(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第4863749号
(P4863749)

(45)発行日　平成24年1月25日(2012.1.25)　(24)登録日　平成23年11月18日(2011.11.18)

(51)Int.Cl.		F I		
G06F 12/00	(2006.01)	G06F	12/00	542K
G06F 12/16	(2006.01)	G06F	12/00	514E
G06F 3/08	(2006.01)	G06F	12/16	310A
G06F 3/06	(2006.01)	G06F	3/08	H
		G06F	3/06	301J

請求項の数 23　(全 32 頁)　最終頁に続く

(21)出願番号　特願2006-92217(P2006-92217)
(22)出願日　平成18年3月29日(2006.3.29)
(65)公開番号　特開2007-265265(P2007-265265A)
(43)公開日　平成19年10月11日(2007.10.11)
審査請求日　平成21年1月8日(2009.1.8)

(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　100064414
弁理士　磯野　道造
(72)発明者　田中　勝也
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
(72)発明者　島田　健太郎
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

審査官　桜井　茂行

最終頁に続く

(54)【発明の名称】フラッシュメモリを用いた記憶装置、その消去回数平準化方法、及び消去回数平準化プログラム

(57)【特許請求の範囲】

【請求項１】

記憶装置コントローラと、記憶媒体として前記記憶装置コントローラに接続される複数のフラッシュメモリ・モジュールとを有する記憶装置であって、

前記フラッシュメモリ・モジュールは、１個以上のフラッシュメモリ・チップと、前記フラッシュメモリ・チップに属するブロックの消去回数を平準化するメモリコントローラとを有し、

前記記憶装置コントローラは、

前記複数のフラッシュメモリ・モジュールを含むウエアレベリング・グループを構成し、

前記ウエアレベリング・グループに属する前記複数のフラッシュメモリ・モジュール毎に、書き込みデータ容量に基づいてフラッシュメモリ・モジュール内の平均消去回数を算出し、

前記フラッシュメモリ・モジュール毎の平均消去回数が所定の条件を満たした場合に、前記ウエアレベリング・グループに属するフラッシュメモリ・モジュール間でデータを移動する

ことを特徴とするフラッシュメモリを用いた記憶装置。

【請求項２】

前記記憶装置コントローラは、

前記フラッシュメモリ・モジュール毎の前記平均消去回数の最大値と最小値の差が所定

値以上の場合に、前記平均消去回数が最大のフラッシュメモリ・モジュールに格納されたデータと前記平均消去回数が最小のフラッシュメモリ・モジュールに格納されたデータを入換える
ことを特徴とする請求項１に記載のフラッシュメモリを用いた記憶装置。
【請求項３】
前記記憶装置コントローラは、
前記ウエアレベリング・グループに属するフラッシュメモリ・モジュールにアクセスするための論理ページアドレスを前記記憶装置コントローラの内部で扱うための仮想ページアドレスに変換し、
前記データを移動した後に、前記仮想ページアドレスと前記論理ページアドレスの対応関係を変更する
ことを特徴とする請求項１に記載のフラッシュメモリを用いた記憶装置。
【請求項４】
前記記憶装置は、複数のウエアレベリング・グループを備え、
前記記憶装置コントローラは、前記複数のウエアレベリング・グループを含むＲＡＩＤグループを構成し、
前記ＲＡＩＤグループは、フラッシュメモリ・モジュール故障時に記録データを修復するための冗長情報を含む
ことを特徴とする請求項１に記載のフラッシュメモリを用いた記憶装置。
【請求項５】
前記ＲＡＩＤグループは、ＲＡＩＤレベル０、ＲＡＩＤレベル１、ＲＡＩＤレベル１＋０、ＲＡＩＤレベル３、ＲＡＩＤレベル５、及びＲＡＩＤレベル６のうちのいずれかひとつの論理グループであり、前記ＲＡＩＤグループを構成する全てのウエアレベリング・グループの容量を等しく設定した
ことを特徴とする請求項１に記載のフラッシュメモリを用いた記憶装置。
【請求項６】
前記フラッシュメモリ・モジュールのシステム運用時の実効的書込み速度と前記記憶装置の製品寿命との積をフラッシュメモリの書換え保証回数で割った商を第１の値とし、
前記フラッシュメモリ・モジュールの連続書込み速度と前記記憶装置の製品寿命との積をフラッシュメモリの書換え保証回数で割った商を第２の値とし、
前記ウエアレベリング・グループの容量を、第１の値以上、かつ、第２の値以下に設定した
ことを特徴とする請求項１又は請求項２に記載のフラッシュメモリを用いた記憶装置。
【請求項７】
前記ＲＡＩＤグループは、ＲＡＩＤレベル２又はＲＡＩＤレベル４の論理グループであり、ＲＡＩＤグループを構成するウエアレベリング・グループの内、冗長情報格納用のウエアレベリング・グループ数をｍ、データ格納用の第１の論理グループ数をｎとするとき、冗長情報格納用のウエアレベリング・グループ容量を、データ格納用のウエアレベリング・グループ容量の１倍以上、かつ、ｎ／ｍ倍以下に設定した
ことを特徴とする請求項４に記載のフラッシュメモリを用いた記憶装置。
【請求項８】
前記論理ページアドレスでアクセスするメモリ領域は、前記仮想ページアドレスでアクセスするメモリ領域より大きく設定した
ことを特徴とする請求項３に記載のフラッシュメモリを用いた記憶装置。
【請求項９】
前記記憶装置コントローラは、前記ウエアレベリング・グループに属するフラッシュメモリ・モジュールにアクセスするための論理ページアドレスと前記記憶装置コントローラの内部で扱うための仮想ページアドレスの対応情報及び前記フラッシュメモリ・モジュール内のブロックの平均消去回数を記憶する
ことを特徴とする請求項１又は請求項４に記載のフラッシュメモリを用いた記憶装置。

【請求項１０】
前記記憶装置コントローラは、複数の前記ＲＡＩＤグループを構成する
ことを特徴とする請求項４に記載のフラッシュメモリを用いた記憶装置。
【請求項１１】
前記記憶装置コントローラは、ＲＡＩＤレベルの異なる複数の前記ＲＡＩＤグループを構成する
ことを特徴とする請求項４に記載のフラッシュメモリを用いた記憶装置。
【請求項１２】
前記記憶装置コントローラは、前記記憶装置が起動された時あるいは前記記憶装置に記録媒体が接続された時に、前記記録媒体がフラッシュメモリであるか否かを判定する
ことを特徴とする請求項１又は請求項４に記載のフラッシュメモリを用いた記憶装置。
【請求項１３】
１個以上のフラッシュメモリ・チップと、前記フラッシュメモリ・チップに属するブロックの消去回数を平準化するためのメモリコントローラとを有するフラッシュメモリ・モジュールと、
複数の前記フラッシュメモリ・モジュールを組み合わせてウエアレベリング・グループを構成し、前記ウエアレベリング・グループに属するフラッシュメモリ・モジュールをアクセスするための論理ページアドレスを前記記憶装置コントローラの内部で扱うための仮想ページアドレスに変換し、前記ウエアレベリング・グループを複数組み合わせてＲＡＩＤグループを構成する記憶装置コントローラと
を備えたフラッシュメモリを用いた記憶装置の消去回数平準化方法であって、
前記記憶装置コントローラは、
前記フラッシュメモリ・モジュール内の所定メモリ空間毎の書込みデータ容量を計数管理するステップと、
前記フラッシュメモリ・モジュール毎に延べ書込みデータ容量を前記フラッシュメモリ・モジュール容量で割った平均消去回数を算出するステップと、
前記平均消去回数の最大値と最小値の消去回数差が所定値以上であるか否かを判定する第１の判定ステップと、
前記第１の判定ステップで、消去回数差が所定値以上の場合に、消去回数差が最大の前記フラッシュメモリ・モジュール間で、書込みデータ容量最大メモリ空間と書込みデータ容量最小メモリ空間との間でデータの入換え及び前記論理ページアドレスと前記仮想ページアドレスの対応情報を変更するステップとを含んで実行する
ことを特徴とするフラッシュメモリを用いた記憶装置の消去回数平準化方法。
【請求項１４】
前記記憶装置コントローラは、前記変更するステップを実行後、データの入換えをしていないフラッシュメモリ・モジュールが複数あるか否かを判定する第２の判定ステップと、
第２の判定ステップで、データの入換えをしていないフラッシュメモリ・モジュールが複数ある場合、第１の判定ステップに戻るステップとを含んで実行する
ことを特徴とする請求項１３に記載のフラッシュメモリを用いた記憶装置の消去回数平準化方法。
【請求項１５】
前記記憶装置コントローラは、前記ウエアレベリング・グループ当たりの延べ書込みデータ容量が、所定の値に達した場合に、第１の判定ステップを実行する
ことを特徴とする請求項１３又は請求項１４に記載のフラッシュメモリを用いた記憶装置の消去回数平準化方法。
【請求項１６】
前記記憶装置コントローラは、前記論理ページアドレスのメモリ空間が、前記仮想ページアドレスのメモリ空間より大きく設定されており、
前記変更するステップが、論理ページアドレスのメモリ空間内でのデータの上書き操作

により、データの入換えを行うステップを含んで実行する
ことを特徴とする請求項１３から請求項１５のうちのいずれか１項に記載のフラッシュメモリを用いた記憶装置の消去回数平準化方法。
【請求項１７】
書込みデータ容量を計数管理した前記フラッシュメモリ・モジュール内の所定メモリ空間毎に、前記メモリコントローラの１回当たりのデータ転送容量に同じ大きさの空き領域が設けられている
ことを特徴とする請求項１３に記載のフラッシュメモリを用いた記憶装置の消去回数平準化方法。
【請求項１８】
書込みデータ容量を計数管理した前記フラッシュメモリ・モジュール内の所定メモリ空間と同じ大きさの空き領域が前記フラッシュメモリ・モジュール内に設けられている
ことを特徴とする請求項１３に記載のフラッシュメモリを用いた記憶装置の消去回数平準化方法。
【請求項１９】
前記記憶装置コントローラは、前記フラッシュメモリ・モジュールを新規のフラッシュメモリ・モジュールへ交換した場合、交換前のフラッシュメモリ・モジュールにおいて計数管理した所定メモリ空間毎の書込みデータ容量を、交換後のフラッシュメモリ・モジュールの所定メモリ空間毎の書込みデータ容量とするステップを含んで実行する
ことを特徴とする請求項１３に記載のフラッシュメモリを用いた記憶装置の消去回数平準化方法。
【請求項２０】
１個以上のフラッシュメモリ・チップと、前記フラッシュメモリ・チップに属するブロックの消去回数を平準化するためのメモリコントローラとを有するフラッシュメモリ・モジュールと、
複数の前記フラッシュメモリ・モジュールを組み合わせてウエアレベリング・グループを構成し、前記ウエアレベリング・グループに属するフラッシュメモリ・モジュールをアクセスするための論理ページアドレスを前記記憶装置コントローラの内部で扱うための仮想ページアドレスに変換し、前記ウエアレベリング・グループを複数組み合わせてＲＡＩＤグループを構成する記憶装置コントローラと
を備えたフラッシュメモリを用いた記憶装置の消去回数平準化方法であって、
前記記憶装置コントローラは、
前記フラッシュメモリ・モジュール内の所定メモリ空間毎の書込みデータ容量を計数管理するステップと、
前記フラッシュメモリ・モジュール毎に、所定時の前回の平均消去回数である第１の平均消去回数と該所定時からの延べ書込みデータ容量を前記フラッシュメモリ・モジュール容量で割った第２の平均消去回数とを加算して平均消去回数を算出するステップと、
前記平均消去回数の最大値と最小値の消去回数差が所定値以上であるか否かを判定する第１の判定ステップと、
前記第１の判定ステップで、消去回数差が所定値以上の場合に、消去回数差が最大の前記フラッシュメモリ・モジュール間で、書込みデータ容量最大メモリ空間と書込みデータ容量最小メモリ空間との間でデータの入換え及び前記論理ページアドレスと前記仮想ページアドレスの対応情報を変更するステップとを含んで実行する
ことを特徴とするフラッシュメモリを用いた記憶装置の消去回数平準化方法。
【請求項２１】
前記記憶装置コントローラは、前記第１の平均消去回数を、第２の平均消去回数により置換するステップを含んで実行する
ことを特徴とする請求項２０に記載のフラッシュメモリを用いた記憶装置の消去回数平準化方法。
【請求項２２】

１個以上のフラッシュメモリ・チップと、前記フラッシュメモリ・チップに属するブロックの消去回数を平準化するためのメモリコントローラとを有するフラッシュメモリ・モジュールと、

複数の前記フラッシュメモリ・モジュールを組み合わせてウエアレベリンググループであるウエアレベリング・グループを構成し、前記ウエアレベリング・グループに属するフラッシュメモリ・モジュールをアクセスするための論理ページアドレスを前記記憶装置コントローラの内部で扱うための仮想ページアドレスに変換し、前記ウエアレベリング・グループを複数組み合わせてＲＡＩＤグループを構成する記憶装置コントローラと

を備えたフラッシュメモリを用いた記憶装置の消去回数平準化プログラムであって、

コンピュータに、

前記フラッシュメモリ・モジュール内の所定メモリ空間毎の書込みデータ容量を計数管理する処理と、

前記フラッシュメモリ・モジュール毎に延べ書込みデータ容量を前記フラッシュメモリ・モジュール容量で割った平均消去回数を算出する処理と、

前記平均消去回数の最大値と最小値の消去回数差が所定値以上であるか否かを判定する第１の判定処理と、

前記第１の判定処理で、消去回数差が所定値以上の場合に、消去回数差が最大の前記フラッシュメモリ・モジュール間で、書込みデータ容量最大メモリ空間と書込みデータ容量最小メモリ空間との間でデータの入換え及び前記論理ページアドレスと前記仮想ページアドレスの対応情報を変更する処理とを

実行させるためのフラッシュメモリを用いた記憶装置の消去回数平準化プログラム。

【請求項２３】

前記コンピュータに、

前記変更する処理を実行後、データの入換えをしていないフラッシュメモリ・モジュールが複数あるか否かを判定する第２の判定処理と、

前記第２の判定処理で、データの入換えをしていないフラッシュメモリ・モジュールが複数ある場合、第１の判定する処理に戻る処理とを

実行させるための請求項２２に記載のフラッシュメモリを用いた記憶装置の消去回数平準化プログラム。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、複数のフラッシュメモリ・モジュールに渡った消去回数の平準化が可能なフラッシュメモリを用いた記憶装置、その消去回数平準化方法、及び消去回数平準化プログラムに関する。

【背景技術】

【０００２】

記憶装置（以下、ストレージ装置という。）は、一般的に、ランダムアクセス可能な不揮発性記憶媒体を備える。ランダムアクセス可能な不揮発性記憶媒体は、例えば、磁気ディスク、光ディスク等である。また、現在、主流ストレージ装置は、小型ディスクドライブを多数備える。

また、半導体技術の進歩に伴って、一括消去可能な不揮発性半導体メモリが開発されている。一括消去可能な不揮発性半導体メモリは、例えば、フラッシュメモリである。フラッシュメモリは、リードオンリメモリ（ＲＯＭ）のように不揮発性でありながら、リードだけでなく、ランダムアクセスメモリ（ＲＡＭ）のようにライトも可能な半導体メモリである。フラッシュメモリを記憶媒体とするストレージ装置は、小型ディスクドライブを多数備えるストレージ装置に比べ、寿命、省電力、アクセス時間等に優れている。

【０００３】

ここで、フラッシュメモリについて説明する。フラッシュメモリは、特性上、データを直接書換えることができない。つまり、フラッシュメモリは、記憶しているデータを書換

える場合、記憶している有効なデータを退避させる。次に、記憶しているデータをブロック単位で消去する。そして、データを消去したブロックにデータを書き込む。なお、ブロックは、データを一括消去する単位の記憶領域である。

【０００４】

フラッシュメモリは、例えば、データを消去した場合には、「１」に、すべて設定される。このため、データの書換え時に、ビット変換して「１」を「０」に書換えることはできる。しかし、データの消去をしないで、直接的に「０」を「１」に書換えることができない。そこで、フラッシュメモリは、データの書換え時に、一旦、ブロックの全体を消去する。このように、フラッシュメモリにおけるデータの書換えは、ブロックの消去が伴う。

【０００５】

フラッシュメモリは、ブロック消去の回数には制限がある。例えば、ブロック当たり１０万回までの消去回数が保証されている。データの書換えが集中して消去回数が増大したブロックは、データ消去ができなくなり使用不能となる問題点がある。そのため、フラッシュメモリを記憶媒体として使用するストレージ装置では、特定ブロックに消去が集中しないような消去回数平準化処理が必須となる。

【０００６】

特許文献１には、消去回数平準化処理方法についての記載がある。ストレージ装置は、ホストコンピュータとフラッシュメモリ間のブロック対応関係に柔軟性を与え、ホストコンピュータがアクセスする論理ブロックによってフラッシュメモリの物理ブロックが一方的に決定されることのないようなマッピング制御方式を導入した。上記従来のストレージ装置は、ホストコンピュータがアクセスする論理ブロック毎の書込み回数と、ストレージ装置が消去する物理ブロック毎の消去回数を計数管理する。そして、書込み回数の多い論理ブロックで、かつ、消去回数の多い物理ブロックと、書込み回数が少ない論理ブロックで、かつ、消去回数の少ない物理ブロックが生じた場合に、書込み回数が多い論理ブロックには消去回数の少ない物理ブロックを対応させ、書込み回数の少ない論理ブロックには消去回数の多い物理ブロックが対応するように、マッピングを変更する。

【特許文献１】特開平８－１６４８２号公報（段落００３８－００４６、図１０）

【発明の開示】

【発明が解決しようとする課題】

【０００７】

一般的に、フラッシュメモリ・モジュール（以下、ＰＤＥＶという。）は、メモリコントローラと複数のフラッシュメモリ・チップで構成され、メモリコントローラがフラッシュメモリ・チップに対して上記従来技術と同様な消去回数平準化処理を行う。大規模ストレージ装置では、記憶媒体として多数のフラッシュメモリ・モジュールを接続して大容量化することが考えられる。このとき、各フラッシュメモリ・モジュール内ではメモリコントローラにより消去回数平準化が行われる。しかし、特定のフラッシュメモリ・モジュールのみにデータ書換えが集中した場合に、書換えが集中したフラッシュメモリ・モジュールは消去回数が増大し、書換え寿命に早く達する。特定モジュールの消去回数増大を防止するためには、複数のモジュールに渡る消去回数の平準化が必要となる。

【０００８】

また、上記従来技術の消去回数平準化方法を、多数のフラッシュメモリ・モジュールを接続したストレージ装置へ適用しようとした場合、フラッシュメモリ・モジュール内のメモリコントローラが、フラッシュメモリ・チップの物理ブロックを隠蔽しているため、ストレージ装置内の上位の記憶制御部（以下、記憶装置コントローラという。）が、物理ブロック毎の消去回数を計数管理することができない問題がある。

【０００９】

さらにまた、フラッシュメモリ・モジュール内のメモリコントローラを使用しない、つまりモジュール毎に消去回数を平準化せずに、ストレージ装置全体で従来技術の消去回数平準化方法を適用する場合は、ストレージコントローラが膨大な数の物理ブロックの消去

回数を一元管理しなければならないため管理負荷が大きく、ストレージ装置の性能低下の要因となるおそれがある。

【００１０】

本発明は、前記の問題点に鑑みてなされたものであり、フラッシュメモリ物理ブロック関連情報を使用せずに、複数のフラッシュメモリ・モジュールに渡った消去回数の平準化が可能なフラッシュメモリを用いた記憶装置、その消去回数平準化方法、及び消去回数平準化プログラムを提供することを目的とする。

【課題を解決するための手段】

【００１１】

本発明によるフラッシュメモリを用いた記憶装置は、記憶装置コントローラと、記憶媒体として記憶装置コントローラに接続される複数のフラッシュメモリ・モジュールとを有する記憶装置であって、フラッシュメモリ・モジュールは、１個以上のフラッシュメモリ・チップと、フラッシュメモリ・チップに属するブロックの消去回数を平準化するメモリコントローラとを有し、記憶装置コントローラは、複数のフラッシュメモリ・モジュールを含むウエアレベリング・グループを構成し、ウエアレベリング・グループに属する複数のフラッシュメモリ・モジュール毎に、書き込みデータ容量に基づいてフラッシュメモリ・モジュール内の平均消去回数を算出し、フラッシュメモリ・モジュール毎の平均消去回数が所定の条件を満たした場合に、ウエアレベリング・グループに属するフラッシュメモリ・モジュール間でデータを移動することを特徴とする。

【００１２】

本発明によるフラッシュメモリを用いた記憶装置の消去回数平準化方法は、１個以上のフラッシュメモリ・チップと、フラッシュメモリ・チップに属するブロックの消去回数を平準化するためのメモリコントローラとを有するフラッシュメモリ・モジュールと、複数のフラッシュメモリ・モジュールを組み合わせて第１の論理グループを構成し、第１の論理グループに属するフラッシュメモリ・モジュールをアクセスするための第１のアドレスを記憶装置コントローラの内部で扱うための第２のアドレスに変換し、第１の論理グループを複数組み合わせて第２の論理グループを構成する記憶装置コントローラとを備えたフラッシュメモリを用いた記憶装置の消去回数平準化方法であって、記憶装置コントローラは、フラッシュメモリ・モジュール内の所定メモリ空間毎の書込みデータ容量を計数管理するステップと、フラッシュメモリ・モジュール毎に延べ書込みデータ容量をフラッシュメモリ・モジュール容量で割った平均消去回数を算出するステップと、平均消去回数の最大値と最小値の消去回数差が所定値以上であるか否かを判定する第１の判定ステップと、第１の判定ステップで、消去回数差が所定値以上の場合に、消去回数差が最大のフラッシュメモリ・モジュール間で、書込みデータ容量最大メモリ空間と書込みデータ容量最小メモリ空間との間でデータの入換え及び第１のアドレスと第２のアドレスの対応情報を変更するステップとを含むことを特徴とする。

【発明の効果】

【００１３】

本発明によれば、複数のフラッシュメモリ・モジュールに渡って消去回数を平準化できるので、ストレージ装置の記憶媒体を長寿命化できる。

【発明を実施するための最良の形態】

【００１４】

以下、本発明の実施の形態について図面を参照して説明する。

《概要》

本発明によるフラッシュメモリを用いた記憶装置は、記憶装置コントローラと、記憶媒体として複数のフラッシュメモリ・モジュールとを有するフラッシュメモリを用いた記憶装置であって、フラッシュメモリ・モジュール（例えば、フラッシュメモリ・モジュールＰ００）は、１個以上のフラッシュメモリ・チップ（例えば、フラッシュメモリ・チップ４０５）と、フラッシュメモリ・チップに属するブロック（例えば、ブロック４０６）の消去回数を平準化するためのメモリコントローラ（例えば、メモリコントローラＭＣ）と

を有し、記憶装置コントローラ（例えば、ストレージコントローラＳＣ）は、複数のフラッシュメモリ・モジュールを組み合わせて第１の論理グループ（例えば、ウエアレベリング・グループＷ００）を構成し、第１の論理グループに属するフラッシュメモリ・モジュールにアクセスするための第１のアドレス（例えば、論理ページアドレス６００）を記憶装置コントローラの内部で扱うための第２のアドレス（例えば、仮想ページアドレス６０４）に変換し、第１の論理グループを複数組み合わせて第２の論理グループ（例えば、ＲＡＩＤグループ）を構成する。

【００１５】

図１は、本発明によるストレージ装置の実施の形態の構成を示すブロック図である。ストレージ装置１００は、ストレージコントローラＳＣ及びフラッシュメモリ・モジュールＰ００～Ｐ３５を備える。

【００１６】

ストレージコントローラＳＣは、チャネルアダプタＣＡ０、チャネルアダプタＣＡ１、キャッシュメモリＣＭ０、キャッシュメモリＣＭ１、ストレージアダプタＳＡ０、ストレージアダプタＳＡ１、相互接続網ＮＷ０、及び相互接続網ＮＷ１を備える。なお、チャネルアダプタＣＡ０、チャネルアダプタＣＡ１、キャッシュメモリＣＭ０、キャッシュメモリＣＭ１、ストレージアダプタＳＡ０、及びストレージアダプタＳＡ１は、二つずつを図示しているが、いくつ備えられていてもよい。

【００１７】

相互接続網ＮＷ０及び相互接続網ＮＷ１は、例えば、スイッチ等であり、ストレージコントローラＳＣを構成する装置を相互に接続する。具体的には、相互接続網ＮＷ０及び相互接続網ＮＷ１は、チャネルアダプタＣＡ０、キャッシュメモリＣＭ０、及びストレージアダプタＳＡ０を相互に接続する。同様に、相互接続網ＮＷ０及び相互接続網ＮＷ１は、チャネルアダプタＣＡ１、キャッシュメモリＣＭ１、及びストレージアダプタＳＡ１を相互に接続する。

【００１８】

チャネルアダプタＣＡ０は、図２で後記するが、チャネルＣ００、チャネルＣ０１、チャネルＣ０２、チャネルＣ０３を介して、外部の上位装置（図示省略）に接続されている。同様に、チャネルアダプタＣＡ１は、チャネルＣ１０、チャネルＣ１１、チャネルＣ１２、チャネルＣ１３を介して、外部の上位装置（図示省略）に接続されている。なお、上位装置は、本実施の形態のストレージ装置１００にデータを読み書きする計算機である。ストレージ装置１００は、上位装置や他のストレージ装置等と接続する場合、ファイバチャンネルスイッチ、ＦＣ－ＡＬ（Ｆｉｂｒｅ Ｃｈａｎｎｅｌ-Ａｒｂｉｔｒａｔｅｄ Ｌｏｏｐ）、ＳＡＳ（Ｓｅｒｉａｌ Ａｔｔａｃｈｅｄ ＳＣＳＩ） Ｅｘｐａｎｄｅｒ等を介して接続する。

【００１９】

キャッシュメモリＣＭ０は、チャネルアダプタＣＡ０及びストレージアダプタＳＡ０から受信したデータを一時的に記憶する。同様に、キャッシュメモリＣＭ１は、チャネルアダプタＣＡ１及びストレージアダプタＳＡ１から受信したデータを一時的に記憶する。

【００２０】

ストレージアダプタＳＡ０は、図３で後記するが、フラッシュメモリ・モジュールＰ００等に接続されている。具体的には、ストレージアダプタＳＡ０は、チャネルＤ００を介して、フラッシュメモリ・モジュールＰ００～Ｐ０５に接続されている。また、ストレージアダプタＳＡ０は、チャネルＤ０１を介して、フラッシュメモリ・モジュールＰ１０～Ｐ１５に接続されている。また、ストレージアダプタＳＡ０は、チャネルＤ０２を介して、フラッシュメモリ・モジュールＰ２０～Ｐ２５に接続されている。また、ストレージアダプタＳＡ０は、チャネルＤ０３を介して、フラッシュメモリ・モジュールＰ３０～Ｐ３５に接続されている。

【００２１】

同様に、ストレージアダプタＳＡ１は、フラッシュメモリ・モジュールＰ００等に接続

されている。具体的には、ストレージアダプタＳＡ１は、チャネルＤ１０を介して、フラッシュメモリ・モジュールＰ００～Ｐ０５に接続されている。また、ストレージアダプタＳＡ１は、チャネルＤ１１を介して、フラッシュメモリ・モジュールＰ１０～Ｐ１５に接続されている。また、ストレージアダプタＳＡ１は、チャネルＤ１２を介して、フラッシュメモリ・モジュールＰ２０～Ｐ２５に接続されている。また、ストレージアダプタＳＡ１は、チャネルＤ１３を介して、フラッシュメモリ・モジュールＰ３０～Ｐ３５に接続されている。なお、具体的には、ストレージアダプタとフラッシュメモリ・モジュールとは、ファイバチャンネルスイッチ、ＦＣ－ＡＬ、ＳＡＳ　Ｅｘｐａｎｄｅｒ等を介して接続されている。

【００２２】

チャネルアダプタＣＡ０及びチャネルアダプタＣＡ１、並びに、ストレージアダプタＳＡ０及びストレージアダプタＳＡ１は、保守端末ＳＶＰに接続されている。保守端末ＳＶＰは、ストレージ装置１００の管理者から入力された設定情報を、チャネルアダプタＣＡ０、チャネルアダプタＣＡ１及び／又はストレージアダプタＳＡ０、ストレージアダプタＳＡ１に送信する。なお、ストレージ装置１００は、ストレージアダプタＳＡ０及びチャネルアダプタＣＡ０に代わって、一つのアダプタを備えていてもよい。この場合、当該アダプタが、ストレージアダプタＳＡ０及びチャネルアダプタＣＡ０の処理を行う。

【００２３】

図２は、チャネルアダプタの構成を示すブロック図である。チャネルアダプタＣＡ０は、ホストチャネル・インターフェース２１、キャッシュメモリ・インターフェース２２、ネットワーク・インターフェース２３、プロセッサ２４、ローカルメモリ２５、及びプロセッサ周辺制御部２６を備える。

【００２４】

ホストチャネル・インターフェース２１は、チャネルＣ００、チャネルＣ０１、チャネルＣ０２、及びチャネルＣ０３を介して、外部の上位装置（図示省略）と接続するインターフェースである。また、ホストチャネル・インターフェース２１は、チャネルＣ００、チャネルＣ０１、チャネルＣ０２、及びチャネルＣ０３上のデータ転送プロトコルと、ストレージコントローラＳＣの内部のデータ転送プロトコルとを相互に変換する。

【００２５】

キャッシュメモリ・インターフェース２２は、相互接続網ＮＷ０及び相互接続網ＮＷ１と接続するインターフェースである。ネットワーク・インターフェース２３は、保守端末ＳＶＰと接続するインターフェースである。なお、ホストチャネル・インターフェース２１とキャッシュメモリ・インターフェース２２とは、信号線２７によって接続されている。

【００２６】

プロセッサ２４は、ローカルメモリ２５に記憶されているプログラムを実行することによって、各種処理を行う。具体的には、プロセッサ２４は、上位装置と、相互接続網ＮＷ０、及び相互接続網ＮＷ１との間のデータ転送を制御する。

【００２７】

ローカルメモリ２５は、プロセッサ２４によって実行されるプログラムを記憶する。また、ローカルメモリ２５は、プロセッサ２４によって参照されるテーブルを記憶する。なお、当該テーブルは、管理者によって設定又は変更される。

【００２８】

この場合、管理者は、テーブルの設定又はテーブルの変更に関する情報を、保守端末ＳＶＰに入力する。保守端末ＳＶＰは、入力された情報をネットワーク・インターフェース２３を介して、プロセッサ２４に送信する。プロセッサ２４は、受信した情報に基づいて、テーブルを作成又は変更する。そして、プロセッサ２４は、当該テーブルを、ローカルメモリ２５に格納する。

【００２９】

プロセッサ周辺制御部２６は、ホストチャネル・インターフェース２１、キャッシュメ

モリ・インターフェース２２、ネットワーク・インターフェース２３、プロセッサ２４、及びローカルメモリ２５間のデータ転送を制御する。プロセッサ周辺制御部２６は、例えば、チップセット等である。なお、チャネルアダプタＣＡ１も、チャネルアダプタＣＡ０と同一の構成である。よって、説明を省略する。

【００３０】

図３は、ストレージアダプタの構成を示すブロック図である。ストレージアダプタＳＡ０は、キャッシュメモリ・インターフェース３１、ストレージチャネル・インターフェース３２、ネットワーク・インターフェース３３、プロセッサ３４、ローカルメモリ３５、及びプロセッサ周辺制御部３６を備える。

【００３１】

キャッシュメモリ・インターフェース３１は、相互接続網ＮＷ０及び相互接続網ＮＷ１と接続するインターフェースである。ストレージチャネル・インターフェース３２は、チャネルＤ００、チャネルＤ０１、チャネルＤ０２、及びチャネルＤ０３と接続するインターフェースである。また、ストレージチャネル・インターフェース３２は、チャネルＤ００、チャネルＤ０１、チャネルＤ０２、及びチャネルＤ０３上のデータ転送プロトコルと、ストレージコントローラＳＣの内部のデータ転送プロトコルとを相互に変換する。なお、キャッシュメモリ・インターフェース３１とストレージチャネル・インターフェース３２とは、信号線３７によって接続されている。ネットワーク・インターフェース３３は、保守端末ＳＶＰと接続するインターフェースである。

【００３２】

プロセッサ３４は、ローカルメモリ３５に記憶されているプログラムを実行することによって、各種処理を行う。

【００３３】

ローカルメモリ３５は、プロセッサ３４によって実行されるプログラムを記憶する。また、ローカルメモリ３５は、プロセッサ３４によって参照されるテーブルを記憶する。なお、当該テーブルは、管理者によって設定又は変更される。

【００３４】

この場合、管理者は、テーブルの設定又はテーブルの変更に関する情報を、保守端末ＳＶＰに入力する。保守端末ＳＶＰは、入力された情報をネットワーク・インターフェース３３を介して、プロセッサ３４に送信する。プロセッサ３４は、受信した情報に基づいて、テーブルを作成又は変更する。そして、プロセッサ３４は、当該テーブルを、ローカルメモリ３５に格納する。

【００３５】

プロセッサ周辺制御部３６は、キャッシュメモリ・インターフェース３１、ストレージチャネル・インターフェース３２、ネットワーク・インターフェース３３、プロセッサ３４、及びローカルメモリ３５間のデータ転送を制御する。プロセッサ周辺制御部３６は、例えば、チップセット等である。なお、ストレージアダプタＳＡ１も、ストレージアダプタＳＡ０と同一の構成である。よって、説明を省略する。

【００３６】

図４は、フラッシュメモリ・モジュールの構成を示すブロック図である。フラッシュメモリ・モジュールＰ００は、メモリコントローラＭＣ及びフラッシュメモリＭＥＭを備える。フラッシュメモリＭＥＭはデータを記憶する。メモリコントローラＭＣは、フラッシュメモリＭＥＭに対してデータを読み書きあるいは消去する。

【００３７】

メモリコントローラＭＣは、プロセッサ（μＰ）４０１、インターフェース部（Ｉ／Ｆ）４０２、データ転送部（ＨＵＢ）４０３、メモリ（ＲＡＭ）４０４、及びメモリ（ＲＯＭ）４０７を備える。

【００３８】

フラッシュメモリＭＥＭは、複数のフラッシュメモリ・チップ４０５を備える。フラッシュメモリ・チップ４０５は、複数のブロック４０６を含み、データを記憶する。ブロッ

ク４０６は、図５で後記するが、メモリコントローラＭＣがデータを消去する単位である。

【００３９】

ブロック４０６は、複数のページを含む。ページは、図５で後記するが、メモリコントローラＭＣがデータを読み書きする単位である。なお、ページは、有効ページ、無効ページ、未使用ページ、又は不良ページのいずれかに分類される。有効ページは、有効なデータを記憶しているページである。無効ページは、無効なデータを記憶しているページである。未使用ページは、データを記憶していないページである。不良ページは、当該ページの記憶素子が壊れている等の理由によって、物理的に使用できないページである。

【００４０】

インターフェース部４０２は、チャネルＤ００を介して、ストレージコントローラＳＣ内のストレージアダプタＳＡ０に接続されている。また、インターフェース部４０２は、チャネルＤ１０を介して、ストレージコントローラＳＣ内のストレージアダプタＳＡ１に接続されている。

【００４１】

インターフェース部４０２は、ストレージアダプタＳＡ０及びストレージアダプタＳＡ１からの命令を受信する。ストレージアダプタＳＡ０及びストレージアダプタＳＡ１からの命令は、例えば、ＳＣＳＩコマンドである。

【００４２】

具体的には、インターフェース部４０２は、ストレージアダプタＳＡ０及びストレージアダプタＳＡ１からデータを受信する。そして、インターフェース部４０２は、受信したデータをメモリ４０４に格納する。また、インターフェース部４０２は、メモリ４０４に格納されているデータを、ストレージアダプタＳＡ０及びストレージアダプタＳＡ１へ送信する。

【００４３】

メモリ４０４は、例えば、ダイナミック型ランダムアクセスメモリであり、高速に読み書きできる。メモリ４０４は、インターフェース部４０２が送受信するデータを一時的に記憶する。また、メモリ４０７は不揮発性メモリであり、プロセッサ４０１によって実行されるプログラムを記憶する。当該プログラムは、プロセッサ４０１が実行可能となるように、ストレージ装置起動時にメモリ４０７からメモリ４０４へコピーされる。また、メモリ４０４は、プロセッサ４０１によって参照されるテーブルを記憶する。当該テーブルは、例えば、フラッシュメモリＭＥＭの論理ページアドレスと物理ページアドレスとの変換テーブルである。論理ページアドレスは、フラッシュメモリ・モジュール外から（例えばストレージアダプタＳＡ０から）、フラッシュメモリの読み書きする単位であるページをアクセスするためのアドレスである。物理ページアドレスは、メモリコントローラＭＣが、フラッシュメモリの読み書きする単位であるページをアクセスするためのアドレスである。

【００４４】

データ転送部４０３は、例えばスイッチであり、プロセッサ４０１、インターフェース部４０２、メモリ４０４、メモリ４０７及びフラッシュメモリＭＥＭを相互に接続し、それらの間のデータ転送を制御する。

【００４５】

プロセッサ４０１は、メモリ４０４に記憶されているプログラムを実行することによって、各種処理を行う。例えば、プロセッサ４０１は、メモリ４０４に記憶されているフラッシュメモリの論理ページアドレスとフラッシュメモリの物理ページアドレスとの変換テーブルを参照して、フラッシュメモリＭＥＭにデータを読み書きする。また、プロセッサ４０１は、フラッシュメモリ・モジュール内のブロック４０６に対して、リクラメーション処理（ブロック再生処理）及びウエアレベリング処理（消去回数平準化処理）を行う。

【００４６】

リクラメーション処理（ブロック再生処理）は、未使用ページが少なくなったブロック

を使用できるように、ブロック４０６内の無効ページを未使用ページに再生する処理である。すなわち、リクラメーション処理の対象となるブロック（対象ブロック）４０６内には、有効ページ、無効ページ、及び未使用ページが含まれ、多くの無効ページが存在しているとする。この場合に、未使用ページを増加させるには、無効ページを消去する必要がある。しかしながら、消去は、ページ単位ではなく、ブロック単位でしか消去できない。このため、有効ページを空きのあるブロックに複写し、その後、対象ブロックを消去して、ブロックを再生する処理が必要となる。具体的には、プロセッサ４０１は、リクラメーション処理の対象となるブロック（対象ブロック）４０６内の有効ページに記憶されているデータを、未使用ブロックへ複写する。そして、プロセッサ４０１は、データを複写した未使用ブロックの論理ブロック番号を、対象ブロックの論理ブロック番号に変更する。そして、対象ブロックのデータをすべて消去し、リクラメーション処理を完了する。

【００４７】

例えば、プロセッサ４０１がブロック４０６にデータを書き込むと、ブロック４０６内の未使用ページが少なくなる。そして、ブロック４０６内の未使用ページが足りなくなると、プロセッサ４０１は、当該ブロック４０６へデータを書き込めなくなる。そこで、プロセッサ４０１は、当該ブロック４０６をリクラメーションすることによって、無効ページを未使用ページに再生する。

【００４８】

また、ウエアレベリング処理（消去回数平準化処理）は、それぞれのブロック４０６のデータ消去回数を平準化する処理である。これによって、フラッシュメモリＭＥＭの寿命を長くできる。なぜなら、フラッシュメモリＭＥＭは、データの消去回数が多くなると、寿命になるからである。フラッシュメモリＭＥＭは、一般的に、約一万回から十万回のデータの消去が保証されている。なお、他のフラッシュメモリ・モジュールＰ０１～Ｐ３５も、フラッシュメモリ・モジュールＰ００と同一の構成である。よって、説明を省略する。

【００４９】

図５は、フラッシュメモリ・モジュールのブロックの構成を示す説明図である。フラッシュメモリ・モジュールＰ００のブロック４０６は、複数のページ５０１を含む。ブロック４０６は、一般的に、数十程度のページ５０１（例えば、３２ページ、６４ページ等）を含む。

【００５０】

ページ５０１は、メモリコントローラＭＣ等がデータを読み出しあるいは書き込む単位である。例えば、ＮＡＮＤ型フラッシュメモリでは、メモリコントローラＭＣ等は、２０～３０μｓ弱／ページの速度でデータを読み出し、０．２～０．３ｍｓ／ページの速度でデータを書き込む。また、メモリコントローラＭＣ等は、２～４ｍｓ／ブロックの速度でデータを消去する。

【００５１】

ページ５０１は、データ部５０２及び冗長部５０３を含む。ページ５０１は、例えば、５１２バイトのデータ部５０２及び１６バイトの冗長部５０３を含む。データ部５０２は、通常のデータを記憶する。

【００５２】

冗長部５０３は、当該ページ５０１の管理情報及びエラー訂正情報を記憶する。管理情報は、オフセットアドレス及びページステータスを含む。なお、オフセットアドレスは、当該ページ５０１が属するブロック４０６内における相対的なアドレスである。また、ページステータスは、当該ページ５０１が有効ページ、無効ページ、未使用ページ、又は処理中のページのいずれであるかを示す。エラー訂正情報は、当該ページ５０１のエラーを検出及び訂正するための情報であり、例えば、ハミングコードである。

【００５３】

図６は、本発明によるストレージ装置の実施の形態における論理グループの構成とアドレス変換の階層を示す説明図である。なお、図６のストレージ装置のハードウエア的な構

成は、図１のストレージ装置と同様であるが、簡単のため、フラッシュメモリ・モジュールＰ００～Ｐ３５と接続するストレージコントローラＳＣのチャネルはＤ００、チャネルＤ０１、チャネルＤ０２、及びチャネルＤ０３のみを示している（チャネルＤ１０、チャネルＤ１１、チャネルＤ１２、チャネルＤ１３は省略）。

【００５４】

本発明のストレージ装置１００では、同一チャネル上に接続したフラッシュメモリ・モジュールを組み合わせてウエアレベリング・グループ（ＷＤＥＶ）を構成する。例えば、チャネルＤ００上のフラッシュメモリ・モジュールＰ００、フラッシュメモリ・モジュールＰ０１、フラッシュメモリ・モジュールＰ０２、及びフラッシュメモリ・モジュールＰ０３により、ウエアレベリング・グループＷ００を構成する。同様に、チャネルＤ０１上のフラッシュメモリ・モジュールＰ１０、フラッシュメモリ・モジュールＰ１１、フラッシュメモリ・モジュールＰ１２、及びフラッシュメモリ・モジュールＰ１３により、ウエアレベリング・グループＷ１０を構成する。チャネルＤ０２上のフラッシュメモリ・モジュールＰ２０、フラッシュメモリ・モジュールＰ２１、フラッシュメモリ・モジュールＰ２２、及びフラッシュメモリ・モジュールＰ２３により、ウエアレベリング・グループＷ２０を構成する。チャネルＤ０３上のフラッシュメモリ・モジュールＰ３０、フラッシュメモリ・モジュールＰ３１、フラッシュメモリ・モジュールＰ３２、及びフラッシュメモリ・モジュールＰ３３により、ウエアレベリング・グループＷ３０を構成する。

【００５５】

各フラッシュメモリ・モジュールは、ストレージコントローラＳＣからモジュールごとに論理ページアドレスでアクセスすることができる。例えば、チャネルＤ００上のフラッシュメモリ・モジュールＰ００、フラッシュメモリ・モジュールＰ０１、フラッシュメモリ・モジュールＰ０２、及びフラッシュメモリ・モジュールＰ０３は、各モジュールの論理ページアドレス６００でアクセスする。同様に、チャネルＤ０１上のフラッシュメモリ・モジュールＰ１０、フラッシュメモリ・モジュールＰ１１、フラッシュメモリ・モジュールＰ１２、及びフラッシュメモリ・モジュールＰ１３は各モジュールの論理ページアドレス６０１でアクセスし、チャネルＤ０２上のフラッシュメモリ・モジュールＰ２０、フラッシュメモリ・モジュールＰ２１、フラッシュメモリ・モジュールＰ２２、及びフラッシュメモリ・モジュールＰ２３は各モジュールの論理ページアドレス６０２でアクセスし、チャネルＤ０３上のフラッシュメモリ・モジュールＰ３０、フラッシュメモリ・モジュールＰ３１、フラッシュメモリ・モジュールＰ３２、及びフラッシュメモリ・モジュールＰ３３は、各モジュールの論理ページアドレス６０３でアクセスする。

【００５６】

ストレージコントローラＳＣは、同じウエアレベリング・グループに属する複数のフラッシュメモリ・モジュールの論理ページアドレスを纏めて、仮想ページアドレスに変換する。例えば、ウエアレベリング・グループＷ００に属するフラッシュメモリ・モジュールＰ００～Ｐ０３の論理ページアドレス６００を纏めて、仮想ページアドレス６０４に変換する。同様に、ウエアレベリング・グループＷ１０に属するフラッシュメモリ・モジュールＰ１０～Ｐ１３の論理ページアドレス６０１を纏めて、仮想ページアドレス６０５に変換し、ウエアレベリング・グループＷ２０に属するフラッシュメモリ・モジュールＰ２０～Ｐ２３の論理ページアドレス６０２を纏めて、仮想ページアドレス６０６に変換し、ウエアレベリング・グループＷ３０に属するフラッシュメモリ・モジュールＰ３０～Ｐ３３の論理ページアドレス６０３を纏めて、仮想ページアドレス６０７に変換する。

【００５７】

このように、ストレージコントローラＳＣは、論理ページアドレスを仮想ページアドレスに変換している。このため、消去回数平準化のためフラッシュメモリ・モジュール間でデータ移動を実施されると、論理ページアドレスが変更されるが、上位系のストレージコントローラＳＣは、論理ページアドレスと仮想ページアドレスとのマッピングを変更できるので、矛盾無くデータにアクセスできるようになる。

【００５８】

本発明の実施の形態のストレージ装置１００では、複数のウエアレベリング・グループを組み合わせてＲＡＩＤグループ（ＶＤＥＶ）を構成する。図６では、４個のウエアレベリング・グループＷ００、ウエアレベリング・グループＷ１０、ウエアレベリング・グループＷ２０、及びウエアレベリング・グループＷ３０を組み合わせて、ＲＡＩＤグループＶ００を構成している。一つのＲＡＩＤグループを構成する各ウエアレベリング・グループの仮想ページアドレス空間における容量は、同じである。一つあるいは複数のＲＡＩＤグループ内の領域を組み合わせて、論理ボリウム６０８を構成する。論理ボリウム６０８は、ストレージコントローラＳＣが上位装置に対して見せる記憶空間である。

【００５９】

また、チャネルＤ００上のフラッシュメモリ・モジュールＰ０４及びフラッシュメモリ・モジュールＰ０５は、予備グループ（ＹＤＥＶ）Ｙ００を構成する。同様に、チャネルＤ０１上のフラッシュメモリ・モジュールＰ１４及びフラッシュメモリ・モジュールＰ１５は、予備グループＹ１０を、チャネルＤ０２上のフラッシュメモリ・モジュールＰ２４及びフラッシュメモリ・モジュールＰ２５は、予備グループＹ２０を、チャネルＤ０３上のフラッシュメモリ・モジュールＰ３４及びフラッシュメモリ・モジュールＰ３５は、予備グループＹ３０を構成する。モジュール交換の詳細は後記する。

【００６０】

図７は、本発明によるストレージ装置の実施の形態におけるＲＡＩＤグループの構成を示す説明図である。ＲＡＩＤグループ７２０は、ウエアレベリング・グループ７００、ウエアレベリング・グループ７０１、ウエアレベリング・グループ７０２、及びウエアレベリング・グループ７０３からなる、ＲＡＩＤレベル５のＲＡＩＤグループである。例えば、ウエアレベリング・グループ７００は、フラッシュメモリ・モジュール７３０とフラッシュメモリ・モジュール７３１からなる。なお、ＲＡＩＤは、機能によってＲＡＩＤレベル０、ＲＡＩＤレベル１のようにレベル分けされている。

【００６１】

ＲＡＩＤグループ７２１は、ウエアレベリング・グループ７０４及びウエアレベリング・グループ７０５からなる、ＲＡＩＤレベル１のＲＡＩＤグループである。同様にＲＡＩＤグループ７２２は、ウエアレベリング・グループ７０６及びウエアレベリング・グループ７０７からなる、ＲＡＩＤレベル１のＲＡＩＤグループである。

【００６２】

本発明のストレージ装置１００では、ＲＡＩＤレベル０、ＲＡＩＤレベル１、ＲＡＩＤレベル３、ＲＡＩＤレベル５、ＲＡＩＤレベル６、あるいはＲＡＩＤレベル１＋０の場合に、ＲＡＩＤグループを構成する各ウエアレベリング・グループの論理ページアドレス空間における容量を等しく設定する。その容量は、上限を（１）式で、下限を（２）式で与えられる計算値で設定する。すなわち、フラッシュメモリ・モジュールの連続書込み速度とシステム製品寿命との積をフラッシュメモリの書換え寿命（書換え保証回数）で割った商を第２の値（上限）とし、フラッシュメモリ・モジュールのシステム運用時の実効的書込み速度とシステム製品寿命との積をフラッシュメモリの書換え寿命（書換え保証回数）で割った商を第１の値（下限）とし、各ウエアレベリング・グループの論理ページアドレス空間における容量（第１の論理グループの容量）を、第１の値以上、かつ、第２の値以下に設定する。例えば、システム製品寿命は５年から１０年程度であり、書換え寿命は１万回から１０万回程度である。（２）式の実効的書込み速度とは、上位機装置からストレージ装置１００への書込みアクセス割合を考慮した場合の、実効的な書込み速度である。

【００６３】

【数１】

$$\text{ウエアレベリング・グループ容量値(上限)} = \frac{\text{モジュール当たり連続書込み速度} \times \text{システム製品寿命}}{\text{書換え寿命}} \quad \cdots (1)$$

$$\text{ウエアレベリング・グループ容量値(下限)} = \frac{\text{モジュール当たり実効的書込み速度} \times \text{システム製品寿命}}{\text{書換え寿命}} \quad \cdots (2)$$

【００６４】

上記（１）式及び（２）式で与えられる範囲内にウエアレベリング・グループの容量を設定する。ウエアレベリング・グループ内フラッシュメモリ・モジュール間で消去回数平準化を行うことにより、ストレージ装置１００のシステム製品寿命期間内でフラッシュメモリ・モジュールの書換え寿命が保証される。

【００６５】

ＲＡＩＤグループ７２３は、ウエアレベリング・グループ７０８、ウエアレベリング・グループ７０９、ウエアレベリング・グループ７１０、及びウエアレベリング・グループ７１１からなる、ＲＡＩＤレベル４のＲＡＩＤグループである。そしてウエアレベリング・グループ７０８、ウエアレベリング・グループ７０９、及びウエアレベリング・グループ７１０は、データ格納用のウエアレベリング・グループであり、ウエアレベリング・グループ７１１は、パリティ（冗長情報）格納用のウエアレベリング・グループである。パリティ格納用のウエアレベリング・グループは、他のデータ格納用のウエアレベリング・グループに対してデータの更新回数が多い。従ってＲＡＩＤレベル４では、ＲＡＩＤグループ内での消去回数平準化のため、パリティ格納用のウエアレベリング・グループの論理ページアドレス空間における容量値を、データ格納用ウエアレベリング・グループ容量値より大きく設定する。例えば、ＲＡＩＤグループを構成するウエアレベリング・グループの数がｎ個の場合、パリティ格納用のウエアレベリング・グループは、データ格納用のウエアレベリング・グループに対して、１倍以上、かつ、（ｎ－１）倍以下の論理ページアドレス空間を持つように設定する。

【００６６】

また、図示していないが、ＲＡＩＤレベル２の場合も冗長情報格納用のウエアレベリング・グループは、データ格納用のウエアレベリング・グループより更新回数が多い。ＲＡＩＤレベル２で、例えば１０Ｄ４Ｐならば、冗長情報格納用のウエアレベリング・グループの論理ページアドレス空間における容量値をデータ格納用ウエアレベリング・グループ容量値の１倍以上、かつ、１０／４＝２．５倍以下に設定する。また、２５Ｄ５Ｐならば、冗長情報格納用のウエアレベリング・グループの論理ページアドレス空間における容量値を、データ格納用ウエアレベリング・グループ容量値の１倍以上、かつ、２５／５＝５倍以下に設定する。

【００６７】

言い換えると、ＲＡＩＤレベル２又はＲＡＩＤレベル４の場合、データ格納用のウエアレベリング・グループ数をｎ、冗長情報格納用のウエアレベリング・グループ数をｍとすると、冗長情報格納用ウエアレベリング・グループは、データ格納用ウエアレベリング・グループに対して、１倍以上、かつ、ｎ／ｍ倍以下の論理ページアドレス空間容量を持つ

ように設定する。

【００６８】

このように、ストレージコントローラＳＣのＲＡＩＤグループは、ウエアレベリング・グループを組み合わせてＲＡＩＤグループを構成している。すなわち、ストレージコントローラＳＣの管理単位は、ウエアレベリング・グループを考慮してＲＡＩＤグループで管理しているので、各ウエアレベリング・グループ内の論理ページアドレスと仮想ページアドレスの対応に係わらず、各ウエアレベリング・グループの仮想ページアドレスは、独立することになる。これにより、ストレージコントローラＳＣは、複数、かつ、レベルの異なるＲＡＩＤグループを接続することができる。

【００６９】

図８は、フラッシュメモリ・モジュールとハードディスクドライブをストレージコントローラＳＣに接続した例を示す構成図である。フラッシュメモリ・モジュール８１０、フラッシュメモリ・モジュール８１１、及びフラッシュメモリ・モジュール８１２でウエアレベリング・グループ８３０を構成する。フラッシュメモリ・モジュール８１３、フラッシュメモリ・モジュール８１４、フラッシュメモリ・モジュール８１５でウエアレベリング・グループ８３１を構成する。ウエアレベリング・グループ８３０とウエアレベリング・グループ８３１を組み合わせて、ＲＡＩＤグループ８４０を構成する。

【００７０】

図６の場合と同様に、ストレージコントローラＳＣは、フラッシュメモリ・モジュール８１０、フラッシュメモリ・モジュール８１１、及びフラッシュメモリ・モジュール８１２をアクセスするため、論理ページアドレス８００を仮想ページアドレス８０２へ変換する。また、ストレージコントローラＳＣは、フラッシュメモリ・モジュール８１３、フラッシュメモリ・モジュール８１４、及びフラッシュメモリ・モジュール８１５をアクセスするため、論理ページアドレス８０１を仮想ページアドレス８０３へ変換する。

【００７１】

ハードディスクドライブ８２０及びハードディスクドライブ８２３を組み合わせてＲＡＩＤグループ８４１を構成する。同様に、ハードディスクドライブ８２１及びハードディスクドライブ８２４を組み合わせてＲＡＩＤグループ８４２を、ハードディスクドライブ８２２及びハードディスクドライブ８２５を組み合わせてＲＡＩＤグループ８４３を、それぞれ構成する。ストレージコントローラＳＣは、論理ブロックアドレス８０４あるいは論理ブロックアドレス８０５で各ハードディスクドライブをアクセスする。ハードディスクドライブからなるＲＡＩＤグループにおいては、消去回数平準化が不要なため、ウエアレベリング・グループを定義しない。フラッシュメモリ・モジュールからなるＲＡＩＤグループに対してのみ、ウエアレベリング・グループを定義し、かつ、論理ページアドレスと仮想ページアドレスの変換を行う。

【００７２】

ストレージコントローラＳＣは、装置起動時あるいは記憶媒体を接続した際に、記憶媒体がフラッシュメモリ・モジュールかハードディスクドライブかによって、アドレス変換の要否を判断及びＲＡＩＤグループの構成方法の決定等、制御方法を変更する。

【００７３】

ストレージコントローラＳＣは、フラッシュメモリ・モジュールからなるＲＡＩＤグループ８４０とハードディスクドライブからなるＲＡＩＤグループ８４１～８４３との一方あるいは両方の領域を組み合わせて、論理ボリウム８０８を構成する。フラッシュメモリ・モジュールからなる記憶領域とハードディスクドライブから記憶領域の使い分け例としては、読み出しアクセスが多く、あまり更新されないデータを主にフラッシュメモリ・モジュールに記憶し、更新する頻度が高いデータをハードディスクドライブに記憶することが考えられる。フラッシュメモリ・モジュールはハードディスクドライブに対してアクセス速度が高速なため、このようなデータのアクセス特性によって記憶領域を使い分けることにより、ストレージ装置が高性能化する。

【００７４】

次に、動作について図面を参照しながら詳細に説明する。

図９から図１４において、本発明の実施の形態のストレージ装置における消去回数平準化方法を説明する。本方法は、複数のフラッシュメモリ・モジュールに渡って消去回数を平準化する点が特徴である。

【００７５】

図９は、複数のフラッシュメモリ・モジュール間の消去回数平準化方法を示すフローチャートである。簡単のため、平準化の対象となるウエアレベリング・グループＷ００は、２個のフラッシュメモリ・モジュールＰ００とフラッシュメモリ・モジュールＰ０４からなる場合を考える。

【００７６】

図１０は、消去回数平準化に伴うデータ入換え前の、仮想ページアドレスと論理ページアドレスの変換テーブルを示す説明図である。

【００７７】

図１１は、消去回数平準化に伴うデータ入換え後の、仮想ページアドレスと論理ページアドレスの変換テーブルを示す説明図である。

【００７８】

図１０と図１１において、仮想ページアドレスと論理ページアドレスの対応関係と共に、オフセット値が示されている。本発明の実施の形態のストレージ装置では、論理ページアドレス空間の大きさ（データ長）を、対応する仮想ページアドレス空間の大きさ（データ長）より大きく設定する。そして、論理ページアドレス空間において、有効データが先頭アドレス側に書き込まれていて、末尾アドレス側に空き領域が存在する場合をオフセット値０、有効データが末尾アドレス側に書き込まれていて、先頭アドレス側に空き領域が存在する場合をオフセット値１と表示する。空き領域の大きさはフラッシュメモリのページのデータ部の整数倍（１倍以上）であり、フラッシュメモリ・モジュール内のメモリコントローラがフラッシュメモリに対して一度に書込み可能なデータの大きさとする。

【００７９】

図１２は、ストレージコントローラにおいて管理するフラッシュメモリ・モジュール毎の消去回数管理テーブルを示す説明図である。ストレージコントローラＳＣは、フラッシュメモリ・モジュール内のデータ入換えの単位となる領域ごとに、延べ書込み容量を記録する。（３）式のように、前回の平均消去回数とモジュール内論理ページアドレス空間に対する所定期間の延べ書込み容量をモジュールの論理ページアドレス空間容量で割った平均処理回数を加算することにより、そのモジュール内フラッシュメモリの平均消去回数を算出することができる。

【数２】

$$\text{平均消去回数} = \text{前回の値} + \frac{\sum \text{論理ページアドレス空間毎の所定期間の延べ書込み容量値}}{\text{モジュール容量値}} \quad \cdots (3)$$

【００８０】

図１２の管理テーブルでは、２個の平均消去回数値を記録している。一つは前回の平準化処理実行時の平均消去回数（ｆ００、ｆ０４）であり、もう一つは現在までの平均消去回数（ｅ００、ｅ０４）である。論理ページアドレス空間の所定領域毎に計数管理する延べ書込み容量値は、前回平準化処理実行時から現在までの容量値を記憶する。現在の平均消去回数値は、（３）式により自動的に計算される。このように期間を分けて延べ書込み容量を計数管理することにより、論理ページアドレス空間領域における最近のアクセス頻度を知ることができる。また、管理テーブルには、移動フラグが設けられており、データ入換え前の場合、移動フラグには０が設定されており、データ入換え後、移動フラグには１が設定される。なお、（３）式は、期間を分けて延べ書込み容量を計数管理した場合であるが、期間をわけずに通期での延べ書込み容量を計数管理した場合は、（４）式のように表現できる。（３）式と（４）式とは、平均消去回数の計算結果は同一となる。

【数３】

$$\text{平均消去回数} = \frac{\sum \text{論理ページアドレス空間毎の通期の延べ書込み容量値}}{\text{モジュール容量値}} \quad \cdots (4)$$

【００８１】
図１０又は図１１のアドレス変換テーブル及び図１２の平均消去回数管理テーブルは、電源断等の障害時やシステム起動時間外でも保持する必要がある。従って、ストレージコントローラＳＣは、各フラッシュメモリ・モジュールの所定領域に、各フラッシュメモリ・モジュールに関連するアドレス変換テーブル及び各モジュールの平均消去回数管理テーブル関連データを保存する。

【００８２】
図９において、ストレージコントローラＳＣは、イベントの発生時、例えば、ウエアレベリング・グループ（ＷＤＥＶ）当たりの延べ書込み容量が所定値に達するか、あるいは所定期間毎に、平均消去回数平準化処理の実行を開始する。また、ウエアレベリング・グループ内のフラッシュメモリ・モジュールの移動フラグを０に設定する（ステップＳ９０１）。

【００８３】
次に、ストレージコントローラＳＣは、図１２の平均消去回数管理テーブルから、移動フラグが０で、かつ、平均消去回数の最大値と最小値をチェックする（ステップＳ９０２）。

【００８４】
ストレージコントローラＳＣは、平均消去回数の最大値と最小値の消去回数差が所定値以上か否かを判定する（ステップＳ９０３）。消去回数差が所定値以上であればステップＳ９０４へ進む。消去回数差が所定値以上でなければ、終了する。

【００８５】
次にストレージコントローラＳＣは、図１２から、平均消去回数が最大であるフラッシュメモリ・モジュール（ＰＤＥＶ）における延べ書込み容量が最大の論理ページアドレス空間と、平均消去回数が最小であるフラッシュメモリ・モジュール（ＰＤＥＶ）における延べ書込み容量が最小の論理ページアドレス空間とを選択する（ステップＳ９０４）。

【００８６】
ストレージコントローラＳＣは、仮想ページ―論理ページアドレス変換テーブルの状態欄を入換え中表示にする。具体的には、ストレージコントローラＳＣは、選択した２つの論理ページアドレス空間の間で、データ入換えと、仮想ページアドレスとのマッピング変更するため、図１０の変換テーブルの状態欄に、入換え操作中を示す値を入力する（ステップＳ９０５）。ストレージコントローラＳＣは、入換え操作中表示がある記憶領域に対してはアクセスを一時的に保留し、データ入換え及びマッピング変更完了を待って、再アクセスする。その際、上位装置からの書込みデータはストレージコントローラＳＣ内のキャッシュメモリに保持する。

【００８７】
次に、ストレージコントローラＳＣは、上記２つの論理ページアドレス空間でデータ入換えを行う（ステップＳ９０６）。入換え操作の詳細は後記する。

【００８８】
ストレージコントローラＳＣは、入換え操作完了後、図１２の消去回数管理テーブルにおいて、入換え対象の領域が属するフラッシュメモリ・モジュール（ＰＤＥＶ）の、平均消去回数（前回）欄に現在の平均消去回数値を置換し、かつ、延べ書込み容量値をクリアする（ステップＳ９０７）。これにより入換え直後の平均消去回数値は、前回値と同じになる。

【００８９】
ストレージコントローラＳＣは、図１１のように、仮想ページアドレスと論理ページア

ドレスのマッピングとオフセット値を変更し、状態欄をクリアし、移動フラグを１に設定する（ステップＳ９０８）。図１２の場合は、平準化の対象となるウエアレベリング・グループＷ００は、２個のフラッシュメモリ・モジュールＰ００とフラッシュメモリ・モジュールＰ０４からなっているので、データ入換えがあると、フラッシュメモリ・モジュールＰ００とフラッシュメモリ・モジュールＰ０４の移動フラグは、共に１に設定される。

【００９０】

ストレージコントローラＳＣは、移動フラグ０の複数のフラッシュメモリ・モジュール（ＰＤＥＶ）があるか否かを判定する（ステップＳ９０９）。移動フラグ０の複数のフラッシュメモリ・モジュールがない場合、平準化処理を終了し、移動フラグ０の複数のフラッシュメモリ・モジュールがある場合は、ステップＳ９０２に戻る。図１２の場合は、２個のフラッシュメモリ・モジュールＰ００とフラッシュメモリ・モジュールＰ０４の移動フラグは１に設定されている場合、平準化処理を終了する。例えば、ウエアレベリング・グループ内のフラッシュメモリ・モジュールが４個以上から構成されている場合は、さらに残りの２個のフラッシュメモリ・モジュールのデータ入換えができるか否かを、ステップＳ９０２でチェックする。

【００９１】

なお、消去回数平準化処理を実行する度に仮想ページアドレスと論理ページアドレスのマッピングが変更され、しかも、平均消去回数値が更新されるので、フラッシュメモリ・モジュールの所定領域に格納したアドレス管理テーブル（図１０又は図１１）と、平均消去回数管理テーブルの図１２を、消去回数平準化処理を実行する度に更新する必要がある。

【００９２】

次にステップＳ９０６におけるデータ入換え操作の詳細を説明する。

図１３は、消去回数平準化に伴うデータ入換え前の仮想ページアドレスと論理ページアドレスのマッピングを示す説明図である。例として、図１３の仮想ページアドレス空間のデータ領域１３０１と仮想ページアドレス空間のデータ領域１３０２のデータ入換えとマッピング変更する場合を説明する。仮想ページアドレス空間のデータ領域１３０１は、論理ページアドレス空間のデータ領域１３０３に対応している。また、仮想ページアドレス空間のデータ領域１３０２は、論理ページアドレス空間のデータ領域１３０４に対応している。論理ページアドレス空間には、データ領域間に空き領域があり、データ領域の後に空き領域がある場合がオフセット値０であり、データ領域の前に空き領域がある場合がオフセット値１と設定される。例えば、図１３の１３０３は、データ領域の前に空き領域があるので、オフセット値１であり、１３０４は、データ領域の後に空き領域があるので、オフセット値０である。よって、論理ページアドレス空間の全体のメモリ容量は、仮想ページアドレス空間の全体のメモリ容量より大きく設定されている。

【００９３】

図１４は、消去回数平準化に伴うデータ入換え後の仮想ページアドレスと論理ページアドレスのマッピングを示す説明図である。仮想ページアドレス空間のデータ領域１４０１は、論理ページアドレス空間のデータ領域の１４０４に対応している。また、仮想ページアドレス空間のデータ領域１４０２は、論理ページアドレス空間のデータ領域の１４０３に対応している。１４０３はオフセット値０であり、１４０４はオフセット値０である。

【００９４】

図１５から図２４は、データ入換えの例として、オフセット値０のデータ領域とオフセット値１の領域でのデータ入換えを段階的に示した図である。図１５から図２４において、左側がオフセット値０の論理ページアドレス空間である。ここでは、論理ページアドレス空間がＥ、Ｆ、Ｇ、Ｈ、－と示しているように５分割されており、ＥからＨが有効データの書き込まれている領域を示し、－が空き領域を示す。

【００９５】

図１５から図２４において、右側がオフセット値１の論理ページアドレス空間である。ここでは、論理ページアドレス空間が－、Ａ、Ｂ、Ｃ、Ｄと示しているように５分割され

ており、ＡからＤが有効データの書き込まれている領域を、－が空き領域を示す。

【００９６】

図１５は、データ入換え操作の入換え前の初期状態を示す説明図である。左側がオフセット値０、右側がオフセット値１の状態である。

【００９７】

図１６は、データ入換え操作の入換え途中の状態を示す説明図である。、左側（オフセット値０）の論理ページアドレス空間Ｅのデータを、右側（オフセット値１）の空き領域－へ上書きする。

【００９８】

図１７は、データ入換え操作の入換え途中の状態を示す説明図である。右側の論理ページアドレス空間Ａのデータを、左側の元の論理ページアドレス空間Ｅへ上書きする。

【００９９】

図１８は、データ入換え操作の入換え途中の状態を示す説明図である。左側の論理ページアドレス空間Ｆのデータを、右側の元の論理ページアドレス空間Ａへ上書きする。

【０１００】

図１９は、データ入換え操作の入換え途中の状態を示す説明図である。右側の論理ページアドレス空間Ｂのデータを、左側の元の論理ページアドレス空間Ｆへ上書きする。

【０１０１】

図２０は、データ入換え操作の入換え途中の状態を示す説明図である。左側の論理ページアドレス空間Ｇのデータを、右側の元の論理ページアドレス空間Ｂへ上書きする。

【０１０２】

図２１は、データ入換え操作の入換え途中の状態を示す説明図である。右側の論理ページアドレス空間Ｃのデータを、左側の元の論理ページアドレス空間Ｇへ上書きする。

【０１０３】

図２２は、データ入換え操作の入換え途中の状態を示す説明図である。左側の論理ページアドレス空間Ｈのデータを、右側の元の論理ページアドレス空間Ｃへ上書きする。

【０１０４】

図２３は、データ入換え操作の入換え途中の状態を示す説明図である。右側の論理ページアドレス空間Ｄのデータを、左側の元の論理ページアドレス空間Ｈへ上書きする。

【０１０５】

図２４は、データ入換え操作の入換え後の最終状態を示す説明図である。左側がオフセット値０、右側がオフセット値０の状態である。

【０１０６】

基本的にフラッシュメモリは、物理アドレス空間において上書き操作ができない半導体デバイスである。すなわち、データを書換えたい場合、物理アドレス空間では、未使用ページにデータを書込み、元のページを無効ページに設定されるのであって、元ページのデータを上書きする操作が実施されていない。

【０１０７】

本発明の実施の形態によれば、論理ページアドレス空間での操作であるため、論理ページ上での上書き操作が可能となっている。このように上書き操作によるデータ入換えに基づいて消去回数平準化を行うことができる。

【０１０８】

図２５は、データ入換え操作のデータ入換え前後でのオフセット値の遷移を示す説明図である。オフセット値０同士の論理ページアドレス空間でデータ入換えを行うと、データ入換え後のオフセット値は０と１になる。オフセット値０の論理ページアドレス空間とオフセット値１の論理ページアドレス空間との間でデータ入換えを行うと、データ入換え後のオフセット値は０と０になる。オフセット値１の論理ページアドレス空間同士でのデータ入換えを行うと、データ入換え後のオフセット値は１と０になる。

【０１０９】

次にデータ入換え操作の詳細をフローチャートを用いて説明する。

図２６は、図１５から図２４で説明したオフセット値０の論理ページアドレス空間とオフセット値１の論理ページアドレス空間との間のデータ入換え手順を示すフローチャートである。ストレージコントローラＳＣは、オフセット値０の論理ページアドレス空間とオフセット値１の論理ページアドレス空間をデータ入換え対象とする（ステップＳ２６０１）。

【０１１０】

ストレージコントローラＳＣは、対象論理ページアドレス空間をｎ分割し、ｉ＝１に設定する（ステップＳ２６０２）。図１５の例ではｎ＝５であり、分割した（ｎ－１）部分に有効データが書き込まれており、残りの１部分が空き領域である。ストレージコントローラＳＣは、オフセット値０の論理ページアドレス空間からオフセット値１の論理ページアドレス空間へのデータ移動（ステップＳ２６０３）と、オフセット値１の論理ページアドレス空間からオフセット値０の論理ページアドレス空間へのデータ移動（ステップＳ２６０４）をし、ｉに1を加算する（ステップＳ２６０５）。そして、ストレージコントローラＳＣは、ｉ＝ｎであるか否かを判定する（ステップＳ２６０６）。ｉ＝ｎと判定されない場合は、ステップＳ２６０３に戻る。ｉ＝ｎの場合に判定された場合は、データ入換えを終了する。

【０１１１】

図２７は、オフセット値０の論理ページアドレス空間とオフセット値０の論理ページアドレス空間との間のデータ入換え手順を示すフローチャートである。オフセット値０の論理ページアドレス空間とオフセット値０の論理ページアドレス空間をデータ入換え対象とする。（ステップＳ２７０１）。

【０１１２】

ストレージコントローラＳＣは、対象論理ページアドレス空間をｎ分割し、ｉ＝ｎに設定する（ステップＳ２７０２）。分割した（ｎ－１）部分に有効データが書き込まれており、残りの１部分が空き領域である。ｉ＝１になるまで、分割した単位ごとのデータ入換えを繰り返す（ステップＳ２７０３～ステップＳ２７０６）。そして、ステップＳ２７０６において、ｉ＝１と判定された場合、データの入換え操作が完了する。

【０１１３】

図２８は、オフセット値１の論理ページアドレス空間とオフセット値１の論理ページアドレス空間との間のデータ入換え手順を示すフローチャートである。オフセット値１の論理ページアドレス空間とオフセット値１の論理ページアドレス空間をデータ入換え対象とする（ステップＳ２８０１）。

【０１１４】

対象データ領域をｎ分割し、ｉ＝２に設定する（ステップＳ２８０２）。分割した（ｎ－１）部分に有効データが書き込まれており、残りの１部分が空き領域である。ｉ＞ｎになるまで、分割した単位ごとのデータ入換えを繰り返す（ステップＳ２８０３～ステップＳ２８０６）。そして、ステップＳ２８０６において、ｉ＞ｎに判定された場合、データの入換え操作が完了する。

【０１１５】

図２９から図３４は、本発明の実施の形態における他の消去回数平準化方法を説明する図である。前述の方法（図１３から図２８）ではデータ入換え用の空き領域をフラッシュメモリ・モジュール内に分散配置したが、ここではモジュール毎に纏めて配置した方法を述べる。

【０１１６】

図２９は、データ入換え前の仮想ページアドレスと論理ページアドレスのマッピングを示す説明図である。図３０は、データ入換え後の仮想ページアドレスと論理ページアドレスのマッピングを示す説明図である。簡単のため、平準化対象のウエアレベリング・グループＷ００は、２個のフラッシュメモリ・モジュールＰ００とフラッシュメモリ・モジュールＰ０４からなる場合を考える。データ入換え前の状態（図２９）において、フラッシュメモリ・モジュールＰ００及びフラッシュメモリ・モジュールＰ０４の論理ページアド

レス空間は、アドレスＡＣ０以上アドレスＡＣ４未満の空間がデータ領域であり、アドレスＡＣ４以上の領域（２９０３、２９０４）がデータ入換え用の空き領域となっている。この空き領域の大きさ（データ長）は、消去回数平準化のためにデータ入換えを行うデータ領域と同じ大きさである。

【０１１７】

図３１は、データ入換え前の仮想ページアドレスと論理ページアドレスの変換テーブルを示す説明図である。図３２は、データ入換え後の仮想ページアドレスと論理ページアドレスの変換テーブルを示す説明図である。データ入換え用の空き領域を纏めて配置したため、図１０、図１１の変換テーブルで必要であったオフセット値管理は不要となる。その代わり、空き領域の位置を管理する必要がある。

【０１１８】

図３３は、データ入換え前の空き領域管理テーブルを示す説明図である。図３４は、データ入換え後の空き領域管理テーブルを示す説明図である。空き領域管理テーブルは、フラッシュメモリ・モジュール毎に、空き領域の先頭論理ページアドレスと大きさ（データ長）を管理する。

【０１１９】

図２９において、仮想ページアドレス空間おけるデータ領域２９０１とデータ領域２９０２のデータを入換え、かつ、仮想ページアドレスと論理ページアドレスのマッピングを変更する手順を説明する。図３１の論理ページアドレスと仮想ページアドレス変換テーブルから、仮想ページアドレス空間におけるデータ領域２９０１に対応する論理ページアドレス空間は、データ領域２９０５であり、仮想ページアドレス空間におけるデータ領域２９０２に対応する論理ページアドレス空間は、データ領域２９０６であることが分かる。また、図３３の空き領域管理テーブルから、フラッシュメモリ・モジュールＰ００におけるデータ入換え用の空き領域は２９０３、フラッシュメモリ・モジュールＰ０４におけるデータ入換え用の空き領域は２９０４であることが分かる。

【０１２０】

次に、データ領域２９０５のデータを空き領域２９０４に書込み、データ領域２９０６のデータを空き領域２９０３へ書込む。そして、図３０に示すように、仮想ページアドレス空間のデータ領域３００１を論理ページアドレス空間のデータ領域３００４へ対応させ、仮想ページアドレス空間におけるデータ領域３００２を論理ページアドレス空間のデータ領域３００３へ対応させる。上記データの入換え完了後、図３２に示すように、仮想ページアドレスと論理ページアドレスの変換テーブルを変更する。また、図３４の空き領域管理テーブルから、フラッシュメモリ・モジュールＰ００におけるデータ入換え用の空き領域は３００５、フラッシュメモリ・モジュールＰ０４におけるデータ入換え用の空き領域は３００６であることが分かる。

【０１２１】

本方法によれば、データ入換え用の空き領域をフラッシュメモリ・モジュール内に分散配置せずに、モジュール毎に纏めて配置して空き領域を設けているため、オフセット値の管理が不要のため、データ入換えの制御が簡単になる。

【０１２２】

次に、フラッシュメモリ・モジュール（ＰＤＥＶ）に障害が発生した場合の対処方法について説明する。図３５から図３９により、本発明の実施の形態において、フラッシュメモリ・モジュールに障害が発生した場合のモジュール交換方法を示す。

【０１２３】

図３５は、モジュール交換の手順を示すフローチャートである。図３６から図３９は、図３５のフローチャートの各ステップの構成を示す説明図である。

【０１２４】

図３６は、フラッシュメモリ・モジュールに障害発生した場合を示す説明図である。図３６には、ＲＡＩＤグループ（ＶＤＥＶ）Ｖ００と、ＲＡＩＤグループＶ００を構成するウエアレベリング・グループ（ＷＤＥＶ）Ｗ００、ウエアレベリング・グループＷ１０、

ウエアレベリング・グループＷ２０、及びウエアレベリング・グループＷ３０が示されている。また、ウエアレベリング・グループＷ００が接続されているチャネルと同じＤ０１上に、予備グループ（ＹＤＥＶ）Ｙ００が接続されている。ウエアレベリング・グループＷ００内のフラッシュメモリ・モジュール（ＰＤＥＶ）Ｐ０１に障害発生（ステップＳ３５０１）した場合を考える。

【０１２５】

次に、ウエアレベリング・グループ（ＷＤＥＶ）Ｗ００が使用できる予備グループ（ＹＤＥＶ）を選定する。ウエアレベリング・グループＷ００と同じチャネルＤ０１上に接続されている予備グループＹ００を選定する（ステップＳ３５０２）。そして予備グループＹ００に属するフラッシュメモリ・モジュールから、フラッシュメモリ・モジュールＰ０１と交代させるモジュールＰ０４を選定する（ステップＳ３５０３）。

【０１２６】

図３７は、フラッシュメモリ・モジュールを入換え後の状態を示す説明図である。図３７に示すように、ウエアレベリング・グループＷ００と予備グループＹ００の間で、フラッシュメモリ・モジュールＰ０１とフラッシュメモリ・モジュールＰ０４を入換える。障害のモジュールＰ０１は、交換待ちの状態となる。

【０１２７】

次に図３８は、フラッシュメモリ・モジュールを入換え後のデータ再生を示す説明図である。図３８に示すように、新しくウエアレベリング・グループＷ００に組み入れたフラッシュメモリ・モジュールＰ０４に対して、フラッシュメモリ・モジュールＰ０１に書き込まれていたデータを再生し、書き込む（ステップＳ３５０４）。このとき、データ再生に使用するデータは、消去回数平準化のため、他のウエアレベリング・グループ内のフラッシュメモリ・モジュール間で分散して格納されていることに注意する。つまり、データ再生は、ウエアレベリング・グループ内で同じ仮想ページアドレスに格納されたデータについて行う。

【０１２８】

図３９は、予備グループのフラッシュメモリ・モジュールを新しく交換した場合を示す説明図である。図３９に示すように、交換待ちフラッシュメモリ・モジュールＰ０１を新しいフラッシュメモリ・モジュールＰ０６と交換し、そのＰ０６を予備グループＹ００に組み入れる（ステップＳ３５０５）。以上で、フラッシュメモリ・モジュールの交換処理が終了する。

【０１２９】

なお、モジュール交換直後における図９の消去回数平準化処理の実行可否は、交換前のモジュールにおける延べ書込み容量値に基づき判定する。新規モジュールでは、その論理ページアドレス空間全領域の延べ書込み容量値がデータ再生に伴う書込み以外はゼロであり、論理ページアドレス空間内所定領域ごとの書込み頻度を知ることができない。従って交換前のモジュールの書込み容量値を用いることにより、論理ページアドレス空間における書込み頻度が分かり、消去回数平準化処理を実行できる。

【産業上の利用可能性】

【０１３０】

本発明は、複数のフラッシュメモリ・モジュールに渡って消去回数を平準化し、フラッシュメモリ・モジュールの長寿命化の用途に適用でき、例えば、複数のフラッシュメモリ・モジュールを使用した大容量のフラッシュメモリを用いた記憶装置、その消去回数平準化方法、及び消去回数平準化プログラムの用途に適用できる。

【図面の簡単な説明】

【０１３１】

【図１】本発明によるストレージ装置の実施の形態の構成を示すブロック図である。

【図２】チャネルアダプタの構成を示すブロック図である。

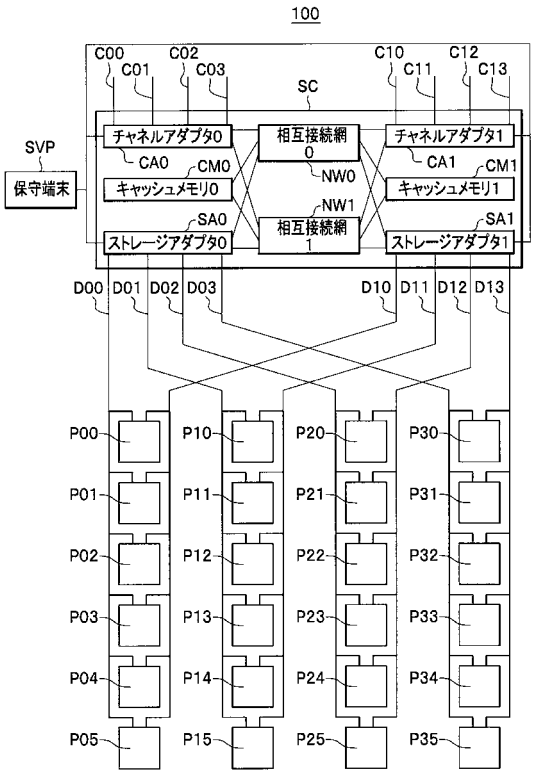
【図３】ストレージアダプタの構成を示すブロック図である。

【図４】フラッシュメモリ・モジュールの構成を示すブロック図である。

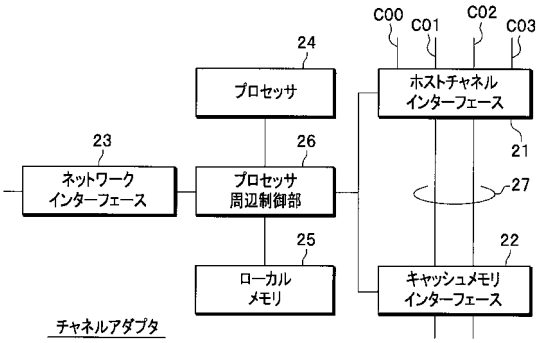
【図５】フラッシュメモリ・モジュールのブロックの構成を示す説明図である。
【図６】本発明によるストレージ装置の実施の形態における論理グループの構成とアドレス変換の階層を示す説明図である。
【図７】本発明によるストレージ装置の実施の形態におけるＲＡＩＤグループの構成を示す説明図である。
【図８】フラッシュメモリ・モジュールとハードディスクドライブをストレージコントローラに接続した例を示す構成図である。
【図９】複数のフラッシュメモリ・モジュール間の消去回数平準化方法を示すフローチャートである。
【図１０】消去回数平準化に伴うデータ入換え前の仮想ページアドレスと論理ページアドレスの変換テーブルを示す説明図である。
【図１１】消去回数平準化に伴うデータ入換え後の仮想ページアドレスと論理ページアドレスの変換テーブルを示す説明図である。
【図１２】ストレージコントローラにおいて管理するフラッシュメモリ・モジュール毎の消去回数管理テーブルを示す説明図である。
【図１３】消去回数平準化に伴うデータ入換え前の仮想ページアドレスと論理ページアドレスのマッピングを示す説明図である。
【図１４】消去回数平準化に伴うデータ入換え後の仮想ページアドレスと論理ページアドレスのマッピングを示す説明図である。
【図１５】データ入換え操作の入換え前の初期状態を示す説明図である。
【図１６】データ入換え操作の入換え途中の状態を示す説明図である。
【図１７】データ入換え操作の入換え途中の状態を示す説明図である。
【図１８】データ入換え操作の入換え途中の状態を示す説明図である。
【図１９】データ入換え操作の入換え途中の状態を示す説明図である。
【図２０】データ入換え操作の入換え途中の状態を示す説明図である。
【図２１】データ入換え操作の入換え途中の状態を示す説明図である。
【図２２】データ入換え操作の入換え途中の状態を示す説明図である。
【図２３】データ入換え操作の入換え途中の状態を示す説明図である。
【図２４】データ入換え操作の入換え後の最終状態を示す説明図である。
【図２５】データ入換え操作のデータ入換前後でのオフセット値の遷移を示す説明図である。
【図２６】図１５から図２４で説明したオフセット値０の論理ページアドレス空間とオフセット値１の論理ページアドレス空間との間のデータ入換え手順を示すフローチャートである。
【図２７】オフセット値０の論理ページアドレス空間とオフセット値０の論理ページアドレス空間との間のデータ入換手順を示すフローチャートである。
【図２８】オフセット値１の論理ページアドレス空間とオフセット値１の論理ページアドレス空間との間のデータ入換手順を示すフローチャートである。
【図２９】データ入換え前の仮想ページアドレスと論理ページアドレスのマッピングを示す説明図である。
【図３０】データ入換え後の仮想ページアドレスと論理ページアドレスのマッピングを示す説明図である。
【図３１】データ入換え前の仮想ページアドレスと論理ページアドレスの変換テーブルを示す説明図である。
【図３２】データ入換え後の仮想ページアドレスと論理ページアドレスの変換テーブルを示す説明図である。
【図３３】データ入換え前の空き領域管理テーブルを示す説明図である。
【図３４】データ入換え後の空き領域管理テーブルを示す説明図である。
【図３５】モジュール交換の手順を示すフローチャートである。
【図３６】フラッシュメモリ・モジュールに障害発生した場合を示す説明図である。

【図３７】フラッシュメモリ・モジュールを入換え後の状態を示す説明図である。
【図３８】フラッシュメモリ・モジュールを入換え後のデータ再生を示す説明図である。
【図３９】予備グループのフラッシュメモリ・モジュールを新しく交換した場合を示す説明図である。
【符号の説明】
【０１３２】
ＣＡ０，ＣＡ１　チャネルアダプタ
ＣＭ０，ＣＭ１　キャッシュメモリ
Ｃ００，Ｃ０１，Ｃ０２，Ｃ０３，Ｃ１０，Ｃ１１，
Ｃ１２，Ｃ１３，Ｄ００，Ｄ０１，Ｄ０２，Ｄ０３，
Ｄ１０，Ｄ１１，Ｄ１２，Ｄ１３　チャネル
ＭＣ　メモリコントローラ
ＭＥＭ　フラッシュメモリ
ＮＷ０，ＮＷ１　相互接続網
Ｐ００，Ｐ０１，Ｐ０２，Ｐ０３，Ｐ０４，Ｐ０５，
Ｐ１０，Ｐ１１，Ｐ１２，Ｐ１３，Ｐ１４，Ｐ１５，
Ｐ２０，Ｐ２１，Ｐ２２，Ｐ２３，Ｐ２４，Ｐ２５，
Ｐ３０，Ｐ３１，Ｐ３２，Ｐ３３，Ｐ３４，
Ｐ３５　フラッシュメモリ・モジュール
ＳＡ０，ＳＡ１　ストレージアダプタ
ＳＣ　ストレージコントローラ
ＳＶＰ　保守端末
２１　ホストチャネル・インターフェース
２２　キャッシュメモリ・インターフェース
２３　ネットワーク・インターフェース
２４　プロセッサ
２５　ローカルメモリ
２６　プロセッサ周辺制御部
２７　信号線
３１　キャッシュメモリ・インターフェース
３２　ストレージチャネル・インターフェース
３３　ネットワーク・インターフェース
３４　プロセッサ
３５　ローカルメモリ
３６　プロセッサ周辺制御部
３７　信号線
１００　ストレージ装置
４０１　プロセッサ
４０２　インターフェース部
４０３　データ転送部
４０４，４０７　メモリ
４０５　フラッシュメモリ・チップ
４０６　ブロック

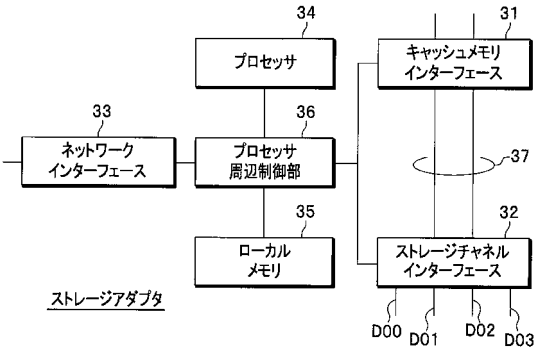
【図１】



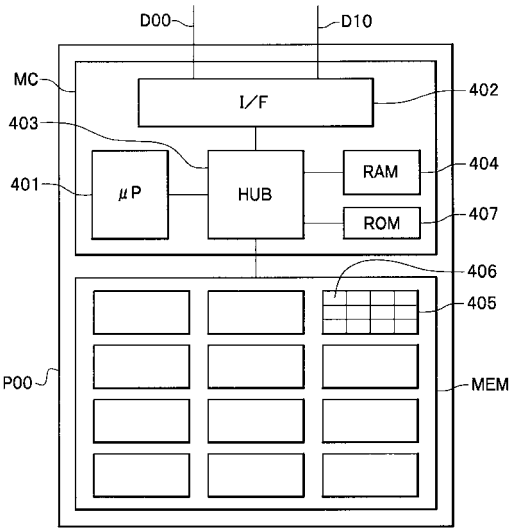
【図２】



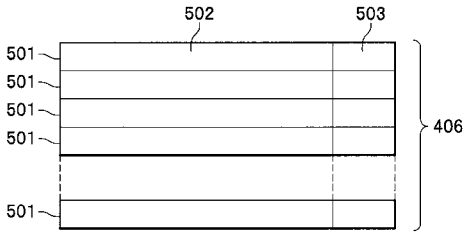
【図３】



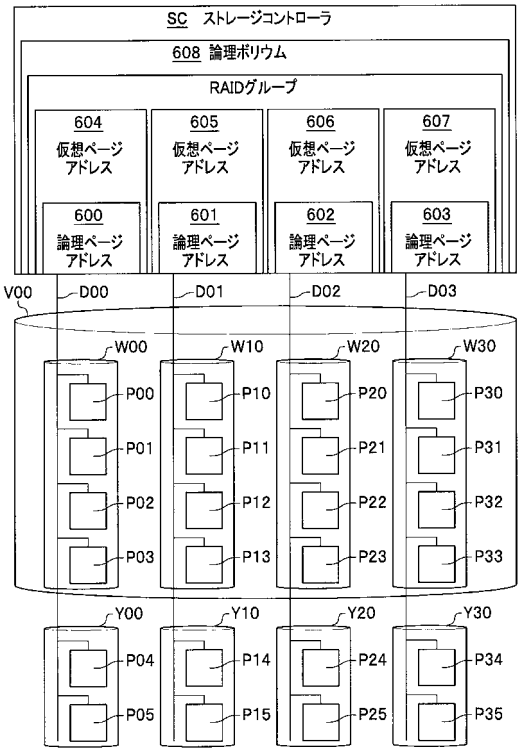
【図４】



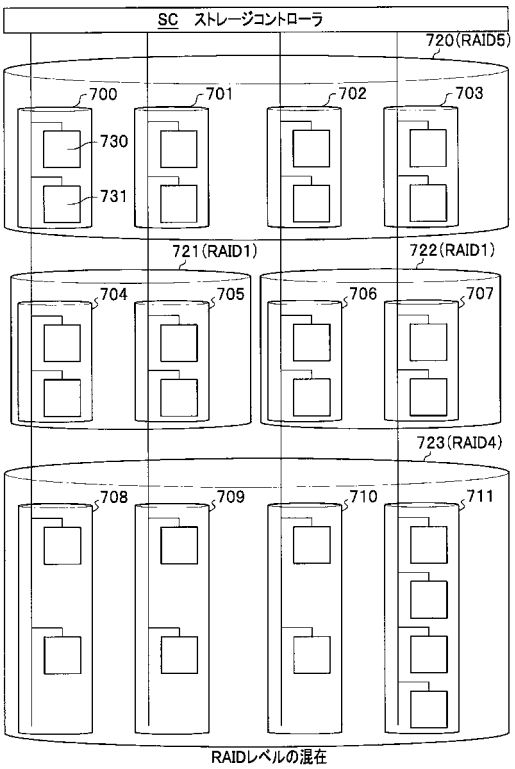
【図５】



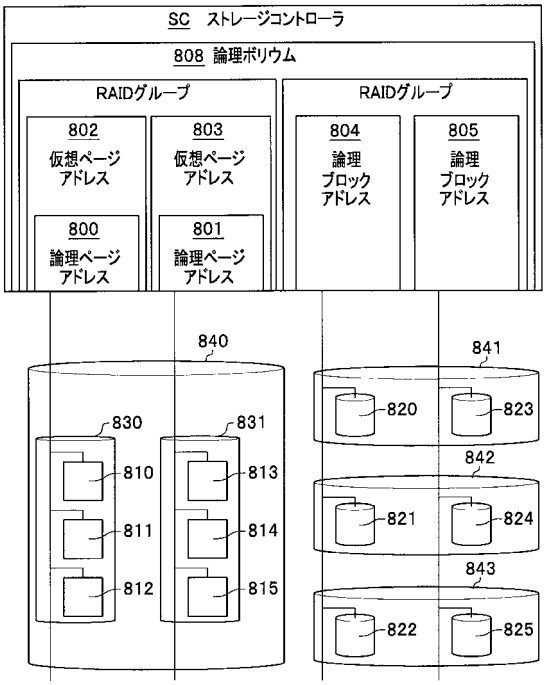
【図６】



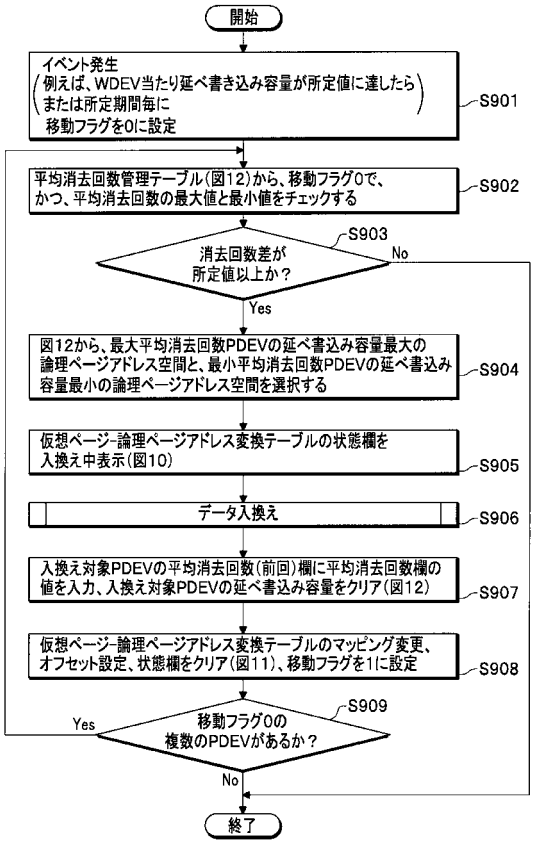
【図７】



【図８】



【図９】



【図１０】

WDEV	仮想ページアドレス	データ長	PDEV	論理ページアドレス	オフセット	状態
W00	AA0		P00	AB0	0	
	AA1		P00	AB1	1	1
	AA2		P00	AB2	0	
	AA3		P00	AB3	0	
	AA4		P04	AB0	1	
	AA5		P04	AB1	1	
	AA6		P04	AB2	0	1
	AA7		P04	AB3	1	

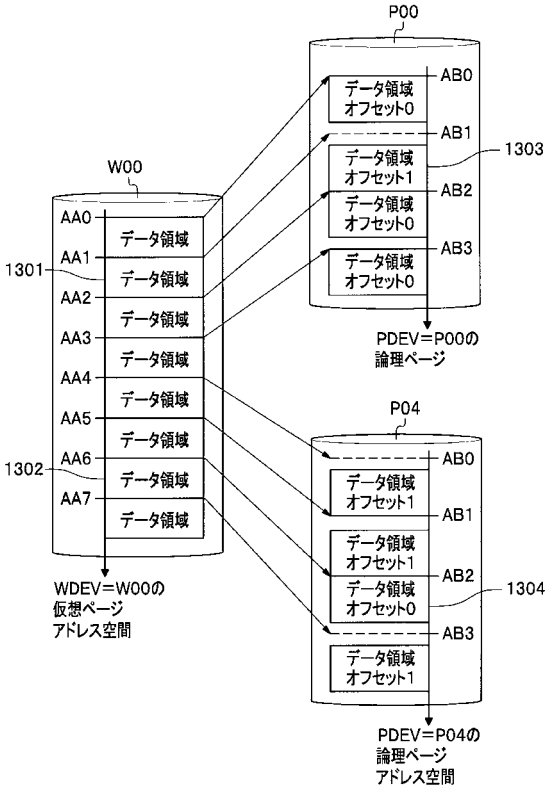
【図１１】

WDEV	仮想ページアドレス	データ長	PDEV	論理ページアドレス	オフセット	状態
W00	AA0		P00	AB0	0	
	AA1		P04	AB2	0	
	AA2		P00	AB2	0	
	AA3		P00	AB3	0	
	AA4		P04	AB0	1	
	AA5		P04	AB1	1	
	AA6		P00	AB1	0	
	AA7		P04	AB3	1	

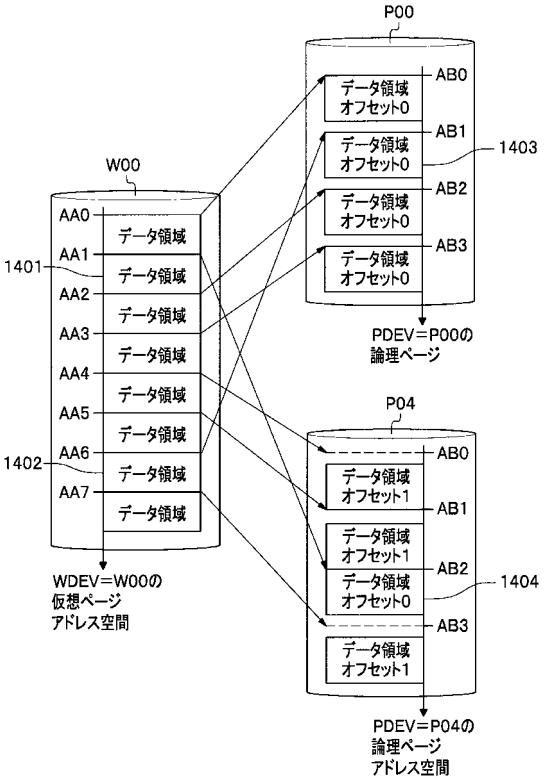
【図１２】

PDEV	PDEV容量	論理ページアドレス	延べ書込み容量	平均消去回数	平均消去回数(前回)	移動フラグ
P00	s00	AB0	a000	e00=f00+(a000+a001+a002+a003)/s00	f00	0
		AB1	a001			
		AB2	a002			
		AB3	a003			
P04	s04	AB0	a040	e04=f04+(a040+a041+a042+a043)/s04	f04	0
		AB1	a041			
		AB2	a042			
		AB3	a043			

【図１３】

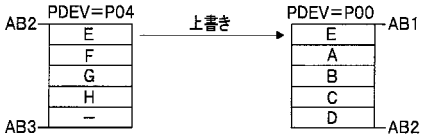


【図１４】

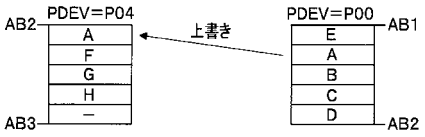


【図１５】

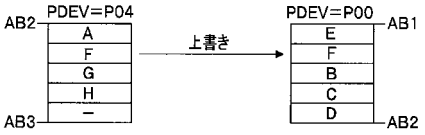
【図１６】



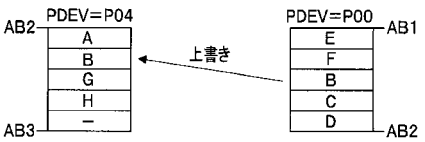
【図１７】



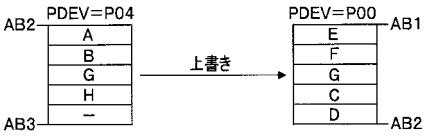
【図１８】



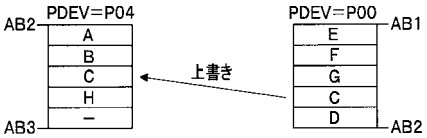
【図１９】



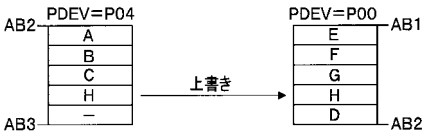
【図２０】



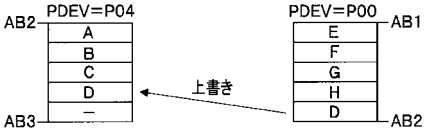
【図２１】



【図２２】



【図２３】

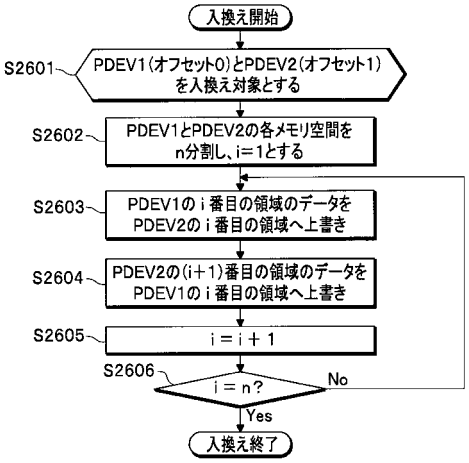


【図２４】

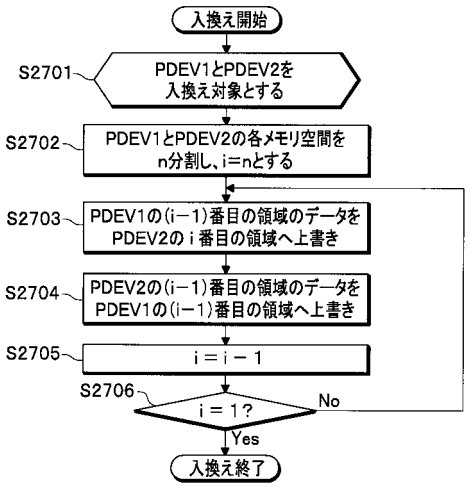
【図２５】

データ入換え前のオフセット値の組合せ		データ入換え後のオフセット値の組合せ	
0	0	0	1
0	1	0	0
1	1	1	0

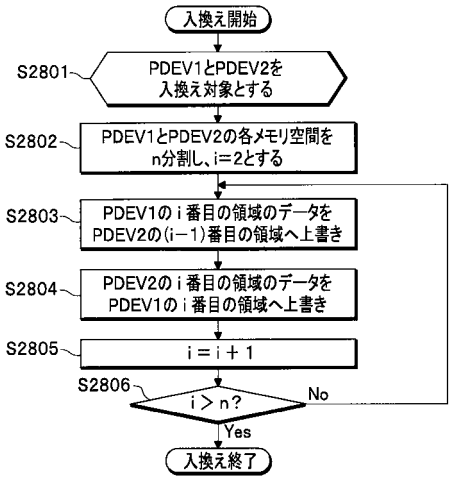
【図２６】



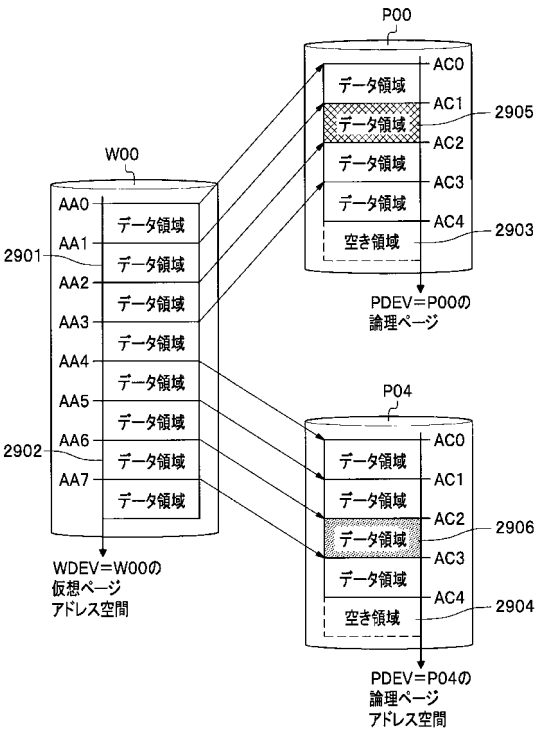
【図２７】



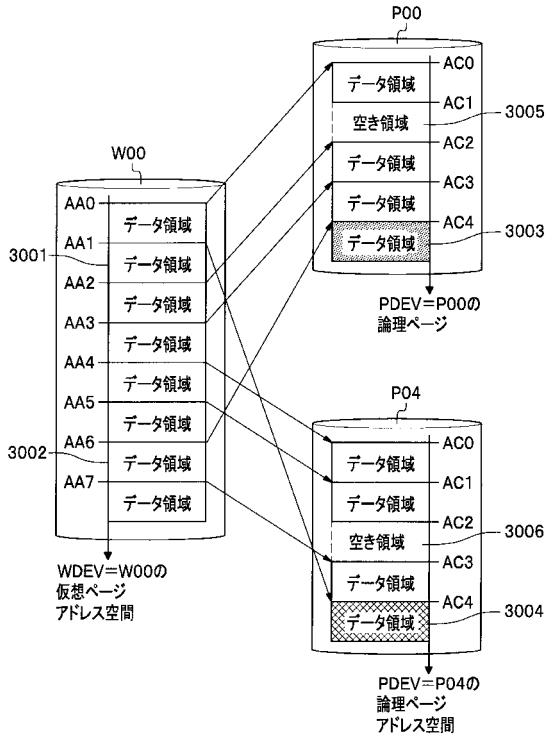
【図２８】



【図２９】



【図３０】



【図３１】

WDEV	仮想ページアドレス	データ長	PDEV	論理ページアドレス	状態
W00	AA0		P00	AC0	
	AA1		P00	AC1	
	AA2		P00	AC2	
	AA3		P00	AC3	
	AA4		P04	AC0	
	AA5		P04	AC1	
	AA6		P04	AC2	
	AA7		P04	AC3	

【図３２】

WDEV	仮想ページアドレス	データ長	PDEV	論理ページアドレス	状態
W00	AA0		P00	AC0	
	AA1		P04	AC4	
	AA2		P00	AC2	
	AA3		P00	AC3	
	AA4		P04	AC0	
	AA5		P04	AC1	
	AA6		P00	AC4	
	AA7		P04	AC3	

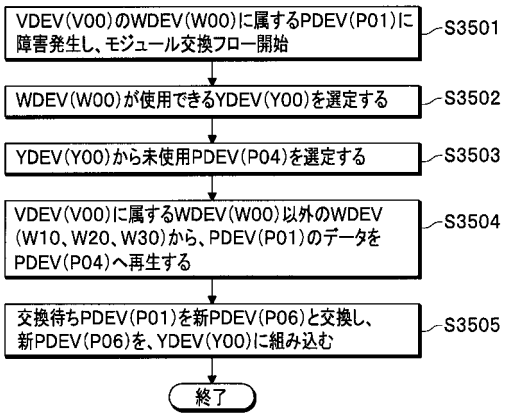
【図３３】

PDEV	論理ページアドレス	データ長
P00	AC4	
P04	AC4	

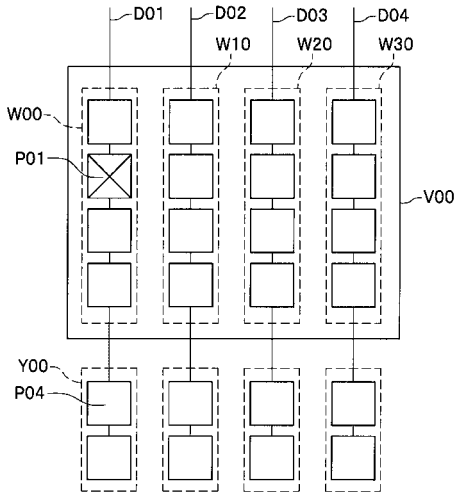
【図３４】

PDEV	論理ページアドレス	データ長
P00	AC1	
P04	AC2	

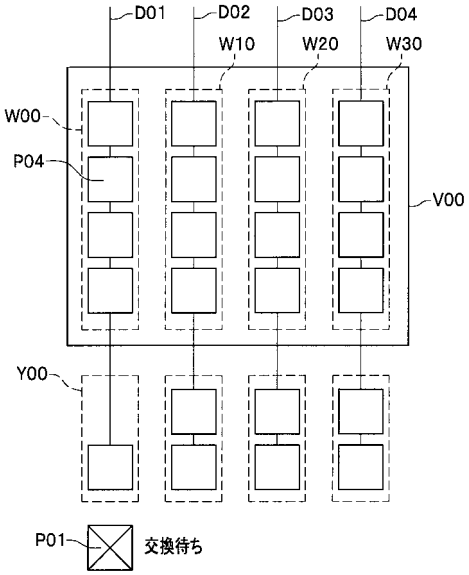
【図３５】



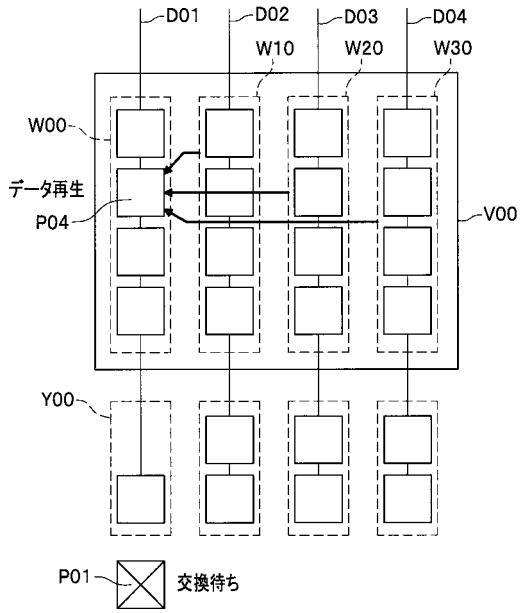
【図３６】



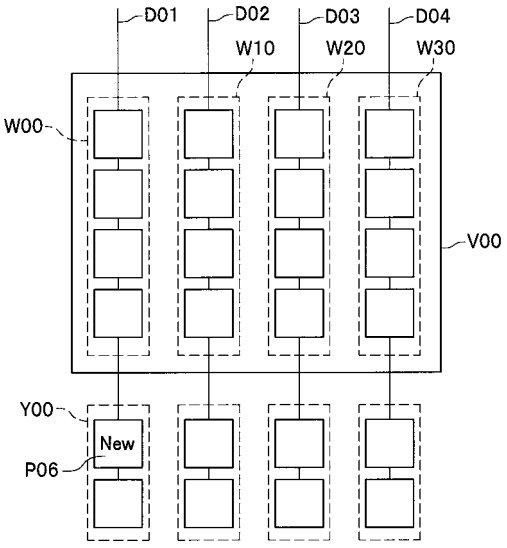
【図３７】



【図３８】



【図３９】



フロントページの続き

(51)Int.Cl. FI
G06F 3/06 305C

(56)参考文献 特開２０００－２０７１３７（ＪＰ，Ａ）
特開平０９－２１８７５４（ＪＰ，Ａ）
特開２００４－０２１８１１（ＪＰ，Ａ）
特開平８－１６４８２（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ １２／００
Ｇ０６Ｆ １２／１６
Ｇ０６Ｆ ３／０６
Ｇ０６Ｆ ３／０８