(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
2015 年 10 月 1 日(01.10.2015)

WIPO | PCT

(10) 国際公開番号
WO 2015/145586 A1

(51) 国際特許分類:
G06F 11/20 (2006.01)

(21) 国際出願番号: PCT/JP2014/058381

(22) 国際出願日: 2014 年 3 月 25 日(25.03.2014)

(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Ｍｕｒａｋｕｍｏ (MURAKUMO CORPORATION) [JP/JP]; 〒1530061 東京都目黒区中目黒一丁目１番４５号 Tokyo (JP).

(72) 発明者: 山田　浩之 (YAMADA, Hiroyuki); 〒1530061 東京都目黒区中目黒一丁目１番４５号 株式会社Ｍｕｒａｋｕｍｏ内 Tokyo (JP).

(74) 代理人: 畑添　隆人 (HATAZOE, Takahito); 〒1020072 東京都千代田区飯田橋二丁目１番４号 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).

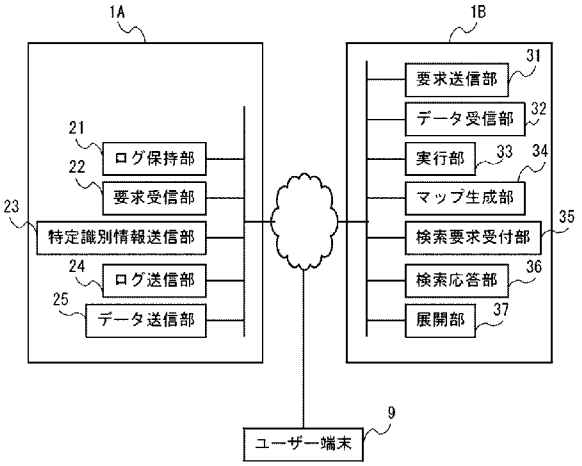
添付公開書類:

— 国際調査報告（条約第 21 条(3)）

WO 2015/145586 A1

(54) **Title**: DATABASE SYSTEM, INFORMATION PROCESSING DEVICE, METHOD, AND PROGRAM

(54) 発明の名称 : データベースシステム、情報処理装置、方法およびプログラム



9 User terminal
21 Log maintaining unit
22 Request receiving unit
23 Specific identification information transmitting unit
24 Log transmitting unit
25 Data transmitting unit
31 Request transmitting unit
32 Data receiving unit
33 Execution unit
34 Map generating unit
35 Search request accepting unit
36 Search response unit
37 Expansion unit

(57) **Abstract**: A database system wherein: a first node maintains a transaction log for a database and also maintains identification information that enables a sequence of instructions to be comprehended; the first node transmits to a second node the identification information, specific identification information that represents the transaction log at a prescribed point in time, and the transaction log after the prescribed point in time; the first node transmits data from the database to the second node from the prescribed point in time onward; and the second node, when data received from the first node is expanded in a memory, executes for the expanded data instructions that pertain to the transaction log that is newer than the transaction log that is represented by the specific identification information.

(57) 要約: データベースシステムにおいて、第一のノードは、データベースのトランザクションログを命令の順序を把握可能な識別情報とともに保持し、所定の時点におけるトランザクションログを示す特定識別情報、所定の時点より後のトランザクションログおよび識別情報を第二のノードに送信し、データベース上のデータを所定の時点以降に第二のノードに送信し、第二のノードは、第一のノードから受信されたデータがメモリに展開された場合に、特定識別情報が示すトランザクションログよりも新しいトランザクションログに係る命令を展開されたデータに対して実行することとした。

明　細　書

発明の名称 ：

データベースシステム、情報処理装置、方法およびプログラム

技術分野

[0001] 本発明は、データベースを管理するための技術に関する。

背景技術

[0002] 従来、自己の障害に備えて他のサーバノードから送信されるトランザクションログを受信し、受信されたトランザクションログをデータベースに反映するデータベース管理システムが提案されている（特許文献１を参照）。

[0003] また、マスタが、複数のトランザクションログを並列に作成し、並列に作成した複数のトランザクションログを他のマスタ又はスレーブに送信し、他のマスタまたはスレーブが、当該複数のトランザクションログをデータベースに並列に適用する技術が提案されている（特許文献２を参照）。

先行技術文献

特許文献

[0004] 特許文献1：特表２０１２－５３２３７６号公報

特許文献2：特開２０１２－１３３４１７号公報

発明の概要

発明が解決しようとする課題

[0005] 従来、ログベースのレプリケーションが行われるデータベースシステムにおいて、複製先のノードにおいてデータベースを構築する方法が種々提案されている。例えば、（１）ノードを停止させてからデータコピーを行う方法や、（２）複製元のノードから複製先のノードへ一旦テーブル全体をコピーして、コピー後に、複製元のノードでコピー中に発生した差分を抽出して反映する方法、（３）複製元のノードにおいてチェックポイントを作成してスナップショットを保存し、このスナップショットを複製先のノードにコピーした後に、チェックポイント以降のトランザクションログを複製先のノード

において反映する方法、等が提案されている。

[0006] しかし、（１）ノードを停止させてからデータコピーを行う方法では、データコピーが終了するまでデータベースを用いることが出来ない。また、（２）コピー後に差分を抽出して反映する方法では、差分の抽出および反映のために多くのリソースが必要となる。更に、（３）チェックポイント以降のトランザクションログを複製先のノードにおいて反映する方法では、チェックポイントを作成してスナップショットを複製元ノードのストレージに保存する処理、および複製先のノードにおいてトランザクションログを反映させる処理に大きなリソースが必要とされる。

[0007] 本発明は、上記した問題に鑑み、データベースシステムにおいて、複製先のノードがサービス提供可能となるまでに必要なリソースを低減させることを課題とする。

課題を解決するための手段

[0008] 本発明は、上記した課題を解決するために、以下の手段を採用した。即ち、本発明は、複数のノードを有するデータベースシステムであって、前記複数のノードのうち、データベースの複製元である第一のノードは、該第一のノードによって管理されているデータベースのトランザクションログを、該トランザクションログに係る命令の順序を把握可能な識別情報とともに保持するログ保持手段と、前記ログ保持手段によって保持されている前記識別情報のうち、所定の時点におけるトランザクションログを示す特定識別情報を、前記複数のノードのうち、データベースの複製先である第二のノードに送信する特定識別情報送信手段と、少なくとも前記所定の時点より後の前記トランザクションログおよび前記識別情報を、互いに関連づけて前記第二のノードに送信するログ送信手段と、前記データベースによって管理されているデータを、前記所定の時点以降に、前記第二のノードに送信するデータ送信手段と、を備え、前記第二のノードは、前記第一のノードから、前記トランザクションログ、該トランザクションログの識別情報、前記特定識別情報および前記データを受信する受信手段と、受信された前記データが該第二のノ

ードのメモリに展開されてデータの検索または処理に供される状態となった場合に、受信された前記トランザクションログのうち、少なくとも前記特定識別情報が示すトランザクションログよりも新しい前記トランザクションログに係る命令を、メモリに展開された前記データに対して実行する実行手段と、を備える、データベースシステムである。

[0009] 本発明に係るデータベースシステムでは、第一のノードは、所定の時点におけるトランザクションログを示す特定識別情報、所定の時点以降のトランザクションログ、および所定の時点以降のデータを第二のノードに送信する。そして、第二のノードは、受信されたデータがメモリに展開されて検索等に供される状態となった場合（例えば、データを含むページが所謂キャッシュとなった場合）に、所定の時点以降のトランザクションログに係る命令を、メモリに展開されたデータ（キャッシュ）に対して実行する。即ち、本発明では、受信されたデータが、検索等の用に供される前にメモリに展開されることを前提として、トランザクションログに係る命令の実行（トランザクションログの反映）を、メモリへの展開まで遅延させることとしている。換言すれば、本発明によれば、第二のノードによるサービス提供を、トランザクションログの反映を待たずに開始させることが出来る。なお、前記データは、テーブル単位またはページ単位で送受信されてよい。

[0010] また、前記データ送信手段は、前記データを、前記トランザクションログに係る命令の順序が互いに依存関係にあるレコードが同一の管理単位に入るように区切られた所定の管理単位毎に送信し、前記実行手段は、前記トランザクションログに係る命令を、前記所定の管理単位毎に実行してもよい。

[0011] また、受信された前記トランザクションログのうち、少なくとも前記特定識別情報が示すトランザクションログよりも新しい前記トランザクションログに基づいて、命令の対象となるデータが収容された前記所定の管理単位と命令の内容との関係を示すマップを生成するマップ生成手段を更に備え、前記実行手段は、前記マップを参照して、前記トランザクションログに係る命令を、前記所定の管理単位毎に実行してもよい。

[0012] また、受信された前記トランザクションログのうち、少なくとも前記特定識別情報が示すトランザクションログよりも新しい前記トランザクションログに基づいて、命令の対象となるデータと命令の内容との関係を示すマップを生成するマップ生成手段を更に備え、前記実行手段は、前記マップを参照して、前記トランザクションログに係る命令を、命令の対象となるデータに対して実行してもよい。

[0013] また、前記データ送信手段は、前記第一のノードのメモリに展開されてデータの検索または処理に供されている状態のデータを、前記第二のノードに送信してもよい。

[0014] また、前記第二のノードは、前記第一のノードに対して、前記データベース中のデータを指定して送信要求を送信する要求送信手段を更に備え、前記第一のノードは、前記第二のノードから、該第一のノードによって管理されているデータの送信要求を受信する要求受信手段を更に備え、前記特定識別情報送信手段は、前記送信要求に応じて、要求されたデータに係る前記特定識別情報を前記第二のノードに送信し、前記データ送信手段は、前記送信要求に応じて、要求されたデータを前記第二のノードに送信してもよい。

[0015] また、前記第二のノードは、ユーザー端末からの検索要求を受け付ける検索要求受付手段と、前記検索要求への応答を前記ユーザー端末に対して返す検索応答手段と、を更に備え、前記要求送信手段は、前記検索要求に応じて、少なくとも前記検索要求に係るデータの送信要求を、前記第一のノードに対して送信し、前記実行手段は、前記送信要求に応じて前記第一のノードから送信され、前記受信手段によって受信されたデータをメモリに展開し、前記トランザクションログに係る命令を該データに対して実行し、前記検索応答手段は、前記検索要求の結果前記第一のノードから得られ、前記トランザクションログに係る命令が実行されたデータに基づいて、前記検索要求への応答を前記ユーザー端末に対して返してもよい。

[0016] また、前記第二のノードは、ユーザー端末からの検索要求を受け付ける検索要求受付手段と、前記検索要求への応答を前記ユーザー端末に対して返す

検索応答手段と、を更に備え、前記実行手段は、前記検索要求に応じて、前記受信手段によって受信されたデータをメモリに展開し、前記トランザクションログに係る命令を該データに対して実行し、前記検索応答手段は、前記検索要求に応じて前記トランザクションログに係る命令が実行されたデータに基づいて、前記検索要求への応答を前記ユーザー端末に対して返してもよい。

[0017] また、前記データ送信手段は、前記データベースの管理情報を、前記第二のノードに送信し、前記要求送信手段は、前記管理情報を参照して、前記第一のノードに対して、前記データベース中のデータを指定して送信要求を送信してもよい。

[0018] また、前記第二のノードは、前記第一のノードから受信したデータを直接メモリに展開して、データの検索または処理に供する展開手段を更に備え、前記実行手段は、受信したデータが前記展開手段によってメモリに展開されたことを受けて、前記トランザクションログに係る命令を前記データに対して実行してもよい。

[0019] また、前記特定識別情報送信手段は、前記特定識別情報として、ログ保持手段によって保持されている前記識別情報のうち、最新のトランザクションログを示す識別情報を送信してもよい。

[0020] また、前記実行手段は、前記トランザクションログに係る命令を、前記識別情報によって把握される命令の順序に応じて、受信された前記データに対して実行してもよい。

[0021] なお、本発明は、コンピューターシステム、情報処理装置、コンピューターによって実行される方法、またはコンピューターに実行させるプログラムとして把握することが可能である。また、本発明は、そのようなプログラムをコンピューターその他の装置、機械等が読み取り可能な記録媒体に記録したものとしても把握できる。ここで、コンピューター等が読み取り可能な記録媒体とは、データやプログラム等の情報を電気的、磁気的、光学的、機械的、または化学的作用によって蓄積し、コンピューター等から読み取ること

ができる記録媒体をいう。

発明の効果

[0022] 本発明によれば、データベースシステムにおいて、複製先のノードがサービス提供可能となるまでに必要なリソースを低減させることが可能となる。

図面の簡単な説明

[0023] [図1]実施形態に係るシステムのハードウェア構成の概略を示す図である。

[図2]実施形態に係るシステムの機能構成の概略を示す図である。

[図3]実施形態において、第一のノードおよび第二のノードによって実行される管理情報送受信処理の流れを示すフローチャートである。

[図4]実施形態において、第一のノードおよび第二のノードによって実行されるデータベース送受信処理の流れを示すフローチャートである。

[図5]実施形態において、第二のノードによって実行されるマップ生成処理の流れを示すフローチャートである。

[図6]実施形態において、第二のノードによって実行される更新実行処理の流れを示すフローチャートである。

[図7]実施形態において、第二のノードによって実行される検索要求対応処理の流れを示すフローチャート（１）である。

[図8]実施形態において、第二のノードによって実行される検索要求対応処理の流れを示すフローチャート（２）である。

[図9]実施形態に係る第一のノードおよび第二のノードの機能構成の概略を示す図である。

[図10]実施形態において、第一のノードおよび第二のノードによって実行されるトランザクションログ送受信開始処理の流れを示すフローチャートである。

[図11]実施形態において、第二のノードによって実行されるオンデマンド処理の流れを示すフローチャートである。

[図12]実施形態において、第二のノードによって実行される削除処理の流れを示すフローチャートである。

発明を実施するための形態

[0024] 以下、本開示に係るシステム、情報処理装置、方法およびプログラムの実施の形態を、図面に基づいて説明する。但し、以下に説明する実施の形態は、実施形態を例示するものであって、本開示に係るシステム、情報処理装置、方法およびプログラムを以下に説明する具体的構成に限定するものではない。実施にあたっては、実施形態に応じた具体的構成が適宜採用され、また、種々の改良や変形が行われてよい。

[0025] 本実施形態では、本開示に係るシステム、情報処理装置、方法およびプログラムを、追記型のデータベースシステムにおいて実施した場合の実施の形態について説明する。追記型のデータベースシステムとは、データの更新の際に、古いデータを新しいデータで上書きすることなく、新しいデータを追記することでデータの更新を行うタイプのデータベースシステムである。但し、本開示に係るシステム、情報処理装置、方法およびプログラムは、複数のノードを有するシステムにおいて、あるノードが管理するデータを他のノードにおいて用いるための技術について広く用いることが可能であり、本開示の適用対象は、本実施形態において示した例に限定されない。

[0026] ＜第一の実施形態＞

はじめに、第一の実施形態について説明する。

[0027] ＜＜システムの構成＞＞

図１は、本実施形態に係るシステムのハードウェア構成の概略を示す図である。本実施形態に係るシステムは、ユーザー端末９からの検索要求（クエリ）に応答するための複数のノード（情報処理装置）１を備える。複数のノード１は、ネットワークを介して互いに通信可能に接続される。複数のノード１のうち、データベースが既に構築されているノード１の何れかを、本実施形態におけるオリジナルのデータベースを有する第一のノード１Ａとすることが出来る。また、本実施形態において第一のノード１Ａからデータを受信する第二のノード１Ｂは、データベース用のソフトウェアがインストールされているが、データベースのコンテンツが未構築のノードである。本実施

形態では、このようなシステムにおいて、第一のノード１Ａが管理するデータベースのコンテンツ（データ）を、第二のノード１Ｂにおける検索の用に供するための技術を説明する。本実施形態では、ノードを区別することなくノード一般について述べる場合には「ノード１」と記載し、ノードを区別して述べる場合には「第一のノード１Ａ」、「第二のノード１Ｂ」のように添字を付して記載する。

[0028] なお、本開示に係るシステムは、階層構造を有するデータベースシステムや、マルチマスターのデータベースシステムに適用可能であるが、この際、データベースシステムにおけるマスターノードおよびスレーブノードの何れであっても、第一のノード１Ａまたは第二のノード１Ｂとすることが出来る。

[0029] 第一のノード１Ａおよび第二のノード１Ｂは、ＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）１１Ａ、１１Ｂ、ＲＡＭ（Ｒａｎｄｏｍ Ａｃｃｅｓｓ Ｍｅｍｏｒｙ）１２Ａ、１２ＢおよびＲＯＭ（Ｒｅａｄ Ｏｎｌｙ Ｍｅｍｏｒｙ）１３Ａ、１３Ｂ等からなる制御部１０Ａ、１０Ｂと、補助記憶装置１４Ａ、１４Ｂと、通信インターフェース１５Ａ、１５Ｂと、を備えるコンピューターである。但し、ノード１の具体的なハードウェア構成に関しては、実施の形態に応じて適宜省略や置換、追加が可能である。また、ノード１は、単一の装置に限定されない。ノード１は、所謂クラウドや分散コンピューティングの技術等を用いた、複数の装置によって実現されてよい。

[0030] なお、本実施形態では、データベースの各レコードは、ページ単位で管理されており、ストレージ（例えば、補助記憶装置１４Ａ、１４Ｂ）とメモリ等（例えば、ＲＡＭ１２Ａ、１２Ｂ）との間のデータのやり取りは、ページ単位で行われる。ここで、メモリ等に展開されてデータの検索または処理に供されている状態のデータは、キャッシュと称される。一方、ストレージ上のデータは、データの検索または処理に供されていない状態にある。即ち、データベースの内容を検索や更新等の処理の対象とする場合、各ノードは、

対象のレコードを含むページをストレージから読み出してキャッシュとし、対象のレコードに係る処理を行う。また、各ノードは、対象のレコードの更新を行った後に、当該レコードを含むキャッシュのページをストレージに書き込むことで、データベースの更新を保存する。なお、ページがストレージに書き込まれるタイミングは、実施の形態に応じて適宜決定される。ページには、１または複数のレコードが格納される。また、ページは、ブロックと称されてもよく、本発明の「所定の管理単位」に相当する。

[0031] 本実施形態では、後述するノード間の送受信処理においても、ページ単位でデータが要求され、ページ単位でデータが送受信される。但し、データの読出／書込および送受信をページ単位で行うのは、本開示を実施する際に採用可能な具体的構成の一例であり、本開示の技術的範囲は、ページ単位での読出／書込および送受信に限定されない。データの読出／書込および送受信は、例えばテーブル単位で行われてもよいし、その他の管理単位で行われてもよい。

[0032] 図２は、本実施形態に係る第一のノード１Ａおよび第二のノード１Ｂの機能構成の概略を示す図である。本実施形態に係る第一のノード１Ａは、ＣＰＵ１１Ａが、ＲＡＭ１２Ａに展開された各種プログラムを解釈および実行して、ノード１Ａに備えられた各種ハードウェアを制御することで、ログ保持部２１、要求受信部２２、特定識別情報送信部２３、ログ送信部２４およびデータ送信部２５を備えるコンピューターとして機能する。本実施形態では、第一のノード１Ａの機能が汎用のＣＰＵ１１Ａによって実行される例について説明しているが、これらの機能は、その一部または全部が、１または複数の専用のプロセッサによって実現されてもよい。

[0033] ログ保持部２１は、第一のノード１Ａによって管理されているデータベースにおけるトランザクションのログ（以下、「トランザクションログ」）を、トランザクションに含まれる命令の時系列上の順序を把握可能な識別情報であるＬＳＮ（Ｌｏｇ　Ｓｅｑｕｅｎｃｅ　Ｎｕｍｂｅｒ）とともに保持する。命令には、データベースの更新に係る更新命令と、トランザクションの

管理に係る管理命令とが含まれる。また、トランザクションログは、ノードにおいてトランザクションが発生する毎に生成され、ＬＳＮが付されてログ保持部２１によって保持される。

[0034] 要求受信部２２は、第二のノード１Ｂによって送信された、第一のノード１Ａによって管理されているデータの送信要求を受信する。

[0035] 特定識別情報送信部２３は、ログ保持部２１によって保持されているＬＳＮのうち、所定の時点におけるトランザクションログを示す特定識別情報（特定ＬＳＮ）を、要求受信部２２によって受信されたデータ送信要求に応じて第二のノード１Ｂに送信する。ここで送信される特定ＬＳＮは、要求されたページ毎に決定されてよい。なお、本実施形態において、特定識別情報送信部２３は、特定ＬＳＮとして、最新のトランザクションログを示すＬＳＮを送信する。特定ＬＳＮを最新のトランザクションログを示すＬＳＮとすることで、ログ送信部２４によって送信すべきトランザクションログの量を低減させ、また、第二のノード１Ｂにおいて反映すべきトランザクションログの量を低減出来る。但し、送信される特定ＬＳＮは、第二のノード１Ｂにおいて反映すべきトランザクションログを判別するための、所定の時点におけるトランザクションログを示すものであればよく、最新のトランザクションログを示すＬＳＮに限定されない。

[0036] ログ送信部２４は、少なくとも、特定ＬＳＮが示す所定の時点より後のトランザクションログおよびＬＳＮを、互いに関連づけて第二のノード１Ｂに送信する。トランザクションログの送信開始後、ログ送信部２４は、新たに発生したトランザクションログについても、順次第二のノード１Ｂに対して送信する。

[0037] データ送信部２５は、要求受信部２２によって受信されたデータ送信要求に応じて、データベースによって管理されているページのうち、要求に係るページであって、少なくとも特定ＬＳＮに係る所定の時点までの内容が反映されたページを、第二のノード１Ｂ宛に送信する。本実施形態では、データ送信部２５は、データを、トランザクションログに基づく更新命令の順序が

互いに依存関係にあるレコードが同一の管理単位に入るように区切られた所定の管理単位（本実施形態では、ページ）毎に送信する。

[0038] なお、本実施形態では、データ送信部２５は、第一のノード１Ａのメモリに展開されてデータの検索または処理に供されている状態のデータ、即ちキャッシュを、第二のノード１Ｂ宛に送信する。送信の際、送信対象のページのキャッシュが無い場合には、第一のノード１Ａは、対象ページをストレージから読み出してキャッシュとする。これは、本実施形態に係るデータベースシステムでは、各ノード１において発生したデータの更新は即座にキャッシュに反映され、キャッシュがデータの最新の状態を表しているために、第二のノード１Ｂ宛に送信されるデータとしてキャッシュを採用することで、送信されるデータを、最新のトランザクションログを示すＬＳＮである特定ＬＳＮに係る所定の時点までの内容が反映されたページとすることが出来るからである。

[0039] 但し、送信されるデータは、少なくとも特定ＬＳＮに係る所定の時点までの内容が反映されたものであればよく、本実施形態において採用されたキャッシュに限定されない。本開示が適用されるデータベースシステムが、ストレージにあるデータに、少なくとも特定ＬＳＮに係る所定の時点までの内容が反映されていることが保証できるものである場合には、ストレージから読み出されたデータがそのまま送信されてもよい。

[0040] また、本実施形態に係る第二のノード１Ｂは、ＣＰＵ１１Ｂが、ＲＡＭ１２Ｂに展開された各種プログラムを解釈および実行して、ノード１Ｂに備えられた各種ハードウェアを制御することで、要求送信部３１、データ受信部３２、実行部３３、マップ生成部３４、検索要求受付部３５、検索応答部３６および展開部３７を備えるコンピューターとして機能する。本実施形態では、第二のノード１Ｂの機能が汎用のＣＰＵ１１Ｂによって実行される例について説明しているが、これらの機能は、その一部または全部が、１または複数の専用のプロセッサによって実現されてもよい。

[0041] 要求送信部３１は、第一のノード１Ａに対して、データベース中のデータ

を指定してデータの送信要求を送信する。本実施形態では、データの指定は、ページ単位で行われる。

[0042] データ受信部３２は、第一のノード１Ａから、トランザクションログ、該トランザクションログのＬＳＮ、ページ毎の特定ＬＳＮおよびデータを受信する。なお、先述の通り、本実施形態では、データの送受信は、ページ単位で行われる。

[0043] 実行部３３は、データの送信要求に応じて第一のノード１Ａから送信され、データ受信部３２によって受信されたデータが、該第二のノード１Ｂのメモリに展開されてデータの検索または処理に供される状態（キャッシュ）となった場合に、受信されたトランザクションログのうち、少なくとも特定ＬＳＮに係るトランザクションログよりも新しいトランザクションログに含まれる更新命令を、メモリに展開されたデータに対して実行する。ここで、トランザクションログに含まれる更新命令は、ＬＳＮによって把握される順序に従って実行される。

[0044] マップ生成部３４は、受信されたトランザクションログのうち、少なくとも特定ＬＳＮが示すトランザクションログよりも新しいトランザクションログに基づいて、トランザクションログに含まれる更新命令の対象となるデータが収容されているページと当該更新命令との関係を示すマップを生成する。

[0045] なお、上述の通り本実施形態では、データはページをもって管理されるため、マップ生成部３４は、更新命令の対象となるデータが収容されたページと更新命令との関係を示すマップを生成するが、マップ生成手段は、その他の管理単位（例えば、テーブル等）と更新命令との関係を示すマップを生成してもよい。また、マップ生成手段は、このような管理単位を用いること無く、更新命令の対象となるレコードと更新命令との関係を示すマップを生成してもよい。

[0046] 検索要求受付部３５は、ユーザー端末９からの検索要求（クエリ）を受け付ける。

[0047] 検索応答部３６は、第一のノード１Ａから得られ、トランザクションログに基づく更新命令が反映されたデータに基づいて、検索要求（クエリ）への応答をユーザー端末９に対して返す。

[0048] 展開部３７は、第一のノード１Ａから受信したデータをメモリに展開して、データの検索または処理に供する。本実施形態において、展開部３７は、クエリや、後述する更新実行処理における求めに応じて、第一のノード１Ａから受信されたデータをメモリに読み出すことでキャッシュとする。

[0049] ＜＜処理の流れ＞＞

次に、本実施形態に係る処理の詳細を説明する。なお、本実施形態において説明される処理の具体的な内容および順序等は、実施する上での一例である。具体的な処理内容および順序等は、実施の形態に応じて適宜選択されてよい。

[0050] 図３は、本実施形態において、第一のノード１Ａおよび第二のノード１Ｂによって実行される管理情報送受信処理の流れを示すフローチャートである。本フローチャートに示された処理は、第二のノード１Ｂにおいて、データベースの構築開始の指示が受け付けられたことを契機として開始される。

[0051] ステップＳ１０１およびステップＳ１０２では、管理情報要求が送受信される。要求送信部３１は、第一のノード１Ａに対して、データベースの管理情報（システムカタログ等）を指定してデータの送信要求を送信する（ステップＳ１０１）。ここで、管理情報は、データベース中のテーブルやレコードの位置を特定可能な情報を含んでいる。本実施形態に係るデータベースはページ単位で管理されているため、管理情報は、データベースのテーブルとページとの関係を特定可能な情報を含む。また、本実施形態では、管理情報の読出／書込および送受信についても、他のテーブルと同様ページ単位で行われる。但し、管理情報がテーブル単位等その他の単位に従って要求されてもよいことは、先述したデータの読出／書込および送受信と同様である。要求受信部２２は、第二のノード１Ｂによって送信された要求を受信する（ステップＳ１０２）。その後、処理はステップＳ１０３へ進む。

[0052] ステップＳ１０３およびステップＳ１０４では、データの送信要求に対する応答（Ａｃｋｎｏｗｌｅｄｇｅｍｅｎｔ。以下、「ＡＣＫ」と称する）が送受信される。このＡＣＫは、管理情報の特定ＬＳＮを含む。先述の通り、ＬＳＮとは、トランザクションログに含まれる命令（更新命令）の順序を把握可能な識別情報であり、特定ＬＳＮとは、ログ保持部２１によって保持されているＬＳＮのうち、所定の時点におけるトランザクションログ（本実施形態では、最新のトランザクションログ）を示すＬＳＮである。即ち、特定識別情報送信部２３は、ログ保持部２１によって保持されているＬＳＮのうち、要求された管理情報に係る特定ＬＳＮを含むＡＣＫを、第二のノード１Ｂに対して送信する（ステップＳ１０３）。データ受信部３２は、第一のノード１Ａから、特定ＬＳＮを含むＡＣＫを受信する（ステップＳ１０４）。その後、処理はステップＳ１０５へ進む。

[0053] ステップＳ１０５およびステップＳ１０６では、トランザクションログの送受信が開始される。ログ送信部２４は、所定の時点より後のトランザクションログおよびＬＳＮを、互いに関連づけて第二のノード１Ｂに送信する（ステップＳ１０５）。以降、第一のノード１Ａにおいて生成されたトランザクションログは、順次、継続的に第二のノード１Ｂに対して送信される。ここで、「所定の時点」とは、ステップＳ１０３で送信された特定ＬＳＮが示すトランザクションログが生成された時点である。データ受信部３２は、第一のノード１Ａによって送信された、トランザクションログおよび該トランザクションログに係るＬＳＮの受信を開始する（ステップＳ１０６）。

[0054] ログの送受信開始後、第一のノード１Ａにおいて新たに発生したトランザクションログは、発生次第、第二のノード１Ｂに対して送信され、第二のノード１Ｂによって受信される。即ち、第二のノード１Ｂは、ログの送信開始後、所定の時点以降のトランザクションログを全て受信することとなる。その後、処理はステップＳ１０７へ進む。

[0055] なお、先述の通り、処理順序は本フローチャートに示した例に限定されない。管理情報送受信処理では、対象のページに関係するトランザクションロ

グのうち、特定ＬＳＮが示す時点以降のトランザクションログが全て第二のノード１Ｂによって受信されればよい。このため、例えば、ステップＳ１０３およびステップＳ１０４の送受信処理と、ステップＳ１０５およびステップＳ１０６の送受信処理との処理順序は、入れ替わってもよい。

[0056] ステップＳ１０７では、管理情報について、マップ生成処理が開始される。マップ生成部３４は、管理情報に関するページを、マップ生成処理の対象とする。即ち、本ステップ以降、管理情報に対して実行される更新命令と、当該更新命令の実行対象（対象ページ、対象テーブルまたは対象レコード等）との組合せがマッピングされたマップの生成が開始される。マップ生成処理は、本フローチャートに示された処理と並行して実行される。マップ生成処理の詳細は、図５を用いて説明する。その後、処理はステップＳ１０８へ進む。

[0057] ステップＳ１０８およびステップＳ１０９では、データベースの管理情報が送受信される。データ送信部２５は、ステップＳ１０２で受信された要求を受けて、データベースの管理情報（システムカタログ等）を、第二のノード１Ｂに対して送信する（ステップＳ１０８）。この際、データ送信部２５は、送信対象のデータのキャッシュを、第二のノード１Ｂ宛に送信する。送信の際、送信対象のデータのキャッシュが無い場合には、第一のノード１Ａは、対象データをストレージから読み出してキャッシュとし、少なくとも特定ＬＳＮに係る所定の時点までの内容が反映されたデータとしてから、第二のノード１Ｂ宛に送信する。第二のノード１Ｂのデータ受信部３２は、第一のノード１Ａから、管理情報を受信する（ステップＳ１０９）。その後、処理はステップＳ１１０へ進む。

[0058] なお、本フローチャートでは、管理情報の送信（ステップＳ１０８）が、マップの生成開始（ステップＳ１０７）よりも後に記載されているが、管理情報は、特定ＬＳＮに係る所定の時点以降のものが送信されればよく、送信のタイミングは本フローチャートに示された例に限定されない。

[0059] ステップＳ１１０では、更新実行処理が開始される。実行部３３は、管理

情報に関するページを、更新実行処理の対象とする。即ち、本ステップ以降、マップ生成処理においてマップに記録された、管理情報に未反映の更新命令が実行される。更新実行処理は、本フローチャートに示された処理と並行して実行される。更新実行処理の詳細は、図６を用いて説明する。その後、処理はステップＳ１１１へ進む。

[0060] ステップＳ１１１では、検索要求（クエリ）の受付が開始される。第二のノード１Ｂは、検索要求対応処理を開始させることで、ユーザー端末９から送信されるクエリの受付を開始する。以降、検索要求受付部３５は、ユーザー端末９からのクエリを受け付け、クエリが受け付けられると、クエリに応じるために必要なページが、管理情報を参照することで取得される。検索要求対応処理は、本フローチャートに示された処理と並行して実行される。検索要求対応処理の詳細は、図７および図８を用いて説明する。その後、本フローチャートに示された処理は終了する。

[0061] なお、本実施形態では、管理情報の複製についても、データベースの他のテーブル同様、マップ生成処理、更新実行処理および検索要求対応処理を用いて複製する例について説明したが、管理情報の複製には、その他の方法が採用されてもよい。管理情報の複製には、例えば、チェックポイントを作成してチェックポイントのスナップショットを複製する従来の方法等を採用することが出来る。

[0062] 図４は、本実施形態において、第一のノード１Ａおよび第二のノード１Ｂによって実行されるデータベース送受信処理の流れを示すフローチャートである。本フローチャートに示された処理は、第二のノード１Ｂが、第一のノード１Ａからデータベースの管理情報を受信することで、第一のノード１Ａによって管理されているデータベースのページ構成を把握可能となったことを契機として開始される。

[0063] ステップＳ２０１およびステップＳ２０２では、データの送信要求が送受信される。要求送信部３１は、管理情報のキャッシュを参照して、第一のノード１Ａに対して、データベース中のページを指定してデータの送信要求を

送信する（ステップＳ２０１）。ページの指定には、例えば、ページ番号が用いられてよい。データの送信要求は、管理情報から目的のデータを含むページが分かる場合には、ページを指定したものであってもよいし、データベース中の具体的なレコードを指定したものであってもよい。本実施形態において、データはページ単位で要求されるが、テーブル単位等その他の単位に従って要求されてもよいことは先述した通りである。要求受信部２２は、第二のノード１Ｂから、第一のノード１Ａによって管理されているデータの送信要求を受信する（ステップＳ２０２）。その後、処理はステップＳ２０３へ進む。

[0064] ステップＳ２０３およびステップＳ２０４では、データの送信要求に対するＡＣＫが送受信される。このＡＣＫは、ページ番号および当該ページの特定ＬＳＮを含む。即ち、特定識別情報送信部２３は、ログ保持部２１によって保持されているＬＳＮのうち、要求されたページに係る特定ＬＳＮを含むＡＣＫを、第二のノード１Ｂに対して送信する（ステップＳ２０３）。データ受信部３２は、第一のノード１Ａから、特定ＬＳＮを含むＡＣＫを受信する（ステップＳ２０４）。その後、処理はステップＳ２０５へ進む。

[0065] ステップＳ２０５では、要求に係るページ等について、マップ生成処理が開始される。マップ生成部３４は、ステップＳ２０１において要求したページを、マップ生成処理の対象とする。即ち、本ステップ以降、要求されたページに対して実行される更新命令と、当該更新命令の実行対象（対象ページ、対象テーブルまたは対象レコード等）との組合せがマッピングされたマップの生成が開始される。ここで開始されるマップ生成処理の内容は、ステップＳ１０７において管理情報に対して開始されたものと概略同様であり（図５を参照）、本フローチャートに示された処理と並行して処理される。その後、処理はステップＳ２０６へ進む。

[0066] なお、本実施形態では、管理情報やページ毎にマップの生成開始タイミングが異なる例について説明しているが（ステップＳ１０７やステップＳ２０５を参照）、マップの生成開始タイミングは、データベース全体において同

時であってもよい。

[0067] ステップＳ２０６およびステップＳ２０７では、要求に係るページが送受信される。データ送信部２５は、ステップＳ２０１で受信された要求に応じて、データベースによって管理されているページのうち要求されたものを、第二のノード１Ｂに送信する（ステップＳ２０６）。この際、データ送信部２５は、送信対象のデータのキャッシュを、第二のノード１Ｂ宛に送信する。送信の際、送信対象のページのキャッシュが無い場合には、第一のノード１Ａは、対象ページをストレージから読み出してキャッシュとし、少なくとも特定ＬＳＮに係る所定の時点までの内容が反映されたページとしてから、第二のノード１Ｂ宛に送信する。第二のノード１Ｂのデータ受信部３２は、第一のノード１Ａから、データを受信する（ステップＳ２０７）。データの送受信単位は、要求に応じたものであればよく、ページ単位、テーブル単位およびレコード単位の何れでもよい。その後、処理はステップＳ２０８へ進む。

[0068] なお、本フローチャートでは、データの送信（ステップＳ２０６）が、マップの生成開始（ステップＳ２０５）よりも後に記載されているが、データは、特定ＬＳＮに係る所定の時点以降のものが送信されればよく、送信のタイミングは本フローチャートに示された例に限定されない。

[0069] ステップＳ２０８では、更新実行処理が開始される。実行部３３は、ステップＳ２０７で受信されたデータを、更新実行処理の対象とする。即ち、本ステップ以降、マップ生成処理でマップに記録された、未実行の更新命令が実行される。その後、本フローチャートに示された処理は終了する。

[0070] なお、本実施形態において、第二のノード１Ｂは、第一のノード１Ａのデータベースにある全てのデータを取得するため、図４に示されたデータベース送受信処理は、管理情報に示された全てのページについて取得が完了するまで、要求するページを順次変更しながら繰り返し実行される。即ち、要求送信部３１は、ステップＳ２０１からステップＳ２０８の処理を繰り返すことで、管理情報を参照して把握される全てのページまたはテーブルについて

、データの送信要求を送信する。但し、第二のノード１Ｂが、第一のノード１Ａのデータベースにある全てのデータを取得する必要がない場合には、要求送信部３１は、必要な一部のページやテーブルについてのみ、データの送信要求を行ってもよい。また、本実施形態では、複数のページを取得する場合に、ステップＳ２０１からステップＳ２０８の処理を繰り返す例について説明したが、複数のページを取得する場合に、ステップＳ２０１において複数ページ分まとめて要求することとしてもよい。

[0071] 本フローチャートに示された処理において第一のノード１Ａから受信されたページには、未反映のトランザクションログが存在する可能性がある。このため、マップ生成処理、更新実行処理および検索要求対応処理（詳細については後述する）が本フローチャートに示された処理と並行して行われることで、未反映のトランザクションログが反映される。これらの処理が開始されるタイミングは、上記説明した通りである。なお、要求に係るページやテーブルに関するトランザクションログは、ステップＳ１０５以降順次受信されているトランザクションログに含まれている。

[0072] なお、本実施形態では、第一のノード１Ａから受信されたデータは、一旦ストレージに保存される。但し、このような処理に代えて、展開部３７は、第一のノード１Ａから受信したデータを、ストレージ（本実施形態では、補助記憶装置１４Ｂ）を介さずに直接メモリに展開して、データの検索または処理に供してもよい。受信データが直接メモリに展開された場合、当該受信データに係る更新命令は、後述するマップ生成処理に従って、即座に反映される。

[0073] 次に、第二のノード１Ｂにおいて実行されるデータベースの管理処理の流れを説明する。本実施形態において、第二のノード１Ｂのデータベースは、マップ生成処理、更新実行処理および検索要求対応処理によって管理される。第二のノード１Ｂは、これらの処理を実行するためのインスタンスを適宜複数立ち上げて並列実行することによって、マップの生成、マッピングされた更新命令の反映、および検索要求（クエリ）の処理の夫々について、並列

処理することを可能としている。

[0074] 図５は、本実施形態において、第二のノード１Ｂによって実行されるマップ生成処理の流れを示すフローチャートである。本フローチャートに示された処理は、マップ生成処理の対象となったデータ（ステップＳ１０７およびステップＳ２０５を参照）に対して、繰り返し実行される。

[0075] ステップＳ３０１では、トランザクションログに含まれる命令が、ＬＳＮを参照して時系列順に参照される。第二のノード１Ｂは、未反映のトランザクションログに含まれる命令を、時系列において最も古い時点のもの（ＬＳＮが最も小さいもの）から順に参照する。ここで参照されるトランザクションログは、ステップＳ１０５およびステップＳ１０６の処理において送受信が開始され、第二のノード１Ｂによって受信されたものである。その後、処理はステップＳ３０２へ進む。

[0076] ステップＳ３０２およびステップＳ３０３では、命令が更新命令であるか否かが判定され、更新命令以外の命令が実行される。第二のノード１Ｂは、ステップＳ３０１で参照された命令が、更新命令であるか否かを判定する（ステップＳ３０２）。参照された命令が更新命令以外の命令（例えば、コミット等の管理命令）であった場合、命令はそのまま実行される（ステップＳ３０３）。一方、参照された命令が更新命令であった場合、処理はステップＳ３０４へ進む。

[0077] ステップＳ３０４では、参照された命令が示すページが、マップ生成処理の対象であるか否かが判定される。第二のノード１Ｂは、対象となっているデータが、ステップＳ１０７やステップＳ２０５等において、マップ生成処理の対象に設定されているか否かを判定する。ここで、対象データが、マップ生成処理の対象に設定されていないと判定された場合、本フローチャートに示された処理は終了する。一方、対象データが、マップ生成処理の対象に設定されていると判定された場合、処理はステップＳ３０５へ進む。

[0078] ステップＳ３０５では、対象データ（ページ）のキャッシュが存在するか否かが判定される。第二のノード１Ｂは、参照した命令が更新命令である場

合、当該命令を参照した時点でその対象がデータベース内に存在するような更新命令（例えば、ＤＥＬＥＴＥ命令）については、当該更新命令の対象のキャッシュが存在するか否かを判定する。一方、当該命令を参照した時点でその対象がデータベース内に存在しないような更新命令（例えば、ＩＮＳＥＲＴ命令）については、当該更新命令の対象が挿入される領域のデータのキャッシュが存在するか否かを判定することで、当該更新命令の対象のキャッシュが存在するか否かを判定する。なお、レコードがページ単位で管理されている場合、第二のノード１Ｂは、更新命令の対象を含むページのキャッシュが存在するか否かを判定する。対象のページのキャッシュが存在する場合、処理はステップＳ３０６へ進む。一方、対象のページのキャッシュが存在しない場合、処理はステップＳ３０７へ進む。

[0079] ステップＳ３０６では、トランザクションログに係る更新命令が実行される。実行部３３は、受信されたデータのうち、展開部３７によってメモリに展開されてキャッシュとなったデータに対して、トランザクションログに係る更新命令を実行する。その後、処理はステップＳ３０８へ進む。

[0080] ステップＳ３０７では、マップが生成される。トランザクションログに含まれる更新命令の対象となっているページのキャッシュが存在しない場合（ステップＳ３０５で「ＮＯ」に進んだ場合）、マップ生成部３４は、トランザクションログに含まれる更新命令の対象となるデータが収容されたページと更新命令との関係を示す情報を、マップに記録することで、更新命令の実行を待機させる。換言すれば、マップ生成部３４は、更新命令の実行を待機させる際に、当該更新命令をマッピング（マップ生成）の対象とし、待機させた更新命令の内容を当該レコードに関連付けて記録するマップを生成する。但し、全ての更新命令をマップ生成の対象としてもよい。本実施形態では、マップには、待機させた更新命令の内容およびＬＳＮのうちの少なくとも一方が、更新命令の対象であるレコードを含むページ（具体的にはページ番号）に関連付けられて記録される。その後、処理はステップＳ３０８へ進む。

[0081] ステップＳ３０８では、マップ生成処理を終了するか否かが判定される。マップ生成処理が終了されない場合、処理はステップＳ３０１へ戻り、トランザクションログにおける時系列順（ＬＳＮの順）で次の命令が参照され、ステップＳ３０４からステップＳ３０７に示された処理の対象となる。通常、第一のノード１Ａにおいて新たに発生したトランザクションログを第二のノード１Ｂに反映するために、マップ生成処理は、第二のノード１Ｂが運用されている間、繰り返し実行される。即ち、本実施形態において、第二のノード１Ｂは、受信されたトランザクションログを、古いものから順に検査し、トランザクションログに含まれる更新命令の対象となっているページがオンメモリ（キャッシュ）である場合に、キャッシュに対してトランザクションログに応じた更新命令を実行する。一方、第二のノード１Ｂが停止される等の理由でマップ生成処理が終了される場合、本フローチャートに示された処理は終了する。

[0082] マップ生成処理で生成されたマップは、対象ページがキャッシュとなった場合に、更新実行処理または検索要求対応処理において参照され、更新実行処理または検索要求対応処理において、当該マップに待機させられた更新命令が実行される。即ち、本実施形態に係るシステムによれば、第一のノードから受信されたデータに対するトランザクションログの反映を、対象データが実際に用いられる時点（換言すれば、キャッシュとなる時点）まで遅延させることが出来る。

[0083] ここで、更新命令の内容を記録したトランザクションログは、ノード１ＢのＲＡＭ１２Ｂまたは補助記憶装置１４Ｂに保持されている。更新実行処理または検索要求対応処理では、当該マップに記録されるＬＳＮに基づいてＲＡＭ１２または記憶装置１４に保持されているトランザクションログへのアクセスが行われ、実行する更新命令の内容が特定される。

[0084] なお、マップ生成処理では、参照した順番に応じて、更新命令の内容またはＬＳＮがマップに記録される。これにより、後述する更新実行処理または検索要求対応処理では、待機させられた更新命令の実行順序を認識すること

ができる。但し、仮に、更新命令の内容またはＬＳＮが参照された順番に応じてマップに記録されなかったとしても、更新実行処理または検索要求対応処理では、待機させられた更新命令のＬＳＮにより、当該更新命令の実行順序を認識することができる。

[0085] 図６は、本実施形態において、第二のノード１Ｂによって実行される更新実行処理の流れを示すフローチャートである。本フローチャートに示された処理は、マップに待機している更新命令が存在する間、繰り返し実行される。

[0086] ステップＳ４０１では、待機している更新命令の有無が判定される。第二のノード１Ｂは、マップを参照し、待機している更新命令があるか否かを判定する。待機している更新命令がマップに記録されていない場合、本フローチャートに示された処理は終了する。一方、待機している更新命令がマップに記録されている場合、第二のノード１Ｂは、更新命令の対象となるデータが含まれる管理単位を、処理の対象として選択して、処理をステップＳ４０２へ進める。

[0087] ステップＳ４０２では、更新対象が存在するか否かが判定される。第二のノード１Ｂは、ステップＳ４０１において選択された処理の対象が、第二のノード１Ｂに存在するか否か、換言すれば、処理の対象となるページが、既に第一のノード１Ａから受信されているか否か、を判定する。更新対象のページが存在しない場合、処理はステップＳ４０１へ戻る。即ち、第二のノード１Ｂは、マップを参照して、第二のノード１Ｂに存在するページについて待機している更新命令を待ち受ける。更新対象のページが存在する場合、処理はステップＳ４０４へ進む。

[0088] ステップＳ４０３では、更新対象ページに関する更新命令がマップから抽出される。第二のノード１Ｂは、ステップＳ４０２において存在すると判定された更新対象のページに対する更新命令を、マップから全て抽出する。このようにすることで、更新対象ページに関するマップ上の更新命令を全て実行し、更新命令が実行されたページをそのままキャッシュとして、データの

検索または処理に供することが出来る。その後、処理はステップＳ４０４へ進む。

[0089] ステップＳ４０４では、展開部３７は、自身が処理する対象を、補助記憶装置１４ＢからＲＡＭ１２Ｂに読み出すことで、キャッシュとする。その後、処理はステップＳ４０５へ進む。

[0090] ステップＳ４０５およびステップＳ４０６では、マップに待機している更新命令および抽出された更新命令が実行され、実行済みの更新命令がマップから削除される。実行部３３は、マップを参照して、自身が選択した処理対象に関連付けられている更新命令を実行する（ステップＳ４０５）。即ち、実行部３３は、ステップＳ４０１で発見された更新命令、およびステップＳ４０３で抽出された更新命令を、所定の管理単位であるページ毎に、マップに記録されている順に実行する。例えば、更新命令が削除命令である場合、実行部３３は、対象のレコードに削除ポインタを付与する。そして、第二のノード１Ｂは、自身が選択した処理対象について待機している更新命令を実行した後に、実行した更新命令に関する記録をマップから削除する（ステップＳ４０６）。その後、処理はステップＳ４０７へ進む。

[0091] ステップＳ４０７では、更新実行処理を終了するか否かが判定される。更新実行処理が終了されない場合、処理はステップＳ４０１へ戻る。即ち、更新実行処理では、処理対象として選択したページに複数の更新命令が待機している場合、マップに記録されている順（ＬＳＮ順）に、当該待機している更新命令が実行される。これにより、そのページに関して、ノード１とレプリケーション元のコンピューターとの間で、レプリケーションが実行されている最中におけるデータベースの整合性を保つことができる。一方、第二のノード１Ｂが停止される等の理由で更新実行処理が終了される場合、本フローチャートに示された処理は終了する。

[0092] なお、第二のノード１Ｂは、更新実行処理を実行するためのインスタンスを複数立ち上げてもよく、このようなインスタンスを生成若しくは削除または起動若しくは停止することで、更新実行処理を実行するインスタンスの数

を増減させてもよい。例えば、第二のノード１Ｂは、マップで記録されている待機させられている更新命令の管理単位の数（本実施形態では、ページ数）に応じて、起動するインスタンスの数を増減させてもよい。この場合は、更新実行処理の処理能力を要求に応じて可変にでき、無駄なインスタンスが存在することを避けることができるため、リソースを効率よく活用することが可能になる。

[0093] また、例えば、更新実行処理を実行するためのインスタンスを、予め一定数起動させておいてもよい。この場合、処理の要求が生じた場合に改めてインスタンスを起動しなくてもよいため、生じた要求に迅速に対応することが可能になる。

[0094] 図７および図８は、本実施形態において、第二のノード１Ｂによって実行される検索要求対応処理の流れを示すフローチャートである。本フローチャートに示された処理は、第二のノード１Ｂが、ユーザー端末９等からクエリを受信したことを契機として開始される。

[0095] ステップＳ５０１では、クエリが受け付けられる。検索要求受付部３５は、ユーザー端末９からのクエリを受け付ける。その後、処理はステップＳ５０２へ進む。

[0096] ステップＳ５０２では、管理情報のキャッシュが存在するか否かが判定される。検索応答部３６は、メモリ上に、管理情報のキャッシュが存在するか否かを判定する。なお、本実施形態では、レコードはページ単位で管理されているため、検索応答部３６は、検索対象での管理情報を含むページのキャッシュが存在するか否かを判定する。

[0097] 本実施形態では、管理情報を含むページがＲＡＭ１２Ｂに読み出された段階で、当該ページに対して、待機させられた更新命令が実行される（図６および後述するステップＳ５０４からステップＳ５０７に示す処理を参照）。また、キャッシュが存在するページに対する更新命令は即座に実行される（図５を用いて説明したマップ生成処理を参照）。このため、本実施形態において、管理情報を含むページのキャッシュが既に存在する場合（ステップＳ

５０２における判定結果が「ＹＥＳ」の場合）、当該ページに対する更新命令の実行が待機されていることはない。従って、管理情報を含むページのキャッシュが存在する場合、検索応答部３６は、記憶領域からの読み出しを行わず、既にキャッシュが存在する管理情報を参照の対象とし、処理はステップＳ５０８へ進む。一方、管理情報のキャッシュが存在しない場合、処理はステップＳ５０３へ進む。

[0098] ステップＳ５０３では、管理情報がストレージから読み出される。本実施形態では、データベースの構築開始の時点で管理情報送受信処理が行われるため（図３を参照）、クエリを受けた時点で管理情報がキャッシュに存在しない場合にも、第二のノード１Ｂは、ストレージに管理情報を有している。展開部３７は、管理情報のキャッシュが存在しない場合、管理情報をストレージ（補助記憶装置１４Ｂ）からＲＡＭ１２Ｂに読み出す。その後、処理はステップＳ５０４へ進む。

[0099] ステップＳ５０４では、待機している更新命令の有無が判定される。第二のノード１Ｂは、マップを参照し、ＲＡＭ１２Ｂに読み出された管理情報のページについて待機している更新命令があるか否かを判定する。ＲＡＭ１２Ｂに読み出されたページについて待機している更新命令がマップに記録されていない場合、処理はステップＳ５０８へ進む。一方、ＲＡＭ１２Ｂに読み出したページについて待機している更新命令がマップに記録されている場合、処理はステップＳ５０５に進む。

[0100] ステップＳ５０５では、更新対象ページに関する更新命令がマップから抽出される。第二のノード１Ｂは、ステップＳ５０３において存在すると判定された更新対象のページ（ここでは、管理情報を含むページ）に対する更新命令を、マップから全て抽出する。このようにすることで、更新対象ページに関するマップ上の更新命令を全て実行し、更新命令が実行されたページをそのままキャッシュとして、データの検索または処理に供することが出来る。その後、処理はステップＳ５０６へ進む。

[0101] ステップＳ５０６およびステップＳ５０７では、マップに待機している更

新命令および抽出された更新命令が実行され、実行済みの更新命令がマップから削除される。実行部３３は、マップを参照して、ＲＡＭ１２Ｂに読み出した管理情報に関連づけられている更新命令を実行する（ステップＳ５０６）。即ち、本実施形態において、第二のノード１Ｂは、クエリに応じて（ステップＳ５０１）、データ受信部３２によって受信された管理情報（ステップＳ２０７）をメモリに展開し（ステップＳ５０３）、マップに基づいて更新命令を管理情報に対して実行する。

[0102] ここで、実行部３３は、マップに記録されている順に、待機している更新命令を実行する。そして、第二のノード１Ｂは、ＲＡＭ１２Ｂに読み出した管理情報について待機している更新命令を全て実行した後に、実行した更新命令に関する記録をマップから削除する（ステップＳ５０７）。その後、処理はステップＳ５０８へ進む。

[0103] ステップＳ５０８では、管理情報が参照されてページが特定される。検索応答部３６は、受信されたクエリにより検索範囲に指定された記憶領域を特定するために、管理情報を参照する。本実施形態では、管理情報を参照することで、クエリを処理するために必要なデータを含むページが特定される。その後、処理はステップＳ５１１へ進む。

[0104] ステップＳ５１１では、検索対象のデータ（ページ）のキャッシュが存在するか否かが判定される。検索応答部３６は、受信したクエリの内容に適合するレコードを抽出するために、当該クエリにより検索範囲に指定された記憶領域にアクセスする。そして、検索応答部３６は、メモリ上に、クエリに係るレコードのキャッシュが存在するか否かを判定する。なお、本実施形態では、レコードはページ単位で管理されているため、検索応答部３６は、検索対象でのレコードを含むページのキャッシュが存在するか否かを判定する。

[0105] 本実施形態では、ページがＲＡＭ１２Ｂに読み出された段階で、当該ページに対して、待機させられた更新命令が実行される（図６および後述する図８のステップＳ５１５からステップＳ５１８に示す処理を参照）。また、キ

ャッシュが存在するページに対する更新命令は即座に実行される（図５を用いて説明したマップ生成処理を参照）。このため、本実施形態において、検索範囲に含まれるページのキャッシュが既に存在する場合（ステップＳ５１１における判定結果が「ＹＥＳ」の場合）、当該ページに対する更新命令の実行が待機されていることはない。従って、対象のページのキャッシュが存在する場合、検索応答部３６は、記憶領域からの読み出しを行わず、既にキャッシュが存在するページを検索処理の対象とし、処理はステップＳ５１９へ進む。一方、対象のページのキャッシュが存在しない場合、処理はステップＳ５１２へ進む。

[0106] ステップＳ５１２では、検索対象のページが第二のノード１Ｂのストレージ（データベース）に存在するか否かが判定される。上述の通り、本実施形態では、ページの送信要求は逐次的に行われるため（ステップＳ２０１を参照）、クエリを受けた時点で、第二のノード１Ｂがクエリに係るページを有していない場合がある。ページがストレージに存在しないと判定された場合、該当するページを第一のノード１Ａから取得するために、処理はステップＳ５１４へ進む。一方、ページがストレージに存在すると判定された場合、処理はステップＳ５１３へ進む。

[0107] ステップＳ５１４では、図４に示されたデータベース送受信処理が実行される。ステップＳ５１２においてページがストレージに存在しないと判定された場合、要求送信部３１は、第一のノード１Ａに対して、データベース中の少なくともクエリに係るデータを指定してデータの送信要求を送信する。本実施形態では、図４に示されたデータベース送受信処理が繰り返されることで管理情報に示されたページが順番に取得されるが、ステップＳ５０１で検索要求を受けて検索対象となったページについては、この順番を無視して、優先的にデータ送信要求される。即ち、本実施形態によれば、検索要求を受けて検索対象となったページを、優先的にリクエスト（データ送信要求）の対象とすることが出来る。その後、処理はステップＳ５１１へ戻る。

[0108] そして、第一のノード１Ａから対象ページが取得されるまで、ステップＳ

５１１からステップＳ５１４に示された処理が繰り返され、ステップＳ５０１で受け付けられたクエリへの対応は、対象ページが第一のノード１Ａから取得されるまで待機される。

[0109] ステップＳ５１３では、検索対象のページがストレージから読み出される。展開部３７は、検索対象のデータ（ページ）のキャッシュが存在しないが、第二のノード１Ｂのストレージに存在する場合、そのページをストレージ（補助記憶装置１４Ｂ）からＲＡＭ１２Ｂに読み出し、検索処理の対象とする。その後、処理はステップＳ５１５へ進む。

[0110] なお、展開部３７は、検索範囲に含まれるページを補助記憶装置１４ＢからＲＡＭ１２Ｂに読み出す場合に、例えば、１つ１つ対象のページを読み出してもよいし、複数のページを一度に読み出してもよい。この場合に、実行部３３は、１または複数のページを読み出す毎に、読み出した１または複数のページに対して、後述するステップＳ５１５以降の処理を実行する。

[0111] ステップＳ５１５では、待機している更新命令の有無が判定される。第二のノード１Ｂは、マップを参照し、ＲＡＭ１２Ｂに読み出されたページについて待機している更新命令があるか否かを判定する。ＲＡＭ１２Ｂに読み出されたページについて待機している更新命令がマップに記録されていない場合、処理はステップＳ５１９へ進む。一方、ＲＡＭ１２Ｂに読み出したページについて待機している更新命令がマップに記録されている場合、処理はステップＳ５１６に進む。

[0112] ステップＳ５１６では、更新対象ページに関する更新命令がマップから抽出される。第二のノード１Ｂは、ステップＳ５１５において存在すると判定された更新対象のページに対する更新命令を、マップから全て抽出する。このようにすることで、更新対象ページに関するマップ上の更新命令を全て実行し、更新命令が実行されたページをそのままキャッシュとして、データの検索または処理に供することが出来る。その後、処理はステップＳ５１７へ進む。

[0113] ステップＳ５１７およびステップＳ５１８では、マップに待機している更

新命令および抽出された更新命令が実行され、実行済みの更新命令がマップから削除される。実行部３３は、マップを参照して、ＲＡＭ１２Ｂに読み出したページに係る更新命令を実行する（ステップＳ５１７）。即ち、本実施形態において、第二のノード１Ｂは、クエリに応じて（ステップＳ５０１）、データ受信部３２によって受信されたデータ（ステップＳ２０７）をメモリに展開し（ステップＳ５１３）、マップに基づいて更新命令を該データに対して実行する。

[0114] ここで、実行部３３は、マップに記録されている順（ＬＳＮ順）に、待機している更新命令を実行する。そして、第二のノード１Ｂは、ＲＡＭ１２Ｂに読み出したページについて待機している更新命令を全て実行した後に、実行した更新命令に関する記録をマップから削除する（ステップＳ５１８）。その後、処理はステップＳ５１９へ進む。

[0115] このように、検索要求対応処理では、更新実行処理と同様、処理対象として選択したページに複数の更新命令が待機している場合、マップに記録されている順（ＬＳＮ順）に、当該待機している更新命令が実行される。これにより、そのページに関して、第一のノード１Ａと第二のノード１Ｂとの間で、データベースの整合性を保つことができる。

[0116] ステップＳ５１９では、検索応答部３６は、第一のノード１Ａから得られ、更新命令が反映されたページに対して、受信したクエリに応じた検索処理を実行する。その後、処理はステップＳ５２０へ進む。

[0117] ステップＳ５２０では、検索要求（クエリ）に対する応答が送信される。検索応答部３６は、クエリに対する応答として、検索処理の結果をユーザー端末９に送信する。その後、本フローチャートに示された処理は終了する。

[0118] なお、第二のノード１Ｂは、検索要求対応処理を実行するためのインスタンスを複数立ち上げてもよく、このようなインスタンスを生成若しくは削除または起動若しくは停止することで、検索要求対応処理を実行するインスタンスの数を増減させてもよい。例えば、第二のノード１Ｂは、ユーザー端末９から受け付けたクエリの数に応じて起動するインスタンスの数を増減させ

てもよい。この場合、検索要求対応処理の処理能力を要求に応じて可変にでき、無駄なインスタンスが存在することを避けることができるため、リソースを効率よく活用することが可能になる。

[0119] また、検索要求対応処理は、複数のインスタンスに分けて実行されてもよい。例えば、検索要求対応処理のうち、ステップＳ５１５からステップＳ５１８の処理を、他の処理とは異なるインスタンスで実行することとしてもよい。この場合、例えば、ステップＳ５１５からステップＳ５１８以外の処理を実行するための第一のインスタンスにおいてページ読み出し（ステップＳ５１３）が行われた後に、ページ更新のため、処理がステップＳ５１５からステップＳ５１８の処理を実行するための第二のインスタンスに引き継がれる。そして、第二のインスタンスにおけるステップＳ５１８の処理が完了すると、処理は第一のインスタンスに戻り、ステップＳ５１９の検索処理が実行される。

[0120] ＜＜第一の実施形態に係るシステムの効果＞＞

本実施形態に係るシステムによれば、データベース全体を一括コピーすること無く、また、チェックポイントを作成すること無く、データベースの部分（ページ等の所定の管理単位）毎にデータを第一のノード１Ａから第二のノード１Ｂに送信出来る。また、第二のノード１Ｂは、トランザクションログに基づく更新命令同士の依存関係を気にすること無く、データベースの部分毎に並列にトランザクションログを反映することが出来る。また、マップを生成することで、トランザクションログに基づく更新命令の実行を遅延させることが出来、受信されたデータがメモリに展開された場合に、マップに基づいて、トランザクションログに基づく更新命令を、所定の管理単位毎に並列に実行することが出来る。

[0121] 結果として、本実施形態に係るシステムによれば、データベース全体の複製完了を待つこと無く、受信したデータを第二のノード１Ｂにおけるクエリに供することが出来る。

[0122] また、本実施形態に係るシステムでは、第一のノード１Ａから第二のノー

ド１Ｂへのデータ送信を、データの送信要求に応じたものとすることで、データ取得を逐次的にする等、データ取得のタイミングを調整することが出来る。更に、第二のノード１Ｂにおいて受け付けられた検索要求に応じてデータの送信要求を行ったり、トランザクションログの反映を行ったりすることで、検索対象となったデータを優先的に取得したり、優先的にトランザクションログを反映したりすることが可能となる。即ち、本実施形態に係るシステムによれば、第二のノード１Ｂが全てのデータを受信していない状態で、ユーザー端末からの検索要求を受け付けることが可能となる。

[0123] ＜第二の実施形態＞

次に、第二の実施形態について説明する。第二の実施形態に係るシステムは、第二のノードにデータベース全体を保持することなく、第二のノードがクエリに応答可能とするために、第一の実施形態に係るシステムに構成を追加したものである。第一の実施形態では、全てのページについてのデータベース送受信処理が完了した場合、第二のノードにデータベース全体が保持されることとなるが、第二の実施形態では、データベース中のデータは、保持対象と判定されたものを除いて、第二のノードに恒久的には保持されない。第二の実施形態に係るシステムのうち、上述した第一の実施形態のシステムと共通する構成には、同一の符号を付し、説明を省略する。

[0124] ＜＜システムの構成＞＞

第二の実施形態に係るシステムのハードウェア構成は、第一の実施形態に係るシステムのハードウェア構成と概略同様であるため、説明を省略する（図１を参照）。但し、第二の実施形態では、第二のノードの機能構成が第一の実施形態とは異なるため、第二のノードに符号「１Ｂ’」を付して説明する。

[0125] 図９は、本実施形態に係る第一のノード１Ａおよび第二のノード１Ｂ’の機能構成の概略を示す図である。本実施形態に係る第一のノード１Ａの機能構成は、第一の実施形態と概略同様であるため説明を省略する。本実施形態に係る第二のノード１Ｂ’は、ＣＰＵ１１Ｂが、ＲＡＭ１２Ｂに展開された

各種プログラムを解釈および実行して、ノード１Ｂ’に備えられた各種ハードウェアを制御することで、要求送信部３１、データ受信部３２、実行部３３、マップ生成部３４、検索要求受付部３５、検索応答部３６、展開部３７、判定部３８、記録部３９および削除部４０を備えるコンピューターとして機能する。本実施形態では、第二のノード１Ｂ’の機能が汎用のＣＰＵ１１Ｂによって実行される例について説明しているが、これらの機能は、その一部または全部が、１または複数の専用のプロセッサによって実現されてもよい。

[0126] 要求送信部３１、データ受信部３２、実行部３３、マップ生成部３４、検索要求受付部３５、検索応答部３６および展開部３７については、第一の実施形態と概略同様であるため、説明を省略する。

[0127] 判定部３８は、検索要求に応じて第一のノード１Ａから得られ、トランザクションに基づく命令が実行されたデータを、第二のノード１Ｂ’における保持対象とするか否かを判定する。

[0128] 記録部３９は、判定部３８によって保持対象と判定されたデータを、不揮発性の記憶装置である補助記憶装置１４Ｂに記録する。

[0129] 削除部４０は、検索要求に応じて第一のノード１Ａから得られ、トランザクションに基づく命令が実行されたデータを、検索応答部３６による応答に用いた後、所定の条件に従って自動的に第二のノードから削除する。

[0130] ＜＜処理の流れ＞＞

次に、本実施形態に係る処理の詳細を説明する。なお、本実施形態において説明される処理の具体的な内容および順序等は、実施する上での一例である。具体的な処理内容および順序等は、実施の形態に応じて適宜選択されてよい。

[0131] なお、第一の実施形態において説明したデータベース送受信処理（図４を参照）、マップ生成処理（図５を参照）および更新実行処理（図６を参照）は、第二の実施形態でも実行される。これらの処理の詳細は、第一の実施形態において説明した通りであるため、説明を省略する。

[0132] 図１０は、本実施形態において、第一のノード１Ａおよび第二のノード１Ｂ’によって実行されるトランザクションログ送受信開始処理の流れを示すフローチャートである。本フローチャートに示された処理は、第一の実施形態において説明した管理情報送受信処理（図３を参照）に代えて実行される処理であり、データベースの構築開始の指示が受け付けられたことを契機として開始される。

[0133] ステップＳ８０１およびステップＳ８０２では、トランザクションログ要求が送受信される。要求送信部３１は、第一のノード１Ａに対して、トランザクションログの送信要求を送信する（ステップＳ８０１）。要求受信部２２は、第二のノード１Ｂ’によって送信された要求を受信する（ステップＳ８０２）。その後、処理はステップＳ８０３へ進む。

[0134] ステップＳ８０３およびステップＳ８０４では、トランザクションログの送受信が開始される。ログ送信部２４は、以降、第一のノード１Ａにおいて新たに発生したトランザクションログおよびＬＳＮを、発生次第、互いに関連づけて第二のノード１Ｂ’に順次送信する（ステップＳ８０３）。第二のノード１Ｂ’のデータ受信部３２は、第一のノード１Ａによって送信された、トランザクションログおよび該トランザクションログに係るＬＳＮの受信を開始する（ステップＳ８０４）。即ち、第二のノード１Ｂ’は、ログの送信開始後、所定の時点以降のトランザクションログを全て受信することとなる。その後、処理はステップＳ８０５へ進む。

[0135] ステップＳ８０５では、検索要求（クエリ）の受付が開始される。第二のノード１Ｂ’は、後述するオンデマンド処理を開始させることで、ユーザー端末９から送信されるクエリの受付を開始する。以降、検索要求受付部３５は、ユーザー端末９からのクエリを受け付け、クエリが受け付けられると、クエリに応じるために必要なページ（管理情報を含む）が取得される。その後、本フローチャートに示された処理は終了する。

[0136] 即ち、第二の実施形態では、第一の実施形態と異なり、データベースの構築開始の時点で管理情報送受信処理が行われない。

[0137] また、図４に示すデータベース送受信処理については、第一の実施形態において「第一のノード１Ａによって管理されているデータベースのページ構成を把握可能となったことを契機として開始される」と説明したが、第二の実施形態では、第一の実施形態とは異なり、後述するオンデマンド処理の中で呼び出されるタイミングで実行される（後述するステップＳ６０５およびステップＳ５１４を参照）。

[0138] 図１１は、本実施形態において、第二のノード１Ｂ’によって実行されるオンデマンド処理の流れを示すフローチャートである。本フローチャートに示された処理は、第二のノード１Ｂ’が、ユーザー端末９等からクエリを受信したことを契機として開始される。即ち、本実施形態では、第一の実施形態において説明した検索要求対応処理（図７および図８を参照）に代えて、オンデマンド処理が実行される。

[0139] 但し、第二の実施形態では、上述の通り、データベースの構築開始の時点で管理情報送受信処理が行われていない。このため、検索要求が受け付けられた時点で、第二のノード１Ｂ’が管理情報を有していない可能性がある。そして、第二のノード１Ｂ’が管理情報を有していない場合、データの送信要求においてクエリに係るページを指定することが出来ない。このため、本実施形態に係るオンデマンド処理では、検索要求が受け付けられると、はじめに、第二のノード１Ｂ’が管理情報を有しているか否かが確認される。そして、第二のノード１Ｂ’が管理情報を有していない場合、これを第一のノード１Ａから取得してから、データベース送受信処理が実行される。以下、本実施形態に係るオンデマンド処理の詳細を、図１１を参照しながら説明する。

[0140] ステップＳ６０１では、クエリが受け付けられる。検索要求受付部３５は、ユーザー端末９からのクエリを受け付ける。その後、処理はステップＳ６０２へ進む。

[0141] ステップＳ６０２では、管理情報のキャッシュが存在するか否かが判定される。検索応答部３６は、メモリ上に、管理情報のキャッシュが存在するか

否かを判定する。なお、本実施形態では、レコードはページ単位で管理されているため、検索応答部３６は、検索対象での管理情報を含むページのキャッシュが存在するか否かを判定する。

[0142] 本実施形態では、管理情報を含むページがＲＡＭ１２Ｂに読み出された段階で、当該ページに対して、待機させられた更新命令が実行される（図６および後述するステップＳ６０６からステップＳ６０９に示す処理を参照）。また、キャッシュが存在するページに対する更新命令は即座に実行される（図５を用いて説明したマップ生成処理を参照）。このため、本実施形態において、管理情報を含むページのキャッシュが既に存在する場合（ステップＳ６０２における判定結果が「ＹＥＳ」の場合）、当該ページに対する更新命令の実行が待機されていることはない。従って、管理情報を含むページのキャッシュが存在する場合、検索応答部３６は、記憶領域からの読み出しを行わず、既にキャッシュが存在する管理情報を参照の対象とし、処理はステップＳ６１０へ進む。一方、管理情報のキャッシュが存在しない場合、処理はステップＳ６０３へ進む。

[0143] ステップＳ６０３では、管理情報を含むページが第二のノード１Ｂ’のストレージ（データベース）に存在するか否かが判定される。上述の通り、本実施形態では、データベースの構築開始の時点で管理情報送受信処理が行われていない可能性があるため、クエリを受けた時点で、第二のノード１Ｂ’がクエリに係るページを特定するための管理情報を有していない場合がある。管理情報がストレージに存在しないと判定された場合、管理情報を第一のノード１Ａから取得するために、処理はステップＳ６０５へ進む。一方、ページがストレージに存在すると判定された場合、処理はステップＳ６０４へ進む。

[0144] ステップＳ６０５では、図４に示されたデータベース送受信処理が実行される。ステップＳ６０３において管理情報がストレージに存在しないと判定された場合、要求送信部３１は、クエリに応じて、第一のノード１Ａに対して、管理情報に係るデータを指定してデータの送信要求を送信する。上述の

通り、本実施形態では、管理情報の送受信についても、他のテーブルと同様ページ単位で行われるため、図４に示されたデータベース送受信処理によって管理情報を取得することが出来る。その後、処理はステップＳ６０２へ戻る。そして、第一のノード１Ａから管理情報が取得されるまで、ステップＳ６０２からステップＳ６０５に示された処理が繰り返される。

[0145] ステップＳ６０４では、管理情報がストレージから読み出される。展開部３７は、管理情報のキャッシュが存在しないが、第二のノード１Ｂ’のストレージに存在する場合、管理情報をストレージ（補助記憶装置１４Ｂ）からＲＡＭ１２Ｂに読み出す。その後、処理はステップＳ６０６へ進む。

[0146] ステップＳ６０６では、待機している更新命令の有無が判定される。第二のノード１Ｂ’は、マップを参照し、ＲＡＭ１２Ｂに読み出された管理情報のページについて待機している更新命令があるか否かを判定する。ＲＡＭ１２Ｂに読み出されたページについて待機している更新命令がマップに記録されていない場合、処理はステップＳ６１０へ進む。一方、ＲＡＭ１２Ｂに読み出したページについて待機している更新命令がマップに記録されている場合、処理はステップＳ６０７に進む。

[0147] ステップＳ６０７では、更新対象ページに関する更新命令がマップから抽出される。第二のノード１Ｂ’は、ステップＳ６０３において存在すると判定された更新対象のページ（ここでは、管理情報を含むページ）に対する更新命令を、マップから全て抽出する。このようにすることで、更新対象ページに関するマップ上の更新命令を全て実行し、更新命令が実行されたページをそのままキャッシュとして、データの検索または処理に供することが出来る。その後、処理はステップＳ６０８へ進む。

[0148] ステップＳ６０８およびステップＳ６０９では、マップに待機している更新命令および抽出された更新命令が実行され、実行済みの更新命令がマップから削除される。実行部３３は、マップを参照して、ＲＡＭ１２Ｂに読み出した管理情報に係る更新命令を実行する（ステップＳ６０８）。即ち、本実施形態において、第二のノード１Ｂ’は、クエリに応じて（ステップＳ６０

１）、データ受信部３２によって受信された管理情報（ステップＳ２０７）をメモリに展開し（ステップＳ６０４）、マップに基づいて更新命令を管理情報に対して実行する。

[0149] ここで、実行部３３は、マップに記録されている順に、待機している更新命令を実行する。そして、第二のノード１Ｂ’は、ＲＡＭ１２Ｂに読み出した管理情報について待機している更新命令を全て実行した後に、実行した更新命令に関する記録をマップから削除する（ステップＳ６０９）。その後、処理はステップＳ６１０へ進む。

[0150] ステップＳ６１０では、管理情報が参照されてページが特定される。検索応答部３６は、受信されたクエリにより検索範囲に指定された記憶領域を特定するために、管理情報を参照する。本実施形態では、管理情報を参照することで、クエリを処理するために必要なデータを含むページが特定される。その後、処理は図８を用いて説明した検索要求対応処理のステップＳ５１１からステップＳ５２０に示す処理の内容と概略同様の処理へ進む。

[0151] 即ち、第一のノード１Ａから対象ページが取得されるまで、ステップＳ５１１からステップＳ５１４に示された処理が繰り返され、ステップＳ６０１で受け付けられたクエリへの対応は、対象ページが第一のノード１Ａから取得されるまで待機される。本実施形態では、検索要求を受けて検索対象となることで、はじめてリクエスト（データ送信要求）の対象となる。その後、ステップＳ５１９およびステップＳ５２０に示された処理が実行され、クエリに対する応答が完了すると、本フローチャートに示された処理は終了する。

[0152] 図１２は、本実施形態において、第二のノード１Ｂ’によって実行される削除処理の流れを示すフローチャートである。本フローチャートに示された処理は、第二のノード１Ｂ’において、メモリに展開されてキャッシュとなっているデータの単位（本実施形態では、ページ単位）毎に、定期的に実行される。

[0153] ステップＳ７０１では、データが削除条件を満たしたか否かが判定される

。削除部４０は、第二のノード１Ｂ’上にあるキャッシュのうち、判定対象のページに係るキャッシュついて、所定の削除条件を満たしたか否かを判定する。ここで用いられる削除条件は、例えば、データ処理に関する時点（第一のノード１Ａから取得された時点や、クエリ応答や更新等のために最後に参照された時点等）から所定の時間が経過したこと、等である。但し、削除条件は、本実施形態における例示に限定されない。削除条件を満たさないと判定された場合、判定対象のページに係るキャッシュは削除されず、本フローチャートに示された処理は終了する。一方、削除条件を満たしたと判定された場合、処理はステップＳ７０２へ進む。

[0154] ステップＳ７０２では、データが保持の対象であるか否かが判定される。判定部３８は、第二のノード１Ｂ’によって保持されているデータのうち、判定対象のページについて、第二のノード１Ｂ’における保持対象とするか否かを判定する。ここで、判定のための条件は、例えば、当該データがクエリにおいて参照される頻度等に基づいて決定されてよい。例えば、参照される頻度の高いデータを保持の対象とし、頻度の低いデータを保持の対象としないことで、クエリへの応答速度を犠牲にすること無く、第二のノード１Ｂ’のストレージやメモリを節約することが出来る。また、例えば、予め管理者等によって指定されたテーブルやページを保持の対象として設定しておくことも可能である。但し、保持対象とするか否かの判定条件は、本実施形態における例示に限定されない。判定条件は、実施の形態に応じて適宜採用することが出来る。判定対象のページが、保持の対象となるための条件を満たさない場合、処理はステップＳ７０４へ進む。一方、判定対象のページが、保持の対象となるための条件を満たしている場合、処理はステップＳ７０３へ進む。

[0155] ステップＳ７０３では、データがストレージに記録される。記録部３９は、判定部３８によって保持対象と判定されたデータ（本実施形態では、ページ単位で判定される）を、ストレージ（本実施形態では、補助記憶装置１４Ｂ）に記録する。その後、処理はステップＳ７０５へ進む。

[0156] ステップＳ７０４では、削除条件を満たしたデータがストレージから削除される。削除部４０は、判定部３８によって保持対象ではないと判定されたデータ（本実施形態では、ページ単位で判定される）が、第二のノード１Ｂ’のストレージに記録されている場合に、これをストレージ（本実施形態では、補助記憶装置１４Ｂ）から削除する。このようにすることで、受信されてキャッシュとなる前に一旦ストレージに記録されたデータについても、ストレージから削除することが出来る。その後、処理はステップＳ７０５へ進む。

[0157] ステップＳ７０５では、削除条件を満たしたデータのキャッシュが削除される。削除部４０は、削除条件を満たしたページのキャッシュを、第二のノード１Ｂ’のメモリから削除する。その後、本フローチャートに示された処理は終了する。

[0158] なお、図１２に示したフローチャートでは、キャッシュとなっているデータについて削除の対象とするか否かを判定している。但し、ストレージに保持されているがキャッシュとなっていないデータについても、定期的に保持対象であるか否かを判定し、保持対象となるための条件（ステップＳ７０２を参照）を満たさなくなった場合にはストレージから削除することとしてもよい。このようにすることで、ストレージにのみ保持されているようなデータについても、時間の経過に従ってストレージから削除し、ストレージの容量を節約することが出来る。

[0159] ＜＜第二の実施形態に係るシステムの効果＞＞

第二の実施形態に係るシステムによっても、第一の実施形態に係るシステムと同様の効果を得ることが出来る。また、第二の実施形態では、更に判定部３８、記録部３９および削除部４０を備えることで、第二のノード１Ｂ’のストレージやメモリを節約しながら、ユーザー端末９に対して、あたかも第二のノード１Ｂ’がデータベース全体に係るデータを有しているかのようにサービスを提供することが出来る。

[0160] また、判定部３８や削除部４０によって用いられる条件を実施の形態に応

じて適宜設定することで、ストレージリソースとネットワークリソースのバランスに応じて、リソースの無駄が少ないシステムを構築することが可能となる。

請求の範囲

[請求項1] 複数のノードを有するデータベースシステムであって、

前記複数のノードのうち、データベースの複製元である第一のノードは、

該第一のノードによって管理されているデータベースのトランザクションログを、該トランザクションログに係る命令の順序を把握可能な識別情報とともに保持するログ保持手段と、

前記ログ保持手段によって保持されている前記識別情報のうち、所定の時点におけるトランザクションログを示す特定識別情報を、前記複数のノードのうち、データベースの複製先である第二のノードに送信する特定識別情報送信手段と、

少なくとも前記所定の時点より後の前記トランザクションログおよび前記識別情報を、互いに関連づけて前記第二のノードに送信するログ送信手段と、

前記データベースによって管理されているデータを、前記所定の時点以降に、前記第二のノードに送信するデータ送信手段と、

を備え、

前記第二のノードは、

前記第一のノードから、前記トランザクションログ、該トランザクションログの識別情報、前記特定識別情報および前記データを受信する受信手段と、

受信された前記データが該第二のノードのメモリに展開されてデータの検索または処理に供される状態となった場合に、受信された前記トランザクションログのうち、少なくとも前記特定識別情報が示すトランザクションログよりも新しい前記トランザクションログに係る命令を、メモリに展開された前記データに対して実行する実行手段と、

を備える、

データベースシステム。

[請求項2] 前記データ送信手段は、前記データを、前記トランザクションログに係る命令の順序が互いに依存関係にあるレコードが同一の管理単位に入るように区切られた所定の管理単位毎に送信し、

前記実行手段は、前記トランザクションログに係る命令を、前記所定の管理単位毎に実行する、

請求項１に記載のデータベースシステム。

[請求項3] 受信された前記トランザクションログのうち、少なくとも前記特定識別情報が示すトランザクションログよりも新しい前記トランザクションログに基づいて、命令の対象となるデータが収容された前記所定の管理単位と命令の内容との関係を示すマップを生成するマップ生成手段を更に備え、

前記実行手段は、前記マップを参照して、前記トランザクションログに係る命令を、前記所定の管理単位毎に実行する、

請求項２に記載のデータベースシステム。

[請求項4] 受信された前記トランザクションログのうち、少なくとも前記特定識別情報が示すトランザクションログよりも新しい前記トランザクションログに基づいて、命令の対象となるデータと命令の内容との関係を示すマップを生成するマップ生成手段を更に備え、

前記実行手段は、前記マップを参照して、前記トランザクションログに係る命令を、命令の対象となるデータに対して実行する、

請求項１に記載のデータベースシステム。

[請求項5] 前記データ送信手段は、前記第一のノードのメモリに展開されてデータの検索または処理に供されている状態のデータを、前記第二のノードに送信する、

請求項１から４の何れか一項に記載のデータベースシステム。

[請求項6] 前記第二のノードは、

前記第一のノードに対して、前記データベース中のデータを指定

して送信要求を送信する要求送信手段を更に備え、

前記第一のノードは、

前記第二のノードから、該第一のノードによって管理されているデータの送信要求を受信する要求受信手段を更に備え、

前記特定識別情報送信手段は、前記送信要求に応じて、要求されたデータに係る前記特定識別情報を前記第二のノードに送信し、

前記データ送信手段は、前記送信要求に応じて、要求されたデータを前記第二のノードに送信する、

請求項１から５の何れか一項に記載のデータベースシステム。

[請求項7] 前記第二のノードは、

ユーザー端末からの検索要求を受け付ける検索要求受付手段と、

前記検索要求への応答を前記ユーザー端末に対して返す検索応答手段と、

を更に備え、

前記要求送信手段は、前記検索要求に応じて、少なくとも前記検索要求に係るデータの送信要求を、前記第一のノードに対して送信し、

前記実行手段は、前記送信要求に応じて前記第一のノードから送信され、前記受信手段によって受信されたデータをメモリに展開し、前記トランザクションログに係る命令を該データに対して実行し、

前記検索応答手段は、前記検索要求の結果前記第一のノードから得られ、前記トランザクションログに係る命令が実行されたデータに基づいて、前記検索要求への応答を前記ユーザー端末に対して返す、

請求項６に記載のデータベースシステム。

[請求項8] 前記第二のノードは、

ユーザー端末からの検索要求を受け付ける検索要求受付手段と、

前記検索要求への応答を前記ユーザー端末に対して返す検索応答手段と、

を更に備え、

前記実行手段は、前記検索要求に応じて、前記受信手段によって受信されたデータをメモリに展開し、前記トランザクションログに係る命令を該データに対して実行し、

前記検索応答手段は、前記検索要求に応じて前記トランザクションログに係る命令が実行されたデータに基づいて、前記検索要求への応答を前記ユーザー端末に対して返す、

請求項１から５の何れか一項に記載のデータベースシステム。

[請求項9] 前記データ送信手段は、前記データベースの管理情報を、前記第二のノードに送信し、

前記要求送信手段は、前記管理情報を参照して、前記第一のノードに対して、前記データベース中のデータを指定して送信要求を送信する、

請求項５に記載のデータベースシステム。

[請求項10] 前記第二のノードは、

前記第一のノードから受信したデータを直接メモリに展開して、データの検索または処理に供する展開手段を更に備え、

前記実行手段は、受信したデータが前記展開手段によってメモリに展開されたことを受けて、前記トランザクションログに係る命令を前記データに対して実行する、

請求項１から９の何れか一項に記載のデータベースシステム。

[請求項11] 前記特定識別情報送信手段は、前記特定識別情報として、ログ保持手段によって保持されている前記識別情報のうち、最新のトランザクションログを示す識別情報を送信する、

請求項１から１０の何れか一項に記載のデータベースシステム。

[請求項12] 前記データは、テーブル単位またはページ単位で送受信される、

請求項１から１１の何れか一項に記載のデータベースシステム。

[請求項13] 前記実行手段は、前記トランザクションログに係る命令を、前記識別情報によって把握される命令の順序に応じて、受信された前記デー

タに対して実行する、

請求項１から１２の何れか一項に記載のデータベースシステム。

[請求項14] データベースを管理する管理手段と、

前記データベースのトランザクションログを、該トランザクションログに係る命令の順序を把握可能な識別情報とともに保持するログ保持手段と、

前記ログ保持手段によって保持されている前記識別情報のうち、所定の時点におけるトランザクションログを示す特定識別情報を、データベースの複製先である他の情報処理装置に送信する特定識別情報送信手段と、

少なくとも前記所定の時点より後の前記トランザクションログおよび前記識別情報を、互いに関連づけて前記他の情報処理装置に送信するログ送信手段と、

前記データベースによって管理されているデータを、前記所定の時点以降に、前記他の情報処理装置に送信するデータ送信手段と、

を備える情報処理装置。

[請求項15] データベースを管理する他の情報処理装置から、該データベースのトランザクションログ、該トランザクションログに係る命令の順序を把握可能な識別情報、前記識別情報のうち所定の時点におけるトランザクションログを示す特定識別情報、および前記データベースによって管理されているデータを受信する受信手段と、

受信された前記データがメモリに展開されてデータの検索または処理に供される状態となった場合に、受信された前記トランザクションログのうち、少なくとも前記特定識別情報が示すトランザクションログよりも新しい前記トランザクションログに係る命令を、メモリに展開された前記データに対して実行する実行手段と、

を備える情報処理装置。

[請求項16] 複数のノードを有するデータベースシステムにおいて、

前記複数のノードのうち、データベースの複製元である第一のノードが、

該第一のノードによって管理されているデータベースのトランザクションログを、該トランザクションログに係る命令の順序を把握可能な識別情報とともに保持するログ保持ステップと、

前記ログ保持ステップで保持された前記識別情報のうち、所定の時点におけるトランザクションログを示す特定識別情報を、前記複数のノードのうち、データベースの複製先である第二のノードに送信する特定識別情報送信ステップと、

少なくとも前記所定の時点より後の前記トランザクションログおよび前記識別情報を、互いに関連づけて前記第二のノードに送信するログ送信ステップと、

前記データベースによって管理されているデータを、前記所定の時点以降に、前記第二のノードに送信するデータ送信ステップと、

を実行し、

前記第二のノードが、

前記第一のノードから、前記トランザクションログ、該トランザクションログの識別情報、前記特定識別情報および前記データを受信する受信ステップと、

受信された前記データが該第二のノードのメモリに展開されてデータの検索または処理に供される状態となった場合に、受信された前記トランザクションログのうち、少なくとも前記特定識別情報が示すトランザクションログよりも新しい前記トランザクションログに係る命令を、メモリに展開された前記データに対して実行する実行ステップと、

を実行する、

方法。

[請求項17] 複数のノードを有するデータベースシステムにおいて、

前記複数のノードのうち、データベースの複製元である第一のノードを、

該第一のノードによって管理されているデータベースのトランザクションログを、該トランザクションログに係る命令の順序を把握可能な識別情報とともに保持するログ保持手段と、

前記ログ保持手段によって保持されている前記識別情報のうち、所定の時点におけるトランザクションログを示す特定識別情報を、前記複数のノードのうち、データベースの複製先である第二のノードに送信する特定識別情報送信手段と、

少なくとも前記所定の時点より後の前記トランザクションログおよび前記識別情報を、互いに関連づけて前記第二のノードに送信するログ送信手段と、

前記データベースによって管理されているデータを、前記所定の時点以降に、前記第二のノードに送信するデータ送信手段と、

として機能させ、

前記第二のノードを、

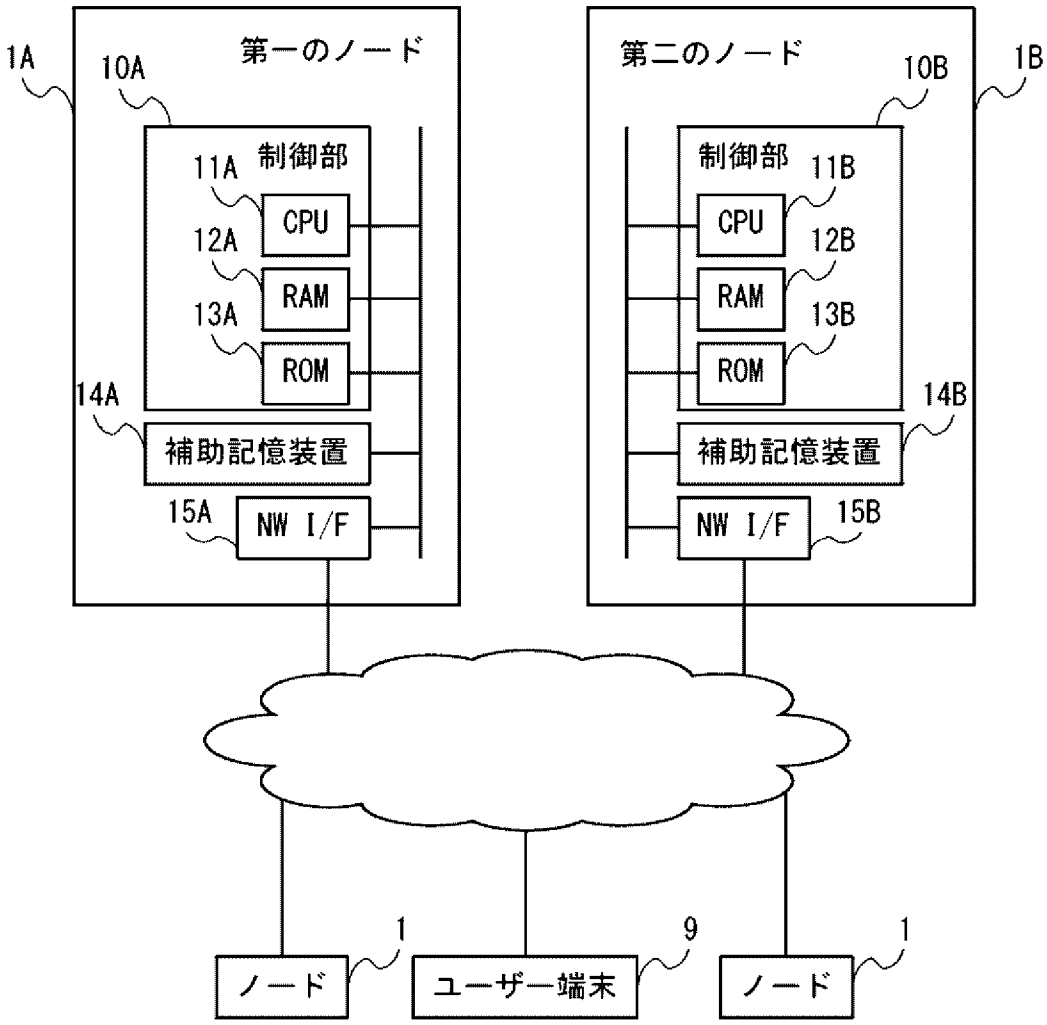
前記第一のノードから、前記トランザクションログ、該トランザクションログの識別情報、前記特定識別情報および前記データを受信する受信手段と、

受信された前記データが該第二のノードのメモリに展開されてデータの検索または処理に供される状態となった場合に、受信された前記トランザクションログのうち、少なくとも前記特定識別情報が示すトランザクションログよりも新しい前記トランザクションログに係る命令を、メモリに展開された前記データに対して実行する実行手段と、

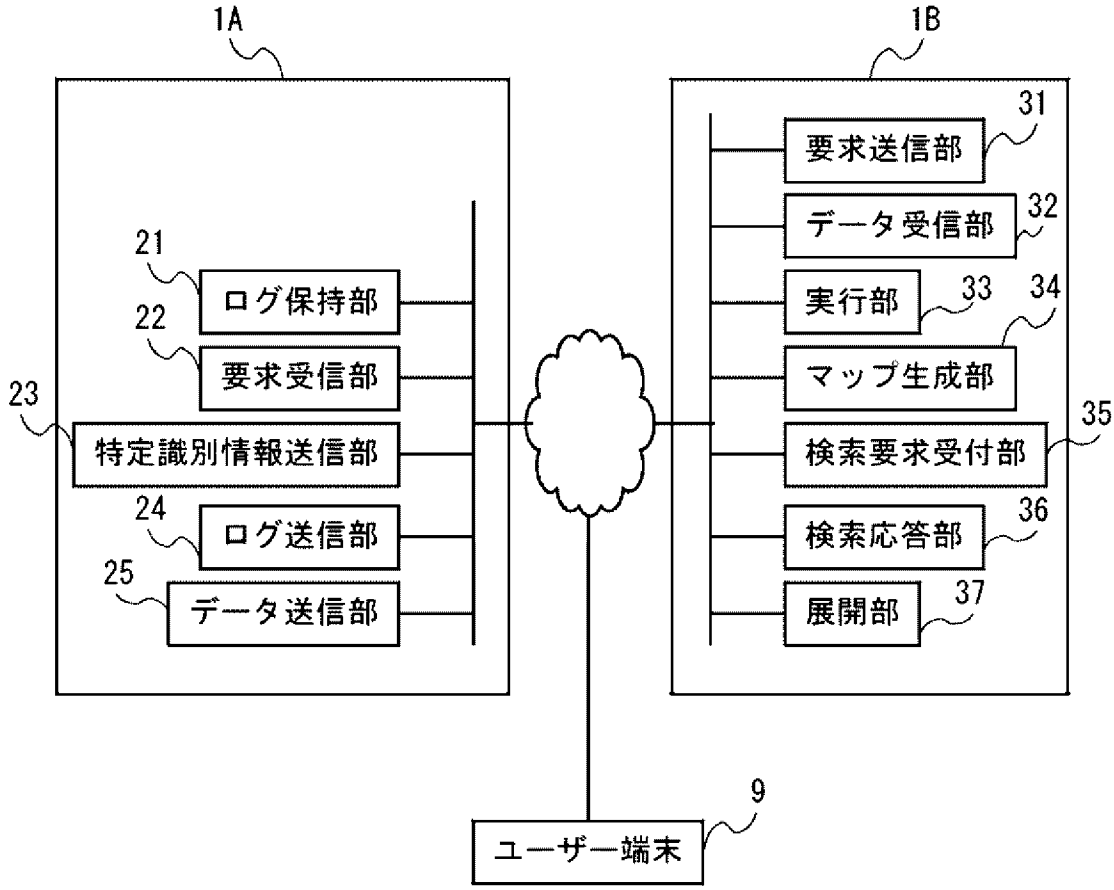
として機能させる、

プログラム。

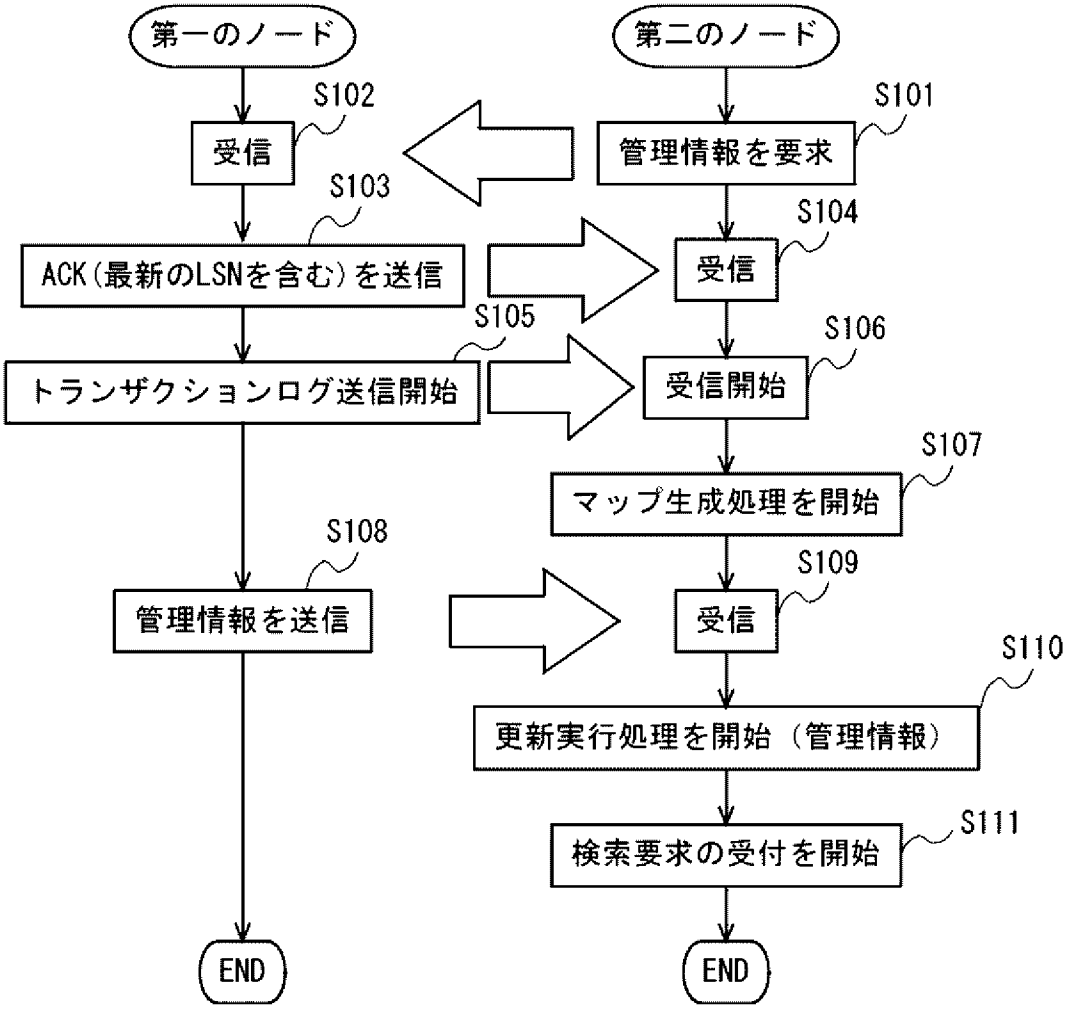
[図1]


1A
10A
第一のノード
制御部
11A
CPU
12A
RAM
13A
ROM
14A
補助記憶装置
15A
NW I/F
第二のノード
10B
1B
制御部
11B
CPU
12B
RAM
13B
ROM
14B
補助記憶装置
NW I/F
15B
1
ノード
9
ユーザー端末
1
ノード

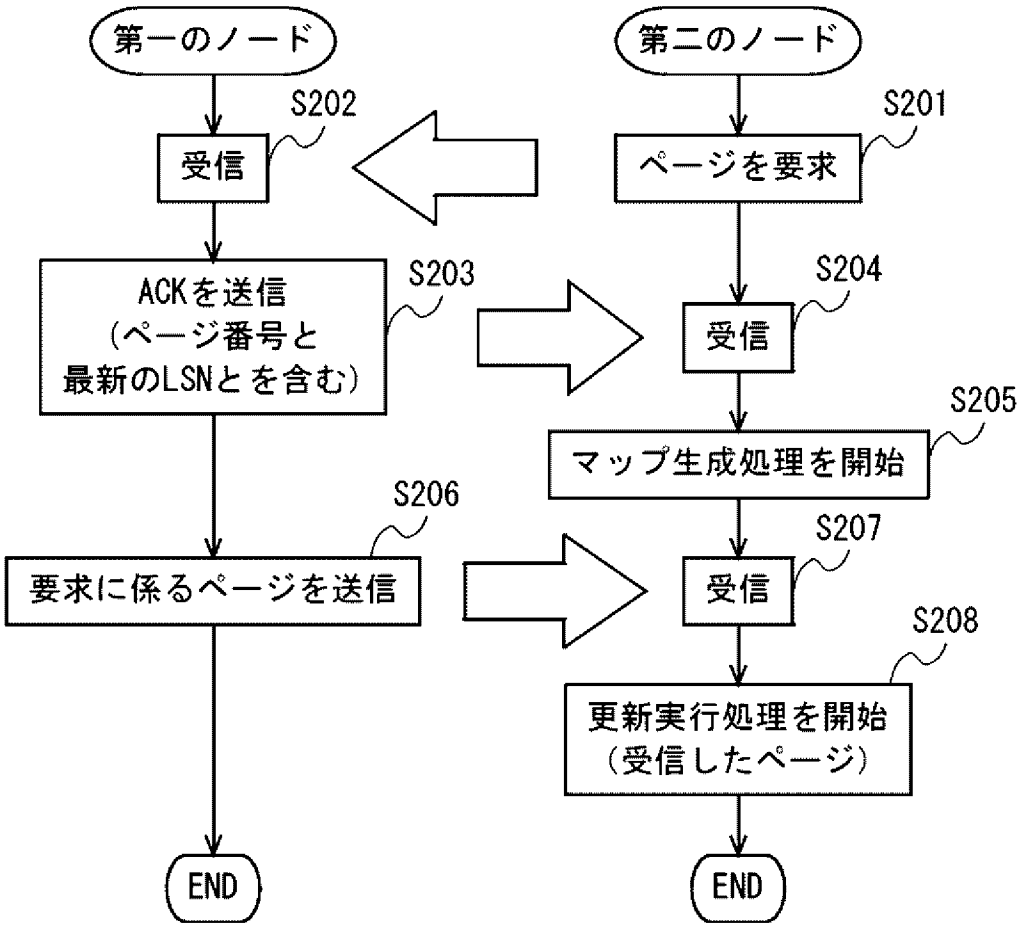
[図2]


1A
1B
21
ログ保持部
22
要求受信部
23
特定識別情報送信部
24
ログ送信部
25
データ送信部
31
要求送信部
32
データ受信部
33
実行部
34
マップ生成部
35
検索要求受付部
36
検索応答部
37
展開部
9
ユーザー端末

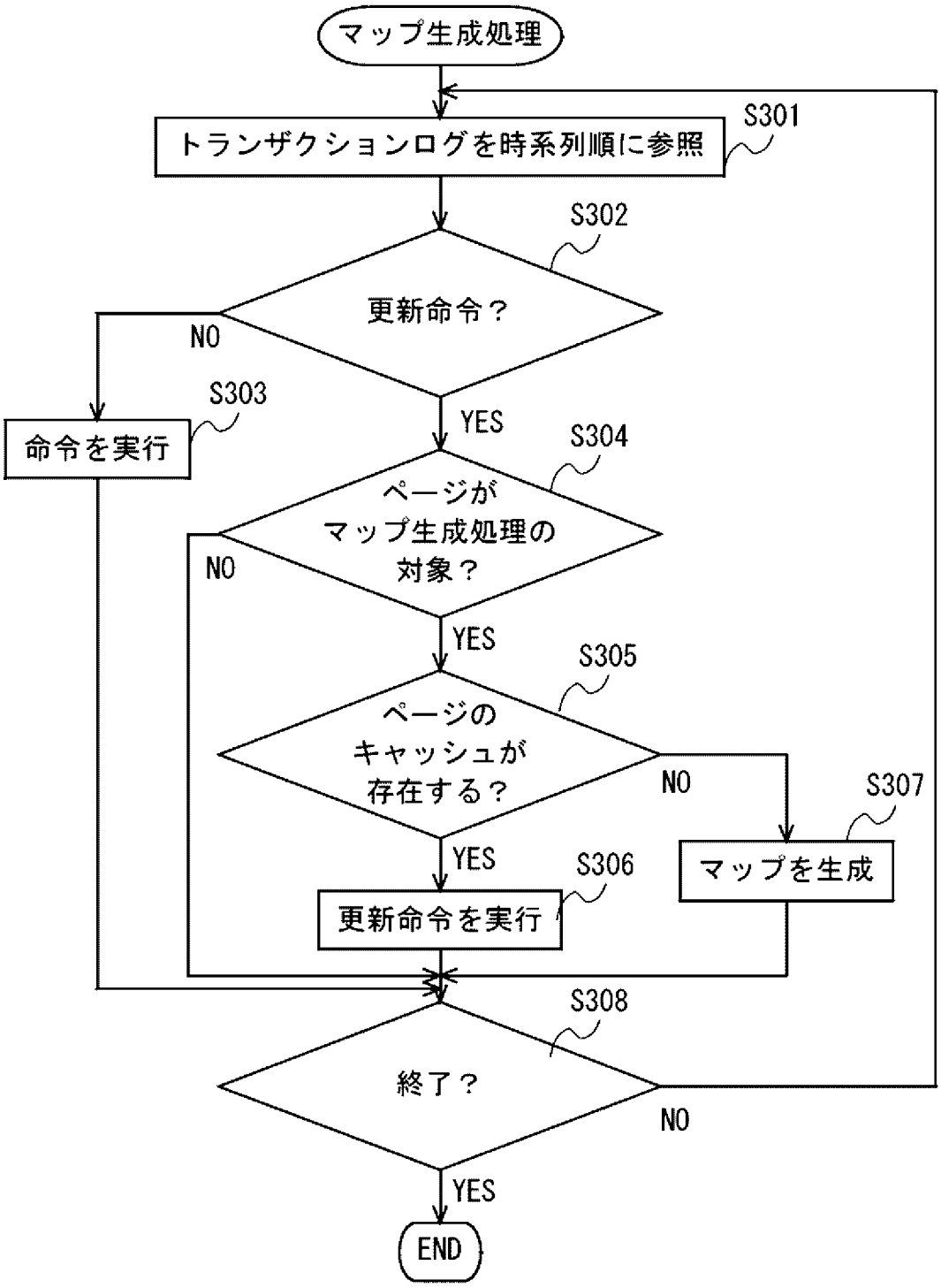
[図3]


第一のノード
第二のノード
S102
受信
S101
管理情報を要求
S103
ACK(最新のLSNを含む)を送信
S104
受信
S105
トランザクションログ送信開始
S106
受信開始
S107
マップ生成処理を開始
S108
管理情報を送信
S109
受信
S110
更新実行処理を開始（管理情報）
S111
検索要求の受付を開始
END
END

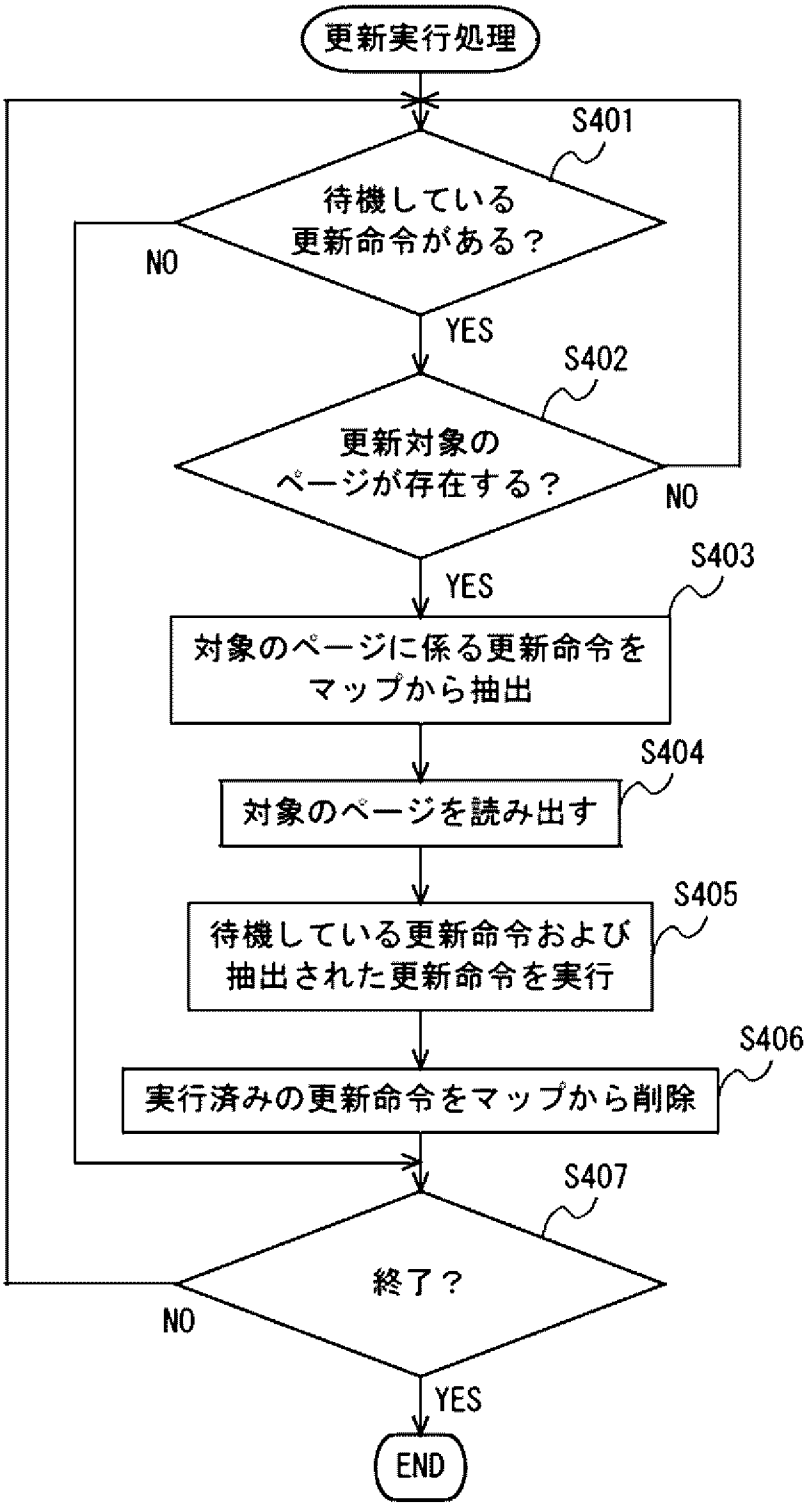
[図4]


第一のノード
第二のノード
S202
S201
受信
ページを要求
ACKを送信
(ページ番号と
最新のLSNとを含む)
S203
S204
受信
S205
マップ生成処理を開始
S206
S207
要求に係るページを送信
受信
S208
更新実行処理を開始
(受信したページ)
END
END

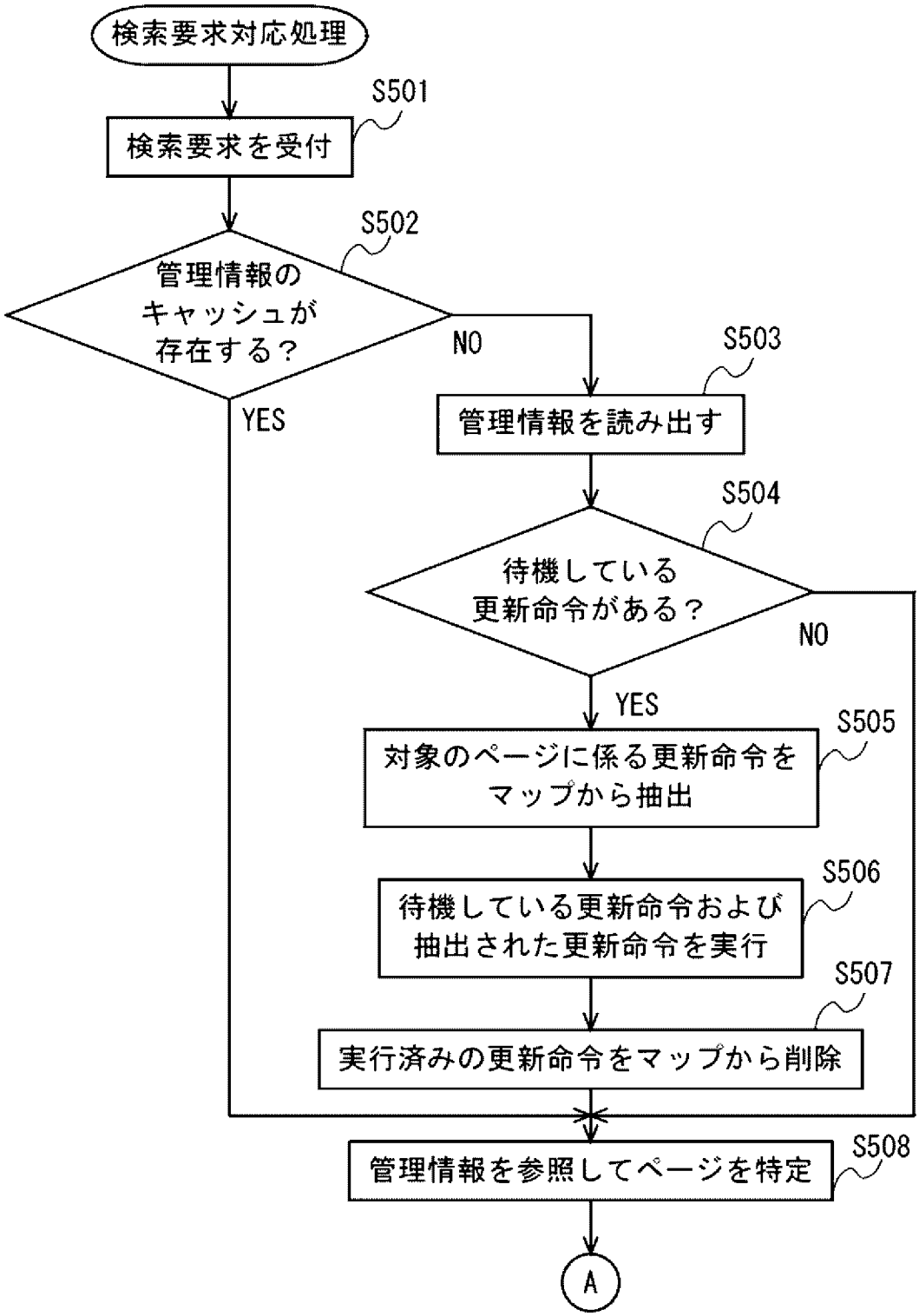
[図5]


マップ生成処理
S301
トランザクションログを時系列順に参照
S302
更新命令？
NO
YES
S303
命令を実行
S304
ページが
マップ生成処理の
対象？
NO
YES
S305
ページの
キャッシュが
存在する？
NO
S307
マップを生成
YES
S306
更新命令を実行
S308
終了？
NO
YES
END

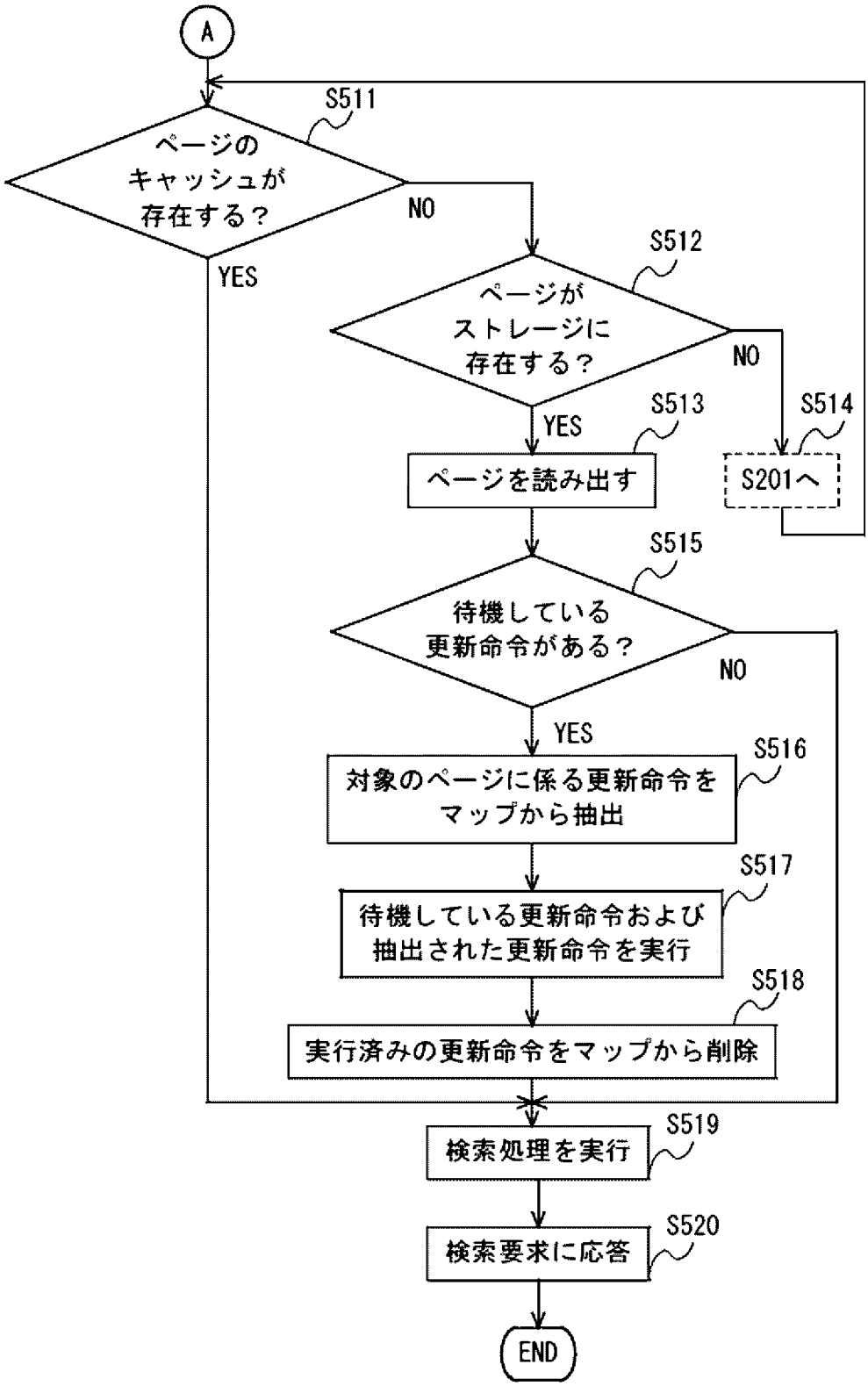
[図6]


更新実行処理
S401
待機している
更新命令がある？
NO
YES
S402
更新対象の
ページが存在する？
NO
YES
S403
対象のページに係る更新命令を
マップから抽出
S404
対象のページを読み出す
S405
待機している更新命令および
抽出された更新命令を実行
S406
実行済みの更新命令をマップから削除
S407
終了？
NO
YES
END

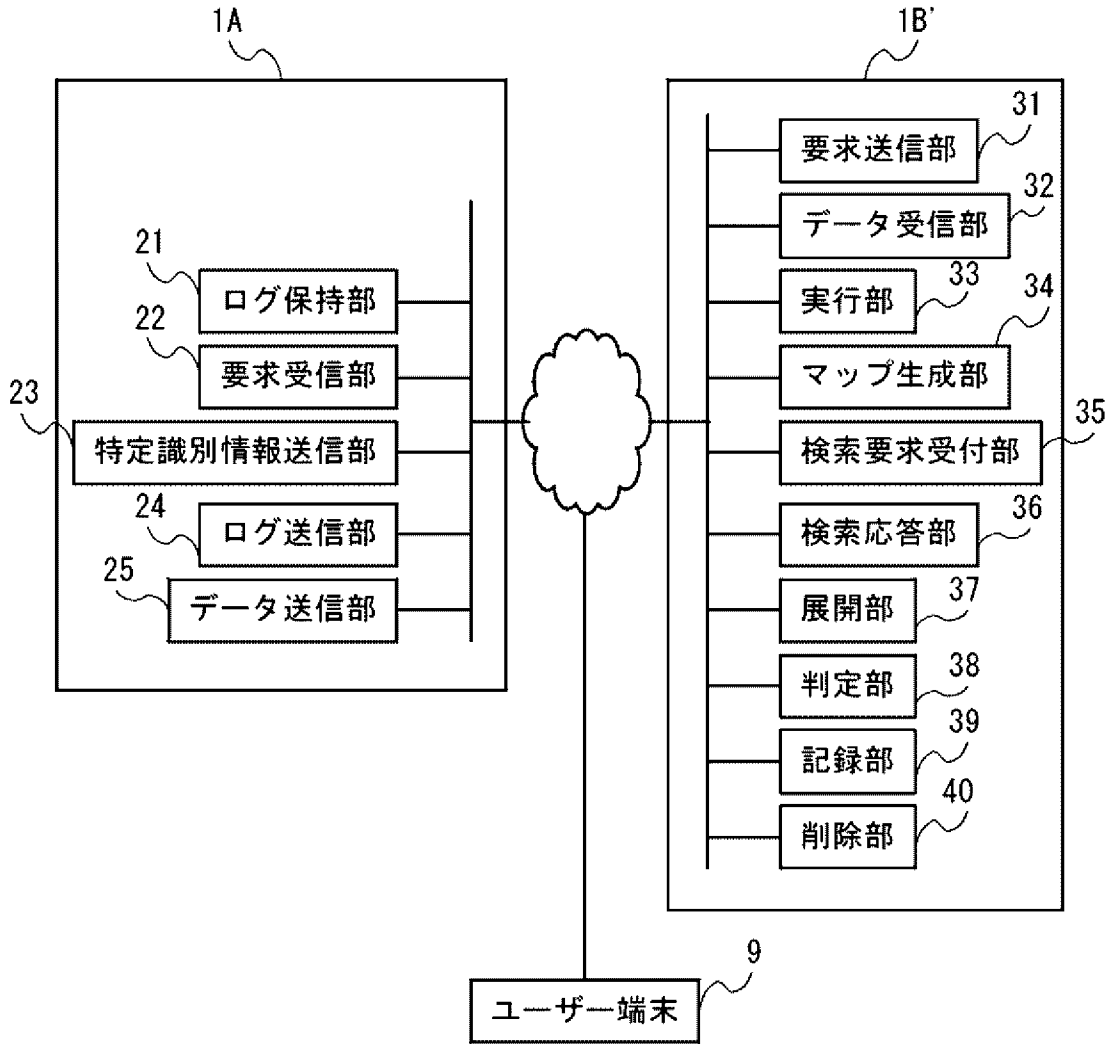
[図7]


検索要求対応処理
S501
検索要求を受付
S502
管理情報の
キャッシュが
存在する？
NO
YES
S503
管理情報を読み出す
S504
待機している
更新命令がある？
NO
YES
S505
対象のページに係る更新命令を
マップから抽出
S506
待機している更新命令および
抽出された更新命令を実行
S507
実行済みの更新命令をマップから削除
S508
管理情報を参照してページを特定
A

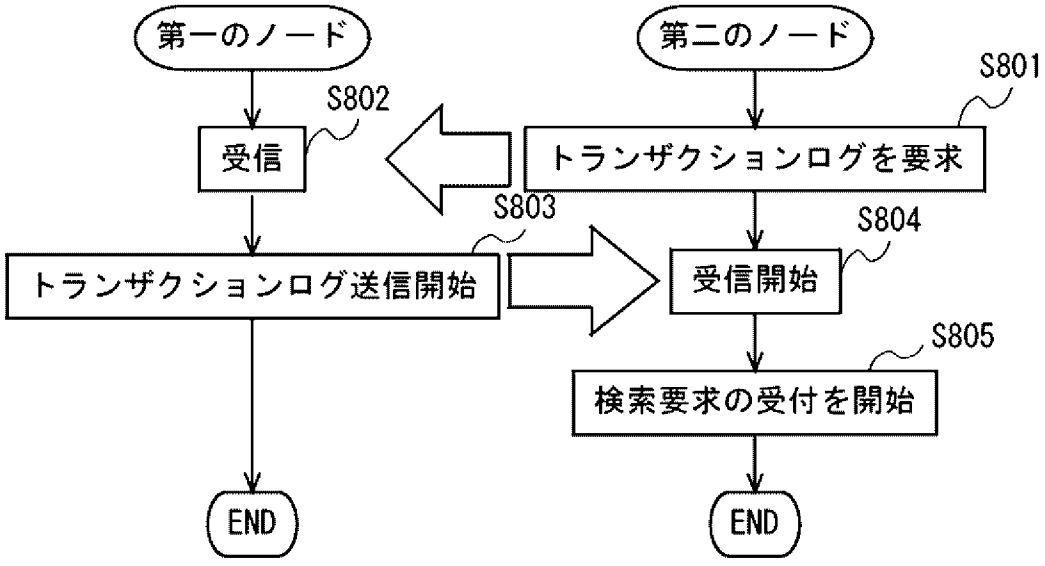
[図8]


A
S511
ページの
キャッシュが
存在する？
NO
YES
S512
ページが
ストレージに
存在する？
NO
YES
S513
ページを読み出す
S514
S201へ
S515
待機している
更新命令がある？
NO
YES
S516
対象のページに係る更新命令を
マップから抽出
S517
待機している更新命令および
抽出された更新命令を実行
S518
実行済みの更新命令をマップから削除
S519
検索処理を実行
S520
検索要求に応答
END

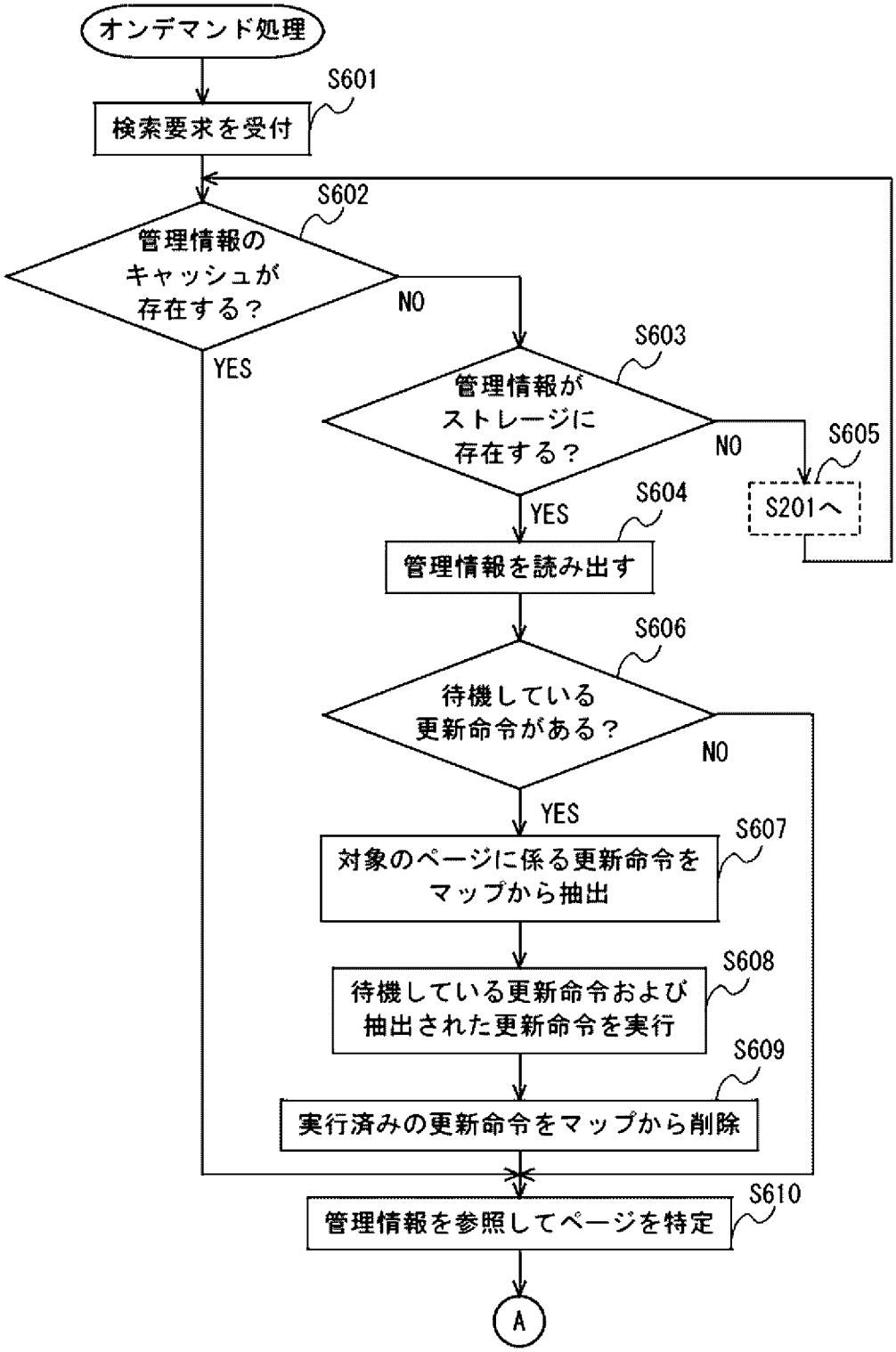
[図9]


1A
1B'
21
ログ保持部
22
要求受信部
23
特定識別情報送信部
24
ログ送信部
25
データ送信部
31
要求送信部
32
データ受信部
33
実行部
34
マップ生成部
35
検索要求受付部
36
検索応答部
37
展開部
38
判定部
39
記録部
40
削除部
9
ユーザー端末

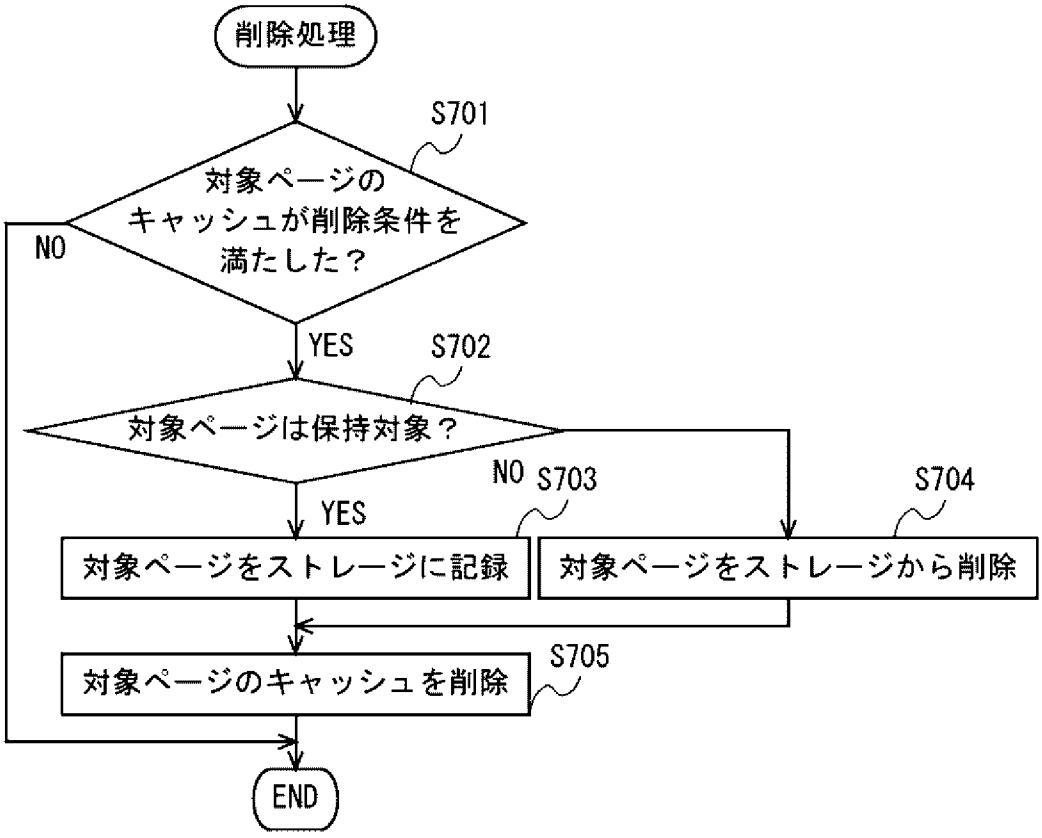
[図10]


第一のノード
第二のノード
S801
S802
受信
トランザクションログを要求
S803
S804
トランザクションログ送信開始
受信開始
S805
検索要求の受付を開始
END
END

[図11]


オンデマンド処理
S601
検索要求を受付
S602
管理情報の
キャッシュが
存在する？
NO
YES
S603
管理情報が
ストレージに
存在する？
NO
S605
S201へ
YES
S604
管理情報を読み出す
S606
待機している
更新命令がある？
NO
YES
S607
対象のページに係る更新命令を
マップから抽出
S608
待機している更新命令および
抽出された更新命令を実行
S609
実行済みの更新命令をマップから削除
S610
管理情報を参照してページを特定
A

[図12]


削除処理
S701
対象ページの
キャッシュが削除条件を
満たした？
NO
YES
S702
対象ページは保持対象？
NO
S703
YES
S704
対象ページをストレージに記録
対象ページをストレージから削除
S705
対象ページのキャッシュを削除
END

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/058381

A. CLASSIFICATION OF SUBJECT MATTER
G06F11/20(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)
G06F11/16-11/20

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)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 10-187520 A (Fujitsu Ltd.), 21 July 1998 (21.07.1998), entire text; all drawings (Family: none)	1-17
A	JP 2004-280337 A (Toshiba Corp.), 07 October 2004 (07.10.2004), paragraphs [0059] to [0071]; fig. 3 to 5 (Family: none)	1-17
A	JP 2013-506892 A (Chicago Mercantile Exchange Inc.), 28 February 2013 (28.02.2013), entire text; all drawings & US 2008/0126833 A1 & EP 2049997 A2 & WO 2008/021713 A2 & CA 2659395 A1	1-17

☒ 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	Date of mailing of the international search report
24 June, 2014 (24.06.14)	08 July, 2014 (08.07.14)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/058381

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 2009-545788 A (TSX Inc.), 24 December 2009 (24.12.2009), entire text; all drawings & US 2008/0126832 A1 & EP 2049999 A1 & WO 2008/014585 A1	1-17
A	US 7464113 B1 (ORACLE INTERNATIONAL CORP.), 09 December 2008 (09.12.2008), entire text; all drawings (Family: none)	1-17
A	JP 2002-024069 A (EMC Corp.), 25 January 2002 (25.01.2002), entire text; all drawings & US 6578160 B1 & GB 2367922 A & DE 10124482 A1	1-17

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１４／０５８３８１

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F11/20(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.　G06F11/16-11/20

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）

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 10-187520 A（富士通株式会社） 1998.07.21，全文，全図（ファミリーなし）	１－１７
A	JP 2004-280337 A（株式会社東芝） 2004.10.07，段落【0059】～【0071】，図３～図５（ファミリーなし）	１－１７
A	JP 2013-506892 A（シカゴ　マーカンタイル　エクスチェンジ，インク.）2013.02.28，全文，全図 & US 2008/0126833 A1 & EP 2049997 A2 & WO 2008/021713 A2 & CA 2659395 A1	１－１７

☑　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の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	国際調査報告の発送日
２４．０６．２０１４	０８．０７．２０１４

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
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）　５Ｂ　３７８１
▲高▼橋　正▲徳▼
電話番号　０３－３５８１－１１０１　内線　３５４４

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 2009-545788 A (ティーエスエックス　インコーポレイテッド) 2009.12.24, 全文, 全図 & US 2008/0126832 A1 & EP 2049999 A1 & WO 2008/014585 A1	１－１７
A	US 7464113 B1 (ORACLE INTERNATIONAL CORPORATION) 2008.12.09 , 全文, 全図（ファミリーなし）	１－１７
A	JP 2002-024069 A (イーエムシー　コーポレイション) 2002.01.25 , 全文, 全図 & US 6578160 B1 & GB 2367922 A & DE 10124482 A1	１－１７