

(51) International Patent Classification:
H04L 12/703 (2013.01)(21) International Application Number:
PCT/US2013/023725(22) International Filing Date:
30 January 2013 (30.01.2013)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.** [US/US]; 11445 Compaq Center Drive W., Houston, Texas 77070 (US).(72) Inventors: **NAZARI, Siamak**; 4245 Technology Drive, Fremont, California 94538 (US). **TAMMA, Roopesh Kumar**; 4245 Technology Drive, Fremont, California 94538 (US). **KERR, Nigel**; Forsyth House, Cromac Square, Belfast, BT2 BLA (IE).(74) Agents: **ORTEGA, Arthur** et al.; Hewlett-Packard Development Company, L.P., Intellectual Property Administra-

tion, Mail Stop 35 3404 E. Harmony Road, Fort Collins, Colorado 80528 (US).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

[Continued on next page]

(54) Title: FAILOVER IN RESPONSE TO FAILURE OF A PORT

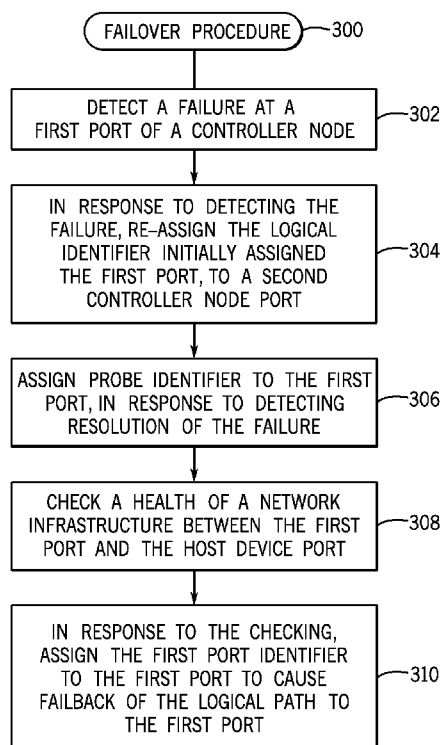


FIG. 3

(57) Abstract: A failure at a first port of the controller node is detected, where the first port is initially assigned a first port identifier and is associated with a logical path through a communications fabric between the first port and a port at a host device. In response to detecting the failure, the first port identifier is assigned to a second port to cause the logical path to be associated with the second port. In response to detecting resolution of the failure, a probe identifier is assigned to the first port. Using the probe identifier, a health of network infrastructure between the first port and the host device is checked. In response to the checking, the first port identifier is assigned to the first port to cause failback of the logical path to the first port.



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

— as to applicant's entitlement to apply for and be granted
a patent (Rule 4.17(ii))

Declarations under Rule 4.17:

— as to the identity of the inventor (Rule 4.17(i))

Published:

— with international search report (Art. 21(3))

FAILOVER IN RESPONSE TO FAILURE OF A PORT

Background

[0001] Host devices are able to access data stored at storage devices. In a network arrangement, access of the storage devices can be managed by controller nodes that are interconnected by a communications fabric to the host devices.

[0002] The host devices are able to submit data requests to the controller nodes. In response to the data requests from the host devices, the controller nodes can submit corresponding data requests to the storage devices to access (read or write) data of the corresponding storage devices.

Brief Description Of The Drawings

[0003] Some embodiments are described with respect to the following figures:

Fig. 1 is a block diagram of an example network arrangement according to some implementations;

Figs. 2A-2B are schematic diagrams illustrating an example failover procedure according to some implementations;

Fig. 3 is a flow diagram of a failover procedure according to some implementations;

Figs. 4A-4B are schematic diagrams illustrating an example failover procedure according to further implementations; and

Fig. 5 is a block diagram of a controller node according to some implementations.

Detailed Description

[0004] Controller nodes that manage access of storage devices can include storage array controllers or other types of controllers. A controller node is coupled to one or multiple storage devices (e.g. disk-based storage devices, integrated circuit

memory devices, etc.). In some arrangements, the storage devices can be part of respective groups of storage devices. For example, the groups of storage devices can include arrays of storage devices.

[0005] In some configurations, each controller node manages access of its respective group of storage device(s). In other configurations, each controller node is able to access multiple groups of storage devices.

[0006] A controller node can receive a data request from a host device through a path in a communications fabric to host devices. The communications fabric can be a storage area network (e.g. Fibre Channel storage area network) or other type of network. In response to a data request from a host device, a controller node submits a corresponding data request to a group of storage device(s) to read or write data of such group.

[0007] The host devices can be server computers, which are coupled to client devices. Client devices can submit data requests to the host devices, which in turn submit corresponding data requests to the controller nodes. In other examples, a host device can itself be a client device that is able to generate data requests (without having to first receive data requests from another client device).

[0008] Failures can occur in a network arrangement that has host devices and controller nodes. The failures can occur at various points in the network arrangement, including the controller nodes, the communications fabric, or at other points. A failure can include any of the following: malfunction or fault of hardware equipment, malfunction or fault of machine-readable instructions (software and/or firmware), a failure caused by an attack by malware (e.g. virus, worm, spyware, etc.), or any other condition that prevents normal operation of a storage system that includes the controller nodes, storage devices, and host devices.

[0009] Traditionally, to respond to a failure that prevents successful communication over a path between a host device and a controller node, logic in the host device can be used to perform a failover from the failed path to a different path. In some examples, such logic can include a multi-pathing module that is able to

selectively perform data communications over any one of multiple paths between a host device and controller nodes. The multi-pathing module can perform load balancing (to balance the data access load across multiple paths), as well as to provide failover support to fail over from one path to another path in case of a detected failure.

[0010] However, employing logic in a host device to perform failover involves making a change at the host device. When failover is performed at the host device, the host device would mark a path associated with the failure as being unavailable. As a result, the host device would no longer be able to use such path, which can reduce input/output communications capacity and can affect load balancing and/or other tasks performed at the host device.

[0011] In accordance with some implementations, instead of performing failover at a host device, failover can instead be performed (at least in part) at a controller node. The failover performed at a controller node can be transparent to a host device. In addition, failback can also be performed at a controller node if the failure condition that caused the failover is later resolved. As discussed further below, the failback is a reliable failback that first performs a health check to ensure that the network infrastructure is healthy prior to performing failback. Failback can refer to a process of returning data communications to the component that had previously failed (and thus caused a failover) and which has subsequently resumed normal operation (in other words, the failure condition has been resolved).

[0012] Fig. 1 illustrates an example network arrangement that includes host devices 102, 104, which are interconnected by a communications fabric 106 to controller nodes 108, 110. The communications fabric 106 can be a Fibre Channel storage area network, or another type of network. Although not shown in Fig. 1, the communications fabric 106 can include one or multiple switching devices to couple the host devices 102, 104 to the controller nodes 108, 110.

[0013] The network arrangement can include two or more host devices 102, and/or two or more controller nodes 108. The controller nodes 108 and 110 manage access of data in storage device groups 112, 114, where a storage device group can

include any group of one or multiple storage devices. In the example of Fig. 1, each controller node 108, 110 is able to access each of the storage device groups 112, 114. In other examples, a controller node is able to access just its associated storage device group.

[0014] As further depicted in Fig. 1, the host devices 102, 104 can be interconnected over a network 116 (e.g. local area network, wide area network, Internet, etc.) to client devices 118. Examples of client devices 118 include desktop computers, notebook computers, tablet computers, smartphones, personal digital assistants, and so forth. Client devices 118 can submit data requests to the host devices 102, 104, which in turn can submit corresponding data requests through the communications fabric 108 to the respective controller nodes 108, 110. In other examples, one or multiple ones of the host devices 102, 104 can themselves be client devices. Thus, as used here, a “host device” can refer to any type of electronic device that is able to send a data request to a controller node to access data in a storage device that is managed by the controller node.

[0015] As further depicted in Fig. 1, communications adapters 120, 122, 124, and 126 are provided in the respective host devices 102, 104, and controller nodes 108, 110. In some examples, the communications adapters 120, 122, 124, and 126 can include host bus adapters (HBAs), such as those used for communications over Fibre Channel links. In other examples, the communications adapters 120, 122, 124, and 126 can include other types of network interfaces. Each communications adapter 120, 122, 124, or 126 can include a respective port 121, 123, 125, or 127 for communicating over the communications fabric 106.

[0016] Although Fig. 1 depicts just one communications adapter in each host device or controller node, it is noted that in other examples, a host device or controller node can include multiple communications adapters with associated ports.

[0017] Each port can be assigned a port identifier, which can identify the respective communications adapter. In some examples, a port identifier can be a port world wide name (WWN). In other examples, a port identifier can be another type of identifier.

[0018] A logical path can be established between a port identifier (e.g. port WWN) of a port in a host device and a port identifier (e.g. port WWN) of a port in a controller node. Communications between a host device and a controller node can occur through the logical path. A logical path differs from a physical path. A physical path can include a specific set of physical links between a specific host device port and a specific controller node port. However, a logical path is defined by port identifiers. If a port identifier of a controller node port is re-assigned to a different controller node port, the logical path remains the same (since it is defined by port identifiers); however, after the port identifier re-assignment, the logical path provides communications between a different pair of ports.

[0019] As further depicted in Fig. 1, failover modules 128 and 130 are provided in the respective controller nodes 108 and 110. Each failover module 128 or 130 is able to detect failure associated with a respective controller node port. The failure can occur in the controller node 108, or the failure can be due to a fault condition that exists in equipment of the communications fabric 106. For example, the failure can be caused by a cable malfunction or defect. Examples of equipment failure can include failure of an HBA or fault of machine-readable instructions such as software and firmware. In addition, a failure can be indicated if data communications exhibits a relatively large number of data errors, such as based on cyclic redundancy check (CRC) detection.

[0020] In other implementations, the failover modules 128 and 130 can be provided outside the respective controller nodes 108 and 110. As described further below, the failover modules 128 and 130 can also perform failback, in case the failure condition that caused a failover is later resolved.

[0021] As an example, upon detecting a failure that prevents the controller node port 125 from communicating over the communications fabric 106 with the host device port 121, the failover module 128 in the controller node 108 is able to initiate a failover procedure. In accordance with some implementations, as depicted in Figs. 2A and 2B, the failover procedure involves re-assigning the port identifier (e.g. WWN1) of the controller node port 125 to another controller node port. In the

example of Figs. 2A and 2B, the port identifier WWN1 initially assigned to the port 125 in the controller node 108 can be re-assigned to the port 127 in the controller node 110. In other examples, if the controller node 108 has multiple communications adapters with respective ports, the failover can cause re-assignment of the port identifier WWN1 from a first port of the controller node 108 to a second port of the controller node 108.

[0022] In the example of Figs. 2A and 2B, it is assumed that the host device port 121 is assigned a port identifier WWNx. In such example, a logical path P1 is established between the WWN1 and WWNx. Prior to re-assignment of the port identifier WWN1 as part of a failover procedure, communications over the logical path P1 can occur between the host device port 121 and the controller node port 125, as depicted in Fig. 2A.

[0023] In Fig. 2A, the controller node port 125 can be considered an active port for logical path P1, while the controller node port 127 can be considered a backup port for logical path P1. After failover has been completed as depicted in Fig. 2B, the controller node port 127 becomes the active port for logical path P1, while the controller node port 125 becomes the backup port for logical path P1.

[0024] After re-assignment of WWN1 from the controller node port 125 to the controller node port 127, communications over the logical path P1 (still defined between WWN1 and WWNx) can occur between the host device port 121 and the controller node port 127, as depicted in Fig. 2B. The controller node port 125 becomes the backup port for logical path P1.

[0025] More generally, prior to re-assigning WWN1, the logical path P1 is associated with the controller node port 125. However, after re-assigning WWN1, the logical path is associated with the controller node port 127.

[0026] If the failure condition at the controller node port 125 is later resolved such that the failure condition no longer exists, the controller node port 125 can be reactivated and WWN1 can be re-assigned from the controller node port 127 back to the controller node 125. Thus, the failover procedure can further perform failback by

re-assigning WWN1 back to the controller node port 125, at which point the logical path P1 is as depicted in Fig. 2A. The failover module 128 or 130 can be notified of the resolution of the failure condition, and this notification can trigger the failback.

[0027] In accordance with some implementations, prior to performing the failback, the controller node port 125 to which the failback is to occur can first be temporarily assigned a probe identifier, which can be another WWN (different from WWN1 and WWN2). The probe identifier can be used for the purpose of checking the health of the network infrastructure (including physical paths and switches) between the controller node port 125 and the host device 102. Checking the health of the network infrastructure avoids a ping-pong failover/failback scenario where a failover of the logical path P1 first occurs from the controller node port 125 to the controller node 127, followed by failback from the controller node port 127 back to the controller node 125, followed further by another failover from the controller node port 125 to the controller node 127 should it be determined that the network infrastructure between the controller node port 125 and the host device 102 is not healthy.

[0028] Checking the health of the network infrastructure can include checking to ensure that components (including physical paths and switches) of the network infrastructure are working properly so that communications can occur between the controller node port 125 and the host device port 121. In addition, checking the health of the network infrastructure can also include checking to ensure that there is a valid physical path from the controller node port 125 to the host device port 121. In an example where there are multiple communications fabrics, the controller node port 125 may have been re-connected to a different communications fabric following the initial failover from the controller node port 125 to the controller node port 127. The host device port 121 may not be connected to the different communications fabric, and thus re-assigning the logical path P1 back to the controller node port 125 may result in a situation where communications is not possible between the controller node port 125 and the host device port 121 over the logical path P1.

[0029] From the perspective of the host device 102, the failover and failback at the controller nodes 108 and 110 is transparent to the host device 102. In the present discussion, a failover procedure can include both failover and failback. After failover and any subsequent failback, the logical path P1 between WWN1 and WWNx remains visible to the host device 102, which can continue to use the logical path P1 for communications.

[0030] In the example of Figs. 2A and 2B, note that, prior to the failover, the controller node port 127 was already assigned a port identifier (e.g. WWN2), which can define another logical path between WWN2 and a host device port identifier. As a result, if the failover procedure causes re-assignment of WWN1 from the controller node port 125 to the controller node port 127, the controller node port 127 would be associated with multiple port WWNs (WWN1 and WWN2) after the failover completes. To allow for a physical port (e.g. controller node port 127) to be associated with multiple port identifiers, N_Port Identifier Virtualization (NPIV) can be provided at the controller nodes 108 and 110. NPIV is a Fibre Channel mechanism that allows multiple port identifiers (e.g. port WWNs) to be associated with a single physical port. In other examples, other mechanisms can be used to allow multiple logical port identifiers to be assigned to the same physical port.

[0031] After failover, the controller node port 127 can potentially communicate over multiple different logical paths, including the logical path P1 between WWNx and WWN1, and another logical path between a host device port identifier and WWN2.

[0032] Fig. 3 is a flow diagram of a failover procedure 300 according to some implementations. The failover procedure can be performed by a failover module (e.g. 128 in Fig. 1) in a controller node or a combination of failover modules (e.g. 128 and 130 in Fig. 1) in multiple controller nodes. The failover procedure 300 includes detecting (at 302) a failure at a first port of a controller node in a storage system. The failure at the first port can be caused by failure in the controller node, or failure in the communications fabric 106 (Fig. 1). The first port is initially assigned a first

port identifier and is associated with a logical path through the communications fabric 106 between the first port and a port at a host device.

[0033] In response to detecting the failure, the failover procedure 300 re-assigns (at 304) the first port identifier to a second port in the storage system to cause the logical path to be associated with the second port. The second port can be another port of the same controller node, or alternatively, the second port can be a port of another controller node.

[0034] Subsequently, in response to detecting resolution of the failure, the failover procedure 300 assigns (at 306) a probe identifier to the first port. The assignment of the probe identifier to the first port can be a temporary assignment. Resolution of the failure can be detected by the failover module 128 or 130; alternatively, resolution of the failure can be indicated by equipment in the communications fabric 106 to the failover module 128 or 130. Using the probe identifier, the failover procedure 300 checks (at 308) a health of a network infrastructure between the first port and the port of the host device. In response to the checking indicating that the network infrastructure is healthy, the failover procedure 300 assigns (at 310) the first port identifier to the first port to cause failback of the logical path to the first port.

[0035] Checking of the health of the network infrastructure can be accomplished as follows, in accordance with some examples. After the first port has been assigned the probe identifier (which can be a probe WWN), the first port attempts to login to the communications fabric 106 using the probe identifier. The login is performed with a server in the communications fabric 106. If login is unsuccessful, that is an indication that the network infrastructure is not healthy, and thus failback of the logical path back to the first port would not be performed.

[0036] If login is successful, the failover module 128 (or another entity associated with the first port) can perform a communications test of the network infrastructure. Login of the first port using the probe identifier allows the first port to perform communications over the communications fabric 106. For example, the test can be a loopback test in which test packets can be sent from the first port to the host

device port, to obtain a response from the host device port. If a response can be obtained from the host device port in response to the test packets, then the network infrastructure is determined to be healthy. More generally, the test can involve performing a test communication in the communications fabric 106 for ascertaining the health of the communications fabric for communications between the first port and the host device port.

[0037] Once the health of the network infrastructure between the first port at the controller node and the host device port has been confirmed, the first port can logout the probe identifier from the communications fabric 106. At this point, the failback performed at 310 can proceed.

[0038] Figs. 4A and 4B depict a failover arrangement according to further implementations. In Fig. 4A, a logical path is established between WWNx (port identifier of the host device port 121), and WWN1 (port identifier initially assigned to the controller node port 125). As further depicted in Fig. 4A, a switch 402 in the communications fabric 106 interconnects the host device port 121 to the controller node port 125.

[0039] Fig. 4B depicts failure (represented by an "X") of the controller node port 125. As a result of this failure, the failover module 128 (Fig. 1) in the controller node 108, either by itself or in cooperation with the failover module 130 in the controller node 110, performs a failover procedure to re-assign WWN1 from the controller node port 125 to the controller node port 127, as depicted in Fig. 4B. As part of this re-assignment, the failover module(s) 130 and/or 132 cooperate with the switch 402 to re-connect the host device port 121 to the controller node port 127. Note that, even though the physical path has been changed by the switch 402, the logical path remains unchanged, since the logical path is still established between WWN1 and WWNx.

[0040] In accordance with some implementations, the switch 402 includes a port connection database (or other data structure) 404. The port connection database 404 has multiple entries, where each entry maps a host device port identifier (e.g. port WWN) to a respective host device physical port, and maps a controller node

port identifier (e.g. port WWN) to a respective controller node physical port. As part of the failover procedure, the port connection database 404 is updated, based on interaction between the failover module(s) 128 and/or 130 and the switch 402. The respective entry of the port connection database 404 is updated to indicate that the port identifier WWN1 is re-assigned to the physical port 127, rather than physical port 125.

[0041] If a failback is subsequently performed in response to resolution of the failure, in which the port identifier WWN1 is assigned back to the physical port 125, then the respective entry of the port connection database 404 can be updated again.

[0042] Fig. 5 is a block diagram of an example controller node 500, which can be the controller node 108 or 110 of Fig. 1, for example. The controller node 500 has a failover module 502, which can be the failover module 128 or 130 of Fig. 1. The failover module 502 can be implemented as machine-readable instructions that are executable on one or multiple processors 504. A processor can include a microprocessor, microcontroller, processor module or subsystem, programmable integrated circuit, programmable gate array, or another control or computing device.

[0043] The processor(s) 504 can be connected to a communications interface 506 (e.g. communications adapter 124 or 126 in Fig. 1) and a storage medium (or storage media) 508. The storage medium (or storage media) 508 can store WWN(s) assigned to port(s) of the controller node 500.

[0044] The storage medium (or storage media) 508 can be implemented as one or multiple computer-readable or machine-readable storage media. The storage media include different forms of memory including semiconductor memory devices such as dynamic or static random access memories (DRAMs or SRAMs), erasable and programmable read-only memories (EPROMs), electrically erasable and programmable read-only memories (EEPROMs) and flash memories; magnetic disks such as fixed, floppy and removable disks; other magnetic media including tape; optical media such as compact disks (CDs) or digital video disks (DVDs); or other types of storage devices. Note that the instructions discussed above can be provided on one computer-readable or machine-readable storage medium, or

alternatively, can be provided on multiple computer-readable or machine-readable storage media distributed in a large system having possibly plural nodes. Such computer-readable or machine-readable storage medium or media is (are) considered to be part of an article (or article of manufacture). An article or article of manufacture can refer to any manufactured single component or multiple components. The storage medium or media can be located either in the machine running the machine-readable instructions, or located at a remote site from which machine-readable instructions can be downloaded over a network for execution.

[0045] In the foregoing description, numerous details are set forth to provide an understanding of the subject disclosed herein. However, implementations may be practiced without some or all of these details. Other implementations may include modifications and variations from the details discussed above. It is intended that the appended claims cover such modifications and variations.

What is claimed is:

- 1 1. A failover method for a storage system, comprising:
2 detecting a failure at a first port of a controller node in the storage system,
3 wherein the first port is initially assigned a first port identifier and is associated with a
4 logical path through a communications fabric between the first port and a port at a
5 host device;
6 in response to detecting the failure, assigning the first port identifier to a
7 second port of the storage system to cause the logical path to be associated with the
8 second port;
9 in response to detecting resolution of the failure, assigning a probe identifier
10 to the first port;
11 using the probe identifier, checking a health of network infrastructure between
12 the first port and the host device; and
13 in response to the checking, assigning the first port identifier to the first port to
14 cause failback of the logical path to the first port.
- 1 2. The method of claim 1, wherein assigning the first port identifier comprises
2 identifying a port world wide name (WWN).
- 1 3. The method of claim 1, further comprising using the probe identifier to login to
2 the communications fabric, wherein checking the health is performed after
3 performing the login.
- 1 4. The method of claim 3, wherein checking the health includes performing a test
2 communication in the communications fabric.
- 1 5. The method of claim 1, wherein assigning the first port identifier to the second
2 port causes the second port to be associated with the first port identifier and a
3 second port identifier.

- 1 6. The method of claim 5, further comprising:
2 communicating, by the second port, data over the logical path using the first
3 port identifier, wherein the logical path is established between the first port identifier
4 and a host device port identifier; and
5 communicating, by the second port, data over a second logical path using the
6 second port identifier, wherein the second logical path is established between the
7 second port identifier and a host device port identifier.
- 1 7. The method of claim 1, wherein the logical path is established between the
2 first port identifier and a second port identifier assigned to the port at the host device.
- 1 8. The method of claim 7, further comprising:
2 prior to the assigning, communicating data through the logical path between
3 the first port and the port at the host device; and
4 after the assigning, communicating data through the logical path between the
5 second port and the port at the host device.
- 1 9. The method of claim 1, wherein the second port is part of the controller node
2 or another controller node.

1 10. A system comprising:
2 at least one storage device;
3 a controller node to manage access of data in the at least one storage device,
4 wherein the controller node has a first port initially assigned a first port identifier,
5 wherein a logical path is established between the first port identifier and a port
6 identifier of a port at a host device; and
7 a failover module to detect failure at the first port, and to initiate a failover
8 procedure in response to the detected failure,
9 wherein the failover procedure includes re-assigning the first port identifier
10 from the first port to a second port to cause the logical path to be associated with the
11 second port after the re-assigning, and
12 wherein the failover procedure further includes:
13 in response to detecting resolution of the failure, use a probe identifier
14 assigned to the first port to perform a test communication in a communications
15 fabric;
16 in response to determining from the test communication that network
17 infrastructure in the communications fabric is healthy, assign the first port identifier
18 back to the first port to perform failback to re-associate the logical path with the first
19 port.

1 11. The system of claim 10, wherein the first port identifier includes a port world
2 wide name.

1 12. The system of claim 10, wherein the first port is indicated as an active port for
2 the logical path and the second port is indicated as a backup port for the logical path
3 prior to the failover procedure, and wherein the first port is indicated as a backup port
4 for the logical path and the second port is indicated as an active port for the logical
5 path after the failover procedure.

1 13. The system of claim 10, wherein the second port is initially assigned a second
2 port identifier prior to the failover procedure, and wherein the second port is assigned
3 the first and second port identifiers after the failover procedure.

1 14. The system of claim 10, wherein the test communication is performed after
2 login using the probe identifier of the first port with the communications fabric, and
3 wherein after determining that the network infrastructure is healthy, the first port is
4 logged out from the communications fabric to allow performance of the failback.

1 15. An article comprising at least one machine-readable storage medium storing
2 instructions that upon execution cause a controller node to:

3 detect a failure at a first port of the controller node that manages access of
4 data in at least one storage device, wherein the first port is initially assigned a first
5 port identifier and is associated with a logical path through a communications fabric
6 between the first port and a port at a host device;

7 in response to detecting the failure, assign the first port identifier to a second
8 port to cause the logical path to be associated with the second port;

9 in response to detecting resolution of the failure, assign a probe identifier to
10 the first port;

11 using the probe identifier, check a health of network infrastructure between
12 the first port and the host device; and

13 in response to the checking determining that the network infrastructure is
14 healthy, assign the first port identifier to the first port to cause failback of the logical
15 path to the first port.

1 / 7

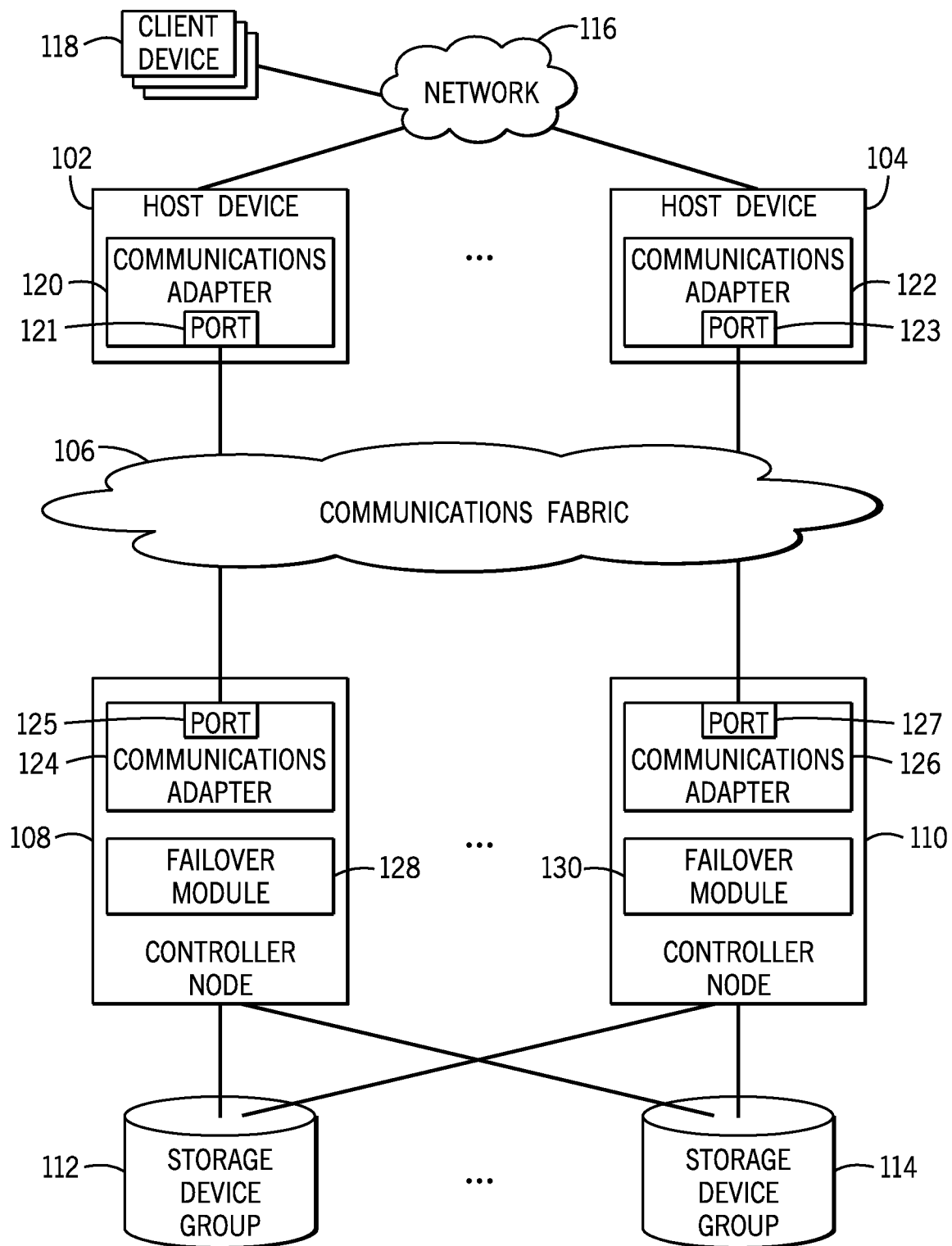


FIG. 1

2 / 7

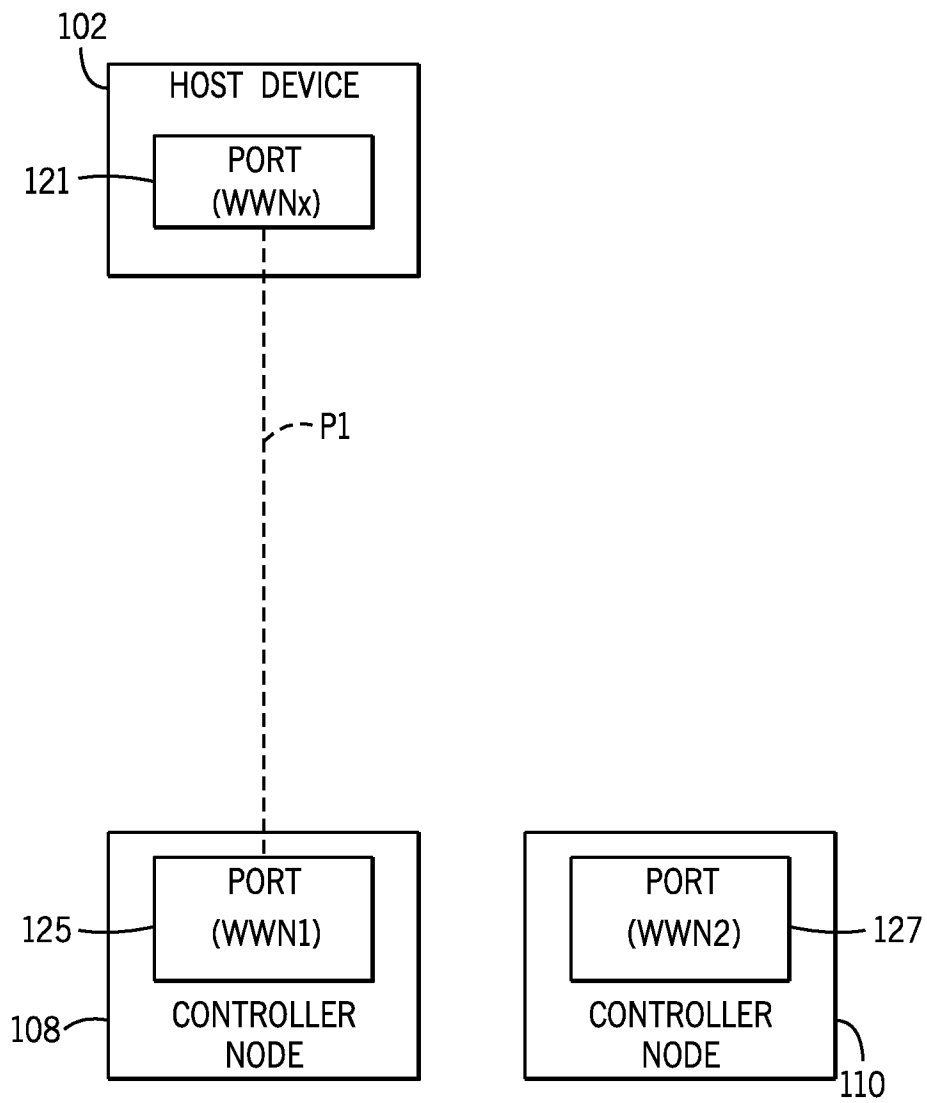


FIG. 2A

3 / 7

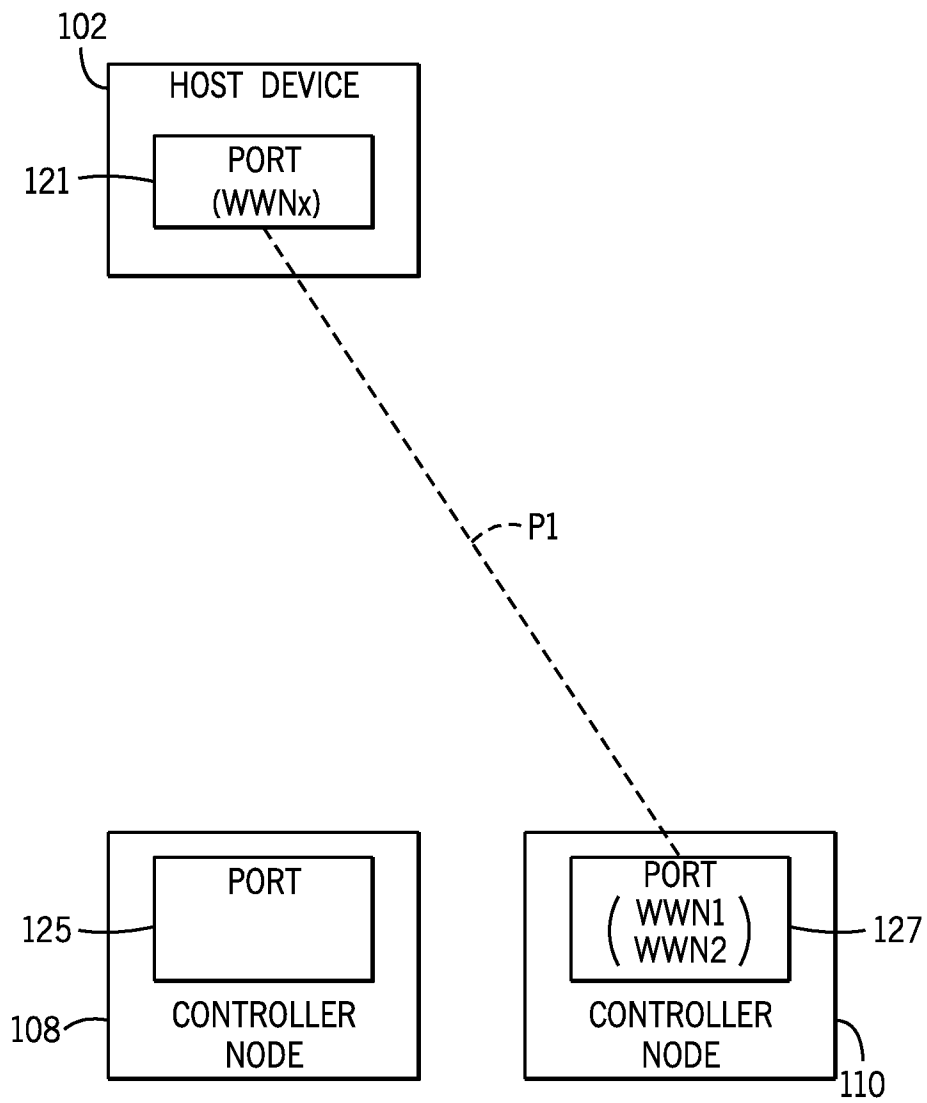


FIG. 2B

4 / 7

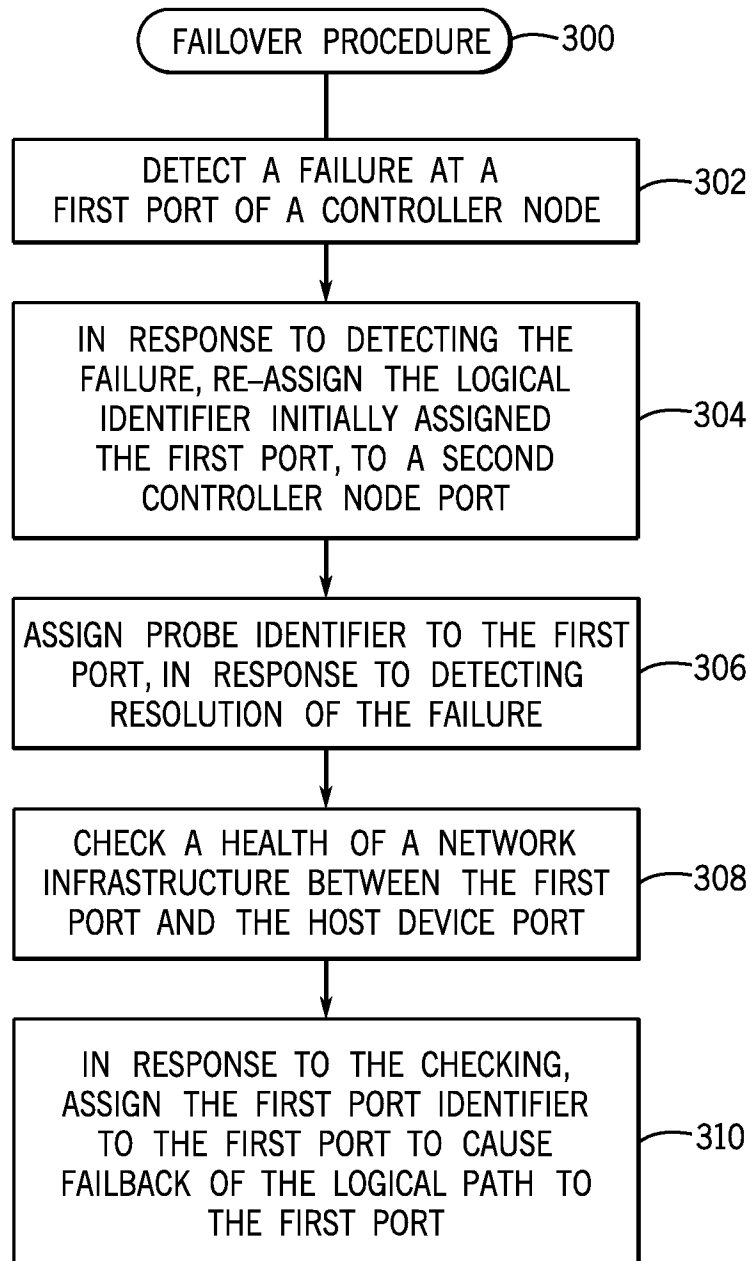


FIG. 3

5 / 7

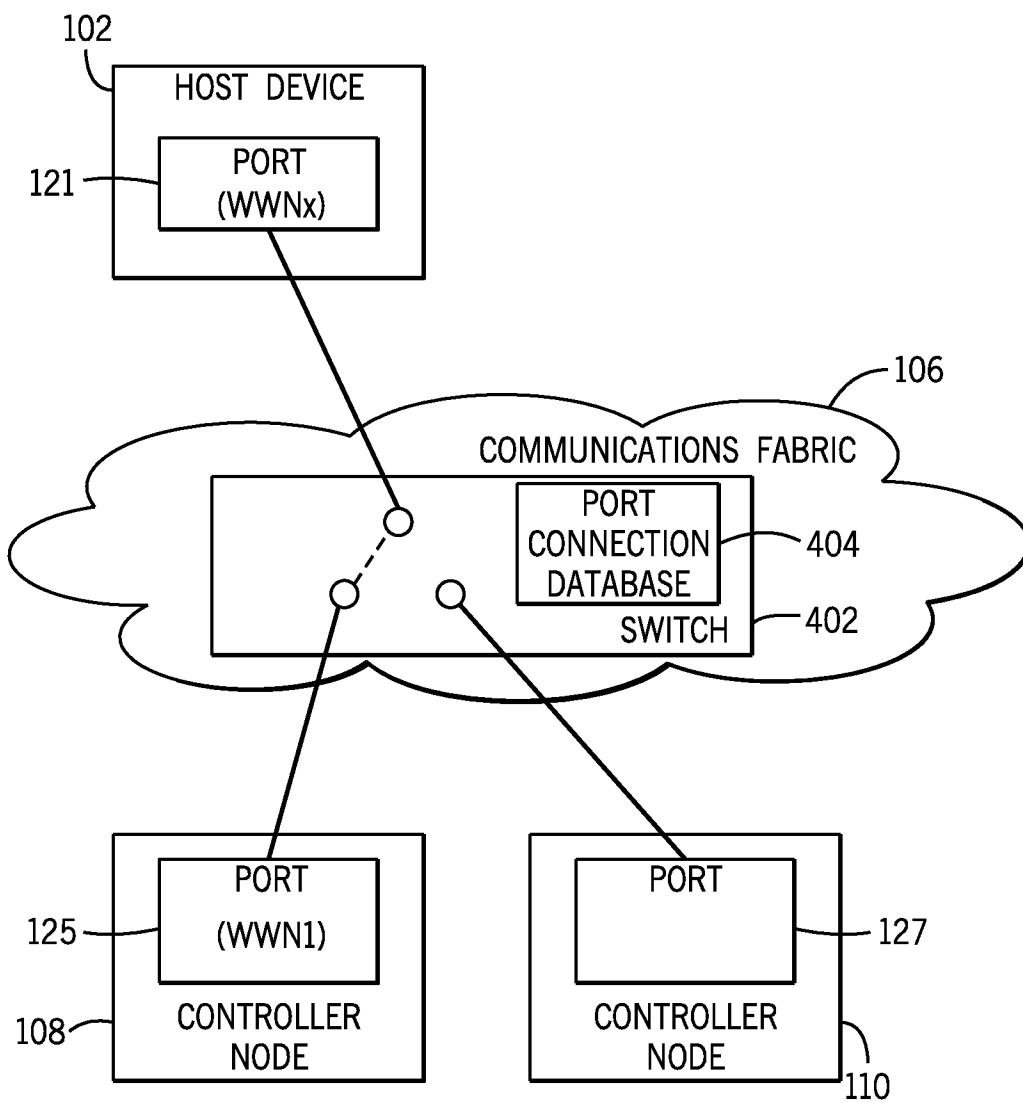


FIG. 4A

6 / 7

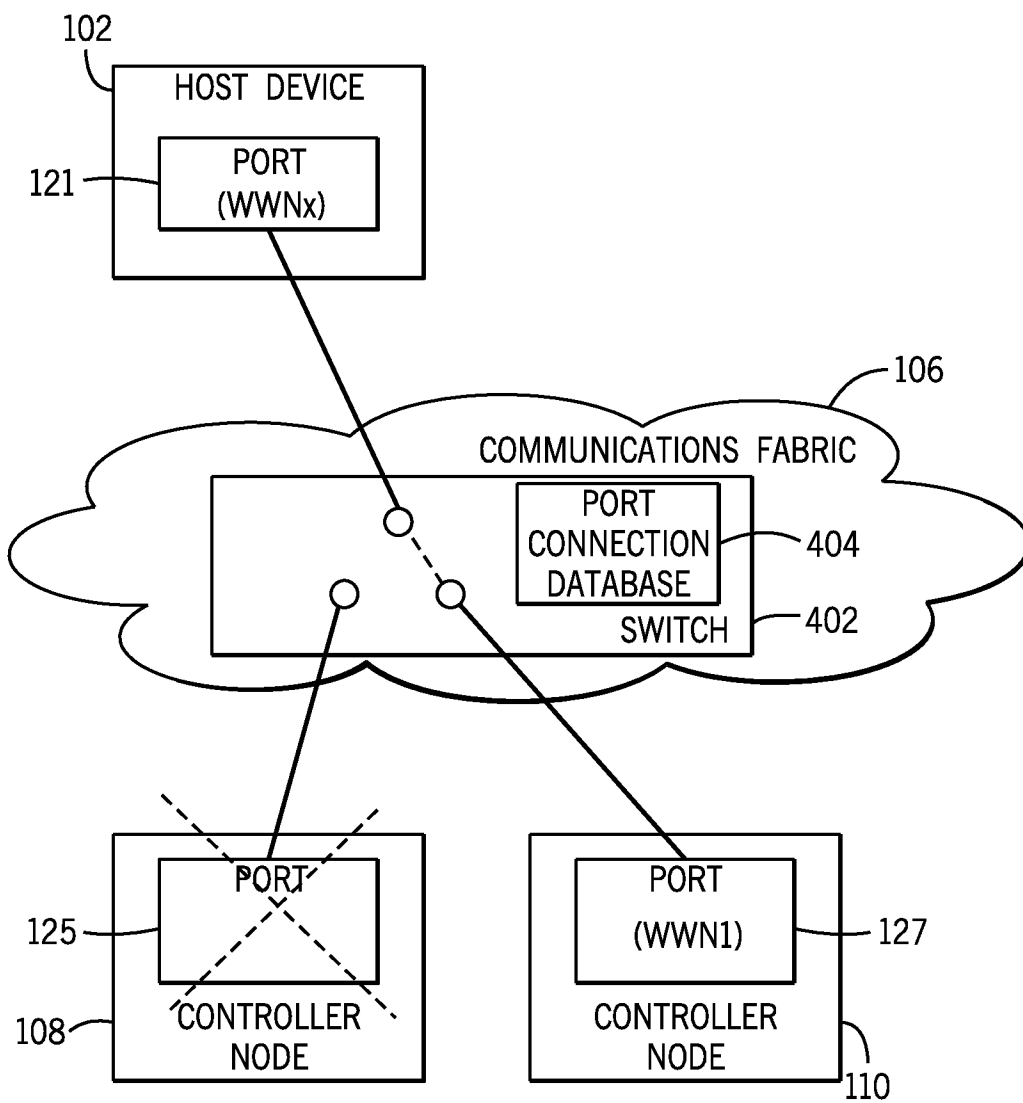


FIG. 4B

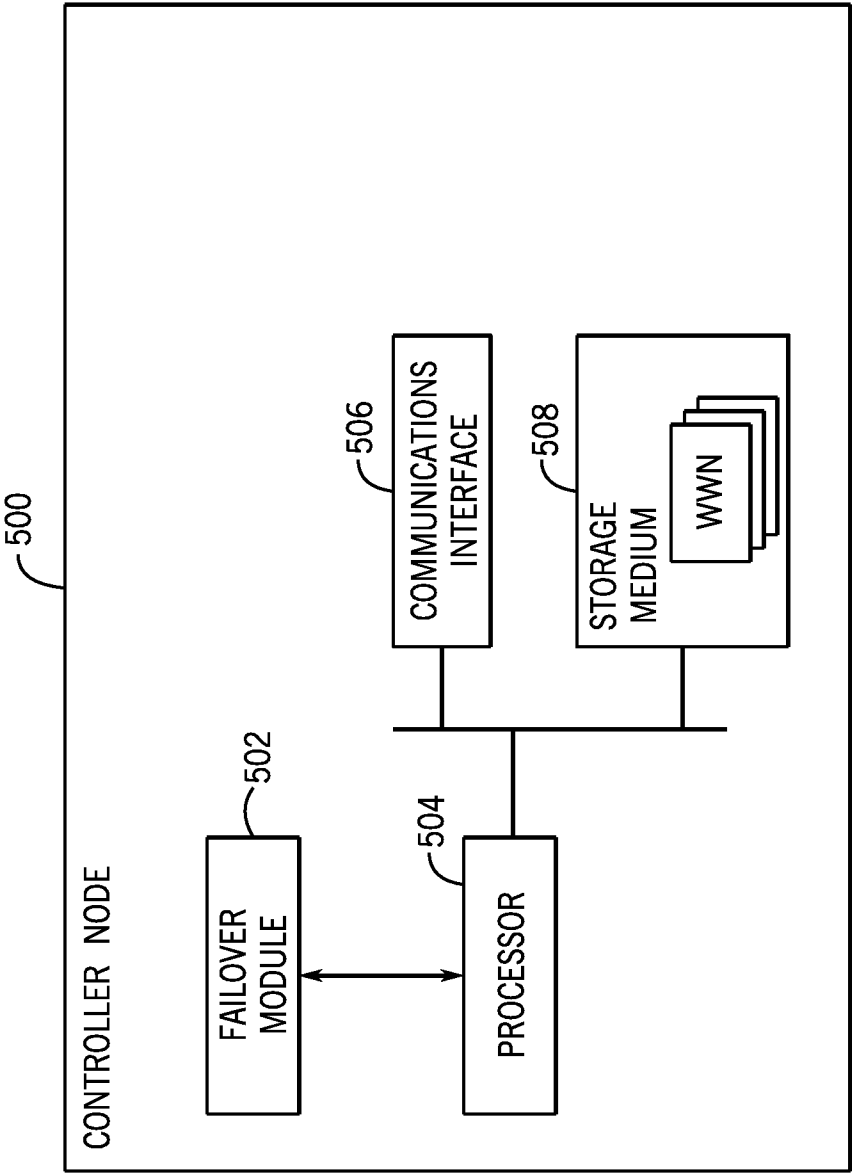


FIG. 5

A. CLASSIFICATION OF SUBJECT MATTER**H04L 12/703(2013.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L 12/703; H04L 12/26; G06F 13/14; G06F 13/36; H04L 12/28; G01R 31/08

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS(KIPO internal) & Keywords: failover, storage system, controller node, port identifier, failure, failback

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2010-0097941 A1 (SCOTT M. CARLSON et al.) 22 April 2010 See paragraphs [0038], [0040]–[0042]; and figures 4, 5.	1–15
A	US 6775230 B1 (NAOKI WATANABE et al.) 10 August 2004 See column 3, lines 51–52; column 5, lines 27–55; and figure 3.	1–15
A	US 2009-0106475 A1 (RICHARD LOUIS ARNDT et al.) 23 April 2009 See paragraphs [0087]–[0090]; and figure 17.	1–15
A	US 6990068 B1 (ALI N. SALEH et al.) 24 January 2006 See paragraphs [0047]–[0048]; and figure 3A.	1–15
A	US 2008-0059664 A1 (ROBERT ALLAN UNGER) 06 March 2008 See paragraphs [0042]–[0053]; and figure 4.	1–15



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

24 October 2013 (24.10.2013)

Date of mailing of the international search report

25 October 2013 (25.10.2013)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City,
302-701, Republic of Korea

Facsimile No. +82-42-472-7140

Authorized officer

KIM, Do Weon

Telephone No. +82-42-481-5560



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2013/023725

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2010-0097941 A1	22/04/2010	CN 102171973 A EP 2361467 A1 EP 2361467 B1 JP 04838400 B1 JP 2012-506184 A KR 10-2011-0086820 A TW 201029384 A US 7944812 B2 WO 2010-046294 A1	31/08/2011 31/08/2011 05/09/2012 14/12/2011 08/03/2012 01/08/2011 01/08/2010 17/05/2011 29/04/2010
US 6775230 B1	10/08/2004	None	
US 2009-0106475 A1	23/04/2009	US 2006-0195674 A1 US 2008-0071960 A1 US 7308551 B2 US 7487326 B2 US 7653801 B2	31/08/2006 20/03/2008 11/12/2007 03/02/2009 26/01/2010
US 6990068 B1	24/01/2006	AU 2000-35820 A1 US 2001-0033548 A1 US 2001-0048660 A1 US 2002-0054572 A1 US 2003-0031127 A1 US 2003-0058804 A1 US 2003-0179700 A1 US 2003-0179701 A1 US 2005-0036442 A1 US 2005-0047327 A1 US 2005-0135234 A1 US 2005-0201380 A1 US 2006-0153066 A1 US 2008-0225696 A1 US 6801496 B1 US 6850486 B2 US 6856627 B2 US 7002917 B1 US 7200104 B2 US 7301895 B2 US 7352692 B1 US 7424035 B2 US 7428212 B2 US 7477594 B2 US 7502313 B2 US 7633854 B2 US 7724655 B2 US 7729337 B2 US 7764596 B2 US 8064481 B2 WO 00-42746 A1	01/08/2000 25/10/2001 06/12/2001 09/05/2002 13/02/2003 27/03/2003 25/09/2003 25/09/2003 17/02/2005 03/03/2005 23/06/2005 15/09/2005 13/07/2006 18/09/2008 05/10/2004 01/02/2005 15/02/2005 21/02/2006 03/04/2007 27/11/2007 01/04/2008 09/09/2008 23/09/2008 13/01/2009 10/03/2009 15/12/2009 25/05/2010 01/06/2010 27/07/2010 22/11/2011 20/07/2000

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2013/023725

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2008-0059664 A1	06/03/2008	AU 2003-220256 A1	13/10/2003
		CN 1910566 A	07/02/2007
		EP 1488325 A1	22/12/2004
		JP 2006-505023 A	09/02/2006
		US 2003-0188007 A1	02/10/2003
		US 7337232 B2	26/02/2008
		US 7668967 B2	23/02/2010
		WO 03-083687 A1	09/10/2003