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

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2480461 C 2010/07/20

(11)(21) 2 480 461

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2003/03/25
(87) Date publication PCT/PCT Publication Date: 2003/10/16
(45) Date de délivrance/Issue Date: 2010/07/20
(85) Entrée phase nationale/National Entry: 2004/09/24
(86) N° demande PCT/PCT Application No.: US 2003/009328
(87) N° publication PCT/PCT Publication No.: 2003/085910
(30) Priorité/Priority: 2002/04/01 (US10/114,568)

(51) Cl.Int./Int.Cl. *H04L 12/56* (2006.01),
H04L 12/46 (2006.01)

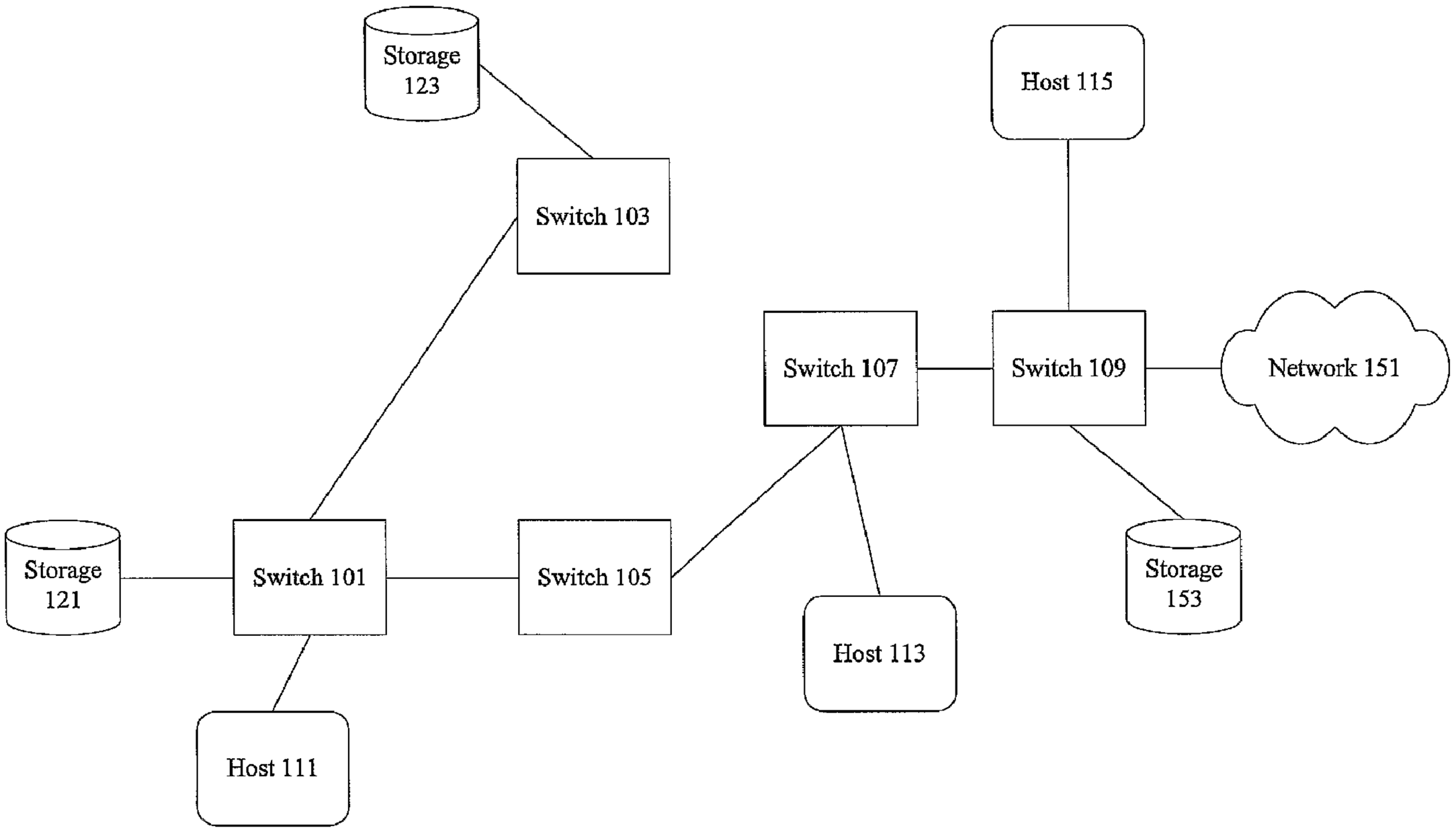
(72) Inventeurs/Inventors:
COMETTO, MAURILIO, US;
LEE, SCOTT S., US

(73) Propriétaire/Owner:
CISCO TECHNOLOGY, INC., US

(74) Agent: GOWLING LAFLEUR HENDERSON LLP

(54) Titre : METHODES ET APPAREILS PERMETTANT LA DISTRIBUTION DE TRAMES DE CANAUX A FIBRE OPTIQUE

(54) Title: METHODS AND APPARATUS FOR FIBRE CHANNEL FRAME DELIVERY



(57) Abrégé/Abstract:
Methods and apparatus are provided for improving fibre channel frame delivery. Techniques are provided for the in order delivery of frames by intelligently delaying or dropping selected fibre channel frames. Other techniques are provided for in order delivery by using label switching and frame labels. The various techniques can be applied during circumstances such as a link state or channel change.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 October 2003 (16.10.2003)

PCT

(10) International Publication Number
WO 2003/085910 A3

(51) International Patent Classification[7]: H04L 12/56

(21) International Application Number:
PCT/US2003/009328

(22) International Filing Date: 25 March 2003 (25.03.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/114,568 1 April 2002 (01.04.2002) US

(71) Applicant *(for all designated States except US)*: **ANDIAMO SYSTEMS, INC.** [US/US]; 375 East Tasman Drive, San Jose, CA 95134 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **COMETTO, Maurilio** [IT/US]; 3456 Outlook Court, San Jose, CA 95132 (US). **LEE, Scott, S.** [US/US]; 409 Calado Avenue, Campbell, CA 95008 (US).

(74) Agent: **WEAVER, Jeffrey, K.**; Beyer Weaver & Thomas, LLP, P.O. Box 778, Berkeley, CA 94704-0778 (US).

(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, EC, 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, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

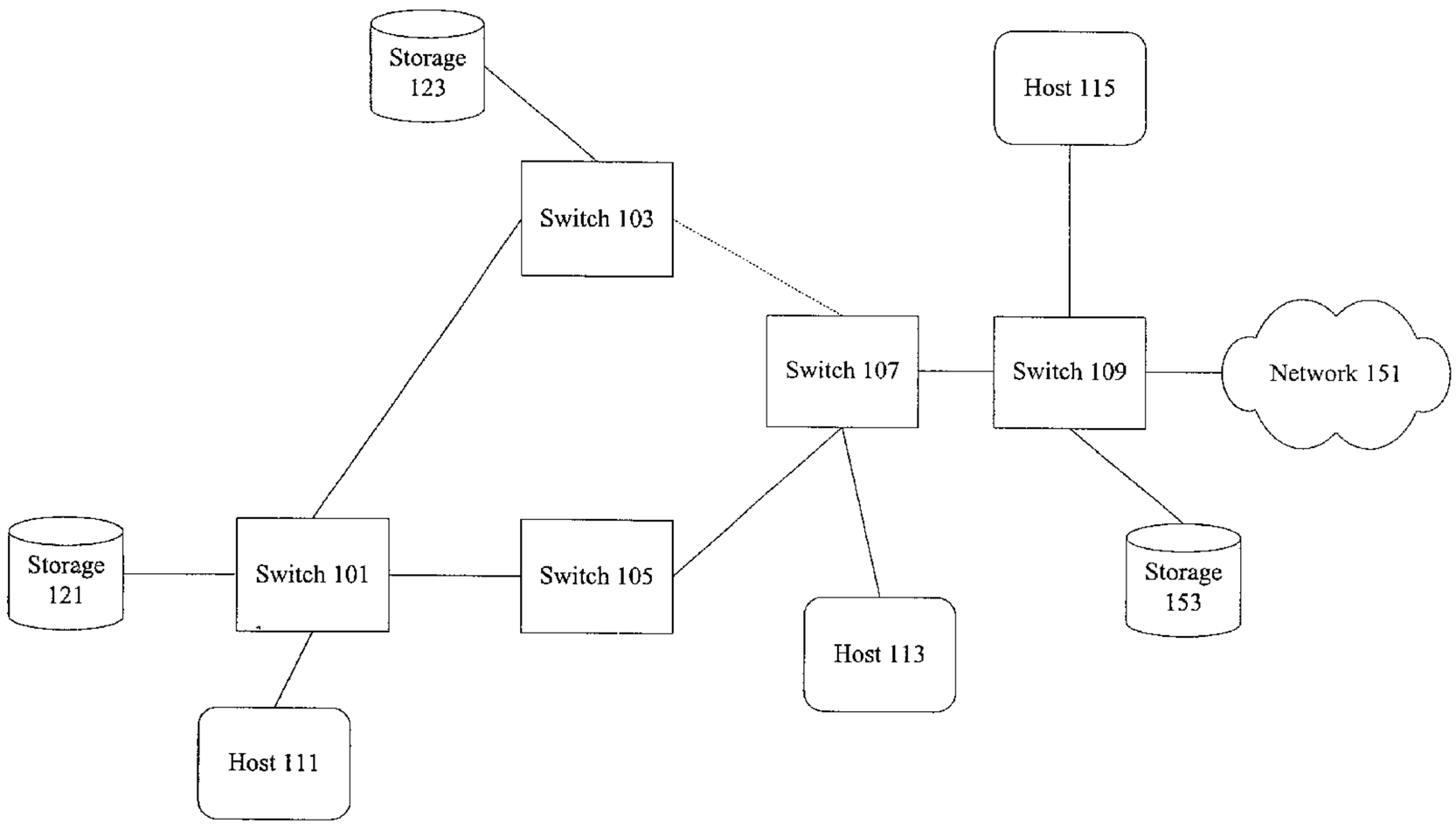
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
5 February 2004

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.

(54) Title: METHODS AND APPARATUS FOR IN-ORDER DELIVERY OF FIBRE CHANNEL FRAMES



(57) Abstract: Methods and apparatus are provided for improving fibre channel frame delivery. Techniques are provided for the in order delivery of frames by intelligently delaying or dropping selected fibre channel frames. Other techniques are provided for in order delivery by using label switching and frame labels. The various techniques can be applied during circumstances such as a link state or channel change.

WO 2003/085910 A3

`METHODS AND APPARATUS FOR FIBRE CHANNEL FRAME DELIVERY

Background of the Invention

1. Field of the Invention.

The present invention relates to fibre channel networks. More specifically, the present invention relates to methods and apparatus for providing in order delivery of fibre channel frames in a fibre channel network during circumstances such as changes in link state or changes in channel state.

2. Description of Related Art

Many conventional network protocols allow the out of order delivery of a packet sequence. A network node in a TCP/IP based network can receive an out of order set of packets and reorder the packets upon receipt. Packets often arrive out of order if they travel along different paths to reach a destination.

However, some fibre channel devices such as disks, disk arrays, and other storage mechanisms can not handle out of order frames. Link and channel state changes are some of the circumstances that may cause out of order delivery of frames in a fibre channel fabric. Multiple links seen as a single link between two fibre channel entities is referred to herein as a channel. Some mechanisms in existing networks call for the flushing of all frames in the network upon a change in link state. Flushing all of the frames can prevent out of order delivery when paths

and routes change in a network. All frames are flushed even if the paths for the frames to the associated destinations are not changed. However, flushing all of the frames either explicitly or implicitly can be very disruptive to network operation, as more frames are dropped than is necessary and network operation is at least temporarily halted.

It is therefore desirable to provide methods and apparatus for improving fibre channel frame delivery and providing in order delivery particularly during link state and channel changes.

Summary of the Invention

Methods and apparatus are provided for improving fibre channel frame delivery. Techniques are provided for the in order delivery of frames by intelligently delaying or dropping selected fibre channel frames. Other techniques are provided for in order delivery by using label switching and frame labels. The various techniques can be applied during circumstances such as a link or channel state change.

According to various embodiments, a method for selectively delivering fibre channel frames in a fibre channel fabric is provided. A set of next hops at a fibre channel entity is identified. The set of next hops is used to forward frames based on a destination identifier. A fibre channel fabric link change is detected. An updated set of next hops is identified. The updated set of next hops is used to forward a frame received at a fibre channel entity based on a destination identifier while accounting for the fibre channel fabric link change. The set of next hops is compared to the updated set of next hops. The forwarding of frames towards the updated set of next hops is prevented for a predetermined period of time if it is determined that the set of next hops is different from the updated set of next hops.

According to another embodiment, a method for selectively delivering frames in a fibre channel fabric is provided. A channel for forwarding frames from a first fibre channel entity to a second fibre channel entity is identified. The channel comprises a plurality of links connecting the first fibre channel entity to the second

fibre channel entity. A change in the channel is detected at the first fibre channel entity. An updated channel for forwarding frames from the first fibre channel entity to the second fibre channel entity is identified. The updated channel is different from the channel. Frames received for forwarding on the updated channel that have not yet been placed into output queues associated with the updated channel are blocked.

In yet another embodiment, a method for selectively delivering frames in a fibre channel fabric is provided. A fibre channel fabric link change is detected at a fibre channel switch. An updated routing table associated with a topology version number is generated. Generating an updated routing table includes determining a next hop, an incoming label, and a destination corresponding to each entry in the updated routing table. A frame is received at the fibre channel switch. The frame includes a first destination and a first label corresponding to a first entry in the updated routing table. It is determined whether the fibre channel switch has received a first outgoing label corresponding to the first entry in the updated routing table, the received first outgoing label having the same topology version number as the updated routing table. The frame is dropped if it is determined that the fibre channel switch has not received the first outgoing label.

These and other features and advantages of the present invention will be presented in more detail in the following specification of the invention and the accompanying figures, which illustrate by way of example the principles of the invention.

Brief Description of the Drawings

The invention may best be understood by reference to the following description taken in conjunction with the accompanying drawings, which are illustrative of specific embodiments of the present invention.

Figure 1 is a diagrammatic representation of a network that can use the techniques of the present invention.

Figure 2 is a diagrammatic representation of a fibre channel fabric undergoing a fibre channel fabric link change.

Figure 3A is a diagrammatic representation of a routing table showing a set of next hops

Figure 3B is a diagrammatic representation of a routing table showing an updated set of next hops.

Figure 4 is a diagrammatic representation showing virtual output queues.

Figure 5 is a process flow diagram showing the blocking of fibre channel frames.

Figure 6 is a process flow diagram showing the dropping of fibre channel frames.

Figure 7 is a diagrammatic representation of possible reordering at a channel.

Figure 8A is a diagrammatic representation of a forwarding channel table.

Figure 8B is a diagrammatic representation of an updated forwarding channel table.

Figure 9 is a process flow diagram showing the forwarding of fibre channel frames upon a change in the channel.

Figure 10 is a diagrammatic representation of label switching routers.

Figure 11 is a diagrammatic representation of label switching routers during a link change.

Figure 12 is a diagrammatic representation of label switching routers while out labels are partially resolved.

Figure 13 is a process flow diagram showing techniques for in order delivery using input and outgoing labels.

Detailed Description of Specific Embodiments

Reference will now be made in detail to some specific embodiments of the invention including the best modes contemplated by the inventors for carrying out the invention. Examples of these specific embodiments are illustrated in the accompanying drawings. While the invention is described in conjunction with these specific embodiments, it will be understood that it is not intended to limit the invention to the described embodiments. On the contrary, it is intended to cover

alternatives, modifications, and equivalents as may be included within the spirit and scope of the invention as defined by the appended claims.

Methods and apparatus of the present invention provide for in order delivery of fibre channel frames. According to various embodiments, several network conditions can lead to the out of order delivery of frames to fibre channel device. The techniques of the present invention provide for delaying, blocking, dropping, and/or labeling certain fibre channel frames to deliver frames to a fibre channel device in order. In one embodiment, frames that will be traversing a new path are blocked to allow frames travelling along the old path to either reach the destination first or be dropped from the network.

Figure 1 is a diagrammatic representation of one example of a network that can use the techniques of the present invention. Figure 1 shows a storage area network implemented using fibre channel. A switch 101 is coupled to switches 103 and 105 as well as to a host 111 and storage 121. In one embodiment, host 111 is a server or client system while storage 121 is any storage subsystem such as a single disk or a redundant array of independent disks (RAID). Switch 105 is coupled to switch 107. Switch 107 is connected to host 113 and switch 103 is connected to storage 123. Switch 109 is connected to host 115, switch 107, host 153, and an external network 151 that may or may not use fibre channel. In order for a host 111 to access network 151, a path going through switch 105 can be used. It should be noted that any apparatus including a processor, memory, and a connection to a fibre channel fabric can be a fibre channel switch.

Ports used to connect switches to each other in a fibre channel network are referred to herein as non F-ports while ports used to connect a switch to a host a referred to herein as F-ports. In one example, non F-ports are used to connect switch 105 to switch 107 while F-ports are used to connect switch 107 to host 113. Similarly, FL-ports are used to connect switch 103 to storage 123. Ports such as F-ports and FL-ports are herein referred to as edge ports. Other ports are referred to as non-edge ports.

According to various embodiments, a frame transmitted from host 111 to a network 151 or to a storage device 153 includes parameters such as the exchange identifier, a sequence, and a sequence number. The exchange identifier can provide information on what exchange the frame belongs to. The sequence can provide information on what portion of the exchange the frame belongs to while the sequence number can provide information on how the frames should be ordered. Sequence numbers can be used to allow for in order delivery of fibre channel frames.

Some fibre channel devices such as certain storage disks and disk arrays require that frames be received in the order in which they were transmitted. Conventional networks such as TCP/IP networks do not have such a requirement, as TCP/IP networks generally have mechanisms for reordering packets upon receipt. If frames with sequence numbers of 191, 192, and 193 are transmitted in order in a fibre channel network, a fibre channel device receiving the frames may expect that the frames are in the same order in which they were transmitted. A fibre channel device may not be able to handle receiving the frames out of order.

In a static fibre channel network, frames will typically be received in the order in which they were transmitted. However, several occurrences can lead to the out of order delivery of fibre channel frames. Links state changes in particular can lead to out of order delivery.

Figure 2 is a diagrammatic representation of a fibre channel fabric undergoing a fibre channel fabric link change. Figure 2 shows one example of a link change that can lead to out of order delivery of fibre channel frames. A new link using non edge-ports is introduced between switch 103 and switch 107. With the introduction of a new link between switch 103 and switch 107, a new version of the routing tables can be generated. A variety of routing table generation algorithms such as Fibre Channel Shortest Path First (FSPF) can be used. Traffic from host 111 travelling through switches 101, 105, and 107 to switch 109 now can travel through switches 101, 103, and 107 to switch 109. The set of next hops available to transmit frames from a switch 111 to a storage device 153 is switch 105 before the introduction of the new link.

A set of neighboring fibre channel entities available to transmit a frame from one fibre channel entity to another fibre channel entity is referred to herein as a set of next hops. After a link change, a set of next hops can be updated in a switch. In one example, the set of next hops for transmitting frames from switch 101 to a network 151 changes from simply switch 105 to both switch 103 and switch 105 after link state change. A set of neighboring entities that can be used to transmit frames from a source to a destination updated after a link change or the generation of updated routing tables is referred to herein as an updated set of next hops. It should be noted that a set of next hops can include one or more neighboring nodes. In one example, a set of next hops is a single neighboring entity. In another example, a set of next hops includes multiple neighboring entities.

Having an updated set of next hops can lead to out of order delivery of fibre channel frames. In one example, the earlier frames transmitted at switch 101 in a sequence may be travelling through switch 105 while later frames in the same sequence may be travelling through a switch 103. A variety of network conditions can cause the later frames travelling through a switch 103 to arrive at switch 109 before the earlier frames travelling through a switch 105 arrive at switch 109. In one example, the earlier frames are slowed at a switch 105 because of congestion at switch 105 while the later frames travel through switch 103 quickly because of a new high bandwidth link between switch 103 and switch 107. A storage device 153 receiving the later transmitted frames before receiving the earlier transmitted frames from host 111 may not be able to handle the out of order frames.

Figures 3A and 3B depict routing tables showing information relating to sets of next hops and updated sets of next hops at a switch 101. Figure 3A shows sets of next hops for a frame received at switch 101 and a network where a link has not yet been established between switch 103 and switch 107. When a frame is received at switch 101, an identifier depicting the destination of the frame can be used to reference an entry in the routing table. In one example, the destination of the frame is switch 107 and the entry 309 can be referenced to determine that the set of next hops is switch 105. If it is determined that the destination of the received frame is switch 101, the routing table can be used to pass the frame to a processor associated

with switch 101. A routing table can also be used to drop frames. In one example, a value such as a null value can be placed in the set of next hops and a frame having a destination associated with the null value can be dropped upon referencing a routing table. An entry directing that a frame with a particular destination be dropped is referred to herein as an adjacency drop.

In one embodiment, a routing table is provided for each virtual storage area network (VSAN) that the switch is a part of. It should be noted that a fibre channel switch can be a part of many different VSANs and a routing table can be provided for each VSAN the switch is associated with.

After a link is added connecting switch 103 to switch 107, the routing table is updated. Figure 3B is a diagrammatic representation of a routing table showing updated sets of next hops. According to various embodiments, a frame having a destination set as switch 107 can be forwarded along either switch 103 or switch 105 based on entry 329. In a stable topology, all frames in a particular flow or exchange follow the same path.

It should be noted that a routing table may allow both paths or it can select a best path. If the best path selected has a next hop of switch 105 the updated set of next hops is the same as the set of next hops before the link change as shown in Figure 3A. If the path selected is switch 103 for a frame with a destination of switch 107, the updated set of next hops is different from the set of next hops in the routing table before the link change. Determining whether an updated set of next hops is different from an original set of next hops can be useful for deciding on whether to block or drop a particular frame. In one example, frames are not blocked or dropped if the path for a particular sequence of frames remains unchanged even after a link change in the fibre channel fabric. If the path for the sequence of frames remains the same, the frames will be delivered in order to a destination. If the path for the sequence of frames has changed, there is risk that frames will be delivered out of order.

One of the mechanisms that can impact the order in which fibre channel frames are delivered are the queues within a fibre channel switch. A frame

transmitted first from a host can remain in a queue associated with a switch 105 while a frame transmitted later from a host can be delivered quickly through a switch 103. Figure 4 is a diagrammatic representation of queues that can be associated with a fibre channel switch, according to various embodiments. Although one particular type of queue will be described, it should be noted that a variety of different input and output queues associated with various input and output ports can be used to implement the techniques of the present invention.

A switch 401 is connected to external nodes 451, 453, 455, and 457. The switch 401 includes a buffer 403 of shared memory associated with each switch port. A buffer 403 is associated with external node 451. Buffers associated with external nodes 453, 455, and 457 are not shown for purposes of clarity. The buffer 403 can hold traffic destined for external nodes 453, 455, 457, and loop back traffic to external node 451.

In typical implementations, frames destined for the various external nodes are all placed in the same buffer 403. Consequently, when a switch 401 receives a large volume of frames destined for a particular node such as external node 453, frames associated with external node 453 can use the entire buffer 403. According to various embodiments, the frames stored in buffer 403 are referenced by pointers in frame descriptor queues 411-447. Each frame descriptor can contain a pointer or reference identifying where the frame is stored in the buffer 403. Pointers or references to a shared buffer are herein referred to as descriptors. Descriptors can also identify other information such as frame priority.

In one example, an arbitrator 405 selects frames using a round-robin methodology. In a first round, a frame destined for external node 453 is selected. In a second round, a frame destined for external node 455 is selected, etc. More particularly, the arbitrator 405 may first select a high priority frame associated with descriptor 411 destined for external node 453, then select a high priority frame associated with descriptor 421 destined for external node 455, then select a high priority frame associated with descriptor 431 destined for external node 457, etc. It should be noted that a variety of techniques for selecting a frame can be used, as will be appreciated by one of skill in the art.

A queuing system having buffers apportioned based on destination is referred to herein as virtual output queuing (VOQ). VOQ is described further in Tamir Y., Frazier G.: "High Performance multi-queue buffers for VLSI communications switches", Proc. Of 15th Ann. Symp. On Comp. Arch., pp.343-354, June 1988, the entirety of which is incorporated by reference for all purposes. An abstraction identifying traffic with particular characteristics between two nodes is herein referred to as a flow. In one example, a flow is referenced by a source identifier, a destination identifier, a priority, a class, and an exchange identifier. Other characteristics are also possible. It should be noted, however, that a flow may also be referenced merely by a source and destination identifier.

According to various embodiments, frames in a particular flow may be blocked because a buffer 403 is full. If an alternative route is provided for frames in the same sequence, later frames may be able to traverse a fibre channel fabric more quickly than a frame remaining in a congested switch. In one implementation, later frames can be blocked along an alternative route to allow earlier frames to reach a destination first. Frames transmitted first are herein referred to as earlier frames while frames transmitted later by a source are referred to herein as later frames. Later frames can be blocked using several mechanisms. In one embodiment, the arbitrator 405 may simply not select frames of the sequence for transmission to external nodes. In another embodiment, the later frames may not be queued for transmit scheduling at all until a period of time has elapsed.

The period of time can be determined in several manners. According to various embodiments, a fibre channel switch is configured to buffer a frame for no more than a fibre channel switch latency before the frame is dropped. It should be noted, however, that some switches may not drop a frame after a fibre channel switch latency has elapsed. In one embodiment, if the frame is held in a buffer in the switch for more than a latency period, a frame is dequeued from the virtual output queue and dropped. The fibre channel switch latency may depend on switching speeds and network congestion. The amount of time the frame can spend in a fibre channel switch before being dropped is referred to herein as the fibre channel switch latency.

The amount of time a frame can spend in a fibre channel network before being dropped is referred to herein as the fibre channel fabric drain latency or network drain latency. According to various embodiments, the fibre channel fabric drain latency or network drain latency is calculated by multiplying the fibre channel switch latency by the maximum number of hops it takes for a frame to traverse the fibre channel fabric. A wide variety of techniques for determining fibre channel switch latency and fibre channel fabric drain latency are contemplated.

Figure 5 is a process flow diagram showing the forwarding of frames arriving from a connected host or disk. At 501, a fibre channel fabric link change is detected. As noted above, the fibre channel fabric link change can lead to the out of order delivery of a frame sequence to a fibre channel device. According to various embodiments, A fibre channel fabric link change can be detected based on the receipt of link update messages or the transmission of link update messages. At 503, updated sets of next hops are generated for each destination based on the new link state information. The updated set of next hops can be generated using algorithms such as Fibre Channel Shortest Path First (FSPF). With new information about network topology, a switch may better be able to determine what is the best path for transmitting a frame to a particular destination. At 505, it is determined for each destination whether the set of next hops is equal to the updated set of next hops. If it is determined at 507 that all sets of next hops are equal to the corresponding updated sets of next hops, no action is taken.

For example, if the original path for transmitting a frame to a destination was through switches 103 and 107 while the new path for transmitting a frame to a destination is also through switches 103 and 107, no action is necessary. However, if the updated path for transmitting a frame to a destination is 105 and 107, the updated set of next hops is not equivalent to be original set of next hops. Blocking and dropping mechanisms can be applied. At 509, for each destination with a set of next hops different from the corresponding updated set of next hops, the queues or virtual output queues for transmission of a frame having that destination are blocked. According to various embodiments, blocking the virtual output queue may entail not transmitting the frame to a next hop. At 511, the routing table is updated

using the updated sets of next hops. The queues with a change in the set of next hops can then be blocked during a fabric drain latency or fabric drain period at 513.

The fabric drain latency allows time for earlier packets that may still remain at congested network switches to either be delivered or dropped from the network. Different fabric drain periods may be associated with each virtual output queue or a single fabric drain period can be applied to all the virtual output queues. After a fabric drain period has elapsed, the blocked queues are unblocked at 515. Later frames in the blocked queues can now be transmitted since earlier frames were either dropped or delivered to the destination. Later frames can be transmitted to a destination without the risk that earlier transmitted frames are still in the network waiting to arrive at the destination out of order.

The techniques described in Figure 5 can be applied at any port in a fibre channel fabric on a per VSAN, per destination basis. According to various embodiments, however, the techniques depicted in Figure 5 are applied at an edge-ports connecting a switch to a host or disk. The techniques for edge-ports focus on blocking traffic to avoid injecting traffic which will potentially be dropped.

Figure 6 is a process flow diagram showing the forwarding of frames between fibre channel switches. At 601, the fibre channel fabric link change is detected. At 603, updated sets of next hops are calculated for each destination. At 605, it is determined for each destination if the set of next hops is equal to the updated set of next hops. If all sets of next hops are equal to the corresponding updated sets of next hops at 607, no action needs to be taken. If the sets of next hops are not equal to all the corresponding updated sets of next hops, frames having destinations associated with a changed set of next hops are dropped at 609.

The routing table can be updated at 611 and the fibre channel switch waits for a corresponding fabric drain period associated with the virtual output queue at 613. After the fabric drain period has elapsed for a particular queue, frames that would be placed in the queue associated with a fabric drain period are now forwarded instead of being dropped.

As noted above, the techniques of the present invention can apply towards the in order delivery of fibre channel frames in a fibre channel fabric upon detection of a link change. Typically, when a new node or new link is added or an old node or an old link is subtracted from the fibre channel topology, there is a risk of out of order delivery of fibre channel frames.

However, changes in network topology are not the only events that can trigger out of order delivery. Changes in a channel between two switches can also lead to out of order frame delivery.

Figure 7 is a diagrammatic representation showing a channel change that can cause out of order delivery. Switches 107 and 109 may be interconnected originally through redundant links 701, 703, and 705 that form a channel 709. Traffic traveling between switches 107 and 109 may be distributed across the different links in the channel 709 based on factors such as fairness and load balancing. In another example, identical hash functions can be used at the transmitting and receiving switches to determine what link to access next. Using identical hash functions can provide for in order delivery.

When a new link 707 is added to form an updated channel 711, more links are now available for transmission between switch 107 and switch 109. It should be noted, however, that no change in network topology occurs. Switches 107 and 109 are still connected in the fibre channel fabric. According to various embodiments, no changes or updates to routing tables are necessary when a channel changes. The routing table may simply instruct that a frame be forwarded from switch 107 to switch 109. The forwarding from switch 107 to switch 109 is not affected by the addition of a new link 707 to the original channel 709 to form the updated channel 711.

However, adding a link to a channel can also cause out of order delivery of fibre channel frames. Using the example above where frames with sequence numbers 1-6 are being transmitted on link 701, 703, or 705, a frame with sequence numbers 7 and 8 may be transmitted along link 707. A link 707 may be an uncongested or a higher bandwidth link that allows the frames with sequence

numbers 7 and 8 to leave switch 109 before the frames 1-6 arrive. The techniques of the present invention provide for the blocking and dropping fibre channel frames to allow in order delivery to switches and nodes downstream of switch 109. Blocking is also herein referred to as delaying.

As noted above, when a channel changes, the routing table may not necessarily change as a link still exists between switch 107 and switch 109. Logic and mechanisms for determining what links are available in a channel for transmitting a frame are referred to herein as forwarding channel tables. It should be noted that forwarding channel tables may change. Figure 8A shows a forwarding channel table for a switch 107. For a next hop 109, entry 803 provides that links 701,703, and 705 can be used to transmit a frame from switch 107 to the next hop 109.

Figure 8B shows a forwarding channel table for a switch 107 after a change in channel. After a change in the channel occurs, entry 813 provides information that links 701,703, 705, and 707 can be used to transmit from a switch 107 to a next hop 109.

Figure 9 is a process flow diagram showing the forwarding the frames during a channel change, according to various embodiments. At 901, a switch detects a change in the channel. The addition or subtraction of one or more links can be a change in the channel. At 903, traffic directed toward the channel based on the routing table next hop is dropped. In one embodiment, traffic destined for transmission through the channel but not yet queued is dropped to allow changes in the forwarding channel tables to accommodate the channel change. However, traffic already queued remains. As noted above, the traffic already queued may reside in virtual output queues associated with particular flows. At 905, the queues for transmitting along the links in the channel are marked and blocked. In one example, marking and blocking the queues includes allowing all frames already in the queues to be transmitted while blocking any new frames for the switch drain latency time period.

At 907, a forwarding channel table can then be updated to add and/or remove links in the channel. At 909, traffic heading toward the channel that was previously dropped at 903 can now be allowed into the queues. A switch can then wait at 911 during a modified fibre channel switch drain latency. According to various embodiments, a modified fibre channel switch drain latency is longer than the standard fibre channel switch drain latency to allow all frames already in the queues to be transmitted instead of dropped. At 913, queues associated with the links in the channel are released and all frames in the queues are either transmitted or dropped if the frame is too old and new frame sequences can now be delivered in order.

Although mechanisms such as delaying frames can be used according to various embodiments, the techniques of the present invention also contemplate using labels to provide for in order delivery of fibre channel frames. One of the many reasons for using labels is that labels provide a fast mechanism for accessing entries in a routing table. Instead of looking at a destination address, the destination identifier can be an "in label" or an incoming label that can be used to quickly reference a routing table entry. Labels can be used for a variety of other reasons as well. Using labels in fibre channel networks is described in concurrently filed U.S. application No. 10/114,394, titled Label Switching In Fibre Channel Networks (Attorney Docket No. ANDIP009) by inventors Scott S. Lee and Dinesh G. Dutt, the entirety of which is incorporated by reference for all purposes. One example of an architecture for implementing label switching is Multiple Protocol Label Switching (MPLS) described in RFC 3031, the entirety of which is also incorporated by reference for all purposes.

Figure 10 is a diagrammatic representation of a fibre channel network including labels switching routers that can forward packets based on labels associated with frames. In addition to containing the destination address, a frame includes as a destination identifier an in label also referred to herein as an incoming label that allows a switch to quickly access an entry in a routing table. For example, a label switching router 1004 can receive a frame with a destination of 2 and an in label of 420. It should be noted that in Figures 10-12, a destination or a next hop of 1-5 refers to a switch 1001-1005. The label switching router 1004 can access its

routing table 1014 to recognize that the next hop is label switching router 1002 and the "out label" should be 220. The out label is also referred to herein as an outgoing label. According to various embodiments, the label switching router 1004 replaces the frame label value of 420 corresponding to the in label in the routing table with a frame label of 220 corresponding to the out label in the routing table 1014.

By replacing the label value, the label switching router 1004 provides label information to the next hop router 1002, to allow the label switching router 1002 to similarly access a routing table entry quickly. It should be noted that although label switching can be provided for fast access of entries in a routing table, label switching can be used for a variety of reasons. The techniques of the present invention provide that frames can be delivered in order by using labels.

When a label switching router 1002 receives a frame from label switching router 1004, the label switching router uses the label 220 to access an entry in the routing table 1012. Using the in label 220, the label switching router 1002 recognizes that the frame has reached the ultimate hop switch and no longer needs to be forwarded to another switch. The frame can then be forwarded to the ultimate destination which may be a host or a disk.

A variety of techniques can be used for generating routing tables with labels. In one embodiment, routing tables are generated upon the receipt of link state update packets under the FSPF protocol. Routing tables can be generated periodically or upon the identification of a change in link state. According to various embodiments, a newly generated routing table is associated with an incarnation number. A combination of all the incarnation numbers in a fibre channel fabric is herein referred to as a topology version number. In one embodiment, every time a new routing table is generated at a switch, the incarnation number is incremented by one. According to various embodiments, each label switching router in a fibre channel network not only generates new forwarding routes toward each destination, but each label switching router also generates new in labels different from the previous set of in labels.

Figure 11 is a diagrammatic representation showing label switching routers during a change in link state. Here, the link between label switching router 1004 and label switching router 1001 is no longer available. Link state update packets or link state records are propagated throughout the fibre channel fabric. To transmit to a label switching router 1001 from a label switching router 1004, label switching router 1004 can no longer forward directly to label switching router 1001. Instead, label switching router 1004 sends frames towards either label switching router 1002 or label switching router 1003. The new routes are reflected in the routing table 1114. The label switching router 1004 generates a new routing table 1114 with the new set of next hops and generates a new in label 411 replacing the old in label 410.

According to various embodiments, the label switching router 1004 sends label switching control messages to the other label switching routers in the fibre channel fabric in order to remove the outdated 410 in label. The label switching control messages sent to other label switching routers in the fibre channel fabric include a topology version number. In one embodiment, the other label switching routers verify that the topology version number of the label switching control messages matches the topology version number of the routing tables. If the topology version number of the label switching control message does not match the topology version number of the routing table in a particular label switching router or if the control message contains an older version number, the label switching control message is discarded.

If the topology version number of the label switching control message is newer than the topology version number of the routing table in a particular label switching router, the label switching control message can be stored and used later when the newer version of the routing table is available.

If the topology version numbers match, the label switching router receiving the label switching control message can remove the outdated label. For example, if a label switching router 1004 generates a new in label 411 replacing an old in label 410 and transmits a label switching control message to withdraw the outdated label to a label switching router 1002, label switching router 1002 will remove the out label 410 associated with a destination ID of 1 and a next hop of 4. According to

various embodiments, the label switching control message to remove an outdated label is referred to herein as a label withdraw message.

While out labels are unresolved, fibre channel frames are dropped to prevent out of order delivery of fibre channel frames. In Figures 11 and 12, a "?" is used to indicate that a label is unresolved. For example, suppose label switching router 1005 transmitted frames 1 and 2 to label switching router 1001 in order to eventually reach label switching router 1004. After the failure of the link between label switching router 1001 and label switching router 1004, label switching router 1005 transmitted frames 3 and 4 to label switching router 1002 in order to eventually reach label switching router 1004. If frames 1 and 2 are not dropped, they could arrive at label switching router 1004 after frames 3 and 4. However, because the out label at label switching router 1001 for destination label switching router 1004 is not resolved after the link failure, frames 1 and 2 are dropped and frames 3 and 4 can arrive in order at label switching router 1004. In other words, because the out label 440 associated with destination 4 in the routing table 1111 is no longer accurate, frames 1 and 2 can be dropped.

To resolve the out labels, label switching router 1004 advertises the new in labels including in label 411 to other label switching routers such as label switching router 1002. The label switching control message to add a new label is referred to as a label mapping message. In one example, label switching router 1004 sends a label mapping message having a label to label switching router 1002. The label mapping message received at label switching router 1002 can instruct the label switching router 1002 to use the new in label 411 as the new out label associated with the next hop of 4 and the destination of 1.

Figure 12 is a diagrammatic representation showing label switching routers in a fibre channel fabric where out labels are not yet completely resolved. Label switching router 1004 has generated a routing table 1214 that has replaced an old set of in labels 410, 420, 430, 440, and 450 with an updated stack of in labels 411, 421, 431, 441, and 451 respectively. The label switching router 1004 has also resolved its out labels and has replaced out labels 110, 110, 220, 330, 250, and 350, with new out labels 211, 311, 221, 331, 251, and 351, respectively. According to various

embodiments, out labels are resolved at label switching router 1004 after the router 1004 has received from all other routers label mapping messages with topology version numbers corresponding to that of the routing table 1214.

Label switching router 1002, however, has not fully result out labels in routing table 1212. Although the routing table 1212 contains a new set of in labels, the set of out labels his only partially resolved. In particular, although the label switching router 1002 has new out labels corresponding to the next hop of 4, it does not have new out labels corresponding to the next hop of 5. This may have resulted from the receipt of a label mapping message from label switching router 1004 but no label mapping message from label switching router 1005. If the label switching router 1002 receives a frame for transmission to a next hop of 5, the frame is dropped because the out labels associated with the next hop of five are unresolved.

As link state update messages and label switch control messages are passed throughout the fibre channel fabric, every switch eventually can resolve the out labels in the same manner label switching router 1004 and label switching router 1001 have resolved the out labels in Figure 12.

Figure 13 is a flow process diagram showing the forwarding of frames using labels. At 1301, a fibre channel fabric link changes detected. At 1303, updated sets of next hops for each destination are calculated based on link state information. According to various embodiments, the topology version number is identified at 1305. The topology version number can be a combination of the incarnation numbers of all of the switches in a fibre channel fabric. In one embodiment, the topology version number includes all of the incarnation numbers appended together. In another embodiment, the topology version number is a check sum of the various incarnation numbers. And in still another embodiment, the topology version number is a unique number meaningful to the switches in the network. At 1307, switches can transmit label switching control messages to withdraw previously advertised labels. At 1309, each label switching router can generate new in or incoming labels and advertise the new in or incoming labels to the other switches in the fabric.

It should be noted that the techniques of the present invention do not necessarily have to be performed in any particular order. For example, in one embodiment a label switching router can generate and advertise an in labels at the same time out labels are withdrawn at the other label switching routers.

At 1311, frames with old labels are dropped. Frames are also dropped while out labels are not resolved. For example, if the out labels associated with a next hop of six are not resolved, frames configured for transmission to a next hop of six are dropped. At 1313, out labels are resolved by using information from the label mapping messages or information from advertisements from the other label switching routers. Frames no longer are dropped, as out labels are resolved at 1315.

While the invention has been particularly shown and described with reference to specific embodiments thereof, it will be understood by those skilled in the art that changes in the form and details of the disclosed embodiments may be made without departing from the spirit or scope of the invention. For example, embodiments of the present invention may be employed with a variety of network protocols and architectures. Instructions such as quench messages can be sent at a variety of different times. It is therefore intended that the invention be interpreted to include all variations and equivalents that fall within the true spirit and scope of the present invention.

Claims:

1. A method for selectively delivering frames in a fibre channel fabric comprising a plurality of fibre channel switches, the method comprising:

detecting a fibre channel fabric link change at a fibre channel switch;

generating an updated routing table (1309) associated with a topology version number (1305), the topology version number comprising a number common to all of the fibre channel switches in the fibre channel fabric, wherein generating an updated routing table comprises determining a next hop, an incoming label, and a destination corresponding to each entry in the updated routing table;

receiving a frame at the fibre channel switch, the frame having a first destination and a first label corresponding to a first entry in the updated routing table;

determining whether the fibre channel switch has received a first outgoing label corresponding to the first entry in the updated routing table, the received first outgoing label having the same topology version number as the updated routing table; and

dropping the frame (1311) if it is determined that the fibre channel switch has not received the first outgoing label.

2. The method of claim 1, wherein the first label is equivalent to a first incoming label associated with the first entry.

3. The method of any of claims 1-2, further comprising: forwarding the frame to the next hop if it is determined that the fibre channel switch has received the first outgoing label.

4. The method of claim 3, wherein the routing table is associated with a particular virtual storage area network.

5. The method of claim 4, wherein the topology version number (1305) is derived using incarnation numbers with each switch in the virtual storage area network.

6. The method of claim 5, wherein the input and outgoing labels are MPLS labels.

7. The method of claim 6, wherein the first label is contained in the EISL header associated with the fibre channel frame.

8. The method of claim 7, wherein the first label in the frame is contained in the fibre channel frame header.

9. The method of any of claims 2-8, further comprising: receiving a first link state control message having the first outgoing label and a topology version number corresponding to the topology version number of the routing table.

10. The method of claim 9, further comprising: advertising the first incoming label to other switches in the fibre channel fabric.

11. Apparatus for selectively delivering frames in a fibre channel fabric comprising a plurality of fibre channel switches, the apparatus comprising:

means for detecting a fibre channel fabric link change at a fibre channel switch;

means for generating an updated routing table (1309) associated with a topology version number (1305), the topology version number being a number common to all of the fibre channel switches in the fibre channel fabric, wherein generating an updated routing table comprises determining a next hop, an incoming label, and a destination corresponding to each entry in the updated routing table;

means for receiving a frame at the fibre channel switch, the frame having a first destination and a first label corresponding to a first entry in the updated routing table;

means for determining whether the fibre channel switch has received a first outgoing label corresponding to the first entry in the updated routing table, the received first outgoing label having the same topology version number as the updated routing table; and

means for dropping the frame (1311) if it is determined that the fibre channel switch has not received the first outgoing label.

12. The apparatus of claim 11, wherein the first label is equivalent to a first incoming label associated with the first entry.

13. The apparatus of any of claims 11-12, further comprising: forwarding the frame to the next hop if it is determined that the fibre channel switch has received the first outgoing label.

14. The apparatus of claim 13, wherein the routing table is associated with a particular virtual storage area network.

15. The apparatus of claim 14, further comprising means for deriving the topology version number (1305) using incarnation numbers with each switch in the virtual storage area network.

16. The apparatus of any of claims 12-15, further comprising: means for receiving a first link state control message having the first outgoing label and a topology version number corresponding to the topology version number of the routing table.

17. The apparatus of claim 16, further comprising: means for advertising the first incoming label to other switches in the fibre channel fabric.

18. A method according to claim 1 the method further comprising:

identifying a set of next hops (101,105,107,109) at a fibre channel entity (111), wherein the set of next hops is used to forward frames based on a destination identifier (115) received with the frame;

identifying an updated set of next hops (101, 105, 107, 109) based on the fibre channel link change (501), wherein the updated set of next hops is used to forward a frame received at a fibre channel entity based on the destination identifier of the frame while accounting for the fibre channel fabric link change;

comparing the set of next hops to the updated set of next hops (505); and

preventing the forwarding of frames (509) towards the updated set of next hops for a predetermined period of time if it is determined that the set of next hops is different from the updated set of next hops.

19. The method of claim 18, wherein the predetermined period of time is a fibre channel fabric drain period (513).

20. The method of any of claims 18-19, wherein preventing the forwarding of frames (509) towards the updated set of next hops comprises:

blocking frames scheduled for forwarding to the updated set of next hops; or

dropping frames (609) scheduled for forwarding to the updated set of next hops.

21. The method of any of claims 18 to 20, further comprising updating a routing table (611) to forward frames to the updated set of next hops.

22. The method of any of claims 18 to 21, wherein the fibre channel entity is a fibre channel host or fibre channel switch comprising a plurality of routing tables, wherein each routing table is associated with a particular virtual storage area network.

23. The method of claim 21 or 22, further comprising:

allowing the forwarding of frames to the updated set of next hops after the predetermined period of time has elapsed and the routing table has been updated to schedule frames for forwarding to the updated set of next hops.

24. The method of claim 23, further comprising:

determining that the fibre channel host is associated with an edge-port before preventing the forwarding of frames to the updated set of next hops.

25. The method of claim 23, further comprising:

dropping frames that have been held in the switch for more than a fibre channel switch drain period.

26. The method of any of claims 23 to 25, further comprising:

determining that the fibre channel switch is associated with a non edge-port before preventing the forwarding of frames to the updated set of next hops.

27. A method according to Claim 18, further comprising:

dropping frames received for forwarding on the updated channel that have not yet been placed into output queues associated with the updated channel.

28. The method of claim 27, further comprising: marking the output queues (905) associated with the links in the updated channel to block all frames after the frames already in the queues are delivered.

29. The method of claim 28, wherein dropping frames comprises updating a forwarding channel table (907) to forward frames to a drop adjacency.

30. The method of claim 28, wherein the output queues are virtual output queues.

31. The method of claim 28, further comprising: updating a forwarding channel table (907) to forward frames on the updated channel.

32. The method of claim 31, further comprising: waiting for a period of time derived using the fibre channel switch latency.

33. The method of claim 32, further comprising: releasing the output queues (913) to allow forwarding of frames on the updated channel after the period of time has elapsed and the forwarding channel table has been updated (907) to forward frames on the updated channel.

34. The method of claim 33, wherein the forwarding-channel table is updated using a hash modulo.

35. Apparatus according to Claim 11, further comprising:

means for identifying a set of next hops (101, 105, 107, 109) at a fibre channel entity (111), wherein the set of next hops is used to forward frames based on a destination identifier (115);

means for identifying an updated set of next hops (101, 103, 107, 109), based on the fibre channel fabric link change (501), wherein the updated set of next hops is used to forward a frame received at a fibre channel entity based on a destination identifier while accounting for the fibre channel fabric link change;

means for comparing the set of next hops to the updated set of next hops; and

means for preventing the forwarding of frames (509) towards the updated set of next hops for a predetermined period of time if it is determined that the set of next hops is different from the updated set of next hops.

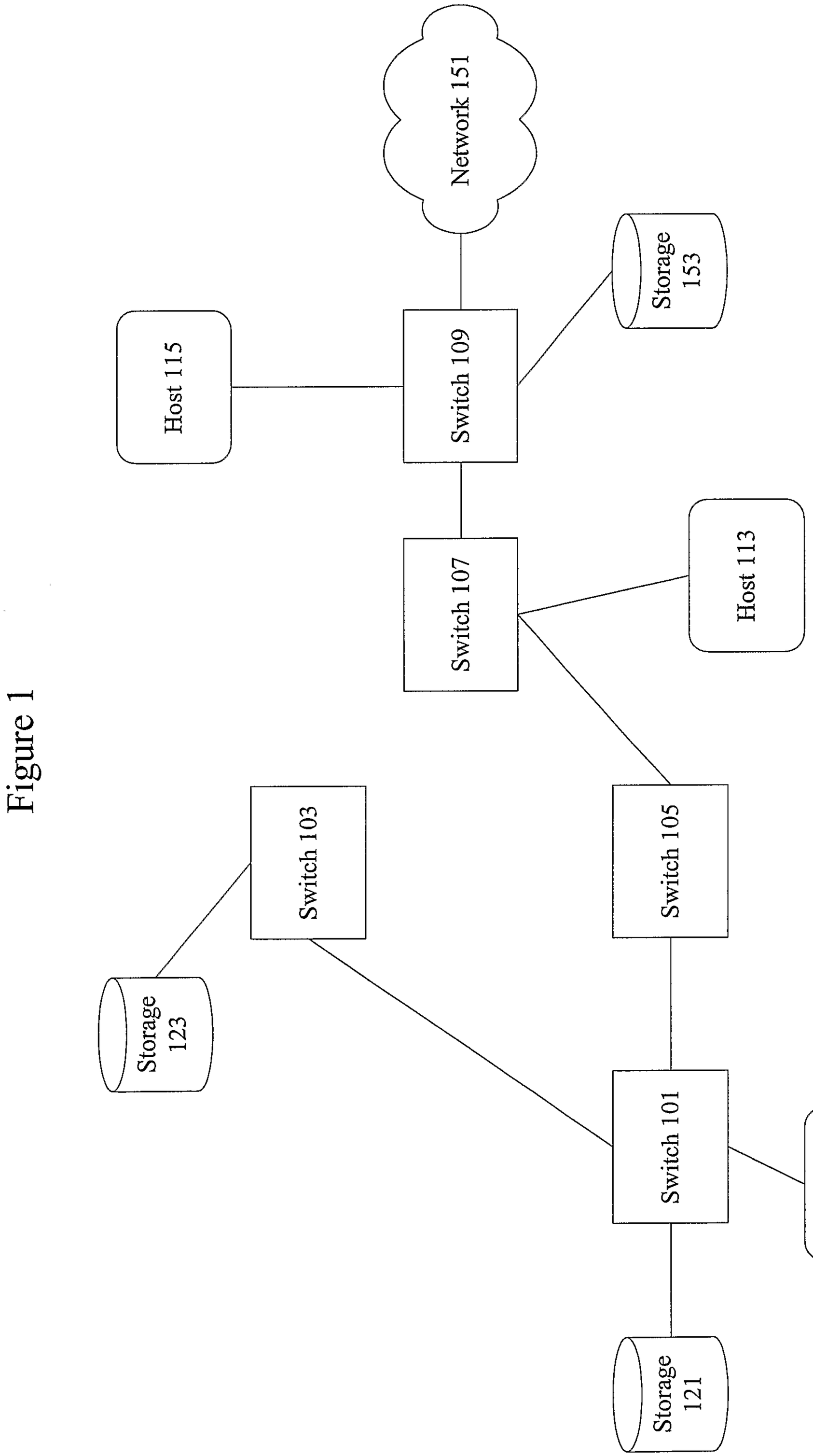

Network 151
Host 115
Switch 109
Storage 153
Switch 107
Host 113
Switch 103
Switch 105
Storage 123
Switch 101
Host 111
Storage 121

Figure 1

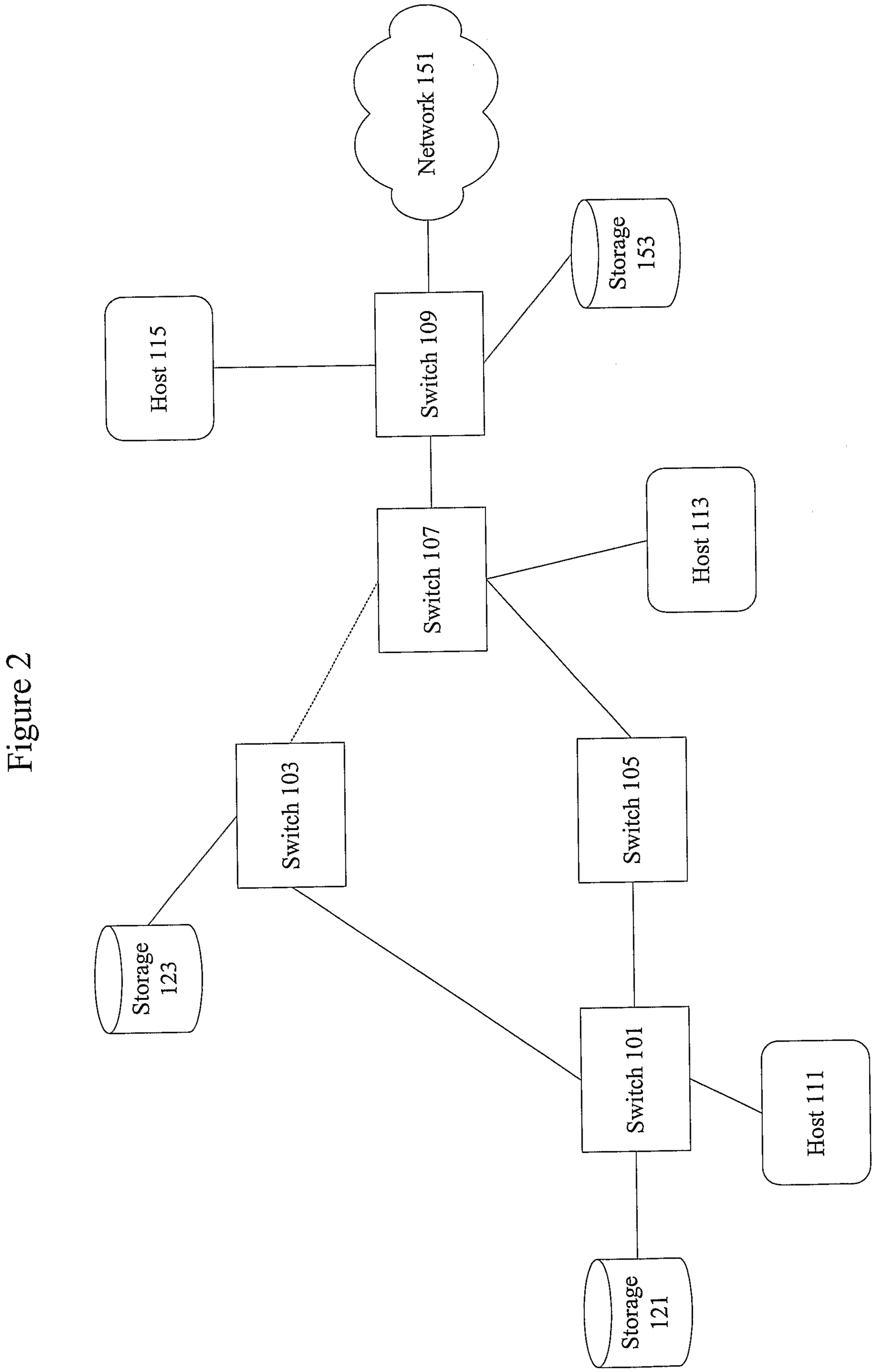

Network 151
Host 115
Switch 109
Storage 153
Switch 107
Host 113
Switch 103
Switch 105
Storage 123
Switch 101
Host 111
Storage 121

Figure 2

Figure 3A

	Routing Table For Switch 101	
	Destination Identifier	Set Of Next Hops
303	Switch 101	Processor
305	Switch 103	Switch 103
307	Switch 105	Switch 105
309	Switch 107	Switch 105
311	Switch 109	Switch 105
	315	317

Figure 3B

	Updated Routing Table For Switch 101	
	Destination Identifier	Updated Set Of Next Hops
323	Switch 101	Processor
325	Switch 103	Switch 103
327	Switch 105	Switch 105
329	Switch 107	Switch 103 Switch 105
331	Switch 109	Switch 103 Switch 105
	335	337

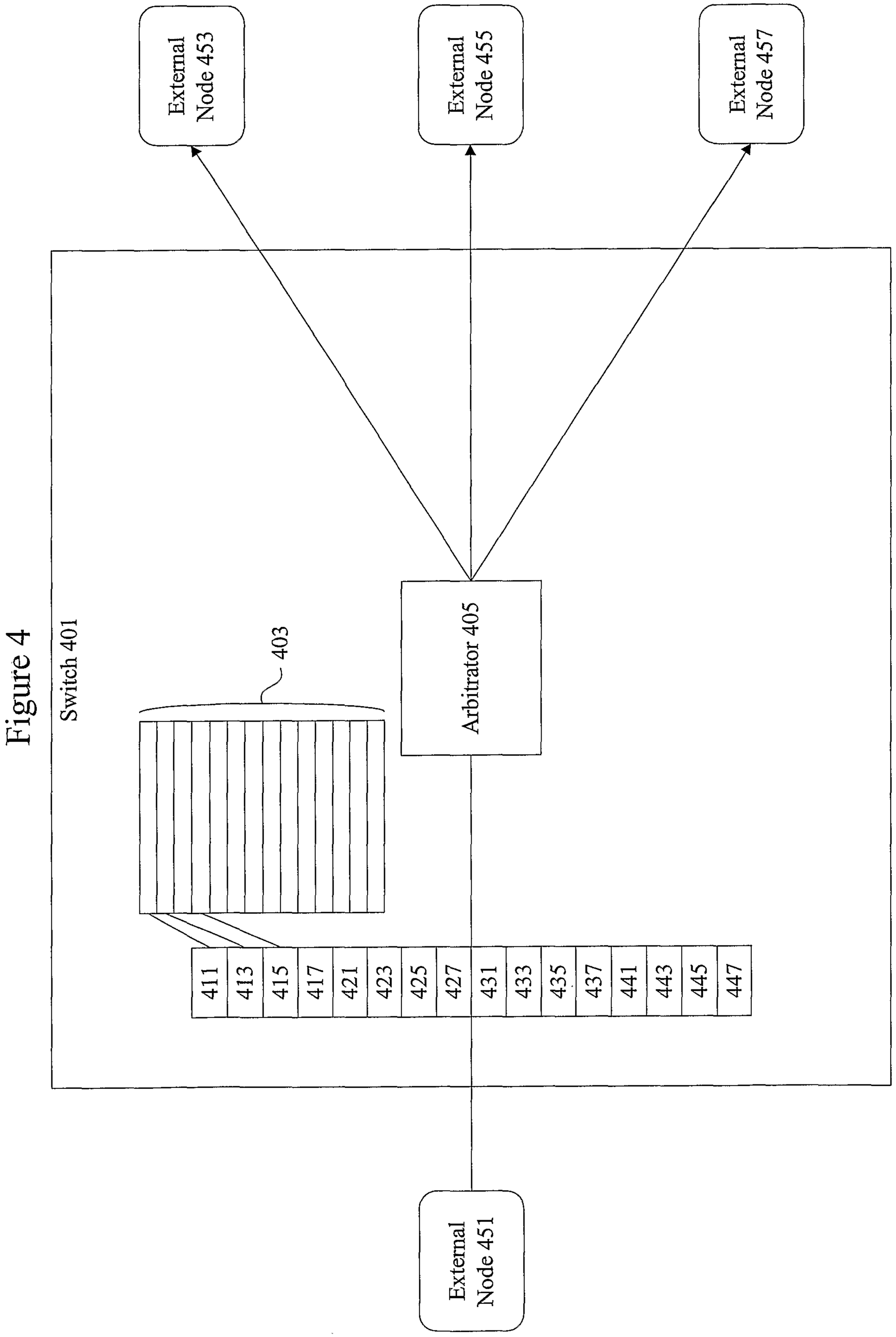

External Node 453
External Node 455
External Node 457
Switch 401
403
Arbitrator 405
411
413
415
417
421
423
425
427
431
433
435
437
441
443
445
447
External Node 451

Figure 4

Figure 5

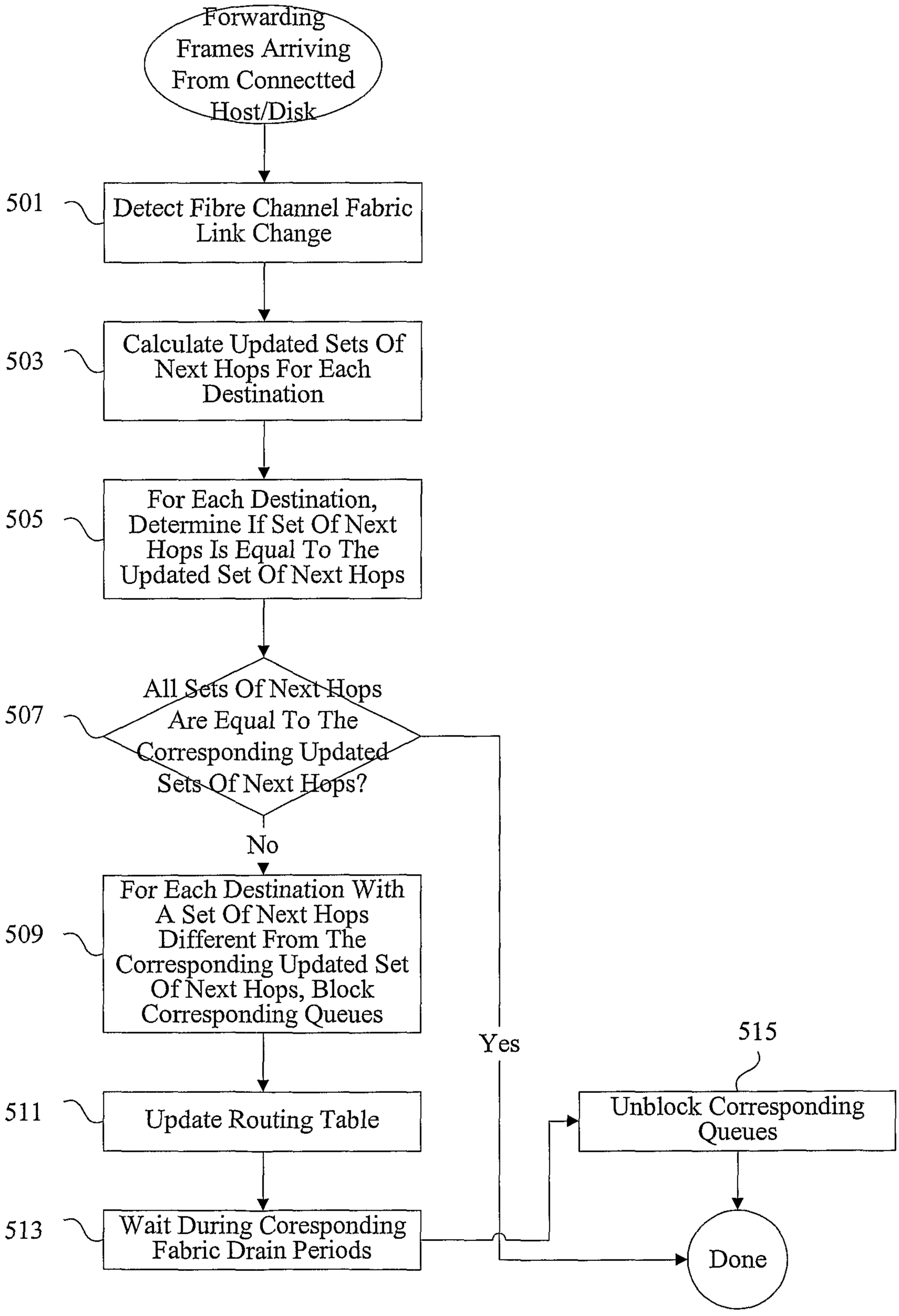
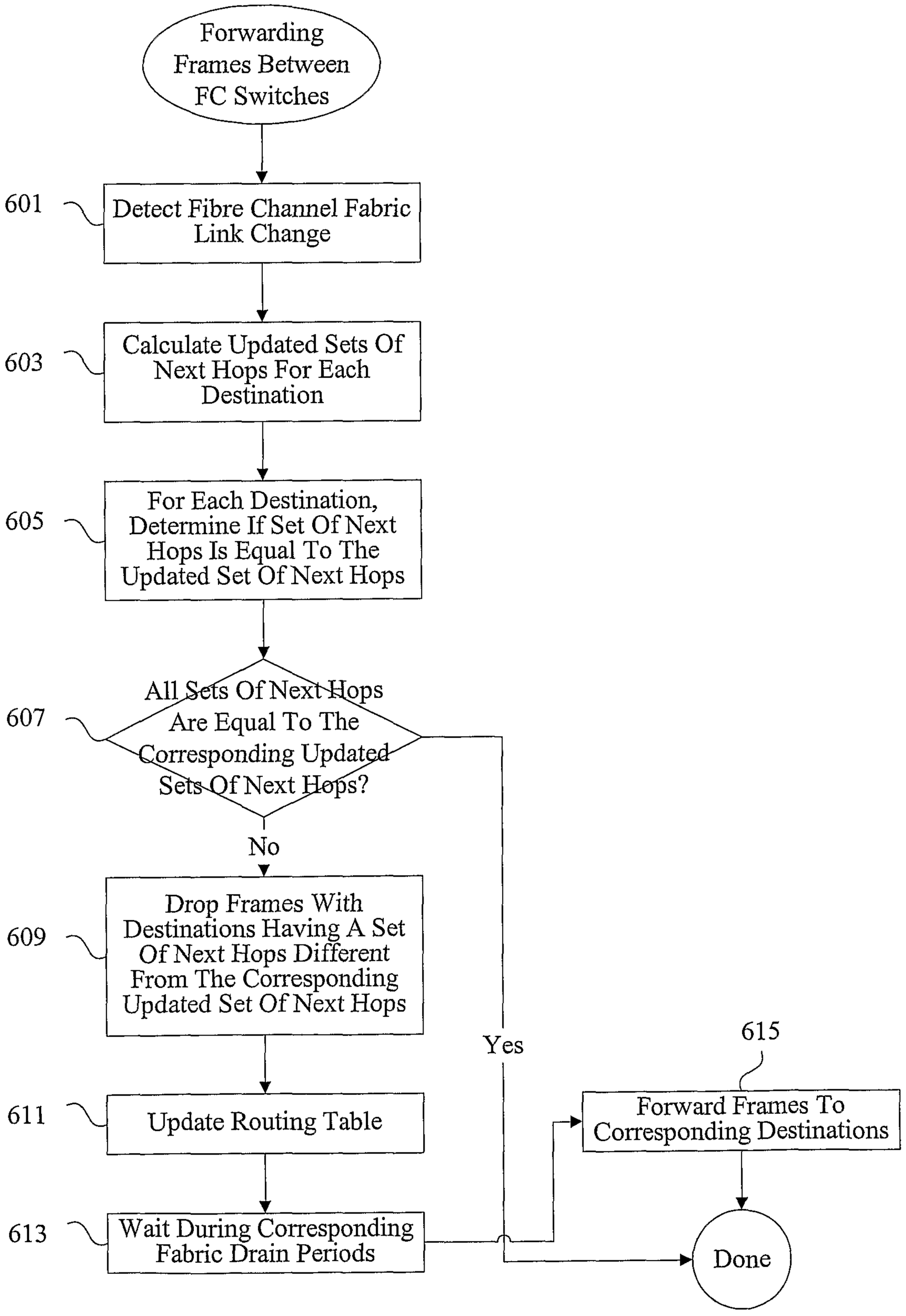

Forwarding Frames Arriving From Connectted Host/Disk
501
Detect Fibre Channel Fabric Link Change
503
Calculate Updated Sets Of Next Hops For Each Destination
505
For Each Destination, Determine If Set Of Next Hops Is Equal To The Updated Set Of Next Hops
507
All Sets Of Next Hops Are Equal To The Corresponding Updated Sets Of Next Hops?
No
Yes
509
For Each Destination With A Set Of Next Hops Different From The Corresponding Updated Set Of Next Hops, Block Corresponding Queues
511
Update Routing Table
513
Wait During Coresponding Fabric Drain Periods
515
Unblock Corresponding Queues
Done

Figure 6


Forwarding Frames Between FC Switches
601
Detect Fibre Channel Fabric Link Change
603
Calculate Updated Sets Of Next Hops For Each Destination
605
For Each Destination, Determine If Set Of Next Hops Is Equal To The Updated Set Of Next Hops
607
All Sets Of Next Hops Are Equal To The Corresponding Updated Sets Of Next Hops?
No
Yes
609
Drop Frames With Destinations Having A Set Of Next Hops Different From The Corresponding Updated Set Of Next Hops
611
Update Routing Table
613
Wait During Corresponding Fabric Drain Periods
615
Forward Frames To Corresponding Destinations
Done

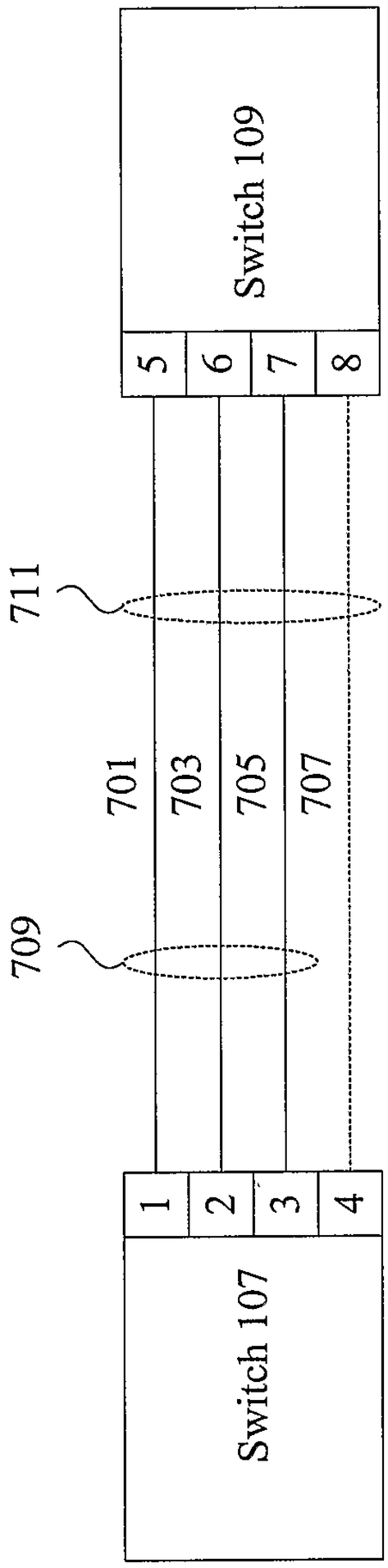


Figure 7

Figure 8A

Forwarding Channel Table For Switch 107	
Next Hop	Links
109	Link 701, 703, and 705
...	...

803

805

807 809

Figure 8B

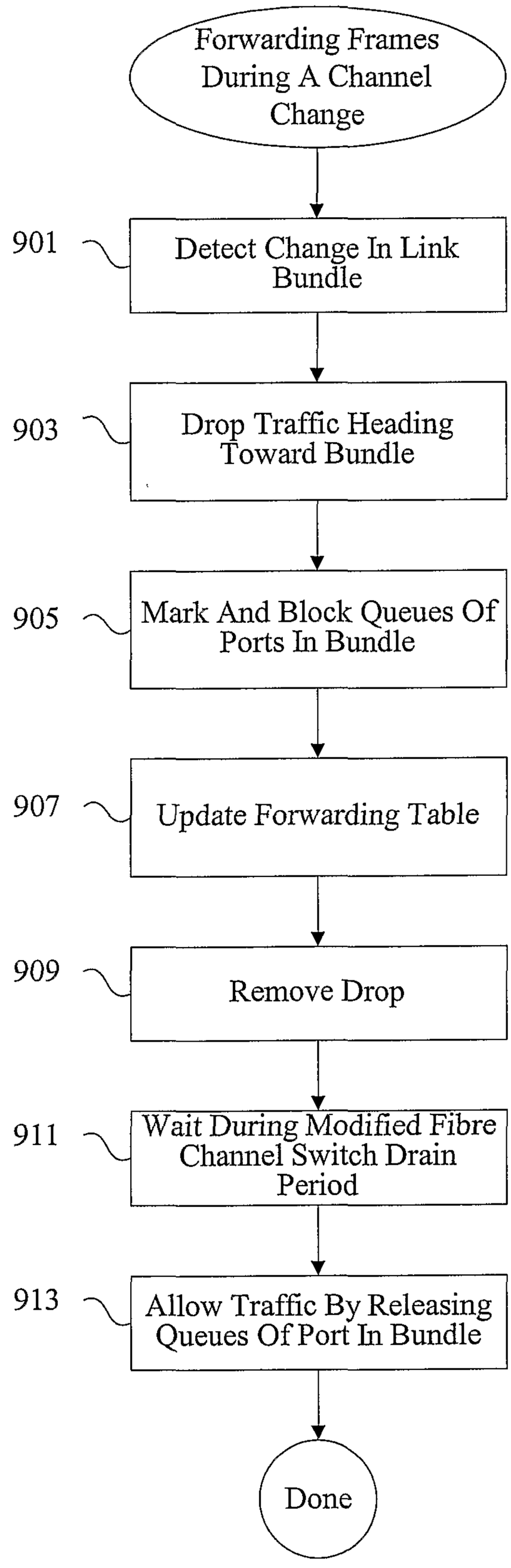
Updated Forwarding Channel Table For Switch 107	
Next Hop	Links
109	Link 701, 703, 705, and 707
...	...

813

815

817 819

Figure 9


Forwarding Frames During A Channel Change
901
Detect Change In Link Bundle
903
Drop Traffic Heading Toward Bundle
905
Mark And Block Queues Of Ports In Bundle
907
Update Forwarding Table
909
Remove Drop
911
Wait During Modified Fibre Channel Switch Drain Period
913
Allow Traffic By Releasing Queues Of Port In Bundle
Done

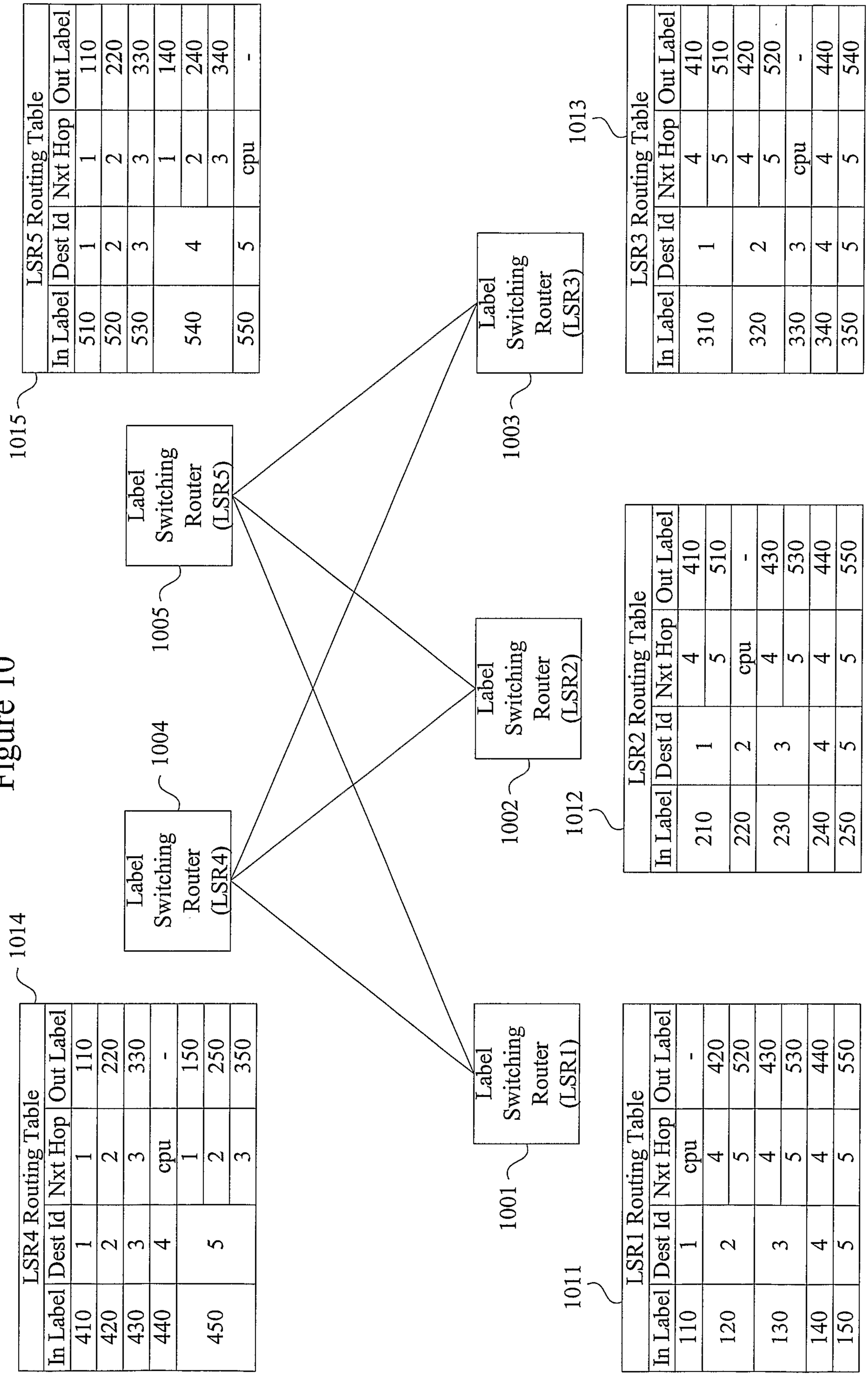
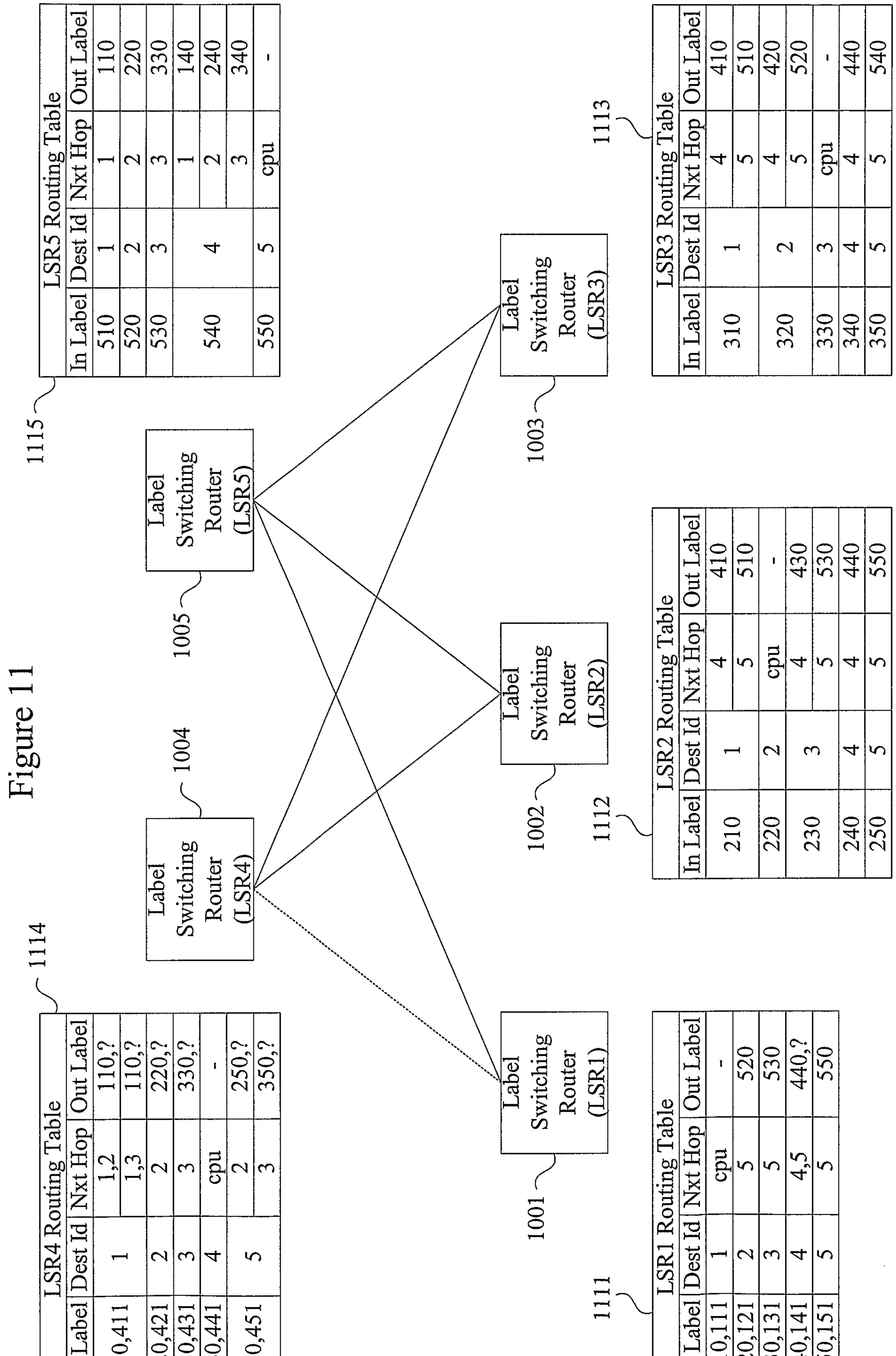

Label Switching Router (LSR5)
1005
Label Switching Router (LSR3)
1003
Label Switching Router (LSR4)
1004
Label Switching Router (LSR2)
1002
Label Switching Router (LSR1)
1001
1015
LSR5 Routing Table
In Label | Dest Id | Nxt Hop | Out Label
510 | 1 | 1 | 110
520 | 2 | 2 | 220
530 | 3 | 3 | 330
540 | 4 | 1 | 140
540 | 4 | 2 | 240
540 | 4 | 3 | 340
550 | 5 | cpu | -
1013
LSR3 Routing Table
In Label | Dest Id | Nxt Hop | Out Label
310 | 1 | 4 | 410
310 | 1 | 5 | 510
320 | 2 | 4 | 420
320 | 2 | 5 | 520
330 | 3 | cpu | -
340 | 4 | 4 | 440
350 | 5 | 5 | 540
1012
LSR2 Routing Table
In Label | Dest Id | Nxt Hop | Out Label
210 | 1 | 4 | 410
210 | 1 | 5 | 510
220 | 2 | cpu | -
230 | 3 | 4 | 430
230 | 3 | 5 | 530
240 | 4 | 4 | 440
250 | 5 | 5 | 550
1011
LSR1 Routing Table
In Label | Dest Id | Nxt Hop | Out Label
110 | 1 | cpu | -
120 | 2 | 4 | 420
120 | 2 | 5 | 520
130 | 3 | 4 | 430
130 | 3 | 5 | 530
140 | 4 | 4 | 440
150 | 5 | 5 | 550
1014
LSR4 Routing Table
In Label | Dest Id | Nxt Hop | Out Label
410 | 1 | 1 | 110
420 | 2 | 2 | 220
430 | 3 | 3 | 330
440 | 4 | cpu | -
450 | 5 | 1 | 150
450 | 5 | 2 | 250
450 | 5 | 3 | 350

Figure 10

Figure 11



LSR5 Routing Table (1115)

In Label	Dest Id	Nxt Hop	Out Label
510	1	1	110
520	2	2	220
530	3	3	330
540	4	1	140
		2	240
		3	340
550	5	cpu	-

LSR3 Routing Table (1113)

In Label	Dest Id	Nxt Hop	Out Label
310	1	4	410
		5	510
320	2	4	420
		5	520
330	3	cpu	-
340	4	4	440
350	5	5	540

LSR2 Routing Table (1112)

In Label	Dest Id	Nxt Hop	Out Label
210	1	4	410
		5	510
220	2	cpu	-
		4	430
230	3	5	530
		4	440
240	4	5	550
250	5		

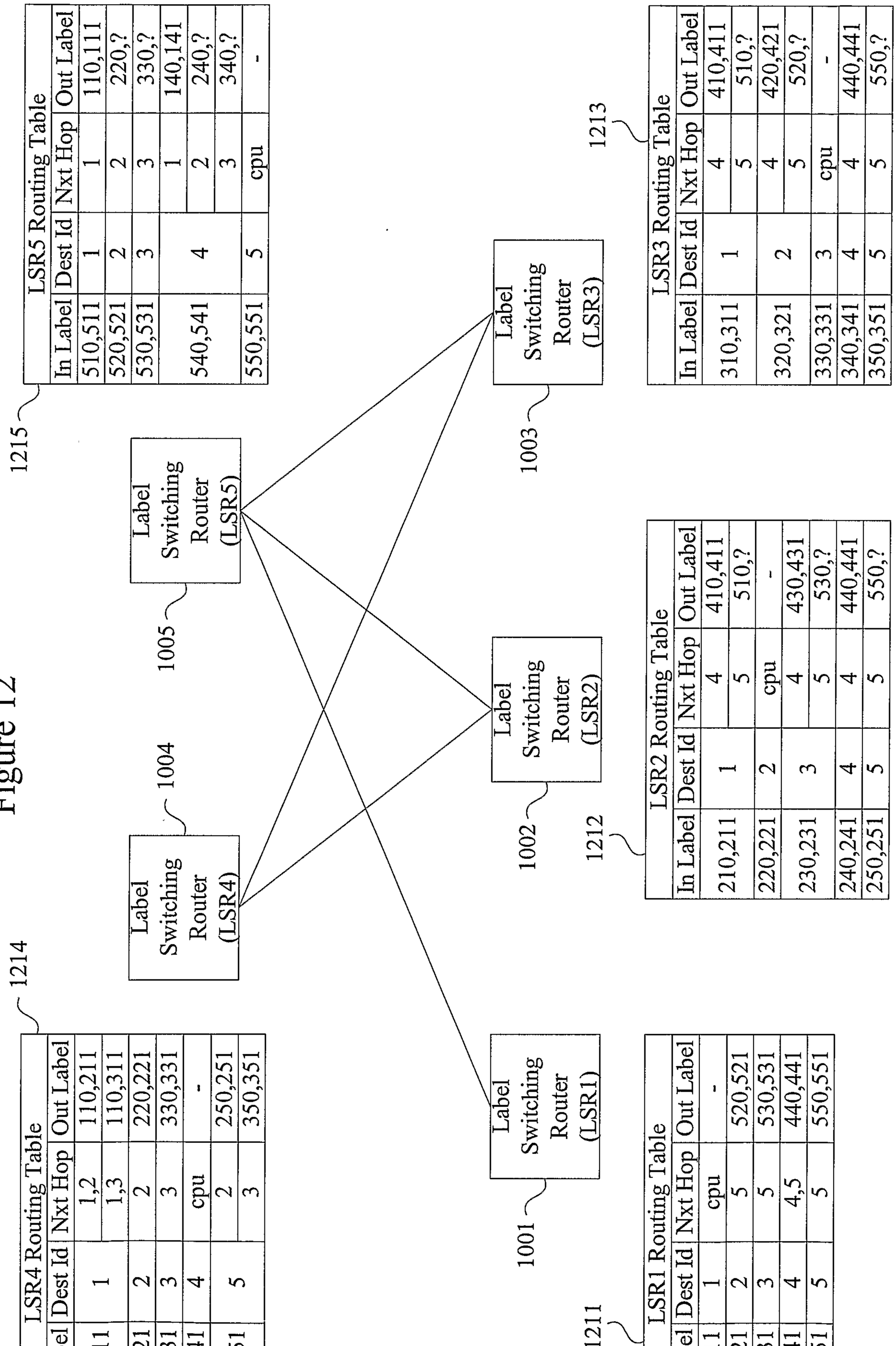
LSR4 Routing Table (1114)

In Label	Dest Id	Nxt Hop	Out Label
410,411	1	1,2	110,?
		1,3	110,?
420,421	2	2	220,?
430,431	3	3	330,?
440,441	4	cpu	-
450,451	5	2	250,?
		3	350,?

LSR1 Routing Table (1111)

In Label	Dest Id	Nxt Hop	Out Label
110,111	1	cpu	-
120,121	2	5	520
130,131	3	5	530
140,141	4	4,5	440,?
150,151	5	5	550

Figure 12



LSR5 Routing Table (1215)

In Label	Dest Id	Nxt Hop	Out Label
510,511	1	1	110,111
520,521	2	2	220,?
530,531	3	3	330,?
540,541	4	1	140,141
		2	240,?
		3	340,?
550,551	5	cpu	-

LSR3 Routing Table (1213)

In Label	Dest Id	Nxt Hop	Out Label
310,311	1	4	410,411
		5	510,?
320,321	2	4	420,421
		5	520,?
330,331	3	cpu	-
340,341	4	4	440,441
350,351	5	5	550,?

LSR4 Routing Table (1214)

In Label	Dest Id	Nxt Hop	Out Label
410,411	1	1,2	110,211
		1,3	110,311
420,421	2	2	220,221
430,431	3	3	330,331
440,441	4	cpu	-
450,451	5	2	250,251
		3	350,351

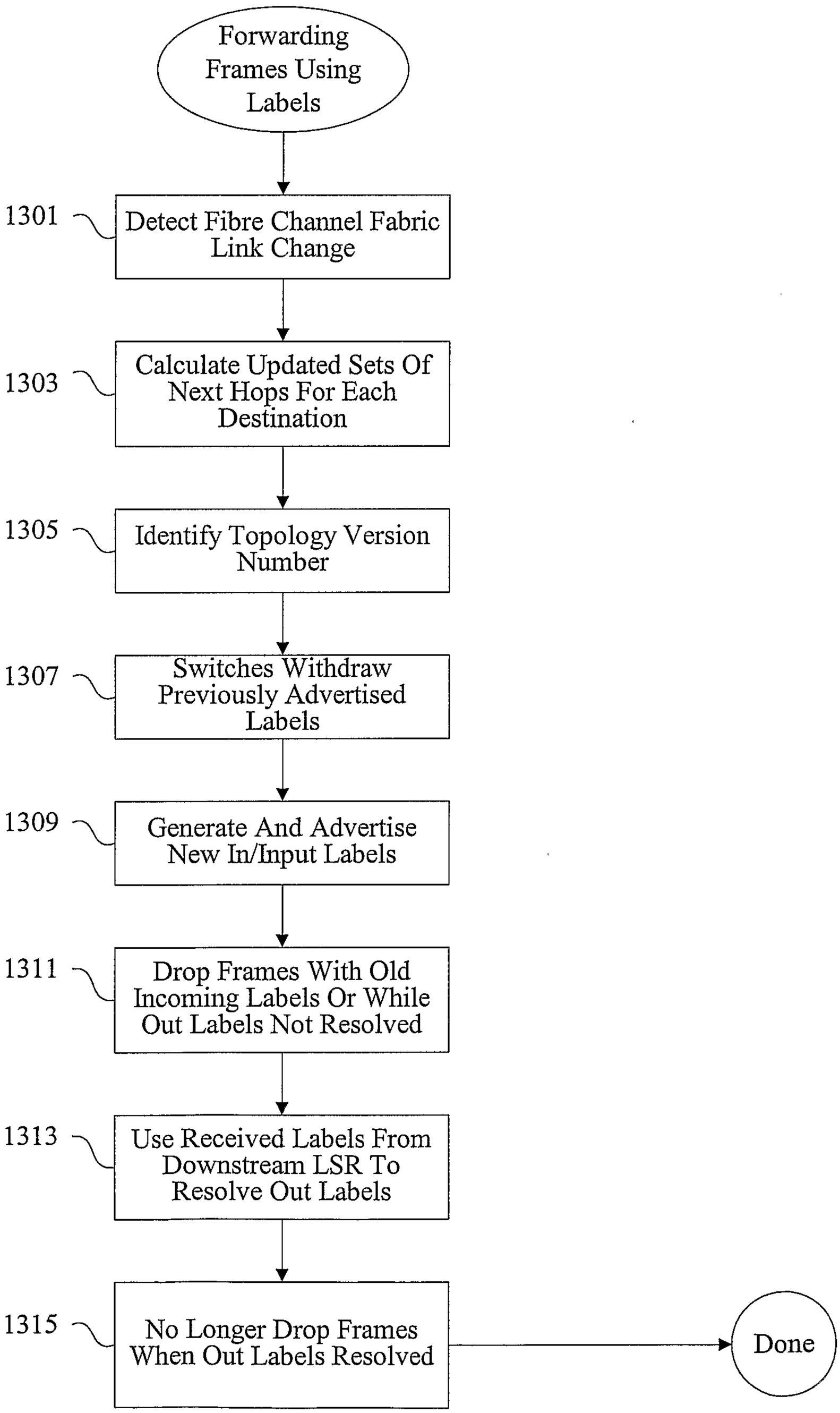
LSR2 Routing Table (1212)

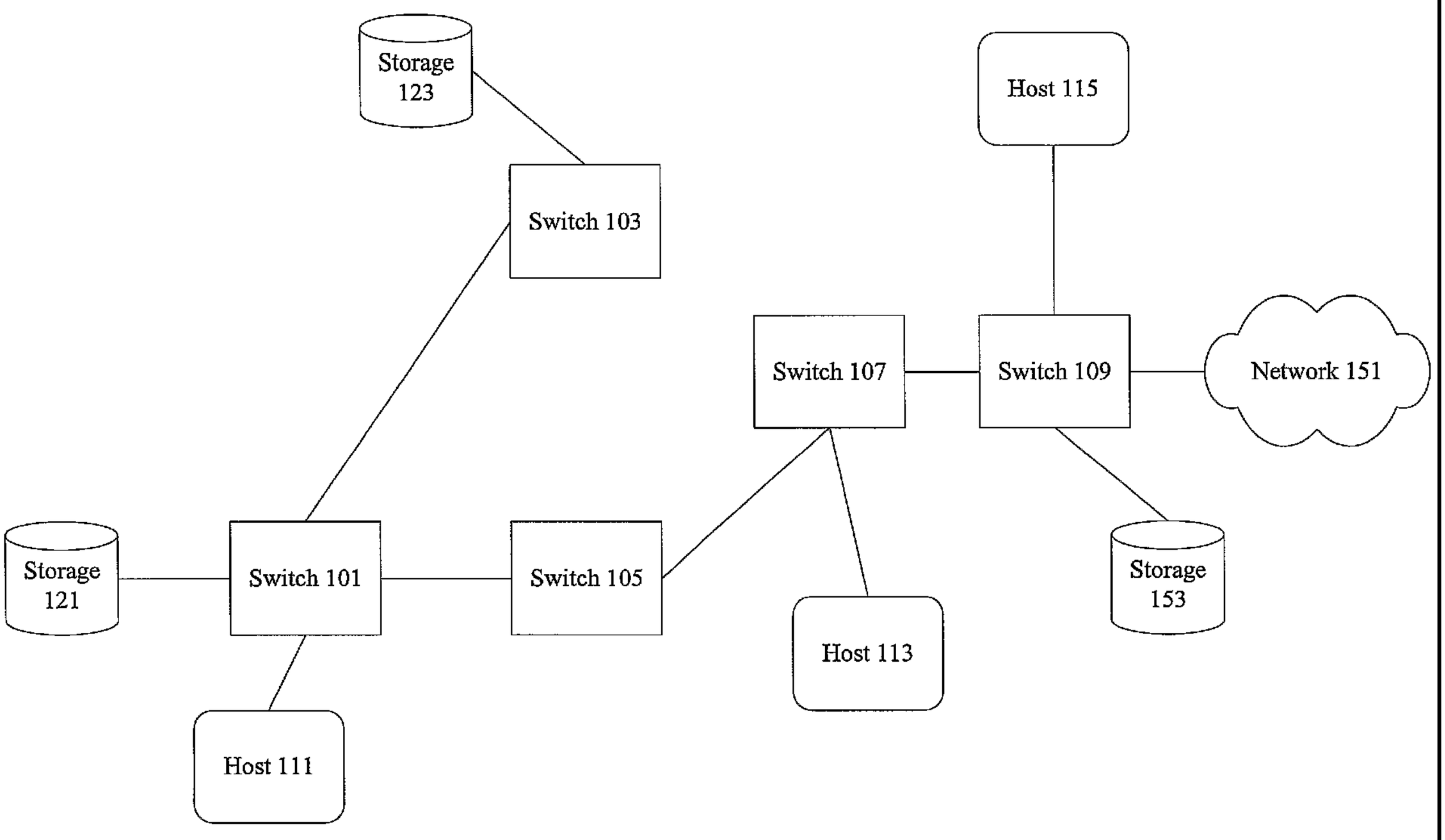
In Label	Dest Id	Nxt Hop	Out Label
210,211	1	4	410,411
		5	510,?
220,221	2	cpu	-
230,231	3	4	430,431
		5	530,?
240,241	4	4	440,441
250,251	5	5	550,?

LSR1 Routing Table (1211)

In Label	Dest Id	Nxt Hop	Out Label
110,111	1	cpu	-
120,121	2	5	520,521
130,131	3	5	530,531
140,141	4	4,5	440,441
150,151	5	5	550,551

Figure 13


Forwarding Frames Using Labels
1301
Detect Fibre Channel Fabric Link Change
1303
Calculate Updated Sets Of Next Hops For Each Destination
1305
Identify Topology Version Number
1307
Switches Withdraw Previously Advertised Labels
1309
Generate And Advertise New In/Input Labels
1311
Drop Frames With Old Incoming Labels Or While Out Labels Not Resolved
1313
Use Received Labels From Downstream LSR To Resolve Out Labels
1315
No Longer Drop Frames When Out Labels Resolved
Done


Storage 123
Switch 103
Host 115
Switch 107
Switch 109
Network 151
Storage 121
Switch 101
Switch 105
Host 113
Storage 153
Host 111