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

特許第6206161号
(P6206161)

(45)発行日　平成29年10月4日(2017.10.4)　(24)登録日　平成29年9月15日(2017.9.15)

(51)Int.Cl.		F I		
G06F 3/06	(2006.01)	G06F	3/06	301A
G06F 3/08	(2006.01)	G06F	3/08	H
G06F 12/16	(2006.01)	G06F	3/06	304Z
		G06F	3/06	301N
		G06F	12/16	310A

請求項の数 9　(全 38 頁)

(21)出願番号　特願2013-261843(P2013-261843)
(22)出願日　平成25年12月18日(2013.12.18)
(65)公開番号　特開2015-118572(P2015-118572A)
(43)公開日　平成27年6月25日(2015.6.25)
審査請求日　平成28年9月5日(2016.9.5)

(73)特許権者　000005223
富士通株式会社
神奈川県川崎市中原区上小田中４丁目１番１号
(74)代理人　100104190
弁理士　酒井　昭徳
(72)発明者　斎藤　博紀
神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

審査官　田中　啓介

最終頁に続く

(54)【発明の名称】ストレージ制御装置、制御方法、および制御プログラム

(57)【特許請求の範囲】
【請求項１】
ストレージ装置の記憶領域に対して所定のデータサイズのブロック単位でデータを冗長化して記憶させる制御を行うストレージ制御装置であって、
データの書込要求から特定される書込先の論理アドレスを、前記ブロック単位に対応付けてグループ分けしたグループ情報を記憶する記憶部と、
受け付けたデータの読出要求に応じて、前記グループ情報に基づき前記読出要求から特定される読出先の論理アドレスを含むグループの読出回数を計数し、計数した前記グループの読出回数に基づき前記読出先の論理アドレスを含む読出要求、または、前記読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行する制御部と、
を有し、
前記記憶領域は、規定回数の読み出しがブロックに対して行われると該ブロックのデータを他のブロックへコピーするコピー制御が行われる半導体メモリの記憶領域であり、
前記制御部は、
計数した前記グループの読出回数が前記規定回数より小さい第１の回数となったことに応じて、前記グループに含まれる論理アドレスのデータを対象とした前記コピー制御を起動させるように、前記グループに含まれる論理アドレスを含む読出要求を発行する処理を行い、
前記処理を行っているときに、データの読出要求を受け付けたことに応じて、当該読出

先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行する、
ことを特徴とするストレージ制御装置。
【請求項２】
前記ストレージ制御装置は、前記半導体メモリを含む前記ストレージ装置と、当該ストレージ装置が記憶するデータの冗長データを記憶する前記ストレージ装置とは異なる他のストレージ装置の制御を行うものであり、
前記制御部は、
計数した前記グループの読出回数が前記第１の回数を超えた後のデータの読出要求に応じて、当該読出要求から特定される読出先の論理アドレスが前記グループに含まれる場合に、当該読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を、前記他のストレージ装置に発行する、
ことを特徴とする請求項１に記載のストレージ制御装置。
【請求項３】
前記制御部は、
計数した前記グループの読出回数が前記第１の回数となったことに応じて、前記グループに含まれる論理アドレスのデータを対象とした前記コピー制御を起動させるように、前記グループに含まれる論理アドレスを含む読出要求を、前記規定回数から前記第１の回数を減じた第２の回数、または、当該読出要求に対する応答時間が所定時間を超えるまで、前記ストレージ装置に発行し、前記グループの読出回数を初期化することを特徴とする請求項１または２に記載のストレージ制御装置。
【請求項４】
前記制御部は、
前記グループに含まれる論理アドレスを含む読出要求を、前記第２の回数、または、当該読出要求に対する応答時間が所定時間を超えるまで発行したことに応じて、前記グループに含まれる論理アドレスを含む読出要求の発行を停止するとともに、前記グループの読出回数を初期化することを特徴とする請求項３に記載のストレージ制御装置。
【請求項５】
前記制御部は、
前記グループに含まれる論理アドレスを含む書込要求を前記ストレージ装置に発行したことに応じて、前記グループの読出回数を初期化することを特徴とする請求項１～４のいずれか一つに記載のストレージ制御装置。
【請求項６】
前記ストレージ制御装置は、前記半導体メモリを含む前記ストレージ装置と、当該ストレージ装置が記憶するデータの冗長データを記憶する前記ストレージ装置とは異なる他のストレージ装置の制御を行うものであり、
前記記憶部は、
書込要求から特定される前記他のストレージ装置の書込先の論理アドレスを、前記ブロック単位に対応付けてグループ分けした前記他のストレージ装置のグループ情報を記憶しており、
前記制御部は、
受け付けたデータの読出要求に応じて、前記他のストレージ装置のグループ情報に基づき当該読出要求から特定される前記他のストレージ装置の読出先の論理アドレスを含むグループの読出回数を計数し、計数した当該グループの読出回数が前記規定回数より小さく前記第１の回数とは異なる第３の回数となったことに応じて、当該グループに含まれる論理アドレスを含む読出要求を、前記規定回数から前記第３の回数を減じた第４の回数、または、当該読出要求に対する応答時間が所定時間を超えるまで前記他のストレージ装置に発行し、当該グループの読出回数を初期化することを特徴とする請求項３に記載のストレージ制御装置。
【請求項７】

前記ストレージ制御装置は、さらに、半導体メモリの記憶領域を所定のデータサイズで分割したブロック単位で書き込みを行う前記ストレージ装置とは異なる移行先のストレージ装置の制御を行うものであり、
前記制御部は、
前記ストレージ装置によりガベージコレクションが実行された場合に、受け付けた書込要求から特定される前記ストレージ装置の書込先の論理アドレスをグループ分けした複数のグループに含まれる論理アドレスを含むデータ移行要求を、前記ストレージ装置から前記移行先のストレージ装置への移行対象データの移行間隔が所定間隔を超えないように前記ストレージ装置に発行し、
発行した前記データ移行要求に含まれる論理アドレスを、前記ブロック単位に対応付けてグループ分けした前記移行先のストレージ装置のグループ情報を生成し、
前記データ移行要求に含まれる論理アドレスを含む読出要求の発行先を前記移行先のストレージ装置へ切り替える、
ことを特徴とする請求項１～６のいずれか一つに記載のストレージ制御装置。
【請求項８】
ストレージ装置の記憶領域に対して所定のデータサイズのブロック単位でデータを冗長化して記憶させる制御を行うストレージ制御装置の制御方法であって、
前記記憶領域は、規定回数の読み出しがブロックに対して行われると該ブロックのデータを他のブロックへコピーするコピー制御が行われる半導体メモリの記憶領域であり、
前記ストレージ制御装置が、
受け付けたデータの読出要求に応じて、データの書込要求から特定される書込先の論理アドレスを、前記ブロック単位に対応付けてグループ分けしたグループ情報に基づき前記読出要求から特定される読出先の論理アドレスを含むグループの読出回数を計数し、計数した前記グループの読出回数に基づき前記読出先の論理アドレスを含む読出要求、または、前記読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行し、
計数した前記グループの読出回数が前記規定回数より小さい第１の回数となったことに応じて、前記グループに含まれる論理アドレスのデータを対象とした前記コピー制御を起動させるように、前記グループに含まれる論理アドレスを含む読出要求を発行する処理を行い、
前記処理を行っているときに、データの読出要求を受け付けたことに応じて、当該読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行する、
処理を実行することを特徴とする制御方法。
【請求項９】
ストレージ装置の記憶領域に対して所定のデータサイズのブロック単位でデータを冗長化して記憶させる制御を行うストレージ制御装置の制御プログラムであって、
前記記憶領域は、規定回数の読み出しがブロックに対して行われると該ブロックのデータを他のブロックへコピーするコピー制御が行われる半導体メモリの記憶領域であり、
前記ストレージ制御装置に、
受け付けたデータの読出要求に応じて、データの書込要求から特定される書込先の論理アドレスを、前記ブロック単位に対応付けてグループ分けしたグループ情報に基づき前記読出要求から特定される読出先の論理アドレスを含むグループの読出回数を計数し、計数した前記グループの読出回数に基づき前記読出先の論理アドレスを含む読出要求、または、前記読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行し、
計数した前記グループの読出回数が前記規定回数より小さい第１の回数となったことに応じて、前記グループに含まれる論理アドレスのデータを対象とした前記コピー制御を起動させるように、前記グループに含まれる論理アドレスを含む読出要求を発行する処理を行い、

前記処理を行っているときに、データの読出要求を受け付けたことに応じて、当該読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行する、

処理を実行させることを特徴とする制御プログラム。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、ストレージ制御装置、制御方法、および制御プログラムに関する。

【背景技術】

【０００２】

従来、不揮発性のメモリとして、フラッシュメモリを採用したストレージ装置がある。フラッシュメモリでは、読み出しを行う際に浮遊ゲートに電圧をかけるため浮遊ゲートの電子量が変化する。このため、あるブロックに対して読み出しが繰り返されることにより、電子の変化量が大きくなってビットエラーが発生する、いわゆるリードディスターブが起きる。リードディスターブを防止するため、フラッシュメモリを含むストレージ装置内部では、読出回数を管理して、一定回数の読み出しが行われるとブロックのデータを別ブロックへコピーする処理を行う。

【０００３】

関連する先行技術として、たとえば、フラッシュメモリの読出頻度に対応してリードをフラッシュメモリから行うかＲＡＭ（Ｒａｎｄｏｍ Ａｃｃｅｓｓ Ｍｅｍｏｒｙ）から行うかを制御し、所定のエリアのデータのリード頻度が一の値を超えたとき、フラッシュメモリのデータをＲＡＭに転送するものがある。また、ホストＰＣ（Ｐｅｒｓｏｎａｌ Ｃｏｍｐｕｔｅｒ）からのアクセスコマンドに基づいて、リードディスターブによりデータが読み出される回数に制限があるフラッシュメモリからのデータ読み出し、および、フラッシュメモリをリフレッシュするための読み出し回数の更新を制御する技術がある。さらに、データ復旧処理において、フラッシュメモリから所定の時間内にデータ読出処理が完了しない場合に、データ読出処理を中止して付加データを設定し、パリティデータを用いて誤りデータを訂正した退避対象データをキャッシュメモリに書き込む技術がある。（たとえば、下記特許文献１～３を参照。）

【先行技術文献】

【特許文献】

【０００４】

【特許文献１】特開２００１－２９０７９１号公報

【特許文献２】特開２００８－１８１３８０号公報

【特許文献３】国際公開第２００９／１０７２１３号

【発明の概要】

【発明が解決しようとする課題】

【０００５】

しかしながら、従来技術によれば、フラッシュメモリを含むストレージ装置内部で行われるリードディスターブを防止する処理により、ストレージ装置への読出要求に対する応答性能が低下することがある。

【０００６】

１つの側面では、本発明は、ストレージ装置への読出要求に対する応答性能の低下を抑制するストレージ制御装置、制御方法、および制御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】

【０００７】

本発明の一側面によれば、データが繰り返し読み出されることによりビットエラーが発生する特性を有する半導体メモリの記憶領域を所定のデータサイズで分割したブロック単位で書き込みを行うストレージ装置と接続されるストレージ制御装置であって、ストレー

ジ装置への書込要求に応じて、書込要求に含まれる書込先の論理アドレスを、同一グループに含まれる論理アドレスに対応するデータサイズが所定のデータサイズを超えないようにグループ分けし、ストレージ装置への読出要求に応じて、読出要求に含まれる読出先の論理アドレスを含むグループの読出回数を計数し、計数したグループの読出回数に基づきグループに含まれる論理アドレスの読み出しを制御するストレージ制御装置、制御方法、および制御プログラムが提案される。

【発明の効果】

【０００８】

本発明の一態様によれば、ストレージ装置への読出要求に対する応答性能の低下を抑制することができるという効果を奏する。

【図面の簡単な説明】

【０００９】

【図１】図１は、本実施の形態にかかるストレージ制御装置の動作例を示す説明図である。

【図２】図２は、ストレージシステムの接続例を示す説明図である。

【図３】図３は、ストレージ制御装置のハードウェア構成例を示すブロック図である。

【図４】図４は、ストレージ制御装置の機能構成例を示すブロック図である。

【図５】図５は、ブロック予測テーブルの記憶内容の一例を示す説明図である。

【図６】図６は、書込時における動作例を示す説明図である。

【図７】図７は、読出時における動作例を示す説明図である。

【図８】図８は、上書書込時における動作例を示す説明図である。

【図９】図９は、ＲＤ防止コピー閾値の設定例を示す説明図（その１）である。

【図１０】図１０は、ＲＤ防止コピー閾値の設定例を示す説明図（その２）である。

【図１１】図１１は、ガベージコレクションが行われた際のブロックとグループとの乖離の一例を示す説明図である。

【図１２】図１２は、リフレッシュ処理の実行前後におけるブロックとグループとの関係の一例を示す説明図である。

【図１３】図１３は、ストレージ装置制御処理手順の一例を示すフローチャートである。

【図１４】図１４は、書込時処理手順の一例を示すフローチャート（その１）である。

【図１５】図１５は、書込時処理手順の一例を示すフローチャート（その２）である。

【図１６】図１６は、ＲＤ防止コピー引起し処理手順の一例を示すフローチャートである。

【図１７】図１７は、リフレッシュ処理手順の一例を示すフローチャートである。

【発明を実施するための形態】

【００１０】

以下に図面を参照して、開示のストレージ制御装置、制御方法、および制御プログラムの実施の形態を詳細に説明する。

【００１１】

図１は、本実施の形態にかかるストレージ制御装置の動作例を示す説明図である。図１に示すストレージシステム１００は、ストレージシステム１００の利用者に対して、ストレージシステム１００が有する記憶領域を提供するシステムである。たとえば、ストレージシステム１００は、ストレージシステム１００の利用者が作成した文書ファイルやデータベースを記憶する。

【００１２】

ストレージシステム１００は、ストレージシステム１００の全体を制御するストレージ制御装置１０１と、ストレージシステム１００が提供する記憶領域を有するストレージ装置１０２－１、２、…、ｎとを有する。ｎは１以上の整数である。

【００１３】

また、ストレージシステム１００が提供する記憶領域への読出要求に対する応答時間を短縮するために、ストレージ装置１０２は、ＳＳＤ（Ｓｏｌｉｄ　Ｓｔａｔｅ　Ｄｒｉｖ

ｅ）が適用される。以下、単に、ストレージ装置１０２と記載した場合、ＳＳＤが適用されたストレージ装置であるとする。また、本実施の形態では、ストレージ装置１０２－１～ｎのうち、少なくとも一つはＳＳＤが適用されたストレージ装置である。図１の例では、ストレージ装置１０２－１が、ＳＳＤが適用されたストレージ装置であるとする。また、ストレージシステム１００は、ＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙｓ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）技術を用いて形成された仮想的なボリュームを利用者に提供する。ストレージ制御装置１０１は、ＲＡＩＤ技術を用いて、ストレージ装置１０２の記憶領域に対して所定のデータサイズのブロック単位でデータを冗長化して記憶させる制御を行う。ＲＡＩＤについては、図２で後述する。

【００１４】

ＳＳＤは、記憶媒体としてフラッシュメモリを用いるドライブ装置である。具体的に、ＳＳＤは、フラッシュメモリと、受け付けた書込データやフラッシュメモリの記憶内容を一時的に記憶するキャッシュメモリと、ＳＳＤコントローラと、を有する。

【００１５】

フラッシュメモリとは、記憶領域を所定のデータサイズで分割したブロック単位で消去、書き込みを行うことができ、電源を切ってもデータを失わない性質を有する半導体メモリである。所定の管理データサイズ（以下、「ブロックデータサイズ」と呼称する）は、フラッシュメモリの製造時に、フラッシュメモリの仕様としてフラッシュメモリの設計者により設定される値である。

【００１６】

具体的に、フラッシュメモリは、１ビットの情報を蓄積するメモリセルを複数有する。メモリセルは、シリコン基板上のＰ型半導体層を挟み込むようにソースとドレインとなる２つのＮ型半導体部分を有し、Ｐ型半導体層の上に絶縁膜を介して浮遊ゲートを配置し、浮遊ゲート上に制御ゲートを配置した構造である。そして、メモリセルは、浮遊ゲートに蓄えられた電子量を用いて情報を記憶する。具体的に、メモリセルは、浮遊ゲートの電子量に応じてソース－ドレイン間に電流が流れるか否かが変化することを用いて、情報を記憶する。電子量に応じて電流が流れるか否かが変化する理由として、浮遊ゲートに電子があればソース－ドレイン間の抵抗が高くなりゲート電圧を高くしなければ電流が流れない。一方、浮遊ゲートに電子がなければソース－ドレイン間の抵抗が低くなり、ゲート電圧が低くても電流が流れる。ソース－ドレイン間の電流を流すために要求されるゲートへの印加電圧は、閾値電圧と呼称される。また、フラッシュメモリには、ＮＯＲ型のフラッシュメモリとＮＡＮＤ型のフラッシュメモリとがある。

【００１７】

ＳＳＤコントローラは、フラッシュメモリおよびキャッシュメモリを制御する。たとえば、ＳＳＤコントローラは、受け付けた論理アドレスに対する書込データを一旦キャッシュメモリに格納する。そして、ＳＳＤコントローラは、キャッシュメモリに格納した書込データをブロックデータサイズで分割して、分割したデータをブロックに順次格納するとともに、格納したブロックと論理アドレスとの対応関係を示す情報を、フラッシュメモリに格納する。書込データと、書込先のブロックとの関係については、ＳＳＤの仕様によって様々であり、たとえば、ＳＳＤへの書込データの発行間隔が、書込データと書込先のブロックとの関係に影響する場合もある。ここで、論理アドレスを指定する方法として、ＳＳＤにおいては、ＬＢＡ（Ｌｏｇｉｃａｌ Ｂｌｏｃｋ Ａｄｄｒｅｓｓｉｎｇ）が採用される。以下、論理アドレスを、単に、ＬＢＡと称する。ＳＳＤの内部動作については、図６～図８に後述する。

【００１８】

ここで、ストレージ制御装置１０１がストレージ装置１０２に発行する書込要求のＬＢＡは、利用者が操作する装置による仮想ボリュームの書込要求のＬＢＡから特定されるものである。同様に、ストレージ制御装置１０１がストレージ装置１０２に発行する読出要求のＬＢＡは、利用者が操作する装置による仮想ボリュームの読出要求のＬＢＡから特定されるものである。たとえば、ストレージ制御装置１０１が、ストレージ装置１０２－１

とストレージ装置１０２－２とにより、ＲＡＩＤ０の仮想ボリュームを形成するとする。このとき、ストレージ制御装置１０１は、ストレージ装置１０２－１のＬＢＡ０、ストレージ装置１０２－２のＬＢＡ０、ストレージ装置１０２－１のＬＢＡ１、ストレージ装置１０２－２のＬＢＡ１、…、という順に仮想ボリュームのＬＢＡ０～３、…を形成する。そして、ストレージ制御装置１０１は、利用者が操作する装置による仮想ボリュームの書込要求を受け付けた場合、書込要求のＬＢＡを２で除した際の商と余りとを用いて、書込要求のＬＢＡと発行先のストレージ装置１０２とを特定する。

【００１９】

読出要求からデータ受領完了までにかかる時間は、ＨＤＤ（Ｈａｒｄ　Ｄｉｓｋ　Ｄｒｉｖｅ）が１０ミリ秒程度であるのに対して、ＳＤＤが数ミリ秒程度である。したがって、ＳＳＤが適用されたストレージ装置は、ＨＤＤが適用されたストレージ装置より読出要求に対応する応答時間を短縮することができる。

【００２０】

しかしながら、ＳＳＤに読出要求を続けると、読出要求に対する応答が遅延する場合がある。読出要求に対する応答が遅延する場合として、ガベージコレクションの動作のタイミングやリードディスターブ（Ｒｅａｄ　Ｄｉｓｔｕｒｂ）によるビットエラーを防止させるため実施される動作のタイミングと、読出要求のタイミングとが重なった場合が挙げられる。

【００２１】

ここで、ガベージコレクションとは、小ブロックを纏まったページに集めてデータ記録域を確保する動作である。また、リードディスターブは、データが繰り返し読み出されることによりビットエラーが発生させる。リードディスターブは、電子を蓄える半導体メモリが有する特性である。リードディスターブは、ＮＡＮＤ型のフラッシュメモリで発生しやすい。リードディスターブによるビットエラーは、書き込んだ値が正しく読み出せなかったことによるエラーである。

【００２２】

リードディスターブの現象をより詳細に説明する。ＳＳＤ内部の記録媒体に用いられるフラッシュメモリの書込単位となるブロックに書き込みを行った時点では、浮遊ゲートに電子がない場合、閾値電圧は低く、読み出し時に印加する電圧とのマージンが十分にある状態である。しかしながら、ブロックのいずれかのセルに読出要求を繰り返し行うと、同一ブロック内の他のセルにも電圧が印加されて浮遊ゲートに保持した電子の量が少しずつ変化して、閾値電圧と読み出し時に印加する電圧とのマージンが減っていき、ビットエラーとなる。したがって、ＳＳＤは、各ブロックの読出回数を管理しておき、規定回数の読み出しが行われるとブロックのデータを別ブロックへコピーするように制御を行う。以降、上述の制御を「ＲＤ防止コピー」と呼称する。

【００２３】

ここで、ＲＤ防止コピーは、ＳＳＤ内部の動作であるから、ＳＳＤにアクセスする装置は、ＲＤ防止コピーがいつ行われるかという情報を得ることが難しい。このため、ＳＳＤがＲＤ防止コピーの実行中に、ＳＳＤにアクセスする装置がＲＤ防止コピーの対象であるブロックに読出要求を行ってしまう可能性がある。この場合、ＳＳＤは、ＲＤ防止コピーが完了した後にデータ読み出しと完了応答を行うため、ＳＳＤにアクセスする装置はＲＤ防止コピーが完了するまで待たされることになる。特に、ＲＤ防止コピーとガベージコレクションとが重なる場合には、ＳＳＤにアクセスする装置は大きく待たされることになり、たとえば、１秒以上完了応答が得られない場合がある。

【００２４】

そこで、ストレージ制御装置１０１は、ストレージ装置１０２に発行する書込要求のＬＢＡをブロック単位でグループ分けしたグループの読出回数を計数する。そして、ストレージ制御装置１０１は、計数したグループの読出回数に応じて、受け付けた読出要求のＬＢＡから特定されるＬＢＡを含む読出要求、または特定されるＬＢＡのデータに対応する冗長データの記憶先のＬＢＡを含む読出要求を発行する。これにより、ストレージ制御装

置１０１は、ＲＤ防止コピーが発生するであろうＬＢＡを予測して、応答遅延が発生するであろうＬＢＡに対する読出要求を避けることができ、読出要求に対する応答性能の低下を抑制することができる。

【００２５】

以下に、書込要求のＬＢＡをブロック単位でグループ分けしたグループの読出回数を計数し、計数したグループの読出回数に応じた読出要求の制御についての具体的な説明を行う。

【００２６】

上述したように、ＳＳＤは、一連の書込データをブロックデータサイズで分割して、分割したデータを順次ブロックに格納する。したがって、ストレージ制御装置１０１は、ストレージ制御装置１０１が受け付けた書込要求から特定される書込先のＬＢＡを、同一グループに含まれるＬＢＡに対応するデータサイズがブロックデータサイズを超えないように（ブロックデータサイズ以下で）グループ分けしたグループ情報を記憶する。ストレージ制御装置１０１は、１つのストレージ装置１０２に対するグループ情報を、ブロック予測テーブル１１１に登録する。ストレージ制御装置１０１は、ストレージ装置１０２に書込要求を発行する前にグループ分けしてグループ情報を生成してもよいし、ストレージ装置１０２に書込要求を発行した後にグループ分けしてグループ情報を生成してもよい。また、ストレージ制御装置１０１は、ストレージ装置１０２に対して既に発行された複数の書込要求に対して、グループ情報を生成してもよい。

【００２７】

ストレージ制御装置１０１は、各ストレージ装置１０２に対応するブロック予測テーブル１１１を管理する。図１の例では、ストレージ制御装置１０１は、ストレージ装置１０２－１に対応するブロック予測テーブル１１１－１を管理する。

【００２８】

図１の例では、ＬＢＡの１０００個分のデータサイズが、ブロックデータサイズと一致するものとし、ストレージ制御装置１０１が、ＬＢＡ０－１００７への書込要求をストレージ装置１０２－１に発行するとする。ストレージ制御装置１０１は、ＬＢＡ０－９９９をグループＧ０にグループ分けし、ＬＢＡ１０００－１００７をグループＧ１にグループ分けしたグループ情報を生成する。また、ストレージ装置１０２－１は、ストレージ制御装置１０１からのＬＢＡ０－１００７への書込要求を受け付けて、ＬＢＡ０－９９９をブロックＢ０に割り当て、ＬＢＡ１０００－１００７をブロックＢ１に割り当てたとする。以下の説明では、説明の簡略化のため、グループＧｘを、単に「Ｇｘ」と記載することがある。同様に、ブロックＢｘを、単に「Ｂｘ」と記載することがある。ｘは０以上の整数である。

【００２９】

次に、ストレージ制御装置１０１は、ストレージ制御装置１０１が受け付けた読出要求１１２に応じて、ブロック予測テーブル１１１－１に基づき読出要求１１２から特定される読出先のＬＢＡを含むグループの読出回数を計数する。計数する契機として、ストレージ制御装置１０１は、読出要求１１２から読出先のＬＢＡを特定した際に計数してもよいし、読出先のＬＢＡを含む読出要求をストレージ装置１０２－１に発行する前、または発行した後に計数してもよい。

【００３０】

図１は、ストレージ装置１０２－１のＬＢＡ０－９９９を読出先とする読出要求１１２が７９００回、ストレージ装置１０２－１のＬＢＡ１０００－１００７を読出先とする読出要求が１０２回の例を示す。このとき、ストレージ制御装置１０１は、Ｇ０の読出回数を７９００［回］と計数し、Ｇ１の読出回数を１０２［回］と計数する。また、ストレージ装置１０２－１は、ストレージ制御装置１０１から、読出先のＬＢＡ０－９９９を含む読出要求を７９００回受け付け、読出先のＬＢＡ１０００－１００７を含む読出要求を１０２回受け付ける。そして、ストレージ装置１０２－１は、Ｂ０の読出回数を７９００［回］とし、Ｂ１の読出回数を１０２［回］とする。

【００３１】
続けて、ストレージ制御装置１０１は、計数したグループの読出回数に基づき、読出要求１１２から特定される読出先のＬＢＡを含む読出要求１１３、または、読出先のＬＢＡのデータに対応する冗長データの記憶先のＬＢＡを含む読出要求１１４を発行する。図１中、点線で括られた読出要求１１３は、実際には発行していないことを示す。図１の例では、ストレージ制御装置１０１は、読出要求１１３を発行せず、読出要求１１４を発行する。
【００３２】
ここで、読出要求１１３、１１４について説明する。ここで、説明を容易にするため、読出要求１１３に含まれる読出先のＬＢＡのデータを、「元データ」と呼称する。
【００３３】
冗長データは、たとえば、読出要求１１３に含まれる読出先のＬＢＡが示す記憶領域に記憶されたデータと同一内容のデータである。また、冗長データは、冗長データを加工することにより、元データと同一内容のデータが得られるデータでもよい。
【００３４】
また、読出要求１１４の発行先について、ストレージ制御装置１０１は、ＲＡＩＤにより、読出要求１１２に含まれるＬＢＡ、または読出要求１１３に含まれるＬＢＡから、読出要求１１４の発行先が特定できるように管理してある。図１では、ＲＡＩＤ１の例を示す。図１に示すｄ１－ｓｒｃ、ｄ２、ｄ１－ｒｄｎは、図２におけるＲＡＩＤの説明時に併せて説明を行う。
【００３５】
次に、計数したグループの読出回数に基づき、読出要求１１３、または、読出要求１１４を発行する例について説明する。たとえば、ストレージ制御装置１０１は、ストレージ装置１０２－１の仕様に応じて設定された規定回数より小さい第１の回数を記憶しておく。第１の回数は、ストレージシステム１００の管理者によって設定される。そして、ストレージ制御装置１０１は、計数したグループの読出回数が第１の回数になるまでは、読出要求１１３を発行する。また、ストレージ制御装置１０１は、計数したグループの読出回数が第１の回数になった場合、読出要求１１４を発行する。以下、グループの読出回数と比較する値を、「ＲＤ防止コピー閾値」と呼称する。ＲＤ防止コピー閾値は、各ストレージ装置１０２に対してそれぞれで異なる値を設定してもよい。
【００３６】
読出要求１１３、または、読出要求１１４を発行した後、ストレージ制御装置１０１は、発行先からの応答に含まれるデータに基づき読出データを取得し、利用者に読出データを通知する。たとえば、読出要求１１３を発行した場合、ストレージ制御装置１０１は、発行先からの応答に含まれるデータと同一内容のデータを利用者が操作する装置に通知する。また、読出要求１１４を発行した場合、ストレージ制御装置１０１は、ＲＡＩＤ技術を用いて、発行先からの応答に含まれる冗長データを用いて元データの内容を復元して、復元したデータを利用者が操作する装置に通知する。図１の例では、ストレージ制御装置１０１は、データｄ１－ｒｄｎと同一内容のデータを、利用者が操作する装置に通知する。ＲＡＩＤ技術による読出データの復元については、図９および図１０で後述する。次に、図２を用いて、ストレージシステム１００の接続例を示す。
【００３７】
図２は、ストレージシステムの接続例を示す説明図である。ストレージシステム１００は、ホストサーバ２０１と、ストレージ制御装置１０１と、ストレージ装置１０２－１、２、…、ｎと、ホットスペア２０２とを有する。ホストサーバ２０１は、ストレージ制御装置１０１に接続する。ストレージ装置１０２－１、２、…、ｎと、ホットスペア２０２とは、ストレージ制御装置１０１に接続する。
【００３８】
ホストサーバ２０１は、ストレージシステム１００の利用者が操作する装置から、ストレージ装置１０２－１、２、…ｎへのアクセス要求を受け付ける装置である。ホットスペ

ア２０２は、ストレージ装置１０２－１、２、…、ｎに対して予備となる装置である。
【００３９】
ここで、ＲＡＩＤ技術について説明する。ストレージシステム１００は、ＲＡＩＤ技術により、１つの仮想的なボリュームを形成する。ここで、仮想的なボリュームを形成するストレージ装置群を、ＲＡＩＤグループに含まれるストレージ装置であると呼称する。
【００４０】
ＲＡＩＤには、仮想的なボリュームの形成の仕方を表すＲＡＩＤレベルが存在する。ＲＡＩＤレベルは、主にＲＡＩＤ０～ＲＡＩＤ６までのＲＡＩＤレベルが存在する。また、ＲＡＩＤには、ＲＡＩＤ０＋１というように、ＲＡＩＤレベルを組み合わせたレベルも存在する。たとえば、ＲＡＩＤ０は、冗長性を持たず、複数のストレージ装置にデータを分散するＲＡＩＤレベルである。また、ＲＡＩＤ１は、２台のストレージ装置に同一のデータを記憶するＲＡＩＤレベルである。
【００４１】
ここで、ＲＡＩＤ１について、図１を用いて説明する。図１は、ＲＡＩＤ１の例を示す。ストレージ装置１０２－１が元データと冗長データとを記憶するならば、ストレージ制御装置１０１は、ストレージ装置１０２－１に読出要求１１４を発行する。具体的には、ストレージ装置１０２－１のＬＢＡ０－９９９は、データｄ１－ｓｒｃを記憶し、ＬＢＡ１０００－１００７は、データｄ２を記憶しているとする。また、ストレージ装置１０２－２のＬＢＡ０－９９９は、データｄ１と同一内容のデータｄ１－ｒｄｎを記憶しているとする。この場合、ストレージ制御装置１０１は、ストレージ装置１０２－２に、読出先のＬＢＡがＬＢＡ０－９９９となる読出要求１１４を発行することになる。
【００４２】
また、ストレージ装置１０２－１～３によりＲＡＩＤ５を形成することもできる。同様に、読出要求１に含まれる読出先のＬＢＡがストレージ装置１０２－１であれば、ストレージ制御装置１０１は、冗長先のストレージ装置１０２－２、３に読出要求１１４を発行することになる。
【００４３】
ＲＡＩＤを形成することにより、ストレージシステム１００は、ＲＡＩＤレベルがＲＡＩＤ０以外であれば、利用者にデータの冗長性による信頼性の高い記憶領域を提供できる。また、ＲＡＩＤを形成することにより、アクセスが複数のストレージ装置１０２に分散することから、ストレージシステム１００は、利用者にアクセス応答が高速な記憶領域を提供できる。
【００４４】
また、複数台のストレージ装置１０２にデータを分散してＩ／Ｏを高速化させる技術はストライピングと呼ばれ、データがＸｂｙｔｅずつ順番にＲＡＩＤグループを形成する仮想的なボリュームに配置される。Ｘｂｙｔｅの各データを、「ストリップ（ｓｔｒｉｐ）」と呼称する。また、ＲＡＩＤグループを形成する各ストレージ装置１０２に配置されたストリップを横つながりに組み合わせた１つのデータセットを、「ストライプ（ｓｔｒｉｐｅ）」と呼称する。
【００４５】
図２の例では、ストレージ制御装置１０１は、仮想的なボリュームを形成して、ホストサーバ２０１に提供するＲＡＩＤコントローラの機能を有する。仮想的なボリュームの提供を受けたホストサーバ２０１は、仮想的なボリュームの論理アドレスを用いてストレージ制御装置１０１にアクセスを行う。ストレージ制御装置１０１は、仮想的なボリュームの論理アドレスをストレージ装置１０２のＬＢＡに変換して、変換したＬＢＡを用いてストレージ装置１０２にアクセスする。以下の記載では、説明の簡略化のため、ホストサーバ２０１からのアクセス先の論理アドレスは、ストレージ装置１０２のＬＢＡに変換済みであることを前提とする。
【００４６】
（ストレージ制御装置１０１のハードウェア）

図３は、ストレージ制御装置のハードウェア構成例を示すブロック図である。図３において、ストレージ制御装置１０１は、ＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）３０１と、ＲＯＭ（Ｒｅａｄ Ｏｎｌｙ Ｍｅｍｏｒｙ）３０２と、ＲＡＭ３０３と、を含む。また、ストレージ制御装置１０１は、ディスクドライブ３０４およびディスク３０５と、チャネルアダプタ３０６と、Ｉ／Ｏコントローラ３０７と、ＳＡＳ（Ｓｅｒｉａｌ Ａｔｔａｃｈｅｄ ＳＣＳＩ）エキスパンダ３０８と、を含む。また、ＣＰＵ３０１～ディスクドライブ３０４、チャネルアダプタ３０６、Ｉ／Ｏコントローラ３０７はバス３０９によってそれぞれ接続される。

【００４７】

ＣＰＵ３０１は、ストレージ制御装置１０１の全体の制御を司る演算処理装置である。また、ＣＰＵ３０１は、複数のプロセッサコアを有するマルチコアであってもよい。ＲＯＭ３０２は、ブートプログラムなどのプログラムを記憶する不揮発性メモリである。ＲＡＭ３０３は、ＣＰＵ３０１のワークエリアとして使用される揮発性メモリである。

【００４８】

ディスクドライブ３０４は、ＣＰＵ３０１の制御に従ってディスク３０５に対するデータのリードおよびライトを制御する制御装置である。ディスクドライブ３０４には、たとえば、磁気ディスクドライブ、ＳＳＤなどを採用することができる。ディスク３０５は、ディスクドライブ３０４の制御で書き込まれたデータを記憶する不揮発性メモリである。たとえばディスクドライブ３０４が磁気ディスクドライブである場合、ディスク３０５には、磁気ディスクを採用することができる。また、ディスクドライブ３０４がＳＳＤである場合、ディスク３０５には、半導体メモリを採用することができる。

【００４９】

チャネルアダプタ３０６は、ホストサーバ２０１に接続するアダプタである。Ｉ／Ｏコントローラ３０７は、ストレージ装置１０２－１、２、…、ｎと、ホットスペア２０２とのデータの入出力を制御する制御装置である。ＳＡＳエキスパンダ３０８は、複数のＳＡＳデバイスを接続可能にする装置である。ストレージ装置１０２－１、２、…、ｎと、ホットスペア２０２は、ＳＡＳ対応のデバイスである。

【００５０】

また、ストレージ制御装置１０１は、ストレージシステム１００の管理者等から直接操作される際を想定して、キーボード、マウスを有してもよい。

【００５１】

（ストレージ制御装置１０１の機能）

次に、ストレージ制御装置１０１の機能について説明する。図４は、ストレージ制御装置の機能構成例を示すブロック図である。ストレージ制御装置１０１は、制御部４０１と、記憶部４０２とを有する。制御部４０１は、グループ分け部４１１と、計数部４１２と、アクセス制御部４１３とを含む。制御部４０１は、記憶装置に記憶された、本実施の形態にかかるストレージ装置の制御プログラムをＣＰＵ３０１が実行することにより、制御部４０１の機能を実現する。記憶装置とは、具体的には、たとえば、図３に示したＲＯＭ３０２、ＲＡＭ３０３、ディスク３０５などである。グループ分け部４１１～アクセス制御部４１３の処理結果は、ＲＡＭ３０３、ディスク３０５などの記憶装置に格納される。

【００５２】

また、ストレージ制御装置１０１は、ブロック予測テーブル１１１を記憶する記憶部４０２にアクセス可能である。記憶部４０２は、ＲＡＭ３０３、ディスク３０５といった記憶装置に格納される。ブロック予測テーブル１１１の記憶内容の一例については、図５で後述する。

【００５３】

グループ分け部４１１は、ストレージ制御装置１０１が受け付けた書込要求に応じて、書込要求から特定される書込先のＬＢＡを、同一グループに含まれるＬＢＡに対応するデータサイズがブロックデータサイズを超えないようにグループ分けする。また、グループ分け部４１１は、さらに、同一グループに含まれるＬＢＡを含む書込要求のストレージ装

置１０２への発行間隔が所定間隔ｄｉを超えないようにグループ分けしてもよい。ここで、所定間隔ｄｉとは、ＳＳＤの仕様に関わる値であり、ある書込要求を受け付けた際に、フラッシュメモリに書き込む前に次の書込要求を待つ時間間隔である。ＳＳＤは、ある書込要求を受け付けてから所定間隔ｄｉを超えると次の書込要求を待たずにフラッシュメモリに書き込む。所定間隔ｄｉは、ストレージシステム１００の管理者により設定される値である。ストレージシステム１００の管理者は、ストレージ装置１０２の仕様を確認して、所定間隔ｄｉを設定する。具体的なグループ分けの一例と所定間隔ｄｉとについては、図６で後述する。

【００５４】

計数部４１２は、ストレージ制御装置１０１が受け付けた読出要求に応じて、読出要求から特定される読出先のＬＢＡを含むグループの読出回数を計数する。具体的なグループの読出回数の計数の例については、図７に後述する。

【００５５】

また、計数部４１２は、グループの読出回数を初期化してもよい。初期化する契機について、アクセス制御部４１３の処理結果に基づくものがあるため、アクセス制御部４１３の説明の後に説明する。

【００５６】

アクセス制御部４１３は、計数したグループの読出回数に基づき読出先のＬＢＡを含む読出要求、または読出先のＬＢＡのデータに対応する冗長データの記憶先のＬＢＡを含む読出要求を発行する。また、アクセス制御部４１３は、計数したグループの読出回数が第１の回数となったことに応じて、グループに含まれるＬＢＡを含む読出要求を、第２の回数、または読出要求に対する応答時間が所定時間を超えるまでストレージ装置１０２に発行してもよい。たとえば、アクセス制御部４１３は、計数したグループの読出回数が第１の回数となった後、グループに含まれるＬＢＡを含む読出要求をストレージ装置１０２に発行する。また、アクセス制御部４１３は、計数したグループの読出回数が第１の回数となり、読出要求の応答をストレージ装置１０２から受け付けた後、グループに含まれるＬＢＡを含む読出要求をストレージ装置１０２に発行してもよい。

【００５７】

第２の回数は、ストレージ装置１０２の仕様に応じて設定された規定回数から第１の回数を減じた値である。所定時間は、ストレージ制御装置１０１とストレージ装置１０２の仕様によって設定される値であり、たとえば、読出要求に対する通常の応答時間の平均値に所定値を加えた値である。たとえば、ストレージ制御装置１０１は、サービス提供前に、ストレージ装置１０２に対し読出要求を何回か発行し、発行してから応答があるまでの時間の平均値を通常の応答時間として、通常の応答時間に所定値を加えた値を、所定時間として設定しておく。所定値は、ストレージシステム１００の管理者が設定する値である。

【００５８】

ここで、アクセス制御部４１３は、グループに含まれるＬＢＡを含む読出要求を、第２の回数、または読出要求に対する応答時間が所定時間を超えるまでストレージ装置１０２に発行した場合、ストレージ装置１０２においてＲＤ防止コピーが行われたと判断する。読出要求を発行することにより、ストレージ装置１０２にＲＤ防止コピーを行わせることができるため、以下、グループの読出回数が第１の回数となった後に発行する読出要求を、「ＲＤ防止コピーを引き起こすコマンド」と呼称する。ＲＤ防止コピーを引き起こすコマンドは、読出要求であるＲｅａｄでもよいし、記憶先が正確であるか否かを確認するＶｅｒｉｆｙでもよい。

【００５９】

また、ＲＤ防止コピーを引き起こすコマンドを第２の回数回発行することにより、ＲＤ防止コピーが行われたと判断する方法を、以下、「第１のＲＤ防止コピー実行判断方法」と呼称する。また、応答時間が所定時間を超えるまでストレージ装置１０２にＲＤ防止コピーを引き起こすコマンドを発行することにより、ＲＤ防止コピーが行われたと判断する

方法を、以下、「第２のＲＤ防止コピー実行判断方法」と呼称する。

【００６０】

第１のＲＤ防止コピー実行判断方法の一例として、たとえば、ＲＤ防止コピー閾値が７７００［回］であり、ＲＤ防止コピーが発生する規定回数が８０００［回］であるとする。このとき、第２の回数は３００［回］となる。このとき、計数したグループの読出回数が７７００回となったことに応じて、ストレージ制御装置１０１は、ＲＤ防止コピーを引き起こすコマンドを３００［回］発行した際に、ＲＤ防止コピーが行われたと判断する。第１のＲＤ防止コピー実行判断方法を採用した場合、ストレージ制御装置１０１は、第２のＲＤ防止コピー実行判断方法と比較して、コマンドの発行回数を比較すればよいため、第２のＲＤ防止コピー実行判断方法より判断にかかる負荷を減らすことができる。

【００６１】

第２のＲＤ防止コピー実行判断方法を採用した場合、ストレージ制御装置１０１は、ストレージ装置１０２のＲＤ防止コピーが発生する回数を記憶していなくてよい。したがって、第２のＲＤ防止コピー実行判断方法を採用した場合、ストレージ制御装置１０１は、ストレージ装置１０２がＲＤ防止コピーを実行する規定回数という仕様が不明であっても実施することができる。

【００６２】

また、アクセス制御部４１３は、ＲＤ防止コピーを引き起こすコマンドを発行する間に、グループに含まれるＬＢＡを含む書込要求をストレージ装置１０２に発行したことに応じて、グループに含まれる論理アドレスを含む読出要求の発行を停止してもよい。

【００６３】

また、あるストレージ装置１０２において、計数したグループの読出回数が第１の回数となった後に、ストレージ制御装置１０１が読出要求を受け付けたとする。ここで、アクセス制御部４１３は、受け付けた読出要求に応じて、読出要求から特定されるＬＢＡが上述のグループに含まれる場合に、読出先のＬＢＡのデータに対応する冗長データの記憶先のＬＢＡを含む読出要求を、他のストレージ装置１０２に発行してもよい。

【００６４】

ここで、他のストレージ装置１０２は、あるストレージ装置１０２が記憶するデータの冗長データを記憶する装置である。たとえば、ストレージ装置１０２－１～３によりＲＡＩＤ５を形成するとする。そして、ストレージ装置１０２－１がデータ１を記憶しており、ストレージ装置１０２－２がデータ２を記憶しており、ストレージ装置１０２－３がデータ１およびデータ２から生成されるパリティデータを記憶するとする。このとき、ストレージ装置１０２が記憶するデータ１の冗長データを記憶する他のストレージ装置１０２は、パリティデータを記憶するストレージ装置１０２－３と、データ２を記憶するストレージ装置１０２－２とである。

【００６５】

また、計数部４１２は、アクセス制御部４１３により、ＲＤ防止コピーが行われたと判断した際に、グループの読出回数を初期化してもよい。また、計数部４１２は、グループに含まれるＬＢＡを含む書込要求をストレージ装置１０２に発行したことに応じて、グループの読出回数を初期化してもよい。また、計数部４１２は、アクセス制御部４１３がＲＤ防止コピーを引き起こすコマンドを発行する間に、グループに含まれるＬＢＡを含む書込要求をストレージ装置１０２に発行したことに応じて、グループの読出回数を初期化してもよい。

【００６６】

次に、あるストレージ装置１０２と他のストレージ装置１０２とのブロックデータサイズが同一の値であり、かつ、あるストレージ装置１０２と他のストレージ装置１０２との規定回数が同一の値であるときに用いる機能について説明する。

【００６７】

このとき、グループ分け部４１１は、他のストレージ装置１０２への書込要求に応じて、書込要求に含まれる書込先のＬＢＡを、同一グループに含まれるＬＢＡに対応するデー

タサイズがブロックデータサイズを超えないようにグループ分けする。また、計数部４１２は、他のストレージ装置１０２への読出要求に応じて、読出要求に含まれる読出先のＬＢＡを含むグループの読出回数を計数する。

【００６８】

また、アクセス制御部４１３は、計数したグループの読出回数が、第３の回数となったことに応じて、グループに含まれるＬＢＡを含む読出要求を、第４の回数、または、読出要求に対する応答時間が所定時間を超えるまで他のストレージ装置１０２に発行する。ここで、第３の回数は、ストレージ装置１０２の仕様に応じて設定された規定回数より小さく前記第１の回数とは異なる値である。また、第４の回数は、規定回数から第３の回数を減じた値である。ブロックデータサイズと規定回数とがそれぞれ同一の値となる例については、図９および図１０に後述する。

【００６９】

次に、ストレージ装置１０２のデータを、ＳＳＤを適用したストレージ装置１０２に移行するときに用いる機能について説明する。データを移行する理由としては、図１１、図１２で後述する。移行先のストレージ装置１０２は、たとえば、未使用であるホットスペア２０２である。以下、移行先のストレージ装置１０２が、ホットスペア２０２であるとして説明する。

【００７０】

このとき、アクセス制御部４１３は、受け付けた書込要求から特定されるストレージ装置１０２の書込先のＬＢＡをグループ分けした複数のグループに含まれるＬＢＡを含むデータ移行要求を、ストレージ装置１０２に発行する。このとき、アクセス制御部４１３は、ホットスペア２０２へのデータの移行間隔が所定間隔ｄｉを超えないように発行する。また、グループ分け部４１１は、発行したデータ移行要求に含まれるＬＢＡを、同一グループに含まれるＬＢＡに対応するデータサイズがホットスペア２０２のブロックデータサイズを超えないようにグループ分けする。

【００７１】

また、グループ分け部４１１は、受け付けた書込要求から特定されるホットスペア２０２の書込先のＬＢＡを、同一グループに含まれるＬＢＡに対応するデータサイズが当該ブロックデータサイズを超えないようにグループ分けする。また、計数部４１２は、受け付けた読出要求に応じて、受け付けた読出要求から特定されるホットスペア２０２の読出先のＬＢＡを含むグループの読出回数を計数する。また、アクセス制御部４１３は、計数したグループの読出回数に基づき読出先の論理アドレスを含む読出要求、または、読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行する。

【００７２】

図５は、ブロック予測テーブルの記憶内容の一例を示す説明図である。ブロック予測テーブル１１１は、ストレージ装置１０２のうち、ＳＳＤが適用されたストレージ装置１０２ごとに有するテーブルである。たとえば、図５に示すブロック予測テーブル１１１－１は、ストレージ装置１０２－１のブロックとＬＢＡ割り当ての関係とを登録するテーブルである。図５に示すブロック予測テーブル１１１－１は、レコード５０１－１～３を有する。

【００７３】

ブロック予測テーブル１１１は、グループ番号と、ＬＢＡと、読出回数という３つのフィールドを有する。グループ番号フィールドは、書込要求に含まれる書込先の論理アドレスを、同一グループに含まれる論理アドレスに対応するデータサイズがブロックデータサイズを超えないようにグループ分けした際の、グループの識別番号を記憶するフィールドである。ＬＢＡフィールドは、該当のグループにグループ分けされたＬＢＡを記憶するフィールドである。読出回数フィールドは、該当のグループの読出回数を記憶するフィールドである。

【００７４】

たとえば、レコード５０１－１は、ＬＢＡ０－９９９が一つのブロックに割り当てられたと予測したグループＧ０にグループ分けされており、グループＧ０の読出回数が７７００［回］であることを示す。

【００７５】

（ＳＳＤへの書込時と読出時との動作例）

次に、図６～図８を用いて、ストレージ装置１０２－１への書込時と読出時との動作例について説明する。ここで、本実施の形態において、ＳＳＤが適用されたストレージ装置１０２における１ブロックのブロックデータサイズは、説明の簡略化のため、全て５１２［Ｋバイト］であるとする。そして、１ブロックには、１０００個のＬＢＡが割当可能であるとする。したがって、１つのＬＢＡは、５１２［バイト］の記憶領域を有する。

【００７６】

図６は、書込時における動作例を示す説明図である。図６では、ＳＳＤが適用されたストレージ装置１０２－１への書込時において、ブロック予測テーブル１１１の更新例について説明する。以下の説明において、書込先のＬＢＡが割り当て済みでない書込要求を「新規書込要求」と称する。また、書込先のＬＢＡが割り当て済みの書込要求を「上書書込要求」と称する。

【００７７】

ここで、一般的なＳＳＤ内部の動作について説明する。ＳＳＤは、ブロック（５１２ＭＢ等）に複数のＬＢＡを割り当てる。割り当て動作は、通常固定ではなく、データが書き込まれた際に順次関連付けが行われる。ＳＳＤ内部にあるＳＳＤコントローラは、ＳＳＤに転送されたデータをＳＳＤ内部にあるキャッシュメモリに一時保存して、纏まった単位でフラッシュメモリへ書き込む。ＳＳＤコントローラは、このデータ纏め作業を優先で実施するが、転送データの発行間隔が所定間隔ｄｉ、たとえば１０秒を超えた場合、電源遮断によるデータ損失を防止するため、以降のデータを待つことなくフラッシュメモリへ書き込む。そして、ＳＳＤコントローラは、新規書込要求に関して、一旦書き込んだブロックに書き込まない。転送されたデータが新たに発生した場合、ＳＳＤコントローラは、次のブロックに書き込む。

【００７８】

上述したＳＳＤの動作から、ストレージ制御装置１０１でも同様の管理を行うことにより、ストレージ制御装置１０１は、ブロックとＬＢＡ割り当てとの対応関係について予測することができる。ストレージ制御装置１０１は、ストレージ装置１０２が同一のブロックに割り当てたであろうＬＢＡ群を、同一のグループに設定する。

【００７９】

具体的な予測例として、ホストサーバ２０１から、以下に示す第１の書込要求～第３の書込要求を受け付けたとして、ブロックとＬＢＡ割り当てとの関係を説明する。第１の書込要求～第３の書込要求は、いずれも新規書込要求であるとする。ストレージ制御装置１０１は、書込要求を受け付けた場合、ブロック予測テーブル１１１を参照して、受け付けた書込要求が新規書込要求か上書書込要求かを判断するが、図６では説明を省略し、図８で説明を行う。また、まだＬＢＡが一つも割り当てられていないストレージ装置１０２に対して、ストレージ制御装置１０１が書込要求に含まれる書込先のＬＢＡを含める割当中のグループが、Ｇ０であるとする。

【００８０】

第１の書込要求は、時刻ｔ０から時刻ｔ１までにかけて発生しており、書込先のＬＢＡがＬＢＡ０－１００７となる書込要求である。ＬＢＡ０－１００７のデータサイズは、５１６，０９６［バイト］に相当する。第２の書込要求は、時刻ｔ１の２秒後である時刻ｔ２から時刻ｔ３までにおけるＬＢＡ６３２８－６３４５に対する書込要求である。ＬＢＡ６３２８－６３４５のデータサイズは、９，２１６［バイト］に相当する。第３の書込要求は、時刻ｔ３の１３秒後である時刻ｔ４から時刻ｔ５にかけて、ＬＢＡ４２３０－４５２９に対する書込要求である。ＬＢＡ４２３０－４５２９のデータサイズは、１５３，６００［バイト］に相当する。第１の書込要求～第３の書込要求を受け付けた後、ストレー

ジ制御装置１０１は、それぞれをストレージ装置１０２に発行する。
【００８１】
時刻ｔ０において、ストレージ制御装置１０１は、第１の書込要求に含まれる書込先のＬＢＡのうちのＬＢＡ０－９９９を、割当中のグループであるＧ０に含めるようにグループ分けする。そして、Ｇ０にＬＢＡの空きがなくなったため、ストレージ制御装置１０１は、割当中のグループを、現在の割当中のグループの次のグループであるＧ１に設定する。次に、ストレージ制御装置１０１は、第１の書込要求に含まれる書込先のＬＢＡのうちのＬＢＡ１０００－１００７を、割当中のグループであるＧ１に含めるようにグループ分けする。
【００８２】
次に、時刻ｔ２において第２の書込要求を受け付けると、ストレージ制御装置１０１は、第１の書込要求の完了から第２の書込要求の発生までの間隔が所定間隔ｄｉを超えないため割当中のグループをＧ１のままとする。そして、ストレージ制御装置１０１は、第２の書込要求に含まれる書込先のＬＢＡ６３２８－６３４５を、割当中のグループであるＧ１に含めるようにグループ分けする。ここで、１０秒経過したか否かの判断方法について、ストレージ制御装置１０１は、時刻ｔ１時点でタイマをスタートしておき、時刻ｔ２時点でのタイマが示す時間が１０秒を超えたか否かで判断することができる。ストレージ制御装置１０１は、タイマを、ストレージ装置１０２ごとに用意しておく。
【００８３】
続けて、時刻ｔ４において、ストレージ制御装置１０１は、第３の書込要求を受け付けると、第２の書込要求の完了から第３の書込要求の発生までの間隔が所定間隔ｄｉを超えるため、割当中のグループを次のグループＧ２に設定する。そして、ストレージ制御装置１０１は、第３の書込要求に含まれる書込先のＬＢＡ４２３０－４５２９を、割当中のグループであるＧ２に含めるようにグループ分けする。
【００８４】
図７は、読出時における動作例を示す説明図である。図７では、ストレージ装置１０２への読出時において、ブロック予測テーブル１１１の更新例について説明する。
【００８５】
ここで、ＲＤ防止コピーの動作は、各グループに対し、何回の読出要求を受けたかで予測することができる。したがって、ストレージ制御装置１０１は、データの読出要求を行う際に、ブロック予測テーブル１１１のＬＢＡを参照し、該当するグループの読出回数を確認する。これにより、ストレージ制御装置１０１は、読出要求に含まれる読出先のＬＢＡが含まれるであろうブロックの読出回数を予測することができる。ここで、ＲＤ防止コピーが行われる読出回数は、ＳＳＤの仕様によって様々である。たとえば、ＳＳＤコントローラは、あるブロックに対して８０００回読み出しが行われた場合に、ＲＤ防止コピーを行う。そこで、ストレージ制御装置１０１は、ＳＳＤコントローラがＲＤ防止コピーを行う読出回数より小さい値をＲＤ防止コピー閾値として、読出回数がＲＤ防止コピー閾値に到達する場合に、ＲＤ防止コピーが行われる可能性があると判断する。
【００８６】
具体的な予測例として、図７では、図６に示した第３の書込要求が発生した後、ホストサーバ２０１から、以下に示す第１の読出要求～第３の読出要求があったとして、ＲＤ防止コピーが行われることを予測する動作を説明する。また、ＲＤ防止コピー閾値を７９００［回］とする。
【００８７】
第１の読出要求は、時刻ｔ６に発生した、読出先のＬＢＡがＬＢＡ０－１００７となる読出要求である。第２の読出要求は、時刻ｔ７に発生した、読出先のＬＢＡがＬＢＡ６３２８－６３４５となる読出要求である。第３の読出要求は、時刻ｔ８に発生した、読出先のＬＢＡがＬＢＡ６３２８－６３２９となる読出要求である。第１の読出要求～第３の読出要求を受け付けた後、ストレージ制御装置１０１は、それぞれをストレージ装置１０２に発行する。また、第２の読出要求と第３の読出要求との間に、ＬＢＡ１０００－１００

７、ＬＢＡ６３２８－６３４５のいずれかに対する読出要求が７８９７回あったものとする。

【００８８】

時刻ｔ６において、第１の読出要求に含まれる読出先のＬＢＡのうちのＬＢＡ０－９９９はＧ０に含まれるため、ストレージ制御装置１０１は、Ｇ０の読出回数をインクリメントする。また、第１の読出要求の読出先のＬＢＡのうちのＬＢＡ１０００－１００７はＧ１に含まれるため、ストレージ制御装置１０１は、Ｇ１の読出回数をインクリメントする。

【００８９】

時刻ｔ７において、第２の読出要求に含まれる読出先のＬＢＡであるＬＢＡ６３２８－６３４５はＧ１に含まれるため、ストレージ制御装置１０１は、Ｇ１の読出回数をインクリメントする。

【００９０】

時刻ｔ８となる前段階でＧ１に対する読出回数は、７８９９［回］になったものとする。そして、時刻ｔ８において、第３の読出要求に含まれる読出先のＬＢＡであるＬＢＡ６３２８－６３４５はＧ１に含まれるため、ストレージ制御装置１０１は、Ｇ１の読出回数をインクリメントする。インクリメントした結果、Ｇ１の読出回数が７９００［回］となり、ＲＤ防止コピー閾値に到達したため、ストレージ制御装置１０１は、ＲＤ防止コピーが行われる可能性があると判断する。

【００９１】

ＲＤ防止コピーが行われる可能性があると判断した後、ストレージ制御装置１０１は、ＲＤ防止コピーを引き起こすコマンドをストレージ装置１０２－１に発行する。図７の例では、ストレージ制御装置１０１は、時刻ｔ８から時刻ｔ９にかけて、ＲＤ防止コピーを引き起こすコマンドをストレージ装置１０２－１に発行する。そして、ストレージ制御装置１０１は、ＲＤ防止コピーが行われたと判断した場合に、Ｇ１の読出回数を０［回］に初期化する。ここで、ＲＤ防止コピーが行われたと判断する方法としては、たとえば、上述した第１のＲＤ防止コピー実行判断方法と、第２のＲＤ防止コピー実行判断方法とがある。

【００９２】

図８は、上書書込時における動作例を示す説明図である。図８では、ストレージ装置１０２への書込時において、既に書込要求を行ったＬＢＡに対して、さらに書込要求があった場合と、ＲＤ防止コピーを引き起こすコマンドの発行中に書込要求があった場合のブロック予測テーブル１１１の更新例について説明する。

【００９３】

ここで、一般的なＳＳＤ内部における上書書込時の動作について説明する。ＳＳＤは、上書書込を行う場合、ブロック単位でリードモディファイライト動作を行う。具体的に、ＳＳＤは、書込先のＬＢＡを含むブロックの記憶領域を読み出して、読み出したデータに上書書込のデータを反映して、反映したデータを書込先のＬＢＡを含むブロックに書き込む。ブロックに書き込む際に、ＳＳＤは、対象のブロックの記憶内容を消去した後、書き込みを行う。ブロックの記憶内容が消去されて書き込みが行われると、浮遊ゲートの電子量に応じた閾値電圧と、読み出し時に印加する電圧との間に十分なマージンが発生することになる。このように、上書書込を行うと、十分なマージンが発生するので、ＳＳＤは、書き込みを行ったブロックの読出回数を初期化する。

【００９４】

上述したＳＳＤの動作から、ストレージ制御装置１０１でも同様の予測を行うことにより、ブロックの予測精度を向上することができる。具体的な更新例として、図８では、図７に示した第３の読出要求が発生した後、ホストサーバ２０１から、第４の書込要求と、第４の読出要求と、第５の書込要求とを受け付けたとして、ブロック予測テーブル１１１の更新例について説明する。

【００９５】

第４の書込要求は、時刻ｔ１０から時刻ｔ１１までにかけて発生しており、書込先のＬＢＡがＬＢＡ０－８９９となる書込要求である。ＬＢＡ０－８９９のデータサイズは、４６０，８００［バイト］に相当する。第４の読出要求は、時刻ｔ１２に発生した、読出先のＬＢＡがＬＢＡ６２３８－６３４５となる読出要求である。第５の書込要求は、時刻ｔ１３から時刻ｔ１４までにかけて発生しており、書込先のＬＢＡがＬＢＡ６２３８－６３４５となる書込要求である。また、第４の書込要求と第４の読出要求との間に、ＬＢＡ１０００－１００７、ＬＢＡ６３２８－６３４５のいずれかに対する読出要求が７８９９回あったものとする。

【００９６】

ストレージ制御装置１０１は、第４の書込要求を受け付けた場合、第４の書込要求が新規書込要求か上書書込要求かを判断するため、ブロック予測テーブル１１１に、第４の書込要求に含まれる書込先のＬＢＡがあるか否かを判断する。図８の例では、Ｇ０にＬＢＡ０－８９９が含まれることから、ストレージ制御装置１０１は、第４の書込要求に含まれる書込先のＬＢＡがあり、第４の書込要求が上書書込要求であると判断する。ここで、図８では、書込要求に含まれる書込先のＬＢＡ全てがブロック予測テーブル１１１にあった例を示したが、書込要求に含まれる書込先のＬＢＡの一部分がブロック予測テーブル１１１にある場合も起こり得る。この場合、ストレージ制御装置１０１は、書込要求のうち、ブロック予測テーブル１１１にあるＬＢＡを上書書込要求として扱い、書込要求のうち、ブロック予測テーブル１１１にないＬＢＡを新規書込要求として扱う。

【００９７】

時刻ｔ１０において、第４の書込要求が上書書込要求であるため、ストレージ制御装置１０１は、第４の書込要求に含まれる書込先のＬＢＡを含むＧ０の読出回数を０［回］に初期化する。また、第４の書込要求を受け付けた後、ストレージ制御装置１０１は、受け付けた第４の読出要求をストレージ装置１０２に発行する。

【００９８】

時刻ｔ１２において、時刻ｔ１２となる前段階でＧ１に対する読出回数は、７８９９［回］になったものとする。そして、時刻ｔ１２において、第４の読出要求に含まれる読出先のＬＢＡであるＬＢＡ６３２８－６３４５はＧ１に含まれるため、ストレージ制御装置１０１は、Ｇ１の読出回数をインクリメントする。インクリメントした結果、Ｇ１の読出回数がＲＤ防止コピー閾値に到達したため、ストレージ制御装置１０１は、ＲＤ防止コピーが行われる可能性があると判断し、ＲＤ防止コピーを引き起こすコマンドをストレージ装置１０２－１に発行する。

【００９９】

時刻ｔ１３において、ＲＤ防止コピーを引き起こすコマンドを発行中に、第５の書込要求を受け付けると、ストレージ制御装置１０１は、第５の書込要求に含まれる書込先のＬＢＡが、ＲＤ防止コピーの対象のグループに含まれるかを判断する。図８の例では、第５の書込要求に含まれる書込先のＬＢＡが、ＲＤ防止コピーの対象であるＧ１に含まれるため、ストレージ制御装置１０１は、ＲＤ防止コピーを引き起こすコマンドの発行を停止する。同時に、ストレージ制御装置１０１は、Ｇ１の読出回数を０［回］に初期化する。

【０１００】

（ＲＤ防止コピー閾値の設定例）

次に、図９と図１０とを用いて、ＲＤ防止コピー閾値の設定例を示す。図９と図１０とで用いる例の前提として、ストレージ装置１０２－１～３が、ブロックデータサイズが同一のサイズであり、ＲＤ防止コピーが起こる回数が同一となるという、同一の特性を有するものとする。この場合、ストレージ制御装置１０１は、ストレージ装置１０２－１～３のＲＤ防止コピー閾値を、それぞれ、７７００［回］、７８００［回］、７９００［回］に設定しておく。また、ＳＳＤが適用されたストレージ装置１０２－１～３により、ＲＡＩＤレベルがＲＡＩＤ５であるＲＡＩＤグループを形成するものとする。そして、ストレージ装置１０２－１のＬＢＡ１０００－１００７と、ストレージ装置１０２－２のＬＢＡ１０００－１００７と、ストレージ装置１０２－３のＬＢＡ１０００－１００７とにより

、１つのストライプが形成されたとする。

【０１０１】

ストレージ装置１０２－１のＬＢＡ１０００－１００７はブロックＢｘに対応付けられており、ストレージ制御装置１０１は、ストレージ装置１０２－１のＬＢＡ１０００－１００７をグループＧｘに対応付けたものとする。同様に、ストレージ装置１０２－２のＬＢＡ１０００－１００７はブロックＢｙに対応付けられており、ストレージ制御装置１０１は、ストレージ装置１０２－２のＬＢＡ１０００－１００７をグループＧｙに対応付けたものとする。また、ストレージ装置１０２－３のＬＢＡ１０００－１００７はブロックＢｚに対応付けられており、ストレージ制御装置１０１は、ストレージ装置１０２－３のＬＢＡ１０００－１００７をグループＧｚに対応付けたものとする。

【０１０２】

さらに、該当のストライプにおいて、ストレージ装置１０２－１、２には実データｄ１、ｄ２が格納されており、ストレージ装置１０２－３には実データｄ１、ｄ２によるパリティデータｐ（ｄ１，ｄ２）が格納されたとする。以下、説明の簡略化のため、実データｄ１、ｄ２、パリティデータｐ（ｄ１，ｄ２）を、単に、ｄ１、ｄ２、ｐ（ｄ１，ｄ２）と記載する。

【０１０３】

図９は、ＲＤ防止コピー閾値の設定例を示す説明図（その１）である。図９に示すストレージシステム１００として、図９に示すブロック予測テーブル１１１－１は、ストレージ装置１０２－１のＧｘの読出回数が７７００［回］であることを記憶する。また、図９に示すブロック予測テーブル１１１－２は、ストレージ装置１０２－２のＧｙの読出回数が７７００［回］であることを記憶する。さらに、図９に示すブロック予測テーブル１１１－３は、ストレージ装置１０２－３のＧｚの読出回数が０［回］であることを記憶する。

【０１０４】

また、ストレージ制御装置１０１は、ストレージ装置１０２－１～３のＲＤ防止処理フラグを「処理していない」に設定したとする。ここで、ＲＤ防止処理フラグは、該当のストレージ装置１０２が、ＲＤ防止コピーを行っているであろうか否かを判断するフラグである。ＲＤ防止処理フラグは、ストレージ装置１０２ごとにあるデータである。ストレージ制御装置１０１は、ＲＤ防止処理フラグを、たとえば、ブロック予測テーブル１１１に対応付けて記憶する。ＲＤ防止処理フラグは、「処理していない」と、「処理中」という２つの識別子のいずれかとなる。「処理していない」とは、ストレージ制御装置１０１が、該当のストレージ装置１０２が、ＲＤ防止コピーを行っていないであろうと判断したことを示す識別子である。「処理中」とは、ストレージ制御装置１０１が、該当のストレージ装置１０２が、ＲＤ防止コピーを行っているであろうと判断したことを示す識別子である。

【０１０５】

ここで、ＲＤ防止コピー閾値が各ストレージ装置１０２間で同一の場合に起こりうる現象について説明する。同一のＲＡＩＤグループに含まれるストレージ装置１０２同士は、ほぼ均等に読出が行われる可能性が高く、読出回数の値が近い値になり易い。読出回数の値が近い値になった結果、あるタイミングにおいてＲＤ防止コピーを引き起こすコマンドを発行中の各ストレージ装置１０２が同時に複数台存在する状態になる可能性がある。上述の状態となり、ＲＤ防止コピーを引き起こすコマンドを発行中でないストレージ装置１０２の数が、データが復元可能な台数を下回ると、ストレージ制御装置１０１は、ＲＤ防止コピーが行われたと判断するまで読出要求に対するデータを取得できなくなる。

【０１０６】

そこで、ストレージ制御装置１０１は、ＲＤ防止コピー閾値をストレージ装置１０２間でずらして設定しておく。図９の例では、ストレージ装置１０２－１のＧｘの読出回数の値が、ストレージ装置１０２－１のＲＤ防止コピー閾値である７７００［回］に到達しており、ストレージ制御装置１０１は、ストレージ装置１０２－１のＧｘへの読出を抑止す

る。また、読出の抑止を行った後、ストレージ制御装置１０１は、ストレージ装置１０２－１のＲＤ防止処理フラグを「処理中」に設定するとともに、ストレージ装置１０２－１にＲＤ防止コピーを引き起こすコマンドを発行する。コマンドを発行してＲＤ防止コピーが行われたと判断した場合、ストレージ制御装置１０１は、ストレージ装置１０２－１のＧｘの読出回数を０［回］に初期化するとともに、ストレージ装置１０２－１のＲＤ防止処理フラグを「処理していない」に設定する。

【０１０７】

また、読出の抑止中において、図９に示すように、ホストサーバ２０１からストレージ装置１０２－１のＬＢＡ１０００－１００７の読出要求があったとする。このとき、ストレージ制御装置１０１は、ストレージ装置１０２－２からｄ２を読み出すとともに、ストレージ装置１０２－３からｐ（ｄ１，ｄ２）を読み出す。読み出したことにより、ストレージ制御装置１０１は、ストレージ装置１０２－２のＧｙの読出回数と、ストレージ装置１０２－３のＧｚの読出回数とをインクリメントする。続けて、ストレージ制御装置１０１は、ｄ２とｐ（ｄ１，ｄ２）とからｄ１を復元して、ホストサーバ２０１にｄ１を返却する。

【０１０８】

図１０は、ＲＤ防止コピー閾値の設定例を示す説明図（その２）である。図１０に示すストレージシステム１００の状態は、図９で示したストレージシステム１００の状態から、下記に示す第１の動作と第２の動作とが実行されたことにより、ストレージシステム１００の状態が変化したものである。

【０１０９】

第１の動作は、ストレージ装置１０２－１に対してＲＤ防止コピーを引き起こすコマンドを複数発行中にストレージ装置１０２－１のＬＢＡ１０００－１００７の読出要求が４９回発行されたという動作である。第１の動作により、ストレージ制御装置１０１は、グループＧｙの読出回数とグループＧｚの読出回数とをそれぞれ４９回インクリメントする。

【０１１０】

第２の動作は、ストレージ装置１０２－１に対してＲＤ防止コピーを引き起こすコマンドを複数発行中と発行終了後との合計で、ストレージ装置１０２－２のＬＢＡ１０００－１００７の読出要求が５０回発行されたという動作である。第２の動作により、ストレージ制御装置１０１は、グループＧｙの読出回数を５０回インクリメントする。

【０１１１】

第１の動作と第２の動作とにより、図１０に示すブロック予測テーブル１１１には、次に示す値が登録される。次に示す値として、具体的に、図１０に示すブロック予測テーブル１１１－１は、ストレージ装置１０２－１のＧｘの読出回数が０［回］であることを記憶する。また、図１０に示すブロック予測テーブル１１１－２には、ストレージ装置１０２－２のＧｙの読出回数が７８００［回］であることを登録される。さらに、図１０に示すブロック予測テーブル１１１－３には、ストレージ装置１０２－３のＧｚの読出回数が５０［回］であることが登録される。

【０１１２】

図１０の例では、ストレージ装置１０２－２のＧｙの読出回数の値がＲＤ防止コピー閾値である７８００［回］に到達したため、ストレージ制御装置１０１は、ストレージ装置１０２－２のＧｙへの読出を抑止する。また、読出の抑止を行った後、ストレージ制御装置１０１は、ストレージ装置１０２－２のＲＤ防止処理フラグを「処理中」に設定するとともに、ストレージ装置１０２－２にＲＤ防止コピーを引き起こすコマンドを複数発行する。コマンドを複数発行してＲＤ防止コピーが行われたと判断した場合、ストレージ制御装置１０１は、ストレージ装置１０２－２のＧｙの読出回数を０［回］に初期化するとともに、ストレージ装置１０２－２のＲＤ防止処理フラグを「処理していない」に設定する。

【０１１３】

また、読出の抑止中において、図１０に示すように、ホストサーバ２０１からストレージ装置１０２－２のＬＢＡ１０００－１００７の読出要求があったとする。このとき、ストレージ制御装置１０１は、ストレージ装置１０２－１からｄ１を読み出すとともに、ストレージ装置１０２－３からｐ（ｄ１，ｄ２）を読み出す。読み出したことにより、ストレージ制御装置１０１は、ストレージ装置１０２－１のＧｘの読出回数と、ストレージ装置１０２－３のＧｚの読出回数とをインクリメントする。続けて、ストレージ制御装置１０１は、ｄ１とｐ（ｄ１，ｄ２）とからｄ２を復元して、ホストサーバ２０１にｄ２を返却する。

【０１１４】

このように、ＲＤ防止コピー閾値をストレージ装置１０２間でずらして設定することにより、ＲＤ防止コピーが行われる期間をストレージ装置１０２間でずらすことができる。ＲＤ防止コピーが行われる期間をずらすことにより、ストレージシステム１００は、ミラーリングからのデータ取得ができない状態、またはパリティを用いたデータ復元ができない状態になることを避けることができる。

【０１１５】

（ガベージコレクションとリフレッシュ処理との説明）

次に、図１１を用いて、ガベージコレクションによりストレージ装置１０２が管理するブロックと、ストレージ制御装置１０１が予測するグループとが乖離した様子について説明する。また、図１２において、リフレッシュ処理により乖離が解消される様子について説明する。ここで、図１１と図１２とで共通して用いる前提として、ストレージ装置１０２－１が、ガベージコレクションを行うものとする。

【０１１６】

図１１は、ガベージコレクションが行われた際のブロックとグループとの乖離の一例を示す説明図である。表１１０１は、ガベージコレクションの実行前の状態における、ストレージ装置１０２が管理するブロックとＬＢＡ割り当てとの関係を示すテーブルである。また、表１１０２は、ガベージコレクションの実行後の状態における、ストレージ装置１０２が管理するブロックとＬＢＡ割り当てとの関係を示すテーブルである。

【０１１７】

ガベージコレクションの実行前の状態として、表１１０１とブロック予測テーブル１１１－１より、ストレージ装置１０２－１のＬＢＡ０－９９９は、ブロックＢ０に対応付けられており、ストレージ制御装置１０１は、ＬＢＡ０－９９９をＧ０に含めたものとする。Ｂ０の読出回数とＧ０の読出回数とは、ともに、２０［回］であるとする。

【０１１８】

また、ストレージ装置１０２－１のＬＢＡ１０００－１００７、６３２８－６３４５は、ブロックＢ１に対応付けられており、ストレージ制御装置１０１は、ＬＢＡ１０００－１００７、６３２８－６３４５をグループＧ１に含めたものとする。Ｂ１の読出回数とＧ１の読出回数とは、ともに、２８［回］であるとする。

【０１１９】

また、ストレージ装置１０２－１のＬＢＡ１００８－１０９９、４２３０－４５２９は、ブロックＢ２に対応付けられており、ストレージ制御装置１０１は、ＬＢＡ１００８－１０９９、４２３０－４５２９をグループＧ２に含めたものとする。Ｂ２の読出回数とＧ２の読出回数とは、ともに、２３６７［回］であるとする。

【０１２０】

また、ストレージ装置１０２－１のＬＢＡ１１００－２０９９は、ブロックＢ１０００に対応付けられており、ストレージ制御装置１０１は、ＬＢＡ１１００－２０９９をグループＧ１０００に含めたものとする。Ｂ１０００の読出回数とＧ１０００の読出回数とは、ともに、６９２［回］であるとする。

【０１２１】

この状態で、ストレージ装置１０２－１が、ガベージコレクションを実行したとする。具体的には、ストレージ装置１０２－１のＳＳＤコントローラは、Ｂ１のＬＢＡ数が２６

［個］であり、Ｂ２のＬＢＡ数が３９２［個］であるから、Ｂ１のＬＢＡとＢ２のＬＢＡとを纏めても、一つのブロックに収まるＬＢＡ数１０００を超えないと判断する。そして、ストレージ装置１０２－１のＳＳＤコントローラは、Ｂ２が示す記憶領域に記憶されたＬＢＡ１００８－１０９９、４２３０－４５２９のデータを、Ｂ１の記憶領域に移行する。

【０１２２】

ガベージコレクションの実行後において、ストレージ装置１０２－１のＬＢＡ１００８－１０９９、４２３０－４５２９は、ＬＢＡ１０００－１００７、６３２８－６３４５と併せてＢ１に対応付けられる。これに対し、ストレージ制御装置１０１は、ＬＢＡ１０００－１００７、６３２８－６３４５をＧ１に、ＬＢＡ１００８－１０９９、４２３０－４５２９をＧ２にグループ分けしており、ＬＢＡの割当内容が乖離したことになる。

【０１２３】

図１２は、リフレッシュ処理の実行前後におけるブロックとグループとの関係の一例を示す説明図である。リフレッシュ処理は、ＬＢＡの割当内容が乖離したストレージ装置を移行元のストレージ装置１０２とし、移行元のストレージ装置１０２が記憶する移行対象データを移行先のストレージ装置１０２に移行させて、グループとＬＢＡとの関係を再定義する処理である。移行先のストレージ装置１０２は、ＬＢＡ割り当てがない、未使用のストレージ装置１０２であることが好ましい。ＬＢＡ割り当てがある、使用中のストレージ装置１０２に移行すると、既にＬＢＡの割当内容が乖離している可能性があるためである。図１２の例では、移行先のストレージ装置がホットスペア２０２であるとする。また、ホットスペア２０２は、ＳＳＤを適用したストレージ装置１０２であるとする。

【０１２４】

ここで、データの移行方法は、たとえば、下記に示す第１の移行方法と、第２の移行方法とがある。第１の移行方法は、ストレージ制御装置１０１が移行元となるストレージ装置１０２にデータ移行要求を発行し、移行元のストレージ装置１０２が、移行先のストレージ装置１０２に直接移行対象データを送信する方法である。また、第２の移行方法は、データ移行要求を受け付けた移行元のストレージ装置１０２が移行対象データをストレージ制御装置１０１に送信し、ストレージ制御装置１０１は、移行対象データを移行先のストレージ装置１０２に送信する方法である。本実施の形態では、第１の移行方法を採用したとする。

【０１２５】

どちらの場合であっても、ストレージ制御装置１０１は、移行先のストレージ装置１０２への移行対象データのうちの一つのグループに含めるようにグループ分けしたデータ内の移行間隔が所定間隔ｄｉを超えないようにする。たとえば、一つのグループに含めるようにグループ分けしたＬＢＡ群のデータ５１２［Ｋバイト］が、２５６［Ｋバイト］、２５６［Ｋバイト］というように分割して移行先のストレージ装置１０２に送信されたとする。さらに、１回目の２５６［Ｋバイト］のデータが移行してから、２回目の２５６［Ｋバイト］のデータが移行するまでの時間が、所定間隔ｄｉを超えたとする。このとき、ストレージ制御装置１０１では一つのグループに含めるようにグループ分けされたＬＢＡ群が、移行先のストレージ装置１０２の２つのブロックに割り当てられてしまい、ＬＢＡ割り当ての内容に乖離が発生してしまう。

【０１２６】

移行間隔が所定間隔ｄｉを超えないようにするために、第１の移行方法の場合には、ストレージシステム１００を構築する際に、ストレージシステム１００の管理者が、ストレージ装置１０２のデータ転送の間隔が所定間隔ｄｉ以内であるかを確認しておけばよい。また、第２の移行方法の場合には、ストレージ制御装置１０１が、たとえば、移行元のストレージ装置１０２から受け付けた移行対象データをバッファリングして、移行先のストレージ装置１０２にブロックデータサイズごとに転送すればよい。

【０１２７】

データ移行要求は、移行対象データが記憶された記憶領域を示すＬＢＡと、移行対象デ

ータの移行先のストレージ装置１０２を特定する情報と、を含む。
【０１２８】
次に、図１２の例を用いて、リフレッシュ処理の動作例を示す。ストレージ制御装置１０１は、ＬＢＡ０－２０９９、４２３０－４５２９、６３２８－６３４５をホットスペア２０２に移行するデータ移行要求を、ストレージ装置１０２－１に発行する。
【０１２９】
そして、ストレージ制御装置１０１は、まだグループ分けしていないＬＢＡのうちの先頭のＬＢＡを選択する。図１２の例では、ストレージ制御装置１０１は、ＬＢＡ０を選択する。次に、ストレージ制御装置１０１は、ホットスペア２０２のブロックデータサイズである、５１２［Ｋバイト］を超えず、選択したＬＢＡを先頭とした、ＬＢＡ０－９９９を、Ｇ０にグループ分けする。
【０１３０】
次に、ストレージ制御装置１０１は、まだグループ分けしていないＬＢＡのうちの先頭となるＬＢＡ１０００を選択する。そして、ストレージ制御装置１０１は、５１２［Ｋバイト］を超えず、選択したＬＢＡを先頭とした、ＬＢＡ１０００－１９９９を、Ｇ１にグループ分けする。
【０１３１】
続けて、ストレージ制御装置１０１は、まだグループ分けしていないＬＢＡのうちの先頭となるＬＢＡ２０００を選択する。そして、ストレージ制御装置１０１は、５１２［Ｋバイト］を超えず、選択したＬＢＡを先頭とした、ＬＢＡ２０００－２０９９、４２３０－４５２９、６３２８－６３４５をＧ２にグループ分けする。全てのＬＢＡをグループ分けしたため、ストレージ制御装置１０１は、リフレッシュ処理を終了する。
【０１３２】
リフレッシュ処理後、ストレージ制御装置１０１は、ホットスペア２０２を、ストレージ装置１０２－１が属するＲＡＩＤグループに追加するとともに、ストレージ装置１０２－１をＲＡＩＤグループから外して新たなホットスペアに設定する。また、ストレージ制御装置１０１は、移行後のホットスペア２０２の記憶内容を、ストレージ装置１０２－１に戻し、ＲＡＩＤグループを変更せずにそのままにしてもよい。
【０１３３】
これにより、ストレージ制御装置１０１が予測するグループとＬＢＡ割り当ての関係と、ホットスペア２０２が管理するブロックとＬＢＡ割り当てとの内容が一致して、ＲＤ防止コピーが発生するＬＢＡの予測精度が向上する。図１２の例では、ホットスペア２０２のグループとＬＢＡ割り当ての予測を示すブロック予測テーブル１１１－ｈｓの内容と、ホットスペア２０２が管理するブロックとＬＢＡ割り当てとの内容を示す表１２０１の内容が一致する。
【０１３４】
リフレッシュ処理を実行する契機としては、ストレージ制御装置１０１は、たとえば、定期的にリフレッシュ処理を実行する。また、ストレージ制御装置１０１は、ＬＢＡの割当内容が乖離したことを判断した場合にリフレッシュ処理を実行してもよい。ＬＢＡの割当内容が乖離したことを判断する方法として、ストレージ制御装置１０１は、たとえば、上述した第１のＲＤ防止コピー実行判断方法および第２のＲＤ防止コピー実行判断方法を組み合わせることにより、ＬＢＡの割当内容が乖離したことを判断する。
【０１３５】
具体的には、第１のＲＤ防止コピー実行判断方法によりストレージ装置１０２のあるグループの読出回数がＲＤ防止コピー閾値に到達し、ストレージ制御装置１０１は、ＲＤ防止コピーを引き起こすコマンドを第２の回数分、ストレージ装置１０２に発行したとする。コマンド発行後、第２のＲＤ防止コピー実行判断方法により、ストレージ制御装置１０１は、ストレージ装置１０２からの応答が遅延しなかった場合に、ストレージ装置１０２においてＲＤ防止コピーが行われておらず、ＬＢＡの割当内容が乖離したと判断する。ＬＢＡの割当内容が乖離したと判断した場合、ストレージ制御装置１０１は、リフレッシュ

処理を実行する。また、ストレージ制御装置１０１は、ＬＢＡの割当内容が乖離したと判断した回数が所定値を超えた場合に、リフレッシュ処理を実行してもよい。

【０１３６】

具体的に、図１２のリフレッシュ処理を行う前の状態を用いて、ＬＢＡの割当内容が乖離したことを判断する例について説明する。ここで、ＲＤ防止コピー閾値を７９００［回］とし、ストレージ装置１０２がＲＤ防止コピーを行う読出回数を８０００［回］とする。そして、図１２に示す表１１０２、ブロック予測テーブル１１１－１の状態となった以降、リフレッシュ処理を行わず、かつ、ホストサーバ２０１から、ストレージ装置１０２－１のＬＢＡ１００８－１０９９の読出要求が、５５３３回あったとする。

【０１３７】

上述した前提となる場合に、ストレージ制御装置１０１は、ブロック予測テーブル１１１－１のＬＢＡ１００８－１０９９が含まれるＧ２の読出回数を５５３３回インクリメントする。インクリメントした結果、Ｇ２の読出回数は、２３６７＋５５３３＝７９００となり、ＲＤ防止コピー閾値に到達したため、ストレージ制御装置１０１は、ストレージ装置１０２－１に対して、ＲＤ防止コピーを引き起こすコマンドを１００［回］発行する。

【０１３８】

しかしながら、ストレージ装置１０２－１は、ＬＢＡ１００８－１０９９の読出要求が５５３３回あった時点で、Ｂ１の読出回数が５５３３［回］であると管理している。したがって、ストレージ装置１０２は、ＲＤ防止コピーを引き起こすコマンドを１００［回］受け付けても、Ｂ１の読出回数が８０００［回］に到達しないため、Ｂ１に対するＲＤ防止コピーを行わない。このため、ストレージ装置１０２－１は、ストレージ制御装置１０１に対して、ＲＤ防止コピーを引き起こすコマンドに対する応答を遅延せずに行うことになる。このように、ストレージ制御装置１０１は、ＲＤ防止コピーを引き起こすコマンドの発行回数と、ＲＤ防止コピーを引き起こすコマンドの発行に対する応答時間を確認することにより、ＬＢＡの割当内容が乖離したことを判断することができる。

【０１３９】

ＬＢＡの割当内容が乖離したと判断したときにリフレッシュ処理を実行することにより、ストレージ制御装置１０１は、定期的にリフレッシュ処理を実行する場合に比べて効率的にリフレッシュ処理を実行することができる。具体的に、ストレージ制御装置１０１は、応答遅延が発生すればリフレッシュ処理を行い応答遅延を抑制することができ、応答遅延が発生しなければ移行先のストレージ装置１０２の書込回数を減らすことができる。ここで、ＳＳＤ内の浮遊ゲートに書き込める回数には上限があり、ある一定回数以上書き込みを行うと、浮遊ゲート内の絶縁膜が劣化して、メモリセルは情報を記憶することができなくなる。したがって、書込回数を減らすことにより、ストレージ制御装置１０１は、移行先のストレージ装置１０２が故障するまでの期間を延ばすことができる。

【０１４０】

次に、図１３～図１７を用いて、ストレージ制御装置１０１が実行するフローチャートについて説明する。

【０１４１】

図１３は、ストレージ装置制御処理手順の一例を示すフローチャートである。ストレージ装置制御処理は、ホストサーバ２０１のアクセスコマンドに応じてストレージ装置１０２を制御する処理である。

【０１４２】

ストレージ制御装置１０１は、ホストサーバ２０１から、アクセスコマンドを受け付ける（ステップＳ１３０１）。次に、ストレージ制御装置１０１は、アクセスコマンドの種別が次に示す要求のいずれに一致するかを判断する（ステップＳ１３０２）。次に示す要求は、書込要求と、読出要求と、である。アクセスコマンドの種別が書込要求である場合（ステップＳ１３０２：書込要求）、ストレージ制御装置１０１は、書込時処理を実行する（ステップＳ１３０３）。書込時処理の詳細については、図１４および図１５で後述する。

【０１４３】
アクセスコマンドの種別が読出要求である場合（ステップＳ１３０２：読出要求）、ストレージ制御装置１０１は、続けて、ＲＤ防止処理フラグが次に示す識別子のいずれに一致するかを判断する（ステップＳ１３０４）。次に示す識別子は、「処理していない」と、「処理中」と、である。ＲＤ防止処理フラグが「処理していない」である場合（ステップＳ１３０４：“処理していない”）、ストレージ制御装置１０１は、ブロック予測テーブル１１１の読出先のＬＢＡが含まれるグループの読出回数がＲＤ防止コピー閾値以上か否かを判断する（ステップＳ１３０５）。
【０１４４】
ブロック予測テーブル１１１の読出先のＬＢＡが含まれるグループの読出回数がＲＤ防止コピー閾値未満である場合（ステップＳ１３０５：Ｎｏ）、ストレージ制御装置１０１は、読出先のストレージ装置１０２に読出要求を発行する（ステップＳ１３０６）。発行した結果、ストレージ制御装置１０１は、読出要求に対するデータを得る。次に、ストレージ制御装置１０１は、ブロック予測テーブルのアクセス対象ＬＢＡを含むグループの読出回数をインクリメントする（ステップＳ１３０７）。
【０１４５】
ブロック予測テーブルの読出先のＬＢＡが含まれるグループの読出回数がＲＤ防止コピー閾値以上である場合（ステップＳ１３０５：Ｙｅｓ）、ストレージ制御装置１０１は、読出先のストレージ装置への読出要求の発行を抑止する（ステップＳ１３０８）。次に、ストレージ制御装置１０１は、ＲＤ防止コピー引起し処理を実行する（ステップＳ１３０９）。ＲＤ防止コピー引起し処理の詳細については、図１６で後述する。
【０１４６】
ここで、ステップＳ１３０９と、後続するステップＳ１３１０との実行の関係において、ストレージ制御装置１０１は、ステップＳ１３０９の処理開始後に、ステップＳ１３１０の処理を実行する。たとえば、ストレージ制御装置１０１は、ＲＤ防止コピー引起し処理を実行するスレッドを起動した後、ストレージ装置制御処理を実行するスレッドと、ＲＤ防止コピー引起し処理を実行するスレッドとをマルチスレッドにより並列に処理してもよい。また、ＣＰＵ３０１がマルチコアである場合、マルチコアのうちのあるコアが、ストレージ装置制御処理を実行するスレッドを実行し、他のコアがＲＤ防止コピー引起し処理を実行するスレッドを実行してもよい。
【０１４７】
ステップＳ１３０９の処理開始後、ストレージ制御装置１０１は、ＲＡＩＤグループ内の読出先のストレージ装置１０２以外のストレージ装置１０２に、読出要求に応じた読出データを復元可能なデータを記憶する記憶領域を示すＬＢＡの読出要求を発行する（ステップＳ１３１０）。ステップＳ１３１０における読出要求は、ホストサーバ２０１からの読出要求である。また、ＲＤ防止処理フラグが「処理中」である場合（ステップＳ１３０４：“処理中”）、ストレージ制御装置１０１は、ステップＳ１３１０の処理を実行する。
【０１４８】
ステップＳ１３１０の処理について、たとえば、読出先のストレージ装置１０２がＲＡＩＤ１のＲＡＩＤグループに属する場合、ストレージ制御装置１０１は、ＲＡＩＤグループのミラーリング先のストレージ装置１０２に読出要求を発行する。そして、ストレージ制御装置１０１は、ミラーリング先のストレージ装置１０２から、ホストサーバ２０１が発行した読出要求に応じたデータと同一内容の読出データを取得する。
【０１４９】
また、読出先のストレージ装置１０２がＲＡＩＤ５のＲＡＩＤグループに属する場合、ストレージ制御装置１０１は、ＲＡＩＤ５のＲＡＩＤグループの読出先のストレージ装置１０２以外の全てのストレージ装置１０２に読出要求を発行する。そして、ストレージ制御装置１０１は、パリティデータと、パリティデータを生成する際に用いたデータ群のうち、読出先のストレージ装置１０２に格納した読出データ以外のデータを取得する。次に

、ストレージ制御装置１０１は、取得したパリティデータと、読出先のストレージ装置１０２に格納した読出データ以外のデータとから、読出先のストレージ装置１０２に格納した読出データを復元する。

【０１５０】

ステップＳ１３０７、または、ステップＳ１３１０の処理終了後、ストレージ制御装置１０１は、ホストサーバ２０１へ、読出要求に応じた読出データを通知する（ステップＳ１３１１）。ステップＳ１３０３、または、ステップＳ１３１１の処理終了後、ストレージ制御装置１０１は、ステップＳ１３０１の処理に移行する。ストレージ装置制御処理を実行することにより、ストレージ制御装置１０１は、ホストサーバ２０１のアクセスコマンドに応じてストレージ装置１０２を制御することができる。

【０１５１】

図１４は、書込時処理手順の一例を示すフローチャート（その１）である。また、図１５は、書込時処理手順の一例を示すフローチャート（その２）である。書込時処理は、ホストサーバ２０１からのアクセスコマンドの種別が書込要求である場合に実行する処理である。

【０１５２】

ストレージ制御装置１０１は、タイマが示す時間が所定間隔ｄｉを超えたか否かを判断する（ステップＳ１４０１）。このとき、書込先のストレージ装置１０２が、ＬＢＡ割り当てがない、まだ何も書き込まれていないストレージ装置１０２である場合、タイマがスタートしていない。この場合、ストレージ制御装置１０１は、ステップＳ１４０１：Ｎｏであると判断する。

【０１５３】

タイマが示す時間が所定間隔ｄｉを超えていない場合（ステップＳ１４０１：Ｎｏ）、ストレージ制御装置１０１は、書込先のストレージ装置１０２のブロック予測テーブル１１１に、書込先のＬＢＡがあるか否かを判断する（ステップＳ１４０２）。書込先のＬＢＡがない場合（ステップＳ１４０２：Ｎｏ）、ホストサーバ２０１からの書込要求は新規書込要求であるとして、ストレージ制御装置１０１は、割当中のグループを、書込先のグループに設定する（ステップＳ１４０３）。次に、ストレージ制御装置１０１は、書込先のグループが書込先のＬＢＡを含むようにグループ分けする（ステップＳ１４０４）。ステップＳ１４０４の処理終了後、ストレージ制御装置１０１は、グループ分けして生成したグループ情報を、ブロック予測テーブル１１１に格納する。一方、書込先のＬＢＡがある場合（ステップＳ１４０２：Ｙｅｓ）、ホストサーバ２０１からの書込要求は上書書込要求であるとして、ストレージ制御装置１０１は、書込先のＬＢＡが含まれるグループを、書込先のグループに設定する（ステップＳ１４０５）。

【０１５４】

タイマが示す時間が所定間隔ｄｉを超えた場合（ステップＳ１４０１：Ｙｅｓ）、ストレージ制御装置１０１は、書込先のストレージ装置１０２のブロック予測テーブル１１１に、書込先のＬＢＡがあるか否かを判断する（ステップＳ１４０６）。書込先のＬＢＡがない場合（ステップＳ１４０６：Ｎｏ）、ホストサーバ２０１からの書込要求は新規書込要求であるとして、ストレージ制御装置１０１は、割当中のグループを、現在割当中のグループの次のグループに設定する（ステップＳ１４０７）。次に、ストレージ制御装置１０１は、割当中のグループを、書込先のグループに設定する（ステップＳ１４０８）。続けて、ストレージ制御装置１０１は、書込先のグループが書込先のＬＢＡを含むようにグループ分けする（ステップＳ１４０９）。ステップＳ１４０９の処理終了後、ストレージ制御装置１０１は、グループ分けして生成したグループ情報を、ブロック予測テーブル１１１に格納する。一方、書込先のＬＢＡがある場合（ステップＳ１４０６：Ｙｅｓ）、ストレージ制御装置１０１は、ホストサーバ２０１からの書込要求は上書書込要求であるとして、書込先のＬＢＡが含まれるグループを、書込先のグループに設定する（ステップＳ１４１０）。

【０１５５】

ステップＳ１４０５、ステップＳ１４１０の処理において、ストレージ制御装置１０１は、書込先のＬＢＡを、同一グループに含まれるＬＢＡに対応するデータサイズがブロックデータサイズを超えないようにグループ分けする。書込先のＬＢＡがブロックデータサイズに収まらない場合、ストレージ制御装置１０１は、書込先のＬＢＡを、ＬＢＡ順に複数のグループにグループ分けする。

【０１５６】

ステップＳ１４０４、ステップＳ１４０５、ステップＳ１４０９、ステップＳ１４１０のうちのいずれかの処理終了後、ストレージ制御装置１０１は、書込先のグループの読出回数を０に初期化する（ステップＳ１４１１）。ホストサーバ２０１からの書込要求が上書書込要求である場合、ステップＳ１４１１を実行することにより、ストレージ制御装置１０１は、書込要求に含まれるＬＢＡが含まれるブロックの読出回数を初期化することになる。

【０１５７】

ステップＳ１４１１の処理終了後、ストレージ制御装置１０１は、ＲＤ防止処理フラグが次に示す識別子のいずれに一致するかを判断する（ステップＳ１５０１）。次に示す識別子は、「処理していない」と、「処理中」と、である。ＲＤ防止処理フラグが「処理していない」である場合（ステップＳ１５０１：“処理していない”）、ストレージ制御装置１０１は、書込先のストレージ装置１０２へ書込データを含む書込要求を発行する（ステップＳ１５０２）。

【０１５８】

一方、ＲＤ防止処理フラグが「処理中」である場合（ステップＳ１５０１：“処理中”）、ストレージ制御装置１０１は、続けて、書込先のＬＢＡがＲＤ防止コピーの対象のグループに含まれるか否かを判断する（ステップＳ１５０３）。

【０１５９】

書込先のＬＢＡがＲＤ防止コピーの対象のグループに含まれない場合（ステップＳ１５０３：Ｎｏ）、ストレージ制御装置１０１は、ＲＤ防止コピーを引き起こすコマンドの発行を一時停止する（ステップＳ１５０４）。次に、ストレージ制御装置１０１は、書込先のストレージ装置１０２へ書込データを含む書込要求を発行する（ステップＳ１５０５）。続けて、ストレージ制御装置１０１は、ＲＤ防止コピーを引き起こすコマンドの発行を再開する（ステップＳ１５０６）。

【０１６０】

一方、書込先のＬＢＡがＲＤ防止コピーの対象のグループに含まれる場合（ステップＳ１５０３：Ｙｅｓ）、ストレージ制御装置１０１は、ＲＤ防止コピーを引き起こすコマンドの発行を停止させる（ステップＳ１５０７）。次に、ストレージ制御装置１０１は、書込先のストレージ装置へ書込データを含む書込要求を発行する（ステップＳ１５０８）。続けて、ストレージ制御装置１０１は、ＲＤ防止処理フラグを、「処理していない」に変更する（ステップＳ１５０９）。

【０１６１】

ステップＳ１５０２、ステップＳ１５０６、ステップＳ１５０９のうちのいずれかの処理終了後、ストレージ制御装置１０１は、タイマをスタートする（ステップＳ１５１０）。ステップＳ１５１０の処理終了後、ストレージ制御装置１０１は、書込時処理を終了する。書込時処理を実行することにより、ストレージ制御装置１０１は、ホストサーバ２０１からの書込要求に対応してストレージ装置１０２を制御することができる。

【０１６２】

図１６は、ＲＤ防止コピー引起し処理手順の一例を示すフローチャートである。ＲＤ防止コピー引起し処理は、ストレージ装置１０２にＲＤ防止コピーを起こさせる処理である。図１６に示す処理について、ブロック予測テーブル１１１とＲＤ防止処理フラグとは、図１３における読出先のストレージ装置１０２におけるブロック予測テーブル１１１とＲＤ防止処理フラグとを指す。説明の簡略化のため、図１６において、ブロック予測テーブル１１１とＲＤ防止処理フラグには、「読出先のストレージ装置１０２における」を省略

して記載する。

【０１６３】

ストレージ制御装置１０１は、ＲＤ防止処理フラグを、「処理中」に変更する（ステップＳ１６０１）。次に、ストレージ制御装置１０１は、読出先のストレージ装置へ、ＲＤ防止コピーを引き起こすコマンドを発行する（ステップＳ１６０２）。続けて、ストレージ制御装置１０１は、ＲＤ防止コピーが行われたか否かを判断する（ステップＳ１６０３）。ＲＤ防止コピーが行われたか否かを判断する方法は、上述した第１のＲＤ防止コピー実行判断方法と、第２のＲＤ防止コピー実行判断方法とがある。

【０１６４】

ＲＤ防止コピーが行われていない場合（ステップＳ１６０３：Ｎｏ）、ストレージ制御装置１０１は、ステップＳ１６０２の処理に移行する。ＲＤ防止コピーが行われた場合（ステップＳ１６０３：Ｙｅｓ）、ストレージ制御装置１０１は、ブロック予測テーブルの読出先のＬＢＡが含まれるグループの読出回数の値を０に初期化する（ステップＳ１６０４）。次に、続けて、ストレージ制御装置１０１は、ＲＤ防止処理フラグを、「処理していない」に変更する（ステップＳ１６０５）。ステップＳ１６０５の処理終了後、ストレージ制御装置１０１は、ＲＤ防止コピー引起し処理を終了する。ＲＤ防止コピー引起し処理を実行することにより、ストレージ制御装置１０１は、ストレージ装置１０２にＲＤ防止コピーを起こさせることができる。

【０１６５】

図１７は、リフレッシュ処理手順の一例を示すフローチャートである。リフレッシュ処理は、移行元のストレージ装置１０２が記憶する移行対象データを移行先のストレージ装置に移行させて、グループとＬＢＡとの関係を再定義する処理である。

【０１６６】

ストレージ制御装置１０１は、移行元のストレージ装置１０２の全てのＬＢＡを、移行先のストレージ装置１０２に移行するデータ移行要求を、移行元のストレージ装置に発行する（ステップＳ１７０１）。次に、ストレージ制御装置１０１は、移行元のストレージ装置１０２のブロック予測テーブル１１１の先頭ＬＢＡを選択する（ステップＳ１７０２）。続けて、ストレージ制御装置１０１は、移行先のストレージ装置１０２の割当中のグループを、先頭のグループに設定する（ステップＳ１７０３）。

【０１６７】

次に、ストレージ制御装置１０１は、移行先のストレージ装置１０２のブロック予測テーブル１１１の割当中のグループに、選択したＬＢＡを先頭として１ブロック分のＬＢＡを含むようにグループ分けする（ステップＳ１７０４）。続けて、ストレージ制御装置１０１は、移行元のストレージ装置１０２の全てのＬＢＡをグループ分けしたか否かを判断する（ステップＳ１７０５）。

【０１６８】

移行元のストレージ装置１０２のＬＢＡのうち、グループ分けしていないＬＢＡがある場合（ステップＳ１７０５：Ｎｏ）、ストレージ制御装置１０１は、移行元のストレージ装置１０２のＬＢＡのうち、まだグループ分けしていない先頭のＬＢＡを選択する（ステップＳ１７０６）。続けて、ストレージ制御装置１０１は、割当中のグループを、現在割当中のグループの次のグループに設定する（ステップＳ１７０７）。ステップＳ１７０７の処理終了後、ストレージ制御装置１０１は、ステップＳ１７０４の処理に移行する。

【０１６９】

一方、移行元のストレージ装置１０２の全てのＬＢＡをグループ分けした場合（ステップＳ１７０５：Ｙｅｓ）、ストレージ制御装置１０１は、リフレッシュ処理を終了する。リフレッシュ処理を実行することにより、ストレージ制御装置１０１は、ストレージ装置１０２が管理するグループとＬＢＡ割り当ての関係と、ストレージ制御装置１０１が予測するグループとＬＢＡ割り当ての関係との乖離を解消することができる。

【０１７０】

以上説明したように、ストレージ制御装置１０１によれば、書込要求のＬＢＡをブロッ

ク単位でグループ分けしたグループの読出回数を計数し、計数したグループの読出回数に基づいて元データか冗長データかを読み出す読出要求をストレージ装置１０２に発行する。これにより、ストレージ制御装置１０１は、読み出しを行うとＲＤ防止コピーが発生するであろうＬＢＡを予測することができる。そして、ストレージ制御装置１０１は、ホストサーバ２０１からの読出要求に対する応答性能の低下を抑制することができる。また、ストレージ制御装置１０１は、ストレージ装置１０２の内部処理とストレージ制御装置１０１からの読出要求とのタイミングが重なることを防ぎ、安定した応答時間を得ることができる。

【０１７１】

また、ストレージ制御装置１０１によれば、ストレージ制御装置１０１が受け付けた書込要求に応じて、書込要求のＬＢＡをブロック単位でグループ分けしてもよい。これにより、ストレージ制御装置１０１は、ストレージ装置１０２への書込要求に追随してグループ分けが行えるため、ＲＤ防止コピーが発生するであろうＬＢＡの予測精度の低下を抑制することができる。

【０１７２】

また、ストレージ制御装置１０１によれば、ストレージ装置１０２への読出要求に応じて、読出要求に含まれる読出先のＬＢＡを含むグループに含まれるＬＢＡの書込要求を発行したことに応じて、グループの読出回数を初期化してもよい。ＳＳＤコントローラは、上書書込が行われると該当するブロックの読出回数を初期化するため、ストレージ制御装置１０１は、予測精度の低下を抑制することができる。

【０１７３】

また、ストレージ制御装置１０１によれば、グループの読出回数がＲＤ防止コピー閾値以上であることに応じて、ＲＤ防止コピーを引き起こすコマンドをストレージ装置１０２に発行して、グループの読出回数を初期化してもよい。ここで、ストレージ制御装置１０１は、ストレージ装置１０２にＲＤ防止コピーを引き起こすコマンドを、ストレージ装置１０２の仕様に応じて設定された規定回数からＲＤ防止コピー閾値を減じた回数分発行する。または、ストレージ制御装置１０１は、ストレージ装置１０２がＲＤ防止コピーを引き起こすコマンドを、コマンドに対する応答時間が所定時間を超えるまで発行する。これにより、ストレージ制御装置１０１は、ＲＤ防止コピーが発生するであろうＬＢＡに対しＲＤ防止コピーを事前に起こしておくことになる。したがって、ホストサーバ２０１からの読出要求を受け付ける前にＲＤ防止コピーが完了していれば、ホストサーバ２０１からの読出要求に対する応答遅延が発生しなくなる。これにより、ストレージ制御装置１０１は、ホストサーバ２０１からの読出要求に対する応答性能の低下を抑制することができる。

【０１７４】

また、ストレージ制御装置１０１によれば、ＲＤ防止コピーを引き起こすコマンドを発行する間に、コマンドのＬＢＡを含むグループに含まれるＬＢＡの書込要求を発行したことに応じて、コマンドの発行を停止するとともにグループの読出回数を初期化してもよい。ＳＳＤコントローラは、上書書込が行われると該当するブロックの読出回数を初期化するため、ストレージ制御装置１０１は、読出回数の予測精度を向上することができる。さらに、上書書込が行われればＲＤ防止コピーを行わなくてよいので、コマンドの発行を停止することにより、ストレージ制御装置１０１は、無駄なコマンドの発行を行わなくて済む。

【０１７５】

また、ストレージ制御装置１０１によれば、同一グループに含まれるＬＢＡを含む書込要求のストレージ装置１０２への発行間隔が所定間隔ｄｉを超えないようにグループ分けしてもよい。これにより、ストレージ制御装置１０１は、ストレージ装置１０２が行うブロックとＬＢＡ割り当てとの関係の予測精度を向上することができる。

【０１７６】

また、本実施の形態にかかるストレージ装置１０２の半導体メモリは、ＮＡＮＤ型のフ

ラッシュメモリであってもよい。ＮＡＮＤ型のフラッシュメモリは、ＮＯＲ型のフラッシュメモリと比較してリードディスターブが発生し易いため、ＮＡＮＤ型のフラッシュメモリを制御するＳＳＤコントローラは、ＲＤ防止コピーを行う回数が多くなることになる。したがって、ＮＡＮＤ型のフラッシュメモリを有するストレージ装置１０２に対して本実施の形態におけるストレージ装置の制御方法を実行することにより、多く発生する応答性能が遅延する状態を起こさないようにして、応答性能の低下を抑制することができる。

【０１７７】

また、ストレージ制御装置１０１によれば、グループの読出回数がＲＤ防止コピー閾値以上であれば、グループに含まれるＬＢＡの読出要求に応じて、読出先のデータを復元可能なデータ先を示すＬＢＡの読出要求を、他のストレージ装置１０２に発行してもよい。これにより、ストレージ制御装置１０１は、ＲＤ防止コピーが発生するであろうストレージ装置１０２を避けて、ホストサーバ２０１の読出要求に応答することができ、読出要求に対する応答性能の低下を抑制することができる。

【０１７８】

また、ストレージ制御装置１０１によれば、同一の特性を有するストレージ装置１０２群によりＲＡＩＤグループを形成する場合、ストレージ装置１０２群の各々のＲＤ防止コピー閾値をずらして設定してもよい。これにより、ストレージ制御装置１０１は、ＲＤ防止コピーが行われる期間をストレージ装置１０２間でずらすことができる。ＲＤ防止コピーが行われる期間をずらすことにより、ストレージ制御装置１０１は、ミラーリングからのデータ取得ができない状態、またはパリティを用いたデータ復元ができない状態になることを避け、読出要求に対する応答性能の低下を抑制することができる。

【０１７９】

また、ストレージ制御装置１０１によれば、移行元のストレージ装置１０２が記憶する移行対象データを移行先のストレージ装置１０２に移行させて、グループとＬＢＡとの関係を再定義してもよい。これにより、ストレージ制御装置１０１は、ストレージ制御装置１０１が予測するグループとＬＢＡ割り当ての関係と移行先のストレージ装置１０２が管理するブロックとＬＢＡ割り当てとの内容が一致して、ＲＤ防止コピーが発生するＬＢＡの予測精度が向上する。

【０１８０】

また、ストレージ制御装置１０１によれば、移行先のストレージ装置１０２に対してもグループ分けや、グループの読出回数の計数を行うことにより、移行先のストレージ装置１０２に対しても、ＲＤ防止コピーが発生するＬＢＡの予測を行うことができる。

【０１８１】

なお、本実施の形態で説明したストレージ装置の制御方法は、予め用意されたプログラムをパーソナル・コンピュータやワークステーション等のコンピュータで実行することにより実現することができる。本ストレージ装置の制御プログラムは、ハードディスク、フレキシブルディスク、光ディスク等のコンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行される。また本ストレージ装置の制御プログラムは、インターネット等のネットワークを介して配布してもよい。

【０１８２】

上述した実施の形態に関し、さらに以下の付記を開示する。

【０１８３】

（付記１）ストレージ装置の記憶領域に対して所定のデータサイズのブロック単位でデータを冗長化して記憶させる制御を行うストレージ制御装置であって、

データの書込要求から特定される書込先の論理アドレスを、前記ブロック単位に対応付けてグループ分けしたグループ情報を記憶する記憶部と、

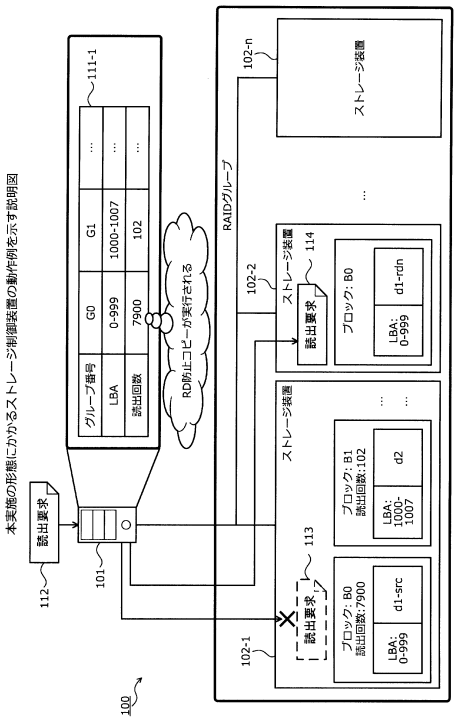
データの読出要求に応じて、前記グループ情報に基づき前記読出要求から特定される読出先の論理アドレスを含むグループの読出回数を計数し、計数した前記グループの読出回数に基づき前記読出先の論理アドレスを含む読出要求、または、前記読出先の論理アドレ

スのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行する制御部と、
を有することを特徴とするストレージ制御装置。
【０１８４】
（付記２）前記記憶領域は、規定回数の読み出しが行われるとブロックのデータを他のブロックへコピーするコピー制御が行われる半導体メモリの記憶領域であることを特徴とする付記１に記載のストレージ制御装置。
【０１８５】
（付記３）前記ストレージ制御装置は、前記半導体メモリを含む前記ストレージ装置と、当該ストレージ装置が記憶するデータの冗長データを記憶する前記ストレージ装置とは異なる他のストレージ装置の制御を行うものであり、
前記制御部は、
計数した前記グループの読出回数が前記規定回数より小さい第１の回数となった後に、データの読出要求に応じて、当該読出要求から特定される読出先の論理アドレスが前記グループに含まれる場合に、当該読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を、前記他のストレージ装置に発行する、
ことを特徴とする付記２に記載のストレージ制御装置。
【０１８６】
（付記４）前記制御部は、
計数した前記グループの読出回数が前記規定回数より小さい第１の回数となったことに応じて、前記グループに含まれる論理アドレスを含む読出要求を、前記規定回数から前記第１の回数を減じた第２の回数、または、当該読出要求に対する応答時間が所定時間を超えるまで前記ストレージ装置に発行し、前記グループの読出回数を初期化することを特徴とする付記２または３に記載のストレージ制御装置。
【０１８７】
（付記５）前記制御部は、
前記グループに含まれる論理アドレスを含む読出要求を、前記第２の回数、または、当該読出要求に対する応答時間が所定時間を超えるまで前記ストレージ装置に発行する間に、前記グループに含まれる論理アドレスを含む書込要求を前記ストレージ装置に発行したことに応じて、前記グループに含まれる論理アドレスを含む読出要求の発行を停止するとともに、前記グループの読出回数を初期化することを特徴とする付記４に記載のストレージ制御装置。
【０１８８】
（付記６）前記制御部は、
前記グループに含まれる論理アドレスを含む書込要求を前記ストレージ装置に発行したことに応じて、前記グループの読出回数を初期化することを特徴とする付記１または２に記載のストレージ制御装置。
【０１８９】
（付記７）前記記憶部は、
書込要求から特定される前記他のストレージ装置の書込先の論理アドレスを、前記ブロック単位に対応付けてグループ分けした前記他のストレージ装置のグループ情報を記憶しており、
前記制御部は、
受け付けた読出要求に応じて、前記他のストレージ装置のグループ情報に基づき当該読出要求から特定される前記他のストレージ装置の読出先の論理アドレスを含むグループの読出回数を計数し、計数した当該グループの読出回数が前記規定回数より小さく前記第１の回数とは異なる第３の回数となったことに応じて、当該グループに含まれる論理アドレスを含む読出要求を、前記規定回数から前記第３の回数を減じた第４の回数、または、当該読出要求に対する応答時間が所定時間を超えるまで前記他のストレージ装置に発行し、当該グループの読出回数を初期化することを特徴とする付記３に記載のストレージ制御装

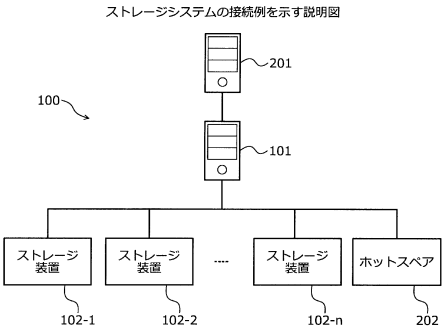
置。
【０１９０】
（付記８）前記ストレージ制御装置は、さらに、半導体メモリの記憶領域を所定のデータサイズで分割したブロック単位で書き込みを行う移行先のストレージ装置の制御を行うものであり、
前記制御部は、
受け付けた書込要求から特定される前記ストレージ装置の書込先の論理アドレスをグループ分けした複数のグループに含まれる論理アドレスを含むデータ移行要求を、前記移行先のストレージ装置へのデータの移行間隔が所定間隔を超えないように前記ストレージ装置に発行し、
発行した前記データ移行要求に含まれる論理アドレスを、前記ブロック単位に対応付けてグループ分けした前記移行先のストレージ装置のグループ情報を生成することを特徴とする付記１～７のいずれか一つに記載のストレージ制御装置。
【０１９１】
（付記９）前記制御部は、
受け付けた書込要求に応じて、当該書込要求から特定される前記移行先のストレージ装置の書込先の論理アドレスを、同一グループに含まれる論理アドレスに対応するデータサイズが当該所定のデータサイズを超えないようにグループ分けし、
受け付けた読出要求に応じて、当該読出要求から特定される前記移行先のストレージ装置の読出先の論理アドレスを含むグループの読出回数を計数し、
計数した当該グループの読出回数に基づき当該読出先の論理アドレスを含む読出要求、または、当該読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行する付記８に記載のストレージ制御装置。
【０１９２】
（付記１０）ストレージ装置の記憶領域に対して所定のデータサイズのブロック単位でデータを冗長化して記憶させる制御を行うストレージ制御装置の制御方法であって、
前記ストレージ制御装置が、
データの読出要求に応じて、データの書込要求から特定される書込先の論理アドレスを、前記ブロック単位に対応付けてグループ分けしたグループ情報に基づき前記読出要求から特定される読出先の論理アドレスを含むグループの読出回数を計数し、計数した前記グループの読出回数に基づき前記読出先の論理アドレスを含む読出要求、または、前記読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行する、
処理を実行することを特徴とする制御方法。
【０１９３】
（付記１１）ストレージ装置の記憶領域に対して所定のデータサイズのブロック単位でデータを冗長化して記憶させる制御を行うストレージ制御装置の制御プログラムであって、
前記ストレージ制御装置に、
データの読出要求に応じて、データの書込要求から特定される書込先の論理アドレスを、前記ブロック単位に対応付けてグループ分けしたグループ情報に基づき前記読出要求から特定される読出先の論理アドレスを含むグループの読出回数を計数し、計数した前記グループの読出回数に基づき前記読出先の論理アドレスを含む読出要求、または、前記読出先の論理アドレスのデータに対応する冗長データの記憶先の論理アドレスを含む読出要求を発行する、
処理を実行させることを特徴とする制御プログラム。
【符号の説明】
【０１９４】
１００　ストレージシステム
１０１　ストレージ制御装置
１０２　ストレージ装置

１１１　ブロック予測テーブル
４０１　制御部
４０２　記憶部
４１１　グループ分け部
４１２　計数部
４１３　アクセス制御部

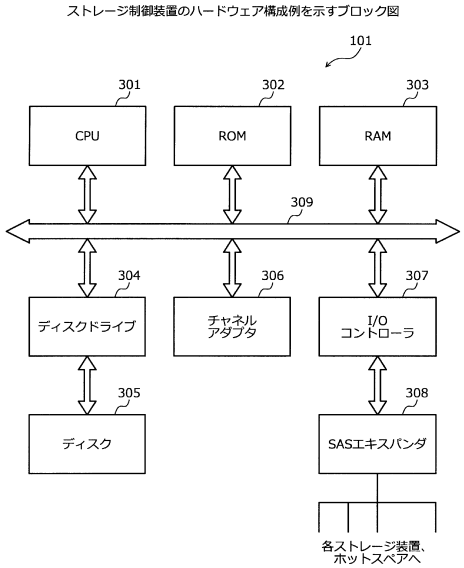
【図１】



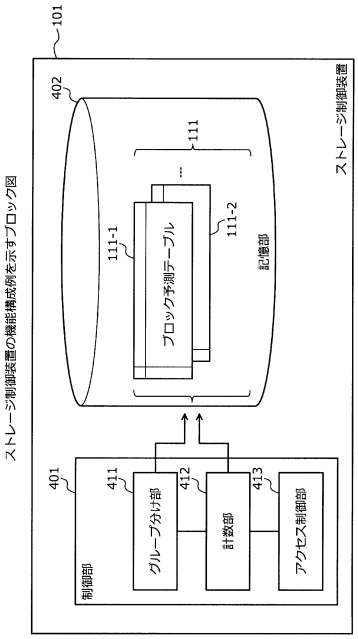
【図２】



【図３】

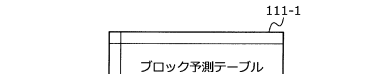


【図４】



【図５】

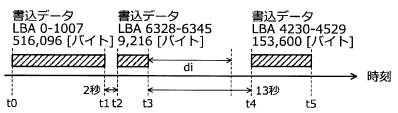
ブロック予測テーブルの記憶内容の一例を示す説明図



	501-1	501-2	501-3	
グループ番号	G0	G1	G2	…
LBA	0-999	1000-1007	(未割当)	…
読出回数	7700	7700	0	…

【図６】

書込時における動作例を示す説明図



割当中

時刻 t0

グループ番号	G0	G1	G2	…
LBA	0-999	1000-1007	(未割当)	…
読出回数	0	0	0	…

割当中

時刻 t2

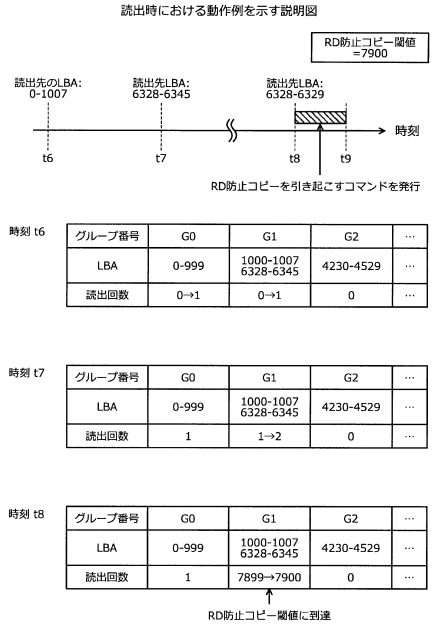
グループ番号	G0	G1	G2	…
LBA	0-999	1000-1007 6328-6345	(未割当)	…
読出回数	0	0	0	…

割当中

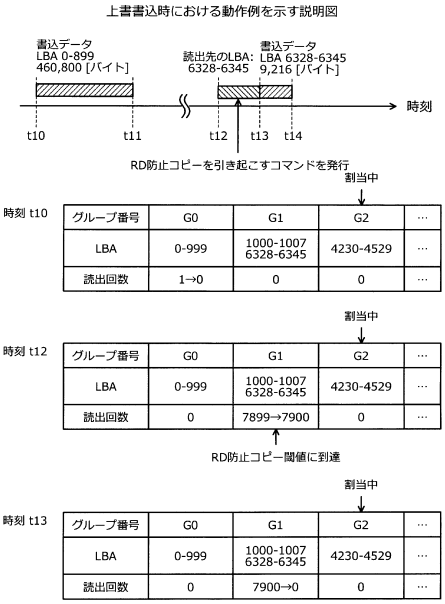
時刻 t4

グループ番号	G0	G1	G2	…
LBA	0-999	1000-1007 6328-6345	4230-4529	…
読出回数	0	0	0	…

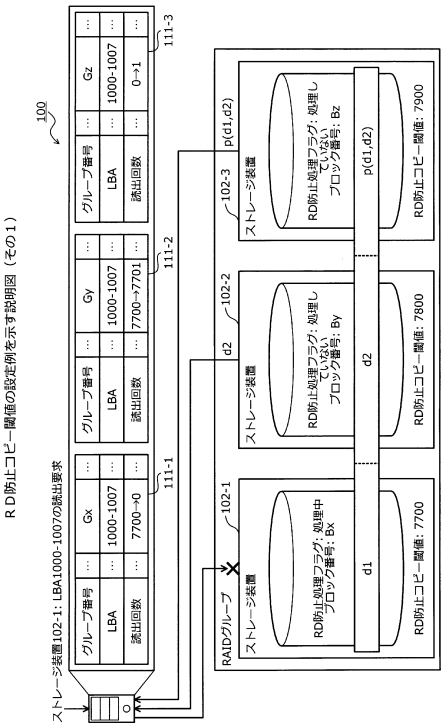
【図７】


読出時における動作例を示す説明図
RD防止コピー閾値
=7900
読出先のLBA:
0-1007
読出先LBA:
6328-6345
読出先LBA:
6328-6329
時刻
t6
t7
t8
t9
RD防止コピーを引き起こすコマンドを発行
時刻 t6
グループ番号 G0 G1 G2 …
LBA 0-999 1000-1007 6328-6345 4230-4529 …
読出回数 0→1 0→1 0 …
時刻 t7
グループ番号 G0 G1 G2 …
LBA 0-999 1000-1007 6328-6345 4230-4529 …
読出回数 1 1→2 0 …
時刻 t8
グループ番号 G0 G1 G2 …
LBA 0-999 1000-1007 6328-6345 4230-4529 …
読出回数 1 7899→7900 0 …
RD防止コピー閾値に到達

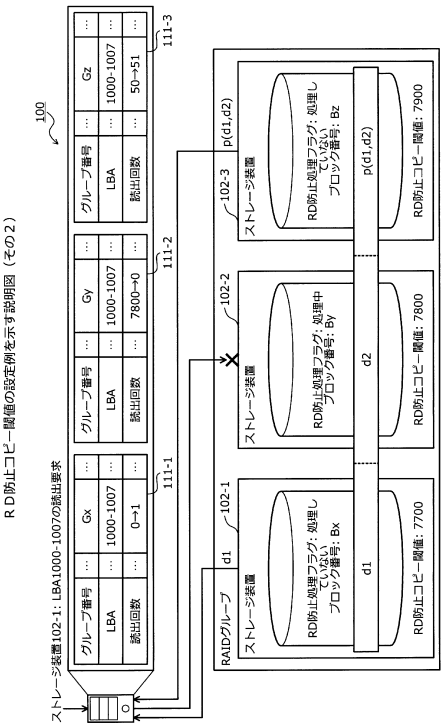
【図８】


上書書込時における動作例を示す説明図
書込データ
LBA 0-899
460,800 [バイト]
読出先のLBA:
6328-6345
書込データ
LBA 6328-6345
9,216 [バイト]
時刻
t10
t11
t12
t13
t14
RD防止コピーを引き起こすコマンドを発行
割当中
時刻 t10
グループ番号 G0 G1 G2 …
LBA 0-999 1000-1007 6328-6345 4230-4529 …
読出回数 1→0 0 0 …
割当中
時刻 t12
グループ番号 G0 G1 G2 …
LBA 0-999 1000-1007 6328-6345 4230-4529 …
読出回数 0 7899→7900 0 …
RD防止コピー閾値に到達
割当中
時刻 t13
グループ番号 G0 G1 G2 …
LBA 0-999 1000-1007 6328-6345 4230-4529 …
読出回数 0 7900→0 0 …

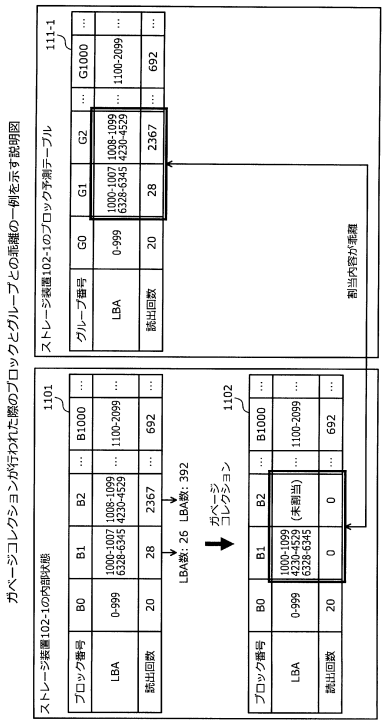
【図９】


RD防止コピー閾値の設定例を示す説明図（その１）
ストレージ装置102-1: LBA1000-1007の読出要求
100
111-3
111-2
111-1
グループ番号 … Gz … …
LBA … 1000-1007 … …
読出回数 … 0→1 … …
グループ番号 … Gy … …
LBA … 1000-1007 … …
読出回数 … 7700→7701 … …
グループ番号 … Gx … …
LBA … 1000-1007 … …
読出回数 … 7700→0 … …
RAIDグループ
ストレージ装置
102-1
102-2
102-3
RD防止処理フラグ: 処理中
ブロック番号: Bx
RD防止処理フラグ: 処理中
ブロック番号: By
RD防止処理フラグ: 処理しない
ブロック番号: Bz
d1
d2
p(d1,d2)
RD防止コピー閾値: 7700
RD防止コピー閾値: 7800
RD防止コピー閾値: 7900

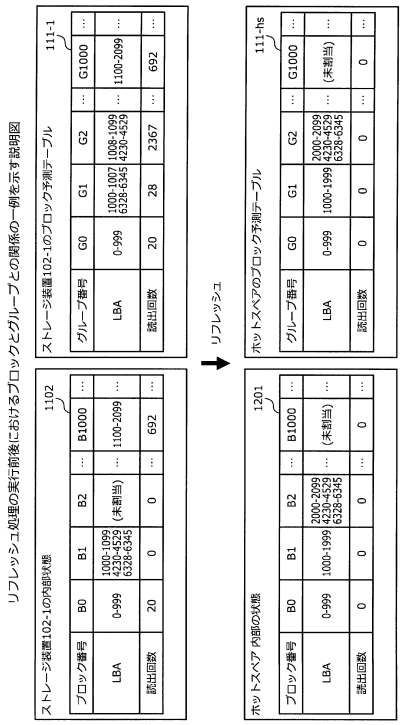
【図１０】


RD防止コピー閾値の設定例を示す説明図（その２）
ストレージ装置102-1: LBA1000-1007の読出要求
100
111-3
111-2
111-1
グループ番号 … Gz … …
LBA … 1000-1007 … …
読出回数 … 50→51 … …
グループ番号 … Gy … …
LBA … 1000-1007 … …
読出回数 … 7800→0 … …
グループ番号 … Gx … …
LBA … 1000-1007 … …
読出回数 … 0→1 … …
RAIDグループ
ストレージ装置
102-1
102-2
102-3
RD防止処理フラグ: 処理中
ブロック番号: Bx
RD防止処理フラグ: 処理中
ブロック番号: By
RD防止処理フラグ: 処理しない
ブロック番号: Bz
d1
d2
p(d1,d2)
RD防止コピー閾値: 7700
RD防止コピー閾値: 7800
RD防止コピー閾値: 7900

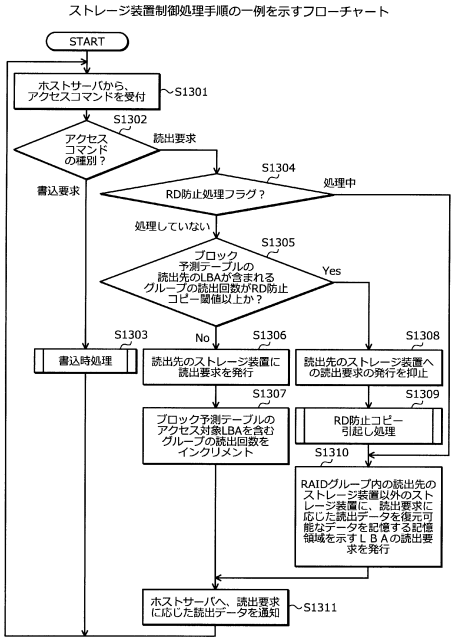
【図１１】



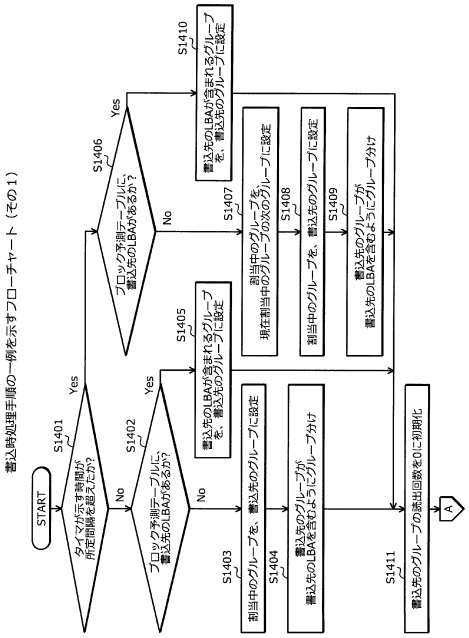
【図１２】



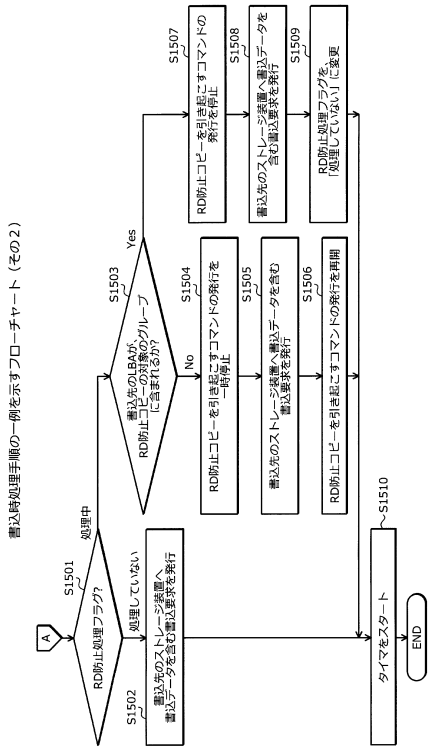
【図１３】



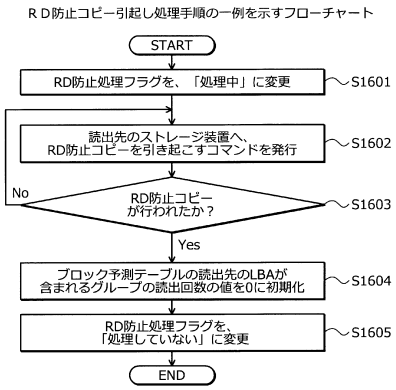
【図１４】



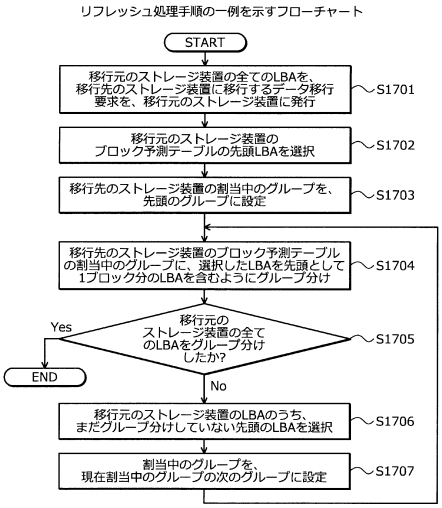
【図１５】


書込時処理手順の一例を示すフローチャート（その２）
A
S1501
RD防止処理フラグ?
処理中
処理していない
S1502
書込先のストレージ装置へ
書込データを含む書込要求を発行
S1503
書込先のLBAが、
RD防止コピーの対象のグループ
に含まれるか?
Yes
No
S1504
RD防止コピーを引き起こすコマンドの発行を
一時停止
S1505
書込先のストレージ装置へ書込データを含む
書込要求を発行
S1506
RD防止コピーを引き起こすコマンドの発行を再開
S1507
RD防止コピーを引き起こすコマンドの
発行を停止
S1508
書込先のストレージ装置へ書込データを
含む書込要求を発行
S1509
RD防止処理フラグを、
「処理していない」に変更
S1510
タイマをスタート
END

【図１６】


ＲＤ防止コピー引起し処理手順の一例を示すフローチャート
START
RD防止処理フラグを、「処理中」に変更
S1601
読出先のストレージ装置へ、
RD防止コピーを引き起こすコマンドを発行
S1602
No
RD防止コピー
が行われたか？
S1603
Yes
ブロック予測テーブルの読出先のLBAが
含まれるグループの読出回数の値を0に初期化
S1604
RD防止処理フラグを、
「処理していない」に変更
S1605
END

【図１７】


リフレッシュ処理手順の一例を示すフローチャート
START
移行元のストレージ装置の全てのLBAを、
移行先のストレージ装置に移行するデータ移行
要求を、移行元のストレージ装置に発行
S1701
移行元のストレージ装置の
ブロック予測テーブルの先頭LBAを選択
S1702
移行先のストレージ装置の割当中のグループを、
先頭のグループに設定
S1703
移行先のストレージ装置のブロック予測テーブル
の割当中のグループに、選択したLBAを先頭として
1ブロック分のLBAを含むようにグループ分け
S1704
Yes
移行元の
ストレージ装置の全て
のLBAをグループ分け
したか?
S1705
END
No
移行元のストレージ装置のLBAのうち、
まだグループ分けしていない先頭のLBAを選択
S1706
割当中のグループを、
現在割当中のグループの次のグループに設定
S1707

フロントページの続き

(56)参考文献 特開２０１２－２３８２５９（ＪＰ，Ａ）
特開２０１２－２０３６４２（ＪＰ，Ａ）
米国特許出願公開第２００９／０１９３１７４（ＵＳ，Ａ１）
米国特許出願公開第２０１３／０２６２７５０（ＵＳ，Ａ１）
米国特許出願公開第２０１１／００３８２０３（ＵＳ，Ａ１）
特開２００９－０３７３１７（ＪＰ，Ａ）
特開２００８－２８７４０４（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ３／０６－３／０８
Ｇ０６Ｆ１２／００－１２／０６、１２／１６
Ｇ０６Ｆ１３／１６－１３／１８