

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4804233号
(P4804233)

(45) 発行日 平成23年11月2日(2011.11.2)

(24) 登録日 平成23年8月19日(2011.8.19)

(51) Int. Cl. F I
G06F 17/30 (2006.01) G O 6 F 17/30 4 1 5
G06F 12/00 (2006.01) G O 6 F 17/30 3 4 O Z
 G O 6 F 12/00 5 1 3 Z

請求項の数 17 (全 38 頁)

(21) 出願番号	特願2006-161233 (P2006-161233)	(73) 特許権者	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22) 出願日	平成18年6月9日(2006.6.9)	(74) 代理人	100075513 弁理士 後藤 政喜
(65) 公開番号	特開2007-328716 (P2007-328716A)	(74) 代理人	100114236 弁理士 藤井 正弘
(43) 公開日	平成19年12月20日(2007.12.20)	(74) 代理人	100120260 弁理士 飯田 雅昭
審査請求日	平成21年1月5日(2009.1.5)	(72) 発明者	櫻山 俊彦 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内
		(72) 発明者	西澤 格 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内 最終頁に続く

(54) 【発明の名称】 ストリームデータ処理方法

(57) 【特許請求の範囲】

【請求項1】

計算機に入力されるストリームデータを1次情報として取得し、該取得した1次情報から予め設定した外部データの情報に基づいて2次情報を生成するストリームデータ処理方法において、

前記計算機にストリームデータの処理を記述するクエリを予め登録する処理と、

前記ストリームデータの値の特徴を表すストリームデータの特性情報を受け付ける処理と、

前記ストリームデータを監視することにより監視情報を取得する処理と、

前記計算機が前記クエリを参照して、前記外部データの一部または全部を前記計算機に
 予め設定したストリームプレディクトキャッシュへ予め読み込む処理と、

前記計算機がストリームプレディクトキャッシュへ読み込んだ外部データと、前記ストリームデータとを突き合わせることにより、前記2次情報を生成する処理と、
 を含み、

前記外部データを前記ストリームプレディクトキャッシュへ予め読み込む処理は、

前記ストリームデータの特性情報と、前記取得した監視情報とを用いて、前記ストリームプレディクトキャッシュへ予め読み込む前記外部データを決定することを特徴とするストリームデータ処理方法。

【請求項2】

請求項1に記載のストリームデータ処理方法において、

10

20

前記外部データをストリームプレディクトキャッシュへ読み込む処理は、
前記予め登録した前記クエリから、前記外部データに関する記述部分を抜き出して、前記ストリームプレディクトキャッシュへ外部データを取得するためのストリームプレディクトキャッシュデータ取得命令を生成する処理をさらに含み、
前記生成したストリームプレディクトキャッシュデータ取得命令を発行して前記外部データを読み込むことを特徴とするストリームデータ処理方法。

【請求項 3】

請求項 2 に記載のストリームデータ処理方法において、
前記予め登録された前記クエリから取得した突き合わせ対象のリレーション名と、前記ストリームプレディクトキャッシュにおける外部データの格納場所を決定するための条件を表す格納場所判別条件と、前記ストリームプレディクトキャッシュにおける外部データの格納場所と、前記ストリームプレディクトキャッシュの外部データに係る情報が更新された時刻情報を表すタイムスタンプと、をストリームプレディクトキャッシュ管理情報に保持する処理をさらに含むことを特徴とするストリームデータ処理システム。

10

【請求項 4】

請求項 3 に記載のストリームデータ処理方法において、
前記ストリームプレディクトキャッシュデータ取得命令を生成する処理は、
前記外部データに関する記述部分を抜き出し、かつ前記ストリームプレディクトキャッシュ管理情報を参照し、前記格納場所判別条件を付加することにより前記ストリームプレディクトキャッシュデータ取得命令を生成することを特徴とするストリームデータ処理方法。

20

【請求項 5】

請求項 4 に記載のストリームデータ処理方法において、
前記特性情報を入力する前記ストリームデータの名前を表すストリーム名と、前記特性情報を入力するカラムの名前を表すカラム名と、タイムスタンプに関する条件を表す時間に関する条件と、前記カラム名に対応する値の特性を表す値に関する条件カラムとを、ストリームデータ特性管理情報に保持する処理をさらに含むことを特徴とするストリームデータ処理方法。

【請求項 6】

請求項 5 に記載のストリームデータ処理方法において、
前記ストリームプレディクトキャッシュデータ取得命令を生成する処理は、
前記ストリームデータ特性管理情報を参照し、前記ストリームデータ特性管理情報のタイムスタンプカラムの時刻情報から、前記ストリームデータ特性管理情報の時間に関する条件カラムの条件を満たす前記ストリームデータ特性管理情報の値に関する条件カラムの条件を適用することを特徴とするストリームデータ処理方法。

30

【請求項 7】

請求項 5 に記載のストリームデータ処理方法において、
前記ストリームプレディクトキャッシュデータ取得命令を生成する処理は、
前記ストリームデータにおける時刻情報に対応する条件である時間付き制約情報を保持するストリームデータ特性管理情報を参照し、入力された前記ストリームデータが前記時間付き制約情報に違反していないか否かを判定することを特徴とするストリームデータ処理方法。

40

【請求項 8】

請求項 4 に記載のストリームデータ処理方法において、
前記監視情報を取得する処理は、
前記ストリームデータの監視対象のストリーム名と、監視対象の前記ストリームデータのカラム名と、前記監視情報を取得した時刻情報を表すタイムスタンプとを、監視情報管理部に保持することを特徴とするストリームデータ処理方法。

【請求項 9】

請求項 8 に記載のストリームデータ処理方法において、

50

前記監視情報管理部を参照し、前記監視情報から前記ストリームデータが必要と予測できる外部データの一部若しくは全部を取得するための条件である予測条件を算出する処理をさらに含むことを特徴とするストリームデータ処理方法。

【請求項 10】

請求項 9 に記載のストリームデータ処理方法において、

前記監視情報管理部を参照し、前記データレートが前記データレート閾値未満の場合は、前記外部データを取得しないことを特徴とするストリームデータ処理方法。

【請求項 11】

請求項 4 に記載のストリームデータ処理方法において、

前記特性情報を入力する前記ストリームデータの名前を表すストリーム名と、前記特性情報を入力するカラムの名前を表すカラム名と、タイムスタンプに関する条件を表す時間に関する条件と、前記カラム名に対応する値の特性を表す値に関する条件カラムとを、ストリームデータ特性管理情報に保持する処理と、

前記ストリームデータの監視対象のストリーム名と、監視対象の前記ストリームデータのカラム名と、前記監視情報を取得した時刻情報を表すタイムスタンプとを、監視情報管理部に保持する処理と、

前記ストリームデータ特性管理情報、及び前記監視情報管理部を参照し、該カラムがシークエンシャル属性を有しているか判定し、前記シークエンシャル属性の場合、前記監視情報管理部の現在の値カラムの値よりも小さな値をストリームプレディクトキャッシュから削除する処理と、

を含むことを特徴とするストリームデータ処理方法。

【請求項 12】

請求項 4 に記載のストリームデータ処理方法において、

予め登録された前記クエリと、前記クエリの名前と、前記クエリが使用可能なメモリ量の上限を表すメモリ上限値とを登録クエリ管理情報に保持する処理を含むことを特徴とするストリームデータ処理方法。

【請求項 13】

請求項 12 に記載のストリームデータ処理方法において、

前記ストリームプレディクトキャッシュ管理情報を参照し、メモリに格納される前記ストリームプレディクトキャッシュにおいて前記格納場所判別条件を満たすメモリ使用量を計算し、前記登録クエリ管理情報を参照して、前記ストリームプレディクトキャッシュのメモリ使用量が前記登録クエリ管理情報のメモリ上限値よりも大きいかなかを判定する処理を含むことを特徴とするストリームデータ処理方法。

【請求項 14】

請求項 13 に記載のストリームデータ処理方法において、

前記ストリームプレディクトキャッシュのメモリ使用量が前記登録クエリ管理情報のメモリ上限値よりも大きいかなかを判定し、前記メモリ上限値よりも大きいと判定した場合には、前記メモリ上限値に対応する外部データ量を計算し、前記外部データの各タブルの距離が小さいものから前記メモリ上限値に対応する前記外部データ量だけ格納することを特徴とするストリームデータ処理方法。

【請求項 15】

請求項 14 に記載のストリームデータ処理方法において、

前記ストリームプレディクトキャッシュ管理情報を参照し、前記ストリームプレディクトキャッシュ管理情報の格納場所判別条件カラムの条件から除外された前記ストリームプレディクトキャッシュの外部データを削除し、前記ストリームプレディクトキャッシュ管理情報の格納場所判別条件カラムの条件に新規に追加されたストリームプレディクトキャッシュの外部データを前記計算機のメモリまたはディスクに格納することを特徴とするストリームデータ処理方法。

【請求項 16】

請求項 15 に記載のストリームデータ処理方法において、

10

20

30

40

50

前記ストリームプレディクトキャッシュは、前記計算機のメモリ上に設定されたストリームプレディクトキャッシュオンメモリと、前記計算機のディスク上に設定されたストリームプレディクトキャッシュオンディスクと、を有し、

前記2次情報を生成する処理は、

前記ストリームプレディクトキャッシュ管理情報を参照し、格納場所判別条件から突き合わせを行う外部データが存在する前記ストリームプレディクトキャッシュの格納場所を判定する処理と、

前記判定結果がストリームプレディクトキャッシュオンメモリの場合には、ストリームプレディクトキャッシュオンメモリから該当する外部データを読み込む処理と、

前記判定結果がストリームプレディクトキャッシュオンディスクの場合には、ストリームプレディクトキャッシュオンディスクから該当する外部データを読み込む処理と、

前記判定結果がストリームプレディクトキャッシュオンメモリ及びストリームプレディクトキャッシュオンディスクの何れにも存在しない場合には、該当する外部データを読み込む処理と、

前記読み込んだ外部データと前記ストリームデータを突き合わせる処理と、を含むことを特徴とするストリームデータ処理方法。

【請求項17】

請求項4に記載のストリームデータ処理方法において、

前記格納場所管理部は、ストリームプレディクトキャッシュクリア命令を受け付け、前記ストリームプレディクトキャッシュに格納された外部データをクリアし、前記外部データを再取得する命令を発行することを特徴とするストリームデータ処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、リアルタイムストリームデータと外部データとの連携が可能なストリームデータ処理システムの構成方法、及び該ストリームデータ処理システムにおけるリアルタイムデータと外部データの連携方法に関する。

【背景技術】

【0002】

ストレージ装置に格納されたデータに対する処理を実行するデータベース管理システム（以下、DBMSとする）に対して、時々刻々と到着するデータをリアルタイム処理するデータ処理システムに対する要求が高まっている。例えば、株式の売買を行うシステムでは、株価の変動にいかにも迅速に反応できるかがシステムの最重要の課題の一つであり、従来のDBMSのように株式のデータを一旦記憶装置に格納してから、該格納データに関して検索を行うような方法では、株価変動のスピードに即応できず、ビジネスチャンスを逃してしまうことになりかねない。例えば、特許文献1では、記憶されているクエリが周期的に実行される機構を開示しているが、株価のようにデータが入ってきた瞬間にクエリを実行することが重要なリアルタイムデータ処理には適用が困難であった。

【0003】

このような時々刻々と到着するデータをストリームデータと定義し、該ストリームデータのリアルタイム処理に好適なデータ処理システムとして、ストリームデータ処理システムが提案されている。例えば、非特許文献1には、ストリームデータ処理システムSTREAMが開示されている。

【0004】

ストリームデータ処理システムでは、従来のDBMSとは異なり、まずクエリ（問い合わせ）をシステムに登録し、データの到来と共に該クエリが継続的に実行される。前記STREAMでは、ストリームデータを効率的に処理するために、ストリームデータの一部を切り取るスライディングウィンドウと呼ばれる概念を導入している。スライディングウィンドウ指定を含むクエリの記述言語の好適な例としては非特許文献1に開示されているCQL（Continuous Query Language）をあげることができる

10

20

30

40

50

。CQLは、DBMSで広く用いられているSQL (Structured Query Language) のFROM句に、ストリーム名に続いて括弧を用いることにより、スライディングウィンドウを指定する拡張が施されている。SQLに関しては、非特許文献2に開示されるものが知られている。上記スライディングウィンドウを指定する代表的な方法としては、(1) 切り取るデータ列の数を指定する方法、そして(2) 切り取るデータ列の時間間隔を指定する方法の2種類をあげることができる。例えば、非特許文献1の第二節に示された“Rows 50 Preceding”は50行分のデータを処理対象として切り取る(1)の好適な例、そして“Range 15 Minutes Preceding”は15分間分のデータを処理対象として切り取る(2)の好適な例である。スライディングウィンドウによって切り取られたストリームデータはメモリ上に保持され、クエリ処理に使用される。

10

【0005】

ストリームデータ処理システムは、ファイナンシャルアプリケーション、交通情報システム、計算機システム管理に代表される、リアルタイム処理が必要とされる応用に対する適用が期待されている。

【特許文献1】米国特許第5495600号公報

【非特許文献1】R. Motwani, J. Widom, A. Arasu, B. Babcock, S. Babu, M. Datar, G. Manku, C. Olston, J. Rosenstein, and R. Varma著: “Query Processing, Resource Management, and Approximation in a Data Stream Management System”, In Proc. of the 2003 Conf. on Innovative Data Systems Research (CIDR), January 2003

20

【非特許文献2】C. J. Date, Hugh Darwen著: “A Guide to SQL Standard (4th Edition)”, Addison-Wesley Professional; 4 edition (November 8, 1996), ISBN: 0201964260

【非特許文献3】Stratis Viglas, Jeffrey F. Naughton著: “Rate-based query optimization for streaming information sources”, In Proc. of SIGMOD Conference 2002, pp. 37-48

30

【発明の開示】

【発明が解決しようとする課題】

【0006】

時々刻々と到着するストリームデータを利用するユーザのために、ストリームデータを有意な情報に変換する必要が生じる場合がある。例えば、ストリームデータが製品の製造工程において、製品に付いているRFIDタグをRFIDリーダーで読み込んだデータ、あるいはRFIDリーダーを管理する製造管理サーバが出力するデータである場合、ストリームデータには製品のコードと製造工程の情報などが含まれている。製品のコードは、ユーザにとって理解しにくい情報であるため、ストリームデータ処理システムがデータベースに問い合わせを行って製品のコードから製品名などの有意な情報を取得し、この有意な情報をユーザへ提供する必要がある。あるいは、ストリームデータが、センサノードが出力するデータ、あるいはセンサネットを管理するセンサネットサーバが出力するデータである場合、ストリームデータとしてセンサノードのIDとセンシングデータが含まれている。センサノードのIDは、ユーザにとって理解しにくい情報であるため、ストリームデータ処理システムがデータベースに問い合わせを行ってセンサノードの所在と種類などの有意な情報を取得し、この有意な情報をユーザへ提供する必要がある。

40

【0007】

上記従来例のストリームデータ処理システムでは、上述のようなストリームデータの変

50

換処理（ストリームデータ処理）をストリームデータの受信のたびに実行することになり、次のような問題が生じる。

【0008】

（1）データベース（以下、外部データという）へのアクセスは、処理に時間がかかるため、ストリームデータ処理システムが膨大な量のストリームデータを処理する場合には、データベースへのアクセスを行う時間がボトルネックとなる。この結果、ストリームデータの量が多くなると、リアルタイムでユーザに有意な情報を提供することができない、という問題がある。

【0009】

（2）ストリームデータ処理システム上に外部データの複製を保存しておきストリームデータを有意な情報に変換する処理を高速化する手法が考えられるが、外部データの複製全部をストリームデータ処理システム上に保存することは、データベースの複製を保存するための計算機リソースが必要になって非効率であり、また、データベースの容量が大きい場合にはストリームデータ処理システムに複製を保存することが不可能であることも考えられる。このため、有意な情報をリアルタイムで提供する処理性能を保ちつつ、外部データの複製量を減少させることも課題となっていた。

【0010】

（3）外部データが更新された場合にも、ストリームデータ処理システム上の外部データの複製に対し更新を反映し、ストリームデータ処理を行えることも課題となっていた。

【0011】

本発明は上記問題点に鑑みてなされたものであり、本発明の第1の目的は、外部データと連携し、かつリアルタイムにストリームデータ処理を行うシステムを提供することである。また、本発明の第2の目的は、リアルタイム処理を保ちつつ、外部データの複製量を減少させるシステムを提供することである。最後に、本発明の第3の目的は、外部データが変更された場合にも、ストリームデータ処理システム上の外部データの複製に対し更新を反映し、ストリームデータ処理を正確に行えるシステムを提供することである。

【課題を解決するための手段】

【0012】

本発明は、前記第1の目的を達成するため、本発明では必要に応じて外部データの一部、若しくは全部を記憶装置に設けたストリームプレディクトキャッシュに複製し、前記複製を用いてストリームデータと連携する機構を設ける。次に、第2の目的を達成するため、本発明ではユーザ、若しくはアプリケーションによってシステムに登録されたストリームデータの該属性値における時間と値域の組合せによる2次元の条件を記述したストリームデータ特性情報を用いて外部データの取得対象を絞り込むことにより、リアルタイム処理を保ちつつ、外部データの複製量を減少させる機構を提供する。

【0013】

さらに、第2の目的を達成するための別の手段として、本発明では前記ストリームデータ特性情報、及びストリームデータの監視を行うことにより取得した監視情報を用いることにより、ストリームデータと連携する必要がある外部データの取得対象を予測する機構を提供する。

【0014】

最後に、第3の目的を達成するため、本発明ではユーザ、若しくはアプリケーションによる外部データの複製の更新命令を受け付け、ストリームデータ処理システム上の外部データの複製に対し更新を反映し、ストリームデータ処理を行える機構を提供する。また、第3の目的を達成する別の手段として、外部データの変更を検知し、変更時間を付与したデータを生成し、ストリームデータ処理システムに送信することにより、ストリームデータ処理システム上の外部データの複製に対し更新を反映し、ストリームデータ処理を行える機構を提供する。

【発明の効果】

【0015】

10

20

30

40

50

本発明を用いることにより、高信頼・高可用のリアルタイムデータ処理が可能なストリームデータ処理システムが実現できる。

【発明を実施するための最良の形態】

【0016】

以下、本発明の実施の形態について説明する。

【0017】

図1は本発明の一実施形態が適用されたストリームデータ処理システム、および関連するシステム構成の概略図である。

【0018】

図1において、記憶装置104を有するデータ管理システム101は、ネットワーク103を介してストリームデータ処理システム102に接続されている。ネットワーク103は、イーサネット（登録商標）、光ファイバ、FDDI（Fiber Distributed Data Interface）で接続されるローカルエリアネットワーク（LAN）、若しくはLANよりも低速なインターネットを含んだワイドエリアネットワーク（WAN）でも差し支えない。また、データ管理システム101、及びストリームデータ処理システム102はパーソナルコンピュータ、ワークステーションなどの任意のコンピュータシステムで構成される。すなわち、データ管理システム101は、データベース（データ114）の管理を行うDBMS（DataBase Management System）を実行する計算機であり、ストリームデータ処理システム102はストリームデータ111を有意な情報に変換してクライアント計算機116へ提供するストリームデータ処理を実行する計算機（またはサーバ）である。

【0019】

そして、ネットワーク103には、ストリームデータをストリームデータ処理システム102へ出力する計算機12が接続される。計算機12は、工場などで生産される製品の管理を行うものである。本実施形態の計算機12では、製品が製造された製造ラインの識別子と製品の識別子と製造された日時を対にした情報をストリームデータとして出力する。また、計算機12が管理する工場では複数の製造ラインで複数種の製品が製造されるものとする。そして、計算機12は、各製造ラインで生産された製品の情報を、RFIDリーダーやバーコードリーダーなどを介して取得し、複数の製造ラインで得られた製品の情報をリアルタイムでストリームデータ処理システム102へ送信する。なお、計算機12では、ストリームデータ111を出力するアプリケーション13が実行される。また、ストリームデータ処理システム102は、図示のようにRFIDリーダー10やセンサノード11が接続され、ストリームデータ111を受信するようにしても良い。

【0020】

ここで、本実施形態で扱うストリームデータは、映像や音声の配信で用いられるストリームとは異なり、ひとつのストリームデータが有意な情報に対応するものである。また、ストリームデータ処理システム102が計算機12からストリームデータ処理システム102を受信するストリームデータは、連続的あるいは間欠的であり、ストリームデータ毎に異なる製品の情報や異なる要素が含まれる。

【0021】

ストリームデータ処理システム102は、クエリ管理部105、ストリームデータ監視情報管理部106、格納場所管理部107、ストリームデータ特性管理部108、記憶装置109、突き合わせ部112から構成される。

【0022】

クエリ管理部105は、ユーザ115の操作によって、クライアント計算機116上で実行されるアプリケーションにより登録されたクエリを管理している。次に、ストリームデータ監視情報管理部106は、前記ストリームデータ処理システム102に入力されるストリームデータ111を監視し、監視情報（ストリームデータ111の統計情報など）を取得し、管理している。次に、ストリームデータ特性管理部108は、ユーザ115、またはクライアント計算機116上で実行されるアプリケーション117により登録された処理時刻以降のストリームデータの値の特徴であるストリームデータの特性情報を管理

10

20

30

40

50

している。

【0023】

格納場所管理部107は、前記クエリ管理部105で管理されている登録クエリ情報、及び前記ストリームデータ監視情報管理部106で管理されているストリームデータ監視情報、前記ストリームデータ特性管理部108で管理されているストリームデータ特性情報の何れか、または両方を用いて、前記データ管理システム101上にある記憶装置104のデータ114のうち、前記ストリームデータ処理システム102に取得するデータを決定する。以下では、ストリームデータ処理システム102がデータ管理システム101から取得したデータを格納する記憶装置109の領域をストリームプレディクトキャッシュ(以下、SPC: Stream Predict Cache)110と呼ぶ。そして、ストリームデータ処理システム102は、前記データ管理システム101に対し、ストリームプレディクトキャッシュ取得命令(SPC取得命令)を発行する。

10

【0024】

前記データ管理システム101は、前記格納場所管理部107において発行されたSPC取得命令を入力とし、記憶装置104中のデータ114において、一部または全部を前記ストリームデータ処理システム102に出力する。前記格納場所管理部107は、取得データを前記SPC110(図2のSPC_OM217及びSPC_OD219)として、記憶装置109上に格納する。記憶装置109は、メモリ、ディスク装置など任意の記憶媒体とする。

【0025】

突き合わせ部112は、前記ストリームデータ111、および前記SPC110の突き合わせを行い、有意な情報としての突き合わせ結果113をクライアント計算機116へ出力する。

20

【0026】

図1において、本発明の概要を説明する。ストリームデータ処理システム102は、ユーザ115が登録したクエリとストリームデータ111の特性に基づいて、入力されるストリームデータ111と、データ管理システム101のデータ114の突き合わせを行う。

【0027】

ユーザ115は計算機116のアプリケーション117により、クエリ管理部105にクエリを登録し、ストリームデータ111の特性をストリームデータ特性管理部108に登録しておく。ここで、ストリームデータの特性とは、例えば、ストリームデータのIDの値は1ずつ増加する、という特性や、ある属性の値は一定範囲内であるという特性や、ある時間帯に到着するストリームデータのある属性の値は一定範囲内である、といった特性を指す。

30

【0028】

ストリームデータ監視情報管理部106は、ストリームデータ処理システム102が受信したストリームデータ111を監視し、監視によって得られたストリームデータ111の傾向に関する情報を監視情報として保持する。ストリームデータ111の傾向に関する情報は、例えば、ストリームデータ111のデータレートや、ある属性の値の分布などの統計情報である。格納場所管理部107は、ストリームデータ特性管理部108に設定されたストリームデータ111の特性と、ストリームデータ監視情報管理部106に保持された傾向に関する情報を参照し、突き合わせを行うためにSPC110に格納すべきデータを決定する。SPC110には、格納場所管理部107が決定したストリームデータ処理システム102のデータ114を予め読み出して、格納しておく。

40

【0029】

突き合わせ部112は、SPC110を参照することで、受信したストリームデータ111と予め先読みしておいたデータ114とを突き合わせることができ、リアルタイムで突き合わせ結果113をユーザ115(計算機116)へ提供することができる。

【0030】

50

すなわち、本発明では、ユーザ 115 によって入力されたストリームデータ 111 の特性と、ストリームデータ監視情報管理部 106 でモニタした統計情報から、突き合わせを行うために SPC 110 へ先読みしておくデータを格納場所管理部 107 で決定する。これにより、突き合わせに必要なデータ 114 の量を最小限としながら、外部データ（データ管理システム 101 のデータ 114）と連携し、かつリアルタイムにストリームデータ処理をすることが可能となるのである。

【0031】

図 2 は、本発明を適用した一実施形態のストリームデータ処理システム、および関連するシステムの構成を示すブロック図である。

【0032】

前記ストリームデータ処理システム 102 において、前記クエリ管理部 105 は、クエリ設定部 203、メモリ上限設定部 204、登録クエリ管理テーブル 205 から構成される。

【0033】

クエリ設定部 203 は、ユーザ 115、またはクライアント計算機 116 上で実行されるアプリケーション 117 からコマンド入力部 229 を通してクエリの登録・変更を行うインターフェース（以下、I/F）を備える。クエリの登録・変更が行われた場合、登録クエリ管理テーブル 205 を更新し、SPC 管理部 212 に変更通知を送信する。

【0034】

メモリ上限設定部 204 は、前記ユーザ 115、または前記クライアント計算機 116 上で実行される前記アプリケーション 117 から前記コマンド入力部 229 を通して該クエリで利用可能なメモリ上限値を設定する I/F を備える。メモリ上限値の入力が行われた場合、登録クエリ管理テーブル 205 を更新し、SPC 管理部 212 に変更通知を送信する。

【0035】

登録クエリ管理テーブル 205 は、前記クエリ設定部 203、及び前記メモリ上限設定部 204 において設定された情報を保持するテーブルである。

【0036】

前記ストリームデータ監視情報管理部 106 は、ストリームデータ監視部 207、データレート閾値・更新頻度設定部 206、監視情報管理テーブル 208 から構成される。

【0037】

ストリームデータ監視部 207 は、前記登録クエリ管理テーブル 205 を参照し、監視が必要なストリームデータ 111 を監視する。ストリームデータ監視部 207 は、監視の結果、監視情報（ストリームデータ 111 の統計情報）を取得した場合、監視情報管理テーブル 208 を更新し、SPC 管理部 212 に変更通知を送信する。ここで、監視するストリームデータ 111 は必要なものに限定せず、すべてのストリームデータの監視を行ってもよい。

【0038】

データレート閾値・更新頻度設定部 206 は、前記監視が必要なストリームデータ 111 に対し、データレートによりリレーションデータ 226 をキャッシュするか判断する閾値を設定し、及び監視情報を更新する頻度（間隔）を登録・変更する I/F を備える。データレート閾値、及び更新頻度が登録・変更された場合、データレート閾値・更新頻度設定部 206 は、監視情報管理テーブル 208 を更新し、SPC 管理部 212 に変更通知を送信する。なお、図 2 において、データ管理システム 101 のリレーションデータ 226 と XML データ 227 が図 1 のデータ 114 に相当する。

【0039】

監視情報管理テーブル 208 は、前記データレート閾値・更新頻度設定部 206、及びストリームデータ監視部 207 において設定または取得した情報を保持するテーブルである。

【0040】

前記ストリームデータ特性管理部 108 は、ストリームデータ特性設定部 214、時間付き制約設定部 213、ストリームデータ特性管理テーブル 215 から構成される。

【0041】

ストリームデータ特性設定部 214 は、前記ユーザ 115、または前記クライアント計算機 116 上で実行される前記アプリケーション 117 から前記コマンド入力部 229 を通じて該ストリームデータ 111 の特性情報を入力する I/F を備える。ストリームデータ 111 の特性情報の入力が行われた場合、ストリームデータ特性管理テーブル 215 を更新し、SPC 管理部 212 に変更通知を送信する。

【0042】

時間付き制約設定部 213 は、ストリームデータ 111 の該当カラムにおける時刻情報に対応する条件である時間付き制約を入力する I/F を備える。時間付き制約条件の入力が行われた場合、時間付き制約設定部 213 はストリームデータ特性管理テーブル 215 を更新し、SPC 管理部 212、及び時間付き制約違反判定部 221 に変更通知を送信する。

10

【0043】

ストリームデータ特性管理テーブル 215 は、前記ストリームデータ特性設定部 214、及び前記時間付き制約設定部 213 において設定された情報を保持するテーブルである。

【0044】

前記格納場所管理部 107 は、SPC 管理部 212、SPC 管理テーブル 211、SPC 取得命令発行部 209、SPC クリア命令発行部 210、SPC 情報設定部 228 から構成される。

20

【0045】

SPC 取得命令発行部 209 は、前記 SPC 管理部 212 から SPC 取得要求を受信した場合、前記データ管理システム 101 に対し、前記ストリームデータ 111 が必要とする外部データ（データ管理システム 101 のリレーションデータ 226）の一部若しくは全部である SPC データを取得する SPC 取得命令を発行する。なお、SPC データは、記憶装置 109 上の SPC_OM 217 または SPC_OD 219 に格納（キャッシュ）されているデータを示す。なお、記憶装置 109 は、メモリ 216 とディスク 218 で構成した例を示す。また、SPC 110 は、アクセス速度は速いが容量の小さいメモリ 216 上の SPC_OM 217 と、アクセス速度は遅いが容量の大きいディスク 218 上の SPC_OD 219 により階層的に構成され、優先度の高いデータを SPC_OM 217 に格納し、優先度の低いデータを SPC_OD 219 に格納する。

30

【0046】

SPC クリア命令発行部 210 は、SPC_OM 217 または SPC_OD 219 にキャッシュされたデータの元であるリレーションデータ 226 がデータ管理システム 101 で更新された場合、前記ユーザ 115、または前記クライアント計算機 116 上で実行される前記アプリケーション 117 から前記コマンド入力部 229 を通じて、または一定間隔により SPC クリア命令を発行する I/F を備え、SPC クリア命令が発行された場合、SPC 管理部 212 に SPC クリア要求を出力する。

40

【0047】

SPC 情報設定部 228 は、前記ユーザ 115、または前記クライアント計算機 116 上で実行される前記アプリケーション 117 から前記コマンド入力部 229 を通じて SPC 情報を入力する I/F を備える。SPC 情報の入力が行われた場合、SPC 情報設定部 228 は、SPC 管理テーブル 211 を更新し、SPC 管理部 212 に変更通知を送信する。

【0048】

SPC 管理部 212 は、前記クエリ設定部 203、メモリ上限設定部 204、データレート閾値・更新頻度設定部 206、ストリームデータ監視部 207、ストリームデータ特性設定部 214、時間付き制約設定部 213、SPC 情報設定部 228 から変更通知を受

50

信した場合に、前記登録クエリ管理テーブル205、及び前記監視情報管理テーブル208、前記ストリームデータ特性管理テーブル215の何れか、または両方を参照し、前記データ管理システム101上にあるリレーションデータ226のうち、前記ストリームデータ処理システム102に読み込むSPCデータを決定する。そして、SPC管理部212は、SPCデータの格納場所を決定し、SPC管理テーブル211を更新する。また、SPC管理部212は、前記変更通知を受信した場合、及びSPCクリア命令発行部210からSPCクリア要求を受信した場合、SPC取得命令発行部209にSPC取得要求を出力する。また、SPC管理部212は、前記データ管理システム101から受信したSPCデータに対し、前記SPC管理テーブル211を参照し、記憶装置109におけるメモリ216上のストリームプレディクトキャッシュオンメモリ(以下、SPC_OM: SPC_OnMemory)217、及びディスク218上のストリームプレディクトキャッシュオンディスク(以下、SPC_OD: SPC_OnDisk)219に格納する。なお、2つのSPC_OM217、SPC_OD219は、リアルタイム処理の必要性が高いSPCデータをアクセス速度の速いSPC_OM217に格納し、データレートが低いストリームデータ111に対しては、アクセス速度の遅いSPC_OD219にSPCデータを格納するよう使い分けることができる。

10

【0049】

また、SPC管理部212中にデータ変換部230があり、データ変換部230は、前記記憶装置104中のデータがXMLデータ227などのリレーションデータでない場合に、レコード形式に変換する機構を備える。これにより、XMLデータ227もリレーションデータ226と同様に利用することを可能にする。

20

【0050】

SPC管理テーブル211は前記SPC管理部212で決定した格納場所情報を保持するテーブルである。

【0051】

前記データ管理システム101は前記記憶装置104、データ変更検知部201、タプル生成部202から構成される。

【0052】

記憶装置104は、リレーションデータ226、XMLデータ227の何れか、または両方から構成される。前記リレーションデータ226は、複数のリレーションデータ226₁、226₂、・・・、226_nから構成され、前記XMLデータ227は、複数のXMLデータ227₁、227₂、・・・、227_nから構成される。ここで、記憶装置104に格納されるデータは、上記で述べたデータの他にCSVファイル等のいかなるデータ形式でもよい。以下では、リレーションデータ226を用いる例を説明する。

30

【0053】

データ変更検知部201は、前記リレーションデータ226の変更を監視し、変更を検知した場合、タプル生成部202にタプル生成要求を出力する。

【0054】

タプル生成部202は、前記データ変更検知部201からタプル生成要求を受信した場合、変更となったリレーションデータに対し、ストリーム化されたタプルを生成し、前記SPC管理部212に送信する。なお、タプルとはリレーションデータ226のレコードに相当するものであり、ストリームデータ111と突き合わせを行うためのデータ(属性値)を示す。

40

【0055】

ここで、前記SPCクリア命令発行部210がある場合、前記データ変更検知部201、及び前記タプル生成部202は必ずしも必要でない。一方、前記データ変更検知部201、及び前記タプル生成部202がある場合、前記SPCクリア命令発行部210は必ずしも必要でない。

【0056】

前記記憶装置109は、メモリ216、ディスク218から構成される。

50

【 0 0 5 7 】

前記メモリ 2 1 6 中に S P C _ O M 2 1 7 があり、前記 S P C 管理部 2 1 2 から送信された S P C データを保持する。また、前記突き合わせ部 1 1 2 に前記 S P C _ O M 2 1 7 のデータを入力する。

【 0 0 5 8 】

前記ディスク 2 1 8 中に S P C _ O D 2 1 9 があり、前記 S P C 管理部 2 1 2 から送信された S P C データを保持する。

【 0 0 5 9 】

また、前記 S P C _ O M 2 1 7、及び前記 S P C _ O D 2 1 9 は、前記 S P C 管理部 2 1 2 からのデータ移動要求を受け、前記メモリ 2 1 6、前記ディスク 2 1 8 間のデータを移動する。本実施例では、前記メモリ 2 1 6、及び前記ディスク 2 1 8 の階層構造としたが、メモリ 2 1 6 のみの構成でもよい。

10

【 0 0 6 0 】

ストリームデータ 1 1 1 は、複数のストリームデータ 1 1 1₁、1 1 1₂、・・・、1 1 1_n から構成される。

【 0 0 6 1 】

前記突き合わせ部 1 1 2 は、突き合わせ実行部 2 2 0、時間付き制約違反判定部 2 2 1 から構成される。

【 0 0 6 2 】

突き合わせ実行部 2 2 0 は、前記メモリ 2 1 6 中の前記 S P C _ O M 2 1 7 のデータを取得し、前記ストリームデータ 1 1 1 との突き合わせを行い、突き合わせ結果 1 1 3 を出力する。

20

【 0 0 6 3 】

また、突き合わせ実行部 2 2 0 中にデータ変換部 2 3 0 があり、データ変換部 2 3 0 は、前記記憶装置 1 0 4 中のデータが X M L データ 2 2 7 などのリレーションデータでない場合に、また前記ストリームデータが X M L 形式などのレコード形式でない場合に、レコード形式に変換する機構を備える。

【 0 0 6 4 】

時間付き制約違反判定部 2 2 1 は、前記ストリームデータ特性管理テーブル 2 1 5 に保持されている時間付き制約条件を参照し、前記ストリームデータ 1 1 1 が条件に違反しているか否かを判定する。条件に違反している場合、ログ 2 2 2 に違反情報を出力する。なお、ログ 2 2 2 は、例えばクライアント計算機 1 1 6 に対して送信すればよい。また、データ管理システム 1 0 1 の記憶装置 1 0 4 等にログ 2 2 を格納してもよい。

30

【 0 0 6 5 】

図 3 は、ストリームデータ 1 1 1 の好適なデータフォーマットの例を模式的に表した図である。

【 0 0 6 6 】

ストリームデータ 1 1 1 はレコード形式であり、レコードを構成する製品 I D カラム 3 0 1、ライン名カラム 3 0 2 がセグメントに相当し、前記製品 I D カラム 3 0 1、及び前記ライン名カラム 3 0 2 の組み合わせをタプル 3 0 3 とする。前記ストリームデータ 1 1 1 が前記ストリームデータ処理システムに入力された場合、到着した時間を示すタイムスタンプが付加される。なお、ストリームデータソースにおいて、タイムスタンプを付加してもよい。

40

【 0 0 6 7 】

図 4 は、連続して入力されたストリームデータ 1 1 1 を例示した図であり、製品ストリーム (S 1) 1 1 1₁ を表している。

【 0 0 6 8 】

製品ストリーム 1 1 1₁ において、表の 1 行が前記タプル 3 0 3 に対応し、各タプルの到着時間であるタイムスタンプカラム 3 0 4 が付加されている。例えば、タプル 3 0 4 は、タイムスタンプカラム 3 0 4 の値が「 1 0 : 0 0 」、製品 I D カラム 3 0 1 の値が「 1

50

000」、ライン名カラム302の値が「ラインA」であるタプルであることを表している。

【0069】

ここで、本実施例では、タイムスタンプを「10:00」のような時間、及び分の形式で表しているが、「2006/2/3 9:00:00 JST」のような日付、及び秒を含めた形式に代表される他の形式でもよい。

【0070】

図5は、リレーションデータ226を例示した図であり、製品管理表(R1)226₁を表している。

【0071】

製品管理表は、製品IDカラム401、製品名カラム402、価格カラム403から構成される。製品管理表の1行がレコードに対応する。

【0072】

例えば、レコード404は、製品IDカラム401の値が「1000」、製品名カラム402の値が「製品A」、価格カラム403の値が「5000」であることを表している。

【0073】

図6は、リレーションデータ226を例示した図であり、注文表(R2)226₂を表している。

【0074】

注文表は、注文IDカラム501、製品IDカラム502、注文者名カラム503から構成される。注文表の1行がレコードに対応する。

【0075】

例えば、レコード504は、注文IDカラム501の値が「10000」、製品IDカラム502の値が「1000」、注文者名カラム503の値が「注文者A」であることを表している。

【0076】

図7は、クエリを前記ストリームデータ処理システム102に登録する際の好適なコマンドの記述例である。

【0077】

登録クエリ1201は、前記ユーザ115、または前記クライアント計算機116上で実行される前記アプリケーション117から前記コマンド入力部229を通して、前記クエリ設定部203において登録される。登録クエリ1201は、前記製品ストリーム(S1)111₁の過去50行([Rows 50])、前記製品管理表(R1)226₁、前記注文表(R2)226₂において、前記製品ストリーム111₁の製品IDカラム301、前記製品管理表226₁の製品IDカラム402、前記注文表226₂の製品IDカラム502が同一の値であり、かつ前記注文表226₂の注文者カラム503が「注文者A」であるもののうち、注文IDカラム501、製品IDカラム(301または402または502)、製品名カラム402、ライン名カラム302、注文者名カラム503を抜き出すクエリであることを表している。

【0078】

以下では、ストリームデータ処理システム102で実行される突き合わせ処理の処理手順を、図8のフローチャートに示し、図2を参照しながら以下に説明する。

【0079】

ストリームデータ処理システム102で実行する突き合わせ処理では、まず、前記SPC管理部212において、前記クエリ設定部203から発行される変更通知により、前記ユーザ115、または前記クライアント計算機116上で実行される前記アプリケーション117から前記コマンド入力部229を通して登録されたクエリの登録、または変更を受け付けたか否かを判定する(S2002)。

【0080】

10

20

30

40

50

前記ステップS 2 0 0 2でy e sと判定された場合、後述するクエリ登録・変更時の処理を行う(S 2 0 0 3)。前記ステップS 2 0 0 3が終了した後、または、前記ステップS 2 0 0 2でn oと判定された場合、前記S P C管理部2 1 2において、前記メモリ上限設定部2 0 4、前記データレート閾値・更新頻度設定部2 0 6、前記ストリームデータ監視部2 0 7、前記ストリームデータ特性設定部2 1 4、前記時間付き制約設定部2 1 3から発行される変更通知により、監視情報、及び設定情報に変更があるか否かを判定する(S 2 0 0 4)。

【0081】

前記ステップS 2 0 0 4でy e sと判定された場合、後述する監視情報、及び設定情報変更時の処理を行う(S 2 0 0 5)。前記ステップS 2 0 0 5が終了した後、または、前記ステップS 2 0 0 4でn oと判定された場合、後述するリレーションデータ更新時の処理を実行し(S 2 0 0 6)、前記突き合わせ実行部2 2 0において、前記ストリームデータ1 1 1が到着したか否かを判定する(S 2 0 0 7)。

【0082】

前記ステップS 2 0 0 7でy e sと判定された場合、前記突き合わせ実行部2 2 0において、後述するように受信したストリームデータ1 1 1とS P C 1 1 0のデータとを突き合わせる突き合わせ処理を行う(S 2 0 0 8)。前記ステップS 2 0 0 8が終了した後、または前記ステップS 2 0 0 7でn oと判定された場合、前記コマンド入力部2 2 9において、システム終了コマンドが投入されたかにより、システムを終了するか否かを判定する(S 2 0 0 9)。

【0083】

前記ステップS 2 0 0 9でy e sと判定された場合、処理を終了し(S 2 0 1 0)、前記ステップS 2 0 0 9でn oと判定された場合、前記ステップS 2 0 0 2に戻り、上述の処理を継続する。

【0084】

以上の処理がストリームデータ処理システム1 0 2で実行される処理のメインルーチンであり、クエリや監視情報、リレーションデータ2 2 6やS P Cデータの更新などを行いながらストリームデータ1 1 1を受け付けて突き合わせ処理を実施する。

【0085】

次に、前記図8に示したクエリ登録・変更時の処理S 2 0 0 3の処理手順を、図9のフローチャートに示し、図2、図10、図11a~図11c、図12を用いて説明する。

【0086】

クエリ登録・変更時の処理では、まず、前記メモリ上限設定部2 0 4において、登録・変更されたクエリのメモリ上限値を受け付け(S 2 1 0 2)、クエリの登録・変更情報、及びメモリ上限値を前記登録クエリ管理テーブル2 0 5に反映させる(S 2 1 0 3)。

【0087】

次に、前記ストリームデータ監視部2 0 7において、前記登録クエリ管理テーブル2 0 5を参照し、登録されたクエリに対し、監視が必要なストリームデータの監視情報を取得し(S 2 1 0 4)、監視が必要なストリームデータの監視の更新間隔を設定し(S 2 1 0 5)、取得したストリームデータ監視情報を前記監視情報管理テーブル2 0 8に反映させる(S 2 1 0 6)。

【0088】

次に、前記ストリームデータ特性設定部2 1 4において、ストリームデータ特性情報を受け付け(S 2 1 0 7)、前記ストリームデータ特性管理テーブル2 1 5に反映させる(S 2 1 0 8)。

【0089】

次に、後述するS P C管理テーブル2 1 1の更新処理を行い(S 2 1 0 9)、後述する突き合わせ処理時のためにリレーションデータ2 2 6をS P C 1 1 0に読み込むS P C取得処理を行う(S 2 1 1 0)。

【0090】

最後に、前記突き合わせ部 112 において、前記ストリームデータ 111 の受付を開始し (S2111)、クエリ登録・変更時の処理を終了する (S2112)。

【0091】

以上の処理により、ユーザ 115 がクライアント計算機 116 から入力したクエリやメモリ上限設定値、更新頻度、ストリームデータ 111 の特性がストリームデータ処理システム 102 の各テーブルに設定される。

【0092】

図 10 は、登録クエリ管理テーブル 205 の構成例を表した図である。

【0093】

クエリ名カラム 601 には、登録されたクエリの名前を格納する。登録クエリカラム 602 には、登録されたクエリを格納する。

【0094】

メモリ上限値カラム 603 には、前記メモリ上限設定部 204 において設定されたメモリ使用量の上限値を格納する。クエリ実行形式格納先カラム 604 には、クエリの実行形式の格納先のアドレス (ポインタ) を格納する。実行形式をあらかじめ作成しておくことにより、クエリの実行のたびにコンパイルを行うオーバーヘッドを削減できる。

【0095】

登録者名カラム 605、登録日時カラム 606 にはそれぞれ、クエリ登録時に取得したクエリの登録者名、クエリの登録日時が格納され、システムのアクセス管理、セキュリティ管理に使用される。

【0096】

実行対象ストリームカラム 607、実行対象リレーション 608 にはそれぞれ、登録クエリから抽出したクエリ実行対象となるストリーム名、及びリレーション名を格納する。

【0097】

例えば、行 609 は、図 7 に示した登録クエリ 1201 に対し、前記メモリ上限設定部 204 において、メモリ上限値が「10MB」と設定された場合の登録クエリ管理テーブル 205 を表している。

【0098】

行 609 は、クエリ名カラム 601 の値が「Q1」、登録クエリカラム 602 の値が「SELECT 注文ID、製品ID、製品名、ライン名、注文者名 FROM S1 [ROWS 50]、R1、R2 WHERE S1.製品ID=R1.製品ID AND R1.製品ID=R2.製品ID AND R2.注文者=注文者A;」、メモリ上限値カラム 603 の値が「10MB」、クエリ実行形式格納先 604 の値が「0x7FFFAEE1」、登録者名カラム 605 の値が「榎山」、登録日時カラム 606 の値が「2006/02/01 13:07:26 JST」、実行対象ストリームカラム 607 の値が「S1」、実行対象リレーション 608 の値が「R1, R2」であることを表している。

【0099】

図 11a ~ 図 11c は、監視情報管理テーブル 208 の構成例を表した図である。監視情報管理テーブル 208₁、208₂、208₃ はそれぞれ、タイムスタンプが 10:00、10:10、10:20 における監視情報管理テーブル 208 を表した図である。

【0100】

図 11a ~ 図 11c において、ストリーム名カラム 801、カラム名カラム 802 にはそれぞれ、前記登録クエリ管理テーブル 205 を参照し、取得した監視対象のストリーム名、及び監視対象のカラム名を格納する。

【0101】

データレートカラム 803 には、前記ストリームデータ監視部 207 において取得した監視情報のうち、監視対象のストリームデータの到着レートを格納する。データレートを取得することにより、今後到着するストリームデータのデータ量を予測できる。データレート閾値カラム 804 には、前記データレート閾値・更新頻度設定部 206 において設定されたデータレート閾値を格納する。データレート閾値を用いることにより、リレーショ

10

20

30

40

50

ンデータ 2 2 6 をキャッシュするか否かを判定する。

【 0 1 0 2 】

現在の値カラム 8 0 5 には、前記ストリームデータ監視部 2 0 7 において取得した監視情報のうち、監視情報を取得した時点における監視対象のカラム名の最新の値を格納する。統計値カラム 8 0 6 には、前記ストリームデータ監視部 2 0 7 において取得した監視情報のうち、前回監視情報を取得した時点からの統計値を格納する。タイムスタンプカラム 8 0 7 には、前記ストリームデータ監視部 2 0 7 において、監視情報を取得した時刻情報を格納する。

【 0 1 0 3 】

更新頻度カラム 8 0 8 には、前記データレート閾値・更新頻度設定部 2 0 6 において設定された監視情報の更新間隔を格納する。

10

【 0 1 0 4 】

例えば、行 8 0 9₁ は、前記データレート閾値・更新頻度設定部 2 0 6 において、データレート閾値が「5 タプル/分」、更新頻度が「10 分」と設定された場合の監視情報管理テーブル 2 0 8₁ を表している。

【 0 1 0 5 】

行 8 0 9₁ は、ストリーム名カラム 8 0 1 の値が「S 1」であり、カラム名カラム 8 0 2 の値が「製品 ID」であり、データレートカラム 8 0 3 の値が「10 タプル/分」であり、データレート閾値カラム 8 0 4 の値が「5 タプル/分」であり、現在の値カラム 8 0 5 の値が「1000」であり、統計値カラム 8 0 6 の値が「-」（データなし）であり、タイムスタンプカラム 8 0 7 の値が「10 : 00」であり、更新頻度カラム 8 0 8 の値が「10 分」であることを表している。

20

【 0 1 0 6 】

そして、図 1 1 b の行 8 0 9₂ は、10 : 10 において、データレートカラム 8 0 3 の値が「9 タプル/分」に、現在の値カラム 8 0 5 の値が「1090」にそれぞれ変化しており、図 1 1 c の行 8 0 9₃ は、10 : 20 において、データレートカラム 8 0 3 の値が「8 タプル/分」に、現在の値カラム 8 0 5 の値が「1180」にそれぞれ変化していることを表している。

【 0 1 0 7 】

ここで、本実施形態では、データレート 8 0 3 に格納するデータとして、タプル/分の例を表したが、これに限定されるものではない。統計値カラム 8 0 6 に格納するデータは、前回監視情報を取得した時点からの統計情報としたが、それ以前から取得している統計情報でも構わない。また、統計値として、中心値、及び平均値を示したが、分散などの他の統計情報でもよい。さらに、後述するストリームデータの特性情報の一部を、前記ストリームデータを監視することにより取得してもよい。

30

【 0 1 0 8 】

また、データレート閾値カラム 8 0 4、更新頻度カラム 8 0 8 はすべての情報が入力される必要はなく、入力されない場合、事前に設定された値を用いてもよい。

【 0 1 0 9 】

また、データレートカラム 8 0 3、現在の値カラム 8 0 5、統計値カラム 8 0 6 に格納する監視情報取得データはすべての情報が入力される必要はない。

40

【 0 1 1 0 】

図 1 2 は、ストリーム特性管理テーブル 2 1 5 の構成例を表した図である。

【 0 1 1 1 】

ストリーム名カラム 7 0 1 には、特性情報を入力するクエリの名前を格納する。カラム名カラム 7 0 2 には、クエリにおける前記特性情報を入力するカラムのカラム名を格納する。時間に関する条件カラム 7 0 3、値に関する条件カラム 7 0 4 にはそれぞれ、タイムスタンプに関する条件、値の特性を示す条件が格納され、組み合わせて使用される。条件の優先度カラム 7 0 5 には、優先度が格納され、複数のストリーム特性に対してどの条件を優先的に使用するか決定する。制約判定フラグカラム 7 0 6 には、前記時間付き制約設

50

定部 2 1 3 で入力された場合、フラグを Y として格納する。確定条件フラグカラム 7 0 7 には、特性情報が必ず満たされる、若しくは満たされると想定される場合、フラグを Y として格納する。登録者カラム 7 0 8 には、特性情報を登録した登録者名を格納する。

【 0 1 1 2 】

例えば、行 7 0 9 は、「S 1」というストリーム名を有するストリームデータにおいて、「1 0 : 0 0 から 1 0 : 2 0」において、「製品 ID」は「1 0 0 0 から 1 5 0 0 まで」の範囲であり、「優先的に処理を行う」ことを示している。また、「条件を違反していないか否かを判定」し、「この条件が必ず満たされる、若しくは満たされると想定される」ことを表している。さらに、上記の内容は「堰山」によって登録されたことを表している。

10

【 0 1 1 3 】

ここで、ストリームデータ特性情報を入力する I / F は、図 1 2 に示した表形式の他に、コマンドによる入力でもよい。また、特性を入力するストリーム名カラム 7 0 1、及びカラム名カラム 7 0 2 は、前記登録クエリ管理テーブル 2 0 5 を参照し、取得してもよい。また、時間に関する条件カラム 7 0 3、条件の優先度カラム 7 0 5、制約判定フラグカラム 7 0 6、確定条件フラグカラム 7 0 7、登録者カラム 7 0 8 において、すべての情報が入力される必要はなく、入力されない場合、事前に設定された値を用いてもよい。

【 0 1 1 4 】

次に、前記図 8 に示した監視情報変更および設定変更時の処理 S 2 0 0 5 の処理手順を、図 1 3、及び図 2 を用いて説明する。

20

【 0 1 1 5 】

監視情報更新および設定変更時の処理では、まず、前記メモリ上限設定部 2 0 4 において、登録クエリがバッファに用いることのできるメモリの上限値を受け付けたか否かを判定する (S 2 2 0 2)。前記ステップ S 2 2 0 2 で y e s と判定された場合、メモリ上限値を前記登録クエリ管理テーブル 2 0 5 に反映させる (S 2 2 0 3)。

【 0 1 1 6 】

前記ステップ S 2 2 0 3 が終了した場合、または前記ステップ S 2 2 0 2 で n o と判定された場合、前記データレート閾値・更新頻度設定部 2 0 6 において、ストリームデータ監視情報の更新間隔変更を受け付けたか、また前記ストリームデータ監視部 2 0 7 において、監視情報を更新したか否かを判定する (S 2 2 0 4)。前記ステップ S 2 2 0 4 で y e s と判定された場合、監視情報の更新間隔、または更新された監視情報を前記監視情報管理テーブル 2 0 8 に反映させる (S 2 2 0 5)。

30

【 0 1 1 7 】

前記ステップ S 2 2 0 5 が終了した場合、または前記ステップ S 2 2 0 4 で n o と判定された場合、前記ストリームデータ特性設定部 2 1 4 において、ストリームデータ特性が変更されたか、また前記時間付き制約設定部 2 1 3 において、時間付き制約が設定されたかを判定する (S 2 2 0 6)。前記ステップ S 2 2 0 6 で y e s と判定された場合、ストリームデータ特性情報、または時間付き制約情報を前記ストリームデータ特性管理テーブル 2 1 5 に反映させる (S 2 2 0 7)。

【 0 1 1 8 】

前記ステップ S 2 2 0 7 が終了した場合、または前記ステップ S 2 2 0 6 で n o と判定された場合、上記図 9 の S 2 1 0 9 と同様に、後述する S P C 管理テーブル更新処理を行い (S 2 2 0 8)、後述する突き合わせ処理時の S P C 取得処理を行い (S 2 2 0 9)、監視情報更新および設定変更時の処理を終了する (S 2 2 1 0)。

40

【 0 1 1 9 】

次に、前記図 8 に示したリレーションデータ更新時の処理 S 2 0 0 6 の処理手順を図 1 4 のフローチャートに示し、図 2 を参照しながら以下に説明する。

【 0 1 2 0 】

リレーションデータ更新時の処理では、まず、前記 S P C クリア命令発行部 2 1 0 において、前記ユーザ 1 1 5、または前記クライアント計算機 1 1 6 上で実行される前記アプ

50

リケーション117から前記コマンド入力部229を通して、または一定間隔により入力されたSPCクリアコマンドを受け付けたか否かを判定する(S2302)。前記ステップS2302でyesと判定された場合、上記図9のS2109と同様に、後述するSPC管理テーブル更新処理を行い(S2303)、後述する突き合わせ処理時のSPC取得処理を行う(S2304)。前記ステップS2304が終了した場合、または前記ステップS2302でnoと判定された場合、リレーションデータ更新時の処理を終了する(S2305)。

【0121】

以上により、SPCクリア命令が発行されるとSPC管理テーブル211とSPC110の更新が行われる。

10

【0122】

次に、前記図9のS2109、図13のS2208、図14のS2303、図23のS2715で呼び出されるSPC管理テーブル更新処理の処理手順を、図15のフローチャートに示し、図2、図16、図17、図18を参照しながら説明する。

【0123】

SPC管理テーブル更新処理では、まず、前記SPC管理部212において、後述するSPC格納場所判別条件の算出処理を行い(S2402)、前記ステップS2402で算出した格納場所判別条件を満たすSPC_OM217に格納されているSPCデータのメモリ使用量を計算する(S2403)。次に、前記SPC管理部212において、前記登録クエリ管理テーブル205を参照し(S2404)、計算したメモリ使用量が、前記登録クエリ管理テーブル205に保持されているメモリ上限値カラム603よりも大きい

20

否かを判定する(S2405)。

【0124】

前記ステップS2405でyesと判定された場合、メモリ上限値分のリレーションレコードのタプル数を計算する(S2406)。そして、各タプルの距離が小さいものから前記ステップS2406で計算したメモリ上限値分のリレーションレコードタプル数だけメモリ216に格納するものとし、SPC_OM217の格納場所判別条件を算出する(S2407)。

【0125】

ここで、タプルの距離は、距離関数を予め定義しておき、距離が小さいかを判定する。例えば、ストリームデータの属性値が現在の値と比較した値を距離とする距離関数を定義した場合、現在の値50に対し、属性値80(距離30)と属性値10(距離40)を比較すると、属性値80の方が小さい距離であると判定できる。その他にも、現在の値よりも小さい値は無限大とするような距離関数や、自身で定義した距離関数(統計結果から導き出した関数など)も適用することができる。

30

【0126】

前記ステップS2407が終了した後、又は前記ステップS2405でnoと判定された場合、S2402で計算された格納場所判別条件を前記SPC管理テーブル211に反映させ(S2408)、SPC管理テーブル更新処理を終了する(S2409)。

【0127】

以上の処理によりSPC管理テーブル211のリレーション名901毎に、現在のストリームデータ111の値に応じた格納場所判別条件902が設定される。

40

【0128】

図16、図17、図18は、SPC管理テーブル211の構成例を表した図である。図16は、タイムスタンプ10:00におけるSPC管理テーブル211₁、図17は、タイムスタンプ10:10におけるSPC管理テーブル211₂、図18は、タイムスタンプ10:20におけるSPC管理テーブル211₃を表している。

【0129】

図16～図18のリレーション名カラム901には、登録クエリから取得した突き合わせ対象のリレーション名を格納する。ここで、前記SPC管理テーブル211に保持され

50

るリレーション名は、すべてのリレーション名でもよい。格納場所判別条件カラム902には、格納場所を決定するための条件を格納し、格納場所カラム903に格納場所を設定する。更新頻度カラム904、更新ポリシーカラム905はそれぞれ、リレーションデータの更新頻度、及びリレーションデータに更新が生じたときのポリシーを格納する。SPCインデクスカラム906には、SPCデータに対し、インデクスを張るカラム名を格納する。取得コストカラム907には、前記リレーションデータ226を前記ストリーム処理システム102上にSPCデータとして取得する際のコストを格納する。なお、取得コストは、リレーションデータ226を取得するまでの時間や、時間に対応する値などで表記することができる。最後に、タイムスタンプカラム908には、更新されたタイムスタンプを格納する。

10

【0130】

ここで、前記更新頻度カラム904、更新ポリシーカラム905、SPCインデクスカラム906、取得コストカラム907の値では、すべての情報が入力される必要はなく、入力されない場合、事前に設定された値を用いてもよい。また、前記SPC情報設定部228において、前記ユーザ115、または前記クライアント計算機116上で実行される前記アプリケーション117から、明示的に外部データを取得しないことを指定してもよい。

【0131】

例えば、図16の前記SPC管理テーブル211₁において、行909₁は、10:00において、「R1」というリレーション名を有するリレーションデータにおける製品IDが1000以上1100未満の場合、メモリ(SPC_OM217)216にSPCデータを格納することを表している。また、リレーションデータの更新頻度が低く、リレーションデータ更新時の処理では、SPCクリア命令発行時にリレーションデータ226を再取得し、SPCデータを入れ替える。また、SPCデータにはインデクスが張られてなく、SPCデータ取得コストが高いことを表している。

20

【0132】

また、行910₁は、R1というリレーション名を有するリレーションデータにおいて、10:00において、製品IDが1100以上1200未満の場合、ディスク(SPC_OD219)にSPCデータを格納することを表している。

【0133】

また、SPC管理テーブル211₂において、行909₂は、10:10において、メモリ(SPC_OM)にSPCデータを格納する条件が、製品IDが1090以上1180未満の場合に更新されたことを表している。行910₂は、10:10において、ディスク(SPC_OD)にSPCデータを格納する条件が、製品IDが1180以上1270未満の場合に更新されたことを表している。

30

【0134】

また、SPC管理テーブル211₃において、行909₃は、10:20において、メモリ(SPC_OM)にSPCデータを格納する条件が、製品IDが1500以上1580未満の場合に更新されたことを表している。行910₃は、10:20において、ディスク(SPC_OD219)にSPCデータを格納する条件として、製品IDが1580以上1660未満の場合に更新されたことを表している。

40

【0135】

次に、前記SPC格納場所判別条件の算出処理S2402の処理手順を、図19のフローチャートに示し、図2、図11a~図11c、図12、図16、図17、図18を参照しながら説明する。

【0136】

SPC格納場所判別条件の算出処理では、まず、前記ストリームデータ特性管理テーブル215を参照し(S2502)、図12に示した時間に関する条件カラム703にデータが格納されているか否かを判定する(S2503)。この処理は、ストリームデータ特性管理テーブル215のストリーム名毎に行われる。

50

【 0 1 3 7 】

前記ステップS 2 5 0 3でy e sと判定された場合、現在のタイムスタンプが満たす条件における値に関する条件カラム7 0 4の条件を適用する（S 2 5 0 4）。前記ステップS 2 5 0 4が終了した後、または前記ステップS 2 5 0 3でn oと判定された場合、値に関する条件カラム7 0 4の条件を適用する（S 2 5 0 5）。ここで、前記値に関する条件7 0 4がシーケンシャル属性など単体で適用できない場合は、処理を行わない。

【 0 1 3 8 】

次に、前記監視情報管理テーブル2 0 8を参照し現在の監視情報を取得する（S 2 5 0 6）。そして、取得した監視情報のうちデータレート8 0 3が、図1 1 a ~ 図1 1 cに示した前記データレート閾値カラム8 0 4の値以上となっているか否かを判定する（S 2 5 0 7）。前記ステップS 2 5 0 7でy e sと判定された場合、ジョイン対象のカラムの値に関する図1 2の条件カラム7 0 4がシーケンシャル属性となっているか否かを判定する（S 2 5 0 9）。

10

【 0 1 3 9 】

上記S 2 5 0 9の判定は、後の突き合わせを行うためのものであり、突き合わせではある属性値をキーにして突き合わせを行うことをジョイン処理とする。上記S 2 5 0 9のジョイン対象のカラムとは、あるテーブルに対してジョイン処理を行うストリームデータのキーのカラムを意味する。例えば、図7に示した登録クエリ1 2 0 1では、S 1とR 1に対し、S 1の製品IDとR 1の製品IDをキーにして突き合わせを行うことを意味する。R 1のジョイン対象のカラムとは「S 1の製品ID」カラムを指しており、S 1の製品IDカラムがシーケンシャル属性を持っている場合S 2 5 0 9の判定がy e sとなる。

20

【 0 1 4 0 】

次に、前記ステップS 2 5 0 9でy e sと判定された場合、前記現在の値カラム8 0 5の値よりも小さな値を格納場所判別条件から除外する（S 2 5 1 0）。

【 0 1 4 1 】

ここで、S 2 5 1 0で行われる除外は、格納場所判別条件について「条件式の範囲を狭くする」ことを意味する。例えば、上記S 2 5 0 5では、値に関する条件範囲が「 $30 \leq \text{属性値} < 100$ 」となっており、S 2 5 1 0で値に関する条件カラム7 0 4の現在の値が5 0となっていた場合、現在の値よりも小さい範囲を除外して、「 $50 \leq \text{属性値} < 100$ 」とすることを意味するものである。

30

【 0 1 4 2 】

前記ステップS 2 5 1 0が終了した後、又は前記ステップS 2 5 0 9でn oと判定された場合、前記図1 1 aの監視情報管理テーブル2 0 8のデータレートカラム8 0 3、前記現在の値カラム8 0 4、前記統計値カラム8 0 5、タイムスタンプカラム8 0 6、更新頻度カラム8 0 7の値を用いて予測条件を算出する（S 2 5 1 1）。

【 0 1 4 3 】

このS 2 5 1 1で求める予測条件は、ユーザから入力された時間に関する条件（S 2 5 0 3）や、シーケンシャル属性（S 2 5 0 9）以外に、ストリームデータの監視で取得した監視情報から上記条件式の範囲を絞り込むことを意味する。例えば、S 2 5 0 5では、範囲が「 $30 \leq \text{属性値} < 100$ 」となっており、ストリームデータからの監視情報から、範囲が「 $40 \leq \text{属性値} < 60$ 」であることが分かった場合、S 2 5 1 1において上記の予測条件でさらに絞り込みを行う。上記で示した例以外にも、監視情報から条件を導き出す場合、上記S 2 5 1 2で予測条件の処理を行うことができる。

40

【 0 1 4 4 】

上記予測条件の演算は、例えば、図1 2の値に関する条件カラム7 0 4にシーケンシャル属性がある場合、予測条件を「現在の値以上、現在の値+データレート×更新頻度未満」とすることができる。ここで、上記で示した予測条件算出法に限定せず、他の予測アルゴリズムを用いてもよい。

【 0 1 4 5 】

前記ステップS 2 5 0 7でn oと判定された場合、リレーションデータをストリームデ

50

ータ処理システム102のSPC110上に取得しないため、メモリ、及びディスクの格納場所判別条件を無効化する(S2508)。

【0146】

ここで、「無効化する」とは「キャッシュしない」ことを意味する。すなわち、図16に示したリレーション名「R4」に対する行が無効化に対応する。例えば、S2505では、範囲が「30 ≤ 属性値 < 100」となっており、データレートが閾値以下(あまりデータが到着していない)の場合には、キャッシュをせず、条件式を「(何もなし)」とする。ちなみに、全範囲がキャッシュされる場合もある。

【0147】

前記ステップS2508が終了した後、または前記ステップS2511が終了した後は、今まで設定したすべての格納場所判別条件を適用し(S2512)、SPC格納場所判別条件の算出処理を終了する(S2513)。

10

【0148】

以上の処理により、SPC管理テーブル211のリレーション名901毎に、現在のストリームデータ111の値に応じた格納場所判別条件902が算出される。

【0149】

次に、図9に示したS2110、図13のS2209、図14のS2304、図23のS2716で行われる突き合わせ処理時のSPCデータ取得処理の処理手順を、図20のフローチャートに示し、図2、図21を参照しながら説明する。

【0150】

20

突き合わせ処理時のSPCデータ取得処理では、まず、前記SPC管理テーブル211を参照し(S2602)、上記図19の処理で格納場所判別条件902が更新された際に、SPC_OM217、及びSPC_OD219において格納場所判別条件902から除外されたSPCデータを削除する(S2603)。次に、前記登録クエリ管理テーブル205を参照し(S2604)、SPCデータ取得命令を次のように発行する(S2605)。

【0151】

SPCデータ取得命令は、前記登録クエリ管理テーブル205に保持されているクエリ602に対し、リレーションデータに関連のある行、及び列を取得する命令を生成し、前記SPC管理テーブル211に保持されている条件(格納場所903など)を付加することにより生成する。ここで、関連のある行は登録クエリ602の選択処理に記述されている条件から生成し、関連のある列は射影、結合処理に記述されている条件から生成する。

30

【0152】

最後に、前記ステップS2605のSPC取得命令で取得したSPCデータを前記SPC管理テーブル211に記述されている格納場所判別条件902により、格納先をメモリ(SPC_OM217)またはディスク(SPC_OD219)のいずれかに振り分け(S2606)、突き合わせ処理時のSPCデータ取得処理を終了する(S2607)。

【0153】

以上の処理により、登録クエリ602と格納場所判別条件902に基づいてSPC取得命令を発行し、データ管理システム101のリレーションデータ226がSPC_OM217またはSPC_OD219に読み込まれる。

40

【0154】

図21は、SPC取得命令の例を表している。SPC取得命令1301₁、1301₂、1301₃はそれぞれ、10:00、10:10、10:20におけるSPC取得命令を表している。

【0155】

例えば、図7に示した前記登録クエリ1201の場合、図4、図5、図6よりSELECT句における「ライン名」、FROM句における「S1[Rows 50]」、WHERE句における「S1.製品ID=R1.製品ID」は前記製品ストリーム111₁に關係する。一方、SELECT句における「注文ID」、「製品ID」、「製品名」、「注文者名」、F

50

ROM句における「R1」、「R2」、WHERE句における「R1.製品ID=R2.製品ID」、「R2.注文者='注文者A'」は前記製品ストリーム111₁に関係せず、前記製品管理表226₁、及び前記注文表226₂に関係する。以上より、リレーションデータに関連のある行、及び列を取得する命令「SELECT 注文ID、製品ID、製品名、注文者名 FROM R1、R2 WHERE R1.製品ID=R2.製品ID AND R2.注文者='注文者A';」を生成できる。

【0156】

また、図16に示した前記SPC管理テーブル211₁に保持されている10:00のメモリ、及びディスク両方の格納場所判別条件「1000<=製品ID<1200」を上記SPC取得命令に付加すると、10:00におけるSPC取得命令13011「SELECT 注文ID、製品ID、製品名、注文者名 FROM R1、R2 WHERE R1.製品ID=R2.製品ID AND R2.注文者='注文者A' AND 1000<=製品ID AND 製品ID<1200;」を生成できる。

10

【0157】

また、10:10、及び10:20における格納場所判別条件902を付加することにより、10:10、10:20におけるSPC取得命令13012、13013を同様に生成できる。

【0158】

ここで、本実施例では、SPC取得命令をSQL文で表現したが、これに限定されるものではない。また、本実施例では、2種類のリレーションデータを結合した結果を、SPCデータとしたが、SPCデータ取得命令において、リレーションデータを個別に取得し、前記ストリームデータ処理システム102中で結合処理を行ってもよい。

20

【0159】

図22は、前記SPC取得命令1301₁、1301₂、1301₃により取得したSPC__OMデータ217₁、217₂、217₃、及びSPC__ODデータ219₁、219₂、219₃を表した図である。

【0160】

図22において、SPC__OMデータ217₁~217₃、及びSPC__ODデータ219₁~219₃は、図5に示した前記製品管理表226₁、及び図6に示した前記注文表226₂を突き合わせた図となっている。図22で注文IDカラム1001は、前記注文表226₂の注文IDカラム501に対応し、製品IDカラム1002は、前記製品管理表226₁の製品IDカラム401、または前記注文表226₂の製品IDカラム502に対応し、製品名カラム1003は、前記製品管理表226₁の製品名カラム402に対応し、注文者名カラム1004は、前記注文表226₂の注文者名カラム503に対応する。

30

【0161】

例えば、行1005は、前記製品管理表226₁の行404、及び前記注文表226₂の行504を突き合わせた行になっている。また、行1006は、同様に行405、及び行505を突き合わせた行になっている。同様にして、SPC__OMデータ217₂、217₃、及びSPC__ODデータ219₂、219₃を取得できる。

40

【0162】

次に、前記図8のS2008で実行される突き合わせ処理の処理手順を図23のフローチャートに示し、図2を参照しながら以下に説明する。

【0163】

突き合わせ処理では、まず、前記突き合わせ部112において、前記ストリームデータ111を取得する(S2702)。次に、前記ストリームデータ特性管理テーブル215を参照し(S2703)、前記ストリームデータ特性管理テーブル215に保持されている制約条件(図12に示した前記制約判定フラグカラム706の値が「Y」となっている時間に関する条件カラム703、及び値に関する条件カラム704の条件)を満たしているか否かを判定する(S2704)。前記ステップ2704でnoと判定された場合、ス

50

トリームデータを破棄し (S 2 7 1 7)、破棄したことを示す前記ログ 2 2 2 を出力し (S 2 7 1 8)、突き合わせ処理を終了する (S 2 7 1 9)。

【 0 1 6 4 】

図 2 4 は、ログ 2 2 2 の例を表した図である。

【 0 1 6 5 】

例えば、ログ 1 4 0 1 は、「 1 0 : 2 1 」にタプル「 (1 0 : 2 1 , 2 5 0 0 , ライン A) 」を破棄したことを表している。また、タプルを破棄した条件は「 檜山 」が登録した「 1 0 : 0 0 <=時間<= 1 1 : 0 0 , 1 5 0 0 <=製品 I D <= 2 0 0 0 」であることを表している。

【 0 1 6 6 】

ここで、上記で示したログ 2 2 2 は、前記ユーザ 1 1 5、または前記クライアント計算機 1 1 6 上で実行される前記アプリケーション 1 1 7 に対するイベント、若しくはメッセージでもよい。

【 0 1 6 7 】

前記図 8 のステップ 2 7 0 4 で y e s と判定された場合、取得したストリームデータ 1 1 1 が制約条件に違反していないので、前記 S P C 管理テーブル 2 1 1 を参照し (S 2 7 0 5)、取得したストリームデータ 1 1 1 と突き合わせ対象の S P C データの格納場所を特定する (S 2 7 0 6)。次に、取得したストリームデータと S P C _ O M 2 1 7 のデータの突き合わせを行う (S 2 7 0 7)。ここで、突き合わせ対象の S P C データが S P C _ O M 2 1 7 にない場合は、処理を行わない。次に、突き合わせ対象の S P C データにおいて、 S P C _ O M 2 1 7 上に存在しない S P C データがあるか否かを判定する (S 2 7 0 8)。前記ステップ 2 7 0 8 で n o と判定された場合、突き合わせ処理を終了する (S 2 7 1 9)。

【 0 1 6 8 】

前記ステップ 2 7 0 8 で y e s と判定された場合、取得したストリームデータ 1 1 1 と S P C _ O D 2 1 9 のデータとの突き合わせを行う (S 2 7 0 9)。ここで、突き合わせ対象の S P C データが S P C _ O D 2 1 9 にない場合は、処理を行わない。次に、突き合わせ対象の S P C データが、 S P C _ O M 2 1 7 上にも S P C _ O D 2 1 9 上にも存在しないかを判定する (S 2 7 1 0)。前記ステップ 2 7 1 0 で n o と判定された場合、前記ステップ S 2 7 0 7 で実行した S P C _ O M 2 1 7 のデータとの突き合わせ結果または前記ステップ S 2 7 0 9 で実行した S P C _ O D 2 1 9 上のデータとの突き合わせ結果を結合する (S 2 7 1 4)。

【 0 1 6 9 】

前記ステップ 2 7 1 0 で y e s と判定された場合、突き合わせを行う S P C データがストリームデータ処理システム 1 0 2 の S P C 1 1 0 上にないので、前記データ管理システム 1 0 1 上にある前記リレーションデータ 2 2 6 を取得する。このため、 S P C データ取得命令を発行し (S 2 7 1 1)、データ管理システム 1 0 1 から該当するリレーションデータ 2 2 6 をストリームデータ処理システム 1 0 2 の S P C _ O M 2 1 7 または S P C _ O D に読み込んで、取得したストリームデータ 1 1 1 と S P C 上のリレーションデータの突き合わせを行う (S 2 7 1 2)。前記ステップ S 2 7 0 7 で実行した S P C _ O M 2 1 7 上の S P C データとの突き合わせ結果と、前記ステップ S 2 7 0 9 で実行した S P C _ O D 2 1 9 上の S P C データとの突き合わせ結果、及び前記ステップ S 2 7 1 2 で実行したリレーションデータ 2 2 6 との突き合わせ結果を結合する (S 2 7 1 3)。つまり、受信したストリームデータ 1 1 1 に、 S P C 1 1 0 上の該当する S P C データの情報を付加し、突き合わせ結果 1 1 3 を生成する。突き合わせ結果 1 1 3 は、ストリームデータ 1 1 1 にユーザ 1 1 5 が理解可能な有意な情報を付加したものに变换される。そして、ストリームデータ処理システム 1 0 2 は、この突き合わせ結果 1 1 3 をクライアント計算機 1 1 6 へ送信する。

【 0 1 7 0 】

前記ステップ S 2 7 1 3、または前記ステップ S 2 7 1 4 が終了した後に、突き合わせ

10

20

30

40

50

対象となるリレーションデータの予測を再び行うため、前記図9のS2109と同様に、SPC管理テーブル更新処理を行い(S2715)、前記突き合わせ処理時のSPCデータ取得処理を行い(S2716)、突き合わせ処理を終了する(S2719)。

【0171】

以上の処理により、ストリームデータ特性管理部109に設定された特性情報と、ストリームデータ監視情報管理部106がモニタした監視情報から、突き合わせに必要な外部データ(リレーションデータ226)を、ストリームデータ111が受信する以前に、SPC110上に読み出しておくことが可能となって、突き合わせをリアルタイムで行うことができる。また、SPC110をメモリ216上のSPC_OM217と、ディスク218上のSPC_OD219のように階層的に分けておくことで、リレーションデータ226の一部または全部をストリームデータ処理システム102のSPC110に予め格納しておくことができる。

10

【0172】

図25は、前記突き合わせ結果113の例を表した図である。

【0173】

突き合わせ結果は、図4に示した製品ストリーム111₁、及び図22に示したSPCデータ(SPC_OMデータ217、またはSPC_ODデータ219)との突き合わせ表となっている。タイムスタンプカラム1101は、前記製品ストリーム111₁におけるタイムスタンプカラム304に対応し、注文IDカラム1102は、前記SPCデータにおける注文IDカラム1001に対応し、製品IDカラム1103は、前記SPCデータにおける製品IDカラム1002に対応し、ライン名カラム1104は、前記製品ストリーム111₁におけるライン名カラム302に対応し、製品名カラム1105は、前記SPCデータにおける製品名カラム1003に対応し、注文者名カラム1106は、前記SPCデータにおける注文者名カラム1004に対応する。

20

【0174】

例えば、行1107は、製品ストリーム111₁における行305、及び前記SPC_OMデータ2171における行1005突き合わせた行になっている。

【0175】

以上、本発明の一実施形態について説明した。

【0176】

本発明は、上記の実施形態に限定されるものではなく、その要旨の範囲内で数々の変形が可能である。

30

【0177】

例えば、上記の実施形態では、クライアント計算機116、及びデータ管理システム101はネットワークで介して接続されているが、同一計算機上にあってもよい。また、記憶装置109として、メモリ216、及びディスク218の2階層構造としたが、メモリ216のみ1階層、若しくは3階層以上の構造でもよい。また、メモリ、ディスクに限定せず、テープ、フラッシュメモリなど他の記憶装置でもよい。

【0178】

また、上記の実施形態では、2種類のリレーションデータを結合した結果を、SPCデータとしたが、SPCデータ取得命令において、リレーションデータを個別に取得し、前記ストリームデータ処理システム102中で結合処理を行ってもよい。上記の場合に、図10に示した前記登録クエリ管理テーブル205において、前記メモリ上限値カラム603の値を突き合わせ対象のリレーションごとにメモリ上限値を設定してもよい。

40

【0179】

図26は、前記メモリ上限値カラム603の値を突き合わせ対象のリレーションごとにメモリ上限値を設定した場合の前記登録クエリ管理テーブル205を表している。行610、及び行611は、図10に示した前記登録クエリ管理テーブル205における行609に対応する。行610は、実行対象リレーションカラムが「R1」に対して、メモリ上限値を「7MB」であることを表し、行611は、実行対象リレーションカラムが「R2

50

」に対して、メモリ上限値を「3MB」であることを表している。

【0180】

また、上記の実施形態では、ストリームデータ特性管理テーブル215、及び監視情報管理テーブル208の両方を用いてSPCデータの格納場所を判定しているが、いずれか一方で判定してもよい。

【0181】

また、上記の実施形態では、前記ストリームデータ特性管理テーブル215に設定されるストリーム特性情報を前記ユーザ115、または前記クライアント計算機116上で実行される前記アプリケーション117が入力しているが、ストリームデータのスキーマ定義において、ストリームデータ特性情報を入力し、スキーマ定義から自動的に前記ストリームデータ特性管理テーブル215を生成してもよい。

10

【0182】

また、上記の実施形態では、前記登録クエリ管理テーブル205、監視情報管理テーブル208、SPC管理テーブル211、ストリームデータ特性管理テーブル215に設定される情報は、前記ユーザ115、または前記クライアント計算機116上で実行される前記アプリケーション117が設定しているが、例えば、Stratis Viglas, Jeffrey F. Naughton著: "Rate-based query optimization for streaming information sources", In Proc. of SIGMOD Conference 2002, pp. 37-48 (非特許文献3)に開示されているクエリの最適化を行う処理機構が設定してもよい。

20

【0183】

また、上記の実施形態では、リレーションデータ更新時に、SPCクリア命令により、SPCデータを更新する例を説明したが、前記データ管理システム101において、前記データ114(226)の変更を監視することにより、前記データ114の変更を検知し、変更のあった前記データ114を前記ストリームデータ処理システム102に送信することにより、SPCデータを更新してもよい。

【0184】

データ管理システム101のリレーションデータ226の変更を検知し、ストリームデータ処理システム102に変更の内容を送信することによってSPCデータの更新を行う処理について説明する。この処理は、図8のS2006で実行されるリレーション更新時の処理の他の例を示し、この処理手順を、図27のフローチャートに示し、図2を参照しながら以下に説明する。

30

【0185】

リレーション更新時の処理では、まず前記SPC管理部212において、前記データ管理システム101において生成されたタブルを受け付けたか否かを判定する(S2802)。前記ステップS2802でyesと判定された場合、前記SPC管理テーブル211を参照し(S2803)、SPCデータの格納場所を特定し(S2804)、該当する場所にあるタブルを変更されたタブルに入れ替える(S2805)。前記ステップS2805が終了した場合、または前記ステップS2802でnoと判定された場合、リレーション更新時の処理を終了する(S2806)。

40

【0186】

ここで、前記データ管理システム101においてリレーションデータ226の変更を検知するには、前記データ変更検知部201において、データベースにおけるトリガを設定する、またはアプリケーションからのデータ更新命令を監視することにより実現できる。そして、前記タブル生成部202において、変更した時間をタイムスタンプとして付加したタブルを生成し、前記ストリームデータ処理システム102に送信する。

【0187】

また、前記突き合わせ実行部220において、前記タブル生成部202において付加されたタイムスタンプを用いることにより、一貫性を完全に損なわない突き合わせを実現で

50

きる。ここで、前記データ管理システム101、前記ストリームデータ処理システム102の時刻は同期しているものとする。

【0188】

また、上記の実施形態では、前記ストリームデータ111がレコード形式である例を説明したが、XMLドキュメント形式やCSVファイル形式などいかなるデータ形式でもよい。さらに、上記の実施形態ではデータ管理システム101の前記記憶装置104中のデータ114がリレーションデータ226である例を説明したが、XMLドキュメント形式やCSVファイル形式などいかなるデータ形式でもよい。レコード形式以外の場合、図2で示した前記SPC管理部212中のデータ変換部230、若しくは前記突き合わせ実行部220中のデータ変換部230において、レコード形式以外のデータ形式からレコード形式へと変換することにより、突き合わせ実行部220において、ストリームデータ111、及び記憶装置104中のデータ114の突き合わせを行うことができる。

10

【0189】

図28は、前記ストリームデータ111がXMLドキュメント形式である場合の好適なデータフォーマットを模式的に表した図である。

【0190】

ストリームデータ303は、「<ストリームデータ>」、及び「</ストリームデータ>」で囲まれた領域で表現され、行1501は図3に示した製品IDカラム301に対応し、行1502は図3に示したライン名カラム302に対応する。前記突き合わせ実行部220中のデータ変換部230において、上記に示した対応関係からレコード形式に変換することにより、突き合わせを行うことができる。

20

【0191】

図29は、前記データ114がXMLドキュメント形式である場合の例を表した図である。

【0192】

図5に示した行404は、「<製品>」、及び「</製品>」で囲まれた領域で表現され、行1503は、図5に示した製品IDカラム401に対応し、行1504は、図5に示した製品名カラム402に対応し、行1505は、図5に示した価格カラム403に対応する。前記SPC管理部212中のデータ変換部230、若しくは前記突き合わせ実行部220中のデータ変換部230において、上記に示した対応関係からレコード形式に変換することにより、突き合わせを行うことができる。

30

【0193】

また、上記の実施形態では、ストリームデータ処理システム102は、任意のコンピュータシステムとして説明したが、前記ストリームデータ処理システム102で行う処理の一部、若しくは全部をストレージ装置で行ってもよい。

【0194】

また、上記の実施形態では、計算機12が製品の情報や環境条件(温度)などをストリームデータ111を出力する例を示したが、これに限定されるものではない。例えば、計算機12に代わって多数のセンサノードを管理するセンサネットサーバが、センサノードの測定値をストリームデータ111と出力し、ストリームデータ処理システム102でセンサノードの測定値を外部データに基づいて、ユーザ115が理解可能な有意な情報を含む突き合わせ結果113に変換し、クライアント計算機116へ提供するようにしてもよい。

40

【0195】

なお、ストリームデータを1次情報として取得し、前記取得した1次情報を予め設定した外部データの情報に基づいて2次情報に変換するストリームデータ処理システムにおいて、

前記ストリームデータを1次情報として取得するストリームデータ入力部と、
前記ストリームデータと突き合わせを行う外部データを格納する外部データ格納部と、
前記ストリームデータに対する処理を記述したクエリを予め登録するクエリ管理部と、

50

前記クエリを参照して、前記外部データの一部または全部を記憶装置に予め設定したストリームプレディクトキャッシュへ予め読み込む格納場所管理部と、

前記ストリームプレディクトキャッシュへ読み込んだ外部データと、前記ストリームデータとを突き合わせることにより、前記2次情報を生成する突き合わせ部と、を備えたことを特徴とするストリームデータ処理システム。

【0196】

また、上記ストリームデータ処理システムにおいて、

前記ストリームデータの値の特徴を表すストリームデータの特性情報を格納するストリームデータ特性管理部と、

前記ストリームデータを監視して監視情報を取得するストリームデータ監視部と、を有し、

前記格納場所管理部は、

前記ストリームデータの特性情報と、前記取得した監視情報とに基づいて、前記ストリームプレディクトキャッシュへ予め読み込む前記外部データを決定することを特徴とするストリームデータ処理システム。

【0197】

なお、請求項1に記載のストリームデータ処理方法において、

処理時刻以降に到着する前記ストリームデータの値の特徴を表すストリームデータの特性情報を入力し、前記ストリームデータの特性情報を用いて、前記ストリームプレディクトキャッシュとして取得する前記外部データを決定することを特徴とするストリームデータ処理方法、としてもよい。

【0198】

また、請求項1に記載のストリームデータ処理方法において、

前記ストリームデータを監視することにより監視情報を取得し、前記取得した監視情報を用いて、前記ストリームプレディクトキャッシュとして取得する前記外部データを決定することを特徴とするストリームデータ処理方法、としてもよい。

【0199】

また、請求項6に記載のストリームデータ処理方法において、

前記ストリームデータに対し複数の前記特性情報がある場合、何れの条件を優先的に使用するか判定する優先度と、前記ストリームデータにおける時刻情報に対応する条件であり、前記ストリームデータが違反していないかチェックを行う時間付き制約情報であるか判定するフラグを表す制約判定フラグカラムと、前記特性情報が必ず満たされる、若しくは満たされると想定されるか判定するフラグを表す確定条件フラグと、前記特性情報を登録した登録者とのうち一部または全部をストリームデータ特性管理テーブルに保持することを特徴とするストリームデータ処理方法、としてもよい。

【0200】

また、請求項6に記載のストリームデータ処理方法において、

前記ストリームデータ特性管理テーブルを参照し、前記ストリームデータ特性管理テーブルの値に関する条件カラムの条件を適用することを特徴とするストリームデータ処理方法、としてもよい。

【0201】

また、請求項8に記載のストリームデータ処理方法において、

前記時間付き制約情報に違反している場合、前記ストリームデータを破棄し、破棄した前記ストリームデータおよび前記時間付き制約情報をログとして出力することを特徴とするストリームデータ処理方法、としてもよい。

【0202】

また、請求項9に記載のストリームデータ処理方法において、

監視対象の前記ストリームデータの到着レートを表すデータレートと、前記データレートにより前記ストリームプレディクトキャッシュ要否を判定する値を表すデータレート閾値と、前記監視情報を取得した時点における監視対象の前記カラム名の最新の値を表す現

10

20

30

40

50

在の値と、前記監視情報を前回取得した時点からの統計値を表す統計値と、前記ストリームデータの監視情報の更新間隔を表す更新頻度カラムとのうち一部または全部を監視情報管理テーブルに保持することを特徴とするストリームデータ処理方法、としてもよい。

【0203】

また、請求項13に記載のストリームデータ処理方法において、

予め登録された前記クエリの実行形式の格納先のアドレスを表すクエリ実行形式格納先と、前記クエリ登録時に取得した前記クエリの登録者と、前記クエリ登録時に取得した前記クエリの登録日時と、前記クエリから抜き出した前記クエリ実行対象となるストリーム名を表す実行対象ストリームと、前記クエリから抜き出した前記クエリ実行対象となるリレーション名を表す実行対象リレーションとのうち、一部または全部を登録クエリ管理テーブルに保持することを特徴とするストリームデータ処理方法、としてもよい。

10

【0204】

また、請求項5に記載のストリームデータ処理方法において、

前記外部データの更新頻度と、前記外部データに更新が生じたときのポリシーを表す更新ポリシーと、前記ストリームプレディクトキャッシュデータに対しインデクスを張るカラム名を表すストリームプレディクトキャッシュインデクスと、前記外部データを前記ストリーム処理システム上にストリームプレディクトキャッシュデータとして取得するコストを表す取得コストとのうち、一部または全部をストリームプレディクトキャッシュ管理テーブルに保持することを特徴とするストリームデータ処理方法、としてもよい。

【0205】

20

また、請求項5に記載のストリームデータ処理方法において、

外部のデータ管理システムにおいて、外部データの変更を検知し、変更した時刻を付与したタプルを生成し、ストリームデータ処理システムに送信することを特徴とするストリームデータ処理方法、としてもよい。

【0206】

また、請求項5に記載のストリームデータ処理方法において、

外部のデータ管理システムにおいて生成されたタプルを受け付け、前記ストリームプレディクトキャッシュ管理テーブルを参照し、ストリームプレディクトキャッシュデータの格納場所を特定し、新規タプルに入れ替えることを特徴とするストリームデータ処理方法、としてもよい。

30

【産業上の利用可能性】

【0207】

以上のように、本発明は時々刻々と変化するストリームデータを外部のデータに基づいて有意な情報に変換し、ユーザへ提供する計算機システムに適用することができる。特に、リアルタイムで処理する必要があるストリームデータの量が膨大な量になる金融機関のシステムや、センサネットワークシステムなどに適用することができる。

【図面の簡単な説明】

【0208】

【図1】本発明の一実施形態に係るストリームデータ処理システム、および関連するシステム構成の概要を示すブロック図である。

40

【図2】本発明の一実施形態が適用されたストリームデータ処理システム、および関連するシステムの構成を示すブロック図である。

【図3】ストリームデータ111の好適なデータフォーマットの例を模式的に表した図である。

【図4】連続して入力されたストリームデータ111を例示した図であり、製品ストリーム(S1)111₁を表している。

【図5】リレーションデータ226を例示した図であり、製品管理表(R1)226₁を表している。

【図6】リレーションデータ226を例示した図であり、注文表(R2)226₂を表している。

50

【図 7】クエリを前記ストリームデータ処理システム 102 に登録する際の好適なコマンドの記述例を示す説明図である。

【図 8】ストリームデータ処理システム 102 を用いた突き合わせ処理の処理手順を表したフローチャートである。

【図 9】クエリ登録・変更時の処理 S2003 の処理手順を表したフローチャートである。

【図 10】登録クエリ管理テーブル 205 の構成例を表した図である。

【図 11 a】時刻 10:00 における監視情報管理テーブル 208 の一例を表した説明図である。

【図 11 b】時刻 10:10 における監視情報管理テーブル 208 の一例を表した説明図である。

【図 11 c】時刻 10:20 における監視情報管理テーブル 208 の一例を表した説明図である。

【図 12】ストリーム特性管理テーブル 215 の構成例を表した図である。

【図 13】監視情報変更および設定変更時の処理 S2005 の処理手順を表したフローチャートである。

【図 14】リレーションデータ更新時の処理 S2006 の処理手順を表したフローチャートである。

【図 15】SPC 管理テーブル更新処理 S2109 の処理手順を表したフローチャートである。

【図 16】SPC 管理テーブル 211 の構成例を表した図であり、タイムスタンプ 10:00 における SPC 管理テーブル 211₁ を表している。

【図 17】SPC 管理テーブル 211 の構成例を表した図であり、タイムスタンプ 10:10 における SPC 管理テーブル 211₂ を表している。

【図 18】SPC 管理テーブル 211 の構成例を表した図であり、タイムスタンプ 10:20 における SPC 管理テーブル 211₃ を表している。

【図 19】SPC 格納場所判別条件の算出処理 S2402 の処理手順を表したフローチャートである。

【図 20】突き合わせ処理時の SPC データ取得処理 S2110 の処理手順を表したフローチャートである。

【図 21】SPC 取得命令の一例を示す説明図。

【図 22】SPC 取得命令 13011、13012、13013 により取得した SPC_OM データ 217₁、217₂、217₃、及び SPC_OD データ 219₁、219₂、219₃ を表した図である。

【図 23】突き合わせ処理 S2008 の処理手順を表したフローチャートである。

【図 24】ログ 222 の例を表した説明図である。

【図 25】突き合わせ結果 113 の例を表した説明図である。

【図 26】メモリ上限値カラム 603 の値を突き合わせ対象のリレーションごとにメモリ上限値を設定した場合の登録クエリ管理テーブル 205 を表す説明図。

【図 27】データの変更を検知し、ストリームデータ処理システムに送信することによる SPC データの更新する場合のリレーション更新時の処理 S2006 の処理手順の他の例を表したフローチャートである。

【図 28】ストリームデータ 111 が XML ドキュメント形式である場合の好適なデータフォーマットを模式的に表した説明図である。

【図 29】データ 114 が XML ドキュメント形式である場合の例を表した説明図である。

【符号の説明】

【0209】

101 データ管理システム

102 ストリームデータ処理システム

10

20

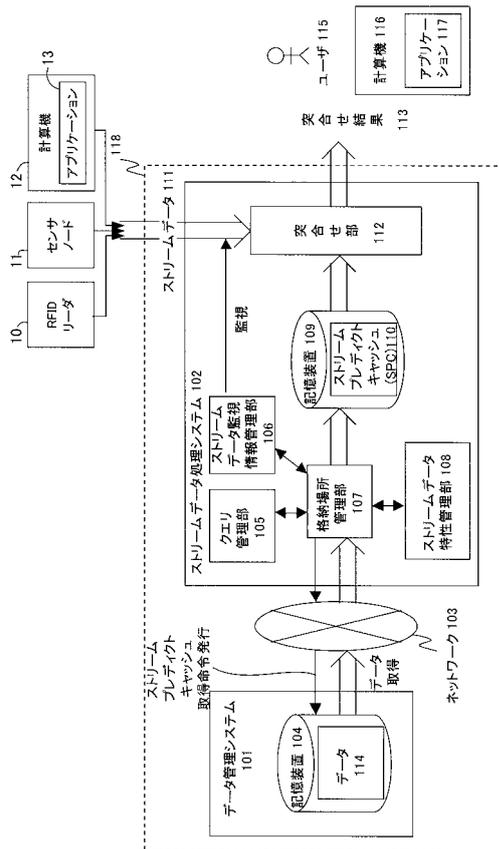
30

40

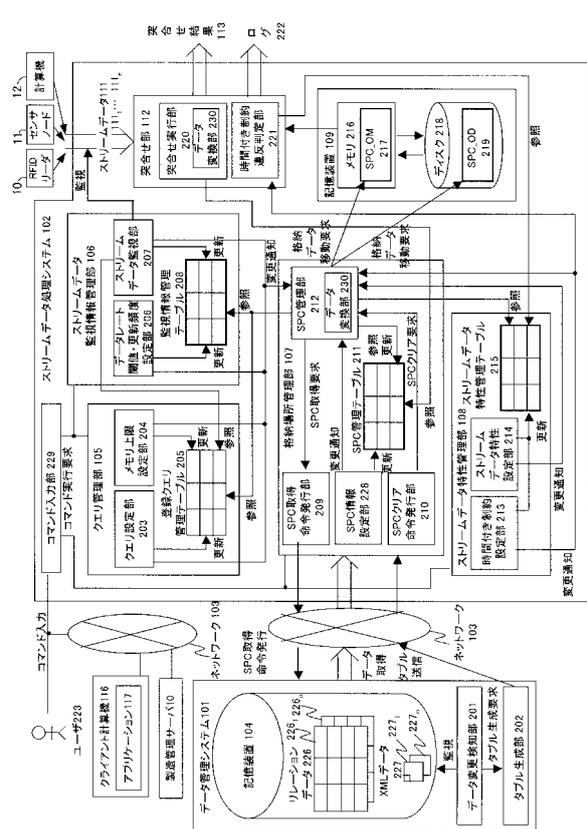
50

- 1 0 3 ネットワーク
- 1 0 4 記憶装置
- 1 0 5 クエリ管理部
- 1 0 6 ストリームデータ監視情報管理部
- 1 0 7 格納場所管理部
- 1 0 8 ストリームデータ特性管理部
- 1 0 9 記憶装置
- 1 1 0 ストリームプレディクトキャッシュ (S P C)
- 1 1 1 ストリームデータ
- 1 1 2 突き合わせ部
- 1 1 3 突き合わせ結果
- 1 1 4 データ

【 図 1 】



【 図 2 】



【図3】

製品ID	ライン名	...
------	------	-----

【図4】

製品ストリーム(S1) 111,

タイムスタンプ	製品ID	ライン名
10:00	1000	ラインA
10:00	1001	ラインA
...
10:00	1009	ラインB
10:01	1010	ラインA
10:01	1011	ラインA
...
10:01	1019	ラインB
...
10:10	1090	ラインA
10:10	1091	ラインB
...
10:10	1099	ラインB
10:10	1100	ラインB
...
10:19	1179	ラインC
10:20	1500	ラインC
10:20	1501	ラインC
...
10:20	1510	ラインA
10:21	1511	ラインB
10:21	2500	ラインA
...

【図5】

製品管理表(R1) 2 2 6₁

製品ID	製品名	価格
1000	製品A	5000
1001	製品A	5000
...
1009	製品A	5000
1010	製品A	5000
1011	製品A	5000
...
1019	製品A	5000
...
1090	製品B	8000
1091	製品B	8000
...
1099	製品B	8000
1100	製品B	8000
...
1199	製品C	2000
...
1500	製品D	10000
1501	製品D	10000
...
1510	製品D	10000
1511	製品D	10000
...

【図6】

注文表(R 2) 2 2 6₂

注文ID	製品ID	注文者名
10000	1000	注文者A
10001	1001	注文者B
...
10009	1009	注文者C
10100	1010	注文者C
10101	1011	注文者C
...
10200	1019	注文者D
...
10300	1090	注文者A
10301	1091	注文者A
...
10399	1099	注文者D
10400	1100	注文者E
...
11000	1199	注文者A
...
12000	1500	注文者E
12001	1501	注文者A
...
12100	1510	注文者C
12100	1511	注文者D
...

【図17】

SPC管理テーブル 211
タイムスタンプ 10:10

リレーション名	格納場所 判別条件	格納場所	更新 頻度	更新 ポリシー	SPC インデク ス	取得 コスト	タイム スタンプ
R1	1000<=ID <1180	メモリ	低	SPCクリ ア命令時	無	高	10:10
R2	1170<=ID <1270	ディスク	低	SPCクリ ア命令時	無	高	10:10
R3	1000<=ID <1180	メモリ	高	即時更新	注文ID	低	10:10
R4	1170<=ID <1270	ディスク	高	即時更新	注文ID	低	10:10
R5	18<=温度<28	メモリ	高	即時更新	湿度	低	10:10
R6	13<=温度<18	ディスク	高	即時更新	湿度	低	10:10
R7	20<=温度<25	---	---	---	---	---	---
...

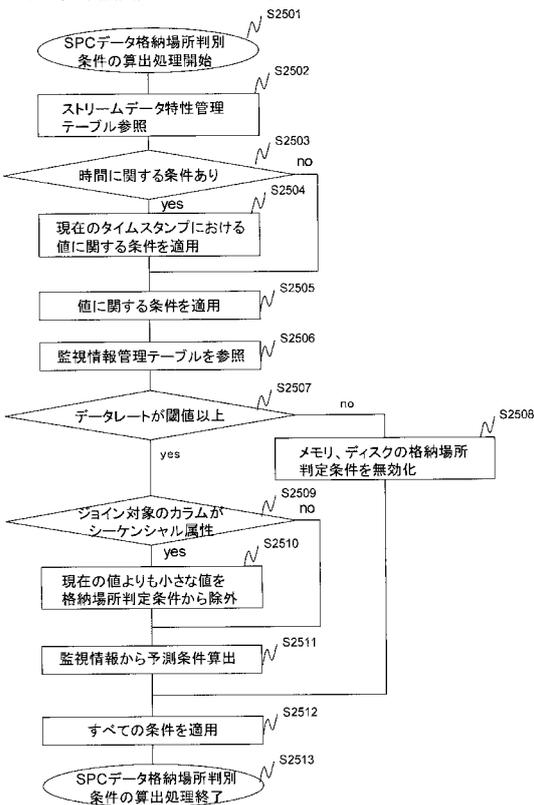
【図18】

SPC管理テーブル 211
タイムスタンプ 10:20

リレーション名	格納場所 判別条件	格納場所	更新 頻度	更新 ポリシー	SPC インデク ス	取得 コスト	タイム スタンプ
R1	1500<=ID <1580	メモリ	低	SPCクリ ア命令時	無	高	10:20
R2	1580<=ID <1660	ディスク	低	SPCクリ ア命令時	無	高	10:20
R3	1500<=ID <1580	メモリ	高	即時更新	注文ID	低	10:20
R4	1580<=ID <1660	ディスク	高	即時更新	注文ID	低	10:20
R5	19<=温度<29	メモリ	高	即時更新	湿度	低	10:20
R6	14<=温度<19	ディスク	高	即時更新	湿度	低	10:20
R7	29<=温度<34	---	---	---	---	---	---
...

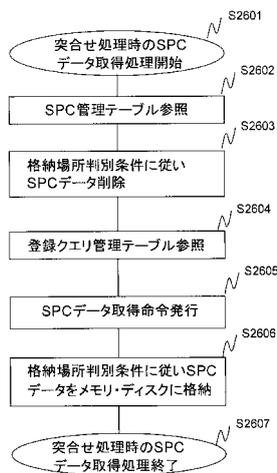
【図19】

SPCデータ格納場所判別条件の算出処理 S2402

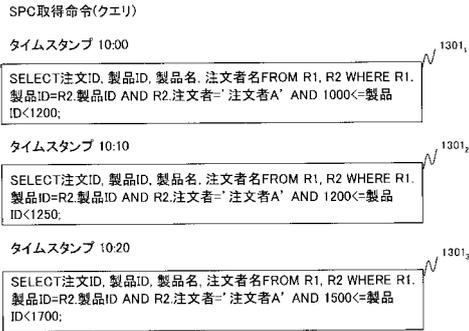


【図20】

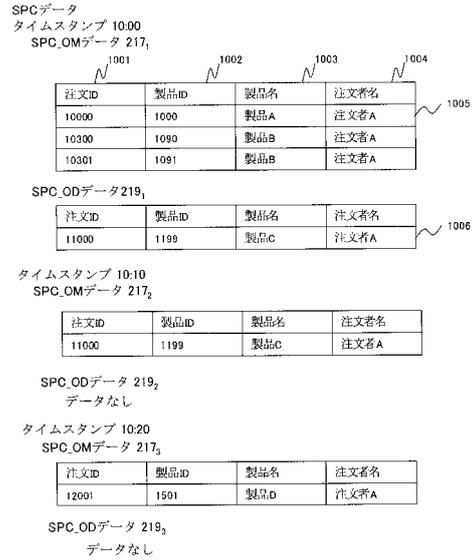
突合せ処理時のSPC取得処理 S2110,S2209,S2304,S2716



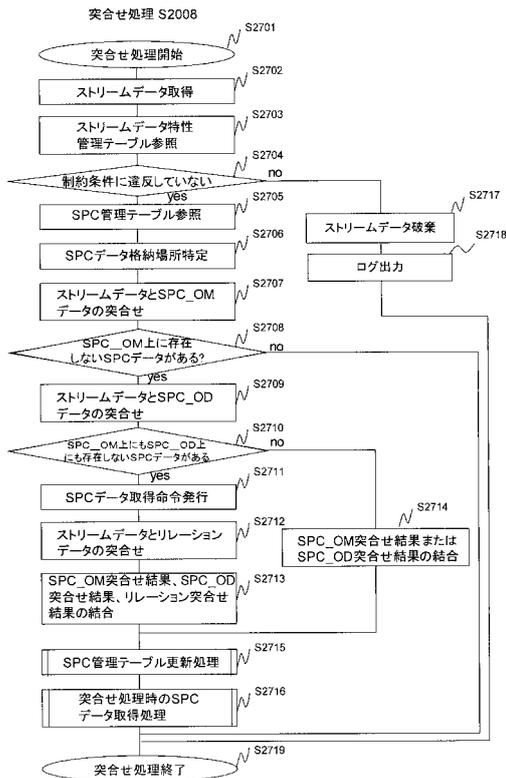
【図 2 1】



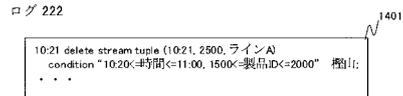
【図 2 2】



【図 2 3】



【図 2 4】



【図 2 5】

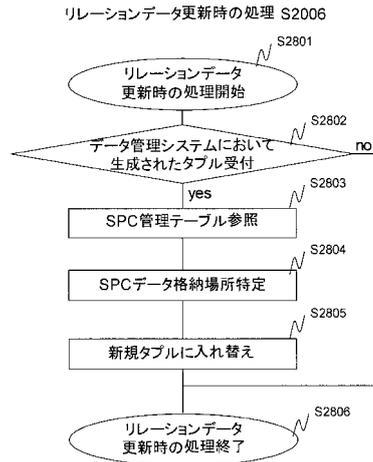


【図 26】

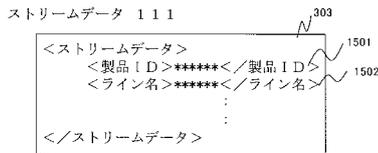
登録クエリ管理テーブル 205

クエリ名	登録クエリ	登録者	登録日時	実行対象ストリーマ	実行対象ストリーマ	メモリ上限
601	602	603	604	605	606	607
01	SELECT注文ID, 製品ID, 製品名, ライン名, 注文者名 FROM S1(FROWS 50), R1, R2 WHERE S1.製品ID=R1.製品ID AND R1.製品ID=R2.製 品ID AND R2.注文者=注文者A;	樺山	2006/02/01 13:07:26 JST	0x7FFFAE1	R1 R2	7MB 3MB
02	SELECT * FROM S2[Range 5 minutes], R3 WHERE S2. IJ=RCID;	樺山	2006/02/02 09:46:15 JST	0x7FFFA71	R3	20MB
...

【図 27】



【図 28】



【図 29】



フロントページの続き

(72)発明者 今木 常之

東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内

審査官 鈴木 和樹

(56)参考文献 渡辺陽介、外2名、問合せ最適化機構を備えたデータストリーム統合システムの開発、第15回データ工学ワークショップ(DEWS 2004) 論文集 [online]、日本、電子情報通信学会データ工学研究専門委員会、2004年 6月18日、p. 1 - 8、[DL from <http://www.ieice.org/iss/de/DEWS/proc/2004/paper/3-C/3-C-04.pdf>]

山田真一、外2名、時系列データベースとストリーム処理の統合のためのデータモデルの提案、情報処理学会研究報告(2005 - DBS - 137 (I I))、日本、社団法人情報処理学会、2005年 7月15日、第2005巻、第68号、p. 299 - 306

Shivnath Babu、外3名、Adaptive Caching for Continuous Queries, Data Engineering, 2005. ICDE 2005. Proceedings. 21st International Conference on [online], IEEE, 2005年 4月 8日、p. 118 - 129、[DL from IEEE Xplore]

(58)調査した分野(Int.Cl., DB名)

G06F 17/30

G06F 12/00

IEEE Xplore

JSTPlus(JDreamII)