

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2016年1月14日(14.01.2016)



(10) 国際公開番号
WO 2016/006098 A1

- (51) 国際特許分類:
G06F 12/00 (2006.01) G06F 17/30 (2006.01)
G06F 9/50 (2006.01)
- (21) 国際出願番号: PCT/JP2014/068579
- (22) 国際出願日: 2014年7月11日(11.07.2014)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: 株式会社日立製作所 (HITACHI, LTD.)
[JP/JP]; 〒1008280 東京都千代田区丸の内一丁目
6番6号 Tokyo (JP).
- (72) 発明者: 森田 悠介(MORITA, Yuusuke); 〒1008280
東京都千代田区丸の内一丁目6番6号 株式会
社日立製作所内 Tokyo (JP). 中野 幸生(NA-
KANO, Yukio); 〒1008280 東京都千代田区丸の内
一丁目6番6号 株式会社日立製作所内 Tokyo
(JP).
- (74) 代理人: 特許業務法人ウィルフォート国際特許
事務所 (WILLFORT INTERNATIONAL PATENT

FIRM); 〒1030016 東京都中央区日本橋小網町 1
9-7 日本橋TCビル 1階 Tokyo (JP).

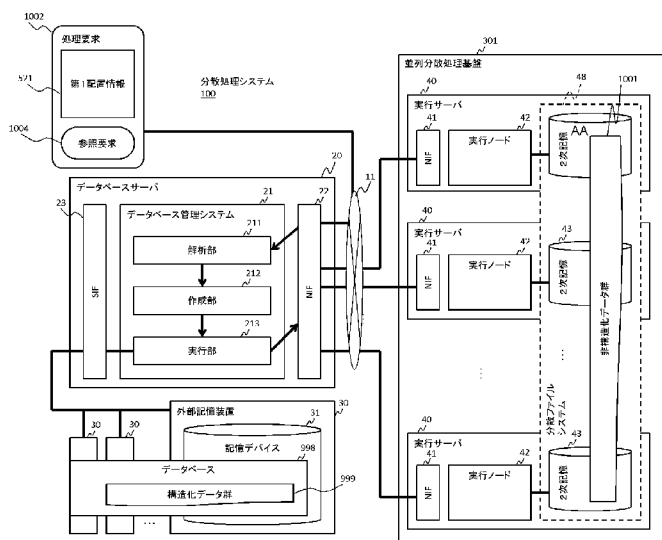
- (81) 指定国 (表示のない限り、全ての種類の国内保
護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA,
BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN,
CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES,
FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN,
IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR,
LS, LT, LU, LY, MA, MD, 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, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保
護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW,
MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシ
ア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ
(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), OAPI
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML,
MR, NE, SN, TD, TG).

[続葉有]

(54) Title: DATA PROCESSING SYSTEM, DATABASE MANAGEMENT SYSTEM, AND DATA PROCESSING METHOD

(54) 発明の名称: データ処理システム、データベース管理システム及びデータ処理方法

FIG. 1



- 20 Database server
- 21 Database management system
- 30 External storage apparatus
- 31 Storage device
- 40 Execution server
- 42 Execution node
- 43 Secondary storage
- 46 Distributed file system
- 100 Distributed processing system
- 211 Analysis unit
- 212 Creation unit
- 213 Execution unit
- 301 Parallel distribution processing base
- 521 First arrangement information
- 998 Database
- 999 Structured data group
- 1001 Unstructured data group
- 1002 Processing request
- 1004 Reference request

(57) Abstract: A DBMS that manages a database including a structured data group which is one or more structured data sets. The DBMS is capable of communicating with a plurality of execution nodes that manage an unstructured data group, which is one or more unstructured data sets, and execute parallel distributed processing. Each of the plurality of execution nodes is capable of: creating processing data that is based on at least one piece of structured data within the structured data group, which is managed by a DMB, and is further based on at least one piece of unstructured data within the unstructured data group, which is managed by the plurality of execution nodes; and processing using the processing data. On the basis of control information (information pertaining to control of the provision of two or more pieces of structured data designated by a reference request to the DBMS), two or more pieces of structured data are acquired from the database by the DBMS and are provided from the DBMS to two or more execution nodes among the plurality of execution nodes.

(57) 要約:

[続葉有]

WO 2016/006098 A1



添付公開書類:

— 国際調査報告 (条約第 21 条(3))

1つ以上の構造化データの集合である構造化データ群を含んだデータベースを管理するDBMSが、1つ以上の非構造化データの集合である非構造化データ群を管理し並列分散処理を実行する複数の実行ノードと通信可能である。複数の実行ノードの各々は、DBMSが管理する構造化データ群のうちの少なくとも1つの構造化データと複数の実行ノードが管理する非構造化データ群のうちの少なくとも1つの非構造化データとを基に処理用データを作成し処理用データを用いて処理を行うことができるノードである。制御情報 (DBMSへの参照要求で指定された2以上の構造化データの提供の制御に関する情報) を基に、2以上の構造化データが、DBMSによりデータベースから取得され、DBMSから、複数の実行ノードのうちの2以上の実行ノードへ提供される。

明 細 書

発明の名称：

データ処理システム、データベース管理システム及びデータ処理方法

技術分野

[0001] 本発明は、概して、データ処理に関し、特に、分散処理に関する。

背景技術

[0002] 企業等は、複数の業務データの集合である業務データ群を管理し、業務データ群は、複数の非構造化データである非構造化データ群と、複数の構造化データである構造化データ群とに大別される。非構造化データ群が業務データ群に占める割合は、一般に、構造化データ群と比較して圧倒的に多い（例えば業務データ群の約80%が非構造化データ群である）と言われている。しかし、非構造化データ群は、ファイルサーバ等に放置され、作成直後以外は活用されないことが多い。このような現状に加え、企業等に蓄積される非構造化データ群の量が更に増加傾向にあることから、非構造化データ群を分析しビジネスへ二次活用したいといったニーズが存在し得る。

[0003] しかし、非構造化データだけを用いて、有意な業務分析を実施することは困難である。

[0004] そこで、非構造化データだけでなく構造化データを用いた分析（以下、組合せ分析）を行うことが知られている。組合せ分析は、一般に、定型的な分析ではなく、非定型的な分析である。非定型的な分析は、例えば、大量の業務データ群の中から有意なデータを探し出し分析に活用する非定型的な分析、具体的には、例えば、様々なデータを組み合わせて新たな分析用データを作成し、作成したデータの中から有意なデータを探し出す分析である。

[0005] 非構造化データ群の管理には、一般に、分散ファイルシステムが用いられる。分散ファイルシステムとは、ネットワークで接続された複数のコンピュータ上にあるフォルダ群を一つの共有フォルダのサブフォルダとして扱うことで、複数のコンピュータのいずれからでも、フォルダ群にアクセス可能と

なるシステムである。実際のフォルダの構造とは異なった仮想的な構造を組み立てることで、分散して存在するファイルやフォルダがまるで一つのコンピュータで管理されているかのように見える。なお、分散ファイルシステムは、同システムが管理するファイル（データ）を用いて分散処理を実行可能な並列分散処理基盤に組み込まれることが一般的である（例えばHadoop）。また、並列処理に関する技術として、データの分析処理をグループ抽出処理（Map処理）とデータ集約処理（Reduce処理）とに単純化したMapReduceのようなプログラミングモデルが知られている（例えば特許文献1）。

[0006] 一方、構造化データ群の管理には、一般に、データベースが用いられ、データベースは、データベース管理システム（以下、DBMS）により管理される。

[0007] 従って、構造化データ群と非構造化データ群のそれぞれから必要なデータを取得して組合せ分析を行うような分析アプリケーションを開発しようとする、分析アプリケーションがDBMSと並列分散処理基盤（分散ファイルシステム）の両方にアクセスしなければならず、故に、分析アプリケーションの開発者が、両システム（DBMSと並列分散処理基盤（分散ファイルシステム））に関する深い知識を有していることが必要となる。加えて、分析アプリケーションの機能として、非構造化データと構造化データを組み合わせて分析用のデータを作成する機能までも提供する必要がある。特に、非定型的な分析業務では、決められた組み合わせ方法に従って特定のデータを用いて分析用のデータを作成すればよいという訳ではない。取り扱うファイル（データ）は特定のファイル（データ）に限らず、また、ファイル（データ）の形式によって組み合わせ方法も動的に変更する必要がある。そのため、分析アプリケーションの開発コストは高い（例えば開発者の負担が大きい）。

先行技術文献

特許文献

[0008] 特許文献1：US2008/0086442

発明の概要

発明が解決しようとする課題

[0009] そこで、組合せ分析については、DBMS及び並列分散処理基盤の一方を使用することが考えられ、特に、並列分散処理基盤を使用することがよいと考えられる。並列分散処理基盤は、大量データを効率的に分散処理するためのソフトウェア基盤であり、コンピュータ（以下、実行ノード）を大量に並べ、並列処理を行うことで、大量データの分析処理を高速化することが可能となる。並列分散処理基盤を使用する分析アプリケーション開発者にとっては、どのような実行ノードが存在し実行ノード間でどのように分散処理が行われるのかを意識する必要がなく、故に、分析アプリケーションの開発コストを低減できる。

[0010] しかし、どのような実行ノードが存在し実行ノード間でどのように分散処理が行われるのかを意識する必要が無いという点は、分析アプリケーションの開発コストの低減に貢献する一方で、分析アプリケーションのレスポンスの遅延の原因となる。特に、非定型的な分析では、短いスパンで様々な異なるデータの組み合わせを試す必要があり、分析アプリケーションのレスポンスが一層遅延することが考えられる。

[0011] この種の問題は、分析アプリケーション以外のアプリケーションからの要求の処理（つまり、分析処理以外の処理）についても存在し得る。

課題を解決するための手段

[0012] 構造化データ群を含んだデータベースを管理するDBMSが、非構造化データ群を管理し並列分散処理を実行する複数の実行ノードと通信可能である。複数の実行ノードの各々は、受けた構造化データ（DBMSが管理する構造化データ群のうちの少なくとも1つの構造化データの一例）と対象の非構造化データ（複数の実行ノードが管理する非構造化データ群のうちの少なくとも1つの非構造化データの一例）とを基に処理用データを作成し処理用データを用いて処理を行うことができるノードである。制御情報（DBMSへの参照要求で指定された2以上の構造化データの提供の制御に関する情報）

を基に、2以上の構造化データが、DBMSによりデータベースから取得され、DBMSから、複数の実行ノードのうち2以上の実行ノードへ提供される。

- [0013] なお、「受けた構造化データ」とは、DBMSから配置された構造化データと、実行ノードがDBMSに対して送信要求を送信することによりDBMSから送信された構造化データのいずれであってもよい。「対象の非構造化データ」とは、受けた構造化データに対して組み合わされる非構造化データ（分析用データの作成に必要な非構造化データ）でよい。「2以上の実行ノード」は、「複数の実行ノード」の全てであってもよいし一部であってもよい。

発明の効果

- [0014] アプリケーションの開発コストの低減と、高い処理性能（アプリケーションのレスポンスの高速化）の両方を実現できる。

図面の簡単な説明

- [0015] [図1]実施例1に係る分散処理システムの概要を示す。
[図2]データベースサーバのハードウェア構成を示す。
[図3]分散処理システム内における命令及びデータの装置間のやり取りを示す。
。
[図4]分散処理システムの分散分析処理の概略の流れを示す。
[図5]解析処理（図4のS405）の流れを示す。
[図6]拡張要求の一部の一例を示す。
[図7]第1配置情報の一例を示す。
[図8]第2配置情報の一例を示す。
[図9]図6の拡張要求中の参照要求に従って取得される行データを含んだitem表の一例を示す。
[図10]図9の表の格納の一例を示す。
[図11]作成処理（図4のS406）の流れを示す。
[図12]配置ジョブ対応表の一例を示す。

- [図13]実行処理（図4のS407）の流れを示す。
- [図14]取得処理（図13のS1503又はS1504）の流れを示す。
- [図15]配置処理（図13のS1503又はS1505）の流れを示す。
- [図16]取得処理及び配置処理の一例の一部を示す。
- [図17]取得処理及び配置処理の一例の別の一部を示す。
- [図18]実施例2に係る第2準備情報の一例を示す。
- [図19]実施例2に係る取得処理の流れを示す。
- [図20]実施例2に係る一時表の一例を示す。
- [図21]実施例3に係る分散処理システムの概要を示す。
- [図22]実施例3に係る負荷情報の一例を示す。
- [図23]実施例3に係る第2配置情報の一例を示す。
- [図24]負荷監視処理の流れを示す。
- [図25]実施例3に係る配置ジョブ対応表の一例を示す。

発明を実施するための形態

- [0016] 以下、図面を参照して幾つかの実施例を説明する。
- [0017] 以下の説明では、種々の対象（例えば、筐体、ポート）のID（識別情報）として、番号、名称等が使用されるが、それらは互いに置換可能であってもよい、それらのうちの少なくとも1つに代えて又は加えて他種の情報が使用されてもよい。
- [0018] また、以下の説明では、「プログラム」を主語として処理を説明する場合があるが、プログラムは、プロセッサ（例えばCPU（Central Processing Unit）によって実行されることで、定められた処理を、適宜に記憶資源（例えばメモリ）および／または通信インターフェイスデバイス（例えば通信ポート）等を用いながら行うため、処理の主語がプロセッサ（又はそのプロセッサを有する装置）とされてもよい。プログラムを主語として説明された処理は、プロセッサあるいはそのプロセッサを有する装置（計算機またはストレージ装置等）が行う処理としてもよい。また、プロセッサは、処理の一部または全部を行うハードウェア回路を含んでもよい。プログラムは、プロ

グラムソースから各コントローラにインストールされてもよい。プログラムソースは、例えば、プログラム配布計算機または記憶メディアであってもよい。

[0019] また、以下の説明では、「構造化データ」とは、構造定義されているデータを言う。構造化データは、表や索引のようなスキーマ（例えば、顧客データ、経理データ、商品データ、在庫データ等の表）でもよいし、表中の値や索引中の値のようなスキーマ部分でもよい。

[0020] 一方、「非構造化データ」とは、構造定義されていないデータ（例えば、主にリレーショナルモデルに適合しないデータモデルに分類されるデータ）を言う。例えば、非構造化データは、報告書等の文章データ、電子メールやソーシャルメディア等のデジタルコンテンツ、コールセンタで記録する通話記録等がよく、或いはそれらのうちの少なくとも1つのうちの一部でもよい。非構造化データは、例えばファイルでよい。なお、完全な構造定義を持たないデータがしばしば「半構造化データ」と呼ばれることがあるが、本明細書では、構造化データ以外のデータが非構造化データであり、構造化データはデータベースで管理されるデータであり、非構造化データは並列分散処理基盤で管理されるデータである。すなわち、本明細書では、分析処理で扱われるデータは、構造化データと非構造化データとに大別される。

実施例 1

[0021] 図1は、実施例1に係る分散処理システムの概要を示す。

[0022] 分散処理システム101は、データ処理システムの一例であり、データベースサーバ20と、並列分散処理基盤301とを有する。並列分散処理基盤301が、複数の実行サーバ40を有する。データベースサーバ20に複数の実行サーバ40が通信可能に接続される。

[0023] データベースサーバ20は、第1のインタフェースデバイスの一例であるネットワークインタフェース（以下、NIF）22と、第2のインタフェースデバイスの一例であるストレージインタフェース（以下、SIF）23と、データベース管理システム（DBMS）21とを有する。NIF22に複

数の実行サーバ40が例えばネットワーク11を介して通信可能に接続され、SIF23に、構造化データ群999（複数の構造化データ）を含むデータベース998を記憶した1以上の外部記憶装置30が通信可能に接続される。各外部記憶装置30が、記憶デバイス31を有し、記憶デバイス31が、データベース998の一部であるデータベース部を記憶する。言い換えれば、1以上のデータベース部の集合が、データベース998である。DBMS21は、データベース998を管理する。DBMS21は、解析部211と、作成部212と、実行部213とを有する。

[0024] 実行サーバ40は、インタフェースデバイスの一例であるNIF41と、実行ノード42と、2次記憶デバイス43とを有する。実行ノード42は、分析処理（例えばグループ抽出処理やデータ集約処理）を行う。複数の実行サーバ40が有する複数の2次記憶デバイス43が、非構造化データ群（複数の非構造化データの集合）1001を記憶し、非構造化データ群1001が分散ファイルシステム48により管理される。複数の実行ノード42の各々は、受けた構造化データと対象の非構造化データとを基に分析用データを作成し分析用データを用いて分析処理を行うことができるノードである。

[0025] 「サーバ」は、物理計算機資源に基づくエンティティの一例である。例えば、サーバは、計算機の一部であり、計算機は、物理計算機資源（例えば、物理記憶デバイス、物理I/Oデバイス及びそれらに接続された物理プロセッサ）を有する物理計算機でもよいし、物理計算機資源が論理的に分割された論理計算機資源が割り当てられた仮想計算機でもよい。全ての実施例において、「サーバ」は、物理計算機（物理サーバ）とする。

[0026] 「ノード」も、物理計算機資源に基づくエンティティの一例である。例えば、ノードは、計算機の一部でもよいし（サーバと同義でもよいし）、計算機（サーバ）でコンピュータプログラムが実行されることにより実現される機能であってもよい。全ての実施例において、「ノード」は、サーバでコンピュータプログラムが実行されることにより実現される機能とする。

[0027] 以下、本実施例で行われる処理の概要を説明する。

[0028] DBMS 21 (解析部 211) は、参照要求 (例えば SQL で書かれたクエリ) 1004 と第 1 配置情報 521 とを含んだ処理要求 1002 を、NIF 22 を介して受信する。第 1 配置情報 521 は、その参照要求 1004 で指定された条件に該当する 2 以上の構造化データの配置先とされ得る 2 以上の実行ノード 42 が指定された情報であり、その 2 以上の実行ノード 42 の各々に関する情報 (例えば、各実行ノードの IP アドレス及びポート番号) を含む。参照要求 1002 と第 1 配置情報 521 はそれぞれ個別に (異なるタイミングで) 受信されてもよいし、第 1 配置情報 521 は、外部 (例えば後述の管理ノード) から受信せず予め DBMS 21 が管理していてもよい。或いは、第 1 配置情報 521 に代えて、後述の第 2 配置情報が、外部から DBMS 21 に送られてもよいし DBMS 21 により予め管理されていてもよい (つまり第 1 配置情報 521 は無くてもよい)。

[0029] DBMS 21 (解析部 211) は、参照要求 1004 と第 1 配置情報 521 を解析し、第 1 配置情報 521 の解析結果としての第 2 配置情報を作成する。第 2 配置情報の内容は、第 1 配置情報の内容と同一でよく、具体的には、参照要求 1004 で指定された条件に該当する 2 以上の構造化データの配置先とされ得る 2 以上の実行ノード 42 に関する情報 (例えば、各実行ノードの IP アドレス及びポート番号) を含んでよい。第 1 配置情報 521 と第 2 配置情報の違いは、情報の形式でよい。例えば、第 1 配置情報 521 が、テキストファイルのような非構造化形式の情報であり、第 2 配置情報が、その非構造化データに基づく表のような構造化形式の情報でよい。第 1 配置情報 521 及び第 2 配置情報のうち少なくとも第 2 配置情報が、制御情報の一例である。

[0030] DBMS 21 (作成部 212) は、参照要求 1004 の解析結果を基に、参照要求の処理手順である実行計画 (例えばクエリ実行計画) を作成する。具体的には、例えば、DBMS 21 は、構造化データの 1 以上の属性のうちの特定期間属性についての条件が参照要求において指定されているか否かと、実行ノード 42 について配置する構造化データに関する配置条件を表す情報を

第2配置情報が含んでいるか否かを基に、実行するジョブの数と、複数のジョブの各々に割り当てる処理を決定し、その決定の結果に従う実行計画を作成する。ジョブは、「スレッド」又は「プロセス」のような処理単位でよい。ジョブとしては、構造化データをデータベース988から取得する取得処理を行う取得ジョブ、取得処理により取得された構造化データを第2配置情報に従い実行ノード42に配置する配置処理を行う配置ジョブ、もしくは、取得処理とその取得処理により取得された構造化データの配置処理との両方を行う取得／配置ジョブのうちの少なくとも1つがある。2以上の取得処理により2以上の構造化データが取得されるが、2以上の取得処理（2以上のジョブ）にそれぞれ異なる2以上のデータベースページ（以下、ページ）が割り当てられる。これにより、2以上の取得処理において取得元のページが重複しない。なお、「ページ」（データベースページ）とは、データベースが特定属性の値に基づき論理的に分割されることにより得られたデータベースエレメントの一例である。

[0031] DBMS 21（実行部213）は、実行計画に基づき、データベースから構造化データを取得する取得処理のジョブを実行する。取得処理のジョブ（取得／配置ジョブ、又は、取得ジョブ）が複数ある場合、DBMS 21（実行部213）は、それら複数のジョブを非同期かつ並列に実行する。また、DBMS 21（実行部213）は、実行計画で指定された（第2配置情報に記載された）実行ノードに対し、取得した構造化データを配置する配置処理のジョブを実行する。配置処理のジョブ（取得／配置ジョブ、又は、配置ジョブ）が複数ある場合、DBMS 21（実行部213）は、それら複数のジョブを非同期かつ並列に実行する。なお、配置処理は、全ての取得対象の構造化データ（参照要求で指定された条件に適合する全ての構造化データ）の取得の完了を待たずに、適宜（例えば、一定量の構造化データが取得される都度に）実行されてもよい。

[0032] 複数の実行ノード42のうちの2以上の実行ノード42がそれぞれ並列に2以上の分析処理を実行する。分析処理では、分析に用いるデータは、非構

造化データ及び構造化データのうち、又は、非構造化データ及び構造化データの組合せである。分析に用いるデータが非構造化データの場合、実行ノード42は、対象の非構造化データを用いた分析処理を実行する。分析に用いるデータが構造化データの場合、実行ノード42は、DBMS21によりその実行ノード42に配置された構造化データ（実行ノード42を有する実行サーバ40内の記憶デバイスに配置されている構造化データ）を用いた分析処理を実行する。分析に用いるデータが非構造化データ及び構造化データの組合せの場合、実行ノード42は、対象の非構造化データと配置された構造化データとを組み合わせた分析処理を実行する。

[0033] DBMS21と並列分散処理基盤301の連携によって実現される分散処理システム101において、下記2つの処理が行われる。

（処理1）分析用データの準備（実行ノードへの構造化データの配置）

（処理2）2以上の実行ノード42での分散分析処理

[0034] この2つの処理をそれぞれ高速化することで、結果的に、分析アプリケーションの性能向上（レスポンス高速化）を実現することが期待できる。

[0035] DBMS21が、参照要求1004と第2配置情報を基に、構造化データの取得処理と、取得した構造化データの実行ノードへの配置処理をそれぞれ並列かつ非同期に実行する。これにより、参照要求1004に適合する構造化データを2以上の実行ノードに分割して配置するまでの処理時間の短縮（高速化）が期待される。これにより、（処理1）の性能向上が期待される。

[0036] また、或る分析要求（例えば、「値Aを含むデータを用いた分析を実行する」という要求）を受信した実行ノード42に対し、DBMS21が、第2配置情報を基に、構造化データ（例えば、値Aを含む行データ）をその実行ノード42に直接配置できる。これにより、実行ノード42が分析処理を実行する際の構造化データへのアクセス性能が向上する。具体的には、例えば、実行ノード間で構造化データを受け渡しするという実行ノード間通信の発生を避けることができる。結果として、（処理2）の性能向上が期待される。

- [0037] また、分散処理システム101（特に、例えば、分散ファイルシステム48並びに並列分散処理基盤301）は信頼性（例えば、障害対応、並びに分析性能の向上）を確保するため、分析に用いるデータ（構造化データ及び非構造化データ）の複製を、複数の実行ノード42に格納することが考えられる。例えば当該複製（オリジナルを含む）は3以上の実行ノード42に格納されることがある。本実施例によれば、実行ノード42が分析処理に用いる構造化データを、その実行ノード42に直接配置することができる。このため、分散ファイルシステム48並びに並列分散処理基盤301の信頼性確保が目的とされても、データの複製（オリジナルの複製）の格納先となる実行ノードは、現用系の実行ノード42と待機系の実行ノード42の2つで充分となる。つまり、本実施例によれば、分散処理システムを構成するハードウェアリソース（記憶デバイス）の削減が可能となる。
- [0038] 図2は、データベースサーバ20のハードウェア構成を示す。
- [0039] データベースサーバ20は、NIF22と、SIF23と、主記憶デバイス51と、2次記憶デバイス52と、光学ドライブ25と、表示デバイス202と、キーボード203と、バス1005を介してそれらに接続されたCPU (Central Processing Unit) 24とを有する。
- [0040] NIF22、SIF23、光学ドライブ25、表示デバイス202及びキーボード203は、それぞれI/Oデバイスの一例である。光学ドライブ25は、DVD (Digital Versatile Disk) のような光学記憶媒体201からデータを読み出すことができる。表示デバイス202は、出力デバイスの一例であり、第2配置情報522の少なくとも一部を表示してよい。キーボード203は、入力デバイスの一例であり、キーボード203を介して第1配置情報が手動で入力されてもよい。
- [0041] ネットワーク11は、例えば、LAN (Local Area Network) のような通信ネットワークであり、NIF22は、LANを介した通信のためのLANC (LANコントローラ) のようなネットワークインタフェースである。
- [0042] 主記憶デバイス51及び2次記憶デバイス52は、それぞれ記憶デバイス

の一例である。主記憶デバイス51は、例えば揮発性（又は不揮発性）のメモリであり、CPU24により実行されるDBMS（解析部211、作成部212及び実行部213）が記憶される。また、主記憶デバイス51は、各処理部の処理を実行する際に頻繁に使用するデータを一時的に格納すること等に用いられるワークエリア26を有する。2次記憶デバイス52は、不揮発性の記憶デバイスであり、例えばHDD（Hard Disk Drive）又はSSD（Solid State Drive）である。2次記憶デバイス52は、外部から入力された第1配置情報521、及び、第1配置情報521の変換後の情報である第2配置情報522（第1配置情報521をDBMS21で管理しやすいように第1配置情報521が表形式に変更された情報）を記憶する。第1配置情報521は外部記憶装置30に格納されてもよい。また、第2配置情報522は、主記憶デバイス51（ワークエリア26）に格納されてもよいし、外部記憶装置30に格納されてもよい。

[0043] SIF23は、ネットワークを介して外部記憶装置30に接続する。ネットワークがSAN（Storage Area Network）である場合には、SIF23はHBA（Host Bus Adapter）でよい。DBMS21が実行されると、必要に応じて、SIF23を介して外部記憶装置30にI/O要求が発行される。ネットワークは、例えば、FC（Fibre Channel）プロトコルが適用されるSANであるが、他種のネットワークでもよい。例えば、ネットワークは、IP（Internet Protocol）ネットワークでよく、SIF23と外部記憶装置30との間がiSCSIプロトコルで通信されてもよい。

[0044] 外部記憶装置30は、データベースサーバ20が管理するデータを格納し、データベースサーバ20からのI/O要求に従うデータのI/Oを行う。ここで言う「データ」は、必ずしも構造化データでなくてもよい。外部記憶装置30は、インタフェースデバイス32、記憶デバイス31及びそれらに接続された制御デバイス（プロセッサ）33を有する。インタフェースデバイス32は、ネットワークに接続される。制御デバイス33は、データベースサーバ20からのI/O要求を処理する。記憶デバイス31は、不揮発性

の記憶デバイスであり、データベースサーバ20からのI/O要求に従って読み書きされるデータを管理するデータベース部311を格納する。データベース部311は、データベース988（図1参照）の一部である。データベース988は、表及び索引を含み、データベース部311は、データベース988内の表の一部としての表1006と、データベース988内の索引の一部としての索引1007とを含む。記憶デバイス31は、単一のHDD、SSD又は他種の記憶デバイスであってもよいし、RAID (Redundant Arrays of Inexpensive (or Independent) Disks) グループのような記憶デバイス群でもよい。

[0045] 図3は、分散処理システム101内における命令及びデータの装置間のやり取りを示す。図4は、分散処理システム101の分散分析処理の概略の流れを示す。以下、図3及び図4を参照して、分散処理システム101の処理全体（分散分析処理）の流れを説明する。

[0046] 並列分散処理基盤301が、複数の実行サーバ40（実行ノード42）とデータベースサーバ20（DBMS21）とに接続されたサーバ（ノード）である管理サーバ49（管理ノード44）を有する。管理サーバ49（管理ノード44）は、アプリケーションサーバ60（分析アプリケーション61）からネットワーク10を介して分析要求を受け付ける。アプリケーションサーバ60は、分析アプリケーション61を実行するサーバである。管理ノード44、実行ノード42及びDBMS21が、異なる複数のサーバでそれぞれ実行されるが、管理ノード44、実行ノード42及びDBMS21のうちの少なくとも2つは、同一サーバで実行されてもよいし、同一物理サーバ上の異なる複数の仮想サーバでそれぞれ実行されてもよい。なお、図1には、並列分散処理基盤301内に管理サーバ49（管理ノード44）は示されていない。なぜなら、図1は、実施例1の概要の説明のための図であり、概要の説明に、管理サーバ49（管理ノード44）の説明は必須ではないためである。

[0047] 分散分析処理の開始（S400）により、管理ノード44は、分析要求を

解析する（図4のS401）。分析要求の解析結果（例えば分析要求の構成等）を基に、管理ノード44は、分析要求が、非構造化データのみを対象にした要求であるか、或いは、少なくとも構造化データを対象にした要求であるかを判断する（図4のS402）。

[0048] 分析要求が、非構造化データのみを対象にした要求の場合、S402の判断結果は否定となる（S402：No）。具体的には、例えば、分散ファイルシステム48内に1以上の非構造化データ（コールセンタへの問合せメール又はテキスト）が格納されている場合に、分析要求が、「過去1年間にコールセンタに問い合わせのあった商品毎に問合せ回数の集計」であれば、S402の判断結果は否定となる。この分析要求によれば、非構造化データから商品名を抽出して商品毎に商品名を集計すればよい。一方、分析要求が、少なくとも構造化データを対象にした要求の場合、S402の判断結果は肯定となる（S402：Yes）。具体的には、例えば、分析要求が、「過去1年間の商品毎の販売台数と、過去1年間にコールセンタに問い合わせのあった商品毎に問合せ回数から、商品毎の利益率を算出」であれば、S402の判断結果は肯定となる。この分析要求によれば、データベースから商品毎に販売台数を取得し、分散ファイルシステム内の1以上の非構造化データ（コールセンタへの問合せメール又はテキスト）から商品名を抽出して商品毎に商品名を集計し、商品毎に販売台数（例えば、販売台数×商品単価＝総費用）と商品名数（例えば、商品名数×所定単価＝サポート費用）とを基に商品毎に1台あたりの利益率を算出すればよい。

[0049] S402：Noの場合、管理ノード44は、2以上の実行ノード42の各々について、分析命令を作成し、作成した分析命令を送信する（S409）。2以上の実行ノード42の各々は、分析命令を受け、分析命令に従い分析処理を行い、分析処理の結果を管理ノード44に送信する（S410）。管理ノード44は、2以上の実行ノード42からそれぞれ受けた2以上の分析処理結果を統合した統合データ（分析要求の処理結果）を作成し、統合データをアプリケーションサーバ60へ送信する（S412）。

[0050] S402: Yesの場合、管理ノード44は、分析要求の解析結果を基に、DBMS21への参照要求1004と、2以上の実行ノード42の各々への分析命令とを作成する(図4のS403及びS404)。また、管理ノード44は、複数の実行ノード42に関する情報(例えば、実行ノード42毎に、IPアドレス、実行ノード名(ホスト名)、ポート番号及びファイルシステム情報(管理している非構造化データ(例えばフォルダ名、ファイル名))等を含んだ情報)を管理しており、その情報と分析要求の解析結果とを基に、第1配置情報521を生成することができるものとする。複数の実行ノード42に関する情報が第1配置情報521それ自体でもよい。管理ノード44は、参照要求1004と第1配置情報521とを含んだ処理要求1002をDBMS21に送信できる。なお、上述したように、管理ノード44は、参照要求1004をDBMS21に送信し、その後、第1配置情報521をDBMS21に送信してもよい。また、第1配置情報521又は第2配置情報522が予めDBMS21で管理されていてもよく、DBMS21への第1配置情報521の送信が必須でなくてもよい。DBMS21で予め管理されている第1配置情報521又は第2配置情報522は、適宜アップデートされてよい。

[0051] 管理ノード44は、参照要求1004と第1配置情報521とをネットワーク11を通じてDBMS21に送信し、且つ、2以上の実行ノード42の各々にネットワーク11を通じて分析命令を送信する。ネットワーク11は、アプリケーションサーバ60と管理ノード44との通信に使用されるネットワーク10でもよいし、ネットワーク10から独立した別のネットワークでもよい。

[0052] 参照要求1004及び第1配置情報521を受信したDBMS21において、解析部211が、解析処理として、例えば、参照要求1004を解析し、その解析結果と第1配置情報521とを基に第2配置情報522を生成する(図4のS405)。次に、作成部212が、作成処理として、例えば、解析結果と第2配置情報522とに基づき、参照要求1004の処理手順で

ある実行計画を作成する（図4のS406）。実行部213は、実行処理として、例えば、実行計画に基づき参照要求1004を実行する（図4のS407）。

[0053] 一方、2以上の実行ノード42の各々が、管理ノード44からの分析命令にตอบสนองして分析処理を実行する（図4のS408）。各実行ノード42は、分析処理において、対象の非構造化データを取得する。各実行ノード42は、分析処理を継続する上で構造化データを扱う必要となった場合、その構造化データがDBMS21によって配置されるのを待つ。各実行ノード42は、DBMS21によって構造化データを配置されている場合、分析処理を継続する（例えば構造化データと非構造化データのマージ処理を行う）。各実行ノード42は、分析命令のตอบสนองとして、分析処理の結果を管理ノード44に送信する。

[0054] すなわち、S402: Yesの場合、管理ノード44はDBMS21に参照要求1004を送信するものの、管理ノード44は、参照要求1004（処理要求1002）のตอบสนองとして、参照要求1004に従い取得された構造化データを受信しない。管理ノード44は、例えば、参照要求1004（処理要求1002）のตอบสนองとして、所定種類のตอบสนอง（例えば、参照要求1004を受信したことを表すตอบสนอง、又は、参照要求1004の処理に成功したか失敗したかを表すตอบสนอง）を受信してよい。参照要求1004に従い取得された構造化データは、DBMS21から実行ノード42に配置される。実行ノード42は、配置された構造化データと取得した非構造化データとを組み合わせた分析用データを生成し、その分析用データに従う分析結果を管理ノード44に送信する。管理ノード44は、分析命令がそれぞれ送信された2以上の実行ノード42の全てから分析結果を受信した場合に、それらの分析結果を統合した統合データ（統合分析結果）を作成し、統合データを分析要求のตอบสนองとしてアプリケーションサーバ60（分析アプリケーション61）に送信する（図4のS412）。

[0055] 以上が、分散処理システム101の処理全体（分散分析処理）の流れであ

る。

[0056] 以上の流れによれば、分析処理で扱うデータの内容（構造化データのみ、または非構造化データのみ、または構造化データと非構造化データ）に関係なく、管理ノード44が、分析アプリケーション61から分析要求を受け付け、その分析要求の応答を分析アプリケーションに返す。これにより、分析アプリケーション61の開発コストの低減が期待できる。

[0057] また、以上の流れによれば、構造化データがDBMS21から管理ノード44の介在無しに直接実行ノード42に配置される。これにより、管理ノード44が各実行ノード42から分析結果を受けるまでの時間の短縮が期待でき、以って、高い分析性能（分析アプリケーションのレスポンスの高速化）が期待できる。

[0058] 従って、以上の流れによれば、分析アプリケーション61の開発コストの低減と、高い分析性能（分析アプリケーション61のレスポンスの高速化）の両方の実現が期待できる。

[0059] 以下、分散処理システム101内のDBMS21が行う処理を詳細に説明する。

[0060] 図5は、解析処理（図4のS405）の流れを示す。

[0061] 解析部211は、参照要求1004を受信し（S501）、受信した参照要求1004を解析する（S502）。そして、解析部211は、解析結果を基に、第1配置情報521を受信予定であるか否かを判断する（S503）。S503の判断結果が肯定の場合（S503:Yes）、解析部211は、第1配置情報521を受信し（S504）、第1配置情報521を基に第2配置情報522を作成し（S505）、作成部212を実行する（S506）。一方、S503の判断結果が否定の場合（S503:No）、解析部211は、S504及びS505無しに作成部212を実行する（S506）。

[0062] 以下、解析処理を詳細に説明する。

[0063] DBMS21への要求の送信には、所定のAPI（Application Programm

ing Interface)、例えば、JDBCのようなAPIを用いることができる。JDBCのようなAPIで提供されるクラスを用いて参照要求1004(例えばSQLで記述されたクエリ(SELECT文))の実行を指示する要求の作成が可能である。そのような一般的な要求によれば、参照要求1004が実行され、参照要求1004の実行結果(参照要求の実行により取得された構造化データ)が、参照要求の送信元に提供される。

[0064] しかし、本実施例では、参照要求1004の実行により取得された構造化データは、DBMS21により実行ノード42に配置されなければならない。そこで、本実施例では、DBMS21への要求として、拡張要求(拡張された要求)が使用される。図6は、拡張要求の一部の一例を示す。拡張要求6001は、図1の説明で述べた処理要求1002の一例である。図6の拡張要求6001において、[2]~[4]が拡張部(拡張された要求部分)である。すなわち、拡張部[2]は、参照要求1004の実行結果を準備しておけばよく実行結果の返信は不要であることの指示である。また、拡張部[2]に、参照要求1004が記述されている。拡張部[3]は、参照要求1004の実行結果の送信先を、本拡張要求6001を発行したノード(本実施例では管理ノード44)にするのではなく指定された配置先とするこの指示である。拡張部[4]は、参照要求1004の実行結果を用意し、その実行結果を指定された配置先に配置することの指示である。なお、DBMS21は、管理ノード44からの要求に対する応答として、例えば、参照要求1004に従い取得された全ての構造化データの配置が終了した場合に成功を意味する応答を返してよいし、少なくとも1つの構造化データについて取得又は配置に失敗した場合に失敗を意味する応答を返してもよいし、管理ノード44から要求を受けてから一定時間が経過した時点で、取得された全ての構造化データの配置が終了していない場合にタイムアウトエラーを返してもよい。タイムアウトエラーは、後述の実施例2で特に有効であると考えられる。なぜなら、実施例1では、構造化データはDBMS21から実行ノード42に配置されるが(プッシュ型:実行ノード42からの要求無しの提

供)、実施例2では、実行ノード42からDBMS21が送信要求を受けた場合にその送信要求に応答して構造化データが実行ノード42に送信されるからである(プル型:実行ノード42からの要求に伴う提供)。

[0065] このように、拡張要求6001により、DBMS21の動作を制御できる。プッシュ型のケース(実施例1及び後述の実施例3)では、拡張要求6001には、DBMS21が実行ノード42に構造化データを配置することの指示が記述される。プル型のケース(実施例2)では、拡張要求6001には、DBMS21が構造化データをどのように保持しておくかの指示が記述され、そのケースでは、DBMS21は、実行ノード42から送信要求を受信した場合に、その送信要求に応答して、保持している構造化データをその実行ノード42に送信する。

[0066] 拡張要求6001の拡張部[3]には、第1配置情報521を埋め込むことができる。第1配置情報521は、例えば図7に示すように、HTML(HyperText Markup Language)又はXML(Extensible Markup Language)のようなマークアップ言語で表すことができる。また、拡張部[3]には、第1配置情報521が分割してDBMS21に送信されることを意味する情報を記述することもできる(すなわち、管理ノード44から第1配置情報521が分割してDBMS21に送信されてもよい)。また、第1配置情報521は、独立したファイル(データ)として管理ノード44からDBMS21に送信されてもよい。

[0067] 第1配置情報521には、2以上の実行ノード42の各々について、実行ノードの構成に関する構成情報(例えば、通し番号、IPアドレス、ホスト名、ポート番号)の他に、図7に例示するように、実行ノード42毎に、その実行ノード42に配置されるべき構造化データの配置条件(例えば、表名、列名及び値のような1以上の構造化データ属性のうち少なくとも1つにより定義された条件)が記述されてもよい。例えば、「配置条件」は、特定の表(本実施例では後述のitem表)と同表を構成する列名(本実施例ではcategory列)と同列に記載される値の組み合わせにより表現される。図7の例に

よれば、実行ノード：1（通し番号「1」の実行ノード）に対して、item表の「category」列の値が「microwave ovens」である行データ（構造化データ）が配置される。なお、配置条件が設定されていない実行ノード（例えば実行ノード：32）については、配置条件が設定されている実行ノードに配置する行データ（構造化データ）以外の行データが配置される。配置条件は必ずしも第1配置情報521に記述されなくてもよい。

[0068] 実行ノード42に配置条件が対応付けられていないケースでは、DBMS21が、実行ノード42間で構造化データの受け渡しが発生しないよう、同一の構造化データを複数の実行ノード42に配置しておくことが考えられる。しかし、そうすると、複数の実行ノード42の記憶領域を多く消費することになり、コストが高くなると考えられる。一方、上述のように、実行ノード42に配置条件を対応付けておくことで、同一の構造化データを無駄に多くの実行ノード42に配置する必要が無い。分散処理システム101の信頼性を確保する目的があるとしても、例えば、現用系の実行ノード42と待機系の実行ノード42の2つの実行ノード42に同一の構造化データが配置されれば充分であると考えられる。

[0069] 解析部211は、図7の第1配置情報521（非構造化データ）を、図8のような第2配置情報522（構造化データ）に変換する。第2配置情報522は、図8のように一体として保持されてもよいし、分割して保持されてもよい。図8の第2配置情報522において、実行ノード毎に行データ（レコード）が存在し、各行データにおいて、NO811が、通し番号であり、IP812が、IPアドレスであり、HOST813が、ホスト名（ノード名）であり、PORT814が、ポート番号であり、TABLE815が、表名であり、COLUMN816が、カラム名であり、VALUE817が、値である。行データにおいて、TABLE815、COLUMN816及びVALUE817が、配置条件を表す情報である。なお、第1配置情報521では、「IP」及び「HOST」のうちの一方を省略可能であり、そのため、第2配置情報522では、IP812及びHOST813の一方が設

定されない行データが存在することがある。また、第1配置情報では、「NO」及び「PORT」は省略可能でよい。

[0070] また、第2配置情報522には、必ずしも配置条件が記述されない。配置条件が記述されていない第1配置情報521があり得るためである。すなわち、このケースでは、第1配置情報は、各実行ノードについて構成情報を含むが配置条件を含まない簡易第1配置情報である。解析部211は、簡易第1配置情報の受信後、簡易第1配置情報から第2配置情報を作成する。簡易第1配置情報から作成された第2配置情報を、以下、簡易第2配置情報とすることができる。

[0071] また、第1配置情報521において、実行ノードの構成情報のうちの少なくとも1つの情報要素が省略されていた場合の第2配置情報522の作成方法は、例えば次の通りである。或る実行ノード42について「NO」が第1配置情報521において省略されていた場合、解析部211が、「NODE」タグの数をカウントし、適切な値（例えばカウント値）を、その実行ノード42に対応したNO811として第2配置情報522に設定する。或る実行ノード42について「PORT」が第1配置情報521において省略されていた場合、解析部211が、予め設定されているデフォルトの値を、その実行ノード42に対応したPORT814として第2配置情報522に設定する。図8の例では、PORT814のデフォルト値は「30000」である。或る実行ノード42について「IP」があるが「HOST」が第1配置情報521において省略されている場合、解析部211は、その実行ノードについてHOST813のフィールドを空欄のまま（例えばNULL値を設定）とする。或る実行ノードについて「IP」が省略されているが「HOST」がある場合、解析部211は、DNS (Domain Name System) サーバ等にアクセスし、当該「HOST」に対応するIPアドレスを割り出し、その実行ノードについてそのIPアドレスをIP812として第2配置情報522に設定する。記述された「HOST」からIPアドレスの割り出しが出来ない場合は、解析部211は、第1配置情報521の送信元にエラー（JDBC

であれば「例外」)を返却する。或る実行ノードについて「I P」及び「H O S T」の両方が第1配置情報521において省略されている場合、解釈部211は、上記と同様エラー(JDBCであれば「例外」)を返却する。

[0072] 第2配置情報522の作成後、解析部211は、作成部212を実行する。

[0073] 上述したように、作成部212は、解析部211による解析結果と、第2配置情報522とを基に、実行するジョブの数と、各ジョブに割り当てる処理とを決定する。その際、2以上の取得処理にそれぞれ対応した2以上の取得元が同一のページとならないように(取得元のページが重複しないように)、取得処理にページが割り当てられる。

[0074] 図9は、図6の拡張要求6001中の参照要求1004に従って取得される行データを含んだitem表の一例を示す。図10は、図9の表の格納の一例を示す。

[0075] 図9のitem表901は、商品毎に、行データを有し、各行データは、itemID911(商品の識別子)、category912(商品のカテゴリ)、maker913(商品の発売元)、price914(商品の価格)及びstock915(商品の在庫数)を有する。

[0076] item表901は、図10に示すように、少なくとも1つの外部記憶装置30が有する複数の記憶デバイス31に跨っている。以下、DBMS21が管理する構造化データを格納する領域を、「データ領域」と総称することがある。「データ領域」は、外部記憶装置30、記憶デバイス31、データベース998、データベース部311、表領域1555、及びページ1112のうちの少なくとも1つを意味する。1つの記憶デバイス31に対し1つのデータベース部311が割り当てられる。また、1つのデータベース部311に対し1つの表領域(1つの表(例えばitem表901)は複数の表領域1555にまたがって配置されているものとする)が割り当てられる。表領域1555は、8つのページ1112から構成される。「ページ」とは、データベースエレメントの一例であり、DBMS21が1度のI/O要求で取得す

るデータの単位である。1つのページ1112は、「itemID」の値に従い分割された複数（図10の例によれば10,000）の行データで構成される。つまり本実施例では、DBMS21は、1度のI/O要求でitem表901から10,000の行データを取得できる。1つの外部記憶装置30は、6つの記憶デバイス31を有する。つまり、本実施例では、4つの外部記憶装置30にわたってitem表901の192万の行データが格納されていることになる。

[0077] 図11は、作成処理（図4のS406）の流れを示す。

[0078] 作成部212は、参照要求1004の解析結果（解析部211により作成された解析結果）から、参照要求に条件が設定されているか否かを判断する（S1301）。参照要求1004に設定されている条件は、本実施例ではWHERE句とする。

[0079] S1301の判断結果が否定の場合（S1301：No）、作成部212は、第2配置情報522に記載された実行ノード毎に配置条件が設定されているか否かを判断する（S1302）。

[0080] S1302の判断結果が否定の場合（S1302：No）、作成部212は、S1301の判断結果が否定（或いは、S1301の判断結果が肯定でも後述のS1305の判断結果が否定）且つS1302の判断結果が否定というパターン（パターン1）に基づき、実行するジョブの数を決定し（S1303）、各ジョブに割り当てる処理を決定し、且つ、決定した処理をジョブに割り当て（S1304）、実行部213を実行する（S1313）。S1303及びS1304の処理により、実行計画が作成される。S1304では、取得処理を行うジョブには取得元のページ1112が割り当てられる。取得処理を行うジョブには、他の取得処理を行うジョブに割り当てられるページ1112と異なるページ1112が割り当てられるようにする。つまり、同一のページ1112が異なるジョブに割り当てられないようにする。これにより、2以上の取得処理が並列且つ非同期に実行可能である。また、S1304では、配置処理を行うジョブに対して配置先の実行ノードの割り当てが行われる。なお、取得処理を行うジョブは、取得ジョブであることも

取得／配置ジョブであることもある。配置処理を行うジョブは、配置ジョブであることも取得／配置ジョブであることもある。

- [0081] S 1 3 0 2 の判断結果が肯定の場合 (S 1 3 0 2 : Y e s) 、作成部 2 1 2 は、S 1 3 0 1 の判断結果が否定 (或いは、S 1 3 0 1 の判断結果が肯定でも後述の S 1 3 0 5 の判断結果が否定) 且つ S 1 3 0 2 の判断結果が肯定というパターン (パターン 2) に基づき、実行するジョブの数を決定し (S 1 3 0 6) 、各ジョブに割り当てる処理を決定し、且つ、決定した処理をジョブに割り当て (S 1 3 0 7) 、実行部 2 1 3 を実行する (S 1 3 1 3) 。 S 1 3 0 6 及び S 1 3 0 7 の処理により、実行計画が作成される。 S 1 3 0 7 では、取得処理を行うジョブに対するページ 1 1 1 2 の割り当てや、配置処理を行うジョブに対する配置先実行ノードの割り当てが行われる。なお、取得処理を行うジョブは、取得ジョブである。配置処理を行うジョブは、配置ジョブである。
- [0082] S 1 3 0 1 の判断結果が肯定の場合 (S 1 3 0 1 : Y e s) 、作成部 2 1 2 は、条件 (WHERE 句) がデータ領域への格納ルールに関係のある特定属性 (本実施例では「itemID」) についての条件を含む否かを判断する (S 1 3 0 5) 。 S 1 3 0 5 の判断結果が否定の場合 (S 1 3 0 5 : N o) 、 S 1 3 0 2 が行われる。
- [0083] S 1 3 0 5 の判断結果が肯定の場合 (S 1 3 0 5 : Y e s) 、作成部 2 1 2 は、S 1 3 0 2 と同じ判断を行う (S 1 3 0 8) 。
- [0084] S 1 3 0 8 の判断結果が否定の場合 (S 1 3 0 8 : N o) 、作成部 2 1 2 は、S 1 3 0 5 の判断結果が肯定且つ S 1 3 0 8 の判断結果が否定というパターン (パターン 3) に基づき、実行するジョブの数を決定し (S 1 3 0 9) 、各ジョブに割り当てる処理を決定し、且つ、決定した処理をジョブに割り当て (S 1 3 1 0) 、実行部 2 1 3 を実行する (S 1 3 1 3) 。 S 1 3 0 9 及び S 1 3 1 0 の処理により、実行計画が作成される。 S 1 3 1 0 では、取得処理を行うジョブに対するページ 1 1 1 2 の割り当てや、配置処理を行うジョブに対する配置先実行ノードの割り当てが行われる。なお、取得処理

を行うジョブは、取得ジョブであることも取得／配置ジョブであることもある。配置処理を行うジョブは、配置ジョブであることも取得／配置ジョブであることもある。

[0085] S 1 3 0 8 の判断結果が肯定の場合 (S 1 3 0 8 : Y e s) 、作成部 2 1 2 は、 S 1 3 0 5 の判断結果が肯定且つ S 1 3 0 8 の判断結果も肯定というパターン (パターン 4) に基づき、実行するジョブの数を決定し (S 1 3 1 1) 、各ジョブに割り当てる処理を決定し、且つ、決定した処理をジョブに割り当て (S 1 3 1 2) 、実行部 2 1 3 を実行する (S 1 3 1 3) 。 S 1 3 1 1 及び S 1 3 1 2 の処理により、実行計画が作成される。 S 1 3 1 2 では、取得処理を行うジョブに対するページ 1 1 1 2 の割り当てや、配置処理を行うジョブに対する配置先実行ノードの割り当てが行われる。なお、取得処理を行うジョブは、取得ジョブである。配置処理を行うジョブは、配置ジョブである。

[0086] 以上のように、特定属性についての条件が参照要求において指定されているか否か (S 1 3 0 1 及び S 1 3 0 5) と、第 2 配置情報が実行ノード毎に配置条件を含んでいるか否か (S 1 3 0 2 、 S 1 3 0 8) を基に、処理が、複数のパターンのうちのいずれかのパターンに該当し、該当したパターンに基づき、実行するジョブの種類 (取得／配置ジョブ、又は、取得ジョブ及び配置ジョブ) と、実行するジョブの数と、複数のジョブの各々に割り当てる処理が決定される。これにより、実行するジョブの種類及び数を適切にすることが期待できる。なお、第 2 配置情報には配置条件が必ずしも含まれなくてもよいので、 S 1 3 0 2 及び S 1 3 0 8 は必ずしも行われなくてもよい。その場合、パターン 2 及びパターン 4 は無くてもよい。また、参照要求の構成が予め決まっていれば S 1 3 0 1 及び S 1 3 0 5 の判断が不要なケースがある場合には、パターン 1 及びパターン 3 が同一パターンでもよいし、パターン 2 及びパターン 4 が同一パターンでもよい。

[0087] 以下、パターン 1 ~ パターン 4 の各々について詳細に説明する。なお、以下の説明では、図 9 及び図 1 0 の例が採用されている。

[0088] <パターン1>

[0089] 参照要求1004は、条件(WHERE句)が設定されていない参照要求(S1301の判断結果が否定となる場合の参照要求)1004、又は、WHERE句は設定されているがデータ領域への格納ルールに関係のある特定属性(本実施例では「itemID」)についてのWHERE句が設定されていない参照要求(S1301の判断結果が肯定となるがS1305の判断結果が否定となる場合の参照要求)1004である。前者の参照要求1004の一例は、下記(a)であり、後者の参照要求1004の一例は、下記(b)である。

(a) SELECT * FROM item

(b) SELECT * FROM item WHERE item.stock >= 10

[0090] そして、パターン1では、第2配置情報522に配置条件が設定されていない。

[0091] パターン1の場合、取得処理と配置処理は単一のジョブ(取得/配置ジョブ)で実行することが可能となる。取得/配置ジョブの採用により、実行するジョブの数を節約できる。取得/配置ジョブの数は、実行ノード42(例えば現用系の実行ノード42)の数以上でも未満でもよい。本実施例では、パターン1について決定される取得/配置ジョブの数は、並列分散処理基盤301が有する実行ノード42(例えば現用系の実行ノード42のように配置先となり得る実行ノード42)の数と同じ32とする。以下、便宜的に、32個の取得/配置ジョブには、通し番号1~32が割り振られているとし、例えば通し番号[1]の取得/配置ジョブを「取得/配置ジョブ[1]」と言う。また、番号[1-1-1]のページ1112を「ページ[1-1-1]」と言う。

[0092] 作成部212は、各取得/配置ジョブに、取得元のページ1112を割り当てる。例えば、各取得/配置ジョブに、同数(同量)のページ1112且つ他のいずれの配置/取得ジョブに割り当てられていないページ1112が割り当てられてよい(各ページ1112のサイズは同じでよい)。具体的には、例えば、取得/配置ジョブ[1]に、6個のページ[1-1-1]~[

1-1-6] が割り当てられ、取得／配置ジョブ [2] には、別の6個のページ [1-1-7]、[1-1-8]、及び [1-2-1] ~ [1-2-4] が割り当てられてよい。1つの取得／配置ジョブには、連続したページ 1 1 1 2 が割り当てられてもよいし非連続のページ 1 1 1 2 が割り当てられてもよい。

[0093] 作成部 2 1 2 は、各取得／配置ジョブに、取得された構造化データの配置先実行ノードを割り当てる。本実施例では、取得／配置ジョブの数と実行ノードの数が同じなので、取得／配置ジョブと実行ノードが 1 : 1 で割り当てられてよい。なお、取得／配置ジョブは、割り当てられたページ 1 1 1 2 から構造化データを取得し、取得した構造化データ（割り当てられた実行ノードに配置されるデータ）の一時格納領域として主記憶デバイス 5 1 上のワークエリア 2 6 を用いることが考えられる。しかし、ワークエリア 2 6 のサイズには限りがある。従って、取得／配置ジョブは、ワークエリア 2 6 に格納されているデータが一定量以上となった時点で、適宜ワークエリア 2 6 から構造化データを実行ノード 4 2 に配置する必要がある。ジョブ毎に、割り当てられる一時格納領域（ワークエリア 2 6 内の領域）の容量や、データの配置を実行するタイミング（例えば、一時格納領域の使用率が Y % 以上になった時）などは、DBMS 2 1 の定義情報に記載されてもよい。取得／配置ジョブに対する一時格納領域の割り当ても、作成部 2 1 2 により実行される。なお、「一時格納領域の使用率」とは、一時格納領域の容量に対する、一時格納領域に格納されている行データの総量、の割合である。

[0094] <パターン 2>

[0095] 該当する参照要求 1 0 0 4 としては、パターン 1 と同じであるが、第 2 配置情報 5 2 2 については、配置条件が設定されている点で、パターン 1 と異なる。

[0096] 配置条件が設定されている場合、取得／配置ジョブに代えて、取得ジョブと配置ジョブが採用される。取得処理を行うジョブと配置処理を行うジョブが互いに独立しているので、各々のジョブ数を調整することで、適切な並列

処理が期待できる。配置ジョブの数は、並列分散処理基盤 3 0 1 が有する実行ノード 4 2（例えば現用系の実行ノード 4 2 のように配置先となり得る実行ノード 4 2）の数と同じ 3 2 でよい。取得ジョブの数は、item 表のページ 1 1 1 2 の数と同じ 1 9 2 でよい。DBMS 2 1 の定義として、同時に実行可能なジョブの数の上限がある場合、ジョブ数の上限から配置ジョブの数を引いた数が取得ジョブの数のように、動的に取得ジョブの数が、作成部 2 1 2 により調整されてよい。また、ジョブの上限数が実行ノード数未満の場合、なるべく並列に処理を実行可能なジョブの数が多くなるよう、配置ジョブの数と実行ジョブの数が均等であってもよい。パターン 2 の説明においては、配置ジョブ数 3 2、取得ジョブ数 6 4 とする。

[0097] 作成部 2 1 2 は、各取得ジョブに対してページ 1 1 1 2 を割り当てられ、また、各配置ジョブに対して配置先実行ノードが割り当てられる。これに関しては、取得／配置ジョブが取得ジョブと配置ジョブに代わっただけで、後はパターン 1 と同様でよい。例えば、各取得ジョブについて、割り当てられるページ 1 1 1 2 の数がほぼ同量で、かつ、他の取得ジョブが取得元のページ 1 1 1 2 が重複しないようにページ 1 1 1 2 が割り当てられてよい。

[0098] 作成部 2 1 2 は、取得ジョブと配置ジョブの対応付け、例えば、どの取得ジョブが取得した行データ（構造化データ）をどの配置ジョブに割り当てられた一時格納領域に格納するのかを決める。作成部 2 1 2 が行う処理としては、配置条件が設定された第 2 配置情報と、配置ジョブへの実行ノードの割り当て結果とを基に、配置ジョブ対応表 1 4 0 0（図 1 2 参照）を作成する。配置ジョブ対応表 1 4 0 0 は、第 2 配置情報（表）に「ジョブ ID」列が追加されたものである。ジョブ ID は、配置ジョブの識別子である。取得ジョブは、配置ジョブ対応表 1 4 0 0 を参照することにより、取得された行データをどの配置ジョブ（どの配置ジョブに対応する一時格納領域に）に格納すればよいかわかる。

[0099] <パターン 3>

[0100] 参照要求 1 0 0 4 は、データ領域への格納ルールに関係のある特定属性（

本実施例では「itemID」) についての条件 (WHERE句) が設定されている参照要求 (S 1 3 0 5 の判断結果が肯定となる場合の参照要求) 1 0 0 4 である。このような参照要求 1 0 0 4 の例は、例えば下記 (c) 及び (d) である。

(c) SELECT * FROM item WHERE item.itemID > 960000 AND item.stock >= 10

(d) SELECT * FROM item WHERE item.itemID > 100000 AND item.itemID <= 260000

[0101] そして、パターン 3 では、第 2 配置情報 5 2 2 に配置条件が設定されていない。

[0102] パターン 3 の処理は、基本的にはパターン 1 の処理と同様の処理となる。例えば、参照要求 (c) によると、条件によって割り当て対象となるページ 1 1 1 2 の数がパターン 1 に比べて半数に減少することで、取得／配置ジョブに割り当てるページ 1 1 1 2 数も減少することになるが、それ以外の違いはない。

[0103] 一方、参照要求 (d) によると、割り当て対象となるページ 1 1 1 2 が、1 6 個のページ [1-2-3] ~ [1-4-2] となる。これは、実行ノードの数 (3 2) 以下である。このような場合、参照要求 (a) ~ (c) のいずれかに基づく実行計画とは異なり、取得ジョブと配置ジョブがそれぞれに取得処理と配置処理を実行することになる。具体的には、例えば、配置ジョブ数を 3 2、取得ジョブ数を 1 6 とする。パターン 3 では、第 1 配置情報 5 2 1 に配置条件が設定されていないため、配置ジョブ対応表 1 4 0 0 を作成することができず、このため、作成部 2 1 2 が取得ジョブと配置ジョブを明示的に対応付ける必要がある。そこで、例えば、1 6 の取得ジョブのうちの第 1 の取得ジョブが、行データの配置先として、3 2 の配置ジョブのうちの第 1 及び第 2 の配置ジョブを交互に切り替えることを表す実行計画を、作成部 2 1 2 が作成する。

[0104] <パターン 4 >

- [0105] 参照要求1004は、データ領域への格納ルールに関係のある特定属性（本実施例では「itemID」）についての条件（WHERE句）が設定されている参照要求（S1305の判断結果が肯定となる場合の参照要求）1004である。そして、パターン4では、第2配置情報522に配置条件が設定されている。
- [0106] 第2配置情報522に配置条件が設定されているので、パターン4の処理は、パターン2の処理と同様である。
- [0107] 上述したように、作成部212は、実行計画作成後、実行部213を実行する（S1313）。
- [0108] 図13は、実行処理（図4のS407）の流れを示す。
- [0109] 実行部213は、実行計画から、指定された数の取得ジョブと配置ジョブ、もしくは取得／配置ジョブを用意する（S1501）。
- [0110] 用意されたジョブが取得／配置ジョブであれば（S1502：Yes）、実行部213は、各取得／配置ジョブを実行する（S1503）。2以上の取得／配置ジョブが並列に実行されてよい。一方、用意されたジョブが取得ジョブ及び配置ジョブであれば（S1502：No）、実行部213は、各取得ジョブと各配置ジョブを実行する（S1504及びS1505）。2以上の取得ジョブと2以上の配置ジョブが並列に実行されてよい。
- [0111] 実行部213は、各ジョブの結果に基づく処理結果を、参照要求1004（処理要求1002）の送信元である管理ノード44に送信する（S1506）。その処理結果は、例えば、成功又は異常を表す結果でよく、少なくとも取得処理によりデータベース998から取得された構造化データを含まない。そのような構造化データはDBMS21（配置処理）によりDBMS21から実行ノード42に直接配置されるからである。
- [0112] 図14は、取得処理（図13のS1503又はS1504）の流れを示す。
- [0113] 取得ジョブは、第2配置情報522に配置条件が設定されていれば（S1600：Yes）、S1601～S1607を行い、第2配置情報522に

配置条件が設定されていなければ（S1600：No）、S1611～S1615を行う。

[0114] <第2配置情報522に配置条件が設定されているケース>

[0115] 取得ジョブは、配置ジョブ対応表からいずれか実行ノードを選択し（S1601）、割り当てられたページ1112から、行データを取得する（S1602）。取得された行データは、参照要求で指定された条件に適合する構造化データの一例である。取得ジョブは、取得された行データが、選択された実行ノードに対応した配置条件に適合するか否かを判断する（S1603）。S1603の判断結果が否定であれば（S1603：No）、取得ジョブは、配置ジョブ対応表から次の実行ノードを選択し（S1604）、再度S1603を行う。S1603の判断結果が肯定であれば（S1603：Yes）、取得ジョブは、選択された実行ノードが割り当てられた配置ジョブの一時格納領域に、行データを格納する（S1605）。取得ジョブに割り当てられたページ1112に未取得の行データがあれば（S1606：Yes）、取得ジョブは、配置ジョブ対応表から次の実行ノードを選択し（S1607）、次の行データを取得する（S1602）。取得ジョブに割り当てられたページ1112に未取得の行データが無ければ（S1606：No）、取得ジョブが終了する。

[0116] <第2配置情報522に配置条件が設定されていないケース>

[0117] 配置条件が設定されていなければ、配置ジョブ対応表1400が存在しない。このため、取得ジョブは、予め作成部212で指定された1以上の実行ノード（配置ジョブ）から1つを選択し（S1611）、割り当てられたページ1112から行データを取得し（S1612）、選択した配置ジョブの一時格納領域に、取得した行データを格納する（S1613）。割り当てられたページ1112に未取得の行データがある場合（S1614：No）、取得ジョブは、先ほど選択した実行ノード（配置ジョブ）とは別の実行ノード（配置ジョブ）を選択し（S1615）、同様に行データを格納する。この処理を繰り返し、予め作成部212で指定された1以上の実行ノード（配

置ジョブ) 全てに一度ずつ行データを格納するまでの繰り返しを1周目とするならば、1周目の処理完了後は2周目、3周目…と、割り当てられたページ1112の全ての行データを取得するまで、処理が繰り返えされる。

[0118] 図15は、配置処理(図13のS1503又はS1505)の流れを示す。

[0119] 配置ジョブは、配置タイミングが指定されていれば(S1701:Yes)、S1702~S1705を行い、配置タイミングが指定されていなければ(S1701:No)、S1706~S1709を行う。配置タイミングは、配置ジョブに対応した一時格納領域の状態について定義されていてもよい(例えば、一時格納領域の使用率がY%以上になった時)、時間長又は時刻等で定義されてもよい。

[0120] <配置タイミングが指定されているケース>

[0121] 配置ジョブは、全ての取得ジョブが終了しているか否かを判断する(S1702)。S1702の判断結果が肯定の場合(S1702:Yes)、配置ジョブは、同配置ジョブに対応した一時格納領域内の全ての行データを、その配置ジョブに割り当てられている実行ノード42に配置する(S1705)。それにより、配置ジョブが終了する。

[0122] S1702の判断結果が否定の場合(S1702:No)、配置ジョブは、配置タイミングが満たされているか否かを判断する(S1703)。S1703の判断結果が肯定の場合(S1703:Yes)、配置ジョブは、同配置ジョブに対応した一時格納領域内の行データを、その配置ジョブに割り当てられている実行ノード42に配置し(S1704)、S1702に戻る。S1703の判断結果が否定の場合(S1703:No)、配置ジョブは、S1704の実行無しにS1702に戻る。

[0123] <配置タイミングが指定されていないケース>

[0124] 配置ジョブは、全ての取得ジョブが終了しているか否かを判断する(S1706)。S1706の判断結果が肯定の場合(S1706:Yes)、配置ジョブは、同配置ジョブに対応した一時格納領域内の全ての行データを、

その配置ジョブに割り当てられている実行ノード42に配置する（S1709）。それにより、配置ジョブが終了する。

[0125] S1706の判断結果が否定の場合（S1706：No）、配置ジョブは、配置ジョブに対応する一時格納領域の状態（例えば使用率）が所定の条件を満たしているか否かを判断する（S1707）。S1707の判断結果が肯定の場合（S1707：Yes）、配置ジョブは、一時格納領域内の行データを、その配置ジョブに割り当てられている実行ノード42に配置し（S1708）、S1706に戻る。S1707の判断結果が否定の場合（S1707：No）、配置ジョブは、S1708の実行無しにS1706に戻る。なお、この段落で言う「所定の条件」とは、DBMS21によって一意に決められている条件であり、「配置タイミング」ではない。

[0126] 以上が、取得ジョブ及び配置ジョブの処理である。取得／配置ジョブの処理は、取得ジョブの処理（図14）と配置ジョブの処理（図15）の組合せとすることができる。ただし、その組合せからは、第2配置情報522に配置条件が設定されているか否かの判断は除かれる。そのような判断は不要のためである。つまり、取得／配置ジョブではS1600の判断結果が否定の場合（S1600：No）のS1611～S1615は実行されない。

[0127] 以上の取得ジョブ及び配置ジョブによれば、一例として、図16及び図17に示すような取得処理及び配置処理が行われる。なお、図16及び図17は、「第2配置情報522に配置条件が設定されていないケース」かつ「配置タイミングが指定されているケース」の一例である。フェーズ1～3、…、フェーズX及びX+1は、時間的な遷移を示す。すなわち、取得ジョブ[1]（取得ジョブ1601）が、取得ジョブ[1]に割り当てられたページ1112から1個ずつ行データを取得して配置ジョブ[1]及び[2]（配置ジョブ1602）のいずれかの一時格納領域1603に格納する。複数の行データが配置ジョブ[1]及び[2]にそれぞれ対応した2つの一時格納領域1603に均等に格納されていく。フェーズX+1において、配置ジョブ[1]について配置タイミング（例えば、一時格納領域の使用量が50%を

超えた)が満たされ、配置ジョブ [1] が、配置ジョブ [1] の一時格納領域 1603 から、行データを、配置ジョブ [1] に割り当てられた実行ノード [1] (実行ノード 42) に配置する。

実施例 2

[0128] 以下、実施例 2 を説明する。その際、実施例 1 との相違点を主に説明し、実施例 1 との共通点については説明を省略又は簡略する。また、実施例 2 における要素が、実施例 1 の要素に対応しかつその要素と全く同じであれば、実施例 1 の要素の参照符号と同一の参照符号を使用し、実施例 1 の要素に対応するがその要素と機能等が異なっていれば、図示の有無に関わらず、実施例 1 の要素の参照符号に子符号「B」を追記するものとする。

[0129] まず、実施例 2 の概要を説明する。

[0130] 実施例 2 では、DBMS 21B が、能動的に構造化データ (行データ) を実行ノード 42B に提供しない。実施例 2 では、DBMS 21B は、構造化データの送信要求を実行ノード 42B から受信し、その送信要求に応答して、構造化データをその実行ノード 42B に送信する。少なくとも実施例 2 では、DBMS が実行ノードからの要求無しに構造化データを実行ノードに提供することを、構造化データの「配置」と言い、DBMS が実行ノードからの要求に応答して構造化データを実行ノードに提供することを、構造化データの「送信」と言う。「配置」及び「送信」はそれぞれ「提供」の一例でよい。

[0131] 概要として、実施例 2 では、管理ノード 44B が、2 以上の実行ノード 42B の各々に、その実行ノード 42B が送信要求において指定すべき識別子を通知する。2 以上の実行ノード 42B の各々は、管理ノード 44B から通知された識別子を指定した送信要求を DBMS 21B へ送信する。DBMS 21B は、その送信要求で指定された識別子から特定される構造化データを、その実行ノード 42B に送信する。「識別子」は、構造化データを直接的又は間接的に特定できる識別子であり、例えば、行データの識別子でもよいし、行データの集合としての構造化データに割り振られた識別子でもよい。

以下の説明では、行データの集合は、「一時表」と呼ばれ、DBMS 21B（実行部 213B）により一時表が作成され、実行ノード 42Bからの送信要求では、その一時表の識別子が指定されるとする。一時表は、実行ノード 42B毎に作成される。例えば、管理ノード 44Bが、DBMS 21Bに対して、実行ノード 42Bの一時表に割り振る識別子を各実行ノード 42Bについて指定し、且つ、2以上の実行ノード 42Bの各々に対して、その実行ノード 42Bについて作成される一時表の識別子を通知する。その通知は、例えば、一時表の識別子を含んだ分析命令を管理ノード 44Bから実行ノード 42Bに送信すること、であってもよい。

[0132] 実施例 2 では、DBMS 21B が構造化データを配置しないため、配置ジョブは不要である。実施例 2 では、実行部 213B が、実行ノード 42B 毎に一時表を作成し、一時表を、ワークエリア 26 又は 2 次記憶デバイス 52 に格納する。このようにして、実行ノード 42B 毎に一時表が DBMS 21B により管理される。その後、実行ノード 42B が、管理ノード 44B から指定された識別子を指定した送信要求を DBMS 21B に送信する。DBMS 21B は、一時表の識別子が指定された送信要求を実行ノード 42B から受信した場合、その識別子から特定される一時表を、その実行ノード 42B に送信する。送信要求の送信及び一時表の送信は、管理ノード 44B の介在無しに、DBMS 21B と実行ノード 42B 間で直接行われる。

[0133] 実施例 1 では、DBMS 21 は、管理ノード 44 から参照要求 1004 を受信する都度にデータベース 998 から行データを取得するが、実施例 2 では、データベース 998 から取得された行データを一時表として管理しておく。このため、実施例 2 では、DBMS 21B が、繰り返し同じ参照要求 1004 を受信しても、一時表が既に用意されているので、参照要求 1004 を受信する都度にデータベース 998 から行データを取得しなくてよい。これにより、実行ノード 42B へ提供する構造化データの準備にかかる時間を短縮できる。一方で、実施例 1 では、実行ノード 42 が、構造化データを用いて分析処理を実行しようとした時点で、既に必要な構造化データが配置済

み（その実行ノード42を有する実行サーバ40に構造化データが存在している）可能性があるため、実行ノード42の分析処理の実行時間の短縮が期待できる。

[0134] 以下、実施例2を詳細に説明する。

[0135] 実施例2では、第1配置情報521に代えて第1準備情報521Bが採用され、第2配置情報522に代えて第2準備情報522Bが採用される。第1及び第2準備情報521B及び522Bのうち少なくとも第2準備情報522Bが、制御情報の一例である。第1準備情報521Bは、管理ノード44Bのような外部からDBMS21Bに入力される非構造化形式の情報であり、第2準備情報522Bは、第1準備情報521Bの変換後の情報（構造化形式の情報）でよい。第1準備情報521Bは、実行ノード42Bの各々に関する情報（例えばIPアドレス、ホスト名及びポート番号）を含んでおらず、一時表の識別子と一時表に含まれる行データのデータ条件とを一時表毎に含む。従って、第2準備情報522Bは、図18に示すように、一時表の識別子（TABLEID（表ID）1811）と一時表に含まれる行データのデータ条件（TABLE（表名）1812、COLUMN（列名）1813、及びVALUE（値）1814）とを一時表毎に含む情報である。データ条件は、送信される行データに関する条件である。

[0136] 図19は、実施例2に係る取得処理の流れを示す。

[0137] 実施例1に係る取得ジョブの処理の説明（図14の説明）において、「第2配置情報522」を「第2準備情報522B」に読み替え、「配置条件」を「データ条件」に読み替え、「一時格納領域への行データの格納」を「一時表への行データの格納」に読み替えることで、当業者であれば、実施例2に係る取得ジョブの処理の説明（図19の説明）を理解することができる。具体的には以下の通りである。

[0138] すなわち、取得ジョブは、第2準備情報522Bにデータ条件が設定されていれば（S2300：Yes）、S2301～S2307を行い、第2準備情報522Bにデータ条件が設定されていなければ（S2300：No）

、S 2 3 1 1 ~ S 2 3 1 5 を行う。

[0139] <第2準備情報5 2 2 Bにデータ条件が設定されているケース>

[0140] 取得ジョブは、第2準備情報5 2 2 BからいずれかのTABLEID（一時表）を選択し（S 2 3 0 1）、割り当てられたページ1 1 1 2から、行データを取得する（S 2 3 0 2）。取得ジョブは、取得された行データが、選択された一時表に対応したデータ条件に適合するか否かを判断する（S 2 3 0 3）。S 2 3 0 3の判断結果が否定であれば（S 2 3 0 3 : N o）、取得ジョブは、第2準備情報5 2 2 Bから次の一時表を選択し（S 2 3 0 4）、再度S 2 3 0 3を行う。S 2 3 0 3の判断結果が肯定であれば（S 2 3 0 3 : Y e s）、取得ジョブは、選択された一時表に、行データを格納する（S 2 3 0 5）。取得ジョブに割り当てられたページ1 1 1 2に未取得の行データがあれば（S 2 3 0 6 : Y e s）、取得ジョブは、第2準備情報5 2 2 Bから次の一時表を選択し（S 2 3 0 7）、S 2 3 0 2を行う。取得ジョブに割り当てられたページ1 1 1 2に未取得の行データが無ければ（S 2 3 0 6 : N o）、取得ジョブが終了する。

[0141] 以上の取得処理により、「第2準備情報5 2 2 Bにデータ条件が設定されているケース」では、図20に示すような一時表2 0 0 0が完成する。図20の一時表2 0 0 0は、図9に示したitem表の一部（ページ1 1 1 2）から取得された行データの集合である。故に、一時表2 0 0の各レコード（行データ）は、図9の行データと同じ属性、すなわち、itemID 2 0 1 1、category 2 0 1 2、maker 2 0 1 3、price 2 0 1 4、及びstock 2 0 1 5を有する。

[0142] <第2準備情報5 2 2 Bにデータ条件が設定されていないケース>

[0143] 取得ジョブは、第2準備情報5 2 2 BからいずれかのTABLEID（一時表）を選択し（S 2 3 1 1）、割り当てられたページ1 1 1 2から、行データを取得する（S 2 3 1 2）。取得ジョブは、取得された行データを、選択した一時表に格納する（S 2 3 1 3）。取得ジョブに割り当てられたページ1 1 1 2に未取得の行データがあれば（S 2 3 1 4 : Y e s）、取得ジョブは、第2準備情報5 2 2 Bから次の一時表を選択し（S 2 3 1 5）、S 2

3 1 2を行う。取得ジョブに割り当てられたページ 1 1 1 2 に未取得の行データが無ければ (S 2 3 1 4 : N o) 、取得ジョブが終了する。このケースでは、複数の一時表に複数の行データが均等に分散される。

実施例 3

[0144] 以下、実施例 3 を説明する。その際、実施例 1 及び 2 との相違点を主に説明し、実施例 1 及び 2 との共通点については説明を省略又は簡略する。また、実施例 3 における要素が、実施例 1 の要素に対応しかつその要素と全く同じであれば、実施例 1 の要素の参照符号と同一の参照符号を使用し、実施例 1 の要素に対応するがその要素と機能等が異なっていれば、図示の有無に関わらず、実施例 1 の要素の参照符号に子符号「C」を追記するものとする。

[0145] まず、実施例 3 の概要を説明する。

[0146] 図 2 1 は、実施例 3 に係る分散処理システムの概要を示す。

[0147] 実施例 3 では、分散処理システム 1 0 1 C において、実行サーバ 4 0 C が、負荷監視部 4 6 を更に有する。DBMS 2 1 C が、負荷取得部 2 1 4 を更に有する。

[0148] 負荷監視部 4 6 は、ハードウェアでもよいし、実行サーバ 4 0 C の CPU (プロセッサ) によりプログラムが実行されることで発揮される機能でもよい。負荷監視部 4 6 は、実行ノード 4 2 を実行する実行サーバ 4 0 C の負荷値 (例えば、実行サーバ 4 0 C の CPU 利用率、メモリ使用量、ネットワークトラフィック、電力使用量等の種々のメトリックのメトリック値) を監視する。負荷監視部 4 6 は、負荷値の監視結果を、DBMS 2 1 C の負荷取得部 2 1 4 に通知する。DBMS 2 1 C の負荷取得部 2 1 4 が、負荷値の監視結果を、ワークエリア 2 6 及び 2 次記憶デバイス 5 2 のうちの少なくとも一方に格納される負荷情報に反映する。負荷情報は、実行ノード (実行サーバ) 毎の負荷値 (少なくとも 1 種類のメトリック値) を表す。負荷情報の形式は、非構造化形式でもよいし構造化形式でもよい。

[0149] 実施例 3 では、実行ノード 4 2 (実行サーバ 4 0 C) 毎の負荷に基づき、実行ノード 4 2 毎に、配置するデータ量 (行データの総量) が制御される。

例えば、負荷の小さい実行サーバ40C内の実行ノード42に、負荷の大きい実行サーバ40C内の実行ノード42よりも多くの行データが配置されるよう制御される。これにより、実行サーバ40Cの負荷に応じた量の行データが実行ノード42に配置されるので、並列分析処理の性能向上が期待できる。

[0150] 以下、実施例3を詳細に説明する。

[0151] 図24は、負荷監視処理の流れを示す。

[0152] 負荷監視処理の開始(S2700)により、負荷監視部46は、DBMS21Cが停止するかもしれない負荷監視部46が配置された実行サーバ40Cが停止するまで、繰り返し(例えば定期的に)負荷値を取得し(S2701)、一定期間の複数の負荷値の平均値を算出する(S2702)。平均値に代えて又は加えて、一定期間の複数の負荷値に基づく他種の値が算出されてもよい。

[0153] 負荷監視部46は、算出された平均値が所定の条件を満たすか否かを判断する(S2703)。「算出された平均値が所定の条件を満たす」とは、例えば、算出された平均値が所定の値以上であることでもよいし、算出された平均値と過去に算出された1つ以上の平均値に基づく値との差分が所定の差分以上でもよいし、算出された平均値と過去に算出された1つ以上の平均値から特定される平均値増加度合が所定の度合以上でもよい。

[0154] S2703の判断結果が肯定の場合(S2703:Yes)、負荷監視部46は、算出された平均値をDBMS21Cの負荷取得部214に送信する。

[0155] 以上が、負荷監視処理である。この処理において、負荷値の取得間隔や、平均値の算出区間(どの時点の負荷値からどの時点の負荷値までの平均値を算出するか)、比較対象の平均値の取得タイミングなどの数値は管理者任意の値でよい。

[0156] 図22は、負荷情報の一例を示す。

[0157] 負荷情報2201は、上述したように、ワークエリア26及び2次記憶デ

バイス52のうちの少なくとも一方に格納される情報であり、実行ノード（実行サーバ）毎の負荷値（少なくとも1種類のメトリック値）を表す。図22の例によれば、負荷情報は、表形式であり、実行ノード42（実行サーバ40C）毎に、NO（通し番号）2211、IP（IPアドレス）2212、及び、負荷値（CPU（CPU利用率）2213及びNETWORK（ネットワークトラフィック）2214等）を有する。

[0158] DBMS21Cの負荷取得部214は、負荷監視部46から負荷値（平均値）を受信した場合、その負荷監視部46を有する実行サーバ40Cの実行ノード42に対応した行に、受信した負荷値（平均値）（CPU利用率及びネットワークトラフィック等）を上書きする。

[0159] 図23は、実施例3に係る第2配置情報の一例を示す。

[0160] 第2配置情報522Cは、実行ノード42毎に、配置条件に代えて（又は加えて）、RATIO（配置割合）2315を有する。実行ノードの配置割合は、その実行ノードの負荷値（負荷情報から特定された負荷値）に基づいて算出された値である。実行ノードの配置割合は、複数の実行ノードに配置される複数の行データの総量に対する、その実行ノードに配置する構造化データの量の割合である。具体的には、例えば、配置割合：Kは、規定回数（例えば100（任意））中、格納を実行する回数を意味する（ $0 \leq K \leq 100$ ）。例えば、K=50であれば、2回に1回は格納がスキップされる。ただし、スキップは、連続した50回の格納後に50回連続して行われてもよく、その辺りの実装についてはDBMS21Cごとに任意でよい。配置割合に基づく制御により、複数の行データの複数の実行ノード42への配分が制御される。

[0161] 次に、実施例3に係る解析部211C、作成部212C及び実行部213C（図21参照）の各々の処理を説明する。

[0162] 解析部211Cと、実施例1に係る解析部211との主な相違点は、次の通りである。解析部211Cは、第2配置情報522Cを作成する際に負荷情報2201（図22参照）を参照し、実行ノード42毎の負荷値に基づいて、実行ノード42毎の配置割合を算出し、実行ノード42毎の配置割合を

含んだ第2配置情報522C（図23参照）を作成する。

- [0163] 作成部212Cと、実施例1に係る作成部212との主な相違点は、次の通りである。作成部212Cは、実行ノード42毎の配置割合を含んだ配置ジョブ対応表1400C（図25参照）を作成する。配置ジョブ対応表1400Cと、実施例1に係る配置ジョブ対応表1400（図14参照）との主な相違点は、RATIO（配置割合）2511の有無である。なお、図25の配置ジョブ対応表1400Cは、配置条件が記述されていない第1配置情報522Cに基づく表であるため、図14の配置ジョブ対応表1400のように配置条件（TABLE1215、COLUMN1216及びVALUE1217）を含んでいない。また、実施例3では、取得ジョブ及び配置ジョブが採用され、取得／配置ジョブは採用されない。
- [0164] 実行部213Cと、実施例1に係る実行部213との主な相違点は、次の通りである。取得ジョブ1601Cは、取得した行データを一時格納領域1603に格納するにあたり、その一時格納領域1603に対応した実行ノード42についての配置割合（配置ジョブ対応表1400Cに記載の配置割合2511）を参照し、その配置割合を基に、一時格納領域1603に行データを格納してよいか否かを判断する。その判断結果が肯定の場合に、取得ジョブ1601Cは、行データを一時格納領域1603に格納する。
- [0165] 以上、幾つかの実施例を説明したが、本発明は、これらの実施例に限定されるものでなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。
- [0166] 例えば、実施例1～3の少なくとも2つを組み合わせることができる。この場合、例えば、実施例1と2の組合せでは、DBMSは、取得ジョブにより取得した行データを含んだ一時表を作成し、実行ノードから送信要求を受信した場合に、その実行ノードに対応したデータ条件（例えば実施例1の第2配置情報522のような制御情報から特定された条件）と、一時表に対応したデータ条件とが一致するか否かを判断してよい。その判断結果が肯定の場合に、DBMSは、一時表を、その実行ノードに送信してよい。

[0167] また、本発明は、分析アプリケーション 61 以外のアプリケーションからの要求の処理（つまり、分析処理以外の処理）にも適用できる。

符号の説明

[0168] 101 : 分散処理システム

請求の範囲

[請求項1]

1つ以上の非構造化データの集合である非構造化データ群を管理し並列分散処理を実行する複数の実行ノードと通信可能であり1つ以上の構造化データの集合である構造化データ群を含んだデータベースを管理するデータベース管理システム（以下、DBMSと呼称）と、

前記複数の実行ノード及び前記DBMSと通信可能であり前記DBMSに参照要求を送信する管理ノードとを有し、

前記複数の実行ノードの各々は、前記DBMSが管理する構造化データ群のうちの少なくとも1つの構造化データと前記複数の実行ノードが管理する非構造化データ群のうちの少なくとも1つの非構造化データとを基に処理用データを作成し前記処理用データを用いて処理を行うことができるノードであり、

前記参照要求で指定された2以上の構造化データの提供の制御に関する情報である制御情報を基に、前記2以上の構造化データが、前記DBMSにより前記データベースから取得され、前記DBMSから、前記複数の実行ノードのうちの2以上の実行ノードへ提供される、データ処理システム。

[請求項2]

前記制御情報は、配置情報であり、

前記配置情報は、前記2以上の構造化データをそれぞれ配置する前記2以上の実行ノードの各々に関する情報を含み、

前記DBMSが、前記配置情報を基に、前記2以上の構造化データを前記2以上の実行ノードに配置する、

請求項1記載のデータ処理システム。

[請求項3]

前記DBMSが、

構造化データの1以上の属性のうち特定の属性についての条件が前記参照要求において指定されているか否かと、前記2以上の実行ノードの各々について配置する構造化データに関する条件である配置

条件を表す情報を前記配置情報が含んでいるか否かとを基に、実行するジョブの数と、複数のジョブの各々に割り当てる処理を決定し、

それぞれ処理が割り当てられた複数のジョブを実行することで、前記2以上の構造化データをそれぞれ取得する2以上の取得処理と、前記取得された2以上の構造化データをそれぞれ前記2以上の実行ノードに配置する2以上の配置処理とを、それぞれ並列かつ非同期に実行する、

請求項2記載のデータ処理システム。

[請求項4]

前記配置情報が前記配置条件を表す情報を含んでいない場合、前記複数のジョブの各々は、構造化データを取得する取得処理と取得した構造化データを実行ノードに配置する配置処理との両方を行う取得／配置処理を実行する

請求項3記載のデータ処理システム。

[請求項5]

前記データベースが前記特定の属性の値を基に複数のデータベースエレメントに分割されており、

前記DBMSは、前記2以上の取得処理を実行する2以上のジョブに、前記複数のデータベースエレメントのうちの、前記参照要求に適合する2以上のデータベースエレメントを、割り当てられるようになっており、

前記2以上の取得処理を実行する2以上のジョブの各々は、そのジョブに割り当てられたデータベースエレメントに前記2以上の構造化データの少なくとも1つがあれば、そのジョブに割り当てられたデータベースエレメントから、構造化データを取得するようになっており、

前記配置情報が前記配置条件を表す情報を含んでいなくても、前記参照要求に前記特定の属性の値に関する条件が指定されていて、その条件に適合する値に対応したデータベースエレメントの数が実行ノード数以下の場合、前記複数のジョブは、2以上の取得ジョブと2以

上の配置ジョブであり、

前記2以上の取得ジョブの各々は、構造化データを取得する取得処理を行うジョブであり、

前記2以上の配置ジョブの各々は、取得された構造化データを実行ノードに配置する配置処理を行うジョブである、

請求項4記載のデータ処理システム。

[請求項6]

前記配置情報が前記配置条件を表す情報を含んでいる場合、前記複数のジョブは、2以上の取得ジョブと2以上の配置ジョブであり、前記2以上の取得ジョブの各々は、構造化データを取得する取得処理を行うジョブであり、前記2以上の配置ジョブの各々は、取得された構造化データを実行ノードに配置する配置処理を行うジョブである、請求項3記載のデータ処理システム。

[請求項7]

前記配置情報は、前記2以上の実行ノードの各々について配置する構造化データに関する配置条件を表す情報を含み、

前記DBMSが、前記配置情報を基に、前記2以上の構造化データの各々を、その構造化データが適合する配置条件に対応付けられた実行ノードに配置する、

請求項2記載のデータ処理システム。

[請求項8]

前記複数の実行ノードの各々は物理計算機資源に基づいており、

前記DBMSが、前記2以上の実行ノードの各々について配置割合を決定し、前記2以上の実行ノードの各々に対応づいた配置割合は、前記2以上の構造化データの総量に対する、その実行ノードに配置する構造化データの量の割合であって、その実行ノードの物理計算機資源の負荷に基づき決定された割合であり、

前記DBMSが、前記2以上の実行ノードの各々への構造化データの配置を、その実行ノードの配置割合を基に制御する、

請求項2記載のデータ処理システム。

[請求項9]

前記制御情報は、準備情報であり、

前記準備情報は、複数の識別子を含み、

前記DBMSは、

前記準備情報が表す識別子毎に、取得された構造化データを関連付け、

前記複数の識別子のうちのいずれかの識別子を指定した送信要求を前記2以上の実行ノードのいずれかから受信した場合に、受信した送信要求で指定されている識別子に関連付けられている構造化データを、その送信要求の送信元の実行ノードに送信する、
請求項1記載のデータ処理システム。

[請求項10] 前記準備情報は、識別子毎に、送信される構造化データに関する条件であるデータ条件を表す情報を含み、

前記DBMSは、前記準備情報が表す識別子毎に、前記取得された2以上の構造化データのうち、その識別子に対応付けられたデータ条件に適合する構造化データを関連付ける、
請求項1記載のデータ処理システム。

[請求項11] 前記管理ノードが、前記2以上の実行ノードの各々に、その実行ノードが送信要求において指定すべき識別子を通知し、

前記2以上の実行ノードの各々は、前記管理ノードから通知された識別子を指定した送信要求を前記DBMSへ送信する、
請求項10記載のデータ処理システム。

[請求項12] 前記管理ノードが、アプリケーションから処理要求を受信し、その処理要求に基づく前記参照要求を前記DBMSに送信し、且つ、前記受信した処理要求に基づく処理命令を前記2以上の実行ノードの各々に送信し、

前記2以上の実行ノードの各々が、前記処理命令を受けて、前記DBMSが管理する構造化データ群のうちの少なくとも1つの構造化データと前記複数の実行ノードが管理する非構造化データ群のうちの少なくとも1つの非構造化データとを基に処理用データを作成し前記処

理用データを用いて処理を行い、その処理の結果を前記管理ノードに通知し、

前記管理ノードが、前記2以上の実行ノードからそれぞれ受けた2以上の並列分散処理結果に従う処理結果を前記分析アプリケーションへ応答する、

請求項1記載のデータ処理システム。

[請求項13]

1つ以上の構造化データの集合である構造化データ群を含んだデータベースを管理するデータベース管理システムであって、

参照要求を解析する解析部と、

前記参照要求で指定された2以上の構造化データの提供の制御に関する情報である制御情報を基に、前記2以上の構造化データを前記データベースから取得し、1つ以上の非構造化データの集合である非構造化データ群を管理し並列分散処理を実行する複数の実行ノードのうちの2以上の実行ノードへ前記2以上の構造化データを提供する実行部と

を有し、

前記複数の実行ノードの各々は、前記DBMSが管理する構造化データ群のうちの少なくとも1つの構造化データと前記複数の実行ノードが管理する非構造化データ群のうちの少なくとも1つの非構造化データとを基に処理用データを作成し前記処理用データを用いて処理を行うことができるノードである、

データベース管理システム。

[請求項14]

1つ以上の非構造化データの集合である非構造化データ群を管理し並列分散処理を実行する複数の実行ノードと通信可能であり1つ以上の構造化データの集合である構造化データ群を含んだデータベースを管理するデータベース管理システム（以下、DBMSと呼称）により、参照要求を受信し、前記複数の実行ノードの各々は、前記DBMSが管理する構造化データ群のうちの少なくとも1つの構造化データと

前記複数の実行ノードが管理する非構造化データ群のうちの少なくとも1つの非構造化データとを基に処理用データを作成し前記処理用データを用いて処理を行うことができるノードであり、

前記DBMSにより、前記参照要求で指定された2以上の構造化データの提供の制御に関する情報である制御情報を基に、前記2以上の構造化データを、前記データベースから取得し、前記2以上の構造化データを、前記複数の実行ノードのうちの2以上の実行ノードへ提供する、

データ処理方法。

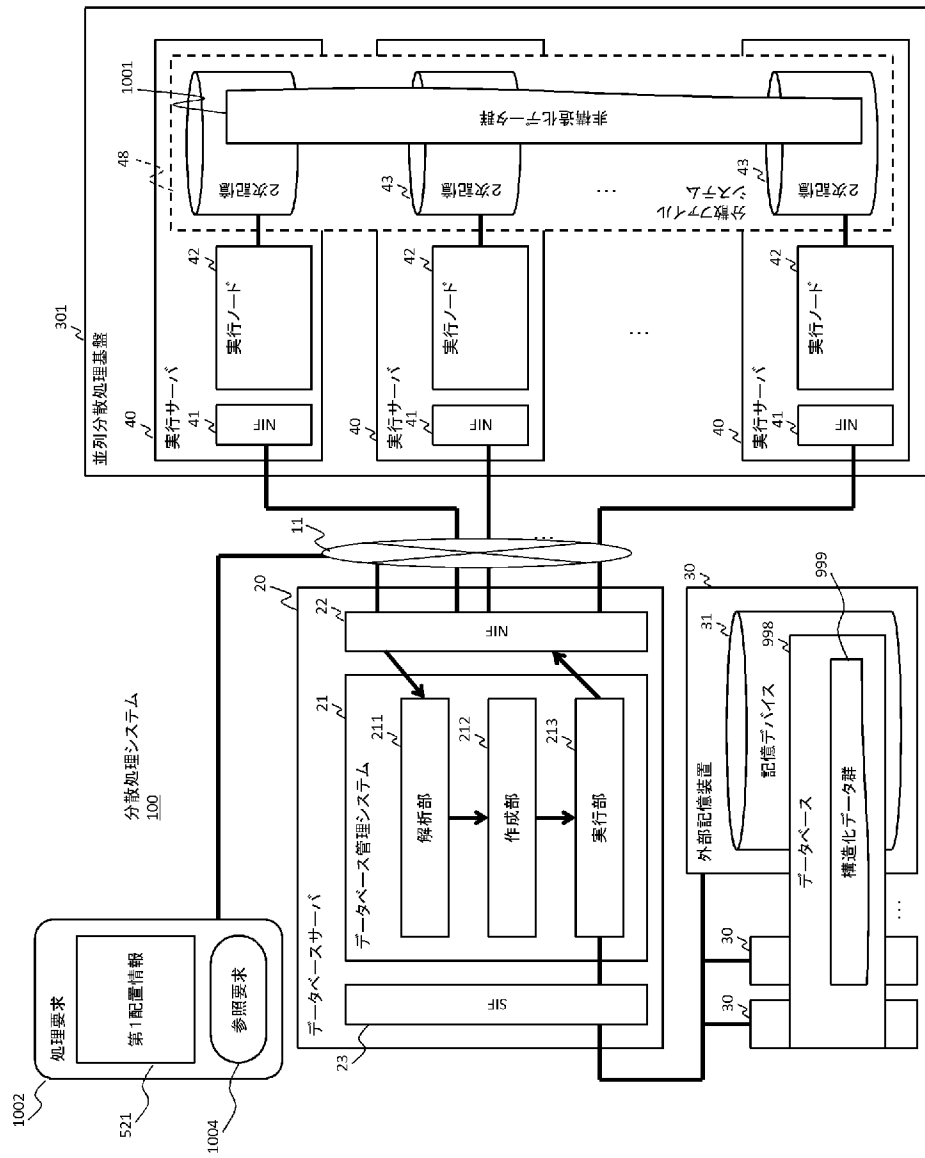
[請求項15]

前記管理ノードにより、アプリケーションから処理要求を受信し、その処理要求に基づく前記参照要求を前記DBMSに送信し、且つ、前記受信した処理要求に基づく処理命令を前記2以上の実行ノードの各々に送信し、

前記管理ノードにより、前記2以上の実行ノードの各々から、その実行ノードにより行われた処理の結果を受信し、前記2以上の実行ノードからそれぞれ受けた2以上の処理結果に従う結果を出力する、請求項14記載のデータ処理方法。

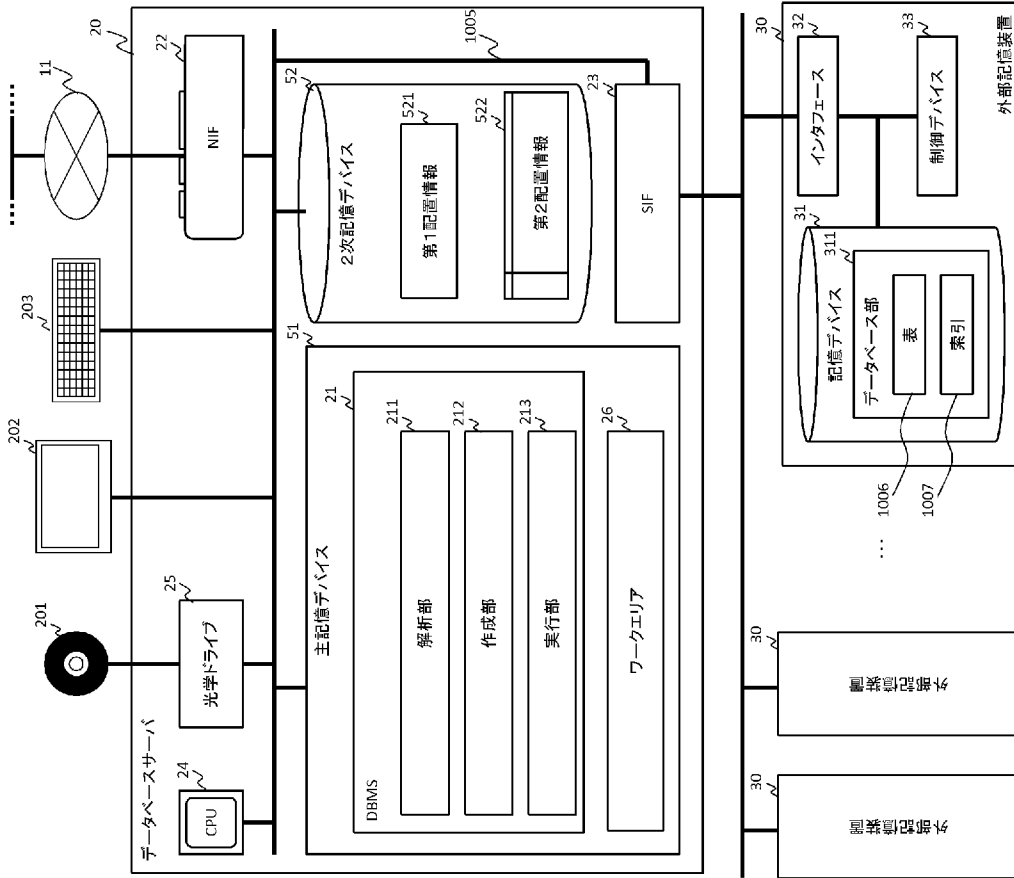
[図1]

FIG. 1



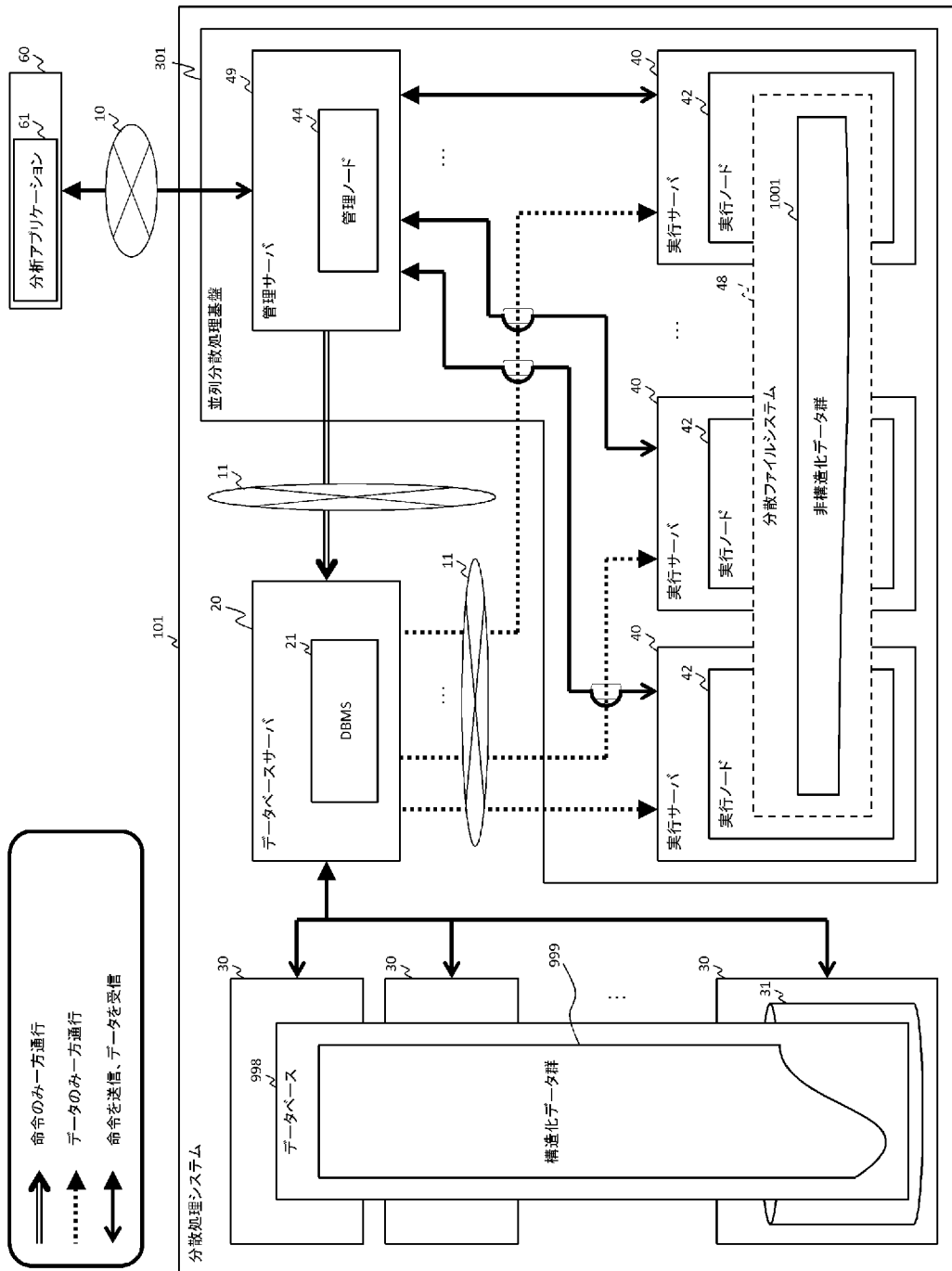
[図2]

FIG. 2



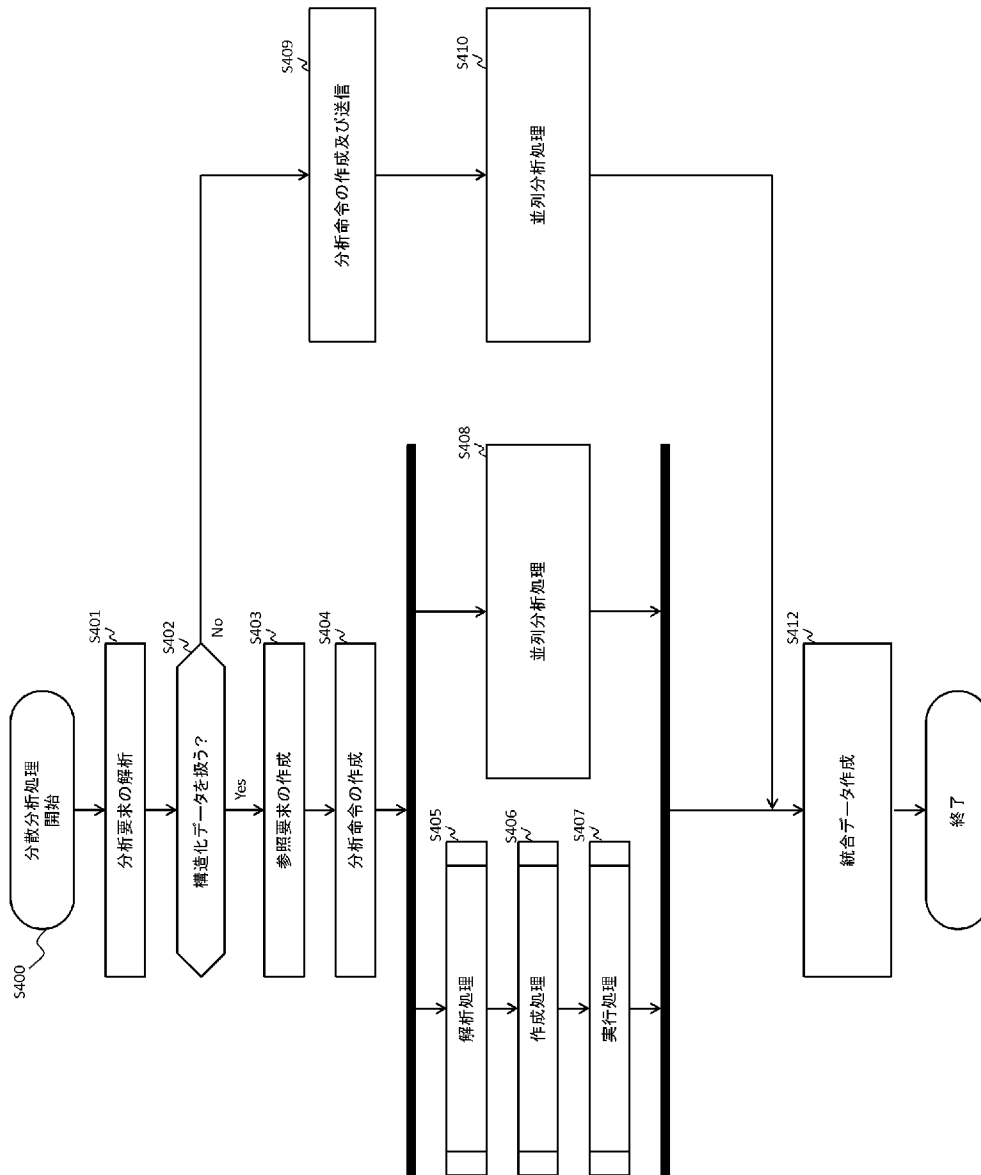
[図3]

FIG. 3



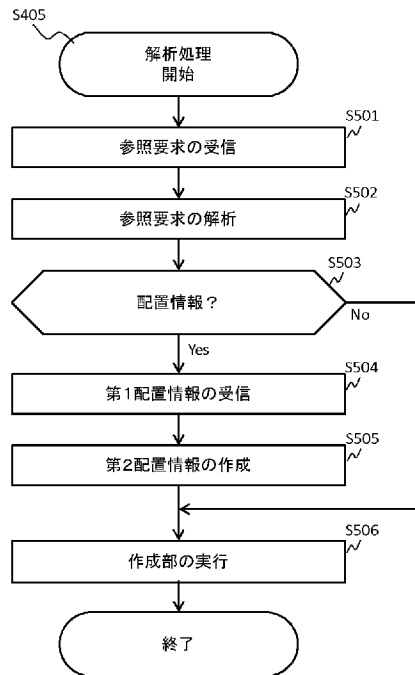
[図4]

FIG. 4



[図5]

FIG. 5



[図6]

FIG. 6

6001 拡張要求

```

// [1] ドライバ読み込み
Connection con = DriverManager.getConnection( url, user, password );
// [2] DBMS接続
PreparedStatement st = con.prepareStatement ( "SELECT * FROM item" );
// [3] 配置情報
st.setPosition();
// [4] データの取得, 配置
st.executeArrange();

```

[図7]

FIG. 7

521 第1配置情報

```

<NODE>
<NO> 1 </NO>
<IP> 100.1.1.1 </IP>
<PORT> 30001 </PORT>
<DATA>
  <TABLE> item </TABLE> <COLMUN> category </COLMUN> <VALUE> microwave ovens </VALUE>
</DATA>
</NODE>

<NODE>
<NO> 2 </NO>
<HOST> SLAVE2 </HOST>
<DATA>
  <TABLE> item </TABLE> <COLMUN> category </COLMUN> <VALUE> microwave ovens </VALUE>
</DATA>
</NODE>

<NODE>
<NO> 3 </NO>
<HOST> SLAVE3 </HOST>
<PORT> 29999 </PORT>
<DATA>
  <TABLE> item </TABLE> <COLMUN> category </COLMUN> <VALUE> refrigerator </VALUE>
</DATA>
</NODE>

<NODE>
<NO> 32 </NO>
<IP> 100.1.1.32 </IP>
<HOST> SLAVE32 </HOST>
</NODE>
    
```

[図8]

FIG. 8

811 NO	812 IP	813 HOST	814 PORT	815 TABLE	816 COLUMN	817 VALUE
1	100.1.1.1		30001	item	category	microwave ovens
2	100.1.1.2	SLAVE2	30000	item	category	microwave ovens
3	100.1.1.3	SLAVE3	29999	item	category	refrigerator
...
32	100.1.1.32	SLAVE32	30000			

522 第2配置情報

[図9]

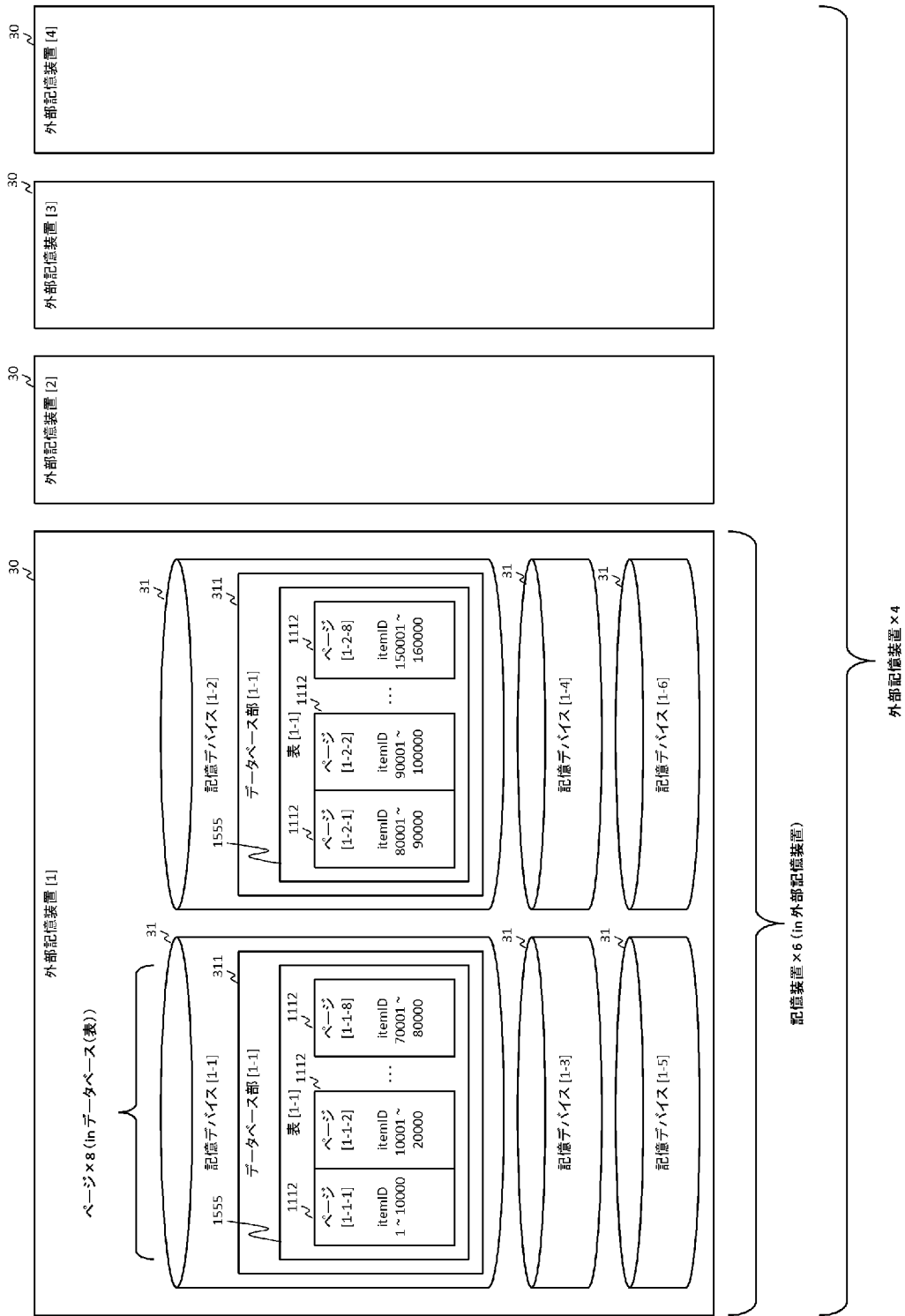
FIG. 9

⁹¹¹ itemID	⁹¹² category	⁹¹³ maker	⁹¹⁴ price	⁹¹⁵ stock
000001	microwave ovens	A社	30,000	10
000002	microwave ovens	A社	50,000	20
000003	microwave ovens	B社	45,000	7
000004	refrigerator	B社	120,000	5
000005	microwave ovens	C社	50,000	35
000006	microwave ovens	A社	80,000	10
000007	refrigerator	C社	110,000	15
000008	rice cooker	A社	25,000	25
000009	refrigerator	A社	65,000	6
000010	microwave ovens	D社	12,000	15
...
000050	rice cooker	C社	10,000	20
000051	refrigerator	B社	110,000	10
...
000100	rice cooker	D社	6,000	22
...

901
item表

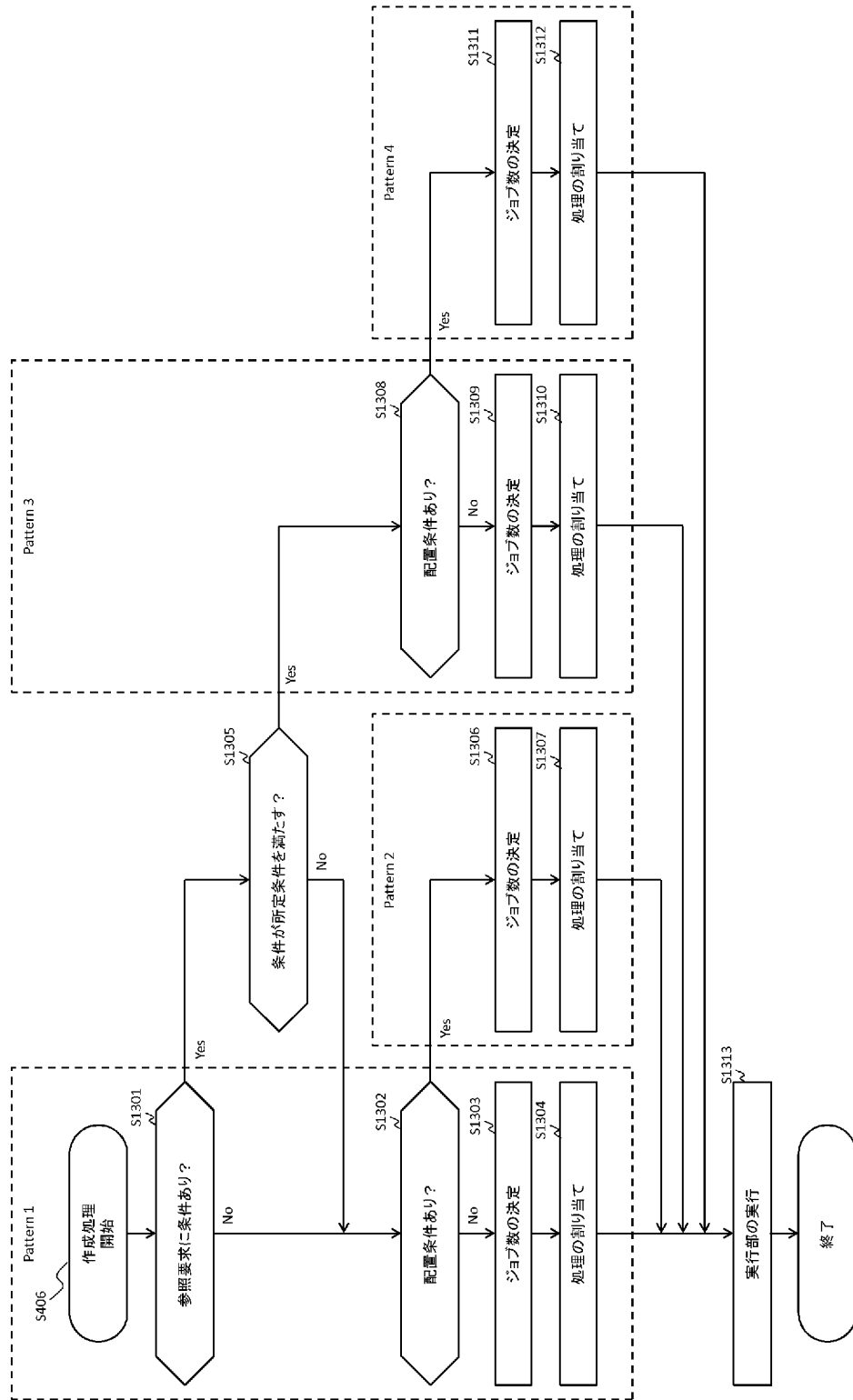
[図10]

FIG. 10



[図11]

FIG. 11



[図12]

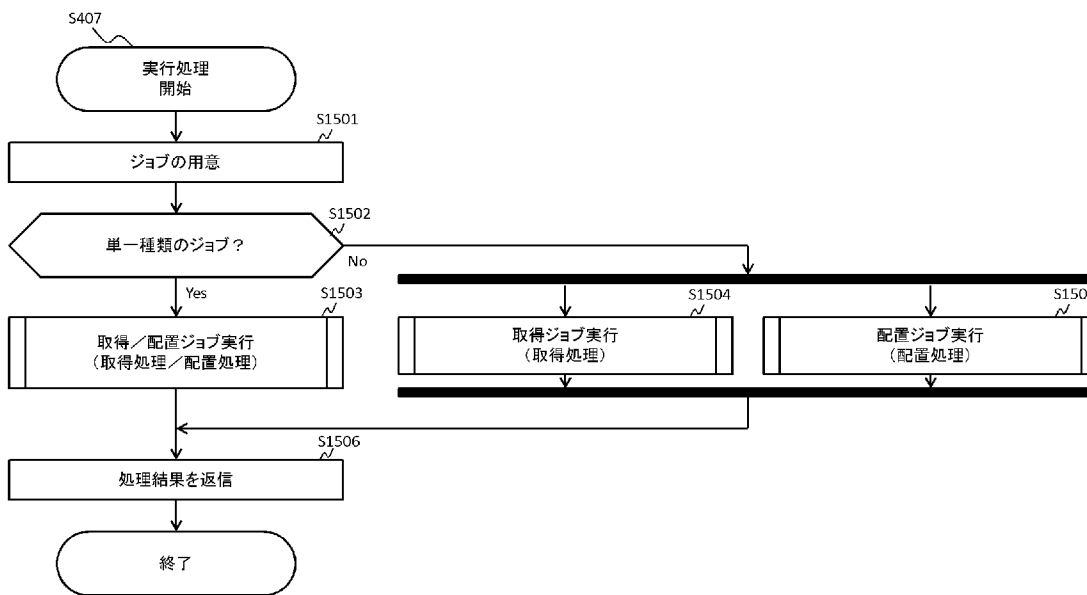
FIG. 12

NO	IP	HOST	PORT	TABLE	COLUMN	VALUE	JOBID
1	100.1.1.1		30001	item	category	microwave ovens	00000001
2	100.1.1.2	SLAVE2	30000	item	category	microwave ovens	00000002
3	100.1.1.3	SLAVE3	29999	item	category	refrigerator	00000003
...
32	100.1.1.32	SLAVE32	30000				00000032

1400
配置ジョブ対応表

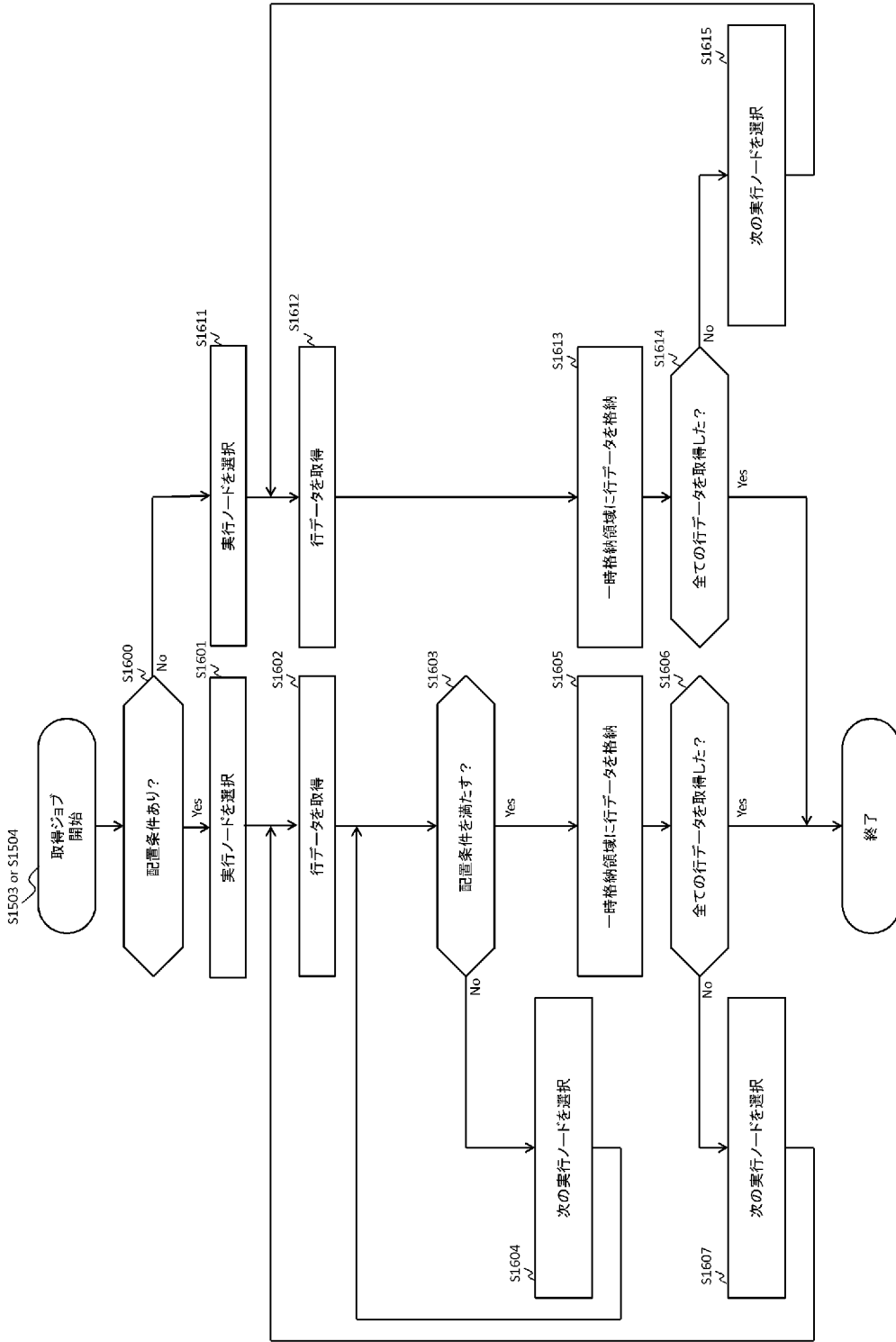
[図13]

FIG. 13



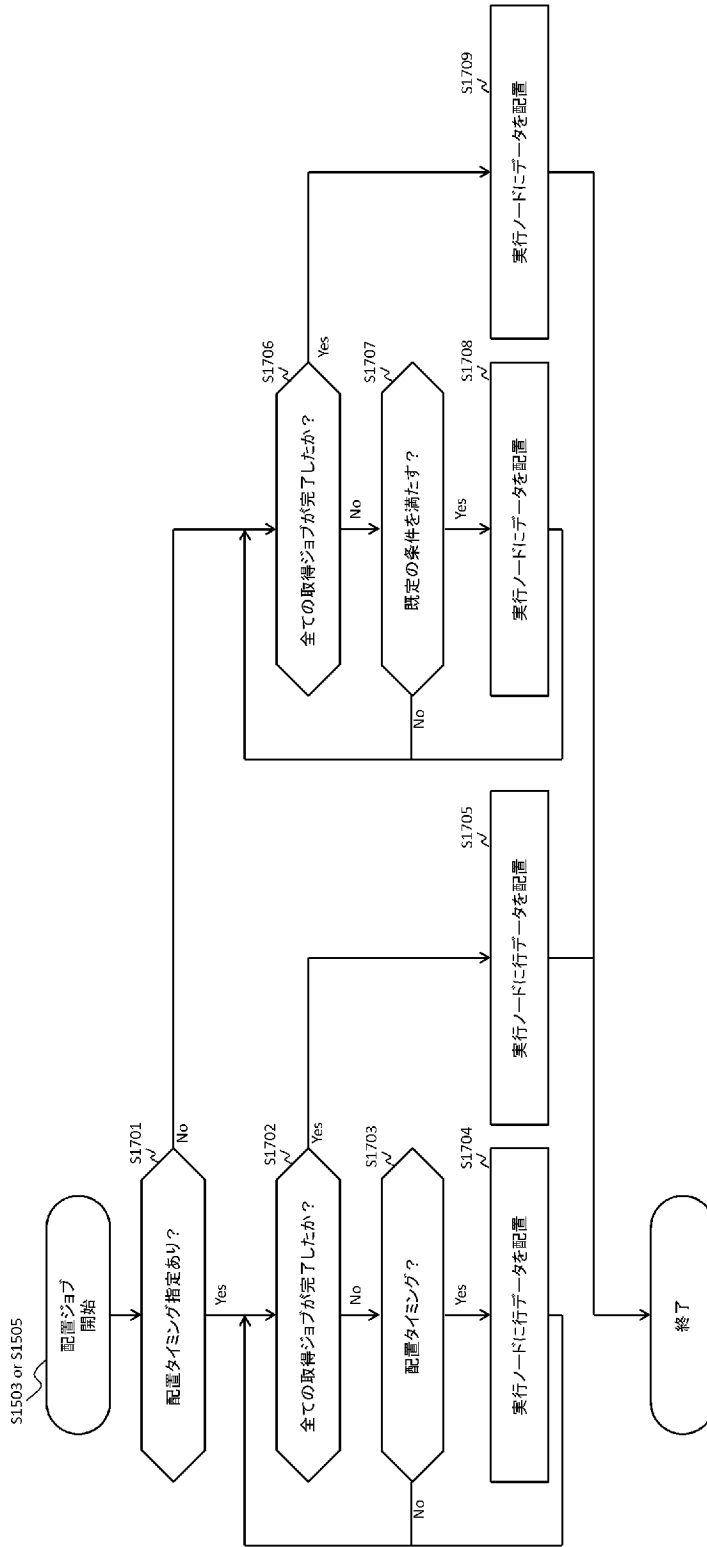
[図14]

FIG. 14



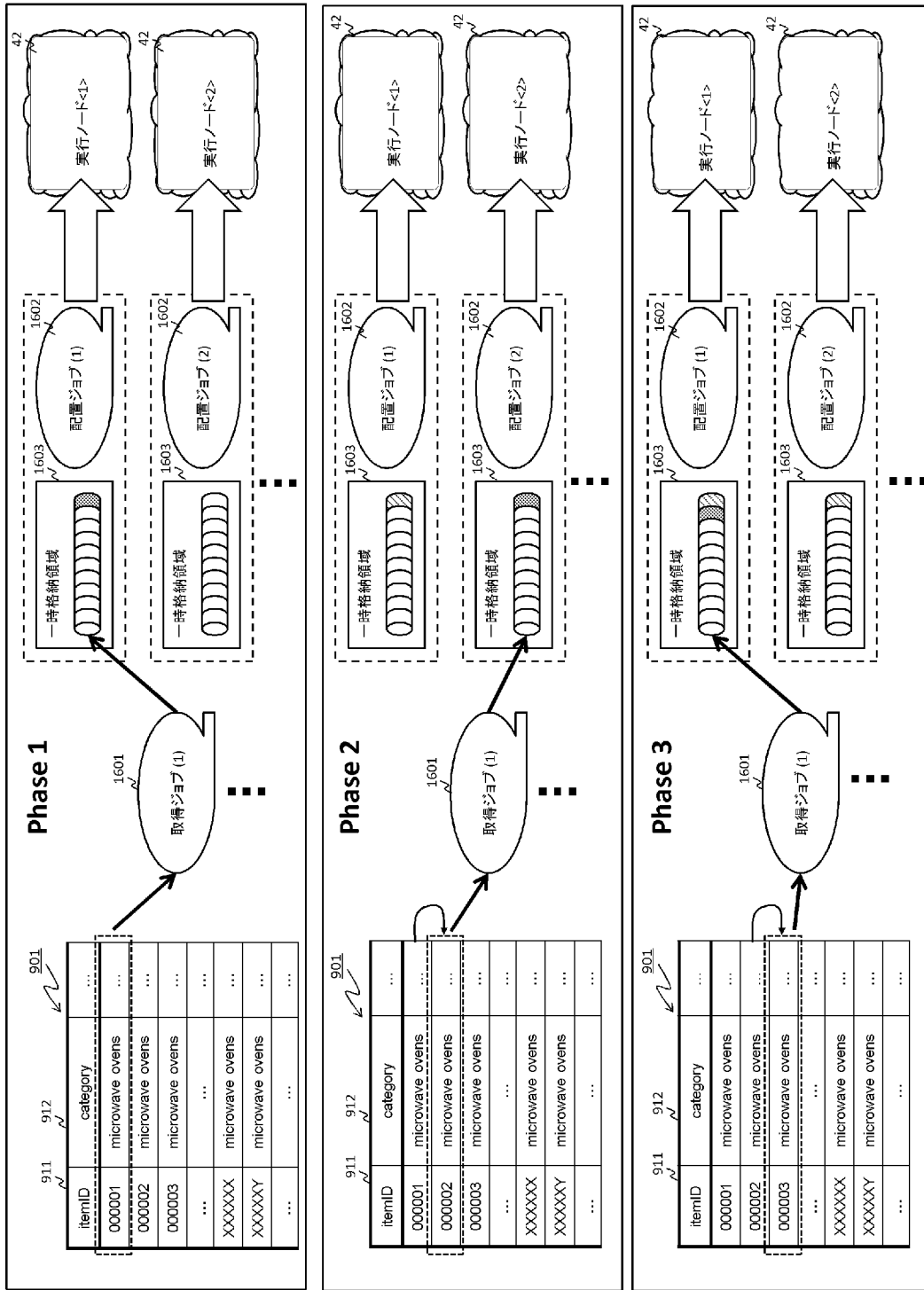
[図15]

FIG. 15



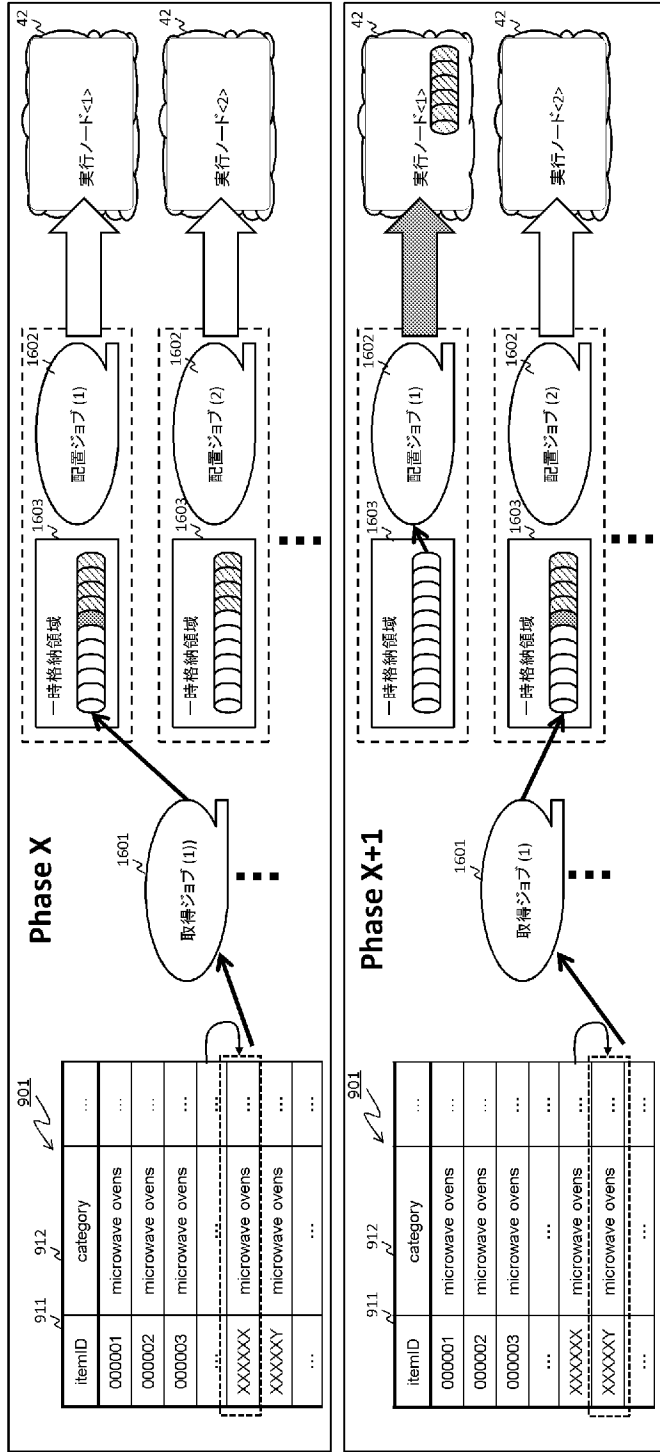
[図16]

FIG. 16



[図17]

FIG. 17



[18]

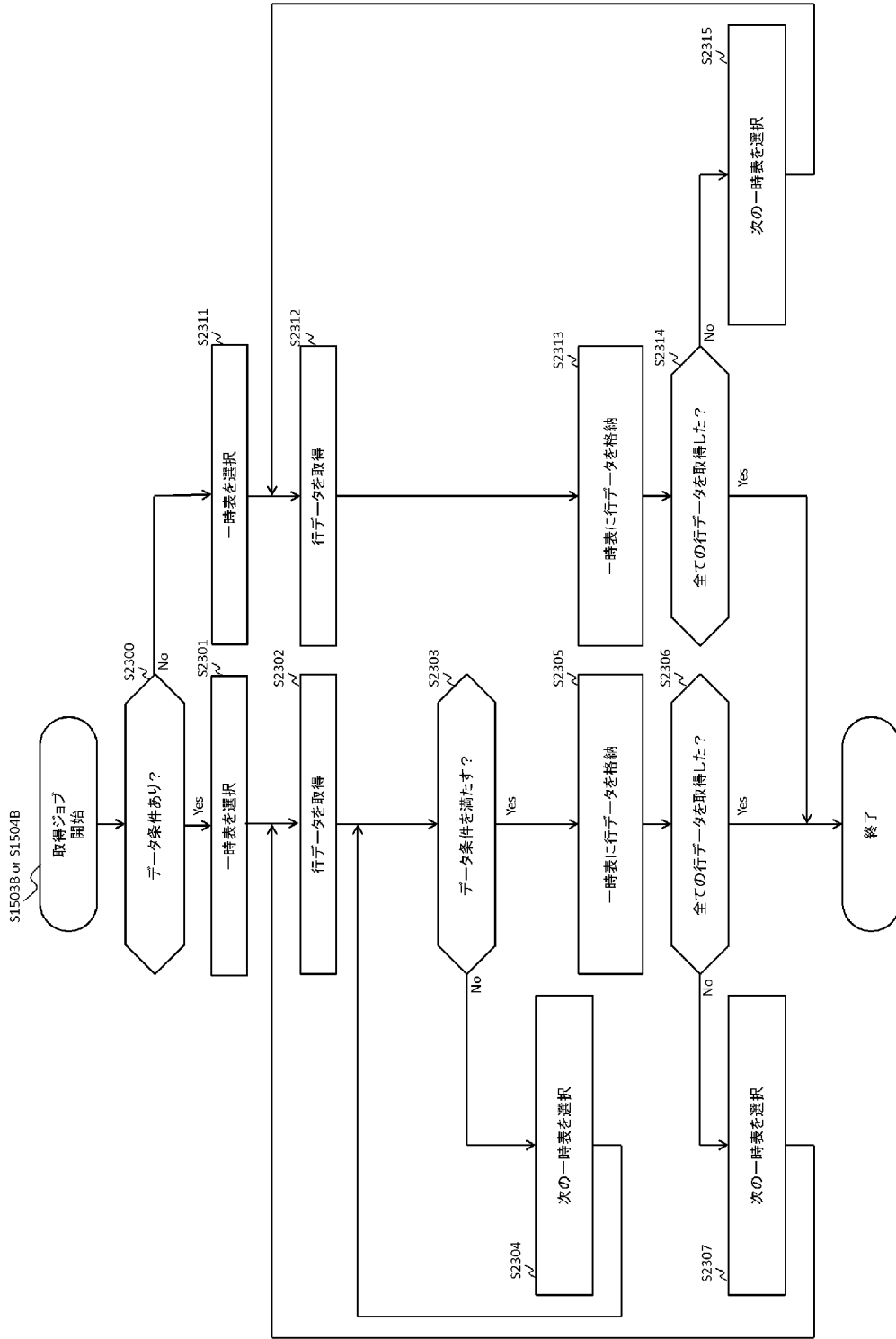
FIG. 18

1811 TABLEID	1812 TABLE	1813 COLUMN	1814 VALUE
T1	item	category	microwave ovens
T2	item	category	microwave ovens
T3	item	category	refrigerator
...
T32			

522B
第2テーブル情報

[図19]

FIG. 19



[図20]

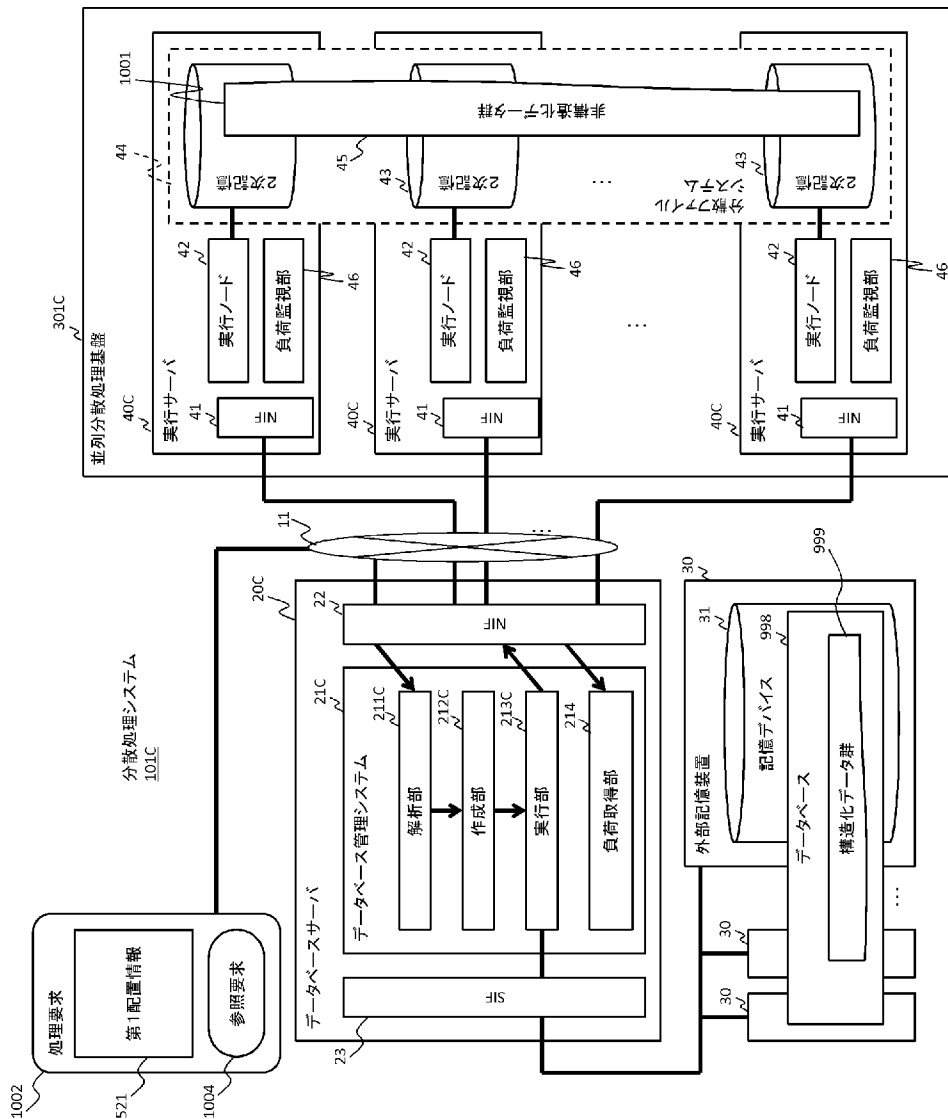
FIG. 20

2011 ~	2012 ~	2013 ~	2014 ~	2015 ~
itemID	category	maker	price	stock
000001	microwave ovens	A社	30,000	10
000003	microwave ovens	B社	45,000	7
000006	microwave ovens	A社	80,000	10
...

20000
一時表

[図21]

FIG. 21



[図22]

FIG. 22

NO	IP	CPU	...	NETWORK
1	100.1.1.1	10	...	12.5
2	100.1.1.2	30	...	15.0
3	100.1.1.3	50	...	55.5
...
32	100.1.1.32	5	...	10.0

2201
負荷情報

[図23]

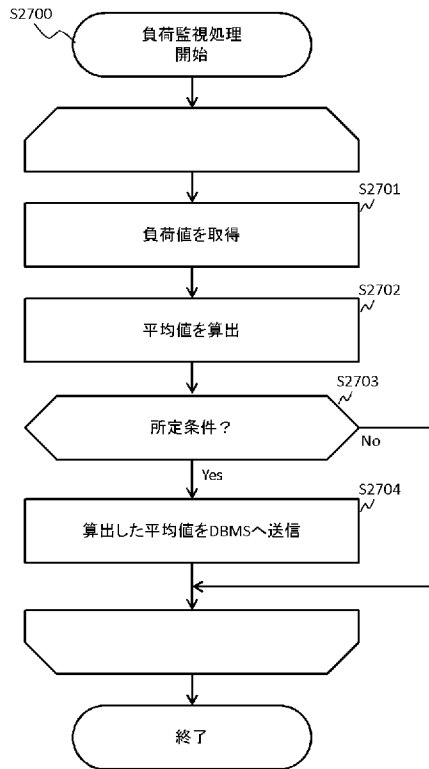
FIG. 23

NO	IP	HOST	PORT	RATIO
1	100.1.1.1		30001	4
2	100.1.1.2	SLAVE2	30000	4
3	100.1.1.3	SLAVE3	29999	2
...
32	100.1.1.32	SLAVE32	30000	3

522C
第2配置情報

[図24]

FIG. 24



[図25]

FIG. 25

NO	IP	HOST	PORT	RATIO	JOBID
1	100.1.1.1		30001	4	00000001
2	100.1.1.2	SLAVE2	30000	4	00000002
3	100.1.1.3	SLAVE3	29999	2	00000003
...
32	100.1.1.32	SLAVE32	30000	3	00000032

1400C
配置ジョブ対応表

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/068579

A. CLASSIFICATION OF SUBJECT MATTER
G06F12/00(2006.01)i, G06F9/50(2006.01)i, G06F17/30(2006.01)i

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)
G06F12/00, G06F9/50, G06F17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2014
Kokai Jitsuyo Shinan Koho	1971-2014	Toroku Jitsuyo Shinan Koho	1994-2014

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JSTPlus(JDreamIII), IEEE Xplore, THE ACM DIGITAL LIBRARY

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	JP 2010-511925 A (Exegy Inc.), 15 April 2010 (15.04.2010), entire text; all drawings & US 2008/0114724 A1 & EP 2092419 A & WO 2008/063974 A2	1, 2, 7-15 3-6
A	JP 2011-81794 A (International Business Machines Corp.), 21 April 2011 (21.04.2011), entire text; all drawings & US 2011/0082873 A1 & CN 102033900 A & KR 10-2011-0037889 A	1-15
A	US 2013/0268532 A1 (Sundeeep DOSHI), 10 October 2013 (10.10.2013), entire text; all drawings & WO 2013/154947 A1	1-15

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 22 October, 2014 (22.10.14)	Date of mailing of the international search report 04 November, 2014 (04.11.14)
--	--

Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2014/068579

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 10-207902 A (Toshiba Corp.), 07 August 1998 (07.08.1998), entire text; all drawings (Family: none)	1-15

A. 発明の属する分野の分類（国際特許分類（IPC））
 Int.Cl. G06F12/00(2006.01)i, G06F9/50(2006.01)i, G06F17/30(2006.01)i

B. 調査を行った分野
 調査を行った最小限資料（国際特許分類（IPC））
 Int.Cl. G06F12/00, G06F9/50, G06F17/30

最小限資料以外の資料で調査を行った分野に含まれるもの
 日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2014年
 日本国実用新案登録公報 1996-2014年
 日本国登録実用新案公報 1994-2014年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）
 JSTPlus(JDreamIII), IEEE Xplore, THE ACM DIGITAL LIBRARY

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	JP 2010-511925 A (エクセジー・インコーポレイテッド) 2010.04.15, 全文・全図	1, 2, 7-15
A	& US 2008/0114724 A1 & EP 2092419 A & WO 2008/063974 A2	3-6
A	JP 2011-81794 A (インターナショナル・ビジネス・マシーンズ・コー ポレーション) 2011.04.21, 全文・全図	1-15
	& US 2011/0082873 A1 & CN 102033900 A & KR 10-2011-0037889 A	

C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー	の日の後に公表された文献
「A」特に関連のある文献ではなく、一般的技術水準を示すもの	「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「O」口頭による開示、使用、展示等に言及する文献	「&」同一パテントファミリー文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願	

国際調査を完了した日 22.10.2014	国際調査報告の発送日 04.11.2014
--------------------------	--------------------------

国際調査機関の名称及びあて先 日本国特許庁（ISA/J P） 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官（権限のある職員） 池田 聡史 電話番号 03-3581-1101 内線 3565	5U	9475
--	---	----	------

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	US 2013/0268532 A1 (Sundeep DOSHI) 2013. 10. 10, 全文・全図 & WO 2013/154947 A1	1 - 1 5
A	JP 10-207902 A (株式会社東芝) 1998. 08. 07, 全文・全図 (ファミリーなし)	1 - 1 5