(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5091481号
(P5091481)
(45)発行日　平成24年12月5日(2012.12.5)　(24)登録日　平成24年9月21日(2012.9.21)

(51)Int.Cl.　　　　　　　　　　F I
G06F 9/38 (2006.01)　　G06F 9/38 310A
　　　　　　　　　　　　　　　G06F 9/38 310F
　　　　　　　　　　　　　　　G06F 9/38 350B

請求項の数 18 (全 22 頁)

(21)出願番号　特願2006-533522(P2006-533522)
(86)(22)出願日　平成16年6月2日(2004.6.2)
(65)公表番号　特表2007-503661(P2007-503661A)
(43)公表日　平成19年2月22日(2007.2.22)
(86)国際出願番号　PCT/US2004/017096
(87)国際公開番号　WO2004/111839
(87)国際公開日　平成16年12月23日(2004.12.23)
審査請求日　平成19年6月1日(2007.6.1)
(31)優先権主張番号　10/458,457
(32)優先日　平成15年6月10日(2003.6.10)
(33)優先権主張国　米国(US)

前置審査

(73)特許権者　591016172
アドバンスト・マイクロ・ディバイシズ・インコーポレイテッド
ＡＤＶＡＮＣＥＤ　ＭＩＣＲＯ　ＤＥＶＩＣＥＳ　ＩＮＣＯＲＰＯＲＡＴＥＤ
アメリカ合衆国、９４０８８－３４５３　カリフォルニア州、サニィベイル、ピィ・オゥ・ボックス・３４５３、ワン・エイ・エム・ディ・プレイス、メイル・ストップ・６８（番地なし）
(74)代理人　100108833
弁理士　早川　裕司

最終頁に続く

(54)【発明の名称】リプレイ機構を備えた読み出し／書き込みユニット

(57)【特許請求の範囲】

【請求項１】

マイクロプロセッサであって、

オペレーションを発行するように構成されたスケジューラと、

前記スケジューラによって発行されたメモリオペレーションを受信するように結合され、前記メモリオペレーションを実行するように構成された読み出し／書き込みユニットとを含み、

前記読み出し／書き込みユニットは、前記読み出し／書き込みユニットに発行された複数のメモリオペレーションを識別する情報を記憶するように構成され、前記読み出し／書き込みユニットは、前記複数のメモリオペレーションの１つに関して前記マイクロプロセッサ内で行われた予測に基づいてのデータ推測が誤っていたことが検出されたことに応答して、前記スケジューラに対して、前記読み出し／書き込みユニット内の複数のメモリオペレーションの少なくとも１つを再発行しなければならないことを示すリプレイ指示を与えるように構成され、前記データ推測は、メモリオペレーションのオペランドの値を予測すること、及び／又はメモリオペレーションの結果値を予測することを含み、

前記スケジューラは、前記読み出し／書き込みユニットからの前記指示に応答して、前記複数のメモリオペレーションのうち前記少なくとも１つを再発行するよう構成される、

マイクロプロセッサ。

【請求項２】

前記読み出し／書き込みユニットは、前記複数のメモリオペレーションの１つのアドレ

スと合致するアドレスを有する複数のメモリオペレーションの各々を識別することによって、前記リプレイの指示を生成するように構成される、
　請求項１記載のマイクロプロセッサ。
【請求項３】
　前記読み出し／書き込みユニットは、複数のメモリオペレーションの１つのアドレスの推測値、又は前記複数メモリオペレーションの前記１つの前記アドレスの前記推測値とは異なる値であって前記推測値より最近に生成された新たな値の何れかと合致するアドレスを有する前記複数のオペレーションの各々を識別することによって、前記リプレイ指示を生成するように構成される、
　請求項１記載のマイクロプロセッサ。
【請求項４】
　前記読み出し／書き込みユニットは、読み出しオペレーションであってかつ前記複数のメモリオペレーションの前記１つのアドレスと合致するアドレスを有する前記複数のメモリオペレーションの各々を識別することによって、前記リプレイ指示を生成するよう構成される、
　請求項１記載のマイクロプロセッサ。
【請求項５】
　前記読み出し／書き込みユニットは、前記複数のメモリオペレーションに含まれる読み出しオペレーションのうちいずれの読み出しオペレーションが書き込みオペレーションのデータを転送したかを追跡するように構成され、書き込みオペレーションのアドレスが誤って推測されたものとして検出された場合、前記読み出し／書き込みユニットは、前記書き込みオペレーションからデータを転送した前記読み出しオペレーションのより最近のオペレーションのいずれをも識別することによって、前記リプレイ指示を生成するよう構成される、
　請求項１記載のマイクロプロセッサ。
【請求項６】
　前記読み出し／書き込みユニットは、前記複数のメモリオペレーションの前記一つが発生する前に前記複数のメモリオペレーションのそれぞれを識別することによって、前記リプレイ指示を生成するよう構成される、
　請求項１記載のマイクロプロセッサ。
【請求項７】
　前記読み出し／書き込みユニットは、データ推測が誤っていたことが検出された、前記複数のメモリオペレーションの前記一つを識別することで前記リプレイ指示を生成するよう構成される、
　請求項１記載のマイクロプロセッサ。
【請求項８】
　前記読み出し／書き込みユニットは、前記複数のメモリオペレーションの前記一つのアドレスの推測値を、前記複数のメモリオペレーションの前記一つのアドレスの前記推測値とは異なる値であって前記推測値より最近に生成された新しい値と比較することで、前記データ推測の誤りを検出するよう構成される、
　請求項１記載のマイクロプロセッサ。
【請求項９】
　前記読み出し／書き込みユニットは、読み出しオペレーションの推測結果と、前記読み出しオペレーションの非推測結果と、を比較することで、前記データ推測の誤りを検出するよう構成される、
　請求項１記載のマイクロプロセッサ。
【請求項１０】
　コンピュータシステムであって、
　メモリと、
　前記メモリに結合されたプロセッサとを含み、

前記プロセッサは、メモリオペレーションを発行するよう構成されたスケジューラと、前記スケジューラによって発行されたメモリオペレーションを受信するよう結合されて前記メモリオペレーションを実行するよう構成された読み出し／書き込みユニットと、を含み、

前記読み出し／書き込みユニットは、当該読み出し／書き込みユニットに発行された複数のメモリオペレーションを識別する情報を記憶するよう構成され、前記複数のメモリオペレーションの１つに関して前記プロセッサ内で行われたデータ推測が誤っていたことが検出されたことに応答して、当該読み出し／書き込みユニット内の前記複数のメモリオペレーションの少なくとも１つが再発行されなければならないこと示すリプレイ指示を前記スケジューラに与えるよう構成され、　前記データ推測は、メモリオペレーションのオペランドの値を予測すること、及び／又はメモリオペレーションの結果値を予測することを含み、

前記スケジューラは、前記読み出し／書き込みユニットからの前記指示に応答して、前記複数のメモリオペレーションの少なくとも一つを再発行するように構成され、前記スケジューラは前記読み出し／書き込みユニットから分離されている、

コンピュータシステム。

【請求項１１】

スケジューラが読み出し／書き込みユニットにオペレーションを発行して当該オペレーションを実行させ、

前記読み出し／書き込みユニットが、前記オペレーションに実行されたデータ推測が誤っていることを検出し、

前記読み出し／書き込みユニットは前記スケジューラから分離しており、前記データ推測は、メモリオペレーションのオペランドの値を予測すること、及び／又はメモリオペレーションの結果値を予測することを含み、前記検出に応答して、前記読み出し／書き込みユニットが、現在前記読み出し／書き込みユニットに発行されている少なくとも１つのオペレーションを識別するリプレイ指示を生成し、

前記生成に応じて、前記スケジューラが、前記読み出し／書き込みユニットに前記少なくとも一つのオペレーションを再発行する、

方法。

【請求項１２】

前記生成では、前記読み出し／書き込みユニットが、前記データ推測が誤っていたオペレーションよりもより最近の前記読み出し／書き込みユニット内の全ての未済のオペレーションを識別して、識別された当該オペレーションをリプレイする、１つあるいはそれ以上のリプレイ指示を生成する、

請求項１１記載の方法。

【請求項１３】

前記生成では、前記読み出し／書き込みユニットが、前記データ推測が誤っていた前記オペレーションのアドレスと合致するアドレスを有する前記読み出し／書き込みユニット内の未済のオペレーションを識別して、識別された当該オペレーションをリプレイする、１つあるいはそれ以上のリプレイ指示を生成する、

請求項１１記載の方法。

【請求項１４】

前記生成では、前記読み出し／書き込みユニットが、前記データ推測が誤っていたオペレーションのアドレスの推測値と合致するアドレスを有する前記読み出し／書き込みユニット内の未済のオペレーションを識別して、識別された当該オペレーションをリプレイする、１つあるいはそれ以上のリプレイ指示を生成する、

請求項１１記載の方法。

【請求項１５】

前記読み出し／書き込みユニット内の未済の読み出しオペレーションのどのオペレーションが比較的古い読み出しオペレーションからデータを転送したかを追跡し、

前記比較的古い読み出しオペレーションのアドレスに対して行われたデータ推測が前記指示によって誤っているとして識別された場合、前記生成では、前記読み出し／書き込みユニットが、前記比較的古い読み出しオペレーションからデータを転送した読み出しオペレーションを識別して、識別された当該オペレーションをリプレイする、１つ以上のリプレイ指示を生成する、

請求項１１記載の方法。

【請求項１６】

前記生成では、前記読み出し／書き込みユニットが、前記データ推測が誤っていたオペレーションを識別して、識別された当該オペレーションをリプレイするリプレイ指示を生成する、

請求項１１記載の方法。

【請求項１７】

前記検出では、前記読み出し／書き込みユニットが、前記オペレーションのアドレスの推測値を前記オペレーションのアドレスの前記推測値とは異なる値であって前記推測値より最近に生成された新しい値と比較する、

請求項１１記載の方法。

【請求項１８】

前記検出では、前記オペレーションの推測結果を前記オペレーションの非推測結果と比較する、

請求項１１記載の方法。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、マイクロプロセッサ技術分野に関し、より具体的には、マイクロプロセッサにおいてデータの投機的実行すなわちデータスペキュレーション（data speculation）を実行する方法に関する。

【背景技術】

【０００２】

スーパースカラーマイクロプロセッサは、複数の命令を並列に実行し、その設計に適合する最短のクロックサイクルを用いることによって高パフォーマンスを発揮する。しかし、命令と命令の間におけるデータ及びコントロールフローの依存性によって、所定の時間内に発行可能な命令数が制限されてしまう。結果として、ある種のマイクロプロセッサではさらなるパフォーマンスゲイン（性能向上）を図るために投機的実行をサポートしている。

【０００３】

このような投機あるいはスペキュレーションの一つのタイプとして、コントロールフロー推測（control flow speculation）が挙げられる。コントロールフロー推測はプログラム制御が進行する方向を予測する。例えば、分岐が発生するかどうかを予測するために分岐予測が用いられる。様々なタイプの分岐予測が利用可能であり、単純に毎回同じ予測を行う方法から、履歴を基にした予測を行うために、プログラム中の過去の分岐の詳細な履歴を維持する方法まである。分岐予測は、ハードウェアの最適化、コンパイラの最適化、またはその両方によって容易になる。分岐予測機構によって提供される予測に基づいて、命令は投機的にフェッチされ、実行される。分岐命令が最終的に評価されるとき、分岐予測を識別することができる。もし予測が誤っていれば、誤った予測に基づいて投機的に実行されたすべての命令は廃棄される。

【０００４】

他の投機のタイプとしては、データ値を予測するデータ推測がある。すでに提案されたデータ推測の形式は、メモリオペレーション用にアドレスを投機的に生成し、コンピュータオペレーションで使用するためにデータ値を投機的に生成する。コントロール推測にともない、投機的にデータ値を生成するために使用された根本的な条件を最終的に評価し、

推測を認証済みとするか、あるいは推測を実行しないようにする。
【０００５】
　投機的実行においては、依存性チェックの完了を待たずに命令の実行を進めることができるので、正しい推測によって得られる効率が不正確な推測によって失われる効率を超えていれば、大幅な性能の向上が達成できる。従って、マイクロプロセッサにおいてデータ推測を実行し、誤推測に対する効率的な復旧機構を提供できるようになることが望ましい。
【発明の開示】
【発明が解決しようとする課題】
【０００６】
　データ推測のマイクロプロセッサのロードストアユニット、即ち、読み出し／書き込みユニット、においてオペレーションをリプレイする方法とシステムの様々な実施形態が開示される。幾つかの実施形態では、マイクロプロセッサはオペレーションを発行するように構成されたスケジューラと、このスケジューラによって発行されたメモリオペレーションを実行するように構成された読み出し／書き込みユニットを含み得る。読み出し／書き込みユニットは、読み出し／書き込みユニットに発行されたメモリオペレーションを識別する情報を記録するように構成される。発行されたメモリオペレーションの１つに関するデータ推測の誤りの検出に応じて、スケジューラにリプレイ指示を与えるように読み出し／書き込みユニットが構成され、読み出し／書き込みユニット内に発行されたメモリオペレーションの少なくとも一つが再発行されなければならないことを指示する。スケジューラは読み出し／書き込みユニットによって識別されたメモリオペレーションを応答可能なように再発行するよう構成される。
【課題を解決するための手段】
【０００７】
　一実施形態では、読み出し／書き込みユニットは誤って投機されたメモリオペレーションのアドレスを識別するアドレスを有するメモリオペレーションの各々をリプレイするように構成される。読み出し／書き込みユニットは、又（あるいは他の例として）、誤って投機されたメモリオペレーションの投機的アドレスを識別するアドレスを有するメモリオペレーションの各々をリプレイするように構成され得る。一実施形態では、読み出し／書き込みユニットは、誤って投機されたメモリオペレーションのアドレスと一致するアドレスを有する読み出しオペレーション（ロードオペレーション）のみをリプレイし得る。幾つか実施形態では、何れの読み出しオペレーションが書き込みオペレーション（ストアオペレーション）からデータを転送したかを読み出し／書き込みユニットがトラッキングし得、書き込みオペレーションのアドレスが誤って投機的推測されれば、読み出し／書き込みユニットが当該書き込みオペレーションからデータを転送した、より若い、つまりより最近の読み出しオペレーションの全てをリプレイし得る。ある種の状況では誤ったデータ推測が検出されたメモリオペレーションをリプレイするように読み出し／書き込みユニットが構成され得る。
【０００８】
　幾つかの実施形態では、メモリオペレーションのアドレスの推測値をメモリオペレーションのアドレスの新たな値と比較し、及び／又はメモリオペレーションの投機的な結果をメモリオペレーションの非推測結果と比較することによって誤ったデータ推測を検出するように読み出し／書き込みユニットが構成され得る。
【０００９】
　方法の様々な実施形態が、実行するように読み出し／書き込みユニットにオペレーションを発行し、この読み出し／書き込みユニットはオペレーションで実行されるデータ推測の誤りの指示を受信し、この指示に応じて読み出し／書き込みユニットは読み出し／書き込みユニット内の突出している少なくとも一つのオペレーションを識別するリプレイ指示を生成し、スケジューラはリプレイ指示によって識別されたオペレーションを応答可能なように再発行する。

【００１０】
本発明は、添付の図面に関連した以下の詳細な説明を参照することによって、よりよく理解することができる。
本発明は様々な変更や代替形態が可能であるが、本発明の個別の実施形態は、各図面において例示として示されたものであり、ここでは詳細に説明されている。しかしながら、ここで行われている実施形態の説明は、本発明を開示された特定の形態に限定することを意図するものではなく、逆に、本発明は、添付の特許請求の範囲により定義されるような本発明の精神や範囲の範疇であるすべての変更、均等物、及び代替物を含むものであることは理解すべきである。本明細書中の見出しは文章構成のためのものに過ぎず、本明細書の説明または特許請求の範囲を限定または解釈のために用いるべきものではないことに注意されたい。さらに、本明細書において用いられる「できる“may”」という用語は、許可的な意味合い（つまり、そのような可能性を持つ、または可能である）に用いられ、強制的な意味合いではない（つまり、must；しなければいけない）ことに注意すべきである。「include;含む」という用語およびその派生語は「含むが、それに限定されない」ことを意味する。「接続される」という用語は「直接的または間接的に接続される」ことを意味し、「結合される」という用語は「直接的または間接的に結合される」ことを意味する。

【発明を実施するための最良の形態】
【００１１】
図１は、マイクロプロセッサ１００の一実施形態のブロック図である。マイクロプロセッサ１００はシステムメモリ２００に記憶された命令を実行するように構成される。これらの命令の多くはシステムメモリ２００に記憶されたデータを操作するものである。システムメモリ２００はコンピュータシステム全体に物理的に配置可能であり、一つあるいはそれ以上のマイクロプロセッサ１００からアクセスできる。
【００１２】
マイクロプロセッサ１００は命令（インストラクション）キャッシュ１０６およびデータキャッシュ１２８を含む。マイクロプロセッサ１００は命令キャッシュ１０６に結合されたプリフェッチユニット１０８を含む。ディスパッチユニット１０４は命令キャッシュ１０６から命令を受信して、スケジューラ１１８にオペレーションを発送するように構成される。１以上のスケジューラ１１８はディスパッチユニット１０４から発送されたオペレーションを受信し、１以上の実行コア１２４にオペレーションを発行するように結合できる。実行コア１２４は、データキャッシュ１２８に対するアクセスを実行するように構成された読み出し／書き込みユニット１２６を含む。実行コア１２４によって生成された結果は、結果バス１３０に出力されうる。これらの結果は続いて発行される命令のためのオペランド値として使用可能であり、および／またはレジスタファイル１１６に記憶することができる。リタイアキュー（retire queue）１０２はスケジューラ１１８およびディスパッチユニット１０４に接続される。このリタイアキューは、発行されたオペレーションのそれぞれがいつリタイア可能であるかを決定するように構成される。一実施形態では、マイクロプロセッサ１００はｘ８６型アーキテクチャと互換性を持つように設計される。マイクロプロセッサ１００はその他多くの要素を含んでいてもよいことに注意すべきである。例えば、マイクロプロセッサ１００は分岐予測ユニット（図示せず）を含んでいてもよい。
【００１３】
命令キャッシュ１０６は命令を、ディスパッチユニット１０４がそれらを受信する前に、一時的に記憶することができる。命令コードは、システムメモリ２００からプリフェッチユニット１０８を介してコードをプリフェッチすることにより、命令キャッシュ１０６に供給される。命令キャッシュ１０６は様々な構成で実装可能である（例えば、セットアソシアティブ、フルアソシアティブ、またはダイレクトマッピング）。幾つかの実施形態では、命令キャッシュ１０６及び／又はデータキャッシュ１２８の多数レベルがあり得る。幾つかのレベルのキャッシュは、図示されるようにマイクロプロセッサ１００と統合さ

れ得、一方、他のレベルのキャッシュはマイクロプロセッサの外部に拡張され得る。

【００１４】

プリフェッチユニット１０８は命令キャッシュ１０６に記憶するためにシステムメモリ２００から命令コードをプリフェッチすることができる。一実施形態では、プリフェッチユニット１０８はシステムメモリ２００から命令キャッシュ１０６にコードをバースト転送するように構成することができる。プリフェッチユニット１０８は様々な特定のコードプリフェッチ技術およびアルゴリズムを採用することができる。

【００１５】

ディスパッチユニット１０４は、オペランドアドレス情報、即値データ（immediate data）および／または変位データ（displacement data）とともに、実行コア１２４が実行可能なビットエンコードされたオペレーションを含む信号を出力する。ある実施形態では、ディスパッチユニット１０４はある種の命令を実行コア１２４内で実行可能なオペレーションにデコードするためのデコード回路（図示せず）を含んでいてもよい。単純な命令は単一のオペレーションに対応する。ある実施形態では、より複雑な命令は複数のオペレーションに対応する。レジスタの更新を含むオペレーションをデコードすると、レジスタファイル１１６中のあるレジスタ位置が推測レジスタ状態を記憶するために予約される（別の実施形態では、リオーダバッファが各レジスタに対する一以上の推測レジスタ状態を記憶するために用いられる）。レジスタマップ１３４は、レジスタのリネームを容易にするために、ソースおよび目的オペランドの論理レジスタ名を物理レジスタ名に変換することができる。レジスタマップ１３４は、レジスタファイル１１６内のどのレジスタが現在割り当てられているか、割り当てられていないかを追跡可能である。

【００１６】

図１のマイクロプロセッサ１００はアウトオブオーダー実行をサポートする。リタイアキュー１０２はレジスタ読み出しおよび書き込みオペレーションのためにオリジナルのプログラムシーケンスの記録、つまりトラッキングを行い、投機的命令実行および分岐予測ミスからの復帰を許可し、精密な例外処理を容易にする。ある実施形態では、リタイアキュー１０２は、有効になったオペレーションがバッファのボトムに移動し、それによって待ち行列のトップに新しいエントリのための空きを作る、先入れ先出し構造として実装される。リタイアキュー１０２は、オペレーションが実行を完了したことに応答してオペレーションを停止し、いずれのオペレーションに対して実行されたいずれの制御スペキュレーション及びデータが、そのオペレーションを含みそのオペレーションまでに実行されたプログラムに関して、照合すなわちベリファイがなされる。リタイアキュー１０２は、物理レジスタの、マイクロプロセッサ１００のアーキテクチャ状態に対する投機的状態について、その物理レジスタでその値を生成したオペレーションがリタイアした場合に、コミットメントを行う。ある実施形態においては、リタイアキュー１０２はリオーダバッファの一部として実装される。このようなリオーダバッファは、レジスタリネーミングをサポートするために、推測レジスタ状態のデータ値の記録部を提供するように構成することもできる。他の形態では、リタイアキュー１０２は、何らデータ値の記録部を提供しないようにすることもできる。また、これに代えて、オペレーションがリタイアする際に、リタイアキュー１０２は、推測レジスタ状態を記憶する必要がなくなったレジスタファイル１１６内のレジスタの割り当てを取り消し、レジスタマップ１３４に対してどのレジスタが現在空いているのかを示す信号を供給する。レジスタファイル１１６内（他の実施形態では、リオーダバッファ内）の推測レジスタ状態をそれらの状態を生成したオペレーションが有効になるまで維持することによって、分岐予測が間違っていた場合、予測ミスされたパスに従った投機的に実行されたオペレーションの結果がレジスタファイル１１６内で無効化される。

【００１７】

特定のオペレーションのにおける要求されているオペランドがレジスタの位置であるときは、レジスタアドレス情報がレジスタマップ１３４（またはリオーダバッファ）に送られる。例えば、ｘ８６アーキテクチャにおいては、８つの３２ビット論理レジスタがある

（例えば、ＥＡＸ，ＥＢＸ，ＥＣＸ，ＥＤＸ，ＥＢＰ，ＥＳＩ，ＥＤＩおよびＥＳＰ）。物理レジスタファイル１１６（またはリオーダバッファ）はこれらの論理レジスタの内容を変更する結果のための記憶装置を含み、アウトオブオーダー実行を許可する。レジスタファイル１１６内の物理レジスタは、デコード時に論理レジスタの一つの内容を修正するために決定される各オペレーションの結果を記憶するために割り当てられる。従って、特定のプログラムを実行している間の任意の時点において、レジスタファイル１１６（または、他の実施形態では、リオーダバッファ）は、与えられた論理レジスタの投機的に実行された内容を含む１以上のレジスタを持つ。

【００１８】

レジスタマップは物理レジスタをあるオペレーションの目的オペランドとして指定された特定の論理レジスタに割り当てる。ディスパッチユニット１０４は、レジスタファイル１１６が、所与のオペレーションにおいてソースオペランドとして指定された論理レジスタに割り当てられた、既に割り当てられた１以上の物理レジスタを持つことを決定する。このレジスタマップ１３４はその論理レジスタに最も最近に割り当てられた物理レジスタのためのタグを供給することができる。このタグは、レジスタファイル１１６内のオペランドのデータ値にアクセスするために用いることができ、または結果バス１３０において転送される結果を介してデータ値を受信することができる。もしオペランドがメモリ位置に対応するならば、そのオペランド値は読み出し／書き込みユニット２２２を介して結果バスに供給される（結果転送および／またはレジスタファイル１１８への記憶のため）。スケジューラ１１８のひとつからオペレーションが発行されると、オペランドデータ値が実行コア１２４に供給される。他の実施形態においては、オペレーションが発送されると、オペランド値が対応するスケジューラ１１８に供給される（そのオペレーションが発行されたときに対応する実行コア１２４に供給する代わりに）。

【００１９】

書き込みユニット１０４の出力で与えられたビット符号化されたオペレーションと即値データ（immediate data）は、一つ又はそれ以上のスケジューラ１１８にルーティングされ得る。本文に記載されたように、スケジューラはオペレーションの実行の準備ができたことを検出し、一つ又はそれ以上のファンクショナルユニットに準備オペレーションを発行するデバイスである。例えば、予約ステーションはスケジューラである。スケジュール又はスケジューラのグループにおけるオペレーションは命令又はオペレーションウィンドー又はスケジューリングウィンドーにおけるオペレーションとして称され得る。各スケジューラ１１８は、実行コア１２４への発行を待つ幾つかの継続中のオペレーションに関するオペレーション情報（例、オペランド値、オペランドタグ、及び／即値データと同様のビット符号化された実行ビット）を保持することができる。ある種の実施形態では、各スケジューラ１１８は、オペランド値を記録することはできない。その代わりに、オペランド値がファンクショナルユニット１２６によって（例えば、レジスタファイル１１６又は結果バス１３０から）読み出し可能となる時点を判定するために、レジスタファイル１１６において利用可能な発行されたオペレーション及び結果を各スケジューラが監視し得る。幾つかの実施形態では、各スケジューラ１１８は専用のファンクショナルユニット１２６に関連付けられ得る。他の実施形態では、シングルスケジューラ１１８が１つより多いファンクショナルユニット１２６にオペレーションを発行し得る。

【００２０】

スケジューラ１１８は、実行コア１２４によって実行されるオペレーション情報を一時記憶するために提供されうる。上述のとおり、各スケジューラ１１８は係属中のオペレーションのためにオペレーション情報を記憶することができる。さらに、各スケジューラは、すでに実行済みであるが再発行される可能性のあるオペレーションのためのオペレーション情報を記憶することができる。実行のために利用可能となった任意の必要とされるオペランドの値に応答して、オペレーションが、実行コア１２４で実行されるよう発行される。従って、オペレーションが実行される順番は、元のプログラム命令シーケンスの順番とは同じでない可能性がある。データ推測を含むオペレーションを、非投機的になるまで

、スケジューラ１１８にとどめ、データ推測が誤っていれば再発行することができる。図１に示したように、読み出し／書き込みユニット１２６Ｃはスケジューラ１１８に再発行される一つあるいはそれ以上のオペレーションを識別するリプレイ指示を提供することができる（例えば、そのようなリプレイ指示は、リプレイされる各オペレーションのタグを含み得る）。スケジューラ１１８は、そのようなリプレイ指示によって識別されたオペレーションを応答可能なように再発行し得る。

【００２１】

一実施形態では、実行コア１２４のそれぞれは幾つかのファンクショナルユニット１２６（例、図１に示されるファンクショナルユニット１２６Ａ～１２６Ｃ）を含み得る。例えば、１２６Ａのような幾つかのファンクショナルユニットは、シフト、ローテート、論理演算、及び分岐オペレーションと同様に加算及び減算の整数演算オペレーションも実行するように構成され得る。例えば、１２８Ｂのような他のファンクショナルユニットは、浮動小数点も含めることができるように構成され得る。例えば、１２６Ｃのようなファンクショナルユニットによって実行される読み出し／書き込みメモリオペレーション用のアドレス生成を行うように一つあるいはそれ以上のファンクショナルユニットが構成され得る。ファンクショナルユニットはデータキャッシュ１２８及び／又はシステムメモリに記録されたデータにアクセスする読み出し／書き込みオペレーションを実行する。一実施形態では、１２６Ｃのようなファンクショナルユニットは継続中の読み出し及び／又は書き込みに関するデータ及びアドレス情報の幾つかのストレージロケーションを備えた読み出し／書き込みバッファで構成され得る。

【００２２】

一つあるいはそれ以上のファンクショナルユニット１２６が分岐予測ユニットに条件付き分岐命令の実行に関する情報も与えることができ、それによってもし分岐予測ミスが発生したときに、分岐予測ユニットが命令プロセッシングパイプラインに入力された予測を誤った分岐に後続する命令をフラッシュすることができ、プリフェッチユニット１０６にリダイレクトすることができるようにする。このリダイレクトされたプリフェッチユニット１０６は、次に命令キャッシュ１０６又はシステムメモリ２００から正しい命令セットのフェッチを開始する。そのような状況では、予測誤りのあった分岐命令の後に発生した元のプログラムシーケンスにおける命令の結果は廃棄され、投機的に実行され、レジスタファイル１１６に一時的に記録されたものを含む。

【００２３】

実行コア１２４内のファンクショナルユニット１２６によって生成された結果は、レジスタの値が更新されていれば、結果バスにおいてレジスタファイル１１６に対して出力される。メモリ位置の内容が変更されていれば、実行コア１２４内で生成された結果は、読み出し／書き込みユニット１２６Ｃに対して供給される。

【００２４】

データキャッシュ１２８は、実行コア１２４とシステムメモリ２００との間で転送されたデータを一時的に記憶するために提供されるキャッシュメモリである。上述の命令キャッシュ１０６と同様に、データキャッシュ１２８はセットアソシアティブ構成を含む、様々な特定のメモリ構成において実装可能である。さらに、データキャッシュ１０６および命令キャッシュ１２８は、ある実施形態においては統一されたキャッシュとして実装可能である。

【００２５】

幾つかの実施形態では、マイクロプロセッサ１００は、マイクロプロセッサがシステムメモリ２００に直接インターフェースすることを可能にする統合されたメモリコントローラ１６０を含む。他の実施形態では、メモリコントローラ１６０はマイクロプロセッサ１００をシステムメモリ２００に間接的に結合するバスブリッジに含まれ得る。

［データ投機］

【００２６】

本明細書では、データ値の誤りが見つけられ、結果として再計算される可能性があれば

、データ値は投機的であるとする。投機的なデータ値とは正しいか又は誤っているかを確定的に識別できないものである。このデータ値は、幾つかのデータ推測が実行されたことに関するオペレーションの結果であるか、あるいはこのデータ値が他の投機的データ値に依存すれば（例えば、データ値が一つ又はそれ以上の投機的なオペランドを有するオペレーションの結果を生成する場合）、データ値は再計算することができる。非推測値とは何れのデータ推測にも依存しない値である（しかし、そのような値はなおコントロール推測に影響される）。

【００２７】

マイクロプロセッサ１００における様々な機構はデータ推測を実行することができる。例えば、ディスパッチユニット１０４、メモリコントローラ１６０、一つ又はそれ以上のファンクショナルユニット１２６の各々が特定的なオペレーションに関するデータ投機を実行可能である。１つのオペレーションの結果が他のオペレーション用の投機的なオペランドとして使用可能であることをディスパッチユニット１０４が検出することができる。例えば、読み出しオペレーションが先行の書き込みオペレーションによってデータキャッシュ１２８に記録されたデータにアクセスすることをディスパッチユニットが予測することができる。書き込みオペレーションのソースとして使用されたレジスタに記録されたデータ値を、ディスパッチユニット１０４が応答可能なように読み出しオペレーションの推測結果として識別することができる。データ推測のこの形式は依存予測として本文において称される。読み出しオペレーションの結果をオペランドとして特定するオペレーション用の推測オペランドとして書き込みオペレーションのソースをリンクすることによって、依存予測がディスパッチユニット１０４に拡張され得る。まだアドレスが未計算である書き込みに関してロード即ち読み出しをバイパスすることを可能にすることによって、即ち、より若い、つまりより最近の読み出しは比較的早期の書き込みには依存しない、と予測することによって、他のタイプの依存予測が読み出し／書き込みユニット１２６Ｃにおいて実行することができる。

【００２８】

マルチプロセッサシステムにおいて、メモリコントローラ１６０がキャッシュコヒーレンシーを維持するようにコヒーレンシー検査を実行可能である。他のマイクロプロセッサのキャッシュとのコヒーレンシー検査を終了する前に、メモリコントローラ１６０がシステムメモリ２００からのキャッシュラインのコピーを投機的に戻し得る。検索するキャッシュラインの正しいコピーが他のプロセッサキャッシュに目下のところ記憶されていることを後続的にコヒーレンシー検査が決定すれば、システムメモリ２００から投機的に検索されたキャッシュラインのコピーは非有効化できる。従って、そのキャッシュラインにアクセスすることによって生成された全ての読み出しオペレーションはコヒーレンシー検査が終了するまで投機的である。この推測の形式は、本文でメモリ予測として称される。

【００２９】

ディスパッチユニット１０４は、オペレーションの結果を予測することによってデータ推測を実行することができる。例えば、ある種のオペレーションは同じ結果を生成し得る傾向があり、それ故、それらのオペレーションの各々が処理される毎に、ファンクショナルユニット１２６によるオペレーションの実際の実行が行われる前にディスパッチユニット１０４によって結果が投機的に生成され得る。このデータ推測の形式は、データ予測として本文に称される。データ予測はマイクロプロセッサの他の部分においても（例えば、読み出し／書き込みユニット１２６Ｃにおいて）実行可能であることに注目されたい。

【００３０】

読み出し／書き込みユニット１２６Ｃはアドレスを投機的に生成し得、この投機的アドレスに基づき、比較的早期に処理された読み出しのパターンに基づきアドレスがまだ未計算の読み出し命令の結果を投機的に生成し得る。例えば、前のＮロードオペレーションが、連続したオフセットＣによって相互に間隙をあけるアドレスＡ１～ＡＮ（例、Ａ１、Ａ２＝Ａ１＋Ｃ、・・・、ＡＮ＝Ａ（Ｎ＋１）＋Ｃ）をターゲットにしたとすれば、読み出し／書き込みユニット１２６Ｃは、読み出しオペレーションの結果として最新にアクセスさ

れたアドレスＡＮ＋連続するオフセットＣでデータを投機的に戻し得る。このデータ投機の形式は、本文においてアドレス予測として称される。アドレス予測の他の形式が多くの実施形態に使用可能であることに注目されたい。

【００３１】

幾つかの投機的な結果を得るために、タイプの異なる複数のデータ推測形式を実行することも可能である。例えば、整数オペレーションの推測結果は、データ推測を使用して生成可能である。この投機的な結果は後に書き込みオペレーションによって記憶することができる。読み出しオペレーションは依存予測を通じてこの書き込みに依存するように予測され得、それ故、この読み出しオペレーションの投機的な結果は整数オペレーションの推測結果となる。

【００３２】

データ推測が実行されたオペレーションの結果に依存するオペレーションも投機的な結果を生成し得る。例えば、アドレス予測が読み出しオペレーションの投機的な結果を生成するように使用されれば、オペランドとして読み出しの推測結果を用いて実行する全ての依存オペレーションが推測結果を生成することができ、これは次に他の依存オペレーションによってオペランドとして使用することが可能である。従って、読み出しオペレーションにおいて根本となる推測が誤っていると決定されれば、依存オペレーションの結果も誤っている可能性があり、それ故、この読み出しに依存するオペレーションの全体的な依存チェーンが正しい結果を生成するために再度実行されることが求められる。言い換えれば、根本となる推測が正しいと決定されれば、（これらの結果が他の投機的な値の何れにも基づかないと仮定して）依存オペレーションの結果は正しい可能性がある。

【００３３】

データ推測が実行される多くオペレーションは、ファンクショナルユニットによってこれらのオペレーションが実行されるときにベリファイ、即ち検証することができる。例えば、あるオペレーション結果を投機的に生成するよう使用されたデータ予測はファンクショナルユニット１２６によって検証することができ、このファンクショナルユニット１２６は実際のオペレーションの結果と投機的な結果を比較することによってこのオペレーションを実行する。正しい結果が既に利用可能であるのでデータ推測が誤っていれば、そのようなオペレーションは再実行されることが求められる。他のオペレーションは、完全に実行されなくとも検証可能である。例えば、未計算のアドレスで読み出しが比較的早期の書き込みの結果を（例えば、依存予測又はアドレス予測に起因して）転送した場合、読み出しの推測結果は、読み出しアドレスが計算されるとき検証可能である。データ推測が誤っていれば、そのようなオペレーションは、正しい結果を生成するために（少なくとも一部）再実行されることが求められる。

【００３４】

データ推測が実行されたオペレーションと、それに従属するオペレーションを再実行することが求められるので、リタイヤキュー１０２は根本となるデータ推測が解決された（resolved）、あるいは推測が正しかったか誤っていたか判明したオペレーションをリタイヤするためのみに構成することができる。

［リプレイ機構を備えた読み出し／書き込みユニット］

【００３５】

データ推測の実行をサポートするマイクロプロセッサでは、１つ又はそれ以上のオペレーションがデータ推測の誤りに起因して再実行されることが求められる。マイクロプロセッサ１００は、オペレーションの再実行を可能とする様々なリプレイ機構を含むことができる。例えば、オペレーションのデータ投機的オペランドの１つが誤っているという指示に応答して、スケジューラ１１８がオペレーションをファンクショナルユニット１２６Ｃに対して再発行するよう構成可能である。同様に、推測が誤っていると判明したオペレーションのデータ推測に応答して、読み出し／書き込みユニットにおいて未済のオペレーションのうちどれをリプレイしなければならないかを検出するリプレイ機構も読み出し／書き込みユニット１２６Ｃは含む。読み出し／書き込みユニット１２６Ｃは、未済のオペレ

ーションをリプレイしなければならないかに関する信号をスケジューラに送信し、後にスケジューラが読み出し／書き込みユニット１２６Ｃにこれらのオペレーションを再発行させる。

【００３６】

図２に、読み出し／書き込みユニット１２６Ｃによって検出されたデータ推測の誤りに起因してオペレーションをリプレイするよう構成される読み出し／書き込みユニット１２６Ｃの一実施形態が示される。データ推測検証ロジック３０５は、読み出し／書き込みユニット１２６Ｃによって実行されたデータ推測のそれぞれの形式を検証するように構成することができる。例えば、読み出し／書き込みユニットアドレス予測（例えば、読み出しオペレーションのアドレスを予測することによる）と依存予測（例えば、未計算のアドレスにおけるより早い時期になされた書き込みの読み出しをバイパス可能にすることによる）を実行すれば、読み出し／書き込みユニットは、アドレスと依存予測の双方を検証するように構成することができる。データ推測検証ロジック３０５も（又は他の例として）、マイクロプロセッサ１００の他の部分によって実行されたデータ推測を検証するよう構成することができる。例えば、ディスパッチユニット１０４が依存予測を実行するように構成されれば（例えば、読み出し結果を比較的早い時期になされた書き込みのソースに投機的にリンクすることによって）、データ推測検証ロジック３０５はその依存予測を検証するように構成することができる。

【００３７】

読み出し／書き込みユニットに発行されているが（例えば、スケジューラ１１８によって）、まだ実行を完了してないオペレーションに関するオペレーションストレージ３０７も、読み出し／書き込みユニット１２６Ｃに含まれる。データ推測が実行された読み出し／書き込みユニットに発行されたオペレーションは、そのオペレーションがデータ推測検証ロジック３０５によって検証されるまでリタイヤすることができない。オペレーションストレージ３０７は、読み出し／書き込みユニット１２６Ｃ内の全ての未済のオペレーションのトラッキング、即ち追跡が可能である。オペレーションストレージ３０７は、未処理の読み出し即ちロード及び書き込み即ちストアのそれぞれについてエントリ３１０を含むことができる。

【００３８】

エントリ３１０は、又、エントリが読み出し又は書き込みに割り当てされたか否かを示す情報３１３を含むことができる（又は、幾つかの実施形態では、エントリがメモリアドレスから読み出された値にオペレートし、メモリアドレスに結果を記憶するオペレーションに対応すれば、エントリは読み出し及び書き込みの双方を含むことを示し得る）。加えて、エントリ３１０がタグ３１５（例、マイクロプロセッサ１００内のオペレーションとその結果を識別する）、アドレス３１７及び／又はデータ３１９を含み得る。幾つかの実施形態では、各エントリのデータフィールド３１９が推測及び非推測データの双方に関するストレージを含む。同様に、幾つかの実施形態では、アドレスフィールド３１７がオペレーションのアドレスの１つの値より多い値（例、アドレス予測によって生成された投機的なアドレスとオペレーションを実行することによって生成された新たなアドレス値）に関するストレージを含み得る。ある種の実施形態では、エントリが、データ推測としてオペレーション及び／又はオペランドを識別するよう追加のフィールドを含むことができる。エントリ３１０は、オペレーションを読み出し／書き込みユニット１２６Ｃに発行するスケジューラ１１８に応じて割り当てられ、このオペレーションの実行を終了する読み出し／書き込みユニット１２６ｃに応じて割り当て解除される。

【００３９】

データ推測検証ロジック３０５は、オペレーションの推測結果をオペレーションの実際の結果と比較することによって（例、データ予測及び／又は幾つかの依存予測の形式）、データ推測のある種の形式を検証することができる。例えば、読み出しオペレーションの推測結果は、オペレーションストレージ３０７における読み出しのエントリ３１０に記憶可能である。この読み出しオペレーションの実際の結果がデータキャッシュ１２８から受

信されると、データ推測検証ロジックが、実際の結果をオペレーションストレージ３０７に記録された投機的結果と比較し得る。データ推測検証ロジック３０５は、オペレーションのアドレスを１つ又はそれ以上のより早期のオペレーションのアドレスと比較することによってデータ推測の他のタイプ（例、ある種の依存予測のタイプ）を検証可能である。データ推測検証ロジック３０５は、結果バス１３０に送信された新たな値に応じてオペレーションの投機的なアドレスと前述の新たな値を比較することによってそのアドレスのデータ推測の他の形式（例、アドレス予測）を検証可能である。投機的なアドレスがアドレスの新たな値と合致即ちマッチングしなければ、データ推測検証ロジック３０５は、アドレスに実行されたデータ推測は誤りだったと判定する。

【００４０】

特定的なオペレーションに関するデータ推測が誤りだったことが検出されたことに応答して、データ推測検証ロジック３０５は、１つ又はそれ以上のオペレーションをリプレイすることを生じさせることができる。オペレーションは、そのオペレーションを識別するリプレイ信号をスケジューラ１１８に送信することによってリプレイされ得る。そのような信号に応じて、スケジューラ１１８がリプレイに関するオペレーションのマーキングを行うことができる（例、オペレーションがリプレイされなければならないことを示すようそのオペレーションと関連付けられたステート情報を変更することによって）。一実施形態では、データ推測検証ロジック３０５がオペレーションをリプレイしなければならないことを示すフラグを伴ってそのオペレーションのタグをスケジューラ１１８に供給することによってリプレイさせることができる。

【００４１】

誤って推測されたオペレーションを再実行することが（例えば、誤った依存予測、又は誤ったアドレス予測に起因して）求められれば、データ推測検証ロジック３０５は、オペレーションのリプレイの実行を行うことができる。誤って推測された結果に対する正しい結果が既に入手可能であれば（例えば、誤った結果が、誤ったデータ又は誤った依存予測に起因すれば）、マイクロプロセッサの他の部分における依存オペレーションが正しい値を用いて再実行可能であるように、データ推測検証ユニットが、読み出し／書き込みユニット１２６Ｃに、誤って推測されたオペレーションの正しい結果を、マイクロプロセッサの他のコンポーネントにブロードキャストさせる。ある種の実施形態では、データ推測検証ロジック３０５がそのようなオペレーションをリプレイすることを生じさせない。読み出し／書き込みユニット１２６Ｃは、たとえこのオペレーションのリプレイを生じさせるとしても、未済のオペレーションの実行を完了可能であることに注目されたい。

【００４２】

誤って推測されたオペレーションをリプレイすることに加えて、必要であれば、データ推測検証ロジック３０５が読み出し／書き込みユニットで現在未済である他のオペレーションをリプレイすることも生じさせ得る。例えば、幾つかの読み出しオペレーションが書き込みオペレーションの投機的なアドレスに基づく書き込みオペレーションからのデータを転送することができる。書き込みオペレーションの推測アドレスが誤っていると決定されれば（例、書き込みオペレーションのアドレスを生成するオペレーションの実行が推測アドレスに匹敵しない新たなアドレス値を生成すれば）、書き込みの推測アドレスに基づきデータを転送した全ての読み出しオペレーションがリプレイすることができる。同様に、書き込みオペレーションに関する新たなアドレスとアドレスが合致する読み出しオペレーションも、それらのオペレーションが書き込みからのデータを転送可能であるように、リプレイされることが可能である。同様に、データ推測検証ロジック３０５は、書き込みオペレーションに関するオペランドが結果バスにリブロードキャストされるときを検出し、又、書き込みオペレーションのオペランドの前の値を転送した依存読み出しオペレーションの全てのリプレイを応答可能なようにするよう構成可能である。

【００４３】

一実施形態では、データ推測検証ロジック３０５は、読み出し／書き込みユニット内の、誤って推測されたオペレーションよりも最近のものである未済のオペレーションの全て

をリプレイさせることによって、依存オペレーションをリプレイすることができる。ある種の実施形態では、コンパレータロジック３０３がオペレーションストレージ３０７内の比較的早期のオペレーションをリプレイするために識別し得る。コンパレータロジック３０３は、値を比較し、及び／又は一致する値を検出するための様々な手段（例えば、コンパレータ、内容アドレス宛先指定可能メモリ等）の全て含む。

【００４４】

他の実施形態では、特定的なオペレーションに関する誤ったデータ推測の検出に応じてある種の比較的若い、即ち、より最近のオペレーションを選択的にリプレイするため、データ推測検証ロジック３０５がコンパレータロジック３０３を使用可能である。例えば、一実施形態では、書き込みオペレーションのアドレスに関する誤ったデータ推測の検出に応じて、コンパラータ論理３０３は、書き込みのアドレスの投機的な値及び新たな値と、オペレーションストレージ３０７におけるより最近のオペレーションのそれぞれのアドレスと比較可能である。より最近のアドレスの何れかが書き込みのアドレスの何れかと合致すれば、合致したより最近のオペレーションがこれらのオペレーションをリプレイするよう求められる指示をスケジューラに与えることによって再実行を生じさせることができる。

【００４５】

幾つかの実施形態では、読み出し／書き込みユニット１２６Ｃは、未済の読み出しのアドレスと比較的古い書き込みのアドレスとを比較することによって書き込みからロードフォワーディングを実装し得る。読み出しのアドレスが比較的古い記録のアドレスと合致すれば、読み出し／書き込みユニット１２６Ｃが、読み出しオペレーションの結果として比較的古い書き込みオペレーションによって記憶されたデータを出力し得る。書き込みオペレーションからデータを転送する度に、転送された読み出しオペレーションと関連付けられたエントリ、フォワーディングトラッキングバッファ３０９における書き込みオペレーションを識別するタグを、読み出し／書き込みユニット１２６Ｃが記憶することができる。これらの実施形態の幾つかでは、コンパレータロジック３０３は、フォワーディングトラッキングバッファ３０９によって与えられた情報を使用可能であり、オペレーションに関する誤ったデータ推測を検出するデータ推測検証ロジック３０６に応じて再実行するためのオペレーションを検証する。例えば、書き込みオペレーションのためのアドレス予測が誤っていると決定されれば、コンパレータロジック３０３が書き込みオペレーションのためのアドレスの新たな値とオペレーションストレージ３０７におけるより最近の読み出しオペレーションの各々のアドレスを比較するよう構成することができる。いずれのマッチングオペレーションも、誤って投機的推測された書き込みアドレスに起因するのではなく、誤って投機予測された書き込みからデータを転送した筈である。従って、読み出し／書き込みユニット１２６Ｃは、これらのマッチングオペレーションの再実行を生じさせることができる。加えて、誤って投機された書き込みオペレーションのタグは、より最近の読み出しに関するフォワーディングトラッキングバッファ３０９に記憶されたタグと比較され得る。誤って投機された書き込みオペレーションがより最近の読み出しのために記憶されたタグと合致すれば、より最近の読み出しが、書き込みの誤って投機されたアドレスに基づきデータを誤って転送したことを示し、読み出し／書き込みユニット１２６Ｃがより最近の読み出しを再実行させる。

【００４６】

図３に誤ったデータの推測の検出に応じて読み出し／書き込みユニット内の未済のオペレーションをリプレイする方法の一実施形態が示される。５０１において、オペレーションに関するデータ推測が誤っているという指示が検出される。例えば、読み出し／書き込みユニットがアドレス生成ユニット（幾つかの実施形態では、読み出し／書き込みユニットの一部であり得る）からオペレーションに関するアドレスの新たな値を受信し得る。新たなアドレス値は、オペレーション又は他のオペレーションの投機的な結果を生成するよう読み出し／書き込みユニットによって使用された投機的なアドレス値とは異なっており、そのアドレスに対して実行された投機的なデータ推測が誤っていることを示す。他の実

施例では、読み出し／書き込みユニットが読み出しの投機的な結果（例、読み出し結果がより早い時期の書き込みのソースに等しいであろうと予測することによって生成された）とデータキャッシュにアクセスすることによって得られた読み出しの実際の結果と比較することによって、依存予測が誤っていることを検出可能である。

【００４７】

図３Ａに示す実施形態では、５０３で示したようにオペレーションに関するデータ推測が誤っていると検出されたことに応答して、読み出し／書き込みユニットが読み出し／書き込みユニット内のより最近の未済のオペレーションの全てを再実行させる。読み出し／書き込みユニットは、リプレイされる各オペレーションのタグをスケジューラに供給することによって、より最近のオペレーションを再実行させる。読み出し／書き込みユニットによって供給されたタグに応答して、スケジューラが、スケジューラにおけるオペレーションのステート[例えば、“発行された(issued)”から“未発行の（non-issued）”、又は“再発行が求められる(need to be reissued)”まで]を更新し得る。そのようなリプレイ機構によれば、より最近のオペレーションがマイクロプロセッサのプロセッシングパイプラインからフラッシュさえることなく、命令キャッシュ１０６からリフェッチされてリプレイできるようになる。

【００４８】

図３Ｂには、読み出し／書き込みユニット内の未済のオペレーションを再実行する方法の他の実施形態が示される。５０５において、書き込みのアドレスのデータ推測が誤っていることの指示が検出される。より最近の未済の読み出し又書き込みが、５０７で決定されたように、誤って投機された書き込みのアドレスの推測値又は新たな値と合致するアドレスを有する場合、マッチングオペレーションを再実行することができる。より最近のオペレーションのアドレスの何れも書き込みオペレーションのアドレスの投機的値及び新たな値と合致しなければ、より最近のオペレーションは５０９で示したようにリプレイされ得ない。

【００４９】

図３Ｃに、読み出し／書き込みユニット内の未済のオペレーションを再実行する方法の更なる他の実施形態が示される。５１５において、書き込みオペレーションのアドレスに関するデータ推測が誤っていることの指示が検出される。最近の読み出しのアドレスの全てが書き込みオペレーションのアドレスの新たな値と合致すれば、合致するより最近の読み出しはリプレイすることができる。５２１で決定したように、より最近の読み出しが、（例えば、より最近の読み出しがデータを転送した書き込みオペレーションのタグと誤って投機された書き込みのタグとを比較することによって決定された）誤って投機された書き込みのデータを転送した場合、合致する、より最近の読み出しをリプレイすることができる。

【００５０】

他の実施形態で、図３Ａ～図３Ｃに示す実施形態とは異なるように実行し得ることに注目されたい。例えば、図３Ｂの実施形態と同様の実施形態において、誤って投機された書き込みのアドレスの推測値又は新たな値とアドレスが合致する、より最近の読み出しオペレーションのみが再実行することができる。同様に、図３Ｃと同様の実施形態では、誤って投機された書き込みのアドレスの新たな値とアドレスが合致する、より最近の読み出し及び書き込みが再実行することが可能である。他の多くの変形が可能である。

［例示的なコンピュータシステム］

【００５１】

図４は、バスブリッジ９０２を介して様々なシステムコンポーネントに結合されたプロセッサ１００を含むコンピュータシステム９００の一実施形態のブロック図を示す。プロセッサ１００は、上述したように読み出し／書き込みユニットの一実施形態を含み得る。コンピュータシステムの他の実施形態が、可能であり、また考えられる。図示されたシステムにおいて、メインメモリ２００は、メモリバス９０６を介してバスブリッジ９０２に結合され、画像コントローラ９０８は、ＡＧＰバス９１０を介してバスブリッジ９０２に

結合される。幾つかのＰＣＩデバイス９１２Ａ～９１２Ｂは、ＰＣＩバス９１４を介してバスブリッジ９０２に結合される。第２バスブリッジ９１６も、ＥＩＳＡ／ＩＳＡバス９２０を介してＥＩＳＡ又はＩＳＡデバイス９１８のうちの１つあるいはそれ以上に電気的インターフェースを適用するように供給されることが可能である。この実施例では、プロセッサ１００は、ＣＰＵバス９２４と任意のＬ２キャッシュ９２８を介してバスブリッジ９０２に結合される。ある種の実施形態では、プロセッサ１００は統合されたＬ１キャッシュ（図示せず）を含むことが可能である。

【００５２】

バスブリッジ９０２は、プロセッサ１００，メインメモリ２００，画像コントローラ９０８と，ＰＣＩバスに接合されたデバイス間のインターフェースを与える。オペレーションがバスブリッジ９０２に接続されたデバイスの１つから受信されると、バスブリッジ９０２がこのオペレーションの目標（例、特定的なデバイス又はＰＣＩバス９１４の場合、目標はＰＣＩバス９１４上にあるデバイス）を識別する。バスブリッジ９０２がオペレーションを目標となるデバイスに伝送する。バスブリッジ９０２は通常、ソースデバイス又はバスによって使用されるプロトコル、あるいはターゲットデバイス又はバスによって使用されるプロトコルのオペレーションを変換する。

【００５３】

ＰＣＩバス９１４のためにＩＳＡ／ＥＩＳＡバスにインターフェースを供給することに加えて、第２バスブリッジ９１６は、更なる機能性を内蔵し得る。第２バスブリッジ９１６の外部あるいは第２バスブリッジに統合されているうちの何れかである入力／出力コントローラ（図示せず）は、コンピュータシステム９００内に含むことができ、キーボード及びマウス９２２と様々なシリアル及びパラレルポートに関するオペレーショナルサポートを提供する。他の実施形態では、外部のキャッシュユニット（図示せず）は、プロセッサ１００とバスブリッジ９０２間のＣＰＵバス９２４にも結合され得る。他の例では、外部のキャッシュはバスブリッジ９０２に結合することができ、外部のキャッシュのキャッシュコントロールは、バスブリッジ９０２に統合可能である。Ｌ２キャッシュ９２８はプロセッサ１００の裏面の構成に示される。Ｌ２キャッシュ９２８は、プロセッサ１００から分離し得、プロセッサ１００でカートリッジ（例、スロット１又はスロットＡ）内に統合可能であり、あるいはプロセッサ１００で半導体基板上にも統合可能である。

【００５４】

メインメモリ２００はアプリケーションプログラムが記録され、プロセッサ１００が主に実行するメモリである。適合可能なメインメモリ２００は、ＤＲＡＭ（Dynamic Random Access Memory）を含み得る。例えば、複数のＳＤＲＡＭ（同期ＤＲＡＭ）あるいはＲａｍｂｕｓ　ＤＲＡＭ（ＲＤＲＡＭ；Rambus 社開発のＤＲＡＭ）が適合可能であり得る。

【００５５】

ＰＣＩデバイス９１２Ａ～９１２Ｂは、ネットワークインターフェースカード、ビデオアクセラレータ、オーディオカード、ハードあるいはフロッピイディスクドライブあるいはドライブコントローラ、ＳＣＳＩ（Small Computer System Interface）アダプター及びテレフォニーカードのような多様な周辺デバイスを示す。同様に、ＩＳＡデバイス９１８はモデム、サウンドカード、及びＧＰＩＢあるいはフィールドバスインターフェースカードのような周辺デバイスの様々な形式を表す。

【００５６】

グラフィックスコントローラ９０８は、モニタ９２６上のテキスト及び画像のレンダリングをコントロールするよう提供される。グラフィックスコントローラ９０８は、メインメモリ２００に、又はメインメモリ２００から有効的にシフト可能である３次元データ構造を表現する技術分野で通常、知られる一般的なグラフィックスアクセラレータを実施可能である。それ故、グラフィックスコントローラ９０８は、バスブリッジ９０２内のターゲットインターフェースへのアクセスを要求可能であり、又受信可能であり、それによってメインメモリ２００へのアクセスを入手するＡＣＰバス９１０のマスターであり得る。専用のグラフィックスバスは、メインメモリ２００からのデータの高速の検索に対応する

。ある種のオペレーションに関して、グラフィックスコントローラ９０８は、ＡＧＰバス９１０条にＰＣＩプロトコルを生成するよう更に構成可能である。それ故、バスブリッジ９０２のＡＧＰインターフェースは、ＰＣＩプロトコルターゲットとイニシエータトランザクションと同様にＡＧＰプロトコルトランザクションの双方を支援するよう機能性を含み得る。モニタ９２６は、画像あるいはテキストが表示可能である電気的モニタの全てである。適切なモニタ９２６は、ブラウン管（“ＣＲＴ；Cathode Ray Tube”）、液晶モニタ（“ＬＣＤ；Liquid Crystal Display”）を含む。

【００５７】

ＡＧＰ、ＰＣＩ及びＩＳＡ又はＥＩＳＡバスは、上述した記載において、使用され、要求されたように全てのアーキテクチャが上書きされ得ることに注目される。コンピュータシステム９００が追加のプロセッサ（例、コンピュータシステム９００の任意のコンポーネントとして示されるプロセッサ１００）を含むマルチプロセッシングコンピュータシステムであることが更に注目される。プロセッサ１００ａは、プロセッサ１００と同様であり得る。より具体的には、プロセッサ１００ａは、プロセッサ１００の同一のコピーであり得る。プロセッサ１００ａは、独立したバスを介して（図４に示されるように）バスブリッジ９０２に接続されるか、あるいはプロセッサ１００とＣＰＵバス９２４を共有することができる。更に、プロセッサ１００ａは、Ｌ２キャッシュ９２８と同様の任意のＬ２キャッシュ９２８ａに結合することができる。

【００５８】

図５を参照すると、上述したように読み出し／書き込みユニットの実施形態を有するコンピュータシステム９００の他の実施形態が示される。他の実施形態が可能であるし、考えられる。図５の実施形態においては、コンピュータシステム９００が幾つかのプロセッシングノード１０１２Ａ、１０１２Ｂ、１０１２Ｃ及び１０１２Ｄを含む。各プロセッシングノードは、各プロセッシングノード１０１２～１０１２Ｄ内に含まれるメモリコントローラ１０１６Ａ～１０１６Ｄを介して各メモリ２００Ａ～２００Ｄに結合される。加えて、プロセッシング１０１２Ａ～１０１２Ｄは、プロセッシングノード１０１２Ａ～１０１２Ｄ間を通信するように使用されるインターフェースロジックを含む。例えば、プロセッシングノード１０１２Ａは、プロセッシングノード１０１２Ｂと通信するためのインターフェースロジック１０１８Ａ、プロセッシングノード１０１２Ｃと通信するためのインターフェースロジック１０１８Ｂ、及び更なる他のプロセッシングノード（図示せず）と通信するための第３インターフェースロジック１０１８Ｃを含む。同様に、プロセッシングノード１０１２Ｂは、インターフェースロジック１０１８Ｄ、１０１８Ｄ及び１０１８Ｆを含み、プロセッシングノード１０１２Ｃは、インターフェースロジック１０１８Ｇ、１０１８Ｈ、及び１０１８Ｉを含み、プロセッシングノード１０１２Ｄは、インターフェースロジック１０１８Ｊ、１０１８Ｋ、及び１０１８Ｌを含む。プロセッシングノード１０１２Ｄは、インターフェースロジック１０１８Ｌを介して複数の入力／出力デバイス（例、デイジーチェーン構成におけるデバイス１０２０Ａ～１０２０Ｂ）と通信するよう結合される。他のプロセッシングノードは、他のＩ／Ｏデバイスと通信し得る。他のプロセッシングノードは、同様の方法で他のＩ／Ｏデバイスと通信し得る。

【００５９】

プロセッシングノード１０１２Ａ～１０１２Ｄは、インタープロセッシングノード通信用にパケットベースのリンクを実装する。本実施形態において、リンクが単方向ライン（ユニディレクタルライン）の一式（例、１０２４Ａがプロセッシングノード１０１２Ａからプロセッシングノード１０１２Ｂへパケットを伝送するように使用され、ライン１０２４Ｂは、プロセッシングノード１０１２Ｂからプロセッシングノード１０１２Ａにパケットを伝送するように使用される）として実装される。一般的に、ライン１０１４の各セットは、図５に表した他のプロセッシングノード間にパケットを伝送するように使用される。一般的に、ライン１０１４の各セットは、１つあるいはそれ以上のラインを含み、このデータラインに応じた１あるいはそれ以上のクロックラインを含み、伝送されるパケット形式を示す１あるいはそれ以上のコントロールラインを含む。リンクは、プロセッシング

ノード間の通信用にキャッシュコヒーレントに、あるいはプロセッシングノードとＩ／Ｏデバイス（あるいはＰＣＩバス又はＩＳＡバスのような従来の構造のＩ／Ｏバスへのバスブリッジ）間の通信用には非コヒーレントに動作させてもよい。更に、リンクは、図示されるようにＩ／Ｏデバイス館のデイジーチェーン構造を使用して非コヒーレントに動作され得る。プロセッシングノードから他のプロセッシングノードに伝送されるパケットは、１つあるいはそれ以上の中間ノードを通じて伝送し得る。例えば、プロセッシングノード１０１２Ａからプロセッシングノード１０１２Ｄまで伝送されたパケットは、図５に示すプロセッシングノード１０１２Ｂあるいはプロセッシングノード１０１２Ｃの何れかを介して伝送し得る。全ての適合可能なルーティングアルゴリズムが使用され得る。コンピュータシステム９００の他の実施形態は、図５に示す実施形態より多いか、あるいは少ないノードを含み得る。

【００６０】

通常、パケットはノード間のラインに１つあるいはそれ以上のビット時間として伝送され得る。ビット時間は、対応するクロックラインのクロック信号の立ち上がり又は立ち下がりであり得る。パケットはトランザクションを開始するためのコマンドパケット、キャッシュコヒーレンシーを維持するためのプローブパケット、及びプローブとコマンドに応答するための応答パケットを含み得る。

【００６１】

メモリコントローラとインターフェースロジックに加えて、プロセッシングノード１０１２Ａ～１０１２Ｄは、１つあるいはそれ以上のプロセッサを含み得る。大まかに言って、ノードは少なくとも１プロセッサを含み、任意的に要求されるようにメモリと他の論理と通信するためのメモリコントローラを含み得る。より具体的には、各プロセッシングノード１０１２Ａ～１０１２Ｄは、１つあるいはそれ以上のプロセッサ１００のコピーを含み得る。外部のインターフェースユニットは、メモリコントローラ１０１６と同様にノード内にインターフェースロジック１０１８を含み得る。

【００６２】

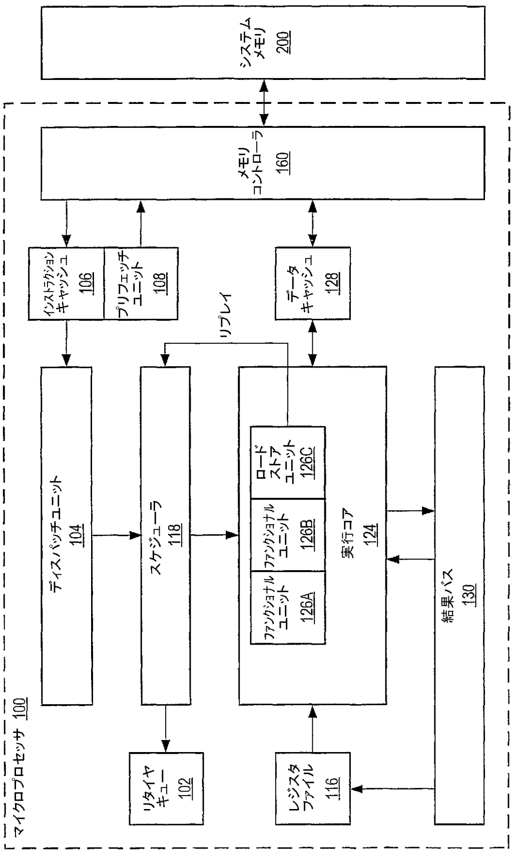
メモリ２００Ａ～２００Ｄは、適合可能な全てのメモリデバイスを含み得る。例えば、メモリ２００Ａ～２００Ｄは、ＲＡＭＢＵＳ　ＤＲＡＭ（ＲＤＲＡＭ），同期ＤＲＡＭ（ＤＤＲＡＭ），静的ＲＡＭ等を含み得る。コンピュータシステム９００のアドレススペースは、メモリ２００Ａ～２００Ｄ間に分割される。各プロセッシングノード１０１２Ａ～１０１２Ｄは、メモリマップを含み得、何れのアドレスが何れの２００Ａ～２００Ｄにマップされるかを決定し、従って、プロセッシングノード１０１２Ａ～１０１２Ｄの何れに特定的なアドレスのメモリ要求がルートされなければならないかのを決定するよう使用されるメモリマップを含み得る。一実施形態において、コンピュータシステム９００におけるアドレスの干渉点はこのアドレスに対応するバイトを記録するメモリに結合されたメモリコントローラ１０１６Ａ～１０１６Ｄである。言い換えれば、メモリコントローラ１０１６Ａ～１０１６Ｄは、対応するメモリ２００Ａ～２００Ｄへのアクセスがキャッシュ干渉方法で生じすることを確実にする役目を担う。メモリコントローラ１０１６Ａ～１０１６Ｄは、２００Ａ～２００Ｄへのインターフェースするためのコントロール回路を含み得る。加えて、メモリコントローラ１０１６Ａ～１０１６Ｄは、メモリ要求をキューするための要求キューを含み得る。

【００６３】

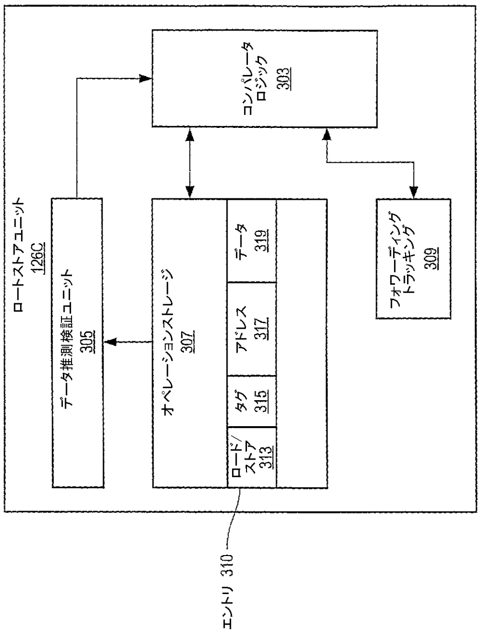
インターフェースロジック１０１８Ａ～１０１８Ｌは、リンクからのパケットを受信し、リンク上に伝送されるパケットを緩衝する様々なバッファを含み得る。コンピュータシステム９００は、パケットを伝送するフローコントロール機構の全てを使用し得る。例えば、一実施形態において、各インターフェースロジック１０１８は、インターフェースロジックが接続されるリンクの他の末端におけるレシーバ内のバッファの各形式の数のカウントを記憶する。インターフェースロジックは、受信するインターフェースロジックがパケットを記録する空いているバッファを有する限り、パケットを伝送しない。受信するバッファが伝送されるパケットをルーティングすることによって解除されるので、受信する

インターフェースロジックが、メッセージを送信ロジックに伝送し、バッファが解除されたことを示す。そのような機構は、“クーポンベースの”システムと称される。
【００６４】
　Ｉ／Ｏデバイス１０２０Ａ～１０２０Ｂは、適合可能なデバイスであり得る。例えば、Ｉ／Ｏデバイス１０２０Ａ～１０２０Ｂが、これらのデバイスが結合され得る他のコンピュータシステム（例、ネットワークインターフェースカードあるいはモデム）と通信するデバイスを含み得る。更に、Ｉ／Ｏデバイス１０２０Ａ～１０２０Ｂは、ビデオアクセラレータ、オーディオカード、ハード又はフロッピーディスク又はドライブコントローラ、ＳＣＳＩ（Small Computer System Interface）アダプタ、及びテレフォニーカード、サウンドカード、及びGRIB又はフィールドバスインターフェースカードのような様々なデータ入手カードを含む。用語“Ｉ／Ｏデバイス”と“周辺のデバイス”は、同義語として本文に意図される。
【００６５】
　ここで使用されるように、“クロックサイクル”の用語は、命令プロセッシングパイプラインの様々なステージがタスクを完了する時間間隔を意味する。命令及び計算された値は、クロックサイクルを定義するクロック信号に従ってメモリ素子（レジスタあるいはアレイなど）によって取得される。例えば、メモリ素子は、クロック信号の立ち上がり又は立ち下がりエッジに従って値を取得することができる。
【００６６】
　上述した明細書は信号を“表された（asserted）”ものとして説明するものである。情報の特定的な一部を示す値を伝送するとき、信号は“表された”ものとして定義され得る。特定的な信号では、バイナリ値が「１」の信号を伝送するとき、あるいはバイナリ値が「０」の値を伝送する場合に、「表された」として定義され得る。
【００６７】
　これまでの開示を完全に理解したならば、当業者には様々な変形及び変更が可能であることが明白であろう。添付の特許請求の範囲はそのような変形及び変更の全てを包含するするように解釈されることを意図したものである。
【産業上の利用可能性】
【００６８】
　本発明は、概してマイクロプロセッサの技術分野に利用可能である。
【図面の簡単な説明】
【００６９】
【図１】一実施形態に従ったマイクロプロセッサを示す図。
【図２】一実施形態に従った読み出し／書き込みユニットのブロック図。
【図３Ａ】一実施形態に従った読み出し／書き込みユニット内にオペレーションをリプレイする方法のフロー図。
【図３Ｂ】他の実施形態に従った読み出し／書き込みユニット内にオペレーションをリプレイする方法のフロー図。
【図３Ｃ】更なる他の実施形態に従った読み出し／書き込みユニット内にオペレーションをリプレイする方法のフロー図。
【図４】一実施形態に従った例示的なコンピュータシステム。
【図５】他の実施形態に従った他の例示的なコンピュータシステム。

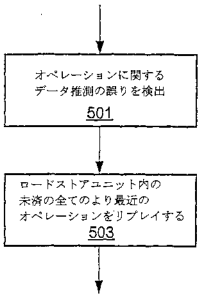
【図１】



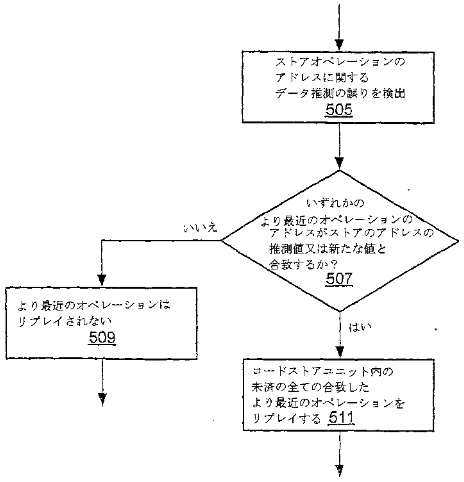
【図２】



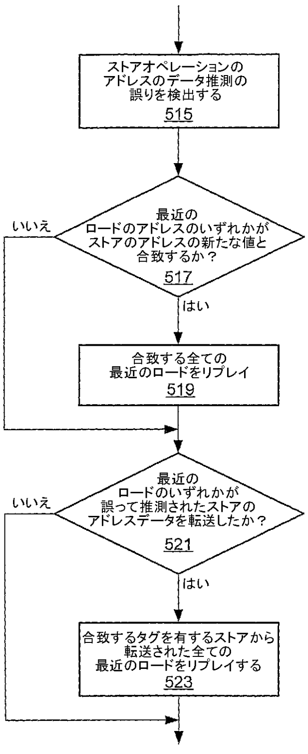
【図３Ａ】



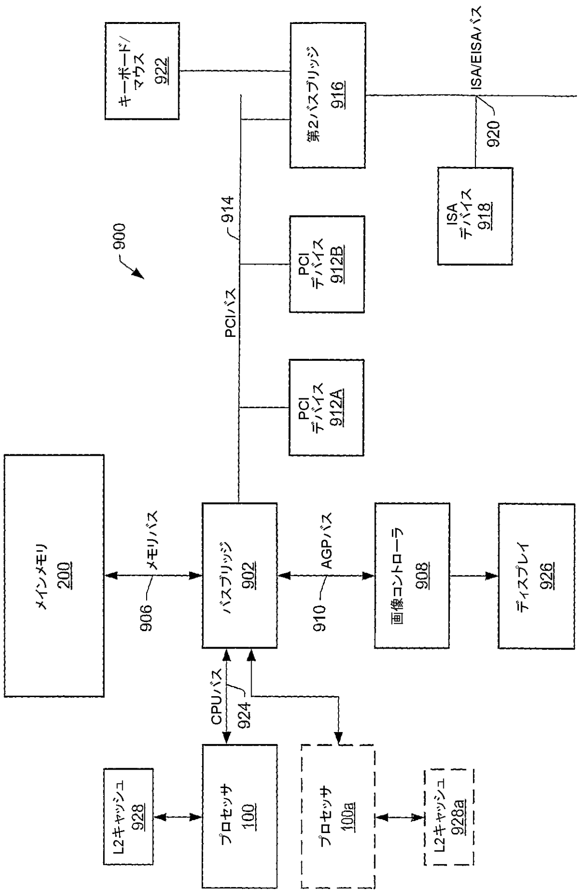
【図３Ｂ】



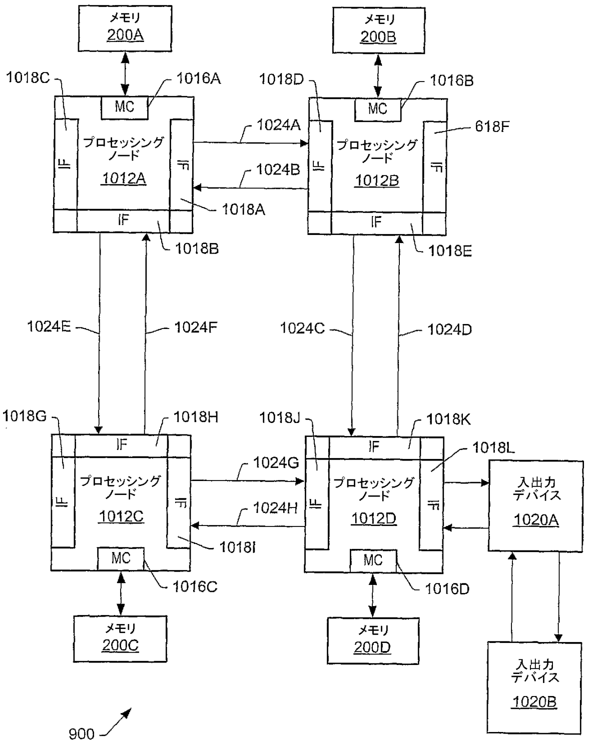
【図３Ｃ】



【図４】


キーボード/マウス 922
第2バスブリッジ 916
ISA/EISAバス
920
ISAデバイス 918
914
PCIバス
PCIデバイス 912B
PCIデバイス 912A
900
メインメモリ 200
メモリバス
906
バスブリッジ 902
AGPバス
910
画像コントローラ 908
ディスプレイ 926
CPUバス
924
L2キャッシュ 928
プロセッサ 100
プロセッサ 100a
L2キャッシュ 928a

【図５】


メモリ 200A
メモリ 200B
1018C
1016A
1018D
1016B
MC
1024A
618F
プロセッシングノード 1012A
プロセッシングノード 1012B
IF
1024B
1018A
1018B
1018E
1024E
1024F
1024C
1024D
1018G
1018H
1018J
1018K
1018L
1024G
プロセッシングノード 1012C
プロセッシングノード 1012D
1024H
1018I
入出力デバイス 1020A
1016C
1016D
メモリ 200C
メモリ 200D
入出力デバイス 1020B
900

フロントページの続き

(72)発明者　マイケル　エイ．　フィリッポ
アメリカ合衆国、テキサス州　７８６５２、マンチャカ、チャパラル　ロード　２０３０
(72)発明者　ジェイムズ　ケイ．　ピケット
アメリカ合衆国、テキサス州　７８７３３、オースティン、パロミノ　リッジ　ドライブ　ナンバー２　１７００
(72)発明者　ベンジャミン　ティー．　サンダー
アメリカ合衆国、テキサス州　７８７３５、オースティン、メディシン　クリーク　５７０１
(72)発明者　ラマ　エス．　ゴパル
アメリカ合衆国、テキサス州　７８７５９、オースティン、ミラモンテ　ドライブ　５９０６

審査官　三坂　敏夫

(56)参考文献　国際公開第０１／０５０２５３（ＷＯ，Ａ１）
特表２００３－５１９８３３（ＪＰ，Ａ）
特開２００２－１６３１０５（ＪＰ，Ａ）
国際公開第０２／０４２９０２（ＷＯ，Ａ１）
米国特許出願公開第２００３／００８８７６０（ＵＳ，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
G06F　9/38