

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-133897

(P2004-133897A)

(43) 公開日 平成16年4月30日(2004.4.30)

(51) Int.Cl.⁷**G06F 3/06****G06F 12/00****G06F 13/00**

F 1

| | | |
|------|-------|------|
| G06F | 3/06 | 305A |
| G06F | 3/06 | 301Z |
| G06F | 12/00 | 514E |
| G06F | 12/00 | 545A |
| G06F | 13/00 | 301H |

テーマコード(参考)

5B065

5B082

5B083

審査請求 未請求 請求項の数 26 O L 外国語出願 (全 42 頁)

(21) 出願番号

特願2003-190750(P2003-190750)

(22) 出願日

平成15年7月3日(2003.7.3)

(31) 優先権主張番号

10/237402

(32) 優先日

平成14年9月6日(2002.9.6)

(33) 優先権主張国

米国(US)

(特許庁注:以下のものは登録商標)

イーサネット

(71) 出願人

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(74) 代理人

100075096

弁理士 作田 康夫

(72) 発明者

山本 政行

アメリカ合衆国カリフォルニア州サニーベ

イル イーストエルカミノリアル 965

(72) 発明者

田口 雄一

アメリカ合衆国カリフォルニア州サニーベ

イル イーストエルカミノリアル 965

F ターム(参考) 5B065 BA01 EA36
5B082 FA07 HA08
5B083 BB02 CD03 CE01 EE08

(54) 【発明の名称】ストレージ・ネットワークにおけるイベント通知

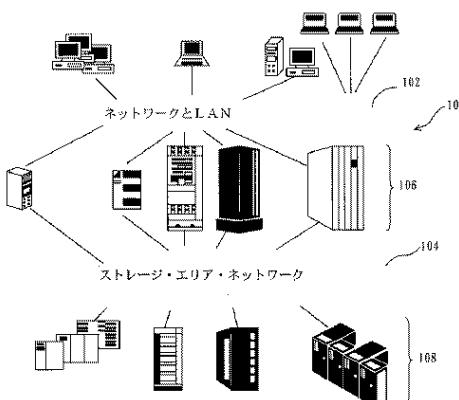
(57) 【要約】 (修正有)

【課題】高性能で、多目的の SAN を適切に利用するためには、イベント通知を効率良く管理する必要がある。

【解決手段】ヘテロジニアスなネットワークは、複数の納入業者からのハードウェアおよびソフトウェア製品に関わるネットワークを含んでいる。ネットワークは、データを記憶する構成のストレージ・システムと、リクエストを処理する構成のサーバ、データ・コミュニケーションのためにストレージ・システムとサーバを結合するスイッチと、障害が発生しているデバイスから受信したイベント・メッセージを解釈するためのイベント辞書を持つネットワーク・マネージャを有する。

【選択図】 図1

図1 システム50aの全体構成図



【特許請求の範囲】**【請求項 1】**

複数の納入業者からのハードウェアとソフトウェア製品を有するヘテロジニアスなネットワークであって、

データを記憶するように構成されたストレージ・システムと、

リクエストを処理するように構成されたサーバと、

データ通信を行うためにストレージ・システムとサーバを結合するスイッチと、ト辞書にあるエラーコード・リストを用いて障害が発生しているコンポーネントを判定障害が発生しているデバイスから受信したイベント・メッセージを解釈するためのイベント辞書を有するネットワーク・マネージャと、

で構成されることを特徴とするネットワーク。

【請求項 2】

ネットワークはストレージ・エリア・ネットワークであることを特徴とする請求項 1 に記載のネットワーク。

【請求項 3】

さらに、ヘテロジニアスなネットワークと結合されたメッセージング・ネットワークで構成されることを特徴とする請求項 1 に記載のネットワーク。

【請求項 4】

イベント辞書は、エラーコード・リストと前記エラーコード・リストに対応するエラー・コンポーネント・リストを含むことを特徴とする請求項 1 に記載のネットワーク。

【請求項 5】

イベント・メッセージはエラーコードを含み、ネットワーク・マネージャは、障害が発生しているコンポーネントを識別するために、イベント辞書にあるエラーコードを参照することを特徴とする請求項 4 に記載のネットワーク。

【請求項 6】

ネットワーク・マネージャが複数の納入業者に関する複数のイベント辞書を持つことを特徴とする請求項 1 に記載のネットワーク。

【請求項 7】

さらに、ネットワークにおける I/O コミュニケーションのトポロジーに関する情報を含むトポロジー・テーブルで構成され、ネットワーク・マネージャは障害の影響を判断するためにトポロジー・テーブルにアクセスすることを特徴とする請求項 1 に記載のネットワーク。

【請求項 8】

ストレージ・エリア・ネットワーク (S A N) であって、

S A N 内部にあるデバイスにおいて、障害が発生したデバイスから受信したイベント・メッセージを解釈するためのイベント辞書を有するネットワーク・マネージャで構成されることを特徴とするストレージ・エリア・ネットワーク (S A N)。

【請求項 9】

S A N は、異なるイベント・メッセージのルールを持つ複数の納入業者からのネットワーク製品を含むことを特徴とする請求項 8 に記載の S A N。

【請求項 10】

ストレージ・エリア・ネットワーク (S A N) を管理するように構成されたマネージメント・サーバであって、

S A N 内部にあるデバイスにおいて、障害が発生したデバイスから受信したイベント・メッセージを解釈するためのイベント辞書を含むネットワーク・マネージャで、

構成されることを特徴とするマネージメント・サーバ。

【請求項 11】

S A N は、複数の納入業者が個別仕様のエラーコードを用いている状態で、複数の納入業者から納入される異なるイベント・メッセージのルールを持つネットワーク製品を使用することを特徴とする請求項 10 に記載のマネージメント・サーバ。

10

20

30

40

50

【請求項 12】

ストレージ・エリア・ネットワーク(S A N)であって、データ・リクエストを処理するように構成された複数のアプリケーション・サーバと、S A Nのマネージメント機能を行うように構成され、複数のコンポーネントを有し、前記複数のコンポーネントの一つに障害が発生しているデバイス、すなわち障害が発生したデバイスからのトラップ・メッセージに含まれるエラーコードを解釈するためのトラップ辞書を含むS A Nマネージャを含むマネージメント・サーバと、データを記憶するように構成された複数のストレージ・サブシステムと、アプリケーション・サーバとストレージ・サブシステムの間でデータを転送するように構成された複数のスイッチと、で構成され、前記S A Nは、異なったエラーコードのルールをもつ複数の納入業者から納入されたネットワーク製品を含むヘテロジニアスなネットワークであることを特徴とする。10

【請求項 13】

S A Nマネージャは、異なった納入業者に対する複数のトラップ辞書を含むことを特徴とする請求項12に記載のS A N。

【請求項 14】

さらに S A Nと結合したメッセージング・ネットワークで構成されていることを特徴とする請求項12に記載のネットワーク。20

【請求項 15】

トラップ辞書は、エラーコード・リストと前記エラーコード・リストに対応するエラー・コンポーネント・リストを含むことを特徴とする請求項12に記載のS A N。20

【請求項 16】

エラーコードにより問題が発生しているデバイス内部の一つのコンポーネントが識別され、前記一つのコンポーネントを識別するためにトラップ辞書が参照されることを特徴とする請求項15に記載のS A N。

【請求項 17】

マネージメント・サーバはさらに、ネットワークにおけるI / Oコミュニケーションのトポロジーに関する情報を含むトポロジー・テーブルで構成され、前記S A N・マネージャは障害の影響を判断するために前記トポロジー・テーブルにアクセスできることを特徴とする請求項12に記載のS A N。30

【請求項 18】

ストレージ・ネットワークの管理方法であって、複数の納入業者で製造された複数のネットワーク製品を提供することと、複数のコンポーネントの一つに障害が発生し、障害が発生している一つのコンポーネントを識別するエラーコードを含むイベント・メッセージを、前記複数のコンポーネントを含むデバイスから受信することと、イベント・メッセージのエラーコードを解釈するために、エラーコード・リストとそれに対応するエラー・コンポーネント・リストを含むイベント辞書にアクセスすることと、イベンすることで、構成されることを特徴とするストレージ・ネットワークの管理方法。40

【請求項 19】

さらに、コンポーネントの障害の影響を判定するためにトポロジー・テーブルにアクセスすることで構成されることを特徴とする請求項18に記載の方法。

【請求項 20】

さらに、ネットワーク・アドミニストレータへイベント通知を伝えることで構成され、前記イベント通知は、障害が発生しているコンポーネントと障害の影響に関する情報を含んでいることを特徴とする請求項19に記載の方法。

【請求項 21】

イベント通知は、ストレージ・ネットワークのトポロジーを図示的に説明するトポロジー・ビューと、障害が発生したコンポーネントと障害の影響に関する情報を提供するイベント・サマリーを含んでいることを特徴とする請求項20に記載の方法。50

【請求項 2 2】

イベント通知はさらに、コンポーネントの障害に関わるデータ・パスを含むことを特徴とする請求項 2 1 に記載の方法。

【請求項 2 3】

ストレージ・エリア・ネットワークのマネージメント・サーバ内にあるネットワーク・マネージャであって、

障害が発生しているデバイス内にあるコンポーネントを識別し、障害の影響を判定するために、ストレージ・エリア・ネットワーク内のデバイスから受信したイベント・メッセージを解釈するための参照テーブルにアクセスするコードで構成されることを特徴とするネットワーク・マネージャ。10

【請求項 2 4】

第1のサーバ上で稼働しているストレージ・エリア・ネットワーク(S A N)マネージャが行うイベント通知方法であって、前記第1のサーバは、ネットワークを経由して、少なくとも第2のサーバ、スイッチ、ストレージ・サブシステムと結合されており、

前記第2のサーバと前記第2のサーバがアクセスするロジカル・ボリュームの間の I / O パスについての情報を前記第2のサーバから受信することと、

構成情報を前記スイッチと前記ストレージ・サブシステムから受信することと、前記第2のサーバからの情報と前記構成情報を用いて S A N についてのトポロジー情報を生成することと、

もし障害が前記スイッチあるいは前記ストレージ・サブシステムで発生したならば、前記スイッチあるいは前記ストレージ・サブシステムからイベント・メッセージを受信することと、20

トポロジー情報を用いて、第2のサーバが障害によってどのロジカル・ボリュームにアクセス出来ないかを判定すること、で構成されることを特徴とする方法。

【請求項 2 5】

さらに、マネージャのイベント辞書を用いてイベント・メッセージを解釈することで構成されることを特徴とする請求項 2 4 に記載のイベント通知方法。

【請求項 2 6】

イベント辞書は、スイッチとストレージ・サブシステムに対するエラーコード・リストと、前記エラーコード・リストに対応するエラー・コンポーネント・リストを含むことを特徴とする請求項 2 5 に記載のイベント通知方法。30

【発明の詳細な説明】**【発明の属する技術分野】**

[0 1] 本発明は、ストレージ・ネットワークに関わり、さらに詳細にはストレージ・ネットワークにおけるイベント通知の方法とシステムに関する。

【従来の技術】

[0 2] データは、全ての計算処理の基本となる基礎的な資源である。最近のインターネットと電子ビジネス(e - ビジネス)の爆発的な発展により、データ・ストレージ・システムに対する需要は飛躍的に増加している。一般に、ストレージ・ネットワークは、つぎの2つのアプリケーションあるいはコンフィギュレーションで構成される。すなはち、ネットワーク結合ストレージ(N A S)あるいはストレージ・エリア・ネットワーク(S A N)である。N A S は、イーサネット上の I P を用いて、ストレージサーバとそれらのクライアントの間でファイルフォーマットのデータ転送を行う。N A S に於いては、ディスクアレイやテープ装置のような統合ストレージ・システムは、イーサネットのようなローカル・エリア・ネットワーク(L A N)インターフェイスを通して、T C P / I P のようなメッセージ通信プロトコルを用いて、メッセージング・ネットワークに直接に接続する。ストレージ・システムは、クライアント・サーバ・システムにおけるサーバのように機能する。40

[0 3] 一般に、S A N は、異種のサーバやストレージ資源の間でデータを転送する専50

用の高性能ネットワークである。NASと異なり、従来型のメッセージング・ネットワーク上のクライアントとサーバの間でトラフィックの競合が発生しないように、専用のネットワークが設けられる。SANでは、ストレージ資源とプロセッサーまたはサーバとの間を直接に接続することが許される。

SANは、サーバ間で共有したり、特定のサーバに専用にしたりすることが出来る。それはまた、1ヵ所の場所に集中させたり、地理的に離れた場所に分散させたりすることも可能である。SANのインターフェイスは各種の異なったプロトコルが可能で、ファイバー・チャネル(FC)やエンタープライズ・システム接続(ESCON)や小型コンピュータ・システム・インターフェイス(SCSI)やシリアル・ストレージ・アーキテクチャ(SSA)や高性能パラレル・インターフェイス(HIPPI)や将来出現するであろう他のプロトコルであっても良い。たとえば、インターネット・エンジニアリング・タスク・フォース(IETF)は、TCP/IPによるブロック・ストレージを可能にする新しいプロトコルあるいは標準ISCSIを開発中であり、一方では、幾つかの会社は、ホスト・プロセッサーからiSCSI-TCP/IPプロトコルスタックをはずして、iSCSIをSANの主要な標準にしようと作業中である。10

[04] 現在は、ファイバーチャンネル(FC)は、SANについての主要な標準であり、プロトコルである。FCは、今日に於いては、1Gb/sや2Gb/sのリンクスピードで最高の性能を持ち、プロトコル・スタックを十分に軽減することで優れた(大変短い)待ち時間を実現する。したがって、ファイバーチャンネルを基本とするSANは高性能のアプリケーションにしばしば用いられる。データセンタにおいては、当分の間2Gb/sでのFCには競合するものがないと予想される。20

【発明が解決しようとする課題】

[05] 高性能であり、多目的のSANを適切に利用するためには、効率よく管理する必要がある。ストレージ・ネットワークにおける一つの重要な管理機能は、イベント通知の管理である。SANは一般に各種のベンダーからの様々なハードウェアやオペレイティングシステムを含み、それぞれが固有のメッセージ言語やルールを持っているので、SANにおけるイベント通知の管理は難しい問題である。30

【課題を解決するための手段】

[06] 本発明の実施例は、ストレージ・エリア・ネットワーク(SAN)のようなストレージ・ネットワークにおけるイベント通知およびイベント管理に関わる。ある実施例では、たとえばSANマネージャのようなネットワーク・マネージャが、ストレージ・ネットワーク内でデバイスから情報を収集する。ネットワーク内では、ネットワーク・マネージャは各デバイスについてトラップ辞書を持っている。30

辞書は、障害が発生したあるいは発生しようとしているデバイスから受信したイベント・メッセージを解釈するために用いられる。ネットワーク・マネージャは、デバイス内で問題を起こした特定の部品を確認し、そしてそのイベントの重要度を判定するように構成されている。ネットワーク・マネージャは、イベントの原因と影響度を示すイベント通知を、中央のマネジメント・コンソール上に表示するように設定されている。

[07] ある実施例では、ヘテロジニアスなネットワークは、複数の納入業者からのハードウェアやソフトウェア製品に関連したネットワークを含んでいる。ネットワークは、データを記憶する構成のストレージ・システム、リクエストを処理する構成のサーバ、データ・コミュニケーションのためにストレージ・システムとサーバを結合するスイッチ、そして障害が発生したデバイスから受信したイベント・メッセージを解釈するためのイベント辞書を持つネットワーク・マネージャを含む。40

[08] 他の実施例では、ストレージ・エリア・ネットワーク(SAN)は、SAN内にあるデバイスである、障害が発生したデバイスから受診したイベント・メッセージを解釈するためのイベント辞書を持つネットワーク・マネージャを含んでいる。

[09] 他の実施例では、ストレージ・エリア・ネットワーク(SAN)を管理する構成のマネジメント・サーバは、SAN内にあるデバイスである、障害が発生したデバイスから受診したイベント・メッセージを解釈するためのイベント辞書を持つネットワーク50

・マネージャを含んでいる。

[10] 他の実施例では、ストレージ・エリア・ネットワーク(SAN)は、データ・リクエストを処理する構成の複数のアプリケーション・サーバを含んでいる。マネージメント・サーバは、SANのマネージメント機能を処理するように構成され、SANマネージャを含んでいる。SANマネージャは、障害が発生したデバイスからのトラップ・メッセージに含まれるエラーコードを解釈するトラップ辞書を含んでいる。デバイスは複数のコンポーネントを有し、その一つに障害が発生している。複数のストレージ・サブシステムはデータを記憶するように構成されている。複数のスイッチは、アプリケーション・サーバとストレージ・サブシステム間でデータを転送するように構成されている。SANは、複数の納入業者からの、エラーコードに対して異なるルールを持つネットワーク製品を含む、ヘテロジニアスなネットワークである。10

[11] さらに他の実施例では、ストレージ・ネットワークを管理する方法は、複数の納入業者で製造された複数のネットワーク製品を提供することを含んでいる。イベント・メッセージは、複数の部品を持つデバイスから受信され、複数の部品の中の一つに障害が発生している。イベント・メッセージは、障害が発生している1個の部品を識別するエラーコードを含んでいる。イベント・メッセージにあるエラーコードを解釈するために、イベント辞書がアクセスされる。イベント辞書は、エラーコード・リストとそれに対応するエラー・コンポーネント・リストを含んでいる。障害が発生した部品の識別は、イベント辞書にあるエラーコード・リストを用いて決定される。20

【発明の実施の形態】

[31] 本発明はストレージ・エリア・ネットワーク(SAN)やネットワーク結合ネットワーク(NAS)のような、ストレージ・ネットワークにおけるイベント通知の管理に関する。特に、本発明は、ヘテロジニアスなハードウェアおよび/あるいはソフトウェア・システムを用いるストレージ・ネットワークにおけるイベント通知の管理に関する。ヘテロジニアスなシステムは、複数の納入業者からのハードウェアまたはソフトウェア製品、あるいは双方の製品を有している。説明の便宜上、本発明の特定の実施例をSANを用いて下記により説明するが、本発明の適用範囲を狭めるものではない。30

[32] ここに用いられる「SAN」あるいは「サブ・ネットワーク」という用語は、メッセージング・ネットワークに結合された、集中的に管理される高速ストレージ・ネットワークを意味し、そしてマルチベンダー・ストレージ・デバイス、マルチベンダー・ストレージ・マネージメント・ソフトウェア、マルチベンダー・サーバ、マルチベンダー・スイッチ、あるいは他のマルチベンダー・ネットワーク関連のハードウェアおよびソフトウェア製品を含む。SANあるいはサブ・ネットワークのヘテロジニアスな性質の範囲は変動する。幾つかのSANあるいはサブ・ネットワークは、上記の全てのネットワーク・デバイスやコンポーネントについてマルチベンダーの製品を有するが、一方では、他のものは、上記のデバイスやコンポーネントの一部についてマルチベンダーの製品を有する。40

[33] ここに用いられる「ストレージ・ネットワーク」という用語は、1ないしはそれ以上のストレージ・システムに結合されたネットワークを意味し、そしてマルチベンダー・ストレージ・デバイス、マルチベンダー・ストレージ・マネージメント・ソフトウェア、マルチベンダー・アプリケーション・サーバ、マルチベンダー・スイッチ、あるいは他のマルチベンダー・ネットワーク関連のハードウェアおよびソフトウェア製品を含む。一般に、「ストレージ・ネットワーク」は、他のネットワーク、たとえばメッセージング・ネットワークに結合され、そしてフロントエンド・サーバ・アプリケーションからバックエンド・ストレージ機能を切り離している。したがって、ストレージ・ネットワークはSANやNAS等を含んでいる。

[34] 図1は複数のサーバ106を複数のストレージ・システム108に接続する1ないしはそれ以上のメッセージング・ネットワーク102とSAN104を含むネットワーク・システム100の説明図である。ネットワーク102は、ローカル・エリア・ネットワークや広域ネットワークやインターネットなどである。もし必要な場合は、ネットワーク102は、管理を容易にしその費用を低減するために、ストレージ・デバイス10850

が集中化し、サーバ106がクラスター化することを可能にしている。

[35] SAN104は、いろいろなやり方で、サーバ106とストレージ・デバイス108の間の直接で高速のデータ転送をサポートしている。データはサーバとストレージ・デバイスの間で転送される。特定のストレージ・デバイスが複数のサーバから逐次にあるいは同時にアクセスされてもよい。データはサーバ間で転送することも出来る。さらに、データをストレージ・デバイス間で転送することも出来、サーバの介入なしにデータを転送することを可能にする。それによりサーバは別の仕事を行うことが出来る。たとえば、ストレージ・システムは、サーバの介入無しに、既定の間隔で、他のストレージ・システムにデータをバックアップすることが出来る。

[36] したがって、ストレージ・デバイスあるいはサブシステム108は特定のサーバ・バス専用ではなく、直接SAN104に接続される。ストレージ・サブシステム108は外部に置かれ、機能的に全組織にわたって分配される。 10

[37] 本発明の一実施例では、SAN104はストレージ・インターフェイスで構成され、サーバ106を経由してネットワーク102と結合される。したがって、SANはサーバの背後にあるネットワークあるいはサブ・ネットワークを意味する。

[38] 他の実施例では、SANは、1台ないしはそれ以上のサーバと、1個ないしはそれ以上のSANスイッチまたはファブリックと、1台ないしはそれ以上のストレージ・システムを含むものとして定義される。さらに他の実施例では、SANは、1台ないしはそれ以上のサーバと、1個ないしはそれ以上のSANスイッチまたはファブリックと、1台ないしはそれ以上のストレージ・システムのポートを含むものとして定義される。したがって、SANという用語は、上記の定義を満足する限り、種々の異なったネットワーク・コンフィギュレーションを意味するために用いられる。 20

[39] 図2Aは、ストレージ・システム(あるいはサブシステム)202、SANスイッチあるいはファブリック204、複数のサーバ206aと206b、マネージメント・サーバ208、そしてマネージメント・ネットワーク210を有するSANシステム200を説明している。以下に詳細を説明するが、マネージメント・サーバ208はSANを管理するSANマネージャ209を含んでいる。単一のストレージ・サブシステムがSANシステム200では説明されているが、他の実施例では複数のストレージ・サブシステムが用いられている。同様に、他の実施例では、他のネットワーク・コンポーネントの数は、説明のサンプルとは異なっている可能性がある。 30

[40] ストレージ・サブシステム202には、マネージメント・エージェント212、複数のディスク・ポート214aと214b、複数のロジカル・デバイス216aと216b、そして複数のキャッシュ218aと218bが含まれる。ディスク・ポート214aと214bは、またディスク・ポートd1およびd2とも呼ばれる。ロジカル・デバイス216aと216bはまた、ロジカル・デバイスv1とv2と呼ばれる。マネージメント・エージェント212は、ストレージ・サブシステムの構成を管理し、マネージメント・サーバ208と交信する。たとえば、エージェント212は、以下に詳細を説明するように、マネージメント・サーバ208に対して、ストレージ・サブシステム202においてデータI/Oバス、ディスク・ポートd1とd2の接続情報、およびコンポーネントで発生する障害を伝える。ディスク・ポート214aと214bは、サーバ206aと206bとデータを送受信するためのSANスイッチ204への接続ポートである。本実施例に用いられる接続プロトコルは、ファイバー・チャネルであるが、他のプロトコル、たとえばSCSIやIP上のFCまたはiSCSIなどが使用されてもよい。 40

[41] この技術分野に詳しい人にはよく知られているように、マネージメント・エージェント212は、ディスク・ポート・テーブル220、デバイス・テーブル222、およびバス・テーブル224を含んでいる(図2B)。これらのテーブルは、「コンフィギュレーション」情報が変わると定期的に更新される。ディスク・ポート・テーブル220は、「ニックネーム」のようなストレージ・サブシステムにおけるディスク・ポートに関する情報を提供するディスク・ポートID226と、各ディスク・ポートに対する独自の識別子を提供するワールド・ワイド・ネーム(WWN)228を含む(図2C)。二 50

ニックネームは、たとえば「d 1」がディスク・ポート214aを意味するように、ストレージ・サブシステムの特定の識別名を意味する。「d 1」という名称は、問題のストレージ・サブシステム内でディスク・ポートを識別するには十分であるが、複数のストレージ・サブシステムがある場合には、他のストレージ・サブシステムのディスク・ポートに同じ名称が割り当てられている可能性があるので、不十分である。一方、独自の識別子（ファイバー・チャネルではワールド・ワイド・ネームという）は、特定のコンポーネントに割り当てられた独自の識別情報である。

[42] デバイス・テーブル222は、ストレージ・サブシステム内のロジカル・デバイスとディスク・ドライブの間の関係についての情報を提供するロジカル・デバイスID230と、ディスク・ドライブ・リスト232を含む（図2D）。バス・テーブル224は、バスについてのニックネームを提供するバスID234、バスに接続されたディスク・ポートを識別するディスク・ポートID236、バスに接続されたキャッシュを識別するキャッシュID238、バスに接続されたロジカル・デバイスのニックネームを提供するロジカル・デバイスID240、バスに接続されたSCSIを識別するSCSI ID242、およびバスに接続されたSCSI LUNについての情報を提供するSCSI LUN244を含む（図2E）。

[43] ロジカル・デバイス216aと216bは、サーバにエクスポートされたボリュームである。ロジカル・デバイスは、単一の物理ディスク・ドライブ、あるいはRAID（Redundant Array of Independent Disks）においては複数の物理ディスク・ドライブで構成される。RAIDストレージ・システムにより、データの可用性を増し、また入出力（I/O）の性能を向上することが出来る。RAIDシステムでは、複数の物理ディスク・ドライブが一つの論理ディスク・ドライブとして構成され、論理ディスク・ドライブへのI/Oリクエストは、ストレージ・システムの内部で物理ディスク・ドライブに分散され、並列に処理される。RAID技術には多くの利点がある。たとえば、RAIDストレージ・システムは大変大きなファイル・システムを収容することが出来るので、大きなファイルを幾つかの小さいファイル・システムに分割することなく、単一のファイル・システムに記憶することが出来る。さらに、RAID技術では、異なった物理ディスク上のデータを並列にアクセス出来るので、I/O性能を向上出来る。ある実施例では、図2Bに示すように、各ロジカル・デバイスは4台の物理ディスク・ドライブdd1、dd2、dd3、そしてdd4を含んでいる。

[44] キャッシュ218aと218bは、ロジカル・デバイス216aと216bに関連したデータ・キャッシュである。これらはデータ処理速度を向上するために設けられている。他の実施例では、ストレージ・サブシステムにはキャッシュが含まれていない。

[45] 図3Aを見ると、SANスイッチ204は、サーバとストレージ・サブシステムを接続している。スイッチ204はサーバとストレージ・サブシステム間のデータ接続を行っている。ある実施例では、スイッチは、ネットワークの対象を拡大するために、ブリッジやルータやその他のネットワーク・ハードウェアと結合される場合もある。スイッチ204は、スイッチの構成を管理するスイッチ・マネージメント・エージェント302と複数のスイッチ・ポート304a、304b、304c、および304dを有している。これらのスイッチ・ポートはまた、図3Aに示すように、それぞれs1、s2、s3、およびs4と呼ばれる。スイッチ・マネージメント・エージェント302は、サーバ208にスイッチ・ポートの接続情報を提供し、もしスイッチ204の内部のコンポーネントに障害が発生した場合は、それをサーバ208に通知することで、マネージメント・サーバ208の行うSAN管理を援助する。マネージメント・エージェント302は、スイッチ（「link」とも呼ばれる）を経由して、サーバとストレージ・サブシステム間の相互関係についての情報を提供するポート・リンク・テーブル306を有している。ポート・リンク・テーブル306は、各スイッチ・ポートの識別情報またはニックネームを提供するスイッチ・ポートID308と、各スイッチ・ポートの独自の識別子を提供するスイッチ・ポート・ワールド・ワイド・ネーム（WWN）310と、各スイッチ・ポートに接続されているターゲット・デバイスの独自の識別子を提供するリンクWWN312を含ん

10

20

30

40

50

でいる（図3B）。

[46] 図4Aはサーバ206aと206bの応用例についてさらに詳細に説明している。本実施例では、別個のサーバが応用と管理機能を実行するために用いられている。各サーバ206は、サーバの構成を管理するサーバ・マネージメント・エージェント402と、データ接続のためのサーバ・ポート404を含んでいる。サーバ・マネージメント・エージェント402は、サーバ・ポートの接続情報をサーバ208に提供し、もしサーバ206の内部のコンポーネントに障害が発生した場合は、それをサーバ208に伝えることで、マネージメント・サーバ208のSAN管理を援助する。エージェント402は、一般には便宜上サーバ内部に設置される。また、エージェント402は、ホスト・ポート・テーブル406とLUNバインディング・テーブル408を含んでいる。ホスト・ポート・テーブルは、サーバのホストまたはサーバ・ポートに関する情報を提供する。10

[47] 図4Bを参照すると、エージェント402aにあるホスト・ポート・テーブル406aは、サーバのポートについての情報を記録するための複数の列を持っている。テーブル406aは、サーバ内の特定のポートに対するデバイス特定識別情報またはニックネームを提供するホスト・ポートID410aと、独自のポート識別情報を提供するワールド・ワイド・ネーム412aと、ネットワーク・アドミニストレータにより特定のポートに割り当てられるSCSI識別情報を提供するSCSI_ID414aを含んでいる。

一般に、SANのサーバ・ポートには単一のSCSI_IDが割り当てられる。ワールド・ワイド・ネーム412aはファイバー・チャネルに関連して用いられる用語であって、異なった接続プロトコルが使用されると、それに相当する用語が使用されることになる。20 図4Cは、エージェント402bに設けられたホスト・ポート・テーブル406bを示す。ホスト・ポート・テーブル406bは、ホスト・ポートID410bと、ワールド・ワイド・ネーム412bと、SCSI_ID414bを含む。

[48] 図4Dを参照すると、エージェント402aに設けられたLUNバインディング・テーブル408aは、「LUNバインディング」とも「バインディング」とも呼ばれる、ホスト・ポートからSCSIロジカル・ユニットへのデータI/Oバスについての情報を提供する。テーブル408aは、バインディングのデバイス特定識別情報またはニックネームを提供するバインディングID416aと、特定のポートのニックネームを提供しテーブル406aのホスト・ポートID410aに対応するホスト・ポートID418aと、バインディングに付属しテーブル406aのSCSI_ID414aに対応するSCSI_ID420aと、バインディングに付属するSCSI_LUNを提供するLUN422aと、サーバがSCSI_INQUIRYコマンドをLUNに発行したときに、LUNから与えられる情報を提供するインクライアリ情報424aを含んでいる。インクライアリ情報は、一般に、納入業者名、製品名、LUNのロジカル・デバイスIDなどを含んでいる。30 図4Eは、エージェント402bに設けられたLUNバインディング・テーブル408bを示す。LUNバインディング・テーブル408bは、バインディングID416bと、ホスト・ポートID418bと、SCSI_ID420bと、LUN422bと、インクライアリ情報424bを含んでいる。

[49] 図5Aは、本発明の一実施例に対応するSANのマネージメント関連機能に専用のマネージメント・サーバ208を示している。他の実施例では、単一のサーバがアプリケーション・サーバとマネージメント・サーバの二重の機能を果たすこともある。40

[50] マネージメント・サーバ208は、ネットワークを効率よく利用出来るようにSANを管理するために用いられるSANマネージャあるいはネットワーク・マネージャ502を含んでいる。マネージャ502は、SAN内部の各種のコンポーネントから得たすべての物理的および論理的な接続情報を含んでいる。したがって、マネージャ502は、マネージメント・ネットワーク210を経由して、それぞれの構成テーブルを得るために、SAN内部で、たとえばスイッチ・マネージメント・エージェント302、サーバ・マネージメント・エージェント402、およびストレージ・システム・マネージメント・エージェント212などのマネージメント・エージェントと交信を行う。したがって、SANマネージャあるいはネットワーク・マネージャ502は、トポロジー・リポジトリ50

04とディスカバリ・リスト506を含んでいる。

[51] トポロジー・リポジトリ504は、SANにおいてI/O通信のトポロジーを提供するトポロジー・テーブル508を含んでいる。トポロジー・テーブル508は、SAN内部のデバイスから得られる、たとえばホスト・ポート・テーブル、LUNバインディング・テーブルなどのテーブルをマージすることで作られる。図5Bを参照すると、トポロジー・テーブルは、SANにおけるサーバについてのバインディングIDとホスト・ポートIDの情報を提供するサーバ・セクション550と、SANにおけるスイッチについての情報を提供するインターフェクト・セクション552と、ディスク・ポートIDとキャッシュIDとロジカル・デバイスIDを含むストレージ・サブシステムについての情報を提供するストレージ・セクション554を含んでいる。10

[52] 図5Cは、本発明の一実施例にしたがって、トポロジー・テーブル508を生成するためにSANマネージャ502が行う処理564を示している。SAN内の全てのデバイスが検出される(ステップ566)。検出された各々のデバイスの構成情報が読み出され、トポロジー・リポジトリに記憶される。各LUNバインディング・エントリーは、すべてのエントリーが検索されるまで検索される(ステップ568)。それぞれのエントリーについて、トポロジー・テーブルに新しいエントリーが作成され、たとえばサーバ名やサーバ・バインディングIDやサーバ・ホスト・ポートIDなどのサーバ情報がそこに記憶される。サーバ(ホスト・ポートID X)とSANスイッチ(スイッチ・ポートID Y)間の接続が検出され、接続情報がエントリーに記憶される(ステップ572)。このステップは、キー「ホスト・ポートID」がホスト・ポートID XであるWWNをホスト・ポート・テーブルから選択すること、キー「リンクWWN」がホスト・ポート・テーブルで選択されたWWN、言い換えればホスト・ポートID XのWWN、と等しいスイッチ・ポートID Yをポート・リンク・テーブルから選択すること、ポート・リンク・テーブルの選択されたポート・リンク・エントリーから、「インターフェクト名」と「インターフェクト・ポートID」をコピーすることを伴う。20

[53] その後で、ロジカル・デバイス情報はエントリーに記憶される(ステップ574)。このステップは、キー「ロジカル・デバイスID」と「SCSI ID」と「SCSI LUN」がLUNバインディング・テーブルのエントリーにあるものと同等であるパスをパス・テーブルから選択すること、パス・テーブルの選択されたパスから「ストレージ名」と「ストレージ・ディスク・ポートID」と「ストレージ・キャッシュ・ID」と「ストレージ・ロジカル・デバイスID」をコピーすることを伴う。30

[54] つぎに、ストレージ(ディスク・ポートID X)とSANスイッチ(スイッチ・ポートID Y)の間の接続が検出され、接続情報がエントリーに記憶される(ステップ576)。このステップは、キー「ディスク・ポートID」がディスク・ポートID XであるWWNをディスク・ポート・テーブルから選択すること、キー「リンクWWN」がディスク・ポート・テーブルにおいて選択されたWWN、言い換えればディスク・ポートID XのWWN、と同等であるスイッチ・ポートID Yをポート・リンク・テーブルから選択すること、ポート・リンク・テーブルにおいて選択されたポート・リンク・エントリーから、右側の「インターフェクト名」と「インターフェクト・ポートID」をコピーすることを伴う。ステップ576が終わると、次のLUNバインディング・エントリーが検索され(ステップ578)、上記のステップが、全てのエントリーが処理されるまで繰り返される。40

[55] ディスカバリ・リスト506は、SANにおける全てのデバイスの情報を含んでいる。SANマネージャ502は、このリストからの情報をSANデバイスにあるマネージメント・エージェントから構成情報を検索するために用いる。図5Dを参照すると、ディスカバリ・リストは、発見されるべきターゲットSANデバイスのニックネームを提供するディスカバリIDセクション556と、ターゲットSANデバイスのデバイス・タイプを識別するデバイス・タイプ・セクション558と、ターゲットSANデバイスについての納入業者情報あるいは他の詳細情報を提供するデバイス情報セクション560と、SANマネージャとターゲットSANデバイス間のコミュニケーションのためにターゲッ50

ト SAN デバイスの IP アドレスを提供する IP アドレス・セクション 562 を含む。本実施例では、使用されるコミュニケーション・プロトコルは TCP / IP である。

[56] SAN マネージャ 502 は、一ないしはそれ以上のトラップ辞書（下記に説明）を、上記のトポロジー・テーブルとディスカバリ・リストと同様に用いて、イベント・マネージメントを行うように構成されている。マネージャ 502 は、問題が発生している SAN のコンポーネントからイベント・メッセージを受信する。イベントはネットワーク・アドミニストレータに通知されて、適当な対策を取るようにする。イベント通知のために使用される一つの共通したプロトコルは、IP をベースにしたプロトコルである SNMP (Simple Network Management Protocol) である。本実施例では、マネージャ 502 は SNMP メッセージを扱うように構成されている。10

[57] 稼働状態では、問題が発生しているデバイスは、マネージャ 502 に SNMP トラップ・メッセージを送信する。マネージャ 502 は、メッセージを受信すると、問題の原因と SAN におけるイベントあるいは問題が結果としてどのように影響するかを判断することができる。たとえば、もし障害が SAN スイッチ 204 のスイッチ・ポート 304a で発生したならば、マネージャ 502 は、スイッチ・ポート 304a に障害が発生したのでイベント・メッセージを受信したことと、この障害によりサーバ 206a がロジカル・デバイス v1 にアクセス出来なくなること、を判断することができる。SAN は、様々なメッセージのルールを持つ、多くの異なる納入業者からのハードウェアとソフトウェアを含んでいるので、在来の SAN マネージャではイベントの原因と影響に関するこのような正確な診断は不可能であった。したがって、同様の状況では、在来の SAN マネージャは、ネットワーク・アドミニストレータに対して単に SAN スイッチ 204 に障害が発生していることくらいしか伝えられない。20

[58] イベントの原因と影響についてこのような正確な診断を行うために、マネージャ 502 は、マネージャ 502 が受信したトラップ・メッセージを解読しあるいは解釈するための一ないしはそれ以上のトラップ辞書（また「イベント辞書」あるいは「ルックアップ・テーブル」とも呼ばれる）を持っている。本実施例では、マネージャ 502 は、種々のハードウェアやソフトウェアの納入業者に対して、複数のトラップ辞書を持っている。トラップ辞書を記憶することに關し幾つかの異なった方法がある。トラップ辞書はデバイス・タイプに依存して記憶されてもよく、SAN スイッチに関するトラップ辞書の全てが一ヵ所に記憶される。別の方針としては、トラップ辞書を納入業者の指定するファイルに従って記憶してもよい。30

[59] 本実施例では、トラップ辞書はデバイス・タイプにしたがって記憶される。したがって、マネージャ 502 には、SAN スイッチ用のトラップ辞書 510 とストレージ・サブシステム用のトラップ辞書 512 がある。スイッチ・トラップ辞書 510 は、特別なイベントが発生したこと通知するトラップ・メッセージに付加されるエラーコード 602 と、問題が発生しているコンポーネントを識別するエラー・コンポーネント 604 を含んでいる（図 6A）。たとえば、もし問題が SAN スイッチ 204 のポート s1 に発生すると、エラーコード「A1」を含むトラップ・メッセージがマネージャ 502 に送られる。マネージャ 502 は、スイッチ・トラップ辞書 510 を参照してエラーコードの意味を判定する。40

[60] 同様に、ストレージ・トラップ辞書 512 は特別なイベントが発生したことを通知するトラップ・メッセージに付加されるエラーコード 606 と、問題が発生しているコンポーネントを識別するエラー・コンポーネント 608 と、コンポーネント ID 情報を示す ID 610 を含んでいる。本発明の一実施例では、マネージメント・サーバは、トラップ・メッセージを受信した時に、適切なトラップ辞書を参照するために用いるディクショナリー・サーバ 512 を含んでいる。

[61] 図 7 は、本発明の一実施例において、SAN における SAN マネージャ 502 を用いたイベント通知の処理を説明するフローチャート 700 である。マネージャ 502 は、障害が発生したデバイスから SNMP トラップ・メッセージを受信する（ステップ 702）。デバイスには、複数のコンポーネントがあり、その中の 1 台に障害が発生してい50

る。トラップ・メッセージは、問題が発生しているコンポーネントそのものを確定するのに適したエラーコードを含んでいる。マネージャ502は、問題のデバイスを識別するために、ディスカバリ・リストを使用してSNMPトラップのIPアドレスを調べる(ステップ704)。もし、デバイスに対するトラップ辞書が存在するならば、メッセージのエラーコードが参照され、そして問題が発生しているデバイス内の特定のコンポーネントが識別される(ステップ706)。問題が発生しているコンポーネントについては、トポロジー・リポジトリのトポロジー・テーブルを使用して参照される(ステップ708)。もしトポロジーの問題が存在するならば、問題がユーザに通知される(ステップ710)。図8Aは、本発明の一実施例における、SNMPトラップ・メッセージ802の例を説明している。トラップ・メッセージは、ヘッダー804と、問題のデバイスの納入業者を識別するためのエンタプライズ・セクション806と、問題のデバイスのIPアドレスを提供するエージェント・セクション808と、特定のイベントに関連したエラーコードについてのバリアブル・バインディング810を含んでいる。

[63] 図8Bは、ストレージ・サブシステム202におけるディスク・ドライブの障害に対応してマネージャ502へ送信されるトラップ・メッセージ812を説明している。メッセージ812は、エンタプライズ806で、障害が発生したデバイスは納入業者Dによって製造されたストレージ・サブシステムであることを示し、エージェント・アドレス808はデバイスのIPアドレスは100.100.100.103であることを示し、バリアブル・バインディング・セクション810はデバイスにおいて問題が発生したコンポーネントはディスク・ドライブdd1であることを示している。マネージャ502は、トポロジー・テーブルを調べ、そしてディスク・ドライブdd1で発生した障害によりロジカル・デバイスv1に障害が発生したと判定する。マネージャ502はまた、この障害の結果として、サーバ206aはロジカル・デバイスv1にアクセスできないと判定する。マネージャ502は、ディスク・ドライブdd1の障害と、サーバ206aがロジカル・デバイスv1にアクセス出来ないという内容のイベント通知をネットワーク・アドミニストレータに送る。このイベント通知は、テキスト形式か、グラフィックのイラストか、もしくは両者の組み合わせでよい。

[64] 図9Aは、本発明の一実施例において、上記で説明したイベントの発生を担当者に伝えるために、ネットワーク・アドミニストレータに送られるイベント通知902を示している。イベント通知は、SANトポロジーのグラフィック・イラストレーションを提供するトポロジー・ビュー904と、イベントに影響されるデータ・パス906と、障害が発生しているコンポーネント908と、障害が発生したコンポーネントとその障害の影響について詳しく説明するイベント・サマリー910とを含んでいる。

[65] 図8Cは、SANスイッチのポートの障害に応じて、マネージャ502へ送信されるトラップ・メッセージ814を説明している。メッセージ814は、エンタプライズ806で、障害が発生しているデバイスは納入業者Cが製造したSANスイッチであることを示し、エージェント・アドレス808はデバイスのIPアドレスが100.100.100.102であることを示し、バリアブル・バインディング・セクション810はデバイスで問題が発生しているコンポーネントがスイッチ・ポートs1であることを示している。マネージャ502は、トポロジー・テーブルを用いて、スイッチ・ポートs1の障害でサーバ206aがロジカル・デバイスv1にアクセス出来ないことを判定する。マネージャ502は、スイッチ・ポートs1の障害と、その影響でサーバ206aがロジカル・デバイスv1にアクセス出来ないという情報を含めたイベント通知を、ネットワーク・アドミニストレータに送る。

[66] 図9Bは、本発明の一実施例において、上記のイベントを通知するために、ネットワーク・アドミニストレータに表示されるイベント通知912を説明している。イベント通知は、SANトポロジーのグラフィック・イラストレーションを提供するトポロジー・ビュー914と、イベントに影響されるデータ・パス916と、障害が発生しているコンポーネント918と、故障したコンポーネントとその障害の影響について詳しく説明するイベント・サマリー920とを含んでいる。

10

20

30

40

50

[67] 以上の詳細な説明は、本発明の特定の実施例を説明するためのものであって、それに限定されるものではない。本発明の範囲において数多くの修正や変形が可能である。したがって、本発明は以下の特許請求項により定義されるものである。

【図面の簡単な説明】

【図 1】図 1 は、本発明の一実施例による、メッセージング・ネットワークに結合されたストレージ・ネットワークを含むネットワークの説明図を示す。

【図 2 A】図 2 A は、本発明の一実施例による、マネージメント・サーバと SAN マネージャを含むストレージ・エリア・ネットワークの説明図を示す。

【図 2 B】図 2 B は、本発明の一実施例による、SAN のストレージ・サブシステムの説明である。
10

【図 2 C】図 2 C は、本発明の一実施例による、図 2 B のストレージ・サブシステムのマネージメント・エージェントに備えられているディスク・ポート・テーブルを示す。

【図 2 D】図 2 D は、本発明の一実施例による、図 2 B のストレージ・サブシステムのマネージメント・エージェントに備えられているデバイス・テーブルを示す。

【図 2 E】図 2 E は、本発明の一実施例による、図 2 B のストレージ・サブシステムのマネージメント・エージェントに備えられているパス・テーブルを示す。

【図 3 A】図 3 A は、本発明の一実施例による、SAN の SAN スイッチの説明図を示す。
20

【図 3 B】図 3 B は、本発明の一実施例による、図 3 A の SAN スイッチのマネージメント・エージェントに備えられているポート・リンク・テーブルを示す。

【図 4 A】図 4 A は、本発明の一実施例による、SAN のアプリケーション・サーバの説明図を示す。

【図 4 B】図 4 B は、本発明の一実施例によるアプリケーション・サーバのホスト・ポート・テーブルの説明図を示す。

【図 4 C】図 4 C は、本発明の一実施例によるアプリケーション・サーバのホスト・ポート・テーブルの説明図を示す。

【図 4 D】図 4 D は、本発明の一実施例によるアプリケーション・サーバの LUN バインディング・テーブルの説明図を示す。

【図 4 E】図 4 E は、本発明の一実施例によるアプリケーション・サーバの LUN バインディング・テーブルの説明図を示す。
30

【図 5 A】図 5 A は、本発明の一実施例による、SAN のマネージメント・サーバの説明図を示す。

【図 5 B】図 5 B は、本発明の一実施例による、SAN マネージャのトポロジー・テーブルの説明である。

【図 5 C】図 5 C は、本発明の一実施例による図 5 B のトポロジー・テーブルの生成手順の説明である。

【図 5 D】図 5 D は、本発明の一実施例による、SAN マネージャのディスカバリ・リストの説明である。

【図 6 A】図 6 A は、本発明の一実施例によるストレージ・サブシステムと SAN マネージャの SAN スイッチに対するトラップ辞書の説明である
40

【図 6 B】図 6 B は、本発明の一実施例によるストレージ・サブシステムと SAN マネージャの SAN スイッチに対するトラップ辞書の説明である

【図 7】図 7 は、本発明の一実施例によるイベント通知方法を説明するフロー図である。

【図 8 A】図 8 A は、本発明の一実施例による、エラーコードを含んだトラップ・メッセージの説明である。

【図 8 B】図 8 B は、本発明の一実施例による、エラーコードを含んだトラップ・メッセージの説明である。

【図 8 C】図 8 C は、本発明の一実施例による、エラーコードを含んだトラップ・メッセージの説明である。

【図 9 A】図 9 A は、本発明の一実施例による、SAN マネージャがネットワーク・アド
50

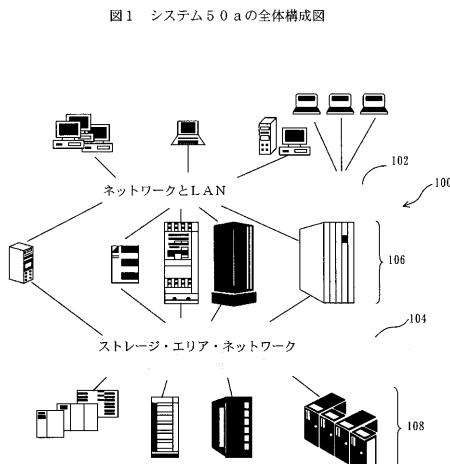
ミニストレータに出すイベント通知の説明図である。

【図9B】図9Bは、本発明の一実施例による、SANマネージャがネットワーク・アドミニストレータに出すイベント通知の説明図である。

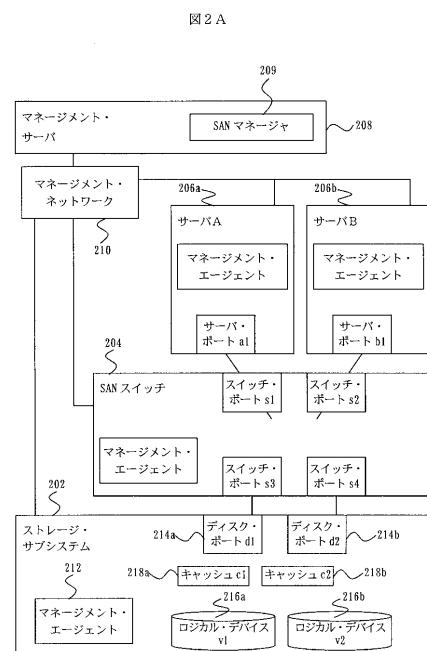
【符号の説明】

102・・・ネットワークとLAN、104・・・ストレージ・エリア・ネットワーク、
 208・・・マネージメント・サーバ、209・・・SANマネージャ、210・・・マ
 ネージメント・ネットワーク、206a・・・サーバA、マネージメント・エージェン
 ト、サーバ・ポートa1、204・・・SANスイッチ、マネージメント・エージェン
 ト、スイッチ・ポートs1、202・・・ストレージ・サブシステム、212・・・マ
 ネージメント・エージェント、214a・・・ディスク・ポートd1、218a・・・
 キャッシュc1、216a・・・ロジカル・デバイスv1、226・・・ディスク・
 ポートID、230・・・ロジカル・デバイスID、222・・・ディスク・ドライブ・
 リスト、234・・・パスID、236・・・ディスク・ポートID、238・・・キャ
 ッシュID、240・・・ロジカル・デバイスID、242・・・SCSI ID、2
 44・・・SCSI LUN、302・・・マネージメント・エージェント、
 304・・・スイッチ・ポート、402a・・・マネージメント・エージェント、40
 6a・・・ホスト・ポート・テープル、408a・・・LUNバインディング・テープル
 、404・・・サーバ・ポートa1、410a・・・ホスト・ポートID、412a..
 .WWN、414a・・・SCSI ID、416a・・・バインディングID、418
 a・・・ホスト・ポートID、420a・・・SCSI ID、422a・・・LUN 10
 20

【図1】

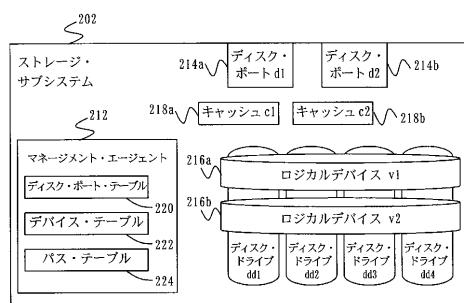


【図2A】



【図2B】

図2B



【図2C】

図2C

| ディスク・ポート ID | WWN |
|-------------|---------|
| d 1 | WWN d 1 |
| d 2 | WWN d 2 |

226

228

202

212

214a

214b

218a

218b

216a

216b

220

222

224

【図2D】

図2D

| ロジカル・デバイス ID | ディスク・ドライブ・リスト |
|--------------|------------------------|
| v 1 | dd 1, dd 2, dd 3, dd 4 |
| v 2 | dd 1, dd 2, dd 3, dd 4 |

230

232

222

【図2E】

図2E

| バス ID | ディスク・ポート ID | キャッシュ ID | ロジカル・デバイス ID | SCSI ID | SCSI LUN |
|-------|-------------|----------|--------------|---------|----------|
| p 1 | d 1 | c 1 | v 1 | 2 | 1 |
| p 2 | d 1 | c 1 | v 2 | 3 | 1 |
| p 3 | d 2 | c 2 | v 2 | 2 | 2 |

234

236

238

240

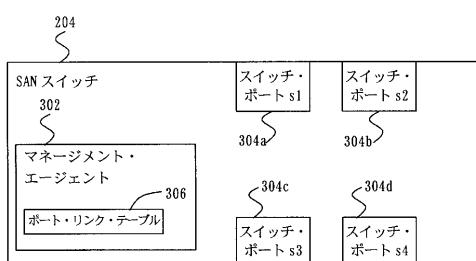
242

244

224

【図3A】

図3A



204

302

306

304a

304b

304c

304d

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

304

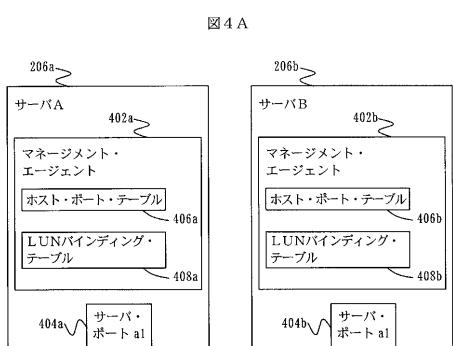
304

304

304

304

【図4A】



【図4B】

図4B

| ホスト・ポートID | WWN | SCSI ID |
|-----------|---------|---------|
| a 1 | WWN a 1 | 2 |

406a
410a
412a
414a

【図4C】

図4C

| ホスト・ポートID | WWN | SCSI ID |
|-----------|---------|---------|
| b 1 | WWN b 1 | 3 |

410b
412b
414b

【図4D】

図4D

| バインディングID | ホスト・ポートID | SCSI ID | LUN | インクアイアリ情報 |
|-----------|-----------|---------|-----|----------------------------|
| LU1 | a 1 | 2 | 1 | 納入業者D、ストレージD、ロジカル・デバイス v 1 |
| LU2 | a 1 | 2 | 2 | 納入業者D、ストレージD、ロジカル・デバイス v 2 |

416a
418a
420a
422a
424a

408a

【図4E】

図4E

| バインディングID | ホスト・ポートID | SCSI ID | LUN | インクアイアリ情報 |
|-----------|-----------|---------|-----|----------------------------|
| LU1 | b 1 | 3 | 1 | 納入業者D、ストレージD、ロジカル・デバイス v 2 |

416b
418b
420b
422b
424b

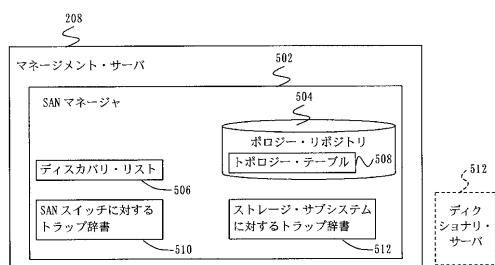
【図5B】

図5B

| サーバ名 | バインディングID | ホスト・ポートID | インターフェクト | | | ストレージ | | | | | |
|----------|-----------|-----------|----------|-------|-------|--------|-----|-------|------------|---------|-------------|
| | | | 名 | 名 | ポートID | 名 | 名 | ポートID | ディスク・ポートID | キャッシュID | ロジカル・デバイスID |
| サーバA LU1 | a 1 | スイッチC | s 1 | スイッチC | s 3 | ストレージD | d 1 | c 1 | v 1 | | |
| サーバA LU2 | a 1 | スイッチC | s 1 | スイッチC | s 4 | ストレージD | d 2 | c 1 | v 2 | | |
| サーバB LU1 | b 1 | スイッチC | s 2 | スイッチC | s 4 | ストレージD | d 2 | c 2 | v 2 | | |

【図5A】

図5A

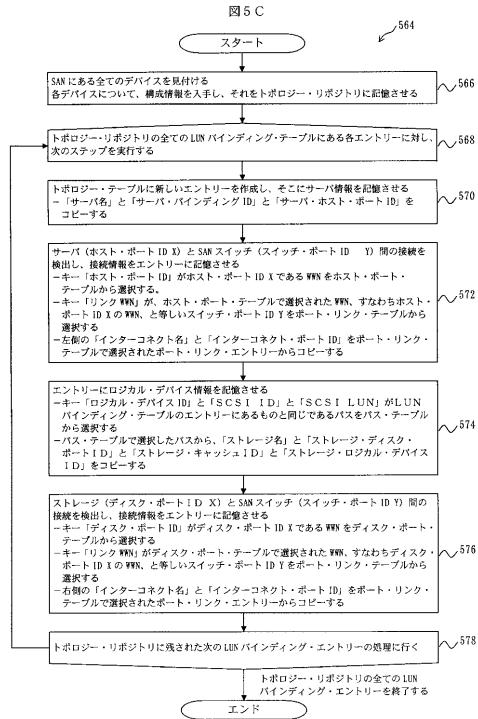


【図5B】

図5B

| サーバ名 | バインディングID | ホスト・ポートID | インターフェクト | | | ストレージ | | | | | |
|----------|-----------|-----------|----------|-------|-------|--------|-----|-------|------------|---------|-------------|
| | | | 名 | 名 | ポートID | 名 | 名 | ポートID | ディスク・ポートID | キャッシュID | ロジカル・デバイスID |
| サーバA LU1 | a 1 | スイッチC | s 1 | スイッチC | s 3 | ストレージD | d 1 | c 1 | v 1 | | |
| サーバA LU2 | a 1 | スイッチC | s 1 | スイッチC | s 4 | ストレージD | d 2 | c 1 | v 2 | | |
| サーバB LU1 | b 1 | スイッチC | s 2 | スイッチC | s 4 | ストレージD | d 2 | c 2 | v 2 | | |

【図 5 C】



【図 5 D】

図 5 D

| ディスクアリ ID | デバイス・タイプ | デバイス情報 | ID アドレス |
|-----------|-----------------|---------------|-----------------|
| 1 | サーバ | ベンダーA, サーバA | 100.100.100.100 |
| 2 | サーバ | ベンダーB, サーバB | 100.100.100.101 |
| 3 | SANスイッチ | ベンダーC, スイッチC | 100.100.100.102 |
| 4 | ストレージ サブシステム | ベンダーD, ストレージD | 100.100.100.103 |

【図 6 B】

図 6 B

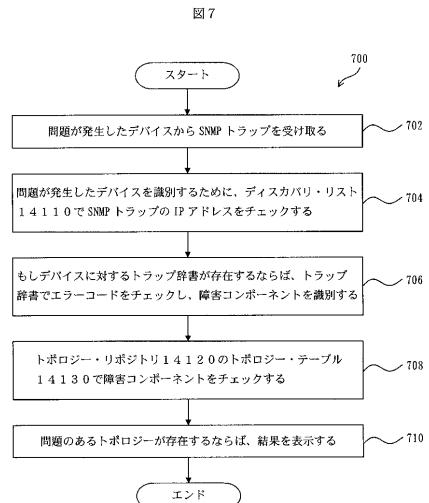
| エラーコード | エラー・コンポーネント | ID |
|---------|-------------|-------|
| 1 0 0 0 | ディスク・ドライブ | d d 1 |
| ... | ... | ... |
| 2 0 0 1 | キャッシュ | c 2 |
| ... | ... | ... |
| 3 0 0 0 | ディスク・ポート | p 1 |
| ... | ... | ... |

【図 6 A】

図 6 A

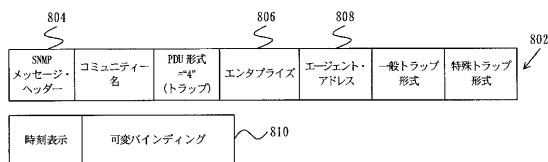
| エラーコード | エラー・コンポーネント |
|--------|-------------|
| A 1 | ポート 1 |
| ... | ... |

【図 7】



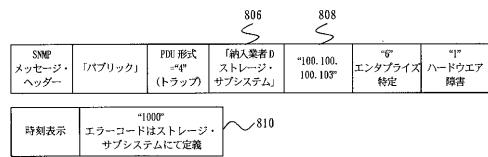
【図 8 A】

図 8 A



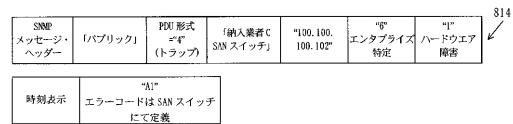
【図 8 B】

図 8 B



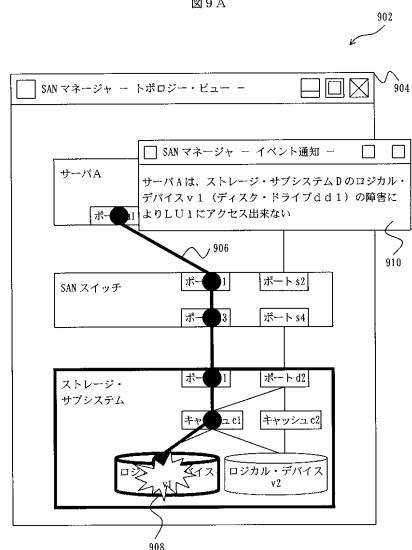
【図 8 C】

図 8 C



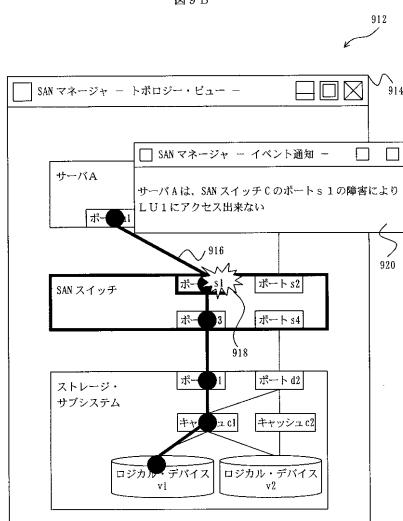
【図 9 A】

図 9 A



【図 9 B】

図 9 B



【外国語明細書】

EVENT NOTIFICATION IN STORAGE NETWORKS

BACKGROUND OF THE INVENTION

[01] The present invention relates to storage networks, more particularly to event
5 notification methods and systems in a storage network.

[02] Data is the underlying resources on which all computing processes are based. With
the recent explosive growth of the Internet and e-business, the demand on data storage
systems has increased tremendously. Generally, storage networking encompasses two
10 applications or configurations: network-attached storage (NAS) or storage area network
(SAN). A NAS uses IP over Ethernet to transports data in file formats between storage
servers and their clients. In NAS, an integrated storage system, such as a disk array or tape
device, connects directly to a messaging network through a local area network (LAN)
interface, such as Ethernet, using messaging communications protocols like TCP/IP. The
15 storage system functions as a server in a client-server system.

[03] Generally, a SAN is a dedicated high performance network to move data between
heterogeneous servers and storage resources. Unlike NAS, a separate dedicated network is
provided to avoid any traffic conflicts between client and servers on the traditional messaging
20 network. A SAN permits establishment of direct connections between storage resources and
processors or servers. A SAN can be shared between servers or dedicated to a particular
server. It can be concentrated in a single locality or extended over geographical distances.
SAN interfaces can be various different protocols, such as Fibre Channel (FC), Enterprise
Systems Connection (ESCON), Small Computer Systems Interface (SCSI), Serial Storage
25 Architecture (SSA), High Performance Parallel Interface (HIPPI), or other protocols as they
emerge in the future. For example, the Internet Engineering Task Force (IETF) is developing
a new protocol or standard iSCSI that would enable block storage over TCP/IP, while some
companies are working to offload the iSCSI-TCP/IP protocol stack from the host processor to
make iSCSI a dominant standard for SANs.

30 [04] Currently, Fibre Channel (FC) is the dominant standard or protocol for SANs. FC is
the performance leader today at 1 Gbps and 2 Gbps link speeds and offers excellent (very

low) latency characteristics due to a fully offloaded protocol stack. Accordingly, Fibre Channel-based SANs are often used in high-performance applications. FC at 2 Gbps is expected to remain unchallenged in the data center for the foreseeable.

- 5 [05] In order to properly utilize the high-performance and versatile SANs, they need to be managed efficiently. One important management function in storage networks is the event notification management. Event notification management in a SAN can be challenging since it generally includes different hardware and operating systems from various vendors with different proprietary messaging languages or rules.

10

BRIEF SUMMARY OF THE INVENTION

- [06] Embodiments of the present invention relates to event notification and event management within a storage network such as a storage area network (SAN). In one embodiment, a network manager, e.g., a SAN manager, collects information from devices within the storage network. The network manager includes a Trap dictionary for each device within the network. The dictionary is used to interpret event messages received from the devices experiencing failure or is about to experience failure. The network manager is configured to identify a specific component within a device with the problem and determine an effect of the event. The network manager is configured to display an event notification on a centralized management console providing the cause and effect of the event.

- [07] In one embodiment, a heterogeneous network includes network related hardware and software products from a plurality of vendors. The network includes a storage system configured to store data, a server configured to process requests, a switch coupling the storage system and the server for data communication, and a network manager including an event dictionary to interpret an event message received from a device experiencing failure.

- [08] In another embodiment, a storage area network (SAN) includes a network manager including an event dictionary to interpret an event message received from a device experiencing failure, the device being provided within the SAN.

- [09] In another embodiment, a management server configured to manage a storage area network (SAN) includes a network manager including an event dictionary to interpret an

event message received from a device experiencing failure, the device being provided within the SAN.

[10] In another embodiment, a storage area network (SAN) includes a plurality of application servers configured to handle data requests. A management server is configured to handle management functions of the SAN and includes a SAN manager. The SAN manager includes a Trap dictionary to interpret an error code included in a Trap message from a device experiencing failure. The device has a plurality of components, where one of the plurality of components is experiencing problem. A plurality of storage subsystems are configured to store data. A plurality of switches are configured to transfer data between the application servers and the storage subsystems. The SAN is a heterogeneous network including network products from a plurality of vendors with different rules for error codes.

[11] Yet in another embodiment, a method of managing a storage network includes providing a plurality of network products manufactured from a plurality of vendors. An event message is received from a device including a plurality of components, wherein one of the components is experiencing failure. The event message includes an error code identifying the one component experiencing the failure. An event dictionary is accessed to interpret the error code in the event message. The event dictionary includes an error code list and a corresponding error component list. An identity of the component experiencing the failure is determined using the error code list in the event dictionary.

BRIEF DESCRIPTION OF THE DRAWINGS

[12] Fig. 1 illustrates a schematic diagram of a network including a storage network coupled to a messaging network according to one embodiment of the present invention.

[13] Fig. 2A illustrates a schematic diagram of a storage area network including a management server and a SAN manager according to one embodiment of the present invention.

30

[14] Fig. 2B illustrates a storage subsystem of a SAN according to one embodiment of the present invention.

[15] Fig. 2C illustrates a disk port table provided in a management agent of the storage subsystem of Fig. 2B according to one embodiment of the present invention.

5 [16] Fig. 2D illustrates a device table provided in a management agent of the storage subsystem of Fig. 2B according to one embodiment of the present invention.

[17] Fig. 2E illustrates a path table provided in a management agent of the storage subsystem of Fig. 2B according to one embodiment of the present invention.

10 [18] Fig. 3A illustrates a schematic diagram of a SAN switch of a SAN according to one embodiment of the present invention.

[19] Fig. 3B illustrates a port link table provided in a management agent of a SAN switch of Fig. 3A according to one embodiment of the present invention.

15 [20] Fig. 4A illustrates a schematic diagram of application servers of a SAN according to one embodiment of the present invention.

20 [21] Figs. 4B and 4C illustrate schematic diagrams of host port tables of an application server according to one embodiment of the present invention.

[22] Figs. 4D and 4E illustrate schematic diagrams of a LUN binding tables of an application server according to one embodiment of the present invention.

25 [23] Fig. 5A illustrates a schematic diagram of a management server of a SAN according to one embodiment of the present invention.

[24] Fig. 5B illustrates a topology table of a SAN manager according to one embodiment of the present invention.

30 [25] Fig. 5C illustrates a process of generating the topology table of Fig. 5B according to one embodiment of the present invention.

[26] Fig. 5D illustrates a discovery list of a SAN manager according to one embodiment of the present invention.

[27] Figs. 6A and 6B illustrate Trap dictionaries for a storage subsystem and SAN switch
5 of a SAN manager according to one embodiment of the present invention.

[28] Fig. 7 is a flow diagram illustrating an event notification method according to one embodiment of the present invention.

10 [29] Figs. 8A-8C illustrate Trap messages including error codes according to one embodiment of the present invention.

[30] Figs. 9A and 9B illustrate schematic event notifications provided to a network administrator by a SAN manager according to one embodiment of the present invention.

15

DETAILED DESCRIPTION OF THE INVENTION

[31] The present invention relates to event notification management in a storage network, such as a storage area network (SAN), network attached network (NAS), or the like. In
20 particular, the present invention relates to event notification management in a storage network using heterogeneous hardware and/or software systems. Heterogeneous systems have hardware or software products, or both from multiple vendors. Specific embodiments of the present invention are described below using SANs for convenience of explanation and should not be used to narrow the scope of the present invention.

25

[32] As used herein, the term “SAN” or “sub-network” refers to a centrally managed, high-speed storage network that is coupled to a messaging network and includes multi-vendor storage devices, multi-vendor storage management software, multi-vendor servers, multi-vendor switches, or other multi-vendor network related hardware and software products. The
30 extent of the heterogeneous nature of the SAN or sub-network varies. Some SANs or sub-networks have multi-vendor products for all of the above network devices and components, while others have multi-vendor products for a portion of the above device and components.

[33] As used herein, the term "storage network" refers to a network coupled to one or more storage systems and includes multi-vendor storage devices, multi-vendor storage management software, multi-vendor application servers, multi-vendor switches, or other multi-vendor network related hardware and software products. The "storage network" 5 generally is coupled to another network, e.g., a messaging network, and provides decoupling of the back-end storage functions from the front-end server applications. Accordingly, the storage network includes the SAN, NAS, and the like.

[34] Fig. 1 schematically illustrates a network system 100 including one or more 10 messaging networks 102 and a SAN 104 connecting a plurality of servers 106 to a plurality of storage systems 108. The network 102 may be a local area network, a wide area network, the Internet, or the like. The network 102 enables, if desired, the storage devices 108 to be centralized and the servers 106 to be clustered for easier and less expensive administration.

15 [35] The SAN 104 supports direct, high-speed data transfers between servers 106 and storage devices 108 in various ways. Data may be transferred between the servers and storage devices. A particular storage device may be accessed serially or concurrently by a plurality of servers. Data may be transferred between servers. Alternatively, data may be transferred between storage devices, which enables data to be transferred without server 20 intervention, thereby freeing server for other activities. For example, a storage system may back up its data to another storage system at predetermined intervals without server intervention.

25 [36] Accordingly, the storage devices or subsystems 108 is not dedicated to a particular server bus but is attached directly to the SAN 104. The storage subsystems 108 are externalized and functionally distributed across the entire organization.

30 [37] In one embodiment, the SAN 104 is constructed from storage interfaces and is coupled to the network 102 via the servers 106. Accordingly, the SAN may be referred to as the network behind the server or sub-network.

[38] In another embodiment, a SAN is defined as including one or more servers, one or more SAN switches or fabrics, and one or more storage systems. In yet another embodiment, a SAN is defined as including one or more servers, one or more SAN switches or fabrics, and

ports of one or more storage systems. Accordingly, the term SAN may be used to referred to various different network configurations as long as the definition provide above is satisfied.

[39] Fig. 2A illustrates a SAN system 200 including a storage system (or subsystem) 202, 5 a SAN switch or fabric 204, a plurality of servers 206a and 206b, a management server 208, and a management network 210. The management server 208 includes a SAN manger 209 that manages the SAN, as explained in more detail later. Although a single storage subsystem is illustrated in the SAN system 200, a plurality of storage subsystems are provided in other embodiments. Similarly, in other embodiments, the number of other 10 network components may be different from the illustrated example.

[40] The storage subsystem 202 includes a management agent 212, a plurality of disk ports 214a and 214b, a plurality of logical devices 216a and 216b, and a plurality of caches 218a and 218b. The disk ports 214a and 214b are also referred to as the disk ports d1 and d2. The 15 logical devices 216a and 216b are also referred to as the logical devices v1 and v2. The management agent 202 manages the configuration of the storage subsystem and communicates with the management server 208. For example, the agent 212 provides the management server 208 with the data I/O path, the connection information of the disk ports d1 and d2, and any failure experienced by the components in the storage subsystem 202, as 20 described in more detail below. The disk ports 214a and 214b are connection ports to the SAN switch 204 to transfer and receive data to and from the servers 206a and 206b. The connection protocol used for the present embodiment is Fibre Channel but other protocols may be used, e.g., SCSI, FC over IP, or iSCSI.

25 [41] As well known by a person skilled in the art, the management agent 212 includes a disk port table 220, a device table 222, and a path table 224 (Fig. 2B). These tables are updated periodically as configuration information changes. The disk port table 220 includes a disk port ID 226 that provides information about the disk ports in the storage subsystems, such as a “nickname,” and a world wide name (WWN) 228 that provides unique identifier for 30 each disk port (Fig. 2C). The nickname refers to a storage subsystem specific identification name, for example, “d1” that refers to the disk port 214a. The name “d1” is sufficient to identify the disk port within the storage subsystem in question but is insufficient when there is a plurality of storage subsystems since disk ports in other storage subsystems may have been assigned that same name. On the other hand, the unique identifier (referred to as the

world wide name in Fibre Channel) is unique identification information assigned to a particular component.

[42] The device table 222 includes a logical device ID 230 that provides information on the relationship between logical devices and disk drives within the storage subsystems and a disk drive list 232 (Fig. 2D). The path table 224 includes a path ID 234 that provides the nickname for the path, a disk port ID 236 that identifies the disk port attached to the path, a cache ID 238 that identifies the cache attached to the path, a logical device ID 240 that provides the nickname of the logical device attached to the path, a SCSI ID 242 that identifies the SCSI attached to the path, and a SCSI LUN 244 that provides information about the SCSI LUN attached to the path (Fig. 2E).

[43] The logical devices 216a and 216b are volumes that are exported to the servers. The logical device may consist of a single physical disk drive or a plurality of physical disk drives in a redundant array of independent disks (RAID). A RAID storage system permits increased availability of data and also increase input/output (I/O) performance. In a RAID system, a plurality of physical disk drives are configured as one logical disk drive, and the I/O requests to the logical disk drive are distributed within the storage system to the physical disk drives and processed in parallel. RAID technology provides many benefits. For example, a RAID storage system can accommodate a very large file system, so that a large file can be stored in a single file system, rather than dividing it into several smaller file systems. Additionally, RAID technology can provide increased I/O performance because data on different physical disk can be accessed in parallel. In one embodiment, each logical device includes four physical disk drives dd1, dd2, dd3, and dd4, as illustrated in Fig. 2B.

25

[44] The caches 218a and 218b are data caches associated with the logical devices 216a and 216b. They are provided to expedite data processing speed. In other embodiments, the storage subsystem does not include any cache.

30

[45] Referring to Fig. 3A, the SAN switch 204 connects the servers and storage subsystems. The switch 204 provides data connection between the servers and storage subsystems. In one embodiment, the switch may be coupled to a bridge, router, or other network hardware to enlarge the network coverage. The switch 204 includes a switch management agent 302 that manages the configuration of the switch and a plurality of switch

ports 304a, 304b, 304c, and 304d. These switch ports also are referred to as s1, s2, s3, and s4, respectively, as indicated by Fig. 3A. The switch management agent 302 assists the management server 208 in managing the SAN by providing the server 208 with the connection information of the switch ports and notifying the server 208 if failure occurs in
5 any component within the switch 204. The management agent 302 includes a port link table 306 that provides information on the interconnect relationship between servers and storage subsystems via switches (also referred to as "link"). The port link table 306 includes a switch port ID 308 that provides identification information or nickname for each switch port, a switch port world wide name (WWN) 310 that provides a unique identifier of each switch
10 port, and a link WWN 312 that provides a unique identifier of the target device that is connected to each switch port (Fig. 3B).

[46] Fig. 4A illustrates the servers 206a and 206b for application use in more detail. In the present embodiment, separate servers are used to perform the application and management
15 functions. Each server 206 includes a server management agent 402 that manages the configuration of the server and a server port 404 for data connection. The server management agent 402 assists the management server 208 in managing the SAN by providing the server 208 with the connection information of the server ports and notifying the server 208 if failure occurs in any component within the server 206. The agent 402 is
20 generally provided within the server for convenience. Also the agent 402 includes a host port table 406 and a LUN binding table 408. The host port table provides the information on the host or server ports in a server.

[47] Referring to Fig. 4B, the host port table 406a, provided in the agent 402a, includes a plurality of columns for storing information on the ports in the server. The table 406a includes a host port ID 410a that provides a device specific identification information or nickname for a particular port within the server, a world wide name 412a that provides a unique port identification information, and a SCSI ID 414a that provides a SCSI
25 identification information assigned to a particular port by an network administrator.
30 Generally, a single SCSI ID is assigned for a server port in the SAN. The worldwide name 412a is a term used in connection with Fibre Channel, so other comparable terms may be used if a different connection protocol is used. Fig. 4C shows the host port table 406b provided in the agent 402b. The host port table 406b includes a host port ID 410b, a world wide name 412b, and a SCSI ID 414b.

[48] Referring to Fig. 4D, the LUN binding table 408a, provided in the agent 402a, provides the information on the data I/O path from the host port to the SCSI Logical Unit, also referred to as "LUN binding" or "binding." The table 408a includes a binding ID 416a that provides a device specific identification information or nickname for the binding, a host port ID 418a, corresponding to the host port ID 410a of the table 406a, that provides a nickname for a particular port, a SCSI ID 420a, corresponding to the SCSI ID 414a of the table 406a, that is attached to the binding, a LUN 422a that provides a SCSI LUN attached to the binding, and an inquiry information 424a that provides the information given by the LUN when servers issue SCSI INQUIRY commands to the LUN. The inquiry information generally includes information such as vendor name, product name, and logical device ID of the LUN. Fig. 4E shows the LUN binding table 408b provided in the agent 402b. The LUN binding table 408b includes a binding ID 416b, a host port ID 418b, a SCSI ID 420b, a LUN 422b, and an inquiry information 424b.

15

[49] Fig. 5A illustrates the management server 208 that is dedicated to the management related functions of the SAN according to one embodiment of the present invention. In another embodiment, a single server may perform the dual functions of the application servers and management servers.

20

[50] The management server 208 includes a SAN manager or network manager 502 that is used to manage the SAN to ensure efficient usage of the network. The manager 502 includes all physical and logical connection information obtained from various components within the SAN. Accordingly, the manager 502 communicates with the management agents, e.g., the switch management agent 302, server management agent 402, and storage system management agent 212, within the SAN to obtain the respective configuration tables via the management network 210. Accordingly, the SAN manager or network manager 502 includes a topology repository 504 and a discovery list 506.

30

[51] The topology repository 504 includes a topology table 508 that provides the topology of the I/O communication in a SAN. The topology table 508 is made by merging the tables, e.g., the host port table, LUN binding table, and the like, obtained from the devices within the SAN. Referring to Fig. 5B, the topology table includes a server section 550 that provides binding ID and host port ID information on the servers in the SAN, an interconnect section

552 that provides information on the switches in the SAN, and an storage section 554 that provides information on the storage subsystems including the disk port ID, cache ID, and logical device ID.

5 [52] Fig. 5C shows a process 564 performed by the SAN manager 502 to generate the topology table 504 according to one embodiment of the present invention. All the devices provided in the SAN are detected (step 566). Configuration information of each detected device is retrieved and stored in the topology repository. Each LUN binding entry is retrieved until all entries are retrieved (step 568). For each entry, a new entry in the topology table is made and server information is stored therein, e.g., server name, server binding ID, and server host port ID (step 570). A connection between a server (host port ID X) and a SAN switch (switch port ID Y) is detected, and the connection information is stored in the entry (step 572). This step involves selecting a WWN from a host port table where the key "host port ID" is host port ID X, selecting a switch port ID Y from a port link table where the key "link WWN" is equal to a selected WWN, i.e., WWN of host port ID X, in a host port table, and copying "interconnect name" and "interconnect port ID" from a selected port link entry in a port link table.

10

15

[53] Thereafter, the logical device information is stored in the entry (step 574). This step involves selecting a path from a path table where the keys "logical device ID," "SCSI ID," and "SCSI LUN" are equal to those in an entry in an LUN binding table, and copying "storage name," "storage disk port ID," "storage cache ID," and "storage logical device ID" from a selected path in a path table.

20

25 [54] Next, a connection between a storage (disk port ID X) and a SAN switch (switch port ID Y) is detected and the connection information is stored in the entry (step 576). This step involves selecting a WWN from a disk port table where the key "disk port ID" is disk port ID X, selecting a switch port ID Y from a port link table where the key "link WWN" is equal to a selected WWN, i.e., WWN of disk port ID X, in a disk port table, and copying "interconnect name" and "interconnect port ID" on the right from a selected port link entry in a port link table. After the step 576, the next LUN binding entry is retrieved (step 578), and the above steps are repeated until all entries have been processed.

30

[55] The discovery list 506 includes the information on all the devices in a SAN. The SAN manager 502 uses information from this list to retrieve the configuration information from the management agents in the SAN devices. Referring to Fig. 5D, the discovery list includes a discovery ID section 556 that provides a nickname of the target SAN device to be discovered, a device type section 558 that identifies the device type of the target SAN device, a device information section 560 that provides vendor information or other detailed information about the target SAN device, an IP address section 562 that provides the IP address of the target SAN device to facilitate communication between the SAN manager and the target SAN device. In the present embodiment, the communication protocol used is
5 TCP/IP.
10

[56] The SAN manager 502 is configured to perform the event management using one or more Trap dictionaries (to be described below) as well the topology table and the discovery list described above. The manager 502 receives an event message from a component in the
15 SAN that is experiencing problem. The event is then notified to a network administrator, so that an appropriate action may be taken. One common protocol used for event notification is Simple Network Management Protocol (SNMP), an IP-based protocol. In the present embodiment, the manager 502 is configured to handle the SNMP messages.

20 [57] In operation, a device that is experiencing problem issues an SNMP Trap message to the manager 502. The manager 502, upon receipt of the message, can determine the cause of the problem and also the consequent effects of the event or problem in the SAN. For example, if failure occurs at the switch port 304a of the SAN switch 204, the manager 502 can determine that an event message has been received because of the switch port 304a's
25 failure and that this failure affects the server 206a from accessing the logical device v1. Such a precise diagnosis of the cause and effect of an event has not been possible in the conventional SAN managers because a SAN includes hardware and software from many different vendors with different messaging rules. Accordingly, the conventional SAN managers, in a similar situation, can merely inform the network administrators that the SAN
30 switch 204 is experiencing problem and little else.

[58] In order to provide such a precise diagnosis of cause and effect of the event, the manager 502 includes one or more Trap dictionaries (also referred to as "event dictionaries" or "look-up tables") to decipher or interpret the Trap messages received by the manager 502.

In one embodiment, the manager 502 includes a plurality of Trap dictionaries for various hardware and software vendors. The Trap dictionaries may be stored in a number of different ways. The Trap dictionaries may be stored according to the device type, so that all the Trap dictionaries relating to SAN switches are stored under a single location. Alternatively, the 5 Trap dictionaries may be stored according to a vendor specific file.

[59] In the present embodiment, the Trap dictionaries are stored according to the device type. Accordingly, the manager 502 includes a Trap dictionary 510 for SAN switches and a Trap dictionary 512 for storage subsystems. The switch Trap dictionary 510 includes an 10 error code 602 that may be attached to a Trap message to notify occurrence of a particular event and an error component 604 that identifies a component that is experiencing problem (Fig. 6A). For example, if problem occurs with a port s1 in the SAN switch 204, a Trap message including an error code "A1" is sent to the manager 502. The manager 502 can determine the meaning of the error code by looking up the switch Trap dictionary 510.

15 [60] Similarly, the storage Trap dictionary 512 includes an error code 606 that may be attached to a Trap message to notify occurrence of a particular event, an error component 608 that identifies a component that is experiencing problem, and an ID 610 that provides the component ID information. In one embodiment, the management server includes a dictionary 20 server 512 that is used to look-up the appropriate Trap dictionaries upon receipt of a Trap message.

[61] Fig. 7 is a flow chart 700 illustrating handling of an event notification in the SAN using the SAN manager 502 according to one embodiment of the present invention. The 25 manager 502 receives a SNMP Trap message from a device experiencing failure (step 702). The device includes a plurality of components, of which one of them is experiencing failure. The Trap message includes an appropriate error code to identify the exact component with the problem. The manager 502 checks the IP address of the SNMP Trap using the discovery list to identify the device in question (step 704). If the Trap dictionary for the device exists, 30 the error code in the message is looked up and the specific component within the device that is having problem is identified (step 706). The component experiencing failure is looked up using the topology table in the topology repository (step 708). If the topology problem exists, the problem is identified to the user (step 710). Fig. 8A illustrates an exemplary SNMP Trap message 802 according to one embodiment of the present invention. The Trap message

includes a header 804, an enterprise section 806 to identify a vendor of the device in question, an agent section 808 to provide an IP address of the device in question, and a variable binding 810 for an error code associated with a particular event.

[63] Fig. 8B illustrates a Trap message 812 transmitted to the manager 502 in response to failure of a disk drive in the storage subsystem 202. The message 812 indicates in the enterprise 806 that the device experiencing failure is a storage subsystem manufactured by vendor D, the agent address 808 indicates that the IP address of the device is 100.100.100.103, and the variable binding section 810 indicates the component experiencing the problem in the device is the disk drive dd1. The manager 502 examines the topology table and determines that the failure in the disk drive dd1 has caused the failure of the logical device v1. The manager 502 also determines that the server 206a cannot access the logical device v1 as a result of this failure. The manager 502 sends an event notification to the network administrator providing information about the failure of disk drive dd1 and the server 206a's inability to access the logical device v1. This event notification may be in the form of text or graphic illustration, or a combination thereof.

[64] Fig. 9A illustrates an event notification 902 provided to a network administrator to inform him or her of the occurrence of the event described above according to one embodiment of the present invention. The event notification includes a topology view 904 providing a graphic illustration of the SAN topology, a data path 906 affected by the event, a component 908 experiencing the failure, and an event summary 910 detailing the component that has failed and the effects of that failure.

[65] Fig. 8C illustrates a Trap message 814 transmitted to the manager 502 in response to the failure of a port in the SAN Switch. The message 814 indicates in the enterprise 806 that the device experiencing failure is a SAN switch manufactured by vendor C, the agent address 808 indicates that the IP address of the device is 100.100.100.102, and the variable binding section 810 indicates the component experiencing the problem in the device is a switch port s1. The manager 502 uses the topology table to determine that the failure in the switch port s1 is preventing the server 206a from accessing the logical device v1. The manager 502 sends an event notification to the network administrator providing information about the failure of the switch port s1 and the resulting effect of the server 206a's failure to access the logical device v1.

[66] Fig. 9B illustrates an event notification 912 displayed to the network administrator to notify the event described above according to one embodiment of the present invention. The event notification includes a topology view 914 providing a graphic illustration of the SAN topology, a data path 916 affected by the event, a component 918 experiencing the failure, 5 and an event summary 920 detailing the component that has failed and the effects of that failure.

[67] The above detailed descriptions are provided to illustrate specific embodiments of the present invention and are not intended to be limiting. Numerous modifications and variations 10 within the scope of the present invention are possible. Accordingly, the present invention is defined by the appended claims.

WHAT IS CLAIMED IS:

1 1. A heterogeneous network including hardware and software products
2 from a plurality of vendors, the network comprising:
3 a storage system configured to store data;
4 a server configured to process requests;
5 a switch coupling the storage system and the server for data communication;
6 and
7 a network manager including an event dictionary to interpret an event message
8 received from a device experiencing failure.

1 2. The network of claim 1, wherein the network is a storage area network.

1 3. The network of claim 1, further comprising:
2 a messaging network coupled to the heterogeneous network.

1 4. The network of claim 1, wherein the event dictionary includes an error
2 code list and an error component list corresponding to the error code list.

1 5. The network of claim 4, wherein the event message includes an error
2 code, wherein the network manager looks up the error code in the event dictionary to identify
3 the component experiencing the failure.

1 6. The network of claim 1, wherein the network manager includes a
2 plurality of event dictionaries for multiple vendors.

1 7. The network of claim 1, further comprising:
2 a topology table including information about a topology of I/O communication
3 in the network, wherein the network manager access the topology table to determine an effect
4 of the failure.

1 8. A storage area network (SAN), comprising:
2 a network manager including an event dictionary to interpret an event
3 message received from a device experiencing failure, the device being provided within the
4 SAN.

1 9. The SAN of claim 8, wherein the SAN includes network products
2 from a plurality of vendors with different event message rules.

1 10. A management server configured to manage a storage area network
2 (SAN) the management server comprising:
3 a network manager including an event dictionary to interpret an event message
4 received from a device experiencing failure, the device being provided within the SAN.

1 11. The management server of claim 10, wherein the SAN uses network
2 products from a plurality of vendors with different event message rules, the plurality of
3 vendors using error codes that are vendor specific.

1 12. A storage area network (SAN), comprising:
2 a plurality of application servers configured to handle data requests;
3 a management server configured to handle management functions of the SAN
4 and including a SAN manager, the SAN manager including a Trap dictionary to interpret an
5 error code included in a Trap message from a device experiencing failure, the device having a
6 plurality of components, where one of the plurality of components is experiencing problem;
7 a plurality of storage subsystems configured to store data; and
8 a plurality of switches configured to transfer data between the application
9 servers and the storage subsystems,
10 wherein the SAN is a heterogeneous network including network products from
11 a plurality of vendors with different rules for error codes.

1 13. The SAN of claim 12, where in the SAN manager includes a plurality
2 of Trap dictionaries for different vendors.

1 14. The network of claim 12, further comprising:
2 a messaging network coupled to the SAN.

1 15. The SAN of claim 12, wherein the Trap dictionary includes an error
2 code list and an error component list corresponding to the error code list.

1 16. The SAN of claim 15, wherein the error code identifies the one
2 component within the device experiencing the problem, wherein the Trap dictionary is looked
3 up to identify the one component.

1 17. The SAN of claim 12, wherein the management server further
2 comprising:
3 a topology table including information about a topology of I/O communication
4 in the network, so that the SAN manager can access the topology table to determine an effect
5 of the failure.

1 18. A method of managing a storage network, comprising:
2 providing a plurality of network products manufactured from a plurality of
3 vendors;
4 receiving an event message from a device including a plurality of components,
5 wherein one of the components is experiencing failure, the event message including an error
6 code identifying the one component experiencing the failure;
7 accessing an event dictionary to interpret the error code in the event message,
8 the event dictionary including an error code list and a corresponding error component list;
9 and
10 determining an identity of the component experiencing the failure using the
11 error code list in the event dictionary.

1 19. The method of claim 18, further comprising:
2 accessing a topology table to determine an effect of the component failure.

1 20. The method of claim 19, further comprising:
2 providing an event notification to a network administrator, the event
3 notification including information about the component experiencing the failure and an effect
4 of the failure.

1 21. The method of claim 20, wherein the event notification includes a
2 topology view that graphically illustrates a topology of the storage network and an event
3 summary that provides information on the component that is experiencing the failure and an
4 effect of that failure.

1 22. The method of claim 21, wherein the event notification further
2 includes a data path effected by the component failure.

1 23. A network manager provided within a management server of a storage
2 area network, the network manager comprising:

3 code for accessing a look-up table to interpret an event message received from
4 a device within the storage area network to identify a component within the device that is
5 experiencing failure and determine an effect of the failure.

1 24. An event notification method performed by a storage area network

2 (SAN) manager running on a first server, wherein the first server is coupled to at least a
3 second server, a switch, and a storage subsystem via a network, the method comprising:

4 receiving from the second server information on an I/O path between the
5 second server and logical volumes that the second server accesses;
6 receiving configuration information from the switch and the storage
7 subsystem;

8 generating topology information on the SAN by using the information from
9 the second server and the configuration information;

10 receiving from the switch or the storage subsystem an event message if a
11 failure occurs at the switch or the storage subsystem; and

12 determining which logical volume the second server cannot access due to the
13 failure using the topology information and the event message.

1 25. The event notification method of claim 24, further comprising;
2 interpreting the event message by using an event dictionary of the manager.

1 26. The event notification method of claim 25, wherein the event
2 dictionary includes an error code list for the switch and the storage subsystem and an error
3 component list corresponding to the error code list.

EVENT NOTIFICATION IN STORAGE NETWORKS

ABSTRACT OF THE DISCLOSURE

[68] A heterogeneous network includes network related hardware and software products from a plurality of vendors. The network includes a storage system configured to store data, a server configured to process requests, a switch coupling the storage system and the server for data communication, and a network manager including an event dictionary to interpret an event message received from a device experiencing failure.

PA 3223217 v2

[図1]

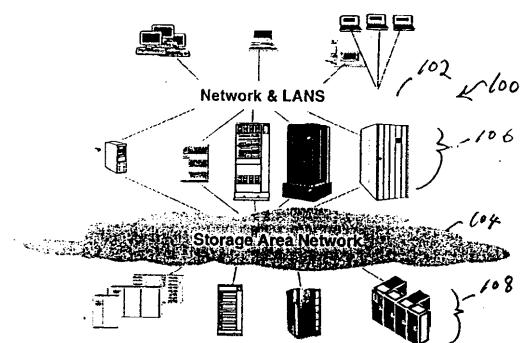


FIG. 1

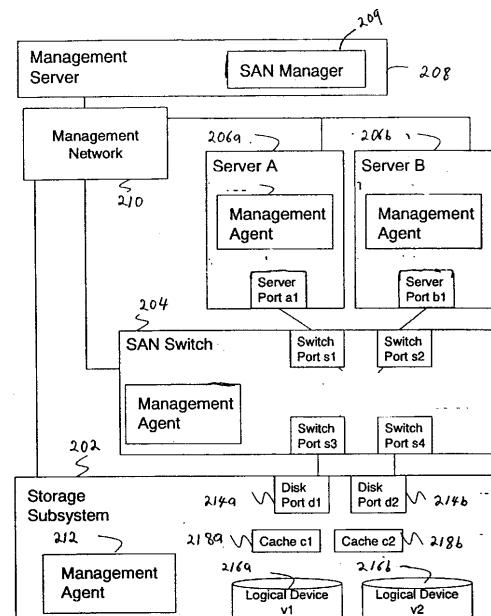


FIG. 2A

[図2A]

[図2B]

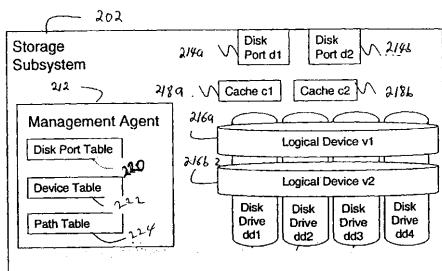


FIG. 2B

[図2C]

| Disk Port ID | WWN |
|--------------|-------|
| d1 | WWNd1 |
| d2 | WWNd2 |

FIG. 2C

[図2D]

| Logical Device ID | Disk Drive List |
|-------------------|--------------------|
| v1 | dd1, dd2, dd3, dd4 |
| v2 | dd1, dd2, dd3, dd4 |

FIG. 2D

[図2E]

| Path ID | Disk Port ID | Cache ID | Logical Device ID | SCSI ID | SCSI LUN |
|---------|--------------|----------|-------------------|---------|----------|
| p1 | d1 | c1 | v1 | 2 | 1 |
| p2 | d1 | c1 | v2 | 3 | 1 |
| p3 | d2 | c2 | v2 | 2 | 2 |

FIG. 2E

[図3A]

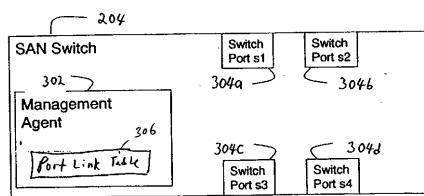


FIG. 3A

[図3B]

| Switch Port ID | Switch Port WWN | Link WWN |
|----------------|-----------------|----------|
| s1 | WWNs1 | WWNa1 |
| s2 | WWNs2 | WWNb1 |
| s3 | WWNs3 | WWNd1 |
| s4 | WWNs4 | WWNd2 |

FIG. 3B

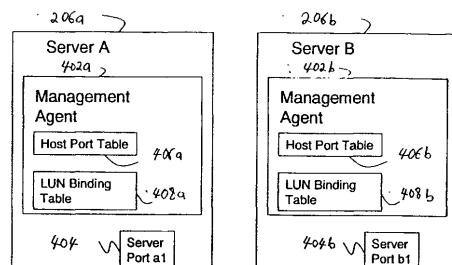


FIG. 4A

[図4B]

| Host Port ID | WWN | SCSI ID |
|--------------|-------|---------|
| a1 | WWNa1 | 2 |

FIG. 4B

[図4A]

| Host Port ID | WWN | SCSI ID |
|--------------|-------|---------|
| b1 | WWNb1 | 3 |

FIG. 4C

[図4D]

| Binding ID | Host Port ID | SCSI ID | LUN | Inquiry Information |
|------------|--------------|---------|-----|--------------------------------------|
| LU1 | a1 | 2 | 1 | VendorD, StorageD, Logical Device v1 |
| LU2 | a1 | 2 | 2 | VendorD, StorageD, Logical Device v2 |

416a 418a 420a 422a 424a

FIG. 4D

[図4E]

| Binding ID | Host Port ID | SCSI ID | LUN | Inquiry Information |
|------------|--------------|---------|-----|--------------------------------------|
| LU1 | b1 | 3 | 1 | VendorD, StorageD, Logical Device v2 |

416b 418b 420b 422b 424b

FIG. 4E

[図5A]

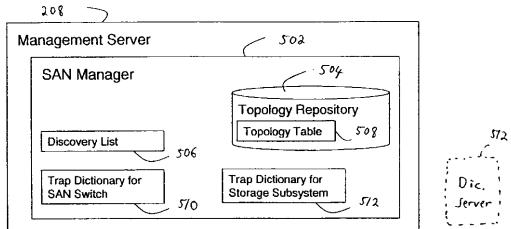


FIG. 5A

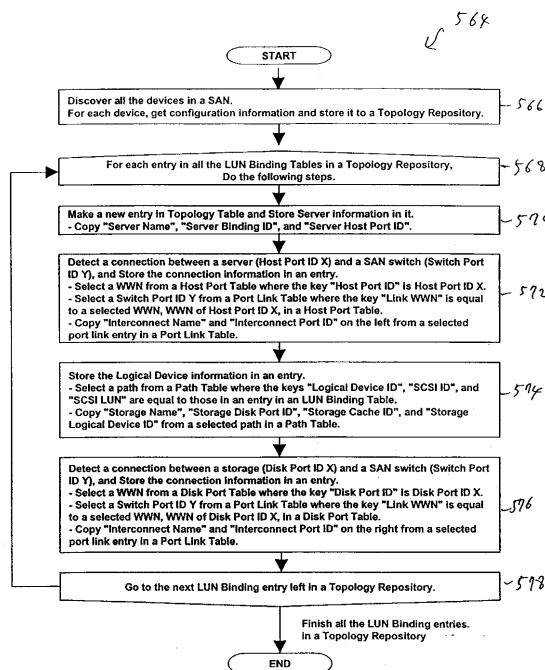


FIG. 5C

[図5D]

[図5B]

| Server | | | Interconnect | | | Storage | | | | |
|---------|------------|--------------|--------------|---------|----------|---------|--------------|--------------|----------|-------------------|
| Name | Binding ID | Host Port ID | Name | Port ID | Name | Port ID | Storage Name | Disk Port ID | Cache ID | Logical Device ID |
| ServerA | LU1 | a1 | Switch C | s1 | Switch C | s3 | Storage D | d1 | c1 | v1 |
| ServerA | LU2 | a1 | Switch C | s1 | Switch C | s4 | Storage D | d2 | c1 | v2 |
| ServerB | LU1 | b1 | Switch C | s2 | Switch C | s4 | Storage D | d2 | c2 | v2 |

FIG. 5B

[図5C]

408b

550 552 554 556

560 562 564

570 572

578 580 582

588 590 592

598 600 602

610 612 614

620 622 624

630 632 634

640 642 644

650 652 654

660 662 664

670 672 674

680 682 684

690 692 694

700 702 704

710 712 714

720 722 724

730 732 734

740 742 744

750 752 754

760 762 764

770 772 774

780 782 784

790 792 794

800 802 804

810 812 814

820 822 824

830 832 834

840 842 844

850 852 854

860 862 864

870 872 874

880 882 884

890 892 894

900 902 904

910 912 914

920 922 924

930 932 934

940 942 944

950 952 954

960 962 964

970 972 974

980 982 984

990 992 994

1000 1002 1004

1010 1012 1014

1020 1022 1024

1030 1032 1034

1040 1042 1044

1050 1052 1054

1060 1062 1064

1070 1072 1074

1080 1082 1084

1090 1092 1094

1100 1102 1104

1110 1112 1114

1120 1122 1124

1130 1132 1134

1140 1142 1144

1150 1152 1154

1160 1162 1164

1170 1172 1174

1180 1182 1184

1190 1192 1194

1200 1202 1204

1210 1212 1214

1220 1222 1224

1230 1232 1234

1240 1242 1244

1250 1252 1254

1260 1262 1264

1270 1272 1274

1280 1282 1284

1290 1292 1294

1300 1302 1304

1310 1312 1314

1320 1322 1324

1330 1332 1334

1340 1342 1344

1350 1352 1354

1360 1362 1364

1370 1372 1374

1380 1382 1384

1390 1392 1394

1400 1402 1404

1410 1412 1414

1420 1422 1424

1430 1432 1434

1440 1442 1444

1450 1452 1454

1460 1462 1464

1470 1472 1474

1480 1482 1484

1490 1492 1494

1500 1502 1504

1510 1512 1514

1520 1522 1524

1530 1532 1534

1540 1542 1544

1550 1552 1554

1560 1562 1564

1570 1572 1574

1580 1582 1584

1590 1592 1594

1600 1602 1604

1610 1612 1614

1620 1622 1624

1630 1632 1634

1640 1642 1644

1650 1652 1654

1660 1662 1664

1670 1672 1674

1680 1682 1684

1690 1692 1694

1700 1702 1704

1710 1712 1714

1720 1722 1724

1730 1732 1734

1740 1742 1744

1750 1752 1754

1760 1762 1764

1770 1772 1774

1780 1782 1784

1790 1792 1794

1800 1802 1804

1810 1812 1814

1820 1822 1824

1830 1832 1834

1840 1842 1844

1850 1852 1854

1860 1862 1864

1870 1872 1874

1880 1882 1884

1890 1892 1894

1900 1902 1904

1910 1912 1914

1920 1922 1924

1930 1932 1934

1940 1942 1944

1950 1952 1954

1960 1962 1964

1970 1972 1974

1980 1982 1984

1990 1992 1994

2000 2002 2004

2010 2012 2014

2020 2022 2024

2030 2032 2034

2040 2042 2044

2050 2052 2054

2060 2062 2064

2070 2072 2074

2080 2082 2084

2090 2092 2094

2100 2102 2104

2110 2112 2114

2120 2122 2124

2130 2132 2134

2140 2142 2144

2150 2152 2154

2160 2162 2164

2170 2172 2174

2180 2182 2184

2190 2192 2194

2200 2202 2204

2210 2212 2214

2220 2222 2224

2230 2232 2234

2240 2242 2244

2250 2252 2254

2260 2262 2264

2270 2272 2274

2280 2282 2284

2290 2292 2294

2300 2302 2304

2310 2312 2314

2320 2322 2324

2330 2332 2334

2340 2342 2344

2350 2352 2354

2360 2362 2364

2370 2372 2374

2380 2382

| Error Code | Error Component | ID |
|------------|-----------------|-------|
| 1000 | Disk Drive | dd1 |
| | | |
| 2001 | Cache | c2 |
| | | |
| 3000 | Disk Port | p1 |
| | | |

S S S
606 608 610

FIG. 6B

[図7]

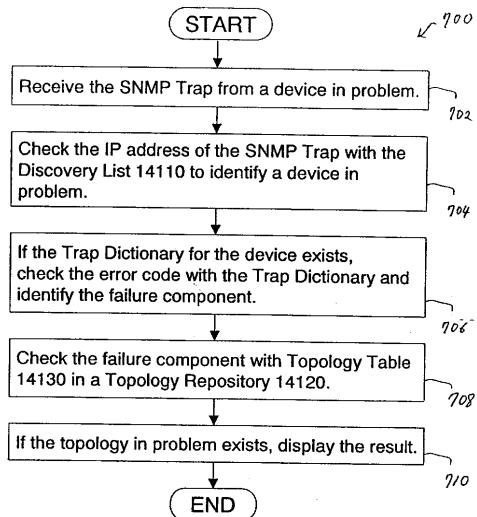


FIG. 7

[図8A]

| SNMP Message Header | Community Name | PDU Type = "4" (TRAP) | Enterprise | Agent Address | Generic Trap Type | Specific Trap Type |
|---------------------|-------------------|-----------------------|------------|---------------|-------------------|--------------------|
| Timestamp | Variable Bindings | | | | | |

FIG. 8A

[図8B]

| SNMP Message Header | "public" | PDU Type = "4" (TRAP) | *Vendor D Storage Subsystem* | *100.100.100.103* | "0" | "1" |
|---------------------|--|-----------------------|------------------------------|-------------------|-----|-----|
| Timestamp | (*100* (Error Code defined in Storage Subsystem)) | | | | | |

FIG. 8B

[図8C]

| SNMP Message Header | "public" | PDU Type = "4" (TRAP) | *Vendor C SAN Switch* | *100.100.100.102* | "0" | "1" |
|---------------------|---|-----------------------|-----------------------|-------------------|-----|-----|
| Timestamp | (*100* (Error Code defined in SAN Switch)) | | | | | |

FIG. 8C

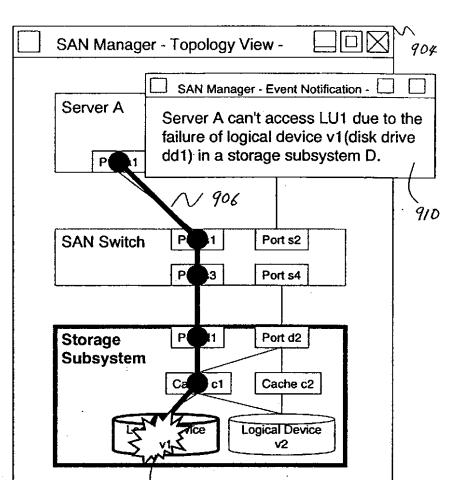


FIG. 9A

[図9A]

[図9B]

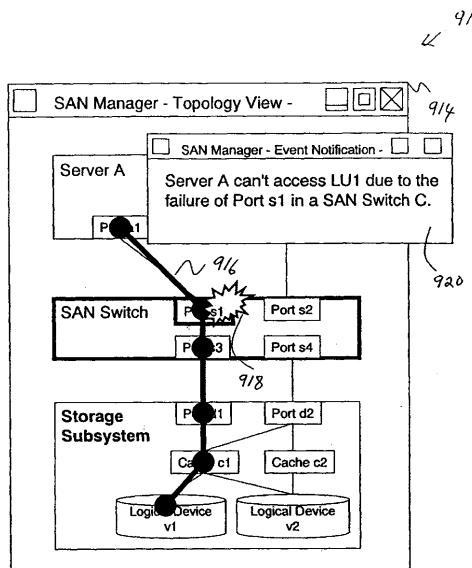


FIG. 9B