

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7028745号
(P7028745)

(45)発行日 令和4年3月2日(2022.3.2)

(24)登録日 令和4年2月21日(2022.2.21)

(51)国際特許分類

G 0 6 F	9/38 (2006.01)	G 0 6 F	9/38	3 7 0 C
G 0 6 F	9/50 (2006.01)	G 0 6 F	9/50	1 5 0 A
G 0 6 F	12/00 (2006.01)	G 0 6 F	12/00	5 7 1 A
G 0 6 F	12/02 (2006.01)	G 0 6 F	12/02	5 8 0 J
G 0 6 F	15/78 (2006.01)	G 0 6 F	15/78	5 6 0

請求項の数 17 (全23頁)

(21)出願番号 特願2018-171047(P2018-171047)
 (22)出願日 平成30年9月13日(2018.9.13)
 (65)公開番号 特開2019-53734(P2019-53734A)
 (43)公開日 平成31年4月4日(2019.4.4)
 審査請求日 令和3年7月21日(2021.7.21)
 (31)優先権主張番号 62/558,745
 (32)優先日 平成29年9月14日(2017.9.14)
 (33)優先権主張国・地域又は機関
米国(US)
 (31)優先権主張番号 15/825,047
 (32)優先日 平成29年11月28日(2017.11.28)
 (33)優先権主張国・地域又は機関
米国(US)
 早期審査対象出願

(73)特許権者 390019839
 三星電子株式会社
 Samsung Electronics
 Co., Ltd.
 大韓民国京畿道水原市靈通区三星路 1 2
 9
 1 2 9 , Samsung - ro , Yeo
 n g t o n g - g u , Su w o n - si
 , Gyeonggi - do , Repub
 lic of Korea
 (74)代理人 110000051
 特許業務法人共生国際特許事務所
 (72)発明者 マラディ,クリシュナ ティ
 アメリカ合衆国,カリフォルニア州 9
 5 1 3 5 ,サン ジョゼ,ロートレク ド
 最終頁に続く

(54)【発明の名称】 高効率ラーニングシステムのためのヘテロジニアスアクセラレータ

(57)【特許請求の範囲】

【請求項 1】

タスクスケジューラによって少なくとも部分的に制御される異種演算環境を含む装置であつて、

前記異種演算環境は、

前記タスクスケジューラと、

命令を実行するよう構成された固定論理回路を有する処理ユニットと、

プロセッシングインメモリ (processing-in-memory) の機能を制御する命令を含む命令を実行するよう構成された再プログラム可能な論理回路を有する再プログラム可能な処理ユニットと、

各々がデータを格納して前記再プログラム可能な処理ユニットによって制御可能な前記プロセッシングインメモリの機能を提供するよう構成された高帯域メモリダイのスタックと、を有し、

前記再プログラム可能な処理ユニットは、少なくとも部分的に前記高帯域メモリダイと積層され、

前記タスクスケジューラは、前記処理ユニットによって実行される演算タスクの第1部分及び前記再プログラム可能な処理ユニットによって実行される演算タスクの第2部分をスケジューリングするよう構成され、

前記タスクスケジューラ、前記処理ユニット、前記再プログラム可能な処理ユニット、及び前記高帯域メモリダイのスタックは、インターポーラダイ上に配置され、

前記再プログラム可能な処理ユニットは、前記演算タスクに関連するデータを前記タスクスケジューラに連結されたシステムメモリ内の中中央メモリから前記高帯域メモリダイに複写する前に、前記演算タスクの第1部分の完了に基づくトリガイベントが発生したか否かを判定するように更に構成されることを特徴とする装置。

【請求項2】

前記処理ユニットは、グラフィクス処理ユニットを含むことを特徴とする請求項1に記載の装置。

【請求項3】

前記処理ユニットは、遠隔直接メモリアクセス(RDMA: remote direct memory accesses)を介して前記高帯域メモリダイに格納されたデータにアクセスするように構成され、

10

前記再プログラム可能な処理ユニットは、前記処理ユニットから前記遠隔直接メモリアクセスをサービスするように構成された直接メモリアクセス回路を含み、

前記遠隔直接メモリアクセスは、前記高帯域メモリダイに対して発生することを特徴とする請求項1に記載の装置。

【請求項4】

前記処理ユニットは、前記高帯域メモリダイからのデータアクセスをステージングするように構成されたレジスタメモリを含むことを特徴とする請求項1に記載の装置。

【請求項5】

前記処理ユニットは、前記演算タスクの第1部分の実行を一時中断し、フラッギング基盤の同期化プロトコルに少なくとも部分的にに基づいて前記演算タスクの第2部分の実行を前記再プログラム可能な処理ユニットにオフロード(offload)するのを待機するように構成されることを特徴とする請求項1に記載の装置。

20

【請求項6】

前記処理ユニットは、前記再プログラム可能な処理ユニットをポーリング(polling)して、前記再プログラム可能な処理ユニットが前記演算タスクの第2部分を実行する準備ができているか否かを判定するように更に構成されることを特徴とする請求項5に記載の装置。

【請求項7】

前記タスクスケジューラは、前記再プログラム可能な処理ユニットが前記演算タスクの第2部分を実行する準備ができていることを前記処理ユニットに通知するように構成されることを特徴とする請求項5に記載の装置。

30

【請求項8】

中央処理ユニットによって少なくとも部分的に制御される異種演算環境を含む装置であつて、

前記異種演算環境は、

固定処理ユニットで演算タスクの第1部分を実行するように割り当て、再プログラム可能な処理ユニットで演算タスクの第2部分を実行するように割り当てるように構成された前記中央処理ユニットと、

命令を実行するように構成された固定論理回路を有する前記固定処理ユニットと、
プロセッシングインメモリ(proc_{essing}-in-memory)の機能を制御する命令を含む命令を実行するように構成された再プログラム可能な論理回路を有する前記再プログラム可能な処理ユニットと、

40

各々がデータを格納して前記再プログラム可能な処理ユニットによって制御可能な前記プロセッシングインメモリの機能を提供するように構成された高帯域メモリダイのスタックと、

前記再プログラム可能な処理ユニットは、少なくとも部分的に前記高帯域メモリダイと積層され、

前記固定処理ユニットに対して確保されたバッファ部分及び前記再プログラム可能な処理ユニットに対して確保されたバッファ部分を含む中央メモリと、

50

前記中央処理ユニット、前記固定処理ユニット、前記再プログラム可能な処理ユニット
、前記高帯域メモリダイのスタック、及び前記中央メモリがその上に配置されたインター
ポーザダイと、を有し、

前記再プログラム可能な処理ユニットは、データに対する前記演算タスクの第2部分を
実行する前に、前記演算タスクに関連するデータを前記中央メモリから前記高帯域メモリ
ダイに複写するように構成され、

前記再プログラム可能な処理ユニットは、前記演算タスクに関連するデータを前記中央
メモリから前記高帯域メモリダイに複写する前に、前記演算タスクの第1部分の完了に基
づくトリガイベントが発生したか否かを判定するように更に構成されることを特徴とする
装置。

10

【請求項 9】

前記中央処理ユニットは、前記演算タスクの第1部分を実行のために前記固定処理ユニット
に割り当てるように更に構成され、

前記固定処理ユニットは、前記演算タスクの第1部分の実行が少なくとも部分的に完了し
た時に、前記中央処理ユニットに前記演算タスクに関連するデータを複写するように構成
され、

前記中央処理ユニットは、前記演算タスクの第2部分を実行のために前記再プログラム可
能な処理ユニットに割り当て、前記演算タスクに関連するデータを前記再プログラム可
能な処理ユニットで利用することができるよう更に構成されることを特徴とする請求項 8
に記載の装置。

20

【請求項 10】

前記中央処理ユニットは、前記演算タスクに関連するデータを前記固定処理ユニットに確
保されたバッファ部分から前記再プログラム可能な処理ユニットに確保されたバッファ部
分に複写するように構成されることを特徴とする請求項 9 に記載の装置。

【請求項 11】

前記中央処理ユニットは、前記固定処理ユニットと前記再プログラム可能な処理ユニット
との間で転送されるデータのタイミングを少なくとも部分的に制御するように更に構成さ
れることを特徴とする請求項 8 に記載の装置。

【請求項 12】

前記異種演算環境は、前記処理ユニットが前記高帯域メモリダイに直接的にアクセスしな
いように構成されることを特徴とする請求項 8 に記載の装置。

30

【請求項 13】

インターポーザダイと、

異種演算環境の処理ユニットに演算タスクを割り当てるように構成された前記インター
ポーザダイ上の中央処理ユニットと、

データを格納するように構成された前記インターポーザダイ上のシステムメモリと、を有
し、

前記異種演算環境は、

命令を実行するように構成された固定論理回路を有する前記インターポーザダイ上の固定
処理ユニットと、

プロセッシングインメモリ (p r o c e s s i n g - i n - m e m o r y) の機能を制御
する命令を含む命令を実行するように構成された再プログラム可能な論理回路を有する前
記インターポーザダイ上の再プログラム可能な処理ユニットと、

各々がデータを格納して前記再プログラム可能な処理ユニットによって制御可能な前記
プロセッシングインメモリの機能を提供するように構成された前記インターポーザダイ上
の高帯域メモリダイのスタックと、を有し、

前記再プログラム可能な処理ユニットは、少なくとも部分的に前記高帯域メモリダイと積
層され、

前記中央処理ユニットは、演算タスクの第1部分を前記固定処理ユニットに割り当て、
前記演算タスクの第2部分を前記再プログラム可能な処理装置に割り当てるよう更に構

40

50

成され、

前記再プログラム可能な処理ユニットは、前記演算タスクに関連するデータを前記システムメモリから前記高帯域メモリダイに複写する前に、前記演算タスクの第1部分の完了に基づくトリガイベントが発生したことを判定するように構成されることを特徴とするシステム。

【請求項 14】

前記固定処理ユニットは、前記固定処理ユニットに割り当てられた前記演算タスクの第1部分の実行が少なくとも部分的に完了した時に、前記システムメモリに前記演算タスクに関連するデータを複写するように構成され、

前記再プログラム可能な処理ユニットは、前記演算タスクに関連するデータを要求する前記再プログラム可能な処理ユニットに割り当てられた前記演算タスクの第2部分を実行する前に、前記データを前記システムメモリから前記高帯域メモリダイに複写するように構成されることを特徴とする請求項13に記載のシステム。

10

【請求項 15】

前記固定処理ユニットは、前記演算タスクに関連するデータを遠隔直接メモリアクセス(R D M A : r e m o t e d i r e c t m e m o r y a c c e s s e s)を介して前記システムメモリに複写するように構成されることを特徴とする請求項14に記載のシステム。

【請求項 16】

前記固定処理ユニットは、グラフィクス処理ユニットを含み、

20

前記再プログラム可能な処理ユニットは、F P G A (F i e l d P r o g r a m m a b l e G a t e A r r a y)を含むことを特徴とする請求項13に記載のシステム。

【請求項 17】

前記固定処理ユニットは、前記演算タスクの第1部分が完了した時刻から前記演算タスクの第2部分が完了する時刻までの間に前記演算タスクの第1部分の実行を停止(h a l t)するように構成されることを特徴とする請求項13に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般的な回路に関するものである。特に、本発明は高効率ラーニングシステムのためのヘテロジニアスアクセラレータ(h e t e r o g e n e o u s a c c e l e r a t o r、異種加速器)に係る装置及びシステムに関する。

30

【背景技術】

【0002】

ディープニューラルネットワーク(d e e p n e u r a l n e t w o r k)のようなラーニングシステム(l e a r n i n g s y s t e m)のごとく、最近生まれたアプリケーションは他のデータセットをトレーニングして、高精度で学習するために大量の演算(c o m p u t a t i o n、c o m p u t i n g、以下、「コンピューティング」とも言う)能力及びメモリ能力を必要とする。さらに、高性能コンピューティング、グラフィクス作業(o p e r a t i o n)などのアプリケーションがデータ集中的及び演算集中的になるにつれて、エネルギー効率性及び低レイテンシが重要になる。

40

「プロセッシングインメモリ(p r o c e s s i n g - i n - m e m o r y)」として知られた技法は、低電力技術プロセスだけではなくデータが存在する場所により近接したメモリ(例えば、動的ランダムアクセスメモリ(D R A M(d y a n a m i c r a n d o m a c c e s s m e m o r y)など)とロジックを混載したダイにおける複雑な作業をスケジューリングして追加的な演算能力を提供することにより、斯かる難問に挑戦する能力を含む。

【0003】

高帯域メモリ(H B M: H i g h B a n d w i d t h M e m o r y)は3次元積層メモリ(例えば、D R A M)のための高性能ランダムアクセスメモリ(R A M)インターフェ

50

ースである。それは、多くのデータをアクセスするネットワーク装置及び高性能グラフィクスアクセラレータと共に使用される。HBMは、一般的に他のDRAM技術（例えば、DDR4、GDDR4など）より、実質的により小さいフォームファクタ（form factor）においてより少ない電力を消費しながら、より広い帯域幅を達成する。これは、多数のメモリダイ（例えば、8個）を一緒に積層することにより、達成する。このような積層は、メモリコントローラを含む選択的な基本ダイ（optional base die）を包含できる。ダイは、TSV（Through Silicon Vias、シリコン貫通ビア）及びマイクロバンプ（micro bump）により連結される。

【先行技術文献】

【特許文献】

10

【0004】

【文献】米国登録特許第8874943B2号公報

米国登録特許第9304730B2号公報

米国登録特許第20140040532A1号公報

米国公開特許第20140181453A1号公報

米国公開特許第20160379115A1号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

本発明の目的は、向上された効率性を有する高効率ラーニングシステムのための異種加速器を提供することにある。

20

【課題を解決するための手段】

【0006】

本発明の一実施例によると、例示的な実施例は、タスクスケジューラにより、少なくとも部分的に制御される異種演算環境（heterogeneous_computing_environment）を含み、前記異種演算環境は命令（instruction）を実行するように構成された固定論理回路を含む処理ユニット（以下、固定処理ユニットという）、プロセッシングインメモリ（processing-in-memory）の機能を制御する命令を含む命令を実行するように構成された再プログラム可能な論理回路を含む処理ユニット（以下、再プログラム可能な処理ユニットという）、及び高帯域メモリダイのスタックを含み、前記高帯域メモリダイの各々はデータを格納し、前記再プログラム可能な処理ユニットにより制御可能なプロセッシングインメモリの機能を提供するように構成され、前記再プログラム可能な処理ユニットは少なくとも部分的に前記高帯域メモリダイと積層され、前記タスクスケジューラは、前記固定処理ユニット及び前記再プログラム可能な処理ユニットの間ににおける演算タスクの分担をスケジューリングするように構成される装置を提供する。

30

【0007】

本発明の他の実施例によると、他の例示的な実施例は、中央処理ユニットにより、少なくとも部分的に制御される異種演算環境を含み、前記異種演算環境は固定処理ユニット及び再プログラム可能な処理ユニットの間ににおいて演算タスクを割り当てるように構成された前記中央処理ユニット、命令を実行するように構成された固定論理回路を含む前記固定処理ユニット、プロセッシングインメモリ（processing-in-memory）の機能を制御する命令を含む命令を実行するように構成された再プログラム可能な論理回路を含む前記再プログラム可能な処理ユニット、高帯域メモリダイのスタック、及び前記固定処理ユニットに専用留保された（reserved）バッファ領域及び前記再プログラム可能な処理ユニットに専用留保されたバッファ領域を含む中央メモリを包含し、前記高帯域メモリダイの各々はデータを格納し、前記再プログラム可能な処理ユニットにより制御可能なプロセッシングインメモリの機能を提供するように構成され、前記再プログラム可能な処理ユニットは少なくとも部分的に前記高帯域メモリダイと積層される装置を提供する。

40

50

【0008】

本発明の他の実施例によると、もう一つの実施例は異種演算環境の処理ユニットに演算タスクを割り当てるように構成された中央処理ユニット、データを格納するように構成されたシステムメモリ、及び前記異種演算環境を含む異種アクセラレータのマルチチップモジュールを含むシステムを提供する。

前記異種アクセラレータのマルチチップモジュールは、命令を実行するように構成された固定論理回路を含む固定処理ユニット、プロセッシングインメモリ (processing - in - memory) の機能を制御する命令を含む命令を実行するように構成された再プログラム可能な論理回路を含む再プログラム可能な処理ユニット、及び高帯域メモリダイのスタックを含み、前記高帯域メモリダイの各々はデータを格納し、前記再プログラム可能な処理ユニットにより制御可能なプロセッシングインメモリの機能を提供するように構成され、前記再プログラム可能な処理ユニットは少なくとも部分的に前記高帯域メモリダイと積層される。

【0009】

1つ又はそれ以上の具現の詳細な説明は、添付された図面及び以下の詳細な説明で掲載される。他の特徴は詳細な説明、図面及び特許請求の範囲から表されることができる。

【0010】

特許請求の範囲でもう少し完全に説明されるように、実質的に図面の少なくとも1つと共に説明されるか、又は図示されるような高効率のラーニングシステムのための異種アクセラレータのシステム及び方法のためのシステム及び／又は方法が提供される。

【発明の効果】

【0011】

本発明の実施例によると、高帯域メモリダイのスタックに積層・直結された再プログラム可能な処理ユニットと固定処理ユニットとに対してCPU(タスクスケジューラ)によりタスクが振り分けられ、好ましくは、各ユニットの入出力データが遠距離直接メモリアクセス(RDMA)により高速で転送されるので、向上された効率性を有するラーニングシステムのための異種アクセラレータが提供される。

【図面の簡単な説明】

【0012】

【図1】図1は、本発明によるシステムの例示的な実施例のブロック図である。

30

【図2】図2は、本発明によるシステムの例示的な実施例のブロック図である。

【図3】図3は、本発明によるシステムの例示的な実施例のブロック図である。

【図4】図4は、本発明によるシステムの例示的な実施例のブロック図である。

【図5】図5は、本発明によるシステムの例示的な実施例のブロック図である。

【図6】図6は、本発明によるシステムの例示的な実施例のブロック図である。

【図7】図7は、本発明の原理によって形成された装置を包含できる情報処理システムの例示的なブロック図である。

【0013】

多様な図面で類似した参照記号は類似した構成を示す。

【発明を実施するための形態】

【0014】

多様な例示的な実施例が一部の例示的な実施例を示す添付された図面を参照してさらに詳細に説明される。しかしながら、本文に開示された内容は多様な他の形態で具現されることができ、本文に説明された例示的な実施例に制限されるものと解釈されてはならない。代わりに、このような例示的な実施例は詳細な説明が完全であり、当業者により本発明の技術的思想が完全に伝達される能够性を有するように提供される。図面において、階層及び領域のサイズ及び相対的なサイズは明確性のために誇張される。

【0015】

構成又は階層が他の構成又は階層と「連結された(on, connected to, or coupled to)」ものと示される場合、これは直接的に(directly

10

20

40

50

) 他の構成又は階層と連結されうるか又は1つ以上の中間(*intervening*)の構成又は階層が存在すると理解されるはずである。対照的に、構成が他の構成又は階層と「直接的に連結された(*directly on, directly connected to, or directly coupled to*)」ものと示される場合、中間の構成又は階層が存在しない。類似した参照番号は、全体的に類似した構成を示す。本文で使用されるように「及び/又は(*and/or*)」は、連関されて羅列された目録中で1つ又はそれ以上の組合せの一部及び全部を含む。

【0016】

「第1(*first*)」、「第2(*second*)」、「第3(*third*)」などの用語は多様な要素(*element*)、構成部品(*component*)、領域、階層、及び/又は区域を説明するために本文で使用されるが、このような要素、構成部品、領域、階層、及び/又は区域は、このような用語に限定されないことがよく理解されるはずである。斯かる用語は1つの要素、構成部品、領域、階層又は区域を他の一つの要素、構成部品、領域、階層、又は区域から区分するためにだけ使用される。即ち、以下で記載される第1要素、構成部品、領域、階層、又は区域は本発明の思想及び範囲から逸脱せずに第2要素、構成部品、領域、又は区域と称することができる。

【0017】

「~の下(*beneath, below, lower, under*)」、「~の上(*above, upper*)」などの空間的に相対的な用語(*spatially relative terms*)は、図面で図示された他の一つの要素又は特徴と1つの要素又は特徴との関連性を容易に説明するため、本文で使用される。空間的に相対的な用語は、図面に図示された指向性に追加的に作業又は使用で、装置の他の指向性(*orientations*)を含むと意図されることをよく理解されるはずである。例えば、図面で装置が上下反転される場合、他の要素又は特徴の「下(*below or beneath or under*)」として説明された要素は、他の要素又は特徴の「上(*above*)」に向くはずである。即ち、「下(*below, under*)」の例示的な用語は上及び下の方向を全て包含できる。装置が他の方向(例えば、90度回転するか、又は他の方向)に指向される場合、本文で使用される空間的に相対的な説明は、その指向方向に合わせて解釈されなければならない。

【0018】

同様に、「ハイ(*high*)」、「ロー(*low*)」、「(ブルアップ(*pull up*)」、「ブルダウン(*pull down*)」、「1」、「0」などの電気的な用語は、図面で図示されたように、他の電圧レベル、他の構成又は特徴と相対的な電圧レベル及び電流を表す説明の便宜のために詳細な説明で使用される。電気的に相対的な用語は、図面に図示された電圧又は電流に追加的に使用又は作業において、装置の他の基準電圧を含むことと意図される。例えば、図面で装置又は信号が反転されるか又は他の基準電圧、電流、又は電荷を使用する場合、「ハイ(*high*)」又は「(ブルアップ(*pull up*)」で説明される構成は、新たな基準電圧又は電流と比較して「ロー(*low*)」又は「ブルダウン(*pull down*)」である。即ち、「ハイ(*high*)」の例示的な用語は、相対的に低いか又は高い電圧又は電流を全て包含できる。装置は、他の電気的なフレームの基準に基づき、詳細な説明で使用される電気的に相対的な説明により解釈できる。

【0019】

詳細な説明で使用される用語は、専ら特定の実施例に対する説明のためであり、本発明の限定を意図しない。詳細な説明で使用されるように、単数形態(*singular forms*)は、明確に別なものとして定義されない限り、複数形態(*plural forms*)を含むように意図される。「包含する(*comprise*)」の用語は、詳細な説明で使用される場合、列挙された特徴、ステップ、作業、要素及び/又は構成部品の存在を特定するが、1つ又はそれ以上の他の特徴、ステップ、作業、要素、構成部品、及び/又はそれらのグループの追加又は存在を排除しない。

【0020】

10

20

30

40

50

例示的な実施例が、理想的な実施例（及び中間構造）の例示的な図面である断面図を参照して詳細な説明で述べられる。このように、例えば、製造技術及び／許容誤差の結果としての図面の形状からの変形が予想されなければならない。即ち、例示的な実施例は、詳細な説明に図示された特定の形状の領域に限定されるものとして解釈されてはならないし、例えば、製造からもたらされる形状の偏差を包含しなければならない。例えば、直四角形で図示された注入された領域は一般的に、注入された領域から注入されない領域への二進変化であるよりは、丸い又は曲線のフィーチャー（feature）及び／又はエッジ（edge）での注入濃度の勾配を有するであろう。同様に、注入によって形成されたベリード（buried）領域はベリード領域と注入が起こる表面間の領域に若干の注入をもたらす。従って、図面に図示された領域は、本質的に概略的であり、その形状はデバイスの領域の実際の形状を説明するためではないし、本発明の範囲を制限しない。

10

【0021】

別に定義されない限り、本文で使用される全ての用語（技術的及び科学的用語を包含する）は当業者により共通的に理解される意味を有する。なお、一般的に使用される辞典で定義されたような用語は、関連技術及び／又は本明細書と関連してその意味と一致する意味を有すると解釈されるべきであり、本文で定義されない限り、理想的や過度に形式的な意味で解釈されてはならない。

【0022】

以下で、例示的な実施例が添付された図面を参照して詳細に説明されるであろう。

20

【0023】

図1は、本発明によるシステム100の例示的な実施例のブロック図である。図示された実施例で、システム100は複数の集積回路（IC：Integrated Circuit）ダイを包含できる。斯かる実施例で、ICはマルチチップモジュール（MCM：Multi-Chip Module）として配列される。

20

【0024】

しばしば、MCMは、通常、統一基板（unifying substrate）上に複数の集積回路（IC又はチップ）、半導体ダイ、及び／又は他の個別部品（discrete component）が集積されて、使用時には単一の構成部品（例えば、より大きいIC）のように扱われる、複数の導電端子、即ちピン（pin）を含むパッケージのような電子組立体（製品）であり得る。しかし、上述された構成は単に説明のための実施例であり、本発明はこれに限定されない。

30

【0025】

図示された実施例においてシステム100は、HBMダイ112上に形成された高帯域幅のメモリ（HBM：High Bandwidth Memories）のスタック108を包含する。上述されたように、HBMは、データを格納（store）してメモリアクセスが標準DRAM又はシステムメモリよりも速く且つ効率的となるように構成できる。一実施例において、HBMダイ112はプロセッシングインメモリ機能（processsing-in-memory functionality）を提供するように構成される。多様な実施例において、HBMダイ112は互いに垂直的に、即ち一つのHBMダイ112が他の一つの上方に積層されているので、TSV（Through-Silicon on Vias）、マイクロバンプ又は他の相互チップ接続（図示せず）を通じてHBMダイ112間で通信を遂行できる。

40

【0026】

多様な実施例において、HBMダイ112のスタック108は、HBMコントローラダイ110を包含する。HBMコントローラダイ110は、標準メモリアクセスの調整機能（例えば、ページテーブル変換、アドレスマッピング、ライト（write）の組合せなど）を遂行するように構成される。多様な実施例において、HBMコントローラダイ110はHBMダイ112向けに特化したメモリ管理ユニット（MMU：Memory Management Unit）を包含する。

【0027】

50

図示された実施例において、システム100はプログラム可能か又は再プログラム可能な処理ユニット106 (PPU又はRPU: Programmable or Reprogrammable Processing Unit、図1ではRPUで代表)を包含する。多様な実施例で、RPU106は動的にプログラムされて多様な機能を遂行するか又は特定の命令を実行するロジック回路を包含すると。一実施例において、RPU106はHBMダイ112のプロセッシングインメモリの機能を制御する命令を含む命令を実行するように構成される。多様な実施例において、RPU106はFPGA (Field-Programmable Gate Array)、ASIC (Application Specific Integrated Circuit)、LUT (Look-Up Table)、PAL (Programmable Array Logic)などの、しかしこれに限定されないデバイスを包含できる。上述の内容は、単に説明のための幾つかの実施例であり、本発明はこれに限定されない。

【0028】

多様な実施例において、RPU106は、HBMダイ112のスタック108の下方に、又はHBMダイ112と共に積層される。斯かる実施例で、RPU106は、スタック108と直接的に通信するように構成され、スタック108へのアクセスをシステム100の他の構成部品（例えば、固定的なGPU104）へ提供するように構成されうる。

【0029】

多様な実施例において、RPU106は粗粒子 (coarse grain) な、機能即ち命令実行の再構成性 (reconfigurability) を提供できる。他の実施例で、RPU106は微粒子 (fine grain) な、機能即ち命令実行の再構成性を提供できる。一部の実施例で、RPU106により遂行される作業は、例えば、しかしこれに限定されないが、乗加算演算 (multiply-accumulate operations)、データシャフリング (data shuffling)、データ転換 (data transposition)などを包含する、HBMダイ112におけるプロセッシングインメモリの機能制御を包含する。上述されたのは、単に説明のための幾つかの実施例であり、本発明はこれに限定されない。多様な実施例で、RPU106自身及びRPU106のHBMスタック108への物理的に近接したアクセスは、固定処理ユニット又は回路、若しくは、相対的に遠隔配置された処理ユニット又は回路より、命令 (instruction)、タスク (task)、又は作業 (operation) のより高速実行を提供できる。

【0030】

図示された実施例において、システム100は変更不能な、即ち固定処理ユニット104 (FPU: Fixed Processing Unit)を包含する。（図中では、FPUの代表例である後述のGPUを並記してFPU/GPUと記載。）このような実施例において、FPU104は多様な命令を実行するか又はロジック作業を遂行するように構成される。FPU104は、FPU104の製造中に静的に形成されて変更不能即ち固定されている論理回路を包含する。但し、多様な構成の設定により固定論理回路の機能具現方法を変更できる。

【0031】

多様な実施例において、FPU104は、例えばグラフィクス処理ユニット (GPU: Graphic Processing Unit)、暗号化処理ユニット (encryption processing unit)、物理処理ユニット (physics processing unit)、マシンラーニング処理ユニット (machine learning processing unit) などの特化された処理ユニット (specialized processing unit)を包含する。上述されたのは、単に説明のための幾つかの実施例であり、本発明がこれに限定されない。多様な実施例において、FPU104はスタック108と直接的に通信しないように（積層ではなく）並置される。

【0032】

10

20

30

40

50

多様な実施例において、システム 100 はインタポーザ (interposer) ダイ又はレイヤ (layer) 120 を包含する。斯かる実施例で、インタポーザ 120 はシリコン (silicon) 又は他の基板 (substrate) を包含でき、FPU104 及び RPU106 (及びスタック 108) の間の通信経路を提供する。斯かる実施例で、インタポーザ 120 は RPU106 及び FPU104 の下方にあって、RPU106 及び FPU104 と互いに連結される。

【0033】

図示された実施例において、システム 100 はパッケージ基板ダイ 122 を包含する。パッケージ基板ダイ 122 は、他の構成部品 (ダイ) 間、及びシステム 110 の外部との間を通信するように構成される。多様なダイ (122, 120, 104, 106 など) は、複数のはんだバンプ (solder bumps) 124 により連結される。

10

【0034】

図 2 は、本発明によるシステム 200 の例示的なブロック図である。図示された実施例で、システム 200 は複数の集積回路 (IC) ダイを包含できる。斯かる実施例で、IC はマルチチップモジュール (MCM) として配列される。

【0035】

図示された実施例において、システム 200 は上述されたように、HBM ダイ 112 及び HBM コントローラ ダイ 110 からなるスタック 108、RPU106、FPU104、インタポーザ 120 及びパッケージ基板 122 を包含する。図示された実施例において、複数の可能な追加されたダイ又は構成部品 (component) が図示される。多様な実施例で、1 つ又はそれ以上の斯かるダイが、以下で説明されるように、システム 200 に包含される。他の実施例で、一つ又はそれ以上の斯かるダイ又は構成部品は、他の IC パッケージの部分として又はより大きいシステムの部分として包含されることができる。例えば、中央処理ユニット (CPU) 及び / 又はシステムメモリがデスクトップ (desk top) 又はラップトップコンピュータ (laptop computer) の部分として包含されうる。

20

【0036】

多様な実施例において、システム 200 は中央処理ユニット (CPU: Central Processing Unit) を包含する。斯かる実施例で、CPU202 はダイに包含され、マイクロプロセッサ及び / 又は複数のプロセッサコアを包含する。CPU202 は命令を処理して管理するシステムの主構成部品 (primary component) である。CPU202 は、OS (operating system、オペレーティングシステム) 及びアプリケーションの実行を主に担当する。多様な実施例で、CPU202 は ARM 社又は Intel 社により生産されるような汎用演算構造 (general purpose computing architecture) を包含する。上述されたのは、単に説明のための幾つかの実施例であり、本発明の範囲はこれに限定されない。図示された実施例において、CPU202 はシステム 200 の異種コンピューティング環境 (heterogeneous computing environment) を制御し、FPU104 及び RPU106 の間で演算作業を分割して割り当てるように構成される。

30

【0037】

多様な実施例において、システム 200 は CPU メモリダイ 212 を包含する。CPU メモリダイ 212 は CPU202 の上部に積層されうる。斯かる実施例で、CPU メモリはデータを格納するように構成され、CPU202 から直接的にアクセスできる。一部の実施例で、CPU メモリ 212 はシステムメモリを包含する。他の実施例で、システムメモリ (例えば、DRAM) は、システム 200 の外部に位置する。

40

【0038】

多様な実施例において、システム 200 は FPU メモリダイ 214 を包含する。一部の実施例で、FPU メモリダイは FPU の上方に積層される。このような実施例で、FPU メモリ 214 はデータを格納するように構成され、FPU104 から直接的にアクセスでき

50

る。

【0039】

多様な実施例において、システム200の三つのプロセッサ(RPU106、FPU104及びCPU202)には、多様なコンピューティングタスク、命令、又はカーネル(kernel)の実行がそれらの間に割り当てられる。このような脈絡で、「カーネル(kernel)」は、タスク又は定義可能なサブタスク(definable sub-task)を実行するように共にグループ化された1つ又はそれ以上の命令として定義される。

カーネルの境界(bounds of kernel)はサブルーチンと一致する場合としない場合があるので、サブルーチンの同義語(synonymous terms)と見做すべきではない。斯かる脈絡で、「カーネル(kernel)」の用語は「コンピュートカーネル(compute kernel)」の用語と少し類似して、「OSカーネル(operating system kernel)」、「カーネル方式(kernel method)」、「イメージカーネル(image kernel)」のようなカーネルの他の定義、又は数学で使用されるカーネルの多様な定義と混同してはならない。

【0040】

このような実施例において、システム200の三つのプロセッサ(RPU106、FPU104、及びCPU202)はタスクを分配して効率性を向上し、電力の使用及びメモリ帯域幅の消耗を低減できる。多様な実施例で、これはARM社のビッグリトル異種演算構造(big.LITTLE heterogeneous computing architecture)と類似するが、同一ではない。ARM社のビッグリトル異種演算構造においては、バッテリ節減的で低速のプロセッサコア(LITTLE)は、相対的にさらに強力で電力消耗が多いプロセッサコア(big)と連結される。このような実施例で、相異なる形態のコア間で(例えば、ビッグ(big)及びリトル(LITTLE)コア、RPU106、FPU104、CPU202などの間で)、作業負荷(workloads)がスワップ(swap、交換)できる。斯かる実施例で、マルチコアシステム200は動的な演算要求に適応してより少ない電力を使用できる(又はより速くなれる)。上述されたのは単に説明のための幾つかの実施例であり、本発明はこれに限定されない。

【0041】

図3は、本発明によるシステム300の例示的な実施例のプロック図である。多様な実施例で、システム300は図1及び/又は図2に図示されたダイ又は他のICから構成される。

【0042】

図示された実施例において、システム300は、固定論理回路を含む処理ユニット(以下、固定処理ユニット(FPU)という)304及び再プログラム可能な論理回路を含む処理ユニット(以下、再プログラム可能な処理ユニット(RPU)という)306、並びにそれらに対して演算(コンピューティング)タスクを割り当てるように構成されたCPU(即ち、ホストプロセッサ又はタスクスケジューラ)302を包含する。多様な実施例で、タスクスケジューラ302は、少なくとも部分的にシステム300を制御するように構成される。多様な実施例で、タスクスケジューラ302はプロセッサ(例えば、CPU)を包含するが、他の実施例で、作業スケジューラ302は、より汎用ではない回路(即ち、専用回路)を包含する。

【0043】

図示された実施例において、システム300は、上述されたように、固定処理ユニット(FPU)304(例えば、GPU、DSP(Digital Signal Processor)など)を包含する。システム300は、上述されたように、再プログラム可能な処理ユニット(RPU)306(例えば、FPGAなど)を包含する。システム300は、上述されたように、更にHBMスタック308を包含する。

【0044】

図示された実施例において、FPU304及びRPU306は制御経路及び/又はデータ

10

20

30

40

50

経路を集積された形で包含するように構成される。斯かる実施例で、コンピューティングタスク（及びそれに連関されたデータ）は、2つのプロセッサ（304, 306）間で転送されるか又はハンドオフ（hand-off）される。以下で説明されるように、多様な実施例で、タスク及びデータのスイッチングは自己調節方式で発現されるか、又はCPU302により調節される。

【0045】

図示された実施例において、FPU304は、プログラマー又はユーザに、公知のFPU（例えば、GPU）プログラミングモデル又はAPI（Application Programming Interfaces）、及びFPUと連関されたデータアクセスプロトコル又はモデルの利用を可能にする。反面、RPU304の再プログラム能力は、効率性を（例えば、電力、時間など）を増加し、動的に変更可能な特定タスク専用の回路（task specific circuitry）を具現できる。

10

【0046】

図4は、本発明によるシステム400の例示的な実施例のブロック図である。多様な実施例で、システム400の各部分は図1及び/又は図2に図示されたダイ又は他のICで構成される。システム400は、図3に簡略に図示されたタスク転送の一実施例を示す。

【0047】

図示された実施例において、システム400は、上述されたように、固定処理ユニット304（例えば、GPU、DSP（Digital Signal Processor）など）を包含する。システム400は、上述されたように、RPU306（例えば、FPGなど）を包含する。システム400は、上述されたように、HBM STACK 308を包含する。

20

【0048】

図示された実施例において、システム400はデータを格納するように構成され、CPU302により又はその制御下に直接的にアクセスされるシステムメモリ412を包含する。システム400はデータを格納し、FPU304により又はその制御下に直接的にアクセスされるように構成されたFPUメモリ414を包含する。

【0049】

図示された実施例において、CPU302は連結バス（interconnect bus）を介しFPU304及びRPU306と連結される。多様な実施例で、連結バス402は、例えば、PCIe（Peripheral Component Interconnect express）プロトコルのような直列プロトコルを使用できる。上述されたのは、単に説明のための実施例であり、本発明がこれに限定されない。

30

【0050】

図示された実施例において、CPU（又はホスト、又はタスクスケジューラ）302は第1タスク、サブタスク、又はタスク490の一部をFPU304に割り当てる。FPU304は例えば第1サブタスクと連関された命令の実行を開始し、タスク490と連関されたデータ404のアクセスを望む。図示された実施例で、データ404はHBM STACK 308に格納されている場合がある。

40

【0051】

図示された実施例において、RPU306は直接メモリアクセス（DMA：Direct Memory Access）、RDMA（Remote DMA）エンジン（又はRDMA回路）464を包含する。RPU306がHBM STACK 308へのアクセスをゲート（制御）できる。多様な実施例で、RPU306がRDMAエンジン464を実行する場合、他のプロセッサはRPU306を介してHBM STACK 308からのDMA呼出又は要請を遂行できる。そこでRDMAエンジン464はDMA要請を提供するか又は遂行できる。

【0052】

図示された実施例において、FPU304はRDMAアクセスの要請をRDMAエンジン464へ提供する。RDMAはHBM STACK 308からデータ404を回収する、即ち

50

、リード(*read*)し、それを F P U 3 0 4 に提供する。多様な実施例で、F P U 3 0 4 は 1 つ又はそれ以上のアクセスレジスタ 4 4 4 を包含する。斯かる実施例で、データ 4 0 4 又はデータ 4 0 4 の一部は、アクセスレジスタ 4 4 4 (例えは、2 5 6 メガバイトの部分) にバッファリングされる、即ち、一時記憶(*stage*)される。以後、データ 4 0 4 は F P U メモリ 4 1 4 に複写される。以後、F P U 3 0 4 はデータ 4 0 4 に対して割り当てられたタスク 4 9 0 の一部を実行できる。

【 0 0 5 3 】

図示された実施例において、C P U (又はタスクスケジューラ) 3 0 2 は第 2 タスク、サブタスク又はタスク 4 9 0 の一部を R P U 3 0 6 に割り当てる。多様な実施例で、R P U 3 0 6 はタスク 4 9 0 の第 2 の部分を R P U 3 0 6 上で効率的に又は好ましい形で実行するロジック回路としてプログラムされうる。例えは、R P U 3 0 6 に割り当てられたタスク 4 9 0 の一部は H B M スタック 3 0 8 の H B M ダイのプロセッシングインメモリの機能の制御を包含する。F P U 3 0 4 がタスク 4 9 0 の第 1 の一部を完了した場合、F P U 3 0 4 は R P U 3 0 6 にカーネルをオフロード(*offload*)すること又はタスク 4 9 0 の残りの実行を伝送するのを中断する。

10

【 0 0 5 4 】

図示された実施例において、R P U 3 0 6 は、タスク 4 9 0 の第 2 部分の処理の開始の準備完了を報らせるレディーフラッグ(*ready flag*) 4 6 2 又は他の何らかの指示を包含する。一部の実施例で、R P U 3 0 6 は多数のレディーフラッグ 4 6 2 を包含し、各々は多様なロジック回路と連関され、該ロジック回路はそれぞれタスクを遂行できる。F P U 3 0 4 は、R P U 3 0 6 を周期的にチェックするか又はポーリングして R P U 3 0 6 へタスク 4 9 0 の処理を転送する時期であるかを確認するように構成されたポーリング回路 4 4 2 を包含する。多様な実施例で、他のフラッグ基盤の同期化プロトコルが採用される。

20

【 0 0 5 5 】

F P U 3 0 4 が、R P U 3 0 6 が準備されたことを認識した場合、タスク 4 9 0 又はデータ 4 0 4 の所有権が 2 つのプロセッサ (3 0 4 、 3 0 6) 間で転送される。以後、R P U 3 0 6 は割り当てられたタスク 4 9 0 の第 2 部分の実行を開始する。多様な実施例で、C P U 3 0 2 は F P U 3 0 4 に R P U 3 0 6 がタスク 4 9 0 の第 2 部分の実行を開始する準備完了を報らせるように構成される。斯かる実施例で、R P U 3 0 6 は、C P U 3 0 2 にこれを報らせることができるか、又は C P U 3 0 2 は、R P U 3 0 6 の状態をモニタする。(例えは、C P U 3 0 2 は、R P U 3 0 6 が他のタスクを完了したことを認識できる。)

30

【 0 0 5 6 】

多様な実施例において、データ 4 0 4 は F P U メモリ 4 1 4 から H B M スタック 3 0 8 へ上述された R D M A プロトコルを使用して転送される。以後、R P U 3 0 6 は H B M スタック 3 0 8 に格納されたデータ 4 0 4 を直接的にアクセスできる。R P U 3 0 6 がタスク 4 9 0 の第 2 部分を実行する間、F P U 3 0 4 は中断又は停止するか又は非同期的に他のタスクを遂行して F P U 3 0 4 の演算電力を浪費しないようにする。上述されたのは、単に説明のための幾つかの実施例であり、本発明はこれに限定されない。

40

【 0 0 5 7 】

一部の実施例において、タスク実行は類似した方式で、しかし逆向きの手順で、R P U 3 0 6 から F P U 3 0 4 (又は C P U 3 0 2) に変更される。斯かる実施例で、F P U 3 0 4 、 R P U 3 0 6 、及び C P U 3 0 2 間のデータ同期化は、上述されたとことと類似に、現出できる。

【 0 0 5 8 】

図 5 は、本発明によるシステム 5 0 0 の例示的な実施例のブロック図である。多様な実施例で、システム 5 0 0 の部分は、図 1 及び / 又は図 2 に図示されたダイ又は他の I C で構成される。システム 5 0 0 は、図 3 に簡略に図示されたタスク転送の一実施例を示す。

【 0 0 5 9 】

図示された実施例において、システム 5 0 0 は、上述されたように、固定処理ユニット 3

50

04(例えば、GPU、DSPなど)を包含する。システム500は、上述されたように、RPU306(例えば、FPGAなど)を包含する。システム500は、上述されたように、HBM STACK308をも包含する。

【0060】

図示された実施例において、システム500はデータを格納するように構成され、FPU304により又はその制御下に直接的にアクセスされるFPUメモリ414を包含する。システム500はFPU304と連関されたデータを格納するように構成されたFPUバッファ又は中央メモリ514を包含する。システム500はRPU306と連関されたデータを格納するように構成されたRPUバッファ又は中央メモリ516をも包含する。このような実施例で、FPUバッファ514及びRPUバッファ516はCPU302により制御されるか又はアクセスされる。多様な実施例で、FPUバッファ514及びRPUバッファ516は、システムメモリ(例えば、図4に図示されたシステムメモリ412)の部分又は領域であるか又は別途のメモリ回路である。

10

【0061】

図示された実施例において、CPU302は連結バス(*interconnect bus*)402を介しFPU304及びRPU306と連結される。多様な実施例で、連結バス402は、例えば、PCIe(*Peripheral Component Interconnect express*)プロトコルのような直列プロトコルを使用できる。上述されたのは、単に説明のための実施例であり、本発明はこれに限定されない。

20

【0062】

上述されたように、CPU(又はタスクスケジューラ)302は多様なタスク590又はタスク590の或る部分を、実行のため、FPU304及びRPU306に分散するか又は割り当てる。多様な実施例で、斯かる割当は、少なくとも部分的に、どのプロセッサが利用可能であり、且つ、より効率的にタスクを実行できるか、という点に基づく。上述されたように、CPU302はFPU304にタスク590の第1部分を割り当てて、タスク590の第2部分をRPU306に割り当てられる。例えば、RPU306に割り当てられた多様なタスク590の部分は、HBM STACK308のHBMダイのプロセッシングインメモリの機能の制御を包含できる。上述されたのは、単に説明のための幾つかの実施例であり、本発明はこれに限定されない。

30

【0063】

図示された実施例において、FPU304は上述されたように、割り当てられたタスクを実行する。斯かる実施例で、FPU304は一時記憶(*temporary*)キャッシュ、即ち、作業用(*working*)メモリとしてFPUメモリ414を使用する。割り当てられたタスク、サブタスク又はタスク590の一部が完了された場合、FPU304はデータをFPUバッファ514に書き込む。

40

【0064】

このような実施例においては、これはFPU304及びFPUバッファ514間で媒介体として作業するか、又は少なくともデータ504がFPUバッファ514に書き込み中であることを報らされているCPU302を包含する。一部の実施例で、これは、メモリ割当機能又は手続呼出(*procedure call*)(例えば、GPUの場合で、「CUDAMalloc()」の呼出が生成される。「CUDAMalloc()」はNVIDIA社が開発したGPU向けの汎用並列コンピューティングプラットフォームで使用される関数である)を通じて行われる。斯かる実施例で、FPU304はCPU302にデータを複写し、CPU302はFPUバッファ514にデータを順次的に複写する。

【0065】

図示された実施例において、以後、CPU302はデータ504をFPUバッファ514からRPUバッファ516へ転送する。斯かる実施例で、CPU302は、FPU304及びRPU306間(少なくともFPUバッファ514と、RPUバッファ516との間)でデータ同期化及びデータ転送のタイミングを制御する。

【0066】

50

図示された実施例において、FPUバッファ514とRPUバッファ516間のデータ504の複写が完了された場合、CPU302はトリガイベント(triggering event)546を開示してデータ504が準備されたことをRPU306に報らせる。多様な実施例で、このようなトリガイベント546は、インタラプト信号、連結バスを介したメッセージ、又はRPU306への信号を包含する。図示された実施例で、RPU306はトリガイベント546を感知して反応するように構成されたポーリング回路542を包含する。また、斯かる実施例で、CPU302はデータ転送のタイミングを制御する。多様な実施例で、ポーリング回路542はRPUと関連されたドライバ又はファームウェアと共に作業する。上述されたのは、単に説明のための実施例であり、本発明はこれに限定されない。

10

【0067】

このような実施例において、トリガイベント546が発生した場合、RPU306はデータ504をRPUバッファ516からHBM STACK308へ複写する。以後、RPU306はそれ自身に割り当てられたタスク、サブタスク又はタスク590の一部を実行する。前に説明されたように、RPU306に割り当てられたタスク590の一部はHBM STACK308のHBMダイのプロセッシングインメモリの機能の制御を包含する。RPU306がタスク590の第2部分を実行する間ち、FPU304は中断又は停止するか又は非同期的に他のタスクを遂行してFPU304の演算電力を浪費しないようにする。上述されたのは、単に説明のための幾つかの実施例であり、本発明がこれに限定されるものではない。

20

【0068】

一部の実施例において、タスク実行は類似した方式で、しかし逆向きの手順で、RPU306からFPU304(又はCPU302)に変更される。斯かる実施例で、FPU304、RPU306、及びCPU302間のデータ同期化は、前に説明されたところと類似に、現出できる。

【0069】

図6は、本発明によるシステム600の例示的な実施例のブロック図である。多様な実施例で、システム500の部分は、図1及び/又は図2に図示されたダイ又は他のICで構成される。システム500は図3に簡略に図示されたタスク転送の一実施例を示す。

30

【0070】

図示された実施例において、システム600は、上述されたように、固定処理ユニット304(例えば、GPU、DSPなど)を包含する。システム600は、上述されたように、RPU306(例えば、FPGAなど)を包含する。システム600は、上述されたように、HBM STACK308をも包含する。

【0071】

図示された実施例において、システム600は、データを格納するように構成され、FPU304により又はその制御下に直接的にアクセスされるFPUメモリ414を包含する。システム600は、上述されたように、システムメモリ412を包含する。

【0072】

図示された実施例において、CPU302は連結バス(interconnector bus)402を介しFPU304及びRPU306と連結される。多様な実施例で、連結バス402は、PCIe(Peripheral Component Interconnect express)プロトコルのような直列プロトコルを使用できる。上述されたのは、単に説明のための実施例であり、本発明がこれに限定されない。

40

【0073】

上述されたように、CPU(又はタスクスケジューラ)302は多様なタスク690又はタスク690の或る部分を、実行のため、FPU304及びRPU306に分散するか又は割り当てる。多様な実施例で、斯かる割当は、少なくとも部分的に、どのプロセッサが利用可能であり、且つ、より効率的にタスク690を実行できるか、という点に基づく。上述されたように、CPU302はFPU304にタスク590の第1部分を割り当てる。

50

、タスク 590 の第 2 部分を RPU306 に割り当てられる。例えば、RPU306 に割り当てられたタスク 590 の部分は、HBM スタック 308 の HBM ダイのプロセッシングインメモリの機能の制御を包含できる。上述されたのは、単に説明のための幾つかの実施例であり、本発明はこれに限定されない。

【0074】

図示された実施例において、FPU304 は、上述されたように、割り当てられたタスクを実行する。斯かる実施例で、FPU304 は一時記憶キャッシュ、即ち、作業用メモリとして FPU メモリ 414 を使用する。割り当てられたタスク、サブタスク又はタスク 690 の一部が完了された場合、FPU304 はデータ 604 をシステムメモリ 412 に書き込む。

10

【0075】

図示された実施例において、FPU304 は RDMA (Remote Direct Memory Access) を通じてシステムメモリ 412 にアクセスする。多様な実施例で、これは、CPU302 又は少なくとも CPU302 の積極的な関与を本質的にバイパス (bypass) する。矢印 614 は FPU304 が、直接 (又は CPU302 の DMA エンジンを使用して) データをシステムメモリ 412 に書き込むアクセスを示す。

【0076】

システムメモリ 412 へのデータの書き込みが完了された場合、データが使用のため準備されたことを RPU306 に報らせるために、トリガイベント 646 (例えば、インタープトなど) が開始される。多様な実施例で、トリガイベント 646 は、上述されたように、CPU302 により生成される。他の実施例で、トリガイベント 646 は FPU304 により生成される。

20

【0077】

図示された実施例において、RPU306 は、トリガイベント 646 が発生した時を感知するポーリング回路 642 を包含する。このような実施例で、以後、RPU306 はデータ 604 をシステムメモリ 412 から HBM スタック 308 へ複写する。(矢印 616 により図示される) 多様な実施例で、これは RDMA を介して行われることができる。斯かる実施例で、RPU306 は、図 4 に図示されたように、RDMA エンジンを包含できる。

【0078】

以後、306 はそれ自身に割り当てられたタスク、サブタスク又はタスク 690 の一部を実行できる。タスク、サブタスク又はタスク 690 の一部は HBM スタック 308 の HBM ダイのプロセッシングインメモリの機能の制御を包含する。RPU306 がタスク 690 の第 2 部分を実行する間、FPU304 は停止か中断か、又は非同期的に他のタスクを遂行して、FPU304 の演算電力の浪費を回避できる。上述されたのは、単に説明のための幾つかの実施例であり、本発明はこれに限定されない。

30

【0079】

一部の実施例において、タスク実行は類似した方式で、しかし逆向きの手順で、RPU306 から FPU304 (又は CPU302 に) に変更される。斯かる実施例で、FPU304、RPU306 及び CPU302 間のデータ同期化は、前に説明されたところと類似に現出できる。

40

【0080】

図 7 は、本発明の理論によって形成された半導体装置を包含できる情報処理システム 700 の例示的なブロック図である。

【0081】

図 7 を参照すると、情報処理システム 700 は、本発明の理論によって構成された 1 つ又はそれ以上の装置を包含する。他の実施例で、情報処理システム 700 は本発明の理論による 1 つ又はそれ以上の技法を使用するか又は実行できる。

【0082】

多様な実施例において、情報処理システム 700 は、例えば、ラップトップ (laptop)、デスクトップ (desktop)、ワークステーション (workstation)

50

)、サーバー、ブレードサーバー(blade server)、個人用デジタル端末機(PDA: Personal Digital Assistant)、スマートフォン、タブレット(tablet)及び他の適切なコンピュータのようなコンピューティング装置、又はそれらの仮想マシン又は仮想コンピューティング装置を包含する。多様な実施例で、情報処理システム700はユーザ(図示せず)により使用され得る。

【0083】

本発明による情報処理システム700は、中央処理ユニット(CPU)、ロジック、又はプロセッサ(以下、プロセッサで代表)710を包含できる。一部の実施例で、プロセッサ710は1つ又はそれ以上の機能ユニットブロック(FUB: Functional Unit Block)又は組合せ論理ブロック(CLB: Combinational Logic Block)715を包含する。このような実施例で、組合せ論理ブロックは、多様なブール論理動作(Boolean Logic Operation)(例えば、NAND、NOR、NOT、XORなどの演算)、双安定ロジック装置(例えば、フリップフロップ、ラッチ)、他のロジック装置、又はそれらの組合せを包含する。斯かる組合せロジック動作は、入力信号を処理して意図された結果を達成するように単純な方式又は複雑な方式で構成されることができる。同期組合せロジック動作の幾つかの例示的な実施例が説明されたが、本発明はこれに限定されず、非同期動作又はそれらの組合せを包含する。一実施例で、組合せロジック動作は複数の相補型金属酸化半導体(CMOS: Complementary Metal Oxide Semiconductor)トランジスタを包含する。多様な実施例で、斯かるCMOSトランジスタは、論理的動作を遂行するゲートを構成するように配列される。他の技術が使用できるが、これは本発明の範囲内に含まれる。

【0084】

本発明による情報処理システム700は、揮発性メモリ720(例えば、RAM(Random Access Memory)をさらに包含する。本発明による情報処理システム700は、不揮発性メモリ730(例えば、ハードドライブ、光学ドライブ、 NAND又はフラッシュメモリ)をさらに包含する。幾つかの実施例で、揮発性メモリ720、不揮発性メモリ730、或いはそれらの部分又は組合せの中で何れか1つは、「格納媒体(storage medium)」と称される。多様な実施例で、揮発性メモリ720及び/又は不揮発性メモリ730は、半永久的、即ち、実質的に永久的な形態でデータを格納するように構成される。

【0085】

多様な実施例において、情報処理システム700は、1つ又はそれ以上のネットワークインターフェース740を包含する。1つ又はそれ以上のネットワークインターフェース740は、情報処理システム700が通信ネットワークの一部分となって通信ネットワークを介して通信するように構成される。

ワイファイ(Wi-Fi)プロトコルの例は、IEEE(Institute of Electrical and Electronics Engineers)802.11g、IEEE802.11nを包含するが、これに限定されない。

セルラー(cellular)プロトコルの例は、IEEE802.16m(別名: 無線-MAN(Metropolitan Area Network) Advanced)、LTE(Long Term Evolution) Advanced、EDGE(Enhanced Data rates for GSM(登録商標)(Global System for Mobile Communications) Evolution)、HSPA+(Evolved High-Speed Packet Access)を包含するが、これに限定されない。有線プロトコルの例は、IEEE 802.3(別名: イーサネット(登録商標)(Ethernet))、ファイバチャンネル(Fiber Channel)、電力線通信(例えば、ホームプラグ(HomePlug)、IEEE1901)を包含するが、これに限定されない。

【0086】

10

20

30

40

50

本発明による情報処理システム 700 はユーザインタフェースユニット 750 (例えば、ディスプレイアダプタ (display adapter) 、ハaptiックインターフェイス (haptic interface) 、ヒューマンインターフェース (human interface) 装置) をさらに包含する。多様な実施例で、このようなユーザインタフェースユニット 750 は、ユーザからの入力を受信するように、及び / 又はユーザに出力を提供するように、構成される。なお、他の種類の装置もまた、ユーザとの相互作用を提供するために使用できる。例えば、ユーザに提供されたフィードバックは何らかの形の感覚フィードバック、例えば、視覚フィードバック、聴覚フィードバック、又は触覚フィードバックであり、ユーザからの入力は音、音声、又は触覚入力を含む、何らかの形で受信される。

10

【 0087 】

多様な実施例において、情報処理システム 700 は 1 つ又はそれ以上の他の装置又はハードウェア構成部品 760 (例えば、ディスプレイ、モニタ、キーボード、マウス、カメラ、指紋認識機又はビデオプロセッサ) を包含する。上述されたのは、単に説明のための幾つかの実施例であり、本発明はこれに限定されない。

【 0088 】

本発明による情報処理システム 700 は 1 つ又はそれ以上のシステムバス 705 をさらに包含する。斯かる実施例で、システムバス 705 は、プロセッサ 710 、揮発性メモリ 720 、不揮発性メモリ 730 、ネットワークインターフェース 740 、ユーザインタフェースユニット 750 及び 1 つ又はそれ以上のハードウェア構成部品 760 を通信的に連結するように構成される。プロセッサ 710 により処理されたデータ又は不揮発性メモリ 730 の外部から入力されたデータは、不揮発性メモリ 730 又は揮発性メモリ 720 の中の何れか 1 つに格納される。

20

【 0089 】

多様な実施例において、情報処理システム 700 は 1 つ又はそれ以上のソフトウェア構成部品 770 を包含するか又は実行できる。幾つかの実施例で、ソフトウェア構成部品 770 は、OS (Operating System 、オペレーティングシステム) 及び / 又はアプリケーションを包含できる。幾つかの実施例で、OS は 1 つ又はそれ以上のサービスをアプリケーションに対して提供し、アプリケーションと、情報処理システム 700 の多様なハードウェア構成部品 (例えば、プロセッサ 710 、ネットワークインターフェース 740) との間で媒介体として世話役となって動作する。

30

斯かる実施例で、情報処理システム 700 は局部的に (例えば、不揮発性メモリ 730 内に) 設置されてプロセッサ 710 により直接実行されるか又は OS と直接的に相互作用するように構成される 1 つ又はそれ以上の基本アプリケーション (native applications) を包含する。斯かる実施例で、基本アプリケーションは、予めコンパイルされたマシン実行可能なコード (pre - compiled machine executable code) を包含する。幾つかの実施例で、基本アプリケーションは、ソースコード又はオブジェクトコードをプロセッサ 710 により実行される、プロセッサ 710 が実行可能なコードに変換するように構成されたスクリプト解釈機 (例えば、csh (C - shell) (登録商標) 、AppleScript (登録商標) 、AutoHotkey (登録商標)) 又は仮想実行マシン (VM : Virtual execution Machine) (例えば、ジャバ仮想マシン (JAVA_Virtual_Machine) (登録商標) 、マイクロソフト共通言語ランタイム (Microsoft_Common_Language_Runtime) (登録商標)) を包含する。

40

【 0090 】

上述された半導体装置は、多様なパッケージング技法を使用して構成されうる。例えば、本発明の理論により構成された半導体装置は、POP (package on package) 技法、BGA (ball grid arrays) 技法、CSPs (chips scale packages) 技法、PLCC (plastic leaded chip carrier) 技法、PDIP (plastic dual in - line pack

50

age) 技法、die_in_waffle pack 技法、die_in_wafer form 技法、COB(chip on board) 技法、CERDIP(ceramic dual in-line package) 技法、PMQFP(plastic metirc quad flat package) 技法、PQFP(plastic quad flat package) 技法、SOIC(small outline IC package) 技法、SSOP(shrink small outline package) 技法、TSOP(thin small outline package) 技法、TQFP(thin quad flat package) 技法、SIP(single in-line package) 技法、MCP(multi-chip package) 技法、WFP(wafer-level fabricated package) 技法、WSP(wafer-level processed stack package) 技法、又は、当業者によく知られた他の技法の中で何れか1つを使用して構成される。
10

【0091】

方法のステップは、コンピュータプログラムを実行する1つ又はそれ以上のプログラム可能なプロセッサにより遂行されて、入力データに対し作業して出力を生成することにより機能を遂行する。方法のステップは特定目的向け論理回路(例えば、FPGA(field programmable gate array)又はASIC(application-specific integrated circuit))により遂行され、装置は特定目的向け論理回路として具現される。

【0092】

多様な実施例において、コンピュータ読み取り可能な媒体は、命令(instruction)を包含する。命令が実行される時、装置は方法のステップの少なくとも一部を遂行する。幾つかの実施例で、コンピュータ読み取り可能な装置は、磁気媒体、光学媒体、他の媒体又はそれらの組合せ(例えば、CD-ROM、ハードドライブ、リード(read)専用メモリ、フラッシュドライブ)に含まれる。斯かる実施例で、コンピュータ読み取り可能な媒体は有形であり、非一時的に具現された製品であり得る。

【0093】

本発明の理論が例示的な実施例を参照して説明されたが、多様な変化及び変更が本発明の思想及び範囲からの逸脱なしに行われることは、当業者には明らかであろう。よって、上述された実施例は発明の範囲を何ら限定するものではなく、単に例示するものとして理解されなければならない。即ち、本発明の範囲は、以下の特許請求の範囲及びその均等物の最も広く許容される解釈により決定されるものであり、前の説明により制限されたり、限定されたりしてはならない。従って、添付された特許請求の範囲は、実施例の範囲内に含まれる全ての変更及び変化を包含するように意図されたものとして理解されるべきである。

【産業上の利用可能性】

【0094】

本発明は、向上された高効率ラーニングシステムを有する異種アクセラレータに有用である。

【符号の説明】

【0095】

100、200、300, 400、500、600 システム
 104 固定処理ユニット、FPU、FPU/GPU
 106 再プログラム可能な処理ユニット、RPU
 108、308 スタック
 110 HBMコントローラ
 112 HBMダイ
 120 インタポーザ、インタポーザダイ、インタポーザレイヤ
 122 基板ダイ、パッケージ基板ダイ
 124 はんだバンプ

10

20

30

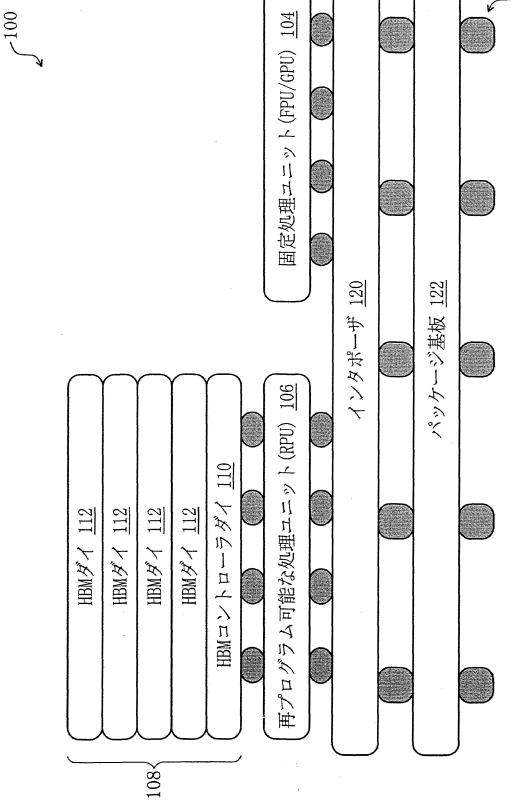
40

50

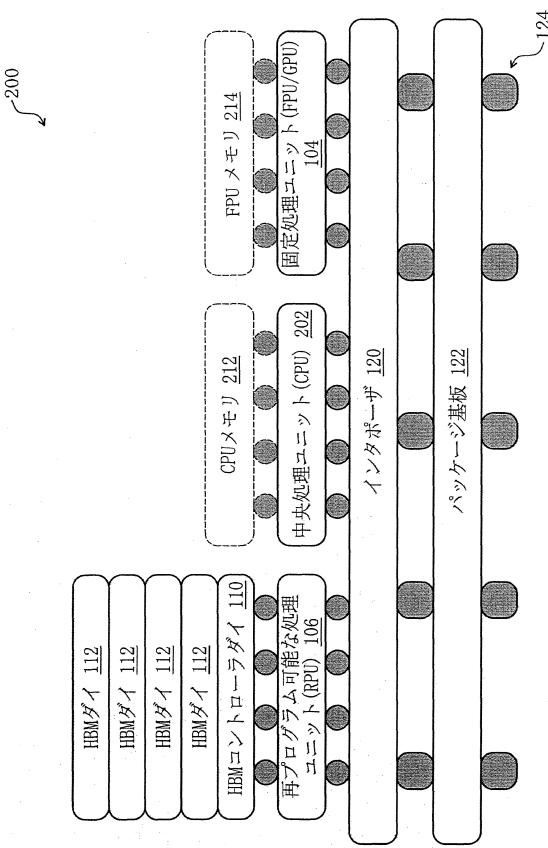
2 0 2	中央処理ユニット、C P U	
2 1 2	C P Uメモリ、C P Uメモリダイ	
2 1 4	F P Uメモリ、F P Uメモリダイ	
3 0 2	タスクスケジューラ、C P U、(ホスト)プロセッサ	
3 0 4	F P U、固定処理ユニット	
3 0 6	R P U、再プログラム可能な処理ユニット	
4 0 2	連結バス	
4 0 4、5 0 4、6 0 4	データ	
4 1 2	システムメモリ	10
4 1 4	F P Uメモリ	
4 4 2	ポーリング回路	
4 4 4	アクセスレジスタ	
4 6 2	レディーフラッグ	
4 6 4	R D M Aエンジン又はR D M A回路	
4 9 0、5 9 0、6 9 0	タスク	
5 1 4	F P Uバッファ、中央メモリ	
5 1 6	R P Uバッファ、中央メモリ	
5 4 2、6 4 2	ポーリング回路	
5 4 6、6 4 6	トリガイベント	
6 1 4、6 1 6	データ(6 0 4)の転送経路	20
7 0 0	情報処理システム	
7 0 5	システムバス	
7 1 0	中央処理ユニット(C P U)、ロジック、又はプロセッサ、プロセッサ	
7 1 5	C L B	
7 2 0	揮発性メモリ	
7 3 0	不揮発性メモリ	
7 4 0	ネットワークインターフェース	
7 5 0	ユーザインターフェースユニット	
7 6 0	ハードウェア構成部品	
7 7 0	ソフトウェア構成部品	30
C L B	組合せ論理ブロック	
C P U	中央処理ユニット	
F P U	変更不能な、即ち固定処理ユニット	
F U B	機能ユニットブロック	
G P U	グラフィックス処理ユニット	
P P U	プログラム可能な処理ユニット	
R P U	再プログラム可能な処理ユニット	

【図面】

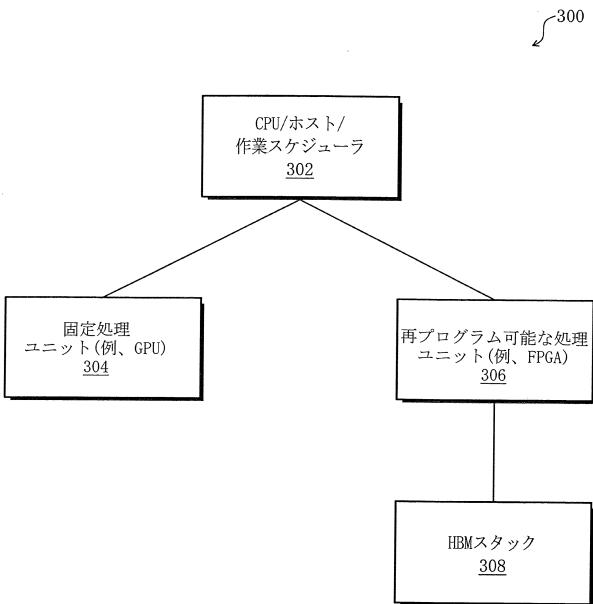
【図 1】



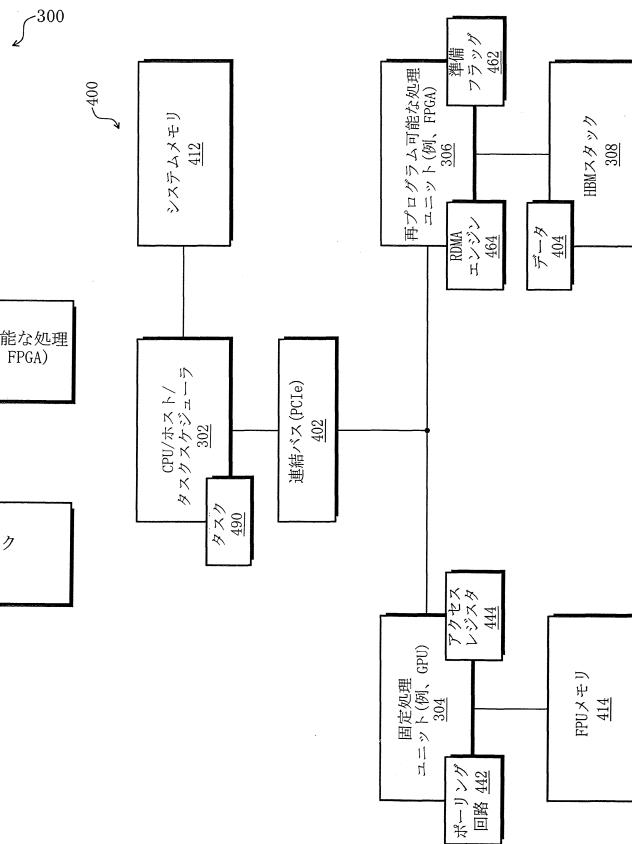
【图2】



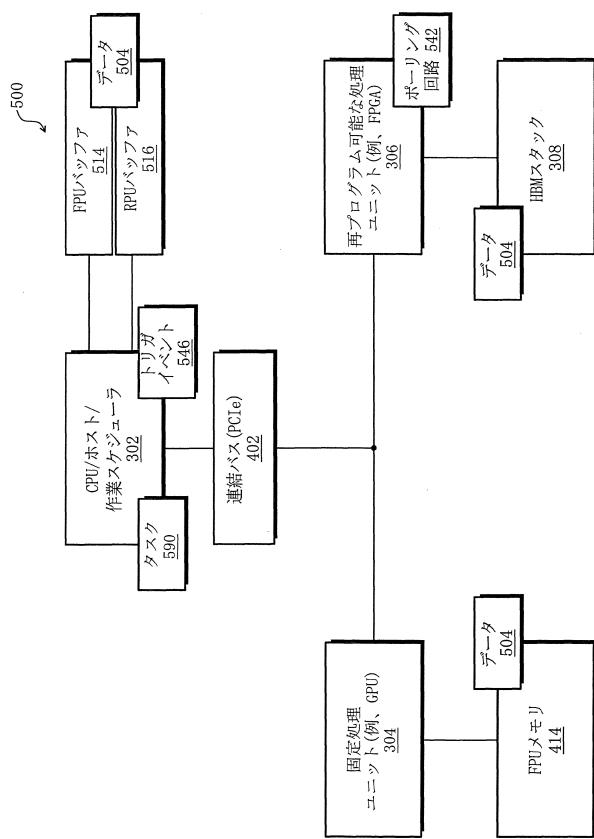
【 义 3 】



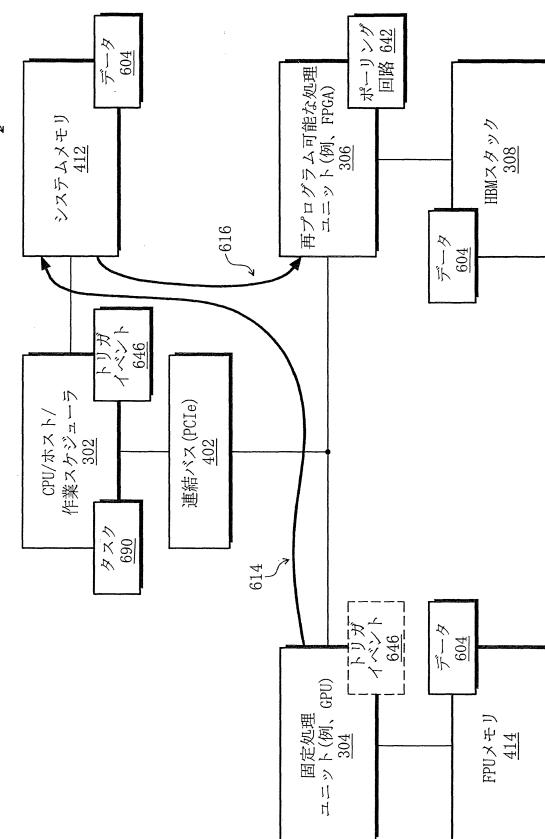
【 四 4 】



【図 5】



【図 6】

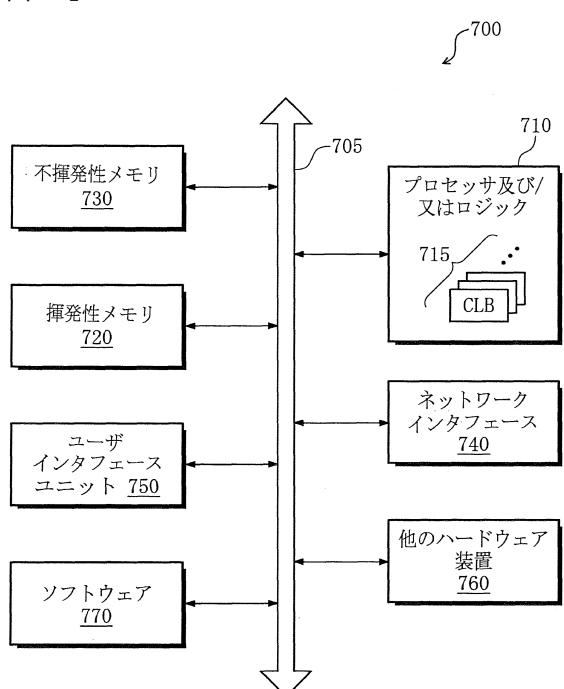


10

20

30

【図 7】



40

50

フロントページの続き

ライブ 4196

(72)発明者 ゼング, ホング ゾング

アメリカ合衆国, カリフォルニア州 95032, ロス ガトス, #6, カールトン アヴェニュー
120

審査官 井上 宏一

(56)参考文献 特開2003-347470 (JP, A)

特開2010-080802 (JP, A)

特表2015-533009 (JP, A)

国際公開第2016/209406 (WO, A1)

(58)調査した分野 (Int.Cl., DB名)

G06F 9/38

G06F 9/50

G06F 12/00

G06F 12/02

G06F 15/78