

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4278452号
(P4278452)

(45) 発行日 平成21年6月17日(2009.6.17)

(24) 登録日 平成21年3月19日(2009.3.19)

| | | | | | |
|-------------------|------------------|-----|------------|------|--|
| (51) Int.Cl. | | F I | | | |
| G06F 12/00 | (2006.01) | | G06F 12/00 | 531M | |
| G06F 3/06 | (2006.01) | | G06F 12/00 | 514E | |
| G06F 13/10 | (2006.01) | | G06F 12/00 | 545A | |
| | | | G06F 3/06 | 304F | |
| | | | G06F 13/10 | 340Z | |

請求項の数 4 (全 14 頁)

(21) 出願番号 特願2003-205616 (P2003-205616)
 (22) 出願日 平成15年8月4日(2003.8.4)
 (65) 公開番号 特開2005-55947 (P2005-55947A)
 (43) 公開日 平成17年3月3日(2005.3.3)
 審査請求日 平成17年12月19日(2005.12.19)

(73) 特許権者 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74) 代理人 100093861
 弁理士 大賀 真司
 (72) 発明者 北村 学
 神奈川県川崎市麻生区王禅寺1099番地
 株式会社日立製作所システム開発研究所
 内
 (72) 発明者 中谷 洋司
 神奈川県川崎市麻生区王禅寺1099番地
 株式会社日立製作所システム開発研究所
 内

最終頁に続く

(54) 【発明の名称】 計算機システム

(57) 【特許請求の範囲】

【請求項1】

ホストと、

前記ホストにスイッチを介して接続されて、前記ホストにファイルアクセスサービスを提供する複数のファイルサーバと、

それぞれに一意な識別子が設定された複数の記憶装置と、前記複数のファイルサーバに接続されて、前記複数のファイルサーバからのアクセスにตอบสนองして前記複数の記憶装置を制御する記憶制御装置を有する記憶装置システムとからなる計算機システムであって、

階層構造を形成する属性情報の集まりとして前記複数の記憶装置に存在し、前記複数のファイルサーバに管理される複数のファイルシステムと前記複数のファイルサーバとの対応関係を示す情報を格納するメモリと、

前記複数のファイルサーバのうち一つのファイルサーバに、バックアップ対象ディレクトリを通知するとともに、バックアップ開始を指示するバックアップサーバとを備え、前記複数のファイルサーバのうち前記バックアップ開始の指示を受けたファイルサーバは

前記メモリに格納された情報を基に、前記ファイルサーバのうち前記バックアップ対象ディレクトリを管理するファイルサーバをバックアップ対象ファイルサーバとして特定し、当該バックアップ対象ファイルサーバに前記バックアップ対象ディレクトリのバックアップを通知し、

前記バックアップの通知を受けた前記バックアップ対象ファイルサーバは、

10

20

前記記憶制御システムに、前記バックアップ対象ディレクトリに属するデータの複製を指示し、その後、前記複製されたデータを記憶する記憶装置の識別子を前記記憶制御システムから取得して、その識別子を前記バックアップ開始の指示を受けたファイルサーバに送信し、

前記バックアップ開始の指示を受けたファイルサーバは、前記バックアップ対象ファイルサーバから受信した識別子を基に前記複製されたデータが格納された記憶装置にパス設定を行い、前記メモリに格納された情報を基に、前記パス設定された記憶装置のデバイスファイルを抽出し、抽出されたデバイスファイルに属する対象ディレクトリのデータをバックアップデータとして前記バックアップサーバに送信してなる、計算機システム。

10

【請求項2】

前記バックアップ対象ファイルサーバは、バッファ上に未書き込みデータが存在するときは、前記バックアップ対象ディレクトリへのデータの書き込みを一時的に抑止し、前記未書き込みデータを全て前記記憶装置へ書き出してなる、請求項1記載の計算機システム。

【請求項3】

ホストと、前記ホストにスイッチを介して接続されて、前記ホストにファイルアクセスサービスを提供する複数のファイルサーバと、それぞれに一意な識別子が設定された複数の記憶装置と、前記複数のファイルサーバに接続されて、前記複数のファイルサーバからのアクセスに回答して前記複数の記憶装置を制御する記憶制御装置を有する記憶装置システムとからなる計算機システムにおいて、バックアップを行うためのバックアップ方法であって、

20

階層構造を形成する属性情報の集まりとして前記複数の記憶装置に存在し、前記複数のファイルサーバに管理される複数のファイルシステムと前記複数のファイルサーバとの対応関係を示す情報を格納するメモリと、

前記複数のファイルサーバのうち一つのファイルサーバに、バックアップ対象ディレクトリを通知するとともに、バックアップ開始を指示するバックアップサーバとを備え、前記複数のファイルサーバのうち前記バックアップ開始の指示を受けたファイルサーバは、前記メモリに格納された情報を基に、前記ファイルサーバのうち前記バックアップ対象ディレクトリを管理するファイルサーバをバックアップ対象ファイルサーバとして特定するステップと、

30

当該バックアップ対象ファイルサーバに前記バックアップ対象ディレクトリのバックアップを通知するステップを実行し、

前記バックアップの通知を受けた前記バックアップ対象ファイルサーバは、前記記憶制御システムに、前記バックアップ対象ディレクトリに属するデータの複製を指示するステップと、

その後、前記複製されたデータを記憶する記憶装置の識別子を前記記憶制御システムから取得して、その識別子を前記バックアップ開始の指示を受けたファイルサーバに送信するステップを実行し、

40

前記バックアップ開始の指示を受けたファイルサーバは、前記バックアップ対象ファイルサーバから受信した識別子を基に前記複製されたデータが格納された記憶装置にパス設定を行うステップと、前記メモリに格納された情報を基に、前記パス設定された記憶装置のデバイスファイルを抽出するステップと、前記抽出されたデバイスファイルに属する対象ディレクトリのデータをバックアップデータとして前記バックアップサーバに送信するステップを実行する、バックアップ方法。

【請求項4】

前記バックアップ対象ファイルサーバは、バッファ上に未書き込みデータが存在するときは、前記バックアップ対象ディレクトリ

50

へのデータの書き込みを一時的に抑止するステップと、前記未書き込みデータを全て前記記憶装置へ書き出すステップを実行する、バックアップ方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、計算機システムにおけるバックアップに関する。

【0002】

【従来の技術】

近年、データの増大に伴い、企業等で使用される計算機システムも大規模化している。しかし、記憶装置を増設すると、バックアップ業務も、複雑でコストがかかるものとなる。例えば、計算機毎にバックアップ装置を設置して、バックアップ処理を実施しなければならない。そのため、ホスト計算機（以下、ホストと略す）、記憶装置システム及びテープ装置を、SAN（Storage Area Network）と呼ばれるネットワークに接続して、バックアップ処理専用のサーバ（以下、バックアップサーバと称する）により全ての記憶装置のデータを一括してバックアップする方法などが考えられている。例えば、特許文献1には、バックアップサーバが、SANに接続された他のサーバが使用している記憶装置内のデータをテープ装置にバックアップする方法や、テープ装置を複数のホストで共用する形態が開示されている。以上により、通常の業務を実行しているサーバに負荷をかけずにバックアップを行い、計算機システムの構築コストを削減することができる。

【0003】

また、複数ホスト間を相互接続するLAN（Local Area Network）に接続されるNAS（Network Attached Storage）と呼ばれる装置が普及している。NASはNFS（Network File System）やCIFS（Common Internet File System）などのネットワークファイルアクセス用プロトコルによるデータアクセスをホストに提供する装置であり、ファイルを記憶する記憶部分と、ホストからのファイルアクセス要求を処理するNASヘッド部分（以下、ファイルサーバと称する）からなる。NASは、複数のホストからのデータ共有や設定作業が容易なことから、中～大規模サイトでも用いられるようになっており、処理能力や信頼性を増すためにファイルサーバを複数有する装置も表れている。

【0004】

商用のNASでは、高性能化・高信頼化のために、各社で独自のOS、独自のファイルシステムを用いる事が多く、通常のホストのように市販のバックアップソフトを導入する事が極めて困難である。そのため、バックアップは、NDMP（Network Data Management Protocol）を用いる方法が多く行われている。NDMPはNAS装置とバックアップソフトを有する外部の装置間でのバックアップ処理用プロトコルであり、NDMPに対応しているNAS装置は、NDMPに対応しているバックアップソフトを有する装置からバックアップやリストアの操作が可能である。すなわち、バックアップソフトを有する1つの装置から、各社のNASのバックアップ操作が行えることから、NDMPは、事実上の標準規格となっている。非特許文献1には、NDMPによるNASのバックアップ方式についていくつかの方法が開示されている。

【0005】

【特許文献1】

特開2002-7304号公報

【非特許文献1】

オライリー・ジャパン社「SAN & NAS ストレージネットワーク管理」 p.193～221

【0006】

【発明が解決しようとする課題】

NASのバックアップでは、記憶装置に格納されているデータをファイルサーバが読み出して、そのNAS固有のバックアップフォーマットを作成して、バックアップサーバないしNASに直接接続されたテープ装置に送る方式をとる。このため、バックアップ処理中はファイルサーバのCPUの負荷が高くなり、オンライン業務に影響が出る。また、複数のファイル

10

20

30

40

50

サーバを有する装置の場合、1台のNASであるにもかかわらず、外部、とくにバックアップサーバからは複数のNASとして認識されるため、管理が複雑になる。

【0007】

本発明の目的は、複数のファイルサーバが存在する環境において、外部には複数のファイルサーバを単一の装置に見せかけつつ、効率的なバックアップ処理を行うシステムを提供することにある。

【0008】

【課題を解決するための手段】

上記課題を解決するために、本発明では、ホストと、前記ホストにスイッチを介して接続されて、前記ホストにファイルアクセスサービスを提供する複数のファイルサーバと、それぞれに一意的識別子が設定された複数の記憶装置と、前記複数のファイルサーバに接続されて、前記複数のファイルサーバからのアクセスにตอบสนองして前記複数の記憶装置を制御する記憶制御装置を有する記憶装置システムとからなる計算機システムであって、階層構造を形成する属性情報の集まりとして前記複数の記憶装置に存在し、前記複数のファイルサーバに管理される複数のファイルシステムと前記複数のファイルサーバとの対応関係を示す情報を格納するメモリと、前記複数のファイルサーバのうち一つのファイルサーバに、バックアップ対象ディレクトリを通知するとともに、バックアップ開始を指示するバックアップサーバとを備え、前記複数のファイルサーバのうち前記バックアップ開始の指示を受けたファイルサーバは、前記メモリに格納された情報を基に、前記ファイルサーバのうち前記バックアップ対象ディレクトリを管理するファイルサーバをバックアップ対象ファイルサーバとして特定し、当該バックアップ対象ファイルサーバに前記バックアップ対象ディレクトリのバックアップを通知し、前記バックアップの通知を受けた前記バックアップ対象ファイルサーバは、前記記憶制御システムに、前記バックアップ対象ディレクトリに属するデータの複製を指示し、その後、前記複製されたデータを記憶する記憶装置の識別子を前記記憶制御システムから取得して、その識別子を前記バックアップ開始の指示を受けたファイルサーバに送信し、前記バックアップ開始の指示を受けたファイルサーバは、前記バックアップ対象ファイルサーバから受信した識別子を基に前記複製されたデータが格納された記憶装置にパス設定を行い、前記メモリに格納された情報を基に、前記パス設定された記憶装置のデバイスファイルを抽出し、抽出されたデバイスファイルに属する対象ディレクトリのデータをバックアップデータとして前記バックアップサーバに送信してなる計算機システムを構築する。

【0009】

前記バックアップ対象ファイルサーバは、バッファ上に未書き込みデータが存在するときには、前記バックアップ対象ディレクトリへのデータの書き込みを一時的に抑止し、前記未書き込みデータを全て前記記憶装置へ書き出してなる。

【0010】

【発明の実施の形態】

(第1の実施形態)

図1は、本発明の計算機システムの構成例を示す図である。尚、ソフトウェアは、楕円で囲んで示している。

【0011】

計算機システムは、ファイルサーバ1A、1B、1C、記憶装置システム2、スイッチ3、バックアップサーバ4、テープ装置5、ホスト6とから構成される。以下、ファイルサーバ1A、1B、1Cを代表するときは、ファイルサーバ1と記載する。尚、ファイルサーバ1と記憶装置システム2を統合してNAS装置としてもよく、NAS装置内の複数のファイルサーバを接続する内部スイッチが存在してもよい。

【0012】

ファイルサーバ1は、LAN I/F11、CPU12、メモリ13、FC I/F14を有し、LAN I/F11を介してスイッチ3に接続され、スイッチ3に接続されるホスト6に対して、NFSやCIFSなどによるファイルアクセスサービスを提供する。ファイルサーバ1Cは、LAN I/F11を介して、バック

10

20

30

40

50

アップサーバ4にも接続される。

【0013】

メモリ13には、デバイスドライバ101、ローカルファイルシステム102、ネットワークファイルシステム103、デバイスマネージャ104、連携プログラム105、バックアッププログラム106等のプログラムが格納される。デバイスドライバ101は、記憶装置システム2にアクセスするためのプログラムである。ローカルファイルシステム102は、後述のファイルシステムへアクセスするためのプログラムである。ネットワークファイルシステム103は、ホスト6から送信されるファイルアクセスプロトコル（NFS/CIFS等）に従ったアクセス要求を処理するためのプログラムである。デバイスマネージャ104は、記憶装置22のデータを他の記憶装置22に複製したり、各記憶装置22から情報の取得を行ったりするためのプログラムである。連携プログラム105は、バックアップを行う装置と連携しながらバックアップを行うためのプログラムである。バックアッププログラム106は、連携プログラム105を実行する他の装置と通信しながら、バックアップを一括して実施するためのプログラムである。

10

【0014】

尚、バックアップサーバ4は、スイッチ3、あるいは全てのファイルサーバ1に直接接続される形態でもよい。そうすれば、ファイルサーバ1Cに障害が起きたとき、ファイルサーバ1A、1Bのどちらかがファイルサーバ1Cに代わってバックアップする形態にできる。あるいは、1台のサーバをスタンバイサーバとしてアイドル状態にしておき、他のファイルサーバに障害が発生したときに、その処理を肩代わりする構成もある。このとき、スタンバイサーバでバックアップ処理を専門に担当させても通常業務に影響を与えない。

20

【0015】

本実施形態では、ファイルサーバ1Cが、ファイルサーバ1A、1Bに代わってバックアップするが、ファイルサーバ1Cに障害が起きたとき、ファイルサーバ1A、1Bのどちらかがファイルサーバ1Cに代わって、バックアップする形態もあり得るので、バックアッププログラム106を有する。同様に、ファイルサーバ1Cは、ファイルサーバ1Cのバックアップ機能に障害が起きたとき、ファイルサーバ1A、1Bのどちらかにバックアップを代わってもらう形態もあり得るので、連携プログラム105を有する。

【0016】

記憶装置システム2は、FC I/F20、記憶制御装置21、複数の記憶装置22、管理端末23とから構成され、FC I/F20を介してファイルサーバ1に接続される。記憶制御装置21は、記憶装置22の複製データを作成する機能（以下、ディスクミラー/スプリット機能（特許文献1に記載されているボリュームコピー機能に相当）と称する）を有する。管理端末23は、記憶装置システム2の外部にあってもよい。また、記憶装置22は、物理的な磁気ディスクドライブであってもよいし、ディスクアレイ装置のように、複数の磁気ディスクドライブをまとめて1つの論理的なディスクドライブに見せかける形態のものであってもよいし、光学的媒体などの他の媒体を利用する装置であってもよい。

30

【0017】

バックアップサーバ4は、ファイルサーバ1Cに接続される他、テープ装置5にも接続され、ファイルサーバ1Cから送信されたデータをテープ装置5へバックアップする。尚、ファイルサーバ1Cとテープ装置5を、FC I/Fを介して接続し、バックアップサーバ4を介さずにバックアップする形態もある。

40

【0018】

通常、各記憶装置22にアクセスするファイルサーバは一意に決まっているため、それぞれの記憶装置22はどれか一つのファイルサーバのみからアクセスできるように設定されるが、複数のファイルサーバから共通に、特定の記憶装置22にアクセス可能とすることもできる。以下、ファイルサーバから、特定の記憶装置22にアクセスできるように設定する処理を、パス設定と称する。

【0019】

図2は、記憶装置システム2の論理的な構成の一例を示す図である。論理的には、記憶装置

50

221、222はファイルサーバ1Aに、記憶装置223、224はファイルサーバ1Bに、記憶装置225は、ファイルサーバ1Cにパス設定されている。更に、未使用の記憶装置221'（プールボリューム）を保持している。必要に応じて記憶装置221'をファイルサーバ1にパス設定して複製データを格納する。記憶装置221'の管理は、ファイルサーバ1のどれかが行ってもよいし、管理端末23を介して管理者が行ってもよい。

【0020】

図3は、図2とは別の、記憶装置システム2の論理的な構成の一例を示す図である。接続関係は、図2と同様であるが、記憶装置221'～225'は、記憶制御装置21のディスクミラー/スプリット機能により、それぞれ記憶装置221～225の複製データが格納される。

【0021】

図4は、ファイルサーバ1で管理されるファイルシステムの階層ディレクトリ構造を示す図である。階層ディレクトリ構造とは、ルートディレクトリdir_a1を頂点にして、複数のファイルやディレクトリがルートディレクトリ下に存在する構造を指す。以下、階層ディレクトリ構造に属するファイル、ディレクトリ、そして、その構造を形成する属性情報の集まりをファイルシステムと呼ぶ。1つのファイルシステムは、複数の記憶装置に跨ることもあるし、1つの記憶装置に複数のファイルシステムが存在することもあるが、本実施形態では、簡単のため、1つのファイルシステムが1つの記憶装置に存在するものとする。

【0022】

図5は、ファイルサーバ1で管理されるファイルシステムの内容を示す図である。ファイルサーバ1は、図2（図3）の記憶装置221、222にそれぞれ1つのファイルシステムを管理しており、それぞれにusr1、usr2という名前を付けている。CPU12は、ローカルファイルシステム102を実行して、ディレクトリイメージ112のように、ルートディレクトリ"/"の配下に、usr1、usr2の2つのディレクトリがあるものとして管理している。これは一般のファイルサーバで管理される通常のファイルシステムの形態である。

【0023】

次に、デバイスマネージャ104の実行により提供される3つの機能について説明する。

【0024】

第一の機能は、パス設定を行う機能である。パス設定は、CPU12あるいは管理端末23から記憶装置システム2に対する指示をもとに行われる。記憶装置22のそれぞれには、記憶装置システム2内で一意な識別子（例えば、図2の記憶装置221には221番という識別子がついている）がついており、パス設定を行う際、CPU12や管理端末23からは識別子が指定される。図2では、初期状態として、ファイルサーバ1Aは記憶装置221、222に、ファイルサーバ1Bは記憶装置223、224に、ファイルサーバ1Cは記憶装置225に、パス設定が為されている。尚、識別子は、物理的な記憶装置に付されてもよいし、論理的な記憶装置に付されてもよい。

【0025】

第二の機能は、記憶制御装置21が有するディスクミラー/スプリット機能を起動して、記憶装置22のうち特定の1台に格納されるデータを、別の記憶装置22に複製する機能である。例えば、CPU12は、コピー元とコピー先の記憶装置の識別子を指定して、複製データ作成指示が出た時点の記憶装置221のデータ（スナップショットイメージ）を記憶装置221'に複製する。

【0026】

第三の機能は、各記憶装置22から、その記憶装置固有の情報取得を行う機能である。図6は、記憶装置221がファイルサーバ1A、1Cに対してパス設定されている場合、ファイルサーバ1A、1Cが記憶装置221にアクセスする時に使用する情報について示す図である。以下、説明の都合上、一部、プログラムを主体として記載することもあるが、実際は、CPU12が行う主体である。

【0027】

CPU12がローカルファイルシステム102を実行して記憶装置221内のファイルシステムにアクセスする場合、記憶装置221に対するアクセス要求は、デバイスドライバ101を実行する

10

20

30

40

50

ことで処理される。このとき、デバイスドライバ101は、ローカルファイルシステム102から、デバイスファイル名と呼ばれる一種のファイル名を受け取ることで、アクセス対象の記憶装置を特定する。図6の例では、記憶装置221内には、"/dev/sd1c"というデバイスファイル名が対応付けられているので、ローカルファイルシステム102は、デバイスドライバ101が実行される際にデバイスファイル名"/dev/sd1c"をパラメータとして渡すことで、アクセス対象の記憶装置221を特定する。デバイスファイル名とファイルシステム名を対応付けた情報は、ファイルサーバ1が有する設定ファイルである/etc/fstabなどに格納されている。この設定ファイルは、メモリ13に格納されてもよいし、実データを格納する記憶装置とは別の設定ファイルを格納するための記憶装置に格納されてもよい。

【0028】

デバイスファイル名は、ファイルサーバ1A、1Cのそれぞれで管理され、その名前付けも独立に行われるため（この名前付けについて明確な規則はない）、ファイルサーバ1Aは記憶装置221に対して"/dev/sd1c"を、ファイルサーバ1Cは"/dev/sd2c"を対応付ける、ということが生じる。また、ローカルファイルシステム102は、この"/dev/sd1c"を指定してアクセス要求を作成し、当該アクセス要求を、デバイスドライバ101に渡す。デバイスドライバ101では、渡されたデバイスファイル名を、ポートIDと論理ユニット番号(LUN)に変換し、ポートIDとLUNをパラメータとして含んだアクセス命令を記憶装置システム2に発行することで、記憶装置221にアクセスする。ただし、ポートIDは、初期化段階で物理インタフェース毎に動的に割り当てられるものであり（ユーザや管理者によって自由に行うことができない）、LUNは、パス設定の際に管理者が管理端末23を介して、あるいは、CPU12がデバイスマネージャ104を実行して決定されるが、ファイルサーバ1Aからアクセスするときの記憶装置221とファイルサーバ1Cからアクセスするときの記憶装置221に対して、必ず同じLUNが割り当てられる保証はない。したがって、ファイルサーバ1Aからアクセスしていた記憶装置221をファイルサーバ1Cからアクセスさせたい場合、ファイルサーバ1Aで使用していた"/dev/sd1c"をファイルサーバ1Cに教えても、そのデバイスファイル名で記憶装置221にアクセスできる保証は無い。

【0029】

そこで、本実施形態では、各記憶装置が一意的な識別子を持っていることを利用して、記憶装置を特定する。例えば、ディレクトリ"/usr1"の属する記憶装置を特定する場合、CPU12は、デバイスマネージャ104を用いて、記憶装置221の識別子を取得する指示を出す。この時、デバイスマネージャ104は、デバイスドライバ101に対してデバイスファイル名"/dev/sd1c"を指定して指示を出す。デバイスドライバ101は、"/dev/sd1c"を、ポートID、LUNの組の情報に変換し、当該ポートID、LUNの組に対応する記憶装置の識別子を取得する指示を記憶装置システム2へと送信する。記憶装置システム2はそれに応答して、記憶装置221の識別子"221"を、ファイルサーバ1Aへ送信する。識別子を送信する方法の一例として、記憶装置システム2においてSCSIのINQUIRYコマンドやMODE SENSEコマンドに対する応答情報の一部に、記憶装置の識別子を入れて送信する方法が考えられる。そして、ファイルサーバ1Aは、当該識別子をファイルサーバ1Cに送信する。一方、ファイルサーバ1Cで識別子"221"が付与された記憶装置221を特定するには、デバイスマネージャ104を実行して、ファイルサーバ1Cの有するデバイスファイル全てに対して識別子の取得を実施し、ファイルサーバ1Aから受信した識別子"221"を返却するデバイスファイルを用いて、記憶装置221にアクセスすればよい。

【0030】

尚、デバイスファイル名とファイルシステムとの対応付けは、ユーザによって付与されたLUNをもとに、それぞれのファイルサーバが行う。また、ファイルサーバ1それぞれが、ファイルシステムと記憶装置の識別子を対応付けている場合、ファイルサーバ1Aは、記憶装置システム2に識別子を問い合わせる必要はなく、ファイルサーバ1Cに対して、識別子を送信する。

【0031】

続いて、バックアップ処理の流れについて、図7を用いて説明する。ここでは、図2の形態

10

20

30

40

50

において、ファイルサーバ1Cが、ファイルサーバ1Aが管理するディレクトリ/usr1以下をバックアップする指示を受け取ったとして説明する。

【0032】

まず、ユーザあるいは管理者が、バックアップサーバ4に対して、バックアップしたいファイルが属するディレクトリ名を指定することで、処理が開始される。通常、バックアップサーバ4は、バックアップ対象のファイルやディレクトリを管理しているファイルサーバに対してバックアップを指示するが、本実施形態では、ファイルサーバ1Cに対して行う。ファイルサーバ1Cは、外部に対して自身の有するファイルシステムが図10のようなディレクトリ構造を持つように見せかけている。よって、バックアップサーバ4は、ディレクトリ名 "/server1/usr1" を指定する。

10

【0033】

次に、バックアップサーバ4は、ファイルサーバ1Cに対して、当該ディレクトリ名を通知して、バックアップ開始を指示する(ステップ71)。ファイルサーバ1Cは、当該指示を受け取ると、バックアッププログラム106を実行して、ディレクトリ名 "/server1/usr1" から、バックアップ対象ディレクトリがファイルサーバ1Aで管理されているものであると特定する(ステップ91)。次に、ファイルサーバ1Cは、ファイルサーバ1Aに対して、"/usr1" のバックアップを行う旨を通知する(ステップ92)。

【0034】

次に、ファイルサーバ1Aは、当該通知を受けると、ローカルファイルシステム102を実行して、ネットワークファイルシステム103の実行による/usr1への書き込みを一時的に抑止し、バッファ上に記憶装置221への未書き込みデータがある場合には、当該データを全て記憶装置221へと書き出す(ステップ81)。尚、バッファは、メモリ13と同じであってもよいし、メモリ13とは別に、ファイルサーバ内に存在してもよい。次に、記憶装置221'から未使用の記憶装置を選択して、選択した記憶装置に対してパス設定を行う(ステップ82)。続いて、デバイスマネージャ104を実行して、記憶装置システム2に対し、記憶装置221のデータを221'に複製する指示を出し、記憶装置システム2では、ディスクミラー/スプリット機能を用いて、ファイルサーバ1Aから指示のあった時点の記憶装置221のデータ(スナップショットイメージ)を記憶装置221'に複製する(ステップ83)。次に、ファイルサーバ1Aは、ステップ81における/usr1への書き込み抑止を解除し、ネットワークファイルシステム103を実行して、通常のアクセスを許可する(ステップ84)。

20

30

【0035】

ステップ81で、/usr1への書き込みを抑止し、バッファ内データを記憶装置221に書き出すのは次の理由による。まずバッファ内データを書き出す理由は、ファイルサーバ1Cから記憶装置221'の内容を吸い上げてバックアップするため、ファイルサーバ1Aのバッファ内に未書き込みデータが残っていると、ファイルサーバ1Cからファイルサーバ1Aのバッファのデータは読み出せないからである(記憶装置221'へのデータ書き込みは、ディスクミラー/スプリット機能によってなされるため、ファイルサーバ1Aは記憶装置221'の書き込みについては意識しない)。次に、/usr1への書き込みを抑止する理由は、バッファ内データを記憶装置221に書き出している途中で新たなデータが入ってきた際、バックアップ指示のあった時点よりも新しいデータ内容が記憶装置に書き込まれることを避けるため、あるいは、例えば記憶装置221と222のそれぞれに格納されるファイルシステムについて同時にバックアップしたい場合(複数のファイルシステムについて同時にバックアップする場合)、記憶装置221'、222'に書き込まれるデータを、双方とも同一時刻のものとしなければならないためである。

40

【0036】

ファイルサーバ1Aは、デバイスマネージャ104を実行して、複製データが格納されている記憶装置221'の識別子を記憶装置システム2から取得し(ステップ85)、その識別子をファイルサーバ1Cに送信する(ステップ86)。

【0037】

ファイルサーバ1Cは、ファイルサーバ1Aから送信された識別子を基に、記憶装置221'にバ

50

ス設定を行い(ステップ93)、記憶装置221'はファイルサーバ1Cではどのデバイスファイル名に関連付けられているか検索する(ステップ94)。検索方法は、前述の図6で説明した通りである。当該デバイスファイルが見つかり、当該ファイルを使用して対象となる記憶装置221'をマウントし、対象ディレクトリのデータを吸い上げる(ステップ95)。バックアップ処理は、具体的にはバックアップサーバ4から指定されるが、一般的にはtarやdumpコマンドを実行することで実現される。次に、tarやdumpコマンドなどで作成されたバックアップデータをバックアップサーバ4へ送信する(ステップ96)。すべてのバックアップデータの送信が完了すると、当該記憶装置221'をアンマウントして(ステップ97)、当該記憶装置のパス設定を解除する(ステップ98)。

【0038】

その後、バックアップサーバ4は、ファイルサーバ1Cから受信したデータを、テープ装置5へバックアップする(ステップ72)。

【0039】

尚、図3のように記憶装置221'などのスナップショットデータを格納する記憶装置が予めパス設定されている構成では、ステップ82、ステップ93、ステップ98は不要となる。

【0040】

以上では、ファイルサーバ1Aのデータのみをバックアップする例を説明したが、例えばバックアップ対象のディレクトリを"/{server1/usr1}、/{server2/user2}"と指定することで、ファイルサーバ1Aの"/usr1と、1Bの"/usr2"のデータのバックアップを同時に行うこともできる。この場合、図9のステップ92では、ファイルサーバ1A、1Bに対して同時にバックアップの指示を出し、データの複製も同時に行う。バックアップデータをテープ装置5にバックアップする処理は、テープ装置5が1台しか存在しない場合には、ファイルサーバ1Aのディレクトリ/usr1のデータ、ファイルサーバ1Bのディレクトリ/usr2のデータを順にテープ装置5に書き出すことになり(あるいは逆)、必ずしも同時にバックアップできるわけではないが、データ複製処理はファイルサーバ1Cからファイルサーバ1A、1Bに同時に指示されているため、いずれも同一時点の複製データを残すことができる。テープ装置5が複数台存在すれば、バックアップデータのテープ装置5への書き込みは同時に実施できる。また、ルートディレクトリ"/"を指定した場合、ファイルサーバ1のすべてのデータ(図10のルートディレクトリ以下のデータ)をバックアップすることができる。そのため、本発明では、ファイルサーバが何台あっても、あたかも1台のファイルサーバだけが存在するかのようにバックアップの操作を実施できる。

【0041】

以上のように、本発明では、特に負荷のかかる記憶装置221'のデータの読み出し、バックアップサーバ4へのデータ送信を、ファイルサーバ1Cが代行するため、ファイルサーバ1A、1Bや、ファイルサーバ1A、1Bとホスト6を結ぶネットワークに殆ど負荷をかけずにバックアップ処理を行う事ができる。また、ファイルサーバが何台あっても、バックアップを実行する際にはファイルサーバ1Cに指示を出すだけなので、ユーザからはファイルサーバ(NAS)が1台だけ存在しているかのようにバックアップ処理を行うことができる。

(第2の実施形態)

図8は、第2の実施形態における計算機システムの構成例を示す図である。この計算機システムは、統合型記憶装置システム800、1台以上のホスト6、バックアップサーバ4がスイッチ3で相互に接続され、またサーバ7、バックアップサーバ4、テープ装置5そして統合型記憶装置システム800とがファイバチャネルスイッチ(図中ではFCスイッチと略記)8で相互接続される構成をとる。

【0042】

統合型記憶装置システム800は、LANアダプタ810、FCアダプタ820、記憶装置アダプタ830、記憶装置22、内部スイッチ825、キャッシュ/制御メモリ835とで構成される。LANアダプタ810は第一の実施形態におけるファイルサーバ1A、1B相当の機能を持つ。FCアダプタ820は、サーバ7やバックアップサーバ4からのリード・ライトコマンドを処理するためのI/O処理プログラム823と、統合型記憶装置システム800内のボリュームのデータを外部記憶装

10

20

30

40

50

置やテープ装置5などに複製したり、外部記憶装置やテープ装置5のデータを統合型記憶装置システム800内のボリュームへ複製するためのボリュームコピープログラム824を有する。記憶装置アダプタ830は、複数台の記憶装置22へのリード・ライト制御、また記憶装置22を複数台まとめて1台以上の論理記憶装置としてLAN I/F811やFC I/F821に見せる機能、ディスクミラー/スプリット機能などを持つ。キャッシュ/制御メモリ835は、サーバ7やLANアダプタ810からのアクセスを高速化するために、記憶装置22のデータや、LANアダプタ810、FCアダプタ820、記憶装置アダプタ830間の通信に使うデータの一部を格納する。LANアダプタ810、FCアダプタ820、記憶装置アダプタ830とも、図中では1つしか示していないが、複数存在しても良い。内部スイッチ825は、LANアダプタ810、FCアダプタ820、記憶装置アダプタ830を相互に接続する。内部スイッチ825は、クロスバースイッチや、共通バスで各アダプタを相互接続するものであってもよい。

10

【 0 0 4 3 】

サーバ7にはデバイスドライバ861、ファイルシステム862、エージェント863、デバイスマネージャ104が格納される。バックアップサーバ4には、デバイスドライバ861、デバイスマネージャ864、そしてバックアッププログラム106が格納される。

【 0 0 4 4 】

図9は、統合型記憶装置システム800の論理的な構成を示す図である。記憶装置アダプタ830は、記憶装置22の何台かをまとめて、あるいは記憶装置22の1台を分割して、NAS用ボリューム840、ミラーボリューム841、FC用ボリューム842を形成する。記憶装置22と各ボリュームが1:1に対応する形態もある。LANアダプタ810やサーバ7からは、記憶装置アダプタ830の存在は意識されない(図9では、記憶装置アダプタは点線で示している)。NAS用ボリューム840、ミラーボリューム841、FC用ボリューム842とも図中では1つずつしか示していないが、複数存在してもよい。NAS用ボリューム841はLANアダプタ810からアクセスできるようにバス設定されており、ミラーボリューム842はディスクミラー/スプリット機能によりNAS用ボリューム840ないしFC用ボリューム842の複製データが格納される。FC用ボリューム842は、FCアダプタ820を介して、サーバ7やバックアップサーバ4からアクセスされるよう定義される。またミラーボリューム841は、図2のように必要に応じてバス設定されてもよいし、図3のように予めNAS用ボリュームの複製データを格納するためにバス設定されていてもよいが、ここでは図2と同じ形態であるとして説明する。

20

【 0 0 4 5 】

第二の実施形態では、ボリュームイメージを直接テープ装置5へとバックアップする、ボリュームレベルのバックアップを行う。バックアッププログラム106を実行すれば、LANアダプタ810の持つボリュームとサーバ7の持つボリュームの両方のバックアップが可能である。ユーザはバックアップ処理を行う際、バックアップ対象のデータを管理するサーバ7ないしLANアダプタ810のホスト名(あるいはIPアドレス)とデバイスファイル名を指定する。バックアップサーバ4は、その指定情報を元にサーバ7ないしLANアダプタ810と通信しながらバックアップ処理を行う。

30

【 0 0 4 6 】

第二の実施形態におけるバックアップ処理は、図7の処理とほぼ同等であり、以下では図7との相違点のみを説明する。まず、ユーザからバックアップデータを管理する装置としてサーバ7やLANアダプタが指定されるため、ステップ91に相当する処理はない。バックアップサーバ4は、コピー元としてミラーボリューム841を指定、コピー先としてテープ装置5を指定して、ボリュームコピープログラム824を実行することで、テープ装置5へのバックアップを実施する。また、ステップ85では、複製データのボリュームの識別子を取得し、ステップ86では、当該ボリュームの識別子をバックアップサーバ4に送信する。ステップ93以降では、ボリュームレベルのバックアップを実施するため、ファイルシステムのマウント処理、アンマウント処理に相当するステップ95、97がない。

40

【 0 0 4 7 】

以上のように、第二の実施形態では、バックアップサーバ4においてLANアダプタ810が使用するNAS用ボリューム840とサーバ7で使用するFC用ボリューム842の両方のバックアップ

50

処理を一括して行うことができる。

【0048】

尚、本発明で説明したプログラムは、CD-ROM等の記憶媒体から移してもよいし、ネットワーク経由で他の装置からダウンロードしてもよい。

【0049】

【発明の効果】

本発明によれば、複数のファイルサーバが存在する環境において、外部には複数のファイルサーバを単一の装置に見せかけつつ、効率的なバックアップ処理を行うシステムを提供することができる。

【図面の簡単な説明】

【図1】 計算機システムの構成例を示す図。

【図2】 記憶装置システムの論理的な構成の一例を示す図。

【図3】 図2とは別の、記憶装置システムの論理的な構成の一例を示す図。

【図4】 ファイルシステムの階層ディレクトリ構造を示す図。

【図5】 ファイルシステムのディレクトリイメージを示す図。

【図6】 ファイルサーバ1が記憶装置221にアクセスする時の情報の流れについて示す図。

【図7】 バックアップの処理フロー。

【図8】 図1とは別の計算機システムの構成例を示す図。

【図9】 図8の記憶装置システムの論理的な構成の一例を示す図。

【図10】 ファイルサーバ1を統合したファイルシステムのディレクトリ構造を示す図。

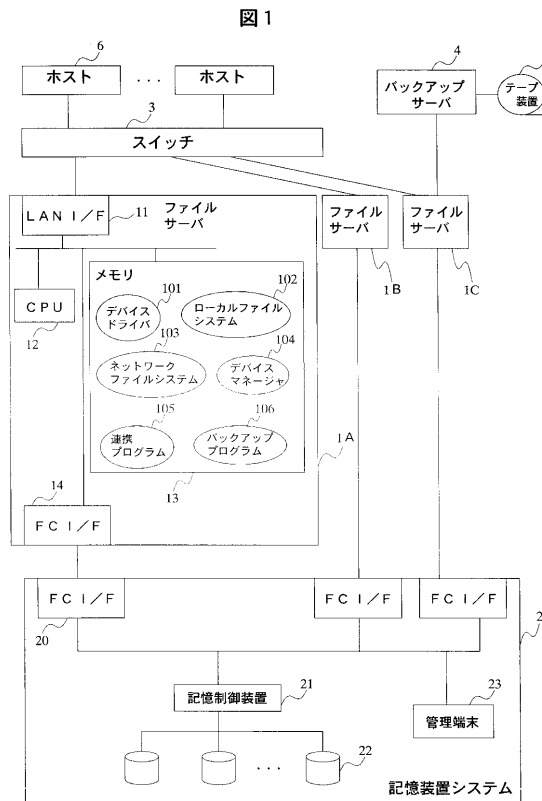
【符号の説明】

1 ... ファイルサーバ、 2 ... 記憶装置システム、 3 ... スイッチ、 4 ... バックアップサーバ、
5 ... テープ装置、 6 ... ホスト、 11 ... LAN I/F、 21 ... 記憶制御装置、 22 ... 記憶装置、
20 ... FC I/F、 23 ... 管理端末

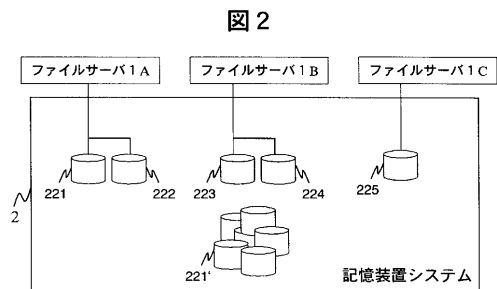
10

20

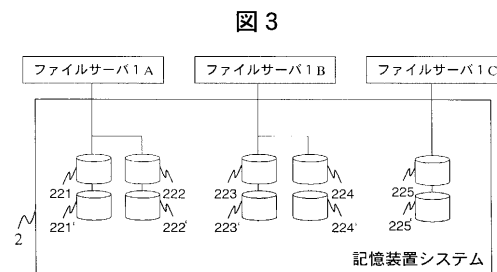
【図1】



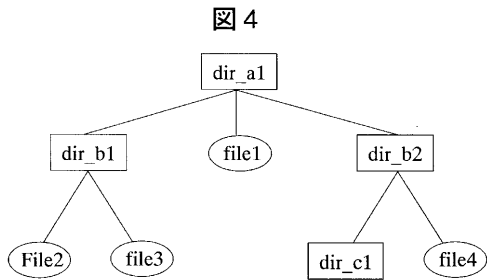
【図2】



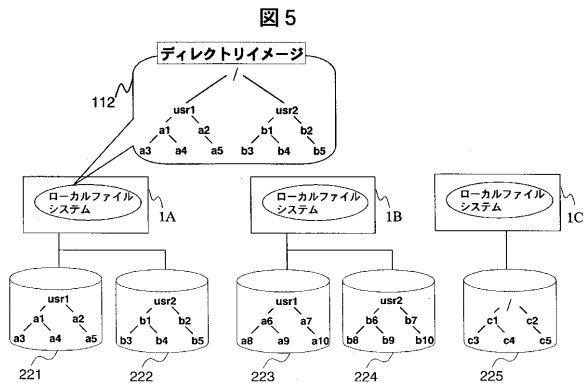
【図3】



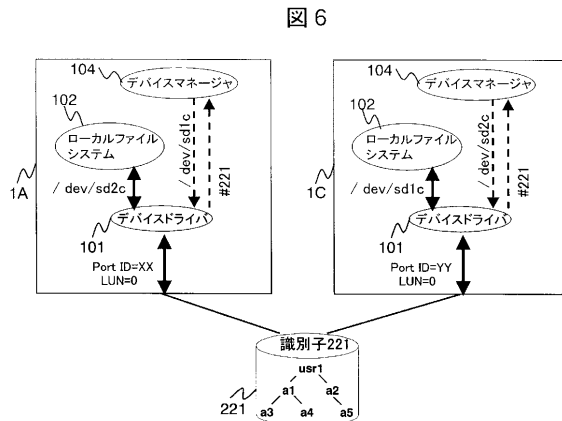
【 図 4 】



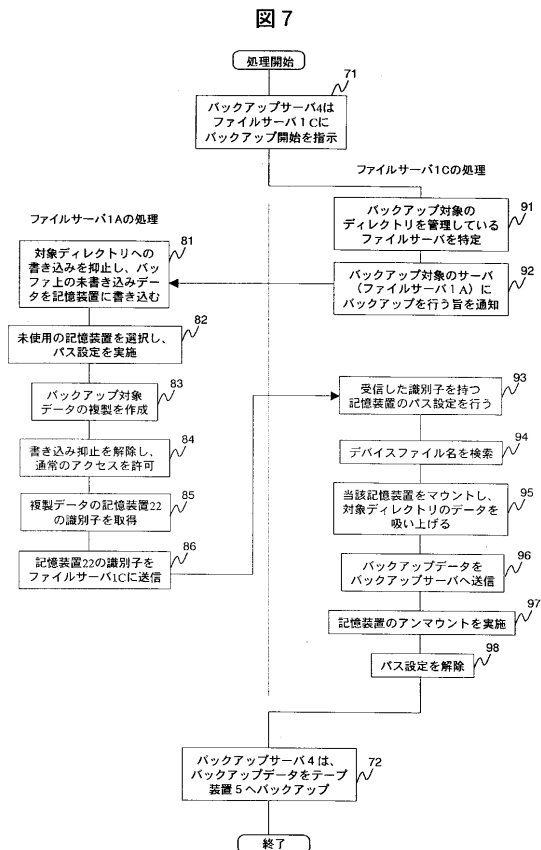
【 図 5 】



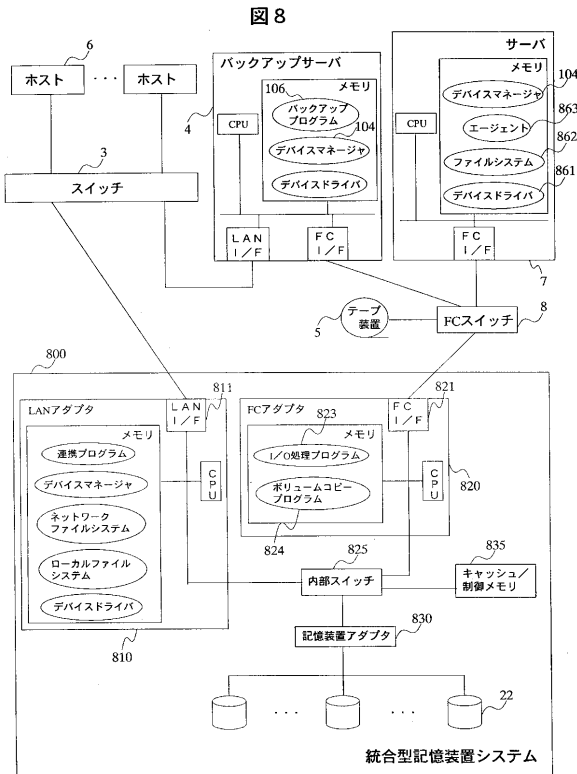
【 図 6 】



【 図 7 】

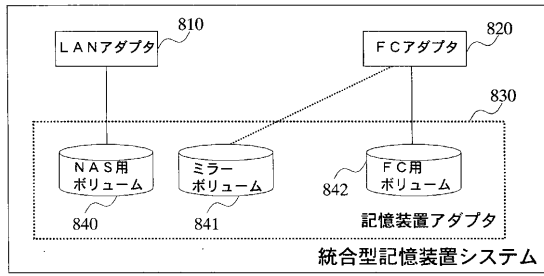


【 図 8 】



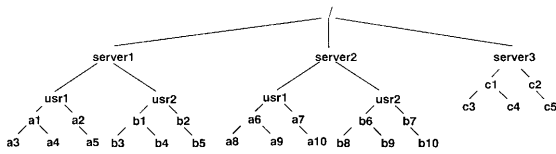
【図9】

図9



【図10】

図10



フロントページの続き

(72)発明者 中野 隆裕

神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

審査官 田川 泰宏

(56)参考文献 特開2002-297456(JP,A)

特開2002-032253(JP,A)

特開2002-312251(JP,A)

特開2003-196136(JP,A)

特開2003-173279(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00

G06F 3/06

G06F 13/10