

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5595530号  
(P5595530)

(45) 発行日 平成26年9月24日 (2014.9.24)

(24) 登録日 平成26年8月15日 (2014.8.15)

(51) Int. Cl.

F I

G06F 12/00 (2006.01)  
G06F 3/06 (2006.01)

G06F 12/00 511Z  
G06F 3/06 301Z  
G06F 12/00 514E  
G06F 12/00 545A  
G06F 3/06 304F

請求項の数 7 (全 40 頁) 最終頁に続く

(21) 出願番号 特願2012-557110 (P2012-557110)  
(86) (22) 出願日 平成22年10月14日 (2010.10.14)  
(65) 公表番号 特表2013-535037 (P2013-535037A)  
(43) 公表日 平成25年9月9日 (2013.9.9)  
(86) 国際出願番号 PCT/JP2010/006117  
(87) 国際公開番号 W02012/049711  
(87) 国際公開日 平成24年4月19日 (2012.4.19)  
審査請求日 平成24年12月20日 (2012.12.20)

(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 110000279  
特許業務法人ウィルフォート国際特許事務所  
(72) 発明者 森下 昇  
神奈川県横浜市戸塚区吉田町292番地  
株式会社日立製作所 横浜研究所内  
(72) 発明者 斎藤 秀雄  
神奈川県横浜市戸塚区吉田町292番地  
株式会社日立製作所 横浜研究所内  
(72) 発明者 江口 賢哲  
神奈川県横浜市戸塚区吉田町292番地  
株式会社日立製作所 横浜研究所内  
最終頁に続く

(54) 【発明の名称】 データ移行システム及びデータ移行方法

(57) 【特許請求の範囲】

【請求項1】

第1の記憶装置システムと、  
前記第1の記憶装置システムに接続された第2の記憶装置システムと、  
前記第1及び第2の記憶装置システムに接続されたホストコンピュータを含む上位装置と  
を有し、  
前記第1の記憶装置システムは、  
前記ホストコンピュータに認識される第1の上位論理ボリュームである移行元ボリュームと、  
前記移行元ボリュームに関連付けられておりシンプロビジョニングに従う仮想的な論理ボリュームであり複数の仮想領域で構成された第1の下位論理ボリュームと、  
それぞれが複数の第1の物理領域で構成されアクセス性能の異なる複数の記憶階層を有する記憶領域である第1のプールと、  
前記移行元ボリュームの第1のシンプロビジョニング情報と  
を有し、  
前記第1のシンプロビジョニング情報は、前記第1の下位論理ボリュームの仮想領域毎にどの第1の物理領域が割り当てられているかを前記第1のプールの記憶階層毎に表す情報を含んだ情報であり、  
前記第1の記憶装置システムは、前記移行元ボリュームを指定したライト要求を受けた

場合、前記移行元ボリュームの前記第 1 の下位論理ボリュームにおけるライト先の仮想領域に前記第 1 のプールから前記第 1 の物理領域を割り当て、その第 1 の物理領域にライト対象のデータを書き込み、前記第 1 のシンプロビジョニング情報を更新し、

前記第 2 の記憶装置システムは、

前記ホストコンピュータに認識される第 2 の上位論理ボリュームである移行先ボリュームと、

前記移行先ボリュームに関連付けられている仮想的な論理ボリュームである第 2 の下位論理ボリュームと、

前記移行先ボリュームとペアを構成し前記移行先ボリュームをコピー元とした第 3 の上位論理ボリュームであるコピー先ボリュームと、

前記コピー先ボリュームに関連付けられておりシンプロビジョニングに従う仮想的な論理ボリュームであり複数の仮想領域で構成された第 3 の下位論理ボリュームと、

それぞれが複数の第 2 の物理領域で構成されアクセス性能の異なる複数の記憶階層を有する記憶領域である第 2 のプールと、

を有し、

( A ) 前記第 2 の記憶装置システムが、ストレージ仮想化技術に従い前記第 2 の下位論理ボリュームに前記移行元ボリュームをマッピングし、

( B ) 前記上位装置が、前記ホストコンピュータからのアクセスパスを、前記移行元ボリュームへのアクセスパスから、前記移行先ボリュームへのアクセスパスへ切り替え、それにより、前記第 2 の記憶装置システムが、前記移行先ボリュームを指定したアクセス要求を前記ホストコンピュータから受信した場合、そのアクセス要求に従うアクセスが、前記第 2 の下位論理ボリュームを介して前記移行元ボリュームに対して行われ、

( C ) 前記第 2 の記憶装置システムが、前記第 1 のシンプロビジョニング情報が有する情報を取得し、

( D ) 前記第 2 の記憶装置システムが、前記移行元ボリュームから前記移行先ボリュームにデータを移行しそのデータを前記移行先ボリュームから前記コピー先ボリュームを通じて前記第 3 の下位論理ボリュームにコピーするコピー処理を実行し、

( E ) 前記第 2 の記憶装置システムが、前記コピー処理の完了後、前記移行先ボリュームに、前記第 2 の下位論理ボリュームに代えて前記第 3 の下位論理ボリュームを関連付け、

前記第 2 の記憶装置システムは、前記コピー処理の最中であっても、前記移行先ボリュームを指定したライト要求を前記ホストコンピュータから受け付けており、前記コピー処理の最中に前記移行先ボリュームを指定したライト要求を前記ホストコンピュータから受けた場合、前記移行元ボリュームを指定したライト要求を前記第 1 の記憶装置システムに転送し、

前記コピー処理において、前記第 2 の記憶装置システムが、前記移行対象のデータが無効なデータであるか否かを判断し、その判断の結果が肯定的の場合に、前記移行対象のデータを前記コピー先ボリュームの第 3 の下位論理ボリュームに書き込むことなく破棄し、

前記コピー処理では、前記 ( C ) で取得した情報を基に、第 1 の物理領域が割り当てられている仮想領域が特定され、その仮想領域に割り当てられている第 1 の物理領域内のデータが移行され、第 1 の物理領域が割り当てられていない仮想領域についてデータは移行されず、

前記特定された仮想領域は、前記コピー処理においてデータが未だ 1 度も移行されていない仮想領域、又は、データが移行された後に前記ホストコンピュータから前記移行先ボリュームを通じて新たにデータの書き込み先とされた仮想領域であり、

前記コピー処理では、前記第 3 の下位論理ボリュームの未割当ての仮想領域がデータのコピー先とされた場合、その仮想領域に前記第 2 のプールから未割当ての第 2 の物理領域が割り当てられ、割り当てられた第 2 の物理領域に移行対象のデータが書き込まれ、

前記 ( C ) と前記 ( D ) の間に、下記の ( P ) が行われ、

( P ) 前記第 2 の記憶装置システムが、前記 ( C ) で取得した情報を基に、前記移行元ボ

10

20

30

40

50

リユームに割当てられた 1 以上の第 1 の物理領域の記憶容量を、前記第 1 のプールの記憶階層毎に特定し、前記第 1 のプールの記憶階層毎の特定された記憶容量を基に、前記第 1 のプールの複数の記憶階層に対応する、前記第 2 のプールの複数の記憶階層の各々について、前記第 2 のプールに必要な空き容量を算出し、前記第 2 のプールの複数の記憶階層の各々について、算出された空き容量分の 1 以上の第 2 の物理領域を予約し、

前記コピー処理では、前記第 2 のプールの複数の記憶階層の各々について、前記 ( P ) で予約された第 2 の物理領域が、前記コピー先ボリュームの仮想領域に割り当てられる、データ移行システム。

【請求項 2】

請求項 1 記載のデータ移行システムであって、

前記 ( P ) では、前記第 2 の記憶装置システムが、前記 ( C ) で取得した情報から、前記第 1 の下位論理ボリュームに割り当てられている第 1 の物理領域の数と、第 1 の物理領域のサイズとを特定し、第 1 の物理領域の数と、第 1 の物理領域のサイズと、第 2 の物理領域のサイズとを基に、前記第 2 のプールに必要な空き容量を算出する、データ移行システム。

【請求項 3】

請求項 1 又は 2 記載のデータ移行システムであって、

前記第 1 及び第 2 の記憶装置システムは、それぞれ、ライト要求に従うライト対象のデータを一時的に記憶するキャッシュメモリを有し、

前記 ( A ) において、前記第 2 の記憶装置システムが、前記移行先ボリュームを指定したライト要求に従うライト対象データを前記第 2 の記憶装置システムのキャッシュメモリに書き込むことであるキャッシュメモリ動作を抑止し、

前記 ( B ) と前記 ( C ) との間で、前記第 2 の記憶装置システムが、前記キャッシュメモリ動作を抑止を解除する、データ移行システム。

【請求項 4】

請求項 1 乃至 3 のうちのいずれか 1 項に記載のデータ移行システムであって、

前記コピー処理では、前記第 2 の記憶装置システムが、前記 ( C ) で取得した情報を基に、前記第 3 の下位論理ボリュームの仮想領域に物理領域を割り当てる、データ移行システム。

【請求項 5】

請求項 1 乃至 4 のうちのいずれか 1 項に記載のデータ移行システムであって、

前記コピー先ボリュームは、前記 ( A ) の後に動的に定義された論理ボリュームであり、

前記 ( E ) において、前記第 2 の記憶装置システムが、前記コピー先ボリュームの定義を解除する、データ移行システム。

【請求項 6】

請求項 1 乃至 5 のうちのいずれか 1 項に記載のデータ移行システムであって、

前記上位装置は、前記ホストコンピュータの他に、前記ホストコンピュータと前記第 1 及び第 2 の記憶装置システムとの間に介在するスイッチ装置を含み、

前記 ( B ) では、前記スイッチ装置が、前記ホストコンピュータからのアクセスパスを、前記移行元ボリュームへのアクセスパスから、前記移行先ボリュームへのアクセスパスへ切り替える、

データ移行システム。

【請求項 7】

第 1 の記憶装置システムとそれに接続された第 2 の記憶装置システムと前記第 1 及び第 2 の記憶装置システムに接続されたホストコンピュータとを含む計算機システムでのデータ移行方法であって、

前記第 1 の記憶装置システムは、

10

20

30

40

50

前記ホストコンピュータに認識される第1の上位論理ボリュームである移行元ボリュームと、

前記移行元ボリュームに関連付けられておりシンプロビジョニングに従う仮想的な論理ボリュームであり複数の仮想領域で構成された第1の下位論理ボリュームと、

それぞれが複数の第1の物理領域で構成されアクセス性能の異なる複数の記憶階層を有する記憶領域である第1のプールと、

前記移行元ボリュームの第1のシンプロビジョニング情報とを有し、

前記第1のシンプロビジョニング情報は、前記第1の下位論理ボリュームの仮想領域毎にどの第1の物理領域が割り当てられているかを前記第1のプールの記憶階層毎に表す情報を含んだ情報であり、

10

前記第1の記憶装置システムは、前記移行元ボリュームを指定したライト要求を受けた場合、前記移行元ボリュームの前記第1の下位論理ボリュームにおけるライト先の仮想領域に前記第1のプールから前記第1の物理領域を割り当て、その第1の物理領域にライト対象のデータを書き込み、前記第1のシンプロビジョニング情報を更新し、

前記第2の記憶装置システムは、

前記ホストコンピュータに認識される第2の上位論理ボリュームである移行先ボリュームと、

前記移行先ボリュームに関連付けられている仮想的な論理ボリュームである第2の下位論理ボリュームと、

20

前記移行先ボリュームとペアを構成し前記移行先ボリュームをコピー元とした第3の上位論理ボリュームであるコピー先ボリュームと、

前記コピー先ボリュームに関連付けられておりシンプロビジョニングに従う仮想的な論理ボリュームであり複数の仮想領域で構成された第3の下位論理ボリュームと、

それぞれが複数の第2の物理領域で構成されアクセス性能の異なる複数の記憶階層を有する記憶領域である第2のプールと、

を有し、

前記データ移行方法では、以下の(A)~(E)を行い、

(A)ストレージ仮想化技術に従い前記第2の下位論理ボリュームに前記移行元ボリュームをマッピングする、

30

(B)前記ホストコンピュータからのアクセスパスを、前記移行元ボリュームへのアクセスパスから、前記移行先ボリュームへのアクセスパスへ切り替え、それにより、前記第2の記憶装置システムが、前記移行先ボリュームを指定したアクセス要求を前記ホストコンピュータから受信した場合、そのアクセス要求に従うアクセスが、前記第2の下位論理ボリュームを介して前記移行元ボリュームに対して行われる、

(C)前記第1のシンプロビジョニング情報が有する情報が前記第2の記憶装置システムに送られる、

(D)前記移行元ボリュームから前記移行先ボリュームにデータを移行しそのデータを前記移行先ボリュームから前記コピー先ボリュームを通じて前記第3の下位論理ボリュームにコピーするコピー処理を実行する、

40

(E)前記コピー処理の完了後、前記移行先ボリュームに、前記第2の下位論理ボリュームに代えて前記第3の下位論理ボリュームを関連付ける、

前記第2の記憶装置システムは、前記コピー処理の最中であっても、前記移行先ボリュームを指定したライト要求を前記ホストコンピュータから受け付けており、前記コピー処理の最中に前記移行先ボリュームを指定したライト要求を前記ホストコンピュータから受けた場合、前記移行元ボリュームを指定したライト要求を前記第1の記憶装置システムに転送し、

前記コピー処理において、前記第2の記憶装置システムが、前記移行対象のデータが無効なデータであるか否かを判断し、その判断の結果が肯定的の場合に、前記移行対象のデータを前記コピー先ボリュームの第3の下位論理ボリュームに書き込むことなく破棄し、

50

前記コピー処理では、前記(C)で取得した情報を基に、第1の物理領域が割り当てられている仮想領域が特定され、その仮想領域に割り当てられている第1の物理領域内のデータが移行され、第1の物理領域が割り当てられていない仮想領域についてデータは移行されず、

前記特定された仮想領域は、前記コピー処理においてデータが未だ1度も移行されていない仮想領域、又は、データが移行された後に前記ホストコンピュータから前記移行先ボリュームを通じて新たにデータの書き込み先とされた仮想領域であり、

前記コピー処理では、前記第3の下位論理ボリュームの未割当ての仮想領域がデータの  
コピー先とされた場合、その仮想領域に前記第2のプールから未割当ての第2の物理領域  
が割り当てられ、割り当てられた第2の物理領域に移行対象のデータが書き込まれ、

10

前記(C)と前記(D)の間に、下記の(P)が行われ、

(P)前記第2の記憶装置システムが、前記(C)で取得した情報を基に、前記移行元ボ  
リュームに割当てられた1以上の第1の物理領域の記憶容量を、前記第1のプールの記憶  
階層毎に特定し、前記第1のプールの記憶階層毎の特定された記憶容量を基に、前記第1  
のプールの複数の記憶階層に対応する、前記第2のプールの複数の記憶階層の各々につ  
いて、前記第2のプールに必要な空き容量を算出し、前記第2のプールの複数の記憶階層の  
各々について、算出された空き容量分の1以上の第2の物理領域を予約し、

前記コピー処理では、前記第2のプールの複数の記憶階層の各々について、前記(P)  
で予約された第2の物理領域が、前記コピー先ボリュームの仮想領域に割り当てられる、  
データ移行方法。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、記憶装置システム間でデータを移行する技術に関する。

【背景技術】

【0002】

第1の記憶装置システムから第2の記憶装置システムへデータを移行するデータ移行技術が知られている。例えば、記憶装置システムが新たに導入された場合、既存の記憶装置システムから新たに導入された第1の記憶装置システムへデータを移行することがある。また、記憶装置システムを使用中に、高負荷の記憶装置システムから低負荷の記憶装置システムへデータを移行することがある。データ移行は、記憶装置システム間の他、記憶装置システム内の記憶装置間でも行うことがある。

30

【0003】

特許文献1には、記憶装置システム間において、ホストからのアクセスを維持したまま、ホストコンピュータ(以下、ホスト)に透過的に、データを移行する技術が開示されている。特許文献1によれば、ホストと第2の記憶装置システム(データ移行元となる既存の記憶装置システム)との間に、第1の記憶装置システム(データ移行先となる記憶装置システム)が接続される。第1の記憶装置システムが、ホストから移行対象データに対するアクセス要求(リード要求及び/又はライト要求)を受け付けながら、第2の記憶装置システムから第1の記憶装置システムへの移行対象データの移行が行われる。第1の記憶装置システムが、未だ移行されていない移行対象データに対しアクセス要求を受領した場合、受領したアクセス要求の処理に必要な部分のデータを読み出すために第2の記憶装置システムへリード要求を発行する。第1の記憶装置システムは、そのリード要求に回答して取得されたデータを用いて、上記受領したアクセス要求を処理する。

40

【0004】

また、例えば、ストレージ仮想化技術が知られている。ストレージ仮想化技術によれば、第1の記憶装置システムに第2の記憶装置システムが接続され、第2の記憶装置システムの記憶領域が、第1の記憶装置システムの記憶領域としてホストに提供される。第1の記憶装置システムが、ホストからアクセス要求を受信した場合、そのアクセス要求に従うアクセス先が第1の記憶装置システムの記憶領域か第2の記憶装置システムの記憶領域か

50

を特定し、特定したアクセス先が第2の記憶装置システムの記憶領域の場合、第2の記憶装置システムへアクセス要求を転送する。

【0005】

また、例えば、シンプロビジョニングが知られている。シンプロビジョニングによれば、複数の仮想ページ（仮想的な記憶領域）で構成された仮想的な論理ボリューム（以下、TPボリューム）と、複数の物理ページ（物理的な記憶領域）で構成されたプールが定義される。未割当ての仮想ページへの書き込みに応答して、その仮想ページに、未割当ての物理ページが割り当てられる。

【0006】

特許文献2は、シンプロビジョニング機能を有する第2の記憶装置システムの記憶領域を、ストレージ仮想化技術を用いて第1の記憶装置システムで利用する技術を開示している。

10

【先行技術文献】

【特許文献】

【0007】

【特許文献1】米国特許第6108748号明細書

【特許文献2】特開2010-055557号明細書

【発明の概要】

【発明が解決しようとする課題】

【0008】

20

以下の説明では、ホストにとって意味の無いデータ（例えば、全ビットが0で構成されたデータであるゼロデータのような、ホストで実行されるアプリケーションプログラムにとってリード及び/又はライトされていない初期値データ）を「無効データ」と言い、ホストにとって意味のあるデータ（例えば、ホストで実行されるアプリケーションプログラムにとってリード及び/又はライトされるデータ）を「有効データ」と言う。

【0009】

TPボリュームからデータを移行する場合、移行元ボリュームの未割当ての仮想ページについては、移行元の記憶装置システムから移行先の記憶装置システムへ、無効データが転送される。このため、TPボリュームに割り当てられている1以上の物理ページが記憶するデータの総量がTPボリュームより少なくても、TPボリュームの容量分のデータが移行されることになる。

30

【0010】

また、移行先ボリュームもTPボリュームの場合、移行先の仮想ページへ転送されるデータが無効データであっても、移行先の仮想ページに、物理ページが割り当てられる。移行先記憶装置システムにおいて消費される記憶容量が無駄に大きくなってしまう。

【0011】

本発明の目的は、第2の記憶装置システム内の外部ボリュームに対応付けられた第1の記憶装置システム内の論理ボリュームから第2の記憶装置システム内のTPボリュームへのデータ移行において、移行対象データの総量、及び、移行先の記憶装置システムで消費される記憶容量を削減し、且つ、そのデータ移行を、ホストから移行対象データへのアクセス要求を受け付けながら実行できるようにすることにある。

40

【課題を解決するための手段】

【0012】

データ移行システムは、第1の記憶装置システムと、第1の記憶装置システムに接続された第2の記憶装置システムと、第1及び第2の記憶装置システムに接続されたホストコンピュータを含む上位装置とを有する。

【0013】

第1の記憶装置システムは、ホストコンピュータに認識される第1の上位論理ボリュームである移行元ボリュームと、移行元ボリュームに関連付けられておりシンプロビジョニングに従う仮想的な論理ボリュームであり複数の仮想領域で構成された第1の下位論理ボ

50

リュームと、複数の第1の物理領域で構成された記憶領域である第1のプールと、移行元ボリュームの第1のシンプロビジョニング情報とを有する。第1のシンプロビジョニング情報は、第1の下位論理ボリュームの仮想領域毎にどの第1の物理領域が割り当てられているかを表す情報を含んだ情報である。第1の記憶装置システムは、移行元ボリュームを指定したライト要求を受けた場合、移行元ボリュームの第1の下位論理ボリュームにおけるライト先の仮想領域に第1のプールから第1の物理領域を割り当て、その第1の物理領域にライト対象のデータを書き込み、第1のシンプロビジョニング情報を更新する。

【0014】

第2の記憶装置システムは、ホストコンピュータに認識される第2の上位論理ボリュームである移行先ボリュームと、移行先ボリュームに関連付けられている仮想的な論理ボリュームである第2の下位論理ボリュームと、移行先ボリュームとペアを構成し移行先ボリュームをコピー元とした第3の上位論理ボリュームであるコピー先ボリュームと、コピー先ボリュームに関連付けられておりシンプロビジョニングに従う仮想的な論理ボリュームであり複数の仮想領域で構成された第3の下位論理ボリュームと、複数の第2の物理領域で構成された記憶領域である第2のプールとを有する。

10

【0015】

データ移行システムでは、以下の処理が行われる。

(A) 第2の記憶装置システムが、ストレージ仮想化技術に従い第2の下位論理ボリュームに前記移行元ボリュームをマッピングする。

(B) 上位装置が、ホストコンピュータからのアクセスパスを、前記移行元ボリュームへのアクセスパスから、前記移行先ボリュームへのアクセスパスへ切り替える。それにより、第2の記憶装置システムが、移行先ボリュームを指定したアクセス要求をホストコンピュータから受信した場合、そのアクセス要求に従うアクセスが、第2の下位論理ボリュームを介して前記移行元ボリュームに対して行われる。

20

(C) 第2の記憶装置システムが、第1のシンプロビジョニング情報が有する情報を取得する。

(D) 第2の記憶装置システムが、コピー処理を実行する。コピー処理では、移行元ボリュームから移行先ボリュームにデータを移行することと、そのデータを移行先ボリュームからコピー先ボリュームを通じて第3の下位論理ボリュームにコピーすることを含む処理が行われる。

30

(E) 第2の記憶装置システムが、コピー処理の完了後、移行先ボリュームに、第2の下位論理ボリュームに代えて第3の下位論理ボリュームを関連付ける。

【0016】

第2の記憶装置システムは、コピー処理の最中であっても、移行先ボリュームを指定したライト要求をホストコンピュータから受け付けており、コピー処理の最中に移行先ボリュームを指定したライト要求をホストコンピュータから受けた場合、移行元ボリュームを指定したライト要求を第1の記憶装置システムに転送する。

【0017】

コピー処理では、上記(C)で取得した情報を基に、第1の物理領域が割り当てられている仮想領域が特定され、その仮想領域に割り当てられている第1の物理領域内のデータが移行される。コピー処理では、第1の物理領域が割り当てられていない仮想領域についてデータは移行されない。特定された仮想領域は、コピー処理においてデータが未だ1度も移行されていない仮想領域、又は、データが移行された後にホストコンピュータから移行先ボリュームを通じて新たにデータの書き込み先とされた仮想領域である。

40

【0018】

記憶装置システム間でデータを移行する1つの方法として、同期型リモートコピーがある。同期型リモートコピーは、例えば次のとおりである。すなわち、或る記憶装置システムXの或る論理ボリュームXと別の記憶装置システムYの或る論理ボリュームYとの間でペアが構成される。記憶装置システムXが、ホストコンピュータから論理ボリュームXを指定したライト要求を受けた場合、そのライト要求に対する応答を、下記の(a)及び(

50

b) が完了したときに、ホストコンピュータに対して返すことができる。

(a) 記憶装置システム X が、論理ボリューム X (又は記憶装置システム X のキャッシュメモリ) にライト対象のデータを書き込む。

(b) 記憶装置システム X が、ライト対象のデータを記憶装置システム Y に転送し、記憶装置システム Y から所定の応答を受領する。

【0019】

同期型リモートコピーが並行して行われるペアの数は、制限されることが好ましい。なぜなら、同期型リモートコピーが並行して行われるペアの数が多いと、記憶装置システム X のアクセス性能 (例えば、単位時間当りに処理できるアクセス要求の数) が低下するからである。また、外部接続と内部コピーの組み合わせの場合、移行元ボリュームを有する記憶装置システムの機能適用状況による移行可否の制限を受けにくいということがある。リモートコピー機能は性能等の理由により、コピーペア数に制限があり、移行元ボリュームに既にリモートコピー機能が適用されていた場合、移行のためにリモートコピー機能を追加適用するのが困難な場合がある。

10

【0020】

そこで、本発明の 1 つの観点によれば、外部接続 (ストレージ仮想化技術に従うボリュームマッピング) と内部コピー (第 2 の記憶装置システムにおける、移行先ボリュームからコピー先ボリュームへのコピー) との組み合わせで、記憶装置システム間のデータ移行が行われる。外部接続の数、及び、内部コピーが並行して行われるペアの数は、同期型リモートコピーが並行して行われるペアの数より多くて良い。このため、ストレージ仮想化環境においてデータ移行が効率的に行われることが期待される。

20

【発明の効果】

【0021】

本発明によれば、第 2 の記憶装置システム内の外部ボリュームに対応付けられた第 1 の記憶装置システム内の論理ボリュームから第 2 の記憶装置システム内の TP ボリュームへのデータ移行において、移行対象データの総量、及び、移行先の記憶装置システムで消費される記憶容量を削減し、且つ、そのデータ移行を、ホストから移行対象データへのアクセス要求を受け付けながら実行することができる。

【図面の簡単な説明】

【0022】

【図 1】本発明の実施例 1 の概要を示す。

【図 2】記憶装置システム 2 a (2 b) の構成を示す。

【図 3】記憶装置システム 2 a (2 b) が有する管理情報を示す。

【図 4】論理ボリューム管理情報 6 0 の構成を示す。

【図 5】仮想ボリューム管理情報 6 1 の構成を示す。

【図 6】ドライブグループ管理情報 6 2 の構成を示す。

【図 7】外部ボリューム管理情報 6 3 の構成を示す。

【図 8】外部ボリュームパス情報 6 3 1 の構成を示す。

【図 9】記憶階層管理情報 6 4 の構成を示す。

【図 10】プール管理情報 6 5 の構成を示す。

【図 11】ページ管理情報 6 5 1 の構成を示す。

【図 12】マッピング管理情報 6 6 の構成を示す。

【図 13】各種管理情報 6 0、6 1、6 2、6 5、6 6 及び 6 3 の関連性を示す。

【図 14】コピー管理情報 6 7 の構成を示す。

【図 15】ゼロデータ破棄管理情報 6 8 の構成を示す。

【図 16】ボリューム移行処理の流れの概要を示す。

【図 17】プール空き容量確認処理の流れを示す。

【図 18】プール空き容量確認メイン処理の流れを示す。

【図 19】コピー処理の流れを示す。

【図 20】通常コピー処理の流れを示す。

30

40

50



【図 2 1】 シンプロビジョニング (TP) 情報活用コピー処理 (個別) の流れを示す。

【図 2 2】 TP 情報活用コピー処理 (一括) の流れを示す。

【図 2 3】 記憶装置システム 2 a (2 b) で実行されるリード処理を示す。

【図 2 4】 通常ボリュームリード処理の流れを示す。

【図 2 5】 TP ボリュームリード処理の流れを示す。

【図 2 6】 ライト処理を示す。

【図 2 7】 通常ボリュームライト処理の流れを示す。

【図 2 8】 TP ボリュームライト処理の流れを示す。

【図 2 9】 非同期ライト処理の流れを示す。

【図 3 0】 TP 情報参照処理を示す。

10

【図 3 1】 ゼロデータ破棄処理の流れを示す。

【図 3 2】 本発明の実施例 2 に係る計算機システムの構成を示す。

【発明を実施するための形態】

【0023】

以下、図面を参照して、本発明の幾つかの実施例を説明する。

【0024】

なお、以下の説明では、「xxx テーブル」或いは「xxx リスト」の表現にて各種情報を説明することがあるが、各種情報は、テーブル及びリスト以外のデータ構造で表現されていてもよい。データ構造に依存しないことを示すために「xxx テーブル」及び「xxx リスト」を「xxx 情報」と呼ぶことができる。

20

【0025】

また、以下の説明では、種々の対象の識別情報として、番号が使用されるが、番号以外種類の識別情報 (例えば、英字や符号を含んだ識別子) も採用可能である。

【0026】

また、以下の説明では、「プログラム」を主語として処理を説明する場合があるが、プログラムは、プロセッサ (例えば CPU (Central Processing Unit)) によって実行されることで、定められた処理を、適宜に記憶資源 (例えばメモリ) 及び/又は通信インタフェース装置 (例えば通信ポート) を用いながら行うため、処理の主語がプロセッサとされてもよい。プログラムを主語として説明された処理は、記憶装置システム或いはそのコントローラが行う処理としても良い。また、プロセッサは、プロセッサが行う処理の一部又は全部を行うハードウェア回路を含んでも良い。コンピュータプログラムは、プログラムソースから各計算機にインストールされても良い。プログラムソースは、例えば、プログラム配布サーバ又は記憶メディアであっても良い。

30

【実施例 1】

【0027】

図 1 は、本発明の実施例 1 の概要を示す。なお、以下の説明では、シンプロビジョニングを「TP」と略す。

【0028】

データの移行元となる移行元記憶装置システム 2 a と、データの移行先となる移行先記憶装置システム 2 b が、ホストコンピュータ (以下、ホスト) 1 と接続されている。また、移行元記憶装置システム 2 a、移行先記憶装置システム 2 b、及びホスト 1 は、管理端末 4 に接続されている。管理端末 4 は、ホスト 1、記憶装置システム 2 a 及び 2 b を管理する。

40

【0029】

ホスト 1 では、アプリケーションプログラム 1 1 とパス切替えプログラム 1 2 が動作する。パス切替えプログラム 1 2 は、アプリケーションプログラム 1 1 からのアクセス要求の発行先となるボリュームを、切り替える。その切り替えは、アプリケーションプログラム 1 1 に透過的に (アプリケーションプログラム 1 1 に意識されることなく)、行われる。

【0030】

50

記憶装置システム 2 a ( 2 b ) は、論理ボリュームに書き込まれるデータ又は論理ボリュームから読み出されたデータを一時的に記憶するキャッシュメモリ 2 1 5 a ( 2 1 5 b ) を有する。

【 0 0 3 1 】

実施例 1 に係るデータ移行の概要は、次の通りである。

( 1 ) 移行元記憶装置システム 2 a の移行元ボリューム 2 3 a が、移行先記憶装置システム 2 b の移行先ボリューム 2 3 b として定義される。具体的には、ストレージ仮想化技術に従い、移行元ボリューム 2 3 a が、外部ボリュームとして、移行先ボリューム 2 3 b の仮想ボリューム 2 4 b にマッピングされる。

( 2 ) ホスト 1 からのアクセス先が、移行元ボリューム 2 3 a から移行先ボリューム 2 3 b に切り替えられる。具体的には、パス切替えプログラム 1 2 が、アプリケーションプログラム 1 1 に認識される L U ( 論理ユニット ) に対応付けられる論理ボリュームを、移行元ボリューム 2 3 a から移行先ボリューム 2 3 b に切り替える。

( 3 ) 移行先ボリューム 2 3 b のコピー先として或る論理ボリューム ( 以下、コピー先ボリューム ) 2 3 c が定義される。

( 4 ) 移行元の記憶装置システム 2 a 内の T P 情報 ( 以下、移行元 T P 情報 ) を基に、移行元ボリューム 2 3 a に関連付けられている仮想ボリューム 2 4 a から、移行先ボリューム 2 3 b に関連付けられている仮想ボリューム 2 4 b に、データが移行される。移行元の T P 情報に基づいてデータが移行されるので、移行されるデータは、有効データのみ ( すなわち、仮想ボリューム 2 4 a に割り当てられている物理ページ内のデータのみ ) である。仮想ボリューム 2 4 a には、移行元の記憶装置システム 2 a 内のプール ( 以下、移行元のプール ) 2 5 a から物理ページが割り当てられる。移行元のプール 2 5 a は、例えば、1 以上の実ボリューム 2 6 a から構成されている。各実ボリューム 2 6 a が、1 以上の物理ページに分割される。なお、実施例 1 において、「実ボリューム」は、記憶装置システム内の物理記憶デバイス ( 以下、ドライブ ) に基づく論理ボリュームである。プールを構成する実ボリュームを「プールボリューム」と言うことがある。また、実施例 1 において、「仮想ボリューム」は、記憶装置システム 2 内のボリューム管理単位である。T P ボリュームである仮想ボリュームもある。T P ボリュームは、T P に従う仮想的な論理ボリュームであり、複数の仮想ページで構成されている。また、「T P 情報」は、どの仮想ページにどの物理ページが割り当てられているかを表す情報を含む。本実施例では、T P ボリュームは、仮想ボリューム 2 4 a 及び 2 4 c である。

( 5 ) 移行先ボリューム 2 3 b に関連付けられている仮想ボリューム 2 3 b に移行したデータが、移行先ボリューム 2 3 b からコピー先ボリューム 2 3 c にコピーされる。それにより、コピー先ボリューム 2 3 c に関連付けられている仮想ボリューム 2 4 c に、データが書き込まれる。それに伴い、仮想ボリューム 2 4 c に、移行先の記憶装置システム 2 b 内のプール ( 以下、移行先のプール ) 2 5 b から物理ページが割り当てられ、その物理ページにデータが書き込まれる。その際、移行先の記憶装置システム 2 b 内の T P 情報 ( 以下、移行先 T P 情報 ) が更新される。

( 6 ) 移行先ボリューム 2 3 b からコピー先ボリューム 2 3 c へのデータのコピーが完了した後、仮想ボリューム切替え ( 論理ボリュームと仮想ボリュームとの関連付けの変更 ) が行われる。具体的には、コピー先ボリューム 2 3 c に関連付けられている仮想ボリューム 2 4 c が、コピー先ボリューム 2 3 c に代えて、移行先ボリューム 2 3 b に関連付けられる。一方、移行先ボリューム 2 3 b に関連付けられている仮想ボリューム 2 4 b が、移行先ボリューム 2 3 b に代えて、コピー先ボリューム 2 3 c に関連付けられる。

【 0 0 3 2 】

以上の ( 1 ) ~ ( 6 ) により、移行元プール 2 5 a における、仮想ボリューム 2 4 a に割り当てられていた物理ページ内のデータが、移行先プール 2 5 b に移行される。

【 0 0 3 3 】

図 2 は、記憶装置システム 2 a ( 2 b ) の構成を示す。

【 0 0 3 4 】

10

20

30

40

50

記憶装置システム 2 a ( 2 b ) は、複数の物理記憶デバイスと、それらに接続されたコントローラとを有する。具体的には、例えば、記憶装置システム 2 a ( 2 b ) は、記憶装置として、ディスクドライブ 2 0 0 を格納する記憶装置 2 0 を有する。また、記憶装置システム 2 a ( 2 b ) は、記憶制御装置として、記憶制御装置 2 1 を有する。記憶制御装置 2 1 は、例えば、

( \* ) ホスト 1 に接続されるインタフェースパッケージ ( フロントエンド ) 2 1 0 、  
 ( \* ) 記憶装置 2 0 に接続されるインタフェースパッケージ ( バックエンド ) 2 1 1 、  
 ( \* ) 1 以上のプロセッサ ( 例えば CPU ) を含んだプロセッサパッケージ 2 1 3 、  
 ( \* ) キャッシュメモリ 2 1 5 をメモリパッケージ 2 1 4 、  
 ( \* ) プロセッサパッケージに接続される保守端末 2 1 6 、  
 ( \* ) パッケージ 2 1 0 、 2 1 1 、 2 1 3 及び 2 1 4 に接続されパッケージ間の接続を切替えるスイッチ 2 1 2 ( 例えばクロスバスイッチ ) 、  
 を有する。スイッチ 2 1 2 に代えて他種の接続部 ( 例えばバス ) が採用されても良い。保守端末 2 1 6 は無くても良い。保守端末 2 1 6 は、例えば、管理端末 4 ( 図 1 参照 ) に接続されていて良い。

10

#### 【 0 0 3 5 】

記憶装置 2 0 は、1 以上のディスクドライブ 2 0 0 を有する。各ドライブ 2 0 0 は、入出力ポートを複数持つ。記憶装置システム 2 a ( 2 b ) に、記憶装置 2 0 は複数存在しても良い。また、ドライブ 2 0 0 に代えて又は加えて、他種の物理記憶デバイス ( 例えば、半導体記憶装置、フラッシュメモリ ) が採用されて良い。

20

#### 【 0 0 3 6 】

フロントエンド 2 1 0 は、ホスト 1 とキャッシュメモリ 2 1 5 の間のデータ転送を制御する。フロントエンド 2 1 0 は、ホスト 1 からアクセス要求を受領する他、他の記憶装置システムに対し、アクセス要求を発行してよい。バックエンド 2 1 1 は、キャッシュメモリ 2 1 5 とドライブ 2 0 0 の間のデータ転送を制御する。フロントエンド 2 1 0 及びバックエンド 2 1 1 の内部は、入出力ポートと、ローカルメモリと、データ転送の制御部とを有して良い。フロントエンド 2 1 0 及びバックエンド 2 1 1 は複数存在しても良い。

#### 【 0 0 3 7 】

プロセッサパッケージ 2 1 3 は、複数のプロセッサ、ローカルメモリ、データ転送の制御部を有して良い。ローカルメモリは、プロセッサで実行されるコンピュータプログラムを記憶することができる。具体的には、例えば、ローカルメモリは、例えば、リードプログラム 5 0 、ライトプログラム 5 1 、TP 情報参照プログラム 5 2 、プール空き容量確認プログラム 5 3 、コピープログラム 5 4 、ゼロデータ破棄プログラム 5 5 、非同期ライトプログラム 5 6 を記憶して良い。

30

#### 【 0 0 3 8 】

メモリパッケージ 2 1 4 は、キャッシュメモリ 2 1 5 と、データ転送の制御部とを有する。キャッシュメモリ 2 1 5 は、記憶装置システム 2 の管理情報を記憶して良い。キャッシュメモリ 2 1 5 は、ドライブ 2 0 0 ( 論理ボリュームの基になる物理記憶デバイス ) に書き込まれるデータ又はドライブ 2 0 0 から読み出されたデータを一時的に記憶する。

#### 【 0 0 3 9 】

プロセッサパッケージ 2 1 3 、メモリパッケージ 2 1 4 は複数存在しても良い。

40

#### 【 0 0 4 0 】

保守端末 2 1 6 は、記憶装置システム 2 の構成を変更したり、記憶装置システム 2 の内部状態を参照したりするために使用されて良い。また、保守端末 2 1 6 は、管理端末 4 から記憶装置システム 2 への指示を中継して良い。保守端末 2 3 は、プロセッサパッケージ 2 1 3 に接続されていて良い。

#### 【 0 0 4 1 】

図 3 は、記憶装置システム 2 a ( 2 b ) が有する管理情報を示す。

#### 【 0 0 4 2 】

管理情報は、論理ボリュームの階層管理に関する情報 ( 例えば、論理ボリュームが属す

50

る階層を表す情報)を含んだ情報である、管理情報は、例えば、下記の情報、

- (\*) 論理ボリュームに関する論理ボリューム管理情報 6 0、
  - (\*) 仮想ボリュームに関する仮想ボリューム管理情報 6 1、
  - (\*) ドライブグループに関するドライブグループ管理情報 6 2、
  - (\*) 外部ボリュームに関する外部ボリューム管理情報 6 3、
  - (\*) TPに関するTP情報 6 9、
  - (\*) コピー管理に関するコピー管理情報 6 7、及び、
  - (\*) ゼロデータ破棄に関するゼロデータ破棄管理情報 6 8、
- を含む。

【0043】

TP情報 6 9は、例えば、下記の情報、

- (\*) 記憶階層に関する記憶階層管理情報 6 4、
  - (\*) プールに関するプール管理情報 6 5、及び、
  - (\*) 仮想ページと物理ページの対応関係に関するマッピング管理情報 6 6、
- を含む。

【0044】

図4は、論理ボリューム管理情報 6 0の構成を示す。

【0045】

論理ボリューム 2 3は、記憶装置システム 2がホスト 1や他の記憶装置システム 2に対して提供する記憶領域の単位である。論理ボリューム管理情報 6 0は、例えば、論理ボリューム毎に存在する情報である。1つの論理ボリューム(図4の説明において「対象論理ボリューム」と言う)を例に採ると、論理ボリューム管理情報 6 0は、例えば、下記の情報、

- (\*) 対象論理ボリュームの番号である「論理ボリューム番号」、
- (\*) 対象論理ボリュームの記憶容量を表す「容量」、
- (\*) 対象論理ボリュームの状態を表す「状態」(「状態」の値として、例えば、“正常”、“閉塞”、“未使用”）、
- (\*) 「ホスト定義情報リスト」、
- (\*) 対象論理ボリュームに関連付けられている仮想ボリュームの番号である「仮想ボリューム番号」、
- (\*) 対象ボリュームに対するリード及び/又はライトを抑止するか否かを表す「リード/ライト抑止フラグ」、
- (\*) 対象ボリュームにおける、リード及び/又はライトが抑止される範囲を表す「リード/ライト抑止範囲」、
- (\*) 対象論理ボリュームが通常ボリュームかプールを構成するプールボリュームかを表す「属性」、

を含む。「ホスト定義情報リスト」は、対象論理ボリュームから見て、アクセス要求元のホスト(図4の説明において「ホスト 1」と言う)を特定するための情報(例えば、ホスト 1の名称やポート特定情報等)を含んで良い。また、「ホスト定義情報リスト」は、ホスト 1から見て、アクセス要求先の対象論理ボリュームを特定するための情報(例えば、対象記憶装置システム 2のポートの識別情報や、論理ボリュームのLUN(Logical Unit Number)等)を含んで良い。

【0046】

図5は、仮想ボリューム管理情報 6 1の構成を示す。

【0047】

仮想ボリューム 2 4は、記憶装置システム 2内のボリューム管理単位である。

【0048】

仮想ボリューム管理情報 6 1は、例えば、仮想ボリューム毎に存在する情報である。1つの仮想ボリューム(図5の説明において「対象仮想ボリューム」と言う)を例に採ると、仮想ボリューム管理情報 6 1は、例えば、下記の情報、

10

20

30

40

50

- (\*) 対象仮想ボリュームを識別するための番号である「仮想ボリューム番号」、
- (\*) 対象仮想ボリュームの記憶容量を表す「容量」、
- (\*) 対象仮想ボリュームの状態を表す「状態」(“正常”、“閉塞”、“未使用”)、
- (\*) 対象仮想ボリュームに対してリード及び/又はライトされるデータをキャッシュメモリ215に保持することが可能か否かを示す「キャッシュメモリ動作状態」、
- (\*) 対象仮想ボリュームに対応する記憶領域が、対象仮想ボリュームを有する記憶装置システムの内部か外部かを示す「記憶位置」、
- (\*) 対象仮想ボリュームがTPボリュームか否かを示す「TP属性」、
- (\*) 対象仮想ボリュームに基づくドライブグループを表す「ドライブグループ番号」、対象仮想ボリュームに対応する外部ボリュームの番号である「外部ボリューム番号」、対象仮想ボリュームに対応するマッピング管理情報68の番号を表す「マッピング管理番号」、
- (\*) 対象仮想ボリュームがドライブグループに基づく場合、そのドライブグループの記憶領域における、対象仮想ボリュームの先頭と終端のアドレスを表す、「先頭アドレス」及び「終端アドレス」、
- (\*) ゼロデータ破棄の実行状態を表す「ゼロデータ破棄実行状態」(“実行中”、“非実行”)、を含む。「ゼロデータ破棄」とは、TPボリュームに割り当てられている物理ページ内のデータがゼロデータか否かを判断し、その判断の結果が肯定的の場合に、その物理ページの割当てを解除することを言う。「ゼロデータ破棄実行状態」の有効な値(例えば“実行中”又は“非実行”)は、対象仮想ボリュームに対応する「TP属性」の値が“有り”の場合に、設定される。

## 【0049】

図6は、ドライブグループ管理情報62の構成を示す。

## 【0050】

「ドライブグループ」とは、論理ボリュームのデータを複数のドライブに分散して格納するための、複数のドライブ200の領域の集合である。ドライブグループは、例えば、RAID (Redundant Array of Independent (or Inexpensive) Disks) グループである。

## 【0051】

ドライブグループ管理情報62は、例えば、ドライブグループ毎に存在する情報である。1つのドライブグループ(図6の説明において「対象ドライブグループ」と言う)を例に採ると、ドライブグループ管理情報62は、例えば、下記の情報、

- (\*) 対象ドライブグループを識別するための番号である「ドライブグループ番号」、
  - (\*) 対象ドライブグループの記憶容量を表す「容量」、
  - (\*) 対象ドライブグループの状態を表す「状態」(“正常”、“閉塞”、“未使用”)、
  - (\*) 対象ドライブグループに基づく論理ボリュームの番号のリストである「論理ボリュームリスト」、
  - (\*) 対象ドライブグループの構成に関する情報(例えば、RAID構成情報(RAIDレベル、データドライブ数とパリティドライブ数、ストライプサイズ等))を含んだ情報であるドライブグループ属性情報、
  - (\*) 対象ドライブグループを構成するドライブの番号のリストである「ドライブ情報リスト」、
  - (\*) 対象ドライブグループを構成している各ドライブの種別を表す「記憶装置種別」(SSD/SAS/SATA等)、
- を含む。ドライブ情報リストは、各ドライブ200において対象ドライブグループに割り当てられた領域の情報であり、例えば、ドライブ番号、対象ドライブ200における領域の開始アドレス、容量等を含む。

## 【0052】

図7は、外部ボリューム管理情報63の構成を示す。

10

20

30

40

50

## 【 0 0 5 3 】

「外部ボリューム」は、仮想ボリュームの実体が、他の記憶装置システム内の仮想ボリュームであるとして定義された仮想ボリュームである。

## 【 0 0 5 4 】

外部ボリューム管理情報 6 3 は、例えば、外部ボリューム毎に存在する情報である。1つの外部ボリューム（図 7 の説明において「対象外部ボリューム」と言う）を例に採ると、外部ボリューム管理情報 6 3 は、例えば、下記の情報、

- (\*) 対象外部ボリュームを識別するための番号である「外部ボリューム番号」、
- (\*) 対象外部ボリュームの記憶容量を表す「容量」、
- (\*) 対象外部ボリュームの状態を表す「状態」（“正常”、“閉塞”、“未使用”）、
- (\*) 対象外部ボリュームに対応する仮想ボリュームの番号である「仮想ボリューム番号」、
- (\*) 対象外部ボリュームの属性を表す「外部ボリューム属性情報」、
- (\*) 対象外部ボリュームが存在する記憶装置システムへ TP 情報を提供できるか否かを表す「TP 情報提供可否」（“可能”、“不可”）、
- (\*) 対象外部ボリュームへのパスを表す情報である「外部ボリュームパス情報リスト」、を含む。「外部ボリューム属性情報」は、例えば、対象外部ボリュームを一意に特定するための情報（例えば、他の記憶装置システムの製造番号、他の記憶装置システムにおけるボリューム番号）を含んで良い。

10

## 【 0 0 5 5 】

図 8 は、外部ボリュームパス情報 6 3 1 の構成を示す。

20

## 【 0 0 5 6 】

外部ボリュームパス情報 6 3 1 は、図 7 に示した「外部ボリュームパス情報リスト」に含まれる情報であり、例えば、外部ボリュームへのパス（以下、外部ボリュームパス）毎に存在する情報である。1つの外部ボリュームパス（図 8 の説明において「対象外部ボリュームパス」と言う）を例に採ると、外部ボリュームパス情報 6 3 1 は、例えば、下記の情報、

- (\*) 対象外部ボリュームを有する記憶装置システムのポートのアドレスを表す「ポートアドレス情報」、
- (\*) 対象外部ボリュームに関連付けられている論理ボリュームの番号である「LUN」、
- (\*) 対象外部ボリュームに対応する仮想ボリュームを有する記憶装置システムのポートの番号である「自記憶装置システムポート番号」、
- (\*) 対象外部ボリュームパスの状態を表す「状態」（“正常”、“閉塞”、“未使用”）、

30

を含む。対象外部ボリュームパスは、例えば、「ポートアドレス情報」、「LUN」及び「自記憶装置システムポート番号」で構成されて良い。

## 【 0 0 5 7 】

図 9 は、記憶階層管理情報 6 4 の構成を示す。

## 【 0 0 5 8 】

記憶階層管理情報 6 4 は、この情報 6 4 を有する記憶装置システムにある記憶装置 2 0、あるいは、その記憶装置システムにストレージ仮想化技術により外部接続されている他記憶装置システム 2 内記憶装置 2 0 について、アクセス性能別に記憶装置 2 0 を分類した情報である。記憶階層管理情報 6 4 は、記憶階層毎に存在する。1つの記憶階層（図 9 の説明において「対象記憶階層」と言う）を例に採ると、記憶階層管理情報 6 4 は、例えば、下記の情報、

40

- (\*) 対象記憶階層を識別するための番号である「記憶階層番号」、
- (\*) 対象記憶階層の状態を表す「状態」（“使用”、“未使用”）、
- (\*) 対象記憶階層に属する記憶装置 2 0 が有するドライブの種別を表す「記憶装置属性」、
- (\*) 対象記憶階層に属する記憶装置 2 0 を特定するための情報である「ドライブグルー

50

プ/外部ボリュームリスト」

を含む。例えば、記憶階層 0 には S S D、記憶階層 1 には S A S ドライブ、記憶階層 2 には S A T A ドライブ、記憶階層 3 には外部ボリュームが分類される。なお、記憶階層（ティア）の種類は、アクセス性能に代えて又は加えて別の観点から定義されても良い。

【 0 0 5 9 】

図 1 0 は、プール管理情報 6 5 の構成を示す。

【 0 0 6 0 】

プール管理情報 6 5 は、プールを構成するプールボリュームの情報と対象プール全体の容量や状態を管理する情報である。プール管理情報 6 5 は、例えば、プール毎に存在する情報である。1 つのプール（図 1 0 の説明において「対象プール」と言う）を例に採ると

、プール管理情報 6 5 は、例えば、下記の情報、

- ( \* ) 対象プールを識別するための番号である「プール番号」、
- ( \* ) 対象プールの使用目的を示す「プール属性」（ T P / スナップショット等）、
- ( \* ) 対象プールを構成する物理ページのサイズである「ページサイズ」、
- ( \* ) 対象プールの状態を表す「状態」（ “ 正常 ” 、 “ 閉塞 ” 、 “ 未使用 ” ） 、
- ( \* ) 「全体情報」、
- ( \* ) 「記憶階層情報」、
- ( \* ) 対象プールを構成するプールボリュームの番号のリストである「プールボリュームリスト」、

( \* ) 対象プールの仮想ボリュームへの割当を管理するための情報（「ページ管理情報空きキューヘッダ」及び「ページ管理情報」、

を含む。「ページ管理情報空きキューヘッダ」は、例えば、未割当ての物理ページ（空きの物理ページ）のキューのヘッダへのポインタである。「ページ管理情報」は、物理ページに関する情報である。「ページ管理情報」の詳細は後述する。

【 0 0 6 1 】

「全体情報」及び「記憶階層情報」は、全体又は記憶階層（この段落の説明において「対象範囲」と言う）について、以下の情報、

- ( \* ) 対象範囲の記憶容量を表す「容量」、
- ( \* ) 対象範囲における 1 以上の割当て済みの物理ページの総記憶容量である「使用量」、
- ( \* ) 対象範囲において将来割当てが予約されている 1 以上の物理ページの総記憶容量である「予約量」、
- ( \* ) 「容量」に対する「使用量」の割合（使用率）の上限である「使用率閾値（上限）」、
- ( \* ) 「容量」に対する「使用量」の割合（使用率）の下限である「使用率閾値（下限）」、

を含む。使用率が「使用率閾値（下限）」を超えると、プールの追加を促すメッセージが保守端末 2 1 6 及び / 又は管理端末 4 へ出力される。使用率が「使用量閾値（上限）」を越えると、プール空き容量が枯渇としているという警告が、保守端末 2 1 6 及び / 又は管理端末 4 へ出力される。

【 0 0 6 2 】

図 1 1 は、ページ管理情報 6 5 1 の構成を示す。

【 0 0 6 3 】

ページ管理情報 6 5 1 は、図 1 0 に示した「ページ管理情報」であり、例えば、物理ページ毎に存在する情報である。ページ管理情報 6 5 1 とマッピング情報 6 6 を基に、プール 2 5 の仮想ボリューム 2 4 への割当が管理される。1 つの物理ページ（図 1 1 の説明において「対象物理ページ」と言う）を例に採ると、ページ管理情報 6 5 1 は、例えば、下記の情報、

- ( \* ) 対象物理ページを識別するための番号である「ページ番号」、
- ( \* ) 対象物理ページの記憶領域の実体（対象物理ページの基になっている物理記憶領域

)のアドレスである「ページ実体アドレス情報」、  
 (\* )対象ページの状態を表す「状態」(“データ有効”、“データ無効”)、  
 (\* )対象物理ページに対応するこのページ管理情報の次のページ管理情報のポインタである「次ページ管理情報ポインタ」、  
 を含む。つまり、ページ管理情報は、キューで管理されている。

【0064】

図12は、マッピング管理情報66の構成を示す。

【0065】

マッピング管理情報66は、TPボリューム毎に存在する情報である。1つのTPボリューム(図12の説明において「対象TPボリューム」と言う)を例に採ると、マッピング管理情報66は、例えば、下記の情報、

(\* )対象TPボリュームに対応するこのマッピング管理情報の状態を表す「状態」(“使用中”、“未使用”)、

(\* )対象TPボリュームの番号である「仮想ボリューム番号」、

(\* )対象TPボリュームが関連付けられているプール(対象TPボリュームへの物理ページの割当元のプール)の番号である「プール番号」、

(\* )対象TPボリュームに割り当てられる物理ページが属する記憶階層の割当優先度を表す「記憶階層割当てポリシー」、

(\* )対象TPボリュームの全体又は記憶階層別の使用率(対象TPボリュームの「容量」に対する、物理ページが割り当てられている1以上の仮想ページの総容量、の割合)を表す「記憶領域使用率」

(\* )対象TPボリュームに割当られた物理ページのページ管理情報のリストである「ページ管理情報アドレスリスト」、

(\* )対象仮想ページ毎のアクセス頻度を表す「ページアクセス頻度情報」、

を含む。「アクセス頻度」は、対象仮想ページに対する単位時間当たりのアクセス(I/O)の回数(IOPS(I/O per Second))である。

【0066】

図13は、各種管理情報60、61、62、65、66及び63の関連性を示す。

【0067】

論理ボリューム管理情報60と仮想ボリューム管理情報61が対応する。仮想ボリュームの属性に応じて、仮想ボリューム管理情報61に、ドライブグループ管理情報62、マッピング管理情報66、外部ボリューム管理情報63のいずれかが対応する。

【0068】

また、マッピング管理情報66とページ管理情報651を基に、仮想ボリューム24にプール25の物理ページが割り当てられる。プール25を構成するプールボリューム26の割当て単位である物理ページ毎に設けたページ管理情報651は、対象物理ページが仮想ボリュームに割り当てられていない状態では、ページ管理情報空きキューにより管理される。ページ管理情報空きキューは、記憶階層毎にあっても良い。また、物理ページがTPボリュームに割当てられた場合、当該物理ページに有効データが格納された後、その物理ページに対応するページ管理情報651の状態が“有効”になる。

【0069】

図14は、コピー管理情報67の構成を示す。

【0070】

コピー管理情報67は、データ移行時に実行するコピー処理の状態や進捗を管理する情報である。コピー管理情報67は、コピーペア(コピー元ボリュームとコピー先ボリュームのペア)毎に存在する情報である。1つのコピーペア(図14の説明において「対象コピーペア」と言う)を例に採ると、コピー管理情報67は、例えば、下記の情報、

(\* )対象コピーペアにおけるコピー元ボリュームの番号である「コピー元ボリューム番号」、

(\* )対象コピーペアにおけるコピー先ボリュームの番号である「コピー先ボリューム番

10

20

30

40

50



号」、

(\*) 対象コピーペアにおけるコピー処理の実行状態を表す「コピー状態」(“実行中”、“中断”、“未使用”)

(\*) コピー対象範囲においてコピーが必要な領域を特定するためのビットマップである「コピーBM」(“ON”：差分あり、“OFF”：差分なし)、

(\*) コピー終了を判定するためのコピーBMのONであるビット数を示す「残差分量」、

(\*) コピー対象がTPボリュームであるか否かを示す「コピーモード」(“通常”、“TP”）、

(\*) コピーモードが“TP”の場合、TP情報の参照形態を特定する「TP情報参照モード」(“個別”、“一括”）、

(\*) コピーモードが“TP”の場合、ゼロデータ破棄の形態を示す「ゼロデータ破棄モード」(“非実施”、“コピー中実施”、“非同期実施”）、

(\*) コピー開始後、コピー元ボリュームを先頭から終端まで未だ一通りコピーできていないことを示す「初期コピーフラグ」、  
を含む。

#### 【0071】

コピーBMを構成する各ビットは、コピー元ボリュームの各論理領域(ブロック或いは仮想ページ)に対応する。コピー元ボリュームの論理領域内のデータとコピー先ボリュームの論理領域内のデータとが異なっている場合、コピー元ボリュームの論理領域に対応したビットが“ON”(差分あり)となる。コピー元ボリュームとコピー先ボリューム間の差分は、ビットマップ以外の形式の情報で管理されても良い。

#### 【0072】

また、「残差分量」は、コピーBMの更新の際に、“OFF”から“ON”への更新に応じて加算され、“ON”から“OFF”への更新に応じて減算されてよい。

#### 【0073】

図15は、ゼロデータ破棄管理情報68の構成を示す。

#### 【0074】

ゼロデータ破棄管理情報68は、データ移行とは非同期に実行するゼロデータ破棄処理(割当て済みの物理ページであってゼロデータを記憶する物理ページの割当てを解除する処理)の実行状態を管理する情報である。ゼロデータ破棄管理情報68は、例えば、

(\*) ゼロデータ破棄の対象のTPボリュームの番号である「ゼロデータ破棄対象ボリューム」、

(\*) ゼロデータ破棄処理の状態を表す「状態」(“実行中”、“中断”、“未使用”)

(\*) ゼロデータ破棄の対象となっている最新の仮想ページのアドレスを表す「ページ進捗ポイント」、

(\*) ゼロデータ確認中に対象仮想ページにライトがあったか否かを示す「ページライト有無フラグ」、

を含む。

#### 【0075】

以下、本実施例で行われる処理の種々の流れを説明する。なお、以下の説明では、移行元記憶装置システム2aが有する要素には、その要素の参照符号に符号「a」を追加し、移行先記憶装置システム2bが有する要素には、その要素の参照符号に符号「b」を追加することがある。

#### 【0076】

図16は、ボリューム移行処理の流れの概要を示す。

#### 【0077】

図16の各ステップでは、管理端末4が、ホスト1、移行元記憶装置システム2a、又は移行先記憶装置システム2bに指示を出す。

10

20

30

40

50

## 【 0 0 7 8 】

まず、移行先記憶装置システム 2 b が、管理端末 4 からの指示に従い、移行元ボリューム 2 3 a を移行先記憶装置システム 2 b の外部ボリュームとして移行先ボリューム 2 3 b を定義し、キャッシュメモリ動作を抑止する（ステップ 1 0 0 0）。具体的には、例えば、管理端末 4 が、移行先記憶装置システム 2 b へ、移行対象ボリュームの情報として、ホスト 1 のアクセス先である移行元ボリューム 2 3 a に関する情報（例えば、移行元記憶装置システム 2 a のポート特定情報や L U N 等）を通知する。移行先記憶装置システム 2 b は、未使用の外部ボリューム管理情報 6 3 b を確保し、管理端末 4 から受領した情報を、外部ボリューム管理情報 6 3 b の「外部ボリュームパス情報リスト」（図 7 参照）へ登録する。さらに、移行先記憶装置システム 2 b は、ホスト 1 がアクセスしている移行元記憶装置システム 2 a のポートの L U N に対し、その L U N が割り当てられている移行元ボリューム 2 3 a に関する情報を要求し、その要求に対する応答（移行元記憶装置システム 2 a からの応答）が有する情報を、外部ボリューム管理情報 6 3 b の「外部ボリューム属性情報」（図 7 参照）に登録する。また、移行先記憶装置システム 2 b は、移行元記憶装置システム 2 a へ移行元ボリューム 2 3 a の容量を問い合わせ、その問合せに対する回答が有する情報（容量、外部ボリューム属性情報）を、外部ボリューム管理情報 6 3 b の「容量」及び「外部ボリューム属性情報」として登録する。移行先記憶装置システム 2 b は、未使用の論理ボリューム管理情報 6 0 b、および、未使用の仮想ボリューム管理情報 6 1 b を確保し、それらを相互に対応付け、更に、それらを外部ボリューム 2 4 b と対応付ける。また、移行先記憶装置システム 2 b は、仮想ボリューム管理情報 6 1 b の「キャッシュメモリ動作状態」を“抑止”に設定する。これにより、移行元ボリューム 2 3 a へのアクセスと移行先ボリューム 2 3 b へのアクセスが混在しても、キャッシュ動作は、一律、移行元記憶装置システム 2 a で実行される。このため、ホスト 1 が誤ったデータにアクセスすることを回避することができる。

10

20

## 【 0 0 7 9 】

次に、プール空き容量確認処理が実行される（ステップ 1 0 0 1）。プール空き容量確認処理については、後に図 1 7 及び図 1 8 を参照して説明する。

## 【 0 0 8 0 】

次に、移行先記憶装置システム 2 b の外部ボリュームとして、移行先ボリューム 2 3 b にホストパスが定義され、それにより、移行先記憶装置システム 2 b がホスト 1 からアクセス要求を受領できるようになる（ステップ 1 0 0 2）。具体的には、移行先記憶装置システム 2 b が、論理ボリューム管理情報 6 1 b の「ホスト定義情報リスト」に、移行先ボリューム 2 3 b へのパス（ホスト 1 を特定する情報、移行先ボリューム 2 3 b のポート特定情報や L U N 等）を登録する。

30

## 【 0 0 8 1 】

次に、ホスト 1 が、使用するアクセスパスを、移行元記憶装置システム 2 a の移行元ボリューム 2 3 a への第 1 のアクセスパスから、移行先記憶装置システム 2 b の移行先ボリューム 2 3 b への第 2 のアクセスパスに切替える（ステップ 1 0 0 3）。その切り替え時に、パス切り替えプログラム 1 2 は、例えば、下記の処理（ 1 ）～（ 3 ）のうちのいずれかの処理、

40

（ 1 ）実行中のアクセス要求（アプリケーションプログラム 1 1 から受領した要求）については、移行元記憶装置システム 2 a との間で処理し、新規にアプリケーションプログラム 1 1 から受領したアクセス要求については、移行先記憶装置システム 2 b へ転送する、

（ 2 ）実行中の要求については、移行元記憶装置システム 2 a との間で処理し、新規に受領した要求については、パス切替えプログラム 1 2 内で保留し、実行中の要求がすべて終了後、保留した要求を移行先記憶装置システム 2 b へ転送する、

（ 3 ）実行中の要求を一旦全てエラー終了させ、第 1 のアクセスパスから第 2 のアクセスパスへの切り替えを行い、リトライ時には、移行先記憶装置システム 2 b への第 2 のアクセスパスを通じてアクセス要求を転送する、

を行うことができる。以下、移行先ボリューム 2 3 b へのアクセスは、移行先記憶装置シ

50

システム 2 b を経由して、移行先ボリューム 2 3 b の外部ボリュームとして定義された移行元ボリューム 2 3 a に到達する。ホスト 1 が複数存在する場合は、順次、全てのホスト 1 についてアクセスパスが切り替えられる。

【 0 0 8 2 】

次に、移行先記憶装置システム 2 b が、ステップ 1 0 0 0 で抑止していた、キャッシュメモリ動作の抑止を、解除する（ステップ 1 0 0 4 ）。

【 0 0 8 3 】

次に、移行先記憶装置システム 2 b が、移行先ボリューム 2 3 b のコピー先となるコピー先ボリューム 2 3 c を、TP ボリュームとして定義する（ステップ 1 0 0 5 ）。具体的には、例えば、移行先記憶装置システム 2 b は、未使用の論理ボリューム管理情報 6 0 b 、および、未使用の仮想ボリューム管理情報 6 1 b を確保し、それらを相互に対応付け、確保した仮想ボリューム管理情報 6 1 b の「TP 属性」を“有り”に設定する。更に、未使用のマッピング管理情報を確保し、記憶領域の割当先となるプールを対応づける。更に、移行先記憶装置システム 2 b は、空きのコピー管理情報 6 7 を確保し、「コピー元ボリューム番号」として移行先ボリューム 2 3 b の番号を設定し、「コピー先ボリューム番号」としてコピー先ボリューム 2 3 c の番号を設定し、「コピー状態」として“中断”を設定する。また、「TP 情報参照モード」及び「ゼロデータ破棄モード」が設定される。

【 0 0 8 4 】

次に、コピー処理が開始される（ステップ 1 0 0 6 ）。コピー処理については、後に図 1 9 ~ 図 2 2 を参照して説明する。

【 0 0 8 5 】

最後に、移行先記憶装置システム 2 b は、コピー先ボリューム 2 3 c の定義を解除する（ステップ 1 0 0 7 ）。コピー処理の結果、コピー先ボリューム 2 3 c には、仮想ボリューム 2 4 b が対応付いており、仮想ボリューム 2 4 b に対応付いた、移行元ボリューム 2 3 a への外部ボリューム定義が、解除される。

【 0 0 8 6 】

図 1 7 は、移行先記憶装置システム 2 b で実行されるプール空き容量確認処理の流れを示す。

【 0 0 8 7 】

管理端末 4 の指示により、プロセッサパッケージ 2 1 3 のプロセッサが、プール空き容量確認プログラム 5 3 を実行する。これにより、図 1 7 に示すプール空き容量確認処理が行われる。

【 0 0 8 8 】

プール空き容量確認プログラム 5 3 が、データ移行処理モードを管理端末 4 から受領する（ステップ 1 1 0 0 ）。具体的には、例えば、プログラム 5 3 は、「TP 情報参照モード」（図 1 4 参照）として、下記の（A）又は（B）、

（A）“個別”（コピー処理移行前に移行元ボリューム 2 3 a の TP 情報の一部を参照し、必要な TP 情報を問い合わせるモード）、

（B）“一括”（コピー処理移行前に移行元ボリューム 2 3 a の TP 情報の全てを参照するモード）、

を選択する（“一括”モードでも初期コピー後は“個別”モードと同様の動きになる）。

【 0 0 8 9 】

次に、プログラム 5 3 は、移行元記憶装置システム 2 a が TP 情報を提供可能か否か判断する（ステップ 1 1 0 1 ）。具体的には、例えば、プログラム 5 3 は、「外部ボリューム属性情報」（図 7 参照）内のメーカー名や製造番号等から、移行元記憶装置システム 2 a が TP 機能を有しているか否かを判断する。移行元記憶装置システム 2 a が TP 機能を有しているか否かの特定が難しい場合、プログラム 5 3 は、有していると判断する。

【 0 0 9 0 】

TP 情報の提供が可能な場合（ステップ 1 1 0 1 の Y ）、プログラム 5 3 は、ステップ 1 1 0 2 を実行し、TP 情報の提供が不可能な場合（ステップ 1 1 0 1 の N ）、ステップ

10

20

30

40

50

1104を実行する。

【0091】

次に、プログラム53は、移行元ボリューム23aのTP情報を移行元記憶装置システム2aに要求し(ステップ1102)、移行元記憶装置システム2aから移行元ボリューム23aのTP情報を受領するまで処理を中断する(ステップ1103)。移行元記憶装置システム2aは、ステップ1102で発行された要求を受け、TP参照処理を行う。TP参照処理については、後に図30を参照して説明する。

【0092】

次に、プログラム53は、移行先記憶装置システム2bのTP情報を参照し(ステップ1104)、プール空き容量を求める。

10

【0093】

次に、プログラム53は、プール空き容量確認メイン処理を実行する(ステップ1105)。

【0094】

図18は、プール空き容量確認メイン処理の流れを示す。

【0095】

プログラム53は、移行元記憶装置システム2aはTP情報提供可能であったか否か判断する(ステップ1200)。ステップ1102で発行した要求についてエラー終了した場合、TP情報提供不可と判断する。尚、プログラム53は、TP情報提供可否判断の結果を、外部ボリューム管理情報63に格納しておく。

20

【0096】

TP情報の提供が不可能な場合(ステップ1200のN)、プログラム53は、移行先記憶装置システム2bにおいて移行元ボリューム23aの移行時に必要となるプール容量が、移行元ボリューム23aの容量であると算出する(ステップ1202)。例えば、移行元ボリューム23aの容量が100GBの場合、プール容量が100GB、記憶階層はデフォルト(例えば記憶階層2(SAS))というような算出がされる。

【0097】

TP情報の提供が可能な場合(ステップ1200のY)、プログラム53は、ステップ1100で管理端末4から受領したTP情報参照モードは“一括”であるか否か判断する(ステップ1201)。

30

【0098】

TP情報参照モードが“一括”でない場合(ステップ1201のN)、プログラム53は、移行先記憶装置システム2bにおいて移行元ボリューム23aの移行時に必要となるプール容量を、移行元ボリューム23aのマッピング管理情報66の記憶階層毎の使用量を基に算出する(ステップ1203)。例えば、移行元ボリューム23aのマッピング管理情報66から、記憶階層0(SSD):10GB、記憶階層1(SASドライブ):30GB、記憶階層2(SATAドライブ):60GBが特定された場合、移行時に必要となるプール容量として、記憶階層0(SSD)については10GB、記憶階層1(SASドライブ)については30GB、記憶階層2(SATAドライブ)については60GBがそれぞれ必要な容量であると算出される。

40

【0099】

尚、移行元記憶装置システム2aと移行先記憶装置システム2bの記憶階層が異なる場合は、移行元記憶装置システム2aの記憶階層管理情報64の「記憶装置属性」を参考として、移行先記憶装置システム2bの「記憶装置属性」に近い記憶階層への移行が想定される。

【0100】

TP情報参照モードが“一括”である場合(ステップ1201のY)、プログラム53は、移行先記憶装置システム2bにおいて移行元ボリューム23aの移行時に必要となるプール容量を、移行元ボリューム23aのマッピング管理情報66(例えば、仮想ボリューム24aに割り当てられている物理ページの数)と、プール管理情報65の「ページサ

50

イズ」とを基に、記憶階層毎に算出する（ステップ1204）。具体的には、例えば、プログラム53は、移行元記憶装置システム2aの現在割当てられている移行元ボリューム23aの全ページ内のデータを、移行先記憶装置システム2bにコピーした場合に必要なプール容量を、移行元記憶装置システム2aと移行先記憶装置システム2bの「ページサイズ」の差異を考慮し算出する。例えば、移行元記憶装置システム2aの「ページサイズ」よりも移行先記憶装置システム2bの「ページサイズ」が大きい場合、X倍（ $X = (\text{移行先記憶装置システム2bのページサイズ}) / (\text{移行元記憶装置システム2aのページサイズ})$ ）のプール容量が必要となる可能性があると考えられる。

【0101】

次に、プログラム53は、移行先記憶装置システム2bのプール管理情報65を参照し、プール空き容量として、各階層のプール空き容量と、全階層のプール空き容量総和とを算出する（ステップ1205）。

10

【0102】

次に、プログラム53は、算出した必要プール容量とプール空き容量とを表す情報（以下、表示情報）を管理端末4に表示し、その情報が管理端末4に表示される（ステップ1206）。

【0103】

具体的には、例えば、TP情報の提供が不可能な場合（ステップ1200のN）、表示情報は、移行元ボリューム23aを移行先記憶装置システム2aの特定の記憶階層に移行することが可能であるか否かを、必要プール容量（移行元ボリューム23aの容量）と各階層のプール空き容量とを比較して示す情報で良い。また、表示情報は、移行対象データ（移行元ボリューム23a内のデータ）を複数の階層に分散すれば移行できるか否かを必要プール容量と全階層のプール空き容量総和を比較して示す情報でも良い。更に、表示情報は、移行対象データを複数の階層に分散して格納する場合の格納案を例示する情報でも良い。例えば、表示情報は、100GBのデータを、移行先記憶装置システム2bの記憶階層2に80%（80GB）、記憶階層3に20%（20GB）に分散するなら格納可能を示す情報でも良い。

20

【0104】

また、例えば、TP情報の提供が可能な場合（ステップ1200のY）、表示情報は、各記憶階層の必要プール容量と各階層のプール空き容量を比較して示す情報で良い。また、表示情報は、移行対象データの格納先の階層を変更することで格納できるか否かを、全階層の必要プール容量総和と全階層のプール空き容量総和を比較して示す情報でも良い。更に、表示情報は、格納先の階層を変更する場合の具体案を例示する情報でも良い。例えば、移行元記憶装置システム2aにおいて、移行対象データが、記憶階層0に10GB、記憶階層1に30GBに存在する場合、表示情報は、移行先記憶装置システム2bでは、記憶階層1に40GBなら格納可能を示す情報でも良い。

30

【0105】

以上で、プール空き容量確認メイン処理が終了する。

【0106】

再び図17を参照する。

40

【0107】

プログラム53は、移行先記憶装置システム2bのプールの空き容量が不足しているか判断する（ステップ1106）。

【0108】

プールの空き容量が不足している場合（ステップ1106のY）、プログラム53は、プール容量の拡大を要求するメッセージを管理端末4に表示（送信）し（ステップ1107）、移行処理を中断する（ステップ1112）。具体的には、例えば、プログラム53は、ステップ1000での処理を元に戻す。

【0109】

プールの空き容量が不足していない場合（ステップ1106のN）、プログラム53は

50

、管理端末 4 から、ステップ 1 2 0 6 での表示に対しての割当ポリシーとしてプール空き容量割当指示を受領する（ステップ 1 1 0 8）。プログラム 5 3 は、受領した指示が表示情報を、マッピング管理情報 6 6 の「記憶階層割当てポリシー」に格納する。

【 0 1 1 0 】

次に、コピー終了後、プログラム 5 3 は、プールの空き容量が容量追加要求閾値以下となるか否か判断する（ステップ 1 1 0 9）。

【 0 1 1 1 】

コピー終了後、プールの空き容量が容量追加要求閾値以下とならない場合（ステップ 1 1 0 9 の N）、プログラム 5 3 は、ステップ 1 1 1 3 を行う。

【 0 1 1 2 】

コピー終了後、プールの空き容量が容量追加要求閾値以下となる場合（ステップ 1 1 0 9 の Y）、プログラム 5 3 は、コピー後にプール容量の拡大が必要となるメッセージを管理端末 4 に表示（送信）し（ステップ 1 1 1 0）、コピー処理を開始するか否かの最終判断の入力を求める（ステップ 1 1 1 1）。

【 0 1 1 3 】

コピー処理を開始しない（中断する）入力を受領した場合（ステップ 1 1 1 1 の N）、プログラム 5 3 は、ステップ 1 1 1 2 を行う。

【 0 1 1 4 】

コピー処理を開始する入力を受領した場合（ステップ 1 1 1 1 の Y）、プログラム 5 3 は、コピー先ボリュームに割当て予定のプール空き容量を予約する（ステップ 1 1 1 3）。具体的には、例えば、コピー処理中のプール枯渇を低減するため、プログラム 5 3 は、プール管理情報 6 5 の「予約量」に必要プール容量を登録し、プール空き容量算出の際、「予約量」を「使用量」として扱う。

【 0 1 1 5 】

図 1 9 は、コピー処理の流れを示す。

【 0 1 1 6 】

管理端末 4 の指示により、プロセッサパッケージ 2 1 3 のプロセッサがコピープログラム 5 4 を実行する。これにより、図 1 9 に示すコピー処理が行われる。

【 0 1 1 7 】

まず、コピープログラム 5 4 は、「コピー B M」（図 1 4 参照）を構成する全てのビットを“ ON ”（差分あり）にする初期化を行う（ステップ 1 3 0 0）。

【 0 1 1 8 】

次に、コピープログラム 5 4 は、「初期コピーフラグ」を“ ON ”に更新する（ステップ 1 3 0 1）。初期コピーフラグは、後述のステップ 1 3 0 7 の T P 情報活用コピー処理（一括）内で参照及び更新される。

【 0 1 1 9 】

次に、コピープログラム 5 4 は、「コピー状態」を“ 実行中 ”に更新する（ステップ 1 3 0 2）。

【 0 1 2 0 】

次に、コピープログラム 5 4 は、外部ボリューム管理情報 6 3 から、移行元記憶装置システム 2 a が T P 情報提供可能か判断する（ステップ 1 3 0 3）。

【 0 1 2 1 】

T P 情報の提供が不可能の場合（ステップ 1 3 0 3 の N）、コピープログラム 5 4 は、通常コピー処理を実行する（ステップ 1 3 0 4）。

【 0 1 2 2 】

T P 情報の提供が可能の場合（ステップ 1 3 0 3 の Y）、コピープログラム 5 4 は、「 T P 情報参照モード」は“ 個別 ”か否か判断する（ステップ 1 3 0 5）。

【 0 1 2 3 】

T P 情報参照モードが“ 個別 ”の場合（ステップ 1 3 0 5 の Y）、コピープログラム 5 4 は、 T P 情報活用コピー処理（個別）を実行する（ステップ 1 3 0 6）。

10

20

30

40

50

## 【 0 1 2 4 】

TP情報参照モードが“個別”でない場合(ステップ1305のN)、コピープログラム54は、TP情報活用コピー処理(一括)を実行する(ステップ1307)。

## 【 0 1 2 5 】

図20は、通常コピー処理の流れを示す。

## 【 0 1 2 6 】

まず、コピープログラム54は、コピー範囲の差分の有無を判断する(ステップ1400)。具体的には、コピープログラム54は、コピー管理情報67の「残差分量」が“0”であるか否かを判断する。

## 【 0 1 2 7 】

コピー範囲の差分が無い場合(ステップ1400のY)、コピープログラム54は、通常コピー処理を終了する。

## 【 0 1 2 8 】

コピー範囲の差分が有る場合(ステップ1400のN)、コピープログラム54は、「コピーBM」を基に、ビット“ON”の位置に対応する領域(コピー元ボリューム(移行元ボリューム23a)内の仮想ページ)におけるデータをコピー対象データに決定する(ステップ1401)。

## 【 0 1 2 9 】

次に、コピープログラム54は、コピー対象とした範囲に対応する、「コピーBM」内のビットを、“OFF”に更新する(ステップ1402)。コピー処理実行前に「コピーBM」を更新するのは、コピー処理中にコピー対象範囲にホストからのライトがあった場合、ライト範囲を再度コピー対象とする必要があるからである。

## 【 0 1 3 0 】

次に、コピープログラム54は、コピー対象データをキャッシュメモリ215aへ格納するよう移行元記憶装置システム2aにリード要求を発行し(ステップ1403)、リード要求が終了するまで待つ(ステップ1404)。

## 【 0 1 3 1 】

次に、コピープログラム54は、「ゼロデータ破棄モード」を参照し、コピー中にゼロデータ破棄を実施するか判断する(ステップ1405)。

## 【 0 1 3 2 】

コピー中にゼロデータ破棄を実施しない場合(ステップ1405のN)、コピープログラム54は、ステップ1407を行う。

## 【 0 1 3 3 】

コピー中にゼロデータ破棄を実施する場合(ステップ1405のY)、コピープログラム54は、コピー対象データにおけるゼロデータをコピー対象から除外する(ステップ1406)。ゼロデータの特定は、ハードウェア又はコンピュータプログラムによって行われて良い。

## 【 0 1 3 4 】

次に、コピープログラム54は、コピー先ボリューム23cに対応するTPボリューム(仮想ボリューム)24cにおいて、コピー対象データのコピー先の全ての仮想ページに物理ページが割当て済みか否かを判断する(ステップ1407)。

## 【 0 1 3 5 】

全ての仮想ページに物理ページが割当て済みの場合(ステップ1407のY)、コピープログラム54は、ステップ1409を行う。

## 【 0 1 3 6 】

全ての仮想ページに未割当ての仮想ページが含まれている場合(ステップ1407のN)、コピープログラム54は、全ての未割当ての仮想ページに物理ページを割り当てる(ステップ1408)。具体的には、コピープログラム54は、ステップ1108で受領したプール空き容量割当指示に従って、コピー先の仮想ページに物理ページを割り当てる。

## 【 0 1 3 7 】

10

20

30

40

50

次に、コピープログラム 5 4 は、キャッシュメモリ 2 1 5 b におけるコピー対象データを、移行先ボリューム 2 3 b を通じてコピー先ボリューム 2 3 c へコピーする（ステップ 1 4 0 9）。コピー先ボリューム 2 3 c にコピーされたデータは、非同期ライト処理で、記憶装置 2 0 b に書き込まれる。その後、ステップ 1 4 0 0 が行われる。

【 0 1 3 8 】

図 2 1 は、TP 情報活用コピー処理（個別）の流れを示す。

【 0 1 3 9 】

通常コピー処理との差異は、コピー対象データのコピー先の記憶領域の割当ての際に、コピープログラム 5 4 が、コピー対象データの TP 情報を移行元記憶装置システム 2 a に要求し、その TP 情報を利用して、コピー先ボリューム 2 3 c の TP ボリューム（仮想ボリューム）2 4 c に物理ページを割り当てる点である。

10

【 0 1 4 0 】

ステップ 1 5 0 0 ~ ステップ 1 5 0 6、および、ステップ 1 5 1 1 は、通常コピー処理のステップ 1 4 0 0 ~ ステップ 1 4 0 6、および、ステップ 1 4 0 9 と同じであるため、説明を省略する。

【 0 1 4 1 】

次に、コピープログラム 5 4 は、コピー先ボリューム 2 3 c に対応する TP ボリューム（仮想ボリューム）2 4 c において、コピー対象データのコピー先の全ての仮想ページに物理ページが割当て済みか否かを判断する（ステップ 1 5 0 7）。

【 0 1 4 2 】

全ての仮想ページに物理ページが割当て済みの場合（ステップ 1 5 0 7 の Y）、コピープログラム 5 4 は、ステップ 1 5 1 1 を行う。

20

【 0 1 4 3 】

全ての仮想ページに未割当ての仮想ページが含まれている場合（ステップ 1 5 0 7 の N）、コピープログラム 5 4 は、コピー対象データの TP 情報を移行元記憶装置システム 2 a に要求し（ステップ 1 5 0 8）、TP 情報要求の終了を待つ（ステップ 1 5 0 9）。

【 0 1 4 4 】

次に、コピープログラム 5 4 は、移行元ボリューム 2 3 a の TP 情報を利用して、コピー先ボリューム 2 3 c の TP ボリューム 2 4 c におけるコピー先の全ての未割当て仮想ページに物理ページを割当てる（ステップ 1 5 1 0）。具体的には、コピープログラム 5 4 は、ステップ 1 1 0 8 で受領したプール空き容量割当指示に従って、コピー先の仮想ページに物理ページを割当てる。

30

【 0 1 4 5 】

図 2 2 は、TP 情報活用コピー処理（一括）の流れを示す。

【 0 1 4 6 】

TP 情報活用コピー処理（個別）との差異は、コピー開始後、コピー元ボリュームを先頭から終端まで未だ一通りコピーできていない初期コピー状態の場合、コピー元ボリュームの TP 情報のマッピング管理情報 6 6 を参照して、未割当ての記憶領域（仮想ページ）からのリードを回避する点である。これにより、コピー対象のデータ量を削減することができる。

40

【 0 1 4 7 】

ステップ 1 6 0 0、ステップ 1 6 0 3 ~ ステップ 1 6 1 3 は、TP 情報活用コピー処理（個別）のステップ 1 5 0 0、ステップ 1 5 0 1 ~ 1 5 1 1 と同じであるため説明を省略する。

【 0 1 4 8 】

コピープログラム 5 4 は、「初期コピーフラグ」（図 1 4 参照）が“ON”であるか否かを判断する（ステップ 1 6 0 1）。「初期コピーフラグ」が“ON”である場合（ステップ 1 6 0 1 の Y）、コピープログラム 5 4 は、移行元ボリューム 2 3 a の TP 情報を基に、物理ページが割り当てられている仮想ページ内のデータを、コピー対象データに決定し、未割当ての仮想ページ内のデータを、コピー対象データから除外する（ステップ 1 6

50



02)。

【0149】

また、コピープログラム54は、コピー処理がボリューム終端まで一通りコピーを実行したか判断する(ステップ1614)。コピー処理がボリューム終端まで一通りコピーを実行していた場合(ステップ1614のY)、コピープログラム54は、「初期コピーフラグ」を“OFF”に設定する(ステップ1615)。

【0150】

再び図19を参照する。

【0151】

コピープログラム54は、移行先ボリューム23bに対応した「リード/ライト抑止フラグ」を“ON”に更新する(ステップ1308)。リード/ライト抑止の範囲は、ボリューム全体である。これにより、当該ボリュームへのリード/ライト処理が一時的に保留される。

10

【0152】

次に、コピープログラム54は、コピー範囲の差分の有無を再度判断する(ステップ1309)。具体的には、コピープログラム54は、コピー管理情報67の「残差分量」が“0”であるか否か判断する。

【0153】

コピー範囲の差分が有る場合(ステップ1309のN)、コピープログラム54は、ホスト1からのライトのすれ違い有りと判断し、再度コピー処理を実行するため、ステップ1303へ戻る。「ライトのすれ違い」というのは、コピー処理中に、コピー元の領域からコピーの領域にデータが移行された後に、コピー元の領域にデータが書き込まれたことである。この場合、コピー元領域からコピー先領域へのデータのコピーが再び必要である。

20

【0154】

コピー範囲の差分が無い場合(ステップ1309のY)、コピープログラム54は、移行先ボリューム23bとコピー先ボリューム23cについて、論理ボリュームと仮想ボリューム間のマッピングを変更する(ステップ1310)。具体的には、コピープログラム54は、論理ボリューム(移行先ボリューム)23b 仮想ボリューム24b、および、論理ボリューム(コピー先ボリューム)23c 仮想ボリューム24cであった関連付けを、論理ボリューム(移行先ボリューム)23b 仮想ボリューム24c、および、論理ボリューム(コピー先ボリューム)23c 仮想ボリューム24bという関連付けに変更する。このステップ1310以降に、移行先記憶装置システム2bが、移行先ボリューム23bを指定したアクセス要求を受けた場合、仮想ボリューム24c(具体的には、仮想ボリューム24cに割り当てられている、プール25b内の物理ページ)に対するアクセスを行なうことになる。

30

【0155】

次に、コピープログラム54は、コピー先ボリューム23cに対応した「リード/ライト抑止フラグ」を“OFF”に更新する(ステップ1311)。これにより、当該ボリュームへのリード/ライト処理の保留が解除される。

40

【0156】

最後、コピープログラム54は、「コピー状態」を“中断”に更新する(ステップ1312)。

【0157】

尚、移行元記憶装置システム2aの移行元ボリューム23aに関するTP情報の全部ではなく一部が移行されても良い。例えば、マッピング管理情報66a内の「ページアクセス頻度情報」が有する情報が、移行元記憶装置システム2aの「ページサイズ」と移行先記憶装置システム2bの「ページサイズ」の差異を基に移行されても良い。例えば、ページサイズが同一ならば、TP情報がそのまま移行され、移行先記憶装置システム2bの「ページサイズ」が移行元記憶装置システム2aの「ページサイズ」の整数倍ならば、移行

50

元記憶装置システム 2 a の複数ページに関する情報 A が、移行先記憶装置システム 2 b の 1 ページに関する情報 B として移行されても良い。つまり、情報 A が、移行先記憶装置システム 2 b において、1 ページに関する情報として管理されて良い。

【 0 1 5 8 】

図 2 3 は、記憶装置システム 2 a ( 2 b ) で実行されるリード処理を示す。

【 0 1 5 9 】

ホスト 1、または、他の記憶装置システム 2 からのリード要求により、プロセッサパッケージ 2 1 3 のプロセッサがリードプログラム 5 0 を実行する。それにより、図 2 3 に示すリード処理が行われる。

【 0 1 6 0 】

リードプログラム 5 0 が、リード要求を受領する ( ステップ 1 7 0 0 )。アクセス要求 ( リード要求及びライト要求 ) は、アクセス先を表す情報 ( アクセス先情報 ) を有する。アクセス先情報は、アクセス先の論理ボリューム番号 ( 例えば L U N ) と、アクセス先の範囲を表す情報 ( 例えば、アドレス ( 例えば L B A ( Logical Block Address ) ) ) と、アクセス対象のデータの長さとの組み合わせ) とを含む。リード処理では、アクセス先の論理ボリュームは、リード対象ボリュームであり、アクセス先の範囲は、リード対象範囲である。

【 0 1 6 1 】

次に、リードプログラム 5 0 が、リード対象範囲について、「リード/ライト抑止」があるか否か判断する ( ステップ 1 7 0 1 )。この判断は、リード対象ボリュームに対応した「リード/ライト抑止フラグ」の値と「リード/ライト抑止範囲」の値とを基に行うことができる。

【 0 1 6 2 】

リード/ライト抑止が無い場合 ( ステップ 1 7 0 1 の N )、リードプログラム 5 0 が、ステップ 1 7 0 3 を行う。

【 0 1 6 3 】

リード/ライト抑止が有る場合 ( ステップ 1 7 0 1 の Y )、リードプログラム 5 0 が、リード/ライト抑止が解除されるまで、リード処理を抑止 ( 保留 ) する ( ステップ 1 7 0 2 )。

【 0 1 6 4 】

次に、リードプログラム 5 0 が、リード対象ボリュームに関連付けられている仮想ボリュームに対応した「記憶位置」が“内部”であり ( つまりその仮想ボリュームが、リード要求を受けた記憶装置システム内に有り )、且つ、その仮想ボリュームに対応した「TP 属性」が“有り”か否か判断する ( ステップ 1 7 0 3 )。

【 0 1 6 5 】

ステップ 1 7 0 3 の判断の結果が否定的の場合 ( ステップ 1 7 0 3 の N )、リードプログラム 5 0 が、通常ボリュームリード処理を実行する ( ステップ 1 7 0 4 )。

【 0 1 6 6 】

ステップ 1 7 0 3 の判断の結果が肯定的の場合 ( ステップ 1 7 0 3 の Y )、リードプログラム 5 0 が、TP ボリュームリード処理を実行する ( ステップ 1 7 0 5 )。

【 0 1 6 7 】

図 2 4 は、通常ボリュームリード処理の流れを示す。

【 0 1 6 8 】

リードプログラム 5 0 が、リード対象範囲内のリード対象データが全てキャッシュメモリ 2 1 5 に存在するか判断する ( ステップ 1 8 0 0 )。

【 0 1 6 9 】

リード対象データが全てキャッシュメモリ 2 1 5 に存在する場合 ( ステップ 1 8 0 0 の Y )、リードプログラム 5 0 が、ステップ 1 8 0 2 を行う。

【 0 1 7 0 】

リード対象データにキャッシュメモリ 2 1 5 に存在しないデータが含まれている場合 (

10

20

30

40

50

ステップ1800のN)、リードプログラム50が、キャッシュメモリ215に無いデータを、リード対象ボリュームの実体に応じて、内部記憶装置又は外部ボリュームからキャッシュメモリ215へリードする(ステップ1801)。

【0171】

次に、リードプログラム50が、キャッシュメモリ215にリードしたデータを、リード要求の発行元(ホスト1、又は、リード要求を受領した記憶装置システムとは別の記憶装置システム)へ転送する(ステップ1802)。

【0172】

次に、リードプログラム50が、リード要求発行元へリード終了を報告する(ステップ1803)。

10

【0173】

次に、リードプログラム50が、リード対象ボリュームのキャッシュメモリ動作可否(リード対象ボリュームの仮想ボリュームに対応した「キャッシュ動作状態」の値)を判断する(ステップ1804)。

【0174】

キャッシュメモリ動作可能の場合(ステップ1804のY)、リードプログラム50が、リード処理を終了する。

【0175】

キャッシュメモリ動作抑止の場合(ステップ1804のN)、リードプログラム50が、キャッシュメモリにリードしたリード対象データを破棄する(ステップ1805)。

20

【0176】

図25は、TPボリュームリード処理の流れを示す。

【0177】

ステップ1901、ステップ1903～ステップ1906は、通常ボリュームリード処理のステップ1801～1805と同じであるため、説明を省略する。

【0178】

リードプログラム50が、リード対象範囲(物理ページが割当て済みの範囲)について、リード対象データが全てキャッシュメモリ215に存在するか判断する(ステップ1900)。

【0179】

また、リードプログラム50が、リード対象範囲(物理ページが未割当ての範囲)について、リード対象データをキャッシュメモリ215のゼロデータ格納領域に設定する(ステップ1902)。

30

【0180】

図26は、記憶装置システム2a(2b)で実行されるライト処理を示す。

【0181】

具体的には、ホスト1、または、他の記憶装置システム2からのライト要求により、プロセッサパッケージ213のプロセッサがライトプログラム51を実行する。これにより、図26に示すライト処理が行われる。ライト処理では、アクセス先の論理ボリュームは、ライト対象ボリュームであり、アクセス先の範囲は、ライト対象範囲である。

40

【0182】

ライトプログラム51は、ライトコマンドを受領する(ステップ2000)。

【0183】

次に、ライトプログラム51は、ライト対象範囲にリード/ライト抑止があるか判断する(ステップ2001)。この判断は、ライト対象ボリュームに対応した「リード/ライト抑止フラグ」の値と「リード/ライト抑止範囲」の値とを基に行うことができる。

【0184】

リード/ライト抑止が無い場合(ステップ2001のN)、ライトプログラム51は、ステップ2003を行う。

【0185】

50

リード/ライト抑止が有る場合（ステップ2001のY）、ライトプログラム51は、リード/ライト抑止が解除されるまで、ライト処理を抑止（保留）する（ステップ2002）。

【0186】

次に、ライトプログラム51は、ライト対象ボリュームに対応する「コピー状態」が“実行中”であるか否か判断する（ステップ2003）。

【0187】

「コピー状態」が“実行中”でない場合（ステップ2003のN）、ライトプログラム51は、ステップ2005を行う。

【0188】

「コピー状態」が“実行中”の場合（ステップ2003のY）、ライトプログラム51は、ライト対象範囲に対応する全ビット（「コピーBM」におけるビット）を“ON”に更新する（ステップ2004）。

【0189】

次に、ライトプログラム51は、ライト対象ボリュームに対応するゼロデータ破棄管理情報68の「処理状態」が“実行中”、かつ、ライト対象範囲にゼロデータ破棄処理中の仮想ページを含むか否か判断する（ステップ2005）。ライト対象範囲にゼロデータ破棄処理中の仮想ページが含まれるか否かは、当該管理情報68内の「ページ進捗ポインタ」を基に判断可能である。

【0190】

ステップ2005の判断の結果が否定的の場合（ステップ2005のN）、ライトプログラム51は、ステップ2007を行う。

【0191】

ステップ2005の判断の結果が肯定的の場合（ステップ2005のY）、ライトプログラム51は、当該管理情報68内の「ページライト有無フラグ」を“ON”に更新する（ステップ2006）。

【0192】

次に、ライトプログラム51は、ライト対象ボリュームに関連付けられている仮想ボリュームに対応した「記憶位置」が“内部”であり（つまりその仮想ボリュームが、ライト要求を受けた記憶装置システム内に有り）、且つ、その仮想ボリュームに対応した「TP属性」が“有り”か否か判断する（ステップ2007）。

【0193】

ステップ2007の判断の結果が否定的の場合（ステップ2007のN）、ライトプログラム51は、通常ボリュームライト処理を実行する（ステップ2008）。

【0194】

ステップ2007の判断の結果が肯定的の場合（ステップ2007のY）、ライトプログラム51は、TPボリュームライト処理を実行する（ステップ2009）。

【0195】

図27は、通常ボリュームライト処理の流れを示す。

【0196】

ライトプログラム51は、ライト要求に従うライト対象データをキャッシュメモリ215に格納する（ステップ2100）。

【0197】

次に、ライトプログラム51は、ライト対象ボリュームのキャッシュメモリ動作可否（ライト対象ボリュームの仮想ボリュームに対応した「キャッシュ動作状態」の値）を判断する（ステップ2101）。

【0198】

キャッシュメモリ動作可能の場合（ステップ2101のY）、ライトプログラム51は、ステップ2104を行う。

【0199】

10

20

30

40

50

キャッシュメモリ動作抑止の場合（ステップ2101のN）、ライトプログラム51は、キャッシュメモリ215に格納したライト対象データを、ライト対象ボリュームの実体に応じて、内部記憶装置又は外部ボリュームに書き込む（ステップ2102）。その際、例えば、ライト対象ボリュームがRAID5又はRAID6のドライブグループに基づく内部の仮想ボリュームであれば、必要に応じて、ライト対象の旧データや旧パリティがキャッシュメモリ215にリードされ、それらを基に新パリティが生成され、ライト対象データと共に新パリティがドライブグループに書き込まれて良い。また、例えば、ライト対象ボリュームがRAID1のドライブグループに基づく内部の仮想ボリュームであれば、ライト対象データが二重化されてドライブグループに書き込まれて良い。

【0200】

10

次に、ライトプログラム51は、キャッシュメモリ215に格納したライト対象データを破棄する（ステップ2103）。

【0201】

次に、ライトプログラム51は、ライト要求の発行元（ホスト1、又は、ライト要求を受領した記憶装置システムとは別の記憶装置システム）へライト終了を報告する（ステップ2104）。

【0202】

図28は、TPボリュームライト処理の流れを示す。

【0203】

ステップ2200、ステップ2201、ステップ2204～ステップ2206は、通常ボリュームライト処理のステップ2100～2104と同じであるため、説明を省略する。

20

【0204】

ライトプログラム51は、ライト対象範囲に属する全ての仮想ページに物理ページが割当て済みか判断する（ステップ2202）。

【0205】

ライト対象範囲に少なくとも1つの未割当ての仮想ページが属している場合（ステップ2202のN）、ライトプログラム51は、全ての未割当ての仮想ページに物理ページを割当て（ステップ2203）。

【0206】

30

図29は、記憶装置システム2a（2b）で実行される非同期ライト処理の流れを示す。

【0207】

具体的には、例えば、記憶装置システム2a（2b）が周期的（又は非周期的）にプロセスパッケージ213の非同期ライトプログラム56を実行する。これにより、図29に示す非同期ライト処理が行われる。

【0208】

非同期ライトプログラム56は、非同期ライト対象ボリュームのキャッシュメモリ動作可否を判断する（ステップ2300）。

【0209】

40

キャッシュメモリ動作抑止の場合（ステップ2300のN）、非同期ライトプログラム56は、非同期ライト処理を終了する。その理由は、全てのダーティデータがホスト1や他の記憶装置システムからのライト要求に同期して、記憶装置や外部ボリュームへライトされているためである。

【0210】

次に、非同期ライトプログラム56は、非同期ライト対象ボリュームのダーティデータ（キャッシュメモリ215における、ドライブに未だ書き込まれていないデータ）から、ライト対象データを決定する（ステップ2301）。

【0211】

次に、非同期ライトプログラム56は、非同期ライト対象ボリュームに関連付けられて

50

いる仮想ボリュームに対応した「記憶位置」が“内部”であり、且つ、その仮想ボリュームに対応した「TP属性」が“有り”か否か判断する（ステップ2302）。

【0212】

ステップ2302の判断の結果が否定的の場合（ステップ2302のN）、非同期ライトプログラム56は、ステップ2305を行う。

【0213】

ステップ2302の判断の結果が肯定的の場合（ステップ2302のY）、非同期ライトプログラム56は、ライト対象範囲に属する全ての仮想ページに物理ページが割当て済みか判断する（ステップ2303）。

【0214】

ライト対象範囲に属する全ての仮想ページに物理ページが割当て済みの場合（ステップ2303のY）、非同期ライトプログラム56は、ステップ2305を行う。

【0215】

ライト対象範囲に少なくとも1つの未割当ての仮想ページがある場合（ステップ2303のN）、非同期ライトプログラム56は、それら未割当ての仮想ページに物理ページを割り当てる（ステップ2304）。

【0216】

次に、非同期ライトプログラム56は、キャッシュメモリ215に格納したライト対象データを、ライト対象ボリュームの実体に応じて、内部記憶装置又は外部ボリュームに書き込む（ステップ2305）。

【0217】

図30は、移行元記憶装置システム2aで実行されるTP情報参照処理を示す。

【0218】

具体的には、例えば、他の記憶装置システム2からの要求により、プロセッサパッケージ213aのプロセッサがTP情報参照プログラム52を実行する。これにより、図30に示すTP情報参照処理が行われる。

【0219】

TP情報参照プログラム52は、TP情報リード要求を受領する（ステップ2400）。

【0220】

次に、TP情報参照プログラム52は、TP情報リード要求の対象ボリュームに対応する仮想ボリューム管理情報61内の「TP属性」が“有る”か否か判断する（ステップ2401）。

【0221】

「TP属性」が“無い”場合（ステップ2401のN）、プログラム52は、ステップ2404を行う。

【0222】

「TP属性」が“有る”場合（ステップ2401のY）、プログラム52は、TP情報リード要求の対象ボリュームに対応するTP情報69を特定し（ステップ2402）、そのTP情報69が有する情報を、TP情報リード要求の発行元へ転送する（ステップ2403）。

【0223】

次に、プログラム52は、TP情報リード要求の発行元へ要求終了を報告する（ステップ2404）。ステップ2401のNの場合、このステップ2404では、エラーが報告されて良い。

【0224】

図31は、記憶装置システム2a（2b）で実行されるゼロデータ破棄処理の流れを示す。

【0225】

具体的には、記憶装置システム2a（2b）が周期的（又非周期的）にプロセッサパッ

10

20

30

40

50

ページ 213 のプロセッサでゼロデータ破棄プログラム 55 を実行する。これにより、図 31 に示すゼロデータ破棄処理が行われる。

【0226】

まず、ゼロデータ破棄プログラム 55 は、ゼロデータ破棄処理の対象の論理ボリューム（以下、図 31 の説明において「対象ボリューム」と言う）に対応する「ページ進捗ポインタ」（図 15 参照）を初期化する（ステップ 2500）。

【0227】

次に、プログラム 55 は、対象ボリュームの終端まで参照したか判断する（ステップ 2501）。

【0228】

対象ボリュームの終端まで参照した場合（ステップ 2501 の Y）、プログラム 55 は、ゼロデータ破棄処理を終了する。

【0229】

対象ボリュームの終端まで参照していない場合（ステップ 2501 の N）、プログラム 55 は、ゼロデータ破棄対象ページについて「ページライト有無フラグ」を初期化（“OFF”）する（ステップ 2502）。

【0230】

次に、プログラム 55 は、ゼロデータ破棄対象ページのリード対象データをページ進捗ポインタから決定する（ステップ 2503）。

【0231】

次に、プログラム 55 は、リード対象データを記憶装置 20 からキャッシュメモリ 215 へリードする（ステップ 2504）。

【0232】

次に、プログラム 55 は、リード対象データがゼロデータであるか確認する（ステップ 2505）。

【0233】

次に、プログラム 55 は、リード対象データが全てゼロデータであるか判断する（ステップ 2506）。

【0234】

リード対象データにゼロデータでない部分がある場合（ステップ 2506 の N）、当該ページ内のデータの破棄は不可であるため、プログラム 55 は、「ページ進捗ポインタ」を次ページに更新し（ステップ 2507）、ステップ 2501 を行う。

【0235】

リード対象データが全てゼロデータである場合（ステップ 2506 の Y）、プログラム 55 は、ゼロデータ確認中に当該ページにライト無しであるかを「ページライト有無フラグ」から判断する（ステップ 2508）。

【0236】

ゼロデータ確認中に当該ページにライト有りの場合（ステップ 2508 の N）、プログラム 55 は、当該ページ内のデータの破棄は不可であるため、「ページ進捗ポインタ」を次ページに更新し（ステップ 2507）、ステップ 2501 を行う。

【0237】

ゼロデータ確認中に当該ページにライト無しの場合（ステップ 2508 の Y）、プログラム 55 は、リード対象範囲がページ境界に到達したか判断する（ステップ 2509）。

【0238】

リード対象範囲がページ境界に到達していない場合（ステップ 2509 の N）、プログラム 55 は、「ページ進捗ポインタ」をリード単位分更新し（ステップ 2510）、ステップ 2503 を行う。

【0239】

リード対象範囲がページ境界に到達している場合（ステップ 2509 の Y）、プログラム 55 は、当該ページのリード/ライトを「リード/ライト抑止フラグ」と「リード/ラ

10

20

30

40

50

イト抑止範囲」を更新することで抑止し（ステップ2511）、ページ内のデータを破棄し（ステップ2512）、当該ページのリード/ライト抑止を解除する（ステップ2513）。その後、ステップ2507が行われる。

【実施例2】

【0240】

以下、本発明の実施例2を説明する。その際、実施例1との相違点を主に説明し、実施例1との共通点については説明を省略或いは簡略する。

【0241】

図32に、本発明の実施例2に係る計算機システムの構成を示す。

【0242】

実施例1と異なる点は、移行元記憶装置システム2aと移行先記憶装置システム2bが、パス切替え部31を持つネットワーク装置3を介してホスト1と接続されている点である。ホスト1のパス切替え時には、管理端末4からの指示でネットワーク装置3がホスト透過に記憶装置システム2へのアクセスパスを切替える。

【0243】

以上、本発明の幾つかの実施例を説明したが、本発明は、これらの実施例に限定されるものでなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【0244】

例えば、本発明において、移行対象ボリュームのボリューム種別は問わない。例えば、移行対象ボリュームがメインフレームボリュームであってもかまわない。メインフレームボリュームにおいては、アクセス性能向上のため、フォーマットに関する制御情報や、特定種別のレコードのレイアウトを変更するケース等が存在するが、データ移行時に移行元記憶装置システム2aから別途情報を受領するか、あるいは、移行先記憶装置システム2bでデータ移行しつつ、新規に制御情報の作成やレコードのレイアウト変更が実施されても良い。

【0245】

また、上述した実施例1及び2では、移行元ボリューム23aに関連付けられている仮想ボリューム24aが、実質的に移行元ボリュームで良く、移行先ボリューム23bに関連付けられている仮想ボリューム24bが、実質的に移行先ボリュームで良く、コピー先ボリューム23cに関連付けられている仮想ボリューム24cが、実質的にコピー先ボリュームで良い。ホスト1に認識されるボリュームは、例えば、LUNで良い。

【0246】

また、図16のステップ1006のコピー処理では、移行先記憶装置システム2bは、TP情報69a（例えばマッピング管理情報66a）から、コピー元に対応する割当て済みの仮想ページを特定し、特定した仮想ページに対応する、TPボリューム24cの仮想ページに、プール25bから物理ページを割り当てて良い。

【符号の説明】

【0247】

1...ホスト 2a...移行元記憶装置システム 2b...移行先記憶装置システム

10

20

30



【図1】

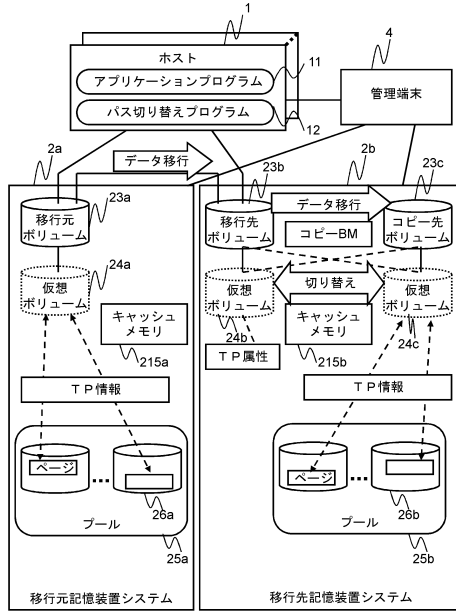


Fig.1

【図2】

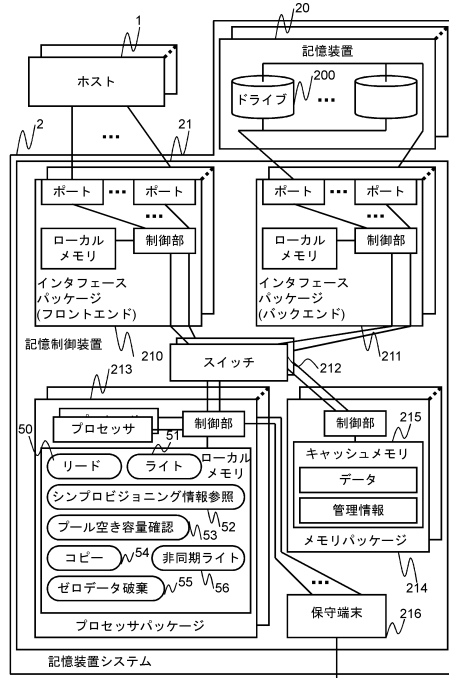


Fig.2

【図3】

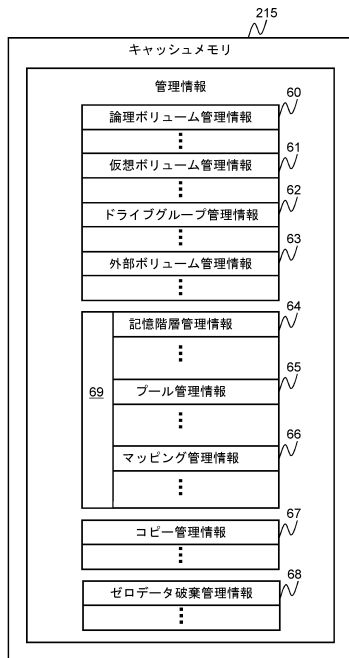


Fig.3

【図4】

論理ボリューム管理情報	60
論理ボリューム番号	
容量	
状態: 正常/閉塞/未使用	
ホスト定義情報リスト	
仮想ボリューム番号	
リード/ライト抑止フラグ: ON/OFF	
リード・ライト抑止範囲	
属性: 通常/プール	

Fig.4

【図5】

仮想ボリューム管理情報	61
仮想ボリューム番号	
容量	
状態: 正常/閉塞/未使用	
キャッシュ動作状態: 可/抑止	
記憶位置: 内部/外部	
TP属性: 有り/無し	
ドライブグループ/外部ボリューム番号 マッピング管理番号	
先頭アドレス	
終端アドレス	
ゼロデータ破壊実行状態: 実行中/非実行	

Fig.5

【図6】

ドライブグループ管理情報 62	
ドライブグループ番号	
容量	
状態: 正常/閉塞/未使用	
論理ボリュームリスト	
ドライブグループ属性情報	
ドライブ情報リスト	
記憶装置種別: SSD/SAS/SATA/...	

Fig.6

【図7】

外部ボリューム管理情報 63	
外部ボリューム番号	
容量	
状態: 正常/閉塞/未使用	
仮想ボリューム番号	
外部ボリューム属性情報	
TP 情報提供可否: 可能/不可	
外部ボリュームバス情報リスト	

Fig.7

【図8】

外部ボリュームバス情報 631	
ポートアドレス情報	
LUN	
自記憶装置システムポート番号	
状態: 正常/閉塞/未使用	

Fig.8

【図11】

ページ管理情報 651	
ページ番号	
ページ実体アドレス情報	
ページ状態: データ有効/無効	
次ページ管理情報ポインタ	

Fig.11

【図12】

マッピング管理情報 66	
状態: 使用中/未使用	
仮想ボリューム番号	
プールの番号	
記憶階層割当てポリシー	
記憶領域 使用率	全体
	記憶階層1
	記憶階層(n-1)
ページ管理情報アドレスリスト	
ページアクセス頻度情報	

Fig.12

【図9】

記憶階層管理情報 64	
記憶階層番号	
状態: 使用/未使用	
記憶装置属性: SSD/SAS/SATA/外部...	
ドライブグループ/外部ボリュームリスト	

Fig.9

【図10】

プール管理情報 65	
プール番号	
属性: シンプロビジョニング/スナップショット	
ページサイズ	
状態: 正常/閉塞/未使用	
全体	容量
	使用量
	予約量
	使用率閾値(上限)
	使用率閾値(下限)
記憶階層1	
...	
記憶階層(n-1)	
プールボリュームリスト	
ページ管理情報空きキューヘッダ	
ページ管理情報	

Fig.10

【図13】

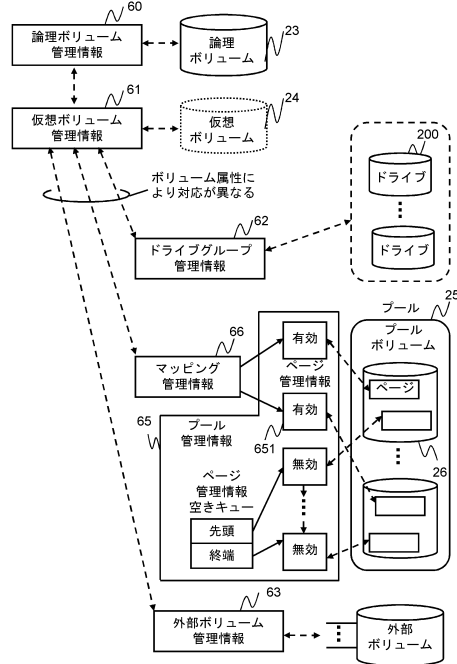


Fig.13

【図14】

コピー管理情報
コピー元ボリューム番号
コピー先ボリューム番号
コピー状態：実行中/中断/未使用
コピーBM：ON(差分あり)/OFF(差分なし)
残差分量
コピーモード：通常/シンプロビジョニング
TP情報参照モード：個別/一括
ゼロデータ破壊モード：非実施/コピー中実施/非同期実施
初期コピーフラグ

Fig.14

【図15】

ゼロデータ破壊管理情報
ゼロデータ破壊対象ボリューム
処理状態：実行中/中断/未使用
ページ進捗ポインタ
ページライト有無フラグ

Fig.15

【図16】

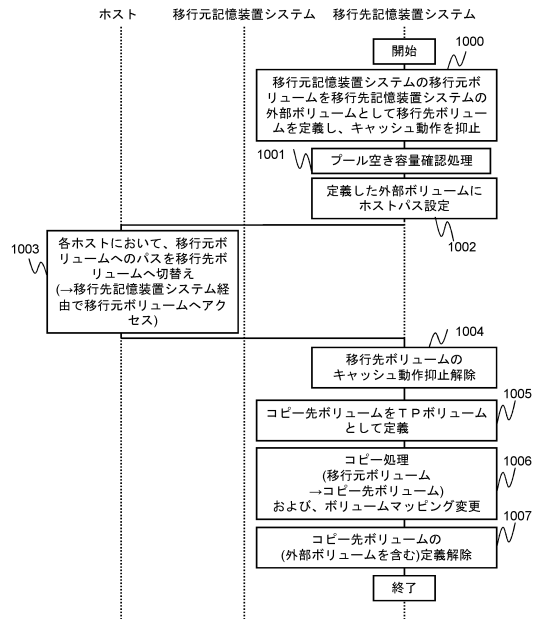


Fig.16

【図17】

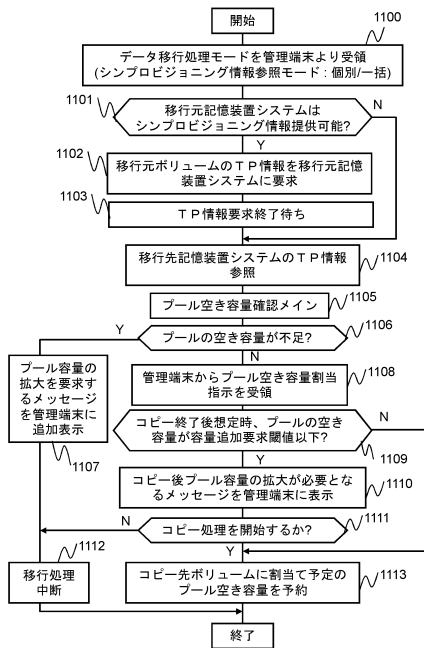


Fig.17

【図18】

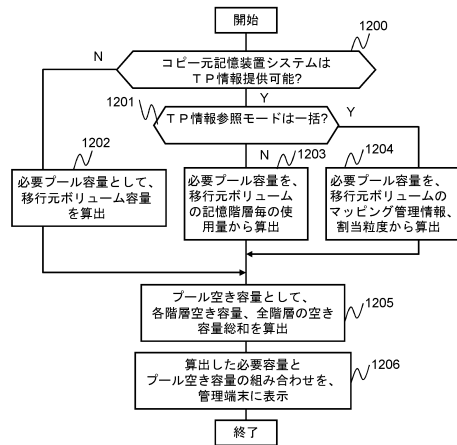


Fig.18

【図19】

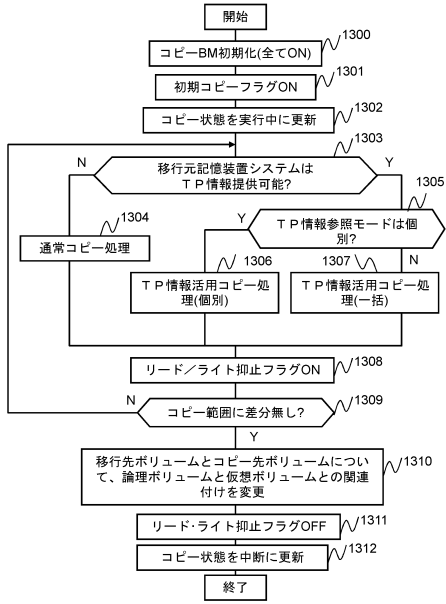


Fig.19

【図20】

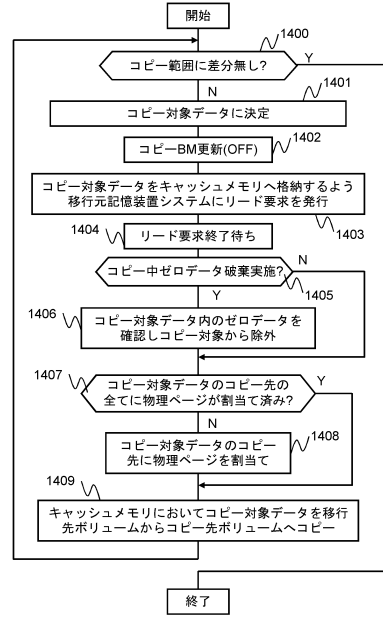


Fig.20

【図21】

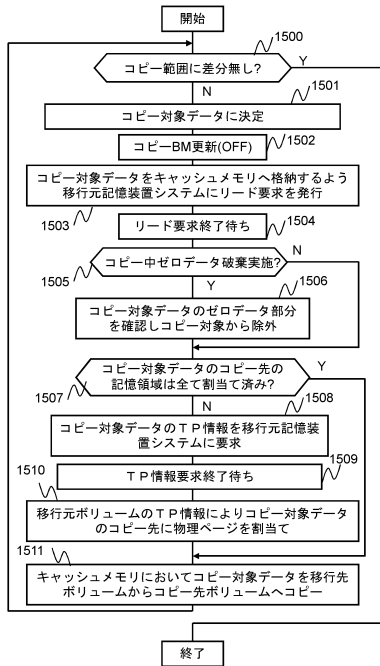


Fig.21

【図22】

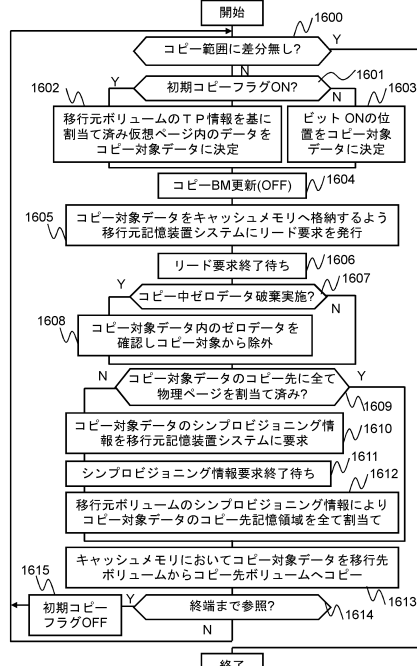


Fig.22

【図 23】

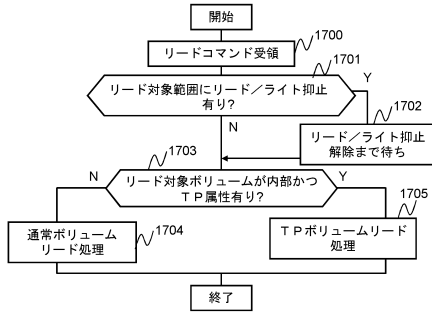


Fig.23

【図 24】

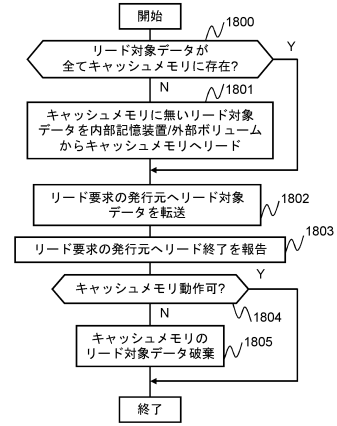


Fig.24

【図 25】

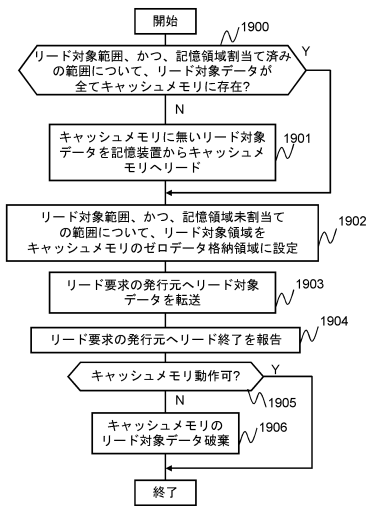


Fig.25

【図 26】

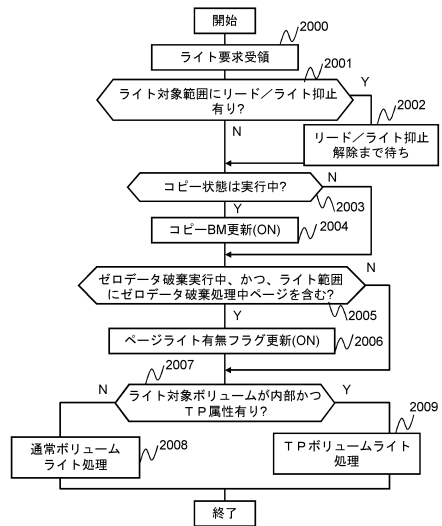


Fig.26

【図 27】

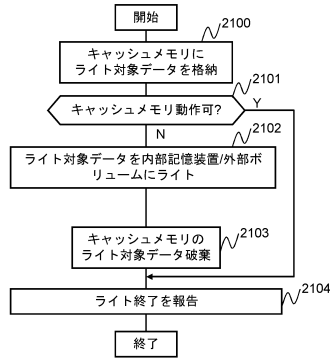


Fig.27

【図 28】

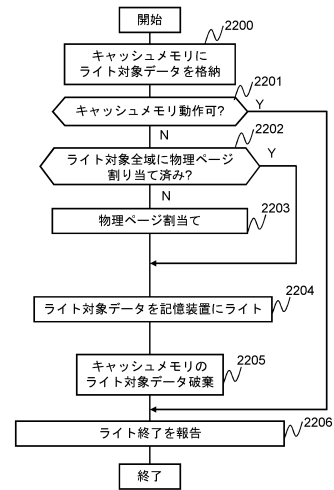


Fig.28

【図 29】

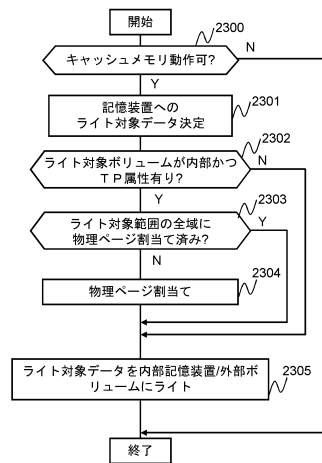


Fig.29

【図 30】

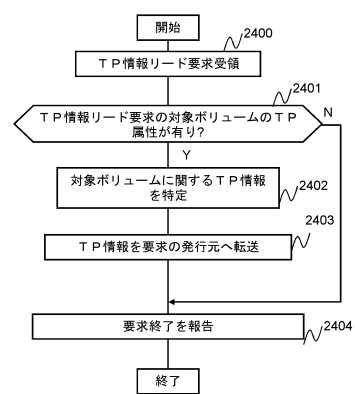


Fig.30

【図 3 1】

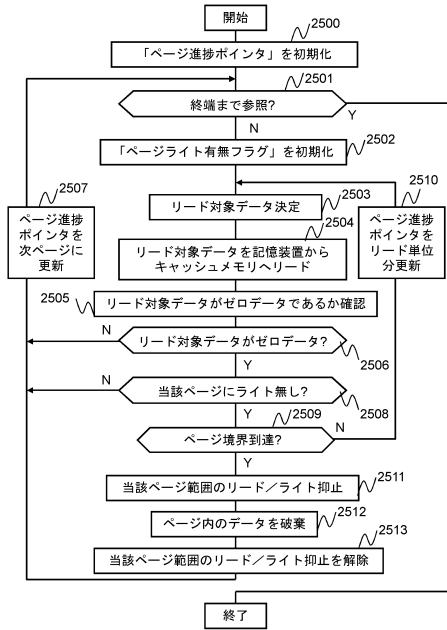


Fig.31

【図 3 2】

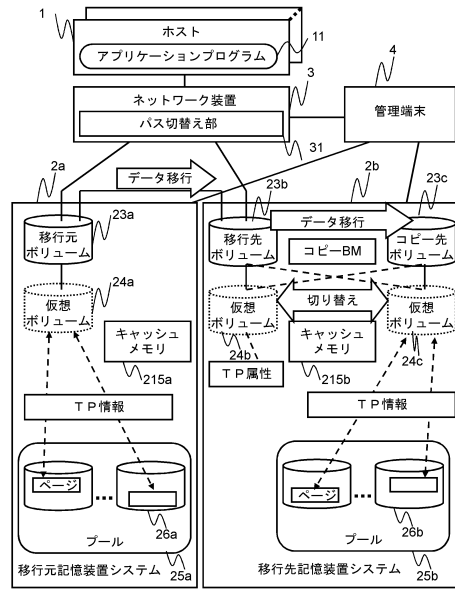


Fig.32

---

フロントページの続き

(51)Int.Cl. F I  
G 0 6 F 3/06 3 0 1 X

(72)発明者 山本 政行  
神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所 横浜研究所内

(72)発明者 山本 彰  
東京都千代田区丸の内一丁目6番1号 株式会社日立製作所 研究開発本部内

審査官 池田 聡史

(56)参考文献 特開2007-310495(JP,A)  
特開2008-282382(JP,A)  
特開2009-282922(JP,A)  
特開2005-011316(JP,A)

(58)調査した分野(Int.Cl., DB名)  
G 0 6 F 1 2 / 0 0  
G 0 6 F 3 / 0 6