Electronic Patent Application Fee Transmittal

Application Number:	
Filing Date:	
Title of Invention:	FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS
First Named Inventor/Applicant Name:	Venkat Konda
Filer:	Venkar Konda
Attorney Docket Number:	S-0038PCT

International Application for filing in the US receiving office Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Transmittal fee	1601	1	300	300
PCT Search Fee- no prior US appl filed	1602	1	1800	1800
Intl Filing Fee (1st-30 Pgs.) PCT Easy	1701	1	1173	1173
Suppl. Intl Filing Fee (each page > 30)	1703	88	14	1232
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Miscellaneous:				
		Total in USD ($)		**4505**

Electronic Acknowledgement Receipt

EFS ID:	3348142
Application Number:	
International Application Number:	PCT/US08/64603
Confirmation Number:	5416
Title of Invention:	FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS
First Named Inventor/Applicant Name:	Venkat Konda
Customer Number:	38139
Correspondence Address:	Venkat Konda Konda Technologies Inc 6278 Grand Oak Way - San Jose CA 95135 US 408-472-3273 venkat@kondatech.com
Filer:	Venkar Konda
Filer Authorized By:	
Attorney Docket Number:	S-0038PCT
Receipt Date:	22-MAY-2008
Filing Date:	
Time Stamp:	23:32:34
Application Type:	International Application for filing in the US receiving office

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes) /Message Digest	Multi Part /.zip	Pages (if appl.)
1	Specification	S-0038PCT.pdf	455158 c24688adfca0a5ac717689fe35cbcee7d49d1e3d	no	95
Warnings:					
Information:					
2	Drawings-only black and white line drawings	S-0038PCT-FIGs.pdf	219915 4a6f2ff640045000a9875f749dea89f0d42afa9e	no	20
Warnings:					
Information:					
3	PCT-Transmittal Letter	S-0038PCT-PTO-1382.pdf	513587 5740ec34a84eebdeb54e6f94d764bcec8f8e9c80	no	1
Warnings:					
Information:					
4	RO/101 - Request form for new IA - Conventional	S-0038PCT-RO-101.pdf	267730 55fc440b99149956009454fef2d0712258eba50e	no	3
Warnings:					
Information:					
5	Fee payment - International Application	S-0038PCT-Fee.pdf	356132 185cb848eb49552820d3cfbb7568c3b0212e3069	no	1
Warnings:					
Information:					
6	Fee Worksheet (PTO-06)	fee-info.pdf	8552 2965a44e9b9127a6c20c983776cd2b8169e4ce5d	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	1821074		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS

Venkat Konda

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, 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 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 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 GENERALIZED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

This application is related to and incorporates by reference in its entirety the PCT Application Docket No. S-0039PCT entitled "FULLY CONNECTED GENERALIZED MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed concurrently, the U.S. Provisional Patent Application Serial No. 60/940, 389 entitled "FULLY CONNECTED GENERALIZED REARRANGEABLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007, the U.S. Provisional Patent Application Serial No. 60/940, 391 entitled "FULLY CONNECTED GENERALIZED FOLDED MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007 and the U.S. Provisional Patent Application Serial No. 60/940, 392 entitled "FULLY CONNECTED GENERALIZED STRICTLY NONBLOCKING MULTI-LINK MULTI-STAGE NETWORKS" by Venkat Konda assigned to the same assignee as the current application, filed May 25, 2007.

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

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.

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 $m \geq \min\left((n-1)(x+r^{1/x})\right)$ 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 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 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

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 $\frac{N}{d}$ 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.

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 $\frac{N}{d}$ 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, 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

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.

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.

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.

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.

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.

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.

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.

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 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 art implementation of the diagram 500A1 of FIG. 5A1; FIG. 5A3 is a diagram 500A3 for one-time programmable integrated circuit prior art implementation of the diagram 500A1 of FIG. 5A1; FIG. 5A4 is a diagram 500A4 for integrated circuit placement and route implementation of the diagram 500A1 of FIG. 5A1.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is concerned with the 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 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 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 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.

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.

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 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 $V_{bft}(N_1, N_2, d, s)$, 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 Serial No. 60/940, 394 that is incorporated 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 in U.S. Provisional Patent Application Serial No. 61/018, 494 that is incorporated by reference above.

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 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).

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.

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

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).

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 $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$.

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 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.

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 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 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(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.

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.

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 $V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$.

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 $i \in \left\{1,2,\ldots,\frac{N}{d}\right\}$, let $I_i = O$, where $O \subset \left\{1,2,\ldots,\frac{N}{d}\right\}$, 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 $I_j = \phi$ 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 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 stage140) 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

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

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), 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 $V_{bft}(N_1, N_2, d, s)$.

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.

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.

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

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*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 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(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.

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.

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 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 $N_2 > N_1$.

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.

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.

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 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.

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 middle stage 130 can be denoted as $(d + d_1) * 2d$. The size of each switch in the root stage (i.e., middle stage140) 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).

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.

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 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 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)).

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).

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).

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 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 $V_{bft}(N_1,N_2,d,s)$.

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.

In the example illustrated in FIG. 1E (or in FIG1E1, 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 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 $d + d_1$ (= $d + p \times d$) 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 $d + d_1$ 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.

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.

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.

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 $V_{bft}(N_1, N_2, d, s)$ 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 $N_1 > N_2$.

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.

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.

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:

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:

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).

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.

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 $d*d$ 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 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).

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).

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 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)).

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 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)).

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

$V_{bft}(N,d,s)$ 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 $V_{bft}(N,d,s)$.

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:

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 $\frac{N}{d}$ 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 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 $ML(2 \times Log_d N - 2{,}1)$ - $ML(2 \times Log_d N - 2, d)$ 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.

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

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.

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 $V_{bft}(N,d,s)$.

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:

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).

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.

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 $d_2 = N_2 \times \frac{d}{N_1} = p \times d$. 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 stage140) 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.

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)).

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)).

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(4,3) 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).

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)).

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)).

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 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 $V_{bft}(N_1, N_2, d, s)$.

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:

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).

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.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(Log_d N_1 - 1,1)$ - $MS(Log_d N_1 - 1, \frac{N_1}{d})$ in the middle stage $130 + 10 * (Log_d N_1 - 2)$ are connected from exactly d switches in middle stage $130 + 10 * (Log_d N_1 - 3)$ through d links and also are connected to exactly d switches in middle stage $130 + 10 * (Log_d N_1 - 1)$ through d links.

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 $V_{bft}(N_1, N_2, d, s)$ 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 $N_2 > N_1$.

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.

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 $d_1 = N_1 \times \frac{d}{N_2} = p \times d$. 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 stage140) 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 $V_{bft}(N_1, N_2, d, s)$, 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 $\frac{d_1}{2}$ 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)).

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 $V_{bft}(N_1, N_2, d, s)$.

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 $ML(2 \times Log_d N_2 - 2,1)$ - $ML(2 \times Log_d N_2 - 2, d)$ to the output switch OS1).

Each of the $\frac{N_2}{d}$ input switches IS1 – IS(N_2/d) are connected to exactly $\frac{d_1}{2}$ 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 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 $MS(Log_d N_2 - 1{,}1)$ - $MS(Log_d N_2 - 1, \frac{N_2}{d})$ 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 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 d switches in middle stage $130 + 10 * (2 * Log_d N_2 - 4)$ 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 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 $N_1 > N_2$.

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.

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.

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)).

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 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 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 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 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

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 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)).

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 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 $V_{mlink-bft}(N,d,s)$.

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.

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.

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

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.

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 $V_{mlink-bft}(N,d,s)$.

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.

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 first type (with fan-out r', $1 \le r' \le \frac{N}{d}$) although the same discussion is applicable to the second type.

To characterize a multicast assignment, for each inlet link $i \in \left\{1,2,...,\frac{N}{d}\right\}$, let $I_i = O$, where $O \subset \left\{1,2,...,\frac{N}{d}\right\}$, 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 $I_j = \phi$ 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.

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 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).

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.

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 middle stage 130 can be denoted as $4d * 2(d + d_2)$. The size of each switch in the root stage (i.e., middle stage140) 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).

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.

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)).

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)).

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).

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 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)).

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)).

Finally the connection topology of the network 300C shown in FIG. 3C 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 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 $V_{mlink-bft}(N_1, N_2, d, s)$.

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.

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

$V_{mlink-bft}(N_1, N_2, d, s)$ 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*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.

Similarly each of the $\frac{N_1}{d}$ middle switches $MS(Log_d N_1 - 1,1)$ - $MS(Log_d N_1 - 1, \frac{N_1}{d})$ in the middle stage $130 + 10 * (Log_d N_1 - 2)$ are connected from exactly d switches in middle stage $130 + 10 * (Log_d N_1 - 3)$ through $2 \times d$ links and also are connected to exactly d switches in middle stage $130 + 10 * (Log_d N_1 - 1)$ through $2 \times d$ links.

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 $V_{mlink-bft}(N_1, N_2, d, s)$ 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 $N_2 > N_1$.

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.

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 $I_j = \phi$ 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 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 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 stage140) 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).

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.

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 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)).

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)).

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).

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 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 $V_{mlink-bft}(N_1, N_2, d, s)$.

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.

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.

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:

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.

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

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 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 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 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.

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 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:

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 $V_{mlink-bft}(N_1, N_2, d, s)$ can also be operated in strictly nonblocking manner for multicast when $s \geq 3$ according to the current invention.

Scheduling Method Embodiments:

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 $130+10*(Log_d N-2)$, 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 out going 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 $130+10*(Log_d N-2)$ are derived. This process is repeated for each available outgoing middle link from the input switch of the multicast connection and separate reachable lists are derived in each middle stage from middle stage 130 to middle stage $130+10*(Log_d N-2)$ for all the available outgoing middle links from the input switch. Then the control goes to act 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 $130+10*(Log_d N-2)$, 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 $130+10*(Log_d N-2)$ 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 $130+10*(Log_d N-2)$ 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 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 "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.

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.

If act 1050 results "no", that is one outgoing link is not available through which all the destinations of the multicast connection are reachable, then the control goes to act 1060. In act 1060, it is checked if all destination links of the multicast connection are reachable using two outgoing middle links from the input switch. According to the current invention, it is always possible to find at most two outgoing middle links from the input switch through which all the destinations of a multicast connection are reachable. So act 1060 always results in "yes", and then the control transfers to act 1080. In act 1080, the multicast connection is setup by traversing from the selected only two outgoing middle links of the input switch in act 1060, to all the destinations. Also the nearest U-turn is taken while setting up the connection. That is at any middle stage if one of the middle switch in the lists derived in acts 1020 and 1030 are common then the connection is setup so that the U-turn is made to setup the connection from that middle switch for all the destination links reachable from that common middle switch. Then the control transfers to act 1090.

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.

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

$V_{bft}(N,d,s)$, $V_{bft}(N_1,N_2,d,s)$, $V_{mlink-bft}(N,d,s)$, 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 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).

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 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 virtal 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.

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.

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

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 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 (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 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 it its immediate succeeding 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 $\frac{N}{d}$ 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 (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 succeeding 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 it 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 $y \geq (\log_d N_1) - 1$ when $N_2 > N_1$, and $y \geq (\log_d N_2) - 1$ 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 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 $x \geq 2$, 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 $\frac{N}{d}$ 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 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 it 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 $N_1 = p * N_2$ 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 $\frac{N}{d}$ 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 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 it 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 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 $\frac{N}{d}$ 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 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 it 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 $N_1 = p * N_2$ 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 $\frac{N}{d}$ 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 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 it 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:

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 **Error! Reference source not found.** 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 **Error! Reference source not found.** 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 **Error! Reference source not found.** 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 **Error! Reference source not found.** 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 **Error! Reference source not found.** wherein any of said acts of checking and setting up are performed recursively.

FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS

Venkat Konda

ABSTRACT OF DISCLOSURE

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 $\frac{N}{d}$ 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 $\frac{N}{d}$ switches with each of them having d outlet links and $s \times d$ incoming links connecting from switches in its immediate succeeding stage. 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

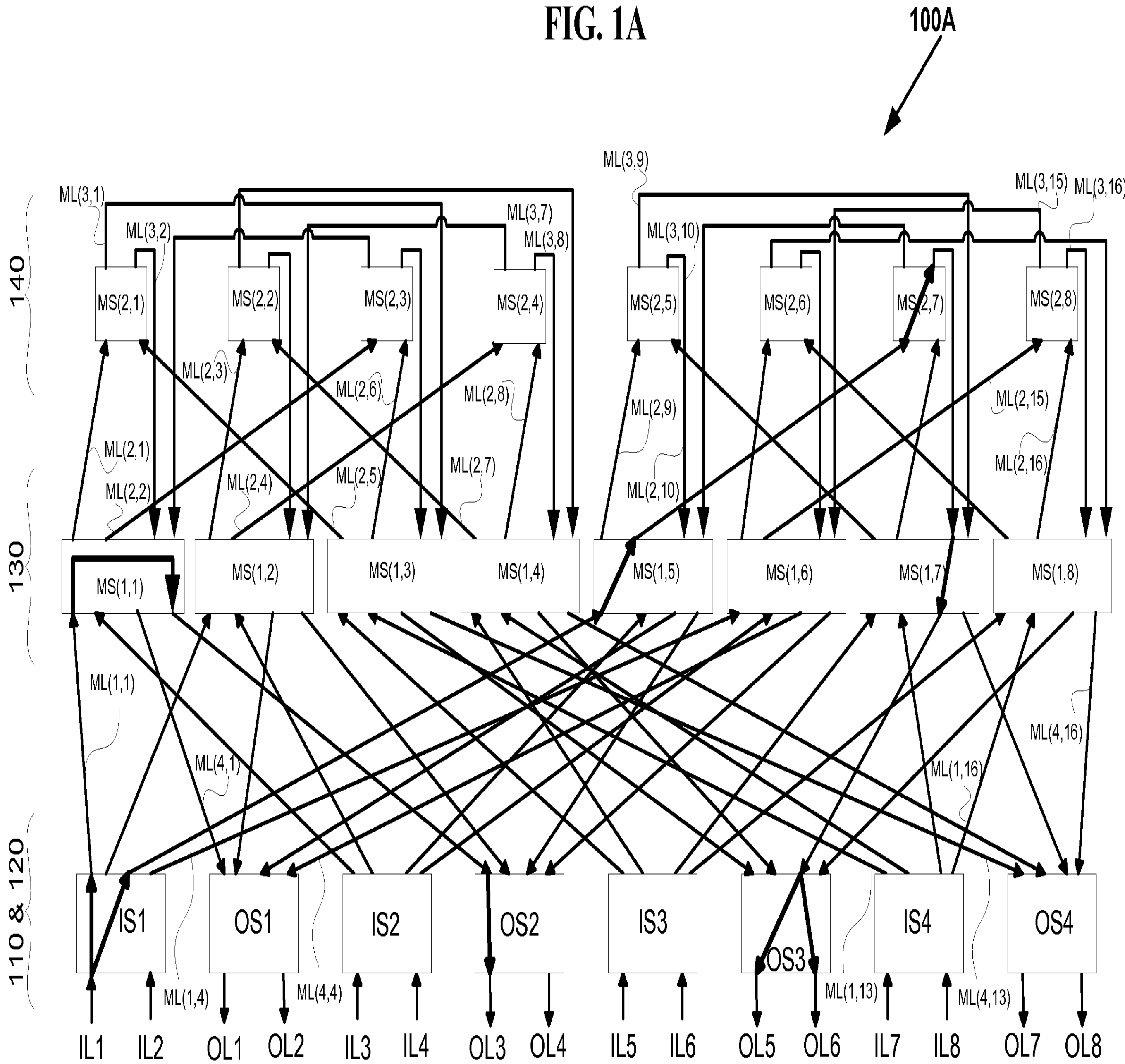
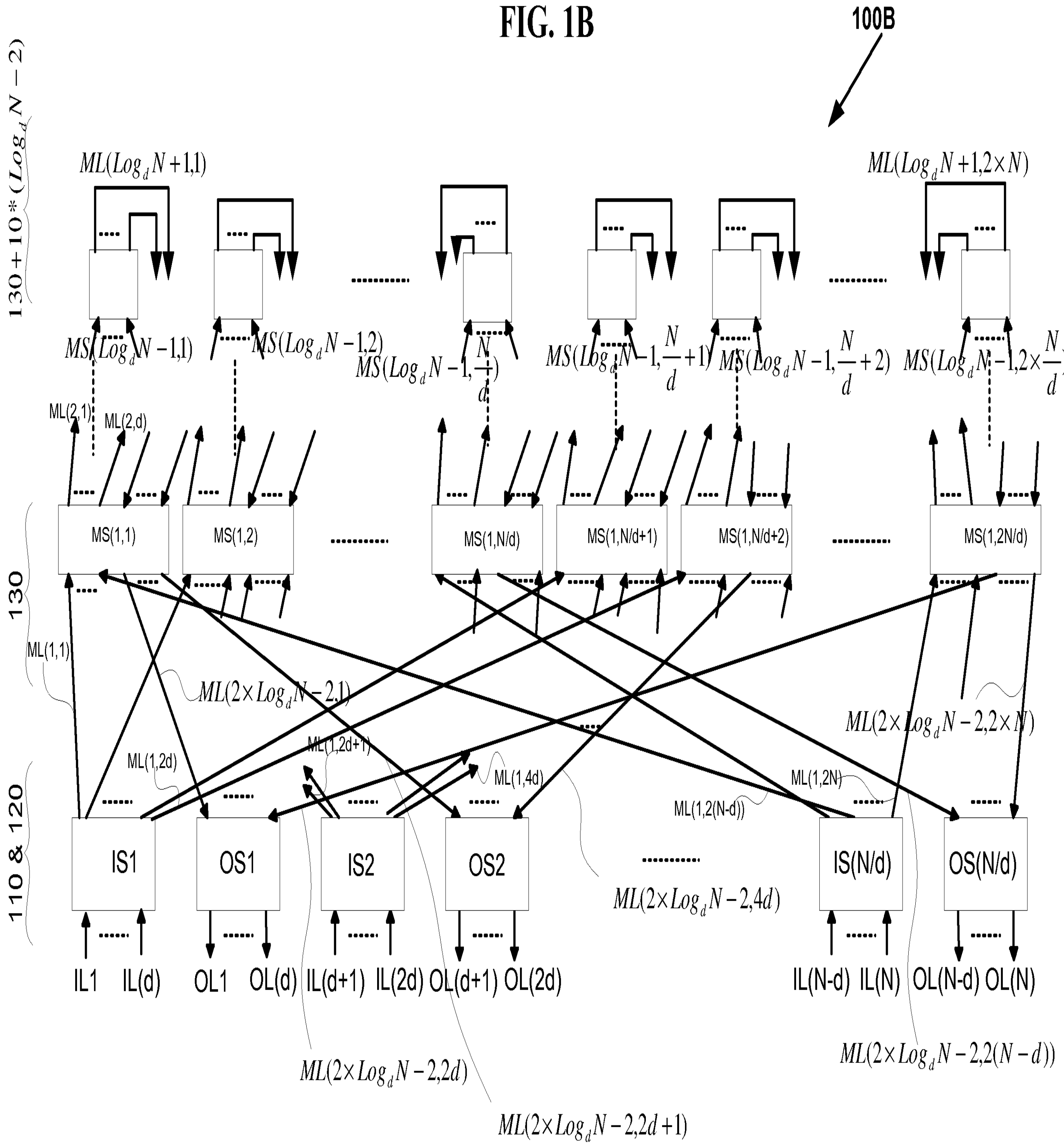

100A
ML(3,1)
ML(3,2)
ML(3,7)
ML(3,8)
ML(3,9)
ML(3,10)
ML(3,15)
ML(3,16)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
MS(2,5)
MS(2,6)
MS(2,7)
MS(2,8)
140
ML(2,3)
ML(2,6)
ML(2,8)
ML(2,9)
ML(2,15)
ML(2,1)
ML(2,16)
ML(2,2)
ML(2,4)
ML(2,5)
ML(2,7)
ML(2,10)
130
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
MS(1,5)
MS(1,6)
MS(1,7)
MS(1,8)
ML(1,1)
ML(4,16)
ML(4,1)
ML(1,16)
110 & 120
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,4)
ML(4,4)
ML(1,13)
ML(4,13)
IL1
IL2
OL1
OL2
IL3
IL4
OL3
OL4
IL5
IL6
OL5
OL6
IL7
IL8
OL7
OL8

FIG. 1B


100B
$130+10*(Log_d N-2)$
130
110 & 120
$ML(Log_d N+1,1)$
$ML(Log_d N+1,2\times N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{d})$
$MS(Log_d N-1,\frac{N}{d}+1)$
$MS(Log_d N-1,\frac{N}{d}+2)$
$MS(Log_d N-1,2\times\frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/d)
MS(1,N/d+1)
MS(1,N/d+2)
MS(1,2N/d)
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,2\times N)$
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2N)
ML(1,2(N-d))
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
$ML(2\times Log_d N-2,4d)$
IL1 IL(d) OL1 OL(d) IL(d+1) IL(2d) OL(d+1) OL(2d)
IL(N-d) IL(N) OL(N-d) OL(N)
$ML(2\times Log_d N-2,2(N-d))$
$ML(2\times Log_d N-2,2d)$
$ML(2\times Log_d N-2,2d+1)$

Fully Connected Generalized Butterfly Fat Tree Networks
Inventor: Venkat Konda
S-0038 PCT

FIG. 1C

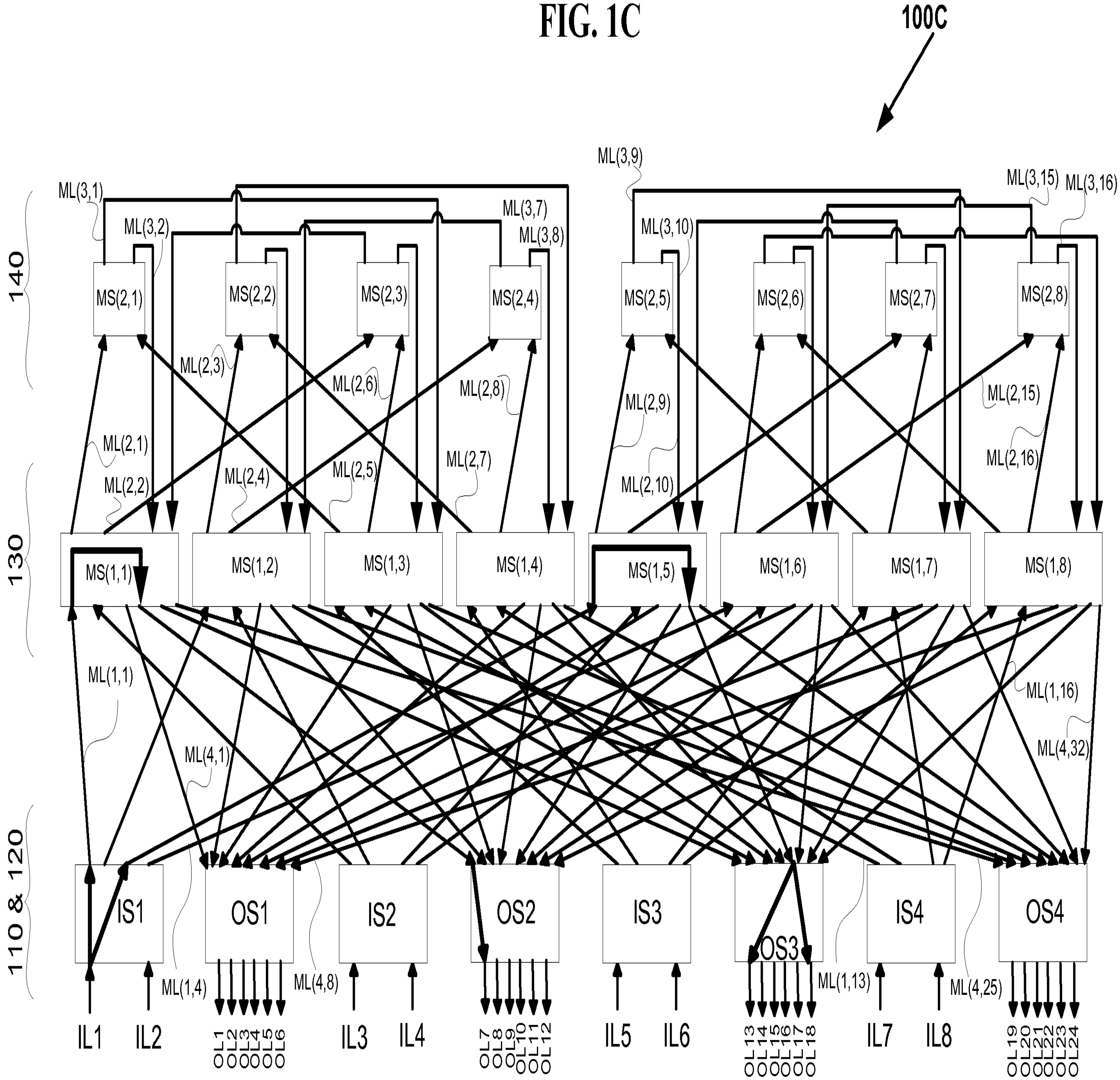


FIG. 1D

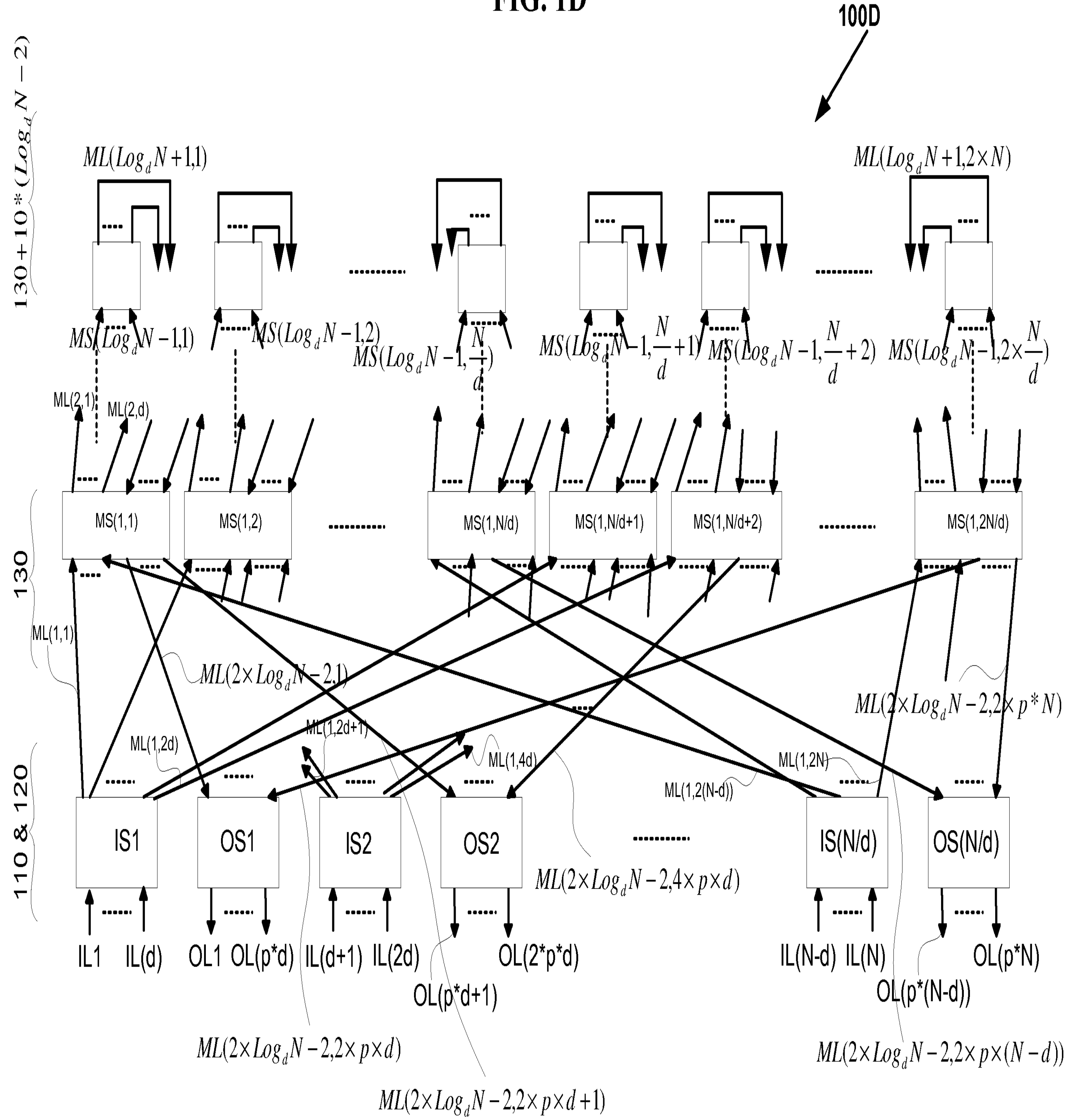

100D
$ML(Log_d N+1,1)$
$ML(Log_d N+1,2\times N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{d})$
$MS(Log_d N-1,\frac{N}{d}+1)$
$MS(Log_d N-1,\frac{N}{d}+2)$
$MS(Log_d N-1,2\times\frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/d)
MS(1,N/d+1)
MS(1,N/d+2)
MS(1,2N/d)
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,2\times p*N)$
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2N)
ML(1,2(N-d))
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
$ML(2\times Log_d N-2,4\times p\times d)$
IL1
IL(d)
OL1
OL(p*d)
IL(d+1)
IL(2d)
OL(2*p*d)
OL(p*d+1)
IL(N-d)
IL(N)
OL(p*N)
OL(p*(N-d))
$ML(2\times Log_d N-2,2\times p\times d)$
$ML(2\times Log_d N-2,2\times p\times d+1)$
$ML(2\times Log_d N-2,2\times p\times (N-d))$
$130+10*(Log_d N-2)$
130
110 & 120

FIG. 1E

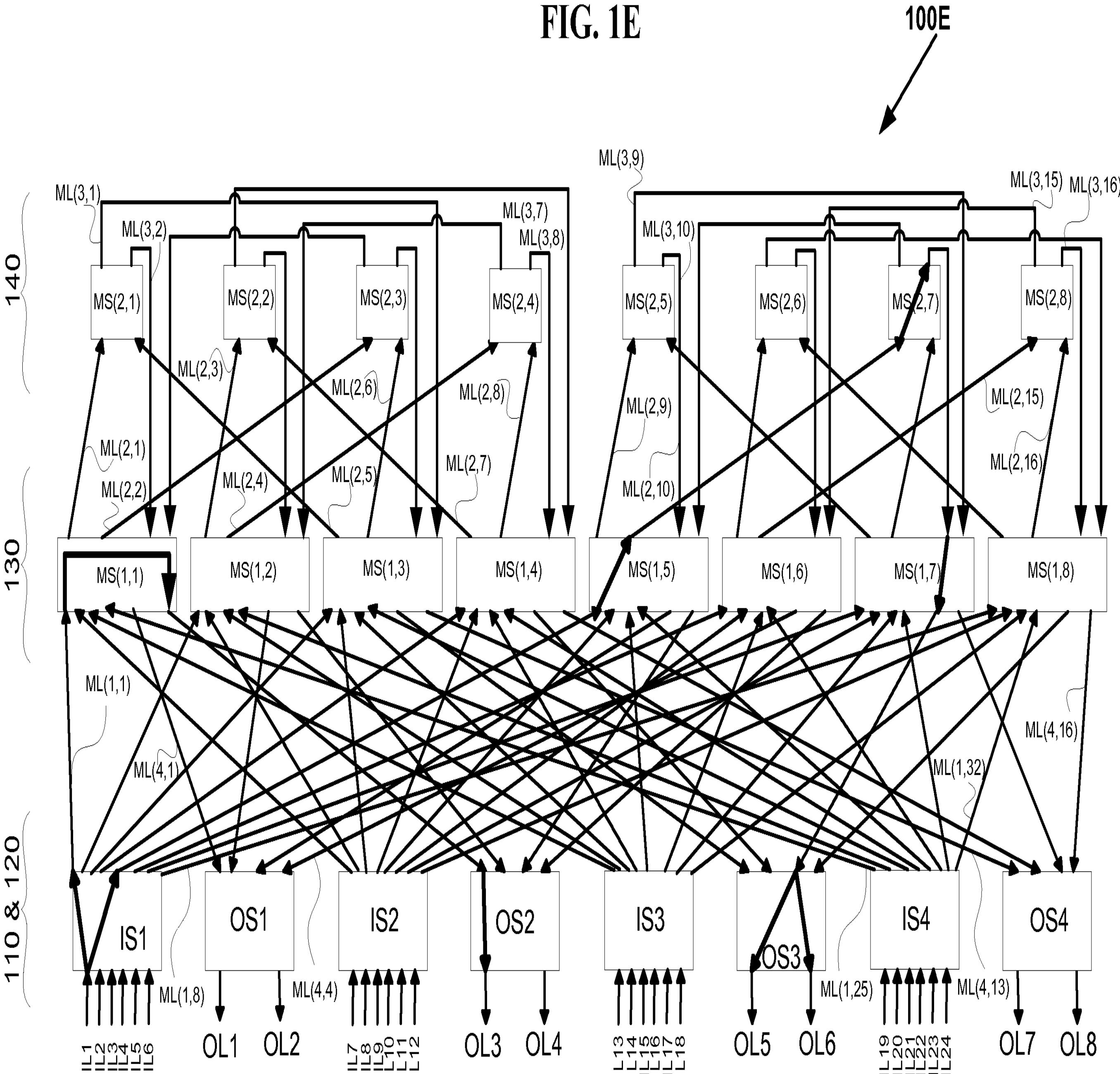

100E
140
130
110 & 120
ML(3,1)
ML(3,2)
ML(3,7)
ML(3,8)
ML(3,9)
ML(3,10)
ML(3,15)
ML(3,16)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
MS(2,5)
MS(2,6)
MS(2,7)
MS(2,8)
ML(2,1)
ML(2,2)
ML(2,3)
ML(2,4)
ML(2,5)
ML(2,6)
ML(2,7)
ML(2,8)
ML(2,9)
ML(2,10)
ML(2,15)
ML(2,16)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
MS(1,5)
MS(1,6)
MS(1,7)
MS(1,8)
ML(1,1)
ML(4,1)
ML(1,32)
ML(4,16)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,8)
ML(4,4)
ML(1,25)
ML(4,13)
IL1 IL2 IL3 IL4 IL5 IL6
OL1 OL2
IL7 IL8 IL9 IL10 IL11 IL12
OL3 OL4
IL13 IL14 IL15 IL16 IL17 IL18
OL5 OL6
IL19 IL20 IL21 IL22 IL23 IL24
OL7 OL8

FIG. 1F

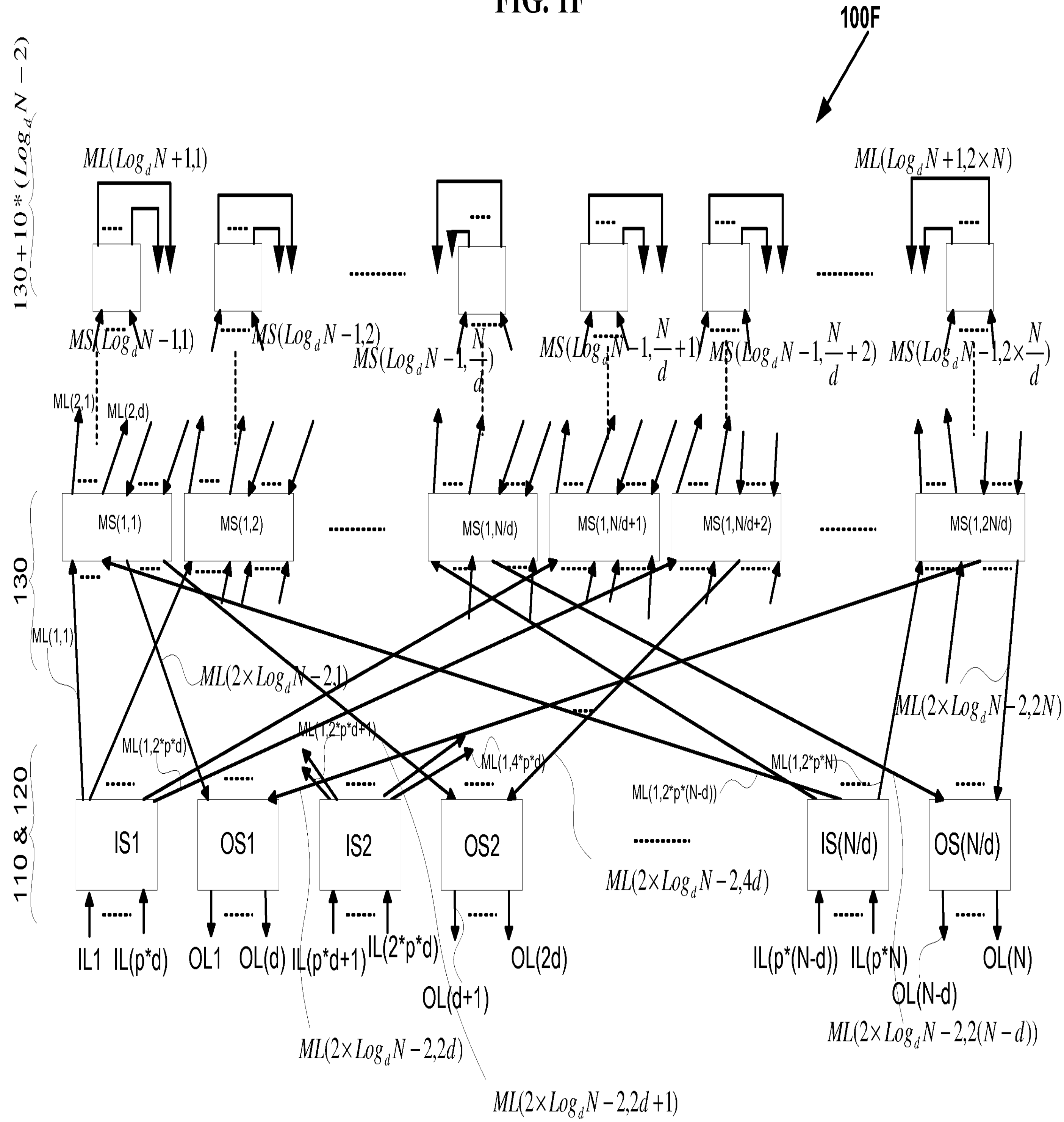

100F
$ML(Log_d N+1,1)$
$ML(Log_d N+1,2\times N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{d})$
$MS(Log_d N-1,\frac{N}{d}+1)$
$MS(Log_d N-1,\frac{N}{d}+2)$
$MS(Log_d N-1,2\times\frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/d)
MS(1,N/d+1)
MS(1,N/d+2)
MS(1,2N/d)
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,2N)$
ML(1,2*p*d)
ML(1,2*p*d+1)
ML(1,4*p*d)
ML(1,2*p*N)
ML(1,2*p*(N-d))
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
$ML(2\times Log_d N-2,4d)$
IL1
IL(p*d)
OL1
OL(d)
IL(p*d+1)
IL(2*p*d)
OL(2d)
OL(d+1)
IL(p*(N-d))
IL(p*N)
OL(N)
OL(N-d)
$ML(2\times Log_d N-2,2(N-d))$
$ML(2\times Log_d N-2,2d)$
$ML(2\times Log_d N-2,2d+1)$
130+10*(Log_d N − 2)
130
110 & 120

FIG. 2A

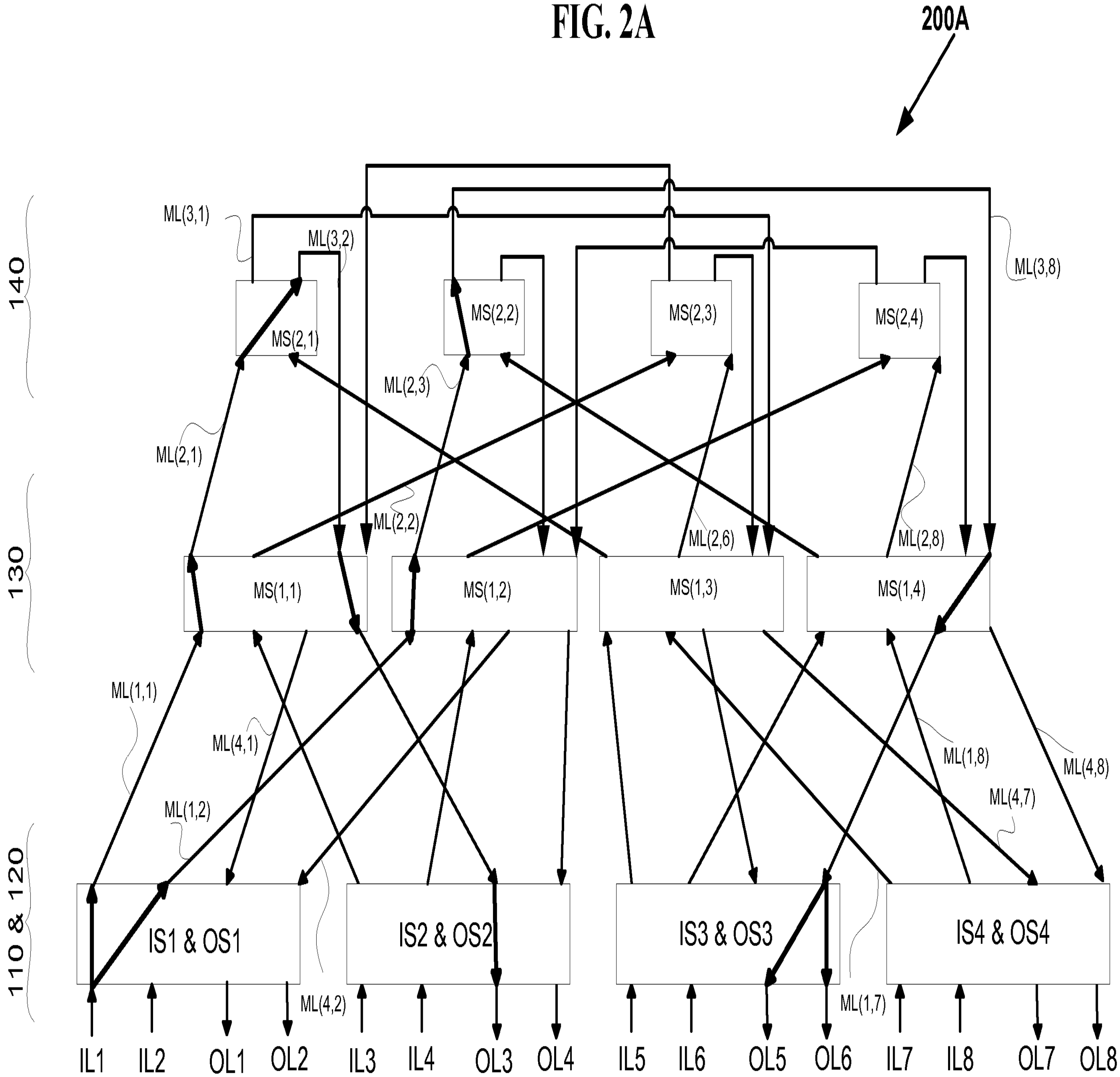

200A
140
130
110 & 120
ML(3,1)
ML(3,2)
ML(3,8)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,3)
ML(2,1)
ML(2,2)
ML(2,6)
ML(2,8)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(4,1)
ML(1,2)
ML(1,8)
ML(4,8)
ML(4,7)
IS1 & OS1
IS2 & OS2
IS3 & OS3
IS4 & OS4
ML(4,2)
ML(1,7)
IL1
IL2
OL1
OL2
IL3
IL4
OL3
OL4
IL5
IL6
OL5
OL6
IL7
IL8
OL7
OL8

FIG. 2B

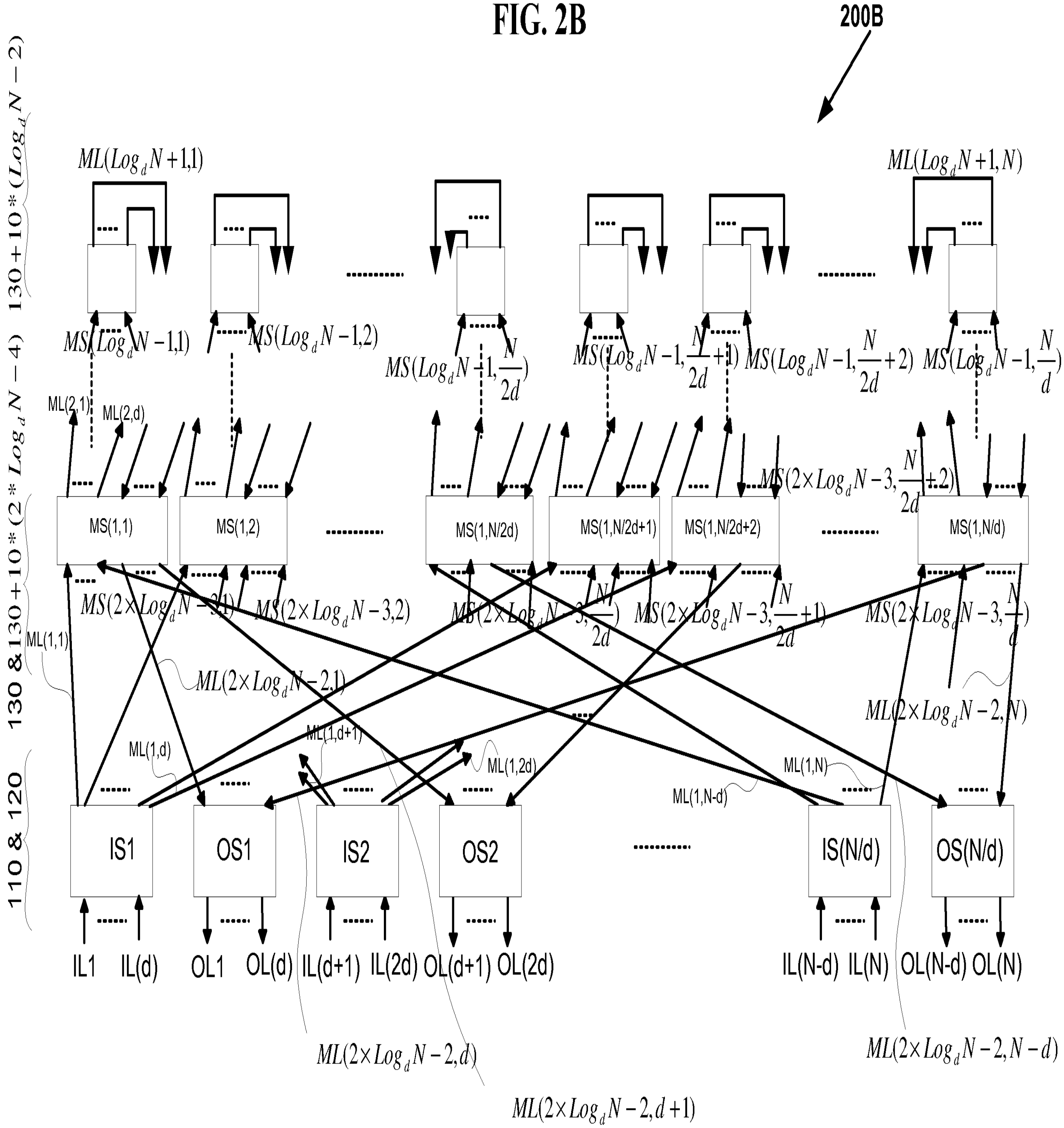

200B
ML(Log_d N+1,1)
ML(Log_d N+1,N)
MS(Log_d N-1,1)
MS(Log_d N-1,2)
MS(Log_d N-1,N/2d)
MS(Log_d N-1,N/2d+1)
MS(Log_d N-1,N/2d+2)
MS(Log_d N-1,N/d)
ML(2,1)
ML(2,d)
MS(2×Log_d N-3,N/2d+2)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
MS(2×Log_d N-3,1)
MS(2×Log_d N-3,2)
MS(2×Log_d N-3,N/2d)
MS(2×Log_d N-3,N/2d+1)
MS(2×Log_d N-3,N/d)
ML(1,1)
ML(2×Log_d N-2,1)
ML(2×Log_d N-2,N)
ML(1,d)
ML(1,d+1)
ML(1,2d)
ML(1,N)
ML(1,N-d)
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
IL1 IL(d) OL1 OL(d) IL(d+1) IL(2d) OL(d+1) OL(2d)
IL(N-d) IL(N) OL(N-d) OL(N)
ML(2×Log_d N-2,d)
ML(2×Log_d N-2,N-d)
ML(2×Log_d N-2,d+1)
110 & 120
130 & 130+10*(2*Log_d N-4)
130+10*(Log_d N-2)

FIG. 2C

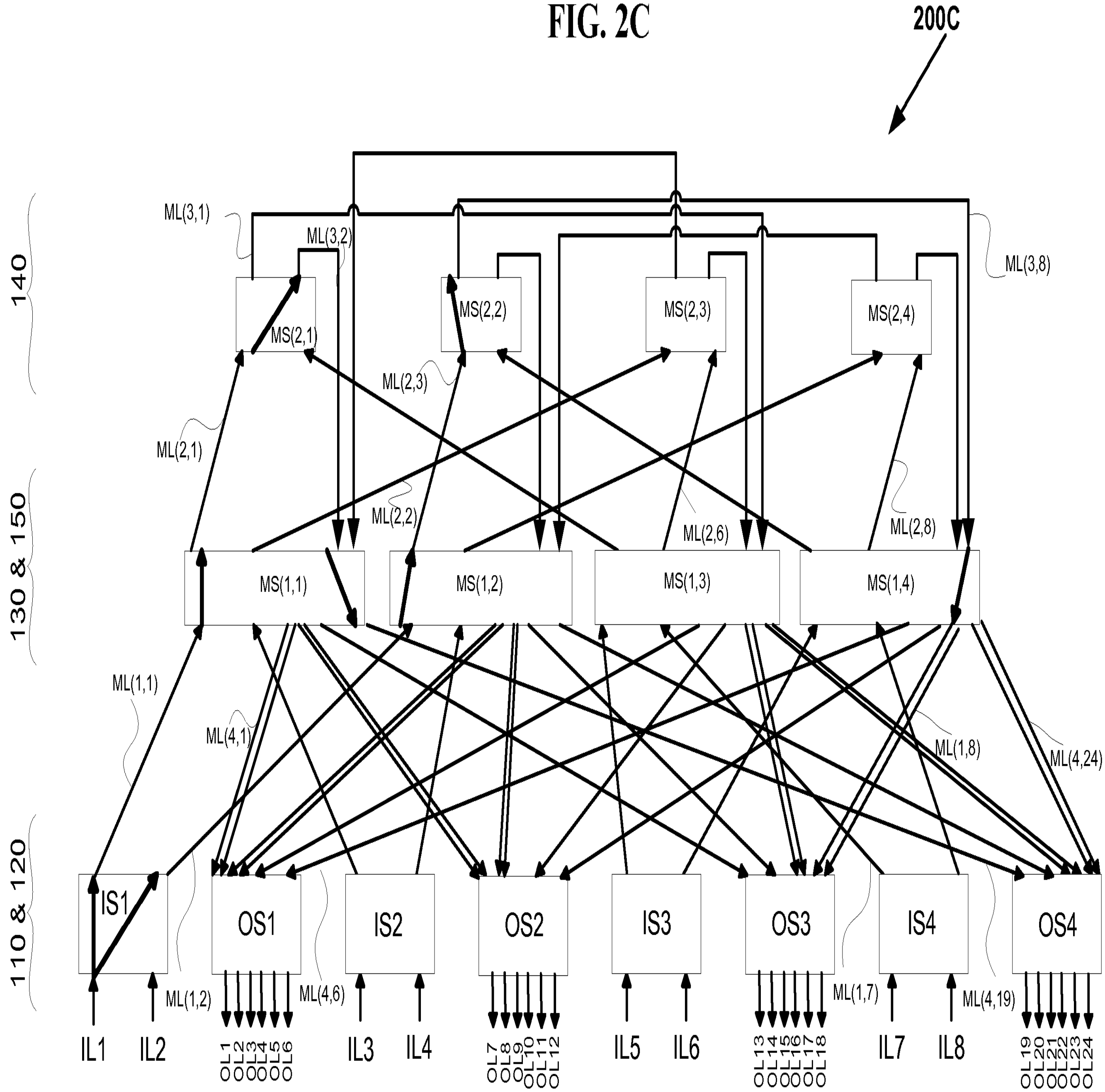

200C
140
130 & 150
110 & 120
ML(3,1)
ML(3,2)
ML(3,8)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,3)
ML(2,1)
ML(2,2)
ML(2,6)
ML(2,8)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(4,1)
ML(1,8)
ML(4,24)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,2)
ML(4,6)
ML(1,7)
ML(4,19)
IL1
IL2
OL1 OL2 OL3 OL4 OL5 OL6
IL3
IL4
OL7 OL8 OL9 OL10 OL11 OL12
IL5
IL6
OL13 OL14 OL15 OL16 OL17 OL18
IL7
IL8
OL19 OL20 OL21 OL22 OL23 OL24

FIG. 2D

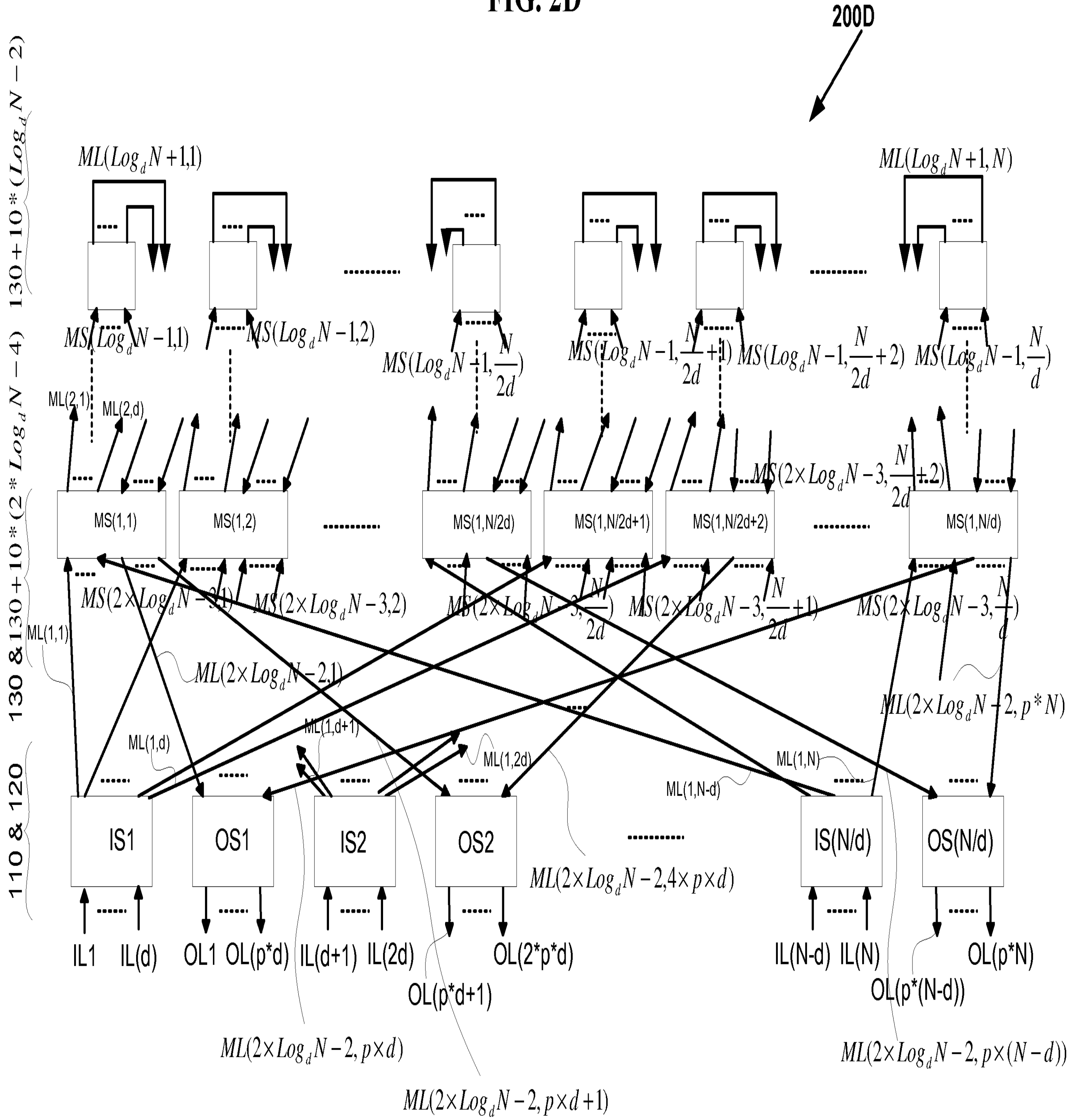

200D
$ML(Log_d N+1,1)$
$ML(Log_d N+1,N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{2d})$
$MS(Log_d N-1,\frac{N}{2d}+1)$
$MS(Log_d N-1,\frac{N}{2d}+2)$
$MS(Log_d N-1,\frac{N}{d})$
ML(2,1)
ML(2,d)
$MS(2\times Log_d N-3,\frac{N}{2d}+2)$
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
$MS(2\times Log_d N-3,1)$
$MS(2\times Log_d N-3,2)$
$MS(2\times Log_d N-3,\frac{N}{2d})$
$MS(2\times Log_d N-3,\frac{N}{2d}+1)$
$MS(2\times Log_d N-3,\frac{N}{d})$
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,p*N)$
ML(1,d)
ML(1,d+1)
ML(1,2d)
ML(1,N)
ML(1,N-d)
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
$ML(2\times Log_d N-2,4\times p\times d)$
IL1
IL(d)
OL1
OL(p*d)
IL(d+1)
IL(2d)
OL(2*p*d)
OL(p*d+1)
IL(N-d)
IL(N)
OL(p*N)
OL(p*(N-d))
$ML(2\times Log_d N-2,p\times d)$
$ML(2\times Log_d N-2,p\times d+1)$
$ML(2\times Log_d N-2,p\times(N-d))$
$130+10*(Log_d N-2)$
$130+10*(2*Log_d N-4)$
130 & 130+10
110 & 120

FIG. 2E

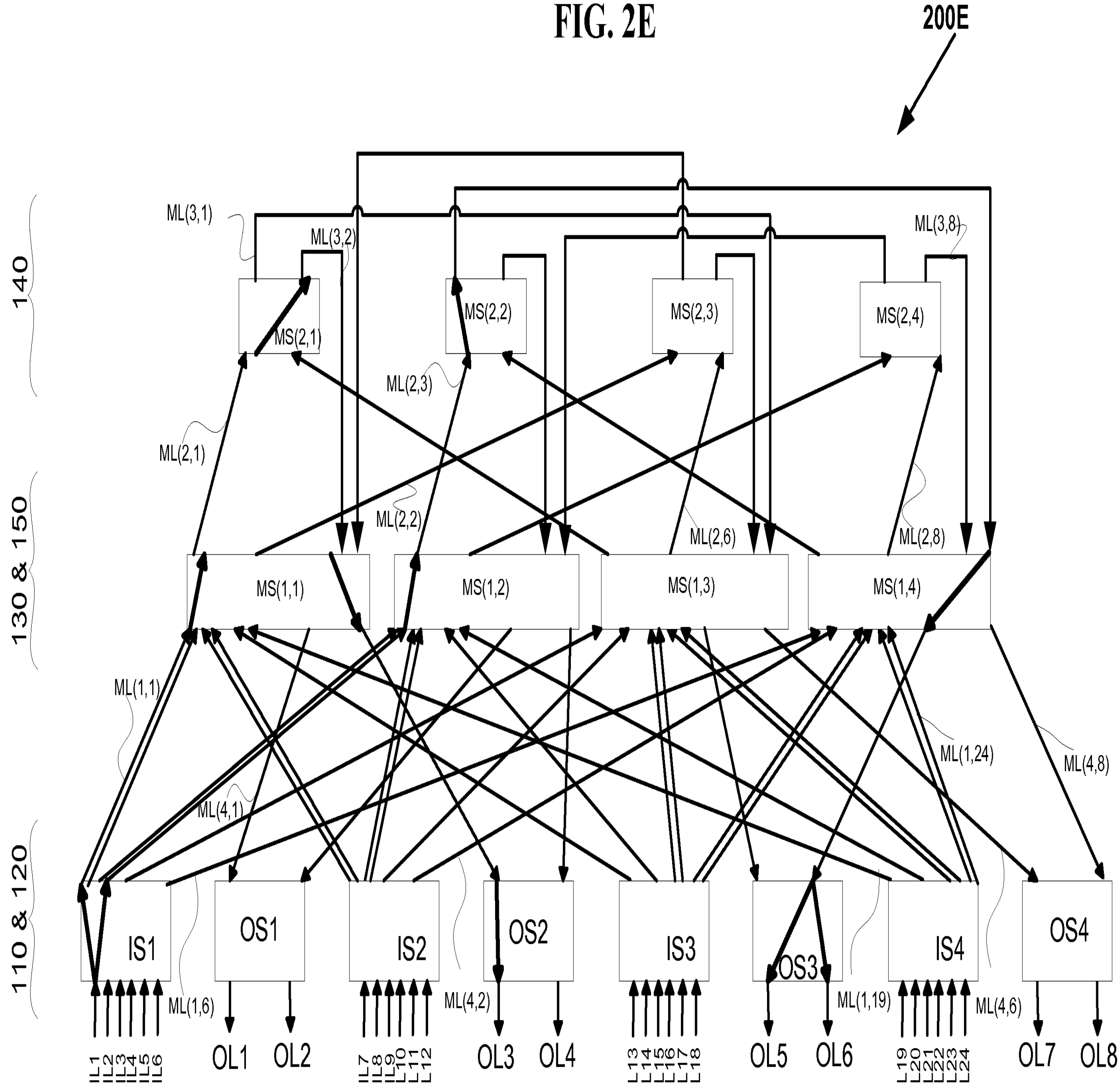

200E
140
130 & 150
110 & 120
ML(3,1)
ML(3,2)
ML(3,8)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,3)
ML(2,1)
ML(2,2)
ML(2,6)
ML(2,8)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(1,24)
ML(4,8)
ML(4,1)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,6)
ML(4,2)
ML(1,19)
ML(4,6)
IL1 IL2 IL3 IL4 IL5 IL6
OL1 OL2
IL7 IL8 IL9 IL10 IL11 IL12
OL3 OL4
IL13 IL14 IL15 IL16 IL17 IL18
OL5 OL6
IL19 IL20 IL21 IL22 IL23 IL24
OL7 OL8

FIG. 2F

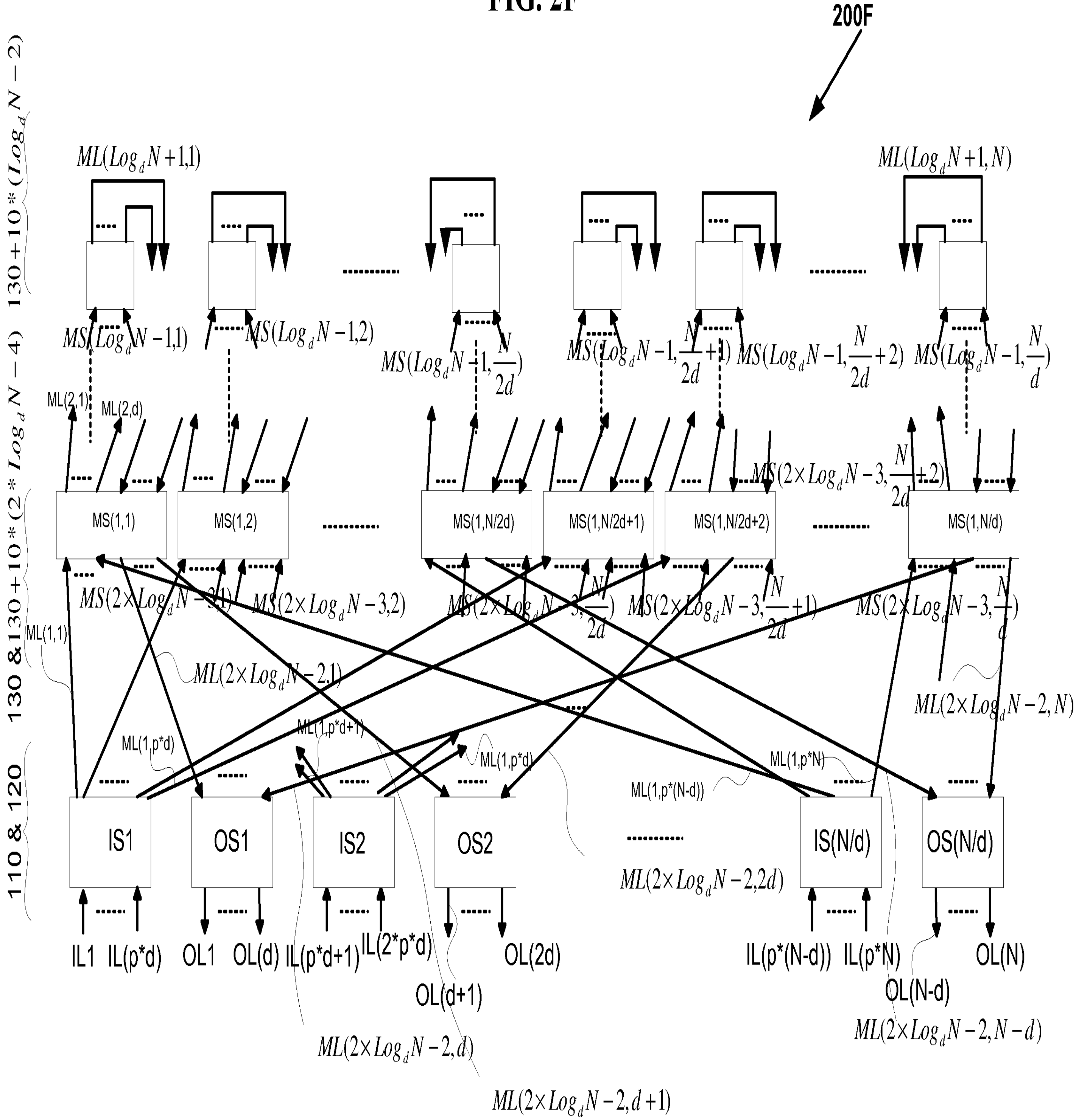

200F
ML(Log_d N+1,1)
ML(Log_d N+1,N)
MS(Log_d N-1,1)
MS(Log_d N-1,2)
MS(Log_d N-1,N/2d)
MS(Log_d N-1,N/2d+1)
MS(Log_d N-1,N/2d+2)
MS(Log_d N-1,N/d)
ML(2,1)
ML(2,d)
MS(2×Log_d N-3,N/2d+2)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
MS(2×Log_d N-3,1)
MS(2×Log_d N-3,2)
MS(2×Log_d N-3,N/2d)
MS(2×Log_d N-3,N/2d+1)
MS(2×Log_d N-3,N/d)
ML(1,1)
ML(2×Log_d N-2,1)
ML(2×Log_d N-2,N)
ML(1,p*d)
ML(1,p*d+1)
ML(1,p*d)
ML(1,p*N)
ML(1,p*(N-d))
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
ML(2×Log_d N-2,2d)
IL1
IL(p*d)
OL1
OL(d)
IL(p*d+1)
IL(2*p*d)
OL(2d)
IL(p*(N-d))
IL(p*N)
OL(N)
OL(d+1)
OL(N-d)
ML(2×Log_d N-2,N-d)
ML(2×Log_d N-2,d)
ML(2×Log_d N-2,d+1)
110 & 120
130 & 130+10*(2*Log_d N-4)
130+10*(Log_d N-2)

FIG. 3A

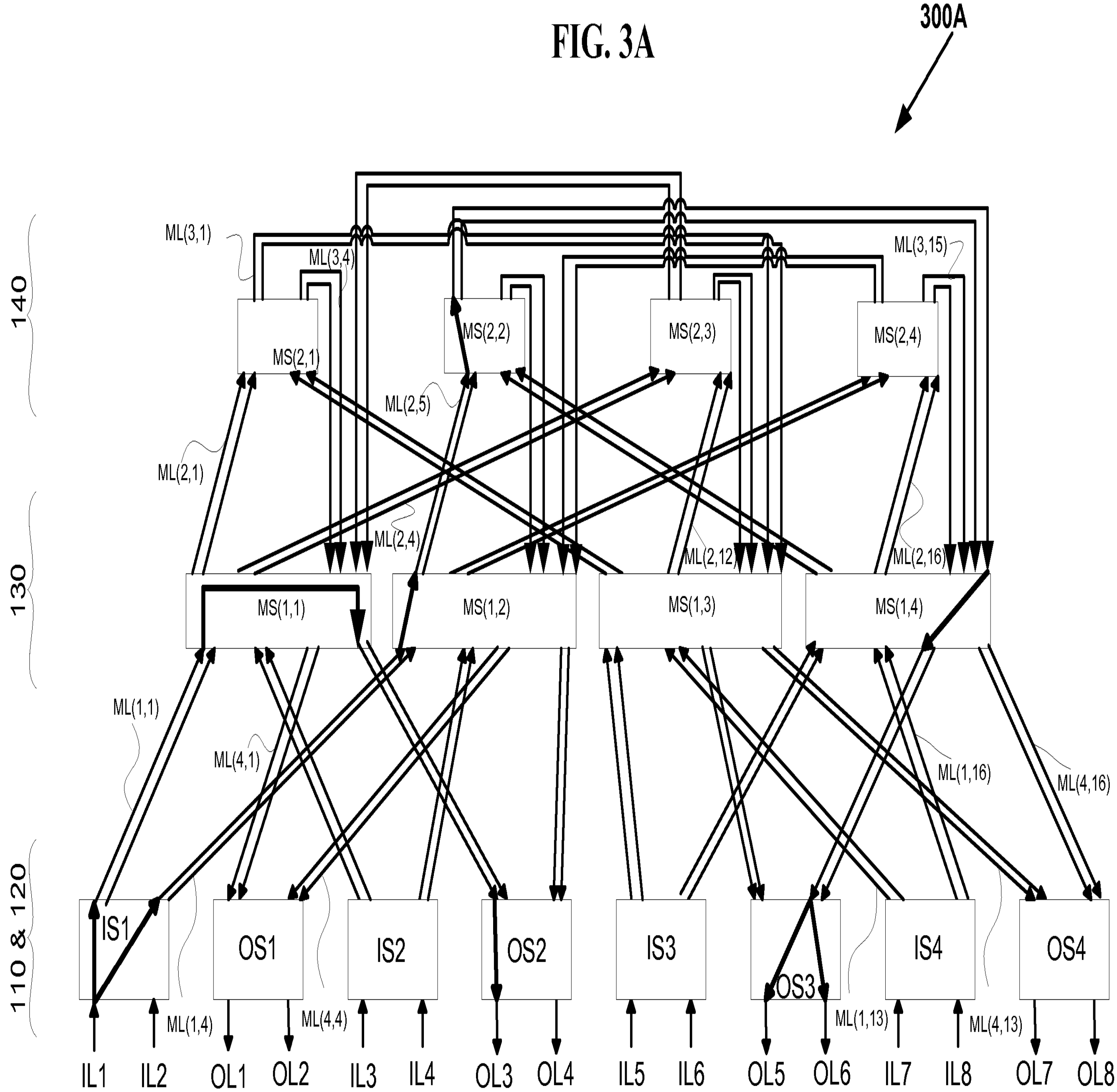

300A
140
130
110 & 120
ML(3,1)
ML(3,4)
ML(3,15)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,5)
ML(2,1)
ML(2,4)
ML(2,12)
ML(2,16)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(4,1)
ML(1,16)
ML(4,16)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,4)
ML(4,4)
ML(1,13)
ML(4,13)
IL1
IL2
OL1
OL2
IL3
IL4
OL3
OL4
IL5
IL6
OL5
OL6
IL7
IL8
OL7
OL8

FIG. 3B

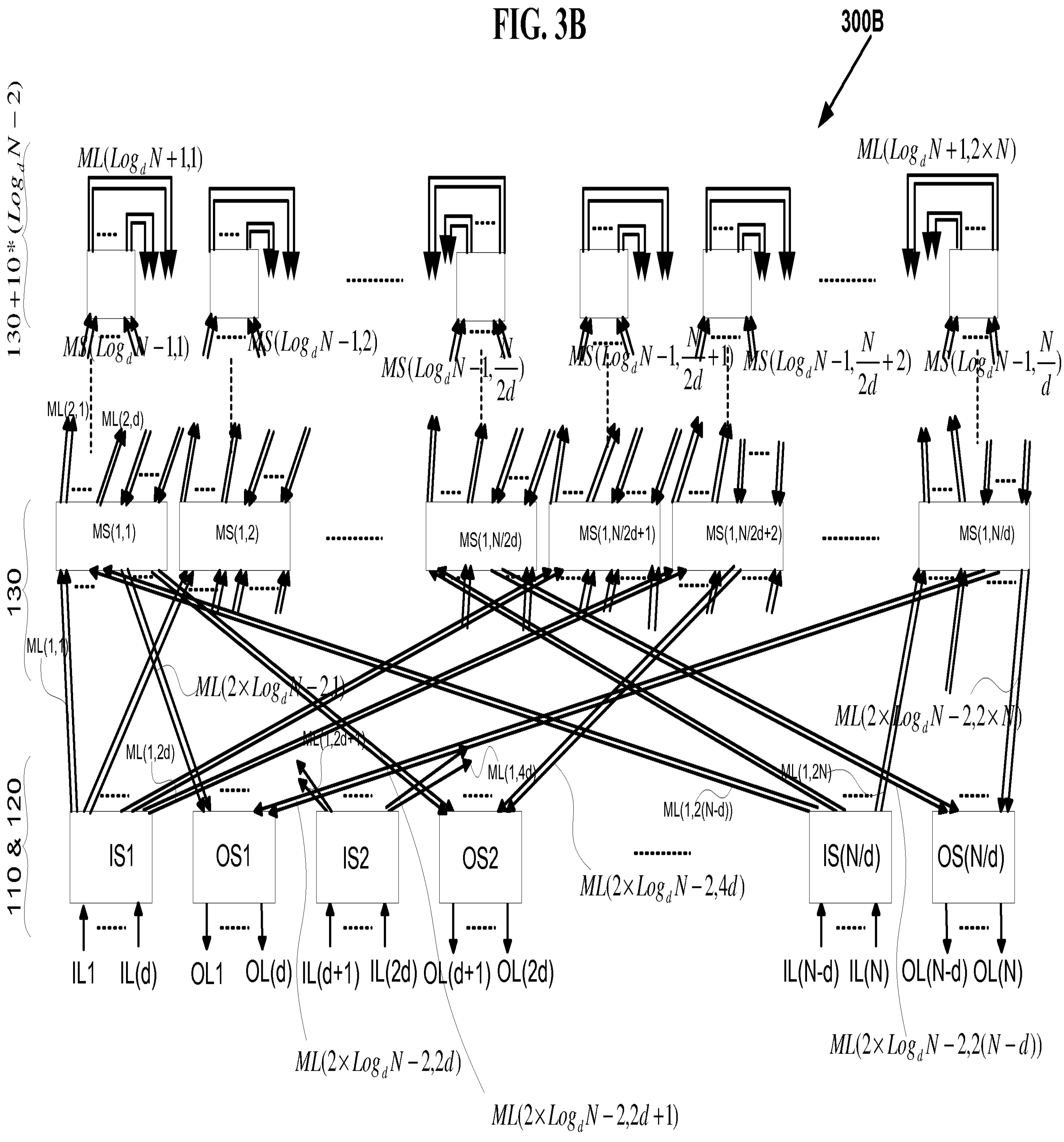

300B
$ML(Log_d N+1,1)$
$ML(Log_d N+1,2\times N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{2d})$
$MS(Log_d N-1,\frac{N}{2d}+1)$
$MS(Log_d N-1,\frac{N}{2d}+2)$
$MS(Log_d N-1,\frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,2\times N)$
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2N)
ML(1,2(N-d))
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
$ML(2\times Log_d N-2,4d)$
IL1 IL(d) OL1 OL(d) IL(d+1) IL(2d) OL(d+1) OL(2d)
IL(N-d) IL(N) OL(N-d) OL(N)
$ML(2\times Log_d N-2,2(N-d))$
$ML(2\times Log_d N-2,2d)$
$ML(2\times Log_d N-2,2d+1)$
$130+10*(Log_d N-2)$
130
110 & 120

FIG. 3C

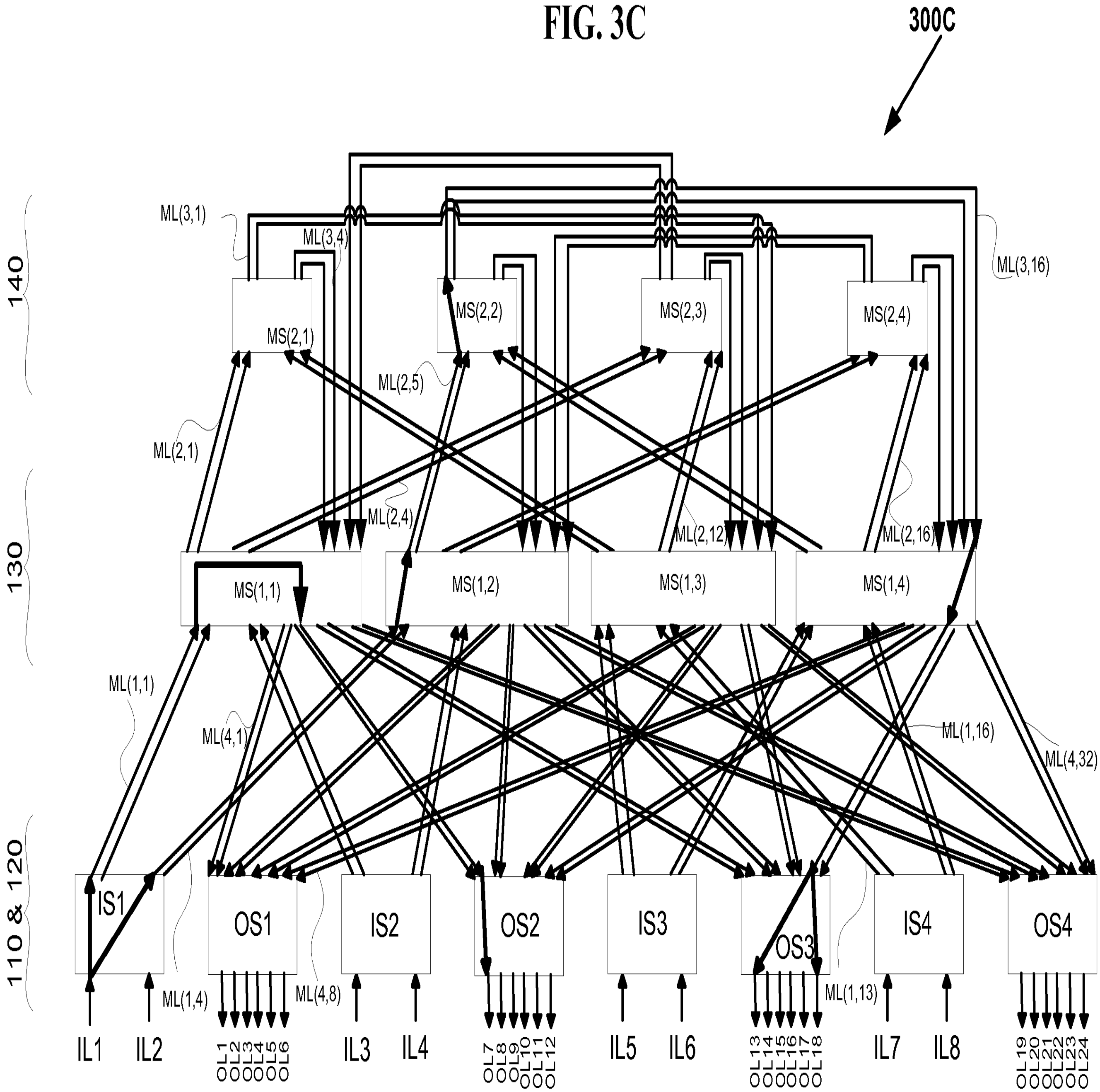

300C
140
130
110 & 120
ML(3,1)
ML(3,4)
ML(3,16)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,5)
ML(2,1)
ML(2,4)
ML(2,12)
ML(2,16)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(4,1)
ML(1,16)
ML(4,32)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,4)
ML(4,8)
ML(1,13)
IL1
IL2
OL1 OL2 OL3 OL4 OL5 OL6
IL3
IL4
OL7 OL8 OL9 OL10 OL11 OL12
IL5
IL6
OL13 OL14 OL15 OL16 OL17 OL18
IL7
IL8
OL19 OL20 OL21 OL22 OL23 OL24

FIG. 3D

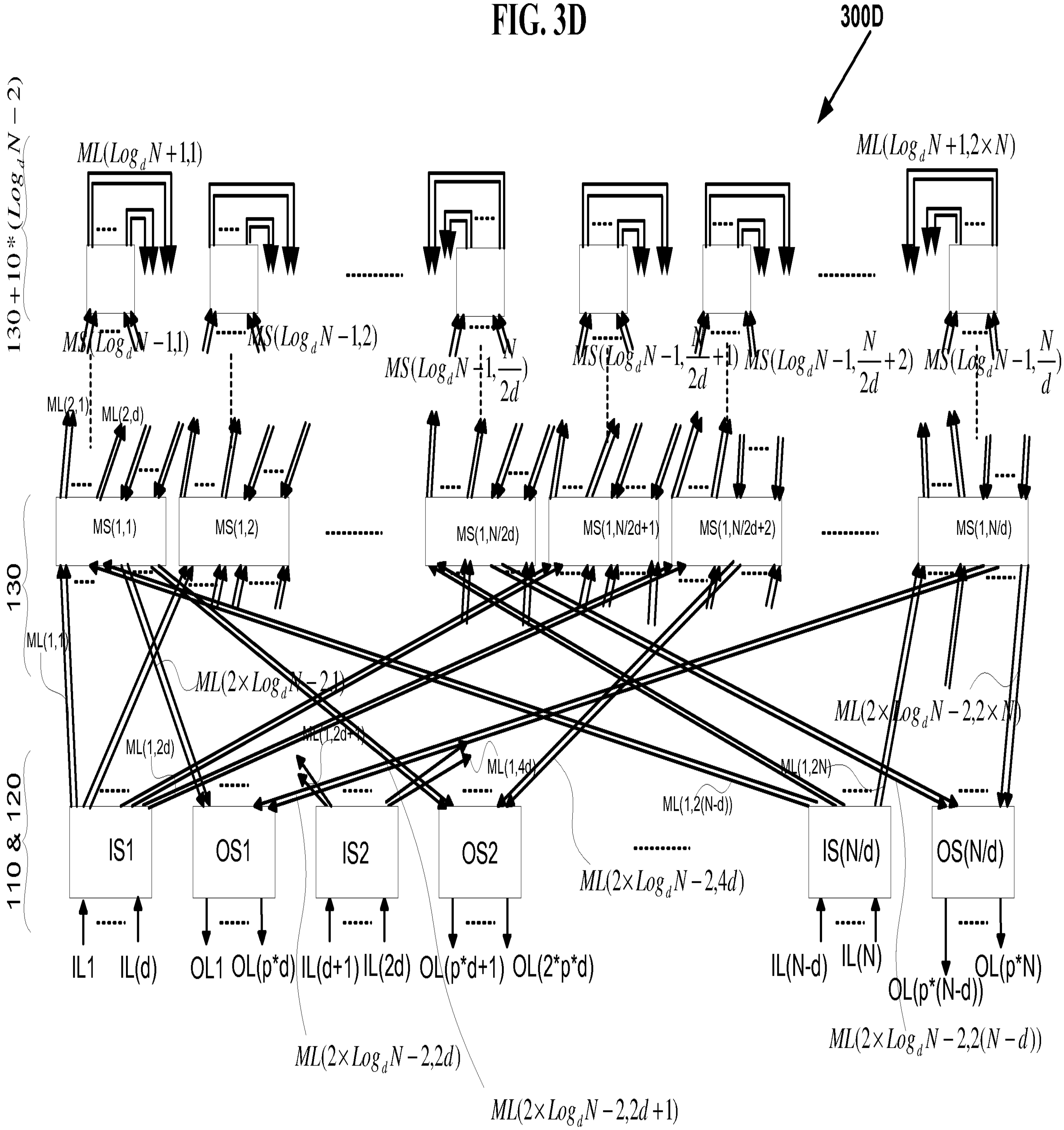

300D
130 + 10 * (Log_d N − 2)
130
110 & 120
ML(Log_d N+1,1)
ML(Log_d N+1,2×N)
MS(Log_d N−1,1)
MS(Log_d N−1,2)
MS(Log_d N−1,N/2d)
MS(Log_d N−1,N/2d+1)
MS(Log_d N−1,N/2d+2)
MS(Log_d N−1,N/d)
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
ML(1,1)
ML(2×Log_d N−2,1)
ML(2×Log_d N−2,2×N)
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2N)
ML(1,2(N-d))
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
ML(2×Log_d N−2,4d)
IL1
IL(d)
OL1
OL(p*d)
IL(d+1)
IL(2d)
OL(p*d+1)
OL(2*p*d)
IL(N-d)
IL(N)
OL(p*(N-d))
OL(p*N)
ML(2×Log_d N−2,2(N−d))
ML(2×Log_d N−2,2d)
ML(2×Log_d N−2,2d+1)

FIG. 3E

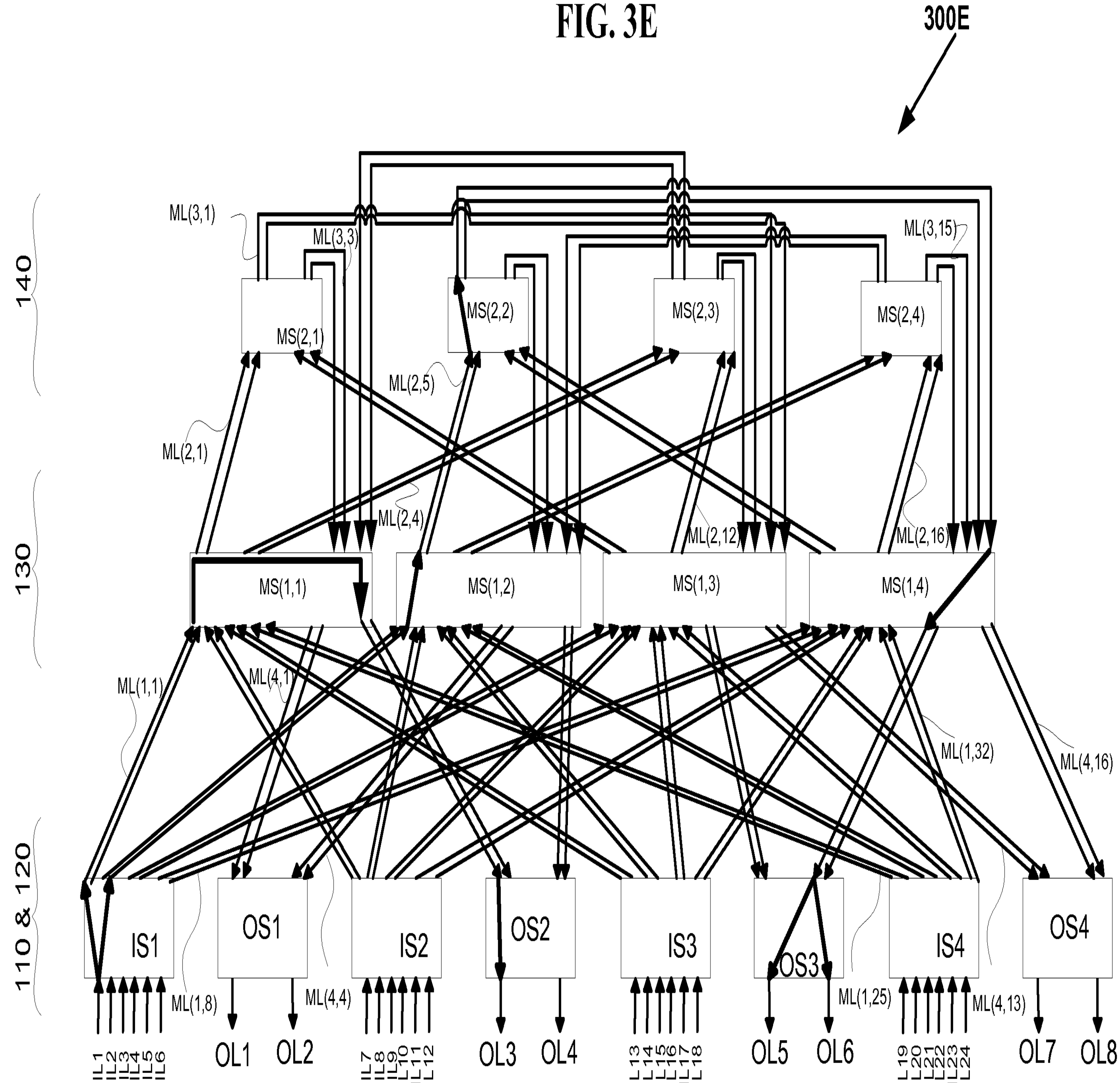

300E
140
130
110 & 120
ML(3,1)
ML(3,3)
ML(3,15)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,5)
ML(2,1)
ML(2,4)
ML(2,12)
ML(2,16)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(4,1)
ML(1,32)
ML(4,16)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,8)
ML(4,4)
ML(1,25)
ML(4,13)
IL1 IL2 IL3 IL4 IL5 IL6
OL1 OL2
IL7 IL8 IL9 IL10 IL11 IL12
OL3 OL4
IL13 IL14 IL15 IL16 IL17 IL18
OL5 OL6
IL19 IL20 IL21 IL22 IL23 IL24
OL7 OL8

FIG. 3F

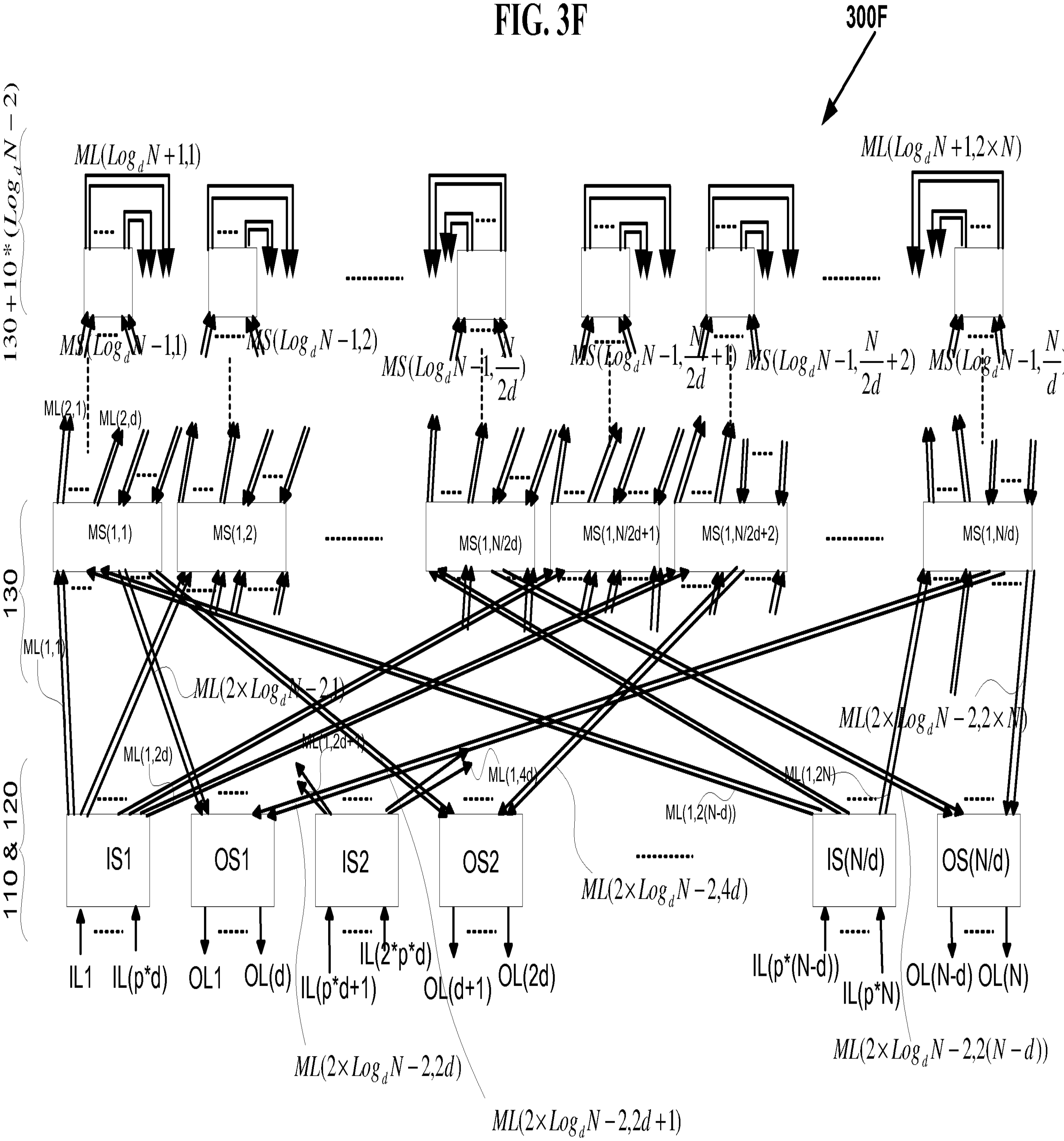

300F
$ML(Log_d N+1,1)$
$ML(Log_d N+1,2\times N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{2d})$
$MS(Log_d N-1,\frac{N}{2d}+1)$
$MS(Log_d N-1,\frac{N}{2d}+2)$
$MS(Log_d N-1,\frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,2\times N)$
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2N)
ML(1,2(N-d))
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
$ML(2\times Log_d N-2,4d)$
IL1
IL(p*d)
OL1
OL(d)
IL(p*d+1)
IL(2*p*d)
OL(d+1)
OL(2d)
IL(p*(N-d))
IL(p*N)
OL(N-d)
OL(N)
$ML(2\times Log_d N-2,2(N-d))$
$ML(2\times Log_d N-2,2d)$
$ML(2\times Log_d N-2,2d+1)$
$130+10*(Log_d N-2)$
130
110 & 120

FIG. 4A

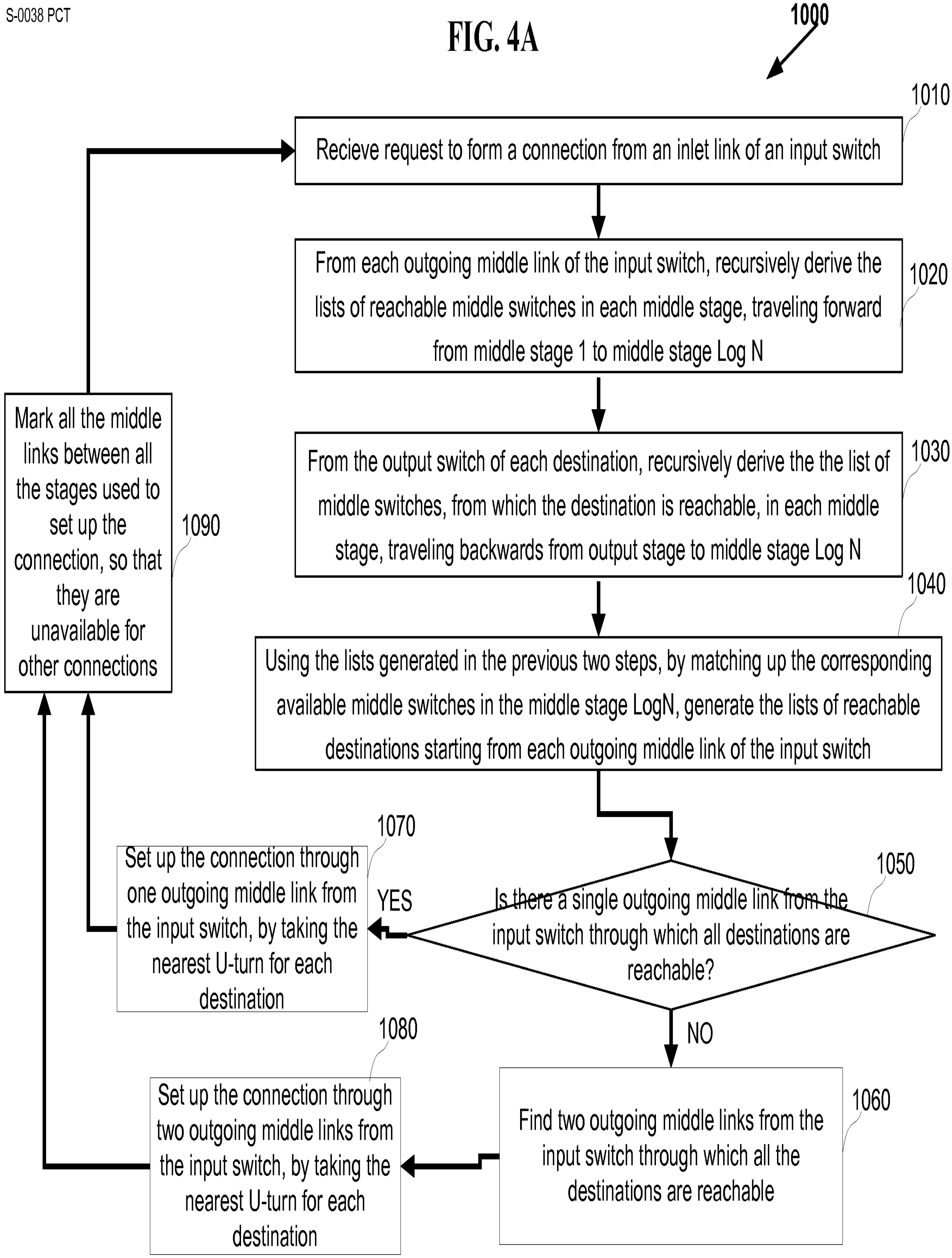

1000
1010
Recieve request to form a connection from an inlet link of an input switch
1020
From each outgoing middle link of the input switch, recursively derive the lists of reachable middle switches in each middle stage, traveling forward from middle stage 1 to middle stage Log N
1030
From the output switch of each destination, recursively derive the the list of middle switches, from which the destination is reachable, in each middle stage, traveling backwards from output stage to middle stage Log N
1040
Using the lists generated in the previous two steps, by matching up the corresponding available middle switches in the middle stage LogN, generate the lists of reachable destinations starting from each outgoing middle link of the input switch
1050
Is there a single outgoing middle link from the input switch through which all destinations are reachable?
YES
NO
1070
Set up the connection through one outgoing middle link from the input switch, by taking the nearest U-turn for each destination
1060
Find two outgoing middle links from the input switch through which all the destinations are reachable
1080
Set up the connection through two outgoing middle links from the input switch, by taking the nearest U-turn for each destination
1090
Mark all the middle links between all the stages used to set up the connection, so that they are unavailable for other connections

FIG. 5A

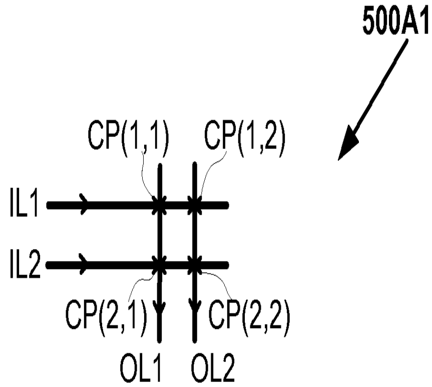


FIG. 5A1
(Prior Art)

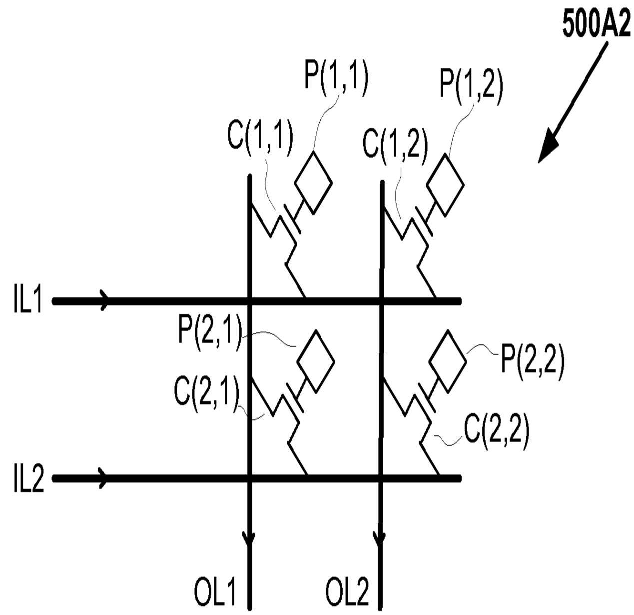


FIG. 5A2
(Prior Art)

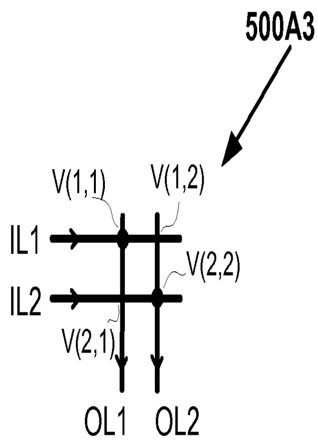


FIG. 5A3
(Prior Art)

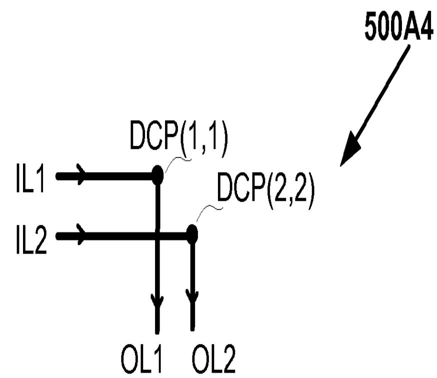


FIG. 5A4

TRANSMITTAL LETTER TO THE UNITED STATES RECEIVING OFFICE

Express Mail mailing number:	Date of deposit: 5/22/2008
File reference no.: S-0038 PCT	International application no. (if known):
Customer Number[1]: 38139	Earliest priority date claimed (Day/Month/Year): 5/25/2007
Title of the invention: Fully Connected Generalized Butterfly Fat Tree Networks	

[1] **Customer Number will allow access to the application in Private PAIR but cannot be used to establish or change the correspondence address.**

☑ **This is a new International Application**

SCREENING DISCLOSURE INFORMATION:

In order to assist in screening the accompanying international application for purposes of determining whether a license for foreign transmittal should and could be granted and for other purposes, the following information is supplied. (check as boxes as apply):

☐ The invention disclose was **not** made in the United States of America.

☐ There is no prior U.S. application relating to this invention.

☑ The following prior U.S. application(s) contain subject matter which is related to the invention disclosed in the attached international application. (*NOTE: priority to these applications may or may not be claimed on the Request (form PCT/RO/101) and this listing does **not** constitute a claim for priority.*)

application no.	60/940, 387	filed on	5/25/2007
application no.	60/940, 390	filed on	5/25/2007

☑ The present international application contains additional subject matter not found in the prior U.S. application(s) identified above. The additional subject matter is found on pages 79-82 and ☑ **DOES NOT ALTER** ☐ **MIGHT BE CONSIDERED TO ALTER** the general nature of the invention in a manner which would require the U.S. application to have been made available for inspection by the appropriate defense agencies under 35 U.S.C. 181 and 37 C.F.R. 5.15.

Itemized list of contents

Sheets of Request form: 3	Check no.:
Sheets of description (excluding sequence listing): 82	Return receipt postcard:
Sheets of claims: 12	Power of attorney:
Sheets of abstract: 1	Certified copy of priority document (specify):
Sheets of drawings: 20	Other (specify):
Sheets of sequence listing:	
Sequence listing diskette/CD:	
Tables related to sequence listing CD:	

The person signing this form is:	☑ Applicant	Venkat Konda Name of person signing
	☐ Attorney/Agent (Reg. No.)	/Venkat Konda/ Signature
	☐ Common Representative	

PCT

REQUEST

The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty.

For receiving Office use only

PCT/US08/64603

International Application No.

22 MAY 2008 (22.05.08)

International Filing Date

PCT INTERNATIONAL APPLICATION RO/US

Name of receiving Office and "PCT International Application"

Applicant's or agent's file reference *(if desired) (12 characters maximum)* S-0038 PCT

Box No. I TITLE OF INVENTION

FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS

Box No. II APPLICANT [X] This person is also inventor

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

Venkat Konda
6278, Grand Oak Way
San Jose, CA
95135

Telephone No. 408-472-3273

Facsimile No. 408-238-2478

Applicant's registration No. with the Office

State *(that is, country)* of nationality: USA

State *(that is, country)* of residence: USA

This person is applicant for the purposes of: [X] all designated States [ ] all designated States except the United States of America [ ] the United States of America only [ ] the States indicated in the Supplemental Box

Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

[ ] Further applicants and/or (further) inventors are indicated on a continuation sheet.

Box No. IV AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE

The person identified below is hereby/has been appointed to act on behalf of the applicant(s) before the competent International Authorities as: [ ] agent [ ] common representative

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)*

Venkat Konda
6278, Grand Oak Way
San Jose, CA
95135

Telephone No. 408-472-3273

Facsimile No. 408-238-2478

Agent's registration No. with the Office

[ ] **Address for correspondence:** Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.

Form PCT/RO/101 (first sheet) (April 2007) *See Notes to the request form*

Box No. V DESIGNATIONS

The filing of this request **constitutes under Rule 4.9(a), the designation** of all Contracting States bound by the PCT on the international filing date, for the grant of every kind of protection available and, where applicable, for the grant of both regional and national patents. However,

☐ DE Germany **is not designated** for any kind of national protection
☐ JP Japan **is not designated** for any kind of national protection
☐ KR Republic of Korea **is not designated** for any kind of national protection
☐ RU Russian Federation **is not designated** for any kind of national protection

(The check-boxes above may only be used to exclude (irrevocably) the designations concerned if, at the time of filing or subsequently under Rule 26bis.1, the international application contains in Box No. VI a priority claim to an earlier national application filed in the particular State concerned, in order to avoid the ceasing of the effect, under the national law, of this earlier national application.)

Box No. VI PRIORITY CLAIM

The priority of the following earlier application(s) is hereby claimed:

Filing date of earlier application *(day/month/year)*	Number of earlier application	Where earlier application is:		
		national application: country or Member of WTO	regional application: regional Office	international application: receiving Office
item (1) 25/5/2007	60/940, 387	US	USPTO	
item (2) 25/5/2007	60/940, 390	US	USPTO	
item (3)				

☐ Further priority claims are indicated in the Supplemental Box.

Transmit certified copy: the receiving Office is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) *(only if the earlier application was filed with the Office which for the purposes of this international application is the receiving Office)* identified above as:

☐ all items ☐ item (1) ☐ item (2) ☐ item (3) ☐ other, see Supplemental Box

Restore the right of priority: the receiving Office is requested to restore the right of priority for the earlier application(s) identified above or in the Supplemental Box as item(s) (________________________). *(See also the Notes to Box No. VI; further information **must** be provided to support a request to restore the right of priority.)*

Incorporation by reference: where an element of the international application referred to in Article 11(1)(iii)(d) or (e) or a part of the description, claims or drawings referred to in Rule 20.5(a) is not otherwise contained in this international application but is completely contained in an earlier application whose priority is claimed on the date on which one or more elements referred to in Article 11(1)(iii) were first received by the receiving Office, that element or part is, subject to confirmation under Rule 20.6, incorporated by reference in this international application for the purposes of Rule 20.6.

Box No. VII INTERNATIONAL SEARCHING AUTHORITY

Choice of International Searching Authority (ISA) *(if two or more International Searching Authorities are competent to carry out the international search, indicate the Authority chosen; the two-letter code may be used)*:

ISA / US . . . .

Request to use results of earlier search; reference to that search *(if an earlier search has been carried out by or requested from the International Searching Authority)*:

Date *(day/month/year)* Number Country *(or regional Office)*

Box No. VIII DECLARATIONS

The following **declarations** are contained in Boxes Nos. VIII (i) to (v) *(mark the applicable check-boxes below and indicate in the right column the number of each type of declaration)*:

	Box	Declaration	Number of declarations
☐	Box No. VIII (i)	Declaration as to the identity of the inventor	:
☐	Box No. VIII (ii)	Declaration as to the applicant's entitlement, as at the international filing date, to apply for and be granted a patent	:
☐	Box No. VIII (iii)	Declaration as to the applicant's entitlement, as at the international filing date, to claim the priority of the earlier application	:
☐	Box No. VIII (iv)	Declaration of inventorship (only for the purposes of the designation of the United States of America)	:
☐	Box No. VIII (v)	Declaration as to non-prejudicial disclosures or exceptions to lack of novelty	:

*RO
**DELETED BY RO

Sheet No. ....

Box No. IX CHECK LIST; LANGUAGE OF FILING

This international application **contains**:

(a) **on paper**, the following number of sheets:

request (including declaration and supplemental sheets) :	*3
description (excluding sequence listing and/or tables related thereto) :	*82
claims :	*12
abstract :	*1
drawings :	*20
Sub-total number of sheets :	*115 / *118
sequence listing :	**
tables related thereto :	**
(for both, actual number of sheets if filed on paper, whether or not also filed in electronic form; see (c) below)	
Total number of sheets :	*115 / *118

(b) ☐ **only in electronic form** (Section 801(a)(i))
- (i) ☐ sequence listing
- (ii) ☐ tables related thereto

(c) ☐ **also in electronic form** (Section 801(a)(ii))
- (i) ☐ sequence listing
- (ii) ☐ tables related thereto

Type and number of carriers (diskette, CD-ROM, CD-R or other) on which are contained the
- ☐ sequence listing: ....
- ☐ tables related thereto: ....

(additional copies to be indicated under items 9(ii) and/or 10(ii), in right column)

This international application is **accompanied by** the following item(s) *(mark the applicable check-boxes below and indicate in right column the number of each item)*: — Number of items

1. ☒ fee calculation sheet :
2. ☐ original separate power of attorney :
3. ☐ original general power of attorney :
4. ☐ copy of general power of attorney; reference number, if any: .... :
5. ☐ statement explaining lack of signature :
6. ☐ priority document(s) identified in Box No. VI as item(s): .... :
7. ☐ translation of international application into *(language)*: .... :
8. ☐ separate indications concerning deposited microorganism or other biological material :
9. ☐ sequence listing in electronic form *(indicate type and number of carriers)*
   - (i) ☐ copy submitted for the purposes of international search under Rule 13*ter* only (and not as part of the international application) :
   - (ii) ☐ *(only where check-box (b)(i) or (c)(i) is marked in left column)* additional copies including, where applicable, the copy for the purposes of international search under Rule 13*ter* :
   - (iii) ☐ together with relevant statement as to the identity of the copy or copies with the sequence listing mentioned in left column :
10. ☐ tables in electronic form related to sequence listing *(indicate type and number of carriers)*
    - (i) ☐ copy submitted for the purposes of international search under Section 802(b-*quater*) only (and not as part of the international application) :
    - (ii) ☐ *(only where check-box (b)(ii) or (c)(ii) is marked in left column)* additional copies including, where applicable, the copy for the purposes of international search under Section 802(b-*quater*) :
    - (iii) ☐ together with relevant statement as to the identity of the copy or copies with the tables mentioned in left column :
11. ☐ other *(specify)*: .... :

Figure of the drawings which should accompany the abstract: FIG. 1B

Language of filing of the international application: English

Box No. X SIGNATURE OF APPLICANT, AGENT OR COMMON REPRESENTATIVE

Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the request)

/Venkat Konda/

For receiving Office use only

1. Date of actual receipt of the purported international application: 22 MAY 2008
2. Drawings: ☐ received: ☐ not received:
3. Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application:
4. Date of timely receipt of the required corrections under PCT Article 11(2):
5. International Searching Authority (if two or more are competent): ISA / US
6. ☐ Transmittal of search copy delayed until search fee is paid

For International Bureau use only

Date of receipt of the record copy by the International Bureau:

Form PCT/RO/101 (last sheet) (April 2007) — *See Notes to the request form*

This sheet is not part of and does not count as a sheet of the international application.

PCT

FEE CALCULATION SHEET

Annex to the Request

For receiving Office use only

PCT/US08/64603

International Application No.

22 MAY 2008 (22.05.08)

Date stamp of the receiving Office

Applicant's or agent's file reference: S-0038 PCT

Applicant

CALCULATION OF PRESCRIBED FEES

1. TRANSMITTAL FEE . . . . $300.00 [T]

2. SEARCH FEE . . . . $1800.00 [S]

   International search to be carried out by ....

   (If two or more International Searching Authorities are competent to carry out the international search, indicate the name of the Authority which is chosen to carry out the international search.)

3. INTERNATIONAL FILING FEE

   Where items (b) and/or (c) of Box No. IX apply, enter **Sub-total number of sheets**
   Where items (b) and (c) of Box No. IX do not apply, enter **Total number of sheets** } ....

   [i1] first 30 sheets . . . . $1194.00 [i1]

   [i2] 88 (number of sheets in excess of 30) x 13 (fee per sheet) = $1144.00 [i2]

   [i3] additional component (only if a sequence listing and/or tables related thereto are filed in electronic form under Section 801(a)(i), or both in that form and on paper, under Section 801(a)(ii)):

   400 x .... (fee per sheet) = ... [i3]

   Add amounts entered at i1, i2 and i3 and enter total at I . . . . [I]

   (Applicants from certain States are entitled to a reduction of 75% of the international filing fee. Where the applicant is (or all applicants are) so entitled, the total to be entered at I is 25% of the international filing fee.)

4. FEE FOR PRIORITY DOCUMENT *(if applicable)* . . . . [P]

5. TOTAL FEES PAYABLE . . . . $4438.00 TOTAL

   Add amounts entered at T, S, I and P, and enter total in the TOTAL box

MODE OF PAYMENT *(Not all modes of payment may be available at all receiving Offices)*

- ☐ authorization to charge deposit account (see below)
- ☐ postal money order
- ☐ cash
- ☐ coupons
- ☐ cheque
- ☐ bank draft
- ☐ revenue stamps
- ☒ other *(specify)*: Credit Card

AUTHORIZATION TO CHARGE (OR CREDIT) DEPOSIT ACCOUNT

(This mode of payment may not be available at all receiving Offices)

- ☐ Authorization to charge the total fees indicated above.
- ☐ *(This check-box may be marked only if the conditions for deposit accounts of the receiving Office so permit)* Authorization to charge any deficiency or credit any overpayment in the total fees indicated above.
- ☐ Authorization to charge the fee for priority document.

Receiving Office: RO/ ....

Deposit Account No.: ....

Date: ....

Name: ....

Signature: ....

Form PCT/RO/101 (Annex) (April 2007)

See Notes to the fee calculation sheet

This sheet is not part of and does not count as a sheet of the international application.

PCT

FEE CALCULATION SHEET
Annex to the Request

For receiving Office use only

International Application No.

Date stamp of the receiving Office

Applicant's or agent's file reference S-0038 PCT

Applicant

CALCULATION OF PRESCRIBED FEES

1. TRANSMITTAL FEE . . . . $300.00 [T]

2. SEARCH FEE . . . . $1800.00 [S]

International search to be carried out by ____________

(If two or more International Searching Authorities are competent to carry out the international search, indicate the name of the Authority which is chosen to carry out the international search.)

3. INTERNATIONAL FILING FEE

Where items (b) and/or (c) of Box No. IX apply, enter **Sub-total number of sheets**
Where items (b) and (c) of Box No. IX do not apply, enter **Total number of sheets** } ________

[i1] first 30 sheets . . . . $1194.00 [i1]

[i2] 88 (number of sheets in excess of 30) x 13 (fee per sheet) = $1144.00 [i2]

[i3] additional component (only if a sequence listing and/or tables related thereto are filed in electronic form under Section 801(a)(i), or both in that form and on paper, under Section 801(a)(ii)):

400 x ________ (fee per sheet) = ________ [i3]

Add amounts entered at i1, i2 and i3 and enter total at I . . . . ________ [I]

(Applicants from certain States are entitled to a reduction of 75% of the international filing fee. Where the applicant is (or all applicants are) so entitled, the total to be entered at I is 25% of the international filing fee.)

4. FEE FOR PRIORITY DOCUMENT *(if applicable)* . . . . ________ [P]

5. TOTAL FEES PAYABLE . . . . $4438.00 TOTAL

Add amounts entered at T, S, I and P, and enter total in the TOTAL box

MODE OF PAYMENT *(Not all modes of payment may be available at all receiving Offices)*

- ☐ authorization to charge deposit account (see below)
- ☐ postal money order
- ☐ cash
- ☐ coupons
- ☐ cheque
- ☐ bank draft
- ☐ revenue stamps
- ☒ other *(specify)*: Credit Card

AUTHORIZATION TO CHARGE (OR CREDIT) DEPOSIT ACCOUNT
(This mode of payment may not be available at all receiving Offices)

- ☐ Authorization to charge the total fees indicated above.
- ☐ *(This check-box may be marked only if the conditions for deposit accounts of the receiving Office so permit)* Authorization to charge any deficiency or credit any overpayment in the total fees indicated above.
- ☐ Authorization to charge the fee for priority document.

Receiving Office: RO/ ________

Deposit Account No.: ________

Date: ________

Name: ________

Signature: ________

Form PCT/RO/101 (Annex) (April 2007)

See Notes to the fee calculation sheet

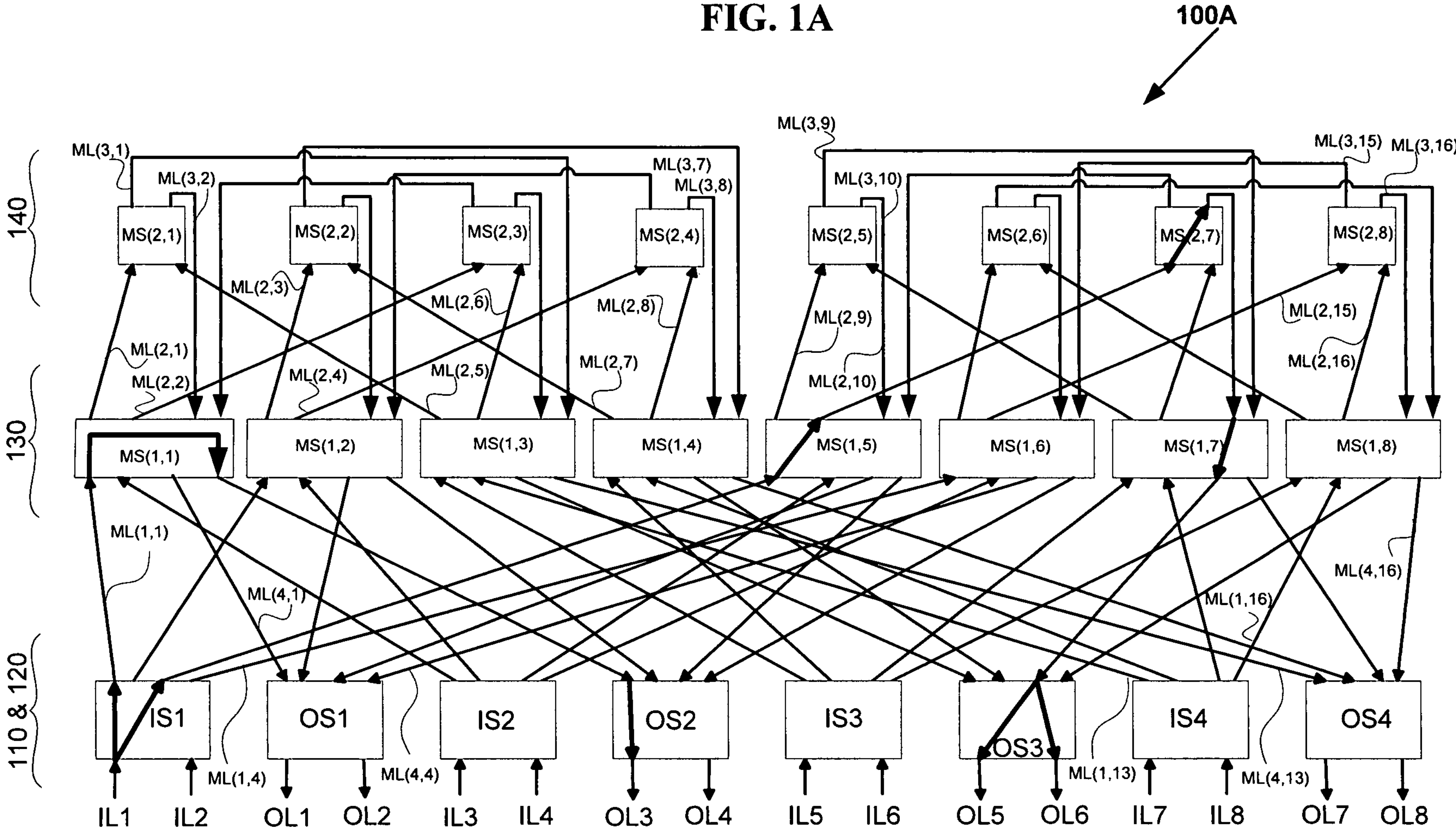

100A
140
130
110 & 120
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
MS(2,5)
MS(2,6)
MS(2,7)
MS(2,8)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
MS(1,5)
MS(1,6)
MS(1,7)
MS(1,8)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(3,1)
ML(3,2)
ML(3,7)
ML(3,8)
ML(3,9)
ML(3,10)
ML(3,15)
ML(3,16)
ML(2,1)
ML(2,2)
ML(2,3)
ML(2,4)
ML(2,5)
ML(2,6)
ML(2,7)
ML(2,8)
ML(2,9)
ML(2,10)
ML(2,15)
ML(2,16)
ML(1,1)
ML(4,1)
ML(1,4)
ML(4,4)
ML(1,13)
ML(4,13)
ML(1,16)
ML(4,16)
IL1
IL2
OL1
OL2
IL3
IL4
OL3
OL4
IL5
IL6
OL5
OL6
IL7
IL8
OL7
OL8

FIG. 1A

FIG. 1B

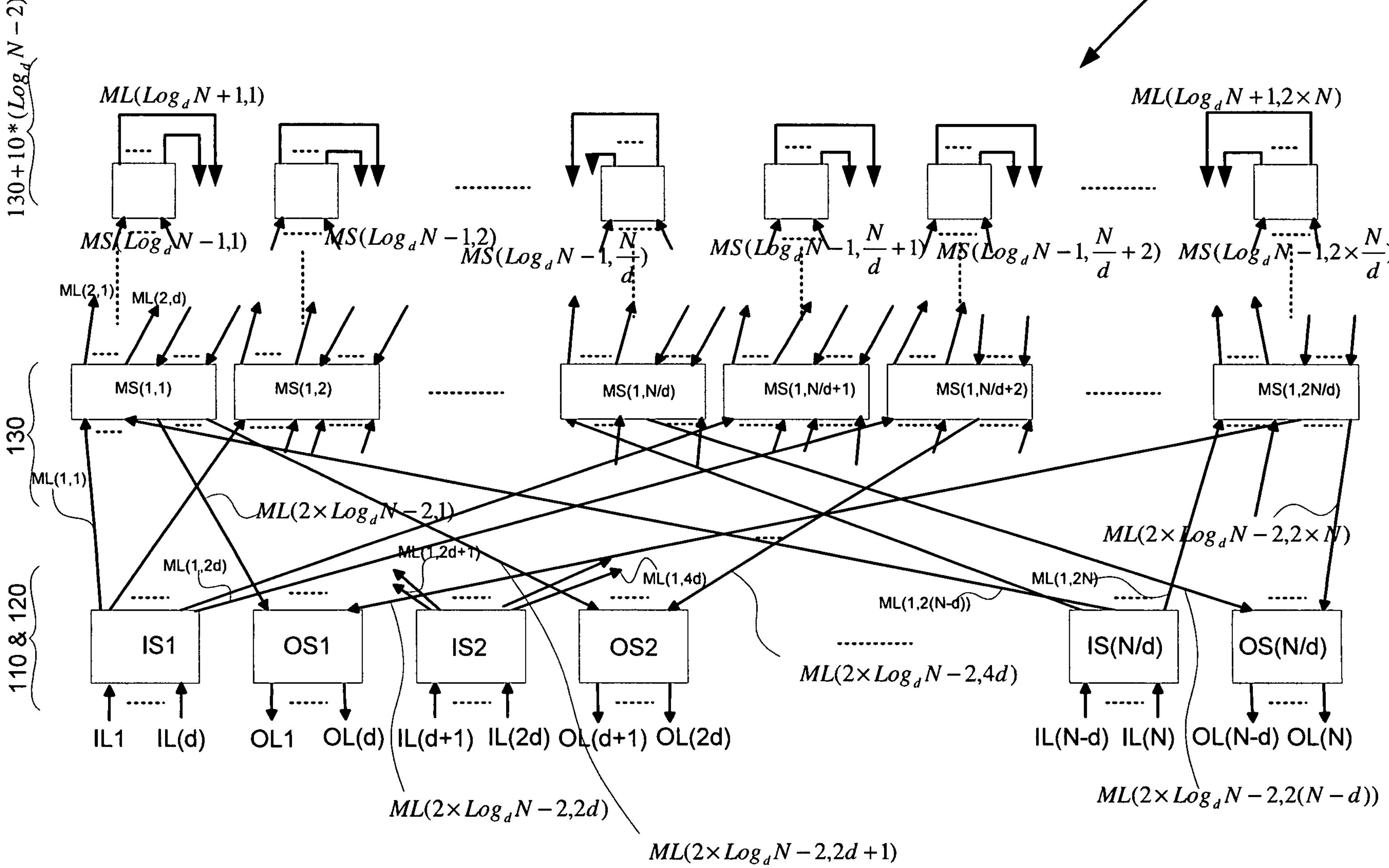

100B
$ML(Log_d N + 1,1)$
$ML(Log_d N + 1,2 \times N)$
$MS(Log_d N - 1,1)$
$MS(Log_d N - 1,2)$
$MS(Log_d N - 1, \frac{N}{d})$
$MS(Log_d N - 1, \frac{N}{d} + 1)$
$MS(Log_d N - 1, \frac{N}{d} + 2)$
$MS(Log_d N - 1,2 \times \frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/d)
MS(1,N/d+1)
MS(1,N/d+2)
MS(1,2N/d)
ML(1,1)
$ML(2 \times Log_d N - 2,1)$
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2N)
ML(1,2(N-d))
$ML(2 \times Log_d N - 2,2 \times N)$
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
$ML(2 \times Log_d N - 2,4d)$
IL1
IL(d)
OL1
OL(d)
IL(d+1)
IL(2d)
OL(d+1)
OL(2d)
IL(N-d)
IL(N)
OL(N-d)
OL(N)
$ML(2 \times Log_d N - 2,2d)$
$ML(2 \times Log_d N - 2,2d + 1)$
$ML(2 \times Log_d N - 2,2(N - d))$
$130 + 10 * (Log_d N - 2)$
130
110 & 120

FIG. 1C

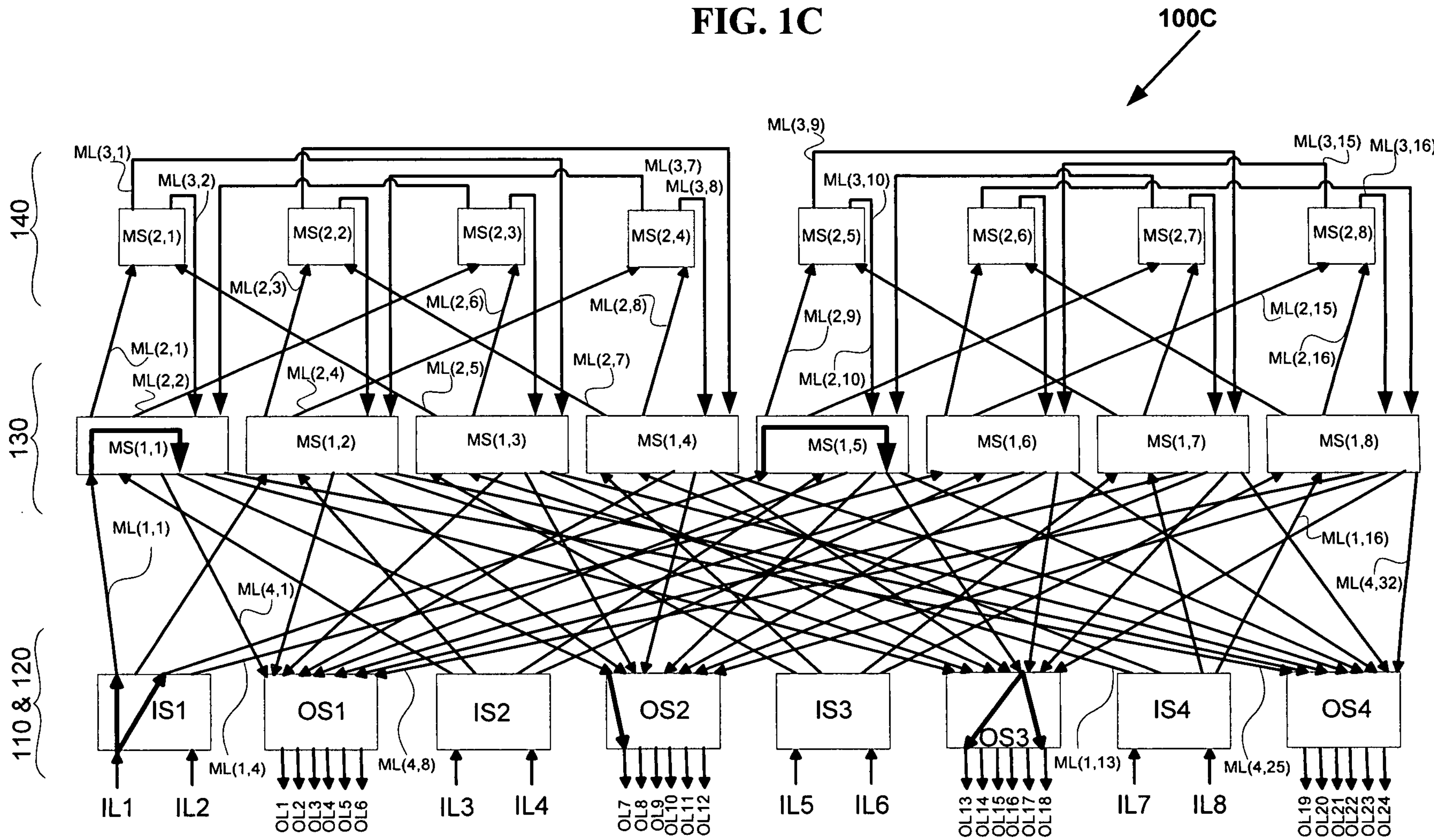

100C
140
130
110 & 120
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
MS(2,5)
MS(2,6)
MS(2,7)
MS(2,8)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
MS(1,5)
MS(1,6)
MS(1,7)
MS(1,8)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(3,1)
ML(3,2)
ML(3,7)
ML(3,8)
ML(3,9)
ML(3,10)
ML(3,15)
ML(3,16)
ML(2,1)
ML(2,2)
ML(2,3)
ML(2,4)
ML(2,5)
ML(2,6)
ML(2,7)
ML(2,8)
ML(2,9)
ML(2,10)
ML(2,15)
ML(2,16)
ML(1,1)
ML(4,1)
ML(1,16)
ML(4,32)
ML(1,4)
ML(4,8)
ML(1,13)
ML(4,25)
IL1
IL2
IL3
IL4
IL5
IL6
IL7
IL8
OL1
OL2
OL3
OL4
OL5
OL6
OL7
OL8
OL9
OL10
OL11
OL12
OL13
OL14
OL15
OL16
OL17
OL18
OL19
OL20
OL21
OL22
OL23
OL24

FIG. 1D

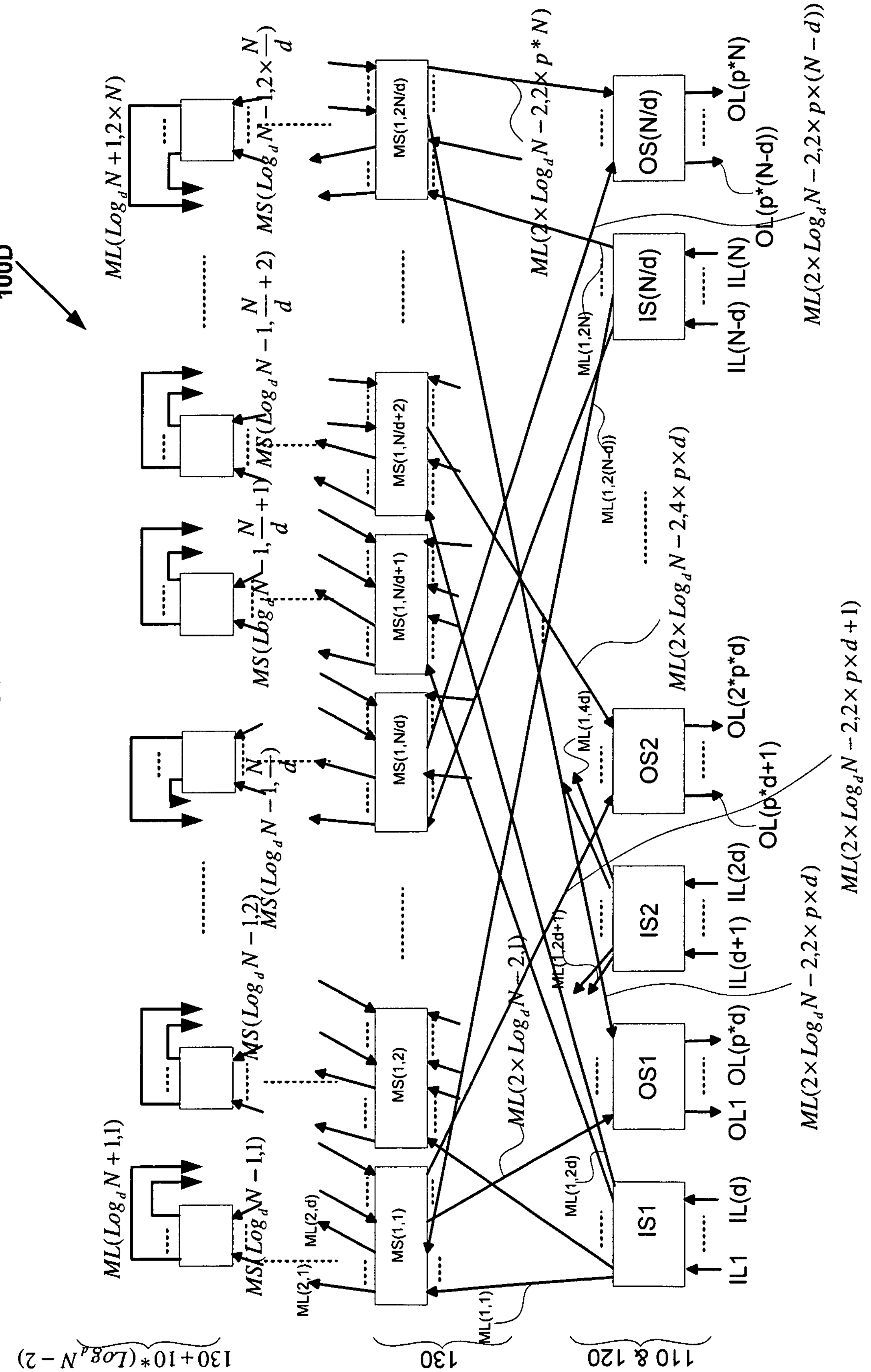

100D
IS1
IS2
IS(N/d)
OS1
OS2
OS(N/d)
MS(1,1)
MS(1,2)
MS(1,N/d)
MS(1,N/d+1)
MS(1,N/d+2)
MS(1,2N/d)
IL1
IL(d)
IL(d+1)
IL(2d)
IL(N-d)
IL(N)
OL1
OL(p*d)
OL(p*d+1)
OL(2*p*d)
OL(p*(N-d))
OL(p*N)
ML(1,1)
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2(N-d))
ML(1,2N)
ML(2,1)
ML(2,d)
110 & 120
130
130+10*(Log_d N − 2)

FIG. 1E

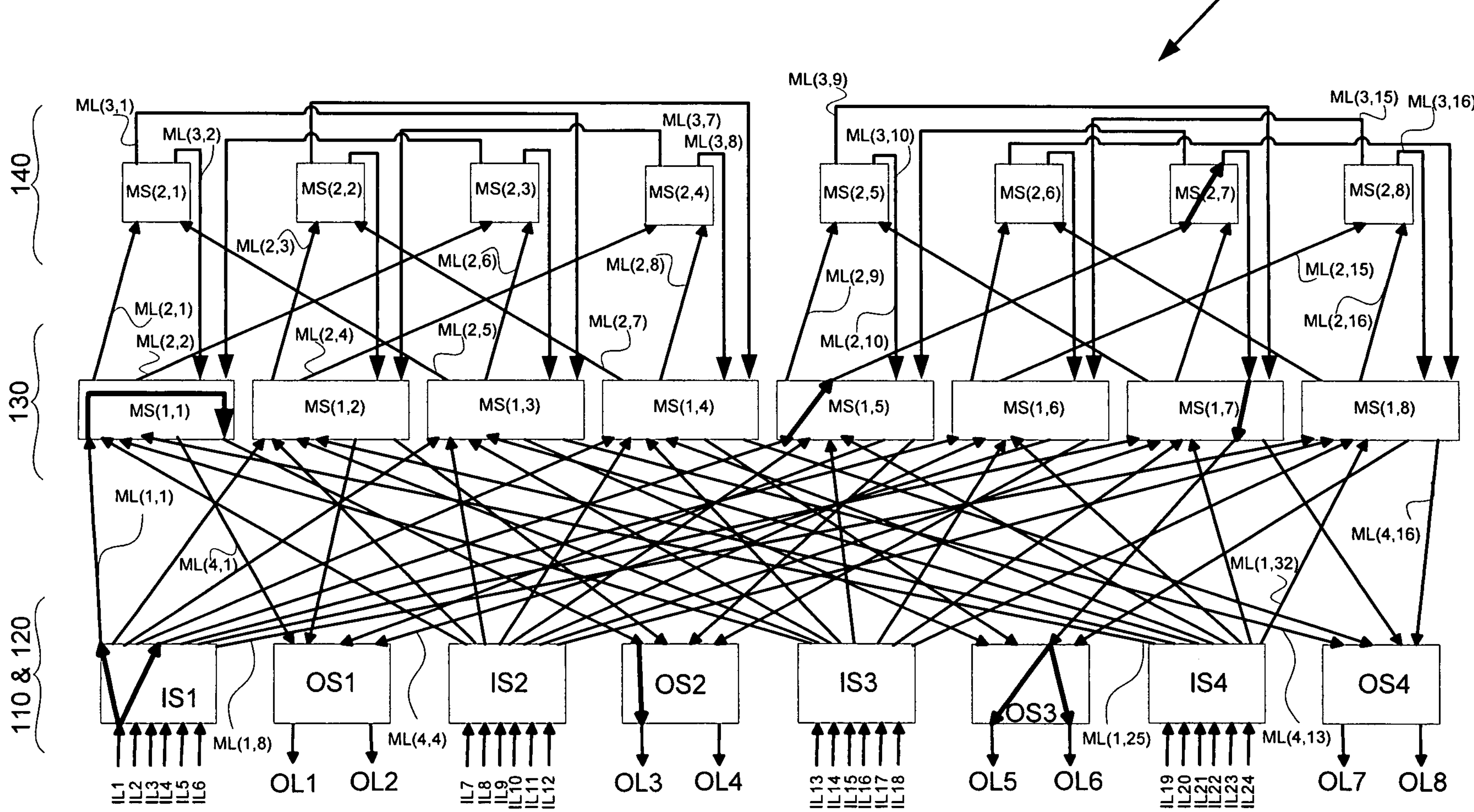

100E
140
130
110 & 120
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
MS(2,5)
MS(2,6)
MS(2,7)
MS(2,8)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
MS(1,5)
MS(1,6)
MS(1,7)
MS(1,8)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(3,1)
ML(3,2)
ML(3,7)
ML(3,8)
ML(3,9)
ML(3,10)
ML(3,15)
ML(3,16)
ML(2,1)
ML(2,2)
ML(2,3)
ML(2,4)
ML(2,5)
ML(2,6)
ML(2,7)
ML(2,8)
ML(2,9)
ML(2,10)
ML(2,15)
ML(2,16)
ML(1,1)
ML(4,1)
ML(1,8)
ML(4,4)
ML(1,25)
ML(4,13)
ML(1,32)
ML(4,16)
IL1 IL2 IL3 IL4 IL5 IL6
IL7 IL8 IL9 IL10 IL11 IL12
IL13 IL14 IL15 IL16 IL17 IL18
IL19 IL20 IL21 IL22 IL23 IL24
OL1 OL2 OL3 OL4 OL5 OL6 OL7 OL8

FIG. 1F

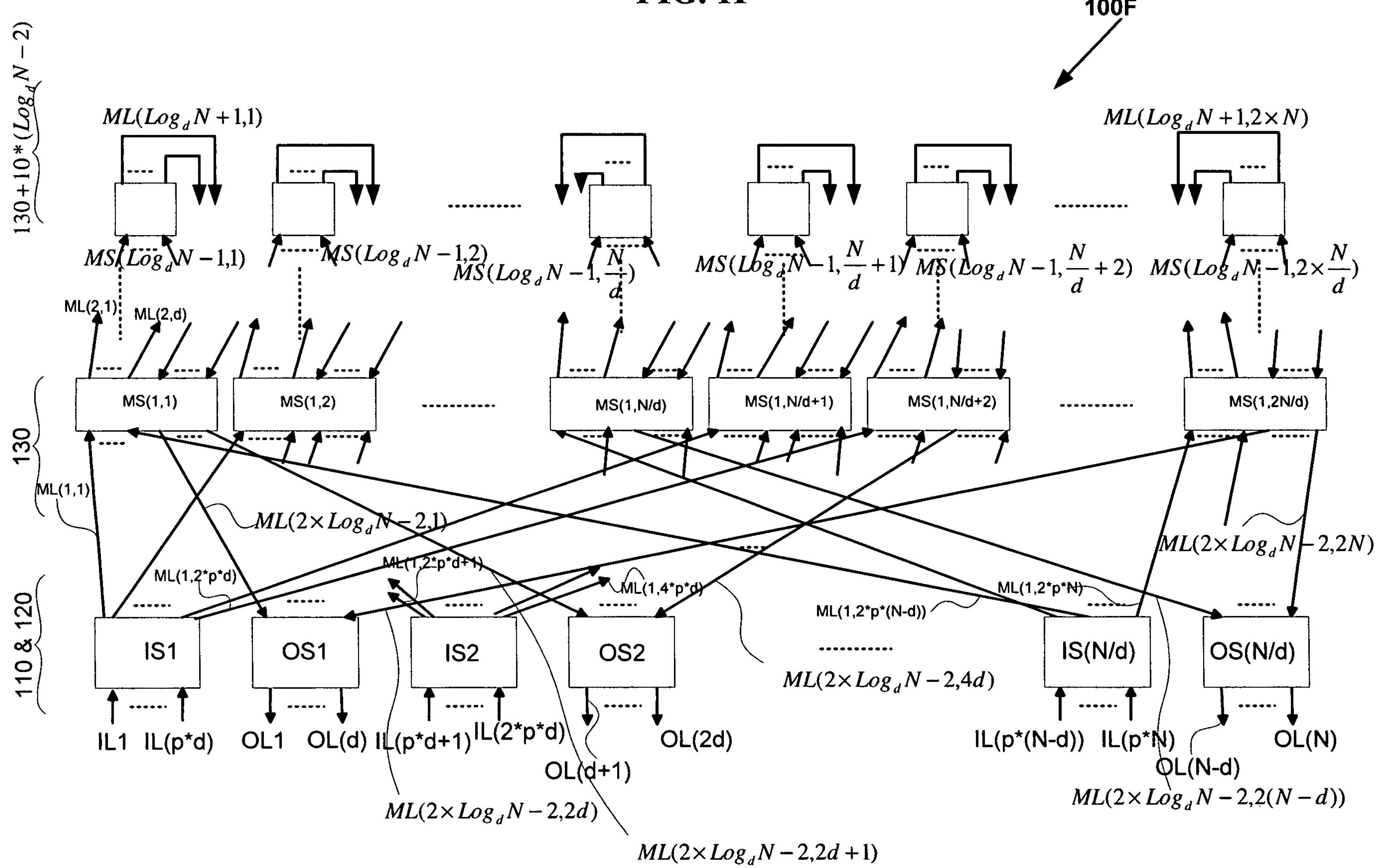

100F
$ML(Log_d N+1,1)$
$ML(Log_d N+1,2\times N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{d})$
$MS(Log_d N-1,\frac{N}{d}+1)$
$MS(Log_d N-1,\frac{N}{d}+2)$
$MS(Log_d N-1,2\times\frac{N}{d})$
130+10*($Log_d N-2$)
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/d)
MS(1,N/d+1)
MS(1,N/d+2)
MS(1,2N/d)
130
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,2N)$
ML(1,2*p*d)
ML(1,2*p*d+1)
ML(1,4*p*d)
ML(1,2*p*N)
ML(1,2*p*(N-d))
110 & 120
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
IL1
IL(p*d)
OL1
OL(d)
IL(p*d+1)
IL(2*p*d)
OL(d+1)
OL(2d)
$ML(2\times Log_d N-2,4d)$
IL(p*(N-d))
IL(p*N)
OL(N-d)
OL(N)
$ML(2\times Log_d N-2,2(N-d))$
$ML(2\times Log_d N-2,2d)$
$ML(2\times Log_d N-2,2d+1)$

FIG. 2A

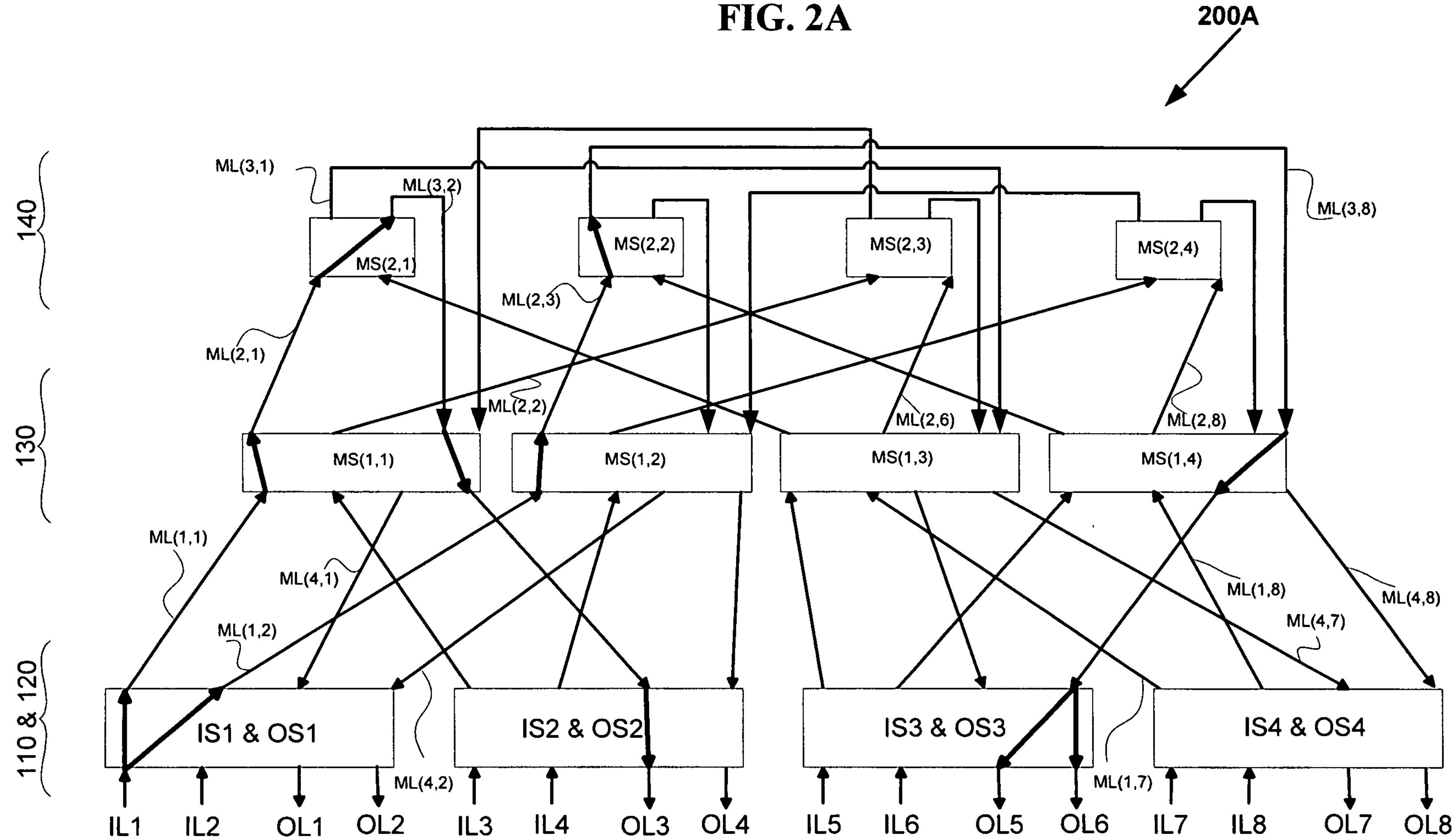

200A
140
130
110 & 120
ML(3,1)
ML(3,2)
ML(3,8)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,1)
ML(2,3)
ML(2,2)
ML(2,6)
ML(2,8)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(4,1)
ML(1,2)
ML(1,8)
ML(4,7)
ML(4,8)
ML(4,2)
ML(1,7)
IS1 & OS1
IS2 & OS2
IS3 & OS3
IS4 & OS4
IL1
IL2
OL1
OL2
IL3
IL4
OL3
OL4
IL5
IL6
OL5
OL6
IL7
IL8
OL7
OL8

FIG. 2B

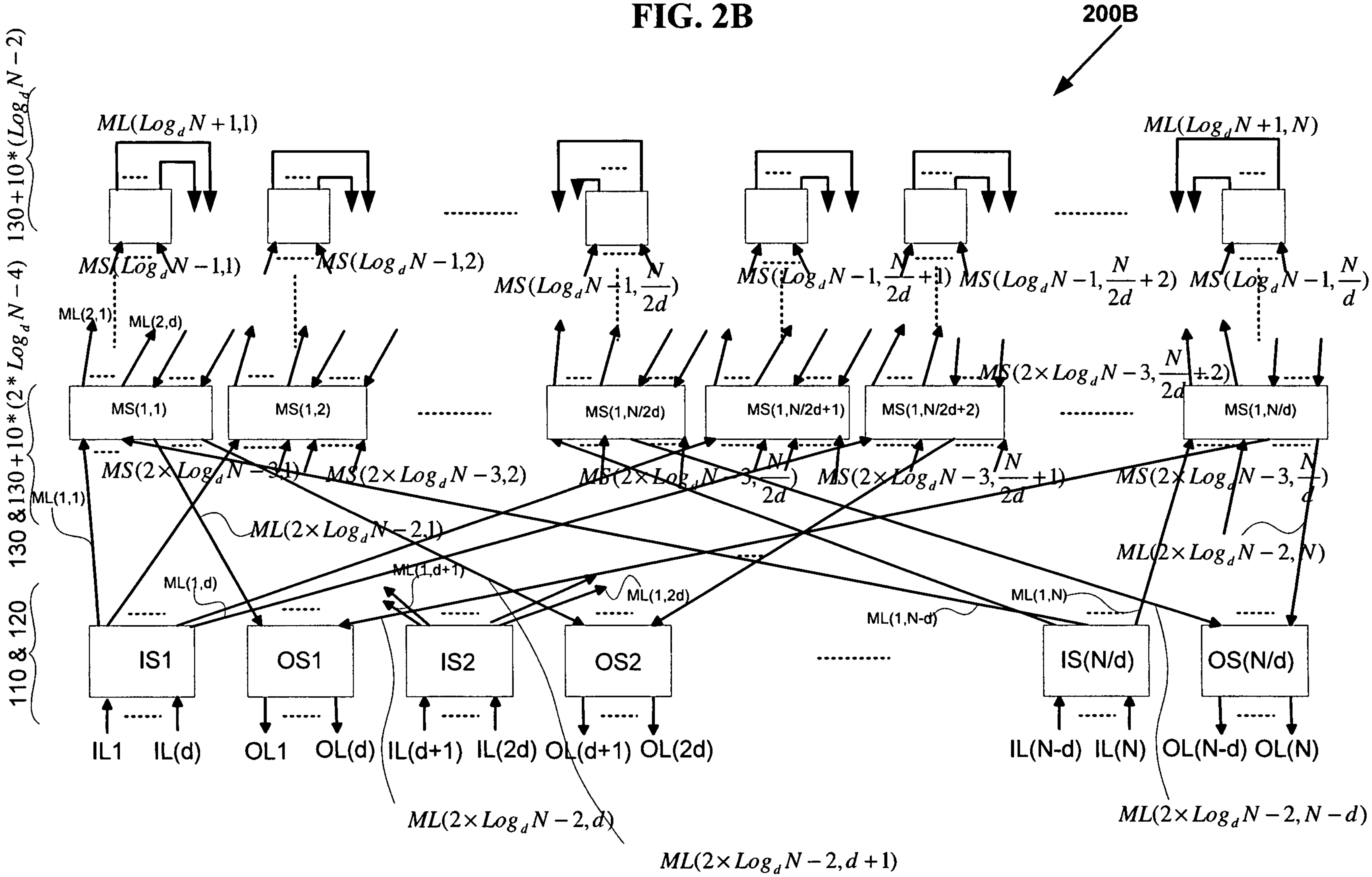

200B
$ML(Log_d N+1,1)$
$ML(Log_d N+1,N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{2d})$
$MS(Log_d N-1,\frac{N}{2d}+1)$
$MS(Log_d N-1,\frac{N}{2d}+2)$
$MS(Log_d N-1,\frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
$MS(2\times Log_d N-3,\frac{N}{2d}+2)$
$MS(2\times Log_d N-3,1)$
$MS(2\times Log_d N-3,2)$
$MS(2\times Log_d N-3,\frac{N}{2d})$
$MS(2\times Log_d N-3,\frac{N}{2d}+1)$
$MS(2\times Log_d N-3,\frac{N}{d})$
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,N)$
ML(1,d)
ML(1,d+1)
ML(1,2d)
ML(1,N)
ML(1,N-d)
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
IL1
IL(d)
OL1
OL(d)
IL(d+1)
IL(2d)
OL(d+1)
OL(2d)
IL(N-d)
IL(N)
OL(N-d)
OL(N)
$ML(2\times Log_d N-2,d)$
$ML(2\times Log_d N-2,N-d)$
$ML(2\times Log_d N-2,d+1)$
110 & 120
130 & 130+10*(2*$Log_d N-4$)
130+10*($Log_d N-2$)

FIG. 2C

200C

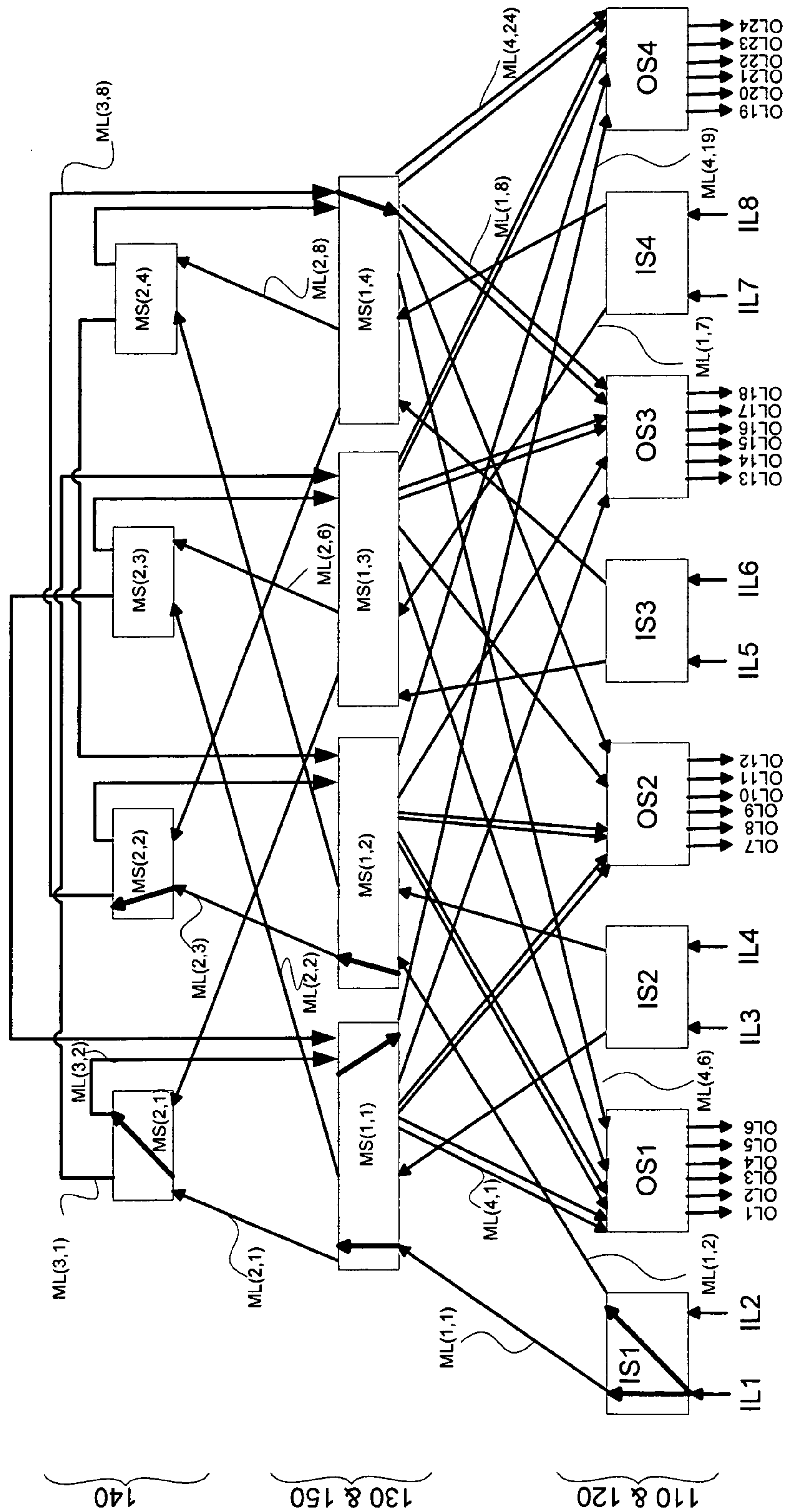

140
130 & 150
110 & 120
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
IS1
IS2
IS3
IS4
OS1
OS2
OS3
OS4
IL1
IL2
IL3
IL4
IL5
IL6
IL7
IL8
OL1
OL2
OL3
OL4
OL5
OL6
OL7
OL8
OL9
OL10
OL11
OL12
OL13
OL14
OL15
OL16
OL17
OL18
OL19
OL20
OL21
OL22
OL23
OL24
ML(1,1)
ML(1,2)
ML(1,7)
ML(1,8)
ML(2,1)
ML(2,2)
ML(2,3)
ML(2,6)
ML(2,8)
ML(3,1)
ML(3,2)
ML(3,8)
ML(4,1)
ML(4,6)
ML(4,19)
ML(4,24)

FIG. 2D

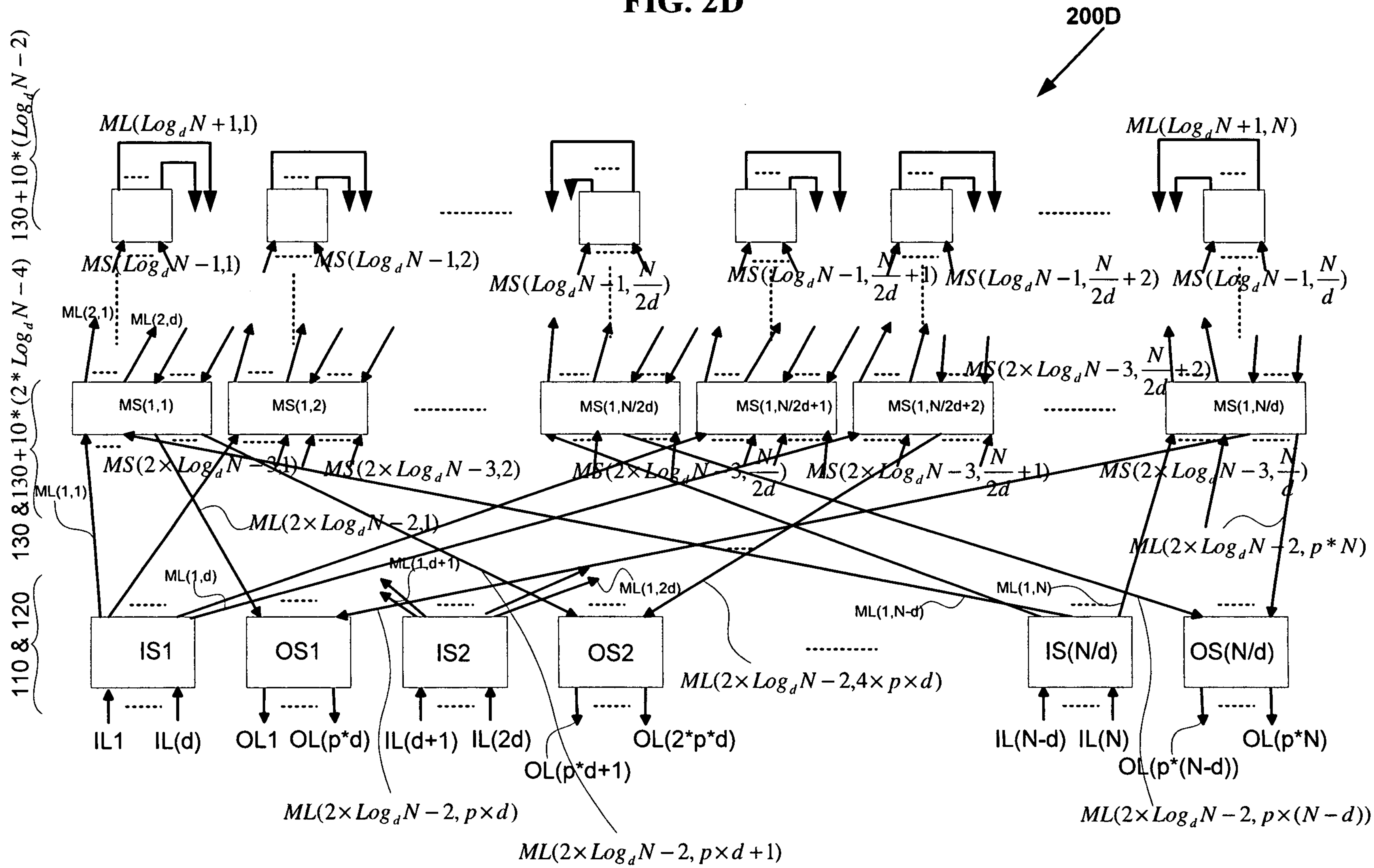

200D
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
IL1
IL(d)
IL(d+1)
IL(2d)
IL(N-d)
IL(N)
OL1
OL(p*d)
OL(p*d+1)
OL(2*p*d)
OL(p*(N-d))
OL(p*N)
ML(1,1)
ML(1,d)
ML(1,d+1)
ML(1,2d)
ML(1,N-d)
ML(1,N)
ML(2,1)
ML(2,d)
$ML(Log_d N+1,1)$
$ML(Log_d N+1,N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{2d})$
$MS(Log_d N-1,\frac{N}{2d}+1)$
$MS(Log_d N-1,\frac{N}{2d}+2)$
$MS(Log_d N-1,\frac{N}{d})$
$MS(2\times Log_d N-3,1)$
$MS(2\times Log_d N-3,2)$
$MS(2\times Log_d N-3,\frac{N}{2d})$
$MS(2\times Log_d N-3,\frac{N}{2d}+1)$
$MS(2\times Log_d N-3,\frac{N}{2d}+2)$
$MS(2\times Log_d N-3,\frac{N}{d})$
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,p\times d)$
$ML(2\times Log_d N-2,p\times d+1)$
$ML(2\times Log_d N-2,4\times p\times d)$
$ML(2\times Log_d N-2,p\times(N-d))$
$ML(2\times Log_d N-2,p*N)$
110 & 120
130 & 130+10*(2*$Log_d N-4$)
130+10*($Log_d N-2$)

FIG. 2E

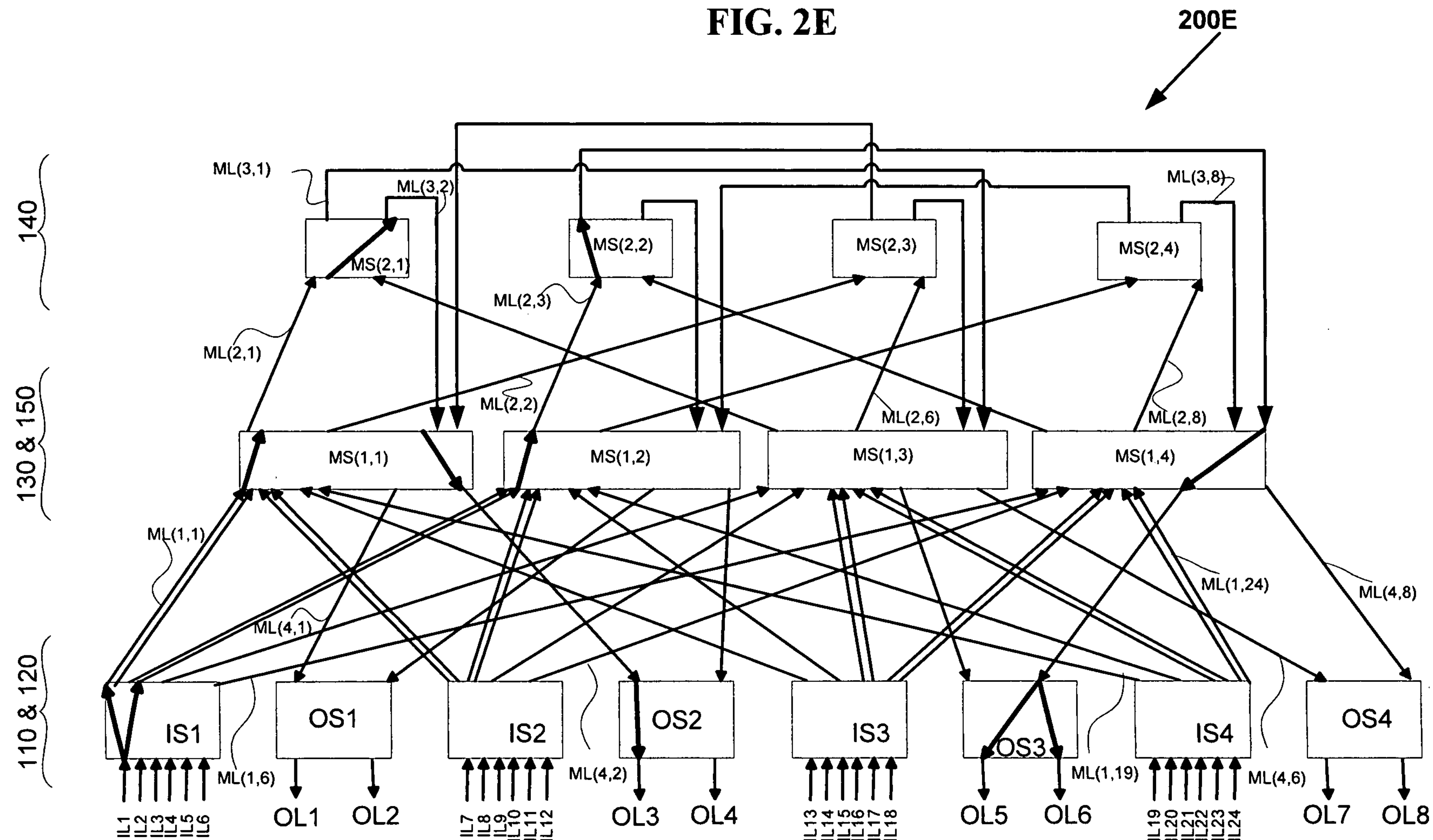

200E
140
130 & 150
110 & 120
ML(3,1)
ML(3,2)
ML(3,8)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,3)
ML(2,1)
ML(2,2)
ML(2,6)
ML(2,8)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(4,1)
ML(1,24)
ML(4,8)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,6)
ML(4,2)
ML(1,19)
ML(4,6)
IL1 IL2 IL3 IL4 IL5 IL6
OL1 OL2
IL7 IL8 IL9 IL10 IL11 IL12
OL3 OL4
IL13 IL14 IL15 IL16 IL17 IL18
OL5 OL6
IL19 IL20 IL21 IL22 IL23 IL24
OL7 OL8

FIG. 2F

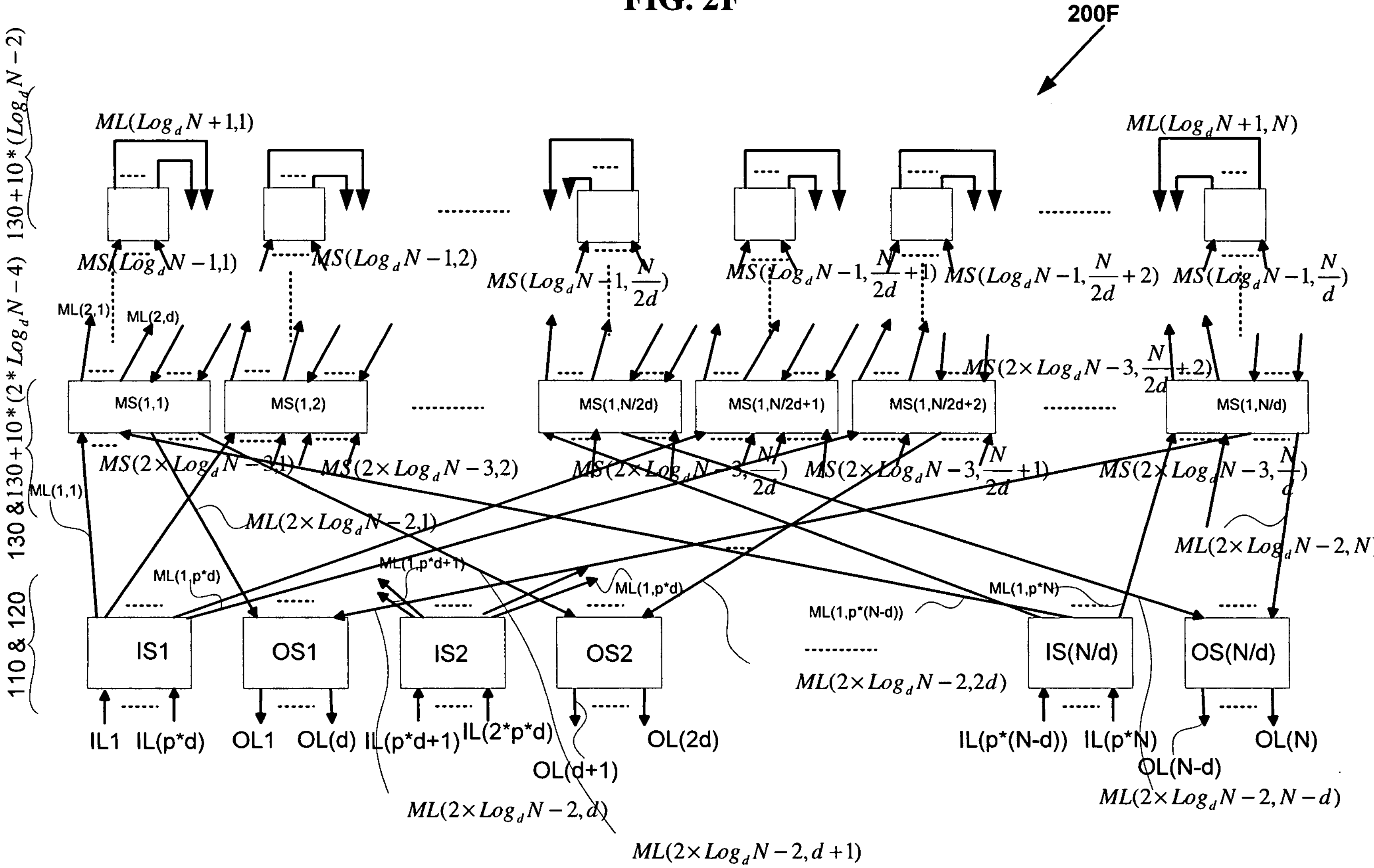

200F
$ML(Log_d N+1,1)$
$ML(Log_d N+1,N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{2d})$
$MS(Log_d N-1,\frac{N}{2d}+1)$
$MS(Log_d N-1,\frac{N}{2d}+2)$
$MS(Log_d N-1,\frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
$MS(2\times Log_d N-3,\frac{N}{2d}+2)$
$MS(2\times Log_d N-3,1)$
$MS(2\times Log_d N-3,2)$
$MS(2\times Log_d N-3,\frac{N}{2d})$
$MS(2\times Log_d N-3,\frac{N}{2d}+1)$
$MS(2\times Log_d N-3,\frac{N}{d})$
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,N)$
ML(1,p*d)
ML(1,p*d+1)
ML(1,p*d)
ML(1,p*N)
ML(1,p*(N-d))
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
$ML(2\times Log_d N-2,2d)$
IL1 IL(p*d) OL1 OL(d) IL(p*d+1) IL(2*p*d)
OL(2d)
OL(d+1)
IL(p*(N-d)) IL(p*N)
OL(N)
OL(N-d)
$ML(2\times Log_d N-2,N-d)$
$ML(2\times Log_d N-2,d)$
$ML(2\times Log_d N-2,d+1)$
$130+10*(Log_d N-2)$
$130+10*(2*Log_d N-4)$
130 & 130
110 & 120

FIG. 3A

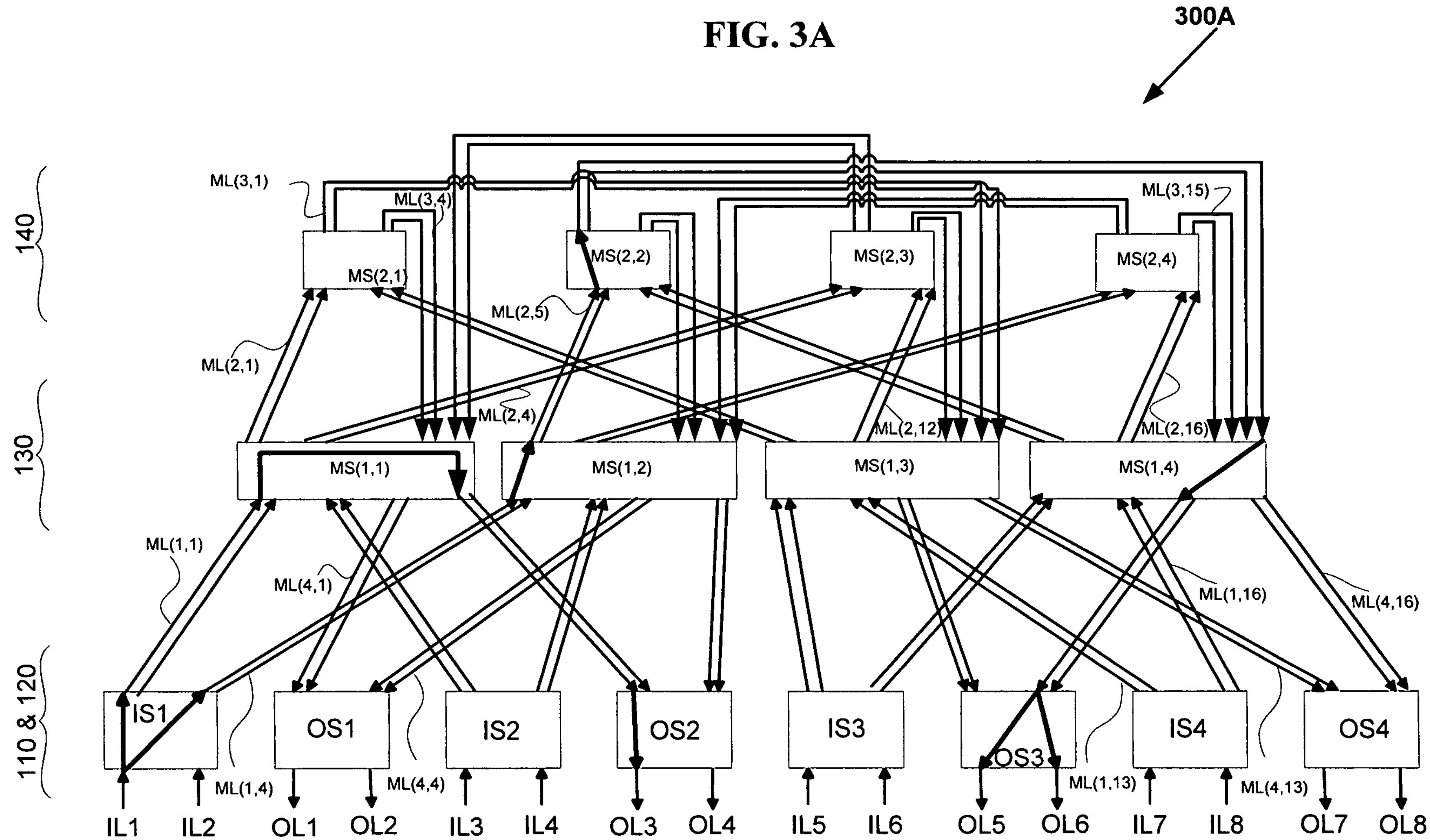

300A
140
130
110 & 120
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(3,1)
ML(3,4)
ML(3,15)
ML(2,1)
ML(2,5)
ML(2,4)
ML(2,12)
ML(2,16)
ML(1,1)
ML(4,1)
ML(1,16)
ML(4,16)
ML(1,4)
ML(4,4)
ML(1,13)
ML(4,13)
IL1
IL2
OL1
OL2
IL3
IL4
OL3
OL4
IL5
IL6
OL5
OL6
IL7
IL8
OL7
OL8

FIG. 3B

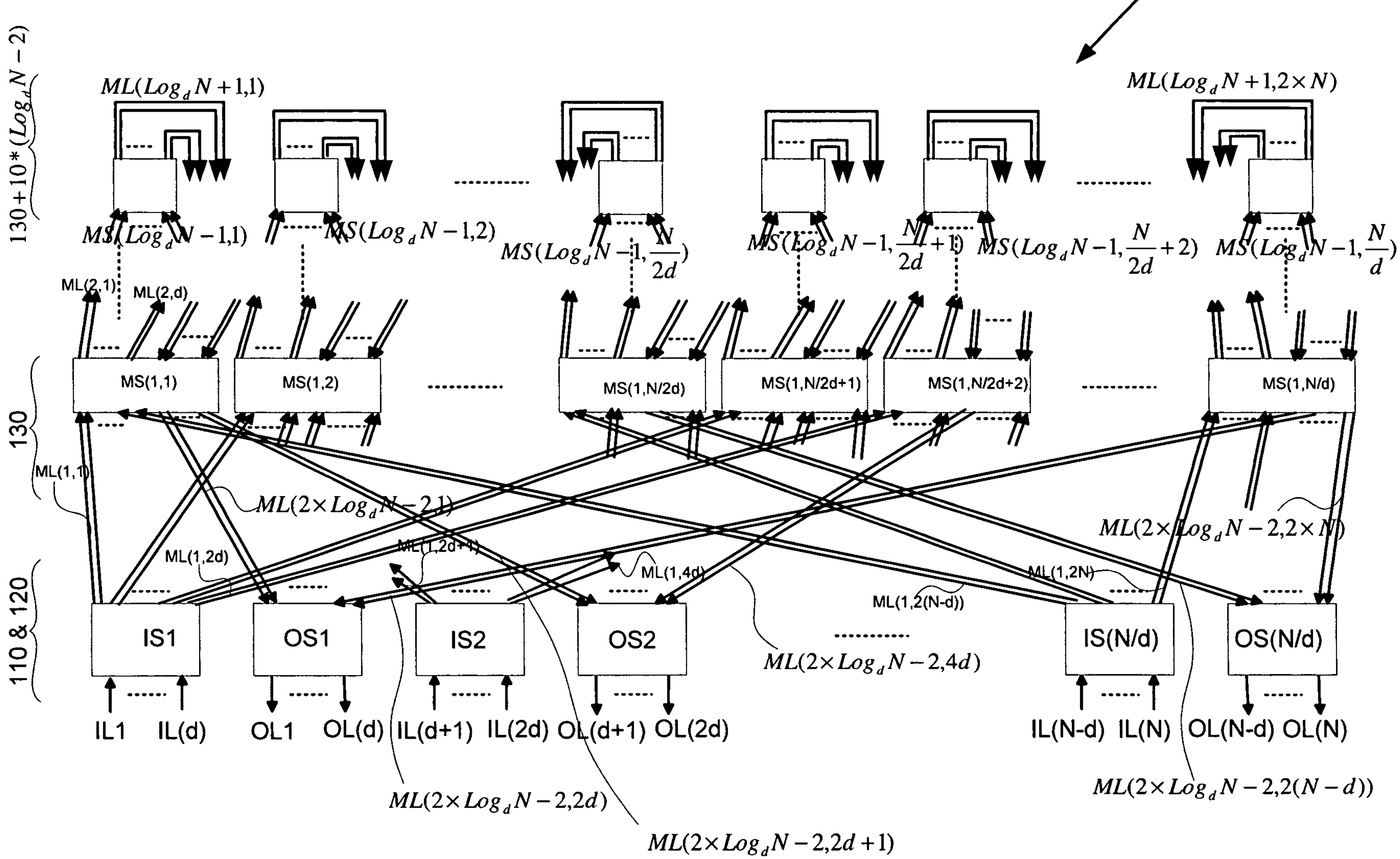

300B
$ML(Log_d N+1,1)$
$ML(Log_d N+1,2\times N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{2d})$
$MS(Log_d N-1,\frac{N}{2d}+1)$
$MS(Log_d N-1,\frac{N}{2d}+2)$
$MS(Log_d N-1,\frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
ML(1,1)
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,2\times N)$
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2N)
ML(1,2(N-d))
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
$ML(2\times Log_d N-2,4d)$
IL1
IL(d)
OL1
OL(d)
IL(d+1)
IL(2d)
OL(d+1)
OL(2d)
IL(N-d)
IL(N)
OL(N-d)
OL(N)
$ML(2\times Log_d N-2,2d)$
$ML(2\times Log_d N-2,2d+1)$
$ML(2\times Log_d N-2,2(N-d))$
$130+10*(Log_d N-2)$
130
110 & 120

FIG. 3C

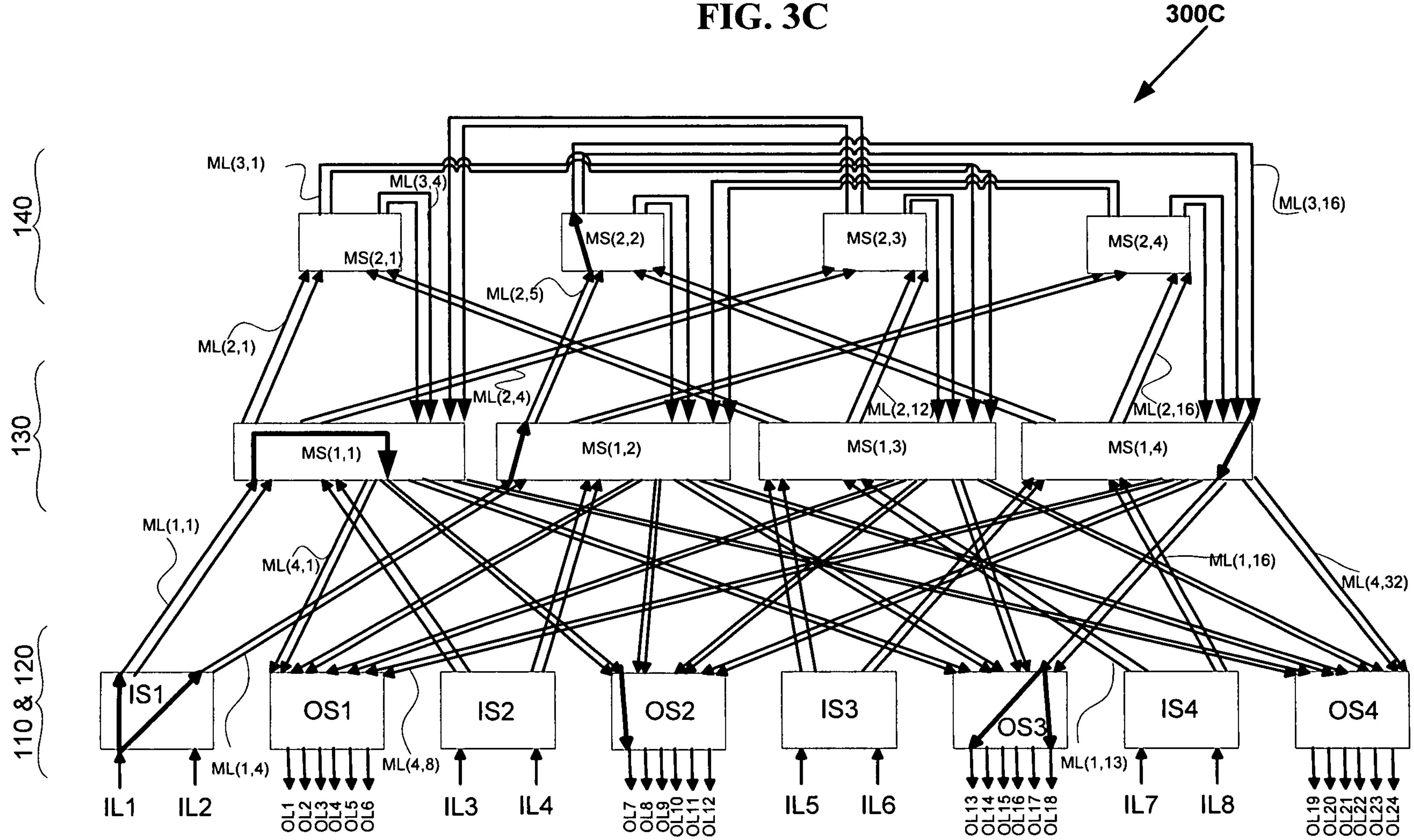

300C
140
130
110 & 120
ML(3,1)
ML(3,4)
ML(3,16)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,1)
ML(2,5)
ML(2,4)
ML(2,12)
ML(2,16)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(4,1)
ML(1,16)
ML(4,32)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,4)
ML(4,8)
ML(1,13)
IL1
IL2
OL1 OL2 OL3 OL4 OL5 OL6
IL3
IL4
OL7 OL8 OL9 OL10 OL11 OL12
IL5
IL6
OL13 OL14 OL15 OL16 OL17 OL18
IL7
IL8
OL19 OL20 OL21 OL22 OL23 OL24

FIG. 3D

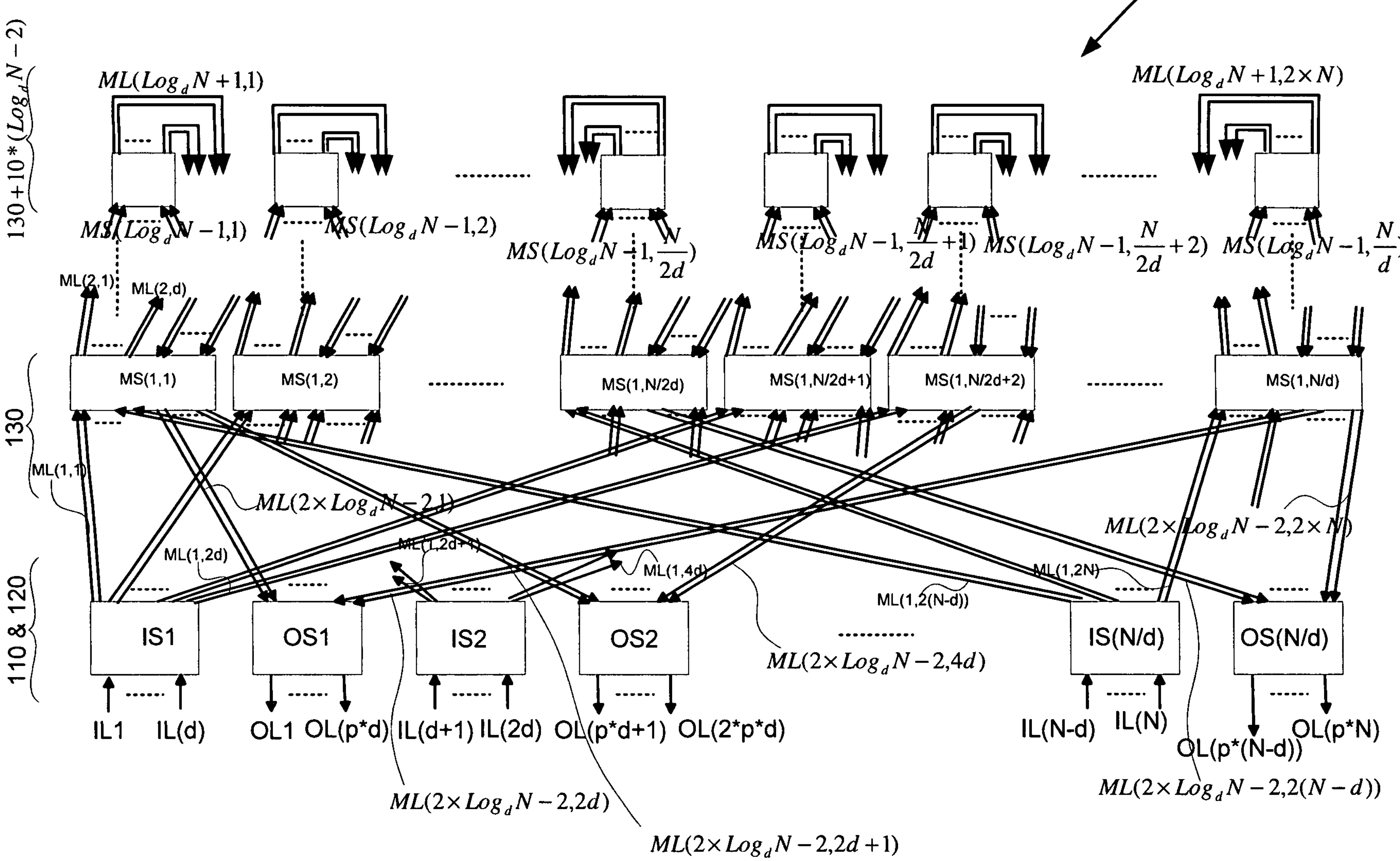

300D
$ML(Log_d N+1,1)$
$ML(Log_d N+1,2\times N)$
$MS(Log_d N-1,1)$
$MS(Log_d N-1,2)$
$MS(Log_d N-1,\frac{N}{2d})$
$MS(Log_d N-1,\frac{N}{2d}+1)$
$MS(Log_d N-1,\frac{N}{2d}+2)$
$MS(Log_d N-1,\frac{N}{d})$
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
ML(1,1)
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2N)
ML(1,2(N-d))
$ML(2\times Log_d N-2,1)$
$ML(2\times Log_d N-2,2\times N)$
$ML(2\times Log_d N-2,4d)$
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
IL1
IL(d)
OL1
OL(p*d)
IL(d+1)
IL(2d)
OL(p*d+1)
OL(2*p*d)
IL(N-d)
IL(N)
OL(p*(N-d))
OL(p*N)
$ML(2\times Log_d N-2,2(N-d))$
$ML(2\times Log_d N-2,2d)$
$ML(2\times Log_d N-2,2d+1)$
$130+10*(Log_d N-2)$
130
110 & 120

FIG. 3E

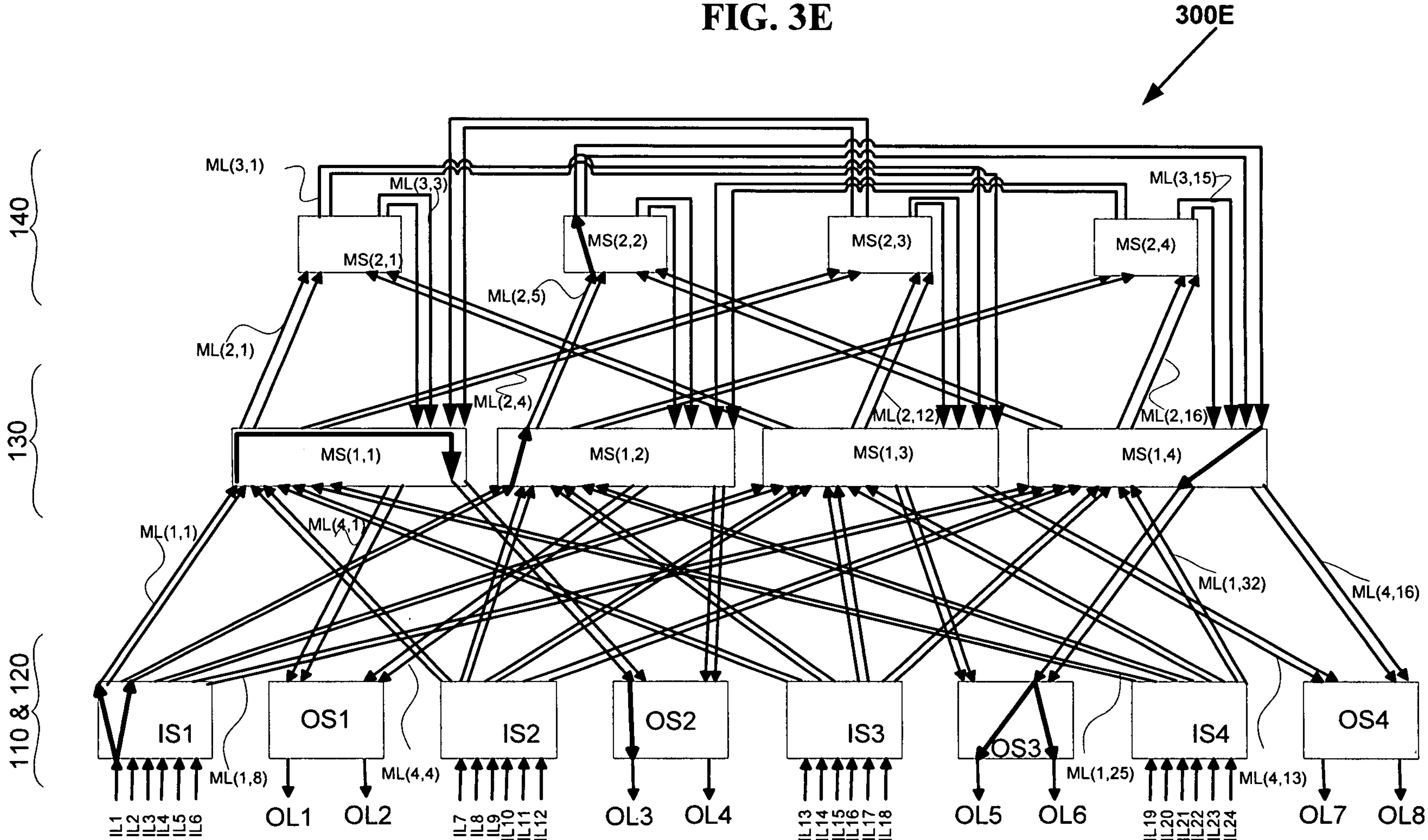

300E
140
130
110 & 120
ML(3,1)
ML(3,3)
ML(3,15)
MS(2,1)
MS(2,2)
MS(2,3)
MS(2,4)
ML(2,5)
ML(2,1)
ML(2,4)
ML(2,12)
ML(2,16)
MS(1,1)
MS(1,2)
MS(1,3)
MS(1,4)
ML(1,1)
ML(4,1)
ML(1,32)
ML(4,16)
IS1
OS1
IS2
OS2
IS3
OS3
IS4
OS4
ML(1,8)
ML(4,4)
ML(1,25)
ML(4,13)
IL1 IL2 IL3 IL4 IL5 IL6
OL1
OL2
IL7 IL8 IL9 IL10 IL11 IL12
OL3
OL4
IL13 IL14 IL15 IL16 IL17 IL18
OL5
OL6
IL19 IL20 IL21 IL22 IL23 IL24
OL7
OL8

FIG. 3F

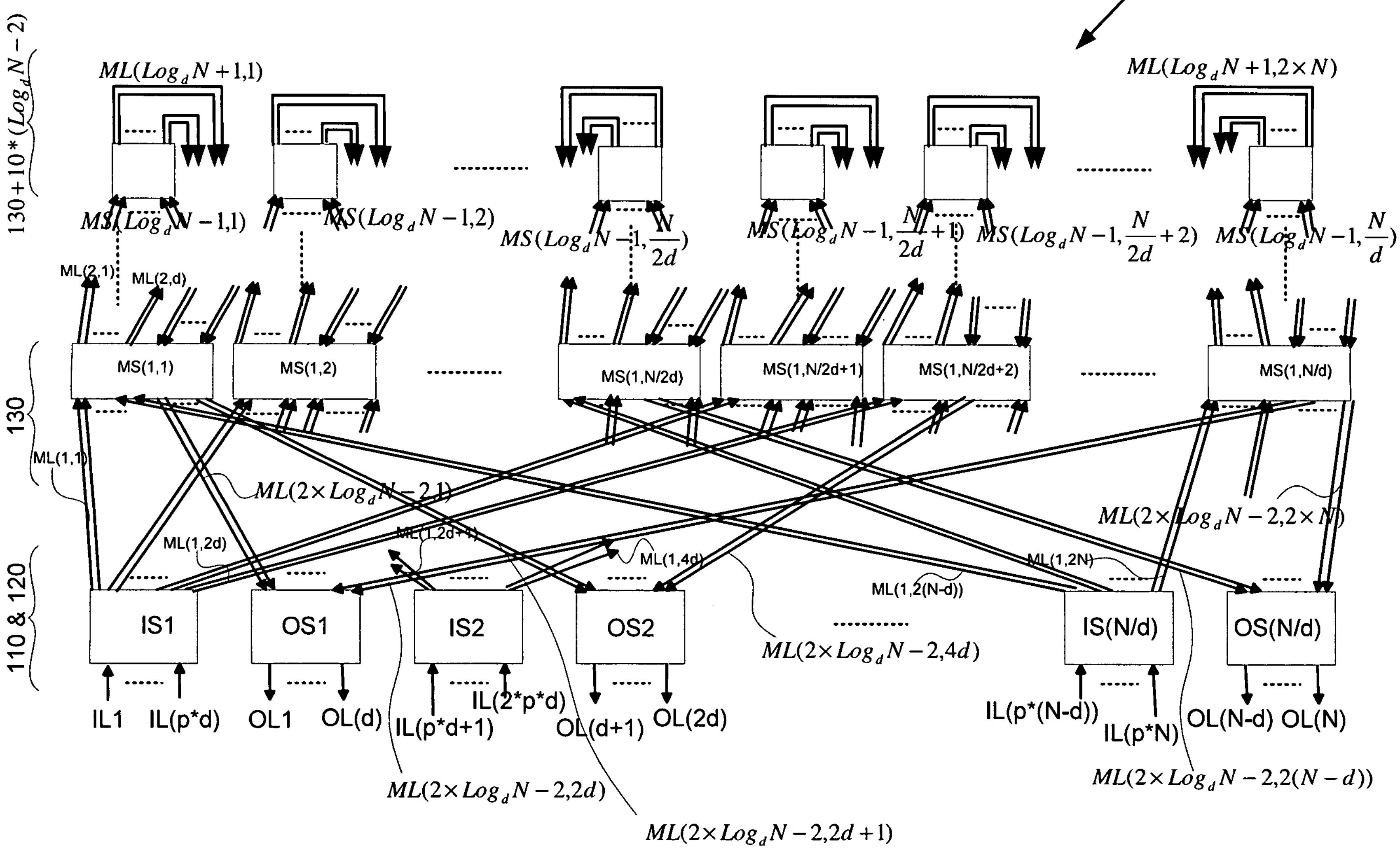

300F
130+10*(Log_d N − 2)
130
110 & 120
ML(Log_d N + 1,1)
ML(Log_d N + 1,2×N)
MS(Log_d N − 1,1)
MS(Log_d N − 1,2)
MS(Log_d N − 1, N/2d)
MS(Log_d N − 1, N/2d + 1)
MS(Log_d N − 1, N/2d + 2)
MS(Log_d N − 1, N/d)
ML(2,1)
ML(2,d)
MS(1,1)
MS(1,2)
MS(1,N/2d)
MS(1,N/2d+1)
MS(1,N/2d+2)
MS(1,N/d)
ML(1,1)
ML(2×Log_d N − 2,1)
ML(1,2d)
ML(1,2d+1)
ML(1,4d)
ML(1,2N)
ML(1,2(N-d))
ML(2×Log_d N − 2,2×N)
IS1
OS1
IS2
OS2
IS(N/d)
OS(N/d)
ML(2×Log_d N − 2,4d)
IL1
IL(p*d)
OL1
OL(d)
IL(p*d+1)
IL(2*p*d)
OL(d+1)
OL(2d)
IL(p*(N-d))
IL(p*N)
OL(N-d)
OL(N)
ML(2×Log_d N − 2,2d)
ML(2×Log_d N − 2,2(N − d))
ML(2×Log_d N − 2,2d + 1)

FIG. 4A

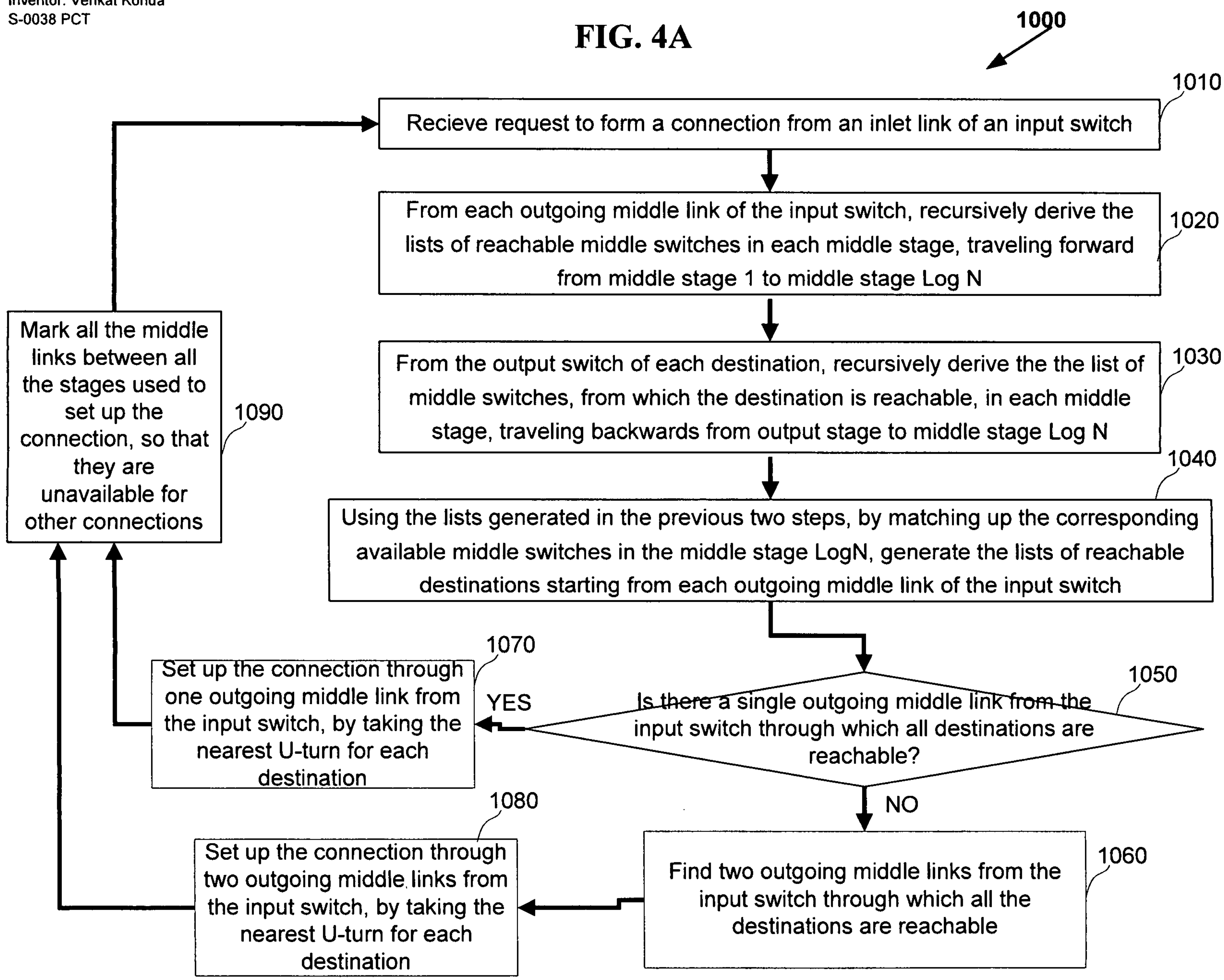

1000
1010
Recieve request to form a connection from an inlet link of an input switch
1020
From each outgoing middle link of the input switch, recursively derive the lists of reachable middle switches in each middle stage, traveling forward from middle stage 1 to middle stage Log N
1030
From the output switch of each destination, recursively derive the the list of middle switches, from which the destination is reachable, in each middle stage, traveling backwards from output stage to middle stage Log N
1040
Using the lists generated in the previous two steps, by matching up the corresponding available middle switches in the middle stage LogN, generate the lists of reachable destinations starting from each outgoing middle link of the input switch
1050
Is there a single outgoing middle link from the input switch through which all destinations are reachable?
YES
NO
1060
Find two outgoing middle links from the input switch through which all the destinations are reachable
1070
Set up the connection through one outgoing middle link from the input switch, by taking the nearest U-turn for each destination
1080
Set up the connection through two outgoing middle links from the input switch, by taking the nearest U-turn for each destination
1090
Mark all the middle links between all the stages used to set up the connection, so that they are unavailable for other connections

FIG. 5A

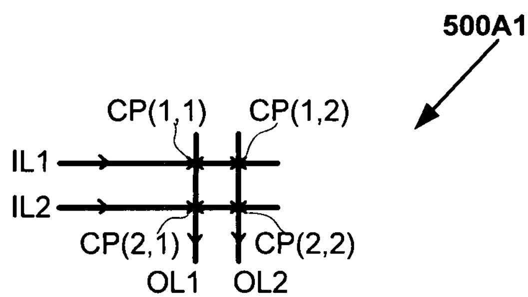


FIG. 5A1
(Prior Art)

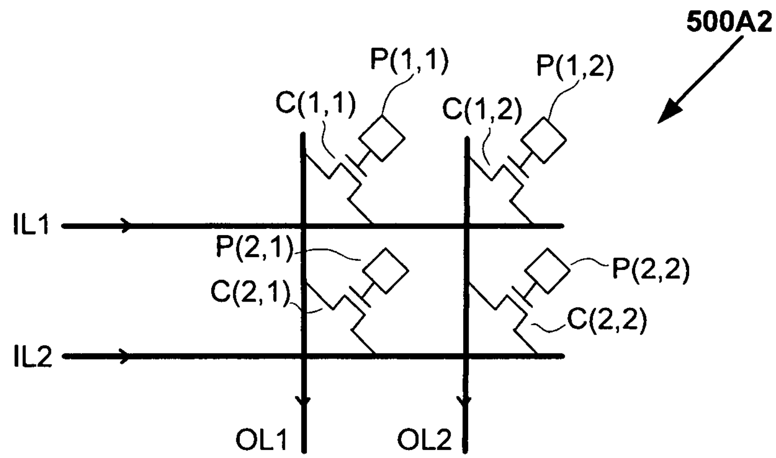


FIG. 5A2
(Prior Art)

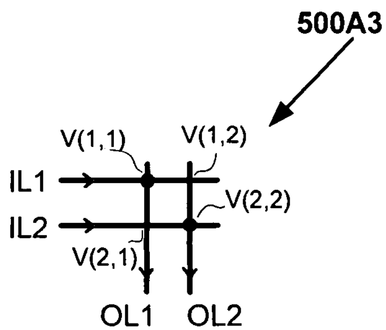


FIG. 5A3
(Prior Art)

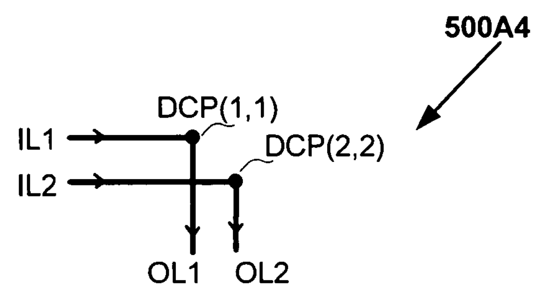


FIG. 5A4

PTO-1382 (Rev. 11-2007)
Approved for use through 02/28/2010. OMB 0651-0021
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL LETTER TO THE UNITED STATES RECEIVING OFFICE

Express Mail mailing number:	Date of deposit: 5/22/2008
File reference no.: S-0038 PCT	International application no. (if known):
Customer Number[1]: 38139	Earliest priority date claimed (Day/Month/Year): 5/25/2007
Title of the invention: Fully Connected Generalized Butterfly Fat Tree Networks	

[1] **Customer Number will allow access to the application in Private PAIR but cannot be used to establish or change the correspondence address.**

☑ **This is a new International Application**

SCREENING DISCLOSURE INFORMATION:

In order to assist in screening the accompanying international application for purposes of determining whether a license for foreign transmittal should and could be granted and for other purposes, the following information is supplied. (check as boxes as apply):

☐ The invention disclose was **not** made in the United States of America.

☐ There is no prior U.S. application relating to this invention.

☑ The following prior U.S. application(s) contain subject matter which is related to the invention disclosed in the attached international application. (*NOTE: priority to these applications may or may not be claimed on the Request (form PCT/RO/101) and this listing does **not** constitute a claim for priority.*)

application no.	60/940, 387	filed on	5/25/2007
application no.	60/940, 390	filed on	5/25/2007

☑ The present international application contains additional subject matter not found in the prior U.S. application(s) identified above. The additional subject matter is found on pages 79-82 and ☑ **DOES NOT ALTER** ☐ **MIGHT BE CONSIDERED TO ALTER** the general nature of the invention in a manner which would require the U.S. application to have been made available for inspection by the appropriate defense agencies under 35 U.S.C. 181 and 37 C.F.R. 5.15.

Itemized list of contents

Sheets of Request form: 3	Check no.:
Sheets of description (excluding sequence listing): 82	Return receipt postcard:
Sheets of claims: 12	Power of attorney:
Sheets of abstract: 1	Certified copy of priority document (specify):
Sheets of drawings: 20	Other (specify):
Sheets of sequence listing:	
Sequence listing diskette/CD:	
Tables related to sequence listing CD:	

The person signing this form is:	☑ Applicant	Venkat Konda Name of person signing
	☐ Attorney/Agent (Reg. No.)	/Venkat Konda/ Signature
	☐ Common Representative	

SEND TO: Mail Stop PCT, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

PATENT COOPERATION TREATY

From the RECEIVING OFFICE

PCT

To:

VENKAT KONDA
6278, GRAND OAK WAY
SAN JOSE, CALIFORNIA 95135

NOTIFICATION CONCERNING PAYMENT OF PRESCRIBED FEES

(PCT Rules 14, 15 and 16 and Administrative Instructions, Sections 102*bis*(c), 304, 323(b), 707(b) and 803)

Date of mailing *(day/month/year)*	10 Jun 2008
Applicant's or agent's file reference: S-0038 PCT	PAYMENT DUE: see item 3 for time limits

International application No.	International filing date/Date of receipt *(day/month/year)*	Priority date *(day/month/year)*
PCT/US2008/064603		25 May 2007

Applicant

KONDA, VENKAT

1. The applicant is hereby notified that this receiving Office has received:

☐ the payment of all the prescribed fees, and ☐ an overpayment, which will be refunded in due course.

☒ no or insufficient payment of the prescribed fees and the applicant is hereby invited to pay the balance due, as summarized under item 2, within the time limit(s) indicated under item 3.

2. Fees and payment calculation:

Total fees payable		Amount paid		Balance
4,505.00	–	0.00	=	4,505.00

☐ The details of the calculation are given in the Annex.

3. Time limit(s) for payment and amount(s) payable (Rules 14.1, 15.4 and 16.1(f)):

☒ within ONE MONTH from the date of receipt of the international application (for the transmittal fee (if any), the search fee and the international filing fee). The amount payable for each fee is the amount applicable on the date of receipt of the international application.

☐ within 16 MONTHS from the priority date (only for the fee for priority document). The applicant's attention is drawn to the fact that the request made by the applicant under Rule 17.1(b) will be considered not to have been made unless the fee is paid within that time limit.

4. Additional observations *(if necessary)*:

☒ The search copy will not be transmitted to the International Searching Authority until the search fee is paid (therefore the start of the international search will be delayed) (Rule 23.1(a) and (b)).

Name and mailing address of the receiving Office	Authorized officer
Mail Stop PCT, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450	MOTLEY NINA
Facsimile No. 571-273-3201	Telephone No. (703)308-9290 X123

Form PCT/RO/102 (January 2004)

ANNEX TO FORM PCT/RO/102
CALCULATION OF THE PRESCRIBED FEES

International application No.
PCT/US2008/064603

[T] Transmittal Fee

Prescribed amount:		300.00 [T]	
Amount paid:	–	0.00	☐ correct amount
Balance:	=	300.00	☐ overpayment
			☒ balance due

[S] Search Fee

Prescribed amount:		1,800.00 [S]	
Amount paid:	–	0.00	☐ correct amount
Balance:	=	1,800.00	☐ overpayment
			☒ balance due

[I] International Filing Fee

Fixed amount for first 30 sheets: . . . . 1,173.00 [i1]

88 (Number of sheets in excess of 30) x 14.00 (Fee per sheet) = 1,232.00 [i2]

Additional component: . . . 400 x 0.00 (Fee per sheet) = 0.00 [i3]

Reduction where the international application is filed *(See PCT Applicant's Guide, Volume I, General Part, for details on the availability of this reduction)*:

using the PCT-EASY software: . . . . . . – 0.00 [r]

or

in electronic form where the text of the description, claims and abstract is not in character coded format: . . . . – 0.00 [r]

or

in electronic form where the text of the description, claims and abstract is in character coded format: . . . . – 0.00 [r]

Sub-total: . . . . = 2,405.00 [i1+i2+i3-r]

Prescribed total amount *(The amount to be entered at I is the sub-total entered at* (i1+i2+i3-r)*, except where the applicant is (or all applicants are) entitled to a reduction of 75%, in which case the amount to be entered at I is 25% of the sub-total* (i1+i2+i3-r)*; certain applicants from certain States are entitled to a reduction of 75% of the international filing fee; see Notes to the Fee Calculation Sheet as annexed to the Request Form, PCT/RO/101, for details)*:	=	2,405.00 [I]	
Amount paid:	–	0.00	☐ correct amount
Balance:	=	2,405.00	☐ overpayment
			☒ balance due

[P] Fee for Priority Document

Prescribed amount:		0.00 [P]	
Amount paid:	–	0.00	☒ correct amount
Balance:	=	0.00	☐ overpayment
			☐ balance due

Form PCT/RO/102 (Annex) (January 2004)

PATENT COOPERATION TREATY

From the RECEIVING OFFICE

To:

VENKAT KONDA
6278, GRAND OAK WAY
SAN JOSE, CALIFORNIA 95135

Confirmation No: 5416

PCT

NOTIFICATION OF THE INTERNATIONAL APPLICATION NUMBER AND OF THE INTERNATIONAL FILING DATE

(PCT Rule 20.2(c))

Date of mailing *(day/month/year)*: 10 Jun 2008

IMPORTANT NOTIFICATION

Applicant's or agent's file reference: S-0038 PCT

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US2008/064603	22 May 2008	25 May 2007

Applicant: KONDA, VENKAT

Title of the invention: FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS

1. The applicant is hereby notified that the international application has been accorded the international application number and the international filing date indicated above.

2. The applicant is further notified that the record copy of the international application:

   ☒ was transmitted to the International Bureau on 10 Jun 2008 .

   ☐ has not yet been transmitted to the International Bureau for the reason indicated below and a copy of this notification has been sent to the International Bureau*:

   ☐ because the necessary national security clearance has not yet been obtained.

   ☐ because *(reason to be specified)*:

* The International Bureau monitors the transmittal of the record copy by the receiving Office and will notify the applicant (with Form PCT/IB/301) of its receipt. Should the record copy not have been received by the expiration of 14 months from the priority date, the International Bureau will notify the applicant (Rule 22.1(c)).

3. FOREIGN TRANSMITTAL LICENSE INFORMATION — Completed by: MN

   ☐ Additional license for foreign transmittal not required. This subject matter is covered by a license already granted or the equivalent U.S. national application. Refer to that license for information concerning its scope.

   ☐ License for foreign transmittal not required. 37 CFR. 5.11(e)(1) or 37 CFR 5.11(e)(2). However, a license may be required for additional subject matter. See 37 CFR 5.15(b).

   ☒ Foreign transmittal license granted. 35 U.S.C. 184; 37 CFR 5.11 on 04 Jun 2008 (date):

   ☐ 37 CFR 5.15(a) ☒ 37 CFR 5.15(b)

Name and mailing address of the receiving Office	Authorized officer
Mail Stop PCT, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450	MOTLEY NINA
Facsimile No. 571-273-3201	Telephone No. (703)308-9290 X123

Form PCT/RO/105 (April 2007)

PATENT COOPERATION TREATY

From the RECEIVING OFFICE

PCT

To:
VENKAT KONDA
6278, GRAND OAK WAY
SAN JOSE, CALIFORNIA 95135

INVITATION TO CORRECT DEFECTS IN THE INTERNATIONAL APPLICATION

(PCT Articles 3(4)(i) and 14(1) and Rule 26)

	Date of mailing *(day/month/year)* 10 Jun 2008
Applicant's or agent's file reference S-0038 PCT	**REPLY DUE** within **TWO MONTHS** from the above date of mailing
International application No. PCT/US2008/064603	International filing date *(day/month/year)* 22 May 2008
Applicant KONDA, VENKAT	

1. [X] The applicant is hereby **invited**, within the time limit indicated above, to correct, **in the international application as filed**, the defects specified on the attached:
   - [ ] Annex A
   - [ ] Annex B1 *(text matter of the international application as filed)*
   - [X] Annex C1 *(drawings of the international application as filed)*

2. [ ] The applicant is hereby **invited**, within the time limit indicated above, to correct, **in the translation of the international application** furnished under Rule 12.3 or 12.4, the defects specified on the attached:
   - [ ] Annex A
   - [ ] Annex B2 *(text matter of the translation of the international application)*
   - [ ] Annex C2 *(drawings of the translation of the international application)*

Additional observations (if necessary):

HOW TO CORRECT THE DEFECTS?

Except where the defect is in the request, any correction must be submitted by filing a replacement sheet embodying the correction and a letter accompanying the replacement sheet, which shall draw attention to the difference between the replaced sheet and the replacement sheet. For a defect in the request, a correction may simply be stated in a letter if it is of such a nature that the correction can be transferred clearly onto the request record copy (Rule 26.4).

ATTENTION

Failure to correct the defects will result in the international application being considered withdrawn by this receiving Office (see Rule 26.5 for further details).

A copy of this Invitation and any attachments has been sent to the International Bureau
[X] and the International Searching Authority.

Name and mailing address of the receiving Office Mail Stop PCT, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-3201	Authorized officer MOTLEY NINA Telephone No. (703)308-9290 X123

Form PCT/RO/106 (April 2007)

ANNEX A TO FORM PCT/RO/106

International application No.
PCT/US2008/064603

The receiving Office has found the following defects in the international application as filed:

1. As to signature of the international application (Rules 4.15, 26.2*bis*(a) and 90.4), the request:

a. ☐ is not signed* by the applicant or, if there is more than one applicant, by at least one of them
b. ☐ is not accompanied by the statement referred to in the check list in Box No. IX of the request explaining the lack of the signature of an applicant for the designation of the United States of America
c. ☐ is signed by what appears to be an agent/common representative but:
   ☐ the international application is not accompanied by a power of attorney appointing him
   ☐ the power of attorney accompanying the international application is not signed by all the applicants
d. ☐ other *(specify)*:

* Although Rule 4.15 requires that all applicants must sign the request (e.g. including all inventors/applicants for the designation of the United States of America), for the purposes of Article 14(1)(a)(i), if there is more than one applicant, it shall be sufficient that the request be signed by one of them (Rule 26.2*bis*(a)).

However, the applicant's attention is drawn to the fact that the national law applied by each designated Office may require, in connection with the processing of the international application in the national phase, that the applicant furnish the confirmation of the international application by the signature of any applicant for the designated State who has not signed the request (Rule 51*bis*.1(a)(vi)).

2. As to indications concerning the applicant* who is entitled, according to Rule 19.1, to file the international application with the receiving Office, the request (Rules 4.4, 4.5 and 26.2*bis*(b)):

a. ☐ does not properly indicate the applicant's name *(specify)*:

b. ☐ does not indicate the applicant's address
c. ☐ does not properly indicate the applicant's address *(specify)*:

d. ☐ does not indicate the applicant's nationality
e. ☐ does not indicate the applicant's residence

☐ Further observations about indications concerning other applicants (if applicable):

* Although Rules 4.4 and 4.5 require indications concerning the applicant, or if there are several applicants, of each of them, for the purposes of Article 14(1)(a)(ii), if there is more than one applicant, it shall be sufficient that the indications required under Rule 4.5(a)(ii) and (iii) be provided in respect of one of them who is entitled according to Rule 19.1 to file the international application with the receiving Office (Rule 26.2*bis*(b)).

However, the applicant's attention is drawn to the fact that the national law applied by each designated Office may require, in connection with the processing of the international application in the national phase, that the applicant furnish any missing indication required under Rule 4.5(a)(ii) and (iii) in respect of any applicant for the designated State (Rule 51*bis*.1(a)(vii)).

3. As to the language of certain elements of the international application, other than the description and claims (Rules 12.1(c) and 26.3*ter*(a) and (c)):

a. ☐ the request is not in a language of publication accepted by this receiving Office; (the) language(s) accepted by this receiving Office is/are:

b. ☐ the text matter of the drawings is not in the language in which the international application is to be published, which is:

c. ☐ the abstract is not in the language in which the international application is to be published, which is:

4. The title of the invention:

a. ☐ is not indicated in Box No. I of the request (Rule 4.1(a))
b. ☐ is not indicated at the top of the first sheet of the description (Rule 5.1(a))
c. ☐ as appearing in Box No. I of the request is not identical with the title heading the description (Rule 5.1(a))

5. As to the abstract (Rules 8 and 26.1(b)):

☐ the international application does not contain an abstract

ANNEX B1 TO FORM PCT/RO/106

International application No.
PCT/US2008/064603

This receiving Office has found that, with regard to the presentation of the text matter of the international application as filed, the physical requirements are not complied with to the extent that compliance therewith is necessary for:

1. ☐ reasonably uniform international publication (Rules 11 and 26.3(a)(i)) *(defects to be specified)*:

	Request	Description	Claims	Abstract
a. ☐ the sheets do not admit of direct reproduction	☐	☐	☐	☐
b. ☐ the element does not commence on a new sheet	☐	☐	☐	☐
c. ☐ sheets are not free from creases, cracks, folds	☐	☐	☐	☐
d. ☐ sheets are not used in the upright position	☐	☐	☐	☐
e. ☐ one side of the sheets is not left unused	☐	☐	☐	☐
f. ☐ the paper of the sheets is not flexible/strong/white/smooth/non-shiny/durable	☐	☐	☐	☐
g. ☐ the sheets are not connected as prescribed (Rule 11.4(b))	☐	☐	☐	☐
h. ☐ sheets are not A4 size (29.7cm x 21cm)	☐	☐	☐	☐
i. ☐ the minimum margins on the sheets are not as prescribed (top: 2cm; left side: 2.5cm; right side: 2cm; bottom: 2cm)		☐	☐	☐
j. ☐ the file reference number indicated on the sheets does not appear in the left-hand corner of the sheets, within 1.5 cm of the top of the sheets		☐	☐	☐
k. ☐ the file reference number exceeds the maximum of 12 characters	☐	☐	☐	☐
l. ☐ the sheets of the description, claims and abstract are not numbered in consecutive Arabic numerals		☐	☐	☐
m. ☐ the sheet numbers are not centered at the top or bottom of the sheets	☐	☐	☐	☐
n. ☐ the sheet numbers are in the margin (see i. above for the size of the margins)		☐	☐	☐
o. ☐ the text matter is not typed or printed	☐	☐	☐	☐
p. ☐ the typing on the sheets is not 1½-spaced		☐	☐	☐
q. ☐ the characters in the text matter on the sheets are less than 0.28 cm high in capital letters	☐	☐	☐	☐
r. ☐ the text matter on the sheets is not in dark, indelible color	☐	☐	☐	☐
s. ☐ the element contains drawings	☐	☐	☐	☐
t. ☐ the sheets contain alterations/overwritings/interlineations/too many erasures	☐	☐	☐	☐
u. ☐ the sheets contain photocopy marks	☐	☐	☐	☐

2. ☐ satisfactory reproduction (Rules 11 and 26.3(b)(i))

Further observations (if necessary):

ANNEX C1 TO FORM PCT/RO/106	International application No. PCT/US2008/064603

This receiving Office has found that, with regard to the presentation of the drawings of the international application as filed, the physical requirements are not complied with to the extent that compliance therewith is necessary for:

1. [X] reasonably uniform international publication (Rules 11 and 26.3(a)(i)) *(defects to be specified)*:

Sheets containing drawings:

- a. [ ] the sheets do not admit of direct reproduction
- b. [ ] the sheets are not free from creases, cracks, folds
- c. [ ] one side of the sheets is not left unused
- d. [ ] the paper of the sheets is not flexible/strong/white/smooth/non-shiny/durable
- e. [ ] the drawings do not commence on a new sheet
- f. [ ] the sheets are not connected as prescribed (Rule 11.4(b))
- g. [ ] the sheets are not A4 size (29.7cm x 21cm)
- h. [ ] the minimum margins on the sheets are not as prescribed (top: 2.5cm; left side: 2.5cm; right side: 1.5cm; bottom: 1cm)
- i. [ ] the file reference number indicated on the sheets does not appear in the left-hand corner of the sheets, within 1.5 cm of the top of the sheets
- j. [ ] the file reference number exceeds the maximum of 12 characters
- k. [ ] the sheets are not free from frames around usable or used surfaces
- l. [ ] the sheets are not numbered in consecutive Arabic numerals (e.g. 1/3, 2/3, 3/3)
- m. [ ] the sheet numbers are not centered at the top or bottom of the sheets
- n. [ ] the sheet numbers are in the margin (see h. above for the size of the margins)
- o. [ ] the sheets contain alterations/overwritings/interlineations/too many erasures
- p. [ ] the sheets contain photocopy marks

Drawings (Rule 11.13):

- a. [ ] do not admit of direct reproduction
- b. [ ] contain unnecessary text matter
- c. [X] contain words so placed as to prevent translation without interference with lines thereof
- d. [ ] are not executed in durable black color; the lines are not uniformly thick and well-defined
- e. [ ] contain cross-sections not properly hatched
- f. [ ] would not be properly distinguishable in reduced reproduction
- g. [ ] contain scales not represented graphically
- h. [X] contain numbers, letters and reference lines lacking simplicity and clarity
- i. [ ] contain lines drafted without the aid of drafting instruments
- j. [ ] contain disproportionate elements of a figure not necessary for clarity
- k. [X] contain numbers and letters of height less than 0.32 cm
- l. [ ] contain letters not conforming to the Latin, and where customary, Greek alphabets
- m. [ ] contain figures on two or more sheets which form a single complete figure but which are not able to be assembled without concealing parts thereof
- n. [ ] contain figures which are not properly arranged and clearly separated
- o. [ ] contain different figures not numbered in consecutive Arabic numerals
- p. [ ] contain different figures not numbered independently of the numbering of the sheets
- q. [ ] are not restricted to reference signs mentioned in the description
- r. [ ] do not contain reference signs that are mentioned in the description
- s. [ ] contain the same feature denoted by different reference signs
- t. [ ] are not arranged in an upright position, clearly separated from one another
- u. [ ] are not presented sideways with the top of the figures at the left side of the sheets

2. [ ] satisfactory reproduction (Rules 11 and 26.3(b)(i))

Further observations (if necessary):

ALL NEW DRAWINGS ARE REQUIRED

PATENT COOPERATION TREATY

From the RECEIVING OFFICE

PCT

To: VENKAT KONDA 6278, GRAND OAK WAY SAN JOSE, CALIFORNIA 95135	NOTIFICATION REGARDING CERTAIN CORRECTIONS MADE *EX OFFICIO* (PCT Administrative Instructions, Sections 319 and 327)
	Date of mailing *(day/month/year)* 10 Jun 2008
Applicant's or agent's file reference S-0038 PCT	REPLY DUE NONE However, see paragraph 3 below
International application No. PCT/US2008/064603	International filing date *(day/month/year)*
Applicant KONDA, VENKAT	

1. The applicant is hereby notified that this receiving Office has corrected formal defects in the international application *ex officio*, as shown on the attached copy of:

   - [x] the request, sheet No.: 3, CHECK LIST
   - [ ] the description, sheet No.:
   - [ ] the claims, sheet No.:
   - [ ] the drawings, sheet No.:
   - [ ] other *(specify)*:

2. If the applicant agrees with these corrections, no further action is required in this regard.

3. In case of disagreement with these corrections, the applicant should promply inform this receiving Office accordingly.

Name and mailing address of the receiving Office Mail Stop PCT, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-3201	Authorized officer MOTLEY NINA Telephone No. (703)308-9290 X123

Form PCT/RO/146 (April 2006)

Electronic Patent Application Fee Transmittal

Application Number:	PCT/US08/64603
Filing Date:	22-May-2008
Title of Invention:	FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS
First Named Inventor/Applicant Name:	VENKAT KONDA
Filer:	Venkar Konda
Attorney Docket Number:	S-0038 PCT

International Application for filing in the US receiving office Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Transmittal fee	1601	1	300	300
PCT Search Fee- no prior US appl filed	1602	1	1800	1800
Intl Filing Fee (1st-30 Pgs.) PCT Easy	1701	1	1173	1173
Suppl. Intl Filing Fee (each page > 30)	1703	88	14	1232
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Miscellaneous:				
		Total in USD ($)		**4505**

Electronic Acknowledgement Receipt

EFS ID:	3489255
Application Number:	
International Application Number:	PCT/US08/64603
Confirmation Number:	5416
Title of Invention:	FULLY CONNECTED GENERALIZED BUTTERFLY FAT TREE NETWORKS
First Named Inventor/Applicant Name:	VENKAT KONDA
Customer Number:	38139
Correspondence Address:	VENKAT KONDA - 6278, GRAND OAK WAY - SAN JOSE CA 95135 US 408-472-3273 -
Filer:	Venkar Konda
Filer Authorized By:	
Attorney Docket Number:	S-0038 PCT
Receipt Date:	19-JUN-2008
Filing Date:	22-MAY-2008
Time Stamp:	21:13:47
Application Type:	International Application for filing in the US receiving office

Payment information:

Submitted with Payment	yes

Payment Type	Credit Card
Payment was successfully received in RAM	$4505
RAM confirmation Number	4442
Deposit Account	
Authorized User	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes) /Message Digest	Multi Part /.zip	Pages (if appl.)
1	Fee Worksheet (PTO-06)	fee-info.pdf	8595 f1fa3c8eed0d5c049c1f25ca43e9df9e065ce27e	no	2

Warnings:

Information:

Total Files Size (in bytes):	8595

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To:	PCT
VENKAT KONDA 6278, GRAND OAK WAY SAN JOSE, CALIFORNIA 95135 PALM Tracking Number = 51864603	NOTIFICATION OF RECEIPT OF SEARCH COPY (PCT Rule 25.1)
	Date of mailing *(day/month/year)* 22 Jul 2008
Applicant's or agent's file reference S-0038 PCT	IMPORTANT NOTIFICATION

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US2008/064603	22 May 2008	25 May 2007

Applicant

KONDA, VENKAT

1. Where the International Searching Authority and the receiving Office are not the same Office:
   The applicant is hereby notified that the search copy of the international application was received by this International Searching Authority on the date indicated below.

   Where the International Searching Authority and the receiving Office are the same Office:
   The applicant is hereby notified that the search copy of the international application was received on the date indicated below.

   22 Jul 2008 *(date of receipt).*

2. ☐ The search copy was accompanied by a nucleotide and/or amino acid sequence listing or tables related thereto in computer readable form.

3. Time limit for establishment of international search report and written opinion of the International Searching Authority
   The applicant is informed that the time limit for establishing the international search report and the written opinion of the International Searching Authority is three months from the date of receipt indicated above or nine months from the priority date, whichever time limit expires later (Rules 42.1 and 43*bis*.1(a)).

4. A copy of this notification has been sent to the International Bureau and, where the first sentence of paragraph 1 applies, to the receiving Office.

Name and mailing address of the ISA/	Authorized officer
Mail Stop PCT, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450	Nina D Motley
Facsimile No. 571-273-3201	Telephone No. (703)308-9290 X123

Form PCT/ISA/202 (January 2004)

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To:
VENKAT KONDA
6278, GRAND OAK WAY
SAN JOSE, CA 95135

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT AND THE WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

Date of mailing *(day/month/year)* 03 SEP 2008	
Applicant's or agent's file reference S-0038 PCT	**FOR FURTHER ACTION** See paragraphs 1 and 4 below
International application No. PCT/US2008/064603	International filing date *(day/month/year)* 22 May 2008
Applicant VENKAT KONDA	

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

   Filing of amendments and statement under Article 19:
   The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

   When? The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.

   Where? Directly to the International Bureau of WIPO, 34 chemin des Colombettes 1211 Geneva 20, Switzerland, Facsimile No.: +41 22 740 14 35

   For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.

3. ☐ **With regard to the protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

   ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

   ☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Reminders**

   Shortly after the expiration of **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90*bis*.1 and 90*bis*.3, respectively, before the completion of the technical preparations for international publication.

   The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. These comments would also be made available to the public but not before the expiration of 30 months from the priority date.

   Within **19 months** from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase **until 30 months** from the priority date (in some Offices even later); otherwise, the applicant must, **within 20 months** from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

   In respect of other designated Offices, the time limit of **30 months** (or later) will apply even if no demand is filed within 19 months.

   See the Annex to Form PCT/IB/301 and, for details about the applicable time limits, Office by Office, see the *PCT Applicant's Guide*, Volume II, National Chapters and the WIPO Internet site.

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 Telephone No. 571-272-7774

Form PCT/ISA/220 (January 2004) *(See notes on accompanying sheet)*

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference S-0038 PCT	**FOR FURTHER ACTION** see Form PCT/ISA/220 as well as, where applicable, item 5 below.	
International application No. PCT/US2008/064603	International filing date *(day/month/year)* 22 May 2008	(Earliest) Priority Date *(day/month/year)* 25 May 2007
Applicant VENKAT KONDA		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 3 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**
   a. With regard to the **language**, the international search was carried out on the basis of:
      ☒ the international application in the language in which it was filed
      ☐ a translation of the international application into ______________, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))
   b. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.
2. ☐ **Certain claims were found unsearchable** (see Box No. II)
3. ☐ **Unity of invention is lacking** (see Box No. III)
4. With regard to the **title**,
   ☒ the text is approved as submitted by the applicant
   ☐ the text has been established by this Authority to read as follows:
5. With regard to the **abstract**,
   ☐ the text is approved as submitted by the applicant
   ☒ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority
6. With regard to the **drawings**,
   a. the figure of the **drawings** to be published with the abstract is Figure No. 1B
      ☒ as suggested by the applicant
      ☐ as selected by this Authority, because the applicant failed to suggest a figure
      ☐ as selected by this Authority, because this figure better characterizes the invention
   b. ☐ none of the figures is to be published with the abstract

Form PCT/ISA/210 (first sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2008/064603

Box No. IV Text of the abstract (Continuation of item 5 of the first sheet)

ABSTRACT OF DISCLOSURE
A generalized butterfly fat tree network comprising (logd N) stages is operated in strictly nonblocking manner for unicast, when s > 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 x 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 x d incoming links connecting from switches in its immediate succeeding stage.

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

☐ 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	Date of mailing of the international search report
22 August 2008	03 SEP 2008

Name and mailing address of the ISA/US	Authorized officer:
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2005)