



(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2017/0033975 A1**

Poojary et al.

(43) **Pub. Date: Feb. 2, 2017**

(54) **METHODS FOR PRIORITIZING FAILOVER OF LOGICAL INTERFACES (LIFS) DURING A NODE OUTAGE AND DEVICES THEREOF**

(52) **U.S. Cl.**
CPC **H04L 41/0668** (2013.01)

(71) Applicant: **NetApp, Inc.**, Sunnyvale, CA (US)

(57) **ABSTRACT**

(72) Inventors: **Pradeep Poojary**, Bangalore (IN);
Sanjeev Manurkar, Bangalore (IN)

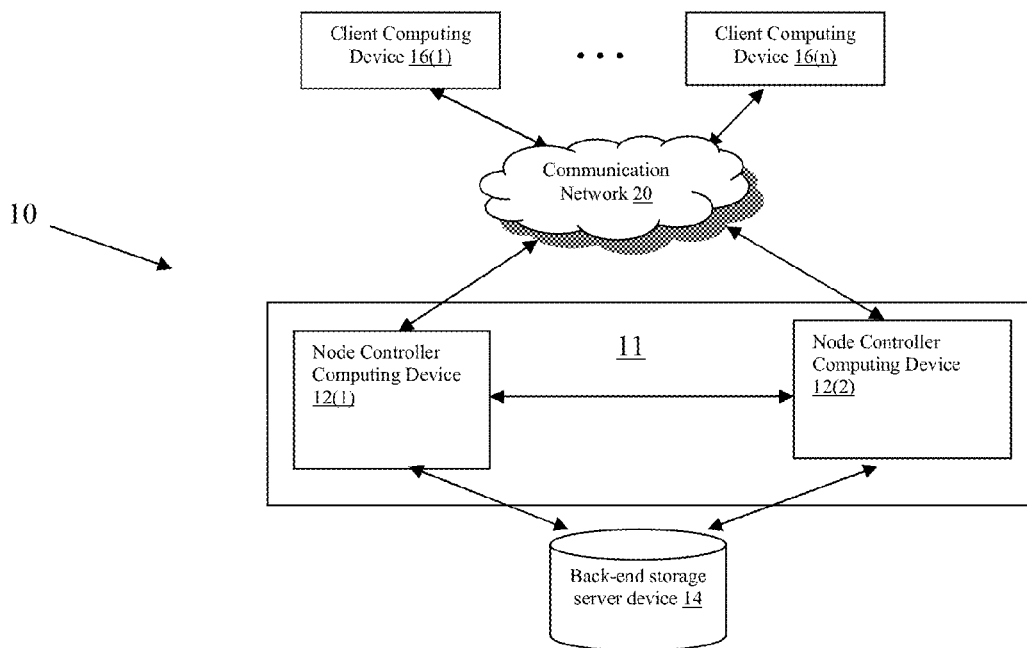
Methods, devices, and non-transitory computer readable media that prioritize failover of logical interfaces (LIFs) include identifying an outage on another one of the plurality of node computing devices in a storage cluster. When the another one of the plurality of node computing devices with the identified outage has provisioned a number of logical interfaces (LIFs) above a stored provisioning threshold is determined. Failover of the LIFs on the another one of the plurality of node computing devices with the outage is managed in an order based on at least one of a plurality of priority weight values assigned to each of the LIFs.

(21) Appl. No.: **14/812,550**

(22) Filed: **Jul. 29, 2015**

Publication Classification

(51) **Int. Cl.**
H04L 12/24 (2006.01)



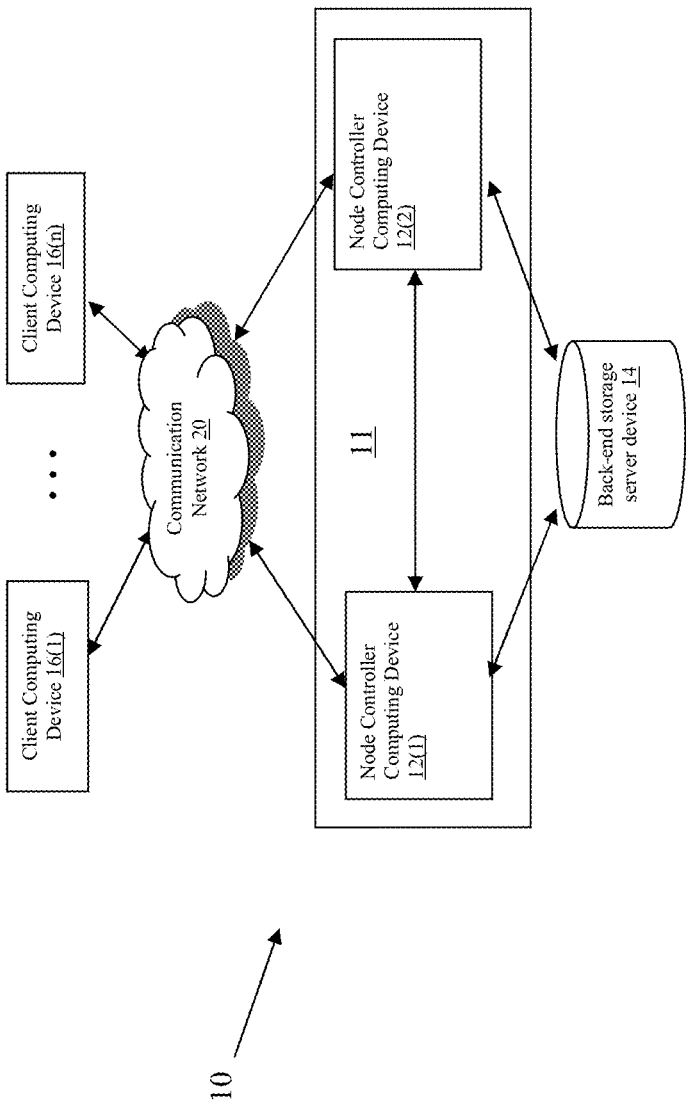


FIG. 1

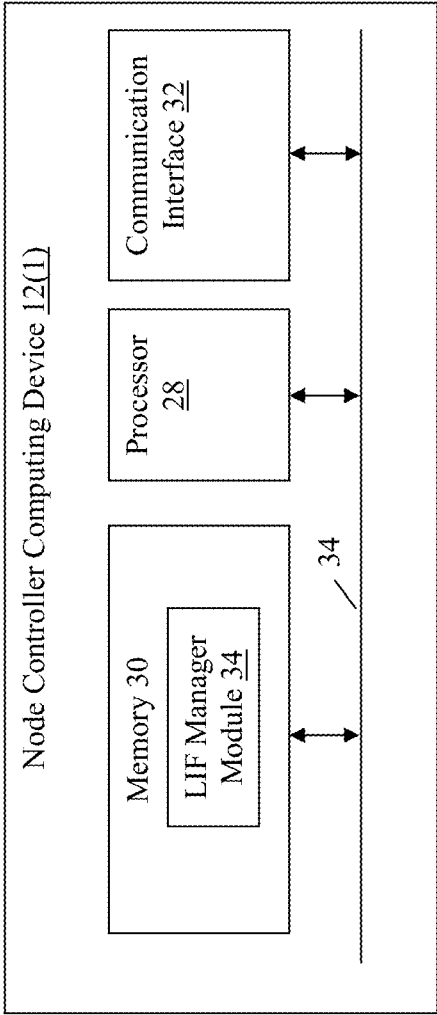


FIG. 2

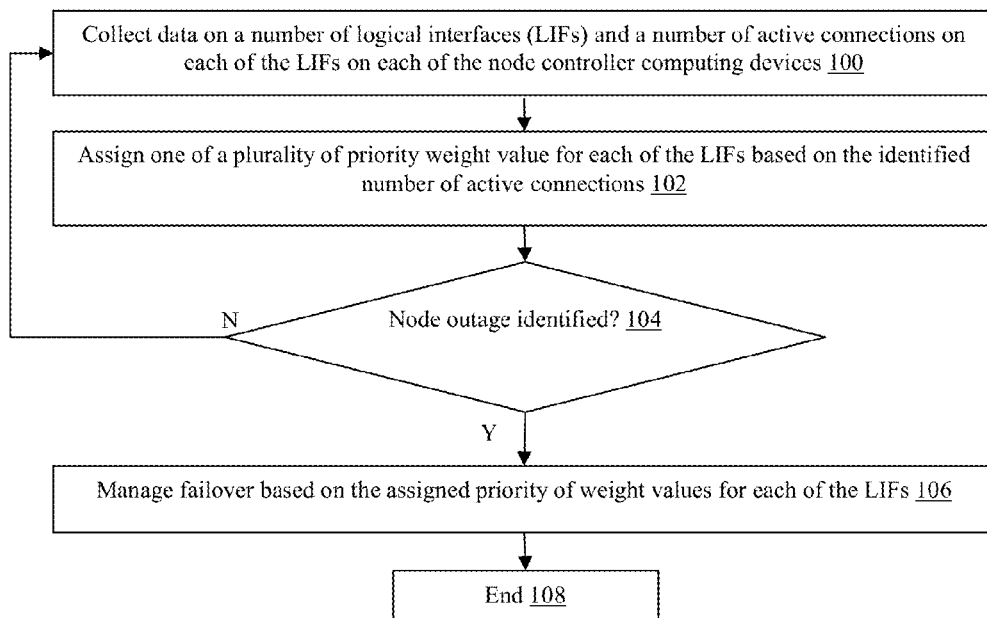


FIG. 3

Node	Logical Interface Name	Number of Active Connections
node controller computing device 12(1) or node 0	datlif1a	3
node controller computing device 12(1) or node 0	datlif2a	6
node controller computing device 12(1) or node 0	datlif3a	5
node controller computing device 12(2) or node 1	datlif1b	3
node controller computing device 12(1) or node 1	datlif2b	3
node controller computing device 12(1) or node 1	datlif3b	5

FIG. 4

METHODS FOR PRIORITIZING FAILOVER OF LOGICAL INTERFACES (LIFs) DURING A NODE OUTAGE AND DEVICES THEREOF

FIELD

[0001] This technology generally relates to methods and devices for managing nodes and, more particularly to methods for prioritizing failover of logical interfaces (LIFs) during a node outage and devices thereof

BACKGROUND

[0002] In a storage management environment, it is possible to over-provision a number of logical interfaces LIFs on one or more of the nodes. In such a situation, when a node outage occurs, such as an unexpected node outage or even during a scheduled node outage for an upgrade events, the failover from the over provisioned node to another node may result in one or more LIFs not being hosted anywhere. If an active LIF is not hosted during a node outage, this can result in a disruption such that client computing devices are unable to access data or other resources from the back end storage device.

SUMMARY

[0003] A method for prioritizing failover of logical interfaces (LIFs) includes identifying by one of a plurality of node computing devices an outage on another one of the plurality of node computing devices in a storage cluster. When the another one of the plurality of node computing devices with the identified outage has provisioned a number of logical interfaces (LIFs) above a stored provisioning threshold is determined by the one of a plurality of node computing devices. Failover of the LIFs on the another one of the plurality of node computing devices with the outage is managed by the one of a plurality of node computing devices in an order based on at least one of a plurality of priority weight values assigned to each of the LIFs.

[0004] A node controller computing device includes a memory coupled to one or more processors which are configured to be capable of executing programmed instructions stored in the memory to identify an outage on another one of the plurality of node computing devices in a storage cluster. When the another one of the plurality of node computing devices with the identified outage has provisioned a number of logical interfaces (LIFs) above a stored provisioning threshold is determined. Failover of the LIFs on the another one of the plurality of node computing devices with the outage is managed in an order based on at least one of a plurality of priority weight values assigned to each of the LIFs.

[0005] A non-transitory computer readable media for prioritizing failover of logical interfaces (LIFs) includes identifying an outage on another one of the plurality of node computing devices in a storage cluster. When the another one of the plurality of node computing devices with the identified outage has provisioned a number of logical interfaces (LIFs) above a stored provisioning threshold is determined. Failover of the LIFs on the another one of the plurality of node computing devices with the outage is managed in an order based on at least one of a plurality of priority weight values assigned to each of the LIFs.

[0006] This technology provides a number of advantages including providing methods, devices, and non-transitory

computer readable media that effectively prioritize failover of logical interfaces (LIFs) during a node outage. With this technology, LIFs with active connections are given a higher priority to failover rather than other LIFs without any connections or with a lower number of active connections. As a result, with this technology during node outages, including unexpected node outages or even during scheduled node outages for upgrade events, LIFs with active connections may either be disruption free or at least any impact on disrupting active connections during the node outage is minimized.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIG. 1 is a block diagram of a network environment with an example of a node controller computing devices in a storage cluster;

[0008] FIG. 2 is a block diagram of an example of one of the node controller computing devices;

[0009] FIG. 3 is a flow chart of an example of a method for prioritizing failover of logical interfaces (LIFs) during a node outage; and

[0010] FIG. 4 is an example of a table of a number of obtained and stored logical interfaces (LIFs) and a number of active connections for the node controller computing devices in the storage cluster shown in FIG. 1.

DETAILED DESCRIPTION

[0011] An example of an environment 10 with a storage cluster 11 comprising node controller computing devices 12(1)-12(2) is illustrated in FIGS. 1-2. In this particular example, the environment 10 includes the node controller computing devices or nodes 12(1)-12(2), a back-end storage server device 14, and client computing devices 16(1)-16(n) coupled via one or more communication networks 20, although the environment could include other types and numbers of systems, devices, components, and/or other elements in other configurations. By way of example only, the environment may include additional numbers of node controller computing devices and/or additional numbers of back-end storage server devices in other configurations. This technology provides a number of advantages including providing methods, non-transitory computer readable media and devices that effectively prioritize failover of logical interfaces (LIFs) during a node outage.

[0012] Referring more specifically to FIGS. 1-2, each of the node controller computing devices 12(1)-12(2) in the storage cluster 11 are used to manage service of input or output (I/O) operations between the back-end storage server device 14 and the client computing devices 16(1)-16(n) by way of example only, although each could perform other types and/or numbers of other operations and other types and/or numbers of nodes could be used. Additionally, in this particular example each of the node controller computing devices 12(1)-12(2) represent a physical machine used to manage these I/O operations, although other configurations, such as a virtual network with virtual machines implementing one or more of the node controller computing devices 12(1)-12(2) could also be used by way of example only.

[0013] For ease of illustration and description, only node controller computing device 12(1) is illustrated in FIG. 2. In this particular example, node controller computing device 12(2) is the same in structure and operation as the node controller computing device, although any of the node

controller computing devices could comprise other types and/or numbers of other physical and/or virtual systems, and/or other devices, components, and/or other elements in other configurations with other operations and/or functions. Additionally, in this particular example the node controller computing device 12(1) includes a processor 28, a memory 30, and a communication interface 32 which are coupled together by a bus 30, although the node controller computing device 12(1) may include other types and/or numbers of physical and/or virtual systems, devices, components, and/or other elements in other configurations.

[0014] The processor 28 of in the node controller computing device 12(1) may execute one or more programmed instructions stored in the memory 30 aspects of this technology as illustrated and described in the examples herein, although other types and numbers of functions and/or other operation can be performed. The processor 28 in the node controller computing device 12(1) may include one or more central processing units and/or general purpose processors with one or more processing cores, for example.

[0015] The memory 30 of in the node controller computing device 12(1) stores the programmed instructions and other data for one or more aspects of the present technology as described and illustrated herein, although some or all of the programmed instructions could be stored and executed elsewhere. A variety of different types of memory storage devices, such as a random access memory (RAM) or a read only memory (ROM) in the system or a floppy disk, hard disk, CD ROM, DVD ROM, or other computer readable medium which is read from and written to by a magnetic, optical, or other reading and writing system that is coupled to the processor 28, can be used for the memory 30. In this particular example, the memory 30 in the node controller computing device 12(1) further includes a LIF manager module 34 with programmed instructions for managing one or more LIF operations, such as those illustrated and described herein, although the memory could comprise other types and/or numbers of modules, programmed instructions and/or other data.

[0016] The communication interface 32 in the node controller computing device 12(1) operatively couples and communicates with the other node controller computing device 12(2), the back-end storage server device 14 and one or more of the client computing devices 16(1)-16(n) which are all coupled together by one or more of the communication networks 20, although other types and numbers of communication networks or systems with other types and numbers of connections and configurations to other devices and elements. By way of example only, the communication networks 20 can use TCP/IP over Ethernet and industry-standard protocols, including NFS, CIFS, SOAP, XML, LDAP, SCSI, and SNMP, although other types and numbers of communication networks, can be used. The communication networks 20 in this example may employ any suitable interface mechanisms and network communication technologies, including, for example, any local area network, any wide area network (e.g., Internet), teletraffic in any suitable form (e.g., voice, modem, and the like), Public Switched Telephone Network (PSTNs), Ethernet-based Packet Data Networks (PDNs), and any combinations thereof and the like.

[0017] In this particular example, each of the client computing devices 16(1)-16(n) may run applications that may provide an interface to make requests for and receive content

and/or other network resources hosted by one or more of the back-end storage server devices 16(1)-16(n) via one or more of the node controller computing devices 12(1)-12(2) in storage cluster 11.

[0018] The back-end storage server device 14 may store and provide data or other network resources in response to requests from the client computing devices 16(1)-16(n) via the one or more of the communication networks 20, for example, although other types and numbers of storage media in other configurations could be used. In particular, the back-end storage server devices 14 may comprise various combinations and types of storage hardware and/or software and represent a system with multiple network server devices in a data storage pool, which may include internal or external networks. Additionally, various network processing applications, such as CIFS applications, NFS applications, HTTP Web Network server device applications, and/or FTP applications, may be operating on the back-end storage server device 14 and transmitting data (e.g., files or web pages) in response to requests from the client computing devices 16(1)-16(n).

[0019] The back-end storage server device 14 and each of the client computing devices 16(1)-16(n) may include a processor, a memory, and a communication interface, which are coupled together by a bus or other link, although other numbers and types of devices and/or nodes as well as other network elements could be used.

[0020] Although the exemplary environment 10 with the node management computing device 12, a storage cluster with the node controller computing devices 12(1)-12(2), back-end storage server device 14, client computing devices 16(1)-16(n) and the communication networks 20 are described and illustrated herein, other types and numbers of systems, devices, components, and/or elements in other configurations or topologies can be used. It is to be understood that the systems of the examples described herein are for exemplary purposes, as many variations of the specific hardware and software used to implement the examples are possible, as will be appreciated by those skilled in the relevant art(s).

[0021] In addition, two or more computing systems or devices can be substituted for any one of the systems or devices in any example. Accordingly, principles and advantages of distributed processing, such as redundancy and replication also can be implemented, as desired, to increase the robustness and performance of the devices and systems of the examples. The examples may also be implemented on computer system(s) that extend across any suitable network using any suitable interface mechanisms and traffic technologies, including by way of example only teletraffic in any suitable form (e.g., voice and modem), wireless traffic media, wireless traffic networks, cellular traffic networks, G3 traffic networks, Public Switched Telephone Network (PSTNs), Packet Data Networks (PDNs), the Internet, intranets, and combinations thereof.

[0022] The examples also may be embodied as a non-transitory computer readable medium having instructions stored thereon for one or more aspects of the present technology as described and illustrated by way of the examples herein, as described herein, which when executed by the processor, cause the processor to carry out the steps necessary to implement the methods of this technology as described and illustrated with the examples herein.

[0023] An example of a method for prioritizing failover of logical interfaces (LIFs) during a node outage will now be illustrated and described with reference to FIGS. 1-4. In step 100, each of the node controller computing devices 12(1)-12(2) may collect, store, and/or share data on a number of logical interfaces (LIFs) and a number of active connections on LIFs on each of the node controller computing devices 12(1)-12(2), although other types and/or amounts of data for prioritizing failover may be collected, stored, and/or shared. In this particular example, the number of logical interfaces (LIFs) and a number of active connections on LIFs on each of the node controller computing devices 12(1)-12(2) is obtained periodically, such as every 60-120 seconds, although other time periods and/or other manners for obtaining these numbers, such as needed or requested, can be used. Additionally by way of example only, the number of active connection can be determined by each of the node controller computing devices 12(1)-12(2) by obtaining a network statistics (netstat) output which includes data on the active number of connections on each of the LIFs. Further by way of example only, in FIG. 4 an example of a stored table with a number of LIFs and a number of active connections for node controller computing device 12(1) or “node 0” and for node controller computing device 12(2) or “node 1” is illustrated. In this particular illustrative example, node controller computing device 12(1) or “node 0” has three LIFs: datalif1a with three (3) active connections); datalif2a with six (6) active connections); and datalif3a with five (5) active connections) and node controller computing device 12(2) or “node 1” has three LIFs: datalif1b with three (3) active connections); datalif2b with three (3) active connections); and datalif3b with five (5) active connections).

[0024] In step 102, each of the node controller computing devices 12(1)-12(2) may assign one of a plurality of priority weight values for each of the identified LIFs based on the identified number of active connections, although other approaches for assigning one of a plurality of weight values may be used. Additionally, each of the node controller computing devices 12(1)-12(2) may assign, store, and/or share the assigned priority weight values, although again other types and/or amounts of data to manage failover may be obtained, stored, and/or shared. By way of example only, in node controller computing device 12(1) or “node 0” the LIF “datalif2a” with six (6) active connections would have a higher assigned priority weight value than LIF “datalif3a” with five (5) active connections which would have a higher assigned priority weight value than LIF “datalif1a” with three (3) active connections. Additionally by way of example only, in node controller computing device 12(2) or “node 1” the LIF “datalif3b” with five (5) active connections) would have a higher assigned priority weight value than LIF “datalif1b” and LIF “datalif2b” which each have three (3) active connections. As illustrated in this second example, one or more other factors relating to the LIFs may need to be utilized by node controller computing device 12(2) or “node 1” to prioritize failover between LIF “datalif1b” and LIF “datalif2b”, such as prioritizing alphabetically or a randomized selection of which to failover in case of a tie by way of example only.

[0025] In step 104, each of the node controller computing devices 12(1)-12(2) may monitor for a node outage in one of the node controller computing devices 12(1)-12(2). By way of example only, the node outage may be an unexpected outage or a scheduled outage. If in step 104, one of the node

controller computing devices 12(1)-12(2) does not identify a node outage in another one of the node controller computing devices 12(1)-12(2), then the No branch is taken back to step 100 collect and store data on a number of logical interfaces (LIFs) and a number of active connections on LIFs on each of the node controller computing devices 12(1)-12(2) as described earlier.

[0026] If in step 104, one of the node controller computing devices 12(1)-12(2) does identify a node outage in another one of the node controller computing devices 12(1)-12(2), then the Yes branch is taken to step 106. In step 106, the one of the node controller computing devices 12(1)-12(2) that detected the outage may manage failover of the LIFs from the another one of the node controller computing devices 12(1)-12(2) based on the assigned one of a plurality of priority weight values for each of the identified LIFs. By way of example only, if node controller computing device 12(1) or “node 0” had the identified node outage, then the node controller computing device 12(2) or “node 1” would failover the LIF “datalif2a” with six (6) active connections), then the LIF “datalif3a” with five (5) active connections which would have a higher assigned priority weight value than LIF “datalif1a” with three (3) active connections based on the assigned one of a plurality of priority weight values as described earlier. Additionally by way of another example only, if the node controller computing device 12(2) or “node 1” had the identified node outage, then the node controller computing device 12(1) or “node 0” would failover the LIF “datalif3b” with five (5) active connections) followed by either LIF “datalif1b” or the LIF “datalif2b” depending in the assigned priority weight value based on the other one or more factors as described earlier.

[0027] In these examples a substantially reduced number of LIFs is illustrated for ease of discussion. If one of the node controller computing devices 12(1) or 12(2) with the node outage was over provisioned for the number of LIFs above a stored threshold, then any of the LIFs with a lower assigned priority weight value, such as a lower priority weight value assigned based on a lower number of active connections or no current active connections, may not be failed over to the other one of the node controller computing devices 12(1) or 12(2). By way of example only, if node controller computing devices 12(1) and 12(2) could each maintain 256 LIFs and had a maximum of stored provisioning threshold of 128 LIFs, but were configured to have 130 LIFs in violation of the maximum of stored provisioning threshold, then during a failover two of the LIFs with the fewest or no active connections resulting a the lowest two assigned priority weight values would not be failed over.

[0028] Accordingly, as illustrated and described by way of the examples herein, this technology provides methods, devices, and non-transitory computer readable media that effectively prioritize failover of logical interfaces (LIFs) during a node outage. With this technology, LIFs with active connections are given a higher priority to failover rather than other LIFs without any connections or with a lower number of active connections. As a result, with this technology during node outages, including unexpected node outages or even during scheduled node outages for upgrade events, LIFs with active connections may either be disruption free or at least any impact on disrupting active connections during the node outage is minimized.

[0029] Having thus described the basic concept of the invention, it will be rather apparent to those skilled in the art

that the foregoing detailed disclosure is intended to be presented by way of example only, and is not limiting. Various alterations, improvements, and modifications will occur and are intended to those skilled in the art, though not expressly stated herein. These alterations, improvements, and modifications are intended to be suggested hereby, and are within the spirit and scope of the invention. Additionally, the recited order of processing elements or sequences, or the use of numbers, letters, or other designations therefore, is not intended to limit the claimed processes to any order except as may be specified in the claims. Accordingly, the invention is limited only by the following claims and equivalents thereto.

What is claimed is:

1. A method for prioritizing failover of logical interfaces (LIFs), the method comprising:

identifying, by one of a plurality of node computing devices, an outage on another one of the plurality of node computing devices in a storage cluster;

determining, by the one of a plurality of node computing devices, when the another one of the plurality of node computing devices with the identified outage has provisioned a number of logical interfaces (LIFs) above a stored provisioning threshold; and

managing, by the one of a plurality of node computing devices, failover of the LIFs on the another one of the plurality of node computing devices with the outage in an order based on at least one of a plurality of priority weight values assigned to each of the LIFs.

2. The method as set forth in claim **1** further comprising periodically obtaining and storing, by the one of a plurality of node computing devices, the number of logical interfaces LIFs on at least the another one of the plurality of node computing devices.

3. The method as set forth in claim **1** further comprising determining, by the one of a plurality of node computing devices, the assigned one of the plurality of priority weight values for each of the LIFs on at least the another one of the plurality of node computing devices.

4. The method as set forth in claim **3** wherein the determining the assigned one of the plurality of priority weight values for each of the LIFs on at least the another one of the plurality of node computing devices further comprises:

identifying, by the one of a plurality of node computing devices, a number of active connections for each of the LIFs on the another one of the plurality of node computing devices; and

assigning, by the one of a plurality of node computing devices, one of the plurality of priority weight value for each of the LIFs on the another one of the plurality of node computing devices based on the identified number of active connections.

5. The method as set forth in claim **1** wherein the identified outage is one of an unexpected outage or a scheduled outage.

6. A node controller computing device, comprising:

one or more processors;

a memory coupled to the one or more processors which are configured to be capable of executing programmed instructions stored in the memory to and that comprise:

identify an outage on another one of the plurality of node computing devices in a storage cluster;

determine when the another one of the plurality of node computing devices with the identified outage has provisioned a number of logical interfaces (LIFs) above a stored provisioning threshold; and

manage failover of the LIFs on the another one of the plurality of node computing devices with the outage in an order based on at least one of a plurality of priority weight values assigned to each of the LIFs.

7. The device as set forth in claim **6** wherein the memory coupled to the one or more processors which are configured to be capable of executing programmed instructions stored in the memory to and that further comprises periodically obtain and store the number of logical interfaces LIFs on at least the another one of the plurality of node computing devices.

8. The device as set forth in claim **6** wherein the memory coupled to the one or more processors which are configured to be capable of executing programmed instructions stored in the memory to and that further comprises determine the assigned one of the plurality of priority weight values for each of the LIFs on at least the another one of the plurality of node computing devices.

9. The device as set forth in claim **8** wherein the memory coupled to the one or more processors which are configured to be capable of executing programmed instructions for the determine the assigned one of the plurality of priority weight values stored in the memory to and that further comprises:

identify a number of active connections for each of the LIFs on the another one of the plurality of node computing devices; and

assign one of the plurality of priority weight value for each of the LIFs on the another one of the plurality of node computing devices based on the identified number of active connections.

10. The device as set forth in claim **6** wherein the identified outage is one of an unexpected outage or a scheduled outage.

11. A non-transitory computer readable medium having stored thereon instructions for prioritizing failover of logical interfaces (LIFs) comprising machine executable code which when executed by a processor, causes the processor to perform steps comprising:

identifying an outage on another one of the plurality of node computing devices in a storage cluster;

determining when the another one of the plurality of node computing devices with the identified outage has provisioned a number of logical interfaces (LIFs) above a stored provisioning threshold; and

managing failover of the LIFs on the another one of the plurality of node computing devices with the outage in an order based on at least one of a plurality of priority weight values assigned to each of the LIFs.

12. The medium as set forth in claim **11** further comprising periodically obtaining and storing the number of logical interfaces LIFs on at least the another one of the plurality of node computing devices.

13. The medium as set forth in claim **11** further comprising determining the assigned one of the plurality of priority weight values for each of the LIFs on at least the another one of the plurality of node computing devices.

14. The medium as set forth in claim **13** wherein the determining the assigned one of the plurality of priority

weight values for each of the LIFS on at least the another one of the plurality of node computing devices further comprises:

identifying a number of active connections for each of the LIFs on the another one of the plurality of node computing devices; and

assigning one of the plurality of priority weight value for each of the LIFs on the another one of the plurality of node computing devices based on the identified number of active connections.

15. The medium as set forth in claim **11** wherein the identified outage is one of an unexpected outage or a scheduled outage.

* * * * *