



(11) **EP 1 642 216 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
22.06.2011 Bulletin 2011/25

(21) Application number: **04756637.7**

(22) Date of filing: **03.07.2004**

(51) Int Cl.:
G06F 17/30 (2006.01)

(86) International application number:
PCT/US2004/021476

(87) International publication number:
WO 2005/008377 (27.01.2005 Gazette 2005/04)

(54) **SNAPSHOTS OF FILE SYSTEMS IN DATA STORAGE SYSTEMS**

SCHNAPPSCHÜSSE VON DATEISYSTEMEN IN DATEISPEICHERSYSTEMEN

IMAGES INSTANTANÉES DE SYSTEMES DE FICHIERS DANS DES SYSTEMES DE STOCKAGE DE DONNEES

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LI LU MC NL PL PT RO SE SI SK TR

(30) Priority: **08.07.2003 US 616128**

(43) Date of publication of application:
05.04.2006 Bulletin 2006/14

(73) Proprietor: **Pillar Data Systems, Inc.**
San Jose, Ca 95134 (US)

(72) Inventors:
• **KAPOOR, Vikram**
Sunnyvale, CA 94085 (US)
• **SHOENS, Kurt, Alan**
Los Altos, CA 94024 (US)

- **SCHULTZ, Mark, Steven**
Los Altos, CA 94024 (US)
- **HAMILTON, Rex, Rilen**
Ben Lomond, CA 95005 (US)

(74) Representative: **Kenrick, Mark Lloyd et al**
Marks & Clerk LLP
Sussex House
83-85 Mosley Street
Manchester
M2 3LG (GB)

(56) References cited:
WO-A2-02/29573 US-A1- 2002 091 670
US-A1- 2003 018 878 US-B1- 6 311 193
US-B1- 6 636 879

EP 1 642 216 B1

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Description

BACKGROUND

- 5 [0001] The present invention relates to snapshots of file systems in data storage systems.
[0002] Further background information may be found in the following:
[0003] U.S. Application No. 10/354,797, Methods and Systems of Host Caching, filed on January 29, 2003 (Attorney Docket No. Pillar 709);
[0004] U.S. Application No. 10/440,347, Methods and Systems of Cache Memory Management and Snapshot Operations, filed on May 16, 2003 (Attorney Docket No. Pillar 713); and
10 [0005] U.S. Application No. Unknown, Systems and Methods of Data Migration in Snapshot Operations, filed on June 19, 2003 (Attorney Docket No. Pillar 711), Express Mail Label No. EJ039579912US.
[0006] Files exist to store information on storage devices (e.g., magnetic disks) and allow the information to be retrieved later. A file system is a collection of files and directories plus operations on them. To keep track of files, file systems have directories. A directory entry provides the information needed to find the blocks associated with a given file. Many file systems today are organized in a general hierarchy (i.e., a tree of directories) because it gives users the ability to organize their files by creating subdirectories. Each file may be specified by giving the absolute path name from the root directory to the file. Every file system contains file attributes such as each file owner and creation time and must be stored somewhere such as in a directory entry.
15 [0007] A snapshot of a file system will capture the content (i.e., files and directories) at an instant in time. A snapshot results in two data images: (1) the active data that an application can be read and write as soon as the snapshot is created and (2) the snapshot data. Snapshots can be taken periodically, hourly, daily, or weekly or on user demand. They are useful for a variety of applications including recovery of earlier versions of a file following an unintended deletion or modification, backup, data mining, or testing of software.
20 [0008] The need for high data availability often requires frequent snapshots that consume resources such as memory, internal memory bandwidth, storage device capacity and the storage device bandwidth. Some important issues for snapshots of file systems is how to manage the allocation of space in the storage devices, how to keep track of the blocks of a given file, and how to make snapshots of file systems work efficiently and reliably.
25 [0009] US2002/091670 describes a method for creating and managing snapshots of a file system.

30 **SUMMARY OF THE INVENTION**

- [0010] The invention provides methods and systems for management of snapshots of a file system.
[0011] In a first aspect of the invention, a snapshot management system performs a method for managing multiple snapshots versions and an active file system by (a) maintaining an index table (195, 196) that contains an entry for each snapshot and the active file system (17), each snapshot having a version number that is unique, monotonic, and not reused; and (b) maintaining a space map block entry (b, e) for each block of usable data storage to indicate snapshots using the block including a first snapshot and a last snapshot, wherein b and e represent index table entries or version numbers, b indicates the first snapshot and e indicates the last snapshot.
35 [0012] In a second aspect of the invention, a snapshot management system includes a processor for maintaining multiple snapshot versions and an active file system, comprising: means for maintaining an index table (195, 196) that contains an entry for each snapshot and the active file system (17), each snapshot having a version number that is unique, monotonic, and not reused; and means for maintaining space map block entry (b, e) for each block of usable data storage to indicate the snapshots using the block, wherein b and e represent index table entries or version numbers, b indicates the first snapshot that uses the block, and e indicates the last snapshot that uses the block.
40
45

BRIEF DESCRIPTION OF THE DRAWINGS

50 **[0013]**

Figure 1 illustrates an embodiment of a data storage system.

Figure 2 illustrates the data storage as a virtualized storage subsystem.

55 Figure 3 illustrates details of the virtualized data storage subsystem organized into file system blocks including, snapshot tables, space map blocks, and usable file system data space.

Figure 4 illustrates details of the space map blocks and a block tracking and index system.

EP 1 642 216 B1

Figure 5 illustrates examples of a snapshot tracking status in the space map blocks.

Figure 6 illustrates details and examples of the snapshot table and snapshot table elements.

5 Figure 7 illustrates a flow diagram for the creation of a file system.

Figure 8 illustrates a flow diagram for file system block management.

10 Figure 9 illustrates a flow diagram for the creation of a file system snapshot.

Figure 10 illustrates a flow diagram of a method to obsolete a file system snapshot.

15 Figure 11 illustrates a flow diagram of a method to delete a snapshot and for a method to clean unneeded snapshots from space map blocks.

Figure 12 illustrates a flow diagram of a method to remove obsolete snapshots from the index table and for a method to remove deleted snapshots from the index table.

20 Figure 13 illustrates a flow diagram of a method to clean obsolete snapshot indexes from space map block entries.

Figure 14 illustrates a flow diagram of a method to clean deleted snapshot indexes from space map block entries.

Figure 15a illustrates a diagram of an active file system with a request to revert to an earlier snapshot.

25 Figure 15b illustrates a diagram of an active file system on hold to obsolete snapshots after the earlier snapshot.

Figure 15c illustrates a diagram of the cleaning of the obsolete snapshots.

30 Figure 15d illustrates a diagram of the file system after reversion to the earlier snapshot.

Figure 16 illustrates a diagram of multiple reversions to earlier snapshots.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

35 **[0014]** The following description includes the best mode of carrying out the invention. The detailed description is made for the purpose of illustrating the general principles of the invention and should not be taken in a limiting sense. The scope of the invention is determined by reference to the claims. Each part is assigned its own part number throughout the specification and drawings.

40 **[0015]** Figure 1 illustrates a data storage system 10, including a plurality of hosts coupled to data storage subsystem (s). Each host is a computer that can connect to client(s), to data storage subsystem(s), and each other. Each host provides software/hardware interfaces such as network interface cards and software drivers to implement Ethernet, Fibre Channel, ATM, and SCSI, and InfiniBand. Hennessy and Patterson, Computer Architecture: A Quantitative Approach (2003), and Patterson and Hennessy, Computer organization and Design: The Hardware /Software Interface (1998) describe computer hardware and software, storage systems, caching, and networks.

45 **[0016]** In an embodiment, the first host includes a motherboard with a CPU-memory bus 14 that communicates with dual processors 12 and 41. The processor used is not essential to the invention and could be any suitable processor such as the Intel Pentium 4 processor. A processor could be any suitable general purpose processor running software, an ASIC dedicated to perform the operations described herein or a field programmable gate array (FPGA). Also, one could implement the invention using a single processor in each host or more than two processors to meet more stringent performance requirements. The arrangement of the processors is not essential to the invention.

50 **[0017]** The first host cache memory 20 includes a cache manager 13, a cache directory 15, and cache lines 16. The cache memory 20 is nonvolatile memory or volatile memory or a combination of both. Nonvolatile memory protects data in the event of a power interruption or a host failure. Data is defined as including user data, instructions, and metadata. Nonvolatile memory may be implemented with a battery that supplies power to the DRAM to make it nonvolatile memory when a conventional external power interrupt circuit detects a power interruption or with inherently nonvolatile semiconductor memory.

55 **[0018]** Each host includes a bus adapter 22 between the CPU-memory bus 14 and an interface bus 24. Each host runs an operating system such as Linux, UNIX, a Windows OS, or another suitable operating system. Tanenbaum,

Modem Operating Systems (2001) describes operating systems in detail . The first host is representative of the other hosts, but this feature is not essential to the invention.

[0019] The first host can communicate with the second host through an interconnect 40, shown as connected to an adapter 25 to the interface bus 24. The PCI bus is one suitable interface bus and the interconnect 40 may be any suitable known bus, SAN, LAN, or WAN technology. In an embodiment, the interconnect 40 is a dedicated Fibre Channel (FC) point-to-point link that connects to FC-PCI bus adapter 25 to provide fast point-to-point communication between the hosts.

[0020] In an alternative embodiment, the interconnect network 30 such as a FC fabric provides extra bandwidth for host-to-host communications. In this embodiment, links 28, 38 connect to the interconnect network 30 and the hosts use link 28 and link 38 when available. FC standard software can set priority levels to ensure high priority peer-to-peer requests, but there will still be some arbitration overhead and latency in claiming ownership of the links. For example, if links 28 and 38 are busy transferring data when a write request arrives, that operation must complete before either link is free for arbitration.

[0021] If the interconnect 40 ever fails, communication between hosts can be handled using the interconnect network 30. The interconnect network 30 can be implemented by interconnects used in data storage systems such as Fibre Channel, SCSI, InfiniBand, or Ethernet, and the type of interconnect is not essential to the invention. In either embodiment, redundant communication between hosts ensures the data storage system has high availability. See Clark, IP SANs: A Guide to iSCSI, iFCP, and FCIP Protocols for Storage Area Networks (2002) and Clark, Designing Storage Area Networks (1999) .

[0022] In an embodiment, the data storage subsystems shown in Figure 1 can be those described in the co-pending U.S. Patent Application No. 10/264,603, entitled, *Systems and Methods of Multiple Access Paths to Single Ported Storage Devices*, filed on October 3, 2002 . It is understood, however, that other storage device(s) or data storage subsystems could be used instead of the data storage subsystem described in that U.S. patent application.

[0023] As shown in Figure 1, the first host connects, or couples, to the first data storage subsystem through the bus adapter 22, the interface bus 24, the adapter 26, the link 28, the interconnection network 30, and the link 32. To connect to the second data storage subsystem, the first host uses the same I/O path except the data passes through link 34, while the second host uses the same type of I/O path plus link 32 to communicate with the first data storage subsystem, or link 34 to communicate with the second data storage subsystem, or link 36 to the data storage subsystem N.

[0024] Figure 2 illustrates that each host of the data storage system can store and retrieve files from a data storage subsystem 35 using an interconnect 28, an interconnect 33, and an interconnect network 30. In an embodiment, each storage device in the data storage subsystem is assigned a logical unit number (LUN) that is an identifier for the storage device. A virtual logical unit number (VLUN) is as an abstraction of the storage device(s) or the virtualization of the data storage subsystems such as a linear array of blocks as it appears to the data storage system users. In various embodiments, the implementation of a VLUN may be striped (i.e., spread) over multiple RAID groups for added performance, spread over sections of a RAID group for flexibility, or copied on multiple RAID groups for reliability. As shown, the storage devices of the data storage subsystem are virtualized as a file system employing contiguous fixed sized blocks 0-n where the size of each block is some value preferably 1-KB to 64-KB.

[0025] Figure 3 illustrates a data storage subsystem for maintaining (e.g., allocating, reading, writing, and deallocating) some blocks for index tables, some for space maps, and others for usable space for data storage. The block(s) allocated to an index table depends on the size of each block and the number of concurrent snapshots supported. For example, three 8-KB blocks may be sufficient space to support an index table of 254 snapshots. As shown, the file system has a pair of index tables 195 and 196 and thus uses six 8-KB blocks so the host can alternate writes to the index tables to ensure recovery in case of a data storage system failure. Thus, if the system fails during a write to one index table, the host can retrieve the unmodified copy of the other index table. While we describe this method of writing to alternate index table copies, other methods, such as write journaling, can be used to protect against system failure during index table writes. The remainder of the storage is allocated to space map blocks with usable space for data storage.

[0026] Each index table includes an index value of the active file system 17 permitting fast location of the active file system. The index table includes a known algorithm to verify the data integrity such as a checksum 18, a cyclic redundancy check, or a digital signature. The index table provides an index to the snapshots and the active file system. Each entry in the index table represents a snapshot or the active file system. As illustrated, the index range is 1-255, but this range is not essential to the invention. In various embodiments, each snapshot and the active file system has one or more associated attributes such as a version number, timestamp and/or image name to identify the snapshot or active file system, an images state, a root block pointer, and an image name as described below. When the data storage system takes a snapshot of the file system it assigns the snapshot a unique version number such as a 32-bit unsigned integer that increases monotonically for each subsequent snapshot. The version number is not reused even as snapshots are deleted or made obsolete to the file system.

[0027] The image state can be one of the following states:

- Active - representing the active file system

EP 1 642 216 B1

- In-use snapshot - representing a snapshot that users can access
- Free - representing available for writing by a snapshot or active file system
- Deleted snapshot - representing a user has deleted the associated snapshot but a cleaner process or thread has not yet removed the space map block entry for this snapshot
- Obsolete snapshot - representing a user has reverted to an earlier snapshot and the cleaner process or thread has not yet removed the space map block entry for this snapshot

[0028] In an embodiment, when the data storage system takes a snapshot of the file system, the host provides a timestamp (e.g., time and date) when the snapshot or active data image was created. The root block pointer provides the address of the root block in the hierarchical structure of the snapshot and the image name is a character string used to easily identify the snapshot to users.

[0029] Referring to Figure 4, the remainder of the data storage subsystem is allocated to space map blocks and usable space for data. Each space map block keeps track of the blocks in its usable space for data storage. For example, a space map block can keep track of 2,047 blocks of usable space. Each space map block entry contains a pair of indexes ("a space map block entry"), for example, 8-bits in length that represent any of 254 snapshots or the active file system. Each space map block entry is an index into the index table shown in Figure 3. Each space map block entry has a beginning value b that indicates the first snapshot (i.e., snapshot b) and an ending value e that indicates the last snapshot (i.e., snapshot e) to use the associated block. Thus, each space map block entry (b, e) in the space map block is used to track the usage of an associated block in the usable space:

Beginning Index, b	Ending Index, e	Meaning
0	0	The block is free to use by a snapshot or the active file system
8-bit Integer	0	The block is in use by one or more snapshots and has not been changed since the first snapshot was taken. The block is also in use by the active file system
a	0	The block is in use by the active file system only and there are no snapshots associated with the block
8-bit Integer	Another 8-bit Integer	The block is in use by a snapshot and is used by all subsequent snapshots through a last snapshot. The data has changed since the last snapshot was created and is, therefore, no longer in use by the active file system

[0030] In an alternative embodiment, each space map block entry contains a pair of version numbers (e.g., 32-bit) that represent snapshots or the active file system. Thus, each version pair (b, e) in the space map block would be used to track the usage of an associated block in the usable space.

[0031] Figure 5 illustrates how a space map block entry is used to track a block in the usable space. In an earlier stage of operation, the space map may contain many entries with beginning and ending values equal to zero indicating that many blocks in the data storage subsystem are free-to-use for snapshots or the active file system. At the same time, other blocks will be in-use that is allocated to snapshots or the active file system. Figure 5 shows one such block used by snapshots assigned version number 10344 through version number 10388 inclusive. The space map block entry (16, 13) indicates the snapshots that use this block. The host associates the space map block entry with the version number of the snapshot by reading the index table. The space map block also contains space map block entry (212, 0) to show that the active file system, for example, assigned version number 65993 is using an associated block.

[0032] Figure 6 illustrates the relationship between the index table 200 and the snapshot attributes and the active file system attributes. The table below illustrates the group of attributes 203, 204, 205, 206, 207, and 208 that are shown in Figure 6:

Index Table Entry	Version Number	Image State	Timestamp	Root block Pointer	Image Name
Snapshot 1	10344	In-use	4 AM, May 8, 2003	1002	Fred
Snapshot 3	10388	In-use	10 AM, May 9, 2003	1004	Smokey

(continued)

Index Table Entry	Version Number	Image State	Timestamp	Root block Pointer	Image Name
Active File System	65993	Active File System	6 PM, May 14, 2003	1006	Hilda
Snapshot 5	12608	Deleted	2 PM, April 2, 2003	1008	Rufus
Snapshot 253	55478	Obsolete	3 AM, May 9, 2003	1010	Vern
Snapshot 254	0	Free	0	0	0

[0033] Figure 7 illustrates a flow diagram of a method for creating a file system. This method can occur during the process of initialization. In an embodiment, the initialization process has proceeded to the point where the next steps relate specifically to file system creation. Although the method is described serially below, the steps can be performed in parallel, for example, asynchronously or in a pipelined manner. There is no requirement the method be performed in the order shown except where indicated. Further, the steps are implemented by computer such as one or more host(s) described earlier. For brevity, we describe the methods as executed by a host.

[0034] Referring to step 105 of Figure 7, the host sets all of the space map block entries equal to (0, 0). This indicates that the blocks corresponding to the entries are free-to-use. At step 106, the host selects an 8-bit index "a" that represents the active file system. At step 108, the host assigns a starting version number to index a. At step 110, the host loads index "a" into the index of the active file system 201 in the directory 200 (Figure 6). At step 112, the host selects the first available space map block and at step 114 loads the beginning index 8-bit index "a" and an 8-bit ending index 0 into the first entry in the selected space map block. At step 116, the host sets the starting version number in the associated attributes for the active file system in index table entry "a". The host further sets the image state to active at step 118, the timestamp to the current date and time at step 120, and a starting root block pointer at step 122, calls an algorithm to verify the data integrity (e.g., checksum) of the snapshot attributes, and stores the results in index table entry "a" at step 124. At step 126, the host may write the index table to nonvolatile storage. In one embodiment, at step 128, the host continues with any other activities such as initialization. In another embodiment, the other activities can precede the creation of the file system.

[0035] Figure 8 illustrates a flow diagram of a method of block management. At step 42, the host receives a request to update the file system data. At step 44, the host reads the space map block entry of the associated received data to determine if that block is used by the active file system only. If yes, as indicated by the space map block entry = (a, 0), the host determines at step 45 whether or not the data is a modification of existing data or an addition to existing data. If the received data is a modification of existing data, the host overwrites the block at step 46 and returns to normal operation at step 58. If the received data is an addition to the existing data, the host determines at step 48 if the space available in the current block is enough to hold all of the received data. If yes, the host adds the received data to the current block at step 50 and returns to normal operation at step 58. If not, the host allocates a free-to-use block at step 52 to hold the additional data and changes the associated space map block entry from (0, 0) --> (a, 0). At step 54, the host adds the received data to the newly allocated block. At step 56, the host updates the file system block pointers to point to the new data. At step 58, the block management routine returns to normal system operation.

[0036] At step 44, if the space map block entry of the block associated with the received data indicates an in-use snapshot uses the block, that is, the space map block entry (b, 0), the host allocates a free-to-use block for the received data at step 60. At step 62, the host adds the received data to the new allocated block. At step 63, the host changes the space map block entry of the new allocated block from (0, 0) --> (a, 0) indicating the new block is used by the active file system only. At step 64, the host updates the file system block pointers to point to the new data. At step 66, the host determines if there are other in-use snapshots pointing to the same old block. If the index b is associated with the latest snapshot version number, there is no other in-use snapshots pointing to the same old block. Therefore, at step 67, the host updates the old space map block entry from (b, 0) --> (b, b) indicating snapshot b is the only snapshot pointing to the associated old block and that the old data has been modified since snapshot b was created. If the index b is not associated with the latest snapshot version number, there is another in-use snapshot pointing to the same old block. Therefore, at step 68, the host updates the old space map block entry from (b, 0) --> (b, e) to indicate that snapshot b is the beginning snapshot and snapshot e is the ending snapshot (i.e., current in-use snapshot with the latest snapshot version number) pointing to the associated old block. In this case, there may be other snapshots with version numbers less than snapshot e and greater than snapshot b pointing to the same old block. In either case, the block management routine returns to normal system operation at step 58.

[0037] Figure 9 illustrates a flow diagram of a method of creating a snapshot. After receiving a request for a new

snapshot of the file system, the host-holds off from updating the active file system at step 140. At step 142, the host searches through the index table for the first snapshot with an image state equal to free. At step 144, if the host searches the entire index table and does not find a free image state, the routine reports no free snapshots and the host resumes normal file system operation at step 164 and if appropriate, the operator can delete some snapshots. If, at step 144, a free image state is found, the host changes the active file system's image state to in-use at step 148. At step 150, the host enters a timestamp of the new in-use snapshot. When this snapshot is created, the version number and the root block pointer remain as they were when the block was allocated to the active file system. At step 151, the host assigns a user-friendly image name for the in-use snapshot. This completes creation of the new snapshot. Next, the host establishes an active file system for normal use. At step 152, the host changes snapshot attributes containing the image state free to the active file system. At step 154, the host assigns the next version number to the new active file system. At step 156, the host enters the current root block pointer (same as the new snapshot) in the attributes of the active file system. At step 158, the host saves a user-friendly image name of the active file system. The snapshot table is written to nonvolatile storage at step 162. The host returns to normal operation at step 164.

[0038] Figure 10 illustrates a flow diagram for a method used to revert to an earlier version of a snapshot. Figures 15a and 15b illustrate diagrams of the method of reversion. At times it is desirable or necessary to revert to an earlier version of the file system. An earlier version of the file system is any previous snapshot of the active file system. Once a previous snapshot is selected, the host discards all versions of the file system after the chosen version, including the current active file system. After receiving a request to revert to the previous snapshot, for example, snapshot p shown in Figure 15a, the host will hold off updates to the current active file system at step 210. At step 212, the host changes the image state of all snapshots after snapshot p from in-use to obsolete and changes the image state of the active file system from active to obsolete (See Figure 15b). At step 214, the host makes a copy of snapshot p, for example, called snapshot c. Snapshot c is used to receive any updates to space map block entry (p, 0) that were held off during the process of reverting to the previous snapshot version. This permits snapshot p to be preserved in its present state after the system is brought back to an active image state, while at the same time, not losing the pending updates. At step 216, the host sets the image state of snapshot c to in-use. At step 218, the host assigns the next unused version to snapshot c. At step 220, the host sets the root block pointer of snapshot c to the same root block pointer of snapshot p. At step 222, the host creates another copy of snapshot p, for example, called copy a. The copy a will become the active file system of the desired image of snapshot p. At step 224, the host sets the image state of copy a to the active file system. At step 226, the host assigns the next version number to copy a. At step 228, the host sets the root block pointer of copy a to the same root block pointer of snapshot p. At step 230 the host records index a of the active file system into the index table 200 (Figure 3). At step 232, the host writes the index table to nonvolatile storage. At step 234, the host returns to normal file system activity.

[0039] Figure 11a illustrates a flow chart for a method to delete a snapshot. At step 75, after receiving a request to delete a snapshot (see also figure 15c) the host searches the index table for the requested snapshot. At step 76, the host sets the image state of the requested snapshot to deleted. At step 77, the host returns to normal file system operation.

[0040] Figure 11b illustrates a high level flow chart for cleaning deleted and obsolete snapshots from the space map blocks and index table of the file system. At step 79, the host determines if any obsolete snapshots exist. If yes, the host goes to reference A in Figure 13 to clean obsolete snapshots from the space map blocks. If not, the host goes to step 81 and determines if any deleted snapshots exist. If not, then no work needs to be done and the method is complete. At step 81, if a deleted snapshot is found, the host goes to reference H in Figure 14 to clean deleted snapshots from the space map blocks and index table of the file system.

[0041] Figure 13 illustrates the detailed flow chart for cleaning obsolete snapshots from space map block entries after a reversion to snapshot p. Step 300 examines each space map block entry one by one in the file system. Step 310 tests whether the beginning index of the space map block entry either matches the snapshot p to which we reverted or precedes snapshot p while the ending index refers to an obsolete snapshot later than p and earlier than c, the snapshot copy created in step 220 on Figure 10. If the space map block entry matches these conditions, step 312 changes the space map block entry to (b, 0) to indicate that the block is now in use by the active file system.

[0042] Step 314 tests if the beginning index of the space map block entry indicates a snapshot later than the reverted-to snapshot p and the ending index indicates an obsolete snapshot earlier than the copy snapshot c. If so, step 316 sets the space map block entry to (0, 0) to indicate that the entry is free-to-use since no snapshot any longer references it.

[0043] If neither of the conditions tested by steps 310 or 314 are true, then step 318 leaves the space map block entry unchanged.

[0044] After executing step 312, 316, or 318, step 306 tests if we have processed the last space map block entry in the file system. If we have processed the last entry, processing continues at Reference J on Figure 12a to removed the index table entry for all the obsolete snapshots. Otherwise, step 308 moves to the next space map block entry and processing continues at step 300.

[0045] After completing the processing of all obsolete snapshots in the space map block, processing continues at Reference J on Figure 12a to remove the index table entries corresponding to obsolete snapshots. Processing begins

at the first index table entry in step 400. Step 410 tests if the index table entry is obsolete. If so, step 420 clears the index table entry by setting the image state to free-to-use, the timestamp to 0, the root block pointer to 0, and by clearing the name reference. Step 430 tests if we have processed the last index table entry. If this is not the last index table entry, step 440 moves to the next index table entry and processing continues at step 410. After all index table entries have been processed, all obsolete snapshots have been completely removed from the file system.

[0046] Returning to Figure 11b, if step 81 detects any deleted snapshots in the index table, processing continues at Reference H on Figure 14. On Figure 14, step 600 begins with the first space map block entry. Step 610 tests if the beginning index of the space map block entry references a deleted snapshot. If yes, then step 620 tests for any active file system or in-use snapshot with a version later than the beginning space map block entry index. If no snapshot or active file system is found later than the space map block entry beginning index, then step 680 sets the entry to (0, 0) to indicate the corresponding block is free to use. If a later snapshot or active file system is found, step 630 sets the beginning index of the space map block entry to the index of the found snapshot or active file system and continues processing at Reference G on the same sheet.

[0047] Step 640 similarly tests the ending index of the space map block entry to see if it references a deleted snapshot. If so, step 650 tests if there is a snapshot with version less than the current ending index and later than or equal to the version of the beginning index. If not, step 680 sets the space map block entry to (0, 0) to indicate that the block is free-to-use. Otherwise, step 660 sets the ending index to the latest in-use snapshot before the current ending index.

[0048] After completion of either step 660 or 680, step 670 tests for another space map block entry. If there are more space map block entries to process, control returns to step 610. After all space map block entries have been processed, control resumes at Reference K on Figure 12b to remove index table entries for deleted snapshots.

[0049] Figure 16 illustrates a diagram of a case where a second reversion to a prior snapshot s has occurred before the background cleaning process for the first reversion to snapshot p has completed. As the file system reverts to a prior snapshot, a pending list is maintained to record the indexes and associated version numbers of the affected snapshots each time the reversion process is invoked. If the file system reverts to more than one snapshot over a span of time where the background cleaning process has not completed, the pending list organizes the cleaning activity by individual reversion event to maintain data reliability of the snapshots. Multiple reversions are likely to occur in large and/or high activity test environments.

Claims

1. A method of snapshot management in a data storage system (10) that maintains multiple snapshot versions and an active file system, comprising:

- (a) maintaining an index table (195, 196) that contains an entry for each snapshot and the active file system (17), each snapshot having a version number that is unique, monotonic, and not reused; and
- (b) maintaining a space map block entry (b, e) for each block of usable data storage to indicate snapshots using the block including a first snapshot and a last snapshot, wherein b and e represent index table entries or version numbers, b indicates the first snapshot and e indicates the last snapshot.

2. The method of claim 1, further comprising:

- (c) allocating a second block;
- (d) writing data in the second block;
- (e) updating file system pointers that pointed to a first block to point to the H second block; and
- (f) repeating the steps (c) through (e) for updating other blocks.

3. The method of claim 1, wherein the index table entry includes one or more attributes selected from a group of attributes comprising a version number (19), an image state (21), a time stamp (23), a root block pointer (27), and an image name (29).

4. The method of claim 2, further comprising (g) reading the associated space map block to determine if the first block is only used for the active file system (17) and if so, updating the first block.

5. The method of claim 4, wherein the space map block entry (a, 0) indicates the first block is only used by the active file system, wherein "a" represents the active file system. (17).

6. The method of claim 1, further comprising:

- (c) reading a first index table entry associated with the active file systems;
- (d) locating (142) a second index table entry that is free-to-use;
- (e) assigning (154) a version number to the second index table entry that is larger than that of the first index table entry;
- (f) copying (156) the root block pointer from the first index table entry to the second index table entry;
- (g) setting (148) the image state of the first index table entry to an in-use snapshot; and
- (h) setting (152) the image state of the second index table entry to the active file system to create a snapshot.

7. The method of claim 1, further comprising:

- (c) maintaining in the index table entries for snapshots an image state (21) that indicates if an entry is deleted or not; and
- (d) marking (76) the image state as deleted to delete a snapshot.

8. A method of claim 1, further comprising:

- (c) ordering the index table (195, 196) by time of snapshot creation;
- (d) reading each space map block entry to determine if the first snapshot was deleted and if so, replacing the index b with the earliest index of a later undeleted snapshot or the active file system, determining if the last snapshot was deleted and if so, replacing the index e with the latest index of an earlier undeleted snapshot to clean deleted snapshots.

9. The method of claim 8, further comprising (e) resetting the image state to free for all index table entries of deleted snapshots when all space map block entries have been processed.

10. The method of claim 8, further comprising (e) setting the space map block entry to indicate that the associated block is free if no undeleted snapshot exists between index b and index e.

11. The method of claim 10, further comprising (f) resetting the image state to free for all index table entries of deleted snapshots when all space map block entries have been processed.

12. The method of claim 1, further comprising:

- (c) maintaining in the index table entries for the snapshots an image state (21) that indicates if an entry is obsolete or not;
- (d) creating a new snapshot c by making a copy of a previous snapshot p;
- (e) creating an active file system a by making a copy of the previous snapshot p;
- and
- (f) setting the image state of all in-use snapshots before snapshot c and after snapshot p to obsolete to revert to the previous snapshot p.

13. The method of claim 12, further comprising:

- (g) reading each space map block entry to determine:
 - (i) if the first snapshot is snapshot p or earlier than snapshot p and the last snapshot is earlier than snapshot c, replacing the index e with a value indicating that the block is still in use by the active file system,
 - (ii) if the first snapshot is later than snapshot p and earlier than snapshot c, setting the space map block entry to indicate that the associated block is free;
- (h) marking the image state of snapshot c as deleted; and
- (i) resetting the image state of all obsolete index table entries to free to clean obsolete snapshots after a revert operation to snapshot p.

14. The method of claim 13, further comprising repeating steps (g) and (h) over a list of snapshots p and of snapshots c.

15. The method of claim 14, wherein the pending list includes the reverted snapshots p and the created snapshots c.

16. A snapshot management system that maintains multiple snapshot versions and an active file system (17), comprising:

5 means for maintaining an index table (195, 196) that contains an entry for each snapshot and the active file system (17), each snapshot having a version number that is unique, monotonic, and not reused; and
means for maintaining space map block entry (b, e) for each block of usable data storage to indicate the snapshots using the block, wherein b and e represent index table entries or version numbers, b indicates the first snapshot that uses the block, and e indicates the last snapshot that uses the block.

10 17. The system of claim 16, further comprising:

means for allocating a second block;
means for writing data in the second block;
means for updating file system pointers that pointed to the first block to point to the second block; and
means for repeating the above steps for updating other blocks.

15 18. The system of claim 16, wherein the index table entry includes one or more attributes selected from a group of attributes comprising a version number (19), an image state (21), a time stamp (23), a root block pointer (27), and an image name (29).

20 19. The system of claim 17, further comprising means for reading the associated space map block to determine if the first block is only used for the active file system (17) and if so, updating the first block.

25 20. The system of claim 19, wherein the space map block entry (a, 0) indicates the first block is only used by the active file system, wherein "a" represents the active file system (17).

30 21. The system of claim 16, further comprising:

means for reading a first index table entry associated with the active file system;
means for locating (142) a second index table entry that is free-to-use;
means for assigning (154) a version number to the second index table entry that is larger than that of the first index table entry;
means for copying (156) the root block pointer from the first index table entry to the second index table entry;
means for setting (148) the image state of the first index table entry to an in-use snapshot; and
means for setting (152) the image state of the second index table entry to the active file system to create a snapshot.

35 22. The system of claim 16, further comprising:

means for maintaining in the index table for the snapshots an image state (21) that indicates if an entry is deleted or not; and
means for marking (76) the image state as deleted to delete a snapshot.

40 23. The system of claim 16, further comprising:

means for ordering the index table by time of snapshot creation; and
means for reading each space map block entry to determine if the first snapshot was deleted and if so, replacing the index b with the earliest index of a later undeleted snapshot or the active file system, determining if the last snapshot was deleted and if so, replacing the index e with the latest index of an earlier undeleted snapshot to clean deleted snapshots.

45 24. The system of claim 23, further comprising means for resetting the image state to free of all index table entries of deleted snapshots when all space map block entries have been processed.

50 25. The system of claim 23, further comprising means for setting the space map block entry to indicate that the associated block is free if no undeleted snapshot exists between index b and index e.

55 26. The system of claim 16, further comprising:

means for maintaining in the index table entries for the snapshots an image state (21) that indicates if an entry is obsolete or not;

means for creating a new snapshot c corresponding to a previous snapshot p;

means for creating an active file system a corresponding to the previous snapshot p; and

5 means for setting the image state of all in-use snapshots before snapshot c and after snapshot p to obsolete to revert to the previous snapshot p.

27. The system of claim 26, further comprising:

10 means for reading each space map block entry to determine:

(i) if the first snapshot is snapshot p or earlier than snapshot p and the last snapshot is earlier than snapshot c, replacing the index e with a value indicating that the block is still in use by the active file system,

15 (ii) if the first snapshot is later than snapshot p and earlier than snapshot c, setting the space map block entry to indicate that the associated block is free;

means for marking the image state of snapshot c as deleted; and

means for resetting the image state of all obsolete index table entries to free to clean obsolete snapshots after a revert operation to snapshot p.

20 28. The system of claim 27, further comprising means for repeating the reading and marking functions over a list of snapshots p and of snapshots c

25 29. The system of claim 28, wherein the list includes the unclean reverted snapshots p and the created snapshots c.

30 30. The snapshot management system of claim 16 further including a processor for maintaining said multiple snapshot versions and said active file system, and further comprising:

a usable space for storing the snapshot versions and the active file system (17).

31. The system of claim 30 wherein the processor is programmed to execute steps of:

allocating a second block;

writing the data in the second block;

35 updating file system pointers that pointed to the first block to point to the second block; and repeating the above steps for updating other blocks.

40 32. The system of claim 30, wherein the index table entry includes one or more attributes selected from a group of attributes comprising a version number (19), an image state (21), a time stamp (23), a root block pointer (27), and an images name (29).

33. The system of claim 31, wherein the processor reads the associated space map block to determine if the first block is only used for the active file system (17) and if so, updates the first block.

45 34. The system of claim 30, wherein the space map block entry (a, 0) indicates the first block is only used by the active file system, wherein "a" represents the active file system (17).

35. The system of claim 30, wherein the processor is programmed to execute steps of:

50 reading a first index table entry associated with the active file system;

locating (142) a second index table entry that is free-to-use;

assigning (154) a version number to the second index table entry that is larger than that of the first index table entry;

copying (156) the root block pointer from the first index table entry to the second index table entry;

55 setting (148) the image state of the first index table entry to an in-use snapshot; and

setting (152) the image state of the second index table entry to the active file system to create a snapshot

36. The system of claim 30, wherein the processor is programmed to execute steps:

maintaining in the index table entries for snapshots an image state (21) that indicates if an entry is deleted or not; and
marking (76) the image state as deleted to delete a snapshot.

5 37. The system of claim 30, wherein the processor is programmed to execute steps:

ordering the index table (195, 196) by time of snapshot creation;
reading each space map block entry to determine if the first snapshot was deleted and if so, replacing the index
10 b with the earliest index of a later undeleted snapshot or the active file system, determining if the last snapshot
was deleted and if so, replacing the index e with the latest index of an earlier undeleted snapshot to clean
deleted snapshots.

15 38. The system of claim 37, wherein the processor is programmed to reset the image state to free for all index table
entries of deleted snapshots when all space map block entries have been processed.

39. The system of claim 37, wherein the processor is programmed to set the space j map block entry to indicate the
associated block is free if no undeleted snapshot exists between index b and index e.

20 40. The system of claim 39 wherein the processor is programmed to reset the image state to free for all index table
entries of deleted snapshots when all space map block entries have been processed.

41. The system of claim 30, wherein the processor is programmed to execute:

25 maintaining in the index table entries for the snapshots an image state that indicates if an entry is obsolete or not;
creating a new snapshot c by making a copy of a previous snapshot p;
creating an active file system a by making a copy of the previous snapshot p, and
setting the image state of all in-use snapshots before snapshot c and after snapshot p to obsolete to revert to
a previous snapshot p.

30

Patentansprüche

35 1. Verfahren für Schnappschussverwaltung in einem Datenspeichersystem (10), das mehrere Schnappschussversionen
und ein aktives Dateisystem führt, umfassend:

(a) Führen einer Indextabelle (195, 196), die einen Eintrag für jeden Schnappschuss und das aktive Dateisystem
(17) enthält, wobei jeder Schnappschuss eine Versionsnummer hat, die eindeutig und monoton ist und nicht
wiederverwendet wird; und

40 (b) Führen eines Raumabbildblock-Eintrags (b, e) für jeden nutzbaren Datenspeicherblock zum Angeben von
Schnappschüssen, die den Block verwenden, der einen ersten Schnappschuss und einen letzten Schnapp-
schuss enthält, wobei b und e Indextabellen-Einträge oder Versionsnummern repräsentieren, b den ersten
Schnappschuss angibt und e den letzten Schnappschuss angibt.

45 2. Verfahren nach Anspruch 1, weiter umfassend:

(c) Zuteilen eines zweiten Blocks;

(d) Schreiben von Daten in den zweiten Block;

(e) Aktualisieren von Dateisystemzeigern, die auf einen ersten Block gezeigt haben, um auf den zweiten Block
zu zeigen; und

50 (f) Wiederholen der Schritte (c) bis (e) zum Aktualisieren anderer Blöcke.

55 3. Verfahren nach Anspruch 1, wobei der Indextabellen-Eintrag ein oder mehr Attribute enthält, die aus einer Gruppe
von Attributen ausgewählt werden, die eine Versionsnummer (19), einen Bildstatus (21), einen Zeitstempel (23),
einen Stammblockzeiger (27) und einen Bildnamen (29) umfasst.

4. Verfahren nach Anspruch 2, weiter umfassend (g) Lesen des assoziierten Raumabbildblocks, um zu bestimmen,
ob der erste Block nur für das aktive Dateisystem (17) verwendet wird, und wenn ja, Aktualisieren des ersten Blocks.

EP 1 642 216 B1

5. Verfahren nach Anspruch 4, wobei der Raumabbildblock-Eintrag (a, 0) angibt, dass der erste Block nur von dem aktiven Dateisystem verwendet wird, wobei "a" das aktive Dateisystem (17) repräsentiert.

6. Verfahren nach Anspruch 1, weiter umfassend:

- 5
- (c) Lesen eines ersten Indextabellen-Eintrags, der mit dem aktiven Dateisystem assoziiert ist;
 - (d) Lokalisieren (142) eines zweiten Indextabellen-Eintrags, der frei zur Verwendung ist;
 - (e) Zuweisen (154) einer Versionsnummer zu dem zweiten Indextabellen-Eintrag, die größer als die des ersten Indextabellen-Eintrags ist;
 - 10 (f) Kopieren (156) des Stammblockzeigers von dem ersten Indextabellen-Eintrag zu dem zweiten Indextabellen-Eintrag;
 - (g) Setzen (148) des Bildstatus des ersten Indextabellen-Eintrags auf einen Schnappschuss in Gebrauch; und
 - (h) Setzen (152) des Bildstatus des zweiten Indextabellen-Eintrags auf das aktive Dateisystem zum Erzeugen eines Schnappschusses.
- 15

7. Verfahren nach Anspruch 1, weiter umfassend:

- (c) Führen, in den Indextabellen-Einträgen für Schnappschüsse, eines Bildstatus (21), der angibt, ob ein Eintrag gelöscht ist oder nicht; und
- 20 (d) Kennzeichnen (76) des Bildstatus als gelöscht, um einen Schnappschuss zu löschen.

8. Verfahren nach Anspruch 1, weiter umfassend:

- (c) Ordnen der Indextabelle (195, 196) nach der Zeit der Schnappschusserzeugung;
 - 25 (d) Lesen jedes Raumabbildblock-Eintrags, um zu bestimmen, ob der erste Schnappschuss gelöscht wurde, und wenn ja, Ersetzen des Indexes b durch den frühesten Index eines späteren nicht gelöschten Schnappschusses oder des aktiven Dateisystems, Bestimmen, ob der letzte Schnappschuss gelöscht wurde, und wenn ja, Ersetzen des Indexes e durch den letzten Index eines früheren nicht gelöschten Schnappschusses, um gelöschte Schnappschüsse zu säubern.
- 30

9. Verfahren nach Anspruch 8, weiter umfassend (e) Zurücksetzen des Bildstatus auf frei für alle Indextabellen-Einträge von gelöschten Schnappschüssen, wenn alle Raumabbildblock-Einträge verarbeitet wurden.

10. Verfahren nach Anspruch 8, weiter umfassend (e) Setzen des Raumabbildblock-Eintrags zum Angeben, dass der assoziierte Block frei ist, wenn kein nicht gelöschter Schnappschuss zwischen Index b und Index e vorhanden ist.

11. Verfahren nach Anspruch 10, weiter umfassend (f) Zurücksetzen des Bildstatus auf frei für alle Indextabellen-Einträge von gelöschten Schnappschüssen, wenn alle Raumabbildblock-Einträge verarbeitet wurden.

12. Verfahren nach Anspruch 1, weiter umfassend:

- (c) Führen, in den Indextabellen-Einträgen für die Schnappschüsse, eines Bildstatus (21), der angibt, ob ein Eintrag veraltet ist oder nicht;
- 45 (d) Erzeugen eines neuen Schnappschusses c durch Erstellen einer Kopie eines vorherigen Schnappschusses p;
- (e) Erzeugen eines aktiven Dateisystems a durch Erstellen einer Kopie des vorherigen Schnappschusses p; und
- (f) Setzen des Bildstatus aller Schnappschüsse in Gebrauch vor Schnappschuss c und nach Schnappschuss p auf veraltet, um den vorherigen Schnappschuss p wiederherzustellen.

13. Verfahren nach Anspruch 12, weiter umfassend:

- (g) Lesen jedes Raumabbildblock-Eintrags, um zu bestimmen:
 - 55 (i) ob der erste Schnappschuss Schnappschuss p oder früher als Schnappschuss p ist und der letzte Schnappschuss früher als Schnappschuss c ist, Ersetzen des Indexes e durch einen Wert, der angibt, dass der Block noch durch das aktive Dateisystem in Gebrauch ist,
 - (ii) ob der erste Schnappschuss später als Schnappschuss p und früher als Schnappschuss c ist, Setzen des Raumabbildblock-Eintrags, um anzugeben, dass der assoziierte Block frei ist;

EP 1 642 216 B1

- (h) Kennzeichnen des Bildstatus von Schnappschuss c als gelöscht; und
- (i) Zurücksetzen des Bildstatus aller veralteten Indextabellen-Einträge auf frei, um veraltete Schnappschüsse nach einer Wiederherstellungsoperation des Schnappschusses p zu säubern.

5 **14.** Verfahren nach Anspruch 13, weiter umfassend Wiederholen der Schritte (g) und (h) über eine Liste von Schnappschüssen p und von Schnappschüssen c.

15. Verfahren nach Anspruch 14, wobei die anstehende Liste die wiederhergestellten Schnappschüsse p und die erzeugten Schnappschüsse c enthält.

10 **16.** Schnappschussverwaltungssystem, das mehrere Schnappschussversionen und ein aktives Dateisystem (17) führt, umfassend:

15 Mittel zum Führen einer Indextabelle (195, 196), die einen Eintrag für jeden Schnappschuss und das aktive Dateisystem (17) enthält, wobei jeder Schnappschuss eine Versionsnummer hat, die eindeutig und monoton ist und nicht wiederverwendet wird; und

Mittel zum Führen eines Raumabbildblock-Eintrags (b, e) für jeden nutzbaren Datenspeicherblock zum Angeben der Schnappschüsse, die den Block verwenden, wobei b und e Indextabellen-Einträge oder Versionsnummern repräsentieren, b den ersten Schnappschuss angibt, der den Block verwendet, und e den letzten Schnappschuss angibt, der den Block verwendet.

20

17. System nach Anspruch 16, weiter umfassend:

25 Mittel zum Zuteilen eines zweiten Blocks;
Mittel zum Schreiben von Daten in den zweiten Block;
Mittel zum Aktualisieren von Dateisystemzeigern, die auf einen ersten Block gezeigt haben, um auf den zweiten Block zu zeigen; und
Mittel zum Wiederholen der obigen Schritte zum Aktualisieren anderer Blöcke.

30 **18.** System nach Anspruch 16, wobei der Indextabellen-Eintrag ein oder mehr Attribute enthält, die aus einer Gruppe von Attributen ausgewählt werden, die eine Versionsnummer (19), einen Bildstatus (21), einen Zeitstempel (23), einen Stammblockzeiger (27) und einen Bildnamen (29) umfasst.

35 **19.** System nach Anspruch 17, weiter umfassend Mittel zum Lesen des assoziierten Raumabbildblocks, um zu bestimmen, ob der erste Block nur für das aktive Dateisystem (17) verwendet wird, und wenn ja, Aktualisieren des ersten Blocks.

20. System nach Anspruch 19, wobei der Raumabbildblock-Eintrag (a, 0) angibt, dass der erste Block nur von dem aktiven Dateisystem verwendet wird, wobei "a" das aktive Dateisystem (17) repräsentiert.

40 **21.** System nach Anspruch 16, weiter umfassend:

Mittel zum Lesen eines ersten Indextabellen-Eintrags, der mit dem aktiven Dateisystem assoziiert ist;
Mittel zum Lokalisieren (142) eines zweiten Indextabellen-Eintrags, der frei zur Verwendung ist;
45 Mittel zum Zuweisen (154) einer Versionsnummer zu dem zweiten Indextabellen-Eintrag, die größer als die des ersten Indextabellen-Eintrags ist;
Mittel zum Kopieren (156) des Stammblockzeigers von dem ersten Indextabellen-Eintrag zu dem zweiten Indextabellen-Eintrag;
Mittel zum Setzen (148) des Bildstatus des ersten Indextabellen-Eintrags auf einen Schnappschuss in Gebrauch;
50 und
Mittel zum Setzen (152) des Bildstatus des zweiten Indextabellen-Eintrags auf das aktive Dateisystem zum Erzeugen eines Schnappschusses.

22. System nach Anspruch 16, weiter umfassend:

55 Mittel zum Führen, in der Indextabelle für die Schnappschüsse, eines Bildstatus (21), der angibt, ob ein Eintrag gelöscht ist oder nicht; und
Mittel zum Kennzeichnen (76) des Bildstatus als gelöscht, um einen Schnappschuss zu löschen.

23. System nach Anspruch 16, weiter umfassend:

5 Mittel zum Ordnen der Indextabelle nach der Zeit der Schnappschusserzeugung; und
Mittel zum Lesen jedes Raumabbildblock-Eintrags, um zu bestimmen, ob der erste Schnappschuss gelöscht wurde, und wenn ja, Ersetzen des Indexes b durch den frühesten Index eines späteren nicht gelöschten Schnappschusses oder des aktiven Dateisystems, Bestimmen, ob der letzte Schnappschuss gelöscht wurde, und wenn ja, Ersetzen des Indexes e durch den letzten Index eines früheren nicht gelöschten Schnappschusses, um gelöschte Schnappschüsse zu säubern.

10 24. System nach Anspruch 23, weiter umfassend Mittel zum Zurücksetzen des Bildstatus auf frei von allen Indextabellen-Einträgen von gelöschten Schnappschüssen, wenn alle Raumabbildblock-Einträge verarbeitet wurden.

15 25. System nach Anspruch 23, weiter umfassend Mittel zum Setzen des Raumabbildblock-Eintrags zum Angeben, dass der assoziierte Block frei ist, wenn kein nicht gelöschter Schnappschuss zwischen Index b und Index e vorhanden ist.

20 26. System nach Anspruch 16, weiter umfassend:

Mittel zum Führen, in den Indextabellen-Einträgen für die Schnappschüsse, eines Bildstatus (21), der angibt, ob ein Eintrag veraltet ist oder nicht;
25 Mittel zum Erzeugen eines neuen Schnappschusses c, der mit einem vorherigen Schnappschuss p korrespondiert;
Mittel zum Erzeugen eines aktiven Dateisystems a, das mit dem vorherigen Schnappschuss p korrespondiert; und
Mittel zum Setzen des Bildstatus aller Schnappschüsse in Gebrauch vor Schnappschuss c und nach Schnappschuss p auf veraltet, um den vorherigen Schnappschuss p wiederherzustellen.

30 27. System nach Anspruch 26, weiter umfassend:

Mittel zum Lesen jedes Raumabbildblock-Eintrags, um zu bestimmen:

- 35 (i) ob der erste Schnappschuss Schnappschuss p oder früher als Schnappschuss p ist und der letzte Schnappschuss früher als Schnappschuss c ist, Ersetzen des Indexes e durch einen Wert, der angibt, dass der Block noch durch das aktive Dateisystem in Gebrauch ist,
(ii) ob der erste Schnappschuss später als Schnappschuss p und früher als Schnappschuss c ist, Setzen des Raumabbildblock-Eintrags, um anzugeben, dass der assoziierte Block frei ist;

Mittel zum Kennzeichnen des Bildstatus von Schnappschuss c als gelöscht; und
Mittel zum Zurücksetzen des Bildstatus aller veralteten Indextabellen-Einträge auf frei, um veraltete Schnappschüsse nach einer Wiederherstellungsoperation des Schnappschusses p zu säubern.

40 28. System nach Anspruch 27, weiter umfassend Mittel zum Wiederholen der Lese- und Kennzeichnungsfunktionen über eine Liste von Schnappschüssen p und von Schnappschüssen c.

45 29. System nach Anspruch 28, wobei die Liste die unsauberen wiederhergestellten Schnappschüsse p und die erzeugten Schnappschüsse c enthält.

50 30. Schnappschussverwaltungssystem nach Anspruch 16, weiter enthaltend einen Prozessor zum Führen der mehreren Schnappschussversionen und des aktiven Dateisystems, und weiter umfassend:

einen nutzbaren Raum zum Speichern der Schnappschussversionen und des aktiven Dateisystems (17).

55 31. System nach Anspruch 30, wobei der Prozessor programmiert ist, um folgende Schritte auszuführen:

Zuteilen eines zweiten Blocks;
Schreiben der Daten in den zweiten Block;
Aktualisieren von Dateisystemzeigern, die auf den ersten Block gezeigt haben, um auf den zweiten Block zu zeigen; und
Wiederholen der obigen Schritte zum Aktualisieren anderer Blöcke.

EP 1 642 216 B1

32. System nach Anspruch 30, wobei der Indextabellen-Eintrag ein oder mehr Attribute enthält, die aus einer Gruppe von Attributen ausgewählt werden, die eine Versionsnummer (19), einen Bildstatus (21), einen Zeitstempel (23), einen Stammblockzeiger (27) und einen Bildnamen (29) umfasst.
- 5 33. System nach Anspruch 31, wobei der Prozessor den assoziierten Raumabbildblock liest, um zu bestimmen, ob der erste Block nur für das aktive Dateisystem (17) verwendet wird, und wenn ja, den ersten Block aktualisiert.
34. System nach Anspruch 30, wobei der Raumabbildblock-Eintrag (a, 0) angibt, dass der erste Block nur von dem aktiven Dateisystem verwendet wird, wobei "a" das aktive Dateisystem (17) repräsentiert.
- 10 35. System nach Anspruch 30, wobei der Prozessor programmiert ist, um folgende Schritte auszuführen:
- Lesen eines ersten Indextabellen-Eintrags, der mit dem aktiven Dateisystem assoziiert ist;
Lokalisieren (142) eines zweiten Indextabellen-Eintrags, der frei zur Verwendung ist;
15 Zuweisen (154) einer Versionsnummer zu dem zweiten Indextabellen-Eintrag, die größer als die des ersten Indextabellen-Eintrags ist;
Kopieren (156) des Stammblockzeigers von dem ersten Indextabellen-Eintrag zu dem zweiten Indextabellen-Eintrag;
Setzen (148) des Bildstatus des ersten Indextabellen-Eintrags auf einen Schnappschuss in Gebrauch; und
20 Setzen (152) des Bildstatus des zweiten Indextabellen-Eintrags auf das aktive Dateisystem zum Erzeugen eines Schnappschusses.
36. System nach Anspruch 30, wobei der Prozessor programmiert ist, um folgende Schritte auszuführen:
- 25 Führen, in den Indextabellen-Einträgen für Schnappschüsse, eines Bildstatus (21), der angibt, ob ein Eintrag gelöscht ist oder nicht; und
Kennzeichnen (76) des Bildstatus als gelöscht, um einen Schnappschuss zu löschen.
37. System nach Anspruch 30, wobei der Prozessor programmiert ist, um folgende Schritte auszuführen:
- 30 Ordnen der Indextabelle (195, 196) nach der Zeit der Schnappschusserzeugung;
Lesen jedes Raumabbildblock-Eintrags, um zu bestimmen, ob der erste Schnappschuss gelöscht wurde, und wenn ja, Ersetzen des Indexes b durch den frühesten Index eines späteren nicht gelöschten Schnappschusses oder des aktiven Dateisystems, Bestimmen, ob der letzte Schnappschuss gelöscht wurde, und wenn ja, Ersetzen
35 des Indexes e durch den letzten Index eines früheren nicht gelöschten Schnappschusses, um gelöschte Schnappschüsse zu säubern.
38. System nach Anspruch 37, wobei der Prozessor programmiert ist zum Zurücksetzen des Bildstatus auf frei für alle Indextabellen-Einträge von gelöschten Schnappschüssen, wenn alle Raumabbildblock-Einträge verarbeitet wurden.
- 40 39. System nach Anspruch 37, wobei der Prozessor programmiert ist zum Setzen des Raumabbildblock-Eintrags zum Angeben, dass der assoziierte Block frei ist, wenn kein nicht gelöschter Schnappschuss zwischen Index b und Index e vorhanden ist.
- 45 40. System nach Anspruch 39, wobei der Prozessor programmiert ist zum Zurücksetzen des Bildstatus auf frei für alle Indextabellen-Einträge von gelöschten Schnappschüssen, wenn alle Raumabbildblock-Einträge verarbeitet wurden.
41. System nach Anspruch 30, wobei der Prozessor programmiert ist zum Ausführen:
- 50 Führen, in den Indextabellen-Einträgen für die Schnappschüsse, eines Bildstatus, der angibt, ob ein Eintrag veraltet ist oder nicht;
Erzeugen eines neuen Schnappschusses c durch Erstellen einer Kopie eines vorherigen Schnappschusses p;
Erzeugen eines aktiven Dateisystems a durch Erstellen einer Kopie des vorherigen Schnappschusses p; und
Setzen des Bildstatus aller Schnappschüsse in Gebrauch vor Schnappschuss c und nach Schnappschuss p
55 auf veraltet, um einen vorherigen Schnappschuss p wiederherzustellen.

Revendications

1. Procédé de gestion d'images instantanées dans un système de stockage de données (10) qui maintient plusieurs versions d'images instantanées et un système de fichiers actif, comprenant les étapes consistant à :

(a) maintenir une table d'index (195, 196) qui contient une entrée pour chaque image instantanée et le système de fichiers actif (17), chaque image instantanée ayant un numéro de version qui est unique, monotone et non réutilisé ; et

(b) maintenir une entrée de bloc de plan d'espace (b, e) pour chaque bloc de stockage de données utilisable pour indiquer des images instantanées utilisant le bloc incluant une première image instantanée et une dernière image instantanée, où b et e représentent des entrées de table d'index ou des numéros de version, b désigne la première image instantanée et e désigne la dernière image instantanée.

2. Procédé selon la revendication 1, comprenant en outre les étapes consistant à :

(c) allouer un deuxième bloc ;

(d) écrire des données dans le deuxième bloc ;

(e) mettre à jour les pointeurs du système de fichiers qui pointaient vers un premier bloc pour qu'ils pointent vers le deuxième bloc ; et

(f) répéter les étapes (c) à (e) pour mettre à jour d'autres blocs.

3. Procédé selon la revendication 1, dans lequel l'entrée de la table d'index comprend un ou plusieurs attributs sélectionnés dans un groupe d'attributs comprenant un numéro de version (19), un état d'image (21), un horodatage (23), un pointeur de bloc racine (27) et un nom d'image (29).

4. Procédé selon la revendication 2, comprenant en outre (g) de lire le bloc de plan d'espace associé pour déterminer si le premier bloc est uniquement utilisé pour le système de fichiers actif (17) et, si tel est le cas, mettre à jour le premier bloc.

5. Procédé selon la revendication 4, dans lequel l'entrée de bloc de plan d'espace (a, 0) indique que le premier bloc est uniquement utilisé par le système de fichiers actif, où « a » représente le système de fichiers actif (17).

6. Procédé selon la revendication 1, comprenant en outre les étapes consistant à :

(c) lire une première entrée de table d'index associée au système de fichiers actif ;

(d) situer (142) une deuxième entrée de table d'index qui est libre d'utilisation ;

(e) affecter (154) un numéro de version à la deuxième entrée de table d'index qui est plus élevé que celui de la première entrée de table d'index ;

(f) copier (156) le pointeur de bloc racine depuis la première entrée de table d'index vers la deuxième entrée de table d'index ;

(g) régler (148) l'état d'image de la première entrée de table d'index à une image instantanée en usage ; et

(h) régler (152) l'état d'image de la deuxième entrée de table d'index au système de fichiers actif afin de créer une image instantanée.

7. Procédé selon la revendication 1, comprenant en outre les étapes consistant à :

(c) maintenir, dans les entrées de table d'index pour les images instantanées, un état d'image (21) qui indique si une entrée est effacée ou non ; et

(d) marquer (76) l'état d'image comme effacé pour effacer une image instantanée.

8. Procédé selon la revendication 1, comprenant en outre les étapes consistant à :

(c) ordonner la table d'index (195, 196) suivant l'heure de la création de l'image instantanée ;

(d) lire chaque entrée de bloc de plan d'espace pour déterminer si la première image instantanée a été effacée et, si tel est le cas, remplacer l'index b par l'index le plus précoce d'une image instantanée non effacée plus tardive ou du système de fichiers actif, déterminer si la dernière image instantanée a été effacée et, si tel est le cas, remplacer l'index e par le dernier index d'une image instantanée non effacée antérieure pour nettoyer les images instantanées effacées.

EP 1 642 216 B1

9. Procédé selon la revendication 8, comprenant en outre (e) de régler à nouveau l'état d'image à libre pour toutes les entrées de table d'index des images instantanées effacées lorsque toutes les entrées de bloc de plan d'espace ont été traitées.

5 10. Procédé selon la revendication 8, comprenant en outre (e) de régler l'entrée de bloc de plan d'espace pour indiquer que le bloc associé est libre si aucune image instantanée non effacée n'existe entre l'index b et l'index e.

10 11. Procédé selon la revendication 10, comprenant en outre (f) de régler à nouveau l'état d'image à libre pour toutes les entrées de table d'index des images instantanées effacées lorsque toutes les entrées de bloc de plan d'espace ont été traitées.

12. Procédé selon la revendication 1, comprenant en outre les étapes consistant à :

15 (c) maintenir, dans les entrées de table d'index pour les images instantanées, un état d'image (21) qui indique si une entrée est obsolète ou non ;

(d) créer une nouvelle image instantanée c en effectuant une copie d'une précédente image instantanée p ;

(e) créer un système de fichiers actif a en effectuant une copie de la précédente image instantanée p ; et

(f) régler l'état d'image de toutes les images instantanées en usage avant l'image instantanée c et après l'image instantanée p à obsolète pour revenir à la précédente image instantanée p.

20

13. Procédé selon la revendication 12, comprenant en outre les étapes consistant à :

(g) lire chaque entrée de bloc de plan d'espace pour déterminer :

25

(i) si la première image instantanée est l'image instantanée p ou une image instantanée antérieure à l'image instantanée p, et si la dernière image instantanée est antérieure à l'image instantanée c, remplacer l'index e par une valeur indiquant que le bloc est toujours utilisé par le système de fichiers actif,

(ii) si la première image instantanée est postérieure à l'image instantanée p, et antérieure à l'image instantanée c, régler l'entrée de bloc de plan d'espace pour indiquer que le bloc associé est libre ;

30

(h) marquer l'état d'image de l'image instantanée c comme effacé ; et

(i) régler à nouveau l'état d'image de toutes les entrées de table d'index obsolètes à libre pour effacer les images instantanées obsolètes après une opération de retour à l'image instantanée p.

35

14. Procédé selon la revendication 13, comprenant en outre de répéter les étapes (g) et (h) sur une liste d'images instantanées p et d'images instantanées c.

15. Procédé selon la revendication 14, dans lequel la liste en cours comprend les images instantanées p revenues et les images instantanées c créées.

40

16. Système de gestion d'images instantanées qui maintient plusieurs versions d'images instantanées et un système de fichiers actif (17), comprenant :

45

un moyen pour maintenir une table d'index (195, 196) qui contient une entrée pour chaque image instantanée et le système de fichiers actif (17), chaque image instantanée ayant un numéro de version qui est unique, monotone et non réutilisé ; et

un moyen pour maintenir une entrée de bloc de plan d'espace (b, e) pour chaque bloc de stockage de données utilisable pour indiquer les images instantanées utilisant le bloc, où b et e représentent des entrées de table d'index ou des numéros de version, b désigne la première image instantanée qui utilise le bloc et e désigne la dernière image instantanée qui utilise le bloc.

50

17. Système selon la revendication 16, comprenant en outre :

55

un moyen pour allouer un deuxième bloc ;

un moyen pour écrire des données dans le deuxième bloc ;

un moyen pour mettre à jour les pointeurs du système de fichiers qui pointaient vers le premier bloc pour qu'ils pointent vers le deuxième bloc ; et

un moyen pour répéter les étapes ci-dessus pour mettre à jour d'autres blocs.

EP 1 642 216 B1

18. Système selon la revendication 16, dans lequel l'entrée de la table d'index comprend un ou plusieurs attributs sélectionnés dans un groupe d'attributs comprenant un numéro de version (19), un état d'image (21), un horodatage (23), un pointeur de bloc racine (27) et un nom d'image (29).
- 5 19. Système selon la revendication 17, comprenant en outre un moyen pour lire le bloc de plan d'espace associé pour déterminer si le premier bloc est uniquement utilisé pour le système de fichiers actif (17) et, si tel est le cas, mettre à jour le premier bloc.
- 10 20. Système selon la revendication 19, dans lequel l'entrée de bloc de plan d'espace (a, 0) indique que le premier bloc est uniquement utilisé par le système de fichiers actif, où « a » représente le système de fichiers actif (17).
21. Système selon la revendication 16, comprenant en outre :
- 15 un moyen pour lire une première entrée de table d'index associée au système de fichiers actif;
un moyen pour situer (142) une deuxième entrée de table d'index qui est libre d'utilisation;
un moyen pour affecter (154) un numéro de version à la deuxième entrée de table d'index qui est plus élevé que celui de la première entrée de table d'index ;
un moyen pour copier (156) le pointeur de bloc racine depuis la première entrée de table d'index vers la deuxième entrée de table d'index ;
20 un moyen pour régler (148) l'état d'image de la première entrée de table d'index à une image instantanée en usage ; et
un moyen pour régler (152) l'état d'image de la deuxième entrée de table d'index au système de fichiers actif afin de créer une image instantanée.
- 25 22. Système selon la revendication 16, comprenant en outre :
- un moyen pour maintenir, dans les entrées de table d'index pour les images instantanées, un état d'image (21) qui indique si une entrée est effacée ou non ; et
un moyen pour marquer (76) l'état d'image comme effacé pour effacer une image instantanée.
- 30 23. Système selon la revendication 16, comprenant en outre :
- un moyen pour ordonner la table d'index suivant l'heure de la création de l'image instantanée ;
un moyen pour lire chaque entrée de bloc de plan d'espace pour déterminer si la première image instantanée a été effacée et, si tel est le cas, remplacer l'index b par l'index le plus précoce d'une image instantanée non effacée plus tardive ou du système de fichiers actif, déterminer si la dernière image instantanée a été effacée et, si tel est le cas, remplacer l'index e par le dernier index d'une image instantanée non effacée antérieure pour nettoyer les images instantanées effacées.
- 35 24. Système selon la revendication 23, comprenant en outre un moyen pour régler à nouveau l'état d'image à libre pour toutes les entrées de table d'index des images instantanées effacées lorsque toutes les entrées de bloc de plan d'espace ont été traitées.
- 40 25. Système selon la revendication 23, comprenant en outre un moyen pour régler l'entrée de bloc de plan d'espace pour indiquer que le bloc associé est libre si aucune image instantanée non effacée n'existe entre l'index b et l'index e.
- 45 26. Système selon la revendication 16, comprenant en outre :
- un moyen pour maintenir, dans les entrées de table d'index pour les images instantanées, un état d'image (21) qui indique si une entrée est obsolète ou non ;
un moyen pour créer une nouvelle image instantanée c correspondant à une précédente image instantanée p ;
un moyen pour créer un système de fichiers actif a correspondant à la précédente image instantanée p ; et
un moyen pour régler l'état d'image de toutes les images instantanées en usage avant l'image instantanée c et après l'image instantanée p à obsolète pour revenir à la précédente image instantanée p.
- 50 27. Système selon la revendication 26, comprenant :
- un moyen pour lire chaque entrée de bloc de plan d'espace pour déterminer :
- 55

EP 1 642 216 B1

(i) si la première image instantanée est l'image instantanée p ou une image instantanée antérieure à l'image instantanée p, et si la dernière image instantanée est antérieure à l'image instantanée c, remplacer l'index e par une valeur indiquant que le bloc est toujours utilisé par le système de fichiers actif,

5 (ii) si la première image instantanée est postérieure à l'image instantanée p, et antérieure à l'image instantanée c, régler l'entrée de bloc de plan d'espace pour indiquer que le bloc associé est libre ;

un moyen pour marquer l'état d'image de l'image instantanée c comme effacé ; et

10 un moyen pour régler à nouveau l'état d'image de toutes les entrées de table d'index obsolètes à libre pour nettoyer les images instantanées obsolètes après une opération de retour à l'image instantanée p.

28. Système selon la revendication 27, comprenant en outre un moyen pour répéter les fonctions de lecture et de marquage sur une liste d'images instantanées p et d'images instantanées c.

15 29. Système selon la revendication 28, dans lequel la liste comprend les images instantanées p non nettoyées revenues et les images instantanées c créées.

30. Système de gestion d'images instantanées selon la revendication 16, comprenant en outre un processeur pour maintenir lesdites multiples versions d'images instantanées et ledit système de fichiers actif, et comprenant en outre :

20 un espace utilisable pour stocker les versions d'images instantanées et le système de fichiers actif (17).

31. Système selon la revendication 30, dans lequel le processeur est programmé pour exécuter les étapes consistant à :

25 allouer un deuxième bloc ;

écrire les données dans le deuxième bloc ;

mettre à jour les pointeurs du système de fichiers qui pointaient vers le premier bloc pour qu'ils pointent vers le deuxième bloc ; et

répéter les étapes ci-dessus pour mettre à jour d'autres blocs.

30 32. Système selon la revendication 30, dans lequel l'entrée de la table d'index comprend un ou plusieurs attributs sélectionnés dans un groupe d'attributs comprenant un numéro de version (19), un état d'image (21), un horodatage (23), un pointeur de bloc racine (27) et un nom d'image (29).

35 33. Système selon la revendication 31, dans lequel le processeur lit le bloc de plan d'espace associé pour déterminer si le premier bloc est uniquement utilisé pour le système de fichiers actif (17) et, si tel est le cas, met à jour le premier bloc.

40 34. Système selon la revendication 30, dans lequel l'entrée de bloc de plan d'espace (a, 0) indique que le premier bloc est uniquement utilisé par le système de fichiers actif, où « a » représente le système de fichiers actif (17).

35. Système selon la revendication 30, dans lequel le processeur est programmé pour exécuter les étapes consistant à :

lire une première entrée de table d'index associée au système de fichiers actif;

situer (142) une deuxième entrée de table d'index qui est libre d'utilisation;

45 affecter (154) un numéro de version à la deuxième entrée de table d'index qui est plus élevé que celui de la première entrée de table d'index ;

copier (156) le pointeur de bloc racine depuis la première entrée de table d'index vers la deuxième entrée de table d'index ;

réglér (148) l'état d'image de la première entrée de table d'index à une image instantanée en usage ; et

50 réglér (152) l'état d'image de la deuxième entrée de table d'index au système de fichiers actif afin de créer une image instantanée.

36. Système selon la revendication 30, dans lequel le processeur est programmé pour exécuter les étapes consistant à :

55 maintenir, dans les entrées de table d'index pour les images instantanées, un état d'image (21) qui indique si une entrée est effacée ou non ; et

marquer (76) l'état d'image comme effacé pour effacer une image instantanée.

37. Système selon la revendication 30, dans lequel le processeur est programmé pour exécuter les étapes consistant à :

ordonner la table d'index (195, 196) suivant l'heure de la création de l'image instantanée ;
lire chaque entrée de bloc de plan d'espace pour déterminer si la première image instantanée a été effacée et,
si tel est le cas, remplacer l'index b par l'index le plus précoce d'une image instantanée non effacée plus tardive
ou du système de fichiers actif, déterminer si la dernière image instantanée a été effacée et, si tel est le cas,
remplacer l'index e par le dernier index d'une image instantanée non effacée antérieure pour nettoyer les images
instantanées effacées.

5

38. Système selon la revendication 37, dans lequel le processeur est programmé pour régler à nouveau l'état d'image
à libre pour toutes les entrées de table d'index des images instantanées effacées lorsque toutes les entrées de bloc
de plan d'espace ont été traitées.

10

39. Système selon la revendication 37, dans lequel le processeur est programmé pour régler l'entrée de bloc de plan
d'espace pour indiquer que le bloc associé est libre si aucune image instantanée non effacée n'existe entre l'index
b et l'index e.

15

40. Système selon la revendication 39, dans lequel le processeur est programmé pour régler à nouveau l'état d'image
à libre pour toutes les entrées de table d'index des images instantanées effacées lorsque toutes les entrées de bloc
de plan d'espace ont été traitées.

20

41. Système selon la revendication 30, dans lequel le processeur est programmé pour exécuter les étapes consistant à :

maintenir, dans les entrées de table d'index pour les images instantanées, un état d'image qui indique si une
entrée est obsolète ou non ;
créer une nouvelle image instantanée c en effectuant une copie d'une précédente image instantanée p ;
créer un système de fichiers actif a en effectuant une copie de la précédente image instantanée p ; et
régler l'état d'image de toutes les images instantanées en usage avant l'image instantanée c et après l'image
instantanée p à obsolète pour revenir à une précédente image instantanée p.

25

30

35

40

45

50

55

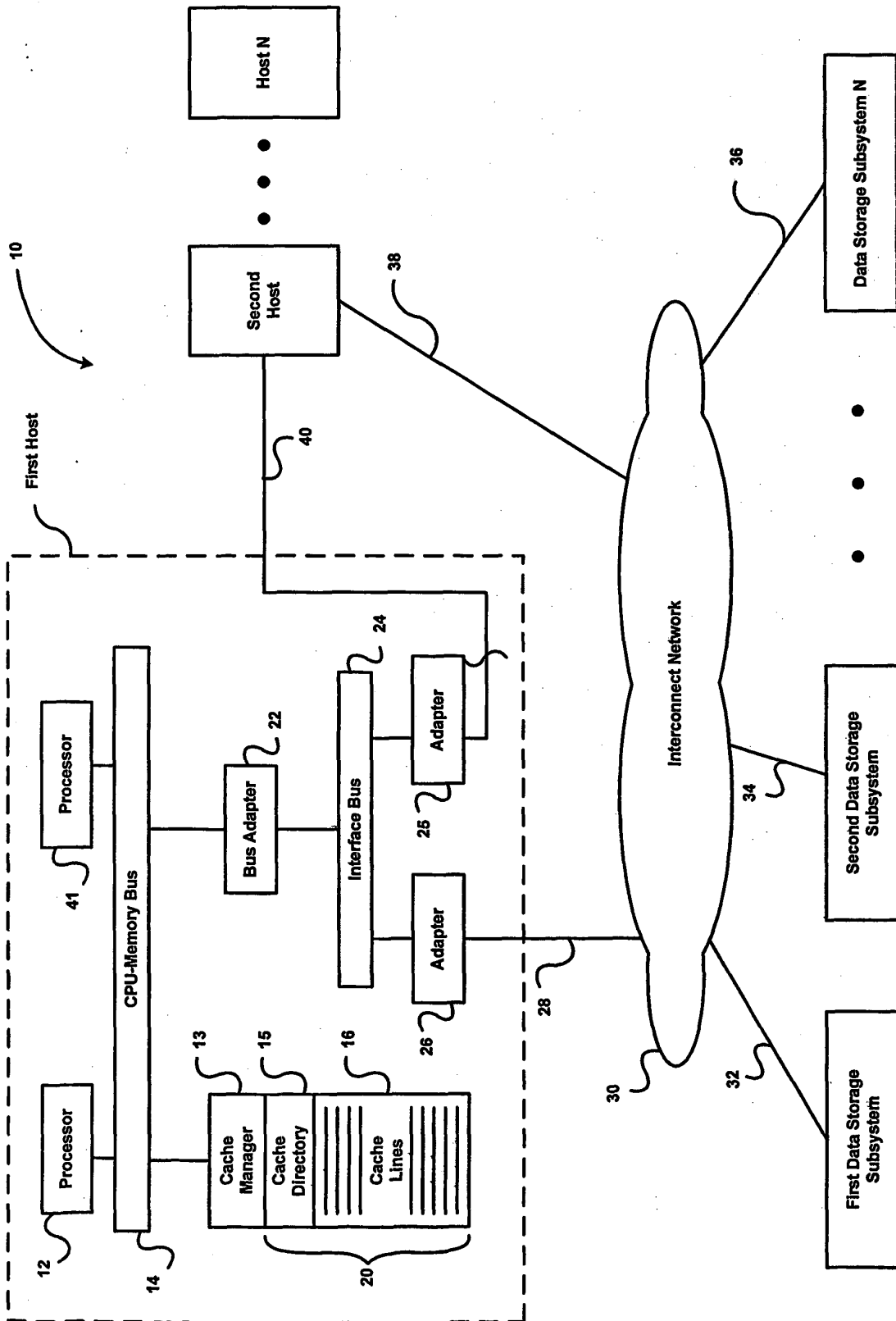


FIGURE 1

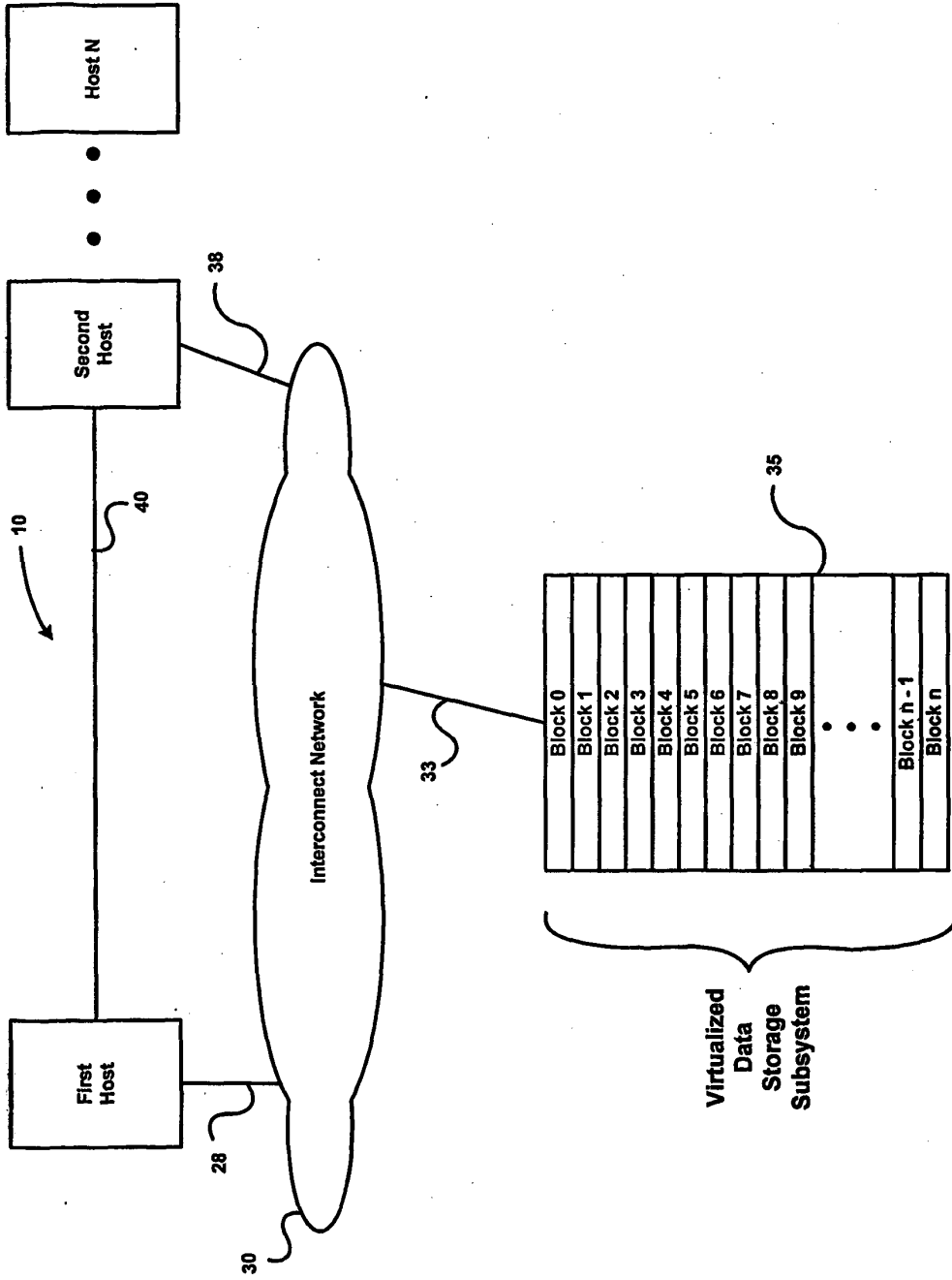


FIGURE 2

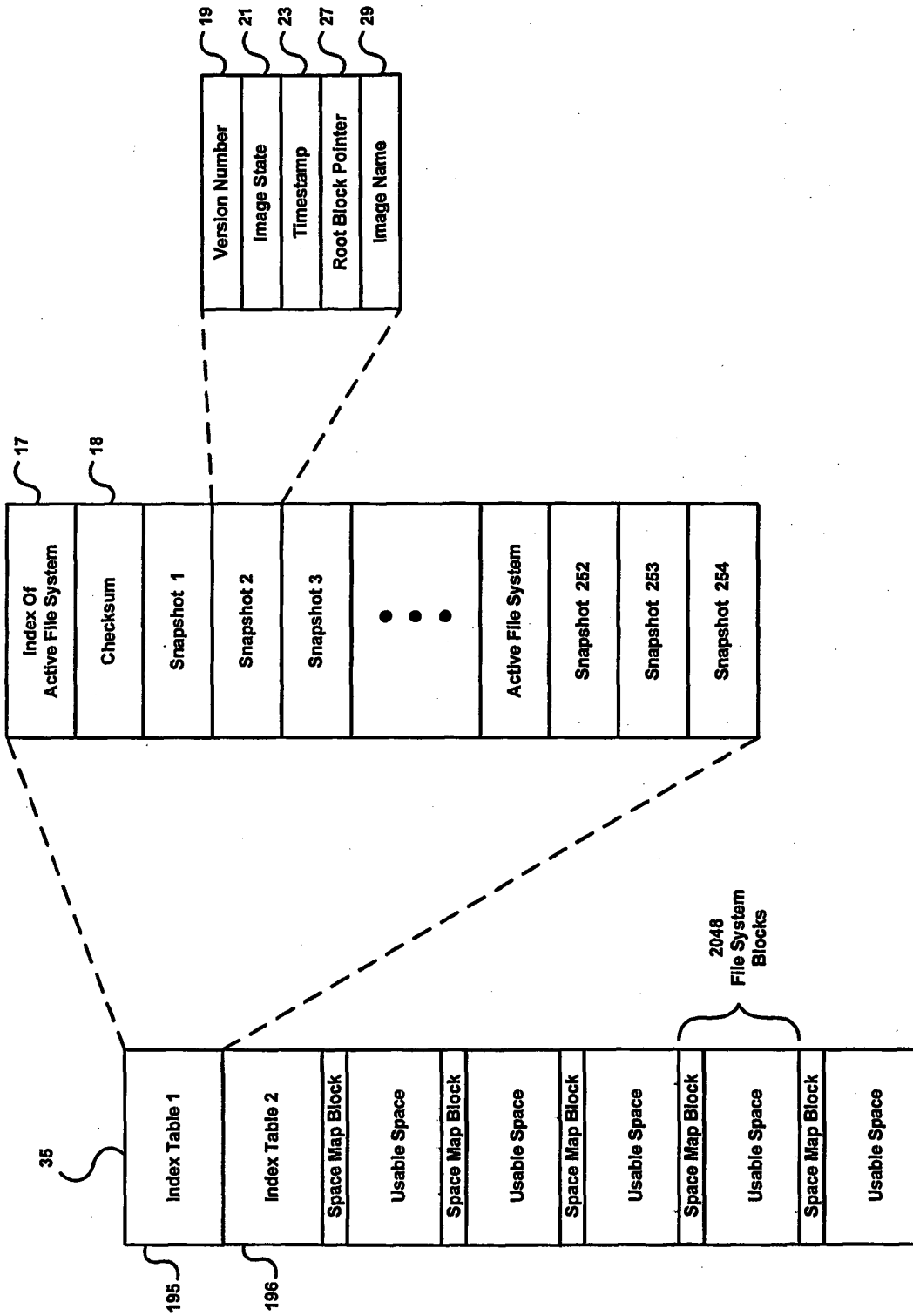


FIGURE 3

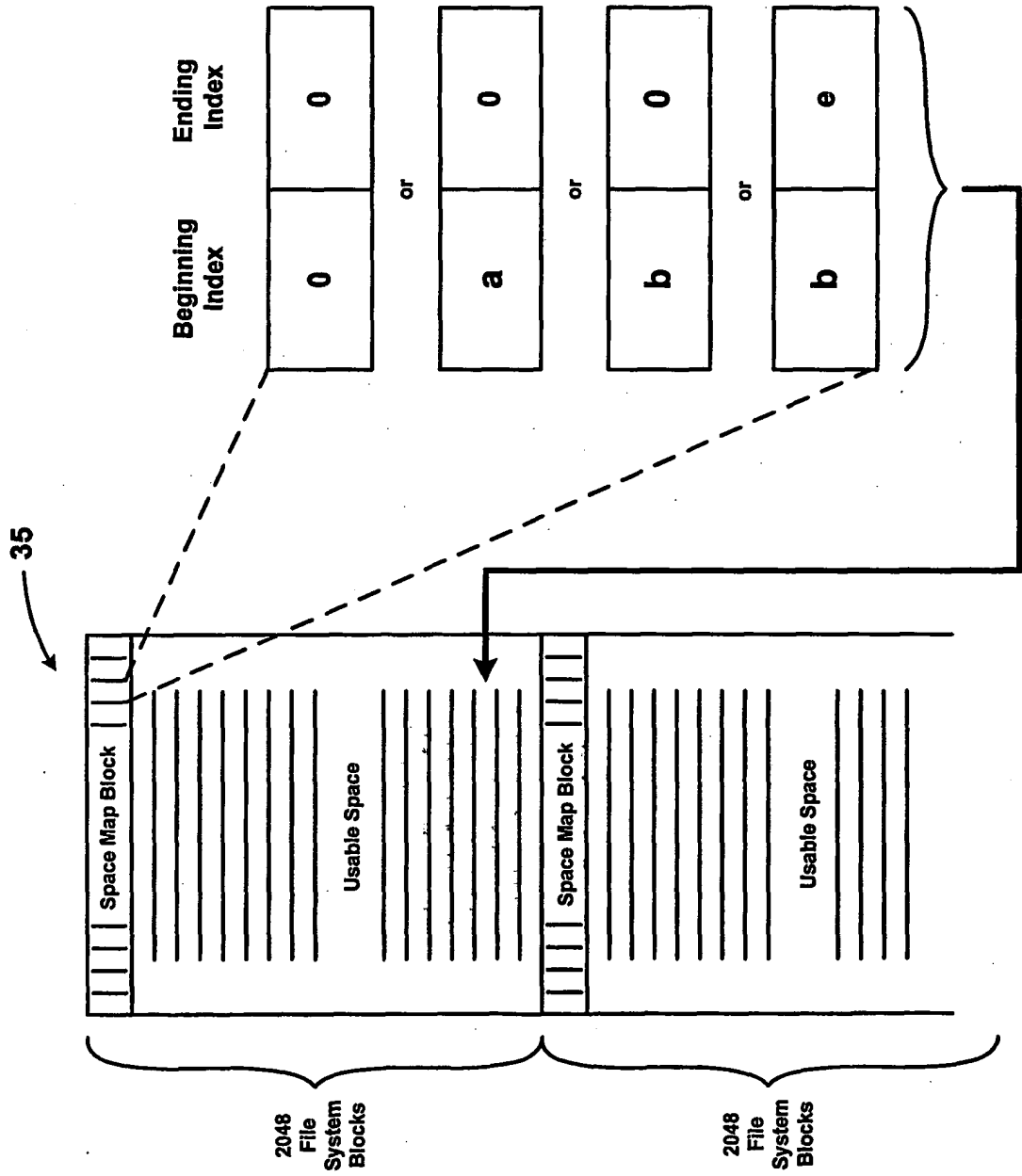


FIGURE 4

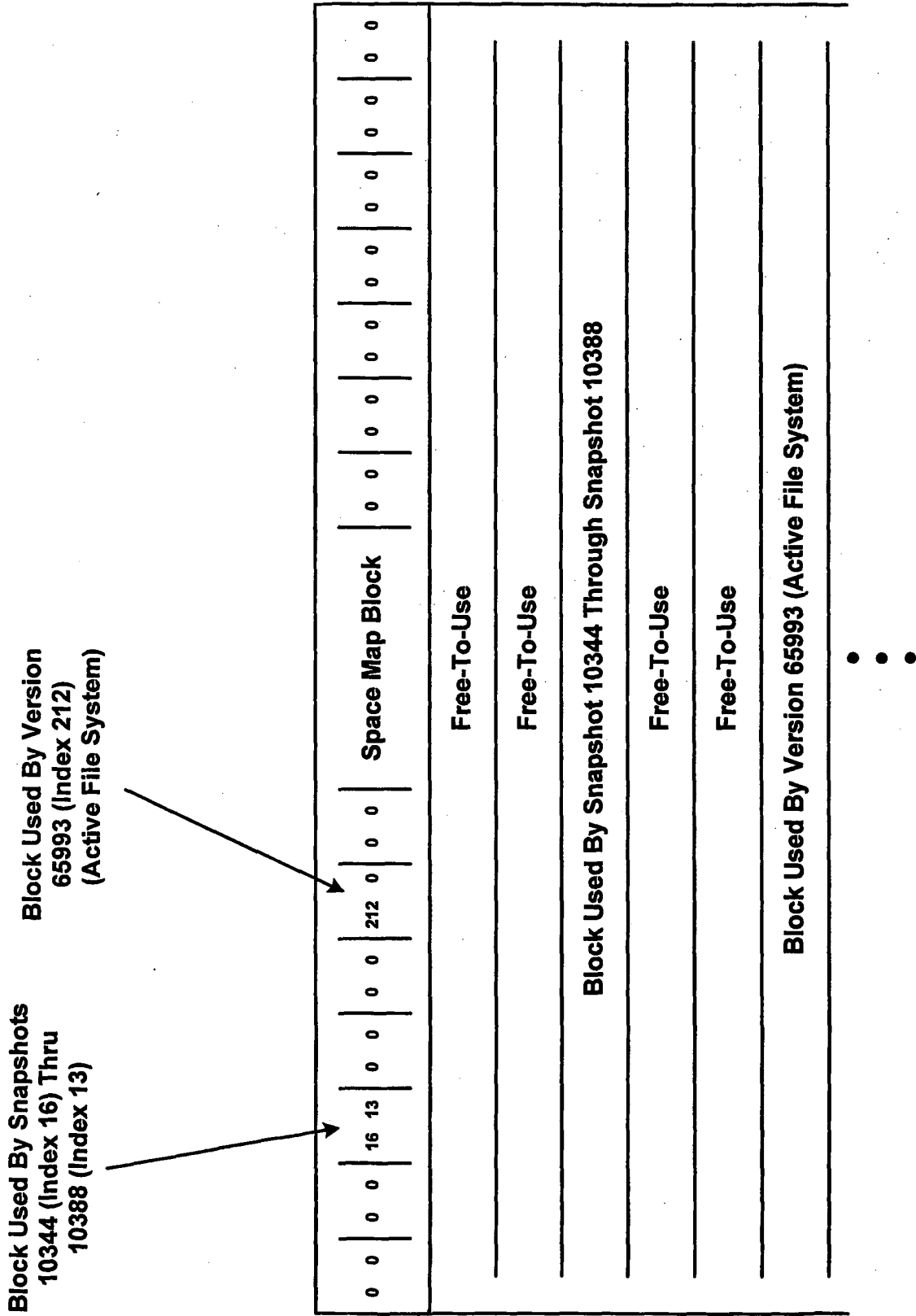


FIGURE 5

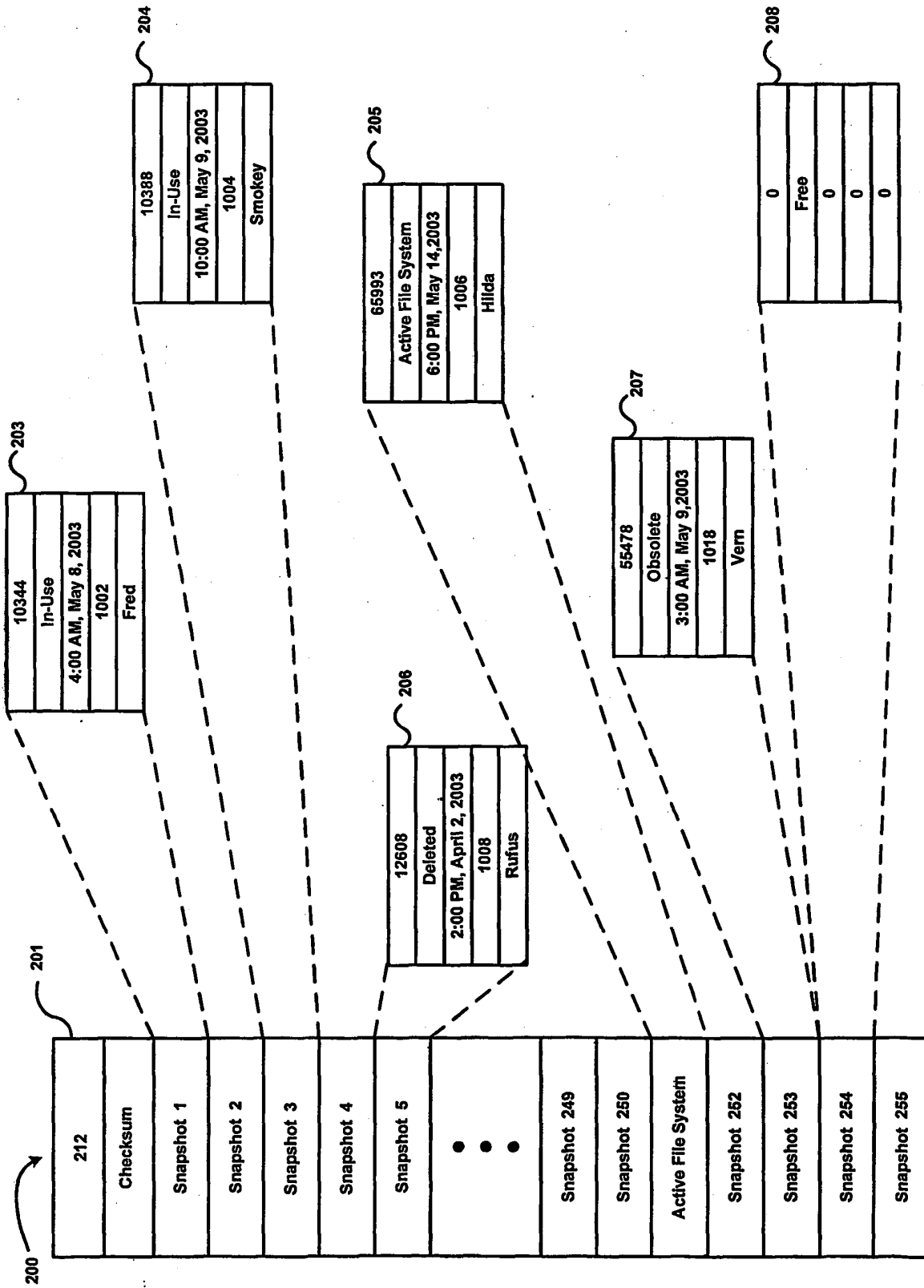


FIGURE 6

Create A File System

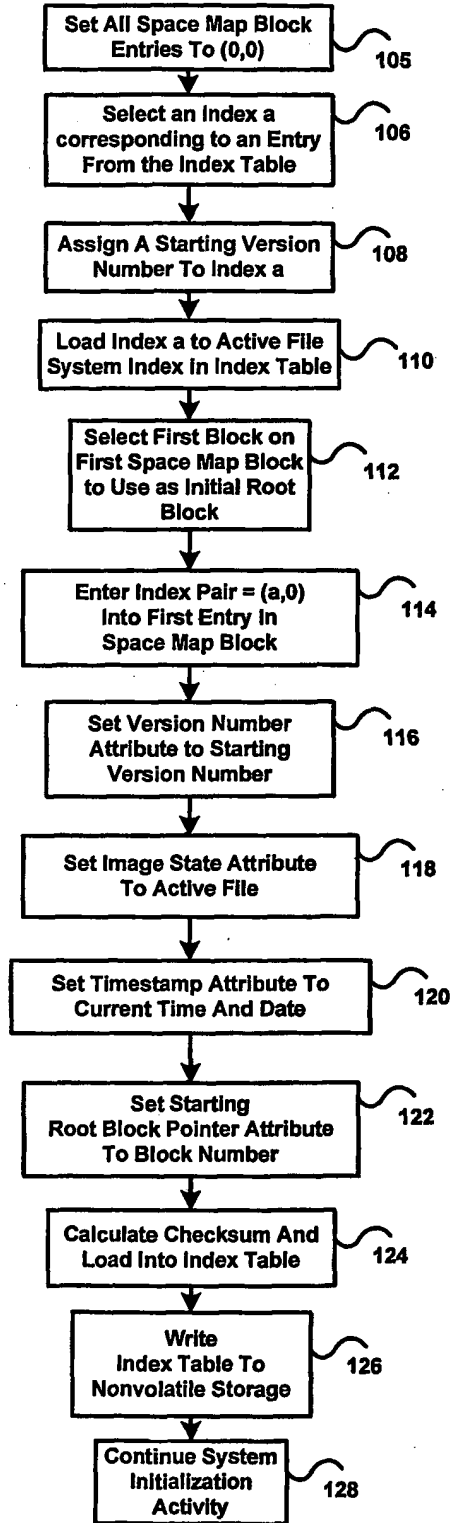


FIGURE 7

Block Management

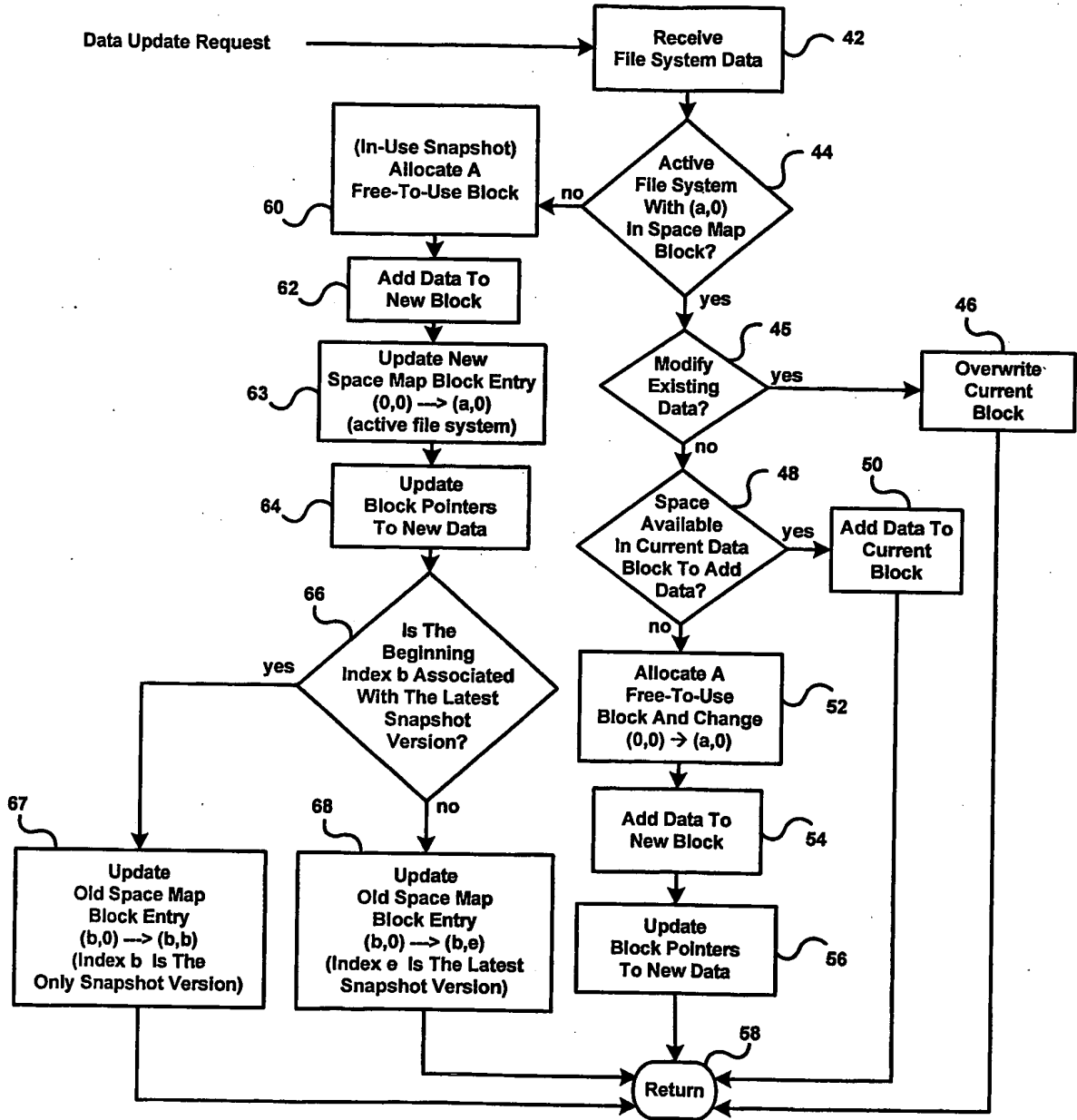


FIGURE 8

Create A File System Snapshot

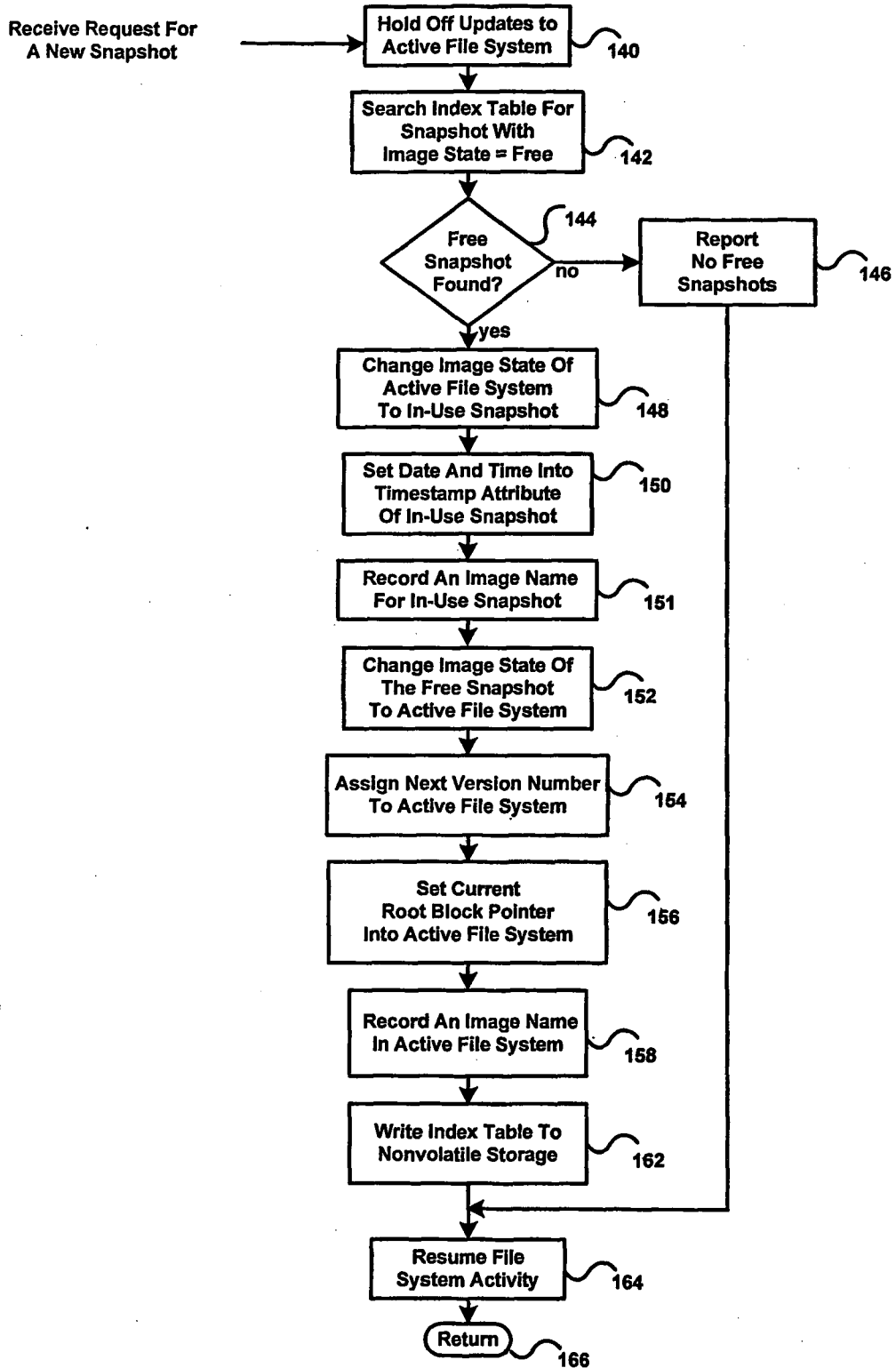


FIGURE 9

Revert To An Earlier Snapshot Version

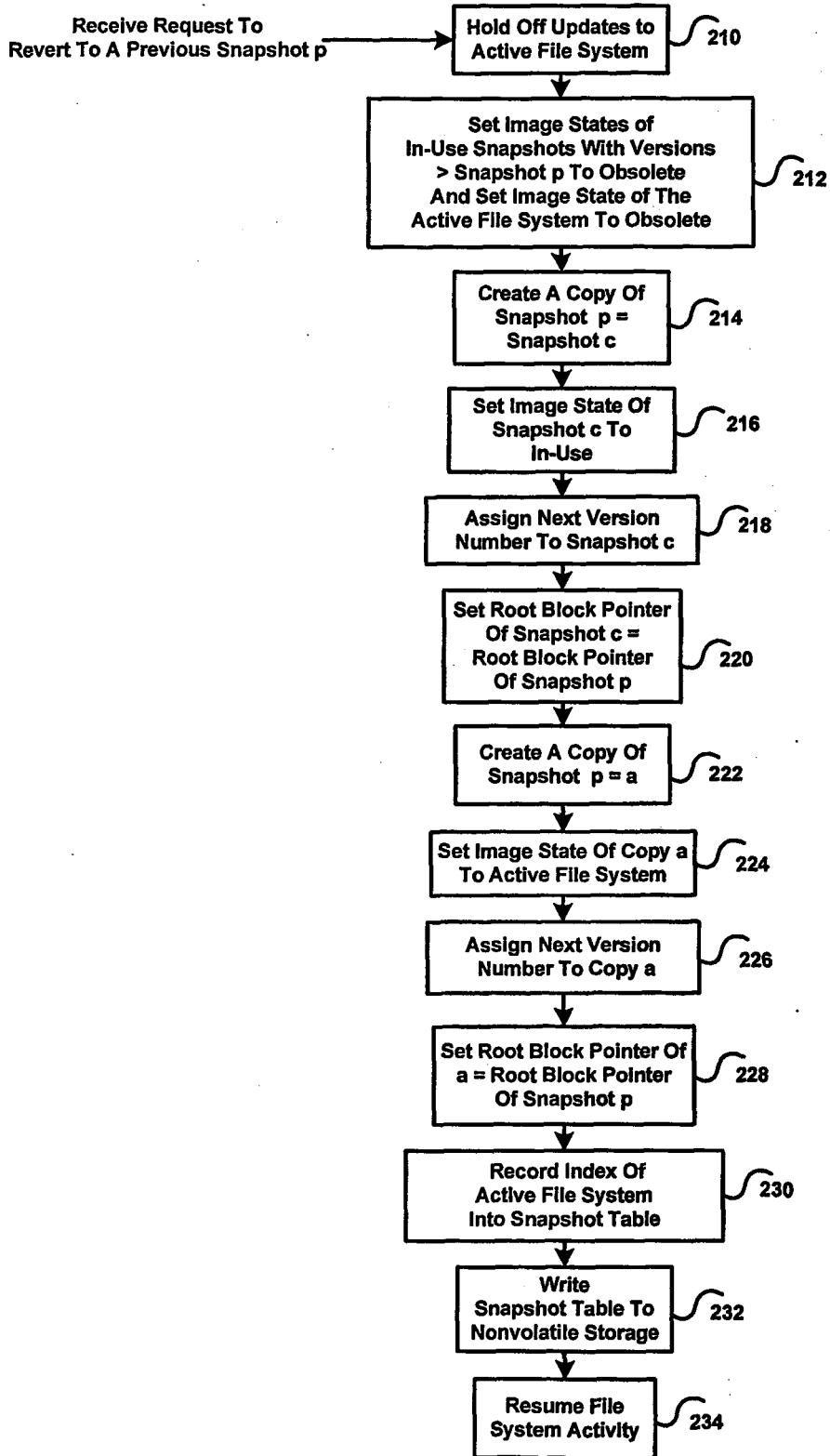


FIGURE 10

Detect And Clean Unneeded Snapshots

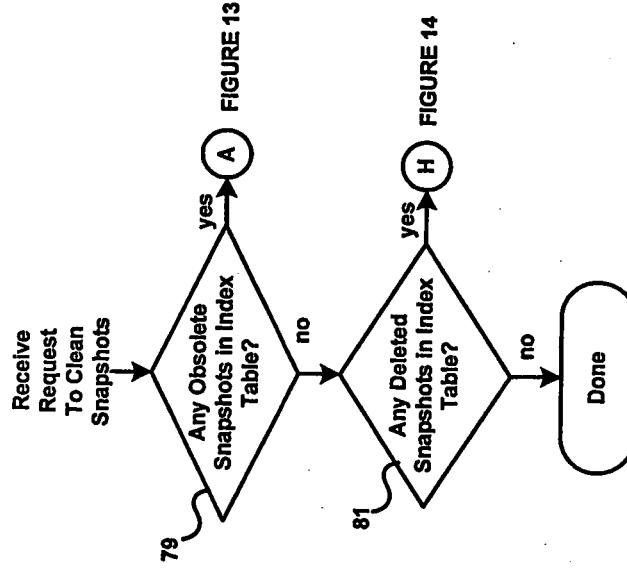


FIGURE 11b

Delete A Snapshot

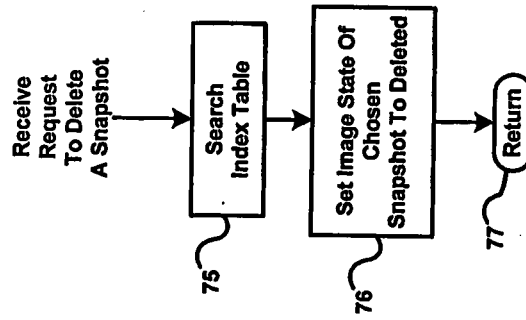


FIGURE 11a

Remove Obsolete Snapshots from Index Table

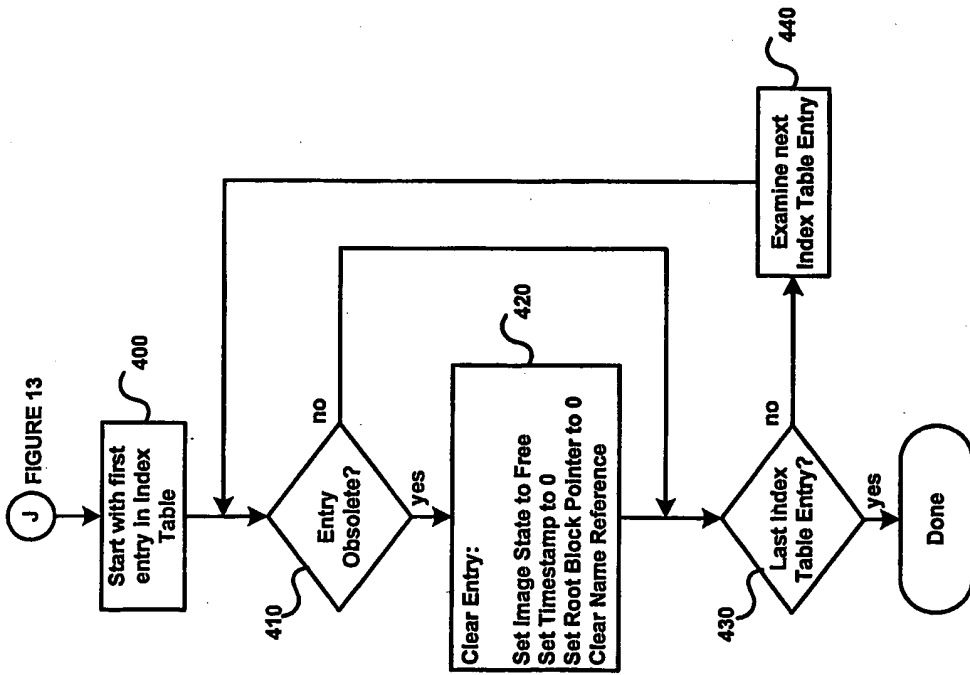


FIGURE 12a

Remove Deleted Snapshots from Index Table

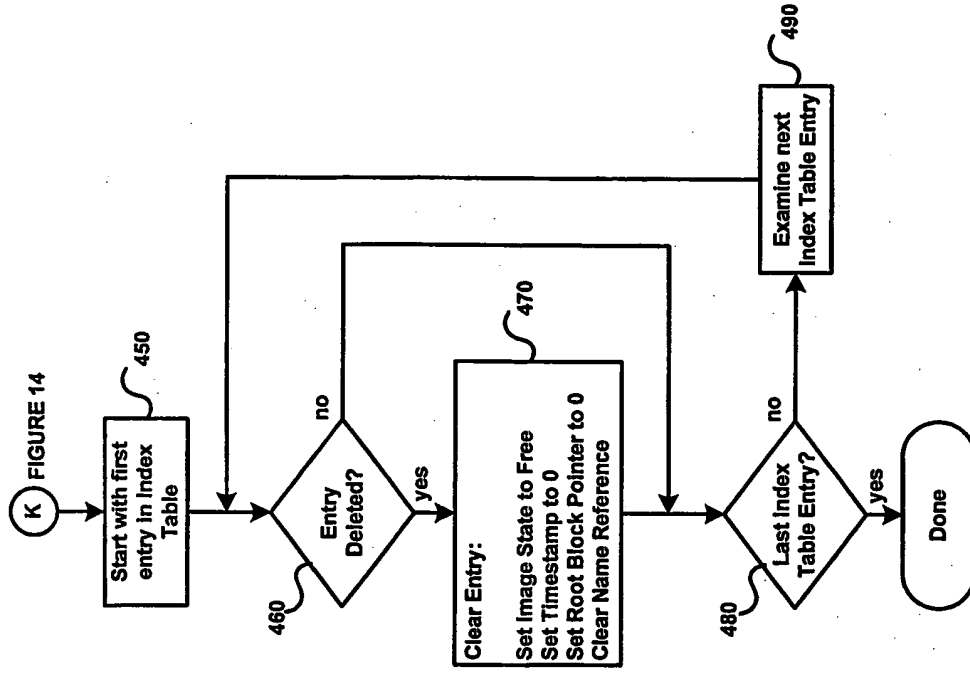


FIGURE 12b

Clean Obsolete Snapshots

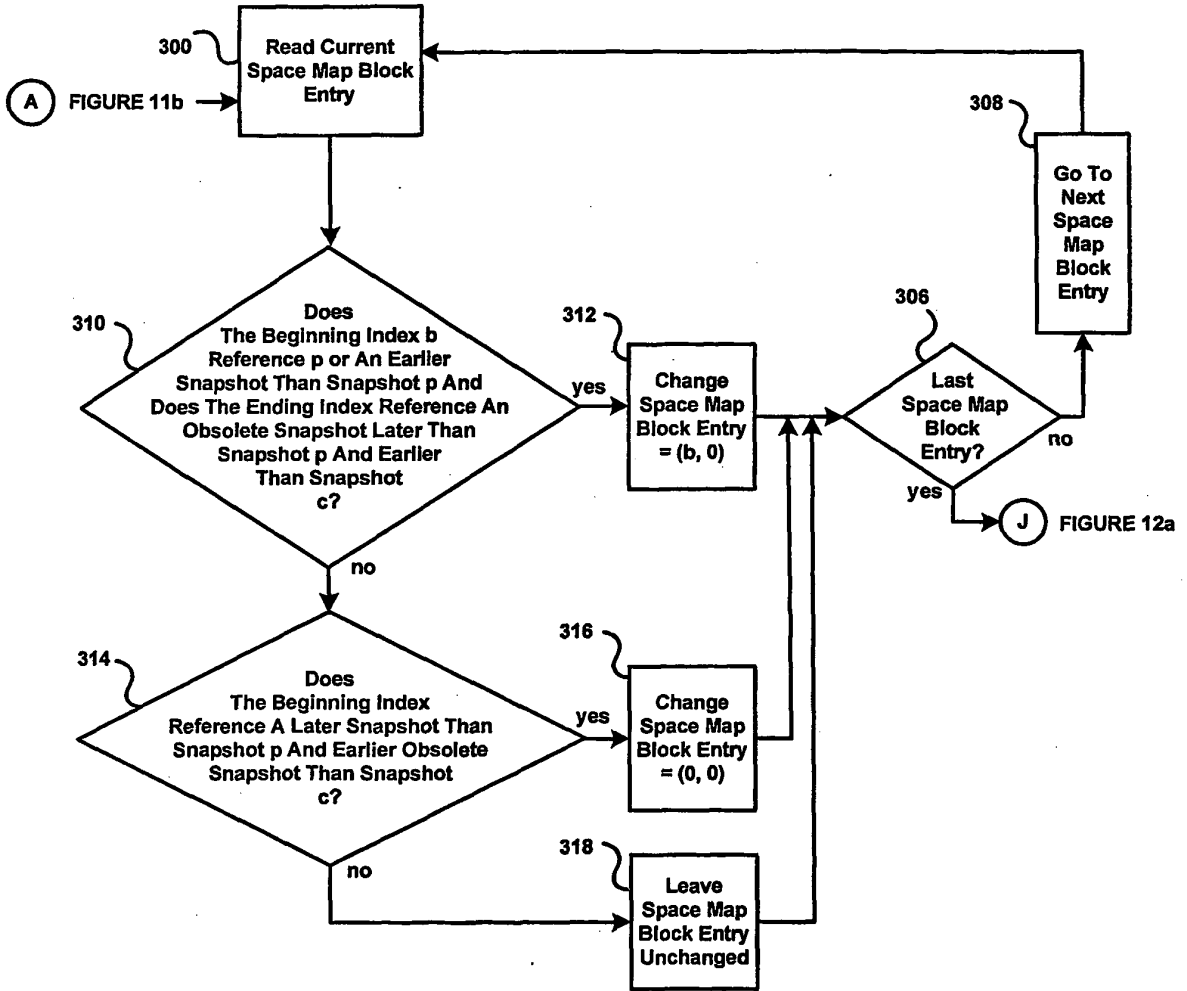


FIGURE 13

Clean Deleted Snapshots

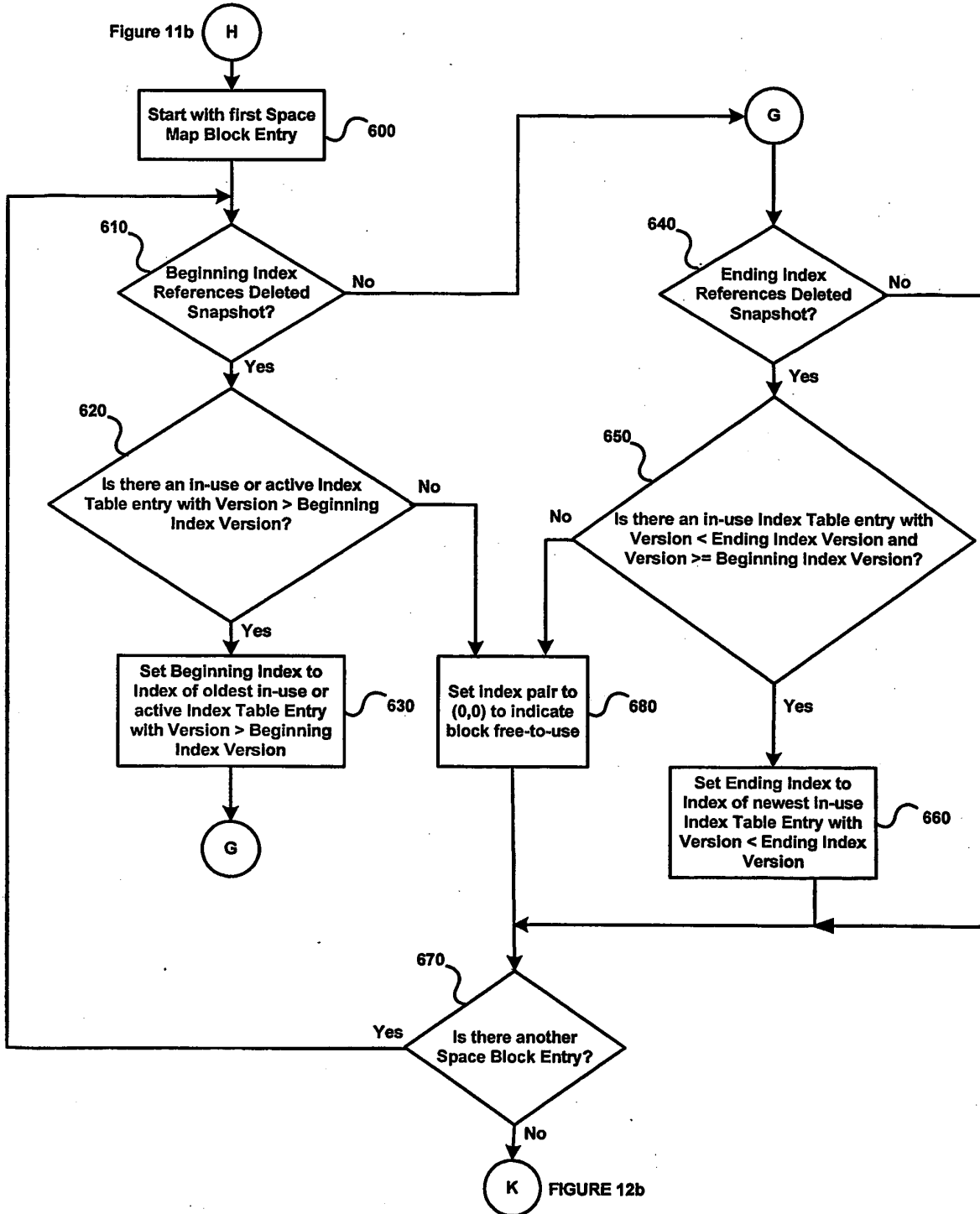
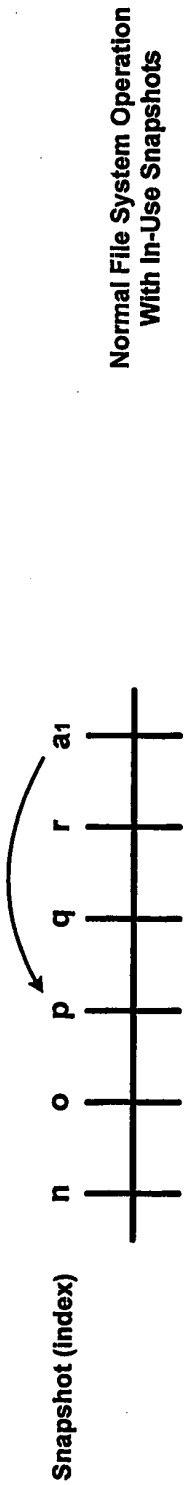
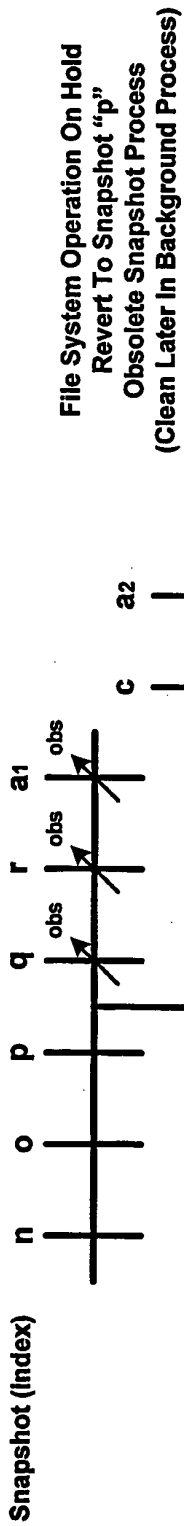


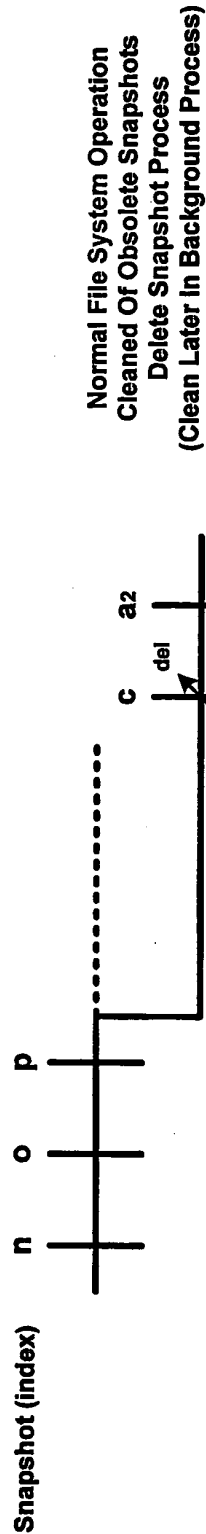
FIGURE 14



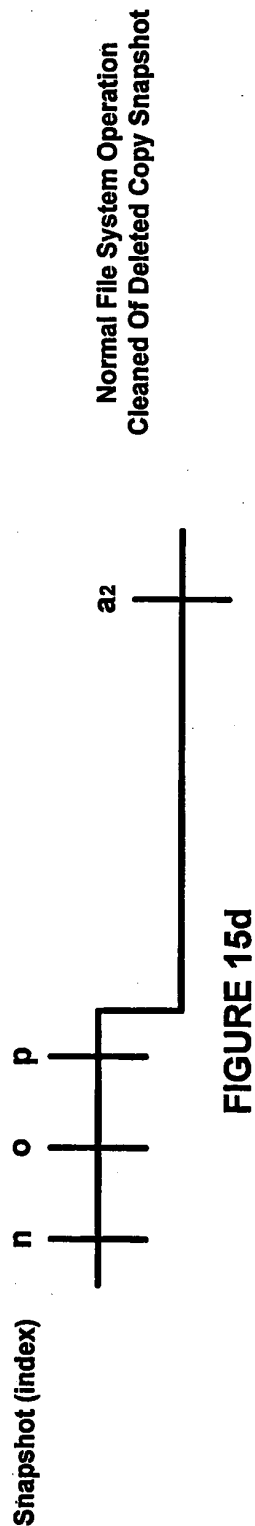
Normal File System Operation
With In-Use Snapshots



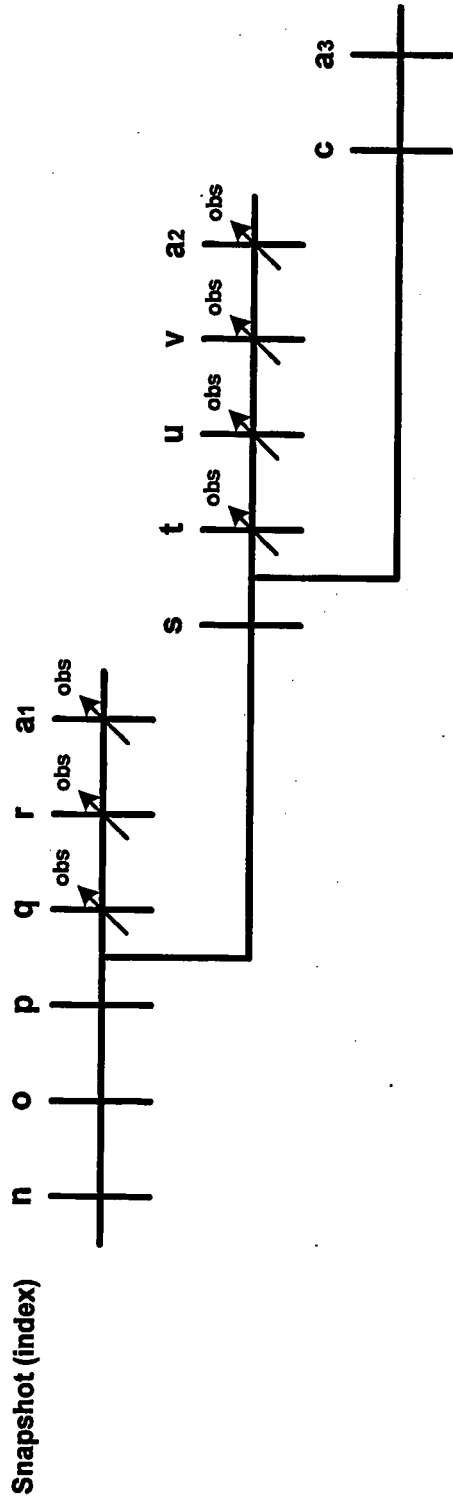
File System Operation On Hold
Revert To Snapshot "p"
Obsolete Snapshot Process
(Clean Later In Background Process)



Normal File System Operation
Cleaned Of Obsolete Snapshots
Delete Snapshot Process
(Clean Later In Background Process)



Normal File System Operation
Cleaned Of Deleted Copy Snapshot



Multiple File System Reversions
To Snapshot "p" And To Snapshot "s"
Prior To Completing Background Cleaning Process

FIGURE 16

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 35479703 A [0003]
- US 44034703 A [0004]
- US 2002091670 A [0009]
- US 26460302 A [0022]

Non-patent literature cited in the description

- **Hennessy ; Patterson.** Computer Architecture: A Quantitative Approach (2003), and Patterson and Hennessy. *Computer organization and Design: The Hardware /Software Interface*, 1998 [0015]
- **Tanenbaum.** *Modem Operating Systems*, 2001 [0018]
- **Clark.** *IP SANs: A Guide to iSCSI, iFCP, and FCIP Protocols for Storage Area Networks*, 2002 [0021]
- **Clark.** *Designing Storage Area Networks*, 1999 [0021]