

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7244551号
(P7244551)

(45)発行日 令和5年3月22日(2023.3.22)

(24)登録日 令和5年3月13日(2023.3.13)

(51)国際特許分類	F I
G 0 6 F 3/06 (2006.01)	G 0 6 F 3/06 3 0 1 Z
G 0 6 F 1/3221(2019.01)	G 0 6 F 1/3221
G 0 6 F 1/3287(2019.01)	G 0 6 F 1/3287
G 0 6 F 13/10 (2006.01)	G 0 6 F 3/06 3 0 1 Y
G 0 6 F 13/14 (2006.01)	G 0 6 F 13/10 3 4 0 A
請求項の数 7 (全48頁) 最終頁に続く	

(21)出願番号	特願2020-572897(P2020-572897)	(73)特許権者	503433420 華為技術有限公司 HUAWEI TECHNOLOGIES CO., LTD. 中華人民共和國 5 1 8 1 2 9 広東省深 チェン 市龍崗区坂田 華為総部 ベ ン 公樓 Huawei Administrat ion Building, Banti an, Longgang Distri ct, Shenzhen, Guang dong 5 1 8 1 2 9, P. R. C hina
(86)(22)出願日	平成30年6月27日(2018.6.27)	(74)代理人	100132481 弁理士 赤澤 克豪
(65)公表番号	特表2021-529393(P2021-529393 A)		
(43)公表日	令和3年10月28日(2021.10.28)		
(86)国際出願番号	PCT/CN2018/093196		
(87)国際公開番号	WO2020/000275		
(87)国際公開日	令和2年1月2日(2020.1.2)		
審査請求日	令和3年1月29日(2021.1.29)		
最終頁に続く			

(54)【発明の名称】 ストレージシステム及びストレージシステムの稼働モードを切り替えるための方法

(57)【特許請求の範囲】

【請求項1】

M個の制御デバイスと、K個のドライブとを備える、ストレージシステムであって、M、及びKが正の整数であり、各制御デバイスは、少なくとも1つの第1のコネクタを有し、各ドライブは、少なくとも1つの第2のコネクタを有し、前記K個のドライブのそれぞれは、対応する第1のコネクタ及び第2のコネクタを介して前記M個の制御デバイスの少なくとも1つに取り外し可能に接続されており、前記対応する第1のコネクタ及び前記第2のコネクタを介して、接続された制御デバイスと通信することができ、

前記M個の制御デバイスは、第1の稼働モード及び第2の稼働モードをサポートし、前記K個のドライブは、前記第1の稼働モード及び前記第2の稼働モードをサポートし、

前記M個の制御デバイスが前記第1の稼働モードから前記第2の稼働モードに切り替わるとき、前記K個のドライブは、前記第1の稼働モードから前記第2の稼働モードに切り替わり、

前記第1の稼働モードが集中型ストレージアーキテクチャで動作するモードであるとき、前記第2の稼働モードは、分散型ストレージアーキテクチャで動作するモードであり、前記第1の稼働モードが分散型ストレージアーキテクチャで動作するモードであるとき、前記第2の稼働モードは、集中型ストレージアーキテクチャで動作するモードであり、前記集中型ストレージアーキテクチャで動作するモードは、前記M個の制御デバイスのうちの制御デバイスが1つのデータを前記K個のドライブのうちの1つに格納するモードであり、前記分散型ストレージアーキテクチャで動作するモードは、前記M個の制御デバイス

のうちの制御デバイスが1つのデータを複数のデータに分割し、かつ前記複数のデータを前記K個のドライブのうちの複数のドライブに分散して格納するモードであり、
前記K個のドライブのそれぞれは、高速信号インターフェース及び低速信号インターフェースと、内部ドライブ制御モジュールとを備え、
前記高速信号インターフェースは、制御デバイスにサービスデータを送信するように構成され、

前記低速信号インターフェースは、ドライブについての情報を取得するために使用される第1の命令と、稼働モードを前記分散型ストレージアーキテクチャで動作する前記モードに変更するように命令するために使用される第2の命令とを受信するように構成され、

各ドライブの前記内部ドライブ制御モジュールは、

前記第2の命令が前記低速信号インターフェースを通じて受信された後、前記各ドライブのP個の第2のコネクタを通信停止状態となるように制御するように構成され、Pは正の整数であり、前記各ドライブの第2のコネクタの数とPとの差は1である、

ストレージシステム。

【請求項2】

前記K個のドライブのうちの第1のドライブは、X個の第2のコネクタを有し、
前記X個の第2のコネクタのうちの少なくとも2つの第2のコネクタは、異なる制御デバイスの第1のコネクタに接続され、Xが正の整数であり、XがM以下である
請求項1に記載のシステム。

【請求項3】

前記M個の制御デバイスのうちの第1の制御デバイスは、
前記第1の制御デバイスに接続されている各ドライブについての情報を取得し、
前記各ドライブについての前記情報に基づいて前記各ドライブの稼働モードを決定し、
前記各ドライブの前記稼働モードが前記第1の制御デバイスの現在の稼働モードと適合すると決定するとき、前記第1の制御デバイスの第1のコネクタのうち、稼働モードが適合するドライブに接続された第1のコネクタに電力を供給する
ように構成される

請求項1又は2に記載のシステム。

【請求項4】

前記第1の制御デバイスが、
前記各ドライブについての前記情報に基づいて前記各ドライブの前記稼働モードを決定した後で、前記各ドライブの前記稼働モードが前記第1の制御デバイスの前記現在の稼働モードと適合しないと決定するとき、前記第1の制御デバイスの前記第1のコネクタのうち、稼働モードが適合しないドライブに接続された第1のコネクタに電力を供給することを停止する

ようにさらに構成される

請求項3に記載のシステム。

【請求項5】

前記第1の制御デバイスが、
前記各ドライブの状態を取得し、
前記第1の制御デバイスに接続されているS個のドライブが切断状態又は欠陥状態にあると決定するとき、前記S個のドライブに接続されているS個の第1のコネクタに電力を供給することを停止する

ように構成される

請求項3に記載のシステム。

【請求項6】

前記K個のドライブのそれぞれは、内部ドライブ制御モジュールをさらに備え、
前記低速信号インターフェースは、稼働モードを前記集中型ストレージアーキテクチャで動作する前記モードに変更するように命令するために使用される第3の命令を受信する
ようにさらに構成され、

10

20

30

40

50

各ドライブの前記内部ドライブ制御モジュールは、

前記第 3 の命令が前記低速信号インターフェースを通じて受信された後で、かつ前記各ドライブの P 個の第 2 のコネクタが通信停止状態にあるとき、前記 P 個の第 2 のコネクタのうち L 個の第 2 のコネクタを制御して制御デバイスと通信することを再開するように構成され、L は正の整数であり、L は P 以下である

請求項 1 に記載のシステム。

【請求項 7】

M 個の制御デバイスと、K 個のドライブとを備えるストレージシステムの稼働モードを切り替えるための方法であって、M 及び K は正の整数であり、前記方法は、

前記 M 個の制御デバイスによって、第 1 の稼働モードで動作し、かつ第 2 の稼働モードをサポートするステップであって、前記第 1 の稼働モードが集中型ストレージアーキテクチャで動作するモードであるとき、前記第 2 の稼働モードは、分散型ストレージアーキテクチャで動作するモードであり、前記第 1 の稼働モードが分散型ストレージアーキテクチャで動作するモードであるとき、前記第 2 の稼働モードは、集中型ストレージアーキテクチャで動作するモードであり、前記集中型ストレージアーキテクチャで動作するモードは、前記 M 個の制御デバイスのうちの制御デバイスが 1 つのデータを前記 K 個のドライブのうち 1 つに格納するモードであり、前記分散型ストレージアーキテクチャで動作するモードは、前記 M 個の制御デバイスのうちの制御デバイスが 1 つのデータを複数のデータに分割し、かつ前記複数のデータを前記 K 個のドライブのうち複数のドライブに分散して格納するモードであり、M が正の整数である、ステップと、

前記 M 個の制御デバイスのうちの第 1 の制御デバイスによって、前記第 1 の制御デバイスに接続されている各ドライブに第 1 の命令を送信するステップであって、前記第 1 の命令は、前記第 2 の稼働モードで動作するように命令するために使用され、各制御デバイスは、少なくとも 1 つの第 1 のコネクタを有し、各ドライブは、少なくとも 1 つの第 2 のコネクタを有し、前記 K 個のドライブのそれぞれは、対応する第 1 のコネクタ及び第 2 のコネクタを介して前記 M 個の制御デバイスの少なくとも 1 つに取り外し可能に接続されており、前記対応する第 1 のコネクタ及び前記第 2 のコネクタを介して、接続された制御デバイスと通信することができる、ステップと、

前記各ドライブによって、前記第 1 の命令に従って、前記第 2 の稼働モードに適合した状態となるように前記各ドライブの P 個の第 2 のコネクタを制御するステップであって、X 及び P が正の整数であり、前記各ドライブの第 2 のコネクタの数と P との差は 1 である、ステップと

を備え、

前記 P 個の第 2 のコネクタを制御する前記ステップは、接続されている制御デバイスと通信することを停止するように前記 P 個の第 2 のコネクタを制御するステップを備え、前記第 2 の稼働モードが前記集中型ストレージアーキテクチャで動作する前記モードであるとき、前記 P 個の第 2 のコネクタを制御する前記ステップは、前記 P 個の第 2 のコネクタのうち L 個の第 2 のコネクタを介して接続されている制御デバイスと通信することを再開するように前記 L 個の第 2 のコネクタを制御するステップを備え、L は正の整数であり、L は P 以下である、方法。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、ストレージ技術の分野に関し、詳細には、ストレージシステム、及びストレージシステムの稼働モードを切り替えるための方法に関する。

【背景技術】

【0002】

ビッグデータ分析技術及びモノのインターネット技術などの技術の発展により、データは関連技術の進化を促進するための中核的な要素になった。従って、データストレージの信頼性を確保することがより重要になっている。

10

20

30

40

50

【 0 0 0 3 】

データストレージシステムは、ドライブコントローラ及びドライブを含む。ストレージシステムのストレージアーキテクチャは、ドライブコントローラがドライブにアクセスする方式に基づいて、集中型ストレージアーキテクチャ及び分散型ストレージアーキテクチャへと分類される。従来技術では、集中型ストレージアーキテクチャと分散型ストレージアーキテクチャの両方が固有の設計である。具体的には、2つのストレージアーキテクチャのうちのストレージアーキテクチャにかかわらず、ドライブコントローラ及びドライブは固定的に接続され、分離されることができない。一方のストレージアーキテクチャが他方のストレージアーキテクチャに切り替えられる必要がある場合、例えば、集中型ストレージアーキテクチャが分散型ストレージアーキテクチャに切り替えられるべきである場合、ストレージアーキテクチャ全体が購入される必要がある。具体的には、統合される分散型ストレージアーキテクチャのドライブコントローラ及び対応するドライブが購入される必要がある。明らかに、コストは相対的に高い。従来技術のストレージアーキテクチャはフレキシブルではなく、フレキシブルな展開に対して不便であることが、承知されることができ

10

【 発明の概要 】

【 0 0 0 4 】

本出願の実施形態は、ストレージアーキテクチャのフレキシビリティを改善するための、ストレージシステムと、ストレージシステムの稼働モードを切り替えるための方法とを提供する。

20

【 0 0 0 5 】

第1の態様によれば、本出願の実施形態はストレージシステムを提供し、ストレージシステムは、M個の制御デバイスと、N個の第1のコネクタと、N個の第2のコネクタと、K個のドライブとを含み、M、N、及びKが正の整数であり、

N個の第1のコネクタがM個の制御デバイスに取り外し可能に接続するように構成され、N個の第2のコネクタが、K個のドライブに取り外し可能に接続するように構成され、又はK個のドライブに固定的に接続するように構成され、K個のドライブの各々がデータを記憶するように構成され、

M個の制御デバイスの各々が、N個の第1のコネクタの全て及びN個の第2のコネクタの全てを使用することによってK個のドライブと通信し、又は、M個の制御デバイスの各々が、N個の第1のコネクタのいくつか及びN個の第2のコネクタのいくつかを使用することによってK個のドライブと通信し、

30

M個の制御デバイスの各々が第1の稼働モード及び第2の稼働モードをサポートし、K個のドライブの各々が第1の稼働モード及び第2の稼働モードをサポートし、M個の制御デバイスが第1の稼働モードから第2の稼働モードに切り替わるとき、K個のドライブが第1の稼働モードから第2の稼働モードに切り替わるように制御され、第1の稼働モードが集中型ストレージアーキテクチャで動作するモードであるとき、第2の稼働モードが分散型ストレージアーキテクチャで動作するモードであり、又は第1の稼働モードが分散型ストレージアーキテクチャで動作するモードであるとき、第2の稼働モードが集中型ストレージアーキテクチャで動作するモードである。

40

【 0 0 0 6 】

前述の技術的解決法において、制御デバイスは第1のコネクタに取り外し可能に接続されることがあり、ドライブは第2のコネクタに取り外し可能に又は固定的に接続されることがあるので、ユーザは、使用上の要件に基づいてドライブと制御デバイスとの間の接続関係をフレキシブルに構成して、異なるストレージアーキテクチャを構成し得る。加えて、制御デバイスとドライブの両方が2つの稼働モードをサポートするので、ストレージアーキテクチャが変更される必要があるとき、制御デバイスとドライブの稼働モードだけが切り替えられる必要があり、又は、ドライブだけが交換される必要があり、新しい制御デバイスと新しいドライブが購入される必要がない。これは、ストレージシステムのフレキシビリティを改善し、ストレージアーキテクチャの切り替えのコストを減らす。

50

【 0 0 0 7 】

ある可能な設計において、M個の制御デバイスの各々が、N個の第1のコネクタの全て又はいくつかに接続され、又は、M個の制御デバイスの各々が、N個の第1のコネクタの全て又はいくつかを含む。

【 0 0 0 8 】

前述の技術的解決法において、制御デバイスは、複数の方式で第1のコネクタに接続され得る。例えば、制御デバイスは、第1のコネクタに電氣的に接続され得る。代替的に、制御デバイスは第1のコネクタを含み得る。例えば、第1のコネクタは制御デバイスに配置される。ストレージシステムのフレキシビリティが改善されることができ。

【 0 0 0 9 】

ある可能な設計において、K個のドライブのうちの1つがX個の第2のコネクタを有し、M個の制御デバイスの各々が、N個の第1のコネクタの全て又はいくつかを使用することによって、K個のドライブの各々の1つの第2のコネクタに取り外し可能に接続され、1つの第1のコネクタが1つの第2のコネクタに接続され、X個の第2のコネクタに接続された第1のコネクタが異なる制御デバイスに接続され、又は、X個の第2のコネクタに接続された第1のコネクタが異なる制御デバイスに含まれ、Xは正の整数であり、XはM以下である。

【 0 0 1 0 】

前述の技術的解決法において、各ドライブは複数の第2のコネクタを含んでもよく、ドライブの複数の第2のコネクタは、複数の第1のコネクタを使用することによって異なる制御デバイスに接続される。このようにして、ドライブに含まれる全ての複数の第2のコネクタが通信可能状態にあるとき、ドライブの複数の第2のコネクタに接続される異なる制御デバイスが全て、ドライブにアクセスすることがあり、集中型ストレージアーキテクチャを構成する。ドライブに含まれる複数の第2のコネクタのうちの1つだけが通信可能状態にあるとき、ドライブの複数の第2のコネクタに接続される異なる制御デバイスのうちの1つだけがドライブにアクセスすることができ、使用上の要件に基づく分散型ストレージアーキテクチャを構成する。ドライブの第2のコネクタの状態は、異なるストレージアーキテクチャを構成するためにフレキシブルに設定される。

【 0 0 1 1 】

ある可能な設計において、1つの第1のコネクタ及び1つの第2のコネクタは、相対応する垂直で直交するコネクタであり、又は、1つの第1のコネクタ及び1つの第2のコネクタは、予め設定された回転角を有する相対応するコネクタである。

【 0 0 1 2 】

前述の技術的解決法において、第1のコネクタ及び第2のコネクタは、空間座標系のX軸、Y軸、及びZ軸の方向における、制御デバイスとドライブとの間の相対的なシフトに打ち勝つために使用され得る。従って、複数の制御デバイスが複数のドライブに確実に接続されることができ、従来技術のバックプレーンが取り除かれることができ、ストレージシステムの設計の複雑さを下げ、バックプレーンの欠陥によりストレージシステム全体が使用されることができなくなる確率を下げ、そうすることによって、ストレージシステムの信頼性を改善する。

【 0 0 1 3 】

さらに、従来技術のバックプレーンが取り除かれ、そうすることによって、通信処理における制御デバイスとドライブとの間でのインピーダンス不整合点が減らされることができ。具体的には、制御デバイスとバックプレーンとの間のインピーダンス不整合、及びバックプレーンとドライブとの間のインピーダンス不整合が減らされる。このことは、信号送信レートを改善することができる。

【 0 0 1 4 】

さらに、従来技術のバックプレーンが取り除かれ、そうすることによって、ストレージシステムにおけるエアフロー欠陥点が減らされ、ストレージシステムの熱放散性能が改善されることができ。従って、ストレージシステムの特定期間において、ストレージシス

10

20

30

40

50

テムの特定熱放散性能により、ストレージシステム内に、より多くのデバイスが許容されることができ、ストレージシステムの密度が改善されることができる。

【0015】

ある可能な設計において、M個の制御デバイスのうちの1つが、K個のドライブの各々についての情報を取得し、次いで、K個のドライブの各々についての情報に基づいてK個のドライブの稼働モードを決定し、K個のドライブの稼働モードが制御デバイスの現在の稼働モードと適合すると決定するとき、K個の第1のコネクタに電力を供給するように構成される。

【0016】

前述の技術的解決法において、制御デバイスは、ドライブの稼働モードが制御デバイスの稼働モードと適合すると決定するときだけ、ドライブに電力を供給する。このようにして、ストレージシステムの電気エネルギーが適切に使用されることができる。

10

【0017】

ある可能な設計において、制御デバイスは、K個のドライブの各々についての情報に基づいてK個のドライブの稼働モードを決定した後で、K個のドライブの稼働モードが制御デバイスの現在の稼働モードと適合しないと決定するとき、K個の第1のコネクタに電力を供給することを停止するようにさらに構成される。

【0018】

ある可能な設計において、制御デバイスは、K個のドライブの各々の状態を取得し、K個のドライブのうちS個のドライブが切断状態又は欠陥状態にあると決定するとき、S個のドライブに接続されるS個の第1のコネクタに電力を供給することを停止するように構成される。

20

【0019】

前述の技術的解決法において、制御デバイスはまず、制御デバイスに接続されるドライブの稼働状態を決定してよく、ドライブの稼働状態が切断状態又は欠陥状態にあると決定するとき、制御デバイスはドライブに電力を供給しない。このようにして、ストレージシステムの電気エネルギーが適切に使用されることができ、欠陥状態にあるドライブが熱を放散することもなく、そうすることによって、ストレージシステムの熱放散が減らされることができる。

【0020】

ある可能な設計において、K個のドライブのうち1つが、高速信号インターフェース及び低速信号インターフェースを含む。高速信号インターフェースは、X個の第2のコネクタ、及びX個の第2のコネクタに接続される第1のコネクタを使用することによって、制御デバイスにサービスデータを送信するように構成され、低速信号インターフェースは、X個の第2のコネクタから入力され、かつドライブについての情報を取得するために使用される、第1の命令を受信するように構成される。

30

【0021】

前述の技術的解決法において、ドライブは、異なるインターフェースを通じて制御デバイスと通信してもよく、そうすることによって、ドライブが制御デバイスによって送信される信号を管理する。

40

【0022】

ある可能な設計において、K個のドライブの各々がさらに、内部ドライブ制御モジュールを含む。K個のドライブの各々に対して、低速信号インターフェースは、X個の第2のコネクタに接続される第1のコネクタに対応する制御デバイスのうちの第1の制御デバイスによって送信され、第2の命令を受信するドライブの稼働モードを分散型ストレージアーキテクチャで動作するモードに変更するために命令するために使用される、第2の命令を受信するようにさらに構成され、内部ドライブ制御モジュールは、第2の命令が低速信号インターフェースを通じて受信された後で、通信停止状態となるようにX個の第2のコネクタのうちP個の第2のコネクタを制御するように構成され、1つの第2のコネクタが通信停止状態にあるということは、ドライブが、第2のコネクタを使用することによ

50

て、第2のコネクタに接続される第1のコネクタに対応する制御デバイスと通信することを停止するように制御されるということを意味し、Pは正の整数であり、XとPの差は1である。

【0023】

前述の技術的解決法において、ドライブは、制御デバイスによって送信される命令に基づいて、通信停止状態となるようにドライブの第2のコネクタを制御してもよく、そうすることによって、ドライブの稼働モードは、分散型ストレージアーキテクチャにおけるモードに切り替えられ、ハードウェア構造及び接続関係を変更することなく、ストレージシステムの稼働モードが切り替えられることができ、ストレージシステムの稼働モードを切り替えるステップが簡略化されることができる。

10

【0024】

ある可能な設計において、K個のドライブの各々はさらに、内部ドライブ制御モジュールを含む。K個のドライブの各々に対して、低速信号インターフェースは、X個の第2のコネクタに接続される第1のコネクタに対応する制御デバイスのうちの第1の制御デバイスによって送信され、第3の命令を受信するドライブの稼働モードを集中型ストレージアーキテクチャで動作するモードに変更するように命令するために使用される、第3の命令を受信するようにさらに構成され、内部ドライブ制御モジュールは、第3の命令が低速信号インターフェースを通じて受信された後で、X個の第2のコネクタのうちのP個の第2のコネクタが通信停止状態にあるとき、L個の第2のコネクタを使用することによって、L個の第2のコネクタに接続されるL個の第1のコネクタに対応する制御デバイスと通信

20

【0025】

第2の態様によれば、本出願は、ストレージシステムの稼働モードを切り替えるための方法を提供し、この方法は、次のステップを含む。ストレージシステムの中のM個の制御デバイスが第1の稼働モードで動作し、M個の制御デバイスが第2の稼働モードをサポートするとき、M個の制御デバイスのうちの少なくとも1つが、ストレージシステムの中のK個のドライブの各々に、第1の命令を受信するドライブに第2の稼働モードで動作するように命令するために使用される第1の命令を送信し、そうすることによって、第1の命令を受信するK個のドライブの各々は、第1の命令に従って、第2の稼働モードに適応した状態となるようにX個の第2のコネクタのうちのP個の第2のコネクタを制御し、第1の稼働モードが集中型ストレージアーキテクチャで動作するモードであるとき、第2の稼働モードが分散型ストレージアーキテクチャで動作するモードであり、又は、第1の稼働モードが分散型ストレージアーキテクチャで動作するモードであるとき、第2の稼働モードが集中型ストレージアーキテクチャで動作するモードである。M個の制御デバイスの各々は、N個の第1のコネクタの全て及びN個の第2のコネクタの全てを使用することによってK個のドライブと通信し、又は、M個の制御デバイスの各々は、N個の第1のコネクタのいくつか及びN個の第2のコネクタのいくつかを使用することによってK個のドライブと通信し、N個の第1のコネクタは、M個の制御デバイスに取り外し可能に接続するように構成され、N個の第2のコネクタは、K個のドライブに取り外し可能に接続するよう

30

40

【0026】

前述の技術的解決法において、ストレージシステムが稼働モードを切り替える必要があるとき、制御デバイスの稼働モードが切り替えられ、制御デバイスが、稼働モード切り替え命令をドライブに送信し、ハードウェア構造及び接続関係を変更することなくストレージシステムの稼働モードを切り替える。ストレージシステムの稼働モードを切り替えるステップが簡略化されることができる。

【0027】

50

ある可能な設計において、制御すること、第2の稼働モードが分散型ストレージアーキテクチャで動作するモードであるとき、第2の稼働モードに適応した状態となるようにP個の第2のコネクタのうちの1つを制御することは、第2のコネクタを使用することによって、第2のコネクタに接続される第1のコネクタに対応する制御デバイスと通信することを停止するようにドライブを制御することを含み、又は、第2の稼働モードが集中型ストレージアーキテクチャで動作するモードであるとき、第2の稼働モードに適応した状態となるようにP個の第2のコネクタのうちの1つを制御することは、P個の第2のコネクタのうちのL個の第2のコネクタを使用することによって、L個の第2のコネクタに接続されるL個の第1のコネクタに対応する制御デバイスと通信することを再開するようにドライブを制御することを含み、Lは正の整数であり、LはP以下である。

10

【0028】

前述の技術的解決法において、ドライブは、制御デバイスによって送信される命令に従ってドライブの第2のコネクタの通信ステータスを制御することによって、ドライブの稼働モードを切り替え得る。例えば、分散型ストレージアーキテクチャで動作するモードが切り替えられる必要があるとき、M個の第2のコネクタのいくつかは、通信停止状態となるように制御され、集中型ストレージアーキテクチャで動作するモードが切り替えられる必要があるとき、通信停止状態にある第2のコネクタのいくつか又は全てが通信を再開する。従って、ドライブの稼働モードは集中型ストレージアーキテクチャにおけるモードに切り替えられ、ドライブの稼働モードはハードウェア構造及び接続関係を変更することなく切り替えられることができる。

20

【0029】

第3の態様によれば、本出願の実施形態は、コンピュータ可読記憶媒体をさらに提供し、コンピュータ可読記憶媒体はコンピュータプログラムを含み、コンピュータプログラムが端末上で動作するとき、ストレージシステムは、第2の態様の方法又は第2の態様の可能な設計のうちの任意の1つを実行することが可能にされる。

【0030】

第4の態様によれば、本出願の実施形態は、コンピュータプログラム製品をさらに提供し、コンピュータプログラム製品が端末上で動作するとき、ストレージシステムは、第2の態様の方法又は第2の態様の可能な設計のうちの任意の1つを実行することが可能にされる。

30

【0031】

第3の態様及び第4の態様の有益な効果並びに第3の態様及び第4の態様の実装については、第2の態様の方法の有益な効果及び第2の態様の実装の説明を参照されたい。

【図面の簡単な説明】

【0032】

【図1A】本出願のある実施形態による適用シナリオの概略図である。

【図1B】本出願のある実施形態による適用シナリオの概略図である。

【図2A】集中型ストレージシステムの例のアーキテクチャ図である。

【図2B】分散型ストレージシステムの例のアーキテクチャ図である。

【図3A】本出願のある実施形態によるストレージシステム300の概略構造図である。

40

【図3B】本出願のある実施形態によるストレージシステム300の別の概略構造図である。

【図3C】本出願のある実施形態によるストレージシステム300の別の概略構造図である。

【図3D】本出願のある実施形態によるストレージシステム300の別の概略構造図である。

【図4】本出願のある実施形態によるストレージシステム300の制御デバイス301に含まれるプロセッサ3011の概略構造図である。

【図5】本出願のある実施形態によるストレージシステム300の制御デバイス301に含まれる複数の個別の構成要素の概略構造図である。

50

【図 6】本出願のある実施形態によるストレージシステム 300 にさらに含まれる検出回路 3021 の概略構造図である。

【図 7】本出願のある実施形態によるストレージシステム 300 にさらに含まれる P C I E スイッチの概略構造図である。

【図 8 A】本出願のある実施形態による分散型ドライブの概略図である。

【図 8 B】本出願のある実施形態による分散型ドライブの別の概略図である。

【図 9 A】本出願のある実施形態による集中型ドライブの概略図である。

【図 9 B】本出願のある実施形態による集中型ドライブの 2 つの第 2 のコネクタ間の離隔の概略図である。

【図 9 C】本出願のある実施形態による集中型ドライブの別の概略図である。

10

【図 10 A】本出願のある実施形態による汎用ドライブの概略図である。

【図 10 B】本出願のある実施形態による汎用ドライブの別の概略図である。

【図 10 C】本出願のある実施形態による、通信停止状態となるように汎用ドライブが第 2 のコネクタ 1002 を制御する概略図である。

【図 10 D】本出願のある実施形態による、汎用ドライブがデータストレージモジュール 1006 を 2 つの独立したアクセス領域へと分割する概略図である。

【図 11】本出願のある実施形態による、分散型ストレージアーキテクチャにおける制御デバイス A に欠陥がある場合の概略図である。

【図 12 A】本出願のある実施形態による集中型ストレージアーキテクチャのストレージシステムの概略図である。

20

【図 12 B】本出願のある実施形態による集中型ストレージアーキテクチャのストレージシステムの側面図である。

【図 13 A】本出願のある実施形態によるドライブ 303 の概略図である。

【図 13 B】本出願のある実施形態による 1 つの第 1 のコネクタ 302 を含む制御デバイス 301 の側面図である。

【図 13 C】本出願のある実施形態による 1 つの第 2 のコネクタを含むドライブ 303 の側面図である。

【図 13 D】本出願のある実施形態による制御デバイス 301 とドライブ 303 との間の接続関係の概略図である。

【図 14 A】本出願のある実施形態による 1 つの第 1 のコネクタ 302 を含む別の制御デバイス 301 の側面図である。

30

【図 14 B】本出願のある実施形態による 1 つの第 2 のコネクタを含む別のドライブ 303 の側面図である。

【図 14 C】本出願のある実施形態による制御デバイス 301 とドライブ 303 との間の別の接続関係の概略図である。

【図 15 A】本出願のある実施形態によるストレージシステム動作方法のフローチャートである。

【図 15 B】本出願のある実施形態によるストレージシステム動作方法のフローチャートである。

【図 15 C】本出願のある実施形態によるストレージシステム動作方法のフローチャートである。

40

【図 16】本出願のある実施形態によるストレージシステムの稼働モードを切り替えるための方法のフローチャートである。

【発明を実施するための形態】

【0033】

本出願の実施形態の目的、技術的解決法、及び利点をより明確にするために、以下は、本明細書の添付の図面及び具体的な実装を参照して、本出願の実施形態の技術的解決法を詳しく説明する。

【0034】

本明細書における「及び/又は」という用語は、関連付けられる物体を記述するための

50

関連付けの関係のみを記述し、3つの関係が存在し得ることを表す。例えば、A及び/又はBは、Aのみが存在する、AとBの両方が存在する、及びBのみが存在するという、3つの場合を表し得る。加えて、本明細書における文字「/」は全般に、別段規定されない限り、関連付けられる物体間の「又は」の関係を示す。

【0035】

加えて、本出願の実施形態における「複数の」は「2つ以上の」を意味することが理解されるべきである。「第1の」及び「第2の」などの語は、区別及び説明のために使用されるだけであり、相対的な重要性のインジケーション又は示唆、或いは、順序のインジケーション又は示唆として理解されるべきではない。

【0036】

本出願の実施形態はストレージシステムを提供する。ストレージシステムは、図1A又は図1Bに示されるデータ読み取り/書き込みシナリオに適用される。図1Aに示されるデータ読み取り/書き込みシナリオでは、クライアントは、ドライブコントローラを使用することによって、ドライブコントローラに接続されるドライブにアクセスする。例えば、クライアントがドライブの中のデータを取得する必要があるとき、クライアントは、データ読み取り命令をドライブコントローラに送信する。データ読み取り命令を受信した後で、ドライブコントローラは、データを記憶するドライブからデータを取得し、データをクライアントに送信する。クライアントがデータを記憶する必要があるとき、クライアントは、データ書き込み命令をドライブコントローラに送信する。データ書き込み命令は、記憶されるべきデータを搬送する。データ書き込み命令を受信した後で、ドライブコントローラは、記憶されるべきデータをデータ書き込み命令から取得し、対応するドライブに記憶されるべきデータを記憶する。

【0037】

図1Aとは異なり、図1Bに示されるデータ読み取り/書き込みシナリオでは、クライアントは、クラウドネットワークを使用することによって、データ読み取り命令又はデータ書き込み命令をドライブコントローラに送信する。クラウドネットワークを使用することによってデータ読み取り命令を受信し、データを対応するドライブから取得した後で、ドライブコントローラは、クラウドネットワークを使用することによってデータをクライアントに送信する。クラウドネットワークを使用することによってデータ書き込み命令を受信した後で、ドライブコントローラは、対応するドライブに、データ書き込み命令から取得された記憶されるべきデータを記憶する。

【0038】

図1A及び図1Bに示されるシナリオにおいて、1つのクライアント、1つのドライブコントローラ、及び1つのドライブの間の相互作用は、説明のための例として使用されるだけであり、本出願の実施形態の適用シナリオに対するいかなる制限も構成すべきではないことに留意されたい。実際のデータ読み取り/書き込みシナリオでは、複数のクライアント、複数のドライブコントローラ、及び複数のドライブがあり得る。例えば、複数のクライアントは、複数のドライブコントローラの任意の1つに接続されるドライブにアクセスしてよく、又は、複数のクライアントは、固有のドライブコントローラに接続されるドライブだけにアクセスしてよい。これは、本出願の実施形態では限定されない。

【0039】

ドライブコントローラ及びドライブは、集中型ストレージアーキテクチャにあることがある。図2Aに示されるように、集中型ストレージアーキテクチャは、ドライブコントローラA、ドライブコントローラB、複数のドライブ、及びバックプレーンを含む。集中型ストレージアーキテクチャでは、1つのデータが複数のドライブのうちの1つに記憶される。例えば、ユーザAのビデオデータはドライブ1に記憶され、ユーザAの音声データはドライブ2に記憶される。動作が、あるデータに対して実行される必要があるとき、ドライブコントローラA又はドライブコントローラBは、命令を送信し、バックプレーンを使用することによって命令をドライブに出力して、対応するドライブからデータを取得し得る。集中型ストレージアーキテクチャでは、ドライブコントローラAとドライブコントロ

10

20

30

40

50

ーラBの両方が、ドライブに記憶されているデータにアクセスすることができ、ドライブコントローラA又はドライブコントローラBのいずれかが欠陥のある状態となった後で、欠陥のあるドライブコントローラのサービスは、他方のドライブコントローラによって引き継がれてよく、他方のドライブコントローラは、欠陥のあるドライブコントローラのサービスを分担し続ける。

【0040】

ドライブコントローラ及びドライブは、分散型ストレージアーキテクチャにあることがある。図2Bは、分散型ストレージアーキテクチャの例を示す。分散型ストレージアーキテクチャは、1つのドライブコントローラ、複数のドライブ、及びバックプレーンを含む。分散型ストレージアーキテクチャでは、データは複数のデータへと分割されてよく、各データに対応する確認コードが生成される。各データ及び対応する確認コードは、複数のドライブの各々に記憶される。例えば、ユーザAのビデオデータはビデオデータ1からビデオデータ3へと分割され、ビデオデータ1からビデオデータ3にそれぞれ対応する確認コード、確認コード1から確認コード3が生成される。次いで、ビデオデータ1及び確認コード1がドライブ1に記憶され、ビデオデータ2及び確認コード2がドライブ2に記憶され、ビデオデータ3及び確認コード3がドライブ3に記憶される。ドライブコントローラがビデオデータを読み取る必要があるとき、ドライブコントローラは、データ読み取り命令を送信し、バックプレーンを使用することによって命令をドライブ1からドライブ3にファンアウトし、ビデオデータ1からビデオデータ3及び確認コード1から確認コード3を3つのドライブから取得し得る。そして、ドライブコントローラは、ビデオデータ1からビデオデータ3及び確認コード1から確認コード3を処理し、例えば、組み合わせて、ビデオデータを最終的に取得する。

【0041】

以下は、添付の図面を参照して、本出願の実施形態において提供されるストレージシステムを詳しく説明する。

【0042】

図3Aは、本出願の一実施形態によるストレージシステム300の概略構造図である。ストレージシステム300は、少なくとも1つの制御デバイス301及び第1のコネクタ302を含む。第1のコネクタ302はドライブに取り外し可能に接続するように構成され、ドライブはデータを記憶するように構成される。少なくとも1つの制御デバイス301は、第1のコネクタ302を使用することによって、第1のコネクタ302に接続されるドライブと通信するように構成される。ドライブは、ソリッドステートディスク(solid state drives、SSD)、例えば、シリアルアドバンスドテクノロジーアタッチメント(serial advanced technology attachment、SATA)インターフェースを伴うSSD、シリアルアタッチトスモールコンピュータシステムインターフェース(serial attached small computer system interface、SAS)インターフェースを伴うSSD、又は不揮発性メモリエクスプレス(non-volatile memory express、NVME)インターフェースを伴うSSDであり得る。代替的に、ドライブは、ハードディスクドライブ(hard disk drive、HDD)、例えば、SATAインターフェースを伴うHDD又はSASインターフェースを伴うHDDであり得る。これは本明細書では限定されない。

【0043】

少なくとも1つの制御デバイス301が、第1のコネクタ302を使用することによって、第1のコネクタ302に接続されるドライブと通信することは、次のように理解され得ることに留意されたい。K個の第1のコネクタ302があり、 $K > 0$ であるとき、及びK個の第1のコネクタ302の各々がドライブに接続される場合、少なくとも1つの制御デバイス301は、K個の第1のコネクタ302を使用することによって、K個の第1のコネクタ302に接続されるドライブと通信する。K個の第1のコネクタ302のいくつかのみがドライブに接続される場合、少なくとも1つの制御デバイス301は、いくつか

10

20

30

40

50

の第1のコネクタ302を使用することによって、いくつかの第1のコネクタ302に接続されるドライブと通信する。

【0044】

前述の技術的解決法において、ドライブが第1のコネクタ302を使用することによって制御デバイス301に取り外し可能に接続され得るので、ユーザは、使用上の要件に基づいてドライブと制御デバイス301との間の接続関係をフレキシブルに構成して、異なるストレージアーキテクチャを構成し得る。例えば、1つの制御デバイス301が、第1のコネクタ302を使用することによって複数のドライブに接続されて、分散型ストレージアーキテクチャを構成する。代替的に、2つの制御デバイス301が、第1のコネクタ302を使用することによって1つのドライブに接続され、集中型ストレージアーキテクチャを構成する。このようにして、ストレージアーキテクチャが変更される必要があるとき、第1のコネクタ302とドライブとの間の接続関係が変更される必要があり、新しい制御デバイスと新しいドライブが購入される必要はない。これは、ストレージシステムのフレキシビリティを改善する。

10

【0045】

以下は、制御デバイス301及び第1のコネクタ302を説明する。

【0046】

ある例では、1つの制御デバイス301と1つの第1のコネクタ302がある。図3Aに示されるように、制御デバイス301及び第1のコネクタ302は2つの独立した構成要素であり得る。制御デバイス301及び第1のコネクタ302は、通信バスを使用することによって接続され得る。例えば、通信バスは、ペリフェラルコンポーネントインターフェースエクスプレス(peripheral component interface express、PCI E又はPCI e)バスであり得る。代替的に、図3Bに示されるように、制御デバイス301は第1のコネクタ302を含み得る。この場合、第1のコネクタ302は制御デバイス301の一部である。例えば、第1のコネクタ302は制御デバイス301のファンアウトコネクタであり得る。ファンアウトコネクタは、1つのノード又は複数のノード上の複数の信号をより広い空間範囲に広げるコネクタである。

20

【0047】

別の例では、複数の第1のコネクタ302があり得る。例えば、第1のコネクタ302は制御デバイス301の一部である。図3Cに示されるように、制御デバイス301は3つの第1のコネクタ302を含み、各第1のコネクタ302はドライブに接続するように構成され得る。

30

【0048】

ある例では、複数の制御デバイス301があり得る。図3Dを参照すると、ストレージシステム300は、M個の制御デバイス301及びN個の第1のコネクタ302を含み、N個の第1のコネクタ302はドライブに接続するように構成され、Mは2以上の整数であり、NはM以上の整数であり、Kは正の整数である。

【0049】

各制御デバイス301と第1のコネクタ302との間の構成関係が図3Aから図3Cに示されており、詳細は再び説明されない。図3Dにおいて、例えば、第1のコネクタ302は制御デバイス301の一部である。全ての制御デバイス301に含まれる第1のコネクタ302の数量は同じであり得る。例えば、M個の制御デバイス301の各々はS個の第1のコネクタ302を含み得る。代替的に、全ての制御デバイス301に含まれる第1のコネクタ302の数量は異なり得る。例えば、M個の制御デバイス301のうちの第1の制御デバイス301は3つの第1のコネクタ302を含み、M個の制御デバイス301のうちの第2の制御デバイス301は2つの第1のコネクタ302を含む、などである。図3Dでは、例えば、Mは3であり、3つの制御デバイス301の各々が2つの第1のコネクタ302を含む。

40

【0050】

以下は制御デバイス301を説明する。以下の説明では、例えば、第1のコネクタ30

50

2は制御デバイス301の一部である。

【0051】

ある例では、図4を参照すると、制御デバイス301はプロセッサ3011を含み得る。プロセッサ3011は、第1のコネクタ302に電氣的に接続される。プロセッサ3011は中央処理装置(central processing unit、CPU)、デジタルシグナルプロセッサ、特定用途向け集積回路、或いは、フィールドプログラマブルゲートアレイ(field-programmable gate array、FPGA)又は別のプログラマブルロジックデバイスであり得る。これは本明細書では限定されない。

【0052】

ある例では、プロセッサ3011は、高速信号ユニット30111及び制御ユニット30112を含み得る。高速信号ユニット30111は、プロセッサ3011とドライブとの間のサービスデータの送信のために構成される。例えば、プロセッサ3011は、高速信号ユニット30111を使用することによって、ドライブに記憶されているデータを取得し、又は、プロセッサ3011は、高速信号ユニット30111を使用することによって、クライアントから受信された記憶されるべきデータをドライブに記憶する。制御ユニット30112は、ドライブの状態を検出し、例えば、ドライブが第1のコネクタ302に接続されているか、又は切断されているかを検出し、ドライブが欠陥状態にあるかどうかを検出し、ドライブについての情報を検出し、ドライブの電源オン状態又は電源オフ状態を制御し、サービスデータを送信するために使用されるドライブのポートのリセット状態又はリセットデアサート状態を制御するように構成される。以下は、高速信号ユニット30111及び制御ユニット30112を別々に説明する。

【0053】

高速信号ユニット30111は、第1のコネクタ302に接続されるドライブの稼働モードがプロセッサ3011の現在の稼働モードと適合すると決定すると、第1のコネクタ302に電力を供給するように構成されてよく、そうすることによって、第1のコネクタ302に接続されるドライブの電源が入れられる。

【0054】

本出願のこの実施形態では、ドライブの稼働モードが高速信号ユニット30111の現在の稼働モードと適合するかどうかを決定する前に、高速信号ユニット30111はまず、第1のコネクタ302がドライブに接続されているか、又は切断されているかを決定し得る。例えば、さらに図4を参照すると、検出回路3021が第1のコネクタ302において配置されてよく、検出回路3021は、高速信号ユニット30111に電氣的に接続される。第1のコネクタ302がドライブから切断されるとき、検出回路3021は第1のレベル、例えば、ローレベルを高速信号ユニット30111に出力する。第1のコネクタ302がドライブに接続されるとき、検出回路3021は、第2のレベル、例えば、ハイレベルを高速信号ユニット30111に出力する。この場合、高速信号ユニット30111において決定条件が予め設定されていてよく、決定条件は次の通りであり得る。検出回路3021によって出力された第2のレベルを受信する場合、高速信号ユニット30111は、検出回路3021に対応する第1のコネクタ302がドライブに接続されると決定し、検出回路によって出力される第1のレベルを受信する場合、高速信号ユニット30111は、検出回路3021に対応する第1のコネクタ302がドライブから切断されていると決定する。例えば、検出回路3021によって出力されるハイレベルを受信するとき、高速信号ユニット30111は、検出回路3021に対応する第1のコネクタ302がドライブに接続されると決定する。

【0055】

複数の第1のコネクタ302がある場合、検出回路3021は、各第1のコネクタ302において配置されてよく、各第1のコネクタ302の検出回路3021によって出力されるレベルに基づいて、第1のコネクタ302がドライブに接続されているかどうかが決まされることに留意されたい。

【0056】

10

20

30

40

50

高速信号ユニット30111はさらに、検出回路3021によって出力されるレベルの変化に基づいて、第1のコネクタ302とドライブとの間の接続が緩いかどうかを決定し得る。例えば、高速信号ユニット30111が、0.5秒の間隔で、検出回路3021により出力されるレベルが第1のレベルから第2のレベルに変化すること、又は第2のレベルから第1のレベルに変化することを決定する場合、高速信号ユニット30111は、第1のコネクタ302とドライブとの間の接続が緩いと決定する。

【0057】

高速信号ユニット30111が、第1のコネクタ302がドライブに接続されていると決定する場合、高速信号ユニット30111はさらに、第1のコネクタ302を使用することによって、第1の情報を第1のコネクタ302に接続されるドライブに送信し得る。第1のコネクタ302に接続されるドライブに欠陥があるかどうかを検出するために、第1の情報が使用される。ドライブに欠陥がない場合、ドライブは、第1のコネクタ302を使用することによって、第1の情報に対応するフィードバック情報を高速信号ユニット30111に送信する。ドライブに欠陥がある場合、ドライブは、フィードバック情報を高速信号ユニット30111に送信することができない。第1の情報及びフィードバック情報は、ドライブと共に高速信号ユニット30111により事前に合意されていてもよい。そして、高速信号ユニット30111が、第1のコネクタ302に接続されるドライブに第1の情報を送信した後で第1のコネクタ302を使用することによって予め設定された時間内にフィードバック情報を受信しない場合、高速信号ユニット30111は、第1のコネクタ302に接続されるドライブに欠陥があると決定する。

【0058】

第1のコネクタ302がドライブに接続され、第1のコネクタ302に接続されるドライブに欠陥がないと、高速信号ユニット30111が決定する場合、高速信号ユニット30111は、第1のコネクタ302に接続されるドライブの稼働モードを取得し得る。

【0059】

本出願のこの実施形態では、高速信号ユニット30111の稼働モードは、第1の稼働モード及び第2の稼働モードを含む。第1の稼働モードは、プロセッサ集中型稼働モードとも呼ばれることがあり、図2Aに示される集中型ストレージアーキテクチャにおけるドライブコントローラと同じ処理モードで高速信号ユニット30111が動作することを示すことがある。例えば、クライアントからデータ書き込み命令を受信した後で、高速信号ユニット30111は、データ書き込み命令において搬送される記憶されるべきデータをドライブに送信する。第2の稼働モードは、プロセッサ分散型稼働モードとも呼ばれることがあり、図2Bに示される分散型ストレージアーキテクチャにおけるドライブコントローラと同じ処理モードで高速信号ユニット30111が動作することを示すことがある。例えば、クライアントからデータ書き込み命令を受信した後、制御デバイス301はまず、データ書き込み命令において搬送される記憶されるべきデータを複数のサブデータへと分割し、各サブデータに対応する確認コードを生成し、各サブデータ及びサブデータに対応する確認コードをドライブに送信する。

【0060】

ドライブの稼働モードは、第3の稼働モード及び第4の稼働モードを含む。第3の稼働モードは、ドライブ集中型稼働モードとも呼ばれることがあり、図2Aに示される集中型ストレージアーキテクチャにおけるドライブと同じ処理モード、例えば、ドライブ共有アクセスモードでドライブが動作することを示すことがある。具体的には、ドライブに接続される複数の制御デバイスが全て、ドライブに対して読み取り/書き込み動作を実行し得る。第4の稼働モードは、ドライブ分散型稼働モードとも呼ばれることがあり、図2Bに示される分散型ストレージアーキテクチャにおけるドライブと同じ処理モード、例えば、ドライブ独占稼働モードでドライブが動作することを示すことがある。具体的には、ドライブに接続される複数の制御デバイスのうちの1つだけが、ドライブに対して読み取り/書き込み動作を実行し得る。

【0061】

10

20

30

40

50

ある特定の実装処理では、高速信号ユニット30111の稼働モードが第1の稼働モードと第2の稼働モードのうちの1つだけを含むことが設定され得る。この場合、高速信号ユニット30111は固定された稼働モードで動作する。高速信号ユニット30111の稼働モードが第1の稼働モードに設定されるとき、高速信号ユニット30111は第1の稼働モードだけで動作することができる。例えば、高速信号ユニット30111は、ユーザの命令に従って、稼働モードのうちの1つで動作することを決定し得る。ある例では、少なくとも1つの稼働モードに対応するシステムコードが、高速信号ユニット30111において設定され得る。例えば、第1の稼働モードに対応する第1のシステムコード及び第2の稼働モードに対応する第2のシステムコードのうちの少なくとも1つが設定され得る。プロセッサ3011の電源が入れられた後で、高速信号ユニット30111は、制御ユニット30112に記憶されている対応するシステムコードを読み取って実行し、そうすることによって、高速信号ユニット30111は、システムコードに対応する稼働モードに入ることができる。システムコードは、異なる稼働モードにおける、ストレージシステムの高信頼性フォルトトレランスポリシー、ドライブの稼働モード、並びに、ドライブのタイプ、製造業者、及びモデルなどの情報を含む。

10

【0062】

本出願のこの実施形態では、高速信号ユニット30111は、第1のコネクタ302に接続されるドライブの稼働モードがプロセッサ3011の現在の稼働モードと適合するかどうかを決定する。決定の方式は、限定はされないが、以下の2つの方式を含む。

【0063】

20

第1の方式では、高速信号ユニット30111がまず、制御ユニット30112を使用することによってドライブについての情報を取得してよく、ドライブについての情報は、ドライブのモデル、容量、及び製造業者などの、少なくとも1つのタイプの関連情報を含む。高速信号ユニット30111は、ドライブについての情報とドライブの稼働モードとの間の対応付けを予め記憶する。この場合、高速信号ユニット30111は、ドライブについての情報とドライブの稼働モードとの間の対応付けに基づいて、ドライブの稼働モードを決定することができる。例えば、ドライブについての情報とドライブの稼働モードとの間の対応付けは、製造業者Aと第3の稼働モードとの間の第1の対応付け、及び製造業者Bと第4の稼働モードとの間の第2の対応付けを含む。制御ユニット30112を使用することによって高速信号ユニット30111により取得されるドライブについての情報が、ドライブの製造業者が製造業者Aであることを示す場合、高速信号ユニット30111は、ドライブの稼働モードが第3の稼働モードであると決定し得る。次いで、高速信号ユニット30111は、ドライブの稼働モードが高速信号ユニット30111の現在の稼働モードと適合するかどうかを決定し得る。

30

【0064】

第2の方式では、高速信号ユニット30111が高速信号ユニット30111の現在の稼働モードを決定する場合、高速信号ユニット30111は対応するドライブ決定条件を設定し得る。ドライブ決定条件は、ドライブの稼働モードが高速信号ユニット30111の現在の稼働モードと適合するとき、ドライブについての情報を含む。ドライブについての情報は第1の方式における対応するコンテンツと同じであり、再び説明されない。例えば、高速信号ユニット30111の現在の稼働モードが第1の稼働モードである場合、ドライブ決定条件は製造業者Aに設定される。次いで、ドライブについての情報がドライブ決定条件と同じであるとき、高速信号ユニット30111は、ドライブの稼働モードが高速信号ユニット30111の現在の稼働モードと適合すると決定する。ドライブについての情報がドライブ決定条件と異なる場合、高速信号ユニット30111は、ドライブの稼働モードが高速信号ユニット30111の現在の稼働モードと適合しないと決定する。

40

【0065】

高速信号ユニット30111の稼働モードがドライブの稼働モードと適合するということは、高速信号ユニット30111の稼働モードが第1の稼働モードであるとき、ドライブの稼働モードが第3の稼働モードであり、又は、高速信号ユニット30111の稼働モ

50

ードが第2の稼働モードであるとき、ドライブの稼働モードが第4の稼働モードであることを意味することに留意されたい。もちろん、高速信号ユニット30111及びドライブの稼働モードがさらに別の稼働モードを含む場合、高速信号ユニット30111の稼働モードがドライブの稼働モードと適合することは、別の対応付けをさらに含み得る。これは本明細書では限定されない。

【0066】

加えて、高速信号ユニット30111が制御ユニット30112を使用することによってドライブについての情報を取得することは、次の通りであり得ることに留意されたい。ドライブについての情報がドライブの中の固定されたストレージ空間に予め記憶され、ドライブが第1のコネクタ302に接続されるとき、高速信号ユニット30111が、制御

10

【0067】

稼働モードが高速信号ユニット30111の現在の稼働モードと適合するドライブを決定した後で、高速信号ユニット30111は、ドライブに接続される第1のコネクタ302に電力を供給する。図4に示されるように、3つの第1のコネクタ302のうちの1番目の第1のコネクタ302及び3番目の第1のコネクタ302がそれぞれドライブA及びドライブBに接続されることと、1番目の第1のコネクタ302に接続されるドライブの稼働モードが高速信号ユニット30111の現在の稼働モードと適合することとを、高速

20

【0068】

稼働モードが高速信号ユニット30111の現在の稼働モードと適合しないドライブを決定した後で、高速信号ユニット30111は、ドライブに接続される第1のコネクタ302に電力を供給することを停止する。図4に示されるように、3つの第1のコネクタ302のうちの1番目の第1のコネクタ302及び3番目の第1のコネクタ302がそれぞれドライブA及びドライブBに接続されることと、3番目の第1のコネクタ302に接続されるドライブの稼働モードが高速信号ユニット30111の現在の稼働モードと適合しないこととを、高速信号ユニット30111が決定するとき、高速信号ユニット30111は、3番目の第1のコネクタ302に電力を供給することを停止する。

30

【0069】

高速信号ユニット30111が第1のコネクタ302に電力を供給することは、次のことを意味することに留意されたい。ドライブに接続される第1のコネクタ302の電源が現在入っている場合、高速信号ユニット30111は、制御ユニット30112の電力供給ステータスを維持することによって第1のコネクタ302に電力を供給し続ける。ドライブに接続される第1のコネクタ302の電源が現在入っていない場合、高速信号ユニット30111は、制御ユニット30112の電力供給ステータスを変更することによって第1のコネクタ302に電力を供給する。高速信号ユニット30111が第1のコネクタ302に電力を供給することを停止することは、次のことを意味することに留意されたい。ドライブに接続される第1のコネクタ302の電源が現在入っている場合、高速信号ユ

40

【0070】

高速信号ユニット30111が電力を少なくとも1つの第1のコネクタ302に供給した後で、例えば、高速信号ユニット30111がM個の第1のコネクタ302に電力を供給した後で、高速信号ユニット30111は、M個の第1のコネクタ302のうちM1個の第1のコネクタ302を使用することによって、M1個の第1のコネクタ302に接

50

続されるドライブのうちの1つ又は複数にデータ読み取り命令を送信し、M1個の第1のコネクタ302を使用することによって、データ読み取り命令に対応するデータを取得してよく、M1は0より大きくM以下の整数である。代替的に、高速信号ユニット3011は、M1個の第1のコネクタ302に接続されるドライブのうちの1つ又は複数にデータ書き込み命令を送信して、データ書き込み命令において搬送される記憶されるべきデータをドライブに記憶し得る。

【0071】

高速信号ユニット3011が、M個の第1のコネクタ302のうちのM1個の第1のコネクタ302を使用することによって、M1個の第1のコネクタ302に接続されるドライブのうちの1つ又は複数にデータ読み取り命令又はデータ書き込み命令を送信した後で、高速信号ユニット3011はさらに、M個の第1のコネクタ302の各々がドライブに接続されるかどうかを検出し続け得る。第1のコネクタ302がドライブから切断されていることを高速信号ユニット3011が決定する場合、高速信号ユニット3011は、第1のコネクタ302に電力を供給することを停止し得る。

10

【0072】

ある例では、制御デバイス301は、複数の個別の構成要素を含み得る。図5を参照すると、制御デバイス301は、プロセッサ、メモリ、クロック信号生成器、及び電力供給モジュールを含み得る。プロセッサは、例えば、中央処理装置(central processing unit、CPU)、FPGA、又は別のプログラマブルロジックデバイスである。これは本明細書では限定されない。プロセッサがCPUである例が以下の説明において使用される。

20

【0073】

制御デバイス301のCPUは、第1のコネクタ302を使用することによって、第1のコネクタ302に接続されるドライブと通信するように構成される。例えば、CPUは、ドライブ送信プロトコルを使用することによってドライブと通信し得る。ドライブ送信プロトコルは、例えば、PCIeプロトコル、シリアルATAインターフェースアタッチメント(serial advanced technology attachment、SATA)プロトコル、又はシリアルアタッチトスモールコンピュータシステムインターフェース(serial attached small computer system interface、SAS)プロトコルである。複数の第1のコネクタ302がある場合、CPUは、複数の第1のコネクタ302の全てを使用することによって、全ての第1のコネクタ302に接続されるドライブと通信する。代替的に、CPUは、複数の第1のコネクタ302のうちのいくつかを使用することによって、いくつかの第1のコネクタ302に接続されるドライブと通信し得る。CPUとドライブとの間の通信は、別のデバイス(例えば、クライアント)から記憶されるべきデータを受信した後で、CPUによって、対応するドライブに記憶されるべきデータを記憶すること、又は、別のデバイスからデータ読み取り命令を受信した後で、ドライブからCPUによって、データ読み取り命令に対応するデータを取得することを含み得る。

30

【0074】

制御デバイス301のメモリは、CPUに接続され、CPUとドライブとの間の通信の間にCPUのためのキャッシュ機能を提供するように構成される。例えば、クライアントから記憶されるべきデータを受信した後で、CPUはまず、メモリに記憶されるべきデータを記憶し、次いで、対応するドライブに記憶されるべきデータを記憶するためにデータ書き込み命令をドライブに送信する。

40

【0075】

制御デバイス301のクロック信号生成器は、CPU及び第1のコネクタ302に別々に接続され、CPU及び第1のコネクタ302に接続されるドライブのためのクロックを提供するように構成される。複数の第1のコネクタ302があるとき、クロック信号生成器は各第1のコネクタ302に接続されることに留意されたい。クロックは、ドライブ送信プロトコルを使用することによってCPUとドライブとの間の通信に必要とされる第1

50

のクロックを含み、及び/又は、CPU及びドライブの動作クロックと、ローカルの位相ロックループクロックとを含む。第1のクロックは、代替的に、高速クロック、又はデータをサンプリングするためにドライブ303によって使用される位相ロックループクロックと呼ばれることもある。ドライブ303によって使用される送信プロトコルが、動作クロック又はローカルの位相ロックループクロックをサポートする場合、クロック信号生成器は、第1のクロックを提供する必要はないことがある。ドライブ送信プロトコルは、例えば、PCI Eプロトコル、シリアルATAインターフェースアタッチメント(serial advanced technology attachment、SATA)プロトコル、又はシリアルアタッチトスモールコンピュータシステムインターフェース(serial attached small computer system interface、SAS)プロトコルである。

10

【0076】

制御デバイス301の電力供給モジュールは、CPU及び第1のコネクタ302に別々に接続され、CPU及び第1のコネクタ302に接続されるドライブの動作電圧を提供するように構成される。複数の第1のコネクタ302があるとき、電力供給モジュールは各第1のコネクタ302に接続されることに留意されたい。

【0077】

ある例では、図5に示されるように、制御デバイス301はさらに、クライアントに接続するように構成される入力/出力インターフェースを含む。入力/出力インターフェースは、CPUに電氣的に接続され、クライアントからデータ読み取り命令又はデータ書き込み命令を受信し、次いで、データ読み取り命令又はデータ書き込み命令をCPUに送信するように構成される。入力/出力インターフェースがクライアントからデータ読み取り命令を取得する場合、入力/出力インターフェースはさらに、クライアントに、データ読み取り命令に対応しCPUによって送信されるデータを送信するように構成される。入力/出力インターフェースは、SAS、PCI Eインターフェース、ギガビットイーサネットインターフェース(gigabit Ethernet、GE)、又は別のインターフェースであり得る。これは本明細書では限定されない。

20

【0078】

ある例では、CPUは、第1のコネクタ302に接続されるドライブの稼働モードがCPUの現在の稼働モードと適合すると決定するとき、第1のコネクタ302に電力を供給するように電力供給モジュールを制御するように構成されてよく、そうすることによって、第1のコネクタ302に接続されるドライブの電源が入れられる。

30

【0079】

本出願のこの実施形態では、ドライブの稼働モードがCPUの現在の稼働モードと適合するかどうかを決定する前に、CPUはまず、第1のコネクタ302がドライブに接続されているか、又は切断されているかを決定し得る。図6を参照すると、この場合、制御デバイス301はさらに検出構成要素を含み得る。検出構成要素は、CPU、第1のコネクタ302、及び電力供給モジュールに別々に接続され、電力供給モジュールはさらに、検出構成要素の動作電圧を提供するように構成される。検出構成要素は、第1のコネクタ302がドライブに接続されているか、又は切断されているかを検出するように構成され得る。複数の第1のコネクタ302があるとき、検出構成要素は各第1のコネクタ302に接続されることに留意されたい。例えば、検出構成要素は、コンプレックスプログラマブルロジックデバイス(complex programmable logic device、CPLD)又は特定用途向け集積回路(application specific integrated circuit、ASIC)であり得る。図6では、検出構成要素がCPLDである例が使用される。

40

【0080】

図6では、CPLDはCPUに直接接続されることに留意されたい。別の例では、CPLDは、代替的に、CPUの周辺ポート拡張チップを使用することによってCPUに接続され得る。図7において、CPUの周辺ポート拡張チップが集積サウスブリッジ(pla

50

t f o r m c o n t r o l l e r h u b、P C H)である例が使用され、P C HとC P L Dはローピンカウント(low pin count、L P C)を使用することによって接続され得る。

【 0 0 8 1 】

さらに図 6 を参照すると、検出回路 3 0 2 1 がさらに第 1 のコネクタ 3 0 2 に配置され得る。第 1 のコネクタ 3 0 2 がドライブから切断されているとき、検出回路 3 0 2 1 は、第 1 のレベル、例えば、ローレベルを C P L D に出力する。第 1 のコネクタ 3 0 2 がドライブに接続されているとき、検出回路 3 0 2 1 は、第 2 のレベル、例えば、ハイレベルを C P L D に出力する。検出回路 3 0 2 1 によって出力される第 2 のレベルが受信される場合、C P L D は、検出回路 3 0 2 1 に対応する第 1 のコネクタ 3 0 2 がドライブに接続されると決定し、検出回路によって出力される第 1 のレベルが受信される場合、C P L D は、検出回路 3 0 2 1 に対応する第 1 のコネクタ 3 0 2 がドライブから切断されていると決定する。例えば、検出回路 3 0 2 1 によって出力されるハイレベルを受信するとき、C P L D は、検出回路 3 0 2 1 に対応する第 1 のコネクタ 3 0 2 がドライブに接続されると決定する。

10

【 0 0 8 2 】

第 1 のコネクタ 3 0 2 がドライブに接続されているか、又は切断されているかを C P L D が決定した後で、C P U は、複数の方式で C P L D の決定結果を取得してよく、決定結果は、第 1 のコネクタ 3 0 2 がドライブに接続されているか、又は切断されているかを示すために使用される。例えば、C P U が C P L D の決定結果を取得する方式は、限定はされないが、以下の 2 つの方式を含む。

20

【 0 0 8 3 】

第 1 の取得方式では、C P L D は決定結果を C P U に送信し得る。複数の第 1 のコネクタ 3 0 2 がある場合、C P L D が決定結果を C P U に送信するとき、第 1 のコネクタ 3 0 2 の識別子が搬送され得る。例えば、複数の第 1 のコネクタ 3 0 2 に対して数字が設定されてよく、C P L D は、第 1 のコネクタ 3 0 2 の数字を決定結果に追加する。

【 0 0 8 4 】

第 2 の取得方式では、C P L D は、決定結果を記憶するように構成されるレジスタを設定してよく、例えば、レジスタは第 1 のレジスタと呼ばれる。C P L D は、第 1 のレジスタに決定結果を記憶してよく、例えば、決定結果を表すために 1 ビットを使用してよい。1 ビットの値が 0 である場合、それは、第 1 のコネクタ 3 0 2 がドライブに接続されていることを示し、1 ビットの値が 1 である場合、それは、第 1 のコネクタ 3 0 2 がドライブから切断されていることを示す。C P U が決定結果を取得する必要がある場合、C P U は、第 1 のレジスタの値を直接読み取り、第 1 のレジスタの値と、第 1 のレジスタにおけるビットの値と接続ステータスとの間の予め設定された対応付けとに基づいて、第 1 のコネクタ 3 0 2 がドライブに接続されているか、又は切断されているかを決定する。第 1 のレジスタの値と接続ステータスとの間の対応付けは次の通りであり得る。第 1 のレジスタのビットの値が 0 である場合、それは、第 1 のコネクタ 3 0 2 がドライブに接続されていることを示し、第 1 のレジスタのビットの値が 1 である場合、それは、第 1 のコネクタ 3 0 2 がドライブから切断されていることを示す。

30

40

【 0 0 8 5 】

第 2 の方式において、複数の第 1 のコネクタ 3 0 2 があるとき、C P L D は、各第 1 のコネクタ 3 0 2 に対する対応する第 1 のレジスタを設定し得ることに留意されたい。C P U は、各第 1 のコネクタ 3 0 2 に対応する第 1 のレジスタの値を読み取り、各第 1 のコネクタ 3 0 2 がドライブから切断されているか、又は接続されているかを決定し得る。代替的に、C P L D は、1 つだけの第 1 のレジスタを設定し、各第 1 のコネクタ 3 0 2 に対応する決定結果を第 1 のレジスタに記憶してよい。具体的には、第 1 のレジスタは複数のビットを含み、1 ビットが、1 つの第 1 のコネクタ 3 0 2 に対応する決定結果を示す。図 6 に示されるように、制御デバイス 3 0 1 は、3 つの第 1 のコネクタ 3 0 2 を含む。C P L D は 1 つの第 1 のレジスタを設定し、全ての第 1 のコネクタ 3 0 2 に対応する決定結果は

50

、第1のレジスタの3ビットを使用することによって示され、第1のレジスタの1ビットは、1つの第1のコネクタ302に対応する決定結果を示すために使用される。例えば、レジスタのビットの値が「001」である場合、それは、3番目の第1のコネクタ302がドライブから切断されており、残りの2つの第1のコネクタ302がドライブに別々に接続されていることを示す。レジスタのビットの値が「000」である場合、それは、3番目の第1のコネクタ302の各々がドライブから切断されていることを示す。

【0086】

CPLDはさらに、第1のコネクタ302を使用することによって、第1のコネクタ302に接続されるドライブに第1の情報を送信し得る。第1の情報は、第1のコネクタ302に接続されるドライブに欠陥があるかどうかを検出するために使用される。ドライブに欠陥がない場合、ドライブは、第1のコネクタ302を使用することによって、第1の情報に対応するフィードバック情報をCPLDに送信する。ドライブに欠陥がある場合、ドライブは、フィードバック情報をCPLDに送信することができない。第1の情報及びフィードバック情報は、CPLDとドライブにより事前に合意され得る。そして、CPLDが第1のコネクタ302に接続されるドライブに第1の情報を送信した後で第1のコネクタ302を使用することによって予め設定された時間内にフィードバック情報を受信しない場合、CPLDは、第1のコネクタ302に接続されるドライブに欠陥があると決定する。それに対応して、CPLDは、検出されるドライブに欠陥があるかどうかについての検出結果も、対応するレジスタ、例えば、第2のレジスタに記憶し得る。

【0087】

CPUが、CPLDの対応するレジスタのビットの値に基づいて、第1のコネクタ302がドライブに接続されており、第1のコネクタ302に接続されるドライブに欠陥がないと決定する場合、CPUは、第1のコネクタ302に接続されるドライブの稼働モードを取得し得る。

【0088】

本出願のこの実施形態では、ドライブの稼働モード及びCPUの稼働モードは上で説明されており、詳細は本明細書では再び説明されない。CPUは、第1のコネクタ302に接続されるドライブについての情報を取得し、ドライブの稼働モードがCPUの稼働モードと適合するかどうかを決定する。具体的な決定方式は既に前に説明されており、本明細書では再び説明されない。

【0089】

本出願のこの実施形態では、CPUがドライブについての情報を取得する方式は、限定はされないが、以下の2つの方式を含む。

【0090】

ドライブについての情報を取得するための第1の方式では、CPUは、第1のコネクタ302に接続されるバスを使用することによって、第1のコネクタ302に接続されるドライブに、ドライブについての情報を取得するために使用される命令を送信し得る。命令を受信した後で、ドライブは、ドライブについての情報をCPUに送信する。この場合、CPUと第1のコネクタ302を接続するバスは、カスタマイズされたバス、又はPCIバスなどの高速信号を送信するように構成されるバスであり得る。高速信号は、CPUがドライブ送信プロトコルを使用することによってドライブと通信するときを使用される対応する信号である。これは本明細書では限定されない。

【0091】

ドライブについての情報を取得するための第2の方式では、図7に示されるように、制御デバイス301はさらにPCIスイッチ(switch)を含む。PCIスイッチは、PCIバスを使用することによってCPUに接続され、PCIスイッチは、I2Cバスを使用することによって第1のコネクタ302に接続される。CPUは、PCIスイッチを使用することによって、第1のコネクタ302に接続されるドライブに、ドライブについての情報を取得するために使用される命令を送信する。命令を受信した後で、ドライブはドライブについての情報をCPUに送信する。

【 0 0 9 2 】

稼働モードがCPUの現在の稼働モードと適合するドライブを決定した後で、CPUは、CP LDに、ドライブに接続される第1のコネクタ302に電力を供給するための命令を送信し、そうすることによって、CP LDは、第1のコネクタ302に電力を供給するように電力供給モジュールを制御する。稼働モードがCPUの現在の稼働モードと適合しないドライブを決定した後で、CPUは、CP LDに、ドライブに接続される第1のコネクタ302に電力を供給することを停止するための命令を送信し、そうすることによって、CP LDは、第1のコネクタ302に電力を供給することを停止するように電力供給モジュールを制御する。

【 0 0 9 3 】

CP LDが第1のコネクタ302に電力を供給するように電力供給モジュールを制御し、CP LDが第1のコネクタ302に電力を供給することを停止するように電力供給モジュールを制御し、これらは上で既に説明されており、本明細書では再び説明されないことに留意されたい。

【 0 0 9 4 】

CPUが少なくとも1つの第1のコネクタ302に電力を供給した後で、CPUは、少なくとも1つの第1のコネクタ302を使用することによってドライブと通信し得る。加えて、CPUはまた、CP LDを使用することによってドライブにリセット命令又はリセットデアサート命令を送信して、リセット状態又はリセットデアサート状態となるように、サービスデータを送信するために使用されるドライブのポートを制御し得る。

【 0 0 9 5 】

CP LDはさらに、第1のコネクタ302に接続されるドライブにCPUによって送信される検出されたデータ読み取り命令又は検出されたデータ書き込み命令に基づいて、第1のコネクタ302に接続されるドライブの稼働状態を決定し得る。例えば、CP LDが、第1のコネクタ302に接続されるドライブにCPUによって送信されるデータ読み取り命令を検出する場合、CP LDは、ドライブが読み取り稼働状態にあると決定する。CP LDが、第1のコネクタ302に接続されるドライブにCPUによって送信されるデータ書き込み命令を検出する場合、CP LDは、ドライブが書き込み稼働状態にあると決定する。もちろん、CP LDは、ドライブの稼働状態を記憶するように構成されるレジスタも設定し得る。具体的な方式は、検出されるドライブに欠陥があるかどうかについての検出結果を対応するレジスタに記憶する前述の方式と同じであり、本明細書では再び説明されない。

【 0 0 9 6 】

CP LDはさらに、第1のコネクタ302とドライブとの間の接続が接続状態から切断状態に変化したと決定した後で、又は、第1のコネクタ302に接続されるドライブの稼働状態などの情報が変化した後で、PCIEスイッチに割り込み信号を送信してよく、PCIEスイッチは割り込み信号をCPUに送信する。割り込み信号を受信した後で、CPU処理が変化する。

【 0 0 9 7 】

さらに図4を参照すると、本出願のこの実施形態において提供されるストレージシステム300は、ドライブ303をさらに含み得る。

【 0 0 9 8 】

本出願のこの実施形態では、ドライブ303は、限定はされないが、次の3つのタイプを含み得る。第1のドライブタイプは、集中型ストレージアーキテクチャのみにおいて使用されることができ得るドライブであり、集中型ドライブと呼ばれ得る。第2のドライブタイプは、分散型ストレージアーキテクチャのみにおいて使用されることができ得るドライブであり、分散型ドライブと呼ばれ得る。第3のドライブタイプは、集中型ストレージアーキテクチャと分散型ストレージアーキテクチャの両方において使用されることができ得るドライブであり、汎用ドライブと呼ばれ得る。3つのタイプのドライブが、以下で別々に説明される。

10

20

30

40

50

【 0 0 9 9 】

第 1 のドライブ / 分散型ドライブ :

図 8 A を参照すると、ドライブ 3 0 3 は、第 2 のコネクタ 8 0 1、存在信号出力モジュール 8 0 2、電源変換モジュール 8 0 3、内部ドライブ制御モジュール 8 0 4、及びデータストレージモジュール 8 0 5 を含む。

【 0 1 0 0 】

第 2 のコネクタ 8 0 1 は、制御デバイス 3 0 1 に接続し、制御デバイス 3 0 1 とドライブ 3 0 3 との間のデータ通信のためのチャンネルを提供するように構成される。例えば、第 1 のコネクタ 3 0 2 に接続された後で、ドライブ 3 0 3 は、第 2 のコネクタ 8 0 1 を使用することによって制御デバイス 3 0 1 からデータ読み取り命令を取得し、次いで、第 2 のコネクタ 8 0 1 を使用することによって、データ読み取り命令に対応するデータを制御デバイス 3 0 1 に送信する。

10

【 0 1 0 1 】

存在信号出力モジュール 8 0 2 は、ドライブ 3 0 3 の存在信号を生成するように構成され、存在信号は、ドライブ 3 0 3 が第 1 のコネクタ 3 0 2 に接続されているか、又は切断されているかを示すために使用される。ある例では、制御デバイス 3 0 1 は、周辺回路及び論理決定モジュールを含み得る。ドライブ 3 0 3 が第 2 のコネクタ 8 0 1 を使用することによって第 1 のコネクタ 3 0 2 に接続されるとき、存在信号出力モジュール 8 0 2 は、第 2 のコネクタ 8 0 1 を使用することによって制御デバイス 3 0 1 の電力供給モジュールのグラウンドケーブルに接続され、そうすることによって、制御デバイス 3 0 1 の周辺回路はローレベル信号を生成し、制御デバイス 3 0 1 の論理決定モジュールは、ローレベル信号に基づいて、ドライブ 3 0 3 が第 1 のコネクタ 3 0 2 に接続されることを決定する。ドライブ 3 0 3 の第 2 のコネクタ 8 0 1 が第 1 のコネクタ 3 0 2 から切断された後で、制御デバイス 3 0 1 の周辺回路がハイレベル信号を生成し、そうすることによって、各制御デバイス 3 0 1 の論理決定モジュールは、ハイレベル信号に基づいて、ドライブ 3 0 3 が第 1 のコネクタ 3 0 2 から切断されていると決定する。

20

【 0 1 0 2 】

電源変換モジュール 8 0 3 は、電力をドライブ 3 0 3 に供給し、第 2 のコネクタ 8 0 1 が第 1 のコネクタ 3 0 2 に接続されるとき、第 2 のコネクタ 8 0 1 からの電源電圧入力を内部ドライブ制御モジュール 8 0 4 及びデータストレージモジュール 8 0 5 の電力供給電圧へと変換し、変換された電圧を内部ドライブ制御モジュール 8 0 4 及びデータストレージモジュール 8 0 5 に出力するように構成される。例えば、電力変換モジュール 8 0 3 は、バック回路 (Buck circuit) などであり得る。

30

【 0 1 0 3 】

ドライブ 3 0 3 はさらに、制御デバイス 3 0 1 が短絡されることにより、又は制御デバイス 3 0 1 によって出力される電圧がドライブ 3 0 3 の普通の動作電圧より低いことにより引き起こされる、ドライブ 3 0 3 への異常な電力供給を避けるために、第 2 のコネクタ 8 0 1 と電力変換モジュール 8 0 3 との間に配置される逆流防止モジュールを含み得る。

【 0 1 0 4 】

内部ドライブ制御モジュール 8 0 4 は、制御デバイス 3 0 1 の信号を処理してドライブ 3 0 3 を管理するという機能を実装するように、例えば、ドライブ 3 0 3 に対する制御デバイス 3 0 1 の制御信号を受信し、データストレージモジュール 8 0 5 上のデータに対する読み取り動作又は書き込み動作を実行するという機能を提供するように構成される。例えば、制御信号は、ドライブ 3 0 3 のモデル、容量、及び製造業者などの関連情報を読み取るために使用される信号、リセット状態にとどまるように第 2 のコネクタ 8 0 1 を制御するために使用される信号、及びリセット状態をデアサートするように第 2 のコネクタ 8 0 1 を制御するために使用される信号である。

40

【 0 1 0 5 】

ある例では、図 8 B に示されるように、ドライブ 3 0 3 はさらに第 1 のストレージモジュール 8 0 6 を含んでもよく、第 1 のストレージモジュール 8 0 6 は、ドライブ 3 0 3 の

50

モデル、容量、及び製造業者などの関連情報を記憶するように構成される。制御デバイス 301 は、ドライブ 303 のモデル、容量、及び製造業者などの関連情報を第 1 のストレージモジュール 806 から直接読み取り得る。例えば、第 1 のストレージモジュール 806 は、電氣的消去可能プログラム可能読み取り専用メモリ (electrically erasable programmable read only memory、EEPROM) チップ及び I2C 回路を含み得る。

【0106】

ドライブの稼働状態を表示するために使用されるインジケータがドライブ 303 に配置される場合、内部ドライブ制御モジュール 804 はさらに、ドライブ 303 のインジケータを制御するように構成され得る。例えば、ドライブ 303 がデータ読み取り状態にあるとき、入力信号インジケータは、第 1 の周波数のインジケーション信号を出力するように制御されてよく、ドライブ 303 がデータ書き込み状態にあるとき、入力信号インジケータは、第 2 の周波数のインジケーション信号を出力するように制御されてよい、などである。

10

【0107】

加えて、高速信号インターフェース及び低速信号インターフェースが内部ドライブ制御モジュール 804 において設定され、内部ドライブ制御モジュール 804 は、異なるインターフェースを通じて制御デバイス 301 と通信し得る。高速信号の定義は上で説明されている。内部ドライブ制御モジュール 804 と制御デバイス 301 との間で送信される通信信号において、高速信号以外の信号は低速信号である。

20

【0108】

もちろん、内部ドライブ制御モジュール 804 はさらに他の機能を有することがあり、本明細書では 1 つずつ例は挙げられない。当業者は使用上の要件に基づいて内部ドライブ制御モジュール 804 の機能を設定し得ることに、留意されたい。これは本明細書では限定されない。

【0109】

ある例では、内部ドライブ制御モジュール 804 は、1 つの個別のデバイス又は複数の個別のデバイスの組合せを使用することによって実装されてよく、例えば、中央処理装置 (central processing unit、CPU)、デジタルシグナルプロセッサ、特定用途向け集積回路、フィールドプログラマブルゲートアレイ (field-programmable gate array、FPGA)、又は別のプログラマブルロジックデバイスであってよい。これは本明細書では限定されない。

30

【0110】

データストレージモジュール 805 は、内部ドライブ制御モジュール 804 の制御下でユーザデータを記憶し、又は読み取るように構成される。ある例では、データストレージモジュールはドライブ媒体であり得る。

【0111】

第 2 のドライブ / 集中型ドライブ:

図 9A を参照すると、ドライブ 303 は、2 つの第 2 のコネクタ (これらは第 2 のコネクタ 901 及び第 2 のコネクタ 902 である)、存在信号出力モジュール 903、電源変換モジュール 904、内部ドライブ制御モジュール 905、及びデータストレージモジュール 906 を含む。

40

【0112】

第 2 のコネクタ 901 及び第 2 のコネクタ 902 の機能はアクセスポート 801 の機能と同じであり、詳細は本明細書では再び説明されない。第 2 のコネクタ 901 及び第 2 のコネクタ 902 は、2 つの独立したアクセスポートであり、2 つの制御デバイス 301 にそれぞれ接続され得る。このようにして、2 つの制御デバイス 301 が第 2 のコネクタ 901 及び第 2 のコネクタ 902 にそれぞれ接続された後で、2 つの制御デバイス 301 は、両方ともドライブ 303 と通信し、例えば、ドライブ 303 のデータストレージモジュール 906 の中のデータを読み取り得る。

50

【 0 1 1 3 】

図 9 A において、ドライブ 3 0 3 は、第 2 のコネクタ 9 0 1 及び第 2 のコネクタ 9 0 2 という 2 つのアクセスポートを含むことに留意されたい。ドライブ 3 0 3 のアクセスポートの数量はまた 2 より多くてもよい。例えば、ドライブ 3 0 3 は、第 2 のコネクタ、第 2 のコネクタ、及び第 3 のアクセスポートを含む。この場合、ドライブ 3 0 3 は、3 つの制御デバイス 3 0 1 に同時に接続され得る。アクセスポートの数量は本出願のこの実施形態では限定されない。

【 0 1 1 4 】

第 2 のコネクタ 9 0 1 及び第 2 のコネクタ 9 0 2 が 2 つの制御デバイス 3 0 1 に同時に接続される必要がある場合、第 2 のコネクタ 9 0 1 と第 2 のコネクタ 9 0 2 との間で十分な離隔が確保される必要があり、この離隔は 2 つの制御デバイス 3 0 1 間の間隔に関係することに留意されたい。例えば、図 9 B に示されるように、2 つの制御デバイス 3 0 1 は両方とも水平に配置され、一方の制御デバイス 3 0 1 が他方の制御デバイス 3 0 1 の右上にある。2 つの制御デバイス 3 0 1 間の間隔が x であり、第 1 のコネクタ 3 0 2 の幅が h である場合、第 2 のコネクタ 9 0 1 と第 2 のコネクタ 9 0 2 との間の離隔は $(x - h)$ 以上である。

10

【 0 1 1 5 】

存在信号出力モジュール 9 0 3 は、ドライブ 3 0 3 の存在信号を生成するように構成され、存在信号は、ドライブ 3 0 3 が第 1 のコネクタ 3 0 2 に接続されているか、又は切断されているかを示すために使用される。ある例では、制御デバイス 3 0 1 は、周辺回路及び論理決定モジュールを含み得る。ドライブ 3 0 3 が第 2 のコネクタ 9 0 1 及び第 2 のコネクタ 9 0 2 を使用することによってそれぞれ 2 つの第 1 のコネクタ 3 0 2 に接続されるとき、存在信号出力モジュール 9 0 3 が第 2 のコネクタ 9 0 1 及び第 2 のコネクタ 9 0 2 を使用することによってそれぞれ 2 つの制御デバイス 3 0 1 の電力供給モジュールのグラウンドケーブルに接続され、そうすることによって、制御デバイス 3 0 1 の周辺回路はローレベル信号を生成し、各制御デバイス 3 0 1 の論理決定モジュールは、ローレベル信号に基づいて、ドライブ 3 0 3 が第 1 のコネクタ 3 0 2 に接続されることを決定する。ドライブ 3 0 3 の第 2 のコネクタ 9 0 1 及び第 2 のコネクタ 9 0 2 がそれぞれ 2 つの制御デバイス 3 0 1 の第 1 のコネクタ 3 0 2 から切断された後で、各制御デバイス 3 0 1 の周辺回路はハイレベル信号を生成し、そうすることによって、各制御デバイス 3 0 1 の論理決定モジュールは、ハイレベル信号に基づいて、ドライブ 3 0 3 が第 1 のコネクタ 3 0 2 から切断されていると決定する。

20

30

【 0 1 1 6 】

電源変換モジュール 9 0 4 は、電力をドライブ 3 0 3 に供給し、第 2 のコネクタ 9 0 1 及び第 2 のコネクタ 9 0 2 が 2 つの第 1 のコネクタ 3 0 2 にそれぞれ接続されるとき、第 2 のコネクタ 9 0 1 及び第 2 のコネクタ 9 0 2 から入力される電源電圧を内部ドライブ制御モジュール 9 0 5 及びデータストレージモジュール 9 0 6 の電力供給電圧へと変換し、変換された電圧を内部ドライブ制御モジュール 9 0 5 及びデータストレージモジュール 9 0 6 に出力するように構成される。例えば、電力変換モジュールはバック回路 (Buck circuit) などであり得る。

40

【 0 1 1 7 】

ドライブ 3 0 3 はさらに逆流防止モジュールを含む。逆流防止モジュールは、逆流防止モジュール A 及び逆流防止モジュール B を含み、それぞれ、第 2 のコネクタ 9 0 1 と電源変換モジュール 9 0 4 との間に、及び第 2 のコネクタ 9 0 2 と電源変換モジュール 9 0 4 との間に配置される。

【 0 1 1 8 】

内部ドライブ制御モジュール 9 0 5 は、制御デバイス 3 0 1 の信号を処理してドライブ 3 0 3 を管理するという機能を実装するように、例えば、ドライブ 3 0 3 に対する制御デバイス 3 0 1 の制御信号を受信し、データストレージモジュール 9 0 6 の中のデータに対して読み取り動作又は書き込み動作を実行するという機能を提供するように構成される。

50

例えば、制御信号は、ドライブ 303 のモデル、容量、及び製造業者などの関連情報を読み取るために使用される信号、リセット状態にとどまるように第 2 のコネクタ 901 及び第 2 のコネクタ 902 を制御するために使用される信号、並びに、リセット状態をデassert するように第 2 のコネクタ 901 及び第 2 のコネクタ 902 を制御するために使用される信号である。

【0119】

ある例では、図 9C に示されるように、ドライブ 303 はさらに第 1 のストレージモジュール 907 を含み得る。第 1 のストレージモジュール 907 は図 8B に示される第 1 のストレージモジュール 806 と同じであり、詳細は本明細書では再び説明されない。

【0120】

ドライブの稼働状態を表示するために使用されるインジケータがドライブ 303 において配置される場合、内部ドライブ制御モジュール 804 はさらに、ドライブ 303 のインジケータを制御するように構成され得る。例えば、ドライブ 303 がデータ読み取り状態にあるとき、入力信号インジケータは、第 1 の周波数のインジケーション信号を出力するように制御されてよく、ドライブ 303 がデータ書き込み状態にあるとき、入力信号インジケータは、第 2 の周波数のインジケーション信号を出力するように制御されてよい、などである。

【0121】

加えて、第 2 のコネクタ 901 のための高速信号インターフェース及び低速信号インターフェース、並びに第 2 のコネクタ 902 のための高速信号インターフェース及び低速信号インターフェースは、内部ドライブ制御モジュール 905 において別々に設定される。内部ドライブ制御モジュール 905 は、第 2 のコネクタ 901 のための高速信号インターフェース及び/又は低速信号インターフェースを使用することによって、第 2 のコネクタ 901 に接続される制御デバイス 301 と通信し得る。内部ドライブ制御モジュール 905 は、第 2 のコネクタ 902 のための高速信号インターフェース及び/又は低速信号インターフェースを使用することによって、第 2 のコネクタ 902 に接続される制御デバイス 301 と通信し得る。高速信号の定義は上で説明される。内部ドライブ制御モジュール 905 と制御デバイス 301 との間で送信される通信信号において、高速信号以外の信号は低速信号である。

【0122】

もちろん、内部ドライブ制御モジュール 905 はさらに他の機能を有してよく、本明細書では 1 つずつ例は挙げられない。当業者は使用上の要件に基づいて内部ドライブ制御モジュール 905 の機能を設定し得ることに、留意されたい。これは本明細書では限定されない。

【0123】

ある例では、内部ドライブ制御モジュール 905 は、1 つの個別のデバイス又は複数の個別のデバイスの組合せを使用することによって実装されてよく、例えば、中央処理装置 (central processing unit、CPU)、デジタルシグナルプロセッサ、特定用途向け集積回路、フィールドプログラマブルゲートアレイ (field-programmable gate array、FPGA)、又は別のプログラマブルロジックデバイスであってよい。これは本明細書では限定されない。

【0124】

データストレージモジュール 906 は、内部ドライブ制御モジュール 905 の制御下でユーザデータを記憶し、又は読み取るように構成される。ある例では、データストレージモジュール 906 はドライブ媒体であり得る。

【0125】

第 3 のドライブ / 汎用ドライブ :

図 10A を参照すると、ドライブ 303 は、2 つの第 2 のコネクタ (これらは第 2 のコネクタ 1001 及び第 2 のコネクタ 1002 である)、存在信号出力モジュール 1003、逆流防止モジュール A、逆流防止モジュール B、電源変換モジュール 1004、内部ド

10

20

30

40

50

ライブ制御モジュール1005、及びデータストレージモジュール1006を含む。ある例では、図10Bを参照すると、ドライブ303はさらに第1のストレージモジュール1007を含み得る。

【0126】

第2のコネクタ1001、第2のコネクタ1002、存在信号出力モジュール1003、電源変換モジュール1004、データストレージモジュール1006、及び第1のストレージモジュール1007はそれぞれ、図9Aの第2のコネクタ901、第2のコネクタ902、存在信号出力モジュール903、電源変換モジュール904、データストレージモジュール906、及び第1のストレージモジュール907と同じであり、詳細は本明細書では再び説明されない。

10

【0127】

以下で、内部ドライブ制御モジュール1005を説明する。

【0128】

内部ドライブ制御モジュール1005はさらに、内部ドライブ制御モジュール905の機能に加えて、ドライブの稼働モードを調整するように構成される。汎用ドライブは、分散型ドライブの稼働モード及び集中型ドライブの稼働モードという2つの稼働モードを有する。以下では、内部ドライブ制御モジュール1005が汎用ドライブの稼働モードを調整する方式を説明する。

【0129】

第1の調整方式：

20

汎用デバイスが集中型ドライブの稼働モードにある場合、内部ドライブ制御モジュール1005が、分散型ドライブの稼働モードに切り替えるようにドライブ303に命令するための、ドライブ303により送信される命令を受信するとき、内部ドライブ制御モジュール1005は、通信停止状態となるように第2のコネクタ1001及び第2のコネクタ1002の1つのアクセスポートを制御し得る。具体的には、内部ドライブ制御モジュール1005は、アクセスポートの機能を論理的に無効にし、内部ドライブ制御モジュール1005は、アクセスポートを通じて入力されるいずれの信号も受信又は処理せず、アクセスポートを通じてどのような信号も出力しない。命令は、高速信号を使用することによって帯域内で制御デバイス301によって送信されてよく、又は、図4に示される制御ユニット30112の信号を送信するために使用される制御信号ケーブルのハイレベル又はローレベルを制御することによって制御デバイス301によって示されてよい。これは本明細書では限定されない。

30

【0130】

図10Cにおいて、例えば、第2のコネクタ1002は通信停止状態となるように制御される。通信停止状態は次のように理解され得る。ドライブは、通信停止状態のコネクタを使用することによって、コネクタに接続される制御デバイス301と通信することができない。従って、ドライブの稼働モードは、集中型ドライブの稼働モードから分散型ドライブの稼働モードに切り替えられる。

【0131】

第2の調整方式：

40

汎用ドライブが集中型ドライブの稼働モードにある場合、内部ドライブ制御モジュール1005が、分散型ドライブの稼働モードに切り替えるようにドライブ303に命令するための、ドライブ303により送信される命令を受信するとき、内部ドライブ制御モジュール1005は、図10Dに示されるように、ドライブ303のデータストレージモジュール1006を2つの独立したアクセスエリアへと分割し得る。例えば、データストレージモジュール1006はドライブ媒体であり、データストレージモジュール1006は合計で12個のドライブ媒体を含み、それぞれドライブ媒体0からドライブ媒体11として標識される。内部ドライブ制御モジュール1005は、ドライブ媒体0からドライブ媒体5を独立したアクセス領域へとグループ化することによって、アクセス領域は、第2のコネクタ1001を使用することによってのみアクセスされることができ、ことを行い、

50

ドライブ媒体 6 からドライブ媒体 11 を別の独立したアクセス領域へとグループ化することによって、アクセス領域は、第 2 のコネクタ 1002 を使用することによってのみアクセスされることができ、ことを行って、集中型ドライブの稼働モードから分散型ドライブの稼働モードにドライブの稼働モードを切り替え得る。

【0132】

第 3 の調整方式：

汎用ドライブが分散型ドライブの稼働モードにある場合、内部ドライブ制御モジュール 1005 が、集中型ドライブの稼働モードに切り替えるようにドライブ 303 に命令するための、ドライブ 303 により送信される命令を受信するとき、内部ドライブ制御モジュール 1005 は、通信停止状態にあるアクセスポートの通信機能を回復させ得る。具体的には、内部ドライブ制御モジュール 1005 はアクセスポートの機能を論理的に有効にし、内部ドライブ制御モジュール 1005 は、アクセスポートを通じて入力される任意の信号を受信又は処理し、アクセスポートを通じて信号を出力する。命令は、高速信号を使用することによって帯域内で制御デバイス 301 によって送信されてよく、又は、図 4 に示される制御ユニット 30112 の信号を送信するために使用される制御信号ケーブルのハイレベル又はローレベルを制御することによって制御デバイス 301 により示されてよい。これは本明細書では限定されない。従って、ドライブの稼働モードは、分散型ドライブの稼働モードから集中型ドライブの稼働モードに切り替えられる。

10

【0133】

第 4 の調整方式：

汎用ドライブが分散型ドライブの稼働モードにある場合、内部ドライブ制御モジュール 1005 が、集中型ドライブの稼働モードに切り替えるようにドライブ 303 に命令するための、ドライブ 303 により送信される命令を受信するとき、内部ドライブ制御モジュール 1005 は、データストレージモジュール 1006 を 2 つの独立したアクセス領域へと分割するという設定を取り消してよく、そうすることによって、ドライブ 303 は 1 つだけの独立したアクセス領域を含み、第 2 のコネクタ 1001 及び第 2 のコネクタ 1002 に接続される 2 つの制御デバイス 301 は、データストレージモジュール 1006 のいずれのドライブ媒体にもアクセスすることができる。従って、ドライブの稼働モードは、分散型ドライブの稼働モードから集中型ドライブの稼働モードに切り替えられる。

20

【0134】

本出願のこの実施形態では、各ドライブ 303 のタイプは、ドライブ 303 のモデル、容量、及び製造業者などの関連情報に従って決定される。具体的な決定方法は上で説明されており、詳細は本明細書では再び説明されない。

30

【0135】

このようにして、制御デバイス 301 は、分散型ストレージアーキテクチャにおけるストレージシステム及び集中型ストレージアーキテクチャにおけるストレージシステムを別々に構成するために、異なるタイプのドライブ 303 に接続される。

【0136】

第 2 のコネクタがドライブ 303 に固定的に接続され得ること、又はドライブ 303 に取り外し可能に接続され得ることに留意されたい。これは本明細書では限定されない。

40

【0137】

制御デバイス 301 が分散型ドライブに接続されるとき、分散型ストレージアーキテクチャにおけるストレージシステムが形成され得る。ある例では、分散型ストレージアーキテクチャにおけるストレージシステムの構造図は、図 3D 及び図 4 に示されるストレージシステムの構造と同じであり得る。図 3D 及び図 4 において第 1 のコネクタ 302 に接続されるドライブ 303 が分散型ドライブであるとき、分散型ストレージアーキテクチャにおけるストレージシステムが形成される。

【0138】

図 3D に示される分散型ストレージアーキテクチャにおけるストレージシステムでは、制御デバイス A、制御デバイス B、及び制御デバイス C は、互いに独立した分散型アクセ

50

スノードであり、各制御デバイスが、制御デバイスに接続されるドライブに独立にアクセスする。制御デバイスの1つに欠陥があるとき、ストレージシステムの別のノードが、欠陥のある制御デバイスに対応するサービスを引き継ぎ、欠陥のある制御デバイスに接続されるドライブの中のデータは、別のドライブの中のデータに基づいて別のノードによって復元され得る。図11に示されるように、制御デバイスAに欠陥がある場合、制御デバイスB又は制御デバイスCが制御デバイスAのサービスを引き継いでよく、制御デバイスAに接続されるドライブの中のデータは、制御デバイスBに接続されるドライブにおいてデータを復元して制御デバイスCに接続されるドライブにおいてデータを復元することを通じて、制御デバイスBに接続されるドライブ及び制御デバイスCに接続されるドライブに復元され得る。

10

【0139】

制御デバイス301が集中型ドライブに接続される時、集中型ストレージアーキテクチャのストレージシステムが形成され得る。図12A及び図12Bは、集中型ストレージアーキテクチャにおけるストレージシステムの例である。ストレージシステム300は2つの制御デバイス301及び2つの集中型ドライブ303を含み、各制御デバイス301は2つの第1のコネクタ302を含む。第1の制御デバイス301の2つの第1のコネクタ302はそれぞれ、2つの集中型ドライブの第2のコネクタ901に接続され、第2の制御デバイス301の2つの第1のコネクタ302はそれぞれ、2つの集中型ドライブの第2のコネクタ902に接続される。説明を簡単にするために、図12Aでは、矢印付きの破線が、第1のコネクタ302と第2のコネクタ901との間の接続ステータスと、第1のコネクタ302と第2のコネクタ902との間の接続ステータスとを示すために使用される。図12Bはストレージシステム300の側面図である。

20

【0140】

図12Bに示される集中型ストレージアーキテクチャにおけるストレージシステムでは、ドライブが2つの制御デバイスと別々に通信し得るので、制御デバイスの1つに欠陥があるか、又は取り除かれるとき、他方の制御デバイスが欠陥のある制御デバイスのサービスを引き継ぎ得る。

【0141】

制御デバイス301は、分散型ドライブ又は集中型ドライブに取り外し可能に接続され、そうすることによって、使用上の要件に基づいて異なるストレージアーキテクチャが形成されることが可能である。例えば、分散型ストレージアーキテクチャが形成される必要があるとき、制御デバイス301は分散型稼働モードにあることが可能にされ、分散型ドライブは第1のコネクタ302へと挿入される。集中型ストレージアーキテクチャに切り替えることが必要とされるとき、制御デバイス301は、集中型稼働モードにあることが可能にされ、分散型ドライブが取り除かれ、集中型ドライブが挿入される。

30

【0142】

ストレージアーキテクチャの切り替え処理を簡単にするために、制御デバイス301は汎用ドライブに直接接続され得る。ストレージシステム300が集中型ストレージアーキテクチャから分散型ストレージアーキテクチャに切り替わる必要があるとき、制御デバイス301は、分散型ストレージアーキテクチャに対応するシステムコードを実行し、ドライブ303に、分散型ドライブに切り替えるようにドライブ303に命令するために使用される命令を送信する。命令を受信した後で、ドライブ303は、上で説明された第1の調整方式又は第2の調整方式でドライブ303を分散型ドライブに合わせて、ストレージシステム300のストレージアーキテクチャを集中型ストレージアーキテクチャから分散型ストレージアーキテクチャに切り替え得る。ストレージシステム300が分散型ストレージアーキテクチャから集中型ストレージアーキテクチャに切り替わる必要があるとき、制御デバイス301は、集中型ストレージアーキテクチャに対応するシステムコードを実行し、ドライブ303に、集中型ドライブに切り替えるようにドライブ303に命令するために使用される命令を送信する。命令を受信した後で、ドライブ303は、上で説明された第3の調整方式又は第4の調整方式でドライブ303を集中型ドライブに合わせて、

40

50

ストレージシステム 300 のストレージアーキテクチャを分散型ストレージアーキテクチャから集中型ストレージアーキテクチャに切り替え得る。この場合、制御デバイス 301 とドライブ 303 との間の接続関係は、集中型ストレージアーキテクチャにおける制御デバイス 301 とドライブ 303 との間の接続関係と同じである。詳しくは、図 12A 及び図 12B を参照されたい。詳細は本明細書では再び説明されない。

【0143】

第 2 のコネクタは第 1 のコネクタ 302 に直接接続され得ることに留意されたい。例えば、第 2 のコネクタが第 1 のコネクタへと挿入され、又は別の方式では第 2 のコネクタが第 1 のコネクタ 302 に接続され得る。例えば、第 2 のコネクタは、通信バスを使用することによって第 1 のコネクタ 302 に接続されてよく、通信バスは、PCIe バス、I²C バスなどであってよい。これは本明細書では限定されない。

10

【0144】

以下は、第 1 のコネクタ 302 及び第 2 のコネクタを説明する。

【0145】

本出願のこの実施形態では、第 1 のコネクタ 302 及び第 2 のコネクタは、相対応する垂直で直交するコネクタであり、又は、第 1 のコネクタ 302 及び第 2 のコネクタは、予め設定された回転角を有し、かつ相対応するコネクタである。

【0146】

例えば、制御デバイス 301 は 1 つの第 1 のコネクタ 302 を含み、ドライブ 303 は 1 つの第 2 のコネクタを含み、第 1 のコネクタ 302 はジャックを有するソケットであり、第 2 のコネクタはピンを有するプラグである。

20

【0147】

本出願のこの実施形態では、例えば、制御デバイス 301 は回路基板であってよく、ドライブ 303 の形状は直方体であってよい。図 13A を参照すると、ドライブ 303 は、それぞれ面 A から面 F である 6 つの面を含む。図 13A に示されるように、ドライブの面 C と面 D の両方が水平面に平行であり、水平面上への面 C 及び面 D の投影は長方形である。長方形の 2 つの長い辺の長さはドライブ 303 に対応する直方体の長さと同じであり、直方体の 2 つの短い辺の長さはドライブ 303 に対応する直方体の高さと同じである。ドライブの面 A、面 F、面 B、及び面 E は、水平面に垂直である。水平面上への面 A 及び面 F の投影は線分であり、線分の長さはドライブ 303 に対応する直方体の長さと同じである。水平面上への面 B 及び面 E の投影も線分であり、線分の長さはドライブ 303 に対応する直方体の高さと同じである。第 2 のコネクタはドライブ 303 の面 E に設定される。

30

【0148】

図 13B は、1 つの第 1 のコネクタ 302 を含む制御デバイス 301 の側面図である。図 13B において、第 1 のコネクタ 302 のジャックは 2 行 × 3 列に並べられ、各行の 2 つの隣接するジャックは同じ間隔を空けられ、第 1 の行の 3 つのジャックは第 2 の行の 3 つのジャックと揃えられる。図 13C は、1 つの第 2 のコネクタを含むドライブ 303 の側面図である。図 13C において、第 2 のコネクタのピンは 3 行 × 2 列に並べられ、各行の 2 つの隣接するピンは同じ間隔を空けられ、加えて、2 行ごとの 2 つのピンは互いに揃えられる。第 2 のコネクタが第 1 のコネクタ 302 へと挿入されるとき、制御デバイス 301 及びドライブ 303 は、図 13D に示される位置関係にある。図 13D は、制御デバイス 301 及びドライブ 303 の側面図である。図 13D において、制御デバイス 301 は水平に設定され、ドライブ 303 の面 C は水平面に平行であり、ドライブ 303 の面 A は水平面に垂直である。

40

【0149】

直方体の高さは通常は直方体の長さ又は幅より短く、そうすることによって、1 つの制御デバイス 301 はより多くのドライブ 303 と通信してよく、ストレージシステムの容量拡張要件が満たされる。

【0150】

図 14A は、1 つの第 1 のコネクタ 302 を含む制御デバイス 301 の側面図である。

50

図14Aでは、第1のコネクタ302のジャックが2行に並べられ、各行の2つの隣接するジャックは同じ間隔を空けられ、第1の行の第1のジャックと第2の行の第1のジャックとの間には特定の距離がある。図14Bは、1つの第2のコネクタを含むドライブ303の側面図である。図14Bでは、第2のコネクタのピンが2つの行に並べられ、各行の2つの隣接するジャックが同じ間隔を空けられ、第1の行の第1のジャックと第2の行の第1のジャックとの間には特定の距離がある。第2のコネクタが第1のコネクタ302へと挿入されるとき、制御デバイス301及びドライブ303は、図14Cに示される位置関係にある。図14Cは、制御デバイス301及びドライブ303の側面図である。図14Cでは、制御デバイス301は水平に設定され、ドライブ303の面Aと水平面との間には、予め設定された挟角がある。図14Cでは、例えば、ドライブ303の面Aと水平面との間の挟角は80度である。もちろん、ドライブ303の面Aと水平面は、代替的に、別の角度で配置されてよく、ドライブ303の面Aと水平面との間の角度は、ドライブ303におけるピンの配置位置に関係し、本明細書では限定されない。

10

【0151】

前述の技術的解決法において、第1のコネクタ302及び第2のコネクタは、垂直で直交するコネクタ、又は予め設定された回転角を有するコネクタであり、そうすることによって、第1のコネクタ302及び第2のコネクタは、空間座標系のX軸、Y軸、及びZ軸の方向における制御デバイス301とドライブ303との間の相対的なシフトに打ち勝つことができる。従って、複数の制御デバイス301は複数のドライブ303に確実に接続されることができる。従来技術のバックプレーンが取り除かれて、ストレージシステムの設計の複雑さを下げ、バックプレーンの欠陥によってストレージシステム全体が使用されることができなくなる確率を下げ、そうすることによって、ストレージシステムの信頼性を改善する。

20

【0152】

さらに、従来技術のバックプレーンが取り除かれ、そうすることによって、通信処理における制御デバイス301とドライブ303との間でのインピーダンス不整合点が減らされることができる。具体的には、制御デバイス301とバックプレーンとの間のインピーダンス不整合、及びバックプレーンとドライブとの間のインピーダンス不整合が減らされる。これは、信号送信レートを改善できる。

【0153】

さらに、従来技術のバックプレーンが取り除かれ、そうすることによって、ストレージシステムにおけるエアフロー欠陥点が減らされ、ストレージシステムの熱放散性能が改善されることができる。従って、ストレージシステムの特定空間において、ストレージシステムの特定熱放散性能で、ストレージシステム内に、より多くのデバイスが許容されることでき、ストレージシステムの密度が改善されることができる。

30

【0154】

当業者は、制御デバイス301とドライブ303の相対的な位置を設定するために、実際の使用上の要件に基づいて異なる第1のコネクタ302及び第2のコネクタを選択し得る。

【0155】

加えて、制御デバイス301が複数の第1のコネクタ302を含むとき、複数の第1のコネクタ302の間に特定の間隔が設定される必要があることに留意されたい。例えば、制御デバイス301及びドライブ303が図14Cに示される第2の接続方式で接続されるとき、制御デバイス301における2つ毎の第1のコネクタ302の間隔は、ドライブ303の高さ以上である。ドライブ303の高さの定義については、図13Aを参照されたい。詳細は本明細書では再び説明されない。

40

【0156】

前述の技術的解決法において、使用上の要件に基づいて、集中型ストレージアーキテクチャにおけるストレージシステム又は分散型ストレージアーキテクチャのストレージシステムは、制御デバイス301を異なる構成を持つドライブ303に接続することによって

50

、フレキシブルに構築され得る。加えて、ストレージアーキテクチャが切り替えられる必要があるとき、制御デバイス301とドライブ303の稼働モードだけが切り替えられる必要があり、又はドライブ303が交換され、制御デバイス301とドライブ303が再び購入される必要がない。これは、ストレージシステムのフレキシビリティを改善し、ストレージアーキテクチャの切り替えのコストを減らす。

【0157】

同じ発明概念に基づいて、本出願のある実施形態は、ストレージシステム動作方法を提供する。方法は、図3D、図4、又は図12Aに示されるストレージシステムに適用される。

【0158】

図15Aから図15Cは、本出願のある実施形態によるストレージシステム動作方法のフローチャートである。フローチャートは次のように説明される。

【0159】

S1：制御デバイスがシステム構成を実行する。

【0160】

ある例では、制御デバイスはクライアントに接続され得る。クライアントにおいて、ユーザは、使用上の要件に基づいてストレージシステムのタイプを設定し得る。例えば、ユーザは、集中型ストレージアーキテクチャ又は分散型ストレージアーキテクチャにストレージシステムを設定し得る。クライアントから命令を受信した後で、制御デバイスは、命令に従ってシステム構成を実行し得る。システム構成は、制御デバイスのシステムコード構成を含む。例えば、制御デバイスのメモリは、集中型ストレージアーキテクチャに対応するシステムコード及び分散型ストレージアーキテクチャに対応するシステムコードを記憶し、対応するシステムコードがクライアントの命令に従ってストレージシステムのために構成される。

【0161】

システム構成はさらにドライブ構成を含み得る。制御デバイスに接続されるドライブが図10Aに示される汎用ドライブであるとき、制御デバイスは、分散型ドライブ又は集中型ドライブにドライブを設定するために命令をドライブに送信し得る。例えば、クライアントからの命令がストレージシステムを集中型ストレージアーキテクチャに設定するように命令するとき、ドライブのアクセスポートの数を2に設定するために、命令がドライブに送信され得る。クライアントからの命令が分散型ストレージアーキテクチャにストレージシステムを設定するように命令するとき、ドライブのアクセスポートの数を1に設定するために、命令がドライブに送信され得る。もちろん、代替的に、ドライブにおけるアクセス領域の数が設定されてよく、詳細は本明細書では説明されない。

【0162】

ステップS1は任意選択のステップであり、言い換えると、ステップS1は実行される必要がないことに留意されたい。例えば、ユーザは、使用上の要件に基づいて、使用されるべきストレージアーキテクチャに対応する制御デバイス及びドライブを直接選択し得る。例えば、ユーザが集中型ストレージアーキテクチャを使用する必要がある場合、ユーザは、集中型ストレージアーキテクチャを実装できる制御デバイスを集中型ドライブに直接接続し得る。この場合、ステップS1は実行される必要はない。

【0163】

S2：制御デバイスがスタートされ初期化される。

【0164】

S3：制御デバイスが、システム構成のタイプに基づいてドライブ決定条件を設定する。

【0165】

本出願のこの実施形態では、ドライブ決定条件は、制御デバイスに接続されるドライブが制御デバイスと適合するかどうかを決定するために使用される。例えば、集中型ストレージアーキテクチャにあるとき、制御デバイスは、ドライブ決定条件に基づいて、制御デバイスに接続されるドライブが集中型ドライブであるかどうかを決定する。ある例では、

10

20

30

40

50

ストレージシステムのために構成されるシステムコードが集中型ストレージアーキテクチャに対応するシステムコードであると、制御デバイスが決定する場合、制御デバイスは、集中型ドライブ決定条件としてドライブ決定条件を構成する。集中型ドライブ決定条件は、集中型ドライブの製造業者及びモデルなどの情報を含み得る。例えば、集中型決定条件は、製造業者 A 又はモデル A であり得る。ストレージシステムのために構成されるシステムコードが分散型ストレージアーキテクチャに対応するシステムコードであると、制御デバイスが決定する場合、制御デバイスは、分散型ドライブ決定条件としてドライブ決定条件を構成する。それに対応して、分散型ドライブ決定条件は、分散型ドライブの製造業者及びモデルなどの情報を含み得る。例えば、分散型決定条件は、製造業者 B 又はモデル B であり得る。もちろん、集中型ドライブ決定条件及び分散型ドライブ決定条件は、代替的に、他のパラメータであってよく、これは本明細書では限定されない。

10

【 0 1 6 6 】

S 4 : 制御デバイスが、各第 1 のコネクタがドライブに接続されているか、又は切断されているかを検出する。

【 0 1 6 7 】

ある例では、制御デバイスの CPU が、 CPLD を使用することによってドライブの存在信号を検出し、各第 1 のコネクタがドライブに接続されているか、又は切断されているかを決定し、決定結果を CPLD の対応するレジスタに記憶し得る。CPU は、各第 1 のコネクタがドライブに接続されているか、又は切断されているかを決定するために、レジスタに記憶されているビットの値を読み取る。第 1 のコネクタがドライブから切断されているとき、制御デバイスは、クライアントのディスプレイを使用することによって警告を出力し得る。例えば、制御デバイスはクライアントにプロンプト情報を出力し、プロンプト情報は第 1 のコネクタのインデックス番号を搬送する。もちろん、代替的に、別の警告モードが使用されてよく、本明細書では限定されない。

20

【 0 1 6 8 】

S 5 : 制御デバイスが、接続状態にあるドライブの数量を決定する。

【 0 1 6 9 】

制御デバイスは、接続状態にあるドライブの数量を決定するために、各第 1 のコネクタがドライブに接続されているか、又は切断されているかを検出する。接続状態にあるドライブの数量が 0 に等しい場合、S 1 3 が実行されることになる。接続状態にあるドライブの数量が 1 以上である場合、S 6 が実行されることになる。

30

【 0 1 7 0 】

S 6 : 制御デバイスが、接続状態にあるドライブについての情報を取得する。

【 0 1 7 1 】

ある例では、制御デバイスの CPU は、PCI E スイッチを使用することによって接続状態にある各ドライブについての情報を取得してよく、又は、接続状態にある各ドライブについての情報を取得するために、PCI E バスを使用することによって接続状態にあるドライブに命令を送信してよい。情報は、ドライブのモデル、容量、及び製造業者などの関連情報を含む。

【 0 1 7 2 】

S 7 : 制御デバイスが、適合するドライブの数量を決定する。

【 0 1 7 3 】

具体的には、接続状態にあるドライブについての情報を取得した後で、制御デバイスの CPU は、S 3 において設定された決定条件に基づいて、接続状態にある各ドライブのタイプを決定する。例えば、S 3 において設定された決定条件が集中型ドライブ決定条件であり、集中型ドライブ決定条件が製造業者 A である場合、CPU は、接続状態にあるドライブの製造業者が製造業者 A であるかどうかを決定する。接続状態にあるドライブの製造業者が製造業者 A である場合、CPU は、ドライブが、適合するドライブであると決定し、接続状態にあるドライブの製造業者が製造業者 A ではない場合、CPU は、ドライブが、適合しないドライブであると決定する。CPU は、接続状態にあるドライブの決定結果

40

50

を、C P L Dの対応するレジスタ、例えば、第2のレジスタに記憶し得る。ドライブが、適合するドライブであると決定するとき、C P Uは、C P L Dのドライブに対応する第2のレジスタにおけるビットの値を1に設定し、ドライブが、適合しないドライブであると決定するとき、C P Uは値を0に設定する。加えて、ドライブが、適合しないドライブであると決定した後で、C P Uはさらに、警告情報をクライアントに出力し得る。

【0174】

次いで、制御デバイスのC P Uは、C P L Dの接続状態にある各ドライブに対応する、第2のレジスタのビットの値を使用することによって、適合するドライブの数量を計算し得る。適合するドライブの数量が0に等しい場合、S 1 3が実行されることになり、適合するドライブの数量が1に等しい場合、S 8が実行されることになり、適合するドライブの数量が1より大きい場合、S 1 0が実行されることになる。

10

【0175】

本出願のこの実施形態では、ステップS 4とS 5、ステップS 6とS 7の実行順序は限定されないことに留意されたい。具体的には、制御デバイスは、最初にステップS 4とS 5を実行し、次いでステップS 6とS 7を実行してよく、又は、最初にステップS 6とS 7を実行し、次いでステップS 4とS 5を実行してよく、又は、ステップS 4とS 5、及びステップS 6とS 7を同時に実行してもよい。当業者は、使用法に基づいて順序を設定し得る。

【0176】

S 8：制御デバイスが、単一ドライブの電源オンスタートアッププロセスを実行する。

20

【0177】

ステップS 6を実行するとき、制御デバイスのC P Uは、ステップS 6においてP C I Eスイッチを使用することによって情報を取得してよく、又は、ステップS 6においてP C I Eバスを使用することによって情報を取得してよく、従って、ステップS 8の実行順序は2つの異なる実装について説明されることに留意されたい。

【0178】

情報がステップS 6においてP C I Eスイッチを使用することによって取得される場合、ステップS 8はステップS 6の後で実行される。情報がステップS 6においてP C I Eバスを使用することによって取得される場合、制御デバイスは、まずステップS 8を実行してドライブの電源をオンにし、次いでステップS 6を実行する必要がある。

30

【0179】

S 9：制御デバイスが、稼働状態となるように対応する適合するドライブを制御する。

【0180】

具体的には、制御デバイスのC P Uは、動作クロック及び/又は高速信号を送信するために使用されるクロックを適合するドライブに提供するようにクロック信号生成器を制御し、そうすることによって、適合するドライブは稼働状態になる。

【0181】

S 1 0：制御デバイスが、インバッチ電源オンのプロセスを実行する。

【0182】

具体的には、適合するドライブの数量が1より多いと決定するとき、制御デバイスのC P Uはまず、予め設定された信頼性ポリシー又はフォールトトレラントポリシーに基づいて複数の適合するドライブのうちの1つ又はいくつかの電源をオンにし、次いで残りのドライブの電源をオンにし得る。具体的なポリシーの内容は本明細書では限定されない。

40

【0183】

S 1 1：制御デバイスが、稼働状態となるように、数回に分けて、対応する適合するドライブを制御する。

【0184】

ステップS 1 1はステップS 9と同じであり、詳細は再び説明されない。

【0185】

S 1 2：対応するアーキテクチャのストレージシステムをスタートする。

50

【 0 1 8 6 】

S 1 3 : 制御デバイスが、ドライブと第 1 のコネクタとの間の接続状態又は切断状態が変化するかどうかを検出する。

【 0 1 8 7 】

具体的には、制御デバイスの P C I E スイッチは、C P L D の割り込み信号を使用することによって、ドライブと第 1 のコネクタの接続状態又は切断状態が変化するかどうかを監視する。P C I E スイッチが検出を通じて C P L D の割り込み信号を取得する場合、P C I E スイッチは、変化が発生すると決定し、S 1 4 が実行されることになる。P C I E スイッチが検出を通じて C P L D の割り込み信号を取得しない場合、P C I E スイッチは、変化が発生しないことを決定し、S 1 8 が実行されることになる。

10

【 0 1 8 8 】

S 1 4 : 制御デバイスが、接続状態にあるドライブの数量が変化するかどうかを決定する。

【 0 1 8 9 】

具体的には、制御デバイスの C P U は、C P L D の第 1 のレジスタのビットの値に基づいて、接続状態にあるドライブの数量が変化するかどうかを決定する。C P U が、第 1 のレジスタのビットの値に基づいて、接続状態にあるドライブの数量が増加すると決定する場合、C P U は、ドライブが挿入されると決定し、ステップ S 5 が実行されることになる。C P U が、第 1 のレジスタの値に基づいて、接続状態にあるドライブの数量が減少すると決定する場合、C P U は、ドライブが取り除かれると決定し、ステップ S 1 5 が実行されることになる。

20

【 0 1 9 0 】

S 1 5 : 制御デバイスが、取り除かれるドライブが欠陥のあるドライブであるかどうかを決定する。

【 0 1 9 1 】

具体的には、制御デバイスの C P L D は、接続状態にある各ドライブの稼働状態をリアルタイムで監視し、C P L D のドライブの欠陥状態を記憶するように構成されるレジスタに監視結果を記憶する。ドライブが取り除かれると決定するとき、C P U は、ドライブの欠陥状態を記憶するように構成され、かつ C P L D のドライブに対応する、レジスタのビットの値を読み取り、ドライブが欠陥のあるドライブであるかどうかを決定する。ドライブが欠陥のあるドライブである場合、ステップ S 1 7 が実行されることになる。ドライブが欠陥のないドライブではない場合、ステップ S 1 6 が実行されることになる。

30

【 0 1 9 2 】

S 1 6 : 制御デバイスが、電源オフ状態となるようにドライブを制御し、ドライブのためのクロック信号を提供することを停止し、C P L D の第 1 のレジスタのビットの値を接続状態に対応する値に設定する。

【 0 1 9 3 】

S 1 7 : 制御デバイスが、ドライブの欠陥状態のレコードをクリアし、C P L D の第 1 のレジスタのビットの値を切断状態に対応する値に設定する。

【 0 1 9 4 】

ドライブは既に取り除かれているので、制御デバイスが、ドライブの欠陥状態の記録をクリアする必要がある。言い換えると、ドライブが欠陥状態にあるかどうかを示すために使用される、C P L D のドライブの欠陥状態を記憶するように構成されるレジスタのビットの値は、初期値に設定される。ある例では、C P L D のドライブの欠陥状態を記憶するように構成されるレジスタのビットの初期値は 0 であり、そうすることによって、ビットの値は 0 に設定される。

40

【 0 1 9 5 】

S 1 8 : 制御デバイスがドライブの欠陥状態を検出する。

【 0 1 9 6 】

C P U は、C P L D のドライブの欠陥状態を記憶するように構成されドライブに対応す

50

る、レジスタのビットの値を読み取り、ドライブが欠陥のあるドライブであるかどうかを決定する。ドライブが欠陥のあるドライブである場合、ステップ S 1 9 が実行されることになる。ドライブが欠陥のあるドライブではない場合、ステップ S 2 0 が実行されることになる。

【 0 1 9 7 】

S 1 9 : 制御デバイスが、電源オフ状態となるようにドライブを制御し、ドライブのためのクロック信号を提供することを停止し、C P L D の第 1 のレジスタのビットの値を切断状態に対応する値に設定する。

【 0 1 9 8 】

S 2 0 : 制御デバイスが、遮断コマンドが受信されるかどうかを検出する。

10

【 0 1 9 9 】

制御デバイスの C P U は、クライアントから遮断コマンドが受信されるかどうかを監視する。遮断コマンドが受信される場合、ステップ S 2 1 が実行されることになり、遮断命令が受信されない場合、ステップ S 1 3 が実行されることになる。

【 0 2 0 0 】

S 2 1 : 停止する。

【 0 2 0 1 】

図 1 5 A から図 1 5 C に示されるストレージシステム動作方法において、スタートアップの前に、ストレージシステムの制御デバイスは、制御デバイスの稼働モード、具体的には、集中型ストレージアーキテクチャで動作するモード又は分散型ストレージアーキテクチャで動作するモードを決定しており、次いで、ストレージシステムは稼働モードで動作する。

20

【 0 2 0 2 】

本出願のある実施形態のストレージシステムでは、稼働モード切り替えはさらに、動作中に実行され得る。図 1 6 は、本出願のある実施形態による、ストレージシステムの稼働モードを切り替えるための方法のフローチャートである。フローチャートは次のように説明される。

【 0 2 0 3 】

ステップ S 1 6 0 1 : ストレージシステムの中の M 個の制御デバイスが第 1 の稼働モードで動作する。

30

【 0 2 0 4 】

本出願のこの実施形態では、M 個の制御デバイスは第 1 の稼働モード及び第 2 の稼働モードをサポートし、第 1 の稼働モード及び第 2 の稼働モードは以下の 2 つの定義を有する。

【 0 2 0 5 】

第 1 の定義 :

第 1 の稼働モードが集中型ストレージアーキテクチャで動作するモードであるとき、第 2 の稼働モードは分散型ストレージアーキテクチャで動作するモードである。

【 0 2 0 6 】

第 2 の定義 :

第 1 の稼働モードが分散型ストレージアーキテクチャで動作するモードであるとき、第 2 の稼働モードは集中型ストレージアーキテクチャで動作するモードである。

40

【 0 2 0 7 】

M は正の整数であることに留意されたい。第 1 の稼働モード及び第 2 の稼働モードの具体的な内容は既に上で説明されており、詳細は本明細書では再び説明されない。

【 0 2 0 8 】

ステップ S 1 6 0 2 : M 個の制御デバイスのうちの少なくとも 1 つの制御デバイスが、ストレージシステムの中の K 個のドライブの各々に第 1 の命令を送信し、K 個のドライブの各々が第 1 の命令を受信する。

【 0 2 0 9 】

本出願のこの実施形態では、第 1 の命令は、第 2 の稼働モードで動作するように、第 1

50

の命令を受信するドライブに命令するために使用される。

【0210】

ある例では、M個の制御デバイスはクライアントに別々に接続されてよく、M個の制御デバイスは、第1の稼働モードから第2の稼働モードに切り替えるための、クライアントにより送信される命令を受信して、第1の命令をK個のドライブに送信する。M個の制御デバイスはまた、予め記憶された動作ポリシー、例えば、ユーザ設定を使用することによって、第1の時間期間において第1の稼働モードで動作し、第2の時間期間において第2の稼働モードで動作してよく、ストレージシステムは、切り替え条件が満たされると決定するとき、第1の命令をK個のドライブに送信する。代替的に、プライマリ制御デバイスとセカンダリ制御デバイスが、M個の制御デバイスにおいて設定され得る。プライマリ制御デバイスは、クライアントから、又は予め記憶された動作ポリシーを使用することによって、ドライブが第1の稼働モードから第2の稼働モードに切り替えられる必要があると決定し、第1の命令をK個のドライブの各々に送信する。もちろん、M個の制御デバイスは、K個のドライブの各ドライブに第1の命令を送信するための別の方式でトリガされてもよく、これは本明細書では限定されない。

10

【0211】

第1の命令は、M個の制御デバイスの各々によって送信されてよく、又はM個の制御デバイスの1つ又はいくつかによって送信されてよいことに留意されたい。これは本明細書では限定されない。

【0212】

ストレージシステムにおいて、M個の制御デバイスの各々が、N個の第1のコネクタの全て及びN個の第2のコネクタの全てを使用することによってK個のドライブと通信してよく、又は、M個の制御デバイスの各々が、N個の第1のコネクタのいくつか及びN個の第2のコネクタのいくつかを使用することによってK個のドライブと通信してよい。N個の第1のコネクタはM個の制御デバイスに取り外し可能に接続するように構成され、N個の第2のコネクタはK個のドライブに取り外し可能に接続するように構成される。K個のドライブの各々がデータを記憶するように構成され、N及びKは正の整数である。

20

【0213】

第1のコネクタの構造、第2のコネクタの構造、第1のコネクタを制御デバイスに接続する方式、第2のコネクタをドライブに接続する方式、並びに制御デバイスとドライブの構造及び機能は、既に上で説明されており、本明細書では再び説明されないことに留意されたい。

30

【0214】

加えて、ステップS1601及びステップS1602の実行順序は本出願のこの実施形態では限定されないことに留意されたい。具体的には、ステップS1601が最初に実行されてよく、次いでステップS1602が実行される、又は、ステップS1602が最初に実行され、次いでステップS1601が実行される、又は、ステップS1601とステップS1602が同時に実行される。

【0215】

ステップS1603：K個のドライブの各々が、第1の命令に従って、第2の稼働モードに適応した状態となるように、各ドライブのX個の第2のコネクタのうちのP個の第2のコネクタを制御する。

40

【0216】

本出願のこの実施形態では、第2の稼働モードは、分散型ストレージアーキテクチャで動作するモードであってよく、又は、集中型ストレージアーキテクチャで動作するモードであってよい。従って、第2の稼働モードに適応した状態となるように各ドライブのX個の第2のコネクタのうちのP個の第2のコネクタを制御することの定義も異なり、具体的には次の通りである。

【0217】

第2の稼働モードが分散型ストレージアーキテクチャで動作するモードであるとき、第

50

2の稼働モードに適応した状態となるようにP個の第2のコネクタのうちの1つを制御することは、

第2のコネクタを使用することによって、第2のコネクタに接続される第1のコネクタに対応する制御デバイスと通信することを停止するようにドライブを制御することを含み、X及びPは正の整数であり、XとPの差は1である。

【0218】

第2の稼働モードが集中型ストレージアーキテクチャで動作するモードであるとき、第2の稼働モードに適応した状態となるようにP個の第2のコネクタを制御することは、

P個の第2のコネクタのうちのL個の第2のコネクタを使用することによって、L個の第2のコネクタに接続されるL個の第1のコネクタに対応する制御デバイスと通信することを再開するようにドライブを制御することを含み、Lは正の整数であり、LはP以下である。

10

【0219】

第2のコネクタを使用することによって、第2のコネクタに接続される第1のコネクタに対応する制御デバイスと通信することを停止するようにドライブを制御する方式、及び、L個の第2のコネクタを使用することによって、L個の第2のコネクタに接続されるL個の第1のコネクタに対応する制御デバイスと通信することを再開するようにドライブを制御する方式は、上で説明されており、詳細は本明細書では再び説明されないことに留意されたい。

【0220】

ドライブが稼働モードを切り替えることを完了した後で、ストレージシステム全体が、第1の稼働モードに対応するストレージアーキテクチャから第2の稼働モードに対応するストレージアーキテクチャに切り替えられ、ストレージアーキテクチャ切り替え処理を完了する。

20

【0221】

前述の技術的解決法において、ストレージシステムが稼働モードを切り替える必要があるとき、制御デバイスの稼働モードが切り替えられ、制御デバイスは、稼働モード切り替え命令をドライブに送信して、ハードウェア構造と接続関係を変更することなくストレージシステムの稼働モードを切り替える。ストレージシステムの稼働モードを切り替えるステップが簡略化されることができる。

30

【0222】

前述の実装の説明により、本出願の実施形態は、ハードウェア、ファームウェア、又はそれらの組合せによって実装され得ることを、当業者は明確に理解し得る。本出願の実施形態がソフトウェアによって実装されるとき、前述の機能は、コンピュータ可読媒体に記憶され、又は、コンピュータ可読媒体において1つ又は複数の命令又はコードとして送られ得る。コンピュータ可読媒体は、コンピュータ記憶媒体又は通信媒体を含む。通信媒体は、ある場所から別の場所へのコンピュータプログラムの送信を容易にする任意の媒体を含む。記憶媒体はコンピュータによってアクセス可能な任意の使用可能な媒体であり得る。これは、例として使用されるが、次のことに限定はされない。コンピュータ可読媒体は、RAM、ROM、電氣的消去可能プログラマブル読み取り専用メモリ (electrically erasable programmable read only memory、EEPROM)、読み取り専用ディスク (compact disc read-only memory、CD-ROM) 又は別の光学ディスクストレージ、磁気ディスク記憶媒体又は別の磁気記憶デバイス、又は、命令又はデータ構造の形で所望のプログラムコードを搬送又は記憶するように構成されることができ、かつコンピュータによってアクセスされることができる、任意の他の媒体を含み得る。加えて、任意の接続が、コンピュータ可読媒体として適切に定義され得る。例えば、ソフトウェアが、ウェブサイト、サーバ、又は別のリモートソースから、同軸ケーブル、光ファイバ/ケーブル、ツイストペア、デジタル加入者線 (Digital Subscriber Line、DSL)、又は赤外線、無線、及びマイクロ波などのワイヤレス技術を使用することによって送信され

40

50

る場合、同軸ケーブル、光ファイバ/ケーブル、ツイストペア、DSL、又は、赤外線、無線、及びマイクロ波などのワイヤレス技術は、それらが属する媒体の固定に含まれる。例えば、本発明の実施形態において使用されるディスク(disk)及びディスク(disc)は、コンパクトディスク(compact disc、CD)、レーザーディスク、光学ディスク、デジタル多用途ディスク(digital video disc、DVD)、フロッピーディスク、Blu-rayディスクを含む。ディスクは一般に磁気手段によってデータを複製し、ディスクはレーザー手段によって光学的にデータを複製する。前述の組合せも、コンピュータ可読媒体の保護範囲に含まれるべきである。

【0223】

要約すると、上で説明されたことは本出願の実施形態にすぎず、本出願の保護範囲を限定することは意図されていない。本出願の開示に従ってなされる任意の修正、等価な置換、又は改善が、本出願の保護範囲内に入るものとする。

10

20

30

40

50

【図面】

【図 1 A】

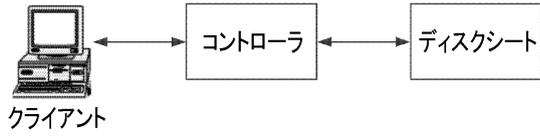


FIG. 1A

【図 1 B】

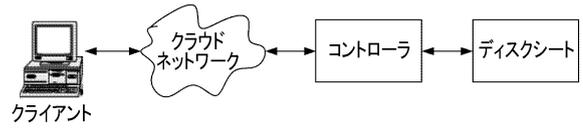


FIG. 1B

10

【図 2 A】

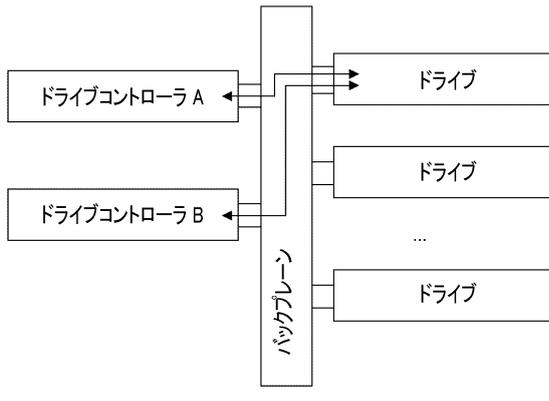


FIG. 2A

【図 2 B】

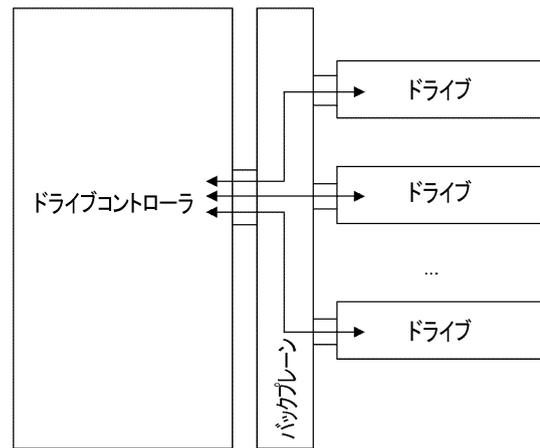


FIG. 2B

20

30

40

50

【図 3 A】

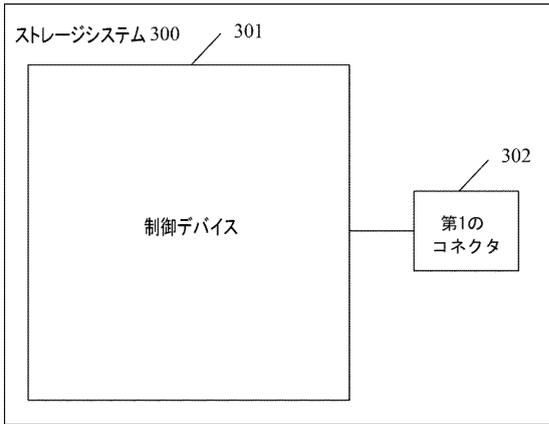


FIG. 3A

【図 3 B】

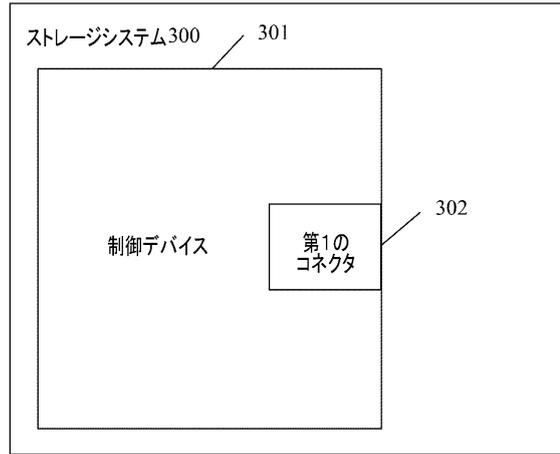


FIG. 3B

【図 3 C】

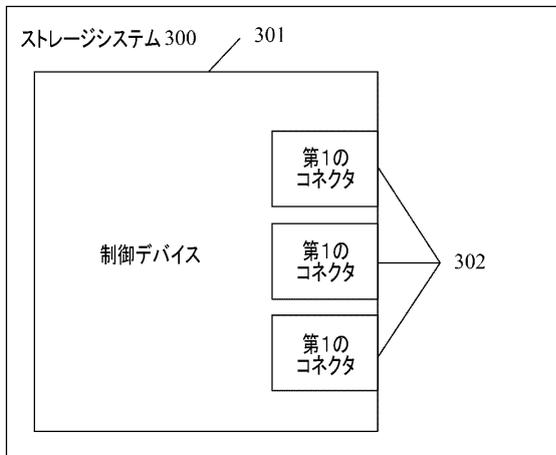


FIG. 3C

【図 3 D】

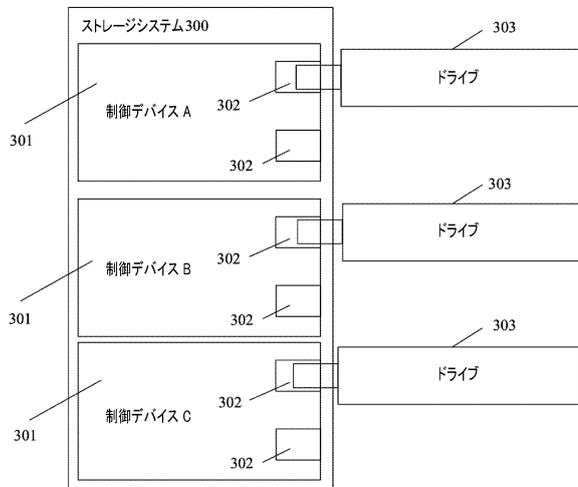


FIG. 3D

10

20

30

40

50

【 図 4 】

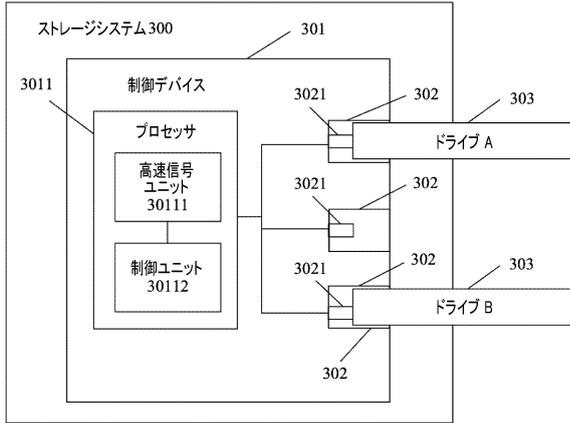


FIG. 4

【 図 5 】

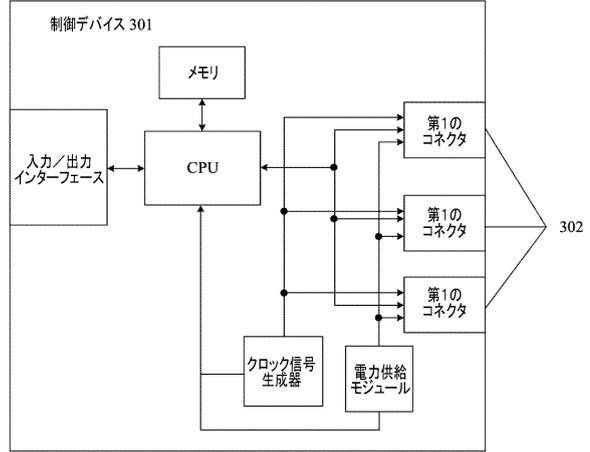


FIG. 5

10

【 図 6 】

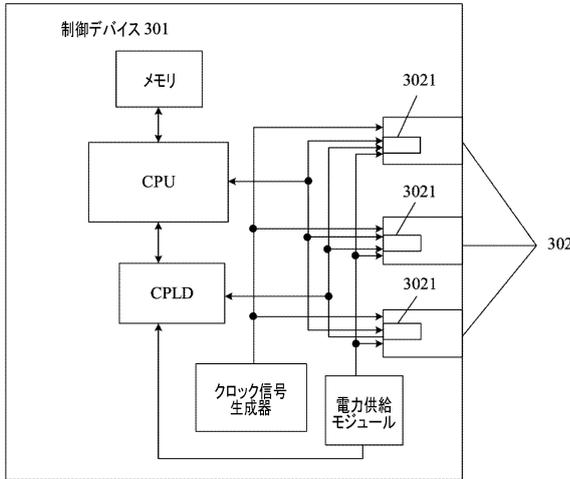


FIG. 6

【 図 7 】

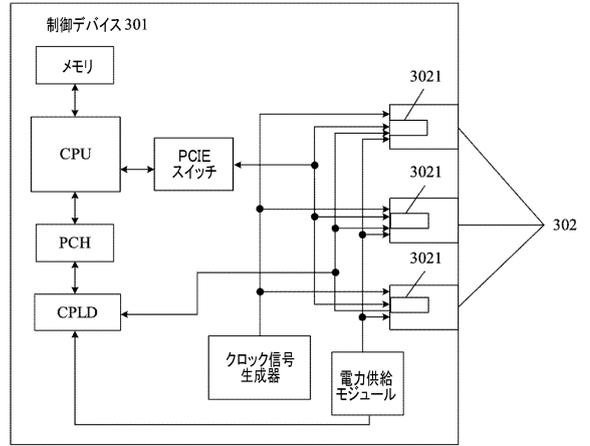


FIG. 7

20

30

40

50

【図 8 A】

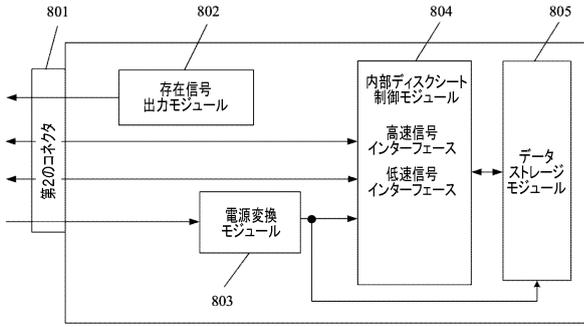


FIG. 8A

【図 8 B】

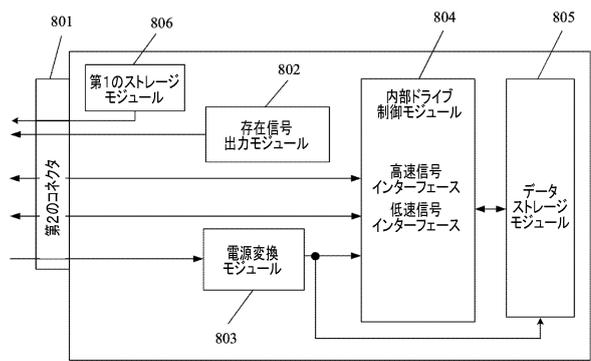


FIG. 8B

10

【図 9 A】

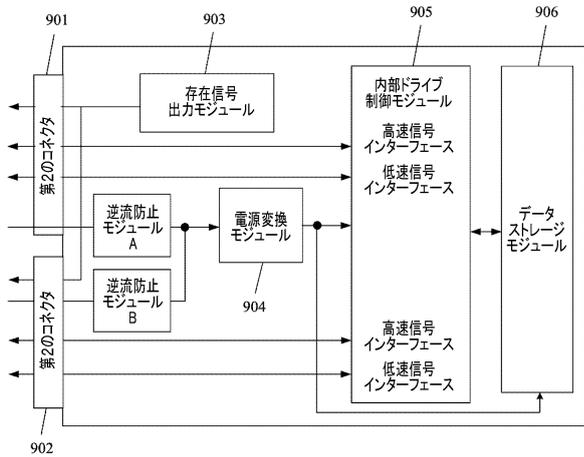


FIG. 9A

【図 9 B】

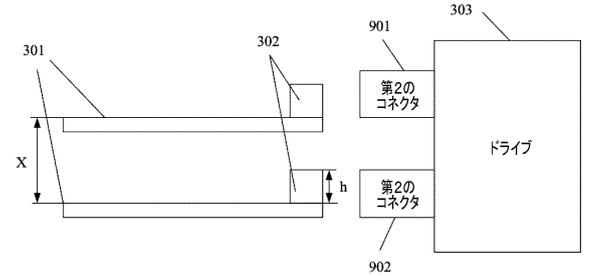


FIG. 9B

20

30

40

50

【図9C】

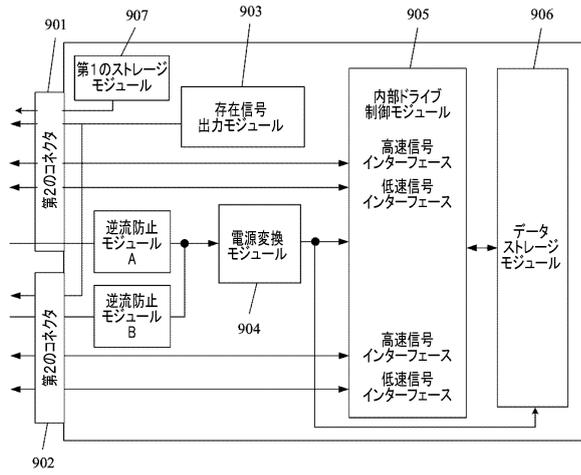


FIG. 9C

【図10A】

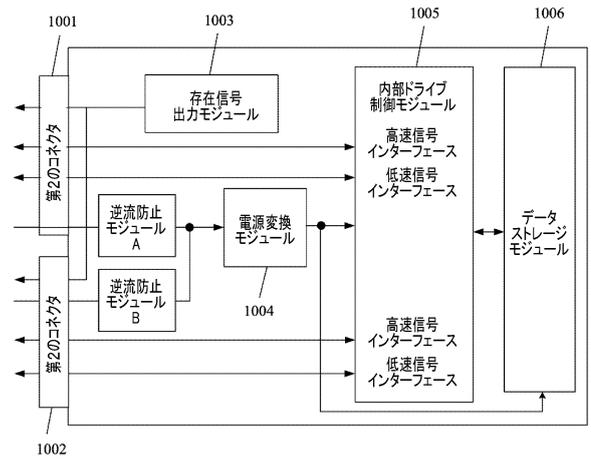


FIG. 10A

【図10B】

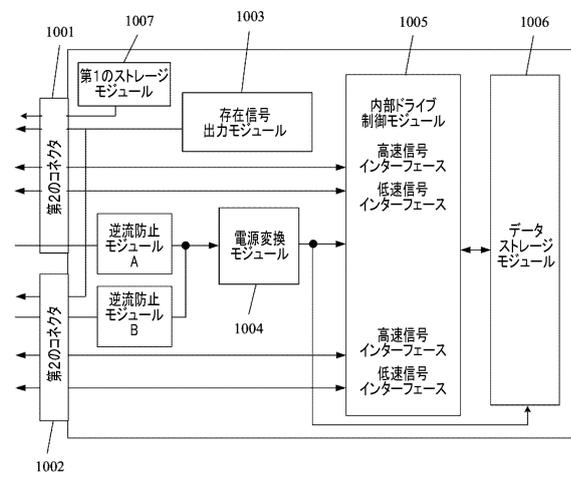


FIG. 10B

【図10C】

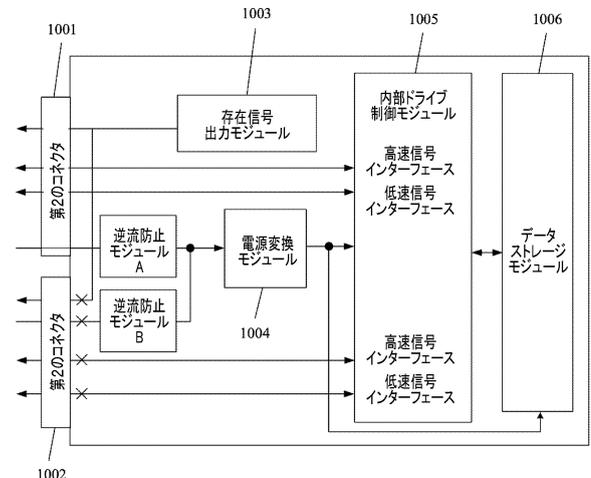


FIG. 10C

10

20

30

40

50

【図10D】

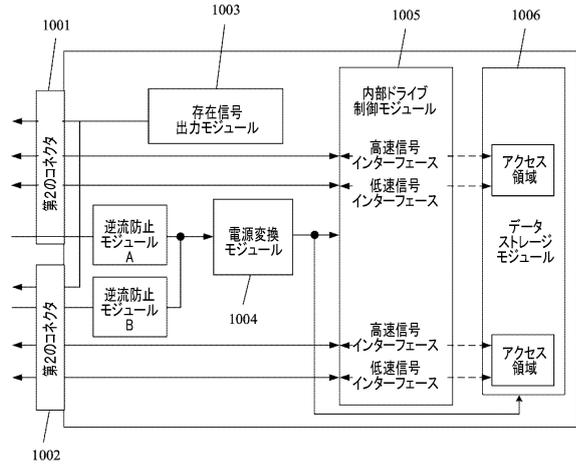


FIG. 10D

【図11】

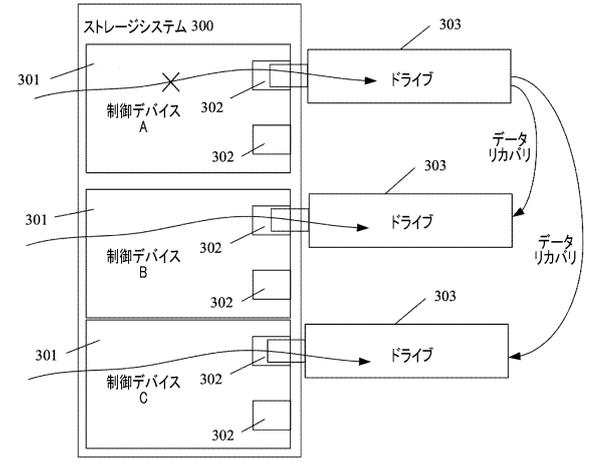


FIG. 11

10

【図12A】

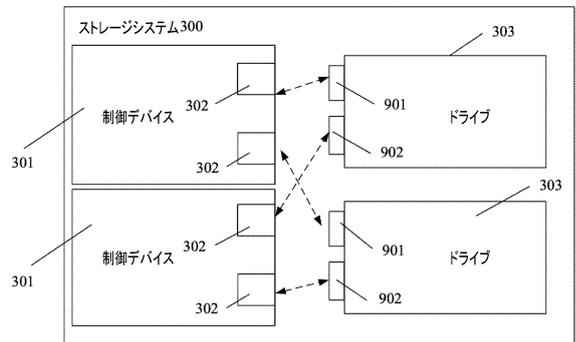


FIG. 12A

【図12B】

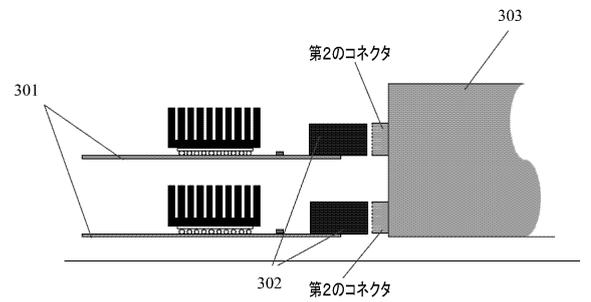


FIG. 12B

20

30

40

50

【図 13 A】

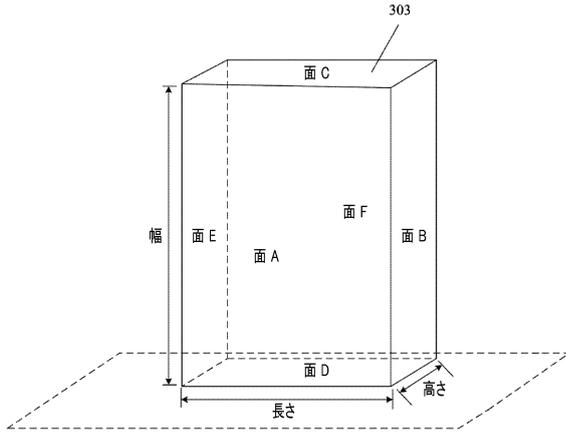


FIG. 13A

【図 13 B】

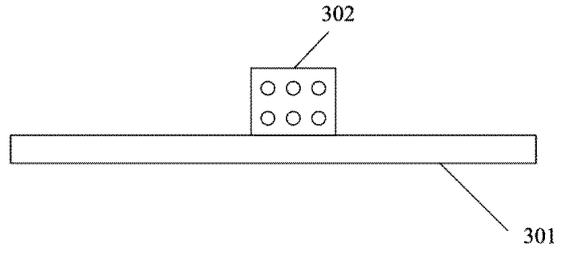


FIG. 13B

【図 13 C】

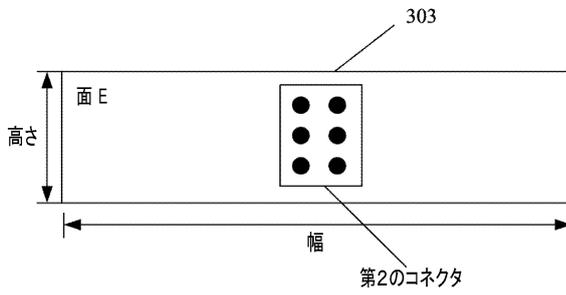


FIG. 13C

【図 13 D】

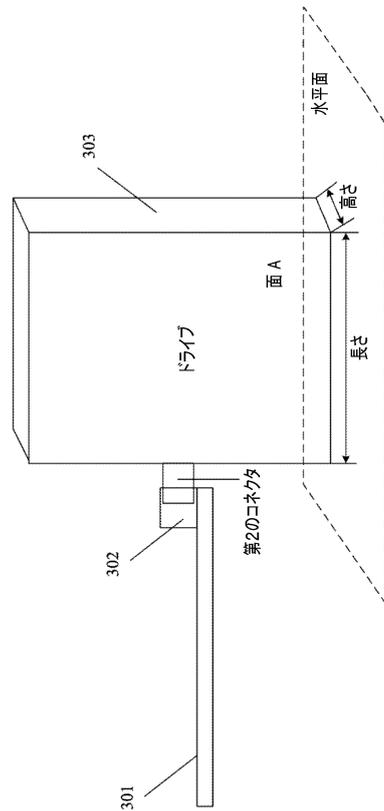


FIG. 13D

10

20

30

40

50

【図14A】

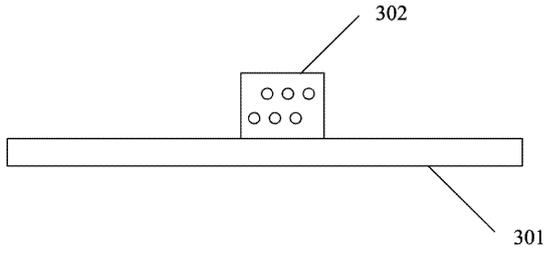


FIG. 14A

【図14B】

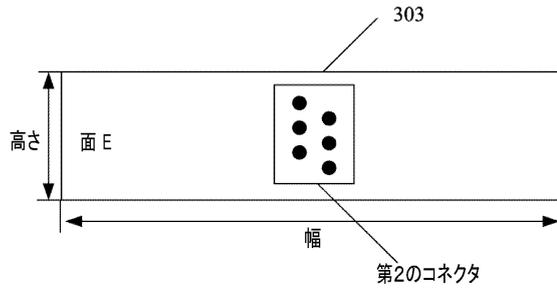


FIG. 14B

10

【図14C】

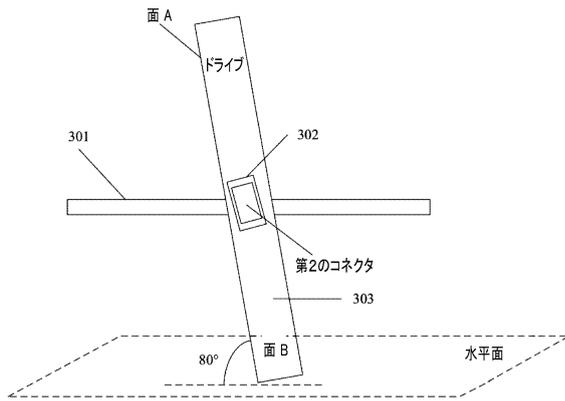


FIG. 14C

【図15A】

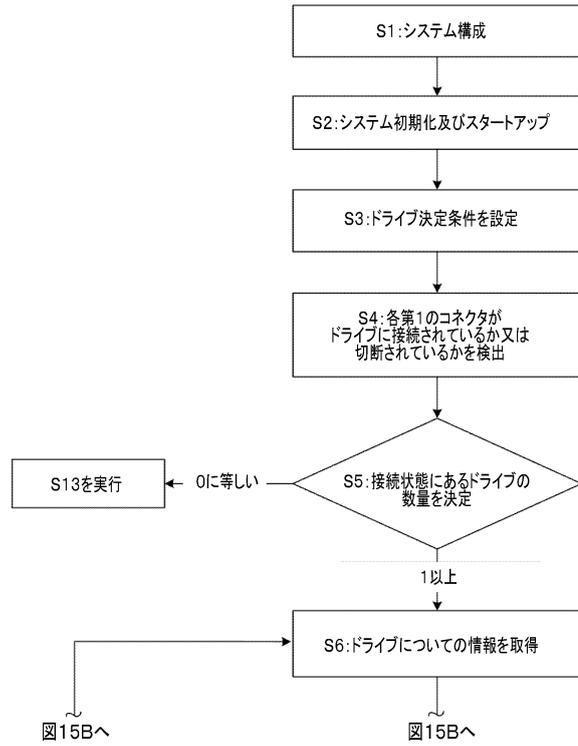


FIG. 15A

20

30

40

50

【 図 1 5 B 】

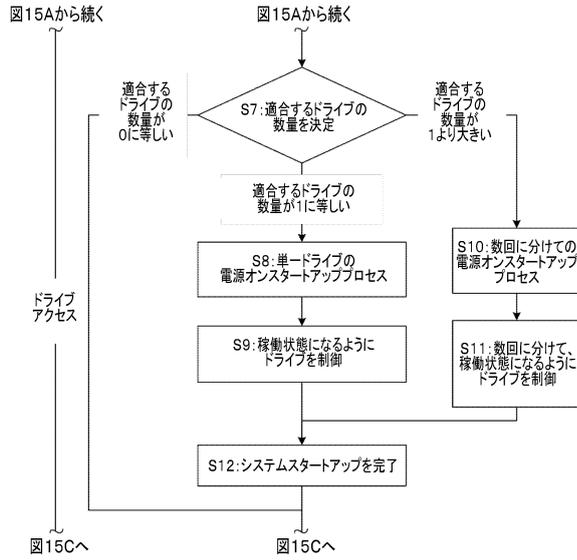


FIG. 15B

【 図 1 5 C 】

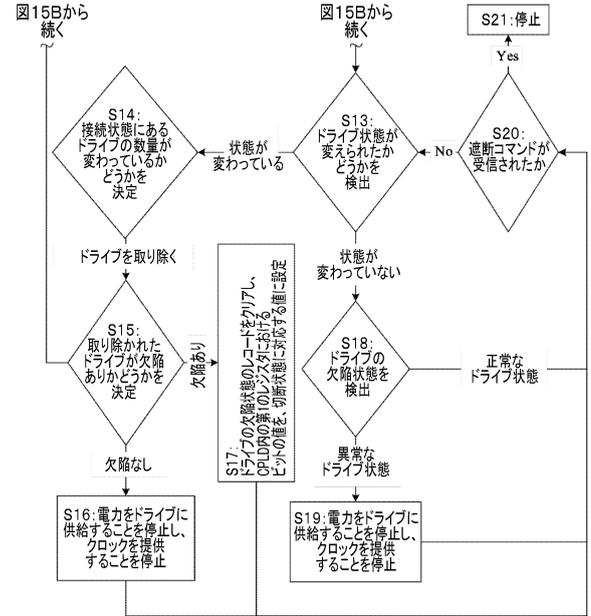


FIG. 15C

【 図 1 6 】

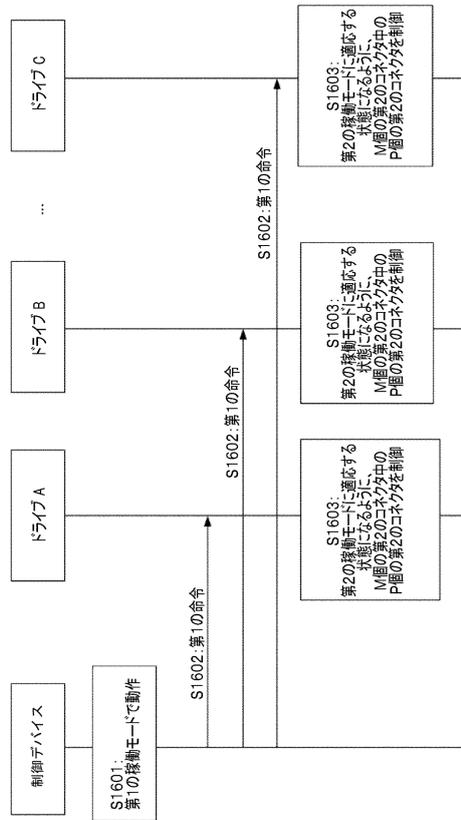


FIG. 16

10

20

30

40

50

フロントページの続き

(51)国際特許分類

F I
G 0 6 F 13/14 3 3 0 E

(74)代理人 100115635

弁理士 窪田 郁大

(72)発明者 陳 燦

中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為総部 ベン 公楼

(72)発明者 李 曉 初

中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為総部 ベン 公楼

(72)発明者 陳 明

中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為総部 ベン 公楼

審査官 打出 義尚

(56)参考文献 国際公開第 2 0 1 8 / 0 5 1 3 8 6 (W O , A 1)

(58)調査した分野 (Int.Cl. , D B 名)

G 0 6 F 3 / 0 6

G 0 6 F 1 / 3 2 2 1

G 0 6 F 1 / 3 2 8 7

G 0 6 F 1 3 / 1 0

G 0 6 F 1 3 / 1 4