



(51) International Patent Classification:
G06F 3/06 (2006.01)

(21) International Application Number:
PCT/US2015/045127

(22) International Filing Date:
13 August 2015 (13.08.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2612/CHE/2015 25 May 2015 (25.05.2015) IN

(71) Applicant: **HEWLETT PACKARD ENTERPRISE DEVELOPMENT LP** [US/US]; 11445 Compaq Center Drive West, Houston, TX 77070 (US).

(72) Inventors: **CHIRUMAMILLA, Narendra**; Sy.No.192, Whitefield Road, Mahadevapura Post, Bangalore 560048 (IN). **MAHESH, Keshetti**; Sy.No.192, Whitefield Road, Mahadevapura Post, Bangalore 560048 (IN). **NAYAKA B, Govindaraja**; Sy.No.192, Whitefield Road, Mahadevapura Post, Bangalore 560048 (IN). **MOHANTA, Taranisen**;

Sy.No.192, Whitefield Road, Mahadevapura Post, Bangalore 560048 (IN). **BASIREDDY, Ranjith Reddy**; Sy.No.192, Whitefield Road, Mahadevapura Post, Bangalore 560048 (IN).

(74) Agents: **ORTEGA, Arthur S.** et al.; Hewlett Packard Enterprise, 3404 E. Harmony Road, Mail Stop 79, Fort Collins, CO 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, 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, SA, 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, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,

[Continued on next page]

(54) Title: STORAGE MANAGEMENT

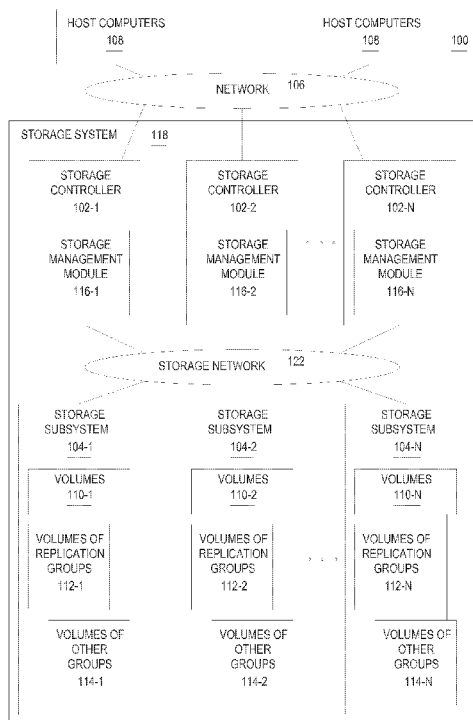


FIG. 1

(57) Abstract: In one example, determine first Input Output Requests Per Second (IOPs) values of IOPs of first groups of volumes of data replication groups assigned to controllers, determine second IOPs values comprising IOPS of second groups of other volumes not of data replication groups assigned to the controllers, determine total IOPs values comprising sum of first and second IOPs values of the controllers, determine usage values based on the total IOPs values of controllers. If difference of usage values of one of controller is greater than a threshold amount, partition volumes of volume sets of first and second groups into first and second partition groups, the difference between sums of the first and second IOPs values is minimum, and volumes of the first groups not transferable to second groups, and reassign ownership of first partition to the first controller and the second partition to the second controller.



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

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

Published:

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

Declarations under Rule 4.17:

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

STORAGE MANAGEMENT

BACKGROUND

[0001] Computer systems include host computers that communicate with storage systems. The storage systems may include storage devices to store data for later retrieval. The host computers may send commands to the storage systems which store data at the storage devices as well as retrieve data from the storage devices.

BRIEF DESCRIPTION OF THE DRAWINGS

[0002] Examples are described in the following detailed description and in reference to the drawings, in which:

[0003] Fig. 1 depicts an example system for storage management in accordance with an example of the techniques of the present application;

[0004] Figs. 2A through 2B depict an example flow chart of processes for storage management in accordance with an example of the techniques of the present application;

[0005] Figs. 3A through 3H depict example diagrams of storage management in accordance with an example of the techniques of the present application; and

[0006] Fig. 4 depicts an example block diagram showing a non-transitory, computer-readable medium that stores instructions for storage management in accordance with an example of the techniques of the present application.

DETAILED DESCRIPTION

[0007] Computer systems include host computers that communicate with storage systems. The storage systems may include storage subsystems with storage devices having storage volumes to store data for later retrieval. The host computers may send commands such as InputOutput (IO) requests to the storage systems to have the storage systems store data at the storage devices as well as retrieve data from the storage devices. Computer systems may include storage controllers to interface between host computers and storage systems

[0008] Some storage systems may be configured to have several storage controllers. In one example, the storage systems may be configured as a dual storage controller arrangement with a first (master) controller that may interact with a second (slave) controller. The storage systems may provide storage volumes to allow storage controllers to store data for later retrieval by the storage controllers. In one example, ownership of volumes may be assigned or distributed in an even manner between storage

- 2 -

controllers to provide a balanced IO workload. In this case, assignment of the volume ownership may be evenly distributed between storage controllers. However, IO workload from IO requests from host computers may not be evenly distributed between the controllers. When a volume owned by a storage controller is accessed, all the IO activity directed to that volume (even via other the controller) is processed by the controller that owns the volume. For example, if host computers and/or applications using the storage systems are mostly accessing the volumes owned by a master controller, then IO processing may occur mostly in the master controller which may use (usage or usage performance) most of the storage controller resources (such as memory, cache, IO bandwidth etc.). In contrast, the resources of the other controller (slave controller) may be underutilized and thus may decrease the performance of the storage system.

[0009] In some examples, the present application discloses techniques which may address this situation and help increase the performance of the storage system. In one example, the techniques provide a load balancing process which may help balance the IO load or workload on volumes on both the controllers while considering volumes of data replication groups and volumes that are not part of the data replication groups. In one example, the process may provide for volume ownership between the two storage controllers (master and slave) to be balanced based on the IO load on individual controllers while considering data replication groups. The data replication groups may include several characteristics or properties. For example, volumes that are associated with data replication groups may be configured to be connected or online to one storage controller for failover purposes (either master or slave controller). For example, if a volume (that is associated with a data replication group) changes ownership from one controller to another controller, then all the volumes associated with the data replication group (including data replication log) which are part of the same data replication changes ownership to the new controller. That is, data replication groups and associated volumes are connected or online to one controller. In another example, if a volume of a data replication groups fails, then the whole or total data replication group fails. In this manner, these techniques consider data replication groups as part of the load balancing process which may improve performance of storage controllers, use of resources and IO response time.

[0010] In one example, the present application discloses storage management techniques which include providing a storage management module to perform a load balancing process. The storage management module is configured to determine first Input Output Requests Per Second (IOPs) values comprising IOPs of first volume groups associated with volumes of data replication groups assigned ownership to first and second

- 3 -

storage controllers. The storage management module is configured to determine second IOPs values comprising IOPS of second volume groups associated with all other volumes not associated with data replication groups assigned ownership to the controllers. The storage management module is configured to determine total IOPs values comprising sum of the first IOPs values and second IOPs values of the controllers. The storage management module is configured to determine usage values based on the total IOPs values of respective controllers.

[00011] The storage management module is configured to check if a difference of the usage values of one of the controllers is greater than a threshold amount. If storage management module determines that there is difference of the usage values of one of the controllers greater than a threshold amount, then it proceeds to partition volumes of volume sets comprising the first volume groups and second volume groups into first and second partition volume groups. In one example, the difference between sums of the first IOPs values and second IOPs values is minimum. In one example, the volumes of the first volume groups are not assigned ownership to the second volume groups. The storage management module is configured to reassign ownership of the first partition volume groups to the first controller and the second partition volume groups to the second controller.

[00012] The storage management module may configure the volumes of the first volume groups to be associated with a data replication process. In one example, the data replication process may include a process of copying volumes of data replication groups from a first storage controller (or storage system) designated as a source controller to volumes at a second controller (or storage system) designated as a destination controller to provide consistency between the source controller and the destination controller in event of failure of the source controller. For example, to protect the data in volumes from loss, the data replication process may replicate data in a storage system implemented on computing device(s) at a first physical location to a storage system implemented on computing device(s) at a second physical location. In such examples, if a failure of volumes occurs or other event prevents retrieval of at least some portion of the data of the storage system at the first location, it may be possible to retrieve the data from the storage system at the second location.

[00013] In this manner, in some examples, the present application discloses techniques to provide load balancing which may help increase the performance of the storage system.

- 4 -

[00014] Fig. 1 depicts an example system 100 for storage management in accordance with an example of the techniques of the present application. The system 100 includes a storage system 118 configured to communicate with host computers 108 over a network 106.

[00015] The storage system 118 includes storage controllers 102 to communicate with host computers 108 over the network 106 and communicate with storage subsystems 104 over a storage network 122.

[00016] The storage controllers 102 (102-1 through 102-n) include respective storage management modules 116 (116-1 through 116-n) to provide storage management functionality as described herein. In one example, storage management modules 116 provide functionality to process IO requests from host computers 108 to access data stored on storage volumes 110 (110-1 through 110-n) of storage subsystems 104 (104-1 through 104-1). The storage volumes 110 may be defined as any electronic means to store data for later retrieval such as storage devices. The storage volumes 110 may be logical units of data that can be defined across multiple storage devices of storage subsystems 104. The IO requests may include requests to read data from volumes 110 and requests to write data to the volumes. The storage system 118 may assign ownership of storage volumes 110 to storage controllers 102. The storage controller that is assigned to control a given volume is referred to as the owner of the given storage volume. In other words, a storage controller that is an owner of a given storage volume performs control of access made to the given storage volume (while other storage controllers do not control access to the given storage volume). The storage system 118 configures storage volumes 110 as volumes of data replication groups 112 (112-1 through 112-n) and as volumes of other groups 114 (114-1 through 114-n) which are not part of the data replication groups. The storage system 118 may configure data replication groups 112 to be part of a data replication process as explained below in further detail.

[00017] In one example, to illustrate operation of the techniques of the present application, storage system 118 may configure the system to have a cluster of storage controllers. In an example, the storage system may configure the system as a dual controller system that includes a first storage controller 102-1 and a second storage controller 102-2 where each of the storage controllers perform functionality of respective storage management modules 116-1 and 116-2. The storage management modules 116 (116-1 and 116-2) may be configured to determine first Input Output Requests Per Second (IOPs) values comprising IOPs of first volume groups associated with volumes 110 of data replication groups 112 assigned ownership to first and second storage controllers. The

- 5 -

storage management module 116 may be configured to determine second IOPs values comprises IOPs of second volume groups associated with all other volumes 114 not associated with data replication groups assigned ownership to the controllers. The storage management module 116 may be configured to determine total IOPs values comprising sum of the first IOPs values and second IOPs values of the controllers. The storage management module 116 may be configured to determine usage values based on the total IOPs values of respective controllers. The IOPs values may be defined as the number of IO requests processed by each of the storage controllers over a period of time. The usage values may be defined as usage performance of storage controllers which include actual performance compared to maximum performance. For example, a storage controller may be configured to perform or have resources to handle a maximum 10,000 IOPs and the actual performance or use is 6,700 IOPs. In this case, usage value is $67\% = (6,700/10,000)$. The host computers 108 send to storage controllers 102 the IO requests which include requests to read data from volumes 110 and write data to the volumes. It should be understood that other measures other than usage or use of IOPs, such as IO latency measurements, may be employed to determine usage or performance or activity of the storage controllers.

[00018] The storage management module 116 may be configured to check if a difference of the usage values of one of the controllers is greater than a threshold amount. The usage values may be defined as the amount of use or utilization of resources by the storage controllers. For example, usage values may include IOPs values of storage controllers. To illustrate, assume that a storage system is configured with a first controller 102-1 and a second controller 102-2 and the system is configured with 10 volumes Vol1 through Vol10 as shown in Fig. 3F. Since volume distribution across controllers is not based on the IO load, it may so happen that first controller performance utilization is very high compared to second controller's performance utilization. The first controller is undergoing more IO activity or workload when compared with second controller. This way second controller processing power is underutilized and may result in overall system performance degradation. Further assume that maximum IO processing power per controller is 10,000 IOPs. The total IOPS processed by the storage system is $13,300 = (7,700 + 5,600)$. This is $66.5\% = (13,300/20,000)$ of the total storage system performance. The storage system should be ready to accept 33.55% more IO requests. However, first controller's performance utilization or usage value is at $77\% = (7,700/10,000)$ whereas second controller's performance utilization or usage value is only $56\% = (5,600/10,000)$. In this case, second controller's processing power is underutilized.

- 6 -

Furthermore, if host computers 108 desire to increase IO requests to the volumes owned by first controller then it may fail due to resource limitation in the first controller, even though there is plenty of free resources available in second controller. This example was to illustrate usage or utilization measurements but does not consider data replication groups.

[00019] If storage management module 116 determines that there is difference of the usage values of one of the controllers is greater than a threshold amount, then it proceeds to partition volumes of volume sets comprising the first volume groups and second volume groups into first and second partition volume groups. The threshold amount may be any amount specified by storage system based on the requirements or criteria of the system. For example, the threshold amount or value may be 10% which is a relatively low number which triggers the partition process. Continuing with the example of Fig. 3F, without considering data replication groups, first controller 102-1 compares its usage (77%) with usage (56%) of second controller 102-2 to determine whether the difference is greater than a threshold amount such as 10%. In one example, the partition process may be such that the difference between sums of the first IOPs values and second IOPs values is to be a minimum. In one example, the partition process may be such that volumes of the first volume groups are not transferable (change of ownership) to the second volume groups. The storage management module 116 may be configured to reassign ownership of the first partition volume groups to the first controller and the second partition volume groups to the second controller, as show in Fig. 3H as explained below in further detail. The storage management module 116 may be configured to configure the volumes of the first volume groups to be associated with a data replication process. In one example, the data replication process may include a process of copying volumes of data replication groups 112 from a first storage controller (or storage system) designated as a source controller to volumes at a second controller (or storage system) designated as a destination controller to provide consistency between the source controller and the destination controller in event of failure of the source controller.

[00020] The storage management module 116 may be configured to communicate with host computers 108 to facilitate access of host computers to data stored in volumes 110 of storage subsystems 104. In one example, first controller 102-1 may be configured to respond to receipt from a host computer 108 of requests to access volumes owned by the first controller. In this case, first storage controller 102-1 may be configured to respond with status of the requests to the host computer. In the case of a read request, first controller 102-1 responds with the requested data from the relevant volume and the status

- 7 -

of the request. In the case of a write request, first controller 102-1 writes the data from host computer to the relevant volume and respond with the status of the request. In another example, first controller 102-1 may be configured to handle a receipt from a host computer 108 a request to access volumes owned by a second controller 102-2. In this case, in response to receipt of the request, first controller 102-1 may cause second storage controller 102-2 to access the volumes and have the second controller send a response with status of the request to the first controller which then in turn responds with the status to the host computer.

[00021] As explained above, storage management module 116 may be configured to configure volumes 110 of data replication groups 112 to be associated with a data replication process. In one example, the data replication process may include a remote replication process which involves continuous copying of data from selected volumes of a source (local) storage system to related volumes on a destination (remote) storage system. In one example, applications or programs may continue to execute on system 100 while data is being replicated in the background. The data replication groups 112 may be defined as a logical group of volumes in a remote replication relationship between two storage systems. In one example, host computers 108 may issue commands to storage system 118 to write data to the volumes associated with data replication groups in a source storage system. In this case, the storage system 118 may copy the data to the volumes in the destination storage system. The storage system 118 may be configured to maintain I/O order across volumes in data replication groups 112. In this manner, this may help ensure data consistency of data on the destination storage system in the event of a failure of the source storage system. Failure events associated with storage systems may be defined as conditions which may impair or prevent the ability of a system to retrieve data from the system such that data may be corrupt. The volumes 110 associated with data replication groups 112 may fail over together and share a write history log to help preserve write order within the data replication groups. In one example, the write history log may include a volume that stores the write data of a data replication group. The log may be created automatically when the data replication group is created. The log may be used to preserve write order within a data replication group. The storage system 118 configures volumes 110 associated with data replication groups 112 to be connected (online) to one storage controller. That is, data replication groups 112 and associated volumes ownership are associated with one storage controller. In one example, if a volume that is associated or part of a data replication group performs a failover operation to another storage

- 8 -

controller then the complete data replication group is configured to failover to that storage controller.

[00022] The storage management module 116 may be configured to perform the data replication process with two replication write modes of operation. For example, storage management modules 116 may employ an asynchronous mode where a source storage system acknowledges I/O completion before data is replicated on the destination storage system. In another example, storage management modules 116 may employ a synchronous mode where a source storage system acknowledges I/O completion after the data is cached on both the local and destination storage systems.

[00023] The storage management modules 116 may be configured to manage volumes as data replication groups 112 and volumes of other groups 114. In one example, data replication groups 112 may include logical group of volumes or Logical Unit Numbers (LUNs). The data replication groups 112 may be configured to be connected or online with respect to a particular controller. That is, ownership of the group and its associated volumes are configured to be associated with only one controller. In one example, volumes of a data replication group failover together as single entity or unit. That is, if a volume associated with a data replication group performs a failover to another storage controller the complete data replication group and its member volumes including the data replication log perform a failover to that storage controller. In this case, even though a single volume failover happens, complete data replication group will perform a failover and ownership of entire data replication group gets changed. In case of non-data replication groups 114 (i.e. regular volumes/normal volumes/ ordinary volumes) that performs failover, ownership of the failover volume will get changed and other volumes in the system will not be impacted. However, in case of data replication groups, ownership of all the member volumes of the data replication get changed.

[00024] The storage management modules 116 may be configured to manage and use metadata related to load balancing during processing. For example, storage management modules 116 may determine the first IOPs values and second IOPs values of the controllers based on metadata of the controllers and then reset the metadata of the controllers after the determination of the IOPs values.

[00025] In this manner, in some examples, the techniques of the present application may provide storage management techniques for load balancing to help improve the performance of storage systems.

[00026] The network 106 may include any means of electronic or data communication between host computers 108 and storage system 118. The network 106

- 9 -

may include a local area network, Internet and the like. The storage network 122 may be any means of electronic or data communication between storage controller 102 and storage subsystems 104. The storage network 122 may include Fibre Channel network, SCSI (Small Computer System Interface) link, Serial Attached SCSI (SAS) link and the like. The storage controllers 102 may be coupled through different sets of network components contained in storage network 122 to corresponding different storage subsystems. The storage network 122 may include expanders, concentrators, routers, and other communications devices. In one example, a storage controller 102 may be coupled over a first set of network components to one storage subsystem, while another storage controller can be coupled by a different set of network components to another storage subsystem.

[00027] The storage controllers 102 may include any electronic device or means of processing requests from host computers 108 to access data from volumes 110. The storage controllers 102 may include network interfaces to allow the storage controllers to communicate over network 106 with host computers 108 and with other storage controllers of system 118. In another example, instead of communicating over network 106, storage controllers 102 may communicate with each other over storage network 122, or through another network.

[00028] The volumes 110 of storage subsystems 104 may be configured as Redundancy Array of Inexpensive Disks (RAID) volumes. A RAID volume may be deployed across multiple storage devices of storage subsystems 104 to provide redundancy. The redundancy can be based on mirroring of data, where data in one storage device is copied to a mirror storage device (which contains a mirror copy of the data). RAID-1 is an example of a mirroring redundancy scheme. In this arrangement, if an error causes data of the source storage device to be unavailable, then the mirror storage device can be accessed to retrieve the data. Another type of redundancy is parity-based redundancy, where data is stored across a group of storage devices, and parity information associated with the data is stored in another storage device. If data within any storage device in the group of storage devices were to become inaccessible (due to data error or storage device fault or failure), the parity information can be accessed to reconstruct the data. Examples of parity-based redundancy schemes include RAID-5 and RAID-6 schemes. If used with RAID volumes, storage controllers 102 in system 100 are RAID controllers. Although reference is made to RAID volumes, it should be understood that other types of volumes can be employed in other examples.

- 10 -

[00029] The storage subsystems 104 include storage volumes 110. As explained above, volumes 110 may be logical units of data that can be defined on one or more storage devices, including an array of storage devices in storage subsystems. The storage device may refer to a physical storage element, such as a disk-based storage element (e.g., hard disk drive, optical disk drive, etc.) or other type of storage element (e.g., semiconductor storage element). In one example, multiple storage devices within a storage subsystem can be arranged as an array configuration.

[00030] The system 100 of Fig. 1 shows example storage controllers 102 and should be understood that other configurations may be employed to practice the techniques of the present application. For example, system 100 may be configured to communicate with a plurality of storage controllers 102 and with a plurality of host computers 108 and with a plurality of storage subsystems 104. The components of system 100 may be implemented in hardware, software or a combination thereof. For example, the functionality of the components of system 100 may be implemented using technology related to Personal Computers (PCs), server computers, tablet computers, mobile computers and the like. The storage controllers 102, host computers 108 and storage systems 118 may communicate using any communications means such as Fibre Channel, Ethernet and the like.

[00031] Fig. 1 shows a storage system 118 to provide storage management. The storage system 118 may include computer-readable storage medium comprising (e.g., encoded with) instructions executable by a processor to implement functionalities described herein in relation to FIG. 1. In some examples, the functionalities described herein in relation to instructions to implement storage management module 116 functions, and any additional instructions described herein in relation to storage medium, may be implemented as engines or modules comprising any combination of hardware and programming to implement the functionalities of the modules or engines, as described below. The functions of storage system 118 may be implemented by computing devices which may be a server, blade enclosure, desktop computer, laptop (or notebook) computer, workstation, tablet computer, mobile phone, smart device, or any other processing device or equipment including a processing resource. In examples described herein, a processor may include, for example, one processor or multiple processors included in a single computing device or distributed across multiple computing devices.

[00032] Figs. 2A through 2B depict an example flow chart 200 of a process for storage management in accordance with an example of the techniques of the present application. To illustrate operation, it may be assumed that system 100 includes storage

- 11 -

system 118 with a first storage controller 102-1 and a second storage controller 102-2 configured in a dual controller configuration. It may also be assumed that storage controllers 102 are configured to communicate with host computers 108 over network 106. It may also be assumed that storage controllers 102-1 and 102-2 are configured to communicate with first storage subsystem 104-1 and second storage subsystem 104-2 over storage network 122. The storage controllers 102 include storage management modules 116 to implement the techniques of the present application and functionality described herein.

[00033] It should be understood the process depicted in Figs. 2A through 2B represents generalized illustrations, and that other processes may be added or existing processes may be removed, modified, or rearranged without departing from the scope and spirit of the present application. In addition, it should be understood that the processes may represent instructions stored on a computer-readable storage medium that, when executed, may cause a processor to respond, to perform actions, to change states, and/or to make decisions. Alternatively, the processes may represent functions and/or actions performed by functionally equivalent circuits like analog circuits, digital signal processing circuits, Application Specific Integrated Circuits (ASICs), or other hardware components associated with the system. Furthermore, the flow charts are not intended to limit the implementation of the present application, but rather the flow charts illustrate functional information to design/fabricate circuits, generate software, or use a combination of hardware and software to perform the illustrated processes.

[00034] The process 200 may begin at block 202, where storage controllers 102 determine first IOPs values comprising IOPs of first volume groups associated with volumes of data replication groups 112 assigned ownership to first and second storage controllers. In one example, first storage management module 116-1 determines first IOPs values comprising IOPs of first volume groups associated with volumes of data replication groups 112-1 which are assigned ownership to the first storage controller. In a similar manner, second storage management module 116-2 determines first IOPs values comprising IOPs of first volume groups associated with volumes of data replication groups 112-2 which are assigned ownership to the second storage controller. Processing proceeds to block 204.

[00035] At block 204, storage controllers 102 determine second IOPs values comprising IOPs of second volume groups associated with all other volumes 114 not associated with data replication groups which are assigned ownership to the controllers. In one example, first storage management module 116-1 determines second IOPs values

- 12 -

comprising IOPS of second volume groups associated with all other volumes 114-1 not associated with data replication groups assigned ownership to the first storage controller. In a similar manner, second storage management module 116-2 determines second IOPS values comprising IOPS of second volume groups associated with all other volumes 114-2 not associated with data replication groups which are assigned ownership to the second storage controller. Processing proceeds to block 206.

[00036] At block 206, storage controllers 102 determine total IOPS values comprising sum of the first IOPS values and second IOPS values of the controllers. In one example, first storage management module 116-1 determines sum of the first IOPS values and second IOPS values of the first storage controller. In a similar manner, second storage management module 116-2 determines sum of the first IOPS values and second IOPS values of the second storage controller. Processing proceeds to block 208.

[00037] At block 208, storage controllers 102 determine usage values based on the total IOPS values of respective controllers. In one example, first storage management module 116-1 determines usage values based on the total IOPS values of the first storage controller. In a similar manner, second storage management module 116-2 determines usage values based on the total IOPS values of the second controller. Processing proceeds to block 210.

[00038] At block 210, storage controllers 102 determine whether difference of the usage values of one of the controllers is greater than a threshold amount. In one example, first storage management module 116-1 determines whether difference of the usage values of the first storage controller is greater than a threshold amount. In a similar manner, second storage management module 116-2 determines whether difference of the usage values of the second storage controller is greater than a threshold amount. If the storage management modules 116 determine a difference of the usage values is greater than a threshold amount, then processing proceeds to block 212. On the other hand, if storage management modules 116 determine a difference of the usage values is not greater than a threshold amount, then processing proceeds to end block to perform some other processing or back to the begin block for further storage controller processing.

[00039] At block 212, storage controllers 102 partition volumes of volume sets comprising first volume groups and second volume groups into first and second partition volume groups. In one example, storage management modules 116 perform the partition such that the difference between sums of the first IOPS values and second IOPS values is minimum. In one example, storage management modules 116 perform the partition

- 13 -

function such that volumes of the first volume groups are not ownership assignable or transferable to the second volume groups. Processing proceeds to block 214.

[00040] At block 214, storage controllers 102 reassign ownership of the first partition volume groups to first controller 102-1 and the second partition volume groups to second controller 102-2. In one example, storage management modules 116 reassign ownership of the first partition volume groups to first controller 102-1 and the second partition volume groups to second controller 102-2. Processing proceeds to block 216.

[00041] At block 216, storage controllers 102 configure the volumes of first volume groups to be associated with a data replication process. For example, first storage management module 116-1 configures the volumes of first volume groups to be associated with a data replication process. In a similar manner, second storage management module 116-2 configures the volumes of first volume groups to be associated with a data replication process. In one example, storage controllers 102 may be configured to perform a data replication process that includes copying volumes of the first volume group of a first controller (or a first storage system 118) designated as a source controller to volumes at a second controller (or second storage system 118) designated as a destination controller to provide consistency between the source controller and the destination controller in event of failure of the source controller. Processing proceeds to end block for other processing.

[00042] In this manner, in some examples, the techniques of the present application may provide storage management techniques for load balancing to help improve the performance of storage systems.

[00043] The process 200 of Figs. 2A through 2B shows an example process and it should be understood that other configurations can be employed to practice the techniques of the present application. For example, process 200 may be configured to communicate with a plurality of additional storage controllers 102, plurality of storage systems 118 and the like.

[00044] Figs. 3A through 3H depicts example diagrams of processes for storage management in accordance with an example of the techniques of the present application. The Figs. 3A through 3H show how the techniques of the present application provide load balancing functionality for storage system 100 of Fig. 1, as an example. To illustrate operation, it may be assumed that system 100 includes storage system 118 with a first storage controller 102-1 and a second storage controller 102-2 configured in a dual controller configuration. It may also be assumed that storage controllers 102 are configured to communicate with host computers 108 over network 106. It may also be assumed that storage controllers 102 are configured to communicate with first storage

- 14 -

subsystem 104-1 and second storage subsystem 104-2 over storage network 122. The storage controllers 102 include storage management modules 116 to implement the techniques of the present application and functionality described herein.

[00045] Fig. 3A shows system 100 where storage system 118 configures a total of ten (10) storage volumes 110 labeled Vol1 through Vol10 to be assigned ownership across first storage controller 102-1 and second storage controller 102-2. In one example, storage system 118 configures storage volumes Vol7 and Vol8 as a first volume group associated with volumes of a first data replication group Group1, as shown in the table. The storage system 118 configures storage volumes Vol1, Vol3, Vol9 as first volume group associated with volumes of a second data replication group Group2, as shown in the table. In addition, data replication groups Group1 and Group2 are associated with respective data replication logs DRI which maintains information about the groups. The storage system 118 configures the other volumes Vol2, Vol4, Vol5, Vol6, Vol10 as second volume groups associated with all other volumes not associated with data replication groups.

[00046] Fig. 3B shows system 100 where storage system 118 processed IO requests from host computers 108 directed to storage volumes Vol1 through Vol10 over a period of time, as shown in the table. The table indicates IOPs associated with each of the volumes such as 1500 IOPs value associated with storage volume Vol1. The system 100 indicates that storage controllers 102 determined or calculated a Total IOPs value of 13,300 for all of the volumes, as shown in the table.

[00047] Fig. 3C shows system 100 where storage system 118 assigns ownership of volumes Vol1 through Vol10 to first storage controller 102-1 and second storage controller 102-2. In this case, the system assigns the volumes without considering the volumes associated with data replication groups Group1 and Group2. In one example, storage system 118 assigns ownership of volumes Vol3, Vol4, Vol5, Vol8, Vol10 to first storage controller 102-1 and assigns ownership of volumes Vol1, Vol2, Vol6, Vol7, Vol9 to second storage controller 102-2. In this case, it may be assumed that maximum IO processing power of each storage controller is a value of 10,000 IOPs for a total value of 20,000 IOPs. In this example, storage system 118 assigns ownership based on controller performance to provide a balanced volume set such that it assigns ownership to first storage controller 102-1 with volumes of total IOPs values of 6600 which results in usage or utilization performance of $66\% = (6,600/10,000)$ for the first controller. In a similar manner, storage system 118 assigns volume ownership to second storage controller 102-2 with volumes of total IOPs values of 6700 which results in usage or utilization performance of $67\% = (6,700/10,000)$ for the second controller. Further, it assumed that

- 15 -

maximum IO processing power per controller is 10,000 IOPs. In this case, the total IOPS processed by the storage system is $13,300 = (6,600 + 6,700)$. This is $66.5\% = (13,300/20,000)$ of the total storage system performance. In this case, the processing power of each of the storage controllers is well utilized.

[00048] However, in this case, storage system 118 assigns volumes ownership without considering the grouping of volumes of data replication groups Group1, Group2. As explained above, data replication groups have characteristics or properties such that the volumes or members of the data replication groups are to be assigned ownership to controllers as single entities. For example, the volumes of data replication group Group1 are to be assigned ownership to first storage controller 102-1 or to second storage controller 102-2 without splitting ownership of the volumes across different controllers. For example, Fig. 3C shows that volumes of data replication group Group1 are assigned ownership or distributed across first controller 102-1 and second controller 102-2. That is, volume Vol7 of data replication group Group1 is assigned ownership to second storage controller 102-2 and volume Vol8 of data replication group Group1 is assigned ownership to first storage controller 102-1. This configuration is undesirable as it prevents successful operation of the data replication process.

[00049] Fig. 3D shows system 100 where storage system 118 assigns ownership of volumes Vol1 through Vol10 to first storage controller 102-1 and second storage controller 102-2. In this case, storage system 118 assigns ownership of first data replication group Group1 and second replication group Group2 to first storage controller 102-1. In this case, however, though storage system 118 has considered data replication groups, the system has an imbalance in workload across first controller 102-1 (which has a workload of total IOPs value of 10,600) and second controller 102-1 (which has a workload of total IOPs value of 2,700).

[00050] In this case, it appears that first controller 102-1 performance utilization is very high compared to second controller 102-2 performance utilization. The first controller 102-1 is undergoing more IO activity or workload when compared with second controller 102-2. This way second controller 102-2 processing power is underutilized and may result in overall system performance degradation. Further assume that maximum IO processing power per controller is 10,000 IOPs. The total IOPS processed by the storage system is $13,300 = (10,600 + 2,700)$. This is $66.5\% = (13,300/20,000)$ of the total storage system performance. The storage system should be ready to accept 33.55% more IO requests. However, first controller's 102-1 performance utilization is over $100\% = (10,600/10,000)$ whereas second controller's 102-2 performance utilization is only $27\% = (2,700/10,000)$. In

- 16 -

this case, second controller's 102-2 processing power is underutilized. Furthermore, if host computers 108 desire to increase IO requests to the volumes owned by first controller 102-1 then it may fail due to resource limitation in the first controller, even though there is plenty of free resources available in second controller 102-2.

[00051] Fig. 3E shows system 100 where storage system 118 assigns ownership of volumes Vol1 through Vol10 to first storage controller 102-1 and second storage controller 102-2. In this case, storage system 118 assigns ownership of first data replication group Group1 and second replication group Group2 to second storage controller 102-2. In this case, similar to the case of Fig. 3D, however, though storage system 118 has considered data replication groups, the system has an imbalance in workload across first controller 102-1 (which has a workload of total IOPs value of 3,400) and second controller 102-2 (which has a workload of total IOPs value of 9,900).

[00052] In this case, it appears that first controller 102-1 performance utilization is low compared to second controller 102-2 performance utilization. The first controller 102-1 is undergoing less IO activity or workload when compared with second controller 102-2. This way first controller 102-1 processing power is underutilized and may result in overall system performance degradation. Further assume that maximum IO processing power per controller is 10,000 IOPs. The total IOPS processed by the storage system is $13,300 = (3,400 + 9,900)$. This is $66.5\% = (13,300/20,000)$ of the total storage system performance. The storage system should be ready to accept 33.55% more IO requests. However, second controller 102-2 performance utilization is at $99\% = (9,900/10,000)$ whereas first controller 102-1 performance utilization is only $34\% = (3,400/10,000)$. In this case, first controller 102-1 processing power is underutilized. Furthermore, if host computers 108 desire to increase IO requests to the volumes owned by second controller 102-2 then it may fail due to resource limitation in the second controller, even though there is plenty of free resources available in first controller 102-1.

[00053] Fig. 3F shows system 100 where storage system 118 assigns ownership of volumes Vol1 through Vol10 to first storage controller 102-1 and second storage controller 102-2. In this case, storage system 118 assigns ownership of second data replication group Group2 to first storage controller 102-1 and assigns ownership of first data replication group Group1 to second storage controller 102-2. In this case, similar to the cases of Figs. 3D and 3E above, however, though storage system 118 has considered data replication groups, the system has an imbalance in workload across first controller 102-1 (which has a workload of total IOPs value of 7,700) and second controller 102-2 (which has a workload of total IOPs value of 5,600).

- 17 -

[00054] In this case, it appears that first controller 102-1 performance utilization is high compared to second controller 102-2 performance utilization. The first controller 102-1 is undergoing more IO activity or workload when compared with second controller 102-2. This way second controller 102-2 processing power is underutilized and may result in overall system performance degradation. Further assume that maximum IO processing power per controller is 10,000 IOPS. The total IOPS processed by the storage system is $13,300 = (7,700 + 5,600)$. This is $66.5\% = (13,300/20,000)$ of the total storage system performance. The storage system should be ready to accept more 33.55% of IO requests. However, first controller 102-1 performance utilization is at $77\% = (7,700/10,000)$ whereas second controller 102-2 performance utilization is less at $56\% = (5,600/10,000)$. In this case, second controller 102-2 processing power is underutilized. Furthermore, if host computers 108 desire to increase IO requests to the volumes owned by first controller 102-1 then it may fail due to resource limitation in the first controller, even though there is plenty of free resources available in second controller 102-2.

[00055] Fig. 3G shows system 100 where storage system 118 assigns ownership of volumes Vol1 through Vol10 to first storage controller 102-1 and second storage controller 102-2, in this case, considering data replication groups Group1 and Group2. In this case, storage system 118 assigns ownership of second data replication group Group2 to second storage controller 102-2 and assigns ownership of first data replication group Group1 to first storage controller 102-1. In this case, similar to the cases of Fig. 3D, 3E, 3F above, however, though storage system 118 has considered data replication groups, the system has an imbalance in workload across first controller 102-1 (which has a workload of total IOPs value of 6,300) and second controller 102-2 (which has a workload of total IOPs value of 7,000).

[00056] In this case, it appears that first controller 102-1 performance utilization is low compared to second controller 102-2 performance utilization. The first controller 102-1 is undergoing less IO activity or workload when compared with second controller 102-2. This way first controller 102-1 processing power is underutilized and may result in overall system performance degradation. Further assume that maximum IO processing power per controller is 10,000 IOPS. The total IOPS processed by the storage system is $13,300 = (6,300 + 7,000)$. This is $66.5\% = (13,300/20,000)$ of the total storage system performance. The storage system should be ready to accept 33.55% more IO requests. However, second controller's 102-2 performance utilization is at $70\% = (7,000/10,000)$ whereas first controller's 102-1 performance utilization is only $63\% = (6,300/10,000)$. In this case, first controller's processing power is underutilized. Furthermore, if host computers 108 desire

- 18 -

to increase IO requests to the volumes owned by second controller 102-2 then it may fail due to resource limitation in the second controller, even though there is plenty of free resources available in first controller 102-1.

[00057] Fig. 3H shows system 100 where storage system 118 assigns ownership of volumes Vol1 through Vol10 to first storage controller 102-1 and second storage controller 102-2, in this case, considering data replication groups Group1 and Group2. In this case, storage system 118 has considered data replication groups, the system provides a well-balanced workload across first controller 102-1 (which has a workload of total IOPs value of 6,600) and second controller 102-2 (which has a workload of total IOPs value of 6,700). In one example, storage management modules 116 employed the load balancing techniques described herein to provide load balancing while considering data replication groups.

[00058] In this case, it appears that first controller 102-1 performance utilization is well balanced compared to second controller 102-2 performance utilization. The first controller 102-1 is undergoing about the same IO activity or workload when compared with second controller 102-2. This way first controller 102-1 and second controller 102-2 processing power is well utilized and may result in overall system performance improvement. Further assume that maximum IO processing power per controller is 10,000 IOPs. The total IOPS processed by the storage system is $13,300 = (6,600 + 6,700)$. This is 66.5% ($13,300/20,000$) of the total storage system performance. The storage system should be ready to accept 33.55% more IO requests. In this case, first controller 102-1 performance utilization is at $66\% = (6,600/10,000)$ which is about the same as second controller 102-2 performance utilization which is $67\% = (6,700/10,000)$. In this case, first controller 102-1 and second controller 102-2 processing power is well utilized. Furthermore, if host computers 108 desire to increase IO requests to the volumes owned by first controller 102-1 or second controller 102-2 then it will be able to accept the request since both controller have plenty of free resources available for IO processing. Furthermore, the techniques of the present application provide load balancing while considering data replication groups which allows for failover conditions. In this manner, these techniques consider data replication groups as part of the load balancing process which may improve performance of storage controllers, use of resources and IO response time.

[00059] The process of Figs. 3A through 3H shows example processes and it should be understood that other configurations can be employed to practice the techniques

- 19 -

of the present application. For example, processes above may be configured to communicate with a different number of storage controllers

[00060] Fig. 4 is an example block diagram showing a non-transitory, computer-readable medium that stores code for operation in accordance with an example of the techniques of the present application. The non-transitory, computer-readable medium is generally referred to by the reference number 400 and may be included in the system in relation to Fig. 1. The non-transitory, computer-readable medium 400 may correspond to any typical storage device that stores computer-implemented instructions, such as programming code or the like. For example, the non-transitory, computer-readable medium 400 may include one or more of a non-volatile memory, a volatile memory, and/or one or more storage devices. Examples of non-volatile memory include, but are not limited to, electrically erasable programmable Read Only Memory (EEPROM) and Read Only Memory (ROM). Examples of volatile memory include, but are not limited to, Static Random Access Memory (SRAM), and dynamic Random Access Memory (DRAM). Examples of storage devices include, but are not limited to, hard disk drives, compact disc drives, digital versatile disc drives, optical drives, and flash memory devices.

[00061] A processor 402 generally retrieves and executes the instructions stored in the non-transitory, computer-readable medium 400 to operate the present techniques in accordance with an example. In one example, the tangible, computer-readable medium 400 can be accessed by the processor 402 over a bus 404. A first region 406 of the non-transitory, computer-readable medium 400 may include storage management module 116 functionality as described herein. The storage management module 116 functionality may be implemented in hardware, software or a combination thereof.

[00062] For example, block 408 provides instructions which may include instructions to determine total IOPS values, as described herein. In one example, the instructions may include instructions to determine first IOPs values comprising IOPs of first volume groups associated with volumes of data replication groups assigned ownership to first and second storage controllers, as described herein. In one example, the instructions may include instructions to determine second IOPs values comprising IOPS of second volume groups associated with all other volumes not associated with data replication groups that are assigned ownership to the controllers, as described herein. In one example, the instructions may include instructions to determine total IOPs values comprising sum of the first IOPs values and second IOPs values of the controllers

[00063] For example, block 410 provides instructions which may include instructions to determine usage, as described herein. In one example, the instructions may

- 20 -

include instructions to determine usage values based on the total IOPs values of respective controllers, as described herein.

[00064] For example, block 412 provides instructions which may include instructions to partition volumes, as described herein. In one example, the instructions may include instructions to partition volumes of volume sets comprising the first volume groups and second volume groups into first and second partition volume groups, wherein the difference between sums of the first IOPs values and second IOPs values is minimum, and the volumes of the first volume groups are not transferable to the second volume groups, as described herein.

[00065] For example, block 414 provides instructions which may include instructions to reassign ownership, as described herein. In one example, the instructions may include instructions to reassign ownership of the first partition volume groups to the first controller and the second partition volume groups to the second controller, as described herein.

[00066] For example, block 416 provides instructions configure volumes for data replication process, as described herein. In one example, the instructions may include instructions to configure the volumes of the first volume groups to be associated with a data replication process, as described herein.

[00067] Although shown as contiguous blocks, the software components can be stored in any order or configuration. For example, if the non-transitory, computer-readable medium 400 is a hard drive, the software components can be stored in non-contiguous, or even overlapping, sectors.

[00068] As used herein, a "processor" may include processor resources such as at least one of a Central Processing Unit (CPU), a semiconductor-based microprocessor, a Graphics Processing Unit (GPU), a Field-Programmable Gate Array (FPGA) configured to retrieve and execute instructions, other electronic circuitry suitable for the retrieval and execution instructions stored on a computer-readable medium, or a combination thereof. The processor fetches, decodes, and executes instructions stored on medium 400 to perform the functionalities described below. In other examples, the functionalities of any of the instructions of medium 400 may be implemented in the form of electronic circuitry, in the form of executable instructions encoded on a computer-readable storage medium, or a combination thereof.

[00069] As used herein, a "computer-readable medium" may be any electronic, magnetic, optical, or other physical storage apparatus to contain or store information such as executable instructions, data, and the like. For example, any computer-readable

- 21 -

storage medium described herein may be any of Random Access Memory (RAM), volatile memory, non-volatile memory, flash memory, a storage drive (e.g., a hard drive), a solid state drive, any type of storage disc (e.g., a compact disc, a DVD, etc.), and the like, or a combination thereof. Further, any computer-readable medium described herein may be non-transitory. In examples described herein, a computer-readable medium or media is part of an article (or article of manufacture). An article or article of manufacture may refer to any manufactured single component or multiple components. The medium may be located either in the system executing the computer-readable instructions, or remote from but accessible to the system (e.g., via a computer network) for execution. In the example of Fig. 4, medium 400 may be implemented by one computer-readable medium, or multiple computer-readable media.

[00070] In examples described herein, storage system 118 may communicate with components implemented on separate system(s) via a network interface device of storage system 118. For example, storage system 118 may communicate with storage subsystem 104 via a network interface device of storage system 118. In examples described herein, a “network interface device” may be a hardware device to communicate over at least one computer network. In some examples, a network interface may be a Network Interface Card (NIC) or the like. As used herein, a computer network may include, for example, a Local Area Network (LAN), a Wireless Local Area Network (WLAN), a Virtual Private Network (VPN), the Internet, or the like, or a combination thereof. In some examples, a computer network may include a telephone network (e.g., a cellular telephone network).

[00071] In some examples, instructions 408-416 may be part of an installation package that, when installed, may be executed by processor 402 to implement the functionalities described herein in relation to instructions 408-416. In such examples, medium 400 may be a portable medium, such as a CD, DVD, or flash drive, or a memory maintained by a server from which the installation package can be downloaded and installed. In other examples, instructions 408-416 may be part of an application, applications, or component(s) already installed on storage system 118 including processor 402. In such examples, the medium 400 may include memory such as a hard drive, solid state drive, or the like. In some examples, functionalities described herein in relation to Figs. 1 through 4 may be provided in combination with functionalities described herein in relation to any of Figs. 1 through 4.

[00072] The foregoing describes a novel and previously unforeseen approach for storage management. While the above application has been shown and described with reference to the foregoing examples, it should be understood that other forms, details, and

- 22 -

implementations may be made without departing from the spirit and scope of this application.

- 23 -

WHAT IS CLAIMED IS:

1. An apparatus for storage management comprising:
a storage management module to:
determine total Input Output Requests Per Second (IOPs) values comprising sum of first IOPs values and second IOPs values of first and second controllers, wherein the first IOPS values comprises IOPs of first volume groups associated with volumes of data replication groups assigned ownership to the controllers, and wherein the second IOPs values comprises IOPs of second volume groups associated with other volumes not associated with data replication groups assigned ownership to the controllers; and
if difference of usage values of one of the controllers is greater than a threshold amount, then
partition volumes of volume sets comprising the first volume groups and second volume groups into first and second partition volume groups, wherein the difference between sums of the first IOPs values and second IOPs values is minimum, and the volumes of the first volume groups are not assigned ownership to the second volume groups, and
reassign ownership of the first partition volume groups to the first controller and the second partition volume groups to the second controller.
2. The apparatus of claim 1, wherein the storage management module further configured to, in response to the first controller receipt from a host computer a request to access a volume owned by the first controller, then the first storage controller to respond with status of the request to the host computer.
3. The apparatus of claim 1, wherein the storage management module further configured to, in response to the first controller receipt from a host computer a request to access a volume owned by the second controller, then the first controller to cause the second storage controller to access the volume, the second controller to send a response with status of the request to the first controller which responds with the status to the host computer.
4. The apparatus of claim 1, wherein the storage management module further configured to identify volumes of data replication groups and assign ownership of the volumes of the data replication groups as single entities to the first or second controller.

- 24 -

5. The apparatus of claim 1, wherein the storage management module to determine the first IOPs values and second IOPs values of the controllers based on metadata of the controllers and then reset the metadata of the controller after the determination of the IOPs values.

6. A method of storage management, the method comprising:

- determining first Input Output Requests Per Second (IOPs) values comprising IOPs of first volume groups associated with volumes of data replication groups assigned ownership to first and second storage controllers;

- determining second IOPs values comprises IOPS of second volume groups associated with all other volumes not associated with data replication groups assigned ownership to the controllers;

- determining total IOPs values comprising sum of the first IOPs values and second IOPs values of the controllers;

- determining usage values based on the total IOPs values of respective controllers; and

- if difference of usage values of one of the controllers is greater than a threshold amount; then

- partitioning volumes of volume sets comprising the first volume groups and second volume groups into first and second partition volume groups, wherein the difference between sums of the first IOPs values and second IOPs values is minimum, and the volumes of the first volume groups are not assigned ownership to the second volume groups,

- reassigning ownership of the first partition volume groups to the first controller and the second partition volume groups to the second controller, and

- configuring the volumes of the first volume groups to be associated with a data replication process.

7. The method of claim 6, further comprising, in response to the first controller receipt from a host computer a request to access a volume owned by the first controller, then the first storage controller responding with status of the request to the host computer.

8. The method of claim 6, further comprising to, in response to the first controller receipt from a host computer a request to access a volume owned by the second controller, then

- 25 -

the first controller to cause the second storage controller accessing the volume, the second controller sending a response with status of the request to the first controller which responds with the status to the host computer.

9. The method of claim 6, further comprising identifying volumes of data replication groups and assign ownership of the volumes of the data replication groups as single entities to the first or second controller.

10. The method of claim 6, wherein the data replication process comprises copying volumes of the first volume group of a first controller designated as a source controller to volumes at a second controller designated as a destination controller to provide consistency between the source controller and the destination controller in event of failure of the source controller.

11. A non-transitory computer-readable medium having computer executable instructions stored thereon for storage management, the instructions are executable by a processor to:

- determine first Input Output Requests Per Second (IOPs) values comprising IOPs of first volume groups associated with volumes of data replication groups assigned ownership to first and second storage controllers;

- determine second IOPs values comprises IOPS of second volume groups associated with all other volumes not associated with data replication groups assigned ownership to the controllers;

- determine total IOPs values comprising sum of the first IOPs values and second IOPs values of the controllers;

- determine usage values based on the total IOPs values of respective controllers;
- and

- if difference of usage values of one of the controllers is greater than a threshold amount, then

- partition volumes of volume sets comprising the first volume groups and second volume groups into first and second partition volume groups, wherein the difference between sums of the first IOPs values and second IOPs values is minimum, and the volumes of the first volume groups are not assigned ownership to the second volume groups, and

- 26 -

reassign ownership of the first partition volume groups to the first controller and the second partition volume groups to the second controller.

12. The non-transitory computer-readable medium of claim 11, further comprising instructions that if executed cause a processor to: in response to the first controller receipt from a host computer a request to access a volume owned by the first controller, then the first storage controller responding with status of the request to the host computer.

13. The non-transitory computer-readable medium of claim 11, further comprising instructions that if executed cause a processor to: in response to the first controller receipt from a host computer a request to access a volume owned by the second controller, then the first controller to cause the second storage controller to access the volume, the second controller to send a response with status of the request to the first controller which responds with the status to the host computer.

14. The non-transitory computer-readable medium of claim 11 further comprising instructions that if executed cause a processor to: identify volumes of data replication groups and assign ownership of the volumes of the data replication groups as single entities to the first or second controller.

15. The non-transitory computer-readable medium of claim 11 further comprising instructions that if executed cause a processor to: determine the first IOPs values and second IOPs values of the controllers based on metadata of the controllers and then reset the metadata of the controller after the determination of the IOPs values.

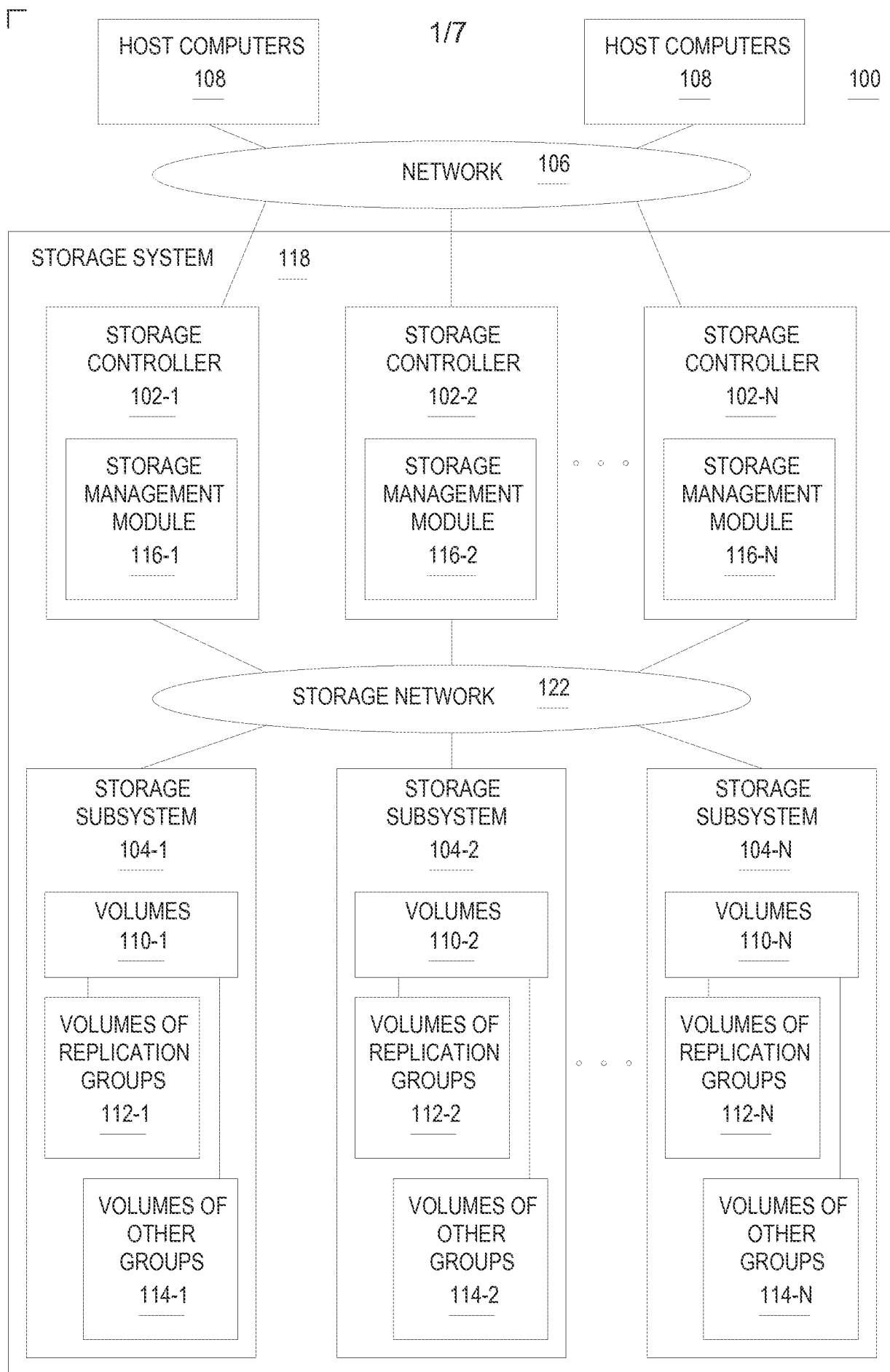


FIG. 1

2/7

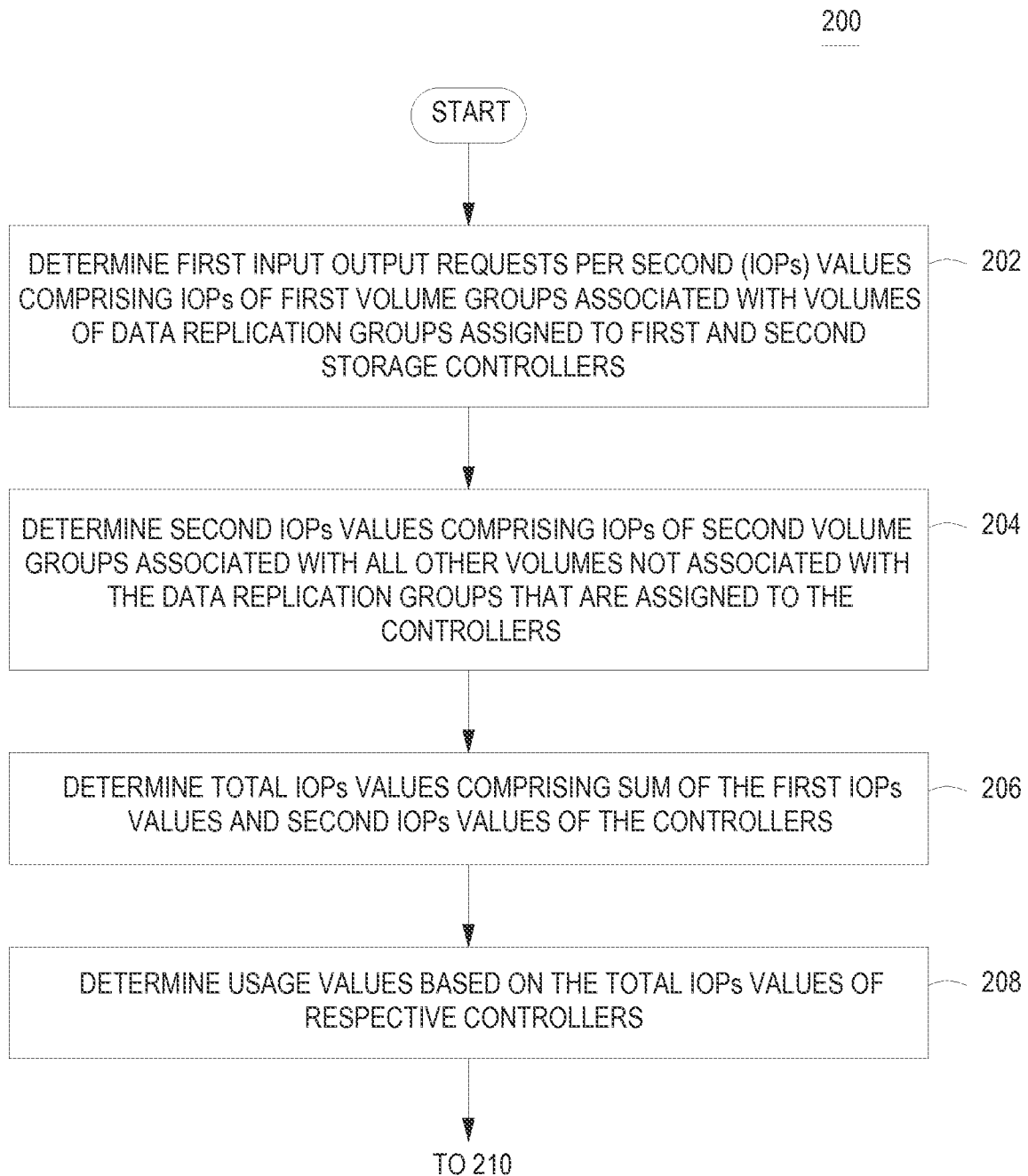


FIG. 2A

3/7

200

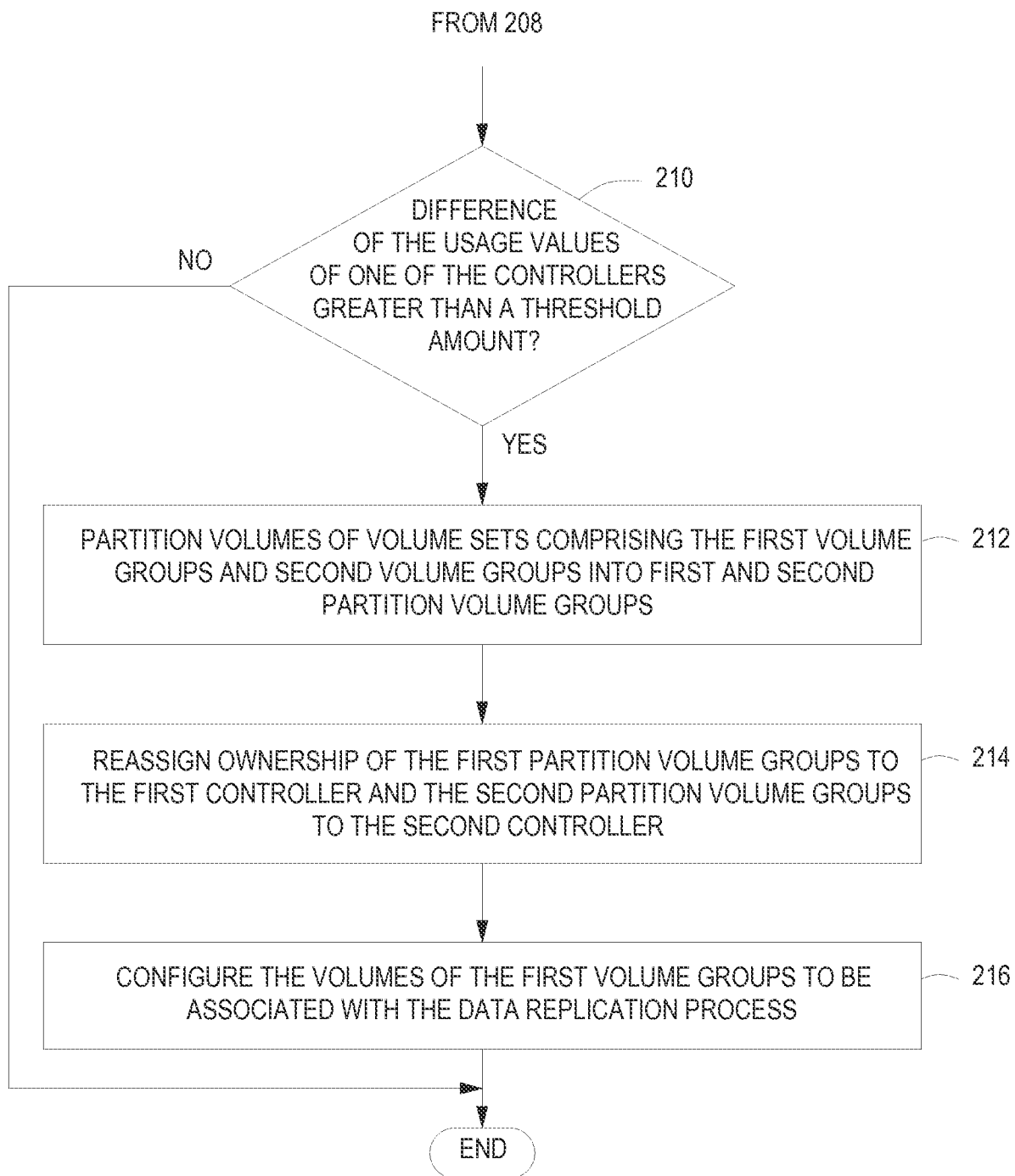


FIG. 2B

┌

FIG. 3A

DATA REPLICATION GROUP ID	VOLUME LIST
1	VOL7, VOL8
2	VOL1, VOL3, VOL9

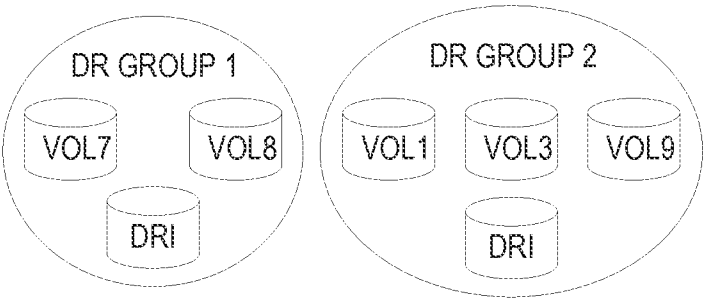


FIG. 3B

TOTAL VOLUME SET IOPs	
VOL1	1500
VOL2	1100
VOL3	1800
VOL4	1200
VOL5	1000
VOL6	1600
VOL7	1500
VOL8	1400
VOL9	1000
VOL10	1200
TOTAL IOPs	13300

FIG. 3C

FIRST CONTROLLER IOPs		SECOND CONTROLLER IOPs	
VOL3	1800	VOL1	1500
VOL4	1200	VOL2	1100
VOL5	1000	VOL6	1600
VOL8	1400	VOL7	1500
VOL10	1200	VOL9	1000
TOTAL IOPs	6600	TOTAL IOPs	6700

└

┌

5/7

FIG. 3D

FIRST CONTROLLER IOPs		SECOND CONTROLLER IOPs	
VOL1	1500	VOL2	1100
VOL3	1800	VOL6	1600
VOL9	1000		
VOL7	1500		
VOL8	1400		
VOL4	1200		
VOL5	1000		
VOL10	1200		
TOTAL IOPs	10600	TOTAL IOPs	2700

FIG. 3E

FIRST CONTROLLER IOPs		SECOND CONTROLLER IOPs	
VOL4	1200	VOL1	1500
VOL5	1000	VOL3	1800
VOL10	1200	VOL9	1000
		VOL7	1500
		VOL8	1400
		VOL2	1100
		VOL6	1600
TOTAL IOPs	3400	TOTAL IOPs	9900

└

┌

6/7

FIG. 3F

FIRST CONTROLLER IOPs		SECOND CONTROLLER IOPs	
VOL1	1500	VOL2	1100
VOL3	1800	VOL6	1600
VOL9	1000	VOL7	1500
VOL4	1200	VOL8	1400
VOL5	1000		
VOL10	1200		
TOTAL IOPs	7700	TOTAL IOPs	5600

FIG. 3G

FIRST CONTROLLER IOPs		SECOND CONTROLLER IOPs	
VOL7	1500	VOL2	1100
VOL8	1400	VOL6	1600
VOL4	1200	VOL1	1500
VOL5	1000	VOL3	1800
VOL10	1200	VOL9	1000
TOTAL IOPs	6300	TOTAL IOPs	7000

FIG. 3H

FIRST CONTROLLER IOPs		SECOND CONTROLLER IOPs	
VOL2	1100	VOL4	1200
VOL5	1000	VOL10	1200
VOL6	1600	VOL1	1500
VOL7	1500	VOL3	1800
VOL8	1400	VOL9	1000
TOTAL IOPs	6600	TOTAL IOPs	6700

└

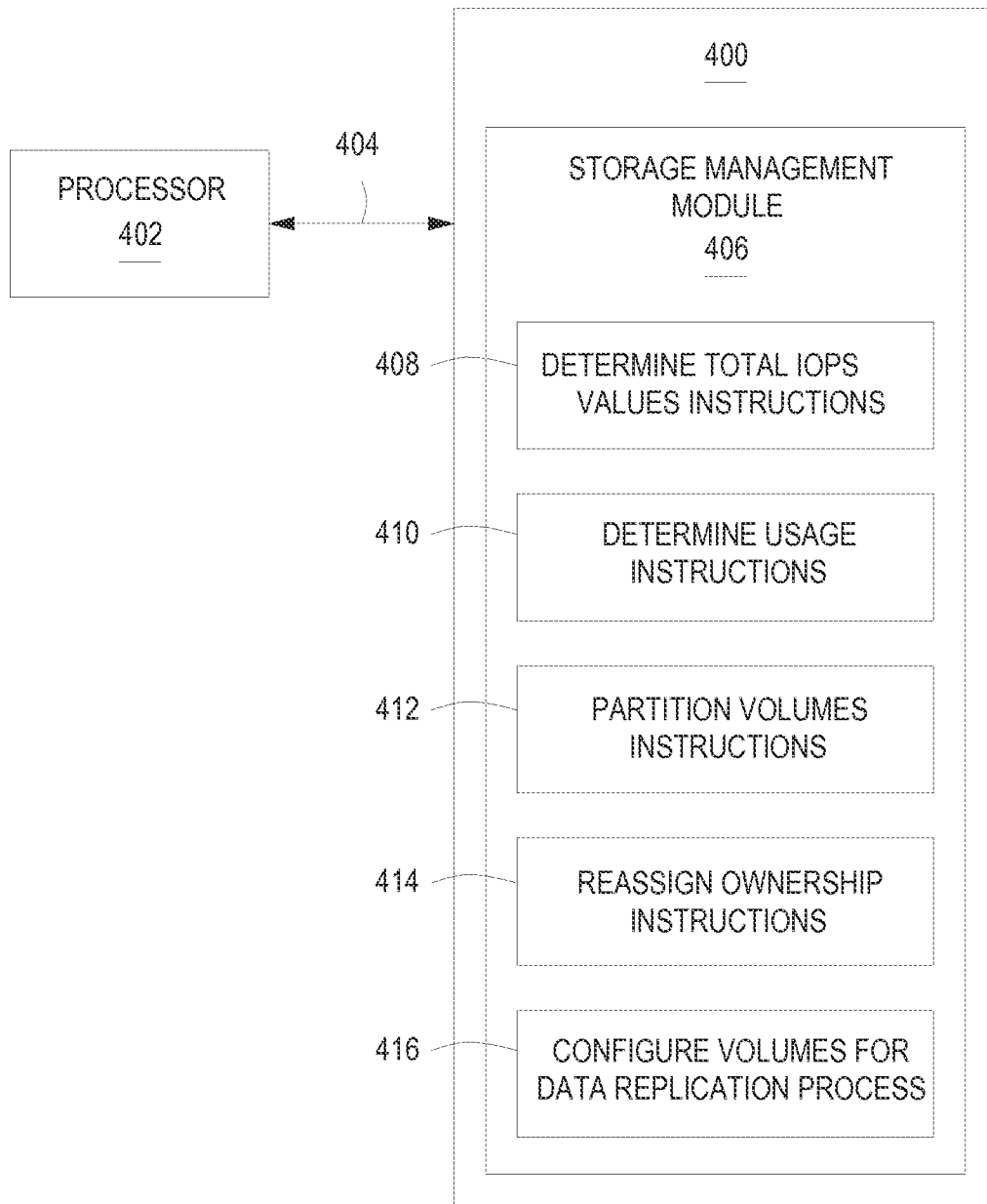


FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2015/045127**A. CLASSIFICATION OF SUBJECT MATTER****G06F 3/06(2006.01)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)

G06F 3/06; G06F 3/00; G06F 7/00; G06F 12/00

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: assign, second, volume, group, request, value, total, balance, storage, controller

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2010-0262772 A1 (DANIEL J. MAZINA) 14 October 2010 See paragraphs [0017], [0020]-[0021], [0027], [0037], [0040]; claims 2, 8, 11; and figure 4.	1-15
Y	US 2009-0271535 A1 (YASUHIKO YAMAGUCHI et al.) 29 October 2009 See paragraphs [0009], [0058]-[0059]; and figures 5-6.	1-15
A	US 2009-0089458 A1 (SADAHIRO SUGIMOTO et al.) 02 April 2009 See paragraph [0042]; and figure 5.	1-15
A	US 2005-0267929 A1 (MANABU KITAMURA) 01 December 2005 See paragraphs [0053], [0057]; and figures 4-5.	1-15
A	EP 2003541 A2 (HITACHI, LTD.) 17 December 2008 See paragraph [0068]; and figure 5.	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 another 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 February 2016 (24.02.2016)

Date of mailing of the international search report

25 February 2016 (25.02.2016)

Name and mailing address of the ISA/KR

International Application Division

Korean Intellectual Property Office

189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-472-7140

Authorized officer

BYUN, Sung Cheal

Telephone No. +82-42-481-8262



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2015/045127

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2010-0262772 A1	14/10/2010	US 8370571 B2	05/02/2013
US 2009-0271535 A1	29/10/2009	EP 1895397 A1	05/03/2008
		JP 2008-059478 A	13/03/2008
		US 2008-0126645 A1	29/05/2008
		US 8051219 B2	01/11/2011
US 2009-0089458 A1	02/04/2009	JP 2009-087175 A	23/04/2009
		US 7711872 B2	04/05/2010
US 2005-0267929 A1	01/12/2005	JP 2006-048627 A	16/02/2006
		JP 2006-228188 A	31/08/2006
		US 2005-0267950 A1	01/12/2005
		US 7209967 B2	24/04/2007
		US 7886299 B2	08/02/2011
EP 2003541 A2	17/12/2008	EP 2003541 A3	24/03/2010
		JP 04434235 B2	17/03/2010
		JP 2008-304963 A	18/12/2008
		US 2008-0307271 A1	11/12/2008
		US 2011-0258402 A1	20/10/2011
		US 7996728 B2	09/08/2011
		US 8397105 B2	12/03/2013