

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4139632号
(P4139632)

(45) 発行日 平成20年8月27日(2008.8.27)

(24) 登録日 平成20年6月13日(2008.6.13)

(51) Int.Cl.

F I

G 0 6 F 12/00 (2006.01)

G 0 6 F 12/00 5 3 5 Z

請求項の数 4 (全 9 頁)

(21) 出願番号 特願2002-188935 (P2002-188935)
 (22) 出願日 平成14年6月28日(2002.6.28)
 (65) 公開番号 特開2004-30487 (P2004-30487A)
 (43) 公開日 平成16年1月29日(2004.1.29)
 審査請求日 平成16年7月13日(2004.7.13)

(73) 特許権者 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (73) 特許権者 000233055
 日立ソフトウェアエンジニアリング株式会
 社
 東京都品川区東品川四丁目12番7号
 (74) 代理人 100075096
 弁理士 作田 康夫
 (72) 発明者 祖敷 潤一郎
 神奈川県横浜市中区尾上町六丁目81番地
 日立ソフトウェアエンジニアリング株式
 会社内

審査官 廣瀬 文雄

最終頁に続く

(54) 【発明の名称】 データベース管理方法およびシステム

(57) 【特許請求の範囲】

【請求項1】

データベースを格納するための記憶手段と備え、実行された複数の検索プロセスからの検索要求の入力に応じて、該検索要求に関連するデータを前記データベースから読み出して一時的に格納するとともに前記複数の検索プロセスの実行により共有されるデータベースバッファを備え、各前記検索プロセスは、当該データベースバッファに格納されたデータから前記問い合わせに関連するデータを格納する前記検索プロセスにそれぞれ対応付けられ、他の検索プロセスとは共有しない固有バッファを備えたデータベース管理システムにおけるデータ管理方法において、

前記検索プロセスからの入力した前記検索要求でページへの2回目以降アクセスの場合、前記固有バッファを参照して次のレコードの取り出しを試み、この結果、前記次のレコードが固有バッファに存在する場合は、次のレコードを検索結果に設定して検索要求に返し、前記固有バッファに検索結果を返すレコードが存在する間は、前記固有バッファの内容を参照しレコードを取り出し、検索結果の設定を繰返し、前記固有バッファ上の全てのレコードを読み終えたら、コピー中断のレコード情報の存在をチェックし、情報がなければ次の検索対象となるページの存在をチェックし、なければ、検索を終了し、前記ページの存在があれば、検索対象となるページを設定し、検索を続行し、

T E S Tモードの排他制御で排他待ちを検知した場合、該当レコードは、前記固有バッファにコピーを行わず、前記固有バッファへのレコードのコピーをそこで一時中断し、中断したレコードの情報は固有バッファ内のコピー中断のレコード情報に記憶しておき、

10

20

取得中の前記DBバッファを解放し、前記固有バッファのレコードの取出しを行い、前記固有バッファにレコードが存在すれば、取り出したレコードを検索結果に設定して検索要求として返し、前記コピー中断のレコード情報が設定されていなければ、次の検索対象となるページが、存在するかチェックし、存在しなければ検索を終了し、存在すれば次ぎの検索対象ページの検索を行い、コピー中断のレコード情報をチェックした結果、前記コピー中断のレコード情報が設定されていれば、コピー中断のレコードに対してWAITモードで排他を掛け、他のトランザクションの排他解除まで排他待ちし、排他待ちが解除されれば、再度前記固有バッファの初期化を行い、当該DBバッファのアクセスを行い、コピー中断のレコードを検知した場合、該当レコードは既に検索要求の条件を満たし、排他済の状態にあるため該当レコードを前記固有バッファにコピーすることを特徴とするデータベース管理方法。

10

【請求項2】

さらに、コピー中断したレコードの後にレコードが存在すれば、レコードの参照、条件の判定、排他制御を行い、ページ内のレコードを前記固有バッファにコピーのループを繰り返して、ページ内のコピーが完了し、前記固有バッファのレコードを取り出すときは、コピー中断したレコードから取り出すためにコピーを行った順に取り出し、ページ内のレコードのコピーが全て完了してループを抜けている場合は、前記固有バッファのレコードを全て検索結果に返した後、次の検索対象ページのアクセス、または、検索終了となり、コピー中断によってループを抜けた場合は、再度、WAITモードの排他で排他待ちして、前記固有バッファの初期化から処理を繰り返すことを特徴とする請求項1記載のデータベース管理方法。

20

【請求項3】

データベースを格納するための記憶手段と備え、実行された複数の検索プロセスからの検索要求の入力に応じて、該検索要求に関連するデータを前記データベースから読み出して一時的に格納するとともに前記複数の検索プロセスの実行により共有されるデータベースバッファを備え、各前記検索プロセスは、当該データベースバッファに格納されたデータから前記問い合わせに関連するデータを格納する前記検索プロセスにそれぞれ対応付けられ、他の検索プロセスとは共有しない固有バッファを備えたデータベース管理システムにおいて、

前記検索プロセスからの入力した前記検索要求でページへの2回目以降アクセスの場合、前記固有バッファを参照して次のレコードの取り出しを試み、この結果、前記次のレコードが固有バッファに存在する場合は、次のレコードを検索結果に設定して検索要求に返し、前記固有バッファに検索結果を返すレコードが存在する間は、前記固有バッファの内容を参照しレコードを取り出し、検索結果の設定を繰り返して、前記固有バッファ上の全てのレコードを読み終えたら、コピー中断のレコード情報の存在をチェックし、情報がなければ次の検索対象となるページの存在をチェックし、なければ、検索を終了し、前記ページの存在があれば、検索対象となるページを設定し、検索を続行する手段と

30

TESTモードの排他制御で排他待ちを検知した場合、該当レコードは、前記固有バッファにコピーを行わず、前記固有バッファへのレコードのコピーをそこで一時中断し、中断したレコードの情報は固有バッファ内のコピー中断のレコード情報に記憶しておき、取得中の前記DBバッファを解放し、前記固有バッファのレコードの取出しを行い、前記固有バッファにレコードが存在すれば、取り出したレコードを検索結果に設定して検索要求として返し、前記コピー中断のレコード情報が設定されていなければ、次の検索対象となるページが、存在するかチェックし、存在しなければ検索を終了し、存在すれば次ぎの検索対象ページの検索を行い、コピー中断のレコード情報をチェックした結果、前記コピー中断のレコード情報が設定されていれば、コピー中断のレコードに対してWAITモードで排他を掛け、他のトランザクションの排他解除まで排他待ちし、排他待ちが解除されれば、再度前記固有バッファの初期化を行い、当該DBバッファのアクセスを行い、コピー中断のレコードを検知した場合、該当レコードは既に検索要求の条件を満たし、排他済の状態にあるため該当レコードを前記固有バッファにコピーする手段とを備えたことを特

40

50

徴とするデータベース管理システム。

【請求項 4】

さらに、コピー中断したレコードの後にレコードが存在すれば、レコードの参照、条件の判定、排他制御を行い、ページ内のレコードを前記固有バッファにコピーのループを繰り返して、ページ内のコピーが完了し、前記固有バッファのレコードを取り出すときは、コピー中断したレコードから取り出すためにコピーを行った順に取り出し、ページ内のレコードのコピーが全て完了してループを抜けている場合は、前記固有バッファのレコードを全て検索結果に返した後、次の検索対象ページのアクセス、または、検索終了となり、コピー中断によってループを抜けた場合は、再度、W A I T モードの排他で排他待ちして、前記固有バッファの初期化から処理を繰り返す手段を備えたことを特徴とする請求項 3 記載のデータベース管理システム。

10

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数のプログラムもしくはプロセッサでデータ操作処理する場合のデータベース管理方法およびシステムに関するものである。

【0002】

【従来の技術】

データベース管理システムにおいて、更新要求や検索要求による D B の参照、更新を行う場合、データベース管理システムで共有する D B バッファを用いてディスクの内容を参照する。更新要求や検索要求によって参照するページが、D B バッファ上に存在する場合、D B バッファの内容を読むことでディスクの I / O 回数を軽減する一般的な手段である。しかし、

20

D B バッファは、データベース管理システムで共有し、その内容を他のプロセスからの D B バッファ参照要求に対して整合性を確保するためシリアルライズする手段をとる。

【0003】

例えば、検索要求における D B バッファの参照要求で、D B 内の要求ページを D B バッファに載せる場合、D B バッファの内容や管理情報を更新するため、D B バッファ全体の整合性を確保するまでの間、D B バッファ全体に対してロック、アンロックを行い、他の D B バッファの参照要求との整合性を確保している。更に、1 回の検索要求におけるレコードの走査では、少なくとも、1 回の D B バッファにアクセスを行い、探索条件を満たす 1 件のレコードを取り出している。その結果、大量にレコードを取り出す検索要求では、大量の D B バッファのアクセスが発生している。

30

【0004】

【発明が解決しようとする課題】

こうした D B バッファアクセスのシリアルライズ手法は、大量に並列実行されることで競合が発生し、D B アクセスの性能に影響を及ぼす。検索要求のたびにシリアルライズを伴う D B バッファアクセスを大量に並列実行した場合、競合が発生し著しく D B の検索性能に影響を及ぼす。

【0005】

本発明は検索要求における並列実行での競合を回避し、性能を得ることを目的としている。

40

【0006】

【課題を解決するための手段】

上記目的を達成するために、検索要求のプロセス単位に検索要求に該当する D B バッファの内容を載せるローカライズされた固有バッファを設け、検索要求に該当する D B バッファ上のページ内のレコードのうち、検索要求の条件を満たす全てのレコードを 1 回の D B バッファアクセスで、固有バッファにコピーし参照する。次の検索要求が同一ページに対するものであれば、固有バッファを参照することによって、D B バッファの参照を行わずに検索要求に該当するレコードの参照を可能にする。更に、検索要求プロセスの固有バ

50

ッファに載せたレコードが、他の更新要求の並列実行によってレコードの内容が書き替えられた場合、固有バッファの内容とDBバッファの内容が不整合な状態になる。これをページ内のレコードに対する排他制御によって検知する。固有バッファ上に検索要求を満たすレコードを載せ、そのレコードに対して排他制御することによって、そのレコードが更新されたか否かを検知する。ただし、排他制御で更新を検知したレコードはその時点で排他待ちを行うのではなく、まず、固有バッファ上に既にコピーしたレコードを検索結果として返し、検索結果を返すレコードが、排他制御によって更新を検知したレコードになった時点で改めて排他待ちする。排他待ち解除の後、もう一度、該当レコードからDBバッファの内容を固有バッファに載せ、DBバッファと固有バッファの内容の整合性を保つ。

【0007】

10

【発明の実施の形態】

以下、本発明を実施する場合の一形態を図面を参照して具体的に説明する。

図1は本発明を実施する場合の情報検索装置を示したブロック図である。

情報検索装置は、各検索プロセス102ごとにローカライズされた固有バッファ103を割当て、検索要求101を受けて固有バッファ103を初期化し、DBバッファ104の内容を走査し、該当するページ内の検索要求の条件を満たす全てのレコードを固有バッファ103上にコピーし、更に、固有バッファ103上を走査し検索要求101に検索結果を返す検索装置と、各プロセスからの要求に該当するページの内容をデータベース管理システムで共用するDBバッファ104と、DBの内容をファイル格納するディスク装置105からなる。

20

固有バッファ103には、固有バッファ内の領域管理、レコード位置の管理、DBバッファのページの該当レコードのコピーを中断したレコード情報を管理する管理情報を有するものである。データベースはディスク装置上に設けられた複数のページ（格納領域）に分割されて格納されている。ページ内のデータがアクセスされた場合、DBバッファへは、そのアクセスされたデータを含むページが格納されることになる。

【0008】

図2は検索中のページに対して他の更新処理によって更新されたレコードが存在するときに、検索プロセスの固有バッファの内容とDBバッファのページの内容を示したブロック図である。検索対象レコードの他のトランザクションによる更新検知は、該当レコードの排他制御によって行う。排他制御には同一排他資源への排他要求に対し、排他待ちを検知できるモード（以下TESTモード）を使用し、排他待ちの検知によって該当レコードの更新を検知する。DBバッファのページ201の検索対象レコードを固有バッファ202にコピー中に、レコードの更新を検知したら（205）、レコードのコピーは一時中断し、それまでにコピーした固有バッファ内のレコードの検索を行う（203）。固有バッファ内の全てのレコードの検索が完了したら、中断したレコード204に対して同一排他資源への排他要求に対して排他待ちを行う排他モード（以下WAITモード）で排他制御を行う。該当レコードの排他待ちが解除されたら、固有バッファの初期化206を行い、中断したレコード204から固有バッファ202へのレコードのコピーを実施し、後続の検索対象のレコードのコピーを行う。DBバッファのページ内の検索対象のレコードを全てコピーが完了したら、固有バッファのレコードをコピーした順に取出し検索結果を返す（207）。本例では、レコード単位の処理で説明しているが、レコード内の特定カラムのデータでも同様であり、複数カラムのデータでも適用可能であり、レコードに限定されるものではない。

30

40

【0009】

図3は検索中のページに対して他の更新処理によって更新されたレコードが存在するときに、検索プロセスの固有バッファの内容とDBバッファの整合性を維持する流れの一例を示した図である。図2は、検索要求の探索条件を満たす3件のレコード格納されているページがDBバッファ上に存在している状態で、検索を実行したときに、2件目にヒットするレコードR2が別の更新トランザクションによって、内容を変更されることを示している。

50

【 0 0 1 0 】

3 0 1 は更新トランザクションの期間と更新処理内容を、3 0 2 は更新トランザクションが掛けるページ内のレコード R 2 に掛ける排他期間を示している。3 0 3 は検索トランザクションの期間と検索処理内容を、3 0 4、3 0 5、3 0 6 はそれぞれ検索トランザクションが掛けるページ内のレコード R 1、レコード R 2、レコード R 3 の排他期間を示している。検索トランザクションは、固有バッファとコピー中断のレコード情報の割当てと初期化を行う (3 1 0 , 3 1 1)。次にページ 3 1 2 の内容を参照しレコード R 1 に T E S T モード排他を掛け、レコード R 1 を固有バッファにコピーする (3 1 4)。検索トランザクションがレコード R 2 を参照する前に、更新トランザクションがレコード R 2 に対して排他を掛け (3 0 2)、レコード R 2 の内容を更新する (3 1 6)。検索トランザクションは、レコード R 2 を参照し、T E S T モードで排他を掛けるが、排他待ちを検知する (3 0 7)。排他待ちを検知したレコード R 2 の情報は、コピー中断のレコード情報に登録し (3 1 5)、固有バッファへのコピーを一時中断する。排他が掛かっている D B バッファのページの排他を解放する (3 2 8)。

10

【 0 0 1 1 】

次に、固有バッファ内の検索を行う。固有バッファにコピーしたレコード R 1 を検索結果として検索要求に返す (3 1 7)。次の検索要求からは、固有バッファ内の検索を行う。固有バッファにコピーした、全てのレコードを検索結果として検索要求に返したら、コピー中断のレコード情報が登録されているか判定する。コピー中断のレコード情報にはレコード R 2 が登録されているので、レコード R 2 に排他制御を行う。このとき、コピー中断のレコードには、W A I T モードで排他待ちさせる (3 0 8)。排他が掛かっている D B バッファのページの排他が解放される (3 2 9)。レコード R 2 の排他待ちが更新トランザクションのコミットで解除されたら (3 0 9)、固有バッファの初期化 (3 1 9) とコピー中断のレコード情報の初期化 (3 1 9) を行い、中断したレコード R 2 から固有バッファへレコードコピーを再開する (3 2 0 , 3 2 1)。レコード R 3 に排他を掛け、固有バッファにコピーが完了 (3 2 2 , 3 2 3) し、D B バッファのページ内の全てのレコードをコピーする。排他が掛かっている D B バッファのページの排他を解放する (3 3 0)。次に、固有バッファ内を検索しレコード R 2 を検索結果として検索要求に返す (2 2 4)。次の検索要求では、次のレコード R 3 を検索結果に返し (2 2 5)、固有バッファと、コピー中断のレコード情報を解放 (2 2 6 , 2 2 7) して検索を終了する。このように、D B バッファから検索してヒットしたレコードを固有バッファにコピーする処理をまとめて行い、D B バッファを解放して検索結果を返しているので、D B バッファの占有時間を短縮することができる。

20

30

【 0 0 1 2 】

図 2 , 図 3 が示すように、ページの内容を 1 回の D B バッファアクセスで、可能な限り固有バッファに複数レコードをコピーし、検索する。しかし、途中で他のトランザクションによるレコードの更新を検知したら、そのレコードが検索要求の対象になるまでは、排他待ちせず、固有バッファ内に貯めたレコードを先に返し、固有バッファ内の全てのレコードの検索が完了した後で排他待ちして整合性を保持する。

【 0 0 1 3 】

図 4 は検索要求を受けた情報検索装置の処理手順を示すフロ - チャ - トである。検索要求でページへのアクセスが 1 回目のアクセスなら (ステップ 4 0 1)、固有バッファの初期化を行う (ステップ 4 0 2)。次に検索要求に該当する D B バッファ内の検索要求ページにアクセスし (ステップ 4 0 3)、D B バッファ上のページの内容を参照しレコードがコピー中断のレコードか判断する (ステップ 4 0 5)。該当レコードがコピー中断のレコードではない場合、該当レコード検索要求の条件を満たすか判断する (ステップ 4 0 6)。該当レコードが検索要求の条件を満たすのであれば、そのレコードに対し排他制御を T E S T モードで行い (ステップ 4 0 7)、T E S T モード排他の結果 (ステップ 4 0 8) 排他待ちを検知しなければ、該当レコードを固有バッファにコピーを行う (ステップ 4 0 9)。また、該当レコードが検索要求の条件を満たさないのであれば、該当レコードは固有

40

50

バッファ上にコピーしない。これを該当ページ内に存在するの全てのレコードに対して行う(ステップ410)。これらの処理を該当ページ内の全てのレコードに対して行った後、該当ページの解放を行う(ステップ412)。次に固有バッファ内のレコードの取出しを行う(ステップ414)。固有バッファのレコードの取出しは固有バッファにコピーした順番に取り出す。レコード取出しの結果を判定し(ステップ414)、固有バッファ内に取り出すレコードが存在すれば、取り出したレコードを検索結果に設定しを検索要求に返す(ステップ417)。

【0014】

2回目の検索要求でページへの2回目以降アクセスになる場合(ステップ401)では、固有バッファを参照し、次のレコードの取出しを行う(ステップ413)。この結果、次のレコードが固有バッファに存在する場合(ステップ414)は、次のレコードを検索結果に設定して検索要求に返す(ステップ417)。固有バッファに検索結果を返すレコードが存在する間は、固有バッファの内容を参照しレコードを取り出し、検索結果の設定を繰返し返す。固有バッファ上の全てのレコードを読み終えたら(ステップ414)、コピー中断のレコード情報の存在をチェックし(ステップ415)、情報がなければ次の検索対象となるページの存在をチェックし(ステップ418)、なければ、検索を終了し、あれば、検索対象となるページを設定しステップ401判定から検索を続行する。こうして、検索要求に該当するレコードの参照におけるDBバッファへのアクセス要求は、ページに格納されているレコードの数に依存することなく、1ページあたり1回になる。また、ステップ407、ステップ408のTESTモードの排他制御で排他待ちを検知した場合、該当レコードは、固有バッファにコピーを行わず、固有バッファへのレコードのコピーはそこで一時中断する。このとき、中断したレコードの情報は固有バッファ内のコピー中断のレコード情報に覚えておき(ステップ411)、取得中のDBバッファを解放し、固有バッファのレコードの取出しを行う(ステップ414)。ここで、固有バッファにレコードが存在をチェックし(ステップ414)存在すれば、取り出したレコードを検索結果に設定して検索要求に返す。

【0015】

固有バッファにレコードが存在をチェックし(ステップ414)、存在しなければ、コピー中断のレコード情報をチェックする(ステップ415)。この結果、コピー中断のレコード情報が設定されていなければ、次の検索対象となるページが、存在するかチェックし(ステップ418)、存在しなければ検索を終了、存在すれば次ぎの検索対象ページの検索を行う。コピー中断のレコード情報のチェック(ステップ415)の結果、情報が設定されていれば、コピー中断のレコードに対してWAITモードで排他を掛け(ステップ416)、他のトランザクションの排他解除まで排他待ちになる。排他待ちが解除されれば、再度固有バッファの初期化を行い(ステップ402)以下DBバッファのアクセスを行う。このとき、ステップ405の判定においてコピー中断のレコードを検知した場合、該当レコードは既に検索要求の条件を満たし、排他済の状態にあるため、ステップ406からステップ408の処理は行わず、該当レコードは固有バッファにコピー(ステップ409)を行う。

【0016】

さらに、コピー中断したレコードの後にレコードが存在すれば、レコードの参照、条件の判定、排他制御を行い、ページ内のレコードを固有バッファにコピーのループを繰り返す。ページ内のコピーが完了し、固有バッファのレコードを取り出すときは、コピーを行った順に取り出す。すなわち、コピー中断したレコードから取り出すことになる。ページ内のレコードのコピーが全て完了してループを抜けている場合は、固有バッファのレコードを全て検索結果に返した後、次の検索対象ページのアクセス、または、検索終了となるが、コピー中断によってループを抜けた場合は、再度、WAITモードの排他で排他待ちして、ステップ402の固有バッファの初期化から処理を繰り返す。

【0017】

このように、排他制御でレコードの更新中を検知し、レコードに対する排他解除待ちする

10

20

30

40

50

ことで、固有バッファの整合性を保つことができる。また、ロックがかかっていないレコードをまとめて処理できるため、DBバッファのアクセスに伴うデータベースが格納されたディスクへのアクセスを減らすことが可能となる。

【 0 0 1 8 】

【発明の効果】

以上、説明したように、本発明によれば、検索要求の並行実行時のDBバッファの競合が軽減され性能向上を実現できる。

【図面の簡単な説明】

【図 1】本発明の実施の一形態の情報検索装置の構成を示すブロック図、

【図 2】検索中のページに対して他の更新処理によって更新されたレコードが存在した場合のレコードコピー方法を示したブロック図、

【図 3】検索中のページに対して他の更新処理によって更新されたレコードが存在するときに、検索プロセスの固有バッファの内容とDBバッファの整合性を維持する流れの一例を示した図

【図 4】検索要求を受けた情報検索装置の処理手順を示すフローチャート、

【符号の説明】

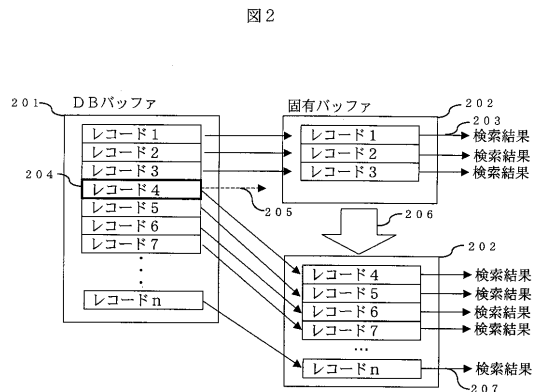
1 0 1・・・検索要求、1 0 2・・・検索プロセス、1 0 3・・・固有バッファ、1 0 4・・・DBバッファ、1 0 5・・・DISK装置、2 0 1・・・DBバッファ、2 0 2・・・固有バッファ、2 0 3・・・固有バッファ内の検索、2 0 4・・・更新を検知したレコード、2 0 5・・・レコードコピーの中断、2 0 6・・・固有バッファの初期化、2 0 7・・・固有バッファ内の再検索、3 0 1・・・更新トランザクションの期間、3 0 2・・・更新トランザクションが掛けるR 2の排他期間、3 0 3・・・検索トランザクションの期間、3 0 4・・・検索トランザクションが掛けるR 1の排他期間、3 0 5・・・検索トランザクションが掛けるR 2の排他期間、3 0 6・・・検索トランザクションが掛けるR 3の排他期間、3 0 7・・・排他待ちを検知しコピー中断のレコード情報の登録、3 0 8・・・R 2の排他で排他待ちが発生している、3 0 9・・・更新トランザクションの排他解除でR 2の排他待ちが解除され検索トランザクションがR 2の排他を掛ける、3 1 0、3 1 5、3 1 9、3 2 6・・・固有バッファのコピー中断レコード情報の内容、3 1 1、3 1 4、4 1 7、3 1 8、3 2 1、3 2 3、3 2 7・・・固有バッファのレコードの内容、3 1 2、3 1 6、3 2 0、3 2 2・・・DBバッファの検索対象となるページの内容、3 2 4、3 2 5・・・固有バッファからレコードの取出し

10

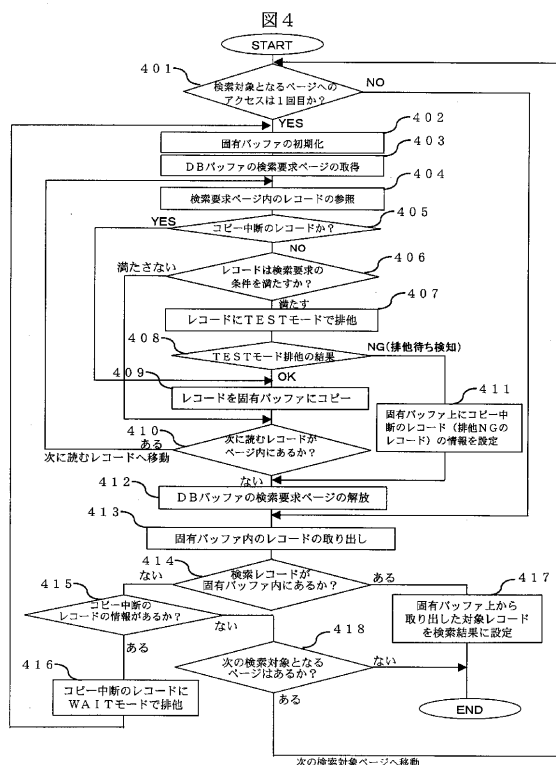
20

30

【圖 2】



【 図 4 】



フロントページの続き

(56)参考文献 特開平 0 7 - 2 6 2 0 7 1 (J P , A)
特開平 0 2 - 2 5 3 4 4 9 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)
G06F 12/00