

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4181044号
(P4181044)

(45) 発行日 平成20年11月12日(2008.11.12)

(24) 登録日 平成20年9月5日(2008.9.5)

(51) Int. Cl. F I
G06F 12/00 (2006.01) G O 6 F 12/00 5 3 1 J
G06F 3/06 (2006.01) G O 6 F 3/06 3 0 4 F

請求項の数 14 (全 30 頁)

| | | | |
|---------------|-------------------------------|-----------|---|
| (21) 出願番号 | 特願2003-550066 (P2003-550066) | (73) 特許権者 | 503093224 イーエムシー コーポレーション EMC CORPORATION アメリカ合衆国 01748 マサチュー セッツ州 ホプキントン サウス ストリ ート 176 |
| (86) (22) 出願日 | 平成14年11月15日(2002.11.15) | (74) 代理人 | 100077481 弁理士 谷 義一 |
| (65) 公表番号 | 特表2005-512191 (P2005-512191A) | (74) 代理人 | 100088915 弁理士 阿部 和夫 |
| (43) 公表日 | 平成17年4月28日(2005.4.28) | (72) 発明者 | リチャード エス. オーラン アメリカ合衆国 89015 ネバダ州 ヘンダーソン アローヘッド トレイル 520 ナンバー 823 |
| (86) 国際出願番号 | PCT/US2002/037309 | | |
| (87) 国際公開番号 | W02003/048941 | | |
| (87) 国際公開日 | 平成15年6月12日(2003.6.12) | | |
| 審査請求日 | 平成16年7月1日(2004.7.1) | | |
| (31) 優先権主張番号 | 09/997,643 | | |
| (32) 優先日 | 平成13年11月29日(2001.11.29) | | |
| (33) 優先権主張国 | 米国 (US) | | |

最終頁に続く

(54) 【発明の名称】 大容量記憶システムの選択されたデータのスナップショットを保存するためのバックアップ方法

(57) 【特許請求の範囲】

【請求項1】

複数のデータブロックでデータを格納し、該データのスナップショットコピーを含むデータ記憶場所にアクセスを有する大容量記憶デバイスを備えたコンピュータシステムにおける、バックアップされるように指定されているデータブロックを前記スナップショットコピーにバックアップする方法であって、

前記大容量記憶デバイス上のデータブロックに対して、前記スナップショットコピーが元データと同期していた時点から新しいデータが書き込まれた元データにおける記憶場所を特定することによって元データに対する変更を追跡するステップと、

指定の期間中に前記大容量記憶デバイスにおいて変更されたデータブロックを特定するステップと、

前記特定されたデータブロック中のバックアップされるように指定されたデータブロックのみを指定し、前記特定されたデータブロック中のバックアップされるように指定されていない他のデータブロックはバックアップされないステップであって、前記指定されたデータブロックが、バックアップされるべき前記データブロックとして指定される前記大容量記憶デバイスの前記データブロックのマップを保持するステップを有し、

前記指定されたデータブロックのコピーを、前記スナップショットコピーを含む前記データ記憶場所に伝送して、前記データブロックの前記伝送されたコピーが前記スナップショットコピーに含められることを可能にするステップと

を有することを特徴とするバックアップ方法。

【請求項 2】

前記変更されたデータブロックを特定するステップは、前記指定の期間中に変更された少なくとも前記データブロックに関するエントリを含むテーブルを保持するステップを有することを特徴とする請求項 1 に記載のバックアップ方法。

【請求項 3】

前記データブロックを指定するステップが、バックアップされない前記他のデータブロックを特定することによって実行されることを特徴とする請求項 1 に記載のバックアップ方法。

【請求項 4】

バックアップされるべき前記データブロックを特定する入力をユーザから受け取るステップを有することを特徴とする請求項 1 に記載のバックアップ方法。

10

【請求項 5】

バックアップされない前記データブロックを特定する入力をユーザから受け取るステップを有することを特徴とする請求項 1 に記載のバックアップ方法。

【請求項 6】

前記変更されたデータブロックを特定するステップが、前記大容量記憶デバイスに対するユーザアクセスの中断なしに実行されることを特徴とする請求項 1 に記載のバックアップ方法。

【請求項 7】

前記スナップショットコピーを含む前記データ記憶場所が、前記大容量記憶デバイスから遠隔に配置されることを特徴とする請求項 1 に記載のバックアップ方法。

20

【請求項 8】

前記スナップショットコピーを含む前記データ記憶場所が、前記大容量記憶デバイス内に含まれることを特徴とする請求項 1 に記載のバックアップ方法。

【請求項 9】

前記指定されたデータブロックのコピーを伝送するステップは、前記期間中に前記指定されたデータブロックの最新の変更を表す前記指定されたデータブロックのコピーだけを伝送して、前記指定されたデータブロックの中間の変更が全く前記スナップショットコピーを含む前記データ記憶場所に伝送されないようにするステップを有することを特徴とする請求項 1 に記載のバックアップ方法。

30

【請求項 10】

データブロックを格納する大容量記憶デバイスを備えたコンピュータシステムにおける、選択されたデータブロックのスナップショットコピーの作成方法であって、

前記スナップショットコピーに含まれる大容量記憶デバイスのデータブロックのサブセットをユーザから指定された指定サブセットの情報として受け取るステップであって、前記スナップショットコピーは、最初の時点において存在するような指定サブセットを保存し、

前記最初の時点の後に変化する大容量記憶デバイスにおけるデータブロックとして、前記大容量記憶デバイスで変化する指定サブセットのデータブロックを特定するステップであって、前記大容量記憶デバイスにおいて、最初の時点の後の変化した指定サブセットの少なくともデータブロックの登録を行なうテーブルを維持し、

40

前記指定サブセットのデータブロックのコピーを保存するステップであって、変化されたデータブロックのコピーが、変化する前の指定サブセットのデータブロックの元のコピーを表すものであり、

前記データブロックの指定サブセットのスナップショットコピーにアクセスするステップと

を有することを特徴とするスナップショットコピーの作成方法。

【請求項 11】

前記スナップショットコピーが、コンピュータシステムにユーザアクセスを中断することなく作成されるもので、ユーザは、スナップショットコピーが作成されるように前記大

50

容量記憶デバイスにI/O要求を連続して行なうことを特徴とする請求項10に記載のスナップショットコピーの作成方法。

【請求項12】

前記スナップショットコピーを保存するステップは、最初の時点に存在する指定サブセットとして、前記データブロックの指定サブセットのバックアップとして、前記スナップショットコピーを保存することを特徴とする請求項10に記載のスナップショットコピーの作成方法。

【請求項13】

前記最初の時点の後に前記大容量記憶デバイスにおいてデータロスを生じた後で、前記スナップショットコピーを用いたデータの指定サブセットを記憶することを特徴とする請求項10に記載のスナップショットコピーの作成方法。

10

【請求項14】

前記スナップショットコピーは、
変化した指定サブセットのデータブロックに対する、その変化したデータブロックの保存コピーと、

前記最初の時点の後に変化しないデータブロックの指定サブセットのデータブロックの元のコピーと

を含むことを特徴とする請求項10に記載のスナップショットコピーの作成方法。

【発明の詳細な説明】

【技術分野】

20

【0001】

本発明は、大容量記憶システムの選択されたデータのスナップショットを保存するためのバックアップシステム及びその方法に関し、より詳細には、1つまたは複数の大容量記憶システム上のいくつかのセクタだけのスナップショットコピーをとるためのバックアップシステム及びその方法に関し、特に、コンピュータデータの保護技術に関する。

【背景技術】

【0002】

コンピュータは、ほとんどのビジネス業務の不可欠な一部分となっている。一部の 경우에는、コンピュータは、あまりにも不可欠となっているため、コンピュータが機能を停止した場合、ビジネス業務を行うことができない。銀行、保険会社、証券会社、財務サービスプロバイダ、及び様々な他のビジネスが、常に変更される情報を格納し、操作し、表示するのにコンピュータネットワークに依拠している。取引の成功または失敗は、正確であるとともに最新である情報の利用可能性にかかっている可能性がある。サービスプロバイダの信用、または存立そのものが、コンピュータネットワーク上に保持される情報の信頼性に依存する可能性がある。世界中のビジネスが、自らのデータの商品価値を認識しており、データを信頼できる形でバックアップすることにより自らのコンピュータネットワーク上に格納された情報を保護する信頼できる費用対効果が大きいやり方を求めている。

30

【0003】

しばしば、スワップファイルなどのファイル、プリンタバッファ、空きセクタテーブル、及び「一時」ファイルが、通常、バックアップされることが所望されていないのに、バックアップされる。不要なファイルをバックアップすることにより、処理サイクル、通信帯域幅、及びバックアップメモリ容量の浪費がもたらされる。不要な情報をバックアップすることを回避するため、バックアップのために個々のファイルを選択することができるシステムが開発されている。それらのシステムは、(1)すべてのファイルシステムコマンドを代行受信するソフトウェアレイヤを追加すること、(2)2つのタイプのファイル、すなわち、追跡ファイル及び非追跡ファイルを定義することによって機能する。追跡ファイルでは、システムは、ファイルのバックアップバージョンと更新バージョンの両方を提供する。バックアップバージョンは、バックアップが開始された時点で存在していたファイルであり、更新バージョンは、バックアップが開始された後に行われたあらゆる変更を含むファイルの現在のバージョンである。非追跡ファイルの場合、現在の(つまり更新

40

50

)バージョンだけが利用可能である。基本的に、システムは、あたかもファイルシステムコマンドを代行受信するソフトウェアレイヤが存在するかのように機能する。

【0004】

ソフトウェアレイヤは、バックアップが開始された時点で存在していた元のファイルデータがファイル変更によって上書きされないようにファイル変更を格納することによってファイルのバックアップバージョン及び更新バージョンを提供するものである。すべてのファイルシステムコマンドを代行受信することにより、システムは、アクセスを要求するプログラムにファイルの適切なバージョンを提供することができる。例えば、バックアッププログラムが読取り要求を行った場合、ソフトウェアレイヤは、要求されたデータのバックアップバージョンを提供する。読取り要求を行う他のプログラムは、ファイルデータの更新バージョンを受け取る。

10

【0005】

ソフトウェアレイヤがすべてのファイルシステムコマンドを代行受信することは、いくつかのファイルだけが追跡される場合に十分な解決策である。しかし、この解決策は、追跡ファイルの数が増加すると実用的でないことが明らかになる。問題は、ソフトウェアレイヤが、基本的に、ファイルシステムの仕事を実行することである。追跡ファイルに関して、オペレーティングシステムによって実行される各ファイル操作が、何らかの形で、ソフトウェアレイヤによっても実行される。ファイルの数が増加すると、ソフトウェアレイヤは、過負荷になり、パフォーマンスが低下してシステムが使用できなくなる。

【0006】

20

また、ソフトウェア代行受信レイヤは、ファイル間に存在する可能性がある関係を見落とす。上述したとおり、ファイルの中に格納されたデータに整合性があるだけでは十分でない。1つのファイルの中に格納されたデータは、他の1つまたは複数のファイルの中に格納されたデータに関連している可能性が高い。従来技術のソフトウェアレイヤは、バックアッププロセス中にファイルがアクセス可能であることを確実にすることだけしかできない。従来技術のソフトウェアレイヤには、バックアップ操作を構成するすべてのファイルにわたって論理的に整合性のあるデータセットを確実にするための備えがない。したがって、上述したタイプのソフトウェアレイヤを使用して行われるバックアップは、格納されたデータの不整合に起因してそれほど有益でない可能性がある。

【0007】

30

したがって、ファイル間の関係を保持しながら、バックアップされることが所望されるデータだけをバックアップするための効率的なシステムを得ることは、当技術分野において重要なものである。

【0008】

【特許文献1】米国特許第5649152号明細書

【発明の開示】

【発明が解決しようとする課題】

【0009】

従来技術における上述した問題は、本発明が克服することに成功しており、本発明は、保護されることが所望されるデータブロックだけに関して元のデータをそのデータのスナップショットコピーにバックアップするためのシステム及び方法を対象とする。

40

【0010】

本発明は、このような状況に鑑みてなされたもので、その目的とするところは、大容量記憶システムの選択されたデータのスナップショットを保存するためのバックアップシステム及びその方法を提供することにある。

【0011】

つまり、本発明のバックアップシステム及びその方法は、従来技術に勝る4つの大きな利点を提供する。第1に、変更されており、バックアップが所望されるファイルとして指定された1次大容量記憶デバイスのデータブロックだけをバックアップすることにより、バックアップを行う必要があるデータの量を縮減する。第2に、バックアップすることが

50

所望されるデータだけをバックアップすることを介してバックアップ用のデータの量を極限まで縮減しているため、記憶領域のより効率的な使用を可能にする。第3に、バックアップが行われる際に1次記憶デバイスが論理的に整合性のある状態にあることを確実にすることにより、バックアップの精度を高める。第4に、バックアップを行うのに必要とされるデータが極限まで縮減されており、論理的に整合性のある状態のバックアップだけが行われるので、バックアップの頻度を高くすることができる。

【課題を解決するための手段】

【0012】

本発明のバックアップ方法は、元のデータとそのデータのスナップショットコピーが、少なくともバックアップに指定されたデータブロックに関して、全く同じデータを含むという前提で開始される。これは、例えば、従来のバックアップ技術、または従来のディスクミラーリング技術を使用して元のデータをスナップショットコピーに完全にコピーすることによって達することができる。元のデータとスナップショットコピーが同一のデータを含むと、本発明は、変更されたすべてのデータブロックをリストするためのマップまたは他のデータ構造を作成し、1次大容量記憶デバイス上のデータブロックに対して行われた変更を追跡し、変更されたデータブロックを特定し、バックアップが所望されないデータブロックから分けてバックアップが所望されるデータブロックを指定する。追跡は、スナップショットコピーが元のデータと同期していた時点から新しいデータが書き込まれた元のデータにおける記憶場所を特定することによって行われる。元のデータに対して行われた変更の特定により、バックアップ記憶デバイスを1次大容量記憶デバイスと同期させるためにスナップショットコピーに対して行う必要がある変更が示される。バックアップ記憶デバイスに対して行う必要がある変更は、リスト上またはテーブル上に登録される。

【0013】

本発明のバックアップシステムは、リストまたはテーブルを識別し、バックアップが所望される情報とバックアップが所望されない情報に分離することを可能にする。この分離は、所望される情報にフラグを設定すること、または所望されない情報にフラグを設定することによって達することができる。テーブル内の情報の識別及び分離により、バックアップ用の情報の量が、所望される情報だけに縮減され、バックアップされる情報の量を縮減することによってバックアッププロセスの速度が向上し、記憶スペースがより効率的に使用される。さらに、識別及び分離により、所望されない情報がバックアップに含められることが防止される。

【0014】

元のデータに対して行う必要がある変更が特定されると、スナップショットコピーにその変更が送られる。すると、スナップショットコピーには、バックアップ記憶デバイスを1次大容量記憶デバイスと同期させるすべてのデータが利用可能である。バックアッププロセス中に元のデータを保存するため、元のデータの静的スナップショットをとることができる。この静的スナップショットは、元のデータに対して行われており、スナップショットコピーに転送される必要がある変更をキャプチャする。バックアップをユーザにトランスペアレントにするため、静的スナップショットは、大容量記憶デバイスに対するユーザアクセスが中断されない形でとられることが好ましい。

【0015】

本発明は、静的スナップショットがいつとられるべきかを決定するため、いつ元のデータが論理的に整合性のある状態にあるかを特定する機構を含む。論理的に整合性のある状態を特定し、その時点までに行われた変更の静的スナップショットを取ることにより、変更がスナップショットコピーに転送された際、スナップショットコピーが論理的に整合性のある状態をキャプチャすることが保証される。連続する論理的に整合性のある状態のスナップショットをキャプチャすることにより、スナップショットコピーは、次々に論理的に整合性のある状態をキャプチャすることができる。このようにして、スナップショットコピーが必要になった場合、スナップショットコピーは、論理的に整合性のある状態にある。スナップショットコピーは、ある論理的に整合性のある状態から他の論理的に整合性

10

20

30

40

50

のある状態に移り、従来技術の諸問題の1つが解消される。

【0016】

本発明は、大容量記憶システムのバックアップに関してデータブロック手法をとり、保護されるように指定されたデータブロックだけがバックアップされるため、本発明は、バックアップを行う際に転送される必要があるデータの量を可能な極限まで最小化する。例えば、大型データベースが変更された5つのレコードを有する場合、従来技術のバックアップシステムは、その大型データベース全体をコピーする。しかし、本発明は、変更された5つのレコードだけをコピーする。データの量が最小限に抑えられるので、本発明は、リモートサイト(remote site)に配置されたバックアップシステムにデータをバックアップするのに特によく適している。本発明は、低い帯域幅の通信リンクを利用してバックアップデータをリモートバックアップサイトに転送することができる。例として、多くの場合、56.6kボーのモデムを備えた従来のダイヤルアップ電話回線で全く十分である。

10

【0017】

本発明のさらなる利点は、以下に説明するように、その説明から明白となり、あるいは本発明の実施によって知ることができる。本発明の利点は、特許請求の範囲で特に指摘する手段(instruments)及び組み合わせを使用して実現し、得ることができる。本発明の以上の特徴及びその他の特徴は、以下の説明及び特許請求の範囲からより十分に明白となり、あるいは以下に提示する本発明の実施によって知ることができる。

【0018】

本発明の上述した利点及び特徴、ならびにその他の利点及び特徴を得ることができるやり方を説明するため、以上に簡単に説明した本発明のより詳細な説明を添付の図面に示した本発明の特定の実施形態に関連して行う。それらの図面は、本発明の典型的な実施形態だけを示しており、したがって、本発明の技術的範囲を限定するものではなく、添付の図面を介して本発明をより具体的に詳細に説明するものである。

20

【発明を実施するための最良の形態】

【0019】

以下、図面を参照して本発明の実施の態様について説明する。

本発明は、大容量記憶手段の中のすべてのデータのスナップショットをとるのではなく、大容量記憶手段からバックアップされることが所望されるデータセクションだけのスナップショットをとるためのバックアップシステムとその方法の両方を技術的範囲に含んでいる。所望されるデータだけのスナップショットがとられるので、本発明は、大容量記憶手段上にあるデータのバックアップコピーを提供する際に時間と記憶スペースをともに最適化する。

30

【0020】

本発明を本発明のバックアップシステム及びその方法を実現するいくつかの実施形態の構造または処理を示す図を使用することによって説明する。本発明を提示するのにこのように図を使用しているが、本発明の技術的範囲を限定しているものではない。本発明は、汎用コンピュータまたは専用コンピュータを使用して実施することができ、すべてのそのようなコンピュータシステムが、本発明の技術的範囲に含まれるべきである。

40

【0021】

また、本発明の技術的範囲内の実施形態は、コンピュータ実行可能命令またはデータ構造がエンコードされているコンピュータ読取可能なメディアも含んでいる。そのようなコンピュータ読取可能なメディアは、汎用コンピュータまたは専用コンピュータがアクセスすることができる任意の利用可能なメディアであることが可能である。例として、限定としてではなく、そのようなコンピュータ読取可能なメディアは、RAM、ROM、EEPROM、CD-ROMまたは他の光ディスクストレージ、磁気ディスクストレージまたは他の磁気記憶デバイス、光磁気記憶デバイス、または所望のコンピュータ実行可能命令及びデータ構造を格納するのに使用することができ、汎用コンピュータまたは専用コンピュータがアクセスすることができる任意の他のメディアを含むことが可能である。また、以

50

上の組み合わせも、コンピュータ読取可能なメディアの範囲に含まれなければならない。同様に、コンピュータ実行可能命令またはデータ構造をデコードし、実行する間、コンピュータ実行可能命令またはデータ構造を格納するCPUまたは他の処理ユニットのレジスタも、コンピュータ読取可能なメディアの範囲に含まれる。コンピュータ実行可能命令は、例えば、汎用コンピュータまたは専用コンピュータがある機能または一群の機能を実行するようにさせる実行可能命令または実行可能データを含んでいる。

【0022】

「データブロック」という用語は、大容量記憶手段に対して書込みまたは読取りが行われるデータのブロックを表すのに使用する。「データブロック」という用語は、広く解釈されるものとし、あらゆるサイズまたは形式のデータが含まれなければならない。例えば、ディスク上の個別セクタの中に格納されたデータがデータブロックと適切に呼ばれる。セクタのグループまたはクラスタの中に格納されたデータ量も、データブロックと適切に呼ぶことができる。大容量記憶手段がRAM、あるいは語またはバイトでアドレス指定可能な記憶デバイスである場合、データブロックという用語は、データのバイト、語、または複数語単位に適用することができる。さらに、「所望されるデータブロック」という用語は、バックアップされるように指定されたデータブロックを表すのに使用し、「所望されないデータブロック」という用語は、バックアップされるように指定されていないデータブロックを表すのに使用する。

【0023】

図1は、本発明のバックアップシステムを説明するためのブロック図で、本発明の適切な動作環境のシステムレベルブロック図である。全体的に示されたシステム10は、いずれのタイプのネットワーク化されたまたはスタンドアロンのコンピュータシステムであってもよいコンピュータシステム12を備えている。例えば、コンピュータシステム12は、コンピュータネットワーク18のようなコンピュータネットワークに接続されたネットワークサーバコンピュータであることが可能である。また、コンピュータシステム12は、スタンドアロンシステムであってもよい。

【0024】

コンピュータシステム12には、複数の記憶場所に複数のデータブロックを格納するための大容量記憶手段が接続されている。記憶場所のそれぞれは、固有アドレスまたは他の機構によって指定される。大容量記憶手段は、データブロックを格納する任意の記憶機構であることが可能である。例えば、そのような大容量記憶手段は、1つまたは複数の磁気ディスクドライブまたは光磁気ディスクドライブを含むことが可能である。図1では、例えば、そのような大容量記憶手段を大容量記憶デバイス20で例示している。

【0025】

大容量記憶デバイス20は、バックアップされることが所望されるデータブロックも、バックアップされることが所望されないデータブロックも含む元のデータ14を備えている。バックアップされることが所望されないデータブロックの例は、スワップファイル、空きセクタテーブル、印刷バッファ、「.tmp」拡張子を有する一時ファイル、及びバックアップされることが所望されないその他のファイルである。

【0026】

大容量記憶デバイス20は、特定の時点で存在していたとおりバックアップされることが所望される元のデータの中のデータブロックのスナップショットコピーも含むことが可能である。したがって、「スナップショット」コピーとは、コピーが、ある時点で存在していたとおり所望のデータブロックをキャプチャしていることを示す。スナップショットコピー16が元のデータと同じ大容量記憶デバイス内に含まれるデータ記憶場所にあるのを示しているが、スナップショットコピー16は、異なる記憶デバイスのデータ記憶場所に配置されることも可能である。一部のケースでは、コンピュータシステム12は、スナップショットデータをコンピュータネットワーク18のような通信メディアを介して異なる記憶デバイスに書き込む。ただし、この実施形態では、スナップショットコピー16は、元のデータ14と同じ大容量記憶デバイス20に格納される。

【 0 0 2 7 】

この実施形態は、バックアッププロセス中、所望されるデータブロックに対応する大容量記憶デバイスのすべてまたは一部のスナップショットコピーを使用する。したがって、所望されるデータブロックに関して特定の時点における大容量記憶手段の静的スナップショットを作成するように前記大容量記憶手段のデータブロックを一時的に格納するための保存メモリ手段を備えている。そのような保存メモリ手段は、RAM、EEPROM、磁気ディスクストレージなどの任意のタイプの書込み可能記憶デバイスを含むことが可能である。また、そのような保存メモリ手段は、大容量記憶デバイス20の一部も含むことが可能である。図1では、そのような保存メモリ手段を、例えば、スナップショット記憶デバイス22で示している。保存メモリ手段については、以下により詳細に説明する。

10

【 0 0 2 8 】

コンピュータシステム12は、いずれのタイプの汎用コンピュータまたは専用コンピュータであってもよく、コンピュータシステム12は、汎用コンピュータまたは専用コンピュータを構成する任意の他のハードウェアも含むことが可能である。例えば、コンピュータシステム12は、プログラマブルコード手段を実行するためのプロセッサ手段も含むことが可能である。プロセッサ手段は、マイクロプロセッサまたは他のCPUデバイスであることが可能である。プロセッサ手段は、デジタル信号プロセッサなどの様々な専用プロセッサも含むことが可能である。コンピュータシステム12は、ユーザに出力を表示するための表示手段、コンピュータシステム12にデータを入力するための入力手段、ハードコピーのプリントアウトを出力するための出力手段、RAM、ROM、EEPROMなどのメモリ手段のような他の従来のコンピュータコンポーネントも含むことが可能である。

20

【 0 0 2 9 】

図2は、本発明のバックアップ方法を説明するためのタイミングを示す図で、図1の元のデータ14のような元のデータを図1のスナップショットコピー16のようなスナップショットコピーにバックアップするのに使用される方法の概要を示している。図2に示したバックアップ方法は、所望されるデータブロックに関する限り、元のデータ14とスナップショットコピー16は最新であるものと想定する。ここで、「最新」とは、スナップショットコピーが、元のデータ14のすべての所望されるデータブロックの最新コピーを含むことを意味する。図2で、スナップショットコピー16は、時刻 T_0 における元のデータ14の最新コピーを有するものと想定する。

30

【 0 0 3 0 】

図2に示したバックアップ方法は、時刻 T_0 から開始して、元のデータ14に対してスナップショットコピー16を最新状態に保つ。そして、連続する論理的に整合性のある状態をキャプチャする。これにより、1つの論理的に整合性のある状態から次の論理的に整合性のある状態に移るスナップショットコピー16がもたらされる。つまり、スナップショットコピー16が、連続する論理的に整合性のある状態をキャプチャすることが可能になる。これにより、バックアップ記憶デバイスを論理的に不整合な状態のままにする可能性がある従来技術のシステムに勝る著しい利点が生じる。バックアップデバイスが論理的に整合性のある状態にあることを確実にすることにより、本発明は、使用可能なスナップショットコピーが常に利用可能であることを確実にする。

40

【 0 0 3 1 】

次に、図2に戻ると、時刻 T_0 から開始して、所望されるデータブロックに対応する元のデータ14に対する変更が追跡される。これを図2にブロック28で示している。変更は、好ましくは、時刻 T_0 から開始して新しいデータが書き込まれており、所望されるデータブロックである大容量記憶デバイスのデータブロックを特定することによって追跡される。これは、時刻 T_0 から開始して新しいデータが書き込まれたデータブロックを特定するマップを保持すること、及び所望されるデータブロックのマップを保持することによって行うことができる。

【 0 0 3 2 】

何らかの時点で、元のデータ14の中の所望されるデータブロックに対して行われた変

50

更をキャプチャし、その変更をスナップショットコピー 16 に転送することが望ましい。好ましい実施形態では、システムは、1次大容量記憶デバイスの論理的に整合性のある状態を識別し、時刻 T_0 以来、変更された少なくとも所望されるデータブロックの静的スナップショットをとる。図2では、論理的に整合性のある状態は、時刻 T_1 として識別され、スナップショットがとられる。

【0033】

静的スナップショットは、特定の時点に存在していたとおりデータを保存して、スナップショット時刻の後に元のデータに対して変更が行われていても、スナップショット時刻に存在していたとおりの状態で、所望されるデータブロックがその特定の時点の後に利用可能であるようにする。そのような静的スナップショットを作成する多くの方法が存在する。いずれのそのような方法も本発明で機能するが、一部の方法が、様々な利点のために他の方法より優先される。どのように静的スナップショットがとられるかの詳細、及び静的スナップショットを作成するための好ましい方法を以下に示す。ただし、この概略では、静的スナップショットを作成するいずれの方法も本発明で使用できることを理解することが重要である。ただし、静的スナップショットは、大容量記憶デバイスに対するユーザの読取りアクセスまたは書込みアクセスを終了させることなく、とられることが好ましい。

10

【0034】

時刻 T_0 の直後、または時刻 T_0 後にコンピューティングリソースが利用可能になった時刻に、バックアップが所望されるデータブロックが、マップまたは他のデータ構造を使用して指定される。バックアップが所望されるデータブロックは、所望されるデータブロックを特定することによって直接に指定すること、またはバックアップ操作が所望されないデータブロックを指定することによって暗黙に示されることが可能である。それらのデータブロックは、ユーザがバックアップされるべき、またはバックアップされるべきでないファイルまたはファイルタイプを特定したことに応答して特定されること、または指定されることが可能である。次に、ファイルシステムを使用してそれらのファイルを特定のデータブロックにマップすることができる。上述した技術は、バックアップされるべきデータブロックを指定するのに役立つ可能性があるが、本発明は、バックアップされるべきデータブロックを特定し、指定するための他の技術を使用して実施することもできる。バックアップ操作が所望されるデータブロックを指定するプロセスは、例えば、図2の期間29の間に行われる。

20

30

【0035】

時刻 T_1 に、時刻 T_0 と時刻 T_1 の間に特定された所望されるデータブロックに対する変更が、その変更をスナップショットコピー 16 に送信することによってバックアップされる。これを図2に矢印30及びブロック32で示している。変更は、時刻 T_0 と時刻 T_1 の間に新しいデータが書き込まれた記憶場所だけに格納され、保護されることが指定された所望のデータブロックに関する元のデータ14のデータブロックを送信することによってスナップショットコピー 16 に送信される。

【0036】

データは、時刻 T_1 にスナップショットによって保存されるので、時刻 T_1 後に新しいデータが大容量記憶デバイスに書き込まれていても、バックアップ記憶デバイスに転送するのに利用可能である。時刻 T_0 と時刻 T_1 の間にいずれの記憶場所にデータが書き込まれたかを追跡するのに使用され、バックアップされることが所望されるデータブロックを指定するのに使用されたマップまたは他の機構が、バックアップ記憶デバイスに転送されるべきデータを特定するのに使用される。したがって、所望されるデータブロックに対する増分の変更だけが送信され、ファイル全体が変更されていない限り、ファイル全体が転送されることはない。さらに、所望されないデータブロックは、そのデータブロックに対する変更が行われていても送信されない。

40

【0037】

時刻 T_1 の直後、または時刻 T_1 後にコンピューティングリソースが利用可能になっ

50

た時刻に、バックアップが所望されるデータブロックが、期間 29 に関連して上述したやり方でマップまたは他のデータ構造を使用して、期間 33 中に指定される。この他に、バックアップされるように以前に指定されている同じデータブロックが、または等価のこととして、バックアップされないように以前に指定されている同じデータブロックが、新しいスナップショットに持ち越されることが可能である。この代替手法では、ユーザは、バックアップされるべきデータブロックを繰り返し指定することを要求されない。上述したとおり以前の指定が新しいスナップショットに持ち越されるかどうかを決める要因には、スナップショットの頻度、ユーザの選好、及び以前のスナップショット以来、ファイル構造が変更されているかどうかが含まれる。

【0038】

バックアップが実行されている最中に新しいデータが時刻 T_1 後に元のデータに書き込まれる可能性があるため、時刻 T_1 後に別のバックアップが行われる場合、ある機構を使用して時刻 T_1 後に行われた変更が特定される。図 2 では、ブロック 34 で示すとおり、時刻 T_1 後の変更が追跡される。これにより、時刻 T_1 後の所望されるデータブロックに対する変更も、スナップショットコピー 16 を何らかの後の時点まで更新するためにスナップショットコピー 16 に転送することが可能になる。

【0039】

図 2 に示すとおり、前述したシーケンスは、時刻 T_2 に繰り返される。これを矢印 36、期間 37、ブロック 38、及びブロック 40 で示している。上述したとおり、時刻 T_2 にとられるスナップショットは、時刻 T_1 と T_2 の間に所望されるデータブロックに対して行われた変更がスナップショットコピー 16 に転送された際、スナップショットコピー 16 が、時刻 T_2 における論理的に整合性のある状態まで更新されるように、論理的に整合性のある状態を表していなければならない。

【0040】

以上のように、本発明は、前回のバックアップ以来、変更された記憶場所に格納されたデータだけをバックアップする。これにより、従来技術に勝る大きな利点がもたらされる。例えば、非常に少ないデータレコードだけが変更されるデータベースを考慮されたい。従来技術のシステムは、変更が行われた場合、データベース全体をバックアップしようと試みる。しかし、本発明は、データベース変更によって実際に変更された少数のデータブロックをバックアップするだけである。また、本発明により、変更されたデータブロックをバックアップが所望されるものとして、あるいは所望されないものとして指定することができるようになる。したがって、第 1 の時点と第 2 の時点の間に変更され、バックアップが所望されるデータブロックだけが、スナップショットコピー 16 に送信される。したがって、メモリ、処理サイクル、及び通信帯域幅が、バックアップされることが所望されないデータブロックのバックアップコピーを格納することで浪費されない。

【0041】

従来技術との重要な相違について上述した。本発明は、スナップショットがとられた時点で存在するとおりのデータをキャプチャする。本発明は、元のデータ 14 に対して行われた時系列の変更をスナップショットコピー 16 に送信しようとはしない。例えば、データベースの単一のレコードが前回のバックアップが行われた時点と今回のバックアップ時点の間に 10 回、変更された場合、いくつかの従来技術のシステムは、10 回の変更をバックアップ記憶デバイスに送信する。しかし、本発明は、単に今回のバックアップ時点の前に行われた前回の変更を送信する。この例では、そのようなスキームにより、バックアップデバイスに送信されるデータの量が 10 分の 1 に縮減される。本発明により、バックアップデバイスに送信されるデータの量が、論理的に整合性のあるバックアップを行うのに必要とされる最小限に抑えられる。したがって、本発明は、スナップショットコピーがコンピュータシステム 12 から遠隔のサイトに配置された実施形態に理想的に適している。バックアップシステムがリモートサイトに配置されている場合、従来のダイヤルアップ電話回線を使用して 1 次システムとバックアップシステムの間でバックアップデータを転送することができる。

10

20

30

40

50

【 0 0 4 2 】

図3は、本発明のバックアップシステムの一実施形態を説明するためのブロック図で、図2に示したバックアップ方法を実施するためのブロック図である。以下の説明は、図3に示す処理ブロックのそれぞれのトップレベルの概要を示す。その後、各処理ブロックの詳細を示す。

【 0 0 4 3 】

コンピュータシステムの通常の動作中、大容量記憶デバイス20のような接続された大容量記憶手段に対してデータの読取りまたは書込みが定期的に行われる。したがって、本発明の実施形態は、大容量記憶デバイスにデータを書き込むための手段、及び大容量記憶デバイスからデータを読み取るための手段を含む。図3では、そのような手段を例として、大容量ストレージ読取り/書込み処理ブロック42で示している。大容量ストレージ読取り/書込み処理ブロック42の詳細を後に提示するが、大容量ストレージ読取り/書込み処理ブロック42の基本的機能は、1次大容量記憶デバイス20上の特定された記憶場所にデータブロックを書き込むこと、または1次大容量記憶デバイス20上の特定された記憶場所からデータブロックを読み取ることである。図3では、特定された記憶場所に対してデータブロックの読取りまたは書込みを行う要求を大容量ストレージ読取り/書込み要求44で示している。読取りまたは書込みが要求された場合はいつでも、大容量ストレージ読取り/書込み処理ブロック42が、大容量ストレージ読取り/書込み応答46によって示す応答を戻すことができる。この応答は、完了コード、または要求された操作の成功または失敗の他のインジケータを含むことが可能であり、読取り要求の場合、要求されたデータを含むことが可能である。

【 0 0 4 4 】

図2に関連して説明したとおり、本発明のバックアップ方法は、第1の時点と第2の時点の間に行われた変更を追跡する。したがって、本発明の実施形態は、第1の時点と第2の時点の間に元のデータ14のいずれの記憶場所に新しいデータが格納されているかを特定するための手段を備えている。そのような場所を特定し、追跡するための任意の方法を本発明で利用することができる。必要とされるのは、前回のバックアップ以来、新しいデータが格納された記憶場所を特定できることだけである。図3では、そのような手段を例として、バックアップマップ48で示している。バックアップマップ48は、1次大容量記憶デバイス20上の各データブロックに関するブルエントリを含むことが可能である。データブロックに新しいデータが書き込まれた際、そのデータブロックに対するエントリを設定することができる。この他に、新しいデータが格納されたデータブロックのリストも保持されることが可能である。必要とされるのは、特定の時点以来、新しいデータが格納されたデータブロックを区別し、特定することができることだけである。

【 0 0 4 5 】

上述したとおり、バックアップが行われる際、少なくとも所望されるデータブロックの静的スナップショットがとられる。したがって、本発明の実施形態は、特定の時点における静的スナップショットを保存するための手段を備えている。静的スナップショットの使用が好ましいのは、静的スナップショットの使用により、変更がバックアップされている最中にユーザが1次大容量記憶デバイス20にアクセスしつづけることができるようになるからである。元のデータ14からスナップショットコピー16に変更を転送するのにある期間を要するため、転送されるべきデータは、転送が済むまで、不変でなければならない。スナップショットコピー16が大容量記憶デバイス20内に配置されていない場合、データが不変であることを確実にする1つのやり方は、1次大容量記憶デバイス20に対するアクセスを防止することである。これにより、1次大容量記憶デバイス20にデータが書き込まれることが防止され、バックアップされるデータが、そのデータをスナップショットコピー16に転送することができるまで、不変であることが確実になる。残念ながら、この解決策は、極めて望ましくない。したがって、スナップショットコピー16に変更が転送される際、少なくとも転送されるデータの静的スナップショットがとられる。そのような静的スナップショットにより、転送されるデータが、そのデータを転送すること

10

20

30

40

50

ができるまで元の状態で保存され、同時に、大容量記憶デバイス20に対してデータの読取りまたは書込みを続けることができるようにデバイス20へのアクセスを継続することが可能になる。

【0046】

静的スナップショットを保存する任意の方法を本発明で使用することができる。ただし、使用される方法は何であれ、1次大容量記憶デバイス20に対するアクセスを中断することなく、静的スナップショットを保存できることが好ましい。言い換えれば、静的スナップショットは、ユーザが、大容量記憶デバイス20に対する読取りまたは書込みを続けることができるような形で保存されることが望ましい。

【0047】

図3では、静的スナップショットを保存するための手段を処理ブロック50で示している。図3に示すとおり、スナップショット処理機構を大容量ストレージ読取り/書込み処理ブロックに組み込むことが理にかなっている可能性がある。スナップショット処理ブロック50の詳細は以下に示すが、1つの好ましい実施形態は、上書きを受ける元のデータ14のデータブロックを元のデータ14からスナップショットストレージ22にコピーし、次に、そのブロックがスナップショットストレージ22の中に保存されたことをスナップショットマップ52の中で示すことにより、静的スナップショットを保存する。スナップショットストレージ22の中にコピーが入れられると、元のデータ14の中のそのデータブロックのコピーを上書きすることができる。

【0048】

図2に関連して前述したとおり、一続きの連続するバックアップが行われる場合、バックアップが行われている最中に元のデータ14に対して行われる変更を追跡することが必要である。言い換えれば、スナップショットがとられた後に元のデータ14に対して行われた変更を追跡する必要がある可能性がある。本発明の実施形態は、スナップショットがとられた時点の後に新しいデータが格納された元のデータ14の記憶場所を特定するための手段を含むことが可能である。特定の時点の後に新しいデータが格納された大容量記憶デバイスの記憶場所を追跡し、特定する任意のタイプの機構を利用することができる。例えば、バックアップマップ48と同様のマップを使用することができる。他の例として、特定の時点の後に新しいデータが格納されたデータ場所のリストも使用することができる。使用されるスナップショット機構のタイプにより、スナップショット機構が、本来的にそのような情報を追跡することも可能である。そのような実施形態では、上述した情報が後の使用のために保存されることが可能である。図3では、そのような手段をスナップショットマップ52で示している。スナップショット機構の1つの実装は、スナップショットがとられた後に新しいデータが格納された記憶場所を、図3のスナップショットマップ52のようなスナップショットマップの中で追跡する。

【0049】

本発明の実施形態は、バックアップされるべきデータブロックをスナップショットコピーに転送するための手段を含む。図3では、そのような手段を例として、1次バックアップ処理ブロック54で示している。1次バックアップ処理ブロック54について、以下により詳細に説明するが、1次バックアップ処理ブロック54の一般的な目的は、バックアップされるべきデータブロックを取り上げ、そのデータブロックをスナップショットコピー16に転送することである。転送されるデータブロックは、前回のバックアップ以来、大容量記憶デバイス上の記憶場所に格納された所望されるデータブロックである。

【0050】

1次バックアップ処理ブロック54は、バックアップを開始し、データをスナップショットコピー16に転送する機能を組み込むことが可能である。どのようにバックアップを開始することができるかの詳細を以下に説明する。

【0051】

図2の説明では、静的スナップショットを使用して、特定の時点における変更された所望されるデータブロックの状態を保存した。次に、変更された所望されるデータブロック

10

20

30

40

50

は、スナップショットコピー 16 にバックアップされる。変更された所望されるデータブロックが静的スナップショットによって保存されている場合、その所望されるデータブロックは、取り出してからでないと、スナップショットコピー 16 に送信することができない。したがって、本発明の実施形態は、静的スナップショットによって保存された所望されるデータブロックを取り出すための手段を備えている。そのような手段は、所望されるデータブロックをスナップショットコピー 16 に転送するための手段の一部であっても、別個であってもよい。図 3 では、静的スナップショットによって保存された所望されるデータブロックを取り出すための手段をバックアップ読取り処理ブロック 56 で示している。バックアップ読取り処理ブロック 56 の一実施形態の詳細を以下に示す。この処理ブロック 56 は、記憶場所から保存済みデータを取り出し、取り出されたデータブロックをスナップショットコピーに転送するために 1 次バックアップ処理ブロック 54 に送る。この機能は、1 次バックアップ処理ブロック 54 に組み込むこともできる。ただし、バックアップ読取り処理ブロック 56 によって実行される機能を際立たせるため、図 3 でそのブロックを他に示している。

【0052】

本発明は、所望されるデータブロックに関してスナップショットコピー 16 において 1 つまたは複数の論理的に整合性のあるバックアップ状態をキャプチャするように設計されている。その論理的に整合性のあるバックアップ状態をキャプチャするため、本発明の実施形態は、論理的に整合性のある状態が達せられた時点特定するための手段を含むことが可能である。論理的に整合性のある状態は、大容量記憶システム上で不適切に終了されたファイルのような論理的に整合性のない状態である。論理的に整合性のある状態は、いくつかの機構によって特定されることが可能である。例えば、論理的に整合性のある状態は、大容量記憶デバイス上の活動を監視することによって特定されることが可能である。大容量記憶デバイス上に活動が存在しない場合、すべての内部データバッファがフラッシュされており、バッファのデータが大容量記憶システムに書き込まれており、大容量記憶システムは、データブロックが更新されている状態にはないものと一般的に考えることができる。さらに、論理的に整合性のある状態が達せられた時点特定するように呼び出すことができる API が存在することも可能である。例えば、オペレーティングシステムまたは他のプログラムが、論理的に整合性のある状態が達せられた時点に戻すように行うことが可能な API コールを有することが可能である。さらに他の例として、システムは、所与の時点でスナップショットがとられるというメッセージをネットワークに接続されたすべてのユーザにブロードキャストすることができる。すると、ユーザは、必要な場合、適切なステップを行って自らのファイルの論理的に整合性のある状態を確実にする。その他の機構を使用してもよい。論理的に整合性のある状態が達せられた時点特定するための手段は、例えば、1 次バックアップ処理ブロック 54 などの図 3 の処理ブロックの 1 つに組み込むことができる。

【0053】

図 4 は、図 3 に示した大容量ストレージ読取り / 書込み処理ブロックの処理フローを示す図である。大容量ストレージ読取り / 書込み処理ブロック 42 の機能は、大容量記憶デバイス 20 に対して読取りまたは書込みを行うことである。さらに、スナップショット処理ブロック 50 が読取り / 書込み処理ブロック 42 に組み込まれているものと想定すると、処理ブロック 42 は、特定の時点における所望されるデータブロックに関する大容量記憶デバイス 20 の静的スナップショットを保存し、保持することも担う。図 3 に示した構成は、スナップショット処理ブロック 50 を一体化した機能として組み込んでいる。ただし、上述したとおり、スナップショット処理ブロック 50 を別個に実装することも可能である。スナップショット処理ブロック 50 を大容量ストレージ読取り / 書込み処理ブロック 42 に組み込むか、またはスナップショット処理ブロック 50 を別に実装するかは、本発明の目的にほとんど重要ではない設計上の選択と考えられる。本発明に重要な態様は、大容量記憶デバイス 20 に対して読取りまたは書込みを行う能力、及び特定の時点における大容量記憶デバイス 20 の少なくとも一部分の静的スナップショットを保存し、保持す

10

20

30

40

50

る能力を含むことである。

【0054】

次に、図4を参照すると、判定ブロック66がまず、スナップショット要求が行われたかどうかを判定する。この判定ブロックは、大容量ストレージ読取り/書込み処理ブロック42に組み込まれたスナップショット処理機能が、図3の大容量記憶デバイス20の少なくとも一部分のスナップショットをとるべきかどうかを明らかにする。スナップショット要求68は、1次バックアップ処理ブロック54によって生成される。1次バックアップ処理ブロック54はまず、論理的に整合性のある状態を特定してからそのようなスナップショット要求を発行する。この他に、論理的に整合性のある状態を特定するための手段を大容量ストレージ読取り/書込み処理ブロック42のスナップショット処理能力に組み込んでよい。その場合、大容量ストレージ読取り/書込み処理ブロック42が、論理的に整合性のある状態を特定し、スナップショットをとる。そのような詳細は設計上の選択であり、本発明の視点からは重要ではない。

10

【0055】

次に、スナップショット要求が受け取られた場合、次のステップは、所望されるデータブロックに対応する大容量記憶デバイス20の少なくとも一部分の静的スナップショットを保存することである。静的スナップショットを保存する任意の手段を本発明で使用することができるが、特定のプロセスを使用して静的スナップショットを保存することが好ましい。好ましい方法をステップ70、72、74、判定ブロック84及び85、ならびにステップ86の説明で要約する。この方法は、「大容量記憶システムに格納されたデータの静的スナップショットを提供するための方法及びシステム」という名称の(特許文献1)で詳細に説明されている。基本的に、静的スナップショットを保存する好ましい方法は、図3のスナップショットストレージ22のようなスナップショットストレージを利用して、新しいデータが上書きされることになる図3の大容量記憶デバイス20のような大容量記憶デバイスのデータブロックを保存する。以下により詳細に説明するとおり、保存されるデータブロックはまず、スナップショットストレージの中にコピーされ、データブロックが保存されたことを示すレコードが更新される。そのようなレコードは、例えば、図3のスナップショットマップ52の中に格納することができる。次に、保存されたデータブロックを失うことなく、新しいデータを大容量記憶デバイス20に書き込むことができる。

20

30

【0056】

判定ブロック66による評価に応じてスナップショットがとられた場合、次のステップは、図4のステップ70で示されるとおり、スナップショットマップをバックアップマップにコピーすることである。上述したとおり、図3のバックアップマップ48のようなバックアップマップを使用して、第1の時点と第2の時点の間にいずれのデータブロックが変更されたかが示される。次に、そのデータブロックがスナップショットコピー16に転送される。図3のスナップショットマップ52は、特定の時点に静的スナップショットが保存されて以来、変更されたデータブロックを特定する。したがって、新しいスナップショットがとられる際、スナップショットマップ52をバックアップマップとして使用することができる。スナップショットマップ52をバックアップマップ48にコピーすることにより、前回のスナップショットがとられた時点と現時点の間に新しいデータが格納されているデータロケーションを特定する所望の機能が実現される。明らかに、スナップショットマップとバックアップマップにコピーすることは必要ない可能性がある。スナップショットマップを単にバックアップマップとして使用して、新しいマップを現在のスナップショットマップとすることができる。

40

【0057】

バックアップマップとして使用することができるようにスナップショットマップが保存された後、次のステップは、現在のスナップショットマップをクリアすることである。このステップを図4にステップ72で示している。スナップショットマップは、スナップショットがとられて以来、新しいデータが格納されているデータブロックの指示を、その変

50

更されたデータブロックがバックアップの所望されるデータブロックであるか、所望されないデータブロックであるかに関わらず、格納するのに使用される。したがって、スナップショットマップは、図4のスナップショットストレージ22のようなスナップショットストレージの中にいずれのデータブロックが格納されているかを示す。新しいスナップショットがとられるので、スナップショットマップはクリアされなければならない。

【0058】

スナップショットマップがステップ72によってクリアされた後、次のステップは、図3のスナップショットストレージ22のようなスナップショットストレージをクリアすることである。これを図4のステップ74で示している。特にこのステップに関して、スナップショットストレージを物理的に消去すること、またはクリアすることは必要ない可能性があると留意されたい。一般に、他のいずれのタイプのストレージでもそうであるように、通常、ストレージに対するインデックスをクリアしてストレージが空であることを示すだけで十分である。したがって、インデックスが図3のスナップショットストレージマップ52のようなスナップショットストレージマップの一部として保持されている場合、ステップ72で実行されるとおりスナップショットストレージマップをクリアするだけで、スナップショットストレージが空であることを示すのに十分である。しかし、スナップショットストレージに対するインデックスがスナップショットストレージマップとは別に保持される場合は、ステップ74によってインデックスを別にクリアする必要がある可能性がある。スナップショットマップ及びスナップショットストレージがクリアされた後、システムは、新しいスナップショットを保存する準備ができている。したがって、図4で示されるとおり、開始に戻る。

【0059】

次に、図4に示した判定ブロック76は、大容量ストレージ読取り/書込み処理ブロック42によって受け取られたメッセージが大容量ストレージ読取り要求であるか、または大容量ストレージ書込み要求であるかをテストする。判定ブロック78に到達した時点では、可能なメッセージは、大容量ストレージ読取り要求か、または大容量ストレージ書込み要求だけである。これは、他のタイプの要求が、判定ブロック78より前に処理されるか、または濾過されて除外されるからである。判定ブロック78は、大容量ストレージ読取り要求と大容量ストレージ書込み要求を区別する。要求が大容量ストレージ読取り要求であった場合、次のステップは、大容量記憶デバイス20から要求されたデータブロックを取り出し、そのデータの要求を行っているプロセスに戻すことである。これをステップ80で示している。しかし、要求が書込み要求であった場合、実行は、判定ブロック82に進む。

【0060】

判定ブロック82は、スナップショットが保存されるべきかどうかを判定する。上述したとおり、好ましい実施形態では、上書きを受けるデータブロックを図3のスナップショットストレージ22のような保存メモリにコピーすることによってスナップショットが保存される。この実施形態では、スナップショットは、基本的に、増分式に保存される。言い換えれば、スナップショットが保存される際、スナップショットストレージが、ステップ72及び74で前述したとおりデータブロックを保存するように準備される。その後、保存されるべきデータに上書きする実際の書込み要求が行われるまで、スナップショットストレージには全くデータが格納されない。したがって、スナップショットがこのように保存される際、スナップショットがとられているか、またはスナップショットデータを保存することに配慮することなく大容量記憶システムに対する書込み要求が行われるべきかを判定することが重要である。判定ブロック82は、スナップショットデータを保存することなく書込み要求が行われるべきか、または書込み要求に関してスナップショットデータが保存されるべきかをテストする。スナップショットデータを保存することなく書込み要求が行われるべき場合、判定ブロック82は、実行がステップ88に進むことを示し、ステップ88で、データブロックが図3の大容量記憶デバイス20のような大容量記憶デバイスに書き込まれる。しかし、スナップショットが保存されるべき場合、実行は、判定

10

20

30

40

50

ブロック 8 4 に進む。

【 0 0 6 1 】

上述したとおり、好ましい実施形態に従ってスナップショットがとられる際、上書きを受けるデータがまず、図 3 のスナップショットストレージ 2 2 のようなスナップショットストレージにコピーされる。スナップショットストレージの中にデータが保存された後、新しいデータブロックを大容量記憶システムに書き込むことができる。スナップショットの目的は、特定の時点で大容量記憶システム上に存在していたとおりデータを保存することである。したがって、スナップショットは、スナップショットの時点で存在していたとおりデータを保存するだけでよい。判定ブロック 8 4 は、スナップショットがとられた時点で大容量記憶システムに格納されていた元のデータブロックが、スナップショットストレージの中に以前に保存されていたかどうかをテストする。言い換えれば、指定された書込み記憶場所に現在、格納されているデータが、スナップショットがとられた時点でその場所に格納されていたデータであり、そのデータをまず保存することなく書込み要求が行われた場合、元のデータは失われる。しかし、スナップショットがとられた時点で格納されていた元のデータが、スナップショットストレージの中に以前に保存済みである場合、元のデータが以前に保存済みであるため、心配なく書込み要求を行い、指定された場所に格納されているデータが何であれ、上書きすることができる。したがって、判定ブロック 8 4 により、元のデータがスナップショットストレージの中にまだ格納されていないと判定された場合、実行は、判定ブロック 8 5 に進む。

10

【 0 0 6 2 】

判定ブロック 8 5 は、データブロックが保護されるようにマーク付けされているかどうかを判定する。保護マップ 5 3 で図 3 に示すとおり、ユーザは、バックアップが所望されるもの、またはバックアップが所望されないものとしてデータブロックを指定して、バックアップが所望されるデータブロックだけが実際にバックアップされるようにすることができる。バックアップが所望されるものとして指定されたデータブロックを本明細書では所望されるデータブロックと呼んでおり、バックアップが所望されるものとして指定されていないデータブロックを所望されないデータブロックと呼んでいる。この特徴により、バックアップに必要とされる時間と記憶スペースの量が削減される。バックアップが所望される / 所望されないデータブロックを指定する要求 4 3 が、大容量ストレージ読取り / 書込み処理 4 2 によって受け取られる。

20

30

【 0 0 6 3 】

スナップショットマップ 5 2 上のデータブロックは、とられた最新の静的スナップショットの結果としてスナップショットストレージ 2 2 の中に格納されたデータブロックを示す。大容量ストレージ読取り / 書込み処理 4 2 が、保護マップ 5 3 上で、バックアップが所望されるデータブロックを示す。この他に、スナップショット処理が、保護マップ 5 3 上で、バックアップが所望されないデータブロックを示し、そのデータブロックに常に最新であるものとしてマークを付けることによってそのデータブロックのバックアップを防止することもできる。また、他の実施形態では、スナップショットマップ 5 2 と保護マップ 5 3 が 1 つのマップであることも可能である。言い換えれば、保護マップ 5 3 上で実行される諸機能をスナップショットマップ 5 2 上で実行することができる。

40

【 0 0 6 4 】

データブロックが保護されるようにマーク付けされていない、つまり言い換えれば、バックアップが所望されない場合、実行は、判定ブロック 8 5 からステップ 8 6 をスキップしてステップ 8 8 に進む。あるいは、データブロックが保護されるようにマーク付けされている場合、実行は、ステップ 8 6 に進み、元のデータブロックがスナップショットストレージ 2 2 の中にコピーされる。

【 0 0 6 5 】

一部の実施形態では、ステップ 8 5 は省略することができる。変更されたデータブロックが、所望されるデータブロックであるか否かには関わらず保存される。次に、図 6 のステップ 1 1 3 が示すとおり、データブロックをスナップショットコピーに送信する際、バ

50

ックアップ読取り処理 5 6 が、保護マップ 5 3 を使用してあらゆる所望されないデータブロックを濾過して除外して、所望されるデータブロックだけがスナップショットコピーに送信されるようにする。

【 0 0 6 6 】

ステップ 8 6 によって元のデータが保存された後、または元のデータが以前に保存済みであるとの判定が判定ブロック 8 4 によって行われた後、またはデータブロックが保護されるようにマーク付けされていないとの判定が判定ブロック 8 5 によって行われた後、実行は、ステップ 8 8 に進み、書込み要求に含まれるデータブロックを大容量記憶デバイス上の指定された記憶場所へ書き込むことによって書込み要求が満たされる。

【 0 0 6 7 】

次に、ステップ 9 0 が、新しいデータを含んでいる記憶場所を特定する。前述したとおり、これは、図 3 のスナップショットマップ 5 2 のようなスナップショットマップの中にエントリを入れることによって達することができる。ステップ 9 0 は、新しいデータが書き込まれている大容量記憶デバイスの記憶場所を特定するための前述した手段の一例に過ぎない。次に、書込み要求を行っているプロセスに応答が戻されることが可能である。そのような応答を送信することを図 4 にステップ 9 2 で示している。そのような応答は、通常、書込み操作の成功または失敗を示すだけでなく、書込み動作の完了を示すためにも、書込み要求を発行したプロセスに送信される。次に、実行は開始に戻り、次の要求が処理される。

【 0 0 6 8 】

図 5 は、図 3 に示した 1 次バックアップ処理ブロックの処理フローを示す図である。1 次バックアップ処理ブロック 5 4 は、スナップショットコピーに転送される必要があるデータブロックを獲得し、その転送を実行することを担う。第 1 に、ステップ 1 0 0 が、論理的に整合性のあるバックアップ状態を特定する。論理的に整合性のある状態が特定された後、バックアップを行うことができるように論理的に整合性のある状態のスナップショットが保存される。スナップショットは、スナップショット処理、例えば、図 3 の大容量ストレージ読取り / 書込み処理ブロック 4 2 に組み込まれたスナップショット処理ブロック 5 0 にスナップショットをとるように通知するステップ 1 0 2 によって保存される。一実施形態では、これにより、スナップショット要求 6 8 が大容量ストレージ読取り / 書込み処理ブロック 4 2 に送信されることになる。前述したとおり、この要求により、とられるスナップショットの準備を行う図 4 のステップ 7 0、7 2、及び 7 4 が実行されるようになる。その後、スナップショットがとられた時点における大容量記憶デバイス 2 0 の中に格納された指定のデータブロックに関する元のデータが、図 4 の判定ブロック 8 4、判定ブロック 8 5、及びステップ 8 6 によって保存される。

【 0 0 6 9 】

図 5 のステップ 1 0 0 によって特定された論理的に整合性のあるバックアップ状態を保存するためにスナップショットがとられた後、図 5 における次のステップは、ステップ 1 0 4 で示すとおりスナップショットコピー 1 6 に転送するためにデータブロックをアセンブルすることである。データブロックをアセンブルしてアセンブル済みデータ 6 4 が形成された後、ステップ 1 0 6 は、アセンブル済みデータ 6 4 をスナップショットコピー 1 6 に送信する。これは、スナップショットコピーに書き込むためにデータを大容量ストレージ読取り / 書込み処理ブロック 4 2 に送信することによって達することができる。次に、実行は、開始に戻り、1 次バックアップ処理ブロック 5 4 によって次の論理的に整合性のある状態が特定されて前述したプロセスが繰り返される。

【 0 0 7 0 】

上述したとおり、ステップ 1 0 4 によってスナップショットコピー 1 6 に送信されるデータブロックは、前回のバックアップ以来、変更されており、バックアップされることが所望されるデータブロックだけである。さらに、データブロックは、スナップショットがとられた時点に存在していたとおりに転送される。したがって、変更されており、図 3 の保護マップ 5 3 のような保護マップの中で保護されるものとして特定されている図 3 のバ

10

20

30

40

50

ックアップマップ48のようなバックアップマップの中で変更されているものとして特定されているデータブロックだけが、転送される。スナップショットは、スナップショットがとられた時点における状態にある所望されるデータブロックを保存する。したがって、1次バックアップブロック54は、スナップショットによって保存された、いくつかのデータブロックを取り出す必要がある。1次バックアップ処理ブロック54が、スナップショット及び/または大容量記憶システムからデータブロックを取り出すのに必要とされる機能を組み込むことも、そのような機能が、別個の処理ブロックに組み込まれることも可能である。その機能を組み込んだ別個の処理ブロックを図3にバックアップ読取り処理ブロック56で示している。

【0071】

図6は、図3に示したバックアップ読取り処理ブロックの処理フローを示す図で、上述したスナップショットによって保存されたデータを回復するように設計されたバックアップ読取り処理ブロック56の処理フローを示している。

【0072】

図6では、判定ブロック112により、バックアップ読取り処理ブロック56が、スナップショットがとられた時点に存在していたデータを取り出すことになる読取り要求だけを扱うことが強調されている。この判定ブロックは、処理の構造及びアーキテクチャにより、そのような読取り要求だけが図3のバックアップ読取り処理ブロック56に送信されることが保証される場合、必要ない可能性がある。同様に、判定ブロック113により、バックアップ読取り処理ブロック56が、バックアップシステムへの最終的な転送が所望されるデータブロックだけを取り出すことが強調されている。あるデータブロックが所望されるデータブロックであるかどうかの検査は、図3の保護マップ53を参照することによって達することができる。

【0073】

適切な状況下では、判定ブロック113を省略してもよい。例えば、図4の判定ブロック85で示すとおり、本発明の一部の実施形態は、保護されるものとしてマークが付けられたデータブロックだけを保存することができる。保護されるデータブロックだけがスナップショットストレージの中に入れられる場合、判定ブロック113を省いてもよい。というのは、データブロックがスナップショットストレージの中に格納されているという判定ブロック114による指示は、データブロックが保護されるようにマーク付けられていることを必然的に意味するからである。

【0074】

スナップショットがとられた時点に存在していたとおりに所望されるデータブロックを取り出すため、どこにそのデータブロックが存在するかを特定しなければならない。図4に関連して前述したとおり、スナップショットがとられた後、所望されるデータブロックに新しいデータブロックが上書きされるときに初めて、所望されるデータブロックが図3のスナップショットストレージ22のようなスナップショットストレージの中にコピーされる。これは、所望されるデータブロックに上書きが全く行われないう場合、大容量記憶デバイスに格納された所望されるデータブロックは、スナップショットがとれた時点で存在していた元の所望されるデータブロックであることを意味する。しかし、所望されるデータブロックに1回または複数回、上書きが行われている場合、元の所望されるデータブロックは、スナップショットストレージの中に格納されている。図6の判定ブロック114により、スナップショットがとられて以来、要求された所望のデータブロックが変更されているかどうか判定される。これは、データブロックが変更されているかどうかを判定するため、図3のスナップショットマップ52のようなスナップショットマップを検査することによって達せられる。上述したとおり、スナップショットマップは、スナップショットがとられて以来、変更されている記憶場所またはデータブロックを特定する。

【0075】

スナップショットがとられて以来、記憶場所に新しいデータが格納されている場合、ステップ116により、データブロックはスナップショットストレージから取り出されるこ

10

20

30

40

50

とが示される。しかし、スナップショットがとられて以来、記憶場所の内容が変更されていない場合、ステップ 118 により、データブロックは、大容量記憶デバイス 20 から取り出されることが示される。いずれの場合でも、保護されるものとして指定されたデータブロックが、ステップ 120 によって要求側プロセスに戻される。

【0076】

バックアップを作成する際の図 3 ~ 図 6 の操作をより詳細に説明するため、図 7 A、図 7 B、図 8 A、図 8 B に詳細な例を示す。図 7 A 及び図 7 B に示す実施形態は、図 7 A 及び図 7 B が、あたかも図 4 に判定ブロック 85 が存在しないかのように機能することにおいて、図 8 A 及び図 8 B に示した実施形態とは異なる。したがって、図 7 A 及び図 7 B は、保護されるデータブロックと保護されないデータブロックをともにスナップショットメモリの中に格納するが、スナップショットメモリからのデータブロックが要求された際、保護されるデータブロックだけを提供する実施形態を示している。対照的に、図 8 A 及び図 8 B は、保護されるデータブロックだけをスナップショットストレージの中に格納する本発明の実施形態を示しており、これは、データブロックがスナップショットストレージの中に出現する場合、そのデータブロックは、必然的に保護されるデータブロックであることを意味する。上述したとおり、本発明は、図 4 の判定ブロック 85 の有無に関わらず実施することができる（保護されるデータブロックだけをスナップショットストレージの中に格納して）。以下の図 7 A 及び図 7 B の説明は、図 4 の判定ブロック 85 が存在せず、したがって、保護されるデータブロックと保護されないデータブロックがともにスナップショットストレージの中に格納されるものと想定する。

【0077】

まず、図 7 A を参照して、大容量記憶デバイス 20 の元のデータ部分 14 の 1 ~ 6 という符号が付けられた記憶場所に格納された一群のデータブロック 122 を考慮されたい。同様に、バックアップマップ 48 が、記憶場所 122 に対応する 6 つのマップ場所 126 を有し、スナップショットマップ 52 が、記憶場所 122 に対応する 6 つのマップ場所 128 を有し、保護マップ 53 も、記憶場所 122 に対応する 6 つのマップ場所 129 を有している。図 7 A に示すとおり、時刻 T_0 に、マップ場所 126、128 がクリアされる。ただし、マップ場所 129 の場所 3 にはマークが付けられ、データブロック 122 のデータブロック 3 が保護されるものとして指定されていないことを示している。

【0078】

図 7 B は、1 ~ 6 の符号が付けられた記憶場所に同様に格納されている一群のデータブロック 124 をやはり有する大容量記憶デバイス 20 のスナップショットコピー 16 部分を示している。ただし、データブロック 3 は、図 7 A 及び図 7 B の対応するデータブロックが対応する符号を有することを確実に示すだけのために示している。図 7 A のマップ場所 129 の場所 3 は、データブロック 122 のデータブロック 3 が、バックアップ中に転送されないことを示しているため、バックアップストレージ 24 中のデータブロック 124 のデータブロック 3 は、必ずしも存在しない。つまり、バックアップストレージ 24 は、1 次大容量ストレージ 20 からのデータブロック 122 のデータブロック 3 に対応するデータブロックを必ずしも有さない。このため、データブロック 3 は、単にグレーで塗りつぶすのではなく、データブロック 124 から完全に除外することができる。時刻 T_0 で、124 に格納されたデータブロックは、少なくとも所望されるデータブロック 1、2、及び 4 ~ 6 に関する限り、122 に格納されたデータブロックと全く同じである。

【0079】

時刻 T_0 の後、データブロック 130 が、記憶場所 122 の場所 3 及び 4 に格納されるものと想定する。次に、データブロック 130 が適切な記憶場所に書き込まれるようにするため、1 つまたは複数の大容量ストレージ書込み要求が図 3 の大容量ストレージ読取り / 書込み処理ブロック 42 に送られる。図 4 を参照すると、大容量ストレージ書込み要求が、以下の形で処理される。

【0080】

判定ブロック 66、76、及び 78 が一緒に、書込み要求が大容量ストレージ読取り /

書込み処理ブロック 4 2 に送られていることを判定する。したがって、実行は、以上 3 つの判定ブロックを通過して判定ブロック 8 2 に進む。上述したとおり、判定ブロック 8 2 は、スナップショットがとられているかどうかをテストする。この例では、この時点で、全くスナップショットはとられていない。したがって、実行は、要求されたデータブロックを図 7 A の大容量ストレージ 2 0 に書き込むステップ 8 8 に進む。したがって、データブロック 1 3 0 が、記憶場所 1 2 2 に格納されて記憶場所 1 3 2 が生成される。図示するのとおり、記憶場所 1 3 2 では、場所 3 及び 4 に格納されていたデータブロックが、3 a 及び 4 a に変更されている。

【 0 0 8 1 】

図 4 に戻ると、次に、ステップ 9 0 により、新しいデータが格納されている記憶場所が変更されているものとして示されるべきことが示される。多くのスナップショット実施形態では、この目的でスナップショットマップを使用することができる。図 7 A で、マップ 1 3 4 が使用され、場所 3 及び 4 がグレーにされて、データが記憶場所 3 及び 4 に格納されていることを示している。マップ場所 1 2 6 によって示されるバックアップマップ 4 8 の中の記憶場所は、この時点で、変更されないままであることに留意されたい。図 4 を参照すると、ステップ 9 2 によって書込み要求応答が戻され、実行は、開始に戻って次の要求を待つ。

【 0 0 8 2 】

次に、図 7 A に戻り、次の要求が、場所 3、4、及び 6 に格納されるべき 3 つのデータブロック 1 3 6 を含むものと想定する。スナップショットはまだとられていないため、この要求は、前の書込み要求と同じ形で処理され、実行は、図 4 の判定ブロック 6 6、7 6、7 8、及び 8 2 を通って図 4 のステップ 8 8 に進む。ステップ 8 8 は、新しいデータが大容量記憶デバイスの中に格納されたことを示し、したがって、図 7 A の記憶場所 1 3 8 は、この時点で、場所 3 に格納されていたデータブロックが 3 b に変更され、場所 4 に格納されていたデータブロックが 4 b に変更され、場所 6 に格納されていたデータブロックが、6 a に変更されていることを示す。前の書込み要求の場合と同様に、次に、マップ場所 1 4 0 が更新されて、場所 3 及び 4 に加えて場所 6 も変更されていることを示す。マップ場所 1 2 6 は変更されないままである。

【 0 0 8 3 】

次に、図 5 を参照して、本例のこの時点で、図 3 のバックアップ処理ブロック 5 4 が、図 5 のステップ 1 0 0 で論理的に整合性のあるバックアップ状態を特定するものと想定する。論理的に整合性のあるバックアップ状態を特定した後、ステップ 1 0 2 は、図 3 のスナップショット要求 6 8 を大容量ストレージ読取り / 書込み処理ブロック 4 2 に送信する。

【 0 0 8 4 】

次に、図 4 を参照すると、そのスナップショット要求が、判定ブロック 6 6 によって処理され、ステップ 7 0、7 2、及び 7 4 が実行されることになる。ステップ 7 0 で、スナップショットマップがバックアップマップにコピーされる。図 7 A で、これは、マップ場所 1 4 0 がバックアップマップ 4 8 のマップ場所 1 4 2 にコピーされることを意味する。したがって、マップ場所 1 4 2 は、場所 3、4、及び 6 に新しいデータが格納されていることを示す。次に、図 4 に戻ると、上述したとおり、ステップ 7 2 により、スナップショットマップがクリアされ、ステップ 7 4 により、スナップショットストレージがクリアされる。次に、図 4 の実行は、開始に戻ってさらなる処理を待つ。

【 0 0 8 5 】

この時点で、書込み要求が大容量ストレージ読取り / 書込み処理ブロック 4 2 に着信して、図 7 A のデータブロック 1 4 4 が、記憶場所 1 3 8 に格納されることを要求する。これは書込み要求であるため、実行は、判定ブロック 6 6、7 6、及び 7 8 を通って図 4 の判定ブロック 8 2 に進む。前の書込み要求とは異なり、図 7 A 及び図 7 B で示すとおり、時刻 T_1 にスナップショットがとられている。したがって、実行は、判定ブロック 8 4 に進む。

10

20

30

40

50

【 0 0 8 6 】

判定ブロック 8 4 により、上書きを受ける記憶場所に格納されたデータが、スナップショットストレージの中に以前に格納済みであるかどうか判定される。この例では、データブロック 1 4 4 は、記憶場所 1 及び 3 に格納される。記憶場所 1 及び 3 は、スナップショットストレージの中にまだ入れられていないため、実行は、ステップ 8 6 に進み、記憶場所 1 3 8 の場所 1 及び 3 が、スナップショットストレージ 2 2 にコピーされる。(この実施形態では、図 4 の判定ブロック 8 5 は存在しないことを思い起こされたい。) 図 7 A では、これを記憶場所 1 4 6 がデータブロック 1 を含み、記憶場所 1 4 8 がデータブロック 3 b を含んでいることで示している。

【 0 0 8 7 】

データブロック 1 及び 3 b がスナップショットストレージ 2 2 の中に保存された後、ステップ 8 8 で、新しいデータブロックが、大容量記憶デバイスに書き込まれる。図 7 A に戻ると、これは、データブロック 1 a によってデータブロック 1 が上書きされ、データブロック 3 c によってデータブロック 3 b が上書きされている記憶場所 1 5 0 を生成するために、データブロック 1 a 及び 3 c が記憶場所 1 3 8 に書き込まれることを意味する。次に、図 5 のステップ 9 0 により、データブロックが変更されているものとして識別される必要があることが述べられている。したがって、スナップショットマップ 5 2 のマップ場所 1 5 2 を変更して、記憶場所 1 及び記憶場所 3 に新しいデータが格納されていることが示される。次に、図 4 のステップ 9 2 によって指示されるとおり、書込み要求応答が戻される。

【 0 0 8 8 】

次に、図 5 に戻ると、図 5 のステップ 1 0 2 によって指示されるとおり、図 3 の大容量ストレージ読取り / 書込み処理ブロック 4 2 によって時刻 T_1 にスナップショットがとられている。次に、ステップ 1 0 4 及び 1 0 6 により、スナップショットがとられる前に変更されていたデータブロックが伝送パケットにアセンブルされて、スナップショットコピー 1 6 に送信されるべきことが示される。転送されるべきデータブロックは、バックアップマップ 4 8 及び保護マップ 5 3 に含まれる情報によって示される。

【 0 0 8 9 】

図 7 A に戻ると、バックアップマップ 4 8 のマップ場所 1 4 2 により、時刻 T_1 にスナップショットがとられるのに先立って記憶場所 3、4、及び 6 が変更されていることが示される。スナップショット場所 1 5 2 の検査により、データブロック 4 及び 6 が大容量記憶システム上にあり、データブロック 3 がスナップショットストレージ 2 2 の中にあることが示される。次に、図 5 のステップ 1 0 4 により、記憶場所 3、4、及び 6 に格納されているデータブロックが図 3 のバックアップ読取り処理ブロック 5 6 によって取り出されることが要求される。

【 0 0 9 0 】

バックアップ読取り処理ブロック 5 6 は、図 6 に示すとおり、1 次バックアップ処理ブロック 5 4 から受け取られた上記の要求を処理する。要求は、記憶場所 3、4、及び 6 に格納されたデータブロックに関する。記憶場所 3 に格納されたデータブロックに関して、判定ブロック 1 1 3 により、記憶場所 3 に格納されているデータブロックが保護されないようにマーク付けられており、したがって、ステップ 1 2 0 で戻されないことが判定される。場所 4 及び 6 に格納されているデータブロックは、保護されないようにマーク付けられていないので、図 6 の判定ブロック 1 1 4 により、ステップ 1 1 8 で大容量記憶デバイスから記憶場所 4 及び 6 に格納されているデータブロックが取り出され、ステップ 1 2 0 で 1 次バックアップ処理ブロック 5 4 に戻される。このプロセスは、記憶場所 1 5 0 からデータブロック 4 b 及び 6 a を取り出すことによってデータブロック 1 5 3 がアセンブルされる図 7 A で図示している。次に、データブロック 1 5 3 は、大容量ストレージ読取り / 書込み処理 4 2 を介してスナップショットコピー 1 6 に転送される。これを図 7 A 及び図 7 B に図示している。図 7 B では、データブロック 1 5 3 がスナップショットコピー 1 6 によって受信され、記憶場所 1 2 4 に印加されて記憶場所 1 5 4 が実現される。記憶場

10

20

30

40

50

所 1 5 4 は、バックアップが所望されないものとして特定されているために除外される記憶場所 3 を除き、元のデータの記憶場所 1 3 8 (図 7 A) と全く同じである。記憶場所 1 3 8 は、スナップショットがとられた時点 T_1 における大容量記憶デバイス 2 0 の状態を表していたことを思い起こされたい。したがって、スナップショットコピー 1 6 を時刻 T_1 における元のデータ 1 4 と同期させるため、時刻 T_0 と時刻 T_1 の間に行われた変更が、現時点で、スナップショットコピー 1 6 にバックアップされている。

【 0 0 9 1 】

次に、図 7 A に戻り、データブロック 1 5 6 が記憶場所 1 5 0 に書き込まれるものと想定する。データブロック 1 5 6 の書込みにより、記憶場所 1、4、及び 6 に格納されたデータブロックに対する変更が生じる。大容量ストレージ読取り / 書込み処理ブロック 4 2 が、上述したとおり、場所 4 及び 6 に格納されるデータブロックの書込みを処理し、時刻 T_1 後にそれらの場所に格納されたデータブロック 1 4 4 (データブロック 4 b 及びデータブロック 6 a) が、スナップショットストレージ 2 2 の中に格納される。次に、新しいデータブロック 4 c 及び 6 b が、大容量記憶デバイス 2 0 に書き込まれる。

【 0 0 9 2 】

記憶場所 1 に格納されるデータブロックに関して、実行は、図 4 で判定ブロック 8 4 まで進む。この判定ブロックは、スナップショットがとられた時点に記憶場所に格納されているデータブロックが、スナップショットストレージの中に以前に保存済みであるかどうかをテストすることを思い起こされたい。記憶場所 1 に格納されているデータブロックに関して、データブロックは、図 7 A のデータブロック 1 4 6 によって示されるとおりスナップショットストレージ 2 2 の中に以前に格納済みである。したがって、図 4 は、ステップ 8 6 がスキップされ、新しいデータが大容量記憶デバイスに書き込まれることを示している。図 7 A で、これにより、データブロック 1 b がデータブロック 1 a に取って代わり、そのデータブロックが失われることになる。

【 0 0 9 3 】

本発明は、前回のバックアップ以来、変更されている記憶場所の所望されるデータブロックだけを転送する。さらに、データブロックは、スナップショットがとられた時点に存在していたとおりに転送される。したがって、前回のバックアップ以来、元のデータにおける特定の記憶場所に 5 つの異なるデータブロックが格納されている場合、最後に格納されたデータブロック (例えば、スナップショットがとられる直前の) だけがスナップショットコピーに転送される。これは、バックアップをとる際にスナップショットコピー 1 6 が、論理的に整合性のあるバックアップだけを保存するからである。言い換えれば、スナップショットコピーは、ある時点における論理的に整合性のある状態から別の時点における論理的に整合性のある状態に移る。別々の時点における所望されるデータブロックの論理的に整合性のある状態を保存することにより、従来技術に勝る大きな利点が提供される。

【 0 0 9 4 】

例えば、元のデータに対して行われたすべての変更を逐一キャプチャする従来技術のシステムを考慮されたい。そのような従来技術のシステムは、すべての書込み操作を元のデータとバックアップコピーの両方に送ろうと試みる。理論的には、これにより、バックアップコピーが大容量記憶デバイスと全く同じコピーになる。しかし、この手法では問題が生じる。具体的には、バックアップコピーにすべての更新を送ることは、比較的大きい帯域幅を要する。単一のデータブロックの複数回の更新を併合して単一の更新にすることにより、本発明は、元のコピーとバックアップコピーの間で転送されなければならないデータの量を少なくする。

【 0 0 9 5 】

さらに、元のデータを含む 1 次システムが書込み更新中にクラッシュした場合、それにより元のデータが論理的に不整合な状態のままになる可能性がある。バックアップコピーが元のデータに対して行われるすべての変更を追跡している場合は、1 次システムがクラッシュした際、バックアップコピーも論理的に不整合な状態のままになる可能性がある。

10

20

30

40

50

この例により、第2の論理的に整合性のある状態が特定される前に既知の論理的に整合性のある状態を離れることの問題が際立つ。本発明は、本発明は、新しい論理的に整合性のある状態が特定されるまで前の論理的に整合性のある状態を保持し、次に、前の論理的に整合性のある状態から次の論理的に整合性のある状態にスナップショットコピーを移すことを、その2つの論理的に整合性のある状態の間で論理的に不整合な状態を通して遷移することなく行うことにより、上述した問題を回避する。

【0096】

図7Aに戻ると、データブロック156が記憶場所150に印加された際、記憶場所158がもたらされる。次に、マップ場所152が更新されて、時刻 T_1 以来、変更されている記憶場所に、現時点で、記憶場所1及び3に加えて記憶場所4及び6が含まれることを示す。これを図7Aでスナップショットストレージ52のマップ場所160で示している。

10

【0097】

次に、大容量記憶デバイス20の第2のバックアップが行われるものと想定する。その場合、バックアップは、図5で上述したとおり行われ、実行は、ステップ100に進み、論理的に整合性のある状態が特定される。図7Aで、その論理的に整合性のある状態は、時刻 T_2 に特定されたものと想定する。すると、図5のステップ102により、時刻 T_2 にスナップショットがとられることが通知される。時刻 T_1 にとられるスナップショットに関連して上述したとおり、大容量ストレージ読取り/書込み処理ブロック42が、図3のスナップショット要求68のようなスナップショット要求を受け取り、ステップ70でスナップショットマップをバックアップマップにコピーする。これを図7Aで示しており、バックアップマップ48のマップ場所162がスナップショットマップ52のマップ場所160と同じになるように変更される。

20

【0098】

次に、図4のステップ72及び74により、スナップショットマップ及びスナップショットストレージがクリアされるべきことが示される。図7Aで、スナップショットマップ52のマップ場所164で示されるとおりスナップショットマップがクリアされる。ただし、スナップショットストレージ22には依然としてデータブロックが格納されているのが示されている。これは、スナップショットストレージ22に対するインデックスがクリアされて、スナップショットストレージ22が全くデータブロックを含くまないように見える限り、データブロックは、スナップショットストレージ22の中に依然として物理的に存在していてもよいことを示すためである。

30

【0099】

時刻 T_2 にスナップショットがとられた後に記憶場所158の中にデータブロックが全く存在しないものと想定すると、データブロック166は、図6で説明するプロセスに従って記憶場所158から読み取られる。記憶場所158の場所3は、図5の判定ブロック113が保護マップ場所129を使用して場所3が保護されないことを判定するため、読み取られないことに留意されたい。したがって、データブロック3cが読み取られて、バックアップシステムに転送されることはない。図5のステップ104及び106で示すとおり、読み取られたデータブロックは、次に、大容量ストレージ読取り/書込み処理ブロック42を介してスナップショットコピーに伝送される。図7Bに示すとおり、データブロック166は、次に、記憶場所168に達するために記憶場所154に印加され、記憶場所168は、バックアップが所望されないものとして特定されているために除外される記憶場所3を除き、元のデータ(図7A)の記憶場所158と全く同じコピーである。

40

【0100】

次に、図8A及び図8Bは、保護されるデータブロックだけをスナップショットストレージの中に格納する本発明の実施形態を示しており、これは、データブロックがスナップショットストレージの中に出現する場合、そのデータブロックは、必然的に保護されるデータブロックであることを意味する。上述したとおり、本発明は、図4の判定ブロック85の有無に関わらず実施することができる(保護されるデータブロックだけをスナップシ

50

ショットストレージの中に格納して)。以下の図 8 A 及び図 8 B の説明は、図 4 の判定ブロック 8 5 が存在し、したがって、保護されるデータブロックだけがスナップショットストレージの中に格納されるものと想定する。図 7 A 及び図 7 B の以上の説明の多くは、8 A 及び図 8 B にも当てはまるので、図 8 A 及び図 8 B の以下の説明は、図 7 A、図 7 B、図 8 A、図 8 B の間の違い、すなわち、スナップショットストレージ 2 2 の操作について中心的に扱う。

【 0 1 0 1 】

図 3 に示すとおり、時刻 T_1 に、スナップショット要求 6 8 が大容量ストレージ読取り / 書込み処理ブロック 4 2 に送られる。図 4 を再び参照すると、そのスナップショット要求が判定ブロック 6 6 によって処理されて、ステップ 7 0、7 2、及び 7 4 が実行されることになる。ステップ 7 0 で、スナップショットマップがバックアップマップにコピーされる。図 8 A で、これは、マップ場所 1 4 0 がバックアップマップ 4 8 のマップ場所 1 4 2 にコピーされることを意味する。したがって、マップ場所 1 4 2 により、場所 3、4、及び 6 に新しいデータが格納されていることが示される。図 4 に戻ると、上述したとおり、ステップ 7 2 により、スナップショットマップがクリアされ、ステップ 7 4 により、スナップショットストレージがクリアされる。次に、図 4 の実行は、開始に戻り、さらなる処理を待つ。

【 0 1 0 2 】

この時点で、図 8 A のデータブロック 1 4 4 が記憶場所 1 3 8 に格納されることを要求する書込み要求が大容量ストレージ読取り / 書込み処理ブロック 4 2 に着信する。これは書込み要求であるため、実行は、図 4 の判定ブロック 6 6、7 6、及び 7 8 を通って判定ブロック 8 2 まで進む。図 8 A 及び図 8 B に示すとおりスナップショットが時刻 T_1 にとられて、実行は、判定ブロック 8 4 に進む。ここまでは、図 7 A 及び図 7 B に関連して説明した処理と全く同じである。

【 0 1 0 3 】

判定ブロック 8 4 により、上書きを受ける記憶場所に格納されたデータが、スナップショットストレージの中に以前に格納済みであるか否かが判定される。この例では、データブロック 1 4 4 は、記憶場所 1 及び 3 に格納されることになる。記憶場所 1 及び 3 はまだスナップショットストレージの中に入れていないため、プロセスは、判定ブロック 8 5 を実行して、バックアップが所望されるデータブロックとバックアップが所望されないデータブロックを区別する。図 8 A の保護マップ場所 1 2 9 のデータブロック 3 は、バックアップが所望されないものとしてマーク付けされているので、実行は、判定ブロック 8 5 からステップ 8 6 をスキップしてステップ 8 8 に進む。記憶場所 1 3 8 のデータブロック 3 b は、スナップショットストレージ 2 2 にコピーされない。しかし、保護マップ場所 1 2 9 のデータブロック 1 は、バックアップが所望されるものとして特定されており、したがって、保護されるものとしてマーク付けされている（すなわち、データブロック 1 は、保護されないようにマーク付けされていない）。したがって、データブロック 1 が処理される際、実行は、判定ブロック 8 5 からステップ 8 6 に進み、記憶場所 1 3 8 のデータブロック 1 がスナップショットストレージ 2 2 にコピーされる。図 8 A で、これは、スナップショットストレージ 2 2 が 1 4 6 という符号の付いたデータブロック 1 を含むことで示されている。上述したとおり、スナップショットストレージ 2 2 が保護されるデータブロックだけを含むことを確実にすることが、図 7 A 及び図 7 B の実施形態と図 8 A 及び図 8 B の実施形態の間の違いである。

【 0 1 0 4 】

データブロック 1 がスナップショットストレージ 2 2 の中に保存された後、ステップ 8 8 で、新しいデータブロックが、元のデータに書き込まれる。図 8 A に戻ると、これは、データブロック 1 a によってデータブロック 1 が上書きされ、データブロック 3 c によってデータブロック 3 b が上書きされた記憶場所 1 5 0 をもたらすためにデータブロック 1 a 及び 3 c が記憶場所 1 3 8 に書き込まれることを意味する。次に、図 5 のステップ 9 0 により、データブロックが変更されているものとして特定される必要があることが述べら

10

20

30

40

50

れている。したがって、スナップショットマップ52のマップ場所152が変更されて、記憶場所1及び記憶場所3に新しいデータが格納されていることを示す。次に、図4のステップ92によって指示されるとおり、書込み要求応答が戻される。

【0105】

この時点で、図8A及び図8Bに示した実施形態は、図7A及び図7Bに示した実施形態と全く同様に機能して進められる。図5に戻ると、図5のステップ102によって指示されるとおり、図3の大容量ストレージ読取り/書込み処理ブロック42によって時刻 T_1 にスナップショットがとられている。次に、ステップ104及び106により、スナップショットがとられる前に変更されたデータブロックがアセンブルされ、スナップショットコピーに送信されるべきことが示される。転送されるべきデータブロックは、バックアップマップ48及び保護マップ53に含まれる情報によって示される。

10

【0106】

図8Aを参照すると、バックアップマップ48のマップ場所142により、時刻 T_1 にスナップショットがとられるのに先立って記憶場所3、4、及び6が変更されていることが示される。スナップショット場所152の検査により、データブロック4及び6が大容量記憶システム上にあり、データブロック3は、保護されるものとしてマーク付けされている場合、スナップショットストレージ22の中にあることが示される(しかし、前述したとおり、データブロック3は、保護されないものとしてマップ場所129においてマーク付けされているため、スナップショットストレージ22の中に格納されていない。)。次に、図5のステップ104により、記憶場所3、4、及び6に格納されたデータブロックが図3のバックアップ読取り処理ブロック56によって取り出されることが要求される。

20

【0107】

バックアップ読取り処理ブロック56は、図6に示すとおり、1次バックアップ処理ブロック54から受け取られたその要求を処理する。要求は、記憶場所3、4、及び6に格納されたデータブロックに関する。記憶場所3に格納されたデータブロックに関して、判定ブロック113により、そのデータブロックが保護されないようにマーク付けされていると判定され、したがって、このデータブロックは、ステップ120で戻されない。場所4及び6に格納されたデータブロックは、保護されないようにマーク付けされていないため、図6の判定ブロック114により、ステップ118で、元のデバイスから記憶場所4及び6に格納されたデータブロックが取り出され、ステップ120で、1次バックアップ処理ブロック54に戻される。このプロセスを図8Aに示しており、記憶場所150からデータブロック4b及び6aを取り出すことによってデータブロック153がアセンブルされる。次に、データブロック153は、大容量ストレージ読取り/書込み処理ブロック42を介してスナップショットコピーに転送される。次に、データブロック153が記憶場所124に印加されて、記憶場所154が達せられる。記憶場所154は、バックアップが所望されないものとして特定されているために除外されている記憶場所3を除き、1次システム(図8A)の記憶場所138と全く同じである。記憶場所138は、スナップショットがとられた時刻 T_1 における大容量記憶デバイス20の状態を表していたことを思い起こされたい。したがって、時刻 T_0 と時刻 T_1 の間に行われた変更は、スナップショットコピーを時刻 T_1 における元のデータと同期させるために、現時点で、スナップショットコピー16にバックアップ済みである。

30

40

【0108】

次に、図8Aに戻ると、データブロック156が次に、記憶場所150に書き込まれるものと想定する。データブロック156の書込みにより、記憶場所1、4、及び6に格納されたデータブロックに対する変更が生じる。大容量ストレージ読取り/書込み処理ブロック42が、場所4及び6に格納されるべきデータブロックの書込みを、時刻 T_1 後にこれらの場所に格納されるデータブロック144に関して前述したのと同様に処理する(データブロック4b及びデータブロック6aが、スナップショットストレージ22に格納されて)。次に、新しいデータブロック4c及び6bが、元のデータ14に書き込まれる。

50

【 0 1 0 9 】

記憶場所 1 に格納されるべきデータブロックに関して、実行は、図 4 で、判定ブロック 8 4 まで進む。この判定ブロックは、スナップショットがとられた時点で記憶場所に格納されているデータブロックが、スナップショットストレージの中に以前に保存済みであるかどうかをテストすることを思い起こされたい。記憶場所 1 に格納されたデータブロックは、図 8 A のデータブロック 1 4 6 によってスナップショットストレージ 2 2 の中に以前に格納済みである。したがって、図 4 は、ステップ 8 6 がスキップされ、新しいデータが単に元のデータに書き込まれることを示している。図 8 A で、これにより、データブロック 1 b がデータブロック 1 a に取って代わり、このデータブロックが失われることになる。データブロック 1 5 6 が記憶場所 1 5 0 に印加された際、記憶場所 1 5 8 がもたらされる。次に、マップ場所 1 5 2 が更新されて、時刻 T_1 以来、変更されている記憶場所に、現時点で、記憶場所 1 及び 3 に加えて記憶場所 4 及び 6 が含まれることを示す。これを図 8 A にスナップショットストレージ 5 2 のマップ場所 1 6 0 で示している。

10

【 0 1 1 0 】

次に、大容量記憶デバイス 2 0 の第 2 のバックアップが行われるものと想定する。その場合、バックアップは、図 5 で上述したとおり行われ、実行は、論理的に整合性のある状態が特定されるステップ 1 0 0 に進む。図 8 A で、時刻 T_2 に論理的に整合性のある状態が特定されたものと想定する。次に、図 5 のステップ 1 0 2 により、時刻 T_2 にスナップショットがとられるべきことが通知される。時刻 T_1 にとられるスナップショットに関連して前述したとおり、大容量ストレージ読取り / 書込み処理ブロック 4 2 が、図 3 のスナップショット要求 6 8 のようなスナップショット要求を受け取り、ステップ 7 0 で、スナップショットマップをバックアップマップにコピーする。これを図 8 A に示しており、バックアップマップ 4 8 のマップ場所 1 6 2 が、スナップショットマップ 5 2 のマップ場所 1 6 0 と同じになるように変更されている。

20

【 0 1 1 1 】

次に、図 4 のステップ 7 2 及び 7 4 により、スナップショットマップ及びスナップショットストレージがクリアされるべきことが示される。図 8 A で、スナップショットマップ 5 2 のマップ場所 1 6 4 によって示されるとおり、スナップショットマップがクリアされる。ただし、スナップショットストレージ 2 2 には依然としてデータブロックが格納されているのが示されている。これは、スナップショットストレージ 2 2 に対するインデックスがクリアされて、スナップショットストレージ 2 2 が全くデータブロックを含くまないように見える限り、データブロックは、スナップショットストレージ 2 2 の中に依然として物理的に存在していてもよいことを示すためである。

30

【 0 1 1 2 】

時刻 T_2 にスナップショットがとられた後に記憶場所 1 5 8 の中に全くデータブロックが存在しないものと想定すると、図 6 で説明したプロセスに従ってデータブロック 1 6 6 が、記憶場所 1 5 8 から読み取られる。図 5 の判定ブロック 1 1 3 が保護マップ場所 1 2 9 を使用して場所 3 が保護されないことを判定するため、記憶場所 1 5 8 の場所 3 は読み取られないことに留意されたい。したがって、データブロック 3 c が読み取られて、バックアップシステムに転送されることはない。図 5 のステップ 1 0 4 及び 1 0 6 に示すとおり、読み取られたデータブロックは、次に、アセンブルされ、大容量記憶読取り / 書込み処理ブロック 4 2 を介してスナップショットコピーに送られる。データブロック 1 6 6 は、記憶場所 1 6 8 に達するために記憶場所 1 5 4 に印加され、記憶場所 1 6 8 は、バックアップが所望されないものとしてユーザによって特定されているために除外される記憶場所 3 を除き、元のデータ (図 8 A) の記憶場所 1 5 8 と全く同じコピーである。

40

【 0 1 1 3 】

上述したように、変更されており、保護されるように指定されているデータブロックだけが、バックアップされる。図 4 のステップ 8 5 で、データブロックが保護されるように指定されていない場合、データブロックは、元のデータブロックをスナップショットストレージに書き込むことなく、単に上書きされることが可能であり、これにより、スナップ

50

ショットストレージメモリが節約される。また、データブロックは、保護されるものとして指定される場合にだけバックアップシステムに伝送されるので、より少ないデータがバックアップシステムに伝送されることにより、バックアップ時間がより高速になる。

【0114】

本発明は、本発明の趣旨または基本的特徴を逸脱することなく、他の特定の形態で実施することができる。説明した実施形態は、すべての点で単に例示的であり、これに限定するものではない。したがって、本発明の技術的範囲は、以上の説明によってではなく、特許請求の範囲によって示される。特許請求の範囲と等価の意味及び範囲に含まれるすべての変更が、特許請求の範囲に含まれるものである。

【図面の簡単な説明】

10

【0115】

【図1】本発明のバックアップシステムを説明するためのブロック図である。

【図2】本発明のバックアップ方法を説明するためのタイミングを示す図である。

【図3】本発明のバックアップシステムの一実施形態を示すシステムレベルのブロック図である。

【図4】図3に示した大容量ストレージ読取り/書込み処理ブロックの処理フローを示す図である。

【図5】図3に示した1次バックアップ処理ブロックの処理フローを示す図である。

【図6】図3に示したバックアップ読取り処理ブロックの処理フローを示す図である。

【図7A】本発明のバックアップ方法の一実施形態を示す図(その1)である。

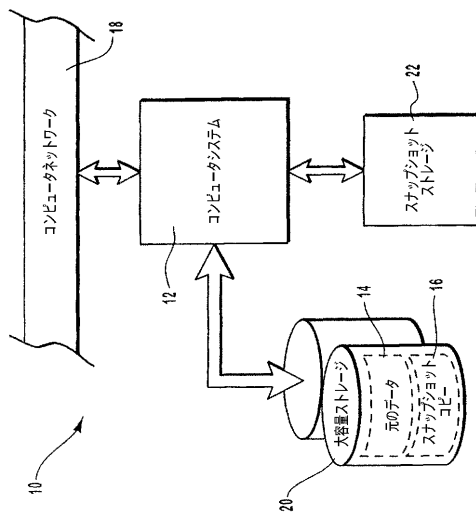
20

【図7B】本発明のバックアップ方法の一実施形態を示す図(その2)である。

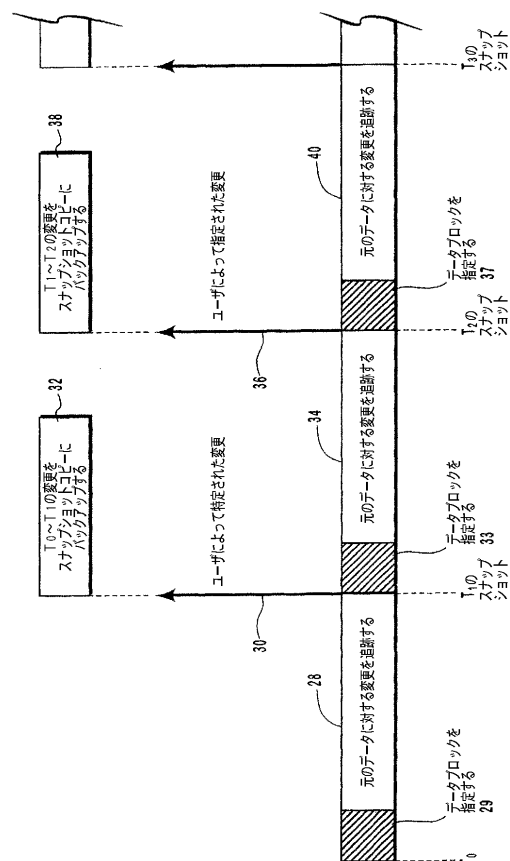
【図8A】本発明のバックアップ方法の一実施形態を示す図(その3)である。

【図8B】本発明のバックアップ方法の一実施形態を示す図(その3)である。

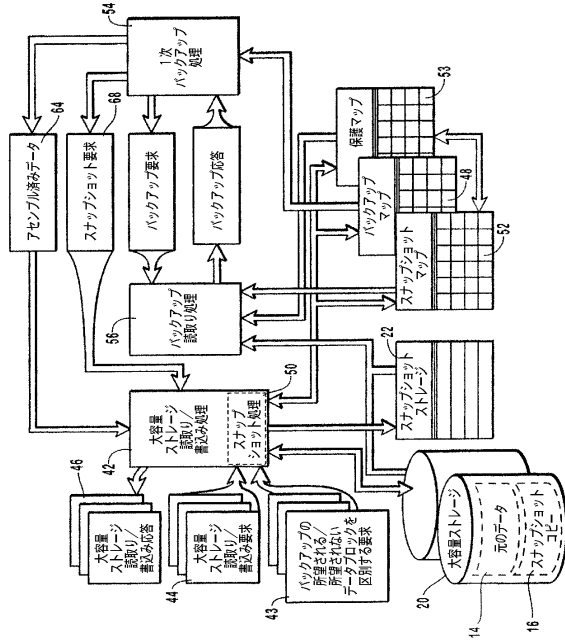
【図1】



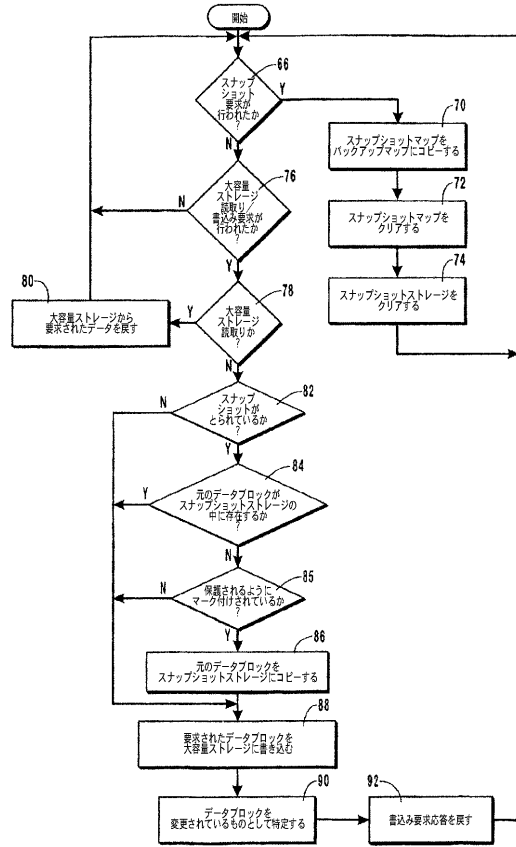
【図2】



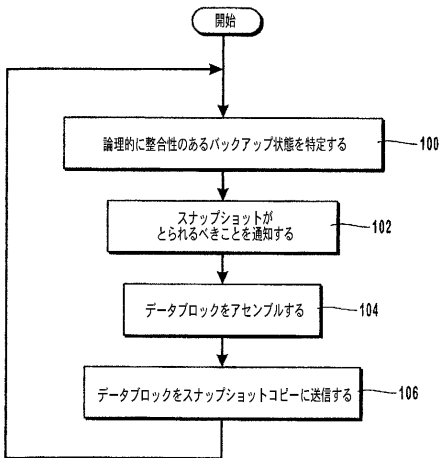
【図3】



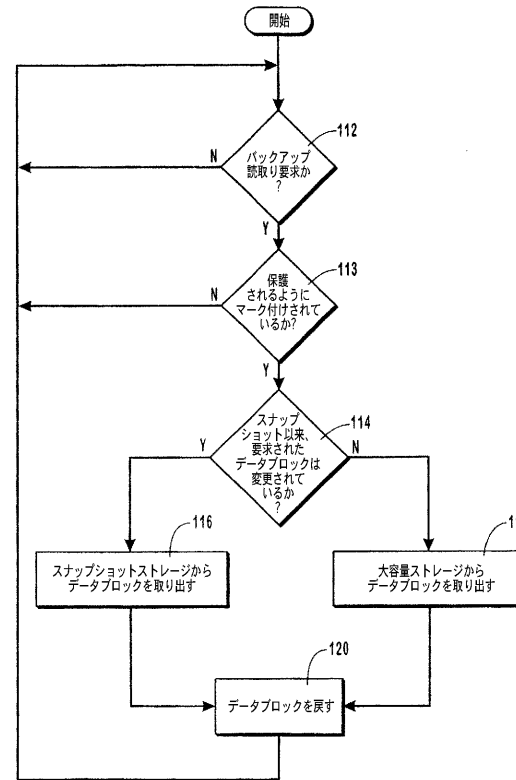
【図4】



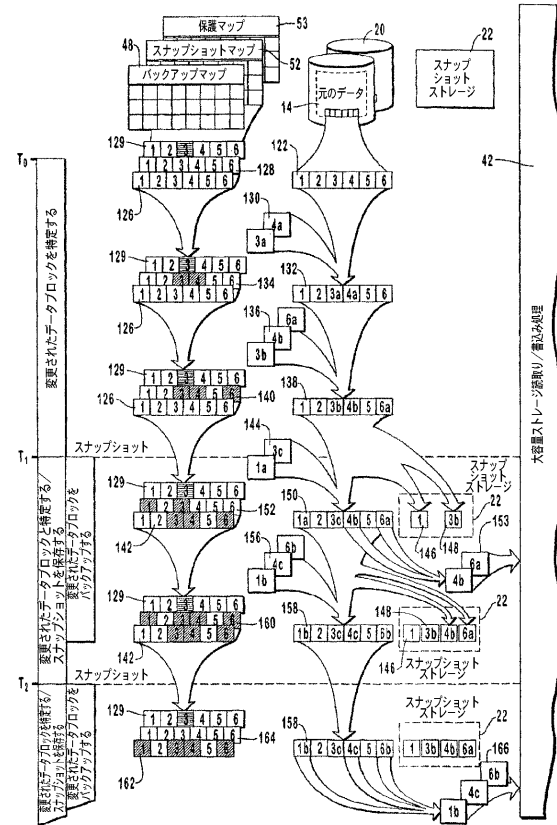
【図5】



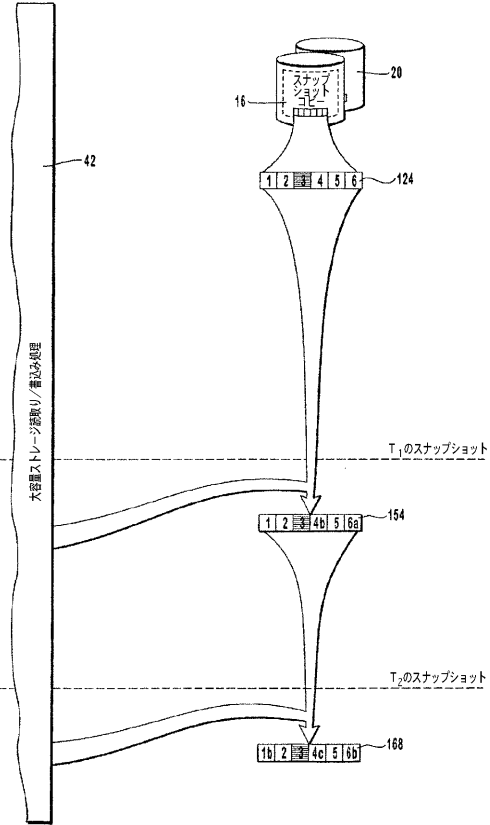
【図6】



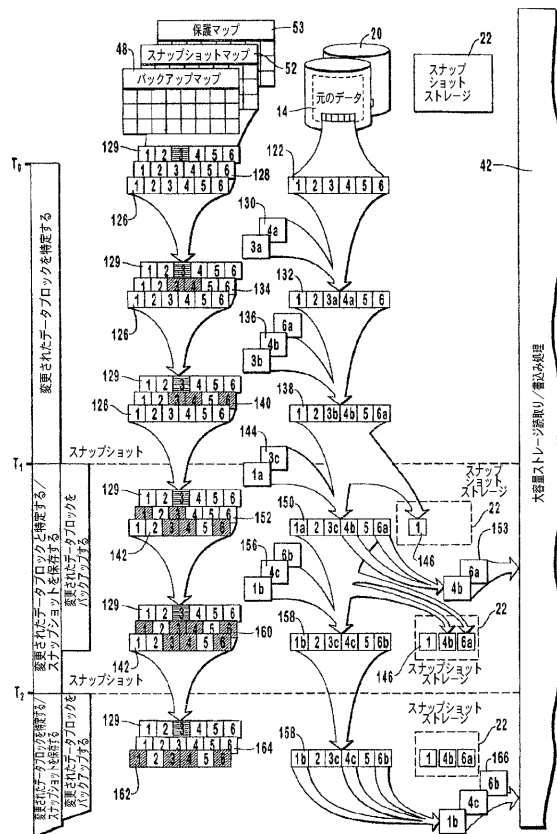
【図7A】



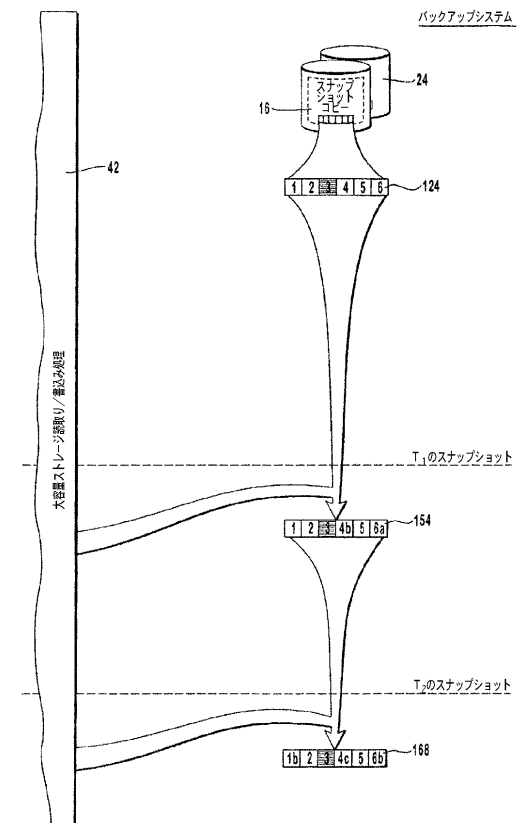
【図7B】



【図8A】



【図8B】



フロントページの続き

審査官 廣瀬 文雄

- (56)参考文献 特開平07-200367(JP,A)
特開2000-347811(JP,A)
特開平11-120055(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00

G06F 3/06