

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-108635

(P2012-108635A)

(43) 公開日 平成24年6月7日(2012.6.7)

(51) Int. Cl.	F 1	テーマコード (参考)
G 0 6 F 17/30 (2006.01)	G 0 6 F 17/30 3 4 0 C	5 B 0 7 5
	G 0 6 F 17/30 1 8 0 D	
	G 0 6 F 17/30 1 1 0 C	

審査請求 未請求 請求項の数 7 O L (全 24 頁)

(21) 出願番号 特願2010-255654 (P2010-255654)
 (22) 出願日 平成22年11月16日 (2010.11.16)

(71) 出願人 000004237
 日本電気株式会社
 東京都港区芝五丁目7番1号
 (74) 代理人 100079164
 弁理士 高橋 勇
 (72) 発明者 並木 悠太
 東京都港区芝五丁目7番1号 日本電気株式会社内
 Fターム(参考) 5B075 KK03 QT06 UU40

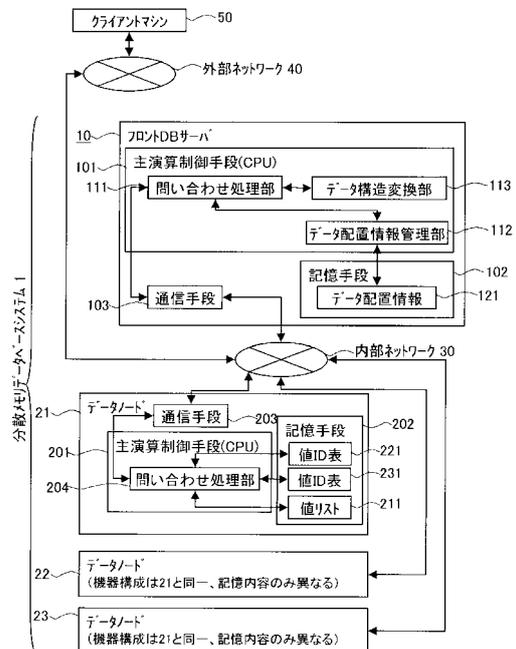
(54) 【発明の名称】 分散メモリデータベースシステム、フロントデータベースサーバ、データ処理方法およびプログラム

(57) 【要約】

【課題】 通信量および処理量を削減して、合計などの集合関数に対する処理を高速に行うことを可能とする分散メモリデータベースシステム等を提供する。

【解決手段】 フロントデータベースサーバ10が、外部から入力される表データを分割して複数の値ID表を生成してこれらを各データノードに分散して記憶させるデータ構造変換部113と、外部のクライアントマシンから発行された集合関数を含むクエリーに基づいて各データノードに表データの中の特定の値IDの出現数を問い合わせると共に、これに応じて各データノードから返された特定の値IDの出現数からクエリーに対応する集合関数の値を計算してクライアントマシンに返送する問い合わせ処理部111とを有し、データ構造変換部が、表データの中で集計軸になり得る列としてあらかじめ指定された複数のデータ項目の各々について個別に複数の値ID表を生成する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

フロントデータベースサーバと複数台のデータノードとが相互に接続されて構成される分散メモリデータベースシステムであって、

前記フロントデータベースサーバが、

外部から入力される表データを分割して各々の実際のデータを値 ID に置換した複数の値 ID 表を生成してこれらを前記各データノードに分散して記憶させるデータ構造変換部と、

外部のクライアントマシンから発行された集合関数を含むクエリーに基づいて前記各データノードに対して前記表データの中の特定の値 ID の出現数を問い合わせると共に、これに応じて前記各データノードから返された特定の値 ID の出現数から前記クエリーに対応する集合関数の値を計算して前記クライアントマシンに返送する問い合わせ処理部とを有し、

前記データ構造変換部が、前記複数の値 ID 表を生成する際に、前記表データの中で集計軸になり得る列としてあらかじめ指定された複数のデータ項目の各々について個別に前記複数の値 ID 表を生成することを特徴とする分散メモリデータベースシステム。

【請求項 2】

前記フロントデータベースサーバの前記データ構造変換部が、前記複数の値 ID 表を生成する際に、前記表データの中で集計軸および集計対象になり得る列としてあらかじめ指定されたデータ項目の各々について実際の値を値 ID に置換すると共に、前記値 ID と前記実際の値との対応を示す値リストを生成する機能を有し、

前記フロントデータベースサーバが、前記値リストおよび前記値 ID 表が前記各データノードの中のいずれに分散されたかをあらかじめ備えられた記憶手段に記憶するデータ配置情報管理部を有することを特徴とする、請求項 1 に記載の分散メモリデータベースシステム。

【請求項 3】

前記フロントデータベースサーバの前記データ構造変換部が、前記複数の値リストを生成する際に、前記実際の値を大小順にソートしてから前記値 ID に置換することを特徴とする、請求項 2 に記載の分散メモリデータベースシステム。

【請求項 4】

前記フロントデータベースサーバの前記データ構造変換部が、前記表データの中で集計軸および集計対象のいずれにもなり得ない列としてあらかじめ指定されたデータ項目の各々について、前記集計軸になり得る列についての前記値 ID 表のいずれか 1 種類に付随して記憶させることを特徴とする、請求項 2 に記載の分散メモリデータベースシステム。

【請求項 5】

複数台のデータノードと相互に接続されて分散メモリデータベースシステムを構成するフロントデータベースサーバであって、

外部から入力される表データを分割して各々の実際のデータを値 ID に置換した複数の値 ID 表を生成してこれらを前記各データノードに分散して記憶させるデータ構造変換部と、

外部のクライアントマシンから発行された集合関数を含むクエリーに基づいて前記各データノードに対して前記表データの中の特定の値 ID の出現数を問い合わせると共に、これに応じて前記各データノードから返された特定の値 ID の出現数から前記クエリーに対応する集合関数の値を計算して前記クライアントマシンに返送する問い合わせ処理部とを有し、

前記データ構造変換部が、前記複数の値 ID 表を生成する際に、前記表データの中で集計軸になり得る列としてあらかじめ指定された複数のデータ項目の各々について個別に前記複数の値 ID 表を生成することを特徴とするフロントデータベースサーバ。

【請求項 6】

フロントデータベースサーバと複数台のデータノードとが相互に接続されて構成される

10

20

30

40

50

分散メモリデータベースシステムにあって、

外部からの表データの入力を前記フロントデータベースサーバのデータ構造変換部が受け付け、

入力された前記表データを、集計軸になり得る列としてあらかじめ指定されたデータ項目の各々について前記フロントデータベースサーバのデータ構造変換部が個別に分割して実際のデータを値IDに置換した複数の値ID表を生成し、

生成された前記複数の値ID表を前記フロントデータベースサーバのデータ構造変換部が前記各データノードに分散して記憶させ、

外部のクライアントマシンから発行された集合関数を含むクエリーを前記フロントデータベースサーバの問い合わせ処理部が受け付け、

受け付けられた前記クエリーに基づいて前記表データの中の特定の値IDの出現数を前記フロントデータベースサーバの問い合わせ処理部が前記各データノードに問い合わせ、

前記各データノードから返された特定の値IDの出現数から前記フロントデータベースサーバの問い合わせ処理部が前記クエリーに対応する集合関数の値を計算して前記クライアントマシンに返送する

ことを特徴とするデータ処理方法。

【請求項7】

フロントデータベースサーバと複数台のデータノードとが相互に接続されて構成される分散メモリデータベースシステムにあって、

前記フロントデータベースサーバが備えるコンピュータに、

外部からの表データの入力を受け付ける手順、

入力された前記表データを、集計軸になり得る列としてあらかじめ指定されたデータ項目の各々について個別に分割して実際のデータを値IDに置換した複数の値ID表を生成する手順、

生成された前記複数の値ID表を前記各データノードに分散して記憶させる手順、

外部のクライアントマシンから発行された集合関数を含むクエリーを受け付ける手順、

受け付けられた前記クエリーに基づいて前記表データの中の特定の値IDの出現数を前記各データノードに問い合わせる手順、

および前記各データノードから返された特定の値IDの出現数から前記クエリーに対応する集合関数の値を計算して前記クライアントマシンに返送する手順

を実行させることを特徴とするデータ処理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は分散メモリデータベースシステム、フロントデータベースサーバ、データ処理方法およびプログラムに関し、特に集合関数に対する処理を迅速に実施することの可能な分散メモリデータベースシステム等に関する。

【背景技術】

【0002】

ある程度以上の規模のコンピュータ装置を利用したシステム、たとえばウェブサービスや業務システム等では、大量のデータを扱うためにデータベース管理システム（DBMS：DataBase Management System）の利用が不可欠である。また近年は、ネットワークによって接続された多数のコンピュータを連携させて、1つの巨大なコンピュータとして処理を行わせる、いわゆるクラウドコンピューティングの技術が確立されてきている。

【0003】

DBMSの中でも特に、データベースに係る処理をネットワークによって接続された多数のコンピュータに分担させるものを、分散メモリデータベースシステムという。分散メモリデータベースシステムは、大量のデータを一括で処理するバッチ用途や、企業の扱う大量のデータから特定の部署が必要とするものを抽出するデータマートの作成などで、特に処理の高速化の効果が発揮される。

10

20

30

40

50

【 0 0 0 4 】

以下、これについて説明する。なお、本明細書では説明を平易なものとするために、ごく少ないデータ個数および項目数について例示するが、実際には膨大な個数および項目数のデータに対して例示したような処理を行うものである。

【 0 0 0 5 】

図 1 4 は、一般的な分散メモリデータベースシステム 9 0 1 の構成を示す説明図である。分散メモリデータベースシステム 1 は、フロントメモリデータベースサーバ 9 1 0 (以後フロント DB サーバ 9 1 0 という) と、複数台のデータノード 9 2 1 ~ 9 2 3 とが内部ネットワーク 9 3 0 を介して相互に接続されて構成される。図 1 では 3 台のデータノード 9 2 1 ~ 9 2 3 を示しているが、もちろんこの台数は 2 台以上何台でもよい。また、フロント DB サーバ 1 0 は、外部ネットワーク 9 4 0 を介して、クライアントマシン 9 5 0 と接続されている。

10

【 0 0 0 6 】

クライアントマシン 9 5 0 はフロント DB サーバ 9 1 0 に対してクエリー (処理要求) を発行するコンピュータである。このクエリーに基づくデータ処理をフロント DB サーバ 9 1 0 とデータノード 9 2 1 ~ 9 2 3 とが連携して行い、フロント DB サーバ 9 1 0 はその検索結果をクライアントマシン 5 0 に返す。その際、フロント DB サーバ 9 1 0 は、クライアントマシン 9 5 0 から発行されたクエリーの各データノード 9 2 1 ~ 9 2 3 に向けての分割と、各データノード 9 2 1 ~ 9 2 3 からの結果の集約とを行う。

20

【 0 0 0 7 】

フロント DB サーバ 9 1 0 では、問い合わせ処理部 9 1 1、データ配置情報管理部 9 1 2、およびデータ構造変換部 9 1 3 が、各々コンピュータプログラムとして後述するそれぞれの機能を実行されるように構成されている。

【 0 0 0 8 】

問い合わせ処理部 9 1 1 は、クライアントマシン 9 5 0 が発行したクエリーを受け付け、このクエリーで処理対象となるデータ項目の所在をデータ配置情報管理部 9 1 2 に対して照会し、この照会に応じてデータ配置情報管理部 1 1 2 から得られた回答を元にしてクライアントマシン 9 5 0 からのクエリーを各データノード 9 2 1 ~ 9 2 3 ごとに分割して、分割されたクエリーを各データノード 9 2 1 ~ 9 2 3 に送信する。そして、送信した各クエリーに対する各データノード 9 2 1 ~ 9 2 3 からの回答を集約してクライアントマシン 9 5 0 に返す。

30

【 0 0 0 9 】

データ配置情報管理部 9 1 2 は、問い合わせ処理部 9 1 1 から照会のあったデータ項目のデータが各データノード 9 2 1 ~ 9 2 3 のうちのいずれに存在するかを、問い合わせ処理部 9 1 1 に返答する。データ構造変換部 9 1 3 は、入力された表構造データを分割して、各データノード 9 2 1 ~ 9 2 3 に記憶させる。

【 0 0 1 0 】

各データノード 9 2 1 ~ 9 2 3 では、問い合わせ処理部 9 1 4 がコンピュータプログラムとして、問い合わせ処理部 9 1 1 で分割されたクエリーに基づいて検索などの処理を行ってその結果をフロント DB サーバ 9 1 0 に返す機能を実行されるように構成されている。

40

【 0 0 1 1 】

図 1 5 は、図 1 4 に示した分散メモリデータベースシステム 9 0 1 に対して入力される表データ 9 6 0 の一例を示す説明図である。図 1 5 で示した表データ 9 6 0 は、日付 9 6 0 a、店 ID 9 6 0 b、売上 9 6 0 c という 3 つの項目のデータを持つ。このうちの日付 9 6 0 a は「8 月 1 0 日」「8 月 1 1 日」の 2 種類の値が使われている。

【 0 0 1 2 】

データ構造変換部 9 1 3 は、このうちの店 ID 9 1 0 b の値「A 1」「D 3」「E 1」を基準として表データ 9 6 0 を分割して、店 ID 別の表データ 9 6 1 ~ 9 6 3 を生成し、これらを各々各データノード 9 2 1 ~ 9 2 3 に送付して記憶させる。図 1 6 は、図 1 5 に

50

示した表データ910をデータ構造変換部913が分割して生成した店ID別の表データ961～963を示す説明図である。

【0013】

図16に示した店ID別の表データ961～963が各々データノード921～923に記憶された状態で、問い合わせ処理部911はクライアントマシン950が発行したクエリーを受け付ける。たとえば、以下の数1に示すクエリーを受け付けた場合、問い合わせ処理部911は各データノード921～923に対して、各々が記憶している店ID別の表データ961～963から店IDごとの合計売上額を算出させ、返信されてきた店IDごとの合計売上額をまとめてクライアントマシン950に返信する。

【数1】

SELECT SUM(売上) FROM 売上表 GROUP BY 店ID

【0014】

これに関連する技術として、以下の各々がある。その中でも特許文献1には、並列コンピュータの複数の処理モジュールで処理されるデータに共通のグローバル次元値番号を付与することによって、プロセッサ間通信の発生を少なくしてデータのソートや集計などを実現するという情報処理システムが記載されている。特許文献2には、ハッシュ値を利用して1つ以上の列の値が共通する行をグループ化する処理を高速化するというグループ化方法が記載されている。

【0015】

特許文献3には、複数の分析問題を含むデータをレイヤに分割することによって、複数の計算機でのデータ分析を高速化するというデータ分析システムが記載されている。特許文献4には、複数の装置にデータを分散させて処理する際にシステムの終了処理時間を短縮するというデータベース処理方法が記載されている。

【0016】

特許文献5には、データにラベルコードを付与して、複数の装置によるデータの加工処理を高速化するというデータ加工システムが記載されている。非特許文献1には、データベースとして多く利用されているオラクル(登録商標)データベースで、完全一致検索を高速化することが可能である「ビットマップインデックス」について記載されている。

【先行技術文献】

【特許文献】

【0017】

【特許文献1】再特W02005/041067号公報

【特許文献2】特開2000-187668号公報

【特許文献3】特開2006-107129号公報

【特許文献4】特開2010-134583号公報

【特許文献5】特開平07-182368号公報

【非特許文献】

【0018】

【非特許文献1】Paul Lane、「データ・ウェアハウスでのビットマップ索引の使用」、Oracle Databaseデータ・ウェアハウス・ガイド11gリリース1より、2007年、日本オラクル株式会社、[平成22年11月8日検索]、インターネット<URL:http://otndnld.oracle.co.jp/document/products/oracle11g/111/doc_dvd/server.111/E05763-01/indexes.htm>

【発明の概要】

【発明が解決しようとする課題】

【0019】

分散メモリデータベースシステムでは特に、処理にかかるコストおよび処理時間を低減するため、ノード間で発生する通信を可能な限り少なくすることが重要である。図14～

10

20

30

40

50

16に示した分散メモリデータベースシステム901の例は、説明を平易なものとするために、ごく少ないデータ個数について例示したものであるが、実際には膨大な量のデータについて処理を行う必要がある。

【0020】

図14～16に示した分散メモリデータベースシステム901では、「店ID」ごとにデータを分割して各データノード921～923に記憶させているので、数1に示したように店IDごとの集計であればデータノード921～923の各々の単体のみで合計などの集合関数の計算ができる。従って、問い合わせ処理部911での処理は単純に各データノード921～923から受信した合計などの数値をまとめるだけでよく、高速に集計処理を行うことができる。

10

【0021】

しかしながら、データの分割の基準とした列（この場合は「店ID」）以外で集計を行おうとすると、各データノード921～923相互間、およびそれらとフロントDBサーバ910との間でデータの交換が必要になる。そのため、集計処理に時間がかかり、またデータの交換に伴う通信コストが発生する。

【0022】

図14～16に示した分散メモリデータベースシステム901の例でいうなら、以下の数2に示すクエリを受け付けた場合には、データノード921～923の各々の単体のみで合計などの集合関数の計算ができないので、全データについて「日付」を基準として再分割を行う（第1の方法）か、もしくは「店ID」ごとにデータを保持している各データノード921～923で「日付」ごとの集計を行い、その結果をフロントDBサーバ910に送付して「日付」ごとの合計を算出する（第2の方法）かのいずれかの方法で算出することとなる。

20

【数2】

SELECT SUM(売上) FROM 売上表 GROUP BY 日付

【0023】

この第1の方法では、一度表データ961～963の全てをフロントDBサーバ910に送付して、元の表データ960に戻してから「日付」を基準として再分割を行った表データを改めて各データノード921～923に送付する必要がある。例えばm台のデータノードがあり、ここに一行あたりlバイトのデータが1台のノードに1日分n件、s(m, n, s, lは各々自然数、簡単のためにs=mとする)店舗分のデータが均等に存在するとした場合、この表データを再分割するには、以下の数3に示す通りのデータ容量の通信が発生することとなる。

30

【数3】

$$1 \text{ ノードあたり } \frac{n}{(s-l)} l \text{ バイト}$$

40

$$\text{システム全体 } m \times \frac{n}{(s-l)} l \text{ バイト}$$

【0024】

また、発生する通信量だけでなく、個々のデータを読み取り、その各々をどのデータノードに移動するかを判定する処理も必要となるので、フロントDBサーバ910および各データノードで必要な処理量も増大する。

【0025】

50

第2の方法では、各データノード921～923から「店ID・日付」ごとに求められた合計データがフロントDBサーバ910に送信される。このため、フロントDBサーバ910への通信量が多くなる。また、フロントDBサーバ910では(数1に示したクエリの場合と違って)単純な差し替えだけでなく、改めて「日付」ごとの合計を求める必要があるため、ここでの処理量も増大する。

【0026】

この問題に対して、インデックス(索引)を用意して異なる集計軸による集計に対応するという方法が既に知られている。しかしながら、この場合であっても、集計操作のために計算対象のデータを他のコンピュータに対して送信する必要は発生する。このため、発生する通信量を削減する効果は小さい。

10

【0027】

また、特許文献1および4～5には、同一列の中で同一の値にID(もしくはラベル)を付与して、これに基づいてグループ分け(いわゆるレンジパーティショニング)を行ったデータを各データノードに保存するという技術が記載されている。これを利用すれば、生のデータそのものを通信するよりも、多少の通信量の削減にはなる。しかしながら、フロントDBサーバ910の側での処理量は軽減されないどころか、IDを実際の値に置換する処理を伴うので、処理量はむしろ増大する。残る特許文献2～4および非特許文献1にも、この問題点について解決しうる技術は記載されていない。

【0028】

本発明の目的は、通信量および処理量を削減して、合計などの集合関数に対する処理を高速に行うことを可能とする分散メモリデータベースシステム、フロントデータベースサーバ、データ処理方法およびプログラムを提供することにある。

20

【課題を解決するための手段】

【0029】

上記目的を達成するため、本発明に係る分散メモリデータベースシステムは、フロントデータベースサーバと複数台のデータノードとが相互に接続されて構成される分散メモリデータベースシステムであって、フロントデータベースサーバが、外部から入力される表データを分割して各々の実際のデータを値IDに置換した複数の値ID表を生成してこれらを各データノードに分散して記憶させるデータ構造変換部と、外部のクライアントマシンから発行された集合関数を含むクエリーに基づいて各データノードに対して表データの中の特定の値IDの出現数を問い合わせると共に、これに応じて各データノードから返された特定の値IDの出現数からクエリーに対応する集合関数の値を計算してクライアントマシンに返送する問い合わせ処理部とを有し、データ構造変換部が、複数の値ID表を生成する際に、表データの中で集計軸になり得る列としてあらかじめ指定された複数のデータ項目の各々について個別に複数の値ID表を生成することを特徴とする。

30

【0030】

上記目的を達成するため、本発明に係るフロントデータベースサーバは、複数台のデータノードと相互に接続されて分散メモリデータベースシステムを構成するフロントデータベースサーバであって、外部から入力される表データを分割して各々の実際のデータを値IDに置換した複数の値ID表を生成してこれらを各データノードに分散して記憶させるデータ構造変換部と、外部のクライアントマシンから発行された集合関数を含むクエリーに基づいて各データノードに対して表データの中の特定の値IDの出現数を問い合わせると共に、これに応じて各データノードから返された特定の値IDの出現数からクエリーに対応する集合関数の値を計算してクライアントマシンに返送する問い合わせ処理部とを有し、データ構造変換部が、複数の値ID表を生成する際に、表データの中で集計軸になり得る列としてあらかじめ指定された複数のデータ項目の各々について個別に複数の値ID表を生成することを特徴とする。

40

【0031】

上記目的を達成するため、本発明に係るデータ処理方法は、フロントデータベースサーバと複数台のデータノードとが相互に接続されて構成される分散メモリデータベースシ

50

テムにあって、外部からの表データの入力をフロントデータベースサーバのデータ構造変換部が受け付け、入力された表データを、集計軸になり得る列としてあらかじめ指定されたデータ項目の各々についてフロントデータベースサーバのデータ構造変換部が個別に分割して実際のデータを値IDに置換した複数の値ID表を生成し、生成された複数の値ID表をフロントデータベースサーバのデータ構造変換部が各データノードに分散して記憶させ、外部のクライアントマシンから発行された集合関数を含むクエリーをフロントデータベースサーバの問い合わせ処理部が受け付け、受け付けられたクエリーに基づいて表データの中の特定の値IDの出現数をフロントデータベースサーバの問い合わせ処理部が各データノードに問い合わせ、各データノードから返された特定の値IDの出現数からフロントデータベースサーバの問い合わせ処理部がクエリーに対応する集合関数の値を計算してクライアントマシンに返送することを特徴とする。

10

【0032】

上記目的を達成するため、本発明に係るデータ処理プログラムは、フロントデータベースサーバと複数台のデータノードとが相互に接続されて構成される分散メモリデータベースシステムにあって、フロントデータベースサーバが備えるコンピュータに、外部からの表データの入力を受け付ける手順、入力された表データを、集計軸になり得る列としてあらかじめ指定されたデータ項目の各々について個別に分割して実際のデータを値IDに置換した複数の値ID表を生成する手順、生成された複数の値ID表を各データノードに分散して記憶させる手順、外部のクライアントマシンから発行された集合関数を含むクエリーを受け付ける手順、受け付けられたクエリーに基づいて表データの中の特定の値IDの出現数を各データノードに問い合わせる手順、および各データノードから返された特定の値IDの出現数からクエリーに対応する集合関数の値を計算してクライアントマシンに返送する手順を実行させることを特徴とする。

20

【発明の効果】

【0033】

本発明は、上述した通り、集計軸になり得る列としてあらかじめ指定されたデータ項目の各々について表データを個別に分割するように構成したので、集計軸になり得るどの列に対しても、他の装置との通信の発生を抑制して、各データノードの内部だけで集合関数に関する処理を行うことが可能となる。これによって、通信量および処理量を削減して、合計などの集合関数に対する処理を高速に行うことを可能であるという優れた特徴を持つ分散メモリデータベースシステム、フロントデータベースサーバ、データ処理方法およびプログラムを提供することが可能となる。

30

【図面の簡単な説明】

【0034】

【図1】本発明の第1の実施形態に係る分散メモリデータベースシステムの構成を示す説明図である。

【図2】図1で説明した分散メモリデータベースシステムに対して入力される表データの一例を示す説明図である。

【図3】図2で示した表データからデータ構造変換部が作成する値リストおよび値ID表を示す説明図である。

40

【図4】図1に示したデータ構造変換部が、図2に示すデータを図3に示すように分割して各データノードに分配して記憶させる処理を示すフローチャートである。

【図5】図1に示したデータ配置情報の、図2および図3に示した各データに対応する例について示す説明図である。

【図6】数4に示したクエリーに対して、図1で説明した分散メモリデータベースシステムで行われる処理を示すフローチャートである。

【図7】図6のステップS403（数4）に示す処理で、データノードからフロントDBサーバに返却される「値ID」ごとの出現数の表を示す説明図である。

【図8】図6のステップS406（数5）に示した処理の結果、クライアントマシンに返却される日付ごとの売上の合計を示す結果データについて示す説明図である。

50

【図 9】図 6 のステップ S 4 0 3 (数 6) に示す処理で、データノードからフロント D B サーバに返却される「値 I D」ごとの出現数の表を示す説明図である。

【図 1 0】図 6 のステップ S 4 0 6 (数 7) に示した処理の結果、クライアントマシンに返却される店 I D ごとの売上の合計を示す結果データについて示す説明図である。

【図 1 1】本発明の第 2 の実施形態に係る分散メモリデータベースシステムの構成を示す説明図である。

【図 1 2】図 1 1 で説明した分散メモリデータベースシステムに対して入力される表データの一例を示す説明図である。

【図 1 3】図 1 2 に示した表データからデータ構造変換部が作成する値 I D 表の例を示す説明図である。

【図 1 4】一般的な分散メモリデータベースシステムの構成を示す説明図である。

【図 1 5】図 1 4 に示した分散メモリデータベースシステムに対して入力される表データの一例を示す説明図である。

【図 1 6】図 1 5 に示した表データをデータ構造変換部が分割して生成した店 I D 別の表データを示す説明図である。

【発明を実施するための形態】

【 0 0 3 5 】

(第 1 の実施形態)

以下、本発明の第 1 の実施形態の構成について添付図 1 ~ 3 に基づいて説明する。

最初に、本実施形態の基本的な内容について説明し、その後でより具体的な内容について説明する。

本実施形態に係る分散メモリデータベースシステム 1 は、フロントデータベースサーバ (フロント D B サーバ 1 0) と複数台のデータノード 2 1 ~ 2 3 とが相互に接続されて構成される分散メモリデータベースシステムである。フロントデータベースサーバ 1 0 は、外部から入力される表データを分割して各々の実際のデータを値 I D に置換した複数の値 I D 表 2 2 1 ~ 2 2 2 , 2 3 1 ~ 2 3 3 を生成してこれらを各データノードに分散して記憶させるデータ構造変換部 1 1 3 と、外部のクライアントマシン 5 0 から発行された集合関数を含むクエリーに基づいて各データノードに表データの中の特定の値 I D の出現数を問い合わせると共にこれに応じて各データノードから返された特定の値 I D の出現数からクエリーに対応する集合関数の値を計算してクライアントマシンに返送する問い合わせ処理部 1 1 1 とを有する。そして、データ構造変換部 1 1 3 は、これら複数の値 I D 表を生成する際に、表データの中で集計軸になり得る列としてあらかじめ指定されたデータ項目の各々について個別に複数の値 I D 表を生成する。

【 0 0 3 6 】

また、フロントデータベースサーバ 1 0 のデータ構造変換部 1 1 3 は、複数の値 I D 表 2 2 1 ~ 2 2 2 , 2 3 1 ~ 2 3 3 を生成する際に、表データの中で集計軸および集計対象になり得る列としてあらかじめ指定されたデータ項目の各々について実際の値を値 I D に置換すると共に、値 I D と実際の値との対応を示す値リスト 2 1 1 ~ 2 1 3 を生成する機能を有し、フロントデータベースサーバ 1 0 が、値リストおよび値 I D 表が各データノードの中のいずれに分散されたかをあらかじめ備えられた記憶手段に記憶するデータ配置情報管理部 1 1 2 を有する。

【 0 0 3 7 】

さらに、フロントデータベースサーバ 1 0 のデータ構造変換部 1 1 3 が、値 I D 表 2 2 1 ~ 2 2 2 , 2 3 1 ~ 2 3 3 を生成する際、実際の値を大小順にソートしてから値 I D に置換する。

【 0 0 3 8 】

以上の構成を備えることにより、分散メモリデータベースシステム 1 は、通信量および処理量を削減して、合計などの集合関数に対する処理を高速に行うことが可能となる。

以下、これをより詳細に説明する。

【 0 0 3 9 】

10

20

30

40

50

図1は、本発明の第1の実施形態に係る分散メモリデータベースシステム1の構成を示す説明図である。分散メモリデータベースシステム1は、フロントデータベースサーバ10（以後フロントDBサーバ10という）と、複数台のデータノード21～23とが内部ネットワーク30を介して相互に接続されて構成される。図1では3台のデータノード21～23を示しているが、もちろんこの台数は2台以上何台でもよい。

【0040】

また、内部ネットワーク30は、外部ネットワーク40を介して、外部のコンピュータ装置であるクライアントマシン50と接続されている。内部ネットワーク30および外部ネットワーク40の、ネットワーク方式やプロトコルなどは任意のものを利用できる。フロントDBサーバ10は、内部ネットワーク30および外部ネットワーク40を通じて、クライアントマシン50からの操作を受け付けることができる。

10

【0041】

クライアントマシン50はフロントDBサーバ10に対してクエリ（処理要求）を発行し、このクエリに基づくデータ処理をフロントDBサーバ10とデータノード21～23とが連携して行い、フロントDBサーバ10はその検索結果をクライアントマシン50に返す。その際、フロントDBサーバ10は、クライアントマシン50から発行されたクエリの各データノード21～23に向けての分割と、各データノード21～23からの結果の集約とを行う。

【0042】

フロントDBサーバ10は、主演算制御手段101、記憶手段102、および通信手段103を備えるコンピュータ装置である。主演算制御手段101はコンピュータプログラムの動作主体となるCPU（Central Processing Unit）であり、記憶手段102は主演算制御手段101が作業中のデータを記憶するRAM（Random Access Memory）などのような主記憶装置である。通信手段103は内部ネットワーク30および外部ネットワーク40を介して、他のコンピュータとのデータ通信を行う。

20

【0043】

主演算制御手段101では、問い合わせ処理部111、データ配置情報管理部112、およびデータ構造変換部113が、各々コンピュータプログラムとして後述するそれぞれの機能を実行されるように構成されている。また、記憶手段102には、後述するデータ配置情報121が記憶されている。

30

【0044】

問い合わせ処理部111は、クライアントマシン50が発行したクエリを受け付け、このクエリで処理対象となるデータ項目の所在をデータ配置情報管理部112に対して照会し、この照会に応じてデータ配置情報管理部112から得られた回答を元にしてクライアントマシン50からのクエリを各データノード21～23ごとに分割して、分割されたクエリを各データノード21～23に送信する。そして、送信した各クエリに対する各データノード21～23からの回答を集約してクライアントマシン50に返す。

【0045】

データ配置情報管理部112は、問い合わせ処理部111から照会のあったデータ項目のデータが各データノード21～23のうちのいずれに存在するかを、データ配置情報121を参照して、問い合わせ処理部111に返答する。

40

【0046】

データ構造変換部113は、この後説明するように、表構造データを本システムに入力する際に、後述する本実施形態に独特のデータ構造に変換して分割し、各データノード21～23に記憶させる。

【0047】

各データノード21～23も、フロントDBサーバ10と同じく一般的なコンピュータ装置としての構成を備えるが、ハードウェアおよびソフトウェアとしては全て同一の構成を有し、ただ各々の記憶している内容が異なっているのみである。従って、図1ではデータノード21についてのみ詳しい構成を示す。データノード21は、フロントDBサーバ

50

10と同様に、主演算制御手段201、記憶手段202、および通信手段203を備えるコンピュータ装置である。

【0048】

主演算制御手段201では、問い合わせ処理部204が、各々コンピュータプログラムとして後述するそれぞれの機能を実行されるように構成されている。また、記憶手段202には、後述する値リスト211~213、および値ID表221~222, 231~233が記憶されている。問い合わせ処理部204は、フロントDBサーバ10で分割生成された値リスト211~213、および値ID表221~222, 231~233のうち、自身に割り当てられたものを記憶手段202に記憶し、これらの値リストおよび値ID表に対して検索などの処理を行い、その結果をフロントDBサーバ10に返す。

10

【0049】

図2は、図1で説明した分散メモリデータベースシステム1に対して入力される表データ210の一例を示す説明図である。このデータ入力はクライアントマシン50から行われてもよいし、フロントDBサーバ10から直接行われてもよい。データ構造変換部113は、この入力された表データ210を、以下で説明する形式に変換して、複数のデータノード21~23に分配して記憶させる。

【0050】

図2で示した表データ210は、日付210a、店ID210b、売上210cという3つの項目のデータを持つ。このうちの「日付210a」および「店ID210b」が「集計軸になり得る列(以後基準列という)」、売上210cが「集計対象になり得る列(以後対象列という)」としてあらかじめ指定されている。なお、表データ210の内容は図15の表データ910と同一である。

20

【0051】

図3は、図2で示した表データ210からデータ構造変換部113が作成する値リスト211~213、および値ID表221~222, 231~233を示す説明図である。図2で示した表データ210で、基準列として指定された項目のうち、日付210aは「8月10日」「8月11日」の2種類、店ID210bは「A1」「D3」「E1」の3種類の値が使われている。

【0052】

そこで、データ構造変換部113は、各列に存在するユニークな値を小さいものから順に整列して並べ、先頭から順にそれぞれの値を特定する番号(値ID)を付与し、「日付」値リスト211、「店ID」値リスト212、「売上」値リスト213という3通りの値リストを作成する。

30

【0053】

「日付」値リスト211は、「日付」値IDと日付との対応を示す。「日付」値ID = 「0」の場合には日付 = 「8月10日」、 「日付」値ID = 「1」の場合には日付 = 「8月11日」となる。

【0054】

「店ID」値リスト212は、「店ID」値IDと店IDの値との対応を示す。「店ID」値ID = 「0」の場合には店ID = 「A1」、 「店ID」値ID = 「1」の場合には店ID = 「D3」、 「店ID」値ID = 「2」の場合には店ID = 「E1」となる。

40

【0055】

「売上」値リスト213は、「売上」値IDと売上の値との対応を示す。「売上」値ID = 「0」~「4」の場合で、売上は各々「800」「1000」「1200」「4800」「12000」となる。

【0056】

そしてデータ構造変換部113は、各々の値IDで表データ210を置き換え、さらに日付210aと店ID210bの値ごとに表データ210を分割して、値ID表221~222, 231~233を作成する。値ID表221は、「日付」値ID = 「0」の場合の「店ID」値IDと「売上」値IDの対応を示す。値ID表222は、「日付」値ID

50

= 「 1 」 の場合の「店 I D」値 I D と「売上」値 I D の対応を示す。

【 0 0 5 7 】

値 I D 表 2 3 1 は、「店 I D」値 I D = 「 0 」 の場合の「日付」値 I D と「売上」値 I D の対応を示す。値 I D 表 2 3 2 は、「店 I D」値 I D = 「 1 」 の場合の「日付」値 I D と「売上」値 I D の対応を示す。値 I D 表 2 3 3 は、「店 I D」値 I D = 「 2 」 の場合の「日付」値 I D と「売上」値 I D の対応を示す。

【 0 0 5 8 】

データ構造変換部 1 1 3 は、以上で作成した値リスト 2 1 1 ~ 2 1 3、および値 I D 表 2 2 1 ~ 2 2 2、2 3 1 ~ 2 3 3 を、各データノード 2 1 ~ 2 3 に分配して記憶させる。図 3 に示した例では、データ構造変換部 1 1 3 は、値 I D 表 2 2 2、値 I D 表 2 3 1、および値リスト 2 1 1 をデータノード 2 1 の記憶手段 2 0 2 に記憶させている。また、値 I D 表 2 2 1、値 I D 表 2 3 2、および値リスト 2 1 2 をデータノード 2 2 の記憶手段 2 0 2 に記憶させている。さらに値 I D 表 2 3 3、および値リスト 2 1 3 をデータノード 2 3 の記憶手段 2 0 2 に記憶させている。

10

【 0 0 5 9 】

ここでデータ構造変換部 1 1 3 が行う処理については、データノード 2 1 ~ 2 3 のうちの特定の 1 つのデータノードに、データの容量および処理量が大きく偏ることがなければ、任意の分割方法を使用することができる。また、図 2 で示した表データ 2 1 0 のデータ構造を設計する際に、基準列および対象列といった各列の属性について、あらかじめ操作者が入力しているものとする。データ構造変換部 1 1 3 は、「基準列」および「対象列」であると設定されたデータ列の全てについて値リストを作成し、かつ「基準列」であると設定されたデータ列の全てについて上記で示したように表データ 2 1 0 を分割する。

20

【 0 0 6 0 】

図 4 は、図 1 に示したデータ構造変換部 1 1 3 が、図 2 に示すデータを図 3 に示すように分割して各データノード 2 1 ~ 2 3 に分配して記憶させる処理を示すフローチャートである。まず、表データ 2 1 0 のデータ構造について、ユーザによる入力を受け付ける（ステップ S 3 0 1）。この際に、どのデータ項目が「基準列」や「対象列」であるかについての入力も同時に行われる。

【 0 0 6 1 】

その次に、ユーザから表データ 2 1 0 の入力を受け付け（ステップ S 3 0 2）、これが完了すると基準列のうちの 1 つについて（ステップ S 3 0 3 ~ 3 0 4）、まずデータ列の値を値 I D に置き換えて（ステップ S 3 0 5）、同時にこの基準列の値ごとに表を分割する（ステップ S 3 0 6）。データ構造変換部 1 1 3 は、このステップ S 3 0 5 ~ 3 0 6 の処理を全ての基準列について繰り返すと共に、置き換えた値と値 I D の対応を示す値リスト 2 1 1 ~ 2 1 3 も同時に作成する。

30

【 0 0 6 2 】

全ての基準列についてこの処理が完了したら、出来上がった値リストおよび値 I D 表を、各データノード 2 1 ~ 2 3 に送信して記憶させる（ステップ S 3 0 7）。

【 0 0 6 3 】

図 5 は、図 1 に示したデータ配置情報 1 2 1 の、図 2 および図 3 に示した各データに対応する例について示す説明図である。データ配置情報 1 2 1 は、各データノード 2 1 ~ 2 3 のコンピュータ名を示すデータノード名 1 2 1 a と、当該コンピュータに記憶される値リスト 2 1 1 ~ 2 1 3、および値 I D 表 2 2 1 ~ 2 2 2、2 3 1 ~ 2 3 3 のデータ種類を示す記憶内容 1 2 1 b との対応を示す。

40

【 0 0 6 4 】

データ配置情報 1 2 1 は、図 5 で示した例のような形に限られるものではなく、計算で使用されるデータ、および値 I D に対応する実際の値が、各データノード 2 1 ~ 2 3 のうちのいずれに記憶されているかを特定できさえすれば、どのようなデータ形式でもよい。

【 0 0 6 5 】

（日付別の売上集計処理）

50

このように各データノード 2 1 ~ 2 3 に分配して記憶されたデータに対して、クライアントマシン 5 0 からフロント DB サーバ 1 0 に対してクエリー (SQL 文) が発行された際の、フロント DB サーバ 1 0 および各データノード 2 1 ~ 2 3 で行われる処理について以下に説明する。

【 0 0 6 6 】

以下に示す数 4 は、図 2 ~ 3 に示した内容のデータ例に対して、クライアントマシン 5 0 がフロント DB サーバ 1 0 に対して発行するクエリーの一例である。これは、表データ 2 1 0 から日付ごとの売上の合計を求めるクエリーである。

【 数 4 】

SELECT SUM(売上) FROM 売上表 GROUP BY 日付

10

【 0 0 6 7 】

図 6 は、数 4 に示したクエリーに対して、図 1 で説明した分散メモリデータベースシステム 1 で行われる処理を示すフローチャートである。数 4 に示したクエリーを受けたフロント DB サーバ 1 0 の問い合わせ処理部 1 1 1 は、データ配置情報管理部 1 1 2 がデータ配置情報 1 2 1 を参照して、「売上表」に対し「日付」別の値 ID 表が各データノード 2 1 ~ 2 3 上に存在することを確認し、分割された各表の存在するノードを特定する (ステップ S 4 0 1)。

【 0 0 6 8 】

そして、フロント DB サーバ 1 0 の問い合わせ処理部 1 1 1 は、データノード 2 1 ~ 2 3 の問い合わせ処理部 2 0 4 に対して、日付別の値 ID 表 2 2 1 ~ 2 2 2 で、「値 ID」によって表現された「売上」列の値を「値 ID」ごとの出現数を数えるよう問い合わせを発行する (ステップ S 4 0 2)。データノード 2 3 は日付別の値 ID 表を記憶していないので、ここでは対象外となる。

20

【 0 0 6 9 】

問い合わせを受け付けたデータノード 2 1 ~ 2 2 の問い合わせ処理部 2 0 4 は、日付別の値 ID 表 2 2 1 ~ 2 2 2 に対し、「売上」列の「値 ID」ごとの出現数を計算して、これをフロント DB サーバ 1 0 に返却する (ステップ S 4 0 3)。図 7 は、図 6 のステップ S 4 0 3 (数 4) に示す処理で、データノード 2 1 ~ 2 2 からフロント DB サーバ 1 0 に返却される「値 ID」ごとの出現数の表 2 4 1 ~ 2 4 2 を示す説明図である。

30

【 0 0 7 0 】

これを受けたフロント DB サーバ 1 0 の問い合わせ処理部 1 1 1 は、「日付」と「売上」の各々の「値 ID」に対応する値リスト 2 1 1 ~ 2 1 3 を、各データノード 2 1 ~ 2 3 から取得する (ステップ S 4 0 4 ~ 4 0 5)。なお、このステップ S 4 0 4 ~ 4 0 5 の処理は、値リスト 2 1 1 ~ 2 1 3 そのものを各データノード 2 1 ~ 2 3 から取得するのではなく、値 ID に対応する値を各データノード 2 1 ~ 2 3 から取得するものでもよい。

【 0 0 7 1 】

そして、フロント DB サーバ 1 0 の問い合わせ処理部 1 1 1 は、ステップ S 4 0 3 で返却された「売上」列の「値 ID」ごとの出現数に、実際の「売上」の数値を適用して、実際の売上金額を結果データ 2 4 3 として算出して、これをクライアントマシン 5 0 に返却する (ステップ S 4 0 6)。

40

【 0 0 7 2 】

以下に示す数 5 は、図 6 のステップ S 4 0 6 に示した処理で、図 2 ~ 3 に示した内容のデータ例に対して数 4 のクエリーによって実際に行われる計算を示す。図 8 は、図 6 のステップ S 4 0 6 (数 5) に示した処理の結果、クライアントマシン 5 0 に返却される日付ごとの売上の合計を示す結果データ 2 4 3 について示す説明図である。

【数5】

$$\begin{aligned} & \text{SUM}(8/10\text{分の売上}) \\ & = (\text{値ID:0}) \times 1 + (\text{値ID:1}) \times 1 + (\text{値ID:4}) \times 1 \\ & = 800 \times 1 + 1000 \times 1 + 12000 \times 1 \\ & = 13800 \end{aligned}$$

$$\begin{aligned} & \text{SUM}(8/11\text{分の売上}) \\ & = (\text{値ID:2}) \times 1 + (\text{値ID:3}) \times 1 \\ & = 1200 \times 1 + 4800 \times 1 \\ & = 6000 \end{aligned}$$

10

【0073】

(店ID別の売上集計処理)

上記で示した処理は、たとえば以下の数6に示すクエリーをクライアントマシン50が発行した場合においても、図6で示した動作によって同様に処理することができる。これは、表データ210から店IDごとの売上の合計を求めるクエリーである。

20

【数6】

SELECT SUM(売上) FROM 売上表 GROUP BY 店ID

【0074】

この場合、問い合わせ処理部111は、データ配置情報管理部112がデータ配置情報121を参照して、「売上表」に対し「店ID」別の値ID表が各データノード21~23上に存在することを確認し、分割された各表の存在するノードを特定する(ステップS401)。

30

【0075】

そして、フロントDBサーバ10の問い合わせ処理部111は、データノード21~23の問い合わせ処理部204に対して、店ID別の値ID表231~233で、「値ID」によって表現された「売上」列の値を「値ID」ごとの出現数を数えるよう問い合わせを発行する(ステップS402)。

【0076】

問い合わせを受け付けたデータノード21~23の問い合わせ処理部204は、店ID別の値ID表231~233に対し、「売上」列の「値ID」ごとの出現数を計算して、これをフロントDBサーバ10に返却する(ステップS403)。図9は、図6のステップS403(数6)に示す処理で、データノード21~23からフロントDBサーバ10に返却される「値ID」ごとの出現数の表251~253を示す説明図である。

40

【0077】

これを受けたフロントDBサーバ10の問い合わせ処理部111は、「店ID」と「売上」の各々の「値ID」に対応する値リスト211~213を、各データノード21~23から取得する(ステップS404~405)。なお、このステップS404~405の処理は、値リスト211~213そのものを各データノード21~23から取得するのではなく、値IDに対応する値を各データノード21~23から取得するものでもよい。

【0078】

そして、フロントDBサーバ10の問い合わせ処理部111は、ステップS403で返

50

却された「売上」列の「値ID」ごとの出現数に、実際の「売上」の数値を適用して、実際の売上金額を結果データ254として算出して、算出して、これをクライアントマシン50に返却する（ステップS406）。

【0079】

以下に示す数7は、図6のステップS406に示した処理で、図2～3に示した内容のデータ例に対して数6のクエリーによって実際に行われる計算を示す。図10は、図6のステップS406（数7）に示した処理の結果、クライアントマシン50に返却される店IDごとの売上の合計を示す結果データ254について示す説明図である。

【数7】

SUM(店ID:0の売上)

10

= (値ID:3)×1+(値ID:4)×1

= 4800×1+12000×1

= 15800

SUM(店ID:1の売上)

20

= (値ID:0)×1+(値ID:1)×1

= 800+1000

= 1800

SUM(店ID:2の売上)

= (値ID:2)×1

30

= 1200

【0080】

（和以外の集計処理）

以上、各データノード21～23に分配して記憶されたデータに対して、クライアントマシン50からフロントDBサーバ10に対して対象列の和（SUM）を求めるクエリーが発行された際の処理について説明したが、これ以外のたとえば最小値（MIN）、最大値（MAX）、出現数（COUNT）、平均値（AVG）を求める動作についても、同一のシステムで可能である。

【0081】

このうち最小値（MIN）もしくは最大値（MAX）を求めるクエリーを受けた場合は、各データノード21～23に記憶されている値ID表221～222，231～233から、問い合わせ処理部204が各々の値IDが最大もしくは最小となるものを選択してフロントDBサーバ10の問い合わせ処理部111に返却する。そしてフロントDBサーバ10の問い合わせ処理部111は、図6のステップS406と同様にして、値リスト211～213から返却された値IDに対応する値をクライアントマシン50に返却する。

40

【0082】

値リスト211～213は、前述したように値の小さいものから順に整列して並べて、その順番に値IDを割り振るものであるので、各データノード21～23の問い合わせ処理部204は、値リスト211～213で実際の値を参照しなくても、値IDが最大もしくは最小となるものを最大値もしくは最小値であると判断することができる。

50

【 0 0 8 3 】

出現数 (C O U N T) を求めるクエリーを受けた場合は、各データノード 2 1 ~ 2 3 に値 I D 表 2 2 1 ~ 2 2 2 , 2 3 1 ~ 2 3 3 に各値 I D ごとの出現数が記憶されているので、フロント D B サーバ 1 0 の問い合わせ処理部 1 1 1 は各データノード 2 1 ~ 2 3 側でその出現数を計算させる。そして、フロント D B サーバ 1 0 の問い合わせ処理部 1 1 1 は各データノード 2 1 ~ 2 3 から返却された出現数を受けて、図 6 のステップ S 4 0 6 と同様にして、その値 I D を実際の値に変換してクライアントマシン 5 0 に返却する。

【 0 0 8 4 】

平均値 (A V G) を求めるクエリーを受けた場合は、図 6 のステップ S 4 0 1 ~ 4 0 6 と同様の処理で、ただステップ S 4 0 6 でフロント D B サーバ 1 0 の問い合わせ処理部 1 1 1 で各々の値 I D を実際の値に変換してから平均値を求めてクライアントマシン 5 0 に返却する点のみが図 6 で示した処理と異なる。

10

【 0 0 8 5 】

(第 1 の実施形態の全体的な動作)

次に、上記の実施形態の全体的な動作について説明する。本実施形態に係るデータ処理方法は、フロントデータベースサーバ (フロント D B サーバ 1 0) と複数台のデータノード 2 1 ~ 2 3 とが相互に接続されて構成される分散メモリデータベースシステム 1 にあって、外部からの表データの入力をフロントデータベースサーバのデータ構造変換部が受け付け (図 4 ・ステップ S 3 0 1 ~ 3 0 2) 、入力された表データを、集計軸になり得る列としてあらかじめ指定されたデータ項目の各々についてフロントデータベースサーバのデータ構造変換部が個別に分割して実際のデータを値 I D に置換した複数の値 I D 表を生成し (図 4 ・ステップ S 3 0 4 ~ 3 0 6) 、生成された複数の値 I D 表をフロントデータベースサーバのデータ構造変換部が各データノードに分散して記憶させ (図 4 ・ステップ S 3 0 7) 、外部のクライアントマシンから発行された集合関数を含むクエリーをフロントデータベースサーバの問い合わせ処理部が受け付け、受け付けられたクエリーに基づいて表データの中の特定の値 I D の出現数をフロントデータベースサーバの問い合わせ処理部が各データノードに問い合わせ (図 6 ・ステップ S 4 0 1 ~ 4 0 2) 、各データノードから返された特定の値 I D の出現数からフロントデータベースサーバの問い合わせ処理部がクエリーに対応する集合関数の値を計算してクライアントマシンに返送する (図 6 ・ステップ S 4 0 4 ~ 4 0 6) 。

20

30

【 0 0 8 6 】

ここで、上記各動作ステップについては、これをコンピュータで実行可能にプログラム化し、これらを前記各ステップを直接実行するコンピュータであるフロント D B サーバ 1 0 に実行させるようにしてもよい。本プログラムは、非一時的な記録媒体、例えば、 D V D 、 C D 、フラッシュメモリ等に記録されてもよい。その場合、本プログラムは、記録媒体からコンピュータによって読み出され、実行される。

この動作により、本実施形態は以下のような効果を奏する。

【 0 0 8 7 】

本実施形態では、全ての基準列、即ち集計軸になり得る列について、表データ 2 1 0 を分割して各データノード 2 1 ~ 2 3 に分配して記憶させている。これによって、どの基準列についてクエリーが発行されたとしても、他の装置とのデータの交換を発生させずにデータノード 2 1 ~ 2 3 のうちの 1 台だけで集合関数に対する集計の処理を行って、その集計結果だけをフロント D B サーバ 1 0 に送信するようにできる。さらに、フロント D B サーバ 1 0 での処理も、単純に値 I D を実際の値に差し替えるだけでよいので、高速に行うことができる。

40

【 0 0 8 8 】

その際、各データノード 2 1 ~ 2 3 が記憶するデータは、値リストと値 I D による表現に変換した上で記憶されるので、各データノード 2 1 ~ 2 3 に記憶されるデータの容量を削減することができる。特に重複する値が多い場合に、そのデータの容量の削減の効果はより顕著なものとなる。さらに、複数の基準列について表データ 2 1 0 を分割する際のデ

50

ータ容量の増大も、最低限に抑制することができる。

【0089】

(第2の実施形態)

本発明の第2の実施形態では、フロントデータベースサーバ(フロントDBサーバ510)のデータ構造変換部が、表データの中で集計軸および集計対象のいずれにもなり得ない列としてあらかじめ指定されたデータ項目の各々について、集計軸になり得る列についての値ID表のいずれか1種類に付随して記憶させる構成とした。

【0090】

この構成によれば、集計軸および集計対象のいずれにもなり得ないデータ項目を含むデータに対しても、第1の実施形態と同一の効果を得ることができる。

10

以下、これをより詳細に説明する。

【0091】

図11は、本発明の第2の実施形態に係る分散メモリデータベースシステム501の構成を示す説明図である。分散メモリデータベースシステム501は、フロントDBサーバ510と複数台のデータノード521~523とが、第1の実施形態と同一の内部ネットワーク30を介して相互に接続されて構成される。外部ネットワーク40、およびクライアントマシン50も、第1の実施形態と同一である。

【0092】

フロントDBサーバ510およびデータノード521~523は、第1の実施形態のフロントDBサーバ10およびデータノード21~23と、ハードウェア的には同一の構成を有する。フロントDBサーバ510の主演算制御手段101で動作する問い合わせ処理部111およびデータ構造変換部113が、各々問い合わせ処理部511およびデータ構造変換部513に置換されている。また、データノード521~523については、各々が記憶している値ID表621~622, 631~632が第1の実施形態と異なる。

20

【0093】

図12は、図11で説明した分散メモリデータベースシステム501に対して入力される表データ610の一例を示す説明図である。表データ610は、項目A610a、項目B610b、項目C610c、項目D610dといったデータ項目を持つが、このうち基準列として項目A610aおよび項目B610b、対象列として項目D610dが指定されているが、項目C610cは基準列および対象列のいずれにも該当しない。なお、項目A610aは同時に主キーにも指定されている。

30

【0094】

図13は、図12に示した表データ610からデータ構造変換部513が作成する値ID表621~622, 631~632の例を示す説明図である。前述のように基準列でも対象列でもない項目C610cが表データ610に含まれる場合、データ構造変換部513は、項目A610aの値ID別の値ID表621~622にのみ項目C610cの値IDを同時に保持し、項目B610bの値ID別の値ID表631~632には項目C610cの値IDを含めないようにする。値リストは、項目A~Dの全てについて、第1の実施形態と同様に作成され、各データノード521~523に分散して記憶される。

【0095】

40

たとえば、項目B610bの特定の値もしくは値IDから、これに対応する項目C610cの値もしくは値IDを知りたい場合には、まず項目B610bの値ID別の値ID表631~632から、項目B610bの特定の値IDに対応する項目A610aの値IDを特定し、そこから項目A610aの値ID別の値ID表621~622を参照してこれに対応する項目C610cの値IDを特定することができる。

【0096】

項目C610cは基準列でも対象列でもなく、従って集合関数による処理の対象とはならないので、全ての値ID表でその対応関係を保持する必要はなく、ただ他のデータ項目の値との対応がわかるようにしておけばよいものである。従って、いずれか1つの基準列(必ずしも主キーである必要はない)についてのみ項目C610cとの対応がわかるよう

50

にしておけば、他の値との対応を辿ることが可能となる。

【0097】

(実施形態の拡張)

以上で説明した第1および第2の実施形態は、その趣旨を改変しない範囲で、様々な拡張が考えられる。

たとえば、作成された値ID表および値リストについて、1つの表を必ず1つのデータノードに記憶する必要はない。1つの値ID表もしくは値リストを、各データノードの記憶容量などのような制約に応じて、複数のデータノードに適宜分割して記憶してもよい。この場合には、データ配置情報管理部112が、どの値ID表もしくは値リストが、どのデータノードに分割して記憶されているかを把握してデータ配置情報121にその旨を記憶するようにすればよい。

10

【0098】

これまで本発明について図面に示した特定の実施形態をもって説明してきたが、本発明は図面に示した実施形態に限定されるものではなく、本発明の効果を奏する限り、これまで知られたいかなる構成であっても採用することができる。

【0099】

上述した各々の実施形態について、その新規な技術内容の要点をまとめると、以下のようになる。なお、上記実施形態の一部または全部は、新規な技術として以下のようにまとめられるが、本発明は必ずしもこれに限定されるものではない。

【0100】

(付記1) フロントデータベースサーバと複数台のデータノードとが相互に接続されて構成される分散メモリデータベースシステムであって、

20

前記フロントデータベースサーバが、

外部から入力される表データを分割して各々の実際のデータを値IDに置換した複数の値ID表を生成してこれらを前記各データノードに分散して記憶させるデータ構造変換部と、

外部のクライアントマシンから発行された集合関数を含むクエリーに基づいて前記各データノードに対して前記表データの中の特定の値IDの出現数を問い合わせると共に、これに応じて前記各データノードから返された特定の値IDの出現数から前記クエリーに対応する集合関数の値を計算して前記クライアントマシンに返送する問い合わせ処理部と

30

を有し、
前記データ構造変換部が、前記複数の値ID表を生成する際に、前記表データの中で集計軸になり得る列としてあらかじめ指定された複数のデータ項目の各々について個別に前記複数の値ID表を生成することを特徴とする分散メモリデータベースシステム。

【0101】

(付記2) 前記フロントデータベースサーバの前記データ構造変換部が、前記複数の値ID表を生成する際に、前記表データの中で集計軸および集計対象になり得る列としてあらかじめ指定されたデータ項目の各々について実際の値を値IDに置換すると共に、前記値IDと前記実際の値との対応を示す値リストを生成する機能を有し、

前記フロントデータベースサーバが、前記値リストおよび前記値ID表が前記各データノードの中のいずれに分散されたかを記憶するデータ配置情報管理部を有することを特徴とする、付記1に記載の分散メモリデータベースシステム。

40

【0102】

(付記3) 前記フロントデータベースサーバの前記データ構造変換部が、前記複数の値ID表を生成する際に、前記実際の値を大小順にソートしてから前記値IDに置換することを特徴とする、付記2に記載の分散メモリデータベースシステム。

【0103】

(付記4) 前記フロントデータベースサーバの前記データ構造変換部が、前記表データの中で集計軸および集計対象のいずれにもなり得ない列としてあらかじめ指定されたデータ項目の各々について、前記集計軸になり得る列についての前記値ID表のいずれか1種

50

類に付随して記憶させることを特徴とする、付記 2 に記載の分散メモリデータベースシステム。

【0104】

(付記 5) 複数台のデータノードと相互に接続されて分散メモリデータベースシステムを構成するフロントデータベースサーバであって、

外部から入力される表データを分割して各々の実際のデータを値 ID に置換した複数の値 ID 表を生成してこれらを前記各データノードに分散して記憶させるデータ構造変換部と、

外部のクライアントマシンから発行された集合関数を含むクエリーに基づいて前記各データノードに対して前記表データの中の特定の値 ID の出現数を問い合わせると共に、これに応じて前記各データノードから返された特定の値 ID の出現数から前記クエリーに対応する集合関数の値を計算して前記クライアントマシンに返送する問い合わせ処理部とを有し、

前記データ構造変換部が、前記複数の値 ID 表を生成する際に、前記表データの中で集計軸になり得る列としてあらかじめ指定された複数のデータ項目の各々について個別に前記複数の値 ID 表を生成することを特徴とするフロントデータベースサーバ。

【0105】

(付記 6) フロントデータベースサーバと複数台のデータノードとが相互に接続されて構成される分散メモリデータベースシステムにあって、

外部からの表データの入力を前記フロントデータベースサーバのデータ構造変換部が受け付け、

入力された前記表データを、集計軸になり得る列としてあらかじめ指定されたデータ項目の各々について前記フロントデータベースサーバのデータ構造変換部が個別に分割して実際のデータを値 ID に置換した複数の値 ID 表を生成し、

生成された前記複数の値 ID 表を前記フロントデータベースサーバのデータ構造変換部が前記各データノードに分散して記憶させ、

外部のクライアントマシンから発行された集合関数を含むクエリーを前記フロントデータベースサーバの問い合わせ処理部が受け付け、

受け付けられた前記クエリーに基づいて前記表データの中の特定の値 ID の出現数を前記フロントデータベースサーバの問い合わせ処理部が前記各データノードに問い合わせ、

前記各データノードから返された特定の値 ID の出現数から前記フロントデータベースサーバの問い合わせ処理部が前記クエリーに対応する集合関数の値を計算して前記クライアントマシンに返送する

ことを特徴とするデータ処理方法。

【0106】

(付記 7) フロントデータベースサーバと複数台のデータノードとが相互に接続されて構成される分散メモリデータベースシステムにあって、

前記フロントデータベースサーバが備えるコンピュータに、

外部からの表データの入力を受け付ける手順、

入力された前記表データを、集計軸になり得る列としてあらかじめ指定されたデータ項目の各々について個別に分割して実際のデータを値 ID に置換した複数の値 ID 表を生成する手順、

生成された前記複数の値 ID 表を前記各データノードに分散して記憶させる手順、

外部のクライアントマシンから発行された集合関数を含むクエリーを受け付ける手順、

受け付けられた前記クエリーに基づいて前記表データの中の特定の値 ID の出現数を前記各データノードに問い合わせる手順、

および前記各データノードから返された特定の値 ID の出現数から前記クエリーに対応する集合関数の値を計算して前記クライアントマシンに返送する手順を実行させることを特徴とするデータ処理プログラム。

【産業上の利用可能性】

10

20

30

40

50

【0107】

本発明はデータベースを利用するコンピュータシステム、特に分散メモリを使用するデータベースシステムに幅広く適用できる。

【符号の説明】

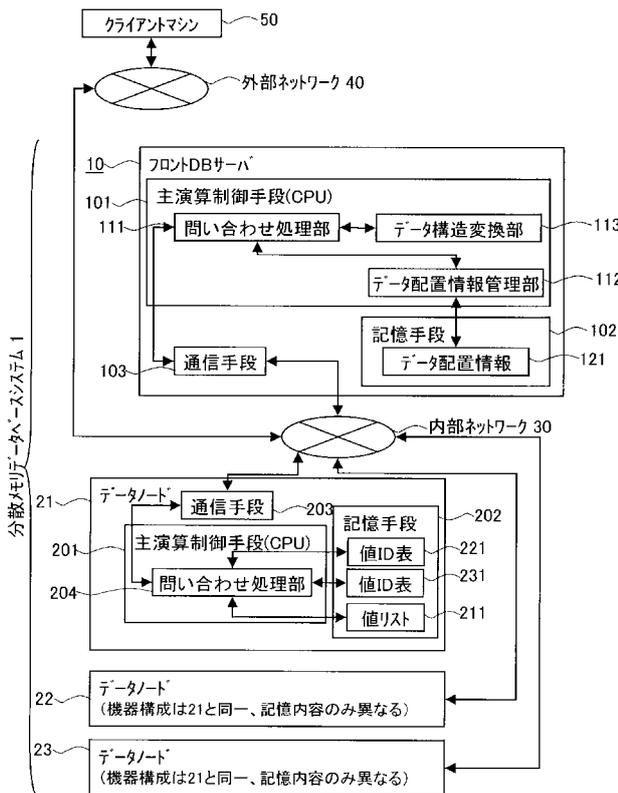
【0108】

- 1、501 分散メモリデータベースシステム
- 10、510 フロントDBサーバ
- 21～23、521～523 データノード
- 30 内部ネットワーク
- 40 外部ネットワーク
- 50 クライアントマシン
- 101、201 主演算制御手段
- 102、202 記憶手段
- 103、203 通信手段
- 111、204、511 問い合わせ処理部
- 112、513 データ配置情報管理部
- 113 データ構造変換部
- 121 データ配置情報
- 210、610 表データ
- 211～213 値リスト
- 221～222、231～33、621～622、631～632 値ID表

10

20

【図1】

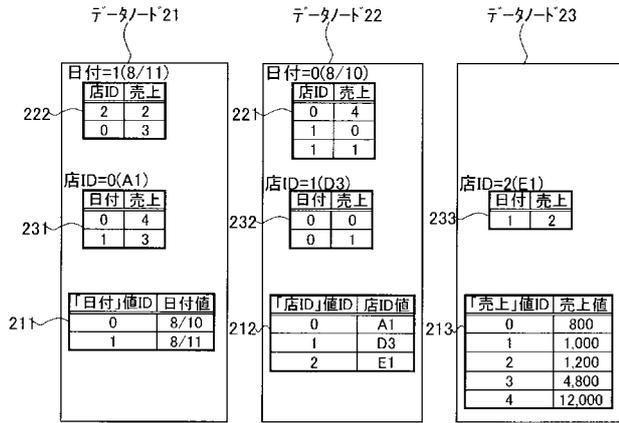


【図2】

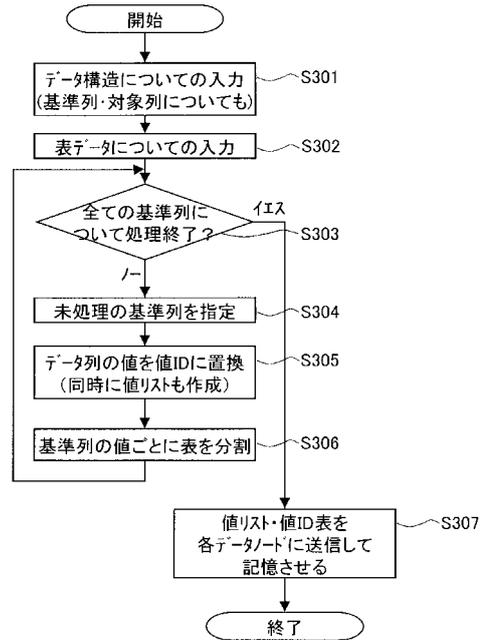
210a	210b	210c
日付	店ID	売上
8/10	A1	12,000
8/10	D3	800
8/10	D3	1,000
8/11	E1	1,200
8/11	A1	4,800

↑ 基準列 ↑ 基準列 ↑ 対象列

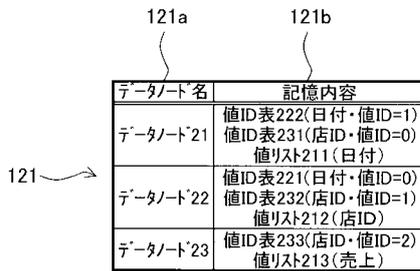
【 図 3 】



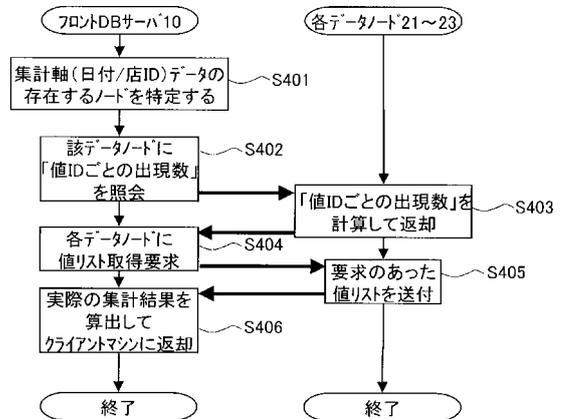
【 図 4 】



【 図 5 】



【 図 6 】



【 図 7 】

241 →

「売上」値ID	出現数
0	1
1	1
4	1

242 →

「売上」値ID	出現数
2	1
3	1

【 図 8 】

243 →

日付	売上合計
8/10	13,800
8/11	6,000

【 図 9 】

251 →

「売上」値ID	出現数
3	1
4	1

252 →

「売上」値ID	出現数
0	1
1	1

253 →

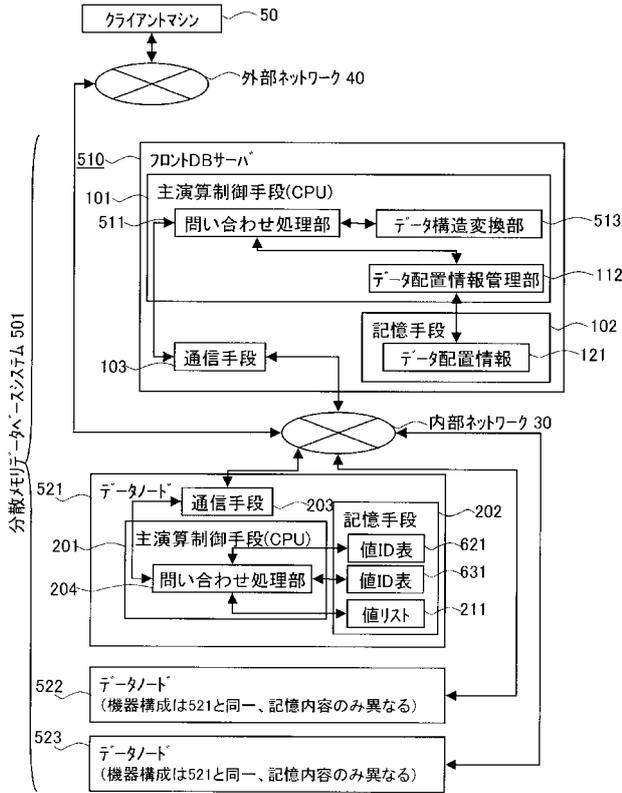
「売上」値ID	出現数
2	1

【 図 1 0 】

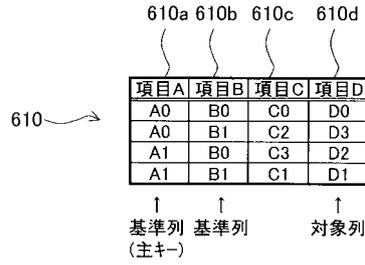
254 →

店ID	売上合計
A1	15,800
D3	1,800
E1	1,200

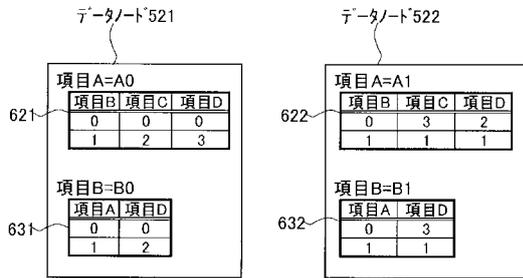
【図11】



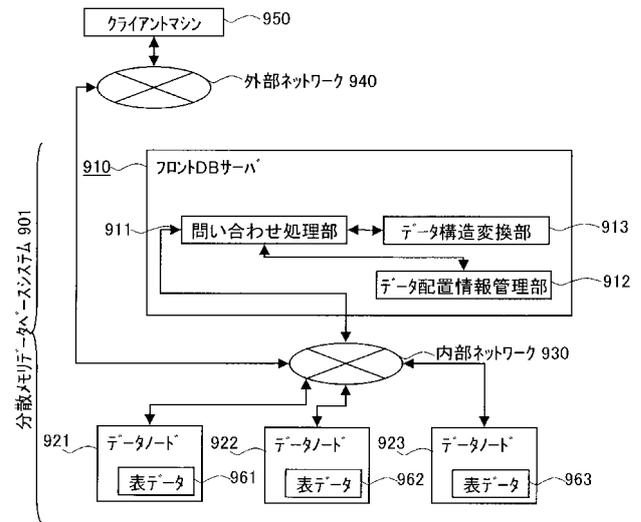
【図12】



【図13】



【図14】



【 図 1 5 】

910a 910b 910c

910

日付	店ID	売上
8/10	A1	12,000
8/10	D3	800
8/10	D3	1,000
8/11	E1	1,200
8/11	A1	4,800

【 図 1 6 】

