Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

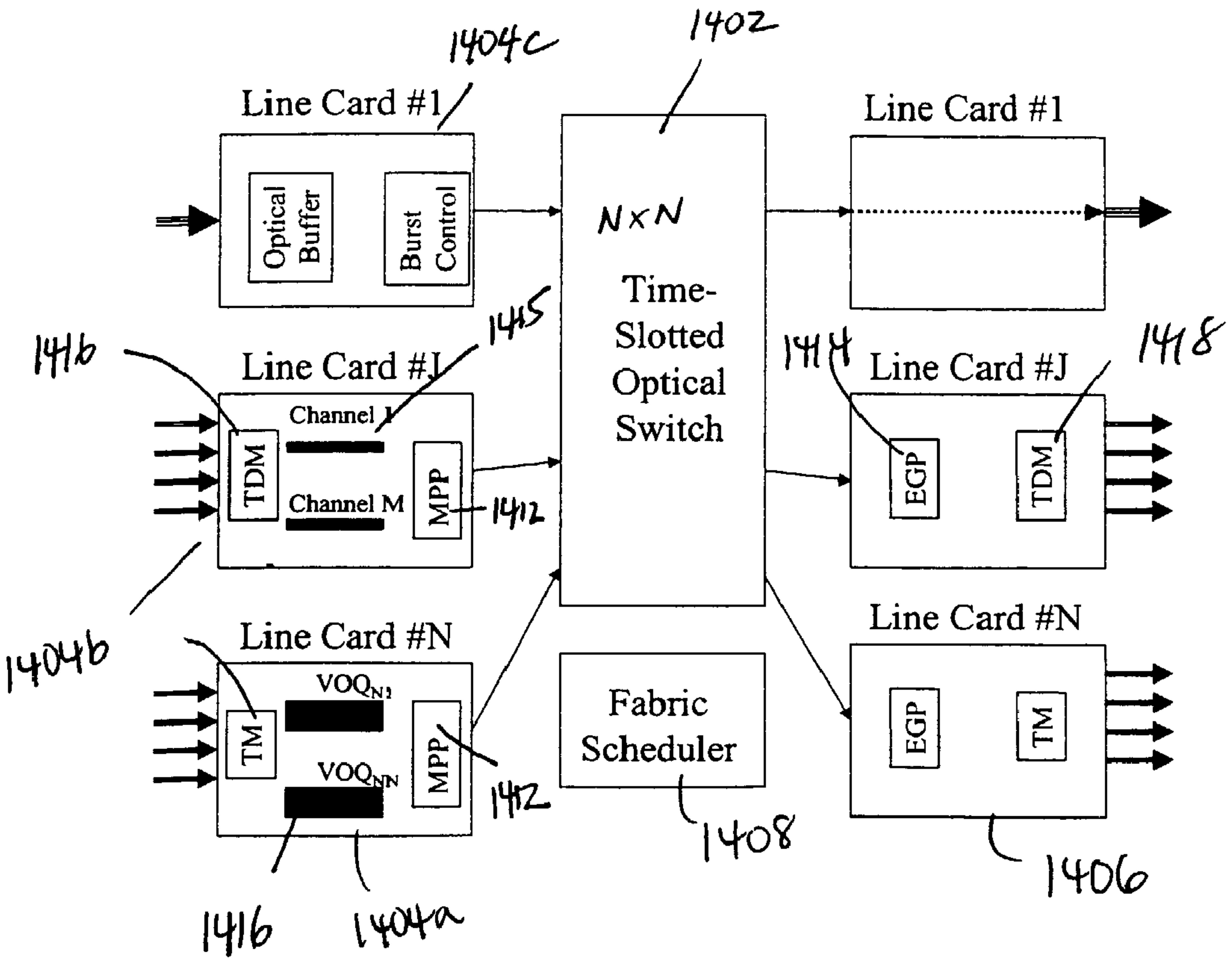
Canadian Intellectual Property Office
An agency of Industry Canada

CA 2411860 A1 2001/12/13
(21) 2 411 860
(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION
(13) A1

(86) **Date de dépôt PCT/PCT Filing Date:** 2001/06/06
(87) **Date publication PCT/PCT Publication Date:** 2001/12/13
(85) **Entrée phase nationale/National Entry:** 2002/12/04
(86) **N° demande PCT/PCT Application No.:** CA 2001/000827
(87) **N° publication PCT/PCT Publication No.:** 2001/095661
(30) **Priorités/Priorities:** 2000/06/06 (2,310,856) CA; 2000/06/06 (2,310,853) CA; 2000/06/06 (2,310,855) CA; 2001/03/06 (2,339,466) CA

(51) **Cl.Int.^{7}/Int.Cl.^{7}** H04Q 11/00, H04L 12/56
(71) **Demandeur/Applicant:**
ACCELIGHT NETWORKS CANADA INC., CA
(72) **Inventeurs/Inventors:**
KIM, HYONG S, US;
LEON-GARCIA, ALBERTO, CA;
ZHANG, JINGYN, CA
(74) **Agent:** DIMOCK STRATTON CLARIZIO LLP

(54) Titre : COMMUTATEUR OPTIQUE MULTISERVICE
(54) Title: A MULTISERVICE OPTICAL SWITCH



(57) Abrégé/Abstract:
A time division multiplexed optical switching system switches a signal from an ingress port to an egress port. The system comprises an optical switching fabric, ingress and egress line cards and a switch control fabric. The optical switching fabric routes an optical signal from one of M ingress fibers to one of N egress fibers. The ingress line card is coupled between the

(57) **Abrégé(suite)/Abstract(continued):**

ingress port and the optical switching fabric for receiving the signal at the ingress port. The ingress line card comprises a buffer for storing the ingress signal and transmits the signal to the optical switching fabric. The egress line card is coupled between the egress port and the optical switching fabric for receiving the optical signal from the egress fiber of the optical switching fabric. The egress line card comprises a buffer for storing the optical signal and transmits the signal to the egress port. The switch control fabric allots time slots to the ingress line card for transmitting to the optical switching fabric.

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

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 December 2001 (13.12.2001)

PCT

(10) International Publication Number
WO 01/095661 A1

(51) International Patent Classification^7: **H04Q 11/00**, H04L 12/56

(21) International Application Number: PCT/CA01/00827

(22) International Filing Date: 6 June 2001 (06.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

2,310,855	6 June 2000 (06.06.2000)	CA
2,310,853	6 June 2000 (06.06.2000)	CA
2,310,856	6 June 2000 (06.06.2000)	CA
2,339,466	6 March 2001 (06.03.2001)	CA

(71) Applicant *(for all designated States except US)*: **ACCELIGHT NETWORKS CANADA, INC.** [CA/CA]; 26 Auriga Drive, Ottawa, Ontario K2E 8B7 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LEON-GARCIA, Alberto** [CA/CA]; 8 Longspur Road, Toronto, Ontario M4B 2Z2 (CA). **ZHANG, Jingyn** [CA/CA]; 26 Auriga Drive, Ottawa, Ontario K2E 8B7 (CA).

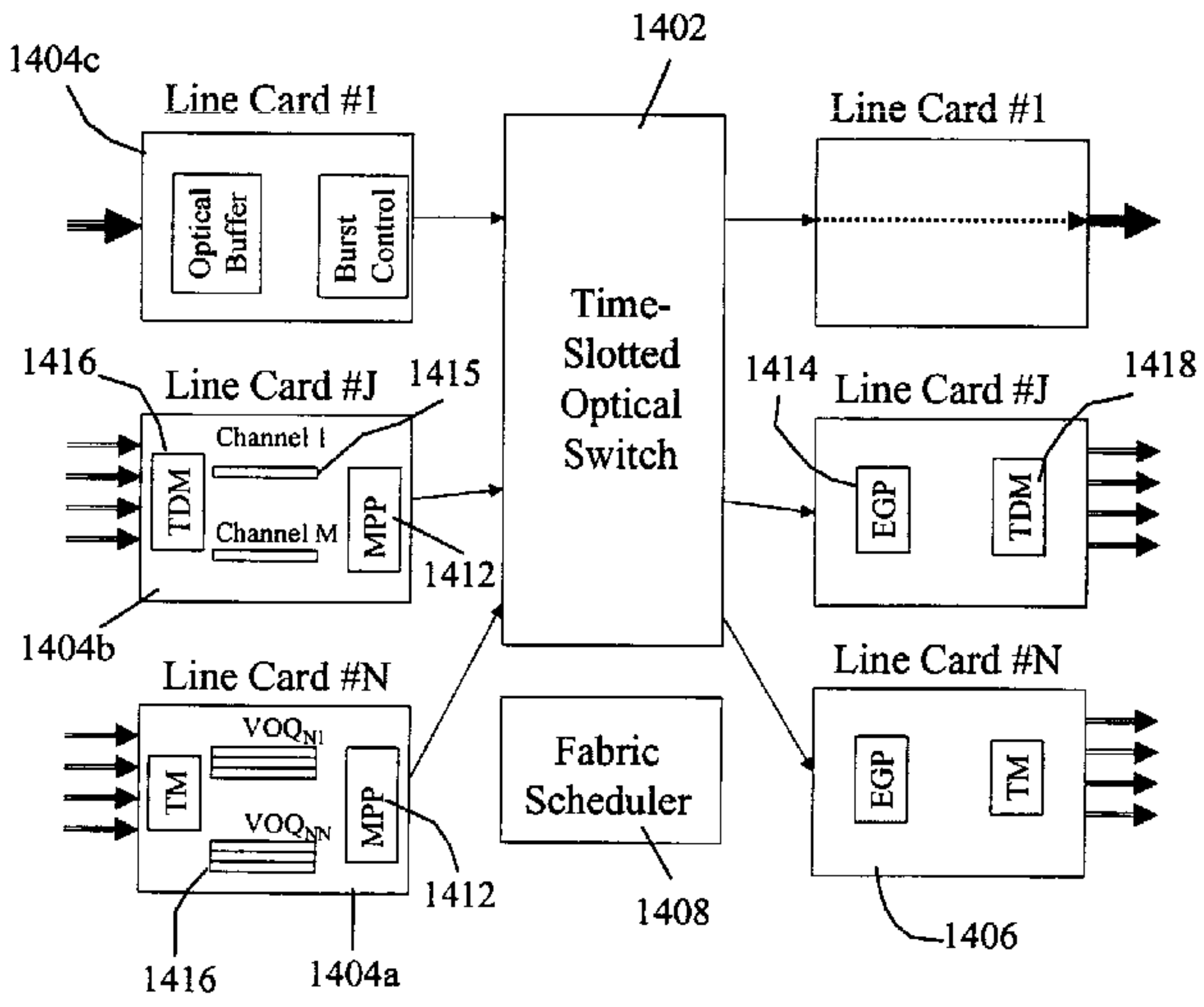
(74) Agent: FASKEN MARTINEAU DUMOULIN LLP; Toronto Dominion Bank Tower, Box 20, Suite 4200, Toronto-Dominion Centre, Toronto, Ontario M5K 1N6 (CA).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian

[Continued on next page]

(54) Title: A MULTISERVICE OPTICAL SWITCH



(57) Abstract: A time division multiplexed optical switching system switches a signal from an ingress port to an egress port. The system comprises an optical switching fabric, ingress and egress line cards and a switch control fabric. The optical switching fabric routes an optical signal from one of *M* ingress fibers to one of *N* egress fibers. The ingress line card is coupled between the ingress port and the optical switching fabric for receiving the signal at the ingress port. The ingress line card comprises a buffer for storing the ingress signal and transmits the signal to the optical switching fabric. The egress line card is coupled between the egress port and the optical switching fabric for receiving the optical signal from the egress fiber of the optical switching fabric. The egress line card comprises a buffer for storing the optical signal and transmits the signal to the egress port. The switch control fabric allots time slots to the ingress line card for transmitting to the optical switching fabric.

WO 01/095661 A1

patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

(48) Date of publication of this corrected version:
19 September 2002

(15) Information about Correction:
see PCT Gazette No. 38/2002 of 19 September 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

A MULTISERVICE OPTICAL SWITCH

The present invention relates to optical switches used in telecommunications and computer networks to switch and route optical signals from one or more ingress ports to one or more egress ports.

BACKGROUND OF THE INVENTION

The transmission of information over optical fiber systems provides advantages of having high transmission rates (measured in bits per second) and low error rates (measured in bits per error). Prior information transmission systems used electronic switches to transfer information from system to system. However, design of electronic switches for optical fiber systems is challenging due to the extremely large volume of information that must be handled electronically. Therefore, optical switches have been developed to improve overall performance of the optical fiber systems. All-optical switches transfer information among optical fiber systems without converting the information streams into electronic form. Hence, they avoid the electronic bottleneck inherent in electronic switches.

The design of an optical switch involves the routing of an incoming optical signal along a desired path. This routing can be accomplished in a number of ways. Mechanical force can be used to move the incoming optical fiber so that it is aligned with the desired outgoing optical fiber as described by S.D. Personick in "Photonic Switching: Technology and Applications," *IEEE Communications Magazine*, May 1987, pp. 5-8.

Mechanical force can also be used to control the incidence angle between an incoming light beam and a mirror in order to reflect the beam to a desired output optical fiber. This approach is used in micro-electromechanical systems as described by T.E. Stern and K. Bala in "*Multiwavelength Optical Networks: A Layered Approach*", Addison-Wesley, Reading, MA. 1999.

Electro-optic effects are also used to control the routing of an optical signal. The index of refraction of a substrate such as lithium niobate can be controlled through the application of an electric field created by a voltage applied across a slab of material. An ion exchange process creates regions of a higher index of refraction in a substrate. A 2x2 optical crosspoint is produced by creating regions of higher refractive index in the shape of two channels or optical waveguides. Voltage control signals are used to direct two incident optical signals to the desired output ports. Larger n input by n output optical switching fabrics are constructed from elementary 2x2 crosspoints using a crossbar arrangement as described by H. Nakajima in "Development on Guided-Wave Switch Arrays" *IEICE Trans. Communications,* Vol. E82-B, No. 2, February 1999, pp. 349-356. Even larger NxN optical switching fabrics are constructed from nxn basic switching fabrics using Clos and Benes multistage switch constructions as described by Joseph Hui in "*Switching and Traffic Theory for Integrated Broadband Networks*", Kluwer Academic Publishers, 1990.

The feasibility of constructing large switches from elementary components such as 2x2 waveguide-based crosspoints is determined by several factors. A loss in signal power incurred in traversing each component determines the maximum number of stages that can be traversed without amplification. Crosstalk that results when the power in one optical signal leaks into another signal affects the integrity of the information that traverses the fabric. The time required to reconfigure each component determines the rate at which the overall switch fabric can be reconfigured. Unfortunately, these factors have limited optical switching networks to be relatively small (due to signal loss) and have high incidences of crosstalk.

Therefore, there is a need for an optical switching fabric that is modular in design and can be built having a small to large number of port counts. The optical switching fabric should be flexible in the type of optical signals that can be carried and rapidly reconfigurable. Furthermore, there is a need for an optical switching fabric that can transmit multiple transmission modes.

SUMMARY OF INVENTION

In accordance with an embodiment of the present invention there is provided a time division multiplexed optical switching system for switching a signal from an ingress port to an egress port. The system comprises an optical switching fabric, ingress and egress line cards and a switch control fabric. The optical switching fabric routes an optical signal from one of *M* ingress fibers to one of *N* egress fibers. The ingress line card is coupled between the ingress port and the optical switching fabric for receiving the signal at the ingress port. The ingress line card comprises a buffer for storing the ingress signal and transmits the signal to the optical switching fabric. The egress line card is coupled between the egress port and the optical switching fabric for receiving the optical signal from the egress fiber of the optical switching fabric. The egress line card comprises a buffer for storing the optical signal and transmits the signal to the egress port. The switch control fabric allots time slots to the ingress line card for transmitting to the optical switching fabric.

BRIEF DESCRIPTION OF THE DRAWINGS

An embodiment of the invention will now be described by way of example only with reference to the following drawings in which:

Figure 1 is a block diagram of a 1xn active splitter;

Figure 2 is a block diagram of $1x2^m$ splitter using binary control signals;

Figure 3 is a block diagram of an active combiner;

Figure 4 is a block diagram of an nxn basic switching unit using active combiners;

Figure 5 is a block diagram of an nxn basic switching unit using passive combiners;

Figure 6 is a block diagram of a 16x16 Benes Switch using identical basic switching units;

Figure 7 is a block diagram of a 16x16 Benes Switch using multiple size basic switching units

Figure 8 is a block diagram of a switch fabric and an associated control unit;

Figure 9 is a block diagram of a strictly non-blocking Clos switch;

Figure 10 is a block diagram of an expanded optical switch using wavelength division multiplexing;

Figure 11 is a schematic diagram of the contents for the interval of a cycle;

Figure 12a is a schematic diagram of a switching schedule;

Figure 12b is a schematic diagram of a traffic matrix and associated switching schedule;

Figure 13 is a block diagram of an on-demand request scheme;

Figure 14 is a schematic diagram of a multiservice switching system;

Figure 15 is a block diagram of a multiservice switching system;

Figure 16 is a block diagram of a dynamic time-slot granting scheme;

Figure 17 is a schematic diagram of a 2Nx2N switching network;

Figure 18 is a block diagram illustrating the offset between two switching fabrics illustrated in figure 17;

Figure 19 is a block diagram of a sample megapacket format;

Figure 20 is a detailed block diagram of an egress megapacker processor;

Figure 21 is a block diagram of a sample megapacket format;

Figure 22 is a detailed block diagram of a megapacket header format;

Figure 23 is a detailed diagram of a sample megapacket format;

Figure 24 is a block diagram illustrating fragmenting a plurality of packets, each smaller than a megapacket, across a plurality of megapackets;

Figure 25 is a block diagram illustrating fragmenting a packet larger than a megapacket across a plurality of megapackets;

Figure 26 is a block diagram illustrating time division multiplexed data in a megapacket;

Figure 27 is a schematic diagram of a multiservice switch using wavelength division multiplexing;

Figure 28 is a schematic diagram of an alternate embodiment of a multiservice switch using wavelength division multiplexing;

Figure 29 is a schematic diagram of an alternate switch that can be used in the multiservice switching system;

Figure 29 is a schematic diagram of an alternate 1x2 switch that can be used in the multiservice switching system;

Figure 30 is a schematic diagram of an alternate 2x2 switch that can be used in the multiservice switching system;

Figure 31 is a schematic diagram of an alternate 4x4 switch that can be used in the multiservice switching system;

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

In the following description, like numerals refer to like structures in the drawings. A recent approach to routing optical signals using an electro-optic effect involves forming of a series of prisms in a segment of a substrate. Such an approach is described by Y. Chiu et al. in "Design and simulation of waveguid electrooptic beam deflectors," *J. Lightwave Technology,* Vol. 13, 1995, pp. 2049-2052, J. Li et al. in "Electrooptic Wafer Beam Deflector in $LiTaO_3$," *IEEE Photonics Technology Letters,* Vol. 8, No. 11, November, 1996, pp. 1486-1488, and Stancil et al. in U.S. Patent No. 5,317,446.

Reversing ferroelectric polarization in triangular-shaped regions in a substrate forms the series of prisms. The value of an applied voltage across the substrate controls the deflection angle of a light beam as it propagates through the substrate, and hence determines the point at which the beam exits the substrate. Hereafter, this component is referred to as an electrooptic wafer beam deflector. The speed at which a light beam's exit point from the electrooptic wafer beam deflector can be reconfigured is limited by the speed of the control voltage signal. The optical signal undergoes very low loss in traversing the component. The electrooptic wafer beam deflector provides a basic building block for an optical switch fabric described herein. The optical switch fabric constructed using multiple electrooptic wafer beam deflectors possesses the desireable properties of low loss, high switching speed, and modular expandibility.

Referring to figure 1, an active splitter is illustrated generally by numeral 100. The splitter comprises an electrooptic wafer beam deflector 12 having a prism segment 13, and a plurality of collimators 15. An incident optical beam 10 enters the electrooptic wafer beam deflector 12, and a voltage V_c is applied to the prism segment 13. The applied voltage V_c determines a deflection angle for selecting an exit point for the deflected optical beam 14. The deflection angle is proportional to the applied voltage V_c. The collimator 15 is arranged so that it directs the deflected optical beam 14 to a corresponding one 16 of a plurality of egress optical fibers 11.

The splitter 100 illustrated in figure 1 is used as a 1x2 splitter by applying a voltage control signal V_c that is either 0 or V volts. The time to switch the optical beam from one position to the other position can be made very small because of the binary nature of the control signal. Referring to figure 2, a $1x2^2$ (or 1x4) active splitter is illustrated generally by numeral 200. Two prism segments 20, under an independent binary control signal 21, are concatenated in a wave beam deflector. The binary nature of the independent control signals enables the $1x2^2$ splitter to have fast transition times. Similarly, a $1x2^m$ splitter comprises *m* concatenated prism segments.

Referring to figure 3, a *n*x1 active combiner is represented generally by numeral 300. Generally the active combiner 300 is obtained by operating the active splitter 100 in reverse. An input optical beam arrives from one of a plurality of ingress optical fibers 30. A collimator 31 directs the optical beam to the wave beam deflector. A voltage signal 32 applied to the prism segment directs the optical beam to a single output fiber 33. An active combiner is more efficient than a passive combiner is for directing the energy in the optical beam to the output fiber.

Referring to figure 4, a 4x4 example of a *n*x*n* switching unit is illustrated generally by numeral 400. The switching unit 400 comprises four 1x4 active splitters 200 and four 4x1 active combiners 300. Each output fiber 40 from each active splitter 200 is coupled with an associated input 41 of each of the active combiners 300. In the present example, the i^{th} output of the j^{th} splitter is coupled with the j^{th} input of the i^{th} combiner.

For each active splitter 200, control voltages V_{c1}, V_{c2}, V_{c3}, and V_{c4} direct the input optical signal to one the output fibers 40. The output fiber 40 propagates the optical signal to a corresponding active combiner 300. The active combiner 300 directs the arriving optical signal to an appropriate output fiber 45 in accordance with one of control voltage signals V_{c5}, V_{c6}, V_{c7}, and V_{c8}.

A consistent set of control voltage signals is used in the *n*x*n* switching unit 400 to direct each of the *n* input optical signals to a distinct set of *n* output ports. The *n*x*n* basic switching unit 400 is equivalent in functionality to a crossbar switch in the sense that it can direct any of the *n* input signals to any output port that is not already in use.

Referring to figure 5, a 4x4 example of a *nxn* switching unit constructed using passive combiners is illustrated generally by numeral 500. The switching unit comprises four 1x4 active splitters 200 and four 4x1 passive combiners 350. Each output fiber 50 from each active splitter is coupled with an associated input 51 of each of the passive combiners. Control voltages V_{c1}, V_{c2}, V_{c3}, and V_{c4} for each active splitter direct the input optical signal to one of the output fibers 53. The output fiber 53 propagates the optical signal to a corresponding passive combiner 350. The passive combiner 350 combines all arriving optical signals and a portion of the energy in the arriving optical signal appears at an output fiber 55. The switching unit 500 provides an acceptable basic switching unit as long as the output signals have an adequate signal-to-noise ratio.

Benes formulated a general method for constructing large switching fabrics from smaller switching fabrics. The term fabric is used to refer to a plurality of switches combined to make a larger switch. The term is synonymous with switching unit, switching network and other terms used in the art. Referring to figure 6 an example of a 16x16 optical switching fabric is illustrated generally by numeral 600. The optical switching fabric 600 comprises three stages of 4x4 optical switching units 61. Each stage in this Benes construction comprises 4 rows of individual 4x4 switching units. Each output 62 from the first stage is coupled via a fiber with an associated input 64 in the second stage. Each output 68 from the second stage is coupled via a fiber with an associated input 66 in the third stage. The i^{th} output 62 from the j^{th} switching unit 61 in the first stage is connected to the j^{th} input 64 of the i^{th} basic switching unit 61 in the second stage. The i^{th} output 68 from the j^{th} switching unit 61 in the second stage is connected to the j^{th} input 66 of the i^{th} basic switching unit 61 in the third stage.

More generally, given a nxn basic switching unit constructed as, it is possible to construct a n^2xn^2 switching fabric using a three-stage construction having the interconnection approach described above. In general a n^2xn^2 three-stage Benes construction requires 3n basic switching units.

A five-stage n^3xn^3 Benes construction for a large switch is obtained as follows. A first stage comprises n^2 rows of nxn basic switching units. A center stage comprises n "central" switches of dimenstion n^2xn^2. A fifth stage also comprises n^2 rows of nxn basic switching units. Each n^2xn^2 central switch can in turn be decomposed into a three-stage array of n rows of nxn basic unit switches, thus resulting in a five-stage switch. For stages one through four, the i^{th} output from the j^{th} basic switch is coupled with the j^{th} input of the i^{th} switch in the following stage. In general, a n^3xn^3 five-stage Benes construction requires $5n^2$ basic switching units. More generally, a n^kxn^k($2k$-1)-stage Benes construction requires $(2k\text{-}1)n^{k-1}$ basic switching units.

A preferred embodiment of the present invention involves the construction of n^2xn^2 and n^3xn^3 Benes constructions of optical switching fabrics using the basic nxn switching units shown in figures 4 and 5. The corresponding three and five-stage switches are feasible because of the low loss property of the basic switching units constructed using the electrooptic wafer beam deflector.

The Benes method also allows the construction of large optical switching fabrics from smaller basic switching units of several sizes. Referring to figure 7 a three-stage 16x16 optical switch fabric constructed from stages of different sizes is represented generally by numeral 700. A first stage and a third stages comprise eight 2x2 basic switching units 71 and a central stage comprising two 8x8 basic switching units 72. In general, an *N*x*N* switch, where *N*=*mn* and *m* and *n* are positive whole numbers, can be constructed in three stages using first and third stages of *m* *n*x*n* basic switching units and a central stage of *n* *m*x*m* basic switching units. Five-stage Benes constructions of dimension *N*x*N*, where *N*=*mnk* and *m*, *n*, and *k* are positive whole numbers. A first stage and a last stage are constructed using *mn* *k*x*k* basic switching units. A middle stage is constructed using k *mn*x*mn* switching units. The k *mn*x*mn* switching units are of the form *N*=*mn* and can therefore be decomposed into a three stage construction as described above.

A preferred embodiment of the present invention involves the construction of three and five stage Benes optical switching fabrics of dimension *N*=*mn* or *N*=*mnk* using the basic switching units illustrated in figures 4 and 5, having sizes n x n, m x m, and/or k x k. The corresponding three and five-stage switches are feasible because of the low loss property of the basic switching units constructed using the electrooptic wafer beam deflector.

All of the Benes switch fabric constructions described above are "non-blocking" in the sense that they can realize any interconnection pattern from any of the N inputs to any of the N distinct outputs. The addition of a new connection to an existing set of less than N existing connections may require the re-arrangement of all connections. For this reason Benes switching fabrics are said to be rearrangeably non-blocking. Various algorithms have been developed for determining the pattern of interconnections within each basic switching units to realize a given overall interconnection pattern in a Benes network.

Clos developed a method for constructing non-blocking multistage fabrics that do not require rearrangement of existing connections when a new connection is set up. The basic Clos construction for an *N*=*pk* switch consists of three-stages. A first and third stage comprises *k* rows of *p*x*m* basic switching units. A central stage comprises *m* *k*x*k* basic switching units. The ith output of the jth switch in the first row is connected to the jth input of the ith central switch. It is well-known that if m=2p-1, then the Clos fabric is strictly non-blocking in the sense that existing connections do not need to be rearranged to establish a new connection from an available input to an available output. Figure 9 shows an example of an 8x8 non-blocking Clos switch constructed from 2x2 and 4x4 basic switching units, illustrated generally by numeral 800. In this example, p=2, k=4 and m=2p-1=3.

A pxm basic switching unit can be constructed by using p of the inputs in an m x m basic switching unit (for m>p). A preferred embodiment of the present invention uses a three-stage arrangement of a Clos switching fabric in which the basic switching units are constructed using the electrooptic wafer beam deflector.

Referring to figure 8, an example of a switching fabric and its associated fabric control unit are illustrated generally by numeral 800. The figure only shows the basic switching units and their associated control signals. Requests for connection patterns are received from elsewhere in the system. The connection matrix request pattern is examined by the fabric control and an algorithm is executed to determine the connection pattern within the basic switching units in the overall switching fabric required to realize the given request pattern. A set of digital control signals c_{ij} is applied to the basic switching units for executing the desired connection patterns. These control signals are converted to voltage levels that cause the optical beams in each basic switching unit to be routed to the appropriate output. The requested interconnection pattern is maintained as long as is necessary by applying the appropriate control voltage signals.

The operation of switching fabrics units as time-slotted optical space switches involves the repetition of a cycle of events as shown in figure 11. Each cycle comprises a guard time t_{config} and a dwell time t_{dwell}. Each cycle is T seconds in duration.

The guard time t_{config} in each cycle provides a time during which the control signals are distributed to the active splitters and combiners and the associated deflection voltages are applied. During the guard time t_{config}, any optical beams present at the inputs may propagate to various outputs in an uncontrolled fashion producing a form of crosstalk. Indicator signals are available at the output of the switches to indicate that the optical signal at the output ports is not valid during the guard time t_{config}. At the end of the guard time t_{config}, each optical beam is deflected from the specified input to the corresponding desired output.

The end of the guard time t_{config}, is followed by the dwell time t_{dwell}. At the beginning of the dwell time t_{dwell}, the input ports are given a signal indicating that switch is ready to switch the bursts of input optical signals. During this interval, the space switch maintains a specific interconnection pattern for directing optical signals from given input ports to specific corresponding output ports. The bursts of input optical signals are transferred to the desired output ports.

The time-slotted optical space switches can operate in a standalone mode and provide transfer of bursts of optical signals from their n input ports to their n output ports. Time-slotted optical space switches of dimension nxm can be obtained by taking a basic switching of a larger size and not using some of the input or output ports.

The configuration of the NxN switch during one cycle T is specified by a matrix $\mathbf{P}(t) = p_{ij}(t)$, where $p_{ij}(t)$ is equal to 1 if input i is connected to output j, and is equal to zero otherwise. $\mathbf{P}(t)$ has the property that each row has exactly one non-zero value, and each column has exactly one non-zero value. The sequence $\mathbf{P}(1)$, $\mathbf{P}(2)$, … $\mathbf{P}(k)$ represents the sequence of interconnection patterns provided for the NxN switch. The number of times an ij^{th} component equals 1 in the sequence $\mathbf{P}(1)$, $\mathbf{P}(2)$, …, $\mathbf{P}(k)$ is the number of time slots allocated to connection ij in k consecutive cycles. Hence, the allocation of transmission opportunities ("bandwidth") among input-output pairs is determined by the sequence of configuration matrices.

The sequence is referred to as a time-division multiplexing (TDM) schedule. Referring to figure 12a there is shown an example of a TDM schedule for a 4x4 switch illustrated generally by numeral 1200. A repetitive pattern of 4 permutation matrices and their associated switch configurations are shown. For the first permutation matrix, $p_{ij}(t)=1$ for $i=j$. Therefore, input 1 is coupled with output 1, input 2 is coupled with output 2, input 3 is coupled with output 3, and input 4 is coupled with output 4. For the second permutation matrix, $p_{ij}(t)$=1 for j=(i+1)mod4. Therefore, input 1 is coupled with output 2, input 2 is coupled with output 3, input 3 is coupled with output 4, and input 4 is coupled with output 1. For the third permutation matrix, $p_{ij}(t)$=1 for j=(i+2)mod4. Therefore, input 1 is coupled with output 3, input 2 is coupled with output 4, input 3 is coupled with output 1, and input 4 is coupled with output 2. For the fourth permutation matrix, $p_{ij}(t)$=1 for j=(i+3)mod4. Therefore, input 1 is coupled with output 4, input 2 is coupled with output 1, input 3 is coupled with output 2, and input 4 is coupled with output 3. Note that in this example the sequence uses only 4 of a possible 4!=24 permutation matrices. Note also that different sequences of permutation matrices can be used to produce TDM schedules.

The sequence of interconnection patterns **P**(1), **P**(2), ... **P**(*k*) can be selected to meet the bandwidth requirements of the traffic that traverses the switch. In the case where the same level of traffic flows between every input and output port and where the traffic flows are relatively steady, a suitable sequence comprises a repetitive interconnection pattern **P**(1), **P**(2), ..., **P**(N-1), **P**(N), **P**(1), **P**(2), ..., **P**(N-1), **P**(N), ... that provides each input-output pair with 1 transmission opportunity per repetition cycle. Each repetition cycle may alternately be referred to as a frame.

In a case where traffic flow differs between various combinations of input/output ports and where the traffic flows are relatively steady, it is preferable to provide a modified schedule accordingly. Such a schedule comprises a repetitive interconnection pattern that provides the input/output pair with a number of transmission opportunities per repetition cycle that is proportional to the relative traffic flow of the input/output pair. Referring to figure 12b an example of a traffic matrix for a 4x4 switch and a corresponding repetitive interconnection pattern that satisfies the traffic demand is illustrated generally by numeral 1250. The ij^{th} entry in the traffic matrix is the proportion of time information is available for transfer from input port i to output port j. The "*x*" in the permutation matrices denote "don't cares" for connections in the switch that have not been assigned. Various algorithms are available for synthesising a repetitive interconnection pattern for a given traffic matrix.

The interconnection pattern can be modified over time to track variations in traffic levels and to deal with temporary surges in traffic. By keeping a running average of the traffic flow between each input/output pair, the variation in the traffic matrix can be tracked and adjustments made. These adjustments may consist of small changes in the permutation matrices or in the repetitive pattern itself through the addition or deletion of one or more permutation matrix. Surges in traffic can be monitored through the backlog of information at the input to the switch. “Don’t cares” in the permutation matrices can be set to help reduce the backlog for certain input-output pairs.

Referring to figure 13, an “on-demand” switching system is illustrated generally by numeral 1300. The switching system comprises an optical switch 1302 coupled to an optical fabric scheduler 1304. The optical switch 1302 has *N* inputs for coupling to *N* input line cards 1306. The optical switch has *N* for coupling with *N* output line cards 1308. In this case, transfer of the time-slotted optical switch is operated “on-demand”, where the transfer for each time slot is computed dynamically. The optical fabric scheduler 1304 accepts requests for packet transfers from the input line cards 1306 through an available signalling system and then executes a scheduling algorithm. The scheduling algorithm determines which input line cards are to be granted permission to transmit to a desired output line card in the next cycle. Algorithms to arbitrate among competing requests from line cards are known in the art.

The complexity of the scheduling algorithm depends on the size *N* of the switch. In general, software implementations of a scheduling algorithm are possible only for switches of small size and/or relatively long-duration time slots. Hardware implementations are required for larger sizes of *N* and smaller values of time slot *T*. For time-slots of duration in the order of microseconds, real-time implementations of the request/grant algorithms are possible for switches with values of *N* in the hundreds.

A combination of pre-allocated and on-demand assignment of transmission opportunities is also possible. A repetitive pattern of the form **P**(1), **P**(2), ..., **P**(N-1), **P**(N), **P**(1), **P**(2),

…, **P**(N-1), **P**(N), … can be used where a subset of cycles are pre-allocated and certain cycles are designated for the request/grant operation. The processing load associated with real-time operation of the request/grant algorithm is lessened by spacing the request/grant cycles evenly in the repetition pattern.

A hybrid operation of a time-slotted optical switching fabric is also possible. The "on-demand" scheduling algorithm is used during all cycles, but pre-allocated traffic is allowed to make requests that are treated with a higher priority than the "on-demand" requests. Thus, pre-allocated traffic is guaranteed to receive its transmission opportunity in its allocated time slot.

In the present embodiment, a multi-service switch based on an optical time-slotted switch is described. Referring to figure 14, a system overview of a multi-service switch is illustrated generally by numeral 1400. The system comprises an *N*x*N* time-slotted optical switch 1402 coupled with a plurality of ingress line cards 1404, egress line cards 1406, and a fabric control system 1408. The ingress line cards 1404 include a megapacket packer (MPP) 1412 and the egress line cards 1406 include an egress megapacket processor (EGP) 1414.

The multi-service switch 1400 can handle packet, TDM, and optical burst traffic flows simultaneously. Each line card 1404 and 1406 in the system 1400 supports one or more services. For example, line card 1404a accepts streams of packets in one or more input lines, transfers the packet streams across the time-slotted optical switch 1402 to an appropriate line card 1406, which transmits them from the multi-service switch in one or more outgoing lines 1410. Line card 1404b accepts streams of time-division multiplexed traffic, for example synchronous optical network (SONET) streams, in one or more ingress lines. The streams are transferred across the time-slotted optical switch 1402 to an appropriate line card 1406, which transmits them from the multi-service switch in one or more egress lines 1410. In the case where packet traffic is transmitted over TDM (SONET) substreams, it is possible to have dual-service line cards that simultaneously handle packet and TDM substreams. Line card 1404c accepts sequences of bursts of optical signals on one or more incoming lines, synchronizes their transfer across the time-slotted optical switch 1402 to an appropriate line card 1406, which transmits them in one or more egress lines 1410.

Referring to figure 15, a system for performing the transfer of information streams across a time-slotted optical fabric is illustrated generally by numeral 1500. Similarly to figure 14, ingress line cards 1404 are coupled to egress line cards 1406 via a switch fabric 1402. Communicating the packet and TDM streams across the fabric 1402 requires conversion of the streams into a format suitable for transfer across a time-slotted optical fabric 1402. A "megapacket" 1502 is used for this purpose. The megapacket 1502 is defined as a block of information that contains control and payload information. In order to transfer a megapacket 1502 across the fabric 1402, it is passed in electronic form to a fabric interface card (FIC) (not shown) and converted into a burst of optical signal. At the beginning of every time slot, synchronized bursts of optical signals are transferred across the fabric in a specific interconnection of inputs and outputs. The size and structure of the megapacket 1502 are selected so that one megapacket 1502 is transmitted serially across the fabric 1402 in one time slot 1504. The payload of the megapacket 1502 comprises the TDM or packet streams that need to be transferred across the multi-service switch. The control section of the megapacket 1502 provides information required for the reconstruction of the TDM and packet streams at each egress line card 1406 after they have traversed the time-slotted optical fabric.

Referring to figure 19, a sample megapacket structure is illustrated generally by numeral 1900. The megapacket comprises a header 1902, followed by a descriptor 1904, a packet 1906, another descriptor 1904, another packet 1906, another descriptor 1904, and padding 1908.

Referring once again to figure 14, the megapacket packer (MPP) 1412 has the task of creating the megapackets 1502. For packet traffic, packets arrive at the system 1400 at the ingress line card 1404a and undergo header processing typical of a packet switching system. This involves extracting packet header fields, carrying out packet classification, and executing processing tasks which may include routing, label switching, or transferring to an "out-of-band" signaling or control system. Typically, packet streams then undergo ingress traffic management, which may include policing/metering and buffer management. Packets are then queued in virtual output queues (VOQs) 1416 for transfer across the fabric 1402. Each VOQ 1416 holds packets destined for a given egress line card. The MPP 1412 takes packets from each VOQ 1416 and forms megapackets that are transferred across the optical fabric to the destination line card. The MPP 1412 maintains queue capacity information and a timer for each VOQ 1416 that stores packets.

The MPP 1412 may employ a variety of strategies for when to make a request for megapacket transmission. For example, when the queue fill exceeds a certain threshold or when the timer expires, the VOQ 1416 may make a request to the fabric scheduler 1402 for a transmission opportunity. In due course, the system receives a transmission grant for a given output port. The MPP 1412 proceeds to construct a megapacket for transmission over the fabric.

At the egress side, megapackets are received by the egress megapacket processor (EGP) 1414. A megapacket buffer (not shown) in the EGP is used to absorb surges in megapacket arrivals. Packets are removed from the EGP megapacket buffer and undergo egress packet traffic processing. The egress processing may include traffic management such as packet transmission scheduling and shaping on the outgoing lines. The egress processing may also include header label processing.

TDM traffic is handled differently than packet traffic in the line cards 1404b. The arriving TDM streams are divided into channels 1415 by an ingress TDM section 1416 and then transferred directly to dedicated buffers in the MPP 1412. Each channel 1415 has a dedicated buffer and its traffic is destined to the same egress line card 1406. A megapacket 1502 is created for a given channel 1415 when the associated buffer reaches a predetermined capacity. The overall time-slotted fabric scheduler 1408 ensures that each channel 1415 receives a transmission opportunity according to a specific predetermined schedule. The number of bytes arriving for a given channel in the time between transmission opportunities for a channel must not exceed the number of bytes that can be transferred in a single megapacket payload. The EGP 1414 receives the sequence of megapackets and transfers the sequence of bytes in the payload to an egress TDM section 1418, which prepares the sequence of bytes transmitted in outgoing lines. In the case of SONET TDM streams, the ingress and egress TDM sections perform the SONET overhead and pointer processing functions.

The manner in which a channel 1415 is defined affects the amount of processing required in the TDM sections 1416, as well as the delay incurred in traversing the time-slotted switch 1402. For example, if the incoming traffic is an OC-192 SONET stream then it is possible to define 192 STS-1 channels. Each channel has its own buffer and fills its own megapacket. Alternatively, in another example, a channel is defined as consisting of all STS-1 substreams destined for the same egress 1406 line card. If the channel comprises *m* STS-1 substreams, then it fills a megapacket *m* times faster than in the case where each STS-1 has its own channel. Although the delay in the latter example is *m* times smaller, it involves more grooming and interleaving of byte substreams by the TDM sections 1416.

TDM channels are assigned regular timeslots for transfer across the optical switch within a repeating fabric frame. Timeslot assignments are calculated by a TDM connection control system, and assignments may be rearranged as new TDM connections are set up. Rearrangement refers to moving an existing connection to a different timeslot to make room for a new connection. In order to maintain continuous transmission out of the egress line cards, the TDM sections 1416 accommodate the worst case rearrangement of a channel's time-slot within a frame. For example, if the ingress line cards into the multieservice switch comprise OC-192 SONET streams, then the frame comprises approximately 192 timeslots per frame. A channel that consists of m STS-1 streams receives m grants in each frame.

Line card 1404c is an example of a line card that handles direct transfer of bursts of optical signals. Bursts of optical signals arrive in the incoming lines. The bursts may be buffered prior to transfer across the optical fabric 1402 by a system of optical delay lines that rearrange arriving bursts so they can be transmitted in an orderly fashion across the fabric 1402. The bursts that are allowed into the optical fabric are also synchronized to the time-slots of the overall multi-service switch.

The fabric scheduler 1408 coordinates the transmission of optical bursts from the various line cards 1404. The scheduler 1408 receives a TDM frame schedule that specifies the time slots when each channel in each line card 1404 is to receive a transmission opportunity. The scheduler 1408 also executes an on-demand scheduling algorithm that determines which packet and burst line cards are allowed to transmit to which output line card at a given time slot. The scheduler 1408 issues grants to the line cards 1406 using a separate transmission link (not shown) that connects the scheduler 1408 to the line cards 1406.

In some implementations, the scheduler 1408 may issue grants for TDM channels 1415. These are tagged as TDM grants, and include the channel number of the ingress card 1404. The MPP 1412 uses the channel number as an index to a table to obtain the egress line card 1406 number, and the channel number for the egress line card 1406. The latter is placed in a field of the megapacket header. The payload is filled with data from the buffer until the payload is full or the buffer is empty.

The time-slotted optical switch 1402 transfers information streams that differ in terms of their traffic properties, for example continuous stream versus bursty arrivals, as well as their delay, delay jitter, and throughput requirements. In order to meet these various demands, the fabric scheduler 1408 handles different types of requests.

Referring to figure 16, a scheme for handling the requests is illustrated generally by numeral 1600. TDM "requests" are received indirectly through the TDM schedule and are accorded strict priority over all other request types. Packet and optical burst traffic send requests that can assume two other "packet" priority levels. The scheduler 1408 maintains request queues for packet and burst traffic and executes an on-demand algorithm that produces the set of grants that are to be issued in any given cycle.

Referring to figure 17, a preferred embodiment of the multiservice switch is represented generally by numeral 1700. The multiservice switch provides $2N$ x $2N$ ports using two NxN time-slotted optical switching fabrics 1702. The switching fabric 1702 is coupled with a plurality of ingress line cards 1404 via a plurality of ingress fabric interface cards (FICs) 1704. Similarly, the switching fabric 1702 is coupled to a plurality of egress line cards 1406 by a plurality of egress FICs 1706. Pairs of ingress line cards 1404 share two ingress FICs 1704 dynamically. Each line card 1404 is capable of transmitting to two FICs simultaneously. Each of the FICs 1704 associated with a pair of ingress line cards is coupled with a different switch fabric 1702. The overall fabric scheduler (not shown) coordinates the transmissions of the line card pairs. The egress line cards 1406 share two egress FICs 1706 dynamically.

In order to avoid sequencing problems, the transmissions across the two fabrics 1702 are staggered by an interval of $T/2$ seconds, where T is the cycle time. Each of the NxN fabrics accepts synchronized bursts of optical signals every T seconds, but an overall system burst transmission time occurs every $T/2$ seconds. The paired line card arrangement results in better utilization of the two NxN optical fabrics and allows line cards to deal better with surges in traffic. The arrangement also allows the multiservice switch to provide protection against a number of fault conditions related to the fabric. The paired arrangement further allows the parallel use of two NxN fabric schedulers. Each scheduler is responsible for determining the set of grants to be issued every other half cycle.

Referring to figure 18, the offset of the optical fibers is illustrated generally by numeral 1800. Although the optical fabrics are offset in time by $T/2$ seconds they remain in frequency and phase alignment. One line card can transmit or receive megapackets on both fabrics concurrently. The two megapackets can be to or from the same line card. Packet ordering between the two fabrics is unambiguous because of the offset.

Encapsulation and segmentation functions in the megapacket and the dual fabric structure of figure 17 require some modification in the operation of the EGP. Referring to figure 20, an EGP is represented generally by numeral 1414. The EGP comprises a first buffer 2002 for receiving information from one FIC and a second buffer 2004 for receiving information from the other FIC. Both buffers 2002 and 2004 are coupled with a megapacket first-in-first-out (FIFO) buffer 2006. The FIFO buffer 2006 is coupled to a megapacket unpacker 2008, which is coupled to an egress traffic manager or egress TDM (not shown).

As previously described, the EGP converts megapackets into packets. This may involve removing padding bytes, and/or assembling packets that spread over multiple megapackets. In the dual fabric structure illustrated in figure 17, a line card may simultaneously receive two megapackets. Megapackets that contain packets are buffered in one of the two FIFOs 2002 and 2004 (one per fabric priority) that can absorb temporary surges in megapacket arrivals. The two FIFOs 2002 and 2004 share a common pool and if the capacity of the pool exceeds a threshold, a backpressure signal 2010 is sent to the fabric scheduler (not shown) so that the megapackets transferred to the line card are stopped. The two FIFOs 2002 and 2004 are serviced in strict priority. When a head-of-line megapacket in the highest priority FIFO is being serviced, its payload is unpacked and the packet stream is passed to the next stage. Long packets that are dispersed over multiple megapackets are further buffered in one of *N* buffers, one per ingress line card. Long packets are transferred to the downstream device when the entire packet has been received at the egress side.

Referring to figure 21, a preferred format of a megapacket is illustrated generally by numeral 2100. The megapacket 2100 begins with a preamble 2102, followed by a header 2104, a payload 2106, and a check 2108. The payload 2106 of the megapacket is filled with packet data from the corresponding packet VOQ or TDM data from the corresponding channel. Any unused space in the payload is filled with zeros. Packet and TDM data are not mixed in a megapacket, nor are packets of different priority type. Packets that cannot be accommodated in a megapacket are segmented and transmission is resumed in the next megapacket.

The preamble 2102 includes physical transmission related fields. During the period before and after the megapacket transmission, each line card transmits a repeating 01 pattern. This is used for clock recovery at the receiver. All data in the megapacket except the preamble 2102 is scrambled to reduce the incidence of long sequences of 1s and 0s that could cause a far end clock recovery circuit to lose lock. For example, a SONET scrambler can be used for this purpose. All data is XORed with a pseudo random binary sequence (PRBS) generated with an x^7+x^6+1 polynomial. The scrambler resets to all ones at the start of each megapacket.

The payload 2106 contains the TDM or packet payload information. The check 2108 is a cyclic redundancy check CRC-16 calculated over the megapacket header and payload (but not the preamble section). The megapacket CRC is used to monitor link quality.

The header 2104 fields are illustrated in figure 22 and defined in table 1. The header 2104 contains source and destination line card addresses. An MP type field specifies the type of the megapacket; TDM, packet, or diagnostic. The MP type field also specifies the priority level of packet-type megapackets. An STS channel field specifies the channel number of the megapacket payload. A Pointer/Length field is defined if the megapacket type is a TDM packet. For packet type, the Pointer points to the start of the first new packet in the megapacket. The field is also used to determine the length of packets in the megapacket. The Pointer/Length field specifies the amount of data in the payload in TDM-type megapackets. A Sequence field is used for recovery in case there are lost or corrupted TDM-type megapackets. SOF Flag and SOF position fields are used by the EGP to reconstruct received SONET frames. The header contains an 8-bit cyclic redundancy check to provide single bit error correction and double bit error detection.

An additional 8-bit cyclic redundancy check is used to detect other uncorrectable errors. CRC is checked after the correction is applied.

Note that packet and TDM-specific fields are undefined for other megapacket types, including for diagnostic megapackets. All undefined and reserved fields are filled with zeros (prior to scrambling).

Referring to figure 23, a format for packets encapsulated within a megapacket is illustrated generally by numeral 2300. A packet length field (PLN) field specifies the length of a packet in bytes, excluding CRC and padding. A reserved (RES) field is set to zero. A forward error correction (FEC) provides single-bit error-correction capability over the PLN and RES fields. The packet in the payload may contain additional control fields. For example, an MPLS field contains an MPLS header for MPLS packets. (Note that ATM or other types of packet traffic can also be encapsulated in a megapacket.) All packets that are encapsulated in a megapacket are padded with zeros to align the packet (including CRC) to 32-bit boundaries. The EGP detects this when it looks for a length field following the last packet. The EGP interprets zero length as the end of a valid payload. A CRC-16 check sum is calculated over the entire encapsulated packet excluding the Pad and CRC fields. RES and FEC bits are set to 0 and the packet is padded to 128 bits. The CRC is calculated over whole 128-bit words. The CRC is then inserted on the nearest 32-bit boundary.

Packets of the same fabric priority that are intended for the same destination are queued together in the same VOQ, then packed into megapackets. Packets are padded to 32 bit boundaries within the megapacket payload. At the end of the megapacket, the last packet may be fragmented. Fragmented packets are always continued in the next megapacket and sent to the same egress line card with the same priority.

Referring to figure 24, an example of packet fracturing is illustrated generally by numeral 2400. In this example, the ingress packets are smaller than a megapacket. Therefore a first megapacket comprises a first ingress packet and a portion of the second ingress packet. A second megapacket comprises the remainder of the second ingress packet and at least a portion of a third ingress packet.

Referring to figure 25, another example of packet fracturing is illustrated generaly by numeral 2500. In this example, the ingress packets are larger than the megapacket. Therefore, a first, second, and third megapacket each comprises a portion of a first ingress packet. Other examples of packet fracturing will be apparent to a person skilled in the art.

For TDM transmissions, each TDM channel is transmitted in separate megapackets. Each megapacket can be partially full. Data is placed in each megapacket in units of 128 bits. The length of the data is transmitted in the MP header. The unused payload space is filled with zeros. An example of a TDM megapacket is illustrated generally in figure 26 by numeral 2600.

The transmission speed across the time-slotted fabric is determined by the optical modulation technique used in the FIC. In one class of examples, the transmission speed need only be slightly higher than the highest speed of the information arriving in the line cards. For example, if the highest input speed into a line card is 10 Gbps, then a transmission speed of 12.5 Gbps across the fabric may suffice to compensate for the overhead incurred in the megapacket headers and the fabric reconfiguration times. However, the inherent transmission capacity of the optical fabric is very high. Therefore, the transmission across the fabric can be increased by using higher transmission speeds in the FIC, for example 40 Gbps, or by introducing wavelength division multiplexing WDM.

The electrooptic wafer beam deflector can route optical signals and maintain high signal quality even when the optical signals are composite and consist of multiple wavelength signals. Consequently, the optical switches described above have the capability of transferring composite optical signals. Referring to figure 10, an optical switching network using wavelength division multiplexing (WDM) is illustrated. A multiplexer 90 concentrates multiple optical signals that occupy non-overlapping wavelengths into a single optical signal that can be switched across the *N*x*N* optical switch. The structure of the switch constrains all components of the composite signal to be switched to the same output port. The composite signal can then either be decomposed into individual components by a multiplexer 91 or the entire composite signal can be transmitted from the switch to an outgoing optical transmission link. Each additional wavelength in the composite signal increases the transmission-carrying capability (measured in bits) in each time-slot. The transmission-carrying capability of the overall switch increases accordingly.

Referring to figure 27, a system implementing WDM multiplexers and demultiplexers is illustrated generally by numeral 2700. Ingress line cards 2702 are coupled via a multiplexer 2704 to an *N*x*N* optical switch fabric 2706. Egress fibers from the optical switch fabric 2706 are coupled via a demultiplexer 2708 to egress line cards 2710. The system is used for concentrating multiple optical signals that occupy non-overlapping wavelengths into a single optical signal that is switched across the *N*x*N* optical switch 2706. The structure of the switch 2706 constrains all components of the composite signal to be switched to the same output port. A group of *n* ingress line cards 2702 simultaneously transmit to group of *n* egress line cards 2710. The wavelength that is assigned to each ingress line card 2702 determines which egress line card 2710 receives its megapacket.

Referring to figure 28, a more versatile and efficient system than that illustrated in figure 27 is illustrated generally by numeral 2800. A group of *m* ingress line cards 2802 is coupled with a first level-1 switch 2804. The first level-1 switch 2804 is coupled with a level-2 switch fabric 2808 via a multiplexer 2806. The level-2 switch fabric 2808 is coupled with a second level-1 switch 2812 via a demultiplexer 2810. The second level-1 switch 2812 is coupled to a group of *m* egress line cards 2814. The level-2 switch fabric 2808 is further coupled with a fabric scheduler 2816.

The operation of the system 2800 is described as follows. The *m* ingress line cards 2808 transfer TDM or packet information to the first level-1 switch 2804. Each first level-1 switch 2804 is coupled to a single multiplexer. The first level-1 switch 2804 buffers and prioritises packet traffic according to a destination egress optical port. The level-1 switch 2804 also grooms TDM traffic according to the destination egress optical port. The level-1 switch 2804 may use one or more of the wavelengths of a given optical port. The first level-1 switch 2804 is a *m*x*n* switch, which may be an optical switch but does not need to be. Similarly, the second level-1 switch 2812 is a *n*x*m* switch, which may be an optical switch but does not need to be.

The fabric scheduler 2816 handles requests for time slots from the level-1 switches 2804, and issues grants for time-slot transmission opportunities from a given input optical port to a given destination output optical port. The level-1 switches 2804 determine how bandwidth is allocated among its attached line cards 2802. The level-2 switch fabric is only concerned with the allocation of bandwidth between its ports. This hierarchical approach simplifies the overall scheduling and makes the system more scalable to larger numbers of line cards.

Various modifications to the connections of the system 2800 will be apparent to a person skilled in the art. For example, it is possible for two groups of line cards to share a multiplexer rather than have a designated one. Furthermore, it is possible for two groups of line cards to share a number of multiplexers.

In an alternate embodiment, a basic switching element other than the electrooptic wafer beam deflector is used. In Canadian Patent Application 2,339,466, Zhang et al. used ferroelectric polarization reversal techniques on new region geometries to develop a new class of optical beam switches. An applied voltage is used to trigger total internal reflection (TIR) at one or more interfaces. Methods are described for the construction of compact switches of format 1 × 2, 1 × 3, 1 × 4, 2 x 2 or 4 × 4, and so on. Once the basic switching blocks are described, it will be apparent to a person skilled in the art how to utilize the switches in the various configurations described in the previous embodiment.

These switches have no moving parts and can be reconfigured at a very high speed. Since switching is implemented by controlling the state of TIR, the cross-talk is extremely low and switching is not sensitive to the variation of an operating electric field. Therefore, an optical switch may be built that has a simple structure, offers low insertion loss, and operates independently of beam polarization and wavelength.

Referring to figure 29, a 1x2 TIR optical switch is illustrated generally by numeral 2900. The switch 2900 comprises an ingress fiber collimator 202, a piece of electrooptic crystal 201 and two egress fiber collimators 208 and 210. The electooptic crystal 201 can be either $LiTaO_3$ or $LiNbO_3$, or other material with high electrooptic coefficient as will be apparent to a person skilled in the art. Inside the crystal 201 simple poled and unpoled structures are created. The structures are interfaced with each other by a straight line at an angle of approximately 1°.

A collimated beam is launched from the input collimator 202 and enters the crystal 201 along the straight path 203. If there is no electrical field applied, the beam will propagate, exit the crystal 201 and enter a 1st egress collimator 208. If there is an electrical field applied, it is applied in such a way that a first portion 206 of the crystal has a higher refractive index than a second portion 204 of the crystal, causing a TIR condition to be true. The beam is then reflected at an interface 205 of the first 206 and second 204 portion, exits the crystal 201, and propagates along path 207 to enter the 2nd collimator 210.

The refractive index is a function of wavelength. For longer wavelengths the values of refractive index become smaller. When a stream of wavelengths are launched into the switch 2900, as long as TIR is maintained for the longest wavelength it is also maintained for all shorter wavelengths. Therefore, the 1×2 optical switch 2900 functions independently of the wavelengths. This is another very attractive and important feature for the switch application in a WDM network system.

Referring to figure 30, an example of a TIR-based 2×2 optical switch is illustrated generally by numeral 3000. This switch 3000 has a symmetrical structure that simultaneously operates on two optical beams using the TIR method. It comprises two ingress fiber collimators 502 and 504, one piece of electrooptic crystal 501, which has a simple sandwich structure of poled and unpoled areas, and two egress fiber collimators 538 and 540. The switch-controlling electrode covers the whole of areas 520,522 and 524. Area 524 is sandwiched between area 250 and 522 and is very thin in order to minimize the beam walk-off as beams are switched.

At a first switch position, no electrical field is applied. A collimated beam emits from the ingress collimator 502, propagates along a straight path 506, penetrates crystal 501 and propagates along a straight path 530 to enter the egress collimator 538. Concurrently, a collimated beam emits from the ingress collimator 504, propagates along a he straight path 508, penetrates crystal 501 and propagates along a straight path 532 to enter the output collimator 540. At a second switch position, a controlling electrical field is applied such that TIR is achieved at the interfaces of area 520 and 524, and 522 and 524. A beam emitted from collimator 502 travels along path 506, is reflected to path 532, and enters collimator 540. A beam emitted from collimator 504 travels along path 508, is reflected to path 530, and enters collimator 538. The switch 3000 is also wavelength independent.

Referring to figure 31, an example of a TIR-based 4x4 optical switch is illustrated generally by numeral 3100. The working principle of this switch is the same as that of the 1x2 switch 2900 and the 2x2 switch 3000. The switch 3100 comprises four ingress fiber collimators 820, 822, 824, and 826, one major piece of electrooptic crystal 800, two pieces of crystal 802 and 804 for setting up an entrance path for beam 830 and 836, four output fiber collimators 880, 882, 884, and 886. The main crystal piece has five poled and unpoled sections, each of which functions as 2x2 switch node. There are 5 pairs of switch controlling electrodes, each of which is placed over the 2x2 poled and unpoled sections. Collimated beams emit from the input collimator 820, 822, 824, 826 and propagate along the straight path 830, 832, 834 and 836, respectively. Along these four paths, light beams intersect with the five 2x2 switch nodes and two air/crystal TIR surfaces, exit the crystal along four output path lines 870, 872, 874 and 876, and enter the four output collimators (880, 882, 884, 886). At any moment, from zero to five pairs of switch controlling electrodes are turned on to set up TIR condition. This leads to 24 non-redundant cross-connection combinations between the four input collimators and four output collimators.

Although the invention has been described with reference to certain specific embodiments, various modifications thereof will be apparent to those skilled in the art without departing from the spirit and scope of the invention as outlined in the claims appended hereto.

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. A time division multiplexed optical switching system for switching a signal from an ingress port to an egress port, said system comprising:
    (a) an optical switching fabric for routing an optical signal from one of *M* ingress fibers to one of *N* egress fibers;
    (b) an ingress line card for coupling between said ingress port and said optical switching fabric for receiving said signal at said ingress port, said ingress line card comprising a buffer for storing said ingress signal, said ingress line card for transmitting said signal to said optical switching fabric;
    (c) an egress line card for coupling between said egress port and said optical switching fabric, for receiving said optical signal from said egress fiber of said optical switching fabric, said egress line card comprising a buffer for storing said optical signal, said egress line card for transmitting said signal to said egress port; and
    (d) a switch control fabric for allotting time slots to said ingress line card for transmitting to said optical switching fabric.

2. A time division multiplexed optical switching system as defined in claim 1, wherein if said signal is a non-optical signal:
    (a) said ingress line card further includes a packer for formatting said stored signal and a fabric interface card for converting said non-optical signal to an optical signal; and
    (b) said egress line card further includes a fabric interface card for converting said optical signal to a non-optical signal and a processor for processing said reassembling said converted signal.

3. A time division multiplexed optical switching system as defined in claim 2, wherein said optical switching fabric is independent of said optical signal's wavelength.

4. A time division multiplexed optical switching system as defined in claim 3, further comprising:
    (a) a wavelength division multiplexer for combining a plurality of optical signals having mutually exclusive wavelengths into a single optical signal; and
    (b) a wavelength division demultiplexer for separating said combined signal optical signals into its constituent signals,

whereby said plurality of signals are transmitted to a common egress fiber of said switching fabric in the same time slot.

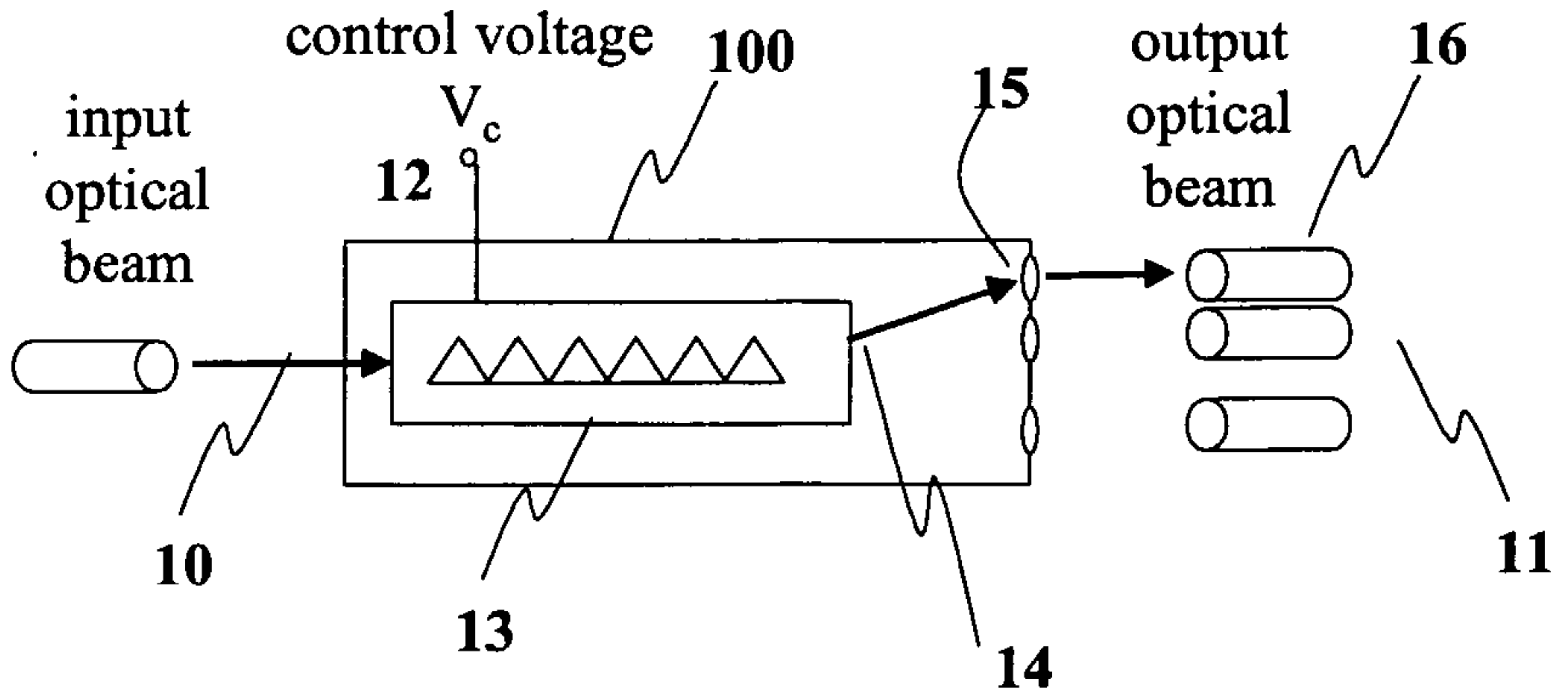


Figure 1

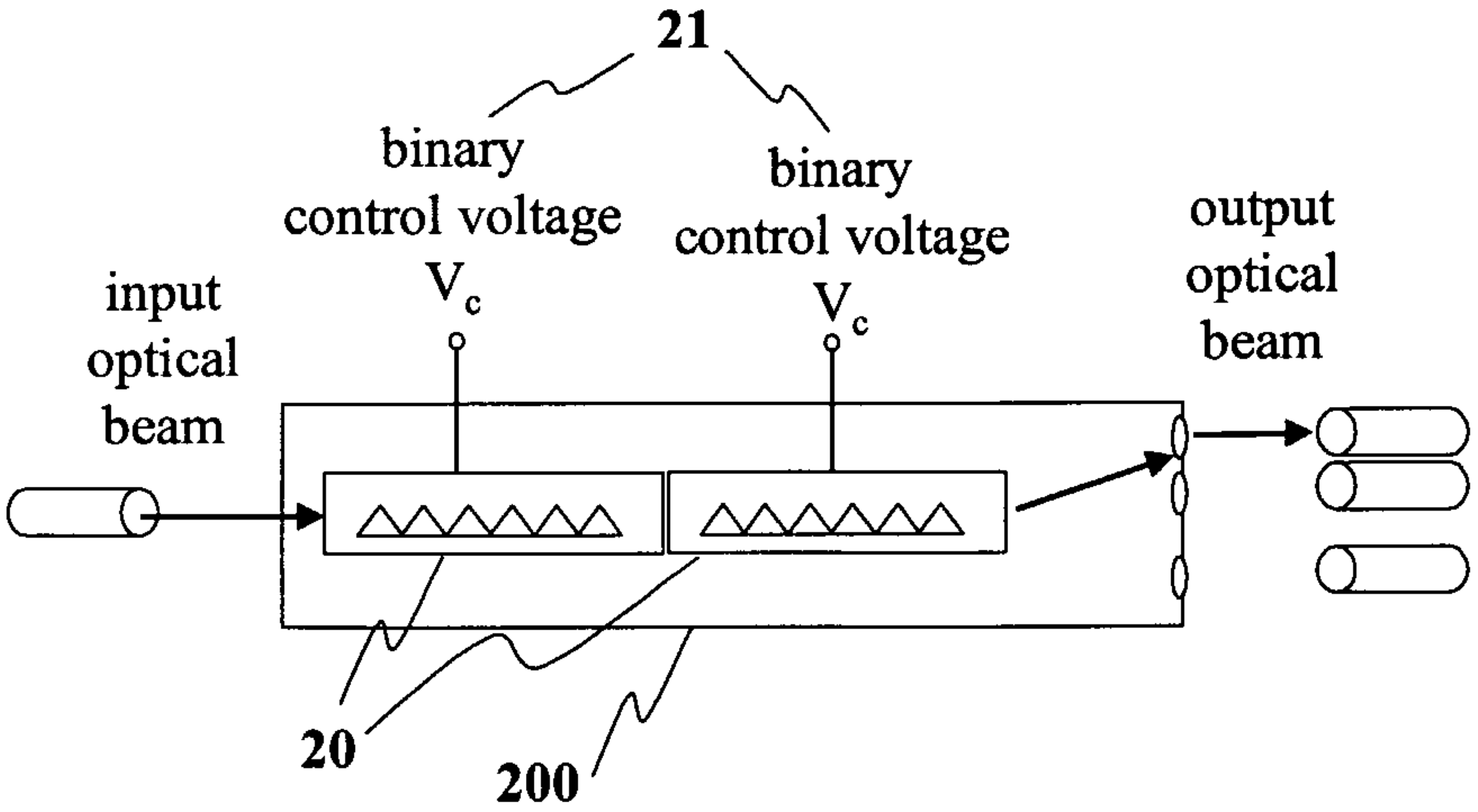


Figure 2

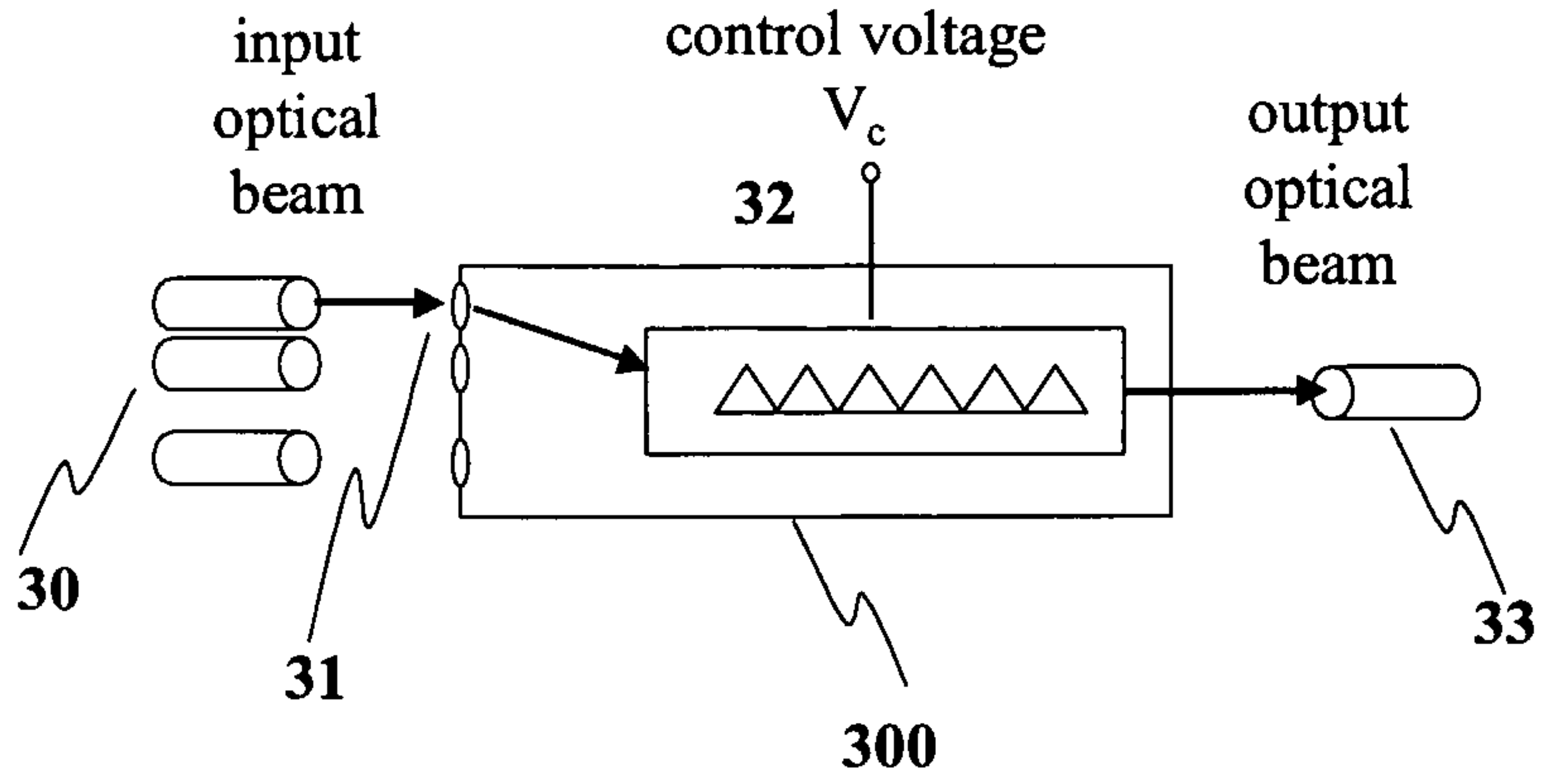


Figure 3

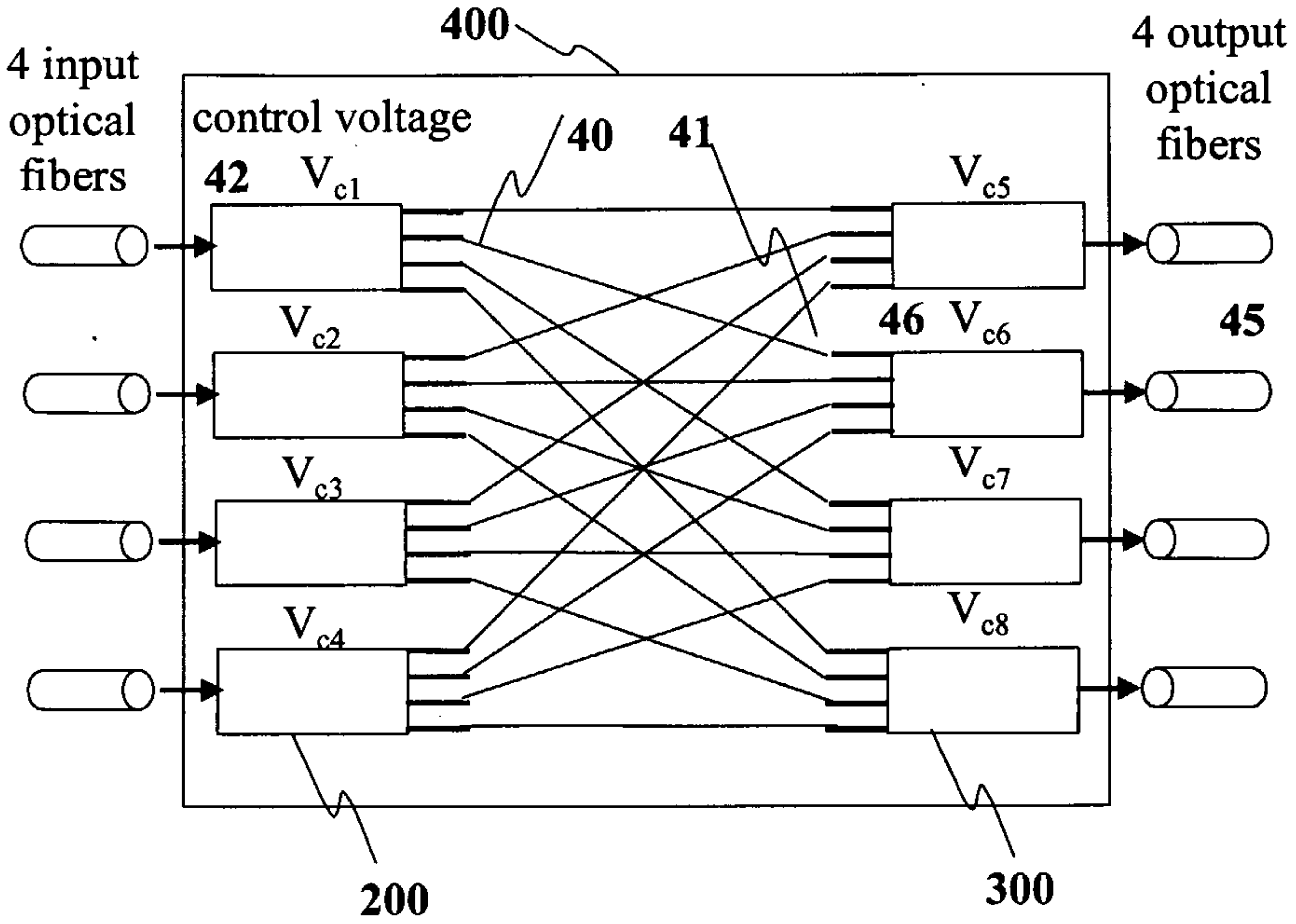


Figure 4

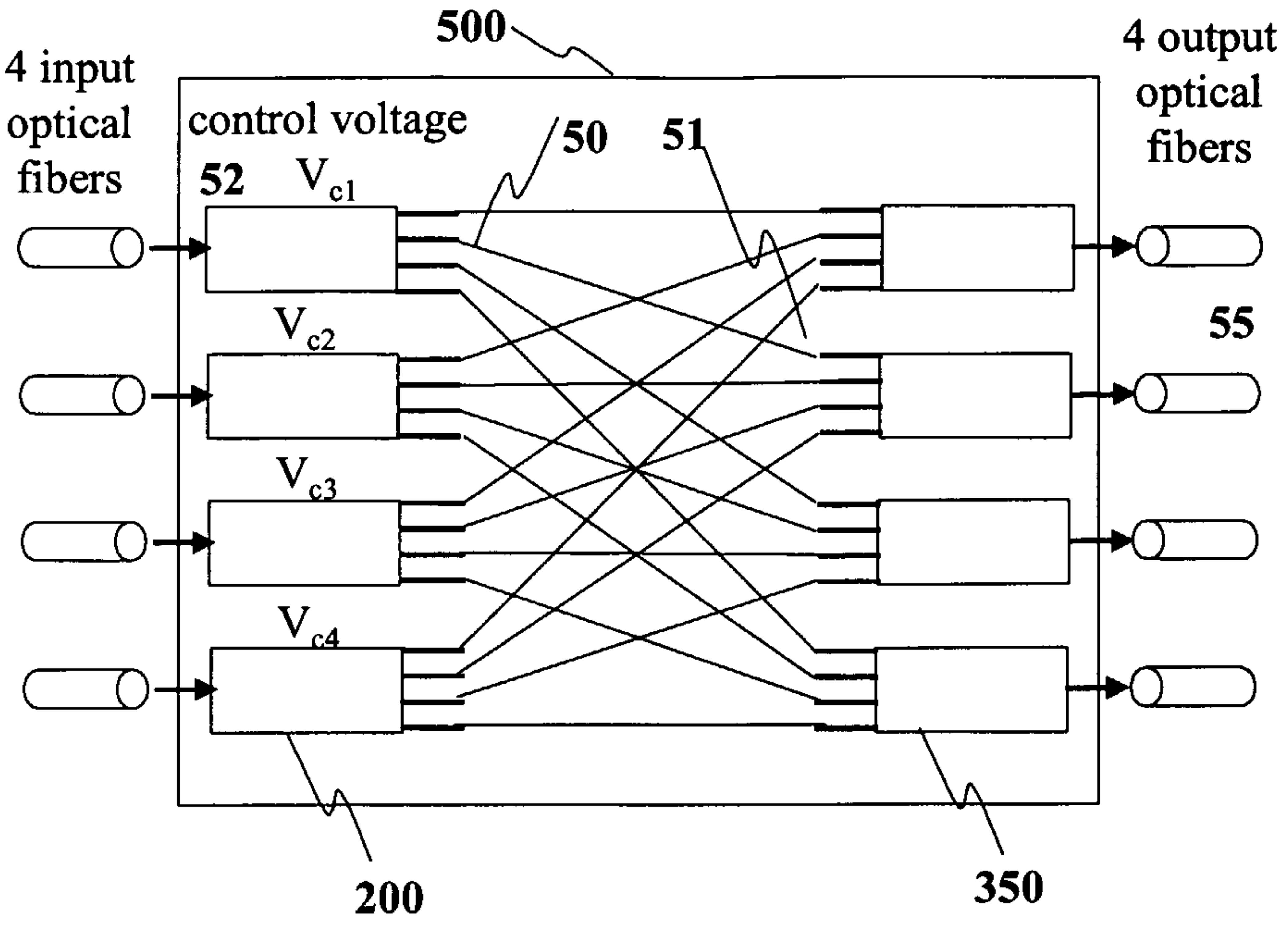


Figure 5

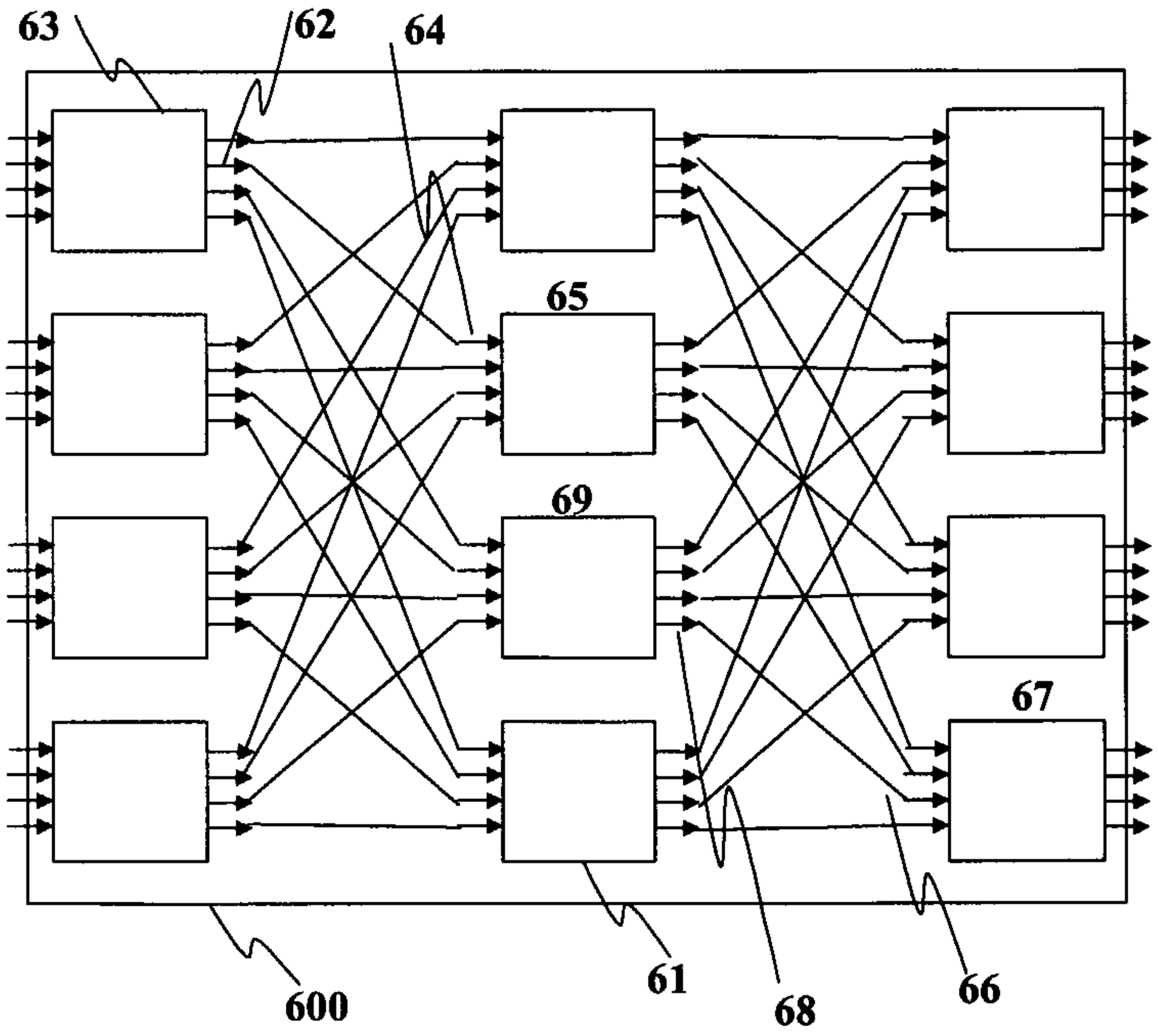


Figure 6

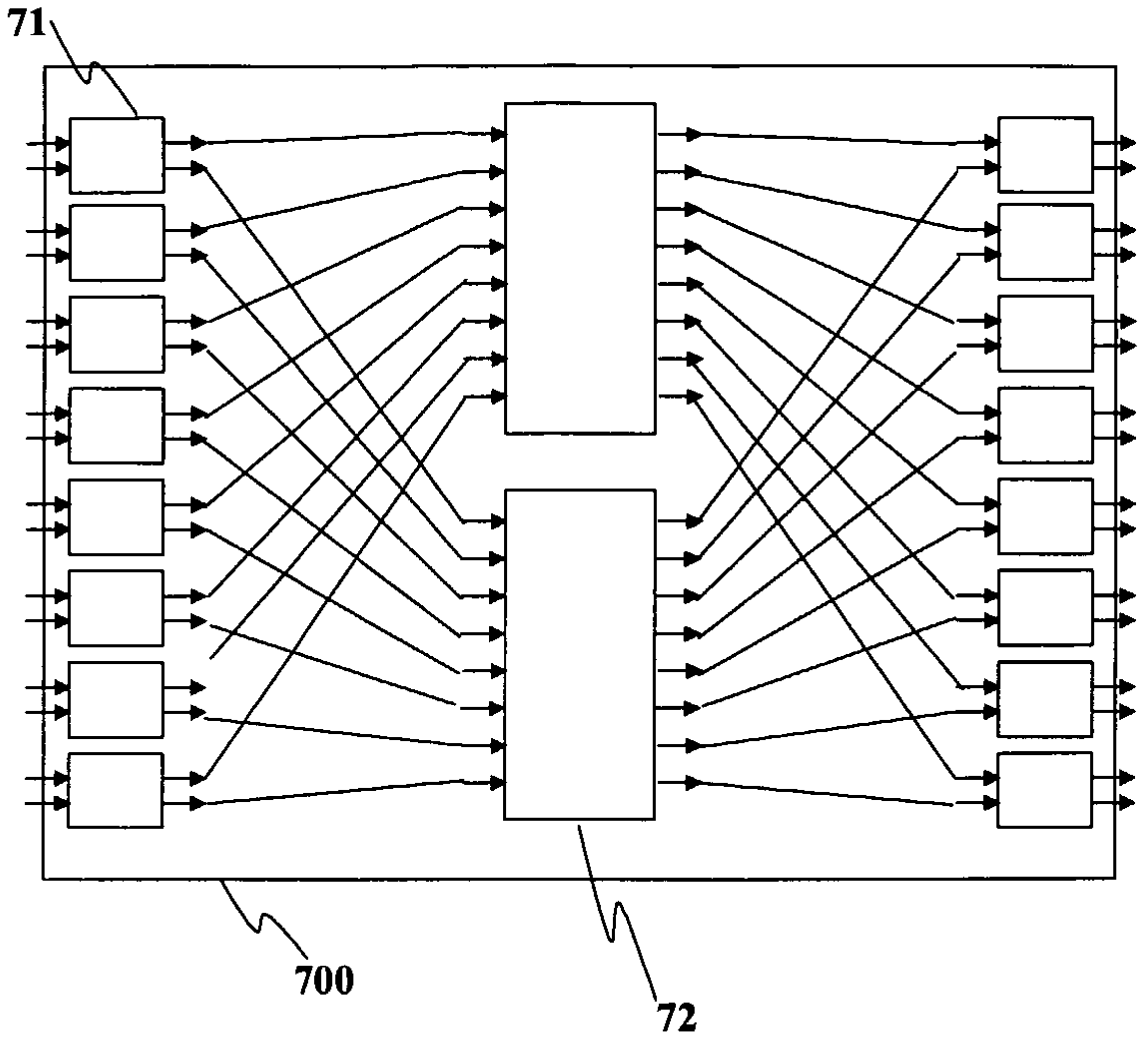


Figure 7

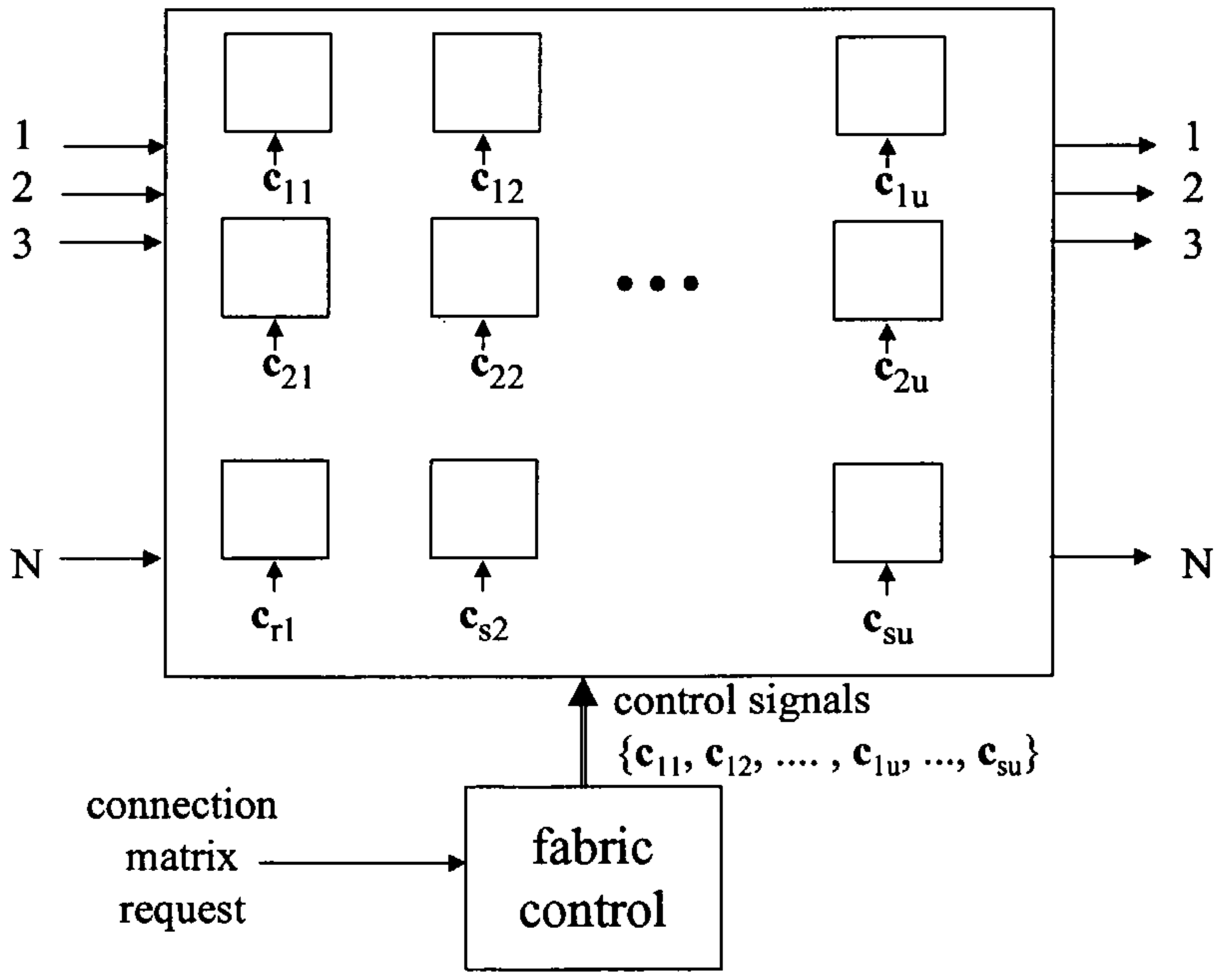


Figure 8

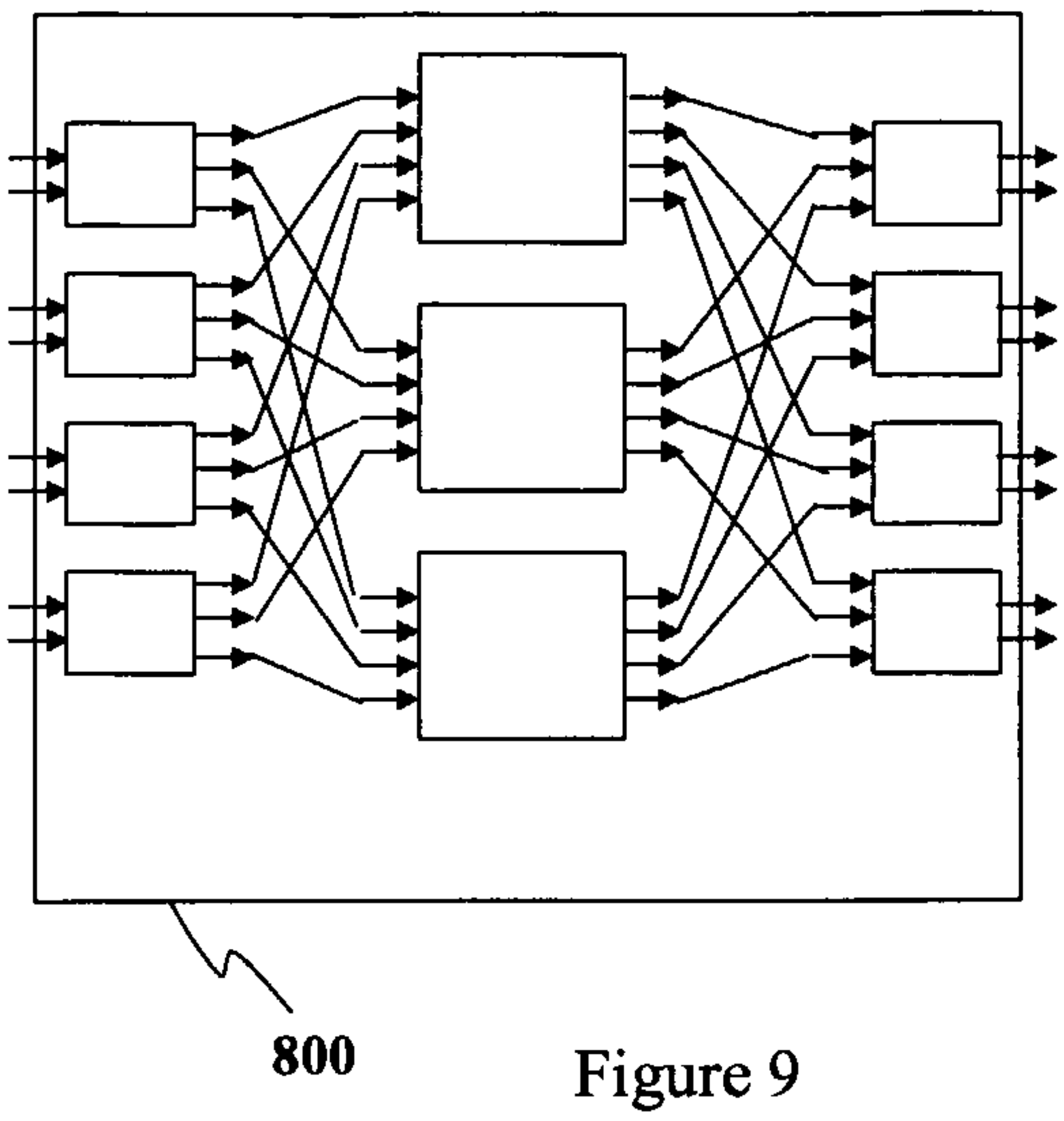


Figure 9

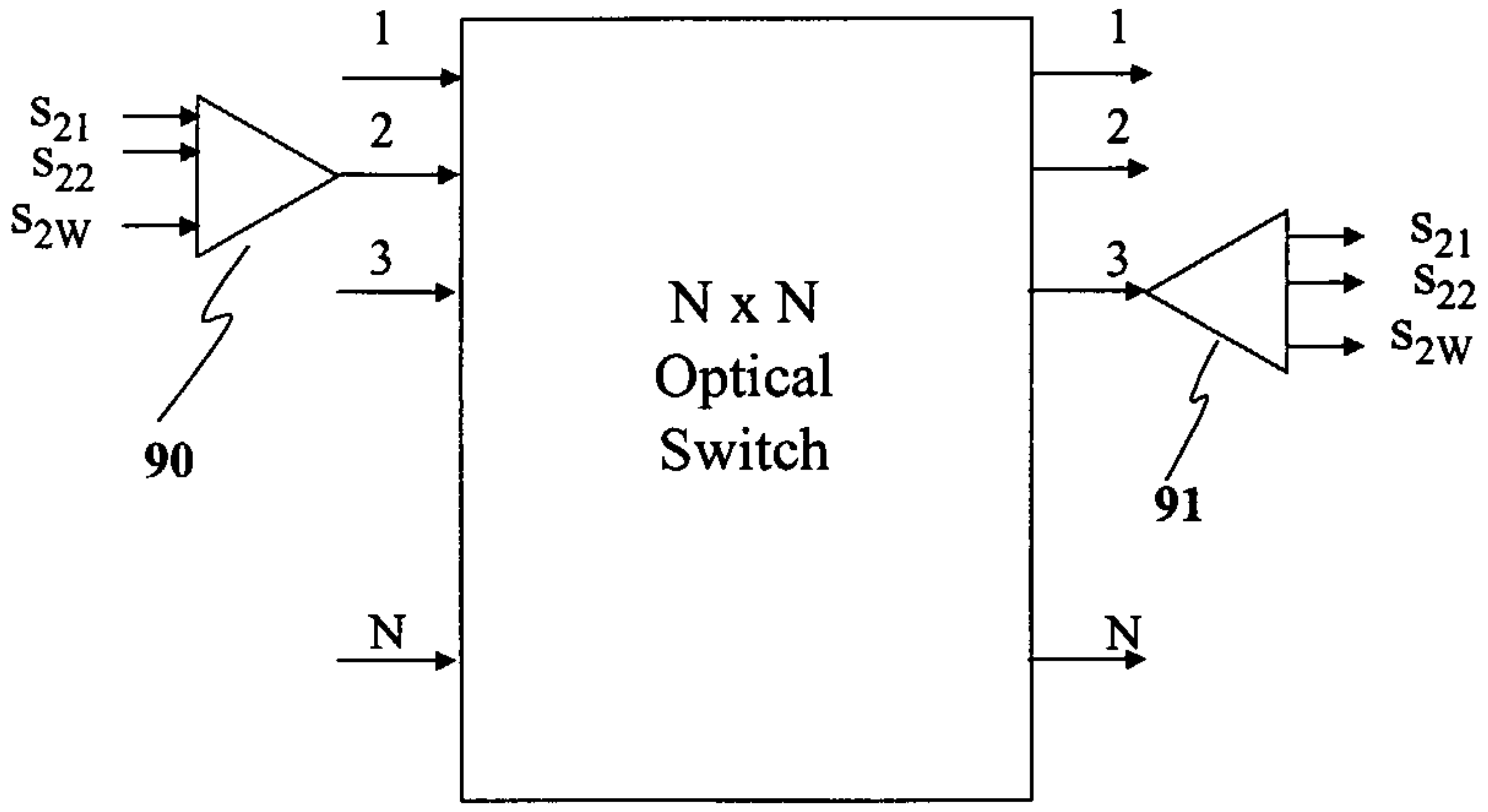


Figure 10

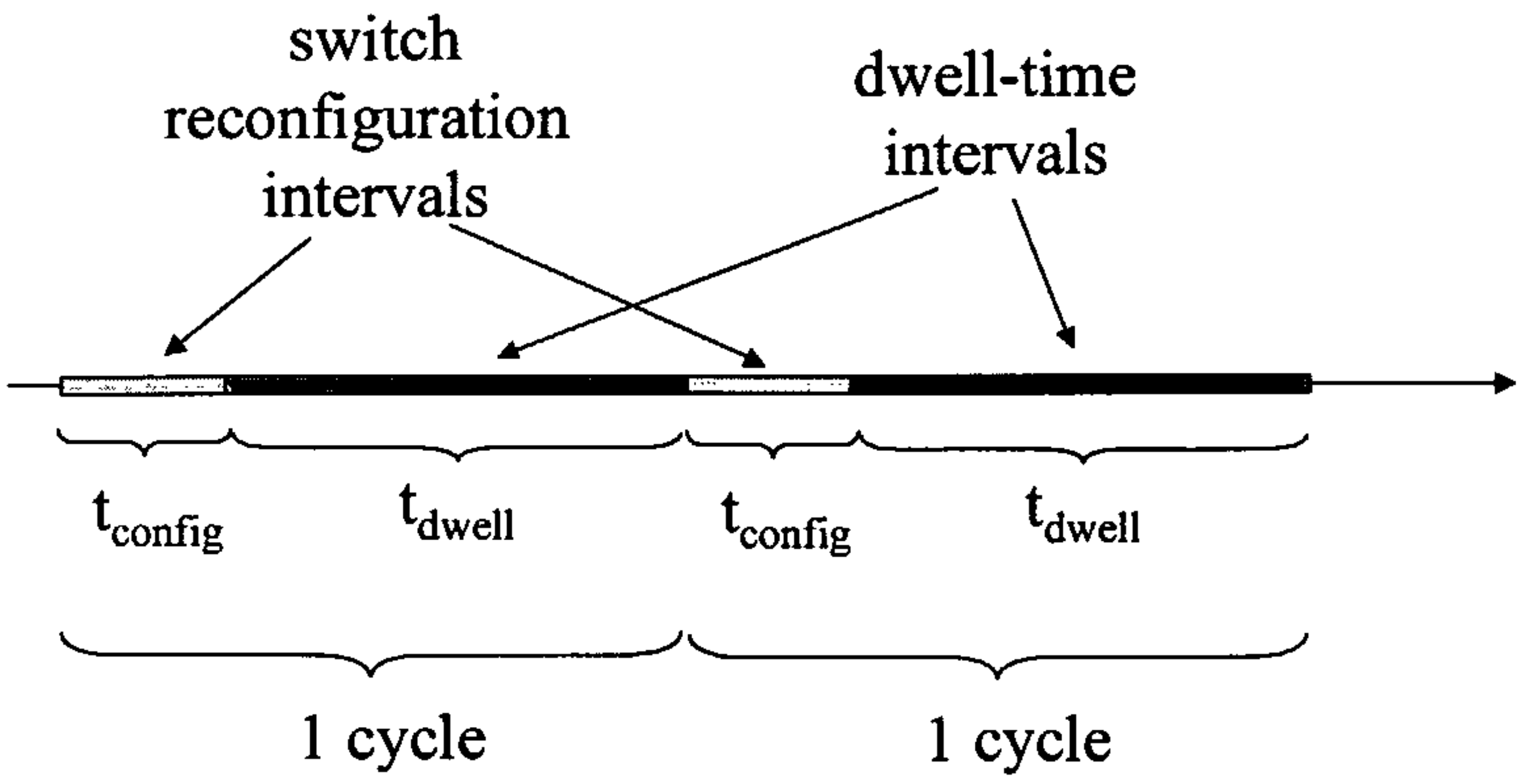


Figure 11

$$\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\quad\begin{pmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{pmatrix}\quad\begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{pmatrix}\quad\begin{pmatrix}0&0&0&1\\1&0&0&0\\0&1&0&0\\0&0&1&0\end{pmatrix}$$

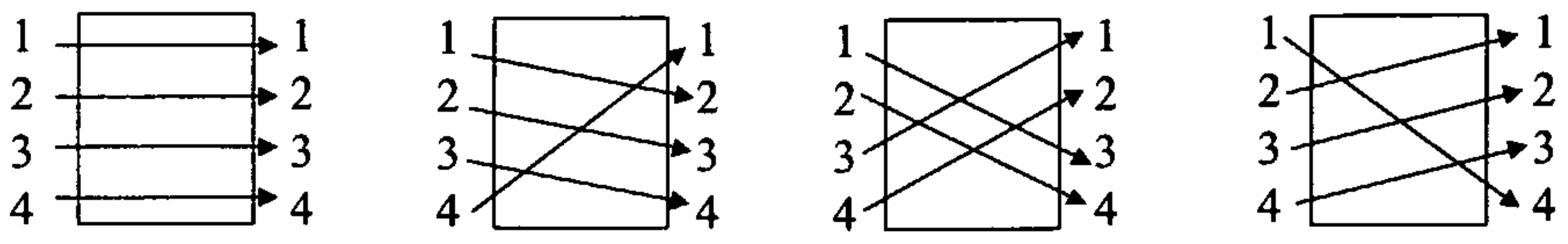


Figure 12a

$$\text{traffic matrix}\quad\begin{pmatrix}0.0&0.4&0.2&0.2\\0.2&0.0&0.4&0.2\\0.2&0.2&0.0&0.2\\0.2&0.4&0.4&0.0\end{pmatrix}$$

$$\begin{pmatrix}0&0&1&0\\0&0&0&1\\0&1&0&0\\1&0&0&0\end{pmatrix}\quad\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}\quad\begin{pmatrix}0&1&0&0\\x&0&0&x\\x&0&0&x\\0&0&1&0\end{pmatrix}\quad\begin{pmatrix}0&0&0&1\\0&0&1&0\\1&0&0&0\\0&1&0&0\end{pmatrix}\quad\begin{pmatrix}x&0&0&x\\0&0&1&0\\x&0&0&x\\0&1&0&0\end{pmatrix}$$

Figure 12b

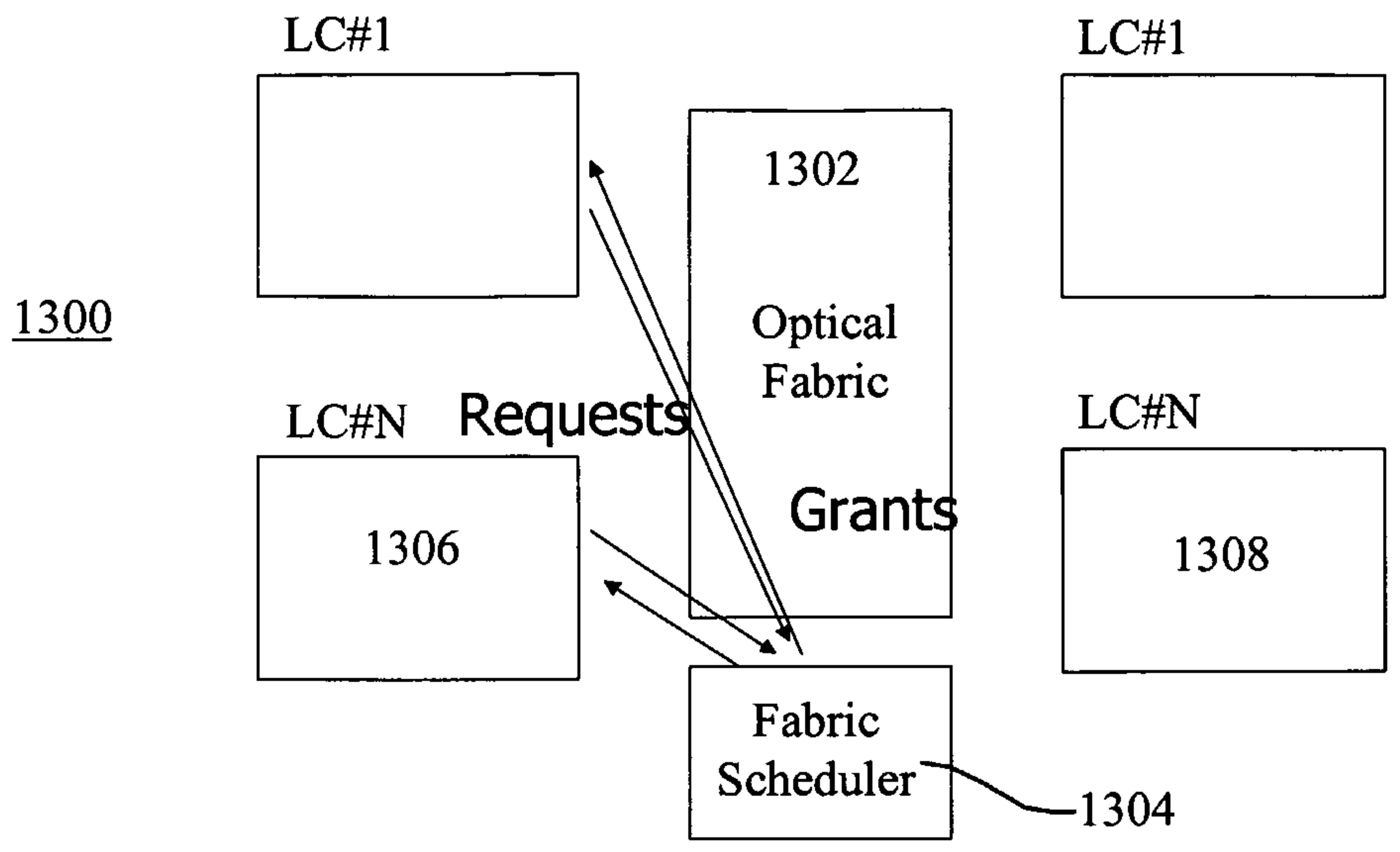


Figure 13

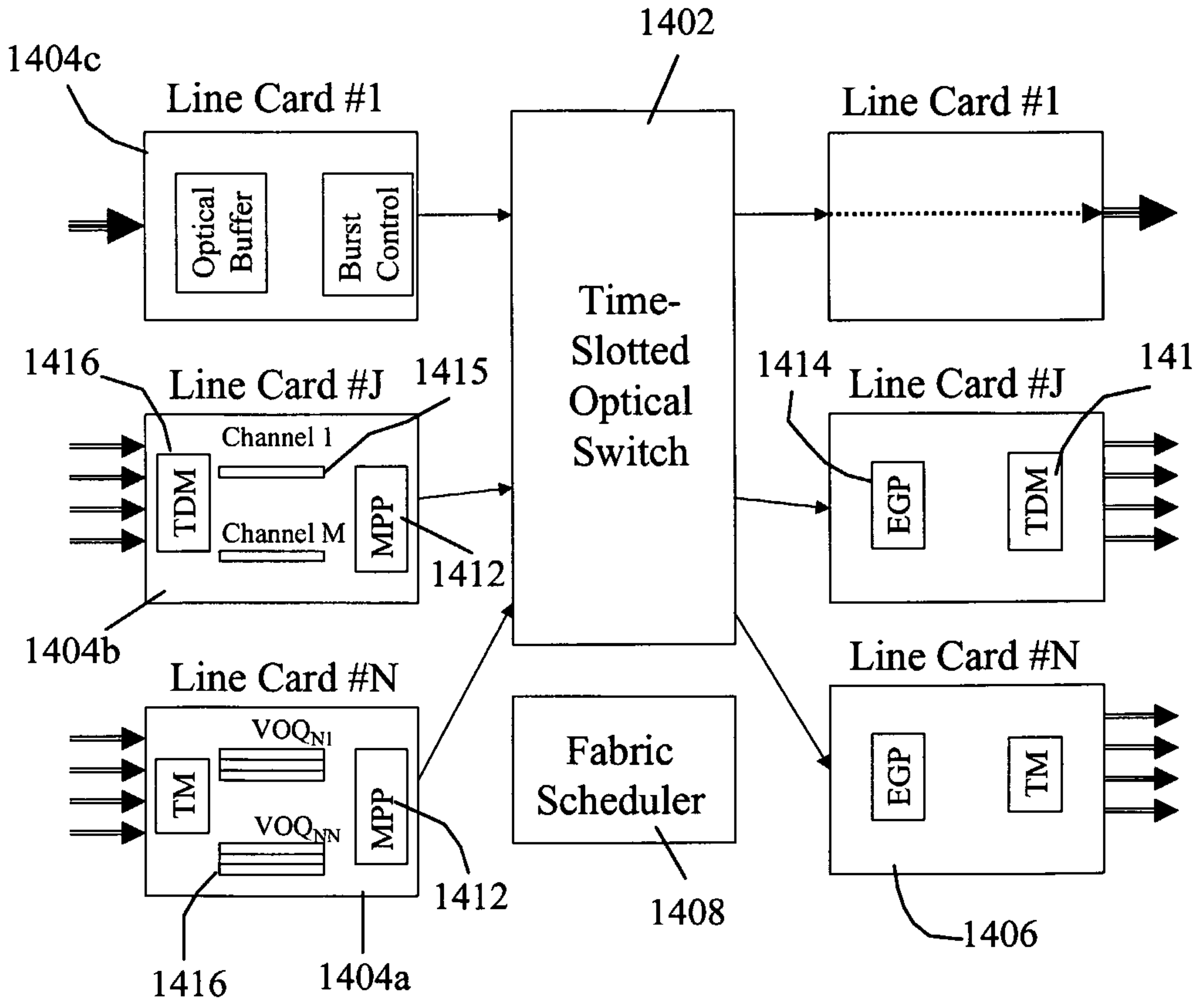


Figure 14

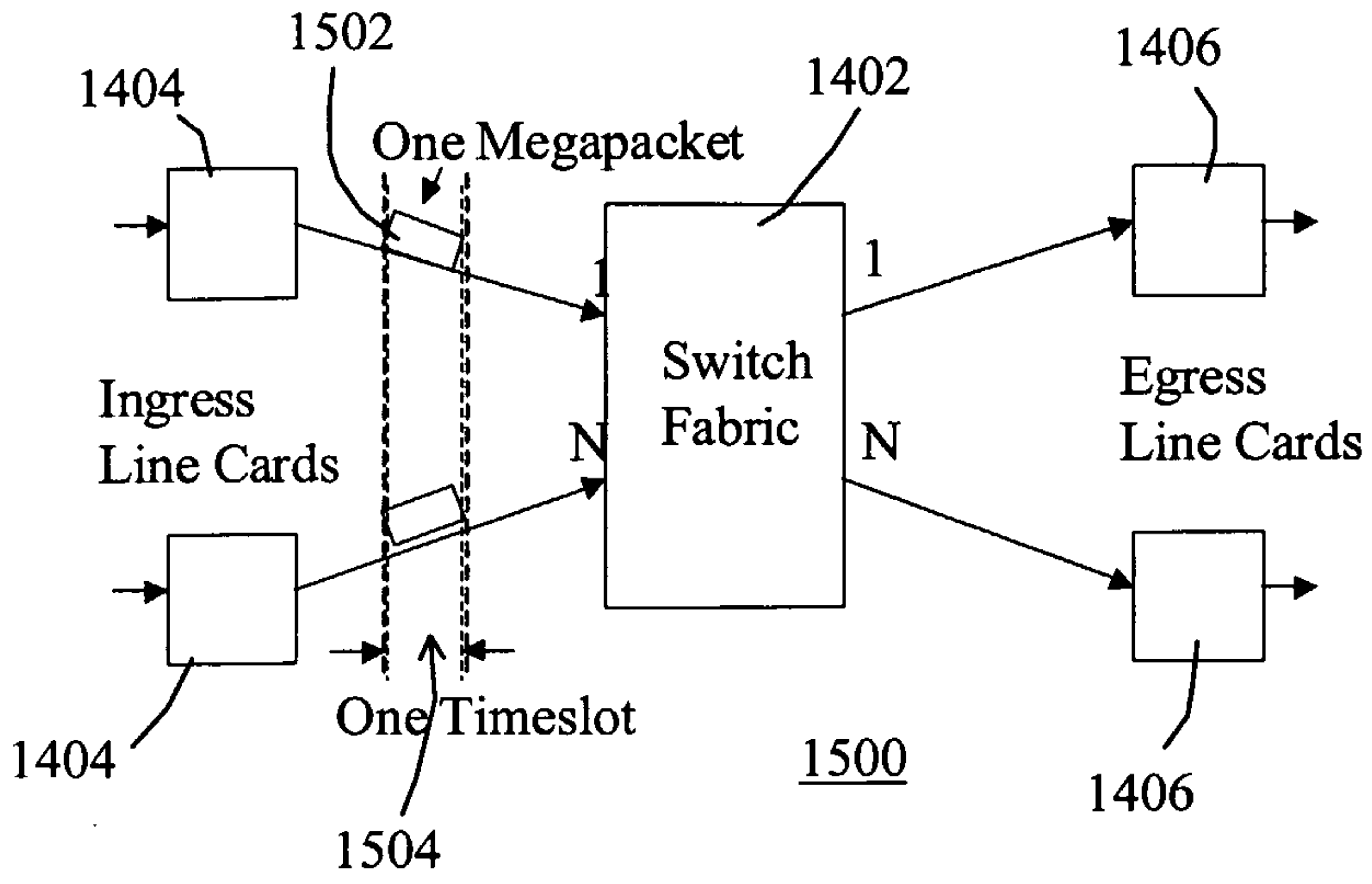


Figure 15

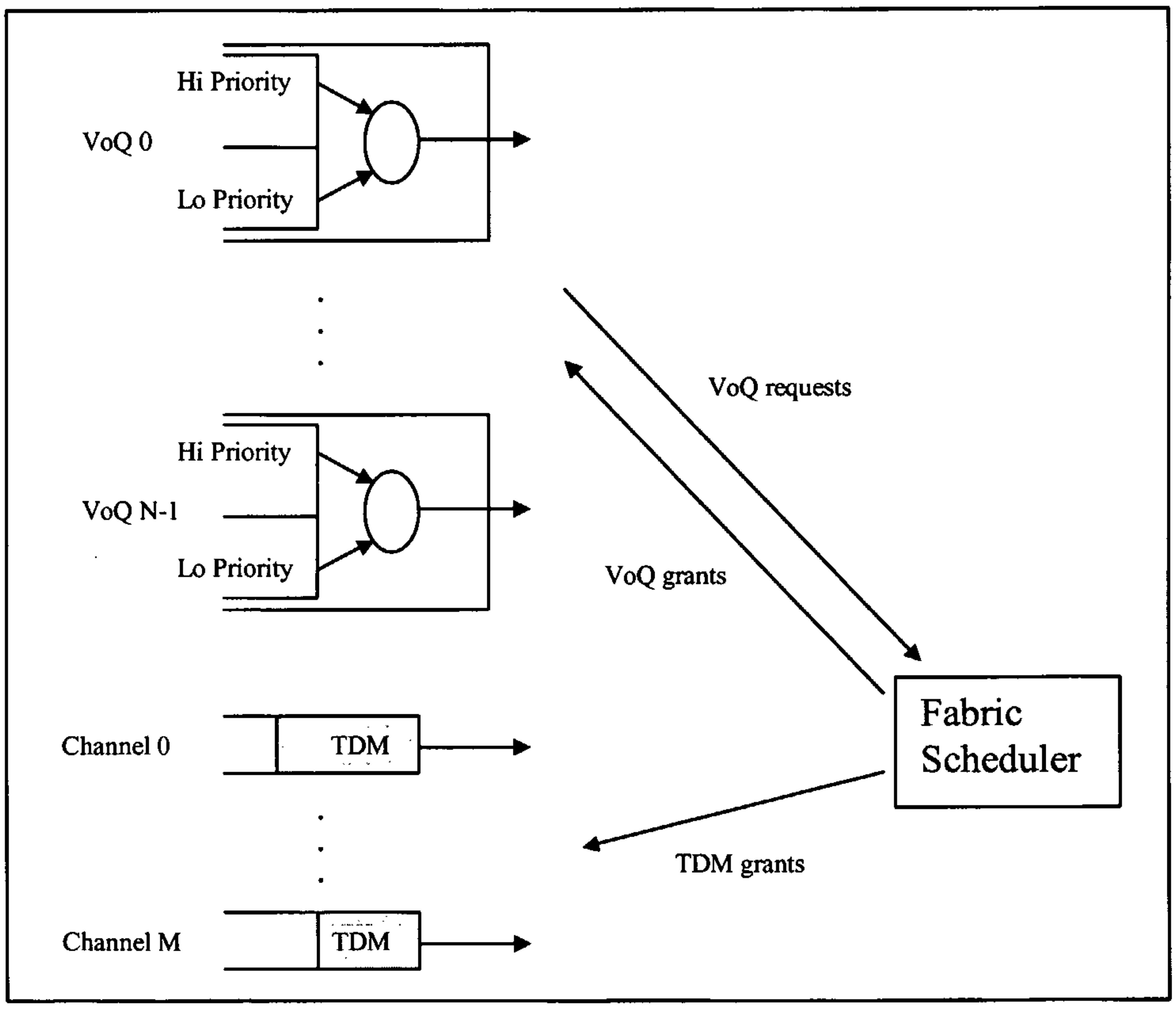


Figure 16

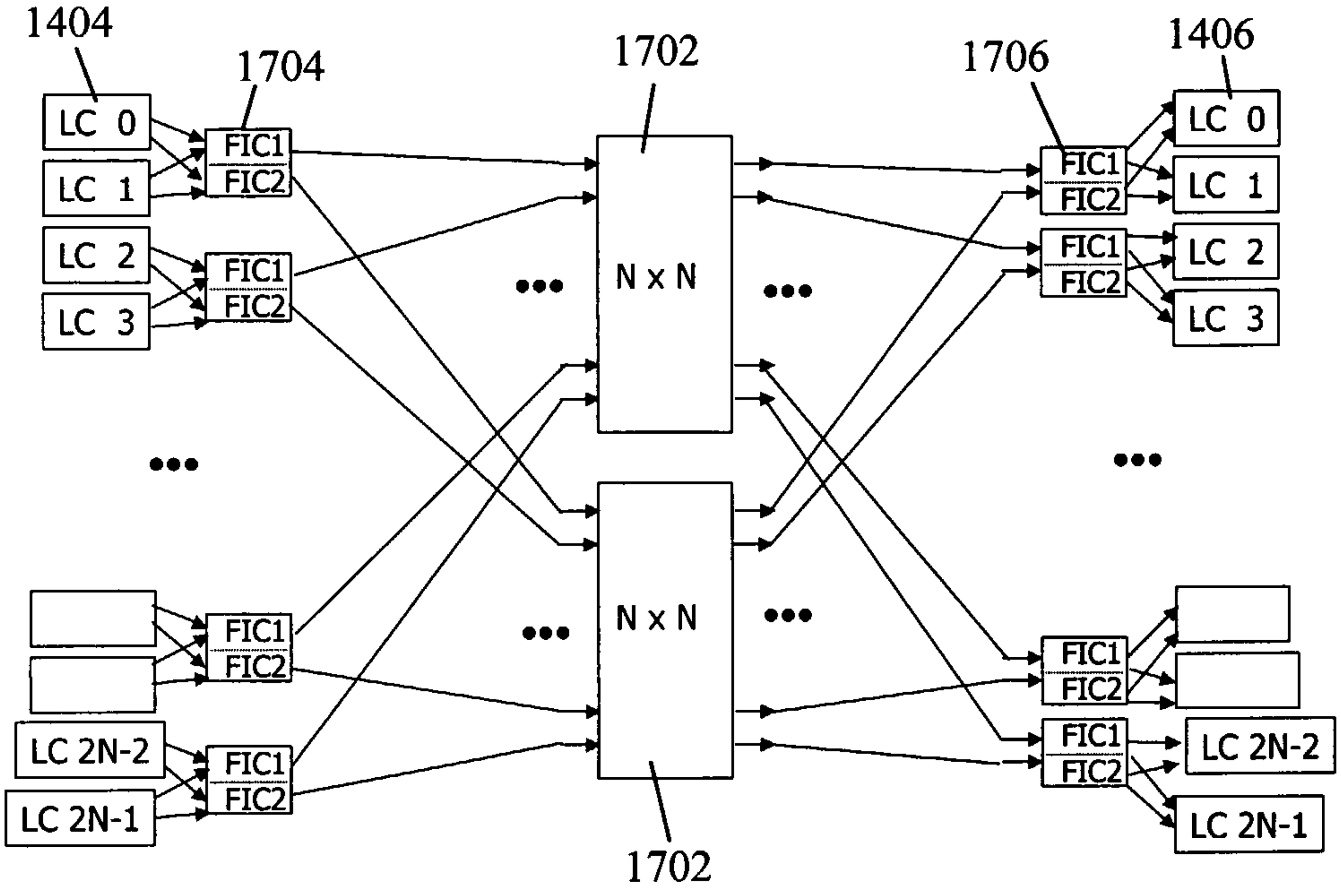


Figure 17

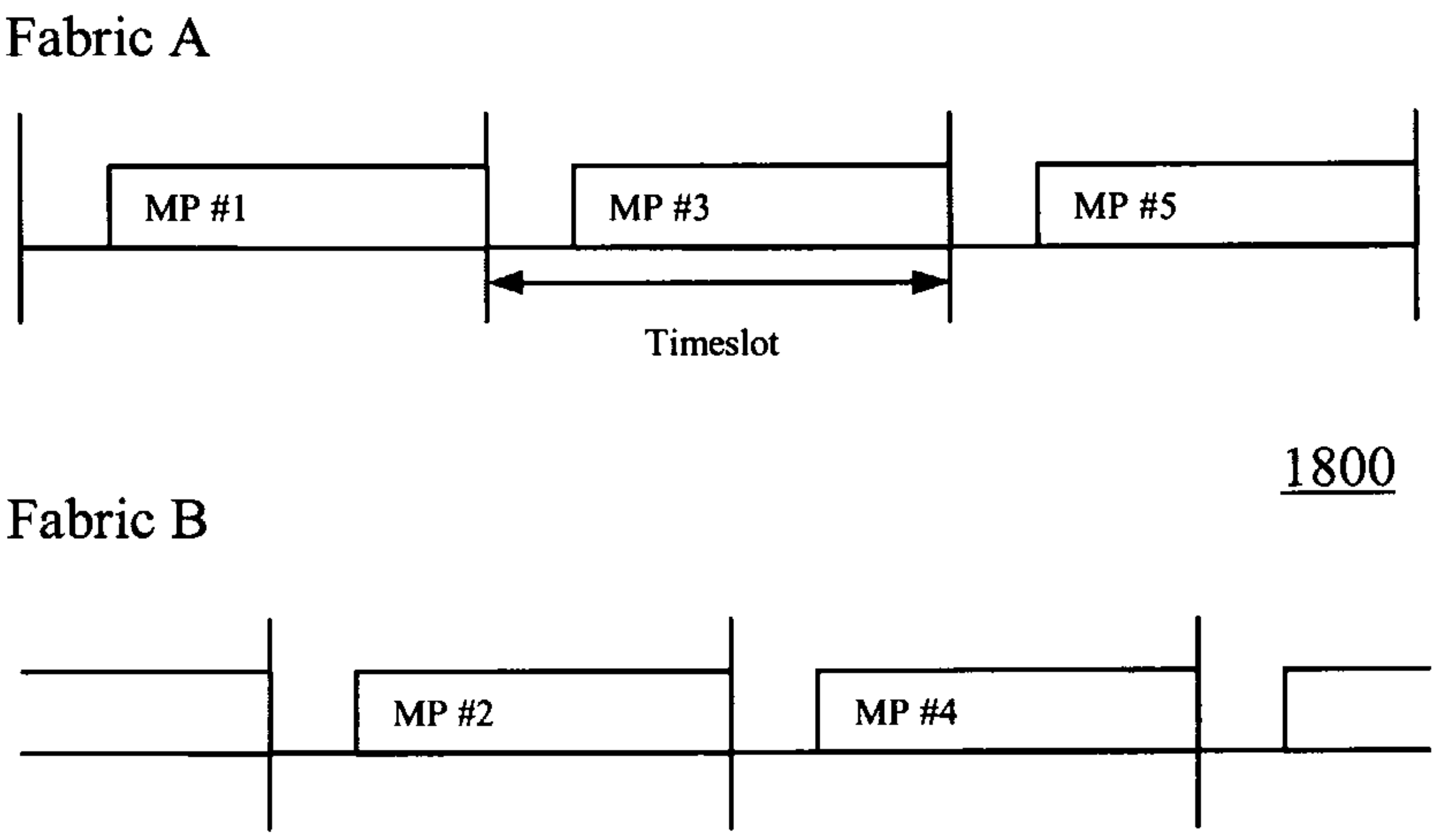


Figure 18

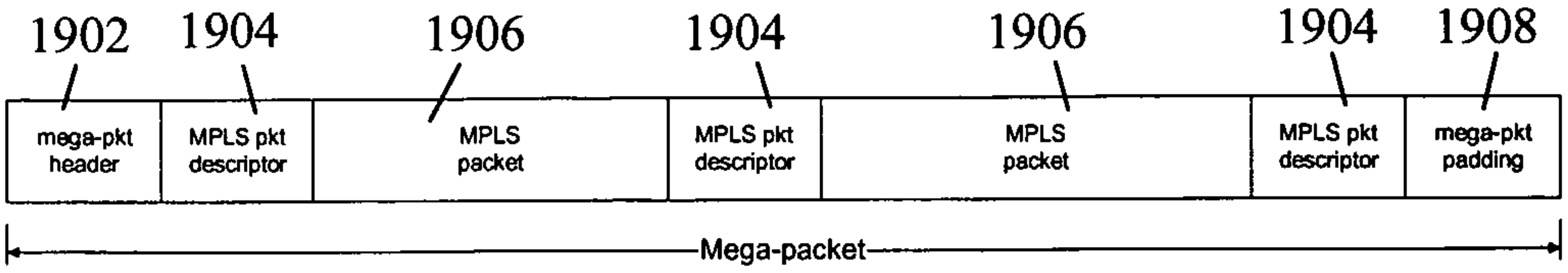


Figure 19

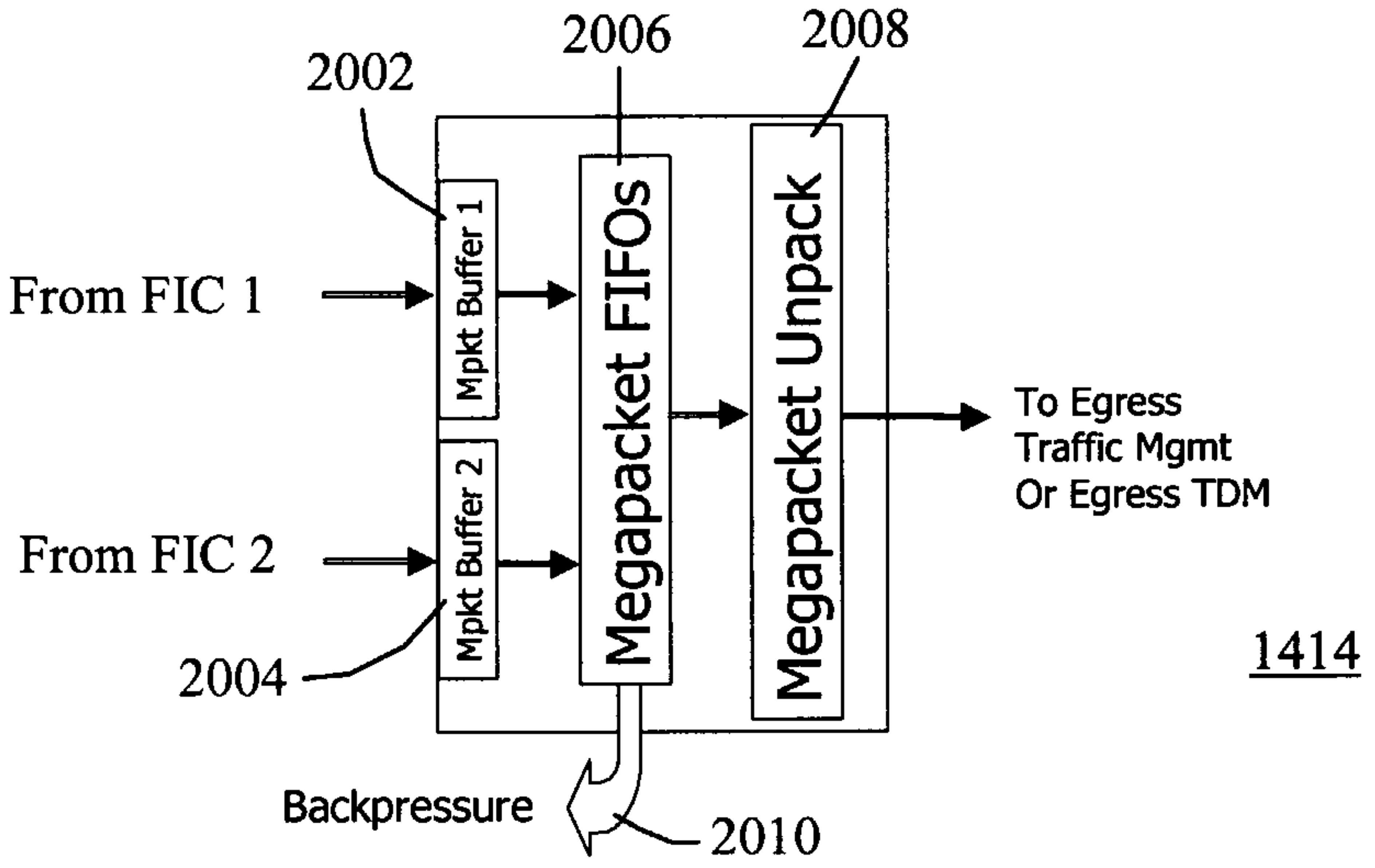


Figure 20

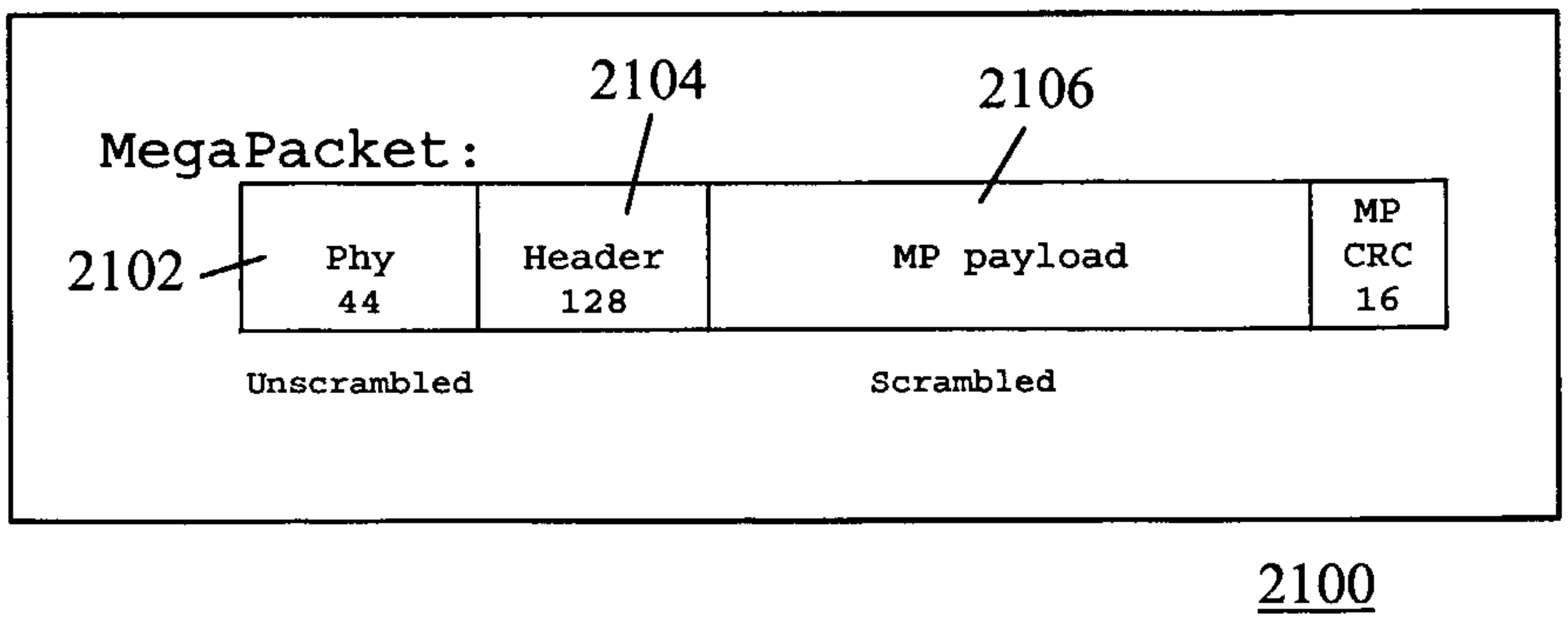


Figure 21

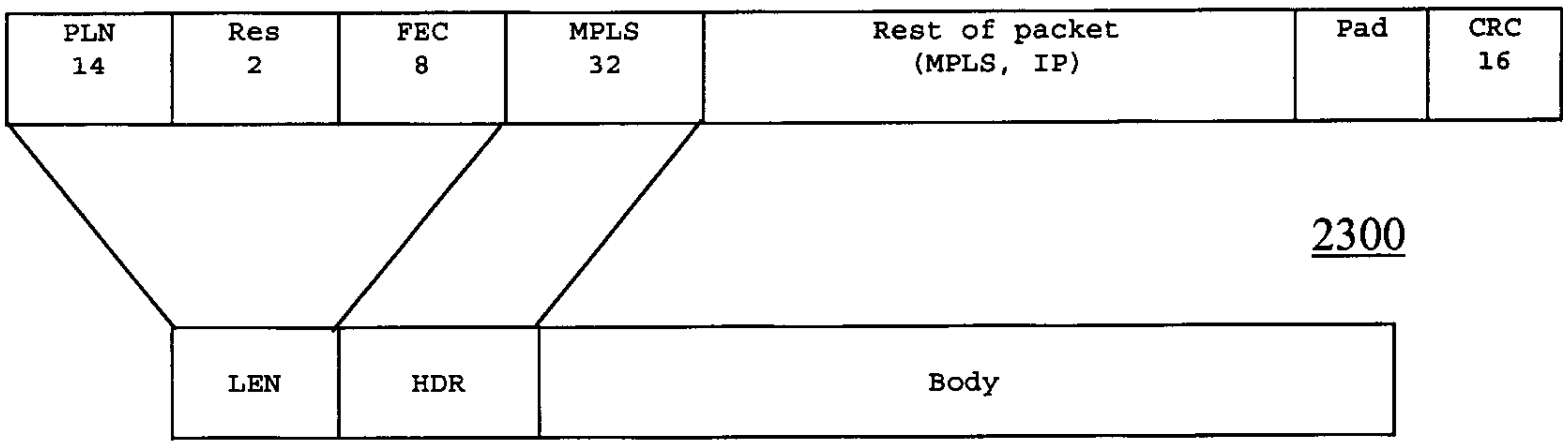
Res	Source	Res	Dest	Res	MP Type	STS Channel	Pointer/ Length	Res
2	9	2	9	1	3	10	12	16

Header – 1st 64 bit Word

Sequence	SOF Flag	SOF Position	Res	CRC	FEC	Res
14	1	13	19	8	8	1

Header – 2nd 64 bit Word

Figure 22



(Simplified diagram used in figures below)

Figure 23

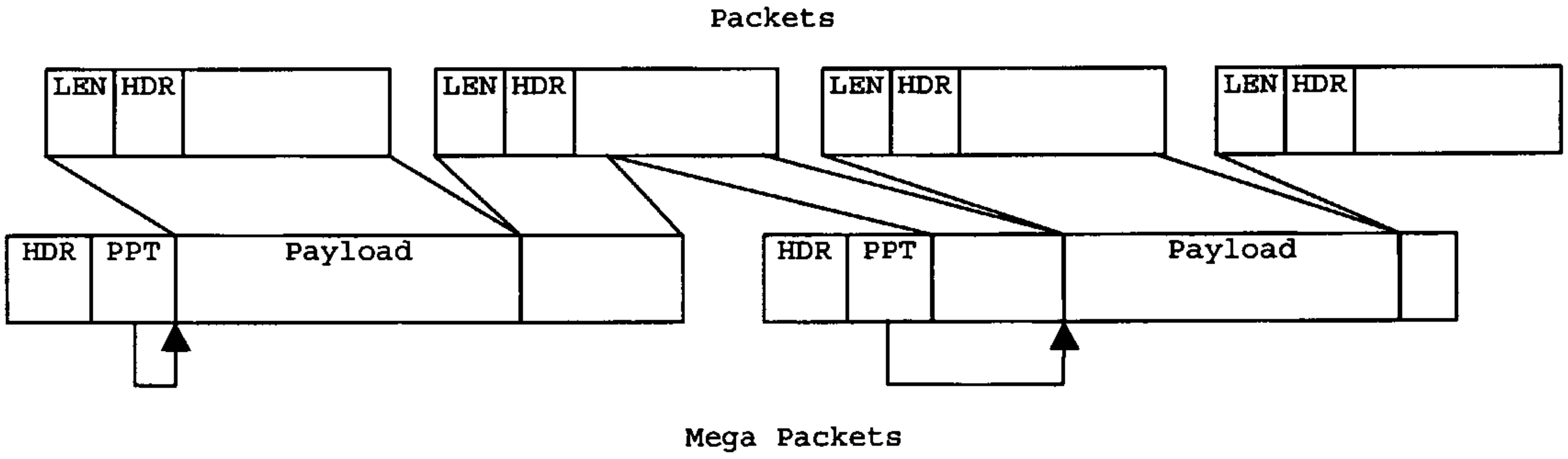


Figure 24 2400

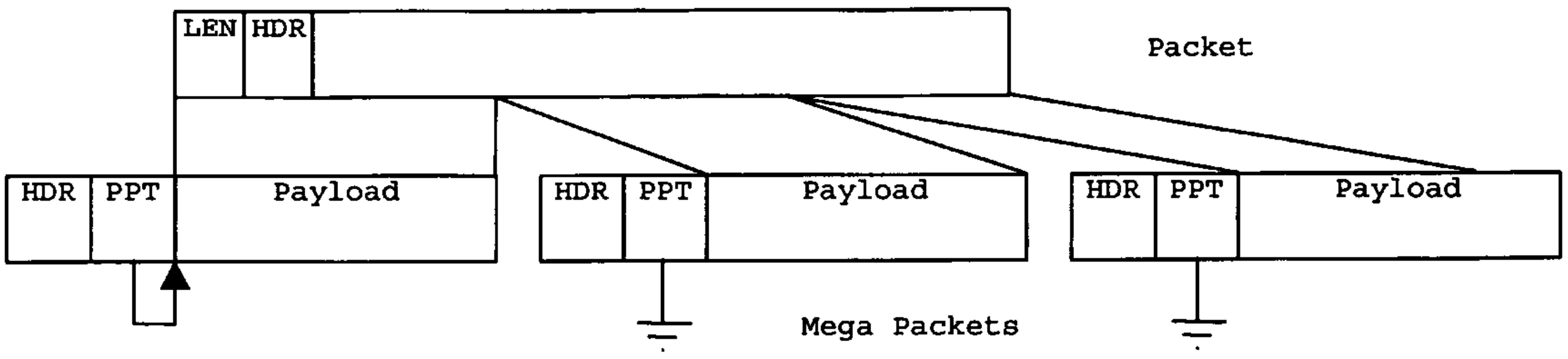


Figure 25 2500

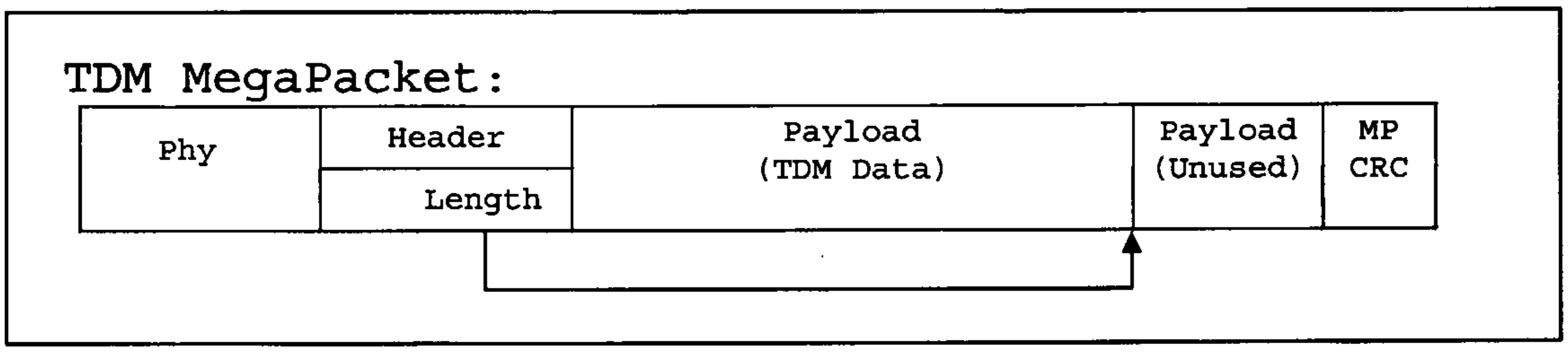


Figure 26 2600

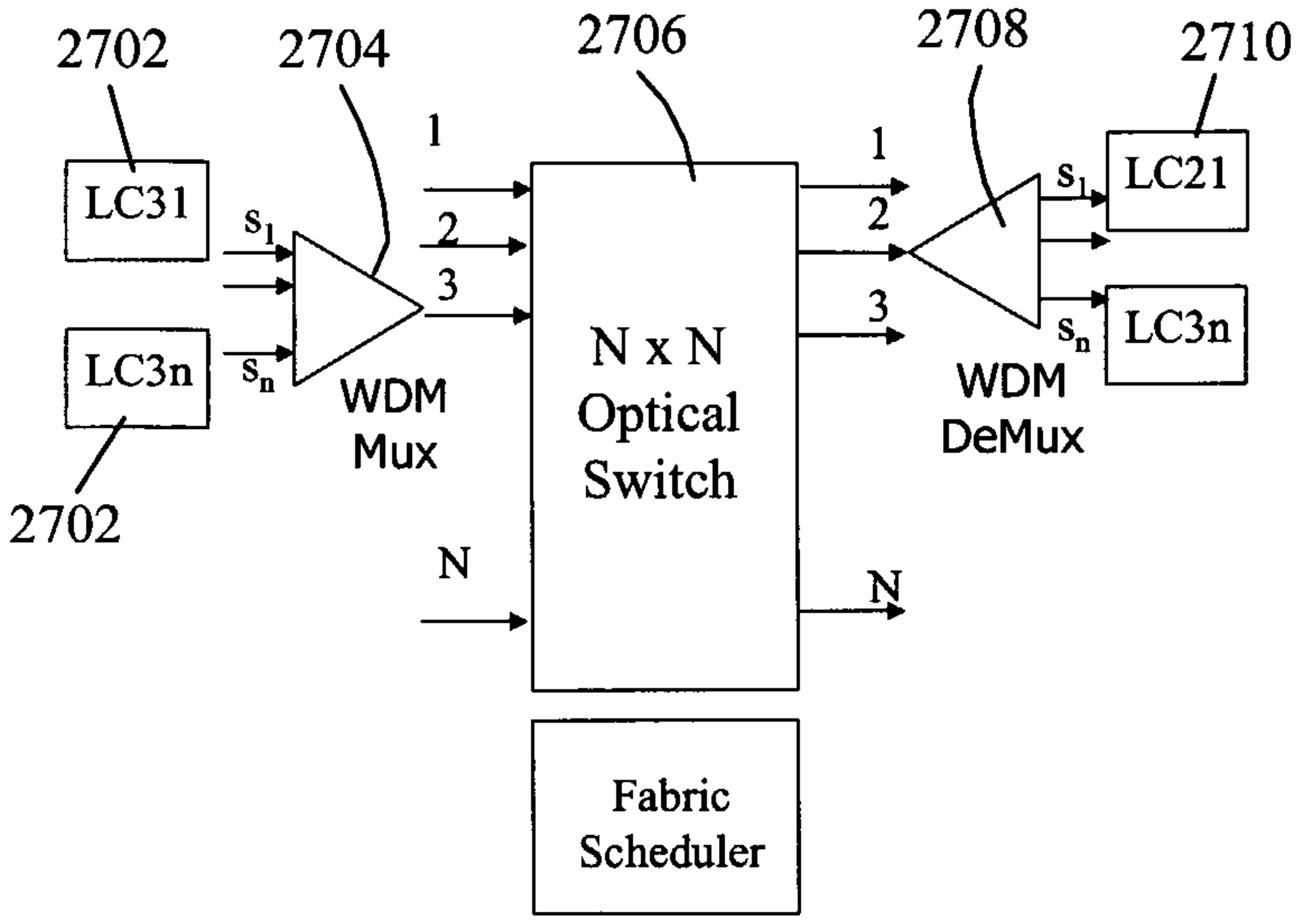


Figure 27 2700

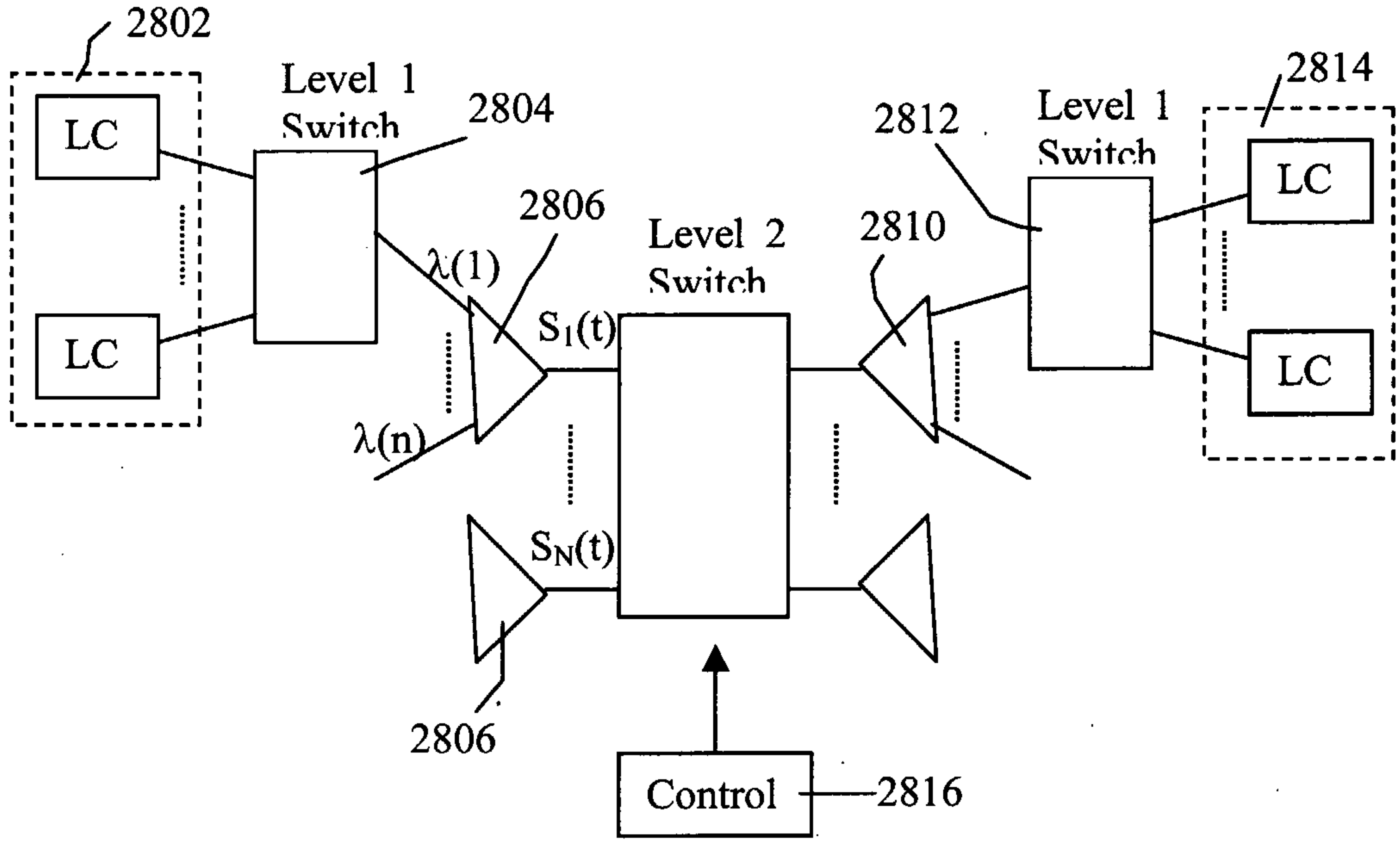


Figure 28 2800

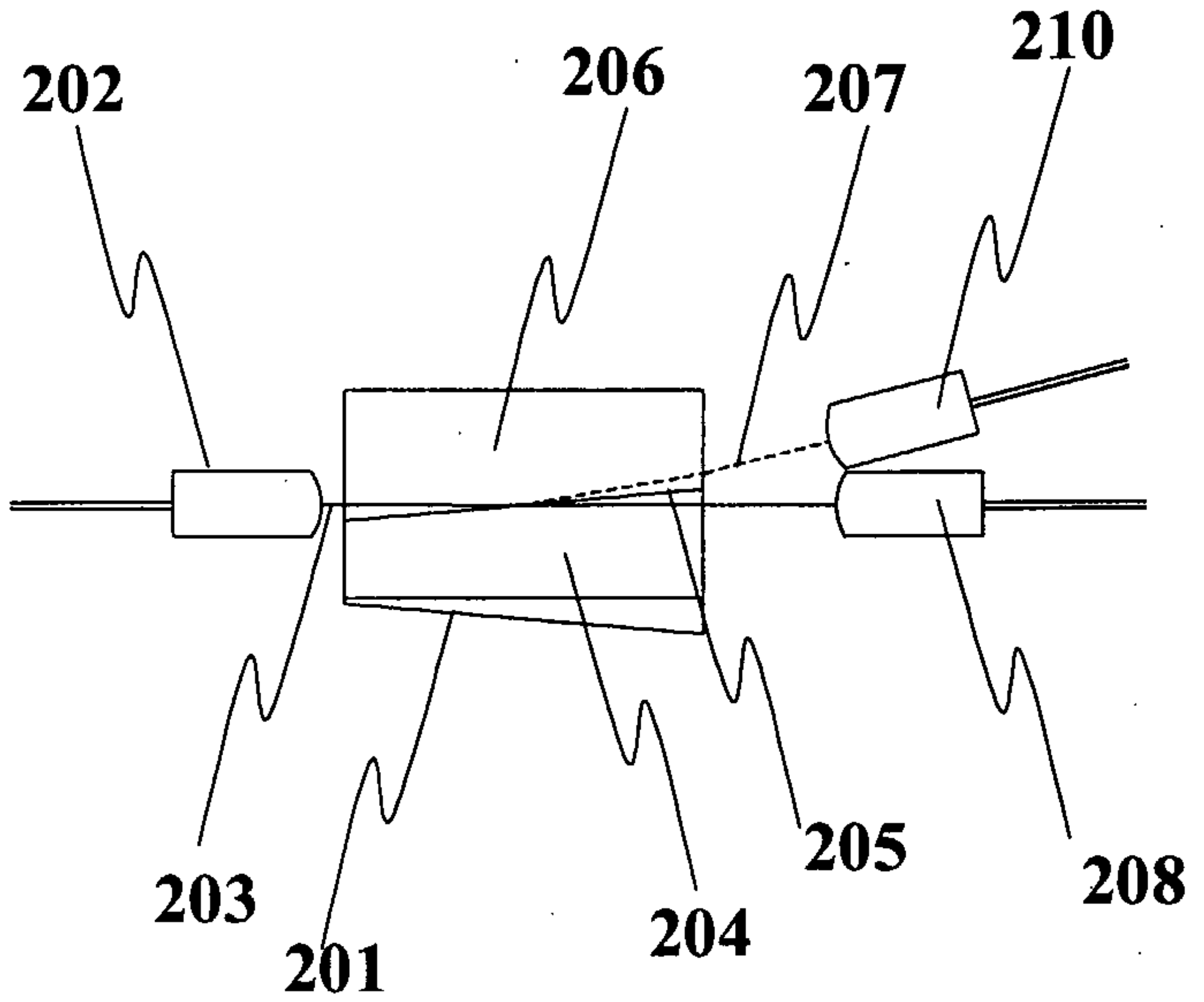


Figure 29 2900

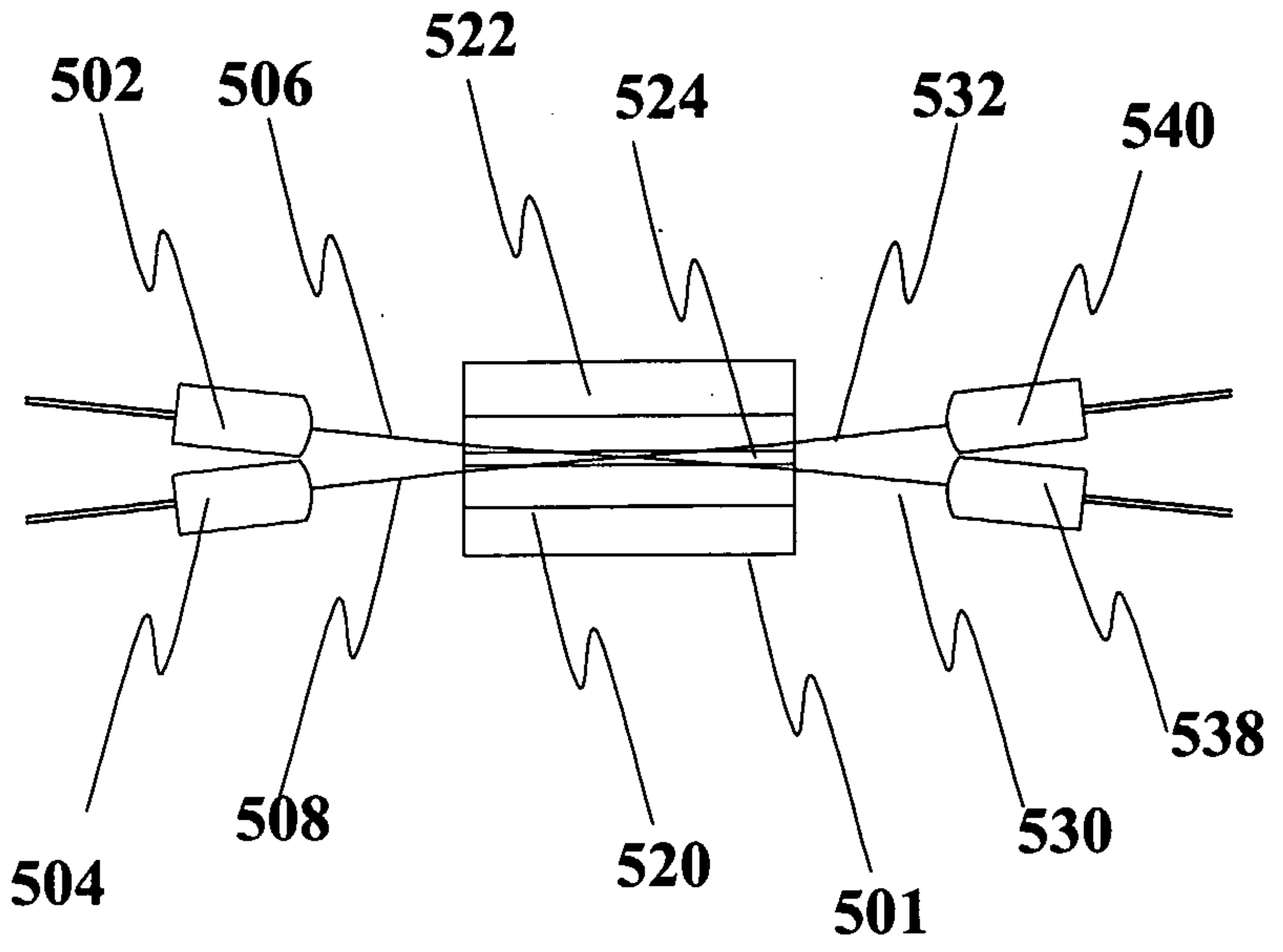


Figure 30

3000

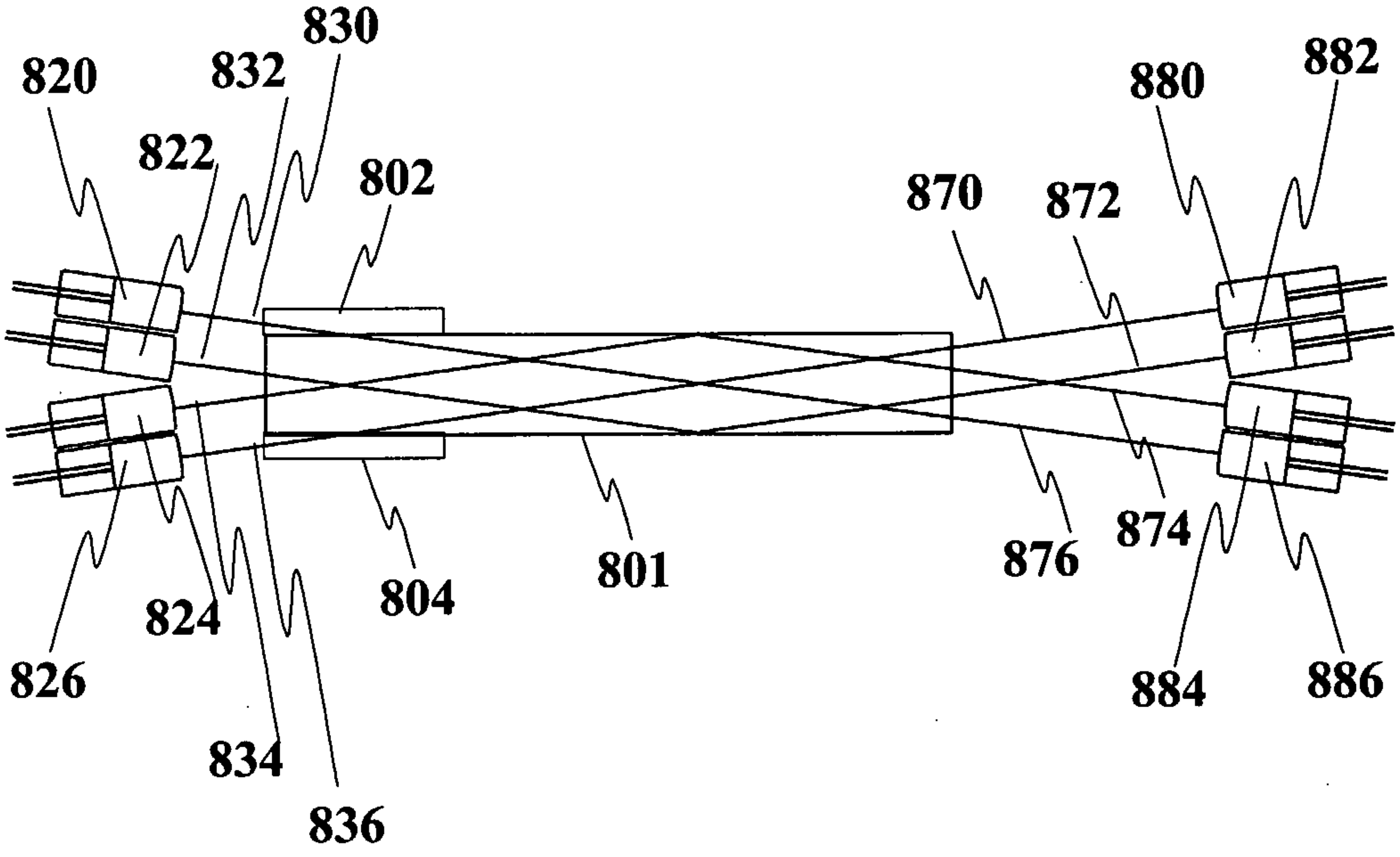
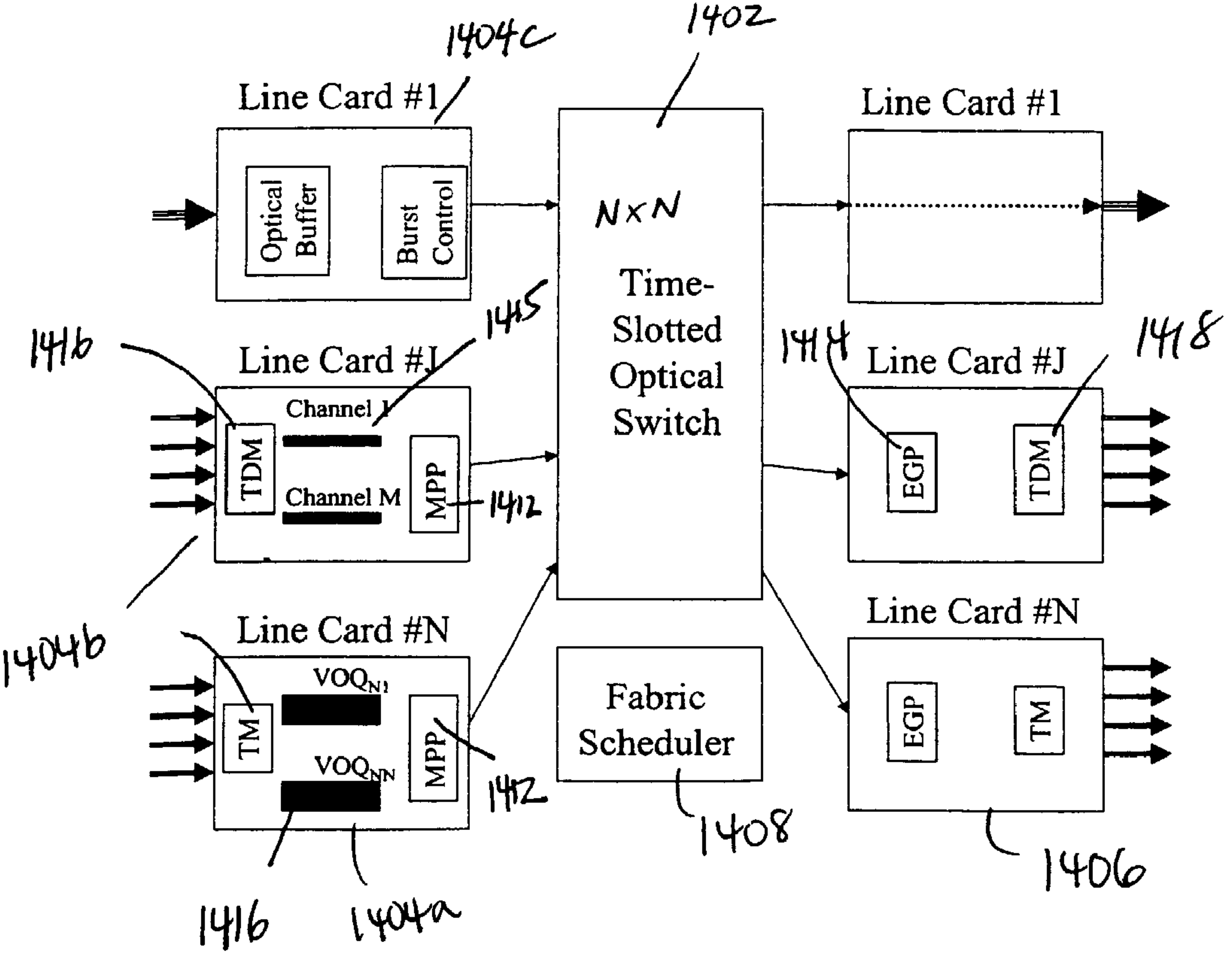


Figure 31

3100


1404c
1402
Line Card #1
Optical Buffer
Burst Control
N×N
Time-Slotted Optical Switch
Line Card #1
1416
1415
Line Card #J
Channel 1
TDM
Channel M
MPP
1412
1414
Line Card #J
1418
EGP
TDM
1404b
Line Card #N
TM
VOQ$_{N1}$
VOQ$_{NN}$
MPP
1412
Fabric Scheduler
1408
Line Card #N
EGP
TM
1406
1416
1404a