(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5124792号
(P5124792)
(45)発行日　平成25年1月23日(2013.1.23)　(24)登録日　平成24年11月9日(2012.11.9)

(51)Int.Cl.　F I
G06F 3/06 (2006.01)　G06F 3/06 302Z
G06F 3/06 540

請求項の数 22　(全 18 頁)

(21)出願番号	特願2009-521760 (P2009-521760)	(73)特許権者	502188642
(86)(22)出願日	平成19年7月17日 (2007.7.17)		マーベル　ワールド　トレード　リミテッド
(65)公表番号	特表2009-545062 (P2009-545062A)		バルバドス国　ビービー１４０２７，セントマイケル、ブリトンズ　ヒル、ガンサイトロード、エル　ホライズン
(43)公表日	平成21年12月17日 (2009.12.17)		
(86)国際出願番号	PCT/US2007/016159		
(87)国際公開番号	WO2008/013695		
(87)国際公開日	平成20年1月31日 (2008.1.31)	(74)代理人	100104156
審査請求日	平成22年7月9日 (2010.7.9)		弁理士　龍華　明裕
(31)優先権主張番号	60/820,180	(74)代理人	100118005
(32)優先日	平成18年7月24日 (2006.7.24)		弁理士　飯山　和俊
(33)優先権主張国	米国(US)	(74)代理人	100143502
(31)優先権主張番号	11/591,325		弁理士　明石　英也
(32)優先日	平成18年11月1日 (2006.11.1)	(74)代理人	100138128
(33)優先権主張国	米国(US)		弁理士　東山　忠義

最終頁に続く

(54)【発明の名称】ＲＡＩＤ（ＲｅｄｕｎｄａｎｔＡｒｒａｙｏｆＩｎｄｅｐｅｎｄｅｎｔＤｉｓｋｓ）システム用のファイルサーバ

(57)【特許請求の範囲】
【請求項１】
ネットワークインターフェースと、
Ｎ個のストレージアレイと、
データ処理モジュールと
前記データ処理モジュールと前記Ｎ個のストレージアレイとの間、および、前記Ｎ個のストレージアレイの間でデータを切り替えるスイッチモジュールと
を備え、
前記Ｎ個のストレージアレイのそれぞれは、
ターゲット処理モジュールと、
１個からＭ個のハードディスクドライブと
を有し、
ＭおよびＮは１よりも大きい整数であり、
前記データ処理モジュールは、第１のデータブロックに対する第１のデータ格納要求を、前記第１のデータブロックに対するエラー検出訂正（ＥＣＣ）データ処理のために、複数の前記ターゲット処理モジュールのうちの１つのターゲット処理モジュールに割り当て、
前記複数のターゲット処理モジュールのうちの前記１つのターゲット処理モジュールは、前記第１のデータブロックの一部を格納し、前記第１のデータブロックの第１の部分および前記第１のデータブロックに対応付けられているＥＣＣデータを、前記複数のターゲ

ット処理モジュールのうちの別の１つのターゲット処理モジュールに送信し、
前記別の１つのターゲット処理モジュールは、前記第１のデータブロックの前記第１の部分及び前記ＥＣＣデータを受信および格納し、
前記データ処理モジュールおよび前記スイッチモジュールは前記Ｎ個のストレージアレイと前記ネットワークインターフェースとの間に接続され、
前記Ｎ個のストレージアレイは前記データ処理モジュールおよび前記スイッチモジュールを介して前記ネットワークインターフェースと通信するＲＡＩＤシステム。
【請求項２】
前記複数のターゲット処理モジュールのうち前記別の１つのターゲット処理モジュールが受信する前記ＥＣＣデータは、前記第１の部分に対応する請求項１に記載のＲＡＩＤシステム。
【請求項３】
前記ネットワークインターフェースは、前記第１のデータブロックを受信して、前記第１のデータブロックを前記データ処理モジュールに転送する請求項１または２に記載のＲＡＩＤシステム。
【請求項４】
前記ネットワークインターフェースは、ギガビットイーサネットネットワークインターフェース、およびデータバスのうち少なくとも１つを有する請求項３に記載のＲＡＩＤシステム。
【請求項５】
前記スイッチモジュールは、マルチポート高速スイッチを有する
請求項１から４の何れか１項に記載のＲＡＩＤシステム。
【請求項６】
前記データ処理モジュールは、第２のデータブロックに対する第２のデータ格納要求を、前記第２のデータブロックに対するＥＣＣデータ処理のために、第２のターゲット処理モジュールに割り当て、前記第２のターゲット処理モジュールは、前記第２のデータブロックのうち第１の部分および前記第２のデータブロックに対応付けられているＥＣＣデータを前記複数のターゲット処理モジュールのうち第３のターゲット処理モジュールに送信する請求項１から５の何れか１項に記載のＲＡＩＤシステム。
【請求項７】
前記第１および第２のデータブロックはそれぞれ、前記複数のターゲット処理モジュールのうち前記１つのターゲット処理モジュールおよび前記第２のターゲット処理モジュールにおいて並列に処理される請求項６に記載のＲＡＩＤシステム。
【請求項８】
前記データ処理モジュールは、前記第１のデータブロックに対してファイルシステム（ＦＳ）プロトコル機能を実行し、
前記ＦＳプロトコルは、ネットワークファイルサーバ（ＮＦＳ）および共通インターネットファイルサーバ（ＣＩＦＳ）のうち少なくとも１つを含む請求項１から７の何れか１項に記載のＲＡＩＤシステム。
【請求項９】
前記データ処理モジュールは、前記第１のデータブロックに適用されるＲＡＩＤストレージレベルを決定する請求項１から８の何れか１項に記載のＲＡＩＤシステム。
【請求項１０】
前記データ処理モジュールは、前記Ｎ個のストレージアレイのうち選択される複数のストレージアレイに前記第１のデータブロックをマッピングして、前記Ｎ個のストレージアレイに対するストレージマップを更新する請求項１から９の何れか１項に記載のＲＡＩＤシステム。
【請求項１１】
前記データ処理モジュールは、データ取得要求を受信すると、前記データ取得要求を前記複数のターゲット処理モジュールのうち第１のターゲット処理モジュールに割り当て、

前記複数のターゲット処理モジュールのうち前記第１のターゲット処理モジュールは、前記複数のターゲット処理モジュールのうちの他の複数のターゲット処理モジュールから前記データ取得要求に対応するデータを取得すると共に、前記データのうちエラーを含む部分に関連するＥＣＣデータを取得し、

前記複数のターゲット処理モジュールのうち前記第１のターゲット処理モジュールは、前記ＥＣＣデータを用いて前記部分に対してデータ復元を実行する請求項１から１０の何れか１項に記載のＲＡＩＤシステム。

【請求項１２】

前記データ処理モジュールは、データ取得要求を受信すると、前記データ取得要求に対応するデータを含む前記複数のターゲット処理モジュールにデータ取得メッセージを送信し、

前記複数のターゲット処理モジュールは、前記データ取得要求に対応する前記データおよび前記データのうちエラーを含む部分に関連するＥＣＣデータを取得し、

前記複数のターゲット処理モジュールは、前記データ取得要求に対応する取得された前記データおよび前記ＥＣＣデータを、前記データ処理モジュールに送信し、前記データ処理モジュールは前記ＥＣＣデータを用いて前記部分に対してデータ復元を実行する請求項１から１１の何れか１項に記載のＲＡＩＤシステム。

【請求項１３】

前記データ処理モジュールは、データ取得要求を受信すると、前記データ取得要求に対応するデータを含む前記複数のターゲット処理モジュールにデータ取得メッセージを送信し、

前記複数のターゲット処理モジュールは、前記データ取得要求に対応する前記データおよび前記データのうちエラーを含む部分に関連するＥＣＣデータを取得し、

前記複数のターゲット処理モジュールは、前記ＥＣＣデータを用いて前記部分に対してデータ復元を実行する請求項１から１２の何れか１項に記載のＲＡＩＤシステム。

【請求項１４】

前記データ処理モジュールは、前記スイッチモジュールと前記ネットワークインターフェースとの間に接続される請求項１から１３の何れか１項に記載のＲＡＩＤシステム。

【請求項１５】

前記ターゲット処理モジュールとは別物である前記データ処理モジュールは、前記ＲＡＩＤシステムとは別物のホストデバイスから前記第１のデータブロックを受信する請求項１から１４の何れか１項に記載のＲＡＩＤシステム。

【請求項１６】

前記データ処理モジュールは、オペレーティングシステム、ファイルシステムプロトコルファンクションのうち少なくとも１つを用いて前記第１のデータブロックを処理し、選択された前記ターゲット処理モジュールの１つにＲＡＩＤ冗長処理および復元処理を割り当てる請求項１から１５の何れか１項に記載のＲＡＩＤシステム。

【請求項１７】

前記データ処理モジュールは、前記Ｎ個のストレージアレイに格納されたデータブロックのマップを管理し、前記ターゲット処理モジュールに前記マップを割り当てる請求項１から１６の何れか１項に記載のＲＡＩＤシステム。

【請求項１８】

前記データ処理モジュールは、前記ターゲット処理モジュールのうちの１つを選択し、

選択された前記ターゲット処理モジュールのうちの１つにＲＡＩＤ冗長処理および復元処理を割り当てる請求項１から１７の何れか１項に記載のＲＡＩＤシステム。

【請求項１９】

前記データ処理モジュールは、エラー訂正コード処理を実行し、前記Ｎ個のストレージアレイへの格納後、前記第１のデータブロックがリカバーされる請求項１から１８の何れか１項に記載のＲＡＩＤシステム。

【請求項２０】

前記ＲＡＩＤストレージのレベルは、前記ＲＡＩＤシステムが分散化、ディスクミラーリング、およびパリティーストレージを実行するか否かを識別する請求項９に記載のＲＡＩＤシステム。

【請求項２１】

前記データ処理モジュールは、要求されたデータに対するデータ要求信号を生成する、請求項１に記載のＲＡＩＤシステム。

【請求項２２】

前記１つのターゲット処理モジュールが、前記第１のデータブロックに対してＥＣＣデータを生成する間、前記データ処理モジュールは、第２のデータブロックを受信し、前記第２のデータブロックを他の前記ターゲット処理モジュールに割り当て、

前記他のターゲット処理モジュールは、前記第２のデータブロックに対してＥＣＣデータを生成する

請求項１から２１のいずれか１項に記載のＲＡＩＤシステム。

【発明の詳細な説明】

【関連出願】

【０００１】

本願は、米国仮特許出願第６０／８２０，１８０号（出願日：２００６年７月２４日）、米国特許出願第１１／７２４，５４９号（出願日：２００７年３月１５日）、および米国特許出願第１１／５９１，３２５号（出願日：２００６年１１月１日）による恩恵を主張する。上記出願の開示内容はすべて、参照により本願に組み込まれる。

【技術分野】

【０００２】

本開示は、ＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ　Ａｒｒａｙ　ｏｆ　Ｉｎｄｅｐｅｎｄｅｎｔ　Ｄｉｓｋｓ）システムに関する。

【背景技術】

【０００３】

本明細書における「背景技術」の説明は、本開示内容がどのような状況の中で考案されているのかを概して説明するためになされる。現時点において名前が挙げられている発明者による研究は、この「背景技術」部分で説明されている範囲内において、説明されていなければ出願時において先行技術としての基準を満たすものでない本明細書の側面と同様に、本開示内容に対する先行技術として、明示または黙示を問わず、認められない。

【０００４】

ＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ　Ａｒｒａｙ　ｏｆ　Ｉｎｄｅｐｅｎｄｅｎｔ　Ｄｉｓｋｓ）システムは、複数のハードディスクにデータを冗長に格納する。一部のＲＡＩＤレベルでは、データブロックを分割して複数の異なるディスクに格納して、データ格納量および検索レイテンシを低減している。また、複数のディスクを用いる場合、平均故障間隔（ＭＴＢＦ）が大きくなると共にフォールトトレランスが高くなる傾向がある。

【０００５】

ＲＡＩＤシステムは、アクセスしているデバイスまたはホストデバイスにとっては、単一の論理ハードディスクドライブに見える。ＲＡＩＤシステムは、各ドライブの格納スペースを複数のユニットに分割することを含むディスクストライピングを採用するとしてもよい。ユニットサイズは、アプリケーションに応じて、セクタ（５１２バイト）から数メガバイトまでの間で、選択される。すべてのディスクのストライプは通常、インターリーブされて順にアドレスが割り当てられる。

【０００６】

ＲＡＩＤシステムには、冗長性を持たないアレイ（ＲＡＩＤ－０）に加えて、多くの種類が存在する。ＲＡＩＤ－０では、データに冗長性を持たせずにストライピングを採用したものである。性能は最も良好であるが、フォールトトレランスがない。ＲＡＩＤ－１では、ストライピングは用いられずにディスクミラーリングが利用され、データの格納時に複製を可能とするべく少なくとも２つのドライブが必要となる。各ディスクを同時に読み

出すことができるので、読出性能は向上する。書込性能は、単一のディスクストレージの場合と変わらない。ＲＡＩＤ－１は、マルチユーザシステムにおいて、性能およびフォールトトレランスが最も高い。

【０００７】

ＲＡＩＤ－２では、複数のディスクにわたってストライピングが利用される。一部のディスクは、エラー検出訂正（ＥＣＣ）情報を格納している。ＲＡＩＤ－３では、ストライピングが用いられて、１つのドライブがパリティ情報の格納専用とされる。埋め込まれたエラー検出訂正（ＥＣＣ）情報を用いてエラーを検出する。データ復元は、残りのドライブに格納されている情報の論理的排他和（ＸＯＲ）を算出することによって実行される。Ｉ／Ｏ動作ではすべてのドライブについて同時にアドレスを指定するので、ＲＡＩＤ－３はＩ／Ｏを重複して行うことができない。このため、ＲＡＩＤ－３は、記録内容が長いアプリケーションの単一ユーザシステムでの利用が最適である。

【０００８】

ＲＡＩＤ４では、大きいストライプが利用される。記録内容は、どの１つのドライブからも読み出すことができる。このため、読出動作においてはＩ／Ｏを重複させて実行することができる。書込動作ではパリティドライブを更新するので、Ｉ／Ｏを重複させることはできない。ＲＡＩＤ－５では、ＲＡＩＤ－４の書込動作に関する制約に対処するべく、回転式パリティアレイを用いている。このため、読出動作および書込動作を共に重複させることができる。ＲＡＩＤ－５では、パリティ情報を格納するが、冗長データは利用しない。しかし、パリティ情報を用いてデータを再構成することができる。ＲＡＩＤ－５は、アレイについて、少なくとも３つ、通常は５つのディスクを必要とする。ＲＡＩＤ－５は、性能が重要視されないマルチユーザシステムまたは書込動作がほとんど行われないマルチユーザシステムに最適である。

【０００９】

ＲＡＩＤ－６は、ＲＡＩＤ－５と同様であるが、複数の異なるドライブにわたって分散される第２のパリティが用いられる。ＲＡＩＤ－６は、フォールトトレランスおよびドライブ故障トレランスが高い。ＲＡＩＤ－７では、リアルタイムの埋め込みオペレーティングシステムおよびコントローラが用いられる。ＲＡＩＤ－７は、高速バスを介したキャッシングおよびスタンドアローンのコンピュータのほかの特性を利用する。

【００１０】

ＲＡＩＤ－１０は、ＲＡＩＤ－０とＲＡＩＤ－１とを組み合わせたものである。ＲＡＩＤ－１０はさらに２つに分類することができ、そのうち一方のＲＡＩＤ－０＋１では、データはストライプとして複数のディスクに渡って編成され、ストライプ状のディスク群をミラーリングさせる。もう一方のＲＡＩＤ－１＋０では、データをミラーリングさせて、ミラーリングの結果にストライピングを施す。

【００１１】

ＲＡＩＤ－５０（またはＲＡＩＤ－５＋０）では、一連のＲＡＩＤ－５群が用いられる。ＲＡＩＤ－５群を、ＲＡＩＤ－０方式に従ってストライピングして、データ保護レベルを下げることなくＲＡＩＤ－５の性能を改善する。ＲＡＩＤ－５３（またはＲＡＩＤ－５＋３）では、ストライピング（ＲＡＩＤ－０方式）をＲＡＩＤ－３の仮想ディスクブロックに用いる。この結果、ＲＡＩＤ－３に比べて性能も高くなるが、コストも同様に大きくなる。

【００１２】

ホストデバイスがストレージに対してデータブロックを送信すると、選択されているＲＡＩＤ方式についてＲＡＩＤ処理が実施される。このようなＲＡＩＤ処理は、選択されているＲＡＩＤレベルおよび／またはその他の処理に対する冗長処理および復元処理（例えば、エラー検出訂正（ＥＣＣ））を含むとしてもよい。

【００１３】

１つの方法を挙げると、単一の中央演算処理装置（ＣＰＵ）が別のデバイスからデータブロックを受信する。ＣＰＵはＥＣＣを含むすべてのＲＡＩＤ処理を実行する。当該方法

によれば、ＥＣＣ関連処理は一定でなく時間がかかり得るので、ＣＰＵによってデータストレージの速度が制限され得ることが多い。つまり、ＣＰＵでの処理によってボトルネックが生じると共にレイテンシが大きくなってしまう場合がある。ＣＰＵが１つの場合、１つのデータブロックのＲＡＩＤ再構成は、後続のデータブロックの処理の前に完了しなければならない。

【発明の概要】

【課題を解決するための手段】

【００１４】

ＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｄｅｐｅｎｄｅｎｔ Ｄｉｓｋｓ）システムは、Ｎ個のストレージアレイを備え、Ｎ個のストレージアレイのそれぞれは、ターゲット処理モジュールと、１個からＭ個のハードディスクドライブとを有し、ＭおよびＮは１よりも大きい整数である。データ処理モジュールは、第１のデータブロックに対する第１のデータ格納要求を、第１のデータブロックに対するエラー検出訂正（ＥＣＣ）データ処理のために、複数のターゲット処理モジュールのうちの１つのターゲット処理モジュールに割り当てる。複数のターゲット処理モジュールのうちの１つのターゲット処理モジュールは、第１のデータブロックの第１の部分および第１のデータブロックに対応付けられているＥＣＣデータを、複数のターゲット処理モジュールのうちの別の１つのターゲット処理モジュールに送信する。

【００１５】

別の特徴によると、複数のターゲット処理モジュールのうち別の１つのターゲット処理モジュールが受信するＥＣＣデータは、第１の部分に対応する。インターフェースは、第１のデータブロックを受信して、第１のデータブロックをデータ処理モジュールに転送する。インターフェースは、ネットワークインターフェース、ギガビットイーサネットネットワークインターフェース、およびデータバスのうち少なくとも１つを有する。スイッチモジュールは、データ処理モジュールとＮ個のストレージアレイとの間、および、Ｎ個のストレージアレイの間でデータを切り替える。スイッチモジュールは、マルチポート高速スイッチを有する。データ処理モジュールは、第２のデータブロックに対する第２のデータ格納要求を、第２のデータブロックに対するＥＣＣデータ処理のために、第２のターゲット処理モジュールに割り当て、第２のターゲット処理モジュールは、第２のデータブロックのうち第１の部分および第２のデータブロックに対応付けられているＥＣＣデータを複数のターゲット処理モジュールのうち第３のターゲット処理モジュールに送信する。第１および第２のデータブロックはそれぞれ、複数のターゲット処理モジュールのうち１つのターゲット処理モジュールおよび第２のターゲット処理モジュールにおいて、重複して処理される。

【００１６】

別の特徴によると、データ処理モジュールは、インターフェース、メモリ、および少なくとも１つのプロセッサを有する。データ処理モジュールは、第１のデータブロックに対してファイルシステム（ＦＳ）プロトコル機能を実行する。ＦＳプロトコルは、ネットワークファイルサーバ（ＮＦＳ）および共通インターネットファイルサーバ（ＣＩＦＳ）のうち少なくとも１つを含む。データ処理モジュールは、第１のデータブロックに適用されるＲＡＩＤストレージレベルを決定する。データ処理モジュールは、Ｎ個のストレージアレイのうち選択される複数のストレージアレイに第１のデータブロックをマッピングして、Ｎ個のストレージアレイに対するストレージマップを更新する。

【００１７】

別の特徴によると、データ処理モジュールは、データ取得要求を受信すると、データ取得要求を複数のターゲット処理モジュールのうち第１のターゲット処理モジュールに割り当てる。複数のターゲット処理モジュールのうち第１のターゲット処理モジュールは、複数のターゲット処理モジュールのほかの複数のターゲット処理モジュールからデータ取得要求に対応するデータを取得すると共に、データのうちエラーを含む部分に関連するＥＣＣデータを取得する。

【００１８】
別の特徴によると、複数のターゲット処理モジュールのうち第１のターゲット処理モジュールは、ＥＣＣデータを用いて部分に対してデータ復元を実行する。データ処理モジュールは、データ取得要求を受信すると、データ取得要求に対応するデータを含む複数のターゲット処理モジュールにデータ取得メッセージを送信する。複数のターゲット処理モジュールは、データ取得要求に対応するデータおよびデータのうちエラーを含む部分に関連するＥＣＣデータを取得する。複数のターゲット処理モジュールは、データ取得要求に対応する取得されたデータおよびＥＣＣデータを、データ処理モジュールに送信し、データ処理モジュールはＥＣＣデータを用いて部分に対してデータ復元を実行する。複数のターゲット処理モジュールは、ＥＣＣデータを用いて部分に対してデータ復元を実行する。

【００１９】
ＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｄｅｐｅｎｄｅｎｔ Ｄｉｓｋｓ）システムは、Ｎ個のストレージアレイを備え、Ｎ個のストレージアレイのうちそれぞれは、ターゲット処理モジュールと、１個からＭ個のハードディスクドライブとを有し、ＭおよびＮは１よりも大きい整数である。データ処理モジュールは、複数のデータブロックに対するエラー検出訂正（ＥＣＣ）処理を複数のターゲット処理モジュールのうち選択される複数のターゲット処理モジュールに、重複しないように、選択的に割り当てる。スイッチモジュールは、データ処理モジュールとＮ個のストレージアレイとの間、および、Ｎ個のストレージアレイのそれぞれとＮ個のストレージアレイのほかの複数のストレージアレイとの間において、通信経路を提供する。

【００２０】
別の特徴によると、データ処理モジュールは、第１のデータブロックに対するデータ格納要求を、第１のデータブロックに対するＥＣＣデータ処理のために、複数のターゲット処理モジュールのうち１つのターゲット処理モジュールに割り当て、複数のターゲット処理モジュールのうち１つのターゲット処理モジュールは、第１のデータブロックの第１の部分および第１のデータブロックに対応付けられているＥＣＣデータを、複数のターゲット処理モジュールのうち別の１つのターゲット処理モジュールに送信する。複数のターゲット処理モジュールのうち別の１つのターゲット処理モジュールが受信するＥＣＣデータは、第１の部分に対応する。インターフェースは、複数のデータブロックを受信して、複数のデータブロックをデータ処理モジュールに転送する。インターフェースは、ネットワークインターフェース、ギガビットイーサネットネットワークインターフェース、およびデータバスのうち少なくとも１つを有する。

【００２１】
別の特徴によると、スイッチモジュールは、マルチポート高速スイッチを有する。スイッチモジュールは、毎秒１ギガビット以上の速度で動作するマルチポートスイッチを有する。スイッチモジュールは、マルチポートギガビットイーサネットスイッチを有する。データ処理モジュールは、第２のデータブロックに対する第２のデータ格納要求を、第２のデータブロックに対するＥＣＣデータ処理のために、第２のターゲット処理モジュールに割り当てる。第２のターゲット処理モジュールは、第２のデータブロックのうち第１の部分および第２のデータブロックに対応付けられているＥＣＣデータを複数のターゲット処理モジュールのうち第３のターゲット処理モジュールに送信する。

【００２２】
別の特徴によると、第１および第２のデータブロックはそれぞれ、複数のターゲット処理モジュールのうち１つのターゲット処理モジュールおよび第２のターゲット処理モジュールにおいて、重複して処理される。データ処理モジュールは、インターフェース、メモリ、および少なくとも１つのプロセッサを有し、データ処理モジュールは、第１のデータブロックに対してファイルシステム（ＦＳ）プロトコル機能を実行する。ＦＳプロトコルは、ネットワークファイルサーバ（ＮＦＳ）および共通インターネットファイルサーバ（ＣＩＦＳ）のうち少なくとも１つを含む。データ処理モジュールは、複数のデータブロックに適用されるＲＡＩＤストレージレベルを決定する。データ処理モジュールは、Ｎ個の

ストレージアレイのうち選択される複数のストレージアレイに複数のデータブロックをマッピングして、Ｎ個のストレージアレイに対するストレージマップを更新する。

【００２３】

別の特徴によると、データ処理モジュールは、データ取得要求を受信すると、データ取得要求を複数のターゲット処理モジュールのうち第１のターゲット処理モジュールに割り当てる。複数のターゲット処理モジュールのうち第１のターゲット処理モジュールは、複数のターゲット処理モジュールのほかのターゲット処理モジュールに対してデータ取得要求に対応するデータを要求すると共にデータのうちエラーを含む部分に関連するＥＣＣデータを要求する。複数のターゲット処理モジュールのうち第１のターゲット処理モジュールは、ＥＣＣデータを用いて部分に対してデータ復元を実行する。

【００２４】

別の特徴によると、データ処理モジュールは、データ取得要求を受信すると、データ取得要求に対応するデータを含む複数のターゲット処理モジュールにデータ取得メッセージを送信する。複数のターゲット処理モジュールは、データ取得要求に対応するデータおよびデータのうちエラーを含む部分に関連するＥＣＣデータを取得する。複数のターゲット処理モジュールは、データ取得要求に対応する取得されたデータおよびＥＣＣデータを、データ処理モジュールに送信し、データ処理モジュールはＥＣＣデータを用いて部分に対してデータ復元を実行する。複数のターゲット処理モジュールは、ＥＣＣデータを用いて部分に対してデータ復元を実行する。

【００２５】

本開示内容はさらに異なる分野でも応用が可能であり、そのような分野は以下の詳細な説明から明らかとなる。本開示内容の好ましい実施形態を説明するためのものとして、詳細な記載および具体的な例を挙げるが、これらは説明を目的としたものに過ぎず本開示内容の範囲を限定するものではないと理解されたい。

【図面の簡単な説明】

【００２６】

以下の詳細な説明および添付図面を参照して本開示内容をより詳しく説明する。添付図面は以下の通りである。

【００２７】

【図１】本開示に係るＲＡＩＤシステムを示す機能ブロック図である。

【００２８】

【図２Ａ】データ処理モジュールを示す機能ブロック図である。

【００２９】

【図２Ｂ】ターゲット処理モジュールを示す機能ブロック図である。

【００３０】

【図２Ｃ】ターゲット処理モジュールの一例を示す、より詳細な機能ブロック図である。

【００３１】

【図３】データブロックの処理を示す機能ブロック図である。

【００３２】

【図４】データブロックの処理を示す機能ブロック図である。

【００３３】

【図５】データブロックの処理を示す機能ブロック図である。

【００３４】

【図６】データブロックの処理のタイミングを示す図である。

【００３５】

【図７】格納されるべきデータブロックの処理方法を示すフローチャートである。

【００３６】

【図８Ａ】データ処理モジュールによってＥＣＣ処理が実施されるデータブロックを取得する方法の一例を示すフローチャートである。

【図８Ｂ】データ処理モジュールによってＥＣＣ処理が実施されるデータブロックを取得

する方法の一例を示すフローチャートである。
【００３７】
【図９Ａ】ターゲット処理モジュールのそれぞれによってＥＣＣ処理が実施されるデータブロックを取得する方法の一例を示すフローチャートである。
【図９Ｂ】ターゲット処理モジュールのそれぞれによってＥＣＣ処理が実施されるデータブロックを取得する方法の一例を示すフローチャートである。
【００３８】
【図１０】ターゲット処理モジュールのうち選択された１つのターゲット処理モジュールによってＥＣＣ処理が実施されるデータブロックを取得する方法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】
【００３９】
以下の説明は、本質的に例示に過ぎず、本開示内容、その用途または利用を限定するものでは決してない。本開示内容を明確に説明するべく、類似の構成要素を図中で指定する際には複数の図面にわたって同一の参照番号を使用する。本明細書で使用する場合、モジュール、回路および／またはデバイスという用語は、１以上のソフトウェアまたはファームウェアプログラムを実行する特定用途集積回路（ＡＳＩＣ）、電子回路、プロセッサ（共有、専用または群）およびメモリ、組み合わせ論理回路、ならびに／または本明細書で記載する機能を提供する上記以外に適切な構成要素を指すものとする。本明細書で言及される場合、「Ａ、ＢおよびＣのうち少なくとも１つ」という表現は、論理演算（ＡまたはＢまたはＣ）、非排他的論理ＯＲを意味すると解釈されたい。尚、方法が含むステップは、本開示内容の原則を変更することなく、別の順序に従って実行され得る。
【００４０】
図１は、ＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｄｅｐｅｎｄｅｎｔ Ｄｉｓｋｓ）システム１００を示す。インターフェース１０４は、ＲＡＩＤシステム１００に格納されるべきデータブロックを受信する。例えば、インターフェース１０４は、ギガビットイーサネット（登録商標）ネットワークインターフェース、データバス等の高速インターフェースであってもよいが、任意のほかの種類のインターフェースを用いるとしてもよい。データ処理モジュール１０８は、ＲＡＩＤ処理の一部を実行する。つまり、データ処理モジュール１０８は、インターフェース１０４からデータブロックを受信して、当該データに対してオペレーティングシステム（ＯＳ）およびファイルシステム（ＦＳ）プロトコル機能を実行する。例えば、ＦＳプロトコルは、ネットワークファイルサーバ（ＮＦＳ）、共通インターネットファイルサーバ（ＣＩＦＳ）、および／または、その他の適切なプロトコルを含むとしてよい。データ処理モジュール１０８は、後述するように、他のターゲット処理デバイスに、冗長処理および復元処理（例えば、エラー検出訂正（ＥＣＣ））を分配する。
【００４１】
データ処理モジュール１０８は、スイッチモジュール１１２と通信する。一例に過ぎないが、スイッチモジュール１１２は、クロスバースイッチ、ギガビットスイッチまたはギガビットイーサネットスイッチといったマルチポート高速スイッチであってもよい。スイッチモジュール１１２は、データパケットとして編成されたデータを切り替えるとしてもよい。想到されることであるが、スイッチモジュール１１２によれば、ハードワイヤード接続に比べ、拡張性および柔軟性が実現される。
【００４２】
スイッチモジュール１１２は、２つ以上のストレージアレイ１２０－１、１２０－２、・・・および１２０－Ｘ（ストレージアレイ１２０と総称する）と通信する。ここで、Ｘは１よりも大きい整数である。ストレージアレイ１２０はそれぞれ、ターゲット処理モジュール１２２－１、１２２－２、・・・および１２２－Ｘ（ターゲット処理モジュール１２２と総称する）と、１以上のハードディスクドライブ（ＨＤＤ）１２４－１１、１２４－１２、・・・および１２４－ＸＹ（ＨＤＤ１２４と総称する）とを有する。ここで、Ｙ

は０よりも大きい整数である。想到されることであるが、ストレージアレイ１２０の数および各ストレージアレイ１２０が有するＨＤＤ１２４の数を変化させてスケーリングを可能とするとしてもよい。

【００４３】

図２Ａは、データ処理モジュール１０８の一例をさらに詳細に示す図である。データ処理モジュール１０８は、インターフェース１０４を介して、データ格納されるべきデータブロックを受信する。データ処理モジュール１０８は、インターフェース１５０、メモリ１５４、および１以上のプロセッサ１５６を有するとしてもよい。

【００４４】

データ処理モジュール１０８は、適用すべきＲＡＩＤストレージレベルを決定して、ＦＳ関連処理を実行して、ストレージアレイに対してデータブロックをマッピングして、ＲＡＩＤ冗長処理および復元処理（例えば、エラー検出訂正（ＥＣＣ））を、選択されるターゲット処理モジュールに割り当てて、ストレージマップを更新する等としてもよい。

【００４５】

ＲＡＩＤ冗長処理および復元処理を実行するように割り当てられているターゲット処理モジュール１２２は、データ処理モジュール１０８から命令を受信する。選択ターゲット処理モジュール１２２は、割り当てられたデータブロックについてエラー検出訂正（ＥＣＣ）を生成する。完了すると、当該ターゲット処理モジュール１２２は、データ処理モジュール１０８が与えたＲＡＩＤ命令に基づいて、他のアレイに格納させるべく、他のターゲット処理モジュールに、データブロックおよび／またはＥＣＣデータの一部分を選択的に送信することによって、データ拡散処理を実行する。一部のデータおよびＥＣＣデータは、ローカルに格納されるとしてもよい。

【００４６】

同時に、他のターゲット処理モジュール１２２には、他のデータブロックについてのＲＡＩＤ冗長処理および復元処理が割り当てられるとしてもよい。他のターゲット処理モジュール１２２は、重複して、他のデータブロックに対するＥＣＣを処理する。データ処理モジュール１０８は、格納されるべきデータブロックのいずれについてもＥＣＣを処理しないので、ボトルネックを生じさせることはない。データ処理モジュール１０８に対応付けられているメモリ１５４は、ストレージアレイ１２０のデータのグローバルドライブマップ１５８を格納および更新するとしてもよい。

【００４７】

図２Ｂを参照しつつ説明すると、各ターゲット処理モジュール１２２は、ＲＡＩＤ構成モジュール１７０およびＲＡＩＤ取得モジュール１６８を有するとしてもよい。ＲＡＩＤ構成モジュール１７０はＥＣＣを処理する。ＲＡＩＤ取得モジュール１６８は、後述するように、ＲＡＩＤ取得要求を処理する。

【００４８】

ＲＡＩＤ構成モジュール１７０は、選択ターゲット処理モジュール１２２に対応付けられているローカルドライブ１２４に格納されるべきデータブロックの一部分についてのＥＣＣを処理する。さらに、ＲＡＩＤ構成モジュール１７０は、リモートストレージアレイ１２０に対応付けられているリモートドライブについてのＥＣＣを処理する。ＲＡＩＤ命令モジュール１７２は、別のターゲット処理についてのＲＡＩＤ命令を生成し、他のターゲット処理モジュール１２２から受信するＲＡＩＤ命令を処理するとしてもよい。ＲＡＩＤ命令モジュール１７２は、ＲＡＩＤ構成モジュール１７０と一体化されるとしてもよい。

【００４９】

選択ターゲット処理モジュール以外のターゲット処理モジュールに対応付けられているリモートストレージアレイ１２０は、選択ターゲット処理モジュールから受信するデータおよび／またはＥＣＣデータを格納する。リモートストレージアレイ１２０は単純に、選択ターゲット処理モジュール１２２から送信されるＲＡＩＤ命令に従うとしてもよい。

【００５０】

想到されるように、リモートストレージアレイ１２０が実行する処理の量は、選択ターゲット処理モジュール１２２が実行するＲＡＩＤ構成処理よりもはるかに少ない。このため、リモートストレージアレイ１２０のターゲット処理モジュール１２２は、重複して他のデータブロックのＲＡＩＤ構成を処理するべく利用され得る。

【００５１】

図２Ｃは、ターゲット処理モジュール１２２の一例をさらに詳細に示す図である。ターゲット処理モジュール１２２は、データ処理モジュールからのＲＡＩＤ構成実行要求および／またはスイッチモジュール１１２を介してリモートターゲット処理モジュールが送信するＲＡＩＤ命令を受信する。ターゲット処理モジュール１２２は、インターフェース１７８、メモリ１８２、および１以上のプロセッサ１８４を有する。

【００５２】

図３を参照しつつ利用について説明すると、第１のデータブロック２００－１はインターフェース１０４を介してデータ処理モジュール１０８で受信される。データ処理モジュール１０８は、当該データブロックに対してＯＳおよびＦＳプロトコル機能を実行する。データ処理モジュール１０８は、当該データブロックを、ストレージアレイ１２０のうちの１つと対応付けられているターゲット処理モジュール１２２に割り当てる。さらに、当該データ処理モジュール１０８は、適用されるべきＲＡＩＤストレージレベルを決定し、ストレージアレイに対してデータブロックをマッピングして、ストレージマップを更新する等してもよい。

【００５３】

例えば、第１のデータブロック２００－１は第１のストレージアレイ１２０－１のターゲット処理モジュール１２２－１に割り当てられるとしてもよい。選択ターゲット処理モジュール１２２－１は、当該データブロックに対してＥＣＣを生成する。ストレージアレイ１２０－１が第１のデータブロックについてＥＣＣを生成している間、データ処理モジュール１０８はインターフェース１０４を介して第２のデータブロック２００－２を受信する。データ処理モジュール１０８は、ＥＣＣ生成のために、ストレージアレイ１２０－２に対応付けられているターゲット処理モジュール１２２－２に第２のデータブロックを割り当てる。

【００５４】

このようにデータブロックについてのＲＡＩＤ構成処理を重複して実行させるが、これは、すべてのターゲット処理モジュールがデータブロックを処理するようになるまで、さらにデータブロック２００－Ｐまで継続することができる。このため、他の方法に比べると、スループットを大幅に向上させ得る。

【００５５】

図４および図５は、データブロック２００－１の処理をさらに詳細に示す図である。処理が終わると、データ処理モジュール１０８は、ストレージアレイ１２０－１のターゲット処理モジュール１２２－１にデータブロック２００－１を送信する。データ処理モジュール１０８はまた、ドライブマップを更新するとしてもよい。ターゲット処理モジュール１２２－１は、当該データブロックについてＥＣＣを処理する。ターゲット処理モジュール１２２－１は、データブロック２００－１に対応付けられているデータのうち一部を、ストレージアレイ１２０－１に対応付けられているローカルドライブ１２４に格納するとしてもよい。さらに、ターゲット処理モジュール１２２－１は、ＲＡＩＤ命令、データ、および／またはＥＣＣデータを、他のストレージアレイに対応付けられているターゲット処理モジュール１２２－２、・・・および１２２－Ｘに送信するとしてもよい。リモートストレージアレイ１２０－２、・・・および１２０－Ｘが有する他のターゲット処理モジュール１２２－２、・・・および１２２－Ｘは、ＲＡＩＤ命令に単純に従っており、処理負荷は制限されている。このため、リモートストレージアレイ１２０－２、・・・および１２０－Ｘが有するターゲット処理モジュール１２２－２、・・・および１２２－Ｘは、他のデータブロックについてＥＣＣを処理することができる。

【００５６】

図５において、ターゲット処理モジュール１２２－１が第１のデータブロック２００－１についてＥＣＣを処理している間に、データ処理モジュール１０８は第２のデータブロック２００－２を受信する。データ処理モジュール１０８は、第２のデータブロック２００－２を、ストレージアレイ１２０－２に対応付けられているターゲット処理モジュール１２２－２に割り当てる。追加のデータブロック２００－Ｐは、その他のストレージアレイ１２０のターゲット処理モジュール１２２に割り当てられるとしてもよい。

【００５７】

図６は、データブロックのＲＡＩＤ処理２５０の一例を概して示す図である。この種のＲＡＩＤ処理はボトルネックを生じさせることがあり、データアクセス時間およびデータ取得時間が短くなる。本開示の一部の実施形態に係るデータ処理を２５２として示す。それぞれのデータブロックに対するＲＡＩＤ構成に係る時間は、一定でないとしてもよい。本開示に係るＲＡＩＤシステムは、格納要求のうちの１つを処理する時間が大幅に長くなっても、データブロックの処理を継続することができる。

【００５８】

図７は、データ格納要求時におけるＲＡＩＤシステム動作方法を示す図である。ステップ３００から開始される。ステップ３０２では、格納されるべきデータブロックがデータ処理モジュール１０８で受信されたか否かを判断する。ステップ３０２が真であれば、ステップ３０４において、データ処理モジュール１０８は、当該データブロックについてのＥＣＣ処理を、複数のターゲット処理モジュール１２２のうち１つに割り当てる。データ処理モジュール１０８はさらに、グローバルドライブマップを更新して、上述したほかの機能を実行するとしてよい。ステップ３０６において、選択ターゲット処理モジュールは、データブロックについてＥＣＣを処理する。選択ターゲット処理モジュールは、リモートストレージアレイに対応付けられるリモートターゲット処理モジュールに、ＲＡＩＤ命令、データ、および／またはＥＣＣデータを送信するとしてもよい。ステップ３１０で終了する。

【００５９】

図８Ａから図１０は、データ取得方法についてさまざまな例を示す。データ取得について説明すると、取得時にエラーが検出されるとＥＣＣ処理が実行されるとしてもよい。エラーは、当該エラーに対応付けられるサブブロックを格納しているハードディスクドライブによって検出されるとしてもよい。エラーが検出されると、ＥＣＣ復元は、同種の複数のターゲット処理モジュール、ターゲット処理モジュールの中から選択される１つのターゲット処理モジュール、および／または、データ処理モジュールによって、ローカルに実行されるとしてもよい。

【００６０】

図８Ａおよび図８Ｂは、データブロックを取得する方法の一例を示すフローチャートである。本実施形態によると、ＥＣＣエラーを含むデータに対するＥＣＣ処理は、データ処理モジュールによって実行される。図８Ａに示す方法はステップ３２０から開始され、ステップ３２２に進んで、データ処理モジュールがデータ取得要求を受信したか否か判断する。ステップ３２２が真であれば、ステップ３２４において、データ処理モジュールは、データ取得要求に対応付けられているデータを有するターゲット処理モジュールすべてに対してブロードキャストメッセージを送信する。これに代えて、データ処理モジュールは、マップを用いて各ターゲット処理モジュールに個別に個別メッセージを送信してもよい。

【００６１】

ステップ３２６において、データ処理モジュールは、エラーを有するターゲット処理モジュールからデータブロック（およびエラーを有するデータに対する対応するＥＣＣデータ）を受信したか否かを判断する。ステップ３２６が真であれば、データ処理モジュールはＥＣＣデータを用いてデータを復元する。ステップ３２６および３２７はステップ３２８に進み、データ処理モジュールは修正されたデータを要求元に送信する。エラーが修正できない場合、データ処理モジュールはエラーメッセージを送信するとしてもよいし、お

よび／または、取得を再試行してもよい。ステップ３２９で終了する。

【００６２】

図８Ｂに示す方法は、ステップ３３０から開始され、ステップ３３２に進む。ステップ３３２において、ターゲット処理モジュールは、データ処理モジュールからデータ取得要求を受信したか否か判断する。ステップ３３４において、ターゲット処理モジュールは、取得要求に関するデータを取得してデータ処理モジュールに送信する。ステップ３３６において、ターゲット処理モジュールはサブブロックにエラーが検出されるか否かを判断する。ステップ３３６が真であれば、ターゲット処理モジュールは、サブブロックに関するＥＣＣデータを、データ処理モジュールに送信する。ステップ３３６および３３７はステップ３３８に進み、データ取得要求に関するデータがすべて送信されたか否かを判断する。送信されていない場合、ステップ３３４に戻る。ステップ３３８が真の場合は、ステップ３３９で終了する。

【００６３】

図９Ａおよび図９Ｂは、データブロックを取得する方法の一例を示すフローチャートである。本実施形態によると、ＥＣＣ処理は、データを格納しているターゲット処理モジュールそれぞれによって実行される。図９Ａに示す方法は、ステップ３４０から開始される。ステップ３４２において、データ処理モジュールはデータ取得要求を受信したか否かを判断する。ステップ３４２が真であれば、ステップ３４４において、データ処理モジュールはすべてのターゲット処理モジュールに対してブロードキャストメッセージを送信する。これに代えて、データ処理モジュールは、マップに基づいて、ターゲット処理モジュールに対して個別メッセージを送信するとしてもよい。ステップ３４８において、データ処理モジュールはデータを受信して要求元に転送する。ステップ３４９で終了する。

【００６４】

図９Ｂに示す方法は、ステップ３５０から開始される。ステップ３５２において、ターゲット処理モジュールはデータ取得要求を受信したか否かを判断する。ステップ３５２が真であれば、ステップ３５４において、ターゲット処理モジュールは、データ処理モジュールに対して、サブブロックで取得要求に関するデータを送信する。ステップ３５６において、サブブロックにおいてエラーが検出されたか否かを判断する。ステップ３５６が真であれば、ステップ３５７に進み、ＥＣＣを処理してデータを復元して、復元されたデータを送信する。データを復元できない場合には、エラーメッセージを送信するとしてもよいし、および／または、再試行してもよい。ステップ３５６および３５７からステップ３５８に進み、データ取得要求に対応付けられているサブブロックすべてが送信されたか否かを判断する。送信されていない場合、ステップ３５４に戻る。送信されている場合には、ステップ３５９で終了する。

【００６５】

図１０は、ＥＣＣ処理が実施されるデータブロックを取得する方法の一例を示すフローチャートである。本実施形態によると、データ復元は、ターゲット処理モジュールのうち選択される１つのターゲット処理モジュールによって実行されるとしてもよい。ステップ３６０から開始され、ステップ３６１に進む。ステップ３６１において、データ処理モジュールはデータ取得要求を受信したか否かを判断する。ステップ３６１が真であれば、ステップ３６２において、データ処理モジュールは、当該データ取得をターゲット処理モジュールのうちの１つに割り当てる。選択ターゲット処理モジュールおよび／またはデータ処理モジュールは、リモートターゲット処理モジュールに対してデータを要求する。

【００６６】

ステップ３６４において、リモートターゲット処理モジュールは、選択ターゲット処理モジュールに対して、取得要求に関連するデータサブブロックを送信する。同様に、選択ターゲット処理モジュールは、ローカルドライブから、取得要求に関連するデータを取得する。これに代えて、リモートターゲット処理モジュールは、データにエラーがなければ、データを直接データ処理モジュールに送信するとしてもよい。エラーがあれば、リモートターゲット処理モジュールは、選択ターゲット処理モジュールにデータを送信してデー

タ復元を行うとしてもよい。

【００６７】

リモートターゲット処理モジュールのそれぞれについて説明すると、ステップ３６６において、リモートターゲット処理モジュールは、データサブブロックのうちのいずれかでエラーが検出されたか否かを判断する。ステップ３６６が真である場合、リモートターゲット処理モジュールは、エラーがあるサブブロックに対応付けられているＥＣＣデータを、選択ターゲット処理モジュールに送信する。エラーがないデータサブブロックは、ターゲット処理モジュールまたはデータ処理モジュールに送信されるとしてもよい。

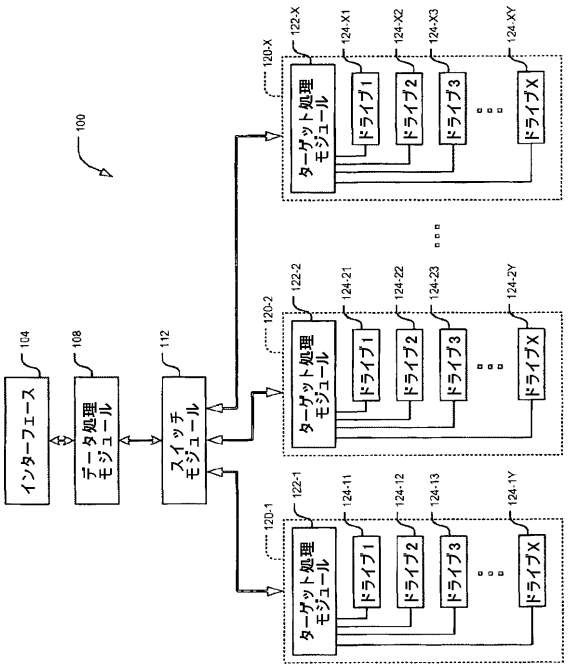
【００６８】

ステップ３６６および３６７からステップ３６８に進み、リモートターゲット処理モジュールに関連する制御において、データサブブロックのすべてが送信されたか否かを判断する。ステップ３７０において、選択ターゲット処理モジュールは、データ復元にＥＣＣデータを用いる、つまり、エラーを修正する。選択ターゲット処理モジュールは、データ処理モジュールにデータを転送する。ステップ３７２において、データ処理モジュールは、復元されたデータを要求元のデバイスに転送する。

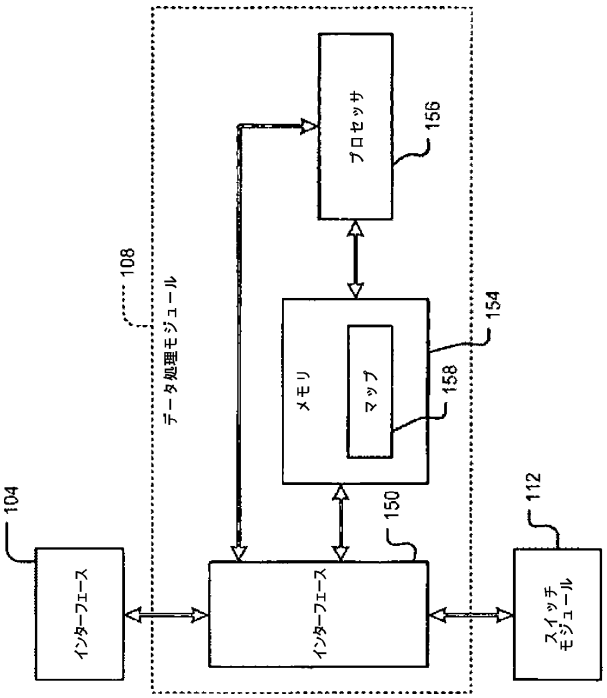
【００６９】

当業者であれば、上述の説明に基づき、本開示の広範囲な教示内容がさまざまな形態で実施され得ることに想到することができる。このため、本開示では具体的な例を紹介したが、本願の図面、明細書および特許請求の範囲を参照することによって当業者はほかの変形例を提案できることが明らかであるので、本開示の真の範囲は記載された具体例に限定されるものではない。

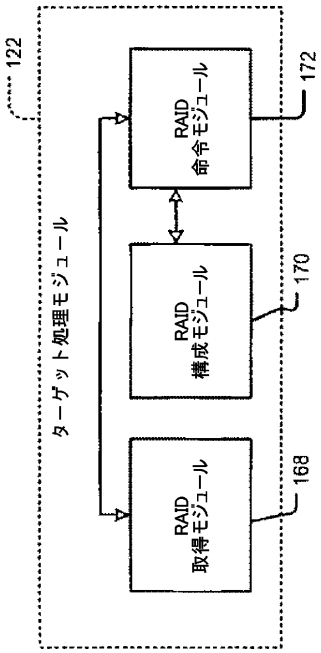
【図１】



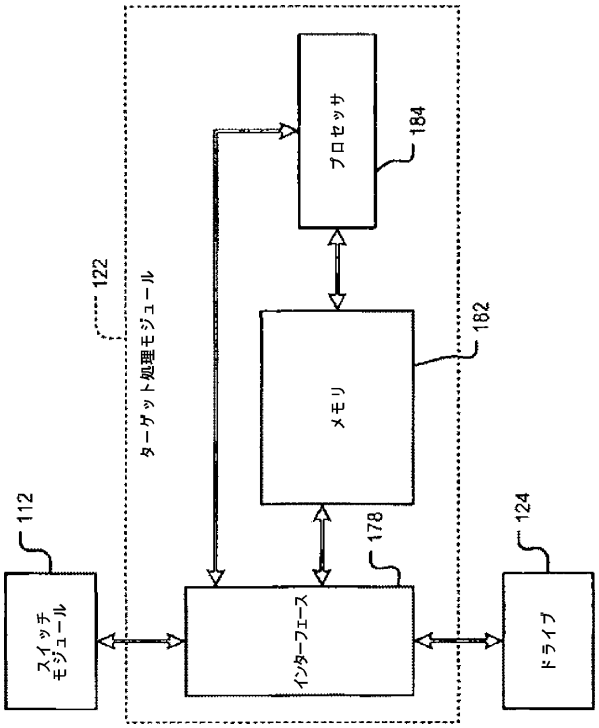
【図２Ａ】



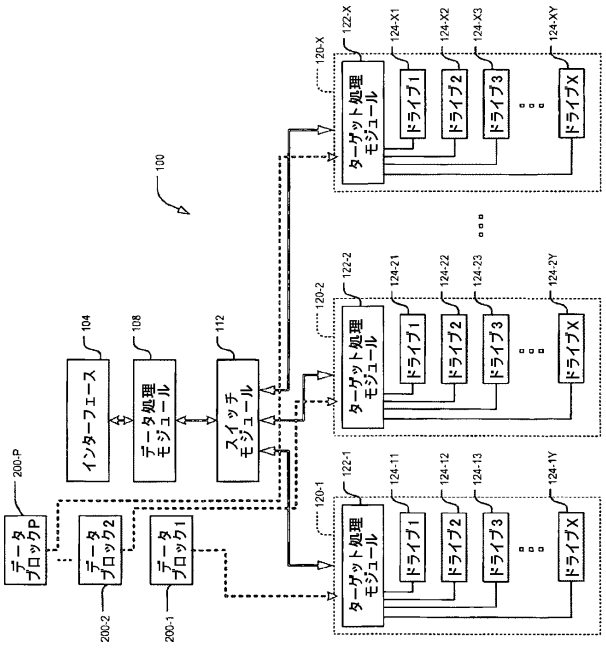
【図２Ｂ】


122
ターゲット処理モジュール
RAID
取得モジュール
168
RAID
構成モジュール
170
RAID
命令モジュール
172

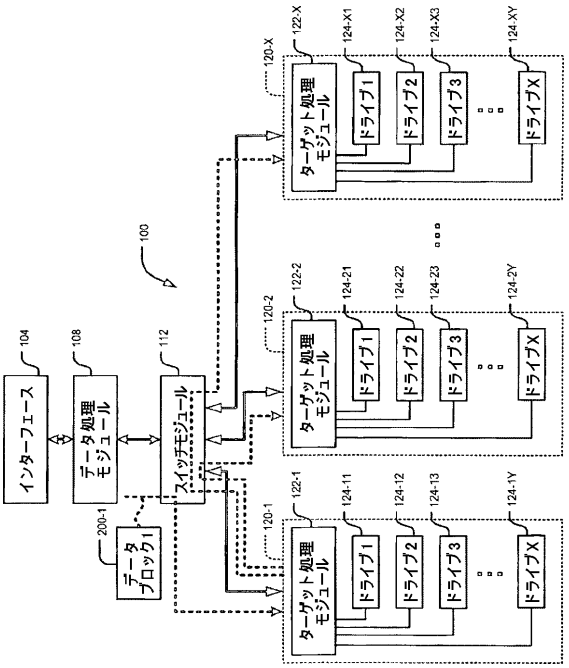
【図２Ｃ】


122
ターゲット処理モジュール
プロセッサ
184
メモリ
182
インターフェース
178
スイッチ
モジュール
112
ドライブ
124

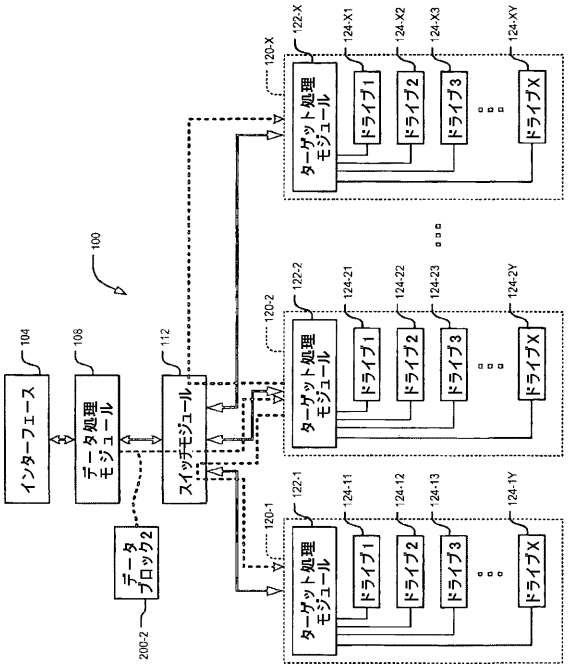
【図３】


100
データ
ブロックP
200-P
データ
ブロック2
200-2
データ
ブロック1
200-1
インターフェース
104
データ処理
モジュール
108
スイッチ
モジュール
112
ターゲット処理
モジュール
120-1
122-1
ドライブ1
ドライブ2
ドライブ3
ドライブX
124-11
124-12
124-13
124-1Y
120-2
122-2
124-21
124-22
124-23
124-2Y
120-X
122-X
124-X1
124-X2
124-X3
124-XY

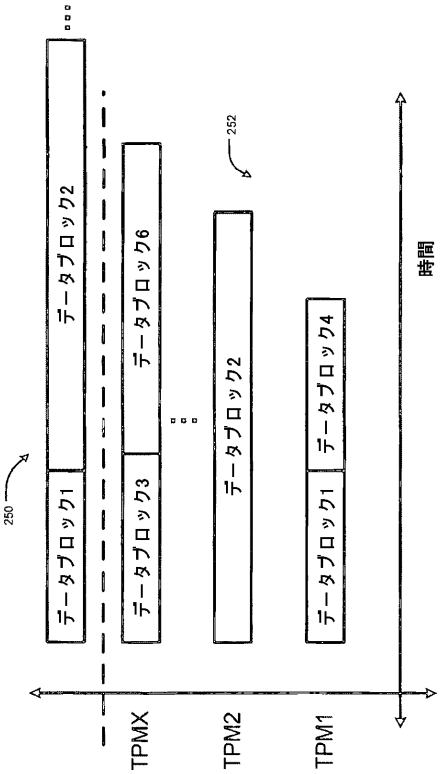
【図４】


100
インターフェース
104
データ処理
モジュール
108
スイッチモジュール
112
データ
ブロック1
200-1
ターゲット処理
モジュール
120-1
122-1
ドライブ1
ドライブ2
ドライブ3
ドライブX
124-11
124-12
124-13
124-1Y
120-2
122-2
124-21
124-22
124-23
124-2Y
120-X
122-X
124-X1
124-X2
124-X3
124-XY

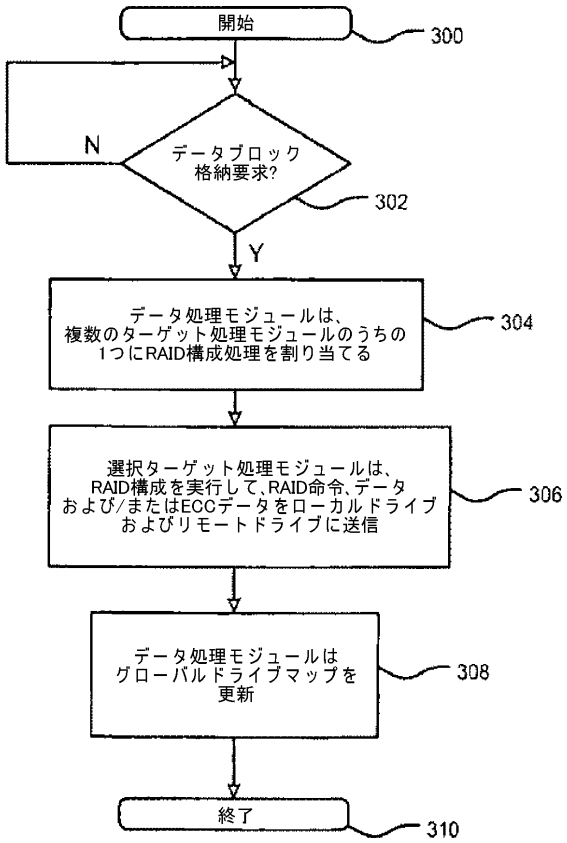
【図５】


100
104
インターフェース
106
データ処理モジュール
112
スイッチモジュール
データブロック2
200-2
120-X
122-X
124-X1
124-X2
124-X3
124-XY
ターゲット処理モジュール
ドライブ1
ドライブ2
ドライブ3
ドライブX
120-2
122-2
124-21
124-22
124-23
124-2Y
120-1
122-1
124-11
124-12
124-13
124-1Y

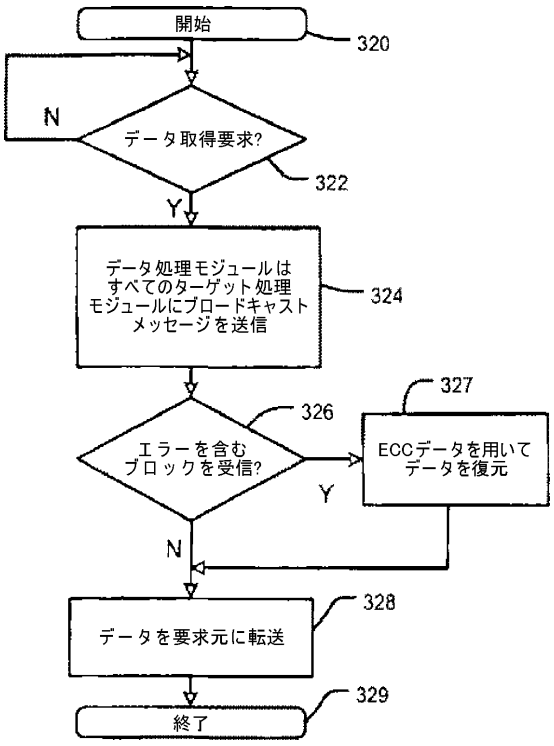
【図６】


250
データブロック2
データブロック6
データブロック3
252
データブロック2
データブロック4
データブロック1
時間
TPMX
TPM2
TPM1

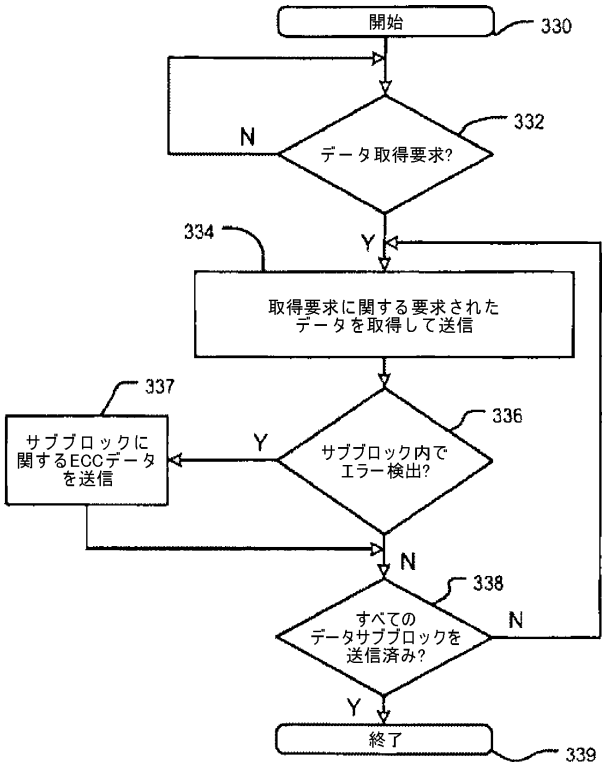
【図７】


開始
300
データブロック格納要求?
302
N
Y
データ処理モジュールは、複数のターゲット処理モジュールのうちの1つにRAID構成処理を割り当てる
304
選択ターゲット処理モジュールは、RAID構成を実行して、RAID命令、データおよび/またはECCデータをローカルドライブおよびリモートドライブに送信
306
データ処理モジュールはグローバルドライブマップを更新
308
終了
310

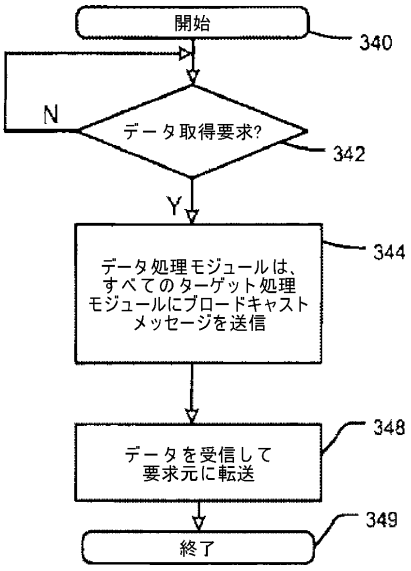
【図８Ａ】


開始
320
データ取得要求?
322
N
Y
データ処理モジュールはすべてのターゲット処理モジュールにブロードキャストメッセージを送信
324
エラーを含むブロックを受信?
326
327
ECCデータを用いてデータを復元
Y
N
データを要求元に転送
328
終了
329

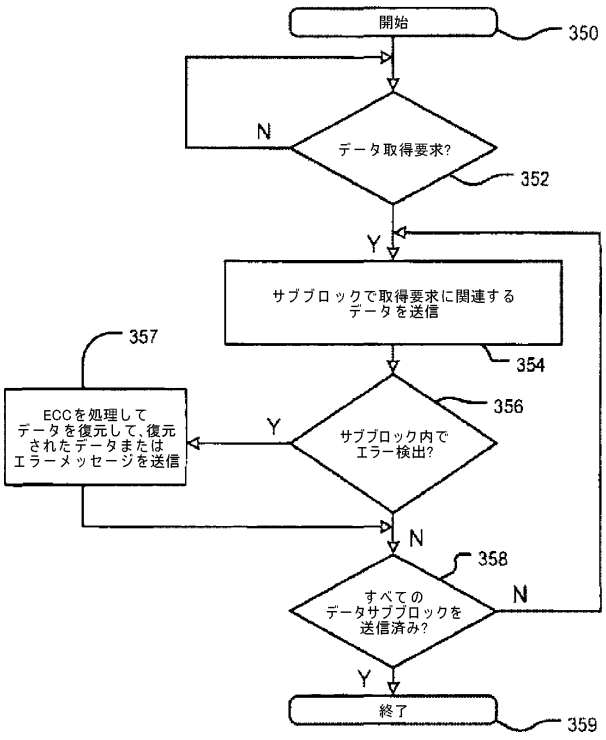
【図８Ｂ】



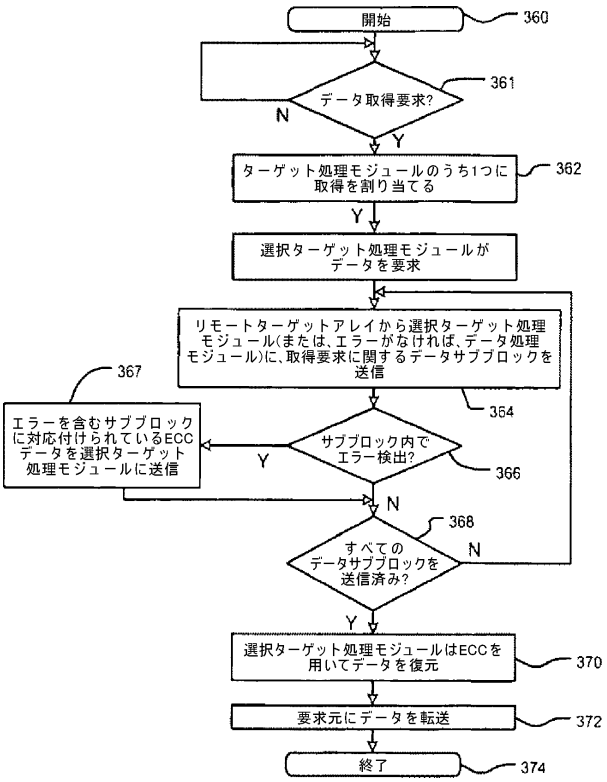
【図９Ａ】



【図９Ｂ】



【図１０】



フロントページの続き

(31)優先権主張番号　11/724,549
(32)優先日　平成19年3月15日(2007.3.15)
(33)優先権主張国　米国(US)

(74)代理人　100112520
弁理士　林　茂則
(74)代理人　100156591
弁理士　高田　学
(72)発明者　スタルジャ、パンタス
アメリカ合衆国、９５０５４　カリフォルニア州、サンタ　クララ、マーベル　レーン　５４８８
マーベル　セミコンダクター　インコーポレイテッド内

審査官　坂東　博司

(56)参考文献　特開平０８－２６３２２４（ＪＰ，Ａ）
特開平０７－２００１８７（ＪＰ，Ａ）
特開平０６－０６７８１４（ＪＰ，Ａ）
特開２００４－１９２４８３（ＪＰ，Ａ）
特開２００５－２７５８２９（ＪＰ，Ａ）
特開２００４－３３４７０６（ＪＰ，Ａ）
欧州特許出願公開第００７３０２２９（ＥＰ，Ａ１）
特表平０５－５０４４３１（ＪＰ，Ａ）
特表２００４－５１４９６８（ＪＰ，Ａ）
米国特許出願公開第２００５／００２２０５２（ＵＳ，Ａ１）
米国特許第０４９８９２０６（ＵＳ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
G06F　3/06