

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号
特許第4988213号
(P4988213)

(45) 発行日 平成24年8月1日(2012.8.1)

(24) 登録日 平成24年5月11日(2012.5.11)

(51) Int.Cl.
G06F 3/06 (2006.01)

F I
G06F 3/06 304R

請求項の数 14 (全 15 頁)

(21) 出願番号	特願2006-16925 (P2006-16925)	(73) 特許権者	390009531
(22) 出願日	平成18年1月25日 (2006.1.25)		インターナショナル・ビジネス・マシーンズ・コーポレーション
(65) 公開番号	特開2006-209773 (P2006-209773A)		INTERNATIONAL BUSINESS MACHINES CORPORATION
(43) 公開日	平成18年8月10日 (2006.8.10)		アメリカ合衆国10504 ニューヨーク州 アーモンク ニュー オーチャードロード
審査請求日	平成20年10月14日 (2008.10.14)		
(31) 優先権主張番号	11/045,163	(74) 代理人	100108501
(32) 優先日	平成17年1月28日 (2005.1.28)		弁理士 上野 剛史
(33) 優先権主張国	米国 (US)	(74) 代理人	100112690
			弁理士 太佐 種一
前置審査		(74) 代理人	100091568
			弁理士 市位 嘉宏
			最終頁に続く

(54) 【発明の名称】 ストレージ・デバイスを保守する装置、該装置の動作方法及びコンピュータ読み取り可能な記憶媒体

(57) 【特許請求の範囲】

【請求項1】

ネットワークを通じてホスト・システムに接続されたアダプタ・モジュールと、
通信チャンネルを通じて前記アダプタ・モジュールに接続された、非ブロッキング・スイッチとして構成された相互接続モジュールであって、前記非ブロッキング・スイッチは、
該非ブロッキング・スイッチを通る全ての通信は該非ブロッキング・スイッチを通る他の全ての通信を妨害しないスイッチである、前記相互接続モジュールと、
前記相互接続モジュールに接続された管理モジュールと、
前記相互接続モジュールに接続されたストレージ・デバイスとを備え、
前記管理モジュールは該管理モジュールに前記ストレージ・デバイスに対する保守操作
を実行するように指示するコマンドを前記相互接続モジュールを通じて前記アダプタ・モ
ジュールから受領し、
前記管理モジュールは前記コマンドに応答し、前記相互接続モジュールを通じて前記ス
トレージ・デバイスに対する保守操作を実行し、
前記アダプタ・モジュールは、前記保守操作の状態について前記相互接続モジュールを
通じて前記管理モジュールに照会し、
前記管理モジュールは、前記照会に応答して、前記保守操作の状態を前記相互接続モジ
ュールを通じて前記アダプタ・モジュールに報告し、
前記管理モジュールは、前記保守操作の完了を前記相互接続モジュールを介して前記ア
ダプタ・モジュールに報告し、

10

20

前記管理モジュールによる前記相互接続モジュールを通じて行われる前記ストレージ・デバイスに対する前記保守操作は、前記相互接続モジュールを通る他の通信を妨害しない、ストレージ・デバイスを保守する装置。

【請求項 2】

前記保守操作は、前記ストレージ・デバイスにバイナリ「0」のパターンを書き込む表面初期化を行う保守操作である、請求項 1 に記載の装置。

【請求項 3】

前記保守操作は、前記ストレージ・デバイスをフォーマットする保守操作である、請求項 1 に記載の装置。

【請求項 4】

前記保守操作は、ヘルス・チェックなる保守操作である、請求項 1 に記載の装置。

【請求項 5】

前記管理モジュールが、
前記相互接続モジュールと交信するインターフェース・モジュールと、
前記インターフェース・モジュールと交信するプロセッサ・モジュールと、
前記インターフェース・モジュールおよび前記プロセッサ・モジュールと交信するメモリ・モジュールとを含む、請求項 1 に記載の装置。

【請求項 6】

(A) ネットワークを介してホスト・システムに接続された第 1 アダプタ・モジュールと、

(B) 前記ネットワークを介して前記ホスト・システムに接続された第 2 アダプタ・モジュールと、

(C) 第 1 格納装置であって、

第 1 通信チャンネルを通じて前記第 1 アダプタ・モジュールに接続され、第 2 通信チャンネルを通じて前記第 2 アダプタ・モジュールに接続され、非ブロッキング・スイッチとして構成された第 1 相互接続モジュールであって、前記非ブロッキング・スイッチは、該非ブロッキング・スイッチを通る全ての通信は該非ブロッキング・スイッチを通る他の全ての通信を妨害しないスイッチである、前記第 1 相互接続モジュールと、

前記第 1 相互接続モジュールに接続された第 1 管理モジュールと、

前記第 1 相互接続モジュールに接続された第 1 ストレージ・デバイスとを備える前記第 1 格納装置と、

(D) 第 2 格納装置であって、

第 3 通信チャンネルを通じて前記第 1 相互接続モジュールに接続され、第 4 通信チャンネルを通じて前記第 1 相互接続モジュールに接続され、前記非ブロッキング・スイッチとして構成された第 2 相互接続モジュールと、

前記第 2 相互接続モジュールに接続された第 2 管理モジュールと、

前記第 2 相互接続モジュールに接続された第 2 ストレージ・デバイスとを備える前記第 2 格納装置とを備え、

前記第 1 管理モジュールは該第 1 管理モジュールに前記第 1 ストレージ・デバイスに対する保守操作を実行するように指示するコマンドを前記第 1 相互接続モジュールを通じて前記第 1 アダプタ・モジュールから受領し、

前記第 1 管理モジュールは前記コマンドに応答し、前記第 1 相互接続モジュールを通じて前記第 1 ストレージ・デバイスに対する保守操作を実行し、

前記第 1 アダプタ・モジュールは、前記保守操作の状態について前記第 1 相互接続モジュールを通じて前記第 1 管理モジュールに照会し、

前記第 1 管理モジュールは、前記照会に応答して、前記保守操作の状態を前記第 1 相互接続モジュールを通じて前記第 1 アダプタ・モジュールに報告し、

前記第 1 管理モジュールは、前記保守操作の完了を前記第 1 相互接続モジュールを介して前記第 1 アダプタ・モジュールに報告し、

前記第 2 アダプタ・モジュールは、前記保守操作の間、前記第 2 相互接続モジュール及

10

20

30

40

50

び前記第1相互接続モジュールを通じて、前記第2ストレージ・デバイスからデータを検索し取得する、ストレージ・デバイスを保守する装置。

【請求項7】

前記保守操作は、前記ストレージ・デバイスにバイナリ「0」のパターンを書き込む表面初期化を行う保守操作である、請求項6に記載の装置。

【請求項8】

前記保守操作は、前記ストレージ・デバイスをフォーマットする保守操作である、請求項6に記載の装置。

【請求項9】

前記保守操作は、ヘルス・チェックなる保守操作である、請求項6に記載の装置。

10

【請求項10】

前記第1管理モジュールが、
前記第1の相互接続モジュールと交信するインターフェース・モジュールと、
前記インターフェース・モジュールと交信するプロセッサ・モジュールと、
前記インターフェース・モジュールおよび前記プロセッサ・モジュールと交信するメモリ・モジュールとを含む、請求項6に記載の装置。

【請求項11】

ネットワークを通じてホスト・システムに接続されたアダプタ・モジュールと、
通信チャンネルを通じて前記アダプタ・モジュールに接続された、非ブロッキング・スイッチとして構成された相互接続モジュールであって、前記非ブロッキング・スイッチは、該非ブロッキング・スイッチを通る全ての通信は該非ブロッキング・スイッチを通る他の全ての通信を妨害しないスイッチである、前記相互接続モジュールと、

20

前記相互接続モジュールに接続された管理モジュールと、
前記相互接続モジュールに接続されたストレージ・デバイスとを備えたストレージ・デバイスを保守する装置の動作方法であって、

前記管理モジュールに前記ストレージ・デバイスに対する保守操作を実行するように指示するコマンドを前記相互接続モジュールを通じて前記アダプタ・モジュールから前記管理モジュールが受領するステップと、

前記管理モジュールが前記コマンドの受領を前記アダプタ・モジュールに確認応答するステップと、

30

前記管理モジュールが前記コマンドに応答し、前記相互接続モジュールを通じて前記ストレージ・デバイスに対する保守操作を実行するステップと、

前記アダプタ・モジュールが、前記保守操作の状態について前記相互接続モジュールを通じて前記管理モジュールに照会するステップと、

前記管理モジュールが、前記照会に応答して、前記保守操作の状態を前記相互接続モジュールを通じて前記アダプタ・モジュールに報告するステップと、

前記管理モジュールが、前記保守操作の完了を前記相互接続モジュールを介して前記アダプタ・モジュールに報告するステップとを含み、

前記管理モジュールによる前記相互接続モジュールを通じて行われる前記ストレージ・デバイスに対する前記保守操作は、前記相互接続モジュールを通る他の通信を妨害しない、ストレージ・デバイスを保守する装置の動作方法。

40

【請求項12】

前記保守操作は、前記ストレージ・デバイスにバイナリ「0」のパターンを書き込む表面初期化を行う保守操作である、請求項11に記載の装置。

【請求項13】

ネットワークを通じてホスト・システムに接続されたアダプタ・モジュールと、
通信チャンネルを通じて前記アダプタ・モジュールに接続された、非ブロッキング・スイッチとして構成された相互接続モジュールであって、前記非ブロッキング・スイッチは、該非ブロッキング・スイッチを通る全ての通信は該非ブロッキング・スイッチを通る他の全ての通信を妨害しないスイッチである、前記相互接続モジュールと、

50

前記相互接続モジュールに接続された管理モジュールと、
前記相互接続モジュールに接続されたストレージ・デバイスとを備えたストレージ・デバイスを保守する装置を動作させるプログラムを記憶したコンピュータ読み取り可能な記憶媒体であって、

コンピュータに、

前記管理モジュールに前記ストレージ・デバイスに対する保守操作を実行するように指示するコマンドを前記相互接続モジュールを通じて前記アダプタ・モジュールから前記管理モジュールが受領する手順と、

前記管理モジュールが前記コマンドの受領を前記アダプタ・モジュールに確認応答する手順と、

前記管理モジュールが前記コマンドに応答し、前記相互接続モジュールを通じて前記ストレージ・デバイスに対する保守操作を実行する手順と、

前記アダプタ・モジュールが、前記保守操作の状態について前記相互接続モジュールを通じて前記管理モジュールに照会する手順と、

前記管理モジュールが、前記照会に応答して、前記保守操作の状態を前記相互接続モジュールを通じて前記アダプタ・モジュールに報告する手順と、

前記管理モジュールが、前記保守操作の完了を前記相互接続モジュールを介して前記アダプタ・モジュールに報告する手順を実行させるためのプログラムであって、前記管理モジュールによる前記相互接続モジュールを通じて行われる前記ストレージ・デバイスに対する前記保守操作は、前記相互接続モジュールを通る他の通信を妨害しない、前記プログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【請求項 14】

前記保守操作は、前記ストレージ・デバイスにバイナリ「0」のパターンを書き込む表面初期化を行う保守操作である、請求項 13 に記載のコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

この発明はストレージ・デバイスの保守操作を実行することに関し、特に保守操作を実行するローカルの管理モジュールに関する。

【背景技術】

【0002】

コンピュータのストレージ・システムは通常、ハード・ディスク駆動装置、磁気テープ駆動装置、光学ストレージ駆動装置などのストレージ・デバイスを複数個含んでいる。各ストレージ・デバイスは通常、通信チャネルを通じて、ここではアダプタ・モジュールと呼ぶ制御デバイスと交信する。アダプタ・モジュールは少なくとも1つのホスト・システムと交信する。さらに、アダプタ・モジュールはホスト・システムによるストレージ・デバイスに対するアクセスを制御すると共に調停する。

【0003】

例えば、ホスト・システムはストレージ・デバイスからデータを検索・取得する要求をアダプタ・モジュールに伝達する。すると、アダプタ・モジュールはストレージ・デバイスに要求されたデータを該アダプタ・モジュールに伝達するように指示する。そして、アダプタ・モジュールは該データをホスト・システムに伝達する。同様に、ホスト・システムはストレージ・デバイスにおける格納のためにアダプタ・モジュールにデータを伝達する。アダプタ・モジュールはストレージ・デバイスに該データを受領し、該データを特定の場所に格納するように指示する。

【0004】

アダプタ・モジュールは通常、ストレージ・デバイスに対する保守操作をも実行する。例えば、アダプタ・モジュールはストレージ・デバイスに対してヘルス・チェックなる保守操作を実行し、当該ストレージ・デバイスが正常に機能していることを確認する。ある

10

20

30

40

50

いは、アダプタ・モジュールは例えば当該ストレージ・デバイスがストレージ・デバイスのアレイに付加された時に、当該ストレージ・デバイスについて初期化操作を実行する。

【0005】

アダプタ・モジュールは通常、複数のストレージ・デバイスを制御する。例えば、アダプタ・モジュールは「RAID」(redundant Array of Independent Drives:独立した駆動装置から成る冗長なアレイ。レイド)を構成する複数のハード・ディスク駆動装置を制御する。これら複数のストレージ・デバイスを制御する一方で、アダプタ・モジュールは同時にホスト・システムのために少なくとも1つのストレージ・デバイスにアクセスすると共に少なくとも1つのストレージ・デバイスに対して保守操作を実行する。あいにく、アダプタ・モジュールはストレージ・デバイスに対するアクセスおよび制御と保守操作との双方を実行する処理帯域幅を欠いている。加えて、保守操作の間における通信チャネルの性能をさらに向上することが望ましい。

10

【発明の開示】

【発明が解決しようとする課題】

【0006】

上述した検討からストレージ・デバイスの保守操作をローカルに実行する装置、システム、および方法を求める必要性が存在することは明らかなはずである。そのような装置、システム、および方法は保守操作を実行しながら、ストレージ・デバイスに対するアクセスを増大せうる点で有利である。

【課題を解決するための手段】

20

【0007】

本発明は当技術分野の現状に呼応して、特に、現在利用可能な保守方法によって未だに完全には解決されていない当技術分野における課題と要求に呼応してなされた。したがって、本発明は当技術分野における上述した弱点の多く、または全てを克服する、保守操作を実行する装置、システム、および方法を提供するためになされた。

【0008】

相互接続モジュールを通じてコマンドを受領する必要なステップと、相互接続モジュールを通じて保守操作を実行する必要なステップとを機能的に実行するように構成された複数のモジュールを含む論理ユニットを用いて保守操作を実行する装置を提供する。説明される実施形態中のこれらのモジュールは相互接続モジュールと管理モジュールを含む。

30

【0009】

相互接続モジュールは非ブロッキング・スイッチとして構成される。ここで使用するように、非ブロッキング・スイッチは当該スイッチを通る全ての通信は当該スイッチを通る他の全ての通信を妨害しないという具合に構成される。相互接続モジュールはストレージ・デバイスおよびアダプタ・モジュールと交信する。アダプタ・モジュールは相互接続モジュールを通じてストレージ・デバイスを制御し、相互接続モジュールを通じてストレージ・デバイスにデータを格納し、そこからデータを検索・取得する。

【0010】

管理モジュールも相互接続モジュールと交信する。管理モジュールはさらに、該管理モジュールにストレージ・デバイスに対する保守操作を実行するように指示するコマンドを相互接続モジュールを通じてアダプタモジュールから受領する。管理モジュールは該コマンドに応答し、相互接続モジュールを通じてストレージ・デバイスに対する保守操作を実行する。本装置は保守操作をローカルに実行し、通信帯域幅とアダプタモジュールの処理帯域幅を解放する。

40

【0011】

本発明のシステムも保守操作を実行するように提示される。本システムはストレージ・システムにおいて実現される。特に、本システムは一実施形態において、アダプタ・モジュールと、ストレージ・デバイス、第1の相互接続モジュール、および管理モジュールから成る第1の格納装置とを含む。

【0012】

50

アダプタ・モジュールは第1の相互接続モジュールを通じて格納装置のストレージ・デバイスと管理モジュールを制御する。アダプタ・モジュールはさらに、複数の格納装置を制御する。一実施形態では、各格納装置の相互接続モジュールは少なくとも1つの他の相互接続モジュールと交信する。アダプタ・モジュール第1の格納装置の第1の相互接続モジュールと交信し、第1の相互接続モジュールを通じて第2の格納装置の第2の相互接続モジュールと交信する。

【0013】

相互接続モジュールは非ブロッキング・スイッチとして構成される。管理モジュールは該管理モジュールにストレージ・デバイスに対する保守操作を実行するように指示するコマンドを相互接続モジュールを通じてアダプタ・モジュールから受領する。管理モジュールは該コマンドに応答し、相互接続モジュールを通じてストレージ・デバイスに対する保守操作を実行する。

10

【0014】

保守操作を実行する、本発明の方法も提示される。開示される実施形態における方法は記述済みの装置とシステムの操作について上で提示した機能を実行するのに必要なステップを実質的に含む。一実施形態では、本方法は相互接続モジュールを通じてコマンドを受領するステップと、相互接続モジュールを通じて保守操作を実行するステップとを含む。

【0015】

まず、管理モジュールが、非ブロッキング・スイッチとして構成された相互接続モジュールを通じてコマンドを受領する。管理モジュールは該コマンドに応答し、相互接続モジュールを通じてストレージ・デバイスに対して保守操作を実行する。一実施形態では、管理モジュールはさらに、相互接続モジュールを通じて保守操作の状態の照会を受領する。管理モジュールは相互接続モジュールを通じて保守操作の状態の報告をも行う。

20

【0016】

この明細書を通じて、特徴、利点、または類似の言い回しに対する言及は、本発明で実現される特徴と利点の全てが本発明の任意の単一の実施形態中に存在するはずである、あるいは存在するということを含意しない。むしろ、特徴と利点に言及する言い回しは、実施形態と共に記述される特定の特徴、利点、または特性は本発明の少なくとも1つの実施形態の中に含まれるということの意味するように理解される。したがって、この明細書を通じて、特徴および利点ならびに類似の言い回しの記述は必然ではないが、同一の実施形態を指している。

30

【0017】

さらに、記述される本発明の特徴、利点、および特性は少なくとも1つの実施形態において任意の好適な態様で組み合わせられる。本発明は特定の実施形態の特定の特徴または利点のうちの少なくとも1つがなくとも実施するというを当業者は認識する。他の例では、本発明の全ての実施形態に存在するわけではない付加的な特徴と利点がいくつかの実施形態には認められる。

【0018】

本発明は相互接続モジュールを通じて受領されるコマンドに応答し、非ブロッキング型の相互接続モジュールを通じてストレージ・デバイスに対する保守操作をローカルに実行する。さらに、本発明は相互接続モジュールを通る通信の帯域幅と、相互接続モジュールと交信するアダプタ・モジュールの処理帯域幅とを解放する。本発明のこれらの特徴と利点は以下の記述と特許請求の範囲から完全に明らかになる、あるいは、以下に提示する本発明の実施によって知得される。

40

【発明を実施するための最良の形態】

【0019】

この明細書で記述される機能ユニット群（以下、複数であることを表現する場合に「群」なる表現を用いる）の多くはモジュールと呼ばれているが、それはそれらの実装の独立性を特に強調するためである。例えば、モジュールはカスタムの超大規模集積（「VLSI」）回路もしくはゲート・アレイ、論理チップ、トランジスタ、または他の個別素子の

50

ような既製の半導体を含むハードウェア回路として実装されうる。モジュールはフィールド・プログラマブル・ゲート・アレイ、プログラマブル・アレイ・ロジック、プログラマブル論理デバイスなどのようなプログラム可能なハードウェア・デバイス中にも実装されうる。

【0020】

モジュールは様々な種類のプロセッサによる実行のためにソフトウェアにおいても実装されうる。実行可能なコードから成る特定のモジュールは例えばコンピュータ命令（これは例えばオブジェクト、プロシージャ、または機能として構成される）から成る少なくとも1つの物理ブロックまたは論理ブロックを含む。それでも、特定のモジュールの実行可能プログラムは物理的に共存している必要はないが、論理的に結合されると当該モジュールを構成し、当該モジュールのための規定された目的を達成する、異なる複数の場所に格納された本質的に異なる命令を含む。

10

【0021】

無論、実行可能なコードから成るモジュールは単一の命令であってもよいし、あるいは多数の命令であってもよい、さらに、いくつかの異なるコード・セグメントに亘って分散していてもよいし、異なる複数のプログラム中に分散していてもよいし、いくつかのメモリ・デバイスの全体に亘って分散していてもよい。同様に、ここでは、オペレーショナル・データはモジュール内にあるように特定されると共に記述される。オペレーショナル・データは単一のデータとして集約されていてもよいし、あるいは、異なる複数の場所（例えば異なる複数のストレージ・デバイス）に亘って分散していてもよいし、さらに、少なくとも部分的にシステムまたはネットワーク上の単なる電子信号として存在していてもよい。

20

【0022】

この明細書を通じて、「一実施形態」、「実施形態」、または類似の用語を参照することは、実施形態と共に説明される特定の特徵、構造、または特性が本発明の少なくとも1つの実施形態に含まれているということを意味する。したがって、この明細書を通じて、「一実施形態において」や「実施形態において」なる句、および類似の言い回しが出現したときは、必然ではないが、すべて同一の実施形態を参照している。

【0023】

さらに、本発明の記述された特徵、構造、または特性は少なくとも1つの実施形態において何らかの好適な態様で組み合わせられうる。以下の記述において、本発明の実施形態の完全な理解を提供するために、多数の特定の細部、例えばプログラミング、ソフトウェア・モジュール、ユーザ設定、ネットワーク・トランザクション、データベースの照会、データベースの構造、ハードウェア・モジュール、ハードウェア回路、ハードウェア・チップなどの例が提示される。しかしながら、本発明は特定の細部のうちの少なくとも1つがなくとも実施されうる、あるいは、他の方法、構成要素、材料を用いてこれを実施されうる、等々のことを当業者は認識しうる。他の例では、本発明の諸側面を曖昧にするのを避けるために、周知の構造、材料、または操作は詳細には示されていない、あるいは記述されていない。

30

【0024】

図1は本発明の保守システム100の一実施形態を示す概略ブロック図である。保守システム100はアダプタ・モジュール125と格納装置120を含む。格納装置120はさらに管理モジュール105、相互接続モジュール110、およびストレージ・デバイス115を含む。

40

【0025】

アダプタ・モジュール125は相互接続モジュール110を通じて格納装置120のストレージ・デバイス115と管理モジュール105を制御する。例えば、アダプタ・モジュール125はストレージ・デバイス115にデータを格納し、それからデータを検索・取得する。アダプタ・モジュール125は通信チャンネル130を通じて相互接続モジュール110と交信する。一実施形態では、通信チャンネル130はワシントン特別区の米国規

50

格協会によって規定されているようなファイバ・チャネル調停ループである。

【0026】

一実施形態では、アダプタ・モジュール125はホスト・システム135と交信している。アダプタ・モジュール125とホスト・システム135はネットワーク140を通じて交信する。アダプタ・モジュール125はホスト・システム135からストレージ・デバイス115にアクセスする要求を受領する。例えば、ホスト・システム135はストレージ・デバイス115の特定の場所からデータを検索・取得するようアダプタ・モジュール125に要求する。アダプタ・モジュール125はストレージ・デバイス115に要求されたデータを検索・取得するように指示し、該要求されたデータをホスト・システム135に伝達する。

10

【0027】

過去においては、アダプタ・モジュール125はストレージ・デバイス115に対する保守操作も実行していた。あいにく、アダプタ・モジュール125がストレージ・デバイス115に対する保守操作を実行すると、アダプタ・モジュール125の処理帯域幅と通信チャネル130の通信帯域幅が低減する。本発明は保守操作を格納装置120内においてローカルに実行して、アダプタ・モジュール125に対する処理要求と通信チャネル130に対する通信要求を低減する。

【0028】

相互接続モジュール110は非ブロッキング・スイッチとして構成される。ここで使用しているように、非ブロッキング・スイッチは当該スイッチを通るすべての通信は当該スイッチを通る他のすべての通信を妨害しないという具合に構成される。一実施形態では、相互接続モジュール110は非ブロッキングのファイバ・チャネル調停ループ型のスイッチとして構成される。

20

【0029】

アダプタ・モジュール125は相互接続モジュール110を通じて管理モジュール105にコマンドを送付する。このコマンドは管理モジュール105にストレージ・デバイス115に対する保守操作を実行するように指示する。管理モジュール105はこのコマンドに応答し、相互接続モジュール110を通じてストレージ・デバイス115に対する保守操作を実行する。アダプタ・モジュール125は保守操作を実行することから解放される。さらに、通信チャネル130は通信チャネル130を通じて保守操作を実行するのに要するアダプタ・モジュール125の帯域幅から解放される。

30

【0030】

図2は本発明の保守装置200の一実施形態を示す概略ブロック図である。保守装置200は相互接続モジュール110と、プロセッサ・モジュール205、メモリ・モジュール210、およびインターフェース・モジュール215を含む管理モジュール105とを含む。保守装置200は1個の管理モジュール105と1個の相互接続モジュール110を備えるように示されているけれども、任意個数の管理モジュール105と任意個数の相互接続モジュール110を使用しうる。

【0031】

相互接続モジュール110は管理モジュール105を含む複数のデバイスおよび通信チャネル130と交信する。さらに、相互接続モジュール110は各デバイスおよび通信チャネル130ごとにポートを備えている。相互接続モジュール110はさらに非ブロッキング・スイッチとして構成される。例えば、相互接続モジュール110は各ポートを他の各ポートに接続するのに十分な複数の内部通信チャネルを備えている。相互接続モジュール110は複数のデバイスと通信チャネルに関係する複数の通信操作を実行する。

40

【0032】

メモリ・モジュール210は少なくとも1つのソフトウェア・プログラムとデータを格納する。プロセッサ・モジュール205は当業者にとって周知のように、ソフトウェア・プログラムを実行し、データを処理する。プロセッサ・モジュール205はインターフェース・モジュール215を通じて相互接続モジュール110と交信する。一実施形態では

50

、プロセッサ・モジュール 205 は相互接続モジュール 110 とインターフェース・モジュール 215 を通じてアダプタ・モジュール 125 からコマンドを受領する。プロセッサ・モジュール 205 はこのコマンドを特定の保守操作を実行するための指令としてデコードする。この特定の保守操作は少なくとも 1 つのソフトウェア・プログラムと少なくとも 1 つデータ・ワードを含む。

【0033】

プロセッサ・モジュール 205 は特定の保守操作のソフトウェア・プログラムのソフトウェア命令を検索・取得し、実行し、そして当該特定の保守操作のデータを処理する。ソフトウェア・プログラムとデータに応答して、プロセッサ・モジュール 205 はインターフェース・モジュール 215 と相互接続モジュール 110 を通じてストレージ・デバイス 115 にコマンドとデータを伝達する。伝達されたコマンドとデータはストレージ・デバイス 115 に特定の保守操作に合致するアクションを実行するように指示する。

10

【0034】

例えば、保守操作のソフトウェア・プログラムとデータに応答して、プロセッサ・モジュール 205 はストレージ・デバイス 115 に表面初期化する保守操作の一部としてストレージ・デバイス 115 のストレージ媒体の少なくとも 1 つの物理的な場所にバイナリ・パターン（例えば、すべてバイナリ「0」）を書き込むように指示する。管理モジュール 105 とストレージ・デバイス 115 間の通信は相互接続モジュール 110 の非ブロッキング性のために、相互接続モジュール 110 を通る他の通信を些かも妨害しない。保守装置 200 は相互接続モジュール 110 の他の通信操作の性能を劣化させることなく、保守操作を実行する。

20

【0035】

図 3 は本発明の複数格納装置型保守システム 300 の一実施形態を示す概略ブロック図である。複数格納装置型保守システム 300 は少なくとも 1 つのアダプタ・モジュール 125、少なくとも 1 つの通信チャンネル 130、および、各々が管理モジュール 105、相互接続モジュール 110、およびストレージ・デバイス 115 を含む少なくとも 1 つの格納装置 120 を含む。図示した複数格納装置型保守システム 300 は 2 つのアダプタ・モジュール 125 と 2 つの格納装置 120 を含むけれども、任意個数のアダプタ・モジュール 125 および格納装置 120 を用いることができる。

【0036】

30

第 1 のアダプタ・モジュール 125 a および第 2 のアダプタ・モジュール 125 b はそれぞれ第 1 の通信チャンネル 130 a および第 2 の通信チャンネル 130 b を通じて第 1 の相互接続モジュール 110 a と交信する。第 1 の格納装置 120 a の第 1 の相互接続モジュール 110 a は第 3 の通信チャンネル 130 b および第 4 の通信チャンネル 130 d を通じて第 2 の格納装置 120 b の第 2 の相互接続モジュール 110 b と交信している。アダプタ・モジュール 125 は第 1 の相互接続モジュール 110 a を通じて第 2 の相互接続モジュール 110 b と交信する。例えば、第 1 のアダプタ・モジュール 125 a が第 1 の相互接続モジュール 110 a を通じて第 1 のストレージ・デバイス 115 a にデータを格納する一方で、第 2 のアダプタ・モジュール 125 b が第 2 の相互接続モジュール 110 b および第 1 の相互接続モジュール 110 a を通じて第 2 のストレージ・デバイス 115 b からデータを検索・取得する。

40

【0037】

アダプタ・モジュール 125 は保守操作を実行するコマンドを管理モジュール 105 に伝達する。例えば、第 1 のアダプタ・モジュール 125 a は第 2 の通信チャンネル 130 b と第 1 の相互接続モジュール 110 a を通じてヘルス・チェックなる保守操作を第 1 の管理モジュール 105 a に伝達する。第 1 の管理モジュール 105 a は第 1 の相互接続モジュール 110 a を通じて、このヘルス・チェックなる保守操作を実行する。

【0038】

相互接続モジュール 110 は非ブロッキング・スイッチとして構成されているから、第 1 の管理モジュール 105 a は例えば第 1 のアダプタ・モジュール 125 a と第 2 のアダ

50

プタ・モジュール125bとの間における第1の相互接続モジュール110aを通る通信を妨害することなく、第1のストレージ・デバイス115aに対する保守操作を実行しうる。例えば、第2のアダプタ・モジュール125bは第1の管理モジュール105aによるヘルス・チェックなる保守操作を妨害することなく、第2の相互接続モジュール110bと第1の相互接続モジュール110aを通じて第2のストレージ・デバイス115bからデータを検索・取得する。

【0039】

さらに、第1の管理モジュール105aのヘルス・チェックなる保守操作は第1のアダプタ・モジュール125aおよび第2のアダプタ・モジュール125bの帯域幅を必要としないし、通信チャンネル130の帯域幅も必要としない。複数格納装置型保守システム300はストレージ・デバイス115に対する保守操作をローカルに実行して、該システム300の他の構成要素に対する要求を低減している。

【0040】

図4は本発明の管理モジュール400の一実施形態を示す概略ブロック図である。プロセッサ・モジュール205、メモリ・モジュール210、およびインターフェース・モジュール215がデジタル電気バス405を共用している。デジタル電気バス405は複数のデータ信号線、複数のアドレス信号線、および少なくとも1つの制御信号線を含む。プロセッサ・モジュール205、メモリ・モジュール210、およびインターフェース・モジュール215は各々、少なくとも1つのレジスタ、メモリ記憶位置、または通信ポートを含み、各々は一意のバイナリ・アドレスを備えている。

【0041】

プロセッサ・モジュール205はアドレス・バスを構成する少なくとも1つのアドレス信号線上のターゲットのバイナリ・アドレス信号線をアサートし、そして読み取り制御信号線をアサートして、ターゲットに該ターゲットのバイナリ値を用いてバス405のデータ信号線を駆動させることにより、ターゲット・レジスタ、メモリ配置位置、または通信ポートからデータを移動させる。プロセッサ・モジュール205はバス405のデータ信号線からのバイナリ値をレジスタ、メモリ配置位置、または通信ポートに書き込むことにより、データを検索・取得する。さらに、プロセッサ・モジュール205はターゲットのバイナリ・アドレス信号線をアサートし、バス405のデータ信号線上のデータのバイナリ値をアサートし、書き込み信号線をアサートして、ターゲットにバイナリ値を受領させることにより、ターゲット・レジスタ、メモリ配置位置、または通信ポートにデータを移動させる。

【0042】

プロセッサ・モジュール205は保守操作を構成する少なくとも1つのソフトウェア・プログラムを実行する。一実施形態では、プロセッサ・モジュール205はインターフェース・モジュール215と相互接続モジュール110を通じてストレージ・デバイス115にコマンドとデータを伝達することにより、保守操作を実行する。プロセッサ・モジュール205はコマンドとデータを含むデータ・グループをインターフェース・モジュール215に宛てる。このデータ・グループはストレージ・デバイス115のアドレスと、ストレージ・デバイス115内にあるレジスタ、メモリ配置位置、または通信ポートの内部アドレスとから成るアドレスを含む。ストレージ・デバイス115は相互接続モジュール110を通じて、インターフェース・モジュールのアドレスとバス405のアドレスとを含むデータ・グループ中のデータと状態情報をも伝達する。

【0043】

以下に示す概略フローチャート図は一般に論理フローチャート図として提示される。したがって、図示される順序と符号が付されたステップは提示される方法の一実施形態を表すものである。図示した方法の少なくとも1つのステップ（または、その部分群）と機能、論理、または効果において等価である他のステップと方法が考えられる。さらに、使用される形式と記号は本方法の論理ステップを説明するために提示されており、本方法の範囲を限定するようには理解されない。本フローチャート図には様々な種類の矢印と線が

使用されているけれども、それら是对應する方法の範囲を限定するようには理解されない。無論、本方法の論理フローのみを示すために、何らかの矢印または他の結合子を使用されうる。例えば、ある矢印は図示する方法の列挙されたステップ間の不特定期間の待機期間または監視期間を表す。さらに、特定の方法が行われる順序は図示された対応するステップの順序に厳密に準拠している場合もあるし、そうでない場合もある。

【0044】

図5は本発明に従う保守方法500の一実施形態を示す概略フローチャート図である。まず、保守方法500の下で、管理モジュール105が相互接続モジュール110を通じてコマンドを受領する(505)。相互接続モジュール110は非ブロッキング・スイッチとして構成されている。一実施形態では、コマンドは保守コマンド識別子を含む。例えば、保守コマンド識別子は保守操作の実行を開始するように構成された実行可能ファイルのファイル名を含む。

10

【0045】

コマンドはさらに少なくとも1つのパラメータを含む。このパラメータは保守操作の実行を変更する。例えば、第1のパラメータと第2のパラメータはストレージ・デバイス115のアドレス範囲を指定する。別の例では、第3のパラメータは保守操作を開始することになる時刻を指定する。一実施形態では、管理モジュール105はコマンドの受領を確認応答する(510)。例えば、アダプタ・モジュール125が管理モジュール105にコマンドを伝達した場合、管理モジュール105はアダプタ・モジュール125に当該コマンドの受領を確認応答する(510)。

20

【0046】

管理モジュール105は相互接続モジュール110を通じてストレージ・デバイス115と通信することにより、保守操作を実行する。一実施形態では、管理モジュール105はフォーマットなる保守操作を実行する(515)。このフォーマットなる保守操作はストレージ・デバイス115にセクタとメタデータの情報を書き込むこと、および、ファイル・システムをサポートしうるようにストレージ・デバイス115を構成することを含む。別の実施形態では、管理モジュール105はストレージ・デバイス115に対して表面初期化なる保守操作を実行する(515)。この表面初期化なる保守操作はストレージ・デバイス115に特定のデータ・パターン(例えばバイナリ「0」から成るパターン)を書き込むことを含む。

30

【0047】

管理モジュール105はストレージ・デバイス115に対して認証なる保守操作をも実行する。一実施形態では、この認証なる保守操作はストレージ・デバイス115のデータの一部に対して冗長データ値(例えばチェックサム)を計算すること、および、計算した冗長データ値を、事前に計算され、ストレージ・デバイス115に格納されている格納済み冗長データ値と比較することを含む。計算された冗長データ値が格納済み冗長データ値に等しい場合、ストレージ・デバイス115は認証される。

【0048】

一実施形態では、管理モジュール105はストレージ・デバイス115に対してヘルス・チェックなる保守操作を実行する(515)。このヘルス・チェックなる保守操作はストレージ・デバイス115の故障の公算を見積もるように構成された少なくとも1つの操作を含む。

40

【0049】

一実施形態では、アダプタ・モジュール125は保守操作の状態について管理モジュール105に照会する(520)。例えば、アダプタ・モジュール125は管理モジュール105によってフォーマットが完了された、ストレージ・デバイス115の量を求めて照会する(520)。

【0050】

一実施形態では、管理モジュール105はアダプタ・モジュール125に保守操作の状態を報告する(525)。管理モジュール105はアダプタ・モジュール125の照会(

50

520) に応答して該状態を報告する(525)。あるいは、管理モジュール105は特定の時間間隔で該状態を報告する(525)。例えば、管理モジュール105は500ミリ秒(500ms)ごとに該状態を報告するさらに、管理モジュール105は保守操作の完了時に該保守操作の状態を報告する(525)。例えば、管理モジュール105はアダプタ・モジュール125に表面初期化する保守操作の完了を報告する(525)。

【0051】

管理モジュール105は非ブロッキング・スイッチとして構成された相互接続モジュール110を通じて保守操作を実行する(515)から、保守方法500は他の操作に対する影響を最小限にした状態で保守操作の実行(515)をローカル化しうる。モジュール(例えばアダプタ・モジュール125や通信チャネル130)のリソースは稀にしか必要にならないから、システム全体の性能が向上する。

10

【0052】

図6は本発明の別の保守システム600の一実施形態を示す概略フローチャート図である。図示するように、保守システム600は複数の格納装置120を含む。第1の格納装置120aは管理モジュール105を含むが、第2の格納装置120bは管理モジュール105を含まない。図示した保守システム600は管理モジュール105を備えた1つの第1の格納装置120aと、管理モジュール105を備えない1つの第2の格納装置120bとを含むけれども、該システムは管理モジュール105を備えた任意個数の格納装置120と管理モジュール105を備えない任意個数の格納装置120とを含むうる。

20

【0053】

アダプタ・モジュール125は管理モジュール105に保守操作を実行する(515)ように指示するコマンドを管理モジュール105に伝達する。このコマンドは管理モジュール105が第1のストレージ・デバイス115a、第2のストレージ・デバイス115b、または、第1のストレージ・デバイス115aおよび第2のストレージ・デバイス115bの双方に対して保守操作を実行する(515)ことを指示する。

【0054】

例えば、アダプタ・モジュール125は管理モジュール105に第2のストレージ・デバイス115bに対してヘルス・チェックなる保守操作を実行するように指示するコマンドを伝達する。管理モジュール105は第1の相互接続モジュール110aと第2の相互接続モジュール110bを通じて第2のストレージ・デバイス115bに対するヘルス・チェックなる保守操作を実行する(515)。本システム600はアダプタ・モジュール125に対する処理要求を低減しながら、保守操作をローカルに実行する。

30

【0055】

本発明は相互接続モジュール110を通じて受領したコマンドに応答してストレージ・デバイスに対する保守操作を非ブロッキングの相互接続モジュール110を通じてローカルに実行する最初のものである。さらに、本発明は相互接続モジュール110を通る通信帯域幅および通信チャネル130を通る通信帯域幅、ならびに、相互接続モジュール110と交信しているアダプタ・モジュール125の処理帯域幅を解放する。本発明はその本旨または本質的特徴を逸脱することなく、他の特定の態様で具体化されうる。上述した実施形態はすべての側面において説明のみを目的としたものであり、限定を目的としていないものとして考えられるべきである。したがって、本発明の範囲は上述した記述によってではなく、特許請求の範囲によって示される。特許請求の範囲の等価の意義と範囲の内に入るすべての変更はその範囲内に包含されることになる。

40

【図面の簡単な説明】

【0056】

【図1】本発明による保守システムの一実施形態を示す概略ブロック図である。

【図2】本発明の保守装置の一実施形態を示す概略ブロック図である。

【図3】本発明の複数格納装置型保守システムの一実施形態を示す概略ブロック図である。

。

【図4】本発明の管理モジュールの一実施形態を示す概略ブロック図である。

50

【図 5】 本発明による保守方法の一実施形態を示す概略フローチャート図である。

【図 6】 本発明の別の保守システムの一実施形態を示す概略ブロック図である。

【符号の説明】

【0057】

105 管理モジュール

105 a 第1の管理モジュール

105 b 第2の管理モジュール

110 相互接続モジュール

110 a 第1の相互接続モジュール

110 b 第2の相互接続モジュール

115 ストレージ・デバイス

115 a 第1のストレージ・デバイス

115 b 第2のストレージ・デバイス

120 格納装置

120 a 第1の格納装置

120 b 第2の格納装置

125 アダプタ・モジュール

125 a 第1のアダプタ・モジュール

125 b 第2のアダプタ・モジュール

135 ホスト・システム

205 プロセッサ・モジュール

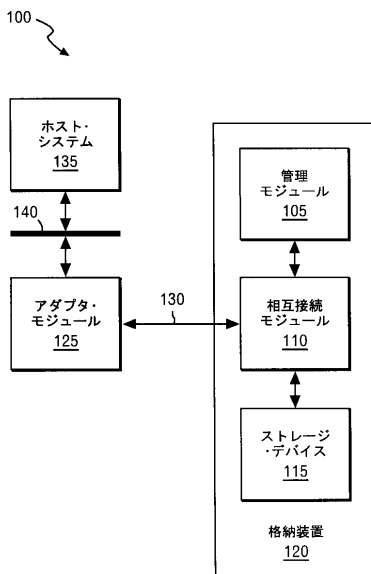
210 メモリ・モジュール

215 インターフェース・モジュール

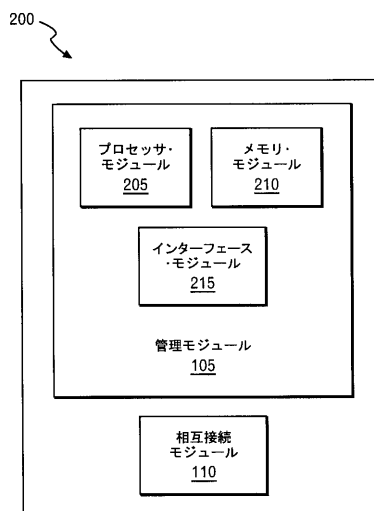
220 管理モジュール

225 相互接続モジュール

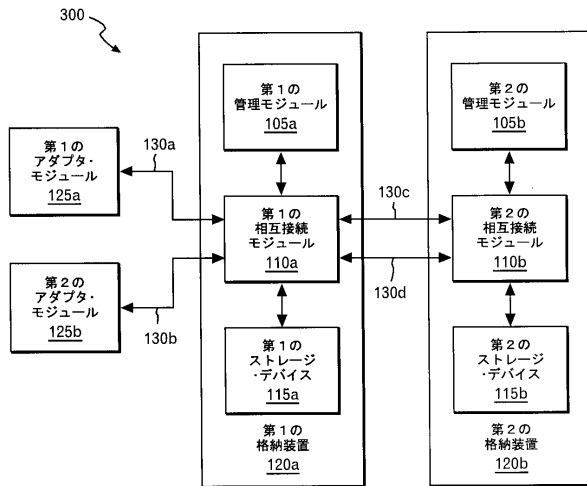
【図 1】



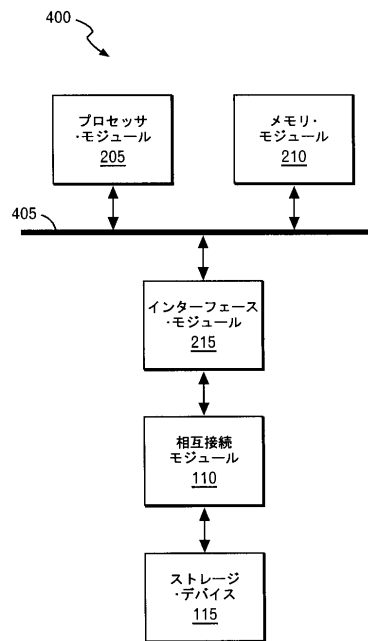
【図 2】



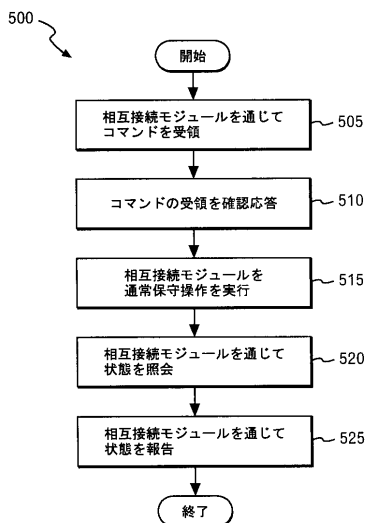
【図 3】



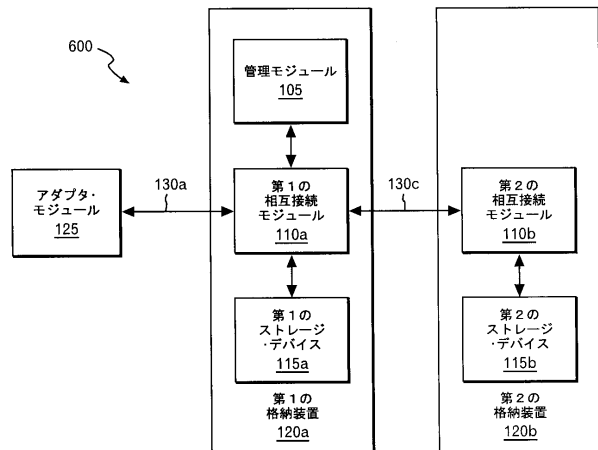
【図 4】



【図 5】



【図 6】



フロントページの続き

- (72)発明者 マシュー・デーヴィッド・ボムホッフ
アメリカ合衆国 85745 アリゾナ州、ツーソン、ノース デザート パーク プレース 1
37
- (72)発明者 ブライアン・ジェームズ・キャグノ
アメリカ合衆国 85715 アリゾナ州、ツーソン、イースト ウインドフラワー ドライブ
8861
- (72)発明者 グレグ・スティーブン・ルーカス
アメリカ合衆国 85748 アリゾナ州、ツーソン、ノース サークル ディー ウエイ 67
5
- (72)発明者 ケニー・ニアン＝ガン・クィウ
アメリカ合衆国 85712 アリゾナ州、ツーソン、ノース ウィルモト ロード 1901、
アパートメント 1001

審査官 菅原 浩二

- (56)参考文献 特開2001-022526 (JP, A)
特開2005-157636 (JP, A)

- (58)調査した分野(Int.Cl., DB名)
G06F 3/06