US009270580B1

(12) **United States Patent**
Abraham et al.

(10) **Patent No.: US 9,270,580 B1**
(45) **Date of Patent: Feb. 23, 2016**

(54) **METHOD AND SYSTEM FOR TRAFFIC ISOLATION IN A NETWORK**

(75) Inventors: **Vineet M. Abraham**, Sunnyvale, CA (US); **Sathish K. Gnanasekaran**, Santa Clara, CA (US); **Shashank R. Tadisina**, San Jose, CA (US); **Daniel Ji Yong Park Chung**, San Jose, CA (US); **Raymond Yimin Lai**, Fremont, CA (US)

(73) Assignee: **BROCADE COMMUNICATIONS SYSTEMS, INC.**, San Jose, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 937 days.

(21) Appl. No.: **12/550,227**

(22) Filed: **Aug. 28, 2009**

(51) **Int. Cl.**
H04L 12/703 (2013.01)
H04L 12/707 (2013.01)

(52) **U.S. Cl.**
CPC ***H04L 45/28*** (2013.01); *H04L 45/22* (2013.01)

(58) **Field of Classification Search**
None
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,633,861	A	5/1997	Hanson	
5,638,359	A	6/1997	Peltola	
5,719,853	A	2/1998	Ikeda	
5,970,048	A	10/1999	Pajuvirta	
6,014,383	A	1/2000	McCarty	
6,091,725	A	7/2000	Cheriton	
6,160,793	A	12/2000	Ghani	
6,185,189	B1	2/2001	Brassier	
6,233,236	B1	5/2001	Nelson	
6,381,642	B1	4/2002	O'Donnell	
6,427,114	B1	7/2002	Olsson	
6,724,722	B1	4/2004	Wang	
6,765,919	B1 *	7/2004	Banks et al.	370/401
6,980,525	B2	12/2005	Banks	
7,120,128	B2	10/2006	Banks	
7,145,868	B2	12/2006	Giroux	
7,167,472	B2	1/2007	Wu	
7,283,486	B2	10/2007	Banks	
7,352,740	B2	4/2008	Hammons	
7,366,194	B2	4/2008	Yu	
7,430,203	B2	9/2008	Millet	
2003/0021223	A1 *	1/2003	Kashyap	370/217
2003/0090997	A1 *	5/2003	Lindstrom	370/228
2003/0195956	A1 *	10/2003	Bramhall et al.	709/223
2004/0078599	A1 *	4/2004	Nahum	713/201
2006/0002292	A1 *	1/2006	Chang et al.	370/225
2006/0002293	A1 *	1/2006	Huck et al.	370/228
2006/0023707	A1 *	2/2006	Makishima et al.	370/389
2006/0215663	A1 *	9/2006	Banerjee et al.	370/395.21
2006/0262784	A1 *	11/2006	Cheethirala et al.	370/389
2007/0070901	A1	3/2007	Aloni	
2007/0253326	A1 *	11/2007	Saha et al.	370/217

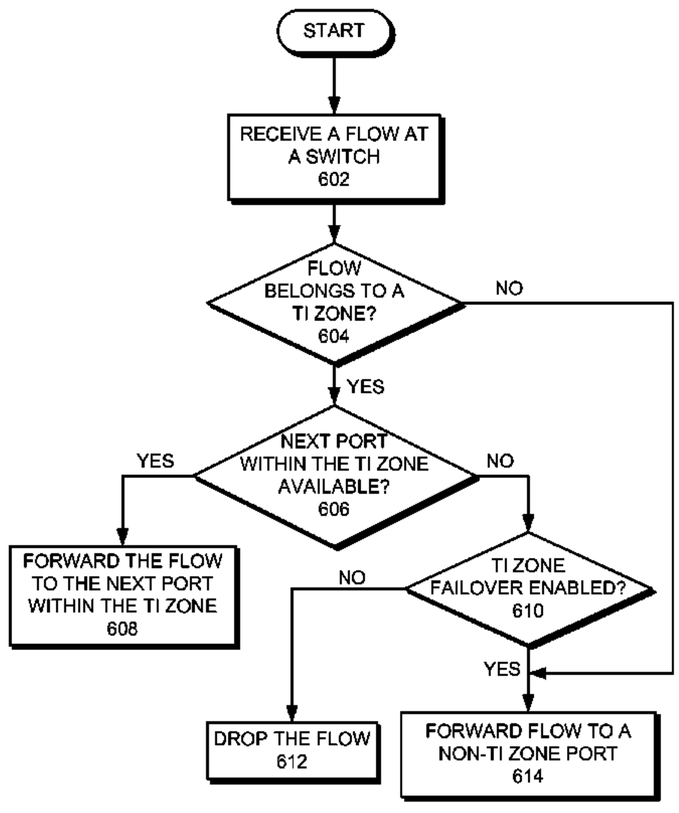
* cited by examiner

Primary Examiner — Huy D Vu
Assistant Examiner — James P Duffy
(74) *Attorney, Agent, or Firm* — Shun Yao; Park, Vaughan, Fleming & Dowler LLP

(57) **ABSTRACT**

One embodiment of the present invention provides a system that facilitates traffic isolation (TI) in a network. During operation, the system configures a set of switch ports as members of a TI zone. The switch ports are part of an end-to-end path across one or more switch domains between a source and a destination. The switch ports within the TI zone and outside the TI Zone belong to a common storage area network (SAN) zone which compartmentalizes data for security purposes. The system then determines whether a data flow entering a switch domain belongs to the TI zone. The system subsequently forwards the data flow to the next port within the TI zone if the data flow belongs to the TI zone.

24 Claims, 8 Drawing Sheets



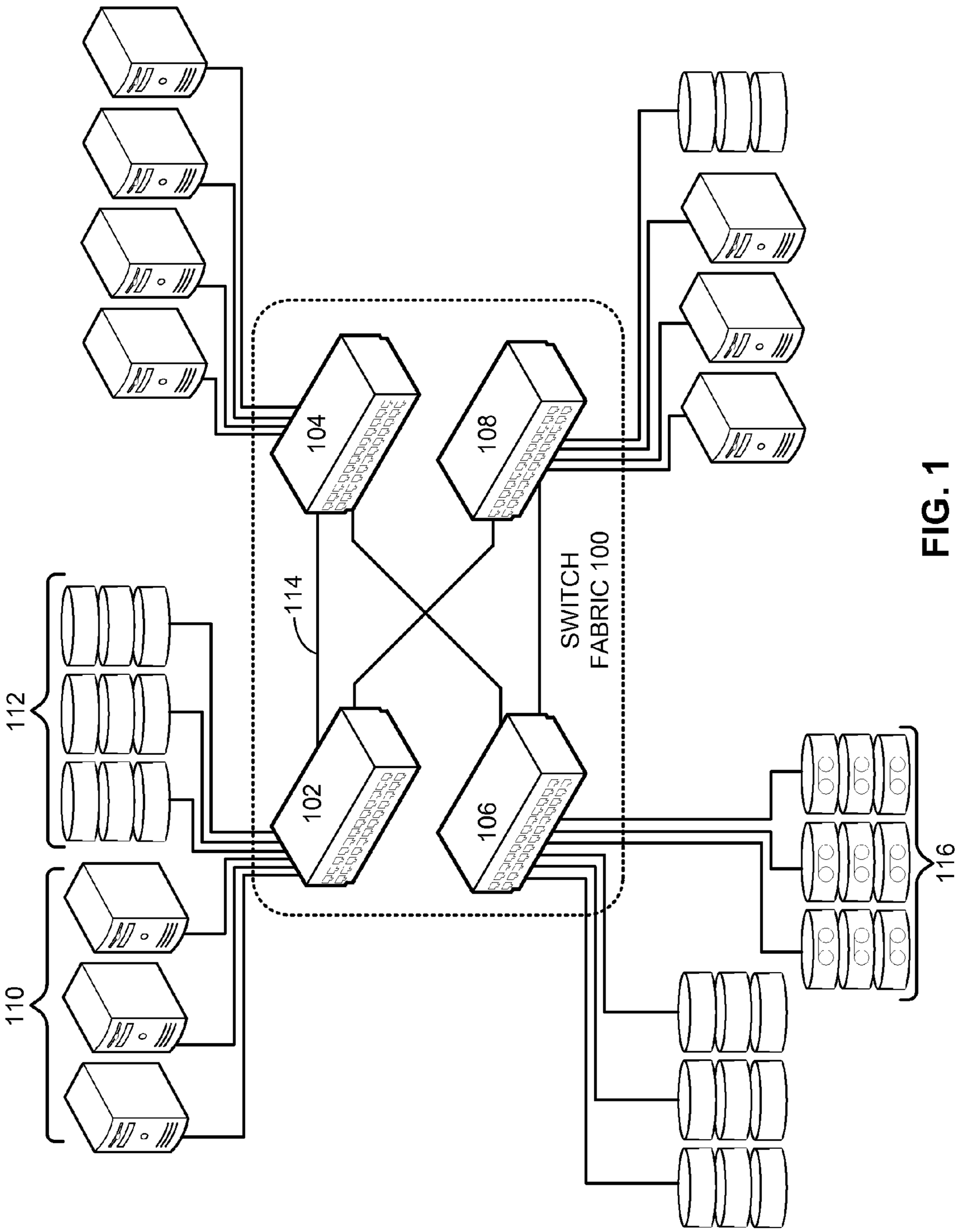


FIG. 1

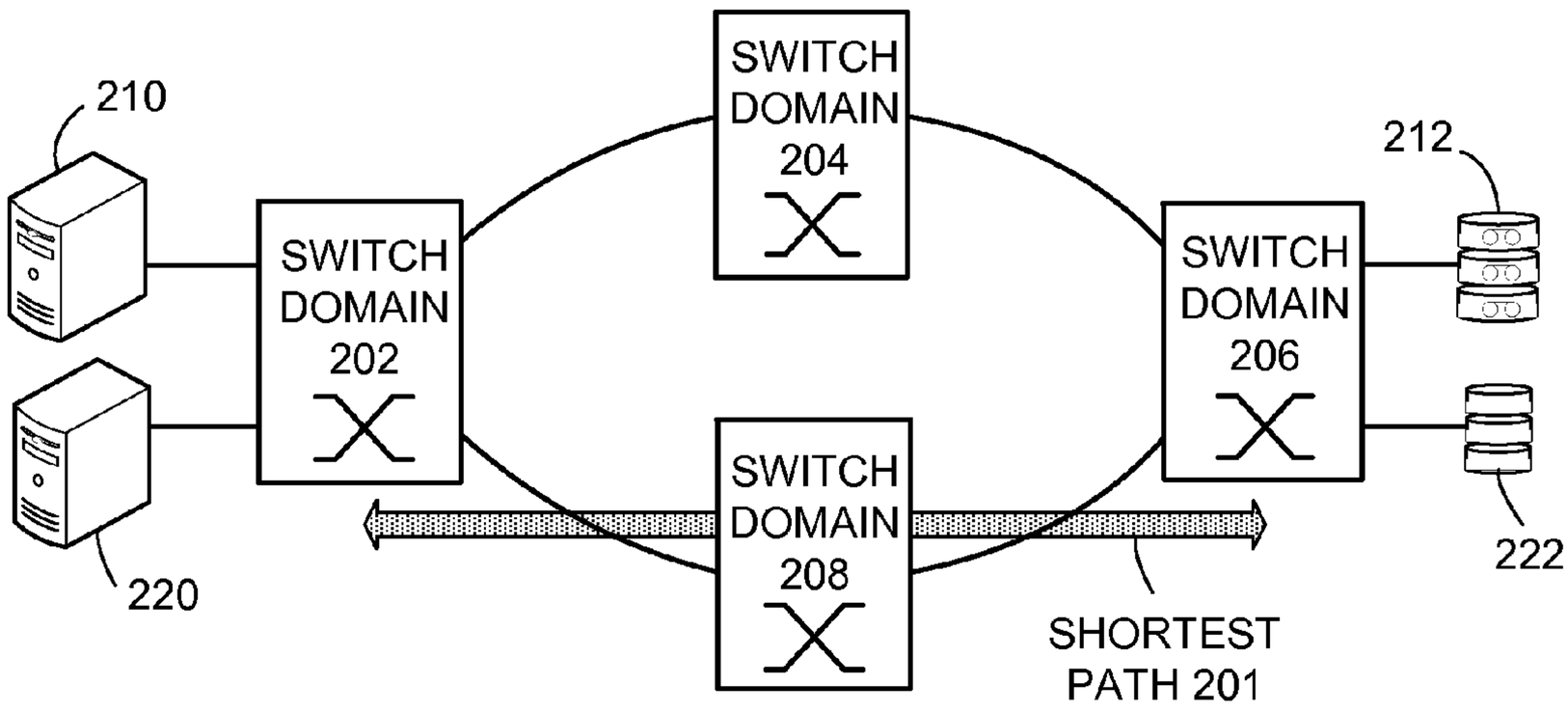

210
SWITCH DOMAIN 204
212
SWITCH DOMAIN 202
SWITCH DOMAIN 206
SWITCH DOMAIN 208
220
222
SHORTEST PATH 201

FIG. 2A

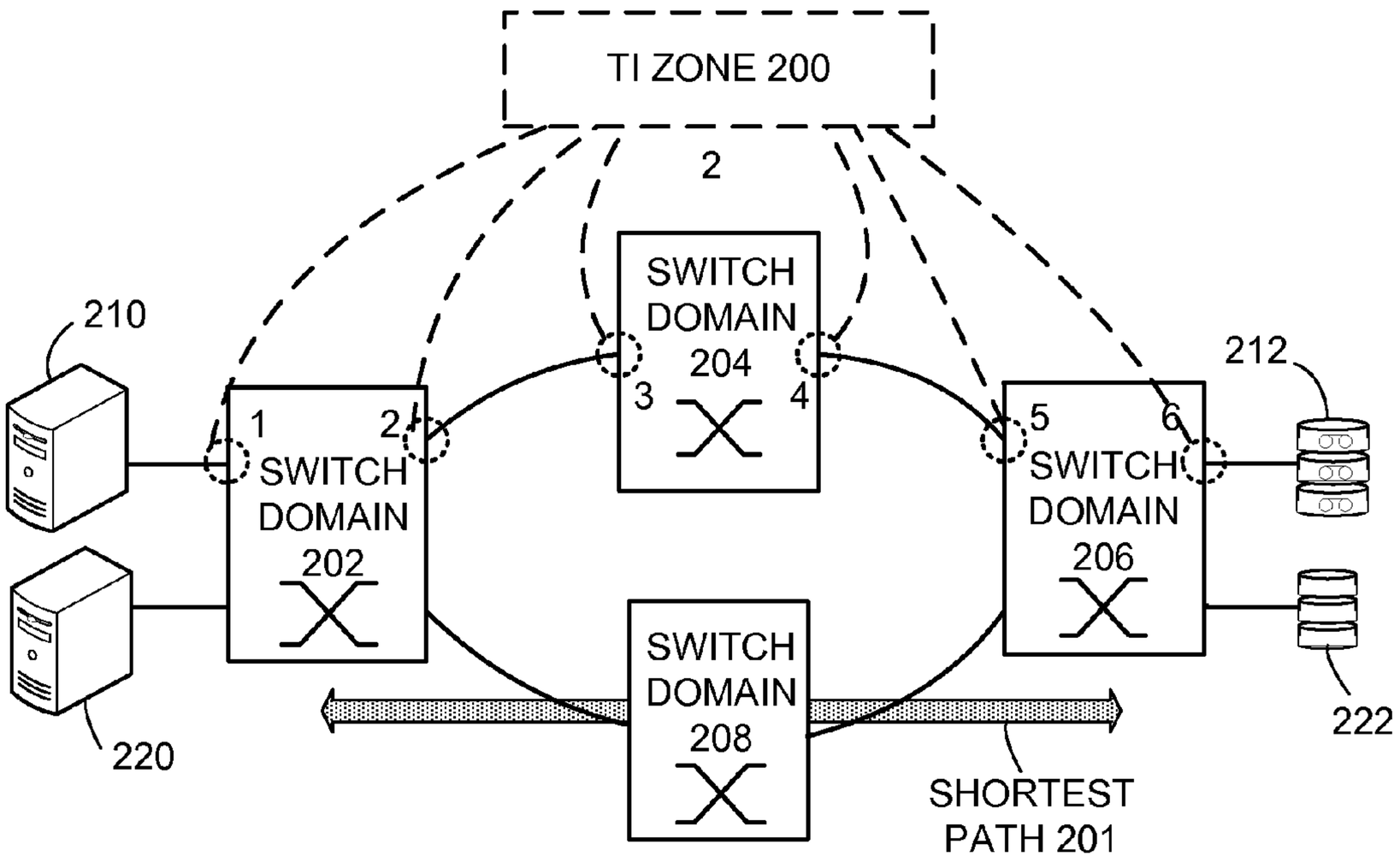

TI ZONE 200
2
210
SWITCH DOMAIN 204
3
4
212
1
2
SWITCH DOMAIN 202
5
6
SWITCH DOMAIN 206
SWITCH DOMAIN 208
220
222
SHORTEST PATH 201

FIG. 2B

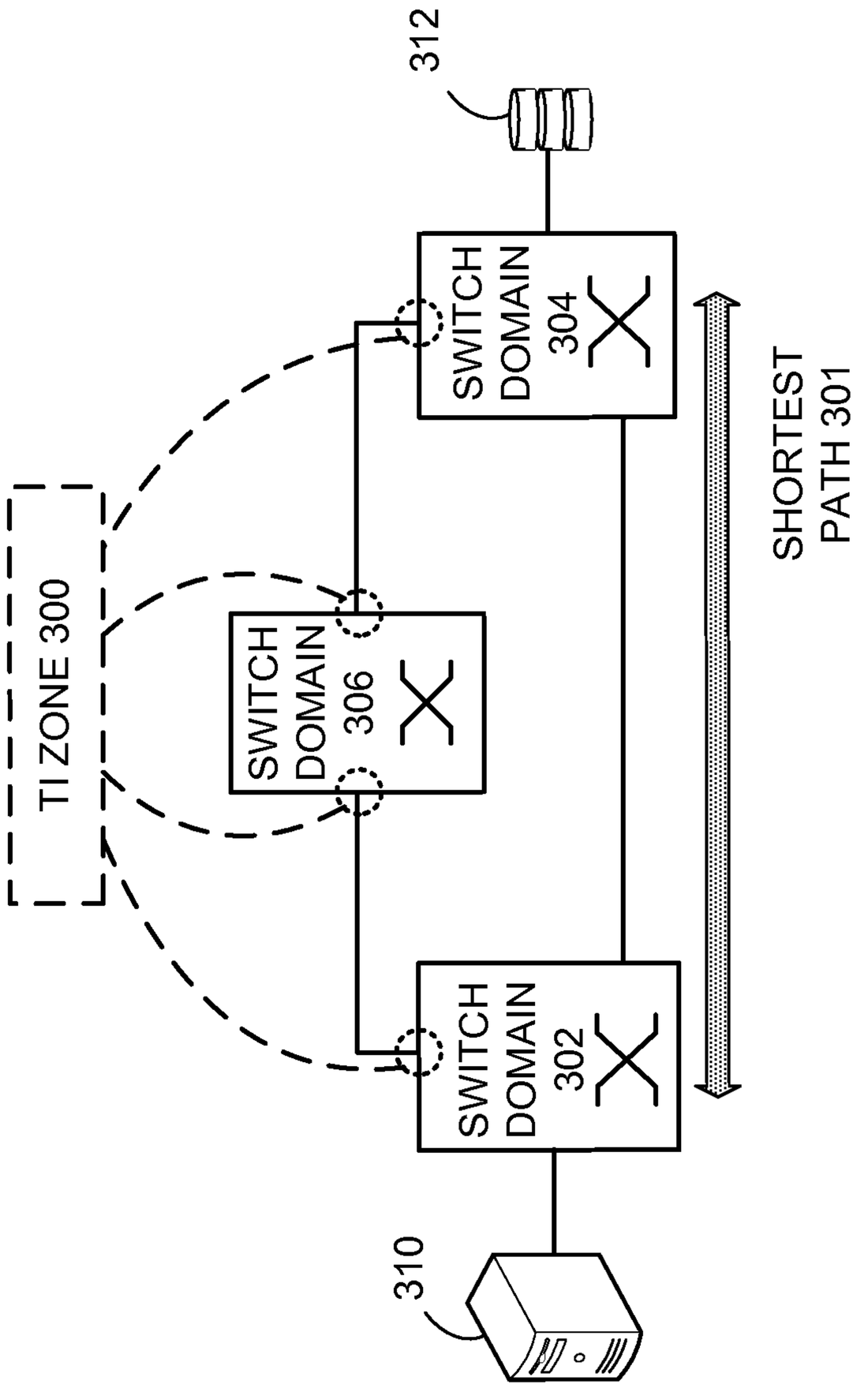

312
SWITCH DOMAIN 304
SHORTEST PATH 301
TI ZONE 300
SWITCH DOMAIN 306
SWITCH DOMAIN 302
310

FIG. 3

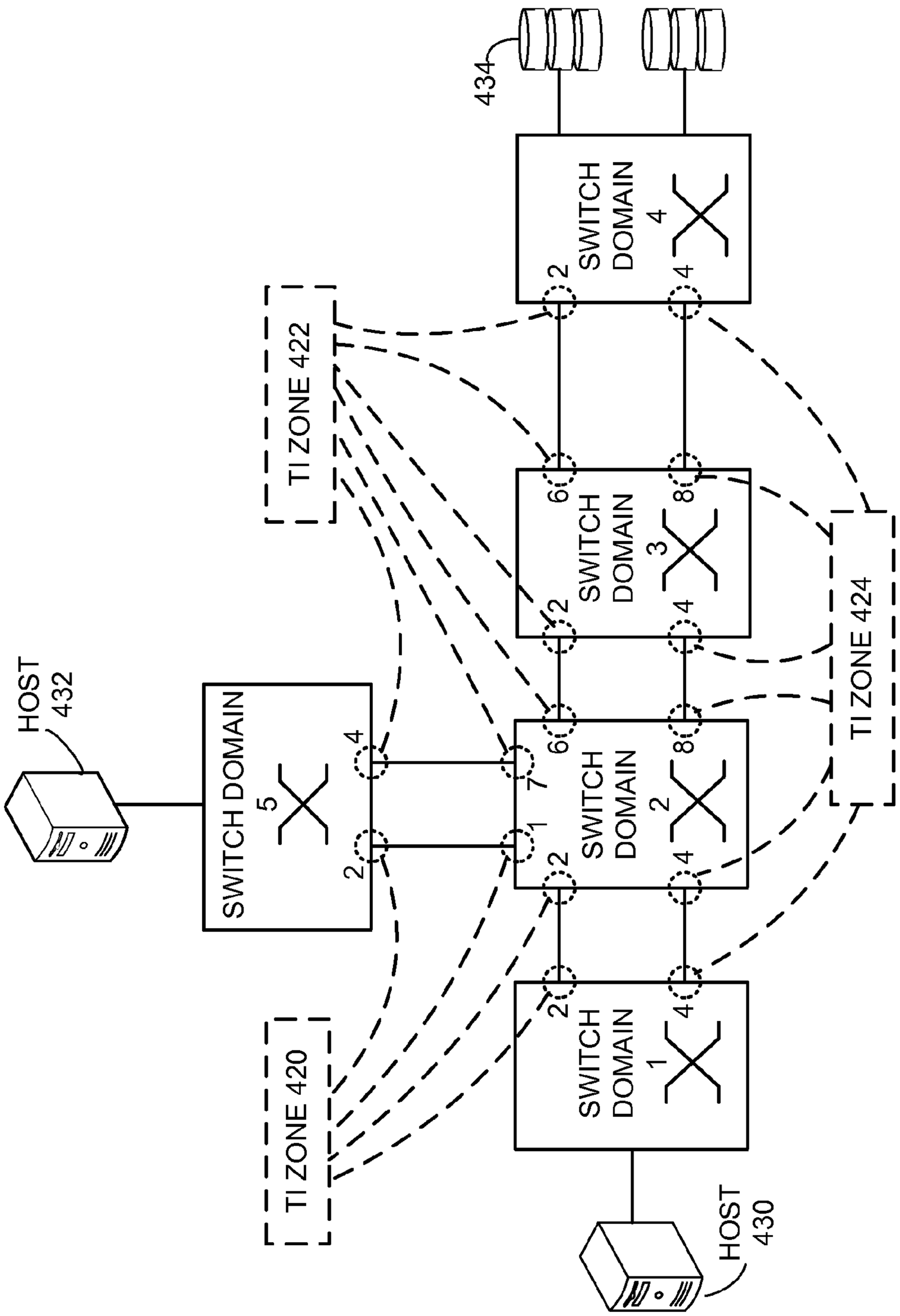

434
SWITCH DOMAIN 4
2
4
TI ZONE 422
6i
8i
SWITCH DOMAIN 3
2
4
HOST 432
SWITCH DOMAIN 5
4
2
6i
8i
SWITCH DOMAIN 2
7
1
2
4
TI ZONE 424
2
4
SWITCH DOMAIN 1
TI ZONE 420
HOST 430

FIG. 4

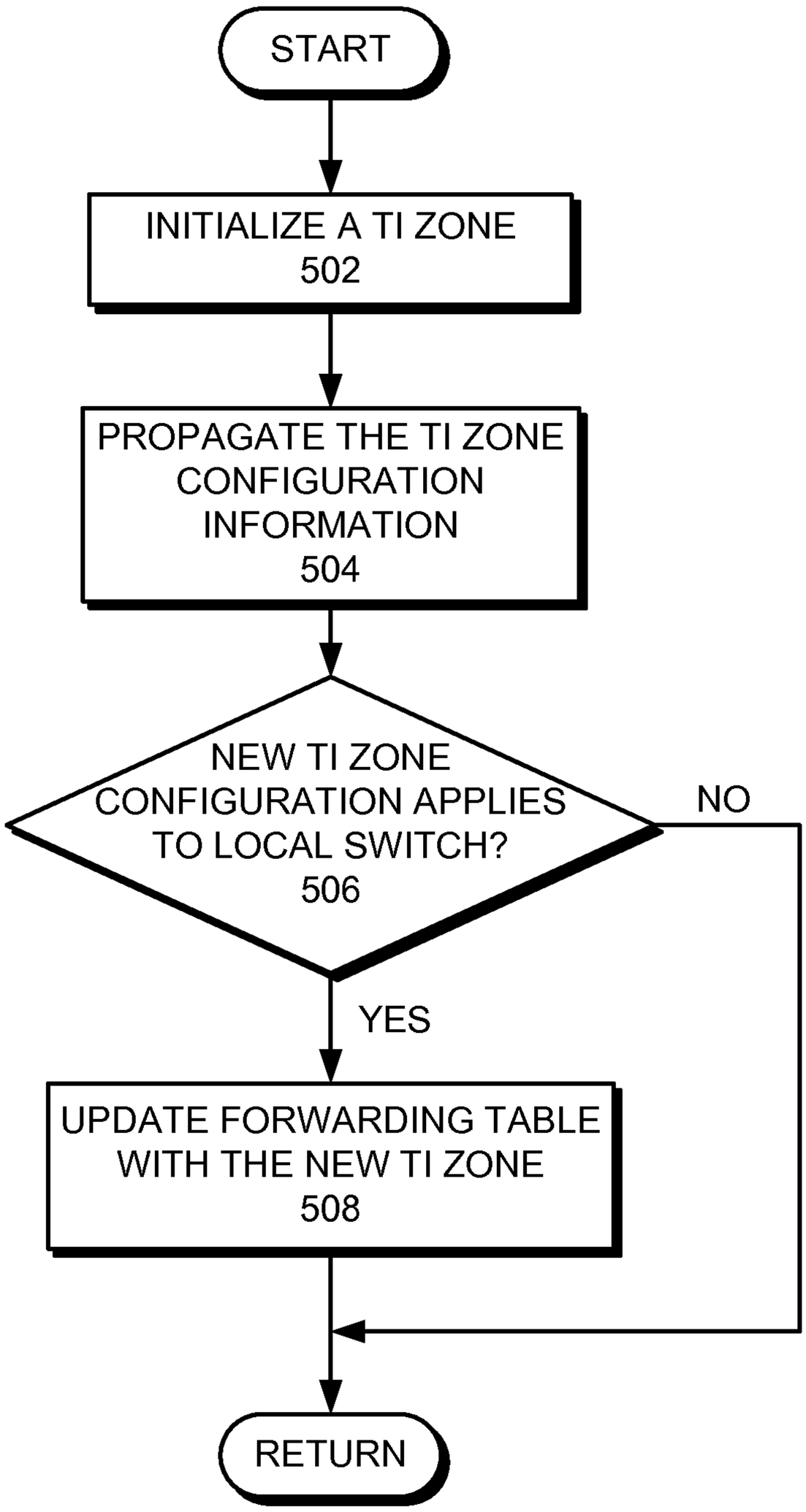

START
INITIALIZE A TI ZONE
502
PROPAGATE THE TI ZONE
CONFIGURATION
INFORMATION
504
NEW TI ZONE
CONFIGURATION APPLIES
TO LOCAL SWITCH?
506
NO
YES
UPDATE FORWARDING TABLE
WITH THE NEW TI ZONE
508
RETURN

FIG. 5

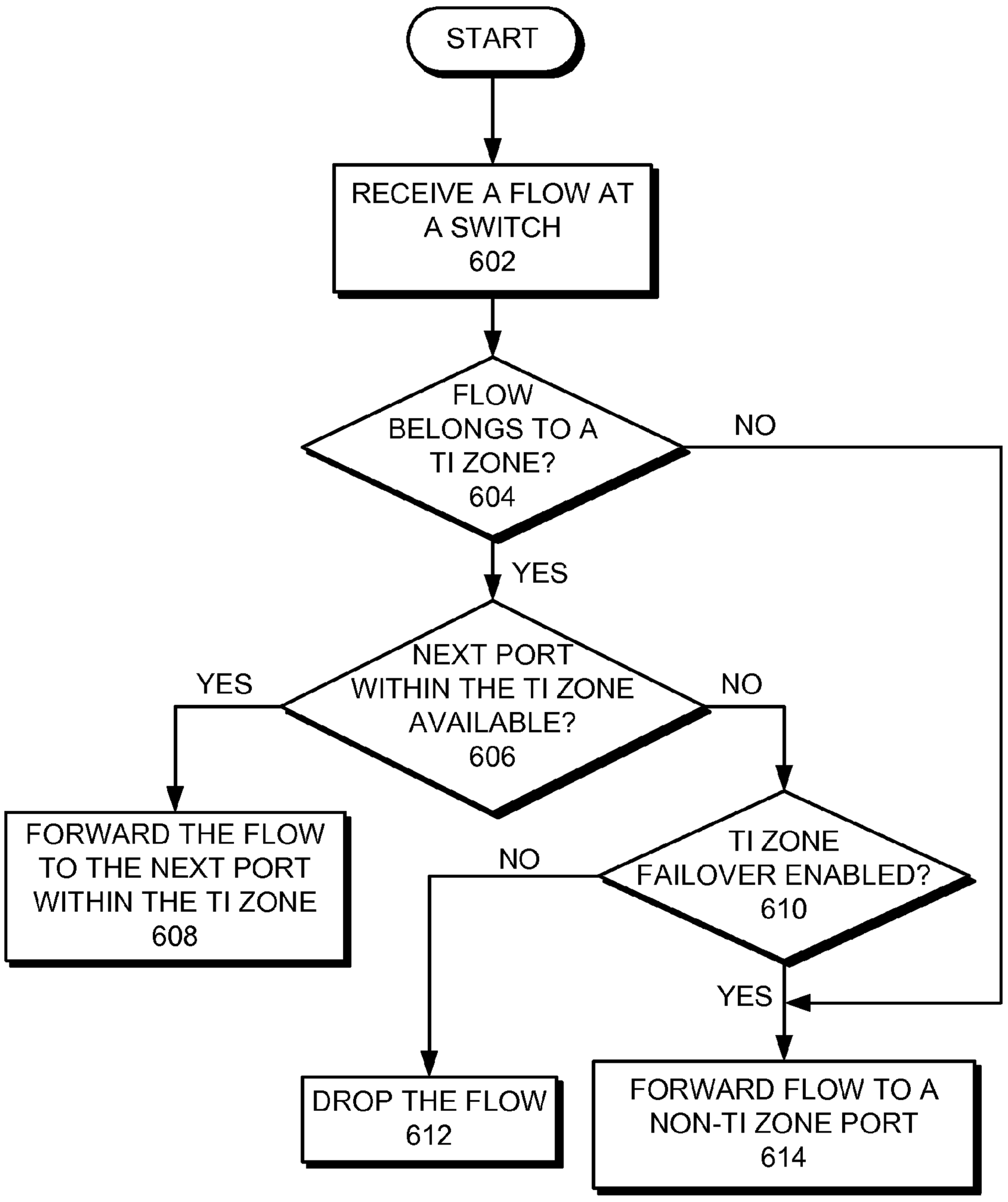

START
RECEIVE A FLOW AT A SWITCH 602
FLOW BELONGS TO A TI ZONE? 604
NO
YES
NEXT PORT WITHIN THE TI ZONE AVAILABLE? 606
YES
NO
FORWARD THE FLOW TO THE NEXT PORT WITHIN THE TI ZONE 608
TI ZONE FAILOVER ENABLED? 610
NO
YES
DROP THE FLOW 612
FORWARD FLOW TO A NON-TI ZONE PORT 614

FIG. 6

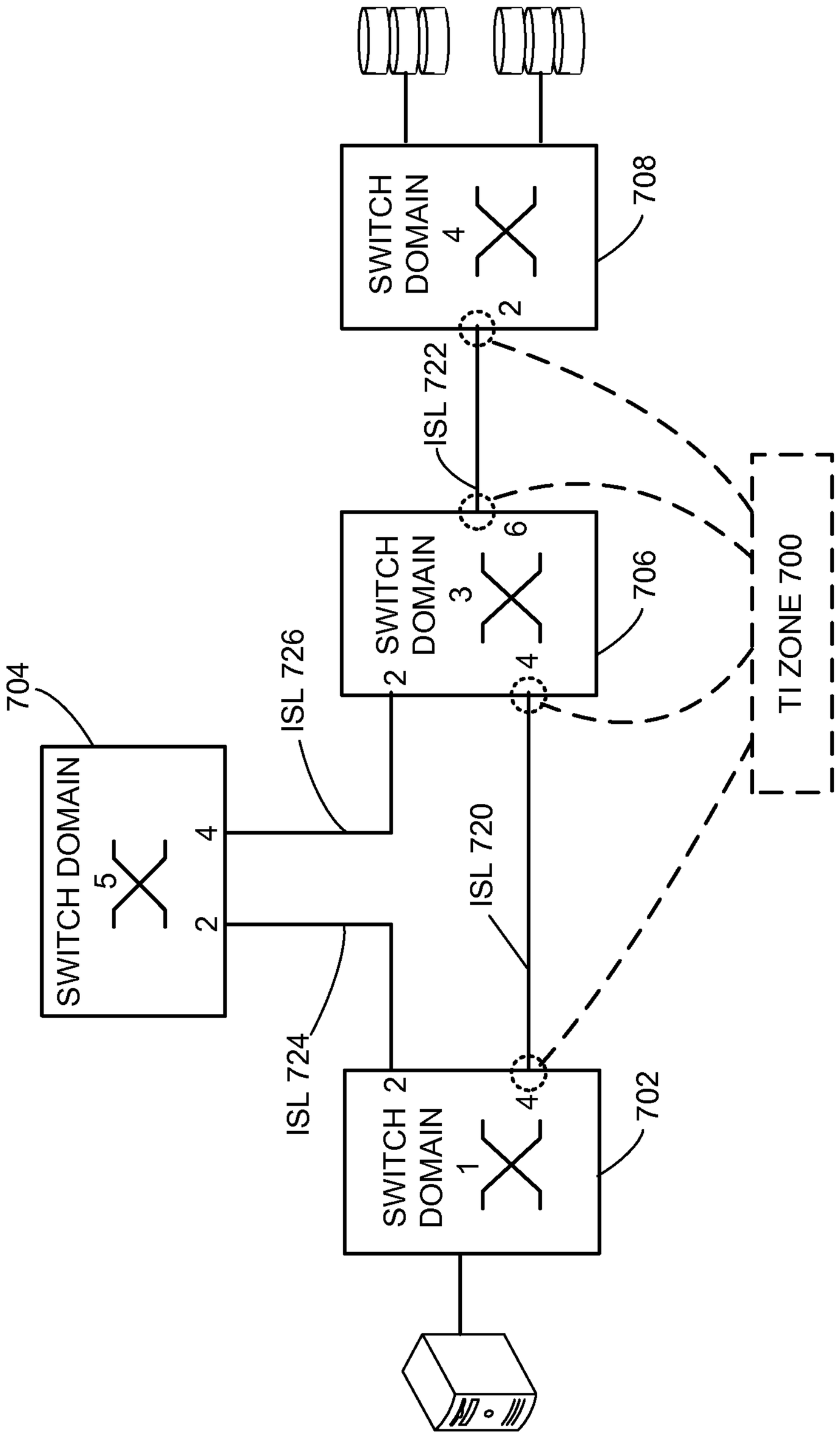

SWITCH DOMAIN 4
708
2
ISL 722
ISL 726
2 SWITCH DOMAIN 3
6
706
4
TI ZONE 700
704
SWITCH DOMAIN 5
4
2
ISL 720
724
2
SWITCH 2 DOMAIN 1
4
702

FIG. 7

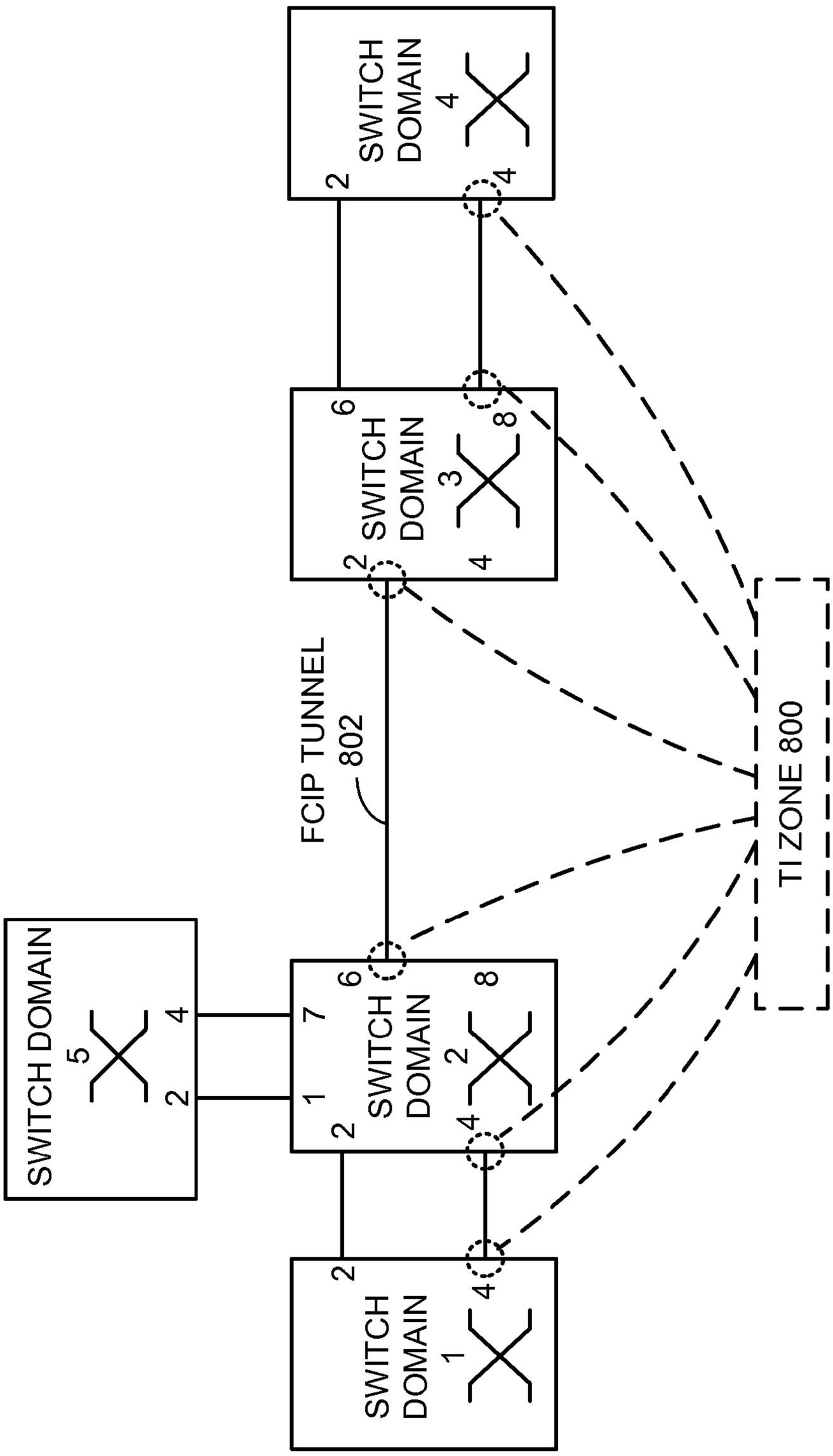

SWITCH DOMAIN
4
2
4
FCIP TUNNEL
802
SWITCH DOMAIN
6
3
8
2
4
TI ZONE 800
SWITCH DOMAIN
5
4
2
SWITCH DOMAIN
7
6
1
2
SWITCH DOMAIN
2
8
4
SWITCH DOMAIN
1
2
4

FIG. 8

METHOD AND SYSTEM FOR TRAFFIC ISOLATION IN A NETWORK

BACKGROUND

1. Field

The present disclosure relates to network management. More specifically, the present disclosure relates to a method and system for facilitating traffic isolation (TI) zones in a network.

2. Related Art

The proliferation of the Internet and e-commerce continues to fuel revolutionary changes in the network industry. Today, a significant number of transactions, from real-time stock trades to retail sales, auction bids, and credit-card payments, are conducted online. Consequently, many enterprises rely on existing storage area networks (SANs), not only to perform conventional storage functions such as data backup, but also to carry out an increasing number of egalitarian network functions such as building large server farms.

Historically, conventional network appliances (e.g., data-center servers, disk arrays, backup tape drives) mainly use a SAN network to transfer large blocks of data. Therefore, the switches provide only basic patch-panel-like functions. In the past decade, however, drastic advances occurred in almost all the network layers, ranging from physical transmission media, computer hardware and architecture, to operating system (OS) and application software.

For example, a single-wavelength channel in an optical fiber can provide 10 Gbps of transmission capacity. With wavelength-division-multiplexing (WDM) technology, a single strand of fiber can provide 40, 80, or 160 Gbps aggregate capacity. Meanwhile, computer hardware is becoming progressively cheaper and faster. Expensive high-end servers can now be readily replaced by a farm of many smaller, cheaper, and equally fast computers. In addition, OS technologies, such as virtual machines, have unleashed the power of fast hardware and provide an unprecedented versatile computing environment.

As a result of these technological advances, a conventional SAN switch fabric faces a much more heterogeneous, versatile, and dynamic environment. The limited network functions in such switches can hardly meet these demands. For instance, applications that require high bandwidth, such as backup applications, may cause congestion in the fabric and impact other critical application performance. Hence, there is a need for options to segregate some applications from others and to assign dedicated paths to certain applications.

SUMMARY

One embodiment of the present invention provides a system that facilitates traffic isolation (TI) in a network. During operation, the system configures a set of switch ports as members of a TI zone. The switch ports are part of an end-to-end path across one or more switch domains between a source and a destination. The switch ports within the TI zone and outside the TI Zone belong to a common storage area network (SAN) zone which compartmentalizes data for security purposes. The system then determines whether a data flow entering a switch domain belongs to the TI zone. The system subsequently forwards the data flow to the next-hop port within the TI zone if the data flow belongs to the TI zone. When the data flow does not belong to the TI zone, the system prevents the data flow from reaching a port within the TI zone.

In a variation on this embodiment, the system uses the TI zone to enforce one or more Quality-of-Service (QoS) parameters for a QoS class

In a variation on this embodiment, the system configures the switch ports as members of the TI zone by first propagating the TI zone configuration across one or more switch domains. The system then determines whether a local switch domain is part of the TI zone. The system subsequently updates a local forwarding table with the TI zone configuration.

In a variation on this embodiment, the system forwards a data flow which does not belong to the TI zone to a port that belongs to the TI-zone, if no port outside the TI zone is available to forward this data flow to its destination.

In a variation on this embodiment, the system forwards a data flow which belongs to the TI zone to a port outside the TI zone, if the next-hop port within the TI zone for this data flow is not available.

In a variation on this embodiment, the system drops a data flow which belongs to the TI zone, if the next port within the TI zone for this data flow is not available.

In a variation on this embodiment, the switch ports configured as part of the TI zone are Fibre Channel ports and comprise one or more of N_Ports, E_Ports, and EX_Ports.

In a further variation, the TI zone comprises a set of inter-switch links (ISLs) coupling adjacent E_Ports that form the end-to-end path.

In a further variation, the TI zone comprises one or more EX_Ports and is configured to traverse fibre channel routers (FCRs) and FCR-coupled fabrics.

BRIEF DESCRIPTION OF THE FIGURES

FIG. **1** illustrates an exemplary FC network that facilitates traffic isolation, in accordance with an embodiment of the present invention.

FIG. **2** illustrates an exemplary use of a TI zone, in accordance with an embodiment of the present invention.

FIG. **3** illustrates another exemplary use of a TI zone, in accordance with an embodiment of the present invention.

FIG. **4** illustrates exemplary configurations of TI zones comprising of E_Ports, in accordance with an embodiment of the present invention.

FIG. **5** presents a flowchart illustrating the process of enabling a TI zone configuration, in accordance with an embodiment of the present invention.

FIG. **6** presents a flowchart illustrating the process of forwarding a flow at a switch, in accordance with an embodiment of the present invention.

FIG. **7** illustrates how a TI zone with a failover option enabled switches over to an alternate path, in accordance with an embodiment of the present invention.

FIG. **8** illustrates exemplary configurations of TI zones comprising of VE_Ports coupled by a FCIP tunnel, in accordance with an embodiment of the present invention.

DETAILED DESCRIPTION

The following description is presented to enable any person skilled in the art to make and use the invention, and is provided in the context of a particular application and its requirements. Various modifications to the disclosed embodiments will be readily apparent to those skilled in the art, and the general principles defined herein may be applied to other embodiments and applications without departing from the spirit and scope of the present invention. Thus, the present

invention is not limited to the embodiments shown, but is to be accorded the widest scope consistent with the claims.

The data structures and code described in this detailed description are typically stored on a computer-readable storage medium, which may be any device or medium that can store code and/or data for use by a computer system. This includes, but is not limited to, application-specific integrated circuits (ASIC), field-programmable gate arrays (FPGA), volatile memory, non-volatile memory, magnetic and optical storage, or other media capable of storing computer-readable media now known or later developed.

Overview

Embodiments of the present invention facilitate traffic isolation (TI) zones in a Fibre Channel (FC) network. TI zones allow users to assign a data flow to specific inter-switch links (ISLs) to isolate traffic within a switch fabric. This configuration provides the option to segregate some applications from others with dedicated, separate paths through the switch fabric. For example, backup applications often require guaranteed high bandwidth, and assigning them to dedicated paths within a TI zone helps reduce congestion and prevent them from impacting other critical application performance. TI zones can also be used to dedicate inter-switch links (ISLs) to high-priority data flows, and control the route for inter-switch traffic. Such configuration can facilitate enforcement of different Quality-of-Service (QoS) classes. For example, a TI zone can be dedicated to a higher-priority QoS classes, so that lower-priority traffic is precluded from consuming the bandwidth allocated to the higher-priority QoS class.

A TI zone can include a set of switch ports used for specific traffic flows. When a TI zone is enforced, a data flow entering a starting switch port is forwarded to a next-hop port within the same TI zone. Under normal conditions, traffic outside the TI zone is precluded from entering the TI zone. Hence, TI zones can effectively facilitate traffic isolation. In come cases, if the next-hop port cannot be reached due to link failure, the data flow can be forwarded to a port on an alternative path outside the zone. This configuration is possible when the failover option is enabled. In case the failover option is disabled, data flows belonging to the zone can be optionally dropped if there is a link failure in the TI zone.

Network Architecture

The heterogeneous nature of modern FC networks imposes many new challenges, among which traffic management deals with controlling and allocating network bandwidth and minimizing congestion at switch ports. Embodiments of the present invention facilitate TI zones as one of the traffic management services, which allocate specific ISLs to data flows. This ensures that a data flow belonging to a particular zone is isolated from traffic from other zones and can enjoy dedicated network resources, which is valuable especially for high-bandwidth traffic.

FIG. **1** illustrates an exemplary FC network to which traffic isolation zones can be applied, in accordance with an embodiment of the present invention. In this example, an FC switch fabric **100** includes four switch modules, **102**, **104**, **106**, and **108**. Each switch module is coupled to a group of network appliances. For example, switch module **102** is coupled to a number of servers **110** and a number of disk arrays **112**.

A respective network appliance can communicate with any appliance (referred to as "target") in the FC network. For example, one of the servers **110** can transfer data to and from one of tape backup devices **116**. Note that, since the switch modules are not coupled in a fully meshed topology, the data frames transferred between servers **110** and tape devices **116** traverse three switch modules **102**, **104**, and **106**. In general, the switch modules are coupled by ISLs, such as ISL **114**. In one embodiment, a network operator can specify and enforce a TI zone, for example, a dedicated switched path coupling server **110** and tape devices **116**, through the switch fabric. This TI zone includes a number switch ports and ISLs. Traffic outside the TI zone is precluded from entering the TI zone. This way, the network operator can guarantee the service quality between server **110** and tape devices **116**.

As shown in FIG. **1**, large-port-count FC switch fabrics often include a number of smaller, interconnected individual switches. The internal connectivity of a switch fabric can be based on a variety of topologies. In this disclosure, the term "switch fabric" refers to a number of inter-coupled FC switch modules. The terms "switch module" and "switch" refer to an individual switch which can be coupled to other switch modules to form a larger port-count switch fabric. The term "edge device" refers to any network appliance, either physical or logical, coupled to a switch. The term "switch domain" refers to a unique identifier for each switch and creates an address for each device coupled to the switch.

A switch typically has two types of ports: a fabric port (denoted as F_Port), which can couple to a network appliance, and an extension port (E_Port), which can couple to another switch. A network appliance communicates with a switch through a host bus adapter (HBA). The HBA provides the interface between an appliance's internal bus architecture and the external FC network. An HBA has at least one node port (N_Port), which couples to an F_Port on a switch through an optical transceiver and a fiber optic link. More details on FC network architecture, protocols, naming/address convention, and various standards are available in the documentation available from the NCITS/ANSI T11 committee (www.t11.org) and publicly available literature, such as "Designing Storage Area Networks," by Tom Clark, 2nd Ed., Addison Wesley, 2003, the disclosure of which is incorporated by reference in its entirety herein.

FIG. **2**A and FIG. **2**B illustrates an exemplary use of a TI zone. Switch domain **202** is coupled to servers **210** and **220**, and switch domain **206** is coupled to backup tape devices **212** and disk array **222**. Assume that the shortest path between switch domain **202** and switch domain **206** is through switch domain **208**, while there exists an alternate path between switch domain **202** and switch domain **206** through switch domain **204**. Assume further that data flows between servers **210** and **220** and backup tape devices **212** and disk array **222** follow a shortest path **201** across switch domains **202**, **208**, and **206**.

In the network shown in FIG. **2**A, shortest path **201** is shared by all applications that may communicate between switch domain **202** and **206**. Hence, when a backup application initiates a sustained high-bandwidth communication session between server **210** and backup tape devices **212**, other applications such as the communication between server **220** and disk array **222** will experience high latency and low performance on the same path **201**.

The application of TI zone provides a graceful solution to this problem in accordance with embodiments of the present invention. As illustrated in FIG. **2**B, a TI zone **200** is created to include ports 1 and 2 on switch domain **202**, ports 3 and 4 on switch domain **204**, and ports 5 and 6 on switch domain **206**. By assigning the backup application between server **210** and backup tape devices **212** to TI zone **200**, the high-bandwidth backup data flow is isolated on a dedicated path from switch domains **202**, through switch domain **204**, to switch domain **206**. Other applications between server **220** and disk array **222** remain on shortest path **201** and are not affected by the sustained backup application.

FIG. **3** illustrates another use for TI zones, where a data flow belongs to a TI zone **300** between a server **310** which is coupled to a switch domain **302**, and a disk array **312** which is coupled to a switch domain **304** via switch domain **306**. This specific path is different from a shortest path **301** and ensures the bidirectional data flow takes the exact same route in both directions. In general, embodiments of the present invention provide a method that controls the ports and path used when routing traffic between fabrics. By isolating traffic, the TI zone can also prevent high-bandwidth traffic from causing congestion and improves fabric utilization.

Traffic Isolation Zones

In accordance with embodiments of the present invention, TI zones allow network administrators to provision a certain set of E_Ports on one or more switches to carry only designated data flows. TI zones assign specific paths to the data flows belonging to the zone, and control the route for inter-switch traffic. TI zones can provide dedicated ISLs to high-priority data flows. TI zones can also be used to force high-volume (but lower priority) data flows onto specific ISLs to limit the impact on other critical applications in the switch fabric. In either case, a TI zone can include a set of switch ports used for specific traffic flows.

When a TI zone is created, a data flow entering a switch from the starting N_Port or E_Port is forwarded to the next E_Port within the zone. If the next E_Port within the TI zone is not available (e.g. due to network congestion or link failure), the data flow can be forwarded to its destination using an E_Port outside the zone if the failover option enabled. In a TI zone with the failover option disabled, when any of the E_Ports within the TI zone goes down, the TI zone is deemed unavailable and data flows belonging to the zone can be optionally dropped.

In one embodiment, an E_Port belonging to a particular TI zone may not carry any other data flows outside the zone, unless that E_Port is the only way to reach the destination and the failover option is enabled. If the failover option is disabled, that E_Port is precluded from carrying any other data flows outside the zone under any circumstance.

In some embodiments, a TI zone can be provisioned to carry traffic of a given QoS class. Dedicating the TI zone to one or more QoS classes can guarantee a certain amount of bandwidth, a minimum delay, and a minimum packet loss rate. In general, a number of QoS parameters, such as bandwidth, end-to-end delay, and in-order packet delivery, can be enforced within a TI zone. In addition, more than one TI zones may be provisioned for one QoS class. In case one TI zone fails (e.g., due to port or link failure), another TI zone can be used the carry the protected traffic.

Ideally, a few general rules may be followed when TI zones are configured. For example, to limit the management overhead, a maximum number (e.g., 255) of TI zones can be created in one switch fabric. Second, a port configured to be in a TI zone may not be a member of other TI zones. In other words, a given port can only be a member of a single TI zone because the port can only be on a single path to any specific domain to ensure successful traffic isolation. This "non-duplication" rule is enforced during zone creation/modification. Lastly, to use a trunk port for a TI zone, all the ports in the trunk group become part of the same TI zone. Mixing different TI zones in a trunk group or configuring only a subset of the ports in the trunk group might result in unpredictable behavior.

In one embodiment, a TI zone can be created for a switch using command-line interface (CLI) with options and a port list of (domain, index) format, where domain is the switch domain ID and index is the port number. The port world-wide name (WWN) can also be used for TI over Fibre Channel router (FCR) support. Below is an example of the command:

zone --create -t ti "redzone" -p "1,1; 2,4; 2,6; 3,8"

The command zone takes the operations such as create, add, remove, delete, activate, deactivate and show. The object type ti after the -t option specifies that the zone is of the traffic isolation type, and redzone is the name of the TI zone configured. The TI zone is identified by a port list (1,1; 2,4; 2,6; 3,8), which comprises port 1 of switch domain 1, port 4 and port 6 of switch domain 2, and port 8 of switch domain 3.

Embodiments of the present invention provide the following CLI syntax to manage TI zones:

Synopsis:

zone --operation -t objtype [-o optionlist] name -p portlist

operation ::=create, add, remove, delete, activate, deactivate or show

objtype ::=ti (traffic isolation zone)

optionlist ::=a (activate), d (deactivate), n (no-failover), f (failover)

portlist :: =D,I (Domain, Index)

TI Zone Create

zone --create -t objtype [-o optionlist] name portlist -p portlist

Create a TI Zone with specified options and the portlist. By default, the zone is created with failover enabled and the zone will be activated.

Examples

Create a Traffic Isolation Zone with failover enabled and activate the zone. zone --create -t ti "redzone" -p "1,1; 2,4; 1,8; 2,6"

Create a Traffic Isolation Zone with failover disabled and deactivated. zone --create -t ti -o dn "redzone" -p "1,1; 2,4; 1,8; 2,6"

TI Zone Add/Remove Members and Options

zone --add [-o optionlist] name portlist -p portlist <<use without portlist to add options>>

zone --remove name portlist -p portlist <<use without portlist to remove options>>

Zone --add command allows users to add portlist members and failover option to an existing TI zone. Zone --remove command allows user to remove portlist members from existing zones. If the last member of a TI zone is removed, the TI zone name will be removed from the defined TI zone lists.

Examples

Add port members to an existing TI zone.

zone --add "redzone" -p "3,4; 3,6"

Add option to disable/enable failover for a TI zone.

zone --add -o n "redzone"

zone --add -o f "greenzone" -p "3,4"

Remove portlist member from an existing TI zone.

zone -remove "bluezone" -p "3,4; 3,6"

TI Zone Activate/Deactivate

zone --activate name.....[name]

zone --deactivate name.....[name]

Zone activate/deactivate command allows user to activate/deactivate TI zone.

Examples

zone --activate redzone

zone --deactivate bluezone

TI Zone Deletion

zone --delete name.....[name]

Zone --delete command will delete TI zones from the defined TI zone lists completely. Users will be prompted to confirm delete action.

Examples

zone --delete bluezone
TI Zone Show
zone --show

The zone --show command without any specified name will display all the TI zones in defined configuration. Zone --show command can be executed for one zone. This command will display the zone name, portlists, failover option and status.

Examples

zone --show
Zone Name: green_zone:
List of port: 2,2; 3,3; 5,3; 4,11;
Failover: Enabled
Status: Activated
Zone Name: blue_zone:
List of port: 1,2; 1,3; 3,3; 4,5;
Failover: Enabled
Status: Activated
Zone Name: red_zone:
List of port: 9,2; 9,3; 8,3; 8,5;
Failover: Disabled
Status: Deactivated
zone --show blue_zone
Zone Name: blue_zone:
List of port: 1,2; 1,3; 3,3; 4,5
Failover: Enabled
Status: Activated

FIG. **4** illustrates an exemplary switch fabric comprising five switch domains 1 to 5, in which three TI zones are created. TI zone **420** facilitates communication between a host **430** and a host **432** though a list of ports (1,2; 2,2; 2,1; 5,2) across switch domains 1, 2, and 5. TI zone **422** includes a list of ports (5,4; 2, 7; 2,6; 3,2; 3,6; 4,2), which are part of a path that traverses switch domains 5, 2, 3, and 4. Applications accessing a disk array **434** from host **432** may send and receive data flows on TI zone **422**. The third TI zone **424** configured in this example provides an end-to-end dedicated path between host **430** and disk array **434** across switch domains 1, 2, 3, and 4. The port list for TI zone **424** is (1,4; 2,4; 2,8; 3,4; 3,8; 4,4). Note that all three TI zones shown in FIG. **4** include only E_Ports. The set of ISLs between the E_Ports form an end-to-end path from the ingress switch domain to the egress switch domain.

For example, TI zone **424** creates a dedicated path from switch domains 1 to 4 through the core switch domains 2 and 3. All data flows belonging to TI zone **424** and entering domain 1 from host **430** will be forwarded to port 4 on switch domain 1. Other data flows from host **430** outside this zone will be routed to port 2 regardless of their destination. Similarly, any traffic entering switch domain 2 on port 2 will be routed to port 6 when heading for switch domain 3 or domain 4 because port 4 and port 8 on switch domain 2 are dedicated to TI zone **424**.

FIG. **5** presents a flowchart illustrating the process of enabling a TI zone after it is created, in accordance with embodiments of the present invention. During operation, the system first initializes a TI zone for a number of switches (operation **502**). The system then propagates the TI zone configuration information to all involved switches (operation **504**). Next, the system determines if the new TI zone configuration applies to a local switch (operation **506**). If so, the forwarding table of the local switch will be updated with the new TI zone information (operation **508**).

A TI zone configuration is interpreted by the local switch and the switch only considers the forwarding update required for its local ports. In one embodiment, a switch is not required to determine whether the TI zones accurately provide dedicated end-to-end paths through the entire switch fabric. This allows routing to be determined at the time TI zones are activated, eliminating a significant amount of overhead that would be required to dynamically route data flows belonging to TI zones in real-time.

In one embodiment, TI zones provide traffic isolation in a "best effort" fashion that works as long as the "lowest-cost path" rule holds in fabric shortest-path first (FSPF) routing. In other words, FSPF routing rules can take precedence over the TI zones. This means that data flows from one TI zone may have to share E_Ports with other data flows when no equal-cost paths can be found. Furthermore, when an E_Port in the preferred TI zone fails, data flows belonging to that TI zone will be switched to a failover path that is the next lowest-cost path to the destination. Similarly, a data flow outside the zone will use an E_Port from this TI zone if no alternative path exists.

In some embodiments, the following rules apply if the TI zone is not the lowest-cost path: when the TI zone path is broken, data flows belonging to this TI zone will switch over to the lowest-cost path which is not part of the TI zone if the failover option is enabled. If the failover option is disabled, the data flows belonging to this TI zone will be blocked. The following rules apply if the TI zone is the only lowest-cost path: if the failover option is enabled, non-TI-zone as well as TI-zone data flows can use the dedicated TI-zone path. If the failover option is disabled, the non-TI zone data flows will be blocked.

FIG. **6** presents a flowchart illustrating the process of making decisions on how to forward a data flow at a switch, in accordance with an embodiment of the present invention. During operation, the system first receives a data flow (operation **602**). The system then determines whether the flow belongs to a TI zone (operation **604**). If not, the data flow is forwarded to a non-TI zone port (operation **614**). If the flow belongs to a TI zone, the system further determines whether the next-hop port within the TI zone is available (operation **606**). If the next-hop port on the TI zone is available, the data flow is forwarded to that port (operation **608**). If the port within that TI zone is not available, the system then decides whether the TI zone failover option is enabled (operation **610**). If so, the data flow is forwarded to a non-TI zone port on the failover path (operation **614**). Otherwise, the data flow is dropped (operation **612**).

FIG. **7** illustrates how a TI zone with the failover option enabled switches over to an alternate path, in accordance with an embodiment of the present invention. Consider a TI zone **700** that traverses an ISL **720** and an ISL **722** in the example in FIG. **7**. When ISL **720** goes offline, data flows belonging to TI zone **700** will automatically switch over to the failover path that spans ISL **724** and ISL **726** between switch domain 1 (**702**) and switch domain 3 (**706**) through switch domain 5 (**704**). If the failover option is disabled, however, TI zone **700** will be deemed unavailable and data flows belonging to the zone are dropped. On the other hand, if a TI Zone is the only path to reach another domain, non-TI zone traffic may be forwarded on the TI zone path as well. For example, in FIG. **7**, if ISL **724** is offline, all the traffic from switch domain 1 (**702**) to switch domain 4 (**708**) will be forwarded via ISL **720**, regardless of whether the traffic belongs to TI zone **700**.

Hence, during the configuration of a TI zone with the failover option disabled, special cautions need to be taken to avoid segmenting a subset of switches from the fabric. Segmentation happens when there is no route to reach switch ports outside the zone from switch ports within the zone. Before deployment of a TI zone with failover disabled, the topology is ideally reviewed to ensure that all switches in the fabric have a path to reach other switches in the fabric. A partially configured TI zone without a full path between a source and a destination might result in fabric-wide operation failure. Recovery from the failure could be difficult since operators may need to take switches offline one by one to locate the broken ISL.

Note that the TI zone in FC networks should be distinguished from the general SAN zoning, which is a method of arranging Fibre Channel devices into logical groups within the fabric. In general SAN zoning, each device may be placed into multiple zones to achieve compartmentalization of data for security purposes. For instance, by dividing up device ports into groups, data access can be limited for some users to specific groups of servers that store confidential data in a SAN network. Traffic within one SAN zone can be strictly prohibited from entering another SAN zone, even when one SAN zone fails. In contrast, traffic within a TI zone is allowed to leave the TI zone under certain circumstances, for example when there is a failure in the TI zone and when the failover option is enabled. Furthermore, traffic outside a TI zone might also be allowed to enter the TI zone under similar circumstances. In general, the designation of TI zones takes place within a SAN zone. In other words, a TI zone and the corresponding non-TI zones all belong to the same SAN zone. A more detailed discussion on general SAN zoning can be found in "Designing Storage Area Networks," by Tom Clark, 2nd Ed., Addison Wesley, 2003.

In contrast, a TI zone isolates data flows over different ISLs to control the ports and path used when routing traffic between fabrics. By assigning specific paths for data flows, TI zones also minimize congestion and improve fabric utilization. General SAN zoning, on the other hand, does not change a data flow's routing; it only partitions the SAN into logical groups and enforces access security between the groups.

TI Zones Over FCR

The traffic isolation feature introduced in the previous section provides capability to isolate traffic between N_Ports or E_Ports across a particular path defined within layer-two fabrics. However, TI zones can also be extended to traverse Fibre Channel routers (FCRs) and FCR-coupled fabrics. One embodiment of the present invention allows a TI zone comprising one or more EX_Ports. EX_Ports couple a Fibre Channel router to a Fibre Channel switch. On the switch side the port looks like a normal E_Port, but on the router side the port is an EX_Port.

A TI zone over FCR typically has two portions: TI zones within the edge and TI zones within network backbone. TI zones within edge device include ports that couple switches to network appliances. TI zones within the network backbone are the portions including E_Ports and EX_Ports that intercouple switches. For example, TI zone **200** in FIG. **2** includes ports 1 and 6 within the edge devices and ports 2, 3, 4, and 5 within the network backbone. While the main focus of TI zones over FCR is in the backbone fabrics, this feature works in conjunction with TI zones within edge devices to achieve desired end-to-end network routing. In other words, a TI zone within edge is used to route traffic between end devices and proxy devices to a particular EX_Port, and a TI zone within network backbone is used to secure a dedicated path within the backbone fabrics.

In another embodiment, the dedicated path within the backbone may comprise a set of EX_Ports or a tunnel based on virtual E_Ports (VE_Ports) across one or more FCRs. FIG. **8** illustrates a TI zone configuration with a Fibre-Channel-over-IP (FCIP) tunnel. TI zone **800** in FIG. **8** includes a list of port (1,4; 2,4; 2,6; 3,2; 3,8; 4,4). A FCIP tunnel **802** couples VE_Port 6 of switch domain 2 to VE_Port 2 of switch domain 3. Hence traffic flows belonging to TI zone **800** are carried over FCIP tunnel **802** between switch domains 2 and 3.

In a summary, embodiments of the present invention facilitate TI zones in a network to manage dedicate route. A TI zone can include a set of switch ports used for specific traffic flows. A data flow that belongs to the TI zone is forwarded to a next-hop port within the zone. TI zones allow users to assign a data flow to specific inter-switch links (ISLs) to isolate traffic and reduce congestion within a switch fabric. TI zones can also be used to dedicate ISLs to high-priority data flows, and control the route for inter-switch traffic.

The foregoing descriptions of embodiments of the present invention have been presented only for purposes of illustration and description. They are not intended to be exhaustive or to limit this disclosure. Accordingly, many modifications and variations will be apparent to practitioners skilled in the art. The scope of the present invention is defined by the appended claims.

What is claimed is:

1. A method for facilitating traffic isolation in a switch, the method comprising:

configuring one or more ports of the switch as members of first a traffic isolation zone, wherein a traffic isolation zone is a subset of ports in the switch and precludes a data flow not belonging to the traffic isolation zone from reaching a port within the traffic isolation zone;

identifying that a data flow belongs to the first traffic isolation zone based on the subset of ports belonging to the first traffic isolation zone;

in response to a port of the first traffic isolation zone being available, associating the data flow with the port as an output port;

in response to the port not being available, determining whether a failover option for the first traffic isolation zone is enabled; and

in response to the failover option being enabled, associating the data flow with a second port as an output port, wherein the second port is a member of a second traffic isolation zone, and wherein the second traffic isolation zone is distinct from the first traffic isolation zone.

2. The method of claim **1**, further comprising:

using the first traffic isolation zone to enforce one or more Quality-of-Service (QoS) parameters for a QoS class.

3. The method of claim **1**, wherein configuring the ports as members of the first traffic isolation zone comprises:

propagating the configuration of the first traffic isolation zone across one or more switches;

determining whether a local switch is part of the first traffic isolation zone; and

updating a local forwarding table with the configuration of the first traffic isolation zone.

4. The method of claim **1**, wherein the method further comprises associating the data flow with the second port as an output port in response to a next-hop port within the first traffic isolation zone for the data flow not being available, wherein a next-hop port within the second traffic isolation zone for the data flow is available.

5. The method of claim 1, wherein the method further comprises dropping the data flow belonging to the first traffic isolation zone in response to the failover option not being enabled.

6. The method of claim 1, wherein the ports configured as members of the first traffic isolation zone are Fibre Channel ports and comprise one or more of N_Ports, E_Ports, and EX_Ports.

7. The method of claim 6, wherein the first traffic isolation zone comprises a set of inter-switch links (ISLs) coupling adjacent E_Ports that form an end-to-end path.

8. The method of claim 6, wherein the first traffic isolation zone comprises one or more EX_Ports and is configurable to traverse Fibre Channel routers (FCRs) and FCR-coupled fabrics.

9. A non-transitory computer-readable medium storing instructions which when executed by a computer cause the computer to perform a method for facilitating traffic isolation in a network, the method comprising:

configuring one or more ports of a switch as members of a first logical group, wherein a logical group is a subset of ports in the switch and precludes a data flow not belonging to the logical group from reaching a port within the logical group;

identifying that a packet belongs to the first logical group based on the subset of ports belonging to the first traffic isolation zone;

in response to a port of the first logical group being available, associating the packet with the port as an output port;

in response to the port not being available, determining whether a failover option for the first logical group is enabled; and

in response to the failover option being enabled, associating the packet with a second port as an output port, wherein the second port is a member of a second logical group, and wherein the second logical group is distinct from the first logical group.

10. The computer-readable medium of claim 9, wherein the method further comprises using the first logical group to enforce one or more Quality-of-Service (QoS) parameters for a QoS class.

11. The computer-readable medium of claim 9, wherein configuring the ports as members of the first logical group comprises:

propagating the configuration of the first logical group across one or more switches;

determining whether a local switch is part of the first logical group; and

updating a local forwarding table with the configuration of the first logical group.

12. The computer-readable medium of claim 9, wherein the method further comprises associating the packet with the second port as an output port in response to a next-hop port within the first logical group for the packet not being available, wherein a next-hop port within the second logical group for the data flow is available.

13. The computer-readable medium of claim 9, wherein the method further comprises dropping the packet belonging to the first logical group in response to the failover option not being enabled.

14. The computer-readable medium of claim 9, wherein the ports configured as members of the first logical group are Fibre Channel ports and comprise one or more of N_Ports, E_Ports, and EX_Ports.

15. The computer-readable medium of claim 14, wherein the first logical group comprises a set of inter-switch links (ISLs) coupling adjacent E_Ports that form an end-to-end path.

16. The computer-readable medium of claim 14, wherein the first logical group comprises one or more EX_Ports and is configurable to traverse Fibre Channel routers (FCRs) and FCR-coupled fabrics.

17. A switch for facilitating traffic isolation in a network, the switch comprising:

a set of ports operable as members of a first logical group, wherein a logical group is a subset of ports in the switch and precludes a data flow not belonging to the logical group from reaching a port within the logical group;

a logical-group mapping module adapted to:

identify that a packet belongs to the first logical group based on the subset of ports belonging to the first logical group; and

determine whether a failover option for the first logical group is enabled in response to a port of the first logical group not being available;

and

a forwarding module adapted to:

associate the packet with the port as an output port in response to the port being available; and

associate the packet with a second port as an output port in response to the failover option being enabled, wherein the second port is a member of a second logical group, and wherein the second logical group is distinct from the first logical group.

18. The switch of claim 17, further comprising a Quality-of-Service (QoS) enforcement module adapted to enforce one or more QoS parameters for a QoS class based on the first logical group.

19. The switch of claim 17, further comprising a logical-group configuration module adapted to:

propagate the configuration of the first logical group across one or more switches;

determine whether a local switch domain is part of the logical group; and

update a local forwarding table with the configuration of the first logical group.

20. The switch of claim 17, wherein the forwarding module is further adapted to associate the packet with the second port as an output port in response to a next-hop port within the first logical group for the packet not being available, wherein a next-hop port within the second logical group for the data flow is available.

21. The switch of claim 17, wherein the forwarding module is further operable to drop the packet belonging to the first logical group port in response to the failover option not being enabled.

22. The switch of claim 21, wherein the ports are adapted to operate as members of the first logical group are Fibre Channel ports and comprise one or more of N_Ports, E_Ports, and EX_Ports.

23. The switch of claim 21, wherein the first logical group comprises a set of inter-switch links (ISLs) coupling adjacent E_Ports that form an end-to-end path.

24. The switch of claim 21, wherein the first logical group comprises one or more EX_Ports and is operable to traverse Fibre Channel routers (FCRs) and FCR-coupled fabrics.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 9,270,580 B1
APPLICATION NO. : 12/550227
DATED : February 23, 2016
INVENTOR(S) : Vineet M. Abraham et al.

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

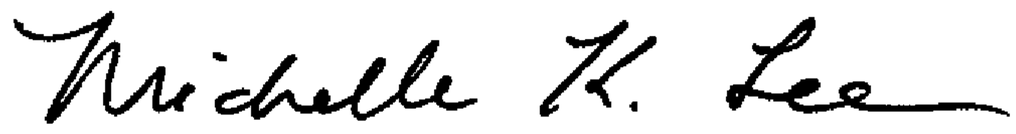
In the Claims:

In claim 6, col. 11, line 7, please delete the term "N_Ports" and replace it with the term --F_Ports--

In claim 14, col. 11, line 66, please delete the term "N_Ports" and replace it with the term --F_Ports--

In claim 22, col. 12, line 56, please delete the term "N_Ports" and replace it with the term --F_Ports--

Signed and Sealed this
Third Day of May, 2016



Michelle K. Lee
Director of the United States Patent and Trademark Office