

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6196310号  
(P6196310)

(45) 発行日 平成29年9月13日 (2017. 9. 13)

(24) 登録日 平成29年8月25日 (2017. 8. 25)

(51) Int. Cl.

F I

G 0 6 F 12/08 (2016.01)

G 0 6 F 12/08 5 1 7 C

G 0 6 F 12/08 5 1 5 B

G 0 6 F 12/08 5 0 7 Z

G 0 6 F 12/08 5 1 1 E

G 0 6 F 12/08 5 7 3

請求項の数 15 (全 23 頁)

(21) 出願番号 特願2015-533311 (P2015-533311)  
 (86) (22) 出願日 平成25年9月26日 (2013. 9. 26)  
 (65) 公表番号 特表2015-534687 (P2015-534687A)  
 (43) 公表日 平成27年12月3日 (2015. 12. 3)  
 (86) 国際出願番号 PCT/US2013/061876  
 (87) 国際公開番号 W02014/052561  
 (87) 国際公開日 平成26年4月3日 (2014. 4. 3)  
 審査請求日 平成28年9月12日 (2016. 9. 12)  
 (31) 優先権主張番号 13/626, 916  
 (32) 優先日 平成24年9月26日 (2012. 9. 26)  
 (33) 優先権主張国 米国 (US)

(73) 特許権者 507364838  
 クアルコム、インコーポレイテッド  
 アメリカ合衆国 カリフォルニア 921  
 21 サン ディエゴ モアハウス ドラ  
 イヴ 5775  
 (74) 代理人 100108453  
 弁理士 村山 靖彦  
 (74) 代理人 100163522  
 弁理士 黒田 晋平  
 (72) 発明者 レズリー・マーク・デブリュイヌ  
 アメリカ合衆国・カリフォルニア・921  
 21・サン・ディエゴ・モアハウス・ドラ  
 イヴ・5775

最終頁に続く

(54) 【発明の名称】 異なるキャッシュ可能性を用いてページ横断命令を管理するための方法および装置

(57) 【特許請求の範囲】

【請求項 1】

キャッシュ可能命令および、キャッシュ可能命令の第1のページからフェッチされた開始部分と、  
キャッシュ不可能命令の第2のページからフェッチされ、キャッシュラインの  
最後でキャッシュライン境界を横断する終了部分とを有する命令を記憶するステップと、  
前記命令の実行許可が拒否されたことを指示するステップであって、前記命令がキャッ  
シュ不可能命令として特定される、ステップと、

前記キャッシュ不可能命令を含む命令のフェッチグループにヒットしたときに前記指示  
を検出するステップであって、前記キャッシュ不可能命令は、前記指示に応答してキャッ  
シュをバイパスして、実行のためにメモリから再フェッチされる、ステップと、  
 を含む方法。

【請求項 2】

前記方法は、異なるキャッシュ可能性を用いてページ横断命令を管理するためのもので  
あり、

前記検出するステップは、確立されたキャッシュラインにヒットすると、前記命令を命  
令キャッシュから実行することができないことを示すためのキャッシュ不可能フラグを設  
定するステップであって、前記命令が受信されるが、前記キャッシュ不可能フラグに基づ  
いて前記キャッシュから実行されない、ステップを含む、

前記再フェッチは、前記命令の前記開始部分と組み合わせるために設定された前記キャ  
ッシュ不可能フラグに 応答してキャッシュをバイパスして、前記命令の少なくとも前記終

了部分をメモリから再フェッチするステップであって、前記命令が実行のために再構築される、ステップを含む、  
請求項1に記載の方法。

【請求項3】

前記命令が、キャッシュ不可能命令を有するページとキャッシュ可能命令を有するページとの間のページ境界にわたってフェッチされる、請求項2に記載の方法。

【請求項4】

前記キャッシュラインが、各キャッシュラインに少なくとも(K/2)ビットに加えてN個のKビットの容量を有する前記キャッシュライン内のアラインされた境界上で命令を記憶する、請求項2に記載の方法。

【請求項5】

前記キャッシュ可能命令が、可変長命令セットアーキテクチャから選択されたKビット命令およびK/2ビット命令を有する命令のセットである、請求項2に記載の方法。

【請求項6】

前記指示が、1つまたは複数のビットに記憶されて、前記命令の実行許可が拒否されたことを示す、請求項2に記載の方法。

【請求項7】

前記キャッシュからフェッチグループにおいて前記命令をフェッチするステップと、  
前記キャッシュ不可能フラグを検出すると、フェッチパイプラインをストールするステップと、

プロセッサが実行のために前記命令に到達したと判断すると、前記フェッチパイプラインをフラッシュするステップと

をさらに含む、請求項2に記載の方法。

【請求項8】

キャッシュ可能命令および、キャッシュ可能命令の第1のページからフェッチされた開始部分と、キャッシュ不可能命令の第2のページからフェッチされ、キャッシュラインの最後でキャッシュライン境界を横断する終了部分とを有する命令を記憶するための手段と

、  
前記命令の実行許可が拒否されたことを指示するための手段であって、前記命令がキャッシュ不可能命令として特定される、手段と、

前記キャッシュ不可能命令を含む命令のフェッチグループにヒットしたときに前記指示を検出するための手段であって、前記キャッシュ不可能命令は、前記指示に応答してキャッシュをバイパスして、実行のためにメモリから再フェッチされる、手段と、  
を備える、装置。

【請求項9】

前記装置は、異なるキャッシュ可能性を用いてページ横断命令の実行を制御するためのものであり、

前記キャッシュ可能命令を記憶するための手段は、キャッシュ可能命令および、キャッシュ可能である開始部分と、キャッシュ不可能であり、キャッシュラインの最後でキャッシュライン境界を横断する終了部分とを有する命令を記憶するように構成可能な命令キャッシュであり、

前記指示するための手段は、前記命令の実行許可が拒否されたという指示を1つまたは複数のビットに記憶するように構成可能なインジケータ回路であって、前記命令がキャッシュ不可能命令として特定される、インジケータ回路であり、

前記検出するための手段は、プロセッサに結合され、前記キャッシュ不可能命令を含む命令のフェッチグループにヒットしたときに前記指示を検出するように構成可能なフェッチパイプラインであって、前記キャッシュ不可能命令が受信されるが、前記指示に応答して前記キャッシュから実行されない、フェッチパイプラインである、  
請求項8に記載の装置。

【請求項10】

前記フェッチパイプラインが、前記キャッシュ不可能命令の前記開始部分と組み合わせ、前記キャッシュ不可能命令を実行のために再構築するために、前記検出された指示にตอบสนองしてキャッシュをバイパスして、前記キャッシュ不可能命令の少なくとも前記終了部分をメモリからフェッチするように動作可能であり、

前記キャッシュラインが、 $K/2$ ビット幅を加えた $N \times K$ ビットである各キャッシュライン内のアラインされた境界上で $K$ ビット幅の命令および $K/2$ ビット幅の命令を記憶し、前記キャッシュラインの $K/2$ ビット部分が、ライン横断命令の $K/2$ ビット終了部分を記憶するか、または、

前記終了部分が、次のキャッシュラインで確立される終了部分の複製である、請求項9に記載の装置。

10

【請求項11】

前記命令は、キャッシュ不可能であり、キャッシュラインの最後でページ境界およびキャッシュライン境界を横断する終了部分を有する、請求項9に記載の装置。

【請求項12】

前記インジケータ回路が前記命令キャッシュの外部にあるか、または、

関連付けられたウェイに記憶された前記キャッシュ不可能なライン横断命令の実行許可が拒否されたというウェイ指示を記憶するように構成可能な、各キャッシュウェイのためのキャッシュウェイインジケータ回路をさらに備える、

請求項11に記載の装置。

20

【請求項13】

プロセッサパイプラインが、

前記キャッシュ不可能命令の前記終了部分と、前記指示が前記プロセッサパイプラインで受信されたことを検出するように構成可能な検出回路と、

前記プロセッサパイプラインへの前記キャッシュ不可能命令を特定するように構成可能なデコード回路と、

前記命令キャッシュから受信された命令をバッファし、前記キャッシュ不可能命令が前記プロセッサパイプラインの中で実行に利用可能な最も古い命令であるかどうかを判断するように構成可能なバッファ及びホールド回路と

を備える、請求項11に記載の装置。

【請求項14】

30

前記キャッシュ不可能命令が、実行に利用可能な最も古い命令であると判断され、前記キャッシュ不可能命令を含む前記プロセッサパイプラインが背後でフラッシュされるか、または、

前記キャッシュ不可能命令の前に受信された分岐命令が、前記プロセッサに、前記キャッシュ不可能命令から分岐させ、前記キャッシュ不可能命令が、前記バッファ及びホールド回路で上書きされる、請求項13に記載の装置。

【請求項15】

実行されると、請求項1乃至7の何れか1項に記載の方法を実施するように動作可能なコンピュータ可読プログラムデータおよびコードで符号化された、非一時的コンピュータ可読記録媒体。

40

【発明の詳細な説明】

【技術分野】

【0001】

〔関連出願の相互参照〕

「Method and Apparatus for Tracking Extra Data Permissions in an Instruction Cache」という名称の2012年6月28日に出願された米国出願第13/535,398号は関連出願であり、その全体が本明細書に組み込まれる。

【0002】

本発明は一般に、様々な長さを有する命令を処理するプロセッサの効率を改善するための技法に関し、より詳細には、キャッシュ可能なメモリとキャッシュ不可能なメモリとの

50

間の境界を横断する命令を特定し、この命令全体を他のキャッシュ可能命令とともにキャッシュラインに記憶することを可能にするための有利な技法に関する。

【背景技術】

【0003】

たとえば、8ビット命令、16ビット命令、32ビット命令、および64ビット命令などの異なる長さの命令を実行するためのいくつかのプロセッサが設計されている。そのようなプロセッサのためのプログラムは、可変長命令セットアーキテクチャから選択された、これらの異なる長さの命令の組合せを含み得る。また、プロセッサは、多重レベルのキャッシュを備えた階層メモリ構成を有してもよく、たとえば、命令キャッシュと、データキャッシュと、システムメモリとを含み得る。命令キャッシュは、キャッシュラインと一緒に複数の命令を記憶し、アクセスするように構成され得る。16ビット命令および32ビット命令をサポートするプロセッサアーキテクチャでは、32ビット命令はアラインされずにキャッシュラインに記憶され得る。16ビットハーフワードアドレス指定を使用すると、その最初の16ビットハーフワードを奇数の16ビットハーフワードアドレスに記憶している32ビット命令は、アラインされていないとみなされる。たとえば、256ビットのキャッシュラインは、8個の32ビット命令、または16個の16ビット命令、または16ビット命令と32ビット命令の両方の組合せを記憶することができる。16ビット命令と32ビット命令の混合を有するキャッシュラインは、2つのキャッシュラインの間を横断する最後の32ビット命令を有することができる。

10

【0004】

20

また、メモリを4キロバイト(4kバイト)ページなどのページに区分する仮想メモリシステムが使用され得る。そのようなシステムでは、2つのキャッシュラインの間を横断するキャッシュライン中の最後の32ビット命令は、ページ境界も横断し得る。各ページに異なる属性を割り当てることができ、属性としては、たとえば、ページ上に記憶された情報がキャッシュ可能であるか、キャッシュ可能でないかを挙げることができる。したがって、異なる長さの混合命令形式を有するキャッシュラインでは、キャッシュラインにわたっておよびページ境界にわたって分割された命令には、ページ属性の競合が生じやすくなることがある。たとえば、キャッシュライン中の最後の命令を除くすべての命令は、キャッシュ可能である属性を有する第1の例示的なページからとすることができるが、キャッシュラインおよびページ境界にわたって分割された最後の命令は、第1の部分はキャッシュ可能であるが、第2の部分はキャッシュ可能ではないことを示す属性を有することがある。最後の命令を分割する境界で特定されたキャッシュライン中の大部分の命令の性能に影響を及ぼすことなく、そのような競合を解消するのは困難であり得る。

30

【発明の概要】

【課題を解決するための手段】

【0005】

本発明の実施形態は、そのいくつかの態様のなかで、最後の命令を分割するページ境界で特定されたキャッシュライン中のキャッシュ可能命令を記憶することによって、性能が改善され得ることを認識する。本発明の一実施形態は、異なるキャッシュ可能性を用いてページ横断命令を管理するための方法が必要とされていることを認識する。キャッシュ不可能命令の第1のページからフェッチされ、キャッシュ可能命令を有するキャッシュライン中のキャッシュ可能命令の第2のページからフェッチされた命令の開始部分によって確立された命令の終了部分についての指示が設定され、命令はキャッシュライン境界を横断する。確立されたキャッシュラインにヒットすると、フェッチパイプライン中で指示が検出されて、命令を命令キャッシュから実行することができないことを示すためのキャッシュ不可能フラグを設定し、命令は受信されるが、キャッシュ不可能フラグに基づいてキャッシュから実行されない。命令の少なくとも終了部分は、命令の開始部分と組み合わせるために、キャッシュ不可能フラグに応答してキャッシュをバイパスしてメモリから再フェッチされ、命令は実行のために再構築される。

40

【0006】

50

別の実施形態は、異なるキャッシュ可能性を用いてページ横断命令の実行を制御するための装置を対象とする。命令キャッシュは、キャッシュ可能命令および、キャッシュ可能である開始部分と、キャッシュ不可能であり、キャッシュラインの最後でキャッシュライン境界を横断する終了部分とを有する命令を記憶するように構成される。インジケータ回路は、命令の実行許可が拒否されたという指示を1つまたは複数のビットに記憶するように構成され、命令はキャッシュ不可能命令として特定される。フェッチパイプラインはプロセッサに結合され、キャッシュ不可能命令を含む命令のフェッチグループにヒットしたときに指示を検出するように構成され、キャッシュ不可能命令は受信されるが、指示に応答してキャッシュから実行されない。

**【0007】**

10

別の実施形態は、異なるキャッシュ可能性を用いてページ横断命令を管理するための装置を対象とする。命令キャッシュは、キャッシュ可能命令および、キャッシュ可能である開始部分と、キャッシュ不可能であり、キャッシュラインの最後でページ境界およびキャッシュライン境界を横断する終了部分とを有する命令を記憶するように構成される。インジケータ回路は、命令の実行許可が拒否されたという指示を記憶するように構成され、命令はキャッシュ不可能命令として特定される。フェッチパイプラインはプロセッサに結合され、キャッシュ不可能命令を含む命令のフェッチグループにヒットしたときに指示を検出するように構成され、キャッシュ不可能命令は、指示に応答してキャッシュをバイパスして、実行のためにシステムメモリから再フェッチされる。

**【0008】**

20

別の実施形態は、システムを動作させるためのコンピュータ可読プログラムデータおよびコードによって符号化された非一時的コンピュータ可読記録媒体を対象とする。キャッシュ不可能命令の第1のページからフェッチされ、キャッシュ可能命令を有するキャッシュライン中のキャッシュ可能命令の第2のページからフェッチされた命令の開始部分によって確立された命令の終了部分についての指示が設定され、命令はキャッシュライン境界を横断する。確立されたキャッシュラインにヒットすると、フェッチパイプライン中で指示が検出されて、命令を命令キャッシュから実行することができないことを示すためのキャッシュ不可能フラグを設定し、命令は受信されるが、キャッシュ不可能フラグに基づいてキャッシュから実行されない。命令の少なくとも終了部分は、命令の開始部分と組み合わせるために、キャッシュ不可能フラグに応答してキャッシュをバイパスしてメモリから再フェッチされ、命令は実行のために再構築される。

30

**【0009】**

別の実施形態は、ページ横断命令を管理するための装置を対象とする。キャッシュ可能命令および、キャッシュ可能命令の第1のページからフェッチされた開始部分と、キャッシュ不可能命令の第2のページからフェッチされ、キャッシュラインの最後でキャッシュライン境界を横断する終了部分とを有する命令を記憶するための手段が利用される。命令の実行許可が拒否されたことを示すための手段が提供され、命令はキャッシュ不可能命令として特定される。キャッシュ不可能命令を含む命令のフェッチグループにヒットしたときに指示を検出するための手段も提供され、キャッシュ不可能命令は、指示に応答してキャッシュをバイパスして、実行のためにシステムメモリから再フェッチされる。

40

**【0010】**

さらなる実施形態は、異なるキャッシュ可能性を用いてページ横断命令の実行を制御するための装置を対象とする。命令キャッシュは、キャッシュ可能命令および、キャッシュ不可能である開始部分と、キャッシュ可能であり、キャッシュラインの最初でキャッシュライン境界を横断する終了部分とを有する命令を記憶するように構成される。インジケータ回路は、命令の実行許可が拒否されたという指示を1つまたは複数のビットに記憶するように構成され、命令はキャッシュ不可能命令として特定される。フェッチパイプラインはプロセッサに結合され、キャッシュ不可能命令を含む命令のフェッチグループにヒットしたときに指示を検出するように構成可能であり、キャッシュ不可能命令は受信されるが、指示に応答してキャッシュから実行されない。

50

## 【0011】

本発明のさらに完全な理解、ならびに本発明のさらなる特徴および利点は、以下の詳細な説明および添付の図面から明らかとなろう。

## 【図面の簡単な説明】

## 【0012】

【図1】 キャッシュライン境界およびページングされたメモリ境界を横断する命令をサポートする命令キャッシュを有するプロセッサ複合体を含むデバイスの特定の実施形態のブロック図である。

【図2】 本発明の一実施形態による、プロセッサ、レベル1命令キャッシュ(L1 Iキャッシュ)、L1データキャッシュ(Dキャッシュ)、レベル2キャッシュ(L2キャッシュ)、およびシステムメモリを有するプロセッサ複合体を示す図である。

10

【図3A】 16ビットおよび32ビットの可変長命令を含む例示的なプログラムセグメントを示す図である。

【図3B】 図3Aのプログラムセグメント300からの命令を含む例示的なL1 Iキャッシュラインを示す図である。

【図4A】 本発明の一実施形態による、命令変換ルックアサイドバッファ(ITLB: instruction translation look aside buffer)および物理メモリを有するページングされた仮想メモリシステムを示す図である。

【図4B】 本発明の一実施形態による、L1 Iキャッシュタグ内にライン横断インジケータを有する仮想-物理アドレス変換サブシステムを示す図である。

20

【図5】 本発明の一実施形態による、ライン横断命令および補助ライン横断インジケータを有する例示的な2ウェイセットアソシアティブ(two way set associative)Iキャッシュ回路を示す図である。

【図6】 本発明の一実施形態による、異なるキャッシュ可能性を用いてページ横断命令を管理するためのプロセスを示す図である。

## 【発明を実施するための形態】

## 【0013】

次に、本発明のいくつかの実施形態が示されている添付の図面を参照しながら、本発明についてより十分に説明する。しかしながら、本発明は、様々な形態で具現化されてもよく、本明細書に記載された実施形態に限定されるものとして解釈されるべきではない。むしろ、これらの実施形態は、本開示が徹底的で完全なものとなり、本発明の範囲を当業者に十分に伝えることになるように提供されている。

30

## 【0014】

本発明の教示による動作を実施する際にまたは実施するために動作させるためのコンピュータプログラムコード、すなわち「プログラムコード」は、C、C++、JAVA(登録商標)、Smalltalk、JavaScript(登録商標)、Visual Basic(登録商標)、TSQL、Perlなどの高級プログラミング言語、または様々な他のプログラミング言語で書くことができる。また、ターゲットプロセッサアーキテクチャのためのプログラムは、ネイティブアセンブラ言語で直接書くことができる。ネイティブアセンブラプログラムは、マシンレベルのバイナリ命令の命令ニーモニック表現を使用する。本明細書で使用するプログラムコードまたは非一時的コンピュータ可読記録媒体は、プロセッサがフォーマットを理解できる、オブジェクトコードなどのマシン言語コードを指す。

40

## 【0015】

図1は、キャッシュライン境界およびページングされたメモリ境界を横断する命令をサポートする命令キャッシュを有するプロセッサ複合体110を含むデバイス100(たとえば、通信デバイス)の特定の実施形態のブロック図である。デバイス100はワイヤレス電子デバイスであってもよく、コンピュータ実行可能命令118を有するシステムメモリ112に結合されたプロセッサ複合体110を含み得る。システムメモリ112は、図2のシステムメモリ227または図4Bのシステムメモリ452を含み得る。プロセッサ複合体110は、プロセッサ111、レベル1命令キャッシュ(L1 Iキャッシュ)122を有する統合メモリサブシステム114、外部タ

50

グ(xTag)回路126、およびキャッシュコントローラ回路128を含み得る。統合メモリサブシステム114は、キャッシュ不可能として指定および特定され得るプログラムメモリ内に1つまたは複数のページを有するページングされたメモリ編成をサポートする。プロセッサ111は、図2のプロセッサ210または図4Bのプロセッサパイプライン442を含み得る。統合メモリサブシステム114は、図2のL1データキャッシュ214およびL2命令/データキャッシュ226または図4BのL2キャッシュ450などの、L1データキャッシュおよびレベル2ユニファイドキャッシュ(図示せず)も含み得る。L1 Iキャッシュ122は、以下でより詳細に説明するように、図2のL1 Iキャッシュ218または図4BのL1 Iキャッシュ448を含み得る。xTag回路126は、図4BのxTag回路447およびxPビット449に関して以下でより詳細に説明するように、命令の実行を制御するオーバーライド指示を与えるための外部許可ビット(xPビット)130も含み得る。

10

#### 【0016】

統合メモリサブシステム114はプロセッサ複合体110に含まれ得るか、またはプロセッサ複合体110の外部にある1つまたは複数の別個のデバイスもしくは回路(図示せず)として実装され得る。例示的な例では、プロセッサ複合体110は、図2、図3B、図4A、図4B、および図5の回路およびシステムのうちのいずれかを含み、図3Aおよび図6に示されるまたは図3Aおよび図6に関連付けられる実施形態のうちのいずれか、またはそれらの任意の組合せに従って動作する。たとえば、図1に示すように、L1 Iキャッシュ122、xTag回路126、およびキャッシュコントローラ回路128はプロセッサ複合体110内でアクセス可能であり、プロセッサ111は、統合メモリサブシステム114のメモリまたはシステムメモリ112に記憶されたデータまたはプログラム命令にアクセスするように構成される。

20

#### 【0017】

カメラインターフェース134はプロセッサ複合体110に結合され、ビデオカメラ136などのカメラにも結合される。ディスプレイコントローラ140はプロセッサ複合体110およびディスプレイデバイス142に結合される。コーデック/デコーデック144もプロセッサ複合体110に結合され得る。スピーカ146およびマイクロフォン148はコーデック144に結合され得る。ワイヤレスアンテナ152およびワイヤレスインターフェース150を介して受信されたワイヤレスデータをプロセッサ111に与えることができるように、ワイヤレスインターフェース150はプロセッサ複合体110およびアンテナ152に結合され得る。

#### 【0018】

プロセッサ111は、システムメモリ112などの非一時的コンピュータ可読記録媒体に記憶されたコンピュータ実行可能命令118を実行するように構成され得、コンピュータ実行可能命令118は、プロセッサ111などのコンピュータに図3Aのプログラムセグメント300などのプログラムを実行させるように実行可能である。コンピュータ実行可能命令118は、プロセッサ111に、統合メモリサブシステム114のメモリおよびシステムメモリ112にアクセスする命令を処理させるようにさらに実行可能である。

30

#### 【0019】

特定の実施形態では、プロセッサ複合体110、ディスプレイコントローラ140、システムメモリ112、コーデック144、ワイヤレスインターフェース150、およびカメラインターフェース134は、システムインパッケージデバイスまたはシステムオンチップデバイス104に含まれる。特定の実施形態では、入力デバイス156および電源158はシステムオンチップデバイス104に結合される。さらに、特定の実施形態では、図1に示すように、ディスプレイデバイス142、入力デバイス156、スピーカ146、マイクロフォン148、ワイヤレスアンテナ152、ビデオカメラ136、および電源158は、システムオンチップデバイス104の外部にある。しかしながら、ディスプレイデバイス142、入力デバイス156、スピーカ146、マイクロフォン148、ワイヤレスアンテナ152、ビデオカメラ136、および電源158の各々は、インターフェースまたはコントローラなどの、システムオンチップデバイス104の構成要素に結合され得る。

40

#### 【0020】

本明細書で説明する実施形態によるデバイス100は、セットトップボックス、エンター

50

テインメントユニット、ナビゲーションデバイス、通信デバイス、携帯情報端末(PDA)、固定ロケーションデータユニット、モバイルロケーションデータユニット、携帯電話、セルラーフォン、コンピュータ、ポータブルコンピュータ、タブレット、モニタ、コンピュータモニタ、テレビジョン、チューナ、ラジオ、衛星ラジオ、音楽プレーヤ、デジタル音楽プレーヤ、ポータブル音楽プレーヤ、ビデオプレーヤ、デジタルビデオプレーヤ、デジタルビデオディスク(DVD)プレーヤ、ポータブルデジタルビデオプレーヤ、データまたはコンピュータ命令を記憶するまたは取り出す任意の他のデバイスなどの様々な電子デバイス、あるいはそれらの任意の組合せに組み込まれてもよい。

#### 【0021】

図2は、メモリ階層204からフェッチされた異なる長さの命令を実行するプロセッサ210を有するプロセッサ複合体200の説明図である。メモリ階層204は、レベル1(L1)データキャッシュ214、命令変換ルックアサイドバッファ(ITLB)217を含むメモリ管理ユニット(MMU)220、L1命令キャッシュ(Iキャッシュ)218、外部タグ(xTag)回路219、キャッシュコントローラ回路221、書込み制御回路222、レベル2命令/データキャッシュ(L2キャッシュ)226、およびシステムメモリ227を含む。xTag回路219は、xTag回路219が関連付けられるIキャッシュ218の外部にあり、Iキャッシュ218中のストレージアレイを変更することなく、xTag回路219の機能をプロセッサ複合体200に追加することを可能にする。プロセッサ複合体200は、プログラムコードを実行するための図1のデバイス100のハードウェア構成要素において適切に使用され得る。本発明の説明をわかりやすくするために、プロセッサ複合体に接続することができる周辺デバイスは図示されていない。特定用途向け集積回路(ASIC)技術、フィールドプログラマブルゲートアレイ(FPGA)技術、または他のプログラマブル論理、ディスクリートゲートもしくはトランジスタ論理、あるいは意図した用途に適した任意の他の利用可能な技術を使用して、プロセッサ複合体200の様々な構成要素を実装することができる。

#### 【0022】

プロセッサ210は、キャッシュから命令およびデータを階層的に取り出す。たとえば、プロセッサ210は、仮想フェッチアドレスおよび動作モードを含むフェッチ出力228を生成することによって、命令をフェッチする。動作モードは、32ビット命令のみのモード、16ビット命令のみのモード、16ビット命令と32ビット命令の混合モード、他の動作モードなどの特定を含み得る。そのようなプロセッサ動作モード状態インジケータは、プロセッサで動作中のプログラムによって制御される。

#### 【0023】

プロセッサの命令セットは、多倍長形式で符号化された命令を含み、従来、より長い命令は、可変長命令セットで利用可能な最も短い命令形式の長さの倍数である。命令は、命令の長さに応じて変わり得る複雑なエンコーディングを有することがあり、異なる長さの命令に隣接することがあるので、フェッチされた命令のアラインおよびデコーディングは、シングルプロセッサパイプラインステージが短い継続時間のクロック周期で行うことができる処理よりも多くの処理を必要とすることがある。命令の複雑さの程度により、プロセッサは、別個の命令アライメントパイプラインステージを含み、デコード動作をプリデコード動作およびデコードパイプラインステージに分割することができる。プリデコード動作は、L1 Iキャッシュミス処理の間にプリデコード動作を行うことによって、通常のパイプライン実行から適切に隠され得る。L1 Iキャッシュミス処理は、フェッチされた命令がL1 Iキャッシュ内で見つからず、メモリ階層のより高いレベルからフェッチされなければならないときに行われる。プリデコード動作は、フェッチされた命令とともにプリデコード情報をL1命令キャッシュに記憶する。そのようなプリデコード動作およびxTag回路219の動作は、書込み制御回路222によって制御される。

#### 【0024】

動作の際、プロセッサ210は、L1 Iキャッシュ218にアクセスするために使用される物理フェッチアドレスにITLB 217によって変換される仮想アドレスを生成して、アドレス指定された命令がL1 Iキャッシュに存在するかどうかを一致機構を使用して判断する。アドレ

ス指定された命令についての一致がL1 Iキャッシュ218に見つからなかった場合、ミスが生じる。ミス情報230はプリデコーダも含み得る書込み制御回路222に送信され、プロセッサ210はL2キャッシュ226に対してアクセス要求232を行う。L2キャッシュ226に命令ヒットがあると、所望の命令を含むL2キャッシュラインが第1のポート(ポートA)234上で書込み制御回路222に出力される。書込み制御回路222は、ミス処理の間、L2キャッシュからフェッチされた命令を部分的にデコードし、命令、命令に関連付けられたプリデコードビット、および実行許可ビットなどのタグ情報を出力238上でL1 Iキャッシュ218に与え、命令はプロセッサ210にも渡される。

#### 【0025】

プロセッサ210がデータをフェッチする必要があるとき、プロセッサ210はL1データキャッシュ214にアクセスして、アドレス指定されたデータが存在するかどうかを判断する。フェッチされたデータについての一致がL1データキャッシュ214に見つからなかった場合、ミスが生じ、L2キャッシュ226が次にアクセスされる。両方のL1キャッシュの場合において、命令またはデータがL1命令またはL1データキャッシュに存在することがわかった場合(キャッシュでヒットすると呼ばれる)、命令およびデータはそのそれぞれのL1キャッシュから出力240および244上で直接読み出される。L2キャッシュアクセスについてミスが生じた場合、命令およびデータはシステムメモリ227によって与えられる。

#### 【0026】

図3Aは、16ビットおよび32ビットの可変長命令を適切に含む例示的なプログラムセグメント300を示す。例示的なプログラムセグメント300は、プロセッサ複合体200のメモリ階層204に適切に記憶され得る。例示のために、プログラムセグメントは1つまたは複数のキャッシュラインから取り出されると仮定されるが、本発明の教示はプログラムセグメントを記憶する任意のメモリデバイスに適用可能であり、命令は記憶セグメント境界にまたがり得ることに留意されたい。キャッシュラインは固定長を有し得るので、プログラムセグメントはキャッシュラインの境界にまたがり得、したがって、キャッシュライン境界にわたって分割される命令を有し得る。

#### 【0027】

プログラムセグメント300は、16ビット命令および32ビット命令からなる可変長命令セットから生じる命令302を含む。たとえば、プロセッサ210は、複数のタイプの命令について16ビット命令形式および32ビット命令形式を使用することができ、命令タイプの使用を指定し制限するいくつかの動作モードをサポートすることができる。たとえば、プロセッサ210は、32ビット命令のみが使用され得ることを指定する第1の動作モードと、16ビット命令と32ビット命令の組合せが使用され得ることを指定する第2の動作モードとを有することができる。プロセッサは複数の動作モードを有し得るが、本発明の説明を明確にするために、例示的なプロセッサ210の説明は主に、上記で説明した第2の動作モードに限定される。

#### 【0028】

様々な実施形態を例示するために、いくつかの例示的な16ビット命令および32ビット命令を図3Aに示す。例示のために、プログラム相対バイトインジケータ304は、命令が始まるキャッシュラインにおけるバイトロケーションを表し、間接的に命令のサイズを示す。たとえば、ADD R5, R4, R3命令306は相対バイト位置00で開始し、バイト位置01で終了する。したがって、ADD R5, R4, R3命令306は16ビット命令である。同様に、ADD命令309も16ビット長である。ロード(LOAD)命令307、LOAD命令308、およびストア(STORE)命令310は32ビット長である。

#### 【0029】

キャッシュラインのサイズは、異なるプロセッサ実装形態において、たとえば、使用される製作技術に基づいてプロセッサおよびメモリ階層の設計において行われる選択に応じて異なり得る。L2キャッシュ226は512ビットキャッシュラインを使用することができ、L1 Iキャッシュ218は、たとえば、128ビットキャッシュラインまたは256ビットキャッシュラインなどの、より小さいキャッシュラインを使用することができる。示されたキャッシ

10

20

30

40

50

ュラインのサイズは例示的なものであり、より大きいまたはより小さいキャッシュラインのサイズは除外されない。例示の目的で、プログラムセグメント300は相対アドレス00で始まることが示されていることにも留意されたい。そのようなプログラムセグメント300は、キャッシュラインにおける様々なポイントで開始して配置されてもよく、複数のキャッシュラインにまたがってもよいことが諒解されよう。

#### 【0030】

図3Bは、図3Aのプログラムセグメント300からの命令を含む例示的なL1 Iキャッシュライン320を示す。例示的な第1のL1 Iキャッシュライン322および例示的な第2のL1 Iキャッシュライン326は、図2のL1 Iキャッシュ218における隣接するキャッシュラインである。第1のL1 Iキャッシュライン322は、16ビットフィールド330、333、334、および336と、16ビット拡張フィールド338とを含む。第1のL1 Iキャッシュライン322は、タグフィールド323および制御フラグCn 324に関連付けられ、制御フラグCn 324は、キャッシュ可能インジケータ(L)と、たとえば、ユーザ実行(Ux)ビットおよび特権実行(Px)ビットなどの実行許可ビットとを含み得る。UxビットおよびPxビットの非アサート状態は、いかなるモードでも実行しないことを示し得る。第2のL1 Iキャッシュライン326は、16ビットフィールド340、342、343、および344と、16ビット拡張フィールド346とを含む。第2のL1 Iキャッシュライン326は、タグフィールド327および制御フラグCn 328に関連付けられ、制御フラグCn 328は、キャッシュ可能インジケータ(L)と、第2のL1 Iキャッシュライン326に記憶された命令に関連付けられたユーザ実行(Ux)ビットおよび特権実行(Px)ビットなどの実行許可ビットとを含み得る。

#### 【0031】

図3Aのプログラムセグメント300の命令は、16ビットフィールド330に記憶された図3Aの16ビットADD R5、R4、R3命令306で始まる第1のL1 Iキャッシュライン322に配置され得る。32ビットLOAD命令307は、2つの16ビットフィールド333および334を含む32ビットフィールド332に記憶される。例示の目的で、16ビットフィールド333はLOAD命令307の上位16ビットを含み、隣接する16ビットフィールド334はLOAD命令307の下位16ビットを含む。

#### 【0032】

第1のL1 Iキャッシュライン322における次の命令は、2つの命令キャッシュラインにわたって記憶された32ビットLOAD命令308である。LOAD命令308の上位16ビットは、第1のL1 Iキャッシュライン322における16ビットフィールド336に記憶される。LOAD命令308の下位16ビットは、第2のL1 Iキャッシュライン326における16ビットフィールド340に記憶される。LOAD命令308の下位16ビットのコピーは、16ビット拡張フィールド338に記憶される。いずれも図3AのADD R8、R6、R7命令309およびSTORE命令310は、第1のL1 Iキャッシュライン322のセグメント330および332と同様に、第2のL1 Iキャッシュライン326における16ビットフィールド342～344に記憶される。本明細書において明確にするために図示されていないプリデコードビットは、キャッシュラインにおける各16ビットフィールドに関連付けられ得ることに留意されたい。

#### 【0033】

図3Bに示すように、16ビット命令および32ビット命令をサポートするプロセッサ複合体における命令キャッシュは、たとえば、1つのK/2形式の命令に加えて、N個のKビット形式でアラインされた命令を記憶することができるキャッシュラインを有して構築され得る。図3Bは例示的なものであり、Kビット命令は8ビットバイトアドレス境界および16ビットハーフワード境界上に記憶され得ることに留意されたい。また、16ビット命令および24ビット命令など、互いの倍数ではない命令形式を有する命令セットアーキテクチャも本発明の実施形態によってサポートされることにさらに留意されたい。

#### 【0034】

たとえば、K=32ビットの場合、ワードアラインされた境界上で8個の32ビット命令を記憶するキャッシュラインは、命令データに関連付けられた256ビット+16ビット=272ビットラインを有するキャッシュで実装される。可変長命令プロセッサの場合、16ビット命令と32ビット命令の混合を有するキャッシュラインは、16ビット拡張フィールド338を有する

第1のキャッシュライン322などの、余分のK/2ビットスペースを利用するキャッシュラインの最後の32ビットロケーションに記憶されるキャッシュライン横断32ビット命令を有し得る。最後の16ビットキャッシュ拡張フィールド338に記憶された32ビットキャッシュライン横断命令の下位16ビット部分は、ビットフィールド340における次の順次キャッシュラインに記憶された16ビット部分の複製である。別の例では、8ビットの倍数である命令を有するプロセッサはライン横断命令も有することができる。そのような命令の場合、ライン横断命令は、たとえば、第2のキャッシュラインに続く命令の1バイト部分、2バイト部分、または3バイト部分を有するバイト境界において分割され得る。第2のキャッシュラインに記憶された1バイト部分、2バイト部分、または3バイト部分はコピーされ、第1のキャッシュラインにおけるライン横断命令の第1の部分に関連付けられた位置に記憶される。この例示的なケースでは、キャッシュラインに対する3バイト拡張が提供される。たとえば、キャッシュ拡張フィールド338は、その現在図示されている16ビットではなく、3バイトビットフィールドに拡張される。他のバイト長命令が可能であり、本発明によって排除されない。キャッシュライン横断命令は、キャッシュ不可能ページへのページ境界も横断することができ、したがって、キャッシュ可能ではない場合があるので、ページ境界(ライン/ページ)横断キャッシュ不可能命令をキャッシュから実行することを防がなければならない。また、単一のライン横断命令の第1の部分を記憶し、第2の命令も記憶するなど、単一の命令の一部分よりも多くを記憶するように拡張フィールド338を拡張することができ、第2の命令は一般に、拡張された拡張フィールドを使用しているキャッシュラインを用いて記憶された次の論理ページに関連付けられる。

#### 【0035】

図4Aは、本発明の一実施形態による、命令変換ルックアサイドバッファ(ITLB)402および物理メモリ404を有するページングされた仮想メモリシステム400を示す。仮想アドレス405は一般に、2つの部分に符号化される。アドレスビットの上側のフィールドは通常、4Kバイトのページなど、選択されたページサイズに基づいて符号化される、仮想ページ番号406を表す。アドレスビットの下側のフィールドは、アドレス指定されたページ内のアドレスを特定する、ページオフセット407である。仮想アドレスから物理アドレスへの変換において、仮想ページ番号は物理ページ番号(Pページアドレス)に変換される。ページオフセットは、仮想アドレスと物理アドレスの両方に対して同じであり、変換されない。

#### 【0036】

仮想アドレスから物理アドレスへの変換システムは、変換プロセスの性能を向上させるために、レベル1とレベル2の命令キャッシュおよびデータキャッシュなどの様々なキャッシュに関連付けられた、1つまたは複数の変換ルックアサイドバッファ(TLB)を含み得る。命令TLB(ITLB)は、エントリ検証およびページがキャッシュ可能命令を含むかまたはキャッシュ不可能命令を含むかなどの記憶されたページの属性とともに、最近の仮想アドレスから物理アドレスへの変換を記憶する、小さいキャッシュである。ITLBは従来、ランダムアクセスメモリ(RAM)回路と結合されるコンテンツアドレス可能メモリ(CAM)回路を含み、比較的小さく、たとえば32個または64個のエントリを有する。各ITLBエントリは、RAM回路中の変換された物理ページ番号に関連付けられた、最近使用された仮想ページ番号を有する、CAM回路中のタグを含む。たとえば、ページングされた仮想メモリシステム400は、ITLB 402と、キャッシュ不可能ページ409などの1つまたは複数のキャッシュ不可能ページと混ざり合ったキャッシュ可能ページ408および410を有する物理メモリ404とを使用する。ITLB 402の各エントリは、有効(V)フラグ、読取り(R)フラグ、書込み(W)フラグを含むフラグ412と、キャッシュ可能インジケータ(L)414と、仮想アドレスタグ416と、関連付けられた物理ページアドレス418とを有する。Lフィールド416は、ページをキャッシュ可能またはキャッシュ不可能として特定するのに適した単一ビットであり得る。ページがキャッシュ可能であるかまたはキャッシュ不可能であるかは、コンパイル中に静的に判断され得、様々な要因に依存し得る。たとえば、メモリマップされた入力および出力(I/O)デバイスがシステムの実際の実装形態で使用される場合、そのようなメモリマップされたロケーションはキャッシュ不可能としてタグ付けされ得る。

## 【0037】

図4Bは、本発明の一実施形態による、仮想アドレスから物理アドレスへの変換サブシステム440を示す。変換サブシステム440は、プロセッサパイプライン442、ITLB 444、物理アドレスバッファ446、xTag回路447、L1 Iキャッシュ448、L2キャッシュ回路450、システムメモリ452、および書込み制御回路454からなる。ITLB 444は、ITLBタグ内にキャッシュ可能インジケータ(L)458を備えるエントリ456を有する。L1 Iキャッシュ448は、各ラインに関連付けられたタグフィールド470、キャッシュされた命令を記憶するための基本エクステンツ471、および各ラインに関連付けられた拡張フィールド472を含む。また、L1 Iキャッシュ448中の例示的な第1のライン457は、第1のライン457の第1の基本エクステンツ473に記憶されたキャッシュ可能命令のセットを含む。図4Aのページ408などの命令のキャッシュ可能ページから、図4Aのページ409などのキャッシュ不可能としてタグ付けされた命令のページにページ境界を横断する32ビット命令の場合、命令のキャッシュ不可能な部分は32ビット命令がキャッシュ不可能命令として扱われることを要求する。この32ビット命令の例では、第1の16ビット部分(Ia)474は、キャッシュ可能ページからフェッチされたキャッシュ可能な部分であるが、第2の16ビット部分(Ib)475は、キャッシュ不可能ページからフェッチされたキャッシュ不可能な部分である。第1の部分Ia 474は第1のライン457の第1の基本エクステンツ473に記憶され、第2の部分Ib 475は第1のライン457に関連付けられた拡張フィールドに記憶される。第1のライン457は、第1の基本エクステンツ473に記憶されたキャッシュ可能命令に関連付けられた1つまたは複数の実行許可ビットを含むタグフィールド470から選択された関連タグも有する。L1 Iキャッシュ448は、Ib(Ib')476の少なくとも第2の部分のコピーのための記憶スペースを含む例示的な第2のライン459も含む。第1の部分Ia 474および第2の部分Ib 475がキャッシュ可能命令を表す場合、第2の部分のコピーIb' 476は第2のライン459に記憶されることになる。キャッシュ可能なライン横断命令を有し、実施中に行われる決定に依存する、そのような場合、第2の部分Ib 475および第2の部分のコピーIb' 476は、両方の部分の内容が同じであるので、位置を切り替えることができる。しかしながら、第1の部分Ia 474および第2の部分Ib 475がキャッシュ不可能命令の部分である例示的な場合には、第2の部分のコピーIb' 476は第2のライン459に記憶されない。ライン/ページ境界にわたって分割されるキャッシュ不可能命令を有するキャッシュラインの例示的なシナリオに対処するために、キャッシュ不可能インジケータフラグは第1のライン457中のこの命令に関連付けられる。以下でより詳細に説明するように、キャッシュ不可能インジケータフラグは、L1 Iキャッシュ448の外部にあるxTag回路447に記憶される。

## 【0038】

変換プロセスは、ITLB 444内で仮想アドレス405から選択された仮想ページ番号406をCAM回路に適用することによって開始する。ITLB 444は、適用された仮想ページ番号406を、一般にCAMタグ460中のエントリタグとともに記憶された最近使用された仮想ページ番号のすべてと並列比較する。一致がある場合、CAM回路は、ITLB 444中のRAM回路における対応するエントリ456にアクセスし、対応するエントリ456は、変換された物理ページアドレス462として出力され、物理アドレスバッファ446に記憶される。変換された物理ページアドレス463は、仮想アドレス405からのページオフセット464と連結された、変換された物理ページアドレス462を含む。

## 【0039】

たとえば、4ギガバイト(4GB)および4Kバイトのページの仮想アドレス空間を有する組込みシステムでは、仮想アドレス405は、ビット[31:12]を有する仮想ページ番号406およびビット[11:0]を有するページオフセット407からなる。同じ組込みシステムにおいて、キャッシュおよびメインメモリのメモリ階層は、512kバイトの物理メモリ空間および4kバイトのページを包含し得る。ITLB 444においてヒットがあると、仮想アドレス405は物理アドレス463に変換される。物理アドレス463は、ビット[28:12]を有する物理ページ番号462およびビット[11:0]を有するページオフセット464からなり、物理ページ番号462のビット[18:12]は512kバイト実施に必要とされる。ITLB 444においてヒットがあると、キャッシ

ュ可能インジケータ(L)458を含むタグも出力され、物理アドレスバッファ446に記憶される。キャッシュ可能インジケータ(L)458およびタグ465の配置は例示的なものである。次いで、物理アドレス463はL1 Iキャッシュ448に適用される。変換サブシステム440の説明を続ける前に、キャッシュライン/ページ横断命令でもあるキャッシュ不可能命令をキャッシュに記憶するという課題について次に説明する。

#### 【0040】

キャッシュ不可能命令の存在を判断したために、キャッシュ可能命令とキャッシュラインを横断するキャッシュ不可能命令とを有するフェッチされたキャッシュラインをキャッシュに記憶することを一般的に除外するという課題に対処するために、余分のK/2ビットフィールドライン横断命令データに関連付けられた属性は、キャッシュライン中で記憶され、残りの命令の属性とは別々に追跡される制御属性で指定され得る。キャッシュライン横断命令でもあるキャッシュ不可能命令をキャッシュラインに記憶した、この例示的な場合における制御属性は、キャッシュ不可能命令をいかなるモードでも実行しないことを示すように設定される。制御属性は、ライン/ページ横断命令を有するキャッシュラインに関連付けられた少なくとも1つの記憶ビットに記憶される。ライン/ページ横断命令の部分がフェッチグループの一部としてキャッシュからフェッチされると、キャッシュ不可能フラグがxTag回路447でアサートされる。xTag回路447などのxTag回路は、ページ横断命令を含み得る各キャッシュラインに対して実装される。また、xTag回路447がプロセッサパイプライン442に転送されるフラグデータについてアクセスされ、この転送は一般に、フェッチされたキャッシュライン命令のそのセットがライン横断命令を含むときのみに生じ得る。また、フェッチグループ中のキャッシュ可能命令に関連付けられた許可ビットも取り出されることに留意されたい。制御属性を有するライン/ページ横断命令またはその部分は、ライン/ページ横断命令をいかなるモードでも実行することを可能にしないように、そのライン/ページ横断命令だけについて、フェッチグループに関連付けられた許可ビットをオーバーライドし得る。そのような動作は、xTag回路447中のキャッシュ不可能フラグによって制御され得る。動作は、このライン/ページ横断命令だけにxTag外部許可ビット(xPビット)449を与えることによって制御され得、xTag外部許可ビット(xPビット)449は、xTag回路447に記憶され、そのライン/ページ横断命令だけについてキャッシュライン許可ビットをオーバーライドする。関連付けられたタグフィールド470からアクセスされたキャッシュ可能命令についての許可ビット、第2の部分Ib 475などの、拡張フィールド472からのライン/ページ横断命令またはその部分、および、たとえば、xTag回路447からライン/ページ横断命令についてxTag 480上でアクセスされるxPビット449は、プロセッサパイプライン442に転送される。

#### 【0041】

プロセッサパイプライン442は、検出(Dt)回路482、第1のデコード(Dc)回路483、バッファ/ホールド(B&H)回路484、再フェッチ回路485、再結合回路486、およびマルチプレクサ487を含む。Dt回路482は、一般に、実行許可が許容されるかどうかをチェックし、実行しないものとして第2の部分Ib 475にタグ付けするパイプラインステージにおいて、第2の部分Ib 475およびアクセスされたxTag 480が受信されたことを検出する。Dc回路483は、第2の部分Ib 475がページ横断命令の一部であるかどうかを特定する。第2の部分Ib 475が、実行許可がチェックされるステージで受信されたにもかかわらず、それだけでは、一般に第2の部分Ib 475がページ横断命令であることを意味しないことに留意されたい。したがって、Dc回路483はデータをデコードし、この例示的な場合では、第2の部分Ib 475がページ横断命令の一部であると判断する。

#### 【0042】

プロセッサパイプライン442の動作は、キャッシュラインから受信した命令をバッファするB&H回路484を用いて続き、第2の部分Ib 475がフェッチグループ中で最も古い命令を表すかどうかを判断する。第2の部分Ib 475はフェッチグループ中で最も古い命令を表していないとB&H回路484が判断した場合、B&H回路484は第2の部分Ib 475をバッファし、最も古い命令を表していると判断されるまで、第2の部分Ib 475をホールドする。プロセッ

サパイプライン442において第2の部分Ib 475が最も古い命令を表していると判断されたとき、第2の部分Ib 475の上にあるプロセッサパイプラインのフラッシュが実行される。キャッシュ不可能命令は、許可失敗問題の解決に関連付けられた既存のデータフローを再利用するシステムメモリ452から再フェッチされる。この特定の実施形態では、第2の部分Ib 475もフラッシュされ得るか、または上書きされることが許容され得る。

#### 【0043】

代替実施形態では、キャッシュライン中の有効な(good)キャッシュ可能データのフラッシュは必要でなくともよく、再フェッチ回路485は、命令キャッシュをバイパスし、たとえば、マルチプレクサ477を介して第2の部分Ib 475をシステムメモリ452から直接取得して、キャッシュ不可能属性を有する第2の部分Ib 475を再フェッチする。再結合回路486は第1の部分Ia 474をシステムメモリ452から受信された第2の部分Ib 475と組み合わせて、完全な命令Ia||Ibを形成し、デコードされるように命令をマルチプレクサ487に通し、組み合わせられた命令を、命令キャッシュからフェッチされることなしに実行することを可能にするパイプライン処理を継続する。実行のための適切なパイプラインプロトコルに従って、組み合わせられた命令での任意の必要なプリデコード動作およびデコード動作を繰り返す必要があり得ることに留意されたい。Dt回路482はフェッチパイプラインステージに関連付けられ、Dc回路483は一般デコードパイプラインステージに関連付けられ、B&H回路484は命令キューに関連付けられ得ることに留意されたい。例示的な回路482~487は、特定の実装形態に従って適切なパイプラインステージに配置され得る。

#### 【0044】

ライン/ページ横断命令に先行する命令は、命令のフローをライン/ページ横断命令からそらす場合があるので、プロセッサパイプラインは、ライン/ページ横断命令に到達したかどうかの判断を行うことができるまで、ライン/ページ横断命令およびライン/ページ横断命令に続く命令の発行をストールする。分岐命令の実行などにより、ライン/ページ横断命令に到達していない場合、標準的分岐動作が続く。一実施形態では、ライン/ページ横断命令に到達した場合、ライン/ページ横断命令およびライン/ページ横断命令に続く命令がフラッシュされ、L1 Iキャッシュ218をバイパスして、少なくともキャッシュ不可能として特定されたライン/ページ横断命令について、キャッシュ不可能要求235がシステムメモリ227に対して行われる。キャッシュ不可能命令は、たとえば、図2のシステムメモリ出力バス236上で返される。代替として、別の実施形態では、前のキャッシュラインで複製されたライン/ページ横断命令の少なくともその部分が再フェッチされ、ライン/ページ横断命令全体が前のフェッチグループに保存された第1の部分から再構築される。両方の実施形態では、ライン/ページ横断命令またはその部分は、キャッシュされていないフェッチされた命令についての適切な属性を用いてシステムメモリから返され、再構築された命令は、キャッシュされることなしに実行され得る。

#### 【0045】

別の実施形態では、固定長命令セットアーキテクチャは、たとえば、固定長命令とともに記憶された可変データ幅のデータを用いるフォンノイマン(Von Neumann)アーキテクチャの使用により、アラインされていない命令を有し得る。固定長命令と混合した幅のデータの組合せは、キャッシュラインを横断し、キャッシュ可能ページとキャッシュ不可能ページとの間のページ境界も横断する任意のアラインされていない命令についての同じ問題および解決策につながり得る。したがって、キャッシュ可能ではない単一のライン/ページ横断命令を有するキャッシュライン中の命令の大部分を実行するためのプロセッサ性能は、そのようなライン/ページ横断命令を有さない任意のキャッシュラインからフェッチされた命令の実行と同じままである。この結果は、キャッシュ不可能データを部分的に含むので本来ならキャッシュから除外されていたはずのラインがキャッシュされるのを可能にすることによって達成される。

#### 【0046】

図4Bに戻ると、実行許可ビットは、L1 Iキャッシュ448中の各ラインに関連付けられたタグに記憶され、基本エクステンツ471に記憶された各キャッシュ可能命令について有効

である。キャッシュ不可能インジケータフラグは、第1のライン457に関連付けられた拡張フィールドに記憶された第2の部分Ib 475に関連付けられた追加の許可ビットに記憶され得る。追加の許可ビットは、L1 Iキャッシュ448の外部にあるxTag回路447に記憶され、たとえば、いかなる理由でも実行しないことを示す。キャッシュラインは、固定数の16ビット命令または固定数の32ビット命令を有するように選ばれる。たとえば、16ビットの拡張フィールドを加えた512ビットのキャッシュラインの場合、4kバイトのページは、0から63の番号が付けられ得る64個のキャッシュラインに対応する。16ビットの拡張フィールドは別個のアレイに記憶され得る。セット63におけるラインのみがページ横断命令を有し得るので、追加の許可ビットを使用すべきかどうかを判断するために、フェッチアドレスがページアドレスの最後と比較される。フェッチアドレスはまた、アドレス指定された命令がキャッシュラインにわたって分割されるかどうかを判断して、その命令をライン横断命令として特定するために比較される。

#### 【0047】

許可ビットは一般に、アクセスされるラインに関連付けられるタグフィールドに書き込まれる。たとえば、Iキャッシュタグ470中の有効フラグによって示されるように、アドレス指定されたキャッシュラインが有効でないとき、フェッチはL2キャッシュ226またはシステムメモリ227に向けられる。一般に、フェッチ要求はL1 Iキャッシュ218にロードするための複数の命令を推測で(speculatively)返し、要求された命令はL1 Iキャッシュ218からプロセッサ210に返される。フェッチ要求に関連付けられるのは、図4AのLビット414などの、要求されたアドレスのキャッシュ可能性属性である。次いで、Lビット414は一般に、フェッチグループに関連付けられ、他のフラグとともに、キャッシュライン中のフェッチグループに関連付けられたタグにロードされる許可ビットに分解される。フェッチグループがキャッシュ可能ではないことをキャッシュ可能性Lビット414が示す場合、フェッチグループはL1 Iキャッシュ448にロードされない。しかしながら、フェッチグループの一部であるページ横断/ライン横断命令の場合、命令の第1の16ビット部分はキャッシュ可能ページからアクセスされ得るが、第2の16ビット部分はキャッシュ不可能ページからアクセスされ得る。一般に、第1のレベル命令キャッシュにミスがあると2つのラインがフェッチされるので、第2のラインのキャッシュ可能性も判断され得る。したがって、ページ/ライン横断命令の第2の16ビット部分は、フェッチグループを記憶する残りのキャッシュラインに関連付けられた拡張フィールドにロードされ得、余分の許可ビット(xPビット)449はxTag回路447に記憶され得る。代替実施形態では、拡張フィールドは、追加の16ビット命令または32ビット命令を記憶することに対応して、16ビットよりも多くを、たとえば32ビットまたは48ビットを記憶するように拡張され得ることに留意されたい。また、プリデコードビットがキャッシュライン中の各命令に関連付けられる場合、拡張フィールドは、記憶される16ビット部分ごとに2または4プリデコードビットを含むように拡張され得る。

#### 【0048】

余分のデータ許可情報は、プリデコードビットのエンコーディングから特定され得る。一般に、余分のデータ許可情報は、ページ横断命令に関連付けられ得る任意の記憶フィールドに記憶され得る。プリデコードビットを使用して、命令を32ビットまたは16ビットとして特定する一実装形態では、ページ横断命令も、サイズおよびアドレスの計算に基づいて特定する代わりに、1つまたは複数の余分のプリデコードビットで特定され得る。また、「いかなる理由でも実行しない」という指示は、プリデコードビットに記憶されて、命令キャッシュに記憶されたキャッシュ不可能データの場合、ページ横断命令を誤った命令として特定し得る。

#### 【0049】

図5は、本発明の一実施形態による、ライン横断命令および補助ライン横断インジケータを有する例示的な2ウェイセットアソシアティブIキャッシュ回路500を示す。本発明はダイレクトマップキャッシュ、フルアソシアティブキャッシュまで含む4ウェイキャッシュ、8ウェイキャッシュなどの他のキャッシュ設計に適用可能であるが、2ウェイセットアソシアティブ命令キャッシュが例示的な命令キャッシュ回路500として示されている。第1

10

20

30

40

50

のウェイ514は、第1のウェイ中の各ラインについての許可ビットとキャッシュラインアドレスタグとを含む第1のウェイタグビットフィールド518を含む。第1のウェイ514は、たとえば、「n」個の命令Ic0～Icnおよび第1の部分Ixaとともに示されるデータのライン519と、第2の部分Ixbを記憶するものとして示される拡張フィールド520とをさらに含む。命令キャッシュ回路500はまた、第2のウェイ中の各ラインについての許可ビットとキャッシュラインアドレスタグとを含む第2のウェイタグビットフィールド522を有する第2のウェイ516のための記憶装置を含む。第2のウェイ516は、「z」個の命令Ib0～Ibzとともに示されるデータのライン523と、占有されていないものとして示される拡張フィールド524とをさらに含む。一般に、各ウェイ中のラインの記憶容量は同じであるが、異なる長さの異なる数の命令を記憶することが可能である。2ウェイIキャッシュの場合、4kバイトのページについてセット63に2つのラインがあり、各ラインはIxaおよびIxbからなる命令などのライン/ページ横断命令を有し得る。また、これらの2つのラインの各々について、余分の許可ビットがxTag回路532および533に別々に記憶され、別々に追跡される。4kバイトのページを有する4ウェイセットアソシアティブキャッシュでは、セット63に4つのxTag回路を有する4つのラインがある。フルアソシアティブキャッシュでは、すべてのキャッシュラインがライン/ページ横断命令を有する可能性があり、次いで、余分の許可ビットは、キャッシュ中の各ライン中のページ横断命令を追跡するために、タグビットフィールド518および522に含まれ得る。

10

#### 【0050】

命令キャッシュが実際に実装される方法は、特定の用途およびシステム全体に課された設計制約に依存する。当業者は、様々な設計の互換性、および各特定の用途について本明細書で説明した機能を実装する最良の方法を認識されよう。たとえば、拡張フィールド520および524は、それに対応するラインアレイに直接関連付けられて示されているが、拡張フィールド520および524はラインアレイとは別のアレイで実装され得る。

20

#### 【0051】

図6は、本発明の一実施形態による、異なるキャッシュ可能性を用いてページ横断命令を管理するためのプロセス600を示す。ブロック602において、キャッシュラインは、キャッシュライン横断命令用の拡張記憶装置および拡張記憶装置用の属性フラグを用いて確立される。属性フラグはキャッシュの外部に記憶されてもよい。ブロック604において、たとえば、拡張記憶装置からのページ横断命令の第2の部分を含む、命令のフェッチグループにおいて、ページ横断命令がフェッチされる。また、ブロック604において、ページ横断命令の第2の部分を用いて属性フラグをプロセッサパイプライン中で追跡するために、属性フラグがxTag回路からキャプチャされる。ブロック606において、ページ横断命令の第2の部分は、プロセッサパイプラインで受信されたこと、ページ横断命令が命令キャッシュから生じたこと、およびページ横断命令がいかなるモードでも実行可能でないものとしてタグ付けされることが検出される。ブロック608において、ページ横断命令は、プロセッサパイプラインに対してページ横断命令を特定し、プロセッサパイプラインの内部で、ページ横断命令を実行しないものとしてタグ付けするようにデコードされる。

30

#### 【0052】

決定ブロック610において、ページ横断命令がプロセッサパイプライン中で最も古い命令であるかどうかの判断が行われる。ページ横断命令がプロセッサパイプライン中で最も古い命令ではない場合、プロセス600はブロック612に進む。ブロック612において、ページ横断命令は、ページ横断命令がプロセッサパイプライン中で最も古い命令になるまでホールドされ、次いでブロック614に進む。決定ブロック610に戻り、ページ横断命令がプロセッサパイプライン中で最も古い命令である場合、プロセス600はブロック614に進む。ブロック614において、許可失敗問題の解決に関連付けられた既存のデータフローを利用する一実施形態では、ページ横断命令を含むプロセッサパイプラインは背後でフラッシュされる。代替実施形態では、フラッシュは実行されず、ページ横断命令またはキャッシュ不可能属性を有するページ横断命令の第2の部分のみがシステムメモリから直接フェッチされる。ブロック616において、ページ横断命令が再フェッチされるか、またはページ横断

40

50

命令の少なくとも第2の部分が命令キャッシュをバイパスしてシステムメモリから再フェッチされる。第2の部分が再フェッチされる場合、ページ横断命令の第1のキャッシュ可能部分はキャッシュ不可能命令を再構築するための動作用に予約される。ブロック618において、ページ横断命令は、必要であれば、キャッシュ可能な第1の部分を、システムメモリから再フェッチされキャッシュ不可能として実行された第2の部分と組み合わせることによって再構築される。

#### 【0053】

本発明は図示された命令フロー論理200に限定されず、プリデコード情報を命令キャッシュに記憶することもできる、可変長命令を有する任意のパイプラインプロセッサにさらに適用可能である。拡張が固有の命令セット使用モードをサポートする場合、可変長プロセッサ命令セットへの拡張は本発明によって対応され得る。たとえば、16ビット命令、32ビット命令、および64ビット命令が動作可能である場合、32ビット命令および64ビット命令が2つのL1 Iキャッシュラインにわたってまたがることができるように動作モードが指定され得る。64ビット命令タイプを使用するプロセッサは、上記で説明した例示的なプロセッサ204の拡張であり得る。拡張されたプロセッサは、たとえば、32ビット命令のみに制限された第1の状態、16ビット命令と32ビット命令の両方についての第2の状態、16ビット命令、32ビット命令、および64ビット命令についての第3の状態、ならびに64ビット命令のみに制限された第4の状態について符号化された動作モード状態を有し得る。Iキャッシュライン中の64ビット命令は、4つの16ビットフィールドに区分され得る。48ビットを有する拡張ビットフィールドは、64ビット命令がライン/ページ横断状況において4つの16ビット部分にわたって分割されることを可能にするように使用され得る。

#### 【0054】

本発明はまた、2の累乗である命令長に限定されない。たとえば、16ビット命令および24ビット命令を有する代替アーキテクチャを考える。この例では、命令キャッシュラインは8ビット命令セクションに区分され得る。24ビット命令は、たとえば、3つの8ビットセクションからなることができる。16ビット命令を記憶する192ビット基本エクステンツキャッシュラインは、12個の16ビット命令および8個の24ビット命令をホールドすることができる。16ビット拡張フィールドは、24ビット命令が3つの8ビット部分に分割されることを可能にする。この例示的なキャッシュのキャッシュラインは、 $192+16=208$ ビットである。

#### 【0055】

一実施形態はまた、図4Bの拡張フィールド472などの、キャッシュラインの最初において構造化された拡張データ記憶部分を用いて構成され得る代替キャッシュを対象とする。キャッシュラインの最初に拡張フィールドがある場合、一実施形態は、キャッシュ不可能である第1の半分およびキャッシュ可能である第2の半分の命令と、キャッシュ可能であるデータを有する残りのキャッシュラインとを対象とする。この代替キャッシュにおいて第1の命令のキャッシュ不可能部分を処理するための手順は、本明細書で説明した、図4Bに示すキャッシュの最後の命令のキャッシュ不可能部分を処理するための手順と同様に動作する。代替キャッシュラインにおける最後の命令へのキャッシュ可能な第2の命令のフェッチは、図4Bに示すキャッシュにおけるN-1命令へのキャッシュ可能な第1の命令のフェッチと同様に動作する。

#### 【0056】

本明細書で開示される実施形態とともに説明される様々な例示的な論理ブロック、構成、モジュール、回路、およびアルゴリズムのステップは、電子ハードウェア、プロセッサによって実行されるコンピュータソフトウェア、またはこれら両方の組合せとして実装され得ることが、当業者にはさらに諒解されよう。様々な例示的な構成要素、ブロック、構成、モジュール、回路、およびステップが、概してそれらの機能に関して、上記で説明されてきた。そのような機能がハードウェアとして実装されるか、プロセッサ実行可能命令として実装されるかは、特定の用途およびシステム全体に課された設計制約に依存する。当業者は、説明した機能を各特定の用途ごとに様々な方法で実装し得るが、そのような実

装の決定は、本開示の範囲からの逸脱を生じるものと解釈すべきではない。

#### 【0057】

本明細書で開示する実施形態に関して説明する方法は、ハードウェアで具現化され、プロセッサによって実行される非一時的信号を記憶するメモリモジュールからソフトウェアによって使用され得る。ソフトウェアは、本明細書で説明したハードウェアの実行をサポートすることができるか、または異なるキャッシュ可能性を用いてページ横断命令を管理するための方法および装置をエミュレートするために使用され得る。ソフトウェアモジュールは、ランダムアクセスメモリ(RAM)、フラッシュメモリ、読取り専用メモリ(ROM)、電氣的プログラマブル読取り専用メモリ(EPROM)、ハードディスク、リムーバブルディスク、テープ、コンパクトディスク読取り専用メモリ(CD-ROM)、または、当技術分野で既知の任意の他の形態の非一時的記憶媒体中に常駐し得る。記憶媒体は、プロセッサがその記憶媒体から情報を読み取り、場合によってはその記憶媒体に情報を書き込むことができるようにプロセッサに結合され得る。プロセッサに結合している記憶媒体は回路実装と一体の直接結合であってもよく、あるいは、直接アクセス、またはダウンローディング技法を使用したデータストリーミングをサポートする1つまたは複数のインターフェースを利用してもよい。

10

#### 【0058】

本発明は現在好ましいコンテキストで開示されているが、本教示は本開示および以下の特許請求の範囲に一致する様々なコンテキストに適合され得ることが認識されよう。

#### 【符号の説明】

20

#### 【0059】

- 100 デバイス
- 104 システムオンチップデバイス
- 110 プロセッサ複合体
- 111 プロセッサ
- 112 システムメモリ
- 114 統合メモリサブシステム
- 118 コンピュータ実行可能命令
- 122 レベル1命令キャッシュ(L1 Iキャッシュ)
- 126 外部タグ(xTag)回路
- 128 キャッシュコントローラ回路
- 130 外部許可ビット(xPビット)
- 134 カメラインターフェース
- 136 ビデオカメラ
- 140 ディスプレイコントローラ
- 142 ディスプレイデバイス
- 144 コーダ/デコーダ(コーデック)
- 146 スピーカ
- 148 マイクロフォン
- 150 ワイヤレスインターフェース
- 152 ワイヤレスアンテナ
- 156 入力デバイス
- 158 電源
- 200 プロセッサ複合体
- 204 メモリ階層
- 210 プロセッサ
- 214 レベル1(L1)データキャッシュ
- 217 命令変換ルックアサイドバッファ(ITLB)
- 218 L1命令キャッシュ(Iキャッシュ)
- 219 外部タグ(xTag)回路

30

40

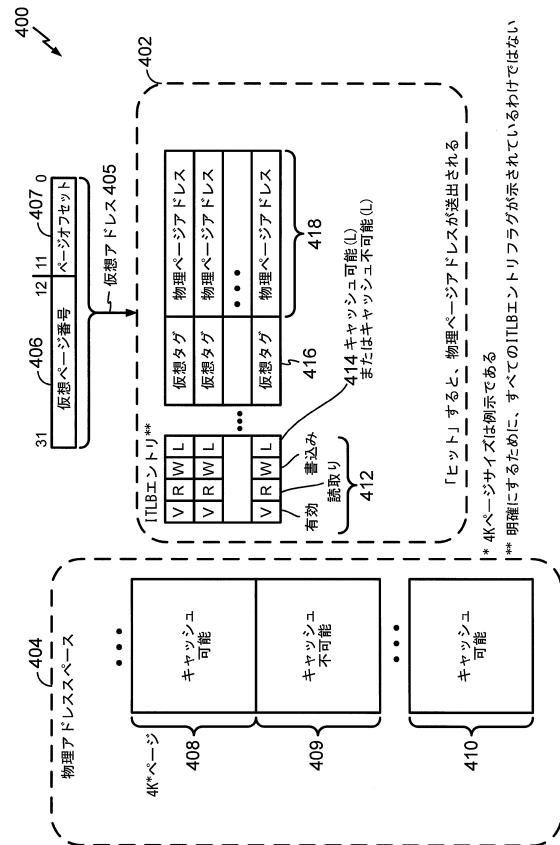
50

|     |                          |    |
|-----|--------------------------|----|
| 220 | メモリ管理ユニット(MMU)           |    |
| 221 | キャッシュコントローラ回路            |    |
| 222 | 書込み制御回路                  |    |
| 226 | レベル2命令/データキャッシュ(L2キャッシュ) |    |
| 227 | システムメモリ                  |    |
| 228 | フェッチ出力                   |    |
| 230 | ミス情報                     |    |
| 232 | アクセス要求                   |    |
| 234 | 第1のポート(ポートA)             |    |
| 235 | キャッシュ不可能要求               | 10 |
| 236 | システムメモリ出力バス              |    |
| 238 | 出力                       |    |
| 240 | 出力                       |    |
| 244 | 出力                       |    |
| 300 | プログラムセグメント               |    |
| 302 | 命令                       |    |
| 304 | プログラム相対バイトインジケータ         |    |
| 306 | ADD R5、R4、R3命令           |    |
| 307 | ロード(LOAD)命令              |    |
| 308 | LOAD命令                   | 20 |
| 309 | ADD命令                    |    |
| 310 | ストア(STORE)命令             |    |
| 320 | L1 Iキャッシュライン             |    |
| 322 | 第1のL1 Iキャッシュライン          |    |
| 323 | タグフィールド                  |    |
| 324 | 制御フラグCn                  |    |
| 326 | 第2のL1 Iキャッシュライン          |    |
| 327 | タグフィールド                  |    |
| 328 | 制御フラグCn                  |    |
| 330 | 16ビットフィールド               | 30 |
| 332 | 32ビットフィールド               |    |
| 333 | 16ビットフィールド               |    |
| 334 | 16ビットフィールド               |    |
| 336 | 16ビットフィールド               |    |
| 338 | 16ビット拡張フィールド             |    |
| 340 | 16ビットフィールド               |    |
| 342 | 16ビットフィールド               |    |
| 343 | 16ビットフィールド               |    |
| 344 | 16ビットフィールド               |    |
| 346 | 16ビット拡張フィールド             | 40 |
| 400 | 仮想メモリシステム                |    |
| 402 | 命令変換ルックアサイドバッファ(ITLB)    |    |
| 404 | 物理メモリ                    |    |
| 405 | 仮想アドレス                   |    |
| 406 | 仮想ページ番号                  |    |
| 407 | ページオフセット                 |    |
| 408 | キャッシュ可能ページ               |    |
| 409 | キャッシュ不可能ページ              |    |
| 410 | キャッシュ可能ページ               |    |
| 412 | フラグ                      | 50 |

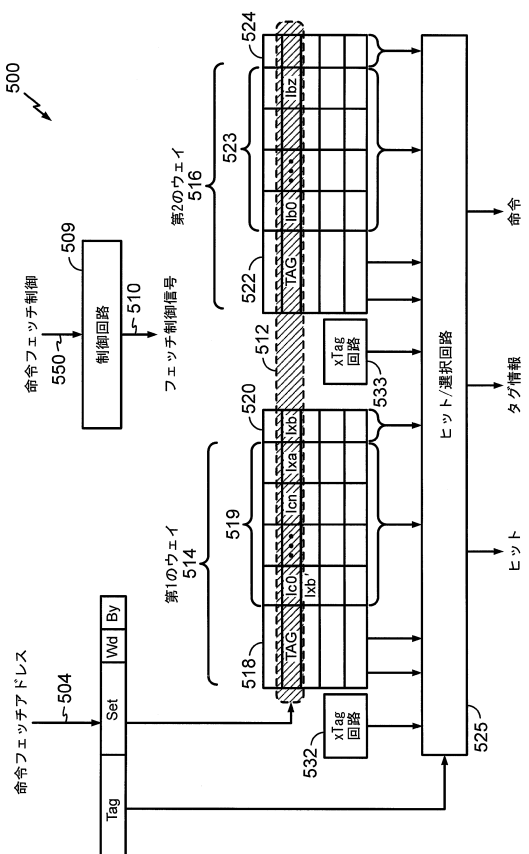
|     |                   |    |
|-----|-------------------|----|
| 414 | キャッシュ可能インジケータ(L)  |    |
| 416 | 仮想アドレスタグ          |    |
| 418 | 物理ページアドレス         |    |
| 440 | 変換サブシステム          |    |
| 442 | プロセッサパイプライン       |    |
| 444 | ITLB              |    |
| 447 | xTag回路            |    |
| 448 | L1 Iキャッシュ         |    |
| 449 | xPビット             |    |
| 450 | L2キャッシュ           | 10 |
| 452 | システムメモリ           |    |
| 454 | 書込み制御回路           |    |
| 456 | エントリ              |    |
| 457 | 第1のライン            |    |
| 458 | キャッシュ可能インジケータ(L)  |    |
| 459 | 第2のライン            |    |
| 460 | CAMタグ             |    |
| 462 | 物理ページ番号、物理ページアドレス |    |
| 463 | 物理アドレス            |    |
| 464 | ページオフセット          | 20 |
| 465 | タグ                |    |
| 470 | Iキャッシュタグ          |    |
| 471 | 基本エクステンツ          |    |
| 472 | 拡張フィールド           |    |
| 473 | 第1の基本エクステンツ       |    |
| 474 | 第1の部分Ia           |    |
| 475 | 第2の部分Ib           |    |
| 476 | Ib(Ib')           |    |
| 480 | xTag              |    |
| 482 | 検出(Dt)回路          | 30 |
| 483 | 第1のデコード(Dc)回路     |    |
| 484 | バッファ/ホールド(B&H)回路  |    |
| 485 | 再フェッチ回路           |    |
| 486 | 再結合回路             |    |
| 487 | マルチプレクサ           |    |
| 500 | 命令キャッシュ回路         |    |
| 514 | 第1のウェイ            |    |
| 516 | 第2のウェイ            |    |
| 518 | タグビットフィールド        |    |
| 519 | データのライン           | 40 |
| 520 | 拡張フィールド           |    |
| 522 | タグビットフィールド        |    |
| 523 | データのライン           |    |
| 524 | 拡張フィールド           |    |
| 532 | xTag回路            |    |
| 533 | xTag回路            |    |
| 600 | プロセス              |    |



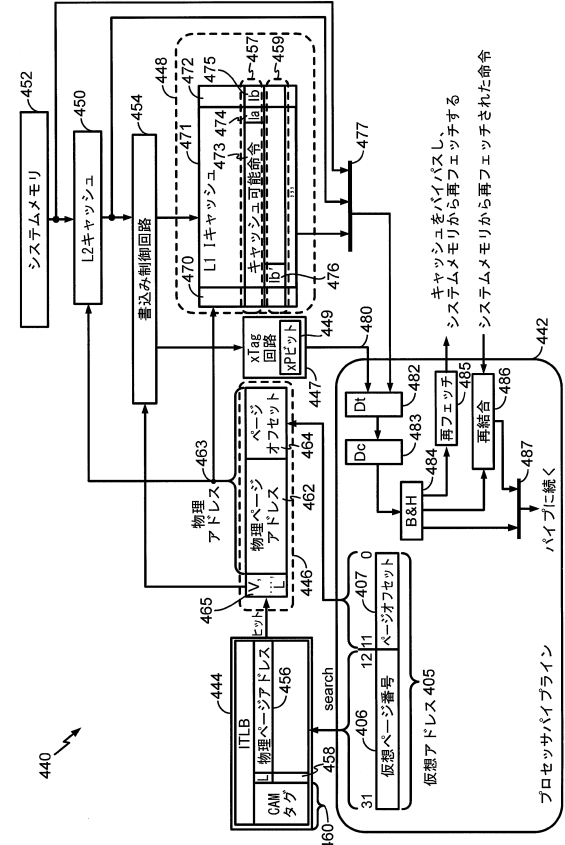
【図 4 A】



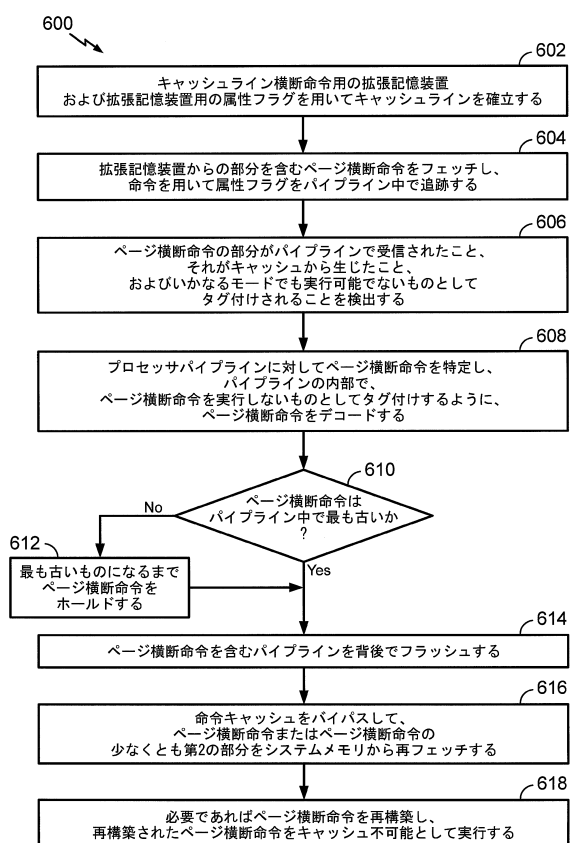
【図 5】



【図 4 B】



【図 6】



---

フロントページの続き

- (72)発明者 ジェームズ・ノリス・ディエフェンダーファー  
アメリカ合衆国・カリフォルニア・92121・サン・ディエゴ・モアハウス・ドライブ・577  
5
- (72)発明者 マイケル・スコット・マッキルヴェイン  
アメリカ合衆国・カリフォルニア・92121・サン・ディエゴ・モアハウス・ドライブ・577  
5
- (72)発明者 ブライアン・マイケル・ステンペル  
アメリカ合衆国・カリフォルニア・92121・サン・ディエゴ・モアハウス・ドライブ・577  
5

審査官 酒井 恭信

- (56)参考文献 米国特許出願公開第2006/0265572 (US, A1)  
特表2008-541314 (JP, A)  
米国特許出願公開第2007/0255905 (US, A1)  
特表2009-535743 (JP, A)  
特開昭60-020255 (JP, A)  
特開平02-054351 (JP, A)  
特開2006-318051 (JP, A)  
米国特許出願公開第2009/0119485 (US, A1)  
特表2011-503719 (JP, A)  
米国特許第7330959 (US, B1)

- (58)調査した分野(Int.Cl., DB名)  
G06F 12/08