

(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
2013 年 11 月 28 日(28.11.2013)



(10) 国際公開番号

WO 2013/175609 A1

- (51) 国際特許分類:
G06F 12/00 (2006.01) G06F 17/30 (2006.01)
- (21) 国際出願番号: PCT/JP2012/063332
- (22) 国際出願日: 2012 年 5 月 24 日(24.05.2012)
- (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(米国を除く全ての指定国について): 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). 国立大学法人東京大学 (THE UNIVERSITY OF TOKYO) [JP/JP]; 〒1138654 東京都文京区本郷七丁目 3 番 1 号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人(米国についてのみ): 吉田 美智子 (YOSHIDA Michiko) [JP/JP]; 〒1858601 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所 中央研究所内 Tokyo (JP). 清水 晃

(SHIMIZU Akira) [JP/JP]; 〒1858601 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所 中央研究所内 Tokyo (JP). 徳田 晴介 (TOKUDA Seisuke) [JP/JP]; 〒1858601 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所 中央研究所内 Tokyo (JP). 茂木 和彦 (MOGI Kazuhiko) [JP/JP]; 〒2448555 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 I T プラットフォーム事業本部内 Kanagawa (JP). 藤原 真二 (FUJIWARA Shinji) [JP/JP]; 〒2448555 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 I T プラットフォーム事業本部内 Kanagawa (JP). 河村 信男 (KAWAMURA Nobuo) [JP/JP]; 〒2448555 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 I T プラットフォーム事業本部内 Kanagawa (JP). 喜連川 優 (KITSUREGAWA Masaru) [JP/JP]; 〒1138654 東京都文京区本郷七丁目 3 番 1 号 国立大学法人東京大学内 Tokyo (JP). 合田 和生 (GODA Kazuo) [JP/JP]; 〒1138654

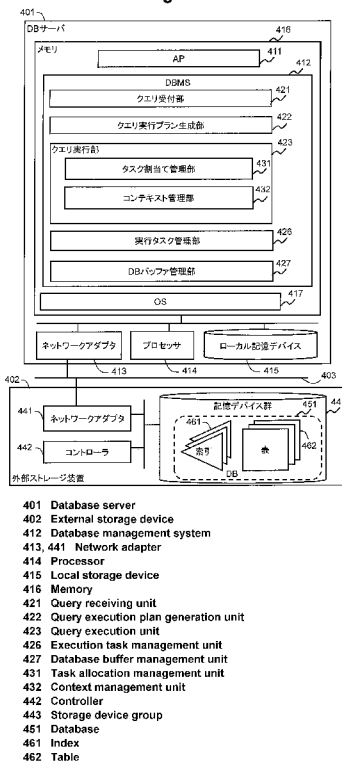
[続葉有]

(54) Title: DATABASE MANAGEMENT SYSTEM, COMPUTER, AND DATABASE MANAGEMENT METHOD

(54) 発明の名称: データベース管理システム、計算機、データベース管理方法

【図10】

Fig. 10



(57) Abstract: A database management system (DBMS) manages a database which is present in a second storage device which has a slower access speed than a first storage device. In an execution of a query, the DBMS dynamically generates a task, and if two or more executable tasks are present, executes at least two of the two or more tasks in parallel. The DBMS generates and manages task commencement information which is information which denotes a description which is to be carried out in an execution in a task, and executes the description which the task commencement information denotes by the task. The task commencement information includes a data address set which is present in the second storage device. On the basis of a management state of the task commencement information, the DBMS controls a movement of one or more data address sets between the first storage database and the second storage database. Additionally, the DBMS selects the task commencement information on the basis of whether the data address set is present in the first storage database.

(57) 要約: データベース管理システム (DBMS) は、第 1 の記憶デバイスよりアクセス速度の遅い第 2 の記憶デバイスに存在するデータベースを管理する。DBMS は、クエリの実行において、タスクを動的に生成し、実行可能な 2 以上のタスクが存在する場合には、2 以上のタスクのうちの少なくとも 2 つを並行して実行するようになっている。DBMS は、前記タスクにおいて実行を行う内容を示す情報であるタスク開始情報を生成して管理し、前記タスクにより、前記タスク開始情報が示す内容を実行する。前記タスク開始情報には、前記第 2 の記憶デバイスに存在するデータアドレス集合を含む。DBMS は、前記タスク開始情報の管理状態に基づいて、前記第 1 の記憶デバイスと前記第 2 の記憶デバイスとの間での 1 以上のデータアドレス集合の移動を制御する。また、DBMS は、前記データアドレス集合が、前記第 1 の記憶デバイスに存在するか否かに基づいてタスク開始情報を選択する。



東京都文京区本郷七丁目 3 番 1 号 国立大学
法人東京大学内 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); 〒
1010052 東京都千代田区神田小川町三丁目 3 番
地 神田小川町トーセイビル I I 7 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, 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, IS,
JP, KE, KG, KM, 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, PE, PG, PH,

PL, PT, QA, RO, RS, RU, RW, 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, ML,
MR, NE, SN, TD, TG).

添付公開書類:

- 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称：

データベース管理システム、計算機、データベース管理方法

技術分野

[0001] 本発明は、データベース管理技術に関する。

背景技術

[0002] 企業活動において、大量に生じる業務データの活用は不可欠になっている。そのため、大量の業務データを蓄積したデータベース（以下、「DB」）を解析処理するシステムが既に考案されている。

[0003] この解析処理において、データベース管理システム（以下、「DBMS」）は、クエリを受け付け、DBを格納する記憶デバイスにデータ読出し要求を発行する。

[0004] 1つのクエリの実行におけるデータ読出しの待ち時間の短縮化を図る技術として、特許文献1に開示の技術が知られている。特許文献1によれば、DBMSは、クエリを実行するために必要な複数のデータベースオペレーション（以下、「DBオペレーション」）を組合せたプラン（以下、「クエリ実行プラン」）を生成し、クエリ実行に要するデータ読出しの都度に前記DBオペレーションを実行するタスクを動的に生成し、前記タスクを並行実行することでデータ読出し要求を多重化する。

先行技術文献

特許文献

[0005] 特許文献1：特開2007-34414号公報

発明の概要

発明が解決しようとする課題

[0006] 一方、DBMSにおいて、クエリ実行に要するデータ読出しの都度にタスクを動的に生成し並行実行するに際して、タスクにおいて実行する内容を示す情報であるタスク開始情報を生成するとともに、その時点で生成可能な数

のタスクを生成し、前記生成したタスク開始情報に基づいて前記生成したタスクを実行する方法が考えられている。ここで、タスク開始情報は、例えば、コンテキストとデータアドレス集合とが関連づけられたものである。コンテキストは、例えば、新たに生成するタスクにおいて実行を開始するDBオペレーションが、クエリ実行プランが表す1以上のDBオペレーションのうちのいずれであるかに関する第1の情報と、その第1の情報が示すDBオペレーションに要するデータのアクセス先に関する第2の情報と、結果を生成するために必要な取得済みのデータに関する第3の情報とを含む情報である。ここで、クエリ実行プランとは、クエリを実行するために必要な1以上のDBオペレーションを表す情報を含む情報である。データアドレス集合とは、前記データのアクセス先に関する第2の情報に基づいて、DBMSが、生成したタスクを実行する際にアクセスするページを特定するための情報である。例えば、データアドレス集合は、DBのページに記憶されているデータであり、キーに対応する索引のページを特定する1以上のエントリの集合、又はキーに対応する表のレコードが格納されたページを特定する1以上のエントリの集合である。このDBMSは、タスクの実行開始時において、タスク開始情報を参照して読み出すべきページを特定し、当該ページを読み出し、当該読み出したページに記憶されているデータを用いてDBオペレーションを実行する。

[0007] DBMSは、タスクの実行開始時においてタスク開始情報を参照する際に、当該タスク開始情報のデータアドレス集合を参照する。DBMSは、ストレージ装置に格納されているページを参照する都度に、当該ページをDBMSが動作する計算機のメモリに読み出すが、一般にDBMSが動作する計算機のメモリの記憶容量は、DBが格納されているストレージ装置の記憶容量より小さい。そのため、メモリに同時に格納できるページは、ストレージ装置に格納されているページの一部に限られる。よって、DBMSがクエリを実行する際には、タスク開始情報が多数生成されるが、その全てのデータアドレス集合を記憶するページをメモリに格納することは困難である。

[0008] このような問題に対して、メモリの空き容量が不足し、データアドレス集合を記憶するページをメモリに格納できない場合に、メモリに格納されているページを一時的に計算機又はストレージ装置のディスクに追い出すことでメモリの空き容量を増加させる方法がある。この方法では、DBMSが、前記ディスクに追い出されたページを参照する際に、当該ページを再度メモリに読み込む必要がある。このように、一時的にページをディスクに追い出す方法を用いる場合には、DBMSがタスク開始情報を参照する際に、当該タスク開始情報のデータアドレス集合を記憶するページをメモリに読み込むためのディスクアクセスが発生し、タスクの処理開始が遅れ、処理時間が伸びるという問題がある。特に、I/Oを多重化してDBオペレーションを実行する場合には、I/Oのレイテンシが大きくなるために、この問題による影響が大きくなる。

[0009] そこで、本発明の目的は、DBMSがタスク開始情報を参照する際に、ディスクに格納されているページの参照回数を低減することで、タスクの実行を効率的に行い、クエリ実行時間を短縮することである。

課題を解決するための手段

[0010] DBMSは、第1の記憶デバイスよりアクセス速度の遅い第2の記憶デバイスに存在するデータベースを管理する。

[0011] DBMSが、クエリ受付部と、クエリ実行プラン生成部と、クエリ実行部と、実行タスク管理部と、DBバッファ管理部とを有する。例えば、DBMSは、コンピュータプログラムであって、計算機で実行されることにより、クエリ受付部と、クエリ実行プラン生成部と、クエリ実行部と、実行タスク管理部と、DBバッファ管理部とを計算機に構築する。

[0012] クエリ受付部は、クエリを受け付ける。クエリ実行プラン生成部は、クエリを実行するために必要な1以上のDBオペレーションを表す情報を含むクエリ実行プランを生成する。実行タスク管理部は、クエリを実行するためのタスクを管理する。

[0013] クエリ実行部は、クエリの実行において、DBオペレーションを実行する

ためのタスクを動的に生成し、動的に生成されたタスクを実行する。具体的には、例えば、クエリ実行部は、クエリの実行において、（a）DBオペレーションを実行するためのタスクを生成すること、（b）生成されたタスクを実行することで、そのタスクに対応したDBオペレーションに必要なデータを読み出すためにDBへデータ読出し要求を発行すること、（c）上記（b）で実行されたタスクに対応したN番目のDBオペレーションの実行結果に基づき（N+1）番目のDBオペレーションを実行する場合には、その実行結果に基づくタスクを新たに生成すること（Nは1以上の整数）、及び、（d）その新たに生成したタスクについて上記（b）及び上記（c）を行うこと、を行い、上記（b）及び（d）において、2以上の実行可能なタスクが存在する場合には、それら2以上のタスクのうちの少なくとも二つのタスクを並行して実行するようになっている。このクエリ実行部の動作は、前述した特許文献1に開示の技術に従う動作で良い。

[0014] クエリ実行部は、クエリの実行において、タスクを新たに生成する場合に（例えば上記（a）又は（c）の場合に）、タスクにおいて実行を行う内容を示す情報であるタスク開始情報を生成し、管理する。クエリ実行部は、タスクを実行開始する場合に、データアドレス集合を記憶するページが前記第1の記憶デバイスに読込まれているタスク開始情報を、データアドレス集合を記憶するページが前記第2の記憶デバイスにのみ格納されているタスク開始情報より優先して選択し、当該タスク開始情報が示す内容をタスクで実行開始する。

[0015] DBバッファ管理部は、クエリの実行において、クエリ実行部により管理されるタスク開始情報の実行状態により、第1の記憶デバイスと第2の記憶デバイスとの間でのページの移動を制御する。

発明の効果

[0016] DBMSがタスク開始情報を参照する際に、ディスクに格納されているページの参照回数を低減することで、タスクの実行を効率的に行うことができ、故に、クエリ実行時間の短縮化が期待できる。

図面の簡単な説明

- [0017] [図1]図 1 は、実施例 1 に係る表 A を示す。
- [図2]図 2 は、実施例 1 に係る表 B を示す。
- [図3]図 3 は、実施例 1 に係る表 C を示す。
- [図4]図 4 は、実施例 1 に係るクエリを示す。
- [図5]図 5 は、実施例 1 に係るクエリ実行プランを示す。
- [図6]図 6 は、実施例 1 に係るクエリ実行時のページに対するアクセス例を示す。
- [図7]図 7 は、実施例 1 に係るタスク開始情報を説明するための図である。
- [図8]図 8 は、クエリ実行時のタスクの実行及びタスク開始情報の管理を説明するための模式図の一例である。
- [図9]図 9 は、実施例 1 に係るクエリ実行時のタスクの実行及びタスク開始情報の管理を説明するための模式図の一例である。
- [図10]図 10 は、実施例 1 に係る計算機システムの構成を示す。
- [図11]図 11 は、実施例 1 に係る計算機システムの一部の詳細構成図である。
- [図12]図 12 は、実施例 1 に係る DB バッファ管理表の構成を示す。
- [図13]図 13 は、実施例 1 に係るクエリ処理の流れを示す。
- [図14]図 14 は、実施例 1 に係るタスク実行処理の流れを示す。
- [図15]図 15 は、実施例 1 に係るタスク実行内容取得処理の流れを示す。
- [図16]図 16 は、実施例 1 に係るタスク開始情報選択処理の流れを示す。
- [図17]図 17 は、実施例 1 に係るページ読込コンテキスト選択処理の流れを示す。
- [図18]図 18 は、実施例 1 に係る DB バッファ読込判定処理の流れを示す。
- [図19]図 19 は、実施例 1 に係るページ読込処理の流れを示す。
- [図20]図 20 は、実施例 1 に係る DB バッファ追い出し処理の流れを示す。
- [図21]図 21 は、実施例 1 に係る追い出し対象 DB バッファ面選択処理の流れを示す。

[図22]図22は、実施例1に係るDBバッファ面解放判定処理の流れを示す。

[図23]図23は、実施例2に係る計算機システムの一部の詳細構成図である。

[図24]図24は、実施例2に係るタスク開始情報格納場所管理表の構成を示す。

[図25]図25は、実施例3に係る計算機システムの構成を示す。

発明を実施するための形態

[0018] 以下、図面を参照しながら、幾つかの実施例を説明する。なお、以下の説明により本発明が限定されるものではない。また、以下の説明では、データベースを「DB」、データベース管理システムを「DBMS」、DBMSを実行するサーバを「DBサーバ」と言う。DBMSに対するクエリの発行元は、DBMSの外部のコンピュータプログラム（例えばアプリケーションプログラム）で良い。外部のコンピュータプログラムは、DBサーバ内で実行されるプログラムでも良いし、DBサーバに接続された装置（例えばクライアント計算機）で実行されるプログラムでも良い。

実施例 1

[0019] まず、本実施例の概要を説明する。

[0020] DBサーバは、DBMSを実行する。DBMSは、クエリを受け付け、受け付けたクエリを実行して、その実行により生成された結果をクエリ発行元に返す。クエリの結果を生成するまでには、DBMSは、1つ以上のDBオペレーションを実行する。DBMSは、それらDBオペレーションのうちの少なくとも1つのDBオペレーションの実行において、DBが格納された記憶デバイスに対して読出し要求の発行を要する場合がある。

[0021] 例えば、DBMSは、図1乃至図3に示す表A、表B、及び表Cと、表A用の索引A、表B用の索引B、及び表C用の索引Cとを含んだDBを、記憶デバイス（例えば、DBサーバと通信できるように接続された外部ストレージ装置）に格納しているとする。表は、1つ以上のレコードの集合であり、

レコードは、1つ以上のカラムから構成される。索引は、表の中の1つ以上のカラムを対象に作成されるデータ構造であり、当該索引が対象とするカラムを含む選択条件による表へのアクセスを高速化する。例えば、索引は、対象とするカラムの値毎にその値を含む表のレコードを特定する情報（レコードID）を保持するデータ構造であり、B木構造等が用いられる。

[0022] また、例えば、DBMSが受け付けたクエリは、図4に示すクエリであるとする。当該クエリは、カラムA__countryの値が「country 5」であり且つカラムA__factory__sizeの値が50より大きい表Aのレコードと、カラムB__factory__idの値が表AのレコードにおけるカラムA__factory__idの値と一致し且つカラムB__product__priceの値が500より大きい表Bのレコードと、カラムC__product__idの値が表BのレコードにおけるカラムB__product__idの値と一致する表Cのレコードとを関連づける。そして、それらをカラムA__factory__name及びカラムB__product__nameの値でグループ化し、それぞれのグループについてカラムA__factory__nameの値と、B__Product__nameの値と、カラムC__countの値の合計値とを抽出するクエリである。

[0023] DBMSは、このようなクエリを実行するために、図5に示すようなクエリ実行プランを生成する。クエリ実行プランは、例えば、データ読出しを伴う1つ以上のDBオペレーションで構成され、DBオペレーションの実行順序には木構造の順序関係が存在する。DBMSは、クエリ実行プランに基づいて、索引Aを用いて、表Aより、カラムA__countryの値がクエリの指定値（図4のクエリでは、country 5）であるレコードを読出し、当該レコードのうちカラムA__factory__sizeの値がクエリの指定値（図4のクエリでは、50）より大きいレコードにおけるカラムA__factory__nameの値とカラムA__factory__idの値とを抽出する。更に、索引Bを用いて、表Bより、カラムB__factory__idの値が前記抽出したA__factory__idの値と一致するレコード

を読み出し、当該レコードのうちカラムB__product__priceの値がクエリの指定値（図4のクエリでは、500）より大きいレコードにおけるカラムB__product__nameの値とカラムB__product__idの値とを抽出する。更に、索引Cを用いて、表Cより、カラムC__product__idの値が前記抽出したカラムB__product__idの値と一致するレコードを読み出し、当該レコードについて、カラムC__countの値を抽出し、前記抽出したカラムA__factory__name及びカラムB__product__nameの値に基づいてグループ化し、それぞれのグループにおけるカラムC__countの値の合計値を算出する。DBMSは、クエリ実行の結果として、前記グループ毎に、カラムA__factory__nameの値、カラムB__product__nameの値、及び算出したカラムC__countの合計値を、クエリ実行結果として生成する。

[0024] 例えば、DBMSは、下記の処理、

（S1）索引Aを用いて、カラムA__countryの値がクエリの指定値である表Aのレコードに対応したレコードIDの集合を検索する、

（S2）S1で見つかったレコードID集合を用いて、表Aの該当するレコードを含むデータをフェッチし、カラムA__factory__sizeの値がクエリの指定値である50より大きいレコードの、カラムA__factory__nameの値とカラムA__factory__idの値とを抽出する、

（S3）索引Bを用いて、カラムB__factory__idの値がS2で抽出したカラムA__factory__idの値と一致する表BのレコードのレコードID集合を検索する、

（S4）S3で見つかったレコードID集合を用いて、表Bの該当するレコードを含むデータをフェッチし、カラムB__product__priceの値がクエリの指定値である500より大きいレコードの、カラムB__product__idとカラムB__product__nameの値とを抽出する、

（S5）索引Cを用いて、カラムC__product__idの値がS4で抽出したカラムB__product__idの値と一致する表Cのレコードのレ

コードID集合を検索する、

(S6) S5で見つかったレコードID集合を用いて、表Cの該当するレコードを含むデータをフェッチし、該当するレコードのカラムC__countの値を抽出する、

(S7) カラムC__countの値を、抽出したカラムA__factory__nameの値及びカラムB__product__nameの値でグループ化して、各グループにおけるカラムC__countの値の合計値を算出する、

(S8) クエリ実行結果として、抽出したカラムA__factory__nameの値、カラムB__product__nameの値、及びS7で算出したカラムC__countの値の合計値を生成し、クエリ発行元に返す、
を行う。

[0025] 図6は、実施例1に係るクエリ実行時のページに対するアクセス例を示す。

[0026] 図6は、図5に示すクエリ実行プランに従ってDBMSがクエリを実行する際の、DBの索引及び表のページに対するアクセス例を示す。ここで、ページとは、DBMSにおけるデータ管理の単位であり、DBMSによるDBからのデータ読み込み及びDBへのデータ書き込みは、ページ単位で実施される。ページは、例えば、索引における下位のページを特定する情報（ページID）を含む1以上のエントリ、又は対応する表のレコードを特定する情報（レコードID）を含む1以上のエントリを格納する。なお、レコードを特定する情報は、当該レコードが格納されているページを特定する情報を含む。

[0027] 図5に示すクエリ実行プランによるクエリの実行時には、まず、DBMSは、ページa11にアクセスする。DBMSは、ページa11のエントリに基づいて、その下位のページa21乃至a25にアクセスする。

[0028] DBMSは、ページa21のエントリに基づいて、表AのページA1にアクセスし、ページa22のエントリに基づいて、表AのページA2及びページA3にアクセスし、ページa23のエントリに基づいて、表AのページA4及びページA5にアクセスし、ページa24のエントリに基づいて、表A

のページA 6にアクセスし、ページa 2 5のエントリに基づいて、表AのページA 7にアクセスする。

[0029] また、DBMSは、ページA 1に格納されたレコードに基づいて、索引Bのページb 1 1にアクセスし、ページA 3に格納されたレコードに基づいて、索引Bのページb 1 1にアクセスし、ページA 4に格納されたレコードに基づいて、索引Bのページb 1 1にアクセスし、ページA 5に格納されたレコードに基づいて、索引Bのページb 1 1にアクセスする。なお、ページA 2のように、索引Bのページへのアクセスに利用されていないものは、クエリで指定された条件に合致するレコードが格納されていないことを示している。

[0030] また、DBMSは、ページA 1に格納されたレコードと索引Bのページb 1 1のエントリに基づいてページb 2 1にアクセスし、ページA 3に格納されたレコードとページb 1 1のエントリに基づいてページb 2 2にアクセスし、ページA 4に格納されたレコードとページb 1 1のエントリに基づいてページb 2 3にアクセスし、ページA 5に格納されたレコードとページb 1 1のエントリに基づいてページb 2 4にアクセスする。

[0031] また、DBMSは、ページA 4に格納されたレコードと索引Bのページb 2 3のエントリに基づいて、表BのページB 1乃至B 3にアクセスする。DBMSは、表BのページB 1に格納されたレコードに基づいて、索引Cのページc 1 1にアクセスする。また、DBMSは、ページB 1に格納されたレコードと索引Cのページc 1 1のエントリに基づいてページc 2 1にアクセスし、ページB 1に格納されたレコードとページc 2 1のエントリに基づいてページc 3 1にアクセスする。また、DBMSは、ページB 1に格納されたレコードと索引Cのページc 3 1のエントリに基づいて、表CのページC 1にアクセスする。

[0032] 図7は、実施例1に係るタスク開始情報を説明するための図である。

[0033] DBMSは、タスクの実行を開始する際に、タスク開始情報TSを参照して、実行するDBオペレーションでアクセスするページを特定する。

[0034] タスク開始情報 T S は、タスクにおいて実行を行う内容を示す情報である。例えば、タスク開始情報は、コンテキスト C T X とデータアドレス集合 A G とが関連づけられたものである。コンテキスト C T X は、例えば、新たに生成するタスクにおいて実行を開始する D B オペレーションが、クエリ実行プランが表す 1 以上の D B オペレーションのうちのいずれであるかに関する第 1 の情報と、その第 1 の情報が示す D B オペレーションに要するデータのアクセス先に関する第 2 の情報と、結果を生成するために必要な取得済みのデータに関する第 3 の情報とを含む情報である。データアドレス集合 A G とは、前記データのアクセス先に関する第 2 の情報に基づいて、D B M S が、生成したタスクを実行する際にアクセスするページを特定するための情報である。例えば、データアドレス集合 A G は、D B 4 5 1 のページに記憶されているデータであり、キーに対応する索引のページを特定する 1 以上のエントリの集合、又はキーに対応する表のレコードが格納されたページを特定する 1 以上のエントリの集合である。

[0035] 外部ストレージ装置 4 0 2 の記憶デバイス群 4 4 3 のページは、アクセスされる都度に、D B M S 4 1 2 によって記憶デバイス群 4 4 3 から D B バッファ 4 3 6 に読み出される。D B バッファ 4 3 6 の記憶容量の大きさはメモリ 4 1 6 の記憶容量の大きさより小さく、一般にメモリ 4 1 6 の記憶容量の大きさは記憶デバイス群 4 4 3 の記憶容量の大きさより小さい。そのため、D B バッファ 4 3 6 に同時に格納することのできるページは、記憶デバイス群 4 4 3 に格納できるページの一部である。同図においては、例えば、タスク開始情報 1 のデータアドレス集合 A G が記憶されているページと、タスク開始情報 3 のデータアドレス集合 A G が記憶されているページとは、記憶デバイス群 4 4 3 から D B バッファ 4 3 6 に読み出されて格納されており、タスク開始情報 2 のデータアドレス集合 A G が記憶されているページは、D B バッファ 4 3 6 に格納されておらず、記憶デバイス群 4 4 3 にのみ格納されている。

[0036] コンテキスト C T X は、データアドレス集合 A G が記憶されているページ

を特定する情報（例えば、ページ I D）、データアドレス集合 A G のエントリのうち次に実行開始されるタスクによってアクセスすべきエントリを示す情報、さらにタスク実行によりアクセスすべきエントリの残数（タスク生成可能数）を格納する。コンテキスト C T X は、クエリ実行部 4 2 3 のコンテキスト管理部 4 3 2 により管理される。

[0037] 図 8 は、D B M S による、クエリ実行時のタスクの実行及びタスク開始情報の管理を説明するための模式図の一例である。図 8 は、図 6 に示すアクセス例におけるタスクの実行及びタスク開始情報の管理を示す。なお、図 8 の表記は、次のルールに従う。

（＊）横軸は、時刻を表す。

（＊）図上部の横に長い五角形は、1 つのタスクによる一連の D B オペレーションを意味する。五角形の左端はタスクにより D B オペレーションを開始する時刻を表し、五角形の右端は当該タスクによる D B オペレーションを終了する時刻を表す。

（＊）図上部の五角形の内部の値は、タスクが D B オペレーションの実行に伴いアクセスする索引のページを示す情報（例えば、ページ I D）、又は、表のレコードを示す情報（例えば、レコード I D）を表す。

（＊）同時に実行可能なタスク数は、「3」とする。

（＊）図中のひし形は、タスク開始情報 T S のデータアドレス集合 A G を表す。ひし形の中の数字は、当該タスク開始情報 T S のコンテキスト C T X に含まれるタスク生成可能数を表す。なお、図上部の五角形内のひし形の位置は、当該データアドレス集合 A G を含むタスク開始情報 T S が生成された時点を示す。

（＊）図中部の D B バッファ 4 3 6 内におけるひし形に続く太い実線は、データアドレス集合 A G を記憶するページが D B バッファ 4 3 6 に格納されていることを示し、太い破線は、データアドレス集合 A G を記憶するページが D B バッファ 4 3 6 に格納されていないことを示す。

（＊）D B バッファ 4 3 6 に同時に格納可能なページ数の上限数は、「2」

とする。

(*) DBバッファ436から記憶デバイス群443への破線矢印は、DBバッファ436から記憶デバイス群443へのページの追い出しを示し、記憶デバイス群443からDBバッファ436への破線矢印は、記憶デバイス群443からDBバッファ436へのページの読み込みを示す。なお、追い出し対象のページが、ダーティページ（記憶デバイス群443に格納されているページのデータと異なるデータが格納されているページ）でない場合には、DBMSは、そのページを、記憶デバイス群443に移動させずに、DBバッファから削除するだけでも良い。

(*) 記憶デバイス群443のアクセス速度は、DBバッファ436を保持するメモリ416のアクセス速度よりも遅い。なお、アクセス速度が速い／遅いとは、単位時間当たりに入出力されるデータの量が多い／少ないであっても良いし、アクセス元（DBMS）からアクセス要求が発行されてからそのアクセス要求に対するレスポンスがアクセス元（DBMS）へ戻るまでのレスポンスタイムの短い／遅いで良い。従って、例えば、メモリ416が、不揮発性半導体メモリであって、記憶デバイス群443を構成する各記憶デバイスも、同じアクセス速度の不揮発性半導体メモリであったとしても、記憶デバイス群443が、外部のストレージ装置に存在する場合には、ネットワーク転送速度が影響するので、記憶デバイス群443のアクセス速度は、メモリ416のアクセス速度よりも遅い。

[0038] (t0)

DBMSは、タスク10Aを生成し、タスク10Aの実行により、記憶デバイス群443上の索引Aのページa11にアクセスする。なお、ページにアクセスする際は、アクセス対象ページがDBバッファ436に格納されている場合は、当該DBバッファ436に格納されているページを参照し、アクセス対象ページがDBバッファ436に格納されていない場合は、ページを記憶デバイス群443からDBバッファ436へ読み込み、DBバッファ436に格納されたページを参照する。

[0039] (t 1)

DBMSは、タスク10Aの実行により、クエリの指定値とページa11のエントリに基づいて、アクセスすべき下位ページがa21～a25であることを特定し、タスク開始情報を生成する。すなわち、コンテキストを生成し、前記アクセスすべきと特定したページa21～a25のアドレスが格納されているページa11をデータアドレス集合11Aとして前記生成したコンテキストと関連付ける。DBバッファに同時に格納可能なページ数の上限は2であり、この時点でアクセスされるページ数は1であるので、データアドレス集合11AはDBバッファ上にある。また、前記コンテキストについて、次にアクセスすべきエントリをa22に設定し、タスク生成可能数を4に設定する。これにより、データアドレス集合11Aとそのデータアドレス集合11Aに関連付けられたコンテキストとのセットとしてのタスク開始情報が完成する。

続いて、DBMSは、タスク10Aの実行により、ページa21にアクセスする。

また、実行可能な同時タスク数は「3」であり、実行中タスクの数は「1」である（タスク10Aのみが実行中である）ため、生成可能なタスクの数は、「3」から「1」を減算した「2」であるので、DBMSは、2つのタスク10B及び10Cを生成する。

また、DBMSは、タスク10Bの実行により、タスク開始情報を参照してタスクを1つ実行開始する。すなわち、データアドレス集合11Aと関連付けられたコンテキストを参照して次にアクセスすべきエントリがa22であることを特定し、前記次にアクセスすべきエントリを次のページであるa23に変更する。そして、タスク生成可能数を1減算して3に変更し、データアドレス集合11Aを参照して前記ページa22のアドレスを取得し、取得したアドレスに基づいてページa22にアクセスする。

また、DBMSは、タスク10Cの実行により、タスク開始情報を参照してタスクを1つ実行開始する。すなわち、データアドレス集合11Aと関連づ

けられたコンテキストを参照して次にアクセスすべきエントリが a 2 3であることを特定し、前記次にアクセスすべきエントリを次のページである a 2 4に変更する。そして、タスク生成可能数を 1 減算して 2 に変更し、データアドレス集合 1 1 A を参照して前記ページ a 2 3 のアドレスを取得し、前記取得したアドレスに基づいてページ a 2 3 にアクセスする。

[0040] (t 2)

DBMS は、タスク 1 0 A の実行により、クエリの指定値とページ a 2 1 のエントリに基づいて、アクセスすべきレコードが A 1 であることを特定し、ページ a 2 1 を参照して前記レコード A 1 のアドレスを取得し、前記取得したアドレスに基づいてページ A 1 にアクセスする。

また、DBMS は、タスク 1 0 B の実行により、クエリの指定値とページ a 2 2 のエントリに基づいて、アクセスすべきレコード A 2 と A 3 とを特定し、タスク開始情報を生成する。すなわち、コンテキストを生成し、前記アクセスすべきと特定したページ A 2 と A 3 のアドレスが格納されているページ a 2 2 をデータアドレス集合 1 1 B として前記生成したコンテキストと関連付け、当該コンテキストについて、次にアクセスすべきエントリを A 3 に設定し、タスク生成可能数を 1 に設定する。これにより、データアドレス集合 1 1 B とそのデータアドレス集合 1 1 B に関連付けられたコンテキストとのセットとしてのタスク開始情報が完成する。続いて、DBMS は、タスク 1 0 B の実行により、ページ A 2 にアクセスする。

また、DBMS は、タスク 1 0 C の実行により、クエリの指定値とページ a 2 3 のエントリに基づいて、アクセスすべきレコード A 4 と A 5 とを特定し、タスク開始情報を生成する。すなわち、コンテキストを生成し、前記アクセスすべきと特定したページ A 4 と A 5 のアドレスが格納されているページ a 2 3 をデータアドレス集合 1 1 C として前記生成したコンテキストと関連付け、当該コンテキストについて、次にアクセスすべきエントリを A 5 に設定し、タスク生成可能数を 1 に設定する。これにより、データアドレス集合 1 1 C とそのデータアドレス集合 1 1 C に関連付けられたコンテキストとの

セットとしてのタスク開始情報が完成する。続いて、DBMSは、タスク10Cの実行により、ページA4にアクセスする。

DBMSは、データアドレス集合11Bを記憶するページと、データアドレス集合11Cを記憶するページとをDBバッファ436に格納することとなるが、DBバッファ436に同時に格納可能なページの数、「2」である。このため、DBMSは、データアドレス集合11B及び11Cを記憶するページのうちの少なくとも1つをDBバッファ436に格納する前に、DBバッファ436上の最も古いデータアドレス集合（DBバッファ436に格納されているデータアドレス集合のうち最初にDBバッファ436に格納されたデータアドレス集合）11Aを記憶するページを、記憶デバイス群443に追い出す。なお、t1においてアクセスされたページa21もバッファに格納されるため上記挙動に影響を与え得る。図8では本発明の要点を簡潔に説明するため、ページa21は一時的にDBバッファに格納され、A1を特定した後はすぐに記憶デバイス群443に追い出されると仮定して説明を行うが、後述の通り、本発明はこの仮定に限定されるものではない。

[0041] (t3)

DBMSは、タスク10Aの実行により、レコードA1がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引Bの検索条件を特定し、索引Bのページb11にアクセスする。

また、DBMSは、タスク10Bの実行により、レコードA2がクエリで指定された条件を満たさないことを判定する。続いて、タスクを実行するため、最初に生成されたタスク開始情報（データアドレス集合11Aとそれに関連づけられたコンテキストから成るタスク開始情報）を参照する。すなわち、データアドレス集合11Aと関連づけられたコンテキストを参照して次にアクセスすべきエントリがa24であることを特定し、前記次にアクセスすべきエントリを次のページであるa25に変更し、タスク生成可能数を1減算して1に変更し、データアドレス集合11Aを参照して前記ページa24のアドレスを取得しようと試みる。しかし、t2においてデータアドレス集

合 1 1 A は、記憶デバイス群 4 4 3 に追い出されており、DBバッファ 4 3 6 上に存在しないため、データアドレス集合 1 1 A を記憶するページを記憶デバイス群 4 4 3 から DBバッファ 4 3 6 に読み込む処理を行う。このため、タスク 1 0 B は、データアドレス集合 1 1 A を記憶するページがメモリ 4 1 6 に読み込まれるのを待つ。ここで、DBMS は、データアドレス集合 1 1 A を記憶するページを DBバッファ 4 3 6 に読み込むと、DBバッファ 4 3 6 に格納されるページの数、格納可能な数「3」を超えてしまうので、DBバッファ 4 3 6 の最も古いデータアドレス集合 1 1 B を記憶するページを、記憶デバイス群 4 4 3 に追い出す。

また、DBMS は、タスク 1 0 C の実行により、レコード A 4 がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引 B の検索条件を特定し、索引 B のページ b 1 1 にアクセスする。

[0042] (t 4)

DBMS は、タスク 1 0 A の実行により、t 3 で特定した索引 B の検索条件とページ b 1 1 のエントリに基づいて、アクセスすべき下位ページが b 2 1 であることを特定し、ページ b 1 1 を参照してページ b 2 1 のアドレスを取得し、前記取得したアドレスに基づいてページ b 2 1 にアクセスする。

また、DBMS は、タスク 1 0 B の実行により、t 3 で記憶デバイス群 4 4 3 から DBバッファ 4 3 6 に読み込む処理を行ったデータアドレス集合 1 1 A を参照してページ a 2 4 のアドレスを取得し、前記取得したアドレスに基づいてページ a 2 4 にアクセスする。

また、DBMS は、タスク 1 0 C の実行により、t 3 で特定した索引 B の検索条件とページ b 1 1 のエントリに基づいて、アクセスすべき下位ページが b 2 3 であることを特定し、ページ b 1 1 を参照してページ b 2 3 のアドレスを取得し、前記取得したアドレスに基づいてページ b 2 3 にアクセスする。

[0043] (t 5)

DBMS は、タスク 1 0 A の実行により、t 3 で特定した索引 B の検索条件

とページ b 2 1 のエントリに基づいて、アクセスすべき下位ページが存在しないことを判定する。続いて、タスクを実行するため、最初に生成されたタスク開始情報（データアドレス集合 1 1 A とそれに関連づけられたコンテキストから成るタスク開始情報）を参照する。すなわち、データアドレス集合 1 1 A と関連づけられたコンテキストを参照して次にアクセスすべきエントリが a 2 5 であることを特定し、タスク生成可能数を 1 減算して 0 になることを確認し、データアドレス集合 1 1 A を参照して前記ページ a 2 5 のアドレスを取得し、当該タスク開始情報を消去し、前記取得したアドレスに基づいてページ a 2 5 にアクセスする。

また、DBMS は、タスク 1 0 B の実行により、クエリの指定値とページ a 2 4 のエントリに基づいて、アクセスすべきレコードが A 6 であることを特定し、ページ a 2 4 を参照して前記レコード A 6 のアドレスを取得し、前記取得したアドレスに基づいてページ A 6 にアクセスする。

また、DBMS は、タスク 1 0 C の実行により、t 3 で特定した索引 B の検索条件とページ b 2 3 のエントリに基づいて、アクセスすべきレコードが B 1 ～ B 3 であることを特定し、タスク開始情報を生成する。すなわち、コンテキストを生成し、前記アクセスすべきと特定したページ B 1 ～ B 3 のアドレスが格納されているページ b 2 3 をデータアドレス集合 1 1 D として前記生成したコンテキストと関連付け、当該コンテキストについて、次にアクセスすべきエントリを B 2 に設定し、タスク生成可能数を 2 に設定する。これにより、データアドレス集合 1 1 D とそのデータアドレス集合 1 1 D に関連付けられたコンテキストとのセットとしてのタスク開始情報が完成する。続いて、DBMS は、タスク 1 0 C の実行により、ページ B 1 にアクセスする。

ここで、DBMS は、データアドレス集合 1 1 D を記憶するページを DB バッファ 4 3 6 に格納することとなるが、DB バッファ 4 3 6 には、その時点には、データアドレス集合 1 1 A 及び 1 1 C を記憶するページが格納されており、DB バッファ 4 3 6 に格納可能なページの数、「2」である。この

ため、DBMSは、データアドレス集合11Dを記憶するページをDBバッファ436に格納する前に、DBバッファ436上の参照されていない最も古いデータアドレス集合11Cを記憶するページを、記憶デバイス群443に追い出す。

[0044] (t6)

DBMSは、タスク10Aの実行により、クエリの指定値とページa25のエントリに基づいて、アクセスすべきレコードがA7であることを特定し、ページa25を参照して前記レコードA7のアドレスを取得し、前記取得したアドレスに基づいてページA7にアクセスする。

また、DBMSは、タスク10Bの実行により、レコードA6がクエリで指定された条件を満たさないことを判定する。続いて、タスクを実行するため、タスク生成可能数が1以上のタスク開始情報の中で最初に生成されたタスク開始情報（データアドレス集合11Bとそれに関連づけられたコンテキストから成るタスク開始情報）を参照する。すなわち、データアドレス集合11Bと関連づけられたコンテキストを参照して次にアクセスすべきエントリがA3であることを特定し、タスク生成可能数を1減算して0になることを確認し、データアドレス集合11Bを参照して前記ページA3のアドレスを取得しようと試みる。しかし、t3においてデータアドレス集合11Bを記憶するページは、記憶デバイス群443に追い出されており、DBバッファ436上に存在しないため、データアドレス集合11Bを記憶するページを記憶デバイス群443からDBバッファ436に読み込む処理を行う。このため、タスク10Bは、データアドレス集合11Bを記憶するページがDBバッファ436に読み込まれるのを待つ。

また、DBMSは、タスク10Cの実行により、レコードB1がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引Cの検索条件を特定し、索引Cのページc11にアクセスする。

[0045] (t7)

DBMSは、タスク10Aの実行により、レコードA7がクエリで指定され

た条件を満たさないことを判定する。続いて、タスクを実行するため、タスク生成可能数が1以上のタスク開始情報の中で最初に生成されたタスク開始情報（データアドレス集合11Cとそれに関連づけられたコンテキストから成るタスク開始情報）を参照する。すなわち、データアドレス集合11Cと関連づけられたコンテキストを参照して次にアクセスすべきエントリがA5であることを特定し、タスク生成可能数を1減算して0になることを確認し、データアドレス集合11Cを参照して前記ページA5のアドレスを取得しようと試みる。しかし、t5においてデータアドレス集合11Cを記憶するページは、記憶デバイス群443に追い出されており、DBバッファ436上に存在しないため、データアドレス集合11Cを記憶するページを記憶デバイス群443からDBバッファ436に読込む処理を行う。このため、タスク10Aは、データアドレス集合11Cを記憶するページがDBバッファ436に読込まれるのを待つ。ここで、DBMSは、データアドレス集合11Cを記憶するページをDBバッファ436に読込むこととなるが、DBバッファ436には、その時点には、データアドレス集合11B及び11Dを記憶するページが格納されており、DBバッファ436に格納可能なページの数、「2」である。このため、DBMSは、DBバッファ436上の参照されていない最も古いデータアドレス集合11Dを記憶するページを、記憶デバイス群443に追い出す。

また、DBMSは、タスク10Bの実行により、t6で記憶デバイス群443からDBバッファ436に読込む処理を行ったデータアドレス集合11Bを参照してページA3のアドレスを取得し、タスク生成可能数が0になった当該タスク開始情報を削除し、前記取得したアドレスに基づいてページA3にアクセスする。

また、DBMSは、タスク10Cの実行により、t6で特定した索引Cの検索条件とページc11のエントリに基づいて、アクセスすべき下位ページがc21であることを特定し、ページc11を参照してページc21のアドレスを取得し、前記取得したアドレスに基づいてページc21にアクセスする

。

[0046] (t 8)

DBMSは、タスク10Aの実行により、t 7で記憶デバイス群443からDBバッファ436に読み込む処理を行ったデータアドレス集合11Cを参照してページA5のアドレスを取得し、タスク生成可能数が0になった当該タスク開始情報を削除し、前記取得したアドレスに基づいてページA5にアクセスする。

また、DBMSは、タスク10Bの実行により、レコードA3がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引Bの検索条件を特定し、索引Bのページb11にアクセスする。

また、DBMSは、タスク10Cの実行により、t 6で特定した索引Cの検索条件とページc21のエントリに基づいて、アクセスすべき下位ページがc31であることを特定し、ページc21を参照してページc31のアドレスを取得し、前記取得したアドレスに基づいてページc31にアクセスする

。

[0047] (t 9)

DBMSは、タスク10Aの実行により、レコードA5がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引Bの検索条件を特定し、索引Bのページb11にアクセスする。

また、DBMSは、タスク10Bの実行により、t 8で特定した索引Bの検索条件とページb11のエントリに基づいて、アクセスすべき下位ページがb22であることを特定し、ページb11を参照してページb22のアドレスを取得し、前記取得したアドレスに基づいてページb22にアクセスする

。

また、DBMSは、タスク10Cの実行により、t 6で特定した索引Cの検索条件とページc31のエントリに基づいて、アクセスすべきレコードがC1であることを特定し、ページc31を参照して前記レコードC1のアドレスを取得し、前記取得したアドレスに基づいてページC1にアクセスする。

[0048] (t 1 0)

DBMSは、タスク10Aの実行により、t 9で特定した索引Bの検索条件とページb 1 1のエントリに基づいて、アクセスすべき下位ページがb 2 4であることを特定し、ページb 1 1を参照してページb 2 4のアドレスを取得し、前記取得したアドレスに基づいてページb 2 4にアクセスする。

また、DBMSは、タスク10Bの実行により、t 8で特定した索引Bの検索条件とページb 2 2のエントリに基づいて、アクセスすべき下位ページが存在しないことを判定する。続いて、タスクを実行するため、タスク生成可能数が1以上のタスク開始情報の中で最初に生成されたタスク開始情報（データアドレス集合1 1 Dとそれに関連づけられたコンテキストから成るタスク開始情報）を参照する。すなわち、データアドレス集合1 1 Dと関連づけられたコンテキストを参照して次にアクセスすべきエントリがB 2であることを特定し、前記次にアクセスすべきエントリを次のページであるB 3に変更し、タスク生成可能数を1減算して1に変更し、データアドレス集合1 1 Dを参照して前記ページB 2のアドレスを取得しようと試みる。しかし、t 7においてデータアドレス集合1 1 Dを記憶するページは、記憶デバイス群4 4 3に追い出されており、DBバッファ4 3 6上に存在しないため、データアドレス集合1 1 Dを記憶するページを記憶デバイス群4 4 3からDBバッファ4 3 6に読込む処理を行う。このため、タスク10Bは、データアドレス集合1 1 Dを記憶するページがDBバッファ4 3 6に読込まれるのを待つ。

また、DBMSは、タスク10Cの実行により、レコードC 1がクエリで指定された条件を満たすことを判定する。続いて、タスクを実行するため、タスク生成可能数が1以上のタスク開始情報の中で最初に生成されたタスク開始情報（データアドレス集合1 1 Dとそれに関連づけられたコンテキストから成るタスク開始情報）を参照する。すなわち、データアドレス集合1 1 Dと関連づけられたコンテキストを参照して次にアクセスすべきエントリがB 3であることを特定し、タスク生成可能数を1減算して0になることを確認

し、データアドレス集合 1 1 D を参照して前記ページ B 3 のアドレスを取得しようと試みる。しかし、t 7 においてデータアドレス集合 1 1 D を記憶するページは、記憶デバイス群 4 4 3 に追い出されており、DBバッファ 4 3 6 上に存在しないため、データアドレス集合 1 1 D を記憶するページを記憶デバイス群 4 4 3 から DBバッファ 4 3 6 に読み込む処理を行う。このため、タスク 1 0 C は、データアドレス集合 1 1 D を記憶するページが DBバッファ 4 3 6 に読み込まれるのを待つ。

[0049] (t 1 1)

DBMS は、タスク 1 0 A の実行により、t 9 で特定した索引 B の検索条件とページ b 2 4 のエントリに基づいて、アクセスすべき下位ページが存在しないことを判定する。続いて、タスクを実行するため、タスク生成可能数が 1 以上のタスク開始情報を検索するが、この時点ではタスク生成可能数が 1 以上のタスク開始情報がないため、タスク 1 0 A を終了する。

また、DBMS は、タスク 1 0 B の実行により、t 1 0 で記憶デバイス群 4 4 3 から DBバッファ 4 3 6 に読み込む処理を行ったデータアドレス集合 1 1 D を参照してページ B 2 のアドレスを取得し、前記取得したアドレスに基づいてページ B 2 にアクセスする。

また、DBMS は、タスク 1 0 C の実行により、t 1 0 で記憶デバイス群 4 4 3 から DBバッファ 4 3 6 に読み込む処理を行ったデータアドレス集合 1 1 D を参照してページ B 3 のアドレスを取得し、タスク生成可能数が 0 になった当該タスク開始情報を削除し、前記取得したアドレスに基づいてページ B 3 にアクセスする。

[0050] (t 1 2)

DBMS は、タスク 1 0 B の実行により、レコード B 2 がクエリで指定された条件を満たさないことを判定する。続いて、タスクを実行するため、タスク生成可能数が 1 以上のタスク開始情報を検索するが、この時点ではタスク生成可能数が 1 以上のタスク開始情報がないため、タスク 1 0 B を終了する。

また、DBMSは、タスク10Cの実行により、レコードB3がクエリで指定された条件を満たさないことを判定する。続いて、タスクを実行するため、タスク生成可能数が1以上のタスク開始情報を検索するが、この時点ではタスク生成可能数が1以上のタスク開始情報がないため、タスク10Cを終了する。

[0051] 以上の処理により、DBへのアクセス処理が終了し、クエリ処理が終了する。

[0052] 実施例1では、図8に示すクエリ処理よりも迅速にクエリ処理を実行できるようにするために、次に示すタスク実行及びタスク開始情報の管理を行う。

[0053] 図9は、実施例1に係るクエリ実行時のタスクの実行及びタスク開始情報の管理を説明するための模式図の一例である。表記のルールは、図8と同様である。

[0054] (t0)～(t2)では、以下に示す点を除いて、それぞれ、図8の(t0)～(t2)と同じ処理が行われる。

(t2)において、DBMSは、DBバッファ436から記憶デバイス群443にデータアドレス集合を記憶するページ追い出すが、ここで、DBバッファ436にデータアドレス集合を記憶するページが複数存在する場合には、DBMSは、単純に最も古いデータアドレス集合(DBバッファ436に格納されているデータアドレス集合のうち最初にDBバッファ436に格納されたデータアドレス集合)を記憶するページを追い出し対象とするのではなく、タスク開始情報の生成スピード(単位時間あたりに生成するタスク開始情報の数)と、タスク開始情報の消費スピード(単位時間あたりに消費するタスク開始情報の数)との相対的な関係、および各タスク開始情報のコンテキストに保持されているタスク生成可能数に基づいて、追い出し対象とするデータアドレス集合を記憶するページを決定する。具体的には、DBMSは、タスク開始情報の生成スピードと、タスク開始情報の消費スピードとの比較結果に基づいて、追い出し対象とするデータアドレス集合を記憶するペ

ージを決定する。互いに比較される生成スピードと消費スピードは、同じ時間帯での生成スピードと消費スピードで良い。

[0055] 具体的には、DBMSは、タスク開始情報の生成スピードが、タスク開始情報の消費スピードよりも大きい場合には、タスク生成可能数が所定の閾値以上のコンテキストと関連づけられているデータアドレス集合を記憶するページを追い出し対象とし、タスク生成可能数が所定の閾値未満のコンテキストと関連づけられているデータアドレス集合を記憶するページを追い出し対象としない。なお、前記閾値は、指定された固定の値でも良く、DBMSによって動的に更新される可変の値でもよい。可変の値としては、例えば、全コンテキストのタスク生成可能数の平均値を計算し、DBMSが任意のタイミングで当該計算した値に更新してもよい。このように、DBMSは、タスク開始情報の総数が増加している傾向がある場合には、タスク生成可能数が多いコンテキストと関連づけられているデータアドレス集合を記憶するページを優先してDBバッファ436から記憶デバイス群443に追い出す。これにより、タスク開始情報を効率よく消費することが期待できる。

[0056] また、DBMSは、タスク開始情報の生成スピードが、タスク開始情報の消費スピード以下の場合には、タスク生成可能数が所定の閾値以上のコンテキストと関連づけられているデータアドレス集合を記憶するページを追い出し対象とせず、タスク生成可能数が所定の閾値未満のコンテキストと関連づけられているデータアドレス集合を記憶するページを追い出し対象とする。このように、DBMSは、タスク開始情報の総数が減少している傾向がある場合には、タスク生成可能数が少ないコンテキストと関連づけられているデータアドレス集合を記憶するページを優先して記憶デバイス群443に追い出す対象として決定し、タスク生成可能数が多いコンテキストと関連づけられているデータアドレス集合を記憶するページをDBバッファ436に優先して残す。これにより、多くのタスクを効率よく実行できることが期待できる。

[0057] (t3)

DBMSは、タスク12Aの実行により、レコードA1がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引Bの検索条件を特定し、索引Bのページb11にアクセスする。

また、DBMSは、タスク12Bの実行により、レコードA2がクエリで指定された条件を満たさないことを判定する。続いて、タスクを実行するため、タスク開始情報を参照する。ここで、実施例1においては、DBMSは、データアドレス集合を記憶するページがDBバッファ436に存在するタスク開始情報を優先して参照する。なお、DBバッファ436にデータアドレス集合を記憶するページが複数ある場合には、DBMSは、タスク開始情報のコンテキストに保持されているタスク生成可能数が少ないタスク開始情報を優先して参照しても良い。DBMSは、タスク12Bの実行により、DBバッファ436に存在するデータアドレス集合13Bと関連づけられたコンテキストを参照して次にアクセスすべきエントリがA3であることを特定し、タスク生成可能数を1減算して0になることを確認し、データアドレス集合13Bを参照して前記ページA3のアドレスを取得し、タスク生成可能数が0になった当該タスク開始情報を削除し、前記取得したアドレスに基づいてページA3にアクセスする。

また、DBMSは、タスク12Cの実行により、レコードA4がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引Bの検索条件を特定し、索引Bのページb11にアクセスする。

[0058] (t4)

DBMSは、タスク12Aの実行により、t3で特定した索引Bの検索条件とページb11のエントリに基づいて、アクセスすべき下位ページがb21であることを特定し、ページb11を参照してページb21のアドレスを取得し、前記取得したアドレスに基づいてページb21にアクセスする。

また、DBMSは、タスク12Bの実行により、レコードA3がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引Bの検索条件を特定し、索引Bのページb11にアクセスする。

また、DBMSは、タスク12Cの実行により、t3で特定した索引Bの検索条件とページb11のエントリに基づいて、アクセスすべき下位ページがb23であることを特定し、ページb11を参照してページb23のアドレスを取得し、前記取得したアドレスに基づいてページb23にアクセスする。

[0059] (t5)

DBMSは、タスク12Aの実行により、t3で特定した索引Bの検索条件とページb21のエントリに基づいて、アクセスすべき下位ページが存在しないことを判定する。続いて、タスクを実行するため、DBバッファ436に存在するデータアドレス集合13Cと関連づけられたコンテキストを参照して、次にアクセスすべきエントリがA5であることを特定し、タスク生成可能数を1減算して0になることを確認し、データアドレス集合13Cを参照して前記ページA5のアドレスを取得し、タスク生成可能数が0になった当該タスク開始情報を削除し、前記取得したアドレスに基づいてページA5にアクセスする。

また、DBMSは、タスク12Bの実行により、t4で特定した索引Bの検索条件とページb11のエントリに基づいて、アクセスすべき下位ページがb22であることを特定し、ページb11を参照してページb22のアドレスを取得し、前記取得したアドレスに基づいてページb22にアクセスする。

また、DBMSは、タスク12Cの実行により、t3で特定した索引Bの検索条件とページb23のエントリに基づいて、アクセスすべきレコードがB1～B3であることを特定し、タスク開始情報を生成する。すなわち、コンテキストを生成し、前記アクセスすべきと特定したページB1～B3のアドレスが格納されているページb23をデータアドレス集合13Dとして前記生成したコンテキストと関連付け、当該コンテキストについて、次にアクセスすべきエントリをB2に設定し、タスク生成可能数を2に設定する。これにより、データアドレス集合13Dとそのデータアドレス集合13Dに関連

付けられたコンテキストとのセットとしてのタスク開始情報が完成する。続いて、DBMSは、タスク12Cの実行により、ページB1にアクセスする。

[0060] (t6)

DBMSは、タスク12Aの実行により、レコードA5がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引Bの検索条件を特定し、索引Bのページb11にアクセスする。

また、DBMSは、タスク12Bの実行により、t4で特定した索引Bの検索条件とページb22のエントリに基づいて、アクセスすべき下位ページが存在しないことを判定する。続いて、タスクを実行するため、DBバッファ436に存在するデータアドレス集合13Dと関連づけられたコンテキストを参照して、次にアクセスすべきエントリがB2であることを特定し、前記次にアクセスすべきエントリを次のページであるB3に変更し、タスク生成可能数を1減算して1に変更し、データアドレス集合13Dを参照して前記ページB2のアドレスを取得し、前記取得したアドレスに基づいてページB2にアクセスする。

また、DBMSは、タスク12Cの実行により、レコードB1がクエリで指定された条件を満たすことを判定し、当該レコードに基づいて索引Cの検索条件を特定し、索引Cのページc11にアクセスする。

[0061] (t7)

DBMSは、タスク12Aの実行により、t6で特定した索引Bの検索条件とページb11のエントリに基づいて、アクセスすべき下位ページがb24であることを特定し、ページb11を参照してページb24のアドレスを取得し、前記取得したアドレスに基づいてページb24にアクセスする。

DBMSは、タスク12Bの実行により、レコードB2がクエリで指定された条件を満たさないことを判定する。続いて、タスクを実行するため、DBバッファ436に存在するデータアドレス集合13Dと関連づけられたコンテキストを参照して、次にアクセスすべきエントリがB3であることを特定

し、タスク生成可能数を1減算して0になることを確認し、データアドレス集合13Dを参照して前記ページB3のアドレスを取得し、タスク生成可能数が0になった当該タスク開始情報を削除し、前記取得したアドレスに基づいてページB3にアクセスする。

また、DBMSは、タスク12Cの実行により、t6で特定した索引Cの検索条件とページc11のエントリに基づいて、アクセスすべき下位ページがc21であることを特定し、ページc11を参照してページc21のアドレスを取得し、前記取得したアドレスに基づいてページc21にアクセスする。

[0062] なお、ここで、データアドレス集合13Dが消費されDBバッファ436にタスク開始情報のデータアドレス集合を記憶するページがなくなるため、DBMSは、記憶デバイス群443からデータアドレス集合13Aを記憶するページをDBバッファ436に読み込む処理を行う。記憶デバイス群443にデータアドレス集合を記憶するページが複数存在する場合には、DBMSは、読み込み対象のデータアドレス集合を、例えば、タスク開始情報の生成スピードとタスク開始情報の消費スピードとの比較結果、およびタスク開始情報のコンテキストに保持されているタスク生成可能数に基づいて決定する。

[0063] 具体的には、DBMSは、タスク開始情報の生成スピードが、タスク開始情報の消費スピードよりも小さい場合には、タスク生成可能数が所定の閾値以上のコンテキストと関連づけられているデータアドレス集合を記憶するページを読み込み対象とし、タスク生成可能数が所定の閾値未満のコンテキストと関連づけられているデータアドレス集合を記憶するページを読み込み対象としない。なお、前記閾値は、指定された固定の値でも良く、DBMSによって動的に更新される可変の値でもよい。可変の値としては、例えば、全コンテキストのタスク生成可能数の平均値を計算し、DBMSが任意のタイミングで当該計算した値に更新してもよい。このように、タスク開始情報の総数が減少している傾向がある場合には、DBMSは、タスク生成可能数が多いコンテキストと関連づけられているデータアドレス集合を記憶するページを

優先して読込む。これにより、多くのタスクを効率よく実行できることが期待できる。

[0064] また、DBMSは、タスク開始情報の生成スピードが、タスク開始情報の消費スピード以上の場合には、タスク生成可能数が所定の閾値以上のコンテキストと関連づけられているデータアドレス集合を記憶するページを読込み対象とせず、タスク生成可能数が所定の閾値未満のコンテキストと関連づけられているデータアドレス集合を記憶するページを読込み対象とする。このように、タスク開始情報の総数が増加している傾向がある場合には、DBMSは、タスク生成可能数が少ないコンテキストと関連づけられているデータアドレス集合を記憶するページを優先して読み込む。これにより、タスク開始情報を効率よく消費することが期待できる。

[0065] 図9の例では、DBMSは、データアドレス集合13Aを記憶するページを、DBバッファ436に読込む。

[0066] (t8)

DBMSは、タスク12Aの実行により、t6で特定した索引Bの検索条件とページb24のエントリに基づいて、アクセスすべき下位ページが存在しないことを判定する。続いて、タスクを実行するため、DBバッファ436に存在するデータアドレス集合13Aと関連づけられたコンテキストを参照して、次にアクセスすべきエントリがa24であることを特定し、前記次にアクセスすべきエントリを次のページであるa25に変更し、タスク生成可能数を1減算して1に変更し、データアドレス集合13Aを参照して前記ページa24のアドレスを取得し、前記取得したアドレスに基づいてページa24にアクセスする。

また、DBMSは、タスク12Bの実行により、レコードB3がクエリで指定された条件を満たさないことを判定する。続いて、タスクを実行するため、DBバッファ436に存在するデータアドレス集合13Aと関連づけられたコンテキストを参照して、次にアクセスすべきエントリがa25であることを特定し、タスク生成可能数を1減算して0になることを確認し、データ

アドレス集合 1 3 A を参照して前記ページ a 2 5 のアドレスを取得し、当該タスク開始情報を消去し、前記取得したアドレスに基づいてページ a 2 5 にアクセスする。

また、DBMS は、タスク 1 2 C の実行により、t 6 で特定した索引 C の検索条件とページ c 2 1 のエントリに基づいて、アクセスすべき下位ページが c 3 1 であることを特定し、ページ c 2 1 を参照してページ c 3 1 のアドレスを取得し、前記取得したアドレスに基づいてページ c 3 1 にアクセスする。

[0067] (t 9)

DBMS は、タスク 1 2 A の実行により、クエリの指定値とページ a 2 4 のエントリに基づいて、アクセスすべきレコードが A 6 であることを特定し、ページ a 2 4 を参照して前記レコード A 6 のアドレスを取得し、前記取得したアドレスに基づいてページ A 6 にアクセスする。

また、DBMS は、タスク 1 2 B の実行により、クエリの指定値とページ a 2 5 のエントリに基づいて、アクセスすべきレコードが A 7 であることを特定し、ページ a 2 5 を参照して前記レコード A 7 のアドレスを取得し、前記取得したアドレスに基づいてページ A 7 にアクセスする。

また、DBMS は、タスク 1 2 C の実行により、t 6 で特定した索引 C の検索条件とページ c 3 1 のエントリに基づいて、アクセスすべきレコードが C 1 であることを特定し、ページ c 3 1 を参照して前記レコード C 1 のアドレスを取得し、前記取得したアドレスに基づいてページ C 1 にアクセスする。

[0068] (t 1 0)

DBMS は、タスク 1 2 A の実行により、レコード A 6 がクエリで指定された条件を満たさないことを判定する。続いて、タスクを実行するため、タスク生成可能数が 1 以上のタスク開始情報を検索するが、この時点ではタスク生成可能数が 1 以上のタスク開始情報がないため、タスク 1 2 A を終了する。

また、DBMS は、タスク 1 2 B の実行により、レコード A 7 がクエリで指

定された条件を満たさないことを判定する。続いて、タスクを実行するため、タスク生成可能数が1以上のタスク開始情報を検索するが、この時点ではタスク生成可能数が1以上のタスク開始情報がないため、タスク12Bを終了する。

また、DBMSは、タスク12Cの実行により、レコードC1がクエリで指定された条件を満たすことを判定する。続いて、タスクを実行するため、タスク生成可能数が1以上のタスク開始情報を検索するが、この時点ではタスク生成可能数が1以上のタスク開始情報がないため、タスク12Cを終了する。

[0069] 以上の処理により、t9までに、クエリ実行プランにおけるDBへのアクセス処理が全て終了し、その後、クエリ処理が終了する。

[0070] 実施例1に係るクエリ処理は、図9に示すように、図8に示すクエリ処理に比して、記憶デバイス群443へのアクセスの回数を低減でき、短時間でクエリ処理を実行することができる。

[0071] 以下、実施例1を詳細に説明する。

[0072] 図10は、実施例1に係る計算機システムの構成を示す。

[0073] DBサーバ401が通信ネットワーク403を介して外部ストレージ装置402に接続されている。通信ネットワーク403を介した通信の Protokol としては、例えば、FC (Fibre Channel)、SCSI (Small Computer System Interface)、又は、TCP/IP (Transmission Control Protocol/Internet Protocol) が採用されて良い。

[0074] DBサーバ401は、計算機、例えば、パーソナルコンピュータ、ワークステーション又はメインフレーム、もしくは、これらのいずれかによって構成される仮想的な計算機 (仮想マシン) である。DBサーバ401は、ネットワークアダプタ413、メモリ416、ローカル記憶デバイス415及びそれらに接続されたプロセッサ (典型的にはマイクロプロセッサ) 414を有する。プロセッサ414は、コンピュータプログラム、例えば、OS (Operating System) 417と、DBMS 412と、DBMS 412にクエリを

発行するAP (Application Program) 411とを実行する。メモリ416は、主記憶装置（第1の記憶デバイス）の一例であり、プロセッサ414によって実行されるプログラムと、プログラムが使用するデータとを一時的に記憶する。ローカル記憶デバイス415は、プログラム、及びプログラムが使用するデータを格納する。ネットワークアダプタ413は、通信ネットワーク403とDBサーバ401とを接続する。AP411は、DBサーバ401ではなく、通信ネットワーク403に接続される図示しない別の計算機で動作しても良い。また、プロセッサ414は、ネットワークアダプタ413及びメモリ416等に接続された制御デバイスに含まれている要素で良い。制御デバイスは、プロセッサ414の他に、専用ハードウェア回路（例えば、データの暗号化及び／又は復号化を行う回路）を含んで良い。

[0075] なお、DBサーバ401は、性能面や冗長性の観点から、プロセッサ414、メモリ416、ローカル記憶デバイス415及びネットワークアダプタ413のうちの少なくとも1つの要素を複数備えていても良い。また、DBサーバ401は、図示しない入力デバイス（例えば、キーボード及びポインティングデバイス）と表示デバイス（例えば液晶ディスプレイ）とを有して良い。入力デバイスと表示デバイスは一体になっていても良い。

[0076] DBサーバ401では、DBMS412が、AP411から発行されたクエリを実行し、そのクエリの実行に伴い、外部ストレージ装置402に格納されたDB451に対するI/O要求をOS417に発行する。OS417が、DBMS412から発行されたI/O要求を、外部ストレージ装置402に送信する。

[0077] 外部ストレージ装置402は、本実施例では、ディスクアレイ装置のような、複数の記憶デバイスを含む記憶デバイス群443を有する装置であるが、それに代えて、単一の記憶デバイスであっても良い。外部ストレージ装置402は、DBサーバ401が使用するデータ及びプログラムを記憶する。外部ストレージ装置402は、DBサーバ401に対する二次記憶装置（第2の記憶デバイス）の一例である。外部ストレージ装置402は、DBサー

バ４０１からＩ／Ｏ要求を受信し、Ｉ／Ｏ要求に対応した処理を実行し、処理結果をＤＢサーバ４０１に送信する。

[0078] 外部ストレージ装置４０２は、ネットワークアダプタ４４１、記憶デバイス群４４３及びそれらに接続されたコントローラ４４２を有する。

[0079] ネットワークアダプタ４４１は、外部ストレージ装置４０２を通信ネットワーク４０３に接続する。

[0080] 記憶デバイス群４４３は、１つ以上の記憶デバイスを含む。記憶デバイスは、不揮発性の記憶媒体であって、例えば、磁気ディスク、フラッシュメモリ、その他半導体メモリである。記憶デバイス群４４３は、ＲＡＩＤ（Redundant ARRAY of Independent Disks）に従い所定のＲＡＩＤレベルでデータを記憶するグループであっても良い。記憶デバイス群４４３の記憶空間に基づく論理的な記憶デバイス（論理ボリューム）がＤＢサーバ４０１に提供されても良い。記憶デバイス群４４３は、ＤＢ４５１を記憶する。

[0081] コントローラ４４２は、例えば、メモリ及びプロセッサを含んでおり、ＤＢサーバ４０１からのＩ／Ｏ要求に従って、ＤＢ４５１を記憶した記憶デバイス群４４３にデータを入出力する。例えば、コントローラ４４２は、ＤＢサーバ４０１からの書込み要求に従う書込み対象のデータを記憶デバイス群４４３に格納したり、ＤＢサーバ４０１からの読出し要求に従う読出し対象のデータを記憶デバイス群４４３から読み出し、そのデータをＤＢサーバ４０１に送信したりする。

[0082] なお、外部ストレージ装置４０２は、性能面や冗長性確保の観点から、コントローラ４４２等の要素を複数備えても良い。

[0083] ＤＢＭＳ４１２は、業務データを含んだＤＢ４５１を管理する。ＤＢ４５１は、１つ以上の表４６２や索引４６１を含む。表は１つ以上のレコードの集合であり、レコードは１つ以上のカラムから構成される。索引は、表の中の１つ以上のカラムを対象に作成されるデータ構造であり、当該索引が対象とするカラムを含む選択条件による表へのアクセスを高速化する。例えば、索引は、対象とするカラムの値毎にその値を含む表のレコードを特定する情

報（レコードID）を保持するデータ構造であり、B木構造等が用いられる。DBの表の構成例や表同士の関連性の一例は、図1乃至図3、及び図6に示した通りである。

[0084] DBMS 412は、クエリ受付部421、クエリ実行プラン生成部422、クエリ実行部423、実行タスク管理部426及びDBバッファ管理部427を含む。

[0085] クエリ受付部421は、AP 411が発行するクエリを受け付ける。クエリは、例えばSQL（Structured Query Language）で記述されている。

[0086] クエリ実行プラン生成部422は、クエリ受付部421が受け付けたクエリについて、当該クエリを実行するために必要な1つ以上のDBオペレーションを有するクエリ実行プランを生成する。クエリ実行プランは、例えば、クエリの実行の際に行うべきDBオペレーションの実行順序を木構造で定義した情報であり、メモリ416に格納される。クエリ実行プランの一例は、図5に示した通りである。

[0087] クエリ実行部423は、クエリ実行プラン生成部422が生成したクエリ実行プランに従って、クエリ実行プランに含まれる情報であるDBオペレーションを実行するためのタスクを動的に生成して実行する。そして、実行した結果をクエリ発行元に返す。例えば、クエリ実行部423は、（a）DBオペレーションを実行するためのタスクを実行タスク管理部426に要求して生成すること、（b）生成されたタスクを実行することで、そのタスクに対応したDBオペレーションに必要なデータを読み出す（DB 451を有する外部ストレージ装置402へOS 417を介して読出し要求を発行すること、（c）上記（b）で実行されたタスクに対応したN番目のDBオペレーションの実行結果に基づき（N+1）番目のDBオペレーションを実行する場合には、その実行結果に基づくタスクを実行タスク管理部426に要求して新たに生成すること（Nは1以上の整数）、及び、（d）その新たに生成したタスクについて上記（b）及び上記（c）を行うこと、を行う。クエリ実行部423は、上記（b）及び（d）において、2以上の実行可能なタ

スクが存在する場合には、それら2以上のタスクのうちの少なくとも二つのタスクを並行して実行する。また、クエリ実行部423は、タスク割当て管理部431と、コンテキスト管理部432とを含む。タスク割当て管理部431は、タスクの実行開始を管理する。コンテキスト管理部432は、タスクにおいて実行を行う内容を示す情報であるタスク開始情報TSのコンテキストCTXを管理する。タスク開始情報の一例は、図7に示した通りである。なお、クエリ実行部423、タスク割当て管理部431、及びコンテキスト管理部432の詳細は、後述する。

[0088] 実行タスク管理部426は、クエリを実行するためのタスクを管理する。例えば、実行タスク管理部426は、タスクを生成するためのメモリ資源を管理しており、クエリ実行部423からのタスクの生成要求に従って、タスク自体に要するメモリ資源を確保してタスクを生成する。また、クエリ実行部423からのタスクの終了要求に従って、タスク自体に要するメモリ資源を解放してタスクを終了する。ここで、タスクとしては、任意のモジュールを採用することができる。例えば、タスクは、OS417が管理するプロセス又はスレッドでも良いし、DBMS412で実装される疑似プロセス又は疑似スレッドでも良い。

[0089] DBバッファ管理部427は、DB451内のページを一時的に格納するための1以上の記憶領域（DBバッファ面）437を管理する。DBバッファ面437は、メモリ416上に構築される。なお、DBバッファ管理部427の詳細は、後述する。

[0090] 図11は、実施例1に係る計算機システムの一部の詳細構成図である。

[0091] クエリ実行部423は、タスク割当て管理部431と、コンテキスト管理部432とを含む。クエリ実行部423は、クエリ実行プラン生成部422が生成したクエリ実行プランに従って、タスクにおいて実行を行う内容を示す情報であるタスク開始情報TSを動的に生成し、前記生成したタスク開始情報に基づいてタスクを実行することでクエリ実行プラン内のDBオペレーションを実行する。コンテキスト管理部432は、1以上のタスク開始情報

T SのコンテキストC T Xを管理する。コンテキストC T Xは、データアドレス集合A Gが記憶されている外部ストレージ装置4 0 2の記憶デバイス群4 4 3におけるページを特定する情報（例えば、ページI D）、データアドレス集合A Gに格納されたエントリのうち次に実行開始されるタスクによりアクセスするエントリを示す情報、さらにタスク実行によりアクセスするエントリの残数（タスク生成可能数）を格納する。コンテキスト管理部4 3 2は、タスク開始情報T Sが生成される都度に、当該タスク開始情報のコンテキストC T Xを管理対象に追加し、タスク開始情報T Sが削除される都度に、当該タスク開始情報のコンテキストC T Xを管理対象から削除する。タスク割当て管理部4 3 1は、クエリ実行部4 2 3がタスクを実行開始する際に参照するタスク開始情報T Sを決定し、当該タスク開始情報のコンテキストC T Xをコンテキスト管理部より取得し、当該コンテキストと関連づけられたデータアドレス集合A Gを参照し、アクセス対象ページを特定する。クエリ実行部4 2 3は、タスクを実行開始し、前記特定されたページにアクセスする。また、タスク割当て管理部4 3 1は、コンテキストC T Xのタスク生成可能数が0になったタスク開始情報T Sを削除する。

[0092] DBバッファ管理部4 2 7は、DB 4 5 1内のページを一時的に格納するための1以上の記憶領域（DBバッファ面）4 3 7を含むDBバッファ4 3 6を管理する。なお、DBバッファ4 3 6におけるDBバッファ面4 3 7の数は、所定の数に制限されている。また、DBバッファ管理部4 2 7は、DBバッファ面4 3 7に格納されているページを管理するためのDBバッファ管理表4 3 5を記憶する。DBバッファ管理部4 2 7は、DB 4 5 1のページの内容を、DBバッファ面4 3 7に読込んだり、DBバッファ面4 3 7に格納されているページを、DBバッファ面4 3 7から追い出したりする。ここで、DBバッファ面4 3 7からデータを追い出すことを、DBバッファ面4 3 7の解放という。なお、DBバッファ管理部4 2 7は、DBバッファ面4 3 7を解放する際に、該当するデータがダーティデータ、すなわち、DB 4 5 1に反映されていないデータである場合には、当該ページをDB 4 5 1

に移動して格納する一方、ダーティデータでない場合には、そのデータをDBバッファ面437から消去するのみとしても良い。

[0093] 前述したクエリ受付部421、クエリ実行プラン生成部422、クエリ実行部423、実行タスク管理部426、DBバッファ管理部427は、DBMS412がプロセッサ414で実行されることによって実現されるが、それらのうちの少なくとも1つの処理部が行う処理の少なくとも一部が、ハードウェアで行われても良い。また、本実施例の説明において、処理部が主語になる場合は、実際にはその処理部を実行するプロセッサ414によって処理が行われるが、処理部の少なくとも一部がハードウェアで実現されている場合は、プロセッサ414に代えて又は加えて、そのハードウェアも、主語とされ得る。DBMS412のようなコンピュータプログラムは、プログラムソースからDBサーバ401にインストールされて良い。プログラムソースは、例えば、DBサーバ401が読み取り可能な記憶メディアで良い。

[0094] また、図10及び図11に示したDBMS412の構成は、一例である。例えば、或る処理部が複数の処理部に分割されたり、複数の処理部の処理を統合した1つの処理部が構築されても良い。

[0095] 図12は、実施例1に係るDBバッファ管理表435の構成を示す。

[0096] DBバッファ管理表435は、DBバッファ面437毎の情報として、バッファID435a、ページID435bを有する。各種情報は、下記の通りである。

(*) バッファID435aは、DBバッファ面437を識別するためのIDを示す。

(*) ページID435bは、対応するDBバッファ面437に格納されているページの外部ストレージ装置402における格納位置を特定することのできる情報（ここでは、ページID）である。なお、対応するDBバッファ面437にページが格納されていない場合には、ページID435bとして空を示す情報が設定される。

[0097] 以下、図13～図22を参照して、本実施例で行われる種々の処理を説明

する。

[0098] 図13は、DBMS412が受け付けたクエリの実行においてタスクを生成し実行するクエリ処理の全体の流れを示す。図13は、1つのクエリ実行の流れを示すが、同時に実行するクエリが2以上ある場合には、DBMS412は、その2以上のクエリについて並行して図13に示す処理を行って良い。

[0099] S1001で、クエリ受付部421が、AP411が発行するクエリを受け付け、例えば、メモリ416に格納する。クエリは、例えばSQL (Structured Query Language) で記述されている。

[0100] S1002で、クエリ実行プラン生成部422が、DB451のデータの統計情報（以下、「DB統計情報」）に基づいて、クエリ受付部421が受け付けたクエリ（例えば、メモリ416に蓄積されているクエリ）についてクエリ実行プランを生成する。この際、1つ以上のクエリ実行プランの候補の各々についてコスト計算を行い、そのうち1つのクエリ実行プランの候補をクエリ実行プランとして選んでも良い。DB統計情報は、例えば、各表462について、レコードの数を表す情報、カラム毎のデータの最大値及び最小値もしくはカラム毎のデータの分布状況を表すヒストグラム等の統計情報を含んで良い。DB統計情報は、ローカル記憶デバイス415又はメモリ416に格納される。DB統計情報は、DBMS412がDB451をチェックすることにより生成されても良いし、DBMS412の外部から入力されても良い。また、クエリ実行プランは、DB統計情報のみならず、クエリの記述に基づいて生成されても良い。

[0101] S1003で、クエリ実行部423がコンテキストを生成し、コンテキスト管理部432が当該生成されたコンテキストを管理する。前記コンテキストは、S1002で生成されたクエリ実行プランの、先頭のDBオペレーションを実行するための内容を示すコンテキストである。

[0102] S1004で、クエリ実行部423が、1以上のタスクを生成する。

[0103] S1005で、クエリ実行部423が、S1004で生成されたタスクを

実行する。

[0104] 図14は、実施例1に係るDBMSによるタスク実行処理の流れを示す。

[0105] S1101で、タスク割当て管理部431が、タスクの実行内容を取得する。タスク実行内容取得処理の詳細は、図15に示す通りである。

[0106] S1102で、クエリ実行部423が、S1101で取得したタスクの実行内容に基づいてタスクを実行することにより、DBオペレーションを実行する。

[0107] S1103で、クエリ実行部423が、S1102でタスクにより実行したDBオペレーションの結果に基づき、コンテキストを生成するか否かを判断する。この判断の結果が肯定的であれば、S1104を行う一方、この判断の結果が否定的であれば、処理を終了する。

[0108] S1104で、クエリ実行部423がコンテキストCTXを生成し、コンテキスト管理部432が、当該生成されたコンテキストを管理する。当該コンテキストCTXには、データアドレス集合AGが記憶されている外部ストレージ装置402の記憶デバイス群443におけるページを特定する情報（例えば、ページID）、データアドレス集合AGに格納されたエントリのうち次に実行開始されるタスクによりアクセスするエントリを示す情報、さらにタスク実行によりアクセスするエントリの残数（タスク生成可能数）が含まれる。

[0109] S1105で、クエリ実行部423からのタスクの生成要求に従って、実行タスク管理部426が、タスクを生成する。

[0110] S1106で、クエリ実行部423が、S1104で生成されたタスクを実行する。

[0111] 図15は、実施例1に係るタスク実行内容取得処理（図14のS1101）の流れを示す。

[0112] S1201で、タスク割当て管理部431は、タスク開始情報選択処理を行う。タスク開始情報TSとは、コンテキストCTXとそれに関連づけられたデータアドレス集合AGのセットである。タスク開始情報選択処理の詳細

は、図 1 6 に示す通りである。

[0113] S 1 2 0 2 で、タスク割当て管理部 4 3 1 は、S 1 2 0 1 で選択したタスク開始情報 T S からタスク実行内容を取得する。タスク実行内容には、タスク開始情報 T S のコンテキスト C T X 及びデータアドレス集合 A G から取得した、アクセス対象ページのアドレスを含む。

[0114] S 1 2 0 3 で、タスク割当て管理部 4 3 1 は、S 1 2 0 1 で選択したタスク開始情報 T S のコンテキスト C T X に保持されている、次に実行開始されるタスクによりアクセスするエントリを次のエントリに更新し、前記コンテキスト C T X のタスク生成可能数を 1 減算する。

[0115] S 1 2 0 4 で、タスク割当て管理部 4 3 1 は、S 1 2 0 1 で選択したタスク開始情報 T S のコンテキスト C T X のタスク生成可能数が 0 であるか否かを判断する。この判断の結果が肯定的であれば、S 1 2 0 5 を行う一方、この判断の結果が否定的であれば、処理を終了する。

[0116] S 1 2 0 5 で、タスク割当て管理部 4 3 1 は、S 1 2 0 1 で選択したタスク開始情報 T S を削除し、コンテキスト管理部 4 3 2 は当該タスク開始情報 T S のコンテキスト C T X を管理対象から削除する。

[0117] 図 1 6 は、タスク開始情報選択処理（図 1 5 の S 1 2 0 1）の流れを示す。

[0118] S 1 3 0 1 で、タスク割当て管理部 4 3 1 は、コンテキスト管理部 4 3 2 が管理しているコンテキストから 1 つを選択し、当該コンテキストと関連づけられているデータアドレス集合を記憶するページ（ここで、この処理の説明においては、このページをページ B という）のページ I D を取得する。

[0119] S 1 3 0 2 で、タスク割当て管理部 4 3 1 は、取得したページ B が D B バッファ 4 3 6 上にあるか否かを判定する。具体的には、タスク割当て管理部 4 3 1 は、ページ I D を含む問合せ要求を D B バッファ管理部 4 2 7 に送信することにより、対応するページ I D のページが D B バッファ 4 3 6 に格納されているか否かを問合せ、その問合せ結果に基づいて、取得したページ B が D B バッファ 4 3 6 上にあるか否かを判定する。D B バッファ管理部 4 2

7は、前記問合せを受けて、DBバッファ管理表435を参照し、該当するページIDのページがあるか否かを判定し、その結果をタスク割当て管理部431に通知する。

[0120] S1303で、ステップS1302の判定の結果、ページBがDBバッファ436にある場合（S1303：Y）には、タスク割当て管理部431は、ステップS1307を行い、ページBがDBバッファ436にない場合（S1303：N）には、ステップS1304を行う。

[0121] S1304で、タスク割当て管理部431は、コンテキスト管理部432の全コンテキストにおいて当該コンテキストに関連づけられているデータアドレス集合を記憶するページがDBバッファ436にないと判定していない場合（S1304：N）には、S1301に進み、次のコンテキストを選択して、以降の処理を行う。一方、コンテキスト管理部432の全コンテキストにおいて当該コンテキストに関連づけられているデータアドレス集合を記憶するページがDBバッファ436にないと判定した場合（S1304：Y）には、タスク割当て管理部431は、ページ読み込みコンテキストを選択する処理（S1305）を行う。

[0122] S1305で、タスク割当て管理部431は、関連づけられているデータアドレス集合を記憶するページを、外部ストレージ装置402からDBバッファ436に読み込むコンテキストを選択するための、ページ読み込みコンテキスト選択処理を実行する。ページ読み込みコンテキスト選択処理の詳細は、図17に示す通りである。図17によれば、ページ読み込みコンテキスト選択処理により、外部ストレージ装置402からDBバッファ436にページを読み込むコンテキストが選択される。ここで、選択したページをページCとし、選択したコンテキストをコンテキストCとする。

[0123] S1306で、タスク割当て管理部431は、DBバッファ管理部427に対し、ページCをDBバッファ436に読み込む、ページ読み込み要求を送信する。DBバッファ管理部427は、前記ページ読み込み要求を受けて、外部ストレージ装置402からページCをDBバッファ436に読み込み、ページ読み

処理が終了したことをタスク割当て管理部431に通知する。ページ読込処理の詳細は、図19に示す通りである。図19によれば、ページ読込処理により、ページCがDBバッファ436に読込まれる。

[0124] S1307で、タスク割当て管理部431は、選択したコンテキストを含むタスク開始情報を選択し、処理を終了する。

[0125] 図17は、ページ読込コンテキスト選択処理（図16のS1305）の流れを示す。

[0126] S1401で、タスク割当て管理部431は、コンテキスト管理部432が管理しているコンテキストから、処理対象とする1つのコンテキストを選択する。

[0127] S1402で、タスク割当て管理部431は、選択したコンテキストと関連づけられているデータアドレス集合を記憶するページをDBバッファ436に読み込み可能であるか否かを判定するための、DBバッファ読込判定処理を実行する。DBバッファ読込判定処理の詳細は、図18に示す通りである。図18によれば、DBバッファ読込判定処理により、選択されたコンテキストと関連づけられているページをDBバッファ436に読込可能であるか否かの判定結果が得られる。

[0128] S1403で、タスク割当て管理部431は、選択したコンテキストと関連づけられているページをDBバッファ436に読込可能か否かを判定し、当該ページをDBバッファ436に読込可能でない場合（S1403：N）には、S1401を行い、次のコンテキストを選択して以降の処理を行う一方、当該ページをDBバッファ436に読込可能である場合（S1403：Y）には、処理を終了する。

[0129] 図18は、DBバッファ読込判定処理（図17のS1402）の流れを示す。

[0130] S1501で、タスク割当て管理部431は、単位時間当たりのタスク開始情報生成数が、単位時間当たりのタスク開始情報消費数よりも小さいか否かを判定し、単位時間当たりのタスク開始情報生成数が、単位時間当たりの

タスク開始情報消費数よりも小さい場合、すなわち、タスク開始情報の総数が減少している傾向がある場合（S 1 5 0 1 : Y）には、タスク生成可能数が多いコンテキストと関連づけられているデータアドレス集合を記憶するページを優先してDBバッファ436に読込むためS 1 5 0 2に進む一方、単位時間当たりのタスク開始情報生成数が、単位時間当たりのタスク開始情報消費数よりも小さくない場合、すなわち、タスク開始情報の総数が増加する傾向にある場合（S 1 5 0 1 : N）には、S 1 5 0 3に進む。

[0131] S 1 5 0 2で、タスク割当て管理部431は、S 1 4 0 1で選択されたコンテキストのタスク生成可能数が所定の閾値以上であるか否かを判定し、その結果、タスク生成可能数が閾値以上である場合（S 1 5 0 2 : Y）には、S 1 5 0 4に進む一方、タスク生成可能数が閾値以上でない場合（S 1 5 0 2 : N）には、S 1 5 0 5に進む。

[0132] S 1 5 0 3で、タスク割当て管理部431は、S 1 4 0 1で選択されたコンテキストのタスク生成可能数が所定の閾値以上であるか否かを判定し、その結果、タスク生成可能数が閾値以上である場合（S 1 5 0 3 : Y）には、S 1 5 0 5に進む一方、タスク生成可能数が閾値以上でない場合（S 1 5 0 3 : N）には、S 1 5 0 4に進む。

[0133] S 1 5 0 4で、タスク割当て管理部431は、選択されたコンテキストと関連づけられているページを読込可と判定し、DBバッファ読込判定処理を終了する。

[0134] S 1 5 0 5で、タスク割当て管理部431は、選択されたコンテキストと関連づけられているページを読込不可と判定し、DBバッファ読込判定処理を終了する。

[0135] このDBバッファ読込判定処理によると、タスク開始情報の総数が減少している傾向がある場合には、タスク生成可能数が多いコンテキストと関連づけられているページがDBバッファ436に優先して読込まれるようにして多くのタスクを効率よく実行でき、タスク開始情報の総数が増加している傾向がある場合には、タスク生成可能数が少ないコンテキストと関連づけられ

ているページがDBバッファに優先して読込まれるようにして、タスク開始情報が効率よく消費されるようにできる。

[0136] 図19は、ページ読込処理（図16のS1306）の流れを示す。

[0137] S1601で、DBバッファ管理部427は、空いているDBバッファ面が無いかな否かを判定し、空きDBバッファ面が無い場合（S1601：Y）には、S1602に進む一方、空きDBバッファ面がある場合（S1601：N）には、S1603に進む。

[0138] S1602で、DBバッファ管理部427は、DBバッファ436からデータを追出すDBバッファ追い出し処理を実行する。DBバッファ追い出し処理の詳細は、図20に示す通りである。図20によれば、DBバッファ追い出し処理により、空きDBバッファ面を生成することができる。DBバッファ追い出し処理終了後には、DBバッファ管理部427は、S1601を行う。

[0139] S1603で、DBバッファ管理部427は、ページのデータを格納するための空きDBバッファ面を確保し、当該DBバッファ面にページのデータを読込む。

[0140] 図20は、DBバッファ追い出し処理（図19のS1602）の流れを示す。

[0141] S1701で、DBバッファ管理部427は、データを追出す対象とするDBバッファ面を選択する、追い出し対象DBバッファ面選択処理を実行する。追い出し対象DBバッファ面選択処理の詳細は、図21に示す通りである。図21によれば、追い出し対象DBバッファ面選択処理により、データを追出す対象とするDBバッファ面を選択することができる。

[0142] S1702で、DBバッファ管理部427は、選択されたDBバッファ面のデータがダーティデータかな否かを判定し、選択されたDBバッファ面のデータがダーティデータである場合（S1702：Y）には、S1703に進み、選択されたDBバッファ面のデータがダーティデータでない場合（S1702：N）には、S1704に進む。

- [0143] S 1 7 0 3 で、DBバッファ管理部 4 2 7 は、選択された DB バッファ面の内容を外部ストレージ装置 4 0 2 の記憶デバイス群 4 4 3 に書き出す。
- [0144] S 1 7 0 4 で、DBバッファ管理部 4 2 7 は、選択された DB バッファ面を開放する。これにより、この DB バッファ面が使用可能となる。
- [0145] なお、この DB バッファ追い出し処理は、例えば、DB バッファ 4 3 6 における空き DB バッファ面数が所定数以下になった場合に、実行するようにしても良い。
- [0146] 図 2 1 は、追い出し対象 DB バッファ面選択処理（図 2 0 の S 1 7 0 1）の流れを示す。
- [0147] S 1 8 0 1 で、DBバッファ管理部 4 2 7 は、追い出し候補とする 1 つの DB バッファ面を選択する。
- [0148] S 1 8 0 2 で、DBバッファ管理部 4 2 7 は、選択された DB バッファ面に格納されているページが、コンテキストと関連づけられているページであるか否かを判定し、当該ページがコンテキストと関連づけられているページである場合（S 1 8 0 2 : Y）には、S 1 8 0 3 に進む一方、当該ページがコンテキストと関連づけられているページでない場合（S 1 8 0 2 : N）には、そのまま処理を終了する。ここで、選択した DB バッファ面に格納されているページが、コンテキストと関連づけられているページであるか否かの判定は、具体的には、DB バッファ管理部 4 2 7 が、ページ ID を含む問合せ要求をコンテキスト管理部 4 3 2 に送信することにより、対応するページ ID のページと関連づけられているコンテキストが、コンテキスト管理部 4 3 2 で管理されているか否かを問合せ、その問合せ結果に基づいて判断する。コンテキスト管理部 4 3 2 は、前記問合せを受けて、管理するコンテキストのなかに、該当するページ ID が関連づけられているコンテキストがあるか否かを判定して、その結果を DB バッファ管理部 4 2 7 に通知する。
- [0149] S 1 8 0 3 で、DBバッファ管理部 4 2 7 は、タスク割当て管理部 4 3 1 により、選択した DB バッファ面を解放可能か否かを判定するための DB バッファ面解放判定処理を実行する。DB バッファ面解放判定処理の詳細は、

図 2 2 に示す通りである。図 2 2 によれば、DBバッファ面解放判定処理により、選択されたDBバッファを解放可能であるか否かの判定結果が得られる。

[0150] S 1 8 0 4 で、DBバッファ管理部 4 2 7 は、選択したDBバッファ面を解放可能であるか否かを判定し、選択したDBバッファを解放可能と判定しなかった場合（S 1 8 0 4 : N）には、S 1 8 0 1 を行い、次のDBバッファ面を選択して以降の処理を行う一方、選択したDBバッファ面を解放可能と判定した場合（S 1 8 0 4 : Y）には、処理を終了する。

[0151] 図 2 2 は、DBバッファ面解放判定処理（図 2 1 の S 1 8 0 3）の流れを示す。

[0152] S 1 9 0 1 で、タスク割当て管理部 4 3 1 は、単位時間当たりのタスク開始情報生成数が、単位時間当たりのタスク開始情報消費数より大きいかなかを判定し、単位時間当たりのタスク開始情報生成数が、単位時間当たりのタスク開始情報消費数よりも大きい場合、すなわち、タスク開始情報の総数が増加している傾向がある場合（S 1 9 0 1 : Y）には、タスク生成可能数が多いコンテキストと関連づけられているデータアドレス集合を記憶するページを優先してDBバッファ 4 3 6 から追い出すため S 1 9 0 2 に進む一方、単位時間当たりのタスク開始情報生成数が、単位時間当たりのタスク開始情報消費数よりも大きくない場合、すなわち、タスク開始情報の総数が減少する傾向にある場合（S 1 9 0 1 : N）には、S 1 9 0 3 に進む。

[0153] S 1 9 0 2 で、タスク割当て管理部 4 3 1 は、図 2 1 の S 1 8 0 1 で選択されたDBバッファ面に格納されたページと関連づけられるコンテキストのタスク生成可能数が所定の閾値以上であるかなかを判定し、その結果、タスク生成可能数が閾値以上である場合（S 1 9 0 2 : Y）には、S 1 9 0 4 に進む一方、タスク生成可能数が閾値以上でない場合（S 1 9 0 2 : N）には、S 1 9 0 5 に進む。

[0154] S 1 9 0 3 で、タスク割当て管理部 4 3 1 は、図 2 1 の S 1 8 0 1 で選択されたDBバッファ面に格納されたページと関連づけられるコンテキストのタ

スク生成可能数が所定の閾値以上であるか否かを判定し、その結果、タスク生成可能数が閾値以上である場合（S 1 9 0 3 : Y）には、S 1 9 0 5に進む一方、タスク生成可能数が閾値以上でない場合（S 1 9 0 3 : N）には、S 1 9 0 4に進む。

[0155] S 1 9 0 4 で、タスク割当て管理部 4 3 1 は、選択された D B バッファを解放可と判定し、D B バッファ面解放判定処理を終了する。

[0156] S 1 9 0 5 で、タスク割当て管理部 4 3 1 は、選択された D B バッファを解放不可と判定し、D B バッファ面解放判定処理を終了する。

[0157] この D B バッファ面解放判定処理によると、タスク開始情報の総数が増加している傾向がある場合には、タスク生成可能数が多いコンテキストと関連づけられるページを格納する D B バッファ面を優先して解放し、タスク生成可能数の少ないコンテキストと関連づけられるページが D B バッファに残るようにすることで、タスク開始情報を効率よく消費することができる。また、タスク開始情報の総数が減少している傾向がある場合には、タスク生成可能数が少ないコンテキストと関連づけられるページを格納する D B バッファ面を優先して解放して、タスク生成可能数が多いコンテキストと関連づけられるページが D B バッファに優先して残るようにすることで、多くのタスクを記憶デバイス群 4 4 3 へのアクセスを行うことなく効率よく実行開始することができる。

実施例 2

[0158] 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。

[0159] 図 2 3 は、実施例 2 に係る計算機システムの一部の詳細構成図を示す。

[0160] 実施例 2 においては、実施例 1 のタスク割当て管理部 4 3 1 の一部の処理と、実施例 1 の D B バッファ管理部 4 2 7 の一部の処理とを、新たに備えたタスク開始情報格納場所管理部 4 3 8 に実行させるようにしたものであり、それに応じて、タスク割当て管理部 4 3 1 及び D B バッファ管理部 4 2 7 の処理が変更されている。

[0161] タスク開始情報格納場所管理部438は、タスク開始情報格納場所管理表439を管理し、タスク割当て管理部431、DBバッファ管理部427からの問い合わせに答える。タスク開始情報格納場所管理部438は、タスク割当て管理部431又はDBバッファ管理部427からの通知に基づいて、タスク開始情報格納場所管理表439のエントリの登録、削除、更新等を行う。タスク開始情報格納場所管理部438は、タスク割当て管理部431からデータアドレス集合を記憶するページがDBバッファにあるか否かの問い合わせがあった場合には、タスク開始情報格納場所管理表439を参照し、対応するデータアドレス集合の格納場所が主記憶（すなわち、DBバッファ）であるか否かを判定し、判定結果をタスク割当て管理部431に通知する。また、タスク開始情報格納場所管理部438は、DBバッファ管理部427からDBバッファを解放しても良いか否かの問い合わせがあった場合には、タスク開始情報格納場所管理表439を参照し、対応するDBバッファのデータを主記憶から追い出しても良いか否かを判定し、判定結果をDBバッファ管理部427に通知する。

[0162] 図24は、実施例2に係るタスク開始情報格納場所管理表の構成を示す。

[0163] タスク開始情報格納場所管理表439は、タスク開始情報毎の情報として、タスク開始情報ID439a、データアドレス集合格納場所439b、主記憶追出可否439cを有する。各種情報は、下記の通りである。

（＊）タスク開始情報ID439aは、タスク開始情報を識別するためのID（タスク開始情報ID）を示す。

（＊）データアドレス集合格納場所439bは、対応するタスク開始情報のデータアドレス集合が格納されている格納場所の情報である。データアドレス集合格納場所439bは、例えば、データアドレス集合がDBバッファ面437に格納されている場合には、「主記憶」となり、DBバッファではなく、外部ストレージ装置402の記憶デバイス群443に格納されている場合には、「二次記憶」となる。

（＊）主記憶追出可否439cは、対応するデータアドレス集合を主記憶装

置、すなわち、DBバッファ面437から解放することが可能であるか否かを示す。

[0164] タスク割当て管理部431は、タスク開始情報を生成した場合、タスク開始情報を削除した場合、タスク開始情報のデータアドレス集合を記憶するページのDBバッファからの解放可否を変更した場合に、その内容をタスク開始情報格納場所管理部438に通知する。本例においては、例えば、タスク割当て管理部431は、所定のタイミングで、各タスク開始情報を対象に、図22に示すDBバッファ解放判定処理を実行し、各タスク開始情報のデータアドレス集合を記憶するページを格納するDBバッファ面の解放可否を示す判定結果をタスク開始情報格納場所管理部438に通知する。

[0165] タスク割当て管理部431は、コンテキストと関連づけられているデータアドレス集合を記憶するページがDBバッファ436にあるか否かを、タスク開始情報格納場所管理部438に問い合わせる。

[0166] DBバッファ管理部427は、タスク開始情報のデータアドレス集合を記憶するページを格納するDBバッファ面を確保した場合、及びそのDBバッファ面を解放した場合に、その内容をタスク開始情報格納場所管理部438に通知する。

[0167] DBバッファ管理部427は、DBバッファ面を開放して良いか否かをタスク開始情報格納場所管理部438に問い合わせる。

[0168] 上記実施例2によると、上記実施例1と同様な効果を得ることができる。

実施例 3

[0169] 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或いは簡略する。

[0170] 図25は、実施例3に係る計算機システムの構成を示す。

[0171] DBサーバ401に、アプリケーションサーバ（以下、APサーバ）3102が通信ネットワーク3112を介して通信できるように接続されており、また、外部ストレージ装置402が、通信ネットワーク403を介して通信できるように接続されている。APサーバ3102に、通信ネットアーク

3111を介してユーザ端末（クライアント端末）3101が通信できるように接続されている。DBサーバ401は、DBを管理するDBMSを実行する。外部ストレージ装置402は、DBを格納する。APサーバ3102は、DBサーバ401で実行されるDBMSにクエリを発行するAPを実行する。ユーザ端末3101は、APサーバ3102で実行されるAPに要求を出す。ユーザ端末3101、又は、APサーバ3102は、複数存在しても良い。

[0172] APサーバ3102、DBサーバ401及び外部ストレージ装置402に、それぞれ、通信ネットワーク3114、3115及び3116をそれぞれ介してAPサーバ管理端末3103、DBサーバ管理端末3104及びストレージ管理端末3105が接続されている。APサーバ管理端末3103は、APサーバ3102を管理する端末であり、DBサーバ管理端末3104は、DBサーバ401を管理する端末であり、ストレージ管理端末3105は、外部ストレージ装置402を管理する端末である。管理端末3103～3105のうちの少なくとも2つが共通（一体）であっても良い。通信ネットワーク3111、3112、3114、3115、3116、403のうちの少なくとも2つが共通（一体）であっても良い。

[0173] 実施例3では、下記の処理の流れが実行されて良い。

（S31）APサーバ3102が、ユーザ端末3101から要求（以下、ユーザ要求）を受信する。

（S32）APサーバ3102が、S31で受信したユーザ要求に従いクエリを生成し、生成したクエリをDBサーバ401に発行する。

（S33）DBサーバ401が、APサーバ3102からのクエリを受け付け、そのクエリを実行し、そのクエリの実行に伴い、データ入出力要求（例えばデータ読出し要求）を外部ストレージ装置402に発行する。DBサーバ401は、1つのクエリの実行において、複数のデータ入出力要求を並行して発行することがある。そのため、1つのクエリの実行において、S33が複数回並行して行われることがある。

(S 3 4) DBサーバ4 0 1が、クエリの実行において、外部ストレージ装置4 0 2からデータ入出力要求に対する応答を受ける。DBサーバ4 0 1は、1つのクエリの実行において、複数のデータ入出力要求を並行して発行することがある。そのため、1つのクエリの実行において、S 3 4が複数回並行して行われることがある。

(S 3 5) DBサーバ4 0 1が、クエリの実行結果をAPサーバ3 1 0 2に送信する。

(S 3 6) APサーバ3 1 0 2が、クエリの実行結果を受信し、その実行結果に従う、S 3 1で受信したユーザ要求に対する回答を、ユーザ端末3 1 0 1に送信する。

[0174] なお、APサーバ3 1 0 2へのユーザ要求、又は、DBサーバへ発行されるクエリは同時に複数あっても良い。

[0175] 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。

[0176] 例えば、上記実施例では、クエリ実行部4 2 3は、DBオペレーションの条件に合致するデータのアドレスが1つの場合、タスクを新たに生成することなく同一のタスクを実行することでそのデータをDB 4 5 1から読み出すが、DBオペレーションの条件に合致するデータのアドレスが1つの場合であっても、そのタスク（親タスク）を元に新たにタスク（子タスク）を生成し、子タスクを実行することで（子タスクにデータアドレスを割り当てることで）、そのデータをDB 4 5 1から読み出しても良い。

[0177] また、DB 4 5 1は、DBサーバ4 0 1におけるメモリ4 1 6以外の、メモリ4 1 6よりもアクセス速度が遅い記憶デバイス、例えばローカル記憶デバイス4 1 5に存在しても良い。

符号の説明

[0178] 4 1 2…データベース管理システム（DBMS）

請求の範囲

[請求項1]

第1の記憶デバイスよりアクセス速度の遅い第2の記憶デバイスに存在する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、

前記データベースへのクエリを受け付けるクエリ受付部と、

前記受け付けたクエリを実行するために必要な1以上のDBオペレーションと前記1以上のDBオペレーションの実行手順とを表す情報を含んだクエリ実行プランを生成するクエリ実行プラン生成部と、

前記生成したクエリ実行プランに基づいて前記クエリを実行し、前記クエリの実行において、前記1以上のDBオペレーションを実行するためのタスクを動的に生成し、実行可能な2以上のタスクが存在する場合には、前記2以上のタスクのうちの少なくとも2つを並行して実行するクエリ実行部と、

前記タスクの実行を管理する実行タスク管理部と、

前記第2の記憶デバイスに存在する前記データベースのデータを、前記第1の記憶デバイスに一時的に読出して管理するDBバッファ管理部と

を有し、

前記クエリ実行部はコンテキスト管理部とタスク割当て管理部とを有し、前記タスクにおいて実行を行う内容を示す情報であるタスク開始情報を生成し、前記タスクにより前記タスク開始情報が示す内容を実行し、

前記タスク開始情報は、コンテキストと前記コンテキストに関連づけられたデータアドレス集合とを含み、

前記コンテキストは、前記新たに生成するタスクにおいて実行を開始するDBオペレーションが、前記クエリ実行プランに含まれる情報である1以上のDBオペレーションのうちのいずれであるかを示す第1の情報と、前記第1の情報が示すDBオペレーションに要するデー

タのアクセス先に関する第2の情報と、前記新たに生成するタスクにより前記1以上のDBオペレーションに関する結果を生成するために必要なデータに関する第3の情報とを含む情報であり、

前記データアドレス集合は、前記第2の記憶デバイスに存在し、前記データのアクセス先に関する第2の情報に基づいて、アクセス先データのアドレスを特定するための情報であり、

前記コンテキスト管理部は、前記生成されたコンテキストを管理し、

前記タスク割当て管理部は、

(A) 前記コンテキストと関連づけられたデータアドレス集合が前記第1の記憶デバイスに存在するか否かに基づいてタスク開始情報を選択し、

前記DBバッファ管理部は、

(B) タスク開始情報の管理状態に基づいて、前記第2の記憶デバイスから前記第1の記憶デバイスへデータアドレス集合を読み出し、又は、前記第1の記憶デバイスに読み出されたデータアドレス集合を削除する、

ことを特徴とする、データベース管理システム。

[請求項2]

前記タスク割当て管理部は、

前記(A)で、前記第1の記憶デバイスに格納されているデータアドレス集合を含んだタスク開始情報が存在するか否かを判定し、その判定の結果が肯定的の場合に、前記第1の記憶デバイスに存在する前記データアドレス集合を含むタスク開始情報を選択し、前記判定が否定的の場合に、前記第2の記憶デバイスに存在するデータアドレス集合を選択し、前記選択されたデータアドレス集合を前記第1の記憶デバイスへ読み出し、前記読み出されたデータアドレス集合を含むタスク開始情報を選択する、

請求項1に記載のデータベース管理システム。

- [請求項3] 前記タスク割当て管理部は、
- 前記（Ａ）で、前記第１の記憶デバイスに複数のデータアドレス集合が存在する場合に、前記複数のデータアドレス集合をそれぞれ含む複数のタスク開始情報のうち、アクセスされていないデータアドレス数が最も少ないデータアドレス集合を含むタスク開始情報を選択する、
- 請求項２に記載のデータベース管理システム。
- [請求項4] 前記タスク割当て管理部は、
- （Ｃ）単位時間あたりに新たに生成される前記タスク開始情報の数である生成スピードと、アクセス先データが全てアクセスされた前記タスク開始情報が単位時間あたりに生じる数である消費スピードとを測定し、
- （Ｄ）前記生成スピードと前記消費スピードとを比較し、
- 前記ＤＢバッファ管理部は、前記（Ｄ）での比較の結果を基に、前記（Ｂ）を行う、
- 請求項１に記載のデータベース管理システム。
- [請求項5] 前記ＤＢバッファ管理部は、前記（Ｄ）での比較の結果が、前記生成スピードが前記消費スピードより速い、である場合に、前記（Ｂ）で、前記第１の記憶デバイスから、アクセスされていないデータアドレス数が所定の値より多いデータアドレス集合を削除する、
- 請求項４に記載のデータベース管理システム。
- [請求項6] 前記ＤＢバッファ管理部は、前記（Ｄ）での比較の結果が、前記生成スピードが前記消費スピードより遅い、である場合に、前記（Ｂ）で、前記第１の記憶デバイスから、アクセスされていないデータアドレス数が所定の値より少ないデータアドレス集合を削除する、
- 請求項４に記載のデータベース管理システム。
- [請求項7] 前記ＤＢバッファ管理部は、前記（Ｄ）での比較の結果が、前記生成スピードが前記消費スピードより速い、である場合に、前記（Ｂ）

で、前記 1 以上のタスク開始情報にそれぞれ対応する 1 以上のデータアドレス集合であって前記第 1 の記憶デバイスに存在していないデータアドレス集合のうち、アクセスされていないデータアドレス数が所定の値より少ないデータアドレス集合を、前記第 2 の記憶デバイスから前記第 1 の記憶デバイスに読み出す、
請求項 4 に記載のデータベース管理システム。

[請求項 8] 前記 DB バッファ管理部は、前記 (D) での比較の結果が、前記生成スピードが前記消費スピードより遅い、である場合に、前記 (B) で、前記 1 以上のタスク開始情報にそれぞれ対応する 1 以上のデータアドレス集合であって前記第 1 の記憶デバイスに存在していないデータアドレス集合のうち、アクセスされていないデータアドレス数が所定の値より多いデータアドレス集合を、前記第 2 の記憶デバイスから前記第 1 の記憶デバイスに読み出す、
請求項 4 に記載のデータベース管理システム。

[請求項 9] 前記 DB バッファ管理部は、
前記第 1 の記憶デバイスに、前記データアドレス集合を格納可能な空き容量が所定量より多い場合に、前記第 2 の記憶デバイス中の 1 以上の前記データアドレス集合を、前記第 2 の記憶デバイスから前記第 1 の記憶デバイスに読み出す、
請求項 1 に記載のデータベース管理システム。

[請求項 10] データベースを管理する計算機であって、
第 1 の記憶デバイスと、
前記第 1 の記憶デバイスに接続された制御デバイスと
を有し、
前記データベースが、第 1 の記憶デバイスよりアクセス速度の遅い第 2 の記憶デバイスに存在しており、
前記制御デバイスが、
(P) 前記データベースへのクエリを受け付け、

(Q) 前記受け付けたクエリを実行するために必要な1以上のDBオペレーションと前記1以上のDBオペレーションの実行手順とを表す情報を含んだクエリ実行プランを生成し、

(R) 前記生成したクエリ実行プランに基づいて前記クエリを実行し、前記クエリの実行において、前記1以上のDBオペレーションを実行するためのタスクを動的に生成し、実行可能な2以上のタスクが存在する場合には、前記2以上のタスクのうちの少なくとも2つを並行して実行し、

(S) 前記第2の記憶デバイスに存在する前記データベースのデータを、前記第1の記憶デバイスに一時的に読出して管理し、

前記制御デバイスが、前記(R)において、前記タスクにおいて実行を行う内容を示す情報であるタスク開始情報を生成し、前記タスクにより前記タスク開始情報が示す内容を実行し、

前記タスク開始情報は、コンテキストと前記コンテキストに関連づけられたデータアドレス集合とを含み、

前記コンテキストは、前記新たに生成するタスクにおいて実行を開始するDBオペレーションが、前記クエリ実行プランに含まれる情報である1以上のDBオペレーションのうちのいずれであるかを示す第1の情報と、前記第1の情報が示すDBオペレーションに要するデータのアクセス先に関する第2の情報と、前記新たに生成するタスクにより前記1以上のDBオペレーションに関する結果を生成するために必要なデータに関する第3の情報とを含む情報であり、

前記データアドレス集合は、前記第2の記憶デバイスに存在し、前記データのアクセス先に関する第2の情報に基づいて、アクセス先データのアドレスを特定するための情報であり、

前記制御デバイスが、

(A) 前記コンテキストと関連づけられたデータアドレス集合が前記第1の記憶デバイスに存在するか否かに基づいてタスク開始情報を選

択し、

(B) タスク開始情報の管理状態に基づいて、前記第2の記憶デバイスから前記第1の記憶デバイスへデータアドレス集合を読み出し、又は、前記第1の記憶デバイスに読み出されたデータアドレス集合を削除する、

ことを特徴とする、計算機。

[請求項11]

前記制御デバイスは、

前記(A)で、前記第1の記憶デバイスに格納されているデータアドレス集合を含んだタスク開始情報が存在するか否かを判定し、その判定の結果が肯定的の場合に、前記第1の記憶デバイスに存在する前記データアドレス集合を含むタスク開始情報を選択する、

請求項10に記載の計算機。

[請求項12]

前記制御デバイスは、

(C) 単位時間あたりに新たに生成される前記タスク開始情報の数である生成スピードと、アクセス先データが全てアクセスされた前記タスク開始情報が単位時間あたりに生じる数である消費スピードとを測定し、

(D) 前記生成スピードと前記消費スピードとを比較し、

前記制御デバイスは、前記(D)での比較の結果を基に、前記(B)を行う、請求項10に記載の計算機。

[請求項13]

第1の記憶デバイスよりアクセス速度の遅い第2の記憶デバイスに存在するデータベースを管理するデータベース管理方法であって、

(P) 前記データベースへのクエリを受け付け、

(Q) 前記受け付けたクエリを実行するために必要な1以上のDBオペレーションと前記1以上のDBオペレーションの実行手順とを表す情報を含んだクエリ実行プランを生成し、

(R) 前記生成したクエリ実行プランに基づいて前記クエリを実行し、前記クエリの実行において、前記1以上のDBオペレーションを実

行するためのタスクを動的に生成し、実行可能な2以上のタスクが存在する場合には、前記2以上のタスクのうちの少なくとも2つを並行して実行し、

(S) 前記第2の記憶デバイスに存在する前記データベースのデータを、前記第1の記憶デバイスに一時的に読出して管理し、

前記(R)で、前記タスクにおいて実行を行う内容を示す情報であるタスク開始情報を生成して管理し、前記タスクにより、前記タスク開始情報が示す内容を実行し、

前記タスク開始情報は、コンテキストと前記コンテキストに関連づけられたデータアドレス集合とを含み、

前記コンテキストは、前記新たに生成するタスクにおいて実行を開始するDBオペレーションが、前記クエリ実行プランに含まれる情報である1以上のDBオペレーションのうちのいずれであるかを示す第1の情報と、前記第1の情報が示すDBオペレーションに要するデータのアクセス先に関する第2の情報と、前記新たに生成するタスクにより前記1以上のDBオペレーションに関する結果を生成するために必要なデータに関する第3の情報とを含む情報であり、

前記データアドレス集合は、前記第2の記憶デバイスに存在し、前記データのアクセス先に関する第2の情報に基づいて、アクセス先データのアドレスを特定するための情報であり、

(A) 前記コンテキストと関連づけられたデータアドレス集合が前記第1の記憶デバイスに存在するか否かに基づいてタスク開始情報を選択し、

(B) タスク開始情報の管理状態に基づいて、前記第2の記憶デバイスから前記第1の記憶デバイスへデータアドレス集合を読み出し、又は、前記第1の記憶デバイスに読み出されたデータアドレス集合を削除する、

ことを特徴とする、データベース管理方法。

[図1]

Fig. 1

表A

<i>A_factory_id</i>	<i>A_factory_name</i>	<i>A_factory_size</i>	<i>A_country</i>
001	factory1	10	country2
002	factory2	100	country5
....			

[図2]

Fig. 2

表B

<i>B_product_id</i>	<i>B_product_name</i>	<i>B_product_price</i>	<i>B_factory_id</i>
0001	product1	50	002
0002	product2	1000	009
0003	product3	200	003
....			

[図3]

Fig. 3

表C

<i>C_order_id</i>	<i>C_product_id</i>	<i>C_count</i>
00000001	0002	100
00000002	0010	50
00000003	0001	500
00000004	0005	10
00000005	0002	350
00000006	0009	200
00000007	0001	100
00000008	0002	500
....		

[図4]

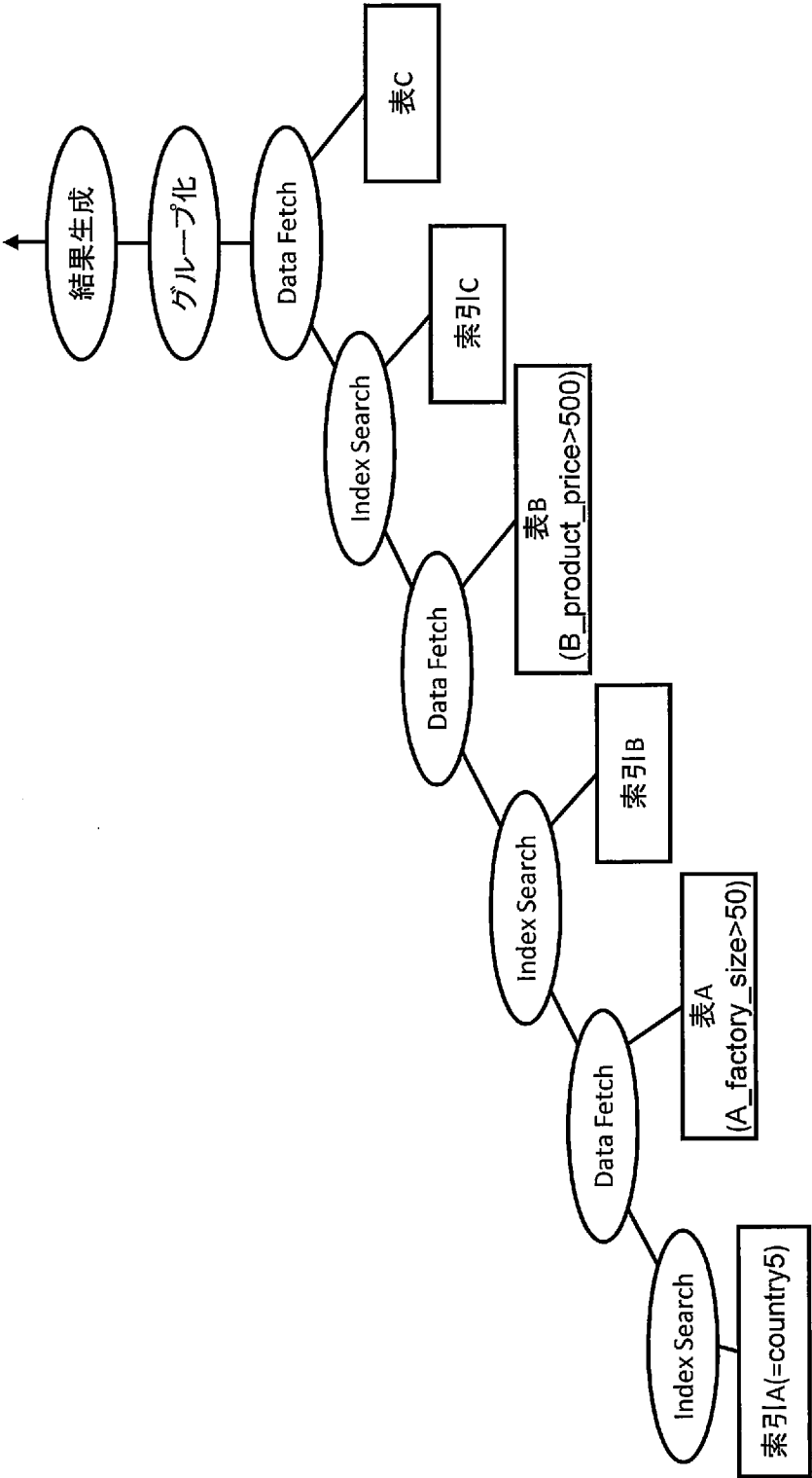
Fig. 4

クエリ

```
SELECT A_factory_name, B_product_name,  
       sum(C_count)  
FROM   A, B, C  
WHERE  A_country = country5  
       AND A_factory_size > 50  
       AND A_factory_id = B_factory_id  
       AND B_product_price > 500  
       AND B_product_id = C_product_id  
GROUP BY A_factory_name, B_product_name
```

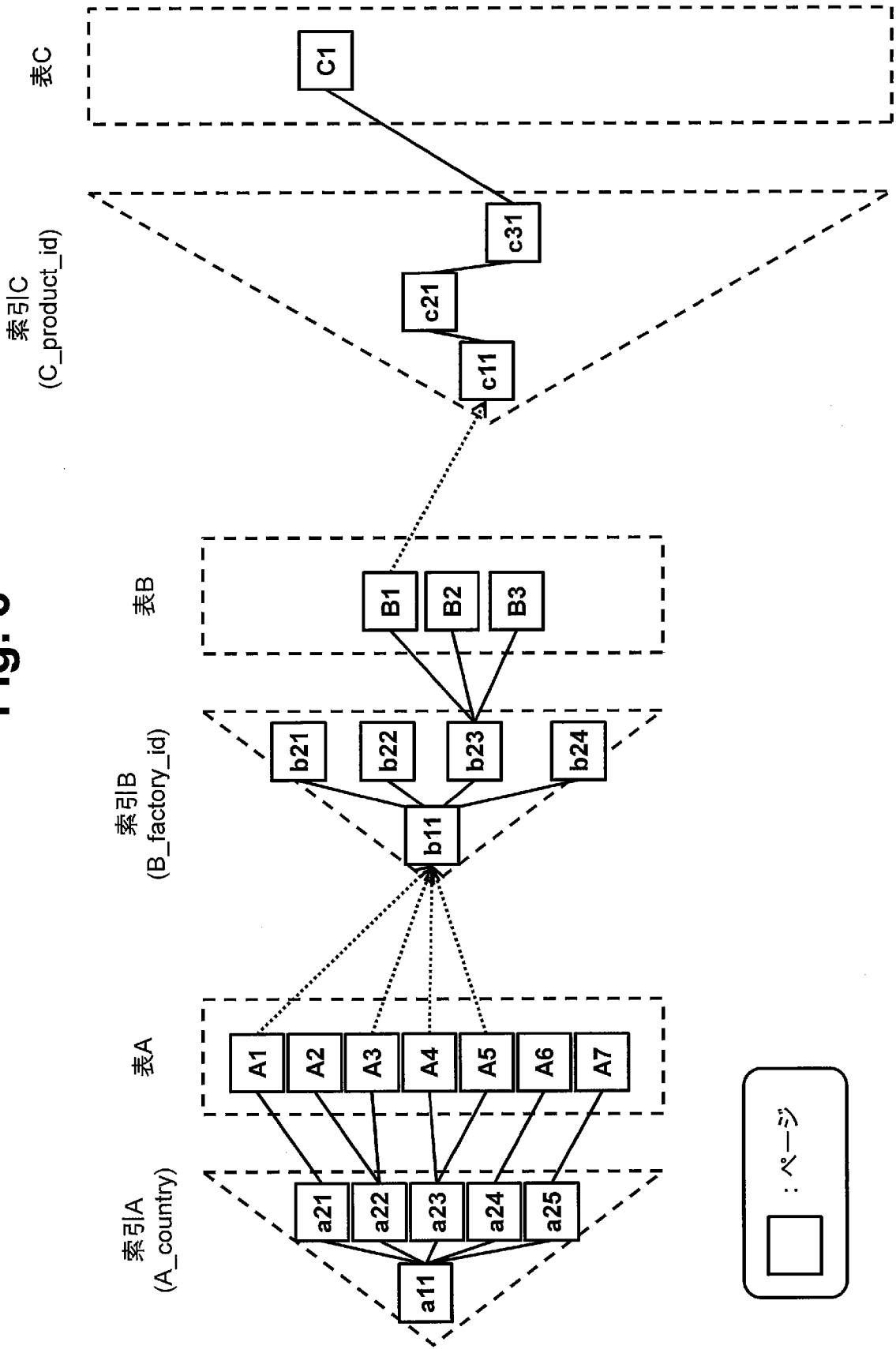
[図5]

Fig. 5



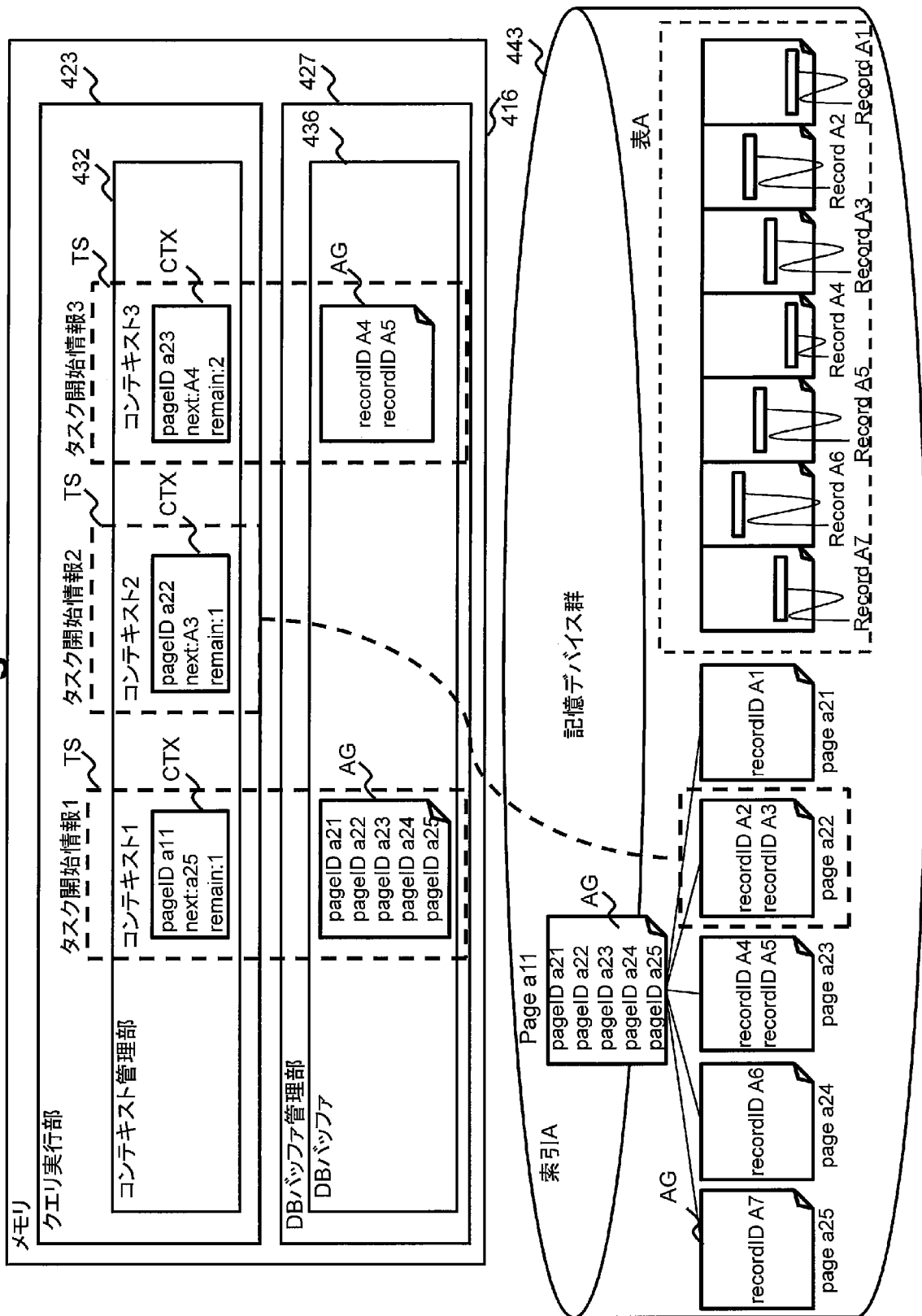
[図6]

Fig. 6



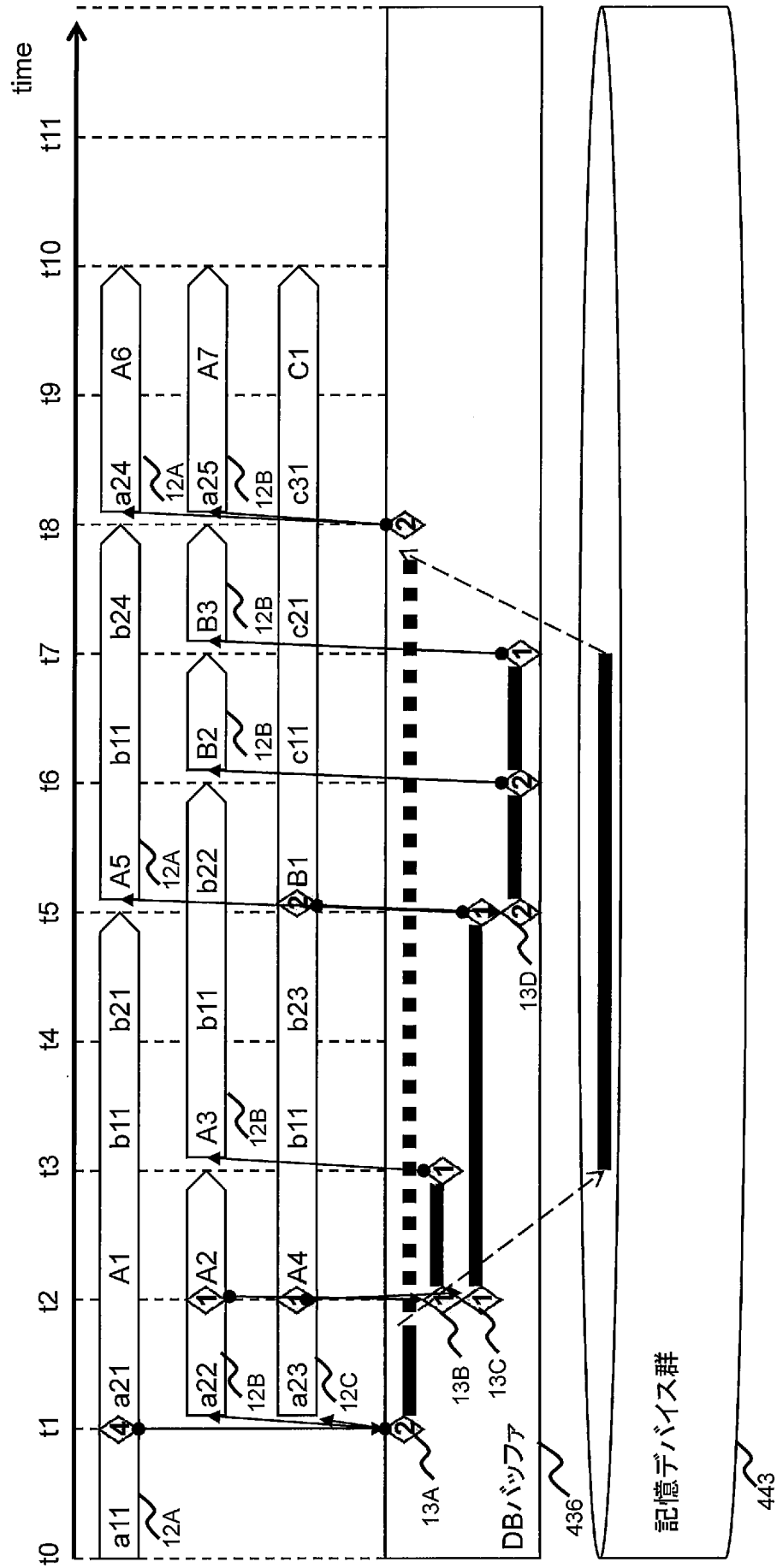
[図7]

Fig. 7



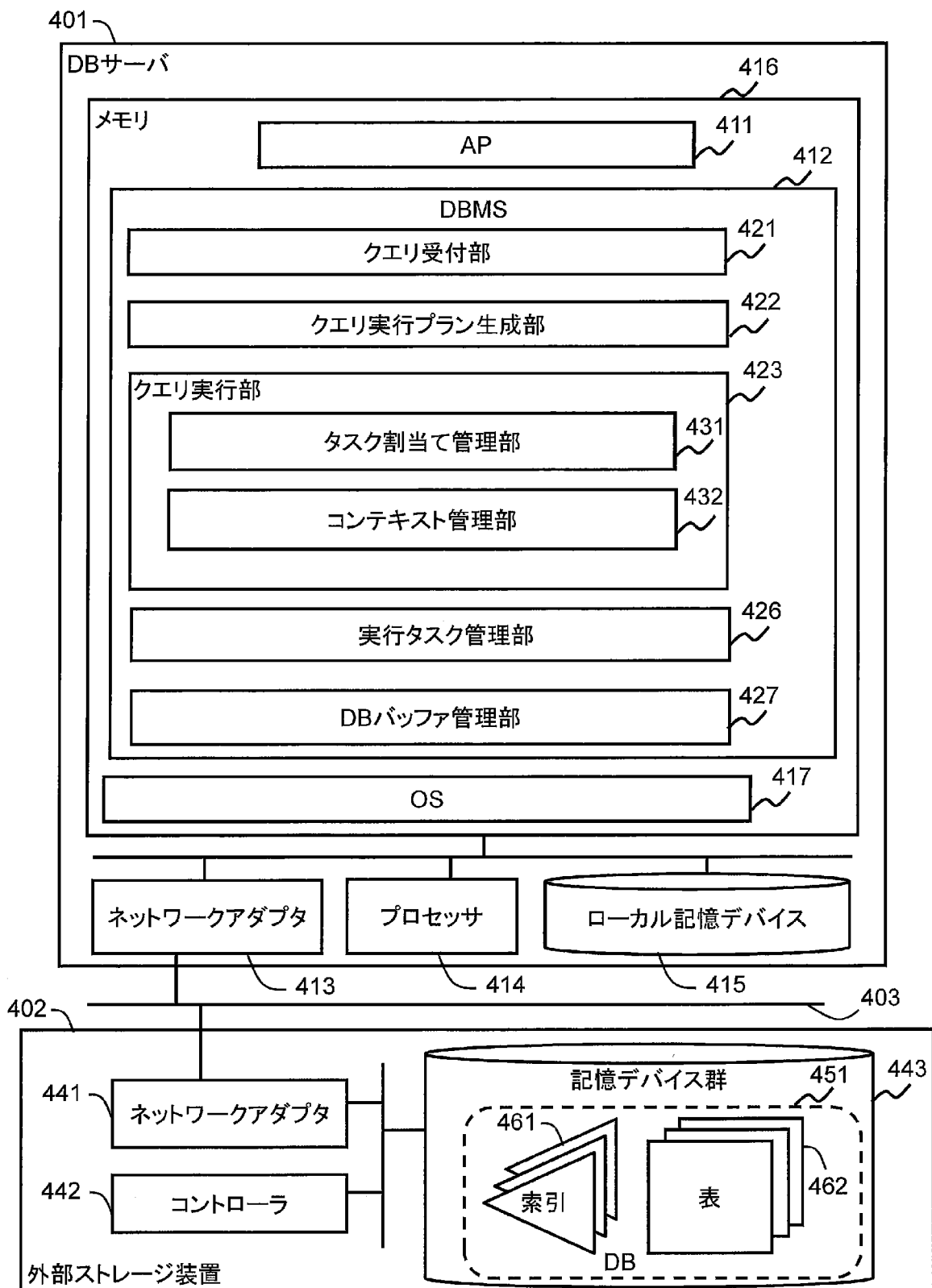
[図9]

Fig. 9

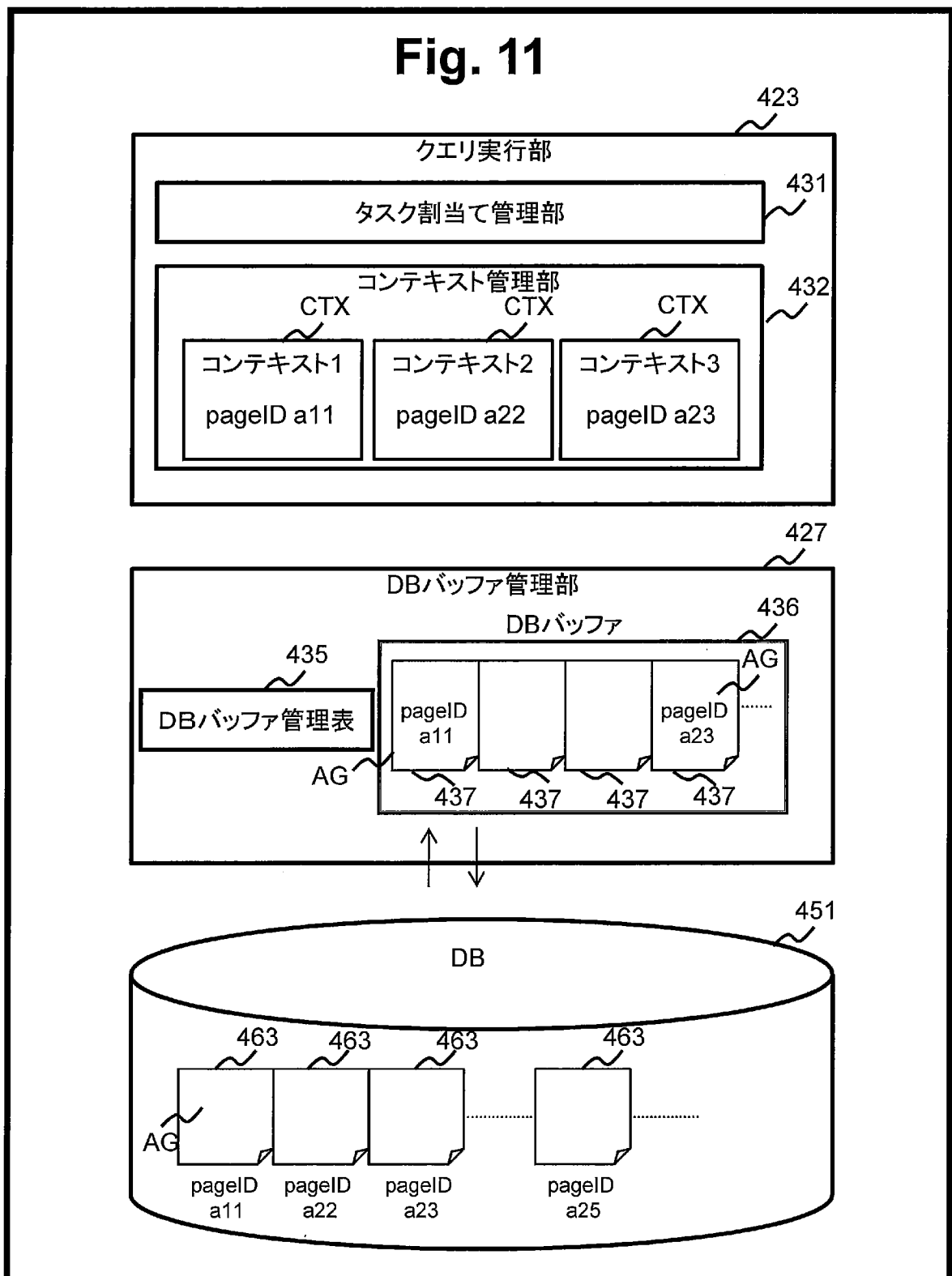


[図10]

Fig. 10



[図11]

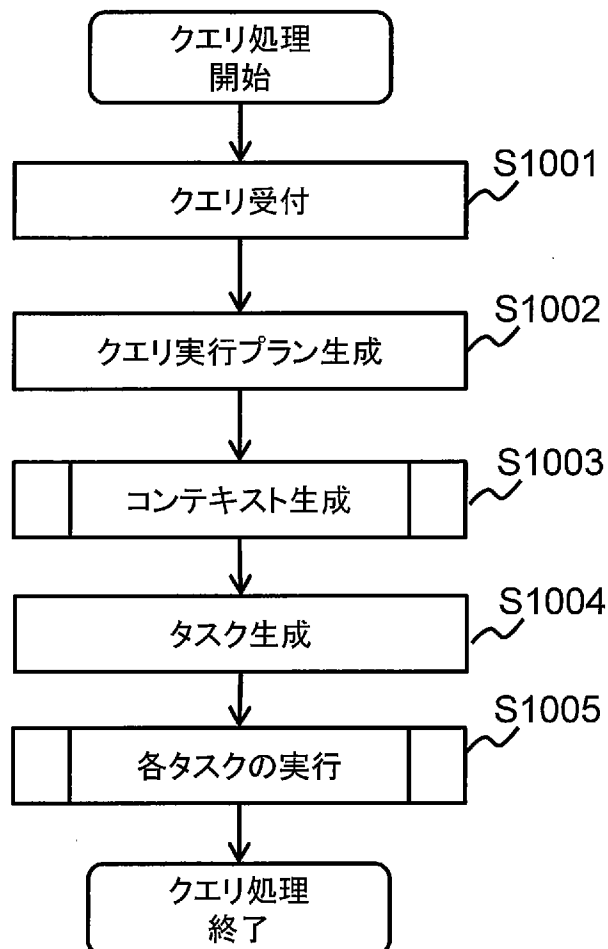


[図12]

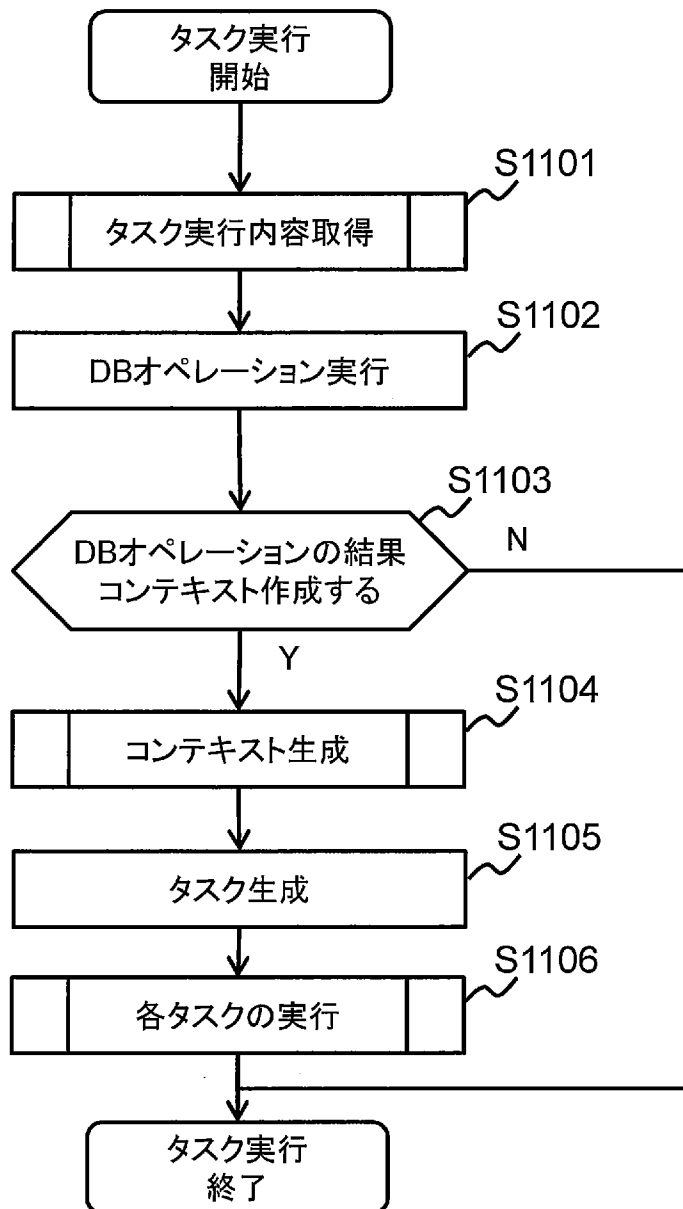
Fig. 12

バッファID	ページID
buff1	pageID a11
buff2	—
buff3	—
buff4	pageID a23
...	...

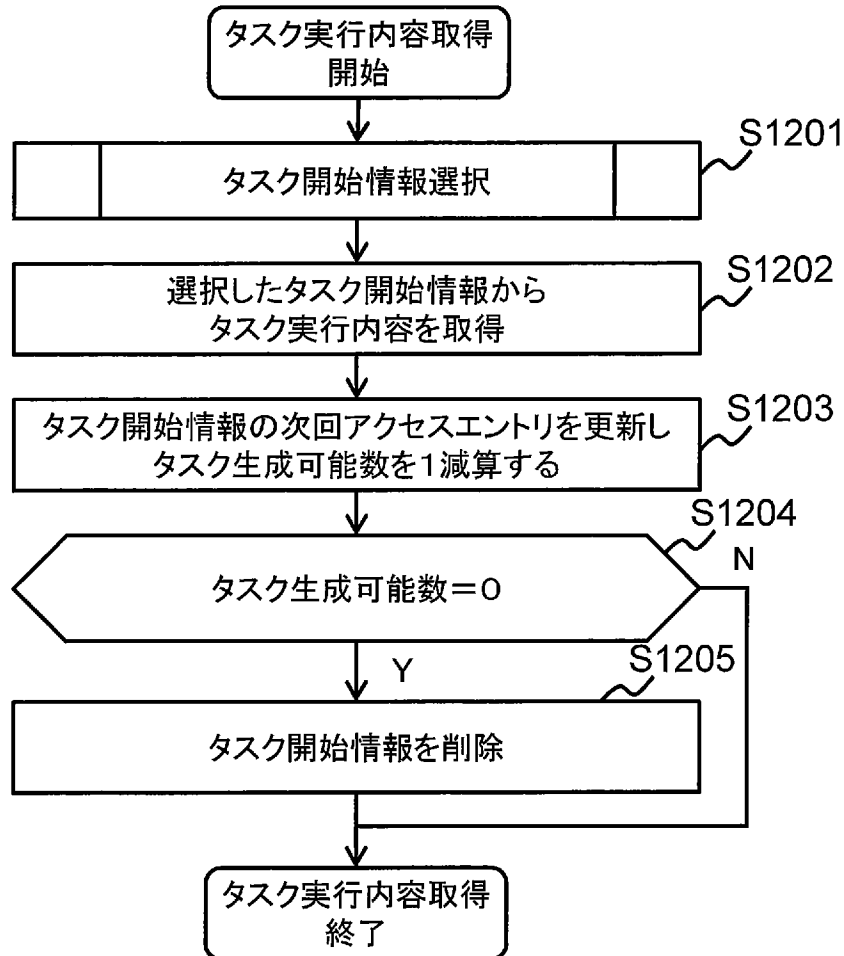
[図13]

Fig. 13

[図14]

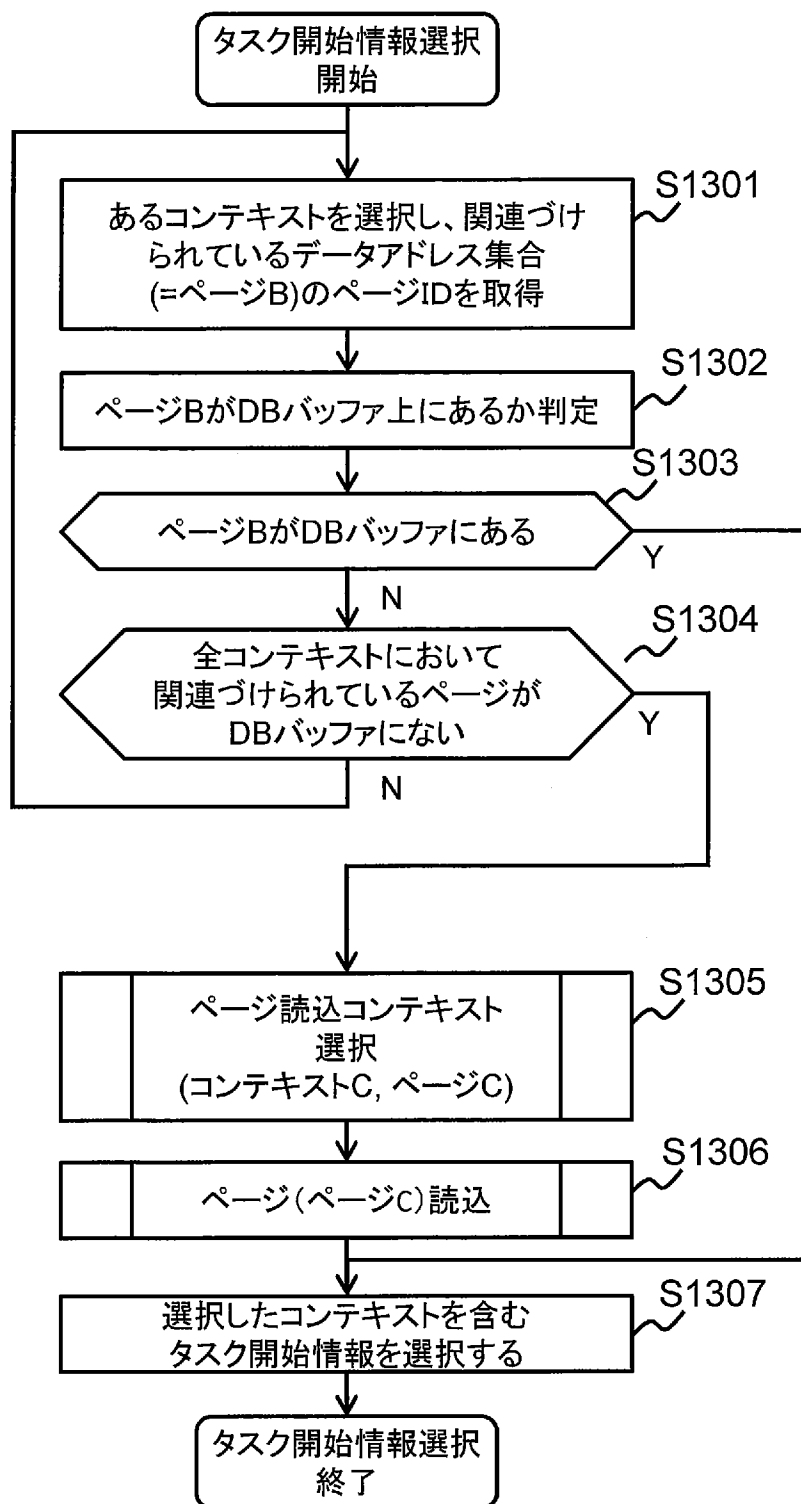
Fig. 14

[図15]

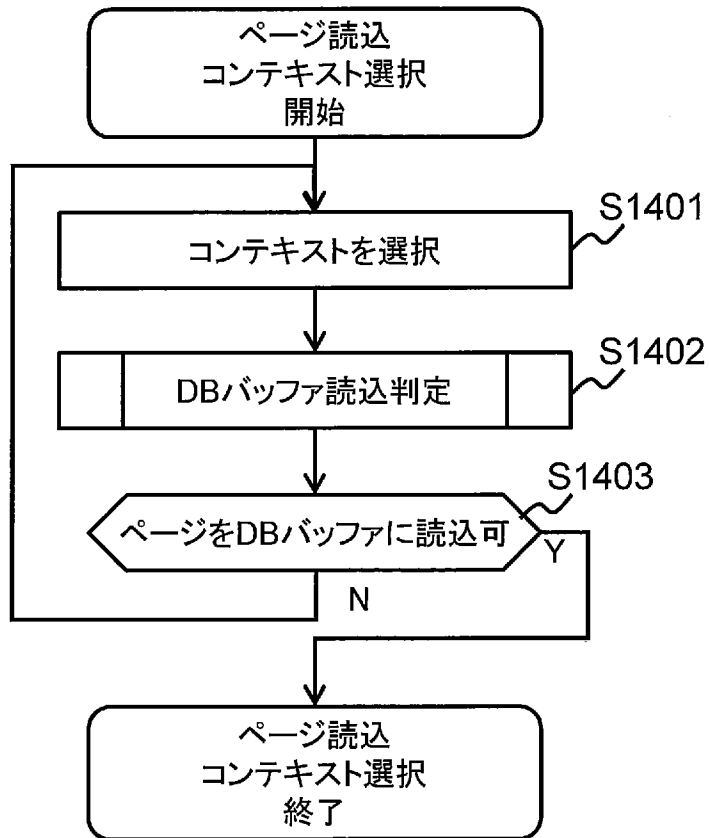
Fig. 15

[図16]

Fig. 16

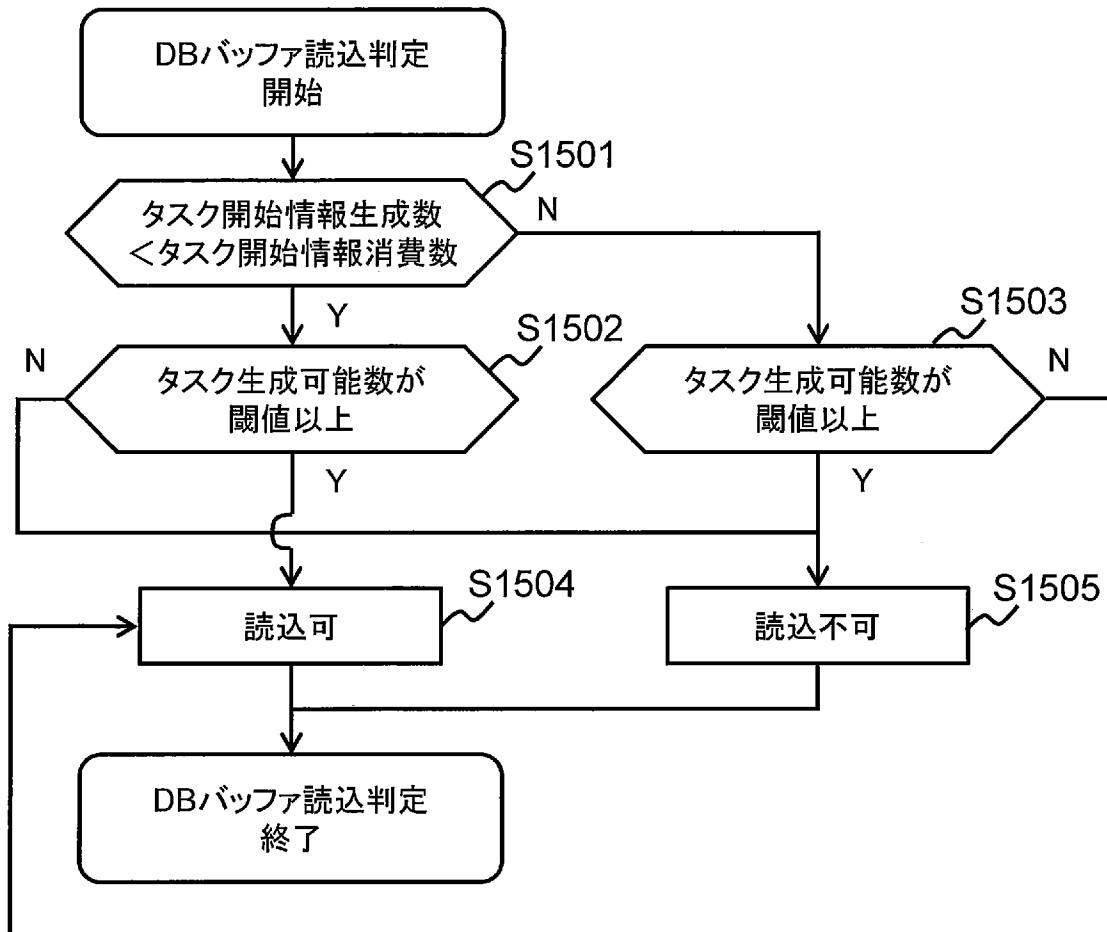


[図17]

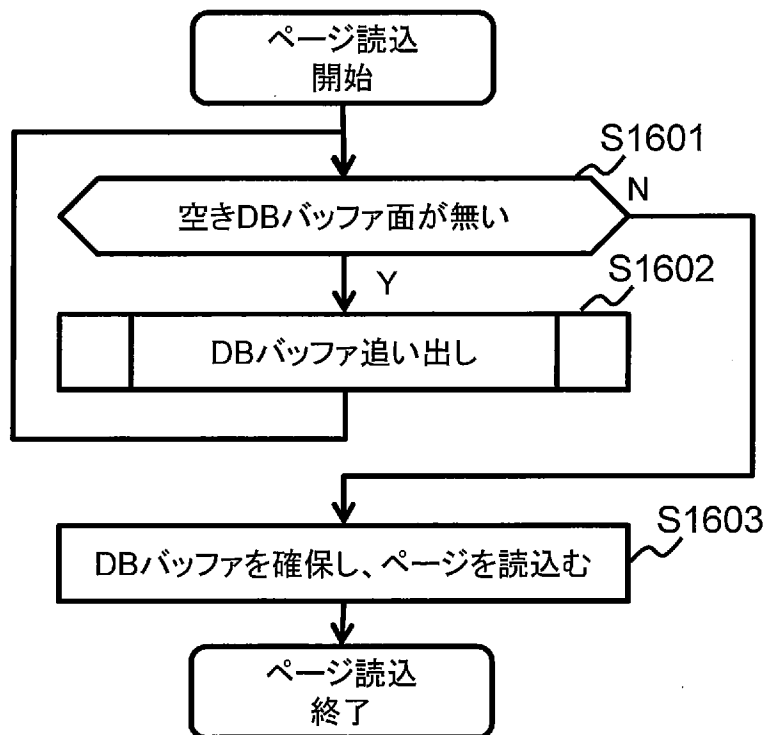
Fig. 17

[図18]

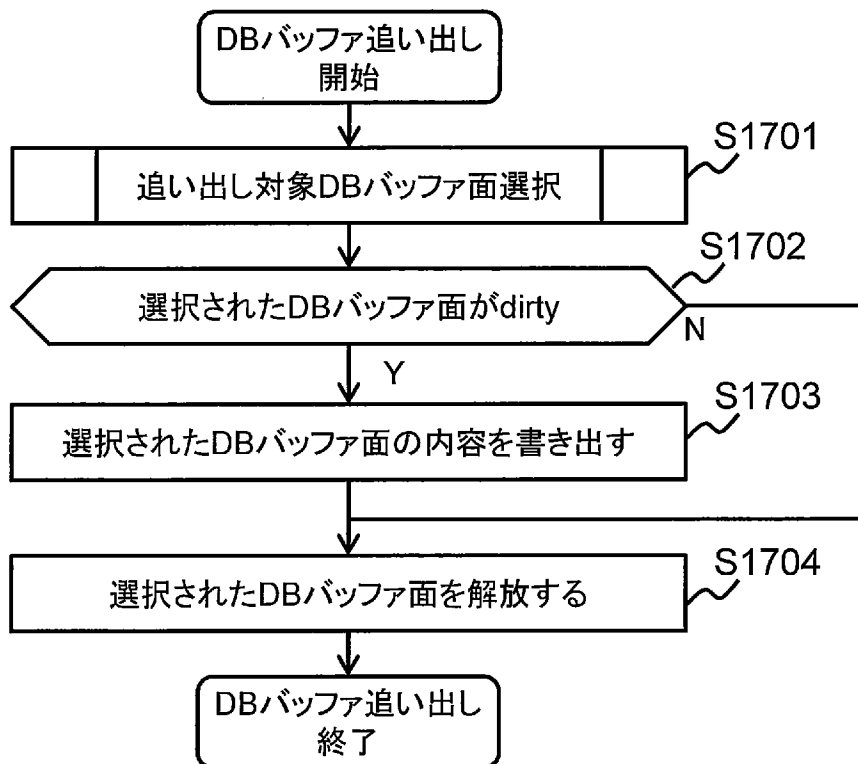
Fig. 18



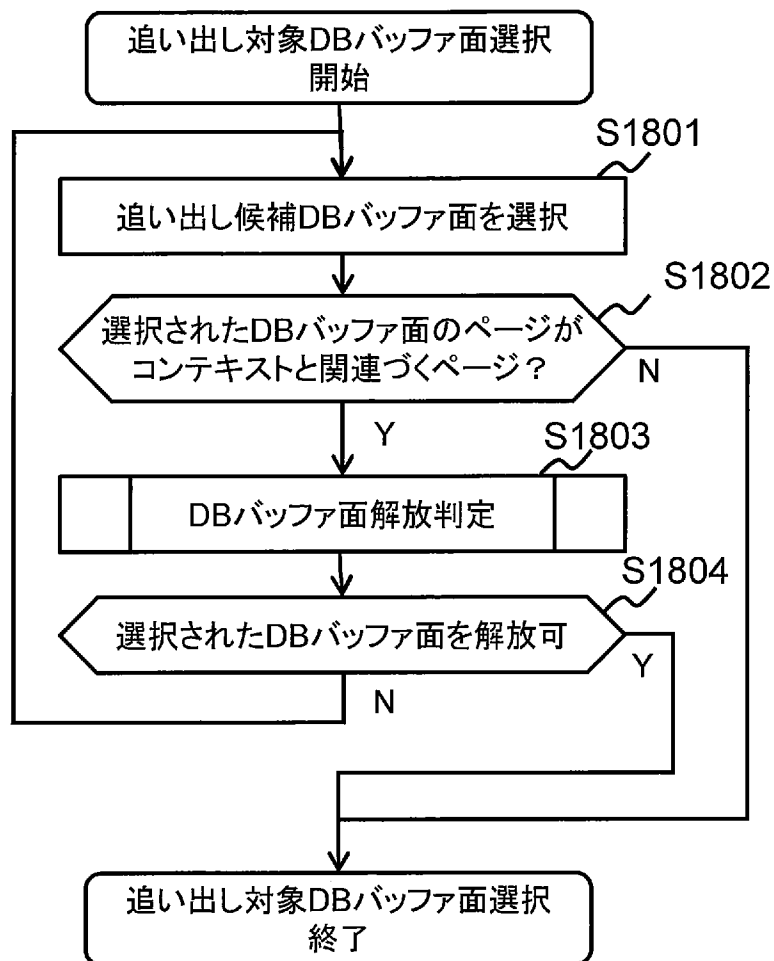
[図19]

Fig. 19

[図20]

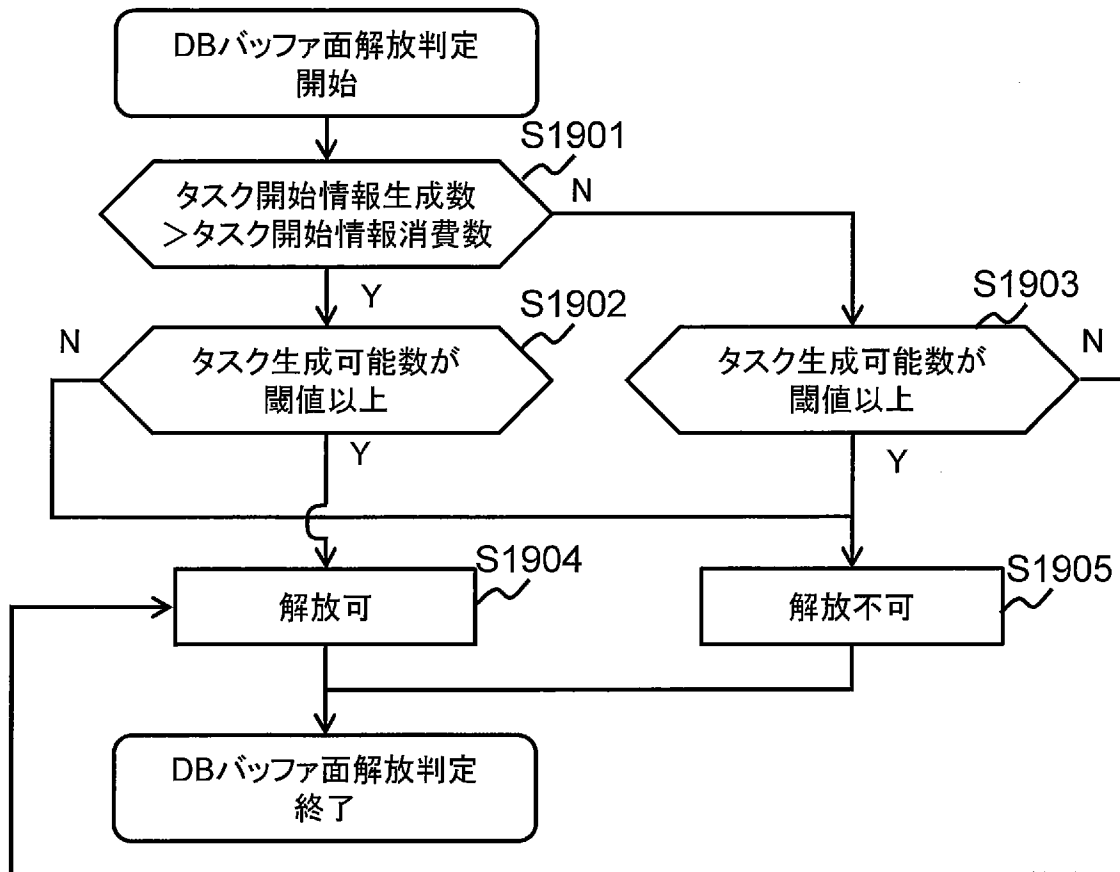
Fig. 20

[図21]

Fig. 21

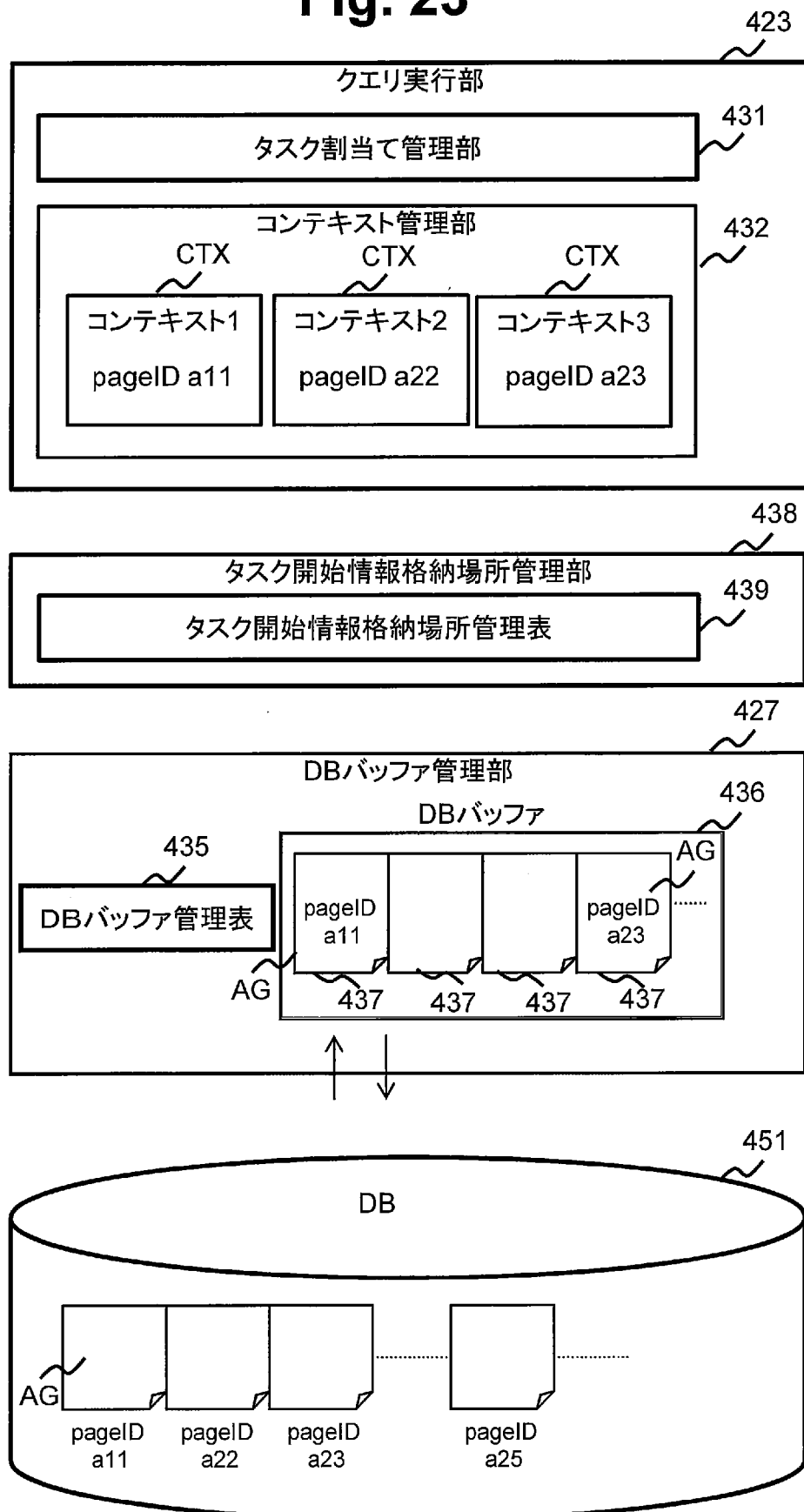
[図22]

Fig. 22



[図23]

Fig. 23



[図24]

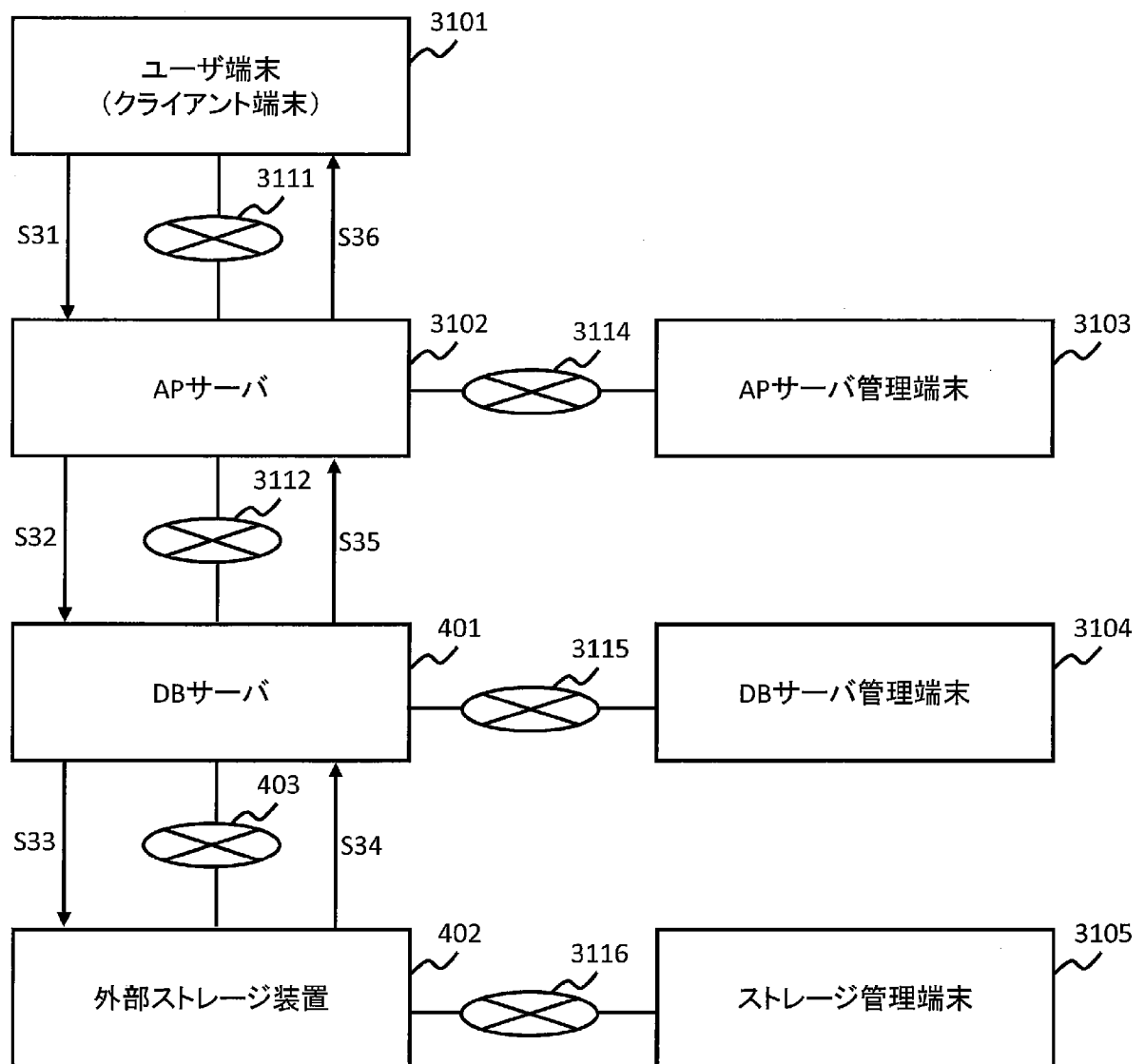
Fig. 24

439
↙

439a ↘	439b ↘	439c ↘
タスク開始情報ID	データアドレス集合 格納場所	主記憶追出可否
taskseed1	主記憶	可
taskseed2	二次記憶	—
taskseed3	主記憶	不可
...	...	

[図25]

Fig. 25



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/063332

A. CLASSIFICATION OF SUBJECT MATTER

G06F12/00 (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, 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	1996-2012
Kokai Jitsuyo Shinan Koho	1971-2012	Toroku Jitsuyo Shinan Koho	1994-2012

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JSTPlus (JDreamII)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2007-34414 A (Masaru KITSUREGAWA), 08 February 2007 (08.02.2007), entire text; all drawings & US 2007/0022100 A1	1-13
A	WO 2012/026140 A1 (Hitachi, Ltd.), 01 March 2012 (01.03.2012), entire text; all drawings (Family: none)	1-13
A	JP 2008-234495 A (Toru FURUYA), 02 October 2008 (02.10.2008), entire text; all drawings & JP 4071816 B	1-13



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

24 July, 2012 (24.07.12)

Date of mailing of the international search report

07 August, 2012 (07.08.12)

Name and mailing address of the ISA/

Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/063332

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 2002-523839 A (Computer Associates Think Inc.), 30 July 2002 (30.07.2002), entire text; all drawings & US 6108653 A & EP 1133741 A & WO 2000/013111 A1	1-13
A	Masaru KITSUREGAWA et al., "Vision and preliminary experiments for out-of-order database engine (OoODE)", Database Society of Japan Ronbunshi, 26 June 2009 (26.06.2009), vol.8, no.1, pages 131 to 136	1-13

A. 発明の属する分野の分類（国際特許分類（I P C）） Int.Cl. G06F12/00(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（I P C）） Int.Cl. G06F12/00, 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 1 9 2 2 - 1 9 9 6 年 日本国公開実用新案公報 1 9 7 1 - 2 0 1 2 年 日本国実用新案登録公報 1 9 9 6 - 2 0 1 2 年 日本国登録実用新案公報 1 9 9 4 - 2 0 1 2 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(JDreamII)		
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
A	JP 2007-34414 A（喜連川優）2007.02.08, 全文・全図 & US 2007/0022100 A1	1 - 1 3
A	WO 2012/026140 A1（株式会社日立製作所）2012.03.01, 全文・全図 （ファミリーなし）	1 - 1 3
A	JP 2008-234495 A（降矢透）2008.10.02, 全文・全図 & JP 4071816 B	1 - 1 3
☒ 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 「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 「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 「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） 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願日の後に公表された文献 「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 「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 「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 「&」 同一パテントファミリー文献		
国際調査を完了した日 2 4 . 0 7 . 2 0 1 2	国際調査報告の発送日 0 7 . 0 8 . 2 0 1 2	
国際調査機関の名称及びあて先 日本国特許庁（I S A / J P） 郵便番号 1 0 0 - 8 9 1 5 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 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 6 5	5 U 9 4 7 5

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	JP 2002-523839 A (コンピュータ アソシエイツ シンク, インコーポレイテッド) 2002.07.30, 全文・全図 & US 6108653 A & EP 1133741 A & WO 2000/013111 A1	1-13
A	喜連川優ほか1名, アウトオブオーダー型データベースエンジン OODEの構想と初期実験, 日本データベース学会論文誌, 2009.06.26, 第8巻 第1号, pp.131~136	1-13