

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
8 November 2007 (08.11.2007)

PCT

(10) International Publication Number  
**WO 2007/127360 A2**

(51) International Patent Classification: **Not classified**

(21) International Application Number:  
PCT/US2007/010222

(22) International Filing Date: 26 April 2007 (26.04.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
11/414,600 28 April 2006 (28.04.2006) US

(71) Applicant (for all designated States except US): **NETWORK APPLIANCE, INC.** [US/US]; 495 East Java Drive, Sunnyvale, CA 94089 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ZHENG, Ling** [CN/US]; 495 East Java Drive, Sunnyvale, CA 94089 (US). **STAGER, Roger** [US/US]; 495 East Java Drive, Sunnyvale, CA 94089 (US). **JOHNSTON, Craig** [US/US]; 495 East Java Drive, Sunnyvale, CA 94089 (US). **TRIMMER,**

**Don** [US/US]; 495 East Java Drive, Sunnyvale, CA 94089 (US). **FRANDZEL, Yuval** [US/US]; 495 East Java Drive, Sunnyvale, CA 94089 (US).

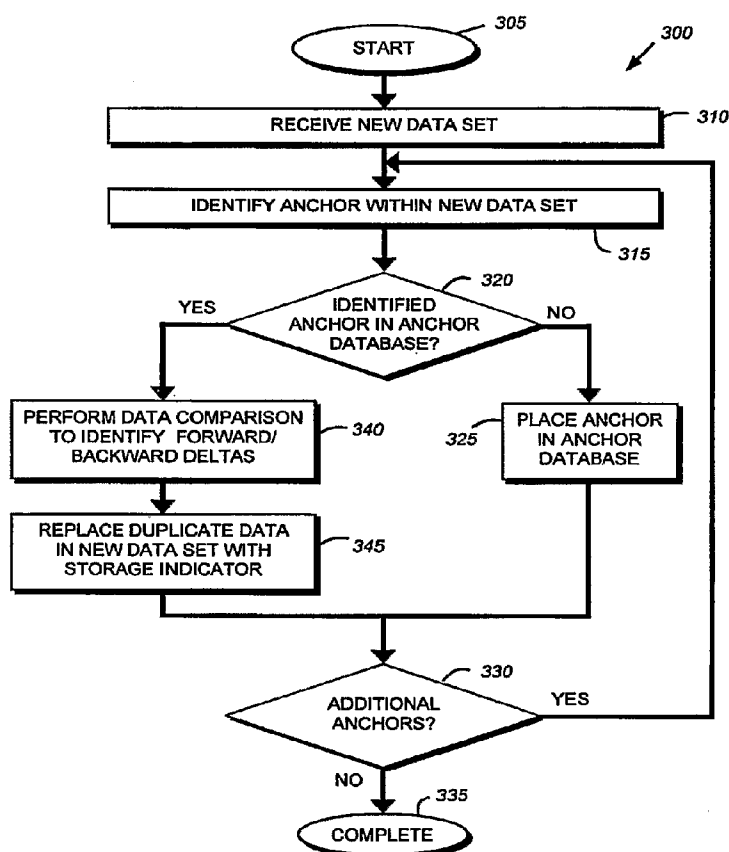
(74) Agent: **BARBAS, Charles, J.**; Cesari And McKenna, LLP, 88 Black Falcon Avenue, Boston, MA 02210 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, 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, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), 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, IS, IT, LT, LU, LV, MC, MT, NL, PL,

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR SAMPLING BASED ELIMINATION OF DUPLICATE DATA



(57) Abstract: A technique for eliminating duplicate data is provided. Upon receipt of a new data set, one or more anchor points are identified within the data set. A bit-by-bit data comparison is then performed of the region surrounding the anchor point in the received data set with the region surrounding an anchor point stored within a pattern database to identify forward/backward delta values. The duplicate data identified by the anchor point, forward and backward delta values is then replaced in the received data set with a storage indicator.



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

*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.*

**Published:**

— *without international search report and to be republished upon receipt of that report*

## SYSTEM AND METHOD FOR SAMPLING BASED ELIMINATION OF DUPLICATE DATA

### FIELD OF THE INVENTION

The present invention relates generally to data storage compression and, more  
5 specifically, sampling based elimination of duplicate data (de-duplication).

### BACKGROUND OF THE INVENTION

A storage system typically comprises one or more storage devices into which in-  
formation may be entered, and from which information may be obtained, as desired. The  
storage system includes a storage operating system that functionally organizes the system  
10 by, *inter alia*, invoking storage operations in support of a storage service implemented by  
the system. The storage system may be implemented in accordance with a variety of  
storage architectures including, but not limited to, a network-attached storage environ-  
ment, a storage area network and a disk assembly directly attached to a client or host  
computer. The storage devices are typically disk drives organized as a disk array,  
15 wherein the term "disk" commonly describes a self-contained rotating magnetic media  
storage device. The term disk in this context is synonymous with hard disk drive (HDD)  
or direct access storage device (DASD).

Data storage is a central part of many industries that operate in archival and com-  
pliance application environments, such as banks, government facilities/contractors and  
20 securities brokerages. In many of these environments, it is necessary to store selected  
data, e.g., electronic-mail messages, financial documents and/or transaction records, in an  
immutable manner, possibly for long periods of time. Typically, data backup operations  
are performed to ensure the protection and restoration of such data in the event of a fail-  
ure. However, backup operations often result in the duplication of data on backup stor-  
25 age resources, such as disks, causing inefficient consumption of the storage space on the  
resources. One form of long term archival storage is the storage of data on electronic  
tape media. A noted disadvantage of physical tape media is the slow data access rate and

the added requirements for managing a large number of physical tapes. In response to these noted disadvantages, several storage system vendors provide virtual tape library (VTL) systems that emulate tape storage devices using a plurality of the disk drives. In typical VTL environments, the storage system serving as the primary storage performs a complete backup operation of the storage system's file system (or other data store) to the VTL system. Multiple complete backups may occur over time thereby resulting in an inefficient consumption of storage space on the VTL system. It is thus desirable to eliminate duplicate data on the storage resources, such as disks associated with a VTL system, and ensure the storage of only single instance of data to thereby achieved storage compression.

One technique for achieving a reduction in data duplication (de-duplication) is described in United States Patent No. 5,990,810, entitled METHOD FOR PARTITIONING A BLOCK OF DATA INTO BLOCKS AND FOR STORING AND COMMUNICATING SUCH SUBBLOCKS, by Ross Williams, issued November 23, 1999 (hereafter "the '810 patent"). The method described in the '810 patent first utilizes a rolling hash function to generate a plurality of sub-blocks of data. The rolling hash utilizes a fixed size window of data that results in a boundary being placed between two sub-blocks. Once a block of data has been partitioned into sub-blocks, the hash value of each sub-block is calculated to form a table of hash values. The hash table is then used to determine if a new sub-block is identical to any sub-block whose hash value has previously been stored in the hash table. To perform this determination, the new sub-block's hash value is calculated and compared with the values contained in the hash table. If a the new sub block's hash value has been previously stored within the hash table, then the sub block identified with the stored hash value is considered as identical with the new sub block. In such a case, the new sub block is replaced with a pointer to the previously stored sub block, thereby reducing the amount of storage space required for the sub block. A noted disadvantage of the technique described in the '810 patent is that it requires performance of an extensive number of computationally intensive hashing calculations, which may affect the overall performance of a storage/VTL system implementing such a method. Another noted disadvantage is that the hash table will become larger as

the size of data set increases and may not scale to large data sets such as terabytes or petabytes of data.

Another technique for eliminating duplicate data is described in United States Patent Application Serial No. 11/105,895, filed on April 13, 2005 entitled METHOD  
5 AND APPARATUS FOR IDENTIFYING AND ELIMINATING DUPLICATE DATA  
BLOCKS AND SHARING DATA BLOCKS IN A STORAGE SYSTEM, by Ling  
Zheng et al. In the system described in this patent application, all data de-duplication operations are performed on fixed size blocks that are illustratively 4 kilobytes (KB) in size. When a new block is to be stored, a hash is computed of the 4 KB block and compared  
10 with a hash table containing hash values of previously stored blocks. Should the new block's hash value be identical to the previously stored block, there is a high degree of probability that the new block is identical to the previously stored block. In such a case, the new block is replaced with a pointer to the previously stored block, thereby reducing storage resource consumption. However, a noted disadvantage of this system is that it is  
15 limited to fixed size blocks. In operation a single bit addition in the middle of a data set may cause misalignment of all subsequent blocks and prevent de-duplication, thereby causing an inefficient use of storage resources.

A third technique, as mentioned in DELTA STORAGE FOR ARBITRARY  
NONTEXT FILES by Chris Reichenberger, In *Proceedings of the 3rd International Work-*  
20 *shop on Software Configuration Management*, Trondheim, Norway, 1214 June 1991 (June 1991), ACM, pp. 144--152, is to use the hashes of data to find the longest common data sequence. However, the above method to compute and compare the hashes of all possible substrings in a data set does not employ any sampling technique. Therefore, it will not be scalable for large data sets.

25

## SUMMARY OF THE INVENTION

The present invention overcomes the disadvantages of the prior art by providing a sampling based technique for eliminating duplicate data (de-duplication) stored on storage resources, such as disks. According to the invention, when a new data set, e.g., a backup data stream, is received by a server, e.g., a storage system or virtual tape library

(VTL) system implementing the invention, one or more anchors are identified within the new data set. As used herein, an anchor is defined as a point within the data set that defines a region of interest for potential data de-duplication. Anchors may be located by performing a rolling hash of the data set. In such cases, an anchor point ("anchor") is located within a fixed size region defining a signature width. Anchors may also be selected based on location within data containers, e.g., a file, or other contextual information, e.g., at predefined offsets within the data set.

For each anchor identified in the new data set, a determination is made whether the anchor has been stored previously in an anchor database (or other data store). If it has not been stored in the anchor database, the anchor is then stored within the anchor database along with information identifying its location within the data set. However, in alternate embodiments, the anchor location may simply be an offset within the data set. If the identified anchor was stored previously within the anchor database, then the server locates the anchor within the pattern database and performs a bit-by-bit (or byte-by-byte) data comparison, forwards and backward from the anchor point between the new data set and the data set containing the anchor stored in the pattern database. The result of this data comparison is a forward delta value and a backward delta value identifying the number of consecutive bits of data that match between the new data set and the stored data set forward and backward from the anchor point, respectively. In the new data set, the duplicate data, i.e., the range of data between the backward delta value and the forward delta value around the anchor point, is replaced with a storage indicator that identifies the anchor and the forward and backward delta values.

### BRIEF DESCRIPTION OF THE DRAWINGS

The above and further advantages of the invention may be better understood by referring to the following description in conjunction with the accompanying drawings in which like reference numerals indicate identical or functionally similar elements:

Fig. 1 is a schematic block diagram of an exemplary storage system environment including a virtual tape library (VTL) system in accordance with an embodiment of the present invention;

Fig. 2 is a schematic block diagram of a VTL system in accordance with an embodiment of the present invention;

Fig. 3 is a flowchart detailing the steps of a procedure for performing a sampling based data de-duplication procedure in accordance with an embodiment of the present invention;

Fig. 4 is a diagram showing the formation of backward and forward deltas around an anchor point in accordance with an embodiment of the present invention; and

Fig. 5 is a block diagram illustrating the replacement of duplicate data with a storage indicator in accordance with an embodiment of the present invention.

## **DETAILED DESCRIPTION OF AN ILLUSTRATIVE EMBODIMENT**

### **A. Storage System Environment**

Fig. 1 is a schematic block diagram of a storage system environment 100 that may be advantageously used with the present invention. The storage system environment 100 comprises a storage system 125 interconnected with a plurality of clients 110 by network 120. Illustratively, the network 120 may be embodied as an Ethernet network or a Fibre Channel (FC) network. Also interconnected with the network 120 is a virtual tape library (VTL) system 200 that illustratively implements the data de-duplication technique of the present invention. Also operatively connected with the VTL system 200 are one or more storage devices, such as disks 130, organized as a storage array 160.

In operation, the storage system 125 services data access requests from the clients 110. Each client 110 may be a general-purpose computer configured to execute applications and interact with the storage system 125 in accordance with a client/server model of information delivery. That is, the client may request the services of the storage system, and the system may return the results of the services requested by the client, by exchanging packets over the network 120. The clients may issue packets including file-based access protocols, such as the Common Internet File System (CIFS) protocol or Network File System (NFS) protocol, over TCP/IP when accessing information, such as data, in the form of data containers, such as files and directories. Alternatively, the client may issue packets including block-based access protocols, such as the Small Computer Sys-

tems Interface (SCSI) protocol encapsulated over TCP (iSCSI) and SCSI encapsulated over Fibre Channel (FCP), when accessing information in the form of data containers, such as blocks.

When an administrator of the storage system 125 desires to perform a backup operation, conventional tape backup software may be utilized with the storage system 125 to perform a backup operation to the VTL system 200. The VTL system 200 appears to the storage system 125 as a remote tape drive; thus, the storage system 125 may perform a conventional tape backup operation to the VTL system 200 using conventional tape backup software. Typically, each backup operation results in the copying of the entire file system stored by the storage system 125 to the VTL system 200, e.g., as a "backup". A noted result of such copying is that the VTL system 200 may include significant amounts of duplicated data, i.e., data that remains common among each of the backups of storage system 125 and/or redundant data between backups of different clients 110. The present invention illustratively eliminates the duplicate data in accordance with the novel sampling based data de-duplication technique, described further below.

#### B. Virtual Tape Library System

Fig. 2 is a schematic block diagram of a VTL system 200 that may be advantageously used with the present invention. Illustratively, the VTL system 200 is a computer that provides storage service relating to the organization of information on storage devices, such as disks 130 of a local storage (disk) array 160. The VTL system 200 illustratively comprises a processor 222, a memory 224, one or more network adapters 226 and one or more storage adapters 228 interconnected by a system bus 225. Each network adapter 226 comprises the mechanical, electrical and signaling circuitry needed to connect the VTL system 200 to storage system 125 over the network 120. The VTL system 200 also includes an operating system 250 that interacts with various modules, such as virtual tape library (VTL) module 260, described further below. The operating system 250 may implement a logical data object store on disks 130 and for use by applications, such as VTL module 260 executing on the VTL system 200. As described herein, the logical data object store is illustratively embodied as an anchor database (DB) 270 and a pattern DB 280.



In the illustrative embodiment, the memory 224 comprises storage locations that are addressable by the processor and adapters for storing software program code, such as operating system 250 and VTL module 260. The processor and adapters may, in turn, comprise processing elements and/or logic circuitry configured to execute the software code and manipulate the data structures. It will be apparent to those skilled in the art that other processing and memory means, including various computer readable media, may be used for storing and executing program instructions pertaining to the inventive technique described herein.

The storage adapter 228 cooperates with the operating system 250 executing on the VTL system 200 to store and retrieve information requested on any type of attached array of writable storage device media such as video tape, optical, DVD, magnetic tape, bubble memory, electronic random access memory, micro-electro mechanical and any other similar media adapted to store information, including data and parity information. However, as illustratively described herein, the information is preferably stored on the disks 130, such as HDD and/or DASD, of local storage array 160. The storage adapter includes input/output (I/O) interface circuitry that couples to the disks over an I/O interconnect arrangement, such as a conventional high-performance, FC serial link topology.

As noted, the operating system 250 illustratively implements a logical data object store on disks 130. It should be also noted that in alternate embodiments, other forms of data organization may be utilized including, for example, a file system. As such, the description of utilizing a data object store should be taken as exemplary only. Furthermore, the VTL module 260 cooperates with the operating system 250 to implement virtual tape library functionality for the VTL system 200. Illustratively, the VTL module 260 stores data from the using the data object store. Notably, the VTL module 260 also implements the data de-duplication technique of the present invention.

### C. Sampling Based Data De-Duplication

The present invention provides a sampling based technique for eliminating duplicate data (de-duplication) stored on storage resources, such as disks 130. According to the invention, when a new data set is received by a server, e.g., storage system 125 or system 200, one or more anchors are identified within the new data set. As used herein,

an anchor is defined as a point within the data set that defines a region of interest for potential data de-duplication. Anchors may be located by performing a rolling hash of the data set. In such cases, an anchor point ("anchor") is located within a fixed size region defining a signature width. Anchors may also be selected based on location within a data container, e.g., a file, or other contextual information, e.g., at predefined offsets within the data set. It should be noted that anchors may be identified using a variety of techniques. As such, the techniques described herein should be taken as exemplary only.

For each anchor identified in the new data set, a determination is made whether the anchor has been stored previously in the anchor database 270. If it has not been stored in the anchor database, the anchor is then stored within the anchor database 270 along with information identifying its location within the data set. Illustratively, data sets are stored within the pattern database 250 disposed over implemented on top of, e.g., the data object store. However, in alternate embodiments, the anchor location may simply be an offset within the data set. If the identified anchor was stored previously within the anchor database, then the server locates the anchor within the pattern database 250 and performs a bit-by-bit data comparison, forward and backward from the anchor point, between the new data set and the data set containing the anchor stored in the pattern database 250. The result of this data comparison is a forward delta value and a backward delta value identifying the number of consecutive bits of data that match between the new data set and the stored data set forward and backward from the anchor point, respectfully. In the new data set, the duplicate data, i.e., the range of data between the backward delta value and the forward delta value around the anchor point, is replaced with a storage indicator that identifies the anchor and the forward and backward delta values.

Fig. 3 is a flowchart detailing the steps of a sampling based data de-duplication procedure 300 in accordance with an illustrative embodiment of the present invention. The procedure 300 begins in step 305 and continues to step 310 where a new data set is received by a server, such as the VTL system 200. In the illustrative environment 100 of Fig. 1, the received data set may comprise a new tape backup data stream directed to the VTL system 200 from the storage system 125. However, in alternate embodiments, any data set may be utilized on any computer including, for example, a file stored within the storage system 125. As such, the description of the data set being a tape backup stream

directed to a VTL system 200 should be taken as exemplary only. As noted, in alternate embodiments, any form of data set may be utilized on any computer. Thus, the teachings of the present invention may be utilized in environments different from the below-described VTL system environment.

5           In response to receiving the new data set, the VTL module 260 identifies one or more anchors within the new data set in step 315. In accordance with the present invention, anchors may be identified in a plurality of ways. One anchor identification technique is to place an anchor at predefined offsets within the data set. These offsets may be fixed or may be based on the type of data within the data set. For example, if the data set  
10 is a tape archival backup data stream, an anchor may be placed at the beginning of each record within the data stream. Another exemplary technique for identifying anchors is to perform a rolling hash, using a fixed size window, along the data set. An anchor may then be placed within the window when the hash value (modulo a constant) equals a predetermined value. For example, a rolling hash using a window size of 512 bits may be  
15 utilized. When the hash value, modulo N, equals 0, an anchor point may be located at the beginning of the window.

Once the anchors have been identified, the VTL module 260 then determines whether the identified anchor is located within the anchor database 270 in step 320. If the identified anchor is not stored within the anchor database, the procedure 300 branches  
20 to step 325 where the VTL module places the anchor within the anchor database. Certain anchors may not be stored in the anchor database, e.g., anchors that occur too frequently to be of use. The VTL module may be configured with these frequently occurring anchors or may, in alternate embodiments, track frequently occurring anchors to identify those that should not be maintained in the anchor database. Then, in step 330, a determination is made whether additional anchors are within the data set. If so, the procedure  
25 300 loops back to step 315 where the VTL module 260 identifies another anchor within the data set. Otherwise, the procedure 300 completes in step 335.

However, if the identified anchor is located within the anchor database 270, the VTL module 260 then begins a bit-by-bit (or byte-by-byte) comparison of the region surrounding the anchor within the pattern database 280 and the anchor point within the re-  
30

ceived data set to identify forward and backward delta values (step 340). The delta values identify the number of conservative bits forward/backward from the anchor point that are identical between the region in the pattern database 280 and the received data set. For example, if the data is identical for a thousand bits in front of the anchor point and 15 bits behind the anchor point, then the backward delta value would be set equal to 15 and the forward delta value would be set equal to 1000. It should be noted that in certain situations, depending upon placement of anchors, a particular delta value may be equal to zero. For example, if the anchor point is selected at the beginning of the window utilized by the rolling hash, no bits prior to the anchor point may match with the pattern database. The backward/forward delta values ("deltas") typically encompass the width of the sliding hash window; however, this is not a requirement of the present invention. By utilizing the anchor point along with forward and backward deltas, a region of duplicate data may be identified i.e., the data that ranges from the backward delta (number of bits) behind the anchor point to the forward delta (number of bits) in front of the anchor point. According to the invention, this identified region of duplicate data may be eliminated from the storage, thereby achieving compression and improving the efficiency of the use of storage resources.

In step 345, the identified region of duplicate data in the new data set is then replaced with a storage indicator. Illustratively, the storage indicator identifies the anchor as well as the forward and backward deltas. Compression of the data set is obtained when the storage indicator is of a smaller size than the duplicate data being replaced. Thus, depending upon implementation details, a minimum size may be required before duplicate data is replaced. For example, if a storage indicator requires 256 bits, unless  $(\text{forward delta value} + \text{backward delta value}) > 256$ , no compression will occur. It should be noted that in alternate embodiments, the system may include an administrator controlled threshold for minimum size before replacement occurs. Such a threshold may enable the administrator to tune the performance of the system based on the nature of the expected data. Thus, the system may include both a static, e.g., size of storage indicator, and dynamic, e.g., administrator set, thresholds. Again, if there are no additional anchors in the data set, the procedure completes in step 335.

During step 340, the region of duplicate data being identified may have additional anchor points in this region. However, those anchor points would not need to be processed by step 315. The next anchor point to be processed will be in the new data after the identified region of duplicate data. Further, the skipped anchor points might not need to be calculated using the anchor identification technique. The set of new anchor points in the duplicate data region will be the same as the set in the original data region. The performance benefit is that relatively few anchor points will have to be calculated and processed when processing long streams of matching data.

Additionally, if duplicate data could be identified before writing the new data to the storage server, the new duplicate data may not need to be written to the VTL system 200 at all. This would have the benefit of reduced I/O operations over the network.

Fig. 4 is a schematic diagram of an environment 400 showing the generation of backward and forward deltas between a previously stored data set 405 and a new data set 410. In the prior data set 405, i.e., a data set previously received and processed by the VTL system 200, three anchors A1, A2, A3 have been identified. In the new data set 410, i.e., the data set currently being processed it is determined that a single anchor A2 exists. In response, the VTL system 200 begins a bit-by-bit data comparison of the region in front of and behind the anchor A2 in both the prior data set 405 and the new data set 410. Assume that data matches until a forward point 420 and to a backward point 415 on each data set. In such a case, the region between points 415, 420 is replaced with a storage indicator in the new data set 410.

Fig 5 is a schematic diagram of an environment 500 showing the replacement of data with a storage indicator in accordance with an embodiment of the present invention. Both prior data set 405 and new data set 410 are centered around anchor A2. Within the prior data set 405, the region around the anchor contains data having values N, C, D, E (the anchor point), F, G, H, O, etc. Similarly, the data surrounding the anchor in new data set 410 are A, B, C, D, E (the anchor point), F, G, H, I, etc. In accordance with the illustrative embodiment, once the anchor point is identified, a bit-by-bit comparison occurs forward and backward from the anchor point. A determination is made that only data F, G and H in front of the anchor point of each data set are identical, i.e., next data has the value of O in the prior data set 405 and a value of I in the new data set 410. Simi-

larly, it is determined that only data C and D in back of the anchor point of each data set are identical. These six areas of data (including the anchor E) within the new data set 410 (namely the region between C and H) are then replaced with a storage indicator 505. Accordingly, the new data set 410 is modified to new data set 410' comprising of data A, B, the storage indicator 505, I, J, K, L, M, etc.

Illustratively, the storage indicator 505 comprises an anchor identifier (ID) field 510, a forward delta field 515 and a backward delta field 520. The anchor ID field 510 identifies the anchor associated with this storage indicator 505. The forward and backward delta fields 515, 520 identify the number of consecutive bits on each side of the anchor that have been replaced. Thus, in this example, the forward delta value would identifies the number of bits contained within data F, G and H, while backward delta identifies the number of bits contained within data C and D. When reconstructing the original data (re-duplication), the system retrieves the data stored within the pattern database identified by the storage indicator and inserts it into the reconstructed data. For example, upon identifying that a storage indicator is present, the system locates the identified anchor and retrieves the data backward/forward from the anchor point identified by the delta values. That region is then re-substituted into the data set in place of the storage indicator.

Anchor database 270 is utilized during step 320 to identify the duplicate data between the prior data set 405 and the new data set 410. For new data set 410 being written to the VTL system 200, the VTL could have the knowledge of where duplicate data in the new data set 410 are likely to reside. For example, the new data for the current complete backup will have the most duplicate data with the data for the last complete backup. The regions of duplicate data will also appear in the same order in two complete backup. Therefore, to identify the duplicate data, the anchor points corresponding to the previous complete backup could be sequentially loaded into memory to compare with the anchors in the new data set 410. This allows for very efficient identifications of duplicate data regions with very small memory requirements.

In an alternate embodiment, an anchor hierarchy may be generated by computer super anchors. A super anchor is illustratively generated by computing a hash based on a

plurality of adjacent anchors. The number of super anchors would be smaller than the number of anchors, thereby increasing the speed at which super anchors may be identified. If a match is identified on a super anchor, the system has identified a region of duplicate data that spans multiple adjacent anchors.

5           Furthermore, while this description has been written in terms of storage de-duplication, the principles of the present invention may be utilized to provide de-duplication over communication links. In such an alternate embodiment, the sending system may perform the above described de-duplication technique to minimize the amount of bandwidth utilized in transmitting data. The source and destination systems will need  
10   to work to ensure that they share synchronized pattern and anchor databases.

          The foregoing description has been directed to specific embodiments of this invention. It will be apparent, however, that other variations and modifications may be made to the described embodiments, with the attainment of some or all of their advantages. For instance, it is expressly contemplated that the teachings of this invention can  
15   be implemented as software, including a computer-readable medium having program instructions executing on a computer, hardware, firmware, or a combination thereof. Furthermore, it should be noted that while this description has been written in terms of VTL system, the teachings of the present invention are not so limited. The present invention may be utilized in any computer environment using any data store, including, e.g., a file  
20   system. Accordingly this description is to be taken only by way of example and not to otherwise limit the scope of the invention. Therefore, it is the object of the appended claims to cover all such variations and modifications as come within the true spirit and scope of the invention.

What is claimed is:

**CLAIMS**

- 1 1. A method for removing duplicate data from a data set, the method comprising the  
2 steps of:  
3 identifying an anchor within the data set;  
4 determining whether the identified anchor exists within an anchor database;  
5 in response to determining that the anchor exists within the anchor database, per-  
6 forming a data comparison between the data set and a stored data set to identify a forward  
7 delta value and a backward delta value relative to the identified anchor; and  
8 replacing a region of the data set identified by the anchor, the forward delta value  
9 and the backward delta value with a storage indicator to form a modified data set.
- 1 2. The method of claim 1 wherein the step of identifying the anchor comprises the  
2 step of performing a rolling hash on the data set.
- 1 3. The method of claim 1 wherein the step of identifying the anchor comprises plac-  
2 ing the anchor at a predetermined location within the data set.
- 1 4. The method of claim 1 wherein the stored data set is stored in a pattern database.
- 1 5. The method of claim 1 wherein the storage indicator comprises an anchor identi-  
2 fier, the forward delta value and the backward delta value.
- 1 6. The method of claim 1 further comprising the step of storing the modified data set  
2 in a data object store.
- 1 7. The method of claim 6 wherein the data object store comprises a file system.
8. The method of claim 1 further comprising the step of finding duplicate data be-  
tween two similar data sets.



9. The method of claim 8 wherein the similar data sets comprise data sets with intermixed common data and unique data.

10 The method of claim 8 wherein duplicate data is found by comparing anchor points of the new data set with the anchor points of similar data.

1 11. The method of claim 10 wherein the anchor points of similar data are identified  
2 based on a pattern of anchor points within the new data set.

1 12. The method of claim 1 further comprising the step of forming an anchor hierarchy  
2 by computing a hash on a plurality of adjacent anchors within the data set.

1 13. A system configured to remove duplicate data from a data set, the system comprising:  
2  
3 means for identifying an anchor within the data set;  
4 means for determining whether the identified anchor exists within an anchor database;  
5  
6 in response to determining that the anchor exists within the anchor database,  
7 means for performing a data comparison between the data set and a stored data set to  
8 identify a forward delta value and a backward delta value relative to the identified anchor; and  
9  
10 means for replacing a region of the data set identified by the anchor, the forward  
11 delta value and the backward delta value with a storage indicator.

1 14. The system of claim 13 wherein the means for identifying the anchor comprises  
2 means for performing a rolling hash on the data set.

1 15. The system of claim 13 wherein the means for identifying the anchor comprises  
2 means for placing the anchor at a predetermined location within the data set.

1 16. The system of claim 13 wherein the stored data set is stored in a pattern database.

1 17. The system of claim 13 wherein the storage indicator for comprises an anchor  
2 identifier, the forward delta value and the backward delta value.

1 18. The system of claim 13 further comprising means for forming an anchor hierarchy  
2 by computing a hash on a plurality of adjacent anchors within the data set.

1 19. A system configured to remove duplicate data from a data set, the system com-  
2 prising:

3 a storage system configured to serve the data set; and

4 a virtual tape library system adapted to receive the data set from the storage sys-  
5 tem, the virtual tape library system adapted to identify an anchor within the data set and  
6 further adapted to determine whether the identified anchor exists within an anchor data-  
7 base.

1 20. The system of claim 19 wherein the virtual tape library system is further adapted  
2 to, in response to, determining that the anchor exists within the anchor database, perform  
3 a data comparison between the data set and a stored data set to identify a forward delta  
4 value and a backward delta value.

1 21. The system of claim 20 further comprising a pattern data base adopted to store the  
2 stored data set.

1 22. The system of claim 20 wherein the virtual tape library system is further adapted  
2 to replace a region of the data set identified by the anchor, the forward delta value and the  
3 backward delta value with a storage indicator to form a modified data set.

1 23. The system of claim 22 wherein the storage indicator comprises an anchor identi-  
2 fier, the forward delta value and the backward delta value.

- 1    24.    The system of claim 19 wherein the anchor is identified by performing a rolling  
2    hash on the data set.
- 1    25.    The system of claim 19 wherein the anchor is identified by placing the anchor at a  
2    predetermined location within the data set.
- 1    26.    The system of claim 19 wherein the data set comprises a backup data stream.
- 1    27.    The system of claim 19 wherein the virtual tape library system is further adapted  
2    to form an anchor hierarchy by computing a hash on a plurality of adjacent anchors  
3    within the data set.
- 1    28.    A computer readable medium for removing duplicate data from a data set, the  
2    computer readable medium including program instructions for performing the steps of:  
3        identifying an anchor within the data set;  
4        determining whether the identified anchor exists within an anchor database;  
5        in response to determining that the anchor exists within the anchor database, per-  
6    forming a data comparison between the data set and a stored data set to identify a forward  
7    delta value and a backward delta value relative to the identified anchor; and  
8        replacing a region of the data set identified by the anchor, the forward delta value  
9    and the backward delta value with a storage indicator to form a modified data set.

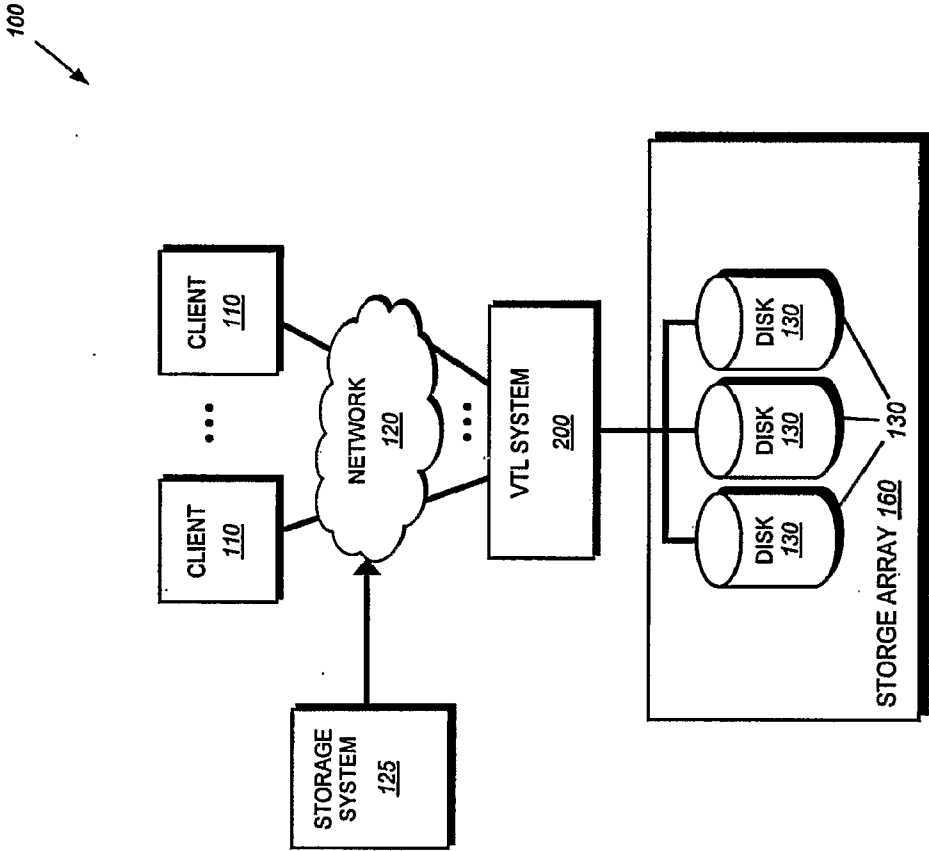


FIG. 1

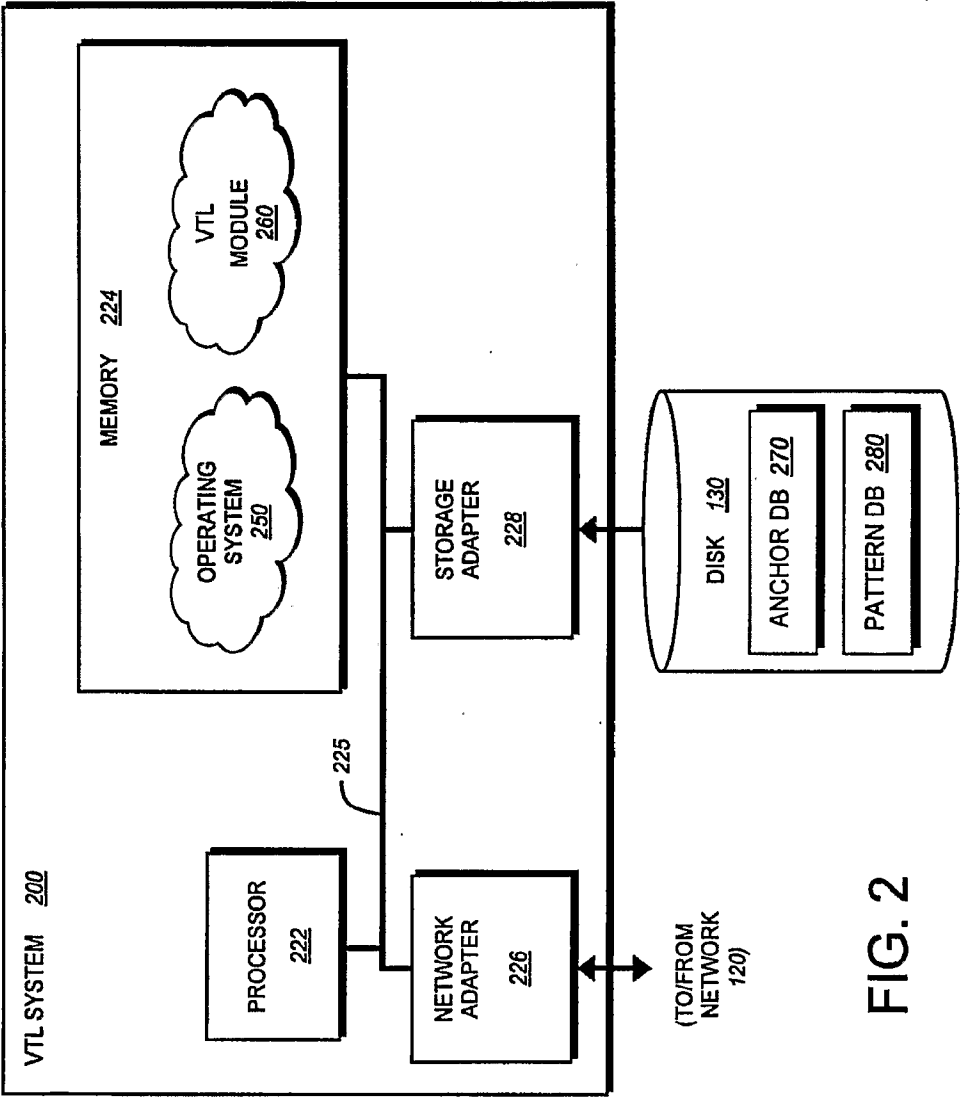


FIG. 2

3/5

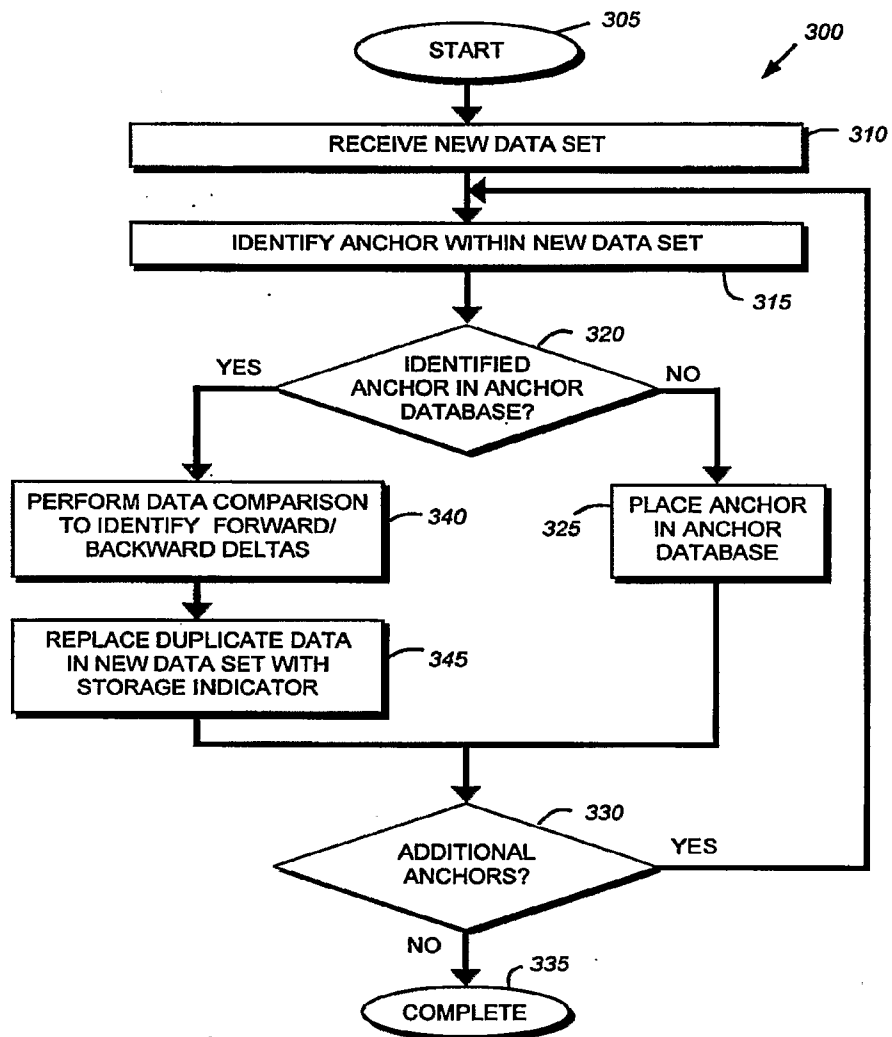


FIG. 3

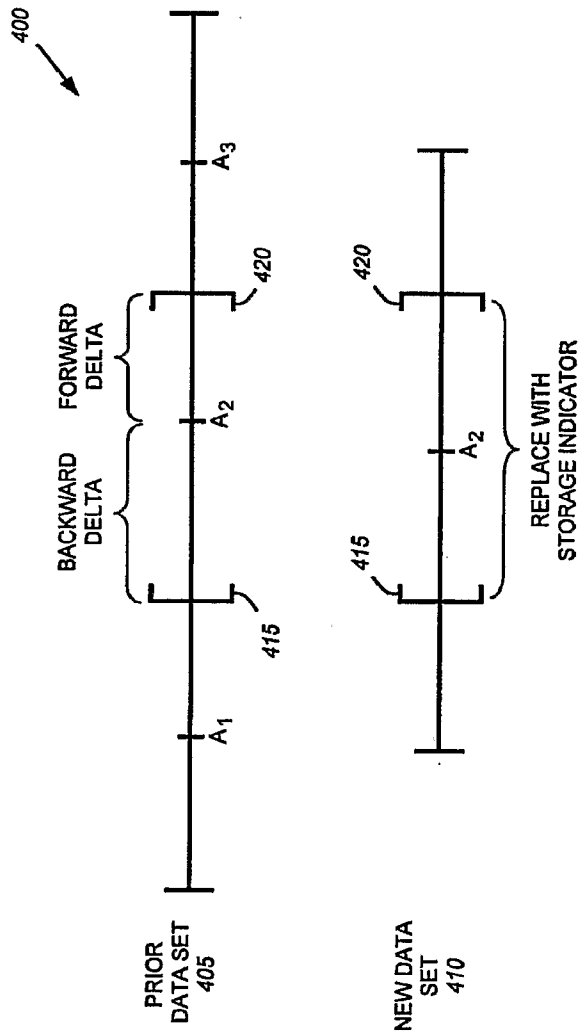


FIG. 4

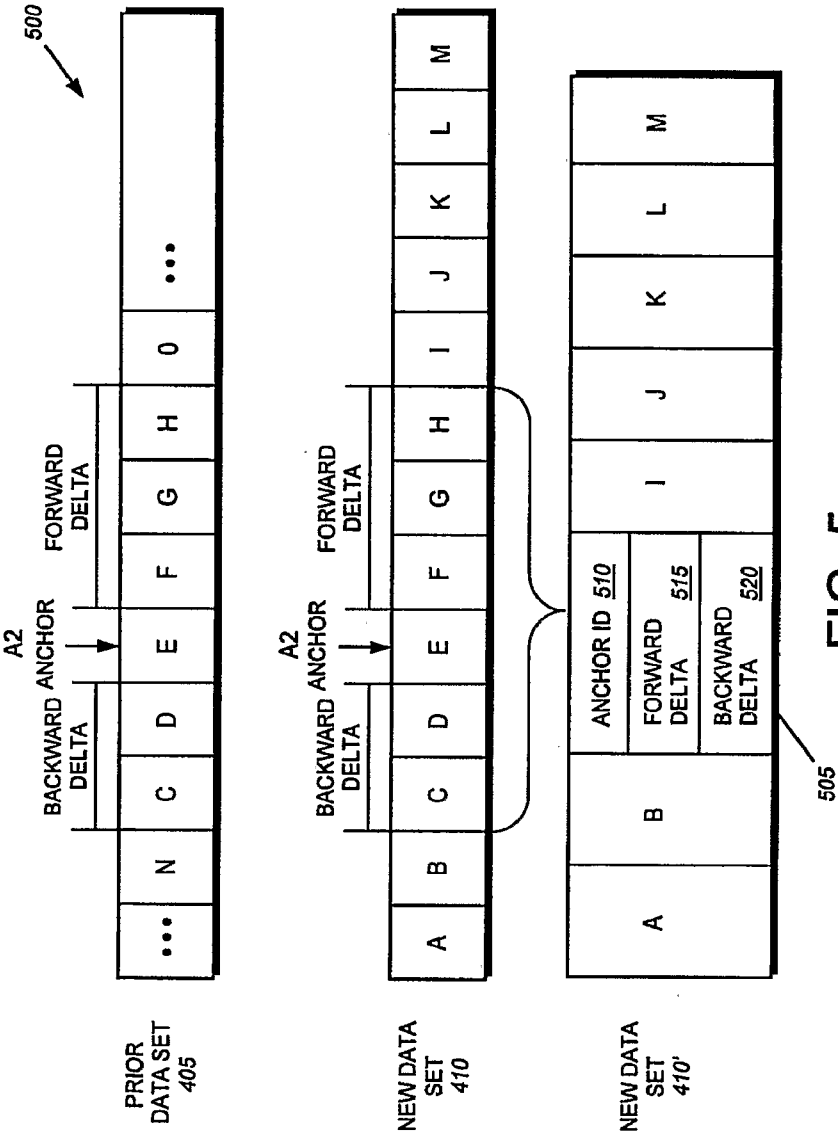


FIG. 5