(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5710782号
(P5710782)
(45)発行日　平成27年4月30日(2015.4.30)　(24)登録日　平成27年3月13日(2015.3.13)

(51)Int.Cl.　　F I
G06F 12/00 (2006.01)　　G06F 12/00 513J
G06F 17/30 (2006.01)　　G06F 17/30 110C
G06F 17/30 180D
G06F 17/30 340C

請求項の数 19　(全 24 頁)

(21)出願番号　特願2013-546104 (P2013-546104)
(86)(22)出願日　平成23年4月15日 (2011.4.15)
(65)公表番号　特表2014-505925 (P2014-505925A)
(43)公表日　平成26年3月6日 (2014.3.6)
(86)国際出願番号　PCT/US2011/032631
(87)国際公開番号　WO2012/087366
(87)国際公開日　平成24年6月28日 (2012.6.28)
審査請求日　平成26年2月3日 (2014.2.3)
(31)優先権主張番号　12/973,668
(32)優先日　平成22年12月20日 (2010.12.20)
(33)優先権主張国　米国 (US)

(73)特許権者　506332063
セールスフォース　ドット　コム　インコーポレイティッド
アメリカ合衆国　カリフォルニア州　９４１０５，サンフランシスコ，ザ　ランドマーク　アット　ワン　マーケット，スイート　３００
(74)代理人　100083806
弁理士　三好　秀和
(74)代理人　100095500
弁理士　伊藤　正和
(74)代理人　100111235
弁理士　原　裕子

最終頁に続く

(54)【発明の名称】マルチテナントストアで横断的ストア結合を行う方法及びシステム

(57)【特許請求の範囲】
【請求項１】
リレーショナルデータストア及び非リレーショナルデータストアを有するマルチテナントデータベースシステムからデータを取得する方法であって、
ホストシステムで前記マルチテナントデータベースシステムに対する要求を受信するステップであって、前記要求は前記マルチテナントデータベースシステムから取得されるデータを特定するステップと、
前記ホストシステムを介する前記要求に基づいて、前記取得されるデータの１つ以上の位置を取得するステップと、
前記ホストシステムで前記要求に基づいて取得される複数のデータ要素を特定するデータベースクエリを生成するステップであって、前記複数のデータ要素は前記非リレーショナルデータストア内に存在する１つ以上のデータ要素及び前記リレーショナルデータストア内に存在する１つ以上の他のデータ要素を含むステップと、
前記データを取得するために前記マルチテナントデータベースシステムに対して前記データベースクエリを実行するステップと
を含み、
前記データベースクエリは結合動作を特定し、
前記結合動作は、
前記非リレーショナルデータストアに対して実行される第１のサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素を識別する第１

のサブクエリと、
前記リレーショナルデータストアに対して実行される第２のサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素を識別する前記第１のサブクエリと前記リレーショナルデータストア内に存在する前記１つ以上の他のデータ要素との間のデータデルタを決定する第２のサブクエリと、
前記リレーショナルデータストア及び前記非リレーショナルデータストアに対して実行される第３のサブクエリであって、前記リレーショナルデータストアから前記非リレーショナルデータストアへと決定されたデータデルタに対応するデータを複製する第３のサブクエリと、
前記非リレーショナルデータストアに対して実行される第４のサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素と前記リレーショナルデータストアから前記非リレーショナルデータストアに複製されることで前記非リレーショナルデータストア内から利用可能な前記１つ以上の他のデータ要素との双方をフェッチすることにより、前記非リレーショナルデータストアから取得される前記データをフェッチする第４のサブクエリと
を含む、方法。
【請求項２】
顧客スキーマが前記取得されるデータの１つ以上の位置を記述し、前記顧客スキーマは前記非リレーショナルデータストア内若しくは前記リレーショナルデータストア内に存在する又は前記非リレーショナルデータストア及び前記リレーショナルデータストアの双方から利用可能である前記データの前記複数のデータ要素の各々を特定し、
前記方法は、前記ホストシステムを介して、前記要求の受信に応答して前記顧客スキーマを取得するステップを更に含む、請求項１に記載の方法。
【請求項３】
前記データベースクエリは、複数のサブクエリを含み、
前記複数のサブクエリの中の少なくとも１つのサブクエリは、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素を前記非リレーショナルデータストアから取得するために行われ、
前記複数のサブクエリの中の少なくとも第２のサブクエリは、前記リレーショナルデータストア内に存在する前記１つ以上の他のデータ要素を前記リレーショナルデータストアから取得するために行われる、請求項１に記載の方法。
【請求項４】
前記マルチテナントデータベースシステムに対して前記データベースクエリを実行するステップは、前記リレーショナルデータストア及び前記非リレーショナルデータストアの双方に記憶されたデータ要素を参照することを含む、請求項１に記載の方法。
【請求項５】
前記結合動作は、前記ホストシステムの最適化エージェントを介して利用可能なクエリ最適化に基づいて生成される複数のサブクエリを含み、前記クエリ最適化は、
前記複数のサブクエリに対する特定の順序付け、
対応するサブクエリの実行に対する目的データストア、
前記取得されるデータに基づく１つ以上の事前クエリ評価、
前記リレーショナルデータストアから前記非リレーショナルデータストアへの複製命令、
前記リレーショナルデータストア及び前記非リレーショナルデータストアの各々から取得される前記複数のデータ要素の少なくとも１つ以上を特定し、且つ前記最適化エージェントにアクセス可能なメモリに配置されるメモリ内結合動作、及び
前記要求を充足するに際し前記最適化エージェントにアクセス可能な前記メモリから前記複数のデータ要素の前記少なくとも１つ以上を取得するための対応するサブクエリを含むグループから選択される、請求項１に記載の方法。
【請求項６】

前記結合動作は、
前記リレーショナルデータストアから２つ以上の関連テーブルを特定する結合動作と、
前記リレーショナルデータストアからの少なくとも１つの関連テーブル及び前記非リレーショナルデータストア内に存在する少なくとも１つ以上のデータ構造を特定する結合動作と、
前記非リレーショナルデータストア内に存在する２つ以上の離れた及び別個のデータ構造を特定する結合動作であって、前記２つ以上の離れた及び別個のデータ構造の各々が重複する共有キーを欠いている結合動作と
を含むグループから選択される、請求項１に記載の方法。
【請求項７】
前記結合動作は、
前記リレーショナルデータストアに対して実行される第５のサブクエリであって、決定されたデータデルタに基づいて前記リレーショナルデータストア内に存在する前記１つ以上の他のデータ要素を取得する第５のサブクエリと
を含む、請求項１に記載の方法。
【請求項８】
前記第５のサブクエリは、
前記リレーショナルデータストアに対して実行される時間ベースのサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素内のデータ要素に対応する任意のタイムスタンプよりも遅いタイムスタンプを有する前記リレーショナルデータストア内のデータ要素に基づいて前記リレーショナルデータストアから取得される前記１つ以上の他のデータ要素を特定する時間ベースのサブクエリと、
前記リレーショナルデータストアに対して実行される記録識別子ベースのサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素内のデータ要素に対応する任意の記録識別子よりも数値が大きい記録識別子を有する前記リレーショナルデータストア内のデータ要素に基づいて前記リレーショナルデータストアから取得される前記１つ以上の他のデータ要素を特定する記録識別子ベースのサブクエリと
の１つを含む、請求項７に記載の方法。
【請求項９】
前記マルチテナントデータベースシステムにおいて複数の新しいトランザクションを受信するステップであって、前記新しいトランザクションの各々は前記非リレーショナルデータストアに書き込まれる新しいデータを特定するステップと、
前記リレーショナルデータストアの追加ログに前記新しいデータを書き込むステップと
を更に含み、決定されたデータデルタに基づいて前記リレーショナルデータストア内に存在する前記１つ以上の他のデータ要素を取得する第５のサブクエリは、前記リレーショナルデータストアの前記追加ログから前記リレーショナルデータストア内に存在する前記１つ以上の他のデータ要素を取得する第５のサブクエリを含む、請求項７に記載の方法。
【請求項１０】
前記追加ログがフラッシュ閾値に到達すると、前記リレーショナルデータストアの前記追加ログに書き込まれた前記新しいデータを前記非リレーショナルデータストアにフラッシュすることを更に含む、請求項９に記載の方法。
【請求項１１】
前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素は、複数の圧縮フラットファイル若しくは複数のバイナリファイル又は前記圧縮フラットファイル及び前記バイナリファイルの組み合わせを含む、請求項１に記載の方法。
【請求項１２】
前記リレーショナルデータストアは、リレーショナルデータベース管理システム（ＲＤＢＭＳ）に従って実装されるリレーショナルデータベースを含み、前記リレーショナルデータベースの複数の関連テーブルは、前記リレーショナルデータベース内の２つ以上の関連テーブルの各々に対して１つ以上の重複する共通の特徴を介して相互に関連付けされる

、請求項１に記載の方法。
【請求項１３】
前記非リレーショナルデータストアは、各々が前記非リレーショナルデータストアに総記憶容量の少なくとも一部を提供する複数の基本ハードウェア記憶装置を有する分散構造型データベースを含み、前記非リレーショナルデータストア内のデータ要素は、主キーに基づいて参照可能であるが、２つ以上の関連テーブルの間の１つ以上の重複する共通の特徴に基づいて参照可能ではない、請求項１に記載の方法。
【請求項１４】
前記リレーショナルデータストアは、Ｏｒａｃｌｅ互換データベースの実装、ＩＢＭ ＤＢ２ Ｅｎｔｅｒｐｒｉｓｅ Ｓｅｒｖｅｒ互換リレーショナルデータベースの実装、ＭｙＳＱＬ互換リレーショナルデータベースの実装、及びＭｉｃｒｏｓｏｆｔ ＳＱＬ Ｓｅｒｖｅｒ互換リレーショナルデータベースの実装を含むグループから選択されるリレーショナルデータベースの実装を含み、
前記非リレーショナルデータストアは、Ｖａｍｐｉｒｅ互換非リレーショナルデータベースの実装、Ａｐａｃｈｅ Ｃａｓｓａｎｄｒａ互換非リレーショナルデータベースの実装、ＢｉｇＴａｂｌｅ互換非リレーショナルデータベースの実装、及びＨＢａｓｅ互換非リレーショナルデータベースの実装を含むグループから選択されるＮｏＳＱＬ非リレーショナルデータベースの実装を含む、請求項１に記載の方法。
【請求項１５】
前記マルチテナントデータベースシステムのインターフェースを介して前記要求を受信するステップは、前記マルチテナントデータベースシステムのウェブサーバを介して前記要求を受信することを含み、前記ウェブサーバは、前記要求の起点となる遠隔に設置されたエンドユーザクライアントマシンとのウェブベースインターフェースを提供し、
前記要求は、前記マルチテナントデータベースシステムに関するホスト組織内で動作する前記マルチテナントデータベースシステムからのサービスに対する要求を含む、請求項１に記載の方法。
【請求項１６】
前記マルチテナントデータベースシステムは、複数の離れた別個の顧客組織によって共有されるハードウェア及びソフトウェアの要素を更に含み、前記離れた別個の顧客組織の各々は前記マルチテナントデータベースシステムが実行されるホスト組織から遠隔に設置される、請求項１に記載の方法。
【請求項１７】
前記非リレーショナルデータストアは複数の分散型コンピュータノードを含み、各コンピュータノードは少なくとも１つのメモリ、１つ以上のプロセッサ及び１つ以上の通信可能にインターフェース接続されたハードディスクドライブを備え、前記分散型コンピュータノードの各々は、中央トランザクション局からの承認又は制御無しで非リレーショナルデータベーストランザクションを読み取り、書き込み、及び更新する機能を有する分離した非リレーショナルデータベースインスタンスを含み、
前記リレーショナルデータストアは、モノリシックリレーショナルデータベースインスタンスへの更新又は変更が前記モノリシックリレーショナルデータベースインスタンスに通信可能にインターフェース接続され且つ制御される持続的記憶装置に対する持続的記憶にコミットされるかどうかを制御する中央トランザクション局と計算リソースを調整するプロセッサ及びメモリを含むモノリシックリレーショナルデータベースインスタンスを含む、請求項１に記載の方法。
【請求項１８】
ホストシステム内のプロセッサによって実行されると方法を実行する命令が記憶された一時的でないコンピュータ可読記憶媒体であって、前記方法は、
マルチテナントデータベースシステムから取得されるデータを特定する要求を受信するステップと、
前記要求に基づいて、前記取得されるデータの１つ以上の位置を取得するステップと、

前記要求に基づいて取得される複数のデータ要素を特定するデータベースクエリを生成するステップであって、前記複数のデータ要素は前記マルチテナントデータベースシステムの非リレーショナルデータストア内に存在する１つ以上のデータ要素及び前記マルチテナントデータベースシステムのリレーショナルデータストア内に存在する１つ以上の他のデータ要素を含むステップと、
前記データを取得するために前記マルチテナントデータベースシステムに対して前記データベースクエリを実行するステップと
を含み、
前記データベースクエリは結合動作を特定し、前記結合動作は、
前記非リレーショナルデータストアに対して実行される第１のサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素を識別する第１のサブクエリと、
前記リレーショナルデータストアに対して実行される第２のサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素を識別する前記第１のサブクエリと前記リレーショナルデータストア内に存在する前記１つ以上の他のデータ要素との間のデータデルタを決定する第２のサブクエリと、
前記リレーショナルデータストア及び前記非リレーショナルデータストアに対して実行される第３のサブクエリであって、前記リレーショナルデータストアから前記非リレーショナルデータストアへと決定されたデータデルタに対応するデータを複製する第３のサブクエリと、
前記非リレーショナルデータストアに対して実行される第４のサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素と前記リレーショナルデータストアから前記非リレーショナルデータストアに複製されることで前記非リレーショナルデータストア内から利用可能な前記１つ以上の他のデータ要素との双方をフェッチすることにより、前記非リレーショナルデータストアから取得される前記データをフェッチする第４のサブクエリと
を含む、一時的でないコンピュータ可読記憶媒体。

【請求項１９】
プロセッサ及びメモリと、
リレーショナルデータストア及び非リレーショナルデータストアが実装されるマルチテナントデータベースシステムとの通信インターフェースと、
前記マルチテナントデータベースシステムから取得されるデータを特定する要求を受信る要求プロセッサと、
前記要求に基づいて、前記取得されるデータの１つ以上の位置を取得する顧客スキーマプロセッサと、
前記要求に基づいて取得される複数のデータ要素を特定するデータベースクエリを生成するサブクエリ生成器であって、前記複数のデータ要素は前記マルチテナントデータベースシステムの前記非リレーショナルデータストア内に存在する１つ以上のデータ要素及び前記マルチテナントデータベースシステムの前記リレーショナルデータストア内に存在する１つ以上の他のデータ要素を含むサブクエリ生成器と、
前記データを取得するために前記マルチテナントデータベースシステムに対して前記データベースクエリを実行するクエリ実行器と
を備え、
前記要求に基づいて前記データベースクエリを生成する前記サブクエリ生成器は、前記要求を充足するために前記要求に基づいて複数のサブクエリを生成するサブクエリ生成器を含み、前記サブクエリは、
前記非リレーショナルデータストアに対して実行される第１のサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素を識別する第１のサブクエリと、
前記リレーショナルデータストアに対して実行される第２のサブクエリであって、前記

非リレーショナルデータストア内に存在する前記１つ以上のデータ要素を識別する前記第１のサブクエリと前記リレーショナルデータストア内に存在する前記１つ以上の他のデータ要素との間のデータデルタを決定する第２のサブクエリと、
前記リレーショナルデータストア及び前記非リレーショナルデータストアに対して実行される第３のサブクエリであって、前記リレーショナルデータストアから前記非リレーショナルデータストアへと決定されたデータデルタに対応するデータを複製する第３のサブクエリと、
前記非リレーショナルデータストアに対して実行される第４のサブクエリであって、前記非リレーショナルデータストア内に存在する前記１つ以上のデータ要素と前記リレーショナルデータストアから前記非リレーショナルデータストアに複製されることで前記非リレーショナルデータストア内から利用可能な前記１つ以上の他のデータ要素との双方をフェッチすることにより、前記非リレーショナルデータストアから取得される前記データをフェッチする第４のサブクエリと
を含む、システム。

【発明の詳細な説明】

【技術分野】

【０００１】

本明細書に記載の主題は、概してコンピュータの分野に関し、より詳細にはマルチテナントストアで横断的ストア結合を行う方法及びシステムに関する。

【０００２】

（優先権の主張）

本願は、出願番号１２／９７３，６６８及び代理人整理番号８９５６Ｐ００６／３２０ＵＳを有し、２０１０年１２月２０日に提出され、“ＭＥＴＨＯＤＳ ＡＮＤ ＳＹＳＴＥＭＳ ＦＯＲ ＰＥＲＦＯＲＭＩＮＧ ＣＲＯＳＳ ＳＴＯＲＥ ＪＯＩＮＳ ＩＮ Ａ ＭＵＬＴＩ－ＴＥＮＡＮＴ ＳＴＯＲＥ”と題された米国特許出願、並びに出願番号６１／３２５，７０９及び代理人整理番号８９５６Ｐ００６Ｚ／３２０ＰＲＯＶを有し、２０１０年４月１９日に提出され、“ＭＥＴＨＯＤＳ ＡＮＤ ＳＹＳＴＥＭＳ ＦＯＲ ＰＥＲＦＯＲＭＩＮＧ ＣＲＯＳＳ ＳＴＯＲＥ ＪＯＩＮＳ ＩＮ Ａ ＭＵＬＴＩ－ＴＥＮＡＮＴ ＳＴＯＲＥ”と題された仮特許出願に関し、これらへの優先権を主張すると共に、全ての内容が参照により本明細書に組み込まれる。本願は、出願番号６１／３２５，９５１及び代理人整理番号８９５６Ｐ００７Ｚ／３２１ＰＲＯＶを有し、２０１０年４月２０日に提出され、“ＭＥＴＨＯＤＳ ＡＮＤ ＳＹＳＴＥＭＳ ＦＯＲ ＯＰＴＩＭＩＺＩＮＧ ＱＵＥＲＩＥＳ ＩＮ Ａ ＭＵＬＴＩ－ＴＥＮＡＮＴ ＳＴＯＲＥ”と題された仮特許出願に更に関し、これへの優先権を主張し、全ての内容が参照により本明細書に組み込まれる。

【０００３】

（著作権表示）

本文書の開示の一部は著作権によって保護される資料を含んでいる。著作権者は、特許文献又は特許開示が特許商標庁の特許出願又は記録に現れるものについては、何人によるこれらの複製に対して異議を持たないが、それ以外はどのようなものであっても全ての著作権を保有する。

【背景技術】

【０００４】

背景欄で検討される主題は、単に背景欄で言及された結果として先行技術であるとみなされるべきではない。同様に、背景欄で言及される又は背景欄の主題と関連する課題は、先行技術で過去に認識されていたとみなされるべきではない。背景欄における主題は、請求項に記載された主題の実施形態に対応し得る異なるアプローチを表しているだけである。

【０００５】

コンピュータ環境内では、データを持続的に記憶するために様々なデータ記憶環境が選

択される場合がある。例えば、データは、ハードドライブにファイルシステムデータを持続的に記憶するオペレーティングシステムによって管理されるファイルシステム内に記憶されてもよい。又はデータは、データベース内に持続的に記憶されてもよい。各々が固有の利点及び欠点を有する様々な種類のデータベースが利用可能である。例えば、所謂リレーショナルデータベースは、各テーブルによって共有される共通の特徴を使用して、様々なデータテーブルをデータベース内で互いに「関連付け」する能力を提供する。例えば、リレーショナルデータベースでは、従業員識別子は１つより多くのテーブルを関連付ける共通の特徴として使用されてもよい。しかしながら、このようなデータベース構造はいくつかの欠点を有している。そのうちの１つは、関連が高水準の計算オーバヘッドコスト及びリレーショナルデータベースのスケーラビリティの範囲を制限する計算の複雑性を伴うという点である。

【０００６】

非リレーショナルデータベースのモデル及び実装も存在し、一般的にはより良好なスケーラビリティを示すが、リレーショナルデータベースのモデル及び実装と関連しない異なる欠点も示す。例えば、非リレーショナルデータベースの実装は、大きなファイル又はオブジェクトを記憶するための改善されたスケーラビリティを示すことが多いが、選択的なデータセットを記憶し又は急速に変化するデータセットに対してデータ保障を実装する等の他の点ではあまり適切でない場合がある。

【０００７】

残念ながら、同時に複数のデータから情報を参照するデータベースクエリは著しく非効率的であり、複数のデータストアの実装から得られる可能性がある利益を損なう。更に、分散データベースモデルの個別の実装を同時に参照するデータベースクエリは、以前のデータベースクエリ機構の使用では完全に実行不可能な場合がある。

【０００８】

実施形態は例示目的で示されており、限定を目的とするものではなく、以下の詳細な説明を参照して図面と併せて検討すればより完全に理解され得る。

【図面の簡単な説明】

【０００９】

【図１】実施形態が動作可能な例示的なアーキテクチャを示す。

【図２】実施形態が動作可能な代替の例示的なアーキテクチャを示す。

【図３】実施形態が動作可能な代替の例示的なアーキテクチャを示す。

【図４】実施形態が動作可能な代替の例示的なアーキテクチャを示す。

【図５】実施形態が動作し、設置され、統合され、又は構成され得るシステムの図表現を示す。

【図６】一実施形態に従ってマルチテナントストアで横断的ストア結合を実行する方法の流れ図を示す。

【図７】一実施形態に従うコンピュータシステムの例示的態様で機械の図表現を示す。

【発明を実施するための形態】

【００１０】

本明細書にはマルチテナントストアで横断的ストア結合を実行するシステム、装置、及び方法が記載される。一実施形態では、このような方法は、リレーショナルデータストア及び非リレーショナルデータストアを有するマルチテナントデータベースシステムからデータを取得することを含む。例えば、このような方法では、マルチテナントデータベースシステムに関するホストシステムは、マルチテナントデータベースシステムから取得されるデータを特定する要求を受信し、要求に基づいて、ホストシステムを介して取得されるデータの１つ以上の位置を取得し、ホストシステムにおいて、要求に基づいて取得される複数のデータ要素を特定するデータベースクエリを生成し、複数のデータ要素は非リレーショナルデータストア内に存在する１つ以上のデータ要素及びリレーショナルデータストア内に存在する１つ以上の他のデータ要素を含み、データを取得するためにマルチテナントデータベースシステムに対してデータベースクエリを実行する。

【００１１】
１つより多くのデータベースを検索又は参照する連合クエリは、異なるデータストアに記憶されたテーブル間の結合動作を要求すること等によって、特にデータベースの最下行レベルでデータを参照する場合に、著しく非効率的である。その理由は、この動作は非常に多くのネットワーク帯域幅を消費するので、このような結合動作はうまくスケールせず、より大きなデータベースの実装で実行できるように実装できないからである。このような結合動作の課題は、例えば、リレーショナルデータベースの実装と非リレーショナルデータベースの実装との結合等、分散モデルで動作するデータベースの複数の実装間でデータ結合を要求する場合に更に悪化する。本明細書に記載の手法は、より大きなデータベースシステムで実行できるように実装可能なように、特に、リレーショナル及び非リレーショナルモデル等の分散動作モデルで動作する複数のデータストアの実装を利用するシステムで実行できるように実装可能なように、このような結合動作を実行する能力を促進する。

【００１２】
例えば、本明細書に記載の手法を用いて、結合動作は、非リレーショナルデータベースに記憶されたオブジェクトに対する非リレーショナルデータベースクエリを開始することによって実行されてもよく、ここでは１つ以上の外部キーペアレントがＯｒａｃｌｅ™等のリレーショナル型データベースの実装で記憶されるオブジェクトである。例えば非リレーショナルデータベースに記憶されたオブジェクトが非リレーショナルデータベースの実装に存在すると共にＯｒａｃｌｅ™に記憶されたオブジェクトがリレーショナルデータベースの実装に存在するにもかかわらず、非リレーショナルデータベースに記憶されたチャイルドテーブルは、マスターテーブルとしてＯｒａｃｌｅ™に記憶された“Ａｃｃｏｕｎｔ”テーブルを有してもよい。

【００１３】
以下の記載では、様々な実施形態の十分な理解を提供するために、多くの特定の詳細が特定のシステム、言語、コンポーネント等の例として説明される。しかしながら、特定の詳細は開示の実施形態を実施するために用いられる必要がないことが当業者には明らかであろう。他の例では、開示の実施形態を不必要に分かりにくくするのを避けるために、周知の材料又は方法は詳細には記載されていない。

【００１４】
図面に描かれた及び本明細書に記載された様々なハードウェア要素に加えて、実施形態は以下に記載の様々な動作を更に含む。このような実施形態に従って記載された動作は、ハードウェア要素によって実行され、又は命令と共にプログラミングされた汎用若しくは専用プロセッサに動作を実行させるために使用可能な機械実行可能命令で具現化されてもよい。代替的に、動作はハードウェア及びソフトウェアの組み合わせによって実行されてもよい。

【００１５】
実施形態は、本明細書に記載の動作を実行するためのシステム又は装置にも関する。開示のシステム又は装置は、要求された目的で特別に構築されてもよく、又はコンピュータに記憶されたコンピュータプログラムによって選択的に起動又は再構成される汎用コンピュータを含んでもよい。このようなコンピュータプログラムは、限定されないが、フロッピー（登録商標）ディスク、光学ディスク、ＣＤ－ＲＯＭ及び磁気光学ディスクを含む任意の種類のディスク、ＲＯＭ（ｒｅａｄ－ｏｎｌｙ　ｍｅｍｏｒｙ）、ＲＡＭ（ｒａｎｄｏｍ　ａｃｃｅｓｓ　ｍｅｍｏｒｙ）、各々がコンピュータシステムバスに結合されるＥＰＲＯＭ、ＥＥＰＲＯＭ、磁気若しくは光学カード又は一時的でない電子命令を記憶するのに適した任意の種類の媒体等の一時的でないコンピュータ可読記憶媒体に記憶されてもよい。一実施形態では、命令が記憶されたコンピュータ可読記憶媒体は、マルチテナントデータベース環境内の１つ以上のプロセッサに本明細書に記載の方法及び動作を実行させる。別の実施形態では、このような方法及び動作を実行するための命令は、後で実行するために一時的でないコンピュータ可読媒体に記憶される。

【００１６】
本明細書に提示されるアルゴリズム及びディスプレイは、本質的に特定のコンピュータ又は他の装置に関連せず、特定のプログラミング言語を参照して記載された実施形態でもない。様々なプログラミング言語が本明細書に開示された実施形態の教示を実装するために利用可能なことが理解されるであろう。
【００１７】
図１は、実施形態が動作可能な例示的なアーキテクチャ１００を示す。アーキテクチャ１００は、ネットワーク１２５を介して複数の顧客組織（１０５Ａ、１０５Ｂ及び１０５Ｃ）と通信可能にインターフェース接続されたホストシステム１１０を表す。ホストシステム１１０内でデータベース機能及びコード実行環境を実装する複数の基本ハードウェア、ソフトウェア及び論理要素１２０を有するマルチテナントデータベースシステム１３０がホストシステム１１０内にあり、マルチテナントデータベースシステム１３０のハードウェア、ソフトウェア及び論理要素１２０は、ネットワーク１２５を介してホストシステム１１０に通信可能にインターフェース接続することによりホストシステム１１０によって提供されるサービスを利用する複数の顧客組織（１０５Ａ、１０５Ｂ及び１０５Ｃ）とは離れており別個である。このような実施形態では、離れた別個の顧客組織（１０５Ａ－１０５Ｃ）の各々は、マルチテナントデータベースシステム１３０が実行されるホストシステム１１０を介して顧客組織（１０５Ａ－１０５Ｃ）にサービスを提供するホスト組織から遠く離れて設置されてもよい。代替的に、１つ以上の顧客組織１０５Ａ－１０５Ｃが、基本データが持続的に記憶されるマルチテナントデータベースシステム１３０を提供する同一のホスト組織内等、ホストシステム１１０と共同設置されてもよい。
【００１８】
一実施形態では、マルチテナントデータベースシステム１３０のハードウェア、ソフトウェア及び論理要素１２０は少なくとも非リレーショナルデータストア１５０及びリレーショナルデータストア１５５を含み、これらはホストシステム１１０内でデータベース機能及びコード実行環境を実装するハードウェア、ソフトウェア及び論理要素１２０に従って動作する。ホストシステム１１０は、ネットワークを介して複数の顧客組織１０５Ａ－１０５Ｃの１つ以上から要求１１５を更に受信してもよい。例えば、着信要求１１５は、マルチテナントデータベースシステム１３０内の顧客組織１０５Ａ－Ｃの１つのためにサービスに対する要求又はデータを取得若しくは記憶する要求に対応してもよい。
【００１９】
図２は、実施形態が動作可能な代替の例示的なアーキテクチャ２００を示す。一実施形態では、ホストシステム１１０は、リレーショナルデータストア１５５及び非リレーショナルデータストア１５０を有するマルチテナントデータベースシステム１３０からデータを取得する方法を実装する。
【００２０】
例えば、このような実施形態では、ホストシステム１１０でマルチテナントデータベースシステム１３０に対する要求１１５が受信される。この要求１１５はマルチテナントデータベースシステム１３０から取得されるデータ２１８を特定する。一部の実施形態では、ホストシステム１１０内で動作する別個のウェブサーバ２１０がネットワーク１２５を介して着信要求１１５を受信する。例えば、ウェブサーバ２１０は、ネットワーク１２５を介して様々な顧客組織１０５Ａ－Ｃから要求１１５を受信することを担当する。ウェブサーバ２１０は、要求１１５の起点となるエンドユーザクライアントマシン（例えば、顧客組織１０５Ａ－Ｃ内に設置されるエンドユーザ装置等）とのウェブベースインターフェースを提供してもよい。この要求１１５は、例えば、遠隔に実装されたクラウドコンピュータサービスを提供するホストシステム１１０等のホスト組織内で動作するマルチテナントデータベースシステム１３０からのサービスに対する要求を構成する。また、最適化エージェント２４５が所定の実施形態に従って事前クエリを展開すること及びデータクエリを最適化すること等の追加の機能を提供してもよい。
【００２１】

一実施形態では、ホストシステム１１０は、要求１１５に基づいて、取得されるデータ２１８の１つ以上の位置２１６を取得する。一実施形態では、顧客スキーマ２４０が取得されるデータ２１８の１つ以上の位置２１６を記述する。ここで、顧客スキーマ２４０は、非リレーショナルデータストア１５０内若しくはリレーショナルデータストア１５５内に存在する又は非リレーショナルデータストア１５０及びリレーショナルデータストア１５５の双方から利用可能である取得されるデータ２１８の複数のデータ要素の各々を特定する。一実施形態では、ホストシステム１１０は要求１１５の受信に応答して顧客スキーマ２４０を取得する。代替的に、ホストシステム１１０は、顧客スキーマ２４０から取得されるデータ２１８の１つ以上の位置２１６を取得する。

【００２２】

例えば、特定の実施形態において、取得されるデータ２１８の１つ以上の位置２１６は、取得されるデータ２１８を構成する複数のデータ要素の各々がマルチテナントデータベースシステム内のどこに設置されているかを特定する顧客スキーマ２４０に記憶され及びこれから取得される。このような顧客スキーマ２４０は、例えば、上記マルチテナント記憶能力を実装又は提供するホストシステム１１０の様々な要素に高速で効率的なアクセスを提供するグローバルキャッシュレイヤを介してアクセス可能であってもよい。代替的な実施形態では、取得されるデータ２１８の１つ以上の位置２１６は、ホストシステム１１０によって、最適化エージェント２４５によって、ホストシステム１１０のクエリレイヤ２６０によって、又は描かれている非リレーショナルデータストア１５０及びリレーショナルデータストア１５５に散在する複数のデータ要素を有するデータ２１８等、分散データベースの実装に散在するマルチテナントデータベースシステム１３０から取得されるデータ２１８の位置２１６を決定することに関与するホストシステム１１０の他の要素によって顧客スキーマ２４０から取得されてもよい。

【００２３】

一実施形態では、ホストシステム１１０は、要求１１５に基づいてデータベースクエリ２１７を生成する。ここで、データベースクエリ２１７は、取得される複数のデータ要素であって、非リレーショナルデータストア１５０内に存在する１つ以上のデータ要素及びリレーショナルデータストア１５５内に存在する１つ以上の他のデータ要素を含む複数のデータ要素を特定する。特定の実施形態では、データベースクエリ２１７は、取得されるデータ２１８の取得される１つ以上の位置２１６に更に基づく。このようなデータベースクエリ２１７は、クエリレイヤ２６０又は最適化エージェント２４５等のホストシステム１１０のサブシステムによって生成するためにホストシステム１１０によって更に生成又は代理されてもよい。

【００２４】

一実施形態では、ホストシステム１１０は、マルチテナントデータベースシステム１３０に対して生成されたデータベースクエリ２１７を実行して、図２によって描かれているようなデータ２１８を取得する。ここで、マルチテナントデータベースシステム１３０の複数の基本ハードウェア、ソフトウェア及び論理要素１２０に向かう下向き矢印は、マルチテナントデータベースシステム１３０の実装機能に渡されるデータベースクエリ２１７を描く。また、上向きに曲がった矢印によって描かれるマルチテナントデータベースシステムによって応答して戻されるデータ２１８は、分散データストア、非リレーショナルデータストア１５０及びリレーショナルデータストア１５５の各々を起点とする複数のデータ要素をホストシステム１１０に返信する。

【００２５】

一実施形態では、データベースクエリ２１７は複数のサブクエリを含む。このような実施形態では、複数のサブクエリの中の少なくとも１つは、非リレーショナルデータストア１５０内に存在する１つ以上のデータ要素を非リレーショナルデータストア１５０から取得するために行われる。また、複数のサブクエリの中の少なくとも第２のサブクエリは、リレーショナルデータストア１５５内に存在する１つ以上の他のデータ要素をリレーショナルデータストア１５５から取得するために行われる。例えば、「非リレーショナルデー

タストアからデータ要素‘ａ’を取得」（例えば、１５０）及び「リレーショナルデータストアからデータ要素‘ｂ’を取得」（例えば、１５５）等の複数のサブクエリ列並びに一般的なＳＱＬ（Ｓｔｒｕｃｔｕｒｅｄ　Ｑｕｅｒｙ　Ｌａｎｇｕａｇｅ）型のクエリを反映して「ｓｅｌｅｃｔ　‘ｘ’　ｆｒｏｍ　‘ｙ’　ｗｈｅｒｅ　‘ｚ’」を記述する別のサブクエリ列が、データベースクエリ２１７の拡大図内に図２によって描かれる。このようなクエリは、選択された実装クエリ言語又は文法に依存して基本データストア（例えば、１５０及び１５５）にクエリを行うのに適切であってもよく、又はそうでなくてもよい。

【００２６】

従って、このような実施形態に従って、マルチテナントデータベースシステム１３０に対してデータベースクエリ２１７を実行することは、必要なデータ２１８を取得するようにリレーショナルデータストア１５５及び非リレーショナルデータストア１５０の双方に記憶されたデータ要素を参照することを含む。

【００２７】

図３は、実施形態が動作可能な代替の例示的なアーキテクチャ３００を示す。特に、所定の実施形態に従ってデータベースクエリ２１７により特定される結合動作が更に描かれる。

【００２８】

例えば、一実施形態によれば、結合動作３０５はデータベースクエリ２１７によって特定される。

【００２９】

特定の実施形態では、結合動作３０５は複数のサブクエリを含む。例えば、このような実施形態では、第１のサブクエリ３０６が非リレーショナルデータストア１５０に対して実行されて、非リレーショナルデータストア１５０内に存在する１つ以上のデータ要素を識別する。これは非リレーショナルデータストア１５０への湾曲した破線の矢印によって描かれている。

【００３０】

このような実施形態では、第２のサブクエリ３０７がリレーショナルデータストア１５５に対して実行されて、非リレーショナルデータストア１５０内に存在する１つ以上のデータ要素を識別する第１のサブクエリとリレーショナルデータストア１５５内に存在する１つ以上の他のデータ要素との間のデータデルタ３１０を決定する。

【００３１】

この実施形態では、第３のサブクエリ３０８がリレーショナルデータストア１５５及び非リレーショナルデータストア１５０に対して実行される。第３のサブクエリはリレーショナルデータストア１５５から非リレーショナルデータストア１５０へと決定されたデータデルタ３１０に対応するデータを複製する。例えば、このような第３のサブクエリ３０８は、リレーショナルデータストア１５５内に存在する１つ以上の他のデータ要素を取得して、例えば、一時テーブル、ファイルに入れ、データを一時的にキャッシュ等してもよい。その後、このような第３のサブクエリ３０８は、非リレーショナルデータストア１５０に対してデータデルタ３１０に対応する取得データの挿入又は書き込みコマンドを発行して、非リレーショナルデータストア１５０にデータデルタ３１０のデータを書き込み、記憶し、又は挿入する。このようにして、複製を終了して、更にリレーショナルデータストア１５５に存在していたが以前に利用できなかったデータ要素を今度は非リレーショナルデータストア１５０から利用可能にする。図３の破線を参照すると、リレーショナルデータストア１５５から非リレーショナルデータストア１５０に識別されたデータデルタ３１０を複製するために、両データストア（リレーショナルデータストア１５５及び非リレーショナルデータストア１５０）に対して実行される第３のサブクエリが描かれている。

【００３２】

一方のデータストアから他方のデータストアにデータを複製又は同期するための決定は、様々な考慮事項に基づいてもよい。例えば、リレーショナルデータストア１５５から非

リレーショナルデータストア１５０に複製するための決定は、初期の場所からの小さなデータセットをより大きなデータセットを有する場所に複製するための決定又はポリシーに基づいてもよい。例えば、要求されたデータの一部である１つ以上のデータ要素は、リレーショナルデータストア１５５から非リレーショナルデータストア１５０の複製を行うためにネットワーク帯域幅の観点からその逆よりも効果的であってもよい。

【００３３】

一部の実施形態では、その反対も同様に当てはまり、データの複製は、非リレーショナルデータストア１５０からリレーショナルデータストア１５５へと反対方向に行ってもよい。このような決定は、例えば、最適化エージェント２４５によって実行され又は行われてもよい。非リレーショナルデータベースの実装（例えば、１５０）を用いる特定の実施形態では、リレーショナルデータベース型のオブジェクトがリレーショナルデータストア１５５（例えば、Ｏｒａｃｌｅ™）に記憶され、結合動作３０５を特定する１つのサブクエリを介して非リレーショナルデータストア１５０に複製される。その後、複製によって非リレーショナルデータストア１５０から全ての必要なデータが利用可能になるので、データ取得動作を特定する別のサブクエリが非リレーショナルデータストア１５０から全ての必要なデータを引き出してもよい。それにもかかわらず、このような例では、データの少なくとも一部は持続的に記憶されており、最初はリレーショナルデータストア１５５（例えば、Ｏｒａｃｌｅ™）からのみ利用可能である。

【００３４】

他の複製の決定及び考慮事項は、最適化エージェント２４５によって同様に考慮及び実装されてもよい。例えば、１つの複製ポリシーは、複製されたデータが常に同期されているか同期されることを保証されているかどうか、又は一部の逸脱が許容可能なリスクであるかどうか等、複製動作が提供する一貫性保証に基づいてもよい。

【００３５】

非リレーショナルデータストア１５０から１０億個のチャイルド行がクエリを行われ得るように、リレーショナルデータストア１５５から非リレーショナルデータストア１５０への１千万“Ａｃｃｏｕｎｔ” テーブルの複製を要求する複製動作を例にする。このような例では、非リレーショナルデータベースのクエリエンジンが、Ｏｒａｃｌｅ™（例えば、リレーショナル）データを取得するためにＪＤＢＣ（Ｊａｖａ（登録商標） Ｄａｔａｂａｓｅ Ｃｏｎｎｅｃｔｉｖｉｔｙ）を介して大量のコールアウトを行うために利用されてもよい。このような例では、Ｏｒａｃｌｅ™ＲＡＣ（Ｏｒａｃｌｅ Ｒｅａｌ Ａｐｐｌｉｃａｔｉｏｎ Ｃｌｕｓｔｅｒ）が利用されてもよい。又は、一貫性保証が重要な考慮事項であり、非常に大きなデータの複製が開始されているＯｒａｃｌｅ™１１ｇデータガードに基づく保証機構が必要な一貫性を提供するために利用されてもよい。データの着信要求に対して応答性のオンザフライでクエリ時にこのようにすることは着信要求への応答又は充足に際し許容できない長い遅延を必要とする場合があるので、このような大きな複製は最も適切には前もって行われてもよい。

【００３６】

逆に、小テーブルと対応する少ないデータ転送を例にする。このような例では、現在のデータを有するリレーショナルデータストア１５５（例えば、Ｏｒａｃｌｅ™）内に存在する１つ又は複数の小テーブルの全内容が、例えば、そのデータの少なくとも一部が小テーブル内に存在し且つリレーショナルデータストア１５５（例えば、Ｏｒａｃｌｅ™）によって持続的に記憶されるデータに対する着信要求に応答して、クエリ時に非リレーショナルデータストア１５０に複製されてもよい。このようなポリシーは、例えば、最適化エージェント２４５によって決定されるこの種のクエリ及び複製の固定コストが重大ではない大規模解析に適切であってもよい。

【００３７】

別の考慮事項は、例えば、特定の組織ごとに関連付けられるテーブル又はオブジェクトのサイズに依存するＯｒｇＩＤベースによるＯｒｇＩＤに関してもよい。例えば、中規模から大規模の組織に対応するデータ（例えば、既定のサイズ閾値に基づく）が前もって複

製されるポリシーが採用されてもよい。一実施形態では、事前複製は、非リレーショナルデータストア１５０内の事前複製テーブル（例えば、事前複製が選択されている）等、リレーショナルデータストア１５５内の特定のオブジェクトへの変更を補足して別のテーブルにこうした変更をプッシュするスキニーテーブル複製コード（ｓｋｉｎｎｙ　ｔａｂｌｅ　ｒｅｐｌｉｃａｔｉｏｎ　ｃｏｄｅ）を利用してもよい。

【００３８】

所定の実施形態では、特定の組織が大量の変更を引き起こす場合があるので、非リレーショナルデータストア１５０で要求された解析が実行されることを許可しつつも、リアルタイム同期は必ずしも変更ごとに要求されず又は適切ではない。従って、このような実施形態では、特定の間隔で更新を同期するオプションが（例えば、以下で検討される最適化エージェント２４５及びハードウェアベースのクエリレイヤエージェント５０１及び７３４を介して）提供される。このような実施形態では、特定の間隔での更新を可能にするポリシーは、結果未決の参照及び許容可能な逸脱であり得る又は採用された複製ポリシー若しくはデータベースクエリの基本目的に依存して後続のデータチェック及び検証を要求し得る他の「乱雑なデータ（ｓｌｏｐｐｙ　ｄａｔａ）」にもかかわらず、より効果的な書き込み及び更新を提供する。

【００３９】

複製に関する別の考慮事項は、テーブルの濃度等のデータストアの統計であってもよい。統計は、最適化エージェント２４５により生成され、利用可能であり、又は集められてもよい。例えば、比較的小規模な行の組が必要とされる所定の実施形態では、処理されるデータベースクエリ全体が大規模であり且つクエリが行われる行の組全体を送信することが合理的な送信コストを必要とすると決定される場合（例えば、既定のサイズ比又は固定閾値等に基づいて）、サブクエリ（例えば、３０６－３０９）がリレーショナルデータストアから１５５要求される行の組全体にクエリを行って、非リレーショナルデータストア１５０にクエリが行われた行の組全体を送信してもよい。このようなポリシーは、データの実際の要求を受信することなく事前複製又は前もって複製を行う必要性を回避してもよく、上記のデータの非一貫性の問題を更に回避してもよい。

【００４０】

データをどこに持続的に記憶すべきか、そしてデータをどこから取得すべきかの考慮事項は、特定のデータストアの基本的な実装ハードウェアに更に基づいてもよい。例えば、非リレーショナルデータストア１５０は、例えば、ギガバイトあたりのコストを単位として費用の安い記憶装置に大規模なフラットファイル及びバイナリファイルを記憶するように最適化されてもよい。非リレーショナルデータストア（例えば、１５０）が圧縮フラットファイル及びバイナリファイルの大量の読み込みに最適化されており、Ｏｒａｃｌｅ™等のリレーショナルモデルのデータストア（例えば、１５５）とは対照的にギガバイトあたりの要求を処理するのに計算コストが余りかからないので、このような基本のハードウェア実装が可能である。Ｏｒａｃｌｅ™等のリレーショナルデータストア１５５は、例えば、全ての記憶されたデータに必須の解析、トランザクション処理及びロールバック機能等の企業レベルのデータ保護を実装するためのリレーショナルデータストア１５５の要件によりギガバイトあたりではより高価な記憶ハードウェアを必要としてもよい。データベーストランザクションが最終的に完了する前に失敗又は中断する非一貫性の状態において、実行可能にデータストアがトランザクション処理を欠いたままであり得る直接挿入モデルとは対照的に、このような企業レベルのデータ保護は、トランザクション処理を介して、故障時に特定のトランザクションの「やり直し（ｒｅｄｏ）」を行う能力を更に可能にする。

【００４１】

従って、所定の実施形態では、記憶されたデータへの最近の編集が、解析、トランザクション処理及びロールバックを実装するリレーショナルデータストアによって処理及び記憶される。また、リレーショナルデータストア１５５に書き込まれた更新の少なくとも一部は、その後、対応するデータを持続的に記憶するギガバイト当たりのコストを減らすた

めに、非リレーショナルデータストア１５０に複製、移転、処理又は移動される。このようにして、ホストシステム１１０は、所定のリレーショナルデータストア１５５に関連付けられる企業レベルデータ保護を利用すると同時に一部の非リレーショナルデータストア１５０を介して利用可能なあまり費用のかからない持続的記憶装置の利益を享受してもよい。

【００４２】

一実施形態では、最適化エージェント２４５が非リレーショナルデータストア１５０及びリレーショナルデータストア１５５の双方から利用可能なデータの「ビュー」を有するので、最適化エージェント２４５は、データ２１８に対する着信要求１１５に対応する所定のデータ要素が１つより多くのソースから１つより多くの手法で取得され得る場合に、「選択的クエリ」をもたらすことができる。最適化エージェント２４５は、このような要求１１５を充足するために発行される複数のサブクエリの改善されたシーケンス又は順序付けを更にもたらすことができる。

【００４３】

従って、所定の実施形態に従って、様々な利用可能な考慮事項を考慮すると、第４のサブクエリ３０９が結合動作３０５内に更に含まれ、非リレーショナルデータストア１５０に対して実行される。ここで、第４のサブクエリ３０９は、非リレーショナルデータストア１５０内に存在する１つ以上のデータ要素とリレーショナルデータストア１５５から非リレーショナルデータストア１５０に複製されることで非リレーショナルデータストア１５０内から利用可能な１つ以上の他のデータ要素との双方をフェッチすることにより、非リレーショナルデータストア１５０から取得されるデータをフェッチする。このようにして、複数のデータ要素３１５は、結合動作３０５によってトリガされたデータ複製の前に複数のデータ要素３１５の一部が非リレーショナルデータストア１５０から最初に利用可能でないにもかかわらず、非リレーショナルデータストア１５０等のデータストアの１つから完全に取得されてもよい。

【００４４】

一部の実施形態では、要求されたデータの１つ以上のデータ要素が持続的に記憶される複数のデータストア間でデータを複製する代わりに、データを持続的に記憶するクエリを受ける２つ以上のデータストア（例えば、１５０及び１５５）の各々から離れた場所に全てのデータ要素を取得するようにポリシーが用いられてもよい。例えば、データは、クエリレイヤ２６０へのメモリ内結合動作を利用して取得され、又はグローバルキャッシュレイヤ（例えば、図５の要素５５０）へのメモリ内結合動作を介して取得されてもよい。このようなメモリ内結合動作は、最適化エージェント２４５から利用可能な既知の統計に基づいて、又は特定されたサイズ閾値（例えば、行の数、サイズ（例えば、データのメガバイト数）を単位とするデータ量、要求されたデータの濃度等）に基づいて選択されてもよい。

【００４５】

例えば、ホストシステム１１０内の既知の統計及び解析に基づいて利用可能な他の考慮事項は、例えば、要素の最大数が知られており、最大又は推定クエリコストが決定可能若しくは知られていて最適化エージェント２４５から利用可能である特定のクエリに対する既知の選択リスト量から導かれる特定のデータベースクエリ２１７又はサブクエリ（例えば、３０６－３０９）に対するクエリコストを含んでもよい。それは更に、既に行われた解析に基づいて知られ、又は複数の利用可能なデータストア（例えば、１５０及び１５５）の中で、最少量の時間で又は最小の計算リソースを利用／消費して最少数を有する結果をもたらすことができる最適化エージェント２４５を介して（例えば、１つ以上の事前クエリを介して）決定可能であってもよい。例えば、大きなデータベースクエリ２１７では、各データストア（例えば、１５０及び１５５）から必要なデータの小部分に対して事前クエリを行って、どの事前クエリがより効果的な結果をもたらすかを決定し、このような決定に基づいて、より効果的なデータストア（例えば、事前クエリの結果によって（１５０又は１５５）を目標にする一次データベースクエリ２１７を充足するために必要とされ

る様々なサブクエリ（３０６－３０９）を生成することが望ましくてもよい。最適化エージェント２４５によって適切な解析が前もって行われる場合、事前クエリを発行しなくても、クエリポリシーが簡単に要求されてもよい。例えば、このような解析決定が行われて、顧客スキーマ２４０を介して１つ以上のデータの場所に対して記憶及び特定されてもよい。

【００４６】

代替的な実施形態では、異なる又は追加の結合動作３０５が行われてもよい。例えば、一実施形態では、データベースクエリ２１７を介してマルチテナントデータベースシステム１３０に対して実行される結合動作３０５は、リレーショナルデータストア１５５から２つ以上の関連テーブルを特定する結合動作３０５、リレーショナルデータストア１５５及び非リレーショナルデータストア１５０内に存在する少なくとも１つ以上のデータ構造から少なくとも１つの関連テーブルを特定する結合動作３０５、及び非リレーショナルデータストア１５０内に存在する２つ以上の離れた及び別個のデータ構造を特定する結合動作３０５からなる結合動作３０５のグループから選択される結合動作３０５を含んでもよい。ここで、２つ以上の離れた及び別個のデータ構造の各々は、非リレーショナルデータストア１５０内で２つの別個のデータ構造を関連付け又は関係付けるための共有された特徴、ストリング、バイナリ又は英数字キー等の重複した共有キーを欠いている。

【００４７】

例えば、一実施形態では、非リレーショナルデータストア１５０は、多くのデータ構造、ファイル、オブジェクト及び他のこうした情報を記憶する能力を提供するが、このようなデータ構造、ファイル、オブジェクト及び他の情報を「関連付け」する機能を実装しない。しかしながら、２つの別個のデータ構造の各々を特定する結合動作３０５は、例えば、所望されるが以前に非関連付けされた全ての情報を有する単一データ構造を形成し、又は代替の場所でこのような結合動作によって特定される所望の情報を取得及び一時的にキャッシュするために、非リレーショナルデータストア１５０に対して適切に形成された結合動作３０５を実行することができるマルチテナントデータベース１３０内の複数の基本ハードウェア、ソフトウェア、及び論理要素１２０等、実装された非リレーショナルデータストア１５０の外部の機能及び論理に依存することにより、各々のこのようなデータ構造を識別及びリンク又は関係付けることができる。

【００４８】

代替的な実施形態では、　特定された結合動作３０５は、非リレーショナルデータストア１５０に対して実行され且つ非リレーショナルデータストア内に存在する１つ以上のデータ要素を取得するための第１のサブクエリ（例えば、３０６）、リレーショナルデータストア１５５に対して実行され且つ非リレーショナルデータストア１５０内に存在する１つ以上のデータ要素とリレーショナルデータストア１５５内に存在する１つ以上の他のデータ要素との間のデータデルタ３１０を決定する第２のサブクエリ（例えば、３０７）、及びリレーショナルデータストア１５５に対して実行され且つ決定されたデータデルタ３１０に基づいてリレーショナルデータストア１５５内に存在する１つ以上の他のデータ要素を取得するための第３のサブクエリ（例えば、３０８）を含む。

【００４９】

このような実施形態では、決定されたデータデルタ３１０に基づいてリレーショナルデータストア１５５内に存在する１つ以上の他のデータ要素を取得する第３のサブクエリ（例えば、３０８）は、時間ベースのクエリフィルタ機構又は記録ベースのフィルタ機構の何れかを含んでもよい。

【００５０】

例えば、一実施形態では、時間ベースのサブクエリは、リレーショナルデータストア１５５に対して実行される。ここで、時間ベースのサブクエリは、非リレーショナルデータストア１５０内に存在する１つ以上のデータ要素内のデータ要素に対応するどのタイムスタンプよりも遅いタイムスタンプを有するリレーショナルデータストア１５５内のデータ要素に基づいてリレーショナルデータストア１５５から取得される１つ以上の他のデータ

要素を特定する。

【００５１】

代替的な実施形態では、記録識別子ベースのサブクエリは、リレーショナルデータストア１５５に対して実行される。ここで、記録ベースのサブクエリは、非リレーショナルデータストア１５０内に存在する１つ以上のデータ要素内のデータ要素に対応するどの記録識別子よりも数値が大きい記録識別子を有するリレーショナルデータストア１５５内のデータ要素に基づいてリレーショナルデータストア１５５から取得される１つ以上の他のデータ要素を特定する。

【００５２】

図４は、実施形態が動作可能な代替の例示的なアーキテクチャ４００を示す。特に、所定の実施形態に従うマルチテナントデータベースシステムによって受信される新しいトランザクションの処理が更に詳細に描かれている。

【００５３】

例えば、所定の実施形態では、持続的な記憶のためにマルチテナントデータベースシステム１３０に書き込み又は挿入される新しい情報は、長期間にわたって非リレーショナルデータストア１５０に現在時に記憶されるように指定されてもよいが、それにもかかわらず一時的にリレーショナルデータストア１５５に書き込まれてもよい。例えば、一時的に一方のデータストア（リレーショナルデータストア１５５等）にデータを書き込んで、後に別のデータストア（非リレーショナルデータストア１５０等）にデータを移行するための考慮事項は、例えば、一方のデータストア対他方の書き込み応答時間の改善、更に代替のデータストからの取得時間の改善を含んでもよい。一方のデータストアは、より低い計算又はより低い動作コストと関連してもよい。非リレーショナルデータストア１５０等の特定のデータストアが、滅多に更新されないが頻繁に取得されるデータでより効果的に動作してもよい。代替的に、リレーショナルデータストア１５５等の他のデータストアが、著しく断片化されたデータ又は滅多に更新されないデータを有する前の例示と比較して非常に頻繁に更新され又は追加されるデータでより良好な動作効率を示してもよい。

【００５４】

従って、所定の実施形態によれば、（例えば、先に描かれた要求１１５内の）マルチテナントデータベースシステム１３０において受信される新しいトランザクションは、非リレーショナルデータストア１５０に書き込まれる新しいデータ４１６を含み又は特定する。一部の実施形態では、新しいデータ４１６は、新しいデータ４１６が非リレーショナルデータストア１５０に書き込まれるという指示にかかわらず、リレーショナルデータストア１５５の追加ログ４１０に書き込まれる。新しいデータ４１６が書き込まれるというこのような指示は、新しいトランザクション４１５によって、例えば、新しいトランザクション４１５内のターゲット４１９の属性によって特定されてもよい。代替的に、例えば、最適化エージェント２４５によって決定される新しいデータ４１６の特徴に基づいて、又は新しいトランザクション４１５に対応するＯｒｇＩＤと関連付けられるフラグ若しくは記憶された設定に基づいて、ホストシステム１１０によって決定が行われてもよい。

【００５５】

一部の実施形態では、決定されたデータデルタ（例えば、３１０）に基づいてリレーショナルデータストア１５５内に存在する１つ以上の他のデータ要素を取得するためのサブクエリを含む結合動作（例えば、３０５）は、リレーショナルデータストア１５５の追加ログ４１０からリレーショナルデータストア１５５内に存在する１つ以上の他のデータ要素を取得するためのサブクエリを含む。例えば、追加ログに書き込まれた新しいデータ４１６が取得されてもよく、又は追加ログに記憶された新しいデータ４１６の要素が取得されてもよい。

【００５６】

一実施形態では、ホストシステム１１０は、追加ログがフラッシュ閾値に到達すると、追加ログ４１０のフラッシュをトリガして、リレーショナルデータストア１５５の追加ログ４１０に書き込まれた新しいデータ４１６を非リレーショナルデータストア１５０にフ

ラッシュする。その結果、例えば、新しいデータはフラッシュされたデータ４１７として非リレーショナルデータストアに存在し、リレーショナルデータストア１５５の追加ログ４１０に以前に存在した新しいデータ４１６に対応することになる。

【００５７】

異なる種類のデータがマルチテナントデータベースシステム１３０によって記憶されてもよい。例えば、一実施形態では、非リレーショナルデータストア１５０内に存在する１つ以上のデータ要素は、複数の圧縮フラットファイル又は複数のバイナリファイル又は圧縮フラットファイル及びバイナリファイルの組み合わせに対応する。このようなファイルは、非リレーショナルデータベースアーキテクチャ（例えば、１５０）を介してより効果的に記憶されてもよい。

【００５８】

別の実施形態では、リレーショナルデータストア１５５は、リレーショナルデータベース管理システム（ＲＤＢＭＳ）に従うリレーショナルデータベースを実装する。ここで、リレーショナルデータベースの複数の関連テーブルは、リレーショナルデータベース内の２つ以上の関連テーブルの各々に対して１つ以上の重複する共通の特徴を介して相互に関連付けされ、これによりリレーショナル型データストア１５５と共通して関連付けられる「リレーションシップ」を形成する。

【００５９】

一実施形態では、非リレーショナルデータストア１５０は、各々が非リレーショナルデータストア１５０に総記憶容量の少なくとも一部を提供する複数の基本ハードウェア記憶装置を有する分散構造型データベースを実装する。このような実施形態では、非リレーショナルデータストア１５０内のデータ要素は、主キーに基づいて参照可能であるが、リレーショナルデータストア１５５内のデータ要素の場合におけるように、２つ以上の関連テーブルの間の１つ以上の重複する共通の特徴に基づいて参照可能ではない。

【００６０】

一実施形態では、リレーショナルデータストア１５５は、Ｏｒａｃｌｅ互換データベースの実装、ＩＢＭ　ＤＢ２　Ｅｎｔｅｒｐｒｉｓｅ　Ｓｅｒｖｅｒ互換リレーショナルデータベースの実装、ＭｙＳＱＬ互換リレーショナルデータベースの実装、及びＭｉｃｒｏｓｏｆｔ　ＳＱＬ　Ｓｅｒｖｅｒ互換リレーショナルデータベースの実装の中から選択されるリレーショナルデータベースモデルを実装する。

【００６１】

一実施形態では、非リレーショナルデータストア１５０は、Ｖａｍｐｉｒｅ互換非リレーショナルデータベースの実装、Ａｐａｃｈｅ　Ｃａｓｓａｎｄｒａ互換非リレーショナルデータベースの実装、ＢｉｇＴａｂｌｅ互換非リレーショナルデータベースの実装、及びＨＢａｓｅ互換非リレーショナルデータベースの実装の中から選択されるＮｏＳＱＬ非リレーショナルデータベースを実装する。

【００６２】

一実施形態では、非リレーショナルデータストア１５０は、各コンピュータノードが少なくとも１つのメモリ、１つ以上のプロセッサ及び１つ以上の通信可能にインターフェース接続されたハードディスクドライブを含む複数の分散型コンピュータノードを含む。このような実施形態では、分散型コンピュータノードの各々は、中央トランザクション局からの承認又は制御無しで非リレーショナルデータベーストランザクションを読み込み、書き込み、及び更新する機能を有する分離した非リレーショナルデータベースインスタンスを更に含んでもよい。

【００６３】

特定の実施形態では、リレーショナルデータストア１５５は、モノリシックリレーショナルデータベースインスタンス（ｍｏｎｏｌｉｔｈｉｃ　ｒｅｌａｔｉｏｎａｌ　ｄａｔａｂａｓｅ　ｉｎｓｔａｎｃｅ）への更新又は変更がモノリシックリレーショナルデータベースインスタンスに通信可能にインターフェース接続され且つ制御される持続的記憶装置に対する持続的記憶にコミットされるかどうかを制御する中央トランザクション局と計

算リソースを調整するプロセッサ及びメモリを含むモノリシックリレーショナルデータベースインスタンスを実装する。

【００６４】

図５は、実施形態が動作し、設置され、統合され、又は構成され得るシステム５００の図表示を示す。

【００６５】

一実施形態では、システム５００は、メモリ５９５及び１つ以上のプロセッサ５９０を含む。例えば、メモリ５９５は実行される命令を記憶してもよく、プロセッサ５９０はこのような命令を実行してもよい。システム５００は、バス５１５と通信可能にインターフェース接続される複数の周辺機器の間でシステム５００内のトランザクション及びデータを転送するためのバス５１５を含む。システム５００は、例えば、要求を受信し、応答を返し、顧客組織１０５ＡＣ内に設置されたクライアント装置等のリモートクライアントとインターフェース接続するウェブサーバ５２５を更に含む。

【００６６】

システム５００は、データベースクエリ及びデータベースサブクエリを最適化して、基本データストアにクエリを行う最適な又は好ましい手法を決定するために選択的に事前クエリを調整するように設計される最適化エージェント５３５を有するように更に描かれる。システム５００は、通信可能にインターフェース接続された装置及びシステムにキャッシュサービスを提供し、特に、顧客スキーマデータ（例えば、メタデータ等）をキャッシュをするグローバルキャッシュレイヤ５５０を更に含む。顧客スキーマデータは、グローバルキャッシュレイヤ５５０と連動して動作可能な顧客スキーマ５３０によって提供され、例えば、必須データ要素がマルチテナントデータベースシステム内のリレーショナルデータベース若しくは非リレーショナルデータベースの実装又はその両方によって記憶されるかどうかを特定し、対応する要求に対するデータセットを構成する１つ以上のデータ要素に関して基本データストア内の場所を特定する。顧客スキーマ５３０は、システム５００内のハードドライブ、持続性データストア又は他の記憶場所に記憶されてもよい。

【００６７】

ハードウェアベースのクエリレイヤエージェント５０１はシステム５００内で別個にあり、要求プロセッサ５７０、顧客スキーマプロセッサ５７５、サブクエリ生成器５８０及びクエリ実行器５８５を含む。一実施形態によれば、要求プロセッサ５７０は、（例えば、ウェブサーバ５２５から、上述のようにホストシステム１１０から、又は接続されたネットワークインターフェースから）取得するデータを特定する要求を受信する。要求プロセッサ５７０は、顧客スキーマプロセッサ５７５と連動して、基本データストアから取得される要求データの１つ以上の位置を取得する。要求プロセッサ５７０は、更にサブクエリプロセッサ５８０と連動して、決定された１つ以上のこのようなデータの位置に基づいて適切な基本データストアから要求された１つ以上のデータ要素を取得するために必要なサブクエリを構築及び生成し、又は一方のデータストアから他方にデータサブセットを同期、フラッシュ又は複製させる結合動作を開始するために必要なサブクエリを生成し、その結果、後続のサブクエリは１つのデータストアから要求されたデータセットを取得することができる。サブクエリ生成器５８０によって生成されるこのようなサブクエリは、最適化エージェント５３５から利用可能な統計及び事前クエリの結果に依存してもよい。クエリ実行器５８５は、通信可能にインターフェース接続されたデータベースの実装に対して生成されたクエリ及びサブクエリを実行する。

【００６８】

図６は、一実施形態に従ってマルチテナントストアで横断的ストア結合を実行する方法６００を示す流れ図であり、所定の実施形態に従ってデータベースクエリ（例えば、２１７）によって特定される結合動作を特定及び実行することを含む。方法６００は、ハードウェア（例えば、回路、専用論理、プログラム可能論理、マイクロコード等）、ソフトウェア（例えば、読み込み、書き込み、更新、最適化、事前クエリの開始、サブクエリの開始等、又はこれらの組み合わせ等の様々なクエリ動作を行うために処理装置上で実行され

る命令）を含み得る論理を処理することによって実行されてもよい。一実施形態では、方法６００は、図５の要素５０１で描かれるハードウェアベースのクエリレイヤ等のハードウェア論理によって実行される。所定の実施形態によれば以下にリストされるブロック及び／又は動作の一部は選択的である。提示されたブロックの番号は明確のためであって、様々なブロックが行われなければならない動作の順番を規定することを意図していない。

【００６９】

方法６００は、ホストシステムでマルチテナントデータベースシステムに対する要求を受信する処理論理で開始する。この要求はマルチテナントデータベースシステムから取得されるデータを特定する（ブロック６０５）。ブロック６１０では、ホストシステムを介する要求に基づいて、処理論理は取得されるデータの１つ以上の位置を取得する。

【００７０】

ブロック６１５では、ホストシステムを介して、処理論理は要求の受信に応答して顧客スキーマを取得する。例えば、顧客スキーマが取得されるデータの１つ以上の位置２１６を記述してもよい。ここで、顧客スキーマは、非リレーショナルデータストア内若しくはリレーショナルデータストア内に存在する又は非リレーショナルデータストア及びリレーショナルデータストアの双方から利用可能であるデータの複数のデータ要素の各々を特定する。

【００７１】

ブロック６２０では、ホストシステムを介して、処理論理は要求に基づいてデータベースクエリを生成する。例えば、データベースクエリは、取得される複数のデータ要素であって、非リレーショナルデータストア内に存在する１つ以上のデータ要素及びリレーショナルデータストア内に存在する１つ以上の他のデータ要素を含む複数のデータ要素を特定してもよい。データベースクエリは複数のサブクエリを更に含んでもよい。一実施形態では、データベースクエリはサブクエリの１つを介して結合動作を特定する。同様に、パージ、フラッシュ、同期又は複製動作がサブクエリを介して特定されてもよい。

【００７２】

ブロック６２５において、処理論理はデータを取得するためにマルチテナントデータベースシステムに対してデータベースクエリを実行する。

【００７３】

ブロック６３０では、処理論理はマルチテナントデータベースシステムで新しいトランザクションを取得する。新しいトランザクションの各々は非リレーショナルデータストアに書き込まれる新しいデータを特定する。ブロック６３５において、処理論理はリレーショナルデータストアの追加ログに新しいデータを書き込む。例えば、一実施形態では、データベースクエリのサブクエリは、取得されるデータがリレーショナルデータストアの追加ログから取得されることを特定する。

【００７４】

ブロック６４０では、追加ログがフラッシュ閾値に到達すると、処理論理はリレーショナルデータストアの追加ログに書き込まれた新しいデータを非リレーショナルデータストアにフラッシュする。

【００７５】

図７は、コンピュータシステムの例示的形態のマシン７００の図形表現を示す。一実施形態によれば、このコンピュータシステムの中で、マシン７００に本明細書で検討された１つ以上の方法を実行させるために、一組の命令が実行されてもよい。代替的な実施形態では、マシンは、ＬＡＮ（Ｌｏｃａｌ　Ａｒｅａ　Ｎｅｔｗｏｒｋ）、イントラネット、エクストラネット、又はインターネットで他のマシンと接続（例えば、ネットワーク化）されてもよい。マシンは、クライアントサーバネットワーク環境におけるサーバ又はクライアントマシンの能力で、又はピアツーピア（又は分散ネットワーク環境）におけるピアマシンとして又はマルチテナントデータベース記憶サービスを提供するオンデマンド環境を含むオンデマンドサービス環境内のサーバ若しくは一連のサーバとして動作してもよい。マシンの所定の実施形態は、パーソナルコンピュータ（ＰＣ）、タブレットＰＣ、セッ

トトップボックス（ＳＴＢ）、パーソナルデジタルアシスタント（ＰＤＡ）、携帯電話、ウェブ装置、サーバ、ネットワークルータ、スイッチ若しくはブリッジ、コンピュータシステム、又はそのマシンによって取り行われる動作を特定する一組の命令（連続的又はそれ以外）を実行することが可能な任意のマシンの形態であってもよい。更に、１つのマシンのみが示されているが、「マシン」という用語は、本明細書で検討された１つ以上の方法を実行する一組（又は複数の組）の命令を個別の又は共同で実行する複数のマシン（例えば、コンピュータ）の集合を含むようにも解釈される。

【００７６】

例示的なコンピュータシステム７００は、プロセッサ７０２、主メモリ７０４（例えば、ＲＯＭ（ｒｅａｄ－ｏｎｌｙ　ｍｅｍｏｒｙ）、フラッシュメモリ、ＳＤＲＡＭ（ｓｙｎｃｈｒｏｎｏｕｓ　ＤＲＡＭ）若しくはＲＤＲＡＭ（Ｒａｍｂｕｓ　ＤＲＡＭ）等のＤＲＡＭ（ｄｙｎａｍｉｃ　ｒａｎｄｏｍ　ａｃｃｅｓｓ　ｍｅｍｏｒｙ）、フラッシュメモリ等のスタティックメモリ、ＳＲＡＭ（ｓｔａｔｉｃ　ｒａｎｄｏｍ　ａｃｃｅｓｓ　ｍｅｍｏｒｙ）、揮発性であるが高データ速度のＲＡＭ等）、及び二次メモリ７１８（例えば、ハードディスクドライブを含む持続性記憶装置及び持続性マルチテナントデータベースの実装）を含み、これらはバス７３０を介して互いに通信する。主メモリ７０４は、（例えば、リレーショナルデータストア及び非リレーショナルデータストアの双方に散在し、ハードウェアベースのクエリレイヤエージェント７３４を介して取得可能なデータ要素の場所等、２つ以上の異なったデータストアの中で特定のデータ又はデータセットを構成するデータ又はデータ要素の１つ以上の位置を特定する）顧客スキーマ７２４を含む。主メモリ７０４は、顧客スキーマ７２４を介して提供される情報の種類等の大規模データセットの複数のデータ要素間のメタデータ及び他の関連又は対応情報を提供するために、システム全体でアクセス可能なグローバルキャッシュレイヤ等のグローバルキャッシュレイヤ７２３を更に含む。主メモリ７０４及びそのサブ要素（例えば、７２３及び７２４）は、本明細書で検討された方法を実行するために処理論理７２６及び７０２と連動して動作可能である。

【００７７】

プロセッサ７０２は、マイクロプロセッサ又は中央処理装置等の１つ以上の汎用処理装置を表す。より詳細には、プロセッサ７０２は、ＣＩＳＣ（ｃｏｍｐｌｅｘ　ｉｎｓｔｒｕｃｔｉｏｎ　ｓｅｔ　ｃｏｍｐｕｔｉｎｇ）マイクロプロセッサ、ＲＩＳＣ（ｅｄｕｃｅｄ　ｉｎｓｔｒｕｃｔｉｏｎ　ｓｅｔ　ｃｏｍｐｕｔｉｎｇ）マイクロプロセッサ、ＶＬＩＷ（ｖｅｒｙ　ｌｏｎｇ　ｉｎｓｔｒｕｃｔｉｏｎ　ｗｏｒｄ）マイクロプロセッサ、他の命令セットを実装するプロセッサ、又は命令セットの組み合わせを実装するプロセッサであってもよい。また、プロセッサ７０２は、ＡＳＩＣ（ａｐｐｌｉｃａｔｉｏｎ　ｓｐｅｃｉｆｉｃ　ｉｎｔｅｇｒａｔｅｄ　ｃｉｒｃｕｉｔ）、ＦＰＧＡ（ｆｉｅｌｄ　ｐｒｏｇｒａｍｍａｂｌｅ　ｇａｔｅ　ａｒｒａｙ）、ＤＳＰ（ｄｉｇｉｔａｌ　ｓｉｇｎａｌ　ｐｒｏｃｅｓｓｏｒ）、又はネットワークプロセッサ等の１つ以上の専用処理装置であってもよい。プロセッサ７０２は、動作及び本明細書で検討された機能を実行するための処理論理７２６を実行するように構成される。

【００７８】

コンピュータシステム７００は、ネットワークインターフェースカード７０８を更に含んでもよい。また、コンピュータシステム７００は、ユーザインターフェース７１０（ビデオディスプレイ装置、ＬＣＤ（ｌｉｑｕｉｄ　ｃｒｙｓｔａｌ　ｄｉｓｐｌａｙ）若しくはＣＲＴ（ｃａｔｈｏｄｅ　ｒａｙ　ｔｕｂｅ）等）、英数字入力装置７１２（例えば、キーボード）、カーソル制御装置７１４（例えば、マウス）及び信号生成装置７１６（例えば、統合スピーカ）を含んでもよい。コンピュータシステム７００は、周辺装置７３６（例えば、無線若しくは有線通信装置、メモリ装置、記憶装置、オーディオ処理装置、ビデオ処理装置等）を更に含んでもよい。コンピュータシステム７００は、データベースクエリ及びサブクエリを管理し、マルチテナントデータベースシステム等の基本データストアとトランザクションを調整するハードウェアベースのクエリレイヤエージェント７３

４を更に含んでもよい。

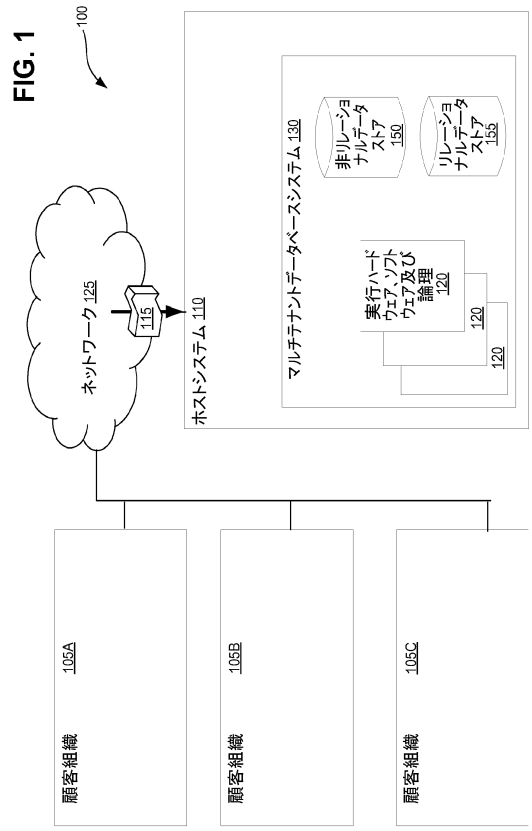
【００７９】

二次メモリ７１８は、本明細書で検討された１つ以上の方法又は機能を具現化する１つ以上の命令の組（例えば、ソフトウェア７２２）が記憶される一時的でないマシン可読記憶媒体（又は、より詳細には、一時的でないマシンアクセス可能記憶媒体）７３１を含んでもよい。また、ソフトウェア７２２が、コンピュータシステム７００によってその実行中に主メモリ７０４内に及び／又はプロセッサ７０２内に完全に又は少なくとも部分的に存在してもよい。ここで、主メモリ７０４及びプロセッサ７０２はマシン可読記憶媒体も構成する。更に、ソフトウェア７２２は、ネットワークインターフェースカード７０８を介してネットワーク７２０上で送信又は受信されてもよい。

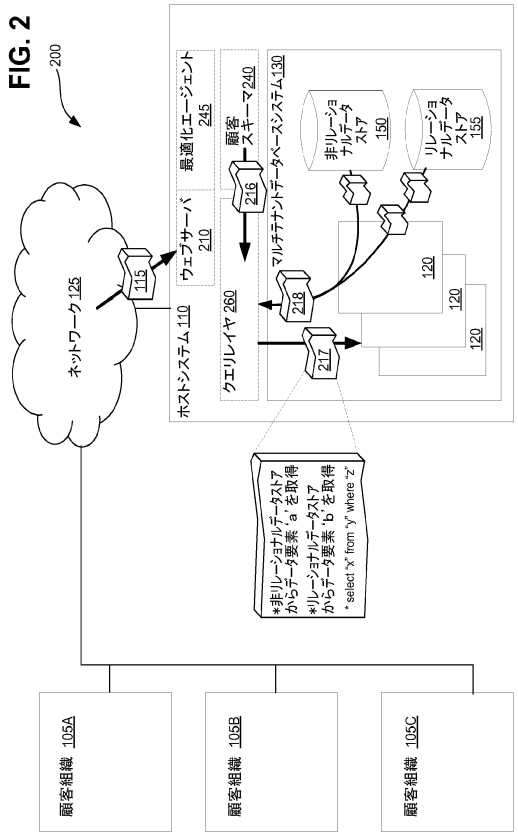
【００８０】

本明細書で検討された主題は例示として特定の実施形態に関して記載されているが、本発明の実施形態は明示的に列挙された開示の実施形態に限定されないことが理解されるべきである。それとは反対に、本開示は当業者に明白な様々な修正及び類似の構成を含むことが意図されている。従って、添付の特許請求の範囲は、全てのこのような修正及び類似の構成を包含するように最も広い解釈に一致するべきである。先の記載は例示的であり限定的ではないことが理解されるべきである。先の記載を読んで理解すると、多くの他の実施形態が当業者には明白であろう。従って、本開示の主題の範囲は、添付の請求項が享受する均等物の全範囲と共に、このような請求項を参照することで決定されるべきである。

【図１】



【図２】



【図３】

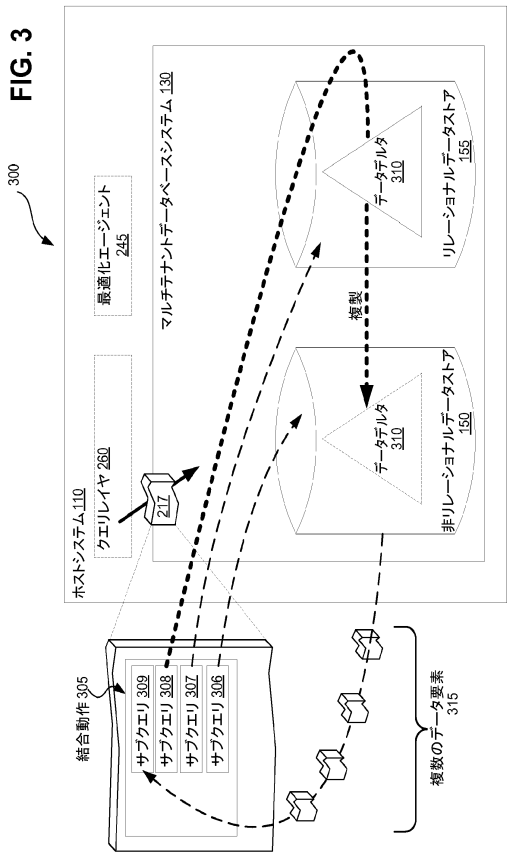


FIG. 3
300
ホストシステム 110
クエリレイヤ 260
217
最適化エージェント 245
マルチテナントデータベースシステム 130
データデルタ 310
リレーショナルデータストア 155
複製
データデルタ 310
非リレーショナルデータストア 150
結合動作 305
サブクエリ 309
サブクエリ 308
サブクエリ 307
サブクエリ 306
複数のデータ要素 315

【図４】

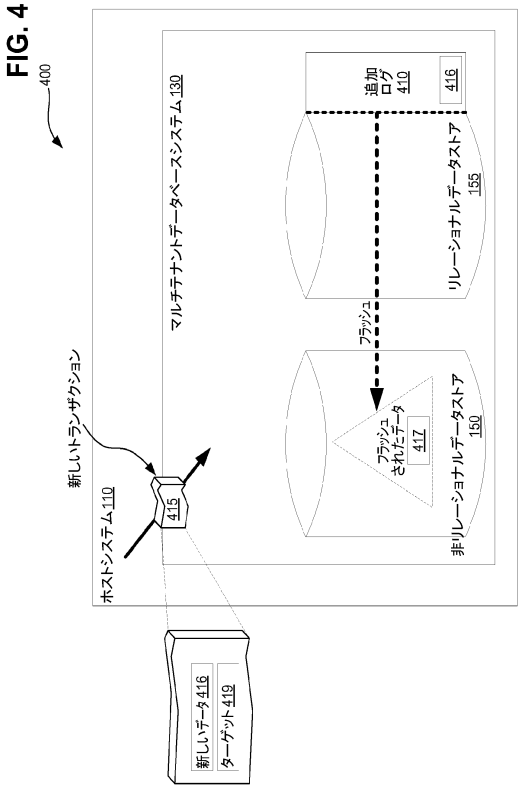


FIG. 4
400
マルチテナントデータベースシステム 130
追加ログ 410
416
リレーショナルデータストア 155
フラッシュ
新しいトランザクション
ホストシステム 110
415
フラッシュされたデータ 417
非リレーショナルデータストア 150
新しいデータ 416
ターゲット 419

【図５】

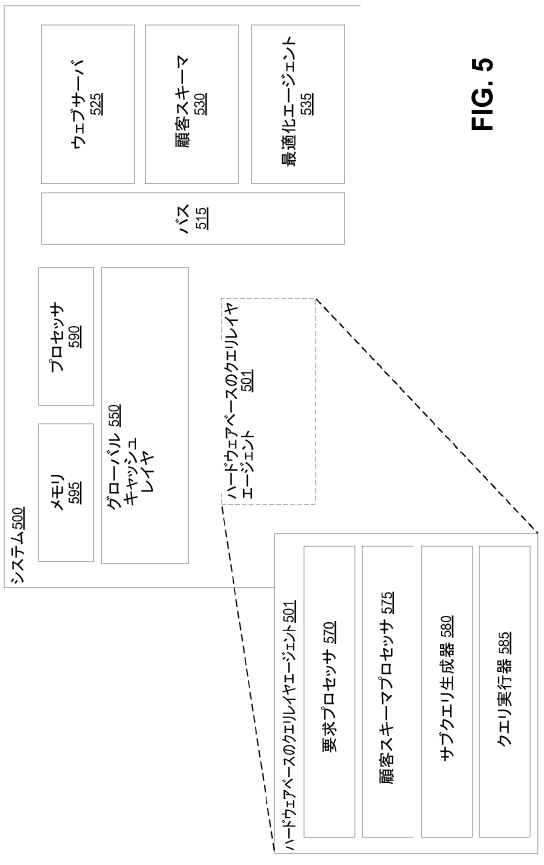


FIG. 5
システム 500
ウェブサーバ 525
顧客スキーマ 530
最適化エージェント 535
バス 515
プロセッサ 590
メモリ 595
グローバル キャッシュ レイヤ 550
ハードウェアベースのクエリレイヤ エージェント 501
ハードウェアベースのクエリレイヤエージェント 501
要求プロセッサ 570
顧客スキーマプロセッサ 575
サブクエリ生成器 580
クエリ実行器 585

【図６】

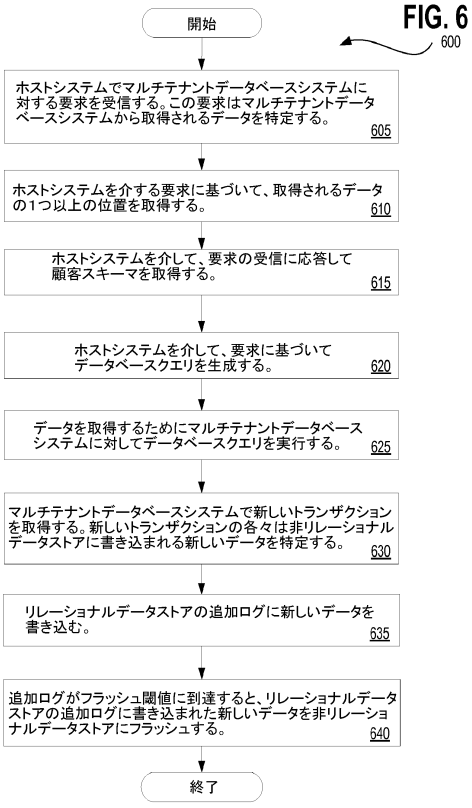


FIG. 6
600
開始
ホストシステムでマルチテナントデータベースシステムに対する要求を受信する。この要求はマルチテナントデータベースシステムから取得されるデータを特定する。 605
ホストシステムを介する要求に基づいて、取得されるデータの1つ以上の位置を取得する。 610
ホストシステムを介して、要求の受信に応答して顧客スキーマを取得する。 615
ホストシステムを介して、要求に基づいてデータベースクエリを生成する。 620
データを取得するためにマルチテナントデータベースシステムに対してデータベースクエリを実行する。 625
マルチテナントデータベースシステムで新しいトランザクションを取得する。新しいトランザクションの各々は非リレーショナルデータストアに書き込まれる新しいデータを特定する。 630
リレーショナルデータストアの追加ログに新しいデータを書き込む。 635
追加ログがフラッシュ閾値に到達すると、リレーショナルデータストアの追加ログに書き込まれた新しいデータを非リレーショナルデータストアにフラッシュする。 640
終了

【図７】

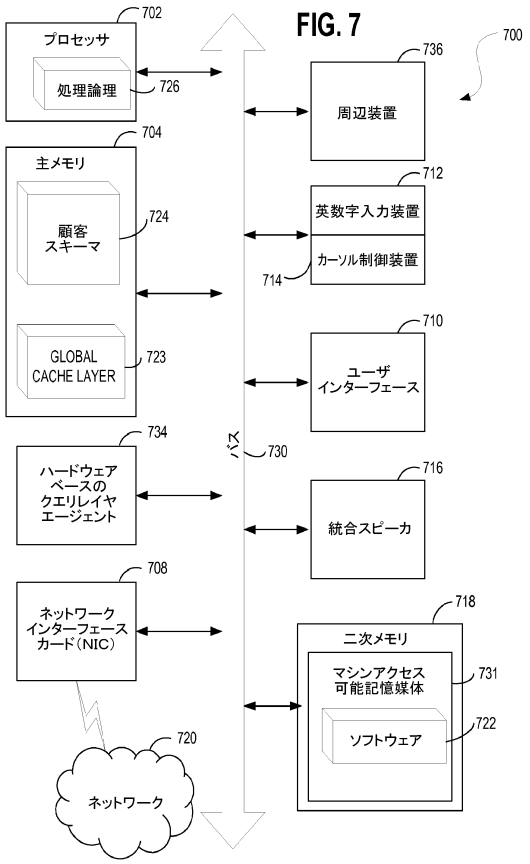


FIG. 7
700
プロセッサ
702
処理論理
726
主メモリ
704
顧客
スキーマ
724
GLOBAL
CACHE LAYER
723
ハードウェア
ベースの
クエリレイヤ
エージェント
734
ネットワーク
インターフェース
カード（NIC）
708
ネットワーク
720
バス
730
周辺装置
736
英数字入力装置
712
カーソル制御装置
714
ユーザ
インターフェース
710
統合スピーカ
716
二次メモリ
718
マシンアクセス
可能記憶媒体
731
ソフトウェア
722

フロントページの続き

(72)発明者 イードスン、 ビル シー.
アメリカ合衆国 ９４３０３ カリフォルニア州 パロアルト コリーナ ウェイ ３８７９
(72)発明者 ワイスマン、 クレイグ
アメリカ合衆国 ９４１１５ カリフォルニア州 サンフランシスコ サクラメント ストリート ２８３８
(72)発明者 オリバー、 ケヴィン
アメリカ合衆国 ９４１２７ カリフォルニア州 サンフランシスコ １５ス アベニュー ２５７９
(72)発明者 テイラー、 ジェームズ
アメリカ合衆国 ９４１１４ カリフォルニア州 サンフランシスコ ２５ス ストリート ４３８５
(72)発明者 フェル、 サイモン ズィー.
アメリカ合衆国 ９４９２５ カリフォルニア州 コーテ マデラ フライング クラウド コース １４
(72)発明者 シュナイダー、 ドノヴァン エー.
アメリカ合衆国 ９４１２７ カリフォルニア州 サンフランシスコ アプトス アベニュー ２５

審査官 原 秀人

(56)参考文献 特開２００１－０５１８７９（ＪＰ，Ａ）
特開２０１０－２２４８２４（ＪＰ，Ａ）
特開２０００－２２２４３０（ＪＰ，Ａ）
特開平０９－１４６８０４（ＪＰ，Ａ）
特開平１１－０９６０５５（ＪＰ，Ａ）
米国特許出願公開第２０１１／０２５８６３０（ＵＳ，Ａ１）
中田 敦，最新のクラウド技術を解剖する設計者が明かす基盤の実像，日経コンピュータ，日本，日経ＢＰ社，２０１０年 ９月 １日，第７６４号，ｐ．６８－７３
萩原 正義，クラウドの設計セオリー，日経ＳＹＳＴＥＭＳ，日本，日経ＢＰ社，２０１０年１１月２６日，第２１２号，ｐ．９９－１０３

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ １２／００
Ｇ０６Ｆ １７／３０