

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4459628号
(P4459628)

(45) 発行日 平成22年4月28日 (2010. 4. 28)

(24) 登録日 平成22年2月19日 (2010. 2. 19)

(51) Int. Cl.

F I

G O 6 F 12/08 (2006. 01)

G O 6 F 12/08 5 0 7 Z

G O 6 F 12/10 (2006. 01)

G O 6 F 12/08 5 6 5

G O 6 F 12/10 5 5 1 Z

請求項の数 2 (全 11 頁)

(21) 出願番号 特願2003-581019 (P2003-581019)
 (86) (22) 出願日 平成15年3月13日 (2003. 3. 13)
 (65) 公表番号 特表2006-505834 (P2006-505834A)
 (43) 公表日 平成18年2月16日 (2006. 2. 16)
 (86) 国際出願番号 PCT/US2003/007710
 (87) 国際公開番号 W02003/083663
 (87) 国際公開日 平成15年10月9日 (2003. 10. 9)
 審査請求日 平成18年3月13日 (2006. 3. 13)
 (31) 優先権主張番号 10/104, 815
 (32) 優先日 平成14年3月22日 (2002. 3. 22)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 593096712
 インテル コーポレーション
 アメリカ合衆国 95052 カリフォル
 ニア州 サンタ クララ ミッション カ
 レッジ ブールバード 2200
 (74) 代理人 100070150
 弁理士 伊東 忠彦
 (74) 代理人 100091214
 弁理士 大貫 進介
 (74) 代理人 100107766
 弁理士 伊東 忠重
 (72) 発明者 ハマーランド, パー
 アメリカ合衆国 97124 オレゴン州
 ヒルズボロ ノースイースト セカンド
 ドライヴ 2601

最終頁に続く

(54) 【発明の名称】 キャッシュメモリにおけるコンテキスト識別子の使用

(57) 【特許請求の範囲】

【請求項 1】

第1のスレッドで使用されるページディレクトリのベースのアドレスと、第2のスレッドで使用されるページディレクトリのベースのアドレスとを比較し、前記第1のスレッドで使用されるページディレクトリのベースのアドレスが、前記第2のスレッドで使用されるページディレクトリのベースのアドレスと一致するかどうかを決定し、

前記第1のスレッドで使用されるページディレクトリのベースのアドレスが、前記第2のスレッドで使用されるページディレクトリのベースのアドレスと一致する場合に、前記第1のスレッド及び前記第2のスレッドの双方に同一のコンテキスト識別子を割り当て、

前記第1のスレッドで使用されるページディレクトリのベースのアドレスが、前記第2のスレッドで使用されるページディレクトリのベースのアドレスと一致しない場合に、前記第1のスレッド及び前記第2のスレッドに異なるコンテキスト識別子を割り当てることを有する方法。

【請求項 2】

プロセッサにより発行された命令に応じて、アドレスを生成するアドレス生成ユニット (AGU) と、

複数のキャッシュラインを保存し、前記複数のキャッシュラインのそれぞれは少なくともコンテキスト識別子と、タグと、データブロックとを有するキャッシュと、

前記 AGU に結合され、前記キャッシュに結合され、前記 AGU によって生成されたアドレス部分と、前記複数のキャッシュラインのそれぞれの各タグとを比較する少なくとも

10

20

1つの比較手段と、

実行中のコンピュータの命令の第1のスレッドに関連付けられたページディレクトリのベースのアドレスを保存する第1のレジスタと、

実行中のコンピュータの命令の第2のスレッドに関連付けられたページディレクトリのベースのアドレスを保存する第2のレジスタと、

前記第1のレジスタと前記第2のレジスタとに結合され、前記第1のレジスタに保存された値と、前記第2のレジスタに保存された値とを比較する他の比較手段と、

少なくとも2つのコンテキスト識別子を生成するコンテキスト識別子生成手段と、

前記コンテキスト識別子生成手段と前記他の比較手段とに結合され、前記少なくとも2つのコンテキスト識別子のうち1つを選択するマルチプレクサと、

前記マルチプレクサに結合され、前記キャッシュに結合され、前記マルチプレクサによって選択されたコンテキスト識別子と、前記複数のキャッシュラインのそれぞれの各コンテキスト識別子とを比較する少なくとも1つの比較手段と、

を有する装置。

【発明の詳細な説明】

【発明の詳細な説明】

【0001】

[技術分野]

本発明はコンピュータ技術に関する。特に、本発明は、コンピュータ装置のキャッシュメモリの性能を改善することに関する。

[背景技術]

コンピュータメモリの階層内で1つ以上のキャッシュメモリシステムを使用することは、コンピュータの性能を増大させる周知の技術である（例えば、1998年Academic PressによるHandy, JimのThe Cache Memory Bookを参照）。図1は典型的なキャッシュメモリアレイ100を示したものである。キャッシュメモリアレイ100は、キャッシュライン110を含む。各キャッシュラインは、タグ120とデータブロック130とを含む。例示のキャッシュライン140は、タグ150とデータブロック160とを含む。参照番号170は、例示のタグ150がメインメモリのアドレス170の一部であることを示す。メインメモリのアドレス170は、データブロック160に対応するメインメモリのアドレスである。

【0002】

プロセッサは、メモリ転送の間、プロセッサの実行コア部とコンピュータメモリの階層との間で、命令とオペランドを往復して転送する。メモリ転送の例は、メモリの階層からプロセッサに命令/オペランドをロードし、プロセッサからメモリの階層に命令/オペランドを保存することである。メモリ転送の間、プロセッサはメインメモリのアドレスを生成する。メインメモリのアドレスの一部が、キャッシュ検索の間にタグ120のエントリと比較され、キャッシュメモリ100が、メモリ転送に対応するエントリを有するかどうか決定する。タグ150とメインメモリのアドレス170との間の関係によって示される通り、キャッシュ検索の処理は、プロセッサが各メインメモリのアドレスの一部のみとタグの各エントリとを比較することを求めることによって加速される。一般的に、キャッシュメモリはプロセッサにより生成された各リニアアドレスの一部を用いて、キャッシュアレイ100に保存されるデータに索引を付ける。

【0003】

スレッドは、コンピュータプログラムの他の部分と関係なく実行し得るコンピュータプログラムの一部である。複数のスレッドがプロセッサで同時に実行されると、プロセッサの性能は高められ得る。複数のスレッドの多様な命令の間の依存関係が検出され、適切に管理されると、複数のスレッドの同時の実行が可能になる。

【0004】

図2は、多くのIntel（商標）プロセッサが、仮想メモリ環境を使用して、多くのリニアアドレス空間が小さい量の物理メモリ（例えば、ランダムアクセスメモリ）によっ

10

20

30

40

50

て対応可能にする方法を示したものである。メモリ転送の間、プロセッサはリニアアドレス210を生成する。リニアアドレス210は、ディレクトリ領域220と、テーブル領域225と、オフセット領域230とを有する。ページディレクトリ235のベースは、制御レジスタCR3240に含まれる。リニアアドレス210のディレクトリエントリ220は、制御レジスタCR3240に含まれる値へのオフセットを与える。ページディレクトリは、ページ領域のベースポインタ245を含む。テーブル領域225は、ページテーブルのベースポインタ245に結びつけられたオフセットを備え、物理アドレス255を含むページのベースを識別する。オフセット領域230は、ページテーブルのエントリと結びつけられ、物理アドレス255を識別する。

【0005】

10

図3は、一部のメインメモリのアドレスのみのキャッシュ検索を実行することに関する欠点を示したものである。プロセッサ(図示なし)は、ロード命令LD0とLD1に応じて、リニアアドレス304と306とを生成する。305と307は、キャッシュ検索を実行するために用いられる各アドレスの一部を示したものである。305と307が同一に見えるが、それぞれ304と306の一部にしか過ぎない。304と306は、各アドレスがそれぞれのディレクトリ領域(320と325)とオフセット領域(330と335)に異なるエントリを有するため、2つの異なる物理アドレスに位置する。図3の通り、プロセッサが複数のスレッドの同時実行に対応する場合は、追加の複雑な状態が持ち込まれる。スレッド0とスレッド1は、それぞれのページディレクトリ(340と345)のベースに異なる値を有し得る。従って、304と306が同じであっても、それらは2

20

[詳細な説明]

本発明は例示の目的で示され、限定の目的で示されるのではない。添付の図面の図において、類似の参照番号は類似の要素を参照する。

【0006】

本願でキャッシュメモリのコンテキスト識別子の使用について説明される。以下の説明において、本発明の実施例の十分な理解を提供するため、多数の特定の詳細が提供される。しかし、関連技術に熟練した人は、1つ以上の特定の詳細を用いずに、又は他の方法や、構成要素や、材料等を用いて、本発明が実行され得ることがわかるだろう。他の場合において、本発明の形態を不明確にすることを回避するために、周知の構造や、材料や、動作が、詳細に示されない、又は説明されない。

30

【0007】

本明細書を通じて、“一実施例”又は“実施例”への言及は、本発明に関連して説明される特定の機能又は構造又は特徴が、本発明の少なくとも1つの実施例に含まれることを意味する。従って、本明細書を通じて多様な場所における“一実施例において”又は“実施例において”という語句が現れることは、必ずしも全てが同じ実施例を参照するのではない。さらに、特定の機能又は構造又は特徴は、1つ以上の実施例において、何らかの適当な方法と組み合わせられ得る。

【0008】

キャッシュメモリシステムは、タグと呼ばれるリニアアドレスの一部でキャッシュラインに索引を付ける。プロセッサは、メモリ転送動作の間にリニアアドレスを生成する。キャッシュ検索の間に、前記リニアアドレスの一部がタグと比較される。多くのIntel(商標)プロセッサは、一部のアドレスとタグとの間の見かけ上の一致が真の一致であることを確保する仕組みを提供する。前記仕組みは、ヒット/ミス決定手段と呼ばれる。ヒット/ミス決定手段は、キャッシュメモリのキャッシュ検索処理と同時に動作する。前記ヒット/ミス決定手段は、発行された命令に関連付けられたリニアアドレスを、対応する物理アドレスに変換し、前記一部のリニアアドレスと一致するタグに関連付けられた完全なアドレスと、物理アドレスとを比較する。前記ヒット/ミス決定手段は、リニアアドレスを物理アドレスに変換しなければならないため、及び、完全なアドレスの比較を実行するため、前記ヒット/ミス決定手段はキャッシュメモリよりかなり遅い。キャッシュメモ

40

50

リが誤った一致を引き起こすことを前記ヒット/ミス決定手段が検出すると、前記ヒット/ミス決定手段はキャッシュメモリによって行われた誤りを訂正する。キャッシュメモリで誤った一致によって持ち込まれた誤りを検出し、それに対応する処理は、プロセッサの性能を低下し得る。

【0009】

図4Aは、コンテキスト識別子を組み込んだキャッシュメモリアレイの一実施例を示したものである。キャッシュメモリアレイ400は、複数のキャッシュライン410を含む。各キャッシュライン410は、コンテキスト識別子405と、タグ420と、データブロック430とを含む。従って、キャッシュライン440Aは、コンテキスト識別子415Aとタグ450Aとを含む。同様に、キャッシュライン440Bは、コンテキスト識別子415Bとタグ450Bとを含む。2つの異なるスレッドがコンテキスト識別子415Aと415Bとを生成する。

10

【0010】

リニアアドレス407Aと470Bがタグ450Aと450Bとに対応する。図4Aと4Bにおいて、リニアアドレスはコンテキスト識別子を含み、類似のリニアアドレスの間の識別するためのコンテキスト識別子の用い方について示す。コンテキスト識別子は、リニアアドレスの一部ではないが、前記リニアアドレスと別に生成される。

【0011】

先行技術に従って実現されたキャッシュメモリシステムは、プロセッサ(図示なし)によって生成された一部のリニアアドレスと、キャッシュアレイ400の各タグ420とを比較することによってキャッシュ検索を行う。プロセッサが、123という一部のリニアアドレスを生成した場合は、前記アドレスは、タグ450Aとタグ450Bとの双方に一致する。前述の通り、ヒット/ミス決定手段は、どちらのキャッシュラインが前記プロセッサにより生成されたメインメモリのアドレスに一致するかどうかを最終的に決定する。先行技術のシステムは、前記ヒット/ミス決定手段が一部のリニアアドレスとタグ420との間の誤った一致を管理して訂正する場合に、性能の不利を受ける。

20

【0012】

図4Bは、前記プロセッサにより生成された一部のリニアアドレスとタグ420との間の誤った一致の頻度を減少させるコンテキスト識別子の使用を示したものである。リニアアドレス470Aと470Bは、図4Aのキャッシュライン440Aと440Bに関連付けられる。発行された命令に応じて、アドレス475がプロセッサによって生成される。アドレス475は、コンテキスト識別子480と一部のリニアアドレス485とを含む。一部のリニアアドレス485がタグ450Bと一致するため、先行技術のキャッシュは475と470Bとの間の誤った一致を識別する場合がある。一実施例において、コンテキスト識別子480がコンテキスト識別子415Bと比較される。また、一実施例において、一部のリニアアドレス485がタグ450Bと比較される。従って、コンテキスト識別子480がコンテキスト識別子415Bと一致しないため、キャッシュアレイ400は475と470Bとの間の誤った一致を見つけない。

30

【0013】

図5は、ページミスハンドラー(PMH)の一部の一実施例を示したものである。一実施例において、PMH500は、コンテキスト識別子の割り当てを生成し、制御する。PMH500は、コンテキスト識別子生成手段510と、比較手段520と、コンテキスト識別子マルチプレクサ530と、インバーター540とを含む。一実施例において、コンテキスト識別子生成手段510は、現在実行するスレッドに対応する1ビットのバイナリ番号を生成する。本発明の他の実施例において、コンテキスト識別子生成手段510は、1ビット以上のバイナリ番号を生成する。比較手段520は、プロセッサで実行する2つのスレッドにより用いられるページディレクトリのベースのアドレスを比較する。異なる実施例において、コンテキスト識別子は2つ以上のビットを有する。

40

【0014】

一実施例によると、PMH500は、同じコンテキスト識別子が2つの異なるスレッド

50

の命令に割り当てられているかどうかを決定する。図5に示される実施例では、PMH500は、前記プロセッサで実行する第1のスレッドに第1のコンテキスト識別子を割り当てる。例えば、スレッド0が前記プロセッサで実行する前記第1のスレッドである場合は、前記プロセッサは、スレッド0のコンテキスト識別子として0を割り当て得る。コンテキスト識別子は、スレッド番号と一致する必要はない。従って、スレッド0は1又は0のコンテキスト識別子を有し得る。本発明の他の実施例において、プロセッサは2以上のスレッドの命令を実行し、1ビット以上のコンテキスト識別子が使用される。

【0015】

第2のスレッドがプロセッサで実行し始めると（例えば、スレッド1）、PMH500は、前記第2のスレッドの命令に、スレッド0の命令に割り当てられたコンテキスト識別子と異なるコンテキスト識別子を割り当てるかどうかを決定する。スレッド0とスレッド1がメインメモリの同じページを使用することが可能である。その場合は、リニアアドレスがタグと同じページに位置することが知られている場合に、タグとリニアアドレスの間の誤った一致があまり生じないため、同じコンテキスト識別子を2つのスレッドの命令に割り当てることが望ましい。比較手段520は、2つのスレッドにより用いられるページディレクトリのベースを比較することにより、スレッド0とスレッド1がメインメモリで同じページを使用しているかどうかを決定する。

【0016】

スレッド0とスレッド1が、それぞれのページディレクトリのベースとして、同じアドレスを使用している場合、マルチプレクサ530は、スレッド0に選択されたものと同じコンテキスト識別子をスレッド1にも選択する。逆に言えば、スレッド0とスレッド1がページディレクトリのベースに同じアドレスを使用していないことを比較手段520が示すと、マルチプレクサ530は、スレッド1のコンテキスト識別子として、インバーター540の出力を選択する。2つのスレッドが同じコンテキスト識別子を共有するかどうかを決定することは、異なる方法で実現され得る。例えば、本発明の実施例においては、2つのスレッドが同じコンテキスト識別子を共有しているかどうかをオペレーティングシステムが決定し、プロセッサに正しいコンテキスト識別子を明示的に知らせ、スレッドに関連づける。本発明の更に他の実施例においては、スレッドのメモリアクセスの傾向を学習する予測手段が、コンテキスト識別子を予測する。一実施例によると、2つの異なるスレッドの命令に同じ値を常に選択するように、又は2つの異なるスレッドの命令に異なるコンテキスト識別子の値を常に選択するようにユーザがPMH500に指示し得る。また、ある実施例においては、ベースディレクトリのアドレスの他に追加の基準を使用し、2つのスレッドが同じコンテキスト識別子を共有するかどうかを決定する。

【0017】

本発明の他の実施例において、前記プロセッサに2つ以上のスレッドが存在する。前記プロセッサにn個のスレッドが存在する実施例においては、前記コンテキスト識別子は、n個の異なるコンテキスト識別子が特定されることを可能にするエンコードの仕組みまで拡張される。nビットのコンテキスト識別子が使用される場合は、単一のビットがキャッシュに位置されるのと同じ方法でnビットがキャッシュに位置されることが、その技術に通常熟練した人はわかるだろう。また、2つの単一のビットのコンテキスト識別子を比較するのはほぼ同じ方法で、nビットのコンテキスト識別子をお互いに比較することが行われ得ることが、その技術に通常熟練した人はわかるだろう。

【0018】

複数のスレッドを備える本発明の実施例は、未使用のコンテキスト識別子のプールを使用する。スレッドがページテーブルのベース（CR3）を変更すると、他のアクティブなスレッドで使用されるページテーブルのベースの値、又は前記プールのコンテキスト識別子の何らかの有効なページテーブルのベースの値と新しい値が比較される。一致している場合には、前記一致を引き起こしたスレッドのコンテキスト識別子の値が取られる。他のアクティブなスレッドと一致しない場合には、前記プールから未使用のコンテキスト識別子が取られる。スレッドがページのベースの値を変更する場合、又は所定のコンテキスト

10

20

30

40

50

識別子を備えたキャッシュの最後のラインが割り当てを外される場合に、コンテキスト識別子は未使用のプールに戻される。後者の構成は、以前のスレッドに持ち込まれたデータでスレッドがヒットすることを可能にする。スレッドのコンテキスト識別子の値を予測する多くの仕組みが使用され得ることが、その技術に通常熟練した人はわかるだろう。

【 0 0 1 9 】

図 6 は、コンテキスト識別子を有するキャッシュメモリシステムからの読み取りを示したものである。キャッシュアレイ 6 1 0 は、キャッシュライン 6 0 5 を含む。各キャッシュライン 6 0 5 は、コンテキスト識別子 6 1 0 A と、タグ 6 1 0 B と、データブロック 6 1 0 C とを含む。キャッシュアレイ 6 1 0 はキャッシュコントローラ 6 2 0 に結合される。キャッシュコントローラ 6 2 0 は、キャッシュアレイ 6 1 0 に含まれるデータへのアクセスを制御する。

10

【 0 0 2 0 】

キャッシュコントローラ 6 2 0 は、ヒット / ミス決定手段 6 7 0 及び比較手段 6 7 5 と 8 0 に結合される。ヒット / ミス決定手段 6 7 0 及び比較手段 6 7 5 と 6 8 0 の動作は、以下に説明する。一実施例において、比較手段 6 7 5 は A G U 6 3 0 と A N D ゲート 6 8 5 とに結合される。A G U 6 3 0 は、メモリ転送を必要とする発行された命令（例えば、ロードと保存の命令）のためのリニアアドレスを生成する。一実施例において、比較手段 6 8 0 は、マルチプレクサ 6 5 0 を通して、A N D ゲート 6 8 5 と P M H 6 4 0 とに結合する。P M H 6 4 0 は、どのコンテキスト識別子がどの発行された命令に割り当てられるかを決定するために用いられる制御ロジックを含む。

20

【 0 0 2 1 】

一般的に、読み取り周期は、発行されたロード命令（L D）6 6 0 によって開始される。L D 6 6 0 は、P M H 6 4 0 と A G U 6 3 0 とに送信される。特に、P M H 6 4 0 はどのスレッドが L D 6 6 0 の出所であるかを決定する。P M H 6 4 0 は、一つには、どのスレッドが L D 6 6 0 の出所であるかに基づいて、L D 6 6 0 にコンテキスト識別子を割り当てる。P M H 6 4 0 は、L D 6 6 0 に関連付けられたコンテキスト識別子を、マルチプレクサ 6 5 0 を通して、比較手段 6 8 0 に発行する。同様に、A G U 6 3 0 は、一部のリニアアドレスを比較手段 6 7 5 に発行する。

【 0 0 2 2 】

比較手段 6 7 5 と 6 8 0 は前記一部のリニアアドレスと L D 6 6 0 のコンテキスト識別子とを用い、L D 6 6 0 が必要とする情報をどのキャッシュアレイ 6 1 0 が含むかを決定する。キャッシュアレイ 6 1 0 が必要な情報を含む場合、キャッシュヒットが結果として生じる。キャッシュアレイ 6 1 0 が必要な情報を含まない場合、キャッシュミスが結果として生じる。

30

【 0 0 2 3 】

特定のキャッシュラインが、L D 6 6 0 のコンテキスト識別子に一致するコンテキスト識別子と、L D 6 6 0 に対応する一部のリニアアドレスに一致するタグとを含むことを、キャッシュヒットは要求する。比較手段 6 7 5 と 6 8 0 は、各キャッシュラインのタグ領域及びコンテキスト識別子領域と、L D 6 6 0 のコンテキスト識別子及び一部のリニアアドレスとを比較する。一実施例によると、比較が単一の平行動作で行われ得るように、キャッシュラインと同じだけの比較手段が存在する。単一のキャッシュラインが、L D 6 6 0 のコンテキスト識別子及び一部のリニアアドレスに一致するコンテキスト識別子及びタグを含む場合、A N D ゲート 6 8 5 がキャッシュヒットの信号を出す。しかし、比較手段 6 7 5 が L D 6 6 0 のアドレスの一部のみと、キャッシュアレイ 6 1 0 A の各タグとを比較すると、キャッシュヒットが誤った一致に基づくという可能性が存在する。

40

【 0 0 2 4 】

ヒット / ミス決定手段 6 7 0 は、キャッシュヒットが誤った一致に基づくかどうかを検出する。ヒット / ミス決定手段は、L D 6 6 0 が命令 / オペランドを要求するメインメモリの位置の完全なリニアアドレスを受信する。一実施例によると、ヒット / ミス決定手段 6 7 0 は、A G U によってそれに伝達されるリニアアドレスを物理アドレスに変換する。

50

ヒット/ミス決定手段 670 は、LD 660 の一部のリニアアドレスに一致するタグに対応する物理アドレスを決定する。ヒット/ミス決定手段 670 は、完全な物理アドレスを比較し、キャッシュヒットが誤った一致に基づくかどうか決定する。

【0025】

図7は、コンテキスト識別子を用いたキャッシュメモリへの書き込みを示したものである。キャッシュアレイ 710 は複数のキャッシュライン 705 を含む。各キャッシュライン 705 はコンテキスト識別子 710A と、タグ 710B と、データブロック 710C とを含む。キャッシュアレイ 710 は、キャッシュコントローラに結合される。キャッシュコントローラ 720 は、どのキャッシュラインにデータのブロックが書き込まれるかを決定することを含む複数の機能を実行する。キャッシュコントローラ 720 は、マルチプレクサ 750 を通して AGU 730 と PMH 740 とに結合される。

10

【0026】

発行された命令 760 がプロセッサ(図示なし)から PMH 740 と AGU 730 とに送信される。一実施例において、発行された命令 760 がメモリ転送を要求する場合に、AGU 730 はリニアアドレスを生成する。AGU 730 は、概してプロセッサの実行コアの一部である。他の実施例において、AGU 730 は物理アドレスを生成し得る。一実施例において、PMH 740 は、マルチプレクサ 750 を制御する少なくとも2つのコンテキスト識別子と制御信号とを生成する。他の実施例において、PMH 740 は何からの数のコンテキスト識別子を生成し得る。

【0027】

20

一般的に、書き込み周期は、発行された保存命令(ST)(例えば、ST 760)によって開始される。ST 760 は、PMH 740 と AGU 730 とに送信される。PMH 740 は、どのスレッドが ST 760 の出所であるか決定する。PMH 740 は、どのコンテキスト識別子が ST 760 に割り当てられるかを更に決定する。一実施例によると、PMH 740 は、3つの動作モード: OFF と NON-SHARED と ADAPTIVE を有する。PMH 740 が OFF モードで動作している場合、PMH 740 は、各コンテキスト識別子に同じ値を書き込み得る。スレッドがキャッシュラインに保存されたコンテキスト識別子と同じコンテキスト識別子を有する場合、2つ以上のスレッドの命令によってキャッシュラインが共有される。

【0028】

30

NON-SHARED モードで動作する場合、PMH 740 は発行された命令を含むスレッドに特有のコンテキスト識別子を割り当て得る。一実施例において、ADAPTIVE モードにより、異なるスレッドから生じる同じ又は異なるコンテキスト識別子を割り当てかどうかを PMH 740 が動的に判断することが可能になる。一実施例において、スレッドがページディレクトリのベースに同じアドレスを共有する場合は、PMH 740 は、スレッドに含まれる命令に同じコンテキスト識別子を動的に割り当てる。ある実施例においてそうでない場合は、PMH 740 は、スレッドに含まれる命令に異なるコンテキスト識別子を動的に割り当てる。

【0029】

一実施例において、2つの異なるスレッドの命令に同じコンテキスト識別子を書き込むかどうか決定する場合に、PMH 740 は、コンピュータがマルチスレッドのモードにあるかどうかと、ページングが可能かどうかとを検討する。発行された命令に割り当てられるコンテキスト識別子の値を決定するために、いくつかの基準が用いられ得る。更に、コンテキスト識別子生成手段と、関連する制御ロジックとが、電子システム内の複数の場所に配置され得る。そして、それらは PMH 又は類似のユニットに含まれる必要はない。

40

【0030】

PMH 740 は、ST 760 に適切なコンテキスト識別子を、マルチプレクサ 750 を通してキャッシュコントローラ 720 に送信する。AGU 730 は ST 760 A がデータを保存するために探すメインメモリの場所に対応したりニアアドレスを生成する。AGU 730 は、キャッシュコントローラ 720 とヒット/ミス決定手段 770 に ST 760 に

50

対応するリニアアドレスを提供する。キャッシュコントローラ720は、どのキャッシュラインがST760のデータを保存するために使用されるかを決定するキャッシュライン交換方法を用いる。キャッシュライン交換方法は、その技術において周知である(キャッシュライン交換方法のより詳細な説明は、1998年Academic PressによるHandy, JimのThe Cache Memory Bookを参照)。

【0031】

適切なキャッシュラインを選択した後、キャッシュコントローラ720は前記選択されたキャッシュライン780のコンテキスト識別子領域780Aに、コンテキスト識別子を書き込む。キャッシュコントローラ720は、キャッシュライン780のタグ領域780BにST760のリニアアドレスの一部を書き込む。一実施例において、前記タグは、32ビットリニアアドレスの23から12ビットを有する。前記タグに使用され得るアドレスビットの多くの組み合わせが存在する。キャッシュコントローラは、選択されたキャッシュライン780のデータブロック領域780CにST760に関連付けられたデータを書き込む。

【0032】

キャッシュ管理は、電子的にアクセス可能な媒体からの命令によって制御され得る。それは、本願で説明した処理を実行するコンピュータ(又は他の電子装置)をプログラムするために用いられ得る。電子的にアクセス可能な媒体は、フロッピー(登録商標)ディスクや、光ディスクや、CD-ROMや、光磁気ディスクや、ROMや、RAMや、EPROMや、EEPROMや、磁気カード又は光カードや、フラッシュメモリや、電子命令を保存するのに適した他の形式の媒体/機械が読み取り可能な媒体を含むが、それに限定されない。さらに、前記命令はまた、コンピュータプログラム製品としてダウンロードされ得る。前記プログラムは、通信リンク(例えば、モデム又はネットワーク接続)を介して搬送波又は他の伝播媒体に組み込まれたデータ信号を用いて、リモートコンピュータから要求中のコンピュータに転送され得る。

【0033】

本発明の示された実施例の前述の説明は、網羅的にすることを意図するものではなく、また、開示された正確な形式に本発明を限定することを意図するものではない。本発明の特定の実施例及び例は、説明の目的で本願において開示されるが、関連技術に熟練した人がわかるように、本発明の範囲ないで多様な同等の改良が可能である。

【0034】

前記改良は、前述の詳細な説明を踏まえて行われ得る。特許請求の範囲で使用される用語は、明細書及び特許請求の範囲で開示された特定の実施例に本発明を限定するために解釈されるべきではない。むしろ、本発明の範囲は、特許請求の範囲によって完全に決定されるべきであり、特許請求の範囲の解釈の確立した原理に従って解釈されるべきである。

【図面の簡単な説明】

【0035】

【図1】先行技術のキャッシュメモリアレイの図である。

【図2】リニアアドレスを物理アドレスに位置付ける概念図である。

【図3】2つの異なる物理アドレスに位置する2つの類似のタグの概念図である。

【図4A】コンテキスト識別子を用いたキャッシュメモリアレイの一実施例の図である。

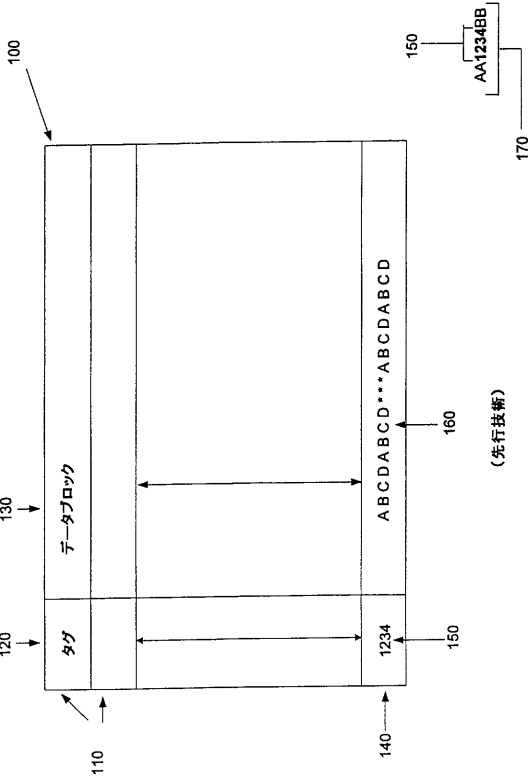
【図4B】コンテキスト識別子の一実施例の概念図である。

【図5】コンテキスト識別子生成手段と関連する制御回路の一実施例の簡略化したブロック図である。

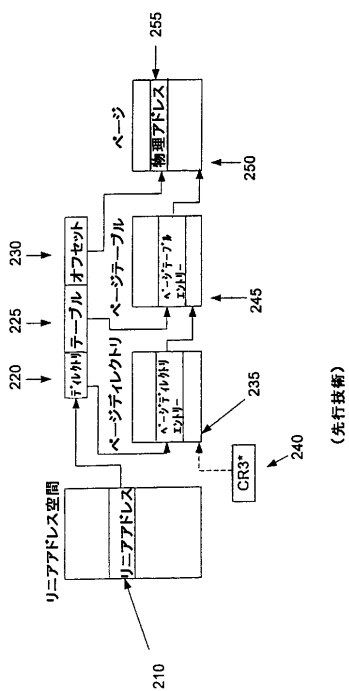
【図6】キャッシュメモリシステムの一実施例のブロック図である。

【図7】キャッシュメモリシステムの一実施例のブロック図である。

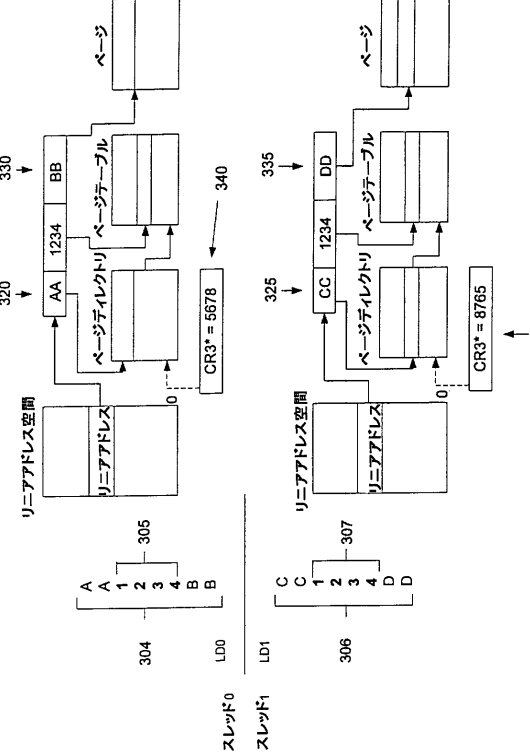
【図 1】



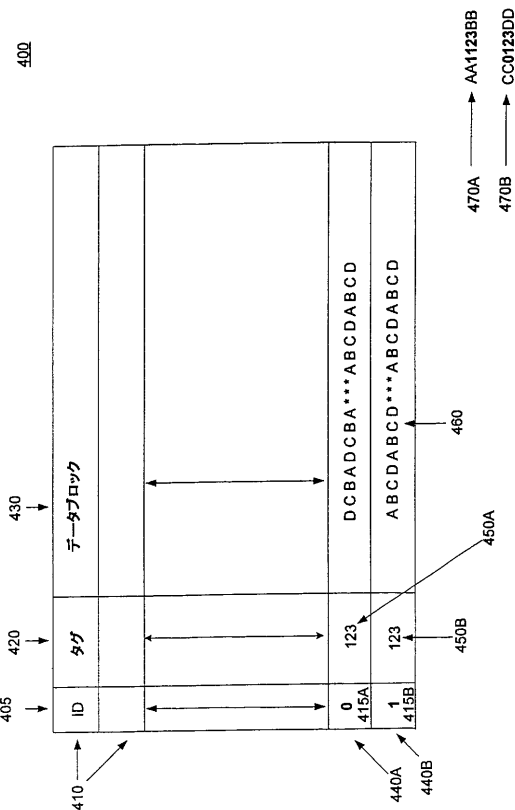
【図 2】



【図 3】



【図 4 A】



【図 4 B】

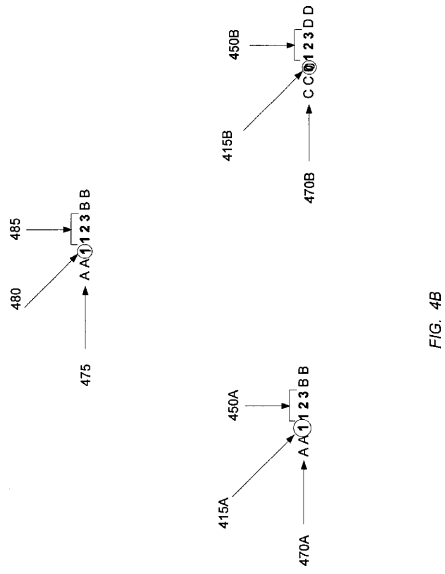
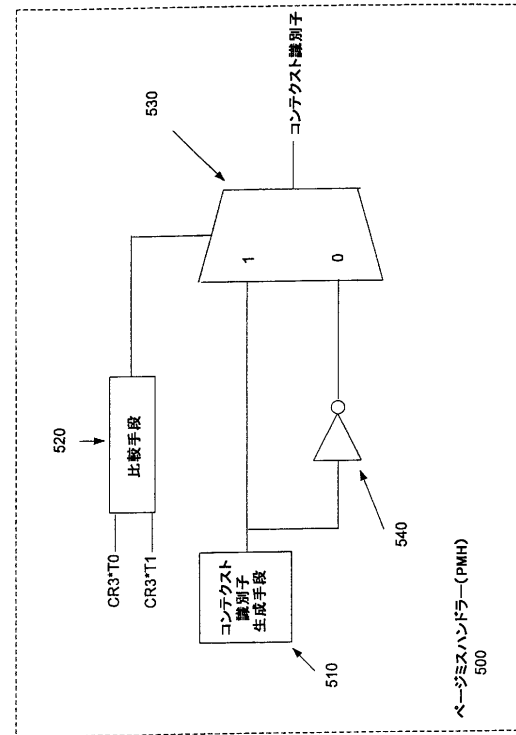
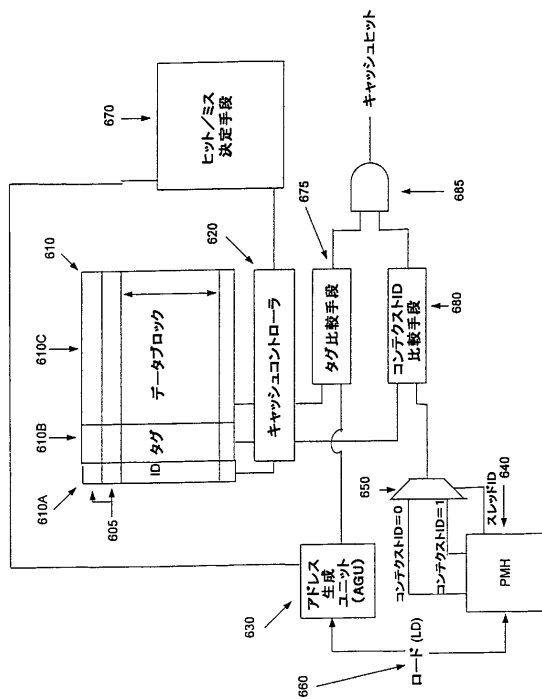


FIG. 4B

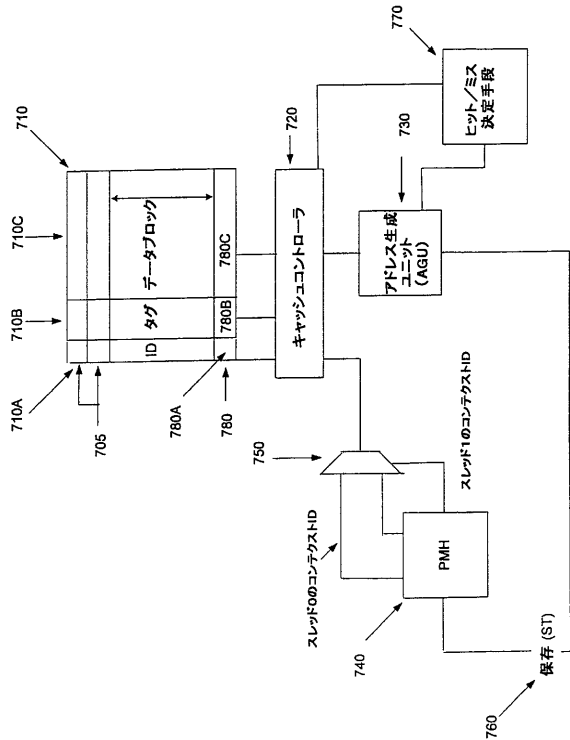
【図 5】



【図 6】



【図 7】



フロントページの続き

- (72)発明者 ヴェンカットラマン, ケイ エス
アメリカ合衆国 97123 オレゴン州 ヒルズボロ サウスイースト ラングウッド ストリート 7219
- (72)発明者 バクサ, アラヴィンド
アメリカ合衆国 97229 オレゴン州 ポートランド ノースウエスト ファルコンリッジレーン 14362
- (72)発明者 アップトン, マイケル
アメリカ合衆国 97210 オレゴン州 ポートランド ノースウエスト 24ス アヴェニュー 1410

審査官 清木 泰

- (56)参考文献 特開平10-187538(JP, A)
特開平03-235144(JP, A)
特開平03-235143(JP, A)

- (58)調査した分野(Int.Cl., DB名)
G06F12/08-12/12