

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7149405号
(P7149405)

(45)発行日 令和4年10月6日(2022.10.6)

(24)登録日 令和4年9月28日(2022.9.28)

(51)国際特許分類	F I
G 0 6 F 9/38 (2006.01)	G 0 6 F 9/38 3 3 0 B
G 0 6 F 12/0875(2016.01)	G 0 6 F 12/0875 1 0 0
G 0 6 F 12/0895(2016.01)	G 0 6 F 12/0895 1 0 0

請求項の数 10 (全28頁)

(21)出願番号	特願2021-500526(P2021-500526)	(73)特許権者	591016172
(86)(22)出願日	令和1年7月3日(2019.7.3)		アドバンスト・マイクロ・デバイス
(65)公表番号	特表2021-530782(P2021-530782 A)		・インコーポレイテッド
(43)公表日	令和3年11月11日(2021.11.11)		ADVANCED MICRO DEVI
(86)国際出願番号	PCT/US2019/040497		CES INCORPORATED
(87)国際公開番号	WO2020/014066		アメリカ合衆国 9 5 0 5 4 カリフォル
(87)国際公開日	令和2年1月16日(2020.1.16)		ニア州、 サンタ クララ、 オーガスティ
審査請求日	令和4年7月1日(2022.7.1)	(74)代理人	ンドライブ 2 4 8 5
(31)優先権主張番号	16/030,031		100108833
(32)優先日	平成30年7月9日(2018.7.9)		弁理士 早川 裕司
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100111615
早期審査対象出願			弁理士 佐野 良太
		(74)代理人	100162156
			弁理士 村雨 圭介
		(72)発明者	トーマス クルーカー

最終頁に続く

(54)【発明の名称】 複数のテーブルの分岐ターゲットバッファ

(57)【特許請求の範囲】

【請求項 1】

命令実行パイプラインと、

第一数の第一分岐ターゲットバッファテーブルエントリを含む第一 B T B テーブルであって、前記第一 B T B テーブルエントリのそれぞれは第一タグ及び第一ターゲットアドレスを含み、前記第一ターゲットアドレスのそれぞれは第一幅を有する、前記第一 B T B テーブルと、

第二数の第二 B T B テーブルエントリを含む第二 B T B テーブルであって、前記第二 B T B テーブルエントリのそれぞれは第二タグ及び第二ターゲットアドレスを含み、前記第二ターゲットアドレスのそれぞれは前記第一幅とは異なる第二幅を有する、前記第二 B T B テーブルと、

所与の予測アドレスについて予測されたターゲットアドレスを提供するように構成される分岐予測器と、

を含む、プロセッサであって、

前記プロセッサは分岐命令の分岐特性に基づいて、前記第一 B T B テーブルの第一 B T B テーブルエントリか、前記第二 B T B テーブルの第二 B T B テーブルエントリかのいずれかに前記第一ターゲットアドレスを含む分岐記述子を格納するように構成される、プロセッサ。

【請求項 2】

前記分岐特性は、分岐ターゲットアドレスサイズである、請求項 1 に記載のプロセッサ。

【請求項 3】

前記分岐特性は分岐タイプである、請求項 1 に記載のプロセッサ。

【請求項 4】

前記第一 B T B テーブルエントリは、ジャンプ分岐命令、呼び出し分岐命令、戻り分岐命令、及び条件分岐命令のうち少なくとも 1 つについて構成される、請求項 3 に記載のプロセッサ。

【請求項 5】

前記分岐特性は前記 B T B エントリに格納されるビットの数である、請求項 1 に記載のプロセッサ。

【請求項 6】

前記分岐特性は前記 B T B エントリのスレッド識別子である、請求項 1 に記載のプロセッサ。

10

【請求項 7】

第一メモリレベルキャッシュ及び第二メモリレベルキャッシュ、をさらに含み、
前記第一 B T B テーブル及び前記第二 B T B テーブルは前記プロセッサの同じメモリレベルキャッシュに含まれる、請求項 1 に記載のプロセッサ。

【請求項 8】

前記第一 B T B テーブルの前記第一 B T B テーブルエントリは、予測ルックアップについて N ウェイアソシアティブであり、

前記第二 B T B テーブルの前記第二 B T B テーブルエントリは、予測ルックアップについて M ウェイアソシアティブであり、

20

前記 M 及び前記 N は 1 以上であり、前記 M は前記 N とは異なる、請求項 1 に記載のプロセッサ。

【請求項 9】

前記第一数の前記第一 B T B テーブルエントリは、前記第二数の前記第二 B T B テーブルエントリとは異なる、請求項 1 に記載のプロセッサ。

【請求項 10】

前記 B T B テーブルエントリのそれぞれは、前記 B T B エントリのそれぞれを、セット内で最近最も使用された (M R U) ステータス、及び最近最も使用されていない (L R U) ステータスのうちの 1 つとしてマーク付けする最近使用されたステータスビットを含み、

30

前記プロセッサは、前記最近使用されたステータスビットの状態に基づいて置き換えポリシーに従い前記セット内の B T B エントリをエビクトするように構成される、請求項 1 に記載のプロセッサ。

【発明の詳細な説明】

【背景技術】

【0001】

プロセッサにおける命令パイプラインは、命令ストリームのさまざまな命令を並行して実行することができる複数のパイプラインステージにおいて命令を処理することにより、命令実行スループットを向上させる。このようなパイプラインは、命令をフェッチ、復号、マッピング、及び実行することに続いて、結果をレジスタなどの別のユニットに書き込むための別個のユニットを含むことが多い。パイプラインの命令フェッチユニットは、プロセッサパイプラインの次のステージに命令ストリームを提供する。一般に、命令フェッチユニットは、パイプラインの残りの部分に命令を継続的に供給し続けるために、命令キャッシュを使用する。

40

【0002】

パイプラインにおける実行ステージにおいて分岐を解消するまでプロセッサが待機してから、命令フェッチステージにおいて次の命令をフェッチする場合に、命令ストリーム内の分岐命令は、パイプラインストールをもたらす場合がある。分岐予測器は、条件分岐が成立、または不成立であるかどうかを予測しようとし得る。一部の実施態様では、分岐予測器は、分岐命令自体を復号して実行することによって分岐命令を計算する前に、分岐タ

50

ターゲット予測を使用して、成立した条件または無条件分岐のターゲットを予測する。分岐ターゲットは、計算されたアドレスからのオフセット、またはレジスタを介した間接参照に基づき得る。

【0003】

従来、分岐ターゲットバッファ(BTB)は、予測された分岐ターゲットを含む分岐情報を格納するプロセッサ内の単一の小さなメモリキャッシュである。予測は、命令アドレスを、BTBに格納されている以前に実行された命令アドレスと比較することを含む。通常、予測が成功することにより、プロセッサがターゲットアドレスを取得するステップの実行をスキップすることができるため、予測は処理時間を節約する。プロセッサは、BTB内で実行の次のステップについてのアドレスをルックアップすることにより、時間を節約する。したがって、BTBがターゲットアドレスについてのヒットを生成する頻度は、プロセッサが命令を実行することができる速度に直接影響する。多くの場合、実行速度は、BTBが格納することができるエン트리数に直接関係している。

10

【0004】

本開示は、より良好に理解され得、その多くの特徴及び利点は、添付図を参照することによって、当業者に明らかになる。異なる図面で同じ参照記号を使用することは、類似または同一の要素ということの意味する。

【図面の簡単な説明】

【0005】

【図1】いくつかの実施形態による命令パイプラインアーキテクチャのブロック図である。

20

【図2】一部の実施形態による処理システムのブロック図である。

【図3】いくつかの実施形態による、ターゲットサイズに基づいて複数の分岐ターゲットバッファ(BTB)テーブルを含むプロセッサのブロック図である。

【図4】いくつかの実施形態による、分岐タイプに基づいて複数のBTBテーブルを含むプロセッサのブロック図である。

【図5】いくつかの実施形態による、異なるレベルのメモリ内に複数のBTBテーブルを含むプロセッサのブロック図である。

【図6】いくつかの実施形態によるBTBテーブルについてのBTBエントリのブロック図である。

【図7】いくつかの実施形態による、BTBテーブルについてのBTBエントリのブロック図である。

30

【図8】いくつかの実施形態による、BTBテーブルについてのBTBエントリのブロック図である。

【図9】いくつかの実施形態による、複数のBTBテーブルのうちの1つにターゲットアドレスサイズごとに分岐ターゲットアドレスを格納する方法を示す流れ図である。

【図10】いくつかの実施形態による、複数のBTBテーブルのうちの1つに分岐タイプごとに分岐ターゲットアドレスを格納する方法を示す流れ図である。

【発明を実施するための形態】

【0006】

プロセッサにおいて分岐ターゲットバッファ(BTB)またはBTBテーブルを使用して分岐情報を格納することにより、プロセッサの速度は実質的に向上する。従来、BTBにおいてエン트리数を増加させる唯一の方法は、バッファサイズを増大させることであった。ただし、BTBなどの一部のコンポーネントについてのローカルストレージを増加させることにより、プロセッサにおける速度及び物理スペースの点で他のコンポーネントが犠牲になるため、BTB容量への制限を含む、プロセッサ内のストレージ要素数への制限がある。さらに、プロセッサのアーキテクチャを設計するときに、バランスを保つためのトレードオフ、及び考慮する要因がある。例えば、実際には、BTBへのいくつかの変更により、プロセッサが動作しているときにプロセッサの速度が低下した。他の状況では、BTBが大きいほど、プロセッサの動作中の消費電力が全体的に高くなり、これは望ましくない。

40

50

【 0 0 0 7 】

従来の B T B の一般的なサイズは、1 0 2 4 個のエントリであり、各エントリは、例えば、ターゲットアドレスに対して 2 0 ビットを含む。ただし、すべての B T B エントリに、保存されるアドレスに対して同じビット数が必要なわけではない。例えば、分岐予測器は、ある特定のアドレスに対して利用可能な 2 0 ビットのうちの 5 ~ 1 0 ビットのみを必要とする場合があるが、ターゲットアドレスの最上位ビットの一部が予測アドレスから変更されていないものとして分岐情報において符号化されることができることから、他のアドレスに対して全 2 0 ビットを必要とする場合がある。B T B の容量及びユーティリティを増大させ、プロセッサ内の B T B のフットプリントをほぼ同じに保つために、少なくとも 1 つの実施形態では、B T B は、2 つ以上のバッファまたはテーブルに分割される。ある特定の実施形態によれば、第一バッファは、短いメモリアドレス、例えば 6、8、または他の数のビットを含むこれらのアドレス用にサイズ設定され、指定される。第二バッファは、より長いメモリアドレス、例えば 1 4、1 6、1 8、または 2 0 ビットを含むこれらのメモリアドレス用にサイズ設定され、指定される。一部の実施態様は、第三バッファを利用し、この第三バッファは、さまざまなメモリアドレス長、例えば 6、8、1 0、1 2、1 4、1 6、1 8、または 2 0 ビットを含むこれらのメモリアドレス長用にサイズ設定され、指定される。また、各 B T B エントリは、タグフィールド及び、いくつかの実施形態では、B T B エントリ内の 1 つ以上の別個のフィールドに配置される他の情報を含む。

10

【 0 0 0 8 】

他の実施形態によれば、プロセッサの設計者及びアーキテクトは、1 つ以上のタイプの命令分岐に基づいて複数の B T B テーブルを作成して指定することによって、それぞれのタイプの分岐ターゲットに対して B T B テーブルを作成する。各 B T B エントリは、B T B エントリが追跡する分岐を探索するコストに関連する値を含む。無条件分岐は、プロセッサ復号器によってプロセッサパイプラインにおいて早期に探索されることができ、条件分岐は、実行時まで探索されない場合がある。プロセッサのいくつかの実施形態は、1 サイクル中に、2 つ以上の分岐予測、そのため 2 つの分岐ターゲットアドレスを送達する。これらの複数の分岐は、単一の B T B エントリに格納される。B T B エントリのさまざまな値を説明するために、第一 B T B は、例えば無条件分岐用に作成されて指定され、第二 B T B は、例えば単一の条件分岐用に作成されて指定され、第三 B T B は、例えば複数の分岐を記録するエントリ用に作成される。第一、第二、及び第三 B T B テーブルのそれぞれは、異なるターゲットアドレスサイズ、したがって異なる全体の B T B テーブル幅を有することにより、単一の B T B が、ある特定の固定幅、及びある特定の固定長（エントリ数）のものであった以前の設計と比較して、プロセッサ内の B T B テーブルのフットプリント全体を減少させるメカニズムを提供し得る。B T B エントリを 2 つ以上の B T B テーブルに分割する、または配置するというこの利点により、プロセッサ設計者は、1 つ以上の B T B テーブルの長さを、相互に、または単一の従来の B T B の従来の長さに比べて増加させる、または減少させることができる。3 つの B T B システムについて、第一、第二、及び第三 B T B テーブルは、第一 B T B、第二 B T B、及び第三 B T B のそれぞれの中のエントリに収まらない大きなターゲットアドレスに対して追加のアドレスビット数を提供するために、各 B T B エントリが指すことができる別個のターゲットテーブルによって補完され得る。動作中に、そしていくつかの実施態様によれば、プロセッサ内の B T B ルックアップは、各 B T B ルックアップについてのすべての B T B バッファ間で実行されることにより、単一の従来の B T B バッファと同様のパフォーマンスを提供する。

20

30

40

【 0 0 0 9 】

ある特定の実施態様では、B T B テーブルは、単一のメモリタイプ内に含まれるか、または、一方で他の実施態様では、B T B テーブルは、プロセッサコアに関するマルチレベルメモリ階層の、複数のメモリ、複数のメモリタイプ、または複数のメモリレベル間で分散される。説明を簡単にするために、本明細書では、プロセッサコアに対する、またはプロセッサに対するメモリのレベルまたは配置にそれぞれ対応する、レベル 1 の B T B、レベル 2 の B T B などを参照する。各 B T B レベルは、単一の B T B テーブル、または複数

50

の B T B テーブルを含んでもよい。他の実施態様では、1 つ以上の B T B テーブルは、単一のプロセッサコアについての単一のメモリキャッシュまたはメモリレベルに組み込まれる。さらに他の実施態様では、B T B テーブルは、プロセッサコアまたはプロセッサ操作間で共有される。特定の実施形態では、複数の B T B レベルのうちの 1 つは、複数のプロセッサコアによって共有され、プロセッサコアのそれぞれは、それ自体に 1 つ以上の B T B テーブルを含む。特定のコアに関連するこれらの B T B テーブルを最初に使用してもよく、必要に応じて B T B エントリを共有 B T B テーブルにエビクトしてもよい。プロセッサコア間で共有される 1 つ以上の B T B テーブルを含むプロセッサの利点は、検索速度、全体的なプロセッサ速度、及び全体的なプロセッサ操作における低下によって相殺されることができる。したがって、高速実装のために、マルチコアプロセッサの別個のプロセッサコアごとに複数の B T B テーブルを提供してもよい。

10

【 0 0 1 0 】

動作中に、最近最も使用されていない (L R U) ものなどの下位の値の B T B エントリは、第一 B T B または第一 B T B レベルから、同じレベルまたは別のレベルのメモリ内の B T B などの別の B T B または B T B レベルに最初にエビクトされる。第二 B T B から第三 B T B へのエビクションも L R U ベースで行われる。B T B テーブルには、さまざまなタイプの連想度を使用してもよい。例として、第一 B T B テーブルは、4 ウェイセットアソシアティブである。別の B T B テーブルは、8 ウェイセットアソシアティブのプロセッサキャッシュの形式を取り、各ウェイはウェイあたり 5 1 2、1 0 2 4、2 0 4 8、またはある他の数のエントリを格納する。一般に、第一 B T B テーブルは、N ウェイセットアソシアティブであることができ、第二 B T B テーブルは、M ウェイセットアソシアティブであることができ、N 及び M は互いに同じ、または異なる整数である。他の実施形態では、1 つ以上の B T B テーブルはフルアソシアティブである。連想度のレベルは、特定のプロセッサについて所望の消費電力レベル、プロセッサが動作するシステム、ならびに特定のプロセッサについての処理負荷の所期のタイプ及び量に応じて選択される。

20

【 0 0 1 1 】

一部の実施態様では、B T B テーブルは、下位レベルの B T B テーブルからの条件分岐に対して最近エビクトされたエントリが投入される限られたエントリ数を含むビクティムバッファとして機能する。所与の領域についてのこのようなビクティムバッファのエントリ数を最大にするために、分岐予測器は、B T B テーブルにいかなるターゲットアドレスも記録しない。分岐予測器がビクティムバッファにおいて分岐を見つけ、条件分岐予測器によって決定されたとおりに分岐が取得されると予測するときに、分岐予測器は、分岐位置に関する情報を復号ユニットに送信し、復号ユニットは、復号操作が完了すると利用可能なターゲットアドレスへのリダイレクトをトリガし、予測される位置が実際に条件分岐であることを確認する。

30

【 0 0 1 2 】

図 1 は、いくつかの実施形態による、アウトオブオーダー命令実行を実装するプロセッサ 1 2 0 の命令パイプラインアーキテクチャ 1 0 0 のブロック図である。プロセッサ 1 2 0 の少数の要素のみを示す。命令キャッシュ 1 0 1 は、命令フェッチユニット 1 0 3 によってアクセスされる。データキャッシュ 1 0 2 は、ロード/ストアユニット 1 1 0 によってアクセスされる。命令キャッシュ 1 0 1 の命令は、データキャッシュ 1 0 2 からのデータを含むデータ上で動作する。命令フェッチユニット 1 0 3 は、1 つ以上の分岐ターゲットバッファ (B T B) テーブル (一般に B T バッファ及び B T B とも称される) 1 0 5 に格納される、または提供される分岐ターゲットアドレスを生成する分岐予測器 1 0 4 を含む。いくつかの実施形態によれば、分岐ターゲットアドレスは、プログラムカウンタ 1 1 3 に関連している。B T B テーブル 1 0 5 は、図 1 の分岐予測器 1 0 4 の内部に示されており、B T B テーブル 1 0 5 は、分岐予測器 1 0 4 または命令フェッチユニット 1 0 3 のある特定の要素に近接してプロセッサ 1 2 0 内に位置していても、位置していなくてもよい。復号器 1 0 6 は、命令キャッシュ 1 0 1 からの命令を含む命令をプロセッサ制御信号に変換する。

40

50

【 0 0 1 3 】

リオーダバッファ 1 0 7 は、リザベーションステーション 1 0 9 などのプロセッサの他のコンポーネントによってアクセスされるレジスタ 1 0 8 に命令を、それらの元のフェッチ順序で格納する。リザベーションステーション 1 0 9 は、レジスタ 1 0 8 などのレジスタをリネームし、動的命令スケジューリングを容易にする。リザベーションステーション 1 0 9 は、データがレジスタに格納されて再読み出しされるのを待機するのではなく、データがフェッチされ次第、または計算され次第すぐに、プロセッサがデータをフェッチして再利用することを可能にする。分岐予測器 1 0 4 によって B T B テーブル 1 0 5 に格納される予測アドレスが正しくない、または命令ストリームに回復不可能な例外が発生する場合、すべての命令のリオーダバッファ 1 0 7 はクリアされ、リザベーションステーション 1 0 9 は再初期化される。リオーダバッファ 1 0 7 は、分岐ターゲットアドレスの誤予測のロールバック制御のためのメカニズムを提供する。リザベーションステーション 1 0 9 は、ロード/ストアユニット 1 1 0、ならびに算術論理演算装置 (A L U)、浮動小数点演算装置 (F P U)、及び整数演算装置 (I U) などの 1 つ以上の機能ユニット 1 1 1 に情報を提供する。合わせて、命令フェッチユニット 1 0 3、復号器 1 0 6、リオーダバッファ 1 0 7、リザベーションステーション 1 0 9、ロード/ストアユニット 1 1 0、及び関連するレジスタは、命令実行パイプラインの一実施形態である。

10

【 0 0 1 4 】

図 2 は、一部の実施形態による処理システム 2 0 0 のブロック図である。処理システム 2 0 0 は、図 1 の命令パイプラインアーキテクチャ 1 0 0 及び B T B テーブル 1 0 5 の一部分の一実施例である。処理システム 2 0 0 は、分岐検出器 2 0 2、条件分岐予測器 2 0 3、戻りアドレス予測器 2 0 4、及び分岐予測トレーナー 2 0 5 を有する分岐予測器 2 0 1 を含む。処理システム 2 0 0 は、当業者によって理解されるように、算術ユニット、スケジューラ、テーブルウォーカーなどの他の図示されていない要素を含む。

20

【 0 0 1 5 】

現在のアドレスごとに、分岐予測器 2 0 1 は、現在のアドレスから開始してフェッチされるバイトのブロックの終了アドレス、予測されたブロック内の分岐タイプ、及び次の予測の開始アドレスを含む予測ウィンドウを提供する。どの分岐が予測ウィンドウの一部分であるかを決定するために、分岐検出器 2 0 2 は、すべてのウェイにわたって、例えば B T B テーブル 1 2 0 6 から B T B テーブル N 2 0 7 として表される複数の B T B テーブルのうち 1 つ以上にわたって、すべての B T B テーブルをルックアップする。所与の B T B エントリ上でのタグマッチングは、エントリに記録される 1 つ以上の分岐が予測ウィンドウに存在することを示す。各 B T B エントリは分岐の位置及びタイプを含む。誤予測された分岐に対してリダイレクトが発生するときに、B T B には分岐予測トレーナー 2 0 5 が投入される。

30

【 0 0 1 6 】

図 3 は、いくつかの実施形態による、複数の B T B テーブルを含むプロセッサのブロック図である。プロセッサ 3 0 0 は、追加の構造及び機能を有する、図 1 のプロセッサ 1 2 0 及び B T B テーブル 1 0 5 の特定の実施例である。プロセッサ 3 0 0 は、第一指定領域 3 0 6 内に複数のプロセッサコアのうち第一プロセッサコア 3 0 5 を含む。第一プロセッサコア 3 0 5 は、分岐予測器 3 0 4 によって使用される第一 B T B テーブル 3 0 1、第二 B T B テーブル 3 0 2、及び第三 B T B テーブル 3 0 3 を含む B T B テーブルと相互運用する。いくつかの実施形態によれば、各プロセッササイクルにおいて、分岐予測器 3 0 4 は、1 つ以上の分岐命令からの 1 つまたは 2 つの予測を提供する。

40

【 0 0 1 7 】

各 B T B テーブル 3 0 1、3 0 2、3 0 3 は、それぞれ、エントリが属する予測アドレスを識別するために使用されるタグフィールド 3 0 7、3 0 8、3 0 9、1 つまたは複数の分岐のターゲットを格納するために使用されるターゲットアドレスフィールド 3 1 0、3 1 1、3 1 2、ならびに予測ウィンドウ内の分岐位置、及び分岐タイプを格納するために使用される情報フィールド 3 1 6、3 1 7、3 1 8 を含む。図 2 の分岐予測トレーナー

50

205などの分岐トレーナーは、それぞれのBTBテーブルに投入され、BTBエントリが第一BTBテーブル301、第二BTBテーブル302、または第三BTBテーブル303に記録されているかどうかを、トレーニングされているBTBエントリごとに決定する。

【0018】

BTBテーブル301～303のそれぞれは、異なるサイズまたは幅313、314、315（特に示されない限り、「サイズ」及び「幅」は本明細書では交換可能に使用される）のターゲットアドレスフィールド310、311、312を含む。例えば、第二BTBテーブル302は、第二アドレスサイズ314のアドレス311を含む。いくつかの実施形態によれば、第二アドレスサイズ314は、第一アドレスサイズ313よりも小さい。第二アドレスサイズ314は、第三BTBテーブル303のターゲットアドレスフィールド312の第三アドレスサイズ315よりも大きい。ある特定の実施形態では、例えば、第一BTBテーブル301のBTBエントリの第一セットにおけるBTBエントリ数と比較して、第三BTBテーブル303のBTBエントリの第三セットには、より多くのBTBエントリがある。これが可能であるのは、ターゲットアドレスフィールド312についてのビット数または幅315を減らすことにより、その他のフィールド309、318のサイズを一定に保つときに、第三BTBテーブル303についてより多くのエントリが可能であることからである。すなわち、各BTBテーブル301、302、303についての同じ量のプロセッサフットプリントに対して、幅315が縮小されることにより、第三BTBテーブル303は、指定領域306の固定サイズに対してより多くのエントリを含むことが可能である。他の実施形態では、すべてのターゲットアドレスに対して一定幅のターゲットアドレスフィールドを有する単一のBTBテーブルを含む以前の実施形態と比較して、より小さな指定領域306が可能である。したがって、この実施例では、第三BTBテーブル303は、その他のBTBテーブル301、302よりも小さなフットプリントをプロセッサ300及び指定領域306では占めてもよいが、第二BTBテーブル302及び第一BTBテーブル301よりもサイズ（エントリ）が大きいと見なされる。

【0019】

第三BTBテーブル303におけるエントリの第三アドレス312、または第二BTBテーブル302におけるエントリの第二アドレス311のような、アドレスビットを保持するために追加のビットが必要とされる場合に、追加のアドレステーブル（図3には示されていない）は、第一、第二、及び第三BTBテーブル301、302、303の間で共有される。アドレスフィールドのビットに格納されるアドレス値の代わりに、追加のアドレステーブルが使用されるときに、複数のアドレスビットは、追加のアドレステーブル内のエントリにインデックス付けされる。すなわち、追加のアドレステーブルのエントリへのポインタは、分岐ターゲットアドレスフィールド310、311、312のうちの対応する1つに指定されるBTBエントリの少なくとも一部分に格納される。

【0020】

図4は、いくつかの実施形態による、分岐タイプに基づいてBTBテーブルを含むプロセッサ400のブロック図である。分岐タイプに基づくBTBテーブル構成は、テーブルを異なる値のBTBエントリに割り当て、テーブルのサイズを適切に設定して、さらなるスペースを高位の値のエントリに割り当てることにより、向上したパフォーマンスを提供する。プロセッサ400は、追加の構造及び機能を有する、図1のプロセッサ120及びBTBテーブル105の特定の実施例である。プロセッサ400は、第一指定領域406内に複数のプロセッサコアのうちの第一プロセッサコア405を含む。第一プロセッサコア405は、分岐予測器404によって使用される第一BTBテーブル401、第二BTBテーブル402、及び第三BTBテーブル403を含むBTBテーブルと相互運用する。いくつかの実施形態によれば、各プロセスサイクルにおいて、分岐予測器404は、1つ以上の分岐命令からの1つまたは2つの予測を提供する。

【0021】

各BTBテーブル401、402、403は、エントリが属する予測アドレスを識別す

10

20

30

40

50

るために使用されるタグフィールド407、408、409、1つまたは複数の分岐のターゲットを格納するために使用されるターゲットアドレスフィールド410、411、412、ならびに予測ウィンドウ内の分岐位置、及び分岐タイプを格納するために使用される情報フィールド416、417、418を含む。図2の分岐予測トレーナー205などの分岐トレーナーは、それぞれのBTBテーブルに投入され、BTBエントリが第一BTBテーブル401、第二BTBテーブル402、または第三BTBテーブル403に記録されているかどうかを、トレーニングされているBTBエントリごとに決定する。さらに、BTBテーブル401～403のそれぞれは、それぞれの分岐タイプについてターゲットアドレスフィールド410、411、412を含む。BTBテーブル401～403のそれぞれは、そこに格納される分岐アドレスのそれぞれ1つ以上のタイプに典型的なターゲットアドレスサイズに基づいて、異なるサイズまたは幅413、414、415のターゲットアドレスフィールド410、411、412を含んでもよい。BTBテーブル401、402、403のそれぞれは、1つ以上の分岐タイプに指定される。例えば、第二BTBテーブル402は、第一BTBテーブル401のエントリとは異なる分岐タイプについてのアドレス411を格納する。ある特定の実施形態では、例えば、第三BTBテーブル403内のエントリのタイプが、BTBエントリに関して正しく予測するときに節約されるサイクル数の観点から、プロセッサ400の運用効率にとってより有益であるため、第二BTBテーブル402のBTBエントリのセット内のBTBエントリの数と比較して、第三BTBテーブル403のBTBエントリの第三セット内により多くのBTBエントリがある。それぞれのBTBテーブル401～403のサイズ(エントリ数)は選択され、BTBテーブル401～403は、そこに格納される分岐アドレスタイプ、及びそれらの相互の運用効率に基づいて構築される。一実施形態では、各BTBエントリは、それぞれのフィールドに少なくともタグ及びターゲットアドレスを含む。各エントリは、スレッドID、及び共有インジケータまたはフラグなどの他の情報も含んでもよい。

【0022】

図5は、いくつかの実施形態による、異なるレベルのメモリ内に複数のBTBテーブルを含むプロセッサのブロック図である。プロセッサ500は、分岐命令を処理する分岐予測器502を有するプロセッサコア501を含む。分岐予測器502は、レベル1のBTBテーブル503、1セットのレベル2のBTBテーブル504、505、506、及びレベル3のBTBテーブル507を含むさまざまな構造を有し、それらと相互運用する。レベル1のBTBテーブル503は、第一レベルのメモリ508にある。レベル2のBTBテーブル504、505、506は、第二レベルのメモリ509にある。レベル3のBTBテーブル507は、第三レベルのメモリ510にある。いくつかの実施形態によれば、1セットのレベル1のBTBテーブル503がフルである場合、またはBTBエントリのすべてのセットがフルである場合、BTBエントリは、レベル1のBTBテーブル503から、複数の、または1セットのレベル2のBTBテーブル504～506のうちの1つに上書きされる、またはエビクトされる。レベル2のBTBテーブル504～506のそれぞれは、特定のターゲットアドレスサイズ、もしくはターゲットアドレスサイズの範囲に対して、または他の図に関連して説明されているような特定の分岐命令タイプに対して作成される。

【0023】

レベル2のBTBテーブル504、505、506のうちの1つがフルである場合、プロセッサ500は、そこにBTBエントリを上書きする、またはそこからBTBエントリを第三レベルのメモリ510内のレベル3のBTBテーブル507にエビクトする。いくつかの実施形態では、最下位のエントリなどのBTBエントリは上書きされる。他の実施形態では、BTBエントリは、図示されていない1つ以上のBTBテーブルのうちの1つにエビクトされる。そのような他のBTBテーブルは、別のレベルのメモリ508～510内のBTBテーブルなどの同じ、もしくは別のメモリに、またはプロセッサ内の別の位置に、またはプロセッサコア501の外部のメモリに、またはバスもしくはブリッジによってプロセッサ500に結合されるメモリ内などのプロセッサ500自体にあってもよい

10

20

30

40

50

。例えば、エビクションは、レベル2のメモリ509内のオーバーフローBTBテーブルに対するものである。いくつかの実施形態によれば、分岐タイプまたは分岐ターゲットアドレスサイズに基づいて、1サブセットのBTBエントリのみ、例えば、BTBテーブル504、505、506のうちの1つの中のそれらのBTBエントリのみをエビクトすることが可能である。

【0024】

例示的な実施形態によれば、レベル2のBTBテーブル504、505、506のそれぞれは、レベル2のメモリ509にある。いくつかの実施形態では、BTBテーブル504、505、506のそれぞれは、4ウェイアソシアティブである1,024個のエントリを含む。各BTBエントリは、サイズが68ビット(68b)であり、最大37ビットのターゲットサイズを記録することができる。37ビットよりも大きいサイズのターゲットは、間接レベルを介して別のテーブルに記録されることができる。本明細書に記載されるサイズは、図3のBTBテーブル301~303、図4のBTBテーブル401~403、及び図5のBTBテーブル503~507などに可能なさまざまなサイズの例示である。いくつかの実施形態では、BTBテーブルは、仮想フェッチアドレスのハッシュを使用してインデックス付け及びタグ付けされる。

10

【0025】

いくつかの実施形態によれば、BTBテーブルは、分岐予測サイクルあたり最大2つの分岐の予測をサポートする。他の実施形態では、プロセスサイクルまたは分岐予測サイクルあたり1つの分岐のみが評価される。好ましくは、BTBテーブルは、すべての分岐タイプ及びターゲットサイズをサポートする。各BTBエントリ、例えば図5のレベル1のBTBテーブル503内のそれらのエントリは、予測アドレスから開始し、アライメントされた64Bの終了、または第一静的分岐の終了、または予測ウィンドウ内の「最後ではない」(NLIPW)第一動的分岐もしくは分岐対のいずれか早い方の終了まで延びる予測ブロックに関する、最大2つのターゲット、及び付随する分岐情報を格納する。

20

【0026】

いくつかの実施形態によれば、プロセッサ設計に応じて、BTBテーブル内の各エントリは、ある特定のフィールドを含む。説明するために、図6は、いくつかの実施形態によるBTBテーブルについてのBTBエントリ600のブロック図である。図6では、以下は、BTBテーブルについての68ビットエントリに関するBTBエントリ600内のそれぞれの位置ビットでのフィールドの記述であり、各BTBエントリ600は、ゼロ(0)から67までインデックス付けされる68ビットを含む。ビットのサイズまたは数は一例に過ぎず、他のサイズ及びビット分布は本明細書に提供されるガイドラインを使用して実装されてもよい。AUXPREDEN601は、位置[67]における1ビットであり、補助予測器の有効化を示す。SHRD602は、位置[66]における1ビットであり、エントリがスレッド間で共有されているか、共有可能であることを示す。TID603は、位置[65]における1ビットであり、スレッド識別子(ID)を示す。TAG604は、位置[64:52]に位置している13ビットであり、分岐エントリについてのタグを格納する。LE605は、位置[51:50]における2ビットであり、符号長または「レングス符号化」を示す。LE605の状態は、これらの値1,1と1,0と0,1に関する状態を含む。PAIR606は、位置[49]における1ビットであり、オーバーフローアドレスバッファと対になることを示す。BRTYPE607は、位置[48:44]内にある5ビットであり、以下でより詳細に説明されるように、分岐タイプの符号化を示す。BR1END608は、位置[43:38]にある6ビットであり、[5:0]の範囲に関して1インクリメントされ得るBranch1のエンドポイントを示す。

30

40

【0027】

LE605及びPAIR606の状態は、合わせて、各(第一レベルまたは第二レベルの)BTBエントリの位置[37:0]におけるビットがどのように使用されるかを示す。USEOVERFLOW(USEOA)609は、位置[37]における1ビットであり、オーバーフローアドレスバッファの使用を示す。例えば、位置[48:32]にある

50

ターゲットは、オーバーフローアドレス識別子 $O A I D X 6 1 0$ （以下を参照）として示されるオーバーフローアドレスバッファのエントリからのものであり、 $P A I R = 0$ の場合に使用される。 $O A I D X 6 1 0$ は、位置 $[3 6 : 3 2]$ にある 5 ビットであり、オーバーフローアドレスバッファのインデックスを示す。 $O A I D X 6 1 0$ は、 $P A I R = 0$ 及び $U S E O A = 1$ の場合に使用される。 $B R O E N D 6 1 1$ は、位置 $[3 7 : 3 2]$ における 6 ビットの代替の割り当て $6 1 2$ であり、 $[5 : 0]$ の範囲に関して 1 インクリメントされ得る $B r a n c h 0$ のエンドポインタを示す。

【 0 0 2 8 】

$B R O T G T 6 1 3$ は、インデックス $(X + 1)$ から 3 1 までの第一範囲 $6 1 5$ にわたる $B r a n c h 0$ ターゲット、または第一分岐ターゲットアドレスであり、ここで X は整数であり、第一分岐ターゲットアドレスは $L E 6 0 5$ の 2 ビットの状態に応じて可変長のものである。例えば、 $B R O T G T 6 1 3$ は、 $L E = 1$, 1 についての位置 $[3 1 : 2 0]$ における 1 2 ビットであり、 $B R O T G T 6 1 3$ は、 $L E = 1$, 0 についての位置 $[3 1 : 1 6]$ における 1 6 ビットであり、そして $B R O T G T 6 1 3$ は、 $L E = 0$, 1 についての位置 $[3 1 : 1 2]$ における 2 0 ビットである。 $L E = 0$, 0 の場合、 $B R O T G T 6 1 3$ は、有効ではない、または使用されない。

10

【 0 0 2 9 】

$B R 1 T G T 6 1 4$ は、位置インデックス $[0]$ から位置インデックス $[X]$ までの第二範囲 $6 1 6$ にわたる $B r a n c h 1$ ターゲット、または第二分岐ターゲットアドレスであり、ここで X は整数である。例えば、 $B R 1 T G T 6 1 4$ は、 $P A I R = 0$ 及び $U S E O A = 0$ の場合に位置 $[3 6 : 0]$ において 3 7 ビットであり、 $B R 1 T G T 6 1 4$ は、 $P A I R = 0$ 及び $U S E O A = 1$ の場合に位置 $[3 1 : 0]$ において 3 2 ビットであり、 $B R 1 T G T 6 1 4$ は、 $P A I R = 1$ 及び $L E = 1$, 1 の場合 ($B R O T G T 6 1 3$ は上記のように 1 2 ビットである場合) に位置 $[1 9 : 0]$ において 2 0 ビットであり、 $B R 1 T G T 6 1 4$ は、 $P A I R = 1$ 及び $L E = 1$, 0 の場合 ($B R O T G T 6 1 3$ は上記のように 1 6 ビットである場合) に位置 $[1 5 : 0]$ において 1 6 ビットであり、そして $B R 1 T G T 6 1 4$ は、 $P A I R = 1$ 及び $L E = 0$, 1 の場合 ($B R O T G T 6 1 3$ は上記のように 2 0 ビットである場合) に位置 $[1 1 : 0]$ において 1 2 ビットである。このようにして、 $B T B$ テーブルの各レベル 1、レベル 2、またはレベル 3 の $B T B$ エントリは、いくつかの実施形態による、2 つの分岐ターゲットアドレス $B R O T G T 6 1 3$ 及び $B R 1 T G T 6 1 4$ を格納することができる。

20

30

【 0 0 3 0 】

分岐タイプ符号化に関して、いくつかの実施形態によれば、5 ビット分岐タイプは、 $B r a n c h 1$ の以下の特性を示す。5 ビット分岐タイプの最初の 2 ビットに対して、各分岐命令についての分岐タイプは、位置 $[1 : 0]$ において次の、値 0 , 0 に関する第一状態が無効に対応し、値 0 , 1 に関する第二状態が $J M P$ (ジャンプタイプ) に対応し、値 1 , 0 に関する第三状態が $C A L L$ (呼び出しタイプ) に対応し、値 1 , 1 に関する第四状態が $R E T$ (戻りタイプ) に対応するように指定される。5 ビット分岐タイプの位置 $[2]$ において、1 ビットは、値 0 についての固定ターゲット、及び値 1 についての可変ターゲットを指定する。分岐は、デフォルトによって固定ターゲットとしてインストールされ、これらの分岐がたとえ $B T B$ エントリに最初にインストールされたものとは異なるターゲットを取得しても、間接分岐のみが可変ターゲットに変換される。

40

【 0 0 3 1 】

5 ビット分岐タイプの最後の 2 ビットに対して、方向性ステータスは、次のスキームに従って位置 $[4 : 3]$ に提供される。直接固定分岐及び間接固定分岐は、位置 $[4 : 3]$ を見ることによって区別される。0 , 0 の値は、 $D E$ (復号) リダイレクトを介して見つかる無条件直接 ($U D$) に対応する。いくつかの実施形態によれば、 $U D$ は静的であり、定義によって固定される。分岐が属しているレベル 2 の $B T B$ のテーブルを知るために、 $U D$ には区別が必要とされる。位置 $[4 : 3]$ における 1 , 0 の値は、 $U D$ ではなく静的に対応する。これらの分岐は、デフォルトによる方向に関して静的としてインストールさ

50

れ、条件分岐が不成立としてリダイレクトされる場合、条件分岐のみが動的方向に変換される。

【 0 0 3 2 】

位置 [4 : 3] における 1 , 1 の値は、「UDではなく、動的で、予測ウィンドウ内の最後 (L I P W) 」の状態に対応する。L I P W は、プロセッサが 6 4 B キャッシュライン内の最後の分岐として識別するいずれかの分岐に設定される。このステータスは、B T B エントリではどの分岐も成立ではない場合に、B T B 読み出しロジックが次のフェッチアドレス (F A) を決定するのに役立つ。本来であれば、次の F A は、B T B エントリ内の最後の分岐の直後の命令アドレスであるが、L I P W が設定されている (そしてどの分岐も成立ではないと仮定する) 場合、次の F A は、シーケンシャルキャッシュラインアド

10

【 0 0 3 3 】

第一分岐が不成立であると予測されることができるときにのみ、分岐ターゲットが対になる値を含むことから、B r a n c h 0 は、動的な固定 J M P タイプの分岐のみであることができる。一実施形態では、分岐ターゲット対は、以下の特性を有する。対あたり 1 つの分岐のみが、可変ターゲットのものであることができる。これらの対は、C a l l / C a l l 、C a l l / R e t 、R e t / C a l l 、または R e t / R e t を含むことができない。これらの対は、前述のように、複数のターゲットサイズの組み合わせをサポートする。対になることが可能であるかどうか、そしてどの組み合わせを使用すべきかを知るために、単一のタイプは、短い分岐予測ターゲット、中間の分岐予測ターゲット、及び長い分岐予測ターゲットの間で区別する。この区別も使用して、別の B T B テーブルからオーバーフローレベル 2 の B T B テーブルまたはレベル 3 の B T B テーブルなどに、分岐ターゲットアドレスをピクティムにする、またはエビクトするとき、これらの B T B テーブルのうちどの B T B テーブルを使用すべきかを選択する。

20

【 0 0 3 4 】

以下の説明は、例示的な一実施形態による、図 4 の B T B テーブル 4 0 1 、4 0 2 、4 0 3 、及び図 5 の B T B テーブル 5 0 3 ~ 5 0 7 などの第一 B T B テーブルの第一 B T B 分岐ターゲットアドレスビットに関してさらなる詳細を提供する。いくつかの実施形態によれば、ターゲットアドレス相当の 2 つの分岐を格納し、分岐タイプなどに基づいて、これら 2 つの分岐に可能なケースの多くをカバーするために、ターゲットフィールドに 3 2 ビットが必要とされる。必要なターゲットサイズは、直接分岐全体で 1 ビットから 4 9 ビットまで変わるが、戻りタイプの分岐は、いかなるターゲットビットをも必要としない。可変ターゲット分岐について、発生した第一ターゲットは、本明細書でさらに説明されるように B T B においてトレーニングされ、他のターゲットは、間接ターゲット予測器においてトレーニングされる。

30

【 0 0 3 5 】

ターゲットフィールドは、ターゲット間で共有され、それぞれは、エントリが 2 つの分岐ターゲットアドレスを格納する場合に、通常、異なる長さを有する。可変長分岐を実装することにはタイミングの観点からコストがかかりすぎるため、ターゲットを格納する、より最適化されたアプローチは、2 つの分岐に対して 3 つのターゲット長のバリエーションを使用し、図 4 の B T B テーブル 4 0 1 、4 0 2 、4 0 3 、及び図 5 の B T B テーブル 5 0 4 、5 0 5 、5 0 6 などのそれぞれの B T B テーブルにおいて、ジャンプタイプ、呼び出しタイプ、戻りタイプなどの分岐タイプを使用して符号化することである。分岐ターゲットが大きすぎて、第二分岐のターゲットと並べて収めることができないケースがある。このようなケースでは、分岐は完全なターゲットアドレスまたはターゲットビットを格納するために独自の B T B エントリを必要とする。2 つのケースがあり、(1) その他の分岐タイプのフィールド及び終了フィールド (上記のとおり) がターゲットビットとして再利用される、エントリストレージあたりの単一の分岐、というケースがある。このメカニズムにより、3 7 ビットのターゲットアドレスにサポートが可能である。そして (2)

40

50

O A I D Xフィールドが指すオーバーフローアドレスエントリによって上位ビットが提供される、37ビットより長いターゲット、というケースがある。第一B T Bにヒットした後にオーバーフローアドレス構造を必要とするターゲットアドレスについてのリダイレクトは、追加のリダイレクトペナルティを負う。B T B置き換えポリシーについての例として、いくつかの実施形態によれば、第一B T Bテーブル401またはB T Bテーブル503は、置き換えポリシーの一部としてラウンドロビンスキームを使用する。第一B T Bテーブル内の256セットのそれぞれは、2ビットのラウンドロビン(R R)カウンタを含む。

【0036】

図1の命令パイプラインアーキテクチャ100、図4のプロセッサ400、または図5のプロセッサ500などにおける、パイプラインに関して、B T Bテーブルは、省電力のために読み出しが抑制されない限り、すべての予測フローについて読み出される。いくつかの実施形態によれば、第一B T B読み出しは、B P 0と指定される第一ステップで開始される。後続のステージは、予測パイプラインにおけるステージのようなB P 1、B P 2などとして設計される。B T Bテーブルは、B P 2からB P 0にリダイレクトする。

【0037】

図4のプロセッサ400、及び図5のプロセッサ500などのプロセッサのいくつかの実施形態によれば、上記のように、B T Bテーブル401、402、403、及びテーブル504、505、506に具現化されるような3つの論理テーブルがある。B T Bテーブルの各テーブルは、図4のような分岐タイプ、または図3のようなターゲットサイズのサブセットをサポートする。B T Bテーブル401~403、503~507に具現化される、そのようなマルチテーブルB T Bは、いくつかの利点を提供する。例えば、そのような配置は、プロセッサでは面積効率がよい。各B T Bテーブル401~403、504~506は、その分岐タイプまたはターゲットサイズのサブセットについてストレージ容量のみを有する。いくつかの実施形態では、各B T Bテーブルは高い連想度を提供するが、ある特定の实施形態では、各B T Bテーブル401~403、504~506はフルL R Uを実装することができる。高い実際の連想度を使用することは、実装するのにコストがかかる。複数のB T Bテーブルの使用は、より高い値の分岐(例えば、対になった分岐、及び動的分岐)を他の分岐と区別する。複数のB T Bテーブル301~303、401~403、503~507は、無制限の分岐ターゲットサイズをサポートするが、複数のB T Bテーブル301~303、401~403のうちの1つのB T Bテーブルのみが全分岐ターゲットアドレスサイズをサポートする必要がある。複数のB T Bテーブルの使用は、プロセッサフロアプランの柔軟性を向上させる。例えば、B T Bテーブル401~403のうちの2つのみが、動的分岐をサポートする必要がある。したがって、これらのB T Bテーブルは、条件予測器の近くに配置されることができる。動作中に、プロセッサの特定の实施形態によれば、B T Bテーブルのうちの2つのみが、著しい数のターゲットアドレスビットを駆動する。

【0038】

以下は、例示的なB T B構成及びサイズ設定スキームに関するさらなる詳細を提供する。第一B T Bテーブル401などの第一B T Bテーブル(インデックス0)は、小さなエントリを格納する。第二B T Bテーブル402などの第二B T Bテーブル(インデックス1)は、D Eリダイレクトを通して探索される分岐(例えば、これらの分岐は低いリダイレクトペナルティを有する)を格納する。第三B T Bテーブル403などの第三B T Bテーブル(インデックス2)は、すべての他のタイプを格納する。例えば、分岐タイプは次のように割り振られる。第一B T Bテーブルは、単一のR E T分岐タイプ、及び単一の静的固定条件付きJ M Pの短い分岐タイプについてのアドレスを格納する。第二B T Bテーブルは、「超長」分岐タイプではない単一の静的固定無条件直接呼び出し、及び超長分岐タイプではない単一の静的固定無条件直接ジャンプを格納する。第三B T Bテーブルは、すべての他の分岐タイプを格納する。他の数のB T Bテーブル、ならびに他のタイプ及びターゲットサイズの割り振りは、図3~5に関連して説明されるものを超えることができ

10

20

30

40

50

る。

【 0 0 3 9 】

B T B テーブルの容量の一実施例は、合計 5 , 6 3 2 個の B T B エントリを含む例示的実施態様では次の通りである。第一 B T B テーブルは、合計 2 , 0 4 8 個の B T B エントリについて、5 1 2 セット及び 4 ウェイを含む。第一 B T B エントリは、小計 7 3 , 7 2 8 ビットについて、エントリあたり 3 6 ビットを含む。第二 B T B テーブルは、合計 1 , 0 2 4 個の B T B エントリについて、2 5 6 セット及び 4 ウェイを含む。第二 B T B エントリは、小計 6 5 , 5 3 6 ビットについて、エントリあたり 6 4 ビットを含む。第三 B T B テーブルは、合計 2 , 5 6 0 個の B T B エントリについて、5 1 2 セット及び 5 ウェイを含む。第三 B T B エントリは、小計 1 8 4 , 3 2 0 ビットについて、エントリあたり 7 2 ビットを含む。合計では、B T B テーブルのセットは、1 3 ウェイ、5 , 6 3 2 個のエントリを含み、図 4 の指定領域 4 0 6 などの 1 つ以上の指定領域内の要素などの 3 2 3 , 5 8 4 ビットを必要とする。

10

【 0 0 4 0 】

この実施例では、マルチテーブルの B T B テーブルについて、オーバーフローアドレスバッファは、図 4 の複数の B T B テーブル 4 0 1 ~ 4 0 3 のセット、または図 5 の B T B テーブル 5 0 4 ~ 5 0 6 のセットの間で共有される。分岐タイプ符号化は、第一 B T B テーブル 4 0 1、5 0 4、第二 B T B テーブル 4 0 2、5 0 5、及び第三 B T B テーブル 4 0 3、5 0 6 の間で異なる。ただし、{ L E、P A I R、T Y P E } に従って定義された B T B 分岐タイプと、{ T a b l e N u m b e r、L E、P A I R、T Y P E } に従って記述された B T B テーブル分岐タイプとの間には直接マッピングがある。これらの B T B 記述 (フィールド) は、スワップ及び B T B エビクション中に使用される。

20

【 0 0 4 1 】

他の実施形態によれば、図 4 の第一 B T B テーブル 4 0 1 などの第一 B T B テーブル内の B T B エントリに対するフィールド及び位置ビットについて、これらのフィールド及び位置ビットは、位置 [3 5 : 3 4] における 2 ビットに関する U N U S E D、位置 [3 3] における 1 ビットに関する S H R D、位置 [3 2] における 1 ビットに関する T I D、位置 [3 1 : 2 0] における 1 2 ビットに関する T A G、位置 [1 9 : 1 4] における 6 ビットに関する B R E N D、位置 [1 3 : 1 2] における 2 ビットに関する B R T Y P E、位置 [1 1 : 0] における 1 2 ビットに関する B R T G T を含む。U N U S E D フィールドは、対応するマクロ位置に未使用のビットがあるかどうかを示す。S H R D フィールドは、B T B エントリがスレッド間で共有されるか、共有可能であるかを示す。T I D フィールドはスレッド I D である。B R E N D フィールドは、[5 : 0] の範囲に関して 1 インクリメントされ得る、分岐エンドポイントである。B R T Y P E フィールドは、分岐タイプ符号化を示す。いくつかの実施形態によれば、第一 B T B テーブル 4 0 1 は、R E T (戻り) タイプの分岐、及び少なくとも 1 つのタイプの J M P 分岐を含むいくつかのタイプの分岐に関する情報を含む。いくつかの実施形態によれば、B R T Y P E フィールドは、戻りタイプについての全分岐タイプを記録する位置 [4 : 0] における B R T G T に関して、以下の、状態 0 , X (X は 0 または 1 であることができる) が無効状態に対応し、値 1 , 0 に関する状態が J U M P に対応し、値 1 , 1 に関する状態が R E T に対応するように符号化される。

30

40

【 0 0 4 2 】

図 7 は、いくつかの実施形態による、第二 B T B テーブルについての B T B エントリ 7 0 0 のブロック図である。以下は、図 4 の第二 B T B テーブル 4 0 2、または図 5 の第二 B T B テーブル 5 0 5 などの第二 B T B テーブル内の B T B エントリについてのフィールド及び位置ビットの記述である。図 7 の各 B T B エントリ 7 0 0 は、図 6 の B T B エントリ 6 0 0 よりも短い、6 4 ビット幅である。図 7 では、B T B エントリ 7 0 0 は、以下の、位置 [6 3] における 1 ビットに関する S H R D 7 0 1、位置 [6 2] における 1 ビットに関する T I D 7 0 2、位置 [6 1 : 4 9] における 1 3 ビットに関する T A G 7 0 3、位置 [4 8 : 4 7] における 2 ビットに関する L E 7 0 4、位置 [4 6 : 4 5] にお

50

る2ビットに関するBRTYPE705、位置[44:39]における6ビットに関するBREND706、位置[38:0]における39ビットに関するBRTGT707というフィールドを含む。SHRD701は、BTBエントリがスレッド間で共有されるか、共有可能であるかを示す。TID702はスレッドIDである。LE704は、レングス符号化を指し、第一BTBテーブル401などの第一BTBテーブルのLE605についてのもと同じ符号化を用いる。BRTYPE705は、分岐タイプ符号化を示す。いくつかの実施形態によれば、第二分岐タイプBTBは、無条件直接分岐のみなど、ただ1つの分岐タイプに関する情報を含む。このフィールドは、JMP、CALL、及び無効の間の区別を符号化するためのものである。いくつかの実施形態による第二分岐タイプのBTBについてのBRTYPEは、以下の、値0,0に関する状態が無効状態に対応し、値0,1に関する状態がJMP(ジャンプ)分岐タイプに対応し、値1,0に関する状態がCALL分岐タイプに対応し、値1,1に関する状態が別の不正状態に対応するように符号化される。第二BTBエントリ700についてのBREND706は、[5:0]の範囲に関して1インクリメントされる分岐エンドポイントである。BRTGT707は、第一BTBエントリ600についてのもと同じまたは同様のフォーマットに従った分岐ターゲットである。オーバーフローアドレスバッファに利用可能なアドレスの追加のビットを必要とする分岐は、図8のBTBエントリ800などでの第三分岐タイプのBTBに割り当てられる。したがって、1ビットのUSEOAフィールドは、アドレスの追加のビットについての必要性に対応するために、いくつかの実施形態による第二分岐タイプのBTBに含まれない。本明細書で言及されるように、OAは、オーバーフローアドレス、またはその中の特定のフィールド、ビット、もしくはフラグを指す。

【0043】

図8は、いくつかの実施形態による、第三BTBテーブルについての第三BTBエントリ800のブロック図である。以下は、図4の第三BTBテーブル403、または図5の第三BTBテーブル506内のBTBエントリなどについての、第三分岐タイプのBTBエントリ800内のエントリについてのフィールド及び位置ビットの記述である。各BTBエントリ800は、ゼロ(0)から71までインデックス付けされた72ビット幅であり、以下の、位置[71]における1ビットに関するUNUSED801、位置[70]における1ビットに関するAUXPREDEN802、位置[69]における1ビットに関するSHRD803、位置[68]における1ビットに関するTID804、位置[67:56]における12ビットに関するTAG805、位置[55:54]における2ビットに関するLE806、位置[53:52]における2ビットに関するBR1LBIAS807、位置[51]における1ビットに関するPAIR808、位置[50:46]における5ビットに関するBRTYPE809、位置[45:40]における6ビットに関するBR1END810、及び位置[39:38]における2ビットに関するBR0LBIAS811というフィールドを含む。フィールド{BREND812, BRTGT814}のグループは、位置[37:0]においてそれぞれの可変長範囲813、815にわたって分散される。例えば、BREND812は、(Y+1)から37に置かれており、BRTGT814は、0から(Y)に置かれており、ここでYは整数である。UNUSED801は、対応するマクロ位置に未使用のビットがあるかどうかを示す。AUXPREDEN802は、ループ出口予測器などの補助予測器の有効化を示す。SHRD803は、BTBエントリがスレッド間で共有されるか、共有可能であるかを示す。TID804はスレッドIDである。LE806は、レングス符号化を指し、第一BTBテーブル401などの第一BTBテーブルの同じLEフィールドについてのもの、及び図6のLE605に関して上述されるようなものと、同じ符号化を使用する。図8では、BR1LBIAS807は、Branch1についてローカルバイアスがあるかどうかを示す。PAIR808は、本明細書の他の箇所でさらに説明されるように使用される。いくつかの実施形態による第三分岐タイプのBTBエントリ800についてのBRTYPE809は、第一BTBエントリ600などの他のBTBテーブルに関するBRTYPEについて説明されるように符号化される。第三分岐タイプのBTBエントリ800についてのBR1EN

10

20

30

40

50

D 8 1 0 は、[5 : 0] の範囲に関して 1 インクリメントされる Branch 1 についての分岐エンドポイントである。B R O L B I A S 8 1 1 は、Branch 0 についてローカルバイアスがあるかどうかを示す。B R E N D 8 1 2 及び B R T G T 8 1 4 は、第一 B T B テーブル 4 0 1 などの第一 B T B テーブルからのフィールド U S E O A、O A I D X、B R O E N D、B R 1 T G T、及び B R O T G T に関連して、それぞれ、Branch 0 及び Branch 1 について 1 つまたは 2 つの分岐ターゲットを格納するために使用される。第一、第二、及び第三分岐タイプの B T B テーブル 4 0 1 ~ 4 0 3、及びそれぞれの B T B エントリ 6 0 0、7 0 0、8 0 0 を参照して記述されるビット（幅）及び位置の数は、一例に過ぎない。プロセッサまたはシステムに複数の B T B テーブルを実装する場合、記述されたフィールド、または他の変数もしくはフィールドについてのビット及び位置の他の数が可能である。

10

【 0 0 4 4 】

例示的な複数の B T B テーブル置き換えポリシーに関して、いくつかの可能な実施形態の 1 つとして以下が提供される。この実施形態では、第一のレベル 1 の B T B テーブルは、1 セットのエントリのうちの 1 つ以上がフルになるまで、すべての B T B エントリを最初に許容し、この実施例でのセットは、エントリのセット関連グループのセットを指す。別の実施例では、第一のレベル 1 の B T B テーブルは、レベル 1 の B T B テーブルがフルになるまで、すべての B T B エントリを最初に許容する。次に、プロセッサまたはプロセッサコアが動作すると、B T B エントリは、複数のレベル 2 の B T B テーブルのうちの 1 つにエビクトされ、これら複数のレベル 2 の B T B テーブルは、図 3 の B T B テーブル 3 0 1 ~ 3 0 3、及び図 4 の B T B テーブル 4 0 1 ~ 4 0 3 などの、ターゲットアドレスサイズまたは分岐タイプに基づいてそれぞれ設計される。例えば、置き換えポリシーは、図 3 の B T B テーブル 3 0 1 ~ 3 0 3、及び図 4 の B T B テーブル 4 0 1 ~ 4 0 3 などの各レベル 2 の B T B テーブルについてのフル L R U である。分岐タイプの B T B 置き換えポリシーは、最近最も使用されていない（L R U）か、最近最も使用された（M R U）かいずれかとしてエントリをマーク付けすることをサポートする。セットのウェイトは、ウェイトがクリアされた後に L R U としてマーク付けされることにより、次回、そのセットに新しい B T B エントリが書き込まれ、空いたエントリが置き換えられ、有効なエントリは上書きされない。ウェイトがインストールされた後、ウェイトは M R U としてマーク付けされることにより、ウェイトは置き換えのためにキューの最後に移動する。以下の表 1 は、レベル 1 がレベル 1 のメモリを指すケース、レベル 2 がレベル 2 のメモリを指すケースなどを詳述する。

20

30

40

50

【表 1】

イベント	新しい値	更新
レベル2のB TBはヒット する（レベル 1のB T Bは ミスする）	MRU	レベル1のB T Bピクティムに置き 換えられる場合、ヒットしたエント リを更新する
レベル2のB TBはヒット する（レベル 1のB T Bは ミスする）	MRU	レベル2のB T Bのヒット位置と異 なる場合、レベル1のB T Bピクテ ィムについてレベル2の位置を更新 する
レベル2のB TBはヒット する（レベル 1のB T Bは ミスする）	LRU	レベル1のB T Bピクティムに置き 換えられていない場合、ヒットした エントリを更新する（その上、レベ ル2のB T Bにおいてヒットしたエ ントリを無効にする）
レベル1のB TB及びレベ ル2のB T B はヒットする	LRU	レベル1のB T B及びレベル2のB T Bにおいてヒットした場合、レベ ル2の位置を更新する（その上、レ ベル2のB T Bにおいてヒットした エントリを無効にする）
チェックをト レーニングす る	MRU	エビクトされたレベル1のB T Bエ ントリについてレベル2のB T Bの 位置を更新する

10

20

【0045】

以下の記述は、いくつかの実施形態によるパイプライン操作に関連するさらなる詳細を提供する。分岐タイプのB T Bは、パワーフィルタによってキャンセルされない限り、予測フローごとに読み出される。読み出しは、第一位置B P 0で開始され、B P 4からB P N 1にリダイレクトされ、ここでNは後続のサイクルを指す。レベル1のB T Bはミスする/レベル2のB T Bはヒットするというスワップケースのパイプラインタイミングを以下の表2に示す。レベル1のB T Bからのピクティムウェイと、レベル2のB T Bからのヒットウェイとの両方が記録される。ピクティムウェイは、特別な配慮を必要とする、ヒットウェイとは異なるレベル2のB T Bテーブル及びバンクを含むことができる。例えば、レベル2のピクティムバンクは、LRU状態に基づいて選択され、上書きされる。レベル1のB T B、及びレベル2のB T Bへの書き込みは、実際の書き込みが位置B P 5にありながら、B P 4にアサートされる書き込みイネーブル(W r E n)信号によって発生する。

30

40

50

【表 2】

B P 0	B P 1	B P 2	B P 3	B P 4	B P 5
— —	— —	レベル1 のBTB はミスす る	レベル2 のBTB はヒット する	レベル1の BTBの WrEn	レベル1の BTBへの Wr
— —	— —	レベル1 のBTB ビクティ ム	— —	レベル2の BTBの WrEn	レベル2の BTBへの Wr

10

【0046】

以下の記述は、いくつかの実施形態による、図4のプロセッサ400などのプロセッサにおけるレベル2のBTB物理構成に関するさらなる詳細を提供する。第一のレベル2のBTBテーブルは、4ウェイに関する512セット、エントリあたり36ビット、セットあたり2個のマクロ、72bのマクロタイプ、及び合計8個のマクロを含む。第二のレベル2のBTBテーブルは、4ウェイに関する256セット、エントリあたり64ビット、セットあたり4個のマクロ、64bのマクロタイプ、及び合計8個のマクロを含む。第三のレベル2のBTBテーブルは、5ウェイに関する512セット、エントリあたり72ビット、セットあたり5個のマクロ、72bのマクロタイプ、及び合計20個のマクロを含む。

20

【0047】

インデックスの1つ以上の上位ビットは、読み出しイネーブルとして使用されるため、1つのマクロの半分だけが第一のレベル2のBTBテーブルについてのウェイあたりの読み出しであり、1つのマクロが第二及び第三のレベル2のBTBテーブルについてのウェイあたりの読み出しである。第一及び第三のレベル2のBTBテーブルが図4のレベル1のBTBテーブル401などのレベル1のBTBテーブルの2倍多いセットを含むことから、タグの下位ビットは、これらのBTBテーブルについてのインデックスの最上位ビット(MSB)として使用されるので、BTBテーブルは、12ビットのタグのみを格納する必要があるのである。

30

【0048】

比較的小型のクライアントデバイスまたは携帯電話用のプロセッサの実施形態では、複数の分岐タイプのBTBテーブルの使用は、パフォーマンスに比べて高い電力コストを有するため、消費電力を減らす追加の方法を実行することができる。例えば、静的な方法は、一部またはすべての時間についてのレベル2のBTBテーブルのパワーゲーティングと、一部またはすべての時間についてのレベル2のBTBテーブルのクロックゲーティングを含む。各ステップでは、BIOSまたはヒューズなどによってレベル2のBTBイネーブル機能を構成することができる。消費電力を減らす動的な方法は、特定のレベル2のBTBテーブルがいつ必要になるかを認識する追加の制御によって適応的にレベル2のBTBテーブルをパワーゲーティングすることを含む。別の動的な方法は、アプリケーションによる電力の好ましい使用に応じて、またはアプリケーションがアクティブであるオペレーティングシステム、ファームウェア、もしくはデバイスの電力設定に応じて、アプリケーションがレベル2のBTBテーブルのその使用を適応させるように、オペレーティングシステム、ファームウェアなどによって実行されるアプリケーションを適応させることを含む。

40

【0049】

以下の記述は、図3の第一BTBテーブル301、第二BTBテーブル302、及び第三BTBテーブル303、ならびに図4の対応する類似のテーブルなどの複数のBTBテ

50

ープルによる使用のためのオーバーフローアドレスバッファに関するさらなる詳細を提供する。予測構造にすべてのビットを格納する必要なしに、最大49ビットのターゲットサイズの分岐をサポートするために、オーバーフローアドレスバッファは、プロセッサに提供される。オーバーフローアドレスバッファを使用して、ある特定のビット数、例えば分岐ターゲットアドレスの先頭の、または1セットの17ビットなどを格納する。いくつかの実施形態によれば、BTBテーブルなどの予測器構造によって参照される先頭の17ビットがポインタを介してオーバーフローアドレスバッファにあることによって、BTBテーブル内のストレージスペースを節約する。オーバーフローアドレスバッファは、多数のターゲットアドレスビットを格納するために必要とされる場合に使用される。テストによれば、トレース分析は、所与のトレースについて、フェッチアドレスの先頭の17ビットが、1セットの命令内のさまざまな分岐に対して限られた数の値のみを取ることを示す。

10

【0050】

いくつかの実施形態によれば、オーバーフローアドレス(OA)バッファは、32個のエントリを含み、各エントリは、本明細書で説明されるように、位置[48:32]におけるBTBエントリ内の位置に仮想アドレスについての単一のフィールドを含む。各オーバーフローアドレスバッファのエントリは、スレッド間で共有される。予測時に、オーバーフローアドレスバッファは、BTBが提供する5ビットのポインタに基づいて、アドレスの先頭の17ビットをBTBに提供する。ポインタが予測器の1つに記録されたことから、予測時に使用されたオーバーフローアドレスバッファのエントリが上書きされた可能性がある。このような場合には、予測は、誤っており、リダイレクトされることにより、予測器が訂正させられる。異なる予測器は、予測パイプの異なるステージにおいてオーバーフローアドレスバッファを読み出すため、オーバーフローアドレスバッファからの複数の同時読み出しを必要とする。

20

【0051】

オーバーフローアドレスバッファは、BTBのトレーニング中に次のように割り振られる。分岐ターゲットが範囲[48:32]についての現在のアドレスと異なるかどうかをトレーニングが最初に判定する場合には、オーバーフローアドレスバッファが必要とされる。オーバーフローアドレスバッファが必要とされる場合、現在のオーバーフローアドレスバッファのエントリを、位置[48:32]におけるターゲットアドレスと比較して、オーバーフローアドレスバッファのエントリが位置[48:32]における所望のターゲットアドレスをすでに含むかどうかを判定する。マッチング時に、マッチングするオーバーフローアドレスバッファのインデックスは、予測器のBTBテーブルに記録される。オーバーフローアドレスバッファのエントリが位置[48:32]におけるターゲットにマッチングしない場合、フルLRU置き換えポリシーを使用して、オーバーフローアドレスバッファに新しいエントリが割り振られる。いくつかの実施形態によれば、書き込みをトレーニングすることは、投機的であり、いくつかのオーバーフローアドレスバッファのエントリは、不適切なパス上にありながら割り振られたために、全く有用ではないものが割り振られる場合がある。

30

【0052】

以下の記述は、BTBトレーニングに関連するさらなる詳細を提供する。予測がリダイレクトされる(予測パイプラインの復号及び実行ステージなどにおいて)ときに、BTBテーブルをトレーニングアルゴリズムによって更新し、この予測がプログラムシーケンス中に再度発生するときに、予測アドレスでの予測精度を向上させる。リダイレクトを受信するときに、リダイレクトソースからのリダイレクトタイプ、分岐ターゲット、及びEnd Addrを、予測履歴バッファからの予測分岐タイプ、分岐ターゲット、End Addr、及び他の状態とともに使用して、どの種類のトレーニング/更新アクションが必要とされるかを決定する。

40

【0053】

次のリダイレクトのケースは、1つのBTBテーブル、または1セットのBTBテーブルをトレーニングする必要がある。ケース1：成立と予測された分岐がいかなる分岐とも

50

マッチングしなかったことは、予測に使用された B T B エントリを無効にすることに対応する。ケース 2 : 予測された分岐がないことは、B T B で新しく見つかった分岐をトレーニングにすることに対応する。ケース 3 : リダイレクトされた分岐 E n d A d d r は、いずれかの予測された分岐の E n d A d d r にマッチングせず、B T B で新しく見つかった分岐をトレーニングにすることに対応する。ケース 4 : リダイレクトされた分岐 E n d A d d r は、予測された分岐の E n d A d d r にマッチングするが、T y p e にはマッチングせず、B T B で分岐タイプを更新することに対応する。この 4 番目のケースは、分岐を動的または可変としてマーク付けするために分岐タイプを変更する必要がある場合のケースを含む。

【 0 0 5 4 】

B T B トレーニングプロセスは、トレーニングパイプを介して行われる。各スレッドは、スレッドについて最後に確認されたリダイレクトに関する情報を捕捉し、B T B 更新が要求されているかどうかを示す誤予測レジスタに関連付けられる。トレーニングパイプは、要求に基づいて 2 つのスレッド間で選択し、両方のスレッドが B T B トレーニングを同時に要求するときに発生する要求競合の場合に、ラウンドロビンを介して調停する。一般に、一方のスレッドについてのトレーニングは、他方のスレッドがすでにトレーニングパイプにある間に開始することができる。ただし、他方のスレッドが同じ B T B インデックスについて現在トレーニングされており、この同じインデックスでの第二トレーニングによって可視である B T B テーブルを時間内に更新しない場合、スレッドは選択されることができない。B T B トレーニングは、投機的に起こり、誤ったパスでリダイレクトが発生することにより、B T B を更新させる可能性がある。ただし、誤予測レジスタが上書きされる場合、スレッドに対して進行中のトレーニングはキャンセルされる。

【 0 0 5 5 】

トレーニングパイプは、トレーニング要求が与えられるサイクル B T N 2 で開始し、サイクル B T N 1、サイクル B T 0、サイクル B T 1、サイクル B T 2、サイクル B T 3、サイクル B T 4 が続き、B T B が書き込まれるサイクル B T 5 で完了する。いくつかの実施形態によれば、図 4 の第一 B T B テーブル 4 0 1 などの第一 B T B テーブルのみが、新しいエントリ、または変更されたエントリによって更新される。トレーニングパイプは、1 サイクルの間、予測パイプをストールする t r a i n - c h e c k - r e a d (T C R) プロセスを開始する。T C R プロセスの 1 つの目的は、エントリがトレーニングルックアップ位置において現在、レベル 1 の B T B に存在するか、レベル 2 の B T B テーブルのうちの一つに存在するかを確認することである。マッチングがレベル 2 の B T B テーブルで見つかる場合、T C R プロセスは、レベル 2 の B T B と、レベル 1 の B T B との間にスワップをトリガした後に、トレーニングフローを繰り返す。

【 0 0 5 6 】

レベル 2 の B T B にヒットしない T C R プロセスに続いて、さまざまなトレーニングタスクが次のように処理される。無効化は、T C R がヒットする場合にのみ実行される。新しい書き込み及び更新について、B T B 更新トレーニングは、誤予測された分岐を B T B エントリヒットとマージする、新しい B T B エントリを作成する、またはこれら両方のアクションを実行する。さまざまなシナリオは次のとおりである。T C R プロセスがレベル 1 の B T B でミスするときに、置き換えポリシーに基づいてウェイを選択すること、及び該当する場合には選択されたエントリをレベル 2 の B T B にエビクトすることなどによって、ルックアップアドレスに新しい B T B エントリを作成する。いくつかの実施形態では、T C R プロセスは、本明細書に記載されるようにミスを考慮し、後続のステップを実行するときに、B T B テーブルの 3 つのレベルのうちの一つすべてのエントリをチェックする。

【 0 0 5 7 】

T C R プロセスが B T B テーブルでヒットするなど、レベル 1 の B T B でヒットするときに、ルックアップアドレスには既存の B T B エントリがある、すなわち、この B T B エントリにはすでに 1 つまたは 2 つの分岐がある。トレーニングタスクが更新である場合、

10

20

30

40

50

誤予測 `End Addr` は、見つかった分岐のうちの1つにマッチングする。それ以外は、新しい分岐を導入する必要がある。BTBに書き戻される必要がある、最大3つの分岐がある。これらの分岐は、`End Addr`に基づいて順序付けされ、`Pos 0`、`Pos 1`及び`Pos 2`とラベル付けされる。次のケースは、以下のとおりに処理される。ケース1は、`Pos 0`のみが有効である場合に発生する。次に、TCRルックアップアドレスにおいて、`Pos 0`分岐をトレーニングすることを実行する。ケース2は、`Pos 0`及び`Pos 1`のみが有効である場合に発生する。`Pos 0`及び`Pos 1`の分岐を対にすることができる場合、この対をトレーニングすることを、TCRルックアップアドレスで実行する。それ以外は、`Pos 1`が誤予測された分岐である場合、TCRルックアップアドレスでは`Pos 0`のみを、予測ウィンドウ内の「最後ではない」(`NLIPW`)としてトレーニングし、`Pos 1`に対して新しいトレーニングフローを`Pos 0 End Addr + 1`ルックアップアドレスで開始する。`Pos 1`が誤予測された分岐ではない場合、トレーニングはドロップされる。ケース3は、`Pos 0`、`Pos 1`及び`Pos 2`が有効である場合に発生する。`Pos 0`分岐及び`Pos 1`分岐を対にすることができる場合、この対をTCRルックアップアドレスでは`NLIPW`としてトレーニングし、`Pos 2`が誤予測された分岐である場合、`Pos 2`に対して新しいトレーニングフローを`Pos 1 End Addr + 1`ルックアップアドレスで開始する。`Pos 2`が誤予測された分岐ではない場合、トレーニングはドロップされる。それ以外は、`Pos 0`のみをTCRルックアップアドレスでは`NLIPW`としてトレーニングする。`Pos 0`が誤予測された分岐ではない場合、誤予測された分岐に対して新しいトレーニングフローを`Pos 0 End Addr + 1`ルックアップアドレスで開始する。ここでも、`Pos 1`、`Pos 2`、または`Pos 1`及び`Pos 2`の両方をドロップする。

【0058】

上記のトレーニングプロセスは、誤予測された分岐の前のすべての分岐を保持しようとし、誤予測された分岐をTCRフローによってインストールしようとする。各リダイレクトは、2つ以下のトレーニングフローをトリガするように制限されているため、誤予測された分岐をトレーニングすることが保証されない。ただし、エントリを`NLIPW`とマーク付けすることにより、次の予測を引き起こし、予測ウィンドウをいくつかの予測に分割するため、新しい分岐を最終的にトレーニングすることができる。各ルックアップアドレスでのTCRフローは、レベル1/レベル2のBTBスワップをトリガし、TCRフローをリプレイさせる。その結果、所与のリダイレクトに対して最大4つのTCRフローが発生することができる。

【0059】

図9は、いくつかの実施形態による、複数のBTBテーブルのうちの1つにターゲットアドレスサイズごとに分岐ターゲットアドレスを格納する方法900である。方法900は、図1に示される命令パイプラインアーキテクチャ100のいくつかの実施形態、図2に示される処理システム200のいくつかの実施形態、及び図3のプロセッサ300のいくつかの実施形態において実施される。図示された実施形態では、プロセッサ命令は、図3に示されるプロセッサ300内のプロセッサコア305などの処理システム内の1つ以上のプロセッサコア上で実行のために、行われているか、割り振られているか、スケジューリングされている。本明細書に考察されるように、プロセッサ命令は分岐命令を含む。

【0060】

ブロック901では、BTBトレーニングロジックは、ターゲットアドレスを分岐命令の予測アドレスと比較することによって、命令ターゲットアドレスサイズを決定する。この方法は、ターゲットアドレスサイズに基づいて、予測されたターゲットアドレスを格納する複数のサイズベースのBTBテーブルのうちの1つを決定する、または選択することを含む。例えば、ブロック902では、プロセッサは、ターゲットサイズが第一サイズ「サイズ1」以下であるかどうかを決定する。そうである場合、ブロック903では、プロセッサは、図3の第一BTBテーブル301などのサイズベースのBTBテーブルの第一サイズベースのBTBエントリにBTBエントリを格納する。そうでない場合、ブロック

10

20

30

40

50

904では、プロセッサは、ターゲットアドレスサイズが、3つのサイズのBTBシステムについての第二サイズ「サイズ2」以下であるかどうかを決定する。そうである場合、ブロック905では、プロセッサは、分岐ターゲットを第二サイズベースのBTBテーブルに格納する。例えば、ターゲットアドレスは、図3の第二BTBテーブル302に格納されている。ブロック904では、分岐命令タイプが第二サイズ「サイズ2」以下でない場合、ブロック906では、プロセッサは、第三サイズベースのBTBに分岐ターゲットを格納する。

【0061】

図10は、いくつかの実施形態による、複数のBTBテーブルのうちの1つに分岐タイプごとに分岐ターゲットアドレスを格納する方法を示す方法1000である。方法1000は、図1に示される命令パイプラインアーキテクチャ100のいくつかの実施形態、図2に示される処理システム200のいくつかの実施形態、図4のプロセッサ400、及び図5のプロセッサ500において実施される。図示された実施形態では、プロセッサ命令は、図4のプロセッサ400内のプロセッサコア405などの処理システム内の1つ以上のプロセッサコア上で実行のために、行われているか、割り振られているか、スケジューリングされている。本明細書に考察されるように、プロセッサ命令は分岐命令を含む。

10

【0062】

方法1000は、3つのBTBテーブルシステムの例示的なコンテキストに記述される。ブロック1001では、プロセッサは、条件命令についてのBTBエントリタイプを決定する。分岐命令タイプに基づいて、ブロック1002では、プロセッサは、分岐命令タイプがBTBテーブルシステムについて高い値のBTBエントリであるかどうかを決定する。そうである場合、ブロック1003では、プロセッサは、図4のBTBテーブル401などの第一BTBテーブルにBTBエントリを格納する。そうでない場合、ブロック1004では、プロセッサは、分岐命令タイプが中間の値のBTBエントリであるかどうかを決定する。そうである場合、ブロック1005では、プロセッサは、図4のBTBテーブル402などの第二BTBテーブルにBTBエントリを格納する。そうでない場合、ブロック1006では、プロセッサは、アドレスを含むBTBエントリを、図4のBTBテーブル403などの第三BTBテーブルのエントリに格納する。

20

【0063】

高い値、中間の値などに関して、ある特定のタイプである条件分岐命令は、高い値、中間の値、または他のタイプのBTBエントリであってもよい。例えば、ジャンプタイプは高い値のエントリであってもよく、戻りタイプの命令は中間の値のエントリであってもよい。いくつかの実施形態によれば、エントリの値は、エントリを保持しないことと比較して、またはBTBもしくはBTBテーブルにセーブされ得る他の可能なエントリと比較して、複数のプロセッササイクルをセーブするエントリに対応する。分岐命令が今後のプロセッササイクルに有用であるなどの価値がある場合、プロセッサは、分岐ターゲットをBTBテーブルに格納する。例えば、ターゲットアドレスは、図4のこのタイプのBTBテーブル401～403のうちの1つに格納されている。分岐命令タイプが有用な分岐タイプでない場合、プロセッサは、ターゲットアドレスをBTBテーブルのエントリに格納しない。

30

40

【0064】

本明細書に開示されるように、いくつかの実施形態では、プロセッサは、命令実行パイプラインと、第一分岐ターゲットバッファ(BTB)テーブルのエントリの第一数を含む第一BTBテーブルであって、各第一BTBテーブルのエントリは第一タグ及び第一ターゲットアドレスを含み、各第一ターゲットアドレスは第一幅を有する、第一BTBテーブルと、第二BTBテーブルのエントリの第二数を含む第二BTBテーブルであって、各第二BTBテーブルのエントリは第二タグ及び第二ターゲットアドレスを含み、各第二ターゲットアドレスは第一幅とは異なる第二幅を有する、第二BTBテーブルと、所与の予測アドレスに対して予測されたターゲットアドレスを提供するように構成される分岐予測器とを含み、プロセッサは、分岐命令の分岐特性に基づいて、第一BTBテーブルの第一B

50

T B テーブルエントリか、第二 B T B テーブルの第二 B T B テーブルエントリかいずれかに第一ターゲットアドレスを含む分岐記述子を格納するように構成される。一態様では、分岐特性は、分岐ターゲットアドレスサイズである。別の態様では、分岐特性は分岐タイプである。さらに別の態様では、第一 B T B テーブルのエントリは、ジャンプ分岐命令、呼び出し分岐命令、戻り分岐命令、及び条件分岐命令のうちの少なくとも1つに対して設定される。

【 0 0 6 5 】

一態様では、分岐特性は、B T B エントリに格納される分岐の数である。別の態様では、分岐特性は、B T B エントリのスレッド識別子である。さらに別の態様では、プロセッサは、第一メモリレベルキャッシュ及び第二メモリレベルキャッシュを含み、第一 B T B テーブル及び第二 B T B テーブルは、プロセッサの同じメモリレベルキャッシュに含まれる。さらに別の態様では、第一 B T B テーブルの第一 B T B テーブルエントリは、予測ルックアップに対してNウェイアソシアティブであり、第二 B T B テーブルの第二 B T B テーブルエントリは、予測ルックアップに対してMウェイアソシアティブであり、M及びNは1以上であり、MはNとは異なる。

10

【 0 0 6 6 】

別の態様では、第一 B T B テーブルエントリの第一数は、第二 B T B テーブルエントリの第二数とは異なる。さらに別の態様では、各 B T B テーブルエントリは、それぞれの B T B エントリを、セット内の最近最も使用された (M R U) ステータス、及び最近最も使用されていない (L R U) ステータスのうちの1つとしてマーク付けする、最近使用されたステータスビットを含み、プロセッサは、最近使用されたステータスビットの状態に基づく置き換えポリシーに従って、セット内の B T B エントリをエビクトするように設定される。

20

【 0 0 6 7 】

本明細書に開示されるように、いくつかの実施形態では、方法は、分岐命令に対する誤予測によって引き起こされるリダイレクトに基づいて、分岐ターゲットバッファ (B T B) エントリに分岐タイプを決定することと、B T B エントリの決定された分岐タイプに基づいて、プロセッサの第一分岐ターゲットバッファ (B T B) テーブルの第一エントリであって、この第一エントリは B T B エントリの第一分岐タイプに対応する第一ターゲットアドレス幅を有する、第一エントリ、及びプロセッサの第二 B T B テーブルの第二エントリであって、第二エントリは第一分岐タイプとは異なる第二分岐タイプに対応する、第一ターゲットアドレス幅とは異なる第二ターゲットアドレス幅を有する、第二エントリのうちの1つに B T B エントリを格納することを含む。一態様では、分岐タイプは、分岐ターゲットアドレスサイズである。別の態様では、B T B エントリの格納は、プロセッサの予測トレーナーユニットによって実行される。

30

【 0 0 6 8 】

一態様では、この方法は、セット内で最近最も使用された (M R U) ステータス、及び最近最も使用されていない (L R U) ステータスのうちの1つとして B T B エントリをマーク付けする最近使用されたステータスビットに基づいて第一 B T B テーブルまたは第二 B T B テーブルの B T B エントリを識別することと、そこにエントリを格納する前に、第一 B T B テーブルまたは第二 B T B テーブルの識別された B T B エントリをエビクトすることを含む。別の態様では、第一 B T B テーブル及び第二 B T B テーブルは、プロセッサの同じメモリレベルキャッシュに含まれる。さらに別の態様では、この方法は、B T B エントリを格納する前に、分岐命令の命令タグに基づいて、第一 B T B テーブル及び第二 B T B テーブルのうちの少なくとも1つを検索することと、B T B エントリを格納する前に、B T B エントリが第一 B T B テーブル及び第二 B T B テーブルのうちの少なくとも1つに見つからないことを識別することを含む。さらに別の態様では、この方法は、B T B エントリのターゲットアドレスのサイズが第一 B T B テーブルまたは第二 B T B テーブルのそれぞれの B T B エントリについてのアドレスサイズを上回る場合に、B T B エントリのターゲットアドレスのオーバーフロービットをオーバーフロー B T B テーブルに格納

40

50

することを含む。

【0069】

本明細書に開示されるように、いくつかの実施形態では、方法は、プロセッサのB T Bの複数の分岐ターゲットバッファ(B T B)テーブルをルックアップすることによって、現在予測されたブロックの終了アドレス、及び次の予測されたブロックの開始アドレスを予測することであって、B T Bは現在の予測アドレスでは第一B T Bテーブル及び第二B T Bテーブルを含む、予測することと、予測された終了アドレス、及び予測された開始アドレスのうちの少なくとも1つに基づいて、プロセッサでの実行のための命令の一部として予測を提供することとを備える。一態様では、第一B T BテーブルのB T Bエントリは、第二B T BテーブルのB T Bエントリのターゲットアドレスフィールドの第二幅よりも小さい第一幅のターゲットアドレスフィールドを含む。別の態様では、現在の予測アドレスは、プログラムカウンタに関連している。

10

【0070】

いくつかの実施形態では、前述の装置及び技法は、図1~7を参照して前述されたシステム、プロセッサ、及びB T Bテーブルなど、1つ以上の集積回路(IC)デバイス(集積回路パッケージまたはマイクロチップとも称される)を含むシステムに実装される。電子設計自動化(EDA)及びコンピュータ支援設計(CAD)ソフトウェアツールは、これらのICデバイスの設計及び製作で使用され得る。それらの設計ツールは典型的には、1つ以上のソフトウェアプログラムとして表される。1つ以上のソフトウェアプログラムは、1つ以上のICデバイスの回路を表すコード上で作動して、回路を製造するための製造システムを設計するかまたは適応させるためのプロセスの少なくとも一部を実行するようにコンピュータシステムを操作するためにコンピュータシステムによって実行可能なコードを含む。このコードは、命令、データ、または命令及びデータの組み合わせを含むことができる。設計ツールまたは製作ツールを表すソフトウェア命令は、典型的には、コンピューティングシステムにアクセス可能なコンピュータ可読記憶媒体に格納される。同様に、ICデバイスの設計または製造の1つ以上の段階を表すコードは、同じコンピュータ可読記憶媒体または異なるコンピュータ可読記憶媒体に格納され、そこからアクセスされ得る。

20

【0071】

コンピュータ可読記憶媒体は、命令及び/またはデータをコンピュータシステムに提供するために、使用中にコンピュータシステムによりアクセス可能な任意の非一時的記憶媒体、または非一時的記憶媒体の組み合わせを含み得る。このような記憶媒体には、光学媒体(例えばコンパクトディスク(CD)、デジタル多用途ディスク(DVD)、ブルーレイディスク)、磁気媒体(例えばフロッピーディスク、磁気テープ、または磁器ハードドライブ)、揮発性メモリ(例えばランダムアクセスメモリ(RAM)またはキャッシュ)、不揮発性メモリ(例えば読み出し専用メモリ(ROM)またはフラッシュメモリ)、または微小電気機械システム(MEMS)ベース記憶媒体が含まれ得るが、これらに限定されない。コンピュータ可読記憶媒体は、コンピューティングシステムに組み込まれてもよく(例えば、システムRAMまたはROM)、コンピューティングシステムに固定して取り付けられてもよく(例えば、磁気ハードドライブ)、コンピューティングシステムに取り外し可能に取り付けられてもよく(例えば、光学ディスクまたはユニバーサルシリアルバス(USB)ベースフラッシュメモリ)、または有線もしくは無線ネットワークを介してコンピュータシステムに結合されてもよい(例えば、ネットワークアクセス可能ストレージ(NAS))。

30

40

【0072】

いくつかの実施形態において、上述される技法の特定の態様は、ソフトウェアを実行する処理システムの1つ以上のプロセッサにより実施され得る。ソフトウェアは、非一時的なコンピュータ可読記憶媒体上に、格納される、またはその他の方法により有形に具現化される、実行可能命令の1つ以上のセットを含む。ソフトウェアが1つ以上のプロセッサによって実行されるときに、1つ以上のプロセッサを操作して、上記の技法の1つ以上の

50

態様を実行する命令及び特定のデータを、ソフトウェアは含むことができる。非一時的なコンピュータ可読記憶媒体には、例えば、磁気ディスクまたは光ディスクの記憶デバイス、フラッシュメモリ、キャッシュ、ランダムアクセスメモリ（RAM）などのソリッドステート記憶デバイス、またはその他の単一または複数の不揮発性メモリデバイスなどが含まれ得る。非一時的なコンピュータ可読記憶媒体上に格納される実行可能な命令は、1つ以上のプロセッサにより、解釈される、またはその他の方法により実行可能である、ソースコード、アセンブリ言語コード、オブジェクトコード、または他の命令フォーマットにあってもよい。

【0073】

一般的な記述の中で上述されるすべてのアクティビティまたは要素が必要なわけではなく、特定のアクティビティまたはデバイスの一部が必要ではない場合があり、記述されるそれらに加えて、1つ以上のさらなるアクティビティが実行され得る、または1つ以上のさらなる要素が含まれてもよい。さらに、アクティビティがリストにされている順序は、必ずしもそれらが実行される順序ではない。また、概念は、特定の実施形態を参照して記載された。しかしながら、当業者は、下記の特許請求の範囲に記載されるように本開示の範囲から逸脱することなく、さまざまな変更及び変形を行うことが可能であることを理解する。したがって、本明細書及び図面は、限定的な意味ではなく例示的な意味で考えられるべきであり、すべてのこれらの変更形態は、本開示の範囲内に含まれることが意図される。

10

【0074】

利益、他の利点、及び問題に対する解決策を、具体的な実施形態に関して、上記にて説明してきた。しかしながら、課題に対する利点、長所、解決策、及び何らかの利点、長所または解決策を生じさせ得る、もしくは顕著にし得る特徴（複数可）は、いずれかまたはすべての請求項の、重要であるか、必要であるか、または本質的特徴として解釈されるというわけではない。さらに、開示された発明の主題が本明細書に教示の利益を有する当業者へ明らかである、異なるが均等な方式において変更され、実施され得るように、上記に開示される特定の実施形態は、例示に過ぎない。下記の特許請求の範囲に記載される以外の、本明細書に示される構成または設計の詳細への制限を意図しない。したがって、上記に開示される特定の実施形態が変更、または修正されてもよく、すべてのこれらのような変形形態が開示された発明の主題の範囲内にあるとみなされる。その結果、本明細書に求められる保護は、下記の特許請求の範囲内に記載されるようなものである。

20

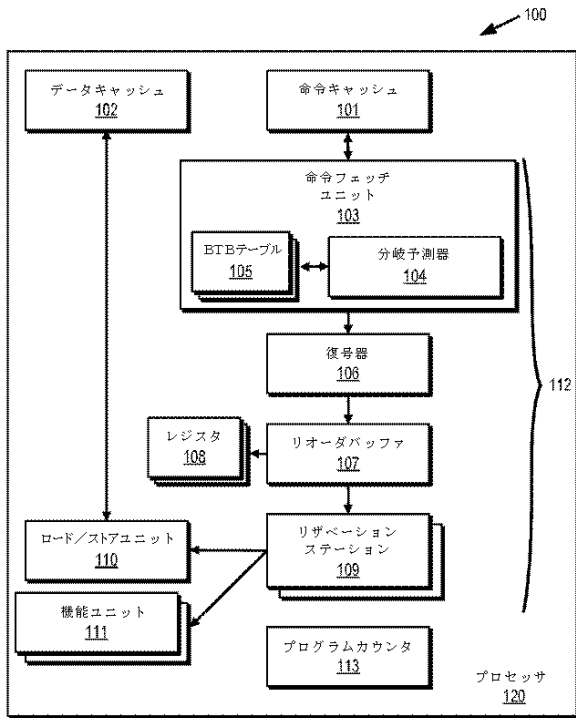
30

40

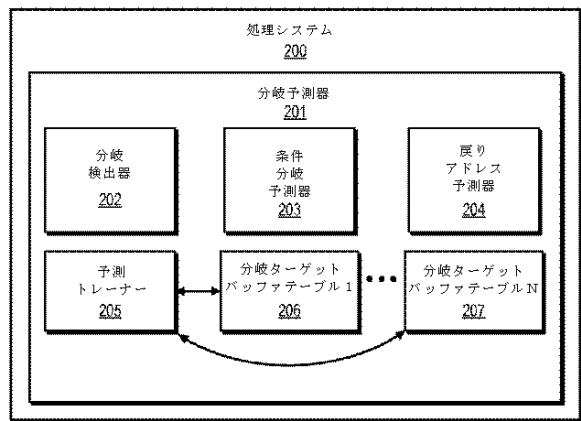
50

【図面】

【図 1】



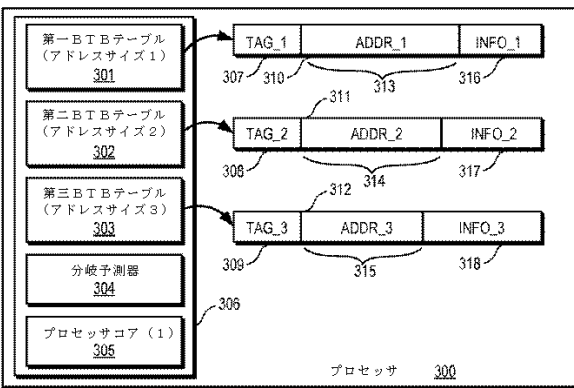
【図 2】



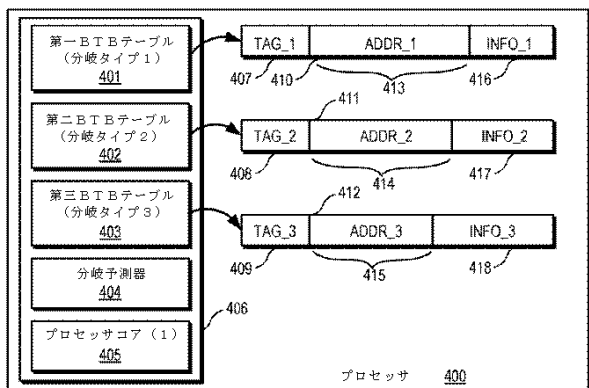
10

20

【図 3】



【図 4】

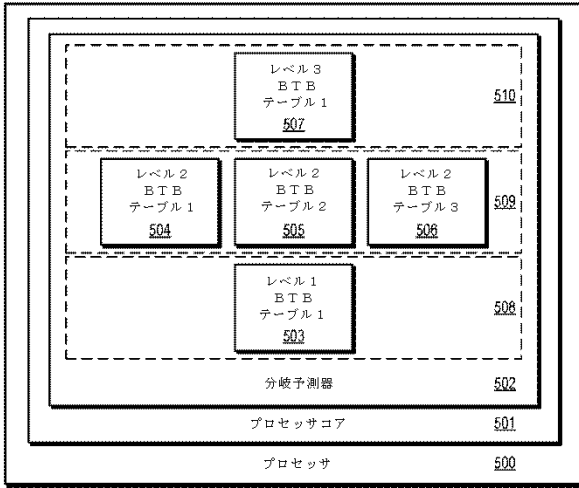


30

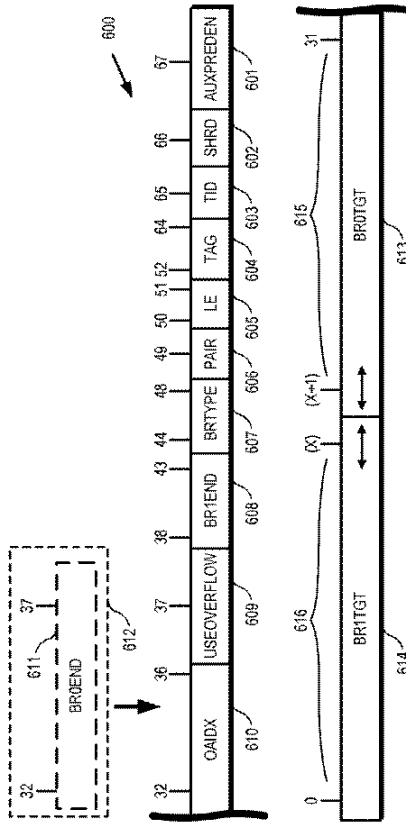
40

50

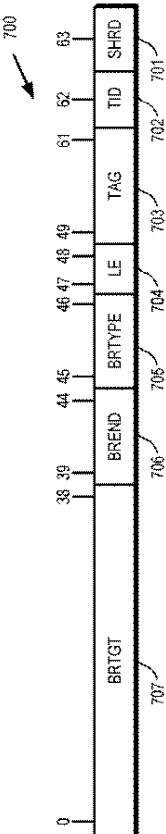
【図 5】



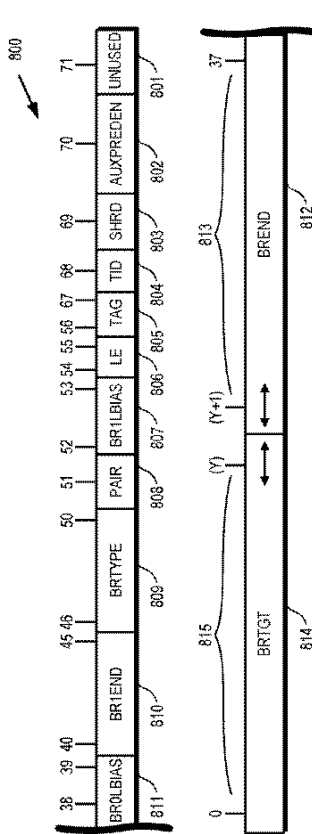
【図 6】



【図 7】



【図 8】



10

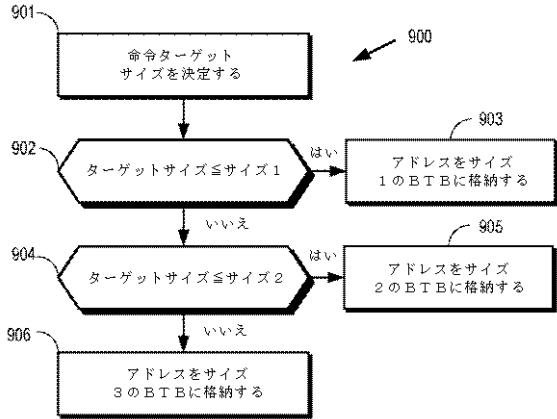
20

30

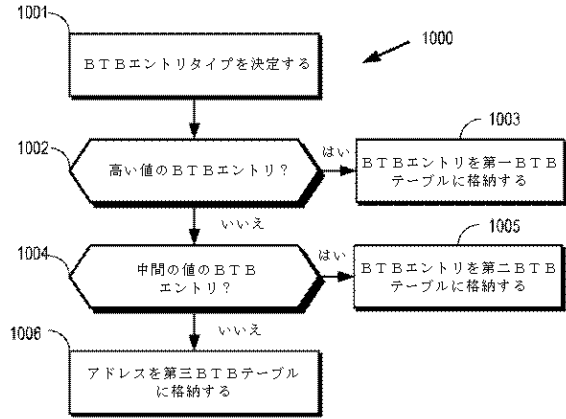
40

50

【図 9】



【図 10】



10

20

30

40

50

フロントページの続き

- アメリカ合衆国 9 5 0 5 4 カリフォルニア州、サンタ クララ、オーガスティン ドライブ 2 4 8 5
(72)発明者 アンソニー ジャービス
アメリカ合衆国 9 5 0 5 4 カリフォルニア州、サンタ クララ、オーガスティン ドライブ 2 4 8 5
審査官 坂庭 剛史
(56)参考文献 特表 2 0 0 8 - 5 4 2 9 1 7 (J P , A)
米国特許出願公開第 2 0 1 1 / 0 0 9 3 6 5 8 (U S , A 1)
米国特許出願公開第 2 0 1 2 / 0 3 2 4 2 0 9 (U S , A 1)
山田祐司、小林良太郎、安藤秀樹、島田俊夫、2 レベル表構成の導入による分岐先バッファの容量削減、並列処理シンポジウム JSPP '99, 日本, 社団法人情報処理学会, 1999年06月09日, Vol.99, No.6, pp.103-110, ISSN1344-0640
(58)調査した分野 (Int.Cl., D B 名)
G 0 6 F 9 / 3 8
G 0 6 F 1 2 / 0 8 7 5
G 0 6 F 1 2 / 0 8 9 5