US 20110320707A1

(19) United States

(12) Patent Application Publication

(10) Pub. No.: US 2011/0320707 A1

(43) Pub. Date: Dec. 29, 2011

Usami

(54) **STORAGE APPARATUS AND STORAGE MANAGEMENT METHOD**

(75) Inventor: **Kazuhiro Usami**, Odawara (JP)

(73) Assignees: **HITACHI COMPUTER PERIPHERALS CO., LTD.**, Kanagawa-ken, (JP); **HITACHI, LTD.**, Chiyoda-ku, Tokyo (JP)

(21) Appl. No.: **12/811,661**

(22) PCT Filed: **Jun. 24, 2010**

(86) PCT No.: **PCT/JP2010/004211**

§ 371 (c)(1),
(2), (4) Date: **Jul. 2, 2010**

Publication Classification

(51) **Int. Cl.**
G06F 12/00 (2006.01)
G06F 12/16 (2006.01)

(52) **U.S. Cl.** **711/114**; 711/111; 711/E12.001; 711/E12.103

(57) **ABSTRACT**

The performance to transfer data to external storage media in thin provisioning is enhanced.

A storage apparatus is connected via a network to a host system writing of data and to a tape management apparatus for managing tape storage media for storing the data. The storage apparatus is designed so that data, for which a write request is issued by the host system, is stored in a virtual volume; a predetermined area in a plurality of storage devices is allocated to a data storage area in the virtual volume storing the data; and if the virtual volume is a backup object to be backed up to the tape storage media, a predetermined area in a specified storage device, from among the plurality of storage devices, is allocated to the data storage area in the virtual volume; and if the virtual volume is not a backup object to be backed up to the tape storage media, a predetermined area in an unspecified storage device, from among the plurality of storage devices, is allocated to the data storage area in the virtual volume.

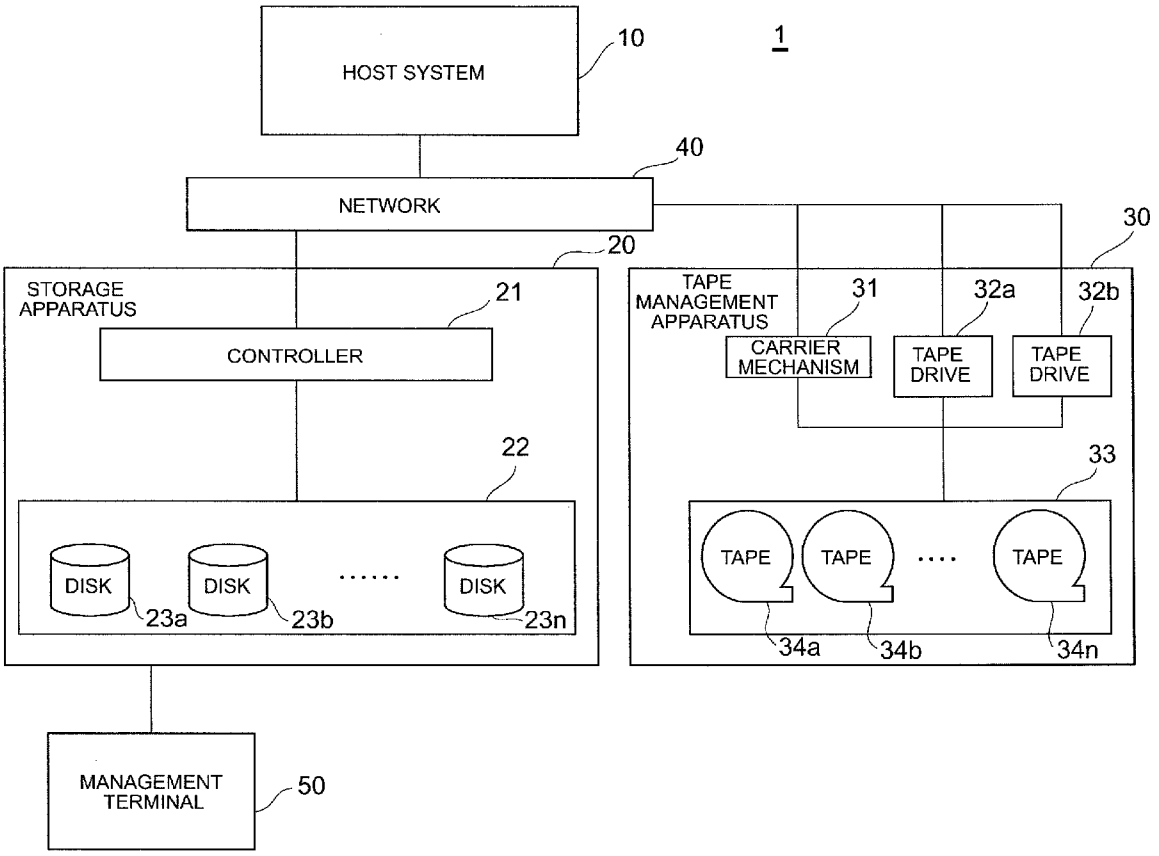


FIG. 1

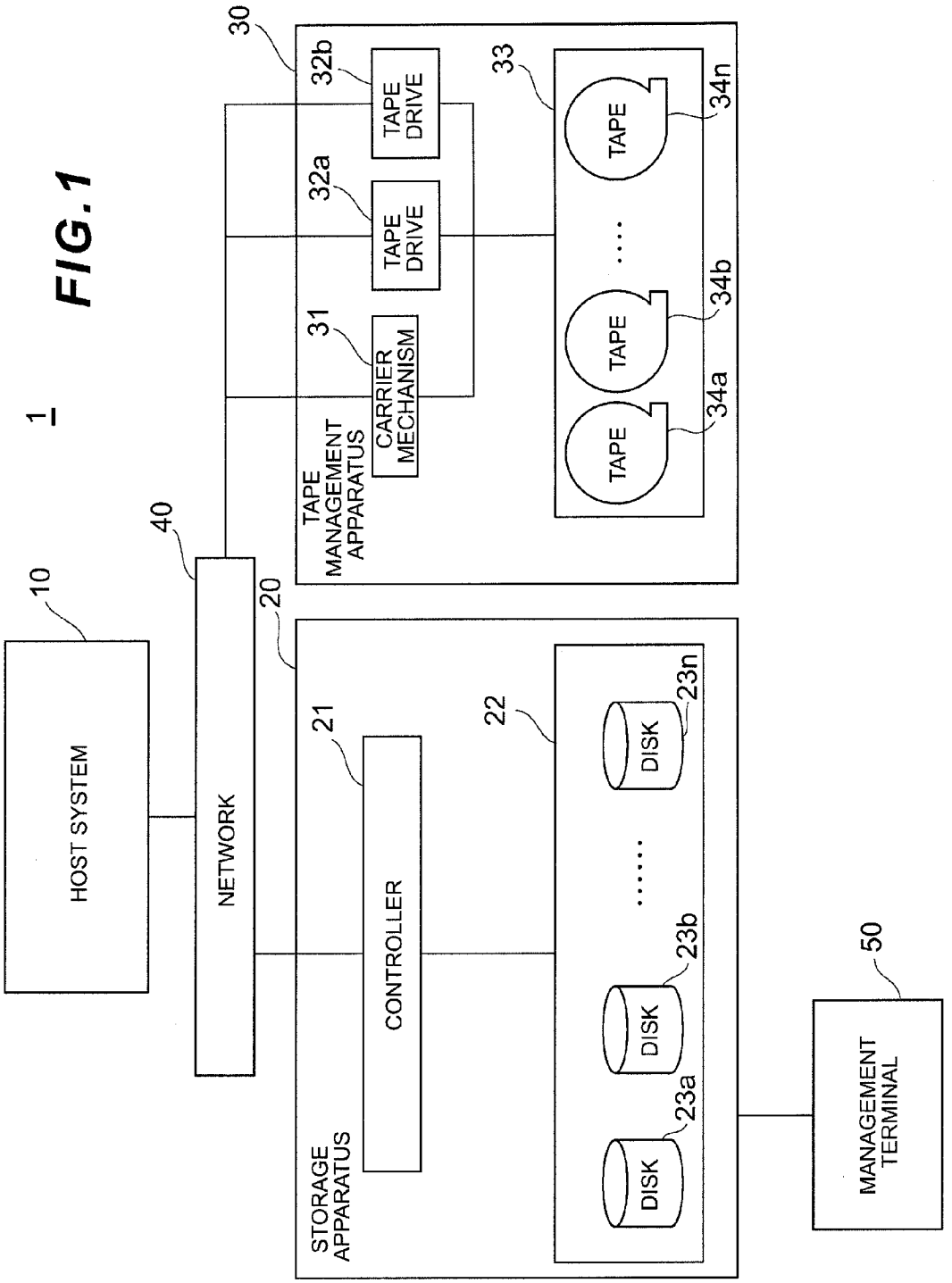

1
HOST SYSTEM
10
NETWORK
40
STORAGE APPARATUS
20
CONTROLLER
21
22
DISK
23a
DISK
23b
DISK
23n
MANAGEMENT TERMINAL
50
TAPE MANAGEMENT APPARATUS
30
CARRIER MECHANISM
31
TAPE DRIVE
32a
TAPE DRIVE
32b
33
TAPE
34a
TAPE
34b
TAPE
34n

FIG. 2

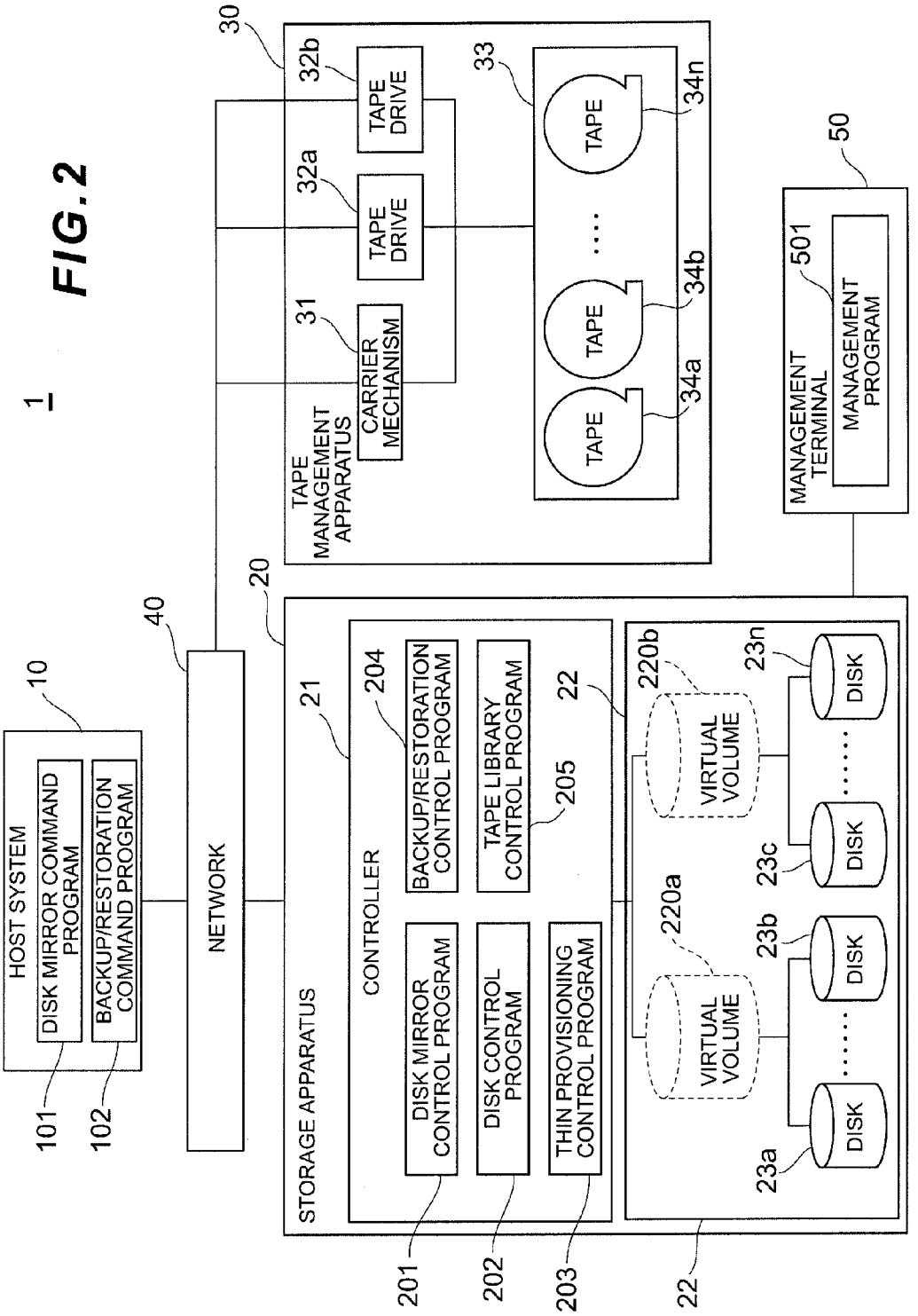

1
HOST SYSTEM
10
DISK MIRROR COMMAND PROGRAM
BACKUP/RESTORATION COMMAND PROGRAM
101
102
NETWORK
40
STORAGE APPARATUS
20
CONTROLLER
21
DISK MIRROR CONTROL PROGRAM
BACKUP/RESTORATION CONTROL PROGRAM
204
DISK CONTROL PROGRAM
TAPE LIBRARY CONTROL PROGRAM
205
THIN PROVISIONING CONTROL PROGRAM
22
201
202
203
VIRTUAL VOLUME
220a
220b
DISK
23a
23b
23c
23n
TAPE MANAGEMENT APPARATUS
30
CARRIER MECHANISM
31
TAPE DRIVE
32a
32b
TAPE
33
34a
34b
34n
MANAGEMENT TERMINAL
50
MANAGEMENT PROGRAM
501

FIG. 3

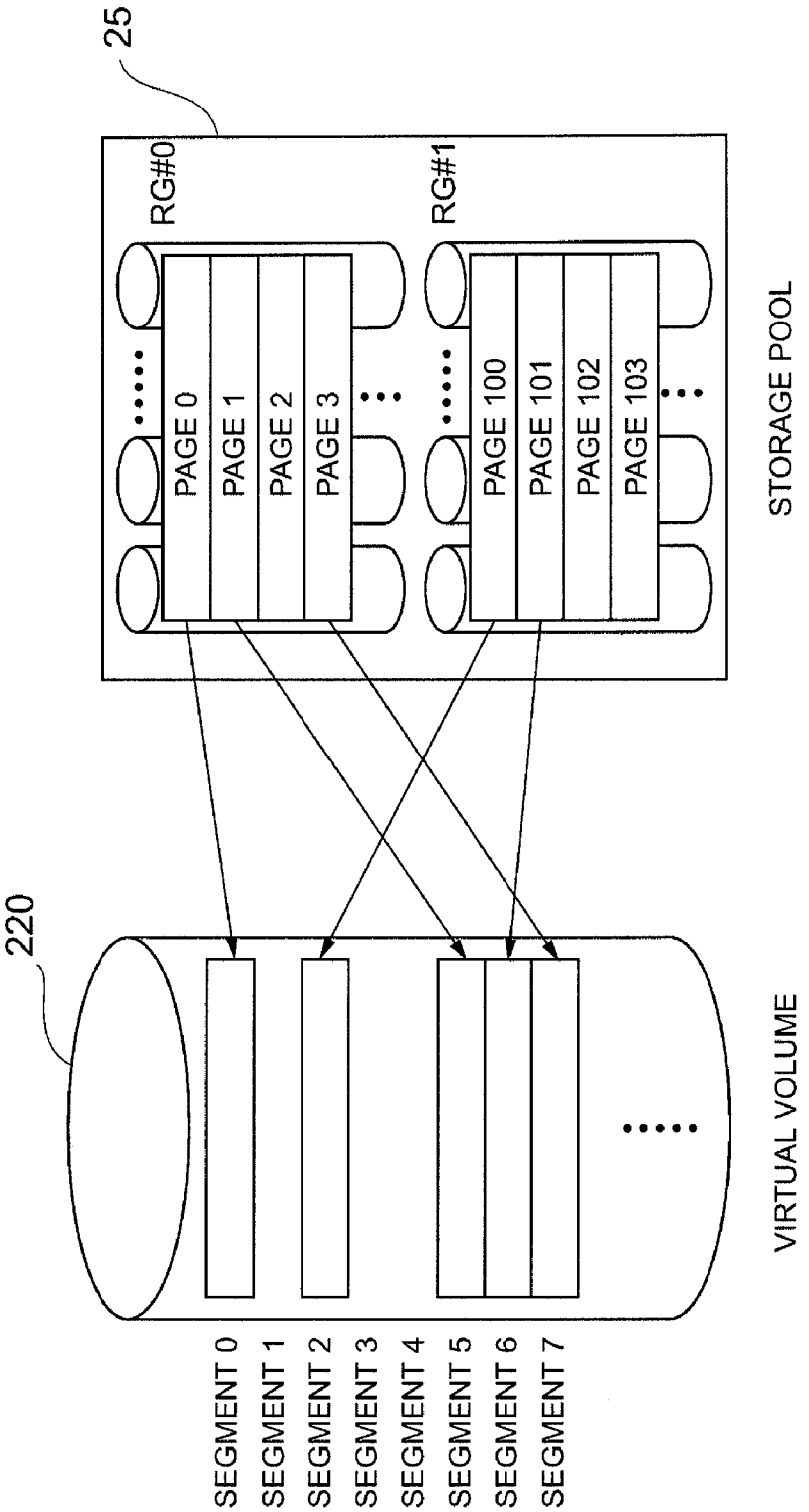
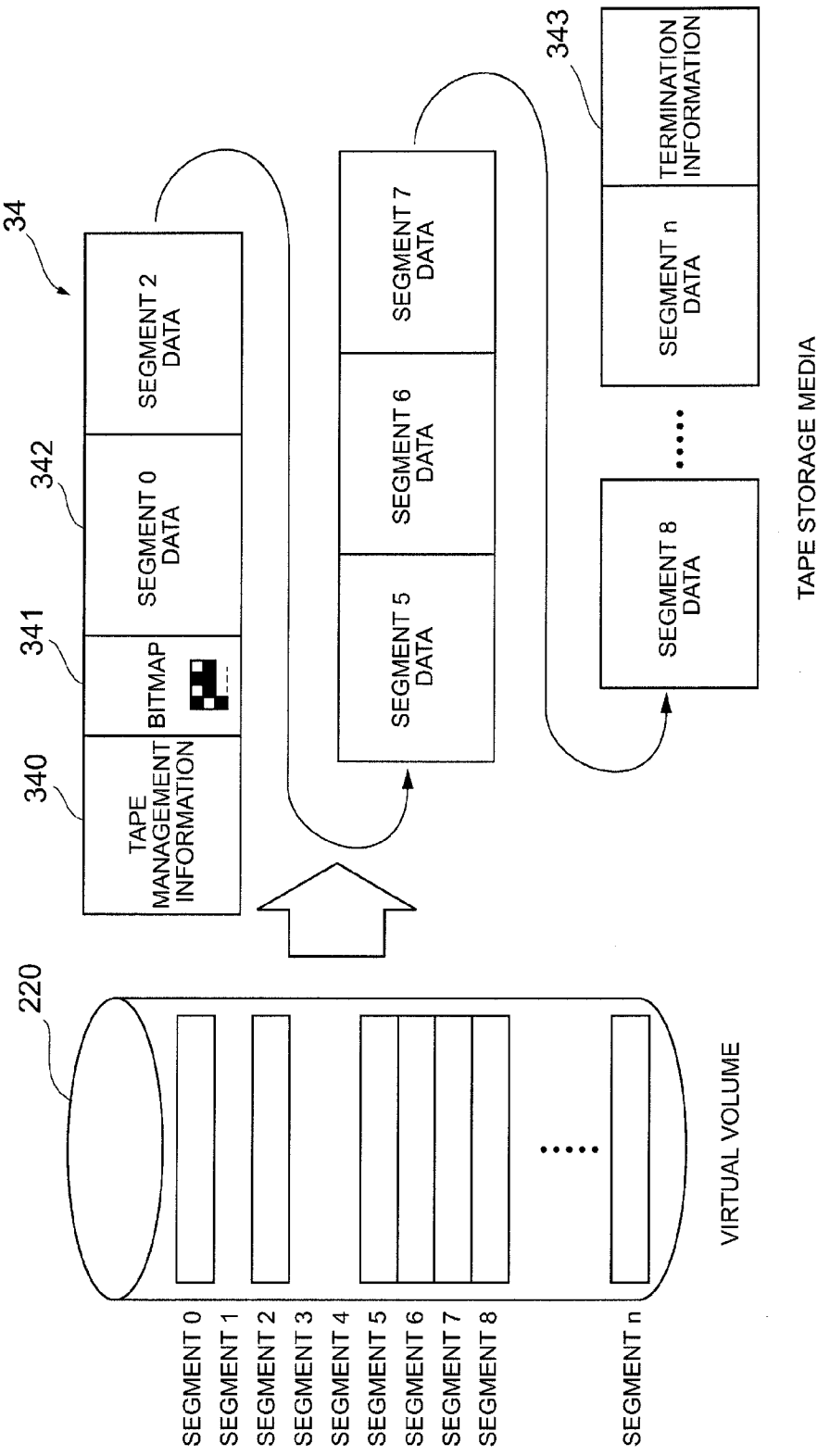

25
RG#0
RG#1
PAGE 0
PAGE 1
PAGE 2
PAGE 3
PAGE 100
PAGE 101
PAGE 102
PAGE 103
STORAGE POOL
220
SEGMENT 0
SEGMENT 1
SEGMENT 2
SEGMENT 3
SEGMENT 4
SEGMENT 5
SEGMENT 6
SEGMENT 7
VIRTUAL VOLUME

FIG. 4


34
340
TAPE MANAGEMENT INFORMATION
341
BITMAP
342
SEGMENT 0 DATA
SEGMENT 2 DATA
SEGMENT 5 DATA
SEGMENT 6 DATA
SEGMENT 7 DATA
SEGMENT 8 DATA
SEGMENT n DATA
343
TERMINATION INFORMATION
TAPE STORAGE MEDIA
220
SEGMENT 0
SEGMENT 1
SEGMENT 2
SEGMENT 3
SEGMENT 4
SEGMENT 5
SEGMENT 6
SEGMENT 7
SEGMENT 8
SEGMENT n
VIRTUAL VOLUME

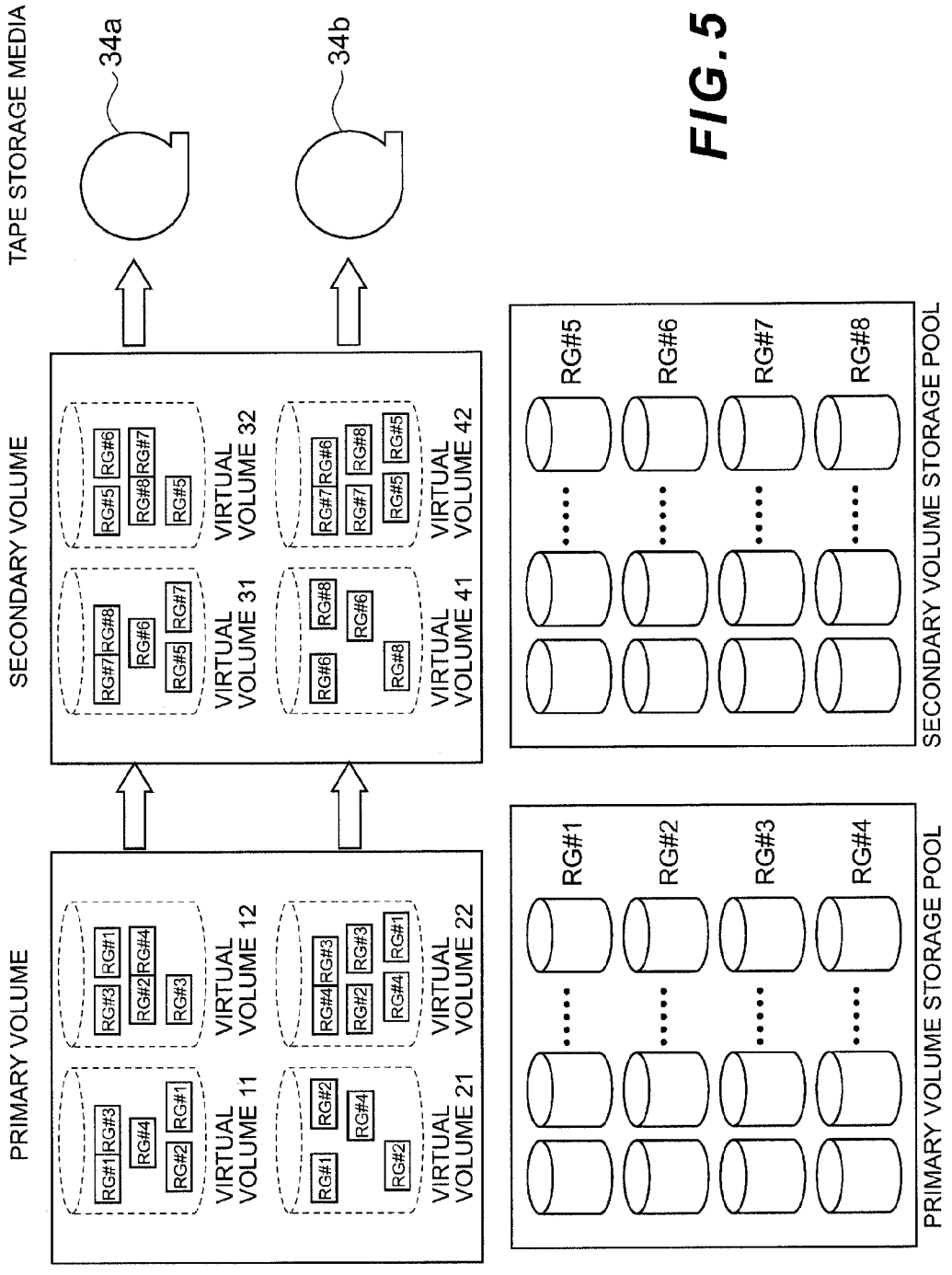

TAPE STORAGE MEDIA
34a
34b
SECONDARY VOLUME
PRIMARY VOLUME
VIRTUAL VOLUME 11
VIRTUAL VOLUME 12
VIRTUAL VOLUME 21
VIRTUAL VOLUME 22
VIRTUAL VOLUME 31
VIRTUAL VOLUME 32
VIRTUAL VOLUME 41
VIRTUAL VOLUME 42
RG#1
RG#2
RG#3
RG#4
RG#5
RG#6
RG#7
RG#8
PRIMARY VOLUME STORAGE POOL
SECONDARY VOLUME STORAGE POOL

FIG. 5

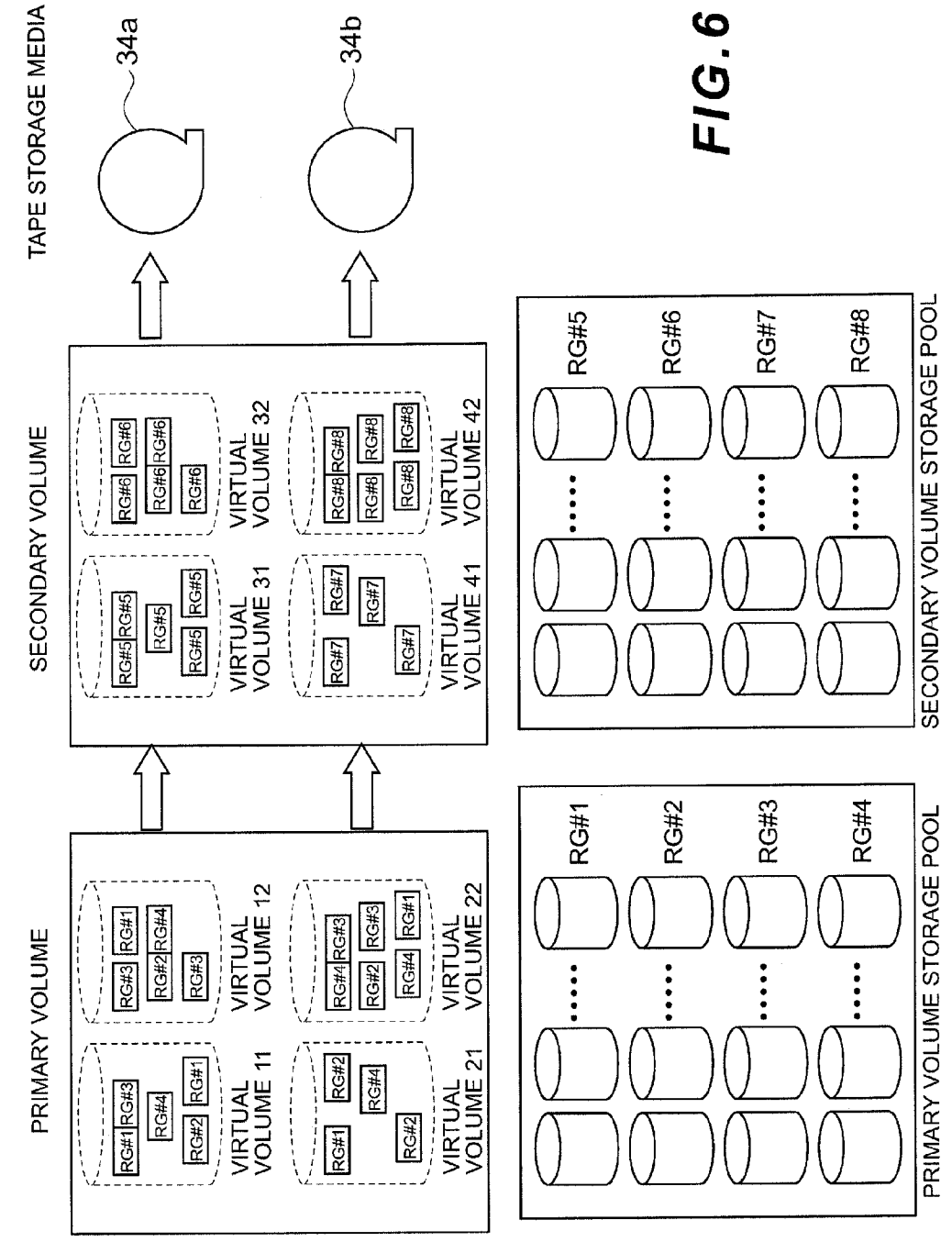

TAPE STORAGE MEDIA
34a
34b
SECONDARY VOLUME
VIRTUAL VOLUME 31
VIRTUAL VOLUME 32
VIRTUAL VOLUME 41
VIRTUAL VOLUME 42
PRIMARY VOLUME
VIRTUAL VOLUME 11
VIRTUAL VOLUME 12
VIRTUAL VOLUME 21
VIRTUAL VOLUME 22
RG#1
RG#2
RG#3
RG#4
RG#5
RG#6
RG#7
RG#8
SECONDARY VOLUME STORAGE POOL
PRIMARY VOLUME STORAGE POOL

FIG. 6

FIG. 7

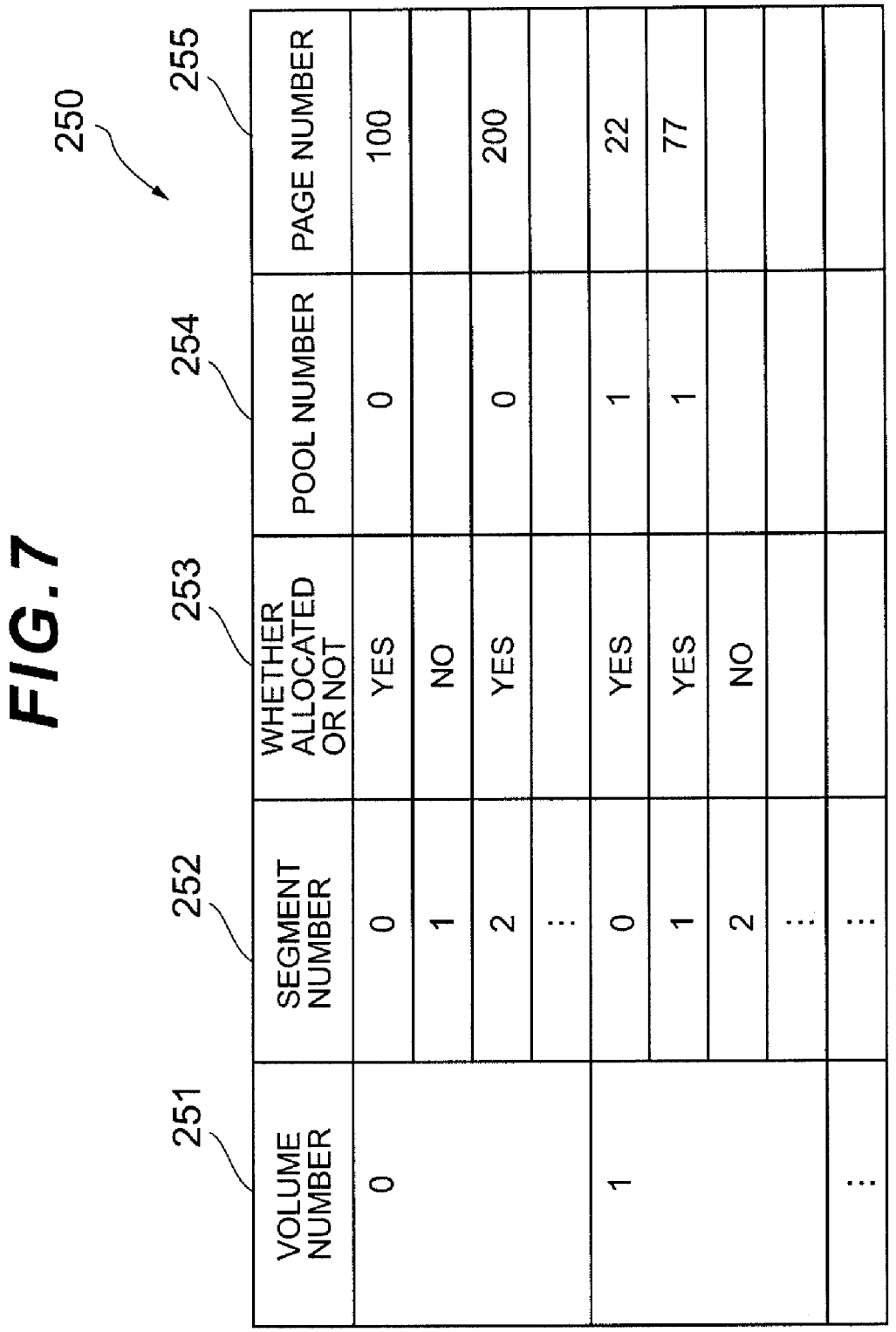

250
251
VOLUME NUMBER
252
SEGMENT NUMBER
253
WHETHER ALLOCATED OR NOT
254
POOL NUMBER
255
PAGE NUMBER
0
0
YES
0
100
1
NO
2
YES
0
200
...
1
0
YES
1
22
1
YES
1
77
2
NO
...
...
...

FIG. 8

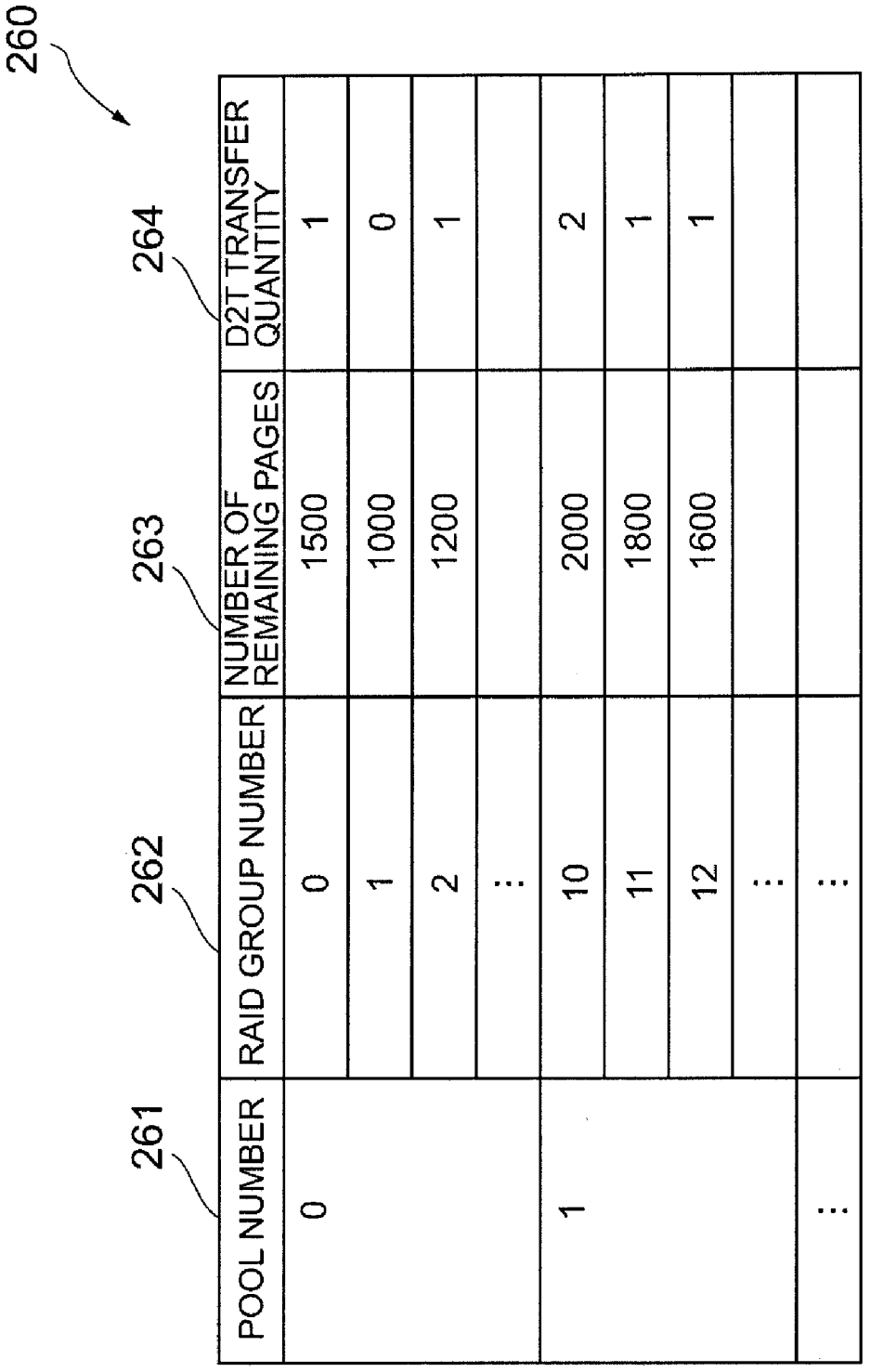

260
261
262
263
264
POOL NUMBER
RAID GROUP NUMBER
NUMBER OF REMAINING PAGES
D2T TRANSFER QUANTITY
0
0
1500
1
1
1000
0
2
1200
1
...
...
...
1
10
2000
2
11
1800
1
12
1600
1
...
...
...
...
...
...

FIG. 9

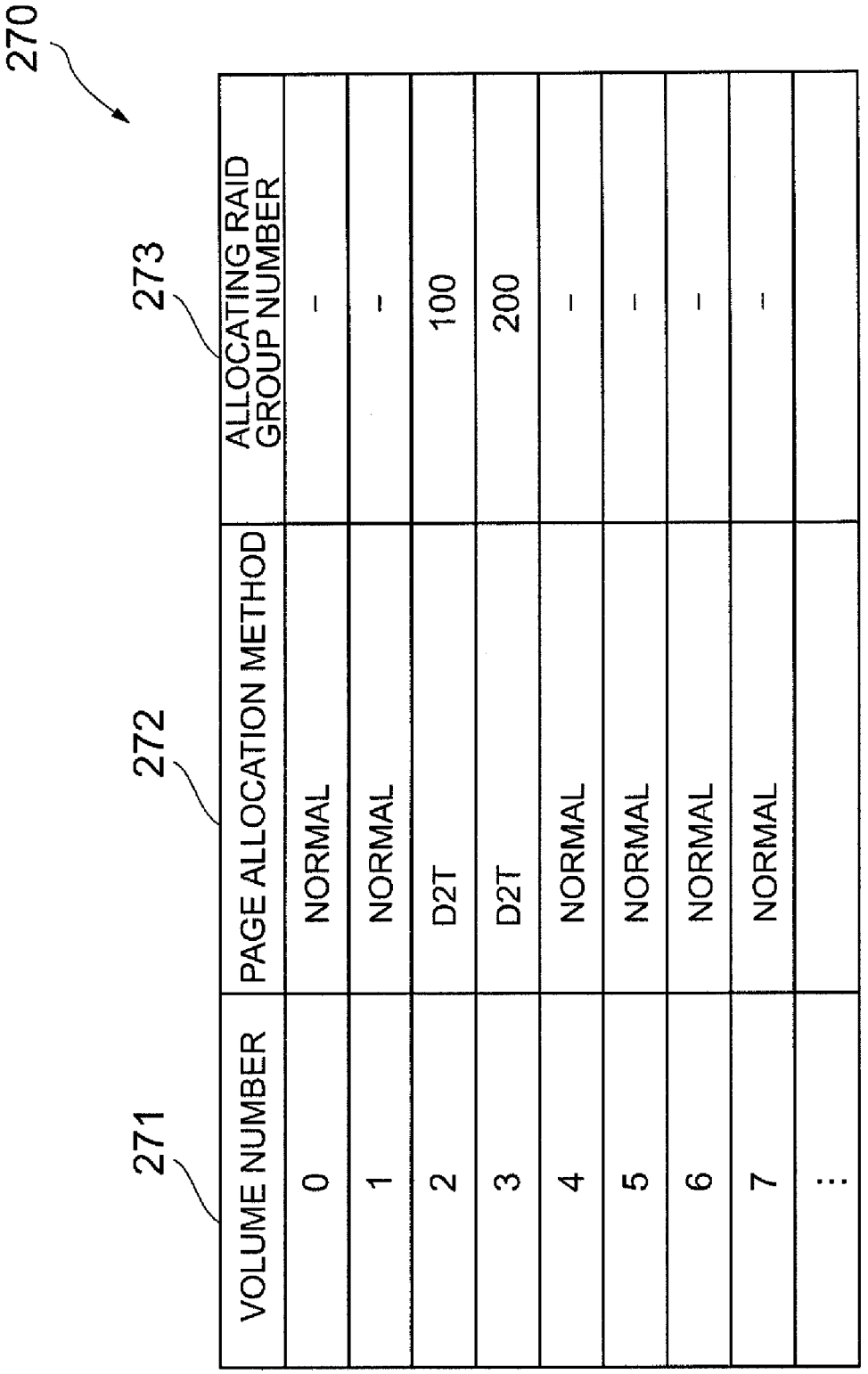

270
271
272
273
VOLUME NUMBER
PAGE ALLOCATION METHOD
ALLOCATING RAID GROUP NUMBER
0
NORMAL
–
1
NORMAL
–
2
D2T
100
3
D2T
200
4
NORMAL
–
5
NORMAL
–
6
NORMAL
–
7
NORMAL
–
⋮

FIG. 10

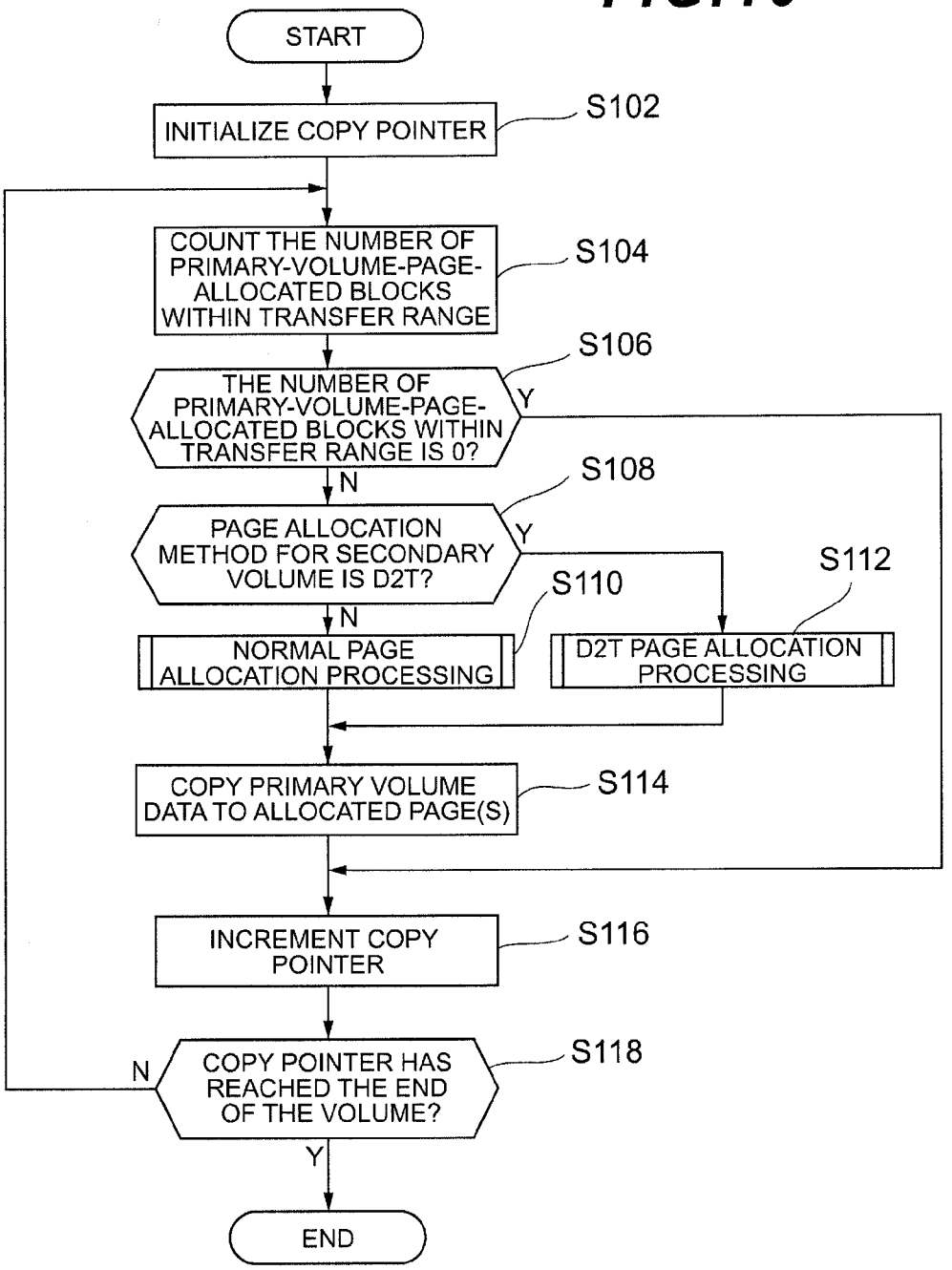

START
INITIALIZE COPY POINTER
S102
COUNT THE NUMBER OF PRIMARY-VOLUME-PAGE-ALLOCATED BLOCKS WITHIN TRANSFER RANGE
S104
S106
THE NUMBER OF PRIMARY-VOLUME-PAGE-ALLOCATED BLOCKS WITHIN TRANSFER RANGE IS 0?
Y
N
S108
PAGE ALLOCATION METHOD FOR SECONDARY VOLUME IS D2T?
Y
N
S110
S112
NORMAL PAGE ALLOCATION PROCESSING
D2T PAGE ALLOCATION PROCESSING
COPY PRIMARY VOLUME DATA TO ALLOCATED PAGE(S)
S114
INCREMENT COPY POINTER
S116
COPY POINTER HAS REACHED THE END OF THE VOLUME?
S118
N
Y
END

FIG. 11

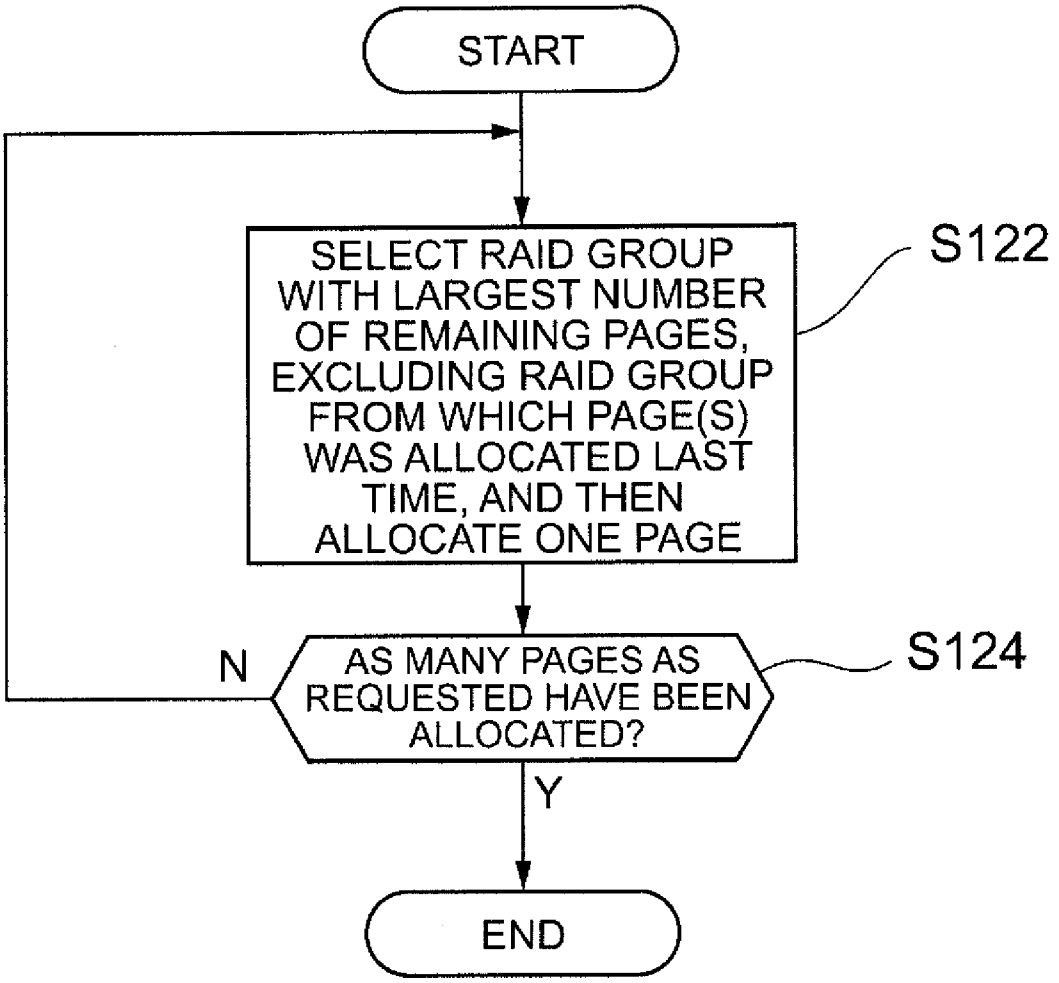

START
SELECT RAID GROUP WITH LARGEST NUMBER OF REMAINING PAGES, EXCLUDING RAID GROUP FROM WHICH PAGE(S) WAS ALLOCATED LAST TIME, AND THEN ALLOCATE ONE PAGE
S122
AS MANY PAGES AS REQUESTED HAVE BEEN ALLOCATED?
S124
N
Y
END

FIG. 12

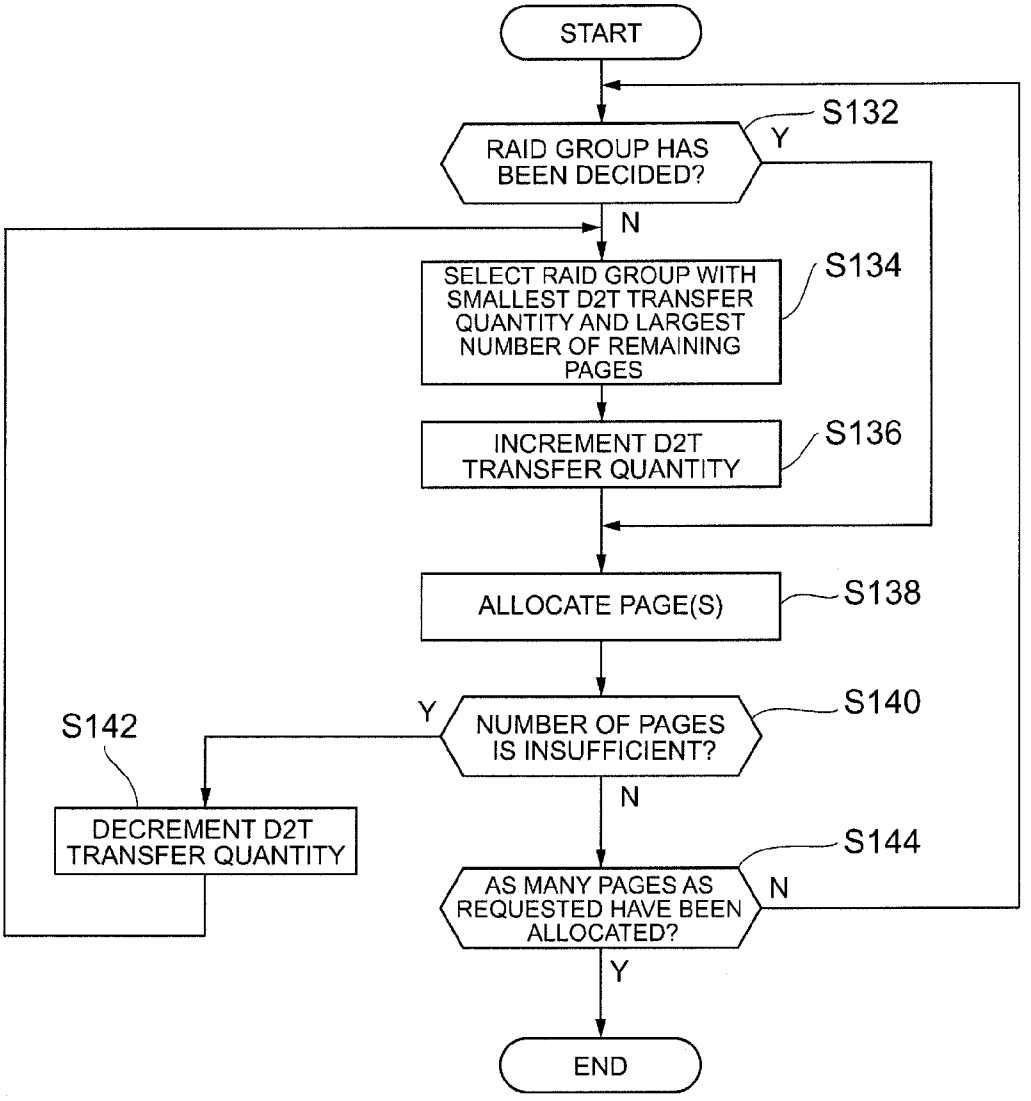

START
RAID GROUP HAS BEEN DECIDED?
S132
Y
N
SELECT RAID GROUP WITH SMALLEST D2T TRANSFER QUANTITY AND LARGEST NUMBER OF REMAINING PAGES
S134
INCREMENT D2T TRANSFER QUANTITY
S136
ALLOCATE PAGE(S)
S138
NUMBER OF PAGES IS INSUFFICIENT?
S140
Y
N
S142
DECREMENT D2T TRANSFER QUANTITY
AS MANY PAGES AS REQUESTED HAVE BEEN ALLOCATED?
S144
N
Y
END

FIG. 13

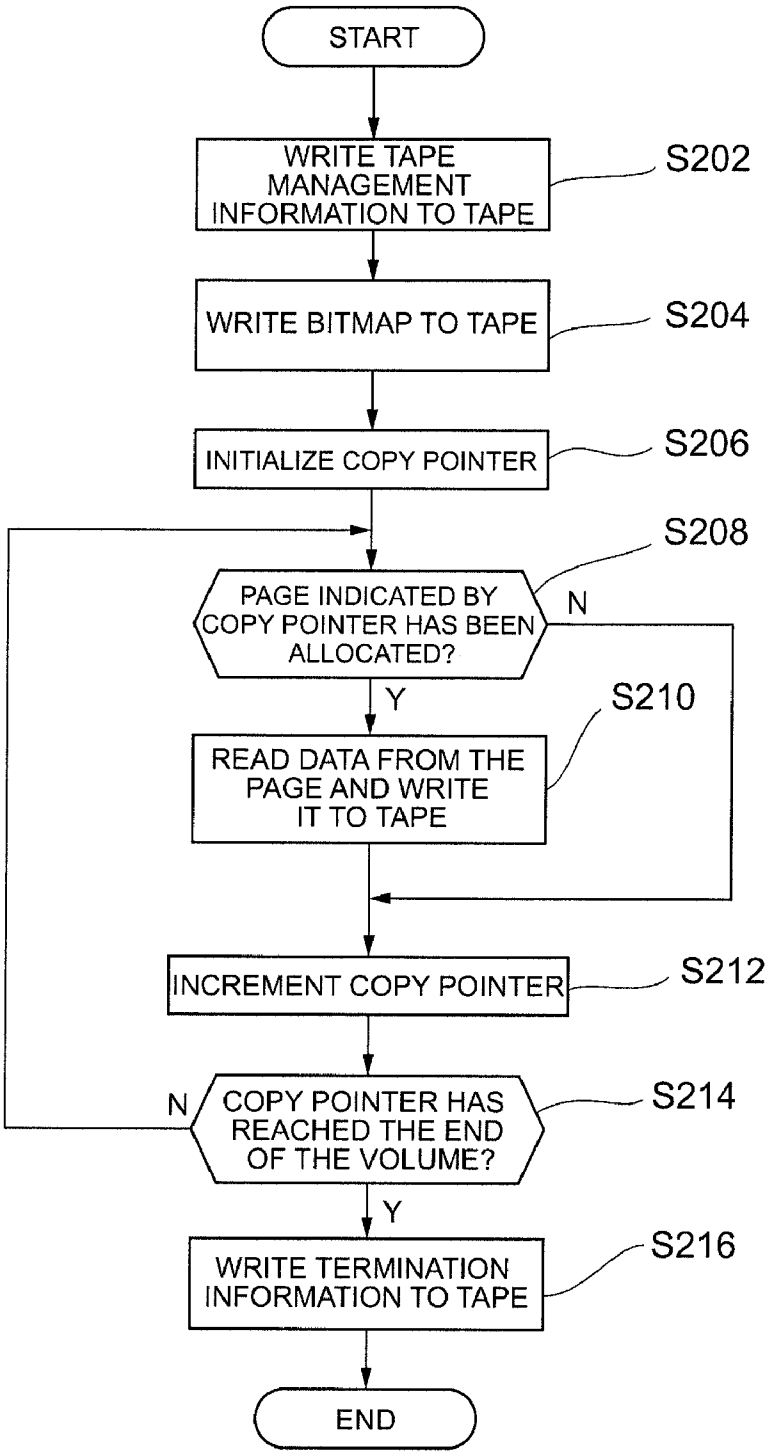

START
WRITE TAPE MANAGEMENT INFORMATION TO TAPE
S202
WRITE BITMAP TO TAPE
S204
INITIALIZE COPY POINTER
S206
S208
PAGE INDICATED BY COPY POINTER HAS BEEN ALLOCATED?
N
Y
S210
READ DATA FROM THE PAGE AND WRITE IT TO TAPE
INCREMENT COPY POINTER
S212
N
COPY POINTER HAS REACHED THE END OF THE VOLUME?
S214
Y
WRITE TERMINATION INFORMATION TO TAPE
S216
END

FIG. 14

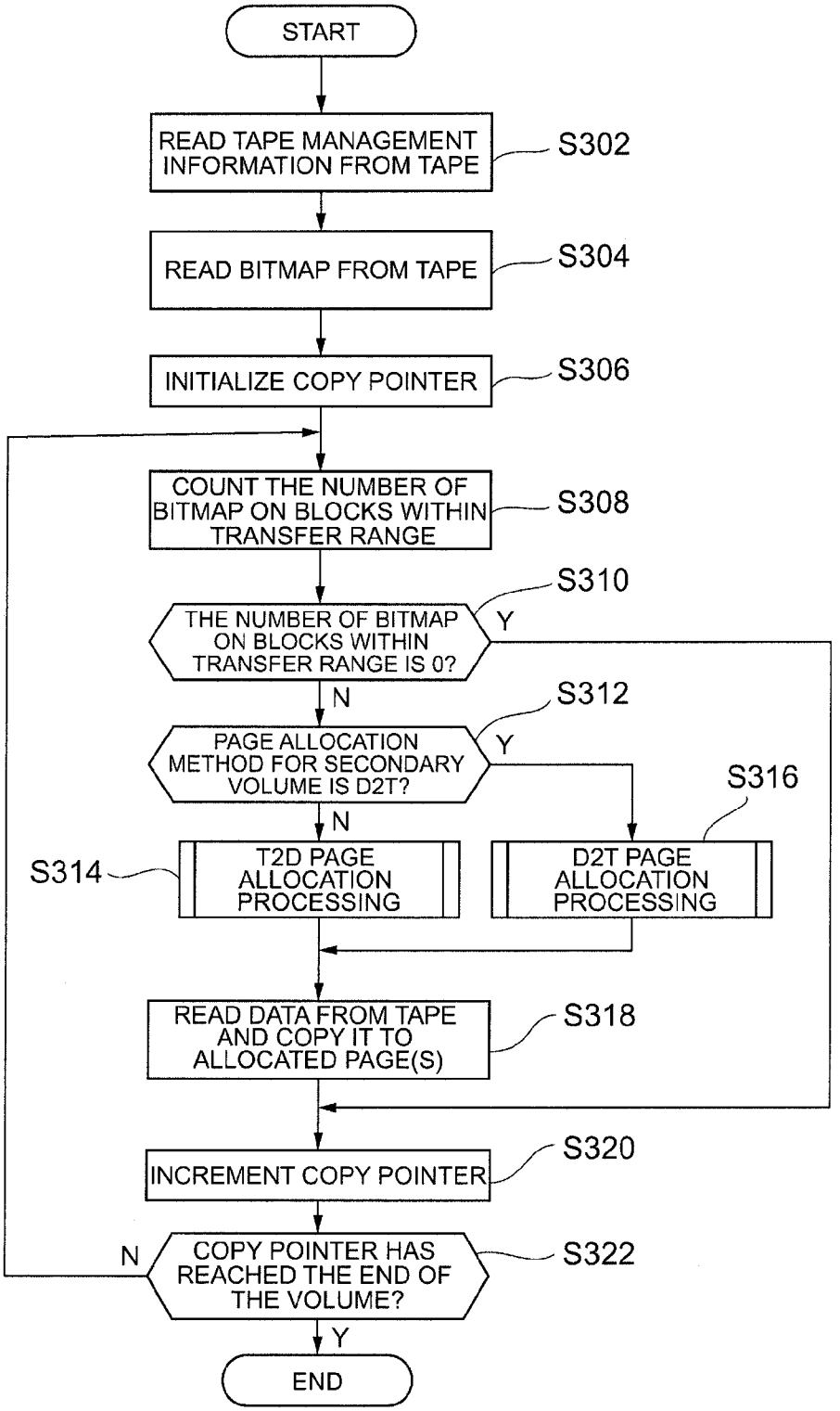

START
READ TAPE MANAGEMENT INFORMATION FROM TAPE
S302
READ BITMAP FROM TAPE
S304
INITIALIZE COPY POINTER
S306
COUNT THE NUMBER OF BITMAP ON BLOCKS WITHIN TRANSFER RANGE
S308
S310
THE NUMBER OF BITMAP ON BLOCKS WITHIN TRANSFER RANGE IS 0?
Y
N
S312
PAGE ALLOCATION METHOD FOR SECONDARY VOLUME IS D2T?
Y
N
S316
S314
T2D PAGE ALLOCATION PROCESSING
D2T PAGE ALLOCATION PROCESSING
READ DATA FROM TAPE AND COPY IT TO ALLOCATED PAGE(S)
S318
INCREMENT COPY POINTER
S320
N
COPY POINTER HAS REACHED THE END OF THE VOLUME?
S322
Y
END

FIG. 15

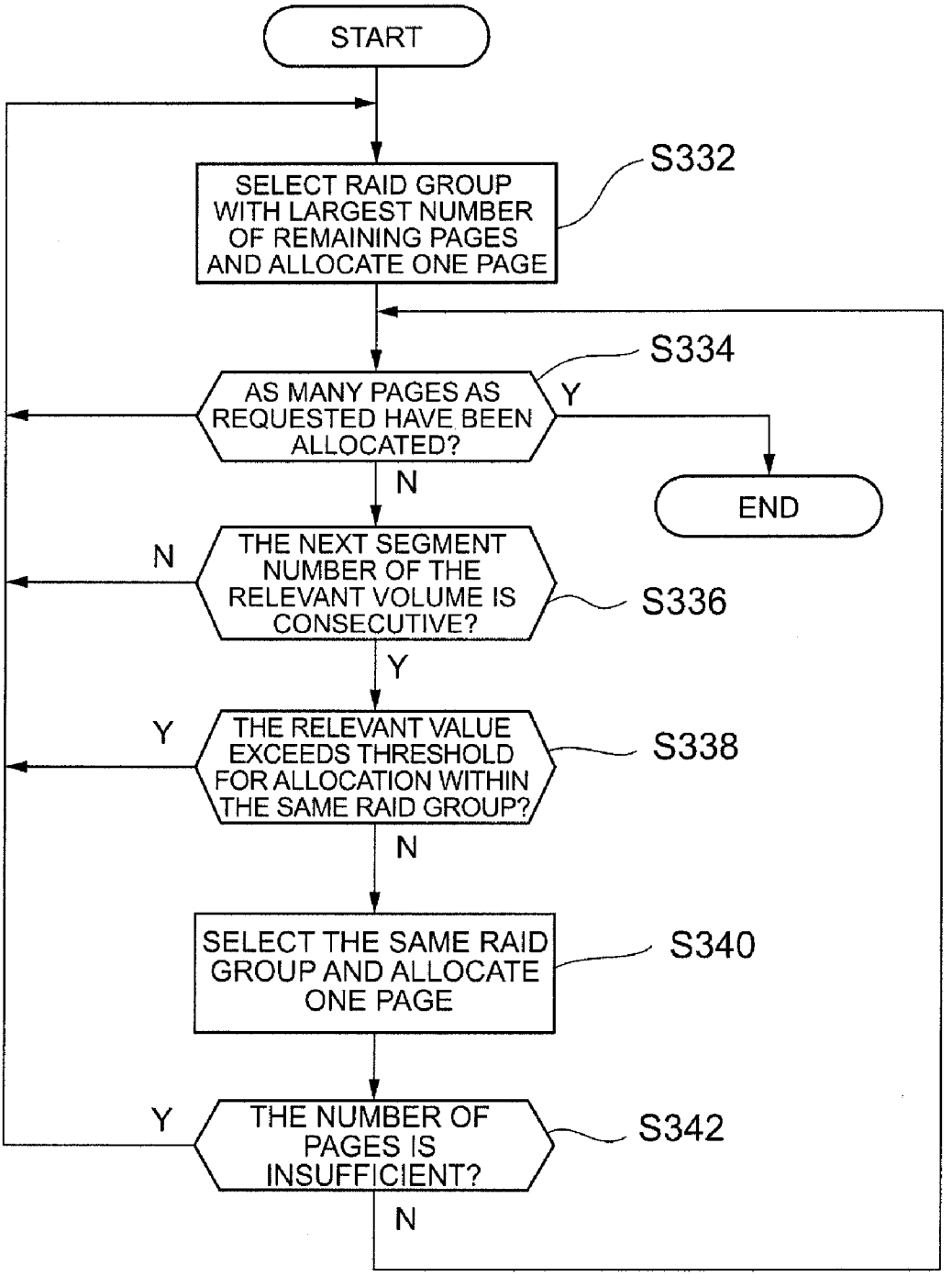
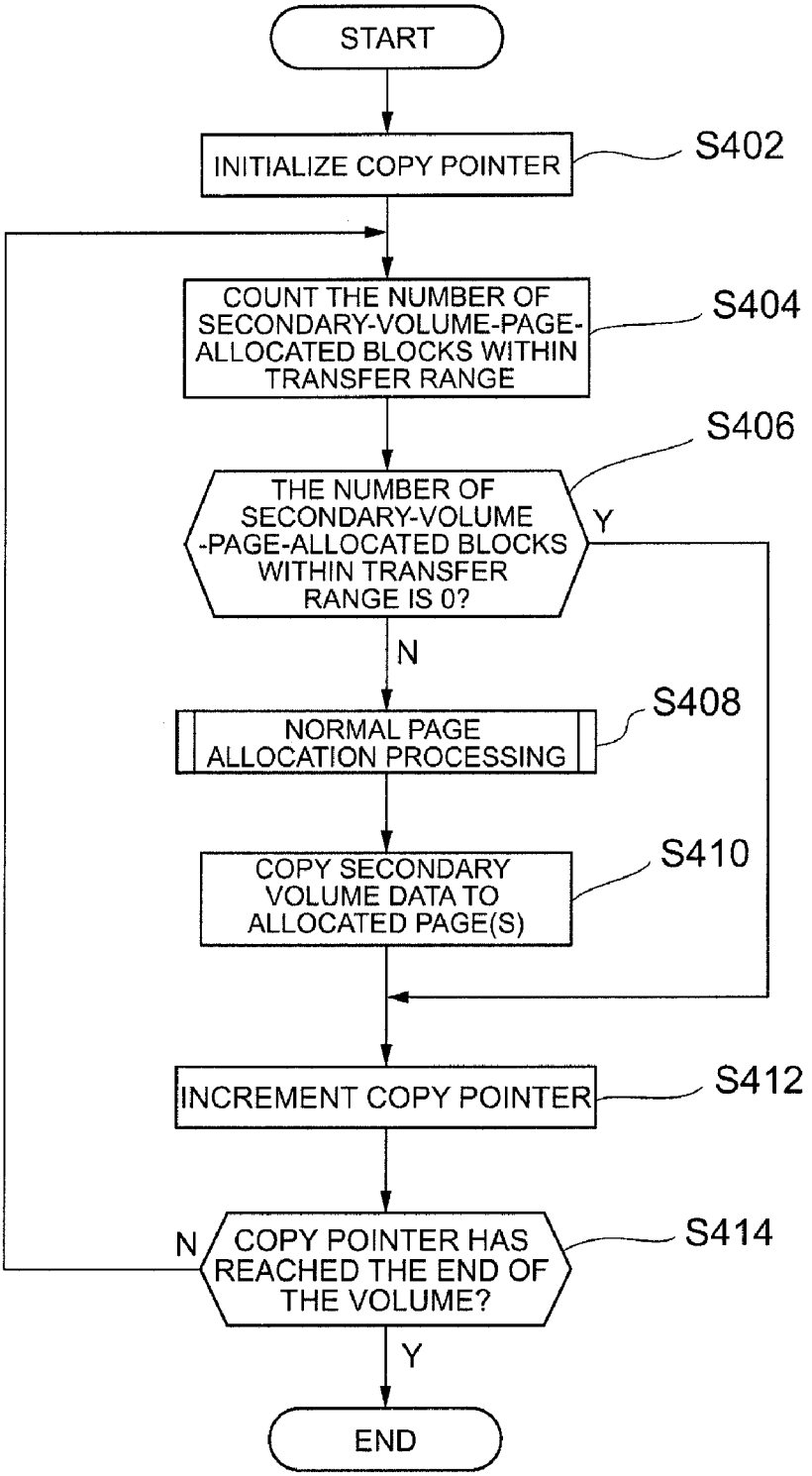

START
SELECT RAID GROUP WITH LARGEST NUMBER OF REMAINING PAGES AND ALLOCATE ONE PAGE
S332
AS MANY PAGES AS REQUESTED HAVE BEEN ALLOCATED?
S334
Y
N
END
THE NEXT SEGMENT NUMBER OF THE RELEVANT VOLUME IS CONSECUTIVE?
S336
N
Y
THE RELEVANT VALUE EXCEEDS THRESHOLD FOR ALLOCATION WITHIN THE SAME RAID GROUP?
S338
Y
N
SELECT THE SAME RAID GROUP AND ALLOCATE ONE PAGE
S340
THE NUMBER OF PAGES IS INSUFFICIENT?
S342
Y
N

FIG. 16


START
INITIALIZE COPY POINTER
S402
COUNT THE NUMBER OF SECONDARY-VOLUME-PAGE-ALLOCATED BLOCKS WITHIN TRANSFER RANGE
S404
S406
THE NUMBER OF SECONDARY-VOLUME -PAGE-ALLOCATED BLOCKS WITHIN TRANSFER RANGE IS 0?
Y
N
NORMAL PAGE ALLOCATION PROCESSING
S408
COPY SECONDARY VOLUME DATA TO ALLOCATED PAGE(S)
S410
INCREMENT COPY POINTER
S412
N
COPY POINTER HAS REACHED THE END OF THE VOLUME?
S414
Y
END

STORAGE APPARATUS AND STORAGE MANAGEMENT METHOD

TECHNICAL FIELD

[0001] The present invention relates to a storage apparatus and a storage management method and is suited for use in, for example, a storage apparatus and storage management method for managing data transfer performance for backup/restoration of data to tape storage media.

BACKGROUND ART

[0002] Recently, along with scale expansion and growing complexity of a storage environment due to an increase of company data, thin provisioning utilizing virtual volumes (hereinafter referred to as the virtual volumes) which do not have their own storage areas has become widely used (for example, see Patent Literature 1) for the purpose of facilitation of operation management and integration of the storage environment.

[0003] With the thin provisioning, a virtual volume(s) is presented to a host system. If the host system makes write access to the virtual volume, it allocates a physical storage area for actually storing the data to the virtual volume. As a result, a volume(s) whose capacity is larger than that of the storage areas in the storage apparatus can be presented to the host system, and the storages areas in the storage apparatus can be used very efficiently.

[0004] One or more HDDs (Hard Disk Drives) constitute one RAID (Redundant Arrays of Independent Disks) group, and one or more logical volumes are defined in storage areas provided by one RAID group. Also, one or more RAID groups constitute one storage pool and each storage pool is associated with one or more virtual volumes. If write access is made from the host system to a virtual volume, a storage area in a unit(s) of a predetermined size (the storage area of this size will be hereinafter referred to as the page) is allocated from any of the logical volumes in a storage pool, which is associated with the relevant volume, to the relevant segment of the virtual volume which has been write-accessed.

[0005] In this case, by means of the conventional thin provisioning, pages may be allocated equally from a plurality of RAID groups constituting pool volumes in order of access to the virtual volume when allocating pages to the relevant segment of the virtual volume. Consequently, data is stored in a plurality of HDDs equally by allocating pages equally from the plurality of RAID groups. In this way, an attempt is made to reduce response time for write access from the host system and enhance access performance.

CITATION LIST

Patent Literature

[0006] PTL 1: Japanese Patent Laid-Open (Kokai) Application Publication No. 2008-282382

SUMMARY OF INVENTION

Technical Problem

[0007] However, if the above-described conventional page allocation method is used, data will be stored randomly in the plurality of RAID groups. Therefore, when backing up the data for each virtual volume to, for example, tape storage media, it is necessary to access each of the RAID groups, from which pages have been allocated to each segment of the virtual volumes, and read the data from each of the RAID groups. As a result, data access may be delayed during data transfer to the tape storage media or access conflicts to the RAID groups may occur, thereby causing a problem of degradation of data transfer performance.

[0008] The present invention was devised in light of the circumstances described above and aims at providing a storage apparatus and storage management method capable of enhancing the performance to transfer data to tape storage media in thin provisioning.

Solution to Problem

[0009] In order to solve the problem described above, a storage apparatus connected via a network to a host system requesting writing of data and to a tape management apparatus for managing tape storage media for storing the data is provided according to this invention. The storage apparatus includes: a storage unit for storing data, for which a write request is issued by the host system, in a virtual volume; and an allocation unit for allocating a predetermined area in a plurality of storage devices to a data storage area in the virtual volume in which the data is stored by the storage unit. The storage apparatus is characterized in that if the virtual volume is a backup object to be backed up to the tape storage media, the allocation unit allocates a predetermined area in a specified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume; and if the virtual volume is not a backup object to be backed up to the tape storage media, the allocation unit allocates a predetermined area in an unspecified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume.

[0010] According to the above-described configuration, the storage apparatus stores data, for which a write request is issued from the host system, in a virtual volume; and allocates a predetermined area in a plurality of storage devices to a data storage area in the virtual volume storing the data. If the virtual volume is a backup object to be backed up to the tape storage media, the storage apparatus allocates a predetermined area in a specified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume; and if the virtual volume is not a backup object to be backed up to the tape storage media, the storage apparatus allocates a predetermined area in an unspecified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume. Therefore, when backing up the data stored in the virtual volume to the tape storage media, it is unnecessary to access the unspecified storage device at the same time. So, the occurrence of access conflicts to the storage devices will be reduced.

Advantageous Effects of Invention

[0011] The present invention can enhance the performance to transfer data to external storage media in thin provisioning.

BRIEF DESCRIPTION OF DRAWINGS

[0012] FIG. **1** is a block diagram showing the overall configuration of a computer system according to an embodiment of this invention.

[0013] FIG. **2** is a block diagram showing the software configuration of each device of the computer system according to the embodiment.

[0014] FIG. **3** is an explanatory diagram explaining page allocation processing according to the embodiment.

[0015] FIG. **4** is an explanatory diagram explaining a data format of a tape storage medium according to the embodiment.

[0016] FIG. **5** is an explanatory diagram explaining conventional page allocation.

[0017] FIG. **6** is an explanatory diagram explaining page allocation according to an embodiment of the invention.

[0018] FIG. **7** is an explanatory diagram explaining the content of a virtual volume management table according to the embodiment.

[0019] FIG. **8** is an explanatory diagram explaining the content of a pool volume management table according to the embodiment.

[0020] FIG. **9** is an explanatory diagram explaining the content of a page allocation management table according to the embodiment.

[0021] FIG. **10** is a flowchart illustrating data backup processing according to the embodiment.

[0022] FIG. **11** is a flowchart illustrating the details of normal page allocation processing according to the embodiment.

[0023] FIG. **12** is a flowchart illustrating the details of D2T page allocation processing according to the embodiment.

[0024] FIG. **13** is a flowchart illustrating processing for backing up data to tape storage media according to the embodiment.

[0025] FIG. **14** is a flowchart illustrating processing for restoring data from tape storage media according to the embodiment.

[0026] FIG. **15** is a flowchart illustrating page allocation processing for data restoration according to the embodiment.

[0027] FIG. **16** is a flowchart illustrating data restoration processing according to the embodiment.

DESCRIPTION OF EMBODIMENTS

[0028] An embodiment of this invention will be explained with reference to the attached drawings.

1. Hardware Configuration of Computer System

[0029] The hardware configuration of a computer system **1** according to this embodiment will be firstly explained with reference to FIG. **1**. As shown in FIG. **1**, the computer system **1** according to this embodiment includes a host system **10**, a storage apparatus **20**, a tape management apparatus **30**, a network **40**, and a management terminal **50**. The host system **10** and the tape management apparatus **30** are connected via the network **40** to the storage apparatus **20**.

[0030] The host system **10** is a computer system equipped with information resources such as a CPU (Central Processing Unit) and a memory and is composed of, for example, a personal computer, a workstation, or a mainframe. The CPU serves as a processor controller and controls the operation of the entire host system **10** according to, for example, programs and operational parameters stored in the memory. Main programs stored in the memory will be explained later in detail. The host system **10** may also include information input devices such as a keyboard, switch, pointing device, and microphone, and information output devices such as a monitor display and speaker.

[0031] The storage apparatus **20** includes, for example, a controller **21** and a storage device group **22**. The controller **21** includes information processing resources such as a CPU and a memory. The CPU serves as a processor controller in the same manner as described above and controls the operation of the entire storage apparatus **20** according to, for example, programs and operational parameters stored in the memory. Main programs stored in the memory will be explained later in detail.

[0032] The storage device group **22** is constituted from a plurality of HDDs **23***a*, **23***b*, and so on to **23***n* (hereinafter sometimes simply referred to as the disk(s) **23**). The storage devices **22** may be composed of a plurality of hard disk drives including, for example, expensive hard disk drives such as FC (Fibre Channel) disks or inexpensive hard disk drives such as SATA (Serial AT Attachment) disks.

[0033] The tape management apparatus **30** includes a CPU (not shown in the drawing) for controlling a carrier mechanism **31** for carrying tape storage media **34***a*, **34***b*, and so on to **34***n* (hereinafter sometimes referred to the tape storage media **34**) which constitute an external storage device **33**, a memory (not shown in the drawing) for storing a control program for controlling the carrier mechanism **31**, the carrier mechanism **31**, tape drives **32***a*, **32***b* (hereinafter sometimes referred to as the tape drive(s) **32**), and the external storage device **33**. The tape management apparatus **30** stores data, which has been transferred from the storage apparatus **20**, in the external storage device **33** (the tape storage media **34**) under the control of the storage apparatus **20** in response to a request from the host system **10**.

[0034] The network **40** connects the host system **10**, the storage apparatus **20**, and the tape management apparatus **30** each other and is composed of, for example, a SAN (Storage Area Network). In this case, communication between the devices is performed according to Fibre Channel Protocol.

[0035] The network **40** may be a LAN (Local Area Network), the Internet, public lines or private lines. If the network **40** is a LAN, communication is performed according to TCP/IP (Transmission Control Protocol/Internet Protocol) protocol.

[0036] The management terminal **50** is a computer device equipped with information processing resources such as a CPU and a memory and manages the disk, etc. of the storage apparatus **20** in accordance with input by, for example, an operator. The management terminal **50** is connected to the storage apparatus **20** via a LAN or similar and is composed of, for example, a personal computer.

2. Software Configuration of Computer System

[0037] Next, the main software configuration of each of the devices constituting the computer system **1** will be explained with reference to FIG. **2**. Particularly, an example of programs stored in the memories for the host system **10** and the storage apparatus **20** will be explained with reference to FIG. **2**.

[0038] As shown in FIG. **2**, the memory for the host system **10** stores, for example, a disk mirror command program **101** and a backup/restoration command program **102**. These programs area executed by the CPU for the host system **10**. The disk mirror command program **101** is a program for giving a command to the storage apparatus **20** to multiplex data (or to duplicate data). In this embodiment, the disk mirror command program **101** gives a command to the storage apparatus **20** to form a pair of virtual volumes **220***a*, **220***b* (hereinafter referred to as the copy pair) to duplicate data by implementing a copy function (hereinafter referred to as the shadow image function) within the same storage apparatus or to separate the copy pair in order to staticize the data and execute a backup.

[0039] The backup/restoration command program **102** gives a command to the tape management apparatus **30** to back up data, that is, to copy data stored in the secondary virtual volume **220***b*, from among the virtual volumes **220***a*, **220***b* constituting the above-described copy pair, to the tape storage media **34** included in the tape management apparatus **30**. Also, the backup/restoration command program **102** gives a command to the tape management apparatus **30** to restore the virtual volume **220***b*, in which a failure occurs, from the data stored in the tape storage media **34**.

[0040] Furthermore, the memory for the storage apparatus **20** stores, for example, a disk mirror control program **201**, a disk control program **202**, a thin provisioning control program **203**, a backup/restoration control program **204**, and a tape library control program **205** as shown in FIG. **2**. These programs are executed by the aforementioned CPU for the storage apparatus **20**.

[0041] The thin provisioning control program **203** is a program for executing processing for virtualizing storage areas. Specifically speaking, the thin provisioning control program **203** executes processing for: generating the virtual volumes **220***a*, **220***b*, which actually do not have storage areas; presenting some virtual volume **220***a* to the host system **10**; and, if write access is made to that virtual volume from the host system **10**, allocating a physical storage area for actually storing data to the virtual volume **220***a* (hereinafter referred to as the thin provisioning processing).

[0042] When copying data written to the virtual volume **220***a* (hereinafter sometimes referred to as the primary virtual volume), which is presented to the host system **10**, to the virtual volume **220***b* (hereinafter sometimes referred to as the secondary virtual volume) which is set to form the copy pair with the primary virtual volume **220***a*, the thin provisioning control program **203** allocates a storage area to the relevant area in the secondary virtual volume **220***b*.

[0043] The outline of a thin provisioning function mounted in the storage apparatus **20** according to this embodiment will be explained with reference to FIG. **3**. FIG. **3** shows the outline of the thin provisioning function mounted in the storage apparatus **20**. In the storage apparatus **20**, one or more HDDs **23** constitute one RAID group and one or more logical volumes (not shown in the drawing) are defined in a storage area(s) provided by one RAID group. The logical volumes provided by a plurality of RAID groups are managed as one pool volume. If write access is made from the host system **10** to the virtual volume **220**, a storage area is allocated on a page basis from the logical volumes in a pool volume associated with the virtual volume **220** and write object data is stored in the relevant page(s).

[0044] For example, if data is stored in segment **0** of the virtual volume **220** as shown in FIG. **3**, page **0** of a RAID group RG #**0** is allocated to the segment **0**. If data is stored in segment **2** of the virtual volume **220**, page **100** of a RAID group RG #**1** is allocated to the segment **2**. In this way, the thin provisioning control program **203** allocates pages sequentially from the plurality of RAID groups in order to enhance data access performance from the host system **10**.

[0045] Referring back to FIG. **2**, the disk mirror control program **201** is a program for controlling disk mirror processing in accordance with a command from the host system **10**. The disk mirror processing is processing for duplicating data or separating a copy pair in order to staticize the data and executes a backup, using the aforementioned shadow image function. Specifically speaking, when data is written to the primary virtual volume **220***a*, the disk mirror control program **201** controls the disk control program **202** so that the data will be copied synchronously or asynchronously to the secondary virtual volume **220***b* which is set to form a copy pair with the primary virtual volume **220***a*.

[0046] The disk control program **202** is a program for controlling each HDD mounted in the storage apparatus **20** and controls reading or writing data, which is read/written to the virtual volume **220**, from or to the relevant HDD **23**, using the thin provisioning function, the disk mirror function, or the backup/restoration function described earlier.

[0047] The backup/restoration control program **204** copies the data, which is stored in the secondary virtual volume, to the tape storage media **34** in the tape management apparatus **30** and executes processing for backing up the data. A data format of the tape storage media **34** to which the data stored in the secondary virtual volume is copied will be explained with reference to FIG. **4**. FIG. **4** is an explanatory diagram explaining the data format of the tape storage media **34**.

[0048] As shown in FIG. **4**, the data format of the tape storage media **34** includes, for example, tape management information **340**, a bitmap **341**, data storage areas **342**, and termination information **343**. The tape management information **340** stores information about data stored in the tape storage media **34** and includes, for example, the name of the tape storage media **34**, data backup start date and time, the number of stored volumes which are backup object virtual volume(s), volume number, volume size, and bitmap size.

[0049] The bitmap **341** stores data storage areas of the virtual volume **220** which is a backup object. It is possible to recognize which segment of the backup object virtual volume **220** stores the relevant data, according to the bitmap **341**. The data storage areas **342** store backup data of the virtual volume **220**. If data is stored in segment **0**, segment **2**, segment **5**, segment **6**, and segment **7** of the virtual volume **220** as shown in FIG. **4**, the segment **0** data, the segment **2** data, the segment **5** data, the segment **6** data, and the segment **7** data are stored in sequential order in the data storage areas **342**.

[0050] The termination information **343** stores a backup termination date and time and the number of stored volumes which are stored in the tape storage media **34**.

[0051] In this way, the backup/restoration control program **204** backs up data on a virtual volume basis. As described above, the data stored in the virtual volume **220** is actually stored in the relevant page(s) in the pool volume. Therefore, the backup/restoration control program **204** reads the data stored in the relevant page allocated to each segment of the virtual volume **220** and transfers it to the tape management apparatus **30**.

[0052] The memory for the management terminal **50** stores, for example, a management program **501**. In response to input by, for example, the operator, the CPU for the management terminal **50** executes the management program **501**. The management program **501** is a program for managing, for example, disks in the storage apparatus **20**. For example, the management program **501** manages association between the virtual volumes and the pool volumes and virtual volume pair setting. Furthermore, the management program **501** executes, for example, an update of data tables described later depending on whether or not the virtual volume is accessed only by the host system **10**, or whether or not the virtual volume is a backup object to be backed up to the tape storage media **34**.

[0053] A method executed by the thin provisioning control program **203** for allocating a page(s) to the secondary virtual

volume will be explained with reference to FIG. **5** and FIG. **6**. FIG. **5** is an explanatory diagram explaining a conventional page allocation method. FIG. **6** is an explanatory diagram explaining a page allocation method according to this embodiment.

[0054] Incidentally, in the following explanation, virtual volumes **11**, **12**, **21**, **22** exist as primary virtual volumes for the disk mirror processing described earlier and a first storage pool constituted from logical volumes created in RAID groups RG #**1** to RG #**4**, respectively, is associated with the above-mentioned primary virtual volumes. Also, virtual volume **31**, **32**, **41**, **42** exist as secondary virtual volumes for the disk mirror processing and a second storage pool constituted from logical volumes created in RAID groups RG #**5** to RG #**8**, respectively, is associated with the above-mentioned secondary virtual volumes.

[0055] Therefore, if write access is made from the host system **10** to a primary virtual volume in the examples shown in FIG. **5** and FIG. **6**, a necessary amount of storage area is allocated on a page basis from any of the logical volumes in the first storage pool to the relevant primary virtual volume. If the data is copied to a secondary virtual volume, a necessary amount of storage area is allocated on a page basis from any of the logical volumes in the second storage pool to the relevant second virtual volume.

[0056] If the conventional page allocation method shown in FIG. **5** is used in the above-described case when the host system **10** writes data to the primary virtual volume and pages are allocated from the first storage pool for the primary virtual volumes to the relevant segment of the primary virtual volume, the pages are allocated in sequential order from each of the RAID groups constituting the first storage pool. For example, pages are allocated sequentially from the RAID groups RG #**1**, RG #**2**, RG #**3**, RG #**4**, which constitute the storage pool for the primary virtual volumes, to the write area in the virtual volume **11** which is the primary virtual volume. In this way, data access performance from the host system **10** is enhanced by reducing response time for data access from the host system **10** by sequentially allocating pages from the plurality of RAID groups.

[0057] When pages are allocated from the second storage pool to the relevant segment of the secondary virtual volume in order to copy data, which is stored in the primary virtual volume, to the secondary virtual volume by the conventional page allocation method, pages are allocated sequentially from each of the RAID groups constituting the second storage pool. For example, as shown in FIG. **5**, arbitrary pages are allocated sequentially from the RAID groups RG #**5**, RG #**6**, RG #**7**, RG #**8**, which constitute the second storage pool, to write areas in the virtual volumes **31**, **32**, **41**, **42** which are the secondary virtual volumes.

[0058] Pages are allocated randomly from each of the RAID groups constituting the first or second storage pool corresponding to the primary virtual volume and the secondary virtual volume by the conventional page allocation method as described above in consideration of the data access performance from the host system **10**. Accordingly, if data stored in the secondary virtual volume is to be stored in the tape storage media, it is necessary to access the plurality of RAID groups, thereby causing a problem of degradation of the data transfer performance Moreover, when data is to be backed up from a plurality of secondary virtual volumes to a plurality of tape storage media, it is sometimes necessary to access the plurality of RAID groups at the same time, thereby causing problems of the occurrence of RAID group conflicts and degradation of the data transfer performance.

[0059] Therefore, according to this embodiment, if data stored in the secondary virtual volume is a data backup object to be backed up to the tape storage media, pages are not allocated sequentially from each of the RAID groups, which constitute the second storage pool associated with the secondary virtual volume, to the relevant segment of the secondary virtual volume, but arbitrary pages are allocated from a specified RAID group. For example, as shown in FIG. **6**, pages are allocated only from the RAID group #**5** constituting the second storage pool to the secondary virtual volume **31** and pages are allocated only from the RAID group #**6** constituting the second storage pool to the secondary virtual volume **32**.

[0060] As a result, when backing up the data stored in the secondary virtual volume **31** and the secondary virtual volume **32** to the tape storage media **34***a*, it is no longer necessary to access the plurality of RAID groups constituting the second storage pool. In other words, when copying the data stored in the secondary virtual volume **31** to the tape storage media **34***a*, it is only necessary to access the RAID group #**5**; and when copying the data stored in the secondary virtual volume **32** to the tape storage media **34***a*, it is only necessary to access the RAID group #**6**.

[0061] However, with regard to the primary virtual volume, pages are allocated sequentially from each of the RAID groups constituting the first storage pool according to this embodiment. As a result, it is possible to maintain the access performance from the host system **10** to the virtual volume and prevent degradation of the performance at the time of a data backup to the tape storage media.

(Conventional Problems and Solutions for Restoration)

[0062] If a failure or similar occurs in a disk **23**, the backup/restoration control program **204** executes data restoration, using the tape storage media **34**. When restoring data in the tape storage media **34** to the secondary virtual volume, the backup/restoration control program **204** executes the page allocation processing in the same manner as the data backup. Specifically speaking, if the secondary virtual volume which is an object to be restored is a data backup object to be backed up to the tape storage media **34**, pages are allocated not sequentially from each of the RAID groups, but from a specified RAID group, to a write area in the virtual volume.

[0063] Furthermore, if the secondary virtual volume which is an object to be restored is not a data backup object to be backed up to the tape storage media **34**, that is, when pages are allocated sequentially from each of the RAID groups, it is possible to execute data restoration at a high speed by considering continuity in the bitmap **341** stored in the tape storage media **34**.

[0064] Specifically speaking, continuity in the bitmap **341** of the tape storage media **34** is judged and the same RAID group is selected to allocate pages to consecutive areas. When selecting the same RAID group to allocate pages, a least common multiple of the amount of data that can be transferred within average response time of the tape management apparatus **30** relative to the host system **10** and the striping size of the RAID group may be used as a threshold for data transfer. The average response time of the tape management apparatus **30** relative to the host system **10** can be maximum occupation time of the RAID group used for data transfer from the tape storage media **34**. The striping size of the RAID

group is a value obtained by multiplying the block size of the disk **23** by the number of the disks **23** in the RAID group.

[0065] The amount of data that can be transferred within the average response time of the tape management apparatus **30** relative to the host system **10** is a value obtained by multiplying the average response time by a speed of data transfer from the tape storage media **34**. For example, assuming that the data transfer speed of the tape storage media **34** by the tape management apparatus **30** is 240 MB/S and the average response time of the tape management apparatus **30** relative to the host system **10** is 5 to 10 ms, the amount of data which can be transferred from the tape storage media **34** within the average response time 5 ms is 1.2 MB and the amount of data which can be transferred from the tape storage media **34** within the average response time 10 ms is 2.4 MB.

[0066] Therefore, the least common multiple of the amount of data ranging from 1.2 MB to 2.4 MB and the striping size of the RAID group can be set as a data transfer threshold for data restoration. For example, the data transfer threshold for each different striping size of the RAID group can be obtained as listed below. If the least common multiple exceeds 2.4 MB, the data transfer threshold shown below is set as 2.4 MB; and if the least common multiple is less than 1.2 MB, the striping size is multiplied by n so that the threshold will be within the range of 1.2 MB to 2.4 MB.

Disk Block Size×Number of Disks in RAID Group=Striping Size→Threshold

512 KB×1=512 KB→2.0 MB

512 KB×2=1.0 MB→2.0 MB

512 KB×3=1.5 MB→1.5 MB

512 KB×4=2.0 MB→2.0 MB

512 KB×5=2.5 MB→2.4 MB

512 KB×6=3.0 MB→2.4 MB [Math.1]

[0067] In this way, it is possible according to this embodiment to achieve high-speed restoration by judging the continuity in the bitmap stored in the tape storage media at the time of restoration and allocating pages from the same RAID group to consecutive areas.

3. Details of Functions of Storage Apparatus

[0068] Next, specific functions of the storage apparatus **20** will be explained with reference to FIG. **7** to FIG. **9**. FIG. **7** to FIG. **9** are tables to which reference is made when explaining the specific functions of the storage apparatus **20**. These tables are stored in the memory for the storage apparatus **20**.

[0069] When data is written by the host system **10** as described above, the storage apparatus **20** allocates an arbitrary page(s) from the storage pool to the relevant segment of the virtual volume. The storage apparatus **20** manages which page of which storage pool is allocated to which volume of the virtual volumes, using a virtual volume management table **250** described below.

[0070] Now, the content of the virtual volume management table **250** for managing the virtual volumes will be explained with reference to FIG. **7**. FIG. **7** is an explanatory diagram explaining the content of the virtual volume management table **250**. As shown in FIG. **7**, the virtual volume management table **250** includes, for example, a volume number column **251**, a segment number column **252**, a whether-allocated-or-not column **253**, a pool number column **254**, and a page number column **255**. The volume number column **251** stores the number for identifying each virtual volume. The segment number column **252** stores the number for identifying a plurality of segments constituting each volume of the virtual volumes.

[0071] The whether-allocated-or-not column **253** stores information for identifying whether or not a page is allocated to each segment (NO/YES). If a page of the pool volume is not allocated, the whether-allocated-or-not column **253** stores NO. If data is written from the host system **10** to an arbitrary segment and an arbitrary page is allocated from the pool volume, the whether-allocated-or-not column **253** stores YES.

[0072] The pool number column **254** stores the number for identifying a pool volume allocated to each segment. The page number column **255** stores the number for identifying a page which is a storage area in the pool volume allocated to each segment. If the storage apparatus **20** allocates a page in the pool volume to any of the virtual volumes, it adds and updates each piece of data stored in the virtual volume management table **250**.

[0073] If the host system **10** issues a data read request to the storage apparatus **20** to read data stored in the virtual volume, the storage apparatus **20** transfers the data, for which the read request is made, to the host system **10**. The storage apparatus **20** refers to the above-described virtual volume management table **250** and reads the relevant page of the pool volume corresponding to the relevant segment of the virtual volume for which the read request is made.

[0074] Furthermore, the storage apparatus **20** manages, for example, the correspondence relationship between the plurality of RAID groups and the pool volumes and the number of remaining pages in each RAID group, using a pool volume management table **260** described below. When allocating a page(s) to a write area in the virtual volume, the storage apparatus **20** refers to the pool volume management table **260** and allocates a page(s) of the pool volume to the virtual volume.

[0075] Now, the pool volume management table **260** will be explained with reference to FIG. **8**. FIG. **8** is an explanatory diagram explaining the content of the pool volume management table **260**. As shown in FIG. **8**, the pool volume management table **260** includes a pool number column **261**, a RAID group number column **262**, a numberof-remaining-pages column **263**, and a number-of-D2T-transfer-object-volumes column **264**.

[0076] The pool number column **261** stores the number for identifying a pool volume. The

[0077] RAID group number column **262** stores the number for identifying a RAID group constituted from a plurality of disks. The number-of-remaining-pages column **263** stores the number of remaining pages in each RAID group. The number of remaining pages means the number of pages in which data is not stored yet. The number-of-D2T-transfer-object-volumes column **264** stores information indicating the number of virtual volume(s) which is a transfer object(s), from among the virtual volumes to which each RAID group is allocated.

[0078] When allocating a page(s) to a write area in the virtual volume, the storage apparatus **20** allocates a page from a RAID group with a small number of volumes from among the data stored in the number-of-D2T-transfer-object-volumes column **264**, or a page from a RAID group with a large

number of pages from among the data stored in the number-of-remaining-pages column **263**.

[0079] If the virtual volume is a backup object volume to be backed up to the tape storage media, it is possible to enhance the performance to transfer data to the tape management apparatus **30** by allocating a page(s) of a RAID group with a small number of volumes in the number-of-D2T-transfer-object-volumes column **264**. If the virtual volume is a backup object volume to be backed up to the tape storage media, it is possible to allocate a page(s) from one RAID group to one virtual volume by allocating the RAID group with a large number of pages in the numberof-remaining-pages column **263**.

[0080] Furthermore, the storage apparatus **20** manages a page allocation method for each virtual volume, using a page allocation management table **270** described below. Examples of the page allocation method can include a normal page allocation method of allocating pages equally from a plurality of RAID groups to the virtual volume and a D2T page allocation method of allocating pages from a fixed RAID group to the virtual volume.

[0081] Now, the page allocation management table **270** will be explained with reference to FIG. **9**. FIG. **9** is an explanatory diagram explaining the content of the page allocation management table **270**. As shown in FIG. **9**, the page allocation management table **270** includes a volume number column **271**, a page allocation method column **272**, and an allocating RAID group number column **273**.

[0082] The volume number column **271** stores the number for identifying each virtual volume in the same manner as the volume number column **251** in the virtual volume management table **250**. The page allocation method column **272** stores information indicating the page allocation method used to allocate a page(s) of the relevant RAID group. The normal allocation method or the D2T allocation method can be indicated as an example of the page allocation method.

[0083] The normal allocation method is to allocate pages sequentially from a plurality of RAID groups to the relevant segment of the virtual volume. For example, if data stored in the virtual volume is accessed by the host system **10**, but it is not a backup object to be backed up to the tape storage media, the normal allocation method is set. The D2T allocation method is to allocate pages from a fixed RAID group to the relevant segment of the virtual volume. For example, if data stored in the virtual volume is not accessed by the host system **10** and is a backup object to be backed up to the tape storage media, the D2T allocation method is set. The information stored in the page allocation method column **272** is set by, for example, the operator via the management terminal **50** described earlier.

[0084] The allocating RAID group number column **273** stores the RAID group number designated when allocating pages from a fixed RAID group to each virtual volume. For example, in the page allocation management table **270**, the page allocation method for the volume number **2** is set to D2T and the allocating RAID group number is set to 100. In this case, pages are allocated from only the RAID group whose RAID group number is 100.

[0085] When the storage apparatus **20** restores data by using the tape storage media **34** as described earlier, it allocates an arbitrary page(s) from the storage pool to the relevant segment of the virtual volume which is an object to be restored. When allocating an arbitrary page(s) of the storage pool to the virtual volume, the storage apparatus **20** executes the page allocation processing depending on whether or not the virtual volume is a backup object to be backed up to the tape storage media.

[0086] Whether the virtual volume is a backup object or not can be judged as described above by referring to the information stored in the page allocation method column **272** in the page allocation management table **270** shown in FIG. **9**. If it is determined that the virtual volume which is an object to be restored is a backup object volume (D2T), the storage apparatus **20** allocates pages from a fixed RAID group to the virtual volume.

[0087] When a page(s) is allocated to the virtual volume from a RAID group, each piece of information stored in the whether-allocated-or-not column **253**, the pool number column **254**, and the page number column **256** associated with the relevant segment number **252** in the virtual volume management table **250** shown in FIG. **7**. Specifically speaking, the information stored in the whether-allocated-or-not column **253** is updated to YES and data stored in the pool number column **254** and the page number column **256** are updated to the pool number and the page number which are allocated to each segment.

[0088] Furthermore, if the storage apparatus **20** refers to the page allocation management table **270** shown in FIG. **9** and determines that the virtual volume which is an object to be restored is not a backup object volume, the storage apparatus **20** allocates pages in consideration of the continuity in the bitmap **341** stored in the tape storage media **34**. Specifically speaking, the storage apparatus **20** judges the continuity in the bitmap **341** in the tape storage media **34**, selects the same RAID group for consecutive areas, and allocates pages from that same RAID group to the consecutive areas. When allocating pages from the same RAID group, a least common multiple of the data transfer unit to the tape management apparatus **30** and the striping size of the RAID group can be used as the allocation threshold as described earlier.

4. Details of Actions of Storage Apparatus

[0089] Next, the details of actions of the storage apparatus **20** will be explained with reference to FIG. **10** to FIG. **16**. Firstly, the page allocation processing for a data backup by the storage apparatus **20** will be explained with reference to FIG. **10** to FIG. **12**. Incidentally, the processing subject of various processing will be explained as a program in the following description. However, needless to say, the controller **21** for the storage apparatus **20** actually executes the processing according to that program.

[0090] FIG. **10** is a flowchart illustrating processing for judging the page allocation method for the data backup processing. As shown in FIG. **10**, the backup/restoration control program **204** firstly initializes a copy pointer (S**102**). The copy pointer is a pointer indicating a specified position of the relevant virtual volume. The virtual volume is divided by the pointer into an arbitrary size. The backup/restoration control program **204** sequentially copies data stored in the virtual volume which is divided by the pointer.

[0091] Next, the backup/restoration control program **204** counts the number of page-allocated blocks of the virtual volume (primary virtual volume) which is a backup object within a transfer range (S**104**). The transfer range herein used means a specified range sectioned by the above-described pointer. The backup/restoration control program **204** can rec-

ognize whether or not data which is a backup object is stored within the transfer range, by counting the number of page-allocated blocks in step S104.

[0092] The backup/restoration control program 204 judges whether the number of page-allocated blocks of the primary virtual volume within the transfer range is 0 or not (S106). If it is determined in step S106 that the number of page-allocated blocks of the primary virtual volume within the transfer range is 0, the backup/restoration control program 204 executes processing in step S116 and subsequent steps described below.

[0093] If it is determined in step S106 that the number of page-allocated blocks of the primary virtual volume within the transfer range is not 0, the thin provisioning control program 203 then judges whether the page allocation method for the secondary virtual volume to which the relevant data is backed up is D2T or not (S108). In step S108, the thin provisioning control program 203 refers to the information stored in the page allocation method column 272 corresponding to the relevant volume in the page allocation management table 270 shown in FIG. 9 and then judges whether the allocation method is D2T or normal. The case where the allocation method is D2T is when the virtual volume is a backup object volume to be backed up to the tape storage media 34 as described earlier.

[0094] If it is judged in step S108 that the page allocation method for the virtual volume is normal, the thin provisioning control program 203 executes normal page allocation processing (S110). The normal page allocation processing executed in step S110 is processing for allocating arbitrary pages sequentially from the plurality of RAID groups. The normal page allocation processing in step S110 will be explained later in detail.

[0095] If it is determined in step S108 that the page allocation method for the virtual volume is D2T, the thin provisioning control program 203 executes D2T page allocation processing (S112). In step S112, pages are allocated from a fixed RAID group from among the plurality of RAID groups. The D2T page allocation processing in step S112 will be explained later in detail.

[0096] After executing the page allocation processing in step S110 or step S112, the backup/restoration control program 204 copies the data, which is stored in the pool volume allocated to the primary virtual volume, to the relevant page (s) of the pool volume allocated to the secondary virtual volume (S114). Then, the backup/restoration control program 204 increments the copy pointer and moves the pointer from the current copy range to the next copy range (S116).

[0097] The backup/restoration control program 204 judges whether the copy pointer has reached the end of the virtual volume (primary virtual volume) (S118). In step S118, whether the data backup to the secondary virtual volume has terminated or not is judged based on whether or not the copy pointer has reached the end of the primary virtual volume. If it is determined in step S118 that the copy pointer has reached the end of the primary virtual volume, the backup/restoration control program 204 terminates the processing. If it is determined in step S118 that the copy pointer has not reached the end of the primary virtual volume, the backup/restoration control program 204 repeats executing the processing in step S104 and subsequent steps.

[0098] Next, the details of the normal page allocation processing in step S100 of FIG. 10 will be explained with reference to FIG. 11. FIG. 11 is a flowchart illustrating the details of the normal page allocation processing. As shown in FIG. 11, the thin provisioning control program 203 selects a RAID group with the largest number of remaining pages, excluding the RAID group from which a page(s) was allocated last time; and then allocates one page from the selected RAID group to the secondary virtual volume (S122).

[0099] Subsequently, the thin provisioning control program 203 judges whether as many pages as requested by the host system 20 have been allocated or not (S124). If it is determined in step S124 that as many pages as requested have been allocated, the thin provisioning control program 203 terminates the processing. If it is determined in step S124 that as many pages as requested have not been allocated, the thin provisioning control program 203 repeats the processing in step S122.

[0100] Next, the details of the D2T page allocation processing in step S112 of FIG. 10 will be explained with reference to FIG. 12. FIG. 12 is a flowchart illustrating the details of the D2T page allocation processing. As shown in FIG. 12, the thin provisioning control program 203 judges whether or not a fixed RAID group has been decided for the virtual volume which is a data write object (S132). In step S132, the thin provisioning control program 203 refers to the information stored in the allocating RAID group number column 273 for the relevant volume in the page allocation management table 270 shown in FIG. 9 and judges whether or not the fixed RAID group has been decided for the virtual volume which is the data write object.

[0101] If it is determined in step S132 that the RAID group has not been decided, the thin provisioning control program 203 selects a RAID group with a smallest D2T transfer quantity and with a largest number of remaining pages, from among the plurality of RAID groups (S134). In step S134, the thin provisioning control program 203 refers to the information stored in the D2T transfer quantity column 264 and the numberof-remaining-pages column 263 in the pool volume management table 260 shown in FIG. 8 and selects a RAID group which satisfies the above-described conditions.

[0102] If it is determined in step S132 that the RAID group has already been decided, the thin provisioning control program 203 executes processing in step S138 and subsequent steps described below. Then, the thin provisioning control program 203 increments the D2T transfer quantity 264 in the pool volume management table 260 shown in FIG. 8 (S136).

[0103] Subsequently, the thin provisioning control program 203 allocates a page(s) from the already selected RAID group or the RAID group selected in step S134 to a write area (segment) of the virtual volume which is the backup object (S138).

[0104] Next, the thin provisioning control program 203 judges whether the number of pages which need to be allocated to the write area (segment) of the virtual volume is insufficient or not (S140). If it is determined in step S140 that the number of pages is insufficient, the thin provisioning control program 203 decrements the D2T transfer quantity 264 in the pool volume management table 260.

[0105] If it is determined in step S140 that the number of pages is not insufficient, the thin provisioning control program 203 judges whether as many pages as requested by the host system 20 have been allocated or not (S144). If it is determined in step S144 that as many pages as requested have been allocated, the thin provisioning control program 203 terminates the processing. If it is determined in step S144 that as many pages as requested have not been allocated, the thin

provisioning control program **203** repeats the processing in step S**138** and subsequent steps.

[0106] If it is determined in step S**140** that the number of pages is insufficient, the thin provisioning control program **203** decrements the D2T transfer quantity (S**142**), returns to the processing in step S**204**, and selects another RAID group that satisfies the conditions in step S**134**.

[0107] Next, the processing for backing up data to the tape storage media **34** will be explained with reference to FIG. **13**. As shown in FIG. **13**, the backup/restoration control program **204** firstly writes the tape management information to a tape (tape storage media **34**) of the tape management apparatus **30** (S**202**). Then, the backup/restoration control program **204** writes the bitmap to the tape storage media **34** to which the tape management information was written in step **202** (S**204**).

[0108] Subsequently, the backup/restoration control program **204** initializes a copy pointer for the virtual volume (secondary volume) which is a backup object to be backed up to the tape storage media **34** (S**206**). The backup/restoration control program **204** then judges whether a page indicated by the copy pointer has been allocated or not (S**208**).

[0109] If it is determined in step S**208** that the page indicated by the copy pointer has been allocated, the backup/restoration control program **204** reads data stored in the relevant page of the pool volume and writes the data to the tape storage media. If it is determined in step S**208** that the page indicated by the copy pointer has not been allocated, the backup/restoration control program **204** executes processing in step S**212** described below.

[0110] Then, the backup/restoration control program **204** increments the copy pointer (S**212**). The backup/restoration control program **204** judges whether the copy pointer has reached the end of the virtual volume (secondary virtual volume) or not (S**214**). If it is determined in step S**214** that the copy pointer has reached the end of the secondary virtual volume, the backup/restoration control program **204** writes the termination information such as the backup terminate date and time and the number of volumes to the tape storage media **34** (S**216**).

[0111] Next, the page allocation processing for data restoration will be explained with reference to FIG. **14** and FIG. **15**. FIG. **14** is a flowchart illustrating processing for judging the page allocation method for data restoration. As shown in FIG. **14**, the backup/restoration control program **204** firstly reads the tape management information from the tape storage media **34** (S**302**). Then, the backup/restoration control program **204** reads the bitmap information from the tape storage media **34** (S**304**). The backup/restoration control program **204** initializes a copy pointer for the virtual volume which is an object to be restored (S**305**).

[0112] Next, the backup/restoration control program **204** counts the number of blocks whose bitmap is ON within the data transfer range in the bitmap read in step S**304** (S**308**). The backup/restoration control program **204** judges whether the number of blocks whose bitmap is ON within the data transfer range, which was counted in step S**308**, is 0 or not (S**310**). If it is determined in step S**310** that the number of blocks whose bitmap is ON within the transfer range is not 0, the thin provisioning control program **203** judges whether the page allocation method for the virtual volume (secondary volume) is D2T or not (S**312**). If it is determined in step S**310** that the number of blocks whose bitmap is ON within the transfer range is 0, the thin provisioning control program **203** executes processing in step S**320** described below.

[0113] In step S**312**, the thin provisioning control program **203** refers to the information stored in the page allocation method column **272** in the page allocation management table **270** shown in FIG. **9** and judges whether the allocation method is D2T or normal. The case where the allocation method is D2T is when the virtual volume is the backup object to be backed up to the tape storage media **34** as described earlier.

[0114] If it is determined in step S**312** that the page allocation method for the secondary virtual volume is not D2T, the thin provisioning control program **203** executes T2D page allocation processing (S**314**). The T2D page allocation processing executed in step S**314** will be explained later in detail. If it is determined in step S**312** that the page allocation method for the secondary virtual volume is D2T, the thin provisioning control program **203** executes the D2T page allocation processing (S**316**). Since the D2T page allocation processing executed in step S**316** is the same as the processing shown in FIG. **12**, its detailed explanation has been omitted.

[0115] Then, the thin provisioning control program **203** reads the data from the tape storage media **34** and copies the data to the page(s) allocated in step S**314** or step S**316** (S**318**). Next, the thin provisioning control program **203** increments the copy pointer (S**320**). The thin provisioning control program **203** then judges whether or not the copy pointer has reached the end of the volume (S**322**).

[0116] If it is determined in step S**322** that the copy pointer has reached the end of the volume, the thin provisioning control program **203** terminates the processing. If it is determined in step S**322** that the copy pointer has not reached the end of the volume, the thin provisioning control program **203** executes the processing in step S**308** and subsequent steps.

[0117] Next, the details of the T2D page allocation processing in step S**314** of FIG. **14** will be explained with reference to FIG. **15**. FIG. **15** is a flowchart illustrating the details of the T2D page allocation processing. As shown in FIG. **15**, the thin provisioning control program **203** selects a RAID group with a largest number of remaining pages from among the RAID groups and allocates one page from the selected RAID group to the virtual volume which is an object to be restored (S**332**).

[0118] The thin provisioning control program **203** judges whether as many pages as requested by the host system **20** have been allocated or not (S**334**). If it is determined in step S**334** that as many pages as requested have been allocated, the thin provisioning control program **203** terminates the processing. If it is determined in step S**334** that as many pages as requested have not been allocated, the thin provisioning control program **203** judges whether the next segment number of the relevant volume is consecutive or not (S**336**).

[0119] If it is determined in step S**336** that the next segment number of the relevant volume is consecutive, the thin provisioning control program **203** judges whether the least common multiple of the amount of data that can be transferred within average response time and the striping size of the RAID group exceeds the threshold for allocation from the same RAID group or not (S**338**). An example of the threshold for allocation from the same RAID group in step S**338** can be the least common multiple of the amount of data that can be transferred within average response time of the tape management apparatus **30** relative to the host system **10** and the striping size of the RAID group as described earlier.

[0120] If it is determined in step S**336** that the next segment number of the relevant volume is not consecutive, the thin

provisioning control program **203** repeats executing the processing in step S**332** and subsequent steps. If it is determined in step S**338** that the least common multiple of the amount of data that can be transferred within average response time and the striping size of the RAID group does not exceed the threshold for allocation from the same RAID group, the thin provisioning control program **203** selects the same RAID group and allocates one page from the selected RAID group (S**340**). If it is determined in step S**338** that the least common multiple of the amount of data that can be transferred within average response time and the striping size of the RAID group exceeds the threshold for allocation from the same RAID group, the thin provisioning control program **203** repeats executing the processing in step S**332** and subsequent steps.

[0121] After one page is allocated from the RAID group in step S**340**, the thin provisioning control program **203** judges whether or not the number of pages for the secondary virtual volume is insufficient (S**342**). If it is determined in step S**342** that the number of pages is insufficient, the thin provisioning control program **203** repeats executing processing in step S**334** and subsequent steps. If it is determined in step S**342** that the number of pages is not insufficient, the thin provisioning control program **203** repeats executing the processing in step S**332** and subsequent steps.

[0122] Next, processing for restoring data from the secondary virtual volume to the primary virtual volume will be explained with reference to FIG. **16**. As shown in FIG. **16**, the backup/restoration control program **204** firstly initializes a copy pointer for the secondary virtual volume (S**402**).

[0123] Subsequently, the backup/restoration control program **204** counts the number of page-allocated blocks of the secondary virtual volume within the transfer range (S**404**). The backup/restoration control program **204** judges whether the number of page-allocated blocks of the secondary virtual volume within the transfer range, which was counted in step S**404**, is 0 or not (S**406**). If it is determined in step S**406** that the number of page-allocated blocks of the secondary virtual volume within the transfer range is 0, the backup/restoration control program **204** executes processing in step S**412** and subsequent steps described below.

[0124] If it is determined in step S**406** that the number of page-allocated blocks of the secondary virtual volume within the transfer range is not 0, the thin provisioning control program **203** executes the normal page allocation processing (S**408**). Since the normal page allocation processing in step S**408** is the same as the normal page allocation processing shown in FIG. **11**, its detailed explanation has been omitted.

[0125] Then, the backup/restoration control program **204** reads the data of the secondary virtual volume and copies the data to the page(s) allocated in step S**408** (S**410**). Subsequently, the backup/restoration control program **204** increments the copy pointer (S**412**). The backup/restoration control program **204** then judges whether or not the copy pointer has reached the end of the volume (S**414**).

[0126] If the backup/restoration control program **204** determines in step S**414** that the copy pointer has reached the end of the volume, it terminates the processing. If the backup/restoration control program **204** determines in step S**414** that the copy pointer has not reached the end of the volume, it executes the processing in step S**404** and subsequent steps.

4. Advantageous Effects of this Embodiment

[0127] According to this embodiment, the storage apparatus **20** stores data, for which a write request is issued from the host system **10**, in a virtual volume and allocates a predetermined area in a plurality of storage devices to a data storage area in the virtual volume storing the data. If the virtual volume is a backup object to be backed up to tape storage media, the storage apparatus **20** allocates a predetermined area in a specified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume; and if the virtual volume is not a backup object to be backed up to the tape storage media, the storage apparatus allocates a predetermined area in an unspecified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume. Therefore, when backing up the data stored in the virtual volume to the tape storage media, it is unnecessary to access the unspecified storage device at the same time. As a result, it is possible to prevent the occurrence of access conflicts to the storage devices and enhance the performance to transfer data to the tape storage media.

5. Other Embodiments

[0128] Preferred embodiments of this invention have been described in detail with reference to the attached drawings. However, this invention is not limited only to these embodiments. It is apparent that a person with an ordinary skill in the art to which the invention pertains could easily think of various change examples or modification examples within the category of technical ideas described in the scope of claims. It is understood as a matter of course that such change examples or modification examples also belong to the technical scope of this invention.

[0129] The aforementioned embodiment has described the case where the secondary virtual volume is backed up to the tape storage media. However, the invention is not limited to this example. For example, if the primary virtual volume is a backup object to be backed up to the tape storage media, an arbitrary page(s) of a specified RAID group constituting a first storage pool associated with the primary virtual volume may be allocated from the first storage pool to the relevant segment of the primary virtual volume.

[0130] Furthermore, in the above-described embodiment, a page allocation method for a pool volume is changed depending on whether or not a plurality of virtual volumes are backed up to the tape storage media. However, the invention is not limited to this example. For example, the page allocation method may be also changed for each virtual volume when virtual volumes to be used for a different purpose exist with other virtual volumes, for example, in a case where some of the virtual volumes are backed up to the tape storage media.

[0131] For example, steps in the processing by the storage apparatus **20** as described in this specification do not necessarily have to be executed chronologically in the order described in the relevant flowchart. In other words, each step in the processing by the storage apparatus **20** may be executed in different processing or those steps may be executed in parallel.

[0132] Furthermore, it is possible to create a computer program for having the hardware such as the CPU, ROM, RAM, etc. contained in the storage apparatus **20** fulfill functions equivalent to those of each component of the storage apparatus **20** described earlier. It is also possible to provide a storage medium in which such a computer program is stored.

INDUSTRIAL APPLICABILITY

[0133] The present invention can be used in a computer system for backing up data, which is stored in a virtual vol-

ume(s) of a storage apparatus, to tape storage media and restoring the data from the tape storage media upon the occurrence of a disk failure.

REFERENCE SIGNS LIST

[0134] **10** Host system
[0135] **20** Storage apparatus
[0136] **21** Controller
[0137] **22** Storage device
[0138] **30** Tape management apparatus
[0139] **31** Carrier mechanism
[0140] **32** Tape drive
[0141] **34** Tape storage media
[0142] **40** Network
[0143] **50** Management terminal

1. A storage apparatus connected via a network to a host system requesting writing of data and to a tape management apparatus for managing tape storage media for storing the data, the storage apparatus comprising:

a storage unit for storing data, for which a write request is issued by the host system, in a virtual volume; and

an allocation unit for allocating a predetermined area in a plurality of storage devices to a data storage area in the virtual volume in which the data is stored by the storage unit;

wherein if the virtual volume is a backup object to be backed up to the tape storage media, the allocation unit allocates a predetermined area in a specified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume; and

wherein if the virtual volume is not a backup object to be backed up to the tape storage media, the allocation unit allocates a predetermined area in an unspecified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume.

2. The storage apparatus according to claim **1**, wherein the allocation unit allocates the storage device, to which a smallest number of virtual volumes are already allocated, to the data storage area in the virtual volume.

3. The storage apparatus according to claim **1**, further comprising a backup object management table for associating identification information for identifying the virtual volume with a flag for judging whether or not the virtual volume is a backup object to be backed up to the tape storage media,

wherein if the allocation unit determines, according to the backup object flag associated with the identification information about the virtual volume allocated to the storage device, that the virtual volume is the backup object to be backed up to the tape storage media, it allocates a predetermined area in a specified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume.

4. The storage apparatus according to claim **1**, wherein the plurality of storage devices constitute a plurality of RAID (Redundant Array of Inexpensive Disks) groups,

wherein if the virtual volume is a backup object to be backed up to the tape storage media, the allocation unit allocates a predetermined area in the storage devices constituting a specified RAID group, from among the plurality of RAID groups, to the data storage area in the virtual volume; and

wherein if the virtual volume is not a backup object to be backed up to the tape storage media, the allocation unit allocates a predetermined area in the storage devices constituting an unspecified RAID group, from among the plurality of RAID groups, to the data storage area in the virtual volume.

5. The storage apparatus according to claim **4**, wherein the allocation unit receives a command from the host system to restore data from the tape storage media to the virtual volume, and

wherein if data storage areas in the virtual volume stored in the tape storage media are consecutive, the allocation unit allocates a predetermined area in the storage devices constituting the same RAID group to the data storage areas.

6. The storage apparatus according to claim **5**, wherein the allocation unit sets a least common multiple of an amount of data that can be transferred within average response time of the tape management apparatus relative to the host system, and an amount of data in the storage devices constituting the RAID group, as a threshold for allocation from the same RAID group.

7. A storage management method using a storage apparatus connected via a network to a host system requesting writing of data and to a tape management apparatus for managing tape storage media for storing the data, the storage management method comprising the steps of:

storing data, for which a write request is issued by the host system requesting writing of the data, in a virtual volume;

if the virtual volume storing the data is a backup object to be backed up to the tape storage media, allocating a predetermined area in a specified storage device, from among the plurality of storage devices, to a data storage area in the virtual volume; and

if the virtual volume is not a backup object to be backed up to the tape storage media, allocating a predetermined area in an unspecified storage device, from among the plurality of storage devices, to the data storage area in the virtual volume.

* * * * *