

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2010-524133

(P2010-524133A)

(43) 公表日 平成22年7月15日(2010.7.15)

(51) Int.Cl. F I テーマコード (参考)
G 0 6 F 9/38 (2006.01) G 0 6 F 9/38 3 5 0 A 5 B 0 1 3

審査請求 未請求 予備審査請求 未請求 (全 18 頁)

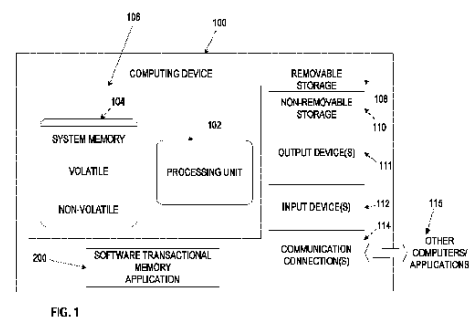
(21) 出願番号	特願2010-503112 (P2010-503112)	(71) 出願人	500046438
(86) (22) 出願日	平成20年3月13日 (2008. 3. 13)		マイクロソフト コーポレーション
(85) 翻訳文提出日	平成21年10月13日 (2009. 10. 13)		アメリカ合衆国 ワシントン州 9805
(86) 国際出願番号	PCT/US2008/056873		2-6399 レッドモンド ワン マイ
(87) 国際公開番号	W02008/127821		クロソフト ウェイ
(87) 国際公開日	平成20年10月23日 (2008. 10. 23)	(74) 代理人	100077481
(31) 優先権主張番号	11/786, 174		弁理士 谷 義一
(32) 優先日	平成19年4月11日 (2007. 4. 11)	(74) 代理人	100088915
(33) 優先権主張国	米国 (US)		弁理士 阿部 和夫
		(72) 発明者	デイビッド デトレフス
			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 インターナショナル パテンツ内

最終頁に続く

(54) 【発明の名称】 バッファリングされた書き込みおよび強制的直列化順序を使用するトランザクショナルメモリ

(57) 【要約】

ソフトウェアトランザクショナルメモリシステムにおいてバッファリングされた書き込みおよび強制的直列化順序をサポートする様々な技術および技法を開示する。オブジェクトのシャドウコピーへの書き込みを行い、コミット中に個々のトランザクションの妥当性を確認した後でコンテンツをオブジェクトに書き戻す、バッファリングされた書き込みプロセスを提供する。特定のトランザクションの書き込みロックが最初に取得されると、特定のオブジェクトのシャドウコピーが作成される。シャドウコピーへの書き込みと該シャドウコピーからの読み込みが行われる。コミット中に特定のトランザクションの妥当性を確認した後、コンテンツは、シャドウコピーから特定のオブジェクトに書き込まれる。トランザクションがコミットされる順序がトランザクションの抽象的直列化順序に一致することを保証する、トランザクション順序付けプロセスを提供する。トランザクションは、該トランザクションのチケット番号が、コミットするべき次のトランザクションをトラッキングするグローバル番号に一致するまで、コミットを許可されない。



【特許請求の範囲】**【請求項 1】**

複数のトランザクションを管理するように動作可能なソフトウェアトランザクショナルメモリシステムを提供するステップと、

オブジェクトのシャドウコピーへの書込みと該シャドウコピーからの読み込みを行い、コミット中に前記複数のトランザクションの個々のトランザクションの妥当性を確認した後にコンテンツを前記オブジェクトに書き戻す、バッファリングされた書込みプロセスを提供するステップと、

前記複数のトランザクションがコミットされる順序が前記複数のトランザクションの抽象的直列化順序に一致することを保証する、トランザクション順序付けプロセスを提供するステップと

を含むステップをコンピュータに実行させるためのコンピュータ実行可能命令を有することを特徴とするコンピュータ読取可能媒体。

【請求項 2】

前記バッファリングされた書込みプロセスは、書込みロックが最初に取得されたときに前記オブジェクトのうちの特定の 1 つの特定のシャドウコピーを作成するように動作可能であることを特徴とする請求項 1 に記載のコンピュータ読取可能媒体。

【請求項 3】

前記バッファリングされた書込みプロセスは、前記特定のシャドウコピーがすでに存在する場合は、前記オブジェクトのうちの前記特定の 1 つの前記特定のシャドウコピーにアクセスするように動作可能であることを特徴とする請求項 2 に記載のコンピュータ読取可能媒体。

【請求項 4】

前記トランザクション順序付けプロセスは、グローバルデータ構造を使用して、前記トランザクションがコミットされている前記順序が前記抽象的直列化順序に一致することを保証するように動作可能であることを特徴とする請求項 1 に記載のコンピュータ読取可能媒体。

【請求項 5】

前記グローバルデータ構造は、入力コミットカウンタと終了コミットカウンタとを含み、前記入力コミットカウンタは、コミット処理に入ったトランザクションの第 1 の数を表し、前記終了コミットカウンタは、コミット処理を終了したトランザクションの第 2 の数を表すことを特徴とする請求項 4 に記載のコンピュータ読取可能媒体。

【請求項 6】

前記個々のトランザクションがコミット処理に入ると、前記個々のトランザクションは、前記入力コミットカウンタをアトミックに読み込み、インクリメントし、前記入力コミットカウンタから読み込まれた値を前記個々のトランザクションのチケット番号として割り当ててことを特徴とする請求項 5 に記載のコンピュータ読取可能媒体。

【請求項 7】

前記個々のトランザクションは、該個々のトランザクションがコミット処理の後に続行することを許可される前、前記終了コミットカウンタが該個々のトランザクションの前記チケット番号に一致するまで待機することを特徴とする請求項 6 に記載のコンピュータ読取可能媒体。

【請求項 8】

前記バッファリングされた書込みプロセス、および前記トランザクション順序付けプロセスの使用は、私有化関連問題が回避されることを保証することを特徴とする請求項 1 に記載のコンピュータ読取可能媒体。

【請求項 9】

ソフトウェアトランザクショナルメモリシステムにおいてバッファリングされた書込みを提供するための方法であって、

複数のトランザクションを管理するように動作可能なソフトウェアトランザクショナル

10

20

30

40

50

メモリシステムを提供するステップと、

特定のトランザクションについて書込みロックを最初に取得したときに、特定のオブジェクトのシャドウコピーを作成するステップと、

前記シャドウコピーへの書込みと前記シャドウコピーからの読み込みを行うステップと、
コミット中に前記特定のトランザクションの妥当性を確認した後に、コンテンツを前記シャドウコピーから前記特定のオブジェクトに書き込むステップと

を含むことを特徴とする方法。

【請求項 10】

前記特定のオブジェクトは、オブジェクトヘッダ内に、ロックタイプおよびバージョンをトラッキングするように動作可能なトランザクショナルメモリワードを有することを特徴とする請求項 9 に記載の方法。

10

【請求項 11】

前記トランザクショナルメモリワードは、前記シャドウコピーへのポインタを表すことを特徴とする請求項 10 に記載の方法。

【請求項 12】

前記特定のトランザクションは、前記特定のトランザクションが直列化されたのと同じ順序でコミット処理を終了することを許可されることを特徴とする請求項 9 に記載の方法

。

【請求項 13】

前記特定のトランザクションは、前記特定のトランザクションのチケット番号が、コミットを許可されるべき次のトランザクションを識別するグローバルカウンタに一致するまで、コミット処理を終了しないことを特徴とする請求項 12 に記載の方法。

20

【請求項 14】

請求項 9 に記載のステップをコンピュータに実行させるためのコンピュータ実行可能命令を有することを特徴とするコンピュータ読取可能媒体。

【請求項 15】

ソフトウェアトランザクショナルメモリシステムにおいてトランザクションが直列化順序でコミットされることを保証するための方法であって、

複数のトランザクションを管理するように動作可能なソフトウェアトランザクショナルメモリシステムを提供するステップと、

30

コミット処理に入ったトランザクションの第 1 の数を表す入力コミットカウンタを提供するステップと、

コミット処理を終了したトランザクションの第 2 の数を表す終了コミットカウンタを提供するステップと、

前記複数のトランザクションのうちの特定のトランザクションがコミット処理に入ったときに、前記入力コミットカウンタをアトミックに読み込み、インクリメントし、前記入力コミットカウンタから読み込まれた値を前記特定のトランザクションのチケット番号として割り当てるステップと、

前記終了コミットカウンタが前記特定のトランザクションの前記チケット番号に一致するときに、前記特定のトランザクションをコミットするステップと

40

を含むことを特徴とする方法。

【請求項 16】

バッファリングされた書込みプロセスは、前記特定のトランザクションがコミットするまでオブジェクトのシャドウコピーへの書込みと該シャドウコピーからの読み込みを行うために使用されることを特徴とする請求項 15 に記載の方法。

【請求項 17】

特定のオブジェクトのシャドウコピーは、書込みロックが最初に取得されたときに作成されることを特徴とする請求項 16 に記載の方法。

【請求項 18】

書込みは、前記特定のオブジェクトの前記シャドウコピーに行われ、読み込みは、前記特

50

定のオブジェクトの前記シャドウコピーから行われることを特徴とする請求項 17 に記載の方法。

【請求項 19】

コンテンツは、前記特定のトランザクションがコミット中に妥当性を確認された後、前記シャドウコピーから前記特定のオブジェクトに書き込まれることを特徴とする請求項 18 に記載の方法。

【請求項 20】

請求項 15 に記載のステップをコンピュータに実行させるためのコンピュータ実行可能命令を有することを特徴とするコンピュータ読取可能媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、バッファリングされた書込み (buffered writes) および強制的直列化順序 (enforced serialization order) を使用するトランザクショナルメモリに関する。

【背景技術】

【0002】

コンピュータの処理ユニットは、命令のプログラムテキストからの命令ストリームを実行する。各命令は、その後続の命令、すなわち、それに続く命令、あるいは分岐またはコールの場合は他の何らかの命令のいずれかを指定する。したがって、プロセッサは、一度に 1 つの命令を実行する (いわゆるパイプライン化された (pipelined) 「アウトオブオーダー (out-of-order)」のプロセッサは、該プロセッサの実装においてはこれに反しているが、これらのセマンティクスは保っている)。プログラムは、通常、区別化された開始命令 (distinguished start instruction) を有するプログラムテキストにコンパイルされる。例えば、C プログラムでは、「メイン」メソッドの第 1 の命令は、区別化された開始命令である。この後に実行される命令のシーケンスを決定する「プロセッサコンテキスト (processor context)」は、しばしば「コントロールスレッド」、または単に「スレッド」と呼ばれる。プログラムは、各プロセスが「仮想」マシンの全てのメモリに対する唯一のアクセスを有するかのように振る舞うことを可能にする、仮想アドレス空間を提供するオペレーティングシステムプロセスにおいて実行する。オペレーティングシステムプロセスは、仮想アドレス空間に加えて、ファイル記述子や 1 つまたは複数のスレッドなどの、様々なプロセスごとのオペレーティングリソースを提供する。従来のプログラムは、シングルスレッド (single-threaded) である。すなわち、従来のプログラムは、単一のコントロールスレッドしか有しないプロセスで実行する。

【0003】

共有メモリマルチプロセッサ (shared-memory multiprocessor) は、同じメモリへのアクセスを共有するいくつかのプロセッサを有しており、あるプロセッサによる書込みが、別のプロセッサによるその後の読込みによって観察されることがある。そのようなマシンは、別々のプロセッサ上でいくつかの異なるプログラムをプロセスごとに実行することによって使用されること可能である。このモードでは、プロセスがそれぞれ別個のアドレス空間を有するので、実際には、共有メモリは使用されない。しかし、別のモードでは、プログラムは、それが実行するプロセスにおいていくつかのコントロールのスレッドを作成することができ、これらのスレッドは、複数のプロセッサ上で同時に実行し、共有メモリを介して通信することができる。(そのようなマルチスレッド、すなわち並列 (concurrent) プログラムは、ユニプロセッサ上で実行することもでき、一般に、プログラムは、存在する利用可能なプロセッサより多くのスレッドを作成することができる。オペレーティングシステムのジョブの 1 つは、利用可能なプロセッサ上で動作可能なスレッドの実行をスケジューリングすることである。したがって、動作中のスレッドを任意の命令において中断して、別のスレッドが再開することを可能にする。)

【0004】

スレッドによってこのように同時にインターリーブされる命令の実行は、並列プログラ

10

20

30

40

50

ミングを非常に困難にする。例として、赤色のカードが全て第1の山に、黒色のカードが全て第2の山になるように分けられている1組のカードを想像されたい。各カードが1つの命令を表し、各山が1つのスレッドを表す。ブリッジのシャフリング技法を使用してこれらの山を混ぜ合わせる。赤色のカードの互いに対する順序は変更されず、黒色のカードの互いに対する順序も変更されないが、カードは交互に配置（インターリーブ）された。これは、まさに、スレッドが並行に実行するときが発生することである。それぞれが可能な実行を表す非常に多くの可能なインタリーブ（interleavings）があることも明らかにはずである。プログラムは、そのような可能な実行の全てについて正確に機能する必要がある。

【0005】

スレッドが並列コンピューティング環境で実行するとき、各スレッドが共有メモリなどのシステムリソースとどのように対話するかを管理するための機構が必要とされる。ソフトウェアトランザクショナルメモリ（STM：software transactional memory）は、並列コンピューティング内の共有メモリへのアクセスを制御するためのデータベーストランザクションと類似の並列処理制御機構（concurrency control mechanism）である。トランザクショナルメモリのコンテキストにおけるトランザクションは、共有メモリへの一連の読み込みと書き込みを実行し、しかもアトミックにトランザクション全体がシステム内で実行している唯一のコントロールスレッドであるかのように実行する、コードの一部である。トランザクション $T \times 1$ が、トランザクション $T \times 2$ によるいずれかの書き込みを観察した場合、トランザクション $T \times 1$ は、次いで $T \times 2$ による全ての書き込みを観察する。トランザクショナルメモリのコンテキストにおけるデータの位置は、単一のオブジェクト、（C++などにおける）キャッシュライン、ページ、単一のワードなどの、アクセスされている共有メモリの特定のセグメントである。トランザクショナルメモリシステムにおける、あるタイプの並列処理制御ロックモードは、楽観的並列処理制御（optimistic concurrency control）、または楽観的ロッキング（optimistic locking）である。

【0006】

楽観的並列処理制御では、システムは、衝突（conflict）が後に検出されることがあるというリスクがあっても進行することを試みる。トランザクショナルメモリシステムは、しばしば、衝突しているトランザクションの1つをロールバックして、それを再実行することによって、そのような衝突の自動的な解決を行う。楽観的オペレーション（optimistic operations）は、読み込みだけであり共有位置への書き込みは伴わないので（すなわち、ロックを取る）ので、悲観的オペレーション（pessimistic operations）と比べると比較的成本はかからない。その名前が暗示するように、楽観的オペレーションに対する期待は衝突がほとんどないことである。これが間違いであると分かった場合は、作業がすでに無駄であり、システムはその作業をやめて衝突の解決を試みる必要がある。

【発明の概要】

【発明が解決しようとする課題】

【0007】

楽観的並列処理制御が明示的に対処していない1つの深刻な問題は、私有化（privatization）シナリオにおいて発生し得る。私有化関連の問題は、プログラムが、並列スレッドに同じ共有メモリロケーションにアクセスするトランザクションを実行させ、これらのトランザクションのうちの1つが、ある共有メモリ位置を私有化するときが発生することがある。私有化は、あるトランザクションが、共有メモリロケーションが該トランザクションに対してのみアクセス可能にされるオペレーションを実行するときが発生する。例えば、あるオブジェクトOへの唯一のリファレンスが、あるグローバルにアクセス可能なキューQに格納されており、スレッド T_1 によって実行されているトランザクション $T \times 1$ が、QからOへのリファレンスを除去するオペレーションを行い、これをローカル変数 T_1 に格納した場合、 $T \times 1$ が、Oを T_1 に私有化することになる。

【0008】

STMの一部の実施形態では、私有化は、予期しない結果が発生させる可能性がある。

10

20

30

40

50

一部のSTM実施形態は、楽観的読込みを「インプレース (in-place)」書込みと結合することによって高い性能を達成しようと試みてきたが、この場合、トランザクショナル書込みは、メモリロケーションに対して直接行われる。これらの技法を使用して、私有化を行うプログラムを実装するときは、以下のシナリオがあり得る。あるグローバルロケーションGが、共有データ構造への一意のポインタを含む。2つのスレッドが、このデータ構造に並列にアクセスしようと試みるトランザクションを実行する。スレッドT1は、Gを読み込むトランザクションTx1を実行し、ポインタ読込みが非ヌルの場合、スレッドT1は、該ポインタが参照するデータ構造内の整数をインクリメントすることを試みる。スレッドT2は、Gをスレッドローカル変数にコピーしてGをヌルにセットするトランザクションTx2を実行する。次いで、スレッドT2は、Gをヌルにセットしたことによってデータ構造を成功裏に「私有化」したと信じて、スレッドローカルポインタ変数を介して該データ構造にアクセスする。しかし、楽観的読込みおよびインプレース書込みでは、ある実行で、Tx1が最初にGを読み込み、非ヌル値を観察することがある。このとき、Tx2は、その全体を実行する。Tx2は、Tx1が読み込んだロケーションG、を書き込むことで、Tx1を途中停止するように「決定づける (dooming)」が、これは、Tx1がコミットを試みるまで発見されない。したがって、Tx1は実行を継続して、データ構造内のフィールドをインクリメントする。このインクリメントは、Tx1がコミットできなかったときは取り消されることになるが、非トランザクショナルコードはスレッドT2でTx2の後に実行するという観点から、この書込み、および「取消し」オペレーションを行う書込みは両方とも「不可解な (inexplicable)」ことであり、これらは理由なく発生して、プログラムを不正確に実行させる可能性がある。

10

20

30

40

50

【0009】

別の種類の私有化関連の問題は、「直列化の異常 (serialization anomalies)」を有することである。前述したように、トランザクションは、並列トランザクションがある一連の順序で実行するという錯覚をプログラムに提供することによって、並列プログラミングを容易にする。具体的には、トランザクションTx2による読込みが、トランザクションTx1による書込みを観察した場合は、Tx2は、Tx1の後に直列化されなければならない。直列化の異常は、トランザクションがそれらの直列的な順序とは異なる順序で完了したときに発生する。プログラムが私有化のイディオムを利用するとき、このことにより、スレッド内でトランザクションの1つが完了した後に実行している非トランザクショナルコードが、「不可解な」書込みを観察する可能性がある。

【課題を解決するための手段】

【0010】

ソフトウェアトランザクショナルメモリシステムにおいて、バッファリングされた書込みと強制的直列化順序をサポートする様々な技術および技法を開示する。トランザクションを管理するように動作可能なソフトウェアトランザクショナルメモリシステムを提供する。オブジェクトのシャドウコピーへの書込みおよび該シャドウコピーからの読込みを行い、コミット中に個々のトランザクションの妥当性を確認した後にコンテンツをオブジェクトに書き戻す、バッファリングされた書込みプロセスを提供する。最初に特定のトランザクションについて書込みロックが取得されると、特定のオブジェクトからシャドウコピーが作成される。シャドウコピーへの書込みが行われる。コミット中に特定のトランザクションの妥当性を確認した後、コンテンツがシャドウコピーから特定のオブジェクトに書き込まれる。

【0011】

トランザクションがコミットされる順序がトランザクションの抽象的直列化順序 (abstract serialization order) に一致することを保証する、トランザクション順序付けプロセスを提供する。コミット処理に入った第1の数のトランザクションを表す、入力コミットカウンタ (entered commit counter) を提供する。コミット処理を終了した第2の数のトランザクションを表す、終了コミットカウンタ (exited commit counter) が提供される。特定のトランザクションがコミット処理に入ると、システムは、入力コミットカウン

タをアトミックに読み込み、インクリメントして、入力コミットカウンタから読み込まれた値を特定のトランザクションのチケット番号として割り当てる。終了コミットカウンタが特定のトランザクションのチケット番号に一致するとき、特定のトランザクションはコミット処理を終了する。

【 0 0 1 2 】

この課題を解決するための手段は、以下の発明を実施するための形態においてさらに詳細に説明される概念の選択を簡略化した形で紹介するために提供したものである。この課題を解決するための手段は、特許請求される対象について重要な特徴または本質的な特徴を特定することは意図されておらず、特許請求される対象の範囲を決定する際の補助として使用されることも意図されていない。

【 図面の簡単な説明 】

【 0 0 1 3 】

【 図 1 】一実施形態のコンピュータシステムを示す概略図である。

【 図 2 】図 1 のコンピュータシステム上で動作する一実施形態のソフトウェアトランザクショナルメモリ・アプリケーションを示す概略図である。

【 図 3 】図 1 のシステムの一実施形態に関する高度なプロセスのフロー図である。

【 図 4 】ソフトウェアトランザクショナルメモリシステムにおいてバッファリングされた書込みを使用するのに必要とされる高度な段階を示す、図 1 のシステムの一実施形態のためのプロセスのフロー図である。

【 図 5 】ソフトウェアトランザクショナルメモリシステムでバッファリングされた書込みを使用するのに必要とされるより詳細な段階を示す、図 1 のシステムの一実施形態のためのプロセスのフロー図である。

【 図 6 】トランザクショナルメモリワード内のバッファリングされた書込み情報をトラッキングするのに必要とされる段階を示す、図 1 のシステムの一実施形態のためのプロセスのフロー図である。

【 図 7 】共有メモリ内のオブジェクトのシャドウコピーおよびオブジェクトへのリファレンスを保持するトランザクショナルメモリワードを示す論理図である。

【 図 8 】トランザクションが抽象的直列化順序と同じ順序でコミットすることを保証するのに必要とされる高度な段階を示す、図 1 のシステムの一実施形態のためのプロセスのフロー図である。

【 図 9 】トランザクションが抽象的直列化順序と同じ順序でコミットすることを保証するのに必要とされるより詳細な段階を示す、図 1 のシステムの一実施形態のためのプロセスのフロー図である。

【 発明を実施するための形態 】

【 0 0 1 4 】

本発明の原理の理解を促進する目的で、図面に図示されている実施形態を参照し、同様に説明するために特定の文言を使用する。しかしながら、このことによって、その範囲を限定することは意図されていないことは理解されよう。説明される実施形態の全ての代替形態およびさらなる変更形態、ならびに本明細書に記載される原理の他のいかなる利用形態も、当業者であれば通常は考え付くであろう。

【 0 0 1 5 】

本システムは、一般的なコンテキストではソフトウェアトランザクショナルメモリシステムとして説明されるが、本システムは、これらに加えて他の目的でも機能する。一実施形態では、本明細書に記載される技法の 1 つまたは複数を、MICROSOFT（登録商標）、NETフレームワークなどのフレームワークプログラムにおける機能として、あるいは、開発者がソフトウェアアプリケーションを開発するためのプラットフォームを提供する任意の他のタイプのプログラムまたはサービスからの機能として実装することができる。別の実施形態では、本明細書に記載の技法の 1 つまたは複数は、並列環境で実行するアプリケーションの開発に取り組む他のアプリケーションの機能として実装される。

【 0 0 1 6 】

一実施形態では、バッファリングされた書込みプロセスを使用して、オブジェクトのシャドウコピーへの書込みと、コミット中に個々のトランザクションの妥当性を確認した後に、コンテンツのオブジェクトへの書き戻しとを行う、ソフトウェアトランザクショナルメモリシステムを提供する。別の実施形態では、トランザクションがコミットされる順序がトランザクションの抽象的直列化順序に一致することを保証するトランザクション順序付けプロセスを提供する。トランザクションは、プログラマにトランザクションが順次それぞれ別々に実行するという錯覚を与えることにより、並列プログラミングを簡単にする。抽象的直列化順序とは、単に、トランザクションが実行するために現われる順序である。この順序は、トランザクションによって行われる読み込みと書き込みによって決定される。 $T \times 1$ および $T \times 2$ が、両方ともコミットするトランザクションであり、 $T \times 2$ による読み込みが $T \times 1$ によって書き込まれた値を観察した場合は、 $T \times 1$ は、抽象的直列化順序で $T \times 2$ に先行しなければならない。それぞれが、他方が書き込んだ値を読み込んだ場合、それらを直列化することはできないため、両方ともコミットすることができず、したがって、少なくとも1つが途中停止しなければならない。両方がコミットしたとき、それぞれが他方に先行しなければならない場合は、1つが途中停止しなければならない。バッファリングされた書込みプロセスおよび/またはトランザクション順序付けプロセスを使用することにより、私有化関連の問題のうちの少なくとも一部は回避される。

10

【0017】

図1に示されているように、システムの1つまたは複数の部分を実装するために使用する例示的なコンピュータシステムは、コンピューティング装置100などのコンピューティング装置を含む。その最も基本的な構成では、コンピューティング装置100は、典型的に、少なくとも1つの処理ユニット102およびメモリ104を含む。コンピューティング装置の正確な構成およびタイプに応じて、メモリ104は、揮発性(RAMなど)、不揮発性(ROM、フラッシュメモリなど)、またはこの2つの何らかの組合せとすることができる。この最も基本的な構成は、図1では破線106で図示されている。

20

【0018】

さらに、装置100は、追加の特徴/機能を有することもできる。例えば、装置100は、これらには限られないが、磁気または光ディスクまたはテープを含め、追加のストレージ(取外し可能および/または取外し不可能)を含むこともできる。このような追加のストレージは、図1では、取外し可能なストレージ108および取外し不可能なストレージ110で図示されている。コンピュータ記憶媒体は、コンピュータ読取可能な命令、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の記憶のために任意の方法または技術で実装される、揮発性および不揮発性媒体、取外し可能な媒体および取外し不可能な媒体を含む。メモリ104、取外し可能なストレージ108、および取外し不可能なストレージ110は全て、コンピュータ記憶媒体の例である。コンピュータ記憶媒体は、RAM、ROM、EEPROM、フラッシュメモリもしくは他のメモリ技術、CD-ROM、デジタル多用途ディスク(DVD)もしくは他の光ストレージ、磁気カセット、磁気テープ、磁気ディスクストレージもしくは他の磁気ストレージ装置、または所望の情報を記憶するために使用可能であって、装置100がアクセス可能な他の任意の媒体も含むが、これらには限定されない。このようないかなるコンピュータ記憶媒体も、装置100の一部とすることができる。

30

40

【0019】

コンピューティング装置100は、コンピューティング装置100が他のコンピュータ/アプリケーション115と通信できるようにする、1つまたは複数の通信接続114を含む。装置100は、キーボード、マウス、ペン、音声入力装置、タッチ入力装置などの入力装置112を有することもできる。また、ディスプレイ、スピーカ、プリンタなどの出力装置111を含むこともできる。これらの装置は、当技術分野ではよく知られているので、ここで詳細に議論する必要はない。一実施形態では、コンピューティング装置100は、ソフトウェアトランザクショナルメモリ・アプリケーション(software transactional memory application)200を含む。ソフトウェアトランザクショナル

50

メモリ・アプリケーション 200 について、図 2 でさらに詳細に説明する。

【0020】

ここで、図 1 を参照しながら図 2 を参照すると、コンピューティング装置 100 上で動作するソフトウェアトランザクショナルメモリ・アプリケーション 200 が図示されている。ソフトウェアトランザクショナルメモリ・アプリケーション 200 は、コンピューティング装置 100 上に存在するアプリケーションプログラムの 1 つである。しかし、ソフトウェアトランザクショナルメモリ・アプリケーション 200 は、代替または追加的に、1 つまたは複数のコンピュータ上で、および / または図 1 に示されているものとは異なる変形形態において、コンピュータ実行可能命令として具現化され得ることが理解されよう。代替または追加として、ソフトウェアトランザクショナルメモリ・アプリケーション 200 の 1 つまたは複数の部分は、他のコンピュータおよび / またはアプリケーション 115 におけるシステムメモリ 104 の一部分、あるいはコンピュータソフトウェアの技術分野において考える他の変形形態におけるシステムメモリ 104 の一部分とすることができる。

10

【0021】

ソフトウェアトランザクショナルメモリ・アプリケーション 200 は、本明細書で説明される技法の一部または全てを実施する役割を担うプログラムロジック 204 を含む。プログラムロジック 204 は、次のロジックを含む。すなわち、複数のトランザクションを管理するように動作可能なソフトウェアトランザクショナルメモリ・システムを提供するためのロジック 206 と、オブジェクトのシャドウコピーへの書込みと該シャドウコピーへの読み込みを行い、コミット中に複数のトランザクションの個々のトランザクションの妥当性を確認した後でコンテンツをオブジェクトに書き戻す、バッファリングされた書込みプロセスを提供するためのロジック 208 と、複数のトランザクションがコミットされる順序が複数のトランザクションの抽象的直列化順序と一致することを保証する、トランザクション順序付けプロセスを提供するためのロジック 210 と、書込みロックが最初に取得されたとき、バッファリングされた書込みプロセスが、オブジェクトのうちの特定の 1 つのオブジェクトの特定のシャドウコピーを作成することと、シャドウコピーがすでに存在する場合はそのシャドウコピーにアクセスすることとを可能にするためのロジック 212 と、トランザクション順序付けプロセスが、グローバルデータ構造を使用してコミットの順序（例えば、入力コミットカウンタと終了コミットカウンタ）を保証することを可能にするためのロジック 214 と、トランザクション順序付けプロセスが、個々のトランザクションがコミット処理に入ると入力コミットカウンタをアトミックに読み込み、インクリメントし、入力コミットカウンタから読み込んだ値を、個々のトランザクションのチケット番号として割り当てることを可能にするためのロジック 216 と、トランザクション順序付けプロセスが、個々のトランザクションがコミット処理後の続行を許可される前は、終了コミットカウンタが個々のトランザクションのチケット番号に一致するまで、該個々のトランザクションを待機させることを可能にするためのロジック 218 と、バッファリングされた書込みプロセスおよび / またはトランザクション順序付けプロセスを使用することによって私有化関連問題を回避するためのロジック 220 と、アプリケーションを動作するための他のロジック 222 とを含む。一実施形態では、プログラムロジック 204 は、プログラムロジック 204 内のプロシージャに対するシングルコールを使用することなど、別のプログラムからプログラマ的に呼び出されるように動作することができる。

20

30

40

【0022】

次に、図 1 ~ 2 を参照し続けながら図 3 ~ 9 を参照すると、ソフトウェアトランザクショナルメモリ・アプリケーション 200 の 1 つまたは複数の実施形態を実装するための複数の段階がさらに詳細に説明されている。図 3 は、ソフトウェアトランザクショナルメモリ・アプリケーション 200 に関する高度なプロセスのフロー図である。一形態において、図 3 のプロセスは、コンピューティング装置 100 のオペレーティングロジックで少なくとも部分的に実装される。本プロセスは、開始点 240 で始まり、複数のトランザクションを管理するように動作可能なソフトウェアトランザクショナルメモリシステムを提供

50

する（段階 2 4 2）。オブジェクトのシャドウコピーへの書込みを行い、複数のトランザクションの個々のトランザクションの妥当性を確認した後でコンテンツをオブジェクトに書き戻す、バッファリングされた書込みプロセスが提供される（段階 2 4 4）。複数のトランザクションがコミットされる順序が複数のトランザクションの抽象的直列化順序に一致することを保証する、トランザクション順序付けプロセスが提供される（段階 2 4 6）。バッファリングされた書込みプロセスおよび / またはトランザクション順序付けプロセスを提供することにより、一部の私有化関連問題を回避することができる（段階 2 4 8）。本プロセスは、終了点 2 5 0 で終了する。

【 0 0 2 3 】

図 4 は、バッファリングされた書込みをソフトウェアトランザクショナルメモリシステムで使用するのに必要とされる高度な段階の一実施形態を図示している。一形態では、図 4 のプロセスは、コンピューティング装置 1 0 0 のオペレーティングロジックで少なくとも部分的に実装される。本プロセスは、開始点 2 6 0 で始まり、複数のトランザクションを管理するように動作可能なソフトウェアトランザクショナルメモリシステムを提供する（段階 2 6 2）。特定のトランザクションについて書込みロックが最初に取得されると、特定のオブジェクトのシャドウコピーが作成される（段階 2 6 4）。シャドウコピーへの書込みと該シャドウコピーからの読み込みが行われる（段階 2 6 6）。下記は、シャドウコピーを使用する一実施形態において、どのように読み込みが行われるかについての説明であるが、他の読み込みおよび書込みのバリエーションを使用することもできよう。読み込みが、書込みロックされていない（non-write-locked）オブジェクトから行われる場合は、読み込みは該オブジェクトから直接行われる。オブジェクトが書込みロックされている場合は、（例えば、トランザクショナルメモリワード内の）該ロックは、何らかのトランザクションのログ内のシャドウコピーをポイントする。該シャドウコピーが何らかの他のトランザクションのログ内にある場合は、衝突が起こるため、何らかの競合管理アクション（contention management action）を取る必要がある。該ログが現在のトランザクションのログの場合は、読み込みは、現在のトランザクションがポイントしているシャドウコピーから行われる。コミット中に特定のトランザクションの妥当性を確認した後、コンテンツが、シャドウコピーから特定のオブジェクトに書き込まれる（段階 2 6 8）。本プロセスは、終了点 2 7 0 で終了する。

【 0 0 2 4 】

図 5 は、バッファリングされた書込みをソフトウェアトランザクショナルメモリシステムで使用するのに必要とされるより詳細な段階の一実施形態を図示している。一形態では、図 5 のプロセスは、コンピューティング装置 1 0 0 のオペレーティングロジックで少なくとも部分的に実装される。本プロセスは、開始点 2 8 0 で始まり、トランザクションを実行しているスレッドが、値 V をオブジェクト O 内のあるロケーションに書き込むことを試みる（段階 2 8 1）。したがって、トランザクションは、オブジェクト O の書込みロックを取ることを望む（段階 2 8 2）。オブジェクトが書込みロックされていない場合は（判断点 2 8 4）、書込みロックが取得され、トランザクションは、シャドウコピーを作成し、該シャドウコピーへのリファレンスを保持する（段階 2 8 7）。オブジェクトが書込みロックされている場合（判断点 2 8 4）、書込みロックが現在のトランザクションによって保持されているときは（判断点 2 8 5）、該トランザクションは、それへのリファレンスを保持する（段階 2 8 8）。これらのイベントのいずれにおいても、トランザクションは、値 V を O のシャドウコピーに書き込む（段階 2 9 0）。しかし、オブジェクトが書込みロックされている場合（判断点 2 8 4）、書込みロックが現在のトランザクションによって保持されておらず、代わりに別のトランザクションによって保持されているとき（判断点 2 8 4）、競合管理アクションが取られる（段階 2 8 6）。本プロセスは、終了点 2 9 2 で終了する。

【 0 0 2 5 】

図 6 は、トランザクショナルメモリワード内のバッファリングされた書込み情報をトラッキングするのに必要とされる段階の一実施形態を図示している。一形態では、図 6 のプ

10

20

30

40

50

ロセスは、コンピューティング装置 100 のオペレーティングロジックで少なくとも部分的に実装される。本プロセスは、開始点 310 で始まり、トランザクショナルメモリワードを特定のオブジェクトのオブジェクトヘッダ内に提供する（段階 312）。図 7 でより詳細に説明されるように、ロックタイプおよびバージョンがトランザクショナルメモリワードを使用してトラッキングされる（段階 314）。書き込みロックされている状態では、トランザクショナルメモリワードは、シャドウコピーを含んでいる書き込みログエントリへのポインタと、オブジェクトのベースアドレスへのポインタとを保持する（段階 316）。シャドウコピーへの書き込みと該シャドウコピーからの読み込みが行われ、トランザクションの妥当性を確認した後、コミット処理においてシャドウコピーの値がオブジェクトに書き戻される（段階 318）。本プロセスは、終了点 320 で終了する。

10

【0026】

図 7 は、オブジェクトがトランザクションログ 406 を所有しているトランザクションによって書き込みロックされていることを示す書き込みログエントリ 408 へのリファレンスを保持している、トランザクショナルメモリワード（transactional memory word）404 を図示している論理図である。書き込みログエントリは、固定サイズのヘッダ 410 で始まり、該ヘッダ 410 は、書き込みロックされる前のオブジェクトの TMW の値である元の TMW 412、およびオブジェクトベース 414、書き込みロックされたオブジェクトへ戻るポインタを含む。該ヘッダの後には、オブジェクトヘッダ 402 に続くオブジェクト 400 のコンテンツで初期化されるシャドウコピー 416 が続く。トランザクションは、このシャドウコピーとの間で書き込みと読み込みを行い、トランザクションがコミットしたときにその更新されたコンテンツをオブジェクト 400 に書き戻す。別の実施形態では、トランザクショナルメモリワードがシャドウコピーへの直接のポインタを含むことがあり、この場合、書き込みログエントリの開始へのポインタを、一定であるヘッダのサイズを引くことによって回復することができる。トランザクショナルメモリワード 404 は、オブジェクト 400 のオブジェクトヘッダ 402 にある。

20

【0027】

オブジェクトが書き込みロックされていないとき、トランザクショナルメモリワードは、バージョン番号、ならびにリーダ（readers）（例えば、悲観的リーダ）のリスト/カウントおよび/またはインジケータを含む。オブジェクトが書き込みロックされているとき、トランザクショナルメモリワードの書き込みロックエントリは、オブジェクトが書き込みロ

ックされる前に読み込んだ TMW 値を保持する。一実施形態では、リーダのリスト/カウントおよび/またはインジケータは、所与の時点で特定の値にアクセスしているリーダ（例えば、悲観的）の数のカウントを含む。別の実施形態では、リーダのリスト/カウントおよび/またはインジケータは、所与の時点で特定の値にアクセスしている特定のリーダ（例えば、悲観的）のリストを含む。さらに別の実施形態では、リーダのリスト/カウントおよび/またはインジケータは、単に、所与の時点で特定の値にアクセスしている 1 つまたは複数のリーダ（例えば、悲観的）が存在することを示すためのフラグまたは他のインジケータである。他のトランザクショナルメモリワードの属性および/またはそれらの組合せを代替または追加的に他の実施形態で使用して、並列処理制御の決定をするトランザクショナルメモリシステムによって使用するためにバージョン番号、書き込みロックなどを示すことができることが認識されよう。

30

40

【0028】

図 8 は、トランザクションが抽象的直列化順序と同じ順序でコミットすることを保証するのに必要とされる高度な段階の一実施形態を図示している。一形態では、図 8 のプロセスは、コンピューティング装置 100 のオペレーティングロジックで少なくとも部分的に実装される。本プロセスは、開始点 420 で始まり、複数のトランザクションを管理するように動作可能なソフトウェアトランザクショナルメモリシステムを提供する（段階 422）。コミット処理に入ったトランザクションの数を表す入力コミットカウンタが提供される（段階 424）。コミット処理を終了したトランザクションの数を表す終了コミットカウンタが提供される（段階 426）。特定のトランザクションは、コミット処理に入る

50

と、入力コミットカウンタをアトミックに読み込み、インクリメントし、入力コミットカウンタから読み込まれた値を特定のトランザクションのチケット番号として割り当てる（段階 4 2 8）。終了コミットカウンタが特定のトランザクションのチケット番号に一致するとき、特定のトランザクションはコミット処理を終了することを許可される（段階 4 3 0）。本プロセスは終了点 4 3 2 で終了する。

【 0 0 2 9 】

図 9 は、トランザクションが抽象的直列化順序と同じ順序でコミットすることを保証するのに必要とされるより詳細な段階の一実施形態を図示している。一形態では、図 9 のプロセスは、コンピューティング装置 1 0 0 のオペレーティングロジックで少なくとも部分的に実装される。本プロセスは、開始点 4 5 0 で始まり、トランザクションを開始する（段階 4 5 2）。トランザクションは、読み込みおよび書込みを行い、制御フローを実行する（段階 4 5 3）。特定のトランザクションが読み込みのみ（read-only）の場合は（判断点 4 5 4）、読み込みのみのトランザクションは、抽象的直列化順序と同じ順序でコミットを終了することに制限されず（段階 4 5 6）、コミットすることができるかどうか確かめるために読み込みの妥当性確認を続けることができる（段階 4 7 1）。このとき、本プロセスは終了点 4 7 2 で終了する。

【 0 0 3 0 】

トランザクションが読み込みのみでない場合（判断点 4 5 4）、トランザクションは、抽象的直列化順序と同じ順序でコミットを終了するように制限されなければならない（段階 4 5 8）。トランザクションは、グローバルな入力コミットカウンタをアトミックに読み込み、インクリメントして、入力コミットカウンタから読み込まれた値を特定のトランザクションのチケット番号として割り当てる（段階 4 6 0）。トランザクションは、コミット処理を行う（段階 4 6 4）。一実施形態では、コミット処理中に次のステップが実行される。すなわち、（１）楽観的読み込みの妥当性を確認し、ロケーション読み込みがまだ書き込みロックされていないことと、T M W が初めに読み込まれたときと同じバージョン番号をそれらが保持することとを保証するステップと、（２）シャドウコピーを書き戻すステップである。トランザクションは、グローバルな終了コミットカウンタがローカルチケット番号に一致するときのみコミット処理を終了して、トランザクションは、終了コミットカウンタをインクリメントする（段階 4 6 8）。トランザクションがコミット処理を終了した後に、書き込みロックおよび悲観的読み込みロックが外される（段階 4 7 0）。本プロセスは、終了点 4 7 2 で終了する。

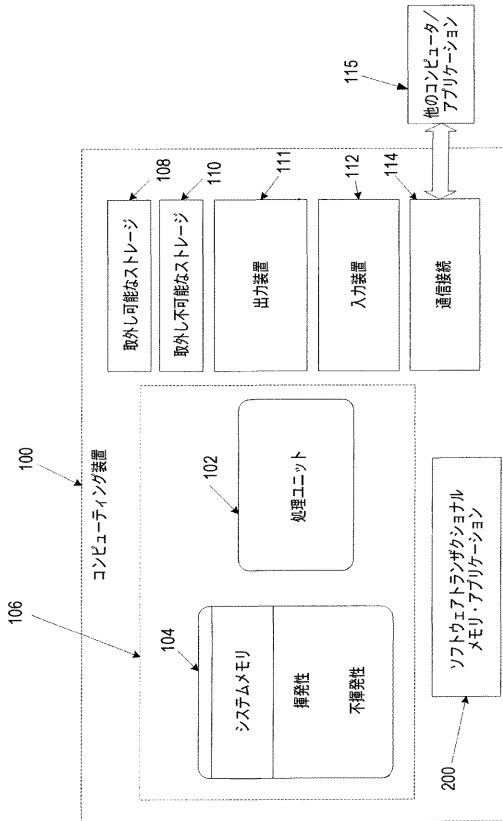
【 0 0 3 1 】

本主題を、構造的特徴および／または方法論的アクトに特有の言語で説明してきたが、添付の特許請求の範囲によって定義される本主題は、必ずしも上述の特有の特徴またはアクトに限定されないことを理解されたい。むしろ、上述の特有の特徴および行為は、特許請求の範囲を実装する例示的な形態として開示されている。本明細書で説明された実施形態および／または添付の特許請求の範囲の精神に含まれる均等な形態、変更形態、および修正形態は全て保護されることが望ましい。

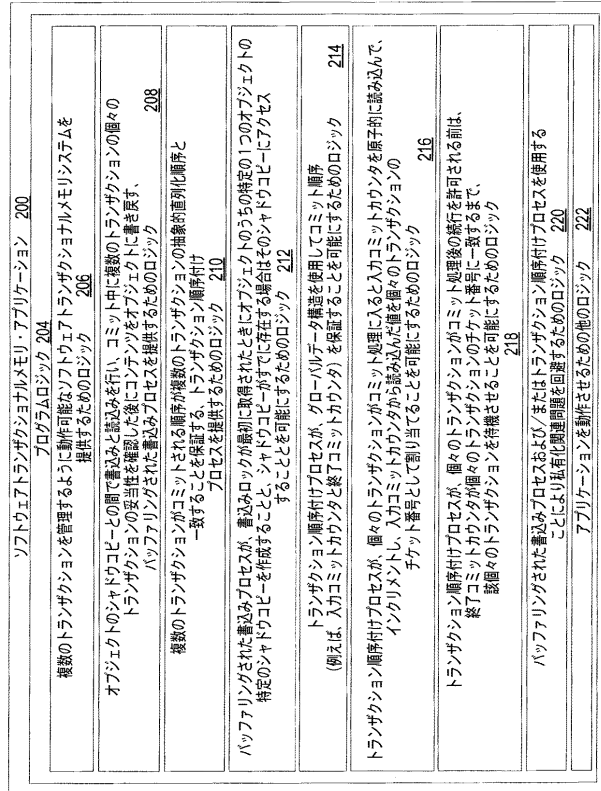
【 0 0 3 2 】

例えば、コンピュータソフトウェア技術分野の当業者には、本明細書で議論された実施例で説明されたようなクライアントおよび／またはサーバ構成、ユーザインターフェーススクリーンコンテンツ、および／またはデータレイアウトを、１つまたは複数のコンピュータにおいて異なるように構成し、実施例で描写されたものより少ないまたは追加のオプションまたは特徴を含むこともできることが理解されよう。

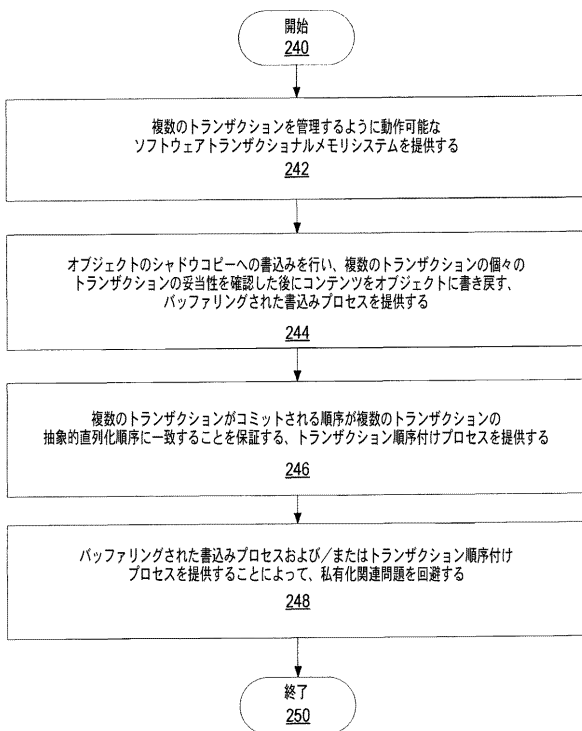
【図 1】



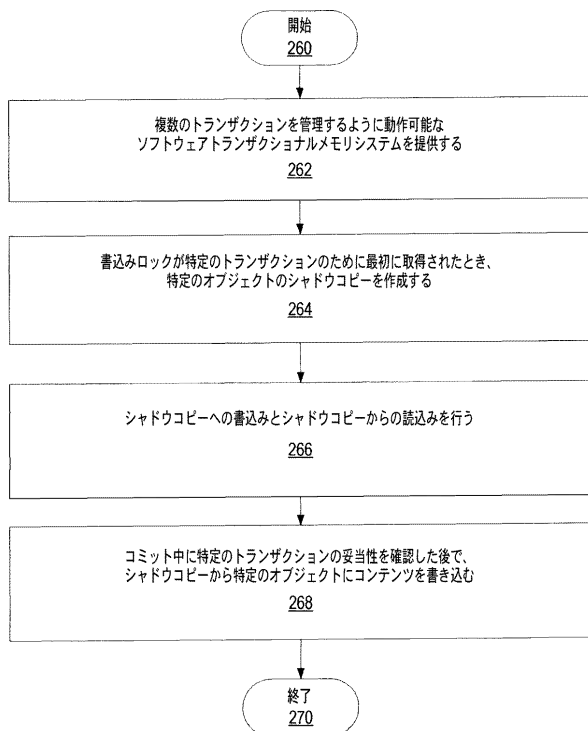
【図 2】



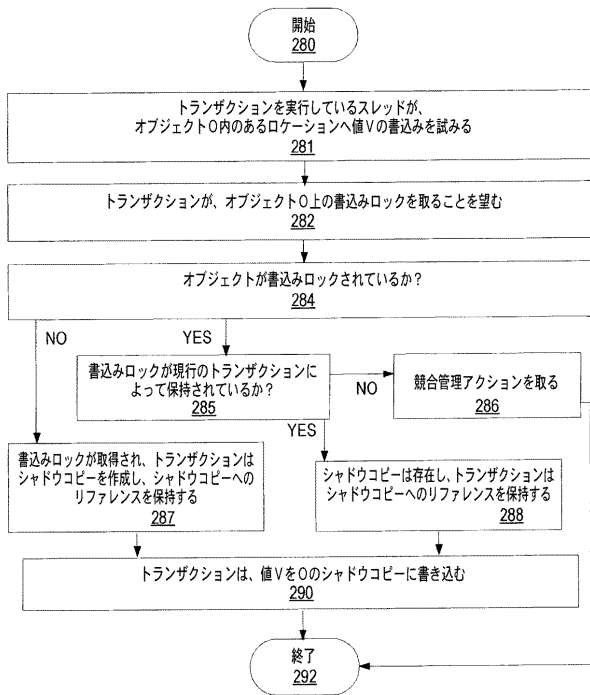
【図 3】



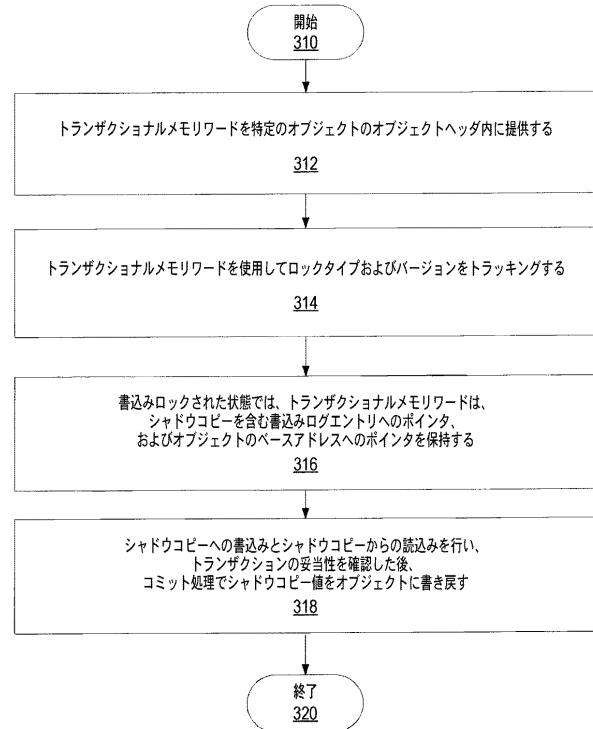
【図 4】



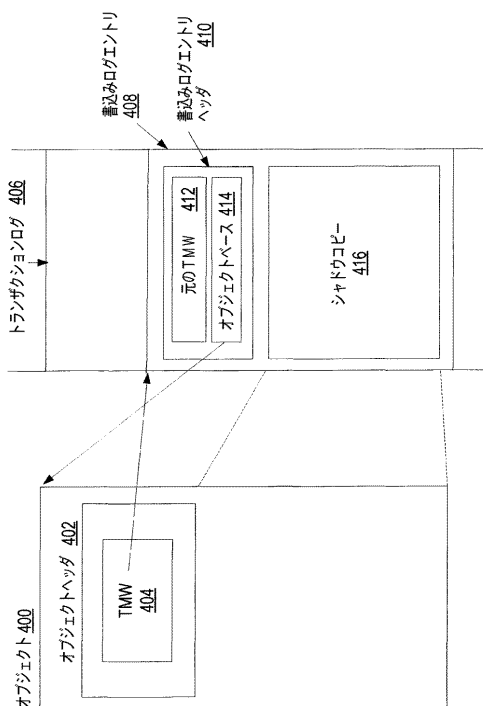
【図 5】



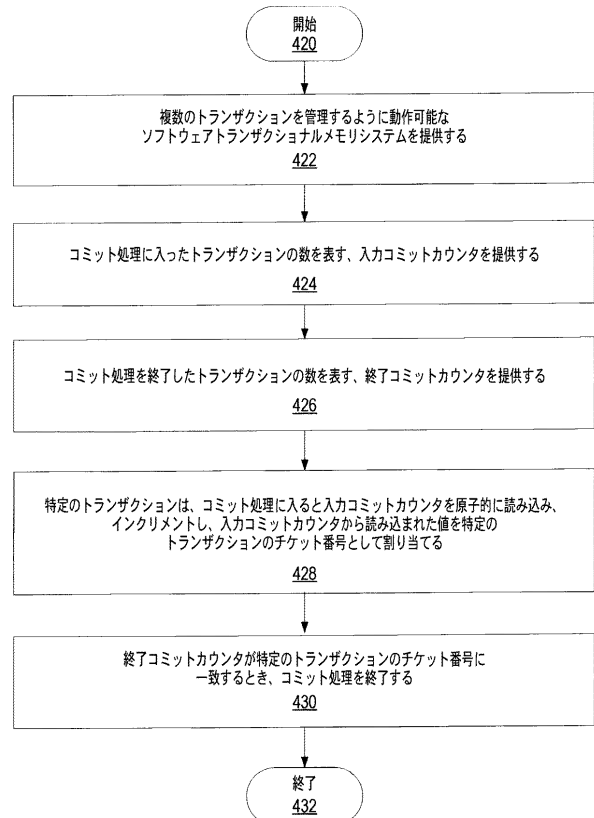
【図 6】



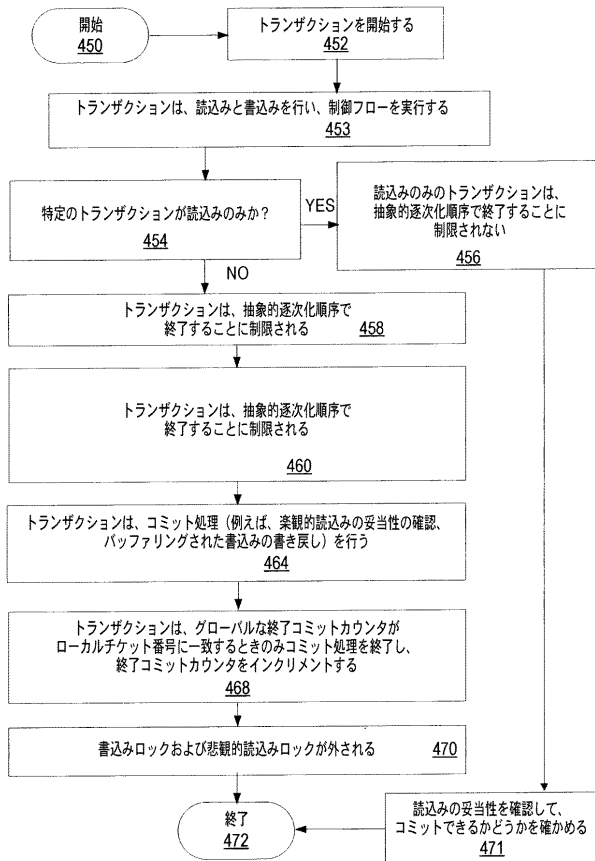
【図 7】





【図 8】



【 図 9 】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2008/056873
A. CLASSIFICATION OF SUBJECT MATTER		
<i>G06F 12/14(2006.01)i, G06F 12/00(2006.01)i</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) IPC 8 : G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models since 1975 Japanese utility models and applications for utility models since 1975		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKIPASS(Kipo Internal), Google, YesKisti keywords: software, buffer*, shadow, transaction*, write, read, copy, memory order*, serial*		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Hammond, L. et al. 'Transactional memory coherence and consistency.' In: Proceedings on Computer Architecture, 31st Annual International Symposium. Edited by IEEE Computer Society, 2004. pages 102-113. See section 2.1.4.1 and figure 3 c).	1-20
Y	Woo, S.K. et al. An effective recovery under fuzzy checkpointing in main memory databases. Information and Software Technology. 25 February 2000, Vol. 42, No. 3, 25, pages 185-196 See section 2.3.	1-20
A	Moore, K.E. et al. 'LogTM: log-based transactional memory.' In: High-Performance Computer Architecture, 12-th International Symposium, February 2006, pages 254-265. See section 1, Introduction.	1-20
A	Garcia-Molina, H. et al. Main memory database systems: an overview. IEEE Transactions on Knowledge and Data Engineering, Dec. 1992, Vol. 4, No. 6, pages 509-516. See section II F, recovery.	1-20
A	US 2004/0225823 A1 (HEYNEMANN, T.A. et al.) 11 NOVEMBER 2004 See Summary.	1-20
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 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 15 JULY 2008 (15.07.2008)		Date of mailing of the international search report 15 JULY 2008 (15.07.2008)
Name and mailing address of the ISA/KR  Korean Intellectual Property Office Government Complex-Daejeon, 139 Seonsa-ro, Seo-gu, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140		Authorized officer YOON, Hye Sook Telephone No. 82-42-481-8370 

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2008/056873

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004-0225823 A1	11.11.2004	NONE	

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(72)発明者 ジョン ジョセフ ダフィー

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 インターナショナル パテント内

(72)発明者 マイケル エム・マグルーダー

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 インターナショナル パテント内

(72)発明者 ゴッツ グレーフェ

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 インターナショナル パテント内

(72)発明者 ビノッド ケー・グローバー

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 インターナショナル パテント内

(72)発明者 ティモシー ローレンス ハリス

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 インターナショナル パテント内

F ターム(参考) 5B013 CC11