

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2017-512338

(P2017-512338A)

(43) 公表日 平成29年5月18日(2017.5.18)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 12/00 (2006.01)	G06F 12/00	520E
G06F 17/30 (2006.01)	G06F 12/00	514M
	G06F 12/00	545A
	G06F 17/30	220B

審査請求 未請求 予備審査請求 未請求 (全 19 頁)

(21) 出願番号 特願2016-552932 (P2016-552932)
 (86) (22) 出願日 平成27年2月18日 (2015. 2. 18)
 (85) 翻訳文提出日 平成28年10月13日 (2016. 10. 13)
 (86) 国際出願番号 PCT/US2015/016409
 (87) 国際公開番号 W02015/126961
 (87) 国際公開日 平成27年8月27日 (2015. 8. 27)
 (31) 優先権主張番号 14/518, 913
 (32) 優先日 平成26年10月20日 (2014. 10. 20)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 61/941, 986
 (32) 優先日 平成26年2月19日 (2014. 2. 19)
 (33) 優先権主張国 米国 (US)

(71) 出願人 516245999
 スノーフレイク コンピューティング インク.
 SNOWFLAKE COMPUTING INC.
 アメリカ合衆国, カリフォルニア州 94401, サン マテオ, サウス エルズワース アベニュー 101, 350号室
 101 South Ellsworth Avenue, #350, San Mateo, California 94401, United States of America

(74) 代理人 100074099
 弁理士 大菅 義之

最終頁に続く

(54) 【発明の名称】 第一クラスデータベース要素としての半構造データの実装

(57) 【要約】

半構造データシステムに対するデータストレージおよびデータアクセスを管理するためのシステム、装置および方法。

【選択図】 図2

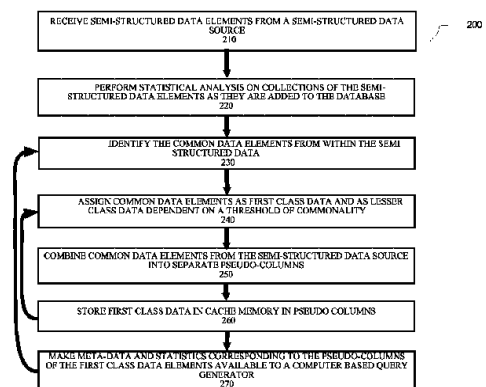


FIG. 2

【特許請求の範囲】**【請求項 1】**

半構造データを格納するための方法であって、
データソースから半構造データ要素を受信することと、
前記データベースに追加されると、前記半構造データ要素の集合に対して統計解析を実施することと、

前記半構造データ内から共通データ要素を同定することと、
共通性の閾値によって、ファーストクラスデータおよびより低いクラスのデータとして前記共通データ要素を割り当てることと、

前記データソースから個別の擬似列に前記共通データ要素をまとめることと、

擬似列内のキャッシュメモリ内にファーストクラスデータを格納し、コンピュータベースのクエリジェネレータに対して利用可能な前記ファーストクラスデータ要素の前記擬似列に対応するメタデータおよび統計を生成することと、
を含む、

方法。

【請求項 2】

前記半構造データ内から共通データ要素を再度同定することと、ファーストクラスデータとして追加の共通データ要素を割り当てることと、キャッシュメモリ内に前記追加データ要素を保存することと、をさらに含む

請求項 1 に記載の方法。

【請求項 3】

共通性の前記閾値未満に低下したファーストクラスデータ要素を同定することと、前記同定されたデータ要素により低いクラスを割り当てることと、キャッシュメモリから除去することをさらに含む、

請求項 2 に記載の方法。

【請求項 4】

共通性の前記閾値は、前記半構造データに、如何なる頻度で前記データ要素が生じるかに関する計数である、

請求項 1 に記載の方法。

【請求項 5】

共通性の前記閾値は、ユーザによって如何なる頻度でデータ要素が要求されるかである、

請求項 1 に記載の方法。

【請求項 6】

ディスクストレージ上の擬似列内に、より低いクラスのデータを格納することをさらに含む、

請求項 1 に記載の方法。

【請求項 7】

オーバーフローシリアライズ列内に、非共通半構造データ要素を格納することをさらに含む、

請求項 2 に記載の方法。

【請求項 8】

前記ファーストクラスデータ要素および前記より低いクラスのデータ要素および前記非共通データをまとめることによって、半構造データを元の形式に再構成することをさらに含む、

請求項 7 に記載の方法。

【請求項 9】

現在の擬似列構造および内容を表すアップデートで、集約されたメタデータを維持することをさらに含む、

請求項 7 に記載の方法。

10

20

30

40

50

【請求項 10】

メインメモリ内に、より低いクラスのデータ要素を格納することをさらに含む、請求項 1 に記載の方法。

【請求項 11】

コンピュータプロセッサと、キャッシュメモリと、ディスクストレージと、コンピュータ命令とを含む半構造データを集約するためのシステムであって、

データソースから半構造データ要素を受信するように構成された受信モジュールと、前記データベースに集合が加えられると、引き出された前記半構造データ要素の前記集合に対応する統計解析データを導出するように構成された統計解析モジュールと、

前記半構造データ内から共通データ要素を同定するように構成され、共通性の閾値によって、ファーストクラスデータおよびより低いクラスデータとして共通データ要素を割り当てるように構成された同定モジュールと、

前記データソースから個別の擬似列に共通データ要素をまとめるように構成された集約モジュールと、

をさらに含む、

前記集約モジュールは、擬似列内のキャッシュメモリ内にファーストクラスデータを格納するように構成され、コンピュータベースのクエリジェネレータに対して利用可能な前記ファーストクラスデータ要素の前記擬似列のメタデータおよび統計を生成するように構成される、

システム。

【請求項 12】

前記同定モジュールは、前記半構造データ内から共通データ要素を再度同定するようにさらに構成され、ファーストクラスデータとして追加の共通データ要素を割り当て、キャッシュメモリ内に前記追加データ要素を保存するようにさらに構成される、請求項 11 に記載のシステム。

【請求項 13】

前記同定モジュールは、共通性の前記閾値未満に低下したファーストクラスデータ要素を同定し、前記同定されたデータ要素に対してより低いクラスを割り当て、キャッシュメモリから除去するようにさらに構成される、

請求項 12 に記載のシステム。

【請求項 14】

共通性の前記閾値は、前記半構造データ内に前記データ要素が如何なる頻度で生じるかに関する計数である、

請求項 11 に記載のシステム。

【請求項 15】

共通性の前記閾値は、データ要素が如何なる頻度でユーザによって要求されるかである、

請求項 11 に記載のシステム。

【請求項 16】

より低いクラスのデータはディスクストレージ上の擬似列内に格納される、

請求項 11 に記載のシステム。

【請求項 17】

非共通半構造データ要素は、オーバーフローシリアライズ列内に格納される、

請求項 11 に記載のシステム。

【請求項 18】

半構造データは、ファーストクラスデータ要素およびより低いクラスデータを再度まとめた元の形式に再構成される、

請求項 17 に記載のシステム。

【請求項 19】

半構造データを集約するための装置であって、

10

20

30

40

50

データソースから半構造データ要素を受信するように構成された受信モジュールと、前記データベースに集合が加えられると、引き出された前記半構造データ要素の前記集合に対応する統計解析データを導出するように構成された統計モジュールと、前記半構造データ内から共通データ要素を同定し、共通性の閾値によって、ファーストクラスデータおよびより低いクラスデータとして共通データ要素を割り当てるための集約手段であって、前記集約手段は、前記データソースから個別の擬似列に共通データ要素をさらにまとめるためのものであり、前記集約手段は、擬似列内のキャッシュメモリ内にファーストクラスデータを更に格納するためのものである、集約手段と、を含む、装置。

10

【請求項 20】

前記集約手段は、前記半構造データ内から共通データ要素をさらに再度同定し、ファーストクラスデータとして追加の共通データ要素をさらに割り当て、キャッシュメモリ内に前記追加データ要素を保存する、請求項 19 に記載の装置。

【請求項 21】

前記集約手段は、共通性の前記閾値未満に低下したファーストクラスデータ要素をさらに同定し、前記同定されたデータ要素により低いクラスをさらに割り当て、キャッシュメモリから除去する、請求項 20 に記載の装置。

20

【請求項 22】

前記集約手段は、現在の擬似列構造および内容を表すアップデートで実質的に継続して、集約されたメタデータをさらに維持する、請求項 19 に記載の装置。

【発明の詳細な説明】**【技術分野】****【0001】****[関連出願に対する参照]**

本出願は、2014年2月19日に提出された“Apparatus and method for enterprise data warehouse data processing on cloud infrastructure”と題された米国仮出願整理番号61/941,986の利益を享受する権利を主張し、米国仮出願整理番号61/941,986の開示は、その全体において参照によって本明細書に組み入れられる。

30

【0002】

本開示は、データストレージおよびコンピューティングリソースを管理するリソース管理システムおよび方法に関する。

【背景技術】**【0003】**

今日、多くの既存のデータ格納・検索(retrieval、取得)システムが利用可能である。例えば、共有されたディスクシステムにおいては、データクラスタ内の全ての処理ノードからアクセス可能な共有ストレージデバイス上に全てのデータが格納される。この種類のシステムにおいては、全てのデータ変更は、データクラスタ内の全ての処理ノードが整合性のあるバージョンのデータにアクセスすることを保証するために共有ストレージデバイスに書き込まれる。多数の処理ノードが共有されたディスクシステム内で増加すると、共有ストレージデバイス(および処理ノードと共有ストレージデバイスとの間の通信リンク)は、データ読み出しおよびデータ書き込み動作を遅延させるボトルネックとなる。このボトルネックは、より多くの処理ノードの追加によってさらに悪化する。このように、既存の共有されたディスクシステムは、このボトルネック問題によって制限されたスケー

40

50

ラビリティを有する。

【0004】

別の既存のデータストレージおよび検索システムは、“シェアード・ナッシング・アーキテクチャ (shared-nothing architecture)” と称される。このアーキテクチャにおいては、データは、複数の処理ノードにわたって分散し、各ノードは、全データベース内のデータのサブセットを格納する。新しい処理ノードが追加、または除去されると、シェアード・ナッシング・アーキテクチャは、複数の処理ノードにわたってデータを再配列しなければならない。このデータの再配列は時間を消費し、データ再配列中に実行されるデータ読み出しおよび書き込み動作に対して混乱を生じさせ得る。また、特定のノードに対するデータのアフィニティは、一般的なデータに対するデータクラスタ上に“ホットスポット (hot spots)” を生成し得る。さらに、各処理ノードは、ストレージ機能も実施するため、このアーキテクチャは、データを格納するために少なくとも一つの処理ノードを必要とする。このように、シェアード・ナッシング・アーキテクチャは、全ての処理ノードが除去される場合にはデータを格納するのに失敗する。さらに、シェアード・ナッシング・アーキテクチャにおけるデータの管理は、多くの異なる処理ノードにわたるデータの分散によって複雑である。

10

【0005】

本明細書に記述されるシステムおよび方法は、既存のシステムの上記で特定された制限を軽減するデータストレージおよびデータ検索に対する改良されたアプローチを提供する。

20

【図面の簡単な説明】

【0006】

本開示の非限定かつ非包括的な実施形態は、以下の図面を参照して記述され、類似の参照番号は、そうでないと特に断らない限りは、種々の図面を通して類似の部分のことを称する。

【図1】半構造データの処理を示す情報フローおよび関係図を示す。

【図2】本明細書に開示される方法の実装を示すプロセスフロー図である。

【図3】本開示の教示によるオペレーティング環境の一実施形態を示すブロック図を示す。

【図4】本開示の教示によるリソースマネージャの実装の一例を示すブロック図を示す。

30

【図5】本開示の教示による実行プラットフォームの実装の一例を示すブロック図を示す。

【図6】本開示の教示による例示的コンピューティングデバイスを示すブロック図を示す。

【発明を実施するための形態】

【0007】

半構造データを管理するための方法、装置およびシステムが本明細書に開示される。例えば、半構造データを管理するための方法の実装は、半構造データソースから半構造データ要素を受信し、データベースに半構造データ要素が加えられると、半構造データ要素の集合に対して統計解析を実施してもよい。さらに、半構造データ内から共通のデータ要素が同定され、共通性の閾値によって、ファーストクラスデータおよび/またはより低いデータとして割り当てられてもよい。実装は、キャッシュメモリ内に格納された個別の擬似列に、半構造データソースからの共通データ要素を更に組み合わせ、その後、データベースのクエリジェネレータに対して利用可能なファーストクラスデータの擬似列に対応するメタデータおよび統計を作成してもよい。

40

【0008】

以下の記述においては、本明細書の一部を形成する添付の図面に対する参照が行われ、その中では、本開示が実践され得る特定の例示の実施形態の例示として示される。これらの実施形態は、本明細書に開示された概念を当業者が実施することを可能にするために十分詳細に記述され、本開示の範囲から逸脱することなく、種々の開示された実施形態に対

50

する改変が行われてもよく、他の実施形態が使用されてもよいことを理解されたい。したがって、以下の詳細な記述は、限定する意味で解釈されるべきではない。

【0009】

本明細書中の“一実施形態 (one embodiment)”、“ある実施形態 (an embodiment)”、“一例 (one example)”または“ある例 (an example)”に対する言及は、実施形態または例に関連して記述される特定の特徴、構造または特性が本開示の少なくとも一実施形態に含まれることを意味する。このように、本明細書中の種々の位置における“一実施形態においては (in one embodiment)”“ある実施形態においては (in an embodiment)”、“一例”または“ある例”という句の出現は、必ずしもすべて同一の実施形態または同一の例に対して言及するものではない。さらに、本明細書に提供される図面は、当業者に対して説明する目的のためのものであって、図面は必ずしも同一の縮尺で描かれているとは限らないことを理解されたい。

10

【0010】

本開示による実施形態は、装置、方法またはコンピュータプログラム製品として具現化されてもよい。したがって、本開示は、全体としてハードウェア構成実施形態、全体としてソフトウェア構成実施形態 (ファームウェア、レジデントソフトウェア、マイクロコードなどを含む) または“回路”“モジュール”もしくは“システム”として本明細書で全て一般的に称され得るソフトウェアおよびハードウェア態様を組み合わせる実施形態の形式をとってもよい。さらに、本開示の実施形態は、媒体中で具現化されるコンピュータ利用可能なプログラムコードを有する表現の任意の有形媒体内で具現化されるコンピュータプログラム製品の形式をとってもよい。

20

【0011】

一つ以上のコンピュータ利用可能またはコンピュータ可読媒体の如何なる組み合わせが使用されてもよい。例えば、コンピュータ可読媒体は、ポータブルコンピュータディスクレット、ハードディスク、ランダムアクセスメモリ (RAM) デバイス、リードオンリーメモリ (ROM) デバイス、消去可能なプログラマブルリードオンリーメモリ (EPROM またはフラッシュメモリ) デバイス、ポータブルコンパクトディスクリードオンリーメモリ (CDROM)、光学ストレージデバイス、磁気ストレージデバイスのうちの一つ以上を含んでもよい。本開示の動作を実行するためのコンピュータプログラムコードは、一つ以上のプログラミング言語の任意の組み合わせで書き込まれてもよい。このようなコードは、コンピュータ可読アセンブリ言語に対するソースコードまたはコードが実行されるであろうデバイスもしくはコンピュータに対して適切なマシンコードからコンパイルされてもよい。

30

【0012】

実施形態は、クラウドコンピューティング環境内でも実装されてもよい。この記述および以下の請求項においては、“クラウドコンピューティング (cloud computing)”とは、仮想化によって迅速にセットアップでき、最小の管理努力またはサービスプロバイダ相互作用でリリースすることができ、その後それによってスケーリングすることができる、構成可能なコンピューティングリソース (例えば、ネットワーク、サーバ、ストレージ、アプリケーションおよびサービス) の共有されたプールに対して、ユビキタスで簡便なオンデマンドネットワークアクセスを可能にするためのモデルとして定義されてもよい。クラウドモデルは、種々の特性 (例えば、オンデマンドセルフサービス、ブロードネットワークアクセス、リソースプーリング、迅速な柔軟性 (rapid elasticity) およびメジャーサービス)、サービスモデル (例えば、ソフトウェア・アズ・ア・サービス (“SaaS”)、プラットフォーム・アズ・ア・サービス (“PaaS”) およびインフラストラクチャ・アズ・ア・サービス (“IaaS”)) および展開モデル (例えば、プライベートクラウド、コミュニティクラウド、パブリッククラウドおよびハイブリッドクラウド) で構成することができる。

40

【0013】

50

添付の図面内のフロー図およびブロック図は、本開示の種々の実施形態によるシステム、方法およびコンピュータプログラム製品の可能性のある実装のアーキテクチャ、機能および動作を図示するものである。この点においては、フロー図またはブロック図内の各ブロックは、モジュール、セグメントまたはコードの一部を表してもよく、それは、特定の（複数の）論理関数を実装するための一つ以上の実行可能な命令を含む。ブロック図および/またはフロー図の各ブロック、ブロック図および/またはフロー図内のブロックの組み合わせは、特定の機能もしくは動作、または専用ハードウェアおよびコンピュータ命令の組み合わせを実施する専用ハードウェアベースシステムによって実装されてもよいことにも留意されたい。コンピュータ可読媒体内に格納された命令が、フロー図および/またはブロック図の一つ以上のブロック内で特定される機能/動作を実装する命令手段を含む製造の態様を作成するように、これらのコンピュータプログラム命令は、コンピュータまたは他のプログラマブルデータ処理装置を特定の方法で機能させることができるコンピュータ可読媒体内にも格納されてもよい。

10

20

30

40

50

【0014】

本明細書に記述されたシステムおよび方法は、新しいデータ処理プラットフォームを使用するフレキシブルかつスケラブルなデータウェアハウスを提供する。幾つかの実施形態においては、記述されたシステムおよび方法は、クラウドベースストレージリソース、コンピューティングリソースなどをサポートするクラウドインフラストラクチャを利用する。例示的なクラウドベースのストレージリソースは、低コストで、オンデマンドで利用可能なかなりのストレージ容量を提供する。さらに、これらのクラウドベースストレージリソースは、フォールト・トレラントかつ高いスケラビリティを有してもよく、個人的なデータストレージシステム内で実現するためには高価であり得る。例示的なクラウドベースのコンピューティングリソースは、オンデマンドで利用可能であり、リソースの実際の利用レベルに基づいて値段設定されてもよい。典型的には、クラウドインフラストラクチャは、迅速な方法で動的に展開され、再構成され、閉鎖される。

【0015】

記述されたシステムおよび方法においては、データストレージシステムは、半構造ベースのリレーショナル（関係）データベースを利用する。しかしながら、これらのシステムおよび方法は、任意のデータストレージアーキテクチャを利用し、データベース内のデータを格納して検索するために任意の言語を利用する任意の種類のデータベースに対して適用可能である。本明細書で用いられるように、半構造データとは、リレーショナルデータベースに関連づけられるデータモデルの典型的な形式的構造とは一致しないが、それにもかかわらず、意味的要素を分離し、データ内のレコードおよびフィールドの階層を強制するためのタグまたは他のマーカを含む構造データの形式を伝達することを意味する。本明細書に記述されるシステムおよび方法は、異なるカスタマー/クライアント間、および同一のカスタマー/クライアント内の異なるユーザ間でコンピューティングリソースおよびデータの分離をサポートするマルチテナントシステムをさらに提供する。

【0016】

大量の半構造データを扱うデータベースおよびデータウェアハウスシステムの性能を顕著に改善する方法およびシステムが本明細書に開示される。既存のデータベースシステムは、リレーショナル（即ち、SQLデータベース）またはキー値ストアのいずれかである。

【0017】

リレーショナルデータベースは、クエリデータアクセスブルーニング（テーブルの特定の列に格納された値について集約されたメタデータに基づいて、検索からデータベースの一部を排除すること）によって効率的なクエリを実施することができる。しかしながら、これは、半構造データを表すために使用することができないデータの厳格な表フォーマットを必要とする。

【0018】

一方、キー値ストアは、よりフレキシブルであるが、ブルーニングの欠如によって深刻

な性能のペナルティを導入する。既存の製品および研究プロジェクトにおけるリレーショナルデータベースに対して半構造データの取り扱いを追加するための多数の方法が存在する。

【0019】

1. シリアライズエンコード。半構造データ記録は、シリアライズ表現として列内に格納される。幾つかのフィールドの値が利用されると常に、その値が抽出されて、基本タイプに変換される。この方法はフレキシブルであるが、このブルーニングによって改善されるべきデータに対するアクセスを行う。シリアライズ表現からの抽出はコストが高く、通常のリレーショナルデータで動作するよりも顕著に多くのCPU時間を必要とする。全てのシリアライズデータの記録は、それらのうちのごくわずかな部分のみ（単一要素など）がクエリ内で使用される場合でさえも、持続性ストレージから読み出され、処理されなければならない。

10

【0020】

2. 取り込みにおける変換。半構造データは、取り込みにおいて、リレーショナルデータに変換される。これは、任意の他のリレーショナルデータに対するアクセスと同程度に迅速に、このデータにアクセスするが、取り込みにおいてデータ構造の厳格な仕様を必要とし、対応するデータベーススキーマが事前に完全に特定されることを必要とする。データベーススキーマを変更する必要があるため、この方法は、変化する構造を有するデータの取り扱いのコストを非常に高くする。ある記録から記録へと変化する構造を有するデータは、この方法を利用して扱うことが不可能である。変換方法は、事前に特定されなければならない、あらゆる重要な変化は、元の半構造データを再度取り込むことを必要とする。

20

【0021】

3. 従来のリレーショナルデータベースに格納されたオブジェクト属性値三つ揃え表現に等価な構造データのリレーショナル様表現。この方法はフレキシブルであるが、実用上、データサブコンポーネントに対するアクセス用のジョイン演算を必要とし、それは、データに依存して非常に遅く成り得る。

【0022】

4. ソースデータの階層反映構造内で異なる濃度を有する列がリンクすることを可能とする、リレーショナルデータモデルに対する非伝統的拡張。このようなデータ表現のためのクエリ生成方法は、十分理解されていない（効率的なクエリ生成は、本技術分野の現在の状態で不可能である）。この方法は、厳格な（とはいっても表ではない）スキーマに一致する入力データを必要とし、このため、任意の半構造データを扱うほど十分フレキシブルではない。

30

【0023】

必要とされるのは、効率的で低コストかつ反応の速い半構造データと動作するためのシステムおよび方法である。なぜなら、それは、より伝統的なデータ構造と同様に処理してクエリすることができる、少なくとも擬似的な列内のデータを管理しながら、半構造データの意味を保存するからである。

【0024】

以下の開示の実装においては、データは、ファイル、ファイルの要素、ファイルの一部などの形式で入力されてもよい。ファイルは、ドキュメントの集合を含んでもよく、データの一部は、ファイル、接続由来の複数のドキュメントおよび/またはドキュメントの一部を含んでもよい。さらに、実装においては、メタデータは、ファイル、ファイルの一部、データの一部に関連付けられてもよい。

40

【0025】

本明細書で用いられるように、“共通データ要素”とは、同一グループに属するデータ要素および論理的に類似の要素の集合を意味することを意図される。

【0026】

図1は、半構造データを処理し、クエリに対するアクセスを増加させるためにデータ要素に対するクラス指定を割り当てるコンピュータシステムにおける半構造データフローの

50

概略を図示する。図面からわかるように、半構造データ 110 は、その中に、共通および非共通データ要素を含んでもよい。半構造データから共通データを抽出するために、解析統計 122 は、半構造データに対して実行されてもよく、共通のデータ要素および非共通データ要素を決定し、各データ要素に対して割り当てられるべきクラス指定を生成する。さらに、ユーザの関心およびユーザの要求 124 は、共通データ要素のためのクラス指定を決定するために使用されてもよいことに留意されたい。図面に示されるように、共通データは、擬似列 120 と呼ばれる一時的な列構造に格納されてもよく、クラス指定に従ってさらに格納されてもよく、ファーストクラスデータ要素は、クエリのためのより良好なアクセスのために格納される。共通であることを判定されなかったデータ要素は、“オーバーフロー”シリアルイズデータ 140 内にシリアルに格納されてもよい。最終的に、ユーザは、ファーストクラスデータ要素からより迅速かつ正確に折丁に対する結果 130 を受信するだろう。

10

【0027】

実装においては、必要とされるデータ要素が擬似列 120 内にない場合、それは“オーバーフロー”シリアルイズデータ 140 から抽出されてもよく、全体の半構造データ記録 150 が再構成されることを要求される場合、擬似列 120 内のファーストクラスデータ、擬似列 120 内のより低いデータおよび“オーバーフロー”データ 140 から再構成され、元の形式へと再度シリアルイズされてもよい。

【0028】

図 2 は、半構造データを取り扱い、データストレージの優先順位を決定するためのデータクラスを割り当てるための方法 200 のフロー図を図示する。図面からわかるように、210 において、半構造データ要素は、半構造データソースまたは複数の半構造データソースから受信される。データ要素が受信されると、それらは共通および非共通として割り当てられてもよい。これは、220 においてデータベースに加えられるときに半構造データの集合の統計解析を実施することによって達成されてもよい。統計解析に基づいて、230 において共通データ要素が同定されてもよい。240 において、共通データ要素は、所望のまたは繰り返し生じるデータに対するファーストクラス、それほど望ましくなく、かつそれほど生じないデータ要素に対するより低いクラス指定など、クラス指定を割り当てられてもよい。方法 200 の命令を実施するシステムは、240 において、共通性の閾値に従って、個別の擬似列における半構造データからファーストクラスデータ要素のストレージをさらにまとめてもよいが、非共通データ要素が 250 において“オーバーフロー”シリアルイズストレージ内に格納される。共通性の閾値は、例えば、半構造データ内にデータ要素が何回現れるかに基づいてもよく、または、ユーザのクエリがデータを何回検索するかに基づいてもよい。これらの要素は、半構造データから抽出されてもよく、ユーザに対して見えない列のフォーマットで個別に格納されてもよい。260 において、ファーストクラスデータは、その望ましさによって、素早くアクセス可能なキャッシュメモリ内の擬似列に格納されてもよく、これらの擬似列のメタデータおよび統計（最小値および最大値、個別値の数など）は、その後、270 においてクエリジェネレータに対して利用可能にされてもよいことを理解されたい。個別の集合（即ち、個別のファイル内に格納されたテーブルの一部）は、抽出されたデータ要素の異なるサブセットを有してもよいことに留意されたい。

20

30

40

【0029】

図 2 からわかるように、共通性の指定およびデータ要素に対するクラスの割り当ては、半構造データ要素のデータベースがクエリに対して最新のものに維持されることを保証するために繰り返されてもよい。

【0030】

実装においては、データを抽出するとき、共通データ要素の値が必要とされる場合、効率的な列アクセスを利用して、対応する擬似列から直接取得されてもよい。

【0031】

実装においては、共通データ要素は、何時でもファーストクラス指定に追加されてもよ

50

いし、同様に、データ要素は、それほど望ましくない、またはそれほど表現していないことが分かったときに、ファーストクラス指定を失うことがある。

【0032】

ユーザにとって、この方法は、シリアライズされた記録を格納することと区別がつかないことがあり、個々のデータ記録の構造に対して何の制約も課さない。しかしながら、大部分の共通データ要素が、従来のリレーショナルデータと同一の方法で格納されるため、それらに対するアクセスが提供されてもよく、全体の半構造記録の読み出しおよび抽出を必要としない。それによって、従来のリレーショナルデータベースの速度の利点を得る。

【0033】

(同一のテーブル由来の)半構造記録の異なる集合は、抽出されたデータ要素の異なる組を有し得るため、クエリジェネレータおよびブルーニングは、部分的に利用可能なメタデータで動作することが可能であるべきである(即ち、テーブルの一部は、特定のデータ要素に対して利用可能なメタデータおよび統計を有してもよいが、他の部分はそれを欠いていてもよい)。

10

【0034】

従来技術に対する利点は、(それほど共通ではない要素のシリアライズストレージおよび共通要素の列ストレージの双方として)ハイブリッドデータストレージ表現を利用するための方法によって提供される性能である。これは、従来のリレーショナルデータベースによって提供されるデータクエリのシリアライズ表現と高性能とを利用して、システムの任意の半構造データを格納するための柔軟性および性能の双方をユーザが達成することを可能とする。

20

【0035】

任意のデータ処理プラットフォームが半構造データを扱うためにこのアプローチを利用することができるのが当業者には理解されるだろう。それは、SQLを実行するDBMSシステムに限定される必要がない。

【0036】

図3に図示されるのは、本明細書に開示された方法を実行するためのコンピュータシステムである。図3に図示されるように、リソースマネージャ302は、複数のユーザ304、306および308に結合される。特定の実装においては、リソースマネージャ302は、データ処理プラットフォーム300にアクセスを望むあらゆる数のユーザをサポートすることができる。ユーザ304-308は、例えば、データストレージおよび検索要求を提供するエンドユーザと、本明細書に記述されたシステムおよび方法、データベースと相互作用するソフトウェアアプリケーションおよびリソースマネージャ302と相互作用する他のコンポーネント/デバイスを管理するシステム管理者と、を含んでもよい。リソースマネージャ302は、データ処理プラットフォーム300内の全システムおよびコンポーネントの動作をサポートする種々のサービスおよび機能を提供する。リソースマネージャ302は、メタデータ310にも結合され、メタデータ310は、データ処理プラットフォーム300中に格納されたデータ全体と関連付けられる。リソースマネージャは、ファイルの組に対応するメタデータと結合されるため、メタデータは、ユーザクエリを生成するために使用されてもよい。幾つかの実施形態においては、メタデータ310は、ローカルキャッシュから利用可能なデータと同様に、遠隔データストレージシステム内に格納されたデータの要約を含む。さらに、メタデータ310は、遠隔データストレージシステムおよびローカルキャッシュ内で如何にしてデータが組織化されるかに関する情報を含んでもよい。メタデータ310は、ストレージデバイスから実際のデータをロードまたはアクセスすることなく、一片のデータが処理される必要があるか否かをシステムおよびサービスが判定することを可能にする。

30

40

【0037】

リソースマネージャ302は、実行プラットフォーム312にさらに結合され、実行プラットフォーム312は、以下により詳細に議論されるように、種々のデータストレージおよびデータ検索タスクを実行する複数のコンピューティングリソースを提供する。実行

50

プラットフォーム 312 は、ストレージプラットフォーム 314 の一部である複数のデータストレージデバイス 316、318 および 320 に結合される。3つのデータストレージデバイス 316、318 および 320 が図 3 に図示されるが、実行プラットフォーム 312 は、あらゆる数のデータストレージデバイスと通信することが可能である。幾つかの実施形態においては、データストレージデバイス 316、318 および 320 は、一つ以上の地理的位置に配置されたクラウドベースのストレージデバイスである。例えば、データストレージデバイス 316、318 および 320 は、パブリッククラウドインフラストラクチャの一部またはプライベートクラウドインフラストラクチャであってもよい。データストレージデバイス 316、318 および 320 は、ハードディスクドライブ (HDD)、ソリッドステートドライブ (SSD)、ストレージクラスまたは任意の他のデータストレージ技術であってもよい。さらに、ストレージプラットフォーム 314 は、(Hadoop 分散ファイルシステム (HDFS) などの) 分散ファイルシステム、オブジェクトストレージシステムなどを含んでもよい。

10

20

30

40

50

【0038】

特定の実施形態においては、リソースマネージャ 302 とユーザ 304 - 308、メタデータ 310 および実行プラットフォーム 312 との間の通信リンクは、一つ以上のデータ通信ネットワークを介して実装される。同様に、実行プラットフォーム 312 とストレージプラットフォーム 314 内のデータストレージデバイス 316 - 320 との間の通信リンクは、一つ以上のデータ通信ネットワークを介して実装される。これらのデータ通信ネットワークは、任意の通信プロトコルおよび任意の種類通信媒体を使用してもよい。幾つかの実施形態においては、データ通信ネットワークは、互いに結合された二つ以上のデータ通信ネットワーク (または、サブネットワーク) の組み合わせである。別の実施形態においては、これらの通信リンクは、任意の種類通信媒体および任意の通信プロトコルを利用して実装される。

【0039】

図 3 に図示されるように、データストレージデバイス 316、318 および 320 は、実行プラットフォーム 312 に関連付けられたコンピューティングリソースから切り離される。このアーキテクチャは、データ処理プラットフォーム 300 にアクセスするユーザおよびシステムのニーズの変化と同様に、データストレージ/検索ニーズの変化に基づいて、データ処理プラットフォーム 300 に対する動的な変化をサポートする。動的変化のサポートによって、データ処理プラットフォーム 300 内のシステムおよびコンポーネントに対する需要の変化に迅速に応じて、データ処理プラットフォーム 300 をスケールアップすることを可能にする。データストレージデバイスからのコンピューティングリソースの切断は、対応する大量のコンピューティングリソースを必要とすることなく、大量のデータのストレージをサポートする。同様に、このリソースの切断は、利用可能なデータストレージリソース内で対応する増加を必要とすることなく、特定の時刻に使用されるコンピューティングリソース内の顕著な増加をサポートする。

【0040】

リソースマネージャ 302、メタデータ 310、実行プラットフォーム 312 およびストレージプラットフォーム 314 は、個々のコンポーネントとして図 3 に図示される。しかしながら、リソースマネージャ 302、メタデータ 310、実行プラットフォーム 312 およびストレージプラットフォーム 314 の各々は、分散システム (例えば、複数の地理的位置において、複数のシステム/プラットフォームにわたって分散される) として実装されてもよい。さらに、リソースマネージャ 302、メタデータ 310、実行プラットフォーム 312 およびストレージプラットフォーム 314 の各々は、ユーザ 304 - 308 から受信された要求の変化およびデータ処理プラットフォーム 300 のニーズの変化によって (互いに独立して) 拡張または縮小することができる。このように、記述された実施形態においては、データ処理プラットフォーム 300 は動的であって、現在のデータ処理ニーズに見合うように規則的な変化をサポートする。

【0041】

図4は、リソースマネージャ302の一実施形態を図示するブロック図である。図3に図示されるように、リソースマネージャ302は、データストレージデバイス406に結合されたアクセスマネージャ402とキーマネージャ404とを含む。アクセスマネージャ402は、本明細書に記述されたシステムに対する認証および承認タスクを扱う。キーマネージャ404は、認証および承認タスク中に、使用されるキーのストレージおよび認証を管理する。要求処理サービス408は、受信されたデータストレージ要求およびデータ検索要求を管理する。管理コンソールデバイス410は、管理者および他のシステムマネージャによる種々のシステムおよびプロセスに対するアクセスをサポートする。

【0042】

リソースマネージャ302は、SQLコンパイラ412、SQLオプティマイザ414およびSQLEグゼキュータ410も含む。SQLコンパイラ412は、SQLクエリを解析し、クエリのための実行コードを生成する。SQLオプティマイザ414は、処理される必要があるデータに基づいてクエリを実行するための最良の方法を決定する。SQLEグゼキュータ416は、リソースマネージャ302によって受信されるクエリのためのクエリコードを実行する。クエリスケジューラおよびコーディネータ418は、コンパイル、最適化および実行プラットフォームに対するディスパッチのために適切なサービスまたはシステムに対して、受信されたクエリを送信する。仮想ウェアハウスマネージャ420は、実行プラットフォーム内で実装される複数の仮想ウェアハウスの動作を管理する。

【0043】

さらに、リソースマネージャ302は、構成およびメタデータマネージャ422を含み、これは、遠隔データストレージデバイス内およびローカルキャッシュ内に格納されたデータに関する情報を管理する。監視および作業負荷アナライザ424は、リソースマネージャ302によって実施されるプロセスを監視し、実行プラットフォーム内の仮想ウェアハウスおよび実行ノードにわたるタスク（例えば、作業負荷）の分散を管理する。構成およびメタデータマネージャ422と、監視および作業負荷アナライザ424は、データストレージデバイス426に結合される。

【0044】

リソースマネージャ302は、トランザクション管理およびアクセス制御モジュール428も含み、これは、種々のタスクと、データストレージ要求およびデータアクセス要求の処理に関連付けられる他のアクティビティとを管理する。例えば、トランザクション管理およびアクセス制御モジュール428は、複数のユーザまたはシステムによる、データに対する一貫し、かつ同期したアクセスを提供する。複数のユーザ/システムは、同時に同一のデータにアクセスしてもよいため、データに対する変更は、現在のバージョンのデータを各ユーザ/システムが扱うことを確実にするために同期されなければならない。トランザクション管理およびアクセス制御モジュール428は、リソースマネージャ302内の単一の中央に集中した位置において、種々のデータ処理アクティビティの制御を提供する。

【0045】

図5は、実行プラットフォームの一実施形態を図示するブロック図である。図5に図示されるように、実行プラットフォーム512は、複数の仮想ウェアハウス502、504および506を含む。各仮想ウェアハウスは、各々がキャッシュとプロセッサとを含む複数の実行ノードを含む。図5に図示された各仮想ウェアハウス502 - 506は3つの実行ノードを含むが、特定の仮想ウェアハウスは、如何なる数の実行ノードを含んでもよい。さらに、仮想ウェアハウス内の実行ノードの数は動的であり、さらなる需要が存在するときには新しい実行ノードが生成され、もう必要とされないときには既存の実行ノードが消去される。

【0046】

各仮想ウェアハウス502 - 506は、図3に図示されたデータストレージデバイス316 - 320のうちの任意のデバイスにアクセスすることが可能である。したがって、仮想ウェアハウス502 - 506は、必ずしも特定のデータストレージデバイス316 - 3

10

20

30

40

50

20に割り当てられるわけではなく、その代わりに、データストレージデバイス316 - 320のうちの任意のデータストレージデバイスからデータにアクセスすることができる。同様に、図5に図示される実行ノードの各々は、データストレージデバイス316 - 320のうちの任意のデータストレージデバイスからデータにアクセスすることができる。幾つかの実施形態においては、特定の仮想ウェアハウスまたは特定の実行ノードは、特定のデータストレージデバイスに一時的に割り当てられてもよいが、仮想ウェアハウスまたは実行ノードは、その後、他の如何なるデータストレージデバイスからデータにアクセスしてもよい。

【0047】

図5の例においては、仮想ウェアハウス502は、3つの実行ノード508、510および512を含む。実行ノード508は、キャッシュ514とプロセッサ516とを含む。実行ノード510は、キャッシュ518とプロセッサ520とを含む。実行ノード512は、キャッシュ522とプロセッサ524とを含む。各実行ノード508 - 512は、一つ以上のデータストレージおよび/またはデータ検索タスクの処理に関連付けられる。例えば、特定の仮想ウェアハウスは、特定のユーザまたはカスタマーに関連付けられたデータストレージおよびデータ検索タスクを扱ってもよい。他の実装においては、特定の仮想ウェアハウスは、特定のデータストレージシステムまたは特定のカテゴリのデータに関連付けられたデータストレージおよびデータ検索タスクを扱ってもよい。

【0048】

上記で議論された仮想ウェアハウス502と同様に、仮想ウェアハウス504は、3つの実行ノード526、528および530を含む。実行ノード526は、キャッシュ532とプロセッサ534とを含む。実行ノード528は、キャッシュ536とプロセッサ538とを含む。実行ノード530は、キャッシュ540とプロセッサ542とを含む。さらに、仮想ウェアハウス506は、3つの実行ノード544、546および548を含む。実行ノード544は、キャッシュ550とプロセッサ552とを含む。実行ノード546は、キャッシュ554とプロセッサ556とを含む。実行ノード548は、キャッシュ558とプロセッサ560とを含む。

【0049】

図5に図示された実行ノードは、1つのキャッシュと1つのプロセッサとを各々含むが、別の実施形態は、如何なる数のプロセッサと如何なる数のキャッシュとを含む実行ノードを含んでもよい。さらに、キャッシュは、異なる実行ノードの間でサイズが変化してもよい。図5に図示されたキャッシュは、ストレージプラットフォーム314(図3)内の一つ以上のデータストレージデバイスから検索されたデータをローカル実行ノードに格納する。このように、キャッシュは、遠隔ストレージシステムから常にデータを検索するプラットフォーム内で生じるボトルネック問題を軽減するか、または排除する。遠隔ストレージデバイスからデータに繰り返しアクセスする代わりに、本明細書に記述されたシステムおよび方法は、有意により高速であって、上記で議論されたボトルネック問題を回避する実行ノード内のキャッシュからデータにアクセスする。幾つかの実施形態においては、キャッシュは、キャッシュされたデータに対する素早いアクセスを提供する高速メモリデバイスを利用して実装される。各キャッシュは、ストレージプラットフォーム314内の

【0050】

さらに、キャッシュリソースおよびコンピューティングリソースは、異なる実行ノードの間で変化してもよい。例えば、一つの実行ノードは、多くのコンピューティングリソースと最小のキャッシュリソースとを含んでもよく、多くのコンピューティングリソースを必要とするタスクに対してその実行ノードを有用にする。別の実行ノードは、多くのキャッシュリソースと、最小のコンピューティングリソースとを含んでもよく、大量のデータをキャッシュすることを必要とするタスクに対してこの実行ノードを有用にする。幾つかの実施形態においては、実行ノードによって実施されるべき期待されるタスクに基づいて

10

20

30

40

50

実行ノードが生成されると、特定の実行ノードに関連付けられたキャッシュリソースとコンピューティングリソースとが決定される。

【0051】

さらに、特定の実行ノードに関連付けられたキャッシュリソースとコンピューティングリソースとは、実行ノードによって実施されるタスクの変化に基づいて、時間につれて変化してもよい。例えば、特定の実行ノードは、実行ノードによって実施されるタスクがプロセッサをより集中的に使用するものになる場合に、より多くの処理リソースが割り当てられてもよい。同様に、実行ノードは、実行ノードによって実施されるタスクがより大きいキャッシュ容量を必要とする場合に、より多くのキャッシュリソースが割り当てられてもよい。

10

【0052】

仮想ウェアハウス502 - 506は、図3の同一の実行プラットフォーム312に関連付けられるが、仮想ウェアハウスは、複数の地理的位置における複数のコンピューティングシステムを利用して実装されてもよい。例えば、仮想ウェアハウス502は、第一の地理的位置においてコンピューティングシステムによって実装することができるが、仮想ウェアハウス504および506は、第二の地理的位置において、別のコンピューティングシステムによって実装される。幾つかの実施形態においては、これらの異なるコンピューティングシステムは、一つ以上の異なるエンティティによって維持されるクラウドベースのコンピューティングシステムである。

【0053】

さらに、各仮想ウェアハウスは、複数の実行ノードを有するものとして、図5に図示される。各仮想ウェアハウスに関連付けられる複数の実行ノードは、複数の地理的位置において複数のコンピューティングシステムを利用して実装されてもよい。例えば、仮想ウェアハウス502の具体的な例は、特定の地理的位置における一つのコンピューティングプラットフォーム上で実行ノード508および510を実装し、別の地理的位置において異なるコンピューティングプラットフォームにおける実行ノード512を実装する。実行ノードを実装するために特定のコンピューティングシステムを選択することは、(例えば、処理リソース要件およびキャッシュ要件などの)特定の実行ノードに対して必要とされるリソースのレベル、特定のコンピューティングシステムで使用可能なリソース、地理的位置内または地理的位置間でのネットワークの通信容量、どのコンピューティングシステムが仮想ウェアハウス内の他の実行ノードを既に実装しているか、などの種々の要因に依存し得る。実行プラットフォーム312もまた、フォールト・トレラントである。例えば、一つの仮想ウェアハウスが機能しない場合、その仮想ウェアハウスは、異なる地理的位置において異なる仮想ウェアハウスと素早く置換される。

20

30

【0054】

特定の実行プラットフォーム312は、あらゆる数の仮想ウェアハウス502 - 506を含んでもよい。さらに、追加の処理および/またはキャッシュリソースが必要とされるときには新しい仮想ウェアハウスが生成されるように、特定の実行プラットフォーム内の仮想ウェアハウスの数は動的である。同様に、既存の仮想ウェアハウスは、仮想ウェアハウスに関連付けられたリソースがもう必要ではない場合には消去されてもよい。

40

【0055】

図6は、例示的なコンピューティングデバイス600を図示するブロック図である。幾つかの実施形態においては、コンピューティングデバイス600は、本明細書で議論されるシステムおよびコンポーネントのうちの一つ以上を実装するために使用される。例えば、コンピューティングデバイス600は、リソースマネージャ302にユーザまたは管理者がアクセスすることを可能にしてもよい。さらに、コンピューティングデバイス600は、本明細書で記述されたシステムおよびコンポーネントのうちの任意のシステムおよびコンポーネントと相互作用してもよい。したがって、コンピューティングデバイス600は、本明細書で議論されたような種々の手順およびタスクを実施するために使用されてもよい。コンピューティングデバイス600は、サーバ、クライアントまたは任意の他のコ

50

ンピューティングエンティティとして機能することができる。コンピューティングデバイス600は、デスクトップコンピュータ、ノートブックコンピュータ、サーバコンピュータ、ハンドヘルドコンピュータ、タブレットなどの種々のコンピューティングデバイスの中の任意のコンピューティングデバイスとすることができる。

【0056】

コンピューティングデバイス600は、一つ以上のプロセッサ602、一つ以上のメモリデバイス604、一つ以上のインターフェイス606、一つ以上のマストレージデバイス608、一つ以上の入力/出力(I/O)デバイス610を含み、その全ては、バス612に結合される。(複数の)プロセッサ602は、(複数の)メモリデバイス604および/または(複数の)マストレージデバイス608に格納された命令を実行する一つ以上のプロセッサまたはコントローラを含む。(複数の)プロセッサ602は、キャッシュメモリなどの種々の種類のコンピュータ可読媒体も含んでもよい。

10

【0057】

(複数の)メモリデバイス604は、揮発性メモリ(例えば、ランダムアクセスメモリ(RAM))および/または不揮発性メモリ(例えば、リードオンリーメモリ(ROM))などの種々のコンピュータ可読媒体を含む。(複数の)メモリデバイス604は、フラッシュメモリなどの再書き込み可能なROMも含んでもよい。

【0058】

(複数の)マストレージデバイス608は、磁気テープ、磁気ディスク、光ディスク、ソリッドステートメモリ(例えば、フラッシュメモリ)などの種々のコンピュータ可読媒体を含む。種々のドライブは、種々のコンピュータ可読媒体からの読み出しおよび/または種々のコンピュータ可読媒体への書き込みを可能とするために、(複数の)マストレージデバイス608に含まれてもよい。(複数の)マストレージデバイス608は、リムーバブル媒体および/または非リムーバブル媒体を含む。

20

【0059】

(複数の)I/Oデバイス610は、コンピューティングデバイス600から、データおよび/または他の情報を入力または検索することを可能とする種々のデバイスを含む。例示的な(複数の)I/Oデバイス610は、カーソル制御デバイス、キーボード、キーパッド、マイク、モニタまたは他のディスプレイデバイス、スピーカ、プリンタ、ネットワークインターフェイスカード、モデム、レンズ、CCDまたは他の画像捕捉デバイスなどを含む。

30

【0060】

(複数の)インターフェイス606は、他のシステム、デバイス、またはコンピューティング環境とコンピューティングデバイス600が相互作用することを可能とする種々のインターフェイスを含む。例示的な(複数の)インターフェイス606は、ローカルエリアネットワーク(LAN)、ワイドエリアネットワーク(WAN)、無線ネットワークおよびインターネットに対するインターフェイスなどの、あらゆる数の異なるネットワークインターフェイスを含む。

【0061】

バス612は、バス612に結合された他のデバイスまたはコンポーネントと同様に、(複数の)プロセッサ602、(複数の)メモリデバイス604、(複数の)インターフェイス606、(複数の)マストレージデバイス608および(複数の)I/Oデバイス610が互いに通信することを可能とする。バス612は、システムバス、PCIバス、IEEE1394バス、USBバスなどの幾つかの種類のバス構造のうちの一つ以上を表す。

40

【0062】

例示の目的のために、プログラムおよび他の実行可能なプログラムコンポーネントは、個別のブロックとして本明細書に図示されるが、このようなプログラムおよびコンポーネントは、様々な時点でコンピューティングデバイス600の異なるストレージコンポーネントにおいて存在してもよく、(複数の)プロセッサ602によって実行されることを理

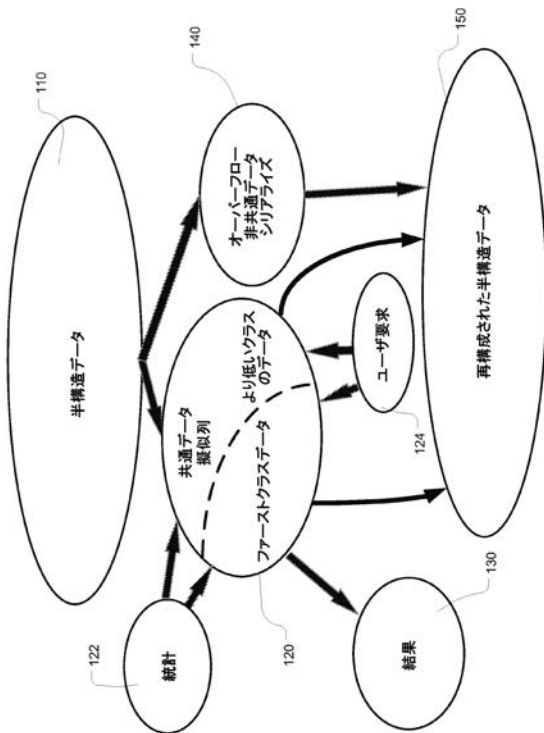
50

解されたい。あるいは、本明細書に記述されたシステムおよび手順は、ハードウェア、またはハードウェアとソフトウェアおよび/またはファームウェアとの組み合わせで実装することができる。例えば、一つ以上の特定用途集積回路（ASIC）は、本明細書に記述されたシステムおよび手順のうちの一つ以上を実行するためにプログラムすることができる。

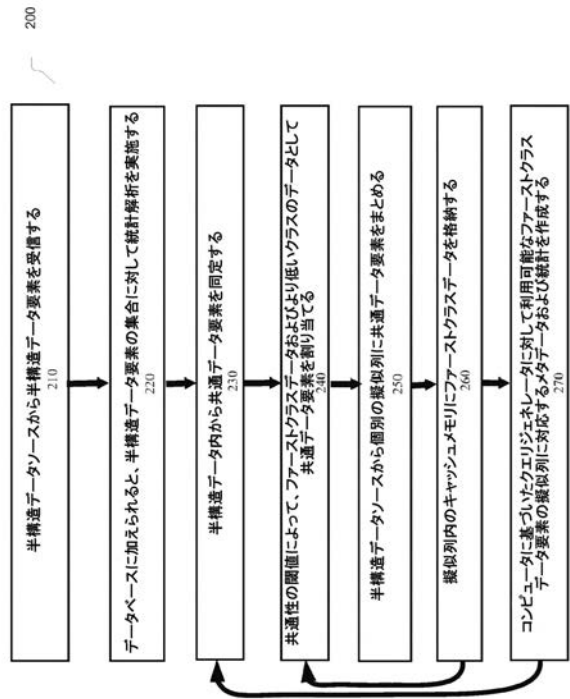
【0063】

本開示は、ある種の好適な実施形態の観点から記述されているが、この開示という恩恵が与えられれば、当業者にとっては、ここに述べた利点と特徴をすべて提供するわけではない実施形態も含めて、他の実施形態も明らかであろうし、それらの他の実施形態もまた、この開示の範囲内である。本開示の範囲から逸脱することなく、他の実施形態が使用されてもよいことを理解されたい。

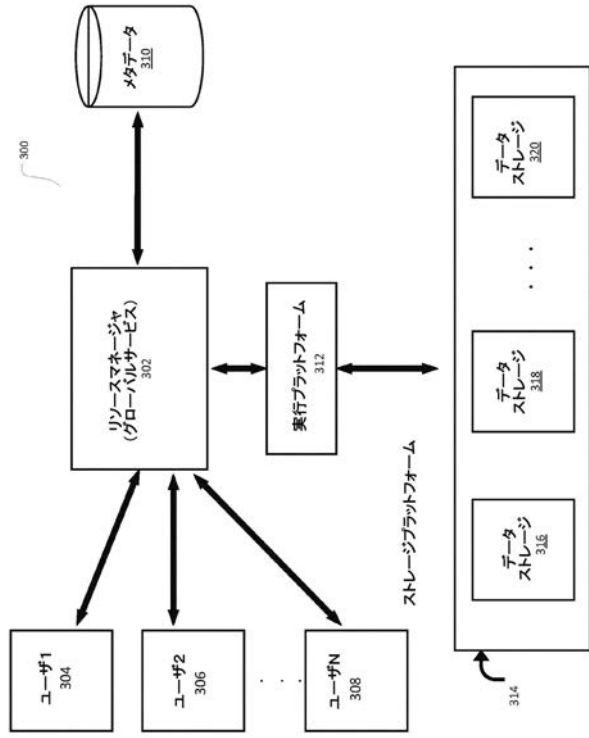
【図1】



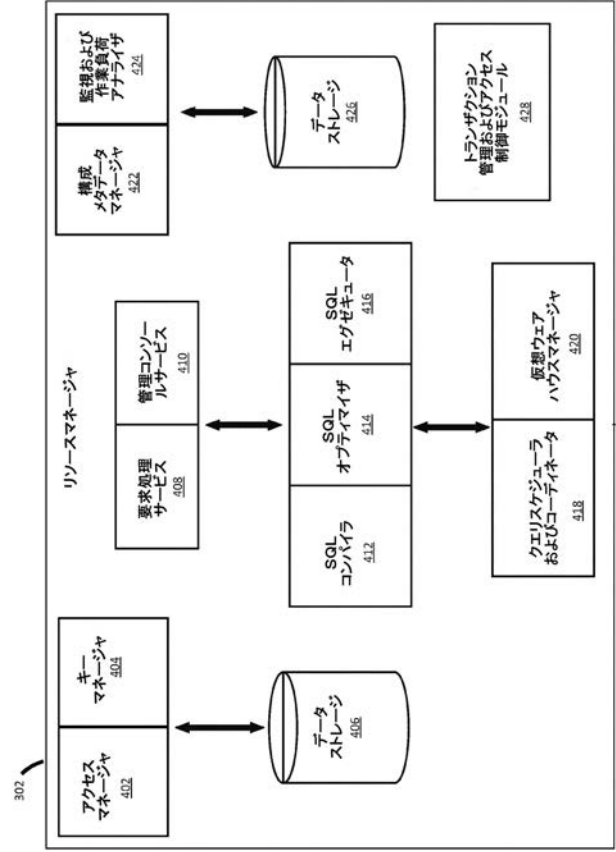
【図2】



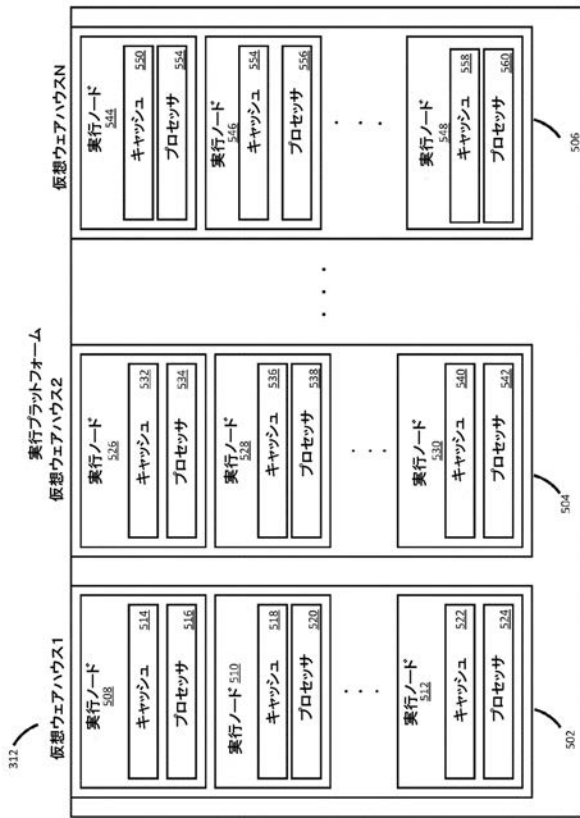
【図 3】



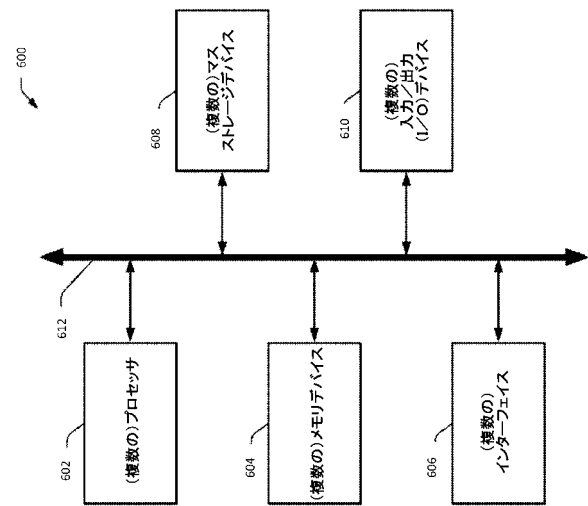
【図 4】



【図 5】



【図 6】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2015/016409
A. CLASSIFICATION OF SUBJECT MATTER IPC(B) - G06F 17/30 (2015.01) CPC - G06F 17/301 (2015.01) According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC(B) - G06F 7/00, 7/06, 9/46, 12/00, 15/16, 17/00, 17/30; G06Q 10/00 (2015.01) USPC - 705/7.11; 707/3, 100, 102, 103, 104.1, 622, 694, 747, 756; 718/101 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched CPC - G06F 17/30, 17/301, 17/30103, 17/30106, 17/30112, 17/3012, 17/30592 (2015.01) (keyword delimited) Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Orbit, Google Patents, Google Scholar. Search terms used: Search terms used: managing data, semi-structured, computer network, data element, common data, overflow column, metadata, statistical analysis, bloom filter, overflow serialized, resource description framework, reconstruct data, pseudo column		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2004/0167904 A1 (WEN et al.) 26 August 2004 (26.08.2004) entire document	1-22
Y	US 2009/0300043 A1 (MACLENNAN) 03 December 2009 (03.12.2009) entire document	1-22
Y	US 2012/0296883 A1 (GANESH et al.) 22 November 2012 (22.11.2012) entire document	1-22
Y	US 2006/0074872 A1 (GORDON) 06 April 2006 (06.04.2006) entire document	3, 10, 13
Y	US 2012/0101860 A1 (EZZAT) 26 April 2012 (26.04.2012) entire document	7-9, 17, 18
Y	US 2013/0124545 A1 (HOLMBERG et al.) 16 May 2013 (16.04.2013) entire document	8, 18
A	US 2008/0027985 A1 (GARRETT et al.) 31 January 2008 (31.01.2008) entire document	1-22
A	US 2002/0120630 A1 (CHRISTIANSON et al.) 29 August 2002 (29.08.2002) entire document	1-22
A	US 2011/0225167 A1 (BHATTACHARJEE et al.) 15 September 2011 (15.09.2011) entire document	1-22
A	US 2012/0323971 A1 (PASUPULETI) 20 December 2012 (20.12.2012) entire document	1-22
A	WO 2006/026659 A2 (ARJUN et al.) 09 March 2006 (09.03.2006) entire document	1-22
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "G" document member of the same patent family		
Date of the actual completion of the international search 27 April 2015		Date of mailing of the international search report 18 MAY 2015
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300		Authorized officer: Blaine R. Copanheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US

(74)代理人 110000132

大菅内外国特許事務所特許業務法人

(72)発明者 ダジュヴィル, ブノワット

アメリカ合衆国, カリフォルニア州 94401, サン マテオ, サウス エルスワース アベニュー 101, 350号室

(72)発明者 アントノフ, ヴァディム

アメリカ合衆国, カリフォルニア州 94401, サン マテオ, サウス エルスワース アベニュー 101, 350号室