

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5283810号
(P5283810)

(45) 発行日 平成25年9月4日 (2013.9.4)

(24) 登録日 平成25年6月7日 (2013.6.7)

(51) Int. Cl.

F I

G 0 6 F 9/38 (2006.01)

G 0 6 F 9/38 3 7 0 C

G 0 6 F 9/30 (2006.01)

G 0 6 F 9/30 3 5 0 A

請求項の数 17 (全 34 頁)

(21) 出願番号 特願2001-503043 (P2001-503043)
 (86) (22) 出願日 平成12年6月15日 (2000.6.15)
 (65) 公表番号 特表2003-502728 (P2003-502728A)
 (43) 公表日 平成15年1月21日 (2003.1.21)
 (86) 国際出願番号 PCT/GB2000/002331
 (87) 国際公開番号 W02000/077627
 (87) 国際公開日 平成12年12月21日 (2000.12.21)
 審査請求日 平成19年6月4日 (2007.6.4)
 審判番号 不服2011-22822 (P2011-22822/J1)
 審判請求日 平成23年10月24日 (2011.10.24)
 (31) 優先権主張番号 99304659.8
 (32) 優先日 平成11年6月15日 (1999.6.15)
 (33) 優先権主張国 欧州特許庁 (EP)

(73) 特許権者 511076424
 ヒューレット・パカード デベロップメン
 ト カンパニー エル. ビー.
 Hewlett-Packard Development Company, L
 . P.
 アメリカ合衆国 テキサス州 77070
 ヒューストン コンパック センタ ド
 ライブ ウェスト 11445
 (74) 代理人 110000039
 特許業務法人アイ・ピー・エス
 (72) 発明者 オルギエッティ・アンドレア
 イギリス、ビーエス16、1エーディー、
 ブリストル、パークレイズ・ヒル 20

最終頁に続く

(54) 【発明の名称】 プロセッサおよびコプロセッサを含むコンピュータ・システム

(57) 【特許請求の範囲】

【請求項 1】

第1のプロセッサと、

前記第1のプロセッサへのコプロセッサとして使用するための第2のプロセッサと、
メモリと、バースト命令に従ってデータ・バーストにおいて前記メモリに書き込みまたは読み取る
データを入れるための、少なくとも1つのデータ・バッファと、

前記バースト命令を実行するためのバースト・コントローラと、

前記バースト・コントローラによる実行のためにバースト命令を順番に提供するための
バースト命令エレメントと、コプロセッサ命令およびバースト命令の実行を、前記コプロセッサ命令およびバースト
命令が実行するデータの可用性により、同期化するための同期化機構と

を含み、

バースト命令が前記第1のプロセッサによって前記バースト命令エレメントへ提供され
、前記バースト・コントローラによって実行されたバースト命令に従って前記少なくとも
1つのデータ・バッファを通じて、データが、前記第2のプロセッサへの入力データとし
て前記メモリから読み取られ、前記第2のプロセッサからの出力データとして前記メモリ
へ書き込まれ、

前記同期化機構は、

特定のバースト命令の実行により増分され、特定のコプロセッサ命令の実行により減分

される第 1 のカウンタを少なくとも含み、

前記第 1 のカウンタを、第 1 の低しきい値を越えてさらに減分することができないとき、前記第 2 のプロセッサの関連付けられた実行のためのコプロセッサ命令がストールあるいは防止され、

前記第 1 のカウンタを、第 1 の高しきい値を越えてさらに増分することができないとき、前記少なくとも 1 つのバッファから前記メモリへのデータの関連付けられた格納のためのバースト命令がストールあるいは防止される

コンピュータ・システム。

【請求項 2】

前記第 2 のプロセッサの実行を順番に制御するためのコプロセッサ命令を提供するためのコプロセッサ命令エレメントをさらに含み、前記コプロセッサ命令が前記第 1 のプロセッサによって提供される、

請求項 1 に記載のコンピュータ・システム。

【請求項 3】

コプロセッサ・コントローラをさらに含み、

前記コプロセッサ・コントローラがコプロセッサ命令を前記コプロセッサ命令エレメントから受信し、前記第 2 のプロセッサの実行を、受信されたコプロセッサ命令に従って制御し、前記コプロセッサと前記少なくとも 1 つのデータ・バッファの間の通信を制御する、

請求項 2 に記載のコンピュータ・システム。

【請求項 4】

前記同期化機構が、前記少なくとも 1 つのデータ・バッファにまだロードされていないデータにおける前記第 2 のプロセッサの実行を必要とするコプロセッサ命令の実行をブロックするように適合され、前記少なくとも 1 つのデータ・バッファから前記メモリへのデータの格納のためのバースト命令の実行を、このようなデータが前記第 2 のプロセッサによって前記少なくとも 1 つのデータ・バッファへ提供されていない場合にブロックするように適合される、

請求項 1 に記載のコンピュータ・システム。

【請求項 5】

前記同期化機構は、

特定のコプロセッサ命令の実行により増分され、特定のバースト命令の実行により減分される第 2 のカウンタをさらに含み、

前記第 2 のカウンタを、第 2 の低しきい値を越えてさらに減分することができないとき、前記少なくとも 1 つのバッファから前記メモリへのデータの関連付けられた格納のためのバースト命令がストールあるいは防止される、

請求項 1 に記載のコンピュータ・システム。

【請求項 6】

前記第 2 のカウンタを、第 2 の高しきい値を越えてさらに増分することができないとき、前記第 2 のプロセッサの関連付けられた実行のためのコプロセッサ命令がストールあるいは防止される、

請求項 5 に記載のコンピュータ・システム。

【請求項 7】

前記バースト命令エレメントは、命令キューである、

請求項 1 に記載のコンピュータ・システム。

【請求項 8】

前記バースト命令エレメントは、さらに加えられたプロセッサである、

請求項 1 に記載のコンピュータ・システム。

【請求項 9】

前記バースト命令エレメントは、プログラム可能な状態機械である、

請求項 1 に記載のコンピュータ・システム。

10

20

30

40

50

【請求項 10】

前記第1のプロセッサは、コンピュータ装置の中央処理装置である、
請求項1に記載のコンピュータ・システム。

【請求項 11】

コンピュータ・システムを動作する方法であって、

第1のプロセッサ、および前記第1のプロセッサへのコプロセッサとして動作する第2のプロセッサによる実行のためのコードを提供することと、

前記第2のプロセッサによって実行されるタスクを提供することとしての、前記コードの一部の識別することと、

前記タスクを提供するコードを、コプロセッサ・コントローラによる実行のためのコプロセッサ命令で置換すること

を含み、

前記コプロセッサ命令は、

前記第2のプロセッサによる前記タスクの実行を制御するように決定され、

前記コードおよび前記タスクから、少なくとも1つのデータ・バッファにより、前記第2のプロセッサによるアクセスのためにデータ・バーストにおいてメイン・メモリからデータを読み取り、そこへ書き込むことができるようにするためのバースト命令を決定することと、

前記少なくとも1つのデータ・バッファと前記メイン・メモリの間でデータの転送を制御するバースト・コントローラによるバースト命令の実行と共に、前記コプロセッサ上で前記タスクを実行することと

を含み、

前記タスクの実行において、コプロセッサ命令の実行とバースト命令の実行の間の同期化が、同期化機構によって達成され、

前記同期化機構は、

特定のバースト命令の実行により増分され、特定のコプロセッサ命令の実行により減分される第1のカウンタを少なくとも含み、

前記第1のカウンタを、第1の低しきい値を越えてさらに減分することができないとき、前記第2のプロセッサの関連付けられた実行のためのコプロセッサ命令がストールあるいは防止され、

前記第1のカウンタを、第1の高しきい値を越えてさらに増分することができないとき、前記少なくとも1つのバッファから前記メモリへのデータの関連付けられた格納のためのバースト命令がストールあるいは防止される

方法。

【請求項 12】

バースト命令を決定する前記ステップが、前記バースト命令を、前記第1のプロセッサによって実行される前記コードの一部内に含めることをさらに含む、

請求項11に記載の方法。

【請求項 13】

バースト命令を決定する前記ステップが、前記コードから前記第2のプロセッサによってアクセスされるメモリ・アドレスを決定すること、および少なくとも1つのデータ・バッファにより、前記第2のプロセッサによるアクセスのためにデータ・バーストにおいてメイン・メモリからデータを読み取り、そこへ書き込むことができるように、前記第2のプロセッサによって行われるメモリ・アクセスを編成することをさらに含む、

請求項11に記載の方法。

【請求項 14】

前記同期化機構が、前記第1の命令の正しい実行のために完了が必要である第2の命令が完了するまで、前記第1の命令をブロックすることを含む、

請求項11に記載の方法。

【請求項 15】

前記コプロセッサ命令エレメントは、命令キューである、
請求項 2 に記載のコンピュータ・システム。

【請求項 16】

前記コプロセッサ命令エレメントは、さらに加えられたプロセッサである、
請求項 2 に記載のコンピュータ・システム。

【請求項 17】

前記コプロセッサ命令エレメントは、プログラム可能な状態機械である、
請求項 2 に記載のコンピュータ・システム。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明は、メイン・プロセッサおよびコプロセッサを含むコンピュータ・アーキテクチャに関し、詳細にはこのようなアーキテクチャにおけるコプロセッサによるメモリ・リソースの使用に関する。

【背景技術】

【0002】

マイクロプロセッサをベースとするコンピュータ・システムは、典型的には、CPUのような汎用マイクロプロセッサに基づいている。このようなマイクロプロセッサは、幅広い範囲の計算タスクを処理するように十分に適合されるが、これらは必然的にすべてのタスクに合わせて最適化されるのではない。タスクが計算に集中する場合（媒体処理など）、CPUはしばしば実行に困難を生じる。

20

【0003】

この問題への標準の手法の 1 つは、個々の計算的に困難なタスクを処理するように特に適合されたコプロセッサを使用することである。このようなコプロセッサは、ASIC（特定用途向け IC）を使用して構築することができる。これらは特定の計算タスクのために構築され、したがってこのようなタスクに合わせて最適化することができる。しかし、これらは使用において柔軟性がなく（これらが特定のタスクのみのために設計されるので）、一般に製造に時間がかかる。解決策としては、FPGA（フィールド・プログラム可能ゲート・アレイ）など、所与の計算タスクに特に適合された構成でプログラムすることができる、柔軟性のあるハードウェアの構築がある。このような構造が構成可能であるだけでなく、再構成可能である場合は、さらなる柔軟性が達成される。このような再構成可能な構造の一例がCHES アレイであり、これは、国際特許出願第GB98/00262号、国際特許出願第GB98/00248号、1998年12月11日出願の米国特許出願第09/209,542号、およびその欧州の相当する欧州特許出願第98309600.9号において論じられている。

30

【発明が解決しようとする課題】

【0004】

このようなコプロセッサの使用は、このような計算の効率を著しく改善することができるが、従来のアーキテクチャの構成がコプロセッサの有効性を抑制する可能性がある。計算をなおより有効にコプロセッサに移すことができる装置を達成することが、特にこれらの計算が大量のデータを処理することを含む場合、望ましい。

40

【課題を解決するための手段】

【0005】

したがって、第 1 のプロセッサ、第 1 のプロセッサへのコプロセッサとして使用するための第 2 のプロセッサ、メモリ、バースト命令に従ってデータ・バーストにおいてメモリへ書き込まれるかあるいはそこから読み取られるデータをバッファに入れるための少なくとも 1 つのデータ・バッファ、バースト命令を実行するためのバースト・コントローラ、バースト・コントローラによる実行のためにバースト命令を順番に提供するためのバースト命令エレメントを含み、それによってバースト命令が第 1 のプロセッサからバースト命令エレメントへ提供され、バースト・コントローラによって実行されたバースト命令に従

50

って少なくとも1つのデータ・バッファを通じて、データが第2のプロセッサによってメモリから読み取られ、そこへ書き込まれるコンピュータ・システムが提供される。

【0006】

この装置は特に、コプロセッサが大きいブロックのデータで作業する場合、特にこのようなブロックのメモリ・アドレスが定期的に変わる場合に有利である。この装置により、このようなブロックを、メイン・プロセッサ（それらを使用するにあまり適していないシステム・コンポーネント・エレメントである）の関与を最小にして、有効にメイン・メモリに出し入れすることができる。

【0007】

コプロセッサがデータ・バッファと類似の方法で制御される場合、特に効率的な構造を達成することができる。これは、第2のプロセッサの実行を制御するためのコプロセッサ命令を順番に提供するためのコプロセッサ命令エレメントにより行うことができる（コプロセッサ命令は初めに第1のプロセッサによって提供される）。コプロセッサ・コントローラがコプロセッサ命令をコプロセッサ命令エレメントから受信し、それに従って第2のプロセッサの実行を制御するのがよい。このコプロセッサ・コントローラは、コプロセッサと少なくとも1つのデータ・バッファの間の通信を制御することができ、たとえば、バスがコプロセッサ・コントローラとデータバッファの間に存在する場合、コプロセッサ・コントローラが、第2のプロセッサからバスに別々のデータ・ストリームの出入りのアクセスを制御することができる。

【0008】

コプロセッサおよびバースト命令が実行するデータの可用性により、コプロセッサとバースト命令の実行を同期化するための同期化機構がある場合、特定の利点を得ることができる。これは、コプロセッサがコプロセッサ命令に基づいて実行する場合、特に十分に実施される。有効な手法は、データ・バッファにまだロードされていないデータに対し第2のプロセッサの実行を必要とするコプロセッサ命令の実行をブロックし、データが第2のプロセッサによってデータ・バッファへ提供されていない場合、データ・バッファからメモリへのデータの格納のためのバースト命令の実行をブロックするように、同期化機構を適合させることである。同期化機構を実行するための特に有効な方法は、カウンタを使用することであり、このカウンタは、適切なバーストおよびコプロセッサ命令を通じて増分あるいは減分することができ、特定の命令をさらに減分できない場合はブロックする。

【0009】

さらなる態様では、本発明はコンピュータ・システムを動作する方法を提供し、これは、第1のプロセッサによる実行のためのコードを提供すること、第1のプロセッサへのコプロセッサとして動作する第2のプロセッサによって実行されるタスクをコードから抽出すること、コードおよびタスクから、少なくとも1つのデータ・バッファにより、第2のプロセッサによるアクセスのためにデータ・バーストにおいてメイン・メモリからデータを読み取り、そこへ書き込むことができるようにするためのバースト命令を決定すること、および少なくとも1つのデータ・バッファとメイン・メモリの間でデータの転送を制御するバースト・コントローラによるバースト命令の実行と共に、コプロセッサ上のタスクの実行を含む。

【0010】

コードからのタスクの抽出に続いて、コプロセッサ・コントローラによる実行のためのコプロセッサ命令が、第2のプロセッサによるタスクの実行を制御するために決定されるのがよい。

【0011】

タスクの実行の場合、コプロセッサ命令の実行とバースト命令の実行の間の同期化が同期化機構によって達成されるのがよい。この同期化機構は、第1の命令の正しい実行のために完了が必要である第2の命令が完了するまで、第1の命令をブロックすることを、有効に含むことができる。この機構は、適切なバーストまたはコプロセッサ命令を通じて増分あるいは減分することができるカウンタを使用することができる。

【 0 0 1 2 】

本発明の特定の実施形態を添付の図面を参照して、以下に説明する。

【 発明を実施するための形態 】

【 0 0 1 3 】

図 1 は、本発明の第 1 の実施形態によるシステムの基本エレメントを示す。本質的に、このシステムはプロセッサ 1 およびコプロセッサ 2 を含み、これらは、最大の計算効率のために、計算をプロセッサ 1 とコプロセッサ 2 の間で区分することができるよう構成されている。プロセッサ 1 は本質的にいかなる汎用プロセッサ（たとえば、i 9 6 0）でもよく、コプロセッサ 2 は本質的に、著しくより高い有効性により計算の一部を処理することができるいかなるコプロセッサでもよい。ここで記載された特定のシステムでは、本質的に計算全体がプロセッサ 1 によるよりもコプロセッサ 2 によって処理されるが、本発明はこの特定の構成に限定されるものではない。

10

【 0 0 1 4 】

特に記載されたシステムでは、コプロセッサ 2 が再構成可能 F P G A の形式であり、これについては以下でさらに述べられるが、コプロセッサ 2 の他の形式、たとえば、A S I C S、D S Pなどを代りに使用することができる（対応する修正を計算モデルに行うことが必要）。プロセッサ 1 およびコプロセッサ 2 は共に D R A M メイン・メモリ 3 へのアクセスを有するが、プロセッサ 1 はより高速なアクセス・メモリ 4 のキャッシュ、典型的には S R A M へのアクセスも有する。D R A M 3 への効率的なアクセスが、情報の「バースト」の効率的なローディングおよび格納のために D R A M と通信するように適合された「バースト・バッファ」メモリ 5 によって提供され、バースト・バッファについては以下でさらに記載される。バースト・バッファ 5 への命令が、バースト命令待ち行列 6 を通じて提供され、バースト・バッファ 5 がバースト・バッファ・コントローラ 7 の制御下で動作する。バースト・バッファのアーキテクチャは、以下に述べられる理由のため、コプロセッサ 2 に関連付けられたアーキテクチャにおいて、ミラーリングされる。コプロセッサ 2 への命令がコプロセッサ命令待ち行列 8 において提供され、コプロセッサがコプロセッサ・コントローラ 9 の制御下で動作する。バースト・バッファおよびコプロセッサの動作、およびそれらの関連付けられた命令待ち行列の同期化が、プロセッサ 1 自体による一般の方法ではなく、特定の機構によって達成される。この実施形態では、この機構がロード / 実行セマフォ (semaphore) 1 0 および実行 / 格納セマフォ 1 1 を含み、これらは以下に記載される方法で動作する（他のこのような同期化機構が可能であり、これらについても以下に述べられる）。

20

30

【 0 0 1 5 】

システム・アーキテクチャにおけるエレメントの説明

このシステムの個々のエレメントが、以下でより詳細に論じられる。プロセッサ 1 は一般に計算を制御するが、計算自体におけるステップのいくつか（または、記載された実施形態では、すべて）がコプロセッサ 2 において実行されるような方法で制御する。プロセッサ 1 が、バースト命令待ち行列 6 を通じて特定のタスクのための命令、すなわち、バースト・バッファ・コントローラ 7 の構成、およびバースト・バッファ・メモリ 5 とメイン・メモリ 3 の間のデータの転送のための命令を提供する。さらに、コプロセッサ命令待ち行列 8 を通じて、プロセッサ 1 はさらなるタスクのための命令、すなわち、コプロセッサ・コントローラ 9 の構成、およびコプロセッサ 2 上の計算の開始のための命令を提供する。コプロセッサ 2 上で実行されるこの計算は、バースト・バッファ・メモリ 5 を通じてデータにアクセスする。

40

【 0 0 1 6 】

コプロセッサ命令待ち行列 8 の使用が有効にプロセッサ 1 をプロセッサ 2 の動作から減結合し、バースト命令待ち行列 6 の使用が有効にプロセッサ 1 をバースト・バッファ 5 から減結合する。この構成の特定の詳細は、以下でより詳細に論じられる。この減結合については、本発明のこの実施形態のための計算モデルに関連して、以下でさらに論じられる。

50

【 0 0 1 7 】

コプロセッサ 2 が、実際の計算のいくつかまたはすべてを実行する。特に適したコプロセッサは C H E S S F P G A 構造であり、これは、国際特許出願第 G B 9 8 / 0 0 2 6 2 号、国際特許出願第 G B 9 8 / 0 0 2 4 8 号、1 9 9 8 年 1 2 月 1 1 日出願の米国特許出願第 0 9 / 2 0 9 , 5 4 2 号、およびその欧州の相当する欧州特許出願第 9 8 3 0 9 6 0 0 . 9 号に記載されており、それらの出願の内容が参照により本明細書に組み込まれる。このコプロセッサは再構成可能であり、4 ビット A L U のチェッカーボード・アレイおよびスイッチング構造を含み、それにより、コプロセッサが構成可能であり、1 つの 4 ビット A L U からの出力を使用して別の A L U へ命令することができる。C H E S S アーキテクチャは特に、パイプライン計算に有効であり、ここでは入力および出力データ・ストリームと対話するように有効に適合される。コプロセッサ・コントローラ 9 (この動作が以下でさらに論じられる) が、高レベルの制御命令 (計算の詳細に関する命令ではなく、コプロセッサ 2 の制御全体のための命令、たとえば「n サイクル実行する」) を、コプロセッサ命令待ち行列 8 から受信する。C H E S S コプロセッサ 2 はコプロセッサ・コントローラ 9 の制御下で動作し、バッファ・バースト 5 との対話を通じてデータを受信し、格納する。したがって、C H E S S コプロセッサ 2 は入力ストリーム上で動作して、出力ストリームを生成する。これは、C H E S S コプロセッサの動作が非常に予測可能なので、効率的なプロセスにすることができる。このモデルに従った計算の詳細な動作が、後で論じられる。

10

【 0 0 1 8 】

20

プロセッサ 1 が、従来の方法による S R A M における高速アクセス・メモリ・キャッシュ 4 へのアクセスを有するが、メイン・メモリは D R A M 3 として提供される。D R A M への有効なアクセスが、バースト・バッファ 5 によって提供される。バースト・バッファは、欧州特許出願第 9 7 3 0 9 5 1 4 . 4 号、および 1 9 9 8 年 1 月 6 日出願の対応する米国特許出願第 0 9 / 3 , 5 2 6 号において記載されており、それらの出願が参照により本明細書に組み込まれる。バースト・バッファ・アーキテクチャについては本明細書で簡単に記載されるが、このアーキテクチャの十分な詳細については、先の出願を参照していただきたい。

【 0 0 1 9 】

30

この実施形態で使用されたバースト・バッファ・アーキテクチャのバージョンの要素 (前記の出願において論じられるように、変形も可能) が、図 2 および図 3 に示される。バースト・バッファ構成要素がプロセッサ 1 と通信する接続 1 2 が提供される。メモリ・バス 1 6 が、メイン・メモリ 3 (図 2 においては図示せず) への接続を提供する。このメモリ・バスはキャッシュ 4 と共有することができ、この場合、メモリ・データバス・アービタ 5 8 が、キャッシュ 4 との通信も可能にする。

【 0 0 2 0 】

この装置におけるバースト・バッファの全体の役割は、計算をコプロセッサ 2 上で実行できるようにすることであり、このコプロセッサ 2 とメイン・メモリ 3 の間のデータ転送を、両方が各システム構成要素の効率を最大にすると同時に全体のシステム効率を最大にする方法で行うことを含む。これは、以下のいくつかの技術の組み合わせによって達成される。

40

【 0 0 2 1 】

- 1) 以下に記載されるようなバースト・バッファ 5 を使用した、D R A M へのバースト・アクセス、
- 2) 「ダブル・バッファリング」とよばれる技術を使用した、コプロセッサ 2 上の計算の同時実行、およびメイン・メモリ 3 とバースト・バッファ・メモリ 5 の間のデータ転送、および
- 3) プロセッサ 1 の実行をコプロセッサ 2 およびバースト・バッファ・メモリ 5 の実行から、命令待ち行列の使用を通じて切り離すことである。

【 0 0 2 2 】

50

「ダブル・バッファリング」は、たとえば、コンピュータ・グラフィックスにおいて知られている技術である。本明細書で使用される形式では、バースト・バッファ・メモリ 5 の一部からのデータを消費すること、すなわち読み取ること、他のデータを同じメモリの異なる領域へ生成すること、すなわち書き込むことを含み、先に書き込まれた領域を読み取り、その逆もできるスイッチング機構を有する。

【0023】

バースト・バッファの利点は、従来のDRAM構造の特徴の有効な利用である。DRAMは、正方行列におけるメモリ位置のアレイを含む。アレイにおけるエレメントにアクセスするには、行が最初に選択され（あるいは「開かれ」）なければならない、その後適切な列の選択が続く。しかし、行が選択された後、その行における列への連続アクセスは、単に列アドレスを提供することによって実行することができる。行を開くことおよびその行にローカルな一連のアクセスを実行することの概念が、「バースト」と呼ばれる。媒体集中計算（典型的には、いかなるデータ依存アドレス指定もなしに長いアレイにアクセスする、規則正しいプログラム・ループを使用するアルゴリズムを含む）におけるように、データが規則正しい方法で構成されるとき、バーストの有効な使用が劇的に計算速度を向上させることができる。バースト・バッファは、バーストの効率的な使用を通じてDRAMからのデータにアクセスするように適合された、新しいメモリ構造である。

【0024】

システムがいくつかのバースト・バッファを含むことができる。典型的には、各バースト・バッファがそれぞれのデータ・ストリームへ割り振られる。アルゴリズムが、異なる数のデータ・ストリームを有するので、固定量のSRAM 26がバースト・バッファ・メモリ領域としてバースト・バッファに使用可能であり、この量が、必要とされるバッファの数に従って分割される。たとえば、固定SRAMの量が2Kバイトである場合、かつ、アルゴリズムが4つのデータ・ストリームを有する場合、メモリ領域を4つの512バイトのバースト・バッファに区分することができる。

【0025】

このタイプのアーキテクチャでは、バーストが、以下によって定義されたアドレスの組を含む。

【0026】

【数1】

$$\text{バースト} = \{ B + S \times i \mid B, S, i \in \mathbb{N}, 0 \leq i < L \}$$

【0027】

ただし、Bは転送の基底アドレスであり、Sはエレメントの間のストライド(stride)であり、Lは長さであり、Nは自然数の組である。この式において明示的に定義されないが、バースト順序が、0からL-1まで増分するiによって定義される。したがって、バーストは、次の3個のエレメントからなる集合によって定義することができる。

【0028】

(base__address, length, stride)

ソフトウェアでは、バーストをエレメントサイズによって定義することもできる。これは、バーストをバイト、ハーフワード、またはワードのサイズにすることができることを意味する。ストライドの単位は、このことを考慮しなければならない。「サイズド・バースト」は、次の形式の4個のエレメントからなる集合によって定義される。

【0029】

(base__address, length, stride, size)

「チャンネル・バースト」は、サイズがメモリへのチャンネルの幅である、サイズド・バーストである。コンパイラが、ソフトウェア・サイズド・バーストからチャンネル・バーストへのマッピングを担う。チャンネル・バーストは、次の4個のエレメントからなる集合によって定義することができる。

【0030】

(base__address, length, stride, width)

チャンネル幅が32ビット（または4バイト）である場合、チャンネル・バーストが常に以下の形式である。

【0031】

(base__address, length, stride, 4)

あるいは、3個の要素からなる集合(base__address, length, stride)に短縮される。

【0032】

このメモリの制御およびバースト・バッファの割り振り（および解放）が、ソフトウェア処理によって高レベルで処理される。この実施形態では、「ダブル・バッファリング」が使用されるが、他の戦略も確かに可能であり、この判断は記憶の効率と簡潔性の間のトレードオフを含む。バースト・バッファ・メモリ領域26がメイン・メモリ3からデータをロードし、そこへデータを格納し、これはメモリ・データバス・アービタ58を通じて行い、これがDMAコントローラ56の制御下で動作し、バースト命令待ち行列6を通じて受信される命令にตอบสนองする。データが、バースト・バッファ・メモリ領域26とプロセッサ1またはコプロセッサ2の間で、接続手段12を通じて交換される。図3に示すように、バースト・バッファ・システム5のための制御インターフェースは、1組のテーブルに基づいている。すなわち、バースト・バッファ・メモリへバーストし、そこからバーストするためのメイン・メモリの領域を記述するメモリ・アクセス・テーブル(MAT)65、およびバースト・バッファ・メモリの領域を記述するバッファ・アクセス・テーブル(BAT)66である。この実施形態では、デュアルポートSRAMの同次領域が、バースト・バッファ・メモリ領域26のために使用される。

【0033】

MATおよびBATを使用しなかったバースト・バッファ構成（これも欧州特許出願第97309514.4号に記載されているようなもの）を、本発明の代替実施形態において使用することができ、MATおよびBATにおいて暗示的に符号化されたパラメータ（ソース・アドレス、宛先アドレス、長さ、ストライド）が、次いで、発行されたバースト転送毎に明示的に指定されなければならない。直接のアドレス、長さおよびストライドではなく、MATおよびBATを使用するための主な理由は、これが全体のコード・サイズを著しく減らすことである。本発明に関連して、これは典型的には重要というよりも有用である。

【0034】

プロセッサ1から発信するバースト命令が、バースト命令待ち行列6により、バースト・バッファ5へ提供される。バースト命令待ち行列6からの命令が、バッファ制御要素54によって処理されて、MAT65およびBAT66におけるスロットが参照される。バッファ・コントローラは、8個のバースト制御レジスタ52からの制御入力も受信する。これらの2つのテーブルに含まれた情報が、実行時間で共に結び付けられて、完全なメイン・メモリ対バースト・バッファのトランザクションが記述される。出力がバッファ・コントローラ54から直接メモリ・アクセス(DMA)コントローラ56へ、よってメモリ・データバス・アービタ58へ提供されて、メイン・メモリ3とバースト・バッファ・メモリ領域26の間のトランザクションが実施される。

【0035】

重要なバースト命令は、データをメイン・メモリ3からバースト・バッファ・メモリ領域26へロードするため、かつ、データをバースト・バッファ・メモリ領域26からメイン・メモリ3へ格納するために使用されるものである。これらの命令は「ロードバースト」および「ストアバースト」である。ロードバースト命令が、データ・ワードのバーストを、メモリ3における決定された位置からバースト・バッファのそれへ転送させる。対応するストアバースト命令もあり、これは、データ・ワードのバーストを、バースト・バッファの1つからメモリ3へ、メモリ3における特定のアドレスで開始して、転送させる。図1のアーキテクチャでは、追加の同期化命令も必要とされ、これらは以下でさらに論じられる。

【 0 0 3 6 】

命令のロードバーストおよびストアバーストは、通常のロードおよび格納命令とは異なり、これらは転送が起こっていかなくとも単一のサイクルで完了する。本質において、ロードバーストおよびストアバースト命令が、メモリ・インタフェース 16 へバーストを実行するように伝えるが、これらはバーストが完了するまで待機しない。

【 0 0 3 7 】

基本動作は、2つのテーブル・エントリへ、メモリ・アクセスおよびバッファ・アクセス・テーブルのそれぞれにおける1つを索引付けする命令を発行することである。メモリ・アクセス・テーブルへの索引が、転送のメモリ側で使用された基底アドレス、エクステンツおよびストライドを検索する。バッファ・アクセス・テーブルへの索引が、バースト・バッファ・メモリ領域内の基底アドレスを検索する。図示の実施形態では、マスキングおよびオフセットが索引値へ、コンテキスト・テーブル（これは欧州特許出願第 9 7 3 0 9 5 1 4 . 4 号においてさらに論じられている）によって提供されるが、代わりに実アドレスを使用することが可能である。直接メモリ・アクセス（DMA）コントローラ 56 がパラメータを2つのテーブルから渡され、これらを使用して、要求された転送を指定する。

【 0 0 3 8 】

表 1 は、可能な命令セットを示す。

【 0 0 3 9 】

【表 1】

演算コード	パラメータ値	コメント
BB_LOADBURST	mat_index (整数)、 bat_index (整数)、 block_increment (ブール)	データのバーストをメイン・メモリからバースト・バッファ・メモリへロードし、任意選択でメイン・メモリにおける基底アドレスを増分する
BB_STOREBURST	mat_index (整数)、 bat_index (整数)、 block_increment (ブール)	データのバーストをバースト・バッファ・メモリからメイン・メモリへ格納し、任意選択でメイン・メモリにおける基底アドレスを増分する
BB_LX_INCREMENT	N/A	LX セマフォの値を増分する
BB_XS_DECREMENT	N/A	XS セマフォの値を減分する
BB_SET_MAT	entry (整数)、 memaddr (整数)、 extent (整数)、 stride (整数)	MAT エントリを所望の値に設定する
BB_SET_BAT	entry (整数)、 bufaddr (整数)、 extent (整数)	BAT エントリを所望の値に設定する

表 1 : バースト・バッファのための命令セット

【 0 0 4 0 】

ストアバースト命令（BB__ストアバースト）が、MAT および BAT におけるパラメータを索引付け (index) し、これが要求された転送の特性を定義する。

block_increment ビットが設定される場合、MAT において索引付けされたエントリの memaddr フィールドが、転送が完了するときに自動的に更新される（以下で論じられるように）。

【 0 0 4 1 】

ロードバースト命令（BB__ロードバースト）も、MAT および BAT におけるパラメ

ータを索引付けし、再度これが要求された転送の特性を定義する。前のように、block_incrementビットが設定される場合、MATにおいて索引付けされたエントリのmemaddrフィールドが、転送が完了するときに自動的に更新される。

【0042】

必要とされた同期化命令が、Load-Execute IncrementおよびExecute-Store Decrement(BB_LX_INCREMENTおよびBB_XS_DECREMENT)として提供される。BB_LX_INCREMENTの目的は、特定のデータのバーストにおけるコプロセッサ2の実行が、必要とされたデータがロードバースト命令の後に続いてバースト・バッファ・メモリ5に到着した後に起こるようにすることである。BB_XS_DECREMENTの目的は、ストアバースト命令の実行が、結果がメイン・メモリ3に格納され戻される(コプロセッサ2上の)計算の完了の後に続くようにすることである。

10

【0043】

この実施形態では、これらの命令が動作する特定の機構は、2つのカウンタの組であり、それぞれ次のものを追跡する。すなわち、a)ストアバーストを受信する用意ができているバースト・バッファ・メモリ5における領域の数、およびb)完了されたロードバースト命令の数、である。

【0044】

コプロセッサ2によるデータの要求が、LXカウンタを減分することによって実行されるが、データの可用性は、XSカウンタを増分することによって信号で通知される。これらのカウンタは、2つの特性を満たさなければならない。すなわち、これらがただ1つのシステム構成エレメントへいかなる所与のときにもアクセス可能でなければならないこと、およびこれらが使用不可能なデータを要求する処理を中断する能力を有していなければならないことである。

20

【0045】

要求されるものにもっとも近く合致する既存の概念はセマフォであり、Dijkstra(「[Dijkstra 1968] E. Dijkstra, 「Co-operating Sequential Processes」, F. Genuys (編集者), Programming Languages, New York: Academic Press, (1968), 43-112ページ)に記載されている。したがって、用語「セマフォ」は、本発明の実施形態において使用されたカウンタを記載するために使用されるが、これらのカウンタがDijkstraによって記載されたセマフォには等しくないが広く類似していることに留意されたい。

30

【0046】

セマフォの基本原理は以下の通りである。セマフォは整数値を含む。Wait()命令をセマフォ上で実行することはこの値をデクリメント(decrement、減分)するが、Signal()命令を実行することはこれをインクリメント(increment、増分)する。Wait()を、値がすでに0であるセマフォ上で実行すると、セマフォの値が増分されるまで、Wait()を実行しようと試みているソフトウェア処理またはハードウェア構成エレメントが停止する。

40

【0047】

この実施形態では、BB_XS_DECREMENT命令がXSセマフォ(図1の11)上でWait()のように動作するが、BB_LX_INCREMENT命令はLXセマフォ(図1の10)上でSignal()のように動作する。以下で記載されるように、コプロセッサ・コントローラ9が、反対に、Wait()をLXセマフォ10上で、Signal()をXSセマフォ11上で実行する。これらの命令の意味は、Dijkstraの論文に記載されたものと同じにすることができるが、Signal()およびWait()動作の構成全体は元の論文に記載されたものとは著しく異なる。システムの正確さに必要な、あるイベントの相対的時間順序が守られるようにするため、これらの命令が適切な順序で(以下でさらに論じられるように)発行される。

50

【 0 0 4 8 】

メモリ・アクセス・テーブル (MAT) 65 が、以下で図 3 を参照して記載される。これは、バースト・トランザクションに含まれたメイン・メモリ位置に関連する情報を保持するメモリ記述子テーブルである。MAT における各エントリは、メイン・メモリへのトランザクションを記述する、索引付けされたスロットである。この実施形態では、MAT 65 が 16 個のエントリを含むが、異なる実施がもちろん可能である。各エントリが以下の 3 個のフィールドを含む。

1. メモリ・アドレス (memaddr)。メイン・メモリにおける関連領域の開始アドレス。この位置が物理メモリ空間にあることが理想的であり、これは、仮想アドレス変換が 2 つの物理ページにまたがるバースト要求の結果となる可能性があり、これがメモリ・コントローラに難点を引き起こすからである。

10

2. エクステント (extent)。転送のエクステント。これは転送の長さであり、ストライドで乗算され、転送された最後のアドレス + 1 を与える。転送の長さは、エクステントをストライドにより除算することによって計算され、これは、転送が完了した後で、関連する BAT 66 (以下参照) の bufsize フィールドへ自動的にコピーされる。

3. ストライド (stride)。転送における連続したエレメントの間の間隔。

【 0 0 4 9 】

memaddr: チャンネル・バーストの第 1 のエレメントの 32 ビット、符号なし、語調整されたアドレスである。

20

【 0 0 5 0 】

extent: extent レジスタにおけるパラメータは、バースト転送の範囲を包含するアドレス・オフセットである。転送が S のストライドによって分離された L 個のエレメントを必要とする場合、エクステントは $S * L$ である。

【 0 0 5 1 】

stride: パラメータ stride は、アクセスの間でスキップされたバイト数である。転送ストライド間隔の値は、1 から 1024 までの範囲に制限される。1024 より大きい値は自動的に 1024 に切り捨てられる。このレジスタの読取りが、バーストのために使用された値を戻す (すなわち、切り捨てが必要であった場合、切り捨てられた値が戻される)。ストライドはメモリ・バス幅の倍数でなければならない、この場合は 4 バイトである。自動切り捨て (丸めなし) が、この位置合わせを実施するために実行される。

30

【 0 0 5 2 】

MAT スロットによって含まれた値の一例は、以下のような可能性がある。

【 0 0 5 3 】

{ 0x1feelbad, 128, 16 }

これは、32 語 (32 個の 4 バイト語) バーストの結果となり、各語が 4 語 (4 個の 4 バイト語) によって分離される。

【 0 0 5 4 】

バースト命令の自動増分標識ビットは、MAT 65 にも関連している。このビットがバースト命令において設定される場合、もしバーストが 32 を越えて継続しているならば、開始アドレス・エントリがポイント・トゥ・ポイントで、次のメモリ位置へ増分される。これは、長いシーケンスのメモリ・アクセスにおいて次のバーストのための開始アドレスを計算することにおける、プロセッサ・オーバーヘッドを節減する。

40

【 0 0 5 5 】

バッファ・アクセス・テーブル (BAT) 66 が、以下で図 3 を参照して記載される。これもまたメモリ記述子テーブルであり、この場合はバースト・バッファ・メモリ領域 26 に関する情報を保持する。BAT 66 における各エントリが、バースト・バッファ・メモリ領域 26 へのトランザクションを記述する。MAT 65 の場合のように、BAT 66 が 16 個のエントリを含むが、もちろん MAT 65 の場合のように変わる可能性がある。この場合は各エントリが以下の 2 つのフィールドを含む。

50

1. バッファ・アドレス (b u f a d d r)。バッファ領域におけるバッファの開始。
2. バッファ・サイズ (b u f s i z e)。最後の転送で使用されたバッファ領域のサイズ。

【 0 0 5 6 】

バッファ・アドレス・パラメータ b u f a d d r は、バッファ領域におけるチャネル・バーストの第 1 のエレメントのためのオフセット・アドレスである。バースト・バッファ領域は、ハードウェアによってプロセッサのメモリ空間の領域へ物理的にマッピングされる。これは、バースト・バッファ領域にアクセスするときにプロセッサが絶対アドレスを使用しなければならないことを意味する。しかし、D M A 転送は単にオフセットを使用し、そのため、ハードウェアが、要求されたいかなるアドレス解決をも管理することが必要である。不当に位置合わせされた値は、切り捨てによって自動的に位置合わせすることができる。このレジスタの読取りが、バーストのために使用された値を戻す（すなわち、切り捨てが必要であった場合、切り捨てられた値が戻される）。デフォルト値は 0 である。

10

【 0 0 5 7 】

パラメータ b u f s i z e は、もっとも最近のバーストによって占有されたバッファ領域内の領域のサイズである。このレジスタは、そのエントリを目標としたバースト転送の完了時に、自動的に設定される。格納された値がバーストの長さであり、これは、0 の値が使用されていないバッファ・エントリを指示するからであることに留意されたい。このレジスタに書き込むことができるが、これは、バッファが保存されリストアされるときに、文脈切り換え後にのみ有用である。デフォルト値は再度 0 である。

20

【 0 0 5 8 】

M A T および B A T エントリをプログラムすることが、B B _ S E T _ M A T および B B _ S E T _ B A T 命令の使用を通じて実行される。エントリ・パラメータが、現在の命令が参照する M A T（または B A T）におけるエントリを決定する。

【 0 0 5 9 】

バースト・バッファ・アーキテクチャおよびその制御のための機構のさらなる詳細が、欧州特許出願第 9 7 3 0 9 5 1 4 . 4 号および対応する米国特許出願第 0 9 / 3 , 5 2 6 号において提供される。上記で提供された詳細は、主として、バースト・バッファ・システムのアーキテクチャエレメントを示し、バースト・バッファ・システムが実施することができる機能効果を、それが提供する入力および出力と共に示すように意図される。バースト・バッファ・システムが、特定のタイプの計算モデルに合わせて最適に適合され、これは、本明細書で、本発明の記載された実施形態のための計算モデルに展開される。この計算モデルについてさらに説明する。

30

【 0 0 6 0 】

バースト命令待ち行列 6 が上で記載された。この実施形態の著しい態様は、命令が類似の方法でコプロセッサへ、コプロセッサ命令待ち行列 8 を通じて提供されることである。コプロセッサ命令待ち行列 8 は、コプロセッサ・コントローラ 9 との接続において動作し、これは、コプロセッサがプロセッサ 1 から命令を受信する方法、およびそれがデータをバースト・バッファ・システム 5 と交換する方法を決定する。

【 0 0 6 1 】

40

コプロセッサ命令待ち行列 8 の使用は、プロセッサ 1 自体が計算自体から切り離される重要な効果を有する。したがって、計算中に、プロセッサ・リソースが他のタスクの実行のために使用可能である。プロセッサ 1 の動作がストールされることにつながる可能性のある唯一の状況は、命令待ち行列 6、8 の一方が命令で満たされることである。この場合は、プロセッサ 1 が、いずれかの待ち行列のための命令を、命令が消費されるよりも速い速度で生成するときに起こる可能性がある。この問題の解決策は入手可能である。事前定義の時間量の後、あるいは、いずれかの待ち行列において占有されたスロットの数が事前定義の量まで減分された事実によってトリガされた割り込みの受信の上で、文脈切り換えを実行してこれら 2 つの待ち行列にサービスするために戻るように、プロセッサ 1 へ要求することによって、有効性を改善することができる。反対に、プロセッサ 1 が、命令が消

50

費される速度に遅れないでいることができないために、2つの待ち行列の一方が空になる場合、これらの命令の消費者（コプロセッサ・コントローラ9またはバースト・バッファ・コントローラ7）は、新しい命令がプロセッサ1によって生成されるまで、ストール（機能停止）する。

【0062】

修正を、プロセッサ1からのそれ以上の関与がまったく必要とされないようにするアーキテクチャにも提供することができ、これらについては本明細書の最後の部分において論じられる。

【0063】

コプロセッサ・コントローラ9の基本機能は、データをバースト・バッファ・メモリ5からコプロセッサ2へ（およびコプロセッサ2からバースト・バッファ・メモリ5へ）取り出すこと、コプロセッサの動作を制御すること、およびコプロセッサ2の実行をバースト・バッファ・メモリ5からの適切なロードまたはそれへの格納と同期化することである。これらの機能を達成するには、コプロセッサ・コントローラを、本質において、ある規則に従ってアドレスを生成することができる、相対的に簡素な状態機械にすることができる。

【0064】

図4は、コプロセッサ・コントローラ9を、アーキテクチャの他の構成エレメントとの関係において示し、その構成エレメント、およびアーキテクチャ全体における他のエレメントとの接続も示す。その厳密な機能は、コプロセッサ2およびその初期化要件（ある場合は）によって必要とされた入力および出力のタイプに依存し、そのため、以下に記載されたものから詳細において変わる可能性がある。CHESコプロセッサの場合、これらの入力および出力が、バースト・バッファ・メモリ5と交換された入力および出力データ・ストリームである。

【0065】

コプロセッサ・コントローラ9は、次の2つの主なタスクを実行する。すなわち、1)コプロセッサ2とバースト・バッファ・メモリ5との間の通信の制御、および2)制御有限状態機械42の使用を通じたシステム状態の維持である。

【0066】

コプロセッサ2がストリームにおけるデータにアクセスし、そのそれぞれがいくつかの制御レジスタ41の1つとの関連付けを与えられる。これらのレジスタ41のためのアドレスが、制御有限状態機械42によってアドレス指定ロジック43と共に、有限状態機械42によって生成されたシーケンスに従って、周期的な様式において生成される。

【0067】

有限状態機械42内のクロックの各チックで、有限状態機械は、レジスタ41の（多くとも）1つが、そのために生成された新しいアドレス、およびレジスタ41がバースト・バッファ・メモリ5をアドレス指定できるようにするために使用されたアドレスを有するための許可を与える。同時に、適切な制御信号が有限状態機械42によって生成され、マルチプレクサ44へ送信されて、適切なアドレスが正しい読み/書き信号と共にバースト・バッファ・メモリ5へ送信されるようにする。特定の読み/書き信号が各レジスタ41に、全体の計算中で変化しない値と共に関連付けられる。

【0068】

レジスタ41のために得られたアドレスがメモリをアドレス指定するために使用された後、一定量がその値へ追加され、これは一般にコプロセッサ2とバースト・バッファ・メモリ5との間の接続の幅と同じである。つまり、この接続の幅が4バイトである場合、カウンタ41に行われた増分が4となる。これは、本質的に、バースト・バッファのプログラミングにおける「ストライド」に比較可能である。

【0069】

上述のコプロセッサ・コントローラの機構は、単一のバスに沿った異なるデータ・ストリームの多重化を可能にする。各データ・ストリームは、それ自体のポートを通じて単一

10

20

30

40

50

の共有バスへアクセスするとみなすことができる。

【 0 0 7 0 】

このシステムが、通信の完全性が保証されるように動作するためには、バスの他方の端でコプロセッサ 2 が、同期した方法でこのバスから読み取り、このバスへ書き込みする用意ができていなければならない。アプリケーション・ソフトウェア（および具体的には、コプロセッサ 2 を構成するアプリケーション・ソフトウェアの一部まで）の責任は、つぎのことを保証することである。すなわち、1) 2 つのストリームが同時にバスにアクセスしようと試みないこと、および 2) コプロセッサ 2 の実行がバースト・バッファ・メモリ 5 とのデータ転送と同期であることである。

【 0 0 7 1 】

この後者の要件は、コプロセッサ 2 が、2 つのデバイスの間の接続上で、バースト・バッファ・メモリ 5 によって配置されたデータを読み取るため、かつその逆のための用意ができることを保証する。

【 0 0 7 2 】

通常は、複数の物理線を C h e s s アレイ 2 とバースト・バッファ・メモリ 5 の間に設けることができるが、一般の多重化の必要性はなお残る。コプロセッサ 2 とバースト・バッファ・メモリ 5 の間の物理接続の数が、コプロセッサ 2 のための論理 I / O ストリームの合計数以上でない限り、2 つ以上の論理ストリームが同じワイヤ上で多重化されなければならないことが常に真となる。（バースト・バッファ・メモリ 5 に有利に使用されるように）高速 S R A M の設計に関連した技術的な理由が、コプロセッサ 2 との複数の接続の使用を防止する。

【 0 0 7 3 】

コプロセッサ・コントローラ 9 は、コプロセッサ 2 を含む C H E S S アレイの実行を制御するようにも動作し、それが指定数のクロック・サイクルで実行するようにする。これは、コプロセッサ 2 におけるパイプラインの内部状態に影響を与えない方法において、C H E S S アレイを、その内部クロックを「ゲーティング」すること（つまり、停止すること）によって「フリーズ」する前に、指定数のサイクルだけチッキングする、制御有限状態機械 4 2 におけるカウンタによって達成される。このチックの数が、以下に記載された C C _ _ S T A R T _ _ E X E C 命令を使用して指定される。

【 0 0 7 4 】

コプロセッサ・コントローラ 9 が、プロセッサ 1 によって、コプロセッサ命令待ち行列 8 の使用を通じてプログラムされる。このコプロセッサ・コントローラ 9 のための可能な命令セットが、以下の表 2 において示される。

【 0 0 7 5 】

10

20

30

【表 2】

演算コード	パラメータ値	コメント
CC_CURRENT_PORT	n (整数)	次の CC_PORT_XXX コマンドが参照するポート番号
CC_PORT_PERIOD	(整数)	ポートの活動の周期
CC_PORT_PHASE_START	start (整数)	ポートの活動の段階開始
CC_PORT_PHASE_END	end (整数)	ポートの活動の段階終了
CC_PORT_TIME_START	t _{start} (整数)	ポートの活動の開始サイクル
CC_PORT_TIME_END	t _{end} (整数)	ポートの活動の終了サイクル
CC_PORT_ADDRESS	addr _{start} (整数)	ポートのための初期アドレス
CC_PORT_INCREMENT	addr _{end} (整数)	ポートのためのアドレス増分
CC_PORT_IS_WRITE	rw (ブール)	読み/書きフラグ
CC_START_EXEC	n _{cycles} (整数)	コプロセッサ 2 の実行を、決定されたサイクルの数だけ開始/再開する
CC_LX_DECREMENT	N/A	LX セマフォの値を減分する
CC_XS_INCREMENT	N/A	XS セマフォの値を増分する

表 2 : コプロセッサ・コントローラ命令セット

【 0 0 7 6 】

前記の命令では、命令フォーマットの異なる選択を行うことができる。1つの可能なフォーマットは32ビットの数字であり、16ビットが演算コードを符号化し、16ビットが上述の任意選択のパラメータ値を符号化する。

【 0 0 7 7 】

個々の命令の意味は、以下の通りである。

- ・ CC_CURRENT_PORT は、ポートの1つを、すべての後続の CC_PORT_XXX 命令の受信者として、次の CC_CURRENT_PORT まで、選択する。
- ・ CC_PORT_PERIOD () は、現在のポートの活動化の周期を、整数パラメータの値に設定する。
- ・ CC_PORT_PHASE_START / CC_PORT_PHASE_END (s_{start} s_{end}) は、現在のポートの活動化の段階の開始/終了を、整数のパラメータ (s_{start} s_{end}) の値に設定する。
- ・ CC_PORT_TIME_START / CC_PORT_TIME_END (t_{start} t_{end}) は、現在のポートの活動の最初/最後のサイクルを設定する。
- ・ CC_PORT_ADDRESS (addr_{start}) は、現在のポートの現在のアドレスを、整数のパラメータ addr_{start} の値に設定する。
- ・ CC_PORT_INCREMENT (addr_{incr}) は、現在のポートのアドレス増分を、整数のパラメータ addr_{incr} の値に設定する。
- ・ CC_PORT_IS_WRITE (rw) は、現在のポートのためのデータ転送方向を、ブール・パラメータ rw の値に設定する。
- ・ CC_START_EXEC n_{cycles} は、コプロセッサ・コントローラ 2 の実行を、関連付けられた整数パラメータ n_{cycles} によって指定されたクロック・サイクルの数だけ開始する。
- ・ CC_LXS_DECREMENT は、LX セマフォの値を (以前に記載されたように、中断の方法において) 減分する。
- ・ CC_XSS_INCREMENT は、XS セマフォの値を増分する。

【 0 0 7 8 】

10

20

30

40

50

カウンタ42の現在値、 t_{cur} が、 $t_{start} \leq t_{cur} < t_{end}$ であり、 $t_{start} \leq (t_{cur} \bmod t_{end}) < t_{end}$ である場合、ポートが能動（つまり、バースト・バッファ・メモリ5との通信の制御を有する）として定義される。これは、たとえば、2つのストリームが、等しい周期、言わば5で存在し、一方が最初の4サイクルのためのBBメモリの制御を有し、他方が残りのサイクルのための制御を有する、システムの可能性を可能にする。

【0079】

このアーキテクチャを使用するアルゴリズムを実行する処理は、最初にコプロセッサ2のプログラミング、次いでコプロセッサ・コントローラ9およびバースト・バッファ・コントローラ7のプログラミングまたは初期化と、その後続くアルゴリズムの実際の実行を含む。

10

【0080】

コプロセッサ2の初期化では、デバイスの実際の実施形態に特定的手段によって、構成がコプロセッサ自体にロードされることが、一般にもっとも直接になる。

【0081】

コプロセッサ・コントローラ9のプログラミングでは、そのステップは次の通りである。

1. メイン・コプロセッサ・コントローラ9が、前に記載されたように、Chessアレイにおいて存在する各論理ストリーム毎の合計数、周期、段階およびアドレス増分に従って構成される。所望の機能を実行するためのコプロセッサ・コントローラ9のプログラミングの一例が、下に示される。

20

2. コプロセッサ・コントローラ9の構成における次のステップは、アドレス構成である。各論理ストリームの特性（周期、段階）がアルゴリズム中で同じであり続ける可能性が高いが、バースト・バッファ・メモリ5におけるコプロセッサ・コントローラ9によってアクセスされた実アドレスは変わる。それはこの可変性であり、これは、バースト・バッファ・コントローラ7がダブル・バッファリングを、バースト・バッファ・アーキテクチャ内で直接の方法で実行できるようにする。このダブル・バッファリングの効果は、先に述べられたように、コプロセッサ2に、それが連続ストリームと対話中である印象を与えることであるが、実際にはバッファが連続的に交換されている。

【0082】

30

バースト・バッファ・コントローラ7も構成される必要がある。これを行うには、メイン・メモリ3からバースト・バッファ・メモリ5へのデータの転送を構成するために、適切なコマンドがバースト命令待ち行列6へ送信されなければならない。これらの命令（ BB_SET_MAT および BB_SET_BAT ）が、 BAT および MAT 内で適切なエントリを、コプロセッサ・コントローラ9のプログラミングに適合する方法で構成する。この実施形態では、 MAT および BAT エントリをプログラムするための命令が、バースト命令待ち行列6を通じて発行される。代替の可能性は、プロセッサ1が読み書きするメモリ・マップ・レジスタの使用となる。この実施形態の場合のように、メモリ・マップ・レジスタから読み取られる可能性がなく（それらが存在しないので）、プロセッサ1がバースト・バッファ・コントローラ7の状態を照会することができないが、これは著しい制限ではない。さらに、この目的のためのバースト命令待ち行列6の使用は、バースト転送の実行により MAT および BAT エントリを構成するための命令をインターリーピングし、したがって、プロセッサ1の監視なしに正しい時間的意味を維持する可能性を可能とする。

40

【0083】

これらのステップが実行された後、Chessアレイの実際の実行を開始することができる。この実施形態では、Chessアレイに指定数のサイクルで実行するように命令することのみが必要である。これは、正確な数のサイクルをパラメータとして、コプロセッサ命令待ち行列8における CC_START_EXEC 命令へ書き込むことによって達成され、次いでこのデータをコプロセッサ・コントローラ9へ渡すことができるようにする

50

。この値がコプロセッサ・コントローラ 9 に転送された 1 クロック・サイクル後、コントローラが値をバースト・バッファ・メモリ 5 とコプロセッサ 2 の C H E S S アレイの間で転送することを開始し、C H E S S アレイの実行を可能にする。

【 0 0 8 4 】

しかし、重要なステップを、計算に関係する命令がそれぞれの命令待ち行列に配置される前に追加しなければならない。これは、必要な同期化機構が、同期化およびダブル・バッファリングの原理をうまく実施するために適切であるようにするためである。この機構における基本エレメントは、コプロセッサ・コントローラ 9 が L X セマフォの値を減分しようと試み、上述のロジックに従って、それがそうできるようになるまでコプロセッサの動作を中断することである。このセマフォの初期値は 0 である。すなわち、したがって、コプロセッサ・コントローラ 9 およびコプロセッサ 2 がこの段階で「フリーズ」される。成功したロードバースト命令の後に L X セマフォの値がバースト・バッファ・コントローラ 7 によって増分されるときにのみ、コプロセッサ 2 がその実行を開始（あるいは再開）できるようになる。この効果を達成するため、C C _ L X _ D E C R E M E N T 命令がコプロセッサ命令待ち行列 8 において、「コプロセッサ 2 の実行を開始する」（C C _ S T A R T _ E X E C）命令の前に挿入される。以下に示されるように、対応する「L X セマフォを増分する」（B B _ L X _ I N C R E M E N T）命令が、バースト命令待ち行列 6 において、対応するロードバースト命令の後に挿入される。

【 0 0 8 5 】

C H E S S 論理ストリームおよびバースト・バッファ・メモリ 5 の間の実際のデータの転送が、先に記載されたようなコプロセッサ・コントローラ 9 のプログラミングに従って実行される。

【 0 0 8 6 】

カウンタ 4 2 が実行しなければならないチックの数は、1 つまたは複数の入力バーストを消費するためにどれだけの時間がかかるかに依存する。システムの正確さを保証することは、アプリケーション・ソフトウェアに任される。カウンタ 4 2 のプログラミングは、バッファが消費された後にコプロセッサ 2 の実行が停止するようにしなければならない。コプロセッサ命令待ち行列 8 における次の命令は、次のデータのバーストがバースト・バッファ・メモリ 5 に到着しているようにするため、同期化命令（つまり、C C _ L X _ D E C R E M E N T）でなければならない。この命令（および場合によっては、必要とされたデータが使用可能になるまでの待機期間）の後に続いて、この新しいデータのバーストの初期アドレスがデータ・ストリームに割り当てられ（C C _ P O R T _ A D D R E S S 命令により）、実行が C C _ S T A R T _ E X E C 命令を介して再開される。この手順は出力ストリームの場合に類似している（重要な違いは、データがメイン・メモリ 3 からバースト・バッファ・メモリ 5 に到着するために必要としたものに等しい待機期間がなくなることである）。

【 0 0 8 7 】

計算モデル

計算モデル全体の例を図 5 を参照して説明する。この例は、アルゴリズムがこのアーキテクチャにおいて使用するためにどのように記録することができるかを指示し、一例として簡素なベクトル加算を使用し、これは従来のマイクロプロセッサ向けに以下のように C でコーディングすることができる。

【 0 0 8 8 】

【表 3】

```
int    a [ 1 0 2 4 ] , b [ 1 0 2 4 ] , c [ 1 0 2 4 ] ;
for ( i = 0 ; i < 1 0 2 4 ; i ++ )
    a [ i ] = b [ i ] + c [ i ] ;
```

【 0 0 8 9 】

図 1 のアーキテクチャ上で元のベクトル加算ループ・ネストと同じ機能性を達成するようプロセッサ 1 を走らせる C コードの一部は、次の通りである。

【 0 0 9 0 】

【表 4】

```

0:  int a[1024], b[1024] c[1024];
1:  int eo, not_eo, k;
2:  /*Port 0 specification: port #, increment, xfer size, period,
3:  phase start, phase end, start time, end time, r/w*/
4:  CIQ_STREAM( 0, 4, 4, 3, 0, 1, 0, 3*BLEN*MAXK+3, 0 );
5:  /*Port 1 specification*/
6:  CIQ_STREAM( 1, 4, 4, 3, 1, 2, 0, 3*BLEN*MAXK+3, 0 );
7:  /*Port 2 specification*/
8:  CIQ_STREAM( 2, 4, 4, 3, 2, 3, 0, 3*BLEN*MAXK+3, 1 );
9:  BIQ_SET_MAT(0, &b[0], BLEN*4, 4);
10: BIQ_SET_MAT(1, &c[0], BLEN*4, 4);
11: BIQ_SET_MAT(2, &a[0], BLEN*4, 4);
12: BIQ_SET_BAT(0, 0x0000, BLEN*4); BIQ_SET_BAT(1, 0x0100, BLEN*4);
13: BIQ_SET_BAT(2, 0x0200, BLEN*4); BIQ_SET_BAT(3, 0x0300, BLEN*4);
14: BIQ_SET_BAT(4, 0x0400, BLEN*4); BIQ_SET_BAT(5, 0x0500, BLEN*4);
15: for( k = 0; k < MAXK; k++ )
16: {
17:   /*Even or odd iteration? - For double buffering*/
18:   eo = k&0x1;
19:   CIQ_LXD(2);
20:   CIQ_SA(0, (BLEN*4*eo));
21:   CIQ_SA(1, ((2*BLEN*4)+BLEN*4*eo));
22:   CIQ_SA(2, ((4*BLEN*4)+BLEN*4*eo));
23:   /*Start Chess*/
24:   CIQ_ST(3*BLEN);
25:   CIQ_XSI(1);
26:   /*BB stuff*/
27:   /*Load A*/
28:   BIQ_FLB(0, eo);
29:   /*Load B*/
30:   BIQ_FLB(2, 2+eo);
31:   BIQ_LXI(2);
32:   if( k >= 1 )
33:   {
34:     not_eo = (eo==0)?1:0;
35:     BIQ_XSD(1);
36:     BIQ_FSB(4, 4+not_eo);

```

10

20

30

```

37: }
38: }
39: eo = MAXK & 0x1;
40: not_eo = (eo==0)?1:0;
41: BIQ_XSD(1);
42: BIQ_FSB(4, 4+not_eo);

```

【 0 0 9 1 】

この構成では、3つのポートがコプロセッサ・コントローラ9において使用され、すなわち、各入力ベクトルのためのもの（bおよびc）、および出力ベクトルのためのもの（a）である。行4、6および8の文は、これらの3つのポートを初期化するためのコード・マクロである。これらは、拡張されるとき、次の表のコマンドになる（これは行4を参照し、他の拡張されたマクロは直接類似している）。

40

【 0 0 9 2 】

【表 5】

```

CC_CURRENT_PORT(0);
CC_PORT_INCREMENT(4);
CC_TRANSFER_SIZE(4);
CC_PORT_PERIOD(3);

```

50

```

CC__PORT__PHASE__START(0);
CC__PORT__PHASE__END(1);
CC__PORT__START__TIME(0);
CC__PORT__END__TIME(3*BL EN*MAXK+3);
CC__PORT__IS__WRITE(0);

```

【0093】

このコードは、ポート0が、カウンタ42の3チック毎、正確にはチック0、3、6、
 . . 3*BL EN*MAXK+3の4バイトのデータを読み取り、それが読み取るアドレ
 スを毎回4バイト増分するという効果を有する。BL EN*MAXKは、合計する2つの
 ベクトルの長さ（この場合、1024）であり、BL ENはDRAMからの単一のデータ
 のバーストの長さ（たとえば、64バイト）である。これらの値で、MAXKは1024
 /64=16に設定される。

10

【0094】

行9から14は、バースト・バッファ転送のためのMATおよびBATを確立し、これ
 らのテーブルにおけるエントリをメイン・メモリ3およびバースト・バッファ・メモリ5
 におけるアドレスに結合する。コマンドBIQ__SET__MAT(0,&b[0],BL
 EN*4,4,TRUE)はコード・マクロであり、これはBB__SET__MAT(0,
 &b[0],BL EN*4,4)に拡張され、MATにおけるエントリ0をアドレス&b
 [0]へ結合し、バースト長をBL EN*4バイトに（つまり、整数が32ビットの場合
 、BL EN整数）、ストライドを4に設定する。後に続く2行は、cおよびaに類似し、
 関係する。行BIQ__SET__BAT(0,0x0000,BL EN*4)は、BB__S
 ET__BAT(0,0x0000,BL EN*4)に拡張され、BATのエントリ0をバ
 ースト・バッファ・メモリ5におけるアドレス0x0000へ結合する。後に続く2行は
 再度類似している。

20

【0095】

この点まで、計算は行われていないが、コプロセッサ・コントローラ9およびバースト
 ・バッファ・コントローラ7が設定されている。行15から38のループ・ネストは、実
 際の計算が行われるところである。このループはMAXK回繰り返され、各反復がBL E
 Nエレメント上で動作し、処理されたMAXK*BL ENエレメントの合計を与える。こ
 のループは、コプロセッサ命令待ち行列8へ送信された命令の組CIQ__x x xで開始し
 て、コプロセッサ2およびコプロセッサ・コントローラ9の動作を制御し、その後、バ
 ースト・バッファ・コントローラ7およびバースト・バッファ・メモリ5を制御すること
 が目的である、バースト命令待ち行列6へ送信された1組の命令が続く。これら2組の相
 対的な順序は、原理においては重要でなく、これは、異なるシステムエレメントの間の同
 期化が、セマフォによって明示的に保証されるからである。互いの後に実行する2つの異
 なるループを有すること（2つの命令待ち行列が十分深かったと仮定して）、または、2
 つの異なる制御のスレッドを有することさえも可能となる。

30

【0096】

CIQ__x x x行は、ソース・コードを書くことを簡約にするコード・マクロである。
 これらの意味は、次の通りである。

40

CIQ__LXD(N):N個のCC__LXS__DECREMENT命令をコプロセッサ
 命令待ち行列8に挿入する。

CIQ__SA(ポート、アドレス):CC__CURRENT__PORT(ポート)およ
 びCC__PORT__ADDRESS(アドレス)命令をコプロセッサ命令待ち行列8に挿
 入する。

CIQ__ST(cycleno):コプロセッサ2にカウンタ42のcyclenoチ
 ックだけ実行させるために、CC__EXECUTE__START(cycleno)命令
 を挿入する。

CIQ__XSI(N):N個のCC__XSS__INCREMENT命令をコプロセッサ
 命令待ち行列8に挿入する。

50

【 0 0 9 7 】

上に示したコードのネットの効果は、以下の通りである。すなわち、1) L X S セマフォ上で対応するロードバーストと同期化すること、2) コプロセッサ 2 上で計算を、カウンタ 4 2 の 3 * B L E N チックだけ開始すること、および 3) X S S セマフォ上で対応するストアバーストと同期化することである。

【 0 0 9 8 】

B I Q _ _ x x x 行は再度、ソース・コードを書くことを簡約にするコード・マクロである。これらの意味は、次の通りである。

B I Q _ _ F L B (m a t e , b a t e) : B B _ _ ロードバースト (m a t e , b a t e , T R U E) 命令をバースト命令待ち行列 6 に挿入する。

B I Q _ _ L X I (N) : N 個の B B _ _ L X _ _ I N C R E M E N T 命令をバースト命令待ち行列 6 に挿入する。

B I Q _ _ F S B (m a t e , b a t e) : B B _ _ ストアバースト (m a t e , b a t e , T R U E) 命令をバースト命令待ち行列 6 に挿入する。

B I Q _ _ X S D (N) : N 個の B B _ _ X S _ _ D E C R E M E N T 命令をバースト命令待ち行列 6 に挿入する。

【 0 0 9 9 】

上に示したコードのネットの効果は、メイン D R A M メモリ 3 からバースト・バッファ・メモリ 5 へ 2 つのバーストをロードし、次いで L X セマフォ 1 0 の値を増分して、コプロセッサ 2 がその実行を上述のように開始できるようにすることである。最初のを除くすべての反復において、コプロセッサ 2 の計算の結果が、次いで、ストアバースト命令を使用してメイン・メモリ 3 に戻すように格納される。2 番目の反復が、最初の反復において実行された計算の結果を格納することを待機する必要は厳密にはないが、これはコプロセッサ 2 とバースト・バッファ・メモリ 5 の間の並行性を向上させる。

【 0 1 0 0 】

2 つの変数 e o および n o t _ _ e o の使用は、先に記載されたダブル・バッファリング効果を可能にするためにここで使用された機構である。

【 0 1 0 1 】

行 3 9 から 4 2 は、バースト・バッファ・メモリ 5 からメイン・メモリ 3 への最後のバースト転送を実行し、ループ本体の最初の反復におけるストアバースト命令の不在を補償する。

【 0 1 0 2 】

結果として生じる時間線は図 6 の通りである。ロードバースト 6 0 1 は最初の動作であり（これらが完了されるまで、コプロセッサ 2 がロード / 実行セマフォによってストールされるので）、これらが完了されるとき、コプロセッサ 2 が 6 0 2 を実行開始することができる。バースト命令待ち行列 6 における次の命令は別のロードバースト 6 0 1 であり、これが、最初の 2 つのロードが終了するとすぐに実行される。次いで、バースト命令待ち行列 6 における次の命令がストアバースト 6 0 3 であり、これは、X S セマフォ 1 1 が、コプロセッサ 2 上の最初の計算が完了したことを信号で知らせるまで待機しなければならない。この処理は、ループ中で継続する。

【 0 1 0 3 】

上に示された例は非常に簡素なアルゴリズムの場合であるが、これは、より複雑である計算において必要とされる基本原理を例示する。当業者は、上に示された手法、原理および技術を、より複雑なアルゴリズムをこのアーキテクチャによる実行に適合させるために図 1 のアーキテクチャをプログラムすることに使用できる。

【 0 1 0 4 】

計算のためのツールチェーン

計算モデルの原理を、ハンド・コーディングによって直接の様式で活用することができる。つまり、手動で C コードを書いて、従来の方法でシステム構成要素の適切な動作をスケジュールするように適合された C P U 上で実行して（命令を適切な待ち行列に配

10

20

30

40

50

置し、記載されたようにシステム構成エレメントを動作に設定し)、コプロセッサのための適切な構成を、そのコプロセッサを構成するための標準の合成ツールに従って提供することである。CHES Sのような、構成可能またはFPGAに基づいたプロセッサでは、このツールが一般にハードウェア記述言語となる。CHES Sに使用するための適切なハードウェア記述言語はJHDLであり、たとえば、Peter BellowsおよびBrad Hutchingsによる1998年4月の「JHDL - An HDL for Reconfigurable Systems」Proceedings of the IEEE Symposium on Field-Programmable Custom Computing Machinesに記載されている。

【0105】

10

好ましい代替手法は、特定のツールチェーンがこの計算アーキテクチャのために使用されることである。このようなツールチェーンおよびその実際の動作のエレメントを次に簡単に説明する。

【0106】

ツールチェーンは、従来の順次コードから、有効な動作に特に適合されたコードへ変換すること、およびシステム構成エレメントの相互運用の機能を有する。例示的ツールチェーンは入力としてCコードを受信し、出力として次のものを提供する。すなわち、1)計算の実行のためのCHES Sコプロセッサ構成、2)データをシステム・メモリとバースト・バッファ・メモリの間で移動するためのバースト・バッファ・スケジュール、および3)データをCHES Sコプロセッサとバースト・バッファ・メモリの間で移動するためのコプロセッサ・コントローラ構成である。

20

【0107】

ツールチェーン自体は2つの構成エレメントを有する。第1のものはフロントエンドであり、これはCコードをその入力として取り、注釈付きのディペンデンスグラフをその出力として提供する。第2の構成エレメントはバックエンドであり、これはフロントエンドによって生成されたディペンデンスグラフを取り、これらからCHES S構成、バースト・バッファ・スケジュール、およびコプロセッサ・コントローラ構成を生成する。

【0108】

フロントエンドの主なタスクは、計算を、それがコプロセッサ2において起こるように、適切に記述するグラフを生成することである。実行された主なステップの1つは、値に基づいた依存解析であり、これはW. PughおよびD. Wonnacottによる1993年12月の「An Exact Method for Analysis of Value-based Array Data Dependences」, University of Maryland, Institute for Advanced Computer Studies - Dept. of Computer Science, University of Marylandに記載されている。生成された出力は、CHES Sアレイにおいて実施されるデータフローの記述、および(ロードバースト命令を介して)入力としてロードされるか、あるいは(ストアバースト命令を介して)出力として格納される必要のあるすべてのアドレスの表現、およびデータがメイン・メモリ3から検索され、それへ格納されなければならない順序の表現である。これは、バースト・バッファ・コントローラ7のための効率的なスケジュールが導出される基礎である。

30

40

【0109】

一例として、4タップFIRフィルタのためのCコードを仮定する場合、次の表のようになる。

【0110】

【表6】

```
int i, j, src[], kernel[], dst[];
for (i=0; i<1000; i++)
    for (j=0; j<4; j++)
```

50

```
dst[i] = dst[i] + src[4+i-j] * kernel[j];
```

【 0 1 1 1 】

テキスト・ファイルとして提供された、フロントエンド、出力への入力として、次の形式を有する。

【 0 1 1 2 】

【 表 7 】

```
loop:0<=i<999 #loop nest description
```

```
loop:0<=j<4
```

```
16:str/0/0/20/ #store instruction
```

```
LOD:
```

```
#Array:d[1/0/0] at line 11
```

```
20:ldc/16/0/0/ #load constant
```

```
22:str/0/0/26/ #store instruction, which
```

```
LOD: 4 <= j #writes its outputs to main
```

```
#Array:d[1/0/0] at line 13 #memory if 4<=j
```

```
26:add/22/27/31/ #addition
```

```
27:lod/26/0/0/ #load instruction, taking its inputs
```

```
Dep(16): [0][0] / Range: j <= 0 #from instruction 16 if j<=0
```

```
Dep(22): [0][1] / Range: 1 <= j #from instruction 22 otherwise
```

```
LID:
```

```
#Array:d[1/0/0] at line 13
```

```
31:mul/26/32/37/ #multiplication
```

```
32:lod/31/0/0/ #load instruction
```

```
Dep(32): [1][1] / Range: 1 <= i && 1 <= j
```

```
LID: i <= 0 || j <= 0 && 1 <= i #which takes its inputs from main
```

```
#Array:src[1/-1/0] at line 13 #memory if i <= 0 || j <= 0 && 1 <= i
```

```
37:lod/31/0/0/
```

```
Dep(37): [1][0] / Range: 1 <= i #load instruction
```

```
LID: i <= 0 #taking its inputs from main memory if
```

```
#Array:kernel[0/1/0] at line 13 #i<=0
```

【 0 1 1 3 】

このテキスト・ファイルは注釈付きのグラフの表現である。グラフ自体が、図 7 に示される。このグラフは、フロントエンド・アルゴリズムによって判明された依存性を明瞭に示す。エッジ 8 1 が、依存性が存在する条件、および適用可能である場合の依存性の距離によりマーク付けされる。記述は、必要とされる機能性を有するハードウェア構成エレメントを生成するに必要なすべての情報を含む。

【 0 1 1 4 】

コンパイル・ツールチェーンのバックエンドは、ある基本機能を有する。1 つは、フロントエンドから得られた、拡張されたディペンデンスグラフをスケジューリングし、時間変更することである。これは、十分に機能的な C H E S S 構成を得るために必要である。スケジューリングは、拡張されたディペンデンスグラフにおけるノード 8 2 のそれぞれが活動化される時点を決定的なことを含み、時間変更は、たとえば、エッジが値を適切な瞬間に伝播するようにするための遅延の挿入を含む。スケジューリングは、シフト線形スケジューリングを使用して実行することができ、これはハードウェア合成において幅広く使用される技術である。時間変更は、ハードウェア合成における共通で静かな直接のタスクであ

10

20

30

40

50

り、適切な数のレジスタを回路に追加して、回路における異なる経路が適切な時点で交わるようにすることのみを含む。この点で、コプロセッサ2（ここでは、CHESスコプロセッサ）の機能性の完全な記述を有する。この記述が図8に示される。次いで、この記述を適切なツールへ渡して、この機能性を有するCHESスコプロセッサをプログラムするために必要な一連の信号（一般に「ビットストリーム」と呼ばれる）を生成することができる。

【0115】

バックエンドの必要とされる別の機能は、バースト・バッファおよびコプロセッサ・コントローラ・スケジュールの生成である。CHES構成が得られた後、これにメイン・メモリからの値を入れる必要があるとき、および値をメイン・メモリへ戻すように格納でき、バースト・バッファ・スケジュールを確立できることが明らかである。したがって、バースト・バッファ・メモリ5にロードされ、そこから格納される必要のあるすべてのデータのアドレス空間を、バースト・バッファ・コントローラ7が動作できる固定のデータのバーストに分割することを含む、ステップが提供される。

【0116】

たとえば、上に提示されたFIRの例では、入力アレイ（ $src[]$ ）が適切なサイズのいくつかのバーストに分割され、アルゴリズムのために必要とされたすべてのアドレス範囲が包含されるようになる。このツールチェーンは長さ B_{len} のバーストを使用して（ B_{len} は2の累乗であり、このツールチェーンへの実行パラメータとして指定される）、できるだけ多くの入力アドレス空間を包含する。このバースト長でそれ以上達成できないとき、ツールチェーンは、長さを減分するバースト、すなわち、 $B_{len}/2$ 、 $B_{len}/4$ 、 $B_{len}/8$ 、 \dots 、2、1を、このアルゴリズムのために必要とされたあらゆる入力アドレスが唯一のバーストに属するまで使用する。

【0117】

これらのバーストのそれぞれについて、ロードされたデータのいずれかが必要とされる、反復空間におけるもっとも早い点が計算される。すなわち、各入力バーストに対して、反復空間において関連付けられた1点があり、そこでは、それより早い反復が、バーストによってロードされたデータのいずれをも必要としないことが保証される。コプロセッサ2の実行が反復空間におけるこの点に到達するときを検出することは容易である。したがって、次のものが作成される。すなわち、1)データをバースト・バッファ・メモリ5に移動するための、関連アドレスのためのロードバースト命令、および2)コプロセッサ2の実行が関連ロードバースト命令と同期化されることを保証するための、対応する同期化点（ $CC_LX_DECREMENT/BB_LX_INCREMENT$ の組）である。

【0118】

計算および通信の効果的なオーバーラップを達成するには、バスを介したデータの転送に関連付けられた待ち時間を隠すために、ロードバースト命令が予め発行されなければならない。

【0119】

アルゴリズムによって包含されなければならないすべての出力アドレス空間が、類似のロジックに従って出力バーストに区分される。再度、出力空間が、可変長のバーストに区分される。

【0120】

ツールチェーンは、1)関連アドレスのためのストアバースト命令、および2)対応する同期化点（ $BB_XS_DECREMENT/CC_XS_INCREMENT$ の組）を作成する。

【0121】

この点で、我々は、次のことに関連する情報を有する。すなわち、1)ロードバーストおよびストアバースト命令の相対的順序、およびそれらの実行のパラメータ（アドレスなど）、および2)コプロセッサ2上で実行される計算に相対的なそれらの位置である。

【0122】

10

20

30

40

50

次いで、この情報が使用されて、上述のFIRの例のように、全体の計算を編成するための適切なCコードが生成される。

【0123】

実際のコード生成段階（つまり、プロセッサ1上で実行するためのCコードの排出）を、`http://www.cs.umd.edu/projects/omega/`で入手可能である、Omega Library of the University of Marylandに含まれたコード生成ルーチンと、その後が続いて、これらのルーチンの総称出力を上述の形式に変換する、カスタマイズされたスクリプトを使用して実施することができる。

【0124】

実験結果 - 画像たたみこみ

画像たたみこみアルゴリズムが、次の表のループ・ネストによって記述される。

【0125】

【表8】

```
for (i=0; i < IMAGE_HEIGHT; i++)
for (j=0; j < IMAGE_WIDTH; j++)
for (k=0; k < KERNEL_HEIGHT; k++)
for (l=0; l < KERNEL_WIDTH; l++)
Dest[i, j] += Source[(i+1)-k, (j+1)-l] * C[k, l];
```

【0126】

境界条件を簡約するために、ソース画像を、垂直方向においてKERNEL_HEIGHT - 1画素、水平方向においてKERNEL_WIDTH - 1画素だけ拡張するために、複製が使用された。2つのカーネル、すなわち共にメディアン・フィルタ（median filtering）を実行する3×3カーネルおよび5×5カーネルが、システム性能を評価することに使用される。

【0127】

図9および図10は、（BBCとして示す）本発明の一実施形態によるアーキテクチャの性能を、バースト・バッファ（BBとして示す）を使用した従来のプロセッサ、および従来のプロセッサおよびキャッシュの組み合わせ（キャッシュとして示す）に対するものとして例示する。2つのバージョンのアルゴリズムが実施され、1つは32ビットの画素により、1つは8ビットの画素による。同じ実験測定値が、異なる画像サイズで、8×8から128×128までの範囲で、異なるバースト長でとられた。

【0128】

図から分かるように、BBC実装が、BBおよびキャッシュ実装に勝る高い性能上の利点を示した。このアルゴリズムは相対的に複雑であり、BBおよびキャッシュ実装におけるシステムの性能全体が非常にコンピュータ制約的（computer bound）であり、アルゴリズムが高度に複雑であるためにCPUがついていけない。本発明の実施形態を使用すると、計算がCHESSEアレイ上で（その固有の並行性により）実行されるので計算が大いに効果的であり、性能はあえていうとIO制約的（IO bound）である。最も、IOもバースト・バッファの有効な使用を通じて効率的である。マルチメディア命令（MIPS MDMXなど）がBBまたはキャッシュ実施におけるCPUの性能を向上させることができ、これは、それらがいくつかの算術命令の並列実行を可能にできるからである。それでもなお、性能向上の結果は、この構成において専用コプロセッサを使用して得られた性能レベルに到達する可能性が低い。

【0129】

修正および変形形態

プロセッサ1をコプロセッサ2およびバースト・バッファ・メモリ5から切り離す機能を、命令待ち行列6、8以外によって達成することができる。有効な代替物は、2つの待ち行列を、図12に記載されたような、命令をバースト・バッファ・メモリ5およびコプロセッサ2へ発行することに完全に専用にされた2つの小型プロセッサ（それぞれが各待

10

20

30

40

50

ち行列用)と置換することである。バースト命令待ち行列が(図1の実施形態を参照して)バースト・コマンド・プロセッサ106によって置換され、コプロセッサ命令待ち行列が、コプロセッサ・コマンド・プロセッサ108によって置換される。これは、これらの2つの構成エレメントによって実行された唯一のタスクとなるので、これらがコプロセッサ2およびバースト・バッファ7からそれぞれ切り離される必要はなくなる。コマンド・プロセッサ106、108のそれぞれが、コマンドをコプロセッサまたはバースト・バッファ(適切のように)へ発行することによって動作でき、次いで、そのコマンドがその実行を完了するまで何も行わず、別のコマンドを発行することなどができる。これは、設計を複雑にするが、メイン・プロセッサ1をその残りの、命令を待ち行列へ発行する単純なタスクから解放する。プロセッサ1によって実行される唯一の作業は、次いで、これらの2つのプロセッサの初期設定となり、これは計算の開始直前に行われる。したがって、計算中に、プロセッサ1が完全にコプロセッサ2およびバースト・バッファ・メモリ5の実行から切り離される。

10

【0130】

2つの従来の、しかしより小型のマイクロプロセッサ(あるいは、別法として、2つの制御の独立スレッドを実行する唯一のプロセッサ)を使用することができ、それぞれが適切なコード(ループ・ネスト)の関連部分を実行する。別法として、外部挙動がコードの関連部分の実行を反映する、2つの汎用状態機械を合成することができる(つまり、これらが同じ命令のシーケンスを提供する)。このような状態機械のハードウェアの複雑さおよびコストは、同等の専用プロセッサのものよりも著しく低くなる。このような状態機械が、メイン・プロセッサ1によって、上述のものに類似した方法でプログラムされる。主な違いは、イベントの繰り返しも符号化されることである。すなわち、これは、プロセッサ1が、多少の(複雑である場合は)命令における1つのアルゴリズムの挙動を符号化できるように必要である。イベントのx回の繰り返しを得るために、プロセッサ1がx個の命令を待ち行列に挿入する必要はないが、この繰り返しパラメータを命令定義において符号化しなければならない。

20

【0131】

上に示したように、特に有効な機構は、有限状態機械(FSM:Finite StateMachine)を待ち行列の代りに使用して、メイン・プロセッサ1の実行をコプロセッサ2およびバースト・バッファ・コントローラ7の実行から切り離すことである。この機構が以下でより詳細に論じられる。

30

【0132】

図1に例示されたアーキテクチャでは、異なるI/Oストリームの実行を駆動するための命令を、コプロセッサ2の実行のための命令と混合することができる。これは、システム構成エレメントの間の相互関係がコンパイル時に知られており、したがって、異なるシステム構成エレメントへの命令を正しい順序でソース・コードにおいてインタリーブすることができるので、可能である。

【0133】

2つの状態機械を、まったく同じ方法による実行のためのこれらの命令を発行するように、構築することができる。このような状態機械の一方は、コプロセッサ2の挙動を制御し、必要とされるようなCC__xxx__xxx命令を発行し、他方はバースト・バッファ・コントローラ7の挙動を制御し、必要とされるようなBB__xxx__xxx命令を発行する。

40

【0134】

このような状態機械を、いくつかの異なる方法において実施することができる。1つの代替手法が図13に示される。上で提示されたベクトル加算の例を参照して、この状態機械150(コプロセッサ2のためのものであるが、バースト・バッファ・コントローラ7のための同等のマシンが直接類似している)が、以下のパターンから構築された命令のシーケンスを実施する。

【0135】

50

【表 9】

CC_LX_DECREMENT、
 CC_LX_DECREMENT、
 CC_START_EXEC、
 CC_XS_INCREMENT

【0136】

メインの状態機械 150 が効果的に、より簡素な状態機械 151、152、153 に分割され、それぞれが 1 種類の命令の実行を制御する。周期および段階（これが、コプロセッサ 2 とバースト・バッファ・コントローラ 7 の間で通信する I/O ストリームに関連付けることができる周期および段階とは、何の関係も有していないことに留意されたい）が、より簡素な状態機械のそれぞれに関連付けられる。状態機械 150 のハードウェアは、典型的には、意図されたアプリケーションの要件を満たすために十分な数の、このようなより簡素な状態機械のアレイを含む。

10

【0137】

イベント・カウンタ 154 が定義される。イベント・カウンタ 154 の役割は、（この場合、コプロセッサ 2 のための）命令が順番に送信されることを可能にすることである。イベント・カウンタ 154 が増分される毎に、 $M * \text{周期}_i + \text{段階}_i = \text{イベント・カウンタ}$ の値であるような値 M が存在する場合、状態機械 i （すなわち、より簡素な状態機械 151、152、153 の 1 つ）が、比較ロジック 155 を通じた実行のために選択され、その命令が実行される。アプリケーション・ソフトウェアの責任は、2 つの異なる状態機械がこの式を満たすことができないようにすることである。この命令の実行が完了されるとき、イベント・カウンタ 154 が再度増分される。このイベントのシーケンスは、次のように要約することができる。

20

【0138】

- 1：イベントカウンタを増分する、すなわち、 $EC++$
- 2： $M * \text{周期}_i + \text{段階}_i = EC$ であるような M が存在する場合、状態機械 i を実行のために選択する
- 3：このような状態機械 i が発見された場合、状態機械 i によって記述された命令を実行する（これは、中断動作を含むことができる）
- 4：1 へ戻る

30

【0139】

命令の実行に関係する多少の余分なパラメータ（読み／書きするためのアドレス、CC_START_EXEC のための実行の長さなど）が、状態機械 150 において符号化されなければならない。複数の状態機械が、典型的には異なるパラメータを有する所与の命令を発行することができることに留意されたい。

【0140】

このシステムは、周期的挙動を生成するために特に十分に動作する。しかし、イベントが一度だけ起こらなければならない場合、これは容易に、無限周期および有限段階を有する簡素な状態機械において符号化することができ、唯一の帰結は、この簡素な状態機械が一度だけ使用されることである。

40

【0141】

この手法はそれ自体を変更することができる。たとえば、この機構に柔軟性を追加するには、可能な選択肢は、1 つまたは複数の簡素な状態機械の実行を所定の「時間ウィンドウ」に制限するために、「開始時間」および「終了時間」パラメータを簡素な状態機械に追加することである。

【0142】

これらの状態機械のプログラミングは、システムの初期化中に、たとえば、プロセッサ 1 によって割り当てられたメモリマップ・レジスタの使用を通じて起こる。代替物は、これらの状態機械をメイン・メモリ 3 の事前定義の領域から、おそらくは専用チャネルおよび直接メモリ・アクセス（DMA）機構の使用を通じて、プログラムするために必要な

50

すべてのパラメータのローディングとなる。

【0143】

2つの専用マイクロプロセッサを使用する、提案された他の代替機構は、図1のアーキテクチャのためのプログラミング・モデルに著しい修正を必要としない。すなわち、メイン・プロセッサ1をプログラムするために使用された同じ技術を、コプロセッサ2のために意図されたコマンドをバースト・バッファ・コントローラ7のために意図されたものから分割する追加ステップと共に、使用することができる。実現可能であるが、この構成は、状態機械の手法に関して不利である可能性がある。これらのプロセッサに、システムの複雑さに加えて、メイン・メモリ3または他のDRAMへのアクセスを提供することが必要となる。システムのコストおよび複雑さは、2つのマイクロプロセッサをこのように追加すること（およびそれらが非常に簡素な計算を実行するためにのみ存在することにおいて、十分に利用しないこと）によっても増大される。

10

【0144】

図1およびその代替物のアーキテクチャを越えた様々な開発も、本発明の本質の原理から逸れることなく行うことができる。3つのこれらの開発の分野が以下に記載される。すなわち、パイプライン、データ依存条件 / 知られていない実行時間、およびメモリへの非アフィン・アクセスである。

【0145】

パイプライン・アーキテクチャは、アプリケーションがそれらの入力データ・ストリーム上で複数の変換が実行されることを必要とする値を有する。たとえば、たたみこみの直後に続いて相関を行うことができる。この種類の構成を収容するために、アーキテクチャおよび計算モデルへの変更が必要となる。アーキテクチャ上では、逐次的にバッファされたCHESSEアレイ、またはより大きい区分されたCHESSEアレイ、または計算段階の間に再構成されたCHESSEアレイを提供することができる。図11Aおよび図11Bは、このようなアプリケーションを処理するために有効であり、複数のCHESSEアレイを含む、異なるパイプライン・アーキテクチャを示す。図11Aは、プロセッサ143から命令された互い違いのCHESSE / バースト・バッファ・パイプラインによる構成、およびメイン・メモリ144とのデータの交換を示し、CHESSEアレイ141がデータを第1の組のバースト・バッファ142から受信して、これを第2の組のバースト・バッファ145へ渡し、この第2の組のバースト・バッファ145がさらにCHESSEアレイ146と対話する（潜在的には、このパイプラインを、さらなる組のCHESSEアレイおよびバースト・バッファにより継続させることができる）。同期化はより複雑になり、近接したCHESSEアレイの間、および近接したバースト・バッファの組の間の通信を含むが、同じ汎用パイプラインを後に続けて、バースト・バッファの効率的な使用、およびCHESSEアレイの間の効率的な同期化を可能にすることができる。すなわち、セマフォを使用して、パイプラインの逐次段階によって実行された計算の正確さを保証することができる。

20

30

【0146】

図11Bは、異なるタイプの計算パイプラインを示し、2つのCHESSEアレイ151、156の間にSRAMキャッシュ155を有し、第1の組のバースト・バッファ152へ提供されたロード、および第2の組のバースト・バッファ157によって提供された格納を有する。プロセッサ153およびメイン・メモリ154の役割は、本質的に他の実施形態から不変である。同期化はこの構成においてそれほど困難でない可能性があるが、この構成が並行性をそれほど効果的でなく活用する可能性がある。

40

【0147】

上述のようなアーキテクチャにおけるコプロセッサの効率的な使用上の1つの制約は、コプロセッサ実施の実行時間が知られるべきであることである（効率的なスケジューリングを可能にするため）。これは、多数の媒体処理ループについて達成可能である。しかし、実行時間がコンパイル時に知られていない場合、ツールチェーンにおけるスケジューリング要件を緩和させる必要があり、プロセッサ、コプロセッサおよびバースト・バッファ

50

の間の同期化および通信プロトコルにおいて、適切な許容が行われる必要がある。コプロセッサ・コントローラは、この状況のための特定の構成も必要とする。

【0148】

別の拡張は、バースト・バッファ・メモリへの非アフィン参照を可能にすることである。上で使用されたバースト・バッファ・モデルでは、すべてのアクセスがA I + Fのタイプであり、ただしAは定数行列、Iは反復ベクトル、Fは定数ベクトルである。この制限されたアクセス・モデルの使用により、コプロセッサ・コントローラおよびプロセッサが予め、どのデータがいずれかの所与の時点で必要とされるかを知ることができ、論理ストリームの効率的な作成が可能となる。このアーキテクチャへのこの有意性は、全体として、どのように非アフィン・アクセスを完全に任意の方法で提供することができるかが不明瞭である（同期化機構が失敗するように思われる）が、非アフィン・アレイ・アクセスを使用してルックアップ・テーブルを参照することが可能になるということである。これは、ルックアップ・テーブルをバースト・バッファにロードすることによって行うことができ、次いで、コプロセッサが、後続アクセスのためのルックアップ・テーブルの開始に相対的なバースト・バッファ・アドレスを生成することができる。このようなアドレスを、それらが使用されるときより十分に前もって生成できるようにすること（場合によっては、これを、同期化機構への精練化によって達成することができる）、およびこのタイプの回帰参照を支持するように論理ストリーム機構を修正することが必要となる。

【0149】

したがって、図1のアーキテクチャへの多数の変形形態および拡張を、特許請求の範囲に記載されたような本発明から逸脱することなく、実行することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態によるシステムの基本エレメントを示す図。

【図2】図1のシステムにおいて使用されたバースト・バッファ構造のアーキテクチャを示す図である。

【図3】図2のバースト・バッファ構造のさらなる機能を示す図である。

【図4】図1のシステムにおいて使用されたコプロセッサ・コントローラの構造、および他のシステム構成エレメントとの関係を示す図である。

【図5】図1のシステム上で使用可能な計算モデルを例示するための一例を示す図である。

【図6】図5の例のための計算およびI/O動作のための時間線を示す図。

【図7】図1のシステムのためのコードを提供するために有用なツールチェーンのフロントエンドからの出力として提供された、注釈付きグラフを示す図。

【図8】図7における仕様から導出された、コプロセッサの内部構成を示す図。

【図9】32ビット・ピクセルを使用する5×5画像たたみこみのための代替アーキテクチャの性能を示す図である。

【図10】8ビット・ピクセルを使用する5×5画像たたみこみのための、図9を生成するために使用された代替アーキテクチャの性能を示す図である。

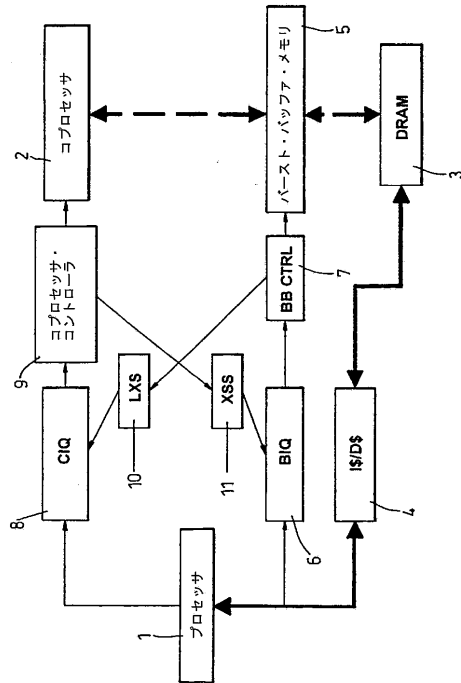
【図11A】本発明のさらなる実施形態を使用する代替パイプライン・アーキテクチャを示す図である。

【図11B】本発明のさらなる実施形態を使用する代替パイプライン・アーキテクチャを示す図である。

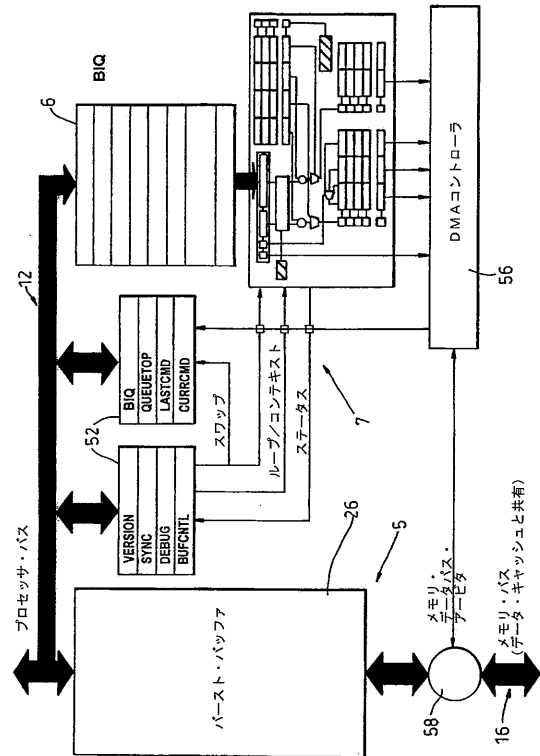
【図12】図1のアーキテクチャにおける、コプロセッサ命令待ち行列およびバースト命令待ち行列の代替として使用可能な、2つの補助プロセッサを示す図。

【図13】図1のアーキテクチャにおけるコプロセッサ命令待ち行列の代替としての状態機械の実施を示す図である。

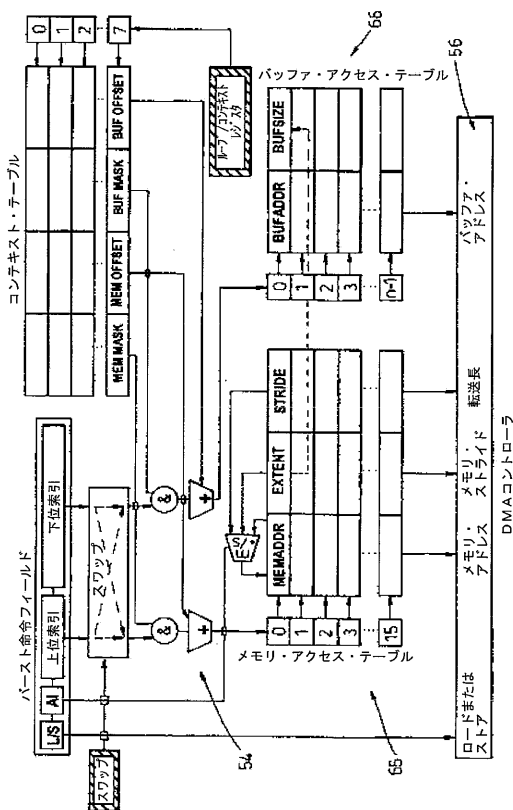
【図 1】



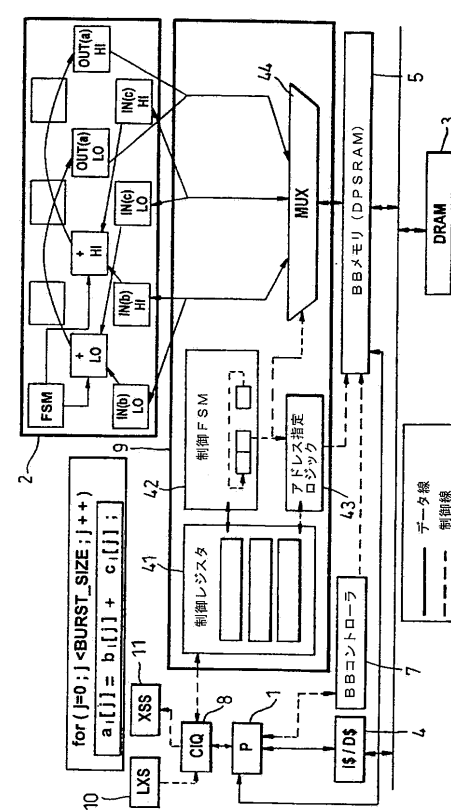
【図 2】



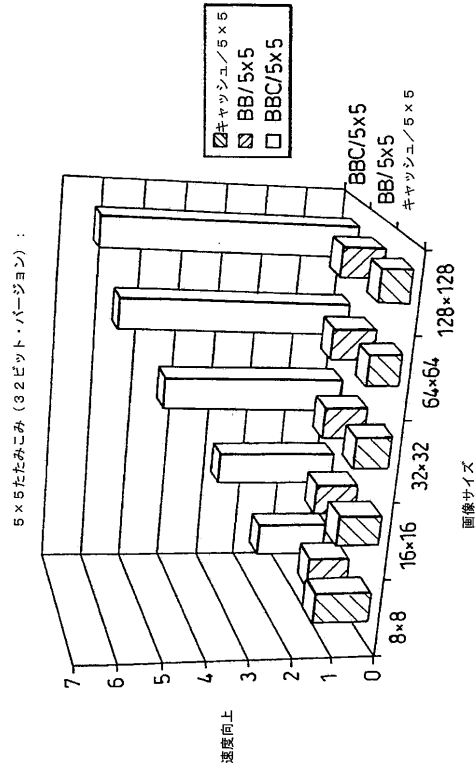
【図 3】



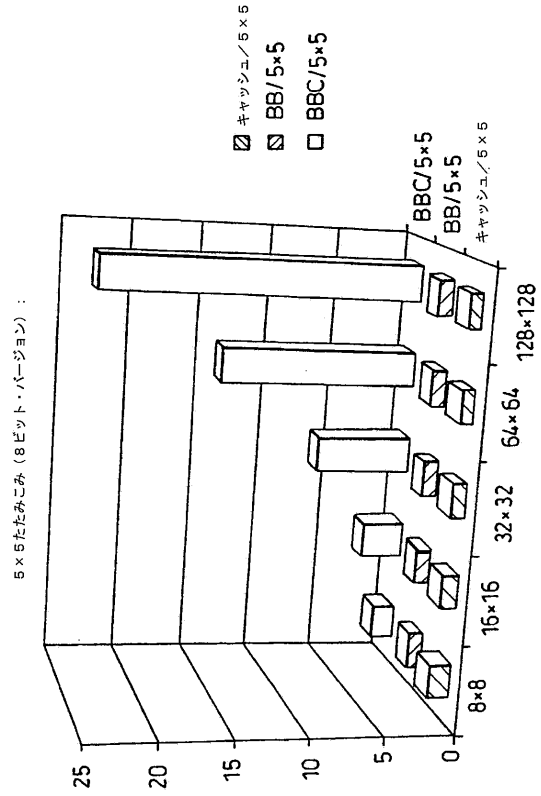
【図 4】



【図 9】



【図 10】



【図 11A】

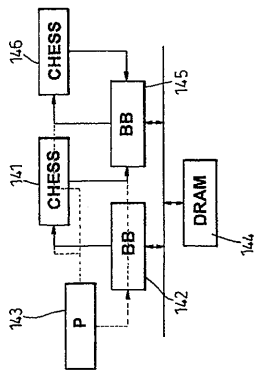


Fig. 11A

【図 11B】

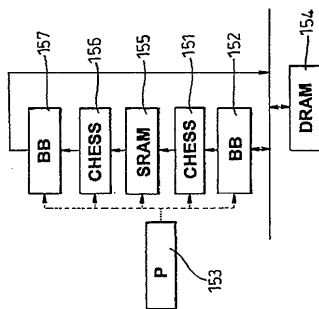
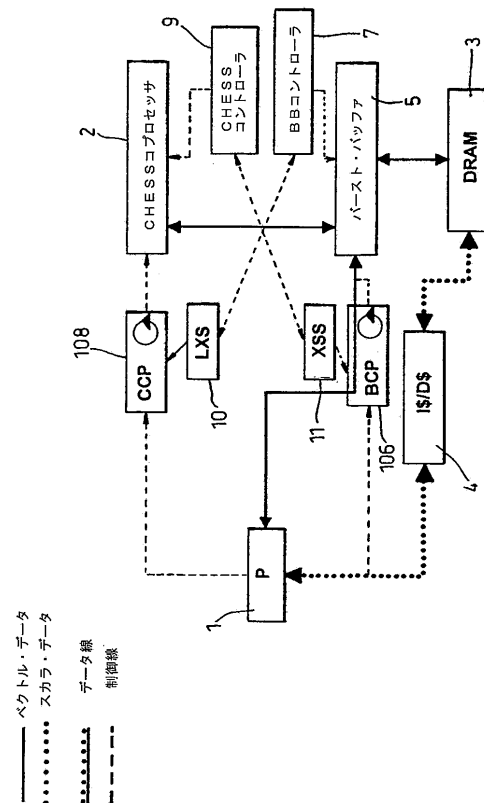
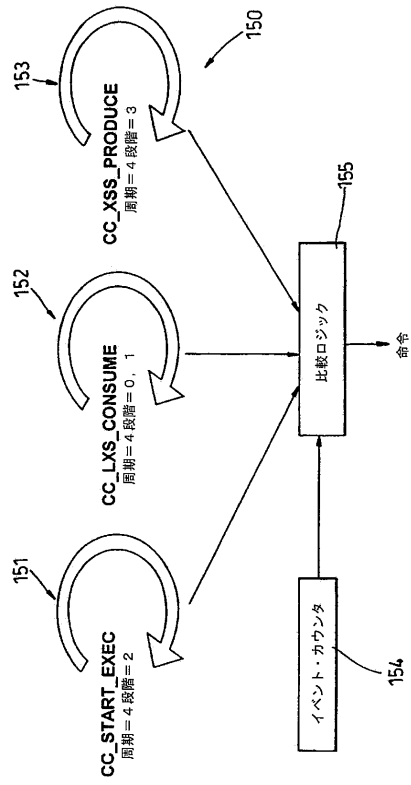


Fig. 11B

【図 12】



【図 13】



フロントページの続き

(72)発明者 マッカーシー・ドミニク・ボール
イギリス、ビーエイチ6、5ディーキュー、ボーンマス、サウスボーン、アメウッド・ロード 1
9

合議体

審判長 金子 幸一

審判官 田中 秀人

審判官 長島 孝志

(56)参考文献 特開平10-232826(JP,A)
特開平11-251442(JP,A)
特開昭56-4862(JP,A)
特開平7-281836(JP,A)
特開平11-65989(JP,A)

(58)調査した分野(Int.Cl., DB名)
G06F9/38