

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5722352号
(P5722352)

(45) 発行日 平成27年5月20日(2015.5.20)

(24) 登録日 平成27年4月3日(2015.4.3)

(51) Int.Cl.

F I

G 0 6 F 12/00 (2006.01)

G 0 6 F 12/00 5 1 4 K

請求項の数 15 (全 18 頁)

(21) 出願番号	特願2012-556228 (P2012-556228)	(73) 特許権者	500046438
(86) (22) 出願日	平成23年3月2日(2011.3.2)		マイクロソフト コーポレーション
(65) 公表番号	特表2013-521579 (P2013-521579A)		アメリカ合衆国 ワシントン州 9805
(43) 公表日	平成25年6月10日(2013.6.10)		2-6399 レッドモンド ワン マイ
(86) 国際出願番号	PCT/US2011/026930		クロソフト ウェイ
(87) 国際公開番号	W02011/109564	(74) 代理人	100140109
(87) 国際公開日	平成23年9月9日(2011.9.9)		弁理士 小野 新次郎
審査請求日	平成26年2月28日(2014.2.28)	(74) 代理人	100075270
(31) 優先権主張番号	12/717, 139		弁理士 小林 泰
(32) 優先日	平成22年3月4日(2010.3.4)	(74) 代理人	100101373
(33) 優先権主張国	米国 (US)		弁理士 竹内 茂雄
		(74) 代理人	100118902
			弁理士 山本 修
		(74) 代理人	100153028
			弁理士 上田 忠

最終頁に続く

(54) 【発明の名称】 データベースサーバのためのバッファプール拡張

(57) 【特許請求の範囲】

【請求項 1】

第1のメモリ、前記第1のメモリよりも低速である第2のメモリ、および前記第2のメモリよりも低速である第3のメモリを含んだコンピュータによって少なくとも部分的に実装される方法であって、

バッファプールの複数のページに関するメタデータにアクセスするステップであって、前記バッファプールは、前記第1のメモリに記憶される1つまたは複数のページ、および前記第2のメモリに記憶される1つまたは複数のページを含む、ステップと、

前記メタデータから第1の閾値と第2の閾値を判定するステップであって、前記第1の閾値と前記第2の閾値は、前記ページを少なくとも3つの分類のうちの1つに分類するのに使用可能であり、

(i) 前記ページの前記メタデータによって表されるデータにアクセスする関数を適用して、前記第1の閾値より小さい値が返される場合、ページは第1の分類に分類され、

(i i) 前記ページの前記メタデータによって表される前記データにアクセスする前記関数を適用して、前記第1の閾値より大きい値が前記第2の閾値より小さい値が返される場合、前記ページは第2の分類に分類され、

(i i i) 前記ページの前記メタデータによって表される前記データにアクセスする前記関数を適用して、前記第2の閾値より大きい値が返される場合、前記ページは第3の分類に分類される、

ステップと、

10

20

前記第 1 および第 2 の閾値に基づいて、

特定のページが前記第 1 の分類に分類される場合、前記特定のページを前記第 2 のメモリから前記第 3 のメモリへ退避させて、前記第 2 のメモリの 1 つまたは複数の他のページのためにスペースを解放するステップと、

前記特定のページが前記第 2 の分類に分類される場合、前記特定のページを前記第 1 のメモリから前記第 2 のメモリへ退避させて、前記第 1 のメモリの 1 つまたは複数の他のページのためにスペースを解放するステップと、

前記特定のページが前記第 3 の分類に分類される場合、前記特定のページを前記第 1 のメモリに残すステップと、

を含むことを特徴とする方法。

10

【請求項 2】

前記バッファプールの前記複数のページに関する前記メタデータにアクセスするステップは、前記バッファプールのランダムに選択されたいくつかのページのメタデータをサンプリングするステップを含み、その数は、前記バッファプールの全てのページより小さいことを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記バッファプールの前記複数のページに関する前記メタデータにアクセスするステップは、前記メタデータから、サンプリングされたそれぞれのページの 1 つまたは 2 つのタイムスタンプを取得するステップを含み、前記 1 つまたは 2 つのタイムスタンプは、前記サンプリングされたページがアクセスされた直近の時間に対応し、前記関数は、少なくとも前記 1 つまたは 2 つのタイムスタンプに対して作用して、前記ページが前記第 1 の分類、前記第 2 の分類、または前記第 3 の分類であるかどうかを示す値を生成することを特徴とする請求項 2 に記載の方法。

20

【請求項 4】

前記第 1 の閾値を判定するステップは、ランダムに選択されたページの予め定められたパーセンタイルがコールドであるような値を判定するステップを含むことを特徴とする請求項 1 に記載の方法。

【請求項 5】

前記第 2 の閾値を判定するステップは、前記バッファプールのランダムに選択されたいくつかのページのメタデータをサンプリングするステップを含み、前記ランダムに選択されたページは前記第 1 のメモリ内に存在することを特徴とする請求項 1 に記載の方法。

30

【請求項 6】

前記第 2 の閾値を判定するステップは、前記ランダムに選択されたページの予め定められたパーセンタイルがウォームであるような値を判定するステップを含むことを特徴とする請求項 5 に記載の方法。

【請求項 7】

前記特定のページを前記第 1 のメモリから前記第 2 のメモリへ退避させるステップは、(i) 前記第 2 のメモリが利用可能なフリースペースを有する場合、および (i i) 前記特定のページのコピーが前記第 2 のメモリ内に無い場合、にのみ、前記特定のページを前記第 2 のメモリにコピーするステップを含むことを特徴とする請求項 1 に記載の方法。

40

【請求項 8】

前記バッファプールのために使用される前記第 2 のメモリの量を変更するステップをさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 9】

前記第 2 のメモリの I / O 閾値に到達したかどうかを判定するステップと、

到達した場合には、前記特定のページを前記第 1 のメモリおよび前記第 2 のメモリ以外のメモリにコピーするステップと、

をさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 10】

前記特定のページを前記第 1 のメモリから前記第 2 のメモリへ退避させるステップは、

50

複数のページを前記第 1 のメモリから前記第 2 のメモリへ 1 回の書き込み操作でコピーするステップを含むことを特徴とする請求項 1 に記載の方法。

【請求項 1 1】

コンピュータ環境におけるシステムであって、

1 つまたは複数のプロセッサと、

第 1 のメモリ、前記第 1 のメモリよりも低速である第 2 のメモリ、および前記第 2 のメモリよりも低速である第 3 のメモリと、

コンピュータ実行可能命令が格納された 1 つまたは複数のコンピュータ可読記憶装置と、を備え、

前記コンピュータ実行可能命令は、前記 1 つまたは複数のプロセッサによって実行されると、

バッファプールの複数のページに関するメタデータにアクセスするステップであって、前記バッファプールは、前記第 1 のメモリに記憶される 1 つまたは複数のページ、および前記第 2 のメモリに記憶される 1 つまたは複数のページを含む、ステップと、

前記メタデータから第 1 の閾値と第 2 の閾値を判定するステップであって、前記第 1 の閾値と前記第 2 の閾値は、前記ページを少なくとも 3 つの分類のうちの 1 つに分類するのに使用可能であり、

(i) 前記ページの前記メタデータによって表されるデータにアクセスする関数を適用して、前記第 1 の閾値より小さい値が返される場合、ページは第 1 の分類に分類され、

(i i) 前記ページの前記メタデータによって表される前記データにアクセスする前記関数を適用して、前記第 1 の閾値より大きい値が前記第 2 の閾値より小さい値が返される場合、前記ページは第 2 の分類に分類され、

(i i i) 前記ページの前記メタデータによって表される前記データにアクセスする前記関数を適用して、前記第 2 の閾値より大きい値が返される場合、前記ページは第 3 の分類に分類される、

ステップと、

前記第 1 および第 2 の閾値に基づいて、

特定のページが前記第 1 の分類に分類される場合、前記特定のページを前記第 2 のメモリから前記第 3 のメモリへ退避させて、前記第 2 のメモリの 1 つまたは複数の他のページのためにスペースを解放するステップと、

前記特定のページが前記第 2 の分類に分類される場合、前記特定のページを前記第 1 のメモリから前記第 2 のメモリへ退避させて、前記第 1 のメモリの 1 つまたは複数の他のページのためにスペースを解放するステップと、

前記特定のページが前記第 3 の分類に分類される場合、前記特定のページを前記第 1 のメモリに残すステップと、

を含む動作を実施することを特徴とするシステム。

【請求項 1 2】

前記第 3 のメモリが 1 つまたは複数のハードディスクを備え、前記第 1 のメモリがランダムアクセスメモリを備え、前記第 2 のメモリがソリッドステートメモリを備えることを特徴とする請求項 1 1 に記載のシステム。

【請求項 1 3】

前記第 1 の分類はアクセス頻度の低いページに適用され、前記第 3 の分類はアクセス頻度の高いページに適用され、前記第 2 の分類は前記第 1 の分類に分類されたページよりもアクセス頻度が高いが前記第 3 の分類に分類されたページよりもアクセス頻度が低いページに適用されることを特徴とする請求項 1 に記載の方法。

【請求項 1 4】

前記第 1 の分類はコールドの分類であり、前記第 2 の分類はウォームの分類であり、前記第 3 の分類はホットの分類であることを特徴とする請求項 1 に記載の方法。

【請求項 1 5】

前記第 1 の分類は最も古くアクセスされたページに適用され、前記第 3 の分類は最も新

しくアクセスされたページに適用され、前記第2の分類は前記第1の分類に分類されたページよりも新しくアクセスされたが前記第3の分類に分類されたページよりも古くアクセスされたページに適用されることを特徴とする請求項1に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データベースシステムに関する。

【背景技術】

【0002】

データベースでは、データベースページの作業用セット(working set)がメインメモリに保持されている時には、良好な性能を発揮する。残念ながら、大きなデータベースの多くがサイズの上で、作業用セットをメインメモリに保持することは実行可能ではない。大きなデータベースでは、これは、ほとんどの場合ディスク等の回転媒体(rotational media)を使用して構築される十分な容量を持つI/Oサブシステムに、データベースのデータページの大部分が存在するということを意味する。そのようなサブシステムは高価であり、かなりの量の空間を占め、かつかなりの量の電力を消費することになる。これらのサブシステムは、回転媒体の性能がメインメモリおよびプロセッサと同じ速さでは進歩していないため、データベースのボトルネックとなることが多い。

【0003】

本明細書において特許請求される主題は、任意の問題を解決する実施形態または上述のような環境においてのみ動作する実施形態に限定されない。むしろ、この背景技術は、本明細書において記載されるいくつかの実施形態を實踐することができる技術分野の一例を例示するために単に提供される。

【発明の概要】

【0004】

簡潔には、本明細書に記載される主題の態様は、データベースシステムのためのバッファプールに関する。態様においては、ソリッドステート記憶装置等の二次的メモリを使用して、データベースシステムのバッファプールを拡張する。ページのアクセス履歴に基づきページを分類するためのホット、ウォーム、コールド等の閾値を、サンプリングアルゴリズムを介して判定することができる。データベースシステムがメインメモリ内のバッファプールでスペースをフリーにする必要がある場合、ページを分類する方法および二次的メモリまたは他の記憶装置の状態に基づき、二次的メモリまたは他の記憶装置内のバッファプールに、ページを退避させることができる。

【0005】

この要約は、以下の詳細な説明でさらに述べる主題のいくつかの態様を簡単に確認するために提供される。この要約は、特許請求の主題の重要な特徴または主要な特徴を確認することを意図しておらず、特許請求の主題の範囲を制限するものとして使用されることも意図していない。

【0006】

「本明細書に記載される主題」という句は、明白に文脈が別段に示さない限り、詳細な説明において記載される主題を指す。用語「態様」は、「少なくとも1つの態様」と読まれるべきである。詳細な説明に記載される主題の態様を確認することは、特許請求される主題の重要な特徴または主要な特徴を確認することを意図していない。

【0007】

上述の態様および本明細書に記載される主題の他の態様は例として示され、同様の数字によって同様の要素が示される添付の図面において限定されない。

【図面の簡単な説明】

【0008】

【図1】本明細書に記載される主題の態様を組み込むことができる例示の汎用コンピュータ環境を表すブロック図である。

10

20

30

40

50

【図 2】本明細書に記載される主題の態様を実装することができる例示のシステムを概略的に表すブロック図である。

【図 3】本明細書に記載される主題の態様に従ってデータベースをホストするシステムの構成要素を例示するブロック図である。

【図 4】本明細書に記載される主題の態様に従って、閾値を判定するためにページをスキャンし、必要に応じてページを退避させる際に発生しうる、いくつかの例示の動作を概略的に表すフロー図である。

【図 5】DBMS がアクセス要求を受け取り、かつバッファプールがフルである場合に発生するようないくつかの例示の動作を概略的に表すフロー図である。

【発明を実施するための形態】

10

【0009】

定義

本明細書で使用される時、用語「含む (includes)」およびその変化形は、「を含むがこれに限定されない (includes, but in not limited to)」を意味する制約の無い用語として読まれるものである。用語「または」は、明白に文脈が別段に示さない限り、「および/または」として読まれるものである。用語「に基づき (based on)」は、「少なくとも部分的に基づき (based at least in part on)」として読まれるものである。用語「一実施形態」および「ある実施形態」は、「少なくとも 1 つの実施形態」として読まれるものである。用語「別の実施形態」は、「少なくとも 1 つの他の実施形態」として読まれるものである。他の定義は、明示的なものも暗黙的なものも、以下に含まれる。

20

【0010】

例示の動作環境

図 1 は、本明細書に記載される主題の態様を実装することができる適切なコンピュータシステム環境 100 の例を例示する。コンピュータシステム環境 100 は適切なコンピュータ環境の単なる一例であり、本明細書に記載される主題の態様の使用または機能性の範囲について何ら制限を示唆することは意図されない。また、コンピュータ環境 100 は、例示の動作環境 100 内に例示される構成要素のうちの任意の 1 つまたは組み合わせに関するいずれの依存関係も要件も有するものとして解釈されるべきではない。

【0011】

本明細書に記載される主題の態様は、多数の他の汎用もしくは専用のコンピュータシステム環境または構成で動作可能である。本明細書に記載される主題の態様と共に使用するのに適した周知のコンピュータシステム、環境、または構成の例には、パーソナルコンピュータ、サーバコンピュータ、ハンドヘルドまたはラップトップ装置、マイクロプロセッサシステム、マイクロコントローラベースのシステム、セットトップボックス、プログラム可能家庭用電化製品、ネットワーク PC、ミニコンピュータ、メインフレームコンピュータ、携帯情報端末 (PDA)、ゲーム機、プリンタ、セットトップやメディアセンタを含む電化製品または他の電化製品、自動車に埋め込むまたは取り付けられるコンピュータ装置、他の自動車用装置、上記のシステムまたは装置のいずれかを含む分散コンピュータ環境、などが含まれる。

30

【0012】

本明細書に記載される主題の態様は、プログラムモジュール等のコンピュータにより実行可能なコンピュータ実行可能命令の一般的文脈で記載することができる。一般的に、プログラムモジュールには、ルーチン、プログラム、オブジェクト、コンポーネント、データ構造、などが含まれ、これらは特定のタスクを実行または特定の抽象データ型を実装する。本明細書に記載される主題の態様はまた、通信ネットワークを介してリンクされるリモート処理装置によりタスクが実行される分散コンピュータ環境において実践することができる。分散コンピュータ環境において、プログラムモジュールは、メモリ記憶装置を含むローカルおよびリモートのコンピュータ記憶媒体のどちらにも配置することができる。

40

【0013】

図 1 を参照すると、本明細書に記載される主題の態様を実装するための例示のシステム

50

には、コンピュータ 110 の形式で汎用コンピュータ装置が含まれる。コンピュータには、命令を実行することが可能な任意の電子デバイスを含むことができる。コンピュータ 110 の構成要素には、プロセッシングユニット 120、システムメモリ 130、および、システムメモリを含む種々のシステム構成要素をプロセッシングユニット 120 に連結させるシステムバス 121 を含むことができる。システムバス 121 は、メモリバスまたはメモリコントローラ、周辺機器用バス、および、様々なバスアーキテクチャの内のいずれかを使用するローカルバス、を含むいくつかのタイプのバス構造の内のいずれかとすることができる。制限ではなく例として、そのようなアーキテクチャには、ISA (Industry Standard Architecture) バス、MCA (Micro Channel Architecture) バス、EISA (Enhanced ISA) バス、VESA (Video Electronics Standards Association) ローカルバス、メザンバスとしても既知の PCI (Peripheral Component Interconnect) バス、PCI-X (Peripheral Component Interconnect Extended) バス、AGP (Advanced Graphics Port)、および PCIe (PCI Express)、が含まれる。

10

【0014】

コンピュータ 110 には典型的には、様々なコンピュータ可読媒体が含まれる。コンピュータ可読媒体は、コンピュータ 110 によりアクセス可能であり、揮発性および不揮発性の媒体、ならびに、着脱可能および着脱不可能な媒体を含む、任意の利用可能な媒体とすることができる。制限ではなく例として、コンピュータ可読媒体には、コンピュータ記憶媒体および通信媒体を含むことができる。

20

【0015】

コンピュータ記憶媒体には、コンピュータ可読命令、データ構造、プログラムモジュール、または他のデータなどの情報を記憶するための任意の方法または技術において実装される、揮発性および不揮発性、着脱可能および着脱不可能な媒体の両方が含まれる。コンピュータ記憶媒体には、RAM、ROM、EEPROM、フラッシュメモリもしくは他のメモリ技術、CD-ROM、DVD (digital versatile disc) もしくは他の光ディスク記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置もしくは他の磁気記憶装置、または、所望の情報を記憶するために使用可能かつコンピュータ 110 によりアクセス可能な任意の他の媒体が含まれる。

30

【0016】

通信媒体は典型的には、コンピュータ可読命令、データ構造、プログラムモジュール、または、搬送波もしくは他の移送機構などの変調データ信号内の他のデータを具現化し、任意の情報配信媒体を含む。用語「変調データ信号」は、信号内の情報を符号化するような様式で設定または変更された信号の特徴の内の 1 つまたは複数を有する信号を意味する。制限ではなく例として、通信媒体には、有線ネットワークまたは直接配線の接続などの有線媒体、ならびに、音響、無線周波数、赤外線および他の無線媒体などの無線媒体が含まれる。上記の内のいずれかの組み合わせも、コンピュータ可読媒体の範囲内に含まれるべきである。

40

【0017】

システムメモリ 130 には、ROM (read only memory) 131 および RAM (random access memory) 132 などの揮発性および / または不揮発性のメモリの形式のコンピュータ記憶媒体が含まれる。基本入出力システム 133 (BIOS) は、起動中などにコンピュータ 110 内の要素間での情報の伝達を支援する基本ルーチンを含み、典型的には、ROM 131 に格納される。RAM 132 は典型的には、プロセッシングユニット 120 が直ちにアクセス可能および / またはプロセッシングユニット 120 により現在操作されている、データおよび / またはプログラムモジュールを含む。制限ではなく例として、図 1 では、オペレーティングシステム 134、アプリケーションプログラム 135、他のプログラムモジュール 136 およびプログラムデータ 1

50

37を例示する。

【0018】

コンピュータ110はまた、他の着脱可能/着脱不可能、揮発性/不揮発性コンピュータ記憶媒体を含むことができる。単なる例として、図1は、着脱不可能、不揮発性磁気媒体に対して読み書きするハードディスクドライブ141、着脱可能、不揮発性磁気ディスク152に対して読み書きする磁気ディスクドライブ151、および、CD-ROMまたは他の光媒体などの着脱可能、不揮発性光ディスク156に対して読み書きする光ディスクドライブ155を例示する。例示の動作環境において使用可能な他の着脱可能/着脱不可能、揮発性/不揮発性コンピュータ記憶媒体には、磁気テープカセット、フラッシュメモリカード、DVD、他の光ディスク、デジタルビデオテープ、ソリッドステートRAM、ソリッドステートROMなどが含まれる。ハードディスクドライブ141は典型的には、インターフェース140などの着脱不可能メモリインターフェースを介してシステムバス121に対して接続され、磁気ディスクドライブ151および光ディスクドライブ155は典型的には、インターフェース150などの着脱可能メモリインターフェースによりシステムバス121に対して接続される。

10

【0019】

ドライブおよびその関連するコンピュータ記憶媒体は、上記で検討され図1に例示されたが、コンピュータ可読命令、データ構造、プログラムモジュール、およびコンピュータ110のための他のデータの記憶装置を提供する。図1において、例えば、ハードディスクドライブ141が、オペレーティングシステム144、アプリケーションプログラム145、他のプログラムモジュール146およびプログラムデータ147を記憶するものとして例示される。なお、これらの構成要素は、オペレーティングシステム134、アプリケーションプログラム135、他のプログラムモジュール136およびプログラムデータ137、と同じでも良く異なっても良い。オペレーティングシステム144、アプリケーションプログラム145、他のプログラムモジュール146およびプログラムデータ147には、本明細書においては異なる番号が与えられて、少なくともそれらが異なるものであることを例示する。

20

【0020】

ユーザは、キーボード162、および、一般的にはマウス、トラックボールまたはタッチパッドと呼ばれるポインティングデバイス161などの入力装置を介してコマンドおよび情報をコンピュータ110に入力する。他の入力装置(図示せず)には、マイク、ジョイスティック、ゲームパッド、パラボラアンテナ、スキャナ、タッチスクリーン、ライティングタブレット、などを含むことができる。これらおよび他の入力装置は、システムバスに連結されるユーザ入力インターフェース160を介してプロセッシングユニット120に対して接続されることが多いが、パラレルポート、ゲームポートまたはUSB(universal serial bus)などの他のインターフェースおよびバス構造により接続されても良い。

30

【0021】

モニタ191または他のタイプのディスプレイ装置も、ビデオインターフェース190などのインターフェースを介してシステムバス121に対して接続される。モニタに加えて、コンピュータは、スピーカ197およびプリンタ196などの他の周辺出力装置を含むこともでき、これらは、出力周辺インターフェース195を介して接続することができる。

40

【0022】

コンピュータ110は、リモートコンピュータ180などの1つまたは複数のリモートコンピュータに対する論理接続を使用して、ネットワーク環境において動作することができる。リモートコンピュータ180は、パーソナルコンピュータ、サーバ、ルータ、ネットワークPC、ピア装置または他の一般的なネットワークノードとすることができ、また、典型的には、コンピュータ110に関して上述した要素の多くまたは全てを含むが、図1にはメモリ記憶装置181のみが例示されている。図1に示す論理接続には、ローカル

50

エリアネットワーク（LAN）１７１およびワイドエリアネットワーク（WAN）１７３が含まれるが、他のネットワークを含むこともできる。そのようなネットワーク環境は、事務所、企業規模のコンピュータネットワーク、イントラネットおよびインターネットにおいて一般的なものである。

【００２３】

LANネットワーク環境において使用される場合、コンピュータ１１０はネットワークインターフェースまたはアダプタ１７０を介してLAN１７１に対して接続される。WANネットワーク環境において使用される場合、コンピュータ１１０には、モデム１７２、またはインターネット等のWAN１７３上で通信を確立するための他の手段を含むことができる。モデム１７２は、内蔵でも外付けも良く、ユーザ入力インターフェース１６０または他の適切な機構を介してシステムバス１２１に対して接続することができる。ネットワーク環境において、コンピュータ１１０に関して示されるプログラムモジュールまたはその一部は、リモートメモリ記憶装置に記憶させることができる。制限ではなく例として、図１では、メモリ装置１８１上に常駐するものとしてリモートアプリケーションプログラム１８５が例示される。図示されるネットワーク通信は例であり、コンピュータ間の通信リンクを確立するための他の手段を使用することができることは理解されるであろう。

【００２４】

データベースバッファプール

上述のように、I/Oサブシステムはデータベースのボトルネックになることが多い。図２は、本明細書に記載される主題の態様を実装することができる例示のシステムを概略的に表すブロック図である。システム２００には、１つまたは複数のプロセッサ２０２、データベース管理システム（DBMS）２０５、回転媒体（rotational media）２１０、２１１を含むことができ、また、他の構成要素を含んでも良い。DBMS２０５はバッファプール２１５内のページを管理することができる。バッファプール２１５を、メインメモリ２２０およびソリッドステート記憶装置２２５（以下では、SSS２２５と称することもある）内に保持することができる。

【００２５】

システム２００は、コンピュータ１１０などの１つまたは複数のコンピュータを使用して実装することができ、また、１つまたは複数のデータベースを実装することができる。システムのプロセッサ２０２は、図１のプロセッシングユニット１２０に対応し、同じコンピュータ上に備えても良く、または、複数のコンピュータに分散させても良い。プロセッサ２０２は、DBMS２０５に対応する命令を実行してデータベースを提供する。

【００２６】

データベースには、リレーショナルデータベース、オブジェクト指向データベース、階層型データベース、ネットワークデータベース、他のタイプのデータベース、上記の組み合わせまたは拡張、などを含むことができる。データベースに記憶されるデータは、テーブル、レコード、オブジェクト、他のデータ構造などの形式で組織化することができる。データベースに記憶されるデータは、専用データベースファイル、専用ハードドライブパーティション、HTMLファイル、XMLファイル、スプレッドシート、フラットファイル、文書ファイル、構成ファイル、他のファイルなどに記憶させることができる。

【００２７】

データベース内のデータには、DBMS２０５などのDBMSを介してアクセスすることができる。DBMS２０５は、データベースのデータの組織、記憶、管理および検索を管理する１つまたは複数のプログラムを含むことができる。DBMS２０５は、要求を受け取りデータベース内のデータにアクセスすることができ、また、このアクセスに必要とされる動作を実行することができる。本明細書で使用される時、アクセスには、データの読み込み、データの書き込み、データの削除、データのアップデート、これらの内の２つ以上の組み合わせなどが含まれる。

【００２８】

本明細書に記載される主題の態様の説明において、便宜上、リレーショナルデータベー

10

20

30

40

50

スに関連する専門用語を本明細書においては使用することがある。リレーショナルデータベースの専門用語を本明細書において使用することがあるが、本明細書の技術は上述したものを含む他のタイプのデータベースにも適用することができる。

【0029】

DBMS 205は、メインメモリおよび二次的メモリのバッファプール内のページを管理するべく、かつ、ページが少なくとも3つの分類に対応するアクセスを有するかどうかの判定に基づきページを退避させるべく、動作可能である。これについては以下でより詳細に説明されるが、簡略的には、ページへのアクセスがパーセンタイルの第1の範囲内にある場合、そのページが第1の分類（例えば「ホット」）にあると考えられる。ページがパーセンタイルの第2の範囲内にアクセスを有する場合、そのページが第2の分類（例えば「ウォーム」）にあると考えられる。ページがパーセンタイルの第3の範囲内にアクセスを有する場合、そのページが第3の分類（例えば「コールド」）にあると考えられる。ページの分類を使用して、メインメモリ220および/またはSSS225からページを退避させることができる。

10

【0030】

効率、性能、または他の理由のため、データベースは、テーブル、インデックス、または他のデータベースデータに対応するデータを回転媒体210などの記憶装置の1つのセット上に、また、ログに対応するデータを回転媒体211などの記憶装置の別のセット上に、配置することができる。あるデータベースでは、記憶装置の1つのセットのみを使用してテーブルとログを記憶することができる。

20

【0031】

バッファプール215には、より早いメモリ（回転媒体210、211に関する）内にある1つまたは複数のデータベースの一部が含まれる。1つまたは複数のデータベースの一部はページと呼ばれることがある。バッファプール215のメモリ（例えばメインメモリ220）のいくつかを揮発性とすることができる。すなわち、メモリへの電源が切られると、メモリに記憶されているどのデータも失われる。メインメモリ220は、RAM、キャッシュ、プロセッサメモリ、または他の揮発性高速メモリとして実装することができる。バッファプール215の他のメモリ（例えばSSS225）を不揮発性とすることができる。すなわち、不揮発性メモリへの電源が切られると、メモリに記憶されているどのデータも保持される。

30

【0032】

バッファプール215に割り当てられるSSS（または他のメモリ）の量は動的に変化させることができる。すなわち、バッファプール215に割り当てられるSSS（または他のメモリ）のバイト数の増加、減少、またはゼロへの設定を、自動的、半自動的、または手動で、DBMS 205の実行の前、最中または後に行うことができる。

【0033】

DBMS 205は、バッファプール215に記憶されるデータベースのページに関するメタデータを保持することができる。各ページに対して、このメタデータには、例えば、ページ上のデータがアクセスされた最も新しい1つまたは2つの時間を記憶するための2つのタイプスタンプ、ページが修正されたかどうかを示すダーティフラグ、ページのコピーがSSS225に記憶されているかどうかを示すフラグ、ページがアクセスされた回数を示すカウンタ、より最近のアクセスに対してはより重みが付けられた、ページへのアクセスの頻度を示す重み値、ページに関する他の情報、などを含むことができる。このメタデータは、バッファプール215（例えばページとして）内に、または何らかの他の場所に記憶させることができる。

40

【0034】

操作において、DBMS 205が、データにアクセスする要求を受け取ると、DBMS 205はまず、データがバッファプール215内にあるかどうかを判定することができる。バッファプール215内にデータが無い場合、DBMSは、回転媒体210からメインメモリ220にデータを読み込むためのスペースがメインメモリ220内にあるかどうか

50

を判定することができる。メインメモリ 220 内にスペースが無い場合、DBMS は、回転媒体 210 からページを読み込むためのスペースを作るために、メインメモリ 220 から退避させることができるページを判定する。

【0035】

上記のステップは、ある程度順次的に説明されたが、他の実装においては、これらのステップは異なる順番でまたは並行して行われても良い。例えば、一実施形態において、DBMS 205 は、メインメモリ 220 内のフリースペースのある部分を維持を試みることができる。このために、周期的に、またはメインメモリ 220 内のフリースペースが閾値以下に減少していると DBMS 205 が判定した時に、DBMS 205 は、メインメモリ 220 をスキャンして、どのページをメインメモリ 220 から退避させてスペースを解放するかを判定することができる。このようにして、DBMS 205 は、回転媒体 210 との間でページを読み書きする用途で使用するための、メインメモリ 220 内にあ

10

【0036】

一実施形態において、メモリプレッシャがあると（例えば、バッファプール内のフリースペースが事前に定義、計算、選択された閾値または他の閾値を下回り、処理がより多くのメモリを要求する、または何らかの他のメモリプレッシャ条件がそろった場合）、DBMS 205 は、スキャンアルゴリズムを実行することができる。スキャンアルゴリズムは、ページを異なるカテゴリに分類するために使用することができる閾値を判定することができる。例えば、一実施形態において、スキャンアルゴリズムは閾値を判定して、ページをページに関するメタデータに基づきホット、ウォーム、コールドのカテゴリに分類

20

【0037】

一実施形態において、スキャンアルゴリズムは、選択された数の無作為のページのメタデータのサンプリングを行うことにより閾値を判定することができる。サンプリングから、少なくとも 2 つの閾値を判定することができる。これらの閾値は、ページへのアクセスの頻度に関するパーセンタイルに対応し得る。2 つの閾値の内の最低位の値よりも低いアクセス特性を有するページはコールドであると考えられる。最低位と上位の閾値の間のアクセス特性を有するページはウォームであると考えられる。上位の閾値よりも高いアクセス特性を有するページはホットであると考えられる。

30

【0038】

用語「ホット」「ウォーム」「コールド」は、本明細書に記載される主題の態様に制限を課すことを意味しない。これらの用語の発想は、ページへのアクセスがパーセンタイル内に含まれるということである。パーセンタイルの高位の範囲に含まれるページはホットであると言うことができる。パーセンタイルの低位と含まれるページはウォームであると言うことができる。パーセンタイルの低位の範囲に含まれるページはコールドであると言うことができる。他の単語、番号、識別子、データ構造、などを、本明細書に記載される主題の態様の主旨または範囲から逸脱することなく、単語ホット、ウォーム、コールドの代わりに使用することができる。さらに、3 つの名称が指定されているだけであるが、他の実施形態においては、3 つより多くの名称を指定することができる。

40

【0039】

ページの分類に使用することができる基準の一例は、ページへのアクセスの頻度である。ページの分類に使用することができる別の基準の例は、ページへのアクセスの履歴である。例えば、アクセス履歴には、ページがアクセスされた最後の 1 回または 2 回の時間を含むことができる。最後のアクセス時間のみを使用してページを分類する場合は、この基準は LRU (last recently used) アルゴリズムと呼ばれることがある。最後のアクセスの 1 回前のアクセス時間を使用してページを分類する場合は、この基準は LRU - 2 アルゴリズムと呼ばれることがある。上記の例では、全てを包括または網羅することは意図されない。実際、本明細書の教示に基づき、当業者はページの分類に使用することができる多くの他の基準を認識するであろう。

50

【 0 0 4 0 】

ー実施形態において、上記の閾値を判定するためのサンプリングをメインメモリ 2 2 0 内のページに対して行うことができ、一方、低位の閾値を判定するためのサンプリングをメインメモリ 2 2 0 および S S S 2 2 5 内のページに対して行うことができる。

【 0 0 4 1 】

これらの閾値が判定された後、D B M S 2 0 5 がページを退避させてメインメモリ 2 2 0 にもっと多くのスペースを確保する必要がある時、D B M S 2 0 5 は、様々な順序、例えば、順次、ラウンドロビン方式、ランダム、L R U 方式、局所性に基づく方式、別の順序などのいずれかでバッファプール内のページをスキャンすることができる。ページがコールド閾値より上でウォーム閾値より下にあり、S S S 2 2 5 内に利用可能なスペースがある場合、D B M S 2 0 5 は、ページを S S S 2 2 5 にコピーすること、または、（例えば、メインメモリ 2 2 0 から S S S 2 2 5 にコピーされるページのキューなどのデータ構造内のページにポインタを置くことにより）実行可能になり次第、ページが S S S 2 2 5 にコピーされることを示すこと、ができる。ページが S S S 2 2 5 にコピーされた後、メインメモリ 2 2 0 内でフリーにされたメモリを使用して別のページを記憶することができる。

10

【 0 0 4 2 】

ウォームおよびコールドであるページに使用できる十分なスペースが S S S 2 2 5 に無い場合、コールドであるページを、フラッシュするかキューに置くかしてディスクにフラッシュする（ダーティである場合）、または、利用可能であるとしてマークする（クリーンな場合）ことができる。ディスクにフラッシュするページは、メインメモリ 2 2 0 からのものでも S S S 2 2 5 からのものでも良い。この場合、一実施形態において、コールドのページを S S S 2 2 5 からフラッシュすることよりもコールドのページをメインメモリ 2 2 0 からフラッシュすることに優先度が与えられてもよい。例えば、1 つまたは複数のデータ構造（例えば、1 つまたは複数のキュー）を、コールドのページをメインメモリ 2 2 0 および S S S 2 2 5 からフラッシュするために維持することができる。メインメモリ 2 2 0 からのコールドのページをディスクにフラッシュした後、S S S 2 2 5 からのコールドのページをディスクにフラッシュすることができる。別の実施形態において、メインメモリ 2 2 0 からのコールドのページのフラッシュを、S S S 2 2 5 からのコールドのページのフラッシュとインターリーブすることができる。ハードウェアサブシステムが適切な設備を提供する場合、コールドのページをメインメモリ 2 2 0 からディスクにフラッシュすること、およびコールドのページを S S S 2 2 5 からディスクにフラッシュすることを、並行して行うことができる。

20

30

【 0 0 4 3 】

適切なハードウェアが利用可能な場合（例えば、ダイレクトメモリアクセス（DMA）ハードウェアなど）、S S S 2 2 5 からディスクへのページのフラッシュは、ページを S S S 2 2 5 からメインメモリ 2 2 0 に読み込まずに実行することができる。そのようなハードウェアを S S S 2 2 5 に利用できない場合、S S S 2 2 5 からディスクへのページのフラッシュは、ページをメインメモリ 2 2 0 に読み込み、そのページをメインメモリ 2 2 0 からディスクにコピーすることにより行うことができる。

40

【 0 0 4 4 】

D B M S 2 0 5 が、メインメモリ 2 2 0 ではなく S S S 2 2 5 にあるページにアクセスする必要がある場合、そのページを S S S 2 2 5 からメインメモリ 2 2 0 にコピーすることができる。

【 0 0 4 5 】

バッファのページが S S S 2 2 5 に書き込まれると、可能であれば、複数の書き込みを単一の書き込みに組み合わせても良い。これにより、I / O スループットを増加させること、また、S S S 2 2 5 の寿命期待値を増加させることができる。

【 0 0 4 6 】

S S S 2 2 5 の I / O 閾値に到達すると、新しい I / O はディスクに向けられうる。I

50

／Ｏ閾値に到達するのは、ＳＳＳ２２５へのアクセスが、ＳＳＳ２２５の読み／書きの帯域幅または何らかの他の定義済みの帯域幅に達する時である。Ｉ／Ｏ閾値に到達したかどうかは、例えば、Ｉ／Ｏ応答時間、Ｉ／Ｏの回数、または何らかの他の要因によって判定することができる。例えば、アクセスの特定のパターンまたは頻度で、ＳＳＳ２２５からＲＡＭ２２０にデータを転送するＩ／Ｏは、ＳＳＳ２２５が利用可能な帯域幅を超えることがある。この場合、ＳＳＳ２２５がより多くのＩ／Ｏに利用可能になるまで、後に続くＩ／Ｏを回転媒体２１０に送信することができる。

【００４７】

図３は、本明細書に記載される主題の態様に従ってデータベースをホストするシステムの構成要素を例示するブロック図である。構成要素３００には、メタデータ３０５、メインメモリバッファページ３１０、ＳＳＳバッファページ３１５、および他の記憶装置３２０が含まれる。メタデータ３０５には、上述したようなページに関するデータが含まれる。メタデータ３０５は、メインメモリ、キャッシュ、または、何らかの他の高速メモリに記憶させることができる。メタデータ３０５は、ページがメインメモリおよび／またはＳＳＳに記憶されることを示す場合がある。

【００４８】

メインメモリバッファページ３１０は、ＲＡＭまたは他の揮発性メモリなどのメインメモリ内に記憶させることができる。そのようなメモリには、機械的動作を伴わずにアクセスすることができる。換言すれば、そのようなメモリでは、主要部の任意の構成要素の物理的動作を伴わずに、データへのアクセスを提供することができる。そのようなメモリは、機械的なタイプの記憶装置より高速であることが多い。

【００４９】

メインメモリバッファページ３１０には、どの程度のスペースが利用可能であるかによって、ホットページ、ウォームページ、およびコールドページを含むことができる。例えば、メインメモリバッファページ３１０には、他の記憶装置３２０から最近検索された、いくつかのコールドページを含むことができる。いくつかのメインメモリバッファページ３１０のコピーを、ＳＳＳバッファページ３１５内に記憶させることができる。メインメモリバッファページ３１０内のいくつかのページが、ＳＳＳバッファ３１５にコピーされないこともある。さらに、ＳＳＳバッファページ３１５は、メインメモリバッファページ３１０内には無いページを含んでも良い。

【００５０】

ＳＳＳバッファページ３１５を、二次的メモリに記憶させることができる。性能向上を図るために、この二次的メモリは記憶装置３２０よりも性能が高くても良い（例えば、より速い応答時間、より広い帯域幅など）。何らかのメモリ（ソリッドステート記憶装置など）は、機械的動作を伴わずにアクセスすることができ、また不揮発性であっても良い。このメモリは記憶装置３２０よりも性能が高くても良いが、メインメモリよりも遅くても良い。

【００５１】

他の記憶装置３２０は、ハードディスク、テープ、他の不揮発性記憶装置などの不揮発性記憶装置を含むことができる。この他の記憶装置３２０は、移動するように（例えば、プラッタ上を前後に、回転、またはその他）動作可能であり、記憶装置３２０の記憶デバイス上のメモリへのアクセスを提供する構成要素（例えば、アーム、プラッタ、または他の媒体など）を含むことができる。安価な実装では、記憶装置３２０は、システム３００のメインメモリおよびＳＳＳよりもスループットが低くても良い。しかしながら、いくつかのシステムにおいては、共同して機能する多くの記憶デバイス（例えば、ディスク）を使用すると、記憶装置３２０のスループットがＳＳＳのスループットに届くかまたはそれを超えることもあるが、応答時間は比較的大きくなりうる。

【００５２】

メタデータ３０５は、メインメモリバッファページ３１０およびＳＳＳバッファページ３１５内のページに関する情報を記憶する。そのような情報には、図２に関連して上述し

10

20

30

40

50

たメタデータを含むことができる。

【 0 0 5 3 】

図 2 から図 3 に例示される構成要素は例であり、必要とされ得るまたは含まれ得る構成要素の全てを包含することを意味しない。他の実施形態において、図 2 から図 3 に関連して記載される構成要素および / または機能は、本明細書に記載される主題の態様の主旨または範囲から逸脱することなく、他の構成要素（図示されるもの、または図示されないもの）に含まれて良く、またはサブ構成要素内に配置されて良い。いくつかの実施形態において、図 2 から図 3 に関連して記載される構成要素および / または機能は、複数の装置に亘って分散させることができる。

【 0 0 5 4 】

図 4 から図 5 は、本明細書に記載される主題の態様に従って発生し得る動作を概略的に表すフロー図である。説明を簡略化するために、図 4 から図 5 に関連して記載される方法論は、一連の動作として示され記載される。本明細書に記載される主題の態様が、例示される動作によりおよび / または動作の順番により制限されないことは理解されかつ認められるべきである。一実施形態において、動作は以下に記載される順序で発生する。しかし、他の実施形態において、動作は並行して、別の順序でおよび / または本明細書に提示も記載もされない他の動作と共に発生することもある。さらに、本明細書に記載される主題の態様に従って方法論を実装するために、例示される全ての動作が要求されるわけではない。加えて、当業者であれば、方法論を、代替えとして状態図を介して相互に関連する一連の状態として、または一連の事象として、表すことができることを理解し分かるであろう。

【 0 0 5 5 】

図 4 は、本明細書に記載される主題の態様に従って、閾値を判定するためにページをスキャンし、必要に応じてページを退避させる際に発生しうる、いくつかの例示の動作を概略的に表すフロー図である。

【 0 0 5 6 】

ブロック 4 1 0 にて、ページをスキャンして閾値を判定する要求が受け取られる。例えば、図 4 を参照すると、メモリプレッシャに応じて、D B M S 2 0 5 は、ページをスキャンする要求をバッファプール構成要素に送信することができる。

【 0 0 5 7 】

ブロック 4 1 5 にて、ページのメタデータがアクセスされる。例えば、図 3 を参照すると、メタデータ 3 0 5 がアクセスされる。メタデータの全てにアクセスする代わりに、データベースのバッファプールのランダムに選択されたいくつかのページの、そのメタデータをサンプリングすることができる。一実施形態において、ページのメタデータのサンプリングには、ページのメタデータから、サンプリングされたそれぞれのページの 1 つまたは 2 つのタイムスタンプを取得することが含まれ、該 1 つまたは 2 つのタイムスタンプはサンプリングされたページがアクセスされた最後の時間に対応する。このメタデータを、ページへのアクセスに対応する値を生成する関数（例えば、分類関数）に与えることができる。サンプリングのメタデータに対するその関数によって生成された値を使用して、閾値を選択することができる。例えば、関数により値を取得した後、値について定義済みの異なるパーセンタイルに対応する閾値が選択される。パーセンタイルの範囲（例えば、0 から 5、5 から 2 5、2 5 から 1 0 0）をホット、ウォーム、コールドに対応させることができる。

【 0 0 5 8 】

ブロック 4 2 0 にて、コールド閾値をメタデータから判定することができる。例えば、コールド閾値は、0 から 2 5 の間のパーセンタイルの範囲内のサンプリングされたページのアクセスの頻度に対応させることができる。コールド閾値が判定されると、上記の関数をページのメタデータにより表されるアクセスデータに適用して、コールド閾値以下の値が返される場合、ページはコールドであると判定することができる。

【 0 0 5 9 】

ブロック 4 2 5 にて、ウォーム閾値をメタデータから判定することができる。例えば、ウォーム閾値は、2 5 0 から 7 5 の間のパーセントイルの範囲内のサンプリングされたページのアクセスの頻度に対応させることができる。ウォーム閾値が判定されると、上記の関数をページのメタデータにより表されるアクセスデータに適用して、コールド閾値より大きくかつウォーム閾値以下の値が返される場合、ページはウォームであると判定することができる。

【 0 0 6 0 】

ブロック 4 3 0 にて、ページを退避させる要求が受け取られる。例えば、図 2 を参照すると、DBMS 2 0 5 の退避構成要素は、1 つまたは複数の他のページのために、バッファプール 2 1 5 から freespace にページを退避させる要求を受け取ることができる。

10

【 0 0 6 1 】

ブロック 4 3 5 にて、バッファプール 2 1 5 のページは、退避を考慮して選択することができる。例えば、図 3 を参照すると、メインメモリバッファページ 3 1 0 の内の 1 つを、退避を考慮して選択することができる。選択は、退避に適したページを見つけるためのページを巡回するために使用されるアルゴリズム、キューまたは他のデータ構造などを使用して行われる。

【 0 0 6 2 】

ブロック 4 4 0 にて、ページが分類される。これは、値を得るために上述の関数をページのメタデータに適用することにより行われうる。次に、この値は上述した閾値と比較され、ページが、ホット、ウォーム、またはコールドであるかどうか判定することができる。

20

【 0 0 6 3 】

ブロック 4 4 5 にて、適切である場合は、ページが退避される。例えば、ページがコールドであると判定される場合、ページはディスクから退避されても良い。ページがウォームである場合、SSS スループットのための閾値に達していなければ、ページを SSS にコピーすることができ、SSS スループットのための閾値に達している場合、ページはディスクから退避されて良い。ページがホットである場合、ページは退避せずにメインメモリ内に残すことができる。

【 0 0 6 4 】

ブロック 4 5 0 にて、他の動作があれば、それを実行させることができる。

30

【 0 0 6 5 】

図 5 は、本明細書に記載される主題の態様に従って、DBMS がアクセス要求を受け取り、かつバッファプールがフルである場合に発生するようないくつかの例示の動作を概略的に表すフロー図である。ブロック 5 0 5 にて、動作が開始される。

【 0 0 6 6 】

ブロック 5 1 0 にて、データベースのデータにアクセスする要求が受け取られる。例えば、図 2 を参照すると、DBMS 2 0 5 はデータベース内のデータにアクセスする要求を受け取ることができる。

【 0 0 6 7 】

ブロック 5 1 5 にて、DBMS は、データがメインメモリの外に存在すると判定する。例えば、図 2 を参照すると、DBMS 2 0 5 は、要求されたデータが SSS 2 2 5 または回転媒体 2 1 0 に存在すると判定することができる。

40

【 0 0 6 8 】

ブロック 5 2 0 にて、DBMS は、メインメモリ内のバッファプールがフルであると判定する。例えば、図 2 を参照すると、DBMS 2 0 5 は、メインメモリ 2 2 0 がページでフルになっており、ページを退避させてデータのためのスペースを作る必要がある、と判定する。

【 0 0 6 9 】

ブロック 5 2 5 にて、退避のためのページが選択される。退避用のページを選択することは、ページへのアクセスの頻度による少なくとも 3 つの分類（例えば、ホット、ウォー

50

ムおよびコールド)に従って、ページを分類する関数(例えば、上述の分類関数)を使用することを含んでも良い。例えば、図2を参照すると、DBMS205は、メインメモリ220内の退避用のページを判定することができる。

【0070】

ブロック530にて、ページが退避させられる。例えば、図2を参照すると、DBMS205は、選択されたページをSSS225または回転媒体210に退避させることができる。ページをSSS225に退避させることは、ページをSSS225にコピーすることを含んでも良い。これにより、ページはバッファプール215に保持され、バッファプールのメインメモリ220には無くなる。このようにして、SSS225はバッファプールを拡張させることができる(メインメモリ220から退避させられたページ用の別個のキャッシュとして動作するだけではない)。

10

【0071】

ブロック535にて、他の動作があれば、それを実行させることができる。

【0072】

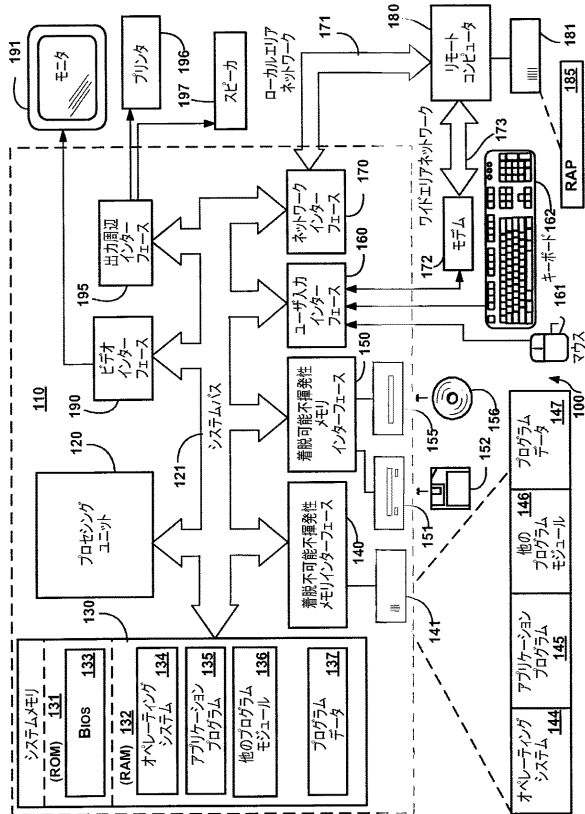
上記の検討では、ソリッドステート記憶装置を使用してバッファプールを拡張することに言及したが、他の実施形態において、ソリッドステート記憶装置以外の記憶装置を使用してバッファプールを拡張することができる。他の記憶装置は、現在存在するものまたは開発中のものでも良い。本明細書に記載される主題の態様に従うと、他の記憶装置は、回転式の記憶装置よりも高いがメインメモリよりも低い、スループット、帯域幅、または何らかの他の特徴を有して良い。

20

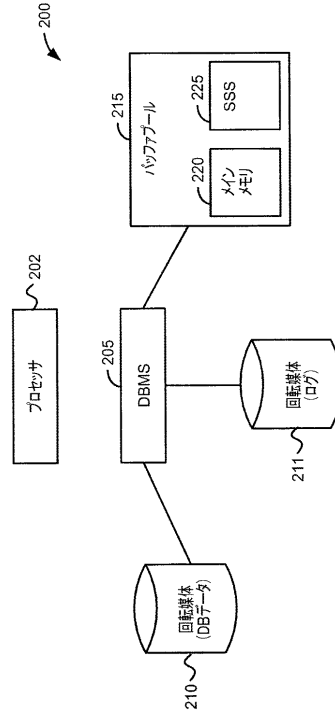
【0073】

前述の詳細な説明から分かるように、データシステムのためのバッファプールに関する態様が記載された。本明細書に記載される主題の態様が種々の修正および代替の構造の影響を受けやすい一方、特定の例示される実施形態が、図面において示され、上記で詳細に説明された。しかしながら、特許請求される主題の態様を開示される特定の形式に制限することは意図されず、逆に、本発明が本明細書に記載される主題の種々の態様の主旨および範囲内にある全ての修正、代替の構造および等価物に及ぶ、ことを理解されたい。

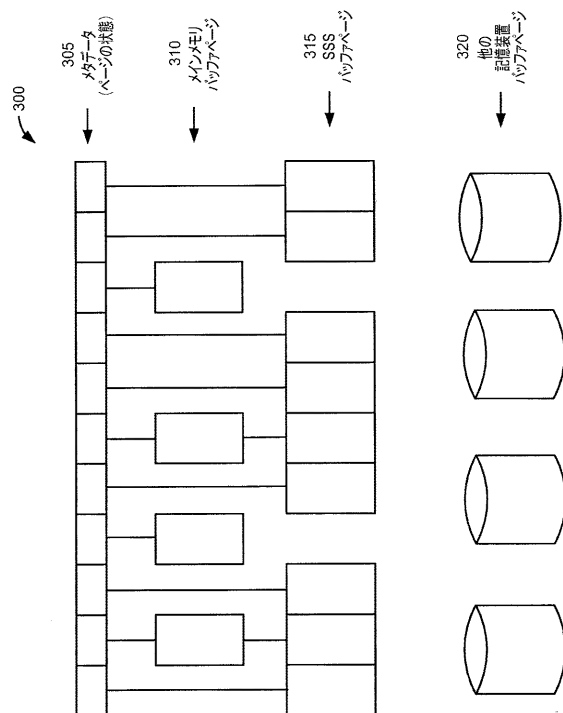
【 図 1 】



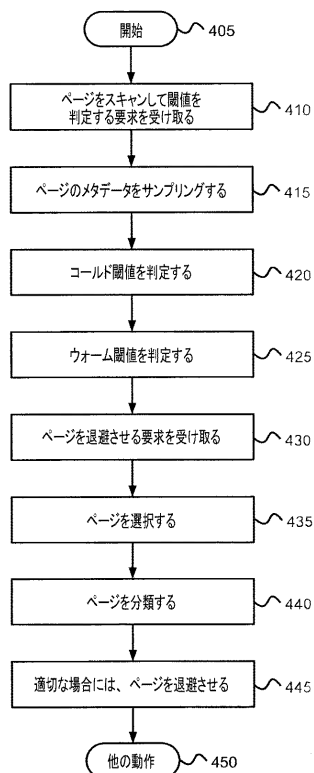
【 図 2 】



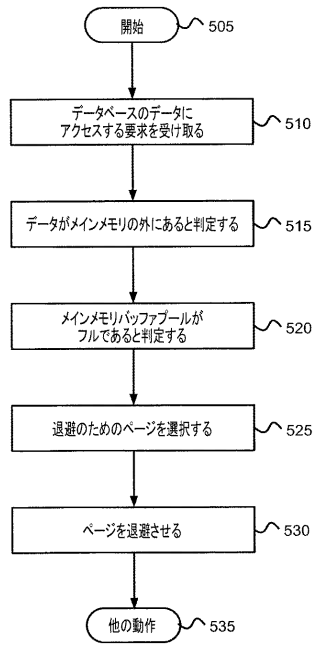
【 図 3 】



【圖 4】



【図 5】



フロントページの続き

- (74)代理人 100120112
弁理士 中西 基晴
- (74)代理人 100147991
弁理士 鳥居 健一
- (74)代理人 100119781
弁理士 中村 彰吾
- (74)代理人 100162846
弁理士 大牧 綾子
- (74)代理人 100173565
弁理士 末松 亮太
- (74)代理人 100138759
弁理士 大房 直樹
- (74)代理人 100091063
弁理士 田中 英夫
- (72)発明者 チェンリアン チャン
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション エルシーエー - インターナショナル パテント内
- (72)発明者 サダシバン クリシュナムルティ
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション エルシーエー - インターナショナル パテント内
- (72)発明者 ゲオルギー アイ・レイニャ
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション エルシーエー - インターナショナル パテント内
- (72)発明者 アレクサンドル パービットスキー
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション エルシーエー - インターナショナル パテント内

審査官 野田 佳邦

- (56)参考文献 特表2005-539309(JP,A)
特開2005-173637(JP,A)
特開2006-185169(JP,A)
特開平08-153014(JP,A)
特開2003-216460(JP,A)
特開2001-282617(JP,A)
特開平08-235042(JP,A)

- (58)調査した分野(Int.Cl., DB名)
G06F 12/00