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

特許第6206160号
(P6206160)

(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	306Z
G06F 3/08	(2006.01)	G06F	3/08	H
G06F 12/16	(2006.01)	G06F	3/06	306B
		G06F	3/06	305C
		G06F	12/16	310Q

請求項の数 5　(全 21 頁)　最終頁に続く

(21)出願番号　特願2013-261803(P2013-261803)
(22)出願日　平成25年12月18日(2013.12.18)
(65)公開番号　特開2015-118561(P2015-118561A)
(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　100074099
弁理士　大菅　義之
(74)代理人　100133570
弁理士　▲徳▼永　民雄
(72)発明者　長　絵美
神奈川県川崎市中原区上小田中４丁目１番１号　株式会社富士通コンピュータテクノロジーズ内

最終頁に続く

(54)【発明の名称】ストレージ制御装置、ストレージ制御プログラム及びストレージ制御方法

(57)【特許請求の範囲】
【請求項１】
複数のＲＡＩＤのグループを形成する複数のメモリブロックを備え、更に、前記複数のメモリブロックの予備として使用する予備ブロックを備える記憶装置との間でデータの入出力を行うインターフェースと、
前記記憶装置中の各メモリブロックが備える複数のメモリセルから、故障したメモリブロックに含まれるメモリセルを読み出し対象に選択すると、前記故障したメモリブロックのデータの修復及び前記予備ブロックへの修復したデータのコピーが完了するまで、前記故障したメモリブロックを有するＲＡＩＤのグループに属するメモリセルを、前記読み出し対象から外し、且つ、
前記故障したメモリブロックを有するＲＡＩＤのグループ以外のメモリセルを前記読み出し対象として選択し、
前記故障したメモリブロックのデータの修復及び前記予備ブロックへの修復したデータのコピーが完了すると、データの修復が完了したメモリブロックを有するＲＡＩＤのグループに属するメモリセル及び前記予備ブロック中のメモリセルを、前記読み出し対象として選択する制御を、前記インターフェースを介して行う制御部と、を備える
ことを特徴とするストレージ制御装置。
【請求項２】
前記制御部は、
前記記憶装置中の全メモリセルが読み出し対象に指定されると、

前記全メモリセルのうち、前記故障したメモリブロックを有するＲＡＩＤのグループ以外に含まれるメモリセルの読み出し処理が完了し、且つ、修復したデータの予備のブロックへのコピーを完了すると、前記故障したメモリブロックを有するＲＡＩＤのグループに属するメモリセル及び前記予備ブロック中のメモリセルを、前記読み出し対象として選択する

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、

前記記憶装置中の全メモリセルが読み出し対象に指定されると、

前記ＲＡＩＤのグループ毎に、１つのメモリブロックに含まれるメモリセルを前記読み出し対象の候補とする

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４】

複数のＲＡＩＤのグループを形成する複数のメモリブロックを備え、更に、前記複数のメモリブロックの予備として使用する予備ブロックを備える記憶装置を読み出し対象とするストレージ制御装置が、

前記記憶装置中の各メモリブロックが備える複数のメモリセルから、故障したメモリブロックに含まれるメモリセルを前記読み出し対象に選択すると、前記故障したメモリブロックのデータの修復及び前記予備ブロックへの修復したデータのコピーが完了するまで、前記故障したメモリブロックを有するＲＡＩＤのグループに属するメモリセルを、前記読み出し対象から外し、且つ、

前記故障したメモリブロックを有するＲＡＩＤのグループ以外のメモリセルを前記読み出し対象として選択し、

前記故障したメモリブロックのデータの修復及び前記予備ブロックへの修復したデータのコピーが完了すると、データの修復が完了したメモリブロックを有するＲＡＩＤのグループに属するメモリセル及び前記予備ブロック中のメモリセルを、前記読み出し対象として選択する

ことを特徴とするストレージ制御方法。

【請求項５】

複数のＲＡＩＤのグループを形成する複数のメモリブロックを備え、更に、前記複数のメモリブロックの予備として使用する予備ブロックを備える記憶装置を読み出し対象とするストレージ制御装置に、

前記記憶装置中の各メモリブロックが備える複数のメモリセルから、故障したメモリブロックに含まれるメモリセルが前記読み出し対象に選択されると、前記故障したメモリブロックのデータの修復及び前記予備ブロックへの修復したデータのコピーが完了するまで、前記故障したメモリブロックを有するＲＡＩＤのグループに属するメモリセルを、前記読み出し対象から外し、且つ、

前記故障したメモリブロックを有するＲＡＩＤのグループ以外のメモリセルを前記読み出し対象として選択し、

前記故障したメモリブロックのデータの修復及び前記予備ブロックへの修復したデータのコピーが完了すると、データの修復が完了したメモリブロックを有するＲＡＩＤのグループに属するメモリセル及び前記予備ブロック中のメモリセルを、前記読み出し対象として選択する制御をする

処理を実行させることを特徴とするストレージ制御プログラム。

【発明の詳細な説明】

【技術分野】

【０００１】

この発明は、メモリからのデータの読み出し方法に関する。

【背景技術】

【０００２】

サーバ装置は、ＳＳＤ（Ｓｏｌｉｄ－Ｓｔａｔｅ　Ｄｒｉｖｅ）やＨＤＤ（Ｈａｒｄ　Ｄｉｓｃ　Ｄｒｉｖｅ）を記憶装置として搭載する。ＨＤＤは、磁性体を塗布した円盤ディスクを高速回転し、磁気ヘッドを移動することで、ディスクに格納された情報を読み出す記憶装置である。ＳＳＤは、ＮＡＮＤ型フラッシュメモリを記憶媒体として用いる記憶装置である。サーバ装置は、ＳＳＤやＨＤＤなどの各メモリを読込むことで、媒体の正常性診断を行うパトロールリード機能を有する。

【０００３】

ＳＳＤは、複数のメモリブロック（ブロック）を有する。各メモリブロックは、複数のページを有する。更に、各ページは、複数のメモリセルを有する。正常性診断の処理を制御する制御部は、メモリセル単位で、正常性の診断をする。制御部は、診断対象のメモリセルの先頭のアドレスを指定する。ＳＳＤは、指定されたアドレスを含むメモリセルからデータを読込み、制御部にデータを返信する。制御部は、ＳＳＤから指定したアドレスを含むメモリセルのデータを受信すると、そのメモリセルは正常であると判定する。制御部は、正常性診断がまだ行われていないメモリセルの先頭のアドレスを順番に指定し、残りのメモリセルに対しても正常性診断を行う。なお、ＳＳＤは、指定されたアドレスを含むメモリセルからデータを読込めない場合、メモリセルが故障していることを制御部に通知する。

【０００４】

ＳＳＤなどＮＡＮＤ型フラッシュメモリを用いた記憶装置は、経年劣化などにより、ブロック単位の故障が発生する。劣化による故障は、予め想定されている。そのため、ＳＳＤなどの媒体は、予備のブロックを有する。制御部は、故障が発生したブロックを故障ブロックとみなし、予備のブロックを故障ブロックと代替する。ＳＳＤは、故障ブロックの容量を予備のブロックで代替することで、媒体の初期容量を保つ。

【０００５】

予備のブロックを故障ブロックと代替するため、制御部は、故障ブロックに格納されていたデータを復旧し、その後、復旧したデータを予備ブロックに配置する。データの復旧方法は、ＥＣＣ（Ｅｒｒｏｒ－Ｃｏｒｒｅｃｔｉｎｇ　Ｃｏｄｅ）やＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ　Ａｒｒａｙ　ｏｆ　Ｉｎｅｘｐｅｎｓｉｖｅ　Ｄｉｓｃ）などである。このような、復旧可能なエラーは、リトライアブルエラーと呼ばれる。

【０００６】

ＳＳＤにおけるデータ修復のための技術として、ＮＡＮＤメモリブロックを１組のグループとし、ｎ－１個のデータを書き込む毎に、パリティデータを生成してｎ個目のデータとして、書き込みをする技術が知られている。（例えば、特許文献１参照）

【０００７】

パトロールリードに関連する技術として、外部からの要求に基づいて、アクセス対象のアドレスを診断し、シーケンシャルリードを実行させる技術が知られている。（例えば、特許文献２参照）

【先行技術文献】

【特許文献】

【０００８】

【特許文献１】特開２０１０－１５２５５１号公報

【特許文献２】特開２０１２－２４７８１７号公報

【発明の概要】

【発明が解決しようとする課題】

【０００９】

上述した背景技術では以下のような問題がある。

一例として、３つのメモリセルＡ、メモリセルＢ、メモリセルＣを有するブロックが故障した場合を想定する。なお、ブロックの故障は、リトライアブルエラーである。制御部が正常性診断を実施し、メモリセルＡの先頭アドレスを指定すると、ＳＳＤは、メモリセルＡに格納されたデータを読込めず、故障を検出する。ＳＳＤは、ブロックが故障してい

ることを、制御部に通知する。ＳＳＤは、故障ブロックのデータを復旧し、予備ブロックに復旧したデータを配置する処理を行う。一方、制御部は、正常性診断の処理を中断せずに、次のメモリセルＢの正常性診断を開始する。故障ブロックのデータ復旧及び予備ブロックへの復旧データの配置が完了していない場合、ＳＳＤは、メモリセルＢに格納されたデータを読込めず、ブロックが故障していることを検出し、制御部に通知する。その後、制御部は、正常性診断を引き続きメモリセルＣに対して行う。すると、ＳＳＤは、メモリセルＣに格納されたデータを読込めず、ブロックが故障していることを検出し、制御部に通知する。

【００１０】

制御部は、故障ブロックに対して正常性診断が行われる場合、１つ目のメモリセルを診断後、故障ブロックのデータ復旧及び予備ブロックへの復旧データの配置が完了するまで、複数回のエラーを通知される。制御部は、所定の回数のエラーを検出した場合に、ＳＳＤを問題のある記憶装置であると判断し、以降、該ＳＳＤを使用しないようにする。しかし、１つのブロックが故障した場合も、リトライアブルエラーである場合は、データが修復され、予備領域が代替で使用されるため、データの修復後、ＳＳＤの機能に何ら影響はない。上記のように、ＳＳＤの機能として、許容するはずのリトライアブルエラーであっても、制御部が、ＳＳＤを問題のある記憶装置と判断することがある。

【００１１】

１つの側面において、本発明の目的は、正常性診断を効率的に行うことである。

【課題を解決するための手段】

【００１２】

ストレージ制御装置は、インターフェースと、制御部とを備える。インターフェースは、複数のＲＡＩＤのグループを形成する複数のメモリブロックを備え、更に、複数のメモリブロックの予備として使用する予備ブロックを備える記憶装置との間でデータの入出力を行う。制御部は、記憶装置中の各メモリブロックが備える複数のメモリセルを読み出し対象とすると、故障したメモリブロックに含まれるメモリセルを読み出し対象に選択することがある。すると、制御部は、故障したメモリブロックのデータの修復及び予備ブロックへの修復したデータのコピーが完了するまで、故障したメモリブロックを有するＲＡＩＤのグループに属するメモリセルを、読み出し対象から外す。且つ、制御部は、故障したメモリブロックを有するＲＡＩＤのグループ以外のメモリセルを読み出し対象として選択する。制御部は、故障したメモリブロックのデータの修復及び予備ブロックへの修復したデータのコピーが完了すると、故障したメモリブロックを有するＲＡＩＤのグループに属するメモリセル及び予備ブロック中のメモリセルを、読み出し対象として選択する。

【発明の効果】

【００１３】

パトロールリードを効率的に行う。

【図面の簡単な説明】

【００１４】

【図１】第１の実施形態に係る正常性診断の処理の例を説明する図である。

【図２】本実施形態におけるストレージシステムの一例を示す。

【図３】第１の実施形態に係る正常性診断の処理の例を説明するシーケンス図である。

【図４Ａ】第２の実施形態に係る正常性診断の処理の例を説明するシーケンス図である。

【図４Ｂ】第２の実施形態に係る正常性診断の処理の例を説明するシーケンス図である。

【図５】中断した正常性診断を再開する処理の例を説明するシーケンス図である。

【図６Ａ】正常性診断に用いる情報テーブルの例を説明する図である。

【図６Ｂ】正常性診断に用いる情報テーブルの例を説明する図である。

【図７】グループ毎の先頭ブロックを正常性診断する処理の例を説明する図である。

【図８】ＳＳＤ及びＨＤＤを備えるストレージ装置の例を説明する図である。

【図９Ａ】ＳＳＤに用いる正常性診断処理の例を説明するフローチャートである。

【図９Ｂ】ＳＳＤに用いる正常性診断処理の例を説明するフローチャートである。

【図１０Ａ】予備ブロックへの修復データの配置が完了するまで実行される正常性診断の処理の例を説明するフローチャートである。
【図１０Ｂ】予備ブロックへの修復データの配置が完了するまで実行される正常性診断の処理の例を説明するフローチャートである。
【図１１】正常性診断をスキップしたグループの正常性診断を行う処理の例を説明するフローチャートである。
【発明を実施するための形態】
【００１５】
以下、本実施形態について、図面を参照しながら詳細に説明する。
図１は、第１の実施形態に係る正常性診断の処理の例を説明する図である。図１のストレージ装置１００は、制御部１０１とＳＳＤ１１０を有する。制御部１０１は、ＳＳＤの正常性診断の処理を制御する。ＳＳＤ１１０は、ＮＡＮＤフラッシュメモリ１１７を記憶媒体として用いる記憶装置であり、複数のブロック１１１（１１１ａ～１１１ｆ）及び予備ブロック１１１ｇを有する。ブロック１１１は、メモリ管理単位である複数のメモリセルを含む。ＳＳＤ１１０は、管理部１１５を有する。管理部１１５は、ＳＳＤ１１０内の各ブロックに格納するデータを管理する。
【００１６】
一例として、図１のＳＳＤ１１０のブロック１１１ａ、ブロック１１１ｂ、ブロック１１１ｃは、１つのグループである。グループは、例えば、ＲＡＩＤ５によるデータの保持が可能である。そのため、管理部１１５は、ＡＡという元データがある場合、ＡＡのデータをＡ１、Ａ２の２つに分割し、更に、ＡＡのデータに対するパリティを生成する。管理部１１５は、Ａ１、Ａ２、ＡＡのデータに対するパリティのデータをブロック１１１ａ、ブロック１１１ｂ、ブロック１１１ｃに格納する。同様に、管理部１１５は、ＢＢのデータをＢ１、Ｂ２の２つに分割し、更に、ＢＢのデータに対するパリティを生成する。管理部１１５は、Ｂ１、Ｂ２、ＢＢのデータに対するパリティのデータをブロック１１１ａ、ブロック１１１ｂ、ブロック１１１ｃに格納する。なお、ＡＡ及びＢＢのデータは、異なるデータである。
【００１７】
Ａ１、Ｂ１、パリティなどのデータは、各ブロック内のメモリセルに保持される。ブロック１１１ａのメモリセル２０１ａは、Ａ１のデータを保持し、メモリセル２０１ｂは、Ｂ１のデータを保持している。ブロック１１１ｂのメモリセル２０２ａは、Ａ２のデータを保持し、メモリセル２０２ｂは、Ｂ２のデータを保持している。ブロック１１１ｃのメモリセル２０３ａは、ＡＡのデータに対するパリティ、メモリセル２０３ｂは、ＢＢに対するパリティを保持している。次に、ブロック１１１ｄ、ブロック１１１ｅ、ブロック１１１ｆは、１つのグループである。ブロック１１１ｄ～ブロック１１１ｆは、ＭＭを分割したＭ１、Ｍ２と、ＭＭのデータに対するパリティを保持している。また、ブロック１１１ｄ～ブロック１１１ｆは、ＮＮを分割したＮ１、Ｎ２と、ＮＮのデータに対するパリティを保持している。そのため、ブロック１１１ａ～ブロック１１１ｃのグループと、ブロック１１１ｄ～ブロック１１１ｆのグループとは、別のデータを記憶している。
【００１８】
以下に、ブロック１１１ｂに故障がある場合の、第１の実施形態に係る正常性診断の処理の例を順に説明する。なお、正常性診断は、メモリセル単位で、順番に処理を行う。
【００１９】
（Ａ１）制御部１０１は、ＳＳＤ１１０の正常性診断を開始する。
【００２０】
（Ａ２）制御部１０１は、ブロック１１１ａのメモリセル２０１aの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する。
【００２１】
（Ａ３）管理部１１５は、メモリセル２０１ａのデータをキャッシュ領域１１６に読

み出す。管理部１１５は、更に、ＲＡＩＤの整合性を保つため、Ａ１の元データＡＡから分割されたＡ２及びパリティのデータをＳＳＤ１１０のキャッシュ領域１１６に読み出す。しかし、ＳＳＤ１１０は、ブロック１１１ｂが故障しているため、ブロック１１１ｂからＡ２のデータを読み出せない。

【００２２】

（Ａ４）管理部１１５は、ブロック１１１ｂからデータを読み出せないことを検出する。

【００２３】

（Ａ５）管理部１１５は、検出した故障がリトライアブルエラーかを判定する。検出した故障がリトライアブルエラーかどうかの判定は、Ａ２のデータを含むブロック１１１ｂに記憶されているデータが、ＥＣＣ又はＲＡＩＤのパリティを用いて修復できるかで判定される。管理部１１５は、リトライアブルエラーが発生したことを制御部１０１に通知する。

【００２４】

（Ａ６）検出したエラーがリトライアブルエラーである場合、管理部１１５は、ブロック１１１ｂのデータを修復し、予備ブロック１１１ｇにデータを配置する。なお、予備ブロック１１１ｇに配置されるデータは、ブロック１１１ｂと同じアドレスを指定してデータを格納する。

【００２５】

（Ａ７）制御部１０１は、故障ブロック１１１ｂを含むグループの正常性診断をスキップする。制御部１０１は、故障ブロック１１１ｂを含まないグループについて、正常性診断を継続する。例えば、制御部１０１は、ブロック１１１ｄのメモリセル２０４ａ、メモリセル２０４ｂなどの順番で、順次メモリセルの先頭アドレスを指定して、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を通知する。管理部１１５は、指定されたアドレスを含むメモリセルのデータを、制御部１０１に送信する。制御部１０１は、指定したアドレスを含むメモリセルは正常であると判定する。

【００２６】

（Ａ８）制御部１０１は、スキップ処理をしたグループ以外の全ブロックの正常性診断が完了したかを判定する。制御部１０１は、スキップ処理をしたグループ以外の全ブロックの正常性診断が完了していると判定し、尚且つ、予備ブロック１１１ｇへのデータの配置が完了している場合、スキップ処理をしたグループ（ブロック１１１ａ～ブロック１１１ｃ）の正常性診断を行う。なお、管理部１１５は、ブロック１１１ｂが保持していたデータを読み出す際には、代替として予備ブロック１１１ｇからデータを読み出す。

【００２７】

（Ａ９）全てのブロック１１１の正常性診断が完了すると、制御部１０１は、正常性診断の処理を終了する。

【００２８】

なお、（Ａ５）の結果、リトライアブルエラーでない場合、制御部１０１は、（Ａ６）（Ａ８）の処理を実行しない。また、制御部１０１は、（Ａ７）の正常性診断で、故障ブロックを検出した場合、（Ａ５）から処理を繰り返す。図１のＳＳＤは、ブロックの数やメモリセルの数を限定するものではない。また、Ａ１～Ｂ２などのデータも、一例であり、何らデータを限定するものではない。第１の実施形態では、ブロック１１１ａ、ブロック１１１ｂ、ブロック１１１ｃのように、３つのブロックを１つのグループとしているものの、更に複数のブロックを１つのグループとしてもよい。

【００２９】

故障を検出したブロックを含むグループに対する正常性診断の処理をスキップすることにより、ある故障ブロックを含むグループに対する正常性診断から検出されるエラーは、１回である。これにより、故障のあるブロックを診断する場合であっても、正常性診断は、故障しているブロックに対して複数回の診断をしないため、１つのブロックについて複数回のエラー検出をしない。

【００３０】
図２は、本実施形態におけるストレージシステムの一例を示す。ストレージシステム１１は、複数のディスク装置１９を搭載したディスク記憶システムである。ホストコンピュータ（以下、ホストと称する）２１からディスク装置１９へのアクセスに関して、冗長化のために、各ディスク装置１９に対して、２本のアクセスパスが存在している。またデータ自体についても、ＲＡＩＤを用いて複数台のディスクにデータが分散されて、冗長化した状態で保存されている。
【００３１】
ストレージシステム１１は、ホストコンピュータ（以下、ホストと称する）２１、ファイバチャネル（ＦＣ）スイッチ２２、コントローラエンクロージャ（ＣＥ）１２、ドライブエンクロージャ（ＤＥ）１８を含む。ストレージシステム１１は、ドライブエンクロージャ（ＤＥ）１８と、ホスト２１とが、ＦＣスイッチ２２、コントローラモジュール（ＣＭ）１３を介して接続されている。
【００３２】
ホスト２１は、ＦＣスイッチ２２を介してコントローラエンクロージャ（ＣＥ）１２と通信をして、ドライブエンクロージャ（ＤＥ）１８が有するディスク装置１９からデータを読み出したり、ディスク装置１９にデータを書き込んだりする。なお、図２では、１つのホスト２１を図示しているが、複数のホスト２１が、コントローラエンクロージャ（ＣＥ）１２に接続されていてもよい。
【００３３】
コントローラエンクロージャ（ＣＥ）１２は、複数のコントローラモジュール（ＣＭ）１３を含む。コントローラモジュール１３は、ディスク装置１９の動作制御を行う。各コントローラモジュール１３は、ストレージ制御装置として機能する。ストレージシステム１１は、運用に使用するコントローラモジュール（ＣＭ）１３を２以上有することで、冗長性を確保する。
【００３４】
コントローラモジュール（ＣＭ）１３は、アクセス指示情報として入出力（Ｉ／Ｏ）コマンドをドライブエンクロージャ（ＤＥ）１８に送信し、ディスク装置１９の記憶領域に対するデータの入出力指令を行う。また、入出力指令からアクセス監視時間が経過しても応答が得られないときは、コントローラモジュール（ＣＭ）１３は、このＩ／Ｏ処理を中断するアボート指示コマンドをドライブエンクロージャ（ＤＥ）１３に送信する。
【００３５】
コントローラモジュール１３は、チャネルアダプタ（ＣＡ）１４、中央演算装置（ＣＰＵ：Ｃｅｎｔｒａｌ　Ｐｒｏｃｅｓｓｉｎｇ　Ｕｎｉｔ）１５、記憶部１６、デバイスアダプタ（ＤＡ）１７を含む。チャネルアダプタ（ＣＡ）１４、ＣＰＵ１５、記憶部１６、デバイスアダプタ（ＤＡ）１７は、内部バスを介して接続されている。なお、ＣＰＵ１５は、制御部１０１として動作する。
【００３６】
チャネルアダプタ（ＣＡ）１４は、ＦＣスイッチ２２を介して、ホスト２１と接続されている。ＣＰＵ１５は、コントローラモジュール１３全体を制御する。
【００３７】
記憶部１６は、例えば、キャッシュメモリ、ＲＯＭ（Ｒｅａｄ　Ｏｎｌｙ　Ｍｅｍｏｒｙ）、ＲＡＭ（Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）等の情報を記録するデバイスである。記憶部１６には、コントローラモジュール１３を動作させるために用いるデータ、及び本実施形態に係るプログラム及び本実施形態で用いるデータ等が格納される。
【００３８】
デバイスアダプタ（ＤＡ）１７は、ドライブエンクロージャ（ＤＥ）１８に接続される。ＣＰＵ１５は、デバイスアダプタ１７を介してドライブエンクロージャ（ＤＥ）１８との間でデータの送受信を行う。
【００３９】

ドライブエンクロージャ（ＤＥ）１８は、１以上のドライブエンクロージャ（ＤＥ）より形成される。ドライブエンクロージャ（ＤＥ）１８は、複数のディスク装置１９を含み、冗長性を考慮したＲＡＩＤ構成を有する。なお、ディスク装置１９は、ハードディスクドライブ（ＨＤＤ）に限らず、例えばＳＳＤ（Ｓｏｌｉｄ　Ｓｔａｔｅ　Ｄｒｉｖｅ）や、光磁気ディスク等であってもよい。また、本実施形態では、記憶装置の一例として、ディスク装置１９を用いたが、これに限定されず、ランダムアクセス及びシーケンシャルアクセスのいずれのアクセス方法も可能な記憶装置であればよい。

【００４０】

図３は、第１の実施形態に係る正常性診断の処理の例を説明するシーケンス図である。なお、図３のシーケンス図の説明に、図１と同じストレージ装置を用いる。制御部１０１は、ブロック１１１ａのメモリセル２０１aの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ１０１）。管理部１１５は、メモリセル２０１ａのデータをキャッシュ領域１１６に読み出す。併せて、管理部１１５は、ＲＡＩＤの整合性を保つため、Ａ１の元データＡＡから分割されたＡ２をキャッシュ領域１１６に読み出すそうとするが、ブロック１１１ｂが故障しているため、ブロック１１１ｂからＡ２のデータを読み出せない。管理部１１５は、リトライアブルエラーが発生したことを制御部１０１に通知する（ステップＳ１０２）。管理部１１５は、故障ブロックのデータ修復を行う（ステップＳ１０３）。管理部１１５は、予備ブロック１１１ｇに、修復したデータを配置する（ステップＳ１０４）。制御部１０１は、リトライアブルエラーが発生したことを、カウントする（ステップＳ１０５）。ここで、Ｓ１０３～Ｓ１０４と、Ｓ１０５は並列して行われる。

【００４１】

制御部１０１は、故障ブロック１１１ｂを含むグループの正常性診断をスキップし、スキップをしたグループ以外のメモリセルを正常性診断の対象とする。制御部１０１は、正常性診断を継続するため、ブロック１１１ｄのメモリセル２０４ａの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ１０６）。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ１０７）。制御部１０１は、ブロック１１１ｄのメモリセル２０４ｂの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ１０８）。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ１０９）。

【００４２】

Ｓ１０４の処理であるデータの修復、及び、予備ブロック１１１ｇへの修復されたデータの配置が、完了する（ステップＳ１１０）。制御部１０１は、スキップ処理をしたグループ以外の全ブロックの正常性診断を完了する（ステップＳ１１１）。ここで、Ｓ１１０とＳ１１１は逆であってもよい。次のＳ１１２の処理は、Ｓ１１０及びＳ１１１の完了後に実行される。

【００４３】

制御部１０１は、スキップ処理をしたグループの正常性診断を行う。制御部１０１は、ブロック１１１ａのメモリセル２０１ａの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ１１２）。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ１１３）。制御部１０１は、ステップＳ１１２及びＳ１１３を繰り返し、全てのブロック１１１の正常性診断が完了すると、正常性診断の処理を終了する。

【００４４】

＜正常性診断の変形例＞

第２の実施形態に係る正常性診断の処理は、リトライアブルエラーにより修復したデータを予備ブロックに配置完了後に、継続していた正常性確認の処理を一旦中断し、正常性

診断をスキップしたグループの正常性診断を行う。

【００４５】

以下に、ブロック１１１ｂに故障がある場合の、第２の実施形態に係る正常性診断の処理の例を順に説明する。第２の実施形態に係る正常性診断は、第１の実施形態に係る正常性診断の処理（Ａ１）～（Ａ６）の後に実行される処理である。第２の実施形態に係る正常性診断の処理の説明に、図１と同じストレージ装置を用いる。

【００４６】

（Ｂ１）制御部１０１は、故障ブロック１１１ｂを含むグループの正常性診断をスキップする。制御部１０１は、故障ブロック１１１ｂを含まないグループのブロック１１１ｄから、正常性診断を継続する。例えば、制御部１０１は、ブロック１１１ｄのメモリセル２０４ａ、メモリセル２０４ｂなどの順番で、順次メモリセルの先頭アドレスを指定して、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を通知する。管理部１１５は、指定されたアドレスを含むメモリセルのデータを、制御部１０１に送信する。制御部１０１は、指定したアドレスを含むメモリセルは正常であると判定する。更に、制御部１０１は、管理部１１５からデータを受信する毎に、予備ブロック１１１ｇへの修復データの配置が完了したかを確認する要求を管理部１１５に通知する。

【００４７】

（Ｂ２）制御部１０１は、予備ブロック１１１ｇへの修復データの配置が完了している場合、実行中の正常性診断処理を中断し、正常性処理をスキップしたグループの正常性診断を行う。

【００４８】

（Ｂ３）正常性診断をスキップしたグループを対象とした正常性診断が完了する。制御部１０１は、中断していた正常性診断を再開する。

【００４９】

（Ｂ４）制御部１０１は、全てのブロック１１１の正常性診断が完了したかを判定する。全てのブロック１１１の正常性診断が完了すると、制御部１０１は、正常性診断の処理を終了する。

【００５０】

図４は、第２の実施形態に係る正常性診断の処理の例を説明するシーケンス図である。図４のシーケンス図の説明に、図１と同じストレージ装置を用いる。図４Ａは、ＳＳＤに対して正常性診断を開始した場合の例である。制御部１０１は、ブロック１１１ａのメモリセル２０１aの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ２０１）。管理部１１５は、メモリセル２０１ａのデータをキャッシュ領域１１６に読み出す。併せて、管理部１１５は、ＲＡＩＤの整合性を保つため、Ａ１の元データＡＡから分割されたＡ２をキャッシュ領域１１６に読み出すそうとするが、ブロック１１１ｂが故障しているため、ブロック１１１ｂからＡ２のデータを読み出せない。管理部１１５は、リトライアブルエラーが発生したことを制御部１０１に通知する（ステップＳ２０２）。管理部１１５は、故障ブロックのデータ修復を行う（ステップＳ２０３）。管理部１１５は、予備ブロック１１１ｇに、修復したデータを配置する（ステップＳ２０４）。制御部１０１は、リトライアブルエラーが発生したことを、カウントする（ステップＳ２０５）。ここで、Ｓ２０３～Ｓ２０４と、Ｓ２０５は並列して行われる。

【００５１】

制御部１０１は、故障ブロックを含むグループの正常性診断をスキップする。制御部１０１は、スキップをしたグループ以外で、次の正常性診断の対象となるメモリセルのアドレスである着目アドレスに、ブロック１１１ｄのメモリセル２０４ａの先頭のアドレスを設定する。更に、制御部１０１は、Ｓ２０４の予備ブロック１１１ｇへのデータの配置が完了したかの確認要求を、管理部１１５へ送信する（ステップＳ２０６）。管理部１１５は、制御部１０１からの確認要求への応答として、予備ブロック１１１ｇへのデータの配置が完了していないことを制御部１０１へ通知する（ステップＳ２０７）。制御部１０１

は、ｗａｉｔ＿ｐａｔｒｏｌのカウントに１加算する（ステップＳ２０８）。
【００５２】
制御部１０１は、故障ブロック１１１ｂを含むグループの正常性診断をスキップし、スキップをしたグループ以外のメモリセルに対して正常性診断を行う。制御部１０１は、正常性診断を継続するため、ブロック１１１ｄのメモリセル２０４ａの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ２０９）。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ２１０）。制御部１０１は、Ｓ２０４の予備ブロック１１１ｇへのデータの配置が完了したかの確認要求を、管理部１１５へ送信する（ステップＳ２１１）。管理部１１５は、制御部１０１からの確認要求への応答として、予備ブロック１１１ｇへのデータの配置が完了していないことを制御部１０１へ通知する（ステップＳ２１２）。制御部１０１は、ｗａｉｔ＿ｐａｔｒｏｌのカウントに１加算する（ステップＳ２１３）。制御部１０１は、ブロック１１１ｄのメモリセル２０４ｂの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ２１４）。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ２１５）。制御部１０１は、Ｓ２０４の予備ブロック１１１ｇへのデータの配置が完了したかの確認要求を、管理部１１５へ送信する（ステップＳ２１６）。管理部１１５は、制御部１０１からの確認要求への応答として、予備ブロック１１１ｇへのデータの配置が完了していないことを制御部１０１へ通知する（ステップＳ２１７）。制御部１０１は、ｗａｉｔ＿ｐａｔｒｏｌのカウントに１加算する（ステップＳ２１８）。
【００５３】
Ｓ２１８後も、Ｓ２０４の予備ブロック１１１ｇへのデータの配置処理が完了するまで、制御部１０１及び管理部１１５は、Ｓ２０９～Ｓ２１０やＳ２１４～Ｓ２１５など正常性診断を継続していく。更に、管理部１１５から指定したアドレスを含むメモリセルのデータを受信する毎に、制御部１０１及び管理部１１５は、Ｓ２１１～Ｓ２１３やＳ２１６～Ｓ２１８などの、予備ブロック１１１ｇへのデータの配置が完了しているかの確認の通信をする。
【００５４】
図４Ｂは、予備ブロックへのデータ配置完了後の正常性診断処理の例を説明するシーケンス図である。制御部１０１は、メモリセルＸＸの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ３０１）。メモリセルＸＸは、予備ブロック１１１ｇへのデータの配置が完了したときに、Ｓ２１８後に継続されている正常性診断が診断しているメモリセルである。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ３０２）。制御部１０１は、Ｓ２０４の予備ブロック１１１ｇへのデータの配置が完了したかの確認要求を、管理部１１５へ送信する（ステップＳ３０３）。管理部１１５は、制御部１０１からの確認要求への応答として、予備ブロック１１１ｇへのデータの配置が完了していることを制御部１０１へ通知する（ステップＳ３０４）。制御部１０１は、ｗａｉｔ＿ｐａｔｒｏｌのカウントに１加算する（ステップＳ３０５）。
【００５５】
制御部１０１は、Ｓ２０９から行われてきた正常性診断を一旦中断する。制御部１０１は、スキップ処理をしたグループの正常性診断を行う。制御部１０１は、ブロック１１１ａのメモリセル２０１ａの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ３０６）。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ３０７）。制御部１０１は、ブロック１１１ａのメモリセル２０１ｂの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ３０８）。管理部１１５は、指定さ

れたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ３０９）。

【００５６】

Ｓ３０９の後も、制御部１０１及び管理部１１５は、着目アドレスに設定したアドレスに達するまで、Ｓ３０６～Ｓ３０８までのような正常性診断を継続する。

【００５７】

図５は、中断した正常性診断を再開する処理の例を説明するシーケンス図である。図５のシーケンス図は、図４Ｂの後に、正常性診断をスキップされたグループの正常性診断が全て完了した後に実行される処理を説明している。図５のシーケンス図の説明に、図１と同じストレージ装置を用いる。制御部１０１は、正常性診断をスキップしたグループの最後のメモリセルの先頭のアドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ４０１）。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ４０２）。制御部１０１は、正常性診断をスキップしたグループの最後のメモリセルからｗａｉｔ＿ｐａｔｒｏｌのカウント個数分後のメモリセルを、次の診断対象とする（ステップＳ４０３）。制御部１０１は、ｗａｉｔ＿ｐａｔｒｏｌのカウントの値をクリアする（ステップＳ４０４）。ここで、正常性診断をスキップしたグループの最後のメモリセルからｗａｉｔ＿ｐａｔｒｏｌのカウント個数分後のメモリセルは、メモリセルＸＸの次のメモリセルである。

【００５８】

制御部１０１は、メモリセルＸＸの次のメモリセルの先頭アドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ４０５）。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ４０６）。制御部１０１は、メモリセルＸＸから２個後のメモリセルの先頭アドレスを指定し、指定したアドレスを含むメモリセルが保持するデータの読み出し要求を管理部１１５に通知する（ステップＳ４０７）。管理部１１５は、指定されたアドレスを含むメモリセルからデータを読み出し、制御部１０１に送信する（ステップＳ４０８）。全てのブロック１１１の正常性診断が完了すると、制御部１０１は、正常性診断の処理を終了する。

【００５９】

図６は、正常性診断に用いる情報テーブルの例を説明する図である。図６Ａは、リトライアブルエラーが発生した際に用いる情報テーブルである。リトライアブルエラーが発生した際に用いる情報テーブルは、リトライアブルエラーが検出されたメモリセルの先頭アドレスの情報と、別のリトライアブルエラーが発生した際に用いる情報テーブルのアドレスとを対応づけて保持する。

【００６０】

制御部１０１は、１回目のリトライアブルエラーの通知を受けた場合、エラーテーブル（ＥＲＲ＿ＴＢＬ）３０１ａを生成する。エラーテーブル３０１ａは、リトライアブルエラーが検出されたメモリセルの先頭アドレスを保持する。制御部１０１は、２回目のリトライアブルエラーの通知を受けた場合、エラーテーブル３０１ｂを生成する。２回目のリトライアブルエラーを検出する場合は、１回目のリトライアブルエラーとは別のブロックに故障が発生している時である。制御部１０１は、エラーテーブル３０１ａに、エラーテーブル３０１ｂのアドレス情報を記憶させ、更に、エラーテーブル３０１ｂに２回目のリトライアブルエラーが検出されたメモリセルの先頭アドレスを保持させる。制御部１０１は、３回目のリトライアブルエラーの通知を受けた場合、エラーテーブル３０１ｃを生成する。３回目のリトライアブルエラーを検出する場合は、１回目及び２回目のリトライアブルエラーとは別のブロックに故障が発生している時である。制御部１０１は、エラーテーブル３０１ｂに、エラーテーブル３０１ｃのアドレス情報を記憶させ、更に、エラーテーブル３０１ｃに３回目のリトライアブルエラーが検出されたメモリセルの先頭アドレスを保持させる。更に複数回のリトライアブルエラーが検出された場合、制御部１０１は、

リトライアブルエラーの回数分のエラーテーブル３０１を生成する。なお、制御部１０１は、エラーテーブル３０１に保持されているアドレスに該当するメモリセルを含むブロックの正常性診断が完了すると、エラーテーブル３０１を削除する。

【００６１】

図６Ｂは、予備ブロックへのデータの配置が完了したかの確認に用いる情報テーブルである。予備ブロックへのデータの配置が完了したかの確認に用いる情報テーブルは、着目アドレスと、確認要求の回数とを対応付けて保持する。更に、予備ブロックへのデータの配置が完了したかの確認に用いる情報テーブルは、別の予備ブロックへのデータの配置が完了したかの確認に用いる情報テーブルのアドレス情報を保持する。

【００６２】

制御部１０１は、１回目のリトライアブルエラー後の１回目の確認要求時に、ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ａを生成する。ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ａは、Ｓ２０６で着目アドレスに設定されたメモリセルの先頭アドレスを保持する。制御部１０１は、１回目のリトライアブルエラー後、確認要求毎に、制御部１０１は、確認要求の回数を加算し、ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ａに保持させる。

【００６３】

次に、制御部１０１は、２回目のリトライアブルエラーに基づいた予備ブロックへのデータ移動に対する１回目の確認要求時に、ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ｂを生成する。制御部１０１は、ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ａに、ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ｂのアドレス情報を記憶させる。ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ｂは、２回目のリトライアブルエラーにより設定された着目アドレスの先頭アドレスを保持する。更に、制御部１０１は、２回目のリトライアブルエラーに基づいた予備ブロックへのデータ移動に対する確認要求毎に、制御部１０１は、確認要求の回数をｗａｉｔ＿ｓｅｑｒｄテーブル３０２ｂに保持させる。

【００６４】

次に、制御部１０１は、３回目のリトライアブルエラーに基づいた予備ブロックへのデータ移動に対する１回目の確認要求時に、ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ｃを生成する。制御部１０１は、ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ｂに、ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ｃのアドレス情報を記憶させる。ｗａｉｔ＿ｓｅｑｒｄテーブル３０２ｃは、３回目のリトライアブルエラーにより設定された着目アドレスの先頭アドレスを保持する。更に、制御部１０１は、３回目のリトライアブルエラーに基づいた予備ブロックへのデータ移動に対する確認要求毎に、制御部１０１は、確認要求の回数を加算し、をｗａｉｔ＿ｓｅｑｒｄテーブル３０２ｃに保持させる。更に複数回のリトライアブルエラーが検出された場合、制御部１０１は、リトライアブルエラーの回数分のｗａｉｔ＿ｓｅｑｒｄテーブル３０２を生成する。なお、制御部１０１は、Ｓ４０５のように、パトロールリードの再開でｗａｉｔ＿ｓｅｑｒｄテーブル３０２が使用されると、使用されたテーブルを削除する。

【００６５】

＜ＳＳＤに用いる正常性診断処理例＞

図７は、グループ毎の先頭ブロックを正常性診断する処理の例を説明する図である。図７で説明する正常性診断処理は、複数のブロックを有するグループ毎に、先頭のブロックを診断し、残りのブロックに対する診断を行わない。また、リトライアブルエラーを検出した場合、図７で説明する正常性診断処理は、第１及び第２の実施形態に示した正常性診断の例を導入する。

【００６６】

図７のＳＳＤ１１０は、複数のブロック１１１と、キャッシュ領域１１６を有する。ＳＳＤ１１０は、ブロック111_{a1}～ブロック111_{a32}を有するＲＡＩＤ５（３１＋１）のグループを備える。同様に、ブロック111_{b1}～ブロック111_{b32}、ブロック111_{c1}～ブロック111_{c32}、が、ＲＡＩＤ５（３１＋１）のグループである。ＲＡＩＤ５（３１＋１）では、１つのデータが３１等分されて、例えば、ブロック１１１

$_{a\,1}$～ブロック１１１$_{a\,3\,1}$に保持されている。図７のＳＳＤ１１０では、１つのデータは、メモリセル０～メモリセル３１に分割して保持されている。また、ＲＡＩＤ５（３１＋１）のグループは、パリティを保持する。

【００６７】

ＳＳＤ１１０のように、複数のＲＡＩＤ５（３１＋１）のグループを有する記憶媒体に対して正常性診断を行う場合、制御部１０１は、ＲＡＩＤグループ毎に、少なくとも１つのブロックを診断すればよい。一例として、ブロック１１１$_{a\,1}$のメモリセル０のデータが読み出された場合、管理部１１５は、ＲＡＩＤの整合性を確認するため、メモリセル０に格納されているデータの元データから分割されたデータを読み出す。管理部１１５は、メモリセル１～メモリセル３１及びパリティのデータをＳＳＤ１１０のキャッシュ領域１１６に読み出す。

【００６８】

ブロック１１１$_{a\,1}$に対する正常性診断が終了し、次に、ブロック１１１$_{b\,1}$の正常性診断を行う場合、制御部１０１は、メモリセル１の診断を行う。制御部１０１が、メモリセル１の先頭アドレスを指定した場合、メモリセル１のデータは、既にキャッシュ領域１１６に読み出されている。そのため、管理部１１５は、メモリセル１のデータの読み出し要求を受信しても、メモリセル１からデータの読み出しを行わず、キャッシュ領域１１６からデータを読み出す。以降、メモリセル２～メモリセル３１に対する制御部１０１からのアドレス指定がなされた場合、管理部１１５は、キャッシュ領域１１６から各データを読み出す。そのため、メモリセル１～メモリセル３１及びパリティを指定する制御部１０１からの読み出し要求は、記憶媒体であるブロックの正常性診断をしていない。

【００６９】

制御部１０１からメモリセル０の読み出し要求を受信した場合、管理部１１５は、メモリセル０をキャッシュ領域１１６に読み出し、更に、メモリセル１～メモリセル３１及びパリティのデータをキャッシュ領域１１６に読み出す。そのため、制御部１０１が、１つのメモリセルに対して読み出し要求を出したことにより、メモリセル０～メモリセル３１及びパリティデータを保持するブロック１１１に対して正常性診断が行われたこととなる。よって、複数のＲＡＩＤ５のグループを有する記憶媒体に対して正常性診断を行う場合、制御部１０１は、ＲＡＩＤ５のグループ毎に、少なくとも１つのブロックを診断すればよい。

【００７０】

なお、第１の形態及び第２の実施形態、及びグループ毎に少なくとも１つのブロックを診断する処理の例は、正常性診断に限られない。第１の形態及び第２の実施形態、及びグループ毎に少なくとも１つのブロックを診断する処理の例は、ＳＳＤ内の全ページやブロックをメモリセル毎に読込む動作が伴うものに対して適用できる。例えば、リビルド処理やコピーバックといった処理である。リビルド処理は、メモリセル毎に読込む動作を、ＲＡＩＤ１、ＲＡＩＤ１＋０などで故障ＳＳＤのミラーからデータを読み出す際に行う。コピーバック処理は、メモリセル毎に読込む動作を、ＲＡＩＤ内の故障ＳＳＤを交換した場合に、他の記憶媒体からデータをコピーする際に行う。

【００７１】

図８は、ＳＳＤ及びＨＤＤを備えるストレージ装置の例を説明する図である。ストレージ装置１００は、制御部１０１ａ及び制御部１０１ｂ、ＳＳＤ１１０（１１０ａ～１１０ｄ）、ＨＤＤ２２０（２２０ａ～２２０ｈ）を備える。

【００７２】

制御部１０１ａは、ＨＤＤ２２０ａ～ＨＤＤ２２０ｄとＳＳＤ１１０ａ～ＳＳＤ１１０ｂを対象とする処理を制御する。制御部１０１ｂは、ＨＤＤ２２０ｅ～ＨＤＤ２２０ｈとＳＳＤ１１０ｃ～ＳＳＤ１１０ｄを対象とする処理を制御する。制御部１０１は、例えば、ＲＡＩＤコントローラである。制御部１０１は、外部装置との通信により、正常性診断、リビルド、コピーバックなどの処理を制御する。制御部１０１は、正常性診断、リビルド、コピーバックなどの処理をする場合、処理の対象がＨＤＤ２２０かＳＳＤ１１０かを

判定する。ＨＤＤ２２０かＳＳＤ１１０かを判断するのは、例えば、ＳＣＳＩコマンドを使用することで、媒体のタイプを特定すればよい。制御部１０１は、処理の対象がＨＤＤ２２０の場合、正常性診断、リビルド、コピーバックなどの処理を実行する。制御部１０１は、処理の対象がＳＳＤ１１０の場合、複数のブロックを有するグループ毎に、先頭のブロックをメモリセル毎に読込む処理をする。制御部１０１は、処理の対象がＳＳＤ１１０であり、故障を検出した場合は、第１及び第２の実施形態の処理を行う。

【００７３】

図９は、ＳＳＤに用いる正常性診断処理の例を説明するフローチャートである。図９のフローチャートの説明に、図１と同じストレージ装置を用いる。制御部１０１は、グループ毎の先頭ブロックを対象に、メモリセル毎の正常性診断を実行する（ステップＳ５０１）。制御部１０１は、管理部１１５からの正常性診断の応答が、リトライアブルエラーの通知であるかを判定する（ステップＳ５０２）。制御部１０１は、故障ブロックを含むグループの正常性診断をスキップし、予備ブロックへの修復データの配置が完了するまで、故障ブロックを含まないグループの正常性診断を継続する（ステップＳ５０３、ステップＳ５０２でＹＥＳ）。制御部１０１は、予備ブロックへの修復データの配置が完了すると、実行中の正常性診断処理を中断し、正常性処理をスキップしたグループの正常性診断を行う（ステップＳ５０４）。制御部１０１は、Ｓ５０４において、リトライアブルエラーを検出した場合、処理をＳ５０３に戻す。

【００７４】

制御部１０１は、着目アドレスに故障したメモリセルの先頭アドレス情報が設定されているかを判定する（ステップＳ５０５、ステップＳ５０２でＮＯ）。制御部１０１は、Ｓ５０４で中断した正常性診断を再開する（ステップＳ５０６、ステップＳ５０５でＹＥＳ）。制御部１０１は、グループ毎の先頭ブロックを対象とした正常性診断が全部完了したかを判定する（ステップＳ５０７、ステップＳ５０５でＮＯ）。制御部１０１は、着目アドレスで指定されたアドレスを有するメモリセルの正常性診断をする（ステップＳ５０８、ステップＳ５０７でＮＯ）。制御部１０１は、管理部１１５からの正常性診断の応答が、リトライアブルエラーの通知であるかを判定する（ステップＳ５０９）。制御部１０１は、故障ブロックを含むグループの正常性診断をスキップし、予備ブロックへの修復データの配置が完了するまで、故障ブロックを含まないグループの正常性診断を継続する（ステップＳ５１０、ステップＳ５０９でＹＥＳ）。制御部１０１は、予備ブロックへの修復データの配置が完了すると、実行中の正常性診断処理を中断し、正常性処理をスキップしたグループの正常性診断を行う（ステップＳ５１１）。Ｓ５１１が終了すると、制御部１０１は、処理をＳ５０５に戻す。制御部１０１は、着目アドレスに、次のメモリセルの先頭アドレスを設定する（ステップＳ５１２、ステップＳ５０９でＮＯ）。制御部１０１は、次のメモリセルの先頭アドレスを、例えば、Ｌｉｎｅ＿ｓｈｉｆｔなどのコマンドにより求める。Ｓ５１２が終了すると制御部１０１は、処理をＳ５０５に戻す。制御部１０１は、正常性診断の処理を終了する（ステップＳ５１３、ステップＳ５０７でＹＥＳ）。

【００７５】

図１０は、予備ブロックへの修復データの配置が完了するまで実行される正常性診断の処理の例を説明するフローチャートである。図１０のフローチャートは、図９のＳ５０３の処理を詳細に説明したフローチャートである。制御部１０１は、管理部１１５からのリトライアブルエラーの通知回数をカウントする（ステップＳ６０１）。制御部１０１は、リトライアブルエラーの通知回数が、所定の閾値に達したかを判定する（ステップＳ６０２）。制御部１０１は、ＳＳＤが故障していると判定し、故障しているＳＳＤの使用を中止する（ステップＳ６０３、ステップＳ６０２でＹＥＳ）。

【００７６】

制御部１０１は、リトライアブルエラーを検出したメモリセルの先頭アドレスの情報をＥＲＲ＿ＴＢＬに格納する（ステップＳ６０４、ステップＳ６０２でＮＯ）。制御部１０１は、予備ブロックへの修復データの配置が完了しているかを判定する（ステップＳ６０５）。予備ブロックへの修復データの配置が完了は、例えば、ＳＳＤのログなどを参照す

る。制御部１０１は、次のグループの先頭ブロックの先頭アドレスを着目アドレスに設定する（ステップＳ６０６）。制御部１０１は、着目アドレスで指定されたアドレスを有するメモリセルの正常性診断をする（ステップＳ６０７）。制御部１０１は、管理部１１５からの正常性診断の応答が、リトライアブルエラーの通知であるかを判定する（ステップＳ６０８）。制御部１０１は、処理をＳ６０２から繰り返す（ステップＳ６０８でＹＥＳ）。制御部１０１は、予備ブロックへのデータの配置が完了したかの確認要求の回数に、１を加算する（ステップＳ６０９、ステップＳ６０８でＮＯ）。制御部１０１は、着目アドレスに、次のメモリセルの先頭アドレスを設定する（ステップＳ６１０）。制御部１０１は、着目アドレスで指定されたアドレスを有するメモリセルの正常性診断をする（ステップＳ６１１）。制御部１０１は、管理部１１５からの正常性診断の応答が、リトライアブルエラーの通知であるかを判定する（ステップＳ６１２）。制御部１０１は、Ｓ６０２から処理を繰り返す（ステップＳ６１２でＹＥＳ）。制御部１０１は、予備ブロックへの修復データの配置が完了しているかを判定する（ステップＳ６１３、ステップＳ６１２でＮＯ）。制御部１０１は、Ｓ６０９から処理を繰り返す（ステップＳ６１３でＮＯ）。制御部１０１は、着目アドレスに設定されている値をクリアする（ステップＳ６１４、ステップＳ６１３でＹＥＳ）。制御部１０１は、予備ブロックへの修復データの配置が完了するまで実行される正常性診断の処理を終了する。

【００７７】

図１１は、正常性診断をスキップしたグループの正常性診断を行う処理の例を説明するフローチャートである。制御部１０１は、ＥＲＲ＿ＴＢＬを用いて、リトライアブルエラーを検出したメモリセルから正常性診断を行う（ステップＳ７０１）。制御部１０１は、管理部１１５からの正常性診断の応答が、リトライアブルエラーの通知であるかを判定する（ステップＳ７０２）。制御部１０１は、故障ブロックを含むグループの正常性診断をスキップし、予備ブロックへの修復データの配置が完了するまで、故障ブロックを含まないグループの正常性診断を継続する（ステップＳ７０３、ステップＳ７０２でＹＥＳ）。制御部１０１は、次のメモリセルの先頭アドレスを含むメモリセルを次の診断対象とする（ステップＳ７０４、ステップＳ７０２でＮＯ）。制御部１０１は、正常性診断をスキップしたグループの診断が全て完了したかを判定する（ステップＳ７０５）。制御部１０１は、処理をＳ７０１から繰り返す（ステップＳ７０５でＮＯ）。制御部１０１は、Ｓ７０１で使用したＥＲＲ＿ＴＢＬが、別のＥＲＲ＿ＴＢＬのアドレスを保持していないことを判定する（ステップＳ７０６）。制御部１０１は、別のＥＲＲ＿ＴＢＬを選択し、次の正常性診断対象とする（ステップＳ７０７）。制御部１０１は、Ｓ７０７が終了すると、処理をＳ７０１から繰り返す。制御部１０１は、ＥＲＲ＿ＴＢＬを初期化する（ステップＳ７０８、ステップＳ７０６でＹＥＳ）。制御部１０１は、Ｓ７０８の処理が終了すると、処理をＳ５０５から実行する。

【００７８】

以上、説明したように、実施形態にかかる方法では、故障を検出したブロックを含むグループに対する正常性診断の処理をスキップすることにより、ある故障ブロックを含むグループに対する正常性診断から検出されるエラーは、１回である。これにより、故障のあるブロックを診断する場合であっても、正常性診断は、故障しているブロックに対して複数回の診断をしないため、１つのブロックについて複数回のエラー検出をしない。

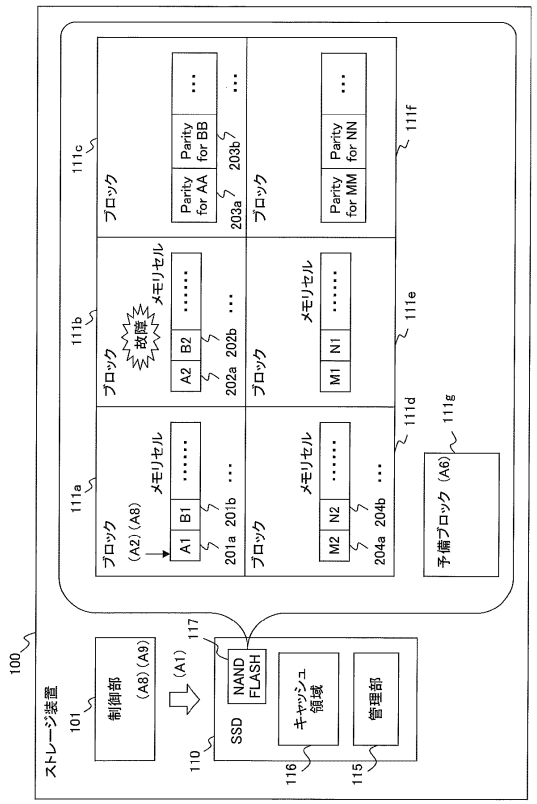
【符号の説明】

【００７９】

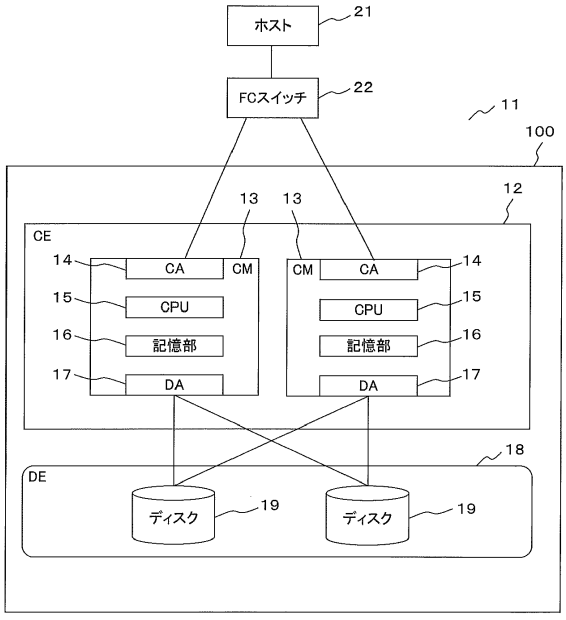
１００　ストレージ装置
１０１　制御部
１１０　ＳＳＤ
１１１（１１１ａ～１１１ｆ）　ブロック
１１１ｇ　予備ブロック
１１５　管理部
１１６　キャッシュ領域

２２０　ＨＤＤ
３０１　エラーテーブル
３０２　ｗａｉｔ＿ｓｅｑｒｄテーブル

【図１】
第1の実施形態に係る正常性診断の処理の例を説明する図

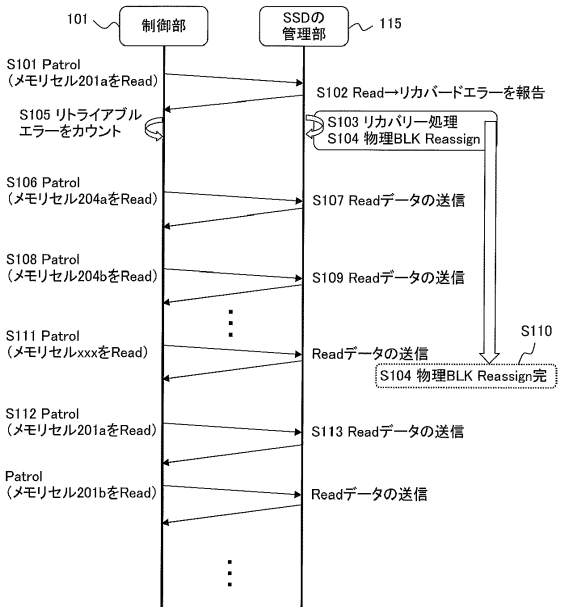


【図２】
本実施形態における
ストレージシステムの一例を示す図



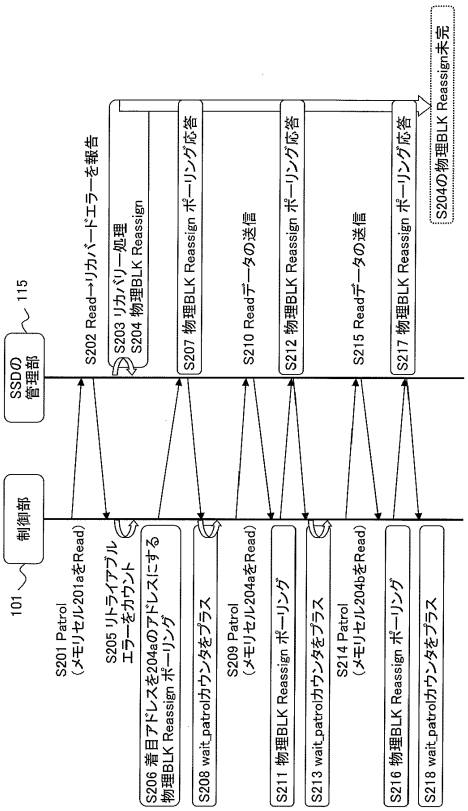
【図３】

第１の実施形態に係る正常性診断の処理の例を説明するシーケンス図



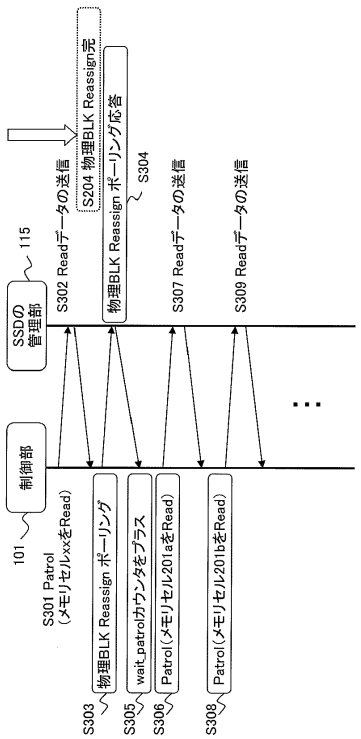
【図４Ａ】

第２の実施形態に係る正常性診断の処理の例を説明するシーケンス図



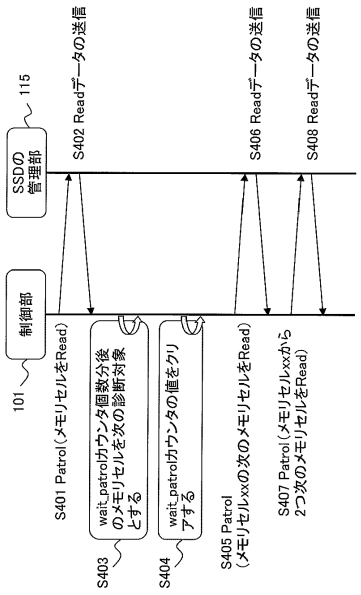
【図４Ｂ】

第２の実施形態に係る正常性診断の処理の例を説明するシーケンス図



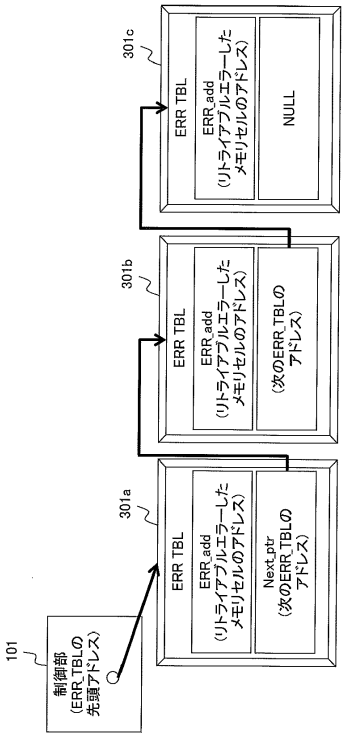
【図５】

中断した正常性診断を再開する処理の例を説明するシーケンス図



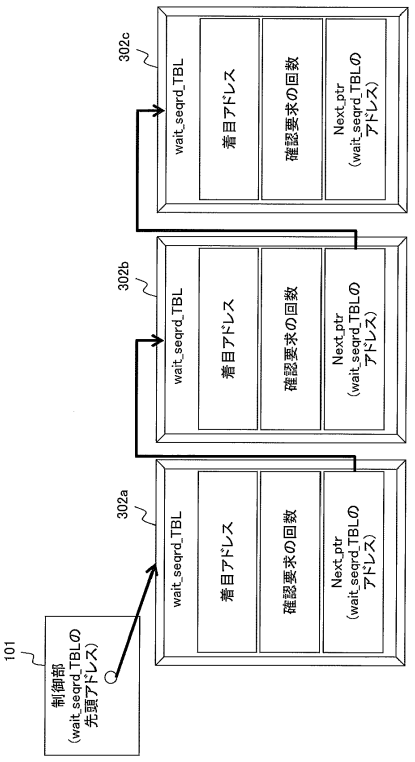
【図６Ａ】

正常性診断に用いる情報テーブルの例を説明する図



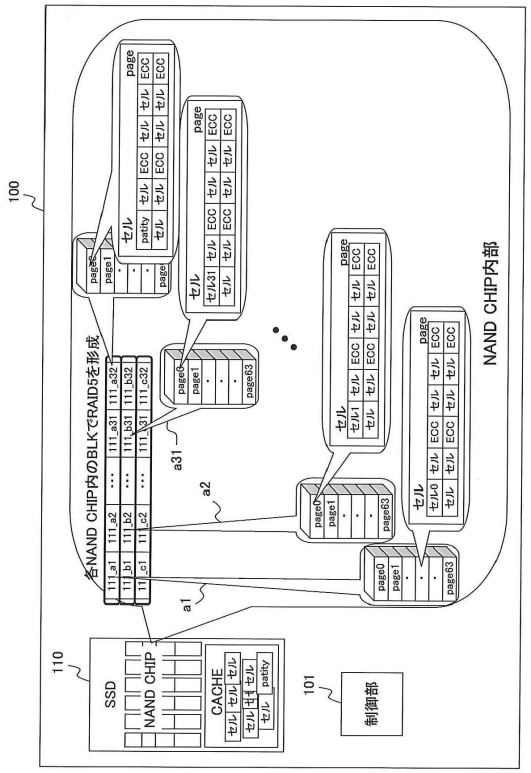
【図６Ｂ】

正常性診断に用いる情報テーブルの例を説明する図



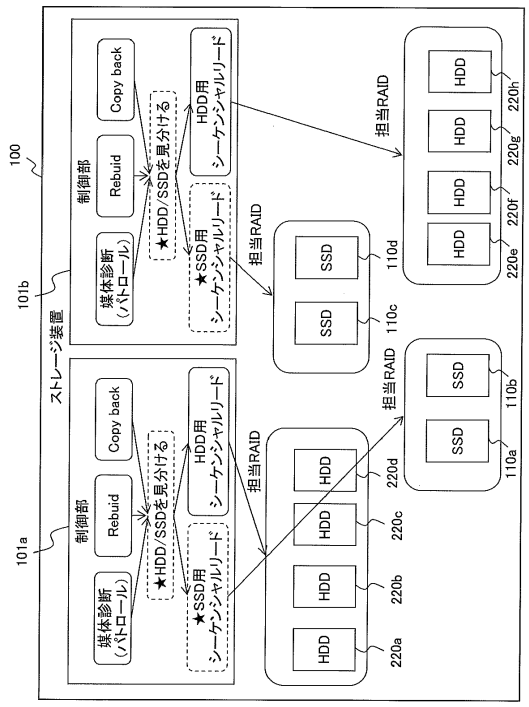
【図７】

グループ毎の先頭ブロックを
正常性診断する処理の例を説明する図



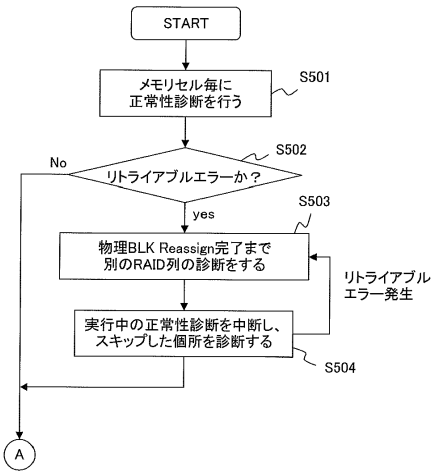
【図８】

SSD及びHDDを備えるストレージ装置の例を説明する図



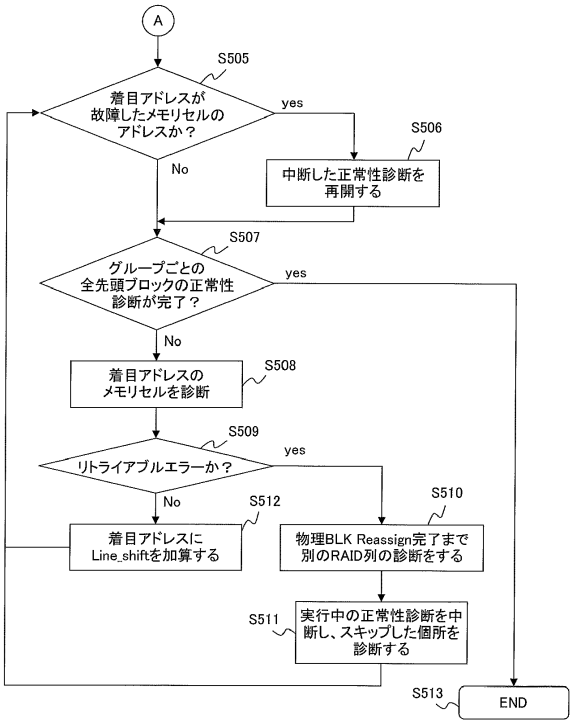
【図９Ａ】

SSDに用いる正常性診断処理の例を説明するフローチャート



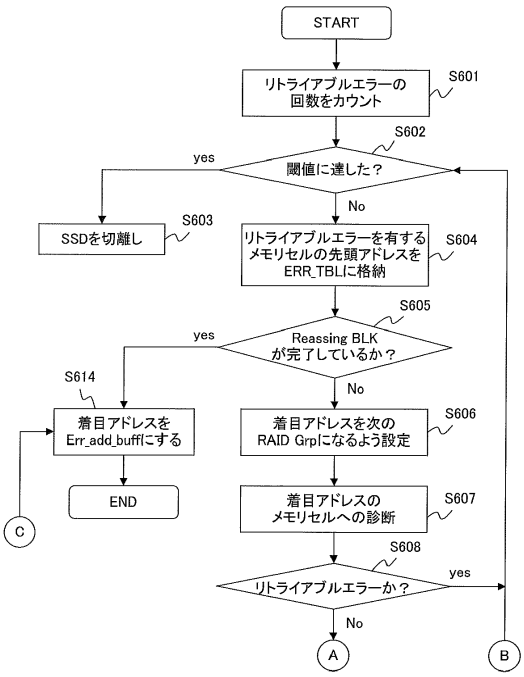
【図９Ｂ】

SSDに用いる正常性診断処理の例を説明するフローチャート



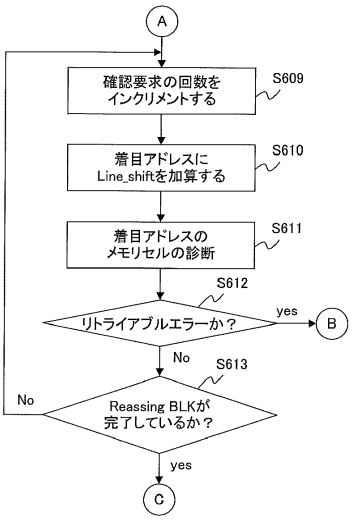
【図１０Ａ】

予備ブロックへの修復データの配置が完了するまで実行される
正常性診断の処理の例を説明するフローチャート



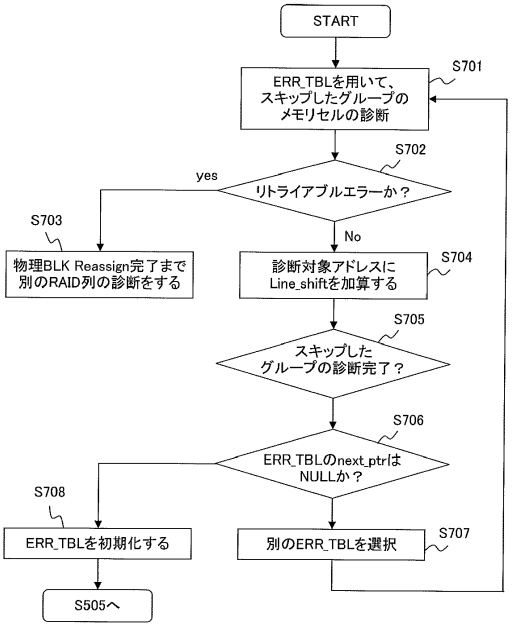
【図１０Ｂ】

予備ブロックへの修復データの配置が完了するまで実行される
正常性診断の処理の例を説明するフローチャート



【図１１】

正常性診断をスキップしたグループの
正常性診断を行う処理の例を説明するフローチャート



フロントページの続き

(51)Int.Cl.　　F I
　　Ｇ０６Ｆ　12/16　　３３０Ｃ
　　Ｇ０６Ｆ　12/16　　３２０Ｋ

(72)発明者　鈴木　利彦
　　神奈川県川崎市中原区上小田中４丁目１番１号　株式会社富士通コンピュータテクノロジーズ内

審査官　田上　隆一

(56)参考文献　特開２００３－３０３０５７（ＪＰ，Ａ）
　　特開２００８－２６２６１４（ＪＰ，Ａ）

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