

(19) 日本国特許庁 (JP)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2012-504808

(P2012-504808A)

(43) 公表日 平成24年2月23日 (2012. 2. 23)

(51) Int.Cl.		F I			テーマコード (参考)
G06F 3/06	(2006.01)	G06F 3/06	304R		5B034
G06F 11/18	(2006.01)	G06F 11/18	310G		5B065

審査請求 未請求 予備審査請求 未請求 (全 18 頁)

(21) 出願番号 特願2011-529553 (P2011-529553) (86) (22) 出願日 平成21年10月1日 (2009. 10. 1) (85) 翻訳文提出日 平成23年3月16日 (2011. 3. 16) (86) 国際出願番号 PCT/EP2009/062757 (87) 国際公開番号 W02010/037814 (87) 国際公開日 平成22年4月8日 (2010. 4. 8) (31) 優先権主張番号 08165694.4 (32) 優先日 平成20年10月2日 (2008. 10. 2) (33) 優先権主張国 欧州特許庁 (EP)	(71) 出願人 390009531 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 INTERNATIONAL BUSINESS MACHINES CORPORATION 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 (74) 代理人 100108501 弁理士 上野 剛史 (74) 代理人 100112690 弁理士 太佐 種一 (74) 代理人 100091568 弁理士 市位 嘉宏
---	---

最終頁に続く

(54) 【発明の名称】 1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための方法、装置、およびプログラム (交

(57) 【要約】

【課題】 1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための方法、装置、およびソフトウェアを提供する。

【解決手段】 ストレージ・システムは、所与のユニットが交換されたことに応答して、自動的にテスト・モードになり、テストが失敗した場合、ストレージ・システムは自動的にサービス・モードにフェイルバックする。

【選択図】 図5

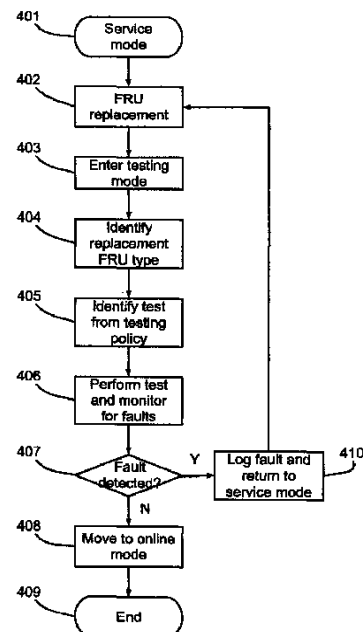


Figure 4

【特許請求の範囲】**【請求項 1】**

1 つまたは複数の交換ユニットのテストを管理するために 1 つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための方法であって、前記方法が、

a) 交換のためにストレージ・システム内の交換可能ユニットを識別し、前記ストレージ・システムをサービス・モードに切り替えるステップと、

b) 前記交換可能ユニットと交換ユニットとの交換に応答して、前記ストレージ・システムを自動的にテスト・モードに切り替えるステップと、

c) 前記交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて前記交換ユニットをテストするステップと、

d) 前記選択されたエレメントの前記テストが成功した場合に、前記ストレージ・システムをオンライン・モードに戻すステップと、

を含む方法。

【請求項 2】

前記ストレージ・システムからの 1 つまたは複数の追加のユニットが、前記テストのために前記交換ユニットと組み合わせて使用するために選択される、請求項 1 記載の方法。

【請求項 3】

前記追加のユニットが、前記交換ユニットの全動作容量をテストするために適切な入出力 (I/O) 経路を提供するために前記交換ユニットと組み合わせて使用するために選択される、請求項 2 記載の方法。

【請求項 4】

いかなる障害も前記テスト時に検出されない場合、前記ストレージ・システムが自動的に前記オンライン・モードに戻され、所定のレベルの障害が前記テスト中に検出された場合、前記ストレージ・システムが自動的に前記サービス・モードに戻される、請求項 1 ないし 3 のいずれかに記載の方法。

【請求項 5】

前記テスト・ポリシーが、前記交換ユニットにテスト・データを排他的に送信することを含む、請求項 1 ないし 4 のいずれかに記載の方法。

【請求項 6】

前記テスト・ポリシーが、前記交換ユニットにテスト・データを送信することを含み、前記テスト・データのレベルが、前記交換ユニットに所定のレベルの入出力を提供するために前記交換ユニットへのユーザ・データのレベルを補完するように構成される、請求項 1 ないし 4 のいずれかに記載の方法。

【請求項 7】

前記テストが 1 つまたは複数のデータ・ストレージ・ユニットにデータを書き込むことを含むときに、前記データが前記ストレージ・ユニットの予約領域に書き込まれる、請求項 1 ないし 6 のいずれかに記載の方法。

【請求項 8】

1 つまたは複数の交換ユニットのテストを管理するために 1 つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための装置であって、前記装置が、

交換のためにストレージ・システム内の交換可能ユニットを識別し、前記ストレージ・システムをサービス・モードに切り替え、

前記交換可能ユニットと交換ユニットとの交換に応答して、前記ストレージ・システムを自動的にテスト・モードに切り替え、

前記交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて前記交換ユニットをテストし、

前記選択されたエレメントの前記テストが成功した場合に、前記ストレージ・システムをオンライン・モードに戻す

10

20

30

40

50

ように動作可能である装置。

【請求項 9】

前記ストレージ・システムからの 1 つまたは複数の追加のユニットが、前記テストのために前記交換ユニットと組み合わせて使用するために選択される、請求項 8 記載の装置。

【請求項 10】

前記追加のユニットが、前記交換ユニットの全動作容量をテストするために適切な入出力 (I/O) 経路を提供するために前記交換ユニットと組み合わせて使用するために選択される、請求項 9 記載の装置。

【請求項 11】

いかなる障害も前記テスト時に検出されない場合、前記ストレージ・システムが自動的に前記オンライン・モードに戻され、所定のレベルの障害が前記テスト中に検出された場合、前記ストレージ・システムが自動的に前記サービス・モードに戻される、請求項 8 ないし 10 のいずれかに記載の装置。

【請求項 12】

前記テスト・ポリシーが、前記交換ユニットにテスト・データを排他的に送信することを含む、請求項 8 ないし 11 のいずれかに記載の装置。

【請求項 13】

前記テスト・ポリシーが、前記交換ユニットにテスト・データを送信することを含み、前記テスト・データのレベルが、前記交換ユニットに所定のレベルの入出力を提供するために前記交換ユニットへのユーザ・データのレベルを補完するように構成される、請求項 8 ないし 11 のいずれかに記載の装置。

【請求項 14】

前記テストが 1 つまたは複数のデータ・ストレージ・ユニットにデータを書き込むことを含むときに、前記データが前記ストレージ・ユニットの予約領域に書き込まれる、請求項 8 ないし 13 のいずれかに記載の装置。

【請求項 15】

コンピュータ・システムにロードされ、そこで実行されたときに、請求項 1 ないし 7 のいずれかに記載の方法のすべてのステップを前記コンピュータ・システムに実行させるコンピュータ・プログラム・コードを含むコンピュータ・プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、1 つまたは複数の交換ユニット (replacement unit) のテストを管理するために 1 つまたは複数の交換可能ユニット (replaceable unit) を含むコンピュータ化ストレージ・システムで使用するための方法、装置、またはソフトウェアに関する。

【背景技術】

【0002】

コンピュータ・システムは一般に、相互接続された 1 組の別個のコンポーネントから構築される。ユニットは、障害のある状態になると、正しく機能するユニットと交換する必要がある。高可用性のストレージ・システムなどの何らかのコンピュータ・システムは、重要なユニットが二重になっている組み込み冗長性を備えている。重要なユニットは、ディスク・アレイ、インターフェース・カード、および接続ケーブルなどのユニットにすることができ、そのいずれも障害のある状態になり、交換を必要とする可能性がある。このようなシステムでは、障害のあるユニットを交換している間、重複ユニットによって提供される冗長性を使用して、連続サービスを提供することができる。たとえば、接続ケーブルを交換する必要があることを示唆する障害またはエラーが検出された場合、交換が実行されている間、ユーザ入出力 (I/O) は重複ユニットに方向転換される。交換ユニットは、所定の位置に置かれると、ユーザ入出力による使用のために使用可能になる。しかし、障害またはエラーが誤診された場合、交換ユニットが障害のある状態である場合、または 2 つ以上のユニットが障害の一因となっている場合、障害またはエラーは再発する可能

10

20

30

40

50

性がある。加えて、検出された障害またはエラーが交換されたユニットによって矯正されない場合、ユーザ入出力が混乱する可能性がある。

【発明の概要】

【発明が解決しようとする課題】

【0003】

したがって、本発明は、第1の態様では、1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための方法を提供し、前記方法は、a) 交換のためにストレージ・システム内の交換可能ユニットを識別し、前記ストレージ・システムをサービス・モードに切り替えるステップと、b) 前記交換可能ユニットと交換ユニットとの交換に応答して、前記ストレージ・システムを自動的にテスト・モードに切り替えるステップと、c) 前記交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて前記交換ユニットをテストするステップと、d) 前記選択されたエレメントの前記テストが成功した場合に、前記ストレージ・システムをオンライン・モードに戻すステップとを含む。

10

【課題を解決するための手段】

【0004】

好ましくは、前記ストレージ・システムからの1つまたは複数の追加のユニットは、前記テストのために前記交換ユニットと組み合わせて使用するために選択される。好ましくは、前記追加のユニットは、前記交換ユニットの全動作容量 (full operational capacity) をテストするために適切な入出力 (I/O) 経路を提供するために前記交換ユニットと組み合わせて使用するために選択される。

20

【0005】

好ましくは、前記追加のユニットは、前記テスト時に使用するために自動的に識別され選択される。

【0006】

好ましくは、いかなる障害も前記テスト時に検出されない場合、前記ストレージ・システムは自動的に前記オンライン・モードに戻される。好ましくは、所定のレベルの障害が前記テスト中に検出された場合、前記ストレージ・システムは自動的に前記サービス・モードに戻される。

30

【0007】

好ましくは、前記テスト・ポリシーは、前記交換ユニットにテスト・データを排他的に送信することを含む。好ましくは、前記テスト・ポリシーは、前記交換ユニットにテスト・データを送信することを含み、前記テスト・データのレベルは、前記交換ユニットに所定のレベルの入出力を提供するために前記交換ユニットへのユーザ・データのレベルを補完するように構成される。

【0008】

好ましくは、前記テストが1つまたは複数のデータ・ストレージ・ユニットにデータを書き込むことを含むときに、前記データは前記ストレージ・ユニットの予約領域に書き込まれる。好ましくは、前記ストレージ・システムは、冗長性を提供するために複数の重複ストレージ・システムを含む。好ましくは、所与のストレージ・システム内の前記交換ユニットがテスト・モードであるときに、すべてのユーザ入出力は前記重複ストレージ・システムに方向転換される。好ましくは、前記ストレージ・システムは、1つまたは複数のファイバ・チャネル・アービトラテッド・ループ (FCAL: Fibre Channel Arbitrated Loop) を含む。

40

【0009】

第2の態様では、1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための装置が提供され、前記装置は、交換のためにストレージ・システム内の交換可能ユニットを識別し、前記ストレージ・システムをサービス・モードに切り替え、前記交換可能ユニットと交換ユニットとの交換に応答して、前記ストレージ・システムを自動的にテスト・モ

50

ードに切り替え、前記交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて前記交換ユニットをテストし、前記選択されたエレメントの前記テストが成功した場合に、前記ストレージ・システムをオンライン・モードに戻すように動作可能である。

【0010】

好ましくは、前記ストレージ・システムからの1つまたは複数の追加のユニットは、前記テストのために前記交換ユニットと組み合わせて使用するために選択される。好ましくは、前記追加のユニットは、前記交換ユニットの全動作容量をテストするために適切な入出力(I/O)経路を提供するために前記交換ユニットと組み合わせて使用するために選択される。好ましくは、前記追加のユニットは、前記テスト時に使用するために自動的に識別され選択される。

10

【0011】

好ましくは、いかなる障害も前記テスト時に検出されない場合、前記ストレージ・システムは自動的に前記オンライン・モードに戻される。好ましくは、所定のレベルの障害が前記テスト中に検出された場合、前記ストレージ・システムは自動的に前記サービス・モードに戻される。

【0012】

好ましくは、前記テスト・ポリシーは、前記交換ユニットにテスト・データを排他的に送信することを含む。好ましくは、前記テスト・ポリシーは、前記交換ユニットにテスト・データを送信することを含み、前記テスト・データのレベルは、前記交換ユニットに所定のレベルの入出力を提供するために前記交換ユニットへのユーザ・データのレベルを補完するように構成される。好ましくは、前記テストが1つまたは複数のデータ・ストレージ・ユニットにデータを書き込むことを含むときに、前記データは前記ストレージ・ユニットの予約領域に書き込まれる。好ましくは、前記ストレージ・システムは、冗長性を提供するために複数の重複ストレージ・システムを含む。好ましくは、所与のストレージ・システム内の前記交換ユニットがテスト・モードであるときに、すべてのユーザ入出力は前記重複ストレージ・システムに方向転換される。好ましくは、前記ストレージ・システムは、1つまたは複数のファイバ・チャネル・アービトラレーテッド・ループ(FCAL)を含む。

20

【0013】

第3の態様では、1つまたは複数のプログラマブル・デバイスからなるグループが第1の態様の方法を実行できるようにするかまたは第2の態様の装置を提供できるようにするように構成された1つまたは複数のプログラムからなるグループが提供される。

30

【0014】

第4の態様では、コンピュータ・システムにロードされ、そこで実行されたときに、第1の態様による方法のすべてのステップを前記コンピュータ・システムに実行させるコンピュータ・プログラム・コードを含むコンピュータ・プログラムが提供される。

【0015】

このコンピュータ・プログラムは、前記プログラムがコンピュータ上で実行されたときに第1の態様の方法を実行するためのソフトウェア・コード部分を含み、デジタル・コンピュータの内部メモリにロード可能なコンピュータ・プログラム(computer program product)の形で提供することができる。

40

【0016】

本発明の一実施形態は、1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための方法を提供し、この方法は、

- a) 交換のためにストレージ・システム内の交換可能ユニットを識別し、ストレージ・システムをサービス・モードに切り替えるステップと、
- b) 交換可能ユニットと交換ユニットとの交換にตอบสนองして、ストレージ・システムを自動的にテスト・モードに切り替えるステップと、
- c) 交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて交換ユニッ

50

トをテストするステップと、

d) 選択されたエレメントのテストが成功した場合に、ストレージ・システムをオンライン・モードに戻すステップと、
を含む。

【0017】

ストレージ・システムからの1つまたは複数の追加のユニットは、テストのために交換ユニットと組み合わせて使用するために選択することができる。追加のユニットは、交換ユニットの全動作容量をテストするために適切な入出力(I/O)経路を提供するために交換ユニットと組み合わせて使用するために選択することができる。追加のユニットは、テスト時に使用するために自動的に識別し選択することができる。いかなる障害もテスト時に検出されない場合、ストレージ・システムは自動的にオンライン・モードに戻すことができる。所定のレベルの障害がテスト中に検出された場合、ストレージ・システムは自動的にサービス・モードに戻すことができる。

【0018】

テスト・ポリシーは、交換ユニットにテスト・データを排他的に送信することを含むことができる。テスト・ポリシーは、交換ユニットにテスト・データを送信することを含むことができ、テスト・データのレベルは、交換ユニットに所定のレベルの入出力を提供するために交換ユニットへのユーザ・データのレベルを補完するように構成される。テストが1つまたは複数のデータ・ストレージ・ユニットにデータを書き込むことを含むときに、データはストレージ・ユニットの予約領域に書き込むことができる。

【0019】

ストレージ・システムは、冗長性を提供するために複数の重複ストレージ・システムを含むことができる。所与のストレージ・システム内の交換ユニットがテスト・モードであるときに、すべてのユーザ入出力は重複ストレージ・システムに方向転換することができる。ストレージ・システムは、1つまたは複数のファイバ・チャネル・アービトレーテッド・ループ(FCAL)を含むことができる。

【0020】

他の実施形態は、1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための装置を提供し、この装置は、
交換のためにストレージ・システム内の交換可能ユニットを識別し、ストレージ・システムをサービス・モードに切り替え、
交換可能ユニットと交換ユニットとの交換に応答して、ストレージ・システムを自動的にテスト・モードに切り替え、
交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて交換ユニットをテストし、
選択されたエレメントのテストが成功した場合に、ストレージ・システムをオンライン・モードに戻す
ように動作可能である。

【0021】

他の一実施形態は、1つまたは複数のプログラマブル・デバイスからなるグループが、1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための方法を実行できるようにするように構成された1つまたは複数のプログラムからなるグループを提供し、この方法は、

a) 交換のためにストレージ・システム内の交換可能ユニットを識別し、ストレージ・システムをサービス・モードに切り替えるステップと、
b) 交換可能ユニットと交換ユニットとの交換に応答して、ストレージ・システムを自動的にテスト・モードに切り替えるステップと、
c) 交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて交換ユニッ

トをテストするステップと、

d) 選択されたエレメントのテストが成功した場合に、ストレージ・システムをオンライン・モードに戻すステップと、
を含む。

【0022】

他の実施形態は、1つまたは複数のプログラマブル・デバイスからなるグループが、1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための装置を提供できるようにするように構成された1つまたは複数のプログラムからなるグループを提供し、この装置は、

10

交換のためにストレージ・システム内の交換可能ユニットを識別し、ストレージ・システムをサービス・モードに切り替え、

交換可能ユニットと交換ユニットとの交換に応答して、ストレージ・システムを自動的にテスト・モードに切り替え、

交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて交換ユニットをテストし、

選択されたエレメントのテストが成功した場合に、ストレージ・システムをオンライン・モードに戻す

ように動作可能である。

【0023】

20

他の一実施形態は、プログラムがコンピュータ上で実行されたときに、1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための方法を実行するためのソフトウェア・コード部分を含み、デジタル・コンピュータの内部メモリにロード可能なコンピュータ・プログラムを提供し、この方法は、

a) 交換のためにストレージ・システム内の交換可能ユニットを識別し、ストレージ・システムをサービス・モードに切り替えるステップと、

b) 交換可能ユニットと交換ユニットとの交換に応答して、ストレージ・システムを自動的にテスト・モードに切り替えるステップと、

c) 交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて交換ユニットをテストするステップと、

30

d) 選択されたエレメントのテストが成功した場合に、ストレージ・システムをオンライン・モードに戻すステップと、

を含む。

【0024】

次に、添付図面に関連して、例としてのみ、本発明の諸実施形態について説明する。

【図面の簡単な説明】

【0025】

【図1】コンピュータ化ストレージ・システムを含むコンピュータ・ネットワークの概略図である。

40

【図2】図1のコンピュータ化ストレージ・システムをより詳細に示す概略図である。

【図3】図2のコンピュータ化ストレージ・システム内のユニットの接続を示す表である。

【図4】図2のコンピュータ化ストレージ・システム用のテスト・ポリシーを示す表である。

【図5】図2のコンピュータ化ストレージ・システムによって実行される処理を示す流れ図である。

【図6】図2のコンピュータ化ストレージ・システムによって実行される他の処理を示す流れ図である。

【発明を実施するための形態】

50

【 0 0 2 6 】

図 1 に関連して説明すると、コンピュータ・システムは、ストレージ・サーバ・アプリケーション・プログラム 1 0 3 がロードされ、ストレージ・システム 1 0 4 に接続された第 1 のコンピュータ 1 0 2 を含む。第 1 のコンピュータ 1 0 2 は、ローカル・エリア・ネットワーク (L A N) 1 0 6 の形のネットワークを介して 2 つの第 2 のコンピュータ 1 0 5 に接続されている。第 2 のコンピュータ 1 0 5 にはそれぞれ、ストレージ・サーバ 1 0 3 を介してストレージ・システム 1 0 4 上にデータを保管し、そのデータにアクセスするように構成されたデータ処理アプリケーション・プログラム 1 0 7 がロードされている。換言すれば、データ処理アプリケーション・プログラム 1 0 7 は、ストレージ・サーバ 1 0 3 およびストレージ・システム 1 0 4 へのユーザ入出力 (I / O) を提供する。

10

【 0 0 2 7 】

図 2 に関連して説明すると、ストレージ・サーバ・アプリケーション 1 0 3 は、より詳細に後述する通り、ユーザ・インターフェースを提供し、ユーザ入出力を処理するように構成されたストレージ・コントローラ 2 0 1 と、ストレージ・システム 1 0 4 の設置、サービス、およびテストを協力して管理するように構成された重複ストレージ・アダプタ 2 0 2、2 0 3 とを含む。ストレージ・システム 1 0 4 は、それぞれが 4 つのディスク 2 0 7 と 2 つのインターフェース・カード 2 0 8、2 0 9 とを含む、3 つのディスク・アレイ 2 0 4、2 0 5、2 0 6 を含む。それぞれのインターフェース・カード 2 0 8、2 0 9 は、アダプタ 2 0 2、2 0 3 に接続するための接続ポートを提供する。それぞれのアダプタ 2 0 2、2 0 3 をそれぞれのディスク・アレイ 2 0 4、2 0 5、2 0 6 の両方のインターフェース・カード 2 0 8、2 0 9 に直列に接続するために、1 組のリンク 2 1 0、2 1 1 が設けられている。図 2 には、リンク 2 1 0、2 1 1 など、何らかの重複項目の出現ごとのラベルは明瞭にするために省略されている。複数の重複アダプタ 2 0 2、2 0 3、リンク 2 1 0、2 1 1、ディスク 2 0 7、およびインターフェース・カード 2 0 8、2 0 9 の設置により、ストレージ・システム 1 0 4 に複数の冗長性が提供されている。この実施形態では、各ユニット 2 0 7、2 0 8、2 0 9、2 1 0、2 1 1 は、事実上、2 つの重複ストレージ・システムを提供することにより、冗長性を提供するために少なくとも 1 回重複されている。1 つの重複ストレージ・システムがオフラインである場合、その重複ストレージ・システムは必要なサービスを提供するように構成されているので、ストレージ・コントローラ 2 0 1 は、重複ストレージ・システムを使用して高可用性を提供するように構成されている。

20

30

【 0 0 2 8 】

この実施形態では、複数のリンク 2 1 0、2 1 1、ディスク 2 0 7、およびインターフェース・カード 2 0 8、2 0 9 のそれぞれは、一般にフィールド交換可能ユニット (F R U : Field Replaceable Unit) と呼ばれ、個別に交換可能な別個のユニットとして設けられている。ストレージ・コントローラ 2 0 1 は、リンク 2 1 0、2 1 1、ディスク 2 0 7、およびインターフェース・カード 2 0 8、2 0 9 内の所定の範囲の障害およびエラーを検出するための機能性を含み、検出されると、エンジニアに警告する。障害のある F R U を交換するために、関連重複ストレージ・システムはオンライン・モードからサービス・モードになり、可用性は冗長重複ストレージ・システムによって維持される。この実施形態では、所与の重複ストレージ・システムがサービス・モードであるときに、すべてのユーザ入出力は冗長重複ストレージ・システムに方向転換される。オンラインまたはサービス・モード切り替えは、両方が同時にサービス・モードにならず、その結果、可用性が維持されることを保証するために、それぞれのストレージ・アダプタ 2 0 2、2 0 3 の間で協力して実行される。

40

【 0 0 2 9 】

この実施形態では、ストレージ・アダプタ 2 0 2、2 0 3 は、サービス・モードとオンライン・モードとの中間モードであるテスト・モードの形の追加のモードによって構成されている。エンジニアが F R U を交換し、重複ストレージ・システムがオンライン・モードに戻れることを示すと、関連ストレージ・アダプタ 2 0 2、2 0 3 は、重複ストレージ

50

・システムを自動的にテスト・モードにするように構成されている。テスト・モードでは、1組の1つまたは複数の所定のテストが交換F R U上で実行され、それが正しく機能していることと、検出された障害が交換によって修正されたことを保証する。この実施形態では、テストは、交換されたF R Uのタイプに依存して選択される。この実施形態では、テスト・モードでは、所与の重複ストレージ・システム上でユーザ入出力が許可される。しかし、テスト・モード切り替えは、両方が同時にテスト・モードにならないことを保証するために、それぞれのストレージ・アダプタ202、203の間で協力して実行される。これは、両方のシステムが同時にサービス・モードにフェイルバックする(fail back)可能性を回避するように構成されている。

【0030】

10

図3に関連して説明すると、テストは、ストレージ・システム104内の各タイプのF R Uごとの適切なテストと、適用可能であれば、そのテストに関する持続時間を識別する、テスト・ポリシー301に定義されている。たとえば、交換されたF R Uがリンクである場合、交換リンクは10分間、ストレージ・テストが行われる。F R Uがディスクである場合、完了するかまたは打ち切られるまでの不確定時間の間、実行される、ディスク検証テストが行われる。所与の重複ストレージ・システムがテスト・モードである間、そのシステムは新しい障害または障害の再発についてモニターされる。任意の障害が検出された場合、重複ストレージ・システムは自動的にサービス・モードに戻される。しかし、テストが完了し、いかなる障害も検出されない場合、重複ストレージ・システムは自動的にオンライン・モードに戻される。

20

【0031】

上記の通り、図3に関連して説明すると、交換リンクについてストレージ・テストが実行される。ストレージ・テストは、リンクを介して適切なレベルのテスト入出力を供給し、テスト期間の間、そのパフォーマンスをモニターすることにより、交換されたリンクに関する最大データ転送速度をテストするように構成されている。しかし、リンクがディスクなど、そのすぐ下流のエレメントより大きい容量を有する場合、ストレージ・テストを実行するための適切な帯域幅を提供するために、いくつかの下流のエレメントを選択する必要がある可能性がある。たとえば、交換されたリンクが2ギガビット/秒(Gb/s)という最大データ転送速度を有し、それぞれが50メガバイト/秒(MB/s)という最大データ転送速度を備えたディスクの上流に接続される場合、2Gb/sのテスト入出力の形のテスト・データを吸収するために、4つのディスクを使用することが必要になるであろう。

30

【0032】

図4に関連して説明すると、それぞれのストレージ・アダプタ202、203は、ストレージ・システム104内の各F R Uに関する固有のID(F R U ID)とともに所与のF R Uのタイプならびにすぐ下流のユニットのF R U IDを記録する、F R U接続表の形の表302を維持するように構成されている。ストレージ・アダプタ202、203は、交換ユニットに関する適切なレベルのテスト入出力の提供を可能にするためにテスト入出力に適切なターゲットを提供する交換ユニットから下流のユニットを識別するためにF R U接続表302を使用するように構成されている。

40

【0033】

次に、それぞれの重複ストレージ・システム上で障害を検出した結果としてサービス・モードになったことに応答してストレージ・アダプタ202、203によって実行される処理について、図5の流れ図に関連してさらに説明する。ステップ401では重複ストレージ・システムがサービス・モードになったことに応答して処理が開始され、処理はステップ402に移行して、障害のあるF R Uの交換を待つ。インストール・エンジニアによって示された通り、障害のあるF R Uが交換されると、処理は重複ストレージ・システムが自動的にテスト・モードに入るステップ403に移行し、さらに処理はステップ404に移行する。ステップ404では交換F R Uのタイプが識別され、処理はステップ405に移行する。ステップ405では、識別されたF R Uタイプを使用して、テスト・ポリシ

50

ー 3 0 1 から適切なテストが識別され、処理はステップ 4 0 6 に移行する。ステップ 4 0 6 では識別されたテストが実行され、処理はステップ 4 0 7 に移行する。ステップ 4 0 7 では障害モニターが実行され、テストの持続時間の間、いかなる障害も検出されない場合、処理はステップ 4 0 8 に移行する。ステップ 4 0 8 では重複ストレージ・システムが自動的にオンライン・モードになり、処理はステップ 4 0 9 で終了する。ステップ 4 0 7 で所定の 1 組の 1 つまたは複数の障害が検出された場合、処理はステップ 4 1 0 に移行する。ステップ 4 1 0 ではそれぞれの障害がログに記録され、重複ストレージ・システムは自動的にサービス・モードに戻される。次に処理はステップ 4 0 2 に移行して追加のサービス・アクションを待ち、上述のように続行する。

【 0 0 3 4 】

次に、ストレス・テストを開始したことに応答してストレージ・アダプタ 2 0 2、2 0 3 によって実行される処理について、図 6 の流れ図に関連してさらに説明する。ステップ 5 0 1 では図 5 の流れ図のステップ 4 0 5 でストレス・テストを識別したことに応答して処理が開始され、処理はステップ 5 0 2 に移行する。ステップ 5 0 2 では交換 F R U に関する最大動作データ転送速度が F R U 接続表 3 0 2 から識別され、処理はステップ 5 0 3 に移行する。ステップ 5 0 3 では、交換 F R U の最大動作データ転送速度で入出力を処理または吸収できるテスト・グループを形成するために、もう一度 F R U 接続表 3 0 2 を使用して、重複ストレージ・システムからの 1 組の接続ユニットが選択される。次に、処理は、ユーザ入出力の現行レベルが決定されるステップ 5 0 4 に移行する。上記の通り、この実施形態では、ユーザ入出力はテスト・モードの重複ストレージ・システム上で許可される。ユーザ入出力のレベルが決定されると、処理はステップ 5 0 5 に移行する。ステップ 5 0 5 では、必要な量の付加テスト入出力 (top-up test I/O)、すなわち、交換 F R U に関する最大動作データ転送速度に測定されたユーザ入出力を付加するために必要なテスト入出力が計算され、処理はステップ 5 0 6 に移行する。ステップ 5 0 6 では、付加テスト入出力が生成され、ストレス・テストに関する定義済み期間の間、ステップ 5 0 3 で選択されたテスト・グループのユニットに経路指定される。定義済み期間の間、テスト入出力が提供されると、処理はステップ 5 0 7 に移行して終了する。

【 0 0 3 5 】

この実施形態では、付加テスト入出力は、その目的のために予約された関連ディスク・ドライブの領域に送信される。換言すれば、任意のユーザ入出力がテスト入出力で上書きされるかまたは破損するのを回避するために、ユーザ入出力は特にテスト入出力ターゲット・ディスク・ドライブの予約領域から除外される。

【 0 0 3 6 】

重複ストレージ・システムがテスト・モードになると、ユーザ入出力の少なくとも一部は、冗長重複ストレージ・システムの使用から調査中の重複ストレージ・システムにフェイルバックするように構成される。ループまたは脱落フレーム (dropped frame) の予期せぬ認識など、任意のエラーが識別された場合、この結果として、テスト・モードの即時失敗およびサービス・モードへのフェイルバックが発生し、それにより、冗長重複ストレージ・システムの下のユーザ入出力だけを再駆動する。この結果、冗長ストレージ・システムはデータ用の良好な経路を維持するので、カスタム・アクセスに対するリスクはまったく発生しない。テスト・モードが失敗した場合、交換 F R U を再調査することができ、追加の修理アクションを実行することができる。これにより、たとえば、修正中の問題が断続的であり、単一 F R U 障害として容易に識別できなかったときに、所与の修理の信頼度が改善される。

【 0 0 3 7 】

他の実施形態では、システム内にいかなる冗長性も提供されず、システムがサービス・モードである間、いかなるユーザ入出力も処理することができない。他の実施形態では、部分冗長性が提供され、すなわち、システム内のすべてのユニットではなく一部のユニットについて冗長性が提供される。他の実施形態では、それぞれの F R U 自体は、全体としてのシステムに関するグローバル・モードではなく、個々のサービス・モード、テスト・

10

20

30

40

50

モード、およびオンライン・モードを有するように構成される。他の一実施形態では、テスト・モード中に所与のシステム上でいかなるユーザ入出力も使用されず、ストレス・テストが実行される場合、テスト・データは、交換 F R U に関する最大動作データ転送速度を提供するように構成される。

【 0 0 3 8 】

当業者には理解されるように、交換ユニットは、既存のユニットまたは障害のあるユニットのための直接交換品ではない可能性があるが、たとえば、障害を修正するためまたは追加の機能性を提供するために、アップグレードされたユニットまたは新たにインストールされシステムに追加されるユニットにすることができる。さらに、所与のサービス・アクション中に 2 つ以上の F R U を交換することもでき、その結果、所与のテストで 2 つ以上の F R U をテストすることもできる。当業者には理解されるように、テスト入出力のフォーマットは、特に、テスト中の運搬機器のタイプに関して構成することができる。たとえば、ファイバ・チャネル・アービトラレーテッド・ループ (F C A L) システムは、低周波数で値が変化するビット・ストリームを含む入出力のパターンによって最も良くストレスを加えることができる。

【 0 0 3 9 】

当業者であれば、本発明の一部または全部を実施する装置が、本発明の一実施形態の一部または全部を提供するように構成されたソフトウェアを有する汎用デバイスにすることができることを理解されるであろう。このデバイスは単一デバイスまたは複数デバイスのグループにすることができ、ソフトウェアは単一プログラムまたは複数プログラムのセットにすることができるであろう。さらに、本発明を実装するために使用されるソフトウェアの一部または全部は、そのソフトウェアを 1 つまたは複数のデバイス上にロードできるように、任意の適切な伝送手段または記憶手段を介して通信することができる。

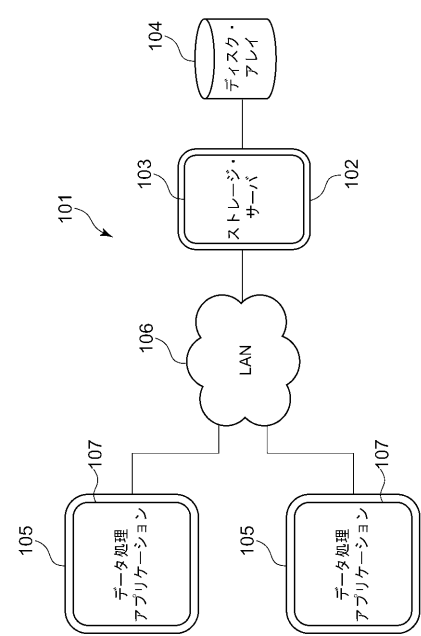
【 0 0 4 0 】

本発明の諸実施形態の説明によって本発明を例示し、諸実施形態についてかなり詳細に説明してきたが、追加の利点および変更例は当業者にとって容易に明らかになるであろう。

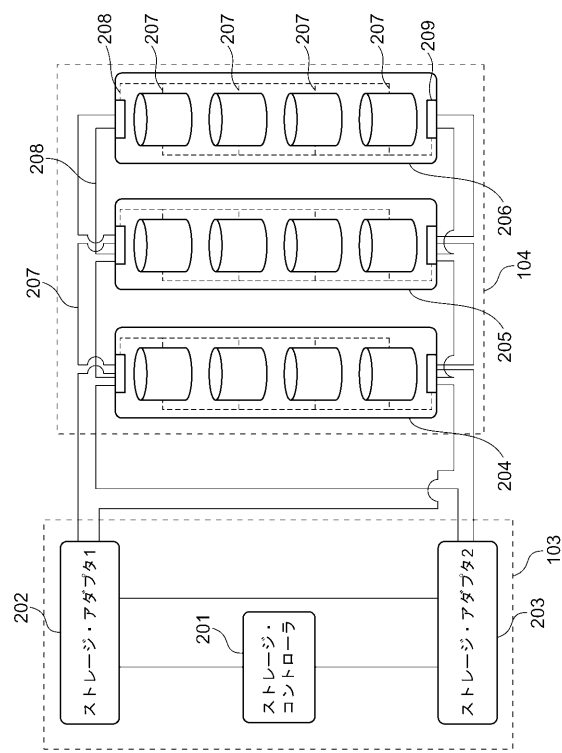
10

20

【 図 1 】



【 図 2 】



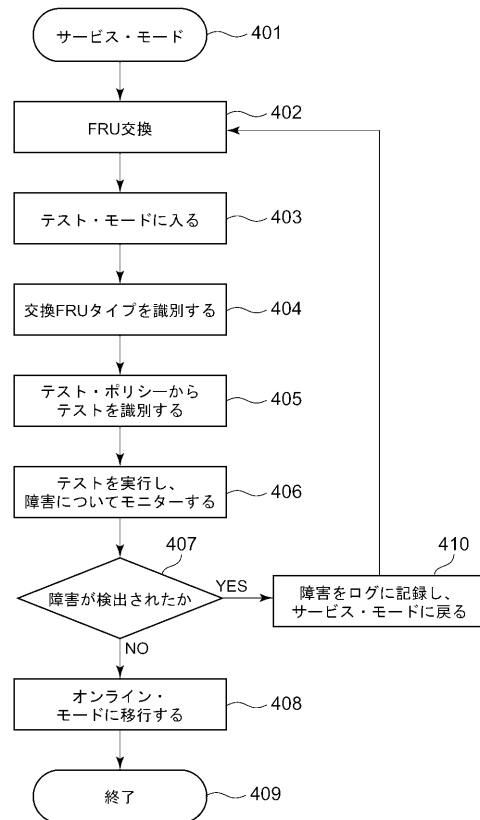
【 図 3 】

テスト・ポリシー		
FRUタイプ	テスト	テスト期間
リンク	ストレス・テスト	10分間
インターフェース・カード	ストレス・テスト	20分間
ディスク	ディスク検証	該当なし

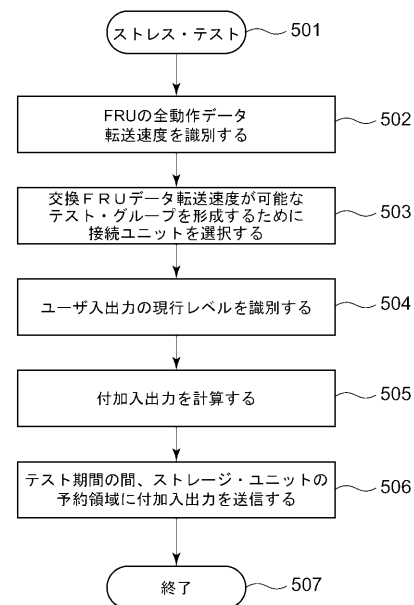
【 図 4 】

FRU接続表			
FRU ID	タイプ	データ転送速度	下流のFRU
FRU ID1	リンク	2Gb/s	FRU ID2
FRU ID2	インターフェース・カード	2Gb/s	FRU ID3
FRU ID3	ディスク	50MB/s	FRU ID4
FRU ID4	リンク	2Gb/s	FRU ID5
FRU ID5	インターフェース・カード	2Gb/s	FRU ID6
FRU ID6	ディスク	50MB/s	FRU ID7
FRU ID7	リンク	2Gb/s	FRU ID8
FRU ID8	インターフェース・カード	2Gb/s	FRU ID9
FRU ID9	ディスク	50MB/s	該当なし

【図 5】



【図 6】



【手続補正書】

【提出日】平成23年4月28日(2011.4.28)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

1 つまたは複数の交換ユニットのテストを管理するために 1 つまたは複数の交換可能ユニットを含むストレージ・システムについて、コンピュータに各ステップを実行させる方法であって、前記方法が、

a) 交換のためにストレージ・システム内の交換可能ユニットを識別し、前記ストレージ・システムをサービス・モードに切り替えるステップと、

b) 前記交換可能ユニットと交換ユニットとの交換にตอบสนองして、前記ストレージ・システムを自動的にテスト・モードに切り替えるステップと、

c) 前記交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて前記交換ユニットをテストするステップと、

d) 選択されたエレメントの前記テストが成功した場合に、前記ストレージ・システムをオンライン・モードに戻すステップと、

を含み、

前記テスト・ポリシーが、前記交換ユニットにテスト・データを送信することを含み、前記テスト・データのレベルが、前記交換ユニットに所定のレベルの入出力を提供するために前記交換ユニットへのユーザ・データのレベルを補完するように構成される、方法。

【請求項 2】

前記ストレージ・システムからの1つまたは複数の追加のユニットが、前記テストのために前記交換ユニットと組み合わせて使用するために選択される、請求項1記載の方法。

【請求項3】

前記追加のユニットが、前記交換ユニットの全動作容量をテストするために適切な入出力（I/O）経路を提供するために前記交換ユニットと組み合わせて使用するために選択される、請求項2記載の方法。

【請求項4】

いかなる障害も前記テストにおいて検出されない場合、前記ストレージ・システムが自動的に前記オンライン・モードに戻され、所定のレベルの障害が前記テスト中に検出された場合、前記ストレージ・システムが自動的に前記サービス・モードに戻される、請求項1ないし3のいずれかに記載の方法。

【請求項5】

前記テスト・ポリシーが、前記交換ユニットにテスト・データを排他的に送信することを含む、請求項1ないし4のいずれかに記載の方法。

【請求項6】

前記テストが1つまたは複数のデータ・ストレージ・ユニットにデータを書き込むことを含むときに、前記データが前記ストレージ・ユニットの予約領域に書き込まれる、請求項1ないし5のいずれかに記載の方法。

【請求項7】

1つまたは複数の交換ユニットのテストを管理するために1つまたは複数の交換可能ユニットを含むストレージ・システムについて、コンピュータが動作を実行する装置であって、前記装置が、

交換のためにストレージ・システム内の交換可能ユニットを識別し、前記ストレージ・システムをサービス・モードに切り替え、

前記交換可能ユニットと交換ユニットとの交換にตอบสนองして、前記ストレージ・システムを自動的にテスト・モードに切り替え、

前記交換ユニットの機能を検証するために、所定のテスト・ポリシーに応じて前記交換ユニットをテストし、

選択されたエレメントの前記テストが成功した場合に、前記ストレージ・システムをオンライン・モードに戻す

ように動作可能であり、

前記テスト・ポリシーが、前記交換ユニットにテスト・データを送信することを含み、前記テスト・データのレベルが、前記交換ユニットに所定のレベルの入出力を提供するために前記交換ユニットへのユーザ・データのレベルを補完するように構成される、装置。

【請求項8】

前記ストレージ・システムからの1つまたは複数の追加のユニットが、前記テストのために前記交換ユニットと組み合わせて使用するために選択される、請求項7記載の装置。

【請求項9】

前記追加のユニットが、前記交換ユニットの全動作容量をテストするために適切な入出力（I/O）経路を提供するために前記交換ユニットと組み合わせて使用するために選択される、請求項8記載の装置。

【請求項10】

いかなる障害も前記テストにおいて検出されない場合、前記ストレージ・システムが自動的に前記オンライン・モードに戻され、所定のレベルの障害が前記テスト中に検出された場合、前記ストレージ・システムが自動的に前記サービス・モードに戻される、請求項7ないし9のいずれかに記載の装置。

【請求項11】

前記テスト・ポリシーが、前記交換ユニットにテスト・データを排他的に送信することを含む、請求項7ないし10のいずれかに記載の装置。

【請求項12】

前記テストが 1 つまたは複数のデータ・ストレージ・ユニットにデータを書き込むことを含むときに、前記データが前記ストレージ・ユニットの予約領域に書き込まれる、請求項 7 ないし 1 1 のいずれかに記載の装置。

【請求項 1 3】

コンピュータ・システムにロードされ、そこで実行されたときに、請求項 1 ないし 6 のいずれかに記載の方法のすべてのステップを前記コンピュータ・システムに実行させるコンピュータ・プログラム・コードを含むコンピュータ・プログラム。

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2009/062757

A. CLASSIFICATION OF SUBJECT MATTER INV. G06F11/27		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EP0-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2004/158424 A1 (MANTEY PAUL JOHN [US] ET AL) 12 August 2004 (2004-08-12) page 1, paragraph 3 - paragraph 4 page 1, paragraph 18 page 3, paragraph 41 - page 4, paragraph 52 figures 1-3	1-15
X	US 5 293 556 A (HILL FLETCHER L [US] ET AL) 8 March 1994 (1994-03-08) the whole document	1-15
☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "A" document member of the same patent family		
Date of the actual completion of the international search 11 February 2010		Date of mailing of the international search report 22/02/2010
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040. Fax: (+31-70) 340-3016		Authorized officer Bauer, Regine

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2009/062757

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004158424 A1	12-08-2004	GB 2398390 A	18-08-2004
US 5293556 A	08-03-1994	NONE	

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LS,MW,MZ,NA,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IS,JP,KE,KG,KM,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LT,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PE,PG,PH,PL,PT,RO,RS,RU,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,ZA,ZM,ZW

(72)発明者 カシュマン、ポール、ニコラス

イギリス国2 ジェイ・エヌ エス・オー2 1 ハンプシャー州ウィンチェスター ハスレー・パーク エム・ピー1 4 8 I B Mユナイテッドキングダムリミテッド内

(72)発明者 セトル、ジョナサン、イアン

イギリス国2 ジェイ・エヌ エス・オー2 1 ハンプシャー州ウィンチェスター ハスレー・パーク エム・ピー1 4 8 I B Mユナイテッドキングダムリミテッド内

(72)発明者 マッカーシー、ティモシー、フィンバル

イギリス国2 ジェイ・エヌ エス・オー2 1 ハンプシャー州ウィンチェスター ハスレー・パーク エム・ピー1 4 8 I B Mユナイテッドキングダムリミテッド内

(72)発明者 ムーア、ロデリック、ガイ

イギリス国2 ジェイ・エヌ エス・オー2 1 ハンプシャー州ウィンチェスター ハスレー・パーク エム・ピー1 4 8 I B Mユナイテッドキングダムリミテッド内

(72)発明者 ショート、ジョナサン、ウィリアム、ルイス

イギリス国2 ジェイ・エヌ エス・オー2 1 ハンプシャー州ウィンチェスター ハスレー・パーク エム・ピー1 4 8 I B Mユナイテッドキングダムリミテッド内

F ターム(参考) 5B034 BB15 CC02 DD01 DD03 DD05

5B065 CA16 EC01

(54)【発明の名称】 1 つまたは複数の交換ユニットのテストを管理するために 1 つまたは複数の交換可能ユニットを含むコンピュータ化ストレージ・システムで使用するための方法、装置、およびプログラム (交換ユニットのテストを管理するために交換可能ユニットを含むコンピュータ化ストレージ・システム)