

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5090941号  
(P5090941)

(45) 発行日 平成24年12月5日(2012.12.5)

(24) 登録日 平成24年9月21日(2012.9.21)

(51) Int.Cl.		F I
<b>G06F 12/00</b>	<b>(2006.01)</b>	G06F 12/00 501A
<b>G06F 3/06</b>	<b>(2006.01)</b>	G06F 12/00 514E
<b>G06F 13/10</b>	<b>(2006.01)</b>	G06F 12/00 545A
		G06F 3/06 301K
		G06F 13/10 340A

請求項の数 20 (全 31 頁)

(21) 出願番号 特願2008-17014 (P2008-17014)  
 (22) 出願日 平成20年1月29日(2008.1.29)  
 (65) 公開番号 特開2009-181148 (P2009-181148A)  
 (43) 公開日 平成21年8月13日(2009.8.13)  
 審査請求日 平成22年7月8日(2010.7.8)

(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 100114236  
 弁理士 藤井 正弘  
 (74) 代理人 100075513  
 弁理士 後藤 政喜  
 (74) 代理人 100098660  
 弁理士 戸田 裕二  
 (72) 発明者 大平 良徳  
 神奈川県川崎市麻生区王禅寺1099番地  
 株式会社日立製作所システム開発研究所  
 内

最終頁に続く

(54) 【発明の名称】 ストレージサブシステム及びストレージシステム

(57) 【特許請求の範囲】

【請求項1】

プロセッサと、  
 メモリと、  
 仮想ボリュームと、

複数の記憶媒体から構成され、前記仮想ボリュームに割り当てられるページ用物理領域を有するページ用ボリュームとを有し、

前記プロセッサは、前記仮想ボリュームのアドレス空間を複数のページに分割し、当該複数のページ各々を少なくとも第一の状態と第二の状態とを含む複数の状態のうちいずれかの状態に分類して管理し、

前記プロセッサは、前記第二の状態に分類されるページをさらに複数のセグメントに分割して管理しており、

前記第一の状態とは、ページにライトデータを書き込むためのライト要求を受けた結果、当該ページに前記ページ用ボリュームからページ用物理領域が割り当てられ、ライトデータが当該ページ用物理領域に格納された状態であり、

前記第二の状態とは、ページ内のセグメントに予め定められた特定パターンデータと一致するライトデータを書き込むためのライト要求を受けた結果、前記メモリ内に当該特定パターンデータと当該セグメントとが対応付けられて管理されている状態であることを特徴とするストレージサブシステム。

【請求項2】

前記複数の状態には、更に第三の状態が含まれ、

前記第三の状態とは、ページに前記ページ用ボリュームからページ用物理領域が割り当てられておらず、かつ当該ページがセグメントに分割されて管理されていない状態であることを特徴とする請求項1記載のストレージサブシステム。

【請求項3】

前記プロセッサは、受信したライト要求に含まれるアドレスで指定されるページが前記第一の状態である場合には、当該ページに割り当てられているページ用物理領域にライトデータを格納することを特徴とする請求項1記載のストレージサブシステム。

【請求項4】

前記第一の状態にある前記ページ内のデータが全て前記特定パターンデータと一致する場合には、前記プロセッサは、前記ページを前記セグメント単位に分割し、各セグメントについて当該セグメントと前記特定パターンデータとを対応付けて前記メモリ内で管理し、

10

前記プロセッサは、前記ページに割り当てられていたページ用物理領域を開放し、当該ページを前記第二の状態のページとして管理することを特徴とする請求項3記載のストレージサブシステム。

【請求項5】

前記プロセッサは、受信したライト要求に含まれるアドレスで指定されるページが前記第二の状態又は前記第三の状態である場合には、ライトデータが前記特定パターンデータと一致するか否かを判定することを特徴とする請求項2記載のストレージサブシステム。

20

【請求項6】

前記プロセッサは、前記ライト要求に含まれる前記アドレスで指定されるページが前記第二の状態であり、前記ライトデータが前記特定パターンデータと一致する場合、前記特定パターンデータと前記アドレスで指定されるセグメントとを対応付けて前記メモリ内で管理し、前記ライトデータを破棄することを特徴とする請求項5記載のストレージサブシステム。

【請求項7】

前記プロセッサは、前記ライト要求に含まれる前記アドレスで指定されるページが前記第二の状態であり、前記ライトデータが前記特定パターンデータと一致しない場合、前記アドレスで指定されるページに、前記ページ用ボリュームからページ用物理領域を割り当て、前記メモリ内で管理されている前記特定パターンデータと当該ページ内のセグメントとの対応関係に基づいて、前記ページに割り当てられた前記ページ用物理領域内の当該セグメントに対応する物理領域に当該特定パターンデータを格納し、当該ライトデータを当該アドレスを用いて特定される前記ページ用物理領域内の物理領域に格納し、

30

前記プロセッサは、前記ページを前記第一の状態のページとして管理することを特徴とする請求項5記載のストレージサブシステム。

【請求項8】

前記プロセッサは、前記ライト要求に含まれるアドレスで指定されるページが前記第三の状態であり、前記ライトデータが前記特定パターンデータと一致する場合、

前記ページを前記セグメント単位に分割して、前記メモリ内に前記アドレスで指定されるセグメントと前記特定パターンデータとを対応付けて記録し、前記ライトデータを破棄し、

40

前記プロセッサは、前記ページを前記第二の状態のページとして管理することを特徴とする請求項5記載のストレージサブシステム。

【請求項9】

前記プロセッサは、前記ライト要求に含まれるアドレスで指定されるページが前記第三の状態であり、前記ライトデータが前記特定パターンデータと一致しない場合、前記ページに、前記ページ用ボリュームからページ用物理領域を割り当て、当該ページ用物理領域に当該ライトデータを格納し、

前記プロセッサは、前記ページを前記第一の状態のページとして管理することを特徴と

50

する請求項 5 記載のストレージサブシステム。

【請求項 10】

前記プロセッサは、前記第一の状態にある前記ページに格納されているデータの内の前記特定パターンデータと一致するデータのデータ数をカウントしており、

前記プロセッサは、前記ライト要求に従って受信した前記ライトデータが前記特定パターンデータと一致する場合には、当該ライトデータを前記ページに割り当てられている前記ページ用物理領域に格納し、前記データ数を 1 増加させ、

前記プロセッサは、前記データ数が予め定められた第一の閾値を超えた場合に、当該ページ内のデータが全て前記特定パターンデータと一致するか否かを判定することを特徴とする請求項 4 記載のストレージサブシステム。

10

【請求項 11】

前記ストレージサブシステムは更に、複数の記憶媒体から構成され、前記仮想ボリュームに割り当てられるセグメント用物理領域を有するセグメント用ボリュームを有し、

前記プロセッサは、前記第一の状態のページ内に格納されている前記特定パターンデータと一致するデータの数が予め定められた第一の閾値を超え、かつ前記ページ内に格納されている前記特定パターンデータとは一致しないデータの数が予め定められた第二の閾値以下である場合、当該ページを前記セグメント単位に分割し、

前記ページ内の各セグメントについて、当該セグメントに前記特定パターンデータと一致するデータが格納されている場合には、前記プロセッサは当該セグメントと当該特定パターンデータを対応付けて前記メモリ内で管理し、当該セグメントに前記特定パターンデータとは一致しないデータが格納されている場合には、前記プロセッサは当該セグメントに対して前記セグメント用ボリュームからセグメント用物理領域を割り当てて当該セグメント用物理領域に当該データを格納し、

20

前記プロセッサは、前記ページに割り当てられていた前記ページ用物理領域を開放し、当該ページを前記第二の状態のページとして管理することを特徴とする請求項 1 記載のストレージサブシステム。

【請求項 12】

前記プロセッサは、前記第二の状態のページに対するライト要求を受信した場合に、当該ライト要求に従って受信したライトデータが前記特定パターンデータと一致しないデータであり、かつ当該ページに格納されている前記特定パターンデータと一致しないデータの数が前記第二の閾値以下である場合は、当該ページに新たに前記セグメント用ボリュームからセグメント用物理領域を割り当て、当該セグメント用物理領域に当該ライトデータを格納することを特徴とする請求項 11 記載のストレージサブシステム。

30

【請求項 13】

前記ストレージサブシステムは更に、複数の記憶媒体から構成される通常ボリュームを有しており、

前記通常ボリュームに格納されているデータを前記仮想ボリュームに移行する場合、

前記プロセッサは、前記通常ボリュームから前記セグメントと同サイズ分のデータを順にリードし、当該リードしたデータ前記仮想ボリュームにライトすることを特徴とする請求項 1 記載のストレージサブシステム。

40

【請求項 14】

前記ストレージサブシステムは、複数の記憶媒体からなる物理ボリュームを有する他のストレージサブシステムに接続され、

前記ストレージサブシステムは、前記他のストレージサブシステムの前記物理ボリュームの先頭セグメントから順にデータをリードし、当該データを前記仮想ボリュームにライトすることを特徴とする請求項 1 記載のストレージサブシステム。

【請求項 15】

前記ストレージサブシステムは、前記物理ボリュームからリードしたデータのうち同じデータがいくつあるかをカウントし、

前記ストレージシステムは、当該カウント数の結果が予め設定されている閾値を超えた

50

場合、当該データを特定パターンデータとして前記メモリ内に登録することを特徴とする請求項 14 のストレージサブシステム。

【請求項 16】

送信元ストレージサブシステムと、  
前記送信元ストレージサブシステムと通信可能に接続された送信先ストレージサブシステムとを有するストレージシステムであって、

前記送信元ストレージサブシステムは、コピー元ボリュームを有し、

前記送信先ストレージサブシステムは、仮想ボリュームと、複数の記録媒体から構成され当該仮想ボリュームに割り当てられるページ用物理領域を有するコピー先物理ボリュームとを有し、

前記送信先ストレージサブシステムは、前記仮想ボリュームのアドレス空間を複数のページに分割し、当該複数のページ各々を、少なくとも第一の状態及び第二の状態を含む複数の状態のいずれかに分類して管理し、

前記送信先ストレージサブシステムは、前記第二の状態にあるページを更に複数のセグメントに分割して管理しており、

前記第一の状態とは、ページにデータを書き込むための要求を受けた結果、当該ページに前記コピー先物理ボリュームからページ用物理領域が割り当てられ、当該データが当該ページ用物理領域に割り当てられた状態であり、

前記第二の状態とは、予め定められた特定パターンデータと一致するデータをページ内のセグメントに書き込むための要求を受けた結果、当該セグメントと当該特定パターンデータとが対応付けられて管理されている状態であり、

前記送信元ストレージサブシステムは、前記コピー元ボリュームからリモートコピーの対象であるデータを読み出し、前記読み出したデータが前記特定パターンデータと一致する場合には、当該特定パターンデータを識別する ID を前記送信先ストレージサブシステムに送信し、

前記送信先ストレージサブシステムは、前記送信元ストレージサブシステムから前記 ID を受信した場合に、当該 ID によって識別される前記特定パターンデータの格納先となる前記仮想ボリューム内のページが前記第一の状態又は前記第二の状態のいずれの状態にあるかに応じて、異なる処理を実行することを特徴とするストレージシステム。

【請求項 17】

前記 ID によって識別される前記特定パターンデータの格納先となる前記仮想ボリューム内のページが前記第一の状態である場合、前記送信先ストレージサブシステムは当該ページに割り当てられているページ用物理領域に当該 ID によって識別される前記特定パターンデータを格納することを特徴とする請求項 16 記載のストレージシステム。

【請求項 18】

前記 ID によって識別される前記特定パターンデータの格納先となる前記仮想ボリューム内のページが前記第二の状態であった場合、前記送信先ストレージサブシステムは当該 ID と前記格納先となる前記セグメントとを対応付けて前記メモリ内で管理することを特徴とする請求項 16 記載のストレージシステム。

【請求項 19】

前記コピー元ボリュームは、前記送信元ストレージサブシステムが有する仮想ボリュームであることを特徴とする請求項 16 記載のストレージシステム。

【請求項 20】

前記コピー元ボリュームは、複数の記憶媒体からなる物理ボリュームであることを特徴とする請求項 16 記載のストレージシステム。

【発明の詳細な説明】

【技術分野】

【0001】

複数のディスク装置を有するストレージサブシステムに関する。特に、ストレージサブシステムが有する記憶領域を有効に使用するための技術に関する。

10

20

30

40

50

## 【背景技術】

## 【0002】

計算機システムに用いられる記憶装置システムとして、複数のディスク装置を有するディスクアレイシステム（以下、ストレージサブシステムとも呼ぶ）がある。ディスクアレイシステムでは、複数のディスク装置を並列に動作させることで、ディスク装置を単体で用いた記憶装置システムに比べ高速化を実現している。

## 【0003】

ディスクアレイシステム内のディスク容量を効率的に利用するための技術として、仮想ボリューム技術がある。本技術は、仮想的な記憶領域を計算機に提供しておき、仮想的な記憶領域の内ライト要求があった領域にのみ物理的な記憶領域を割り当てる技術である。

10

## 【0004】

また、ディスク容量を効率的に利用するための別の技術として重複排除技術がある。本技術は、ライト要求を受けたデータと同じデータが、当該ライト要求を受信する以前にストレージサブシステムに格納されている場合には、当該既に格納されているデータと同一データである旨のみを記憶し、ライト要求を受けたデータそのものはストレージサブシステムに格納しない技術である。

## 【0005】

重複排除技術には、ストレージサブシステムに格納する全てのデータに対して重複排除制御を行うものと、予め登録した特定パターンデータに対してのみ重複排除制御を行うものがある。前者はデータを格納するために必要とされる記憶容量の削減効果（以下容量削減効果と呼ぶ）を高くすることができるが、処理時間が長くなる上に管理情報が多くなる。一方、後者は管理情報を少なくすることができるが、容量削減効果が限定される。

20

## 【0006】

特許文献1では、複数の物理記憶装置の集合からなる物理記憶スペース上の領域をある単位長の物理ブロックで分割し、物理ブロックによって物理記憶スペースを割り当てることにより論理記憶装置を構成するストレージ装置において、前記論理記憶装置上の領域毎に、前記物理ブロックが割り当てられているか又は所定のデータ配置パターンから生成されるデータに従うかを示す識別情報を記録する論理記憶管理テーブルを備え、前記論理記憶装置の領域に対するデータの書き込み要求があった場合、書き込み要求データのデータ配置パターンが予め登録されたデータ配置パターンと一致する場合に、一致するデータ配置パターンの識別情報を、前記論理記憶管理テーブルの書き込み要求のあった領域に対応させて記録し、前記書き込み要求データを破棄する技術が記載されている。

30

## 【0007】

【特許文献1】特開2005-135116号公報

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0008】

特許文献1では、仮想ボリューム技術と重複排除技術の双方が用いられているが、仮想ボリュームの管理サイズと重複排除の管理サイズが同一である。仮想ボリューム技術及び重複排除技術による容量削減効果を高くするために両者の管理サイズを小さくした場合、管理情報が大きくなり、全ての管理情報をキャッシュ等の高速アクセスできる記憶媒体に格納しようとする、システムが高価になる。一方で、管理サイズを大きくすると管理情報は減るが、容量削減効果は低くなる。すなわち、管理情報量と要領削減効果との間にはトレードオフの関係がある。従って、仮想ボリューム技術及び重複排除技術のための管理情報量を抑えつつ、ストレージサブシステムの記憶容量を効率的に使用するための技術が求められている。

40

## 【課題を解決するための手段】

## 【0009】

ストレージサブシステムは、プロセッサと、メモリと、仮想ボリュームと、複数の記憶媒体から構成され仮想ボリュームに割り当てられるページ用物理領域を有するページ用ボ

50

リユームとを有している。そして、プロセッサは仮想ボリュームのアドレス空間を複数のページに分割し、複数のページ各々を少なくとも第一の状態と第二の状態とを含む複数の状態のうちいずれかの状態に分類して管理する。プロセッサは第二の状態に分類されるページをさらに複数のセグメントに分割して管理している。ここで、第一の状態とは、ページにライトデータを書き込むためのライト要求を受けた結果、当該ページにページ用ボリュームからページ用物理領域が割り当てられ、ライトデータがページ用物理領域に格納された状態を指す。また、第二の状態とは、ページ内のセグメントに予め定められた特定パターンデータと一致するライトデータを書き込むためのライト要求を受けた結果、メモリ内に特定パターンデータとセグメントとが対応付けられて管理されている状態を指す。

【発明の効果】

10

【0010】

ボリューム仮想化技術及び重複排除技術のための管理情報量を抑えつつ、ストレージサブシステムの記憶容量を効率的に使用することができる。

【発明を実施するための最良の形態】

【0011】

以下に、本発明の実施の形態を説明する。

【実施例1】

【0012】

以下に本発明の実施形態の一例を説明する。尚、本発明は以下に記載する実施形態のみに限定されるものではない。

20

【0013】

図1は、本発明を適用するストレージシステムのハード構成の一例を示している。ストレージシステムは、ストレージサブシステム0000、ホストコンピュータ(ホスト)0001を有しており、ストレージサブシステム0000とホスト0001は、ホスト・インタフェース(ホストI/F)0002によって接続される。ストレージサブシステム0000は、ホスト・インタフェース0002、管理インタフェース(管理I/F)0003、メモリとCPUから構成されるコントローラ0004、キャッシュメモリ0005、ディスク・インタフェース(ディスクI/F)0007、複数のハードディスク(HDD)0006から構成され、HDD以外は内部バス0008で接続されている。ストレージサブシステム0000の構成管理は、管理端末0009上で動作する管理ソフトウェア0010によって行われる。管理端末0009とストレージサブシステム0000とは、管理インタフェース0003によって接続される。

30

【0014】

尚、図示しないが、管理端末0009及びホスト0001は各々、CPU、メモリ、ストレージサブシステム0000との通信を行うための通信I/F等を有する一般的な計算機であり、上述の管理ソフトウェア0010は管理端末0009内のメモリに格納されて管理端末0009のCPUによって実行される。

【0015】

図2は、ストレージサブシステム0000の論理的なシステム構成の一例を示している。各種プログラム及び各種テーブルの詳細については後述する。尚、これらのプログラム及びテーブルは図1に示すコントローラ0004内のメモリに格納されており、コントローラ0004のCPUによって実行される。ただし、0113、0114、0117、0118のプログラムについては、実施例1では使用しないため、実施例2以降で詳細に説明する。

40

【0016】

0104は通常ボリュームであり、ストレージサブシステム0000は複数の通常ボリューム0104を有する。各通常ボリューム0104は、複数のHDD0006によって構成されるRAIDグループ等のHDDグループ上に構成された記憶領域であって、ホスト0001及び管理端末0009からは1つの論理的な記憶領域として認識される。

【0017】

50

0102は仮想ボリュームであり、ストレージサブシステム0000は複数の仮想ボリューム0102を有する。仮想ボリューム0102は通常ボリューム0104と同様のアドレス空間を持ち、ホスト0001からは通常ボリューム0104と同様に1つの論理的な記憶領域として認識され、通常ボリューム0104と同様にデータのリード、またはライトが行われる。仮想ボリューム0102と通常ボリューム0104との相違点は、通常ボリューム0102にはデータのライト前から物理領域が割り当てられているのに対し、仮想ボリューム0104にはデータがライトされたことを契機に物理領域が割り当てられる点である。仮想ボリューム0104を用いれば、ライトがない場合は物理領域を割り当てずに済むため、ストレージに実装する物理領域を少なくすることができ、記憶容量の有効活用が可能となる。尚、仮想ボリュームに割り当てられる物理領域は、ページ用プール0105から確保される。ページ用プール0105にはページ用ボリューム0103が複数含まれる。ページ用ボリューム0103は通常ボリューム0104と同様、複数のHDD0006によって構成されるRAIDグループ等のHDDグループ上に構成された記憶領域である。従って仮想ボリュームには、ページ用プール0105に属するページ用ボリューム0103上の物理領域が割り当てられることとなる。

10

**【0018】**

図3は、仮想ボリューム0102のアドレス空間の一例を示している。仮想ボリューム0102のアドレス空間は複数個のページに分割されて管理される。ページの管理が容易になるため、通常はページのサイズを固定サイズとする。仮想ボリュームのアドレス空間0201上のページには、(a1)(a2)(a3)の3つの状態があり、それぞれの状態に応じて異なる管理が行われる。さらに(a2)の状態であるページは、セグメントという単位で複数個に分割されて管理される。ここで、図3ではページ内を8個のセグメントに分割しているが、ページ内のセグメント数は2個以上であれば個数は問わない。また、セグメントの管理が容易となるため、通常はセグメントのサイズを固定サイズとする。

20

**【0019】**

(a1)(a2)(a3)の状態について、詳細に説明する。(a1)の状態は、ページ内にホスト0001からのライト要求があり、その結果としてページに物理領域が割り当てられた状態である。

**【0020】**

(a2)の状態は、ページ内にホスト0001からライト要求があったが、ライトデータがストレージサブシステム0000の管理する特定パターンのデータと一致したため、特定パターンデータである旨のみが管理情報として後述するセグメント管理テーブル0122に記録され、ページに物理領域が割当てられていない状態である。即ちこの状態のページは、先述の重複排除技術によって記憶容量の削減が行われている。また(a2)の状態にあるページをセグメント単位で管理するのは、上述の通り重複排除制御を行う際には管理単位をより小さくした方が重複排除による記憶容量の削減効果が高いからである。即ち、重複排除制御が実施されるページ((a2)状態のページ)についてはよりサイズの小さいセグメントを管理単位とすることでストレージサブシステムにおける記憶容量の削減効果を高め、重複排除制御が実施されないページ((a1)及び後述の(a3)状態のページ)についてはセグメント単位ではなくページ単位で管理を行うことによって管理情報及び処理負荷の増大を防止しているのである。

30

40

**【0021】**

(a3)の状態は、ページ内に未だホスト0001からのライト要求がないため、物理領域が割当てられていない状態である。この状態のページは先述の仮想ボリューム技術によって記憶容量が削減されている。

**【0022】**

以上のように、ホストからライト要求があったかどうか、及びライト要求があった場合にはライト要求に含まれるライトデータが特定パターンのデータであるか否かに応じて、仮想ボリューム0102のアドレス空間上の各ページを(a1)(a2)及び(a3)の状態に分類して管理することにより、管理情報量を抑えつつ容量削減効果を高くすること

50

が可能となる。

【 0 0 2 3 】

図 4 は、仮想ボリューム 0 2 0 1 のアドレス空間を管理するために用いられる、仮想ボリューム管理情報 0 2 0 2 の一例を示す。仮想ボリューム管理情報 0 2 0 2 は、仮想ボリューム 0 1 0 2 の作成時に、仮想ボリューム 0 1 0 2 ごとに作成される。各仮想ボリューム管理情報 0 2 0 2 は、仮想ボリュームを識別する仮想ボリューム番号（仮想ボリューム # ） 0 2 1 1 の欄と、1つのページ管理テーブル 0 1 2 1 を有しており、更に仮想ボリューム # 0 2 1 1 で識別される仮想ボリュームに状態が（ a 2 ）であるページが存在する場合には、状態が（ a 2 ）であるページの個数分セグメント管理テーブル 0 1 2 2 を有する。

10

【 0 0 2 4 】

ページ管理テーブル 0 1 2 1 は、仮想ボリューム 0 2 0 1 のアドレス空間上の各ページの状態を管理するテーブルである。セグメント管理テーブル 0 1 2 2 は、（ a 2 ）の状態であるページ内のセグメントを管理するテーブルであり、（ a 2 ）の状態であるページの個数分作成される。

【 0 0 2 5 】

まず、ページ管理テーブル 0 1 2 1 の説明をする。ページ管理テーブル 0 1 2 1 のページ番号（ページ）欄 0 2 0 3 には、仮想ボリューム 0 1 0 2 が有するページの番号が記録される。状態欄 0 2 0 4 には、該ページの状態が記録される。本実施例では、（ a 1 ）の状態に対応する値を 0、（ a 2 ）の状態に対応する値を 1、（ a 3 ）の状態に対応する値を 2 とする。領域番号（領域 # ）欄 0 2 0 5 には、当該ページにページ用プール 0 1 0 5 内のどの物理領域が割り当てられたかを示す情報が記録される。従って、状態 0 2 0 4 が 0 である場合、すなわちページの状態が（ a 1 ）である場合には、該ページに割り当てられた物理領域を示す識別番号が領域 # 0 2 0 5 に記録されるが、状態 0 2 0 4 が 1 若しくは 2 である場合、即ちページ状態が（ a 2 ）若しくは（ a 3 ）である場合には、当該ページには物理領域が割り当てられていないので領域 # 0 2 0 5 には null が記録される。尚、ページ用プール 0 1 0 5 内の物理領域は後述するプール管理テーブル 0 1 2 3 によって管理されている。テーブル番号（テーブル # ）欄 0 2 0 6 には、当該ページのセグメントを管理するセグメント管理テーブル 0 1 2 2 を示す識別番号が記録される。従って状態 0 2 0 4 が 1 である場合、即ちページの状態が（ a 2 ）である場合には当該ページに対応するセグメント管理テーブル 0 1 2 2 の識別番号がテーブル # 0 2 0 6 に記録されるが、それ以外の場合（ページ状態が（ a 1 ）若しくは（ a 3 ）の場合）にはページはセグメント単位で管理されていないためセグメント管理テーブル 0 1 2 2 も存在せずテーブル # 0 2 0 6 には null が記録される。特定パターンカウント数欄 0 2 0 7 には該ページに対するライト要求の内、ライトデータがストレージサブシステムが管理する特定パターンデータと一致した回数を示す数が記録される。尚、特定パターンカウント数 0 2 0 7 は、状態が（ a 1 ）であるページについて記録され、状態が（ a 2 ）若しくは（ a 3 ）であるページについては null 値が記録される。ページの状態が（ a 2 ）である場合、当該ページに書き込まれたライトデータは全て特定パターンデータと一致しているはずであり、ページの状態が（ a 3 ）である場合は未だホストからライト要求を受信していないため、特定パターンと一致するデータの数をカウントする必要がないからである。

20

30

40

【 0 0 2 6 】

次にセグメント管理テーブル 0 1 2 2 を説明する。セグメント管理テーブル 0 1 2 2 のセグメント番号欄（セグメント # ） 0 2 0 8 には、一つのページ内に存在する各セグメントに付けられた識別番号が記録される。各セグメントは、特定パターンデータをライトするようホストからライト要求を受け特定パターンデータが書き込まれた旨の記録を行っている状態と、ホストから未だライト要求を受けていない状態の 2 つの状態を取りうる。フラグ欄 0 2 0 9 には、上記セグメントの状態を識別する値が記録される。本実施例では特定パターンデータが書き込まれた旨の記録を行っている状態を 0、セグメントに対してライト要求を未だ受け取っていない状態を 1 とする。パターン ID 0 2 1 0 にはセグメントに

50

対してホストからライト要求のあったデータのパターンを示す識別番号が記録される。従って、フラグ0209が0の場合にのみ、パターンID0210欄に特定データパターンを識別するIDが記録されることとなり、フラグ0209が1の場合にはパターンID0210にはnullが記録される。尚、ホストからライトを要求されたライトデータが特定パターンデータであるか否かを識別するために使用する、特定パターンデータを管理するテーブル0120については後述する。

#### 【0027】

図5はページ用プール0105を構成する記憶領域（物理領域）を管理するプール管理テーブル0123の一例を示す。領域番号（領域#）1231欄にはページ用プール0105を構成する物理領域の識別番号が記録される。この領域#は先述のページ管理テーブル0121の領域#0205から参照されている。割り当てフラグ1232欄には、当該物理領域が仮想ボリューム0102に割り当てられているか否かを示すフラグが記録される。本実施例では、領域が未割り当ての場合は0を、割り当て済みの場合は1を記録する。

10

プールLU番号（プールLU#）欄1234には、領域#1231が示す物理領域が属するページ用ボリューム0103の識別番号が記録され、開始アドレス欄1235には当該物理領域のページ用ボリューム0103内における位置を特定するための開始アドレスが記録される。即ち、ページ用プール0105を構成する各物理領域は、プールLU#1234によって識別されるページ用ボリューム0103内の、開始アドレス1235からページサイズと等しいサイズ分の記憶領域に相当する。

20

#### 【0028】

図6は仮想ボリューム0102内へライト要求のあったライトデータが特定パターンデータであるか否かを判断する際に使用する、特定パターンデータを管理するための特定パターン管理テーブル0120の一例である。特定パターン管理テーブル0120においては、特定パターンデータを識別するIDがパターンID1201に記録され、パターンデータ1202の列には、該特定パターンのデータが記録される。記録される特定パターンのデータサイズは、セグメントサイズと同一とする。

#### 【0029】

図7は、ページの状態（a1）から（a2）へ遷移させるか否かを判断する際に用いられる閾値を管理する閾値管理テーブル0124の一例である。尚、ページの状態（a1）から（a2）へ遷移させる処理は図13を用いて後述する。閾値管理テーブル0123では、仮想ボリューム0102を識別する仮想ボリューム番号が仮想ボリューム#1241欄に記録され、仮想ボリュームごとに特定パターンカウント数の閾値1242が記録されている。ユーザは管理ソフトウェア0010上で、各々の仮想ボリューム0102に対して閾値を指定することが可能である。そして、先述のページ管理テーブル0121で管理されている特定パターンカウント数0207が閾値に達した場合に、ページ状態（a1）にあるページに実際に書き込まれているデータが調べられ、当該データが全て特定パターンデータであると認められた場合に、当該データの状態が（a1）から（a2）へ遷移されることとなる。

30

#### 【0030】

図8は、仮想ボリューム作成プログラム0110の処理フロー0301の一例を示す。まずユーザは、管理ソフトウェア0010上で、仮想ボリューム0102の作成をストレージサブシステム0000の仮想ボリューム作成プログラム0110に指示する。仮想ボリューム作成プログラム0110は管理ソフトウェア0010から仮想ボリューム作成コマンドを受領し（ステップ0302）、仮想ボリューム0102を作成し、併せて仮想ボリューム管理情報0202を作成する（ステップ0303）。即ち仮想ボリューム作成プログラム0110は、作成した仮想ボリュームに識別番号を割り当てて、仮想ボリューム管理情報0202の仮想ボリューム#0211に記録する。また、このとき作成された仮想ボリューム0102内のページに対しては未だホストからのライト要求を受け付けていないため、該仮想ボリュームに対して作成されたページ管理テーブル0121の状態02

40

50

04には、全てのページについて(a3)を示す2を記録し、領域#0205、テーブル#0206、特定パターンカウント数0207にはnull値を記録する。更にセグメント管理テーブル0122は作成しない。次に、仮想ボリューム作成プログラム0110は閾値管理テーブル0124の仮想ボリューム#1241に、作成した仮想ボリューム0102を識別する番号を記録し、ユーザが管理ソフトウェア0010上で任意に指定した閾値を閾値1242に記録する(ステップ0304)。そして仮想ボリューム作成プログラム0110は、管理ソフトウェア0010に仮想ボリューム作成終了を応答して(ステップ0305)処理を終了する。

【0031】

図9は、特定パターン設定プログラム0115の処理フロー0401の一例を示す。ユーザは、管理ソフトウェア0010上で特定パターンとして追加(または削除)するデータを指定し、特定パターン設定プログラム0115に、特定パターン管理テーブル0120への特定パターンデータの追加(または削除)を指示する。特定パターン設定プログラム0115は、管理ソフトウェア0010から特定パターンデータの設定コマンドを受領する(ステップ0402)。設定コマンドが特定パターンデータとして新たなデータを追加するよう指示するものである場合には、特定パターン設定プログラム0115はコマンドで指定されるデータにIDを割り当てて、該IDと共に指定されたデータを特定パターン管理テーブル0120に記録する(ステップ0403)。

【0032】

設定コマンドが特定パターン管理テーブル0120からの特定パターンデータの削除を指示するものである場合には、特定パターン設定プログラム0115は、設定コマンドで指定されるIDから削除対象の特定パターンデータを識別し、特定パターン管理テーブル0120から該IDと識別された特定パターンデータを削除する(ステップ0403)。

【0033】

特定パターン管理テーブルの更新が終了すると、特定パターン設定プログラム0115は、設定終了を管理ソフトウェア0010に応答して(ステップ0404)、処理を終了する。

【0034】

図10は、ページの状態を(a1)から(a2)へと変遷させるか否かを判断する際に用いられる閾値を変更する閾値変更プログラム0116の処理フロー0116の一例を示す。ユーザは、管理ソフトウェア0010上で閾値を変更したい仮想ボリュームを識別する番号と変更後の閾値を指定し、管理ソフトウェア0010は閾値変更プログラム0116に、閾値管理テーブル0124上の閾値1242の変更を指示する。閾値変更プログラム0116は管理ソフトウェア0010から閾値変更コマンドを受領し(ステップ0502)、閾値管理テーブル0124にて指定された仮想ボリュームの閾値を指定された値に変更する(ステップ0503)。閾値変更プログラム0124は閾値の変更が終了すると、閾値変更終了を管理ソフトウェア0110に応答する(ステップ0504)。

【0035】

図11は、リードプログラム0112の処理フロー1001の一例である。リードプログラム0112はホスト0001から仮想ボリューム0102に対するリード要求を受領すると(ステップ1001)、指定された仮想ボリュームのページ管理テーブル0121を参照し、リード要求があったアドレスに対応するページの状態0204をチェックする(ステップ1002)。尚、リード要求の中にはリード対象の仮想ボリュームを指定するための仮想ボリューム番号と、当該仮想ボリュームにおけるリード対象領域のアドレス情報が含まれている。一方仮想ボリュームを構成する各ページは所定のサイズを有するため、リード要求に含まれるアドレス情報と各ページのサイズからリード対象のページを特定することができる。

【0036】

ページの状態が(a1)である場合、該ページには物理領域が割り当てられているため、リードプログラム0112はページ管理テーブル0121の領域#0205を参照し、

10

20

30

40

50

該ページに対応する物理領域の識別番号を特定する。更に特定された物理領域の識別番号に基づいてプール管理テーブル0123を参照し、リードプログラム0012はリード対象のページ用ボリューム0103の番号1234と当該ページ用ボリューム0103内におけるリード対象領域の開始アドレス1235を取得する(ステップ1004)。そして、リードプログラム0112は取得したページ用ボリュームの番号とアドレスで指定される物理領域から、リード対象のデータをリードし、リードしたデータをホスト0001へ送信する(ステップ1005)。

【0037】

ページの状態が(a2)である場合、該ページはさらにセグメントに分割されて管理されているため、リードプログラム0112は指定された仮想ボリューム0102についてのページ管理テーブル0121のテーブル#0206を参照し、該ページに対応するセグメント管理テーブル0122を取得する(ステップ1006)。そして、リードプログラム0112は取得したセグメント管理テーブル0122を参照してリード対象のセグメントのフラグ0209を確認する(ステップ1007)。尚、ページを構成するセグメントも所定のサイズを有するため、リード要求に含まれるアドレス情報と各セグメントのサイズからリード対象のページを特定することができる。

10

【0038】

該フラグが0であった場合、該セグメントには特定パターンのデータが割り当てられているため、リードプログラム0112はパターンID0210を参照し、該IDで識別されるデータを特定パターン管理テーブル0120より取得し該データをホスト0001へ送信する(ステップ1009)。一方、フラグ0209が1であった場合は、該セグメントには未だデータが書き込まれていないため、所定のフォーマットデータをホスト0001へ送信する(ステップ1010)。尚、一般的なストレージサブシステムではフォーマットデータをゼロとすることが多いが、この限りではない。また、フォーマットデータは、仮想ボリュームごとにユーザが指定することも可能である。

20

【0039】

ページの状態が(a3)である場合は、該ページには未だデータが書き込まれていない。従ってリードプログラム0112はステップ1010と同様フォーマットデータをホスト0001へ送信する(ステップ1003)。

【0040】

図12は、ライトプログラム0111の処理フロー0601の一例である。ライトプログラム0111は、仮想ボリューム0102に対するライト要求をホスト0001から受領すると、指定された仮想ボリュームについてのページ管理テーブル0121を参照し、ライト要求があったアドレスに対応するページの状態0204をチェックする(ステップ1002)。尚、ライト対象のページの指定方法は、先述のリード処理の際と同様である。

30

【0041】

ライトプログラム0111はページの状態が(a1)である場合はWrite処理(1)(ステップ0604)、ページの状態が(a2)である場合はWrite処理(2)(ステップ0605)、ページの状態が(a3)である場合はWrite処理(3)(ステップ0606)を実行する。Write処理(1)(2)(3)についての詳細は図13、図14、図15を用いて説明する。

40

【0042】

図13は、ライトプログラム0111がライト要求を受けた際、ライト対象のページの状態が(a1)である場合に実行される、先述のWrite処理(1)(ステップ0604)の一例を示す。ライトプログラム0111はライト対象の仮想ボリュームについてのページ管理テーブル0121を参照し、ライト対象のページに割り当てられている物理領域の領域#0205を特定する。ライトプログラム0111は、特定された物理領域のプールLU#1234と開始アドレス1235をプール管理テーブル0123から取得する(ステップ0701)。そして、取得したプールLU#1234及び開始アドレス1235で

50

指定される物理領域にライトデータを書き込み（ステップ0702）、ライト完了報告をホスト0001に送信する（ステップ0703）。

【0043】

その後、ライトプログラム0111は特定パターン管理テーブル0120を参照して、ライト要求のあったライトデータが特定パターンデータのいずれかと一致しているか否かを判定する（ステップ0704）。ライトデータが特定パターンデータ1202と一致しない場合は、ライト処理を終了する。ライトデータが特定パターンデータ1202に一致する場合はステップ0705に進み、ページ管理テーブル0121を参照して、ライト要求を受けたページの特定パターンカウント数0207に1を加える。

【0044】

ステップ0705の後ステップ0706に進み、ライトプログラム0111は、閾値管理テーブル0124を参照し、ライト対象の仮想ボリュームに対応する閾値1242を特定し、ステップ0705にて1を加えた後の特定パターンカウント数0207が該閾値を超えているか否かを判定する。該閾値を超えていない場合はライト処理を終了する。特定パターンカウント数が該閾値を超えている場合にはステップ0707に進み、ライトプログラム0111は該ページ内に書き込まれているデータが全て特定パターンデータと一致するデータであるか否かを判定する。該ページに書き込まれているデータが1つでも特定パターンデータと一致しない場合には、ライト処理を終了する。該ページに書き込まれているライトデータが全て特定パターンデータと一致する場合には、ステップ0708に進み、ライトプログラム0111はライト対象のページについてセグメントテーブル0122を新規に作成する。このとき当該ページについては全データが特定パターンデータであるため、該セグメントテーブル0122のフラグ0209には全て0が設定される。そしてライトプログラム0111は、特定パターン管理テーブル0120を参照して、当該ページの各セグメントに格納されている特定パターンデータを識別するIDを特定してパターンID0210に記録する（ステップ0708）。その後、ライト対象のページに割り当てていたページ用プール0105内の物理領域を開放する（ステップ0709）。ステップ0708、0709により、ライト対象のページの状態を（a1）から（a2）へと遷移させたため、ライトプログラム0111は、ページ管理テーブル01021の該ページに対応する状態0204を0から1に変更する（ステップ0710）。また、ページ管理テーブル0121のテーブル#0206に、ステップ0708で新規に作成されたセグメント管理テーブル0122を識別する番号を記録する（ステップ0711）。

【0045】

以上Write処理（1）によれば、一旦物理領域が割り当てられていたページであっても、特定パターンデータのみで構成されることになったページについては、割り当てていた物理領域を開放し、どの特定パターンデータがホストから書き込まれたかを示す管理情報だけを記録すれば良く、容量削減効果を高くすることが可能となる。

【0046】

図14は、ライトプログラム0111が、ライト対象のページの状態が（a2）である場合に実行する、先述のWrite処理（2）（ステップ0605）の一例を示す。ライトプログラム0111は、特定パターン管理テーブル0120を参照して、パターンデータ1202の中にライト要求を受けたライトデータと一致するデータがあるか否かを判定する（ステップ0801）。

【0047】

ステップ0801にて一致する特定パターンデータが存在する場合はステップ0802に進み、ライトプログラム0111はライトデータと一致する特定パターンデータを識別するIDを、特定パターン管理テーブル0120のパターンID1201から特定する。そしてライト対象のページのセグメント管理テーブル0122を参照し、ライト対象のセグメントのパターンID0210に特定したパターンデータのIDを記録する（ステップ0802）。尚、ライト対象のセグメントはリード対象のセグメントと同様の方法にて特定することができる。その後、ライトプログラム0111は、ライト完了報告をホスト0001に

10

20

30

40

50

送信して(ステップ0803)、ライトデータを解放し(ステップ0804)、ライト処理を終了する。

【0048】

ステップ0801にてライトデータが特定パターンデータと一致しない場合はステップ0805に進み、ライトプログラム0111はライト対象のページにページ用プール0105から物理領域を割り当てるため、プール管理テーブル0123を参照する。そして、ライトプログラム0111は未割当の物理領域を示す割り当てフラグ1232が0である領域を確保する(ステップ0805)。即ち、ライトプログラム0111は当該未割当物理領域のフラグ1232を1に変更すると共に、ライト対象の仮想ボリュームのライト対象ページのページ管理テーブル0121に、領域#0205として確保した物理領域の識別番号を記録する。

10

【0049】

ステップ0805にて確保した物理領域に、ライトプログラム0111は、ライト対象のページに元々割り当てられている(即ちホストからライト要求を受け、実際の物理領域へのデータの格納は行わずに特定パターンデータが書き込まれた旨の管理のみをしている)特定パターンデータを格納する(ステップ0806)。即ちライトプログラム0111は、ライト対象の仮想ボリュームについてのページ管理テーブル0121を参照し、ライト要求を受けたページに対応するテーブル#0206から該ページのセグメント管理テーブル0122を取得する。そして、該セグメント管理テーブル0122においてパターンID0210が記録されている各セグメントについて、ライトプログラム0111は特定パターン管理テーブル0120を参照し、該パターンIDで識別される特定パターンデータを取得して、該セグメントに対応する確保した物理領域上の記憶領域に該特定パターンデータを格納する。尚この時物理領域上の書き込まれた特定パターンデータの数を、ライトプログラム0111は、当該ページの特定パターンカウント数0207としてページ管理テーブル0121に記録することとしても良い。

20

【0050】

さらにライトプログラム0111は、ライト対象セグメントに対応する、確保した物理領域上の記憶領域に、ライトデータを書き込む(ステップ0807)。

【0051】

ステップ0805~0807により、ライト要求を受けたページの状態を(a2)から(a1)へと遷移させたため、ライトプログラム0111は、ページ管理テーブル01021の該ページに対応する状態0204を1から0に変更し、テーブル#0206にnullを記録すると共に、当該ページに関するセグメント管理テーブル0122を破棄する(ステップ0808)。その後ライトプログラム0111はホスト0001にライト完了報告を送信して(ステップ0809)ライト処理を終了する。

30

【0052】

図15は、ライト要求対象のページの状態が(a3)である場合にライトプログラム0111が実行するWrite処理(3)の一例を示す。ライトプログラム0111は、特定パターン管理テーブル0120を参照して、特定パターンデータ1202の中にホストから受信したライトデータと一致するデータがあるか否かを判定する(ステップ0901)。

40

【0053】

ステップ0901にて一致するデータが存在する場合はステップ0902に進み、ライトプログラム0111はライト要求対象のページのセグメントテーブル0122を新規に作成する(ステップ0902)。そしてライト対象のセグメントについてフラグ0209に0を設定し、さらに特定パターン管理テーブル0120を参照して、ライトデータと同じ特定パターンデータを識別するIDを特定し、特定したIDをパターンID0210に記録する(ステップ0903)。ステップ0902、0903によりライト対象のページの状態が(a3)から(a2)へと変遷したため、ライトプログラム0111は該ページのページ管理テーブル0121を参照し、該ページの状態0204を2から1に変更する(ステップ0904)。また、ページ管理テーブル0121のテーブル#0206にステップ0

50

902で新規に作成されたセグメント管理テーブル0122を識別する番号を記録する(ステップ0905)。最後にライトプログラム0111は、ライト完了応答をホスト0001に送信して(ステップ0906)、ライトデータを解放し(ステップ0907)、ライト処理を終了する。

#### 【0054】

ステップ0901にて一致する特定パターンデータが存在しない場合はステップ0906に進み、ライトプログラム0111はライト対象のページにページ用プール0103から物理領域を割り当てるため、プール管理テーブル0123を参照して、未割当の物理領域を示す割り当てフラグ1232が0である物理領域を確保する(ステップ0906)。物理領域の確保の方法は、先述のWrite処理(2)のステップ0805と同様である。そして、ライト要求に含まれる仮想ボリュームの識別番号とライト対象領域のアドレスに基づいて、対応する物理領域上の記憶領域を割り出し、割り出された記憶領域にライトデータを書き込む(ステップ0907)。ステップ0906、0907によりライト対象のページの状態が(a3)から(a1)へと変遷したため、ライトプログラム0111は該ページのページ管理テーブル0121を参照し、該ページの状態0204を2から0に変更する(ステップ0908)。最後に、ライトプログラム0111はホスト0001にライト完了コ報告を送信して(ステップ0909)ライト処理を終了する。

10

#### 【実施例2】

#### 【0055】

本実施例では、通常ボリューム0104に格納されているデータを仮想ボリューム0102に移行することによって、通常ボリューム0104に格納されている特定パターンデータの重複を排除し、ストレージサブシステムのデータ量を削減する方法を説明する。

20

#### 【0056】

図16は、図2に示す通常ボリュームの重複排除プログラム0114のフローの一例を示す。まず通常ボリュームの重複排除プログラム0114は、通常ボリューム0104と同じ容量を有する仮想ボリューム0102をデータ移行先のボリュームとして選択する(ステップ2401)。尚、重複排除プログラム0114はデータ移行先の仮想ボリューム0102を仮想ボリューム作成プログラム0110に指示して作成させても良い。次に通常ボリュームの重複排除プログラム0114は、通常ボリュームからセグメントサイズと同じサイズ分のデータを順にリードし(ステップ2402)、リードしたデータを実施例1で説明したライトプログラム0111を使って仮想ボリューム0102に対してライトさせる(ステップ2403)。

30

#### 【0057】

ライト処理の方法は、実施例1で説明した図12~図15と同様であるため説明を省略するが、データの移行先仮想ボリューム0102は、ライトプログラム0111の処理によって、先述の状態(a1)、(a2)のいずれかの状態を有するページに分類される。状態(a1)のページについては物理領域が割り当てられ、当該物理領域にデータが書き込まれるが、状態(a2)のページには重複排除技術が用いられて物理領域が割り当てられない。この結果ストレージシステムのデータ容量を削減することができる。

#### 【0058】

以上のように、本発明によれば通常ボリューム0102内のデータを仮想ボリューム0102に移行する際にも容量削減効果の高い重複排除が可能となる。

40

#### 【実施例3】

#### 【0059】

本実施例では、図17に示す様に、実施例1で説明したストレージサブシステム0000がネットワーク2101~2103を介して複数台接続されている場合に、該複数のストレージサブシステム0000間で容量削減効果の高いリモートコピーを行う方法を説明する。

#### 【0060】

リモートコピーには図2に示すリモートコピー通信プログラム0118を使用し、送信

50

元となるストレージサブシステムではパターン送信プログラム 2102 とデータ送信プログラム 2104 を使用し、送信先となるストレージサブシステムではパターン受信プログラム 2103 とデータ受信プログラム 2105 を使用してストレージサブシステム 0000 間におけるリモートコピーを行う。各プログラム (2102 ~ 2105) の詳細は後述する。本実施例では、一方のストレージサブシステムの通常ボリューム 0104 内のデータを他方のストレージサブシステムの仮想ボリューム 0102 にリモートコピーする場合を考える。

#### 【0061】

リモートコピーの事前準備として、まず送信元ストレージサブシステムと送信先ストレージサブシステムがそれぞれ有する特定パターン管理テーブル 0120 のデータを同期する。このためにパターン送信プログラム 2102 とパターン受信プログラム 2104 とが用いられる。

10

#### 【0062】

図 18 (a) は、送信元ストレージサブシステムで使用されるパターン送信プログラムの処理フローの一例を示す。パターン送信プログラム 2102 は、送信元ストレージサブシステムが有する特定パターンテーブル 0120 をシリアルライズする (ステップ 2301)。シリアルライズの方法については複数通りの方式が考えられるが、本明細書では詳細な記述は行わない。尚、シリアルライズとはソフトウェア内部で扱っているデータを丸ごとファイルで保存したり、ネットワークを介して送受信したりするために変換することをいう。次にパターン送信プログラム 2012 は、シリアルライズ後の特定パターンテーブル 0120 のシリアルデータを送信先ストレージサブシステムのパターン受信プログラム 2014 に送信する (ステップ 2302)。そして、送信先ストレージサブシステムからの応答を受信して終了する (ステップ 2303)。

20

#### 【0063】

図 18 (b) は、送信先ストレージサブシステムで使用されるパターン受信プログラムの処理フローの一例を示す。パターン受信プログラム 2104 は、パターン送信プログラム 2102 から送信されたシリアルデータを受信し (ステップ 2304)、該シリアルデータを特定パターンテーブル 0120 のデータに変換する (ステップ 2305)。そして、変換後のデータを送信先ストレージサブシステムの特定パターン管理テーブル 0120 と置き換える (ステップ 2306)。置き換えが完了したら送信元ストレージサブシステムに対して完了応答を送信する (ステップ 2307)。

30

#### 【0064】

次に、複数のストレージサブシステム間においてリモートコピーを実施する際のデータの送受信方法について説明する。図 19 は、送信元ストレージサブシステムから送信先ストレージサブシステムに送信されるパケットの実装例を示す。パケットは、送信先ストレージサブシステムを特定するアドレス情報を格納するヘッダ部 1801 と、送信データを格納するデータ部とで構成される。データ部は、データ識別フラグ 1802 と仮想ボリューム情報 1803 とデータ格納部 1804 とで構成されている。データ識別フラグ 1802 とは、パケットを用いて送信されるデータが、実データであるのかパターン ID であるのかを示すフラグである。本実施例では、該フラグが 0 の場合、データ格納部 1804 には実データが格納されているものとする。一方、該フラグが 1 の場合、データ格納部 1804 はパターン ID が格納されている。仮想ボリューム情報 1803 は、送信データが書き込まれる、送信先ストレージサブシステム内の仮想ボリュームの番号とアドレスである。

40

#### 【0065】

図 20 (a) は、送信元ストレージサブシステムで使用されるデータ送信プログラム 2103 の処理フローの一例を示す。データ送信プログラム 2103 は、リモートコピーのコピー元である送信元ストレージサブシステムの通常ボリューム (以下コピー元通常ボリュームと呼ぶ) からデータを読み出す (ステップ 1701)。そしてステップ 1702 にて、データ送信プログラム 2103 は、送信元ストレージサブシステムが有する特定パターン管理テーブル 0120 を参照し、読み出したデータが特定パターンデータと一致する

50

か否かを判定する。特定パターンデータのいずれとも一致しない場合はステップ1704に進み、データ送信プログラム2103は、読み出した該データをそのままの形式でパケットとして送信先ストレージサブシステムに送信する。一方、ステップ1702にて、特定パターンデータのいずれかとステップ1701で読み出されたデータとが一致した場合、データ送信プログラム2103は、特定パターン管理テーブル0120のパターンID1201に基づいて、該データをパターンIDに変換し(ステップ1703)、該パターンIDを先述のパケットを用いて送信先ストレージサブシステムに送信する(ステップ1710)。

**【0066】**

図20(b)は、送信先ストレージサブシステムで使用されるデータ受信プログラム2105の処理フローの一例を示す。データ受信プログラム2105は、送信元ストレージサブシステムからパケットを受信すると(ステップ1705)、該パケット内のデータ識別フラグ1802を確認する(ステップ1706)。該データ識別フラグ1802が0であった場合、ステップ1708に進み、受信パケットの仮想ボリューム情報1803が示す仮想ボリュームの記憶領域に実データをライトするようライトプログラム0111に指示する。この結果、実施例1で図12~図15を用いて説明したのと同様の処理にてライトプログラム0111によりデータがリモートコピーの送信先であるコピー先仮想ボリュームに格納される。データ受信プログラム2105は、ライト処理が終了すると送信元ストレージサブシステムに完了応答を送信して終了する(ステップ1709)。ステップ1706にて、データ識別フラグ1802が1であった場合、ステップ1708に進み、データ受信プログラム2105は送信先ストレージサブシステムが有する特定パターン管理テーブル0120を参照して、データ格納部1804に格納されているパターンIDを実データに変換する(ステップ1707)。その後、データ受信プログラム2105は、先述と同様ステップ1708、ステップ1709の処理を実行して終了する。

**【0067】**

以上の方法によれば、コピー元通常ボリュームから読み出されたデータが特定パターンデータと一致する場合には、送信元ストレージサブシステムから送信先ストレージサブシステムへは実データではなくパターンIDが送信されることになるため、リモートコピーに要するデータのコピー量が低減できる。更に、送信先ストレージサブシステムにおいては、実施例1において説明したライトプログラム0111を用いてコピーデータが送信先仮想ボリュームに書き込まれるため、特定パターンデータについてはデータの重複を排除することができる。このためコピー先ストレージサブシステムの必要記憶容量を削減することができる。

**【0068】**

上述した方法では、リモートコピーのコピー元となる送信元ストレージサブシステム内のコピー元ボリュームが通常ボリューム0104であるということを前提として説明した。しかしコピー元ボリュームは仮想ボリューム0102であっても良い。コピー元ボリュームが仮想ボリューム0102であった場合には、送信元ストレージサブシステムにおいて当該コピー元仮想ボリューム0102をページ管理テーブル0121を用いてページ毎に管理しているため、送信元ストレージサブシステムは各ページの状態を把握している。従ってページ状態が(a2)である場合には、図20(a)のステップ1701及び1702の処理を実行せずとも、当該ページには特定パターンデータが対応付けられていることが明らかである。従って、図20(a)のステップ1701、ステップ1702、及びステップ1703を省略することもできる。この場合にはまず、送信元ストレージサブシステムのデータ送信プログラム2103はコピー元仮想ボリューム0102の各ページの状態を送信元ストレージサブシステムが有するページ管理テーブル0121を参照して判定する。判定の結果、該状態が(a1)である場合にはステップ1702に進み、該状態が(a2)である場合にはステップ1702、1703を省略してステップ1710に進み、該状態が(a3)である場合にはステップ1704に進む。それぞれその後のステップは先述のコピー元が通常ボリュームである場合と同様である。

## 【 0 0 6 9 】

更に上述の方法では、送信先ストレージサブシステムにおいて受信したデータがパターンIDであった場合に(即ち図20(b)のステップ1706でフラグの値が1であった場合に)、一旦パターンIDを実際の実データ(即ち特定パターンデータ)に直してから(図20(b)のステップ1707)、ライトプログラム0111を用いて当該特定パターンデータを仮想ボリュームに書き込んでいる。ライトプログラム0111が特定パターンデータを仮想ボリュームに書き込む際に、この特定パターンデータは再度パターンIDに変換されるため(図14のステップ0801及びステップ0802)、特定パターンデータとパターンIDとの間に無駄な変換処理が発生している。この無駄を省くために、ライトプログラム0111がライトデータの代わりにパターンIDを受け付けることができるようにしても良い。この結果、データ受信プログラム2105は図20(b)のステップ1707においてパターンIDを実データに変換することなくライトプログラム0111にパターンIDを渡せばよく、これを受信したライトプログラム0111は受信したパターンIDで指定される所定の特定パターンデータのライトを要求されたと解釈して、ライト処理を実行すれば良い。

10

## 【 0 0 7 0 】

以上のような方法でリモートコピーをすることにより、ネットワークの通信量を削減したデータ送受信を可能とし、さらに複数のストレージサブシステム間で容量削減効果の高い重複排除を行うことが可能となる。

## 【 0 0 7 1 】

尚、以下のような形態であっても構わない。

20

## 【 0 0 7 2 】

送信元ストレージサブシステムと、送信元ストレージサブシステムと通信可能に接続された送信先ストレージサブシステムとを有するストレージシステムにおいて、送信元ストレージサブシステム及び送信先ストレージサブシステムは、論理的な記憶領域を有する。送信元ストレージサブシステムは、ボリュームからリモートコピーの対象であるデータを読み出し、読み出したデータが特定パターンデータと一致する場合には、当該特定パターンデータを識別するIDを送信先ストレージサブシステムに送信する。このような構成により、ネットワークの通信料を削減したデータ送受信が可能となる。

## 【 実施例 4 】

30

## 【 0 0 7 3 】

本実施例では、図17に示す構成と同様ストレージサブシステム0000がネットワークを介して複数台接続されている場合に、該複数のストレージサブシステム0000間で容量削減効果の高いデータ移行を行う方法を説明する。

## 【 0 0 7 4 】

本実施例では、移行元ストレージサブシステムが有するデータ移行プログラム0113を用いて、移行元ストレージサブシステム内の全て又は一部の通常ボリューム0104のデータを移行先ストレージサブシステムの仮想ボリューム0102にコピーする。

## 【 0 0 7 5 】

本実施例ではまず、移行先ストレージサブシステムが特定パターンの分析プログラム0117を使って、移行元ストレージサブシステム内のコピー元通常ボリューム0104内に格納されているデータを分析し、特定パターンデータとしてどのようなパターンのデータを用いれば容量削減効果が高いかを調べる。後述の通り特定パターンの分析プログラム0117を実行すると、データ移行前に一旦コピー元通常ボリュームからデータを読み出してコピー先ストレージサブシステムへ送信する必要があるため、本プログラム0117の実行は必須でない。しかし、事前にコピー元通常ボリューム0104に格納されているデータを分析することで、移行先ストレージサブシステムが既に管理している特定パターンデータの他にも容量削減効果が高い特定パターンデータを発見することができれば、効率よく重複排除を行うことが可能となる。

40

## 【 0 0 7 6 】

50

図 2 1 には、特定パターンの分析プログラム 0 1 1 7 が用いるパターンカウントテーブル 2 0 0 1 の一例を示す。パターンカウントテーブル 2 0 0 1 は、特定パターンの分析プログラム 0 1 1 7 が移行元ストレージサブシステム内のコピー元通常ボリューム 0 1 0 4 のデータを分析した際、当該データのパターンと当該パターンの出現個数を登録しておくために用いられる。パターンID欄 2 0 1 1 には特定パターンの分析プログラム 0 1 1 7 がコピー元通常ボリューム 0 1 0 4 から読み出されたデータに付与したIDが記録される。各々のパターンIDに対応するデータパターンがパターンデータ欄 2 0 1 2 に記録され、同じデータパターンが移行元ストレージサブシステム内のコピー元通常ボリューム 0 1 0 4 の中からいくつ発見されたかをカウントした結果がカウント数欄 2 0 1 3 に記録される。パターンID 2 0 1 1 には、ハッシュ値などを利用することでパターンカウントテーブル 2 0 0 1 の容量を小さくすることができる。

10

**【 0 0 7 7 】**

図 2 2 は、特定パターンの分析プログラム 0 1 1 7 の処理フローの一例を示す。まず、特定パターンの分析プログラム 0 1 1 7 は、パターンカウントテーブル 2 0 0 1 を作成する（ステップ 2 2 0 7）。そして、特定パターンの分析プログラム 0 1 1 7 は、移行元ストレージサブシステムからコピー元通常ボリューム各々についてその構成情報を読み出す（ステップ 2 2 0 8）。コピー元通常ボリュームの構成情報とは、ボリュームのサイズなどの情報を指す。そして各コピー元通常ボリュームに対して、以下に示す処理を実行する。特定パターンの分析プログラム 0 1 1 7 は、移行元ボリュームのアドレス空間をセグメントに分割し、各セグメントについて順に移行元ストレージサブシステムへリード要求を出す（ステップ 2 2 1 1）。次に、特定パターンのリード要求の結果移行元ストレージサブシステムからリードされたデータをデータIDに変換する。具体的には、リードしたデータがパターンカウントテーブル 2 0 0 1 に存在するかどうかを確認し（ステップ 2 2 1 2）、データがパターンカウントテーブル 2 0 0 1 上のパターンデータ 2 0 1 2 に存在する場合は、該パターンデータに対応するカウント数 2 0 1 3 を 1 増加させる（ステップ 2 2 1 3）。一方ステップ 2 2 1 2 にて、リードしたデータがパターンカウントテーブル 2 0 0 1 上に存在しない場合、該データに新しいパターンIDを割り当てて、当該データと当該パターンIDとを共にパターンカウントテーブル 2 0 0 1 に新規に追加し、カウント数を 1 とする（ステップ 2 2 1 4）。

20

**【 0 0 7 8 】**

次に、ステップ 2 2 0 8 にて特定パターンの分析プログラム 0 1 1 7 は、当該コピー元通常ボリューム内の全てのセグメントについてデータをリードしたか否かを判定し（ステップ 2 2 1 5）、全てのセグメントについてリードを終了している場合にはステップ 2 2 1 6 に進み、データ移行の対象となっている全てのコピー元通常ボリュームについてステップ 2 2 0 8 からステップ 2 2 1 5 までの処理が終了しているか否かを判定し（ステップ 2 2 1 6）、終了している場合にはステップ 2 2 1 7 に進む。一方、ステップ 2 2 1 5 の判定の結果、コピー元通常ボリューム内の全てのセグメントについてリードを終了していない場合には、全てのセグメントについて終了するまでステップ 2 2 1 1 ~ 2 2 1 4 を残りのセグメントについて繰り返す。また、ステップ 2 2 1 6 の判定の結果、データ移行対象の全コピー元通常ボリュームについて読み込みが終了していない場合には、全てのコピー元通常ボリュームについてリードを終了するまでステップ 2 2 0 8 ~ 2 2 1 5 を残りのコピー元通常ボリュームについて繰り返す。

30

40

**【 0 0 7 9 】**

ステップ 2 2 1 7 では、特定パターンの分析プログラム 0 1 1 7 は、パターンカウントテーブル 2 0 0 1 を使って容量削減効果の高い特定パターンデータはどれかを分析する。容量削減効果の高い特定パターンデータを分析する方法には複数通り考えられるが、本明細書では閾値を使用する方法について示す。閾値はあるパターンのデータが特定パターンデータであるか否かを判断する場合に用いられる閾値である。即ち、特定パターンの分析プログラム 0 1 1 7 を実行することによって、ある所定のボリュームの中にある所定のパターンを有するセグメントデータが幾つ格納されているかが、カウント数 2 0 1 3 と

50

してパターンカウントテーブル2001に記録される。このカウント数が閾値を超えた場合には、当該データパターンが特定パターンとして認定され、移行先ストレージサブシステムが有する特定パターン管理テーブル0120に追加される。閾値を大きく設定した場合は、重複排除する特定パターン数が少なくなるため、容量削減効果が得られにくい。一方、閾値を小さく設定した場合は、容量削減効果を得られやすいが、識別する特定パターンの個数が大きくなり、管理情報量が大きくなる。

**【0080】**

特定パターンの分析が終わると、データ移行プログラム0113は、移行元ストレージサブシステム内のコピー元通常ボリューム0104からデータをリードし、該データをライトプログラム0111を使用して移行先ストレージサブシステム内の仮想ボリューム0102に書き込む。

10

**【0081】**

図23は、データ移行プログラム0113の処理フローの一例である。データ移行プログラム0113は、特定パターンの分析プログラム0117と同様に、移行元ストレージサブシステム内のデータ移行対象の各コピー元通常ボリュームについての構成情報を読み出し(ステップ2201)、当該コピー元通常ボリュームに対して以下に示す処理を実行する。尚、特定パターンの分析プログラム0117を先に実行している場合には、ステップ2201の処理は省略して、特定パターンの分析プログラム0117が読み出した構成情報を用いることとしても良い。

**【0082】**

20

まず、データ移行プログラム0113は、コピー元通常ボリュームと同じサイズの仮想ボリューム0102を移行先ストレージサブシステム内に作成し(ステップ2203)ここで作成された仮想ボリュームをデータの移行先ボリュームとする。次に、移行元通常ボリュームの先頭セグメントから順にデータを読み出すべく、移行元ストレージサブシステムにリード要求を出す(ステップ2205)。リード要求の結果移行元ストレージサブシステムから受信したデータを、データ移行プログラム0113は移行先ボリュームにライトするようライトプログラム0111に指示する(ステップ2206)。この結果、実施例1で図12~図15を用いて説明したのと同様の処理にて、ライトプログラム0111によりデータが移行先である仮想ボリュームに格納される。次に、ステップ2207に進み、ステップ2201にて構成情報を読み出したコピー元通常ボリューム内のすべてのセグメントに関してステップ2205、2206が終了しているかを判定する(ステップ2207)。当該コピー元通常ボリューム内にまだステップ2205及び2206の処理が実行されていないセグメントが存在する場合には、データ移行プログラム0113はステップ2205に戻って残りのセグメント分も終了するまでステップ2205、2206を繰り返す。当該コピー元通常ボリューム内の全セグメントについてステップ2205及び2206の処理が終了していればステップ2208に進む。ステップ2208では、データ移行対象となっている全てのコピー元通常ボリュームについてステップ2201~2207が終了したか否かを判定し、終了していない場合には、ステップ2201に戻って残りのコピー元通常ボリューム分も終了するまでステップ2201~2207を繰り返す。ステップ2208にてデータ移行対象の全てのコピー元ボリュームについてステップ2201~2207の処理が終了している場合には、データ移行プログラム0113は終了する。

30

40

**【0083】**

以上の方法によれば、送信先ストレージサブシステムにおいては、実施例1において説明したライトプログラム0111を用いてデータが移行先仮想ボリュームに書き込まれるため、特定パターンデータについてはデータの重複を排除することができ、移行先ストレージサブシステムの必要記憶容量を削減することができる。

**【0084】**

尚、特定パターンの分析プログラム0117を実行した後にデータ移行プログラム0113を実行すると、両方のプログラムによって合計2回データがコピー元通常ボリューム

50

から読み出されて移行元ストレージサブシステムから移行先ストレージサブシステムへコピーされることになり、ネットワークにおける通信負荷が増加する。そこでこれを防止するために、特定パターンの分析プログラム0117が図22のステップ2211でコピー元ストレージサブシステムから読み出したデータを、コピー先ストレージサブシステムのキャッシュメモリ0005又は通常ボリューム0104に一旦保存しておいても良い。そうすれば、データ移行プログラム0113は、図23のステップ2205で述べた様にコピー元ストレージサブシステムのコピー元通常ボリュームから再度データを読み出す代わりに、コピー先ストレージサブシステムのキャッシュメモリ0005若しくは通常ボリューム0104に格納されているデータを用いればよく、コピー元ストレージサブシステムからコピー先ストレージサブシステムへのデータ送信量を低減することができる。

10

**【実施例5】****【0085】**

本実施例では、実施例1で説明した重複排除機能に加え、状態が(a2)であるページに特定パターンデータ以外のデータ(非特定パターンデータと呼ぶ)も存在させて、ページ内の非特定パターンデータのセグメント数によってページの状態を細かく変遷させることにより、さらに効率的にストレージサブシステムの容量削減を行う方法について説明する。主に、実施例1と異なる点を説明する。

**【0086】**

図24は、本実施例における、ストレージサブシステム1111の論理的なシステム構成の一例を示したものである。図24に示すストレージサブシステム1111には、実施例1とは異なり、状態が(a2)であるページに非特定パターンデータを格納するために記憶領域を割り当てる際に用いられるセグメント用プール0107が存在する。セグメント用プール0107には、セグメント用ボリューム0106が複数存在する。セグメント用ボリューム0106は通常ボリューム0104やページ用ボリューム0103と同様、複数のHDD0006で構成されるRAIDグループ等のディスクグループ上に形成された記憶領域である。

20

**【0087】**

本実施例におけるストレージサブシステム1111は更に、ページ用プール0105を管理するプール管理テーブル0123とは別に、セグメント用プール0107を管理するためのセグメント用プール管理テーブル0125を有する。セグメント用プール管理テーブルの構成は、実施例1で図5を用いて説明したプール管理テーブル0123の構成と同様であるため図示はしない。尚、セグメント用プール管理テーブル0125においてプールLU#は当該セグメントに割り当てられる物理領域を有するセグメント用ボリューム0106の識別番号である。また、セグメントに割り当てられる各物理領域のサイズはセグメントのサイズと等しいため、領域#で識別されるセグメントに割り当てられる物理領域は、プールLU#で識別されるセグメント用ボリューム0106の開始アドレスで示されるアドレスからセグメントサイズ分の記憶領域である。

30

**【0088】**

本実施形態におけるストレージサブシステム1111では、セグメント管理テーブル0126の構成が実施例1におけるセグメント管理テーブル0122とは異なる。図25は、本実施例で使用するセグメント管理テーブル0126の一例を示した図である。実施例1で説明した図4のセグメント管理テーブル0122におけるフラグ欄0209には、特定パターンデータが書き込まれた旨の記録を行っている状態「0」と、セグメントに対してライト要求を未だ受け取っていない状態を「1」のいずれかの値が記録されたが、本実施例においてはこれらの状態に加え、セグメントに非特定パターンデータが格納されている状態「2」がセグメント管理テーブル0126のフラグ欄1262に記録され得る。更に本実施形態におけるセグメント管理テーブル0126は領域番号(領域#)欄1263を有する。領域#欄1263はフラグ欄1262に「2」が記録されている場合に用いられ、当該セグメントに割り当てられているセグメント用プール0107内の物理領域の領域番号が記録される。

40

50

## 【0089】

本実施形態においては、閾値管理テーブル0127も実施例1における閾値管理テーブル0124とは異なる。本実施形態で用いられる図26は閾値管理テーブル0127の一例を示す図である。本実施例では、ページが(a1)から(a2)、又は(a2)から(a1)に状態遷移する際に、非特定パターンデータ数に関する閾値、閾値が用いられる。状態が(a2)であるページにおいて、非特定パターンデータの数が増加して閾値を超えた場合、ページの状態を(a2)から(a1)へと変遷させる。また、状態が(a1)であるページにおいては、非特定パターンデータの数が減少して閾値を下回った場合、ページの状態を(a1)から(a2)へと変遷させる。閾値、閾値は閾値変更プログラム0116を利用して変更することが可能である。閾値を変更する処理フローについては、実施例1と同様であるため省略する。尚、本実施例における閾値は2以上であることとする。

10

## 【0090】

次に、本実施例におけるライトプログラム0128の処理フローについて、実施例1との相違点を示す。

## 【0091】

まず、ライト対象のページの状態が(a1)である場合に実行される、Write処理(1)が実施例1とは異なる。図27には、本実施形態におけるWrite処理(1)の処理フローの一例を示す。本処理フローでは、(a1)から(a2)へページの状態を遷移させるために、ページ内の非特定パターンデータ数を調べるステップが加わる点で実施例1におけるWrite処理(1)とは異なる。

20

## 【0092】

ステップ0701からステップ0706までは実施例1と同様の処理が行われるので説明を省略する。ステップ1301でライトプログラム0128は、非特定パターンデータの数をチェックして、閾値を下回っているかを確認する。非特定パターンデータの数は、ライト対象ページに格納されている全データをチェックすることで確認しても良いし、ページ管理テーブルに新たに非特定パターンカウント数を記録するための欄を設け、非特定パターンデータが書き込まれる度にこの欄に記録されているカウンタを1ずつ増加させることで確認しても良い。

## 【0093】

非特定パターンデータの数が閾値を下回った場合、まずライトプログラム0128は、該ページに対応するセグメント管理テーブル0126を作成する(ステップ1302)。そして、以下の手順に従って該ページ内のセグメントの情報を作成したセグメント管理テーブル0126に記録する。

30

## 【0094】

ライトプログラム0128は、特定パターン管理テーブル0120を参照して、セグメント内のデータが特定パターンのいずれかと一致するかどうかをページ内のセグメントデータを順にリードして判定する(ステップ1303)。該セグメントに書き込まれたデータが特定パターンデータのいずれかと一致する場合、該セグメントに対応するフラグ欄1262に0を設定し、対応するIDをパターンID欄1264に記録する(ステップ1304)。

40

## 【0095】

ステップ1302にて、リードしたセグメントのデータが特定パターンのいずれとも一致しない場合、ライトプログラム0128はセグメント用プール0107から物理領域を確保し(ステップ1306)、確保した領域に該データを書き込む(ステップ1307)。そして、セグメント管理テーブル0126のフラグ欄1262に2を記録し、確保したセグメント用プール0107の領域#1263を記録する。特定パターンデータではないためパターンID1264にはnullを記録する(ステップ1308)。

## 【0096】

ステップ1311では、ページ内の全セグメントについてステップ1303からステッ

50

プ 1 3 0 8 までの処理が終了したか否かを判定し、終了していない場合にはステップ 1 3 0 3 に戻って残りのセグメントについて上述の処理を繰り返す。

【 0 0 9 7 】

以上のステップにより、ページの状態が ( a 1 ) から ( a 2 ) に変遷したため、ページ管理テーブル 0 1 2 1 の該ページの状態 0 2 0 4 は 0 から 1 に変更される ( ステップ 1 3 0 9 ) 。

【 0 0 9 8 】

最後に、該ページに割り当てられていたページ用プール 0 1 0 5 内の物理領域を解放する ( ステップ 1 3 1 0 ) 。

【 0 0 9 9 】

以上の図 2 7 のフローによれば、一旦物理領域が割り当てられていたページであっても、該ページ内に特定パターンデータが多く、非特定パターンが少ない状態になれば、非特定パターンデータのみ物理領域を割り当てるようにするため、容量削減効果の高い管理が可能となる。

【 0 1 0 0 】

本実施形態においては、ライト対象のページの状態が ( a 2 ) である場合にライトプログラム 0 1 2 8 によって実行される Write 処理 ( 2 ) も実施例 1 とは異なる。図 2 8 には、本実施形態における Write 処理 ( 2 ) の処理フローの一例を示す。

【 0 1 0 1 】

ライトデータが特定パターンデータのいずれかと一致する場合の処理フロー ( ステップ 0 8 0 1 からステップ 0 8 0 4 まで ) は、実施例 1 と同様であるため説明を省略する。

【 0 1 0 2 】

ステップ 0 8 0 1 においてライトデータが特定パターンのいずれとも一致しない場合、ライトプログラム 0 1 2 8 は、ライト対象のページ用のセグメント管理テーブル 0 1 2 6 を取得し、ライト対象のセグメントのフラグ欄 1 2 6 2 を 2 に設定する ( ステップ 1 4 0 4 ) 。そして、ステップ 1 4 0 5 に進み、ライトプログラム 0 1 2 8 はセグメント管理テーブル 0 1 2 6 のフラグ欄 1 2 6 2 を参照して、ライト対象のページにフラグが 2 であるセグメントがいくつ存在するかカウントする。更にライトプログラム 0 1 2 8 は閾値管理テーブル 0 1 2 7 を参照して該ページに対応する閾値 1 2 7 2 の値を取得し、先にカウントしたフラグが 2 であるセグメントの数 ( 即ち非特定パターンデータの数 ) が閾値の値を超えているか否かを判定する ( ステップ 1 4 0 5 ) 。ステップ 1 4 0 5 にて、非特定パターンデータの数が閾値の値を下回っている場合の処理は、ステップ 1 4 0 6 ~ 1 4 0 8 となるが、図 2 7 のステップ 1 3 0 6 ~ 1 3 0 8 と同様であるため説明を省略する。ステップ 1 4 0 5 にて、非特定パターンデータの数が閾値の値を超えている場合には、ライトプログラム 0 1 2 8 はステップ 1 4 1 0 に進み、ライト対象のページにページ用プール 0 1 0 5 から物理領域を割り当てる ( ステップ 1 4 1 0 ) 。そして、該ページに対応するセグメント管理テーブル 0 1 2 6 を参照し、フラグ 1 2 6 2 が 0 であるセグメントについては、該セグメントに対応するパターン ID 1 2 6 4 を参照し、パターン管理テーブル 0 1 2 1 を取得して対応する特定パターンデータを確保領域の該セグメントに格納する。フラグ 1 2 6 2 が 1 であるセグメントについては、確保した領域の該セグメントに対応する領域にフォーマットデータを格納する。フラグ 1 2 6 2 が 2 であった場合は、領域 # 1 2 6 3 を参照し、プール管理テーブル 0 1 2 3 から該セグメントに対応するプール L U # 1 2 3 4 と開始アドレス 1 2 3 5 を取得し、該アドレスに格納されたデータを確保領域の該セグメントにコピーする ( ステップ 1 4 1 1 ) 。そして、ライトプログラム 0 1 2 8 は、該ページに対応するセグメント管理テーブル 0 1 2 6 を解放し ( ステップ 1 4 1 2 ) 、該ページに対応するページ管理テーブル 0 1 2 1 の状態 0 2 0 4 を 1 から 0 に変更して終了する ( ステップ 1 4 1 3 ) 。

【 0 1 0 3 】

以上の図 2 8 のフローによれば、物理領域を割り当てていない領域に非特定パターンデータのライト要求を受けた場合であっても、非特定パターンデータ数が少ないうちはライ

10

20

30

40

50

ト要求を受けたページ全体に物理領域を割り当てず、非特定パターンデータにのみ物理領域を割り当てることによって、ストレージの容量削減効果を高くすることが可能となる。

【0104】

尚状態が(a3)であるページにライト要求を受けた場合のライトプログラム0128の処理フローは、実施例1で説明したWrite処理(3)と同様であるため説明を省略する。

【図面の簡単な説明】

【0105】

【図1】実施例1、実施例2及び実施例5におけるストレージシステムのハード構成の一例を示した説明図である。 10

【図2】ストレージサブシステムの論理的なシステム構成の一例を示した説明図である。

【図3】仮想ボリュームのアドレス空間の一例を示した説明図である。

【図4】仮想ボリューム管理情報の一例を示した説明図である。

【図5】プール管理テーブルの一例を示した説明図である。

【図6】特定パターン管理テーブルの一例を示した説明図である。

【図7】閾値管理テーブルの一例を示した説明図である。

【図8】仮想ボリューム作成プログラムの処理フローの一例を示した説明図である。

【図9】特定パターン設定プログラムの処理フローの一例を示した説明図である。

【図10】閾値変更プログラムの処理フローの一例を示した説明図である。

【図11】リードプログラムの処理フローの一例を示した説明図である。 20

【図12】ライトプログラムの処理フローの一例を示した説明図である。

【図13】Write処理(1)の処理フローの一例を示した説明図である。

【図14】Write処理(2)の処理フローの一例を示した説明図である。

【図15】Write処理(3)の処理フローの一例を示した説明図である。

【図16】重複排除プログラムの処理フローの一例を示した説明図である。

【図17】実施例3及び実施例4におけるストレージシステムのハード構成の一例を示した説明図である。

【図18】(a)パターン送信プログラムの処理フローの一例を示した説明図である。

(b)パターン受信プログラムの処理フローの一例を示した説明図である。

【図19】ストレージサブシステムに送信されるパケットの実装例を示した説明図である 30

【図20】(a)データ送信プログラムの処理フローの一例を示した説明図である。(

b)データ受信プログラムの処理フローの一例を示した説明図である。

【図21】パターンカウントテーブルの一例を示した説明図である。

【図22】特定パターンの分析プログラムの処理フローの一例を示した説明図である。

【図23】データ移行プログラムの処理フローの一例を示した説明図である。

【図24】実施例5におけるストレージサブシステムの論理的なシステム構成の一例を示した説明図である。

【図25】実施例5におけるセグメント管理テーブルの一例を示した説明図である。

【図26】実施例5における閾値管理テーブルの一例を示した説明図である。 40

【図27】実施例5におけるWrite処理(1)の処理フローの一例を示した説明図である。

【図28】実施例5におけるWrite処理(2)の処理フローの一例を示した説明図である。

【符号の説明】

【0106】

0000 ストレージサブシステム

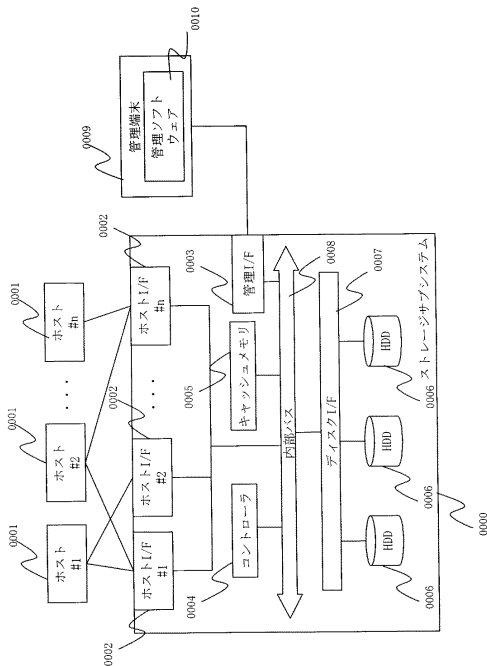
0110 仮想ボリューム作成プログラム

0111 ライトプログラム

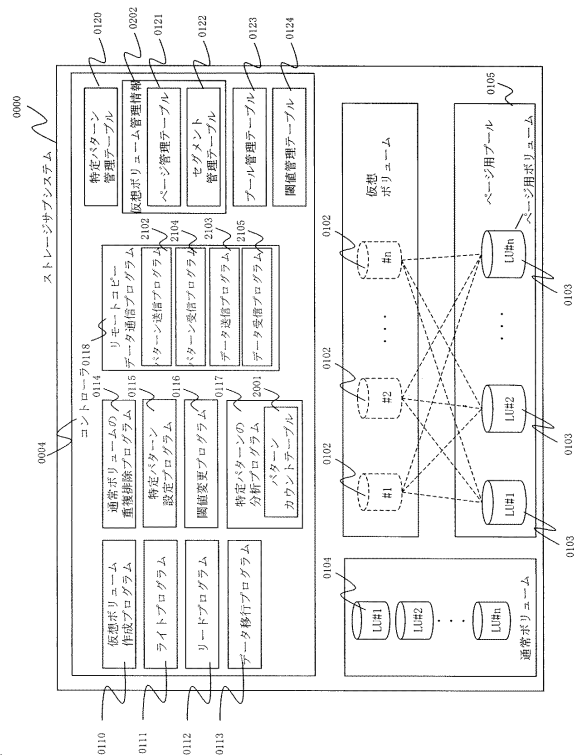
0112 リードプログラム 50

- 0 1 1 3 データ移行プログラム
- 0 1 1 4 通常ボリュームの重複排除プログラム
- 0 1 1 5 特定パターン設定プログラム
- 0 1 1 6 閾値変更プログラム
- 2 0 0 1 特定パターンの分析プログラム
- 0 1 1 8 リモートコピーデータ通信プログラム
- 0 1 2 0 特定パターン管理テーブル
- 0 2 0 2 仮想ボリューム管理情報システム本部
- 0 1 2 3 プール管理テーブル
- 0 1 2 4 閾値管理テーブル
- 0 1 2 1 ページ管理テーブル
- 0 1 2 2 セグメント管理テーブル
- 0 1 0 2 仮想ボリューム
- 0 1 0 3 ページ用ボリューム
- 0 1 0 4 通常ボリューム

【図1】



【図2】



【図3】

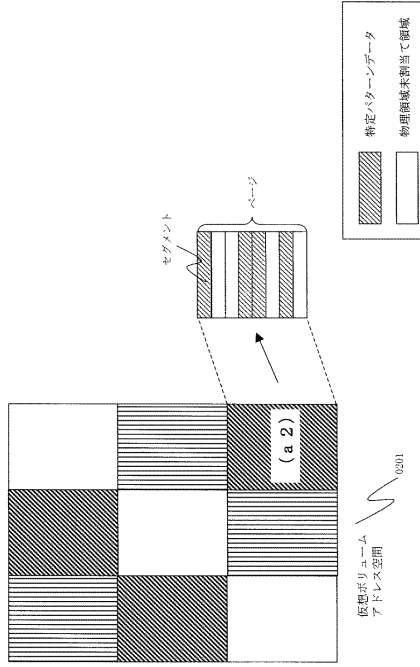


図3

【図4】

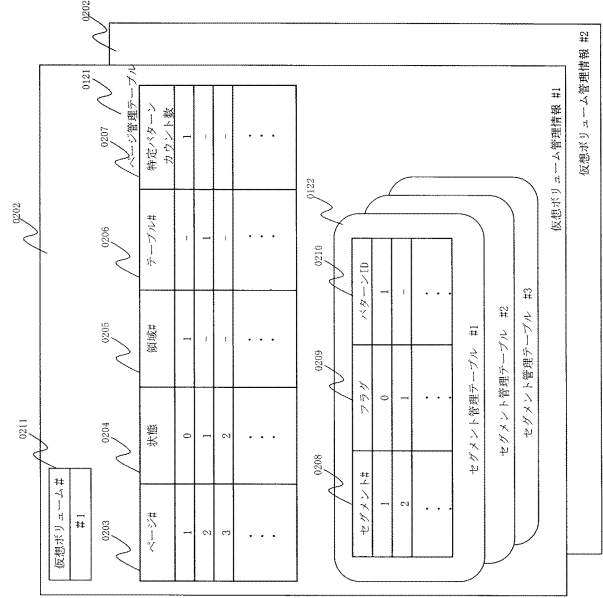


図4

【図5】

領域#	割当てフラグ	テーブル#	開始アドレス
1	0	1	AAA
2	1	1	BBB

図5

【図6】

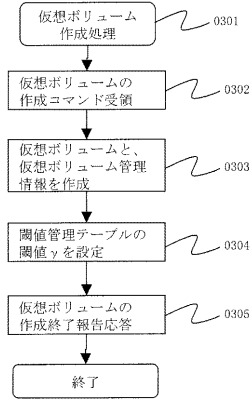
パターンID	パターンデータ
1	00 . . . . . 00
2	abc . . . . . xyz
.	.
.	.
n	ff . . . . . ff

【図7】

仮想ボリユーモア#	閾値γ
1	5
2	6
.	.
.	.

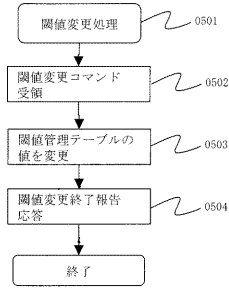
【図8】

図8



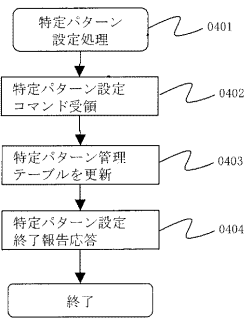
【図10】

図10



【図9】

図9



【図11】

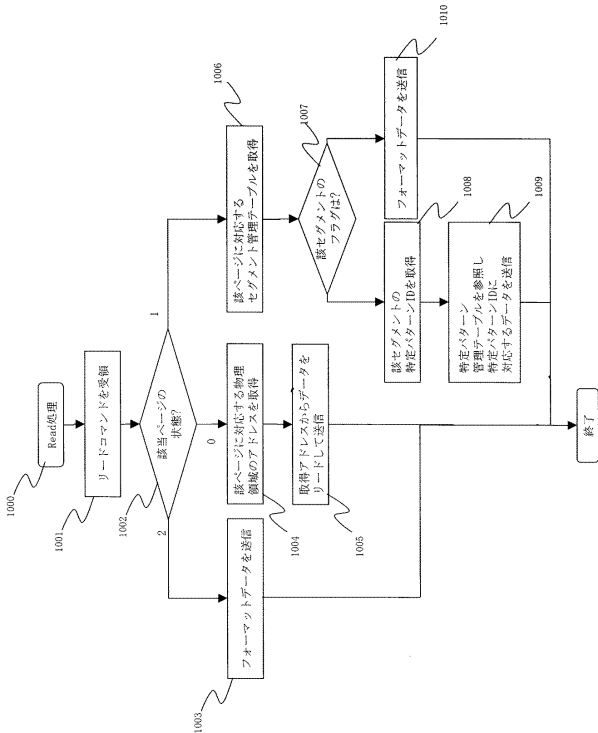
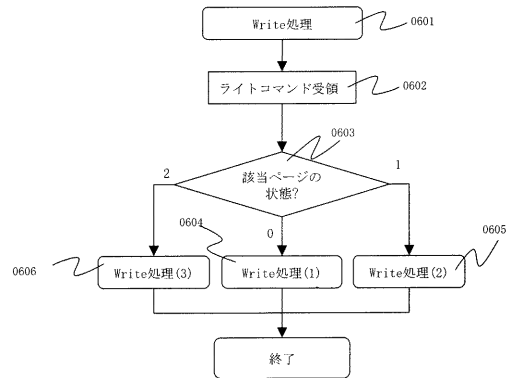


図11

【図12】

図12





【図 17】

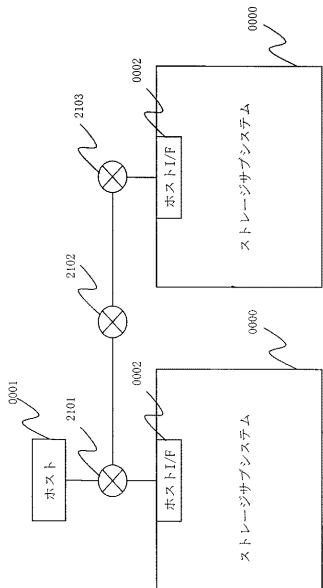


図 17

【図 19】

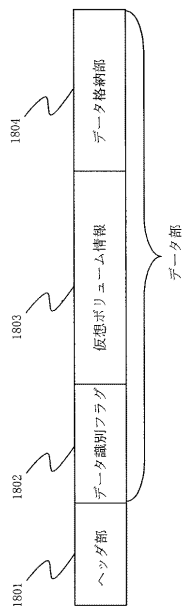


図 19

【図 18】

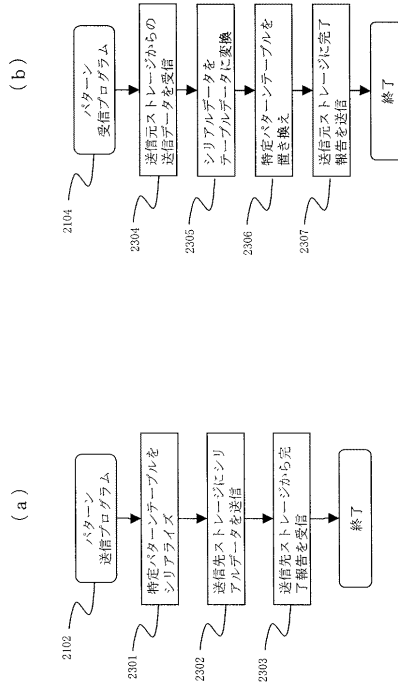


図 18

【図 20】

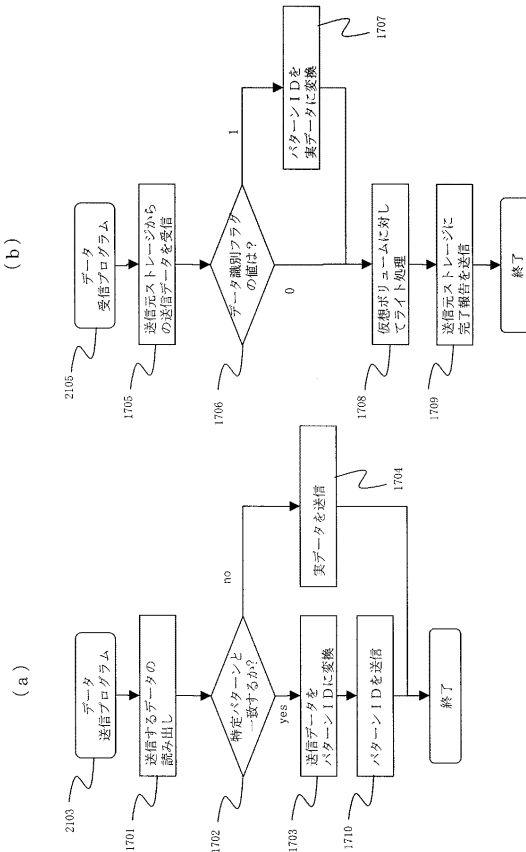


図 20

【図21】

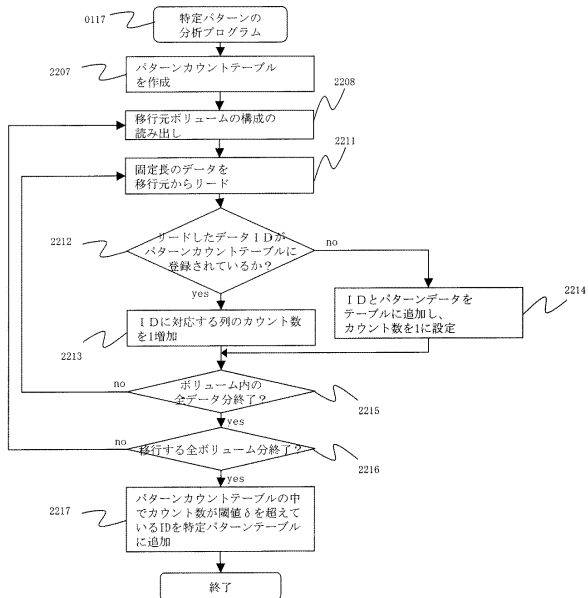
図21

パターンカウントテーブル

パターンID	パターンデータ	カウント数
1	00.....00	2
2	abc...xyz	3
⋮	⋮	⋮
n	ff.....ff	1

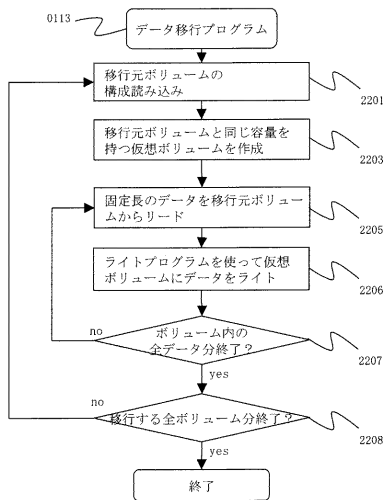
【図22】

図22



【図23】

図23



【図24】

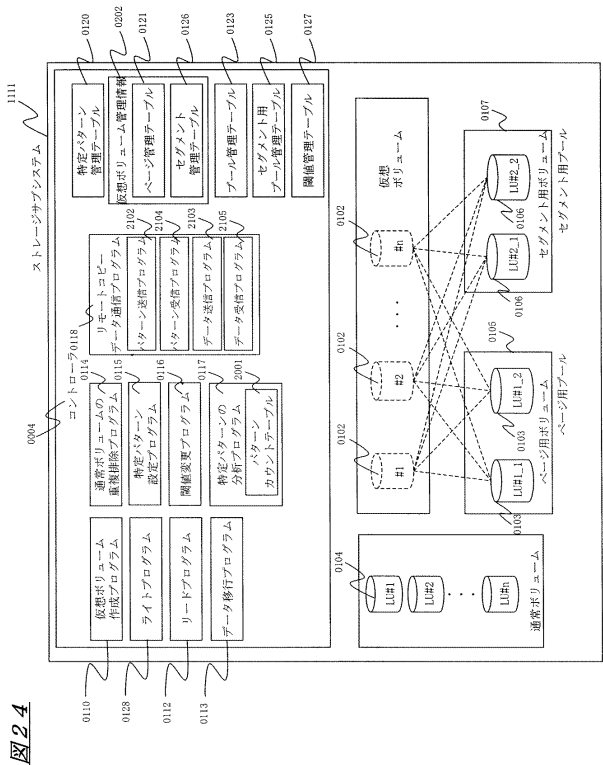


図24

【図25】

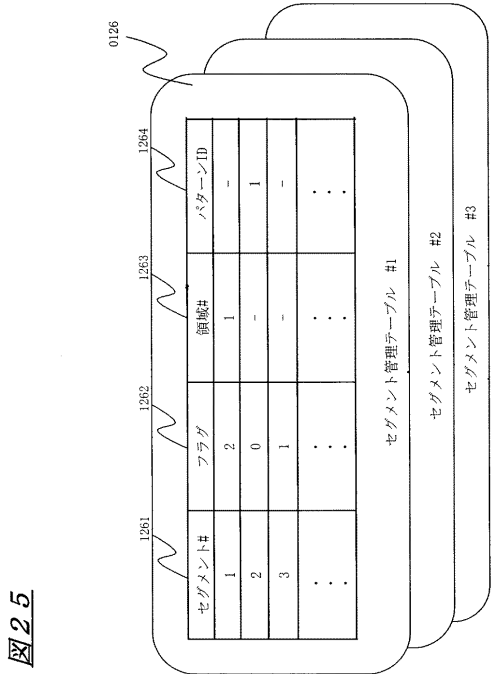


図25

【図26】

図26

仮想ボリューム#	閾値α	閾値β	閾値γ
1	5	2	5
2	10	3	6
...	...	...	...

閾値管理テーブル

【図27】

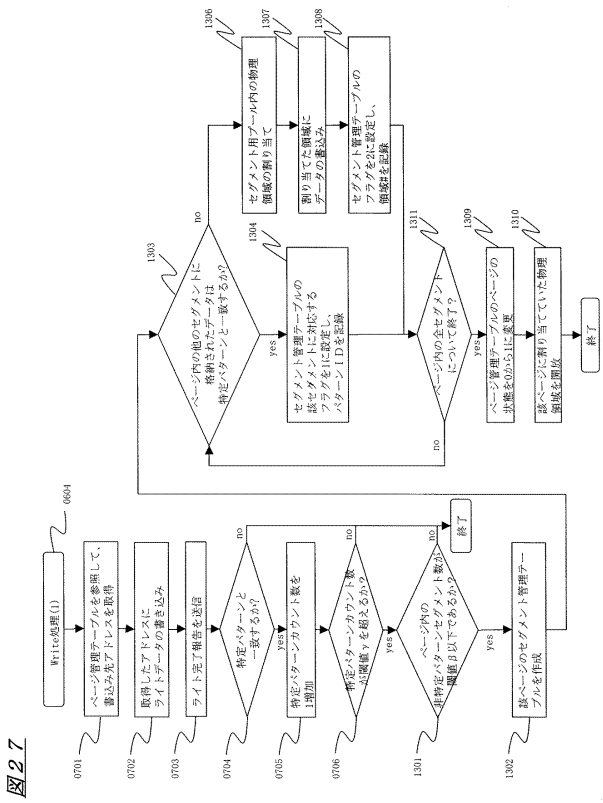


図27

【図28】

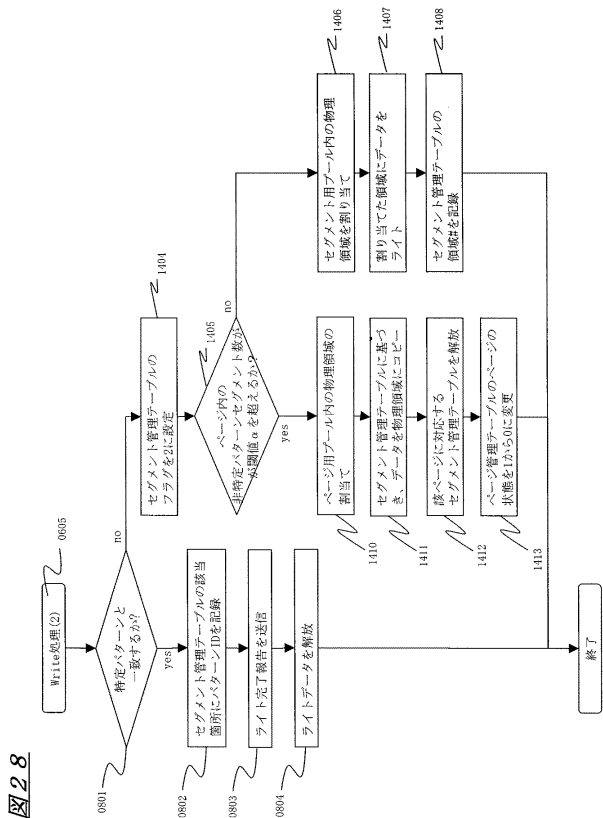


図28

---

フロントページの続き

- (72)発明者 児玉 昇司  
神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内
- (72)発明者 志賀 賢太  
神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内
- (72)発明者 江口 賢哲  
神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

審査官 田川 泰宏

- (56)参考文献 特開 2 0 0 5 - 1 3 5 1 1 6 ( J P , A )  
国際公開第 0 8 / 0 0 5 2 1 1 ( W O , A 1 )

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