

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4806168号
(P4806168)

(45) 発行日 平成23年11月2日(2011.11.2)

(24) 登録日 平成23年8月19日(2011.8.19)

(51) Int.Cl.

G06F 12/00 (2006.01)

F 1

G06F 12/00 512
G06F 12/00 517
G06F 12/00 518A

請求項の数 12 (全 20 頁)

(21) 出願番号 特願2004-152121 (P2004-152121)
 (22) 出願日 平成16年5月21日 (2004.5.21)
 (65) 公開番号 特開2004-348744 (P2004-348744A)
 (43) 公開日 平成16年12月9日 (2004.12.9)
 審査請求日 平成19年5月21日 (2007.5.21)
 (31) 優先権主張番号 10/445,095
 (32) 優先日 平成15年5月22日 (2003.5.22)
 (33) 優先権主張国 米国(US)

前置審査

(73) 特許権者 500046438
 マイクロソフト コーポレーション
 アメリカ合衆国 ワシントン州 9805
 2-6399 レッドmond ワン マイ
 クロソフト ウェイ
 (74) 代理人 110001243
 特許業務法人 谷・阿部特許事務所
 (74) 復代理人 100115624
 弁理士 濱中 淳宏
 (74) 復代理人 100162950
 弁理士 久下 範子
 (72) 発明者 ガナバシー クリシュナムーアシー
 アメリカ合衆国 98029 ワシントン
 州 イサコア 249 アベニュー サウ
 スイースト 4033

最終頁に続く

(54) 【発明の名称】テーブルに対して行われる変更を識別する識別方法及びそのシステム

(57) 【特許請求の範囲】

【請求項 1】

選択されたトランザクションの一部としてのデータ操作ステートメントによって基本テーブルから削除された行を識別するため、少なくとも一部はコンピュータによって実装された方法であって、

マーカトランザクション番号及びマーカシーケンス番号を有する削除マーカを生成するステップであって、前記基本テーブルは、複数の行を有し、前記各行は、行トランザクション番号および行シーケンス番号を有し、前記行トランザクション番号は、前記各行の現在のバージョンがコミットされたトランザクションを識別する番号であり、前記行シーケンス番号は、トランザクションにおいて変更が行われたシーケンスを識別する番号であり

前記マーカトランザクション番号は、前記基本テーブルの前記選択されたトランザクションの前記トランザクション番号が設定され、前記マーカシーケンス番号は、前記ステートメントの実行の前に、前記選択されたトランザクションの一部として変更された最後の行のシーケンス番号が設定されるステップと、

前記ステートメントの実行の前に、前記トランザクションの一部として変更された前記最後の行により設定された前記削除マーカで前記基本テーブル内の各行を走査するステップであって、

前記削除マーカを用いて前記基本テーブル内の各行を走査するステップは、

(a) 前記行トランザクション番号が、前記マーカトランザクション番号に等しいか否かを判定するステップと、

10

20

(b) 前記行トランザクション番号が、前記マーカトランザクション番号に等しくない場合に、前記行を読み出すステップと、

(c) 前記行トランザクション番号が、前記マーカトランザクション番号に等しい場合に、

前記行シーケンス番号が、前記マーカシーケンス番号以下であるか否かを判定するステップと、

前記マーカシーケンス番号以下である場合に、前記行を読み出すステップと、

前記マーカシーケンス番号以下でない場合に、

前記行の前のバージョンを取り出すステップと、

前記ステップ(a)に戻り、前記行の前のバージョンに対して処理を続けるステップと有するステップと、

デルタテーブルに格納された削除のリストで前記走査の結果をフィルタリングするステップであって、前記デルタテーブルには、前記ステートメントにより実施された削除のリストが格納されているステップと、

前記フィルタリングした前記基本テーブルから選択された各行のリストを削除テーブルに投入するステップと

を有することを特徴とする方法。

【請求項2】

前記行の前のバージョンを取り出すステップは、バージョン記憶装置から前記行の前のバージョンを取り出すステップを有し、前記前のバージョンは、前記バージョン記憶装置において、前記行トランザクション番号及び前記行シーケンス番号によって識別されることを特徴とする請求項1に記載の方法。

【請求項3】

選択されたトランザクションの一部としてのデータ操作ステートメントによって基本テーブルに挿入された行を識別するための、少なくとも一部はコンピュータによって実装された方法であって、

マーカトランザクション番号及びマーカシーケンス番号を有する挿入マーカを生成するステップであって、前記基本テーブルは、複数の行を有し、前記各行は、行トランザクション番号および行シーケンス番号を有し、前記行トランザクション番号は、前記各行の現在のバージョンがコミットされたトランザクションを識別する番号であり、前記行シーケンス番号は、トランザクションにおいて変更が行われたシーケンスを識別する番号であり
前記マーカトランザクション番号は、前記基本テーブルの前記選択されたトランザクションの前記トランザクション番号が設定され、前記マーカシーケンス番号は、前記ステートメントの実行の間に、前記選択されたトランザクションの一部として変更された前記最後の行のシーケンス番号が設定されるステップと、

前記ステートメントの実行の間に、前記トランザクションの一部として変更された最後の行により設定された前記挿入マーカで前記基本テーブル内の各行を走査するステップであって、

前記挿入マーカを用いて前記基本テーブル内の各行を走査するステップは、

(a) 前記行トランザクション番号が、前記マーカトランザクション番号に等しいか否かを判定するステップと、

(b) 前記行トランザクション番号が、前記マーカトランザクション番号に等しくない場合に、前記行を読み出すステップと、

(c) 前記行トランザクション番号が、前記マーカトランザクション番号に等しい場合に、

前記行シーケンス番号が、前記マーカシーケンス番号以下であるか否かを判定するステップと、

前記マーカシーケンス番号以下である場合に、前記行を読み出すステップと、

前記マーカシーケンス番号以下でない場合に、

前記行の前のバージョンを取り出すステップと、

10

20

30

40

50

前記ステップ(a)に戻り、前記行の前のバージョンに対して処理を続けるステップと有するステップと、

デルタテーブルに格納された挿入のリストで前記走査の結果をフィルタリングするステップであって、前記デルタテーブルには、前記ステートメントにより実施された挿入のリストが格納されているステップと、

前記フィルタリングした前記基本テーブルから選択された各行のリストを挿入テーブルに投入するステップと

を有することを特徴とする方法。

【請求項 4】

前記行の前のバージョンを取り出すステップは、バージョン記憶装置から前記行の前のバージョンを取り出すステップを有し、前記前のバージョンは、前記バージョン記憶装置において、前記行トランザクション番号及び前記行シーケンス番号によって識別されることを特徴とする請求項3に記載の方法。 10

【請求項 5】

選択されたトランザクションの一部としてのデータ操作ステートメントによって基本テーブルから削除された行を識別する識別システムであって、

コンピュータ実行可能命令を実行するように動作可能なプロセッサと、

コンピュータ実行可能命令を格納したメモリとを備え、

前記命令は、

マーカトランザクション番号及びマーカシーケンス番号を有する削除マーカを生成するステップであって、前記基本テーブルは、複数の行を有し、前記各行は、行トランザクション番号および行シーケンス番号を有し、前記行トランザクション番号は、前記各行の現在のバージョンがコミットされたトランザクションを識別する番号であり、前記行シーケンス番号は、トランザクションにおいて変更が行われたシーケンスを識別する番号であり、前記マーカトランザクション番号は、前記基本テーブルの前記選択されたトランザクションの前記トランザクション番号が設定され、前記マーカシーケンス番号は、前記ステートメントの実行の前に、前記選択されたトランザクションの一部として変更された最後の行のシーケンス番号が設定されるステップと、 20

前記ステートメントの実行の前に、前記トランザクションの一部として変更された前記最後の行により設定された前記削除マーカで前記基本テーブル内の各行を走査するステップであって、 30

前記削除マーカを用いて前記基本テーブル内の各行を走査するステップは、

(a) 前記行トランザクション番号が、前記マーカトランザクション番号に等しいか否かを判定するステップと、

(b) 前記行トランザクション番号が、前記マーカトランザクション番号に等しくない場合に、前記行を読み出すステップと、

(c) 前記行トランザクション番号が、前記マーカトランザクション番号に等しい場合に、

前記行シーケンス番号が、前記マーカシーケンス番号以下であるか否かを判定するステップと、 40

前記マーカシーケンス番号以下である場合に、前記行を読み出すステップと、

前記マーカシーケンス番号以下でない場合に、

前記行の前のバージョンを取り出すステップと、

前記ステップ(a)に戻り、前記行の前のバージョンに対して処理を続けるステップと有するステップと、

デルタテーブルに格納された削除のリストで前記走査の結果をフィルタリングするステップであって、前記デルタテーブルは、前記ステートメントにより実施された削除のリストが格納されているステップと、

前記フィルタリングした前記基本テーブルから選択された各行のリストを削除テーブルに投入するステップと 50

を有することを特徴とする識別システム。

【請求項 6】

前記行の前のバージョンは、バージョン記憶装置において、前記行トランザクション番号及び前記行シーケンス番号によって識別されることを特徴とする請求項5に記載の識別システム。

【請求項 7】

選択されたトランザクションの一部としてのデータ操作ステートメントによって基本テーブルに挿入された行を識別する識別システムであって、

コンピュータ実行可能命令を実行するように動作可能なプロセッサと、

コンピュータ実行可能命令を格納したメモリとを備え、

前記命令は、

マーカトランザクション番号及びマーカシーケンス番号を有する挿入マーカを生成するステップであって、前記基本テーブルは、複数の行を有し、前記各行は、行トランザクション番号および行シーケンス番号を有し、前記行トランザクション番号は、前記各行の現在のバージョンがコミットされたトランザクションを識別する番号であり、前記行シーケンス番号は、トランザクションにおいて変更が行われたシーケンスを識別する番号であり、前記マーカトランザクション番号は、前記基本テーブルの前記選択されたトランザクションの前記トランザクション番号が設定され、前記マーカシーケンス番号は、前記ステートメントの実行の間に、前記選択されたトランザクションの一部として変更された最後の行のシーケンス番号が設定されるステップと、

前記ステートメントの実行の間に、前記トランザクションの一部として変更された前記最後の行により設定された前記挿入マーカで前記基本テーブル内の各行を走査するステップであって、

前記挿入マーカを用いて前記基本テーブル内の各行を走査するステップは、

(a) 前記行トランザクション番号が、前記マーカトランザクション番号に等しいか否かを判定するステップと、

(b) 前記行トランザクション番号が、前記マーカトランザクション番号に等しくない場合に、前記行を読み出すステップと、

(c) 前記行トランザクション番号が、前記マーカトランザクション番号に等しい場合に、

前記行シーケンス番号が、前記マーカシーケンス番号以下であるか否かを判定するステップと、

前記マーカシーケンス番号以下である場合に、前記行を読み出すステップと、

前記マーカシーケンス番号以下でない場合に、

前記行の前のバージョンを取り出すステップと、

前記ステップ(a)に戻り、前記行の前のバージョンに対して処理を続けるステップとを有するステップと、

デルタテーブルに格納された挿入のリストで前記走査の結果をフィルタリングするステップであって、前記デルタテーブルには、前記ステートメントにより実施された挿入のリストが格納されているステップと、

前記フィルタリングした前記基本テーブルから選択された各行のリストを挿入テーブルに投入するステップと

を有することを特徴とする識別システム。

【請求項 8】

前記行の前のバージョンは、バージョン記憶装置において、前記行トランザクション番号及び前記行シーケンス番号によって識別されることを特徴とする請求項7に記載の識別システム。

【請求項 9】

選択されたトランザクションの一部としてのデータ操作ステートメントによって基本テーブルから削除された行を識別するための、コンピュータ実装可能命令を含むコンピュー

10

20

30

40

50

タ可読記録媒体であって、

前記コンピュータ可読記録媒体は、プロセッサにより実行されたときに、以下のステップを実施するコンピュータ可読命令を格納し、前記命令は、

マーカトランザクション番号及びマーカシーケンス番号を有する削除マーカを生成するステップであって、前記基本テーブルは、複数の行を有し、前記各行は、行トランザクション番号および行シーケンス番号を有し、前記行トランザクション番号は、前記各行の現在のバージョンがコミットされたトランザクションを識別する番号であり、前記行シーケンス番号は、トランザクションにおいて変更が行われたシーケンスを識別する番号であり、前記マーカトランザクション番号は、前記基本テーブルの前記選択されたトランザクションの前記トランザクション番号が設定され、前記マーカシーケンス番号は、前記ステートメントの実行の前に、前記選択されたトランザクションの一部として変更された最後の行のシーケンス番号が設定されるステップと、

前記ステートメントの実行の前に、前記トランザクションの一部として変更された前記最後の行により設定された前記削除マーカで前記基本テーブル内の各行を走査するステップであって、

前記削除マーカを用いて前記基本テーブル内の各行を走査するステップは、

(a) 前記行トランザクション番号が、前記マーカトランザクション番号に等しいか否かを判定するステップと、

(b) 前記行トランザクション番号が、前記マーカトランザクション番号に等しくない場合に、前記行を読み出すステップと、

(c) 前記行トランザクション番号が、前記マーカトランザクション番号に等しい場合に、

前記行シーケンス番号が、前記マーカシーケンス番号以下であるか否かを判定するステップと、

前記マーカシーケンス番号以下である場合に、前記行を読み出すステップと、

前記マーカシーケンス番号以下でない場合に、

前記行の前のバージョンを取り出すステップと、

前記ステップ(a)に戻り、前記行の前のバージョンに対して処理を続けるステップとを有するステップと、

デルタテーブルに格納された削除のリストで前記走査の結果をフィルタリングするステップであって、前記デルタテーブルには、前記ステートメントにより実施された削除のリストが格納されているステップと、

前記フィルタリングした前記基本テーブルから選択された各行のリストを削除テーブルに投入するステップと

を含むことを特徴とするコンピュータ可読記録媒体。

【請求項10】

前記行の前のバージョンを取り出すステップは、バージョン記憶装置から前記行の前のバージョンを取り出すステップを有し、前記前のバージョンは、前記バージョン記憶装置において、前記行トランザクション番号及び前記行シーケンス番号によって識別されることを特徴とする請求項9に記載のコンピュータ可読記録媒体。

【請求項11】

選択されたトランザクションの一部としてのデータ操作ステートメントによって基本テーブルから挿入された行を識別するための、コンピュータ実装可能命令を含むコンピュータ可読記録媒体であって、

前記コンピュータ可読記録媒体は、プロセッサにより実行されたときに、以下のステップを実施するコンピュータ可読命令を格納し、前記命令は、

マーカトランザクション番号及びマーカシーケンス番号を有する挿入マーカを生成するステップであって、前記基本テーブルは、複数の行を有し、前記各行は、行トランザクション番号および行シーケンス番号を有し、前記行トランザクション番号は、前記各行の現在のバージョンがコミットされたトランザクションを識別する番号であり、前記行シーケンス番号は、トランザクションにおいて変更が行われたシーケンスを識別する番号であり、前記マーカトランザクション番号は、前記基本テーブルの前記選択されたトランザクションの前記トランザクション番号が設定され、前記マーカシーケンス番号は、前記ステートメントの実行の前に、前記選択されたトランザクションの一部として変更された最後の行のシーケンス番号が設定されるステップと、

10

20

30

40

50

ンス番号は、トランザクションにおいて変更が行われたシーケンスを識別する番号であり、前記マーカトランザクション番号は、前記基本テーブルの前記選択されたトランザクションの前記トランザクション番号が設定され、前記マーカシーケンス番号は、前記ステートメントの実行の間に、前記選択されたトランザクションの一部として変更された最後の行のシーケンス番号が設定されるステップと、

前記ステートメントの実行の間に、前記トランザクションの一部として変更された前記最後の行により設定された前記挿入マーカで前記基本テーブル内の各行を走査するステップであって、

前記挿入マーカを用いて前記基本テーブル内の各行を走査するステップは、

(a) 前記行トランザクション番号が、前記マーカトランザクション番号に等しいか否かを判定するステップと、

(b) 前記行トランザクション番号が、前記マーカトランザクション番号に等しくない場合に、前記行を読み出すステップと、

(c) 前記行トランザクション番号が、前記マーカトランザクション番号に等しい場合に、

前記行シーケンス番号が、前記マーカシーケンス番号以下であるか否かを判定するステップと、

前記マーカシーケンス番号以下である場合に、前記行を読み出すステップと、

前記マーカシーケンス番号以下でない場合に、

前記行の前のバージョンを取り出すステップと、

前記ステップ(a)に戻り、前記行の前のバージョンに対して処理を続けるステップとを有するステップと、

デルタテーブルに格納された挿入のリストで前記走査の結果をフィルタリングするステップであって、前記デルタテーブルには、前記ステートメントにより実施された挿入のリストが格納されているステップと、

前記フィルタリングした前記基本テーブルから選択された各行のリストを挿入テーブルに投入するステップと

を含むことを特徴とするコンピュータ可読記録媒体。

【請求項 1 2】

前記行の前のバージョンを取り出すステップは、バージョン記憶装置から前記行の前のバージョンを取り出すステップを有し、前記前のバージョンは、前記バージョン記憶装置において、前記行トランザクション番号及び前記行シーケンス番号によって識別されることを特徴とする請求項 1 1 に記載のコンピュータ可読記録媒体。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、データベース管理の分野を対象とし、テーブルに対して行われる変更を識別する識別方法及びそのシステムに関し、より詳細には、基本テーブルからの削除及び基本テーブルへの挿入される行 (row) の識別を行なう識別方法及びそのシステム並びにこのような変更の格納を行なう記録媒体に関する。

【背景技術】

【0 0 0 2】

データベース管理システム (DBMS) では、データの現在のバージョンは、概して「基本」テーブルに格納される。基本テーブル中の各行は、変更することができる。このような変更は、基本テーブル中の行の削除でも、挿入でも、修正でもよい。修正とは、行の前のバージョンの削除及び行の現在のバージョンの挿入である。行が変更されると、その変更は、1組の一時的な、メモリに常駐する「遷移 (transition)」テーブルに格納することができる。このような1組の遷移テーブルは、概して自動的に作成され、DBMS によって管理される。遷移テーブルの組は、基本テーブルから削除された行を格納する「削除」テーブルと、基本テーブルに挿入された行を格納する「挿入」テーブルと

10

20

30

40

50

を含むことができる。行が修正されると、行の前のバージョンは削除テーブルに格納され、行の現在のバージョンは挿入テーブルに格納される。

【0003】

遷移テーブルにより、DBMSは、変更の影響を調べ、トリガ動作の条件を設定することが可能になる。具体的には、遷移テーブルは、テーブル間の参照整合性を拡大し、ある特定のビューの基礎をなす基本テーブル中のデータを変更し、違いを検出し、検出した違いに基づいて動作を起こし、また、変更の前と後でのテーブルの状態の違いを見つけ、違いに基づいて動作を起こす。

【0004】

従来のDBMSでは、基本テーブル中の行が変更される前にそのような行を完全にコピーすることによって、遷移テーブルに投入を行うことができる。したがって、例えば、行が削除されると、その行の前のバージョンがコピーされ、このコピーは次いで、削除テーブルに格納される。行をコピーすることの欠点は、特にコピーされた行が、例えば、コードや画像など大規模フィールドのバイナリラージオブジェクト(BLOB)データを含む場合、このようなコピーを生成して格納するのにコストがかかることがある。

10

【0005】

遷移テーブルに投入を行う他の方法は、ログファイルを走査することである。ただし、このようなログの走査を用いることの欠点は、BLOBデータはしばしば、ログの走査の使用を実装するのには複雑すぎることである。さらに、ログの走査は、ディスク入出力(I/O)の使用を必要とする。

20

【発明の開示】

【発明が解決しようとする課題】

【0006】

したがって、当該分野において、テーブルに対して行われる変更を効率的に識別して格納するシステム及び方法が必要である。このようなシステム及び方法は、大量のBLOBデータを含む行を低コストで識別し、適切な遷移テーブルに格納できるようにすることが望ましい。このようなシステム及び方法は、ディスクI/Oを非効率に使用することなく動作することがさらに望ましい。

【0007】

本発明は、このような問題に鑑みてなされたもので、その目的とするところは、テーブルに対して行われる変更を効率的に識別して格納するようにした識別方法及びそのシステムを提供することにある。

30

【課題を解決するための手段】

【0008】

したがって、本発明は、テーブルに対して行われる変更を識別して格納するシステム及びその方法を開示する。データの現在のバージョンは、基本テーブルに格納される。基本テーブル中の各行の前のバージョンは、バージョン記憶装置に格納される。デルタテーブルが、基本テーブルに対して行われた変更の一次キーを識別し、このような変更が挿入であるかまたは削除であるか指定する。1組の遷移テーブルが、基本テーブルに対して行われた各変更を格納する。遷移テーブルの組は、基本テーブルから削除された行を格納する削除テーブルと、基本テーブルに挿入された行を格納する挿入テーブルとを含むことができる。

40

【0009】

行が基本テーブルにコミットされると、その行には、好ましくはトランザクション(transaction)識別子(XID)及びシーケンス番号(SEQ)が割り当てられる。XIDは、変更が行われるトランザクションの一義的な識別子である。SEQは、好ましくはトランザクションにおいて行われる各変更に割り当てられる、単調に増加し続ける値である。バージョン記憶装置内にある行の前のバージョンは、基本テーブル中の行の現在のバージョンのXID及びSEQによって識別することができる。

【0010】

50

遷移テーブルには、トランザクションの一部としてのステートメント (statement) によって変更された行を識別することによって、投入を行うことができる。基本テーブルは、削除マーカ及び挿入マーカを用いて走査される。削除マーカは、ステートメントを実行する前にトランザクションの一部として基本テーブルにコミットされた最終の変更を識別し、挿入マーカは、ステートメントの実行中にトランザクションの一部として基本テーブルにコミットされた最終の変更を識別する。削除マーカの走査の結果は、デルタテーブルに格納されている削除を用いてフィルタリングされ、削除テーブルには、フィルタリングされた行が投入される。挿入走査の結果は、デルタテーブルに格納されている挿入を用いてフィルタリングされ、挿入テーブルには、フィルタリングされた行が投入される。

10

【発明を実施するための最良の形態】

【0011】

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。

なお、本明細書に示す説明は例示のために過ぎず、本発明の技術的範囲を限定するものではない。また、明細書中、同じ構成要素については、同一の符号を付してある。

【0012】

<コンピュータ環境>

図1は、本発明に係る識別システムの一実施形態を説明するためのブロック図で、実施形態及び/またはその部分を組み込むことができる汎用コンピュータシステムを示すブロック図である。

20

【0013】

以下の説明では、本発明及び/またはその一部を実装することができる適切な計算機環境の簡潔かつ全体的な説明を提供することを意図している。また、クライアントワークステーションやサーバなどのコンピュータによって実行されるプログラムモジュールなどのコンピュータ実行可能命令という一般的な状況において本発明を説明する。プログラムモジュールは、特定のタスクを実施し、または特定の抽象データタイプを実装するルーチン、プログラム、オブジェクト、構成要素、データ構造などを備えている。さらに、本発明及び/またはその一部は、可搬型装置、マルチプロセッサシステム、マイクロプロセッサベースの家電製品またはプログラム可能な家電製品、ネットワークPC、ミニコンピュータ、メインフレームコンピュータなど他のコンピュータシステム構成と共に実施できる。本発明は、通信ネットワークを介してリンクされるリモート処理装置によって特定のタスクが実施される分散型計算機環境でも実施することができる。分散型計算機環境では、プログラムモジュールは、ローカル及びリモートメモリ記憶装置の両方に置くことができる。

30

【0014】

図1に示すように、例示的な汎用計算機システムは、従来のパーソナルコンピュータ120などを備え、このコンピュータ120は、処理装置121とシステムメモリ122とシステムメモリなど様々なシステム構成要素を処理装置121に結合するシステムバス123とを備えている。システムバス123は、様々なバスアーキテクチャのどれをも使用するメモリバスまたはメモリコントローラ、周辺バス、及びローカルバスなどいくつかのタイプのバス構造のいずれでもよい。システムメモリは、読み出し専用メモリ(ROM)124及びランダムアクセスメモリ(RAM)125を備えている。基本入出力システム126(BIOS)は、例えば、起動中にパーソナルコンピュータ120内部の要素間の情報の転送を助ける基本ルーチンを備え、ROM124に格納されている。

40

【0015】

パーソナルコンピュータ120は、ハードディスク(図示せず)からの読み出し及びそこへの書き込みを行うハードディスクドライブ127と、取外し可能磁気ディスク129からの読み出しありはそこへの書き込みを行う磁気ディスクドライブ128と、例えば、CD-ROMや他の光学媒体などの取外し可能光ディスク131からの読み出しありはそこへの書き込みを行う光ディスクドライブ130をさらに備えている。ハードディスクドライブ127

50

と磁気ディスクドライブ 128 と光ディスクドライブ 130 は、それぞれハードディスクドライブインターフェース 132 と磁気ディスクドライブインターフェース 133 と光ドライブインターフェース 134 によって、システムバス 123 に接続されている。こうしたドライブ及びそれに関連するコンピュータ読取可能な記録媒体は、コンピュータ可読命令、データ構造、プログラムモジュール、及び他のデータの不揮発性記憶装置をパーソナルコンピュータ 120 に提供する。

【0016】

本明細書で説明する例示的な環境では、ハードディスクと、取外し可能な磁気ディスク 129 と、取外し可能な光ディスク 131 を使用するが、コンピュータによるアクセスが可能なデータを格納することができる、コンピュータ可読な他のタイプの媒体も、例示的な動作環境において使うことができる。このような他のタイプの媒体には、磁気カセット、フラッシュメモリカード、デジタルビデオディスク、ベルヌーイカートリッジ、ランダムアクセスメモリ (RAM) 、読み出し専用メモリ (ROM) などがある。

【0017】

オペレーティングシステム 135 と、1つまたは複数のアプリケーションプログラム 136 と、他のプログラムモジュール 137 と、プログラムデータ 138 などいくつかのプログラムモジュールを、ハードディスクと、磁気ディスク 129 と、光ディスク 131 と、ROM 124 または RAM 125 に格納することができる。ユーザは、キーボード 140 などの入力装置、及びマウスなどの指示装置 142 を介して、コマンド及び情報をパーソナルコンピュータ 120 に入力することができる。他の入力装置 (図示せず) には、マイクロホン、ジョイスティック、ゲーム用パッド、衛星パラボラアンテナ、スキャナなどがあり得る。こうした及び他の入力装置は、しばしば、システムバスに結合されたシリアルポートインターフェース 146 を介して処理装置 121 に接続されるが、他のインターフェース、例えば、並列ポート、ゲームポート、ユニバーサルシリアルバス (USB) によって接続することもできる。モニタ 147 または他のタイプの表示装置も、映像アダプタ 148 などのインターフェースを介してシステムバス 123 に接続される。モニタ 147 に加えて、パーソナルコンピュータは通常、他の周辺出力装置 (図示せず) 、例えば、スピーカ及びプリンタを含む。図 1 の例示的なシステムは、ホストアダプタ 155 と小型計算機システムインターフェース (SCSI) バス 156 と SCSI バス 156 に接続された外部記憶装置 162 も備えている。

【0018】

パーソナルコンピュータ 120 は、リモートコンピュータ 149 など1つまたは複数のリモートコンピュータへの論理接続を使用してネットワーク接続された環境において動作することができる。リモートコンピュータ 149 は、他のパーソナルコンピュータ、サーバ、ルータ、ネットワーク PC 、ピア装置または他の共通ネットワークノードでよく、通常、パーソナルコンピュータ 120 に関連して上述した要素の多くまたはすべてを含むが、図 1 にはメモリ記憶装置 150 のみを示した。図 1 に示した論理接続は、ローカルエリアネットワーク (LAN) 151 及びワイドエリアネットワーク (WAN) 152 を含む。このようなネットワーク環境は、会社、企業規模のコンピュータネットワーク、インターネット及びインターネットにおいてよく見られる。

【0019】

LAN ネットワーク環境において使われる場合、パーソナルコンピュータ 120 は、ネットワークインターフェースまたはアダプタ 153 を介して LAN 151 に接続される。WAN ネットワーク環境において使われる場合、パーソナルコンピュータ 120 は、通常、モデム 154 、または、例えば、インターネットなどのワイドエリアネットワーク 152 を介した通信を確立する他の手段を備えている。モデム 154 は、内部にあっても外部にあってもよく、シリアルポートインターフェース 146 を介してシステムバス 123 に接続される。ネットワーク接続された環境では、パーソナルコンピュータ 120 に関連して図示したプログラムモジュールまたはその一部は、リモートメモリ記憶装置に格納することができる。図示したネットワーク接続は例示的なものであり、コンピュータ間の通信

10

20

30

40

50

リンクを確立する他の手段も使うことができる。

【0020】

<本発明の例示的なシステム及び方法>

データの現在のバージョンは、基本テーブルに格納される。基本テーブル中の各行の前のバージョンは、バージョン記憶装置に格納される。デルタテーブルが、基本テーブルに対して行われた変更の一次キーを識別し、このような変更が挿入であるかまたは削除であるか指定する。1組の遷移テーブルが、基本テーブルに対して行われた各変更を格納する。遷移テーブルの組は、基本テーブルから削除された行を格納する削除テーブルと、基本テーブルに挿入された行を格納する挿入テーブルとを備えている。

【0021】

10

図2は、本発明による例示的なデータベース管理システム(DBMS)を示すブロック図である。DBMS200は、基本テーブル210と、バージョン記憶装置220と、デルタテーブル230と、削除テーブル240及び挿入テーブル242を含む1組の遷移テーブル240, 242とを備えている。

【0022】

データの現在のバージョンは、基本テーブル210に格納される。ただ1つの基本テーブル210を示してあるが、DBMS200は、任意の数の追加の基本テーブルを備えることができる。基本テーブル210は、好ましくは、データ列、トランザクション識別子(XID)列、及びシーケンス番号(SEQ)列を備えている。データ列は、実質的なデータを格納する。XID列は、各行の現在のバージョンがコミットされたトランザクションの識別子を格納する。SEQ列は、トランザクションにおいて変更が行われたシーケンスを識別する。SEQは、好ましくは、DBMS200によって自動的に維持される、単調に増加し続ける値である。DBMS200内のすべての基本テーブルがXID列及びSEQ列を含む必要があるわけではない。望ましくは、本発明による、識別された変更を格納する基本テーブルが、このようなXID列及びSEQ列を含む必要がある。例示的な基本テーブル210を、以下の表1に示す。

20

【0023】

【表1】

表1

顧客ID	名前	XID	SEQ
101	Andy	30	1
102	Bob	30	2
103		30	3
104	Dave	30	4
105	Ed	30	5

30

【0024】

表1は、顧客に関する実質的なデータを格納するデータ列「顧客ID」及び「名前」を備えている。行101～105は、トランザクション「30」の一部としてコミットされた。行103は、削除された記録であるゴースト記録である。したがって、行103はプレースホルダ(placeholder)として表1に残っているが、行103に対応する「名前」データはない。

40

【0025】

バージョン記憶装置220は、基本テーブル210中の行の前のバージョンを格納する。バージョン記憶装置220内にある行は、それと置き換わった行のバージョンのXID及びSEQによって指標づけされる。したがって、現在の各行のXID値及びSEQ値は、前のバージョンへのポインタとして働く。バージョン記憶装置220内の行を示す例示的なテーブルを、以下の表2に示す。

【0026】

【表2】

表2

変更 XID	変更 SEQ	顧客 ID	名前	XID	SEQ
30	3	103	Chris	20	3
30	2	102	Betty	20	2
30	1	101	Alex	20	1
20	1	101	Adam	10	1

【0027】

10

表2は、各行の変更XID値及び変更SEQ値を識別する列「変更XID」及び「変更SEQ」によって指標づけされる。例えば、表1に示す行101の現在のバージョンには、XID30及びSEQ1が割り当てられる。現在のバージョンと置き換わった行101の1つ前のバージョンは、表2において、変更XID30及び変更SEQ1を有して示される。1つ前のバージョンには、XID20及びSEQ1が割り当てられており、したがって、トランザクション20の一部としてコミットされた。1つ前のバージョンと置き換わった行101の2つ前のバージョンは、表2において変更XID20及び変更SEQ1を有して示される。2つ前のバージョンにはXID10及びSEQ1が割り当てられており、したがって、トランザクション10の一部としてコミットされた。

【0028】

20

図3a～図3eは、本発明による例示的な基本テーブル及びバージョン記憶装置を示すブロック図で、基本テーブル210及びバージョン記憶装置220にある、行の例示的なバージョンのブロック図である。例えば、図3aに示すように、行101の現在のバージョンは、基本テーブル210に格納されている。現在のバージョンには、XID30及びSEQ1が割り当てられ、これは、バージョン記憶装置220に格納されている、行の1つ前のバージョンをポイント(point)する。行の1つ前のバージョンにはXID20及びSEQ1が割り当てられ、これは、やはりバージョン記憶装置220に格納されている、行の2つ前のバージョンをポイントする。

【0029】

また、図3bに示すように、行102の現在のバージョンは、基本テーブル210に格納される。現在のバージョンには、XID30及びSEQ2が割り当てられ、これは、バージョン記憶装置220に格納されている、行の1つ前のバージョンをポイントする。

30

【0030】

また、図3cに示すように、ゴースト記録である行103の現在のバージョンは、基本テーブル210に格納される。ゴースト記録は、実質的なデータを含まない削除された記録のためのプレースホルダである。ゴースト記録には、XID30及びSEQ3が割り当てられ、これは、バージョン記憶装置220に格納されている、行の1つ前のバージョンをポイントする。

【0031】

40

図3d及び図3eにそれぞれ示すように、行104及び105の現在のバージョンは、基本テーブル210に格納される。このような行は新しく挿入されるので、バージョン記憶装置220にはこうした行の前のバージョンがない。他の実施形態では、所定の一義的なSEQ値を有する一義的なマーカを、行が新しく挿入されたことを示すのに使うことができる。このような一義的なマーカは、行の前のバージョンがバージョン記憶装置220に格納されていないことを示す。

【0032】

DBMS200が破損した後にバージョン記憶装置220が存続する必要はない。したがって、バージョン記憶装置220は、DBMS200の一時的なデータベースに格納することができる。行の前のバージョンは、好ましくは、このような行にアクセスするクエリを実行することができる間のみ格納され、したがって、前のバージョンは、バージョン

50

記憶装置 220 から周期的に削除される。

【0033】

バージョン記憶装置 220 は、例えば、b ツリーや b + ツリーなどのツリー構造に従って編成することができる。b ツリーに対するキーは、好ましくは各変更の XID 及び SEQ である。b ツリーに従って編成された行は、その XID 値及び SEQ 値に従って行全体を削除することができる。例えば、XID 及び SEQ の指定された範囲内の XID 及び SEQ を有する変更は、全体を削除することができる。行は、トランザクション中の単一の DML ステートメントに統いて、あるいは完全なトランザクションまたは一群のトランザクションに統いて削除することができる。

【0034】

あるいは、バージョン記憶装置 220 は、各行の前のバージョンを識別するための、物理行 ID またはディスクアドレスを使って編成することができる。このような行 ID 方式では、好ましくは XID 及び SEQ の他に、各行の前のバージョンの行 ID が格納される。前のバージョンは好ましくは、「ヒープ (h e a p) 」と呼ばれる 1 組のディスクページに格納される。新しいヒープは、指定された期間に生成することができる。例えば、ユーザは、60 秒ごとに新しいヒープを生成するよう要求することができる。ヒープは、ヒープに格納されているバージョンを生成したすべてのトランザクションが完了した後で削除することができる。

【0035】

バージョン記憶装置 220 の編成に加え、このようなツリー構造及びヒープは、基本テーブル 210 を編成するのに使うことができる。例えば、基本テーブル 210 中の行がバイナリラージオブジェクト (LOB) データを含む場合、このような LOB データは、フラグメントに分解し、その行に対応する別個の LOB 基本テーブルに格納することができる。このような LOB 基本テーブルは、DBMS 200 によって維持されるツリー構造またはヒープに従って編成することができる。

【0036】

デルタテーブル 230 は、基本テーブル 210 中で行われる変更を識別する。例示的なデルタテーブルは、好ましくは 2 つの列を有する。すなわち、基本テーブル 210 の一次キーの列、及び対応する変更が挿入であるかまたは削除であるかを示す変更列である。

【0037】

デルタテーブル 230 は、好ましくは、対応するデータ操作言語 (DML) ステートメントによって行われた変更を識別する。例えば、DML ステートメントが変更 {XID 30, SEQ 2} から {XID 30, SEQ 4} を行った場合、対応するデルタテーブル 230 は、好ましくは以下の表 3 に示すように生成される。

【0038】

【表 3】

表 3

一次キー	変更
102	削除
102	挿入
103	削除
104	挿入

【0039】

表 3 は、変更 {XID 30, SEQ 2} から {XID 30, SEQ 4} の一次キーの値を備え、対応する変更が挿入であるかまたは削除であるか示す。表 3 は、行 102 が修正されたことを示し、この修正は、削除及び挿入の両方を含む。表 3 はまた、行 103 が削除され、行 104 が挿入されたことを示す。

【0040】

10

20

30

40

50

遷移テーブル 240 及び 242 は、基本テーブル 210 に対して行われた変更を格納する。削除テーブル 240 は、基本テーブル 210 から削除された行を格納する。上述した例示的な DML ステートメントによって削除された行を示す例示的な削除テーブル 240 を、以下の表 4 に示す。

【0041】

【表 4】

表 4

顧客ID	名前
102	Betty
103	Chris

10

【0042】

挿入テーブル 242 は、基本テーブル 210 に挿入された行を格納する。上述した例示的な DML ステートメントによって挿入された行を示す例示的な挿入テーブル 242 を、以下の表 5 に示す。

【0043】

【表 5】

表 5

顧客ID	名前
102	Bob
104	Dave

20

【0044】

遷移テーブル 240 及び 242 には、トランザクションの一部としてのステートメントによって変更された行を識別することによって、投入が行われる。基本テーブル 210 は、削除マーカ及び挿入マーカを用いて走査される。削除マーカは、ステートメントの実行の前にトランザクションの一部として基本テーブル 210 にコミットされた最終の変更を識別し、挿入マーカは、ステートメントの実行中に基本テーブル 210 にコミットされた最終の変更を識別する。削除マーカによる走査の結果は、デルタテーブル 230 に格納されている削除を用いてフィルタリングされ、削除テーブル 240 には、フィルタリングされた行が投入される。挿入走査の結果は、デルタテーブル 230 に格納されている挿入を用いてフィルタリングされ、挿入テーブル 242 には、フィルタリングされた行が投入される。

30

【0045】

図 4 は、変更を識別して格納する例示的な方法を説明するためのフローチャートを示す図である。まず、ステップ 410 で、削除マーカ及び挿入マーカの XID が、トランザクションの識別子に設定される。例えば、上述した例示的なステートメントはトランザクション「30」の一部として実行されており、したがって、削除マーカ及び挿入マーカの XID は 30 に設定される。

40

【0046】

次に、ステップ 412 で、削除マーカの SEQ が、トランザクションの一部としてコミットされた最終の変更の SEQ に設定される。例えば、例示的なステートメントの実行の前にトランザクション「30」の一部としてコミットされた最終の変更には、SEQ「1」が割り当てられた。したがって、削除マーカは、{ (XID 30, SEQ 1) } に設定される。

【0047】

次に、ステップ 414 で、DBMS200 は、データ操作言語 (DML) ステートメントを実行する。このステートメントは、基本テーブル 210 中の行を挿入し、削除し、か

50

つ／または修正するのに使うことができる。例えば、上述した例示的なステートメントは、行 102 を修正し、行 103 を削除し、行 104 を挿入した。

【0048】

ステートメントがある特定の行を変更すると、その行の現在のバージョンは、バージョン記憶装置 220 に格納されている、行の前のバージョンを伴って、基本テーブル 210 に格納される。行の現在のバージョンは、基本テーブル 210 にコミットされると、行の前のバージョンへのポインタとして働く XID 値及び SEQ 値を割り当てられる。各変更に割り当てられる XID 値は、ステップ 410 で設定された、トランザクションに対する XID 値である。SEQ 値は、トランザクションにおいて変更が行われたシーケンスを識別する。SEQ は好ましくは、DBMS 200 によって自動的に維持される、単調に増加し続ける値である。例えば、例示的なステートメントによって変更された第 1 の行、つまり、行 102 に割り当てられた XID 値及び SEQ 値は、上述した表 1 に示した { (XID 30, SEQ 2) } である。次の変更された行、つまり、行 103 には { (XID 30, SEQ 3) } が割り当てられる。次の変更された行、つまり、行 104 には { (XID 30, SEQ 4) } が割り当てられる。

【0049】

次に、ステップ 416 で、挿入マーカの SEQ が、ステートメントの実行中にトランザクションの一部としてコミットされた最終の変更の SEQ に設定される。例えば、ステートメントの実行中にトランザクション「30」の一部として変更された最後の行は行 104 であり、SEQ「4」が割り当てられた。したがって、挿入マーカは、{ (XID 30, SEQ 4) } に設定される。

【0050】

次に、ステップ 418 で、基本テーブル 210 が、削除マーカを用いて走査される。基本テーブル 210 を走査する例示的な方法は、図 5 を参照して後で詳しく説明する。このような走査を表 1 に対して削除マーカ { (XID 30, SEQ 1) } を用いて実施した結果を、以下の表 6 に示す。

【0051】

【表 6】

表 6

顧客ID	名前	XID	SEQ
101	Andy	30	1
102	Betty	20	2
103	Chris	20	3

【0052】

次に、ステップ 420 で、削除走査の結果が、デルタテーブル 230 に格納されている削除を用いてフィルタリングされ、ステップ 422 で、削除テーブル 240 に、フィルタリングされた行が投入される。例えば、表 6 中の行が、表 3 に格納されている削除を用いてフィルタリングされると、削除テーブル 240 には、以下の表 7 に示すように投入が行われる。

【0053】

【表 7】

表 7

顧客ID	名前	XID	SEQ
102	Betty	20	2
103	Chris	20	3

【0054】

10

20

30

40

50

次に、ステップ 424 で、基本テーブル 210 が、挿入マーカを用いて走査される。基本テーブル 210 を走査する例示的な方法は、後で図 5 を参照して詳しく説明する。このような走査を表 1 に対して削除マーカ { (XID 30, SEQ 4) } を用いて実施した結果を、以下の表 8 に示す。

【0055】

【表 8】

表 8

顧客ID	名前	XID	SEQ
101	Andy	30	1
102	Bob	30	2
103		30	3
104	Dave	30	4

10

【0056】

ステップ 426 で、挿入走査の結果が、デルタテーブル 230 に格納されている挿入を用いてフィルタリングされ、ステップ 428 で、挿入テーブル 242 に、フィルタリングされた行が投入される。例えば、表 8 中の行が、表 5 に格納されている挿入を用いてフィルタリングされると、挿入テーブル 242 には、以下の表 9 に示すように投入が行われる。

20

【0057】

【表 9】

表 9

顧客ID	名前	XID	SEQ
102	Bob	30	2
104	Dave	30	4

30

【0058】

したがって、図 4 に関連して上述した処理は、DML ステートメントによって行われた変更を識別し、遷移テーブル 240 ~ 242 にこのような変更を投入する。この処理は、挿入マーカ及び削除マーカを用いて基本テーブル 210 を走査することによって、このような変更が識別されることを可能にする。

【0059】

図 5 は、テーブルを走査する例示的な方法を説明するためのフローチャートを示す図である。走査は、マーカの SEQ においてまたはその前に起こる、異なるトランザクションの一部として行われた変更、及び同じトランザクションの一部として行われた変更を読み出す。

【0060】

図 5 に示すように、ステップ 510 で、行の XID がマーカの XID に等しいか否かを判定する。等しくない場合、ステップ 516 で、その行が読み出される。等しい場合、ステップ 512 で、その行の SEQ がマーカの SEQ 以下であるか否かを判定する。マーカの SEQ 以下である場合、ステップ 516 で、その行が読み出される。マーカの SEQ 以下でない場合、ステップ 514 で、その行の前のバージョンがバージョン記憶装置 220 から取り出され、方法はステップ 510 に戻る。

40

【0061】

図 4 に関連して上述したように、行は読み出された後、走査結果に追加され、デルタテーブル 230 に格納されている行を用いてフィルタリングされる。削除走査によって読み出された行がデルタテーブル 230 中にも削除として記載されている場合、その行は削除テーブル 240 に格納される。挿入走査によって読み出された行がデルタテーブル 230

50

中にも挿入として記載されている場合、その行は挿入テーブル 242 に格納される。図 4 及び図 5 に関連して上述したテーブルは、例えば、メモリやデータベースなど、どの記憶装置にも格納することができる。

【0062】

以上、本発明を図面を参照して好ましい実施形態について説明したが、本発明から逸脱することなく、本発明の同じ機能を実施するために、他の類似の実施形態を用いることも、説明した実施形態に修正及び追加を行うこともできる。したがって、本発明は、どの 1 つの実施形態にも限定されるべきではなく、特許請求の範囲による広さ及び技術的範囲内で解釈されるべきである。

【図面の簡単な説明】

10

【0063】

【図 1】本発明に係る識別システムの一実施形態を説明するためのブロック図である。

【図 2】本発明による例示的なデータベース管理システムを示すブロック図である。

【図 3 a】本発明による例示的な基本テーブル及びバージョン記憶装置を示すブロック図(その 1)である。

【図 3 b】本発明による例示的な基本テーブル及びバージョン記憶装置を示すブロック図(その 2)である。

【図 3 c】本発明による例示的な基本テーブル及びバージョン記憶装置を示すブロック図(その 3)である。

【図 3 d】本発明による例示的な基本テーブル及びバージョン記憶装置を示すブロック図(その 4)である。

20

【図 3 e】本発明による例示的な基本テーブル及びバージョン記憶装置を示すブロック図(その 5)である。

【図 4】変更を識別して格納する例示的な方法を説明するためのフローチャートを示す図である。

【図 5】テーブルを走査する例示的な方法を説明するためのフローチャートを示す図である。

【符号の説明】

【0064】

120 パーソナルコンピュータ

30

121 処理装置

122 システムメモリ

123 システムバス

124 読出し専用メモリ(ROM)

125 ランダムアクセスメモリ(RAM)

126 基本入出力システム(BIOS)

127 ハードディスクドライブ

128 磁気ディスクドライブ

129 磁気ディスク

130 光ディスクドライブ

40

131 光ディスク

132 ハードディスクドライブインターフェース

133 磁気ディスクドライブインターフェース

134 光ドライブインターフェース

135 オペレーティングシステム

136 アプリケーションプログラム

137 他のプログラムモジュール

138 プログラムデータ

140 キーボード

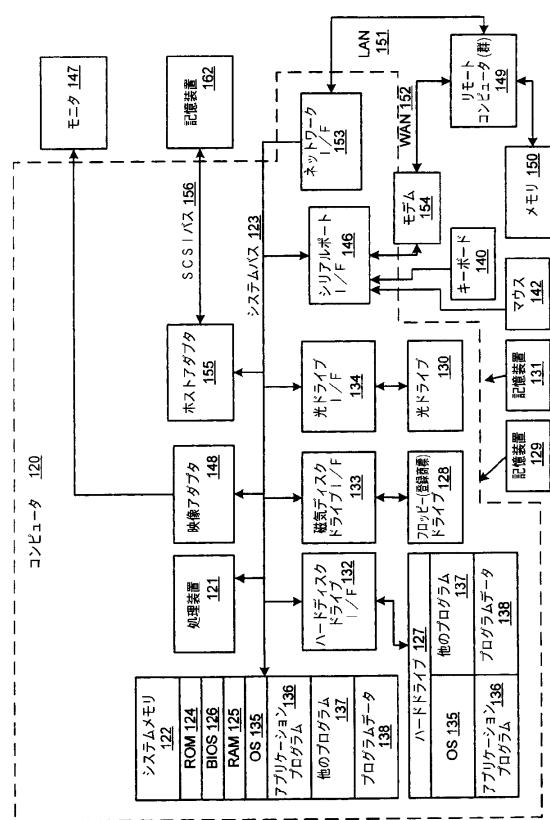
142 指示装置

50

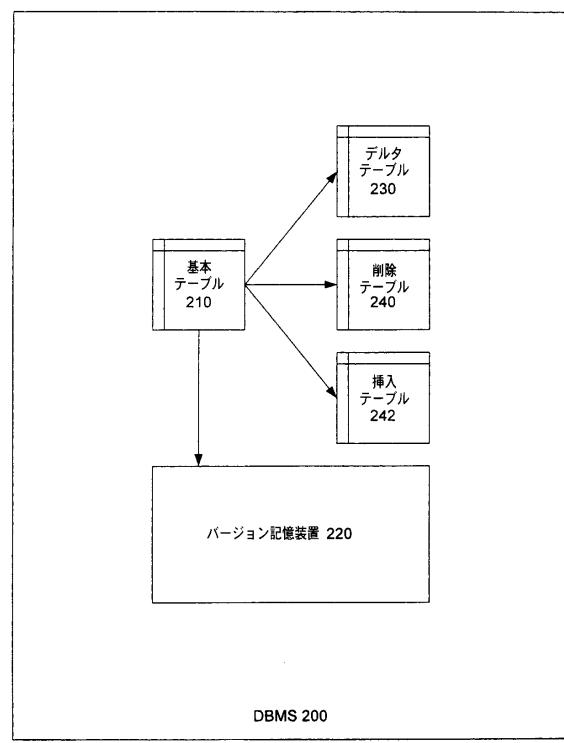
- 1 4 6 シリアルポートインターフェース
 1 4 7 モニタ
 1 4 8 映像アダプタ
 1 4 9 リモートコンピュータ
 1 5 0 メモリ記憶装置
 1 5 1 ローカルエリアネットワーク (LAN)
 1 5 2 ワイドエリアネットワーク (WAN)
 1 5 3 ネットワークインターフェース、ネットワークアダプタ
 1 5 4 モデム
 1 5 5 ホストアダプタ
 1 5 6 小型計算機システムインターフェース (SCSI) バス
 1 6 2 外部記憶装置
 2 0 0 データベース管理システム (DBMS)
 2 1 0 基本テーブル
 2 2 0 バージョン記憶装置
 2 3 0 デルタテーブル
 2 4 0 遷移テーブル、削除テーブル
 2 4 2 遷移テーブル、挿入テーブル

10

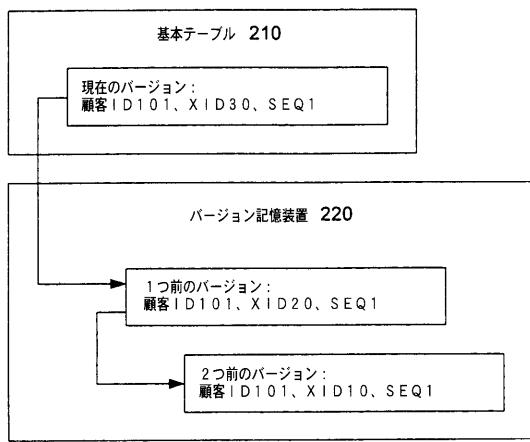
【図1】



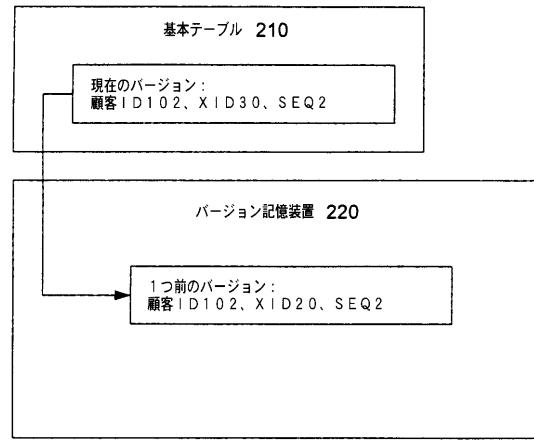
【図2】



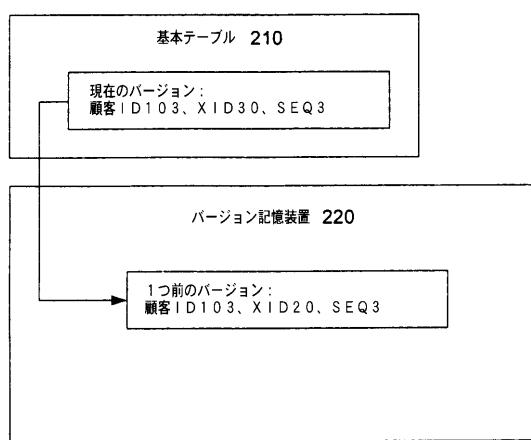
【図3 a】



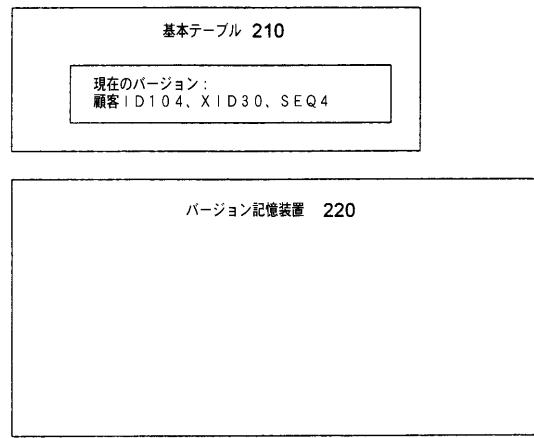
【図3 b】



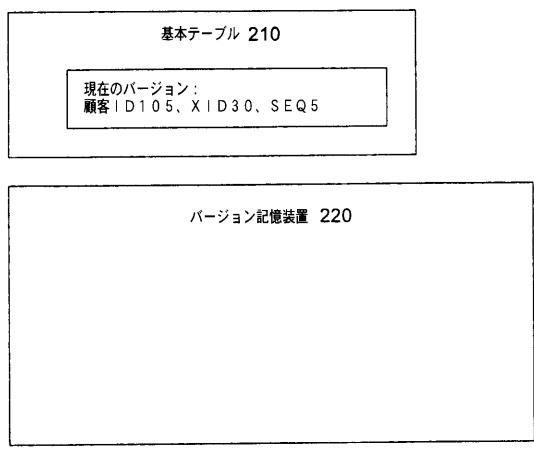
【図3 c】



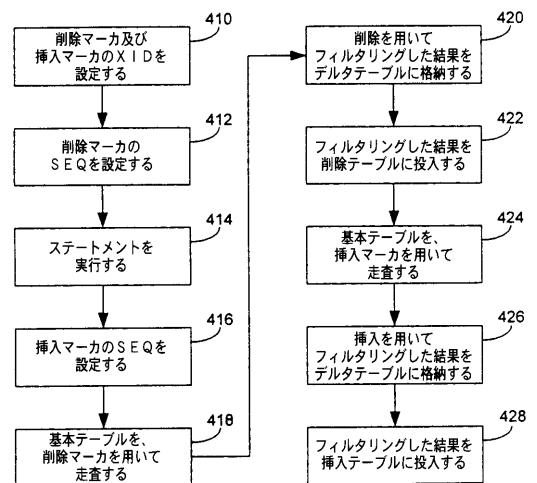
【図3 d】



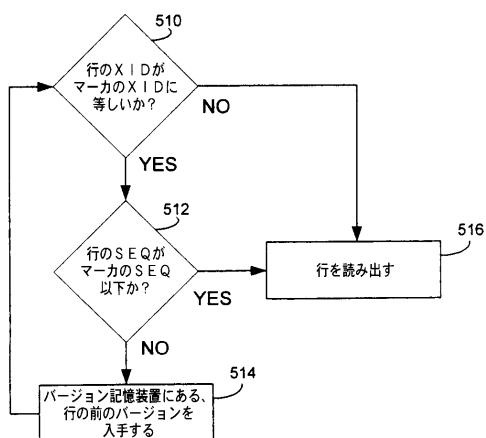
【図3e】



【図4】



【図5】



フロントページの続き

(72)発明者 シャオ ウェイ
アメリカ合衆国 98052 ワシントン州 レッドモンド ノースイースト 84 コート 1
4316

(72)発明者 ナイジェル アール.エリス
アメリカ合衆国 98053 ワシントン州 レッドモンド 194 アベニュー ノースイースト
ト 12125

(72)発明者 ロジャー エヌ.クライン
アメリカ合衆国 98072 ワシントン州 ウッディンビル ノースイースト 160 ストリ
ート 17901

(72)発明者 サミート エイチ.アガーウォル
アメリカ合衆国 98052 ワシントン州 レッドモンド 149 プレイス ノースイースト
8127

審査官 北村 学

(56)参考文献 特開平05-040681 (JP, A)
特開平06-083682 (JP, A)
特開平09-106364 (JP, A)
特開2002-032248 (JP, A)
米国特許第06275832 (US, B1)
米国特許第06339772 (US, B1)

(58)調査した分野(Int.Cl., DB名)

G 06 F 12/00