

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4681247号
(P4681247)

(45) 発行日 平成23年5月11日(2011.5.11)

(24) 登録日 平成23年2月10日(2011.2.10)

(51) Int.Cl.

F I

G O 6 F 3/06 (2006.01)

G O 6 F 12/00 (2006.01)

G O 6 F 3/06 3 O 1 Z

G O 6 F 3/06 3 O 4 F

G O 6 F 3/06 5 4 O

G O 6 F 12/00 5 O 1 A

請求項の数 3 (全 18 頁)

(21) 出願番号 特願2004-114189 (P2004-114189)
 (22) 出願日 平成16年4月8日(2004.4.8)
 (65) 公開番号 特開2005-301499 (P2005-301499A)
 (43) 公開日 平成17年10月27日(2005.10.27)
 審査請求日 平成19年3月2日(2007.3.2)

(73) 特許権者 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74) 代理人 100075513
 弁理士 後藤 政喜
 (74) 代理人 100084537
 弁理士 松田 嘉夫
 (74) 代理人 100114236
 弁理士 藤井 正弘
 (72) 発明者 永田 幸司
 神奈川県小田原市中里322番2号 株式
 会社日立製作所 R A I Dシステム事業部
 内

審査官 木村 貴俊

最終頁に続く

(54) 【発明の名称】 ディスクアレイ装置及びディスクアレイ装置の制御方法

(57) 【特許請求の範囲】

【請求項 1】

上位装置に接続され、前記上位装置からデータを受ける上位インタフェースと、
 前記上位インタフェースに接続され、前記上位装置との間でやり取りされるデータ、及び前記上位装置との間でやり取りされるデータに関する制御情報を保存するメモリと、
 前記メモリに接続され、前記上位装置との間でやり取りされるデータを、前記メモリに読み書きするように制御する複数のディスクインタフェースと、
 前記複数のディスクインタフェースに接続され、前記複数のディスクインタフェースの制御のもとに、前記上位装置から送られたデータが格納される複数のディスクドライブと、

前記複数のディスクドライブの記憶領域を用いて生成される第1の論理ボリュームへのデータの読み書きを制御し、前記第1の論理ボリュームに格納された過去のデータを世代毎の差分データとして第2の論理ボリュームに書き込むように制御し、前記第2の論理ボリュームに格納されている前記世代毎の差分データの関係を管理するスナップショット管理テーブルを前記メモリの管理領域に設けることによって前記差分データを管理する制御プロセッサと、を有し、

前記制御プロセッサは、

前記第2の論理ボリュームに格納されたデータ量を管理し、

前記第2の論理ボリュームに格納されたデータ量が前記第2の論理ボリュームの容量の第1の割合を超えた場合、前記管理領域に空きがあるか否かを判定し、前記管理領域に空

きがある場合、前記第2の論理ボリュームの容量を増加するための条件が満たされると判定し、前記第2の論理ボリュームの容量を増加し、前記管理領域にてスナップショット管理テーブルの容量を増加し、

前記第2の論理ボリュームに格納されたデータ量が前記第2の論理ボリュームの容量の第2の割合より少なくなった場合、前記第2の論理ボリュームの容量を減少させ、該減少させた前記第2の論理ボリュームの容量を他用途に使用可能とするディスクアレイ装置。

【請求項2】

前記制御プロセッサは、

前記第2の論理ボリュームとして使用可能な論理ボリュームが登録されるプール管理テーブルと、前記第2の論理ボリュームに追加可能な論理ボリュームが登録されるプール追加対象管理テーブルとを管理し、

前記プール管理テーブルをグループに分けて設けることによって、前記第1の論理ボリューム及び前記第2の論理ボリュームをグループ毎に管理し、

前記第2の論理ボリュームの容量を増加するための条件が満たされると判定された場合、前記プール追加対象管理テーブルを参照し、前記各グループに追加される論理ボリュームを選択し、

前記選択された論理ボリュームが追加された場合に、前記各グループの論理ボリュームの記憶容量が所定の最大量を超えるか否かを判定し、前記各グループの論理ボリュームの記憶容量が所定の最大量を超える場合、前記論理ボリュームの容量を増加せず、

一方、前記選択された論理ボリュームが追加された場合に、前記プール追加対象管理テーブルに登録される論理ボリュームの容量が所定値未満になるか否かを判定し、前記プール追加対象管理テーブルに登録される論理ボリュームの容量が所定値未満になる場合、前記論理ボリュームの容量を増加せず、

それ以外の場合に、前記プール管理テーブルに新たな論理ボリュームを追加することによって、前記第2の論理ボリュームの容量を増加する請求項1に記載のディスクアレイ装置。

【請求項3】

前記制御プロセッサは、前記第2の論理ボリュームに格納されたデータ量が第2の所定値より所定期間連続して少なくなり、かつ、前記第2の論理ボリュームを削除しても、前記第2の論理ボリュームに格納されるデータ量が前記第3の所定値より少ない場合、論理ボリュームを前記プール管理テーブルから削除することによって、前記第2の論理ボリュームの容量を減少させる請求項1又は2に記載のディスクアレイ装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ディスクアレイ装置に関し、特にスナップショットにおける差分LUの管理技術に関する。

【背景技術】

【0002】

コンピュータシステムにおいては、ストレージ装置に記憶されるデータをバックアップする場合など、元のデータを記憶した記憶ボリュームの複製を生成する場合がある、この場合、元のデータを記憶した記憶ボリュームと複製データを記憶した記憶ボリュームとは内容が一致し整合性が保たれる必要がある。そのためには複製の生成が完了するまでの間、元のデータを記憶した記憶ボリュームの更新を停止させる必要がある。しかし高可用性が求められるコンピュータシステムなどでは、元のデータを記憶した記憶ボリュームに対するアクセスを停止することができない場合がある。

【0003】

そこで、元のデータを記憶した記憶ボリュームと複製データを記憶した記憶ボリュームとの間である時点での整合性を保ち、ある時点の後に元のデータが更新されたとしてもある時点の元のデータを参照できるようにするスナップショットと呼ばれる技術が開発され

10

20

30

40

50

ている。

【 0 0 0 4 】

スナップショットの技術によれば、整合性を保つべき時点の後に元のデータが更新される場合には、整合性を保つべき時点のデータを別の記憶ボリュームに記憶するようにする。つまり元のデータは、更新されなければそのままであり、更新される場合には整合性を保つべき時点のデータが別の記憶ボリュームに記憶される。

【特許文献 1】特開 2 0 0 1 - 3 0 6 4 0 7 号公報

【発明の開示】

【発明が解決しようとする課題】

【 0 0 0 5 】

しかし、スナップショット機能は、正 L U の容量よりも差分 L U の容量を減らせることが利点であるが、ホストからのデータの書き込みが多くなると差分データが多くなる。そして、差分 L U の使用量が 1 0 0 % になるとスナップショットを維持できなくなる。このため、事前に差分 L U として多くの L U を設定すると、データの書き込みが多く発生しなかった場合に他の用途に使用することができない。

【 0 0 0 6 】

本発明は、スナップショットにおいて使用される差分 L U の容量を増減することによって、ストレージ装置のリソースを有効に活用することを目的とする。

【課題を解決するための手段】

【 0 0 0 7 】

上位装置に接続され、前記上位装置からデータを受ける上位インタフェースと、前記上位インタフェースに接続され、前記上位装置との間でやり取りされるデータ、及び前記上位装置との間でやり取りされるデータに関する制御情報を保存するメモリと、前記メモリに接続され、前記上位装置との間でやり取りされるデータを、前記メモリに読み書きするように制御する複数のディスクインタフェースと、前記複数のディスクインタフェースに接続され、前記複数のディスクインタフェースの制御のもとに、前記上位装置から送られたデータが格納される複数のディスクドライブと、前記複数のディスクドライブの記憶領域を用いて生成される第 1 の論理ボリュームへのデータの読み書きを制御し、前記第 1 の論理ボリュームに格納された過去のデータを世代毎の差分データとして第 2 の論理ボリュームに書き込むように制御し、前記第 2 の論理ボリュームに格納されている前記世代毎の差分データの関係を管理するスナップショット管理テーブルを前記メモリの管理領域に設けることによって前記差分データを管理する制御プロセッサと、を有し、前記制御プロセッサは、前記第 2 の論理ボリュームに格納されたデータ量を管理し、前記第 2 の論理ボリュームに格納されたデータ量が前記第 2 の論理ボリュームの容量の第 1 の割合を超えた場合、前記管理領域に空きがあるか否かを判定し、前記管理領域に空きがある場合、前記第 2 の論理ボリュームの容量を増加するための条件が満たされると判定し、前記第 2 の論理ボリュームの容量を増加し、前記管理領域にてスナップショット管理テーブルの容量を増加し、前記第 2 の論理ボリュームに格納されたデータ量が前記第 2 の論理ボリュームの容量の第 2 の割合より少なくなった場合、前記第 2 の論理ボリュームの容量を減少させ、該減少させた前記第 2 の論理ボリュームの容量を他用途に使用可能とするディスクアレイ装置。

【発明の効果】

【 0 0 0 8 】

本発明によると、プール領域に登録されている L U の容量を増減することによって、ストレージ装置のリソースを有効に活用することができる。

【発明を実施するための最良の形態】

【 0 0 0 9 】

以下、本発明の実施の形態を図面を参照して説明する。

【 0 0 1 0 】

図 1 は、本発明の実施の形態のディスクアレイ装置の構成を示すブロック図である。

10

20

30

40

50

【 0 0 1 1 】

本発明の実施の形態のディスクアレイ装置 1 は、ディスクアレイコントローラ 1 0 及びディスク 2 0 を含んで構成されている。また、ディスクアレイ装置 1 は、S A N 2 を介して複数のホスト 3 に接続されており、L A N 4 を介して管理用端末装置 5 に接続されている。

【 0 0 1 2 】

ディスクアレイコントローラ 1 0 は制御プログラム 1 0 3 の動作によって、ディスク 2 0 に対するデータの入出力を制御する。また、ディスク 2 0 によって R A I D (Redundant Array of Independent Disks) が構成されており、記憶されるデータに冗長性を持たせている。このため、ディスクの一部に障害が生じて、記憶されたデータが消失しないようになっている。

10

【 0 0 1 3 】

ディスクアレイコントローラ 1 0 には、C P U 1 0 1、メモリ 1 0 2、データ転送コントローラ 1 0 4、フロントエンドインターフェース 1 0 5、バックエンドインターフェース 1 0 6、キャッシュメモリ 1 0 7、及び L A N インターフェース 1 0 8 が設けられている。

【 0 0 1 4 】

メモリ 1 0 2 には制御プログラム 1 0 3 (図 2 参照) が記憶されており、C P U 1 0 1 が制御プログラム 1 0 3 を呼び出して実行することによって各種処理が行われる。

【 0 0 1 5 】

20

データ転送コントローラ 1 0 4 は、C P U 1 0 1、フロントエンドインターフェース 1 0 5、バックエンドインターフェース 1 0 6、及びキャッシュメモリ 1 0 7 の間でデータを転送する。

【 0 0 1 6 】

フロントエンドインターフェース 1 0 5 は、S A N 2 に対するインターフェースであって、例えば、ファイバチャネルプロトコルによって、ホスト 3 との間でデータや制御信号を送受信する。

【 0 0 1 7 】

バックエンドインターフェース 1 0 6 は、ディスク 2 0 に対するインターフェースであって、例えば、ファイバチャネルプロトコルによって、ディスク 2 0 との間でデータや制御信号を送受信する。

30

【 0 0 1 8 】

キャッシュメモリ 1 0 7 には、フロントエンドインターフェース 1 0 5 とバックエンドインターフェース 1 0 6 との間で送受信されるデータが一時的に記憶されるキャッシュが設けられている。

【 0 0 1 9 】

すなわち、データ転送コントローラ 1 0 4 は、S A N 4 を介してディスクに読み書きされるデータをインターフェース 1 0 5、1 0 6 間で転送する。さらに、これらのディスクに読み書きされるデータをキャッシュメモリ 1 0 7 に転送する。

【 0 0 2 0 】

40

L A N インターフェース 1 0 8 は、L A N 4 に対するインターフェースであって、例えば、T C P / I P プロトコルによって、管理用端末装置 5 との間でデータや制御信号を送受信することができる。

【 0 0 2 1 】

S A N 2 は、例えばファイバチャネルプロトコルのような、データの転送に適するプロトコルで通信可能なネットワークである。

【 0 0 2 2 】

ホスト 3 は、C P U、メモリ、記憶装置、インターフェース、入力装置及び表示装置が備わるコンピュータ装置であり、ディスクアレイ装置 1 から提供されるデータを利用して、データベースサービスやウェブサービス等を利用可能にする。

50

【 0 0 2 3 】

L A N 4 は、ディスクアレイ装置 1 の管理用に用いられるもので、例えば、T C P / I P プロトコルによって、コンピュータ間でデータや制御情報を通信可能であり、例えばイーサネット（登録商標、以下同じ）が用いられる。

【 0 0 2 4 】

管理用端末装置 5 は、C P U、メモリ、記憶装置、インターフェース、入力装置及び表示装置が備わるコンピュータ装置である。管理用端末装置 5 では管理プログラムが動作しており、該管理プログラムによってディスクアレイ装置 1 の動作状態を把握し、ディスクアレイ装置 1 の動作を制御する。なお、管理用端末装置 5 では w e b ブラウザ等のクライアントプログラムが動作しており、ディスクアレイ装置 1 から C G I（Common Gateway I 10 nterface）等によって供給される管理プログラムによってディスクアレイ装置 1 の動作を制御するようにしてもよい。

【 0 0 2 5 】

図 2 は、本発明の実施の形態の制御プログラム 1 0 3 の説明図である。

【 0 0 2 6 】

ホスト 3 の通常 I / O 処理プログラム 3 0 1 から送られたデータ入出力要求は、ディスクアレイ装置 1 の制御プログラム 1 0 3 の R / W コマンド解析プログラム 1 1 1 によって解析され、スナップジョブプログラム 1 2 1 に送られる。

【 0 0 2 7 】

スナップジョブプログラム 1 2 1 は、正 L U に対するデータ書き込み要求を受信すると 20 、正 L U 内の更新前のデータを差分 L U に複写し、この複写後に正 L U の内容を更新する。更に、スナップジョブプログラム 1 2 1 は、データが更新された正 L U 内のブロックに対応する仮想 L U 内のブロックを、スナップショット生成要求の受信時点で正 L U のデータ（すなわち更新前のデータ）が格納されている差分 L U 上のブロックと対応付けるようスナップショット管理テーブル（差分情報管理ブロック 2 0 4 ）を更新する。

【 0 0 2 8 】

また、スナップリストアジョブプログラム 1 2 2 は、仮想 L U に対するアクセス要求を受信すると、スナップショット管理テーブルを参照して、仮想 L U のブロックと対応付けられている正 L U のブロック又は差分 L U のブロックにアクセスする。このようにして、ディスクアレイ装置 1 はスナップショットイメージを提供することが可能になる。そして 30 、ホスト 3 が、通常 I / O 処理プログラム 3 0 1 によって、仮想 L U にアクセスすることによって、スナップショット生成要求発行時点での正 L U 内の情報を利用することができる。

【 0 0 2 9 】

また、通常 I / O 処理プログラム 3 0 1 から送られた制御コマンドは、その他コマンド解析プログラム 1 1 2 によって解析され、構成情報制御プログラム 1 7 0 に送られる。

【 0 0 3 0 】

構成情報制御プログラム 1 7 0 のペア情報管理プログラム 1 7 1 は、スナップショット生成要求を受信すると、まずスナップショット管理テーブルに新しい仮想 L U の識別情報を登録し、ペア情報管理テーブルにスナップショットのペアを新規に登録し、差分ビット 40 マップ 2 0 2、正 L U アドレステーブル 2 0 3、差分情報管理ブロック 2 0 4 を確保、初期化する。この仮想 L U のブロックは、最初はスナップショット管理テーブルによって、正 L U のブロックと一対一で対応付けられている。

【 0 0 3 1 】

プール L U 管理プログラム 1 7 2 は、後述するように、プール領域に登録された L U の追加及び削除を管理する。

【 0 0 3 2 】

プール管理プログラム 1 5 0 は、スナップショット差分管理テーブル（差分情報管理ブロック 2 0 4 ）の領域の確保、解放及び空きキューへの遷移等の管理を行う。

【 0 0 3 3 】

10

20

30

40

50

WEBプログラム160は、スナップショットで作成される各ペアの状態（ペア情報管理テーブル内の情報等からペアの障害有無）をWEBブラウザに提供する。

【0034】

ディスクアレイ装置1の制御プログラム103のRAIDマネージャプログラム131は、ホスト3のRAIDマネージャプログラム302と通信可能に接続されている。このRAIDマネージャプログラム121、302によって、スナップショットの生成やペア状態の変更などを行うことができる。

【0035】

また、DAMPインターフェースプログラム132は、仮想LUの削除処理等のディスクアレイの各種設定を行うユーザインターフェースを提供する。DAMPインターフェースプログラム132は、管理用端末装置5のDAMPプログラム501と通信可能に接続されている。このDAMPインターフェースプログラム132によって、管理用端末装置5のDAMPプログラム501との通信が行われて、ディスクアレイ装置1のRAIDの構成の管理、プールへの自動追加、削除の設定が行われる。

【0036】

図3は、本発明の実施の形態のスナップショットの管理方法の説明図である。

【0037】

正LU201は、通常の運用に供され、ホスト3からのデータ入出力の対象となる論理ユニット（P-VOL：Primary Volume）である。

【0038】

差分ビットマップ202は、キャッシュメモリ107の管理領域に設けられており、正LU201のブロック（例えば、64kバイト/ブロック）に対応するビットを有する。この差分ビットマップ202は、後述する正LUアドレステーブル203及び差分情報管理ブロック204と共にスナップショット管理テーブルを構成する。

【0039】

正LU201のブロックアドレスに対応する差分データが差分LU206に記録されている場合に、差分ビットマップ202中の正LU201に対応するビットが「1」となっている。従って、正LU201に対する書き込みが行われるときに、差分ビットマップ202を参照することによって、更新前データを差分LUにコピーする必要があるかを判定することができる（すなわち、ビットが「1」であれば、スナップショット生成時のデータが既に差分LU206に書き込まれているので、正LU201のデータを差分LU206にコピーする必要がない）。

【0040】

正LUアドレステーブル203は、キャッシュメモリ107の管理領域に設けられており、差分ビットマップ202のビットと対応して差分情報管理ブロック204のアドレスが記録されている。

【0041】

差分情報管理ブロック204は、差分LUと同じ容量を有し、キャッシュメモリ107の管理領域に設けられている。差分情報管理ブロック204は、差分LU206のブロック（例えば、64kバイト/ブロック）毎に区切られて、その各々に管理テーブルが設けられている。この管理テーブルには、差分LU206のブロックに対応する位置に記録された差分データがどの世代のスナップショットのデータであるかが記録されている。また、正LU201の当該ブロックに対応する他の差分データがあれば、その差分情報管理ブロック204上のアドレスへのリンク情報が記録されている。すなわち、差分情報管理ブロック204に記録されたアドレスをたどることによって、複数世代の差分データを参照することができる。

【0042】

なお、差分情報管理ブロック204の使用されていない領域は、空きキューとしてリンクが設定されている。そして、この空きキューの量はキャッシュメモリ107に設けられた空きキューカウンタ205によって管理されている。

【 0 0 4 3 】

差分ＬＵ２０６は、プール領域に登録されたＬＵによって構成される。この差分ＬＵ２０６は、スナップショット作成時点における正ＬＵ２０１のデータが複写されている。そして、差分ＬＵ２０６のデータがどの世代の差分データかは差分情報管理ブロック２０４の世代管理ビットマップによって知ることができる。

【 0 0 4 4 】

よって、正ＬＵにデータを書き込むときは、まず、差分ビットマップ２０２を参照して、更新前データを差分ＬＵにコピーする必要があるかを判定する。そして、差分ビットマップ２０２の対応ビットが「１」であれば、更新前データを差分ＬＵにコピーする必要がないと判定し、正ＬＵにデータを書き込む。一方、差分ビットマップ２０２の対応ビットが「０」であれば、更新前データを差分ＬＵにコピーした後に、正ＬＵにデータを書き込む。

10

【 0 0 4 5 】

そして、正ＬＵのブロックに対応する差分情報管理ブロック２０４のブロックに、新たに設定された差分データに対するリンクアドレスを設定する。そして、必要に応じて正ＬＵアドレステーブルに、差分情報管理ブロック２０４のブロックのアドレスを設定する。そして、差分ビットマップ２０２の対応ビットを「１」にし、差分ＬＵ２０６の当該更新前データを書き込んだアドレスに対応する差分情報管理ブロック２０４の世代管理ビットマップを設定する。さらに、差分情報管理ブロック２０４の空きキューを使用したので、空きキューカウンタ２０５を更新する。

20

【 0 0 4 6 】

また、仮想ＬＵ（Ｖ－ＶＯＬ：Virtual Volume）にアクセスするときには、正ＬＵアドレステーブル２０３を参照して、アクセス対象となる仮想ＬＵのブロックアドレス（正ＬＵのブロックアドレスに等しい）によって差分情報管理ブロック２０４のアドレスを特定して、差分情報管理ブロック２０４の当該アドレスの世代管理ビットマップによって、アクセス対象の世代の差分データがあるかを特定する。そして、所望の世代の差分データがあれば、当該差分情報管理ブロック２０４のアドレスに対応する差分ＬＵ２０６のアドレスから差分データを読み出して、仮想ＬＵのイメージを提供する。一方、所望の世代の差分データでなければ、他の差分データに対するリンクアドレスを参照して、所望の世代の差分データを探す。そして、いずれの差分データも所望の世代のものでなければ、現在、正ＬＵに記録されているデータを仮想ＬＵのデータとして提供する。

30

【 0 0 4 7 】

図４は、本発明の第１の実施の形態のＬＵ追加処理のフローチャートであり、ホスト３からデータの書き込み要求を受信したときに、プールＬＵ管理プログラム１７２によって実行される。

【 0 0 4 8 】

まず、空きキューカウンタ２０５を参照して、プール領域（差分ＬＵ）の空き容量（既使用されている容量でもよい）を確認する（Ｓ１０１）。

【 0 0 4 9 】

次に、ステップＳ１０１で取得したプール領域の使用量と予め定めたしきい値とを比較し、プール領域の使用量がしきい値を超えているか否かを判定する（Ｓ１０２）。このしきい値は、プール領域の使用量が１００％に到達しないように余裕を持たせて７０％程度とすることが望ましいが、ストレージ装置の稼働状態（例えば、データ入出力要求の量）によってしきい値を変更することができる。

40

【 0 0 5 0 】

また、ＬＵのプール領域への追加処理の実行中は通常時の値より高いしきい値を設定することによって（例えば、通常時の７０％に対して７５％）、プール領域の使用量がしきい値を超えたことによってＬＵ追加処理が実行されているにもかかわらず、新たにＬＵ追加処理の実行を防止することもできる。

【 0 0 5 1 】

50

そして、プール領域の使用量がしきい値を超えていなければ、ＬＵを追加する必要がないと判定し、ＬＵを追加することなくこの処理を終了する。一方、プール領域の使用量がしきい値を超えていれば、ステップＳ１０３に進む。

【００５２】

ステップＳ１０３では、プール自動追加対象管理テーブル１４３（図５）を参照して、プール自動追加対象にＬＵが登録されているか否かを判定する。そして、プール自動追加対象にＬＵが登録されていなければ、追加すべきＬＵが存在しないので、ＬＵを追加することなくこの処理を終了する。一方、プール自動追加対象にＬＵが登録されていれば、ステップＳ１０４に進む。

【００５３】

ステップＳ１０４では、プール管理テーブル１４６（図５）を参照して、プール領域に登録されているＬＵが最大数に達しているか否かを判定する。このプール領域の登録数の最大値はプール管理テーブル１４６の容量によって定められる。そして、プール管理テーブル１４６に登録されているＬＵの数が既に最大数であれば、プール管理テーブル１４６に新たにＬＵを追加することができないので、ＬＵを追加することなくこの処理を終了する。一方、プール管理テーブル１４６に登録されているＬＵの数が最大数に満たなければ、ステップＳ１０５に進む。

【００５４】

ステップＳ１０５では、キャッシュメモリ１０７の管理領域に空きがあるか否かを判定する。これは、プール領域にＬＵを追加することに伴い差分情報管理ブロック２０４の容量を増加することが必要なためである。そして、キャッシュメモリ１０７の管理領域に空きがなければ、差分情報管理ブロック２０４の領域を増やすことができないので、ＬＵを追加することなくこの処理を終了する。一方、キャッシュメモリ１０７の管理領域に空きがあれば、ステップＳ１０６に進む。

【００５５】

Ｓ１０２～Ｓ１０５の処理によってＬＵを追加する条件が確認されたら、移動するＬＵをプール自動追加対象管理テーブル１４３から削除し、プール管理テーブル１４６に追加して、差分ＬＵとして使用可能にする（Ｓ１０６）。

【００５６】

この移動するＬＵとして、複数のＬＵをプール管理テーブルに事前に指定しておくことができる。このようにすることで、必要な分だけＬＵを追加することが可能となる。例えば、１００ＧＢのＬＵを１０個登録しておく、プール領域の容量が３００ＧＢ不足した場合に１００ＧＢのＬＵを３個だけプール領域に追加して、あとの７個のＬＵは、別の用途に使用可能となる。

【００５７】

その後、従来から存在した差分情報管理ブロック２０４の最後の空きキューに、新たに増加した差分情報管理ブロック２０４の領域の空きキューのアドレスを設定し、新たに増加した空きキューを従来からの空きキューにつなげる（Ｓ１０７）。

【００５８】

なお、このＬＵ追加処理はプール領域の使用量が予め定めたしきい値を超えた場合に実行されるものであるが、これと異なる更に高い第２のしきい値を設けて、この第２のしきい値を超えたときに、プール領域に記憶された差分データを消去して、緊急にプール領域の空き容量を確保することもできる。

【００５９】

図５は、本発明の第１の実施の形態のＬＵ追加処理の説明図である。

【００６０】

正ＬＵ（ＬＵ０）に対して、スナップショット生成時のデータのイメージを提供する仮想ＬＵ（ＬＵ２）が設けられている。このＬＵ２は、正ＬＵのデータと、プール領域にある差分データとによって構成される仮想的な論理ユニットである。すなわち、ＬＵ２に対するアクセスがあると、そのデータが正ＬＵにあるのか、プール領域にある差分データな

10

20

30

40

50

のかを判断して、差分ＬＵに対するアクセスを実現している。

【００６１】

プール領域に追加されるＬＵは、プール自動追加対象管理テーブル１４３に登録されているもの（ＬＵ６、ＬＵ７）を、プール自動追加対象管理テーブル１４３から削除し、プール管理テーブル１４６に追加する。これによって、追加されたＬＵが差分ＬＵとして使用可能となる。

【００６２】

このように本発明の第１の実施の形態では、プール領域の容量を増加するためのＬＵを事前に複数用意しておき、プール領域の使用量があるしきい値を越えた場合に、自動的にＬＵをプール領域に追加し、プール領域の容量を増加して、プール領域の使用量が１００％になるのを防ぐ。また、プール領域の使用量が少なく、ＬＵを追加する必要がない場合は、ＬＵを他の用途に使用することができる。よって、不要なディスクをプール領域として事前に登録することなく、ディスクの容量を有効に活用することができる。

10

【００６３】

次に、本発明の第２の実施の形態について説明する。

【００６４】

本発明の第２の実施の形態は、前述した第１の実施の形態と異なり、正ＬＵがグループに分けられており、グループ毎にＬＵをプール領域に追加する条件を設定することができる。

【００６５】

20

図６は、本発明の第２の実施の形態のプール管理テーブルの説明図である。

【００６６】

プール管理テーブルは、グループ１の領域（１４１）、グループ２の領域（１４２）に分けて設けられている。

【００６７】

プール管理テーブルは、主ボリュームである正ＬＵが登録される「対象正ＬＵ」欄と、既にプール領域に登録され差分ＬＵとして使用されているＬＵが登録される「プール登録ＬＵ」欄、プール自動追加対象となっているＬＵが登録される「プール自動追加対象ＬＵ」欄、及びプール領域への自動追加の条件が登録される「プール自動追加条件」欄が設けられている。すなわち、プール管理テーブルには、どのＬＵを、どのタイミングで追加するかが規定されている。

30

【００６８】

具体的には、グループ１は、ＬＵ０及びＬＵ３が正ＬＵとして属し、ＬＵ１１及びＬＵ１２が既にプール領域に追加されている。そして、グループプール自動追加対象としてＬＵ３１、ＬＵ３２及びＬＵ３３が登録されている。すなわちグループ毎に追加可能なＬＵを別個に定義することによって、特定のグループ（特定の正ＬＵ）に多量のＬＵがプール領域に追加されることを防止することができる。

【００６９】

また、グループ１の「プール自動追加条件」欄は（２）が選択されているので、各グループで共通のものとして設定された条件によってＬＵが追加される。また、（１）の各グループ個別が選択されると、各グループ毎にＬＵ追加のしきい値を変えて定義することもできる。また、グループ２のように、（３）の各グループのプール領域の最大量によって、当該グループに追加可能なＬＵの最大量を規定することができ、特定のグループ（特定の正ＬＵ）に多量のＬＵがプール領域として追加されることを防止することができる。また、（４）の自動追加対象のＬＵの残り容量を指定することによって、特定のグループに追加するときは残り容量を大きく確保し、その残ったＬＵを他のグループに追加することもできる。

40

【００７０】

以上説明したプール管理テーブルでは、正ＬＵに対応してグループを設けているが、ホスト３に対応してＬＵのグループを設けたり、ホスト３で動作しているアプリケーション

50

プログラムに対応してLUのグループを設けてもよい。

【0071】

図7は、本発明の第2の実施の形態のLU追加処理のフローチャートであり、ホスト3からデータの書き込み要求を受信したときに、プールLU管理プログラム172によって実行される。また、LU追加処理はグループ毎に実行される。

【0072】

まず、空きキュー数カウンタ205を参照して、プール領域(差分LU)のうち既に使用されている容量を確認し(S111)、LUを追加する条件を確認する(S112~S115)。この処理は、第1の実施の形態のLU追加処理(図4のS101~S105)と同じである。

10

【0073】

次に、LUの自動追加条件が共通であるか、個別であるかを判定する(S116)。そして、LUの自動追加条件が共通であれば、グループ個別の条件が設定されていないので、ステップS120に移行する。一方、LUの自動追加条件が個別であれば、そのグループに設定されているLUの追加条件を確認する。

【0074】

まず、プール管理テーブル146等の「プール自動追加対象LU」欄を参照して、そのグループに追加するLUを選択する(S117)。そして、そのグループのプール領域の最大量を確認して、ステップS117で選択されたLUを追加しても、当該最大値を超えないか否かを判定する(S118)。そして、そのグループのプール領域の最大値を超え

20

るときは、当該グループに過大なLUが割り当てられていると判断し、LUを追加することなくこの処理を終了する。一方、そのグループのプール領域の最大値を超えなければ、ステップS119に進む。

【0075】

次に、ステップS119では、ステップS117で選択されたLUを追加しても、プール自動追加対象のLUの残り容量が所定値未満にならないか否かを判定する。そして、プール自動追加対象のLUの残り容量が所定値未満となるときは、他のグループの運用に影響があると判定し、LUを追加することなくこの処理を終了する。一方、プール自動追加対象の残り容量が所定値未満でなければステップS120に進む。

【0076】

その後、LUを追加する条件が確認されたら、移動するLUをプール自動追加対象管理テーブルから削除し、プール管理テーブル146に追加して、差分LUとして使用可能にする(S120)。

30

【0077】

その後、従来から存在した差分情報管理ブロック204の最後の空きキューに、新たに増加した差分情報管理ブロック204の領域の空きキューのアドレスを設定し、新たに増加した空きキューを従来からの空きキューにつなげる(S121)。

【0078】

図8は、本発明の第2の実施の形態のLU追加処理の説明図である。

【0079】

正LU(LU0)に対して、スナップショット生成時のデータのイメージを提供する仮想LU(LU2)が設けられている。同様に、正LU(LU3)に対して仮想LU(LU4)が、正LU(LU5)に対して仮想LU(LU6)が、正LU(LU7)に対して仮想LU(LU8)が設けられている。

40

【0080】

そして、プール領域に追加されるLUは、プール自動追加対象管理テーブル143に登録されている。そして、プール自動追加対象管理テーブル143からLU31~LU33を削除し、LU31及びLU32をグループ1のプール管理テーブル146追加し、LU33をグループ2のプール管理テーブル142に追加する。これによって、追加されたLUが各グループの差分LUとして使用可能となる。

50

【 0 0 8 1 】

このように本発明の第 2 の実施の形態では、複数の L U をグループ化し、そのグループ単位でプール領域の容量を自動的に追加することによって、グループ毎の信頼性を変えた設定をすることができる。

【 0 0 8 2 】

また、各ホストからのデータ入出力の影響を局所的に止めることができる。すなわち、ディスクアレイコントローラ 1 0 単位でプール領域を管理しているが、ホスト 3 が暴走して、そのホストの配下の L U に偏ってデータ書き込み命令が発行され続けると、そのホストの配下の L U に対するプール領域の容量が急激に増大し、他の L U がスナップショットを維持できなくなる。このため、L U を複数のグループに分け、グループ毎にプール自動追加の条件を設定することによって、特定のグループ（特定の正 L U ）に多量の L U がプール領域に追加されることを防止することができる。

10

【 0 0 8 3 】

次に、本発明の第 3 の実施の形態について説明する。

【 0 0 8 4 】

本発明の第 3 の実施の形態は、前述した第 2 の実施の形態でプール領域に追加された L U を削除するものである。なお、前述した第 2 の実施の形態のように、L U がグループ化された場合について説明するが、一つのためのグループを考えることによって前述した第 1 の実施の形態に適用することもできる。

【 0 0 8 5 】

図 9 は、本発明の第 3 の実施の形態のプール管理テーブルの説明図である。

20

【 0 0 8 6 】

プール管理テーブル（L U 削除用）は、グループ 1 の領域（1 4 4）、グループ 2 の領域（1 4 5）に分けて設けられている。なお、ここではグループ 1 の領域（1 4 4）についてのみ説明する。

【 0 0 8 7 】

プール管理テーブル（L U 削除用）は、プール領域からの削除対象となっている L U が登録される「プール削除対象 L U」欄、及びプール領域からの自動削除の条件が登録される「自動削除対象」欄が設けられている。すなわち、プール管理テーブル（L U 削除用）には、どの L U を、どのタイミングで削除するかが規定されている。

30

【 0 0 8 8 】

具体的には、グループ 1 からは L U 3 1、L U 1 2 の順序で L U を削除するように L U の削除順序が規定されている。なお、L U の削除順序を規定する他、L U を登録した逆の順序で（すなわち、最近登録された L U から先に）削除することや、プール領域に登録されている L U の大きさの順序で（例えば、小容量の L U を優先して、又は、大容量の L U を優先して）削除することもできる。

【 0 0 8 9 】

また、グループ 1 では、1 週間継続してプール領域の使用量が 3 0 % に満たない場合に、プール領域から L U を削除するように L U の削除条件が定められている。なお、プール領域の使用量によって L U の削除を判断する他、L U 削除後のプール領域使用量を推定し、推定された L U 削除後のプール領域使用量と設定値とを比較した結果に基づいて、L U を削除するかの判断をしてもよい。なお、L U の使用量を監視する期間は任意に設定することができる。

40

【 0 0 9 0 】

図 1 0 は、本発明の第 3 の実施の形態の L U 削除処理のフローチャートであり、プール L U 管理プログラム 1 7 2 によって実行される。

【 0 0 9 1 】

プール L U 管理プログラム 1 7 2 では、所定のタイミングで、空きキュー数カウンタ 2 0 5 を参照して、プール領域（差分 L U ）のうち既に使用されている容量を確認する（S 1 3 1）。次に、ステップ S 1 0 1 で取得したプール領域の使用量と予め定めたしきい値

50

とを比較し、プール領域の使用量がしきい値を下回っているか否かを判定する（S 1 3 2）。例えば、1 時間毎に空きキュー数カウンタ 2 0 5 の値を確認し、1 週間連続してプール領域の使用量が 3 0 % 未満であるかを判定する。このしきい値等の判定条件はプール管理テーブル 1 4 4 等に登録されている。

【 0 0 9 2 】

そして、プール領域の使用量がしきい値を下回っていなければ、L U を削除する必要がないと判断し、この処理を終了する。一方、プール領域の使用量がしきい値を下回っていれば、削除する L U を選択する（S 1 3 3）。この削除する L U は、プール管理テーブル 1 4 4 等を参照して選択される。例えば、プール管理テーブルに、（ 1 ）削除する L U の順序が登録されている場合には、その順序に従って L U を削除する。また、（ 2 ）L U を登録した逆の順序で（すなわち、最近登録された L U から先に）削除することもできる。また、（ 3 ）プール領域に登録されている L U の大きさの順序で（例えば、小容量の L U を優先して、又は、大容量の L U を優先して）削除することもできる。

10

【 0 0 9 3 】

なお、L U を自動的に削除するのではなく、管理者から指示によって L U を削除する場合には、このステップ S 1 3 3 で削除する L U が指定される。

【 0 0 9 4 】

その後、削除対象 L U の空きキューを検索し、空きビットマップを作成する（S 1 3 4）。この空きビットマップは、削除対象 L U からのデータコピー（S 1 3 8）において参照され、使用中の（空きキューでない）ブロックのみをコピーするために用いられる。また、空きビットマップを作成することによって、削除対象 L U の使用量（削除対象 L U からコピーが必要なデータ量）を知ることができる。

20

【 0 0 9 5 】

そして、差分情報管理ブロック 2 0 4 から削除対象 L U の空きキューを削除する（S 1 3 5）。すなわち、他の L U の空きキューから削除対象 L U へのアドレスのリンクを解除して、削除対象 L U の空きキューと他の L U の空きキューとを切り離して、削除対象 L U の空きキューが新規に割り当てられることを防止する。

【 0 0 9 6 】

そして、削除対象 L U が使用中の領域をロックして、新たにデータが更新されないようにする（S 1 3 6）。

30

【 0 0 9 7 】

そして、L U 削除後もプール領域に残る L U に、削除対象 L U からデータをコピーするために必要な空きキューを確保する（S 1 3 7）。その後、削除対象 L U から、L U 削除後もプール領域に残る L U にデータをコピーする（S 1 3 8）

その後、削除対象 L U からコピーされたデータに関し、正 L U アドレステーブル 2 0 3 に記録される差分情報管理ブロック 2 0 4 のアドレス情報を更新し、差分情報管理ブロック 2 0 4 に記録される差分データ（正 L U 2 0 1 の当該ブロックに対応する他の差分データ）からのアドレスのリンクを更新する（S 1 3 9）。

【 0 0 9 8 】

そして、データコピーが終了したブロックに関する、空きビットマップのビットを更新する（S 1 4 0）。これによって、コピーの進捗状況が把握可能となる。

40

【 0 0 9 9 】

その後、全データのコピーが終了すると、削除対象 L U が使用中の領域のロックを解除する（S 1 4 1）。そして、移動する L U をプール管理テーブルから削除し、プール自動追加対象管理テーブルに追加する（S 1 4 2）。

【 0 1 0 0 】

図 1 1 は、本発明の第 3 の実施の形態の L U 削除処理の説明図である。

【 0 1 0 1 】

正 L U（L U 0）に対して、スナップショット生成時のデータのイメージを提供する仮想 L U（L U 2）が設けられている。同様に、正 L U（L U 3）に対して仮想 L U（L U

50

４）が設けられている。

【０１０２】

そして、プール領域に追加されるＬＵは、プール自動追加対象管理テーブル１４３に登録されている。そして、グループ１のプール管理テーブル１４４からＬＵを削除し、プール自動追加対象管理テーブル１４３に追加する。これによって、削除されたＬＵを他のグループのプール領域に追加することによって、他のグループの差分ＬＵとして使用可能となる。

【０１０３】

このように本発明の第３の実施の形態では、プール領域に登録されているＬＵをプール自動追加対象へ移動するので、ＬＵが他の用途に使用可能となり、リソースを有効に活用

10

【０１０４】

すなわち、スナップショットのプール領域に実際に必要な量より多くのＬＵが登録されていた場合、またホストのＩ／Ｏ（ホストで動作しているアプリケーション）が変化して、プール領域（差分ＬＵ）の必要量が少なくなった場合に、従来技術ではプール領域の容量を減らすことは考慮されていないところ、プール領域に登録されているＬＵをプール自動追加対象へ移動する際に、現在使用されている領域を別のＬＵの空きキューにコピーし、移動対象ＬＵの全領域が未使用となったとき、プール自動追加対象へ移動するので、ＬＵが他の用途に使用可能となる。よって、ディスク容量を有効に活用することができる。

【０１０５】

20

また、設定条件を参照して、ＬＵをプール領域からプール自動追加対象へ移動するので、管理者の手を煩わせることなく、ディスク容量を有効に活用することができる。

【図面の簡単な説明】

【０１０６】

【図１】本発明の実施の形態のディスクアレイ装置のブロック図である。

【図２】本発明の実施の形態の制御プログラムの説明図である。

【図３】本発明の実施の形態の差分ＬＵの管理方法の説明図である。

【図４】本発明の第１の実施の形態のＬＵ追加処理のフローチャートである。

【図５】本発明の第１の実施の形態のＬＵ追加処理の説明図である。

【図６】本発明の第２の実施の形態のプール管理テーブルの説明図である。

30

【図７】本発明の第２の実施の形態のＬＵ追加処理のフローチャートである。

【図８】本発明の第２の実施の形態のＬＵ追加処理の説明図である。

【図９】本発明の第３の実施の形態のプール管理テーブルの説明図である。

【図１０】本発明の第３の実施の形態のＬＵ削除処理のフローチャートである。

【図１１】本発明の第３の実施の形態のＬＵ削除処理の説明図である。

【符号の説明】

【０１０７】

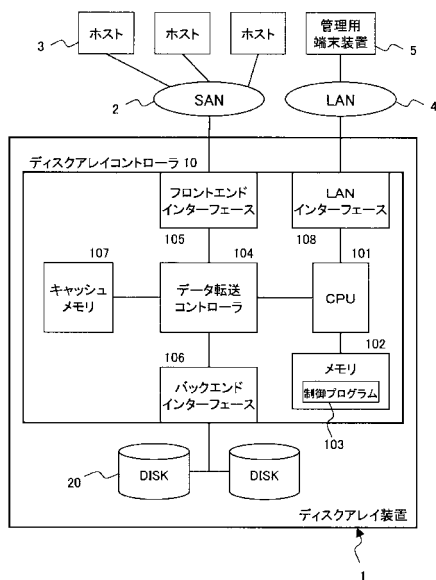
- １ ディスクアレイ装置
- ２ ＳＡＮ
- ３ ホスト
- ４ ＬＡＮ
- ５ 管理用端末装置
- １０ ディスクアレイコントローラ
- ２０ ディスク
- １０１ ＣＰＵ
- １０２ メモリ
- １０３ 制御プログラム
- １０４ データ転送コントローラ
- １０５ フロントエンドインターフェース
- １０６ バックエンドインターフェース

40

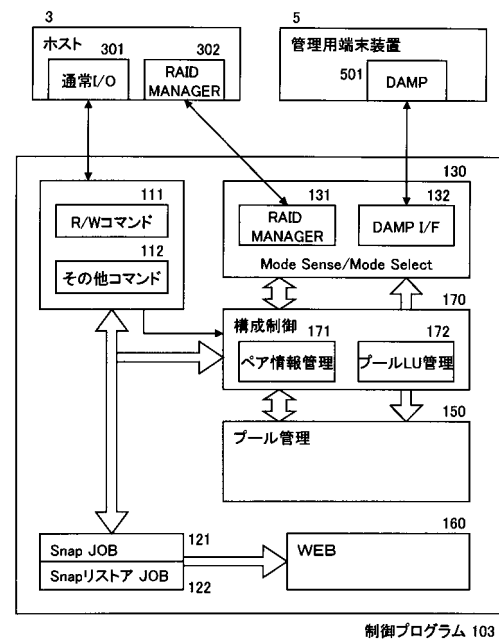
50

- 107 データバッファ
108 LANインターフェース

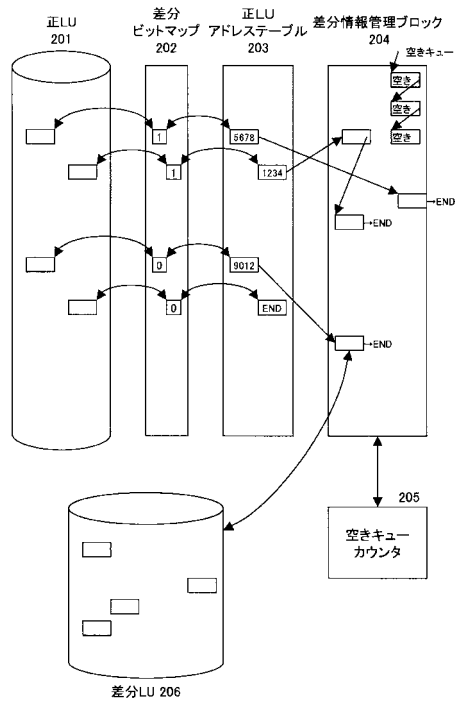
【図1】



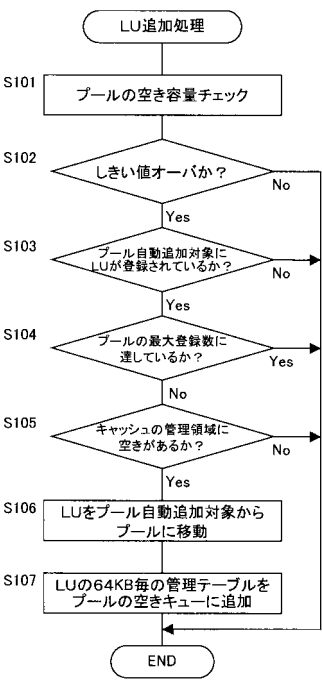
【図2】



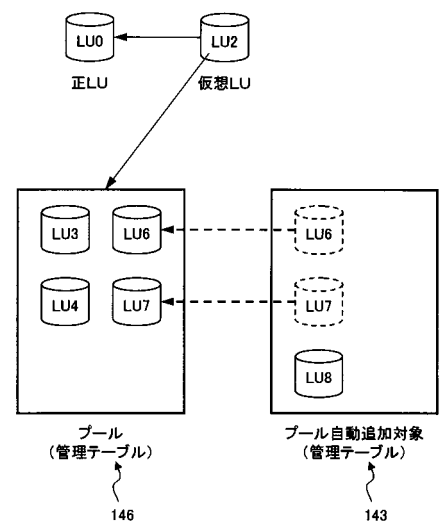
【図 3】



【図 4】



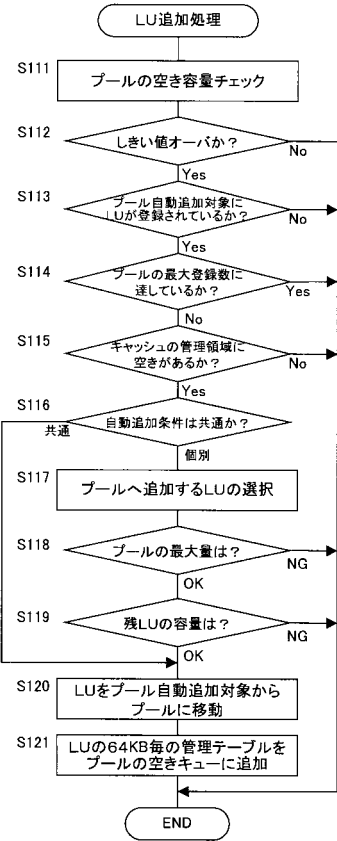
【図 5】



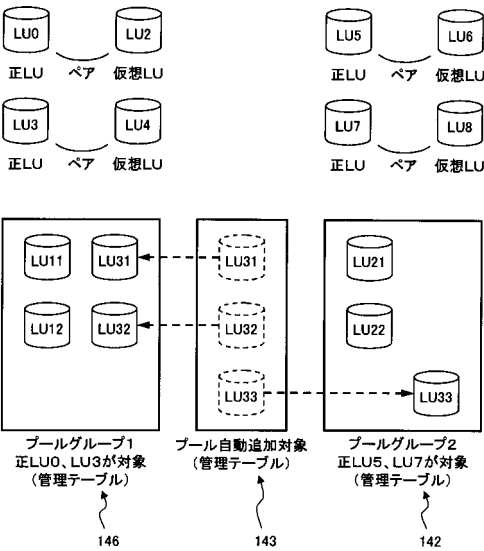
【図 6】

項目	グループ1		グループ2	
	グループ1		グループ2	
対象正LU	LU0、LU3		LU6、LU7	
プール登録LU	LU11、LU12		LU21、LU22	
プール自動追加対象LU	LU31、LU32、LU33		LU31、LU32、LU33	
プール自動追加条件(下記を選択可)	(2)		(3) Max 1TB	
(1)各グループ毎別				
(2)各グループ共通				
(3)各グループのプールの最大量指定				
(4)自動追加対象の残り容量の最小値指定				

【図 7】



【図 8】

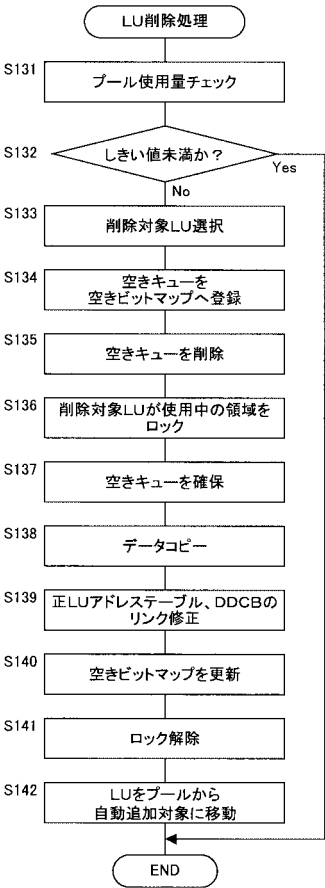


【図 9】

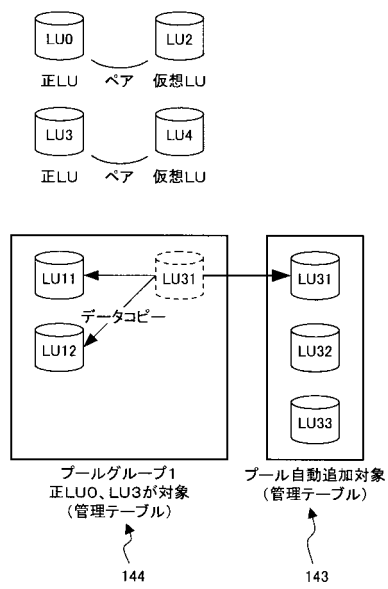
項目	グループ1	グループ2
プール削除対象LU(下記を選択可)	LU31, LU12	...
(1)LU、削除前登録		
(2)登録した範囲に削除		
(3)空き容量に削除		
自動削除条件(下記を選択可)	1週間継続して30%未満	...
(1)プールの現使用量が設定値以下		
(2)LU削除後のプール使用量が設定値以下		
上記は監視する期間も設定可		

プール管理テーブル(LU削除用)

【図 10】



【図 11】



フロントページの続き

(56)参考文献 特開2005-050024(JP,A)
特開2004-127301(JP,A)
特開2002-278819(JP,A)
米国特許出願公開第2004/0030951(US,A1)
特開2002-222061(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06 - 3/08
G06F 12/00
G06F 13/10 - 13/14