

(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2007-527572
(P2007-527572A)

(43) 公表日 平成19年9月27日(2007.9.27)

(51) Int.C1.	F 1	テーマコード (参考)
GO6F 12/00	(2006.01) GO6F 12/00	531M 5B014
GO6F 3/06	(2006.01) GO6F 12/00	545A 5B065
GO6F 13/10	(2006.01) GO6F 12/00	514E 5B082
	GO6F 3/06	304F
	GO6F 13/10	340A

審査請求 未請求 予備審査請求 有 (全 36 頁)

(21) 出願番号	特願2006-534090 (P2006-534090)	(71) 出願人	506107265 セパトン、インコーポレイテッド
(86) (22) 出願日	平成16年9月30日 (2004.9.30)		アメリカ合衆国、マサチューセッツ州 01752, マルボロー、ニッカーソン ロード 400—セカンド フロア
(85) 翻訳文提出日	平成18年5月25日 (2006.5.25)	(74) 代理人	100094318 弁理士 山田 行一
(86) 國際出願番号	PCT/US2004/032122	(74) 代理人	100123995 弁理士 野田 雅一
(87) 國際公開番号	W02005/033945	(72) 発明者	サンドルフィ、ミクロス アメリカ合衆国、マサチューセッツ州 02035, フォックスボロ、マッケンジー レーン 14
(87) 國際公開日	平成17年4月14日 (2005.4.14)		
(31) 優先権主張番号	60/507,329	F ターム (参考)	5B014 EB04 5B065 BA01 EA24 EA33
(32) 優先日	平成15年9月30日 (2003.9.30)		
(33) 優先権主張国	米国(US)		
(31) 優先権主張番号	10/911,987		
(32) 優先日	平成16年8月5日 (2004.8.5)		
(33) 優先権主張国	米国(US)		

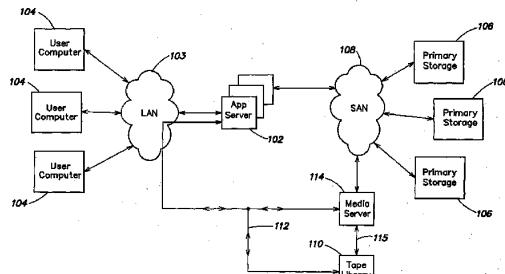
最終頁に続く

(54) 【発明の名称】インスタントボリュームの復旧を支援するエミュレーティッドストレージシステム

(57) 【要約】

バックアップストレージシステムにおいて、バックアップデータセットに対応するデータボリュームをホストコンピュータにマウンティングする装置及び方法であつて、前記方法は、バックアップストレージシステムに保存された最も最近バックアップされたバージョンの一つ以上のデータファイルに対応するデータファイルを一つ以上含むデータボリュームをホストコンピュータにマウンティングする段階、及び前記最も最近バックアップされたバージョンの一つ以上のデータファイルを保存する間、前記バックアップストレージシステムに保存された前記最も最近バックアップされたバージョンの一つ以上のデータファイルより更に最近の第2バージョンの一つ以上のデータファイルに対応するデータを前記バックアップストレージシステムに保存する段階を含む。

【選択図】図2



【特許請求の範囲】**【請求項 1】**

バックアップストレージシステムに保存された最も最近バックアップされた一つ以上のデータファイルに対応するデータファイルを、一つ以上含むデータボリュームをホストコンピュータ上にマウンティングする段階、及び

前記最も最近バックアップされたバージョンの一つ以上のデータファイルを保全する間、前記バックアップストレージシステムに保存された前記最も最近バックアップされたバージョンの一つ以上のデータファイルより更に最近の第2バージョンの一つ以上のデータファイルに対応するデータを前記バックアップストレージシステムに保存する段階を含むことを特徴とする方法。
10

【請求項 2】

前記最も最近バックアップされたバージョンの一つ以上のデータファイルと、前記第2バージョンの一つ以上のデータファイルとをリンクする段階を更に含むことを特徴とする請求項1に記載の方法。

【請求項 3】

前記最も最近バックアップされたバージョンの一つ以上のデータファイルと、前記第2バージョンの一つ以上のデータファイルとを識別するデータ構造を生成する段階を更に含むことを特徴とする請求項1に記載の方法。

【請求項 4】

前記第2バージョンの一つ以上のデータファイルは、前記最も最近バックアップされたバージョンの一つ以上のデータファイルの修正されたバージョンであることを特徴とする請求項3に記載の方法。
20

【請求項 5】

前記データボリュームをマウンティングする段階は、NFSマウンティングとCIFSマウンティングのうちいずれか一つの遂行を含むことを特徴とする請求項1に記載の方法。

【請求項 6】

前記データボリュームをマウンティングする段階は、前記最も最近バックアップされたバージョンの一つ以上のデータファイルに関連したメタデータを含むファイルディスクリプタの生成を含み、前記メタデータは、前記最も最近バックアップされたバージョンの一つ以上のデータファイルのバックアップストレージ媒体上のストレージ位置を識別する識別子を含むことを特徴とする請求項1に記載の方法。
30

【請求項 7】

バックアップデータセットを保存するためのバックアップストレージ媒体、及び、請求項1の方法を具現する命令のセットを実行するよう構成された一つ以上のプロセッサを含む制御器を含むことを特徴とするバックアップストレージシステム。

【請求項 8】

前記バックアップデータセットは、総合フルバックアップデータセットであることを特徴とする請求項7に記載のバックアップストレージシステム。

【請求項 9】

一つ以上のプロセッサで実行時、請求項1の方法を具現する複数の命令でインコーディングされたことを特徴とするコンピュータ判読可能媒体。
40

【請求項 10】

前記プロセッサは、バックアップストレージシステムに含まれることを特徴とする請求項9に記載のコンピュータ判読可能媒体。

【請求項 11】

一つ以上のデータファイルを含むバックアップデータセットに対応するシステムファイルを独自的に識別する第1識別子、及び

バックアップデータセットの最新バージョンの一つ以上のデータファイルそれぞれが保存されたストレージ媒体上の個々のストレージ位置を識別する一つ以上の第2識別子を含むデータ構造が保存されたことを特徴とするコンピュータ判読可能媒体。
50

【発明の詳細な説明】**【発明の分野】****【0001】**

本発明はデータストレージに関するものである。特に既存のフルバックアップ(full back-up)及び後続の増分的バックアップ(incremental back-up)を使って、フルバックアップの同等なことを提供するため、テープストレージシステムをエミュレーティング(emulating)して、エンドユーザーが上記のバックアップからデータを復旧し得るようにする装置及び方法に関するものである。

【関連技術の説明】**【0002】**

ほとんどのコンピューターシステムは、一つ以上のホストコンピューターとこのホストコンピューターによって使われたデータを保存する、一つ以上のデータストレージシステムを含む。このホストコンピューターとストレージシステムは、一般的にファイバチャネルネットワーク、イーサネットネットワーク、又は他の形態の通信ネットワーク等を使って、共にネットワーキング(networking)される。ファイバチャネル(fibre channel)は、チャネル基盤の転送方式の速度とネットワーク基盤の転送方式の柔軟性を組み合わせて、マルティプルイニシエーター(multiple initiator)がネットワークを通じてマルティプルターゲット(multiple target)と通信し得るようにする標準であり、上記のイニシエーターと上記のターゲットは、ネットワークに連結される任意の装置であり得る。ファイバチャネルは、一般的に光ファイバケーブル等の速い電送媒体を使って具現されることによって、大容量データを転送するストレージシステムネットワークに広く選ばれている。

【0003】

図1は様々なホストコンピューターとバックアップストレージシステムを含む一般的にネットワーキングされたコンピューティング環境の一例を示す。一つ以上のアプリケーション・サーバー(application server)102は、近距離通信網(LAN)103を通じて複数のユーザー・コンピュータ104に連結される。アプリケーション・サーバー(application server)102及びユーザー・コンピュータ104は皆“ホストコンピューター”と見なし得る。アプリケーション・サーバー(application server)102は、SAN(storage area network)108を通じて、一つ以上の第1ストレージ装置106に連結される。第1ストレージ装置106は例えば、EMC Corporation、IBM Corporation等で利用され得るディスク・アレイであり得る。代案としてバス(未図示)又はその他ネットワークリンクはアプリケーション・サーバーと第1ストレージシステム106間の相互接続を提供し得る。バス及び/又はファイバチャネルのネットワーク連結はホストコンピューター[例えば、アプリケーション・サーバー102とストレージシステム106の間に電送されたパケットのフォーマットを指示する、SCSI(Small Component System Interconnect)プロトコル等のプロトコルを使って作動し得る。

【0004】

図1に図示されたネットワーキングされたコンピューティング環境は、例えば大型の金融機関又は大企業によって使われ得る大型システムの典型的な例である。ほとんどのネットワーキングされたコンピューティング環境が図1に図示された要素を全て含む必要はない。例えば、小さくネットワーキングされたコンピューティング環境は、ストレージシステムに直接又はLANを通じて連結されるホストコンピューターを簡単に含み得る。又は、図1にはユーザー・コンピュータ104、アプリケーション・サーバー102及び媒体サーバーが別々に図示されているが、これらの機能は一つ以上のコンピューターに結合され得る。

【0005】

第1ストレージ装置106だけでなく、ほとんどのネットワーキングされたコンピューティング環境は、一つ以上の第2又はバックアップストレージシステム110を含む。バックアップ・ストレージシステム110は大容量ではあるが、信頼性がある第2ストレージシステムが使われ得ると言っても、一般的にテープ・ライブラリ(tape library)となり

10

20

30

40

50

得る。一般的に、第2ストレージシステムは第1ストレージ装置より低速であるが、オフサイト(off-site)で保存及び削除が可能な何らかの形態の分離可能な媒体(例えば、テープ、磁気ディスク、又は光ディスク)を含む。

【0006】

図示された例において、アプリケーション・サーバー102は例えば、イーサネット又は他の通信リンク112を通じて、バックアップ・ストレージシステム110と直接、通信し得る。しかし、このような連結は比較的遅く、プロセッサ・タイム又はネットワーク帯域等のリソースを消耗する可能性もある。従って、図示されたようなシステムは例えば、SAN108とバックアップ・ストレージシステム110の間で、ファイバチャネルを使う通信リンクを提供し得る一つ以上の媒体サーバーを含み得る。

10

【0007】

媒体サーバー114はホストコンピューター(ユーザー・コンピュータ104、媒体サーバー114、及び/又はアプリケーション・サーバー102等)、第1ストレージ装置106、及びバックアップストレージシステム110の間でデータの転送を制御する、バックアップ/復旧アプリケーションを含むソフトウェアを実行し得る。バックアップ/復旧アプリケーションの例としてVeritas、Legato社等の製品が挙げられる。データの保護のため、ネットワーキングされたコンピューティング環境内の様々なホストコンピューター及び/又は第1ストレージ装置からのデータは、公知のバックアップ/復旧アプリケーションを使うバックアップストレージシステム110に周期的にバックアップされ得る。

20

【0008】

もちろん、上記したように、ほとんどのネットワーキングされたコンピューティング環境は、図1に図示された例示的なネットワーキングされたコンピューティング環境より小さい、更に少ない構成要素を含み得る。従って、媒体サーバー114は又、実質的に単一のホストコンピューター内のアプリケーション・サーバー102と結合され得、バックアップ/復旧アプリケーションはバックアップ・ストレージシステム110にネットワークを通じて、直接的又は間接的に連結される任意のホストコンピューター上で実行され得ると認識されるべきである。

20

【0009】

典型的なバックアップ・ストレージシステムの一例は多数のテープ・カートリッジ、一つ以上のテープ・ドライブ及びテープ・ドライブへのカートリッジのローディングとアンロードを制御するロボットメカニズムを含むテープ・ライブラリである。バックアップ/復旧アプリケーションは、ロボットメカニズムが特定テープカートリッジ、例えばテープの番号0001の位置を決定して、テープ・ドライブにテープ・カートリッジをローディングすることによって、データがテープ上に記録されるように指示する。また、バックアップ/復旧アプリケーションはデータがテープ上に記録されるフォーマットを制御する。一般的に、バックアップ/復旧アプリケーションはSCSI命令、又はその他標準化された命令を使ってロボットメカニズムに指示し、テープ・ドライブを制御して、テープ上にデータを記録し、テープから記録されたデータを前もって復旧させる。

30

【0010】

従来のテープ・ライブラリのバックアップシステムは、速度、信頼性及び固定された容量を含む、いろいろな問題を有している。ほとんどの大企業は毎週、テラバイトのデータをバックアップする必要がある。しかし、高費用であるにも関わらず、ハイエンドテープ(high-end tape)は一般的に、時間当り約50ギガバイト(GB/hr)に変換する秒当り30~40メガバイト(MB/s)の速度でのみ、データを判読/記録し得る。従って、1又は2テラバイトのデータをテープ・バックアップシステムにバックアップするための連続データの転送時間は少なくとも10~20時間になり得る。

40

【0011】

又は、ほとんどのテープメーカーは、テープが落ちたり(又はロボットメカニズムがテープを運んだり、ローディング動作中に落としたりする可能性があるため、典型的なテープ・ライブラリにおいて比較的頻繁に発生し得る)極度の温度及び湿度等の非理想的

50

な環境条件にテープが露出されたりする場合に、テープで又はテープから、データを保存又は復旧し得るように保障してくれない。従って、調整された環境で保存テープの保存には、相当な注意が必要である。また、複雑な構造のテープ・ライブラリ(ロボットメカニズムを含む)は維持費が高く、それぞれのテープ・カートリッジは比較的高価であり、寿命が制限されている。

【発明の概要】

【0012】

本発明の実施形態は従来のテープ・ライブラリシステムが有する問題点の一部又は全部を軽減させたり克服したりし、従来のテープ・ライブラリシステムに比べ、更に信頼できるバックアップ・ストレージシステムを提供する。

10

【0013】

全体を概括して見ると、本発明の実施形態は、バックアップ/復旧アプリケーションが装置及び媒体を物理的なテープ・ライブラリと同一に見なすよう、従来のテープ・バックアップ・ストレージシステムをエミュレーティングするランダム・アクセス基盤ストレージシステムを提供する。本発明のストレージシステムはソフトウェアとハードウェアを使って、物理的なテープ媒体をエミュレーティングし、一つ以上のランダム・アクセスのディスク・アレイ、トランスレイティング・テープ・フォーマット(translating tape format)、線形、一連のデータをディスクに保存するに適合なデータに代替させる。また、ハードウェア及び/又はソフトウェアで具現されたアプリケーションは、バックアップ・ストレージシステムに保存されたデータを復旧させるために提供される。

20

【0014】

本発明の様々な実施形態によると、一連のテープフォーマットされたデータをランダム・アクセスI/Oに適合したフォーマットに変換するメカニズムが提供される。一実施形態において、NFS(network file system)又は、CIFS(common Internet file system)マウンティッドボリューム(mounted volume)としてのホストコンピューター上のテープフォーマットされたデータの、変換された表現をマウンティングするため、メカニズムが提供される。

20

【0015】

本発明の他の実施形態によると、マウンティングされたファイルシステムに対する記録をセイフストレージ(safe storage)に転換することによって、オリジナルデータが変更されていない状態のままにするためのメカニズムが提供される。一実施形態において、ランダム・アクセスI/Oができるよう、オリジナルデータに対する実時間変化を追跡するためのメカニズムが提供される。他の実施形態において、新しく記録されたデータバック(data back)を一連のテープ・特定I/Oに適合したテープフォーマットされたデータに変換するためのメカニズムが提供される。

30

【0016】

一実施形態においての方法は、バックアップ・ストレージシステムに保存された、最も最近のバックアップされたバージョンの一つ以上のデータファイルに対応する、一つ以上のデータファイルを含むデータボリュームをホストコンピューター上にマウンティングする段階及び、最も最近にバックアップされたバージョンの一つ以上のデータファイルを保存する間、バックアップ・ストレージシステムに保存された、最も最近にバックアップされたバージョンの一つ以上のデータファイルより、更に最近の第2バージョンの一つ以上のデータファイルに対応するデータをバックアップ・ストレージシステムに保存する段階を含む。上記の方法は、最も最近にバックアップされたバージョンの一つ以上のデータファイルと第2バージョンの一つ以上のデータファイルのリンク(linking)を含むこともできる。一例において、上記の方法は、最も最近にバックアップされたバージョンの一つ以上のデータファイルと第2バージョンの一つ以上のデータファイルを同一なものと見なす、データ構造の生成を含むこともできる。他の例において、第2バージョンの一つ以上のデータファイルは、最も最近にバックアップされた、バージョンの一つ以上のデータファイルの修正されたバージョンである可能性がある。

40

50

【0017】

他の実施形態において、バックアップストレージシステムは、バックアップデータセットを保存するためのバックアップストレージ媒体及び、上記の方法を具現する指示のセットを実行するために構成された一つ以上のプロセッサを含む制御器を含む。

【0018】

他の実施形態によると、データ構造が保存されているコンピューター判読可能な媒体が提供され、上記のデータ構造は一つ以上のデータファイルを含むバックアップデータセットに対応するシステムファイルを独自的に識別する第1識別子及び、バックアップデータセットにおいての一つ以上のデータファイル各々の最近のバージョンが保存されたストレージ媒体上の個別のストレージ位置を識別する一つ以上の第2識別子を含む。

10

【0019】

添付図面は一定の比率で図示されていない。図面において、様々な図面に図示された、各々の同一の又は、ほぼ同一の構成要素は、同一の参照番号で示した。明確にするため、全ての図面に図示された、全ての構成要素ごとに参照番号を付与してはいない。

【詳細説明】

【0020】

多様な実施形態を、添付図面を参照して、さらに詳細に説明する。本発明は図面に図示されたり、後述される説明においての構成要素の配置及び、構造の詳細事項に限定されない。本発明は多様な方法と形態で実施され得る。また、ここで使われた表現及び用語は、本発明を限定しようとするのではなく、説明のためのものである。“含む”、“有する”、“構成される”、“成る”などの表現は、後述されるアイテムと同等のものだけでなく、追加的なアイテムを含む意味である。

20

【0021】

この明細書で使われた“ホストコンピューター”という用語は、ストレージシステム又は、他のホストコンピューターと通信できるパーソナルコンピューター、ワークステーション、メインフレーム、ネットワーキングされたクライアント、サーバーなどのよう、一つ以上のプロセッサを有する任意のコンピューターを意味する。ホストコンピューターはユーザー-コンピューター(ユーザー-ワークステーション、PC、メインフレームなどになり得る)だけでなく、媒体サーバー及び、アプリケーションサーバー(図1を参照して、上記したとおりである)を含み得る。また、この明細書内で“ネットワーキングされたコンピューター環境”という用語は、ストレージシステムが各々のホストコンピューターと通信できる方法で、一つ以上の共有したストレージシステムに複数のホストコンピュータが連結された任意のコンピューティング環境を含む。ファイバチャネルは本発明の実施形態に使われ得る通信ネットワークの一例である。しかし、このネットワークはファイバチャネルに限られず、多様なネットワーク構成要素はファイバチャネルの代わりに又はこれに追加的にトーカンリング、イーサネット等の任意のネットワークを通じて又は、他のネットワーク連結の組み合わせを通じて、互いに通信ができると理解されるべきである。この発明の実施形態はSCSI又は、並列SCSIのようなバストポロジーに使われ得もする。

30

【0022】

本発明の多様な実施形態によると、分離可能な媒体基盤ストレージシステムをエミュレーティングするため、一つ以上のディスクアレイを使うことができる仮想分離の可能な媒体ライブラリーバックアップ・ストレージシステムが提供される。本発明の実施形態によると、ユーザーが既存のバックアップ手順を修正または調整したり、新たなバックアップ/復旧アプリケーションを購入する必要なく分離可能な媒体(テープ、磁気ディスク、光ディスク等)にデータをバックアップするのに使われるものと同一のバックアップ/復旧アプリケーションを使い、ディスクアレイにデータがバックアップできる。上記の一実施形態において、テープがエミュレーティングされた分離可能な媒体はテープであり、この発明のバックアップ・ストレージジシステムはテープ及び、従来のテープライブラリーシステムでテープのハンドリングに使われた、ロボットメカニズムを含むテープライブラリーシステムをエミュレーティングする。

40

50

【0023】

本発明の実施形態によるストレージシステムは、ホストコンピューター(バックアップ/復旧アプリケーションを駆動する)とバックアップ・ストレージ媒体と一緒にインターフェースするハードウェアとソフトウェアを含む。ストレージシステムはテープまたは、他の形態の分離可能なストレージ媒体をエミュレーティングし、バックアップ/復旧アプリケーションが装置及び、媒体を物理的テープライブラリーと同一と見做すようになり、線形、一連のテープフォーマットデータをランダムアクセスディスク上に保存するのに適するデータに変換するよう設計することができる。このような方式で、本発明のストレージシステムは新たなバックアップ/復旧アプリケーションソフトウェアまたは、政策を必要とせず、向上した機能(以下で説明するように、ユーザーが個人的にバックアップされたユーザーファイルを検索できるようにする等の機能)を提供できる。

10

【0024】

図2は本発明の実施形態によるバックアップ・ストレージシステム170を含むネットワーキングされたコンピューティング環境の一実施形態のブロック図を表す。図示されたように、ホストコンピューター120はネットワーク連結121を通じてストレージシステム170に繋がる。このネットワーク連結121は例えば、ホストコンピューター120とストレージシステム170間の高速データの転送が可能なファイバチャネル連結などとなり得る。ホストコンピューター120は一つ以上のアプリケーションサーバー102(図1)及び/または、媒体サーバー114(図1)となったり、含むことができ、ネットワーキングされたコンピューティング環境内に存在する任意のコンピューターまたは、第一ストレージシステム110(図1)からデータのバックアップを可能にできるものとして認識されるべきである。また、一つ以上のユーザーコンピューター136はイーサネット連結などの他のネットワーク連結138を通じて、ストレージシステム170に繋がることもできる。後述のとおり、ストレージシステムはユーザーコンピューター136のユーザーがストレージシステムからバックアップされたユーザーファイルを見て、選択的な復旧を可能にすることもできる。

20

【0025】

ストレージシステムは例えば、以下により詳しく説明されたような一つ以上のディスクアレイになり得るバックアップ・ストレージ媒体126を含む。バックアップ・ストレージ媒体126はホストコンピューター120からバックアップしたデータのための実際の保存空間を提供する。しかし、ストレージシステム170はテープライブラリーのような分離可能な媒体ストレージシステムをエミュレーティングし、ホストコンピューター120上にバックアップ/復旧アプリケーションを実行することで、従来の分離可能なストレージ媒体にデータがバックアップされたように見えるようにする、追加的なハードウェア及び、ソフトウェアを含むこともできる。故に、図2に図示されたようにストレージシステム170は例えば、テープのような仮想または、エミュレーティングされた分離可能なストレージ媒体を意味する“エミュレーティングされた媒体”134が含まれる。この“エミュレーティングされた媒体”134はストレージシステム・ソフトウェア及び/または、ハードウェアによりホストコンピューターに提供され、物理的なストレージ媒体としてホストコンピューターに現れる。実際のバックアップ・ストレージ媒体126とエミュレーティングされた媒体134間のインターフェーシングは、以下の詳細な説明のとおりホストコンピューター120からデータを受け取り、バックアップ・ストレージ媒体126にデータを保存するスイッチングネットワーク132及び、ストレージシステム制御器(未図示)となり得る。このような方式で、ストレージシステムは従来のテープ・ストレージシステムをホストコンピューター120にエミュレーティングする。

30

【0026】

一実施形態によると、ストレージシステムはストレージシステム170上のホストコンピューター120からバックアップされたユーザーデータと関連したメタデータ(metadata)を保存する“ラジカルメタデータキャッシュ”242を含み得る。ここで使われた“メタデータ”という用語はユーザーデータに関する情報を表し、実際のユーザーデータの特

40

50

性を記述するデータを意味する。ラジカルメタデータキャッシュ 242 はユーザー及び/または、ソフトウェアアプリケーションがバックアップされたユーザーファイルをランダムに配置し、互いにユーザーファイルを比べ、或いはバックアップされたユーザーファイルにアクセスし、調整できるようにする検索可能なデータの集まりを意味する。ラジカルメタデータキャッシュ 242 内に保存されたデータが使えるソフトウェアアプリケーションの二つの例は、より詳しく後述されるが、エンドユーザー復旧アプリケーション 300 及び、総合フルバックアップ・アプリケーション 240 を含む。

【0027】

要するに、総合フルバックアップ・アプリケーション 240 は既存の一つ以上のフルバックアップ・データセットと一つ以上の増分的バックアップ・データセットから総合フルバックアップ・データセットが生成できる。総合フルバックアップは周期的(例えば、毎週)フルバックアップを遂行する必要がないので、時間とネットワーク・リソースを相当に節約できる。総合フルバックアップ・アプリケーション 240 は後で詳しく説明する。エンドユーザー復旧アプリケーション 300 は、エンドユーザー(例えば、ユーザーコンピューター 136 のオペレータ)がストレージシステム 170 から、前もってバックアップされたユーザーファイルをブラウジング、ロケイティング、ビューアリング及び/又は、復旧できるようにする。これについても後で詳しく説明する。

【0028】

前述のとおり、ストレージシステム 170 はホストコンピューター 120 とバックアップ・ストレージ媒体 126 をインターフェースさせるハードウェア及びソフトウェアを含む。本発明の実施形態によるハードウェア及びソフトウェアは、従来のテープライブラリ・バックアップシステムをエミュレーティングし、ホストコンピューター 120 の観点ではテープ上にデータがバックアップされたように見えるが、実際には複数のディスクアレイのような他のストレージ媒体上にバックアップがされる。

【0029】

図 3 は本発明の実施形態によるストレージシステム 170 の一実施形態を表すブロック図である。一実施形態において、ストレージシステム 170 のハードウェアはストレージシステム制御器 122 及びバックアップ・ストレージ媒体 126 にストレージシステム制御器 122 を連結するスイッチングネットワーク 132 を含む。ストレージシステム制御器 122 はストレージシステム・ソフトウェアの全部または、一部を駆動できるプロセッサ 127 (单一プロセッサまたは、複数のプロセッサとなり得る)及び、メモリー 129 (RAM、ROM、PROM、EEPROM、フラッシュメモリー及び、その組み合わせなど)を含む。メモリー 129 はバックアップ・ストレージ媒体 126 に保存されたデータに関するメタデータを保存するのに使われることもある。本発明の実施形態を実行するプログラミングコードを含むソフトウェアは一般的に RAM、ROM、光ディスク、磁気ディスクまたは、テープなどのコンピューターが記録及び/又は判読できる不揮発性の記録媒体に保存され、以降のプロセッサ 127 によって実行することができるメモリー 129 にコピーされる。このようなプログラミングコードは複数のプログラミング言語、例えば Java、Visual Basic、C、C#、C++、Fortran、Pascal、Eiffel、Basic、COBAL または、その組み合わせのいずれか一つで記録することができ、本発明は特定のプログラミング言語に限定されない。一般的に、動作時にプロセッサ 127 は本発明の実施形態を実行するコードと同じデータが不揮発性の記録媒体から不揮発性の記録媒体よりプロセッサによって、情報に速くアクセスできるようにする RAM のような他の形態のメモリーで判読されるようにする。

【0030】

図 3 に図示したように、制御器 122 は制御器 122 をホストコンピューター 120 及び、スイッチングネットワーク 132 に連結する多数のポートアダプター 124a、124b、124c を含む。図示されたように、ホストコンピューター 120 は例えば、ファイバチャネル・ポートアダプターなどのポートアダプター 124a を通じてストレージシステムに繋がる。ストレージシステム制御器 122 を通じてホストコンピューター 120 はデータをバックアップ・ストレージ媒体 126 にバックアップし、バックアップ・ストレ

10

20

30

40

50

ージ媒体 126 からデータを復旧できる。

【0031】

図示した例において、スイッチングネットワーク 132 は一つ以上のファイバチャネル・スイッチ 128a、128b を含み得る。ストレージシステム制御器 122 はストレージシステム制御器をファイバチャネル・スイッチ 128a、128b に繋ぐ複数のファイバチャネル・ポートアダプター 124b、124c を含む。ファイバチャネル・スイッチ 128a、128b を通じてストレージシステム制御器 122 はデータがバックアップストレージ媒体 126 にバックアップされるようとする。図 3 に図示したように、スイッチングネットワーク 132 はイーサネット・ポートアダプター 125a、125b を通じてストレージシステム制御器 122 に繋がった一つ以上のイーサネットスイッチ 130a、130b をさらに含み得る。一例において、ストレージシステム制御器 122 は例えば、LAN 103 に繋がってストレージシステム 170 が後述のとおり、ホストコンピューター(例えば、ユーザーコンピューター)との通信を可能にする他のイーサネット・ポートアダプター 125c をさらに含む。

【0032】

図 3 に図示した例において、ストレージシステム制御器 122 は二つのファイバチャネル・スイッチと二つのイーサネットスイッチを含む、スイッチングネットワークを通じてバックアップ・ストレージ媒体 126 に繋がる。ストレージシステム 170 内の二つ以上の各々の形態のスイッチの提供は、システム内の全ての単一ポイントの失敗を除去する。即ち、一つのスイッチ(例えば、ファイバチャネル・スイッチ 128a)が失敗しても、ストレージシステム制御器 122 は依然として他のスイッチを通じてバックアップ・ストレージ媒体 126 と通信できる。このような配列は信頼度及び、速度の面でメリットがある。例えば、上記したように、余分の構成要素の提供と単一ポイントの失敗を除去することで信頼度が向上する。また、幾つかの実施形態において、ストレージシステム制御器は並列ファイバチャネル・スイッチの全部または、一部を使ったバックアップストレージ媒体 126 上にデータをバックアップできるので、全体のバックアップ速度が速くなる。しかし、システムは二つ以上の各々の形態のスイッチを含んだり、スイッチングネットワークがファイバチャネル及びイーサネットスイッチを含む必要がない。また、バックアップ・ストレージ媒体 126 が单一ディスクアレイを含む例においてはスイッチが全く必要ない。

【0033】

前述のとおり、一実施形態において、バックアップ・ストレージ媒体 126 は一つ以上のディスクアレイを含み得る。一つの望ましい実施形態において、バックアップ・ストレージ媒体 126 は複数の ATA または、SATA ディスクを含む。このようなディスクは市中で容易に手に入れることのできる製品で、EMC、IBM などの製造社の従来の保存アレイ製品に比べて、比較的低廉である。また、このような分離可能な媒体は(例えば、テープ)の価格とこのような媒体が限定された寿命を有するという事実を念頭に置く場合、このような媒体は価格面で従来のテープ基盤のバックアップ・ストレージシステムに匹敵する。また、このようなディスクはテープに比べて、高速判読/記録が可能である。例えば、單一ファイバチャネル連結を通じてテープのバックアップ速度より確実に速い(例えば、十倍ほど)約 540 GB/hr に換算される、少なくとも 150 MB/s の速度でデータをディスク上にバックアップすることができる。その上、一部のファイバチャネル連結は並列で具現することができるので、一層速くなる。本発明の実施形態によると、バックアップ・ストレージ媒体は、複数の RAID(Redundant Array of Independent Disks) 方式を具現するように構成することができる。例えば、一実施形態においてバックアップ・ストレージ媒体は RAID-5 具現として構成することができる。

【0034】

前述のとおり、本発明による実施形態は、テープカートリッジを物理的なバックアップ・ストレージ媒体として交代するように、ディスクアレイを使った従来のテープライブラリー・バックアップシステムをエミュレーティングすることによって、“仮想テープライ

10

20

30

40

50

ブラー"を提供する。従来のテープライブラリーに提供され物理的なテープカートリッジは"仮想カートリッジ"という用語により取り替えられる。"仮想テープライブラリー"という用語は例えば、一つ以上のディスクアレイとしてソフトウェア及び/又は、物理的なハードウェアで具現することができるエミュレーティングされたテープライブラリーを意味するものとして認識されるべきである。ここでは、主にエミュレーティングされたテープを言及しているが、ストレージシステムはCD-ROM、DVD-ROMなどの他のストレージ媒体をエミュレーティングすることができ、"仮想カートリッジ"という用語は一般的にエミュレーティングされたテープまたは、エミュレーティングされたCDなどのエミュレーティングされたストレージ媒体を意味するものと認識されるべきである。一実施形態において、仮想カートリッジは、実際には一つ以上のハードディスクに対応する。

10

【0035】

故に、一実施形態において、ソフトウェア・インターフェースはテープライブラリーをエミュレーティングするよう提供され、バックアップ/復旧アプリケーションにおいて、データがテープにバックアップされるように見えるようになる。しかし、実際のテープライブラリーはこのディスクアレイ上にデータが実際にバックアップされるようにするが、一つ以上のディスクアレイによって取り替えられる。以下、ストレージシステム170に含まれたソフトウェアの多様な形態、特性及び動作を説明する。

【0036】

ソフトウェアがストレージシステム170に"含まれる"とも説明でき、ストレージシステム制御器122(図3)のプロセッサ127によって実行されるとも説明できるが、ストレージシステム制御器122上で全てのソフトウェアが実行される必要はない。総合フルバックアップ・アプリケーション及び、エンドユーザー復旧アプリケーションなどのソフトウェアプログラムはホストコンピューター及び/又は、ユーザーコンピューターで実行することができ、ストレージシステム制御器、ホストコンピューター及び、ユーザーコンピューターの全部または、一部を経てこの部分が分配され得る。故に、ストレージシステム制御器がコンピューターなどの含まれた物理的エンティティである必要はない。ストレージシステム170は媒体サーバー114または、アプリケーションサーバー102などのホストコンピューター上に存在するソフトウェアと通信できる。また、ストレージシステムは同一または、相違するホストコンピューター上に存在したり、このホストコンピューターで駆動することができ、幾つかのソフトウェア・アプリケーションを含み得る。ストレージシステム170は一部の実施形態において、分離した装置として実施することができても、分離した装置に限られない。一例に、ストレージシステム170は従来のテープライブラリー・バックアップシステム"プラグ・アンド・プレー"の代わりとして作用する独立ユニットとして提供することができる。(即ち、既存のバックアップ手順及び、政策を修正する必要はない。)このようなストレージシステム・ユニットは従来のバックアップシステムを含むネットワーキングされたコンピューティング環境に使われ、余分または、追加的な保存容量を提供することもできる。

20

30

【0037】

前述のとおり、一実施形態によるとホストコンピューター120(例えば、アプリケーションサーバー102又は媒体サーバー114になり得る。図1参照)はこのホストコンピューター120をストレージシステム170に繋ぐネットワーク・リンク(例えば、ファイバチャネル・リンク)121を通じて、バックアップ・ストレージ媒体126上にデータがバックアップできる。主にエミュレーティングされた媒体上にデータをバックアップすることについて後述されるが、この原理はエミュレーティングされた媒体からバックアップデータを復旧するのにも適用されるものと認識されるべきである。ホストコンピューター120とエミュレーティングされた媒体134間のデータの流れは前述のとおり、バックアップ/復旧アプリケーションによって制御できる。バックアップ/復旧アプリケーションの観点では、データが物理的バージョンのエミュレーティングされた媒体上に実際にバックアップされたように見られることもある。

40

【0038】

50

図4に図示したとおり、ストレージシステム・ソフトウェア150はエミュレーティングされた媒体を意味し、ホストコンピューター120上に存在するバックアップ/復旧アプリケーション140とバックアップ・ストレージ媒体126間のインターフェースを提供する一つ以上の論理的抽象層(logical abstraction layer)を含む。ソフトウェア150はバックアップ/復旧アプリケーション140からテープフォーマットデータを受け取って、ランダムアクセスディスク(例えば、ハードディスク、光ディスク等)上に保存するに適合したデータに変換する。一例において、このソフトウェア150はストレージシステム制御器122のプロセッサ127上で実行され、メモリー129(図3)上に保存できる。

【0039】

10

一実施形態によると、上記のソフトウェア150はテープ、テープドライブ及び、テープをテープドライブへ転送したりテープドライブから転送されるのに使われる、ロボットメカニズムのSCSIエミュレーションを提供できる仮想テープライブラリー(VTL)層142を意味する層を含み得る。バックアップ/復旧アプリケーション140は例えば、矢印144で表示されたSCSI命令などを使ってVTL142と通信(例えば、エミュレーティングされた媒体にデータをバックアップまたは、記録)できる。故に、VTLは他のストレージシステム・ソフトウェア及び、ハードウェアとバックアップ/復旧アプリケーション間のソフトウェア・インターフェースが形成できるので、エミュレーティングされたストレージ媒体134をバックアップ/復旧アプリケーションに提供して、エミュレーティングされた媒体が従来の分離可能なバックアップ・ストレージ媒体としてバックアップ/復旧アプリケーションに現れるようとする。

20

【0040】

ファイルシステム層146として言及された、第二ソフトウェア層はエミュレーティングされたストレージ媒体(VTLと表される)と物理的バックアップ・ストレージ媒体126間のインターフェースが提供できる。一例において、ファイルシステム146は小さな運営システムとして動作して、矢印148で表示したSCSI命令などを使って、バックアップ・ストレージ媒体126と通信することで、バックアップ・ストレージ媒体126に、またはそこからデータが判読及び、記録できる。

【0041】

30

一実施形態において、VTLは一般的なテープライブラリー支援を提供し、任意のSCSI媒体 チェンジャー (SCSI media changer)を支援できる。エミュレーティッドテープ装置はIBMLT0-1、LT0-2テープ装置、Quantum SuperDLT320テープ装置、Quantum P3000テープライブラリー・システムまたは、StorageTekL180テープライブラリー・システムなどを含むことができるが、これに限られない。VTL内の各々の仮想カートリッジはデータが保存されるにつれ、動的に増えていくことのできるファイルである。これは固定したサイズを有する、従来のテープカートリッジとは全く違う。一つ以上の仮想カートリッジは、図5を参照して後述するシステムファイルに保存することができる。

【0042】

40

図5は、本発明の実施形態によるシステムファイル200を表したファイルシステム・ソフトウェア146内のデータ構造の一例を表した図面である。この実施形態において、システムファイル200はヘッダー202及び、データ204を含む。ヘッダー202はシステムファイルに保存された各仮想カートリッジを識別する情報を含み得る。ヘッダー202は仮想カートリッジが記録防止の可否、仮想カートリッジの生成/修正日などの情報を含み得る。一例において、ヘッダー202は各仮想カートリッジを独自的に識別し、ストレージシステムに保存された他の仮想カートリッジから各仮想カートリッジを区別する情報を含む。例えば、この情報は仮想カートリッジの名前及び、識別番号(例えば、ロボットメカニズムによってテープが識別できるように、一般的に物理的テープに提供されるバーコードに対応する)を含み得る。ヘッダー202は各仮想カートリッジの容量、最終修正日などの追加的な情報を含み得る。

【0043】

50

本発明は一実施形態によると、ヘッダー 202 のサイズはシステムが追跡可能なデータの独特的なセットの数と保存されたデータの形態(例えば、一つ以上のホストコンピューター・システムからデータバックアップを表す仮想カートリッジ)を表すように極大化できる。例えば、テープ・ストレージシステムに一般的にバックアップしたデータは多数のシステム及び、ユーザーファイルを表す大型データセットによって、一般的に特徴づけられる。データセットが大きいので、これに対して追跡する、非連続データファイルの数は少なくなり得る。故に、一実施形態において、ヘッダー 202 のサイズは効果的に追跡するには多すぎるデータを保存する場合(即ち、ヘッダーが大きすぎること)と十分な数のカートリッジ識別子を保存するスペースが足りない場合(即ち、ヘッダーが小さすぎること)での折衷を通じて、選択できる。例となる一実施形態において、ヘッダー 202 はシステムファイル 200 の最初の 32 MB を活用する。しかし、ヘッダー 202 はシステムの必要及び、このシステムの必要と容量による特徴に基づいた様々なサイズを有し得、ヘッダー 202 のための様々なサイズが選択できるものと認識されるべきである。

【0044】

バックアップ/復旧アプリケーションの観点では、仮想カートリッジは全て同じ属性と特徴を有する物理的テープカートリッジとして現れる。即ち、バックアップ/復旧アプリケーションにおいて、仮想カートリッジは一連の記録されたテープとして現れる。しかし、一つの望ましい実施形態において、仮想カートリッジに保存されたデータはバックアップ・ストレージ媒体 126 上に一連のフォーマットに保存されない。返って、仮想カートリッジ上に記録されたものと現れるデータは、実際にはランダムアクセスの可能なディスクフォーマット・データで、ストレージシステム・ファイル内に保存される。メタデータは保存されたデータを仮想カートリッジにリンクして、バックアップ/復旧アプリケーションがカートリッジフォーマットでデータを判読及び、記録するのに使われる。

【0045】

故に、望ましい一つの実施形態を概括すると、ユーザー及び/又は、システムデータ(“ファイルデータ”を意味する)はホストコンピューター 120 からストレージシステム 170 によって受信され、バックアップ・ストレージ媒体 126 を成すディスクアレイに保存される。後述のとおり、ストレージシステムのソフトウェア 150(図 4)及び/または、ハードウェアはこのファイルデータをシステムファイルの形でバックアップ・ストレージ媒体 126 に記録する。メタデータはストレージシステム制御器によってバックアップされたファイルデータから抽出され、バックアップされたユーザー及び/又は、システムファイルの属性を追跡する。例えば、各ファイルについてこのメタデータはファイル名、ファイルの生成日または、最終修正日、ファイルについての暗号化情報(encryption information)及び、その他の情報を含み得る。また、メタデータは仮想カートリッジにファイルをリンクする各ファイル毎にストレージシステムによって生成できる。このようなメタデータを使って、ソフトウェアはホストコンピューターにテープカートリッジのエミュレーションを提供するが、ファイルデータは実際にはテープフォーマットで保存されず、返って、後述のとおりシステムファイルに保存される。一連のカートリッジフォーマットよりは、システムファイルにデータを保存するのが、特定のファイルを探すため一連のデータを通じてスキャンする必要なく、個々のファイルに高速、効率的なランダムアクセスができるというメリットがある。

【0046】

前述のとおり、一実施形態によるとファイルデータ(即ち、ユーザー及び/又は、システムデータ)は、システムファイルとしてバックアップ・ストレージ媒体に保存され、各システムファイルは実際のユーザー及び/又は、システムファイルのデータとヘッダーを含む。各システムファイル 200 のヘッダー 202 はユーザー及び/又は、システムファイルを仮想カートリッジにリンクするメタデータを含んだテープディレクトリ 206 を含む。“メタデータ”という用語はユーザー及び/又は、システムファイルデータではない実際のユーザー及び/又は、システムデータの属性を表すデータを意味する。一例によると、テープディレクトリはバイトレベル以下の仮想カートリッジ上のデータレイアウトを規

定し得る。一実施形態において、テープディレクトリ 206 は図 6 に図示したとおり、テーブル構造を有する。上記のテーブルは保存された情報タイプについてのコラム 220 (例えば、データ、ファイルマーク (FM) など)、バイトで使われたディスクブロックのサイズについてのコラム 222、及びファイルデータが保存されたディスクブロックの数を反映するコラム 224 を含む。故に、テープディレクトリは制御器がバックアップ・ストレージ媒体 126 に保存された任意のデータファイルに、ランダム (連続の反対) アクセスできるようとする。例えば、図 6 に図示したとおり、テープディレクトリはファイルのデータ 226 が、システムファイル 200 の始まりから一つのブロックを始めることを指示するので、データファイル 226 は仮想テープ上に迅速に配置できる。この一つのブロックはファイルマーク (FM) に対応するので、サイズを有しない。ファイルマークはシステムファイルに保存されない。即ち、ファイルマークはゼロデータ (zero data) に対応する。従来のテープ及び、バックアップ/復旧アプリケーションによって使われることで、(テープディレクトリはファイルマークを含むのに) これは、テープファイルと共にファイルマークを記録し、仮想カートリッジを見る時、ファイルマークをも見たがるためである。従って、ファイルマークはテープディレクトリ内で追跡を行う。しかし、ファイルマークは任意のデータを示さないので、システムファイルのデータセクション内に保存されない。ファイルのデータ 226 は矢印 205 で表示したシステムファイル・データセクションの最初の部分から始まり、長さは 1024 バイトである (即ち、一つのディスクブロックはサイズが 1024 バイトである)。他のファイルデータはデータの量、即ちデータファイルのサイズによって 1024 バイトではない、他のブロックのサイズで保存できるものと認識されるべきである。例えば、もっと大きいデータファイルは効率のため、さらに大きいブロックサイズを使って保存できる。

10

20

30

40

50

【0047】

一例において、テープディレクトリはストレージシステムにバックアップされた、各データファイルに関する “ファイルディスクリプタ” に含まれ得る。ファイルディスクリプタはストレージシステムに保存されたデータファイル 204 に関する、メタデータを含む。一実施形態において、ファイルディスクリプタは、大抵のユニックス基盤のコンピューターシステムに使われるテープアーカイブ (タール)・フォーマットのようなスタンダードフォーマットで具現できる。各ファイルディスクリプタはユーザーファイルに対応する名前、ユーザーファイルの生成/修正日、ユーザーファイルのサイズ及び、ユーザーファイルへのアクセス制限可否などの情報を含み得る。ファイルディスクリプタに保存された追加情報は、データがコピーされたディレクトリ構造を説明する情報をさらに含み得る。故に、ファイルディスクリプタは後述のとおり、対応するデータファイルに関する検索可能なメタデータを含み得る。

【0048】

バックアップ/復旧アプリケーションの観点では、任意の仮想カートリッジはファイルディスクリプタに対応する複数のデータファイルを含み得る。ストレージシステム・ソフトウェアの観点では、データファイルが特定のバックアップ作業にリンクできるシステムファイルに保存される。例えば、特定の時間に一つのホストコンピューターによって実行されたバックアップは、一つ以上の仮想カートリッジに対応する一つのシステムファイルが生成できる。仮想カートリッジは任意のサイズであり得、仮想カートリッジに保存されるユーザーファイルの増加によって動的に増えていくことができる。

【0049】

上記の図 3 を参考すると、ストレージシステム 70 は総合フルバックアップ・ソフトウェアアプリケーション 240 を含み得る。一実施形態において、ホストコンピューター 120 は、エミュレーティングされた媒体 134 上にデータをバックアップして一つ以上の仮想カートリッジを形成する。幾つかのコンピューター環境において、“フルバックアップ” 即ち、ネットワーク内の第一ストレージシステム (図 1) に保存された全てのデータのバックアップコピーは周期的に (例えば、毎週) 作成される。この処理は一般的にコピーするデータが大容量なので、とても時間がかかる。従って、大抵のコンピューティング環境に

おいて、追加的なバックアップ、一名増分的バックアップは、連続的なフルバックアップ例えば、毎日のフルバックアップ中に遂行できる。増分的バックアップは一つの処理なので、増分的バックアップであれ、フルバックアップであれ、最後のバックアップが遂行された以後に変わったデータだけがバックアップされる。一般的に、ファイル内の多くのデータが頻りに変更されなくても、変更されたデータはファイル基盤にバックアップされる。故に、増分的バックアップはフルバックアップのケースより少ないので、高速で果たされる。大抵の環境では一般的に毎週一回ずつフルバックアップを実行し、増分的バックアップは毎日実行するが、このような時間フレームが使われる必要がないということを認識するべきである。例えば、ある環境では一日に何回も増分的バックアップが必要となる。この発明の原理はどれほど頻りに実行されるかとは関係なく、フルバックアップ(選択的な増分的バックアップ)を使う全ての環境に適用される。

10

【0050】

フルバックアップ手順が実行される間、ホストコンピューターは複数のデータファイルから成るバックアップされたデータを含む一つ以上の仮想カートリッジが生成できる。明確性のため、後述の説明ではフルバックアップが単に一つの仮想カートリッジを生成すると仮想する。しかし、フルバックアップは一つ以上に仮想カートリッジを生成し、本発明の原理は仮想カートリッジの数に限定されないと認識されるべきである。

【0051】

一実施形態によると、一つの既存のフルバックアップ・データセットと一つ以上の増分的バックアップ・データセットから、総合フルバックアップ・データセットを生成する方法が提供される。この方法は周期的(例えば、毎週)フルバックアップを遂行する必要がないので、ユーザーの時間とネットワークリソースを相当に節約できる。また、当業者に自明などあり、例えば、最新バージョンのファイルが増分的バックアップに存在する場合、バックアップ/復旧アプリケーションは一般的に最後のフルバックアップに基づいたファイルを復旧して、増分的バックアップからの全ての変更を適用するため、フルバックアップに基づいた復旧データと一つ以上の増分的バックアップは時間を消費する処理となり得る。故に、総合フルバックアップの提供は、バックアップ復旧アプリケーションが、フルバックアップと一つ以上の増分的バックアップから重ねて復旧する必要なく、総合フルバックアップにのみ基づいて、データファイルをさらに迅速に復旧できるようにする、追加的なメリットを有することができる。“最新バージョン”という用語は、ファイルが新たなバージョン番号を有するのとは関係なく、一般的にデータファイルの最も最近のコピー(即ち、データファイルが保存された最も最近の時間)を意味するものと認識されるべきである。“バージョン”という用語は、幾つかの方法で修正できる、または幾度も保存できる同一のファイルのコピーを意味する。

20

【0052】

図7は総合フルバックアップ手順を概略的に表した図面である。ホストコンピューター120は最初の時間、例えば週末にフルバックアップ230を実行できる。ホストコンピューター120は連続的な増分バックアップ232a、232b、232c、232d、232eを例えば、一週間毎日実行できる。続いて、ストレージシステム170は、後述するように総合フルバックアップ・データセット234が生成できる。

30

【0053】

一実施形態によると、ストレージシステム170は総合フルバックアップ・アプリケーション240(図3)として、言及されたソフトウェア・アプリケーションを含み得る。総合フルバックアップ・アプリケーション240は、ストレージシステム制御器122(図2)または、ホストコンピューター120上で駆動できる。総合フルバックアップ・アプリケーション240は総合フルバックアップ・データセット234の生成に必要なソフトウェア命令とインターフェースを含む。一例として、総合フルバックアップ・アプリケーションはフルバックアップ・データセット230と増分的バックアップデータセット232各々のメタデータ表現の論理的な併合を遂行して、総合フルバックアップ・データセット234を含む新たな仮想カートリッジが生成できる。

40

50

【0054】

例えば、図8に図示したとおり、既存のフルバックアップ・データセットはユーザーファイル(F1、F2、F3、F4)を含み得る。第一増分的バックアップ・データセット232aは、ユーザーファイルF2の修正されたバージョンであるF2及び、F3の修正されたバージョンであるF3を含み得る。第二増分的バックアップ・データセット232bは、ユーザーファイルF1の修正されたバージョンであるF1、F2のさらに修正されたバージョンであるF2及び、新たなユーザーファイルであるF5を含み得る。故に、総合フルバックアップ・データセット234はフルバックアップ・データセット230と二つの増分的データセット232a、232bの論理的併合から形成され、各ユーザーファイル(F1、F2、F3、F4、F5)の最終バージョンを含む。故に、図8に図示したとおり、総合フルバックアップ・データセットはユーザーファイルF1、F2、F3、F4及び、F5を含む。

10

【0055】

図3、図4に図示したとおり、ファイルシステム・ソフトウェア146はエミュレーティングされた媒体134に保存された、各ユーザーファイルに関するメタデータを保存した論理的メタデータキャッシュ242が生成できる。論理的メタデータキャッシュは物理的データキャッシュである必要はないが、代わりにストレージ媒体126に保存されたデータの検索可能なコレクションであり得る。他の例において、論理的メタデータキャッシュ242はデータベースとして、具現できる。メタデータがデータベースに保存された場合、従来のデータベース命令(例えば、SQL命令)はフルバックアップ・データセットと一つ以上の増分的バックアップ・データセットの論理的併合を遂行して、総合フルバックアップ・データセットを生成することができる。

20

【0056】

上記のように、エミュレーティングされた媒体134に保存された各データファイルは、データファイルに関するメタデータが含まれたファイルディスクリプタを含み、バックアップストレージ媒体126上のファイルの位置を含み得る。一実施形態では、ホストコンピュータ120上で駆動されるバックアップ/復旧アプリケーションは、エミュレーティングされた媒体134上にストリーミングテープフォーマットでデータを保存する。図9は、このテープフォーマットを表すデータ構造250の例を図示した図面である。上記のように、システムファイルデータ構造は、データファイルに対するファイルディスクリプタ、ファイルの生成、及び/又は、修正日、セキュリティー情報、ファイルの出所であるホストシステムのディレクトリ構造のみならず、その他の仮想カートリッジにファイルをリンクする情報のようなデータファイルに関する情報を有するヘッダーを含む。このようなヘッダーは、ホストコンピュータ、第1ストレージシステム等からバックアップ(コピー)された実際のユーザー、及び、システムファイルであるデータ254と関係をする。システムファイルデータ構造は、次のヘッダーをブロック境界に適切に整列できるパッド256を選択的に含み得る。

30

【0057】

図9に図示されたように、一実施形態では、ヘッダーデータは、論理的メタデータキャッシュ242に配置され、他の一連のテープデータフォーマットに対する高速検索、及び、ランダムアクセスを可能にする。ストレージシステム制御器122上にファイルシステムソフトウェア148を使用することで、具現された論理的メタデータキャッシュの使用は、エミュレーティングされた媒体134に保存された線形、一連のテープデータフォーマットをバックアップストレージ媒体126を構成する物理的ディスク上に保存されたランダムアクセスデータフォーマットに変換できるようにする。論理的メタデータキャッシュ242は、データファイルに対するファイルディスクリプタを含むヘッダー252、データファイルへのアクセスを制御するのに使用され得るセキュリティー情報、及び、以下に論じるようなポインター256を仮想カートリッジ、及び、バックアップストレージ媒体126上のデータファイルの実際の位置に保存する。一実施形態では、論理的メタデータキャッシュはフルバックアップデータセット230と各増分的データセット232にバッ

40

50

クアップされた全てのデータファイルに関するデータを保存する。

【0058】

一実施形態によると、総合フルバックアップアプリケーションソフトウェア240は、論理的メタデータキャッシュに保存された情報を使用して、総合フルバックアップデータセットを生成する。続いて、この総合フルバックアップデータセットは、総合フルバックアップアプリケーション240によって生成された総合仮想カートリッジにリンクされる。バックアップ/復旧アプリケーションにおいて、総合フルバックアップデータセットは、この総合仮想カートリッジ上に保存されるように見える。上記のように、総合フルバックアップデータセットは、既存のフルバックアップデータセットと増分的バックアップデータセットの論理的併合を行うことで生成され得る。このような論理的併合は、それぞれの既存フルバックアップデータセットと増分的バックアップデータセットに含まれたそれぞれのデータファイルの比較、及び、図8を参照にして説明された最終修正されたバージョンの各ユーザーファイルの混合の生成を含み得る。

10

【0059】

一実施形態によると、図10に図示されたように、総合仮想カートリッジ260は、他の仮想カートリッジ、特に、既存のフルバックアップデータセットと増分的バックアップデータセットを含んだ仮想カートリッジ上のデータファイルの位置をポインティングするポインターを含む。上記図8に関して、挙げられた例を考慮すると、総合仮想カートリッジ260は仮想カートリッジ262上の既存フルバックアップデータセット内のユーザーファイル(F4)(既存フルバックアップデータセットは最新バージョンのユーザーファイル(F4)を含むため)の位置と、例として、仮想カートリッジ264上の増分的データセット232a内のユーザーファイル(F3)の位置をポインティング(矢印268で表示される)するポインター266を含む。

20

【0060】

総合仮想カートリッジは、ポインター266がポインティングするデータを含む全ての仮想カートリッジの識別番号を含んだリスト270も含む。この従属カートリッジリスト270は、実データの位置追跡と従属仮想カートリッジの削除防止のために重要であり得る。この実施形態では、総合フルバックアップデータセットは、実際のユーザーファイルを含まないで、バックアップストレージ媒体126上のユーザーファイルの位置を示すポインターのセットを含む。従って、実際のユーザーファイル(他の仮想カートリッジ上に保存された)の削除を防ぎ得る。これは、データを含んだ仮想カートリッジの記録(従属カートリッジリスト270)を維持し、各仮想カートリッジの上書き(over-written)、または、削除を防ぐことで部分的に達成できる。総合仮想カートリッジは、総合仮想カートリッジの大きさ、バックアップストレージ媒体126上の総合仮想カートリッジの位置と同じカートリッジデータ272を含み得る。また、総合仮想カートリッジは識別番号、及び/又は、名前274を有し得る。

30

【0061】

他の実施形態によると、総合仮想カートリッジは、ポインターと実際に保存されたユーザーファイルの組み合わせを含み得る。図11に図示されたように、一例において、総合仮想カートリッジは、仮想カートリッジ262上の既存フルバックアップデータセット230内のデータファイル(図9を参照にして説明したような最新バージョン)の位置をポインティングするポインター266を含む。総合仮想カートリッジは、矢印280で示された増分的データセット232からコピーされた実データファイルを含むデータ278を含むことができる。このような方法で、増分的バックアップデータセットは、総合フルバックアップデータセット276が生成された後に削除され得るので、ストレージ空間が節約される。上記の総合仮想カートリッジは、全てのユーザーファイルのコピーではなく全体、または、一部ポインターを含む総合仮想カートリッジより小さい。

40

【0062】

総合フルバックアップは、ポインターと保存されたファイルデータの組み合わせを含み、上記例に限定されないということが認識されなければならない。例として、総合フルバ

50

ックアップは、ある増分的、及び/又は、フルバックアップに保存された多数のファイルに関するデータファイルに対するポインターを含み得、他の既存のフル、及び/又は、増分的バックアップからコピーされて保存されたファイルデータを含み得る。また、その代案として、総合フルバックアップはどのポインターも含まないで、適切なフル及び/又は、増分的バックアップからコピーされた最新バージョンの実ファイルデータを含む全ての関連増分的バックアップ、及び、以前のフルバックアップに基づいて生成され得る。

【0063】

一実施形態では、総合フルバックアップアプリケーションソフトウェアは、それぞれの既存フルバックアップデータセットと増分的バックアップデータセットに対するユーザー、及び、システムファイルメタデータを比較して、最新バージョンのデータファイルそれが位置された所を決定できるようする。ディファレンシングアルゴリズム(*differencing algorithm*)を含み得る。例えば、ディファレンシングアルゴリズムは、他のバックアップセット内の同一のデータファイルの相異したバージョンの間で生成日、及び/又は、修正日等を比較して最新バージョンのデータファイルを選択するのに使用できる。しかし、ユーザーは、しばしば、ファイル内の任意のデータを実際に変更せず、ユーザーファイルをオープンし、ファイルを保存することができる(従って、その修正のデータを変更する)。従って、システムはシステム、或は、ユーザーファイル内のデータを分析して、データが実際に変更されたかを決定できる更に向上したディファレンシングアルゴリズムを具現し得る。このようなディファレンシングアルゴリズムの変形、及び、他の形態の比較アルゴリズムは、当業者に自明なものである。また、上記のように、メタデータがデータベースフォーマットで保存される場合、SQL命令等のデータベース命令は、論理的併合を行う時に使用され得る。本発明は、総合フルバックアップデータセットを正確に生成できるように最新、または、最終バージョンの各ユーザーファイルが、全体比較された既存バックアップセットから選択され得るようにする全てのアルゴリズムに適用され得る。

【0064】

当業者に自明であるように、総合フルバックアップアプリケーションは、ホストコンピュータが物理的フルバックアップを行う必要なく、フルバックアップデータセットが生成され、利用可能にする。データをバックアップストレージシステムに転送するプロセッサの負担により、ホストコンピュータに負担をかけないようにするだけでなく、総合フルバックアップアプリケーションがストレージシステムに実行される実施形態において、ネットワーク帯域幅の活用を相当減少させる。図7に図示されたように、第1総合フルバックアップデータセット234と一連の増分的バックアップデータセット236を使用して追加的な総合フルバックアップデータセットが生成される。これは、頻繁に修正されず、頻繁にコピーされないファイル、または、対象に相当な時間的効率を与えることができる。その代わりに、総合フルバックアップデータセットは、単に一回コピーされたファイルに対するポインターを維持することができる。

【0065】

図3を参照にして上記のように、ストレージシステムは、エンドユーザー復旧アプリケーション300としてのソフトウェアアプリケーションを含み得る。従って、他の実施形態によると、エンドユーザーがITスタッフの干渉なく、既存バックアップ/復旧手順、及び/又は、政策を変更する必要なくバックアップデータを探して復旧する方法が提供される。典型的バックアップストレージシステムにおいて、ホストコンピュータ120で駆動されるバックアップ/復旧アプリケーションは、ITスタッフによって制御され、エンドユーザーがITスタッフによる干渉なくバックアップされたデータにアクセスすることは不可能であったり、相当難しくなり得る。本発明の実施形態によると、ストレージシステムソフトウェアは、エンドユーザーが、例えば、バックアップストレージ媒体126とのウェブ基盤または他のインターフェースを通じて自分のファイルを探して復旧することを提供する。

【0066】

総合フルバックアップアプリケーション240と同様にエンドユーザー復旧アプリケー

10

20

30

40

50

ション300は、ストレージシステム制御器122または、ホストコンピュータ120上で駆動できるということが認識されなければならない。エンドユーザー復旧アプリケーションは、認証されたユーザーが論理的メタデータのキャッシュを検索して、バックアップストレージ媒体126からバックアップされたファイルを探し、選択的に復旧させるのに必要なソフトウェア命令とインターフェースを含む。

【0067】

一実施形態によると、ユーザーコンピュータ136上で設置、及び/又は、実行されるユーザーインターフェースを含むソフトウェアが提供される。ユーザーインターフェースは、ユーザーがバックアップストレージ媒体上のファイルを探すようにする全ての形態のインターフェースになり得る。例えば、ユーザーインターフェースは、グラフィックユーザーインターフェース、ウェブ基盤、又は、テキストインターフェース等になり得る。ユーザーコンピュータは、例えば、イーサネット連結のようなネットワーク連結138を通じてストレージシステム170に連結される。このネットワーク連結138を通じて、ユーザーコンピュータ136のオペレータはストレージシステム170に保存されたデータにアクセスできる。

【0068】

一例において、エンドユーザー復旧アプリケーション300は、ユーザー認証、及び/又は、認証特徴を含む。例えば、ユーザーは、ユーザー名とパスワードを使用するユーザーコンピュータ上のユーザーインターフェースを通じてログインを要請されえる。ユーザーコンピュータは、適切なユーザー立証メカニズムを使用して、ユーザーがストレージシステムにアクセスしたか否かを決定できるストレージシステム（例えば、エンドユーザー復旧アプリケーション）でユーザー名とパスワードを転送することができる。ユーザー立証メカニズムに含められるが、これに限定されないいくつかの例としてはMicrosoft Active Directory server、Unix "yellow pages" server、又は、Lightweight Directory Access Protocol等がある。ログイン/ユーザー立証メカニズムは、エンドユーザー復旧アプリケーションと通信してユーザー権限を転換することができる。例えば、いくつかのユーザーは自分が生成したファイルのみを検索できるようにすることもでき、又は、所定の権限を持ったり、オーナー（owner）として識別され得もする。例えば、システムオペレータ、または、管理者のような他のユーザーは、バックアップされたファイル全てに対してアクセスが許容され得る。

【0069】

一実施形態によると、エンドユーザー復旧アプリケーションは、論理的メタデータキャッシュを使用して、バックアップストレージ媒体上にバックアップされた全てのデータファイルに対する情報を得る。エンドユーザー復旧アプリケーションは、例えば、バックアップ時間、バックアップ日、ユーザー名、オリジナルユーザーのコンピュータディレクトリ構造（ファイルがバックアップされた場合に得ることができる。）、又は、他のファイル特性等によって分類されたユーザーファイルの階層的ディレクトリ構造を、ユーザーインターフェースを通じてユーザーに提供する。一例において、ユーザーに提供されるディレクトリ構造は、ユーザーに与えられた権限によって変わることができる。エンドユーザー復旧アプリケーションは、ブラウジング要請（すなわち、ユーザーインターフェースを通じて、ユーザーがディレクトリ構造をブラウジングして所望のファイルを探す）を受けたり、ユーザーが名前、日等によってファイルを検索することができる。

【0070】

一実施形態によると、ユーザーはストレージシステムからバックアップされたファイルを復旧することができる。例えば、ユーザーが所望のファイルを探すと、上記のように、ユーザーは、ネットワーク連結138を通じて、ストレージシステムから上記のファイルをダウンロードすることができる。一例において、このようなダウンロードの手順は、当業者に公知されたように、ウェブ基盤ダウンロードに匹敵する方式で具現され得る。

【0071】

ビューアリング/ダウンロードに関する許可を有したエンドユーザーが、ファイルにアク

10

20

30

40

50

セスできるようにし、ユーザーインターフェースを通じてこのアクセスを可能にさせることで、エンドユーザー復旧アプリケーションは、ユーザーが自分のファイルをバックアップ政策、又は、手順を変更せずに検索、及び、復旧させることができる。

【0072】

他の実施形態によると、ユーザーがバックアップストレージ媒体126上に保存されたバックアップデータセットのビュー(view)が付着されたネットワークを“マウンティング”し得る方法、及び、メカニズムが提供される。これは、ユーザーが自分のコンピュータに連結された任意の他のローカル、又は、ネットワークドライブ上のデータをビューイングし、アクセスするのと同様で、マウンティングされたデータセット内のデータをビューイングし、アクセスできるようにする、従って、例えば、ユーザーは媒体サーバー144(図1)を通じた復旧処理を実行せずにアプリケーションサーバー[例えば、システム第1ストレージ装置106(図1)が失敗した場合]で、データを有効に復旧することができる。上記のようなマウンティング手順を使用したアプリケーションサーバーへのデータ復旧は、ボリューム復旧が容易な典型的媒体サーバーに比べて、数十倍早い速度でなり得る。“マウンティング”という用語は、ネットワークドライブ等のネットワーク構成要素、または、データボリュームをホストコンピュータの運営システムで利用可能にすることを意味するものと認識されなければならない。データボリュームは、例えば、单一データファイル、または、システムファイル、複数のファイル、または、複数のファイルを含むディレクトリ構造等を含むことができる。カモンマウンティングプロトコル(common mounting protocol)は、NFS(network file system)、または、CIFS(common internet file system)シェアーリング(sharing)を含む。このようなプロトコルは、ホストコンピュータが、リモートリソース(remote resource)がホストコンピュータ上に局部的に提供されるものと現れるインターフェースを通じて、ネットワーク連結を経て、他のコンピュータ上のリソースにアクセスできるようにする。

【0073】

図12は、本発明の一実施形態によるボリュームマウントを行なう方法を示した順序図である。第1段階290において、ユーザーはデータボリュームを選択してマウンティングし、バックアップストレージシステム制御器122にボリュームマウント要請を伝達する(図3)。一般的に、ユーザーは、バックアップされた情報の全体的、及び、正確な表現をキャプチャーし得るように、フルバックアップデータセット(増分的バックアップデータセットでなく)から、データを復旧するのを所望することがある。現在のフルバックアップデータセットが存在しない場合(例えば、ネットワークマネージャが、フルバックアップを毎週実行することにより、ユーザーが週中にデータを復旧するのを希望しても、現在のフルバックアップを利用できない場合)、総合フルバックアップが生成され、選択されたデータの復旧に使用され得る。

【0074】

一実施形態によると、バックアップストレージシステム170は、データボリュームマウントと復旧手順を行なう方法を制御して具現できるボリューム復旧アプリケーション310(図13)であるソフトウェアアプリケーションを含むことができる。総合フルバックアップ、及び、エンドユーザー復旧アプリケーションと類似したボリューム復旧アプリケーション310は、ホストコンピュータ、及び/又は、ユーザーコンピュータ上で実行することができ、その一部は、ストレージシステム制御器、ホストコンピュータ、及び、ユーザーコンピュータの全体、または、一部に分配され得る。

【0075】

上記図12を再び参照すると、ボリュームマウンティングが要請された後、ボリューム復旧アプリケーションは、現在のフルバックアップデータセットが利用可能であるか否かを照会することができる(段階292)。利用ができない場合、ボリューム復旧アプリケーションは、総合フルバックアップアプリケーション240と通信して総合フルバックアップ処理を行ない(図1参照)、現在のバックアップデータセットを生成することができる(段階294)。ボリューム復旧アプリケーションは、正規のフルバックアップデータ

セット、または、総合フルバックアップデータセットをエクスポートинг(exporting)し、要請されたボリュームマウンティングをNFS、または、CIFSシェアによって行なうことができる。特に、ボリューム復旧アプリケーションは、論理的メタデータキャッシュ242を照会して、段階290で識別され選択されたフルバックアップボリュームを示す適切なメタデータを探す。

【0076】

一実施形態によると、マウント要請(段階290)は、ボリューム復旧アプリケーションが一つ以上のファイルディスクリプタ構造を生成し、NFSまたはCIFSシェアによるマウンティングに対するボリュームのエクスポートингを容易にする(段階296)。図14は、ボリューム復旧アプリケーションによって生成され得るファイルディスクリプタ構造320の一実施形態を示した図面であり、ファイルディスクリプタ320は、テープフォーマットにおいてシステムファイル(例えば、システムファイル322、図15参照)に対応する。上記のように、ファイルディスクリプタは、ストレージシステムに保存されたシステムファイルとデータファイルに対応する検索可能なメタデータを含む。ファイルディスクリプタ320は、例えば、マウンティングされるボリュームに含まれたデータファイルに対するファイルパーミッション(アクセス制御ファイル)324とファイル名322等の情報を含む複数のフィールドを含むことができる。また、ファイルディスクリプタは、データファイルのソースデータの位置(すなわち、ストレージ媒体126上にデータファイルが保存された位置を識別するために)、データファイルの長さ328に対する一つ以上のポインター326、及び、リンクされたリストファイルディスクリプタ構造内の次のエントリ(entry)(例えば、次のデータファイル)に対するポインター330を含む。例えば、参照番号331によって示された“次”的フィールドがナル(null)である場合、データファイルがファイルディスクリプタ320によって示されたシステムファイルに知られている最新のデータファイル(例えば、最近リンクされたリストエントリである。)であるということを示す。マウンティングされるデータボリューム内に含まれた各システムファイルは、図14に図示されたようなファイルディスクリプタ構造によって表現される。要請されたボリューム内の各システムファイルが、生成されたファイルディスクリプタ320を有する場合、ファイルディスクリプタはNFS、または、CIFS要請に答える関連データファイルを探してエクスポートингすることに使用され得る。

【0077】

上記のように、一実施形態において、ファイルディスクリプタは、大部分のユニックス基盤コンピュータシステムに使用されるテープアーカイブ(タール)フォーマット等の標準化されたフォーマットによって具現され得る。図15は、テープ(例えば、タール)データストリームのセグメントによるテープフォーマットで記録された典型的システムファイル332を示した図面である。図16は、システムファイル332に対する対応ファイルディスクリプタ340を示した図面である。図15に図示されたように、テープフォーマットで記録されたファイルは、システムファイル332に保存された実データ338とヘッダー336を含む。データ338は、一つ以上のデータファイルに対応することができる。図示された例において、システムファイル332の長さは1032バイトであるが、上記ファイルはファイルの大きさと記録されたフォーマットによって任意の長さを有し得る。

【0078】

ファイル332に対するファイルディスクリプタ340は、ヘッダー336に含まれる。図16に図示されたように、そして、図14に図示された一般的な例と同様に、ファイルディスクリプタ340は、ファイル名341、保安情報344、システムファイルに知られている各データの保存されたデータに対するポインター342、対応データファイルの長さ346、及び、図示された例においてナル(null)348であるシステムファイルに知られている次のデータファイルを識別する“次”的エントリを含む。

【0079】

上記図12を再び参照すると、マウンティングされたデータボリューム内のファイルに

10

20

30

40

50

対する全てのファイルディスクリプタが生成された場合、ボリューム復旧アプリケーションは、生成されたファイルディスクリプタに基づいたファイルシステムをユーザーが特定のマウントポイントにNFS、または、CIFSシェアによってエクスポートしている(段階298)。このポイントで、マウントが完了し(段階299)、マウンティングされたデータボリュームは後述するように、ユーザーがデータを判読、及び/又は、記録するのに利用可能である。

【0080】

一実施形態によると、NFSまたはCIFS判読動作 [すなわち、ユーザーがマウンティングされたデータボリューム内のデータをビューアリング(viewing)することを所望] は、ファイル特定をマッチングさせるためのファイルディスクリプタを通じて検索することでサービスされる。一実施形態によると、ユーザーは、自分が直接ファイルディスクリプタを実際に検索する必要がないということが認識されなければならない。その代わりに、ボリューム復旧アプリケーションは、例えば、典型的なディレクトリ構造フォーマット内でユーザーにデータを提供するユーザーインターフェースを含むことができる。ボリューム復旧アプリケーションは、特定ファイルに対するユーザー要請を論理的メタデータキャッシュにアクセスする検索命令に変換して、マッチングシステムファイルに対するファイルディスクリプタ320を検索するソフトウェアを含むことができる。ファイルを探すことができた場合、ユーザーコンピュータへのデータ転送は、リンクされたリストをフォローイング(following) [すなわち、実データを探すためにファイルディスクリプタに保存されたポインターをフォローイング] することで達成され、要請したユーザーに送られ得るファイルデータのためにバッファを生成する。

【0081】

他の実施形態によると、また、ユーザーがマウンティングされたボリュームに新たなデータを記録するためにメカニズムが提供され得る。上記のように、マウンティングされたボリュームデータは、ユーザーに普通のネットワークドライブ、または、他のネットワーク - 保存されたデータとして見られることがある。しかし、実際には、オリジナルマウンティングされたボリュームデータは、一般的に少なくとも他のバックアップデータセットが生成するまでには、保護される必要がある実際のバックアップデータである。従って、ユーザーがオリジナルバックアップデータを実際に修正し得るようにするには望ましくなかったりもする。ユーザーがマウンティングされたボリュームに対応するデータを修正できるようになっている間、バックアップデータの修正を防止するために、後述するように、他のストレージ媒体への記録に転換するメカニズムが提供される。

【0082】

図17は、本発明の一実施形態による記録要請を処理する方法を示した順序図である。最初の段階350において、ユーザーはNFSまたはCIFS記録動作(一般的にデータファイルをエディッティングまたはビューアリングする間、“保存”オプションを選択することによる。)を要請する。ボリューム復旧アプリケーションは、利用可能なストレージ空間を探してその空間にデータを記録し、及び、新たに記録されたデータを参照するための適合したファイルディスクリプタを更新することで記録要請を実行する。

【0083】

一実施形態によると、ボリューム復旧アプリケーションは、データを記録するためのストレージ空間がすでに分配されたか否かを照会し(段階352)、分配されていない場合、ストレージ空間を分配する(段階354)。ストレージ空間は、バックアップストレージ媒体126に分配され得る(図13)。分配された保存空間は、記録データのみをホールディングするために特別に示されることができる(関連したメタデータは選択的である。)。

【0084】

図18は、バックアップストレージ媒体126に保存されたNFS、または、CIFS記録データの一例を示した図面である。記録されたデータ360は、ボリューム復旧アプリケーションによってサービスされた記録命令の結果として発生した保存されたデータに対応す

10

20

30

40

50

る、例えば、二つの記録された部分である $w1_362$ 、 $w2_364$ を含む。例えば、 $w1$ 及び $w2$ は、マウンティングされたデータボリューム内に含まれた修正されたデータファイルに対応することができる。二つの記録要請に對応して図示されてはいるが、本発明の原理は、記録要請の数に限定されず適用されることができ、ファイルは記録要請の数によって適合するよう変更され得るということが認識されなければならない。記録されたデータ 360 は、オリジナルデータ（例えば、ファイル 332 ）と新たに記録されたデータ 360 間の磁気表示関係を形成するメタデータを含むヘッダーも含む。特に、図 19 を更に参照して後述するように、ヘッダーは記録されたデータ部分 $w1$ 、 $w2$ が、オリジナルデータに関連して論理的に存在するところを示すオフセット情報を含み得る。

【0085】

図 19 は、二つの記録要請がサービスされた以後のシステムファイルレイアウトの一例を示した図面である。オリジナルシステムファイル 332 は、バックアップストレージ媒体 126 （図 13 ）に保存され、上記のマウンティング手順を通じてユーザーに提供される。図 19 に図示されたシステムファイル 332 は、データフォーマット内にあり、データ部分 338 は、複数のデータファイル（例えば、ユーザーファイル）を含むことができる。データは、オフセットゼロバイト（ポイント 370 ）から始まり、後でポイント 372 で終了する。記録されたファイル 360 は、ユーザーの要請に對応してファイル 332 にデータを記録する。例えば、ユーザーは、システムファイル 332 内に含まれた二つのデータファイルを修正することができ、結果的に、記録されたファイル 360 は、 $w1$ 及び $w2$ を含む。上記のように、この記録されたファイル 360 は、オリジナルバックアップデータを変更しないようにストレージ媒体上のファイル 332 から分離され、保存されることができる。論理的に修正されたシステムファイル 380 が図示され、記録要請を通して使用者による変更（すなわち、記録されたファイル 360 ）を含むファイル 332 を示す。すなわち、修正されたシステムファイル 380 において、 $w1$ 及び $w2$ （使用者修正データファイル）は、バックアップされたデータを除去せずに、オリジナルシステムファイル 332 のデータ部分内に含まれたオリジナルデータファイルを代替することに使用され得る。

【0086】

図 19 に図示されたように、修正されたシステムファイルは、オリジナルシステムファイル 332 と記録されたファイル 360 の論理的併合（summation）に對応する。図示されたように、オリジナルシステムファイルデータ 338 は、オリジナルファイル内のオフセットゼロで始まる。オフセット 64 （参照番号 384 ）で、修正されたデータの第 1 部分（ $w1$ ）が始まり、 9 バイトが追加されたオフセット 73 （参照番号 386 ）で終了する。従って、ユーザーの記録要請によるユーザー修正データファイルである $w1$ は、オリジナルシステムファイル 332 内のオフセット 64 に位置されたオリジナルデータファイルを代替するのに使用することができる。 $w1$ は、記録されたファイル 360 内のオフセット 0 （ 390 ）から存在し、記録されたファイル 360 内のオフセット 9 （ 392 ）で終了するため、 $w1$ の長さは 9 バイトになる。修正されたファイル内 $w1$ のスタート位置（図示された例において、オフセット 64 ）は、ヘッダー 366 に保存された情報、即ち、記録されたファイル 360 とオリジナルファイル 332 との間の相対的関係により決定される。 $w2$ 部分も修正されたファイル 380 内に含まれ、オフセット 1032 （ファイルのオリジナルエンド、参照番号 372 ）で始まり、論理的にファイルを 100 バイト延長する。又、 $w2$ の長さは、ヘッダー 366 に位置された情報から決定される。ファイルの新しい終了ポイントは参照番号 388 で表示される。

【0087】

修正されたファイルは、論理的に生成され、ユーザー修正バージョンのオリジナルファイルで表現されるが、ファイル 360 により表現された新たに記録されたデータは、オリジナルファイル 332 の一部として実際に保存されない。代わりに、上記したように、新たに記録されたデータは、データを記録するために識別されたストレージ媒体上の特定の位置に保存される。このような方式で、一般的なローカル又はネットワークドライブと同

10

20

30

40

50

様に、ユーザーがマウンティングされたボリュームに外観上、記録が可能な反面、オリジナルバックアップデータの保全が維持される。

【0088】

修正されたファイル380は、修正されたファイルを示すファイルディスクリプタを含んだヘッダー382を含む。図20は、このようなファイルディスクリプタ400の例を示した図面である。ファイルディスクリプタ400は、修正されたファイル380のファイル名を識別するネイムフィールド(name field)402、及び修正されたファイル380の許容属性を識別する保安フィールド(security field)404を含む。ファイルディスクリプタ400は、それぞれのオリジナルファイルと記録されたファイルに保存されたデータをキャプチャーするためのオリジナルファイル332に対するポインタ、及び記録されたファイル360に対するポインタを含む複数のデータフィールドも含む。ファイルディスクリプタ400に与えられたポインタのリンクされたリストを連続的にフォローイングすることによって、修正されたファイル380の表示が与えられる。

【0089】

図19及び図20には、修正されたファイルに対するファイルディスクリプタの一例が図示されて説明されている。第1データフィールド406において、図19で参照番号408で識別されるオフセットゼロバイトにある修正されたファイル380内の第1データファイル位置に対するポインタが位置する。フォローイングフィールド410は、ポインタ406により位置が特定されたデータファイルの長さを表示する。図示された例において、図19で見るように、長さは64バイトである(ゼロオフセットポイント408と64バイトのオフセット384との間でデータが延長される)。次のフィールド412は、図19に図示されたように、修正されたファイル380内の次のデータファイルがW1であることを表示する。従って、ポインタ414は、W1に対応したデータの位置が、ゼロオフセットポインタ(参照番号390、図19)で新たに記録されたファイル360に保存されるということを表示する。長さフィールド416は、図19で見るように、W1は修正されたファイル380内でオフセット64(384)とオフセット73(386)との間で延長され、W1の長さが9バイトであるということを表示する。次のフィールド418は、修正されたファイル380内の次のデータファイルが、オリジナルシステムファイル332からのデータファイルであるということを表示する。フィールド420内のポインタは、次のデータファイルが、修正されたファイル380内のオフセット73(図19の参照番号386)に位置されるということを表示する。フィールド422は、図19に図示されたように、データファイルの長さが959バイトということを表示する。次のフィールド424は、フォローイングデータファイルがW2であることを表示する。また、フィールド426内のポインタは、W2の位置、即ち、図19に図示されたように、新たに記録されたファイル360のオフセット9を表示する。フィールド428は、W2の長さが100バイトであるということ、次のフィールド430はナルを含むということ、図19に図示されたように、W2が、修正されたファイル380内の最終データファイルであることを表示する。従って、ファイルディスクリプタ400は、修正されたファイル380の構造、及び修正されたファイル380に含まれたデータの位置を示す「ロードマップ(roadmap)」を含む。

【0090】

上記のボリューム復旧アプリケーション、及び方法は、一連のテープフォーマットデータをNFS又はCIFS等のランダムアクセスI/Oシステムに適合する形態で表示する。ファイルディスクリプタ400のようなリンクされたリストファイルディスクリプタは、特定タールストリーム(tar stream)内の各データファイルのストレージ媒体上の位置と共に、例えば、タールストリーム内の他のデータファイルに関連したタールストリーム内の各データファイルの位置を記録することによって、一連のテープフォーマットデータを、ランダムアクセスが可能なデータに変換するのに使用され得る。また、一実施形態によれば、ボリューム復旧アプリケーションは、バックアップ/復旧アプリケーションが、上記の普通の方式でデータにアクセスできるよう、変更された(即ち、記録された)データバックをテー

10

20

30

40

50

プ(例えば、タール)フォーマットで表示するプロビジョン(provision)を含み得る。一実施形態によれば、インスタンストレージシステムは、ファイルシステムソフトウェアについて、上記した方式でテープヘッダー、パッド、データ、及びファイルマークで適切にフォーマットされた仮想カートリッジを生成する設備を含む。他の実施形態において、ボリューム復旧アプリケーションは、ファイルシステムソフトウェアとインターフェースされて新たに記録、及び修正されたファイルを含む上記のような仮想カートリッジを生成することができる。

【0091】

本発明において、総合フルバックアップアプリケーション、エンドユーザー復旧アプリケーション、及びボリューム復旧アプリケーションのようなソフトウェアの用語が主に使用されたが、ソフトウェア、ハードウェア又はファームウェア、又はその組み合せで他の形態が選択的に具現できることが認識されなければならない。従って、本発明の実施形態は、ストレージシステムのプロセッサで少なくとも一部が実行され、前記したような総合フルバックアップアプリケーション及び/又はエンドユーザー復旧アプリケーションの機能を遂行する場合、コンピュータプログラムでインコーディングされた全てのコンピュータ判読可能媒体(例えば、コンピュータメモリ、フロッピーディスク、コンパクトディスク、テープ等)が含まれ得る。

【0092】

要するに、本発明による実施形態は、従来のテープバックアップシステムをエミュレートするが、エンドユーザーがバックアップされたファイルをビュー又は復旧するようにし、総合バックアップを生成できるような向上した機能性を提供できるストレージシステム、及び方法を含む。しかし、本発明による多様な形態は、コンピュータデータのバックアップ以外に使用され得る。本発明によるストレージシステムは、保存されたデータが、ハードディスクアクセス時間において、連続的でなくランダムにアクセスできる大容量データを経済的に保存するのに使用され得もし、本発明による実施形態は、従来のバックアップストレージシステム以外でも具現できる。例えば、本発明による実施形態は、映画と音楽の幅広い選択を意味するビデオ及び/又はオーディオ・オンデマンドが可能なビデオ又はオーディオデータを保存するのに使用され得る。

【0093】

本発明の一つ以上の実施形態の幾つかの様相に関する詳細な説明により、当業者が、多様な変形、修正、及び改良を行えることを認識しなければならない。このような変形、修正、及び改良は、この詳細な説明の一部として意図され、本発明の思想内で意図されたものである。従って、上記説明と図面は例示のみのためのものである。

【図面の簡単な説明】

【0094】

【図1】バックアップストレージシステムを含む大型のネットワーキングされたコンピューティング環境の一例を示したブロック図である。

【図2】本発明によるストレージシステムを含む、ネットワーキングされたコンピューティング環境の一実施形態のブロック図である。

【図3】本発明によるストレージシステムの一実施形態のブロック図である。

【図4】本発明によるストレージシステムの一実施形態の仮想レイアウトを示したブロック図である。

【図5】本発明の実施形態によるシステムファイルの一例の概略的なレイアウトである。

【図6】本発明の実施形態によるテープディレクトリ構造の一例を示した図面である。

【図7】本発明の実施形態による総合フルバックアップを生成する方法の一例を示した図面である。

【図8】本発明の実施形態による総合フルバックアップを含むバックアップ・データセットのシリーズの一例の概略的な図面である。

【図9】メタデータキャッシュ構造(metadata cache structure)の一例を示した図面である。

10

20

30

40

50

【図10】総合フルバックアップ・データセットを保存する仮想カートリッジの一例を示した図面である。

【図11】総合フルバックアップ・データセットを保存する仮想カートリッジの他の例を示した図面である。

【図12】本発明の実施形態によるバックアップ・ストレージシステムからデータを復旧するための方法の、一実施形態のフローチャートである。

【図13】本発明の実施形態によるバックアップ・ストレージシステムを含むネットワーキングされたコンピューティング環境の他の実施形態のブロック図である。

【図14】本発明の実施形態によるファイルディスクリプタ構造(file descriptor structure)の一例を示した図面である。

【図15】ファイルデータがテープフォーマットで保存され得る方法の一例を示した図面である。

【図16】図15に図示されたファイルに対するファイルディスクリプタを示した図面である。

【図17】本発明の一実施形態によってマウンティングされた、データボリュームにデータ

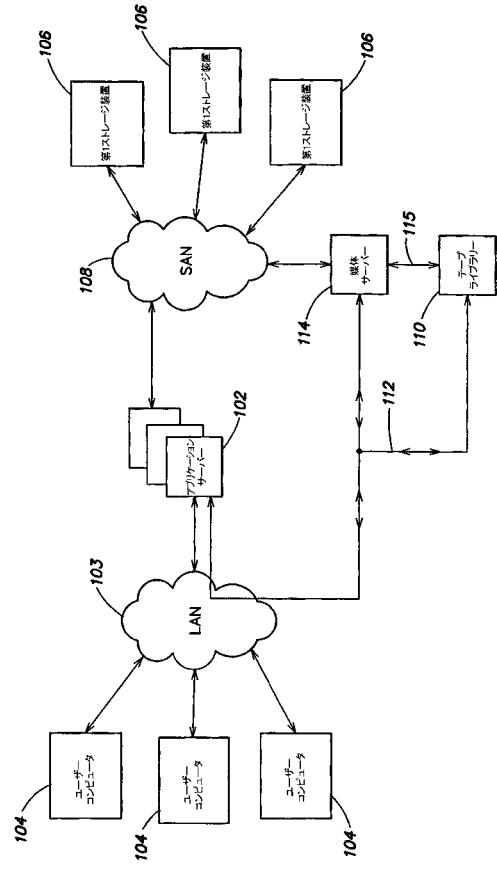
【図1.8】新しく記録されたファイルの一例を示した図面である。

【図1-8】本発明の一実施形態によるオリジナルスタイル「新」の例を示した図面である。

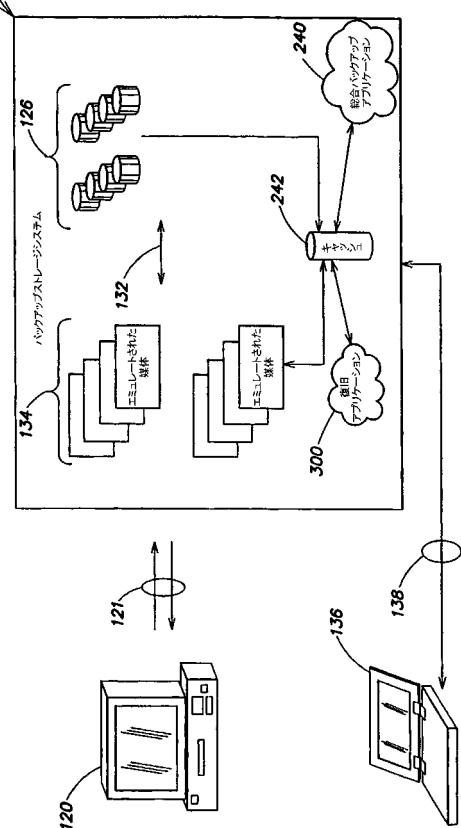
【図19】本発明の実施形態によるオブジェクトファイル、新しく記録されたファイル及び、最終的に修正されたファイル間の関係に対する一例を示した図面である。

【図20】図19に図示した修正されたファイルを示す。ファイルディスクリソラの一例を示した図面である。

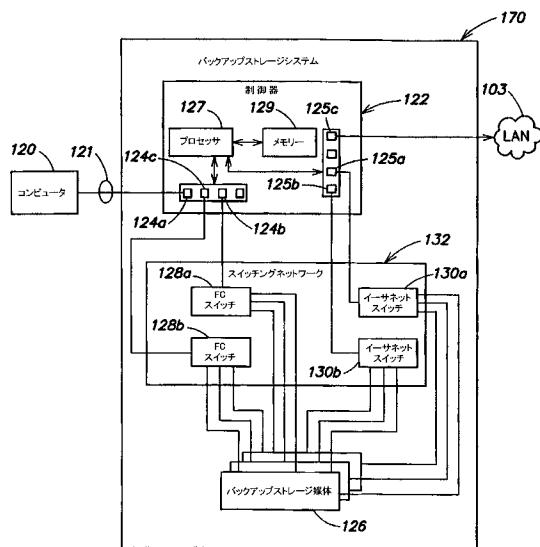
【 四 1 】



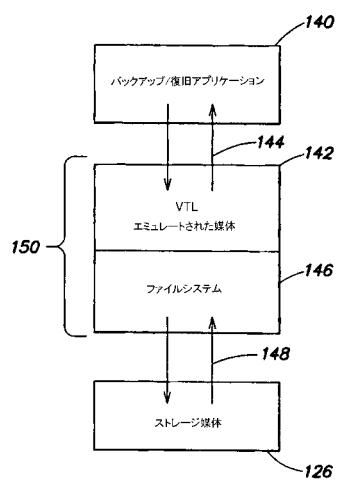
【 図 2 】



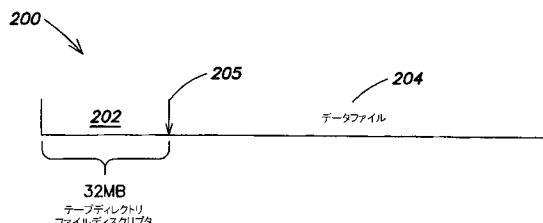
【図3】



【図4】



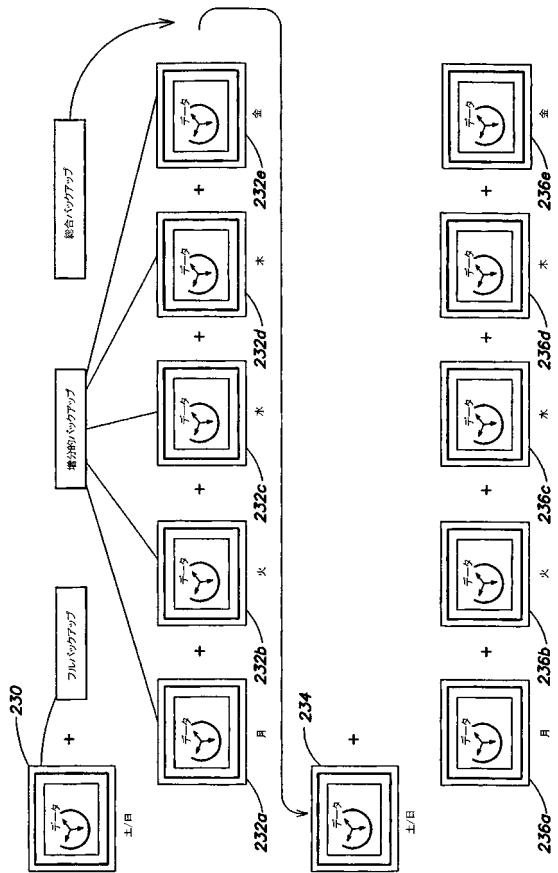
【図5】



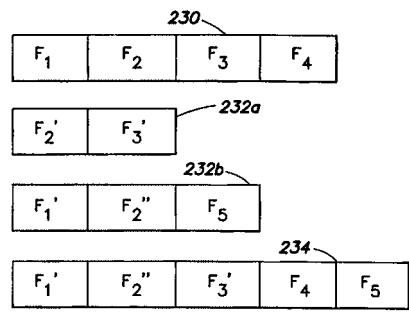
【図6】

形態	サイズ	ブロックカウント
FM	-	1
データ	1024	4
データ	256k	100,000
FM	-	1
データ	1024	1
...
...

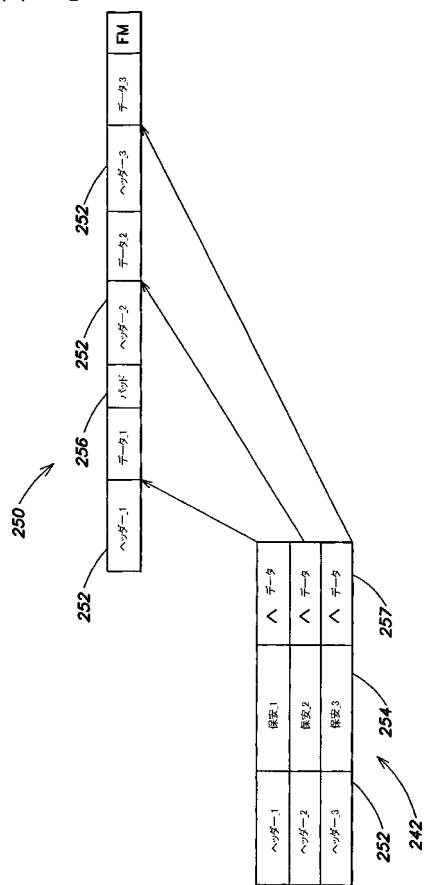
【図7】



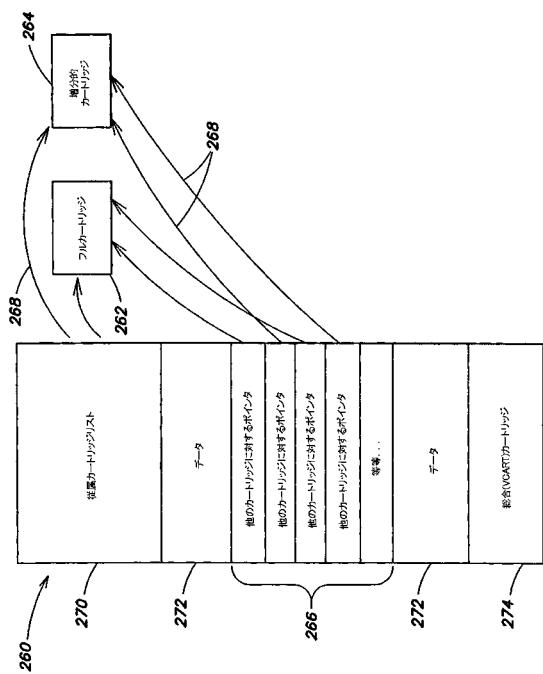
【図8】



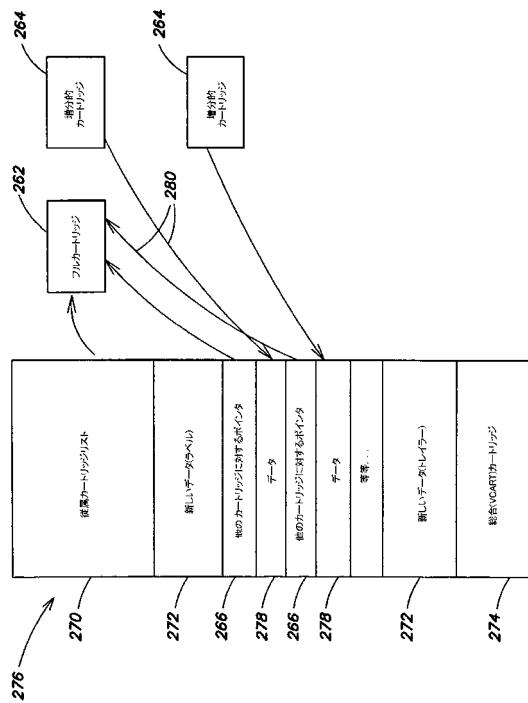
【図9】



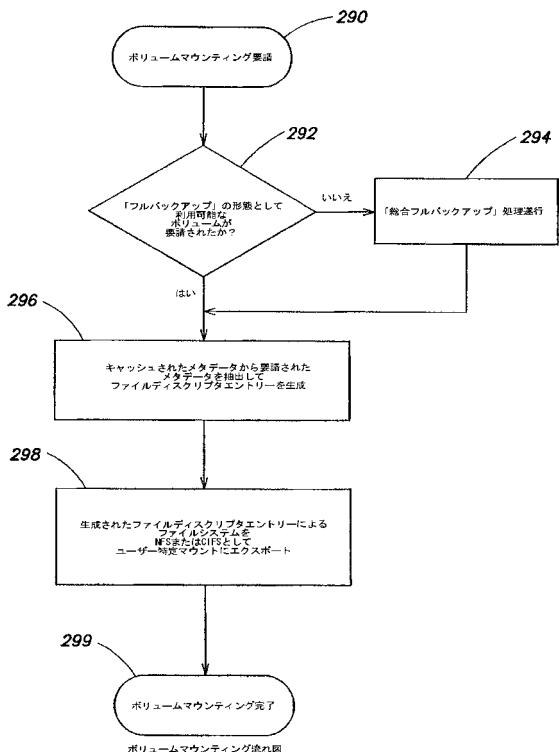
【図10】



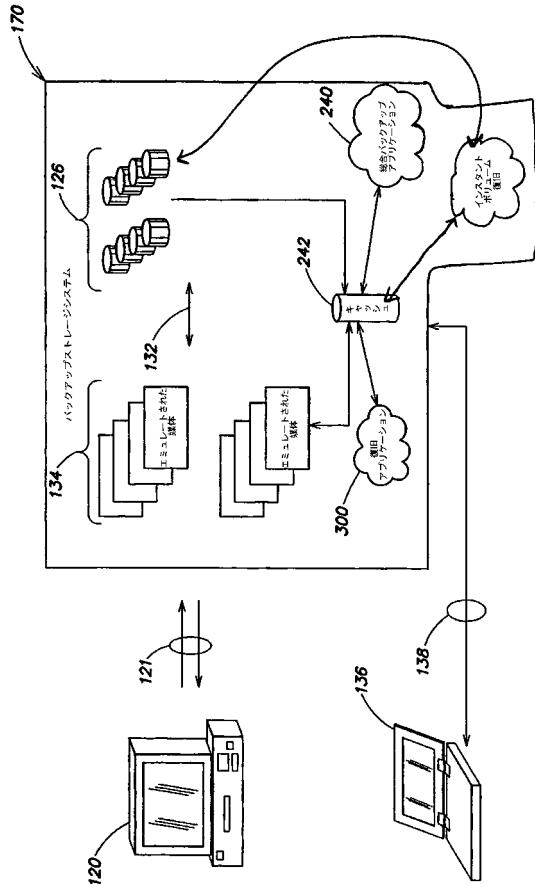
【図11】



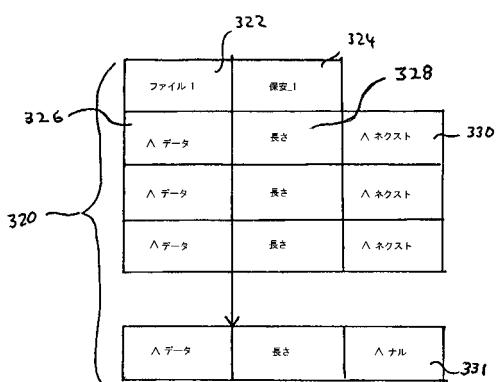
【図12】



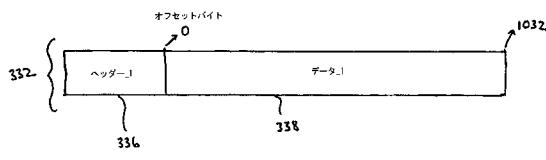
【図13】



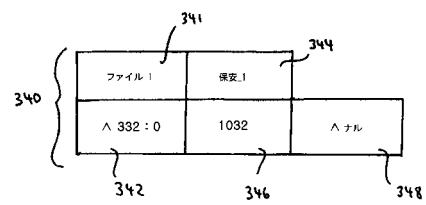
【図14】



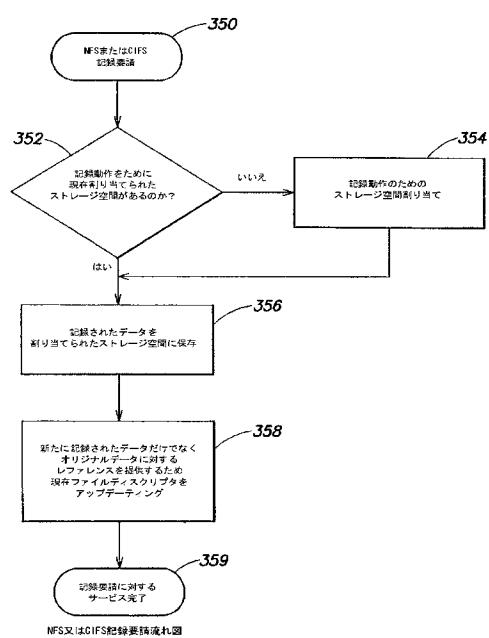
【図15】



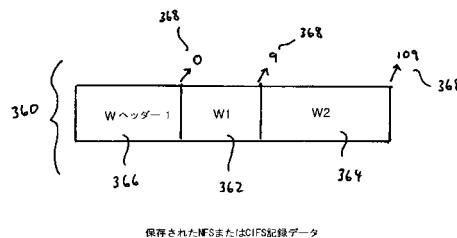
【図16】



【図17】

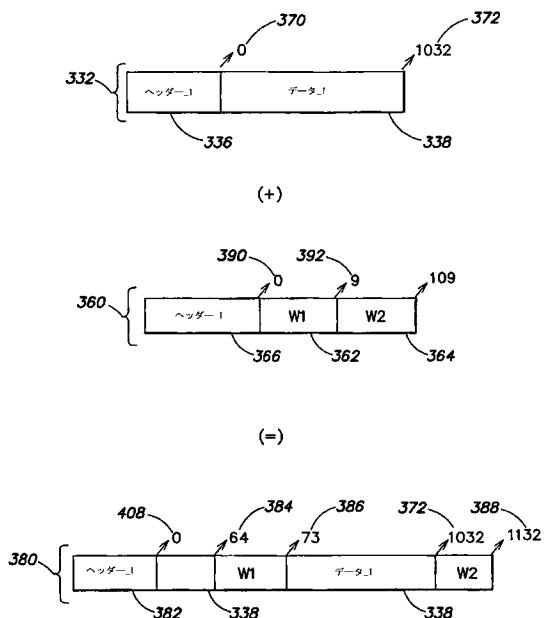


【図18】

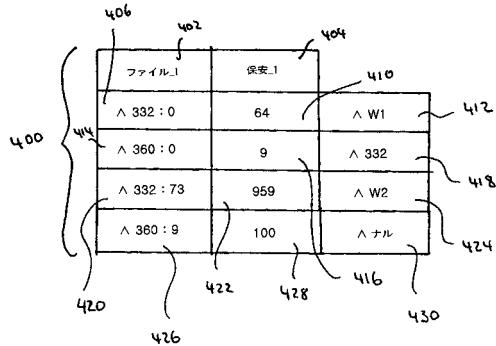


NFS又はCIFS記録要請流れ図

【図19】



【図20】



【手続補正書】

【提出日】平成17年4月15日(2005.4.15)

【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】

バックアップストレージシステムに保存された最も最近バックアップされた一つ以上のデータファイルに対応するデータファイルを、一つ以上含むデータボリュームをホストコンピュータ上にマウントイングする段階、及び

前記最も最近バックアップされたバージョンの一つ以上のデータファイルを維持する間、前記バックアップストレージシステムに保存された前記最も最近バックアップされたバージョンの一つ以上のデータファイルより更に最近の第2バージョンの一つ以上のデータファイルに対応するデータを前記バックアップストレージシステムに保存する段階を含むことを特徴とする方法。

【請求項2】

前記最も最近バックアップされたバージョンの一つ以上のデータファイルと、前記第2バージョンの一つ以上のデータファイルとをリンクする段階を更に含むことを特徴とする請求項1に記載の方法。

【請求項3】

前記最も最近バックアップされたバージョンの一つ以上のデータファイルと、前記第2バージョンの一つ以上のデータファイルとを識別するデータ構造を生成する段階を更に含むことを特徴とする請求項1に記載の方法。

【請求項 4】

前記第2バージョンの一つ以上のデータファイルは、前記最も最近バックアップされたバージョンの一つ以上のデータファイルの修正されたバージョンであることを特徴とする請求項3に記載の方法。

【請求項 5】

前記データボリュームをマウンティングする段階は、NFSマウンティングとCIFSマウンティングのうちいずれか一つの遂行を含むことを特徴とする請求項1に記載の方法。

【請求項 6】

前記データボリュームをマウンティングする段階は、前記最も最近バックアップされたバージョンの一つ以上のデータファイルに関連したメタデータを含むファイルディスクリプタの生成を含み、前記メタデータは、前記最も最近バックアップされたバージョンの一つ以上のデータファイルのバックアップストレージ媒体上のストレージ位置を識別する識別子を含むことを特徴とする請求項1に記載の方法。

【請求項 7】

バックアップデータセットを保存するためのバックアップストレージ媒体、及び、

請求項1の方法を具現する命令のセットを実行するよう構成された一つ以上のプロセッサを含む制御器を含むことを特徴とするバックアップストレージシステム。

【請求項 8】

前記バックアップデータセットは、総合フルバックアップデータセットであることを特徴とする請求項7に記載のバックアップストレージシステム。

【請求項 9】

一つ以上のプロセッサで実行時、請求項1の方法を具現する複数の命令でインコーディングされたことを特徴とするコンピュータ判読可能媒体。

【請求項 10】

前記プロセッサは、バックアップストレージシステムに含まれることを特徴とする請求項9に記載のコンピュータ判読可能媒体。

【請求項 11】

一つ以上のデータファイルを含むバックアップデータセットに対応するシステムファイルを独自的に識別する第1識別子、及び

バックアップデータセットの最新バージョンの一つ以上のデータファイルが保存されたストレージ媒体上の個々のストレージ位置を識別する一つ以上の第2識別子を含むデータ構造が保存されたことを特徴とするコンピュータ判読可能媒体。

【請求項 12】

前記バックアップデータセットは、バックアップされたバージョンの複数のデータファイルを含み、前記一つ以上の第2識別子は、前記バックアップデータセットの最新バージョンの複数のデータファイルそれぞれが保存された前記ストレージ媒体上のそれぞれのストレージ位置を独自的に識別する複数の第2識別子を含むことを特徴とする請求項11に記載のコンピュータ判読可能媒体。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0075

【補正方法】変更

【補正の内容】

【0075】

上記図12を再び参照すると、ボリュームマウンティングが要請された後、ボリューム復旧アプリケーションは、現在のフルバックアップデータセットが利用可能であるか否かを照会することができる（段階292）。利用ができない場合、ボリューム復旧アプリケーションは、総合フルバックアップアプリケーション240と通信して総合フルバックアップ処理を行ない（図2参照）、現在のバックアップデータセットを生成することができる（段階294）。ボリューム復旧アプリケーションは、正規のフルバックアップデータ

セット、または、総合フルバックアップデータセットをエクスポートинг (exporting) し、要請されたボリュームマウンティングをNFS、または、CIFSシェアによって行なうことができる。特に、ボリューム復旧アプリケーションは、論理的メタデータキャッシュ 2 4 2 を照会して、段階 2 9 0 で識別され選択されたフルバックアップボリュームを示す適切なメタデータを探す。

【手続補正3】

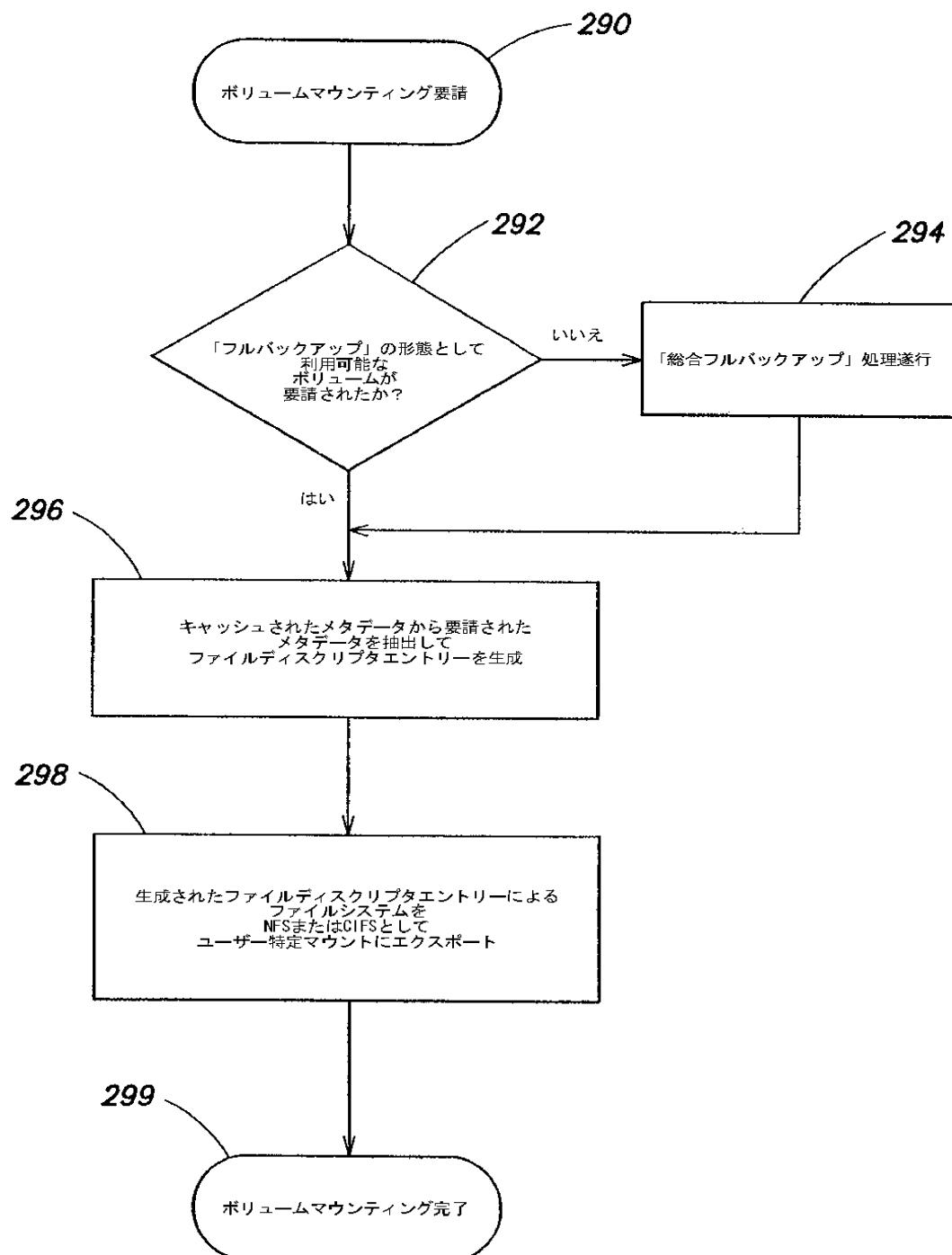
【補正対象書類名】図面

【補正対象項目名】図 1 2

【補正方法】変更

【補正の内容】

【図12】



【手続補正4】

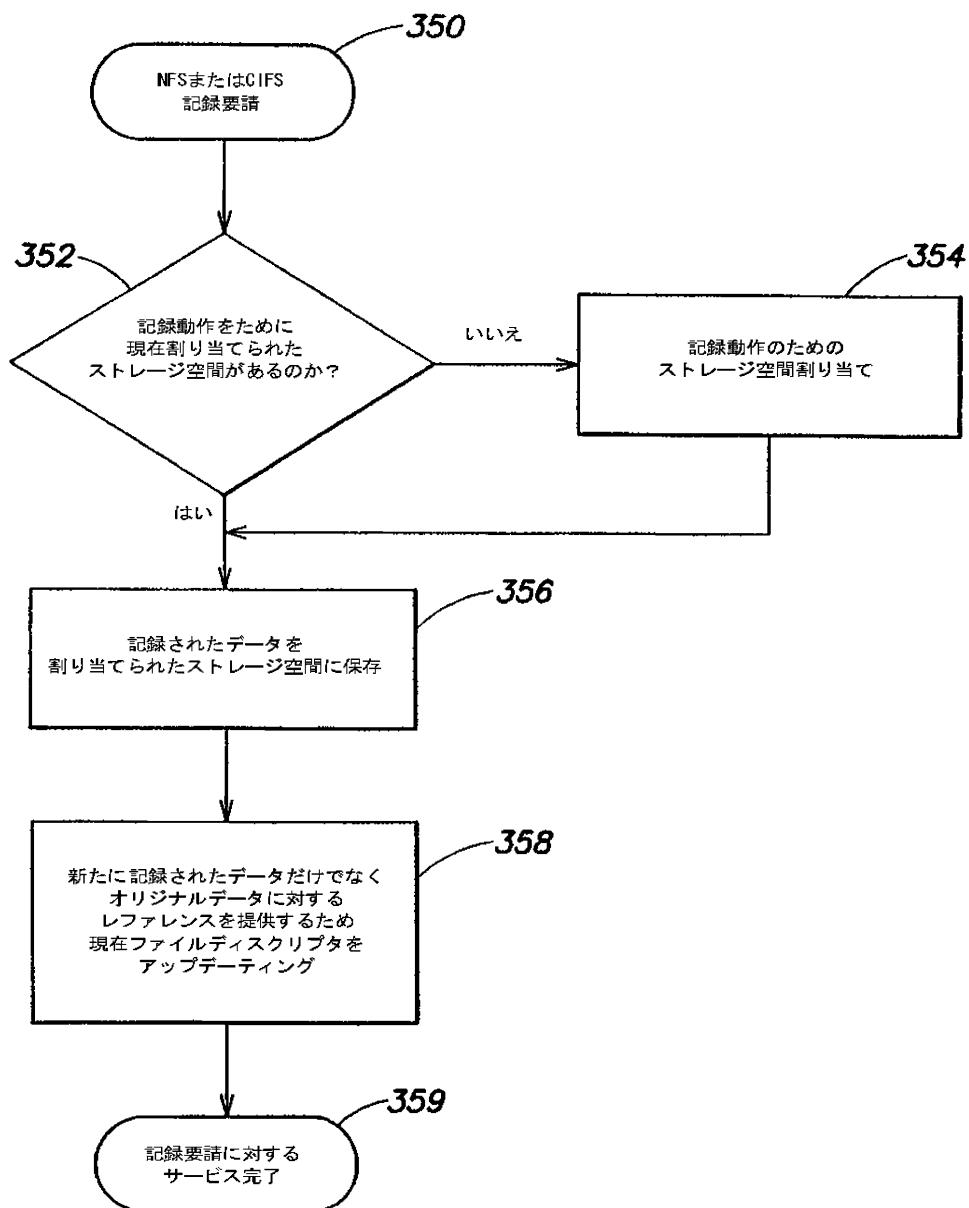
【補正対象書類名】図面

【補正対象項目名】図17

【補正方法】変更

【補正の内容】

【図17】



【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US04/32122												
A. CLASSIFICATION OF SUBJECT MATTER IPC(7) : G06F 12/00 US CL : 711/161,162,165; 707/202,203,204 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) U.S. : 711/161,162,165; 707/202,203,204														
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 practicable, search terms used) EAST SEARCH														
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">Category *</th> <th style="width: 80%;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="width: 10%;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>US 2003/0105912 A1 (NOREN) 05 June 2003 (05.06.2003), see abstract, col. 1, lines 20-27, col. 1, lines 34-55, col. 2, lines 44-62, col. 3, lines 15-45, col. 5, lines 15-19.</td> <td>1-4, 6-11 ----- 5</td> </tr> <tr> <td>---</td> <td></td> <td></td> </tr> <tr> <td>Y</td> <td>US 6,230,190 B1 (EDMONDS et al.) 08 May 2001 (08.05.2001), see col. 4, lines 30-40.</td> <td>5</td> </tr> </tbody> </table>			Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	X	US 2003/0105912 A1 (NOREN) 05 June 2003 (05.06.2003), see abstract, col. 1, lines 20-27, col. 1, lines 34-55, col. 2, lines 44-62, col. 3, lines 15-45, col. 5, lines 15-19.	1-4, 6-11 ----- 5	---			Y	US 6,230,190 B1 (EDMONDS et al.) 08 May 2001 (08.05.2001), see col. 4, lines 30-40.	5
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
X	US 2003/0105912 A1 (NOREN) 05 June 2003 (05.06.2003), see abstract, col. 1, lines 20-27, col. 1, lines 34-55, col. 2, lines 44-62, col. 3, lines 15-45, col. 5, lines 15-19.	1-4, 6-11 ----- 5												

Y	US 6,230,190 B1 (EDMONDS et al.) 08 May 2001 (08.05.2001), see col. 4, lines 30-40.	5												
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>		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 application or patent 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														
Date of the actual completion of the international search 30 January 2005 (30.01.2005)		Date of mailing of the international search report 14 FEB 2005												
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703) 305-3230		Authorized officer <i>Mitchell R. Edman</i> DONALD A SPARKS Telephone No. (571) 272-4201												

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,HU,IE,IT,LU,MC,NL,PL,PT,RO,SE,SI,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BW,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,M,A,MD,MG,MK,MN,MW,MX,MZ,NA,NI,NO,NZ,OM,PG,PH,PL,PT,RO,RU,SC,SD,SE,SG,SK,SL,SY,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,YU,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

1. イーサネット
2. J A V A
3. U N I X
4. フロッピー

F ターム(参考) 5B082 DE06 FA04 GA05