

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4939234号
(P4939234)

(45) 発行日 平成24年5月23日(2012.5.23)

(24) 登録日 平成24年3月2日(2012.3.2)

(51) Int.Cl.

G06F 12/16 (2006.01)
G06F 3/08 (2006.01)

F 1

G06F 12/16 310B
G06F 12/16 320K
G06F 3/08 H

請求項の数 20 (全 26 頁)

(21) 出願番号 特願2007-3050 (P2007-3050)
 (22) 出願日 平成19年1月11日 (2007.1.11)
 (65) 公開番号 特開2008-171163 (P2008-171163A)
 (43) 公開日 平成20年7月24日 (2008.7.24)
 審査請求日 平成21年6月26日 (2009.6.26)

(73) 特許権者 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74) 代理人 100093861
 弁理士 大賀 真司
 (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 中村 崇仁
 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 野田 佳邦

最終頁に続く

(54) 【発明の名称】フラッシュメモリモジュール、そのフラッシュメモリモジュールを記録媒体として用いたストレージ装置及びそのフラッシュメモリモジュールのアドレス変換テーブル検証方法

(57) 【特許請求の範囲】

【請求項 1】

フラッシュメモリコントローラと少なくとも一個のフラッシュメモリチップと、
 からなり、

前記フラッシュメモリコントローラは、

前記フラッシュメモリチップの論理ページアドレスと物理ページアドレスを変換するア
 ドレス変換テーブルを格納するメモリ備え、

前記フラッシュメモリコントローラは、データの読み書き時の第1の動作と、動作停止
 、電源電圧低下及び動作周波数低減のいずれかにより前記第1の動作より低消費電力で動
 作する待機状態時の第2の動作とを制御する

フラッシュメモリモジュールにおいて、

前記第2の動作時に前記アドレス変換テーブルのデータ検証を行うことを特徴とするフ
 ラッシュメモリモジュール。

【請求項 2】

請求項1記載のフラッシュメモリモジュールにおいて、

前記第2の動作時は、前記フラッシュメモリコントローラは、前記アドレス変換テーブ
 ルのデータ検証を行う第1の期間と、前記フラッシュメモリコントローラの動作を停止さ
 せる第2の期間を交互に実行し、

前記第1の期間が前記第2の期間より短いことを特徴とするフラッシュメモリモジュ
 ル。

10

20

【請求項 3】

請求項 1 記載のフラッシュメモリモジュールにおいて、

前記論理ページアドレスと前記物理ページアドレスは、エラー訂正情報を付加して前記アドレス変換テーブルに格納し、

前記アドレス変換テーブルのデータ検証時には、前記アドレス変換テーブルに格納されている前記論理ページアドレスと前記物理ページアドレスと前記エラー訂正情報を検証することによりエラー有無を判定し、

前記アドレス変換テーブルにエラーが有りと判定された場合に、

前記エラー訂正情報により訂正可能なエラーならば、前記エラー訂正情報から訂正した値を前記アドレス変換テーブルに格納し、

前記エラー訂正情報のみで訂正不能ならば、前記フラッシュメモリチップから必要なデータを読み出し、前記アドレス変換テーブルを再構築することを特徴とするフラッシュメモリモジュール。

【請求項 4】

請求項 1 記載のフラッシュメモリモジュールにおいて、

前記フラッシュメモリコントローラと接続された上位装置からの指示により前記第 1 の動作から前記第 2 の動作へ移行することを特徴とするフラッシュメモリモジュール。

【請求項 5】

請求項 1 記載のフラッシュメモリモジュールにおいて、

前記フラッシュメモリコントローラ内の制御により前記第 1 の動作から前記第 2 の動作へ移行することを特徴とするフラッシュメモリモジュール。

【請求項 6】

請求項 1 記載のフラッシュメモリモジュールにおいて、

前記第 2 の動作時に前記アドレス変換テーブルのデータ検証を行うか否かを設定することを特徴とするフラッシュメモリモジュール。

【請求項 7】

請求項 1 記載のフラッシュメモリモジュールにおいて、

前記フラッシュメモリコントローラの消去回数平準化処理完了を待って、前記第 1 動作から前記第 2 の動作への移行及び前記アドレス変換テーブルのデータ検証を開始することを特徴とするフラッシュメモリモジュール。

【請求項 8】

請求項 1 記載のフラッシュメモリモジュールにおいて、

前記メモリは、S R A Mであることを特徴とするフラッシュメモリモジュール。

【請求項 9】

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へのデータの読み書きを制御するフラッシュメモリコントローラとを備えるフラッシュメモリモジュールと、

前記フラッシュメモリチップへ読み書きされるデータを一時的に格納するキャッシュメモリを備えるストレージコントローラと、
を備え、

前記フラッシュメモリコントローラは、

前記フラッシュメモリチップの論理ページアドレスと物理ページアドレスを変換するアドレス変換テーブルを格納するメモリを備え、

データの読み書き時の第 1 の動作と、動作停止、電源電圧低下及び動作周波数低減のいずれかにより前記第 1 の動作より低消費電力で動作する待機状態時の第 2 の動作とを制御する、

ストレージ装置において、

前記フラッシュメモリコントローラは、前記第 2 の動作時に前記アドレス変換テーブルのデータ検証を行うことを特徴とするストレージ装置。

【請求項 10】

請求項 9 記載のストレージ装置において、

10

20

30

40

50

前記ストレージコントローラからのコマンド受信により前記フラッシュメモリコントローラは前記第1の動作から前記第2の動作へ移行する
ことを特徴とするストレージ装置。

【請求項11】

請求項9記載のストレージ装置において、
前記フラッシュメモリコントローラ内部の制御により、前記フラッシュメモリコントローラは前記第1の動作から前記第2の動作へ移行する
ことを特徴とするストレージ装置。

【請求項12】

請求項9記載のストレージ装置において、
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RAIDグループを構成し、
前記RAIDグループ内の1つのフラッシュメモリモジュールが前記第1の動作から前記第2の動作へ移行すると、前記RAIDグループ内の他のフラッシュメモリモジュールも前記第2の動作へ移行するように、前記ストレージコントローラがコマンドを発行することを特徴とするストレージ装置。

【請求項13】

請求項12記載のストレージ装置において、
前記ストレージコントローラへの接続時にアドレス変換テーブルを構築後、前記第2の動作で前記構築されたアドレス変換テーブルのデータ検証を行うスペアのフラッシュメモリモジュールを備え、

前記複数のフラッシュメモリモジュールのいずれかに障害が生じたときは、その障害の生じたフラッシュメモリモジュールに代えて、その障害が発生したフラッシュメモリモジュールと同じRAIDグループのフラッシュメモリモジュールから前記スペアのフラッシュメモリモジュールへデータをリビルドすることを特徴とするストレージ装置。

【請求項14】

請求項13記載のストレージ装置において、
前記フラッシュメモリモジュールは、
前記論理ページアドレスと前記物理ページアドレスは、エラー訂正情報を付加して前記アドレス変換テーブルに格納し、

前記アドレス変換テーブルのデータ検証時に、前記アドレス変換テーブルに格納されている前記論理ページアドレスと前記物理ページアドレスと前記エラー訂正情報を検証することによりエラー有無を判定し、

前記アドレス変換テーブルにエラーが有りと判定された場合に、

前記エラー訂正情報により訂正可能なエラーならば、前記エラー訂正情報から訂正した値を前記アドレス変換テーブルに格納し、

前記エラー訂正情報のみで訂正不能ならば、前記フラッシュメモリチップから必要なデータを読み出し、前記アドレス変換テーブルを再構築するように構成され、

前記フラッシュメモリモジュールに障害が生じたか否かは、前記エラー訂正情報のみで訂正不能なエラーが、前記データ検証時に所定回以上あったか否かで判定されることを特徴とするストレージ装置。

【請求項15】

請求項9記載のストレージ装置において、
前記フラッシュメモリモジュールは、前記フラッシュメモリチップへ読み書きするデータを送受信するための第1のネットワークと、前記第2の動作を行う消費電力制御のためのデータを送受信するための第2のネットワークで、前記ストレージコントローラと接続することを特徴とするストレージ装置。

【請求項16】

フラッシュメモリチップと、このフラッシュメモリチップの論理ページアドレスと物理ページアドレスを変換するアドレス変換テーブルを格納するメモリを備え、前記フラッシュメモリチップへのデータの読み書きを制御するとともに、前記データの読み書き時の第

10

20

30

40

50

1の動作と、動作停止、電源電圧低下及び動作周波数低減のいずれかにより前記第1の動作より低消費電力で動作する待機状態時の第2の動作とを制御するフラッシュメモリコントローラとを備えるフラッシュメモリモジュールと、

前記フラッシュメモリチップへ読み書きされるデータを一時的に格納するメモリを備えるストレージコントローラと、

を備えるストレージ装置のデータ検証方法において、

前記第2の動作時に前記アドレス変換テーブルのデータ検証を行うステップを備えることを特徴とするストレージ装置のデータ検証方法。

【請求項17】

請求項16記載のストレージ装置のデータ検証方法において、

10

前記ストレージコントローラからのコマンド受信により前記第1の動作から前記第2の動作へ移行するステップを

備えることを特徴とするストレージ装置のデータ検証方法。

【請求項18】

請求項16記載のストレージ装置のデータ検証方法において、

前記フラッシュメモリコントローラ内部の制御により、前記第1の動作から前記第2の動作へ移行するステップを

備えることを特徴とするストレージ装置のデータ検証方法。

【請求項19】

請求項16記載のストレージ装置のデータ検証方法において、

20

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構成されたRAIDグループ内の1つのフラッシュメモリモジュールが前記第1の動作から前記第2の動作へ移行する、前記RAIDグループ内の他のフラッシュメモリモジュールも前記第2の動作へ移行するステップを

備えることを特徴とするストレージ装置。

【請求項20】

請求項19記載のストレージ装置のデータ検証方法において、

前記複数のフラッシュメモリモジュールのいずれかに障害が生じたときは、その障害の生じたフラッシュメモリモジュールに代えて、その障害が発生したフラッシュメモリモジュールと同じRAIDグループのフラッシュメモリモジュールからアドレス変換テーブルの構築後前記第2の動作でそのアドレス変換テーブルのデータ検証を行うスペアのフラッシュメモリモジュールへデータをリビルトするステップを

30

備えることを特徴とするストレージ装置のデータ検証方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、フラッシュメモリモジュール、そのフラッシュメモリモジュールを記録媒体として用いたストレージ装置及びそのフラッシュメモリモジュールのアドレス変換テーブル検証方法に関し、特に、フラッシュメモリモジュール内に記録されたデータにアクセスするために使用するアドレス変換テーブルを検証するものに適用しても好適なものである。

40

【背景技術】

【0002】

ストレージ装置は、一般的に、ランダムアクセス可能な不揮発性記録媒体を備える。このランダムアクセス可能な不揮発性記録媒体として、例えば、磁気ディスク又は光ディスク等が用いられる。

【0003】

また、半導体技術の進歩に伴って、例えば、フラッシュメモリのような一括消去可能な不揮発性半導体記録媒体が開発されている。フラッシュメモリモジュールは、モジュール内のフラッシュメモリコントローラが論理ページアドレスを物理ページアドレスにアドレ

50

ス変換テーブルを参照し変換してフラッシュメモリチップにデータを記録する。このフラッシュメモリモジュールを記録媒体とするストレージ装置は、磁気ディスクのような小型ディスクドライブを多数備えるストレージ装置と比べ、寿命、省電力及びアクセス時間等に優れていると考えられている。

【0004】

なお、フラッシュメモリモジュールにおけるアドレス変換テーブルの不正マッピング防止方法が知られている（特許文献1参照）。この従来技術は、有効な論理ページアドレスが無効な論理ページアドレスにデータ化けした場合の、不正マッピングを防止する技術である。

【特許文献1】特開2003-337757号公報

10

【発明の開示】

【発明が解決しようとする課題】

【0005】

フラッシュメモリモジュールを記録媒体として用いたストレージ装置は、既述のように磁気ディスク等を用いたストレージ装置に比べ、省電力を図ることができると考えられている。しかしながら、現在市販されている2.5インチのハードディスクドライブと、2.5インチハードディスクドライブ互換フラッシュメモリモジュールの消費電力値の例を比較すると、データ読み書き動作時（R/W時）は、フラッシュメモリモジュールの消費電力値が2.9ワット、ハードディスクドライブの消費電力値が2.3ワットであり、フラッシュメモリモジュールの消費電力値がハードディスクドライブの消費電力値より少し大きい。また、アイドル時は、フラッシュメモリモジュールの消費電力値が2.2ワット、ハードディスクドライブの消費電力値が1.2ワットであり、フラッシュメモリモジュールの消費電力値がハードディスクドライブの消費電力値の約2倍となっているという結果が得られている。

20

【0006】

フラッシュメモリチップ自体は、アイドル時のチップ当たりの消費電力値はサブミリワット級であり、非常に低消費電力である。従って、アイドル時のフラッシュメモリモジュールの消費電力の大部分は、フラッシュメモリコントローラによるものであると考えられる。このため、フラッシュメモリチップの低消費電力特性を活かすためには、アイドル時のフラッシュメモリコントローラの低消費電力化が必須となる。

30

【0007】

フラッシュメモリコントローラは、通常、CMOS LSI（Complementary Metal Oxide Semiconductor Large Scale Integration）である。CMOS LSIは、電源電圧あるいは動作周波数を下げるにより、低消費電力化が可能である。従って、フラッシュメモリコントローラの低消費電力化を行うには、フラッシュメモリコントローラの電源低電圧化、動作停止あるいは動作周波数の低下が考えられる。

【0008】

一方、フラッシュメモリコントローラ内のアドレス変換テーブルはRAMに格納されるが、このRAMは高速アクセス特性が要求されるため例えばSRAM（Static Random Access Memory）が用いられる。SRAMは、非常に小さい確率ではあるが、放射線等の影響によりソフトエラーが発生する場合がある。メモリコントローラの低消費電力化を図ると、その低消費電力動作時には、アドレス変換テーブルに書き込まれたデータのソフトエラーの発生率は増加する。

40

【0009】

このように発生するソフトエラーに対処するため、一般的には、アドレス変換テーブルに格納されるデータにエラー訂正情報を付加し、データの読み出し時にアドレスデータとそのデータに付加されたエラー訂正情報を検証することにより、ソフトエラーからデータを保証している。言い換えると、SRAM内のアクセスしたデータに関しては、エラー訂正情報によりデータ保証されるが、アクセスしないデータに関しては、エラー保証がされないことになる。したがって、フラッシュメモリコントローラにおいて、動作停止による

50

低消費電力動作時には、S R A Mに格納したアドレス変換テーブルにアクセスしないため、低消費電力動作時に発生したソフトエラーを検出・訂正するデータ検証を行うことができない。また、動作周波数を低下させた低消費電力動作の場合も、アドレス変換テーブル内のデータにアクセスしない限り、ソフトエラーは検出できない。

【0010】

低消費電力動作時に発生したソフトエラーは、通常動作に復帰後に、アドレス変換テーブル内のエラーデータにアクセスして初めて検出することができる。このため低消費電力動作が長期間に及ぶ場合は、アドレス変換テーブル内にソフトエラーが蓄積し、エラー訂正情報からはエラー訂正ができなくなる場合もある。このような場合、フラッシュメモリコントローラに接続されているフラッシュメモリチップにアクセスしてアドレス情報を読み出し、アドレス変換テーブルをS R A M内に再構築する必要がある。このアドレス変換テーブルを再構築するためにはフラッシュメモリ1チップ当たり1~2秒程度は必要となる。このため、フラッシュメモリコントローラに多数のチップを接続している場合は、アドレス変換テーブルの再構築に長時間要することになる。低消費電力動作から通常動作への復帰にアドレス変換テーブルの再構築のための時間がかかることは、高速アクセスを必要とされるストレージ装置にとって許容できない。

10

【0011】

上述の特許文献1では、低消費電力化に対する考慮がない。さらに、アドレス変換テーブル内に起こりうる一般的なソフトエラーに対処できない。つまり、特許文献1に記載された技術は有効な論理ページアドレスが無効な論理ページアドレスにデータ化けした場合の不正マッピングを防止する技術であるが、一般的には有効な論理ページアドレスが無効な論理ページアドレスにデータ化けするとは限らない。例えば、有効な論理ページアドレスが別の有効な論理ページアドレスにデータ化けすることも考えられる。この場合、特許文献1の技術では、アドレス変換テーブルのエラーを検証することができない。

20

【0012】

本発明は、以上の点を考慮してなされたもので、低消費電力動作時にフラッシュメモリモジュール内のアドレス変換テーブルのデータを検証することにより、低消費電力動作から通常動作へ瞬時に復帰が可能なフラッシュメモリモジュール、そのフラッシュメモリモジュールを記録媒体として用いたストレージ装置及びそのフラッシュメモリモジュールのアドレス変換テーブル検証方法を提案しようとするものである。

30

【課題を解決するための手段】

【0013】

本発明は、フラッシュメモリコントローラと少なくとも一個のフラッシュメモリチップとからなり、フラッシュメモリコントローラは、フラッシュメモリチップの論理ページアドレスと物理ページアドレスを変換するアドレス変換テーブルを格納するメモリ備えるとともにフラッシュメモリコントローラは、データの読み書き時の第1の動作と、動作停止、電源電圧低下及び動作周波数低減のいずれかにより第1の動作より低消費電力で動作する待機状態時の第2の動作とを制御するフラッシュメモリモジュールにおいて、第2の動作時にアドレス変換テーブルのデータ検証を行うことを特徴とするものである。

【0014】

40

このフラッシュメモリモジュールにおいては、動作停止、電源電圧低下及び動作周波数低減のいずれかによる低消費電力動作時にアドレス変換テーブルのデータの検証を行うことにより、アドレス変換テーブル内のデータエラーを検出・訂正することができ、エラー訂正情報により訂正不可能なエラーの発生を防止して低消費電力動作から通常動作へ瞬時に復帰ができる。

【発明の効果】

【0015】

本発明によれば、低消費電力動作時にフラッシュメモリモジュール内のアドレス変換テーブルのデータを検証することにより、低消費電力動作から通常動作へ瞬時に復帰が可能なフラッシュメモリモジュール、そのフラッシュメモリモジュールを記録媒体として用い

50

たストレージ装置及びそのフラッシュメモリモジュールのアドレス変換テーブル検証方法を提供できる。

【発明を実施するための最良の形態】

【0016】

以下、本発明を適用したフラッシュメモリモジュール、ストレージ装置及びフラッシュメモリモジュールのアドレス変換テーブル検証方法についての一実施の形態について図面を参照して説明する。

【0017】

図1は、ストレージ装置100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からP33を備える。

10

【0018】

ストレージコントローラSCは、チャネルアダプタCA0、CA1、キャッシュメモリCM0、CM1、ストレージアダプタSA0、SA1及び相互接続網NW0、NW1を備える。なお、チャネルアダプタCA0、CA1、キャッシュメモリCM0、CM1及びストレージアダプタSA0、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。

【0019】

相互接続網NW0及び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、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。

20

【0020】

チャネルアダプタCA0は、図2で後述するが、チャネルC00、C01、C02、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、C11、C12、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、例えば、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。

【0021】

30

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

【0022】

ストレージアダプタSA0は、図3で後述するが、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～P03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～P13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～P23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～P33に接続されている。

40

【0023】

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～P03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～P13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～P23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～P33に接続されている。

【0024】

50

チャネルアダプタ C A 0、C A 1 及びストレージアダプタ S A 0、S A 1 は、保守端末 S V P に接続されている。保守端末 S V P は、ストレージ装置 1 0 0 の管理者から入力された設定情報を、チャネルアダプタ C A 0、C A 1 及び / 又はストレージアダプタ S A 0、S A 1 に送信する。

【 0 0 2 5 】

なお、ストレージ装置 1 0 0 は、ストレージアダプタ S A 0 及びチャネルアダプタ C A 0 に代わって、一つのアダプタを備えていても良い。この場合、当該一つのアダプタが、ストレージアダプタ S A 0 及びチャネルアダプタ C A 0 の処理を行う。

【 0 0 2 6 】

V D E V 0 ~ V D E V 3 は、R A I D (Redundant Arrays of Inexpensive Disks) グループである。例えば R A I D グループ V D E V 0 は、フラッシュメモリモジュール P 0 0、P 1 0、P 2 0、P 3 0 からなる。R A I D グループに属するフラッシュメモリモジュールの一つ、例えばフラッシュメモリモジュール P 0 0 でエラーが発生してデータを読み出せなくなった場合は、同じ R A I D グループに属する他のフラッシュメモリモジュール P 1 0、P 2 0、P 3 0 からデータをリビルドすることができる。

【 0 0 2 7 】

フラッシュメモリモジュール P 0 0 ~ P 3 3 はネットワーク N 0 0 を介してストレージアダプタ S A 0 と接続している。同様に、フラッシュメモリモジュール P 0 0 ~ P 3 3 はネットワーク N 0 1 を介してストレージアダプタ S A 1 と接続している。ストレージコントローラ S C は、ネットワーク N 0 0 及び N 0 1 を介して、フラッシュメモリモジュール P 0 0 ~ P 3 3 の電力制御を行う。また、ストレージアダプタ S A 0 及び S A 1 は、フラッシュメモリモジュール P 0 0 ~ P 3 3 の低消費電力動作の条件設定及び低消費電力動作状況の監視を行うことができる。

【 0 0 2 8 】

図 2 は、チャネルアダプタ C A 0 の構成を示すブロック図である。チャネルアダプタ C A 0 は、ホストチャネルインターフェース 2 1、キャッシュメモリインターフェース 2 2、ネットワークインターフェース 2 3、プロセッサ 2 4、ローカルメモリ 2 5 及びプロセッサ周辺制御部 2 6 を備える。

【 0 0 2 9 】

ホストチャネルインターフェース 2 1 は、チャネル C 0 0、C 0 1、C 0 2、C 0 3 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 2 1 は、チャネル C 0 0、C 0 1、C 0 2、C 0 3 上のデータ転送プロトコルと、ストレージコントローラ S C の内部のデータ転送プロトコルとを相互に変換する。

【 0 0 3 0 】

キャッシュメモリインターフェース 2 2 は、相互結合網 N W 0、N W 1 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 2 3 は、保守端末 S V P 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 2 1 とキャッシュメモリインターフェース 2 2 とは、信号線 2 7 によって接続されている。

【 0 0 3 1 】

プロセッサ 2 4 は、ローカルメモリ 2 5 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 2 4 は、上位装置（図示省略）と相互結合網 N W 0、N W 1 との間のデータ転送を制御する。

【 0 0 3 2 】

ローカルメモリ 2 5 は、プロセッサ 2 4 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 2 5 は、プロセッサ 2 4 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。

【 0 0 3 3 】

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 S V P に入力する。保守端末 S V P は、入力された情報を、ネットワークインターフェース

10

20

30

40

50

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、当該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。

【0034】

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。

【0035】

図3は、ストレージアダプタSA0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。

【0036】

キャッシュメモリインターフェース31は、相互結合網NW0、NW1と接続するインターフェースである。

【0037】

ストレージチャネルインターフェース32は、チャネルD00、D01、D02、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、D01、D02、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。

【0038】

ネットワークインターフェース33は、ストレージアダプタSA0と保守端末SVPとフラッシュメモリモジュールP00～P33とを接続するインターフェースである。

【0039】

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。

【0040】

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。

【0041】

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。

【0042】

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。

【0043】

図4は、フラッシュメモリモジュールP00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、フラッシュメモリコントローラMC、フラッシュメモリMEM及び電源(P.S.)408を備える。フラッシュメモリMEMはデータを記憶する。フ

10

20

30

40

50

ラッシュメモリコントローラMCは、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。電源408は、ラッシュメモリモジュールP00外部から供給された電力を、ラッシュメモリモジュールP00内部で使用する電圧に変換または調整して、ラッシュメモリコントローラMC、ラッシュメモリMEM等へ供給する。

【0044】

ラッシュ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(DKCI/F)402、内部バス(BUS)403、メモリ(RAM)404、メモリ(ROM)407、電力制御部(PCTL)409、ラッシュメモリインターフェース部(FMI/F)410及びデータ転送部(DMA)411を備える。

【0045】

ラッシュメモリMEMは、複数のラッシュメモリチップ405を備える。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で後述するが、ラッシュメモリコントローラMCがデータを一括消去する単位である。

【0046】

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。

【0047】

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、Serial-ATAコマンドやSCSIコマンドである。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。

【0048】

メモリ404は、例えば、SRAM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100の起動時にメモリ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。

【0049】

当該テーブルは、例えば、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、ラッシュメモリモジュールP00外から(例えばストレージアダプタSA0から)、ラッシュメモリMEMの読み書きする単位であるページに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、ラッシュメモリコントローラMCが、ラッシュメモリMEMの読み書きする単位であるページにアクセスするためのアドレスである。

【0050】

内部バス403は、プロセッサ401、インターフェース部402、メモリ404、メモリ407、データ転送部411及びラッシュメモリインターフェース部410を相互に接続し、データの転送路として機能する。

【0051】

電力制御部409は、ラッシュメモリコントローラMCの消費電力を制御する。電力制御部409は、ネットワークN00及びN01を介してストレージアダプタSA0及びSA1と接続している。

【0052】

10

20

30

40

50

フラッシュメモリインターフェース部 410 は、フラッシュメモリコントローラ MC とフラッシュメモリチップ MEM を接続するインターフェースである。

【0053】

データ転送部 411 は、プロセッサ 401 の命令により、フラッシュメモリチップ MEM に対する読み書きデータの転送を制御する。なお、データ転送部 411 の機能をプロセッサ 401 が実行する場合は、データ転送部 411 を省略することができる。

【0054】

プロセッサ 401 は、メモリ 404 に格納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 に格納されているフラッシュメモリ MEM の論理ページアドレスとフラッシュメモリ MEM の物理ページアドレスとのアドレスを変換するアドレス変換テーブル（図 7 を参照して後述する。）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。

【0055】

図 7 は、メモリ 404 に格納するアドレス変換テーブルを示す図である。論理ページアドレス 701 と、各論理ページアドレス 701 に対応する物理ページアドレス 702 と、エラー訂正情報 703 が、アドレス変換テーブル T1 を構成する。エラー訂正情報 703 は、例えばハミングコード (Hamming code) であり、アドレス変換テーブル T1 に格納した論理ページアドレス 701 と物理ページアドレス 702 とエラー訂正情報 703 により発生したエラーの検出と訂正が可能である。

【0056】

また、プロセッサ 401 は、フラッシュメモリモジュール P00 内のブロック 406 に対して、後述するリクラーメーション処理（ブロック再生処理）及び後述するウェアアレベリング処理（消去回数平準化処理）を行う。

【0057】

なお、他のフラッシュメモリモジュール P01 ~ P33 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。

【0058】

図 5 は、フラッシュメモリモジュール P00 のブロック 406 の説明図である。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 64 ページ等）を含む。

【0059】

ページ 501 は、フラッシュメモリコントローラ MC 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、フラッシュメモリコントローラ MC は、(20 ~ 30 μ s 弱 / ページ) の速度でデータを読み出し、(0.2 ~ 0.3 ms / ページ) の速度でデータを書き込む。また、フラッシュメモリコントローラ MC は、(2 ~ 4 ms / ブロック) の速度でデータを消去する。

【0060】

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、ページ当たり容量が 2112 バイトであり、2048 バイトのデータ部 502 及び 64 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。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。

【0061】

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。

【0062】

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。

【0063】

10

20

30

40

50

ページ内冗長部 503 は、通常、フラッシュメモリコントローラ MC のみがアクセス可能であり、ストレージアダプタ SA0、SA1 からは、ページデータ部 502 のみがアクセス可能である。言い換えると、論理ページアドレスはページ内データ部 502 のメモリ空間をマッピングする。

【0064】

なお、フラッシュメモリモジュール P00 は、特性上、ブロック 406 内のデータ「1」を「0」に書き換えることはできるが、「0」を「1」に書き換えることができない。つまり、データを直接書き換えることができない。フラッシュメモリコントローラ MC は、ブロック内の未使用ページにデータを書き込む。未使用ページが無くなり書き込みができなくなったブロックの有效データを他のブロックへ退避させ、該ブロックを消去して再びデータを書き込める状態にすることを、リクラメーション処理（ブロック再生処理）という。

【0065】

このように、フラッシュメモリモジュール P00 におけるデータの書換えは、ブロック 406 の消去が伴う。しかし、ブロック 406 を 1 つ消去するのにかかる時間は、1 ページ 501 の書き込みにかかる時間に比べて約 1 桁長い。従って、ページ 501 のデータの書き換えのために毎回ブロック 406 を 1 つ消去すると、フラッシュメモリモジュール P00 のデータ書換え性能は悲観的に悪くなる。つまりフラッシュメモリモジュール P00 では、消去時間を隠蔽できるアルゴリズムでデータを書き込むことが必要である。

【0066】

また、フラッシュメモリモジュール P00 において、上記ブロック 406 を消去する回数には制限がある。例えば、ブロック 406 は、1 つ当たりの消去回数は、10 万回までしか保証されていない。したがって、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 406 は、データ消去ができなくなり使用不能となる。このため、フラッシュメモリモジュール P00 を記録媒体として使用するストレージ装置 100 では、特定のブロック 406 に消去が集中しないように消去回数平準化処理が行われている。

【0067】

以上述べた消去時間隠蔽と消去回数平準化のため、フラッシュメモリモジュール P00 内ではデータ書き込みの際に、論理ページアドレスから物理ページアドレスにアドレスを変換するアドレス変換処理が行われている。具体的には、フラッシュメモリコントローラ MC は、特定の物理ページアドレスに書き込みが集中するのをさけるため、上位装置（図示省略）からデータ書き込み先アドレスとして受信した論理ページアドレスを、アドレス変換テーブル T1 を参照して実際のフラッシュメモリチップ 405 での書き込み先アドレスである物理ページアドレスへ変換している。このため、フラッシュメモリモジュール P00 では、アドレス変換テーブル T1 のデータ保証が常に求められている。

【0068】

上述のリクラメーション処理やウエアレベリング処理のため、フラッシュメモリモジュール P00 に書き込んだデータは、ストレージコントローラ SC からの指示とは独立に、フラッシュメモリモジュール P00 内で移動することがある。このデータ移動の結果を、フラッシュメモリコントローラ MC が前述の論理ページアドレスと物理ページアドレスのアドレス変換テーブルに正しく反映しないと、ストレージコントローラ SC が正しいデータにアクセスできなくなる。

【0069】

ところが、放射線等によるアドレス変換テーブル T1 のデータ化けにより、アドレス変換テーブル T1 のデータにソフトエラーが生じると、フラッシュメモリモジュール P00 内において、データを書き込む先あるいは読み出し先のアドレスにエラーが発生することになり、ストレージコントローラ SC が正しいデータを読み出すことができなくなる。

【0070】

次にフラッシュメモリコントローラ MC の消費電力制御について説明する。図 6 は、フラッシュメモリコントローラ MC の消費電力制御に関わる構成要素の説明図である。電力

10

20

30

40

50

制御部 409 は、低消費電力化時の動作条件を設定するレジスタ 606 を有する。レジスタ 606 に格納するデータは後述する(図8、図9)。レジスタ 606 に格納するデータは、ネットワーク N00 と N01 を介して、ストレージアダプタ SA0 と SA1 から設定及び監視が可能である。

【0071】

電源 408 は、レジスタ 606 に格納された設定条件を、信号線 604 を介して読み取り、その設定条件に従ってフラッシュメモリコントローラ MC 内のプロセッサ 401、インターフェース部 402、メモリ 404、メモリ 407、フラッシュメモリインターフェース部 410、データ転送部 411 へ、チップ電源配線 601 を介して電力を供給する。また電源 408 は、基板バイアス制御部 608 へ、チップ電源配線 603 を介して基板バイアス用電力を供給する。そして、基板バイアス制御部 608 は、レジスタ 606 に格納された設定条件を、信号線 607 を介して読み取り、その設定条件に従ってフラッシュメモリコントローラ MC 内のプロセッサ 401、インターフェース部 402、メモリ 404、メモリ 407、フラッシュメモリインターフェース部 410、データ転送部 411 へ、チップ電源配線 605 を介して基板バイアス用電力を分配する。

10

【0072】

クロック制御部 615 は、レジスタ 606 に格納された設定条件を、信号線 611 を介して読み取り、その設定条件に従って信号線 616 から入力した基準クロック信号を変換し、フラッシュメモリコントローラ MC 内のプロセッサ 401、インターフェース部 402、電力制御部 409、フラッシュメモリインターフェース部 410、データ転送部 411 へ、信号線 612、613、614 を介して変換したクロック信号を分配する。

20

【0073】

プロセッサ 401 は、電力制御部 409 に対して割込み許可あるいは不許可を判定するための信号を、信号線 610 を介して伝送する。また、電力制御部 409 は、信号線 609 を介してプロセッサ 401 に対して動作停止信号あるいは動作再開信号を伝送する。

【0074】

インターフェース部 402 は、インターフェース部 402 に対する低消費電力化コマンド、例えば Serial - ATA の S l u m b e r コマンド等を受信した場合に、信号線 602 を介して低消費電力化コマンド受信を電力制御部 409 へ通知する。

30

【0075】

図8と図9は、電力制御部(P C T L) 409 内のレジスタ 606 に格納するデータを説明する図である。電力制御部 409 内のレジスタ 606 は、詳細にはレジスタ A(図8)とレジスタ B(図9)の2種類がある。

【0076】

図8は、レジスタ A に格納するデータを説明する図である。レジスタ A は、主に、フラッシュメモリモジュール P00 ~ P33 の低消費電力動作への移行条件を制御するパラメータを格納する。パラメータ 801 は、フラッシュメモリモジュール P00 ~ P33 が自発的に低消費電力動作に移行するか、あるいはストレージアダプタ SA0, SA1 等からのコマンド受信により低消費電力動作に移行するか、あるいはそれら両方を低消費電力化動作の契機とするか、を設定する。

40

【0077】

パラメータ 802 は、低消費電力動作時に、アドレス変換テーブル T1 のデータ検証を行うか否かを設定する。アドレス変換テーブル T1 のデータ検証を行わない場合は、データ検証を行う場合よりも低消費電力となるが、通常動作に復帰する際には、フラッシュメモリチップ 405 からのアドレス情報再読み込みが必要となる。

【0078】

パラメータ 803 は、パラメータ 801 において低消費電力動作への移行を自発モードと設定した場合の、移行条件を格納する。例えば、パラメータ 803 に格納した時間内にフラッシュメモリモジュール P00 ~ P33 へのアクセスが無く、且つパラメータ 801 が自発モードに設定されていた場合に、フラッシュメモリモジュール P00 ~ P33 は自

50

動的に低消費電力動作へ移行する。

【0079】

パラメータ804は、現状の動作が通常動作あるいは低消費電力動作かを示す。パラメータ804を調べることにより、ストレージアダプタSA0, SA1等は、フラッシュメモリモジュールP00～P33の動作状態を知ることができる。また、パラメータ805は、アドレス変換テーブルT1に発生した、エラー訂正情報703だけでは訂正不能なエラーの発生回数を示す。

【0080】

図9は、レジスタBに格納するデータを説明する図である。レジスタBは、フラッシュメモリモジュールP00～P33の低消費電力時の動作条件を制御するパラメータを格納する。レジスタBにおける項目901から906は、フラッシュメモリコントローラMC内部の電力制御対象を示す。また、項目907から909は、低消費電力化のために変化させる動作条件を示す。項目907と908は、低消費電力化時にプロセッサ401、インターフェース部402、メモリ404、メモリ407、フラッシュメモリインターフェース部410、データ転送部411へ供給する電源電圧と基板バイアスの条件を示す。電源電圧と基板バイアスに関しては後述する。項目909は、低消費電力化時のプロセッサ401、インターフェース部402、メモリ404、メモリ407、フラッシュメモリインターフェース部410、データ転送部411の動作周波数を示す。動作周波数は低くするほど低消費電力化効果が大きいが、動作を停止させる場合には周波数を“0”と設定する。項目902と903は、メモリ(RAM404, ROM407)に対する制御パラメータである。メモリ(RAM404, ROM407)に対しては、アクセスする信号の周波数で動作周波数が決まるため、低消費電力化時の動作周波数の設定は不要である。

10

20

30

【0081】

図10は、電源電圧と基板バイアスの説明図である。インバータ回路INVには、正のチップ内電源VDD1あるいはVDD2と、負のチップ内電源VSS1あるいはVSS2と、正の基板バイアスVSUB1あるいはVSUB2と、負の基板バイアスVSUB3あるいはVSUB4を供給する。ここで、VDD1、VSS1、VSUB1、VSUB3は、通常動作時に回路へ供給する電源電圧を表すこととする。そしてVDD2、VSS2、VSUB2、VSUB4は、低消費電力動作時に回路へ供給する電源電圧を表すこととする。

【0082】

例えば、通常動作時は、VDD1=VSUB1、VSS1=VSUB3となる所定の電圧で電力を回路へ供給する。それに対して低消費電力動作時は、VDD1>VDD2かつVDD2<VSUB2、VSS2>VSS1かつVSS2>VSUB4となる条件の電圧で電力を供給する。電源電圧低下は回路の動作電力を低下させ、適切な基板バイアス印加は、リーク電流低減による消費電力を低下させる効果がある。

【0083】

図11は、ストレージアダプタSA0あるいはSA1内で管理するフラッシュメモリモジュール管理テーブルT2を説明する図である。ストレージアダプタSA0とSA1は、ネットワークN00またはN01を介してフラッシュメモリモジュールP00～P33へアクセスし、レジスタA、Bの内容を監視または設定変更することができる。図11のフラッシュメモリモジュール管理テーブルT2は、各フラッシュメモリモジュールP00～P33のレジスタA、B内に格納したパラメータを管理するため、フラッシュメモリモジュール識別子1001、各フラッシュメモリモジュールのレジスタAの内容1002、及び各フラッシュメモリモジュールのレジスタBの内容1003を格納する。

40

図12は、ストレージアダプタSA0あるいはSA1内で管理するRAIDグループ管理テーブルT3を説明する図である。

【0084】

図12のRAIDグループ管理テーブルT3は、RAIDグループ識別子1201と、各RAIDグループを構成するフラッシュメモリモジュール識別子の集合1202を格納

50

する。このR A I D グループ管理テーブルT 3は、図17、図18を用いて説明するフローチャートにおいて使用する。

【0085】

図13は、フラッシュメモリモジュールP 0 0の低消費電力動作を説明するフローチャートである。なお、他のフラッシュメモリモジュールP 0 1～P 3 3については、フラッシュメモリモジュールP 0 0と同様な動作を行うために説明を省略する。図13のフローチャートでは、フラッシュメモリモジュールP 0 0の低消費電力動作への移行条件が、外部からのコマンド受信を契機とする場合を示す。

【0086】

ステップS T 1 0 1において、インターフェース部4 0 2は、例えばストレージアダプタS A 0あるいはS A 1から、低消費電力化コマンドを受信する。このコマンドは、例えばS - A T AのS l u m b e rコマンドである。インターフェース部4 0 2は、低消費電力化コマンド受信を電力制御部4 0 9へ通知する。

【0087】

ステップS T 1 0 2において、電力制御部4 0 9は、プロセッサ4 0 1が割込み許可状態か否かを信号線6 1 0によりチェックする。

【0088】

ステップS T 1 0 3において、割込み不許可ならば(S T 1 0 3: N O)、再びステップS T 1 0 2に戻り割込み許可を待つ。割込み不許可の要因としては、例えば、プロセッサ4 0 1がフラッシュメモリチップ4 0 5の消去回数平準化処理を実行中であることがある。なぜならば、プロセッサ4 0 1が読み書きのアクセスが無い期間において、消去回数平準化処理を実行する場合があるからである。本実施の形態のフラッシュメモリモジュールP 0 0は、消去回数平準化処理の実行中は、低消費電力動作へ移行せず、消去回数平準化処理完了まで待機する。すなわち、以下で説明するように、フラッシュメモリモジュールP 0 0は、消去回数平準化処理が完了した後、通常動作から低消費電力動作への移行及びアドレス変換テーブルT 1のデータ検証が行われる。

【0089】

ステップS T 1 0 3において、プロセッサ4 0 1が割込み許可であった場合は(S T 1 0 3: Y E S)、ステップS T 1 0 4に進む。ステップ1 2 0 4において、電力制御部4 0 9はプロセッサ4 0 1に対して動作停止指示を、信号線6 0 9を介して送信する。

【0090】

ステップS T 1 0 5において、電力制御部4 0 9は、レジスタ6 0 6の設定内容に従い、電源電圧、基板バイアス電圧、クロック周波数を変更する。

【0091】

ステップS T 1 0 6において、プロセッサ4 0 1はアドレス変換テーブルT 1のデータ検証を実行する。このステップS T 1 0 6の内容の詳細は図14のフローチャートを用いて後述する。

【0092】

ステップS T 1 0 7において、インターフェース部4 0 2が低消費電力動作から通常動作への復帰コマンド(アクティヴコマンド)を受信したか否かをチェックする。アクティヴコマンドを受信した場合は(S T 1 0 7: Y E S)、ステップS T 1 0 9へ進む。

【0093】

ステップS T 1 0 9において、電力制御部4 0 9は電源電圧、基板バイアス電圧、クロック周波数を通常動作のための状態へ戻す。ステップS T 1 1 0において、電力制御部4 0 9はプロセッサ4 0 1に対して動作再開を指示する。

【0094】

ステップS T 1 0 7においてアクティヴコマンドの受信がなかった場合は(S T 1 0 7: N O)、ステップS T 1 0 8において、所定時間待機する。所定時間内はステップS T 1 0 7を繰り返し、一定の周期でステップS T 1 0 6のアドレス変換テーブルT 1のデータ検証を実行する。電力制御部4 0 9の消費電力がプロセッサ4 0 1の消費電力より十分

10

20

30

40

50

小さい場合に、プロセッサ401によるアドレス変換テーブルT1のデータチェックのための動作期間の割合を少なくすることで、フラッシュメモリコントローラMCの低消費電力化を図ることができる。

【0095】

図15は、アドレス変換テーブルT1のデータ検証の低デューティ実行を説明するタイムチャートである。プロセッサ401は、アドレス変換テーブルT1のデータ検証を一部分ずつ(1/n)、テーブル検証期間より停止期間の方が長くなるように、つまり、低デューティで実行する。

【0096】

図15に示すように、プロセッサ401がアドレス変換テーブルT1のデータ検証を所定範囲ずつ行う周期をTとする。そして、その周期Tの内、プロセッサ401がアドレス変換テーブルT1の所定範囲をデータ検証する期間をtとする。周期Tと期間t、あるいは周期Tと所定範囲は、以下の3つの条件が成立するように設定する。

【0097】

先ず、1つ目の条件である周期Tの設定条件について説明する。本実施の形態のフラッシュメモリモジュールP00～P33が、アドレス変換テーブルT1のデータ検証を伴う低消費電力動作から通常動作へ移行するのに要する時間は、周期Tと、プロセッサ401動作時にフラッシュメモリモジュールP00～P33がアクティヴコマンドを受信してから通常動作へ移行するのに要する時間の和である。その和が、ストレージ装置100の仕様として許容できる待ち時間以下となるように、周期Tを設定する。

【0098】

次に、2つ目の条件である期間tあるいは期間t内データ検証範囲の設定条件aについて説明する。プロセッサ401が所定範囲のデータ検証に要する時間と消費する電力を考慮して、アドレス変換テーブルT1のデータの検証を伴う低消費電力動作中の平均消費電力がストレージ装置100として許容できる消費電力以下となるように、周期Tと期間t、あるいは周期Tと期間t内にプロセッサ401が実行するデータ検証の範囲を決定する。

【0099】

次に、3つ目の条件である期間tあるいは期間t内データ検証範囲の設定条件bについて説明する。ストレージ装置100に搭載した全てのフラッシュメモリモジュールP00～P33内のアドレス変換テーブルT1に、統計的に1ビットのソフトエラーが起こりうると予想される期間内に、アドレス変換テーブルT1内の全データの検証を少なくとも1回完了するように、周期Tと期間t、あるいは周期Tと期間t内にプロセッサ401が実行するデータ検証の範囲を決定する。

【0100】

図14は、本発明の実施の形態のフラッシュメモリモジュールP00が、低消費電力動作中に行うアドレス変換テーブルT1のデータ検証(ステップST106)を説明するフローチャートである。

【0101】

ステップST201において、電力制御部409はプロセッサ401とメモリ404の動作条件を通常動作に変更し、プロセッサ401へ動作再開を指示する。なお、プロセッサ401とメモリ404が低消費電力動作のままでもアドレス変換テーブルT1のデータ検証が可能ならば、動作条件を変更しなくても良い。ステップST202において、プロセッサ401はデータ検証のアドレス変換テーブルT1内データへのポインタを、レジスタへセットする。

【0102】

ステップST203においてテーブルに格納されている論理ページアドレスと物理ページアドレスとエラー訂正情報をチェックし、ステップST204においてエラーの有無を判定する。このステップST204においてエラーが無い場合は(ST204: NO)ステップST208へ進む。ステップST208において、所定のデータ量のチェックが完

10

20

30

40

50

了した場合は (S T 2 0 8 : Y E S) ステップ S T 2 0 9 へ進む。

【 0 1 0 3 】

本実施の形態のフラッシュメモリモジュール P 0 0 では、一度にアドレス変換テーブル T 1 内全てのデータを連続して検証するのではなく、複数回に分けてデータ検証を行うことにより、平均的な消費電力を低減している。ステップ S T 2 0 9 において、電力制御部 4 0 9 はプロセッサ 4 0 1 へ対して、動作停止を指示する。このとき、ステップ S T 2 0 1 においてプロセッサ 4 0 1 とメモリ 4 0 4 を通常動作へ変更していた場合は、再びプロセッサ 4 0 1 とメモリ 4 0 4 を低消費電力動作へ変更する。

【 0 1 0 4 】

ステップ S T 2 0 8 において所定量のデータ検証が完了していない場合は (S T 2 0 8 : N O) 、ステップ S T 2 1 0 において、次のアドレス変換テーブル T 1 内のデータを指すポインタをレジスタにセットし、ステップ S T 2 0 3 に戻る。

【 0 1 0 5 】

ステップ S T 2 0 4 においてエラーが検出された場合は (S T 2 0 4 : Y E S) 、ステップ S T 2 0 5 においてエラー訂正可能性を判定する。データに付加しているエラー訂正情報からエラーが訂正可能な場合は (S T 2 0 5 : Y E S) 、ステップ S T 2 0 7 へ進み、エラー訂正を実行する。

【 0 1 0 6 】

エラー訂正情報からではエラー訂正不能な場合は、例えばエラー発生ビット数が多い場合は (S T 2 0 5 : N O) 、ステップ S T 2 0 6 のテーブル再構成処理を実行する。そして、処理は、ステップ S T 2 0 2 へ戻る。

【 0 1 0 7 】

次に、テーブル再構成処理 (ステップ S T 2 0 6) について説明する。図 1 6 はテーブル再構成処理を説明するフローチャートである。

【 0 1 0 8 】

ステップ S T 3 0 1 において、電力制御部 4 0 9 はレジスタ A のパラメータ 8 0 5 の、訂正不能エラー発生回数をインクリメントする。そして、ステップ S T 3 0 2 において、パラメータ 8 0 2 の値が所定の値以下であるか否かを判定する。エラー発生回数が所定回数に達した場合は (S T 3 0 2 : Y E S) 、処理はステップ S T 3 0 3 へ進む。

【 0 1 0 9 】

ステップ S T 3 0 3 は、エラー発生回数が所定回数に達したフラッシュメモリモジュール (障害モジュール) を交換し、スペアのフラッシュメモリモジュール (図示を省略) にデータをリビルドする。このリビルドは、障害モジュールと同じ R A I D グループに属するフラッシュメモリモジュールのデータを、ストレージアダプタ S A 0 又は S A 1 が読み出して排他的論理和計算を行い、障害モジュールに格納されていたデータを再生する処理である。

【 0 1 1 0 】

また、スペアのフラッシュメモリモジュールは、予めストレージアダプタ S A 0 、 S A 1 に接続しておくと、リビルドを早く実行できる。スペアのフラッシュメモリモジュールは、ストレージアダプタ S A 0 、 S A 1 へ接続した際に、そのフラッシュメモリモジュール内のフラッシュメモリコントローラにアドレス変換テーブル T 1 を作成しておき、作成後は図 1 3 のアドレス変換テーブル T 1 のデータ検証 (S T 1 0 6 , 図 1 4) の実行を伴う低消費電力動作で待機状態にしておく。このようにスペアのフラッシュメモリモジュールを待機状態にしておくことにより、必要になったときに直ぐに使用可能となる。

【 0 1 1 1 】

ステップ S T 3 0 2 でエラー回数が所定回数以下のは場合は (S T 3 0 2 : N O) 、処理はステップ S T 3 0 4 へ進む。ステップ S T 3 0 4 において、プロセッサ 4 0 1 は、フラッシュメモリモジュール P 0 0 の先頭物理ページアドレスへのポインタをレジスタにセットする。

【 0 1 1 2 】

10

20

30

40

50

ステップST305において、プロセッサ401は該当物理ページアドレスから論理ページアドレス情報を読み出し、アドレス変換テーブルT1へアドレス情報を格納する。

【0113】

ステップST307において、プロセッサ401のポインタが、最終物理ページアドレスに達したかを判定する。ステップST307において最終物理ページアドレスへ達していない場合は(ST307: NO)、ステップST306で、次の物理ページアドレスにアクセスするため、ポインタを進める。

【0114】

ステップST307において、最終物理ページアドレスまでアドレス調査が完了した場合は(ST307: YES)、処理を完了する。

10

【0115】

また、レジスタAのパラメータ805のエラー回数が所定回数に達する場合としては、過去のデータ検証結果を含む累積で所定の回数に達する場合と、あるいは一回のチェックで複数のエラーが発生し所定回数に達する場合が考えられる。

【0116】

次に、本実施の形態のストレージ装置100が、RAIDグループ(VDEV0~VDEV3)内の全フラッシュメモリモジュールに対して低消費電力動作へ移行させる方法を説明する。図17はこの方法を説明するフローチャートである。

【0117】

ストレージアダプタSA0又はSA1は、RAIDグループ内の複数フラッシュメモリモジュール(例えば、VEDV0: フラッシュメモリモジュールP00, P10, P20, P30)に対して、ほぼ同期してデータの読み書きを実行する。従って、低消費電力化するか否かもRAIDグループ内のフラッシュメモリモジュールは同期する。例えば、フラッシュメモリモジュールP00の低消費電力化が自発モードに設定されていた場合を考える。

20

【0118】

ステップST401においてストレージアダプタSA0又はSA1は、ある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の低消費電力動作への移行を検出する。これは、ストレージアダプタSA0, SA1が各フラッシュメモリモジュールP00~P33のレジスタAを監視することにより可能となる。

30

【0119】

ステップST402においてストレージアダプタSA0又はSA1は、図12で説明したRAIDグループ管理テーブルT3にアクセスして、低消費電力動作へ移行したフラッシュメモリモジュールP00と同じRAIDグループ(VDEV0)に属するフラッシュメモリモジュールP10, P20, P30を調査する。

【0120】

ステップST403において、該当RAIDグループ(VDEV0)内で低消費電力動作へ移行していないフラッシュメモリモジュールの有無を判定する。

【0121】

ステップST403において、低消費電力動作へ移行していないフラッシュメモリモジュールがある場合は(ST403: YES)、ステップST404へ進み、ストレージアダプタSA0又はSA1は、該当RAIDグループ内の低消費電力動作へ移行していないフラッシュメモリモジュールに対して、低消費電力化コマンドを発行する。そして、処理はステップST403へ戻る。

40

【0122】

ステップST403において、低消費電力動作へ移行していないフラッシュメモリモジュールがない場合、すなわち、RAIDグループ(VDEV0)内全てのフラッシュメモリモジュールP10, P20, P30が低消費電力動作へ移行した場合は(ST403: NO)、処理を終了する。

【0123】

50

低消費電力動作時から通常動作へ復帰する場合についても、R A I D グループ (V D E V 0 ~ V D E V 3) 内のフラッシュメモリモジュールは同期する。従って、図 17 と実質的に同様のフローチャートで、R A I D グループ内全てのフラッシュメモリモジュールの通常動作への復帰が可能となる。

【 0 1 2 4 】

続いて、R A I D グループ単位で低消費電力動作から通常動作へ移行する方法について説明する。図 18 はこの方法を示すフローチャートである。

【 0 1 2 5 】

ステップ S T 5 0 1 においてストレージアダプタ S A 0 又は S A 1 は、R A I D グループのあるフラッシュメモリモジュール (例えは、V E D V 0 : フラッシュメモリモジュール P 0 0) に対して低消費電力動作から通常動作への移行を指示する。つまり、ストレージアダプタ S A 0 又は S A 1 がフラッシュメモリモジュール P 0 0 へアクティヴコマンドを発行する。

【 0 1 2 6 】

ステップ S T 5 0 2 において、ストレージアダプタ S A 0 又は S A 1 は、図 12 で説明した R A I D グループ管理テーブル T 3 にアクセスして、通常動作への移行を指示したフラッシュメモリモジュールと同じ R A I D グループに属するフラッシュメモリモジュールを調査する。例えは、フラッシュメモリモジュール P 0 0 が通常動作への移行を指示されたのであれば、フラッシュメモリモジュール P 1 0 , P 2 0 , P 3 0 が調査される。

【 0 1 2 7 】

ステップ S T 5 0 3 において、当該 R A I D グループ内で低消費電力動作から通常動作へ移行していないフラッシュメモリモジュールの有無を判定する。

【 0 1 2 8 】

ステップ S T 5 0 3 において、当該 R A I D グループ内で低消費電力動作から通常動作へ移行していないフラッシュメモリモジュールがある場合は (S T 5 0 3 : Y E S) 、ステップ S T 5 0 4 において、ストレージアダプタ S A 0 又は S A 1 は、該当 R A I D グループ内の低消費電力動作から通常動作へ移行していないフラッシュメモリモジュールに対してアクティヴコマンドを発行する。そして、処理はステップ S T 5 0 3 へ戻る。

【 0 1 2 9 】

ステップ S T 5 0 3 において、当該 R A I D グループ内で低消費電力動作から通常動作へ移行していないフラッシュメモリモジュールがない場合は、すなわち、R A I D グループ内全てのフラッシュメモリモジュール P 1 0 , P 2 0 , P 3 0 が通常動作へ移行した場合は (S T 5 0 3 : N O) 、処理を終了する。

【 0 1 3 0 】

このように、R A I D グループ内の複数のフラッシュメモリモジュールを同期して低消費電力化することにより、ストレージアダプタ S A 0 又は S A 1 は R A I D グループ単位でフラッシュメモリモジュールの電力を管理できるので、消費電力低減効果が大きく、またストレージアダプタ S A 0 又は S A 1 の管理付加が軽減する。

【 0 1 3 1 】

以上説明したように、本実施の形態のフラッシュメモリモジュール P 0 0 ~ P 3 3 は、フラッシュメモリコントローラ M C とフラッシュメモリチップ 4 0 5 からなり、フラッシュメモリコントローラ M C は、フラッシュメモリチップ 4 0 5 の論理ページアドレスと物理ページアドレスを変換するアドレス変換テーブル T 1 を格納するメモリ (S R A M) 4 0 4 を備え、フラッシュメモリコントローラ M C は、データの読み書き時の通常動作である第 1 の動作と、動作停止、電源電圧低下及び動作周波数低減のいずれかによりデータ読み書き時の通常動作より低消費電力で動作する待機状態時の低消費電力動作とを制御する。

【 0 1 3 2 】

このフラッシュメモリモジュール P 0 0 ~ P 3 3 は、低消費電力動作時にアドレス変換テーブル T 1 のデータの検証を行うことにより、低消費電力動作時にアドレス変換テーブ

10

20

30

40

50

ル T 1 内のデータに発生したソフトエラーを検出・訂正することができる。すなわち、低消費電力動作時にフラッシュメモリモジュール P 0 0 ~ P 3 3 の低消費電力化とフラッシュメモリコントローラ M C 内のアドレス変換テーブル T 1 のデータ保護が可能となる。このため、低消費電力動作から通常動作に復帰する際に、エラーが蓄積した場合に必要となるアドレス変換テーブル T 1 の再構成を行う必要がなくなり、低消費電力動作しているフラッシュメモリモジュール P 0 0 ~ P 3 3 が通常動作へ瞬時に復帰することができる。

【 0 1 3 3 】

また、フラッシュメモリコントローラ M C は、アドレス変換テーブル T 1 のデータ検証を行うデータ検証期間 t と、フラッシュメモリコントローラ M C の動作を停止させる動作停止期間を交互に実行し、データ検証期間 t を動作停止期間より短くしている。このため 10 、データの検証は低デューティで実行される。

【 0 1 3 4 】

また、スペアのフラッシュメモリモジュールをストレージアダプタ S A 0 又は S A 1 を接続させておくことにより、フラッシュメモリモジュール P 0 0 ~ P 3 3 のいずれかにエラー訂正情報によってはエラー訂正不可能なフラッシュメモリモジュール（障害モジュール）が生じた場合にも、同じ R A I D グループの他のフラッシュメモリモジュールからデータをスペアのフラッシュメモリモジュールに移行してリビルドを行うことができ、直ぐに通常動作へ復帰することができる。

【 0 1 3 5 】

なお、上述の実施の形態においては、本発明を図 5 及び図 6 のように構成されたフラッシュメモリモジュール P 0 0 ~ P 3 3 及びそのフラッシュメモリモジュール P 0 0 ~ P 3 1 を用いたストレージ装置 1 0 0 に適用するようにして説明した場合について述べたが、本発明はこれに限られず、この他種々の構成のフラッシュメモリモジュール、そのフラッシュメモリモジュールを記録媒体として用いたストレージ装置及びそのフラッシュメモリモジュールのアドレス変換テーブル検証方法に広く適用することができる。 20

【 産業上の利用可能性 】

【 0 1 3 6 】

本発明は、種々のフラッシュメモリモジュール、そのフラッシュメモリモジュールを記録媒体として用いたストレージ装置及びそのフラッシュメモリモジュールのアドレス変換テーブル検証方法に広く適用することができる。 30

【 図面の簡単な説明 】

【 0 1 3 7 】

【 図 1 】本発明に係わるストレージ装置の構成を示すブロック図。

【 図 2 】本発明に係わるチャネルアダプタの構成を示すブロック図。

【 図 3 】本発明に係わるストレージアダプタの構成を示すブロック図。

【 図 4 】本発明に係わるフラッシュメモリモジュールの構成を示すブロック図。

【 図 5 】本発明に係わるフラッシュメモリモジュールにおけるブロックの説明図。

【 図 6 】本発明に係わるフラッシュメモリコントローラの消費電力制御に係わる構成要素の説明図。

【 図 7 】本発明に係わるアドレス変換テーブルを示す図。 40

【 図 8 】本発明に係わるレジスタに格納するデータの一例を示す図。

【 図 9 】本発明に係わるレジスタに格納するデータの一例を示す図。

【 図 1 0 】本発明に係わる電源電圧と基板バイアスの説明図。

【 図 1 1 】本発明に係わるフラッシュメモリモジュール管理テーブルを示す図。

【 図 1 2 】本発明に係わる R A I D グループ管理テーブルを示す図。

【 図 1 3 】本発明に係わるフラッシュメモリモジュールの低電力動作を説明するフローチャート。

【 図 1 4 】本発明に係わる低電力動作中に行うアドレス変換テーブルのデータ検証を説明するフローチャート。

【 図 1 5 】本発明に係わるアドレス変換テーブルのデータ検証の低デューティ実行を説明 50

する図。

【図16】本発明に係わるテーブル再構成処理について説明するフローチャート。

【図17】本発明に係わるRAIDグループ内の全てのフラッシュメモリモジュールを低消費電力動作へ移行させる方法を説明するフローチャート。

【図18】本発明に係わるRAIDグループ内の全てのフラッシュメモリモジュールを通常動作へ移行させる方法を説明するフローチャート。

【符号の説明】

【0138】

SC...ストレージコントローラ、

CA0、CA1...チャネルアダプタ

10

CM、CM0、CM1...キャッシュメモリ

SA、SA0、SA1...ストレージアダプタ

P00、P01、P02、P03、P10、P11、P12、P13、P20、P21、
P22、P23、P30、P31、P32、P33...フラッシュメモリモジュール

MC...フラッシュメモリコントローラ

MEM...フラッシュメモリ

C00、C01、C02、C03、C10、C11、C12、C13、D00、D01、
D02、D03、D10、D11、D12、D13...チャネル

VDEV0、VDEV0、VDEV0、VDEV0...RAIDグループ

N00、N01...ネットワーク

20

NW0、NW1...相互結合網

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

30

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...内部バス

40

404、407...メモリ

405...フラッシュメモリチップ

406...ブロック

408...電源

409...電力制御部

410...フラッシュメモリインターフェース部

411...データ転送部

501...ページ

502...ページデータ部

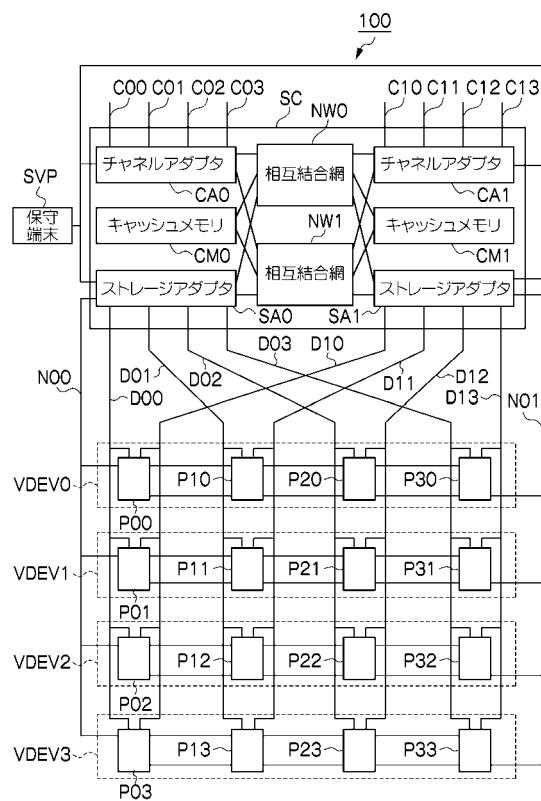
503...ページ冗長部

50

6 0 1、6 0 3、6 0 5...チップ電源配線
6 0 2、6 0 4、6 0 7、6 0 9、6 1 0、6 1 1、6 1 2、6 1 3、6 1 4...信号線
6 0 8...基板バイアス制御部
6 0 6...レジスタ
6 1 5...クロック制御部
6 1 6...基準クロック配線
V D D、V S S...チップ内電源
V S U B 1、V S U B 2...基板バイアス
I N V...インバータ回路
T 1...アドレス変換テーブル
T 2...フラッシュメモリモジュール管理テーブル
T 3...R A I D グループ管理テーブル

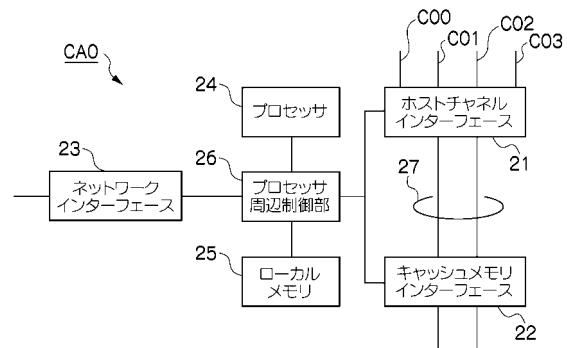
【 四 1 】

1



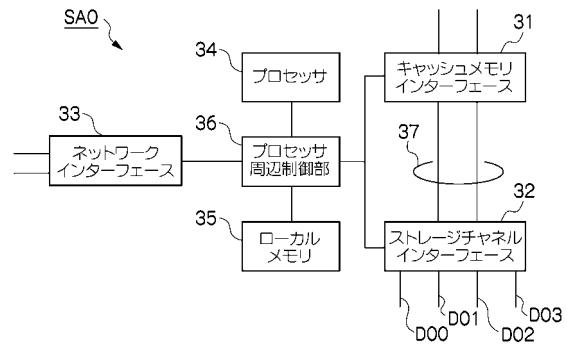
【図2】

2



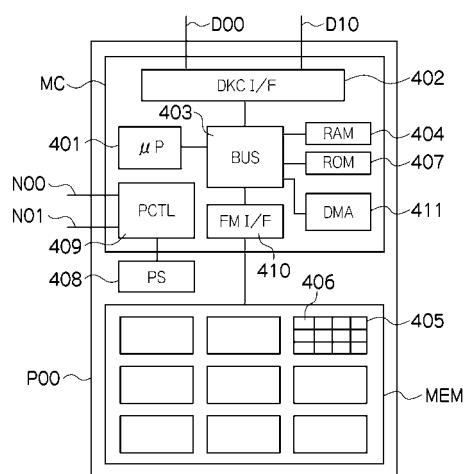
【 义 3 】

3



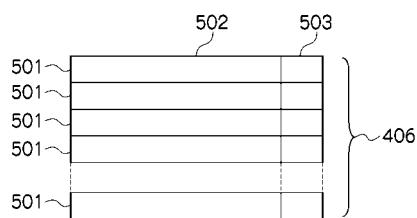
【図4】

図4



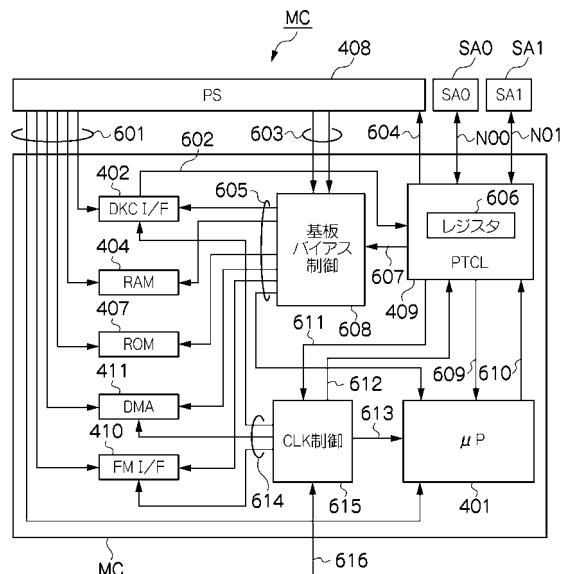
【図5】

図5



【図6】

図6



【図7】

図7



【図9】

図9

レジスタB	901	902	903	904	905	906
パラメータ	DKC I/F	RAM	ROM	DMA	FM I/F	μP
電源電圧						
基板バイアス						
動作周波数						

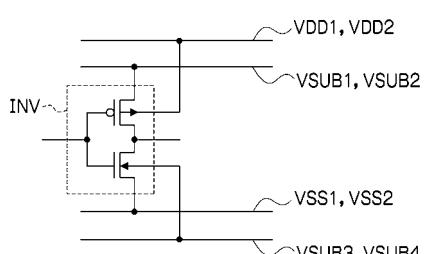
【図8】

図8

レジスタA	801	802	803	804	805
マスター/スレーブ	チェック有無	判定時間	ステータス	訂正不能エラー回数	

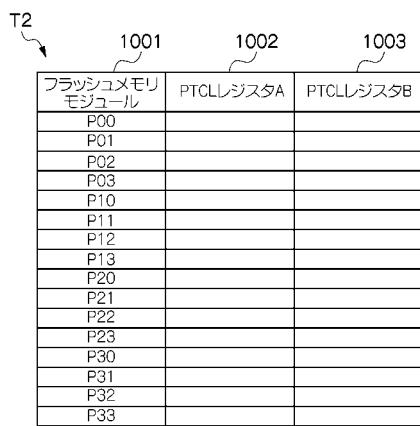
【図10】

図10



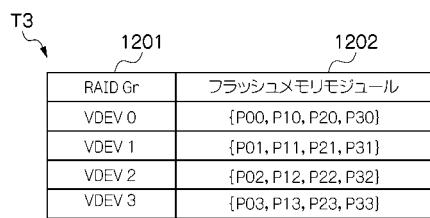
【図11】

図11



【図12】

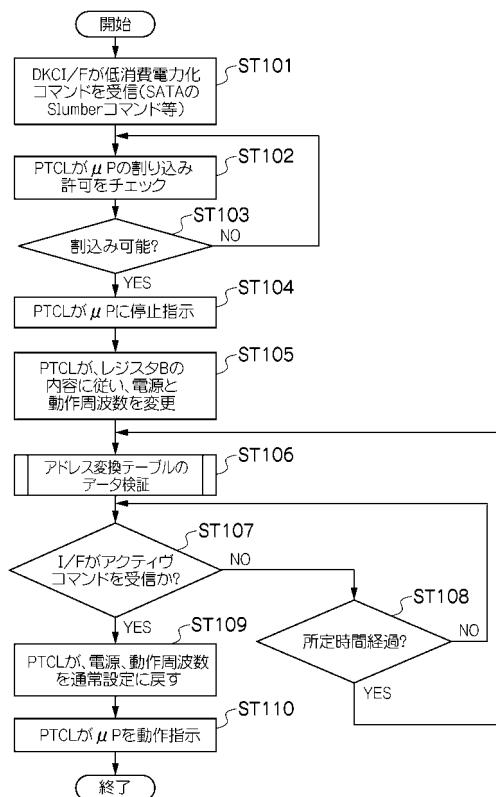
図12



【図13】

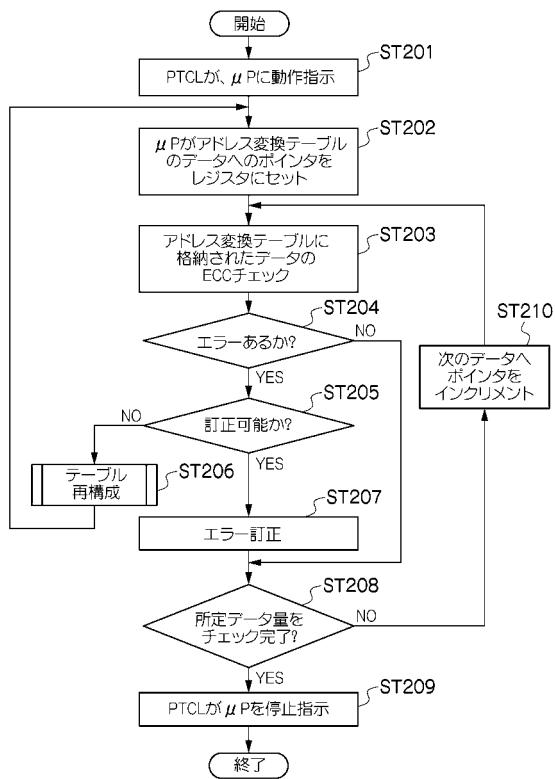
【図13】

図13



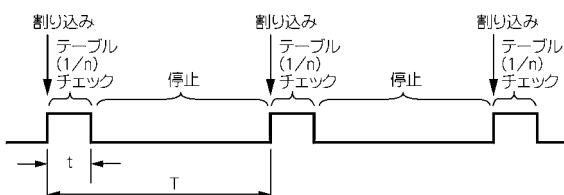
【図14】

図14



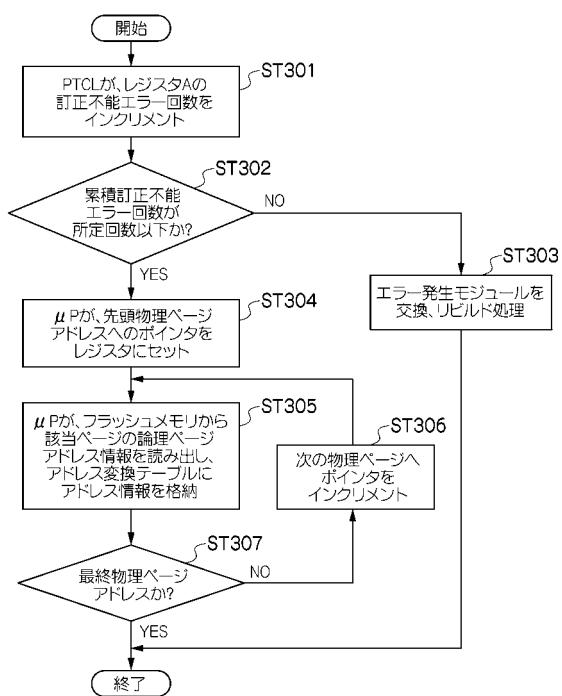
【図15】

図15



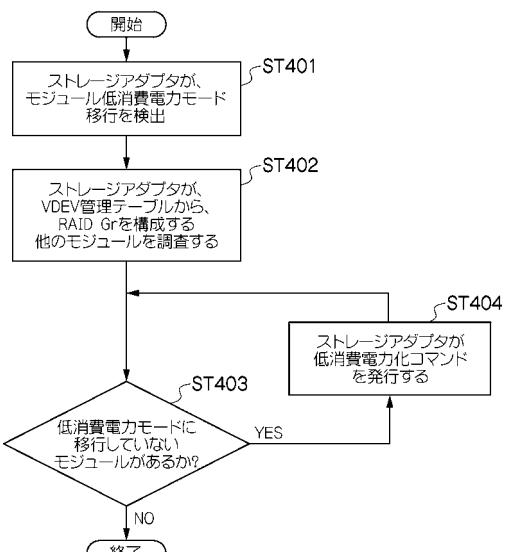
【図16】

図16



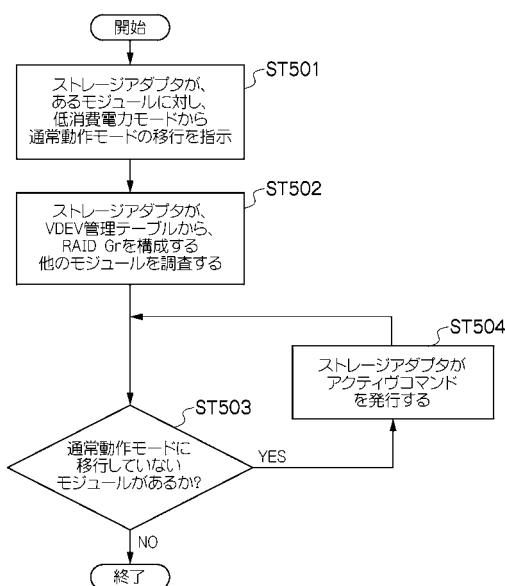
【図17】

図17



【図18】

図18



フロントページの続き

(56)参考文献 特開平07-056817(JP, A)
特開平11-053267(JP, A)
特開昭63-056751(JP, A)
特開2003-167795(JP, A)
特開平07-056694(JP, A)

(58)調査した分野(Int.Cl., DB名)

G 06 F 12 / 16
G 06 F 3 / 08
G 11 C 29 / 42