

(12) UK Patent

(19) GB

(11) 2580263

(45) Date of B Publication

(13) B

17.11.2021

(54) Title of the Invention: Dynamic data relocation using cloud based ranks

(51) INT CL: **G06F 3/06** (2006.01)

(21) Application No: 2004106.7

(22) Date of Filing: 20.08.2018

Date Lodged: 20.03.2020

(30) Priority Data:
(31) 15704775 (32) 14.09.2017 (33) US

(86) International Application Data:
PCT/IB2018/056280 En 20.08.2018

(87) International Publication Data:
WO2019/053534 En 21.03.2019

(43) Date of Reproduction by UK Office 15.07.2020

(56) Documents Cited:
CN 104820575 A CN 101719106 A
US 20120016838 A1

(58) Field of Search:
As for published application 2580263 A viz:
INT CL **G06F, H04L**
Other: **CNABS; VAN; CPRSABS; CNKI; CNTXT; IEEE**
updated as appropriate

Additional Fields
INT CL **H04L**
Other: **WPI, EPODOC**

(72) Inventor(s):

**Lokesh M Gupta
Matthew Borlick**

(73) Proprietor(s):

**International Business Machines Corporation
(Incorporated in USA - New York)
New Orchard Road, Armonk, New York 10504,
United States of America**

(74) Agent and/or Address for Service:

**IBM United Kingdom Limited
Intellectual Property Department, Hursley Park,
Winchester, Hampshire, SO21 2JN, United Kingdom**

GB

2580263

B

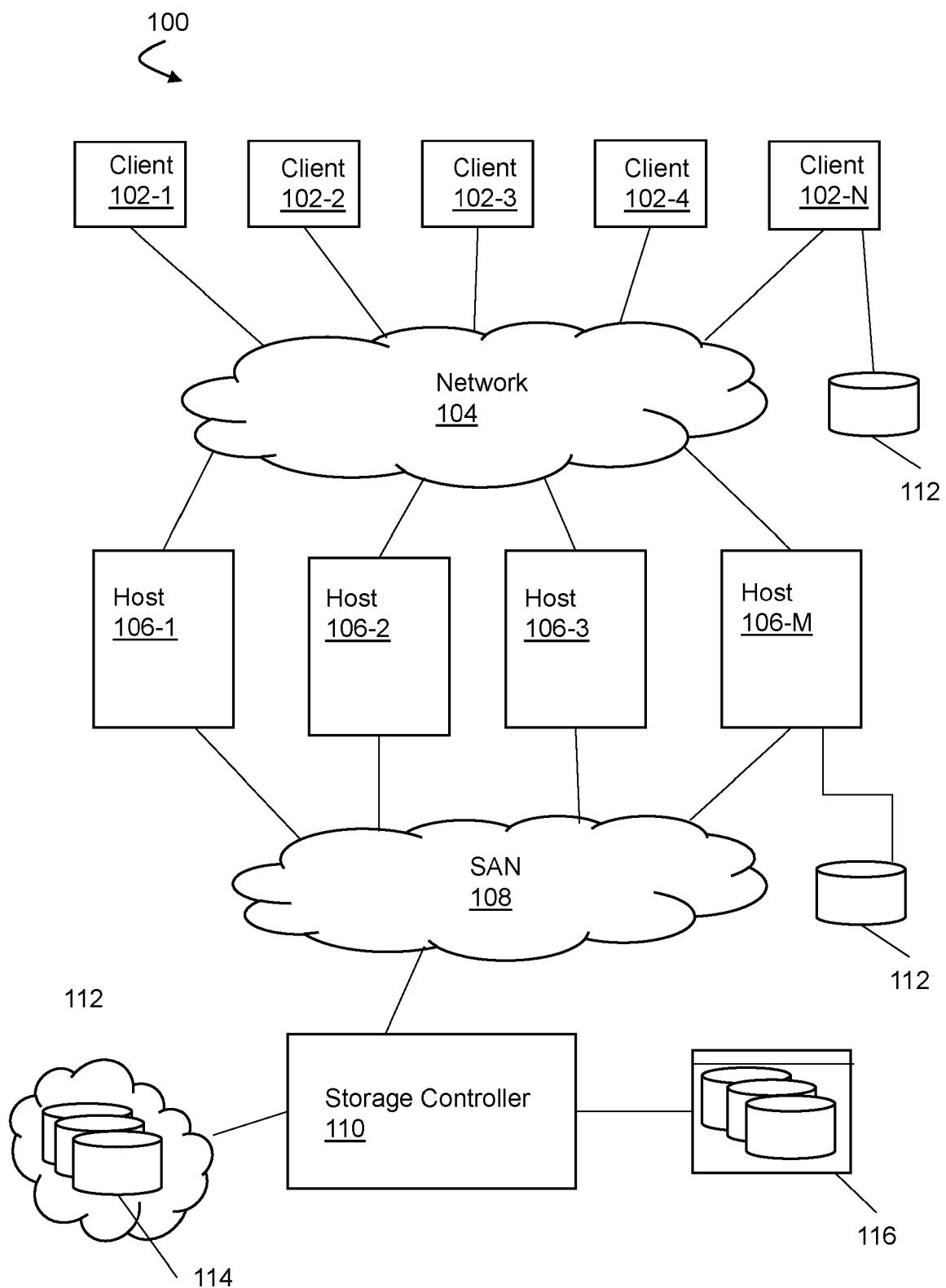
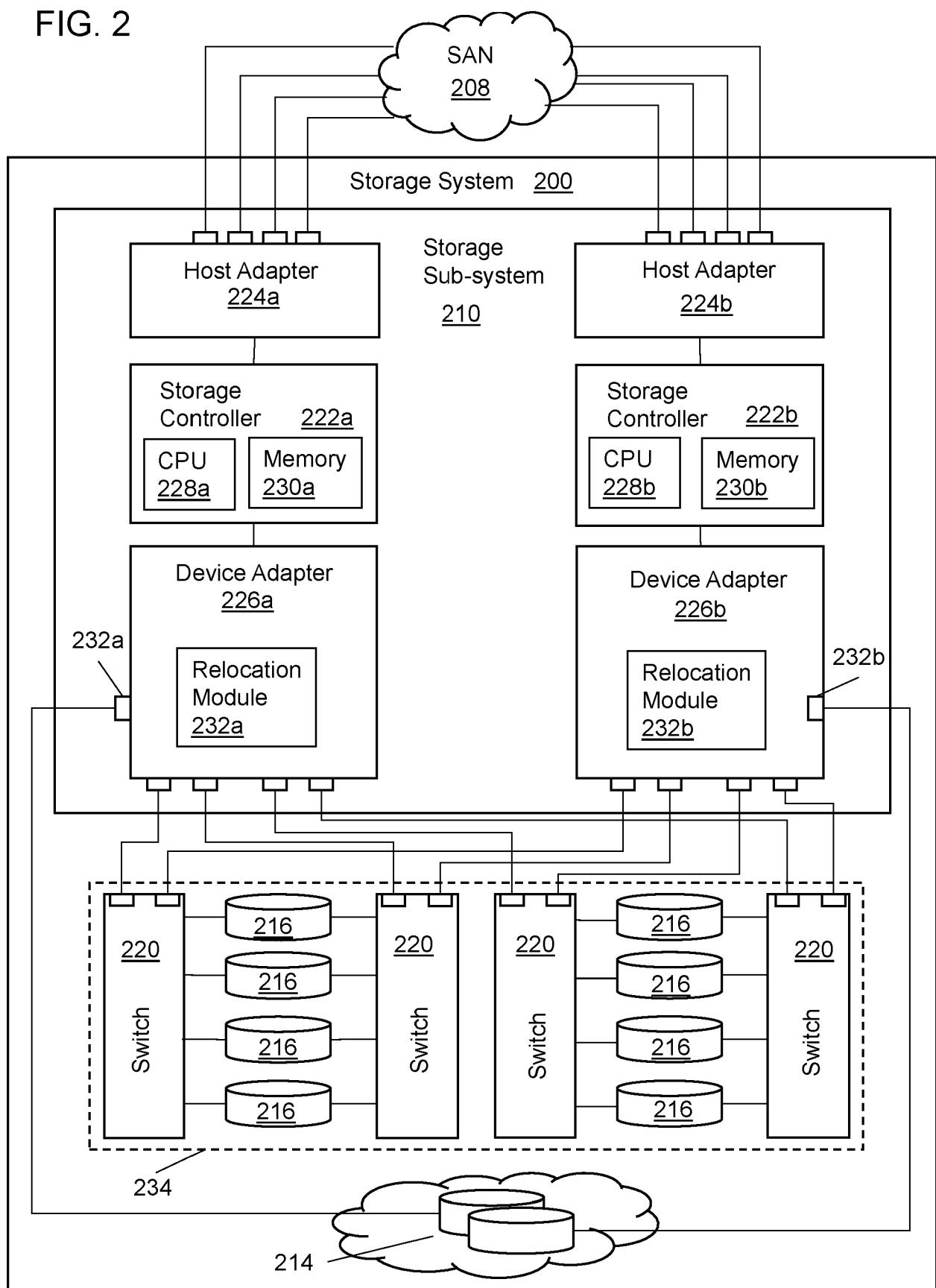


FIG. 1

FIG. 2



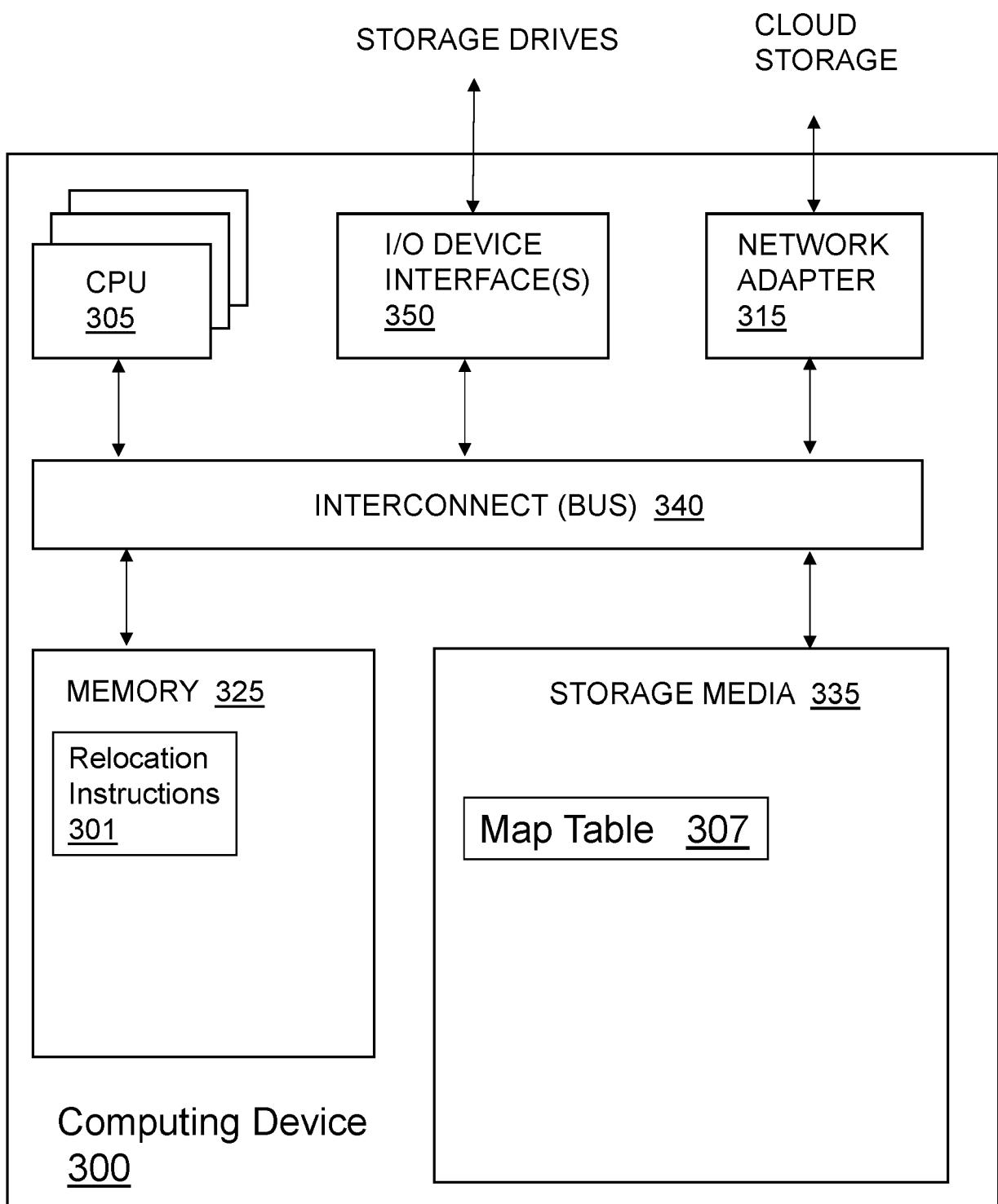


FIG. 3

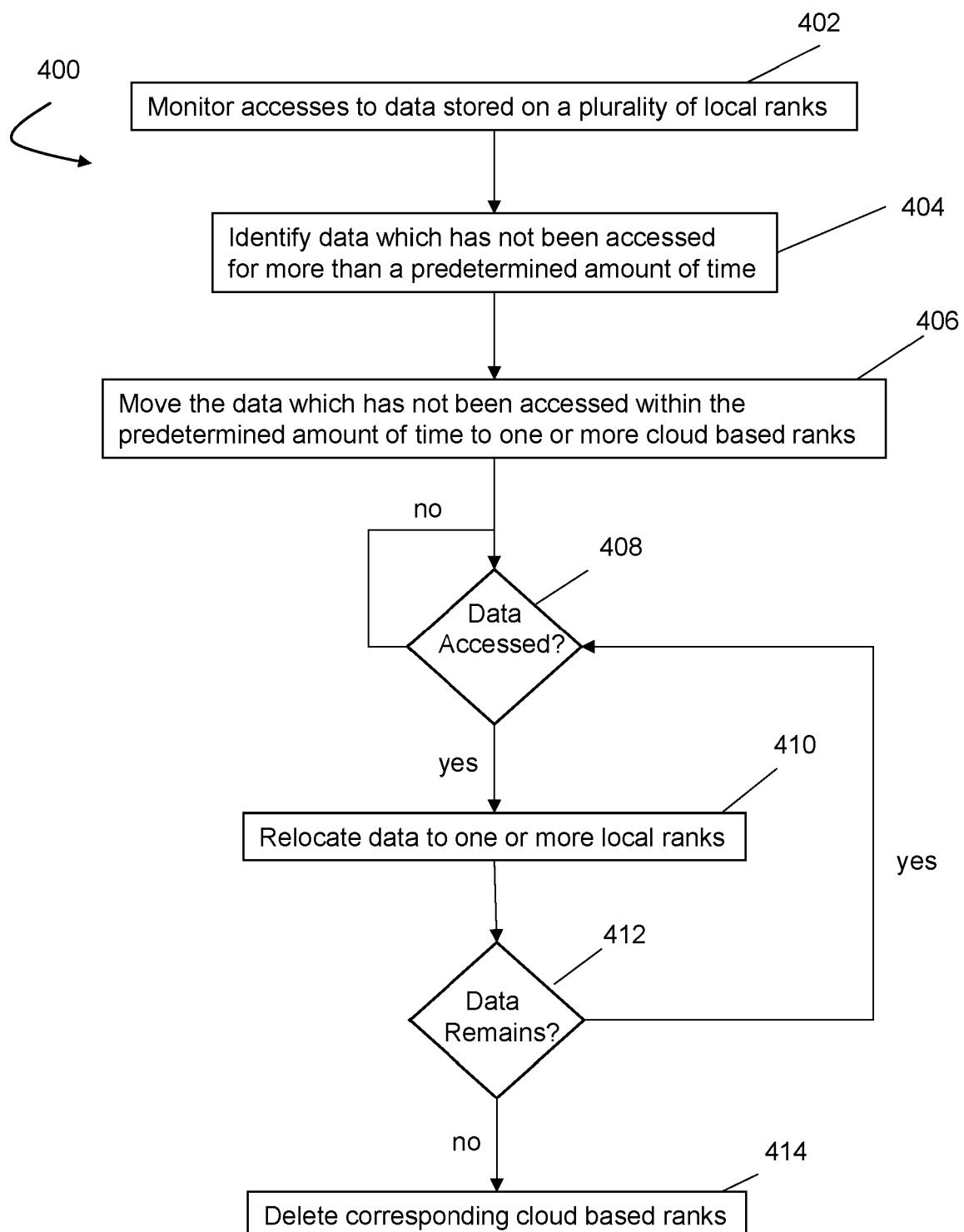


FIG. 4

DYNAMIC DATA RELOCATION USING CLOUD BASED RANKS

BACKGROUND

[0001] Storage networks, such as storage area networks (SANs), are used to interconnect different types of data storage systems with different types of servers (also referred to herein as "host systems"). Some servers involve various hardware such as data storage media, storage controllers, memories, and the accompanying power systems, cooling systems, etc.

[0002] Storage controllers control access to data storage media and memories in response to read and write requests. The storage controllers may direct the data in accordance with data storage devices such as RAID (redundant array of independent disks), JBOD (just a bunch of disks), and other redundancy and security levels. As an example, an IBM® ESS (Enterprise Storage Server) such as a DS8000 series has redundant clusters of computer entities, cache, non-volatile storage, etc.

SUMMARY

[0003] Aspects of the present invention may include a computer implemented method, computer program product, computing device, and system for dynamic data relocation using cloud based ranks. An example method embodying the present invention comprises monitoring accesses to data stored on a plurality of local ranks of an enterprise storage system; identifying data which has not been accessed for a predetermined amount of time based on the monitored accesses; and moving the data which has not been accessed for the predetermined amount of time to one or more cloud based ranks of the enterprise storage system, wherein each cloud based rank comprises storage space on one or more cloud storage devices, the storage space on the one or more cloud storage devices mapped to corresponding virtual local addresses that are grouped as a virtual local rank.

[0004] The above summary is not intended to describe each illustrated embodiment or every implementation of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] Understanding that the drawings depict only exemplary embodiments of the present invention and are not therefore to be considered limiting in scope, the exemplary embodiments of the present invention will be described with additional specificity and detail through the use of the accompanying drawings, in which:

[0006] FIG. 1 is a high-level block diagram depicting an example network environment;

[0007] FIG. 2 is a high-level block diagram depicting an example storage system.

[0008] FIG. 3 is a block diagram of an example device adapter.

[0009] FIG. 4 is a flow chart depicting an example method of dynamic data relocation.

[0010] In accordance with common practice, the various described features are not drawn to scale but are drawn to emphasize specific features relevant to the exemplary embodiments of the present invention.

DETAILED DESCRIPTION

[0011] In the following detailed description, reference is made to the accompanying drawings that form a part hereof, and in which is shown by way of illustration specific illustrative embodiments of the present invention. However, it is to be understood that other embodiments of the present invention may be utilized and that logical, mechanical, and electrical changes may be made. Furthermore, the method presented in the drawing figures and the specification is not to be construed as limiting the order in which the individual steps may be performed. The following detailed description is, therefore, not to be taken in a limiting sense.

[0012] As used herein, the phrases "at least one", "one or more," and "and/or" are open-ended expressions that are both conjunctive and disjunctive in operation. For example, each of the expressions "at least one of A, B and C", "at least one of A, B, or C", "one or more of A, B, and C", "one or more of A, B, or C" and "A, B, and/or C" means A alone, B alone, C alone, A and B together, A and C together, B and C together, or A, B and C together. Additionally, the term "a" or "an" entity refers to one or more of that entity. As such, the terms "a" (or "an"), "one or more" and "at least one" can be used interchangeably herein. It is also to be noted that the terms "comprising," "including," and "having" can be used interchangeably. The term "automatic" and variations thereof, as used herein, refers to any process or operation done without material human input when the process or operation is performed. Human input is deemed to be material if such input directs or controls how or when the process or operation is performed. A process which uses human input is still deemed automatic if the input does not direct or control how or when the process is executed.

[0013] The terms "determine", "calculate" and "compute," and variations thereof, as used herein, are used interchangeably and include any type of methodology, process, mathematical operation or technique. Hereinafter, "in communication" or "communicatively coupled" shall mean any electrical connection, whether wireless or wired, that allows two or more systems, components, modules, devices, etc. to exchange data, signals, or other information using any protocol or format. Furthermore, two components that are communicatively coupled need not be directly coupled to one another, but can also be coupled together via other intermediate components or devices.

[0014] FIG. 1 is a high-level block diagram depicting an example network architecture 100. The network architecture 100 is presented only by way of example and not limitation. Indeed, the systems and methods disclosed herein may be applicable to a wide variety of different network architectures in addition to the network architecture 100 shown in FIG. 1.

[0015] As shown, the network architecture 100 includes one or more clients or client computers 102-1...102-N, where N is the total number of client computers, and one or more hosts 106-1 ... 106-M, where M is the total number of hosts (also referred to herein as "server computers" 106, "host systems" 106, or "host devices" 106). It is to be understood that although five clients 102 are shown in FIG. 1, other numbers of clients 102 can be used in other embodiments of the present invention. For example, in some embodiments only one client 102 is implemented. In other embodiments of the present invention, more than five or fewer than 5 clients 102 are used. Similarly, it is to be understood that although four hosts 106 are shown in FIG. 1, any suitable number of hosts 106 can be used. For example, in some embodiments of the present invention, only a single host 106 is used. In other embodiments of the present invention, more than four or fewer than four storage hosts 106 can be used.

[0016] Each of the client computers 102 can be implemented as a desktop computer, portable computer, laptop or notebook computer, netbook, tablet computer, pocket computer, smart phone, or any other suitable type of electronic device. Similarly, each of the hosts 106 can be implemented using any suitable host computer or server. Such servers can include, but are not limited to, IBM System z® and IBM System i® servers, as well as UNIX servers, Microsoft Windows servers, and Linux platforms.

[0017] The client computers 102 are communicatively coupled to hosts 106 via a network 104. The network 104 may include, for example, a local-area-network (LAN), a wide-area-network (WAN), the Internet, an intranet, or the like. In general, the client computers 102 initiate communication sessions, whereas the server computers 106 wait for requests from the client computers 102. In certain embodiments of the present invention, the computers 102 and/or servers 106 may connect to one or more internal or external direct-attached storage systems 112 (e.g., arrays of hard-disk drives, solid-state drives, tape drives, etc.). These computers 102, 106 and direct-attached storage systems 112 may communicate using protocols such as ATA, SATA, SCSI, SAS, Fibre Channel, or the like.

[0018] The network architecture 100 may, in certain embodiments of the present invention, include a storage network 108 behind the servers 106, such as a storage-area-network (SAN) 108 or a LAN 108 (e.g., when using network-attached storage). In the example shown in FIG. 1, the network 108 connects the servers 106 to one or more storage sub-systems 110. Although only one storage sub-system 110 is shown for purposes of illustration, it is to be understood that more than one storage sub-system 110 can be used in other embodiments of the present invention. The storage sub-system 110 manages connections to arrays of storage devices 116. The arrays of storage devices 116 can include arrays of hard-disk drives and/or solid-state drives. In addition, in the example

shown in FIG. 1, the storage sub-system 110 is configured to connect to and create storage arrays from cloud storage such that the cloud storage appears as a local storage array which enables advanced data management features, as described in more detail below.

[0019] In addition, in conventional systems data cold data (e.g. data which is not accessed frequently) can be placed on relatively slower storage media (e.g. spinning disks) with data accessed more frequently located on faster media (e.g. solid state disks). As the data categorized as cold data grows and ages, conventional systems leave the cold data alone until more disk space is needed and then additional storage space is added which can be an expensive and time intensive process. If the amount of cold data then subsequently decreases, there is more storage space than what is needed. By enabling the creation of cloud based ranks which appear as local ranks to the enterprise storage system, the embodiments of the present invention described herein enable more efficient dynamic relocation of data, as described in more detail below. In particular, an additional category of data based on data accesses is created. This additional category is defined to include data which has not been accessed for an extended period of time. This new category of data can then be moved or relocated to the cloud based ranks, as described in more detail below. Thus, new ranks (i.e. cloud based ranks) are created as needed and can be deleted when not needed which saves space for relatively faster local ranks, while still enabling the advanced management features not typically available for remote storage.

[0020] To access a storage sub-system 110, a host system 106 may communicate over physical connections from one or more ports on the host 106 to one or more ports on the storage sub-system 110. A connection may be through a switch, fabric, direct connection, or the like. In certain embodiments, the servers 106 and storage sub-systems 110 may communicate using a networking standard such as Fibre Channel (FC) or iSCSI.

[0021] FIG. 2 is a high-level block diagram of a storage system 200. Storage system 200 includes one or more arrays of storage drives (e.g., hard-disk drives and/or solid-state drives). As shown, the storage system 200 includes a storage sub-system 210, a plurality of switches 202, and a plurality of storage drives 216 such as hard disk drives and/or solid-state drives (such as flash-memory-based drives). The storage sub-system 210 may enable one or more hosts (e.g., open system and/or mainframe servers) to access data in the plurality of storage drives 216.

[0022] In some embodiments of the present invention, the storage sub-system 210 includes one or more storage controllers 222. In the example shown in FIG. 2, the storage sub-system includes storage controller 222a and storage controller 222b. Although only two storage controllers 226 are shown herein for purposes of explanation, it is to be understood that more than two storage controllers can be used in other embodiments of the present invention. The storage sub-system 210 in FIG. 2 also includes host adapters 224a, 224b and device adapters 226a, 226b to connect the storage sub-system 210 to host devices and storage drives 204, respectively. Multiple

storage controllers 222a, 222b provide redundancy to help ensure that data is available to connected hosts. Thus, when one storage controller (e.g. storage controller 222a) fails, the other storage controller (e.g. 222b) can pick up the I/O load of the failed storage controller to ensure that I/O is able to continue between the hosts and the storage drives 204. This process can be referred to as a "failover."

[0023] Each storage controller 222 can include respective one or more processors 228 and memory 230. The memory 230 can include volatile memory (e.g., RAM) as well as non-volatile memory (e.g., ROM, EPROM, EEPROM, flash memory, etc.). The volatile and non-volatile memory can store software modules that run on the processor(s) 228 and are used to access data in the storage drives 204. The storage controllers 222 can host at least one instance of these software modules. These software modules can manage all read and write requests to logical volumes in the storage drives 204.

[0024] In particular, each storage controller 222 is communicatively coupled to the storage drives 204 via a respective device adapter 226. Each device adapter 226 is configured to manage Input/Output (I/O) accesses (also referred to herein as data access requests or access requests) to the storage drives 216. For example, the device adapters 226 logically organize the storage drives 216 and determine where to store data on the storage drives 216. The storage drives 216 (also referred to as disk drive modules (DDM)) can include groups of different types of drives having different performance characteristics. For example, the storage drives 216 can include a combination of (relatively) slow 'nearline' disks (e.g. 7,200 revolutions per minute (RPM) rotational speed), SAS disk drives (e.g. 10k or 15k RPM) and relatively fast solid state drives (SSD).

[0025] The device adapters 226 are coupled to the storage drives 216 via switches 220. Each of the switches 220 can be fiber switches coupling the storage drives 216 to the device adapters via fiber optic connections. The device adapters 226 logically group the storage drives 216 into array sites 234. For purposes of illustration, a single array site 234 comprised of storage drives 216 is depicted in FIG. 2. However, it is to be understood that more than one array site comprised of storage drives 216 can be included in other embodiments of the present invention. The array site 234 can be formatted as a Redundant Array of Independent Disks (RAID) array 234. It is to be understood that any type of RAID array (e.g. RAID 0, RAID 5, RAID 10, etc.) can be used. Each RAID array is also referred to as a rank. Each rank is divided into a number of equally sized partitions referred to as extents. The size of each extent can vary based on the implementation. For example, the size of each extent can depend, at least in part, on the extent storage type (e.g. Fixed Block (FB) or count key data (CKD)) is dependent on the type of host coupled to the storage sub-system (e.g. open-systems host or mainframe server). The extents are then grouped to make up logical volumes.

[0026] The storage sub-system 210 can enable various management features and functions, such as, but not limited to, full disk encryption, non-volatile storage (NVS) algorithms (e.g. thresholding, stage, destage), storage

pool striping (rotate extents), dynamic volume expansion, dynamic data relocation, intelligent write caching, and adaptive multi-stream prefetching. One example of a storage system 210 having an architecture similar to that illustrated in FIG. 2 is the IBM DS8000™ series enterprise storage system. The DS8000™ is a high-performance, high-capacity storage sub-system providing disk and solid-state storage that is designed to support continuous operations. Nevertheless, the embodiments of the present invention disclosed herein are not limited to the IBM DS8000™ series enterprise storage system, but can be implemented in any comparable or analogous storage system or group of storage systems, regardless of the manufacturer, product name, or components or component names associated with the system. Thus, the IBM DS8000™ is presented only by way of example and is not intended to be limiting.

[0027] Additionally, in the embodiment of the present invention shown in FIG. 2, each of the device adapters 226 includes a respective network port 232, such as an Ethernet port, which communicatively couples the device adapter 226 to cloud storage devices 214 via a network, such as the internet. In the example shown in FIG. 2, each device adapter 226 further includes a respective relocation module 232 which is configured to allocate and group cloud storage devices 214 into virtual RAID arrays, such that the cloud storage devices 214 appear to the storage controllers 222 as a local RAID array or rank. In this way, the features and functions of the storage controllers 226 that are available for local ranks, such as RAID array 234, are also available for the cloud rank 214.

[0028] As described in more detail below with respect to FIGs. 3 and 4, the relocation module 232 is configured to convert between storage controller commands and/or I/O accesses and cloud interface commands and/or I/O accesses. It is to be noted that although a relocation module 232 is included in the device adapters 226 in this example, the relocation module 232 can be included in storage controllers 222 in other embodiments. In particular, in some embodiments, each storage controller 222 includes a respective relocation module that does the conversion for commands to the respective device adapter 226.

[0029] Thus, the embodiments of the present invention described herein enable advantages over conventional cloud storage systems. For example, conventional cloud storage systems typically enable relatively basic functionality, such as remote archiving, backup, and retrieval. However, such conventional systems are unable to perform advanced management functions on the data stored in the cloud, such as the management functions mentioned above (e.g. NVS algorithms such as thresholding, stage, and destage). Thus, through the use of the relocation module 232, discussed in more detail below, the embodiments of the present invention described herein enable the performance of advanced management features on data stored on cloud storage devices which is not available for conventional cloud storage systems. In particular, through the use of the relocation module 232, the storage controllers 222 and device adapters 226 are able to access and utilize the virtual RAID arrays or ranks comprised of cloud storage as if the virtual RAID arrays were local drives coupled to the device adapters 226 rather than as remote storage. In this way, the same management features/functionality available for local drives, such as

those mentioned above, are available for the remote cloud storage without modifying the underlying code and/or hardware associated with implementing those management features.

[0030] Furthermore, by creating virtual RAID arrays that appear as local storage to the storage sub-system 210, the embodiments described herein provide a solution to a problem of efficient data relocation. In particular, the relocation module 232 is configured to monitor data accesses to data on the plurality of storage devices of one or more RAID arrays (e.g. array 234) in the storage system. Based on the monitored accesses, the relocation module 232 can identify a plurality of categories of data. For example, in some embodiments, the relocation module 232 identifies a hot data category, a cold data category and an extremely cold data category. The hot data category corresponds to data which has data accesses within a first time period. The cold data category corresponds to data which has not had a data access within the first time period, but has had a data access within a second time period that is longer or greater than the first time period. The extremely cold data category corresponds to data which has not had a data access within the first or second time periods. In other words, data in the extremely cold data category has not had a data access for more than the second time period. As described in more detail below, the relocation module 232 is configured to move data in the extremely cold category to one or more of the cloud based ranks. Thus, the relocation module 232 can be configured to create the cloud based ranks if space is not available on an existing cloud based rank of the storage system 200.

[0031] FIG. 3 is a block diagram of an example computing device 300 which can be implemented as a device adapter, such as device adapters 226 or a storage controller, such as storage controllers 222. For purposes of explanation, computing device 300 is described herein with respect to a device adapter. In the example shown in FIG. 3, the device adapter 300 includes a memory 325, storage 335, an interconnect (e.g., BUS) 340, one or more processors 305 (also referred to as CPU 305 herein), an I/O device interface 350, and a network adapter or port 315.

[0032] Each CPU 305 retrieves and executes programming instructions stored in the memory 325 and/or storage 335. The interconnect 340 is used to move data, such as programming instructions, between the CPU 305, I/O device interface 350, storage 335, network adapter 315, and memory 325. The interconnect 340 can be implemented using one or more busses. The CPUs 305 can be a single CPU, multiple CPUs, or a single CPU having multiple processing cores in various embodiments. In some embodiments of the present invention, a processor 305 can be a digital signal processor (DSP). Memory 325 is generally included to be representative of a random access memory (e.g., static random access memory (SRAM), dynamic random access memory (DRAM), or Flash). The storage 335 is generally included to be representative of a non-volatile memory, such as a hard disk drive, solid state device (SSD), removable memory cards, optical storage, or flash memory devices.

[0033] In some embodiments of the present invention, the memory 325 stores relocation instructions 301 and

the storage 335 stores map table 307. However, in various embodiments of the present invention, the relocation instructions 301 and the map table 307 are stored partially in memory 325 and partially in storage 335, or they are stored entirely in memory 325 or entirely in storage 335.

[0034] When executed by the CPU 305, the relocation instructions 301 cause the CPU 305 to utilize the map table 307 to implement the relocation module discussed above with respect to FIG. 2. It is to be noted that although the relocation instructions 301 and map table 307 are depicted as being stored in and executed/utilized by a device adapter 300, in other embodiments of the present invention the relocation instructions 301 and map table 307 can be stored on and executed/utilized by a storage controller such as storage controller 222a and/or storage controller 222b shown in FIG. 2. The relocation instructions 301 cause the CPU 305 to allocate space on cloud storage devices, such as cloud storage devices 214 depicted in FIG. 2. The space can be allocated statically or on demand as need arises. For example, the space can be allocated *a priori* or at run time. Furthermore, the cloud storage ranks can be created with different storage capacity.

[0035] The relocation instructions 301 further cause the CPU 305 to group the allocated storage into one or more virtual ranks and to store a mapping between the cloud storage devices and the one or more virtual ranks in the map table 307. In particular, the relocation instructions 301 cause the CPU 305 to generate the map table 307 which maps the allocated storage space to corresponding virtual local addresses and groups the virtual local addresses to create one or more virtual local ranks or RAID arrays. In this way, the virtual ranks of cloud storage appear as local direct attached ranks to a storage controller communicatively coupled to the device adapter 300 via the I/O device interfaces 350. The I/O device interfaces 350 also communicatively couple the device adapter 300 to local ranks of storage devices, such as solid state drives and nearline drives (e.g. storage drives 216 discussed above). For example, the I/O device interfaces 350 can include fiber optic ports. As used herein, a local rank is a rank or RAID array comprised of storage devices that are directly connected to the device adapter 300 without an intervening wide area network, such as the internet.

[0036] When an I/O access (e.g. a read or write request) is received, the cloud conversion instructions 301 cause the CPU 305 to determine if the request is directed to data stored on a virtual rank of cloud storage. When the request is directed to data stored on a virtual rank of cloud storage, the relocation instructions 301 convert the I/O access (also referred to herein as a data access request or access request) for transmission to the cloud storage device via a cloud interface. For example, the relocation instructions 301 can convert the I/O access using commands, format, device address, etc. used by the cloud interface to access the cloud storage devices. As used herein, the terms I/O access, read/write access, and data access can be used interchangeably. Exemplary cloud interfaces can include, but are not limited to, the IBM® Cloud Manager or the Amazon® Simple Storage Service (Amazon S3) interface. Thus, as discussed above, the relocation instructions 301 transparently makes cloud storage available to a storage controller similar to other local storage devices.

[0037] In addition, the relocation instructions 301 cause the CPU 305 to dynamically relocate data which is not accessed frequently to improve efficiency of the storage system. In particular, the relocation instructions 301 cause the CPU 305 to dynamically relocate data to cloud based ranks that have not been accessed for more than a predetermined amount of time. For example, the relocation instructions 301 can be configured to cause the CPU 305 to perform the dynamic relocation described in more detail with respect to FIG. 4.

[0038] In addition, the relocation instructions 303 are configured, in some embodiments of the present invention to cause the CPU 305 to assign a service level to the cloud based ranks. In some such embodiments of the present invention, there are three levels of service. However, in other embodiments of the present invention, providing multiple levels of service, two or more than 3 levels of service can be provided. In this example, three levels of service are utilized and the selection of the level of service is based on the compressibility of the data being mirrored, a respective input/output data rate for the virtual local ranks, and a service level agreement. For example, if a service level agreement indicates a low quality of service, the I/O data rate for the virtual local rank is below a threshold, and the data being accessed is compressible, then a first level of service is selected. A low quality of service can be any quality of service below a pre-defined threshold level of service. The first level of service is the lowest level of service from the three options in this example. For example, it can include higher latencies and lower throughput than the other two levels of service. If the service level agreement indicates a low quality of service, the I/O data rate for the virtual local rank is below a threshold, and the data is not compressible, then the second level of service is selected. The second level of service has greater throughput and/or less latency than the first level of service. The last or third level of service is used for all other data (e.g. the SLA indicates a level of service above the pre-defined threshold and/or the I/O data rate is above a threshold). The third level of service has greater throughput and/or less latency than both the first and second levels of service.

[0039] By providing differing levels of service, the device adapter 326 is able to leverage the virtual ranks of cloud storage to provide greater flexibility in meeting the customer needs for data storage and access. It is to be noted that although the example first, second, and third levels are described as differing in latency and throughput, other factors can be used to differentiate the levels of service. For example, in some embodiments of the present invention, the three levels of service have the same latency and throughput, but differ in cost and redundancy level.

[0040] FIG. 4 is a flow chart depicting a method 400 of dynamic data relocation. The method 400 can be implemented by a device adapter, such as device adapters 226, or a storage controller, such as storage controllers 222. For example, the method 400 can be implemented by a CPU, such as CPU 305 in computing device 300, executing instructions, such as relocation instructions 301. It is to be understood that the order of actions in example method 400 is provided for purposes of explanation and that the method can be performed in a different order in other embodiments of the present invention. Similarly, it is to be understood that some actions can be omitted or additional actions can be included in other embodiments of the present invention.

[0041] At block 402, accesses to data stored on a plurality of local ranks of an enterprise storage system is monitored. For example, device adapter and/or storage controller can be configured to count the number of accesses to data as well as track the time at which the data accesses occur. At block 404, data which has not been accessed for more than a predetermined amount of time. In some embodiments of the present invention, the predetermined amount of time is a specified number of weeks. In other embodiments of the present invention, the predetermined amount of time is a specified number of days. In other embodiments of the present invention, other amounts of time are used.

[0042] In some embodiments of the present invention, the data stored on the plurality of local ranks or RAID arrays is categorized into a plurality of categories based on the monitored data accesses. For example, in some such embodiments of the present invention, three categories of data are used, as discussed above. In other embodiments of the present invention, 4 or more categories are used. Each of the categories can be defined by the frequency of data accesses. For example, in some embodiments, the most frequently accessed (or hot data) is data which is accessed within a first, shortest time period. Similarly, data which is not accessed within the first, shortest time period, but is accessed within a second time period longer than the first time period is categorized into a second category (e.g. cold data). Thus, differing time periods can define the different categories of data. In other embodiments, the number of accesses which occur within a given amount of time can be used to define the different categories. Regardless of the manner of defining the different categories, identifying the data which has not been accessed for the predetermined amount of time involves identifying the data which is least frequently accessed. For example, if multiple time periods are used to define a plurality of categories, the predetermined amount of time used at block 404 is the longest period of time.

[0043] At block 406, the data which has not been accessed within the predetermined amount of time is moved to one or more cloud based ranks of the enterprise storage system. Each cloud based rank comprises storage space on one or more cloud storage devices. The storage space on the one or more cloud storage devices is mapped to corresponding virtual local addresses that are grouped as a virtual local rank, as discussed above.

[0044] Moving the data to a cloud based rank can include creating the cloud based rank. For example, in response to determining that a cloud based rank is not available that has sufficient space for the data identified at block 404 (e.g. the existing cloud based ranks do not have enough space or no cloud based ranks currently exist), the device adapter or storage controller can create one or more new cloud based ranks. Creating one or more new cloud based ranks includes allocating storage space on one or more corresponding cloud storage devices via a cloud interface. A cloud interface is an interface configured to enable access to the cloud storage devices. For example, the cloud interface can be implemented as an application programming interface (API). Example cloud interfaces are discussed above. Allocating the cloud storage can include requesting via the cloud interface a specified amount of storage. This storage space can be requested statically before needed or dynamically as

storage space is needed. Additionally, as used herein, cloud storage devices are storage devices accessed via a wide area network, such as the internet, as opposed to a local area network or direct communication link.

[0045] Creating the new cloud based ranks further includes mapping the allocated storage space to corresponding virtual local addresses. For example, in some embodiments of the present invention, the device adapter or storage controller assigns the allocated storage space to corresponding virtual local addresses. As used herein, a virtual local address is a memory address which appears as an address of a storage device coupled to the device adapter via a local connection. A local connection refers to a connection that is not over a wide area network or remote network, such as the internet.

[0046] Creating the new cloud based ranks further includes grouping the virtual local addresses as virtual local ranks. For example, the device adapter or storage controller is configured to group or arrange the virtual local addresses so that they appear to be directly connected to the device adapter as a local RAID array of storage devices.

[0047] Furthermore, in some embodiments of the present invention, moving the data identified at block 404 includes waiting to move the data until a predetermined condition is met. For example, in some such embodiments of the present invention, moving the data identified at block 404 includes determining the size of the data identified at block 404 and moving the data in response to determining that the size of the data exceeds a predetermined size threshold. For example, the size threshold can be set to 1 gigabyte, 10 gigabytes, or 100 gigabytes, etc., in some embodiments. In other embodiments, the size is determined as a function of extent sizes on the one or more local RAID arrays.

[0048] In other embodiments of the present invention, the device adapter or storage controller are configured to wait to move the data identified at block 404 based on the amount of available space on the one or more local RAID arrays. For example, in some such embodiments of the present invention, moving the data identified at block 404 includes determining the amount of available space remaining on the plurality of local ranks and comparing the amount of available space to a predetermined available space threshold. Once the available space on the plurality of local ranks drops below the threshold, the device adapter or storage controller directs moving the data to the cloud based ranks.

[0049] At block 408, it is determined if a data access request for at least part of the data on the cloud based ranks has been received. A data access request for data on the cloud based ranks includes an access request directed to the corresponding virtual local addresses for at least part of the data moved, at block 406, to the cloud based ranks. The device adapter or storage controller is configured to convert the access request to a cloud data access request configured for an application programming interface (API) corresponding to the one or more cloud

storage devices of the one or more cloud based ranks.

[0050] If a data access request has not been received for at least part of the data moved to the cloud based ranks, the method 400 loops at block 408 until a data access request is received. If a data access request is received for at least part of the data moved to the cloud based ranks, then the requested data is relocated to one or more storage drives of one or more local RAID arrays, at block 410.

[0051] At block 412, it is determined if data still remains on the cloud based ranks. If data still remains on each of the cloud based ranks, then method 400 loops to block 408. If data does not remain on one or more of the cloud based ranks, then the corresponding cloud based ranks which do not have any more data stored thereon are deleted at block 414. Deleting the cloud based ranks includes deallocating the storage space and deleting corresponding entries in the map table. In this way, the cloud based ranks can be dynamically created and deleted as storage space is needed.

[0052] As stated above, the order of actions in example method 400 is provided for purposes of explanation and method 400 can be performed in a different order and/or some actions can be omitted or additional actions can be included in other embodiments. Similarly, it is to be understood that some actions can be omitted or additional actions can be included in other embodiments of the present invention. For example, in some embodiments of the present invention, blocks 412 and 414 are omitted. In such embodiments of the present invention, the cloud based ranks are not deleted even if no more data remains on one or more of the cloud based ranks after relocating data to one or more of the local RAID arrays. Additionally, it is to be understood that acts do not need to be performed serially. For example, the monitoring and identifying described in blocks 402 and 404 can occur at the same time or overlap with the other acts described in method 400.

[0053] Embodiments of the present invention may be a system, a method, and/or a computer program product at any possible technical detail level of integration. The computer program product may include a computer readable storage medium (or media) having computer readable program instructions thereon for causing a processor to carry out aspects of embodiments of the present invention.

[0054] The computer readable storage medium can be a tangible device that can retain and store instructions for use by an instruction execution device. The computer readable storage medium may be, for example, but is not limited to, an electronic storage device, a magnetic storage device, an optical storage device, an electromagnetic storage device, a semiconductor storage device, or any suitable combination of the foregoing. A non-exhaustive list of more specific examples of the computer readable storage medium includes the following: a portable computer diskette, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or Flash memory), a static random access memory (SRAM), a portable compact disc

read-only memory (CD-ROM), a digital versatile disk (DVD), a memory stick, a floppy disk, a mechanically encoded device such as punch-cards or raised structures in a groove having instructions recorded thereon, and any suitable combination of the foregoing. A computer readable storage medium, as used herein, is not to be construed as being transitory signals per se, such as radio waves or other freely propagating electromagnetic waves, electromagnetic waves propagating through a waveguide or other transmission media (e.g., light pulses passing through a fiber-optic cable), or electrical signals transmitted through a wire.

[0055] Computer readable program instructions described herein can be downloaded to respective computing/processing devices from a computer readable storage medium or to an external computer or external storage device via a network, for example, the Internet, a local area network, a wide area network and/or a wireless network. The network may comprise copper transmission cables, optical transmission fibers, wireless transmission, routers, firewalls, switches, gateway computers and/or edge servers. A network adapter card or network interface in each computing/processing device receives computer readable program instructions from the network and forwards the computer readable program instructions for storage in a computer readable storage medium within the respective computing/processing device.

[0056] Computer readable program instructions for carrying out operations of embodiments of the present invention may be assembler instructions, instruction-set-architecture (ISA) instructions, machine instructions, machine dependent instructions, microcode, firmware instructions, state-setting data, configuration data for integrated circuitry, or either source code or object code written in any combination of one or more programming languages, including an object oriented programming language such as Smalltalk, C++, or the like, and procedural programming languages, such as the "C" programming language or similar programming languages. The computer readable program instructions may execute entirely on the user's computer, partly on the user's computer, as a stand-alone software package, partly on the user's computer and partly on a remote computer or entirely on the remote computer or server. In the latter scenario, the remote computer may be connected to the user's computer through any type of network, including a local area network (LAN) or a wide area network (WAN), or the connection may be made to an external computer (for example, through the Internet using an Internet Service Provider). In some embodiments of the present invention, electronic circuitry including, for example, programmable logic circuitry, field-programmable gate arrays (FPGA), or programmable logic arrays (PLA) may execute the computer readable program instructions by utilizing state information of the computer readable program instructions to personalize the electronic circuitry, in order to perform aspects of embodiments of the present invention.

[0057] Aspects of embodiments of the present invention are described herein with reference to flowchart illustrations and/or block diagrams of methods, apparatus (systems), and computer program products according to embodiments of the invention. It will be understood that each block of the flowchart illustrations and/or block diagrams, and combinations of blocks in the flowchart illustrations and/or block diagrams, can be implemented by

computer readable program instructions.

[0058] These computer readable program instructions may be provided to a processor of a general purpose computer, special purpose computer, or other programmable data processing apparatus to produce a machine, such that the instructions, which execute via the processor of the computer or other programmable data processing apparatus, create means for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks. These computer readable program instructions may also be stored in a computer readable storage medium that can direct a computer, a programmable data processing apparatus, and/or other devices to function in a particular manner, such that the computer readable storage medium having instructions stored therein comprises an article of manufacture including instructions which implement aspects of the function/act specified in the flowchart and/or block diagram block or blocks.

[0059] The computer readable program instructions may also be loaded onto a computer, other programmable data processing apparatus, or other device to cause a series of operational steps to be performed on the computer, other programmable apparatus or other device to produce a computer implemented process, such that the instructions which execute on the computer, other programmable apparatus, or other device implement the functions/acts specified in the flowchart and/or block diagram block or blocks.

[0060] The flowchart and block diagrams in the Figures illustrate the architecture, functionality, and operation of possible implementations of systems, methods, and computer program products according to various embodiments of the present invention. In this regard, each block in the flowchart or block diagrams may represent a module, segment, or portion of instructions, which comprises one or more executable instructions for implementing the specified logical function(s). In some alternative implementations, the functions noted in the blocks may occur out of the order noted in the Figures. For example, two blocks shown in succession may, in fact, be executed substantially concurrently, or the blocks may sometimes be executed in the reverse order, depending upon the functionality involved. It will also be noted that each block of the block diagrams and/or flowchart illustration, and combinations of blocks in the block diagrams and/or flowchart illustration, can be implemented by special purpose hardware-based systems that perform the specified functions or acts or carry out combinations of special purpose hardware and computer instructions.

EXAMPLES

[0061] Example 1 includes a computer-implemented method for dynamic data relocation. The method comprises monitoring accesses to data stored on a plurality of local ranks of an enterprise storage system; identifying data which has not been accessed for a predetermined amount of time based on the monitored accesses; and moving the data which has not been accessed for the predetermined amount of time to one or more

cloud based ranks of the enterprise storage system, wherein each cloud based rank comprises storage space on one or more cloud storage devices, the storage space on the one or more cloud storage devices mapped to corresponding virtual local addresses that are grouped as a virtual local rank.

[0062] Example 2 includes the computer-implemented method of example 1, further comprising, in response to an access request directed to the corresponding virtual local addresses for at least part of the moved data, converting the access request to a cloud data access request configured for an application programming interface (API) corresponding to the one or more cloud storage devices of the one or more cloud based ranks; and relocating the at least part of the moved data from the one or more cloud based ranks to at least one of the plurality of local ranks.

[0063] Example 3 includes the computer-implemented method of example 2, further comprising determining that all of the moved data has been relocated to at least one of the plurality of local ranks; and deleting the one or more cloud based ranks.

[0064] Example 4 includes the computer-implemented method of any of examples 1-3, wherein the predetermined amount of time comprises one of a predetermined number of weeks or a predetermined number of days.

[0065] Example 5 includes the computer-implemented method of any of examples 1-4, wherein moving the data comprises, in response to determining that a cloud based rank having sufficient storage space is not available, creating one or more new cloud based ranks. Creating each of the one or more new cloud based ranks comprises allocating storage space on one or more corresponding cloud storage devices via a cloud interface; mapping the allocated storage space to corresponding virtual local addresses; and grouping the virtual local addresses as a virtual local rank.

[0066] Example 6 includes the computer-implemented method of any of examples 1-5, wherein moving the data comprises moving the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that the size of the data which has not been accessed for the predetermined amount of time exceeds a predetermined size threshold.

[0067] Example 7 includes the computer-implemented method of any of examples 1-5, wherein moving the data comprises moving the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that available space on the plurality of local ranks is less than a predetermined available space threshold.

[0068] Example 8 includes a storage system. The storage system comprises a host adapter having one or more ports configured to communicatively couple the host adapter to one or more host devices; a storage controller comprising a processor and a memory, the storage controller communicatively coupled to the host adapter; and a device adapter comprising a processor and a memory, the device adapter communicatively coupled to the storage controller. The device adapter further comprises a plurality of ports communicatively coupled to a plurality of local storage drives and at least one network port communicatively coupled to a plurality of cloud storage devices via a network. One of the device adapter and the storage controller is further configured to implement a data relocation module. The data relocation module is configured to monitor accesses to data stored on a plurality of local redundant array of independent disks (RAID) arrays of an enterprise storage system; identify data which has not been accessed for a predetermined amount of time based on the monitored accesses; and move the data which has not been accessed for the predetermined amount of time to one or more cloud based RAID arrays of the enterprise storage system, wherein each cloud based RAID array comprises storage space on one or more cloud storage devices, the storage space on the one or more cloud storage devices mapped to corresponding virtual local addresses that are grouped as a virtual local RAID array.

[0069] Example 9 includes the storage system of example 8, wherein the relocation module is further configured to, in response to an access request directed to the corresponding virtual local addresses for at least part of the moved data, relocate the at least part of the moved data from the one or more cloud based ranks to at least one of the plurality of local ranks.

[0070] Example 10 includes the storage system of example 9, wherein the relocation module is further configured to delete the one or more cloud based ranks in response to determining that all of the moved data has been relocated to at least one of the plurality of local ranks.

[0071] Example 11 includes the storage system of any of examples 8-10, wherein the relocation module is configured to move the data by creating one or more new cloud based ranks in response to determining that a cloud based rank having sufficient storage space is not available. The relocation module is configured to create each of the one or more new cloud based ranks by allocating storage space on one or more corresponding cloud storage devices via a cloud interface; mapping the allocated storage space to corresponding virtual local addresses; and grouping the virtual local addresses as a virtual local rank.

[0072] Example 12 includes the storage system of any of examples 8-11, wherein the relocation module is configured to move the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that the size of the data which has not been accessed for the predetermined amount of time exceeds a predetermined size threshold.

[0073] Example 13 includes the storage system of any of examples 8-12, wherein the relocation module is configured to move the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that available space on the plurality of local ranks is less than a predetermined available space threshold.

[0074] Example 14 includes a computer program product comprising a computer readable storage medium having a computer readable program stored therein. The computer readable program, when executed by a processor, causes the processor to monitor accesses to data stored on a plurality of local redundant array of independent disks (RAID) arrays of an enterprise storage system; identify data which has not been accessed for a predetermined amount of time based on the monitored accesses; and move the data which has not been accessed for the predetermined amount of time to one or more cloud based RAID arrays of the enterprise storage system, wherein each cloud based RAID array comprises storage space on one or more cloud storage devices, the storage space on the one or more cloud storage devices mapped to corresponding virtual local addresses that are grouped as a virtual local RAID array.

[0075] Example 15 includes the computer program product of example 14, wherein the computer readable program is further configured to cause the processor to, in response to an access request directed to the corresponding virtual local addresses for at least part of the moved data, convert the access request to a cloud data access request configured for an application programming interface (API) corresponding to the one or more cloud storage devices of the one or more cloud based ranks; and relocate the at least part of the moved data from the one or more cloud based ranks to at least one of the plurality of local ranks.

[0076] Example 16 includes the computer program product of example 15, wherein the computer readable program is further configured to cause the processor to delete the one or more cloud based ranks in response to determining that all of the moved data has been relocated to at least one of the plurality of local ranks.

[0077] Example 17 includes the computer program product of any of examples 14-16, wherein the computer readable program is further configured to cause the processor to move the data by creating one or more new cloud based ranks in response to determining that a cloud based rank having sufficient storage space is not available. The relocation module is configured to create each of the one or more new cloud based ranks by allocating storage space on one or more corresponding cloud storage devices via a cloud interface; mapping the allocated storage space to corresponding virtual local addresses; and grouping the virtual local addresses as a virtual local rank.

[0078] Example 18 includes the computer program product of any of examples 14-17, wherein the computer readable program is further configured to cause the processor to move the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that the size of

the data which has not been accessed for the predetermined amount of time exceeds a predetermined size threshold.

[0079] Example 19 includes the computer program product of any of examples 14-18, wherein the computer readable program is further configured to cause the processor to move the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that available space on the plurality of local ranks is less than a predetermined available space threshold.

[0080] Example 20 includes a computing device. The computing device comprises a network adapter configured to communicatively couple the computing device to one or more cloud storage devices via a network; a storage medium configured to store data; and a processor communicatively coupled to the network adapter and to the storage medium. The processor is configured to monitor accesses to data stored on a plurality of local ranks of an enterprise storage system; identify data which has not been accessed for a predetermined amount of time based on the monitored accesses; and move the data which has not been accessed for the predetermined amount of time to one or more cloud based ranks of the enterprise storage system, wherein each cloud based rank comprises storage space on one or more cloud storage devices, the storage space on the one or more cloud storage devices mapped to corresponding virtual local addresses that are grouped as a virtual local rank.

[0081] Example 21 includes the computing device of example 20, wherein the processor is further configured to, in response to an access request directed to the corresponding virtual local addresses for at least part of the moved data, convert the access request to a cloud data access request configured for an application programming interface (API) corresponding to the one or more cloud storage devices of the one or more cloud based ranks; and relocate the at least part of the moved data from the one or more cloud based ranks to at least one of the plurality of local ranks.

[0082] Example 22 includes the computing device of example 21, wherein the processor is further configured to delete the one or more cloud based ranks in response to determining that all of the moved data has been relocated to at least one of the plurality of local ranks.

[0083] Example 23 includes the computing device of any of examples 20-22, wherein the processor is further configured to move the data by creating one or more new cloud based ranks in response to determining that a cloud based rank having sufficient storage space is not available. The processor is configured to create each of the one or more new cloud based ranks by allocating storage space on one or more corresponding cloud storage devices via an application programming interface (API) associated with the corresponding one or more cloud storage devices; and generating a map table which maps the allocated storage space to corresponding virtual local addresses and groups the virtual local addresses as a corresponding virtual local rank.

[0084] Example 24 includes the computing device of any of examples 20-23, wherein the processor is further configured to move the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that the size of the data which has not been accessed for the predetermined amount of time exceeds a predetermined size threshold.

[0085] Example 25 includes the computing device of any of examples 20-24, wherein the processor is further configured to move the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that available space on the plurality of local ranks is less than a predetermined available space threshold.

[0086] Although specific embodiments of the present invention have been illustrated and described herein, it will be appreciated by those of ordinary skill in the art that any arrangement, which is calculated to achieve the same purpose, may be substituted for the specific embodiments of the present invention shown. Therefore, it is manifestly intended that this invention be limited only by the claims and the equivalents thereof.

CLAIMS

1. A computer-implemented method for dynamic data relocation, the method comprising:
 - monitoring accesses to data stored on a plurality of local ranks of an enterprise storage system;
 - based on the monitored accesses, identifying data which has not been accessed for a predetermined amount of time;
 - moving the data which has not been accessed for the predetermined amount of time to one or more cloud based ranks of the enterprise storage system, wherein each cloud based rank comprises storage space on one or more cloud storage devices, the storage space on the one or more cloud storage devices mapped to corresponding virtual local addresses that are grouped as a virtual local rank;
 - wherein the virtual local addresses are memory addresses which appear as addresses of a storage device coupled to a storage controller via a local connection;
 - in response to an access request directed to the corresponding virtual local addresses for at least part of the moved data, converting the access request to a cloud data access request configured for an application programming interface (API) corresponding to the one or more cloud storage devices of the one or more cloud based ranks;
 - relocating the at least part of the moved data from the one or more cloud based ranks to at least one of the plurality of local ranks;
 - determining that all of the moved data has been relocated to at least one of the plurality of local ranks; and deleting the one or more cloud based ranks.
2. The method of claim 1, wherein the predetermined amount of time comprises one of a predetermined number of weeks or a predetermined number of days.
3. The method of claim 1, wherein moving the data comprises:
 - in response to determining that a cloud based rank having sufficient storage space is not available, creating one or more new cloud based ranks;
 - wherein creating each of the one or more new cloud based ranks comprises:
 - allocating storage space on one or more corresponding cloud storage devices via a cloud interface;
 - mapping the allocated storage space to corresponding virtual local addresses; and
 - grouping the virtual local addresses as a virtual local rank.
4. The method of claim 1, wherein moving the data comprises moving the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that the size of the data which has not been accessed for the predetermined amount of time exceeds a predetermined size threshold.

5. The method of claim 1, wherein moving the data comprises moving the data which has not been accessed for the predetermined amount of time to the one or more cloud based ranks in response to determining that available space on the plurality of local ranks is less than a predetermined available space threshold.

6. The method of claim 1, wherein the ranks comprise RAID arrays.

7. A storage system comprising:

a host adapter having one or more ports configured to communicatively couple the host adapter to one or more host devices;

a storage controller comprising a processor and a memory, the storage controller communicatively coupled to the host adapter; and

a device adapter comprising a processor and a memory, the device adapter communicatively coupled to the storage controller;

wherein the device adapter further comprises a plurality of ports communicatively coupled to a plurality of local storage drives and at least one network port communicatively coupled to a plurality of cloud storage devices via a network;

wherein one of the device adapter and the storage controller is further configured implement a data relocation module configured to perform the steps of a method as claimed in any of claims 1 to 6.

8. A computer program product comprising a computer readable storage medium having a computer readable program stored therein, wherein the computer readable program, when executed by a processor, causes the processor to perform the steps of the method as claimed in any of claims 1 to 6.

9. A computing device comprising:

a network adapter configured to communicatively couple the computing device to one or more cloud storage devices via a network;

a storage medium configured to store data; and

a processor communicatively coupled to the network adapter and to the storage medium, wherein the processor is configured to perform the steps of the method as claimed in any of claims 1 to 6.