

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 October 2004 (21.10.2004)

PCT

(10) International Publication Number
WO 2004/090742 A1

(51) International Patent Classification⁷: G06F 15/173

KUMAR, Manoj; 2140 Aldrin Road, Apt. 2B, Ocean, NJ 07712 (US). NGO, David; 152 Dorchester Way, Shrewsbury, NJ 07702 (US). DEVASSY, Varghese; 2 Norcross Crescent, Winnipeg, Manitoba R3X 1H8 (CA).

(21) International Application Number:
PCT/US2004/010503

(22) International Filing Date: 5 April 2004 (05.04.2004)

(74) Agent: WOODS, James, J.; Brown, Raysman, Millstein, Felder & Steiner, LLP, 900 Third Avenue, New York, NY 10022 (US).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/460,234 3 April 2003 (03.04.2003) US

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

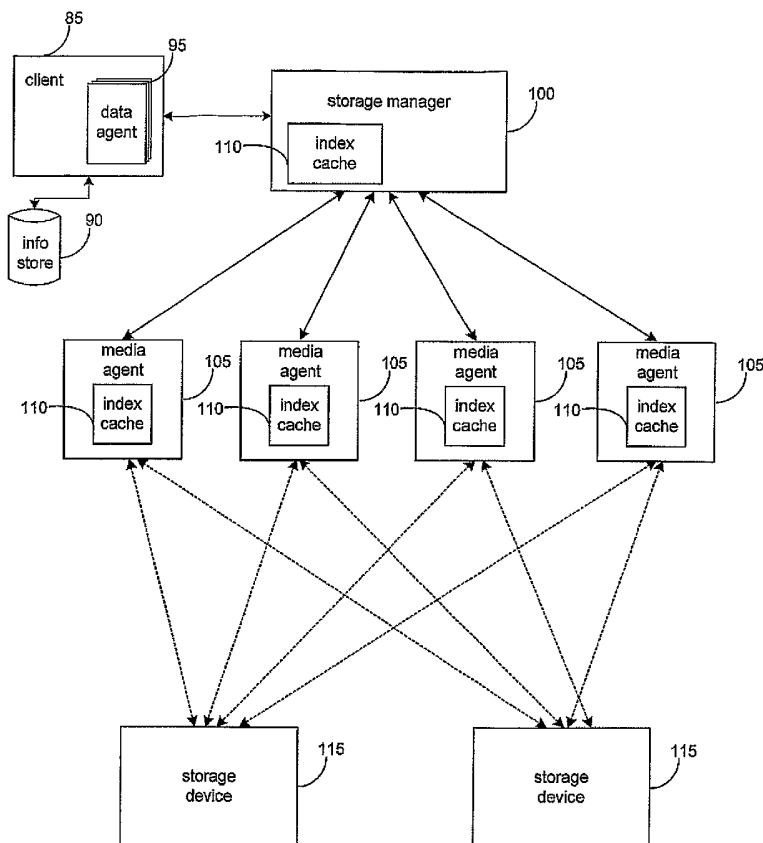
(71) Applicant: COMMVault SYSTEMS, INC. [US/US];
Two Crescent Place, Oceanport, NJ 07757-0090 (US).

(72) Inventors: KOTTOMTHARAYIL, Rajiv; 1508 Garden Drive, Apt. 2, Ocean, NJ 07712 (US). GOKHALE, Parag; 21 Oak Mews Court, Ocean, NJ 07712 (US). PRAHLAD, Anand; 3 Bucknell Drive, East Brunswick, NJ 08816 (US).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR DYNAMICALLY SHARING STORAGE VOLUMES IN A COMPUTER NETWORK



(57) Abstract: A computerized method for dynamically sharing a magnetic media volume in a network, the method comprising removing, in response to a volume sharing request, an association between a first media management component (105) and the magnetic media volume, and associating, according to a set of selection logic, a second media management component (105) with the magnetic media volume.

WO 2004/090742 A1



Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

Published:

— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

SYSTEM AND METHOD FOR DYNAMICALLY SHARING
STORAGE VOLUMES IN A COMPUTER NETWORK

COPYRIGHT NOTICE

A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosures, as it appears in the Patent and Trademark Office patent files or records, but otherwise reserves all copyright rights whatsoever.

RELATED APPLICATIONS

This application claims priority to Provisional Application number 60/460,234, titled SYSTEM AND METHOD FOR PERFORMING STORAGE OPERATIONS IN A COMPUTER NETWORK, filed April 3, 2003 which is hereby incorporated herein by reference in its entirety.

This application is also related to the following patents and pending applications, each of which is hereby incorporated herein by reference in its entirety:

- U.S. Patent No. 6,418,478, titled PIPELINED HIGH SPEED DATA TRANSFER MECHANISM, issued July 9, 2002, attorney docket number 4982/6;
- Application Serial No. 09/610,738, titled MODULAR BACKUP AND RETRIEVAL SYSTEM USED IN CONJUNCTION WITH A STORAGE AREA NETWORK, filed July 6, 2000, attorney docket number 4982/8;
- Application Serial No. 09/744,268, titled LOGICAL VIEW AND ACCESS TO PHYSICAL STORAGE IN MODULAR

DATA AND STORAGE MANAGEMENT SYSTEM, filed January 30, 2001, attorney docket number 4982/10; and Application Serial No. 60/409,183, titled DYNAMIC STORAGE DEVICE POOLING IN A COMPUTER SYSTEM, filed September 9, 2002, attorney docket number 4982/18PROV.

BACKGROUND OF THE INVENTION

The invention disclosed herein relates generally to performing storage operations on electronic data in a computer network. More particularly, the present invention relates to selecting, in response to the initiation of a storage operation and according to selection logic, a media management component and a network storage device to perform storage operations on electronic data.

Storage of electronic data has evolved through many forms. During the early development of the computer, storage of this data was limited to individual computers. Electronic data was stored in the Random Access Memory (RAM) or some other storage medium such as a hard drive or tape drive that was an actual part of the individual computer.

Later, with the advent of networked computing, storage of electronic data gradually migrated from the individual computer to stand-alone storage devices accessible via a network. These individual network storage devices soon evolved in the form of networkable tape drives, optical libraries, Redundant Arrays of Inexpensive Disks (RAID), CD-ROM jukeboxes, and other devices. Common architectures included drive pools, which generally are logical collections of drives with associated media groups including the tapes or other storage media used by a given drive pool.

Serial, parallel, Small Computer System Interface (SCSI), or other cables, directly connected these stand-alone storage devices to individual computers that were part of a network of other computers such as a Local Area Network (LAN) or a Wide Area Network (WAN). Each individual computer on the network controlled the storage devices that were physically attached to that computer and could also access the storage devices of the other network computers to perform backups, transaction processing, file sharing, and other storage-related operations.

Network Attached Storage (NAS) is another storage scheme using stand-alone storage devices in a LAN or other such network. In NAS, a storage controller computer still "owns" the storage device to the exclusion of other computers on the network, but the SCSI or other cabling directly connecting that storage device to the individual controller or owner computer is eliminated. Instead, storage devices are directly attached to the network itself.

Yet another network storage scheme is modular storage architecture which is more fully described in Application No. 09/610,738 and Application No. 09/744,268. An example of such a software application is the Galaxy™ system, by CommVault Systems of Oceanport, New Jersey. The Galaxy™ system is a multi-tiered storage management solution which includes, among other components, a storage manager, one or more media agents, and one or more storage devices. The storage manager directs storage operations of client data to storage devices such as magnetic and optical media libraries. Media agents are storage controller computers that serve as intermediary devices managing the flow of data from client information stores to individual storage devices. Each storage device is uniquely associated with a particular media agent and this association is tracked by the storage manager.

A common feature shared by all of the above-described network architectures is the static relationship between storage controller computers and storage devices. In these traditional network architectures, storage devices can each only be connected, virtually or physically, to a single storage controller computer. Only the storage controller computer to which a particular device is physically connected has read/write access to that device. A drive pool and its associated media group, for example, can only be controlled by the computer to which it is directly connected. Therefore, all backup from other storage controller computers needs to be sent via the network before it can be stored on the storage device connected to the first storage controller computer.

At times, storage solutions in some of the above-described network architectures including LAN, NAS, and modular storage systems may cause overloading of network traffic during certain operations associated with use of storage devices on the network. The network cable has a limited amount of bandwidth that must be shared among all the computers on the network. The capacity of most LAN or network cabling is measured in megabits per second (mbps) with 10mbps and 100mbps being standard. During common operations such as system backups, transaction processing, file copies, and other similar operations, network traffic often becomes overloaded as hundreds of megabytes (MB) and gigabytes (GB) of information are sent over the network to the associated storage devices. The capacity of the network computers to stream data over the network to the associated storage devices in this manner is greater than the bandwidth capacity of the cabling itself so ordinary network activity and communication slows to a crawl. As long as the storage devices are attached to the LAN or other network, this bandwidth issue remains a problem.

The Storage Area Network (SAN) is a highly-evolved network architecture designed to facilitate transport of electronic data and address this bandwidth issue. SAN architecture requires at least two networks. First, there is the traditional network described above which is typically a LAN or other such network designed to transport ordinary traffic between network computers. Then, there is the SAN itself which is a second network that is attached to the servers of the first network. The SAN is a separate network generally reserved for bandwidth-intensive operations such as backups, transaction processing, and the like also described above. The cabling used in the SAN is usually of much higher bandwidth capacity than that used in the first network such as the LAN and the communication protocols used over the SAN cabling are optimized for bandwidth-intensive traffic. Most importantly, the storage devices used by the network computers for the bandwidth-intensive operations are attached to the SAN rather than the LAN. Thus, when the bandwidth-intensive operations are required, they take place over the SAN and the LAN remains unaffected.

CommVault's proprietary DataPipe™ mechanism further described in U.S. Patent No. 6,418,478 is used with a SAN to further reduce bandwidth constraints. The DataPipe™ is the transport protocol used to facilitate and optimize electronic data transfers taking place over a Storage Area Network (SAN) as opposed to those taking place over a LAN using NAS.

None of these solutions, however, address the static relationship between individual storage controller computers and individual storage devices. LANs, WANs, and even SANs using a DataPipe™ all require a static relationship between storage controller computer and storage device since each storage device on the network is uniquely owned by a storage controller computer. As discussed, when

a storage device in this traditional architecture is assigned to a storage controller computer, that storage controller computer owns the device indefinitely and to the exclusion of other computers on the network. This is also true with both logical and physical storage volumes. One computer cannot control the drive pool and media group being that is controlled by another. Requests to store and retrieve data from such a drive pool and media group would have to first pass through the controlling computer. Such a static relationship between storage controller computer and storage device often leads to an inefficient use of resources.

For example, if each storage controller computer needs access to two storage devices and there are five storage controller computers in the network, then a total of ten storage devices will be required. The actual amount of work each of the ten storage devices performs might be much less than the workload capacity of each storage device. Such underutilization of storage device resources cannot be solved when a static relationship is required between storage device and storage controller computer.

If the static relationship were dynamic, however, and storage controller computers could actually share devices, then this underutilization can be addressed. Assuming in the above example that each of the five storage controller computers only uses ten percent of each device's workload capacity, then if all the storage controller computers could actually share the same two storage devices, eight of the storage devices could be eliminated without loss of performance or capability.

Furthermore, none of these existing solutions provide access to storage devices in the event of a storage controller failure. For example, if a storage controller computer were unavailable due to a hardware or software malfunction, then other computers on the network would not be able to access data stored on any storage

device associated with the storage controller computer. Until the storage controller computer was brought back online, the data contained on any associated storage device would be effectively unrecoverable. If the association between the storage controller computer and a storage device were not static, however, then another storage controller computer could bypass the unavailable storage controller computer and access the storage device to retrieve the data.

There is thus also a need for a system which enables dynamic association of storage controller computers and storage devices

SUMMARY OF THE INVENTION

The present invention addresses, among other things, the problems discussed above performing storage operations on electronic data in a computer network.

In one embodiment, a computerized method for dynamically sharing a magnetic media volume in a network is presented, the method comprising removing, in response to a volume sharing request, an association between a first media management component and the magnetic media volume, and associating, according to a set of selection logic, a second media management component with the magnetic media volume. In various embodiments, volume sharing requests may comprise requests in response to a user preference, requests according to a storage policy, requests initiated by a storage manager component, and requests initiated by a client component. In other embodiments, removing an association between the first media management component and the magnetic media volume comprises removing an association between a media agent and the magnetic media volume or removing an index entry association between the first media management component and the magnetic media volume. In some embodiments, associating according to a set of

selection logic comprises associating according to a user preference, a storage policy, or a request initiated by a storage manager component. In some embodiments, associating a second media management component and the magnetic media volume comprises associating a media agent and the magnetic media volume. In other embodiments, associating a second media management component and the magnetic media volume comprises creating an index entry association between the second media management component and the magnetic media volume.

In one embodiment, a system for dynamically sharing a magnetic media volume in a network is presented, the system comprising, a first media management component communicatively coupled to a magnetic media volume and a storage manager component, and a second media management component communicatively coupled to the magnetic media volume and the storage manager component, wherein the storage manager component is programmed to remove, in response to a volume sharing request, an association between the first media management component and the magnetic media volume, and associate, according to a set of selection logic, the second media management component with the magnetic media volume.

In one embodiment, a computerized method for dynamically sharing a magnetic media volume in a network is presented, the method comprising directing, in response to a volume sharing request, a first media management component controlling a magnetic media volume to go into a deactivated state with respect to the magnetic media volume; directing, according to a set of selection logic, a second media management component to go into an activated state with respect to the magnetic media volume; and mounting the magnetic media volume to the second media management component.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention is illustrated in the figures of the accompanying drawings which are meant to be exemplary and not limiting, in which like references are intended to refer to like or corresponding parts, and in which:

Fig. 1 is a block diagram of a network architecture for a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention;

Fig. 2 is a block diagram of an exemplary tape library storage device for a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention; and

Fig. 3 is a block diagram of an exemplary magnetic media storage device for a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention;

Fig. 4 is a flow chart of a method for performing storage operations on electronic data in a computer network according to an embodiment of the invention;

Fig. 5 is a flow chart of a method to archive electronic data in a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention;

Fig. 6 is a flow chart of a method for restoring or auxiliary copying electronic data in a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention;

Fig. 7 is a flow chart of a method to restore a storage index in a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention; and

Fig. 8 is a flow diagram of a method to perform dynamic volume sharing according to one embodiment of the invention.

DETAILED DESCRIPTION

With reference to Figs. 1 through 7, embodiments of the invention are presented. Fig. 1 presents a block diagram of a network architecture for a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention. As shown, the system includes a storage manager 100 and one or more of the following: a client 85, an information store 90, a data agent 95, a media agent 105, an index cache 110, and a storage device 115. The system and elements thereof are exemplary of a three-tier backup system such as the CommVault Galaxy backup system, available from CommVault Systems, Inc. of Oceanport, NJ, and further described in Application Number 09/610,738 which is incorporated herein by reference in its entirety.

A data agent 95 is generally a software module that is generally responsible for archiving, migrating, and recovering data of a client computer 85 stored in an information store 90 or other memory location. Each client computer 85 has at least one data agent 95 and the system can support many client computers 85. The system provides a plurality of data agents 95 each of which is intended to backup, migrate, and recover data associated with a different application. For example, different individual data agents 95 may be designed to handle Microsoft Exchange data, Lotus Notes data, Microsoft Windows 2000 file system data, Microsoft Active Directory Objects data, and other types of data known in the art.

If a client computer 85 has two or more types of data, one data agent 95 is generally required for each data type to archive, migrate, and restore the client computer 85 data. For example, to backup, migrate, and restore all of the data on a

Microsoft Exchange 2000 server, the client computer 85 would use one Microsoft Exchange 2000 Mailbox data agent 95 to backup the Exchange 2000 mailboxes, one Microsoft Exchange 2000 Database data agent 95 to backup the Exchange 2000 databases, one Microsoft Exchange 2000 Public Folder data agent 95 to backup the Exchange 2000 Public Folders, and one Microsoft Windows 2000 File System data agent 95 to backup the client computer's 85 file system. These data agents 95 would be treated as four separate data agents 95 by the system even though they reside on the same client computer 85.

The storage manager 100 is generally a software module or application that coordinates and controls the system. The storage manager 100 communicates with all elements of the system including client computers 85, data agents 95, media agents 105, and storage devices 115, to initiate and manage system backups, migrations, and recoveries.

A media agent 105 is generally a software module that conducts data, as directed by the storage manager 100, between the client computer 85 and one or more storage devices 115 such as a tape library, a magnetic media storage device, an optical media storage device, or other storage device. The media agent 105 is communicatively coupled with and controls the storage device 115. For example, the media agent 105 might instruct the storage device 115 to use a robotic arm or other means to load or eject a media cartridge, and to archive, migrate, or restore application specific data. The media agent 105 generally communicates with the storage device 115 via a local bus such as a SCSI adaptor. In some embodiments, the storage device 115 is communicatively coupled to the data agent 105 via a Storage Area Network ("SAN").

Each media agent 105 maintain an index cache 110 which stores index data the system generates during backup, migration, and restore storage operations as further described herein. For example, storage operations for Microsoft Exchange data generate index data. Index data provides the system with an efficient mechanism for locating user files for recovery operations. This index data is generally stored with the data backed up to the storage device 115, and the media agent 105 that controls the storage operation also writes an additional copy of the index data to its index cache 110. The data in the media agent 105 index cache 110 is thus readily available to the system for use in storage operations and other activities without having to be first retrieved from the storage device 115.

The storage manager 100 also maintains an index cache 110. Index data is also used to indicate logical associations between components of the system, user preferences, management tasks, and other useful data. For example, the storage manager 100 might use its index cache 110 to track logical associations between media agents 105 and storage devices 115.

Index caches 110 typically reside on their corresponding storage component's hard disk or other fixed storage device. Like any cache, the index cache 110 has finite capacity and the amount of index data that can be maintained directly corresponds to the size of that portion of the disk that is allocated to the index cache 110. In one embodiment, the system manages the index cache 110 on a least recently used ("LRU") basis as known in the art. When the capacity of the index cache 110 is reached, the system overwrites those files in the index cache 110 that have been least recently accessed with the new index data. In some embodiments, before data in the index cache 110 is overwritten, the data is copied to an index cache 110 copy in a storage device 115. If a recovery operation requires data that is no longer stored in

the index cache 110, such as in the case of a cache miss, the system recovers the index data from the index cache 110 copy stored in the storage device 115.

In some embodiments, components of the system may reside and execute on the same computer. In some embodiments, a client computer 85 component such as a data agent 95, a media agent 105, or a storage manager 100 coordinates and directs local archiving, migration, and retrieval application functions as further described in Application Number 09/610,738. This client computer 85 component can function independently or together with other similar client computer 85 components.

Fig. 2 presents a block diagram of an exemplary tape library storage device 120 for a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention. The tape library storage device 120 contains tapes 140 and drives 125, 130, and 135. The tapes 140 store electronic data containing backups of application data, user preferences, system information, and other useful information known in the art. Drives 125, 130, and 135 are used to store and retrieve electronic data from the tapes 140. In some embodiments, drives 125, 130, and 135 function as a drive pool, as further described in Application Serial No. 60/409,183 which is hereby incorporated herein by reference in its entirety. A drive pool is generally a logical concept associated with a storage policy. Storage policies representing storage patterns and preferences are more fully discussed in Application Serial No. 09/744,268 which is hereby incorporated by reference herein in its entirety.

The drive pool is identified by a set of drives within a tape library storage device 120 as pointed to by one or more media agents 105. For example, a drive pool known as DP1 consisting of drives 125 and 130 in tape library 120 known

as LIB1 could be associated by a storage policy with a first media agent 105 MA1 in an index cache 110 entry as follows: LIB1/MA1/DP1. A second drive pool consisting of drives 130, 135, and 140 within the tape library storage device 120 associated with the same media agent 105 might be expressed in the index cache 110 as follows: LIB1/MA1/DP2.

Previously, however, drive pools had a static association with a particular media agent 105. A drive pool associated with a first media agent 105 could not be logically associated with a second drive pool associated with a second media agent 105. An index cache 110 entry associating a drive pool with a media agent 105 and other system components might, for example, specify a media library or media group, a media agent, and a drive pool. Only one of each component could be specified in a particular index cache 110 entry. Thus, such drive pools were logically exclusive and separate entries in an index cache 110 entry for a storage policy and could be logically represented as:

LIB1/MA1/DP1
LIB1/MA2/DP2

As further described herein, the present invention permits logical association of drive pools associated with different media agents 105. Multiple drive pools, media agents, and other system components can be associated in a single index cache 110 entry. Thus, for example, an index cache 110 entry for a storage policy, according to the present invention, may combine the two previous entries instead and thus be logically represented as:

LIB1/MA1/DP1/DP2
LIB1/MA2/DP2/DP1

In addition and as further described herein, tapes 140 are associated by the system with drive pools or storage policies, and not with individual drives 125, 130, and 135. The recording format used to archive electronic data is a property of both its media group and its associated drive pool. A media group is a collection of tapes 140 or other storage media assigned to a specific storage policy. The media group dynamically points to different drive pools, even to those with different recording formats since the system updates the recording format of the media group in a media group table stored in the index cache 110 of the storage manager 100. Previously, tapes 140 associated with drive pools could share individual drives 125, 130, and 135, but tapes 140 associated with each drive pool remained logically separate by, among other factors, media groups. Thus, a given set of tapes 140 associated with a particular drive pool and storing data associated with a first storage policy copy could not also store data from a second storage policy copy. An index cache 110 entry would thus associate different tape sets 140 with different media agents 105, storage policies, drive pools, and other system components. For example, two different tape sets might be associated in two index cache 110 entries as follows:

storage policy1:media agent1:drive pool1:tape set1
storage policy2:media agent2:drive pool2:tape set2

All components are thus uniquely associated and require separate index cache 110 entries.

By contrast, the present invention, as further described herein, permits data associated with a particular storage policy copy to be stored on and share tapes 140, and other pieces or removable media such as optical discs, associated with and sharing one or more drive pools or storage policy copies. Data from each storage policy copy is appended to tapes 140 shared by other storage policy copies. Thus, a

storage policy copy is shared between many media agents 105 in a dynamic drive pooling setting with tapes 140 also being shared by the different media agents 105 and storage policies. Tapes 140 can be located in any storage device 115 and tapes 140 for a given storage policy copy can even be spread across multiple storage devices 115. Thus, an index cache entry would associate multiple tape sets 140 with multiple media agents, storage policies, drive pools, and other system components. For example, two different tape sets from the previous example of index entries might be associated in a single index cache 110 entry as follows:

storage policy1:media agent1:drive pool1:tape set1:tape set2
storage policy2:media agent2:drive pool2:tape set1:tape set2

In addition to tape sets 140, a single index cache 110 entry can also specify and associate multiple media agents 105, storage policies, drive pools, network pathways, and other components.

Similarly, different media agents 105 and storage policies can also be associated with the same volumes on magnetic media. For example, turning to Fig. 3, a block diagram is presented of an exemplary magnetic media storage device for a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention. A magnetic media storage device 150, such as a hard disk, is shown. The magnetic media storage device 150 is divided into two volumes 155 and 160 which are associated with a first media agent 165 and a second media agent 170. Previously, each volume on a magnetic media storage device 150 maintained a static relationship with a given media agent 105. For example, if the first volume 155 was associated with the first media agent 165, then the second media agent 170 would only be able to conduct storage operations with the

first volume 155 by passing any associated electronic data through the first media agent 165.

The present invention, however, permits media agents 105 to share the same volume(s) on a given magnetic storage device. Thus, as further described herein, a first media agent 105 can perform, on a given magnetic media volume, storage operations on data associated with a first storage policy copy, and a second media agent can perform, on the same magnetic media volume, storage operations on different data associated with a second storage policy copy. For example, media agent 165 and 170 can both perform storage operations on the first volume 155 or the second volume 160.

Fig. 4 presents a flow chart of a method for performing storage operations on electronic data in a computer network according to an embodiment of the invention. Selection of desired storage components for storage operations is performed dynamically. The system initiates a storage operation, step 175, in response to a scheduled procedure or as directed by a user. For example, the system might initiate a backup operation or a restore operation at a specific time of day or in response to a certain threshold being passed as specified in a storage policy. The system selects a media agent 105 according to selection logic further described herein, step 180. Some examples of selection logic include the ability to conduct a LAN-free storage operation, such as using a SAN, and the desire to optimize storage operations via load balancing. For example, an index entry in the storage manager 100 index cache 110 might associate certain media agents 105, storage devices 115, or other components with LAN-free storage operations either via user input, network topology detection algorithms known in the art, or other methods. As another example, the system might select a free media agent 105 to optimize storage operations via load

balancing when a default media agent 105 or other media agent 105 specified in a storage policy is already performing other storage operations or otherwise occupied. The system also selects an appropriate drive pool in a network storage device according to selection logic further described herein, step 185. Once the system has selected an appropriate media agent and drive pool, the storage operation is performed, step 190 using the selected storage components.

Fig. 5 presents a flow chart of a method to archive electronic data in a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention. More specifically, Fig. 5 presents a method for dynamically selecting a drive pool according to an embodiment of the invention. The system makes a call to reserve one or more archive streams, step 195. Archive streams are generally data paths with permit storage operations to be performed in parallel on electronic data. An archive stream generally has a one-to-one correlation with a media group, such as a media agent 105 and storage device 115.

Thus, according to one embodiment of the invention, the number of archive streams allowed on a storage policy is the sum of all drives in all available drive pools. For example, a system with three drive pools composed of ten total drives could reserve ten archive streams to copy data to each of the drives simultaneously thus increasing storage efficiency and achieving other useful goals. In one embodiment, the system uses the number of drives in a selected drive pool as the default number of streams to reserve. In an alternate embodiment, the system uses magnetic storage libraries as storage devices 115 instead of tape libraries containing drives. Here, the maximum number of streams would equal the number of readers and writers on all mount paths to the magnetic storage libraries.

The system selects one or more appropriate storage locations for the data to be archived, step 200. In some embodiments, a resource manager module associated with the storage manager 100 selects the storage location(s). For example, a resource manager determines the appropriate storage device 115, media agent 105, and drive pool combination based on the pool of available storage network components and other criteria.

Additional criteria used in the selection process generally consider *improving efficiency of the storage operation to be performed*. Load balancing, for example, is one exemplary consideration. The system may contain a given number of drives, but some of those drives may be used by other jobs and thus unavailable. The system thus selects from among available drives as one selection criterion. Alternatively, the network path to a particular storage component may be experiencing heavy traffic and thus a less trafficked path offering greater bandwidth may be selected as desirable.

Another exemplary selection criterion is whether selection of a given drive or set of drives would enable LAN-free or auxiliary archiving. For example, in a given network, certain drives might be accessible via a SAN or other alternate storage route which would reduce the network traffic caused by an archiving operation. Preference, is thus given to selection of these drives over drives which would instead increase the network load.

Yet another exemplary selection criterion is in the case of a storage component failover situation. For example, where an entire media agent 105 or storage device 115 as a whole is offline or if a certain number of drives in a storage device are offline, then the system, in some embodiments, dynamically selects an alternate media agent 105 or drive pool to perform storage operations. In some

embodiments, the alternate media agent 105 or drive pool in the case of failover is specified according to preferences associated with a storage policy. For example, a storage policy may contain a list of failover candidates or selection logic, as described herein, for selecting a storage location to serve as a failover candidate. In some embodiments, the failover candidates are expressed as a triplet indicating the media agent 105, the storage device 115, and, provided the storage device is not a magnetic media storage device 150, the drive pool. Thus, the triplet "LIB1/MA2/DP1" might be used to represent a failover candidate storage path using media agent 115 MA2, storage device 115 LIB1, and drive pool DP1. In other embodiments, a user specifies the alternate storage device 115, media agent 105, or drive pool directly.

The system reserves the selected storage candidates, step 205, and returns the storage IDs of successful reservations to the storage manager 100, step 210. In some embodiments, the storage IDs are unique identifiers enabling components of the system to identify and communicate with the storage candidates. For example, in some embodiments, the storage IDs comprise a unique component name coupled with a network path such as a Uniform Naming Convention ("UNC") entry. The storage IDs of the successful reservations are used to update a media group table stored in the storage manager 100 index cache 110 or other similar locations accessible to components of the system, step 215. The reserved components are thus accessible to other components of the system to perform the archive operation or other storage operations.

Fig. 6 presents a flow chart of a method to restore or auxiliary copy electronic data in a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention. A restore or

auxiliary copy operation is initiated and the electronic data to be restored or copied is identified, step 220.

The system locates the media on which the electronic data to be restored or copied is stored, step 225. Thus, index data stored at the storage manager 100, a media agent 105, or another location is consulted to determine the media ID where the archive file of the electronic data starts. In some embodiments, the system consults a slot map table contained in the index cache 110 to determine the media ID where the archive file of the electronic data starts.

The system determines the library/media agent/drive pool combination for the source tape storing the electronic data being restored or copied, step 230. Alternatively, the system determines the library/media agent/magnetic storage media volume combination for the electronic data being restored or copied. As previously described, in some embodiments the system determines the media agent 105 according to user preferences, selection logic for increasing network efficiency, or other factors. For example, if the electronic data to be restored or copied is located on a particular tape 140 in a given storage device 115, there may be a finite set of media agents 105 that are associated with the storage device 115 due to network pathways, component failures, user preferences, or other reasons.

The system establishes a network connection to the selected media agent 105 and other components, step 235. In some embodiments, the system establishes a high speed connection, such as a DataPipe™ connection manufactured by CommVault Systems, and further described in U.S. Patent No. 6,418,478 which is hereby incorporated herein by reference in its entirety. Electronic data is thus transferred over the established connection from the storage device 115 to its intended

destination such as a network client 85, an information store 90, or other network component, step 240.

In some embodiments, the entire archive file is generally created by a single attempt of backup and is contained within and accessible to the same media agent 105. Thus, a media agent 105 is identified only when initially opening the archive file. When media spanning occurs in the middle of the archive file, however, such as in the case of a file spanning multiple tapes 140 or volumes, the subsequent pieces of media containing the remaining electronic data must be located and requested, step 245. In the case of media spanning, control thus returns to step 225 to locate the additional media and continue the storage operation.

In some embodiments, the initially selected media agent 105 is first queried to determine whether it is associated with the additional media, and if so, to continue the storage operation. If the initially selected media agent 105 is not associated with the additional media, the system queries the other media agents 105 in the network and attempts to locate the additional media. For example, the system may search index caches 110 associated with the system media agents 105 to locate the additional media. Alternatively, if the system cannot locate the additional media, the user is prompted to manually import or otherwise make available the media holding the additional electronic data required to complete the storage operation.

Fig. 7 presents a flow chart of a method to restore a storage index in a system to perform storage operations on electronic data in a computer network according to an embodiment of the invention. In larger storage networks, scalability problems sometimes occasion the loss of index cache 110 directories from media agents 105. For example, in a network with many media agents 105 and a great deal of storage operations being performed, media agents 105 may load and unload

different index caches 110 depending on the electronic data subject to the storage operations being performed. Such loss of an index cache 110 directory from the memory of a media agent 105 requires that the index cache 110 directory be reloaded from stored media such as tapes 140 in a storage device. Reloading an index in this manner, however, often places strain on the network drives, and also results in high system resource usage through the instantiation of related processes and threads such as, for example, ifind, createindex, pipelines, and other actions. For example, when a media agent 105 performs a storage operation, the media agent index cache 110 is updated using a createindex process to indicate any new network pathways, changes to the files system, version information, and other information useful in performing storage operations. The index cache 110 directory must first be restored before the createindex process can be run to reflect these changes. The present invention, thus provides, in some embodiments, a method for efficiently restoring an index cache 110 directory.

The system determines whether there is an index in the cache 110 of a selected media agent, step 255. The system may query media agents 105 directly to determine if there is a copy of the index in a local cache 110. Alternatively, the system may look for the index in other locations, such as in a shared index system as further described herein.

If the index is not in the cache 110, then the system determines the media ID of the archive file containing the index by consulting reference tables stored in the media agent 105 or the storage manager 100 index cache 110, step 260. These tables identify the location of the index file and identify the actual storage media holding the index file such as a particular tape 140, volume 155, or other storage media. In the case of a tape 140 or other similar removable storage media, the system

reserves a drive for accessing the media, step 265, and performs a storage operation to restore the index to the local cache 110 of the selected media agent 105, step 270. In some embodiments, such as in the case of multiple media agents 105 having access to the index, preference is given to a LAN-free or otherwise preferred media agent 105. This is advantageous in the case of backups since the createindex has a high chance of being on the same machine as the tail of the pipeline.

Once the system determines that the index is in the local cache 110 of the selected media agent 105, the index is updated or otherwise accessed, such as through a createindex process, to perform the storage operation at hand, step 275. The storage group table is updated to reflect the storage operations performed, step 280.

In some embodiments, the system employs a shared index cache 110 location. A shared index might be employed, for example, to make the index available to all media agents 105 that may need to participate in a storage operation. Multiple media agents 105, for example, might be candidates for load balancing or failover operations, and all need access to the network path of the index so that the index will be available to each of them. Otherwise, extra index restores would be required when different media agents 105 were used for subsequent storage operations. Thus, in some embodiments, the index cache location is indicated by a UNC path universally accessible via a username and password.

Each media agent 105 is also associated with a unique username and password that permit component authentication, access control, and other similar functions. The username, password, and UNC path to the shared index location is stored in memory associated with each media agent 105. In some embodiments, an indexing module of the media agent 105 employs user impersonation before accessing the index cache.

In an alternate embodiment, the system employs a shared index cache 110 in which a pool of network UNC paths is designated for each media agent 105 as a secondary storage area for the index. At the end of each backup, a media agent 105 copies the index to this secondary area which is accessible to all media agents 105 in the pool. Thus, when a media agent 105 requires the index, the media agent 105 queries both its local cache 110 and the pool of UNC paths to locate the correct index. If the index is located in the secondary area and not in the local cache 105, then the index is copied to the local area for immediate use. Upon completion of a storage operation, such as a backup or a restore, the index is copied back to the shared area so that other media agents 105 and processes can access the index.

In some embodiments, the system determines which media agent 105, for a given browse of a client 85 at a point in time, is most likely to already have a useable index version already in its local cache 110. For example, in some embodiments, the storage manager 100 tracks which media agent 105 performed the last storage operation for a particular client 85. A media agent 105 is selected for a client 85 browse at time T such that the last backup in the full backup cycle at time $\geq T$ was done with indexing at that media agent 105.

Fig. 8 presents a flow diagram showing how dynamic volume sharing is accomplished according to one embodiment of the invention. A client application or other application initiates a request to the storage manager 100 to perform a storage operation on electronic data on a storage device 115, such as a magnetic media storage device 150, in the network, and the storage manager 100 processes this request by requesting access to the volume on which the data is storage, step 285. When a client computer 85 is configured, client data that is to be subject to storage operations is associated with a particular media agent 115. When that client data is

stored or retrieved in the future, the client computer 85 passes storage operation requests on to the associated media agent 115. The media agent 115 associates this client data with a particular storage media, such as a volume on a magnetic media storage device 150. Using dynamic volume sharing, one or more media agents can store and retrieve data among multiple volumes spanning multiple magnetic media storage devices 150.

When the media sharing request is received, the storage manager 100 verifies that a storage device 115 is available that can be switched to accommodate the request, step 290. The storage manager 100 tracks storage device 115 availability in the storage manager index cache 110 populated with information regarding available storage devices 115 and their respective media agent 105 controllers. Access paths across the network to media agents 105 and then on to appurtenant storage devices 115 are also stored in the storage manager index cache 110.

Upon identifying an appropriate storage device 115, the storage manager 100 directs the media agent 105 controlling the storage device 115 to go into a deactivated state with respect to that storage device, step 295. Even though, in some embodiments, there are multiple media agents executing on various hosts for the same storage device 115, the relationship is static and only one of them can control a storage device 115 at a given instant. The other media agents 105 are said to be in a deactivated state with respect to that storage device 115.

The deactivated media agents 105 run a listening process waiting for a message from the storage manager 100 directing them to become active with respect to a storage device 115. Once the first media agent 105 has been deactivated with respect to the storage device 115, the storage manager communicates to the listening process of a second media agent 105 on which the storage device 115 will be mounted

to change from a deactivated state to an activated state with respect to the storage device 115, step 300. At this point the storage manger 100 also updates the storage manager cache 110 to reflect that control of the storage device 115 has been shifted from the first media agent 105 to the second media agent 105, and that the first media agent is now deactivated and that the second media agent is now activated with respect to that storage device, step 305.

The second media agent 105 communicates with the storage device 115 and executes procedures necessary to mount storage device 115 and any associated volumes to the second media agent, step 310. In some embodiments, the second media agent 105 mounts one or more of the volumes associated with the storage device 115, and volumes in the same storage device 115 not mounted by the second media agent 105 may be mounted or otherwise associated with other media agents 105. Once the mount is performed the storage device 150 and its associated volumes 150 are logically connected to the second media agent 105, and this access path is stored by the second media agent 105 in its index cache 110, step 315. The media agent 105 stores the access path to the storage device 115 in the media agent index cache 110 because a storage device 115 connected to multiple media agents 105 may have multiple access paths. Mounting the storage device 115 to the media agent 105 and the resultant access path produced is in large part related to the hardware configuration of the media agent 105. The media agent 105 is generally best-suited to store and delegate management of the access path to the storage device that it controls. In some alternate embodiments, the storage manager 100 stores and tracks the individual hardware configuration of all the network media agents 105 in the storage manager index cache 110 and then passes the resultant access paths to the network storage devices 115 on to the media agents 105 when necessary. In other

embodiments, media agent 105 hardware configurations and resultant access paths to the network storage devices 115 are stored in a shared index location as further described herein.

Once the media agent 105 has completed the mount of the storage device 115 (and any associated volumes) and stored the access path to the storage device 115 in its own media agent index cache 110 or other location, then the access path to the storage device 115 is returned by the media agent 105 to the storage manager 100 where it is also stored in the storage manager index cache 110 for future recall, step 320. While media agents 115 generally communicate with storage devices 115 and the storage manager 100, the storage manager 100 generally communicates with client applications. In some embodiments, the storage manager 100 returns the storage device access path to a client application or other application and initiates a storage operation as appropriate, step 325.

Systems and modules described herein may comprise software, firmware, hardware, or any combination(s) of software, firmware, or hardware suitable for the purposes described herein. Software and other modules may reside on servers, workstations, personal computers, computerized tablets, PDAs, and other devices suitable for the purposes described herein. Software and other modules may be accessible via local memory, via a network, via a browser or other application in an ASP context, or via other means suitable for the purposes described herein. Data structures described herein may comprise computer files, variables, programming arrays, programming structures, or any electronic information storage schemes or methods, or any combinations thereof, suitable for the purposes described herein. User interface elements described herein may comprise elements from graphical user interfaces, command line interfaces, and other interfaces suitable for the purposes

described herein. Screenshots presented and described herein can be displayed differently as known in the art to input, access, change, manipulate, modify, alter, and work with information.

While the invention has been described and illustrated in connection with preferred embodiments, many variations and modifications as will be evident to those skilled in this art may be made without departing from the spirit and scope of the invention, and the invention is thus not to be limited to the precise details of methodology or construction set forth above as such variations and modification are intended to be included within the scope of the invention.

WHAT IS CLAIMED IS:

1. A computerized method for dynamically sharing a magnetic media volume in a network, the method comprising:

removing, in response to a volume sharing request, an association between a first media management component and the magnetic media volume; and associating, according to a set of selection logic, a second media management component with the magnetic media volume.

2. The method of claim 1, wherein removing in response to a volume sharing request comprises removing in response to a user preference.

3. The method of claim 1, wherein removing in response to a volume sharing request comprises removing according to a storage policy.

4. The method of claim 1, wherein removing in response to a volume sharing request comprises removing in response to a request initiated by a storage manager component.

5. The method of claim 1, wherein removing in response to a volume sharing request comprises removing in response to a request initiated by a client component.

6. The method of claim 1, wherein removing an association between a first media management component and the magnetic media volume comprises removing an association between a media agent and the magnetic media volume.

7. The method of claim 1, wherein removing an association between a first media management component and the magnetic media volume comprises removing an index entry association between the first media management component and the magnetic media volume.

8. The method of claim 1, wherein associating according to a set of selection logic comprises associating according to a user preference.

9. The method of claim 1, wherein associating according to a set of selection logic comprises associating according to a storage policy.

10. The method of claim 1, wherein associating according to a set of selection logic comprises associating according to a request initiated by a storage manager component.

11. The method of claim 1, wherein associating a second media management component and the magnetic media volume comprises associating a media agent and the magnetic media volume.

12. The method of claim 1, wherein associating a second media management component and the magnetic media volume comprises creating an index entry association between the second media management component and the magnetic media volume.

13. A system for dynamically sharing a magnetic media volume in a network, the system comprising:

a first media management component communicatively coupled to a magnetic media volume and a storage manager component; and

a second media management component communicatively coupled to the magnetic media volume and the storage manager component;

wherein the storage manager component is programmed to:

remove, in response to a volume sharing request, an association between the first media management component and the magnetic media volume; and

associate, according to a set of selection logic, the second media management component with the magnetic media volume.

14. The system of claim 13, wherein the storage manager component is programmed to remove in response to a user preference.

15. The system of claim 13, wherein the storage manager component is programmed to remove according to a storage policy.

16. The system of claim 13, wherein the storage manager component is programmed to remove in response to a request initiated by a client component.

17. The system of claim 13, wherein the first media management component comprises a media agent.

18. The system of claim 13, wherein the association between the first media management component and the magnetic media volume comprises an index entry.

19. The system of claim 13, wherein the set of selection logic comprises a user preference.

20. The system of claim 13, wherein the set of selection logic comprises a storage policy.

21. The system of claim 13, wherein the second media management a media agent.

22. The system of claim 13, wherein the association between the second media management component and the magnetic media volume comprises an index entry.

23. A computerized method for dynamically sharing a magnetic media volume in a network, the method comprising:

directing, in response to a volume sharing request, a first media management component controlling a magnetic media volume to go into a deactivated state with respect to the magnetic media volume;

directing, according to a set of selection logic, a second media management component to go into an activated state with respect to the magnetic media volume; and

mounting the magnetic media volume to the second media management component.

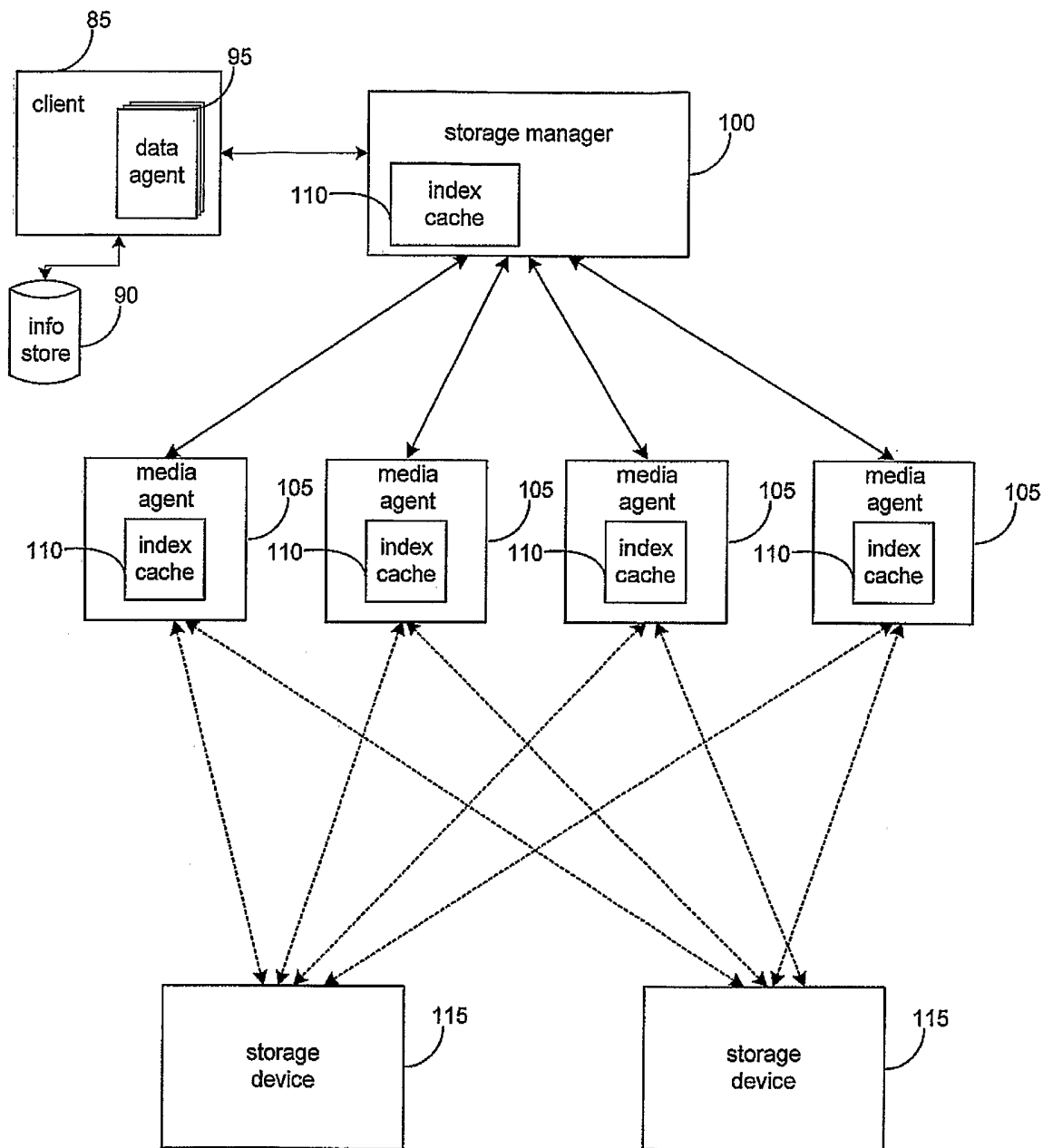


Fig. 1

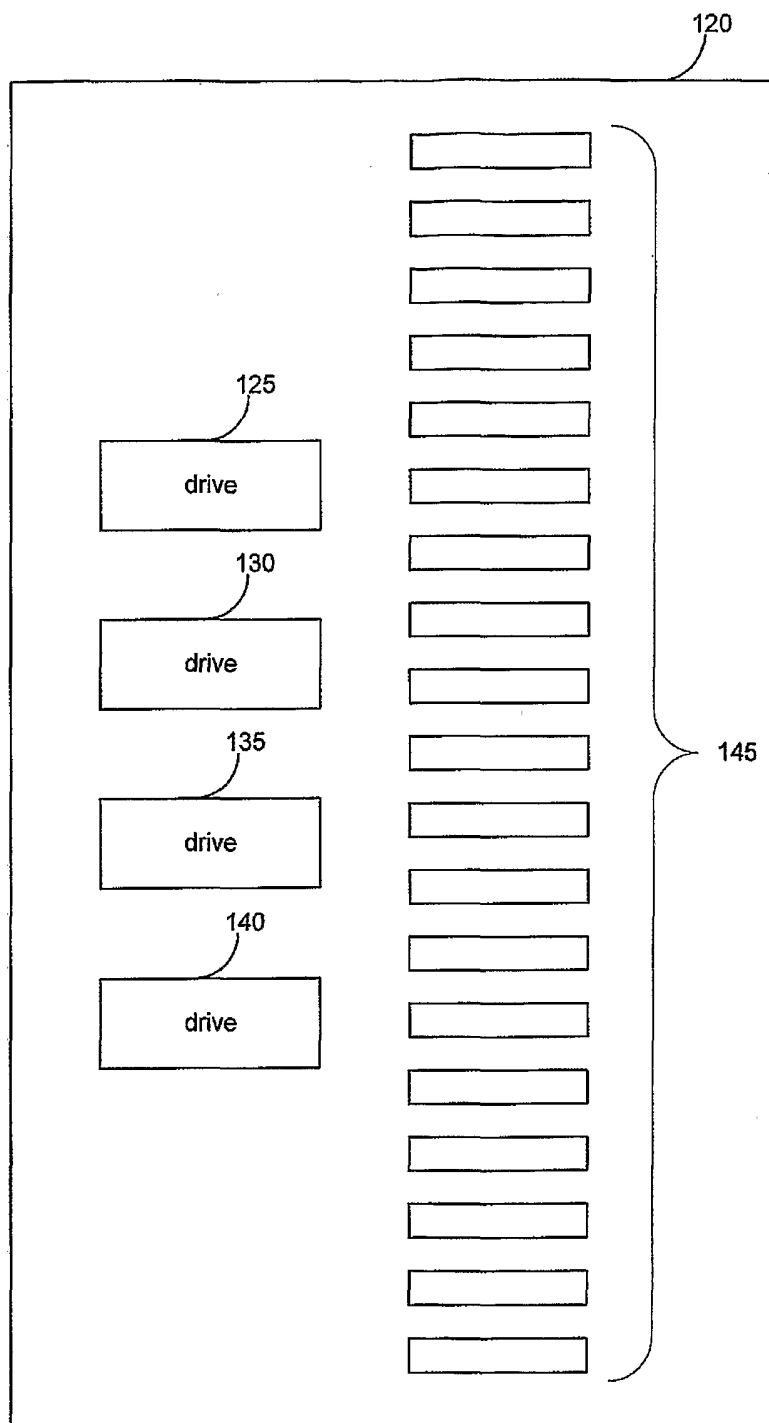


Fig. 2

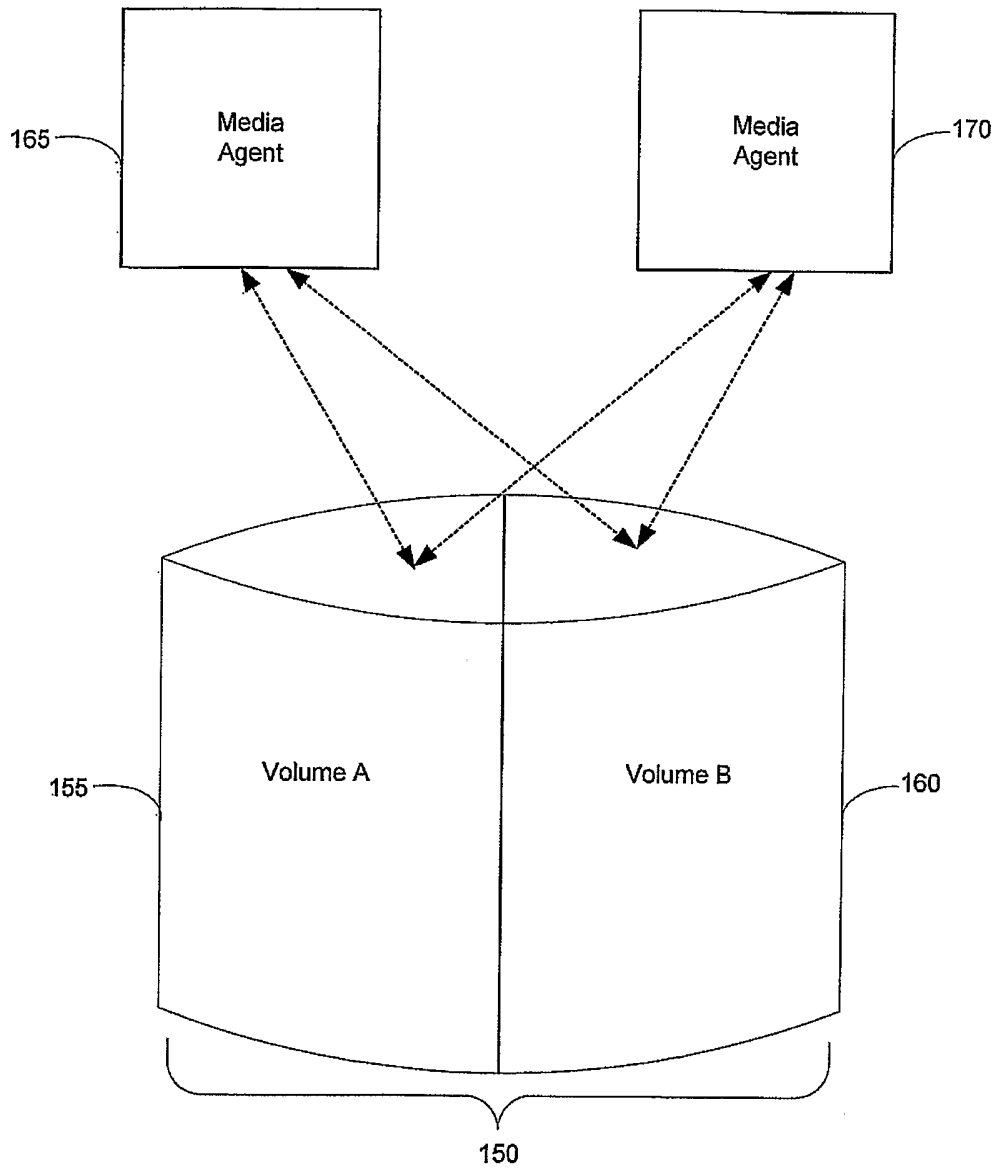


Fig. 3

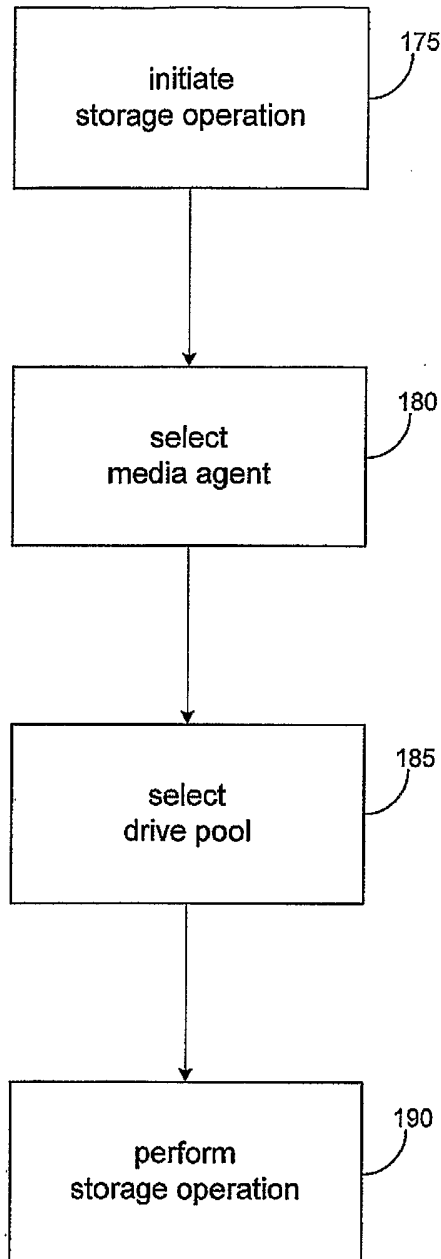


Fig. 4

5 / 8

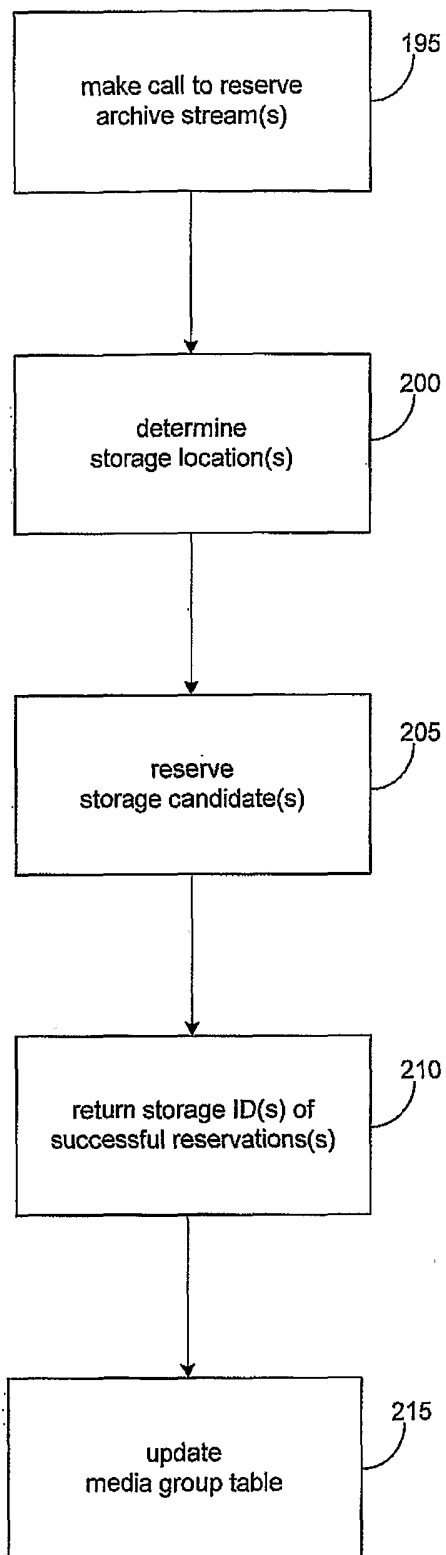


Fig. 5

6 / 8

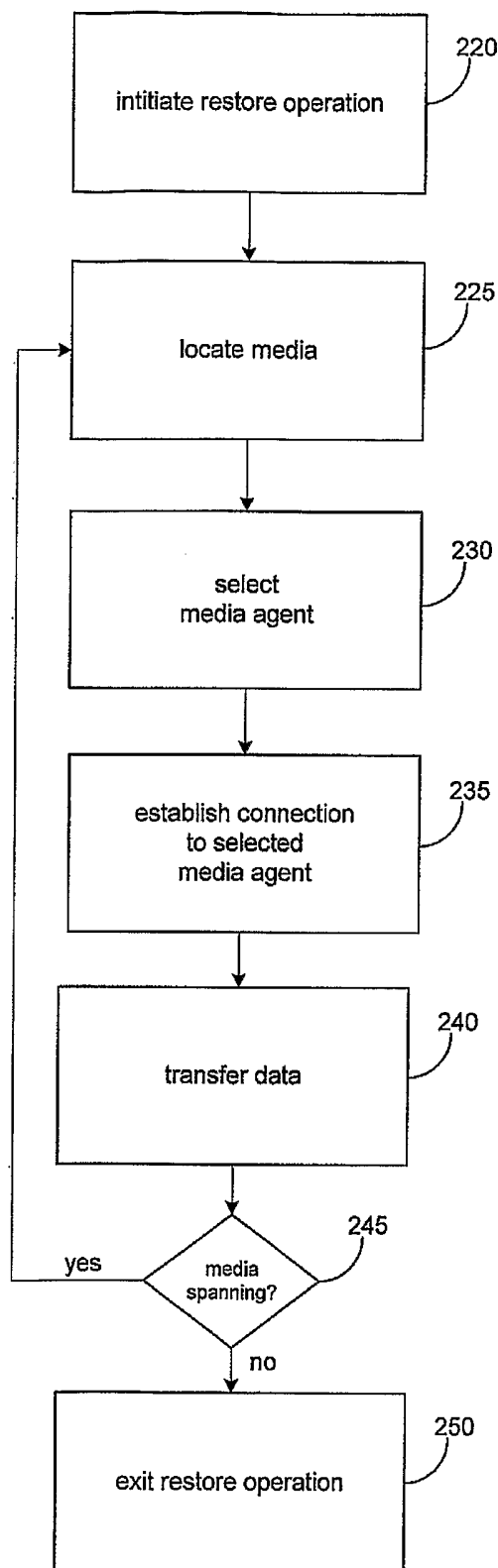


Fig. 6

7 / 8

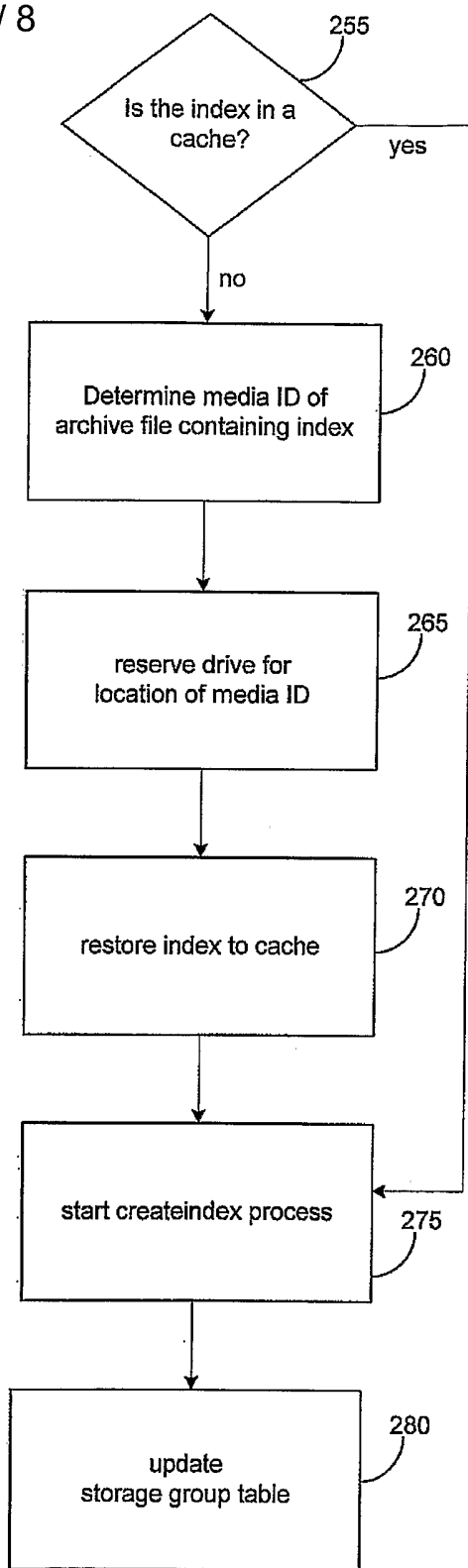


Fig. 7

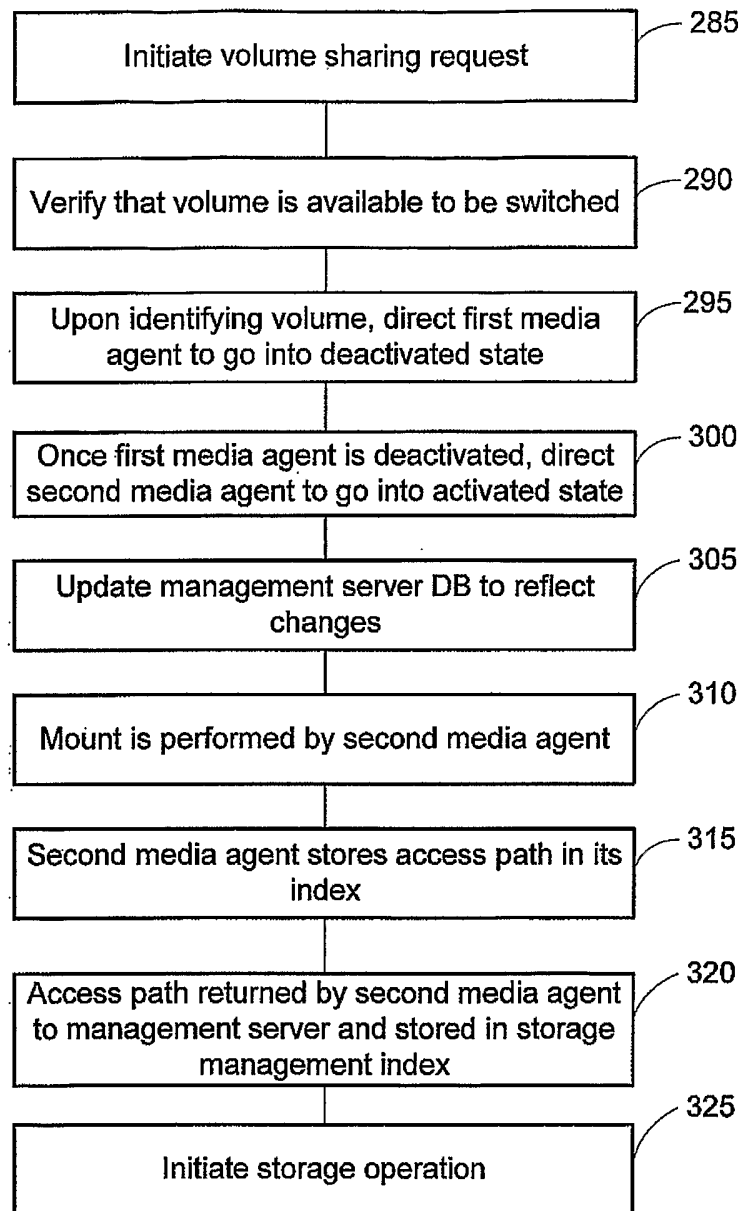


Fig. 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US04/10503

A. CLASSIFICATION OF SUBJECT MATTER
 IPC(7) : G06F 15/173
 US CL : 709/223, 226; 711/111
 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)
 U.S. : 709/223, 226; 711/111

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2002/0004883 A1 (NGUYEN et al) 10 January 2002 (10.01.2002), abstract, page 2 paragraph 14 and 22, pages 3-4 paragraphs 25, 29 and 35, and page 5 claim 4.	1-23
A	US 6,154,787 A (UREVIG et al) 28 November 2000 (28.11.2000), see entire document.	1-23

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 07 September 2004 (07.09.2004)	Date of mailing of the international search report 20 SEP 2004
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703)305-3230	Authorized officer <i>for Michelle L. Egan</i> Hosain T. Alam Telephone No. (703) 305-3900

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US04/10503

Continuation of B. FIELDS SEARCHED Item 3:

EAST

dynamic, automatic, sharing, volume, disk, tape, storage device, storage manager, controller...