

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5070315号
(P5070315)

(45) 発行日 平成24年11月14日 (2012.11.14)

(24) 登録日 平成24年8月24日 (2012.8.24)

(51) Int.Cl.

F I

G 0 6 F 12/00 (2006.01)

G 0 6 F 3/06 (2006.01)

G 0 6 F 12/00 5 0 1 B

G 0 6 F 3/06 3 0 2 J

G 0 6 F 3/06 3 0 4 N

G 0 6 F 3/06 3 0 1

G 0 6 F 12/00 5 0 1 A

請求項の数 8 (全 28 頁)

(21) 出願番号 特願2010-103931 (P2010-103931)
 (22) 出願日 平成22年4月28日 (2010.4.28)
 (65) 公開番号 特開2011-233009 (P2011-233009A)
 (43) 公開日 平成23年11月17日 (2011.11.17)
 審査請求日 平成24年3月1日 (2012.3.1)

(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 100093861
 弁理士 大賀 真司
 (74) 代理人 100129218
 弁理士 百本 宏之
 (72) 発明者 中戸川 雄太
 神奈川県横浜市戸塚区戸塚町5030番地
 株式会社日立製作所 ソフトウェア事業
 部内

審査官 田川 泰宏

最終頁に続く

(54) 【発明の名称】 ストレージ装置及びストレージ装置におけるデータ階層管理方法

(57) 【特許請求の範囲】

【請求項1】

ホスト装置に仮想ボリュームを公開し、前記仮想ボリュームに対する前記ホスト装置からのライト要求に応じて、当該仮想ボリュームに動的に記憶領域を割り当てるストレージ装置において、

複数種別の複数の記憶装置と、

前記仮想ボリュームの前記単位領域ごとのアクセス頻度を監視し、当該アクセス頻度が多い前記単位領域ほど、より応答性能が高い前記記憶装置が提供する前記記憶領域を割り当てるように、対応する前記記憶装置を制御することにより、必要なデータ移行を実行させる制御部と

を備え、

前記制御部は、

前記複数種別の複数の記憶装置のうち、書換え可能回数に制限がある第1の記憶装置が提供する第1の記憶領域から他の第2の記憶装置が提供する第2の記憶領域にデータを移行させる際に、当該データを第1の記憶領域から消去することなく第2の記憶領域に移行させるように、対応する前記記憶装置を制御し、

第1の記憶領域から第2の記憶領域に移行したデータの当該第2の記憶領域に移行された後の更新部分を管理し、

第1の記憶領域から第2の記憶領域に移行させたデータを第2の記憶領域から第1の記憶領域に戻すときには、当該データのうちの前記第2の記憶領域に移行された後に更新さ

れた更新部分を、第 1 の記憶領域に残存する対応するデータ上に上書きさせるように、対応する前記記憶装置を制御する

ことを特徴とするストレージ装置。

【請求項 2】

前記第 1 の記憶装置は、SSD (Solid State Drive) である

ことを特徴とする請求項 1 に記載のストレージ装置。

【請求項 3】

前記制御部は、

前記第 1 の記憶領域の使用容量が予め定められた閾値を超えた場合に、予め与えられた前記仮想ボリュームの前記単位領域ごとの優先度に基づいて、格納されていたデータが前記第 2 の記憶領域に移行された後も当該データが消去されずに残されている単位領域であって、優先度が最も低い単位領域に残されているデータを、前記第 1 の記憶領域の使用容量が前記閾値以下となるまで前記第 1 の記憶領域から消去するように、対応する前記記憶装置を制御する

10

ことを特徴とする請求項 1 に記載のストレージ装置。

【請求項 4】

前記仮想ボリュームの前記単位領域ごとの優先度は、

当該単位領域に対する単位時間当たりの更新量と、当該単位領域に格納されたデータについて予測される前記第 1 及び第 2 の記憶領域間の移行頻度とに基づいて決定された

ことを特徴とする請求項 1 に記載のストレージ装置。

20

【請求項 5】

ホスト装置に仮想ボリュームを公開し、前記仮想ボリュームに対する前記ホスト装置からのライト要求に応じて、当該仮想ボリュームに動的に記憶領域を割り当てるストレージ装置におけるデータ階層管理方法において、

前記ストレージ装置は、

複数種別の複数の記憶装置と、

前記仮想ボリュームの前記単位領域ごとのアクセス頻度を監視し、当該アクセス頻度が多い前記単位領域ほど、より応答性能が高い前記記憶装置が提供する前記記憶領域を割り当てるように、対応する前記記憶装置を制御することにより、必要なデータ移行を実行させる制御部とを有し、

30

前記制御部が、前記複数種別の複数の記憶装置のうち、書換え可能回数に制限がある第 1 の記憶装置が提供する第 1 の記憶領域から他の第 2 の記憶装置が提供する第 2 の記憶領域にデータを移行させる際に、当該データを第 1 の記憶領域から消去することなく第 2 の記憶領域に移行させるように、対応する前記記憶装置を制御する第 1 のステップと、

第 1 の記憶領域から第 2 の記憶領域に移行したデータの当該第 2 の記憶領域に移行された後の更新部分を管理する第 2 のステップと、

第 1 の記憶領域から第 2 の記憶領域に移行させたデータを第 2 の記憶領域から第 1 の記憶領域に戻すときに、当該データのうちの前記第 2 の記憶領域に移行された後に更新された更新部分を、第 1 の記憶領域に残存する対応するデータ上に上書きさせるように、対応する前記記憶装置を制御する第 3 のステップと

40

を備えることを特徴とするデータ階層管理方法。

【請求項 6】

前記第 1 の記憶装置は、SSD (Solid State Drive) である

ことを特徴とする請求項 5 に記載のデータ階層管理方法。

【請求項 7】

前記制御部が、前記第 1 の記憶領域の使用容量が予め定められた閾値を超えた場合に、予め与えられた前記仮想ボリュームの前記単位領域ごとの優先度に基づいて、格納されていたデータが前記第 2 の記憶領域に移行された後も当該データが消去されずに残されている単位領域であって、優先度が最も低い単位領域に残されているデータを、前記第 1 の記憶領域の使用容量が前記閾値以下となるまで前記第 1 の記憶領域から消去するように、対

50

応する前記記憶装置を制御する第４のステップを備える

ことを特徴とする請求項５に記載のデータ階層管理方法。

【請求項８】

前記仮想ボリュームの前記単位領域ごとの優先度は、

当該単位領域に対する単位時間当たりの更新量と、当該単位領域に格納されたデータについて予測される前記第１及び第２の記憶領域間の移行頻度とに基づいて決定された

ことを特徴とする請求項５に記載のデータ階層管理方法。

【発明の詳細な説明】

【技術分野】

【０００１】

10

本発明は、ストレージ装置及びストレージ装置におけるデータ階層管理方法に関し、特に、記憶装置としてＳＳＤ（Solid State Drive）が搭載され、かつデータ管理方式としてストレージ階層管理方式が適用されたストレージ装置に適用して好適なものである。

【背景技術】

【０００２】

従来、企業等においては、ＳＳＤやハードディスク装置（ＨＤＤ：Hard Disk Drive）等の記憶装置が複数搭載されたディスクアレイサブシステムと呼ばれるストレージ装置にデータを保存することが一般的に行われている。

【０００３】

この種のストレージ装置では、１又は複数の同一種別の記憶装置によりＲＡＩＤ（Redundant Array of Inexpensive Disks）グループが構成され、１つのＲＡＩＤグループを構成する記憶装置が提供する記憶領域上にボリュームと呼ばれる論理的な記憶領域が定義される。そしてホストコンピュータ（以下、これをホスト装置と呼ぶ）からのデータは、このボリュームに読み書きされる。

20

【０００４】

なお、ＳＳＤは、ハードディスク装置と比較してより高い信頼性及び応答性能を有するというメリットがあるため、ストレージ装置の記憶装置としてＳＳＤの採用が今後増えていくと考えられる。一方で、ＳＳＤは、半導体記憶素子の劣化に起因する書換え可能回数の制限があり、さらにハードディスク装置と比較してビット単価が高いといったデメリットがある。

30

【０００５】

ところで、近年、企業等の組織で管理されるデータの量は年々増加する一方であり、ストレージ装置の有効利用が重要課題となっている。この課題を解決するために、ストレージ階層管理やシン・プロビジョニングといったデータ管理技術が従来から提案されている。

【０００６】

例えば、特許文献１には、ストレージ装置内に定義されたボリュームを、記憶領域を提供する記憶装置の信頼性及び応答性能といった特性に応じてストレージ階層と呼ばれるグループに分類し、データを、その価値や単位時間当たりのアクセスＩ／Ｏ数などに応じて最適なストレージ階層のボリュームに配置するストレージ階層管理技術が開示されている。このようなストレージ階層管理技術によれば、記憶装置の有効利用を図ることができる。

40

【０００７】

またシン・プロビジョニング技術は、ストレージ装置の容量利用効率を向上するための技術である。シン・プロビジョニングによらない通常のボリュームにおいては、ボリューム作成時にその全容量に相当する記憶領域が確保されるため、その分の記憶領域を予め用意しておく必要がある。

【０００８】

一方、シン・プロビジョニング技術では、複数のボリュームによりプール（以下、これを仮想プールと呼ぶ）を構成し、その上に仮想的なボリューム（以下、これを仮想ボリュームと呼ぶ）を構築する。そして、その仮想ボリュームをホスト装置に公開し、当該仮想

50

ボリュームのうちのデータが書き込まれた単位領域（以下、これをデータブロックと呼ぶ）に対してだけ、物理的な記憶領域を仮想プール内のボリュームから割り当てる。

【0009】

このようなシン・プロビジョニング技術によれば、大容量の仮想ボリュームを作成しても、その仮想ボリュームの全容量分の記憶領域を予め用意しておく必要がなく、実際にホスト装置から使用される分だけの記憶容量を用意しておけば良いため、ストレージ装置の容量利用効率を向上させることができる。

【先行技術文献】

【特許文献】

【0010】

10

【特許文献1】特開2006-099748号公報

【発明の概要】

【発明が解決しようとする課題】

【0011】

ところで、以上のような従来のストレージ階層管理技術と、シン・プロビジョニング技術とを組み合わせた新たなストレージ階層管理技術を採用することで、記憶装置の有効利用がより促進されると考えられる。

【0012】

すなわち、仮想プールを構成するボリュームを応答性能によって複数のストレージ階層に分類し、仮想ボリュームのデータブロックに書き込まれたデータを、データブロックに対する単位時間当たりのアクセスI/O数に応じて他のストレージ階層に属するボリュームに移行（再配置）することで、ユーザはストレージ階層管理技術とシン・プロビジョニング技術の両方の利点を享受することができるものと考えられる。これにより、ストレージ装置導入時のコストを抑えつつ、ストレージ装置の性能向上を図ることが期待できる。

20

【0013】

ところがボリュームに対する単位時間当たりのアクセスI/O数は、ボリューム内の各データブロックに対するアクセスI/O数の総和であるため、ボリューム内のアクセスが集中するデータブロックが変化しても、ボリュームに対する単位時間当たりのアクセスI/O数の変化は小さい場合が考えられる。

【0014】

30

データブロックに対するアクセスI/O数は、ボリューム全体に対するアクセスI/O数より短時間で変化すると考えられるため、ボリュームを管理単位とする従来のストレージ階層管理技術と比較して、上述のようなデータブロックを管理単位とする新たなストレージ階層管理技術ではボリューム間におけるブロックデータの移行が頻発する可能性がある。

【0015】

そして、例えばSSDが提供する記憶領域上に定義されたボリューム（以下、これをSSDボリュームと呼ぶ）と、ハードディスク装置が提供する記憶領域上に定義されたボリューム（以下、これをHDDボリュームと呼ぶ）との間でデータブロックのデータの移行が頻発した場合、データ移行に伴うSSDへのデータ書き込み回数とSSDからのデータ消去回数とが増加し、SSDを短命化させることになる。このような事態は、ストレージ装置のメンテナンスコストを増加させるばかりでなく、メンテナンスに要する労力をも増加させる問題がある。

40

【0016】

本発明は以上の点を考慮してなされたもので、メンテナンスに要するコスト及び労力を抑制しながら、記憶装置の有効利用を促進させ得るストレージ装置及びストレージ装置におけるデータ階層管理方法を提案しようとするものである。

【課題を解決するための手段】

【0017】

かかる課題を解決するため本発明においては、ホスト装置に仮想ボリュームを公開し、

50

前記仮想ボリュームに対する前記ホスト装置からのライト要求に応じて、当該仮想ボリュームに動的に記憶領域を割り当てるストレージ装置において、複数種別の複数の記憶装置と、前記仮想ボリュームの前記単位領域ごとのアクセス頻度を監視し、当該アクセス頻度が多い前記単位領域ほど、より応答性能が高い前記記憶装置が提供する前記記憶領域を割り当てるように、対応する前記記憶装置を制御することにより、必要なデータ移行を実行させる制御部とを設け、前記制御部が、前記複数種別の複数の記憶装置のうち、書換え可能回数に制限がある第1の記憶装置が提供する第1の記憶領域から他の第2の記憶装置が提供する第2の記憶領域にデータを移行させる際に、当該データを第1の記憶領域から消去することなく第2の記憶領域に移行させるように、対応する前記記憶装置を制御し、第1の記憶領域から第2の記憶領域に移行したデータの当該第2の記憶領域に移行された後の更新部分を管理し、第1の記憶領域から第2の記憶領域に移行させたデータを第2の記憶領域から第1の記憶領域に戻すときには、当該データのうちの前記第2の記憶領域に移行された後に更新された更新部分を、第1の記憶領域に残存する対応するデータ上に上書きさせるように、対応する前記記憶装置を制御するようにした。

10

【0018】

また本発明においては、ホスト装置に仮想ボリュームを公開し、前記仮想ボリュームに対する前記ホスト装置からのライト要求に応じて、当該仮想ボリュームに動的に記憶領域を割り当てるストレージ装置におけるデータ階層管理方法において、前記ストレージ装置は、複数種別の複数の記憶装置と、前記仮想ボリュームの前記単位領域ごとのアクセス頻度を監視し、当該アクセス頻度が多い前記単位領域ほど、より応答性能が高い前記記憶装置が提供する前記記憶領域を割り当てるように、対応する前記記憶装置を制御することにより、必要なデータ移行を実行させる制御部とを有し、前記制御部が、前記複数種別の複数の記憶装置のうち、書換え可能回数に制限がある第1の記憶装置が提供する第1の記憶領域から他の第2の記憶装置が提供する第2の記憶領域にデータを移行させる際に、当該データを第1の記憶領域から消去することなく第2の記憶領域に移行させるように、対応する前記記憶装置を制御する第1のステップと、第1の記憶領域から第2の記憶領域に移行したデータの当該第2の記憶領域に移行された後の更新部分を管理する第2のステップと、第1の記憶領域から第2の記憶領域に移行させたデータを第2の記憶領域から第1の記憶領域に戻すときに、当該データのうちの前記第2の記憶領域に移行された後に更新された更新部分を、第1の記憶領域に残存する対応するデータ上に上書きさせるように、対応する前記記憶装置を制御する第3のステップとを設けるようにした。

20

30

【発明の効果】**【0019】**

本発明によれば、第2の記憶領域から第1の記憶領域にデータを戻す際に第1の記憶装置において書き換えられるデータ量を必要最小限にすることができる。従って、第1の記憶装置において本来書き換えが不要なデータ部分についてまで書き換えが行われることを防止し、データの書き換えが頻発することに起因する第1の記憶装置の短命化を有効に防止することができる。かくするにつき、第1の記憶装置が短命化することに起因するメンテナンスに要するコスト及び労力を抑制しながら、記憶装置の有効利用を促進することができる。

40

【図面の簡単な説明】**【0020】**

【図1】第1の実施の形態による計算機システムの全体構成を示すブロック図である。

【図2】第1の実施の形態によるデータ階層管理方式の説明に供する概念図である。

【図3】ストレージ階層管理処理の説明に供する概念図である。

【図4】第1の実施の形態によるデータ階層管理方式の説明に供する概念図である。

【図5】第1の実施の形態によるデータ階層管理方式の説明に供する概念図である。

【図6】第1の実施の形態によるデータ階層管理方式の説明に供する概念図である。

【図7】第1の実施の形態によるデータ階層管理方式の説明に供する概念図である。

【図8】第1の実施の形態によるデータ階層管理方式の説明に供する概念図である。

50

【図 9】データブロックの優先度の説明に供する図表である。

【図 10】データブロック管理テーブルの説明に供する概念図である。

【図 11】ボリューム管理テーブルの説明に供する概念図である。

【図 12】ページ管理テーブルの説明に供する概念図である。

【図 13】データブロック優先度管理テーブルの説明に供する概念図である。

【図 14】データブロック I/O 履歴テーブルの説明に供する概念図である。

【図 15】データブロック移行履歴テーブルの説明に供する概念図である。

【図 16】ライト処理の処理手順を示すフローチャートである。

【図 17】データブロック移行制御処理の処理手順を示すフローチャートである。

【図 18】SSD ボリューム空き領域監視処理の処理手順を示すフローチャートである。

【図 19】第 1 の実施の形態によるデータブロック優先度決定処理の処理手順を示すフローチャートである。

【図 20】第 2 の実施の形態による計算機システムの全体構成を示すブロック図である。

【図 21】第 2 の実施の形態によるデータブロック優先度決定処理の処理手順を示すフローチャートである。

【発明を実施するための最良の形態】

【0021】

以下図面について、本発明の一実施の形態を詳述する。

【0022】

(1) 第 1 の実施の形態

(1-1) 本実施の形態による計算機システムの構成

図 1 において、1 は全体として本実施の形態による計算機システムを示す。この計算機システム 1 は、ホスト装置 2、ストレージ装置 3 及び管理サーバ 4 を備えて構成される。これらホスト装置 2、ストレージ装置 3 及び管理サーバ 4 は、管理用 LAN (Local Area Network) 5 を介して接続され、ホスト装置 2 及びストレージ装置 3 は、SAN (Storage Area Network) などから構成されるネットワーク 6 を介して接続されている。

【0023】

ホスト装置 2 は、CPU (Central Processing Unit) 10 及びメモリ 11 等の情報処理資源を備えたコンピュータ装置であり、例えばパーソナルコンピュータや、ワークステーション、メインフレームなどから構成される。ホスト装置 2 には、ユーザの業務内容に応じたアプリケーションソフトウェアが実装されており、このアプリケーションソフトウェアが使用するデータがネットワーク 6 を介してストレージ装置 3 に読み書きされる。

【0024】

ストレージ装置 3 は、ホスト装置 2 に対して仮想ボリュームを公開し、当該仮想ボリュームに対するホスト装置 2 からのデータの書き込みに応じて動的に仮想ボリュームに記憶領域を割り当てるシン・プロビジョニング機能が搭載されたもので、CPU 20、メモリ 21、ディスクインタフェース 22 及び複数の記憶装置 23 を備えて構成される。

【0025】

CPU 20 は、ストレージ装置 3 全体の動作制御を司るプロセッサである。またメモリ 21 は、主として制御プログラム及び制御データを記憶するために用いられる。後述するデータブロック移行制御プログラム 24、データブロック管理プログラム 25、データブロック管理テーブル 26、ボリューム管理テーブル 27 及びページ管理テーブル 28 もこのメモリ 21 に格納されて保持される。メモリ 21 に格納された制御プログラムを CPU 20 が実行することにより、ストレージ装置 3 全体としての各種処理が行われる。

【0026】

ディスクインタフェース 22 は、記憶装置 23 に対するインタフェースであり、例えばファイバチャネルプロトコルに従って記憶装置 24 との間でデータ及びコマンドなどを送受する。

【0027】

記憶装置 23 は、例えば SSD などの半導体メモリと、SAS (Serial Attached SCSI

10

20

30

40

50

）ディスク等の高性能かつ高価なハードディスク装置及びSATA（Serial AT Attachment）ディスク等の低性能かつ安価なハードディスク装置とから構成される。

【0028】

同じ種別（SSD、SASディスク又はSATAディスク等）の1又は複数の記憶装置23によりRAIDグループが形成され、1つのRAIDグループが提供する記憶領域がページと呼ばれる所定の大きさ単位で仮想ボリュームに動的に割り当てられる。RAIDグループ、データブロック及びページ等の関係については後述する。

【0029】

管理サーバ4は、CPU30及びメモリ31等の情報処理資源を備えたコンピュータ装置であり、例えばパーソナルコンピュータや、ワークステーション、メインフレームなどから構成される。

10

【0030】

CPU30は、管理サーバ4全体の動作制御を司るプロセッサである。またメモリ31は、主として制御プログラム及び制御データを記憶するために用いられる。後述するデータブロック優先度決定プログラム32、データブロック優先度管理テーブル33、データブロックI/O履歴テーブル34及びデータブロック移行履歴テーブル35もこのメモリ31に格納されて保持される。CPU30が、メモリ31に格納された制御プログラムを実行することにより、管理サーバ4全体としての各種処理が行われる。

【0031】

（1-2）本実施の形態によるデータ階層管理方式

20

図2は、ストレージ装置3の論理構成を示す。本ストレージ装置3では、上述のように同一種別の1又は複数の記憶装置23によりRAIDグループRG（RG1～RG3）が形成され、1つのRAIDグループRGを構成する各記憶装置23が提供する記憶領域上に1又は複数のボリュームVOL（VOL1～VOL3）が定義される。また同一種別の1又は複数のボリュームVOL（同一種別の記憶装置23が提供する記憶領域上に定義された1又は複数のボリュームVOL）により1つの種別のストレージ階層41が形成され、複数種別のストレージ階層41により仮想プールVPOOLが形成される。

【0032】

例えば図2の例では、複数のSSDにより第1のRAIDグループRG（RG1～RG3）が構成され、この第1のRAIDグループRGが提供する記憶領域上にSSDボリュームVOL1が定義され、1又は複数のSSDボリュームVOL1により第1のストレージ階層（以下、これをSSD階層と呼ぶ）HC1が形成されている。また複数のSASハードディスク装置により第2のRAIDグループRG2が構成され、この第2のRAIDグループRG2が提供する記憶領域上にHDDボリューム（以下、これをSAS_HDDボリュームと呼ぶ）VOL2が定義され、1又は複数のSAS_HDDボリュームVOL2により第2のストレージ階層（以下、これをSAS_HDD階層と呼ぶ）HC2が形成されている。さらに複数のSATAハードディスク装置により第3のRAIDグループRG3が構成され、この第3のRAIDグループRG3が提供する記憶領域上にHDDボリューム（以下、これをSATA_HDDボリューム）VOL3が定義され、1又は複数のSATA_HDDボリュームVOL3により第3のストレージ階層（以下、これをSATA_HDD階層と呼ぶ）HC3が形成されている。そしてかかるSSD階層HC1、SAS_HDD階層HC2及びSATA_HDD階層HC3により仮想プールVPOOLが構成されている。

30

40

【0033】

この場合、SSD、SASハードディスク装置及びSATAハードディスク装置の中では、SSDが最も信頼性及び応答性能が高く、SATAハードディスク装置が最も信頼性及び応答性能が低い。従って、仮想プールVPOOLを構成するボリュームVOL1～VOL3のうち、SSDボリュームVOL1の信頼性及び応答性能が最も高く、SAS_HDDボリュームVOL2の信頼性及び応答性能がその次に高く、SATA_HDDボリュームVOL3の信頼性及び応答性能が最も低ことが分かる。

50

【0034】

一方、ストレージ装置3内には、1又は複数の仮想ボリュームVVOLが作成され、この仮想ボリュームVVOLがホスト装置2に公開される。各仮想ボリュームVVOLには、それぞれ固有の識別子（以下、これをLUN（Logical Unit number）と呼ぶ）が付与される。また仮想ボリュームVVOLの記憶領域は所定大きさのブロック（以下、これをデータブロックと呼ぶ）DBLに区分され、これら各データブロックDBLにそれぞれ固有の番号（以下、これをLBA（Logical Block Address）と呼ぶ）が付与される。そしてホスト装置2から仮想ボリュームVVOLに対するデータの読み書きは、LUNとLBAとを組み合わせたものをアドレスとして、当該アドレスを指定したデータブロック単位で行われる。

10

【0035】

各仮想ボリュームVVOLは、ストレージ装置3内に定義されたいずれか1つの仮想プールVPOOLと対応付けられる。そして仮想ボリュームVVOLに対するデータのライト要求がホスト装置2からストレージ装置に与えられた場合、そのライト要求においてデータの書き込み先として指定された仮想ボリュームVVOL内のデータブロックに記憶領域が割り当てられていないときには、その仮想ボリュームVVOLと対応付けられた仮想プールVPOOLから必要量の記憶領域がページPGと呼ばれる所定大きさ単位でそのデータブロックに割り当てられる。

【0036】

ここで、ページPGは、仮想プールVPOOLを構成するボリュームVOL（VOL1～VOL3）から対応する仮想ボリュームVVOLに対して割り当てられる記憶領域の基本単位である。このページPGは、ホスト装置2が仮想ボリュームVVOLに対してデータを読み書きする単位であるデータブロックDBLよりも小さく、従って1つのデータブロックDBLに複数のページPGが割り当てられることになる。そしてホスト装置2からのデータは、当該データの書き込み先として指定された仮想ボリュームVVOL内のデータブロックDBLに割り当てられたページPGに書き込まれ、これ以降、そのデータブロックDBLに対するデータの読み書きは結果的にそのページPGに対して行われることになる。

20

【0037】

他方、ストレージ装置3は、上述のような仮想ボリュームVVOLに対する記憶領域の割り当てと並行して、仮想ボリュームVVOLの各データブロックDBLに対する単位時間当たりのアクセスI/O数を監視する。そしてストレージ装置3は、単位時間当たりのアクセスI/O数の平均に応じて、定期的に又は不定期に、仮想ボリュームVVOLのデータブロックDBLにページPGを割り当てる割当て元のボリュームVOL（VOL1～VOL3）を、必要に応じて他のストレージ階層HC（HC1～HC3）のボリュームVOLに変更（再割当て）するストレージ階層管理処理を実行する。

30

【0038】

具体的に、ストレージ装置3は、図3に示すように、単位時間当たりのアクセスI/O数の平均が予め定められた第1の閾値SH1を超えるデータブロックDBLに対してはSSD階層HC1のSSDボリュームVOL1からページPGを割り当て、単位時間当たりのアクセスI/O数の平均が第1の閾値SH1以下で予め定められた第2の閾値SH2（SH2<SH1）よりも多いデータブロックDBLに対してはSAS_HDD階層HC2のSAS_HDDボリュームVOL2からページPGを割り当て、単位時間当たりのアクセスI/O数の平均が第2の閾値SH2以下のデータブロックDBLに対してはSATA_HDD階層HC3のSATA_HDDボリュームVOL3からページPGを割り当てるように、必要に応じてSSDボリュームVOL1、SAS_HDDボリュームVOL2及びSATA_HDDボリュームVOL3間でデータを移行させる。

40

【0039】

加えて本実施の形態の場合、ストレージ装置3は、図4に示すように仮想ボリュームVOLに対してSSDボリュームVOL1からページPGを割り当てている状態において

50

、ページPGの割当て元をSSDボリュームVOL1からSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3に切り替える場合に、図5に示すように、SSDボリュームVOL1に格納されている対応するデータを、SSDボリュームVOL1から消去することなく残したまま、対応するSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3に移行（コピー）することを特徴の1つとしている。

【0040】

そしてこのようなSSDボリュームVOL1からSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3へのデータ移行が行われた場合、その後、ホスト装置2から対応する仮想ボリュームVVOLの対応するデータブロックDBLに対するアクセスがあったときには、図6に示すように、そのアクセスはデータが移行されたSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3に対して行われるが、この際、ストレージ装置3は、かかるデータブロックDBLに格納されたデータが更新された場合には、当該データブロックDBLに割り当てたページPGのうちのその更新が行われたページPGを管理する。

10

【0041】

そしてストレージ装置3は、この後、かかるデータブロックDBLに対する単位時間当たりのアクセスI/O数の平均が第1の閾値SH1（図3）よりも多くなって、当該データブロックDBLのデータをSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3からSSDボリュームVOL1に戻す（移行する）ことになった場合には、図7に示すように、かかるデータブロックDBLのデータのうちの更新が行われたページPGのデータのみを、SSDボリュームVOL1における移行前のデータ（SSDボリュームVOL1からSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3に移行後もSSDボリュームVOL1に残されていたデータのことであり、以下、これを移行前データと呼ぶ）の対応部分上に上書きコピーする。またストレージ装置3は、この後、図8に示すように、そのデータブロックDBLのデータを移行元のSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3から削除する。

20

【0042】

そしてこのようなSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3からSSDボリュームVOL1へのデータ移行が行われた場合、その後、ホスト装置2から対応する仮想ボリュームVVOLの対応するデータブロックDBLに対するアクセスがあったときには、図8に示すように、そのアクセスはSSDボリュームVOL1に対して行われる。

30

【0043】

ただし、以上のような本実施の形態によるデータ階層管理方式によると、SSDボリュームVOL1からSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3に移行されたデータの移行前データがいつまでもSSDボリュームVOL1内に残ることになるため、SSDボリュームVOL1の容量が枯渇するおそれがある。

【0044】

そこで本実施の形態の場合、管理サーバ4がストレージ装置3のSSDボリュームVOL1に残された各移行前データが格納されている仮想ボリュームVVOLのデータブロックDBLに対して優先度をそれぞれ決定すると共に、決定した各データブロックDBLの優先度をストレージ装置3に通知し、ストレージ装置3が、SSDボリュームVOL1の空き容量が予め定められた閾値を超えた場合に、優先度の低いデータブロックDBLに格納されているデータから順番に移行前データをSSDボリュームVOL1から削除するようになされている。

40

【0045】

この場合において、管理サーバ4は、このようなデータブロックDBLの優先度を、そのデータブロックDBLに格納されたデータの単位時間当たりの更新量（ライトアクセス量）と、そのデータについて予測されるその後の単位時間当たりのボリューム間の移行頻度（以下、これを予測移行頻度と呼ぶ）とに基づいて決定する。

50

【0046】

すなわち、データブロックDBLに対するライトアクセス量が多いほどSSDボリュームVOL1に残された移行前データと、SAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3に移行されたデータとの間の差分も大きくなることが予想される。このような場合、SAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3からSSDボリュームVOL1に移行される更新部分のページ数が多くなることが予想されるだけでなく、データを一旦消去してから書き込みを行うSSD固有の上書き動作により、データ移行に多くの時間を要する可能性がある。

【0047】

これに対して、データブロックDBLに対するライトアクセス量が少なければ、SAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3からSSDボリュームVOL1に戻すページ数が小さくなるため、SSDボリュームVOL1におけるデータの書き換えが行われるページ数を最小限に抑えることができ、さらにデータ移行に要する時間も短縮できると考えられる。

【0048】

またデータの予測移行頻度が高いデータブロックDBLについては、SSDボリュームVOL1からSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3にデータを移行させた場合においても、その後、そのSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3からSSDボリュームVOL1にデータを戻すデータ移行が発生する可能性が高い。

【0049】

以上の点を考慮して、本実施の形態の場合、管理サーバ4は、図9のようにデータブロックの優先度を決定する。すなわち、本実施の形態の場合、SSDボリュームVOL1に移行前データが残されているデータブロックDBLのうち、そのデータブロックDBLに格納されたデータに対する単位時間当たりのライトアクセス量が予め定められた閾値（以下、これをライトアクセス量閾値と呼ぶ）以下であり、かつ予測移行頻度が予め定められた閾値（以下、これを予測移行頻度閾値と呼ぶ）よりも大きいデータブロック（図9においてライトアクセス量が「小」、かつ予測移行頻度が「大」となるデータブロック）DBLの優先度を「高」に決定する。

【0050】

また管理サーバ4は、ライトアクセス量がライトアクセス量閾値以下であり、かつ予測移行頻度が予測移行頻度閾値以下のデータブロック（図9においてライトアクセス量が「小」、かつ予測移行頻度が「小」となるデータブロック）DBLと、ライトアクセス量がライトアクセス量閾値よりも大きく、かつ予測移行頻度も予測移行頻度閾値よりも大きいデータブロック（図9においてライトアクセス量が「大」、かつ予測移行頻度が「大」となるデータブロック）DBLとの優先度を「中」に決定する。さらに管理サーバ4は、ライトアクセス量がライトアクセス量閾値よりも大きく、かつ予測移行頻度が予測移行頻度閾値以下のデータブロック（図9においてライトアクセス量が「大」、かつ予測移行頻度が「小」となるデータブロック）DBLの優先度を「低」に決定する。

【0051】

そして管理サーバ4は、以上のようにして決定した対応するデータブロックDBLの優先度をストレージ装置3に通知する。

【0052】

なお、予測移行頻度PFは、データブロックDBLに格納されたデータがSASハードディスク装置又はSATAハードディスク装置からSSDに移行される平均的な時間間隔をAとし、単位時間をTとして、次式

【数1】

$$PF = H(H) \times T / A \quad \dots (1)$$

10

20

30

40

により求めることができる。ここで、H (H) は、ホスト装置 2 のオンライン状態に応じて決定する定数であり、ホスト装置 2 がオンラインの場合には「1」、オフラインの場合には「0」の値をとる関数である。

【0053】

一方、ストレージ装置 3 は、SSD ボリューム VOL 1 の使用容量を常時監視し、使用容量がその SSD ボリューム VOL 1 について予め設定された閾値（以下、これを使用容量閾値と呼ぶ）を超えた場合には、優先度が低いデータブロック DBL から順番（優先度が「低」、「中」又は「高」に決定されたデータブロック DBL の順番）で、SSD ボリューム VOL 1 に残されているそのデータブロック DBL に格納されたデータの移行前データを SSD ボリューム VOL 1 から削除する。

10

【0054】

以上のような本実施の形態によるデータ階層管理機能を実現するための手段として、図 1 に示すように、ストレージ装置 3 のメモリ 21 には、データブロック移行制御プログラム 24 及びデータブロック管理プログラム 25 と、データブロック管理テーブル 26、ボリューム管理テーブル 27 及びページ管理テーブル 28 とが格納され、管理サーバ 4 のメモリ 31 には、データブロック優先度決定プログラム 32 と、データブロック優先度管理テーブル 33、データブロック I/O 履歴テーブル 34 及びデータブロック移行履歴テーブル 35 とが格納されている。

【0055】

このうちデータブロック移行制御プログラム 24 は、ボリューム VOL 間でのデータ移行を制御するプログラムである。またデータブロック管理プログラム 25 は、仮想ボリューム VVOL 内の各データブロック DBL に対する単位時間当たりのアクセス I/O 数に応じてこれらデータブロック DBL に格納されたデータを必要に応じて他のストレージ階層 HC のボリューム VOL に移行させる等の制御を行うプログラムである。ストレージ装置 3 の CPU 20（図 1）は、データブロック移行制御プログラム 24 に基づいてディスクインタフェース 22（図 1）を介して対応する記憶装置 23 を制御することにより、ボリューム間のデータ移行処理を実行する。また CPU 20 は、データブロック管理プログラム 25 に基づいて後述するライト処理、データ移行制御処理及び SSD ボリューム空き領域関し処理を実行する。

20

【0056】

またデータブロック管理テーブル 26 は、各仮想ボリューム VVOL の各データブロック DBL にそれぞれ格納されたデータの格納場所やそのデータの移行に関する情報等を管理するために用いられるテーブルであり、図 10 に示すように、データブロック ID 欄 26A、データブロック格納場所情報欄 26B、移行前データブロック格納場所情報欄 26C、移行実施日時欄 26D 及び優先度欄 26E から構成される。

30

【0057】

そしてデータブロック ID 欄 26A には、各仮想ボリューム VVOL の各データブロック DBL にそれぞれ付与された識別子（データブロック ID）が格納され、データブロック格納場所情報欄 26B には、対応するデータブロック DBL に書き込まれたデータの実際の格納場所を示すアドレス情報（LUN 及び LBA）が格納される。

40

【0058】

また移行前データブロック格納場所情報欄 26C には、対応するデータブロック DBL のデータが SSD ボリューム VOL 1 から SAS_HDD ボリューム VOL 2 又は SATA_HDD ボリューム VOL 3 に移行された場合に、SSD ボリューム VOL 1 におけるそのデータの移行前データの格納場所を示すアドレス情報（LUN 及び LBA）が格納される。

【0059】

さらに移行実施日時欄 26D には、対応するデータブロック DBL のデータが SSD ボリューム VOL 1 から SAS_HDD ボリューム VOL 2 又は SATA_HDD ボリューム VOL 3 に移行された場合に、そのデータ移行が最初に行われた日時が格納され、優先

50

度欄 2 6 E には、管理サーバ 4 から通知されたそのデータブロック D B L の優先度が格納される。

【0060】

従って、図 1 0 の例では、例えば「DataBlock#1-2」というデータブロック D B L については、現在の優先度は「Mid (中)」であり、「2009/12/01 19:00」に「SSDVolume#1」という S S D ボリューム V O L 1 から現在の「HDDVolume#1」という S A S _ H D D ボリューム V O L 2 又は S A T A _ H D D ボリューム V O L 3 にデータ移行が行われており、その移行前データは「SSDVolume#1」という場所に残されていることが示されている。

【0061】

さらにボリューム管理テーブル 2 7 は、ストレージ装置 3 内に定義された各ボリューム (S S D ボリューム V O L 1、S A S _ H D D ボリューム V O L 2 及び S A T A _ H D D ボリューム V O L 3) を管理するために用いられるテーブルであり、図 1 1 に示すように、ボリューム I D 欄 2 7 A、容量欄 2 7 B、使用容量欄 2 7 C 及び使用容量閾値欄 2 7 D から構成される。

10

【0062】

そしてボリューム I D 欄 2 7 A には、ストレージ装置 3 内に定義された各ボリューム V O L の識別子 (L U N) が格納され、容量欄 2 7 B には、対応するボリューム V O L の容量が格納される。また使用容量欄 2 7 C には、そのボリューム V O L の現在の使用容量が格納され、使用容量閾値欄 2 7 D には、そのボリューム V O L について予め設定された使用容量閾値が格納される。

20

【0063】

従って、図 1 1 の例では、「SSDVolume#1」というボリューム V O L は、容量が「700GB」であり、現在の使用容量が「350GB」であって、使用容量閾値が「80%」であることが示されている。

【0064】

さらにページ管理テーブル 2 8 は、ストレージ装置 3 内に定義された各仮想ボリューム V V O L のデータブロック D B L うち、ページ P G が割り当てられている各データブロック D B L について、当該データブロック D B L に割り当てられたページ P G に書き込まれたデータが更新されたか否かを管理するために用いるテーブルであり、図 1 2 に示すように、データブロック I D 欄 2 8 A、ページ I D 欄 2 8 B 及び更新欄 2 8 C から構成される。

30

【0065】

そしてデータブロック I D 欄 2 8 A には、各仮想ボリューム V V O L におけるページ P G が割り当てられたデータブロック D B L の識別子 (データブロック I D) が格納され、ページ I D 欄 2 8 B には、仮想プール V P O L から対応するデータブロック D B L に割り当てられた各ページ P G の識別子 (ページ I D) が格納される。また更新欄 2 8 C には、対応するページ P G の更新が行われたか否かを示すフラグ (更新されたていないときには「0」、更新されたときには「1」) が格納される。

【0066】

従って、図 1 2 の例では、「DataBlock#1-3」というデータブロック D B L には、「Page#1-3-1」、「Page#1-3-2」、……というページ P G が割り当てられており、このうち少なくとも「Page#1-3-2」というページ P G に格納されたデータは更新されていることが示されている。

40

【0067】

一方、データブロック優先度決定プログラム 3 2 は、上述のようにストレージ装置 3 内に定義された各仮想ボリューム V V O L の各データブロック D B L の優先度をそれぞれ決定するためのプログラムである。管理サーバ 4 の C P U 3 0 (図 1) は、このデータブロック優先度決定プログラム 3 2 に従って各データブロック D B L の優先度をそれぞれ決定し、決定結果をストレージ装置 3 に通知する。

【0068】

50

またデータブロック優先度管理テーブル33は、データブロック優先度決定プログラム32により決定された各データブロックDBLの優先度を管理するために用いるテーブルであり、図13に示すように、ホストID欄33A、オンライン状態欄33B、仮想ボリュームID欄33C、データブロックID欄33D及び優先度欄33Eから構成される。
【0069】

そしてホストID欄33Aには、計算機システム1内に存在する各ホスト装置2にそれぞれ付与された識別子（ホストID）が格納され、仮想ボリュームID欄33Cには、対応するホスト装置2がアクセスを許可された各仮想ボリュームVVOLの識別子（LUN）が格納される。

【0070】

10

またデータブロックID欄33Dには、対応する仮想ボリュームVVOLの各データブロックDBLの識別子（データブロックID）が格納され、優先度欄33Eには、データブロック優先度決定プログラム32により決定された対応するデータブロックDBLの優先度が格納される。本実施の形態の場合、データブロックDBLの優先度は、図8について上述したように、優先度が高い順に「High（高）」、「Mid（中）」及び「Low（低）」の3段階に分けられている。

【0071】

さらにオンライン状態欄33Bには、対応する仮想ボリュームVVOLに対する対応するホスト装置2のオンライン状態（オンラインのときには「Online」、オフラインのときには「Offline」）が格納される。この情報は、データブロック優先度決定プログラム32により管理用LAN5（図1）を介して各ホスト装置2からそれぞれ収集される。

20

【0072】

従って、図13の例では、例えば「Host#1」というホスト装置2は、「Volume#1」、「Volume#2」、……という仮想ボリュームVVOLに対するアクセスが許可され、これらの仮想ボリュームVVOLに対するオンライン状態は「Online」となっており、「Volume#1」という仮想ボリュームVVOLの「DataBlock#1-1」というデータブロックDBLについては優先度が「High」に決定され、「DataBlock#1-2」というデータブロックDBLについては優先度が「Mid」に決定され、「DataBlock#1-3」というデータブロックDBLについては優先度が「Low」に決定されていることが示されている。

【0073】

30

またデータブロックI/O履歴テーブル34は、ストレージ装置3内の各仮想ボリュームVVOLの各データブロックDBLに対するホスト装置2からのアクセス履歴を管理するためのテーブルであり、図14に示すように、データブロックID欄34A、ライトアクセス量欄34B及び日時欄34Cから構成される。

【0074】

そしてデータブロックID欄34Aには、ストレージ装置3内に存在する各データブロックDBLの識別子（データブロックID）が格納され、日時欄34C及びライトアクセス量欄34Bには、それぞれ対応するデータブロックDBLに対してライトアクセスがあった日時及びそのときのライトアクセス量がそれぞれ格納される。

【0075】

40

従って、図14の例では、例えば「DataBlock#1-2」というデータブロックDBLについては、「2009/12/01 18:00」、「2009/12/01 19:00」、「2009/12/01 20:00」、……にそれぞれライトアクセスがあり、そのときのアクセス量（書き込まれたデータ量）がそれぞれ「1355KB」、「2300KB」、「1420KB」、……であったことが示されている。

【0076】

データブロック移行履歴テーブル35は、ストレージ装置3内の各仮想ボリュームVVOLのデータブロックDBLごとのデータ移行の履歴を管理するためのテーブルであり、図15に示すように、データブロックID欄35A、移行元ボリューム欄35B、移行先ボリューム欄35C及び移行実施日時欄35Dから構成される。

【0077】

50

そしてデータブロック I D 欄 3 5 A には、ストレージ装置 3 内の各仮想ボリューム V V O L の各データブロック D B L のうち、仮想プール V P O L からページ P G が割り当てられたデータブロック D B L の識別子（データブロック I D）が格納され、移行元ボリューム欄 3 5 B 及び移行先ボリューム欄 3 5 C には、それぞれ対応するデータブロック D B L のデータが移行されたときの移行元のボリューム V O L 及び移行先のボリューム V O L の各識別子（L U N）がそれぞれ格納される。また移行実施日時欄 3 5 D には、対応するデータ移行が行われた日時が格納される。

【0078】

従って、図 1 5 の例の場合、例えば「DataBlock#1-2」というデータブロック D B L については、「2009/12/01 00:00」に「HDDVolume#3」というボリューム V O L から「SSDVolume#1」というボリューム V O L にデータの移行が行われ、その後、「2009/12/01 19:00」に「SSDVolume#1」というボリューム V O L から「HDDVolume#1」というボリューム V O L にデータの移行（再配置）が行われたことが示されている。

【0079】

なお、データブロック I / O 履歴テーブル 3 4 及びデータブロック移行履歴テーブル 3 5 に格納された各情報は、データブロック優先度決定プログラム 3 2 が定期的に管理用 L A N 5（図 1）を介してストレージ装置 3 にアクセスすることにより収集した情報である。

【0080】

（1-3）本実施の形態によるデータ階層管理機能に関する各種処理

次に、本実施の形態によるデータ階層管理機能に関する各種処理の具体的な処理内容について説明する。なお、以下においては、各種処理の処理主体を「プログラム」として説明するが、実際上は、そのプログラムに基づいて、ストレージ装置 3 の C P U 2 0（図 1）又は管理サーバ 4 の C P U 3 0 がその処理を実行することは言うまでもない。

【0081】

（1-3-1）ライト処理

図 1 6 は、ホスト装置 2 から仮想ボリューム V V O L に対するライト要求を受信したときにデータブロック管理プログラム 2 5 により実行されるライト処理の処理手順を示す。

【0082】

データブロック管理プログラム 2 5 は、ホスト装置 2 からのライト要求が与えられるとこのライトアクセス処理を開始し、まず、そのライト要求において指定されたデータの書込み先が、仮想ボリューム V V O L 内の S A S _ H D D ボリューム V O L 2 又は S A T A _ H D D ボリューム V O L 3 からページ P G が割り当てられているデータブロック D B L であるか否かを判断する（S P 1）。

【0083】

この判断において否定結果を得ることは、データの書込み先としてライト要求において指定された仮想ボリューム V V O L 内のデータブロック D B L に未だ対応する仮想プール V P O L からページ P G が割り当てられておらず、又は、データの書込み先としてライト要求において指定された仮想ボリューム V V O L 内のデータブロック D B L に既に S S D ボリューム V O L 1 からページ P G が割り当てられていることを意味する。

【0084】

かくして、このときデータブロック管理プログラム 2 5 は、ライト要求において指定された仮想ボリューム V V O L 内のデータブロック D B L にページ P G が割り当てられていない場合には、そのデータブロック D B L に S S D ボリューム V O L 1 からページ P G を割り当てた上で、そのページ P G に書込み対象のデータを書き込む（S P 2）。またデータブロック管理プログラム 2 5 は、データの書込み先としてライト要求において指定された仮想ボリューム V V O L 内のデータブロック D B L に既に S S D ボリューム V O L 1 からページ P G が割り当てられている場合には、その S S D ボリューム V O L 1 に書込み対象のデータを書き込む（S P 2）。そしてデータブロック管理プログラム 2 5 は、この後、このライト処理を終了する。

【0085】

これに対してステップSP1の判断において肯定結果を得ることは、データの書込み先としてライト要求において指定された仮想ボリュームVVOL内のデータブロックDBLに既にSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3からページPGが割り当てられていることを意味する。

【0086】

かくしてこのときデータブロック管理プログラム25は、データブロック管理テーブル26を参照して、かかるページPGに格納されているデータがSSDボリュームVOL1から移行されたものであって、そのデータの移行前データがSSDボリュームVOL1に残っているか否かを判断する（SP3）。

10

【0087】

そしてデータブロック管理プログラム25は、この判断において否定結果を得ると、ステップSP5に進み、これに対して肯定結果を得ると、ページ管理テーブル28におけるその書込み対象のデータを書き込む各ページPGに対応する更新欄28C（図12）のフラグを、更新が行われたことを意味する「1」に変更する（SP4）。

【0088】

続いて、データブロック管理プログラム25は、ライト要求において指定された仮想ボリュームVVOL内のデータブロックDBLにSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3から割り当てられているページPGに書込み対象のデータを格納し（SP5）、この後、このライト処理を終了する。

20

【0089】

（1-3-2）データブロック移行制御処理

図17は、データブロック管理プログラム25により実行されるデータブロック移行制御処理の処理手順を示す。データブロック管理プログラム25は、この図17に示す処理手順に従って、仮想ボリュームVVOLのデータブロックDBLに書き込まれたデータを必要に応じて他のストレージ階層HCのボリュームVOLに移行させる。

【0090】

すなわちデータブロック管理プログラム25は、ホスト装置2に公開されている各仮想ボリュームVVOL内のデータブロックDBLのうち、対応する仮想プールVPOOLからページPGが割り当てられている各データブロックDBLに対する単位時間当たりのアクセスI/O数を常時監視する（SP10）。

30

【0091】

そしてデータブロック管理プログラム25は、いずれかのデータブロックDBLに対する単位時間当たりのアクセスI/O数の平均が変動し、データの保存先を他のストレージ階層HCのボリュームVOLに切り替える（そのデータブロックDBLのデータを他のストレージ階層HCのボリュームVOLに移行する）必要があることを検知すると、そのデータブロックDBLのデータの移行元及び移行先の各ボリュームVOLを決定する（SP11）。

【0092】

實際上、データブロック管理プログラム25は、まず、単位時間当たりのアクセスI/O数の平均に上述の変動があったデータブロックDBLのデータの現在の格納場所を、データブロック管理テーブル26上で検索する。またデータブロック管理プログラム25は、図3について上述したルールと、ボリューム管理テーブル27に登録された移行先候補の各ボリュームVOLの使用容量及び使用容量閾値とに基づいて、かかるデータブロックDBLのデータの移行先となるボリュームVOLを決定する。この際、データブロック管理プログラム25は、移行先とすべき優先度のボリュームVOLが複数存在する場合には、それらのボリュームVOLの中からアクセスI/O数の平均が最も少ないボリュームVOLを選択する。ただし、同じ優先度の複数のボリュームVOLの中から上述とは異なる方法により移行先のボリュームVOLを選択するようにしても良い。

40

【0093】

50

続いて、データブロック管理プログラム25は、ステップSP11での決定結果に基づいて、かかるデータブロックDBLのデータの移行がSSDボリュームVOL1からSAS__HDDボリュームVOL2又はSATA__HDDボリュームVOL3への移行であるか否かを判断する（SP12）。

【0094】

そしてデータブロック管理プログラム25は、この判断において肯定結果を得ると、データブロック移行制御プログラム24（図1）を起動する。かくしてデータブロック移行制御プログラム24は、ディスクインタフェース22（図1）を介して対応する記憶装置23（図1）を制御することにより、かかるデータブロックDBLのデータをステップSP11において決定した移行元のボリュームVOL（SSDボリュームVOL1）から移行先のボリュームVOL（SAS__HDDボリュームVOL2又はSATA__HDDボリュームVOL3）に移行させる（SP13）。

10

【0095】

またデータブロック管理プログラム25は、データブロック管理テーブル26の対応するデータブロック格納場所情報欄26Bに格納された情報を、かかるデータの実際の移行先のアドレスに変更することにより、これ以降にホスト装置2からそのデータブロックDBLに対するアクセスがあった場合の参照場所を変更し（SP15）、さらにステップSP13及びステップSP14の処理内容に応じてデータブロック管理テーブル26及びページ管理テーブル28を更新する（SP15）。そしてデータブロック管理プログラム25は、この後、ステップSP10に戻る。

20

【0096】

一方、データブロック管理プログラム25は、ステップSP12の判断において否定結果を得ると、かかるデータブロックDBLのデータの移行が、SAS__HDDボリュームVOL2又はSATA__HDDボリュームVOL3からSSDボリュームVOL1への移行であるか否かを判断する（SP16）。

【0097】

データブロック管理プログラム25は、この判断において肯定結果を得ると、データブロック管理テーブル26を参照して、そのデータが、以前にSSDボリュームVOL1からそのSAS__HDDボリュームVOL2又はSATA__HDDボリュームVOL3に移行されたものであり、かつその移行前データがSSDボリュームVOL1に残されているか否かを判断する（SP17）。

30

【0098】

そして、データブロック管理プログラム25は、この判断において否定結果を得るとステップSP22に進む。これに対してデータブロック管理プログラム25は、このステップSP16の判断において肯定結果を得ると、ページ管理テーブル28を参照して、かかるデータが格納されたSAS__HDDボリュームVOL2又はSATA__HDDボリュームVOL3内のデータブロックDBLのうち、SSDボリュームVOL1からそのSAS__HDDボリュームVOL2又はSATA__HDDボリュームVOL3に移行された後にデータが更新されたページ（ページ管理テーブル28の更新欄28C（図12）に格納されたフラグが「1」となっているページ）PGを検出する。そしてデータブロック管理プログラム25は、この後、データブロック移行制御プログラム24を起動する。かくしてデータブロック移行制御プログラム24は、ディスクインタフェース22を介して対応する記憶装置23（図1）を制御することにより、そのページPGのデータのみを当該SSDボリュームVOL1の対応するページPGに上書きコピーさせる（SP18）。

40

【0099】

続いてデータブロック管理プログラム25は、データブロック管理テーブル26の対応するデータブロック格納場所情報欄26Bに格納された情報を、かかるデータの実際の移行先のアドレスに変更することにより、これ以降にホスト装置2からそのデータブロックDBLに対するアクセスがあった場合の参照場所を変更し（SP19）、さらにディスクインタフェース22を介して対応する記憶装置23を制御することにより、そのデータの

50

移行元のSAS__HDDボリュームVOL 2又はSATA__HDDボリュームVOL 3内からそのデータを削除する（SP 20）。

【0100】

さらにデータブロック管理プログラム25は、ステップSP 19及びステップSP 20の処理内容に応じてデータブロック管理テーブル26及びページ管理テーブル28を更新し（SP 21）、この後、ステップSP 10に戻る。

【0101】

他方、データブロック管理プログラム25は、ステップSP 16の判断において否定結果を得ると、データブロック移行制御プログラム24を起動する。かくしてデータブロック移行制御プログラム24は、ディスクインタフェース22を介して対応する記憶装置23（図1）を制御することにより、かかるデータブロックDBLのデータをステップSP 11において決定した移行元のボリュームVOL（SAS__HDDボリュームVOL 2又はSATA__HDDボリュームVOL 3）から移行先のボリュームVOL（SAS__HDDボリュームVOL 2又はSATA__HDDボリュームVOL 3）に移行させる（SP 22）。

10

【0102】

続いて、データブロック管理プログラム25は、データブロック管理テーブル26の対応するデータブロック格納場所情報欄26Bに格納された情報を、かかるデータの実際の移行先のアドレスに変更することにより、これ以降にホスト装置2からそのデータブロックDBLに対するアクセスがあった場合の参照場所を変更する（SP 23）。またデータブロック管理プログラム25は、ディスクインタフェース22を介して対応する記憶装置23を制御することにより、そのデータの移行元のSAS__HDDボリュームVOL 2又はSATA__HDDボリュームVOL 3からそのデータを削除し（SP 24）、この後、ステップSP 10に戻る。

20

【0103】

（1-3-3）SSDボリューム空き容量監視処理

図18は、データブロック管理プログラム25により実行されるSSDボリューム空き容量監視処理の処理手順を示す。データブロック管理プログラム25は、この図18に示す処理手順に従って、SSDボリュームVOL 1の空き容量を監視し、必要に応じてSSDボリュームVOL 1に残されている移行前データのうちの優先度の低いデータを当該SSDボリュームVOL 1から削除する。

30

【0104】

すなわちデータブロック管理プログラム25は、ストレージ装置3内のSSDボリュームVOL 3の空き容量を常時監視する（SP 30）。そしてデータブロック管理プログラム25は、やがていずれかのSSDボリュームVOL 1の空き容量が、ボリューム管理テーブル27において規定されたそのSSDボリュームVOL 1の使用容量閾値を超えたことを検知すると、データブロック管理テーブル26を参照して、そのSSDボリュームVOL 1からSAS__HDDボリュームVOL 2又はSATA__HDDボリュームVOL 3に移行済みのデータであって、その移行前データがSSDボリュームVOL 1に残っているデータがあるか否かを判断する（SP 31）。

40

【0105】

そしてデータブロック管理プログラム25は、この判断において否定結果を得ると、図17のステップSP 11以降の処理を実行する。この際、データブロック管理プログラム25は、ステップSP 30において空き容量不足を検知したSSDボリュームVOL 1の空き容量不足が解消するまで、図17のステップSP 11以降の処理を繰り返し実行する。またデータブロック管理プログラム25は、このストレージ階層管理処理に伴い、データブロック管理テーブル26、ボリューム管理テーブル27及びページ管理テーブル28を更新する（SP 32）。そしてデータブロック管理プログラム25は、この後、ステップSP 30に戻る。

【0106】

50

これに対してデータブロック管理プログラム25は、ステップSP31の判断において肯定結果を得ると、ディスクインタフェース22を介して対応する記憶装置23を制御することにより、ステップSP30において空き容量不足を検知したSSDボリュームVOL1からSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3に移行済みのデータであって、未だそのSSDボリュームVOL1に残存しているデータのうち、管理サーバ4から通知された優先度が最小のデータブロックDBLのデータをSSDボリュームVOL1から削除させる。またデータブロック管理プログラム25は、これに伴い、データブロック管理テーブル26及びボリューム管理テーブル27を更新する（SP33）。

【0107】

続いてデータブロック管理プログラム25は、ステップSP30において空き容量不足を検知したSSDボリュームVOL1の空き容量不足が解消されたか否かを判断する（SP34）。そしてデータブロック管理プログラム25は、この判断において否定結果を得るとステップSP31に戻り、この後ステップSP31～ステップSP34の処理を繰り返す。

【0108】

そしてデータブロック管理プログラム25は、やがてステップSP30において空き容量不足を検知したSSDボリュームVOL1の空き容量不足が解消されることによりステップSP34において肯定結果を得ると、ステップSP30に戻り、この後、同様の処理を繰り返す。

【0109】

（1-3-4）データブロック優先度決定処理

一方、図19は、管理サーバのデータブロック優先度決定プログラム32により実行されるデータブロック優先度決定処理の処理手順を示す。データブロック優先度決定プログラム32は、この図19に示す処理手順に従って、仮想ボリュームVVOL内の各データブロックDBLの優先度を決定してストレージ装置3に通知する。

【0110】

すなわちデータブロック優先度決定プログラム32は、定期的にこのデータブロック優先度決定処理を開始し、まず、データブロック優先度管理テーブル33（図13）に登録されたデータブロックDBLの中から未処理のデータブロックDBLを1つの選択する（SP40）。

【0111】

続いてデータブロック優先度決定プログラム32は、ステップSP40において選択したデータブロックDBLについて、データブロック移行履歴テーブル35（図15）を参照して、上述した（1）式により予測移行頻度を計算する（SP41）。

【0112】

次いでデータブロック優先度決定プログラム32は、ステップSP41の計算結果と、データブロックI/O履歴テーブル34（図14）に基づき認識される対応するデータブロックDBLに対する単位時間当たりのライトアクセス量とに基づき、図9について上述した優先度決定ルールに従って、そのデータブロックDBLの優先度を決定する（SP42）。

【0113】

そしてデータブロック優先度決定プログラム32は、ステップSP42の決定結果に基づいて、データブロック優先度管理テーブル33における対応するエントリの優先度欄33E（図13）を更新し（SP43）、この後、データブロック優先度管理テーブル33に登録されたすべてのデータブロックDBLについて同様の処理を実行し終わったか否かを判断する（SP44）。

【0114】

データブロック優先度決定プログラム32は、この判断において否定結果を得ると、ステップSP40に戻り、この後、ステップSP40～ステップSP44の処理を繰り返す

10

20

30

40

50

。そしてデータブロック優先度決定プログラム 32 は、やがてデータブロック優先度管理テーブル 33 に登録されたすべてのデータブロック DBL について同様の処理を実行し終えることによりステップ SP44 において肯定結果を得ると、データブロック優先度管理テーブル 33 に登録されている、各データブロック DBL の最新の優先度を、データブロック管理テーブル 26 (図 10) の更新命令と共にストレージ装置 3 に送信する (SP45)。

【0115】

かくしてストレージ装置 3 のデータブロック管理プログラム 25 (図 1) は、この更新命令を受信すると、データブロック管理テーブル 26 に登録された各データブロック DBL の優先度を、当該更新命令と共に管理サーバ 4 から送信される各データブロック DBL の最新の優先度に基づき必要に応じて更新する。

10

【0116】

そしてデータブロック優先度決定プログラム 32 は、この後、このデータブロック優先度決定処理を終了する。

【0117】

(1-4) 本実施の形態の効果

以上のように本実施の形態による計算機システム 1 によれば、ストレージ装置 3 において、SAS_HDD ボリューム VOL2 又は SATA_HDD ボリューム VOL3 から SSD ボリューム VOL1 にデータを戻す際に SSD ボリューム VOL1 において書き換えられるデータ量を必要最小限にすることができる。

20

【0118】

従って、SSD ボリューム VOL1 において本来書き換えが不要なデータ部分についてまで書き換えが行われることを防止し、データの書き換えが頻発することに起因する SSD の劣化を有効に防止することができる。かくするにつき、SSD が短命化することに起因するメンテナンスに要するコスト及び労力を抑制しながら、記憶装置 23 の有効利用を促進することができる。

【0119】

(2) 第 2 の実施の形態

(2-1) 第 2 の実施の形態におけるデータブロックの優先度の決定方法

第 1 の実施の形態による計算機システム 1 においては、SAS_HDD ボリューム VOL2 又は SATA_HDD ボリューム VOL3 から SSD ボリューム VOL1 にデータを戻す際に、上書きせずに再利用できる移行前ページの合計が多いデータブロック DBL ほど、SSD が短命化することに起因するメンテナンスに要するコスト及び労力を抑制しながら、記憶装置 23 の有効利用を促進するという効果をより多く得ることができる。

30

【0120】

従って、ある一定期間で発生する SAS_HDD ボリューム VOL2 又は SATA_HDD ボリューム VOL3 から SSD ボリューム VOL1 へのデータ移行において、上書きせずに再利用する移行前ページの合計がより多いデータブロック DBL ほど、より高く優先度を設定して、そのデータを SSD ボリューム VOL1 内に残すべきであると考えられる。

40

【0121】

すなわち、データブロック DBL の優先度 P は、1 回の移行で再利用可能なページ数の平均値を ANP、予測移行頻度を PF として、次式のように表すことができる。

【数 2】

$$P \propto ANP \times PF \quad \dots\dots (2)$$

【0122】

ここで、1 回の移行で再利用可能なページ数の平均値 ANP は、データブロック DBL のページ数を NPD、SAS_HDD ボリューム VOL2 又は SATA_HDD ボリューム VOL3 上で生じるライトアクセスによる更新済みページ数の平均値を AN、データブ

50

ロックDBLのデータサイズをDSD、_HDDボリュームVOL2又はSATA_HDDボリュームVOL3上で生じるライトアクセスによる更新済みデータサイズの平均をADS、ページサイズをDSPとして、次式

【数3】

$$\begin{aligned} ANP &= NPD - AN \\ &= (DSD - ADS) / DSP \end{aligned} \quad \dots (3)$$

で与えられる。

10

【0123】

またHDDボリュームVOL2又はSATA_HDDボリュームVOL3上で生じるライトアクセスによる更新済みデータサイズADSは、単位時間当たりのライトアクセス量Wと、データのHDDボリュームVOL2又はSATA_HDDボリュームVOL3上での滞在時間Bの積に比例すると考えられるため、1回の移行で再利用可能なページ数の平均値ANPは、C及びC'を定数として、次式

【数4】

$$\begin{aligned} ANP &= (DSD - C \times W \times B) / DSP \\ &\propto DSD \times C' - W \times B \end{aligned} \quad \dots (4)$$

20

で与えられる。

【0124】

実用上は、負の数になる場合を考慮して、1回の移行で再利用可能なページ数の平均値ANPを次式のように表すことができる。なお、次式において記号[]は、括弧内が負の数になる場合は0とすることを意味する。

【数5】

$$ANP = [DSD \times C' - W \times B] \quad \dots (5)$$

30

【0125】

この(5)式において「 $DSD \times C'$ 」は定数であり、閾値と考えることができる。従って、「 $C' = 1$ 」とすれば、(5)式は、「HDDボリュームVOL2又はSATA_HDDボリュームVOL3上でデータブロックに対してデータブロックのブロックサイズ以上のライトアクセスが発生する見込みがある場合には、優先度を0とする」ことを意味する。

【0126】

一方、一定期間内で発生するHDDボリュームVOL2又はSATA_HDDボリュームVOL3からSSDボリュームVOL1へのデータの予測移行頻度PFは、(1)式のように表すことができる。

40

【0127】

よって、データブロックDBLの優先度Pは、(1)式、(2)式及び(5)式から、次式

【数6】

$$P \propto [D \times C' - W \times B] H(H) \times T / A \quad \dots (6)$$

のように表すことができる。ただし(6)式においてTは定数であるため、データブロックDBLの優先度Pは次式

【数 7】

$$P \propto [D \times C' - W \times B] H(H) / A \quad \dots (7)$$

により表すことができる。

【0128】

よって、各データブロック DBL の優先度 P の高低は、次式

【数 8】

$$P = [D \times C' - W \times B] H(H) / A \quad \dots (8)$$

10

により個々のデータブロック DBL の優先度を算出し、算出結果を比較することによって求めることができる。

【0129】

なお、(8) 式において、「 $D \times C'$ 」は上述のように閾値として値を設定する。また (8) 式において、単位時間当たりのライトアクセス量 W はデータブロック I/O 履歴テーブル 34 (図 14) から求めることができ、データの HDD ボリューム VOL2 又は SATA_HDD ボリューム VOL3 上での滞在時間 B は、データブロック移行履歴テーブル 35 (図 15) から求めることができる。さらにデータブロック DBL に格納されたデータが SAS ハードディスク装置又は SATA ハードディスク装置から SSD に移行される平均的な時間間隔 A は、データブロック移行履歴テーブル 35 から求めることができ、関数 $H(H)$ は、データブロック優先度管理テーブル 33 (図 13) のオンライン状態欄 33E (図 13) を参照することにより値を決定することができる。

20

【0130】

(2-2) 第 2 の実施の形態によるデータブロック優先度決定処理

図 1 との対応部分に同一符号を付して示す図 20 は、第 2 の実施の形態による計算機システム 40 を示す。この計算機システム 40 は、管理サーバ 41 のデータブロック優先度決定プログラム 42 により実行される各データブロック DBL の優先度の決定方法が異なる点を除いて第 1 の実施の形態による計算機システム 1 (図 1) と同様に構成されている。

30

【0131】

図 21 は、図 19 について上述したデータブロック優先度決定処理に代えて、管理サーバ 41 のデータブロック優先度決定プログラム 42 により実行される第 2 の実施の形態によるデータブロック優先度決定処理の処理手順を示す。

【0132】

データブロック優先度決定プログラム 42 は、定期的にこのデータブロック優先度決定処理を開始し、まず、データブロック優先度管理テーブル 33 (図 13) に登録されたデータブロック DBLの中から未処理のデータブロック DBLを 1つの選択する (SP50)。

【0133】

40

続いてデータブロック優先度決定プログラム 42 は、ステップ SP50 において選択したデータブロック DBL について、データブロック I/O 履歴テーブル 34、データブロック移行履歴テーブル 35 及びデータブロック優先度管理テーブル 33 を参照して、単位時間当たりのライトアクセス量 W と、データの HDD ボリューム VOL2 又は SATA_HDD ボリューム VOL3 上での滞在時間 B と、データブロック DBL に格納されたデータが SAS ハードディスク装置又は SATA ハードディスク装置から SSD に移行される平均的な時間間隔 A と、関数 $H(H)$ の値とを求める (SP51)。

【0134】

続いてデータブロック優先度決定プログラム 42 は、ステップ SP51 において求めた各値を利用して、そのデータブロック DBL の優先度を数値として算出し (SP52)、

50

算出結果に基づいてデータブロック優先度管理テーブル 33 における対応する優先度欄 33E (図 13) を更新する (SP53)。

【0135】

次いでデータブロック優先度決定プログラム 42 は、データブロック優先度管理テーブル 33 に登録されたすべてのデータブロック DBL について同様の処理を実行し終えたか否かを判断する (SP54)。

【0136】

データブロック優先度決定プログラム 42 は、この判断において否定結果を得ると、ステップ SP50 に戻り、この後、ステップ SP50 ～ステップ SP54 の処理を繰り返す。そしてデータブロック優先度決定プログラム 42 は、やがてデータブロック優先度管理
10
テーブル 33 に登録されたすべてのデータブロック DBL について同様の処理を実行し終えることによりステップ SP54 において肯定結果を得ると、データブロック優先度管理
テーブル 33 に登録されている、各データブロック DBL の最新の優先度を、データブ
ロック管理テーブル 26 (図 10) の更新命令と共にストレージ装置 3 に送信する (SP55)。

【0137】

かくしてストレージ装置 3 のデータブロック管理プログラム 25 (図 1) は、この更新
命令を受信すると、データブロック管理テーブル 26 に登録された各データブロック DB
L の優先度を、当該更新命令と共に管理サーバ 4 から送信される各データブロック DBL
の最新の優先度に基づき必要に応じて更新する。
20

【0138】

そしてデータブロック優先度決定プログラム 42 は、この後、このデータブロック優先
度決定処理を終了する。

【0139】

(2-2) 本実施の形態の効果

以上のように本実施の形態による計算機システム 40 では、データブロック優先度管理
テーブル 33 に登録された各データブロック DBL の優先度を数値として算出するため、
例えば図 18 について上述した SSD ボリューム空き容量監視処理のステップ SP33 に
おいて優先度が最小のデータブロック DBL のデータを SSD ボリューム VOL1 から削
除する際に、不要なデータブロック DBL のデータをより精度良く選択して SSD ボリュ
ーム VOL1 から削除することができる。
30

【0140】

よって本実施の形態による計算機システム 40 によれば、ストレージ装置 3 において、
SAS_HDD ボリューム VOL2 又は SATA_HDD ボリューム VOL3 から SSD
ボリューム VOL1 にデータを戻す際に SSD ボリューム VOL1 において書き換えられ
るデータ量を必要最小限にすることができるため、SSD が短命化することに起因するメ
ンテナンスに要するコスト及び労力を抑制しながら、記憶装置 23 の有効利用をより一層
と促進することができる。

【0141】

(3) 他の実施の形態

なお上述の第 1 及び第 2 の実施の形態においては、本発明を図 1 のように構成されたス
トレージ装置 3 に適用するようにした場合について述べたが、本発明はこれに限らず、こ
の他種々の構成のストレージ装置に広く適用することができる。
40

【0142】

また上述の第 1 及び第 2 の実施の形態においては、ストレージ装置 3 内に定義された各
仮想ボリューム VVOL の各データブロック DBL の優先度を管理サーバ 4, 41 におい
て決定するようにした場合について述べたが、本発明はこれに限らず、同様の機能をス
トレージ装置 3 に搭載し、かかる各データブロック DBL の優先度をストレージ装置 3 にお
いて決定するようにしても良い。

【0143】

10

20

30

40

50

さらに上述の第1及び第2の実施の形態においては、SSDボリュームVOL1からSAS_HDDボリュームVOL2又はSATA_HDDボリュームVOL3にデータを移行し、その後、そのデータをSSDボリュームVOL1に戻す際に本発明を適用するようにした場合について述べたが、本発明はこれに限らず、SSD以外の書換え可能回数に制限がある記憶デバイスが提供する記憶領域から他の記憶デバイスが提供する記憶領域にデータを移行し、その後、そのデータを書換え回数に制限がある記憶デバイスが提供する記憶領域に戻す場合に広く適用することができる。

【0144】

さらに上述の第1及び第2の実施の形態においては、仮想ボリュームVVOLのデータブロックDBLごとのアクセス頻度を監視し、当該アクセス頻度が多いデータブロックDBLほど、より応答性能が高い記憶装置23が提供する記憶領域を割り当てるように、対応する記憶装置23を制御することにより、必要なデータ移行を実行させる制御部を、ストレージ装置3全体の動作制御を司るCPU20及びデータブロック管理プログラム25により構成するようにした場合について述べたが、本発明はこれに限らず、このような機能を有する構成要素をCPU20とは別個にストレージ装置3に設けるようにしても良い。

10

【産業上の利用可能性】

【0145】

本発明は、データ管理方式としてストレージ階層管理方式が適用されたストレージ装置に適用できる。

20

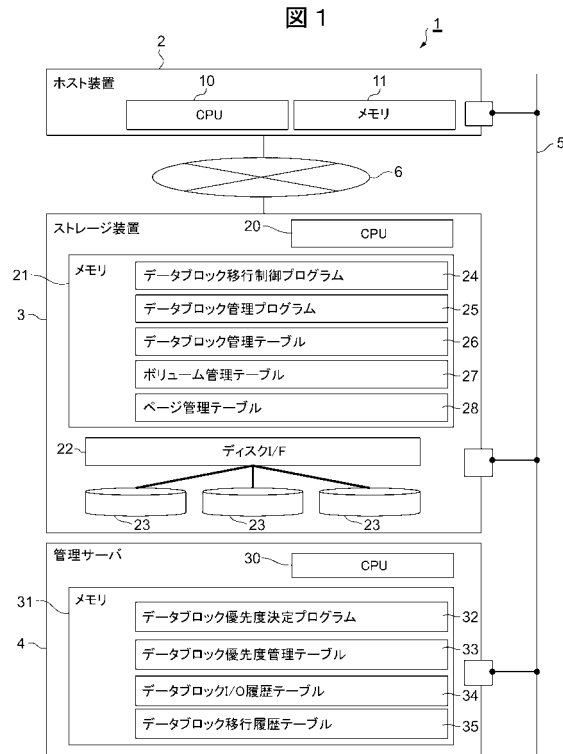
【符号の説明】

【0146】

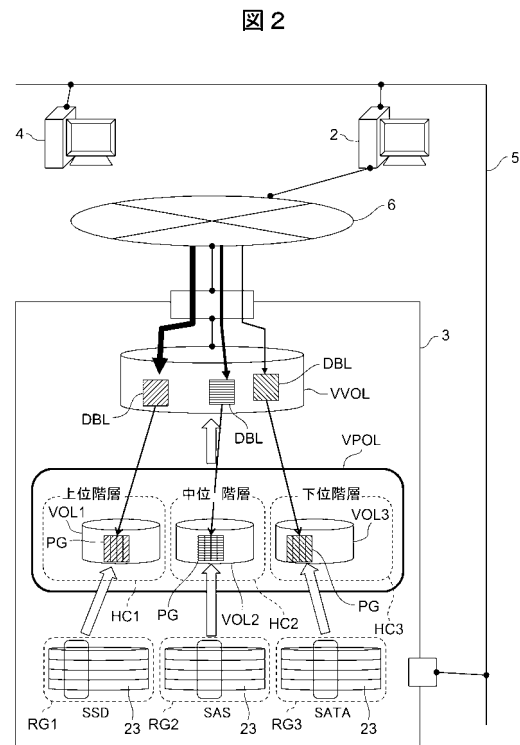
1, 40……計算機システム、2……ホスト装置、3……ストレージ装置、4, 41……管理サーバ、10, 20, 30……CPU、11, 21, 31……メモリ、23……記憶装置、24……データブロック移行制御プログラム、25……データブロック管理プログラム、26……データブロック管理テーブル、27……ボリューム管理テーブル、28……ページ管理テーブル、32, 42……データブロック優先度決定プログラム、33……データブロック優先度管理テーブル、34……データブロックI/O履歴テーブル、35……データブロック移行履歴テーブル、DBL……データブロック、HC, HC~HC3……ストレージ階層、PG……ページ、VOL, VOL1~VOL3……ボリューム、VPOL……仮想プール、VVOL……仮想ボリューム。

30

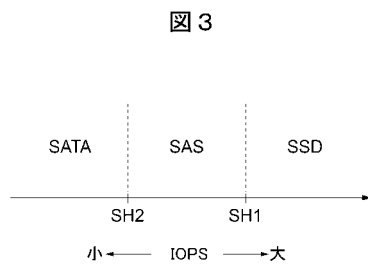
【図 1】



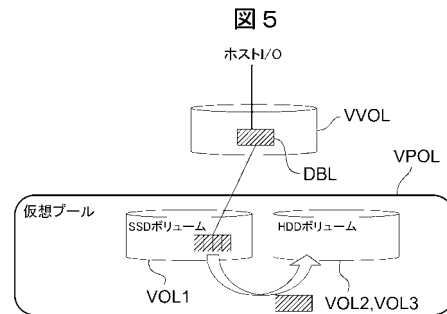
【図 2】



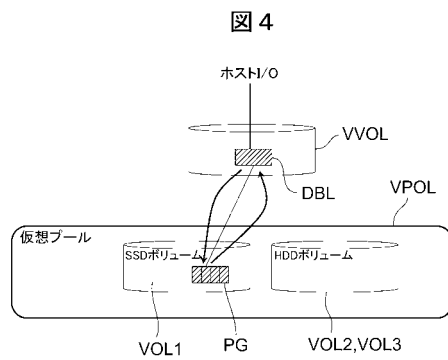
【図 3】



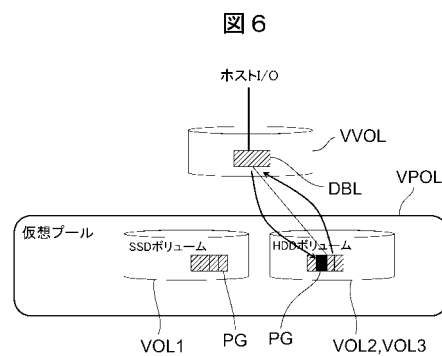
【図 5】



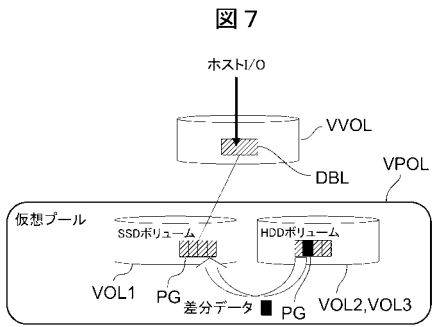
【図 4】



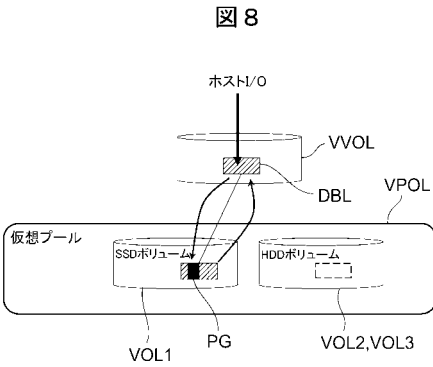
【図 6】



【図 7】



【図 8】



【図 9】

図 9

データブロックの 優先度		ライトアクセス量	
		小	大
予測 移行頻度	大	高	中
	小	中	低

【図 10】

図 10

データブロックID	データブロック 格納場所情報	移行元データブロック 格納場所情報	移行実施日時	優先度
DataBlock#1-1	SSDVolume#3	SSDVolume#1	2009/12/01 19:00	High
DataBlock#1-2	HDDVolume#1	SSDVolume#2	2009/12/01 18:00	Mid
DataBlock#1-3	HDDVolume#3	SSDVolume#2	2009/12/01 18:00	Low
DataBlock#2-1	HDDVolume#1	SSDVolume#1	2009/12/02 14:00	High
DataBlock#2-2	SSDVolume#2	SSDVolume#2	2009/12/02 14:00	Low
DataBlock#2-3	HDDVolume#1	SSDVolume#1	2009/12/01 09:00	Low
DataBlock#3-1	HDDVolume#1	SSDVolume#1	2009/12/01 09:00	Low
DataBlock#3-2	HDDVolume#2	SSDVolume#2	2009/12/01 09:00	Mid
DataBlock#3-3	SSDVolume#1	SSDVolume#1	2009/12/01 09:00	Mid

【図 11】

図 11

ボリュームID	容量	使用容量	使用容量閾値
SSDVolume#1	700 GB	350 GB	80 %
SSDVolume#2	500 GB	300 GB	80 %
...	...	...	...

【図 12】

図 12

データブロックID	ページID	更新
DataBlock#1-2	Page#1-2-1	0
	Page#1-2-2	0
	...	...
DataBlock#1-3	Page#1-3-1	0
	Page#1-3-2	1
	...	...
DataBlock#2-3	Page#2-3-1	1
	Page#2-3-2	0
	...	...
DataBlock#3-1	Page#3-1-1	1
	Page#3-1-2	1
	...	...

【図 13】

ホストID	オンライン状態	仮想ボリュームID	データブロックID	優先度
Host#1	Online	Volume#1	DataBlock#1-1	High
			DataBlock#1-2	Mid
			DataBlock#1-3	Low
		Volume#2	DataBlock#2-1	High
			DataBlock#2-2	High
			DataBlock#2-3	Low
		Volume#3	DataBlock#3-1	Low
			DataBlock#3-2	Mid
			DataBlock#3-3	Mid
Host#2	Offline	Volume#4	DataBlock#4-1	Low
			DataBlock#4-2	Low
			DataBlock#4-3	Low
		Volume#5	DataBlock#5-1	Low
			DataBlock#5-2	Low
			DataBlock#5-3	Low

図 13

【図 14】

図 14

データブロックID	ライトアクセス量	日時
DataBlock#1-1	255 KB	2009/12/01 18:00
	0 KB	2009/12/01 19:00
	20 KB	2009/12/01 20:00
DataBlock#1-2	1355 KB	2009/12/01 18:00
	2300 KB	2009/12/01 19:00
	1420 KB	2009/12/01 20:00
DataBlock#1-3	22542 KB	2009/12/01 18:00
	35144 KB	2009/12/01 19:00
	28755 KB	2009/12/01 20:00

【図 15】

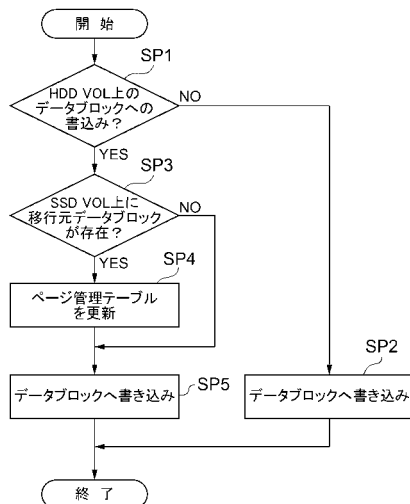
図 15

データブロックID	移行元ボリューム	移行先ボリューム	移行実施日時
DataBlock#1-1	SSDVolume#1	HDDVolume#3	2009/12/01 10:00
	HDDVolume#3	SSDVolume#1	2009/12/01 15:00
	SSDVolume#1	HDDVolume#2	2009/12/01 18:00
	HDDVolume#2	SSDVolume#3	2009/12/01 20:00
DataBlock#1-2	HDDVolume#3	SSDVolume#1	2009/12/01 00:00
	SSDVolume#1	HDDVolume#1	2009/12/01 19:00
DataBlock#1-3	SSDVolume#1	HDDVolume#3	2009/12/01 18:00

35

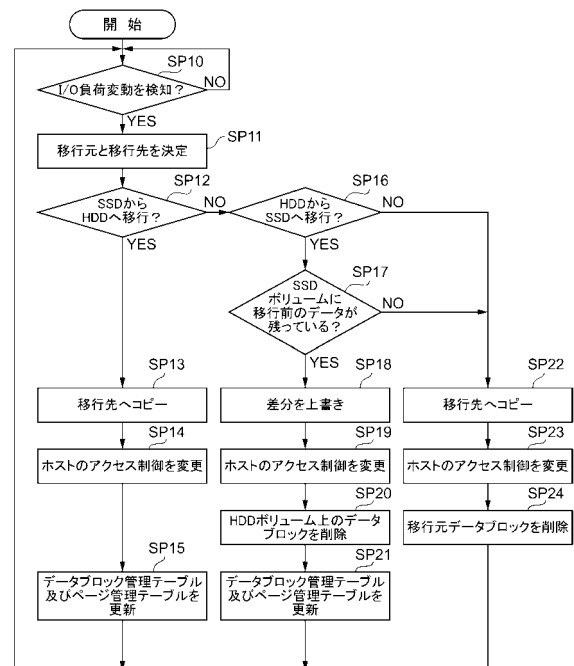
【図 16】

図 16



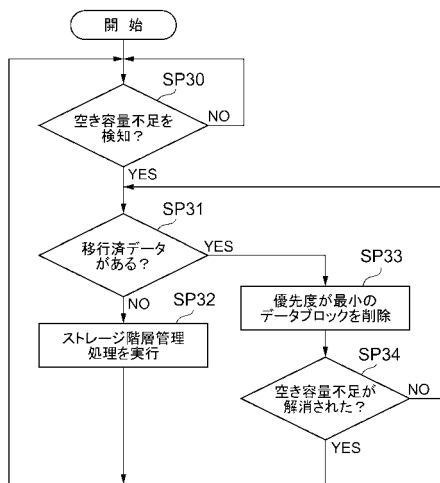
【図 17】

図 17



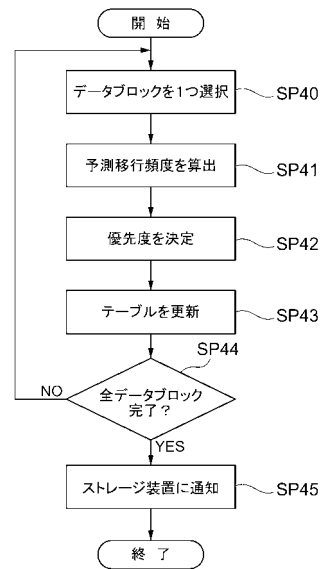
【図 18】

図 18



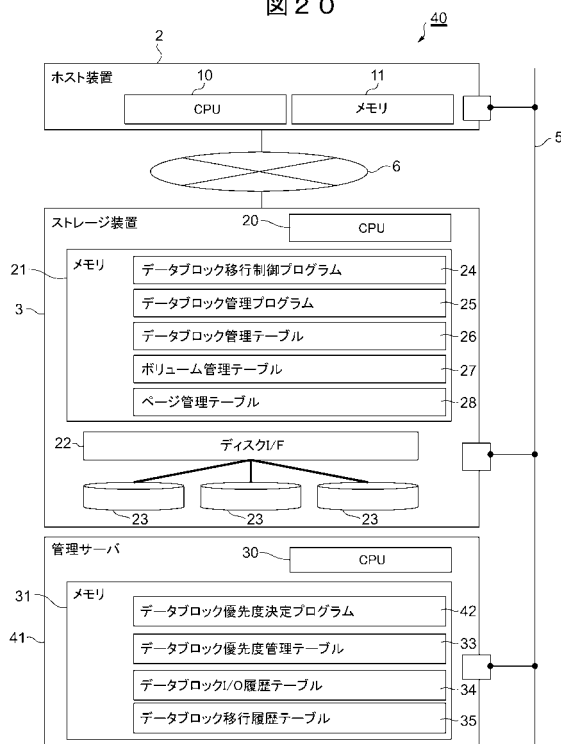
【図 19】

図 19



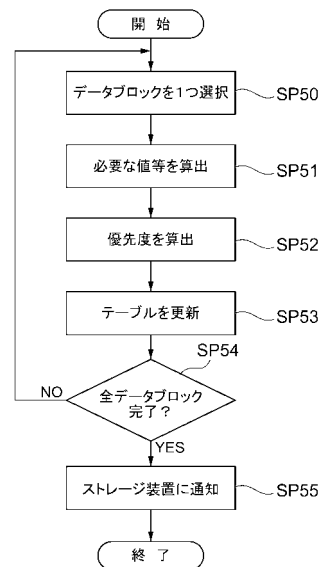
【図 20】

図 20



【図 21】

図 21



フロントページの続き

(56)参考文献 国際公開第2009/102425 (WO, A1)

特開2006-092124 (JP, A)

特開平01-021554 (JP, A)

特開平09-128302 (JP, A)

特開平04-336351 (JP, A)

特開2003-216460 (JP, A)

特開平08-305614 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00

G06F 3/06