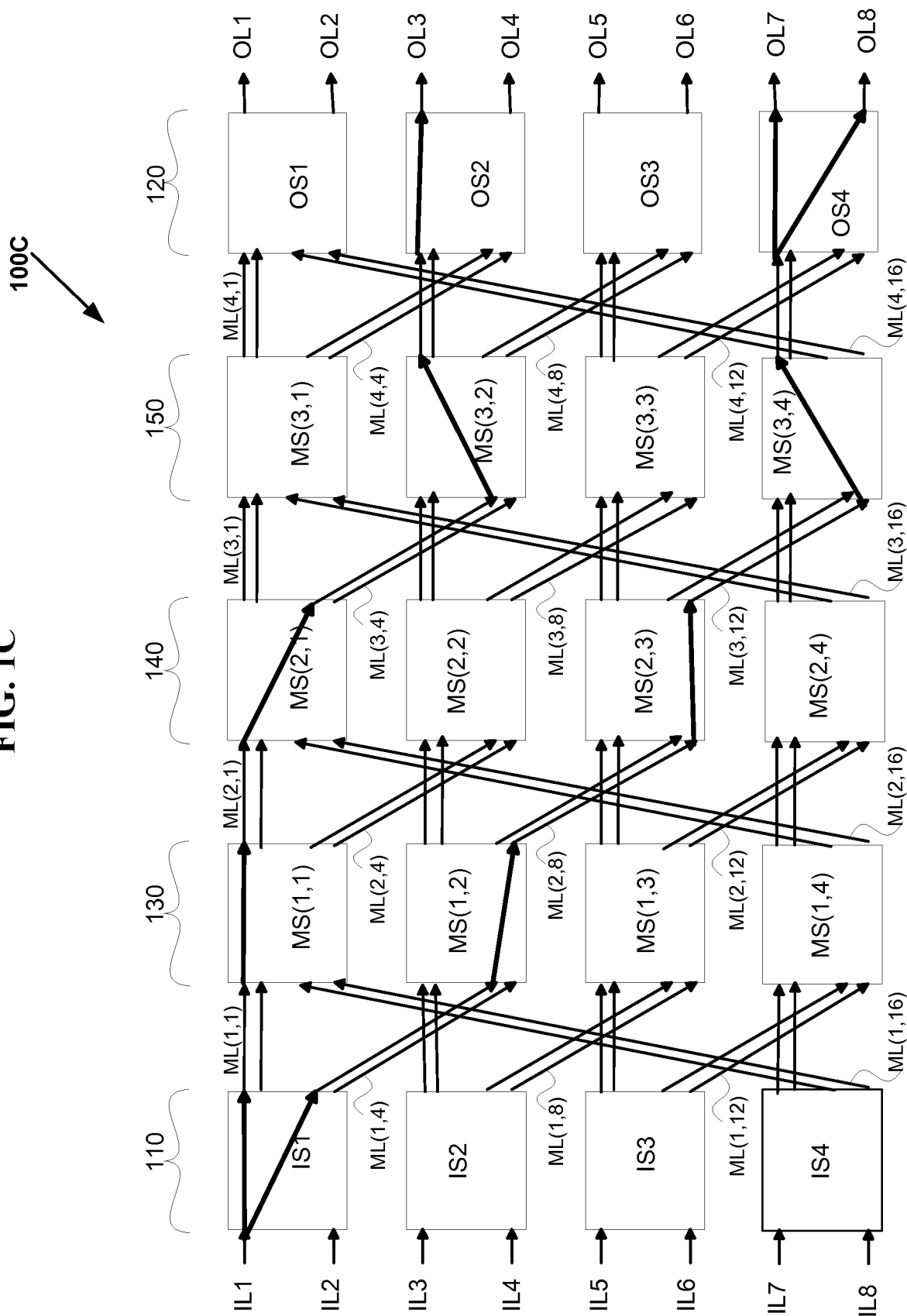


FIG. 1D

100D

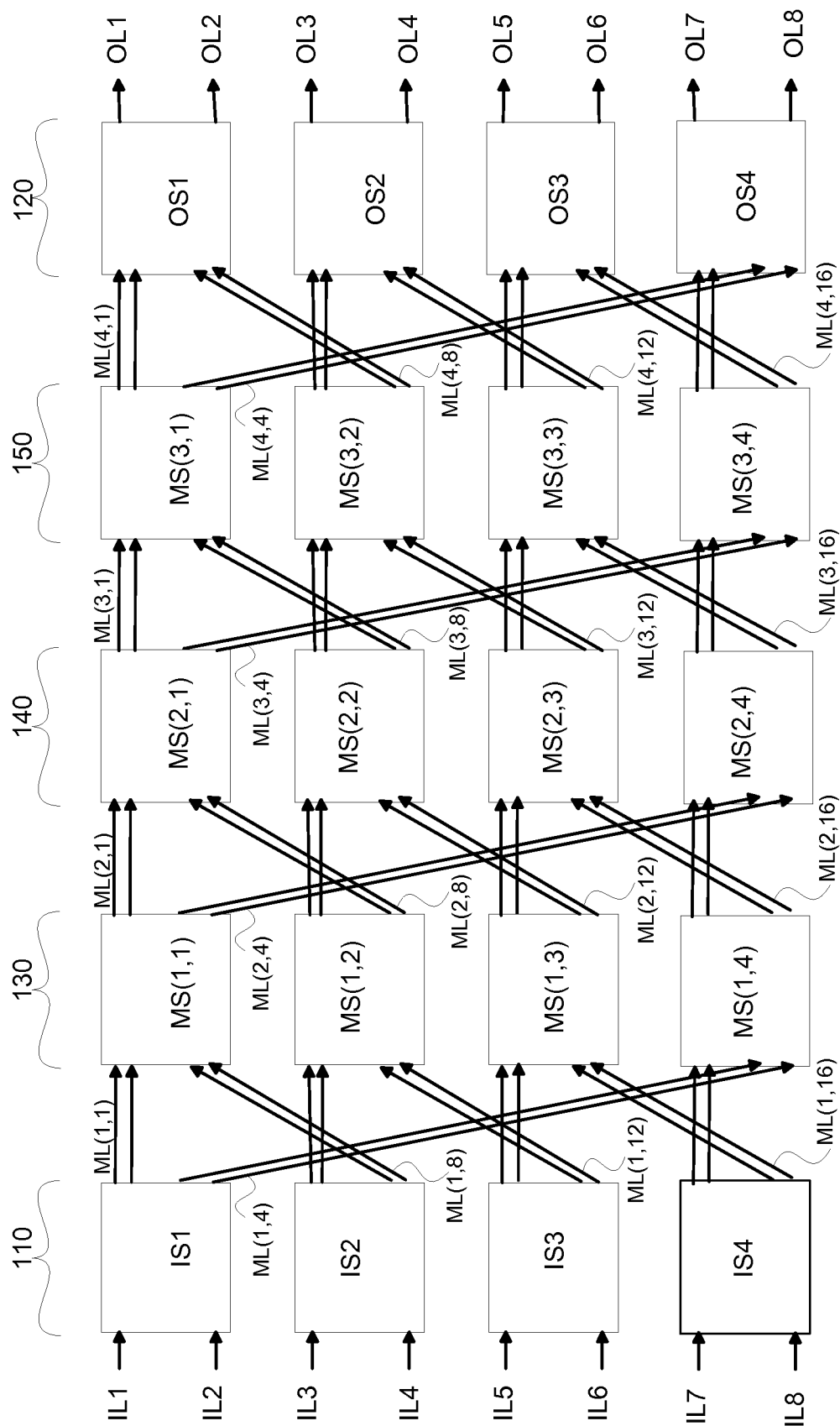


FIG. 1E

100E

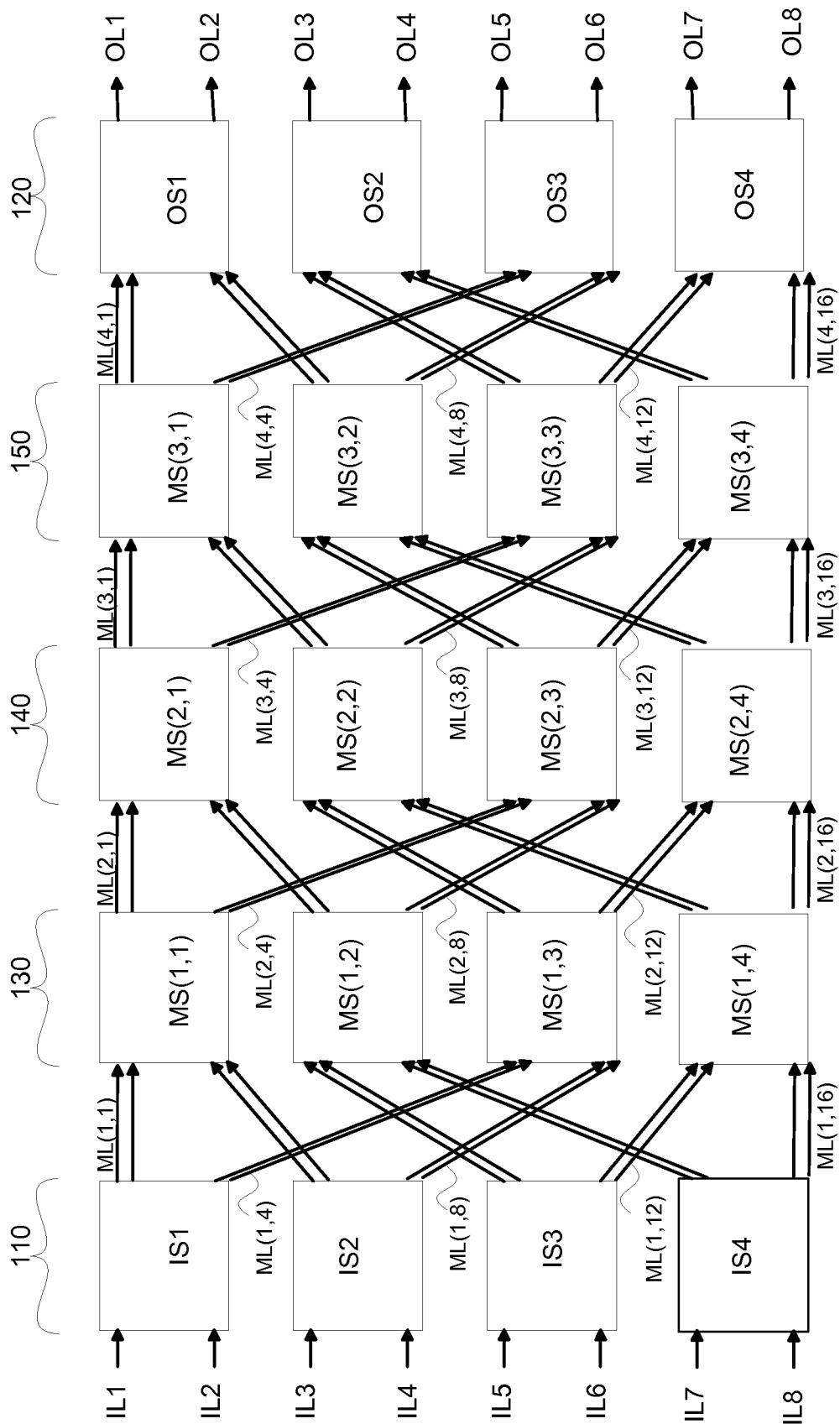


FIG. 1F

100F

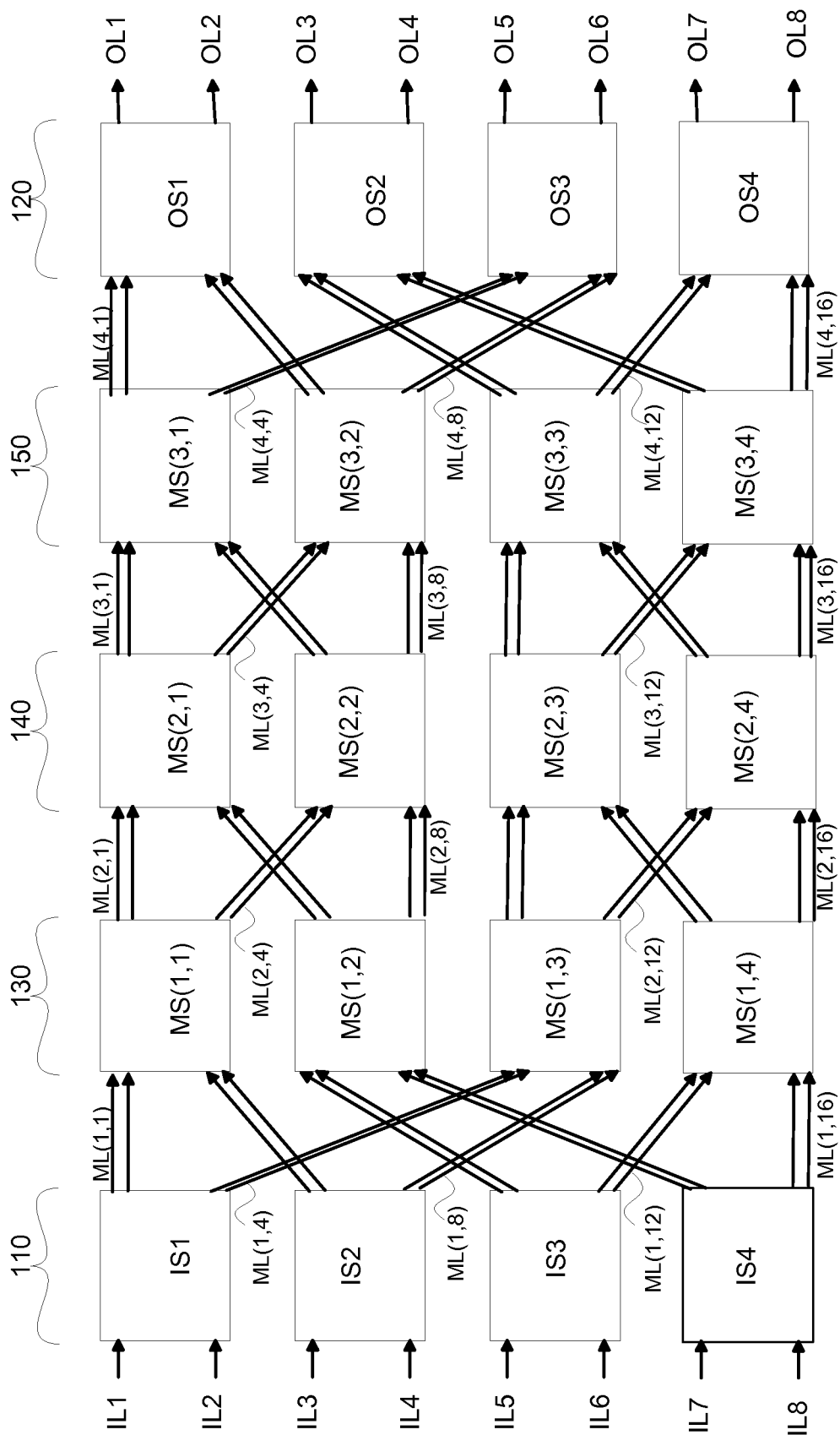


FIG. 1G

100G

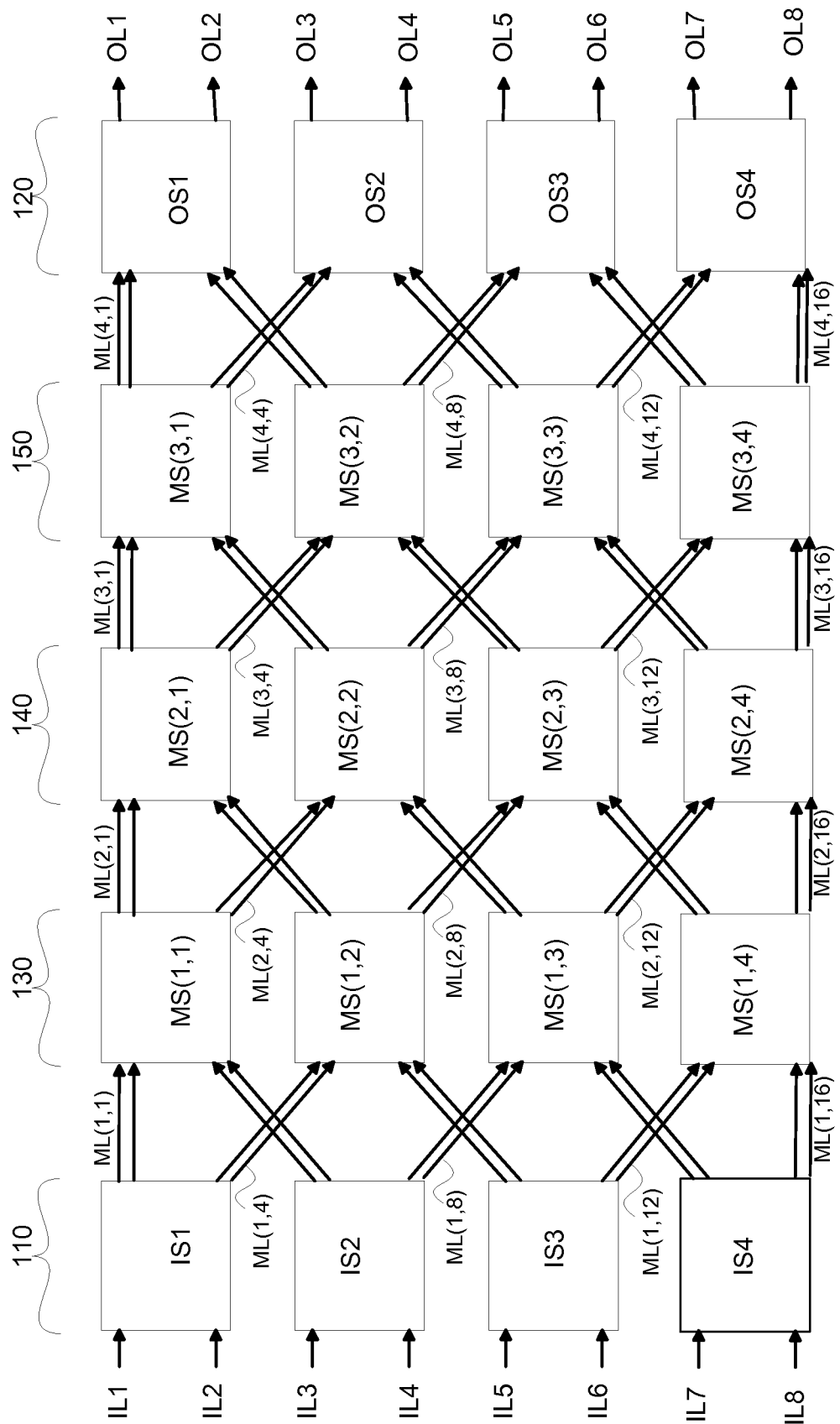


FIG. 1H

100H

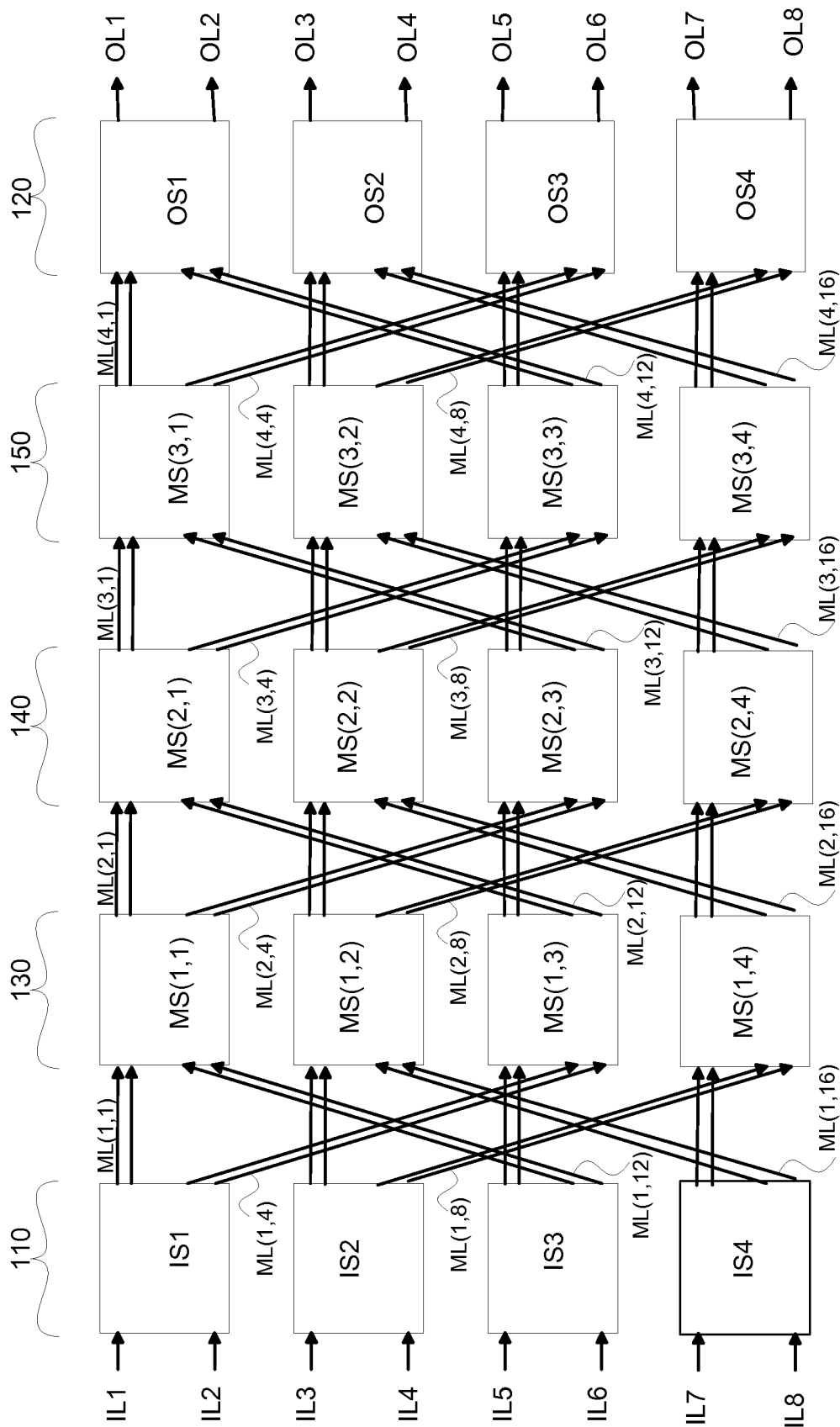


FIG. 11

100I

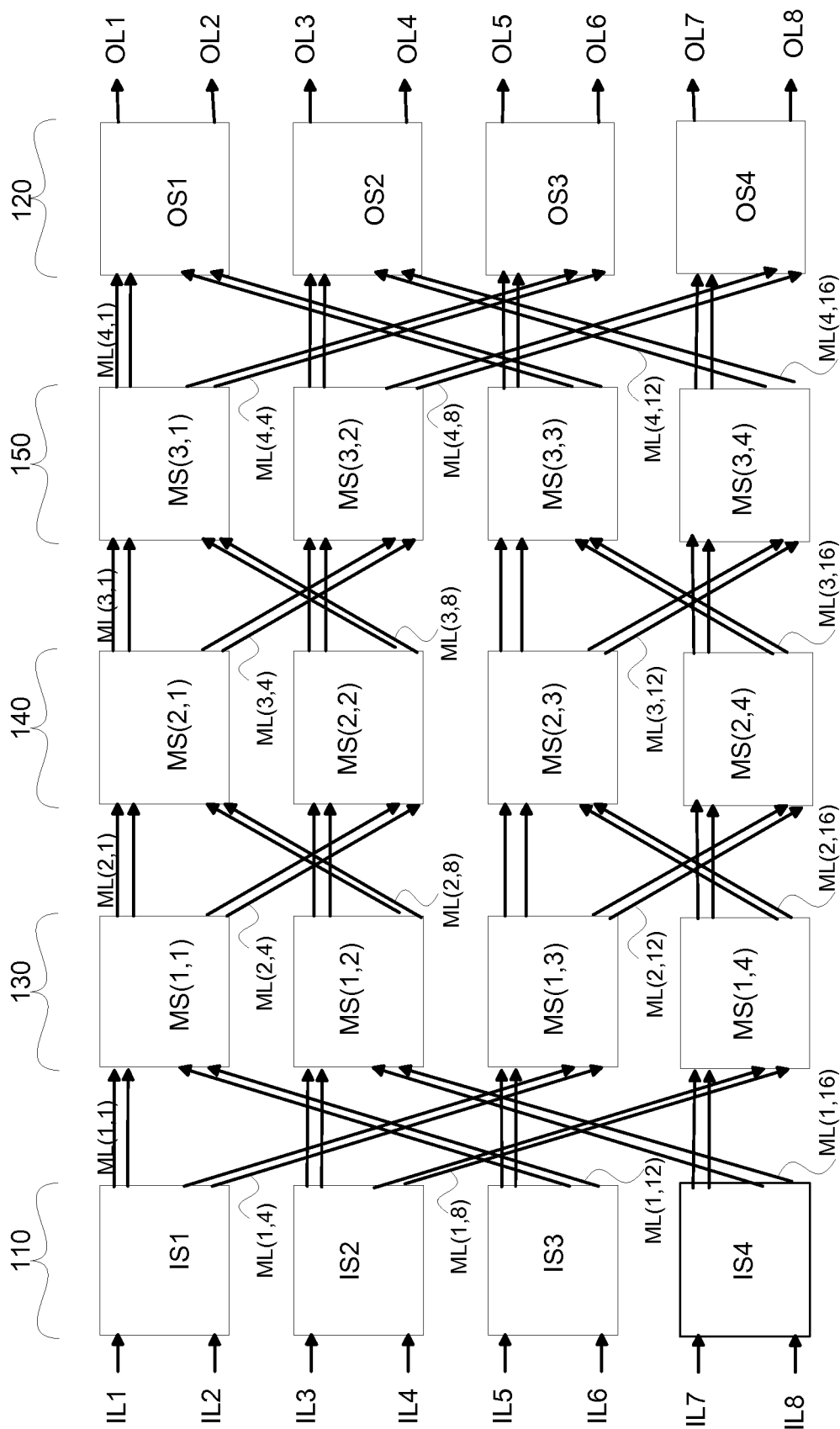


FIG. 1J

100J

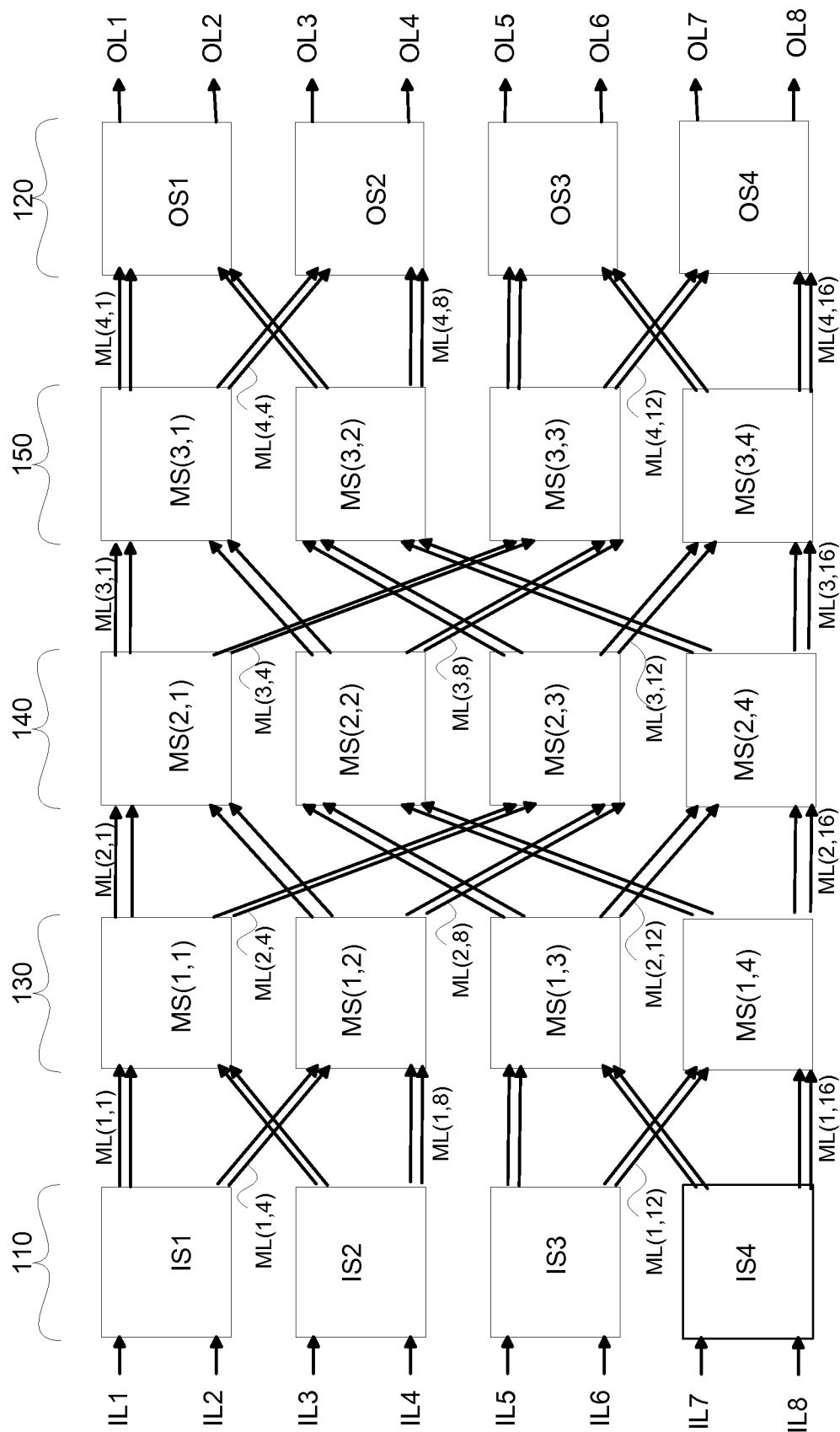


FIG. 1K

100K

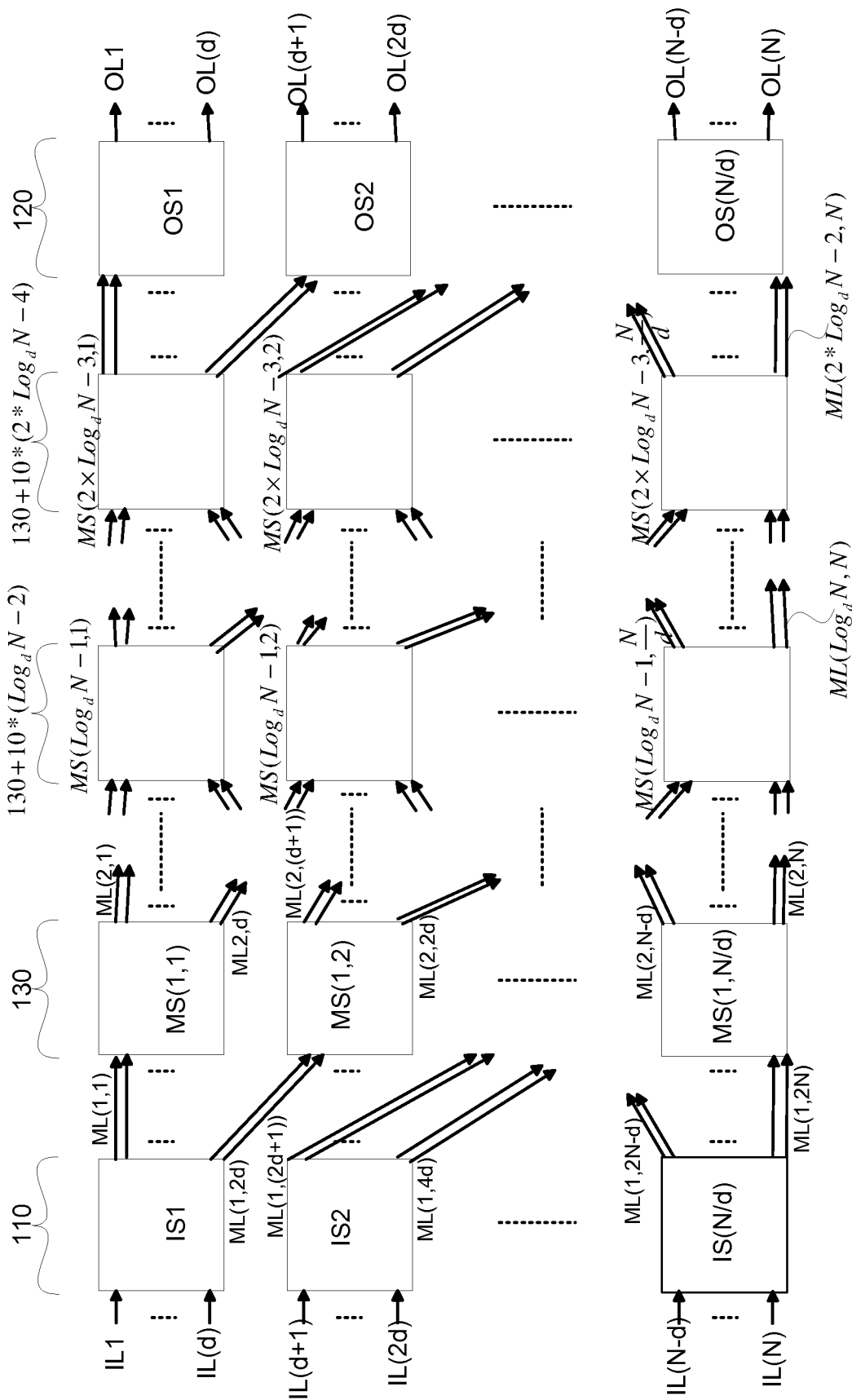


FIG. 1A1

100A1

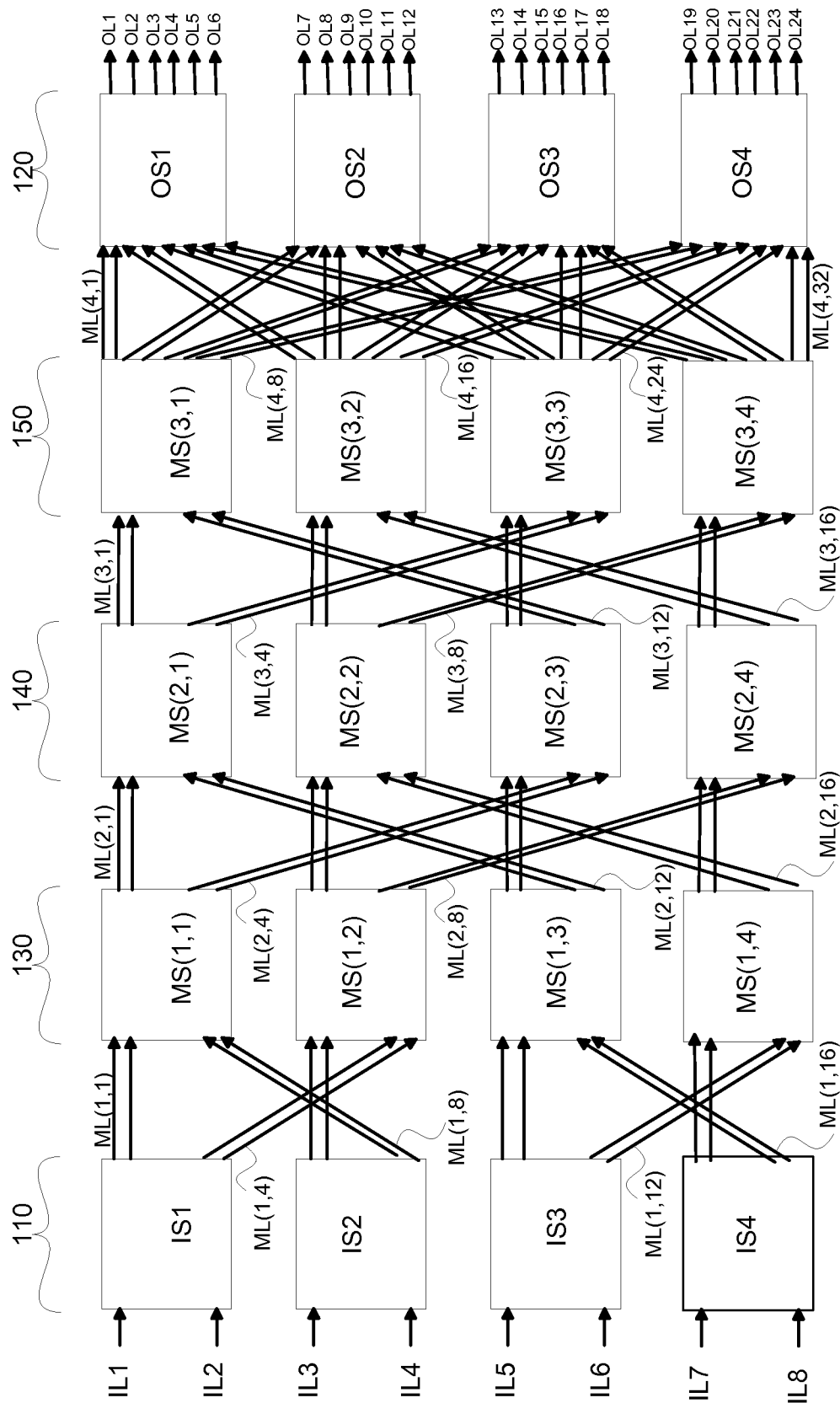


FIG. 1B1

100B1

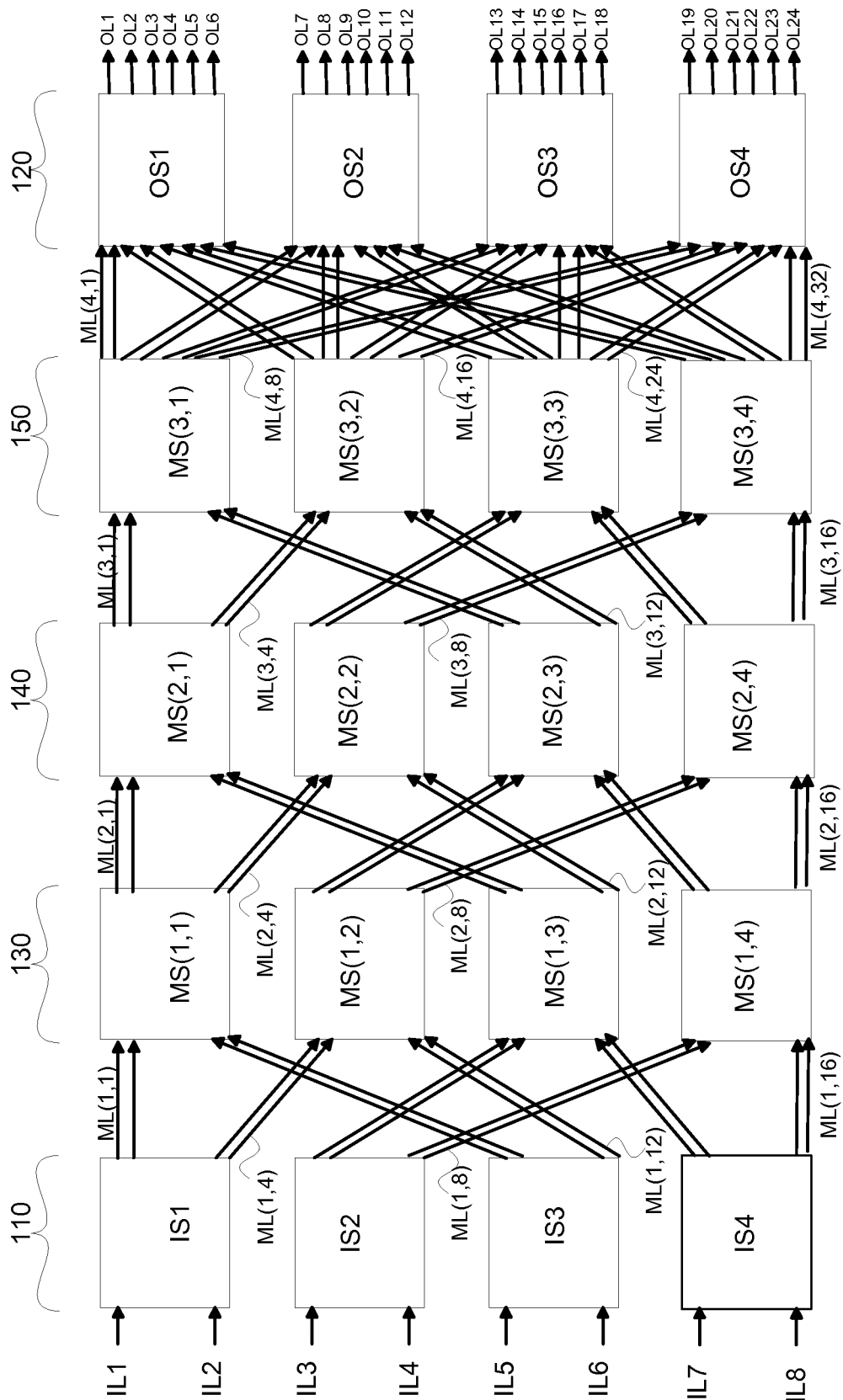


FIG. 1C1

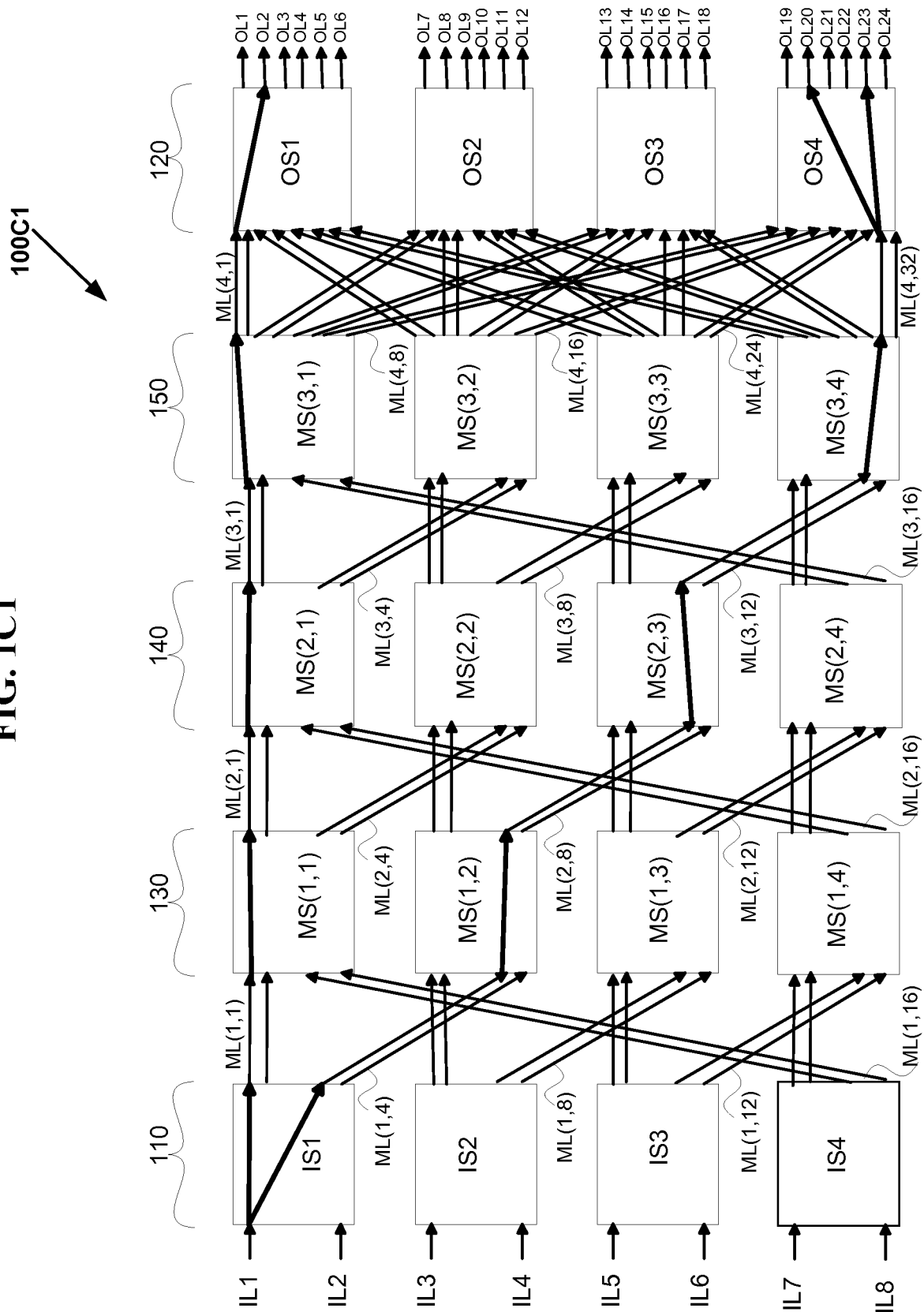


FIG. 1D1

100D1

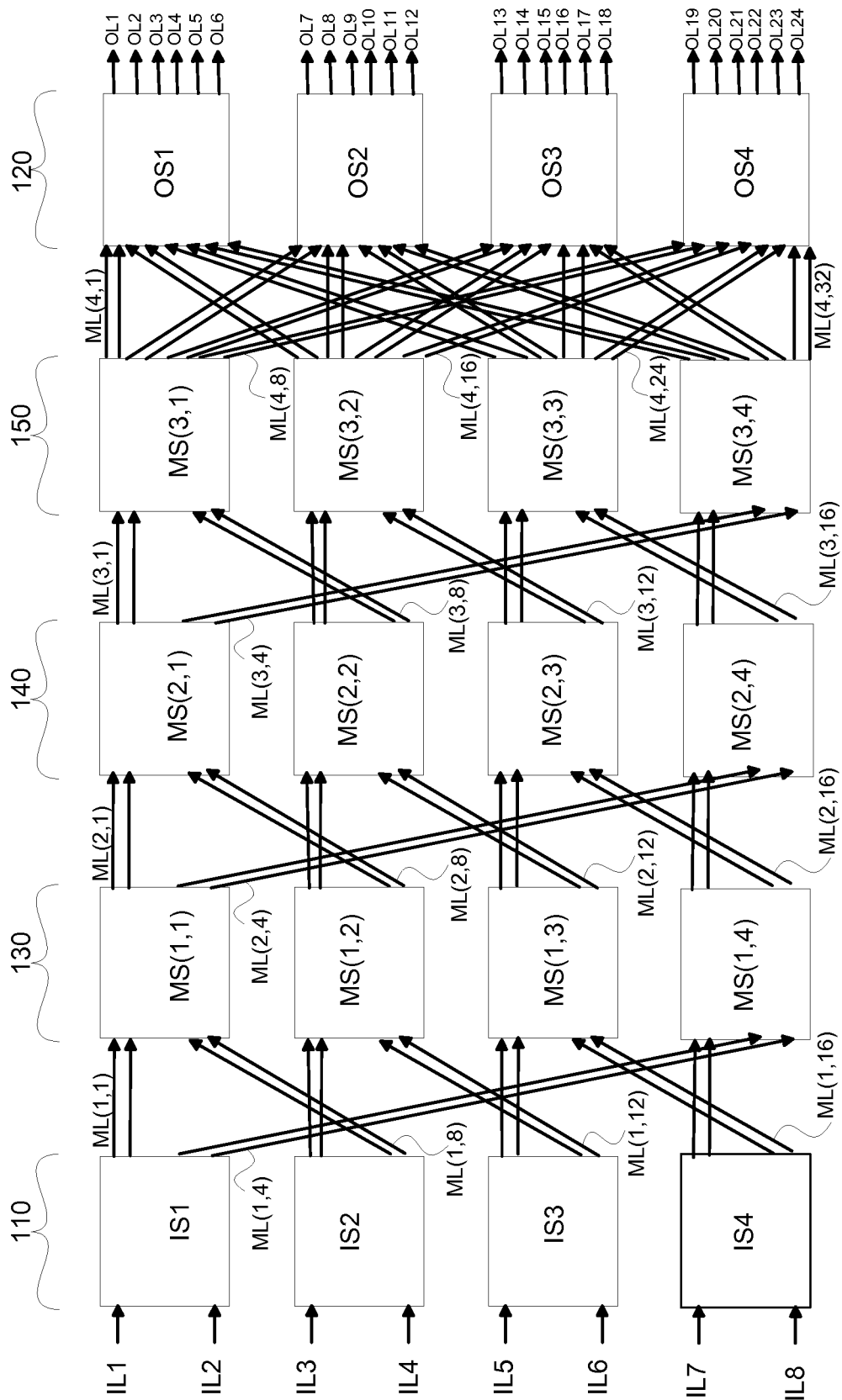


FIG. 1E1

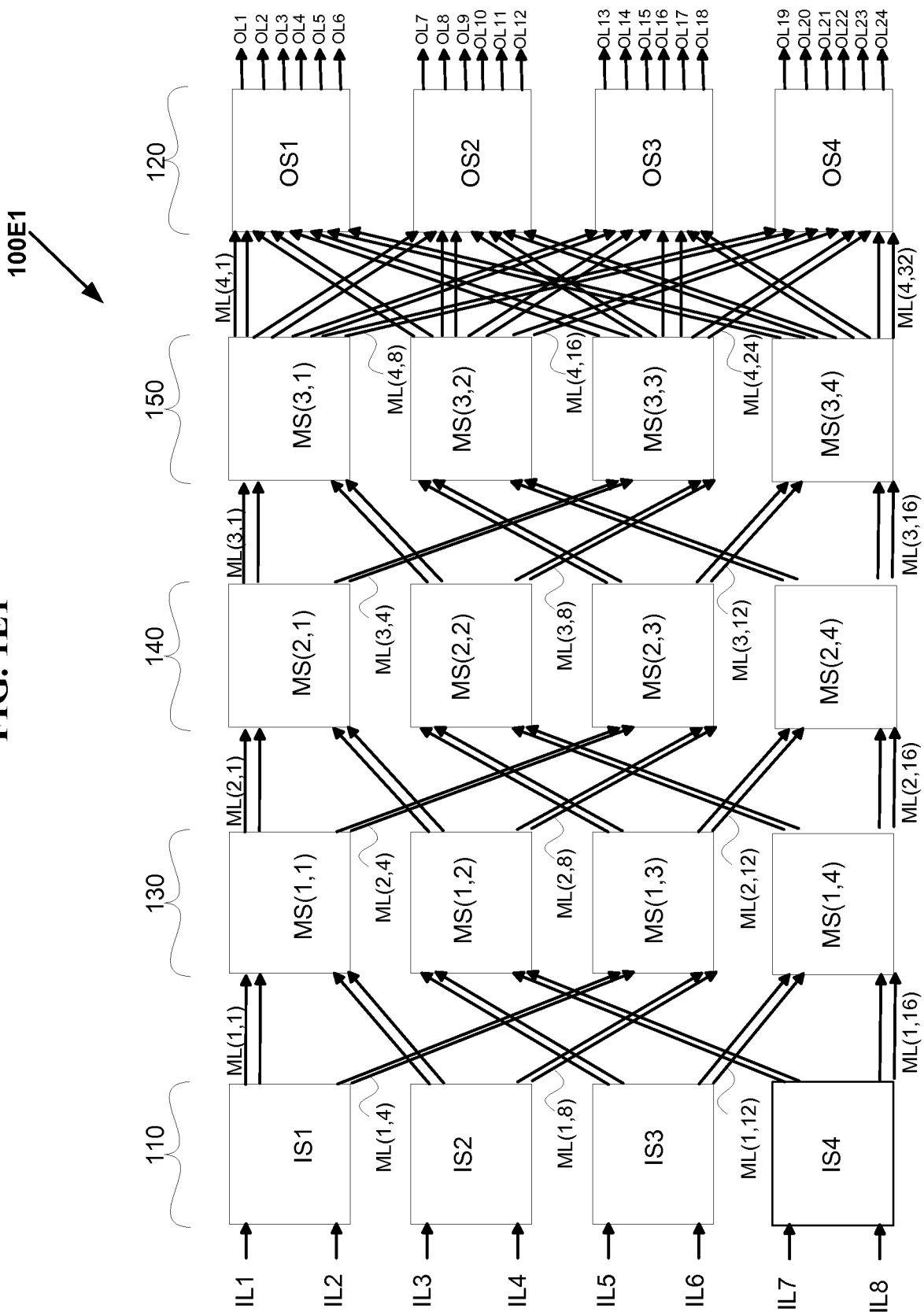


FIG. 1F1

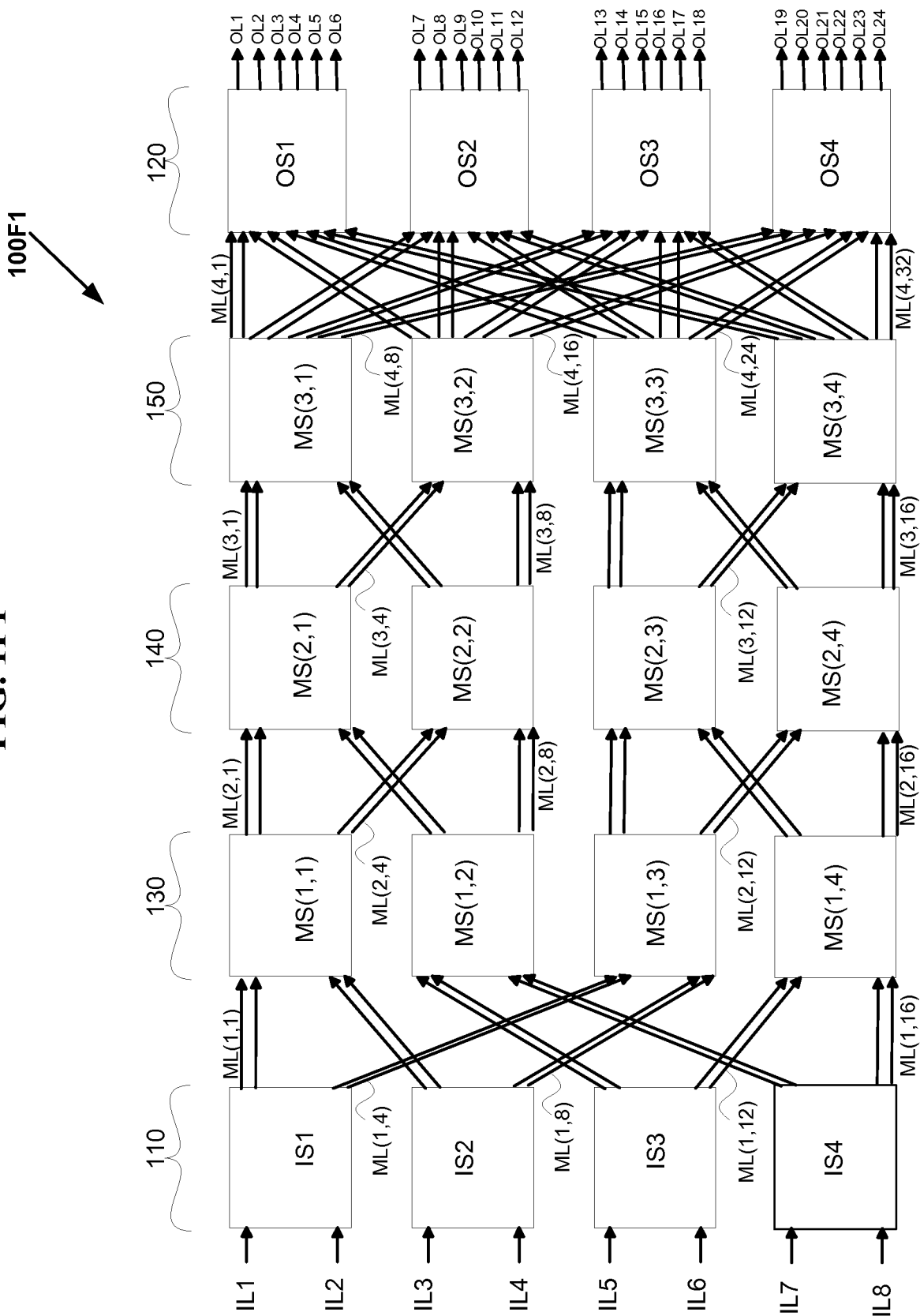


FIG. 1G1

100G1

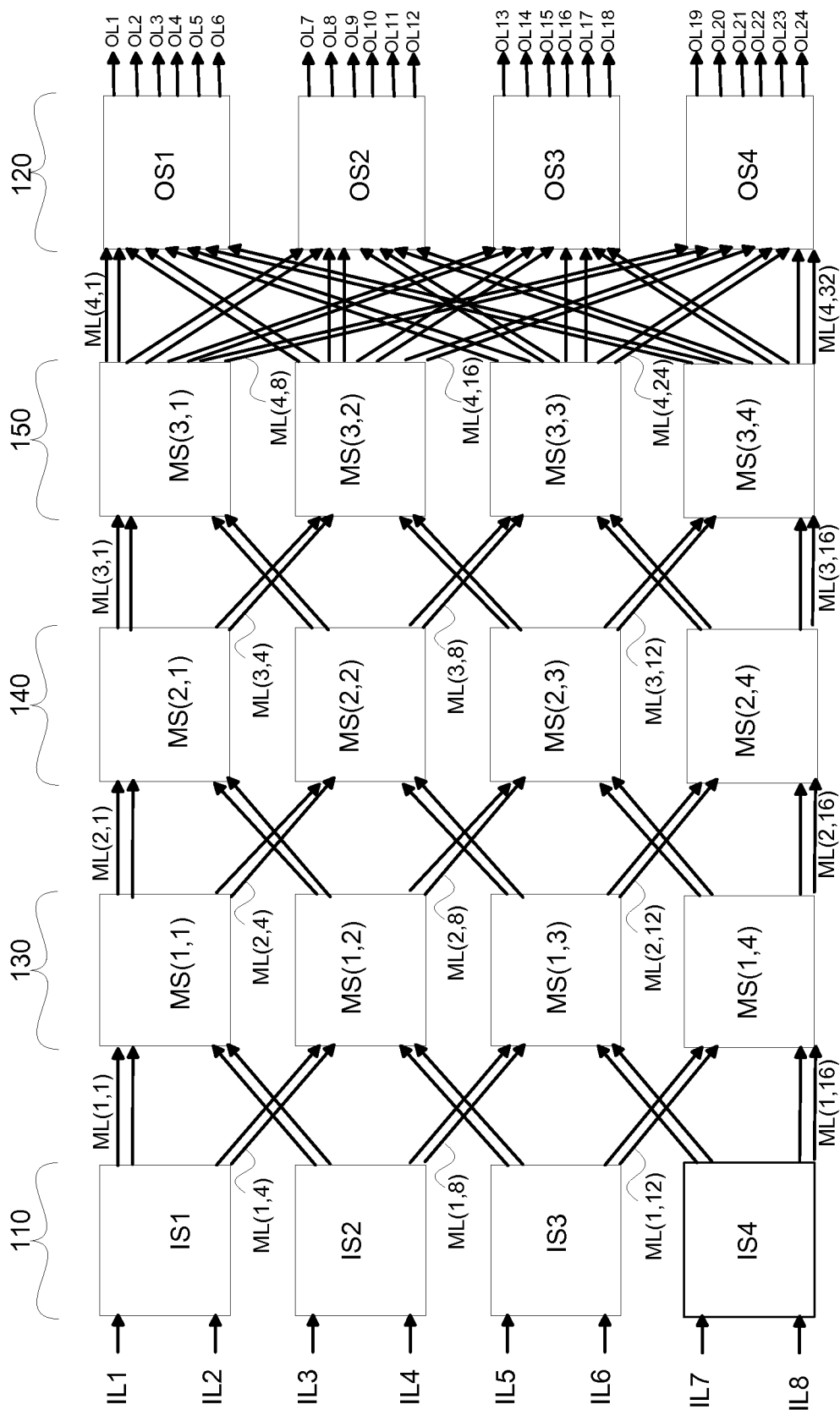


FIG. 1H1

100H1

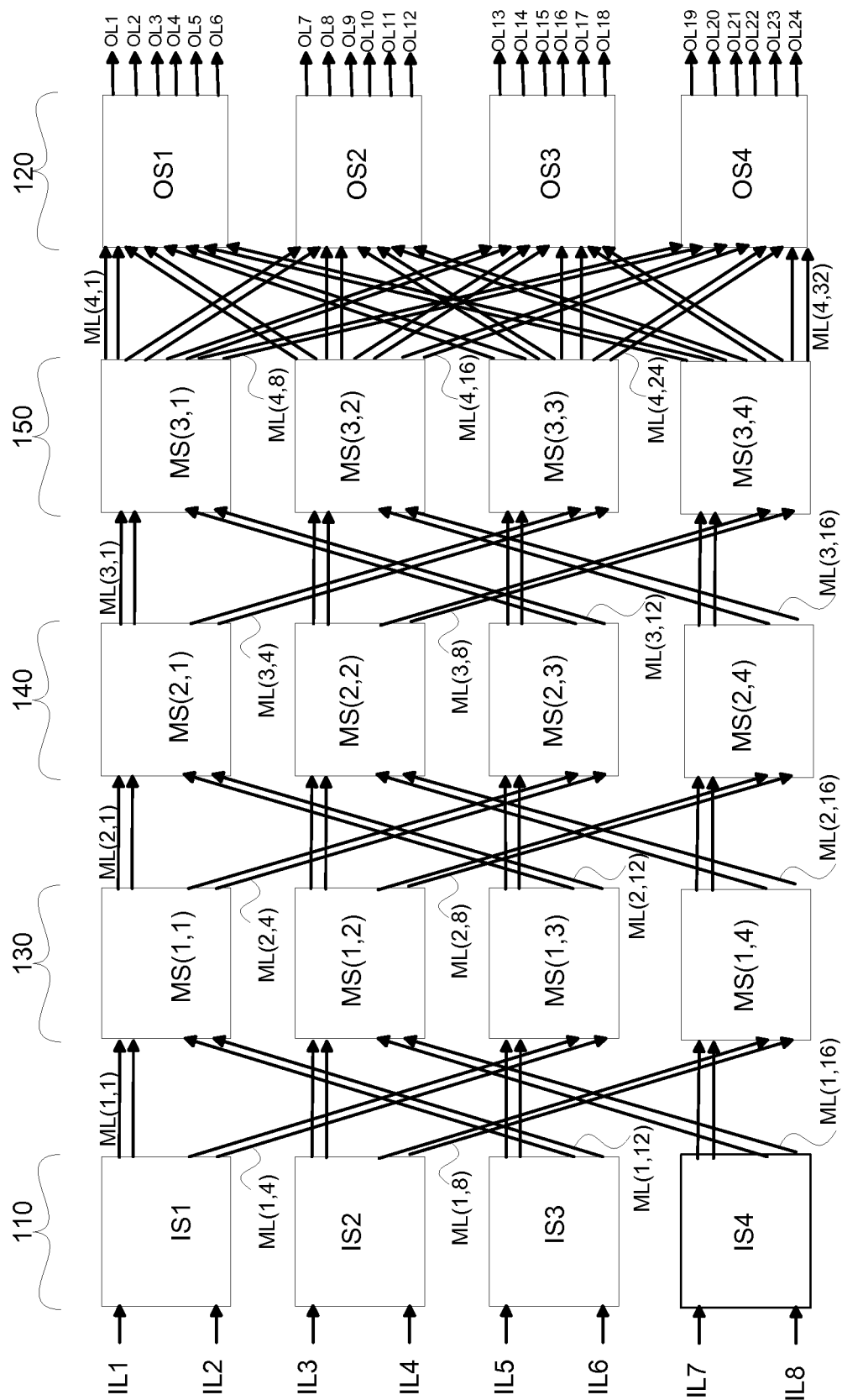


FIG. 111

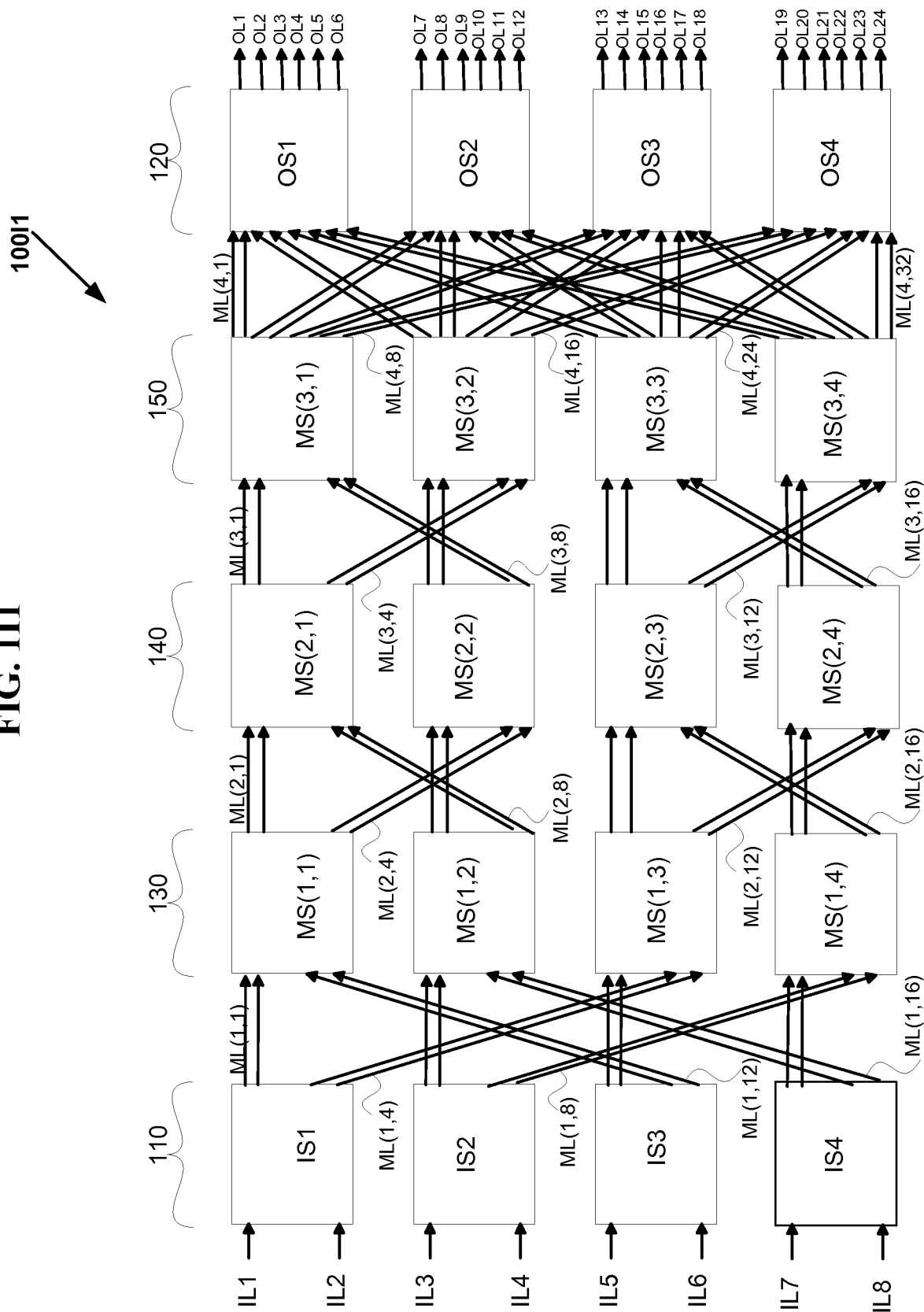


FIG. 1J1

100J1

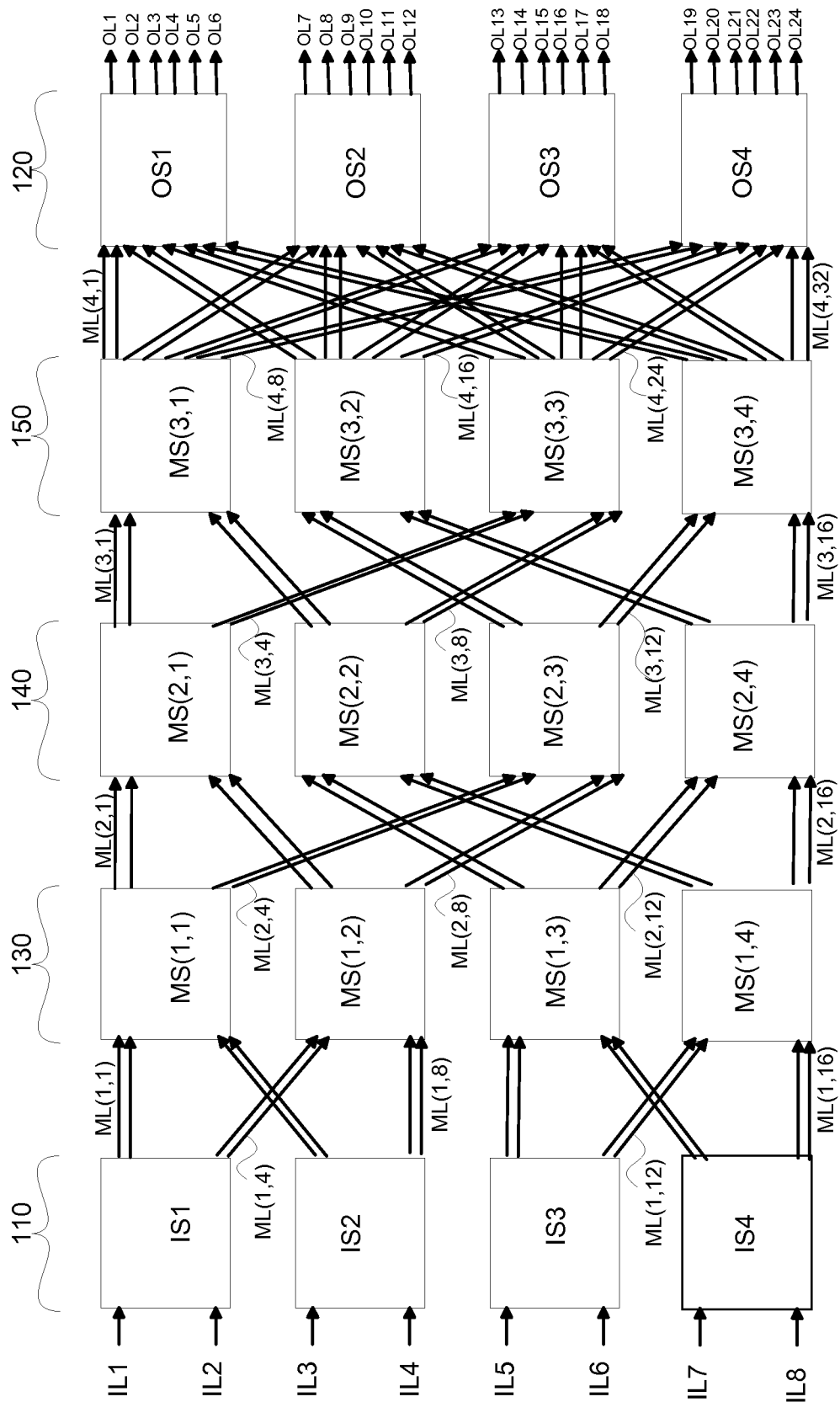


FIG. 1K1

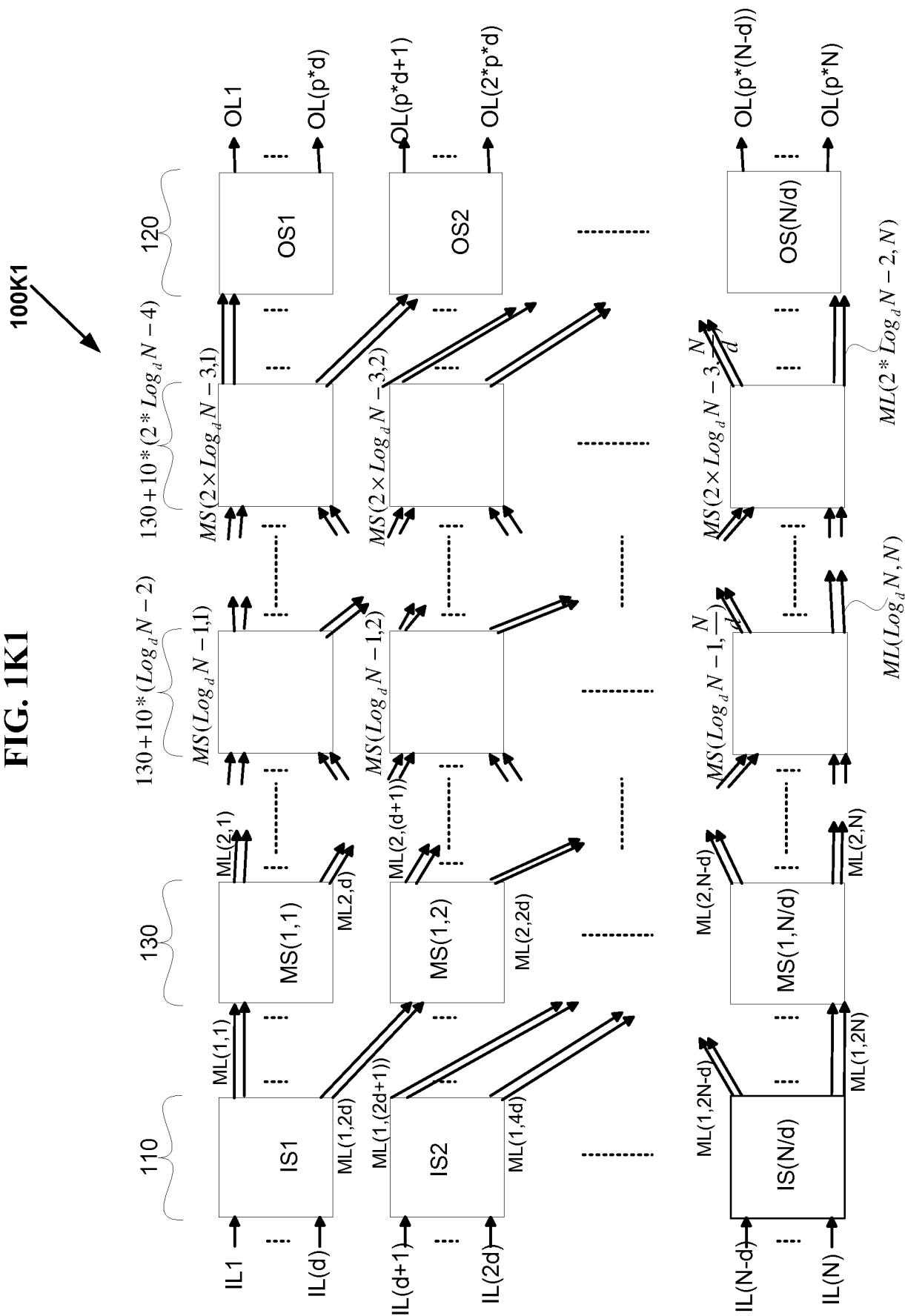


FIG. 1A2

100A2

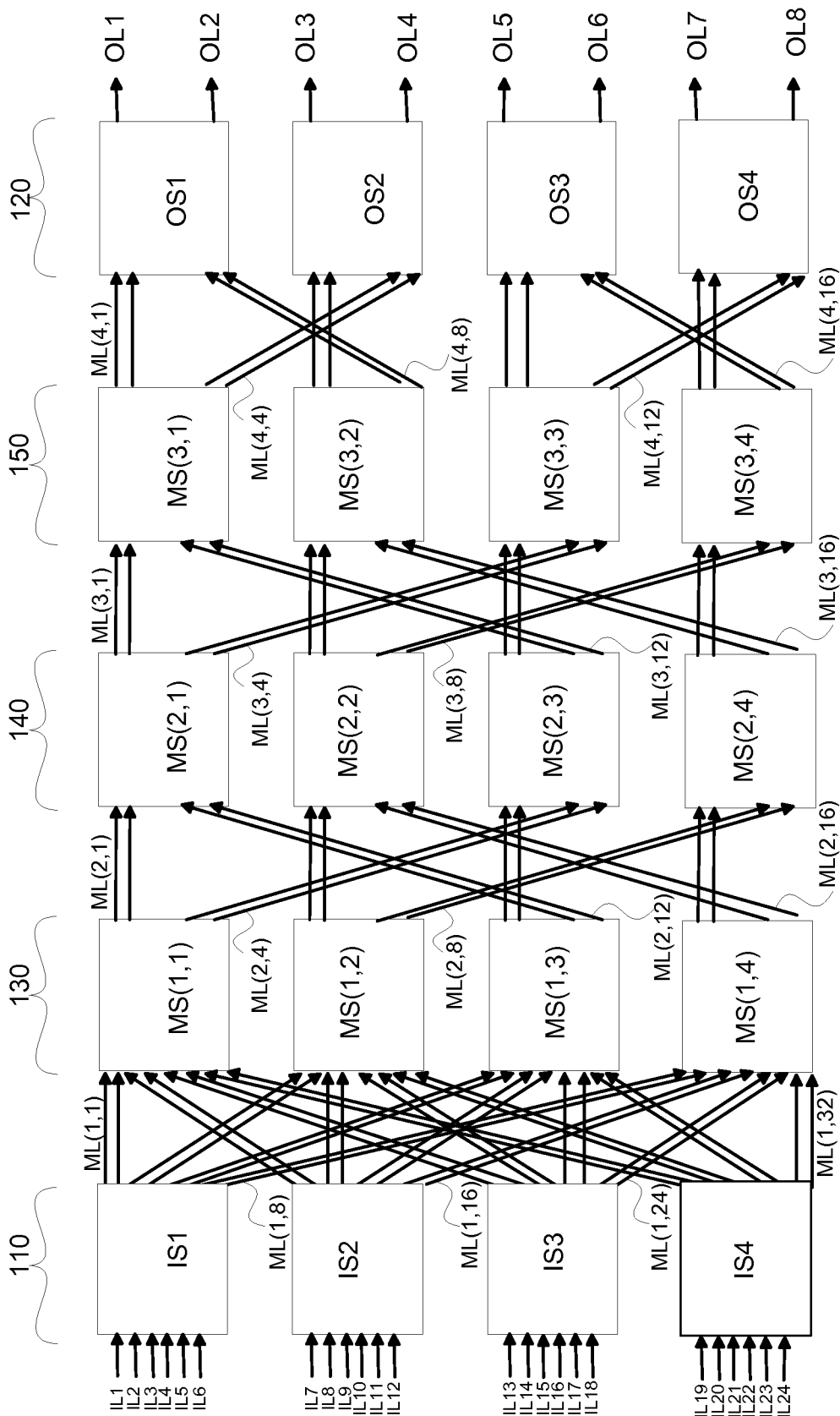


FIG. 1B2

100B2

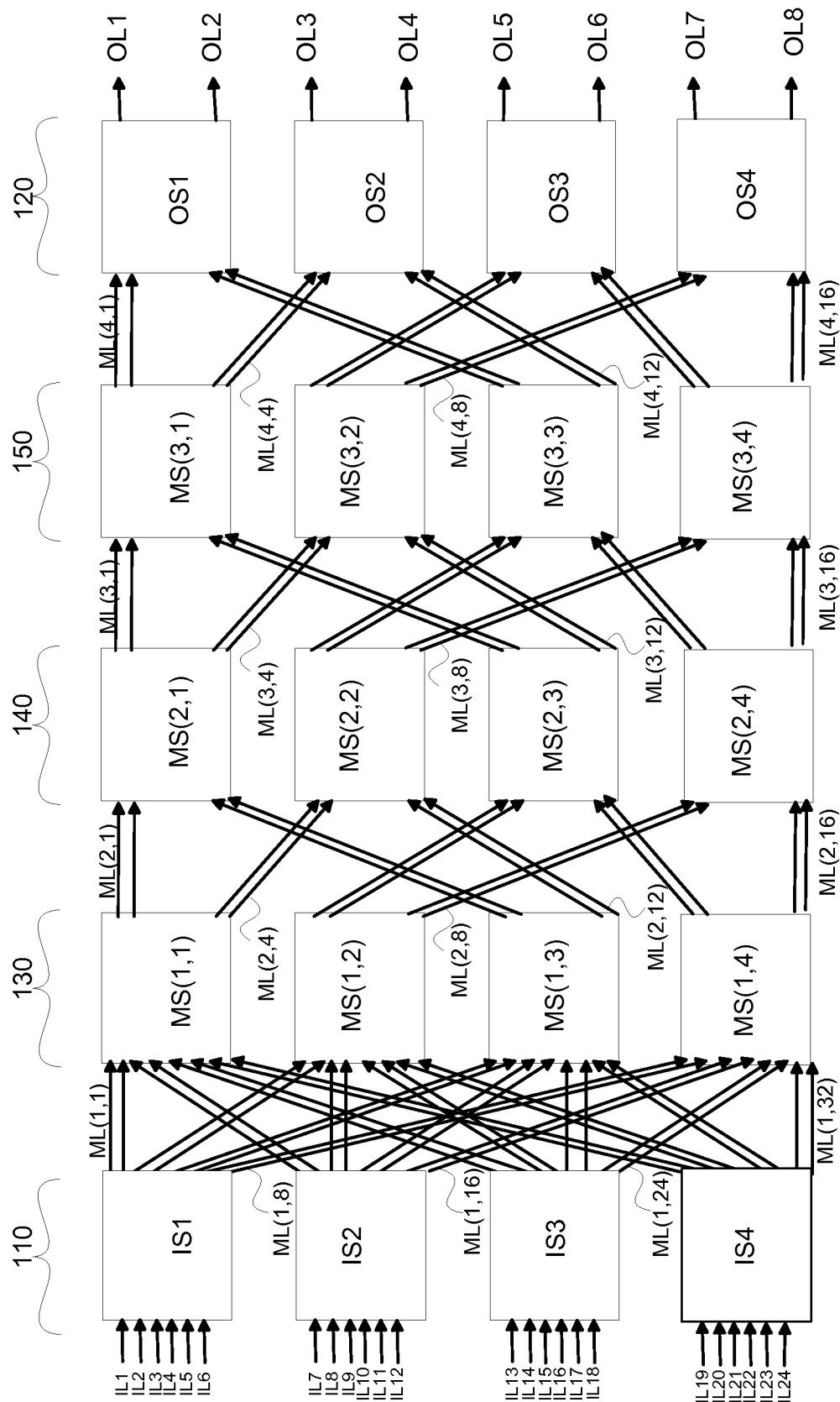


FIG. 1C2

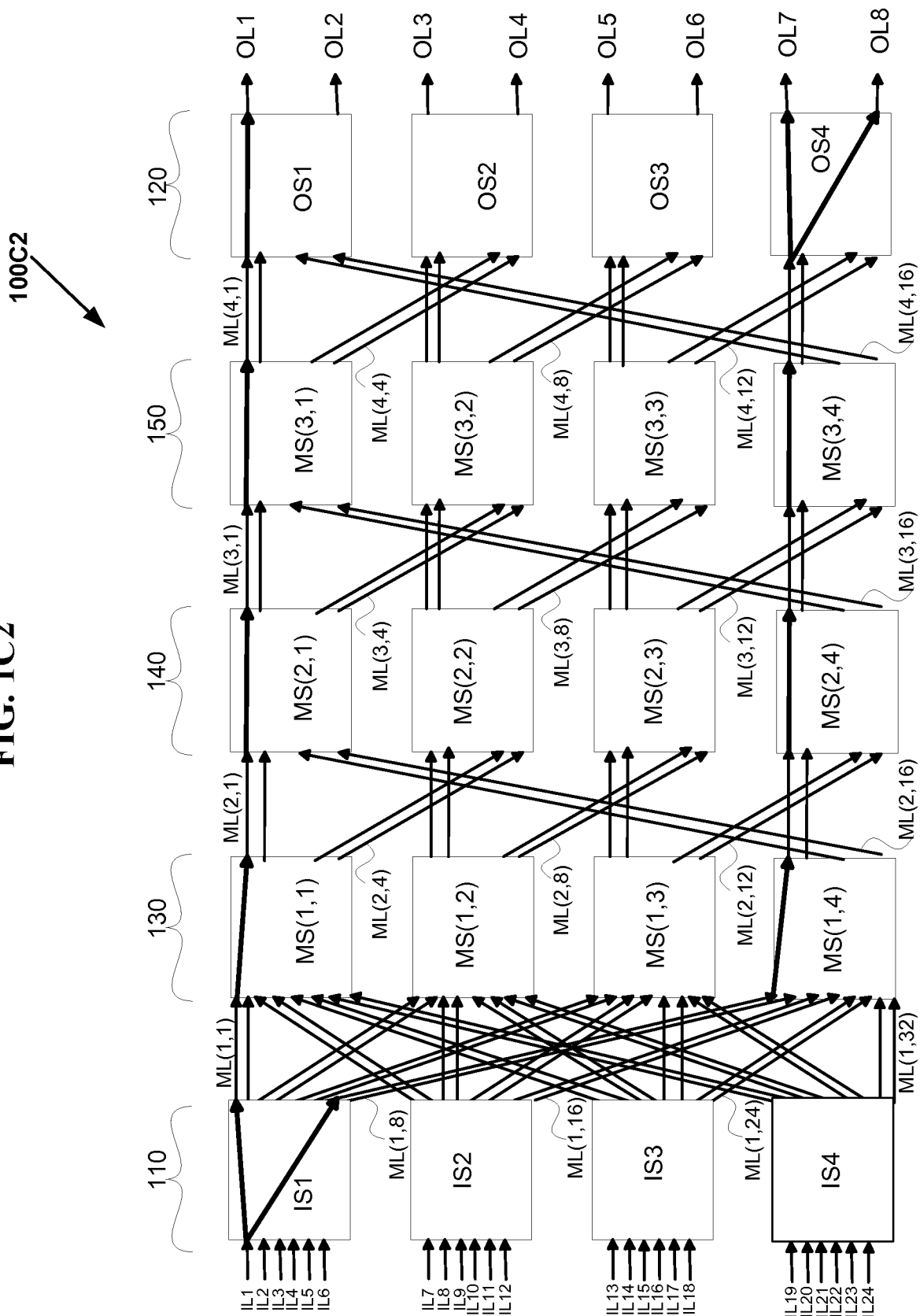


FIG. 1D2

100D2

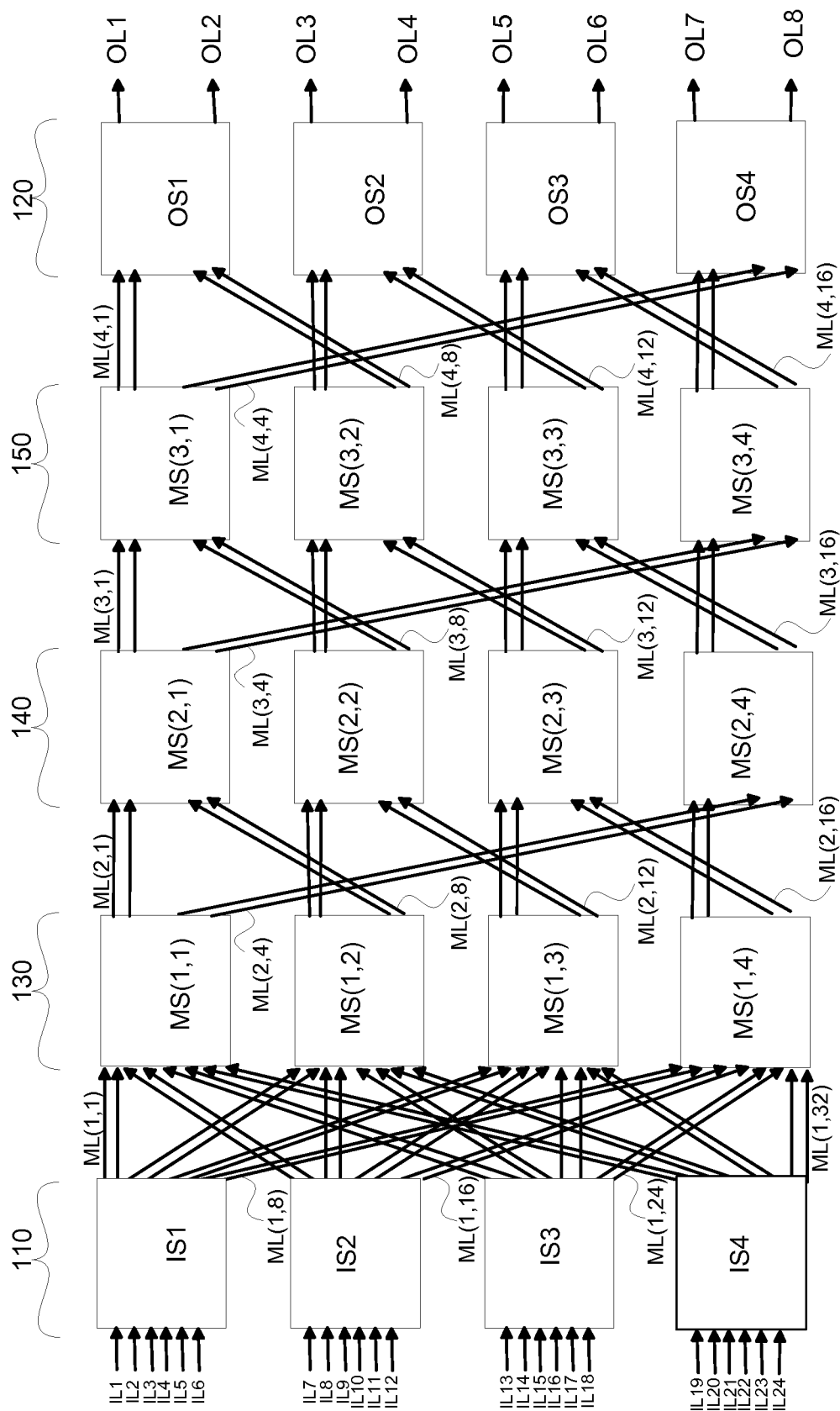


FIG. 1E2

100E2

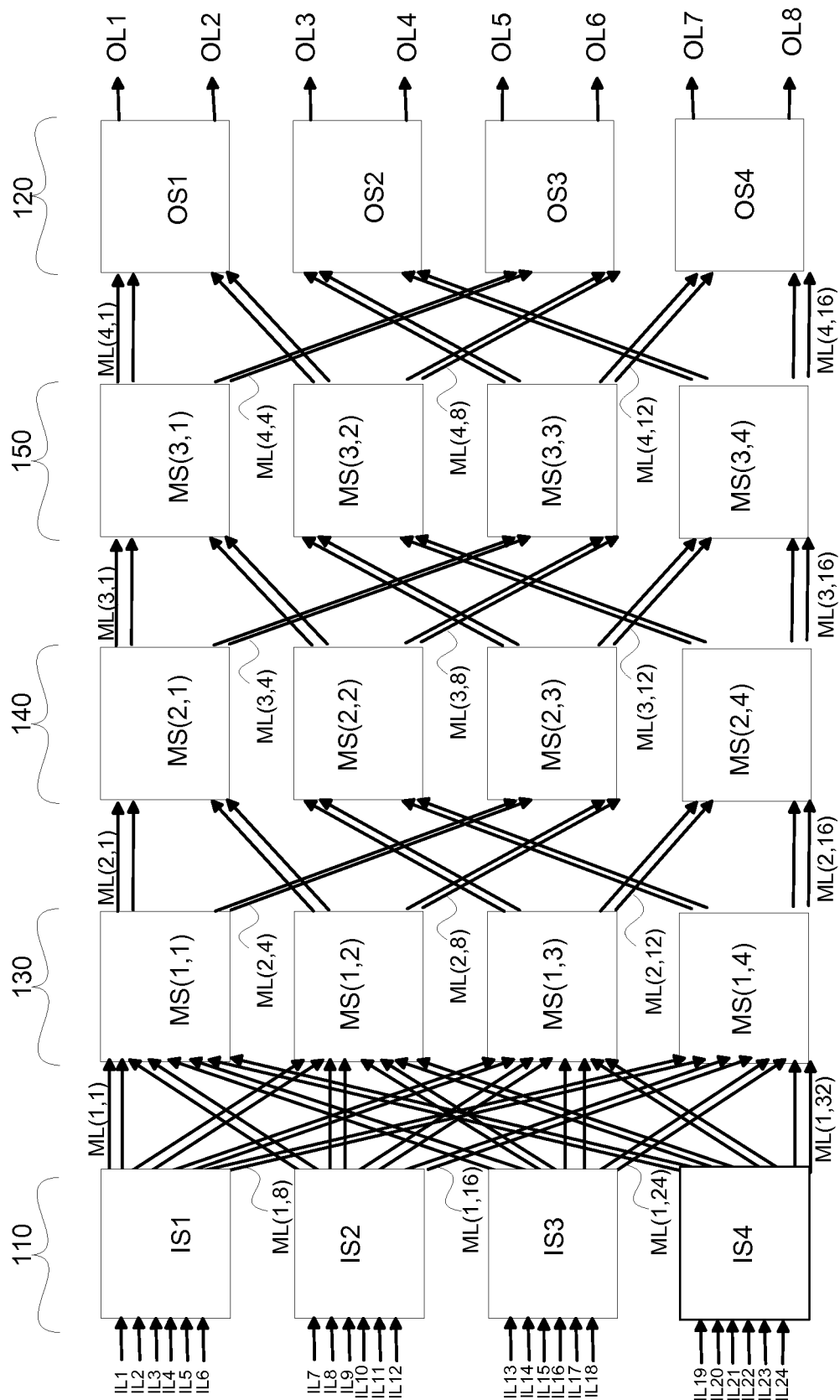


FIG. 1F2

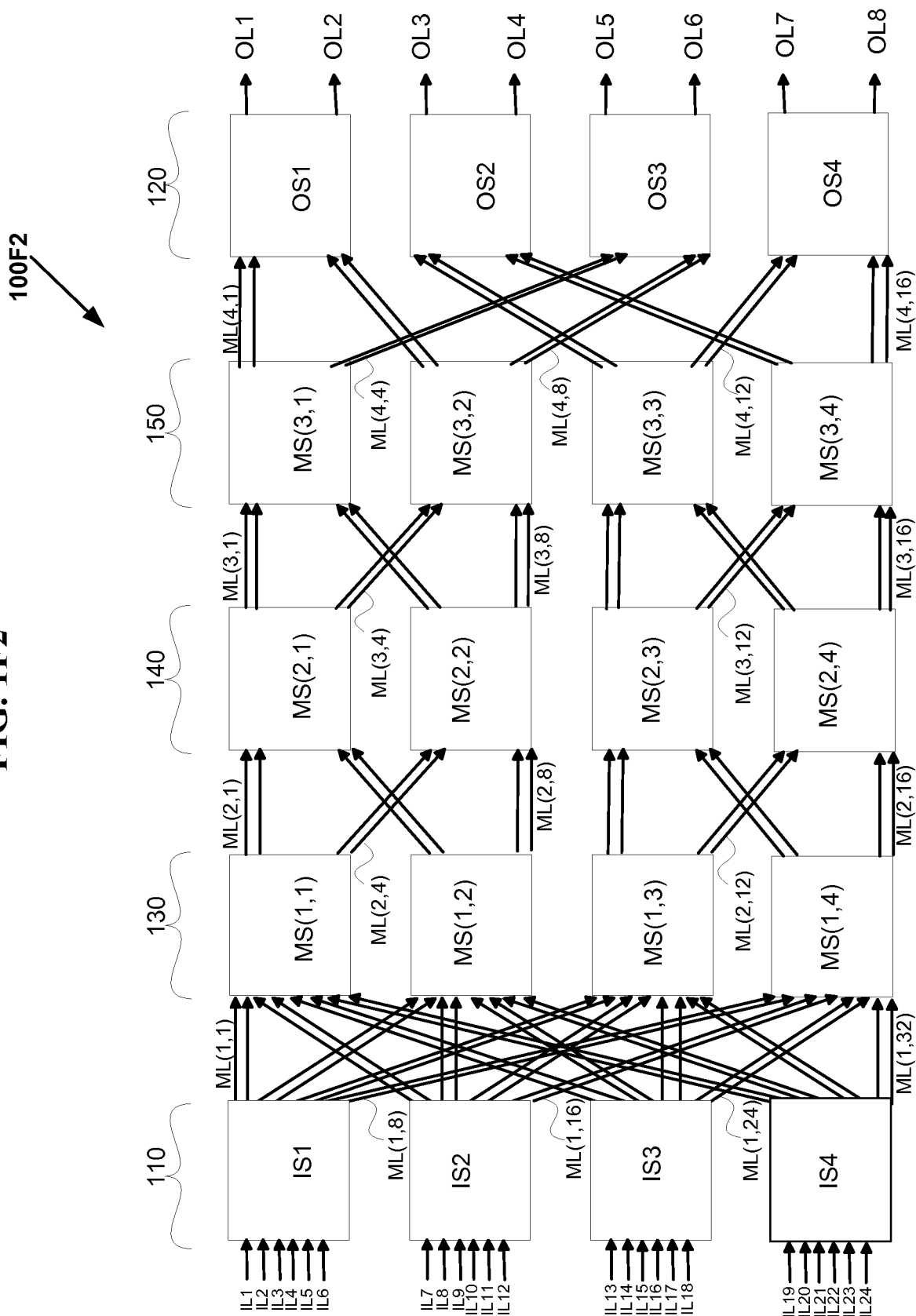


FIG. 1G2

100G2

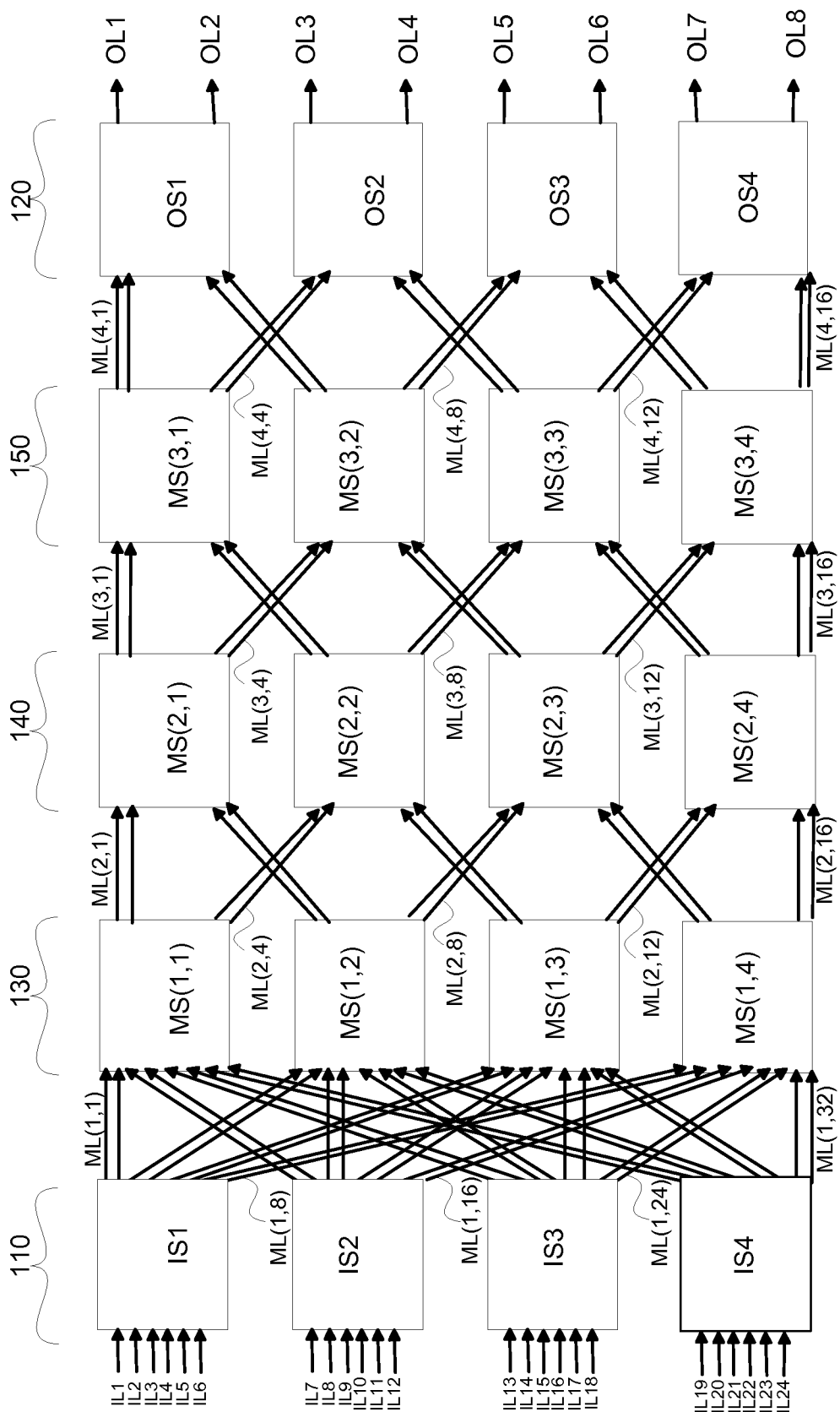


FIG. 1H2

100H2

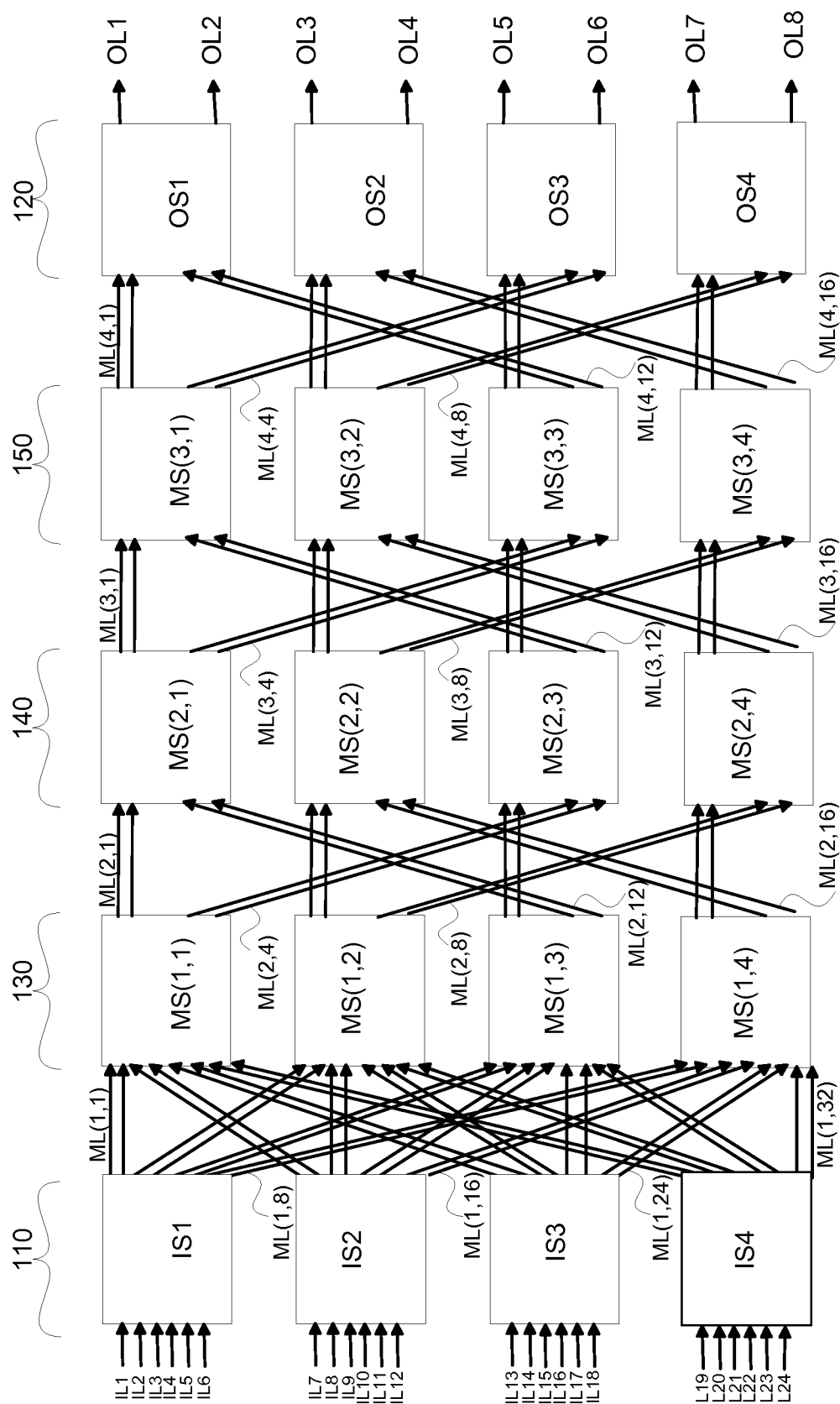


FIG. 112

10012

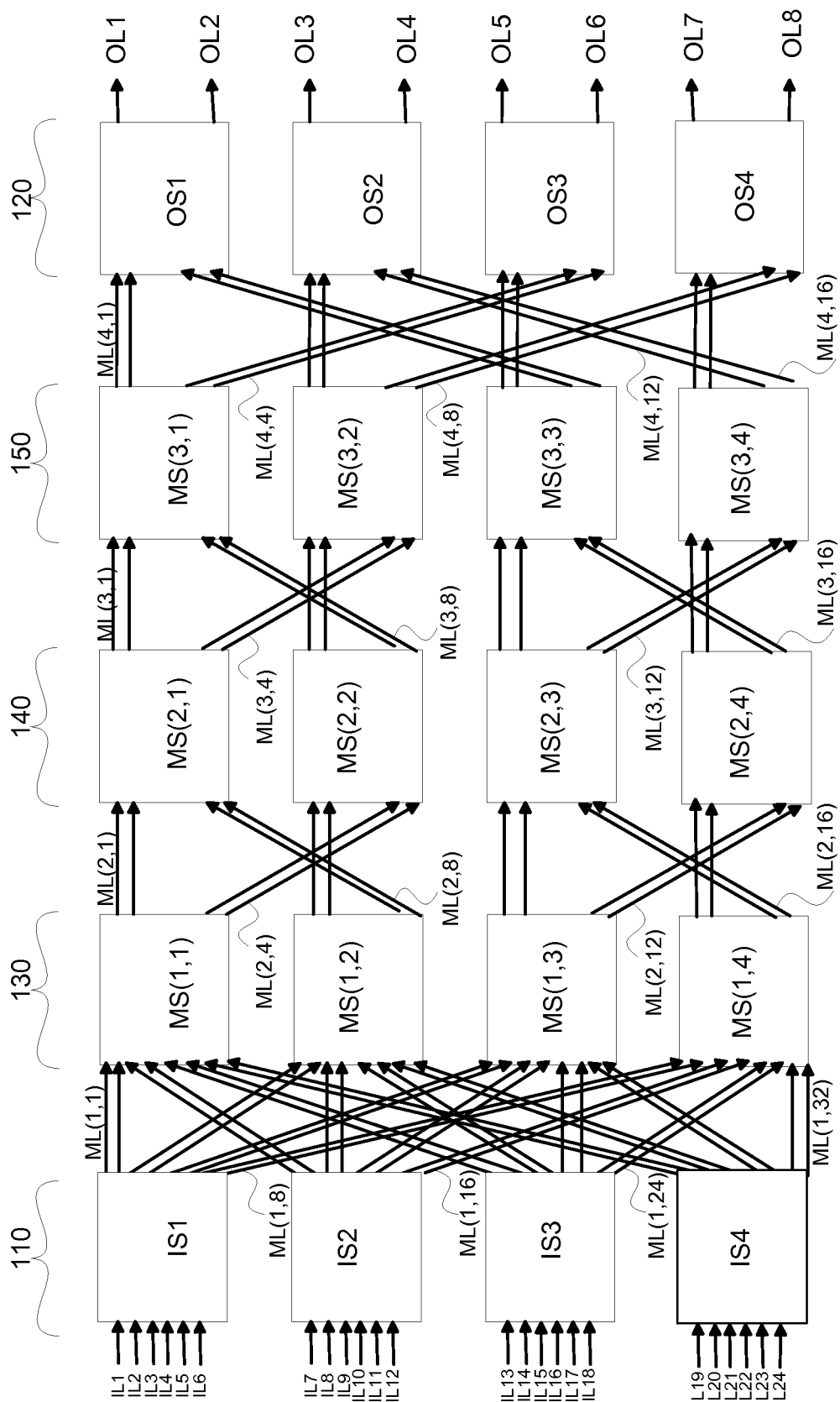


FIG. 1J2

100J2

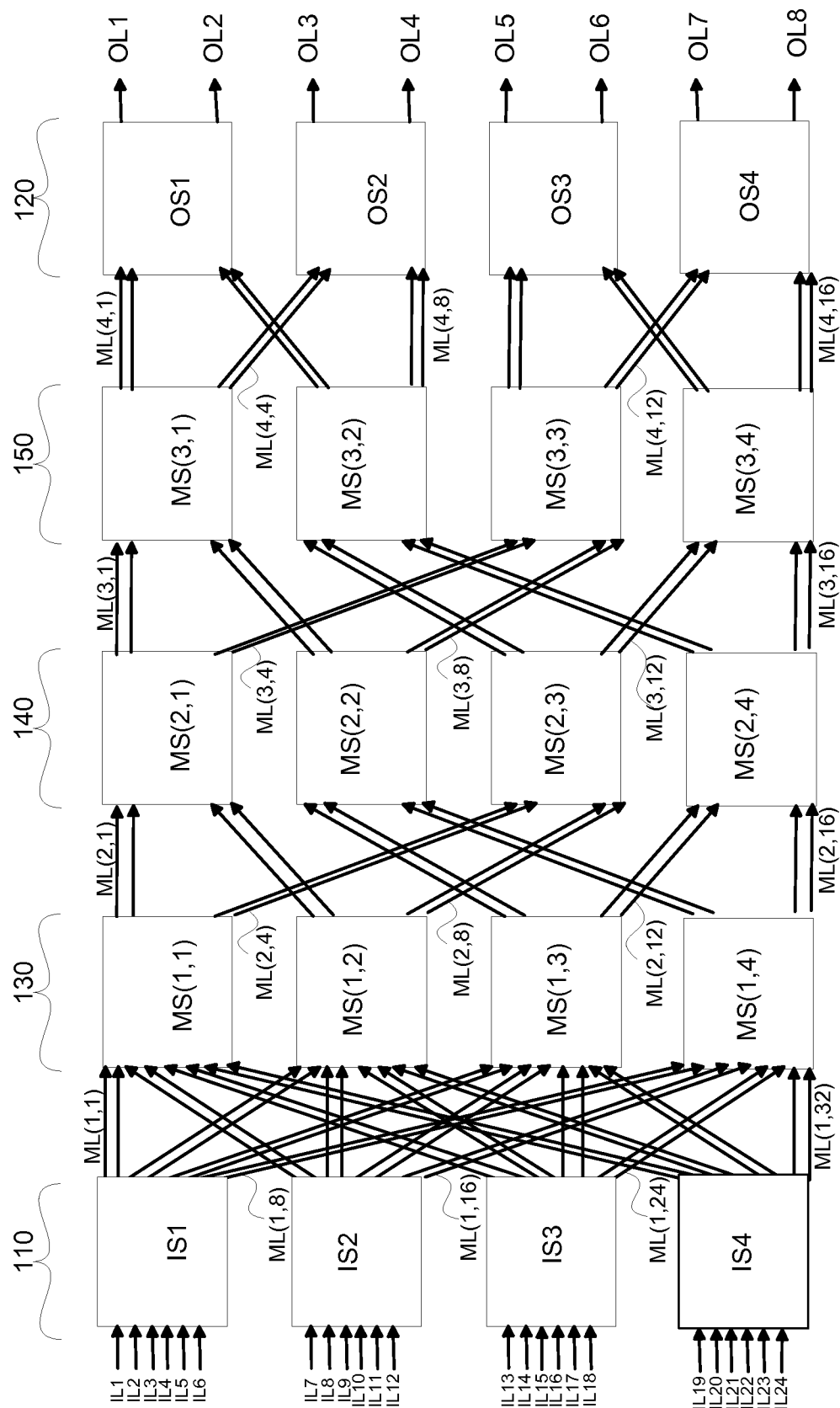
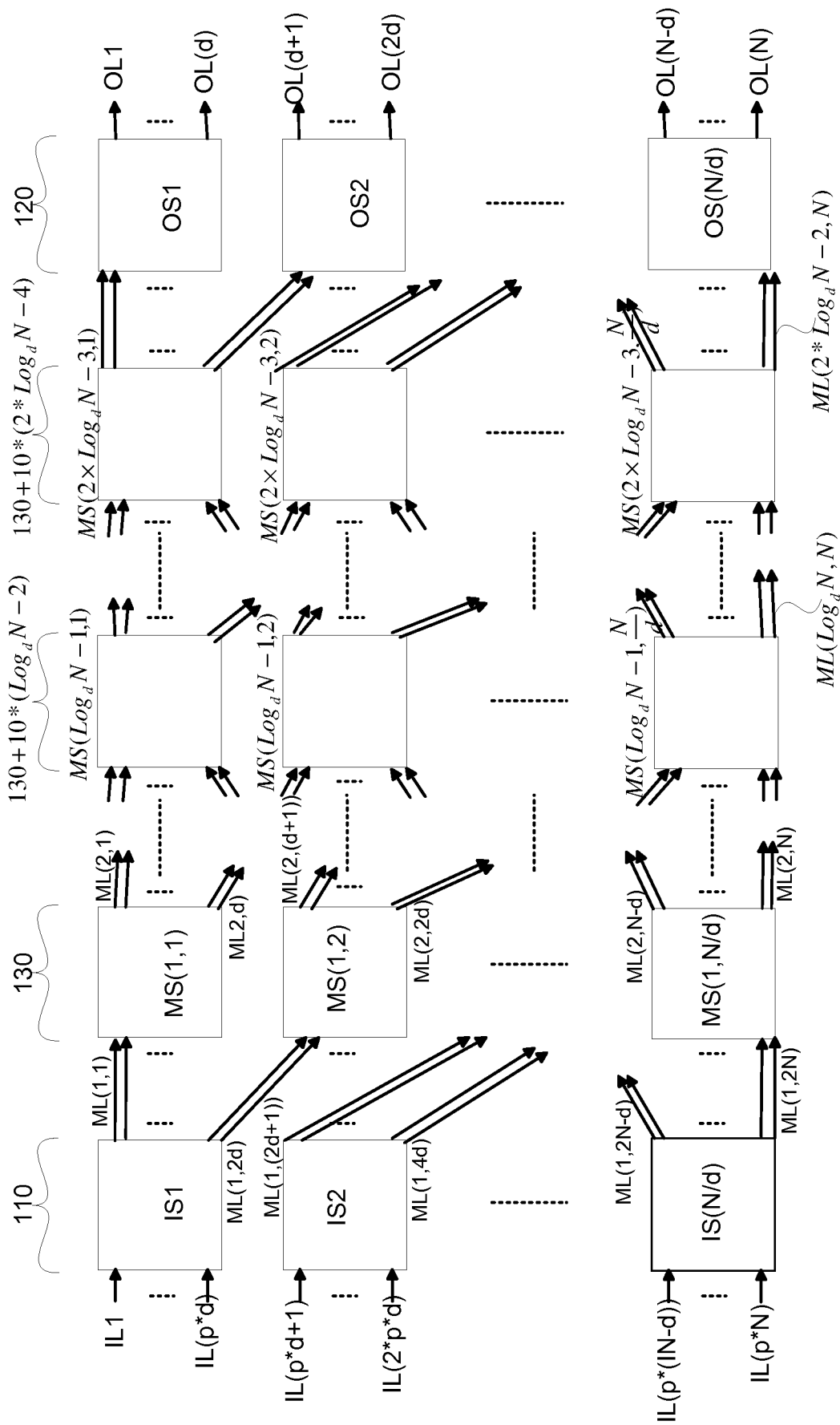


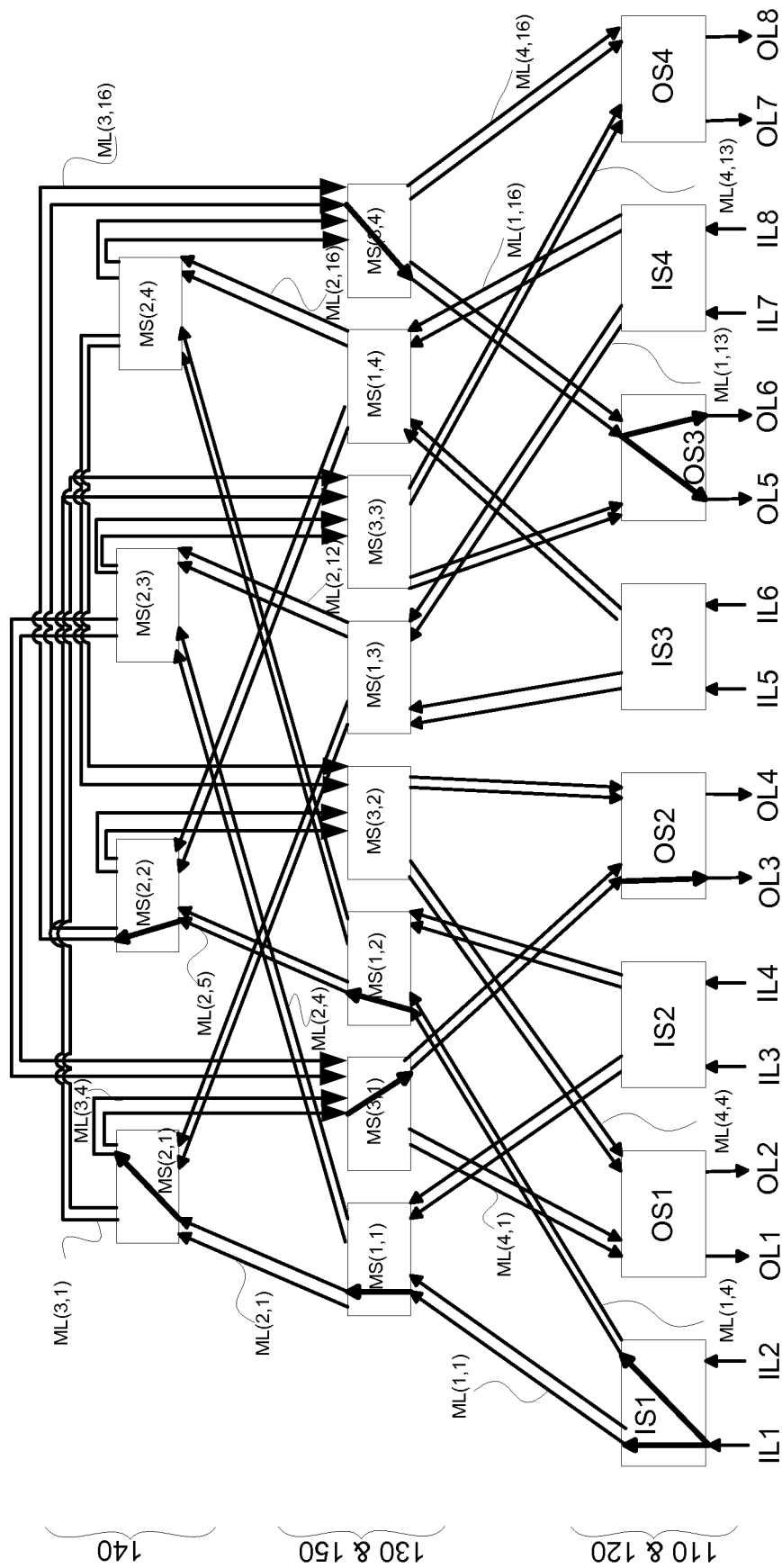
FIG. 1K2

100K2



200A

FIG. 2A



140

130 & 150

110 & 120

FIG. 2C

200C

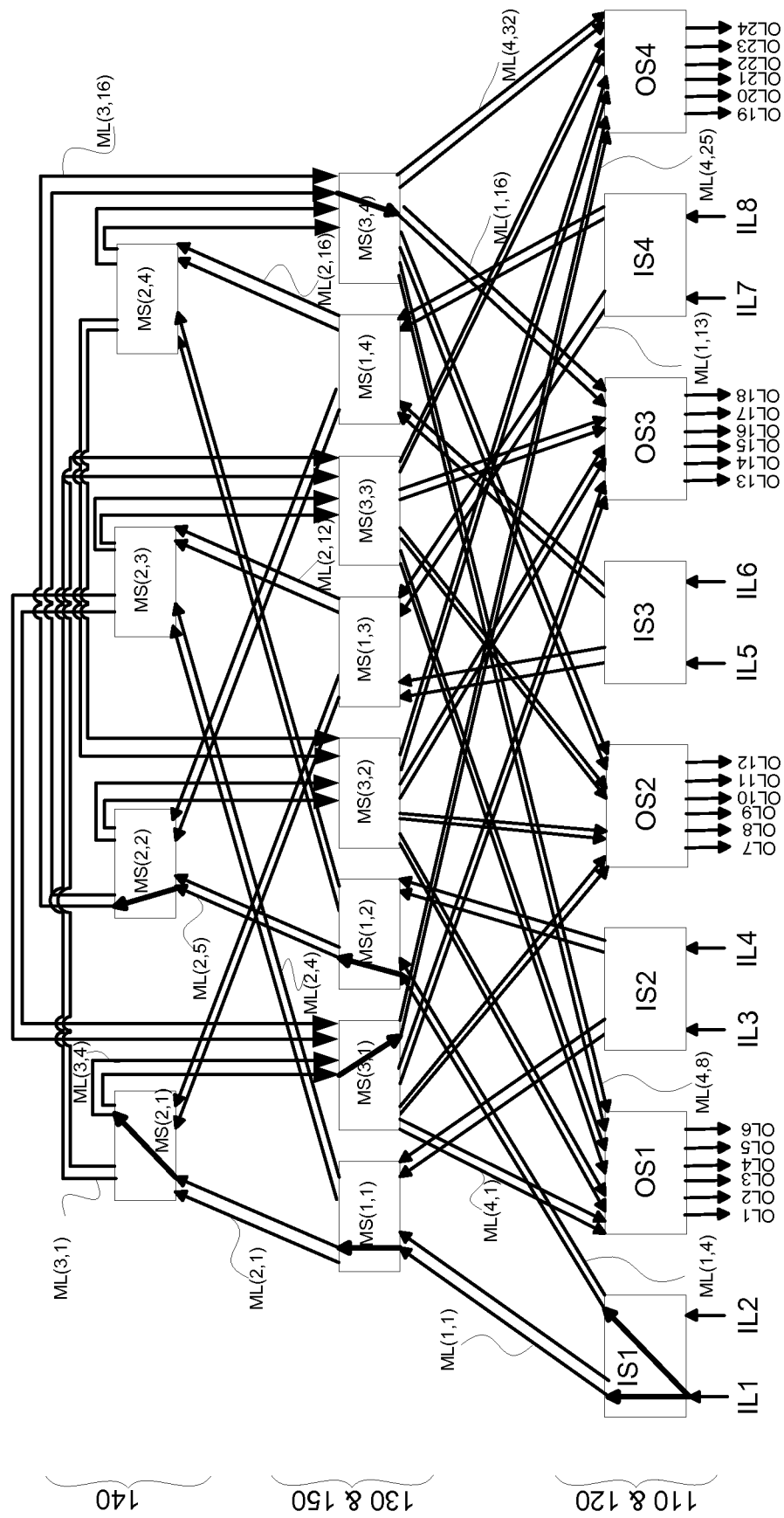


FIG. 2D

200D

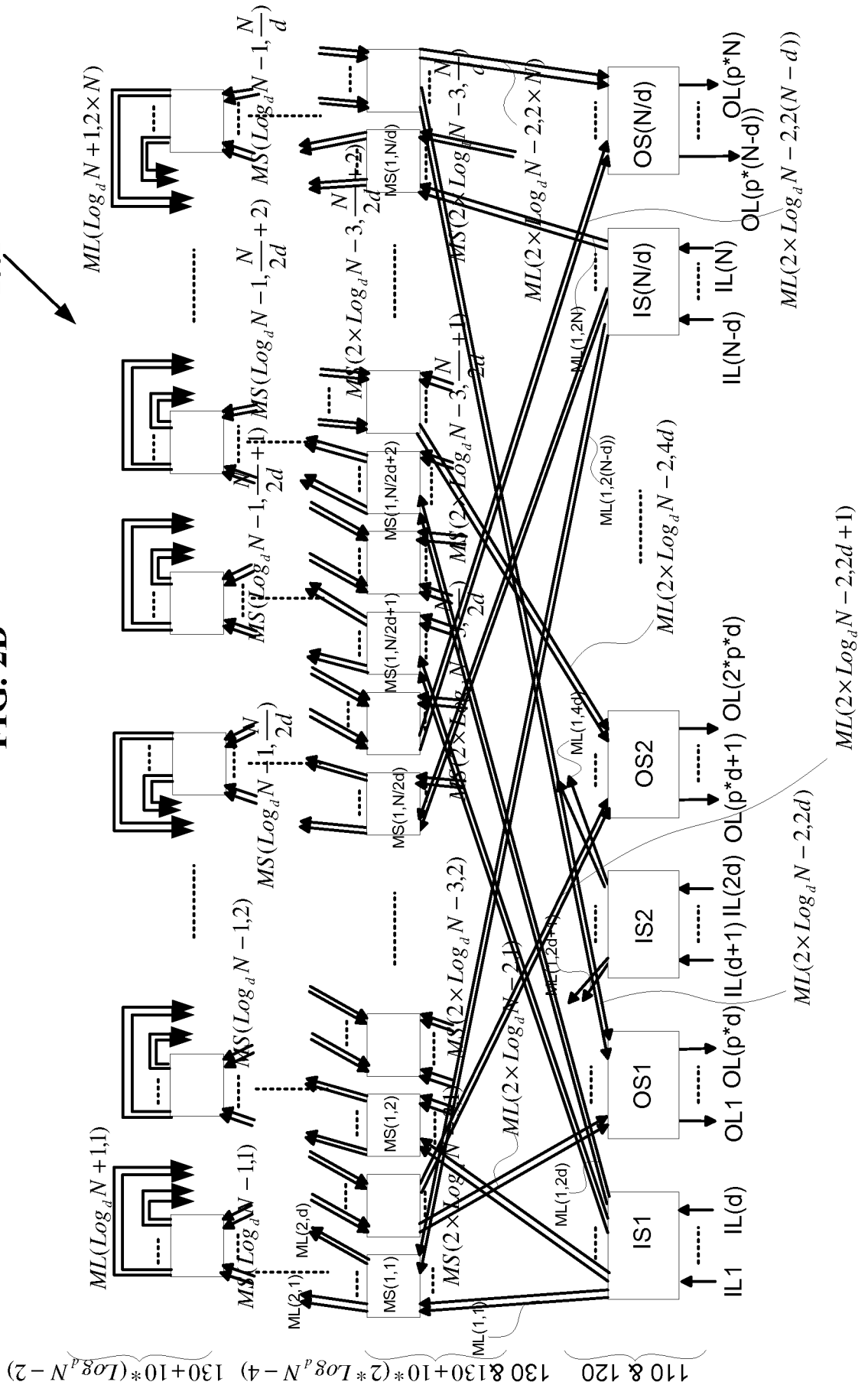


FIG. 2E

200E

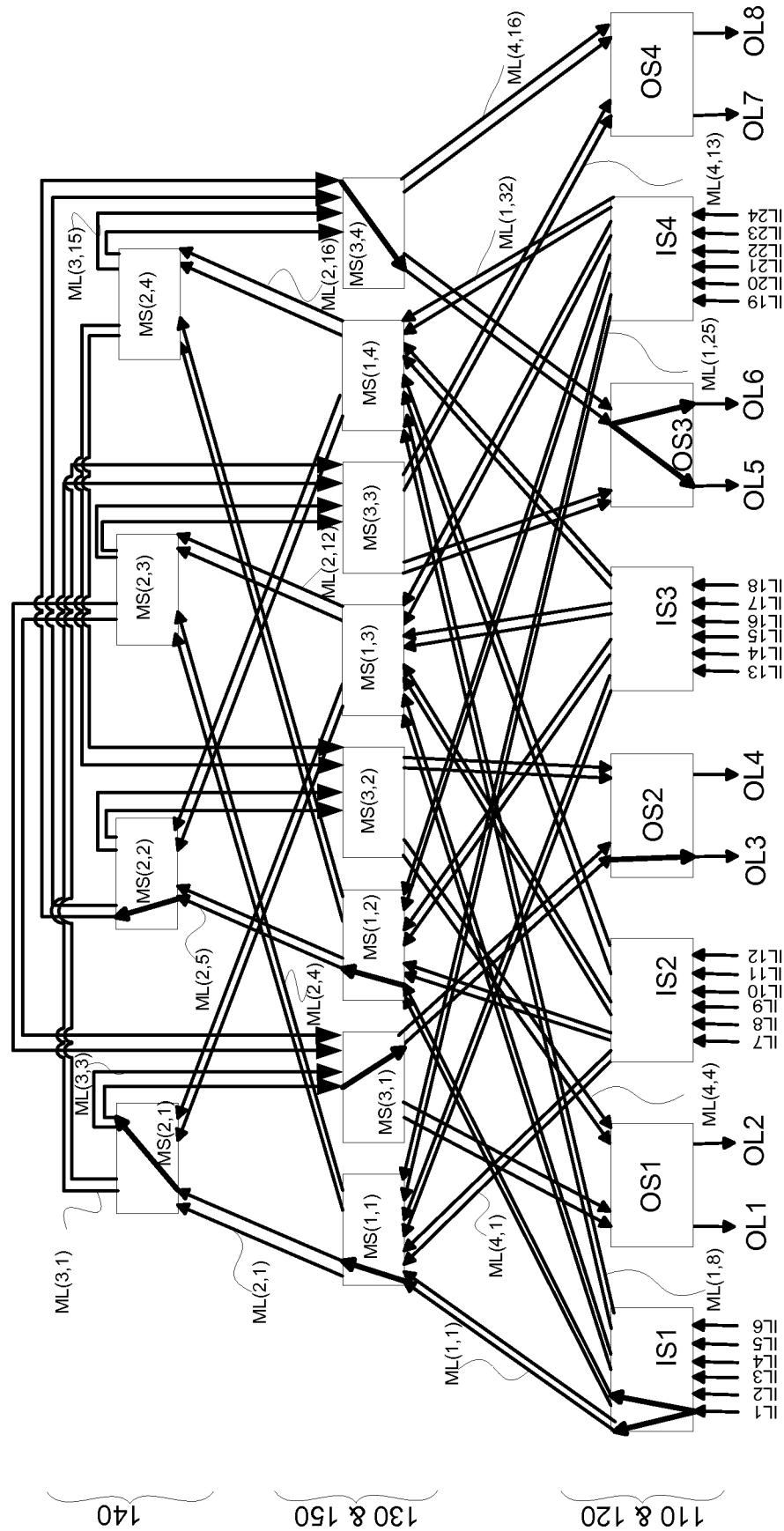


FIG. 2F

200F

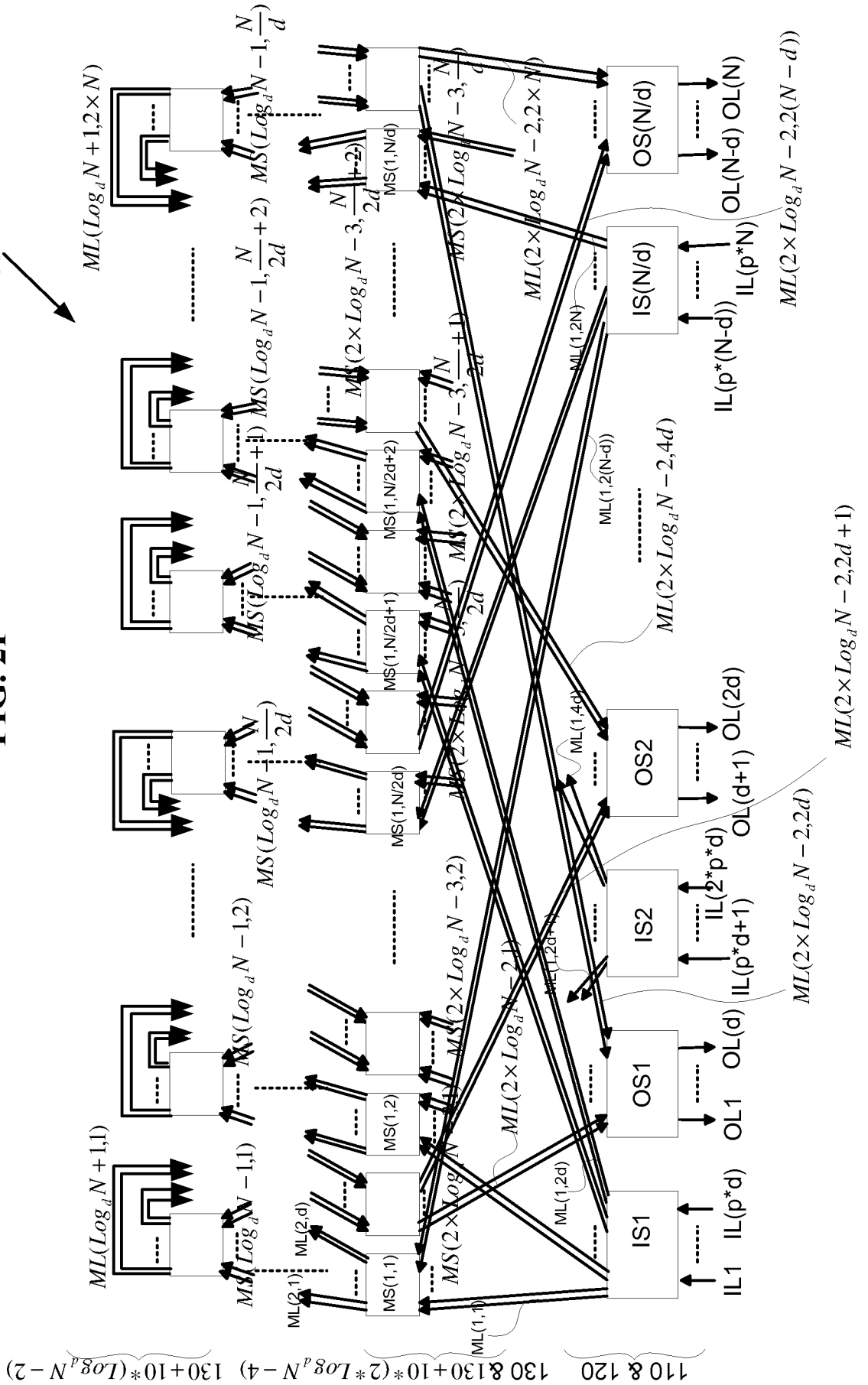


FIG. 3A

300A

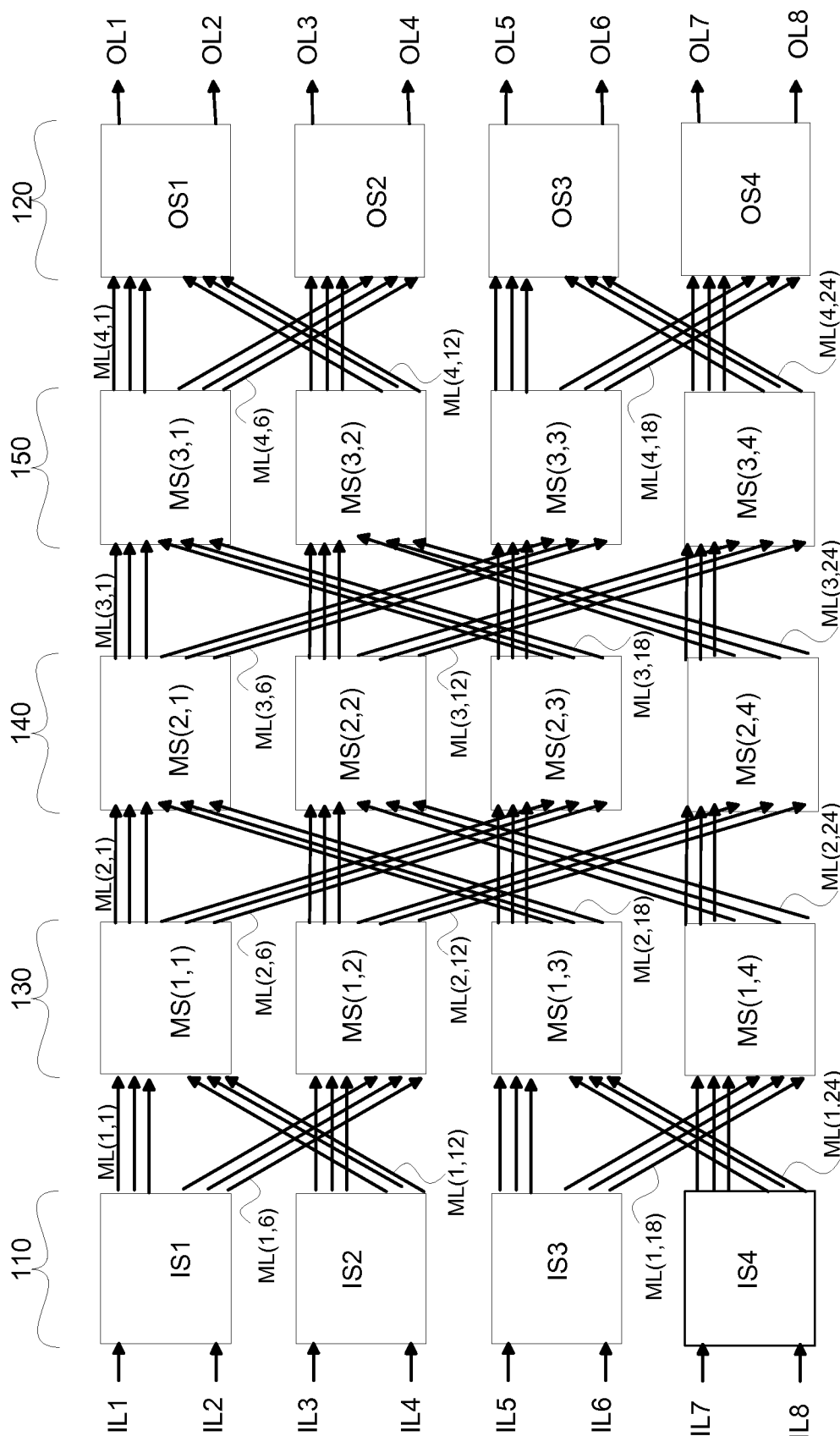


FIG. 3B

300B

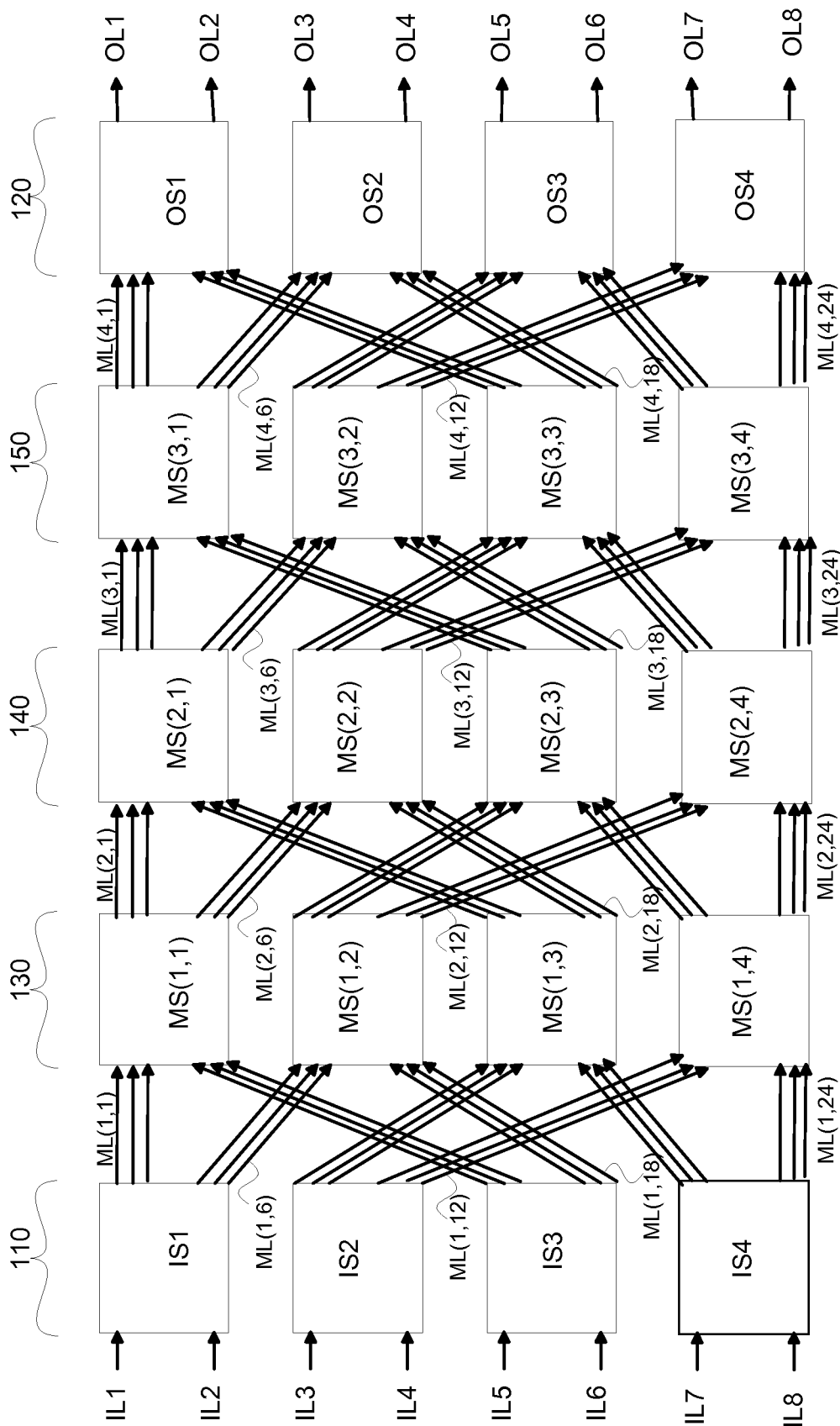


FIG. 3C

300C

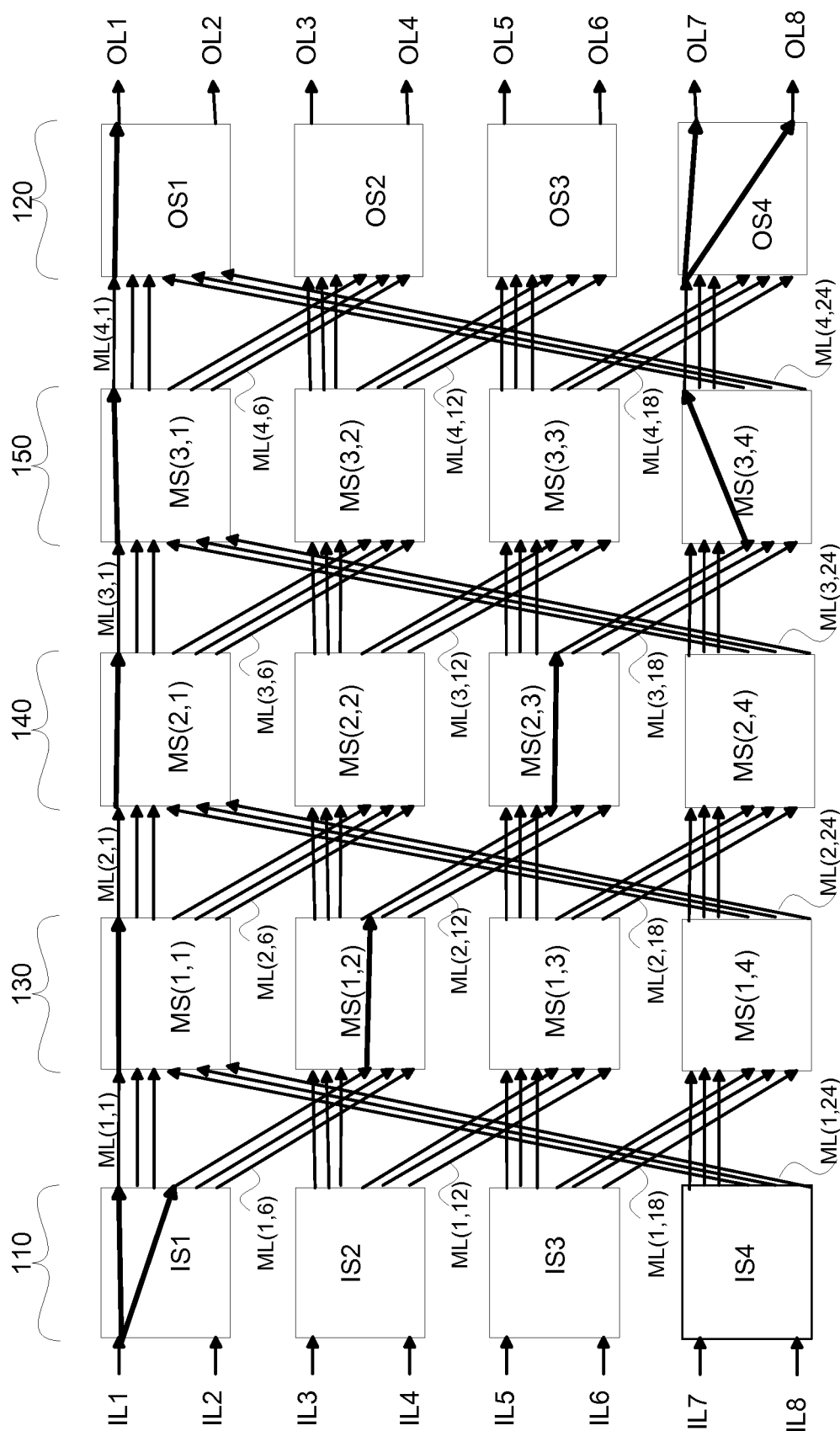


FIG. 3D

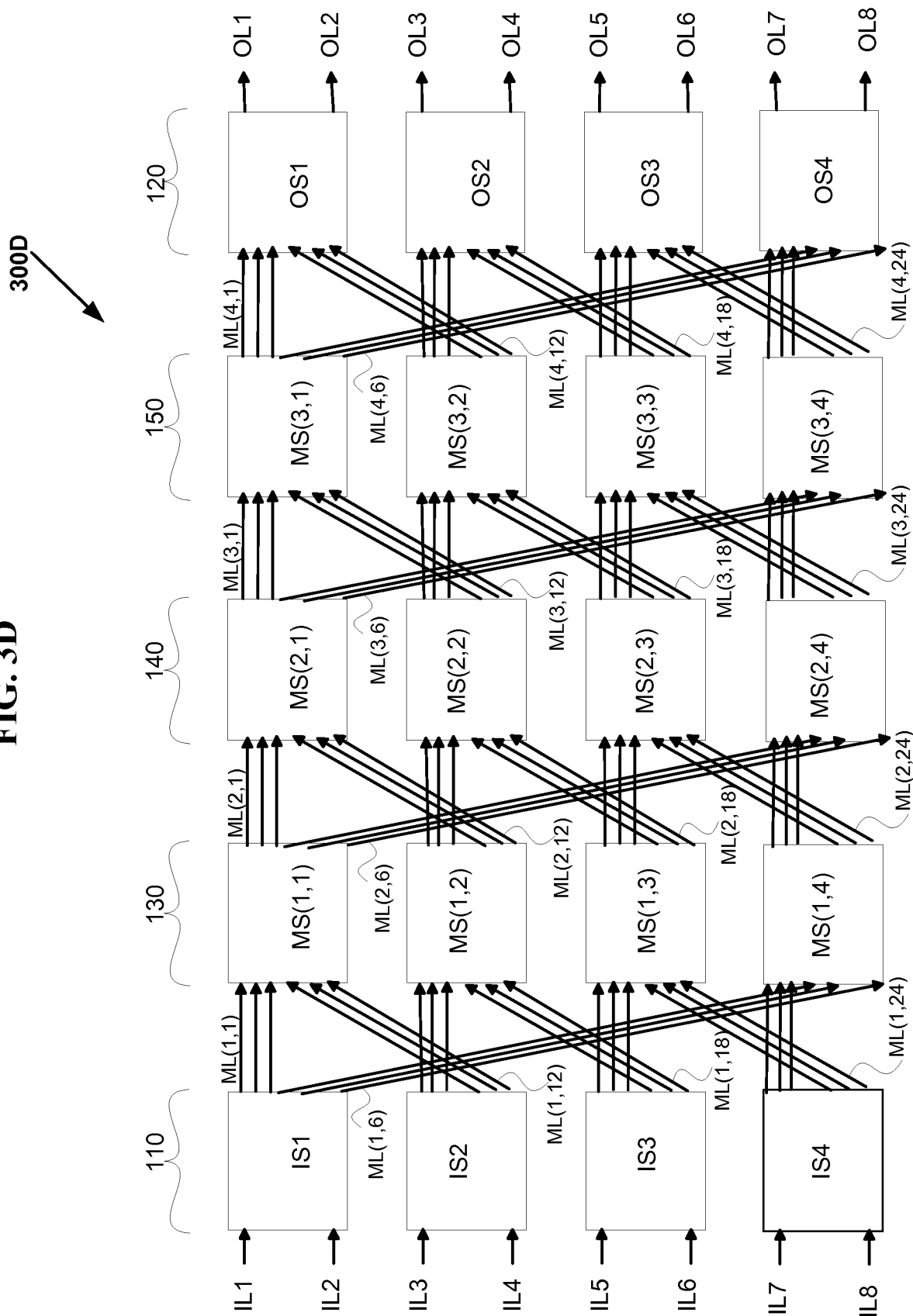


FIG. 3E

300E

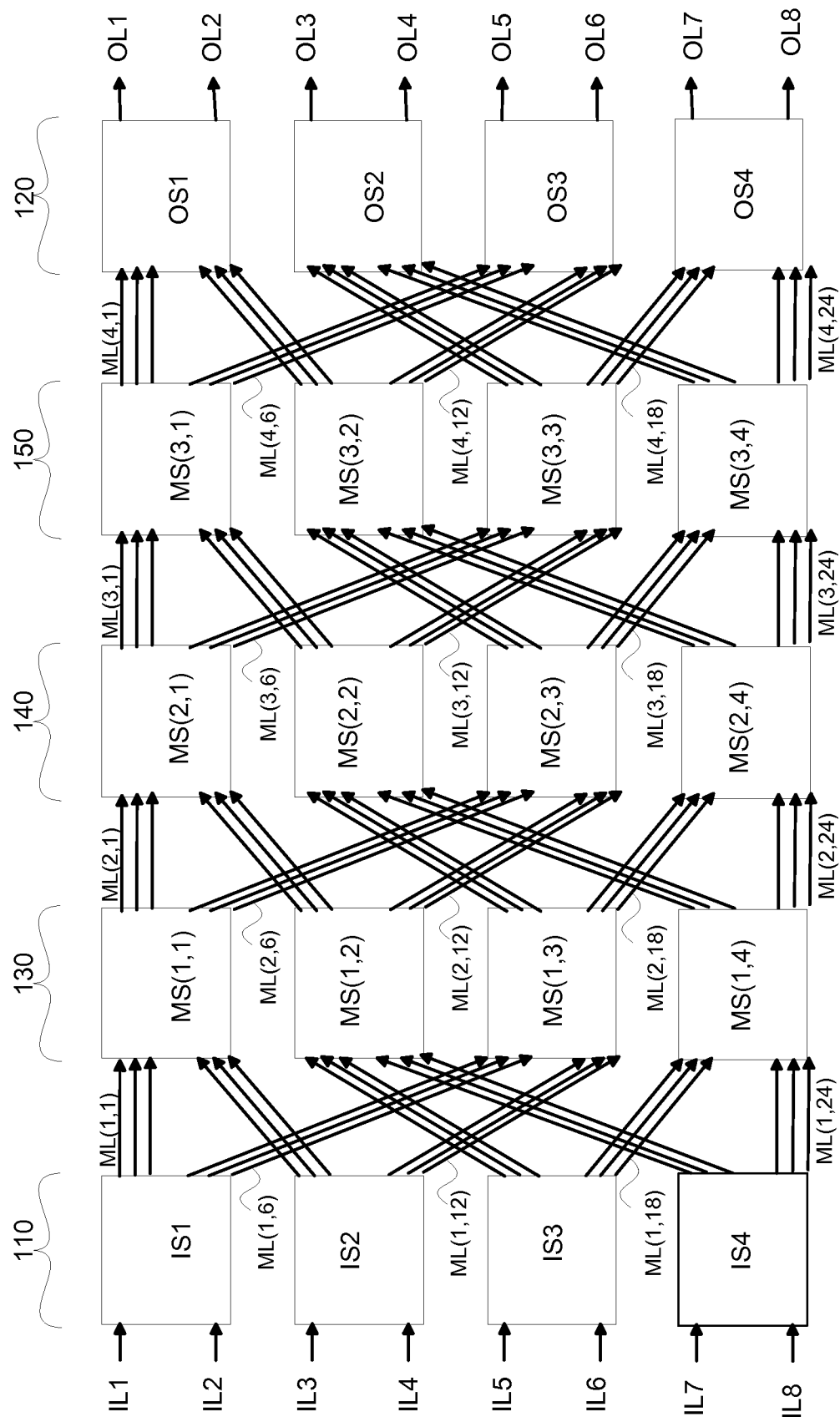


FIG. 3F

300F

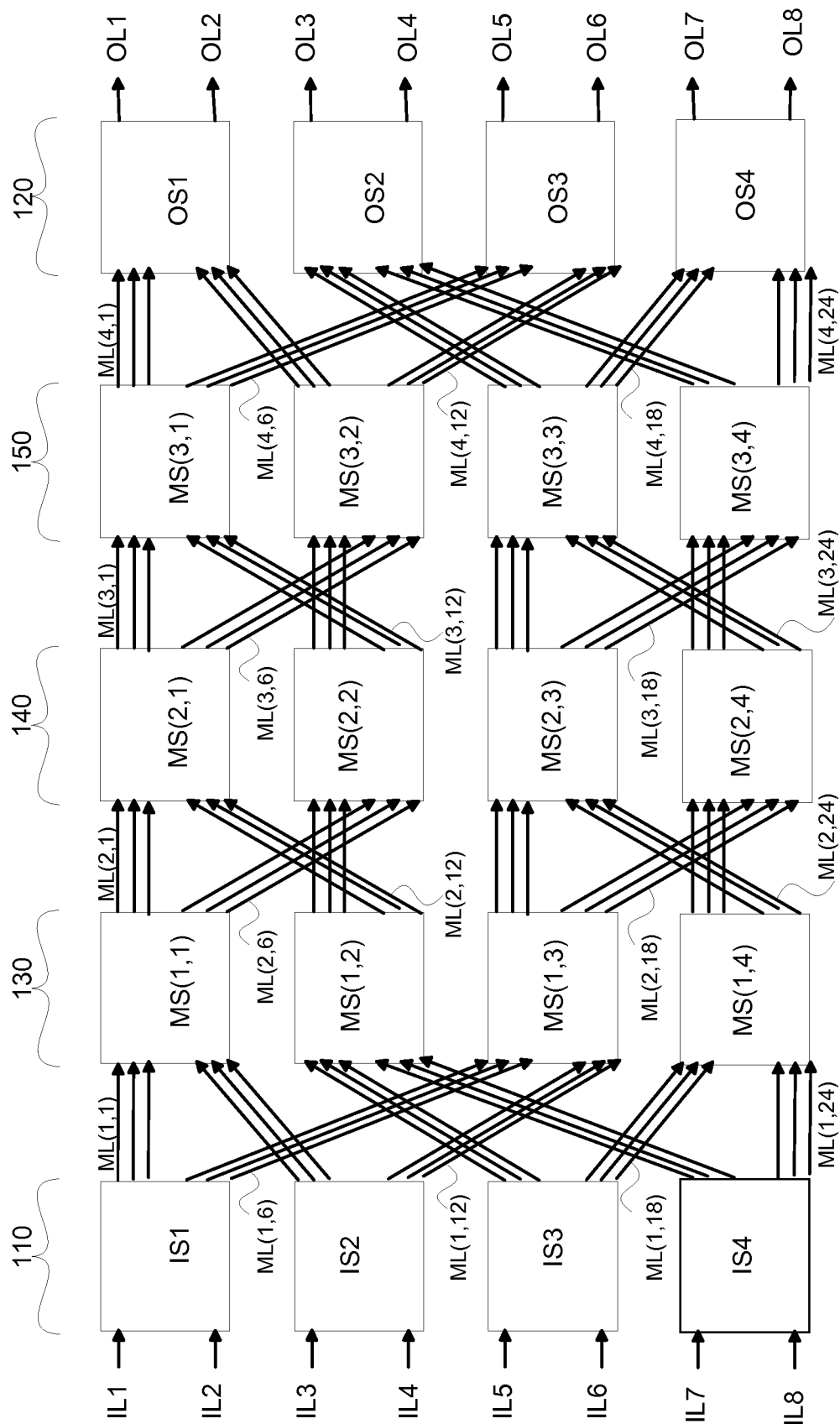


FIG. 3G

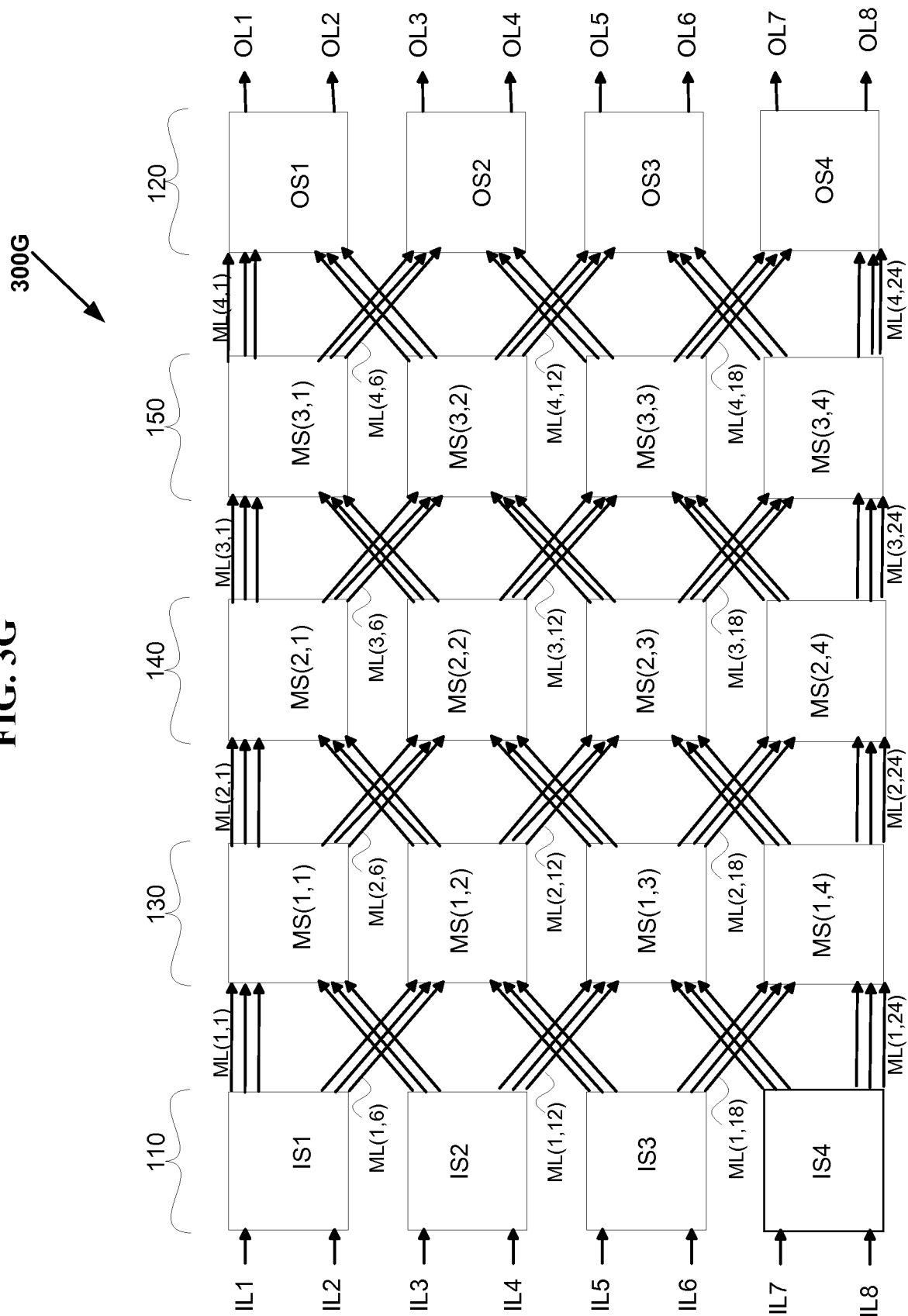


FIG. 3H

300H

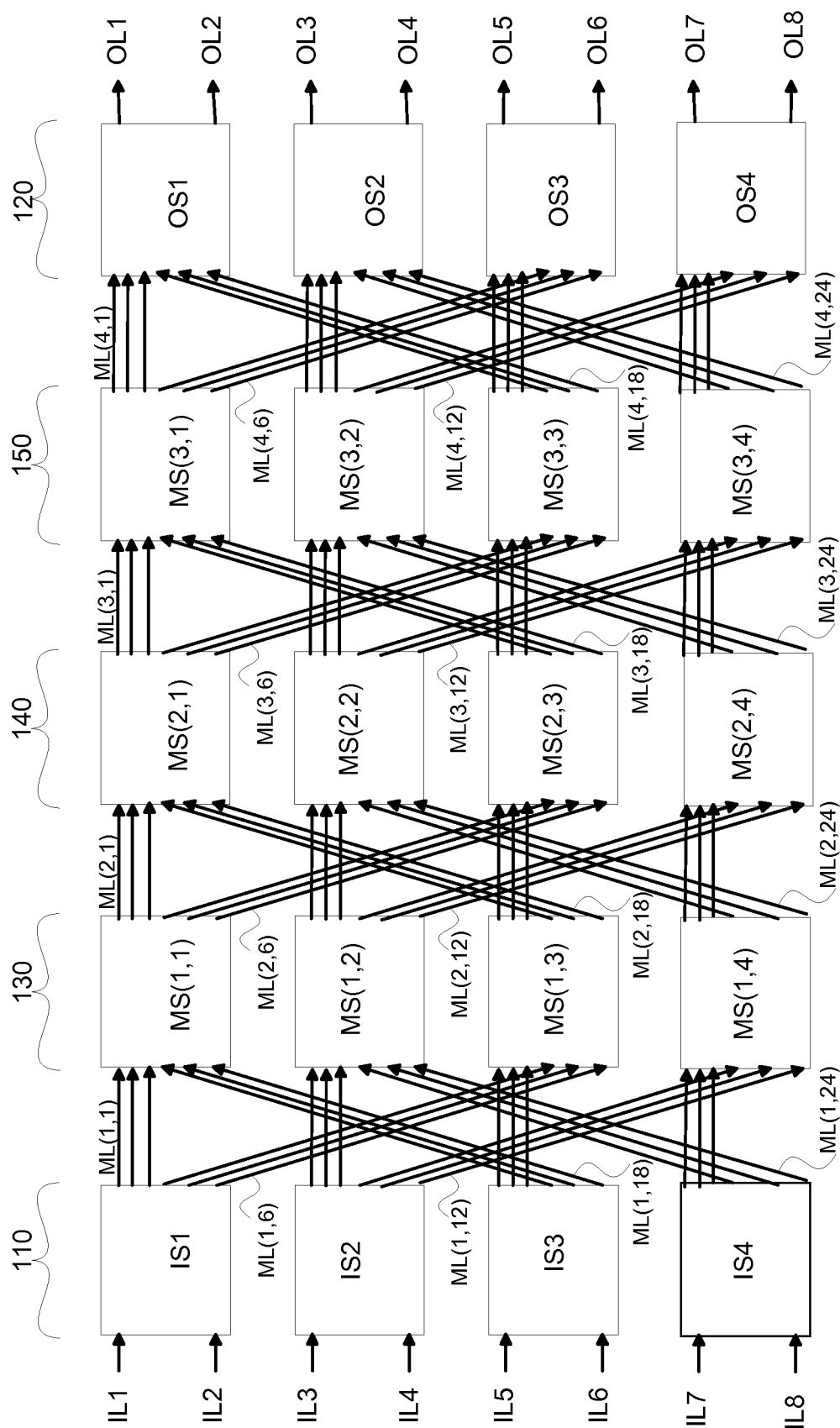


FIG. 3I

300I

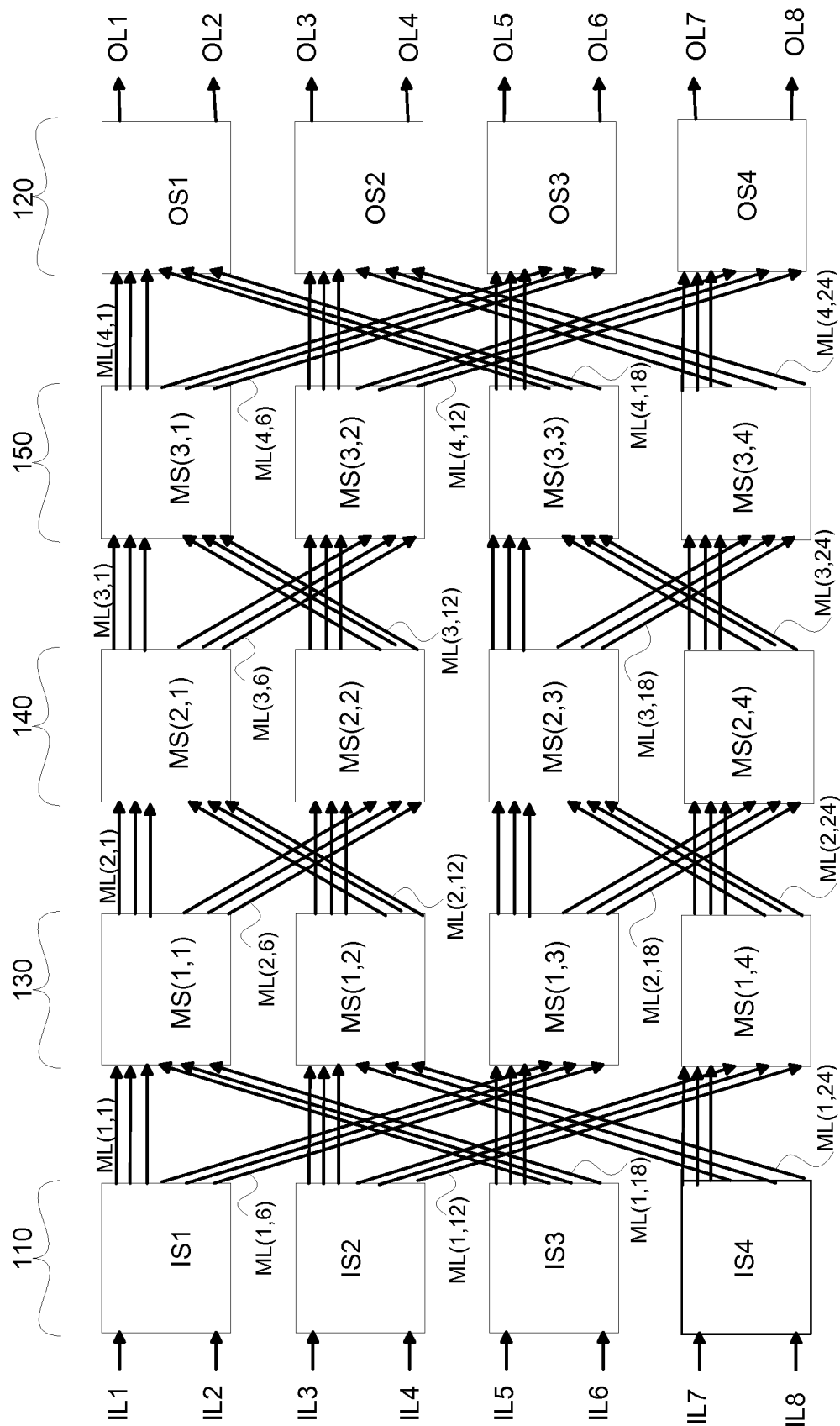


FIG. 3J

300J

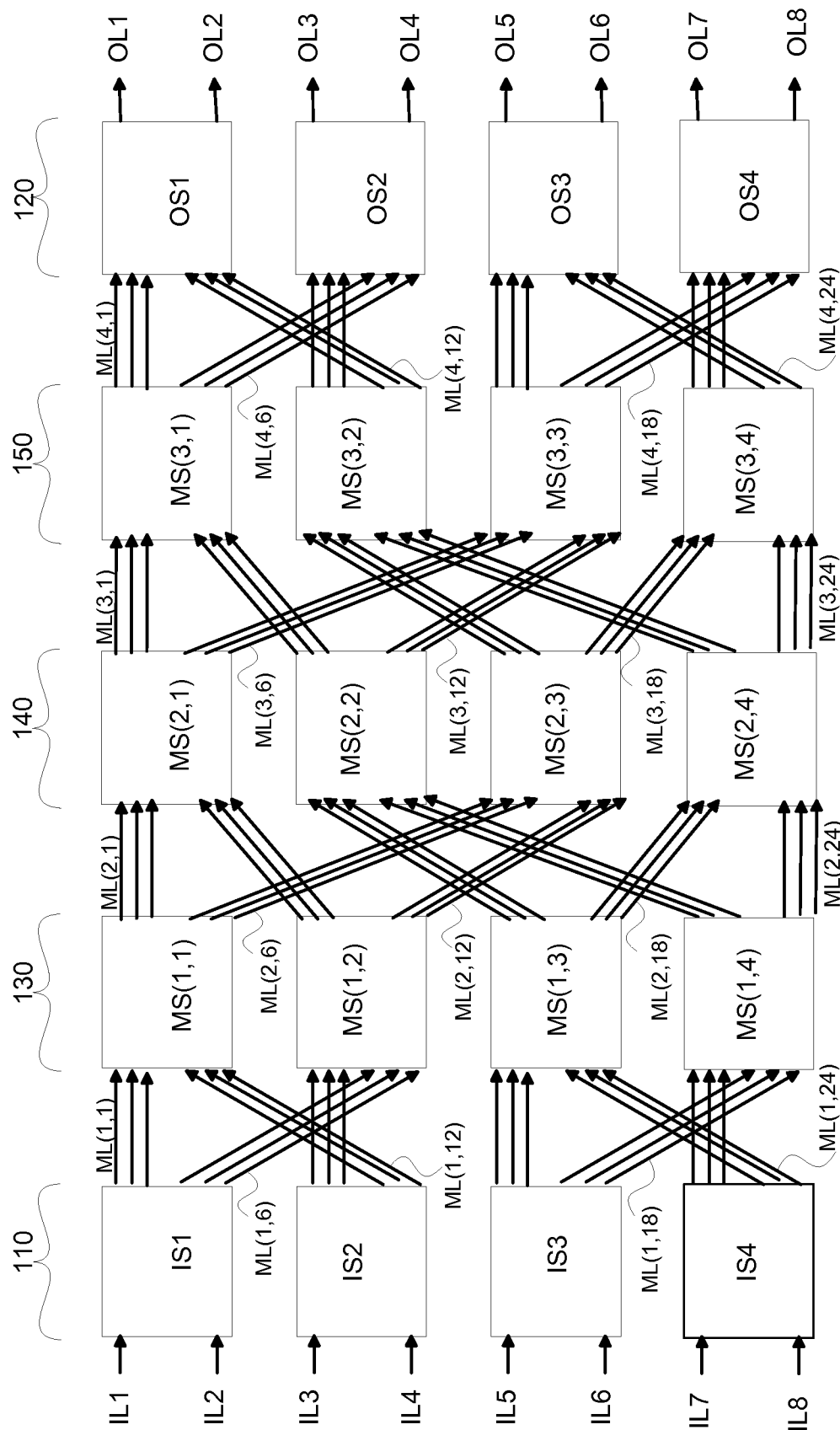


FIG. 3K

300K

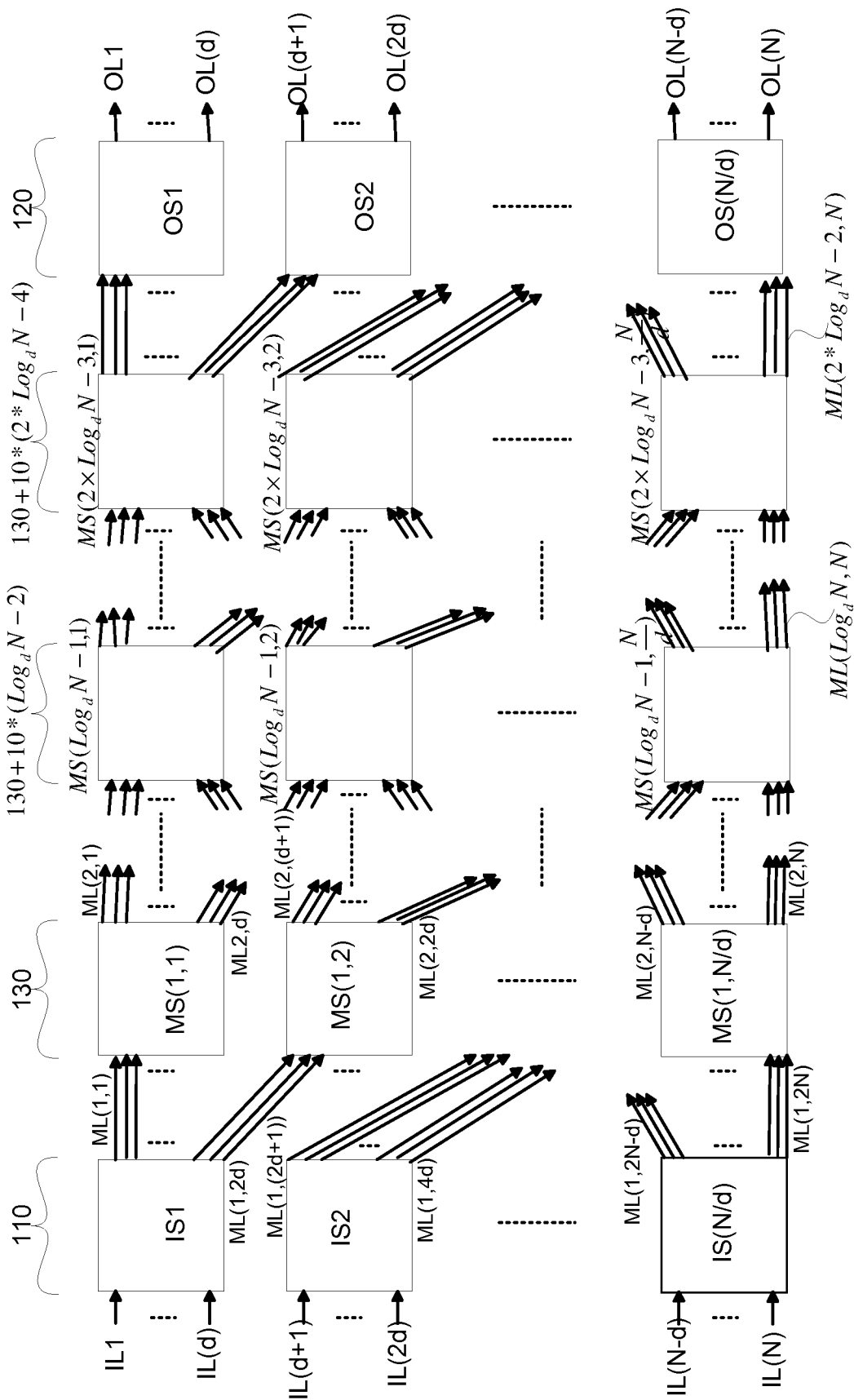


FIG. 3A1

300A1

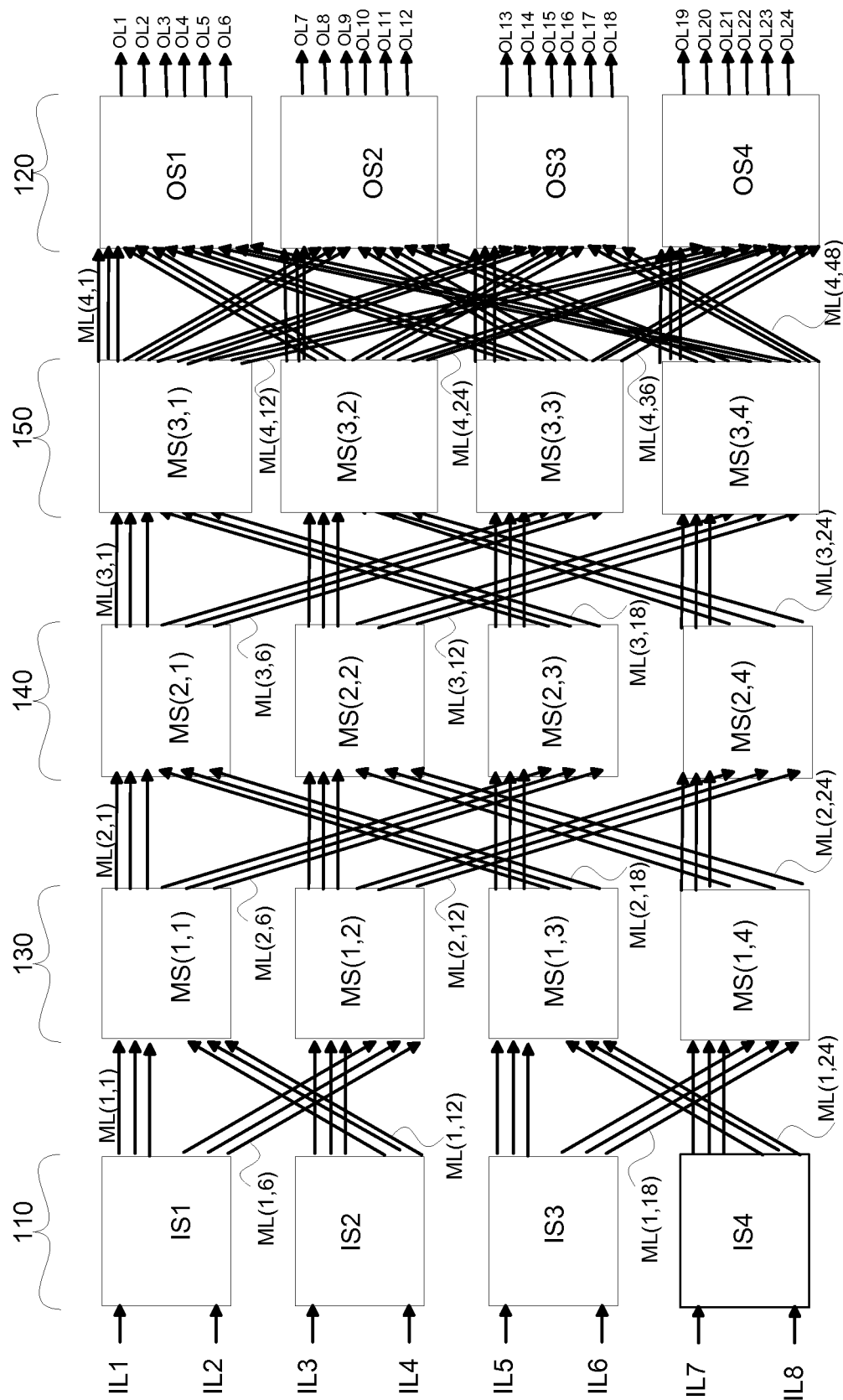


FIG. 3B1

300B1

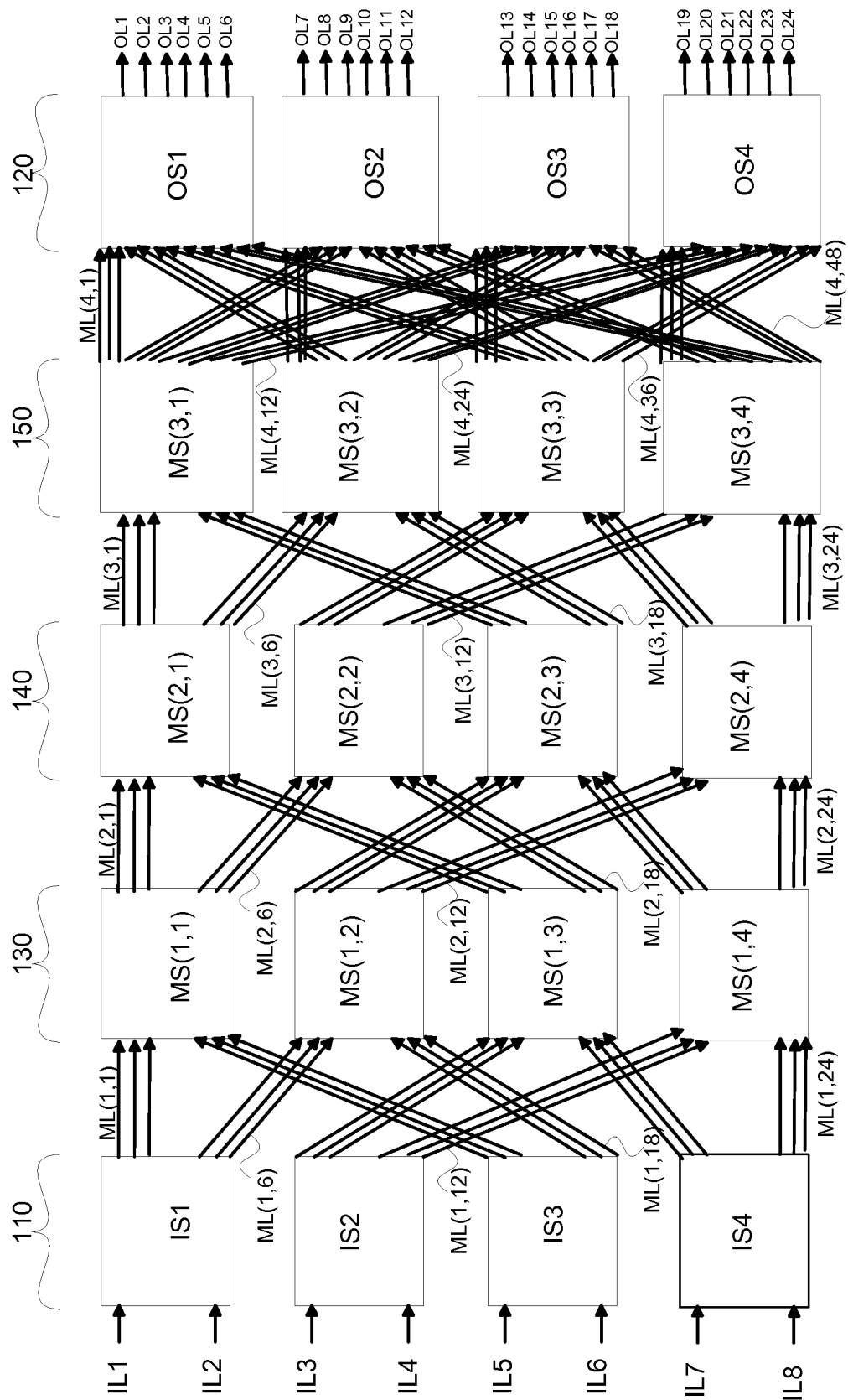


FIG. 3C1

300C1

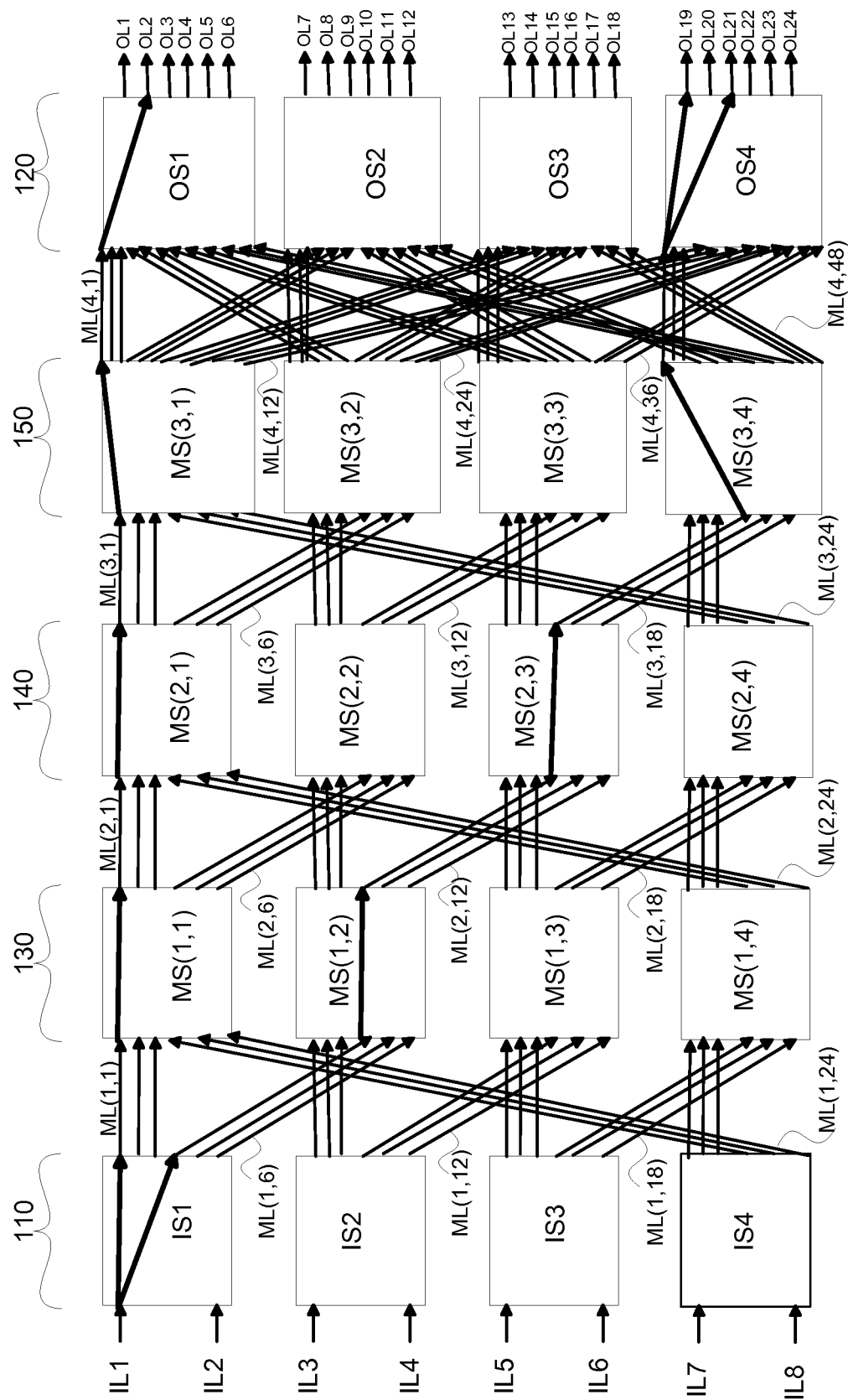


FIG. 3D1

300D1

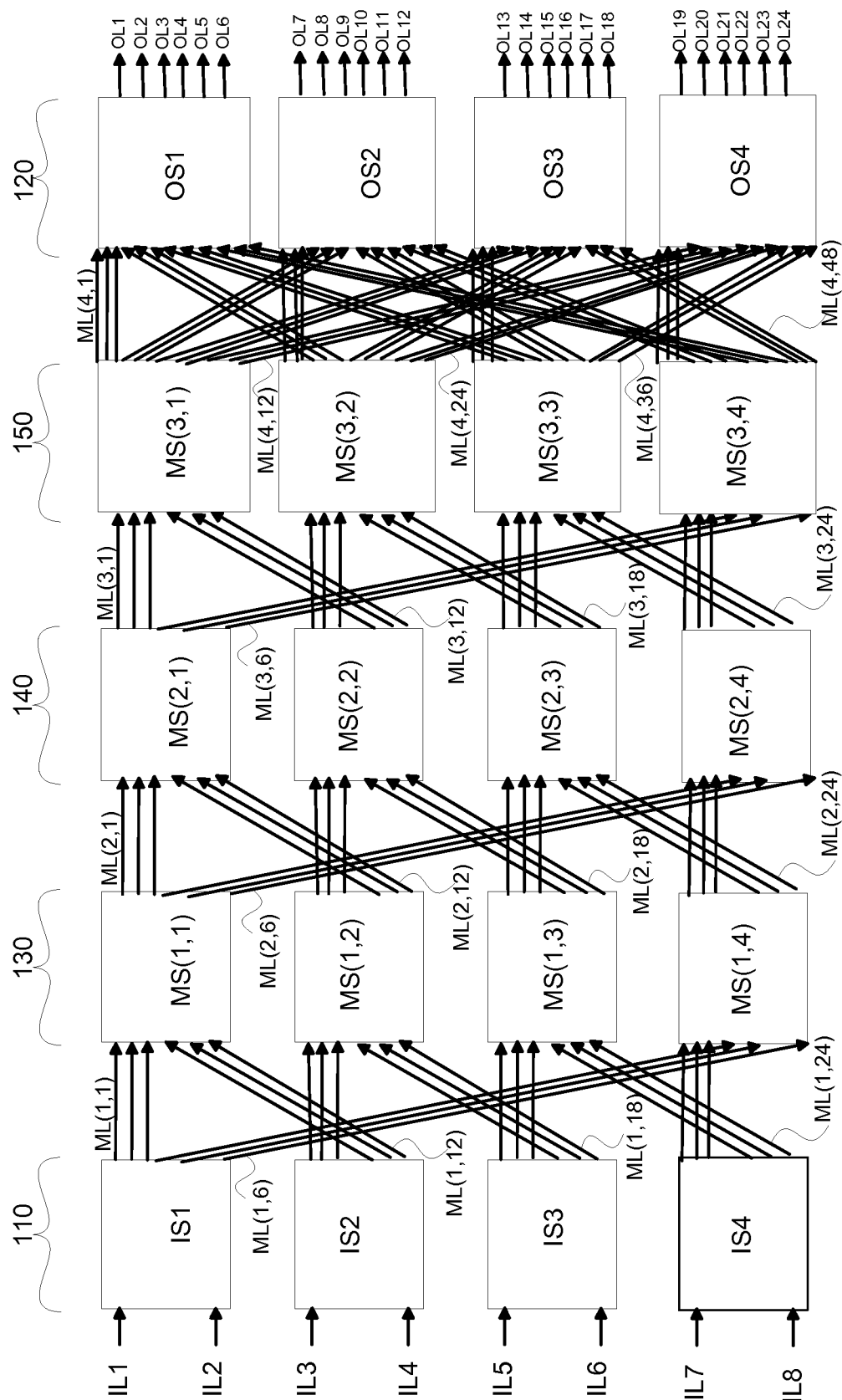


FIG. 3E1

300E1

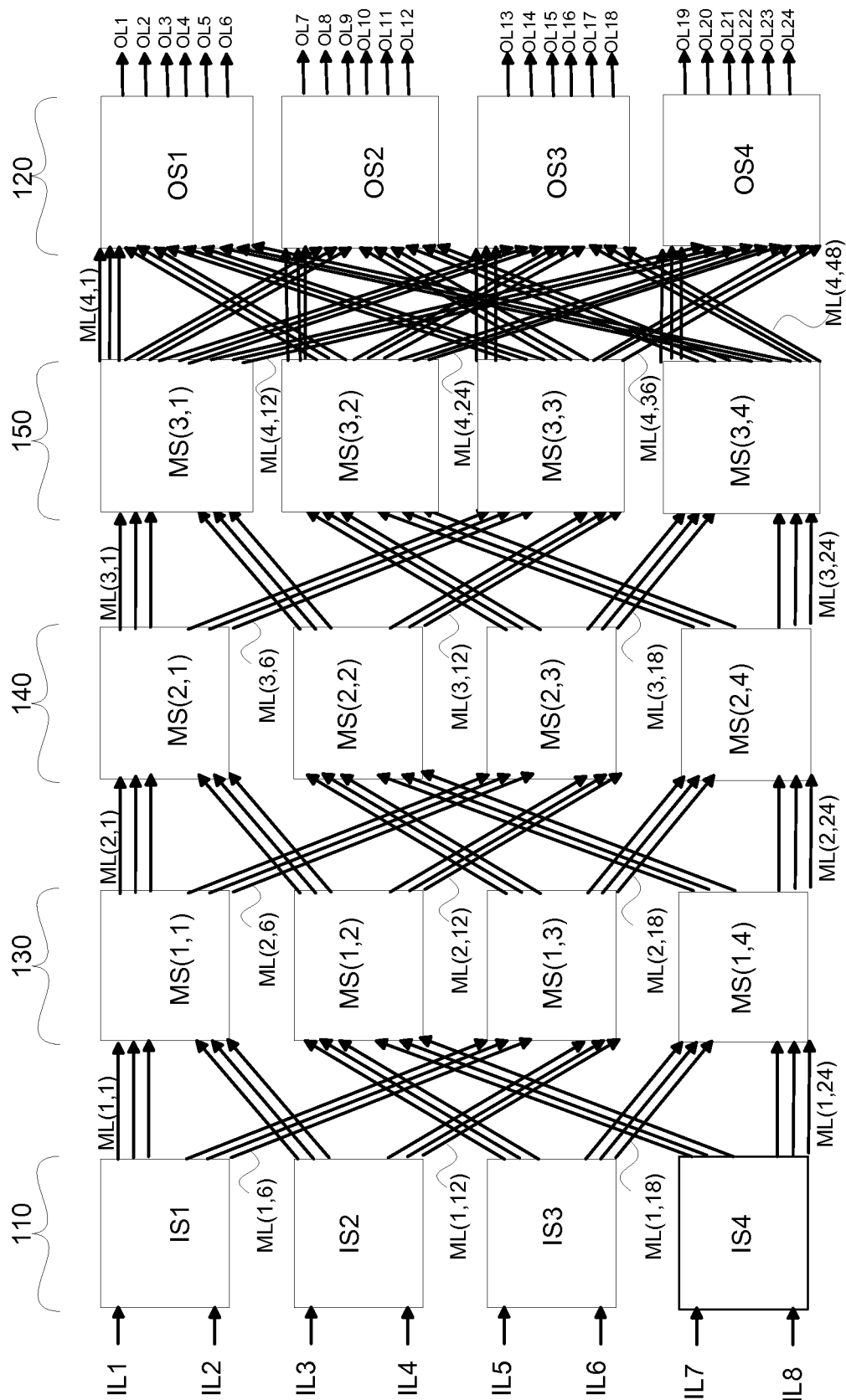


FIG. 3F1

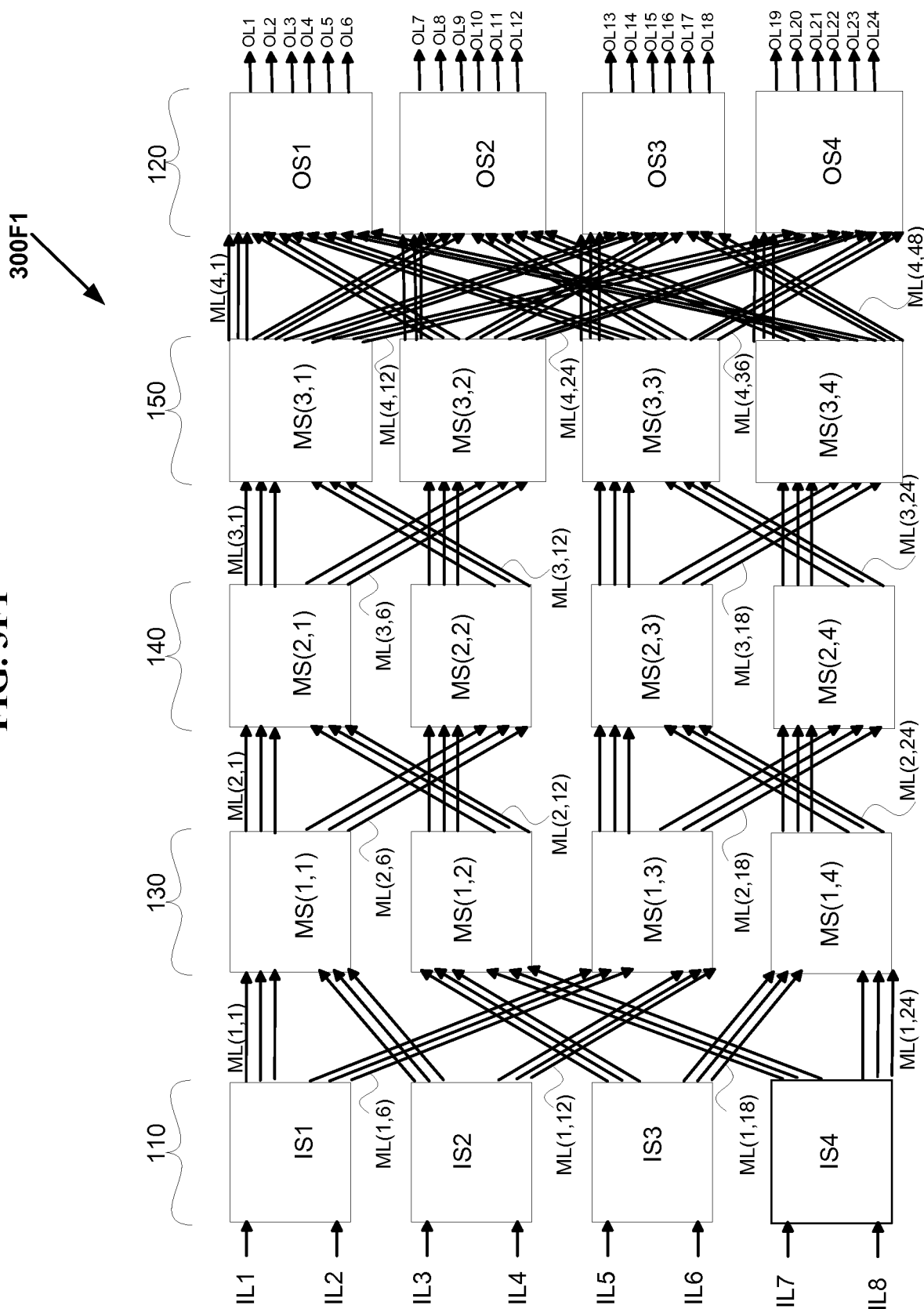


FIG. 3G1

300G1

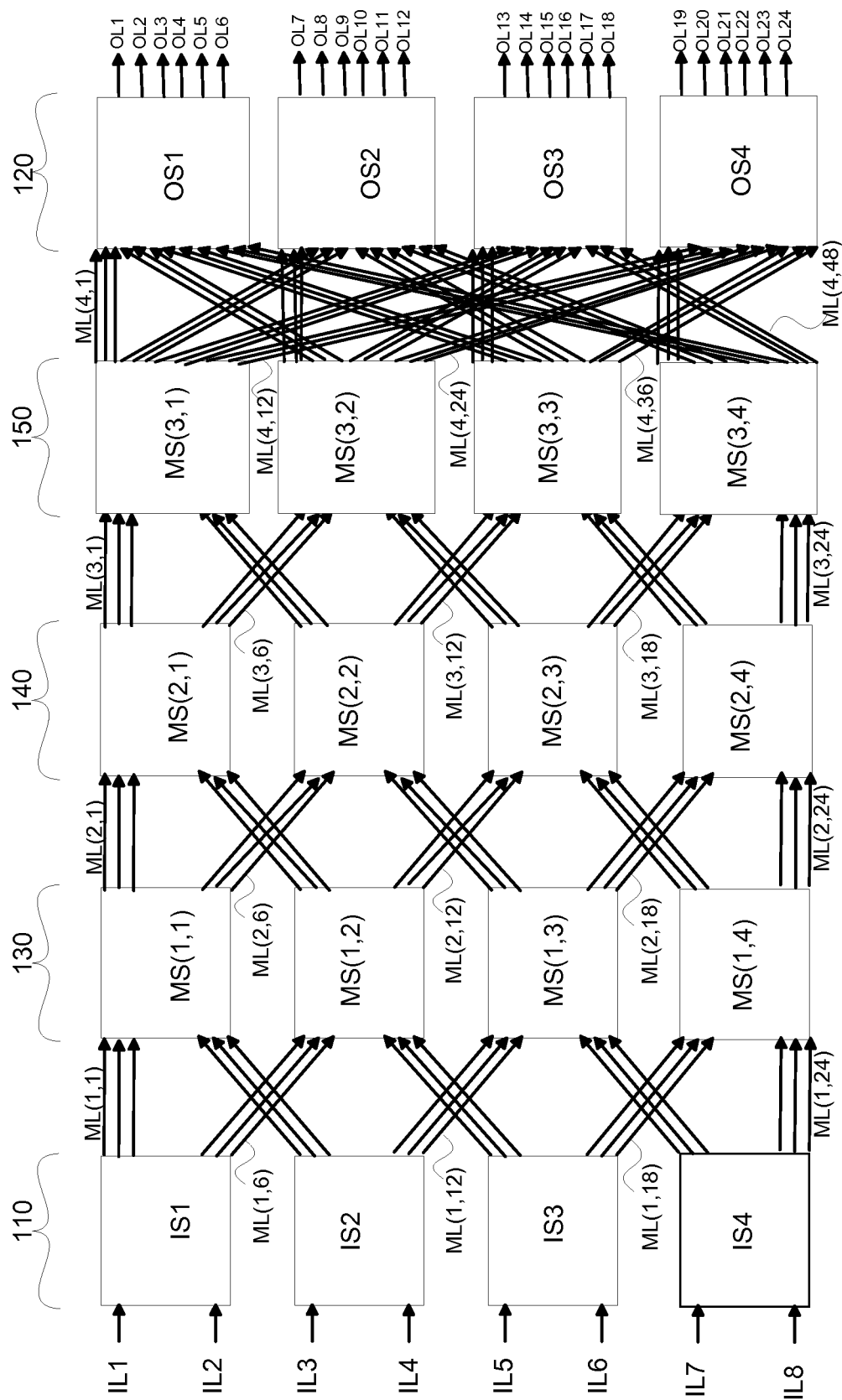


FIG. 3H1

300H1

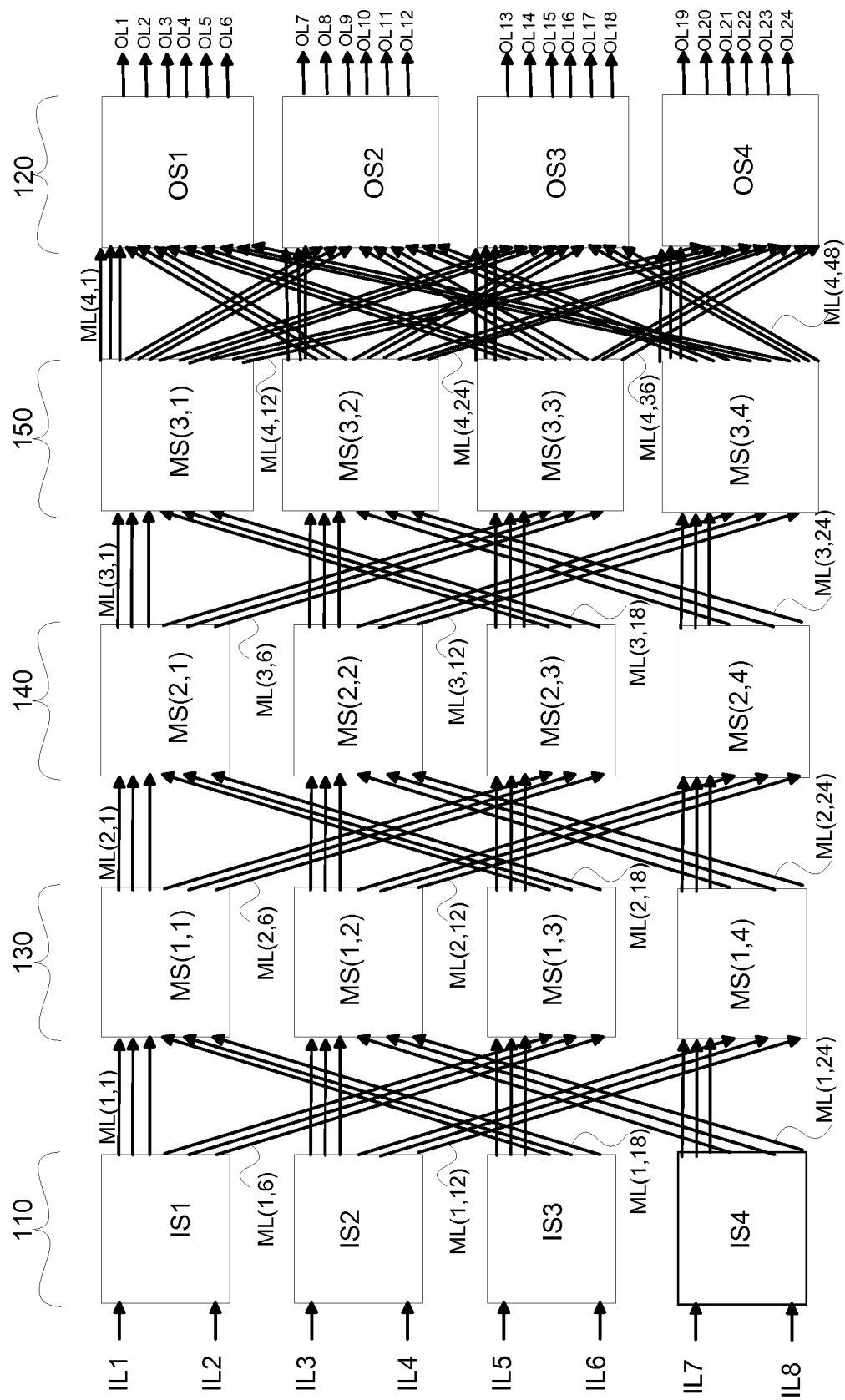


FIG. 3H

30011

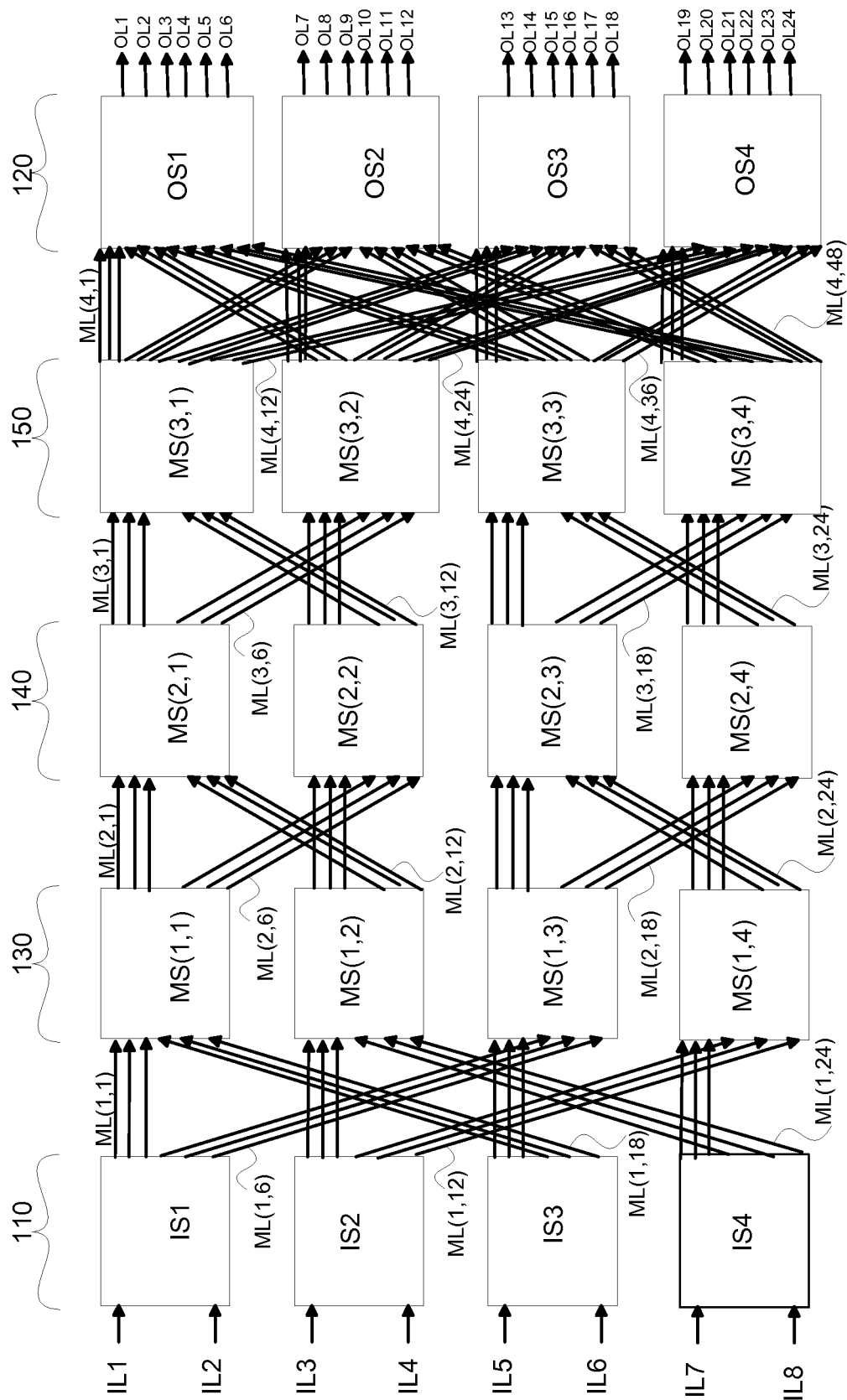


FIG. 3J1

300J1

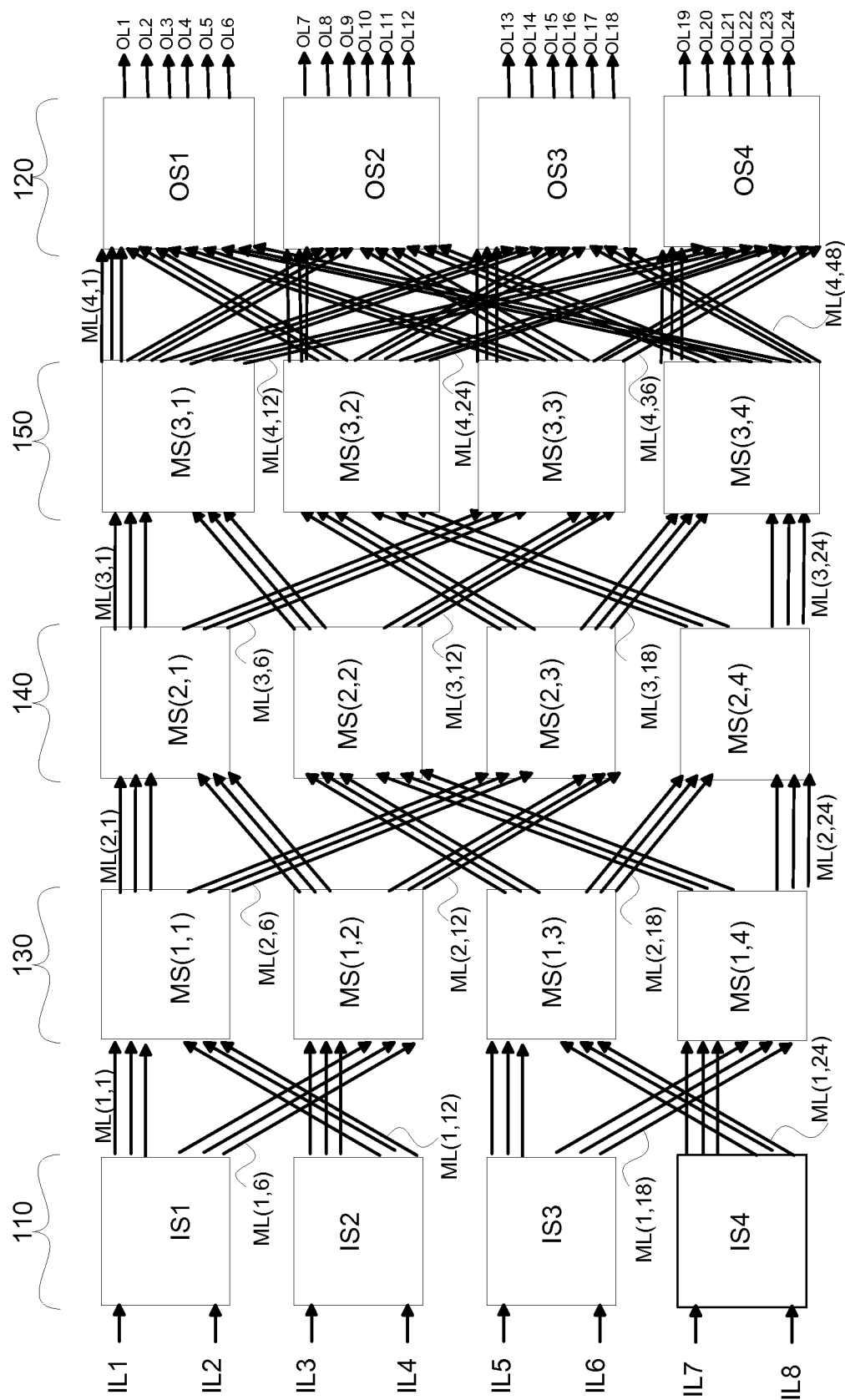


FIG. 3K1

300K1

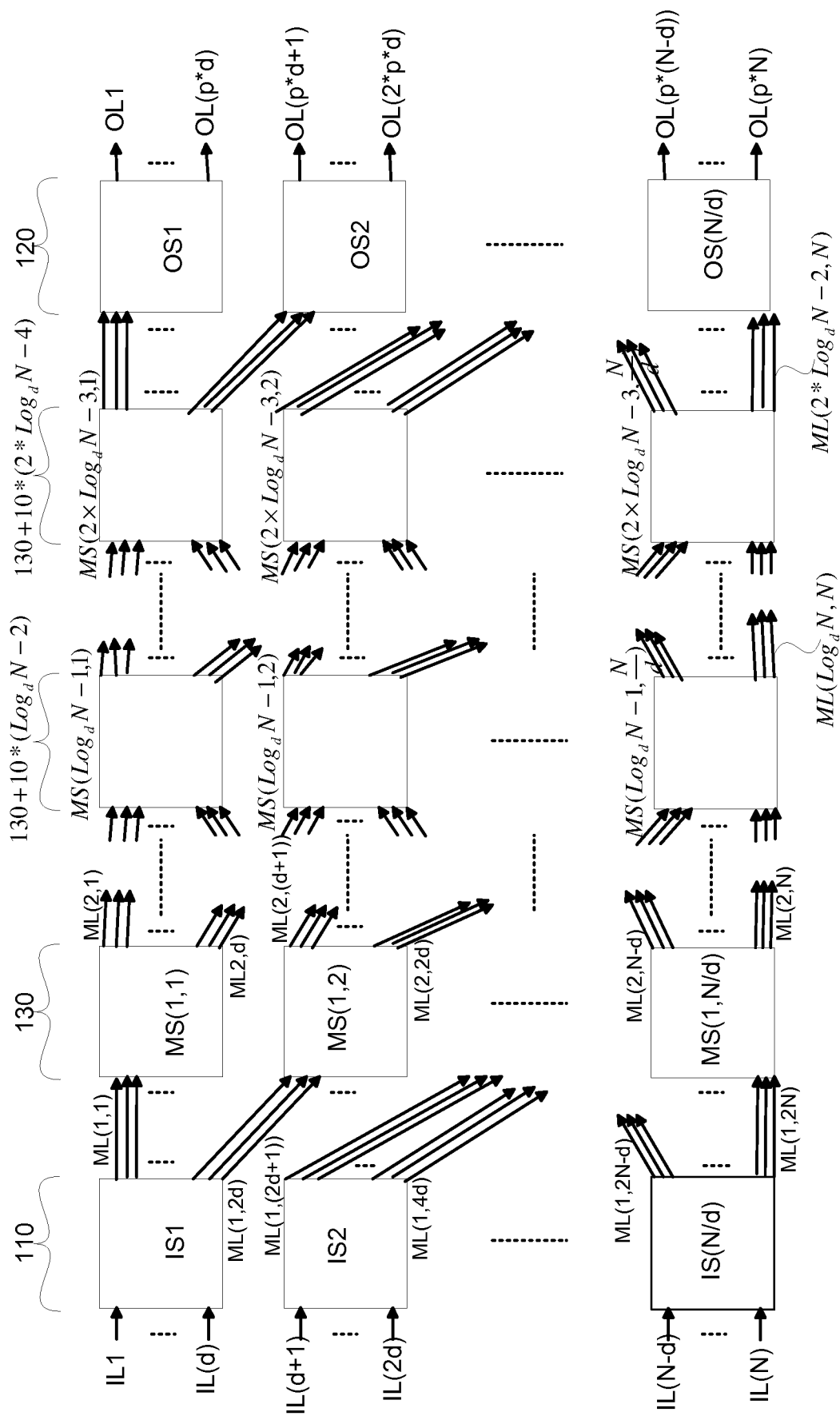


FIG. 3A2

300A2

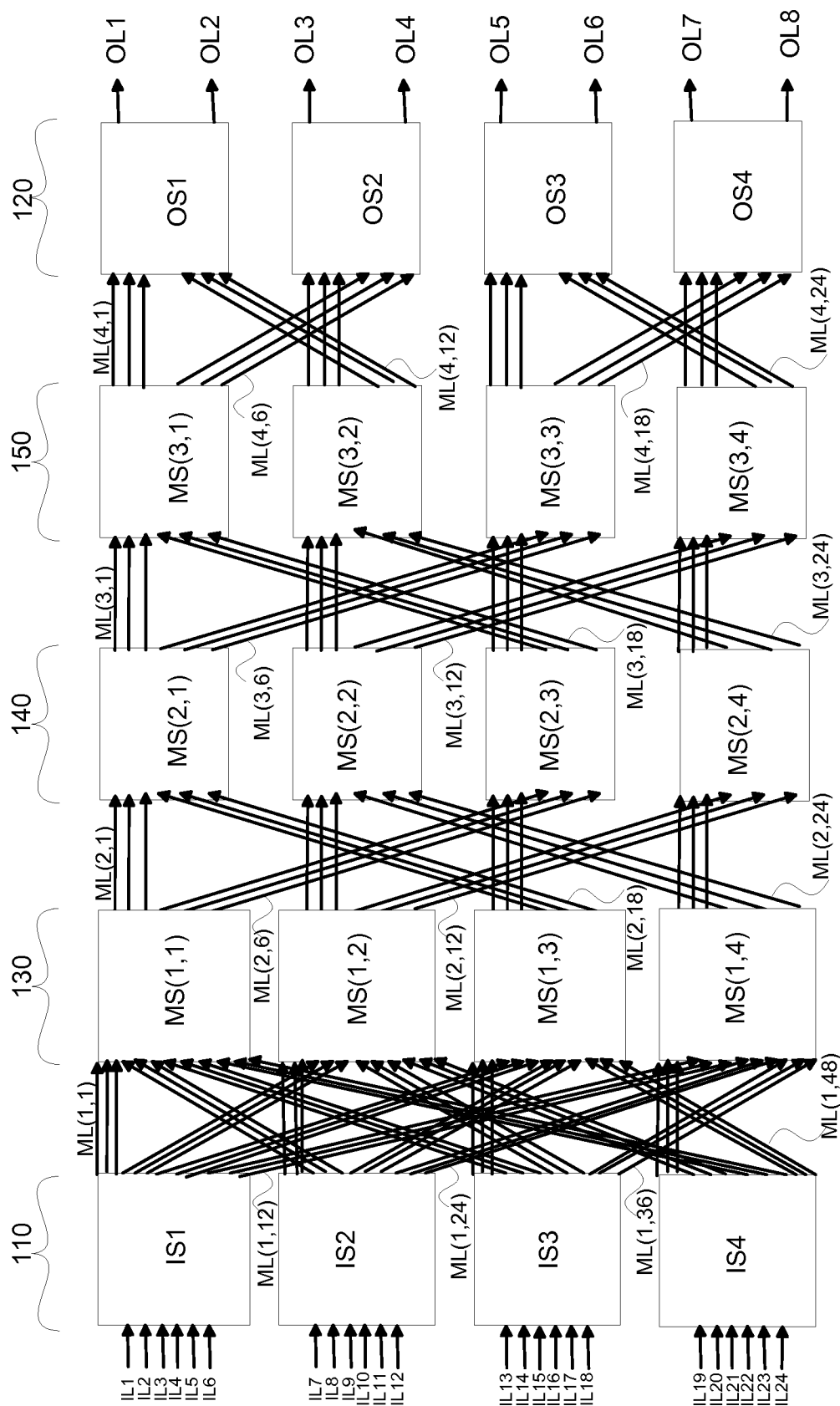


FIG. 3B2

300B2

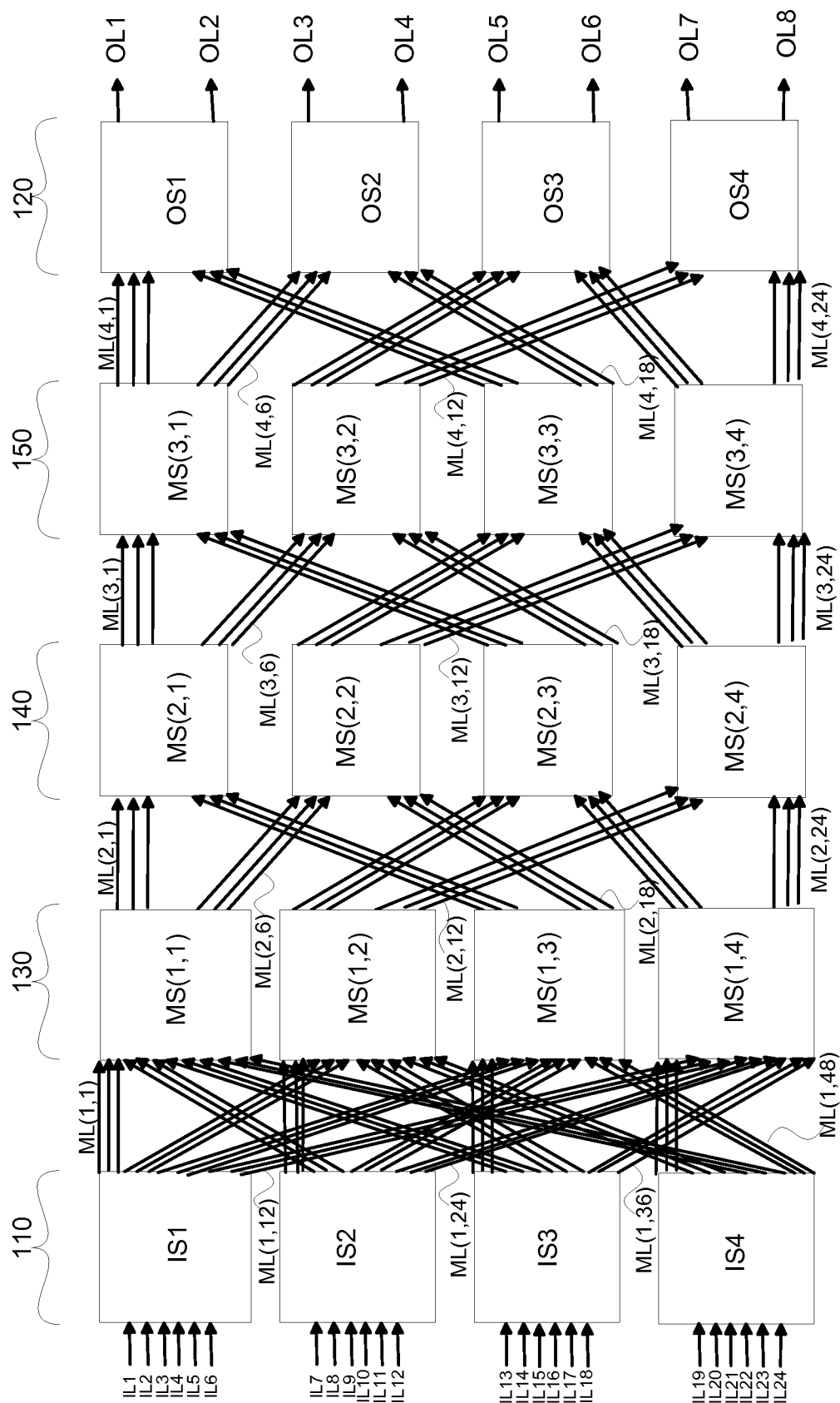


FIG. 3C2

300C2

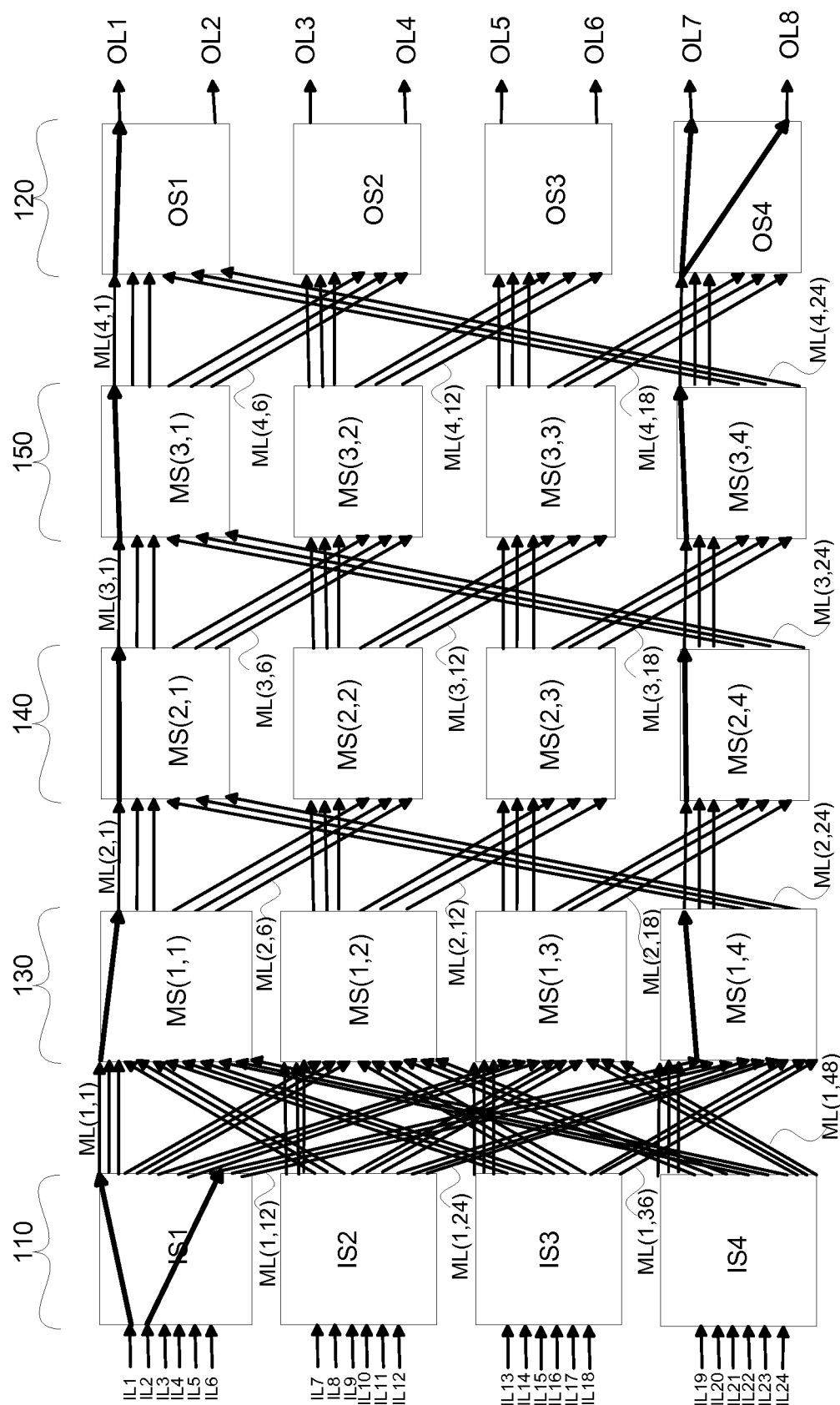


FIG. 3D2

300D2

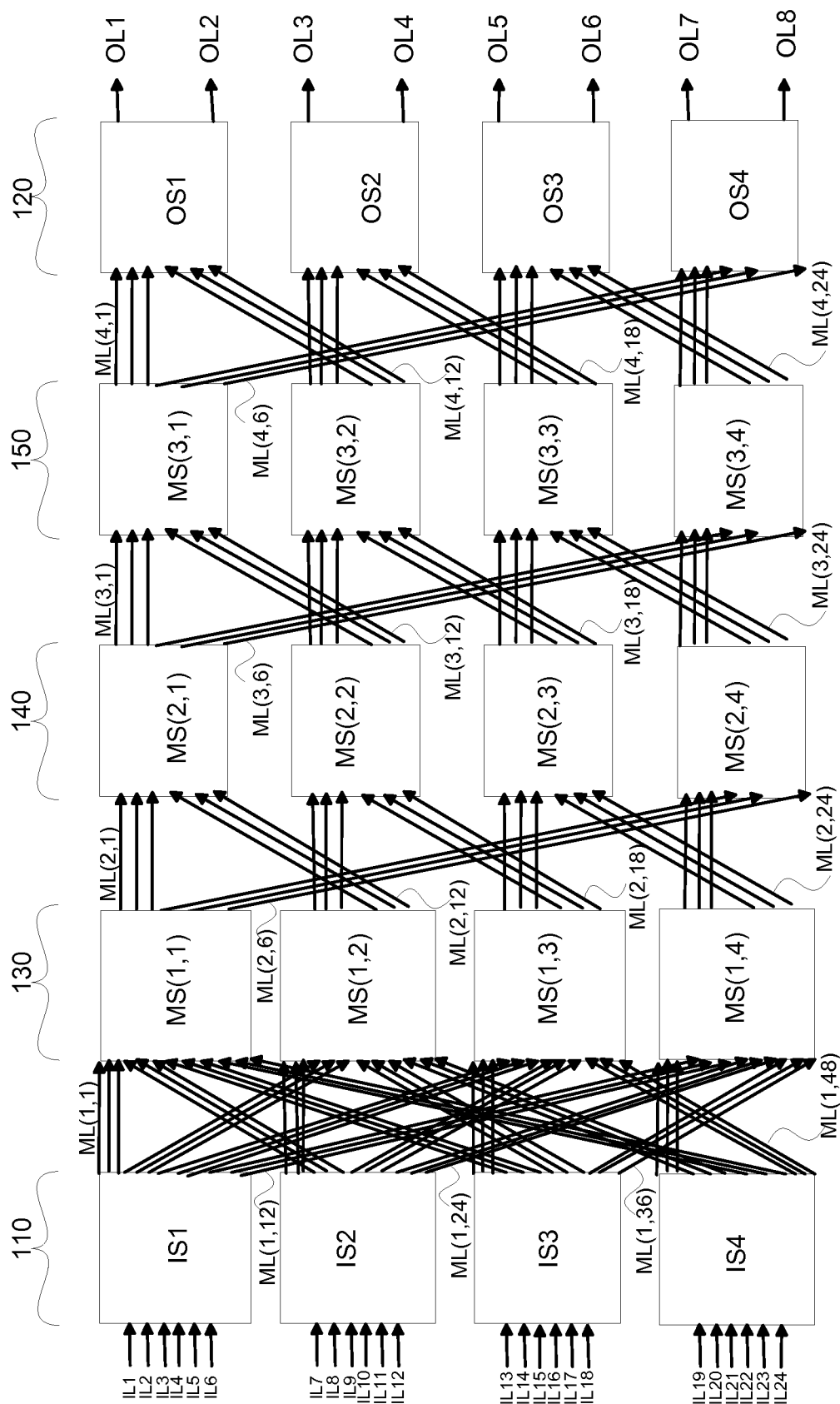


FIG. 3E2

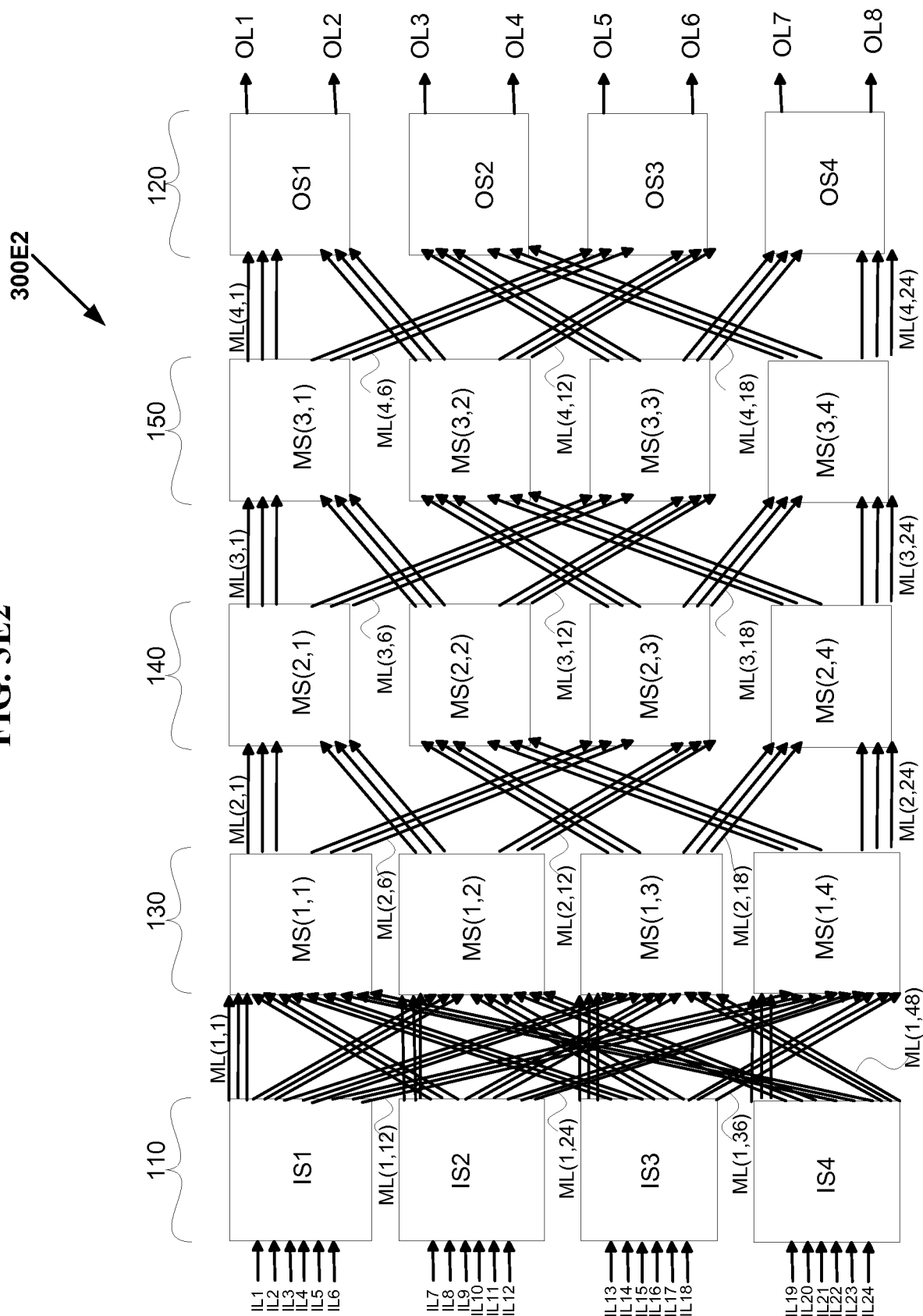


FIG. 3F2

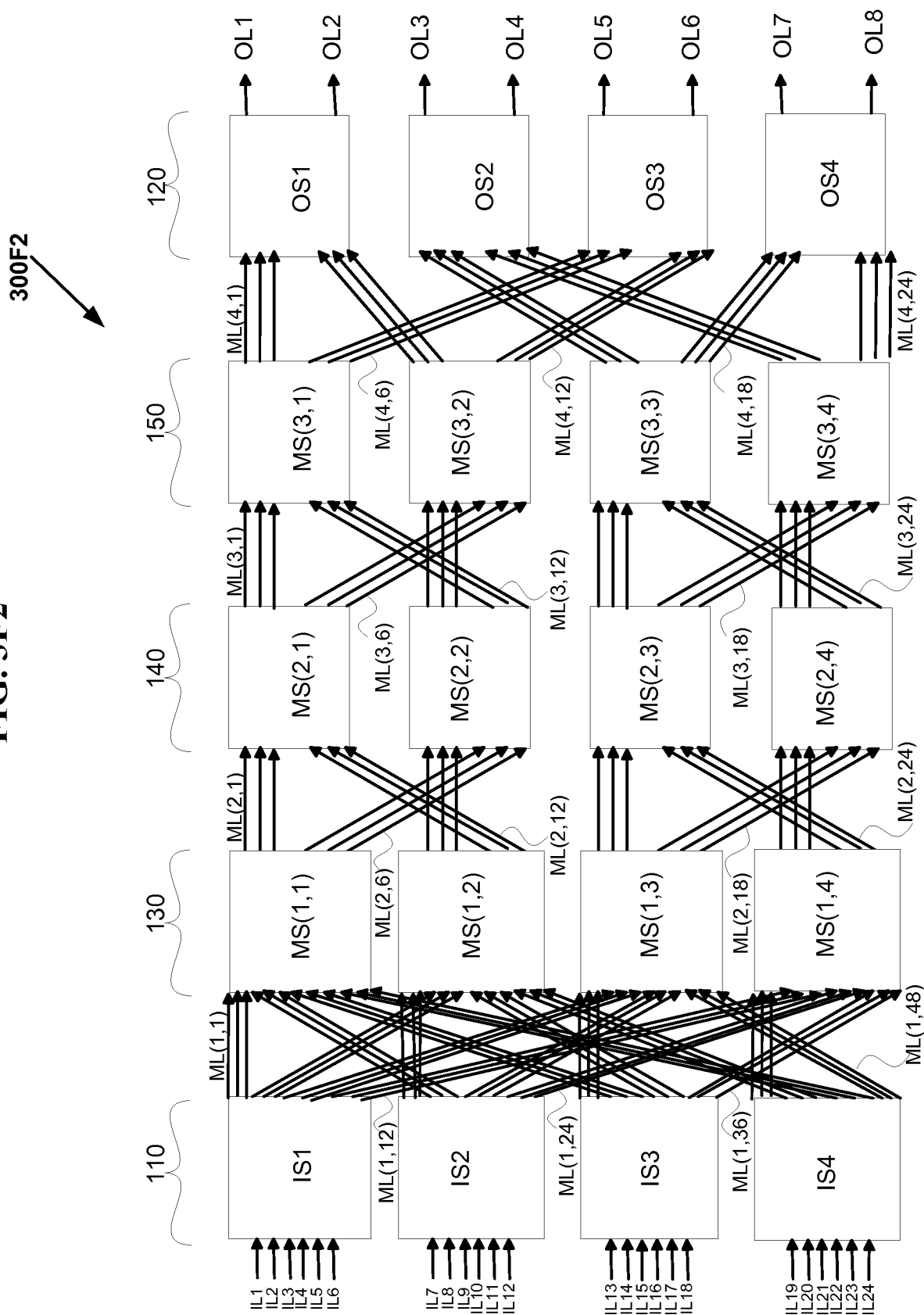


FIG. 3G2

300G2

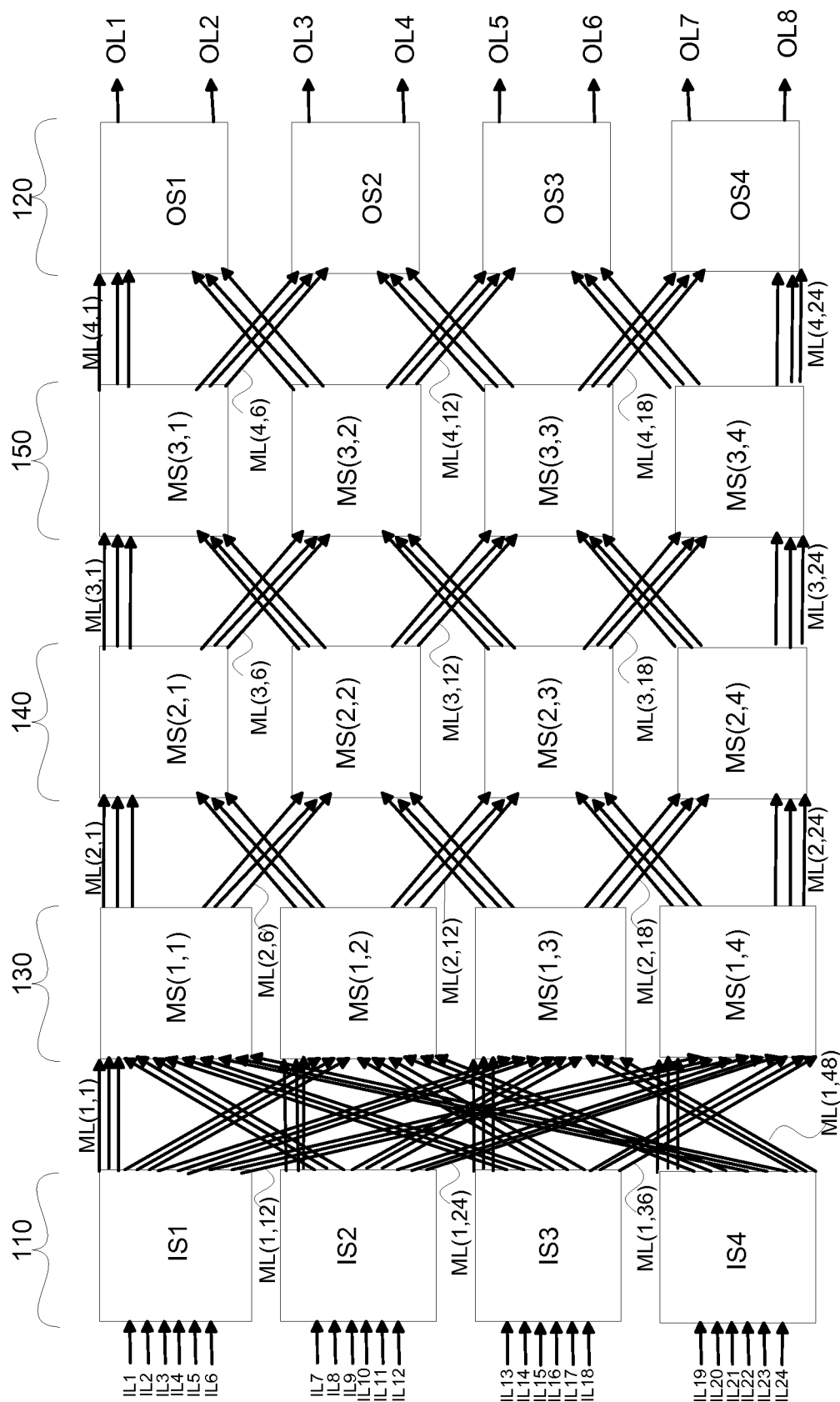


FIG. 3H2

300H2

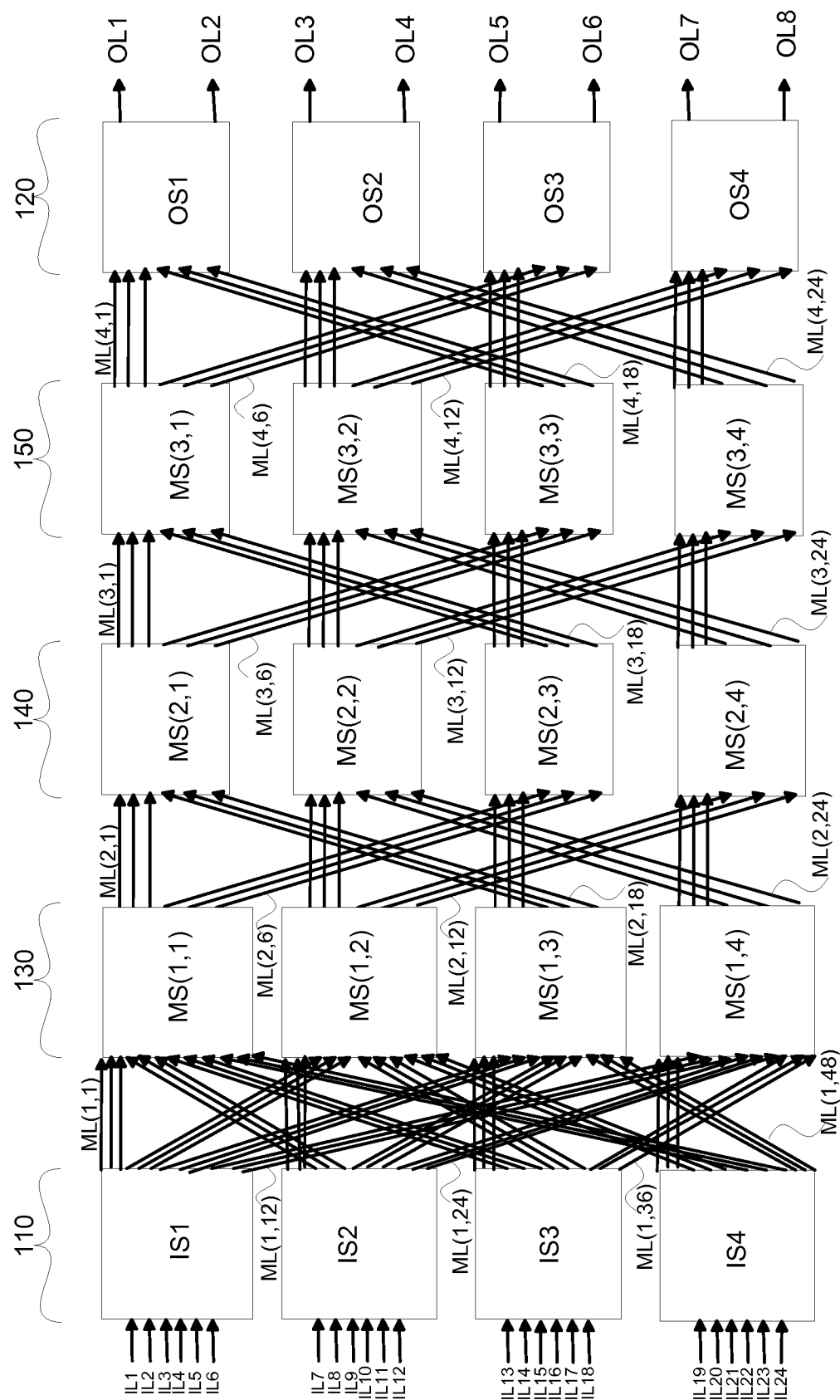


FIG. 3I2

300I2

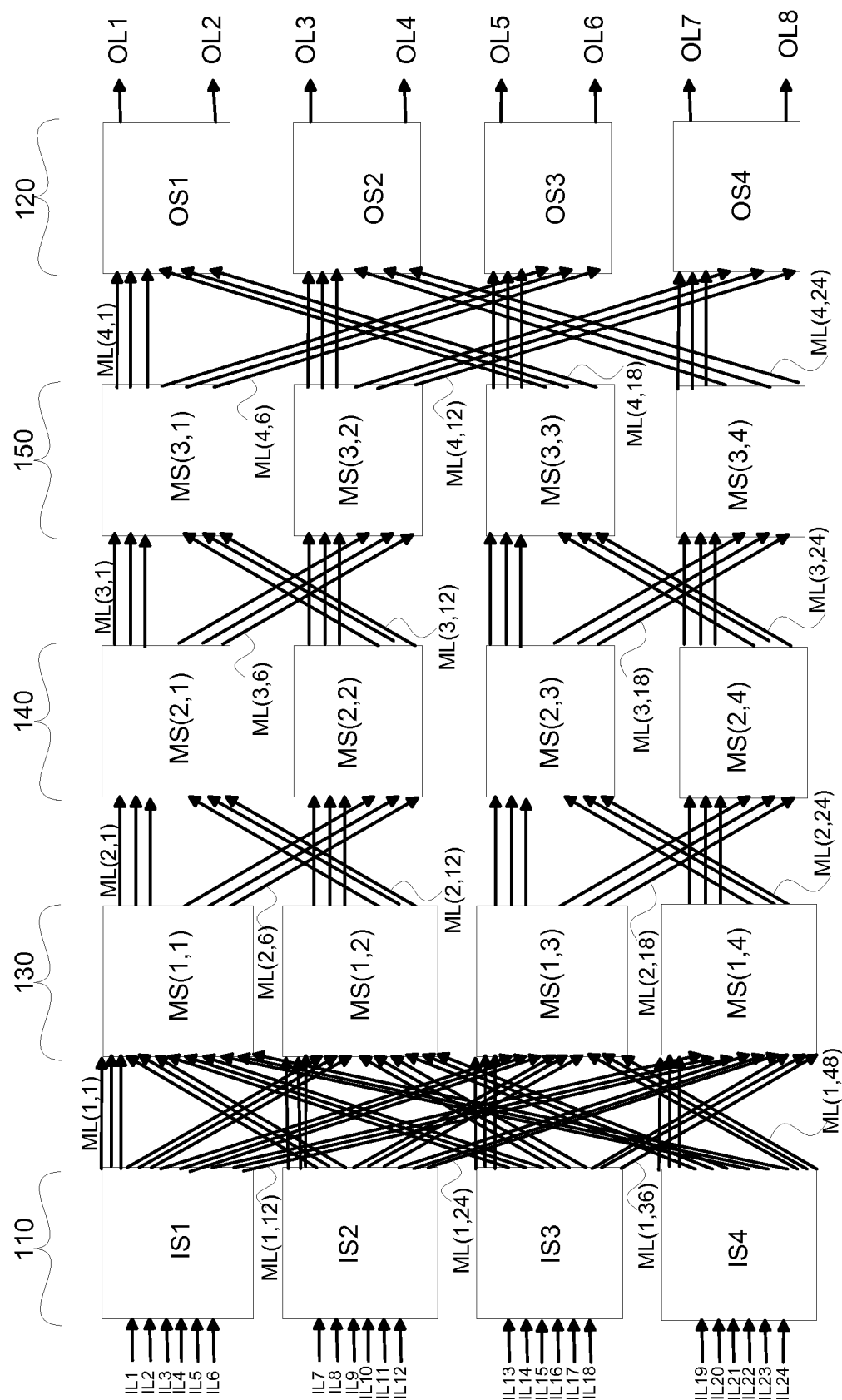


FIG. 3J2

300J2

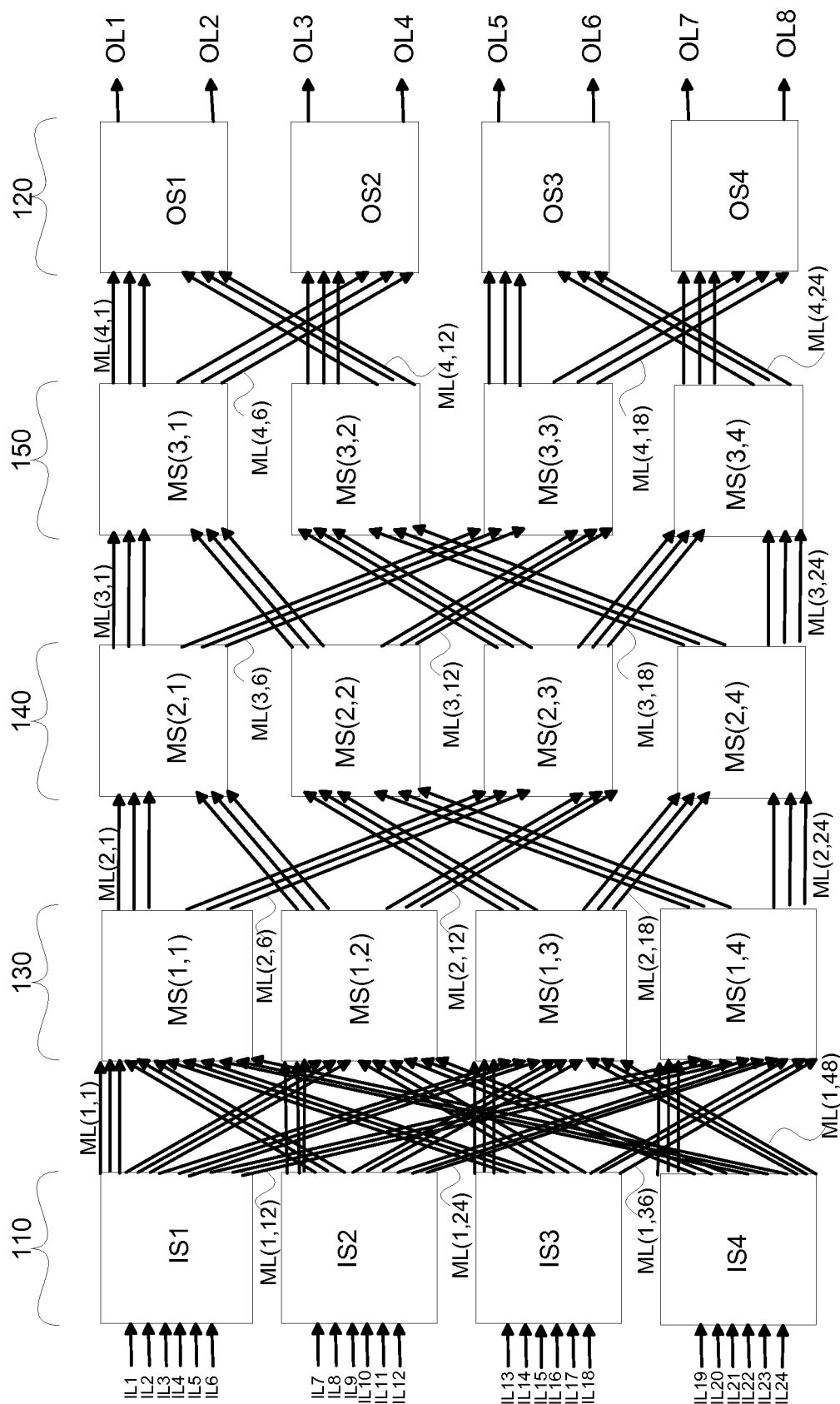
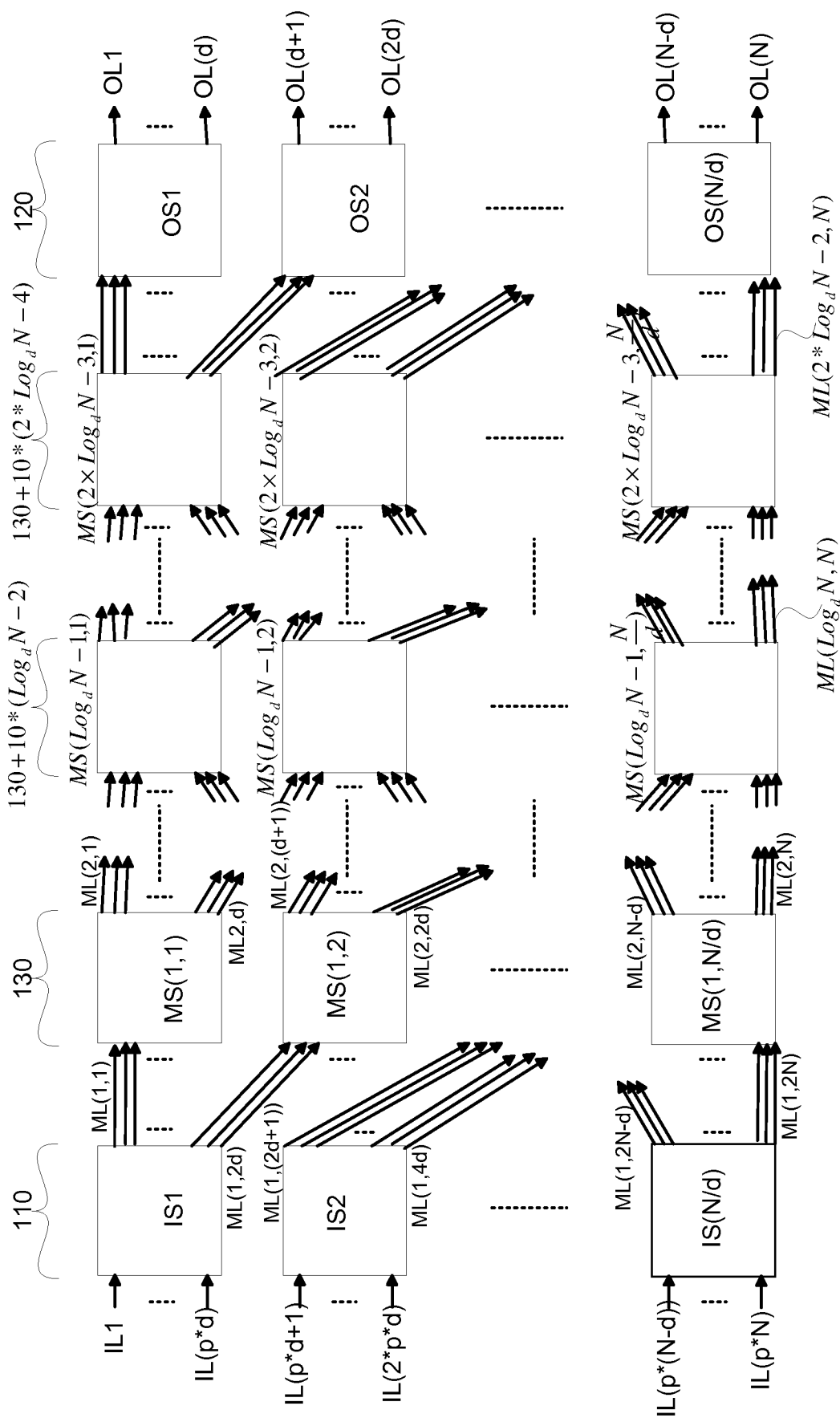


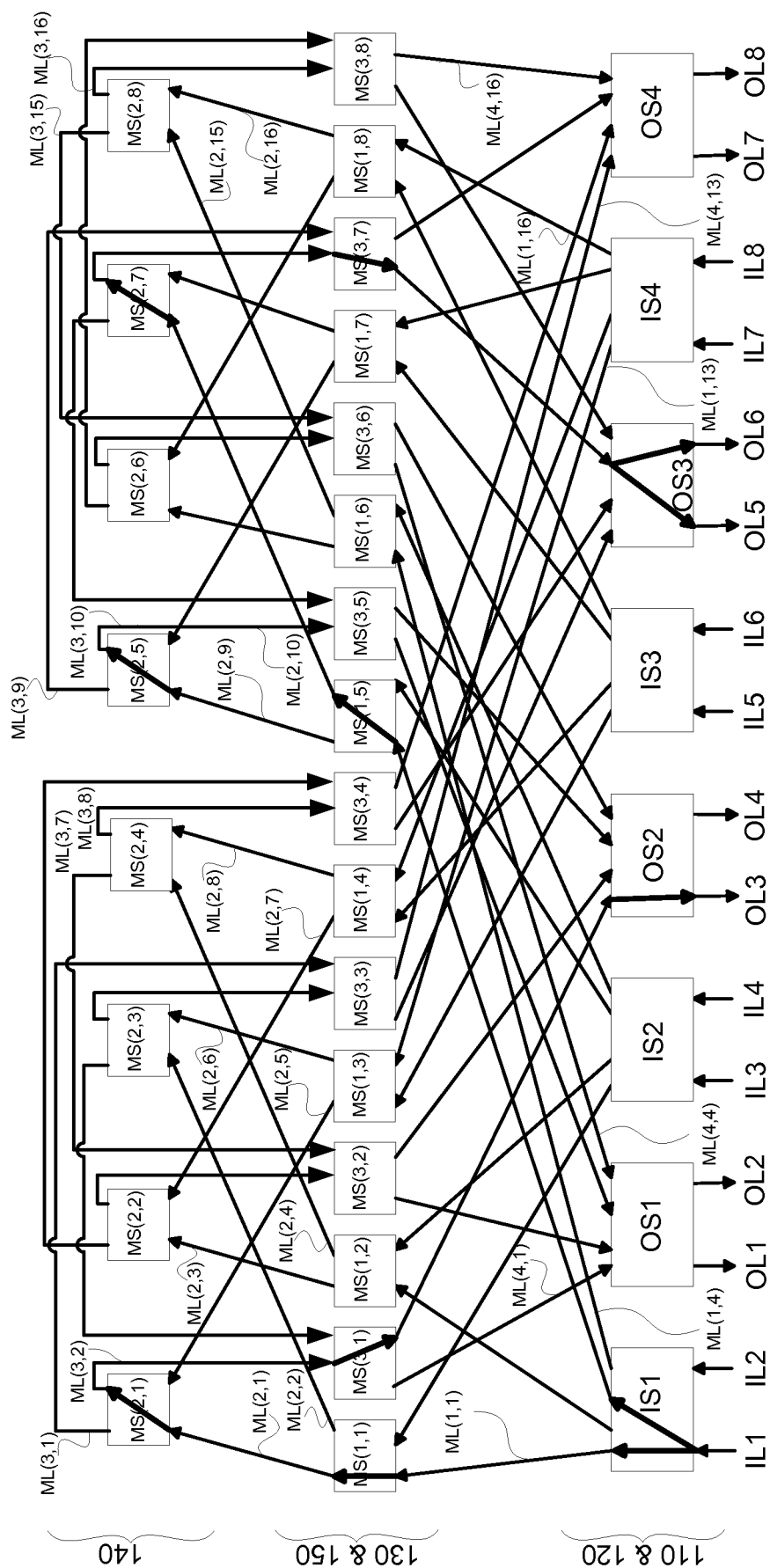
FIG. 3K2

300K2



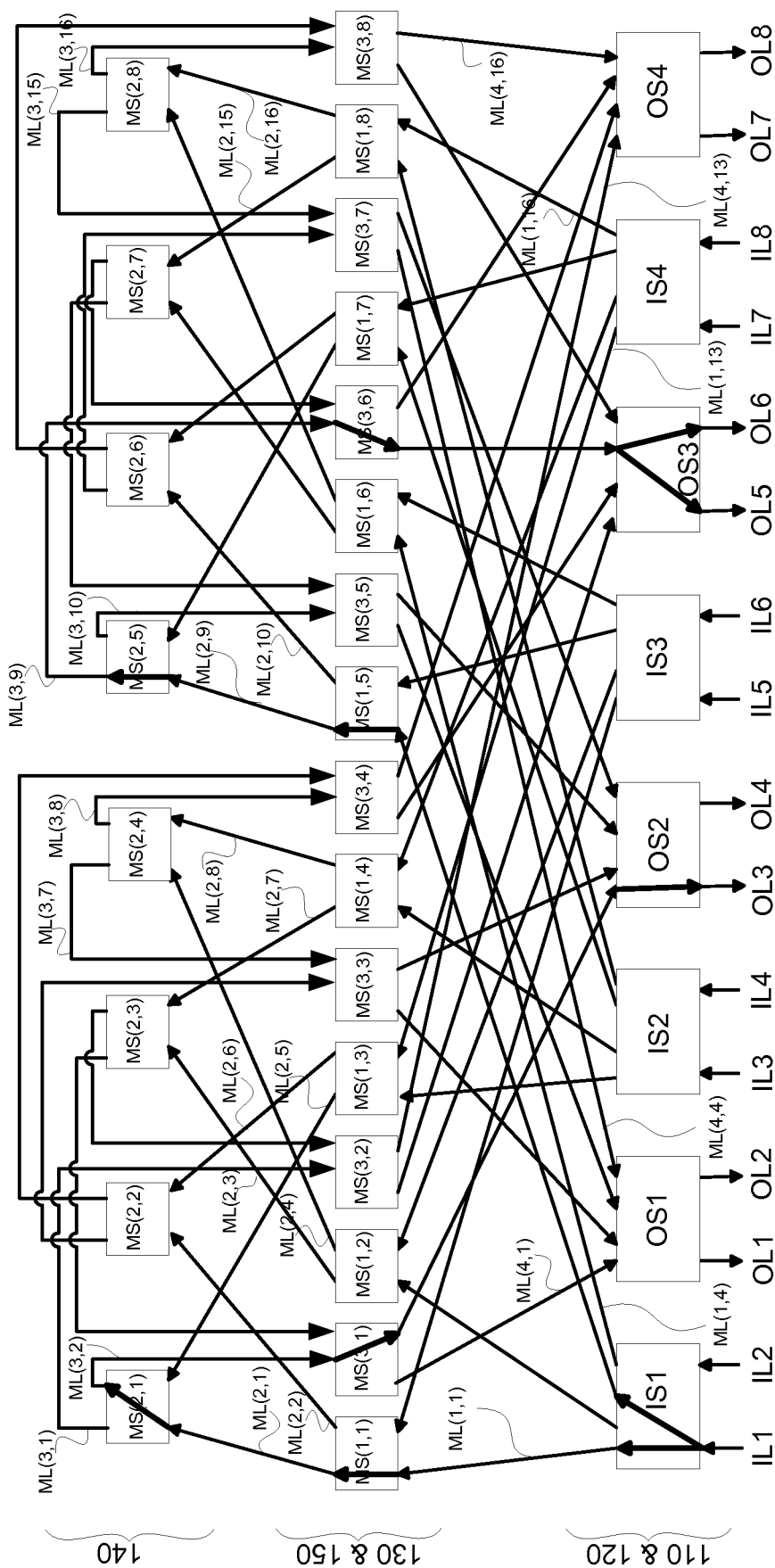
400A

FIG. 4A



400A1

FIG. 4A1



400A2

FIG. 4A2

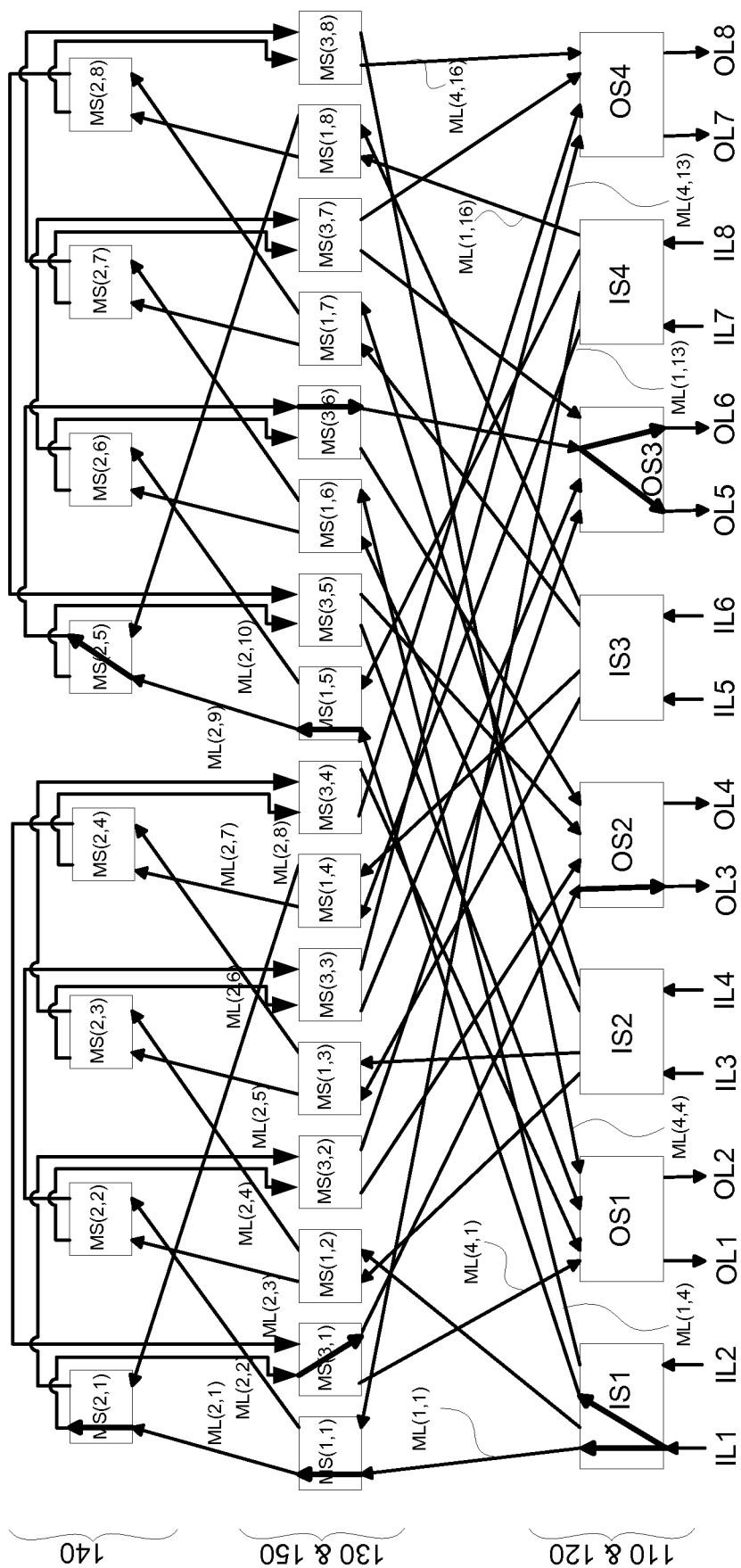
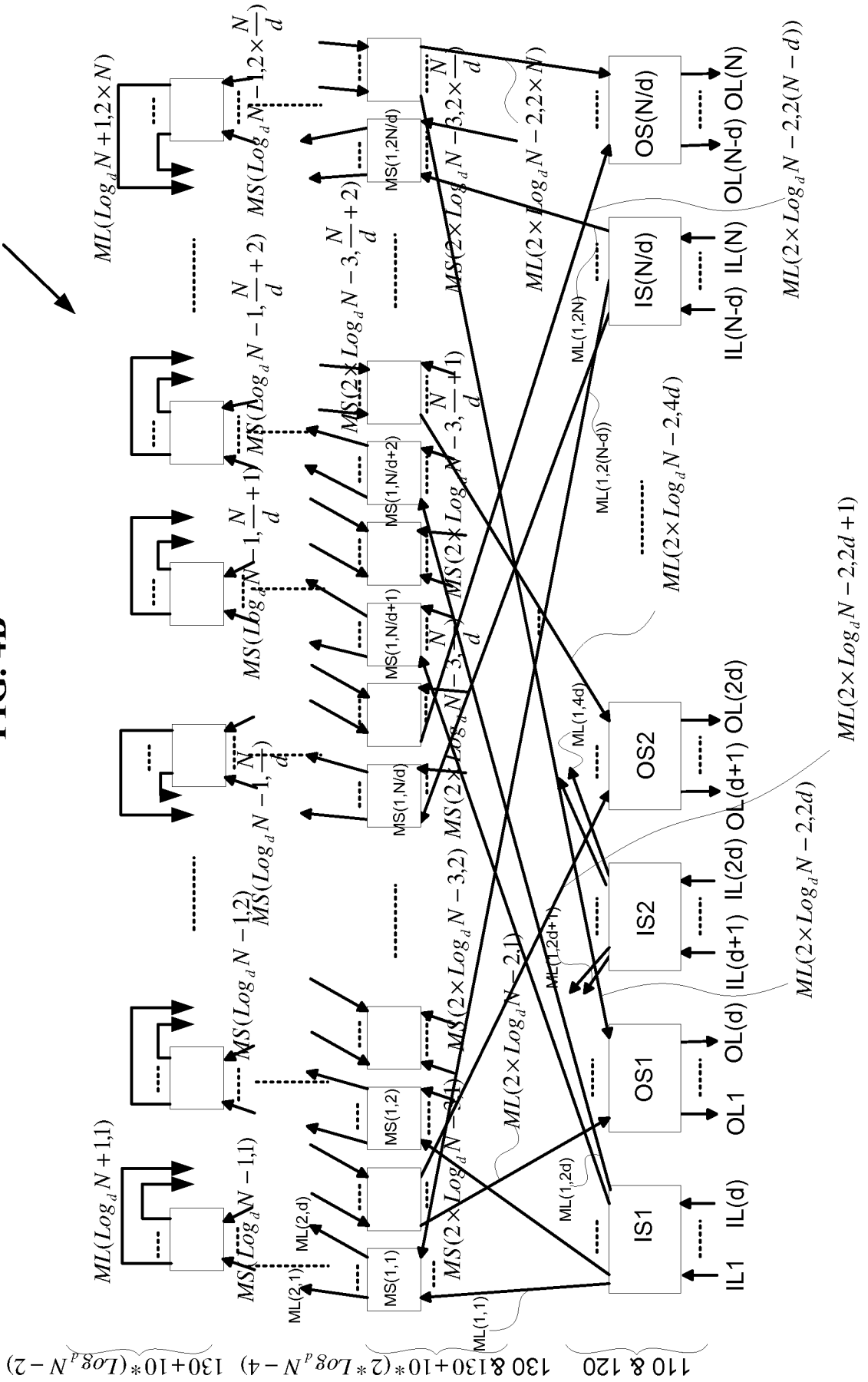


FIG. 4B

400B



400C

FIG. 4C

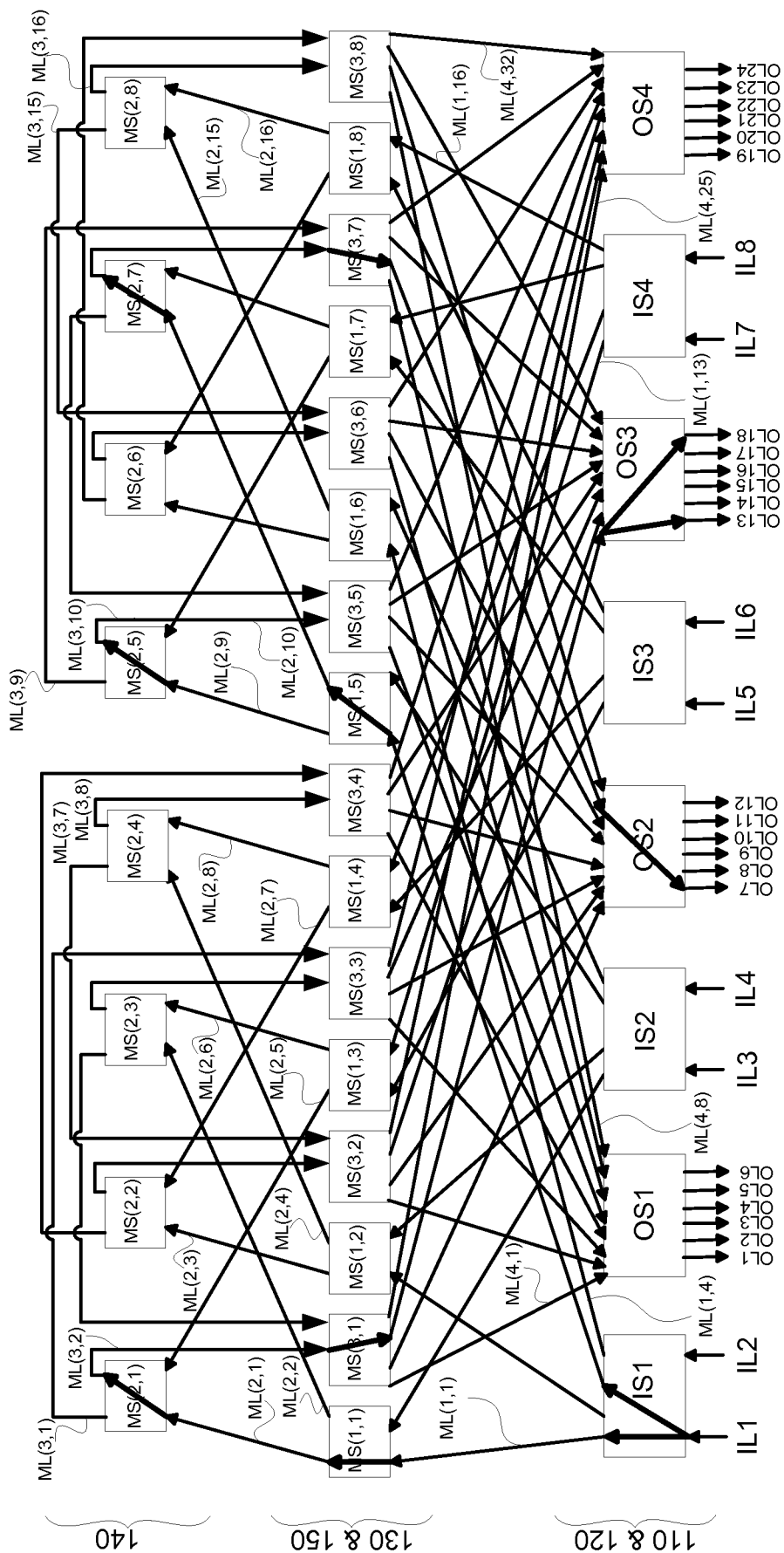
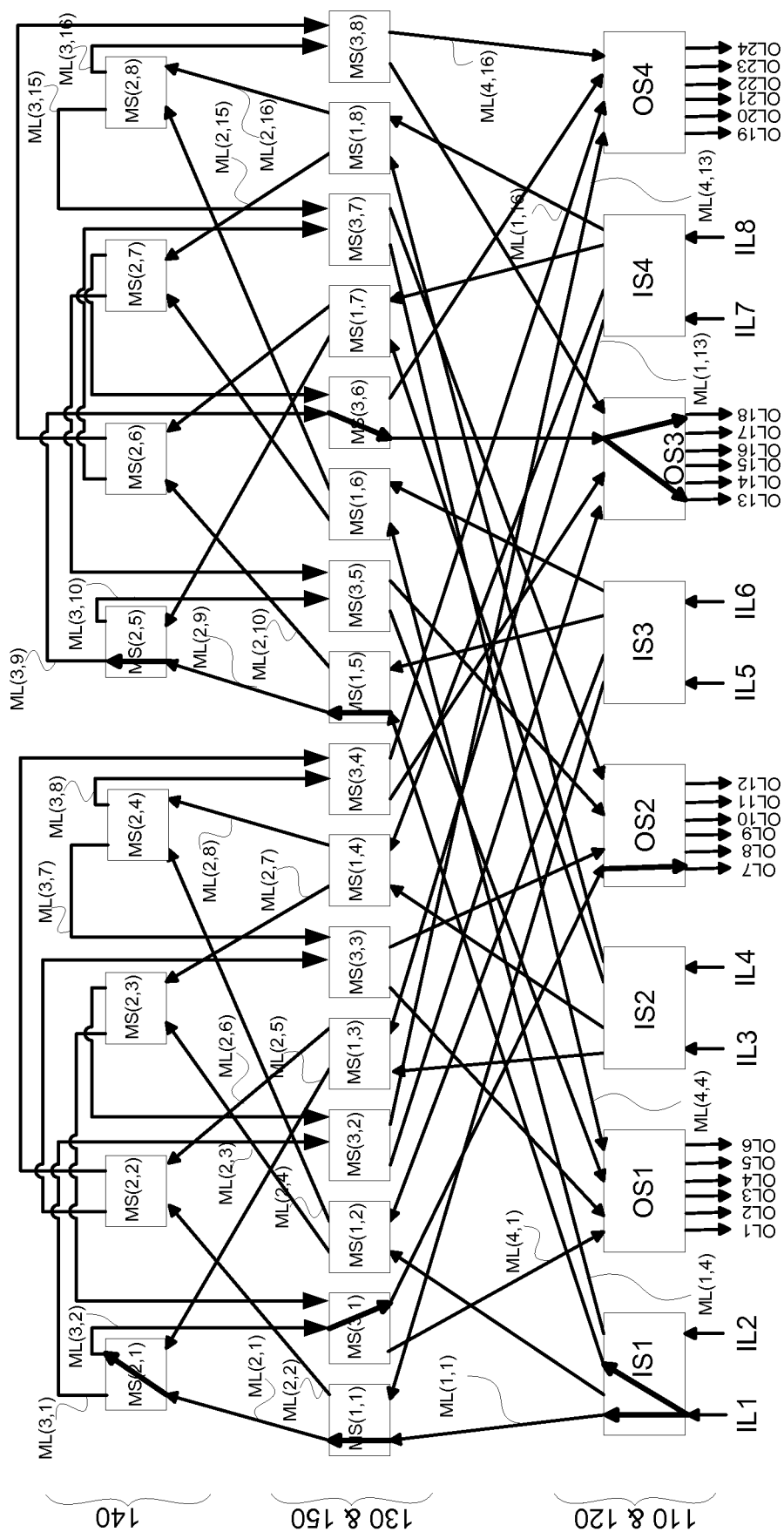


FIG. 4C1

400C1



140

130 & 150

110 & 120

400C2

FIG. 4C2

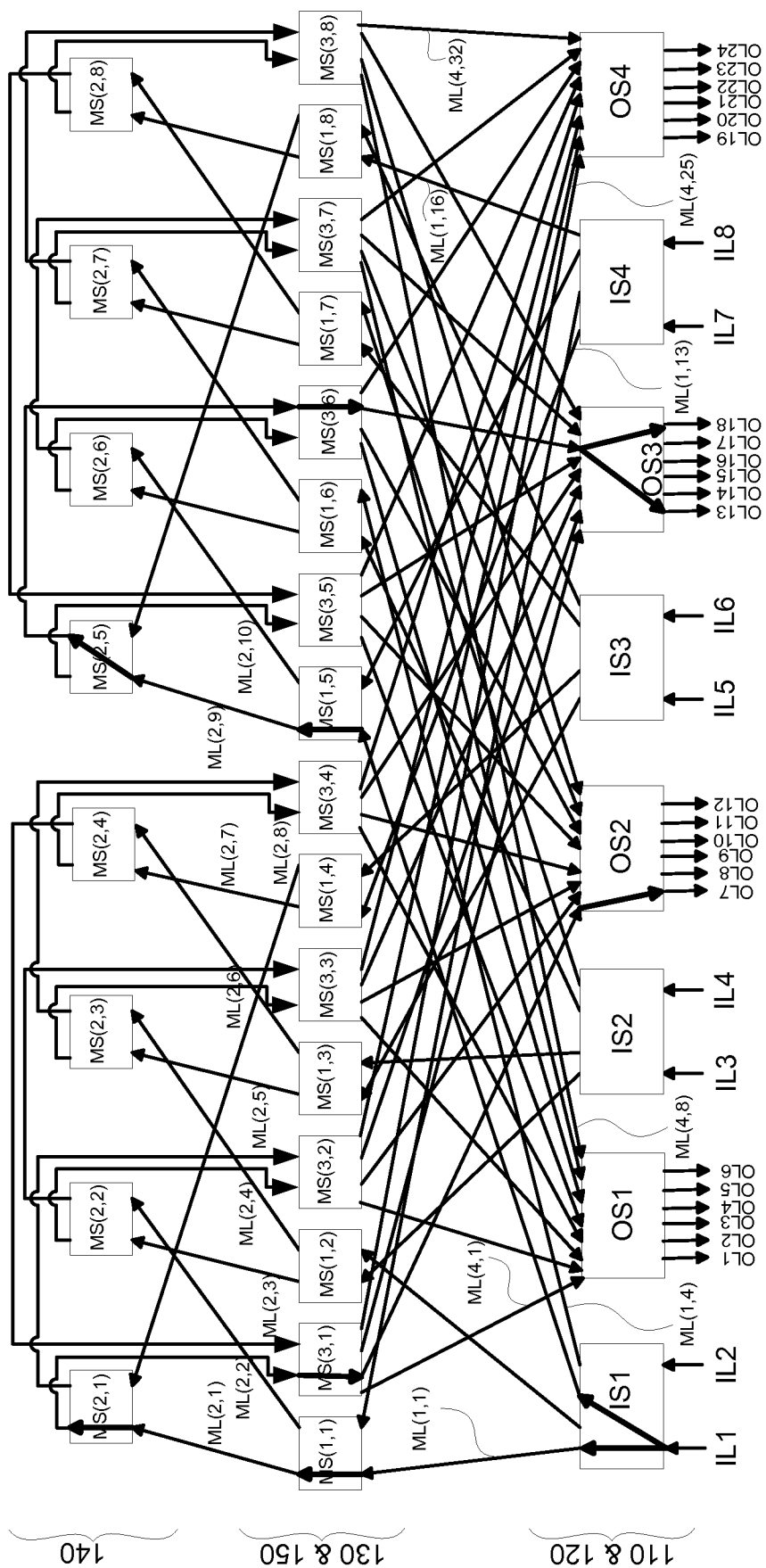
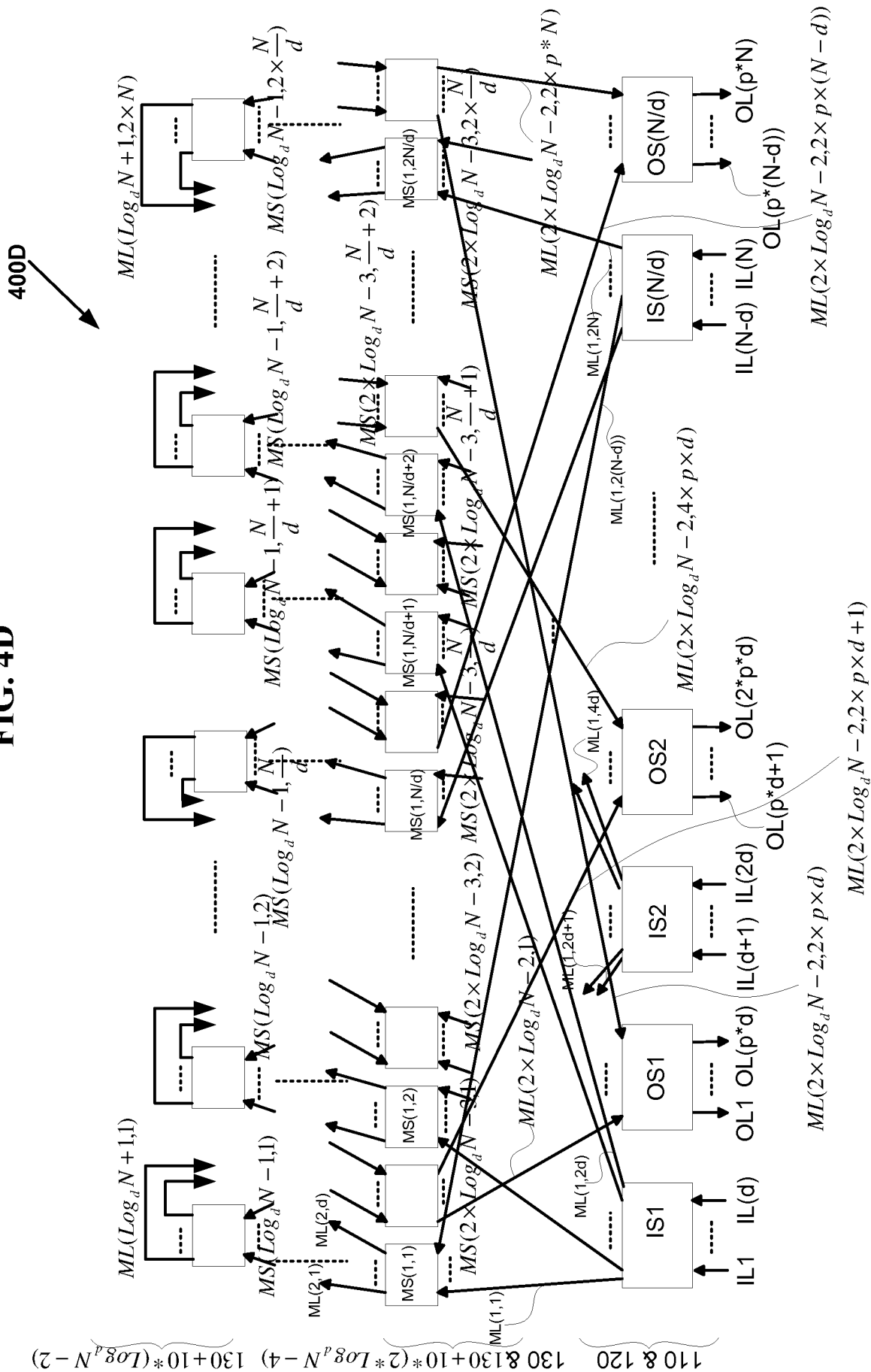


FIG. 4D



110 & 120
130 & 130 + 10 * (2 * Log_d N - 4)
130 + 10 * (Log_d N - 2)

FIG. 4E

400E

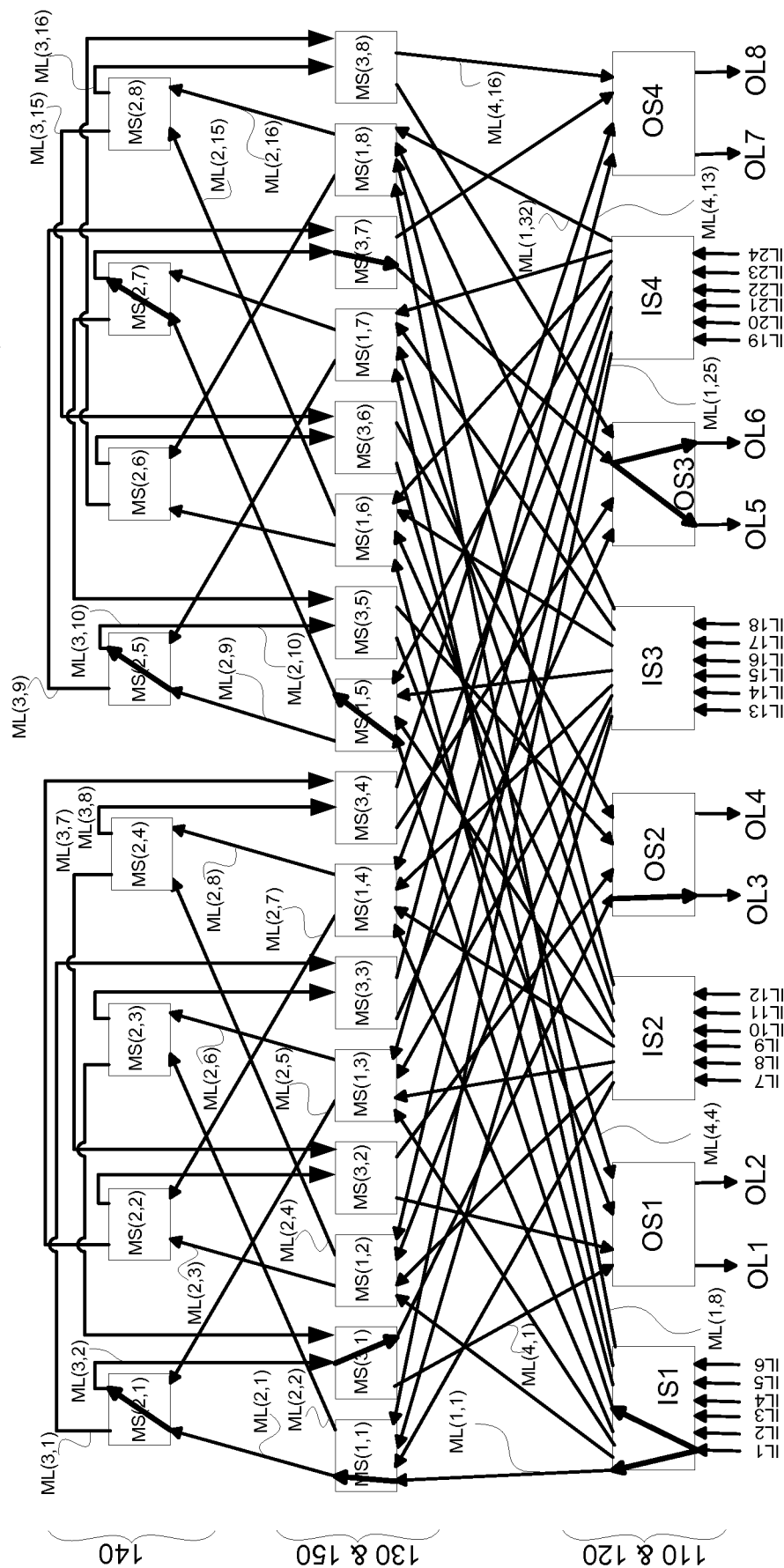


FIG. 4E1

400E1

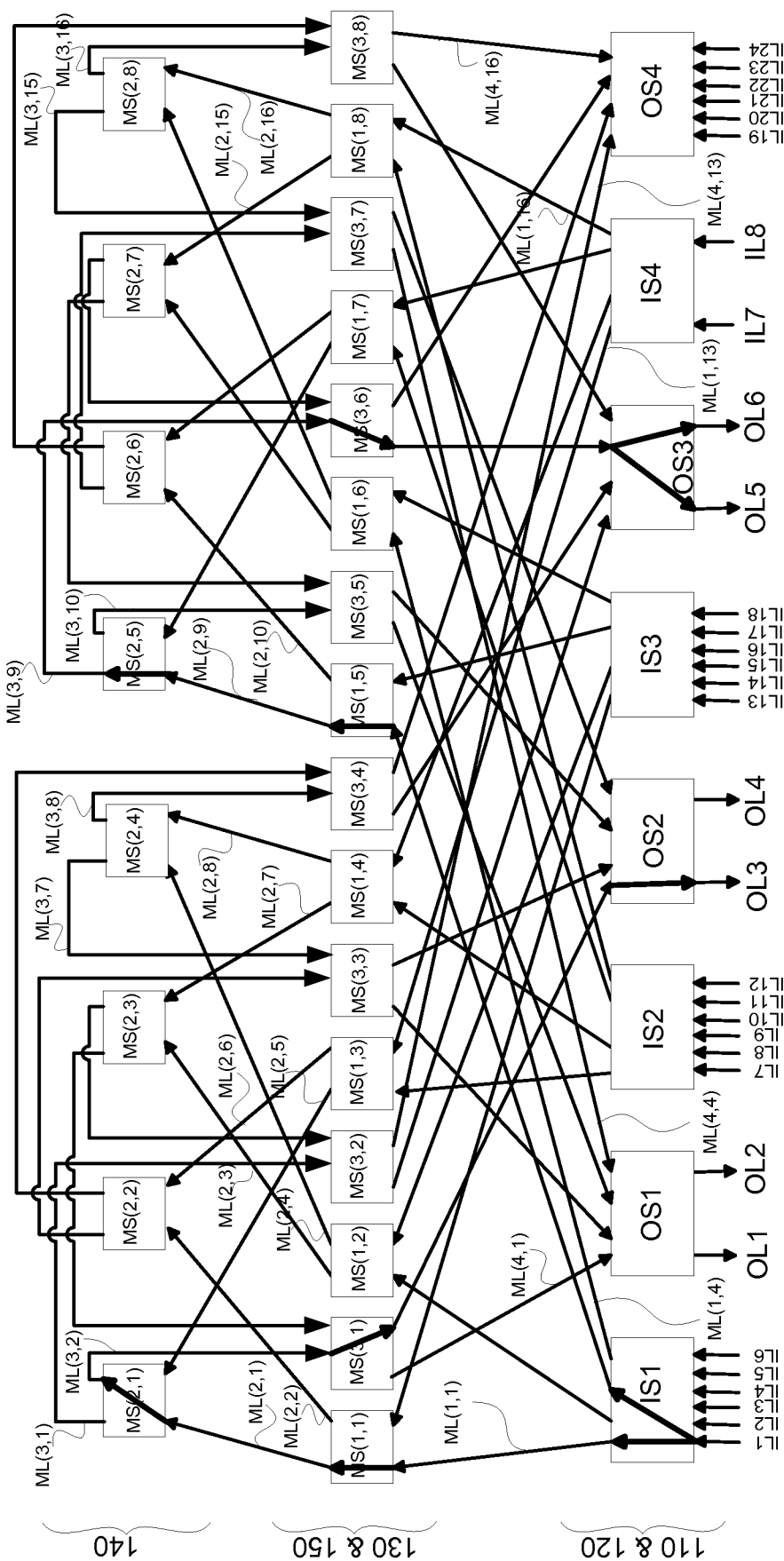
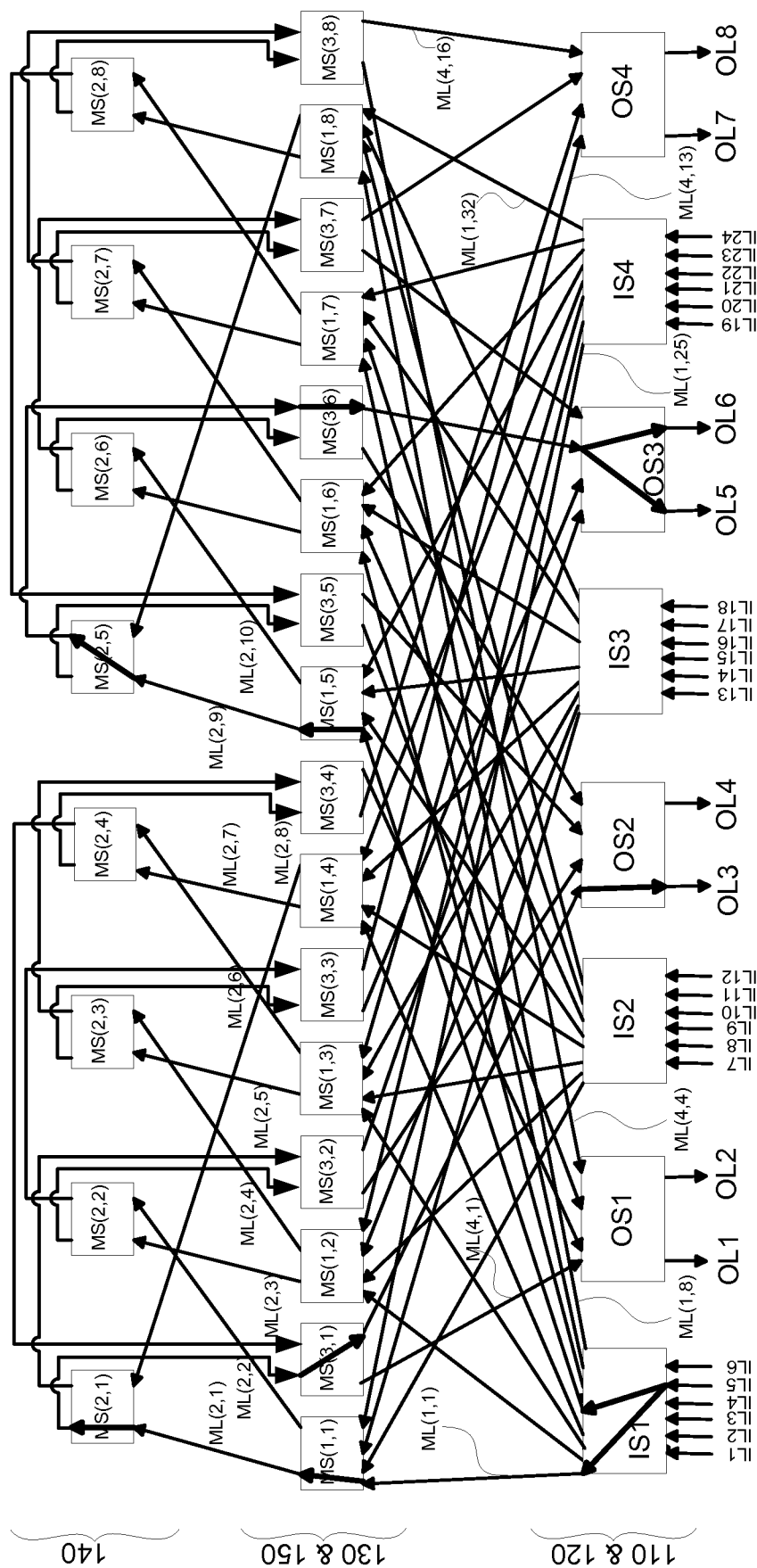


FIG. 4E2

400E2

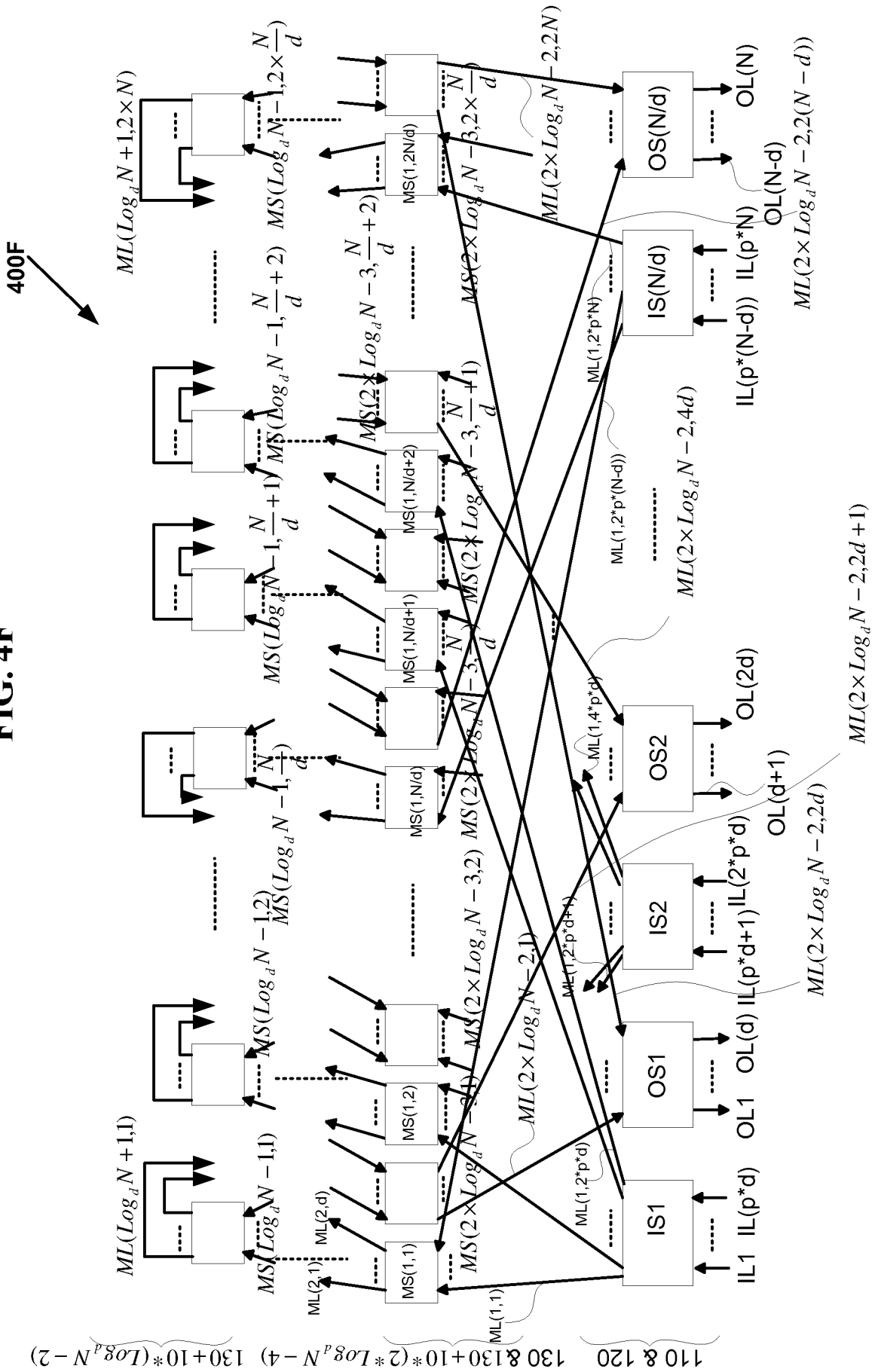


140

130 & 150

110 & 120

FIG. 4F



110 & 120 130 & 130+10* (2*Log_d N - 4) 130+10*(Log_d N - 2)

500A

FIG. 5A

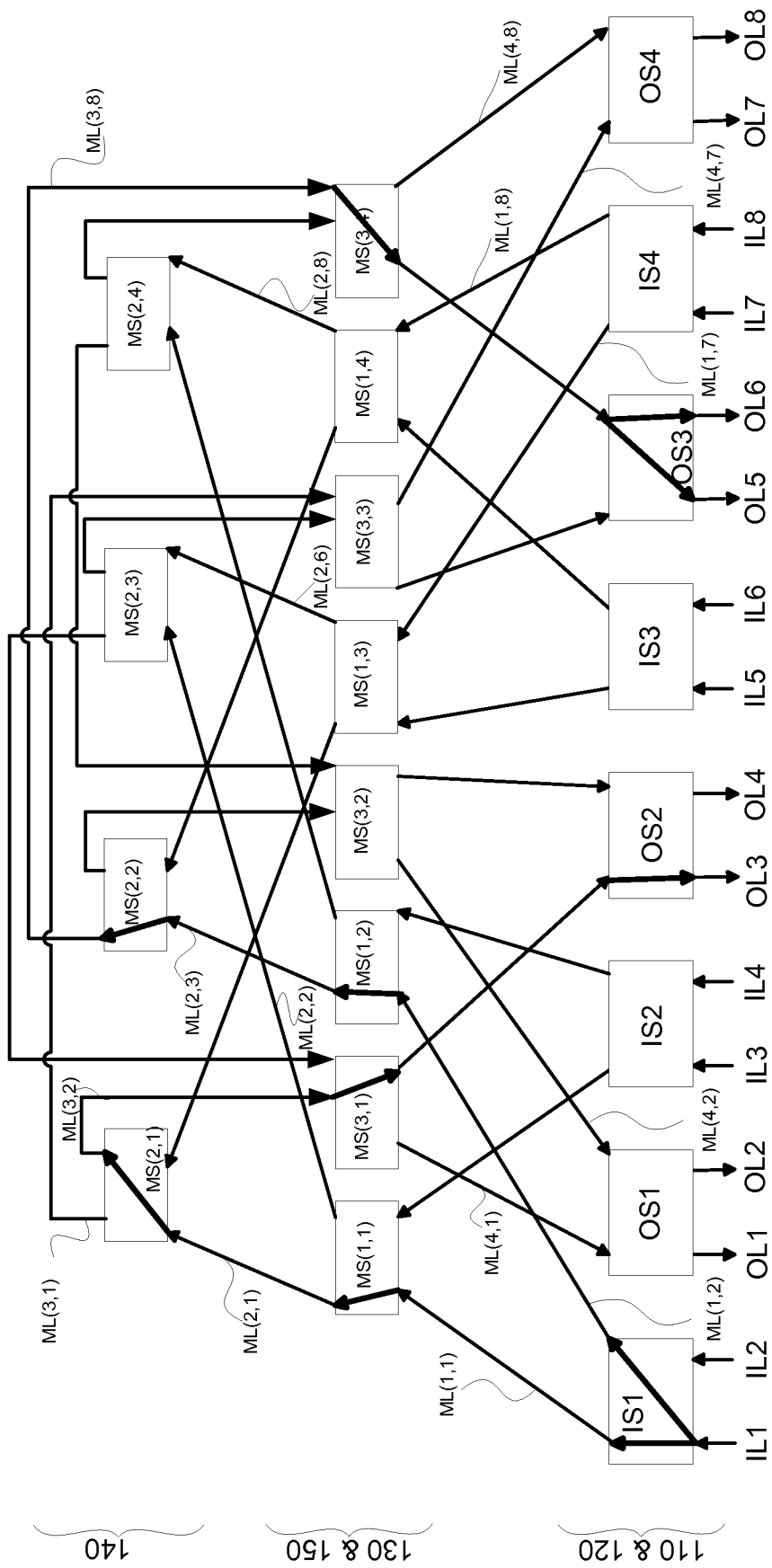
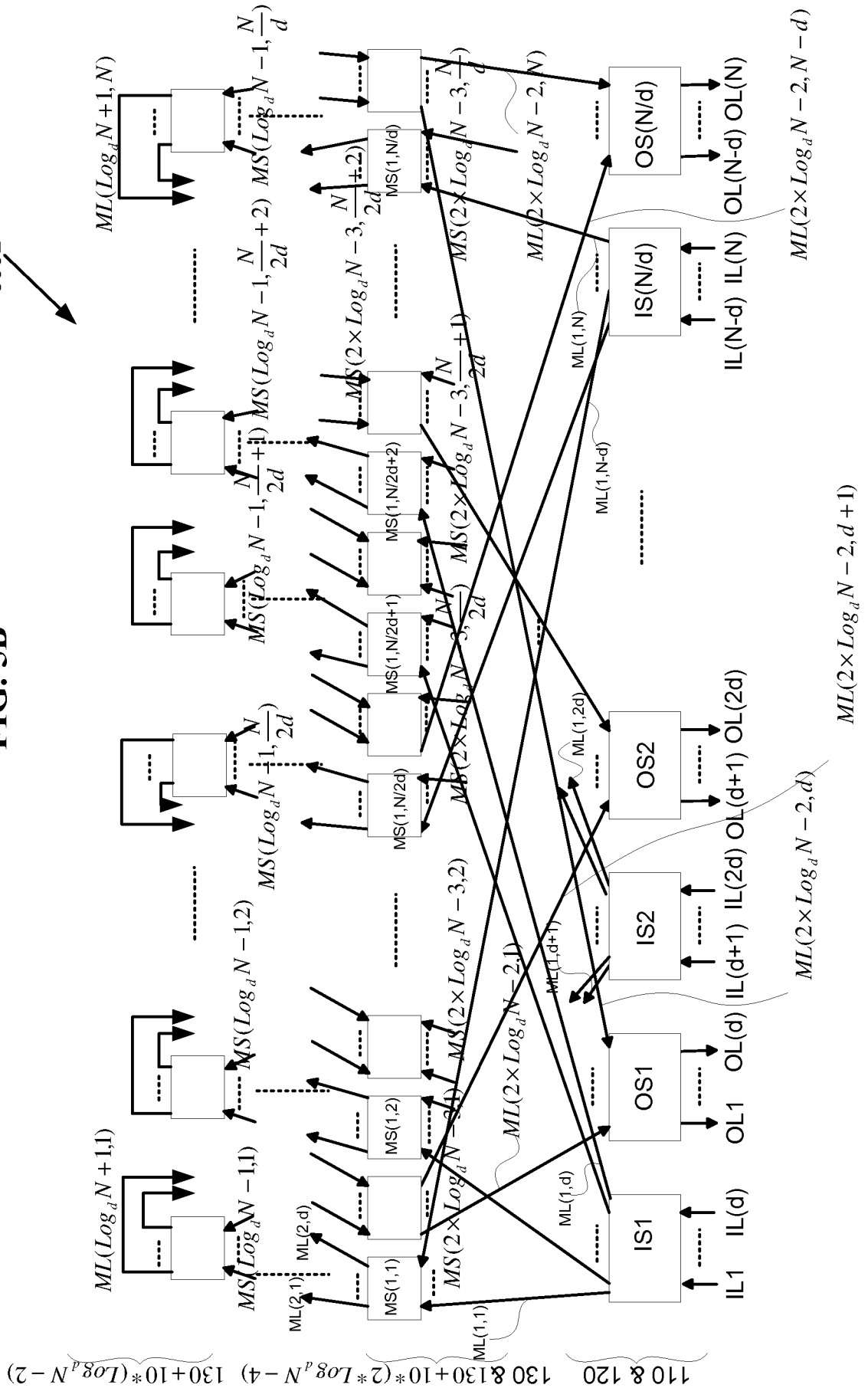


FIG. 5B

500B



500C

FIG. 5C

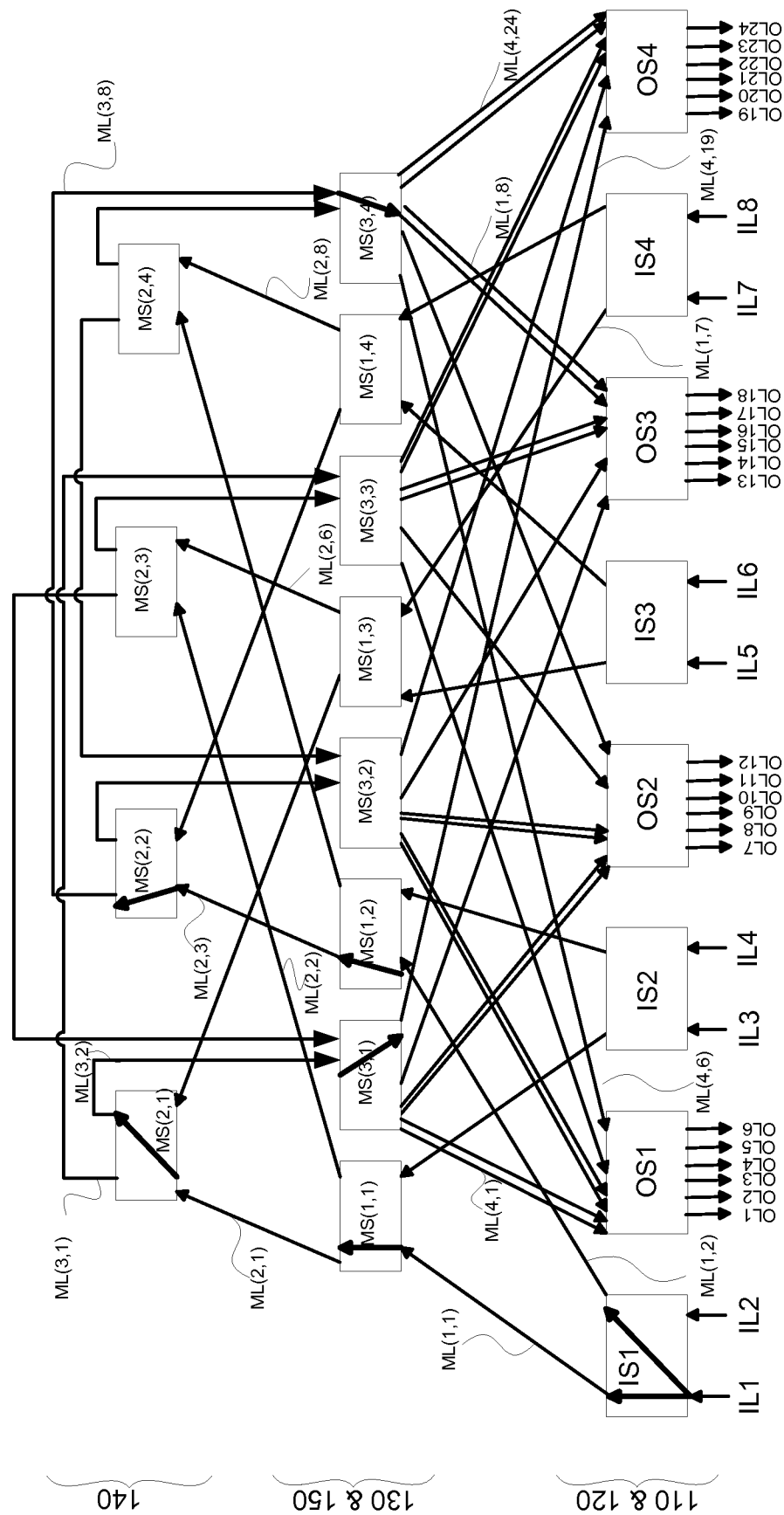


FIG. 5D

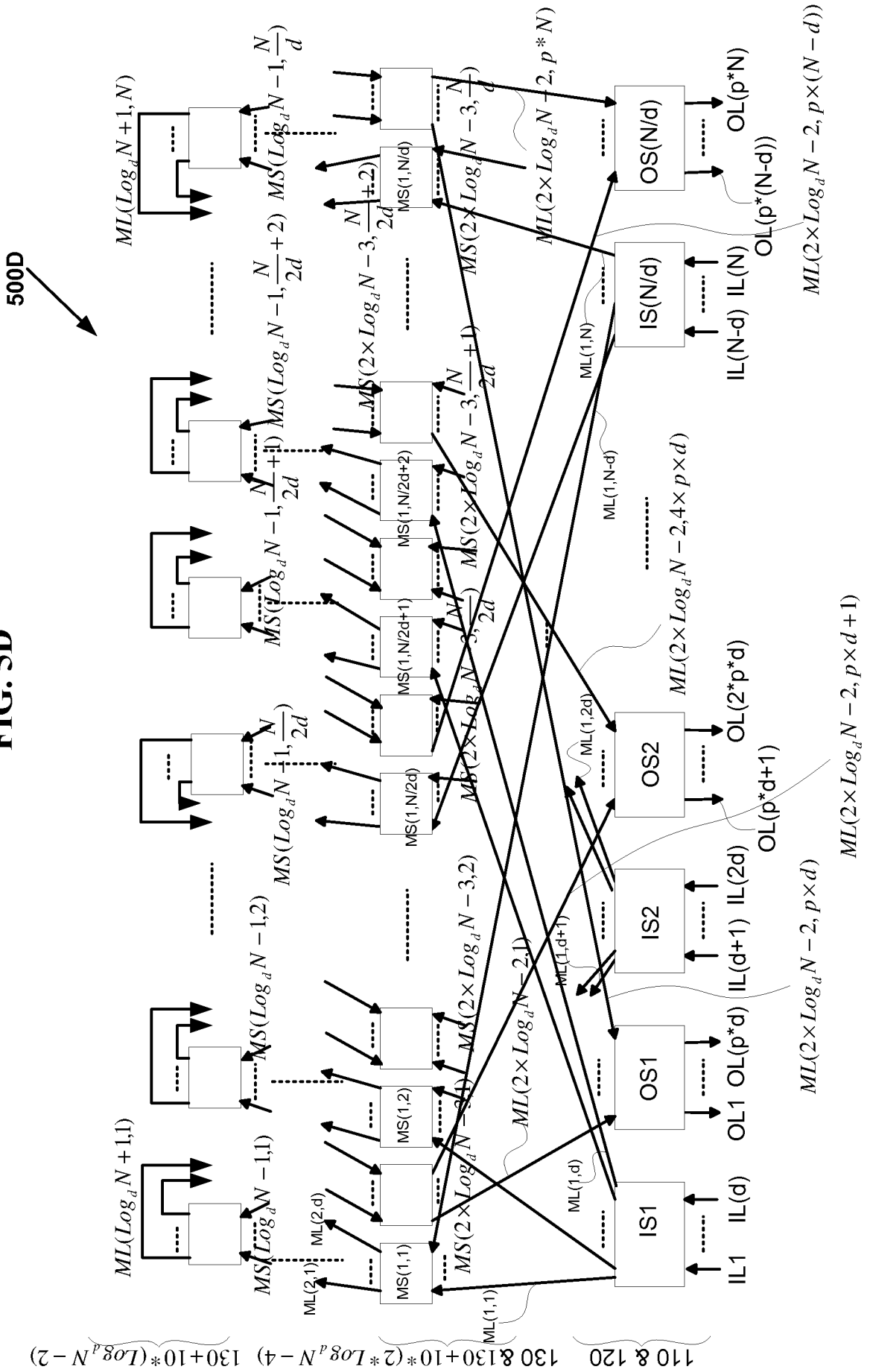


FIG. 5E

500E

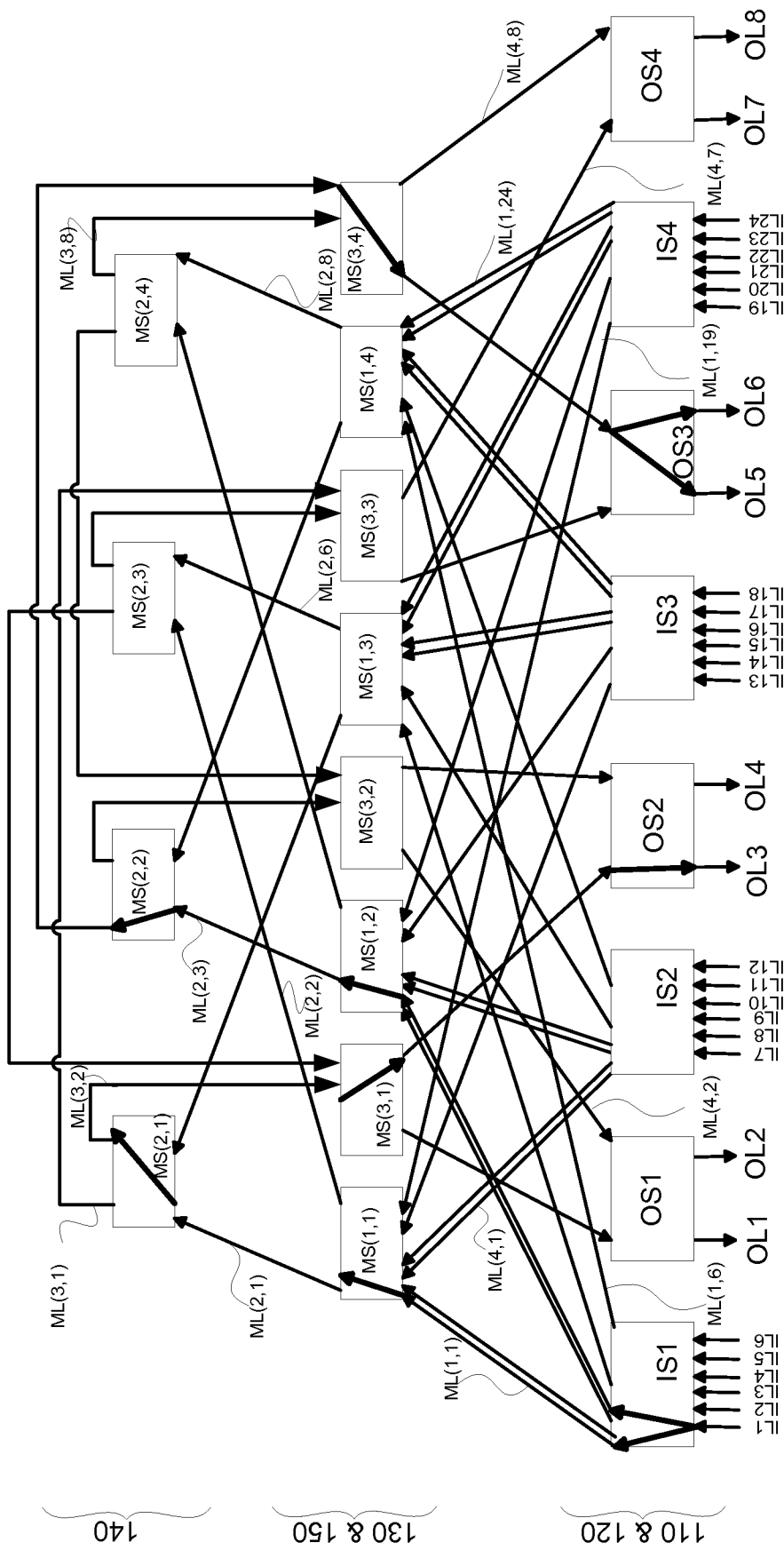


FIG. 5F

500F

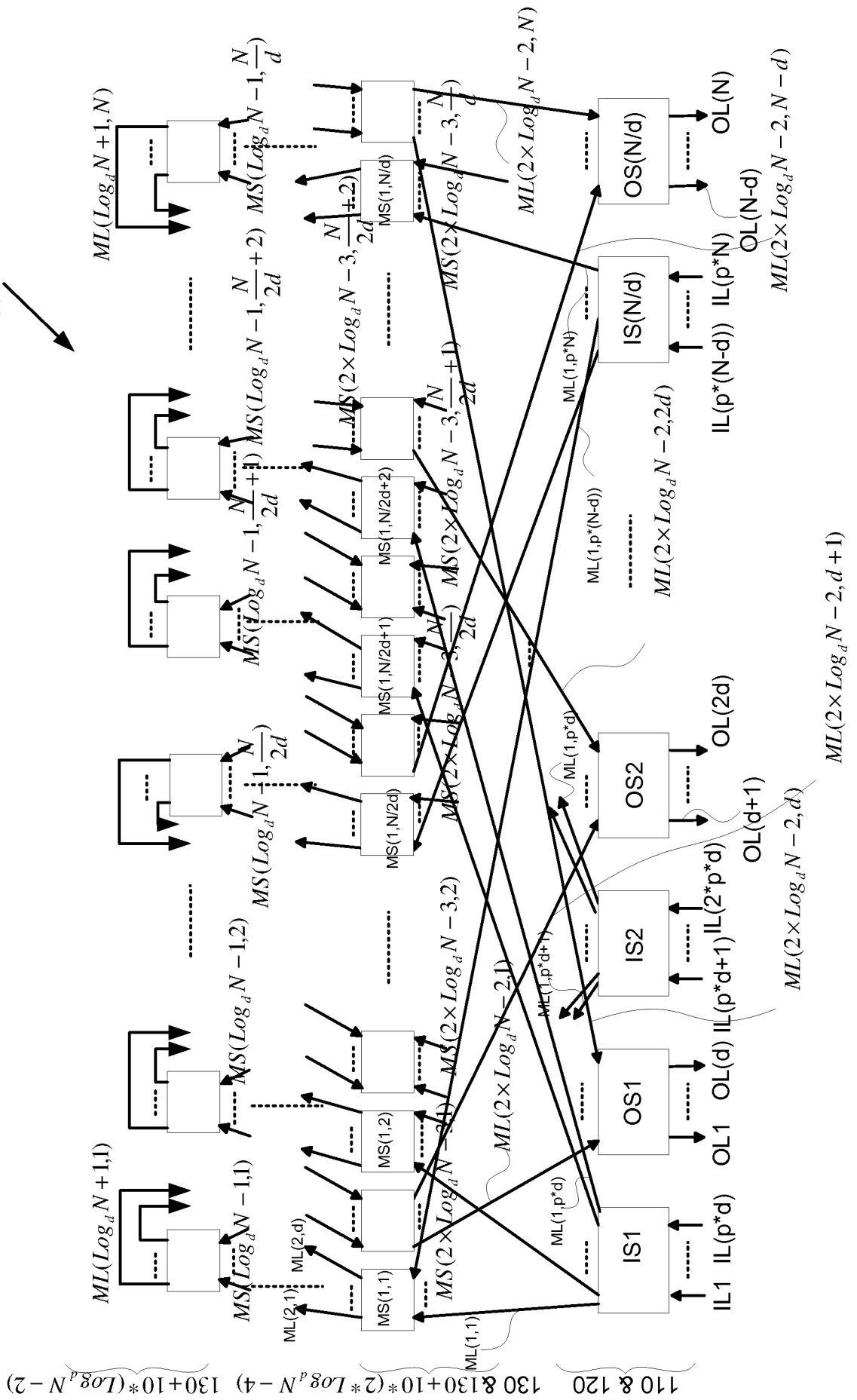


FIG. 6A

600A

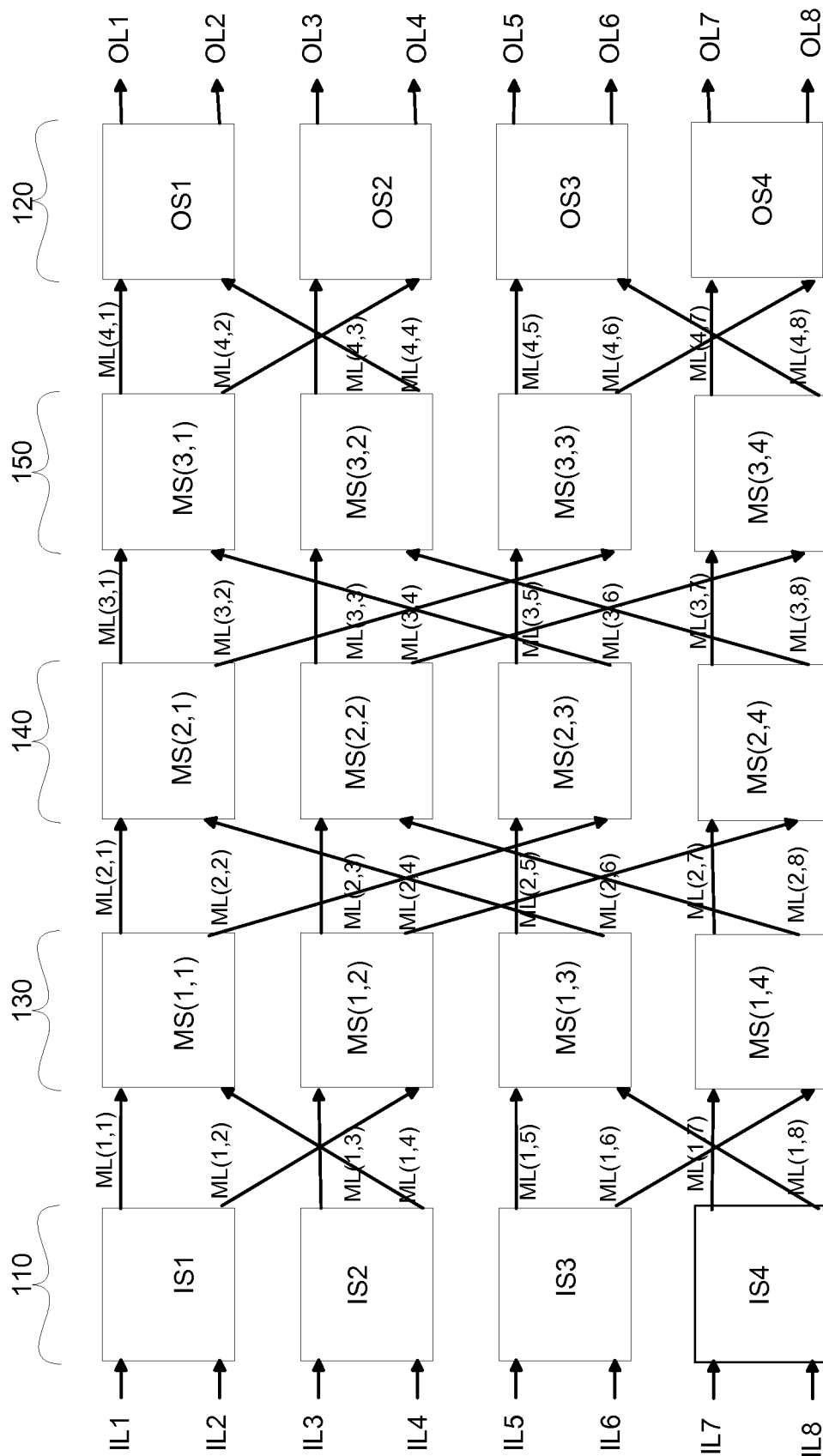


FIG. 6B

600B

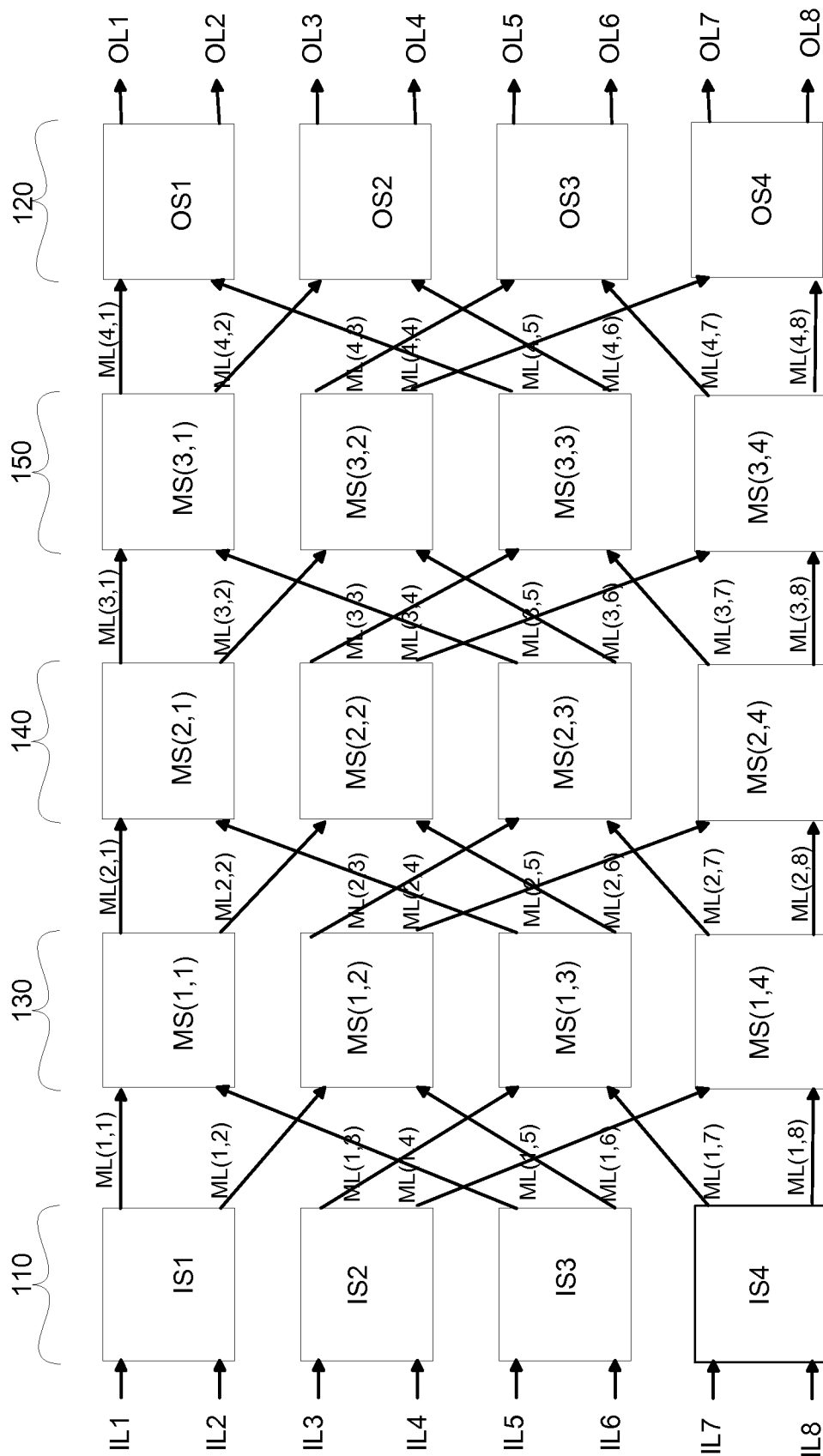


FIG. 6C

600C

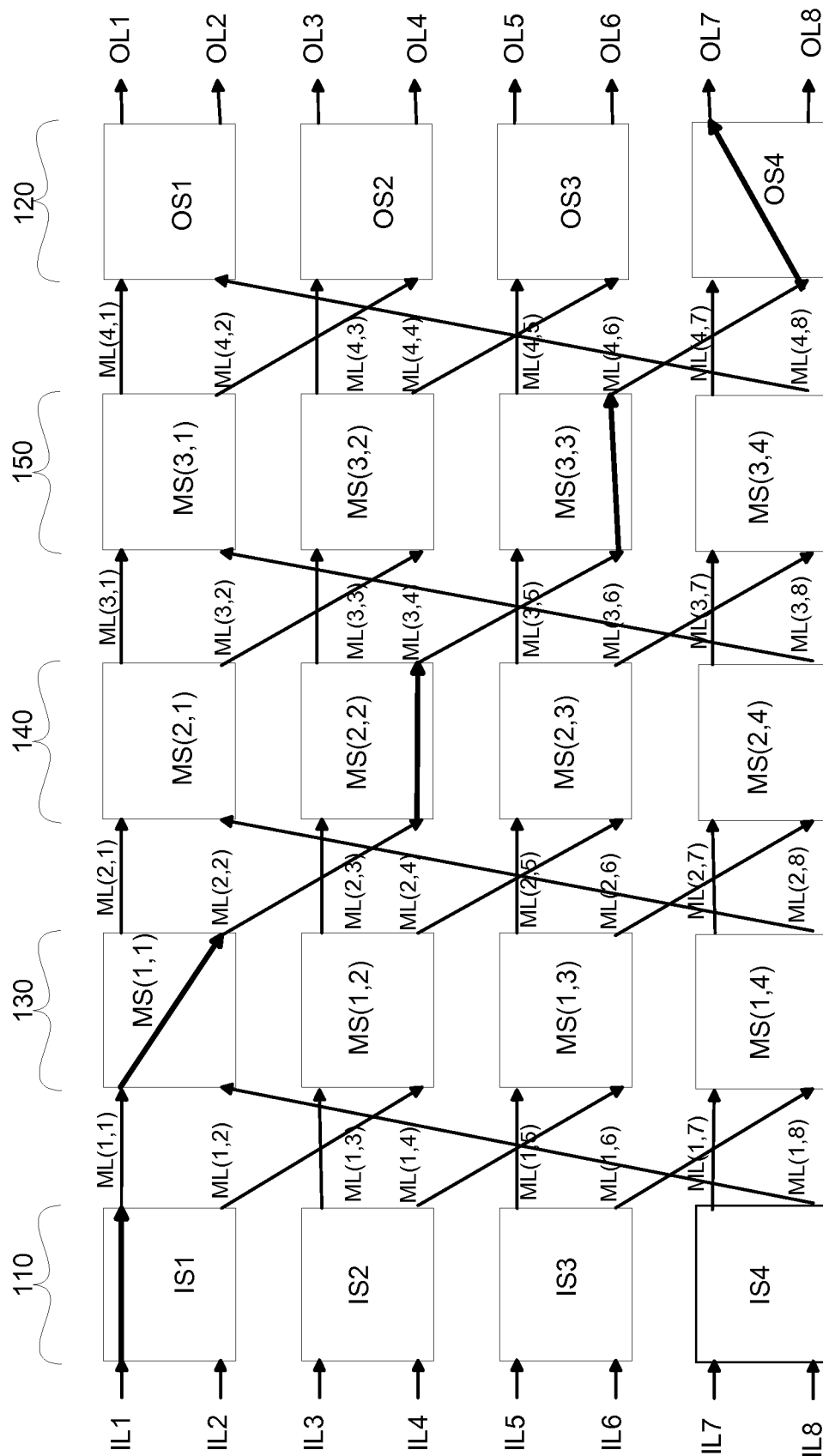


FIG. 6D

600D

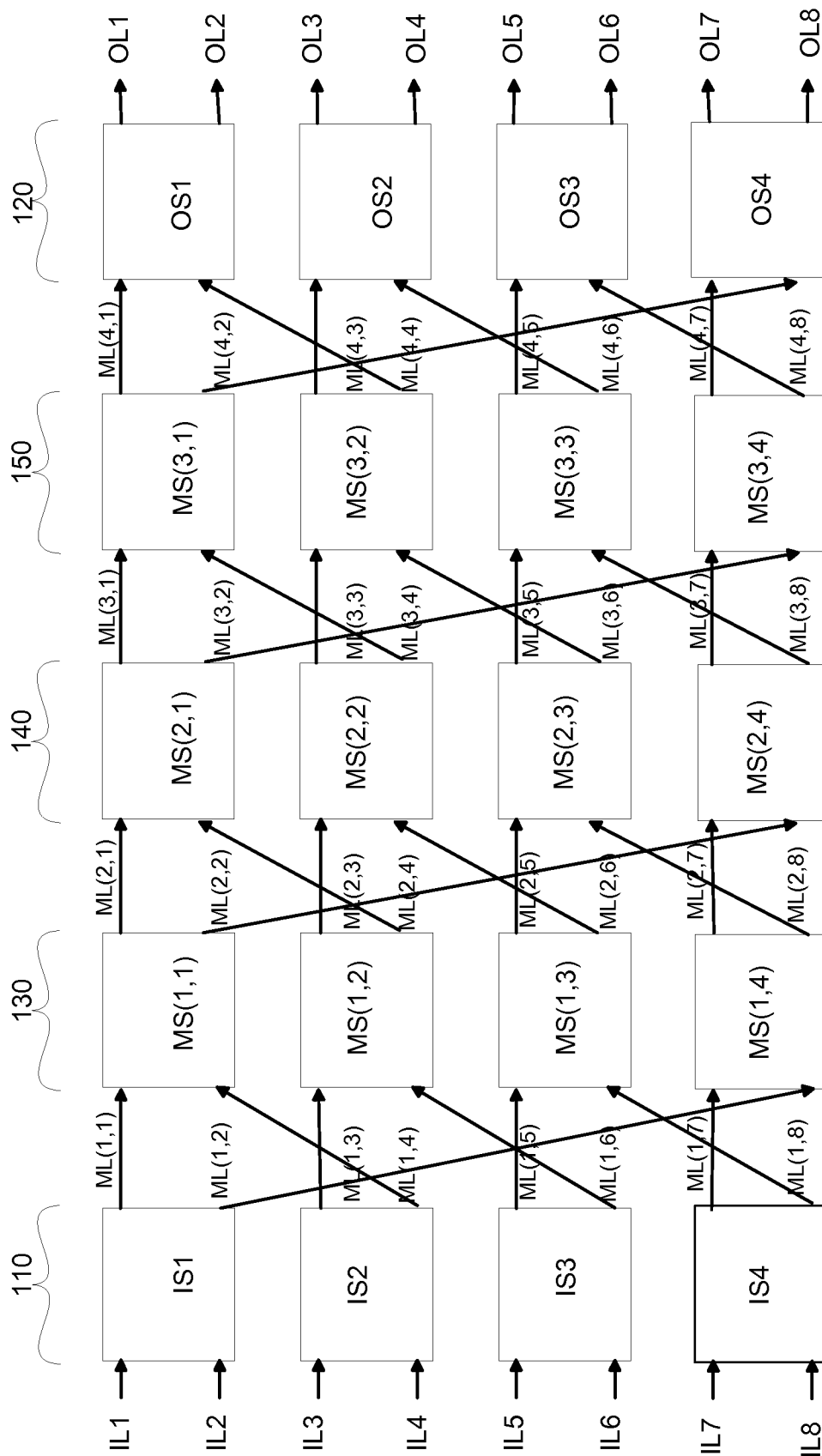


FIG. 6E

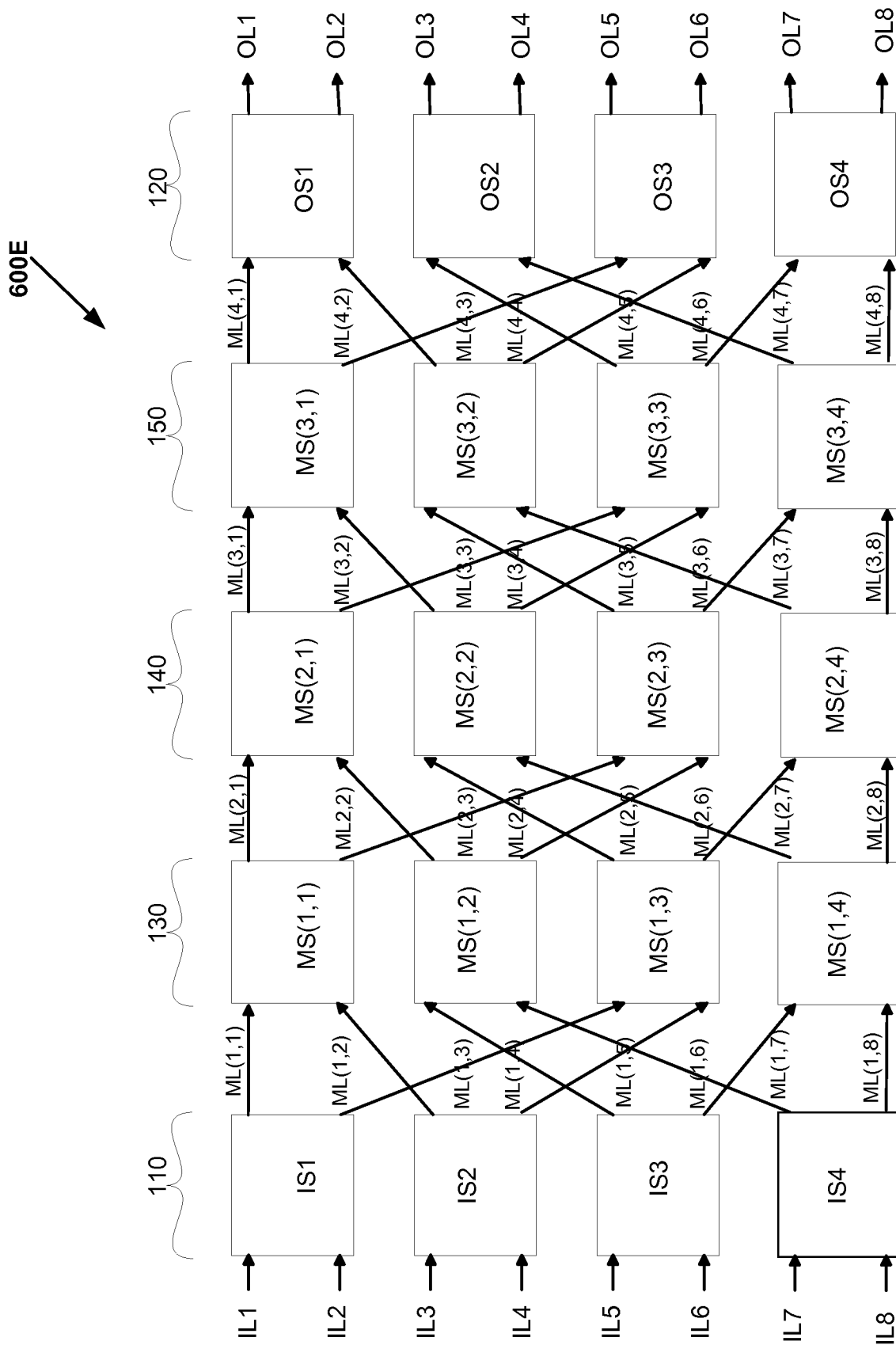


FIG. 6F

600F

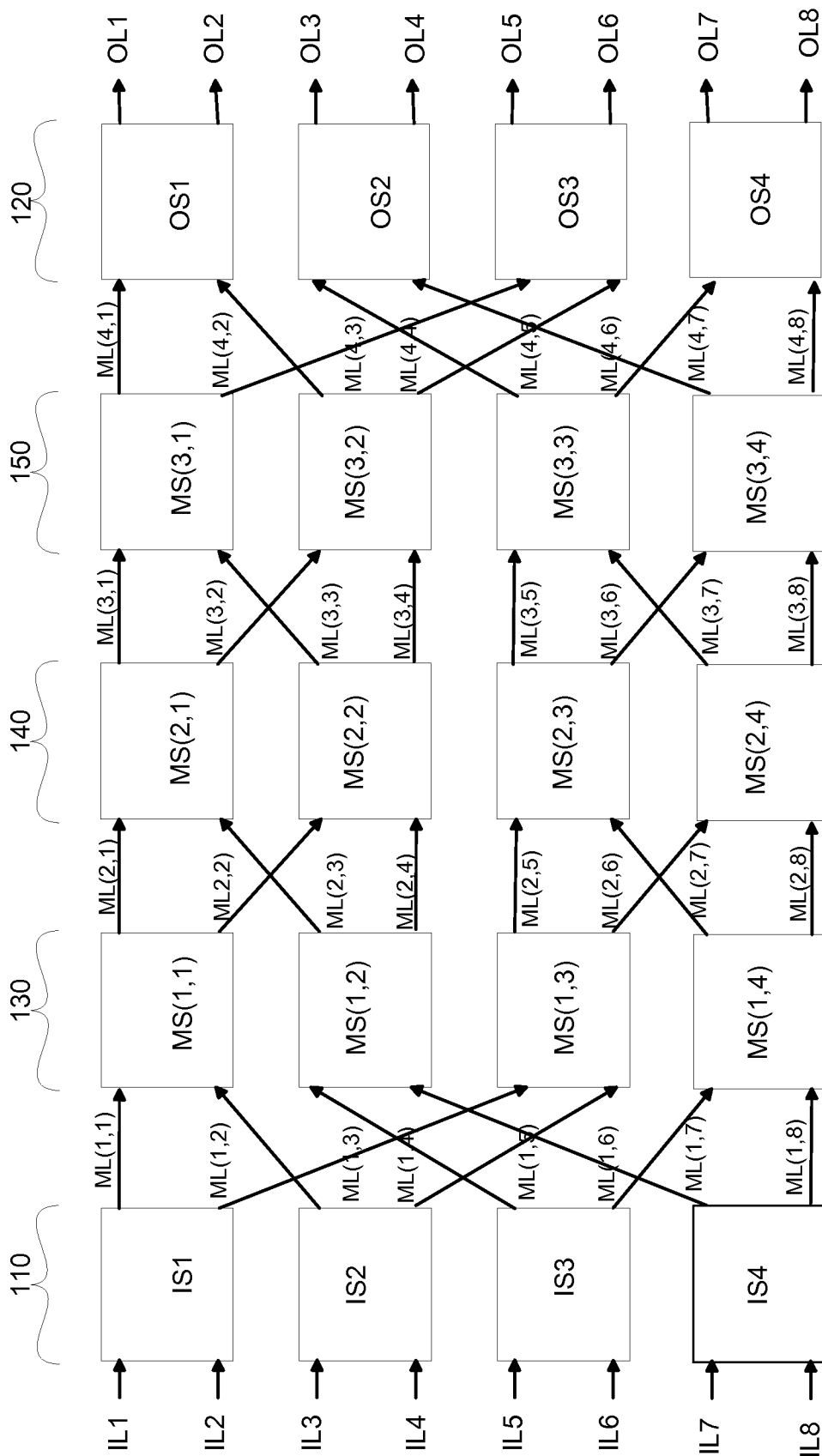


FIG. 6G

600G

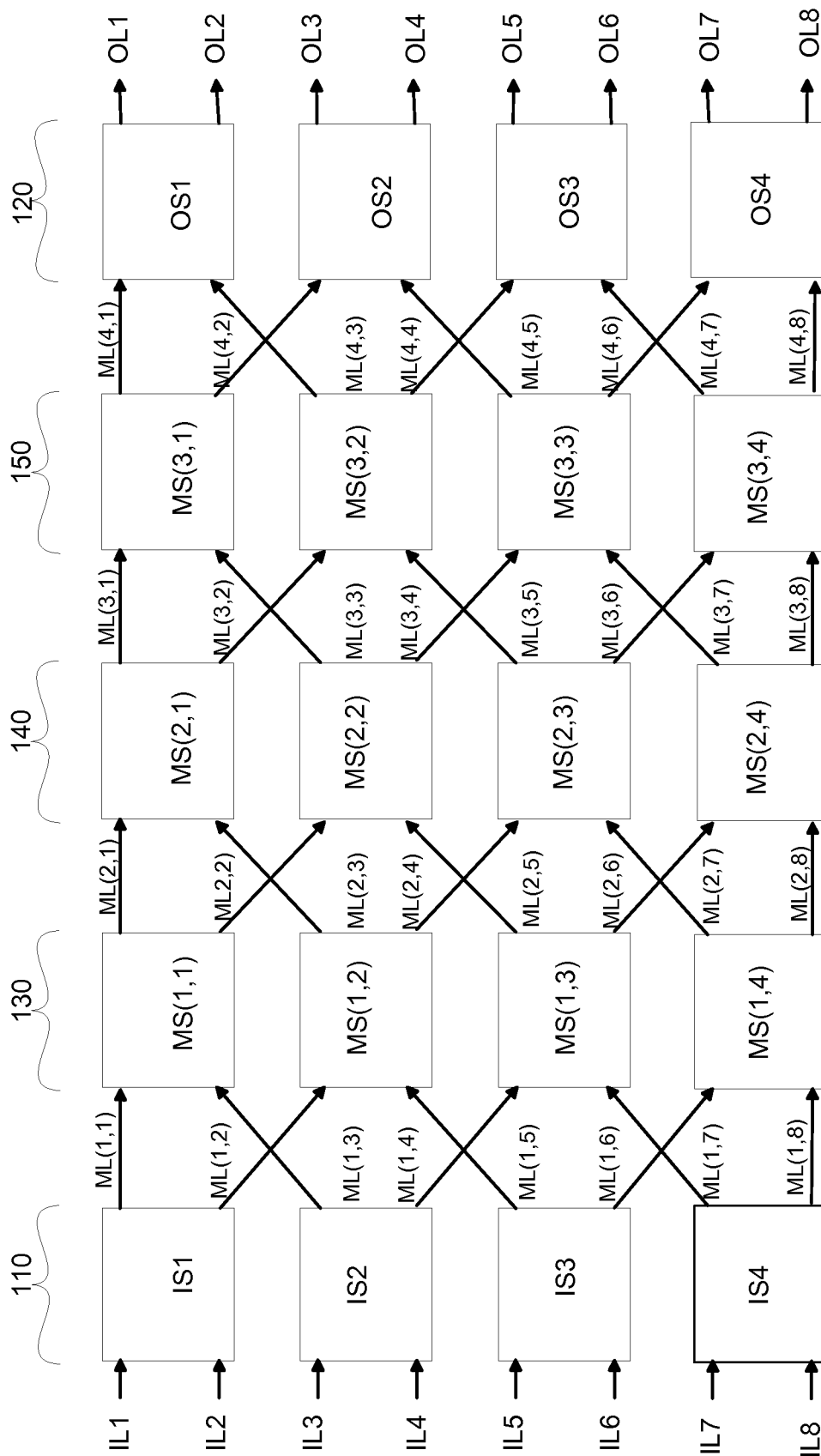


FIG. 6H

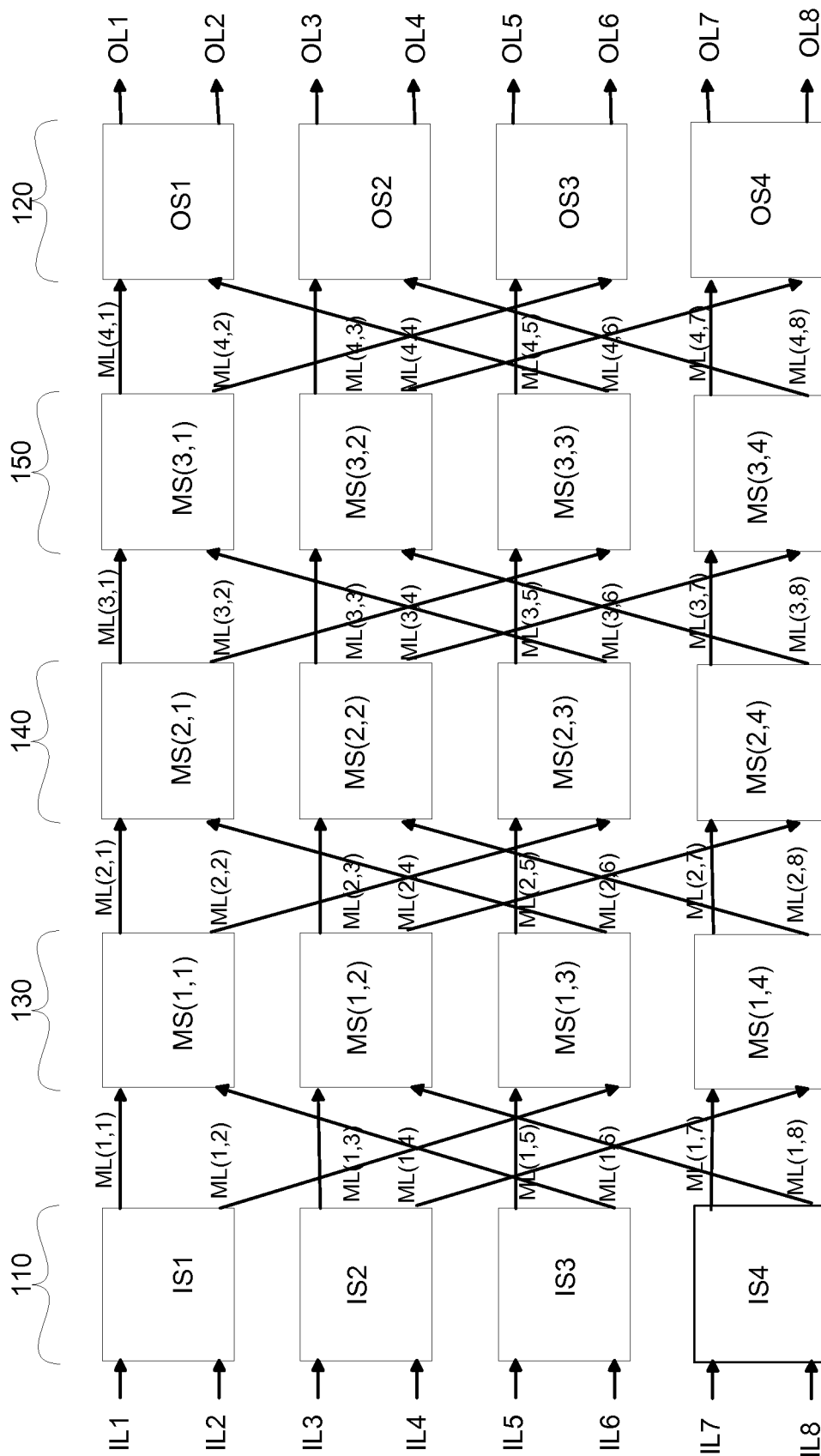


FIG. 6I

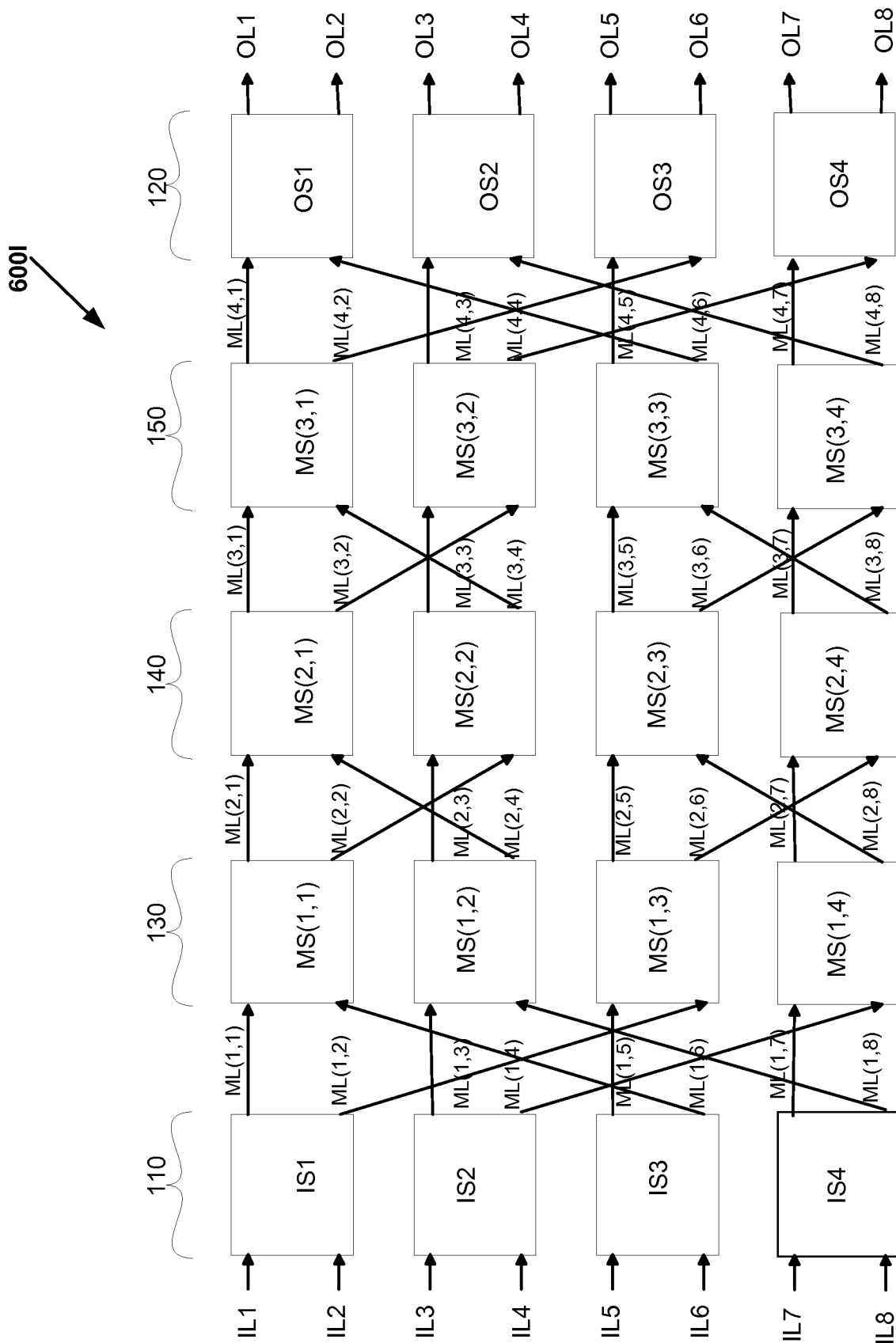


FIG. 6J

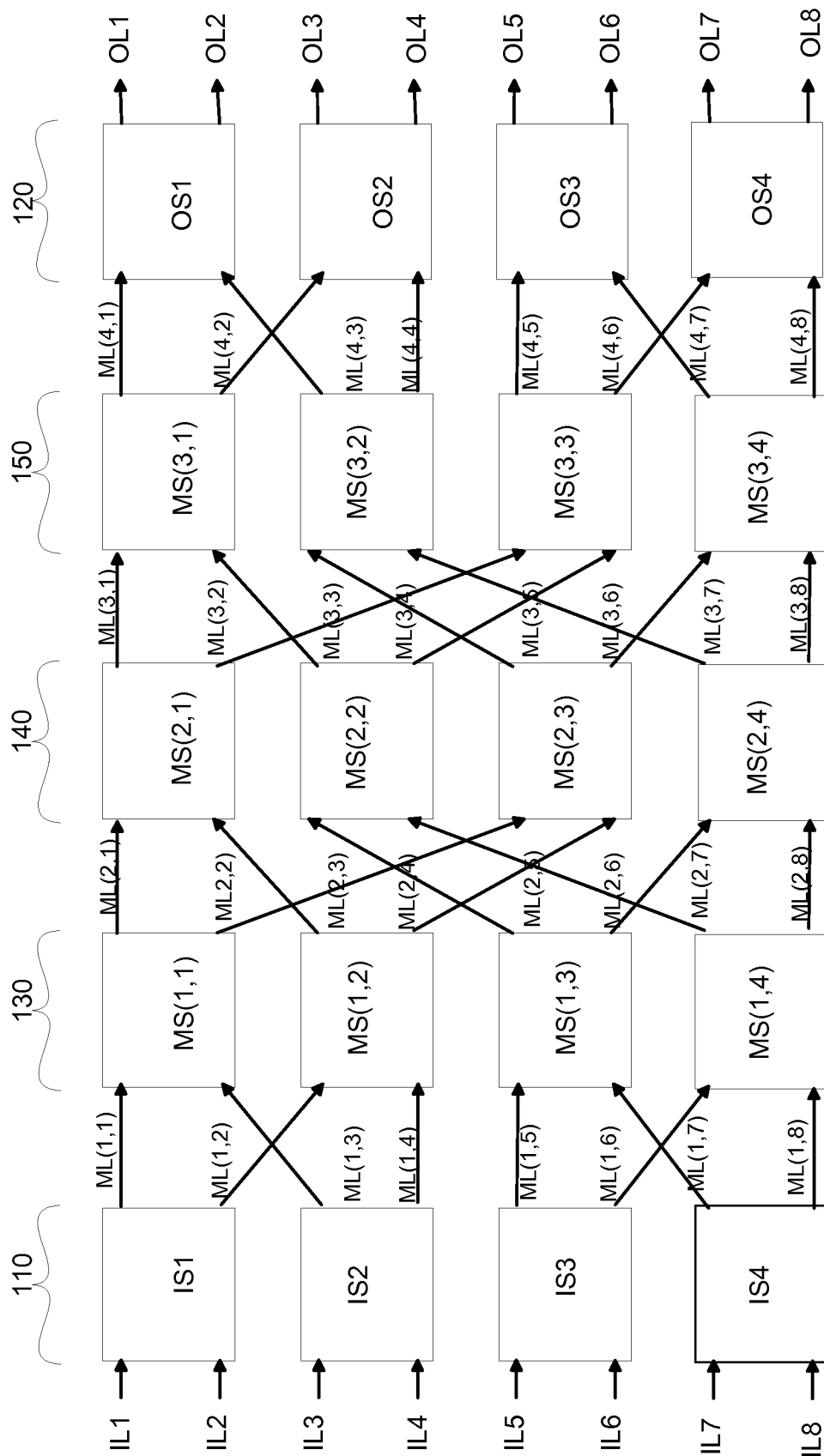


FIG. 6K

600K

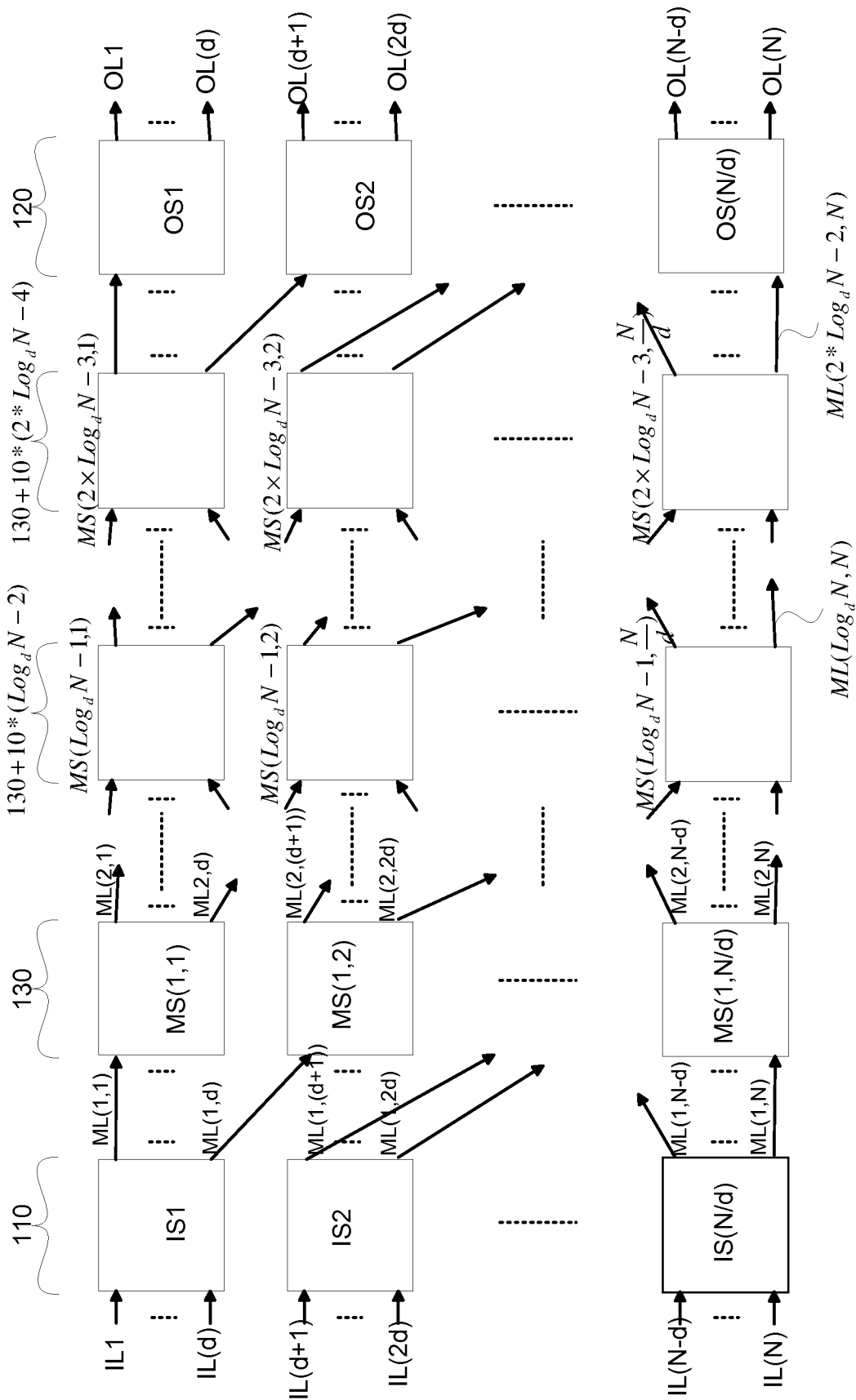


FIG. 6A1

600A1

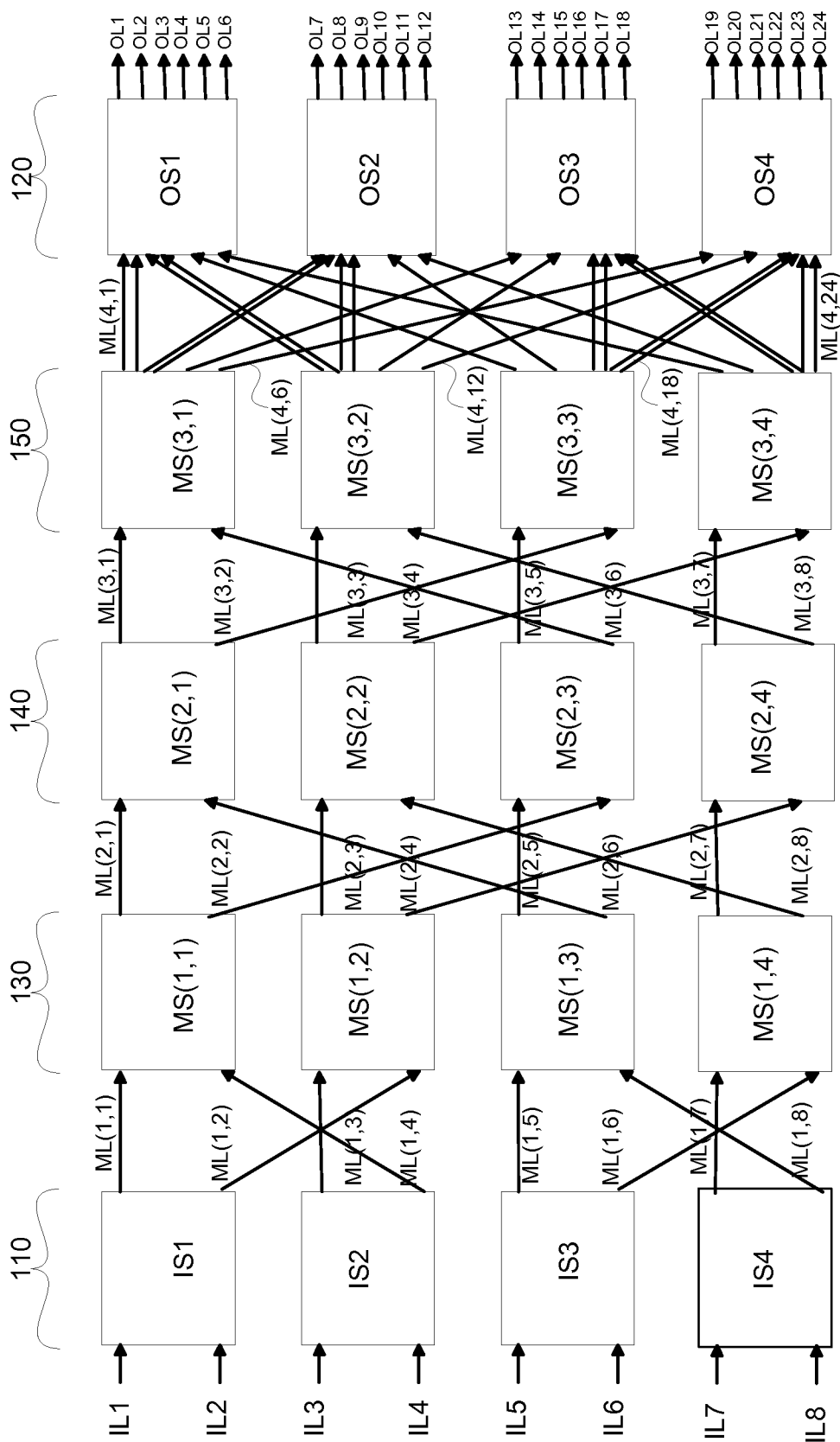


FIG. 6B1

600B1

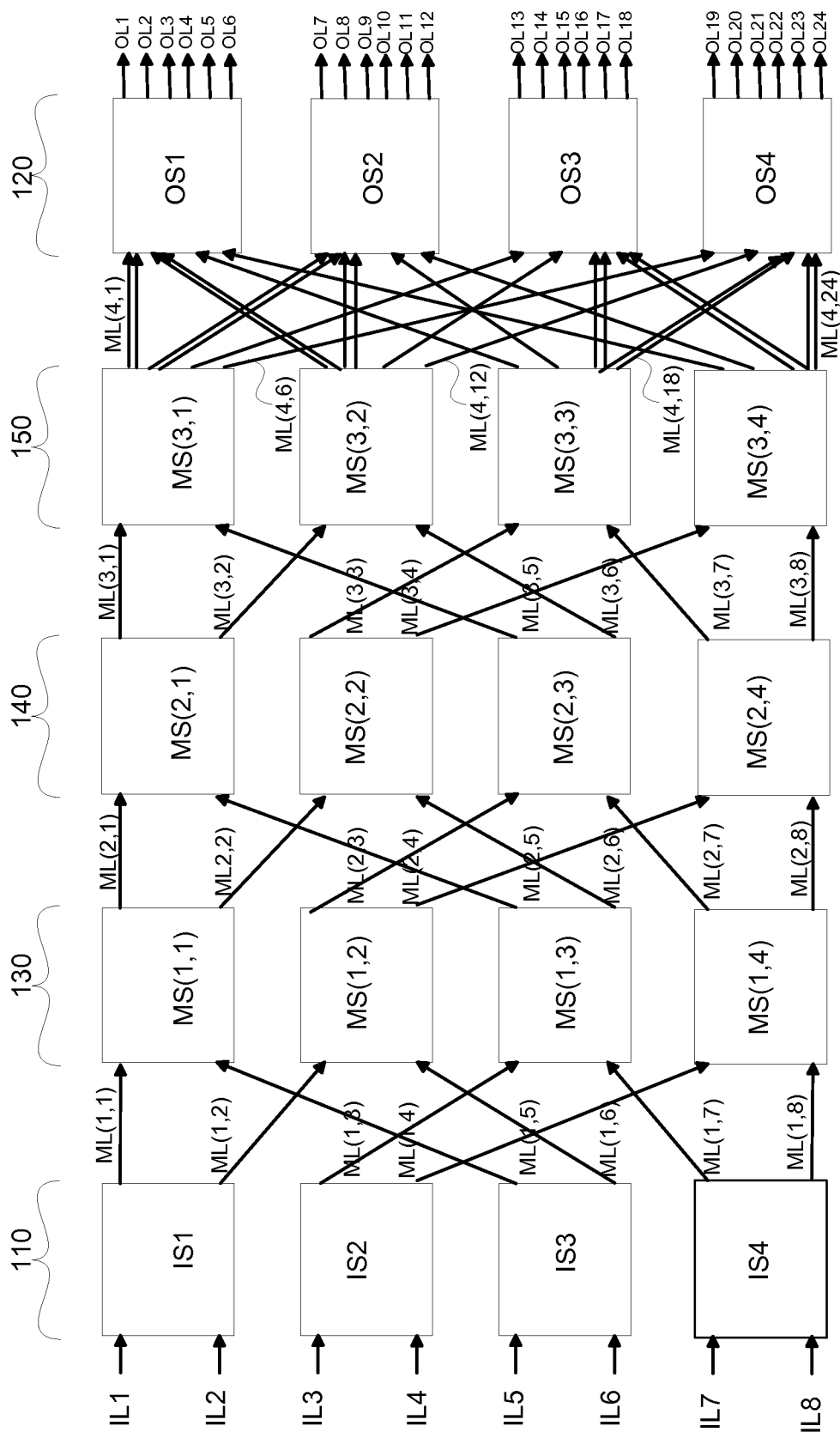


FIG. 6C1

600C1

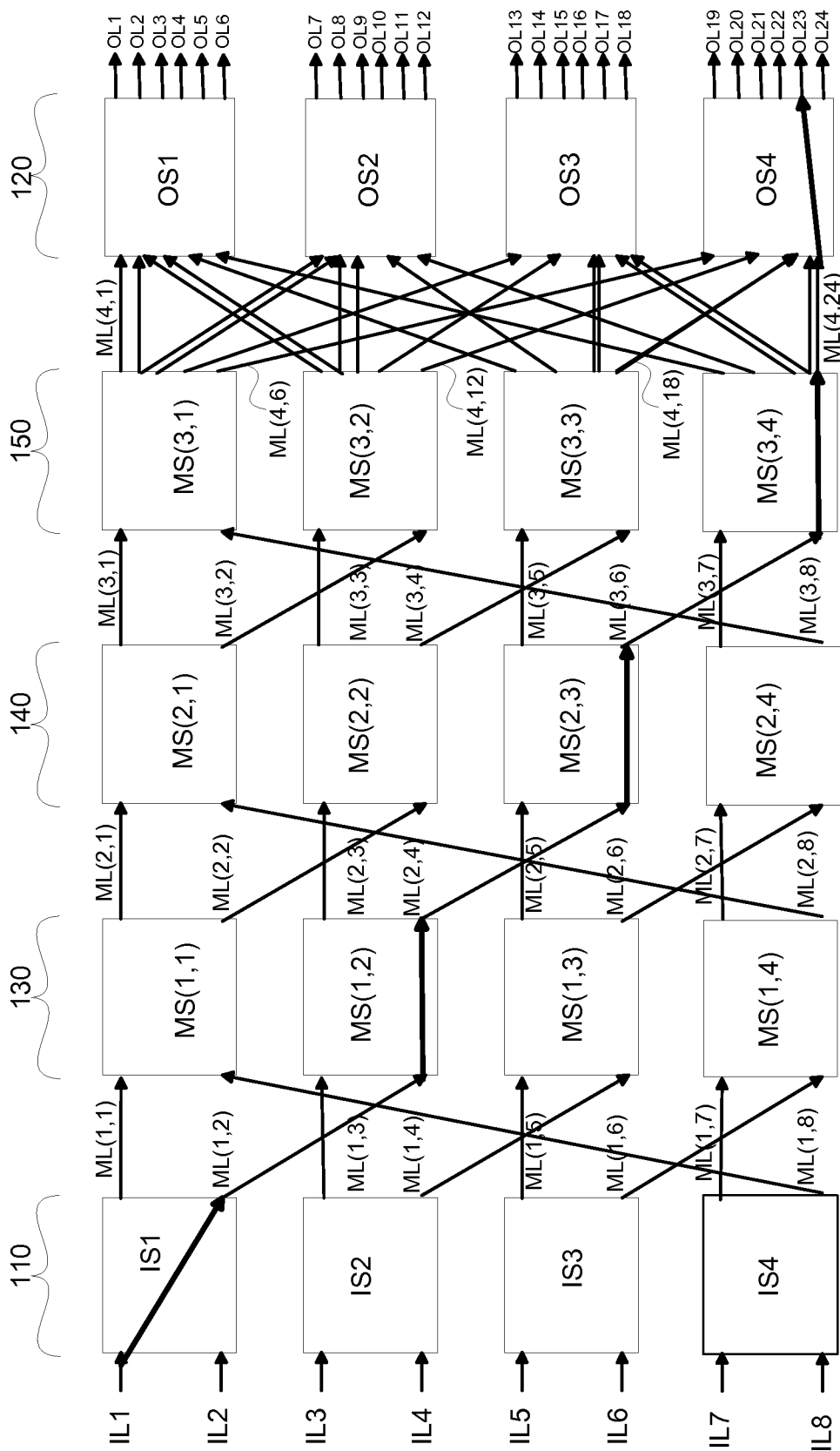


FIG. 6D1

600D1

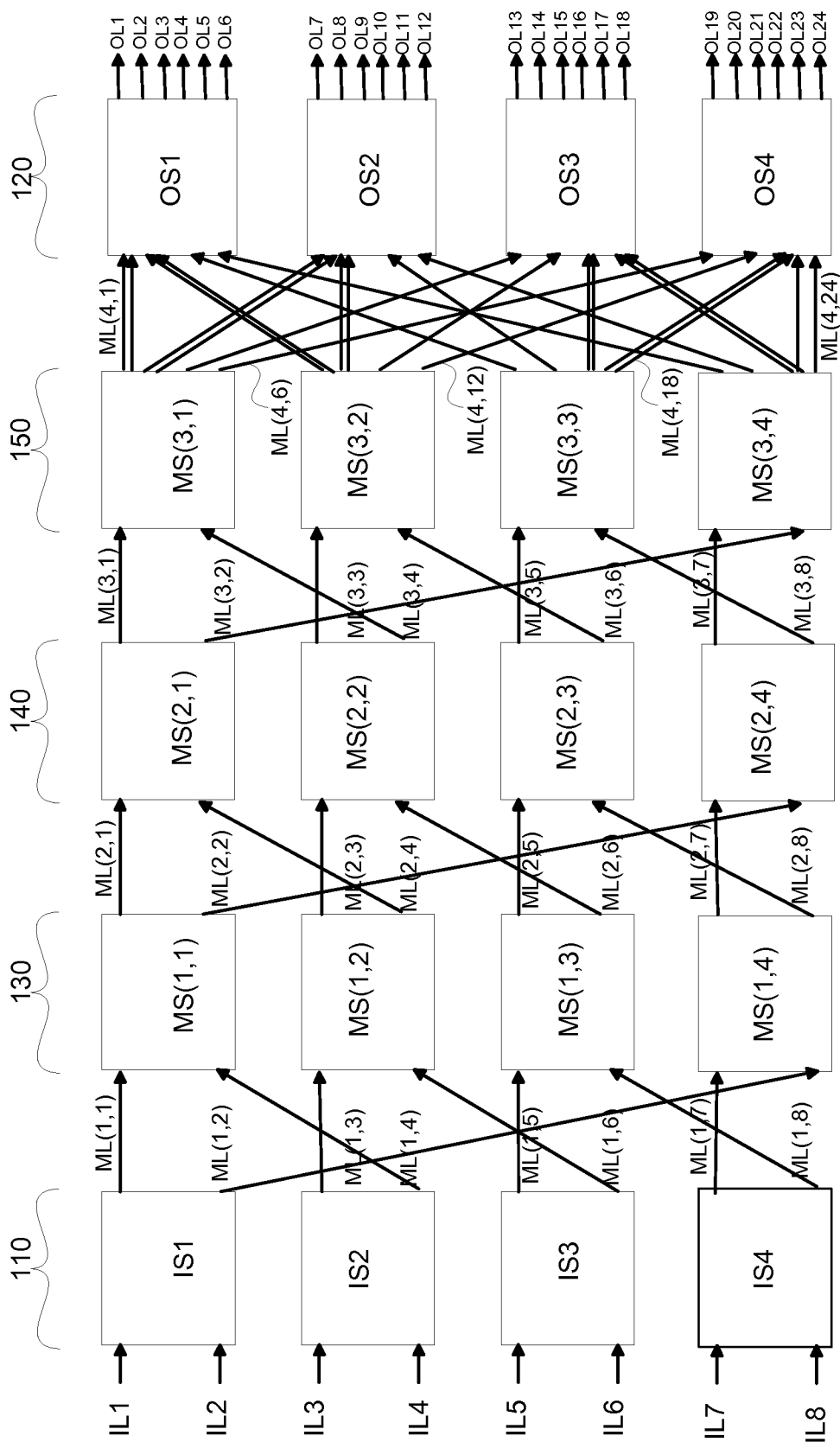


FIG. 6E1

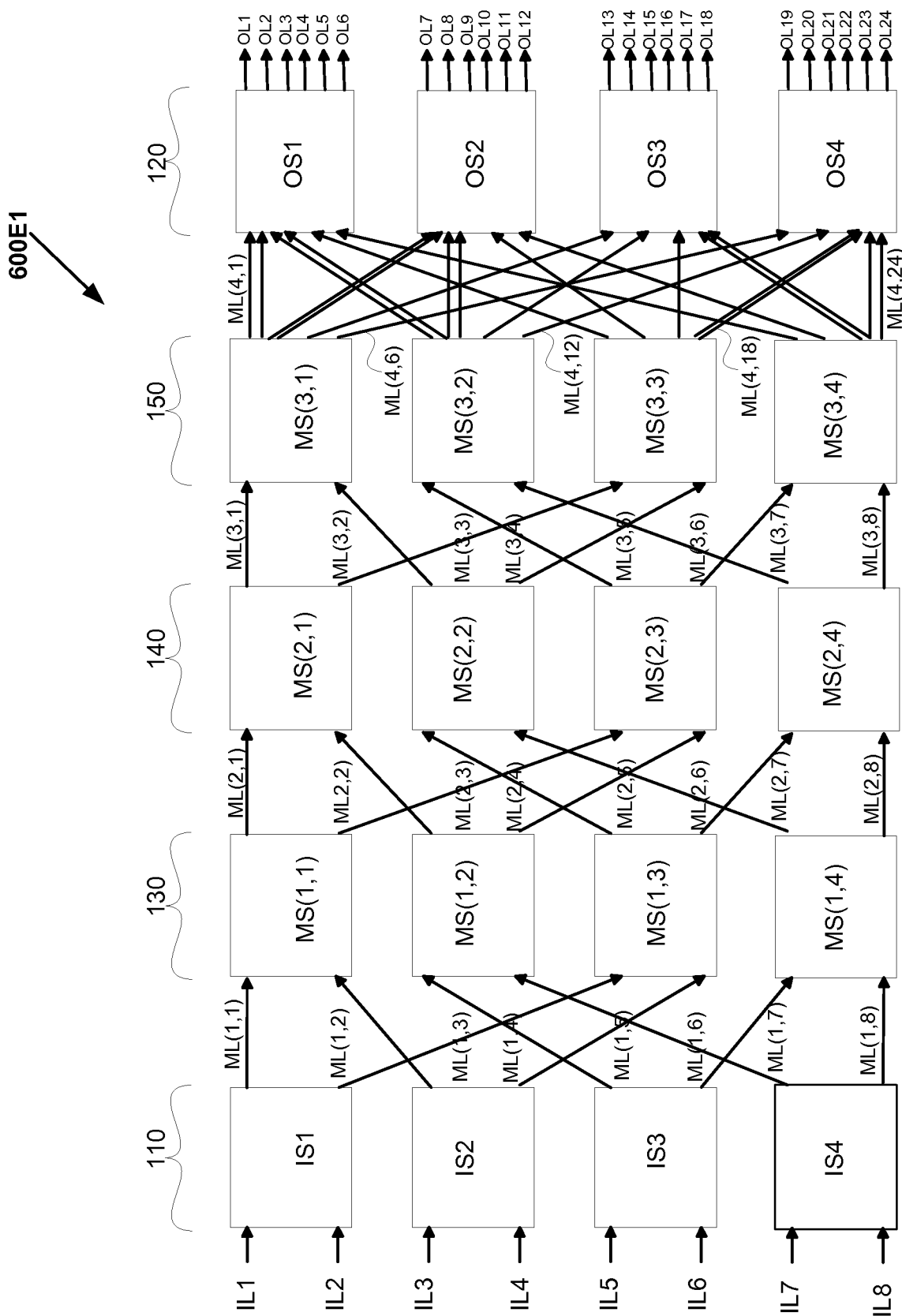


FIG. 6F1

600F1

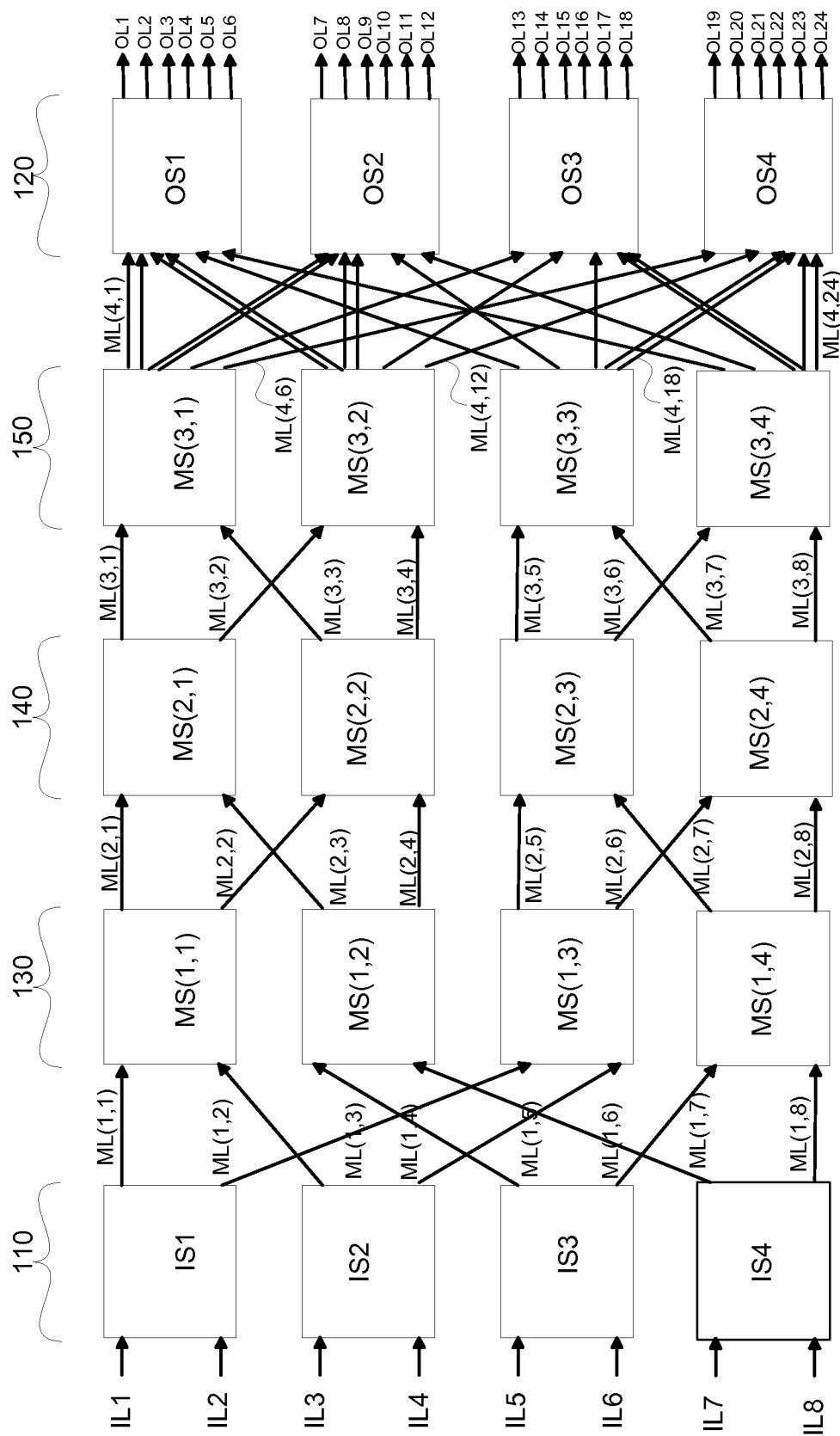


FIG. 6G1

600G1

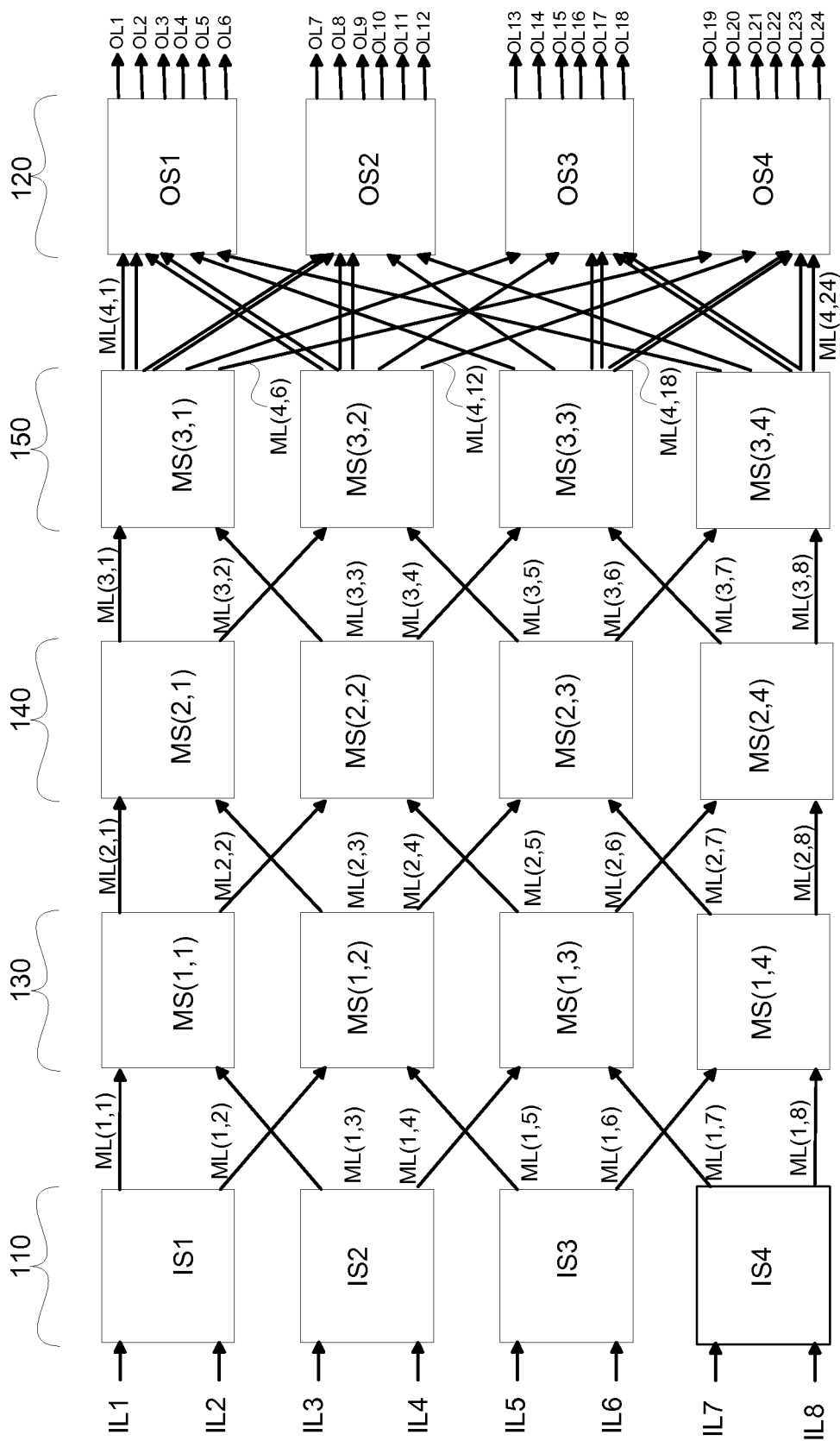


FIG. 6H1

600H1

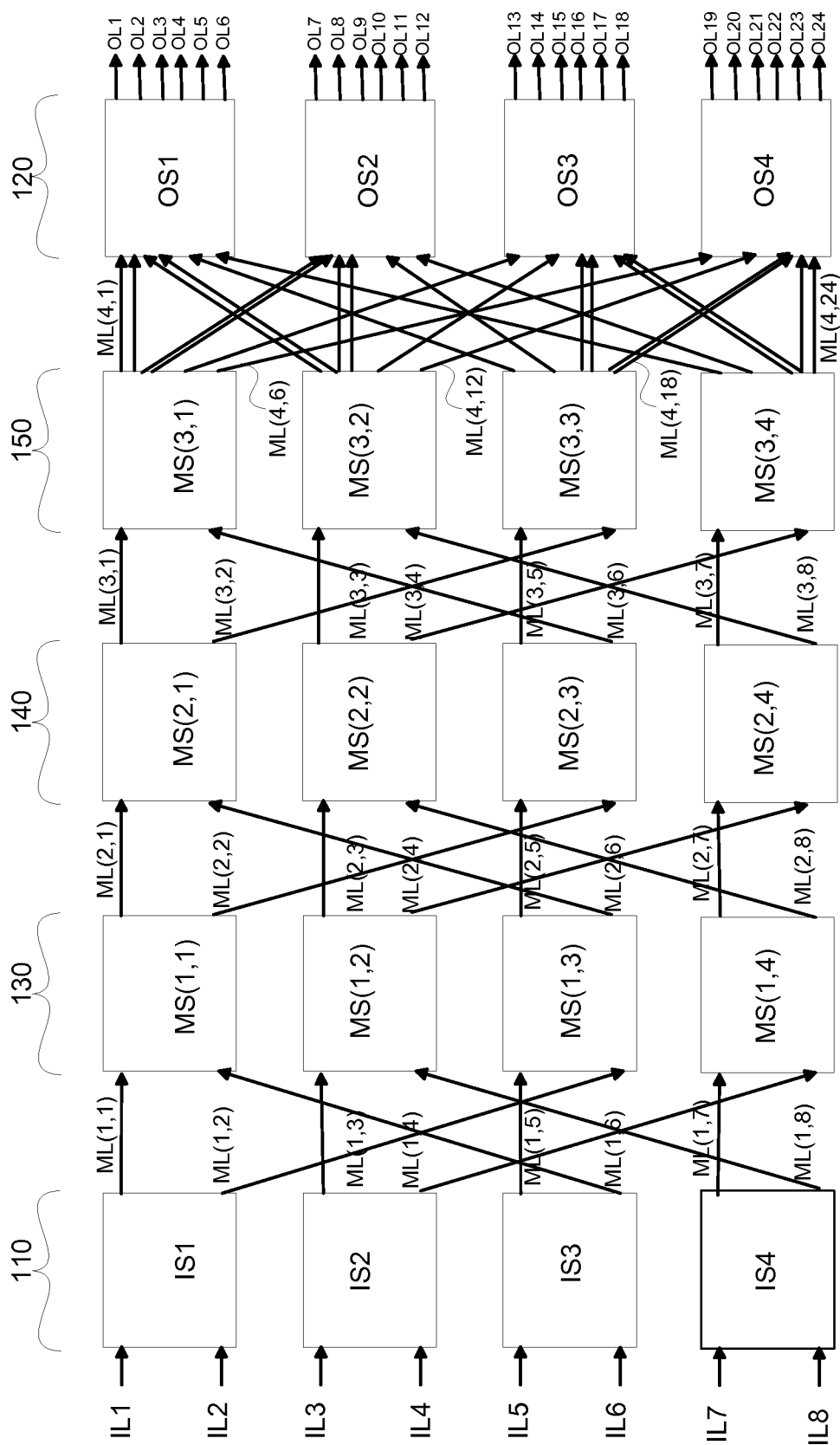


FIG. 6I1

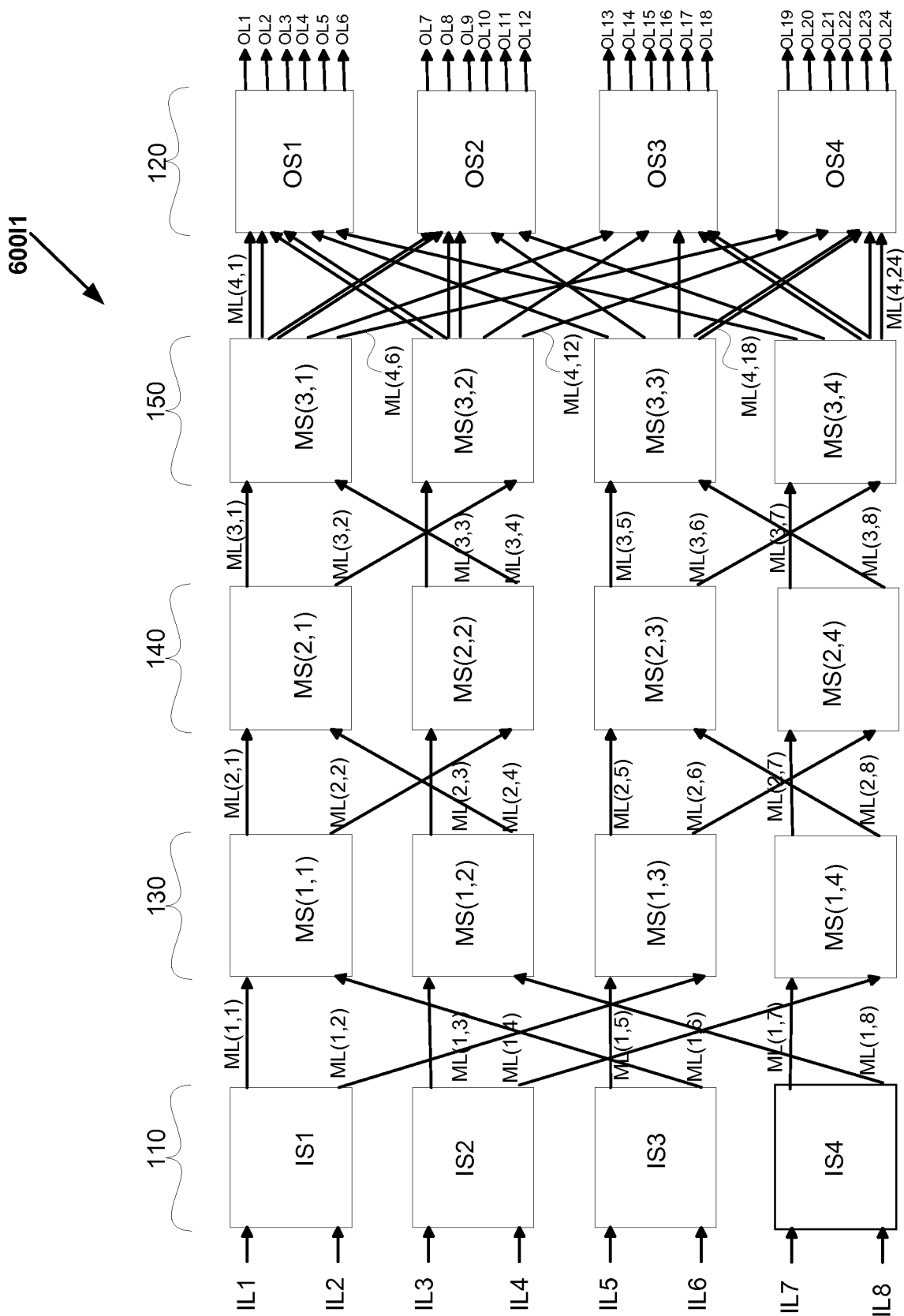


FIG. 6J1

600J1

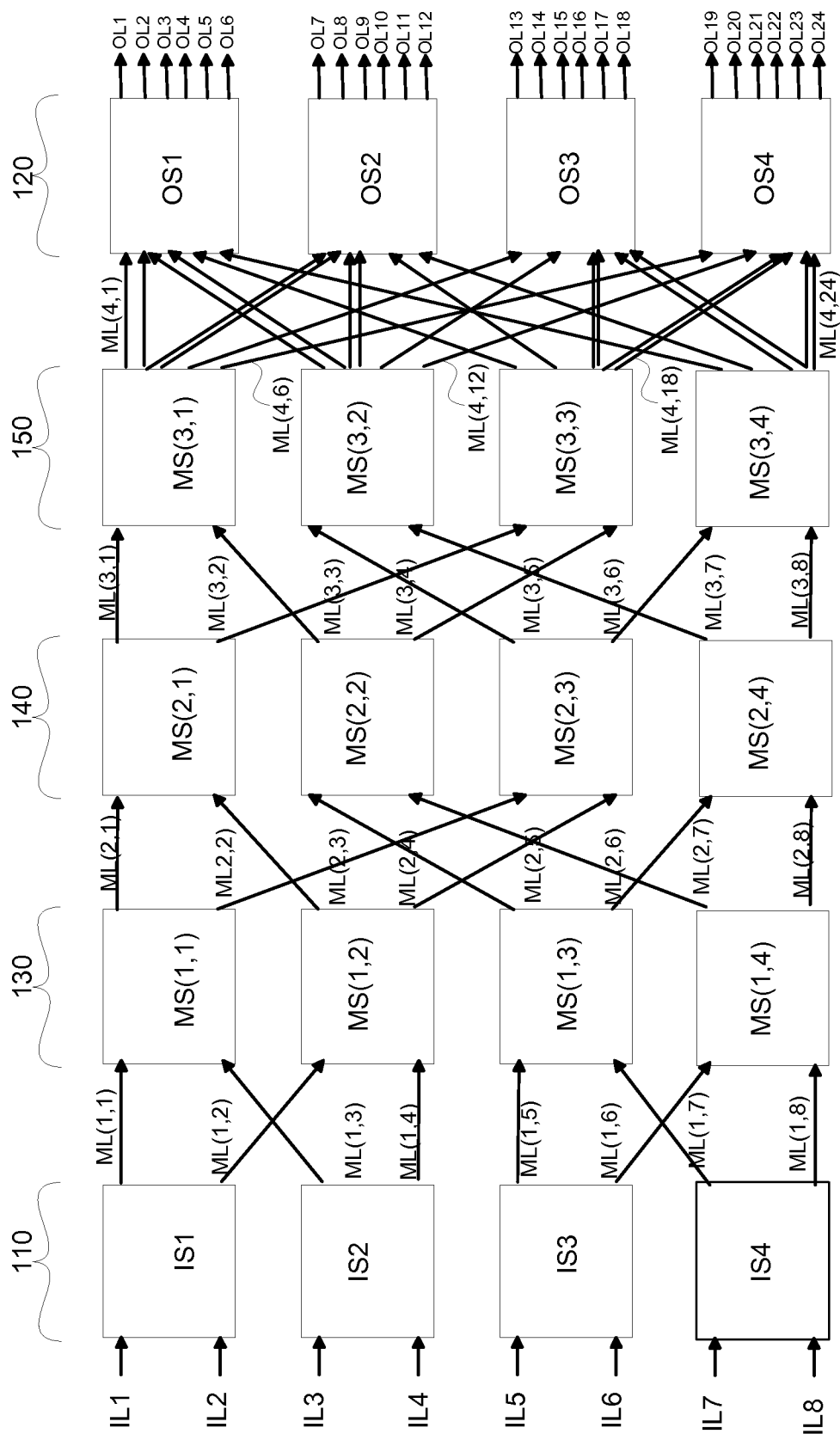


FIG. 6K1

600K1

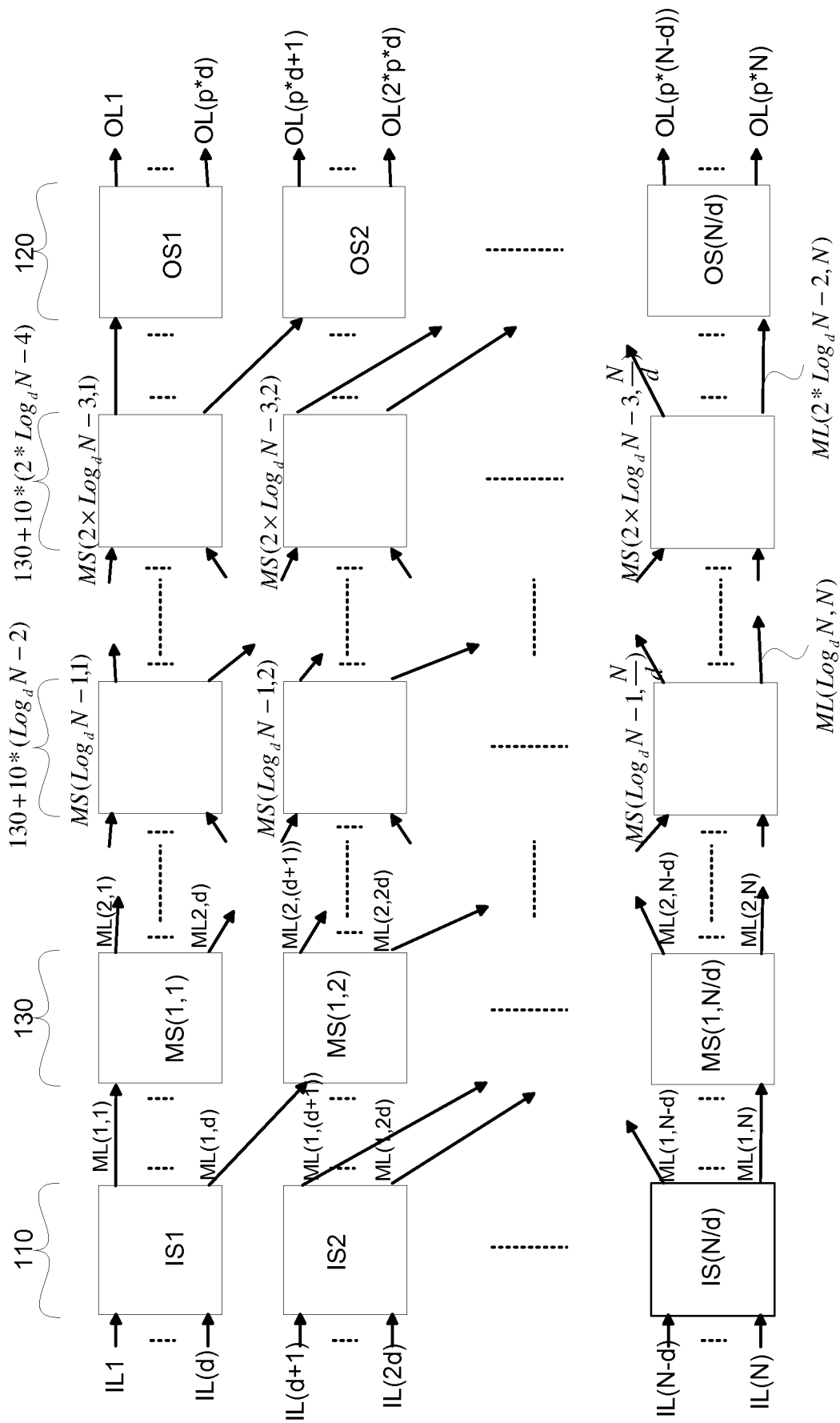


FIG. 6A2

600A2

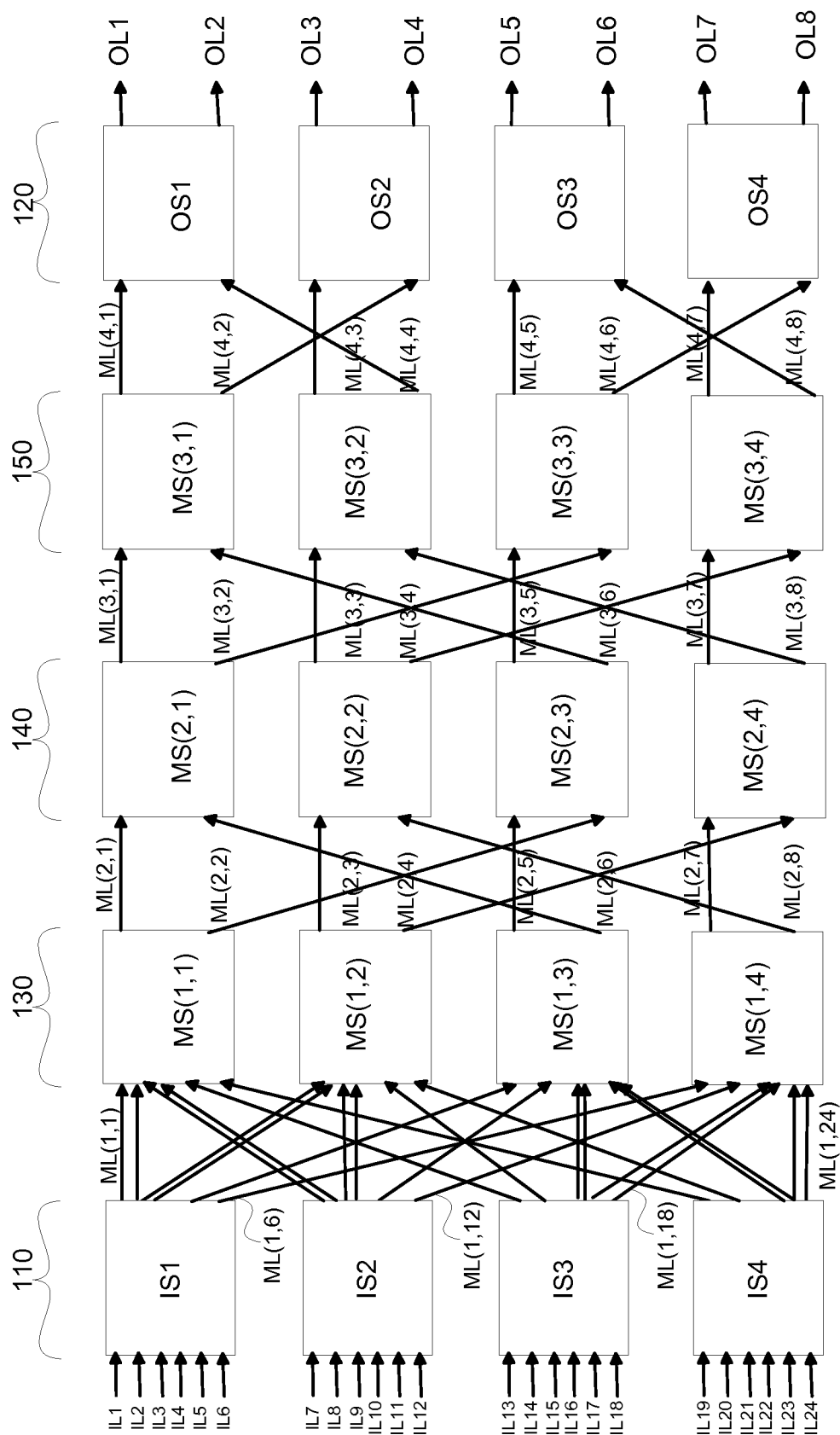


FIG. 6B2

600B2

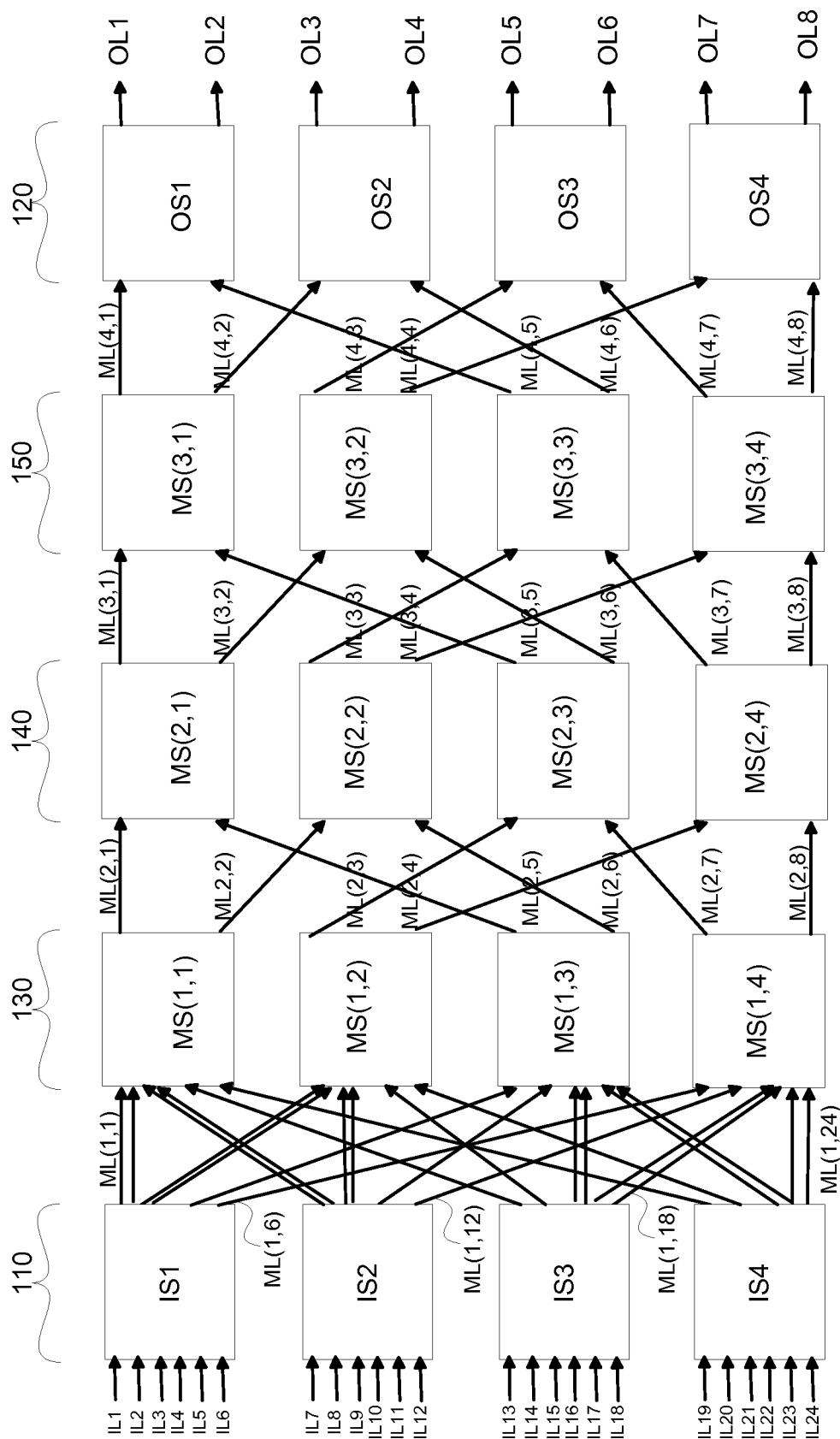


FIG. 6C2

600C2

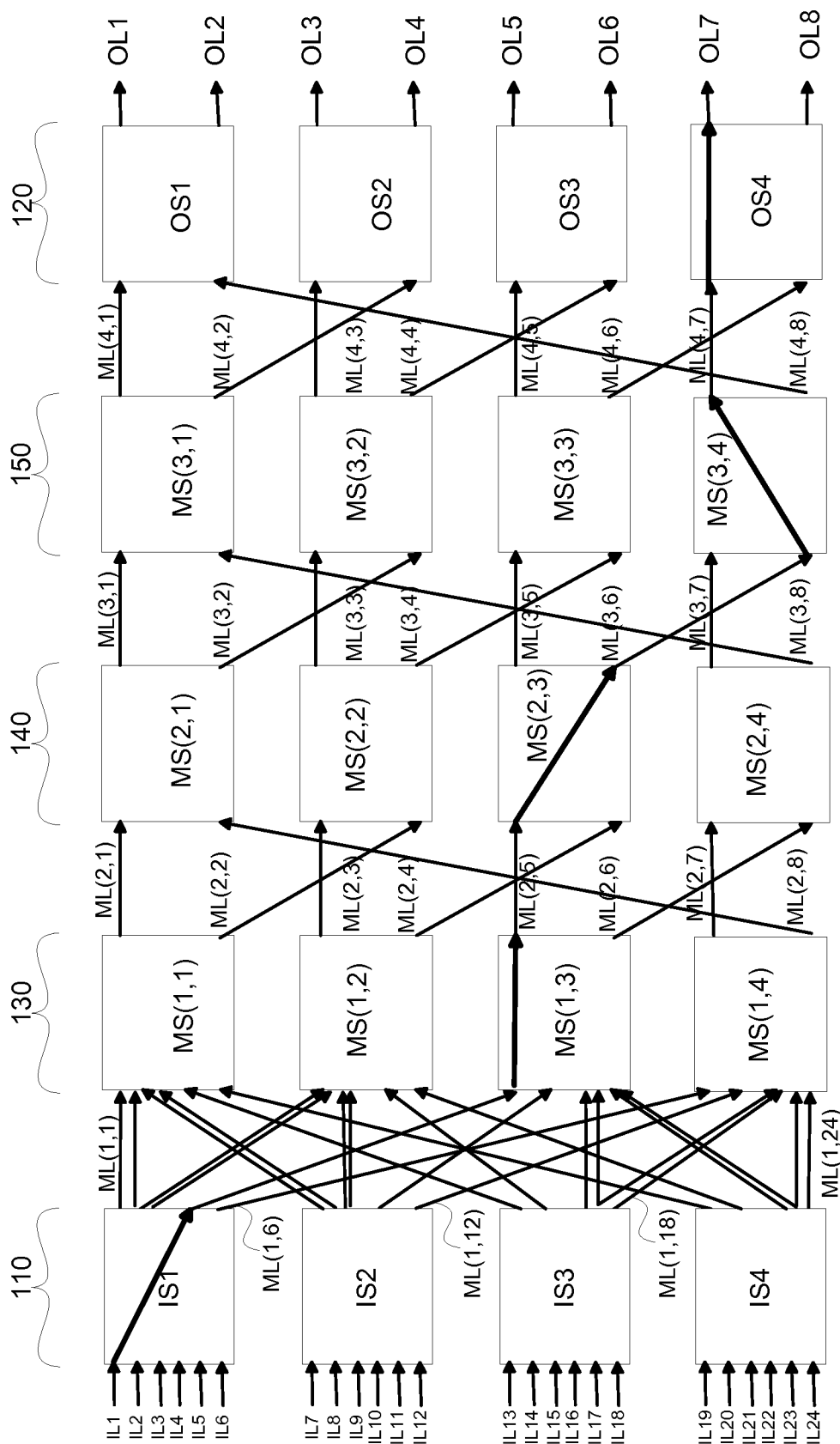


FIG. 6D2

600D2

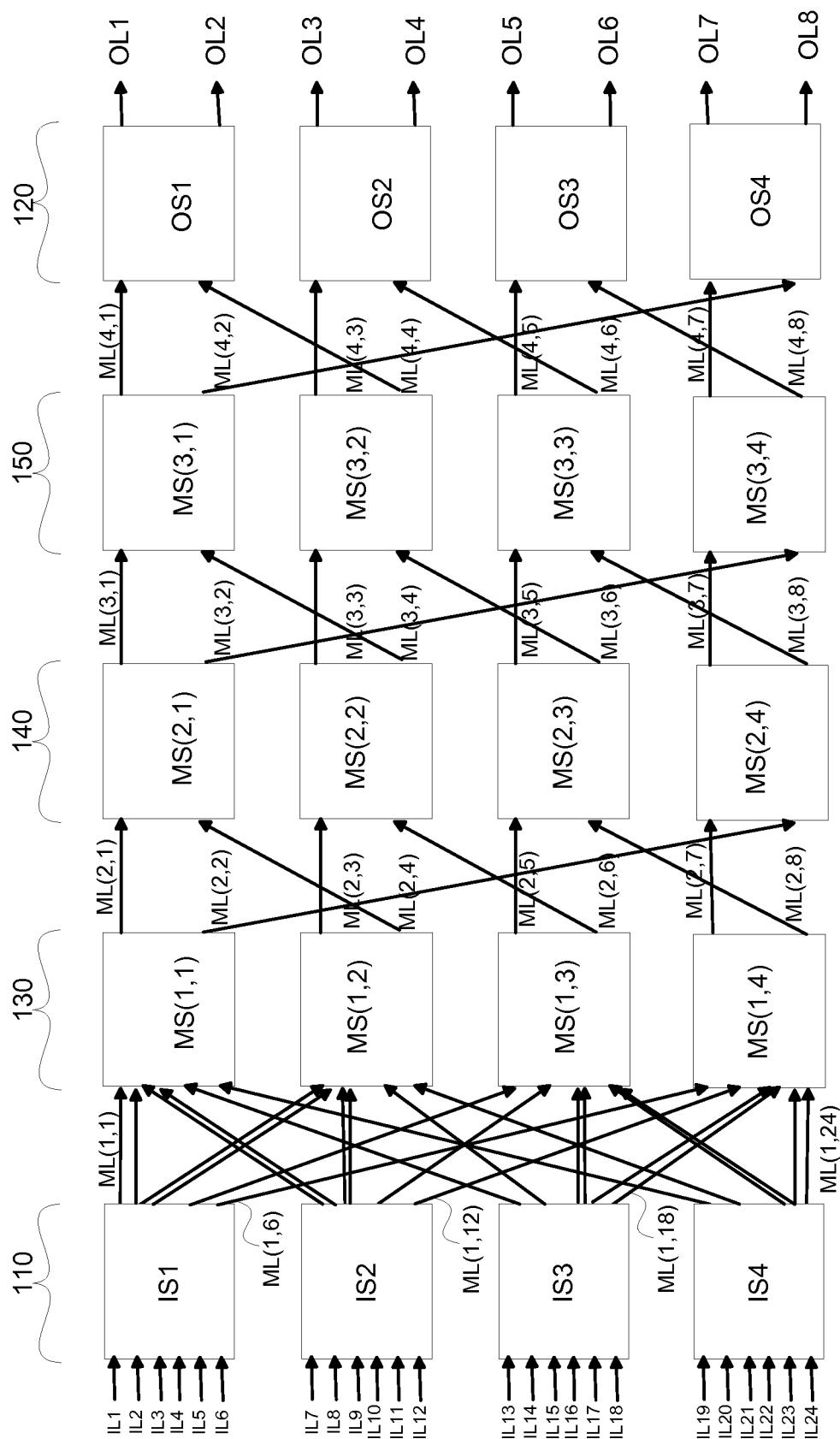


FIG. 6E2

600E2

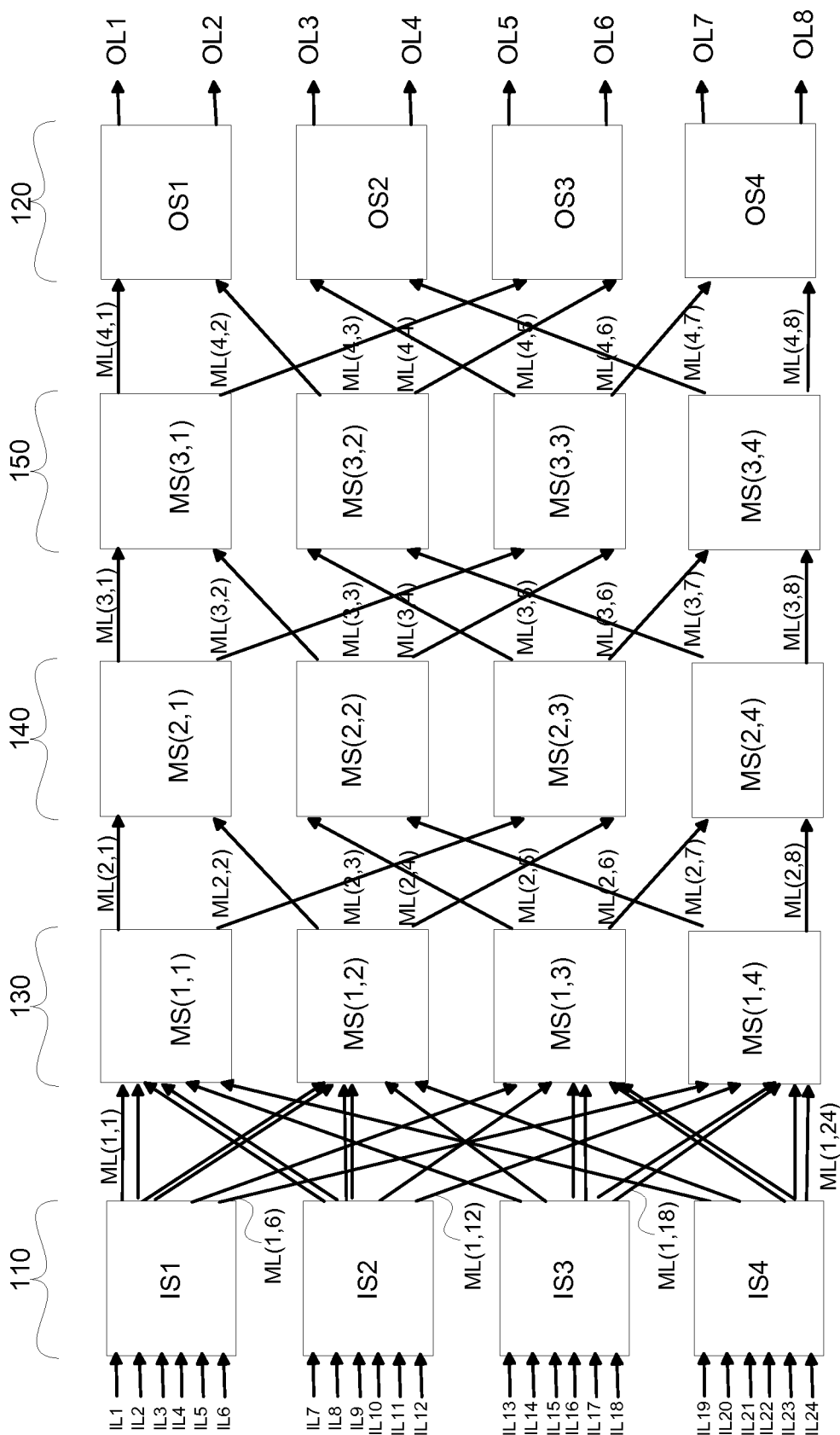


FIG. 6F2

600F2

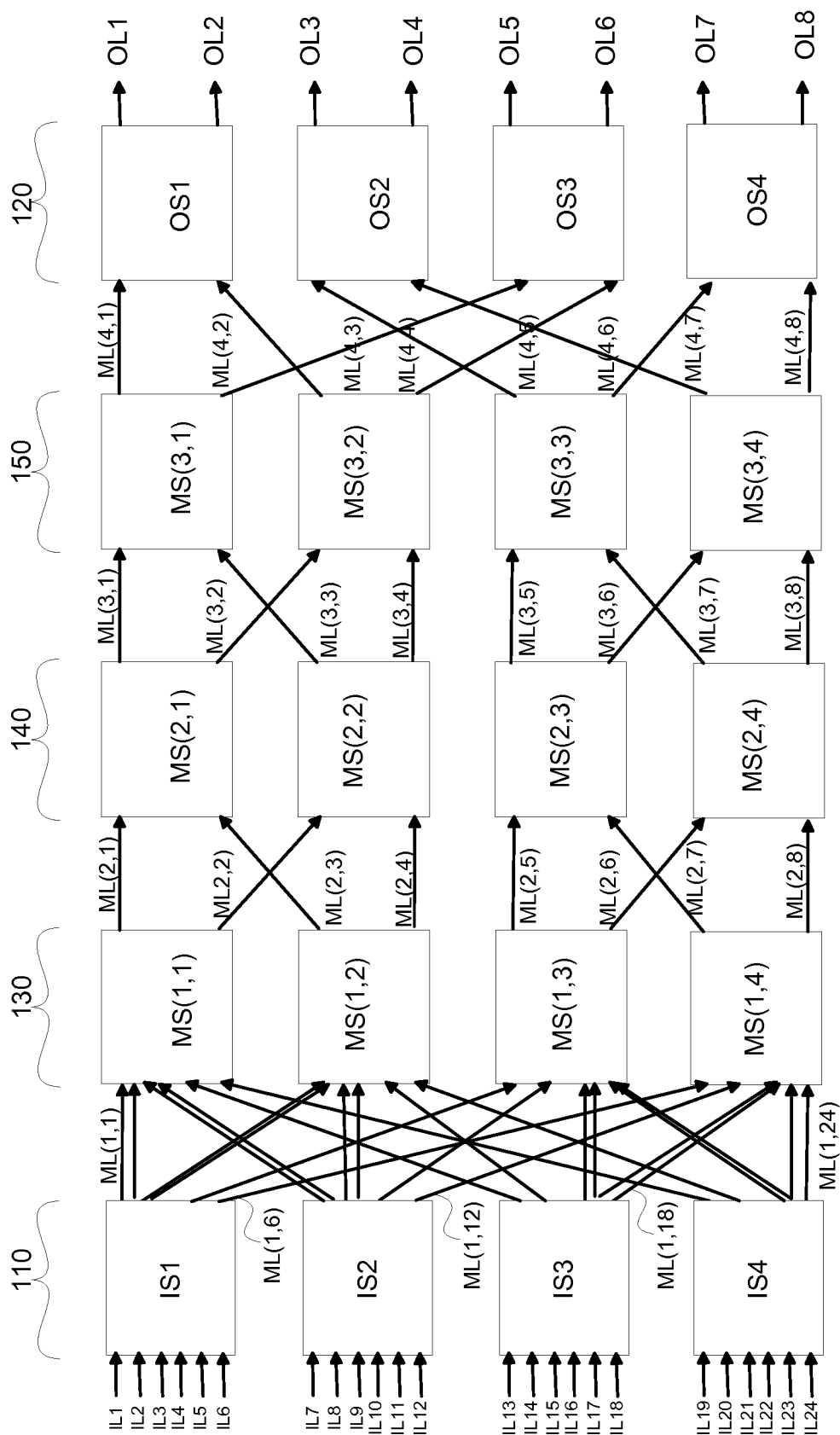


FIG. 6G2

600G2

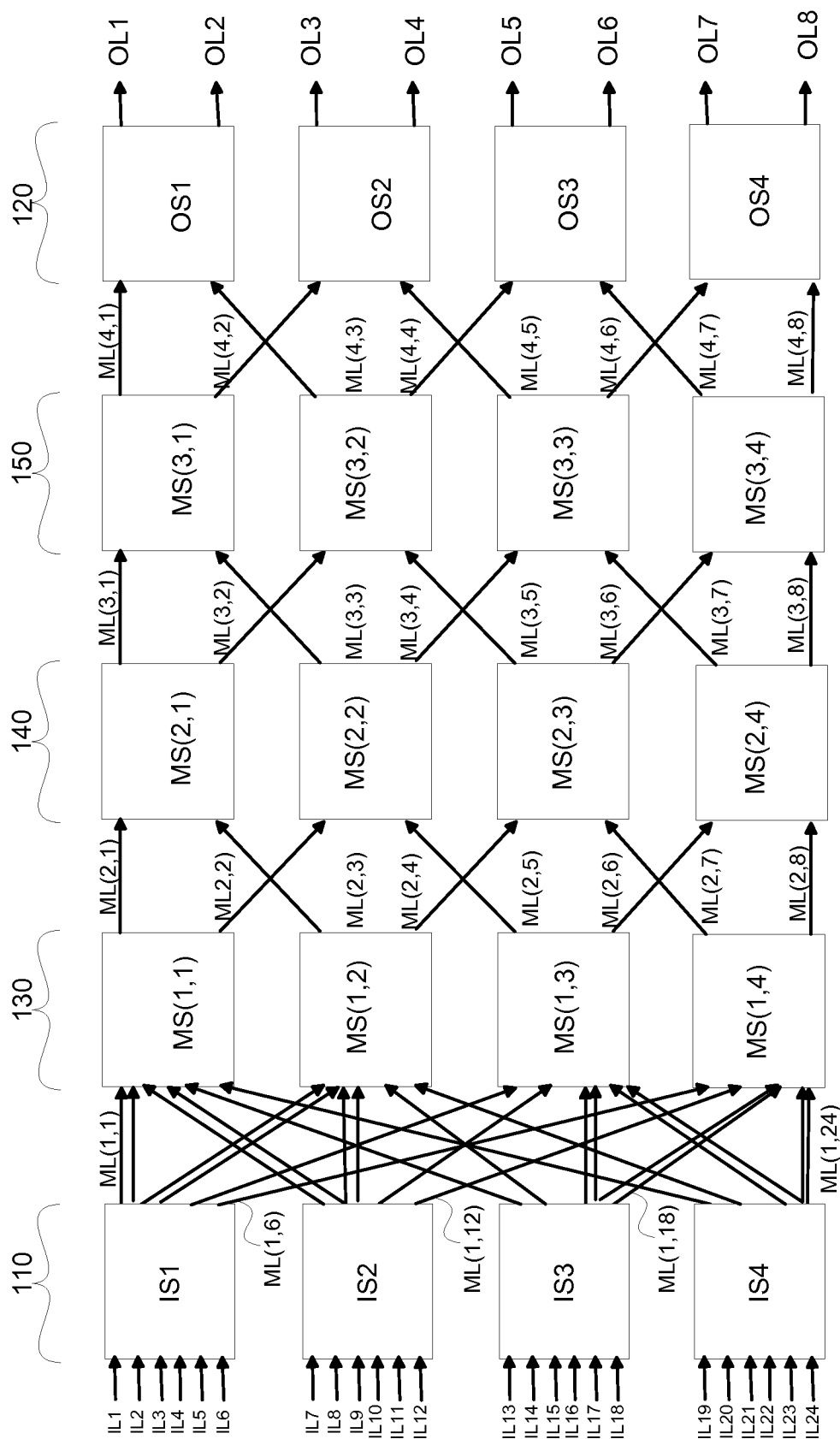


FIG. 6H2

600H2

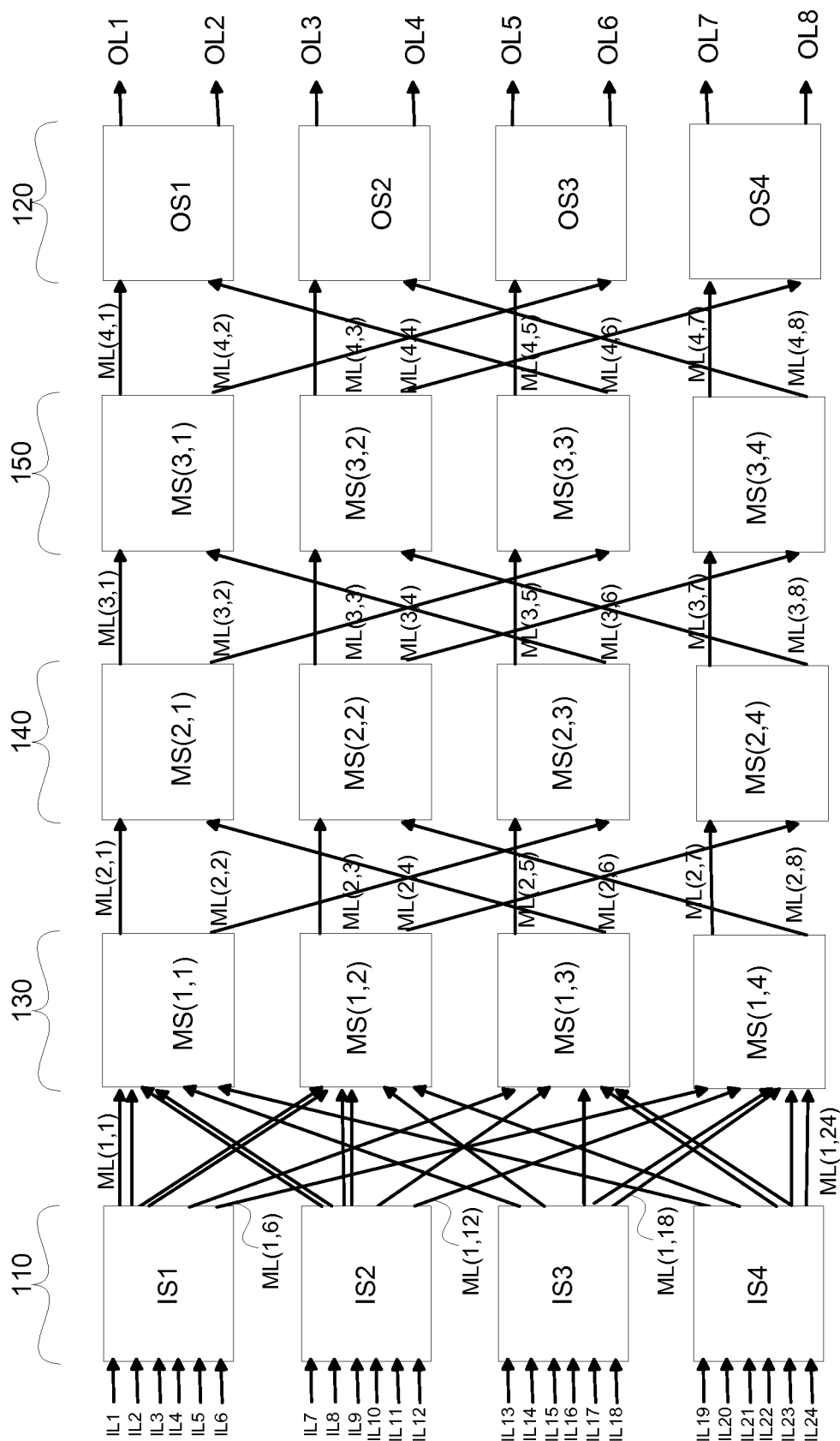


FIG. 6I2

600I2

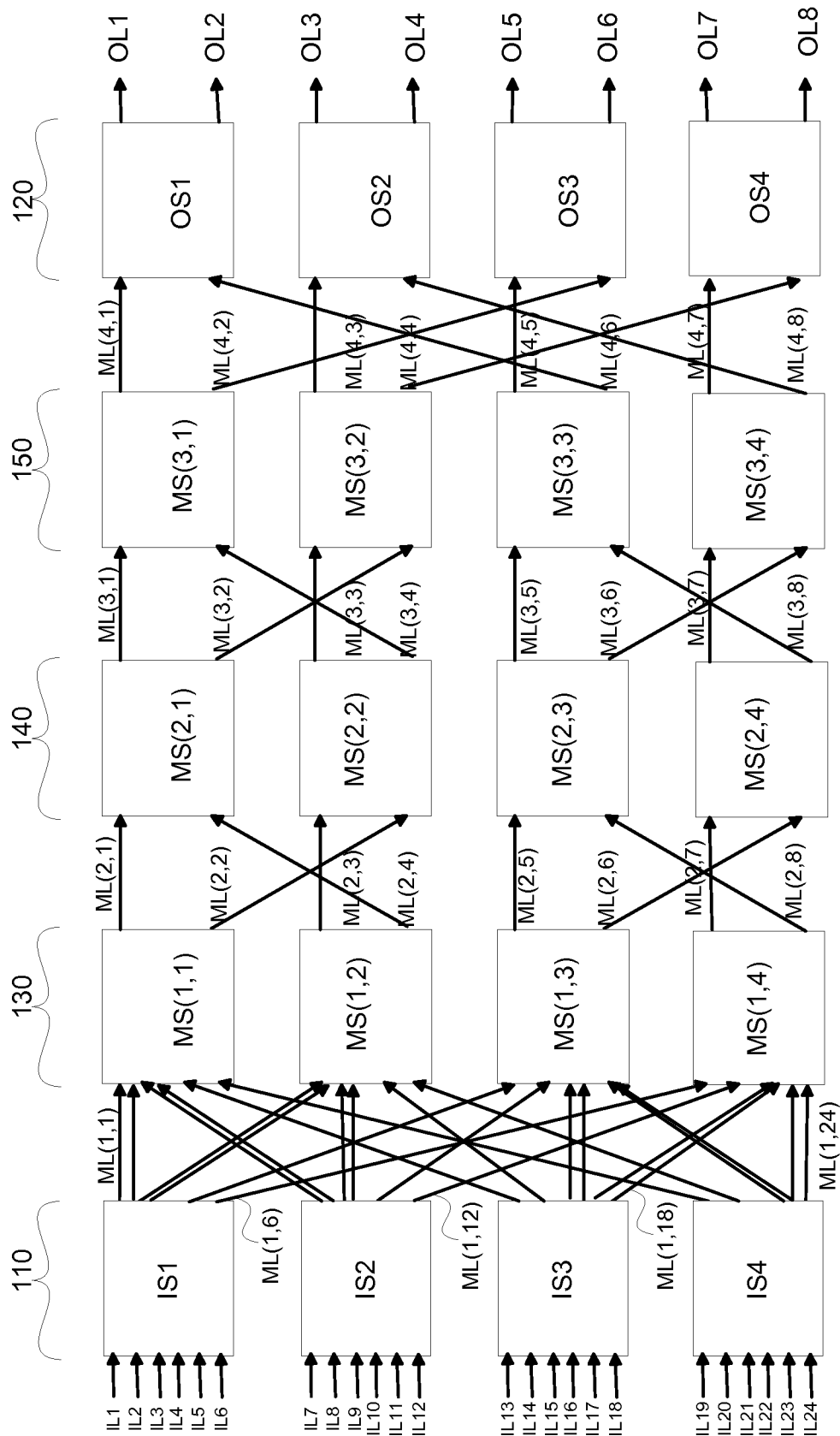


FIG. 6J2

600J2

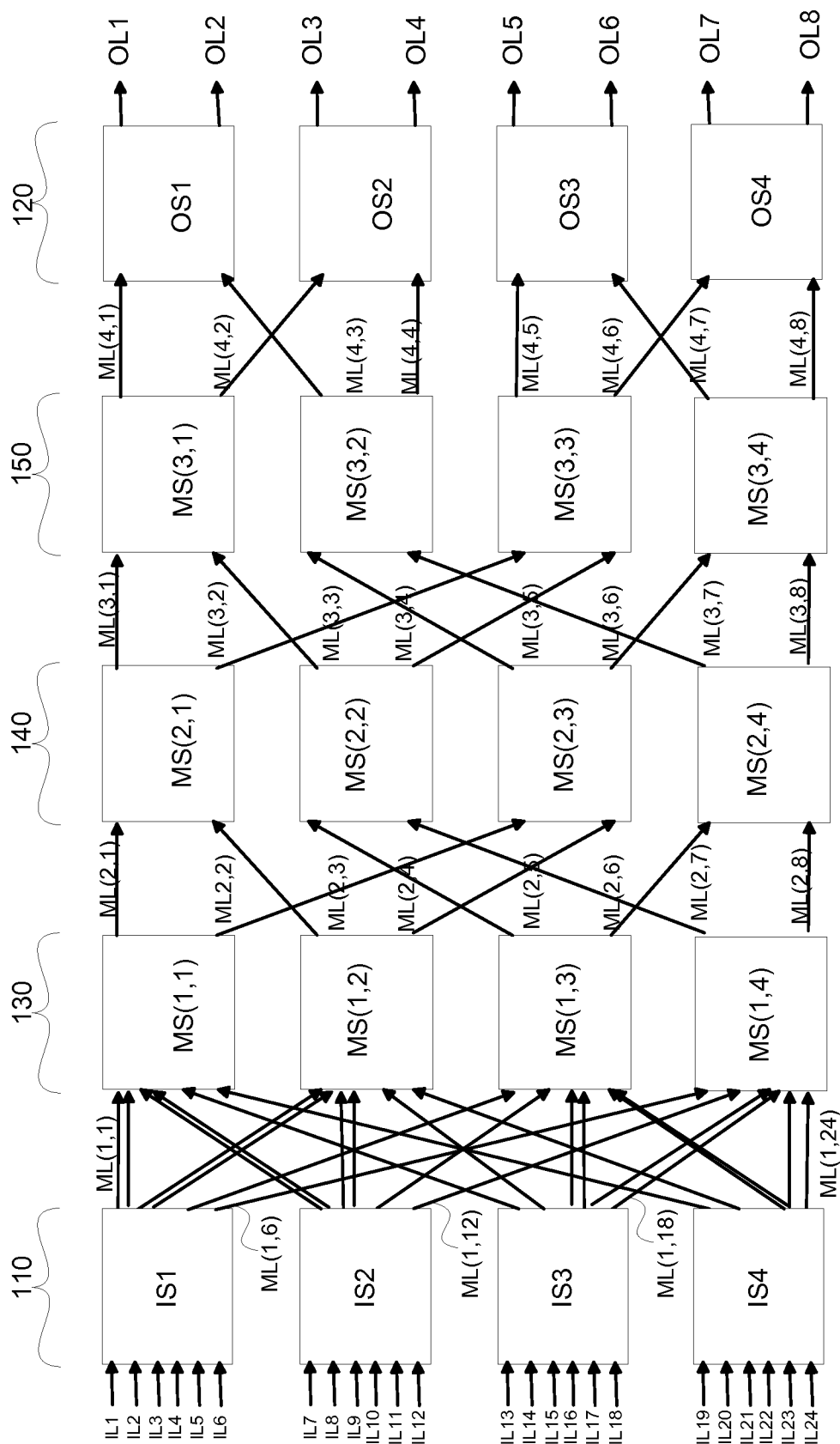


FIG. 6K2

600K2

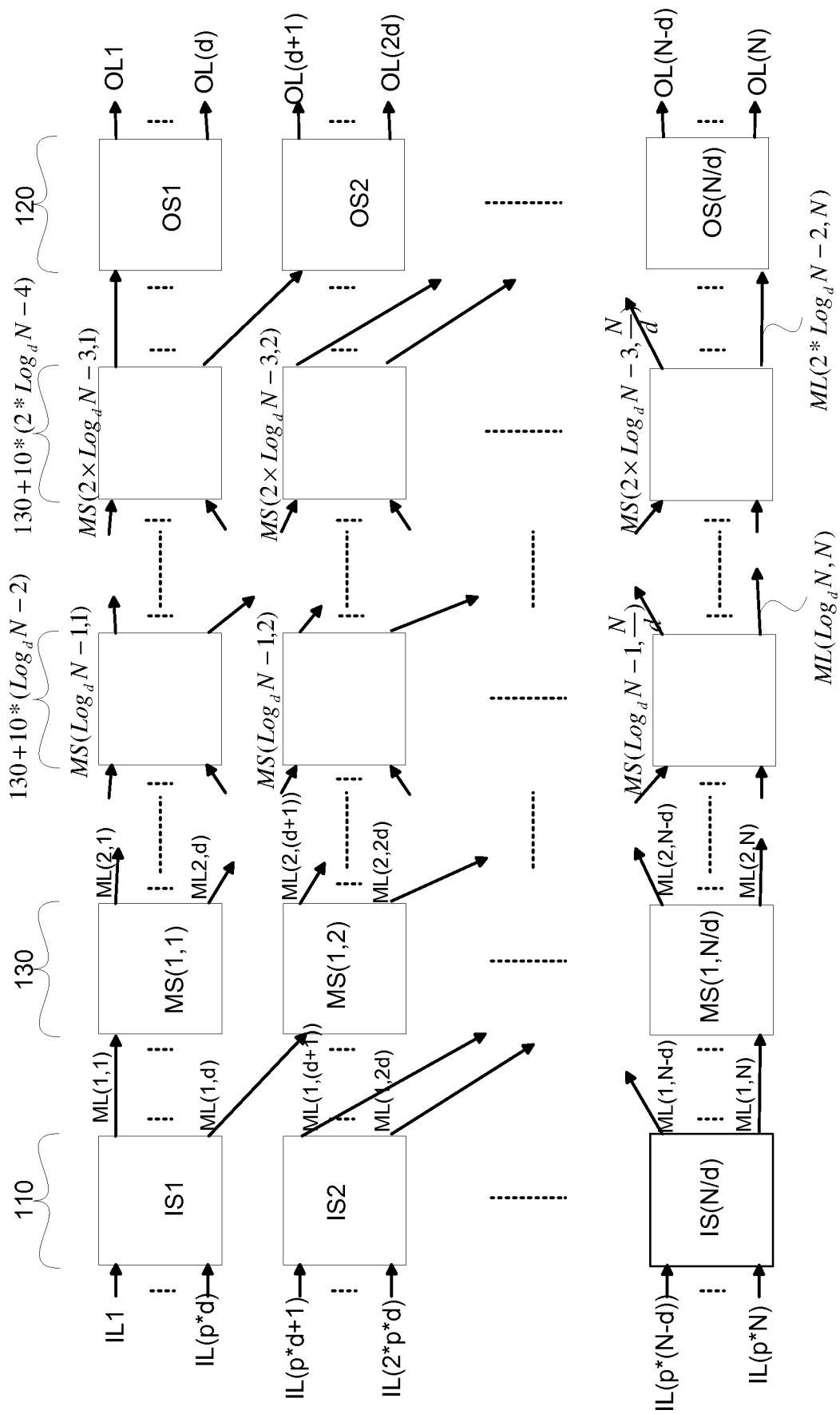


FIG. 7A

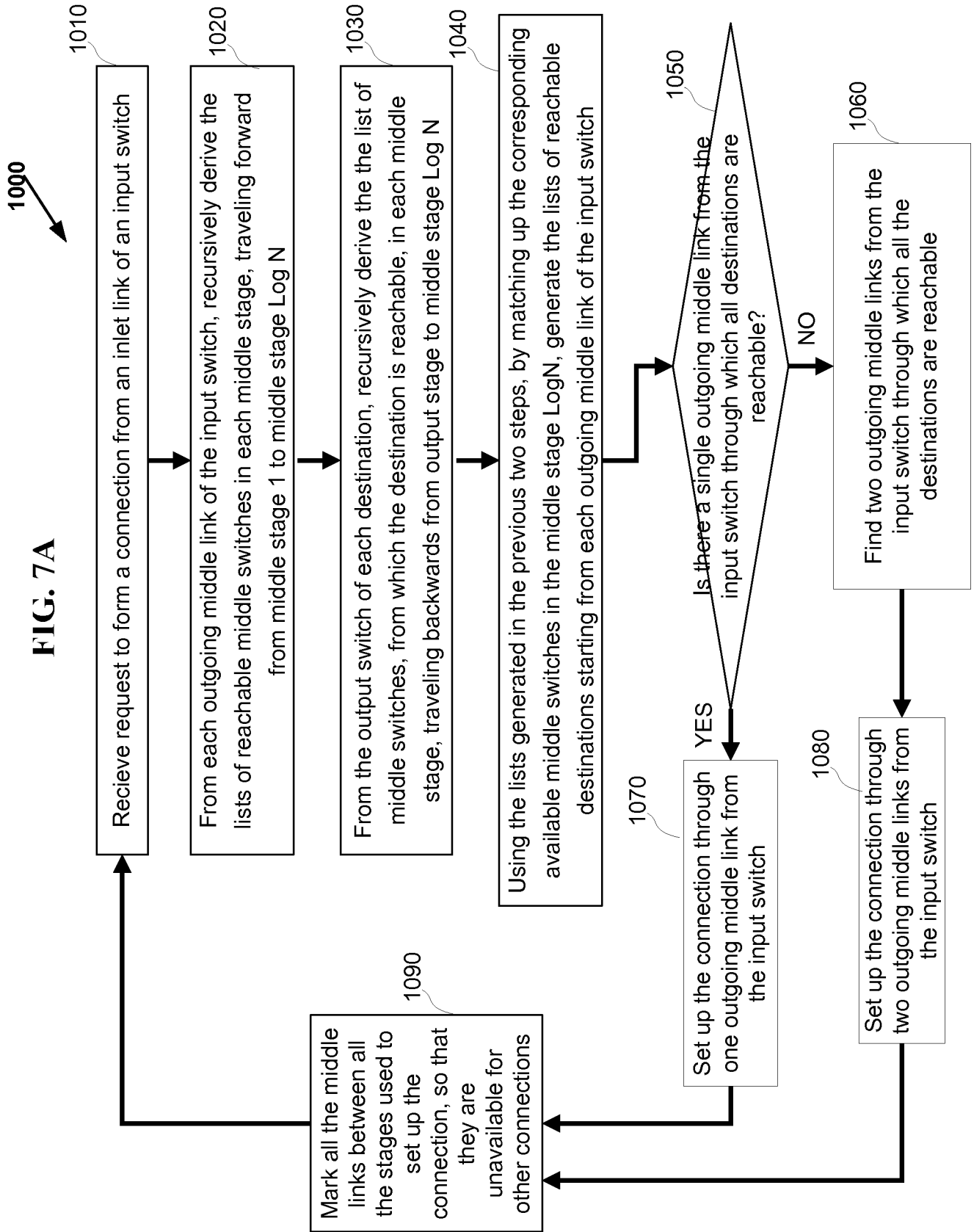


FIG. 8A

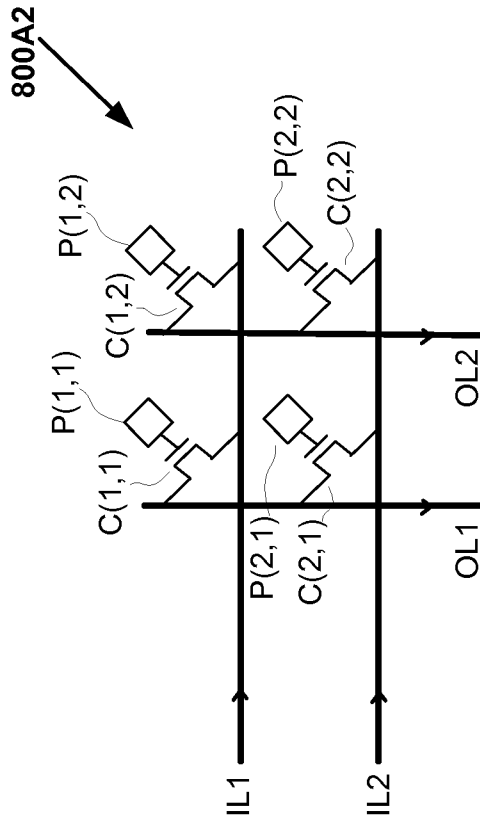


FIG. 8A2
(Prior Art)

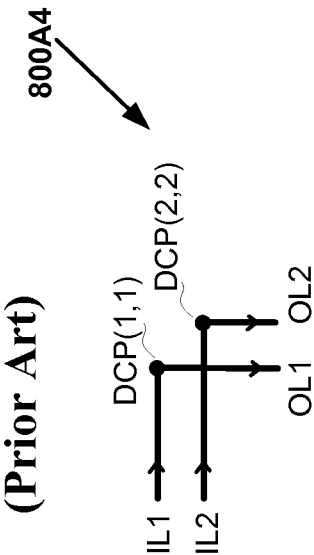


FIG. 8A4

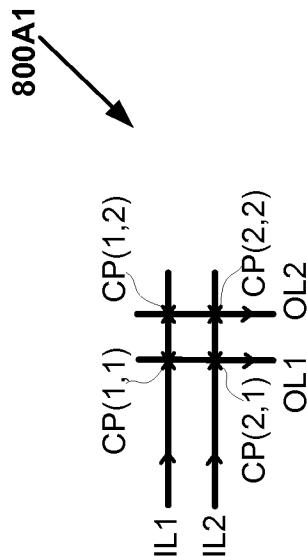


FIG. 8A1
(Prior Art)

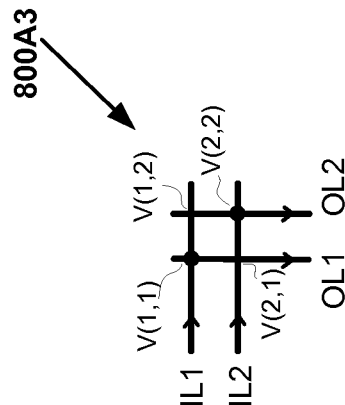


FIG. 8A3
(Prior Art)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 08/64604

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - H04L 12/50 (2008.04)

USPC - 370/388

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): H04L 12/50 (2008.04)

USPC: 370/388

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

USPC: 370/254, 351, 360, 388; 709/220, 223, 227, 228

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Electronic databases: USPTO WEST (PGPB, USPT, EPAB, JPAB); Google Scholar

Search Terms Used: multi-stage or multistage switch or network, multicast or multi-cast network or connections, input or output or middle or penultimate stages, checking or fanning connections, inlet or incoming or outlet or outgoing links or paths et

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2006/0165085 A1 (Konda) 27 July 2006 (27.07.2006) (abstract, and para [0009], [0022]-[0036], [0041]-[0051], [0087]-[0106], [0110]-[0117])	1-22
A	US 2006/0013207 A1 (McMillen et al.) 19 January 2006 (19.01.2006)	1-22
A	US 2005/0105517 A1 (Konda) 19 May 2005 (19.05.2005)	1-22
A	US 6,816,487 B1 (Roberts et al.) 09 November 2004 (09.11.2004)	1-22
A	US 6,473,428 B1 (Nichols et al.) 29 October 2002 (29.10.2002)	1-22

 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

11 August 2008 (11.08.2008)

Date of mailing of the international search report

15 AUG 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:

Lee W. Young

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774