

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6738579号
(P6738579)

(45) 発行日 令和2年8月12日(2020.8.12)

(24) 登録日 令和2年7月22日(2020.7.22)

(51) Int. Cl.	F I					
G06F 7/57	(2006.01)	G06F	7/57	202		
G06F 9/30	(2018.01)	G06F	9/30	372		
G06F 9/38	(2006.01)	G06F	9/30	310A		
G06F 9/302	(2006.01)	G06F	9/38	370C		
G06F 9/318	(2006.01)	G06F	9/302	A		
請求項の数 18 (全 41 頁) 最終頁に続く						

(21) 出願番号	特願2017-527720 (P2017-527720)	(73) 特許権者	591003943 インテル・コーポレーション
(86) (22) 出願日	平成27年11月23日(2015.11.23)		アメリカ合衆国 95054 カリフォル ニア州・サンタクララ・ミッション カレ ッジ ブレーバード・2200
(65) 公表番号	特表2018-507453 (P2018-507453A)	(74) 代理人	110000877 龍華国際特許業務法人
(43) 公表日	平成30年3月15日(2018.3.15)	(72) 発明者	コーベル、ジーザス
(86) 国際出願番号	PCT/US2015/062056		アメリカ合衆国 95054 カリフォル ニア州・サンタクララ・ミッション カレ ッジ ブレーバード・2200 インテル ・コーポレーション内
(87) 国際公開番号	W02016/105754		
(87) 国際公開日	平成28年6月30日(2016.6.30)		
審査請求日	平成30年11月19日(2018.11.19)		
(31) 優先権主張番号	14/581, 815		
(32) 優先日	平成26年12月23日(2014.12.23)		
(33) 優先権主張国・地域又は機関	米国 (US)		
最終頁に続く			

(54) 【発明の名称】 命令フローを最適化するチェックを実行するための装置および方法

(57) 【特許請求の範囲】

【請求項1】

1 または複数のソースオペランドを用いて複数の数学的命令を実行するための演算論理装置 (ALU) と、

現在の数学的命令のための前記1 または複数のソースオペランドを評価し、前記評価に基づいて、前記ALUによって前記現在の数学的命令を実行することを含むデフォルト演算シーケンスを実行するか否か、または特定のタイプのソースオペランドを有する前記数学的命令の結果を、前記デフォルト演算シーケンスより効率的に提供する代替的な演算シーケンスにジャンプするか否かを判断するための命令チェックロジックとを備え、

前記1 または複数のソースオペランドの前記評価を実行すると、前記命令チェックロジックは、前記デフォルト演算シーケンスまたは前記代替的なシーケンスを実行するか否かを示すベクトル出力と、前記代替的なシーケンスで処理されるべき要素をシグナリングするマスク出力と、前記1 または複数のソースオペランドに対する数学的オペレーションの実行から生じた1 または複数の例外を示すべく出力された計算毎の例外フラグとを生成する

プロセッサ。

【請求項2】

前記現在の数学的命令は、前記1 または複数のソースオペランドにより指定された分子および分母を有する除算命令を含み、

前記命令チェックロジックは、非正規化オペランドであるか、無限大に等しいか、非数

(NaN)オペランドであるか、およびゼロによる除算をもたらすかのうちの少なくとも1つである分子または分母のいずれかに応答して、前記代替的な演算シーケンスへのジャンプを生じさせる、請求項1に記載のプロセッサ。

【請求項3】

前記命令チェックロジックは、ゼロ(ZE)による除算、無効な演算(IE)、および非正規化オペランド(DE)のうちの少なくとも1つを含む、1または複数の例外フラグを前記プロセッサ内で設定する、請求項2に記載のプロセッサ。

【請求項4】

前記現在の命令は、平方根演算を実行するソースオペランド値を有する平方根命令を含み、

前記命令チェックロジックは、負の数であるか、非正規化オペランドであるか、無限大に等しいか、および非数(NaN)オペランドであるかのうちの少なくとも1つである前記ソースオペランドに応答して、前記代替的な演算シーケンスへのジャンプを生じさせる、請求項1～3のいずれか1項に記載のプロセッサ。

【請求項5】

前記現在の命令は、平方根演算を実行するソースオペランド値(x)を有する平方根命令を含み、

前記命令チェックロジックは、xを用いて実行される比較に応答して前記代替的な演算シーケンスへのジャンプを生じさせる、請求項1～4のいずれか1項に記載のプロセッサ。

【請求項6】

前記デフォルト演算シーケンスは、デフォルトシーケンスの命令またはマイクロオペレーションを含み、

前記代替的な演算シーケンスは、代替的なシーケンスの命令またはマイクロオペレーションを含む、請求項1～5のいずれか1項に記載のプロセッサ。

【請求項7】

現在の数学的命令のための1または複数のソースオペランドを取得する段階と、

前記現在の数学的命令のための前記1または複数のソースオペランドを評価する段階と、

前記評価に基づいて、前記現在の数学的命令を実行することを含むデフォルト演算シーケンスを実行するか否か、または特定のタイプのソースオペランドを有する前記数学的命令の結果を、前記デフォルト演算シーケンスより効率的に提供する代替的な演算シーケンスにジャンプするか否かを判断する段階と、

前記1または複数のソースオペランドの前記評価を実行すると、前記デフォルト演算シーケンスまたは前記代替的なシーケンスを実行するか否かを示すベクトル出力を生成する段階と、

前記代替的なシーケンスで処理されるべき要素をシグナリングするマスク出力を生成する段階と、

前記1または複数のソースオペランドに対する数学的オペレーションの実行から生じた1または複数の例外を示すべく出力された計算毎の例外フラグを生成する段階と、を備える、方法。

【請求項8】

前記現在の数学的命令は、前記1または複数のソースオペランドにより指定された分子および分母を有する除算命令を含み、

前記方法は、非正規化オペランドであるか、無限大に等しいか、非数(NaN)オペランドであるか、およびゼロによる除算をもたらすかのうちの少なくとも1つである分子または分母のいずれかに応答して、前記代替的な演算シーケンスにジャンプする、請求項7に記載の方法。

【請求項9】

ゼロ(ZE)による除算、無効な演算(IE)、および非正規化オペランド(DE)の

10

20

30

40

50

うちの少なくとも1つを含む、1または複数の例外フラグを前記方法において設定する段階を更に備える、請求項8に記載の方法。

【請求項10】

前記現在の命令は、平方根演算を実行するソースオペランド値を有する平方根命令を含み、

前記方法は、負の数であるか、非正規化オペランドであるか、無限大に等しいか、および非数(NaN)オペランドであるかのうちの少なくとも1つである前記ソースオペランドにตอบสนองして、前記代替的な演算シーケンスにジャンプする段階を更に備える、請求項7~9のいずれか1項に記載の方法。

【請求項11】

前記現在の命令は、平方根演算を実行するソースオペランド値(x)を有する平方根命令を含み、

前記方法は、xを用いて実行される比較にตอบสนองして前記代替的な演算シーケンスにジャンプする段階を更に備える、請求項7~10のいずれか1項に記載の方法。

【請求項12】

前記デフォルト演算シーケンスは、デフォルトシーケンスの命令またはマイクロオペレーションを含み、

前記代替的な演算シーケンスは、代替的なシーケンスの命令またはマイクロオペレーションを含む、請求項7~11のいずれか1項に記載の方法。

【請求項13】

数学的命令およびグラフィックス命令を含む命令およびデータを格納するためのメモリと、

前記数学的命令を実行して前記データを処理するための複数のコアと、

前記グラフィックス命令にตอบสนองしてグラフィックスオペレーションを実行するためのグラフィックスプロセッサユニットと、

ネットワークを介してデータを受信および送信するためのネットワークインタフェースと、

マウスまたはカーソル制御デバイスからユーザ入力を受信するためのインタフェースと、

1または複数のソースオペランドを用いて複数の数学的命令を実行するための演算論理装置(ALU)と、

現在の数学的命令のための前記1または複数のソースオペランドを評価し、前記評価に基づいて、前記ALUにより前記現在の数学的命令を実行することを含むデフォルト演算シーケンスを実行するか否か、または特定のタイプのソースオペランドを有する前記数学的命令についての結果を、前記デフォルト演算シーケンスより効率的に提供する代替的な演算シーケンスにジャンプするか否かを判断する命令チェックロジックとを備え、

前記1または複数のソースオペランドの前記評価を実行すると、前記命令チェックロジックは、前記デフォルト演算シーケンスまたは前記代替的なシーケンスを実行するか否かを示すベクトル出力と、前記代替的なシーケンスで処理されるべき要素をシグナリングするマスク出力と、前記1または複数のソースオペランドに対する数学的オペレーションの実行から生じた1または複数の例外を示すべく出力された計算毎の例外フラグとを生成する

システム。

【請求項14】

前記現在の数学的命令は、前記1または複数のソースオペランドにより指定された分子および分母を有する除算命令を含み、

前記命令チェックロジックは、非正規化オペランドであるか、無限大に等しいか、非数(NaN)オペランドであるか、およびゼロによる除算をもたらすかのうちの少なくとも1つである分子または分母のいずれかにตอบสนองして、前記代替的な演算シーケンスへのジャンプを生じさせる、請求項13に記載のシステム。

10

20

30

40

50

【請求項 15】

前記命令チェックロジックは、ゼロ (Z E) による除算、無効な演算 (I E)、および非正規化オペランド (D E) のうちの少なくとも1つを含む、1または複数の例外フラグをプロセッサ内で設定する、請求項 14 に記載のシステム。

【請求項 16】

前記現在の命令は、平方根演算を実行するソースオペランド値を有する平方根命令を含み、

前記命令チェックロジックは、負の数であるか、非正規化オペランドであるか、無限大に等しいか、および非数 (N a N) オペランドであるかのうちの少なくとも1つである前記ソースオペランドに回答して、前記代替的な演算シーケンスへのジャンプを生じさせる、請求項 13 ~ 15 のいずれか1項に記載のシステム。

10

【請求項 17】

前記現在の命令は、平方根演算を実行するソースオペランド値 (x) を有する平方根命令を含み、

前記命令チェックロジックは、xを用いて実行される比較に回答して前記代替的な演算シーケンスへのジャンプを生じさせる、請求項 13 ~ 16 のいずれか1項に記載のシステム。

【請求項 18】

前記命令チェックロジックは、マイクロオペレーションを実行するためのものである
請求項 1 ~ 6 のいずれか1項に記載のプロセッサまたは請求項 13 ~ 17 のいずれか1
項に記載のシステム。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、概ねコンピュータプロセッサの分野に関する。より具体的には、本発明は、命令フローを最適化するチェックを実行するための方法および装置に関する。

【背景技術】

【0002】

命令セットまたは命令セットアーキテクチャ (I S A) は、ネイティブなデータタイプ、命令、レジスタアーキテクチャ、アドレス指定モード、メモリアーキテクチャ、割り込みおよび例外処理、ならびに外部入出力 (I / O) を含むプログラミングに関連するコンピュータアーキテクチャの一部である。本明細書において「命令」という用語は、一般に、マクロ命令をデコードするプロセッサデコーダの結果であるマイクロ命令またはマイクロオブとは異なり、実行のためにプロセッサに提供される命令であるマクロ命令を指すことに留意されたい。マイクロ命令またはマイクロオブは、プロセッサ上の実行ユニットにマクロ命令に関連するロジックを実装するためのオペレーションの実行を命令するように構成され得る。

30

【0003】

I S A は、命令セットを実装するべく用いられるプロセッサ設計技術のセットであるマイクロアーキテクチャとは区別される。異なるマイクロアーキテクチャを有するプロセッサは、共通の命令セットを共有し得る。例えば、インテル (登録商標) P E N T I U M (登録商標) 4 プロセッサ、インテル (登録商標) C o r e (商標) プロセッサ、およびカリフォルニア州サンノゼにある A d v a n c e d M i c r o D e v i c e s , I n c のプロセッサは、ほぼ同一のバージョンの x 8 6 命令セット (より新しいバージョンに追加されたいいくつかの拡張を伴う) を実装するが、異なる内部設計を有する。例えば、I S A の同一のレジスタアーキテクチャは、専用物理レジスタ、レジスタリネームメカニズム (例えば、レジスタエイリアステーブル (R A T)、リオーダバッファ (R O B)、およびリタイアメントレジスタファイルの使用) 等を用いる1または複数の動的に割り当てられる物理レジスタを含む周知技術を用いて、異なる態様で異なるマイクロアーキテクチャに実装され得る。別途指定されない限り、レジスタアーキテクチャ、レジスタファイ

40

50

ル、およびレジスタという文言は、本明細書においてソフトウェア/プログラマに可視であるもの、および命令がレジスタを指定する態様に言及するために用いられる。区別が必要とされる場合、「論理的」、「アーキテクチャの」、または「可視のソフトウェア」という形容詞は、レジスタアーキテクチャにおけるレジスタ/ファイルを示すために用いられるが、異なる形容詞が所与のマイクロアーキテクチャ（例えば、物理レジスタ、リオーダーバッファ、リタイアメントレジスタ、レジスタプール）におけるレジスタを指すために用いられる。

【0004】

命令セットは、1または複数の命令フォーマットを含む。所与の命令フォーマットは、とりわけ、実行されるオペレーションおよび当該オペレーションが実行されるオペランドを指定する様々なフィールド（ビットの数、ビットの位置）を定義する。いくつかの命令フォーマットは、命令テンプレート（またはサブフォーマット）の定義により更に分類される。例えば、所与の命令フォーマットの命令テンプレートは、異なるサブセットの命令フォーマットのフィールドを有するものと定義され（含まれるフィールドは、通常は同じ順序であるが、少なくともいくつかは、より少ないフィールドが含まれているので、異なるビット位置を有する）、および/または異なる解釈をされる所与のフィールドを有するものと定義され得る。所与の命令は、所与の命令フォーマットを用いて（および定義されている場合には、当該命令フォーマットの命令テンプレートのうちの所与の1つで）表され、オペレーションおよびオペランドを指定する。命令ストリームは、命令の特定のシーケンスであり、シーケンスにおける各命令は、命令フォーマットにおける命令の発生である（定義される場合、当該命令フォーマットの命令テンプレートのうちの所与の1つ）。

【図面の簡単な説明】

【0005】

本発明のより良い理解は、以下の図面と併せて以下の詳細な説明から得られ得る。

【0006】

【図1A】本発明の実施形態による一般的ベクトル向け命令フォーマットおよびその命令テンプレートを示すブロック図である。

【図1B】本発明の実施形態による一般的ベクトル向け命令フォーマットおよびその命令テンプレートを示すブロック図である。

【0007】

【図2A】本発明の実施形態による、例示的な特定ベクトル向け命令フォーマットを示すブロック図である。

【図2B】本発明の実施形態による、例示的な特定ベクトル向け命令フォーマットを示すブロック図である。

【図2C】本発明の実施形態による、例示的な特定ベクトル向け命令フォーマットを示すブロック図である。

【図2D】本発明の実施形態による、例示的な特定ベクトル向け命令フォーマットを示すブロック図である。

【0008】

【図3】本発明の一実施形態によるレジスタアーキテクチャのブロック図である。

【0009】

【図4A】本発明の実施形態による、例示的なインオーダーフェッチ、デコード、リタイアパイプライン、および例示的なレジスタリネーム、アウトオブオーダー発行/実行パイプラインの両方を示すブロック図である。

【0010】

【図4B】本発明の実施形態による、プロセッサに含まれるべきインオーダーフェッチ、デコード、リタイアコアの例示的な実施形態、および例示的なレジスタリネーム、アウトオブオーダー発行/実行アーキテクチャコアの両方を示すブロック図である。

【0011】

【図5A】オンダイ相互接続ネットワークとの接続を伴うシングルプロセッサコアのプロ

10

20

30

40

50

ック図である。

【0012】

【図5B】本発明の実施形態による図5Aのプロセッサコアの一部の拡大図を示す。

【0013】

【図6】本発明の実施形態による統合メモリコントローラおよびグラフィックスを用いるシングルコアプロセッサおよびマルチコアプロセッサのブロック図である。

【0014】

【図7】本発明の一実施形態によるシステムのブロック図を示す。

【0015】

【図8】本発明の実施形態による第2のシステムのブロック図を示す。

10

【0016】

【図9】本発明の実施形態による第3のシステムのブロック図を示す。

【0017】

【図10】本発明の実施形態によるシステムオンチップ(SoC)のブロック図を示す。

【0018】

【図11】本発明の実施形態による、ソース命令セットのバイナリ命令を、ターゲット命令セットのバイナリ命令に変換するソフトウェア命令コンバータの使用と対比するブロック図を示す。

【0019】

【図12】本発明の実施形態が実装され得る例示的なプロセッサを示す。

20

【0020】

【図13】命令の入力をチェックし、これに応じて次の命令シーケンスを決定するのに使用できる出力のセットを生成するための命令チェックロジックを示す。

【0021】

【図14】異なるタイプの入力値に対する除算チェック命令に対して生成される出力を示す。

【0022】

【図15】異なるタイプの入力値のための平方根チェック命令に対して生成される出力を示す。

【0023】

30

【図16】チェックオペレーションを実行し、第1の命令シーケンスまたは第2の命令シーケンスを選択するための方法の一実施形態を示す。

【発明を実施するための形態】

【0024】

以下の説明において、説明の目的のために、下記の本発明の実施形態の完全な理解を提供するべく、多数の具体的な詳細が記載される。しかし、当業者には、本発明の実施形態がこれらの具体的な詳細のいくつかを用いることなく実施され得ることが明らかであろう。本発明の実施形態の基礎となる原理を不明瞭にするのを避けるべく、他の例において、周知の構造およびデバイスがブロック図の形態で示される。

【0025】

40

例示的なプロセッサアーキテクチャおよびデータタイプ

命令セットは、1または複数の命令フォーマットを含む。所与の命令フォーマットは、とりわけ、実行されるオペレーション(オペコード)および当該オペレーションが実行されるオペランドを指定する様々なフィールド(ビットの数、ビットの位置)を定義する。いくつかの命令フォーマットは、命令テンプレート(またはサブフォーマット)の定義により更に分類される。例えば、所与の命令フォーマットの命令テンプレートは、異なるサブセットの命令フォーマットのフィールド(含まれるフィールドは、通常は同じ順序であるが、少なくともいくつかは、より少ないフィールドが含まれているので、異なるビット位置を有する)を有するものと定義され、および/または異なる解釈をされる所与のフィールドを有するものと定義され得る。従って、ISAの各命令は、所与の命令フォーマット

50

トを用いて（および定義される場合には、当該命令フォーマットの命令テンプレートのうちの所与の1つで）表され、オペレーションおよびオペランドを指定するためのフィールドを含む。例えば、例示的なADD命令は、特定のオペコード、ならびに当該オペコードを指定するオペコードフィールドおよびオペランド（ソース1/デスティネーション、およびソース2）を選択するオペランドフィールドを含む命令フォーマットを有する。命令ストリームにおけるこのADD命令が生じることにより、特定のオペランドを選択するオペランドフィールドに特定の内容を有する。Advanced Vector Extensions (AVX) (AVX1およびAVX2)と呼ばれ、ベクトル拡張(VEX)符号化スキームを用いるSIMD拡張のセットが、リリースおよび/または公開されている（例えば、Intel（登録商標）64 and IA-32 Architecture Software Developers Manual, October 2011およびIntel（登録商標）Advanced Vector Extensions Programming Reference, June 2011を参照されたい）。

10

【0026】

例示的な命令フォーマット

本明細書に説明される命令の実施形態は、異なるフォーマットで実施され得る。更に、例示的なシステム、アーキテクチャ、およびパイプラインが以下に詳述される。命令の実施形態は、そのようなシステム、アーキテクチャ、およびパイプライン上で実行され得るが、詳述されるものに限定されない。

20

【0027】

A. 一般的ベクトル向け命令フォーマット

ベクトル向け命令フォーマットは、ベクトル命令に好適な命令フォーマットである。（例えば、ベクトルオペレーションに固有の一定のフィールドが存在する）。ベクトルおよびスカラーオペレーションの両方がベクトル向け命令フォーマットによりサポートされる実施形態が説明されるが、代替的な実施形態は、ベクトル向け命令フォーマットによるベクトルオペレーションのみを用いる。

【0028】

図1A～図1Bは、本発明の実施形態による一般的ベクトル向け命令フォーマットおよびその命令テンプレートを示すブロック図である。図1Aは、本発明の実施形態による、一般的ベクトル向け命令フォーマット、およびそのクラスA命令テンプレートを示すブロック図である。図1Bは、本発明の実施形態による、一般的ベクトル向け命令フォーマット、およびそのクラスB命令テンプレートを示すブロック図である。具体的には、クラスAおよびクラスB命令テンプレートは、一般的ベクトル向け命令フォーマット100に対して定義され、これらの両方は、非メモリアクセス105の命令テンプレートおよびメモリアクセス120の命令テンプレートを含む。ベクトル向け命令フォーマットの文脈における一般的という用語は、いずれの特定の命令セットにも関係しない命令フォーマットを指す。

30

【0029】

ベクトル向け命令フォーマットが、32ビット（4バイト）または64ビット（8バイト）のデータ要素幅（またはサイズ）を有する64バイトのベクトルオペランド長（またはサイズ）（従って、64バイトのベクトルは、16個のダブルワードサイズの要素、または代替的に8クワッドワードサイズの要素のいずれかからなる）、16ビット（2バイト）または8ビット（1バイト）のデータ要素幅（またはサイズ）を有する64バイトのベクトルオペランド長（またはサイズ）、32ビット（4バイト）、64ビット（8バイト）、16ビット（2バイト）、または8ビット（1バイト）のデータ要素幅（またはサイズ）を有する32バイトのベクトルオペランド長（またはサイズ）、および32ビット（4バイト）、64ビット（8バイト）、16ビット（2バイト）、または8ビット（1バイト）のデータ要素幅（またはサイズ）を有する16バイトのベクトルオペランド長（またはサイズ）をサポートする本発明の実施形態が説明されるが、代替的な実施形態は、

40

50

より多い、より少ない、または異なるデータ要素幅（例えば、128ビット（16バイト）のデータ要素幅）を有するより多い、より少ない、および/または異なるベクトルオペランドサイズ（例えば、256バイトのベクトルオペランド）をサポートし得る。

【0030】

図1AにおけるクラスA命令テンプレートは、1)非メモリアクセス105の命令テンプレート中に示される非メモリアクセス、フルラウンド制御タイプオペレーション110の命令テンプレート、および非メモリアクセス、データ変換タイプオペレーション115の命令テンプレート、ならびに2)メモリアクセス120の命令テンプレート中に示されるメモリアクセス、一時的125の命令テンプレート、およびメモリアクセス、非一時的130の命令テンプレートを含む。図1BのクラスB命令テンプレートは、1)非メモリアクセス105の命令テンプレート中に示される非メモリアクセス、ライトマスク制御、部分的ラウンド制御タイプオペレーション112の命令テンプレート、および非メモリアクセス、ライトマスク制御、VSIZETYPタイプオペレーション117の命令テンプレート、ならびに2)メモリアクセス120の命令テンプレート中に示されるメモリアクセス、ライトマスク制御127の命令テンプレートを含む。

10

【0031】

一般的ベクトル向け命令フォーマット100は、図1A～図1Bにおいて示される順序で以下に列挙される、次のフィールドを含む。

【0032】

フォーマットフィールド140。このフィールドにおける特定の値（命令フォーマット識別子の値）は、ベクトル向け命令フォーマット、従って、命令ストリーム中のベクトル向け命令フォーマットにおける命令の発生を一意に識別する。従って、このフィールドは、一般的ベクトル向け命令フォーマットのみを有する命令セットに必要とされないという意味で任意選択である。

20

【0033】

ベースオペレーションフィールド142。その内容は、異なるベースオペレーションを区別する。

【0034】

レジスタインデックスフィールド144。その内容は、レジスタ内であれ、メモリ内であれ、直接に、またはアドレス生成により、ソースオペランドおよびデスティネーションオペランドの位置を指定する。これらは、 $P \times Q$ （例えば、 32×512 、 16×128 、 32×1024 、 64×1024 ）のレジスタファイルからN個のレジスタを選択するのに十分な数のビットを含む。一実施形態においては、Nは、最大で3つのソースおよび1つのデスティネーションレジスタであり得るが、代替的な実施形態は、より多いかまたはより少ないソースおよびデスティネーションレジスタをサポートし得る（例えば、最大で2つのソースをサポートし得、この場合、これらのソースのうちの1つは、デスティネーションとしても機能し、最大で3つのソースをサポートし得、この場合、これらのソースのうちの1つは、デスティネーションとしても機能し、最大で2つのソースおよび1つのデスティネーションをサポートし得る）。

30

【0035】

修飾子フィールド146。その内容は、一般的ベクトル命令フォーマット中におけるメモリアクセスを指定する命令の発生とメモリアクセスを指定しない命令の発生とを、すなわち、非メモリアクセス105の命令テンプレートとメモリアクセス120の命令テンプレートを区別する。メモリアクセスオペレーションは、メモリ階層を読み出し、および/またはこれに書き込むが（いくつかの場合には、レジスタにおける値を用いて、ソースアドレスおよび/またはデスティネーションアドレスを指定する）、非メモリアクセスオペレーションは、これを行わない（例えば、ソースおよびデスティネーションは、レジスタである）。また、一実施形態において、このフィールドは、3つの異なる態様を選択し、メモリアドレス計算を実行するが、代替的な実施形態は、より多い、より少ない、または異なる態様をサポートし、メモリアドレス計算を実行し得る。

40

50

【0036】

追加オペレーションフィールド150。その内容は、ベースオペレーションに加えて、種々様々な異なるオペレーションのうちどれが実行されるかを区別する。このフィールドは、コンテキストに固有である。本発明の一実施形態において、このフィールドは、クラスフィールド168、アルファフィールド152、およびベータフィールド154に分割される。追加オペレーションフィールド150は、オペレーションの共通グループが2、3、または4つの命令ではなく、単一の命令で実行されることを可能にする。

【0037】

スケールフィールド160。その内容は、メモリアドレス生成のための（例えば、 $2^{\text{scale}} * \text{index} + \text{base}$ を用いるアドレス生成のための）インデックスフィールド内容のスケールリングを可能にする。

10

【0038】

変位フィールド162A。その内容は、メモリアドレス生成の一部として（例えば、 $2^{\text{scale}} * \text{index} + \text{base} + \text{displacement}$ を用いるアドレス生成に）用いられる。

【0039】

変位係数フィールド162B（変位係数フィールド162Bの直ぐ上に変位フィールド162Aを並置することにより、一方または他方が使用されることを示すことに留意されたい）。その内容は、アドレス生成の一部として用いられる。変位係数フィールド162Bは、メモリアクセス(N)のサイズに対して調整される変位係数を指定する。Nは、（例えば、 $2^{\text{scale}} * \text{index} + \text{base} + \text{scaled displacement}$ を用いるアドレス生成のための）メモリアクセスにおけるバイトの数である。冗長下位ビットは、無視され、従って、変位係数フィールドの内容は、有効なアドレスを計算するとき使用される最終的変位を生成するべく、メモリオペランドの合計サイズ(N)で乗算される。Nの値は、フルオペコードフィールド174（本明細書において後述される）およびデータ操作フィールド154Cに基づいて、ランタイムでプロセッサハードウェアにより決定される。変位フィールド162Aおよび変位係数フィールド162Bは、非メモリアクセス105の命令テンプレートに用いられず、および/または異なる実施形態が1つのみまたは2つのうちいずれも実装しない場合があるという意味で任意選択である。

20

【0040】

データ要素幅フィールド164。その内容は、（いくつかの実施形態において全ての命令に対して、他の実施形態において命令のうちいくつかのみに対して）いくつかのデータ要素幅のうちどれが用いられるかを区別する。このフィールドは、1つのデータ要素幅のみがサポートされ、および/またはオペコードのいくつかの態様を用いて、データ要素幅がサポートされる場合には必要とされないという意味で任意選択である。

30

【0041】

ライトマスクフィールド170。その内容は、データ要素位置ベースで、デスティネーションベクトルオペランドにおける当該データ要素位置がベースオペレーションおよび追加オペレーションの結果を反映するか否かを制御する。クラスA命令テンプレートは、マーキング・ライトマスキングをサポートするが、クラスB命令テンプレートは、マーキングおよびゼロ書き込みマスキングの両方をサポートする。マーキングする場合、ベクトルマスクは、デスティネーションにおける要素の任意のセットが（ベースオペレーションおよび追加オペレーションにより指定された）任意のオペレーションの実行中に更新から保護されることを可能にする。他の一実施形態では、対応するマスクビットが0を有するデスティネーションの各要素の古い値を保持する。対照的に、ゼロ書き込みする場合、ベクトルマスクは、デスティネーションにおける要素の任意のセットが（ベースオペレーションおよび追加オペレーションにより指定された）任意のオペレーションの実行中にゼロにされることを可能にする。一実施形態において、対応するマスクビットが0の値を有する場合、デスティネーションの要素は、0に設定される。この機能のサブセットは、実行されるオペレーションのベクトル長を制御する能力である（すなわち、要素のスパンは、最

40

50

初のものから最後のものに変更される)。しかし、変更される要素が連続している必要はない。従って、ライトマスクフィールド170は、ロード、ストア、演算、論理等を含む部分的ベクトルオペレーションを可能にする。本発明の実施形態は、ライトマスクフィールド170の内容が、用いられるべきライトマスクを含むいくつかのライトマスクレジスタのうちの一つを選択する(従って、ライトマスクフィールド170の内容が実行されるべき当該マスクングを間接的に識別する)ものとして説明されているが、代替的な実施形態はこれに代えて、または更に、マスクライトフィールド170の内容が実行されるべきマスクングを直接に指定することを可能にする。

【0042】

即値フィールド172。その内容は、即値の指定を可能にする。このフィールドは、即値をサポートしない一般的なベクトル向けフォーマットの実装において存在せず、即値を用いない命令中に存在しないという意味で任意選択である。

10

【0043】

クラスフィールド168。その内容は、命令の異なるクラスを区別する。図1A~図1Bを参照すると、このフィールドの内容は、クラスA命令またはクラスB命令を選択する。図1A~図1Bにおいて、角が丸い四角形は、特定の値がフィールド内に存在することを示すべく用いられる(例えば、図1Aおよび図1Bにおけるクラスフィールド168のクラスA168AおよびクラスB168Bの各々)。

【0044】

クラスAの命令テンプレート

20

クラスAの非メモリアクセス105の命令テンプレートの場合に、アルファフィールド152は、RSフィールド152Aとして解釈され、その内容は、異なる追加オペレーションタイプのうちのどれが実行されるかを区別するが(例えば、ラウンド152A.1およびデータ変換152A.2は、各々、非メモリアクセス、ラウンドタイプオペレーション110、および非メモリアクセス、データ変換タイプオペレーション115の命令テンプレートに対して指定される)、ベータフィールド154は、指定されたタイプのオペレーションのうちいずれが実行されるかを区別する。非メモリアクセス105の命令テンプレートにおいて、スケールフィールド160、変位フィールド162A、および変位スケールフィールド162Bは、存在しない。

【0045】

30

非メモリアクセス命令テンプレート フルラウンド制御タイプオペレーション 非メモリアクセスのフルラウンド制御タイプオペレーション110の命令テンプレートにおいて、ベータフィールド154は、ラウンド制御フィールド154Aとして解釈され、その内容は、静的ラウンドを提供する。本発明の説明される実施形態において、ラウンド制御フィールド154Aは、抑圧全浮動小数点例外(SAE)フィールド156およびラウンドオペレーション制御フィールド158を含み、代替的な実施形態は、これら両方のコンセプトをサポートおよびエンコードして同一のフィールドとすることができ、またはこれらのコンセプト/フィールドのうち一方または他方のみを有し得る(例えば、ラウンドオペレーション制御フィールド158のみを有し得る)。

【0046】

40

SAEフィールド156。その内容は、例外イベント報告を無効にするか否かを区別する。SAEフィールド156の内容が、抑圧が有効にされたことを示す場合、所与の命令は、いずれの種類の浮動小数点例外フラグも報告せず、いずれの浮動小数点例外ハンドラも立ち上げない。

【0047】

ラウンドオペレーション制御フィールド158。その内容は、ラウンドオペレーション(例えば、ラウンドアップ、ラウンドダウン、ゼロへのラウンド、および近似値へのラウンド)のグループのうちどれが実行されるべきかを区別する。このように、ラウンドオペレーション制御フィールド158は、命令ベースでラウンドモードの変更を可能にする。プロセッサがラウンドモードを指定する制御レジスタを含む本発明の一実施形態におい

50

て、ラウンドオペレーション制御フィールド150の内容は、当該レジスタの値を上書きする。

【0048】

非メモリアクセス命令テンプレート データ変換タイプオペレーション 非メモリアクセスのデータ変換タイプオペレーション115の命令テンプレートにおいて、ベータフィールド154は、データ変換フィールド154Bとして解釈され、その内容は、いくつかのデータ変換のうちのどれが実行されるかを区別する（例えば、非データ変換、スウィズル、ブロードキャスト）。

【0049】

クラスAのメモリアクセス120の命令テンプレートの場合、アルファフィールド152は、エビクションヒントフィールド152Bとして解釈され、その内容は、エビクションヒントのうちのどれが用いられるかを区別するが（図1Aにおいて、一時的152B.1および非一時的152B.2は、各々、メモリアクセス、一時的125の命令テンプレート、およびメモリアクセス、非一時的130の命令テンプレートに対して指定される）、ベータフィールド154は、データ操作フィールド154Cとして解釈され、その内容は、いくつかのデータ操作オペレーション（プリミティブとしても知られる）のうちのどれが実行されるかを区別する（例えば、操作なし、ブロードキャスト、ソースのアップコンバージョン、およびデスティネーションのダウンコンバージョン）。メモリアクセス120の命令テンプレートは、スケールフィールド160を含み、変位フィールド162Aまたは変位スケールフィールド162Bを任意選択で含む。

【0050】

ベクトルメモリ命令は、変換サポートを用いて、メモリからのベクトルロードおよびメモリへのベクトルストアを実行する。通常のベクトル命令の場合のように、ベクトルメモリ命令は、データ要素の様式で、メモリから/にデータを転送し、実際に転送される要素は、ライトマスクとして選択されるベクトルマスクの内容により規定される。

【0051】

メモリアクセス命令テンプレート 一時的 一時的データは、キャッシュから利益を得るのに十分なほど速やかに再利用される可能性が高いデータである。しかし、これはヒントであり、異なるプロセッサは、ヒントを完全に無視することを含め、異なる態様で一時的データを実装してもよい。

【0052】

メモリアクセス命令テンプレート 非一時的 非一時的データは、レベル1のキャッシュにおけるキャッシュから利益を得るのに十分なほど速やかに再利用される可能性が低いデータであり、追い出しの優先権を与えられるべきである。しかし、これはヒントであり、異なるプロセッサは、ヒントを完全に無視することを含め、異なる態様で非一時的データを実装してもよい。

【0053】

クラスBの命令テンプレート クラスBの命令テンプレートの場合、アルファフィールド152は、ライトマスク制御（Z）フィールド152Cとして解釈され、その内容は、ライトマスクフィールド170により制御されるライトマスキングがマーキングであるべきか、またはゼロ書き込みであるべきかを区別する。

【0054】

クラスBの非メモリアクセス105の命令テンプレートの場合、ベータフィールド154の一部は、RLフィールド157Aとして解釈され、その内容は、異なる追加オペレーションタイプのどれが実行されるかを区別するが（例えば、ラウンド157A.1およびベクトル長（VSIZE）157A.2は各々、非メモリアクセス、ライトマスク制御、部分的ラウンド制御タイプオペレーション112の命令テンプレート、および非メモリアクセス、ライトマスク制御、VSIZEタイプオペレーション117の命令テンプレートに対して指定される）、ベータフィールド154の残りは、指定されたタイプのどのオペレーションが実行されるかを区別する。非メモリアクセス105の命令テンプレートにお

10

20

30

40

50

いて、スケールフィールド160、変位フィールド162A、および変位スケールフィールド162Bは、存在しない。

【0055】

非メモリアクセス、ライトマスク制御、部分的ラウンド制御タイプオペレーション110の命令テンプレートにおいて、ベータフィールド154の残りは、ラウンドオペレーションフィールド159Aとして解釈され、例外イベント報告は、無効にされる（所与の命令は、いずれの種類の浮動小数点例外フラグも報告せず、いずれの浮動小数点例外ハンドラも立ち上げない）。

【0056】

ラウンドオペレーション制御フィールド159A。ちょうどラウンドオペレーション制御フィールド158のように、その内容は、ラウンドオペレーション（例えば、ラウンドアップ、ラウンドダウン、ゼロへのラウンド、および近似値へのラウンド）のグループのうちどれが実行されるべきかを区別する。このように、ラウンドオペレーション制御フィールド159Aは、命令ベースでラウンドモードの変更を可能にする。プロセッサがラウンドモードを指定する制御レジスタを含む本発明の一実施形態において、ラウンドオペレーション制御フィールド150の内容は、そのレジスタの値を上書きする。

10

【0057】

非メモリアクセス、ライトマスク制御、VSIZEタイプオペレーション117の命令テンプレートにおいて、ベータフィールド154の残りは、ベクトル長フィールド159Bとして解釈され、その内容は、いくつかのデータベクトル長のうちどれが実行されるかを区別する（例えば、128、256、または512バイト）。

20

【0058】

クラスBのメモリアクセス120の命令テンプレートの場合、ベータフィールド154の一部は、ブロードキャストフィールド157Bとして解釈され、その内容は、ブロードキャストタイプのデータ操作オペレーションが実行されるべきか否かを区別するが、ベータフィールド154の残りは、ベクトル長フィールド159Bとして解釈される。メモリアクセス120の命令テンプレートは、スケールフィールド160を含み、変位フィールド162Aまたは変位スケールフィールド162Bを任意選択で含む。

【0059】

一般的ベクトル向け命令フォーマット100に関連して、フォーマットフィールド140、ベースオペレーションフィールド142、およびデータ要素幅フィールド164を含む、フルオペコードフィールド174が示される。一実施形態として、フルオペコードフィールド174がこれらのフィールドの全てを含むものが示されているが、これら全てをサポートしない実施形態において、フルオペコードフィールド174は、これらのフィールド全てより少ないものを含む。フルオペコードフィールド174は、オペレーションコード（オペコード）を提供する。

30

【0060】

追加オペレーションフィールド150、データ要素幅フィールド164、およびライトマスクフィールド170は、一般的ベクトル向け命令フォーマットにおける命令ベースで、これらの機能が指定されることを可能にする。

40

【0061】

ライトマスクフィールドおよびデータ要素幅フィールドを組み合わせることで、マスクが異なるデータ要素幅に基づいて適用されることを可能にするように、型付き命令を生成する。

【0062】

クラスAおよびクラスB内に見出される様々な命令テンプレートは、異なる状況において有益である。本発明のいくつかの実施形態において、異なるプロセッサまたはプロセッサ内の異なるコアは、クラスAのみ、クラスBのみ、または両方のクラスをサポートし得る。例えば、汎用演算用の高性能汎用アウトオブオーダーコアは、クラスBのみをサポートし得、主にグラフィックスおよび/またはサイエンティフィック（スループット）演算用

50

のコアは、クラスAのみをサポートし得、両方用のコアは、両方をサポートし得る（勿論、両方のクラスの全てのテンプレートおよび命令ではないが、両方のクラスのテンプレートおよび命令のいくつかのミックスを有するコアは、本発明の範囲内である）。また、シングルプロセッサは、複数のコアを含み得、これらの全てが同じクラスをサポートし、または異なるコアが異なるクラスをサポートする。例えば、別個のグラフィックスコアおよび汎用コアを有するプロセッサにおいて、主にグラフィックスおよび/またはサイエンティフィック演算用のグラフィックスコアのうちの1つは、クラスAのみをサポートし得るが、汎用コアのうちの1つまたは複数は、クラスBのみをサポートする汎用演算用のアウトオブオーダー実行およびレジスタリネームを用いる高性能汎用コアであり得る。別個のグラフィックスコアを有しない別のプロセッサは、クラスAおよびクラスBの両方をサポートする、もう1つの汎用インオーダーまたはアウトオブオーダーのコアを含み得る。勿論、本発明の異なる実施形態において、あるクラスの機能は、他のクラスにおいても実装され得る。高水準言語で書かれたプログラムは、1) 実行のためにターゲットプロセッサによりサポートされるクラスの命令のみを有する形式、または2) 全てのクラスの命令の異なる組み合わせを用いて書かれた代替的なルーチンを有し、現在、コードを実行しているプロセッサによりサポートされる命令に基づいて、実行するルーチンを選択する制御フローコードを有する形式を含む、種々様々な実行可能な形式にされる（例えば、実行時コンパイルまたは静的コンパイル）。

10

【0063】

B. 例示的な特定ベクトル向け命令フォーマット

20

図2A～図2Dは、本発明の実施形態による例示的な特定ベクトル向け命令フォーマットを示すブロック図である。図2A～図2Dは、フィールドの位置、サイズ、解釈、および順序、ならびにそれらのフィールドのいくつかに対する値を指定するという意味で具体的な特定ベクトル向け命令フォーマット200を示す。特定ベクトル向け命令フォーマット200は、x86命令セットを拡張するために用いられ得、従ってフィールドのうちのいくつかは、既存のx86命令セットおよびその拡張（例えば、AVX）において用いられるものと類似するか、または同じである。このフォーマットは、拡張された既存のx86命令セットのプレフィックスエンコードフィールド、リアルオペコードバイトフィールド、MOD R/Mフィールド、SIBフィールド、変位フィールド、および即値フィールドとの整合性を保つ。図2A～図2Dがフィールドにマッピングされる図1A～図1Bのフィールドが示される。

30

【0064】

本発明の実施形態は、例示的目的で一般的ベクトル向け命令フォーマット100の文脈において、特定ベクトル向け命令フォーマット200を参照して説明されるが、本発明は、特許請求される場合を除き、特定ベクトル向け命令フォーマット200に限定されないことを理解されたい。例えば、一般的ベクトル向け命令フォーマット100は、様々なフィールドのために考えられる様々なサイズを企図するが、特定ベクトル向け命令フォーマット200は、特定サイズのフィールドを有するものとして示される。具体的な例として、データ要素幅フィールド164は、特定ベクトル向け命令フォーマット200における1つのビットフィールドとして示されるが、本発明は、そのようなには限定されない（すなわち、一般的ベクトル向け命令フォーマット100は、他のサイズのデータ要素幅フィールド164を企図する）。

40

【0065】

一般的ベクトル向け命令フォーマット100は、図2Aにおいて示される順序で以下に列挙される、次のフィールドを含む。

【0066】

EVEXPレフィックス（バイト0～3）202。4バイト形式でエンコードされる。

【0067】

フォーマットフィールド140（EVEXPバイト0、ビット[7:0]）。最初のバイト（EVEXPバイト0）は、フォーマットフィールド140であり、0x62を含む（本

50

発明の一実施形態において、ベクトル向け命令フォーマットを区別するべく用いられる一意な値)。

【0068】

第2～第4のバイト(EVEXバイト1～3)は、特定の能力を提供するいくつかのビットフィールドを含む。

【0069】

REXフィールド205(EVEXバイト1、ビット[7-5])は、EVEX.Rビットフィールド(EVEXバイト1、ビット[7]-R)、EVEX.Xビットフィールド(EVEXバイト1、ビット[6]-X)、およびEVEXバイト1、ビット[5]-B)からなる。EVEX.R、EVEX.X、およびEVEX.Bのビットフィールドは、対応するVEXビットフィールドと同一の機能性を提供し、1の補数形式を用いてエンコードされる。すなわち、ZMM0は、1111Bとしてエンコードされ、ZMM15は、0000Bとしてエンコードされる。命令の他のフィールドは、当技術分野で既知のレジスタインデックスの下位の3つのビット(rrr、xxx、およびbbb)をエンコードし、従って、Rrrr、xxxx、およびBbbbは、EVEX.R、EVEX.X、およびEVEX.Bを加えることにより形成され得る。

【0070】

REX'フィールド110。これはREX'フィールド110の第1の部分であり、拡張された32個のレジスタセットの上位の16個または下位の16個のいずれかをエンコードするために用いられるEVEX.R'ビットフィールド(EVEXバイト1、ビット[4]-R')である。本発明の一実施形態において、このビットは、以下に示される他のものと共に、(周知のx86の32ビットモードで)BOUND命令から区別するべく、ビット反転フォーマットで格納される。BOUND命令のリアルオペコードバイトは62であるが、MOD R/Mフィールド(下記)にはMODフィールドにおける11の値を受け付けない。本発明の代替的な実施形態は、これおよび以下に示される他のビットを反転フォーマットで格納しない。1の値は、下位の16個のレジスタをエンコードするべく用いられる。換言すると、R'Rrrrは、EVEX.R'、EVEX.R、および他のフィールドの他のRRRを組み合わせることにより形成される。

【0071】

オペコードマップフィールド215(EVEXバイト1、ビット[3:0]-mmmm)。その内容は、暗黙に示される先頭オペコードバイト(0F、0F38、または0F3)をエンコードする。

【0072】

データ要素幅フィールド164(EVEXバイト2、ビット[7]-W)は、EVEX.Wという表記により表される。EVEX.Wは、データタイプの粒度(サイズ)(32ビットのデータ要素または64ビットのデータ要素のいずれか)を定義するべく用いられる。

【0073】

EVEX.vvvv220(EVEXバイト2、ビット[6:3]-vvvv)。EVEX.vvvvの役割は、以下を含み得る。1)EVEX.vvvvは、反転(1の補数)形式で指定された第1のソースレジスタオペランドをエンコードし、2またはそれより多いソースオペランドを用いる命令に対して有効である。2)EVEX.vvvvは、一定の複数のベクトルシフトに対して1の補数形式で指定されたデスティネーションレジスタオペランドをエンコードする。または3)EVEX.vvvvは、いずれのオペランドもエンコードせず、フィールドは予約され、1111bを含むはずである。従って、EVEX.vvvvフィールド220は、反転(1の補数)形式で格納された第1のソースレジスタ指定子の4つの下位ビットをエンコードする。命令に応じて、指定子サイズを32個のレジスタに拡張するべく、追加の異なるEVEXビットフィールドが用いられる。

【0074】

EVEX.U168クラスフィールド(EVEXバイト2、ビット[2]-U)。EV

10

20

30

40

50

EVEX.U = 0である場合、クラスAまたはEVEX.U0を示す。EVEX.U = 1である場合、クラスBまたはEVEX.U1を示す。

【0075】

プレフィックスエンコードフィールド225 (EVEXバイト2、ビット[1:0] - pp)は、ベースオペレーションフィールドのために追加ビットを提供する。EVEXプレフィックスフォーマットでレガシSSE命令のためのサポートを提供することに加えて、これは、SIMDプレフィックスを圧縮するという利益も有する (EVEXプレフィックスは、SIMDプレフィックスを表すバイトを必要とするのではなく、2ビットのみを必要とする)。一実施形態において、レガシフォーマットおよびEVEXプレフィックスフォーマットの両方でSIMDプレフィックス (66H、F2H、F3H)を用いるレガシSSE命令をサポートするべく、これらのレガシSIMDプレフィックスは、SIMDプレフィックスエンコードフィールドへとエンコードされ、(PLAが変更なしにレガシフォーマットおよびこれらのレガシ命令のEVEXフォーマットの両方を実行し得るように)ランタイム時にデコーダのPLAに提供される前にレガシSIMDプレフィックスへと拡張される。より新しい命令は、EVEXプレフィックスエンコードフィールドの内容をオペコード拡張として直接に用い得るが、ある実施形態は、整合のために同様に拡張されるが、これらのレガシSIMDプレフィックスにより指定される異なる意味を可能にする。代替的な実施形態は、2ビットのSIMDプレフィックスエンコードをサポートするようにPLAを再設計し得、従って拡張を必要としない。

10

【0076】

アルファフィールド152 (EVEXバイト3、ビット[7] - EH。EVEX.EH、EVEX.rs、EVEX.RL、EVEX.ライトマスク制御、およびEVEX.Nとしても知られる。でも示される)。上記のように、このフィールドは、コンテキストに固有である。

20

【0077】

ベータフィールド154 (EVEXバイト3、ビット[6:4] - SSS、EVEX.s₂₋₀、EVEX.r₂₋₀、EVEX.r_{r1}、EVEX.LL0、EVEX.LLBとしても知られる。でも示される)。上記のように、このフィールドは、コンテキストに固有である。

【0078】

REX'フィールド110。これは、REX'フィールドの残りであり、拡張された32個のレジスタセットの上位の16個または下位の16個をエンコードするべく用いられるEVEX.V'ビットフィールドのいずれかである (EVEXバイト3、ビット[3] - V')。このビットは、ビット反転フォーマットで格納される。1の値は、下位の16個のレジスタをエンコードするべく用いられる。換言すると、V'VVVVは、EVEX.V'、EVEX.vvvvを組み合わせることにより形成される。

30

【0079】

ライトマスクフィールド170 (EVEXバイト3、ビット[2:0] - kkk)。その内容は、上記のライトマスクレジスタにおけるレジスタのインデックスを指定する。本発明の一実施形態において、特定の値EVEX.kkk = 000は、特定の命令のために非ライトマスクが用いられることを暗に示す、特別な動作を有する (これは、全ての1に対するハードワイヤされたライトマスクの使用、またはマスキングハードウェアを迂回するハードウェアの使用を含む、様々な態様で実装され得る)。

40

【0080】

リアルオペコードフィールド230 (バイト4)は、オペコードバイトとしても知られる。オペコードの一部は、このフィールド内に指定される。

【0081】

MOD R/Mフィールド240 (バイト5)は、MODフィールド242、Regフィールド244、およびR/Mフィールド246を含む。上記のように、MODフィールド242の内容は、メモリアクセスオペレーションおよび非メモリアクセスオペレーショ

50

ンを区別する。Regフィールド244の役割は、2つの状況に要約され得る。すなわち、デスティネーションレジスタオペランドまたはソースレジスタオペランドのいずれかをエンコードし、またはオペコード拡張として扱われ、任意の命令オペランドをエンコードするためには用いられない。R/Mフィールド246の役割は、メモリアドレスを参照する命令オペランドをエンコードし、またはデスティネーションレジスタオペランドもしくはソースレジスタオペランドのいずれかをエンコードすることを含み得る。

【0082】

スケール、インデックス、ベース(SIB)バイト(バイト6)。上記のように、スケールフィールド150の内容は、メモリアドレス生成に用いられる。SIB.xxx254およびSIB.bbb256。これらのフィールドの内容は、レジスタインデックスXxxxおよびBbbbに関連して既に言及された。

10

【0083】

変位フィールド162A(バイト7~10)。MODフィールド242が10を含む場合、バイト7~10は、変位フィールド162Aであり、これはレガシ32ビット変位(disp32)と同様に機能し、バイト粒度で機能する。

【0084】

変位係数フィールド162B(バイト7)。MODフィールド242が01を含む場合、バイト7は、変位係数フィールド162Bである。このフィールドの位置は、バイト粒度で機能するレガシx86命令セットの8ビット変位(disp8)の位置と同じである。disp8は、符号拡張されるので、-128~127バイトのオフセットのみをアドレス指定し得る。64バイトのキャッシュラインに関しては、disp8は、4つの本当に有用な値である-128、-64、0、および64のみに設定され得る8ビットを用いる。多くの場合に、より広いレンジが必要とされるので、disp32が用いられる。しかし、disp32は、4バイトを必要とする。disp8およびdisp32とは対照的に、変位係数フィールド162Bは、disp8の再解釈である。変位係数フィールド162Bを用いる場合、実際の変位は、メモリオペランドアクセスのサイズ(N)で乗算される変位係数フィールドの内容により決定される。このタイプの変位は、disp8*Nと称される。これにより、平均命令長(単一バイトだが、はるかに広いレンジの変位に用いられる)を小さくする。そのような圧縮された変位は、有効な変位がメモリアクセスの粒度の倍数であり、従って、アドレスオフセットの冗長下位ビットは、エンコードされる必要がないという前提に基づく。換言すると、変位係数フィールド162Bは、レガシx86命令セットの8ビット変位に置き換わる。従って、disp8がdisp8*Nにオーバーロードされることのみを例外として、変位係数フィールド162Bは、x86命令セットの8ビット変位と同じ態様でエンコードされる(従って、ModRM/SIBエンコードルールに変更はない)。換言すると、エンコードルールまたはエンコードの長さに変更はないが、(バイト的アドレスオフセットを得るべく、メモリオペランドのサイズにより変位を調節する必要がある)ハードウェアによる変位値の解釈のみには変更がある。

20

30

【0085】

即値フィールド172は、上記のように動作する。

40

【0086】

フルオペコードフィールド

図2Bは、本発明の一実施形態による、フルオペコードフィールド174を構成する特定ベクトル向け命令フォーマット200のフィールドを示すブロック図である。具体的には、フルオペコードフィールド174は、フォーマットフィールド140、ベースオペレーションフィールド142、およびデータ要素幅(W)フィールド164を含む。ベースオペレーションフィールド142は、プレフィックスエンコードフィールド225、オペコードマップフィールド215、およびリアルオペコードフィールド230を含む。

【0087】

レジスタインデックスフィールド

50

図2Cは、本発明の一実施形態による、レジスタインデックスフィールド144を構成する特定ベクトル向け命令フォーマット200のフィールドを示すブロック図である。具体的には、レジスタインデックスフィールド144は、REXフィールド205、REX'フィールド210、MODR/M.r e gフィールド244、MODR/M.r / mフィールド246、VVVVフィールド220、xxxフィールド254、およびbbbフィールド256を含む。

【0088】

追加オペレーションフィールド

図2Dは、本発明の一実施形態による、追加オペレーションフィールド150を構成する特定ベクトル向け命令フォーマット200のフィールドを示すブロック図である。クラス(U)フィールド168が0を含む場合、EVEX.U0(クラスA168A)を意味する。1を含む場合、EVEX.U1(クラスB168B)を意味する。U=0、かつMODフィールド242が11を含む場合(非メモリアクセスオペレーションを意味する)、アルファフィールド152(EVEXバイト3、ビット[7]-EH)は、RSフィールド152Aとして解釈される。RSフィールド152Aが1(ラウンド152A.1)を含む場合、ベータフィールド154(EVEXバイト3、ビット[6:4]SSS)は、ラウンド制御フィールド154Aとして解釈される。ラウンド制御フィールド154Aは、1ビットのSAEフィールド156および2ビットのラウンドオペレーションフィールド158を含む。RSフィールド152Aが0(データ変換152A.2)を含む場合、ベータフィールド154(EVEXバイト3、ビット[6:4]SSS)は、3ビットのデータ変換フィールド154Bとして解釈される。U=0であり、かつMODフィールド242が00、01、または10を含む場合(メモリアクセスオペレーションを意味する)、アルファフィールド152(EVEXバイト3、ビット[7]EH)は、エビクションヒント(EH)フィールド152Bとして解釈され、ベータフィールド154(EVEXバイト3、ビット[6:4]SSS)は、3ビットのデータ操作フィールド154Cとして解釈される。

【0089】

U=1である場合、アルファフィールド152(EVEXバイト3、ビット[7]-EH)は、ライトマスク制御(Z)フィールド152Cとして解釈される。U=1であり、かつMODフィールド242が11を含む場合(非メモリアクセスオペレーションを意味する)、ベータフィールド154(EVEXバイト3、ビット[4]-S₀)の一部は、RLフィールド157Aとして解釈される。1(ラウンド157A.1)を含む場合、ベータフィールド154(EVEXバイト3、ビット[6-5]S₂₋₁)の残りは、ラウンドオペレーションフィールド159Aとして解釈されるが、RLフィールド157Aが0(VSIZ E157.A2)を含む場合、ベータフィールド154(EVEXバイト3、ビット[6-5]S₂₋₁)の残りは、ベクトル長フィールド159B(EVEXバイト3、ビット[6-5]L₁₋₀)として解釈される。U=1であり、MODフィールド242が00、01、または10を含む場合(メモリアクセスオペレーションを意味する)、ベータフィールド154(EVEXバイト3、ビット[6:4]-SSS)は、ベクトル長フィールド159B(EVEXバイト3、ビット[6-5]L₁₋₀)およびブロードキャストフィールド157B(EVEXバイト3、ビット[4]B)として解釈される。

【0090】

C. 例示的なレジスタアーキテクチャ

図3は、本発明の一実施形態による、レジスタアーキテクチャ300のブロック図である。示される実施形態において、512ビット幅の32個のベクトルレジスタ310が存在する。これらのレジスタは、zmm0~zmm31として参照される。下位の16個のzmmレジスタの下位の256ビットは、レジスタymm0~15上にオーバーレイされる。下位の16個のzmmレジスタの下位の128ビット(ymmレジスタの下位の128ビット)は、レジスタxmm0~15上にオーバーレイされる。以下の表に示されるよ

10

20

30

40

50

うに、特定ベクトル向け命令フォーマット 200 は、これらのオーバーレイされたレジスタファイルで動作する。

【表 1】

調整可能なベクトル長	クラス	オペレーション	レジスタ
ベクトル長フィールド 159B を含まない命令テンプレート	A (図 1A ; U=0)	110、115 125、130	zmmレジスタ (ベクトル長は64 バイト)
	B (図 1B ; U=1)	112	zmmレジスタ (ベクトル長は64 バイト)
ベクトル長フィールド 159B を含む命令テンプレート	B (図 1B ; U=1)	117、127	ベクトル長フィールド 159B に応じて zmm、ymm、ま たはxmmレジスタ (ベクトル長は64 バイト、32バイト、 または16バイト)

10

20

【0091】

換言すると、ベクトル長フィールド 159B は、最大長、および 1 もしくは複数の他のより短い長さのうちから選択され、そのような各々のより短い長さは、先述の長さの半分の長さである。ベクトル長フィールド 159B を用いない命令テンプレートは、最大ベクトル長で動作する。更に、一実施形態において、特定ベクトル向け命令フォーマット 200 のクラス B 命令テンプレートは、パックドもしくはスカラ単精度 / 倍精度浮動小数点データおよびパックドもしくはスカラ整数データで動作する。スカラオペレーションは、zmm / ymm / xmm レジスタにおける最下位のデータ要素位置で実行されるオペレーションである。より高位のデータ要素位置は、命令前と同じままであるか、または実施形態に応じてゼロにされる。

30

【0092】

ライトマスクレジスタ 315。示される実施形態において、8つのライトマスクレジスタ (k0 ~ k7) が存在し、各々は64ビットのサイズである。代替的な実施形態において、ライトマスクレジスタ 315 は16ビットのサイズである。上記のように、本発明の一実施形態において、ベクトルマスクレジスタ k0 は、ライトマスクとして使用され得ない。通常、k0 を示すエンコードがライトマスクに用いられる場合、これは、0xFFFFのハードワイヤードライトマスクを選択し、その命令に対するライトマスキングを実質的に無効にする。

40

【0093】

汎用レジスタ 325。示される実施形態において、メモリオペランドをアドレス指定する既存の x86 のアドレス指定モードと共に用いられる 16 個の 64 ビット汎用レジスタが存在する。これらのレジスタは、RAX、RBX、RCX、RDX、RBP、RSI、RDI、RSP、および R8 ~ R15 の名称により参照される。

【0094】

スカラ浮動小数点スタックレジスタファイル (x87スタック) 345 上に、MMX パックド整数フラットレジスタファイル 350 がエイリアスされ、示される実施形態において、x87スタックは、x87命令セット拡張を用いて、32 / 64 / 80 ビット浮動小

50

数点データにスカラ浮動小数点オペレーションを実行するべく用いられる8つの要素のスタックである。MMXレジスタは、64ビットパックド整数データにオペレーションを実行すると共に、MMXレジスタとXMMレジスタとの間で実行されるいくつかのオペレーションのためのオペランドを保持するべく用いられる。

【0095】

本発明の代替的な実施形態は、より広いか、またはより狭いレジスタを用い得る。更に、本発明の代替的な実施形態は、より多いか、より少ないか、または異なるレジスタファイルおよびレジスタを用い得る。

【0096】

D. 例示的なコアアーキテクチャ、プロセッサ、およびコンピュータアーキテクチャ
 プロセッサコアは、異なる態様で異なる目的のために異なるプロセッサに実装され得る。例えば、そのようなコアの実装は、1)汎用演算用の汎用インオーダーコア、2)汎用演算用の高性能汎用アウトオブオーダーコア、3)主にグラフィックスおよび/またはサイエンティフィック(スループット)演算用の専用コアを含み得る。異なるプロセッサの実装は、1)汎用演算用の1もしくは複数の汎用インオーダーコア、および/または汎用演算用の1もしくは複数の汎用アウトオブオーダーコアを含むCPU、ならびに2)主にグラフィックスおよび/またはサイエンティフィック(スループット)用の1もしくは複数の専用コアを含むコプロセッサを含み得る。そのような異なるプロセッサは、異なるコンピュータシステムアーキテクチャをもたらし、異なるコンピュータシステムアーキテクチャは、
 1)CPUの別個のチップ上のコプロセッサ、2)CPUと同一のパッケージにおける別個のダイ上のコプロセッサ、3)CPUと同一のダイ上のコプロセッサ(この場合、そのようなコプロセッサは、場合によっては統合グラフィックスおよび/またはサイエンティフィック(スループット)ロジック等の専用ロジック、または専用コアとして言及される)、および4)同一のダイ上に、説明されたCPU(場合によっては、アプリケーションコアもしくはアプリケーションプロセッサとして言及される)、上記のコプロセッサ、および追加の機能性を含み得るシステムオンチップを含み得る。例示的なコアアーキテクチャが次に説明され、その次に例示的なプロセッサおよびコンピュータアーキテクチャの説明が続く。

【0097】

図4Aは、本発明の実施形態による、例示的なインオーダーパイプラインおよび例示的なレジスタリネーム、アウトオブオーダー発行/実行パイプラインの両方を示すブロック図である。図4Bは、本発明の実施形態によるプロセッサに含まれる、インオーダーアーキテクチャコアの例示的な実施形態および例示的なレジスタリネーム、アウトオブオーダー発行/実行アーキテクチャコアの両方を示すブロック図である。図4A~図4Bにおける実線ボックスは、インオーダーパイプラインおよびインオーダーコアを示すが、破線ボックスの任意選択の追加は、レジスタリネーム、アウトオブオーダー発行/実行のパイプラインおよびコアを示す。インオーダーの態様がアウトオブオーダーの態様のサブセットであることを考慮して、アウトオブオーダーの態様が説明される。

【0098】

図4Aにおいて、プロセッサパイプライン400は、フェッチステージ402、長さデコードステージ404、デコードステージ406、アロケーションステージ408、リネームステージ410、スケジューリング(ディスパッチまたは発行としても知られる)ステージ412、レジスタ読み出し/メモリ読み出しステージ414、実行ステージ416、ライトバック/メモリライトステージ418、例外処理ステージ422、およびコミットステージ424を含む。

【0099】

図4Bは、実行エンジンユニット450に結合されたフロントエンドユニット430を含むプロセッサコア490を示し、これら両方はメモリユニット470に結合されている。コア490は、縮小命令セットコンピューティング(RISC)コア、複合命令セットコンピューティング(CISC)コア、超長命令語(VLIW)コア、またはハイブリッ

10

20

30

40

50

ドもしくは代替的なコアタイプであり得る。なおも別の選択肢として、コア490は、例えば、ネットワークコアもしくは通信コア、圧縮エンジン、コプロセッサコア、汎用コンピューティンググラフィックス処理ユニット(GPGPU)コア、グラフィックスコア等のような専用コアであってもよい。

【0100】

フロントエンドユニット430は、命令キャッシュユニット434に結合された分岐予測ユニット432を含み、命令キャッシュユニット434は、命令トランシェーションルックアサイドバッファ(TLB)436に結合され、TLB436は、命令フェッチユニット438に結合され、命令フェッチユニット438は、デコードユニット440に結合される。デコードユニット440(もしくはデコーダ)は、命令をデコードして、出力として1または複数のマイクロオペレーション、マイクロコードエントリポイント、マイクロ命令、他の命令、または元の命令からデコードされ、もしくは別の方法で元の命令を反映し、もしくは元の命令から派生した他の制御信号を生成し得る。デコードユニット440は、様々な異なるメカニズムを用いて実装され得る。好適なメカニズムの例としては、ルックアップテーブル、ハードウェア実装、プログラマブルロジックアレイ(PLA)、マイクロコードリードオンリメモリ(ROM)等が挙げられるが、これらに限定されない。一実施形態において、コア490は、特定のマクロ命令用のマイクロコードを(例えば、デコードユニット440内、またはそうでなければフロントエンドユニット430内に)格納するマイクロコードROMまたは他のメディアを含む。デコードユニット440は、実行エンジンユニット450におけるリネーム/アロケータユニット452に結合される。

10

20

【0101】

実行エンジンユニット450は、リタイアメントユニット454に結合されたリネーム/アロケータユニット452と、1または複数のスケジューラユニット456のセットを含む。スケジューラユニット456は、リザベーションステーション、中央命令ウィンドウ等を含む任意の数の異なるスケジューラを表す。スケジューラユニット456は、物理レジスタファイルユニット458に結合される。物理レジスタファイルユニット458の各々は、1または複数の物理レジスタファイルを表し、これらの異なるものが、スカラ整数、スカラ浮動小数点、パックド整数、パックド浮動小数点、ベクトル整数、ベクトル浮動小数点、状態(例えば、実行されるべき次の命令のアドレスである命令ポインタ)等のような1または複数の異なるデータタイプを格納する。一実施形態において、物理レジスタファイルユニット458は、ベクトルレジスタユニット、ライトマスクレジスタユニット、およびスカラレジスタユニットを備える。これらのレジスタユニットは、アーキテクチャベクトルレジスタ、ベクトルマスクレジスタ、および汎用レジスタを提供し得る。(例えば、リオーダバッファおよびリタイアメントレジスタファイルを用い、フューチャファイル、履歴バッファ、およびリタイアメントレジスタファイルを用い、レジスタマップおよびレジスタのプールを用いる等して)レジスタリネームおよびアウトオブオーダー実行が実装され得る様々な態様を示すべく、物理レジスタファイルユニット458は、リタイアメントユニット454と重ね合わされている。リタイアメントユニット454および物理レジスタファイルユニット458は、実行クラスタ460に結合される。実行クラスタ460は、1または複数の実行ユニット462のセット、および1または複数のメモリアクセスユニット464のセットを含む。実行ユニット462は、様々なタイプのデータ(例えば、スカラ浮動小数点、パックド整数、パックド浮動小数点、ベクトル整数、ベクトル浮動小数点)に対して様々なオペレーション(例えば、シフト、加算、減算、乗算)を実行し得る。いくつかの実施形態は、特定の関数または関数のセットに専用のいくつかの実行ユニットを含み得るが、他の実施形態は、1つの実行ユニットのみ、または全てがあらゆる関数を実行する複数の実行ユニットを含み得る。スケジューラユニット456、物理レジスタファイルユニット458、および実行クラスタ460は、場合によっては複数のものとして示される。なぜなら、ある実施形態は、一定のタイプのデータ/オペレーションのための別個のパイプライン(例えば、各々が自身のスケジューラユニット、物理

30

40

50

レジスタファイルユニット、および/または実行クラスタを有し、別個のメモリアクセスパイプラインの場合に、このパイプラインの実行クラスタのみがメモリアクセスユニット464を有する一定の実施形態が実装される) スカラ整数パイプライン、スカラ浮動小数点/パックド整数/パックド浮動小数点/ベクトル整数/ベクトル浮動小数点パイプライン、および/またはメモリアクセスパイプラインを生成するからである。また、別個のパイプラインが用いられる場合に、これらのパイプラインのうちの1または複数は、アウトオブオーダー発行/実行であり、残りはインオーダーであり得ることを理解されたい。

【0102】

メモリアクセスユニット464のセットは、メモリユニット470に結合される。メモリユニット470は、データキャッシュユニット474に結合されたデータTLBユニット472を含み、データキャッシュユニット474は、レベル2(L2)キャッシュユニット476に結合される。例示的な一実施形態において、メモリアクセスユニット464は、ロードユニット、ストアアドレスユニット、およびストアデータユニットを含み得、これらの各々は、メモリユニット470内のデータTLBユニット472に結合される。命令キャッシュユニット434は、メモリユニット470内のレベル2(L2)キャッシュユニット476に更に結合される。L2キャッシュユニット476は、1つまたは複数の他のレベルのキャッシュに結合され、最終的にはメインメモリに結合される。

【0103】

例として、例示的なレジスタリネーム、アウトオブオーダー発行/実行コアアーキテクチャは、パイプライン400を以下のように実装し得る。1) 命令フェッチ438は、フェッチステージ402および長さデコードステージ404を実行する。2) デコードユニット440はデコードステージ406を実行する。3) リネーム/アロケータユニット452は、アロケーションステージ408およびリネームステージ410を実行する。4) スケジューラユニット456は、スケジューリングステージ412を実行する。5) 物理レジスタファイルユニット458およびメモリユニット470は、レジスタ読み出し/メモリ読み出しステージ414を実行し、実行クラスタ460は、実行ステージ416を実行する。6) メモリユニット470および物理レジスタファイルユニット458は、ライトバック/メモリライトステージ418を実行する。7) 様々なユニットは、例外処理ステージ422に参与してもよい。8) リタイアメントユニット454および物理レジスタファイルユニット458は、コミットステージ424を実行する。

【0104】

コア490は、本明細書において説明される命令を含む、1または複数の命令セット(例えば、x86命令セット(より新しいバージョンを追加された、いくつかの拡張を伴う)、カリフォルニア州サニーベールのMIPS TechnologiesのMIPS命令セット、カリフォルニア州サニーベールのARM HoldingsのARM命令セット(NEON等の任意選択の追加拡張を伴う)をサポートし得る。一実施形態において、コア490は、パックドデータ命令セット拡張(例えば、AVX1、AVX2)をサポートするためのロジックを含み、それにより、多くのマルチメディアアプリケーションにより用いられるオペレーションが、パックドデータを用いて実行されることを可能にする。

【0105】

コアは、(オペレーションまたはスレッドの2またはそれより多い並列セットを実行する) マルチスレッディングをサポートし得、時分割マルチスレッディング、同時マルチスレッディング(物理コアが同時にマルチスレッディングするスレッドの各々のための論理コアを、単一の物理コアが提供する)、またはこれらの組み合わせ(例えば、時分割フェッチおよびデコードを行い、その後にインテル(登録商標)ハイパースレッディング技術等の同時マルチスレッディングを行う)を含む様々な態様でこれを実行し得ることを理解されたい。

【0106】

レジスタリネームは、アウトオブオーダー実行の文脈で説明されているが、レジスタリネームは、インオーダーアーキテクチャにおいて用いられ得ることを理解されたい。プロセッ

10

20

30

40

50

サの示される実施形態は、別個の命令およびデータキャッシュユニット434/474、ならびに共有L2キャッシュユニット476も含むが、代替的な実施形態は、例えば、レベル1(L1)内部キャッシュまたは複数のレベルの内部キャッシュ等の命令およびデータの両方に対する単一の内部キャッシュを有し得る。いくつかの実施形態において、システムは、内部キャッシュ、ならびにコアおよび/またはプロセッサの外部にある外部キャッシュの組み合わせを含み得る。あるいは、キャッシュの全てがコアおよび/またはプロセッサの外部にあってもよい。

【0107】

図5A~図5Bは、より具体的な例示的なインオーダーコアアーキテクチャのブロック図を示し、このコアは、チップにおける(同じタイプの他のコアおよび/または異なるタイプを含む)いくつかの論理ブロックのうちの一つである。アプリケーションに応じて、論理ブロックは、ある固定機能ロジック、メモリI/Oインタフェース、および他の必要なI/Oロジックを用いる高帯域幅の相互接続ネットワーク(例えば、リングネットワーク)を介して通信する。

10

【0108】

図5Aは、本発明の実施形態によるシングルプロセッサコアのブロック図であり、オンダイ相互接続ネットワーク502への接続に加え、レベル2(L2キャッシュ504のローカルサブセットを有する。一実施形態において、命令デコーダ500は、パックドデータ命令セット拡張を用いるx86命令セットをサポートする。L1キャッシュ506は、スカラユニットおよびベクトルユニット内のキャッシュメモリへの低レイテンシアアクセスを可能にする。一実施形態において(設計を簡略化するべく)、スカラユニット508およびベクトルユニット510は、別個のレジスタセット(各々、スカラレジスタ512およびベクトルレジスタ514)を用い、それらの間で転送されるデータは、メモリに書き込まれ、次にレベル1(L1)キャッシュ506からリードバックされる。本発明の代替的な実施形態は、異なるアプローチ(例えば、単一のレジスタセットを用い、またはライトバックおよびリードバックされることなく、2つのレジスタファイルの間で、データが転送されることを可能にする通信パスを含む)を用い得る。

20

【0109】

L2キャッシュ504のローカルサブセットは、プロセッサコア毎に1つずつ、別個のローカルサブセットに分割される全体的なL2キャッシュの一部である。各プロセッサコアは、L2キャッシュ504の自身のローカルサブセットへの直接のアクセス経路を有する。プロセッサコアにより読み出されたデータは、そのL2キャッシュサブセット504に格納され、他のプロセッサコアが自身のローカルL2キャッシュサブセットにアクセスすると並列して迅速にアクセスされ得る。プロセッサコアにより書き込まれたデータは、自身のL2キャッシュサブセット504に格納され、必要であれば他のサブセットからフラッシュされる。リングネットワークは、共有データの coherence を保証する。リングネットワークは、双方向であり、プロセッサコア、L2キャッシュ、および他の論理ブロック等のエージェントがチップ内で互いに通信することを可能にする。各リングのデータパスは、1方向毎に1012ビット幅である。

30

【0110】

図5Bは、本発明の実施形態による図5Aのプロセッサコアの一部の拡大図である。図5Bは、L1キャッシュ504のL1データキャッシュ506A部分、ならびにベクトルユニット510およびベクトルレジスタ514に関する更なる詳細を含む。具体的には、ベクトルユニット510は、16ワイドのベクトル処理ユニット(VPU)(16ワイドのALU528を参照)であり、これは整数命令、単精度浮動小数点命令、および倍精度浮動小数点命令のうちの一つまたは複数を実行する。VPUは、スウィズルユニット520によるレジスタ入力のスウィズル、数値変換ユニット522A~Bによる数値変換、およびメモリ入力時のレプリケーションユニット524によるレプリケーションをサポートする。ライトマスクレジスタ526は、結果として生じるベクトル書き込みを予測(predicting)することを可能にする。

40

50

【0111】

図6は、本発明の実施形態による、2つ以上のコアを有し得、統合メモリコントローラを有し得、統合グラフィックスを有し得るプロセッサ600のブロック図である。図6の実線ボックスは、シングルコア602A、システムエージェント610、1または複数のバスコントローラユニット616のセットを有するプロセッサ600を示すが、破線ボックスの任意選択の追加は、複数のコア602A~N、システムエージェントユニット610における1または複数の統合メモリコントローラユニット614のセット、および専用ロジック608を有する代替的なプロセッサ600を示す。

【0112】

従って、プロセッサ600の異なる実装は、1)統合グラフィックスおよび/またはサイエンティフィック(スループット)ロジック(1または複数のコアを含み得る)である専用ロジック608、および1または複数の汎用コア(例えば、汎用インオーダーコア、汎用アウトオーダーコア、2つの組み合わせ)であるコア602A~Nを用いるCPU、2)主にグラフィックスおよび/またはサイエンティフィック(スループット)用の多数の専用コアであるコア602A~Nを有するコプロセッサ、ならびに3)多数の汎用インオーダーコアであるコア602A~Nを用いるコプロセッサを含み得る。従って、プロセッサ600は、例えば、ネットワークプロセッサもしくは通信プロセッサ、圧縮エンジン、グラフィックスプロセッサ、GPGPU(汎用グラフィックス処理ユニット)、(30またはそれより多いコアを含む)高スループット多集積コア(MIC)コプロセッサ、エンベデッドプロセッサ等のような汎用プロセッサ、コプロセッサもしくは専用プロセッサであつてもよい。プロセッサは、1または複数のチップ上に実装され得る。プロセッサ600は、1または複数の基板の一部であり得、および/または例えば、BiCMOS、CMOS、またはNMOS等のいくつかの処理技術のいずれかを用いてこれらの基板上に実装され得る。

【0113】

メモリ階層は、コア内の1または複数のレベルのキャッシュ、1セットまたは1もしくは複数の共有キャッシュユニット606、および統合メモリコントローラユニット614のセットに結合された外部メモリ(図示せず)を含む。共有キャッシュユニット606のセットは、レベル2(L2)、レベル3(L3)、レベル4(L4)等の1または複数の中間レベルのキャッシュ、または他のレベルのキャッシュ、ラストレベルキャッシュ(LLC)、および/またはこれらの組み合わせを含み得る。一実施形態において、リングベースの相互接続ユニット612は、統合グラフィックスロジック608、共有キャッシュユニット606のセット、およびシステムエージェントユニット610/統合メモリコントローラユニット614を相互接続し、代替的な実施形態は、そのようなユニットを相互接続するための任意の数の周知の技術を用い得る。一実施形態において、コヒーレンシは、1または複数のキャッシュユニット606とコア602A~Nとの間で保持される。

【0114】

いくつかの実施形態において、コア602A~Nのうちの1または複数は、マルチスレディングすることができる。システムエージェント610は、コア602A~Nを調整および動作させるそれらのコンポーネントを含む。システムエージェントユニット610は、例えば、電力制御ユニット(PCU)およびディスプレイユニットを含み得る。PCUは、コア602A~Nおよび統合グラフィックスロジック608の電力状態を調整するのに必要とされるロジックおよびコンポーネントであるか、またはこれらを含み得る。ディスプレイユニットは、1または複数の外部接続ディスプレイを駆動するためのものである。

【0115】

コア602A~Nは、アーキテクチャ命令セットの観点からは同種または異種であり得る。すなわち、コア602A~Nのうちの2またはそれより多くのものは、同じ命令セットを実行することができる場合があるが、他のものは、当該命令セットまたは異なる命令セットのサブセットのみを実行することができる場合がある。

10

20

30

40

50

【 0 1 1 6 】

図7～図10は、例示的なコンピュータアーキテクチャのブロック図である。ラップトップ、デスクトップ、ハンドヘルドPC、情報携帯端末、エンジニアリングワークステーション、サーバ、ネットワークデバイス、ネットワークハブ、スイッチ、エンベデッドプロセッサ、デジタル信号プロセッサ(DSP)、グラフィックスデバイス、ビデオゲームデバイス、セットトップボックス、マイクロコントローラ、携帯電話、ポータブルメディアプレーヤ、ハンドヘルドデバイス、および様々な他の電子デバイスの技術分野で既知の他のシステム設計および構成も好適である。一般に、本明細書において開示されるプロセッサおよび/または他の実行ロジックを組み込むことができる多種多様なシステムまたは電子デバイスが、概ね好適である。

10

【 0 1 1 7 】

ここで図7を参照すると、本発明の一実施形態によるシステム700のブロック図が示される。システム700は、1または複数のプロセッサ710、715を含み得、これらは、コントローラハブ720に結合される。一実施形態において、コントローラハブ720は、グラフィックスメモリコントローラハブ(GMCH)790および入出力ハブ(I/OH)750(別個のチップ上にあり得る)を含む。GMCH790は、メモリコントローラおよびグラフィックスコントローラを含み、これらにメモリ740およびコプロセッサ745が結合される。I/OH750は、入出力(I/O)デバイス760をGMCH790に結合する。あるいは、メモリおよびグラフィックスコントローラの一方または両方が(本明細書に説明される)プロセッサ内に集積され、メモリ740およびコプロセッサ745は、プロセッサ710と、I/OH750を有する単一のチップのコントローラハブ720とに直接に結合される。

20

【 0 1 1 8 】

追加のプロセッサ715の任意選択の性質が図7に破線を用いて示されている。各プロセッサ710、715は、本明細書に説明される処理コアのうちの1または複数を含み得、プロセッサ600のいくつかのバージョンであり得る。

【 0 1 1 9 】

メモリ740は、例えば、ダイナミックランダムアクセスメモリ(DRAM)、相変化メモリ(PCM)、またはこれら2つの組み合わせであってもよい。少なくとも一実施形態においては、コントローラハブ720は、フロントサイドバス(FSB)等のマルチドロップバス、Quick Path相互接続(QPI)等のポイントツーポイントインタフェース、または類似の接続795を介してプロセッサ710、715と通信する。

30

【 0 1 2 0 】

一実施形態において、コプロセッサ745は、例えば、高スループットMICプロセッサ、ネットワークプロセッサもしくは通信プロセッサ、圧縮エンジン、グラフィックスプロセッサ、GPGPU、エンベデッドプロセッサ等のような専用プロセッサである。一実施形態において、コントローラハブ720は、統合グラフィックスアクセラレータを含み得る。

【 0 1 2 1 】

物理リソース710と物理リソース715との間には、アーキテクチャ、マイクロアーキテクチャ、熱、電力消費の特性等を含む幅広い価値基準に関して様々な違いが存在し得る。

40

【 0 1 2 2 】

一実施形態において、プロセッサ710は、一般的なタイプのデータ処理オペレーションを制御する命令を実行する。コプロセッサ命令は、命令中に埋め込まれ得る。プロセッサ710は、取り付けられたコプロセッサ745により実行されるべきタイプとしてこれらのコプロセッサ命令を認識する。従って、プロセッサ710は、コプロセッサバスまたは他の相互接続においてこれらのコプロセッサ命令(またはコプロセッサ命令を表す制御信号)をコプロセッサ745に発する。コプロセッサ745は、受信したコプロセッサ命令を受け取って実行する。

50

【 0 1 2 3 】

ここで図 8 を参照すると、本発明の実施形態による第 1 のより具体的な例示的システム 8 0 0 のブロック図が示される。図 8 に示されるように、マルチプロセッサシステム 8 0 0 はポイントツーポイント相互接続システムであり、ポイントツーポイント相互接続 8 5 0 を介して結合される第 1 のプロセッサ 8 7 0 および第 2 のプロセッサ 8 8 0 を含む。プロセッサ 8 7 0 および 8 8 0 の各々は、プロセッサ 6 0 0 のいくつかのバージョンであり得る。本発明の一実施形態において、プロセッサ 8 7 0 および 8 8 0 は各々、プロセッサ 7 1 0 および 7 1 5 であるが、コプロセッサ 8 3 8 は、コプロセッサ 7 4 5 である。別の実施形態において、プロセッサ 8 7 0 および 8 8 0 は各々、プロセッサ 7 1 0 およびコプロセッサ 7 4 5 である。

10

【 0 1 2 4 】

統合メモリコントローラ (I M C) ユニット 8 7 2 および 8 8 2 を各々含むプロセッサ 8 7 0 および 8 8 0 が示される。プロセッサ 8 7 0 は、そのバスコントローラユニットの一部としてポイントツーポイント (P P) インタフェース 8 7 6 および 8 7 8 も含む。同様に、第 2 のプロセッサ 8 8 0 は、P P インタフェース 8 8 6 および 8 8 8 を含む。プロセッサ 8 7 0、8 8 0 は、P P インタフェース回路 8 7 8、8 8 8 を用いて、ポイントツーポイント (P P) インタフェース 8 5 0 を介して情報を交換し得る。図 8 に示されるように、I M C 8 7 2 および 8 8 2 は、プロセッサを各メモリ、すなわち、各プロセッサにローカルに取り付けられたメインメモリの一部であり得るメモリ 8 3 2 およびメモリ 8 3 4 に結合する。

20

【 0 1 2 5 】

プロセッサ 8 7 0、8 8 0 は各々、ポイントツーポイントインタフェース回路 8 7 6、8 9 4、8 8 6、8 9 8 を用い、個々の P P インタフェース 8 5 2、8 5 4 を介してチップセット 8 9 0 と情報を交換し得る。任意選択で、チップセット 8 9 0 は、高性能インタフェース 8 3 9 を介してコプロセッサ 8 3 8 と情報を交換し得る。一実施形態において、コプロセッサ 8 3 8 は、例えば、高スループット M I C プロセッサ、ネットワークプロセッサもしくは通信プロセッサ、圧縮エンジン、グラフィックスプロセッサ、G P G P U、エンベデッドプロセッサ等のような専用プロセッサである。

【 0 1 2 6 】

共有キャッシュ (図示せず) は、どちらかのプロセッサに含まれ、または両方のプロセッサの外部にあり得るが、プロセッサが低電力モードにされると、どちらかまたは両方のプロセッサのローカルキャッシュ情報が共有キャッシュに格納され得るように、P P 相互接続を介してプロセッサとなおも接続され得る。

30

【 0 1 2 7 】

チップセット 8 9 0 は、インタフェース 8 9 6 を介して第 1 のバス 8 1 6 に結合され得る。一実施形態において、第 1 のバス 8 1 6 は、周辺構成要素相互接続 (P C I) バス、または P C I E x p r e s s バス等のバス、または別の第 3 世代 I / O 相互接続バスであり得るが、本発明の範囲はそのように限定されない。

【 0 1 2 8 】

図 8 に示されるように、様々な I / O デバイス 8 1 4 が、第 1 のバス 8 1 6 を第 2 のバス 8 2 0 に結合するバスブリッジ 8 1 8 と共に、第 1 のバス 8 1 6 に結合され得る。一実施形態において、コプロセッサ、高スループット M I C プロセッサ、G P G P U、アクセラレータ (例えば、グラフィックスアクセラレータまたはデジタル信号処理 (D S P) ユニット等)、フィールドプログラマブルゲートアレイ、またはその他のプロセッサ等、1 または複数の追加のプロセッサ 8 1 5 は、第 1 のバス 8 1 6 に結合される。一実施形態において、第 2 のバス 8 2 0 は、低ピンカウント (L P C) バスであり得る。様々なデバイスは、一実施形態において、例えば、キーボードおよび/またはマウス 8 2 2、通信デバイス 8 2 7、ならびに命令 / コードおよびデータ 8 3 0 を含む得るディスクドライブもしくは他の大容量ストレージデバイス等のストレージユニット 8 2 8 を含む第 2 のバス 8 2 0 に結合され得る。更に、オーディオ I / O 8 2 4 は、第 2 のバス 8 2 0 に結合され得る

40

50

。他のアーキテクチャが可能であることに留意されたい。例えば、図 8 のポイントツーポイントアーキテクチャに代えて、システムは、マルチドロップバスまたは他のそのようなアーキテクチャを実装し得る。

【 0 1 2 9 】

ここで図 9 を参照すると、本発明の実施形態による第 2 のより具体的な例示的システム 9 0 0 のブロック図を示す。図 8 および図 9 の同一の要素は、同一の参照番号を有し、図 9 の他の態様を不明瞭にするのを避けるべく、図 8 の特定の態様は、図 9 から省略されている。

【 0 1 3 0 】

図 9 は、プロセッサ 8 7 0、8 8 0 が統合メモリおよび I / O 制御ロジック (「 C L 」) 8 7 2 および 8 8 2 を各々含み得ることを示す。従って、C L 8 7 2、8 8 2 は、統合メモリコントローラユニットを含み、I / O 制御ロジックを含む。図 9 は、メモリ 8 3 2、8 3 4 が C L 8 7 2、8 8 2 に結合されていることのみならず、I / O デバイス 9 1 4 も制御ロジック 8 7 2、8 8 2 に結合されていることを示す。レガシ I / O デバイス 9 1 5 は、チップセット 8 9 0 に結合される。

【 0 1 3 1 】

ここで図 1 0 を参照すると、本発明の一実施形態による S o C 1 0 0 0 のブロック図が示される。図 6 における類似の要素は、同一の参照番号を有する。また、破線ボックスは、より高度な S o C の任意選択の特徴である。図 1 0 において、相互接続ユニット 1 0 0 2 は、1 もしくは複数のコア 6 0 6 A ~ N および共有キャッシュユニット 6 0 6 のセットを含むアプリケーションプロセッサ 1 0 1 0、システムエージェントユニット 6 1 0、バスコントローラユニット 6 1 6、統合メモリコントローラユニット 6 1 4、統合グラフィックスロジック、画像プロセッサ、オーディオプロセッサ、およびビデオプロセッサを含み得る 1 もしくは複数のコプロセッサ 1 0 2 0 のセット、スタティックランダムアクセスメモリ (S R A M) ユニット 1 0 3 0、ダイレクトメモリアクセス (D M A) ユニット 1 0 3 2、および 1 または複数の外部ディスプレイを結合するディスプレイユニット 1 0 4 0 に結合される。一実施形態において、コプロセッサ 1 0 2 0 は、例えば、ネットワークプロセッサもしくは通信プロセッサ、圧縮エンジン、G P G P U、高スループット M I C プロセッサ、エンベデッドプロセッサ等のような専用プロセッサを含む。

【 0 1 3 2 】

本明細書に開示されるメカニズムの実施形態は、ハードウェア、ソフトウェア、ファームウェア、またはそのような実装アプローチの組み合わせで実装され得る。本発明の実施形態は、少なくとも 1 つのプロセッサ、ストレージシステム (揮発性および不揮発性メモリ、ならびに / またはストレージ要素を含む)、少なくとも 1 つの入力デバイス、および少なくとも 1 つの出力デバイスを有するプログラマブルシステム上で実行するコンピュータプログラムまたはプログラムコードとして実装され得る。

【 0 1 3 3 】

図 8 に示されるコード 8 3 0 等のプログラムコードは、本明細書に説明される機能を実行して出力情報を生成するべく、入力命令に適用され得る。出力情報は、既知の様式で 1 または複数の出力デバイスに適用され得る。本願の目的のために、処理システムは、例えば、デジタル信号プロセッサ (D S P)、マイクロコントローラ、特定用途向け集積回路 (A S I C)、またはマイクロプロセッサ等のプロセッサを有する任意のシステムを含む。

【 0 1 3 4 】

プログラムコードは、高水準手続き型プログラミング言語またはオブジェクト指向プログラミング言語で実装され、処理システムと通信し得る。所望であれば、プログラムコードは、アセンブリ言語または機械語でも実装され得る。実際には、本明細書に説明されるメカニズムは、範囲においていずれの特定のプログラミング言語にも限定されない。いずれの場合においても、言語は、コンパイル型言語またはインタープリタ型言語であってもよい。

10

20

30

40

50

【 0 1 3 5 】

少なくとも一実施形態の1または複数の態様は、プロセッサ内の様々なロジックを表す、機械可読媒体上に格納された代表的命令により実装され得、命令は、機械により読み出されると、当該機械に本明細書に説明される技術を実行するためのロジックを生成させる。「IPコア」として知られるそのような表現は、有形機械可読媒体上に格納され、様々な顧客または製造設備に供給され、ロジックまたはプロセッサを実際に作成する製造機械にロードされ得る。

【 0 1 3 6 】

そのような機械可読ストレージ媒体としては、ハードディスク、フロッピー（登録商標）ディスク、光ディスク、コンパクトディスクリードオンリメモリ（CD ROM）、書き換え可能コンパクトディスク（CD RW）、および光磁気ディスク等、その他のタイプのディスクを含むストレージ媒体、リードオンリメモリ（ROM）、ダイナミックランダムアクセスメモリ（DRAM）、スタティックランダムアクセスメモリ（SRAM）等のランダムアクセスメモリ（RAM）、消去可能プログラマブルリードオンリメモリ（EPROM）、フラッシュメモリ、電氣的消去可能プログラマブルリードオンリメモリ（EEPROM）、および相変化メモリ（PCM）等の半導体デバイス、磁気もしくは光カード、または電子命令を格納するのに好適なその他のタイプの媒体を含む、機械またはデバイスにより製造または形成される、非一時的で有形な構成の物品が挙げられ得るが、これらに限定されない。

【 0 1 3 7 】

従って、本発明の実施形態は、命令を含み、または本明細書に説明される構造体、回路、装置、プロセッサ、および/またはシステム機能を定義するハードウェア記述言語（HDL）等の設計データを含む非一時的有形機械可読媒体も含む。そのような実施形態は、プログラム製品と称され得る。

【 0 1 3 8 】

いくつかの場合に、命令コンバータは、ソース命令セットからターゲット命令セットへと命令を変換するべく用いられ得る。例えば、命令コンバータは、（例えば、静的バイナリ変換、動的コンパイルを含む動的バイナリ変換を用いて）命令を、コアにより処理されるべき1または複数の他の命令に翻訳し、モーフィングし、エミュレートし、または別の方法で変換し得る。命令コンバータは、ソフトウェア、ハードウェア、ファームウェア、またはこれらの組み合わせで実装され得る。命令コンバータは、プロセッサにあり、プロセッサから離れ、またはプロセッサ上の一部であり、プロセッサから離れた一部であり得る。

【 0 1 3 9 】

図11は、本発明の実施形態による、ソース命令セットのバイナリ命令を、ターゲット命令セットのバイナリ命令に変換するソフトウェア命令コンバータの使用と対比するブロック図である。示された実施形態において、命令コンバータは、ソフトウェア命令コンバータであるが、あるいは、命令コンバータは、ソフトウェア、ファームウェア、ハードウェア、またはこれらの様々な組み合わせで実装され得る。図11は、高水準言語1102によるプログラムが、少なくとも1つのx86命令セットコア1116を有するプロセッサによりネイティブに実行され得るx86バイナリコード1106を生成するべく、x86コンパイラ1104を用いてコンパイルされ得ることを示す。少なくとも1つのx86命令セットコア1116を有するプロセッサは、少なくとも1つのx86命令セットコアを用いるインテル（登録商標）プロセッサと実質的に同一の結果を実現するべく、（1）インテル（登録商標）x86命令セットコアの命令セットの実質的部分、または（2）少なくとも1つのx86命令セットコアを有するインテル（登録商標）プロセッサ上で起動することを目的とする、オブジェクトコードバージョンのアプリケーションまたは他のソフトウェアを互換的に実行し、または別の方法で処理することにより、少なくとも1つのx86命令セットコアを有するインテル（登録商標）プロセッサと実質的に同一の機能を実行し得る任意のプロセッサを表す。x86コンパイラ1104は、x86バイナリコー

ド 1 1 0 6 (例えば、オブジェクトコード)を生成するように動作可能なコンパイラを表す。x 8 6 バイナリコード 1 1 0 6 は、追加のリンク処理を用いるか、または用いることなく少なくとも1つのx 8 6 命令セットコア 1 1 1 6 を有するプロセッサにおいて実行され得る。同様に、図 1 1 は、高水準言語 1 1 0 2 によるプログラムが、少なくとも1つのx 8 6 命令セットコア 1 1 1 4 を有しないプロセッサ(例えば、カリフォルニア州サニーベールの M I P S T e c h n o l o g i e s の M I P S 命令セットを実行し、および/またはカリフォルニア州サニーベールの A R M H o l d i n g s の A R M 命令セットを実行するコアを有するプロセッサ)によりネイティブに実行され得る代替的な命令セットバイナリコード 1 1 1 0 を生成するべく、代替的な命令セットコンパイラ 1 1 0 8 を用いてコンパイルされ得ることを示す。命令コンバータ 1 1 1 2 は、x 8 6 バイナリコード 1 1 0 6 を、x 8 6 命令セットコア 1 1 1 4 を有しないプロセッサによりネイティブに実行され得るコードに変換するべく用いられる。この変換済みコードは、代替的な命令セットバイナリコード 1 1 1 0 と同じである可能性が低い。なぜなら、これが可能な命令コンバータは、作成することが困難だからである。しかし、変換済みコードは、一般的なオペレーションを実現し、代替的な命令セットの命令から構成される。従って、命令コンバータ 1 1 1 2 は、ソフトウェア、ファームウェア、ハードウェア、またはこれらの組み合わせを表し、これらは、エミュレーション、シミュレーション、またはその他の処理により、x 8 6 命令セットプロセッサまたはコアを有しないプロセッサまたは他の電子デバイスが x 8 6 バイナリコード 1 1 0 6 を実行することを可能にする。

10

【 0 1 4 0 】

20

命令フローを最適化するチェックを実装するための方法および装置

除算および平方根のような I E E E の正確な丸め関数は、2つの態様で実装され得る。すなわち、基数ベースのハードウェアディバイダを用いて機能をネイティブに提供するか、またはソフトウェア命令/マイクロコード (u c o d e) シーケンス、通常は、ニュートン・ファフソナルゴリズムの形式/変形を実装し、最終的結果を得るべくシードに対してイテレートするかである。

【 0 1 4 1 】

ソフトウェアシーケンスに関する1つの問題は、大抵の入力についてアルゴリズムは、非常に効率的に処理され得るのに対して、(例えば、入力非正規化を処理する場合に)一桁遅いシーケンスをもたらすいくつかの特定のコーナーケースには、追加に処理が必要とされるということである。更に、入力がトリビアルな結果を有する場合に結果をパッチするべく、最終的な後処理(ゼロによる除算等)が必要とされ、これによりシーケンスの限界レイテンシが増加する。

30

【 0 1 4 2 】

一実施形態において、命令シーケンスは、最良の場合を想定しており、入力においてコーナーの場合が存在するときには、代替的または「低速」のシーケンスにジャンプし、デフォルトシーケンスのクリティカルパスからチェックアウトする。本発明の実施形態は、主要のデフォルト計算とは別個のパスでこの条件付きジャンプを予め計算する「チェック」命令またはマイクロオペレーション (u o p) のタイプを含む。いくつかの実施形態において、命令チェックオペレーションは、トリビアルな入力に対するデフォルト応答も提供する。下記の実施形態は、除算および平方根 (s q r t) の実装に焦点を当てるが、本発明の基礎となる原理は、いずれの特定のタイプの数学関数にも限定されない。より具体的には、一実施形態は、(例えば、プレディケーション/マスクを用いる)パックス平方根/除算の A V X 5 1 2 実装に基づくが、このコンセプトは、いずれの命令セットアーキテクチャにも拡張され得る。

40

【 0 1 4 3 】

「チェック」命令 / u o p の一実施形態は、入力に基づいて、デフォルト/ファーストフローを用いるアルゴリズムを処理するか否か、または代替的にスローフローにジャンプし、プリアンベクトルをマスクに残しておくべきかを決定する。更に、一実施形態は、計算毎の例外フラグを生成し(無効、非正規、ゼロによる除算)、マスクされない場合に

50

は、実行をアボートする。最終的には、本発明の一実施形態は、トリビアルな入力についてはデフォルト応答を予め計算する（例えば、無限大で割る等）。

【0144】

図12に示されるように、本発明の実施形態が実装され得る例示的なプロセッサ1255は、命令チェック命令をデコードするためのプロセッサパイプラインのデコードステージ1230内における命令チェックデコードロジック1231と、命令チェック命令を実行するためのパイプラインの実行ステージ1240内における命令チェック実行ロジック1241とを含む。一実施形態において、命令チェックデコードロジック1231は、命令チェック命令を複数のマイクロオペレーションにデコードするためのハードウェアを備え、複数のマイクロオペレーションは次に、命令チェック実行ロジック1241により実行される。しかし、本発明の基礎となる原理は、マイクロコード化プロセッサに限定されない。言及されたように、一実施形態において、命令は、除算チェック命令(DIVCHK())および平方根チェック(SQRTCHK())を含む。しかし、本発明の基礎となる原理は、これら特定の実装に限定されない。

10

【0145】

図12に示される他のプロセッサコンポーネントは、汎用レジスタ(GPR)1205のセット、ベクトルレジスタ1206のセット、マスクレジスタ1207のセット、および制御レジスタ1208のセットを含む。図12に明示的に示されていないが、制御レジスタ1208は、別個の制御レジスタバス(CRBus)インタフェースを介してアクセスされ得る。一実施形態において、マスクレジスタ1207は、(例えば、上記のマスクレジスタk0~k7として実装される)ベクトルレジスタ1206に格納された値にビットマスクングオペレーションを実行するために用いられる8個の64ビットオペランドのマスクレジスタを含む。しかし、本発明の基礎となる原理は、いずれの特定のマスクレジスタサイズ/タイプにも限定されない。一実施形態において、複数のベクトルデータ要素は、2個の256ビット値、4個の128ビット値、8個の64ビット値、16個の32ビット値等を格納するための512ビット幅を有し得る各ベクトルレジスタ1206にパックされ得る。しかし、本発明の基礎となる原理は、いずれの特定のサイズ/タイプのベクトルデータにも限定されない。

20

【0146】

シングルプロセッサコア(「コア0」)の詳細は、簡略化するために図12に示される。しかし、図12に示される各コアは、コア0と同じセットのロジックを有し得ることが理解されよう。例えば、各コアは、指定されたキャッシュ管理ポリシーに従って命令およびデータをキャッシュするための専用レベル1(L1)キャッシュ1212およびレベル2(L2)キャッシュ1211を含み得る。L1キャッシュ1212は、命令を格納するための別個の命令キャッシュ1220と、データを格納するための別個のデータキャッシュ1221とを含む。様々なプロセッサキャッシュに格納された命令およびデータは、固定サイズ(例えば、64バイト、128バイト、512バイトの長さ)であり得るキャッシュラインの粒度で管理される。この例示的な実施形態の各コアは、メインメモリ1200および/または共有レベル3(L3)キャッシュ1216から命令をフェッチするための命令フェッチユニット1210と、命令をデコードする(例えば、プログラム命令をマイクロオペレーションまたは「マイクロオブ」にデコードする)ためのデコードユニット1220と、命令を実行するための実行ユニット1240と、命令をリタイアして結果をライトバックするためのライトバックユニット1250とを有する。

30

40

【0147】

命令フェッチユニット1210は、メモリ1200(またはキャッシュのうちの1つ)からフェッチされるべき次の命令のアドレスを格納するための次の命令ポインタ1203と、最近用いられた仮想・物理命令アドレスのマッピングを格納してアドレス変換の速度を向上させるための命令トランслーションルックアサイドバッファ(ITLB)1204と、命令分岐アドレスを推論的に予測するための分岐予測ユニット1202と、分岐アドレスおよびターゲットアドレスを格納するための分岐ターゲットバッファ(BTB)120

50

1 とを含む、様々な周知のコンポーネントを含む。フェッチされると、命令は次に、デコードユニット 1 2 3 0、実行ユニット 1 2 4 0、およびライトバックユニット 1 2 5 0 を含む命令パイプラインの残りのステージにストリーミングされる。これらのユニットの各々の構造および機能は、当業者により良く理解されており、本発明の異なる実施形態の関連する態様を不明瞭にするのを避けるべく、ここでは詳細に説明されない。

【 0 1 4 8 】

図 1 3 は、現在の命令に関連する複数の入力値 1 3 0 1 ~ 1 3 0 3 (例えば、除算命令または平方根命令のための入力)に基づいて、特定の命令シーケンス(例えば、デフォルト/高速シーケンスまたは代替的/低速シーケンス)を選択するためのベクトル出力およびマスク出力 1 3 1 1 を生成する命令チェックロジック 1 3 0 0 を示す。除算命令および平方根命令のためのベクトル出力およびマスク出力の様々な具体的な例が以下に提供される(例えば、図 1 4 (除算)および図 1 5 (平方根)を参照)。更に、いくつかの実施形態において、命令チェックロジック 1 3 0 0 は、特定のタイプの入力(例えば、無限大で割るなどしたトリビアな入力)についての結果 1 3 1 2 を生成し得、後続の命令シーケンス 1 3 1 1 が使用できる例外フラグ 1 3 1 3 を生成し得る。例として、例外フラグは、ゼロによる除算(ZE)、無効な演算(IE)、および非正規化オペランド(DE)等の浮動小数点例外フラグを含み得る。

10

【 0 1 4 9 】

上述のように、本明細書に説明される技術を実装する命令チェックロジック 1 3 0 0 は、命令を複数の u o p にデコードするデコードステージ 1 2 3 0 (例えば、命令チェックデコードロジック 1 2 3 1)と、u o p を実行する実行ステージ 1 2 4 0 (例えば、命令チェック実行ロジック 1 2 4 1)内のロジックとを含み得る。しかし、本発明の基礎となる原理は、マイクロコード化プロセッサアーキテクチャに限定されないことに留意されたい。

20

【 0 1 5 0 】

除算命令に関連する入力および他の変数にチェックを実行し、これに応じて適切な/効率的な命令シーケンスを選択して除算オペレーションを実装する除算チェック(DIVCHK)命令の一実施形態が、ここで説明される。一実施形態において、除算のためのマイクロコードフロー手順は、以下の通りである。

【 数 1 】

30

```

division(x,y)
{
    start_fast_sequence(x,y);
    if(DIVCHK())
        jump slow_flow;
    else
        finish_fast_flow(x,y)
    return res;
slow_flow:
    do_slow_flow(x,y)
    return res;
}

```

40

【 0 1 5 1 】

上記のコードにおいて、DIVCHKは、入力(x,y)を評価し、代替的な「slow_flow」命令シーケンスにジャンプするか否か、または通常の「finish_fast_flow」シーケンスを完了するか否かを決定する。一実施形態において、以下

50

のDIVCHK仕様が使用される。

【0152】

一実施形態において、マイクロオペUOP_{JDIVCHK}{PS,PD}が使用され、第1のパラメータ(Param1)は、除算入力をチェックし、第2のパラメータ(Param2)は、オペレーションが単精度入力を伴うか、または倍精度入力を伴うかを示す。UOPは、第1のソース(srcA)を分子として、第2のソース(srcB)を分母として用い得、除算シーケンスに対する入力の特性をチェックして、3つの異なる情報を返す。1. 特別の場合の入力のデフォルト応答によるベクトル出力(例えば、ゼロによる除算は、INFを返す)。一実施形態において、これは、スローパスアルゴリズムを用いて実際の出力が生成される必要がある場合に、特別なインジケーション(例えば、+1.0)を出力する。2. ファーストパスアルゴリズムを用いて処理され得る要素をシグナリングするマスク出力。一実施形態において、ゼロのマスクは、結果がトリビアルなものであるか、またはスローフローシーケンス(+1.0等の特別なインジケーションにより指定された)により計算が行われる必要があるかのいずれかを意味する。3. 計算毎の例外フラグの生成。これは、例えば、ゼロによる除算(ZE)、無効な演算(IE)、および非正規化オペランド(DE)等の浮動小数点例外フラグを含み得る。

10

【0153】

図14に示される表は、2つの異なる入力タイプ(XはINPUT1であるが、YはINPUT2である)の関数における3つの出力を要約する。この表において、X入力には示され、Y入力は、列に示されている。結果として生じる出力は、(kdst, vdst, flags)として構成され、「kdst」はマスク出力を含み、「vdst」はベクトル出力を含み、「flags」は例外フラグを含む。入力「denorm」は非正規化オペランドを表し、「normal」は通常オペランドを表し、「Inf」は無限を表す。NaNは「非数(not a number)」を表し、QNaNはquiet NaNを表し、SNaNはsignaling NaNを表す。QNaNは、最上位の端数ビットを1にセットしたNaNであり、SNaNは、最上位の端数ビットをクリアしたNaNである。

20

【0154】

一実施形態において、両方の入力正規数である場合(表において「以下を参照」で示される)は、以下のように処理される。

30

【数2】

```
xExp = biased exponent of x; // 指数のメモリ形式
yExp = biased exponent of y; // 指数のメモリ形式
eDiff = xExp - yExp;
if ( (eDiff ≥ bias + 1) || // 商がオーバーフローする可能性がある
      (eDiff ≤ (1 - bias)) || // 商がアンダーフローする可能性がある
      (yExp ≥ (2*bias - 1)) || // 逆数がアンダーフローする可能性がある
      (xExp ≤ 2*precision) ) // 剰余がアンダーフローする可能性がある
){
    return (0,+1.0,None);
} else
    return (1,+0.0,None);
}
```

40

従って、上記のコードは、商のオーバーフロー/アンダーフロー、逆数アンダーフロー、および剰余のアンダーフローについてテストする。これらの条件のいずれかが検出されると、ベクトル出力は+1.0であり、そうでなければベクトル出力は+0.0である。

50

剰余テストにおいては、I S A から生じる入出力の非正規数が内部計算においてフラッシュされ得る。

【 0 1 5 5 】

一実施形態において、D I V C H K オペレーションは、以下の形式を取る。

【 数 3 】

```
// UOP_DIVCHK_PS (.INPUT1(src_a), .INPUT2(src_b))
for(n=0; n<2; n+=1)
{
    i = n*32
    {MASK_OUTPUT[n], VECT_OUTPUT[i+31:i]}
    = divchk(INPUT1[i+31:i], INPUT2[i+31:i]) // (x, y)
}
```

10

【 0 1 5 6 】

例外は、単精度および倍精度の浮動小数点データ要素の両方に対する非正規 (D A Z = 0 の場合)、無効、およびゼロによる除算に対して生成され得る。

【 0 1 5 7 】

平方根命令に関連する入力および他の変数にチェックを実行し、これに応じて適切な / 効率的な命令シーケンスを選択して平方根演算を実装する平方根チェック (S Q R T C H K) 命令の一実施形態が、ここで説明される。一実施形態において、平方根のマイクロコードフロー手順は、以下のとおりである。

20

【 数 4 】

```
sqrt(x)
{
    start_fast_sequence(x);
    if(SQRTCHK())
        jump slow_flow;
    else
        finish_fast_flow(x)
        return res;
Slow_flow:
    do_slow_flow(x)
    return res;
}
```

30

40

【 0 1 5 8 】

上記のコードにおいて、S Q R T C H K は、入力 x を評価し、代替的な「s l o w _ f l o w」命令シーケンスにジャンプするか否か、または通常の「f i n i s h _ f a s t _ f l o w」シーケンスを完了するか否かを決定する。一実施形態において、以下の S Q R T C H K 仕様が使用される。

【 0 1 5 9 】

U O P _ { J S Q R T C H K } { P S , P D } は、一実施形態において使用されるマイクロオペレーションであり、第 1 のパラメータ { S Q R T C H K } は、平方根命令入力をチェックし、第 2 のパラメータ { P S , P D } は、オペレーションが単精度入力または倍精度入力を伴うかを示す。U O P は、平方根シーケンスに対する入力の特性をチェックし

50

、3つの異なる情報を返す。1. 特別の場合の入力のデフォルト応答によるベクトル出力（例えば、負の数の平方根は、NaNを返す）。一実施形態において、これは、ファーストパスアルゴリズムを用いて実際の出力が生成される必要がある場合に、特別なインジケーション（例えば、+1.0）を出力する。2. スローパスアルゴリズムを用いて処理されなければならない要素をシグナリングするマスク出力。3. 計算毎の例外フラグの生成。これは、例えば、無効な演算（IE）、および非正規化オペランド（DE）等の浮動小数点例外フラグを含み得る。

【0160】

図15に示される表は、x（第1の列）に対する様々な異なる入力タイプおよび（kdst, vdst, flags）として構成された、結果として生じる出力を列挙し、「kdst」はマスク出力を含み、「vdst」はベクトル出力を含み、「flags」は例外フラグを含む。入力は、-normを含む（-無限大+/-非正規化オペランド、+/-0、+無限大、qNaN、およびsNaNを含む）。制限BiasExponent(x) < 2*の精度（表において条件(a)として識別される）は、d = x - s*sの計算におけるアンダーフローを防止することが意図される。

【数5】

```
// UOP_SQRTCHK_PS (.INPUT1(src_b))
for(n=0; n<2; n+=1)
{
    i = n*32
    {MASK_OUTPUT[n], VECT_OUTPUT[i+31:i]}
    = sqrtchk(INPUT1[i+31:i]);
}
```

【0161】

例外は、単精度および倍精度の浮動小数点データ要素の両方に対する非正規（DAZ=0の場合）および無効に対して生成され得る。

【0162】

本発明の一実施形態による方法は、図16に示される。本方法は、上記のアーキテクチャの文脈内で実装され得るが、いずれの特定のアーキテクチャにも限定されない。

【0163】

1601において、数学関数から入力取得される。例えば、数学関数が除算である場合、入力は、分子(x)および分母(y)を含む。数学関数が平方根である場合、入力は、平方根が必要とされる数(x)を含む。1602において、数学関数に応じて入力の様々な特性が評価され、代替的/低速の命令シーケンスにジャンプを行うべきか否かを示す出力を生成する。例えば、上述のように、出力は、特別の場合の入力のデフォルト応答によるベクトル出力（例えば、ゼロによる除算は、INFを返す）、（例えば、マスクK2における）ファーストパスアルゴリズムを用いて処理され得る要素をシグナリングするマスク出力、および計算毎の例外フラグを含み得る。更に、1602において、数学関数のうちの少なくともいくつかの結果がトリビアルなものか否かについて判断される。

【0164】

1603において少なくともいくつかの結果がトリビアルなものであることが判断された場合、次に1604において、トリビアルな結果が計算され、返される。いくつかの例において、トリビアルな結果以外の結果は、必要とされない。更なる結果が必要とされる場合、1605において1602の結果に基づいて、代替的/低速の命令シーケンスにジャンプするか否かについて判断される。唯一の結果が1604におけるトリビアルな結果である場合、代替的/低速シーケンスは保証されない。代替的/低速シーケンスが保証される場合、1606において、代替的/低速の命令シーケンスが実行される。1605に

10

20

30

40

50

おける判断が代替的/低速の命令シーケンスは、保証されないというものである場合、1607において、通常/高速命令シーケンスが実行される（例えば、通常のオペランド値を用いて除算、平方根、または他の数学関数の結果を計算する）。

【0165】

上記の明細書において、本発明の実施形態は、その特定の例示的な実施形態を参照して説明されている。しかし、添付の特許請求の範囲に記載される本発明のより広い趣旨および範囲を逸脱することなく、様々な修正および変更がなされ得ることは明らかであろう。従って、明細書および図面は、限定的意味ではなく、例示の意味で顧慮されるものである。

【0166】

本発明の実施形態は、上記の様々な段階を含み得る。段階は、汎用または専用プロセッサに段階を実行させるべく用いられ得る機械実行可能命令で実施され得る。あるいは、これらの段階は、段階を実行するためのハードワイヤードロジックを含む特定のハードウェアコンポーネントにより、またはプログラミングされたコンピュータコンポーネントおよびカスタムハードウェアコンポーネントの任意の組み合わせにより、実行され得る。

【0167】

本明細書において説明されるように、命令は、特定のオペレーションを実行するよう構成され、または予め定められた機能または非一時的コンピュータ可読媒体で実施されるメモリに格納されたソフトウェア命令を有する特定用途向け集積回路（ASIC）等のハードウェアの特定の構成を指し得る。従って、図面に示される技術は、1または複数の電子デバイス（例えば、終端局およびネットワーク要素等）上に格納され、実行されるコードおよびデータを用いて実装され得る。そのような電子デバイスは、非一時的コンピュータ機械可読ストレージ媒体（例えば、磁気ディスク、光ディスク、ランダムアクセスメモリ、リードオンリメモリ、フラッシュメモリデバイス、相変化メモリ）および一時的コンピュータ機械可読通信媒体（例えば、電氣的、光、音響、もしくは搬送波、赤外線信号、デジタル信号等の他の形態の伝搬信号）等のコンピュータ機械可読媒体を用いてコードおよびデータを格納し、（内部でおよび/またはネットワークを介する他の電子デバイスと）通信する。更に、そのような電子デバイスは通常、1もしくは複数のストレージデバイス（非一時的機械可読ストレージ媒体）、ユーザ入出力デバイス（例えば、キーボード、タッチスクリーン、および/またはディスプレイ）、ならびにネットワーク接続等、1もしくは複数の他のコンポーネントに結合された1もしくは複数のプロセッサのセットを含む。プロセッサおよび他のコンポーネントのセットの結合は通常、1または複数のバスおよびブリッジ（バスコントローラとも呼ばれる）を介して行われる。ストレージデバイスおよびネットワークトラフィックを搬送する信号は、各々、1または複数の機械可読ストレージ媒体および機械可読通信媒体を表す。従って、所与の電子デバイスのストレージデバイスは通常、当該電子デバイスの1または複数のプロセッサのセット上で実行するためのコードおよび/またはデータを格納する。勿論、本発明の実施形態の1または複数の部分は、ソフトウェア、ファームウェア、および/またはハードウェアの異なる組み合わせを用いて実装され得る。この詳細な説明を通じて、説明の目的のために、多数の具体的な詳細が、本発明の完全な理解を提供するべく記載された。しかし、当業者には本発明がこれら具体的な詳細のいくつかがなくても実施され得ることが明らかであろう。特定の例において、周知の構造および機能は、本発明の主題を不明瞭にするのを避けるべく、精巧詳細に説明されていない。従って、本発明の範囲および趣旨は、以下の特許請求の範囲の観点から判断されるべきである。

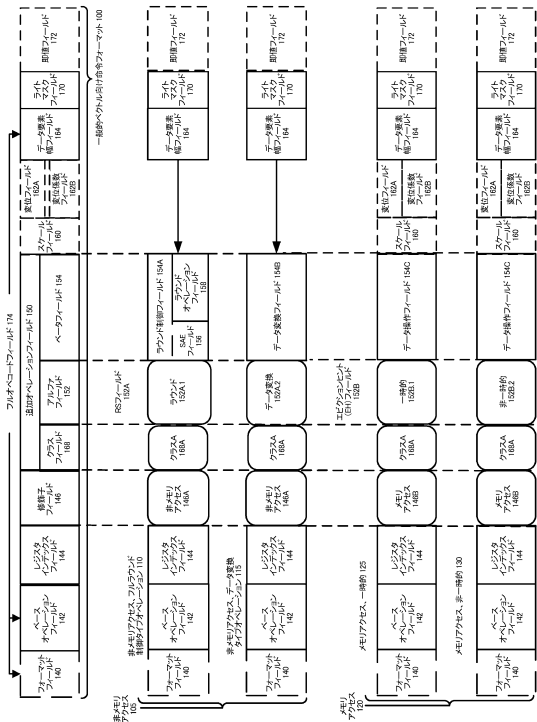
10

20

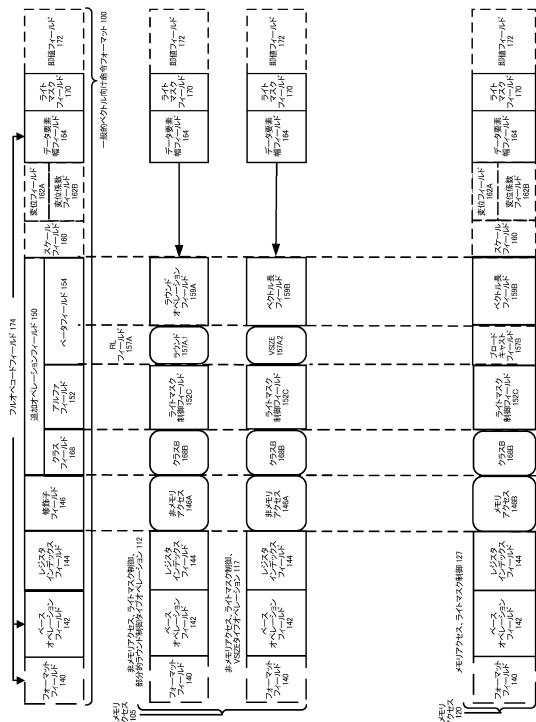
30

40

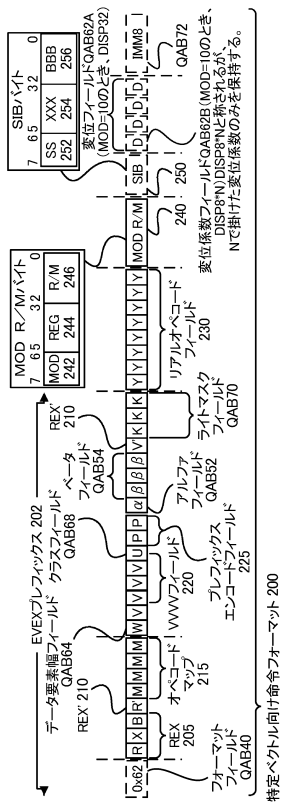
【図1A】



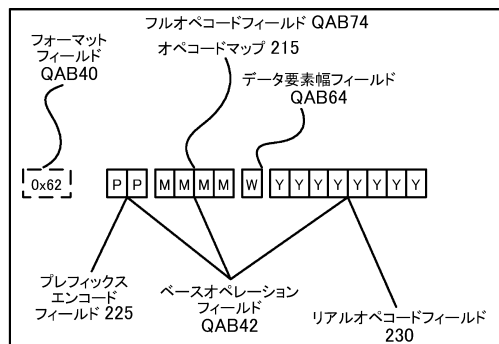
【図1B】



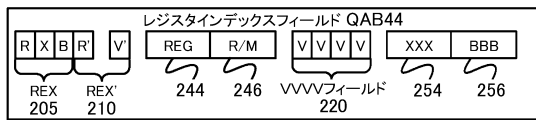
【図2A】



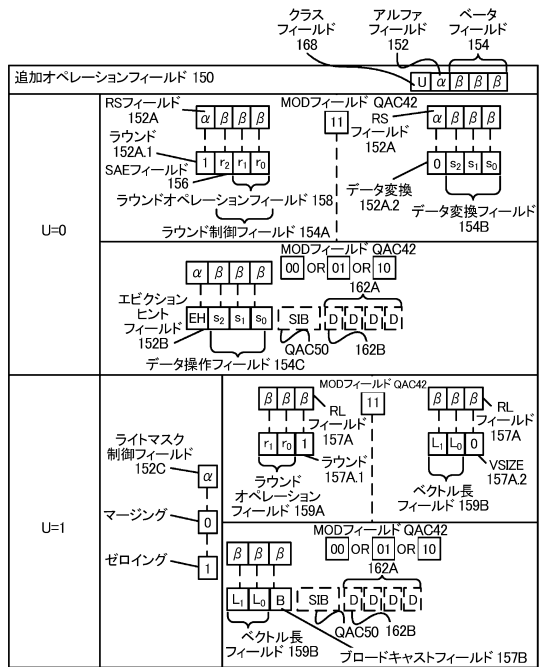
【図2B】



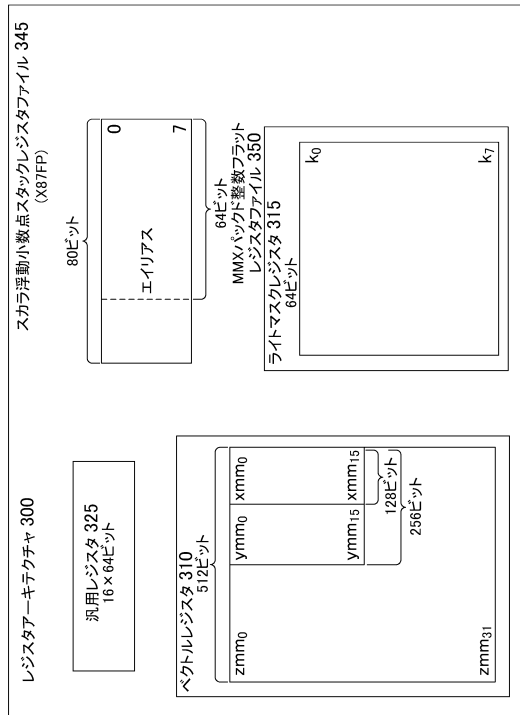
【図2C】



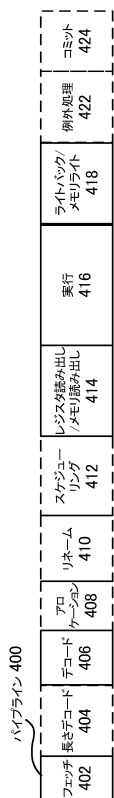
【図 2 D】



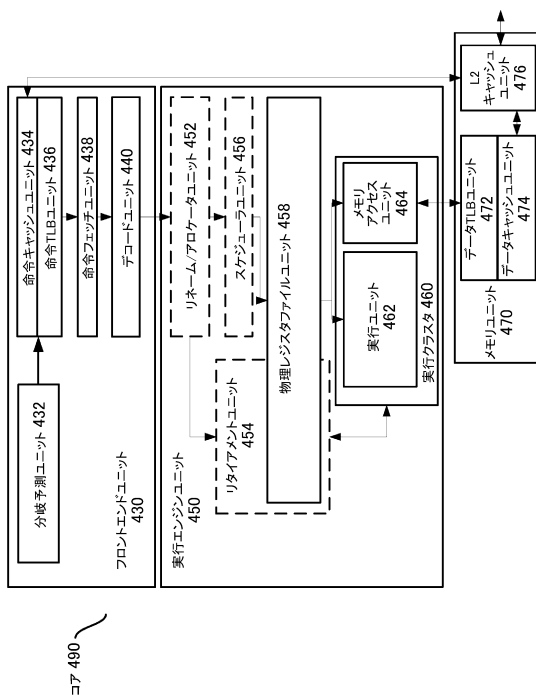
【図 3】



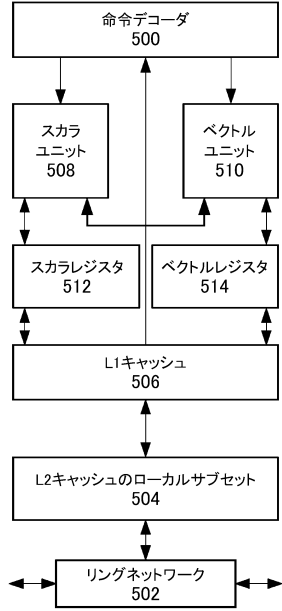
【図 4 A】



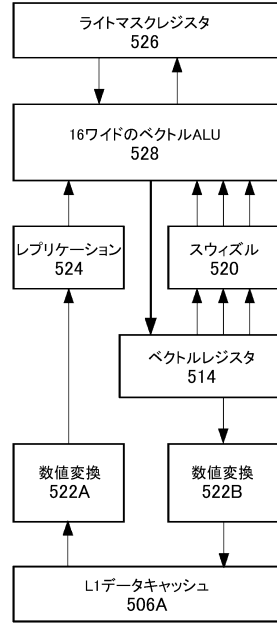
【図 4 B】



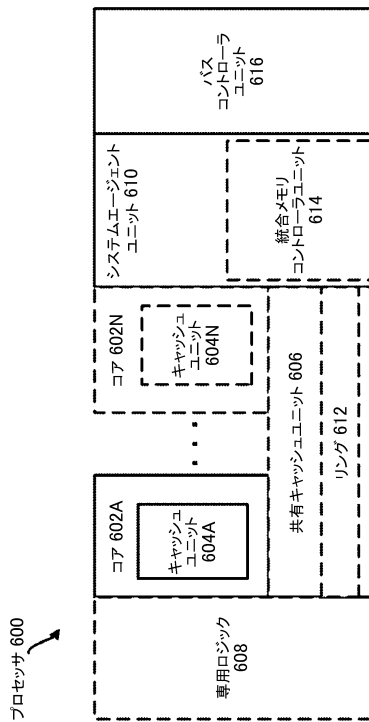
【図5A】



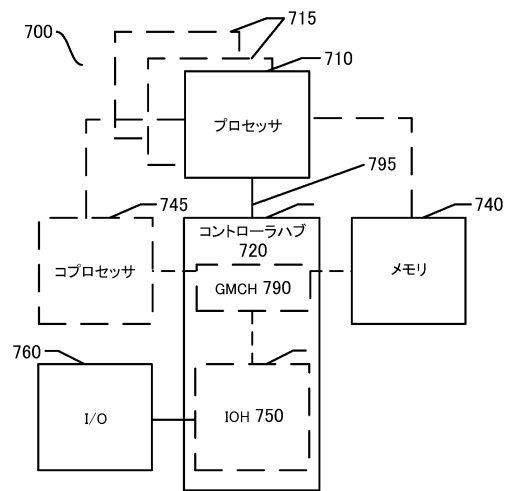
【図5B】



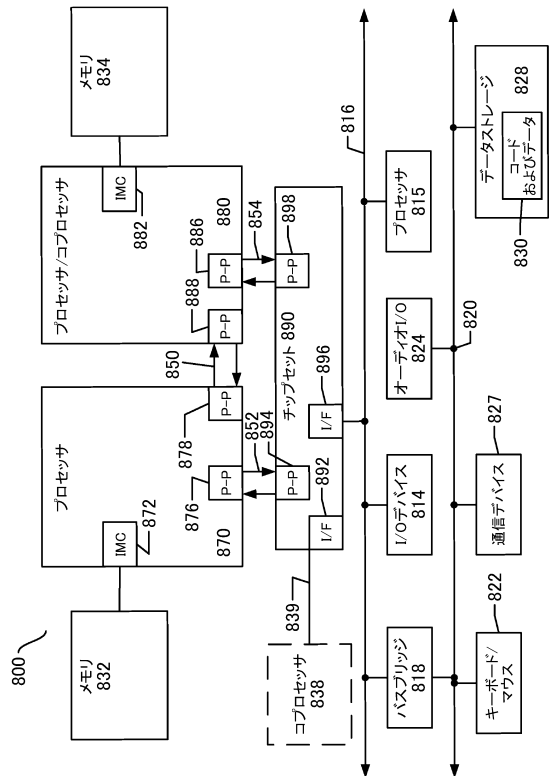
【図6】



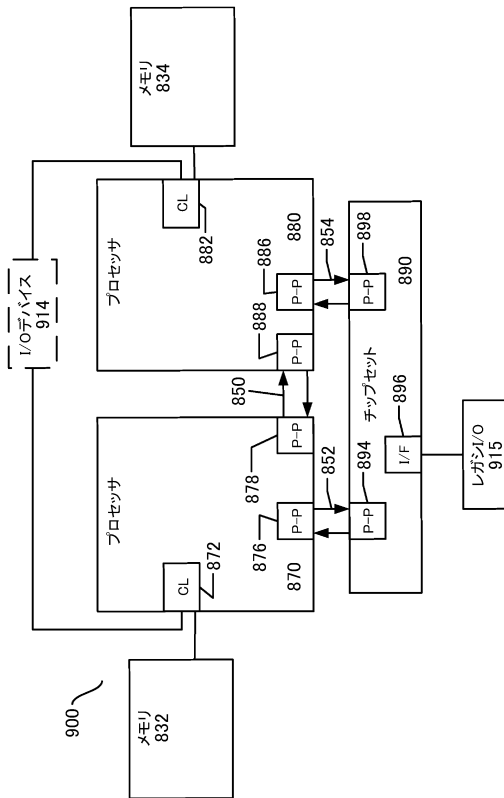
【図7】



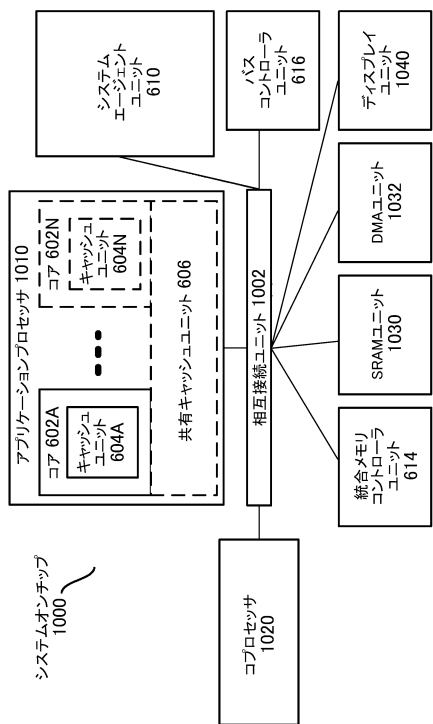
【 図 8 】



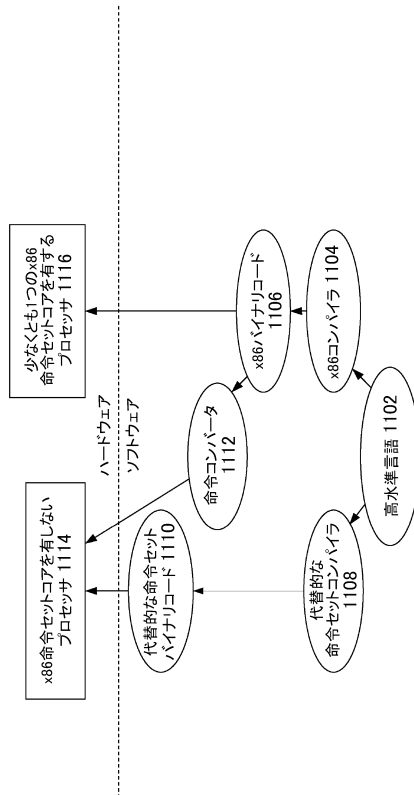
【 図 9 】



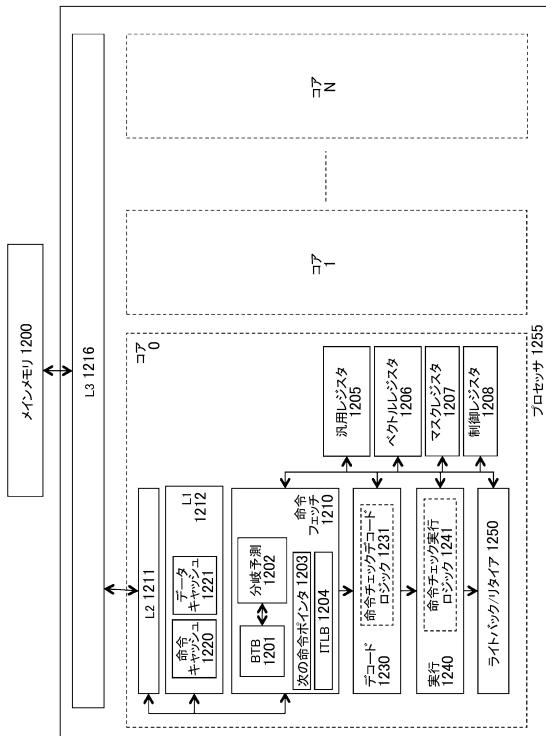
【 図 10 】



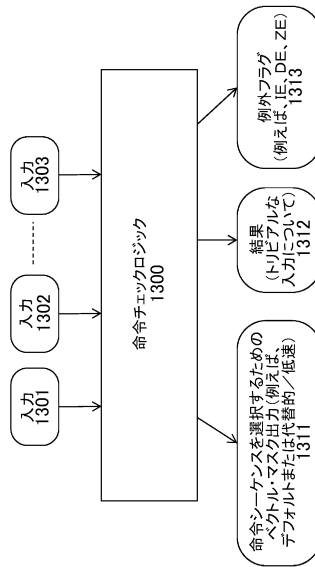
【 図 11 】



【 図 1 2 】



【 図 1 3 】



【 図 1 4 】

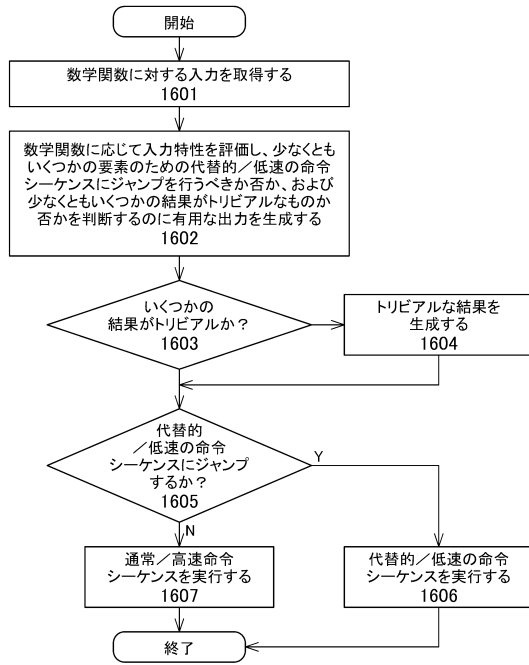
xy	sd	sdnorm	snf	qNaN	sNaN
sd	(0, QNaN, Ind, #E)	(0, sd, #NE)	(0, sd, None)	(0, y, #None)	(0, QNaN(f), #E)
sdnorm	(0, snf, ZED)	(0, -1, #NE)	(0, sd, #NE)	(0, y, #None)	(0, QNaN(f), #E)
snorm	(0, snf, ZED)	(0, -1, #DE)	See below	(0, y, #None)	(0, QNaN(f), #E)
snf	(0, snf, #None)	(0, snf, #NE)	(0, QNaN, Ind, #E)	(0, y, #None)	(0, QNaN(f), #E)
qNaN	(0, s, #None)	(0, s, #None)	(0, s, #None)	(0, s, #None)	(0, s, #E)
sNaN	(0, QNaN(s), #E)	(0, QNaN(s), #E)	(0, QNaN(s), #E)	(0, QNaN(s), #E)	(0, QNaN(s), #E)

Fig. 14

【 図 1 5 】

入力(x)	kdst, vdst, flags	タイプ
-norm (including -infinity)	(0, NaN, #IE)	特殊数
-denorm	(0, NaN, #IE)	特殊数
0	(0, x, None)	特殊数
+0	(0, x, None)	特殊数
+denorm	(0, +1.0, #DE)	スローフロー
+infinity	(0, x, None)	特殊数
qNaN	(0, x, None)	特殊数
sNaN	(0, qNaN(x), #E)	特殊数
BiasExponent(x) ≤ 2 ^{precision}	(0, +1.0, None)	スローフロー
BiasExponent(x) = 2 ^{precision}	(0, +1.0, None)	スローフロー
Other	(1, -0.0, None)	フーストフロー

【図16】



フロントページの続き

- (51)Int.Cl. F I
G 0 6 F 9/32 (2006.01) G 0 6 F 9/302 C
 G 0 6 F 9/318 A
 G 0 6 F 9/32 3 2 0 F
- (72)発明者 ハネク、ロバート エヌ .
 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・ミッション カレッジ ブレー
 バード・2 2 0 0 インテル・コーポレーション内
- (72)発明者 ファーガスン、ワーレン イー .
 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・ミッション カレッジ ブレー
 バード・2 2 0 0 インテル・コーポレーション内
- (72)発明者 バーラミ、タラネー
 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・ミッション カレッジ ブレー
 バード・2 2 0 0 インテル・コーポレーション内
- (72)発明者 テベット、アヴィ エイ .
 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・ミッション カレッジ ブレー
 バード・2 2 0 0 インテル・コーポレーション内
- (72)発明者 ブラッドフォード、デニス アール .
 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・ミッション カレッジ ブレー
 バード・2 2 0 0 インテル・コーポレーション内
- (72)発明者 フェリー、マイケル
 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・ミッション カレッジ ブレー
 バード・2 2 0 0 インテル・コーポレーション内
- (72)発明者 ジャン、ジンヴェ
 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・ミッション カレッジ ブレー
 バード・2 2 0 0 インテル・コーポレーション内

審査官 豊田 真弓

- (56)参考文献 特表2 0 0 2 - 5 0 8 8 6 4 (J P , A)
 特開昭6 2 - 2 3 6 0 7 3 (J P , A)
 特開2 0 0 1 - 2 4 3 0 6 3 (J P , A)
 特開平1 0 - 1 4 3 3 5 5 (J P , A)
 特開平0 7 - 2 4 8 8 9 7 (J P , A)
 特開2 0 0 0 - 1 2 2 8 5 0 (J P , A)
 特開2 0 0 0 - 0 4 0 0 8 1 (J P , A)

(58)調査した分野(Int.Cl. , D B名)

G 0 6 F 7 / 5 7
 G 0 6 F 9 / 3 0
 G 0 6 F 9 / 3 0 2
 G 0 6 F 9 / 3 1 8
 G 0 6 F 9 / 3 2
 G 0 6 F 9 / 3 8