

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6941082号
(P6941082)

(45) 発行日 令和3年9月29日(2021.9.29)

(24) 登録日 令和3年9月7日(2021.9.7)

(51) Int.Cl.

F 1

G 06 F	12/00	(2006.01)	G 06 F	12/00	5 6 0 F
G 06 F	12/06	(2006.01)	G 06 F	12/00	5 6 4 A
G 11 C	5/04	(2006.01)	G 06 F	12/00	5 9 7 R
G 11 C	5/06	(2006.01)	G 06 F	12/06	5 5 0 A
			G 11 C	5/04	2 2 0

請求項の数 20 (全 23 頁) 最終頁に続く

(21) 出願番号 特願2018-152980 (P2018-152980)
 (22) 出願日 平成30年8月15日 (2018.8.15)
 (65) 公開番号 特開2019-53725 (P2019-53725A)
 (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14日 (2021.7.14)
 (31) 優先権主張番号 62/558, 726
 (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 62/558, 732
 (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)

(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 129
129, Samsung-ro, Yeon
g t o n g - g u, Suwon-si, G
yeonggi-do, Republic
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 ティー
アメリカ合衆国, 95135, カリフォルニア州, サンノゼ, ロートレック
ドライブ 4196
最終頁に続く

(54) 【発明の名称】高帯域メモリシステム

(57) 【特許請求の範囲】

【請求項 1】

高帯域メモリ(HBM)システムであって、
PIM(Processing-In-Memory)動作をサービスするPIM機能
を含むHBMメモリ装置と、

命令語及びアドレス(CA)バス並びにデータバスを含みホスト装置に連結される第1
インターフェースと、内部バスを介して前記HBMメモリ装置に連結される第2インターフェースとを含み、前記第1インターフェースの前記CAバスを介して前記ホスト装置から第1命令語と、前記データバスを介して前記ホスト装置から第2命令語とを受信し、前記受信された第2命令語を、前記第2インターフェースを介して前記HBMメモリ装置に
伝送される、前記PIM動作のためのPIM命令語に変換し、前記PIM動作のための内部
バス命令語が前記PIM命令語の実行に重なるように、前記内部バス上で前記PIM動作
のためのオーバーラップタイミングを有する少なくとも1つの命令語を発行する論理回路と、を備え、

前記HBMシステムが前記ホスト装置から前記第1命令語を受信して前記ホスト装置から他の命令語を受信する準備ができるまでの遅延(latency)は、決定論的(deterministic)であることを特徴とする高帯域メモリシステム。

【請求項 2】

前記第2命令語は、前記HBMメモリ装置内の単一のアドレスに対するPIM動作又は前記HBMメモリ装置内の同一の行の複数のアドレスに対するPIM動作のためのもので

あることを特徴とする請求項 1 に記載の高帯域メモリシステム。

【請求項 3】

前記ホスト装置から受信された第 2 命令語は、前記 H B M メモリ装置内の同一のチャンネル内の一つ以上のバンク内の P I M 動作のためのものであり、

前記第 2 命令語が前記ホスト装置から受信されて前記 H B M システムが前記ホスト装置から他の命令語を受信する準備ができるまでの遅延は、決定論的であることを特徴とする請求項 2 に記載の高帯域メモリシステム。

【請求項 4】

前記ホスト装置から受信された第 2 命令語は、前記 H B M メモリ装置内の異なるバンクに亘る P I M 動作のためのものであり、

前記第 2 命令語が前記ホスト装置から受信されて前記 H B M システムが前記ホスト装置から他の命令語を受信する準備ができるまでの遅延は、決定論的であることを特徴とする請求項 1 に記載の高帯域メモリシステム。

【請求項 5】

前記論理回路は、前記 C A バスを介して前記ホスト装置から前記 H B M メモリ装置内における P I M 動作のための第 3 命令語と、前記データバスを介して前記ホスト装置から前記第 3 命令語に続く第 4 命令語とを更に受信し、

前記第 4 命令語は、前記第 4 命令語が前記ホスト装置から受信されて前記 H B M システムが前記ホスト装置から他の命令語を受信する準備ができるまでの遅延に関する、決定論的部分及び推定部分を含む第 1 推定情報を要求することを特徴とする請求項 1 に記載の高帯域メモリシステム。

【請求項 6】

前記論理回路は、前記 C A バスを介して前記ホスト装置から前記 H B M メモリ装置内における P I M 動作のための第 3 命令語と、前記データバスを介して前記ホスト装置から前記第 3 命令語に続く第 4 命令語とを更に受信し、

前記第 4 命令語は、前記第 4 命令語が前記ホスト装置から受信されて前記 H B M システムが前記ホスト装置から他の命令語を受信する準備ができるまでの遅延に関する、決定論的部分及びクレジット基盤部分を含む推定情報を要求することを特徴とする請求項 1 に記載の高帯域メモリシステム。

【請求項 7】

前記論理回路は、前記 C A バスを介して前記ホスト装置から前記 H B M メモリ装置内における P I M 動作のための第 3 命令語と、前記データバスを介して前記ホスト装置から前記第 3 命令語に続く第 4 命令語とを更に受信し、

前記第 4 命令語は、前記第 4 命令語が前記ホスト装置から受信されて前記 H B M システムが前記ホスト装置から他の命令語を受信する準備ができるまでの遅延に関する、決定論的部分及び再試行部分を含む推定情報を要求することを特徴とする請求項 1 に記載の高帯域メモリシステム。

【請求項 8】

前記第 2 命令語は、前記 H B M メモリ装置内の同一のチャンネル内の一つ以上のバンク内の P I M 動作のためのものであることを特徴とする請求項 1 に記載の高帯域メモリシステム。

【請求項 9】

前記第 2 命令語は、前記 H B M メモリ装置内の異なるバンクに亘る P I M 動作のためのものであることを特徴とする請求項 1 に記載の高帯域メモリシステム。

【請求項 10】

並列タイミング経路を有する命令語は、リード命令語及びライト命令語を含むことを特徴とする請求項 1 に記載の高帯域メモリシステム。

【請求項 11】

高帯域メモリ (H B M) システムであって、

高帯域メモリ (H B M) 装置と、

10

20

30

40

50

ホスト装置に連結される第1インターフェースと、前記HBM装置に連結される第2インターフェースとを含み、前記ホスト装置から一つ以上の命令語を受信し、前記受信された各命令語を、前記第2インターフェースを介して前記HBM装置に伝送される少なくとも一つの対応するPIIM(Processing-In-Memory)命令語に変換する論理回路と、を備え、

前記論理回路は、前記ホスト装置から前記HBM装置内におけるPIIM動作のための第1命令語と、前記ホスト装置から前記第1命令語に続く第2命令語とを更に受信し、

前記第2命令語は、前記第2命令語が前記ホスト装置から受信された時点と前記HBMシステムが前記ホスト装置から他の命令語を受信する準備ができた時点との間にに関する、決定論的期間及び非決定論的期間を含む第1時間推定情報を要求することを特徴とする高帯域メモリシステム。 10

【請求項12】

前記非決定論的期間は、推定期間を含むことを特徴とする請求項11に記載の高帯域メモリシステム。

【請求項13】

前記論理回路は、前記ホスト装置から前記HBM装置内におけるPIIM動作のための第3命令語と、前記ホスト装置から前記第3命令語に続く第4命令語とを更に受信し、

前記第4命令語は、前記第3命令語が前記ホスト装置から受信された時点と前記HBMシステムが前記ホスト装置から他の命令語を受信する準備ができた時点との間にに関する、決定論的期間及びクレジット基盤期間を含む第2時間推定情報を要求することを特徴とする請求項11に記載の高帯域メモリシステム。 20

【請求項14】

前記論理回路は、前記ホスト装置から前記HBM装置内におけるPIIM動作のための第5命令語と、前記ホスト装置から前記第5命令語に続く第6命令語とを更に受信し、

前記第6命令語は、前記第6命令語が前記ホスト装置から受信された時点と前記HBMシステムが前記ホスト装置から他の命令語を受信する準備ができた時点との間にに関する、決定論的期間及び再試行期間を含む第3時間推定情報を要求することを特徴とする請求項13に記載の高帯域メモリシステム。

【請求項15】

前記論理回路は、前記ホスト装置から前記HBM装置内におけるPIIM動作のための第7命令語を更に受信し、 30

前記第7命令語が前記ホスト装置から受信された時点と前記HBMシステムが前記ホスト装置から他の命令語を受信する準備ができた時点との間の時間は、決定論的であることを特徴とする請求項14に記載の高帯域メモリシステム。

【請求項16】

前記第7命令語は、前記HBM装置内の単一のアドレスに対するPIIM動作又は前記HBM装置内の同一の行の複数のアドレスに対するPIIM動作のためのものであることを特徴とする請求項15に記載の高帯域メモリシステム。

【請求項17】

前記論理回路は、前記ホスト装置から前記HBM装置内の同一のチャンネル内の一つ以上のバンク内のPIIM動作のための第8命令語を更に受信し、 40

前記第8命令語が前記ホスト装置から受信された時点と前記HBMシステムが前記ホスト装置から他の命令語を受信する準備ができた時点との間の時間は、決定論的であることを特徴とする請求項16に記載の高帯域メモリシステム。

【請求項18】

前記論理回路は、前記ホスト装置から前記HBM装置内の異なるバンクに亘るPIIM動作のための第9命令語を更に受信し、

前記第9命令語が前記ホスト装置から受信された時点と前記HBMシステムが前記ホスト装置から他の命令語を受信する準備ができた時点との間の時間は、決定論的であることを特徴とする請求項17に記載の高帯域メモリシステム。 50

【請求項 19】

前記非決定論的期間は、クレジット基盤期間を含むことを特徴とする請求項11に記載の高帯域メモリシステム。

【請求項 20】

前記非決定論的期間は、再試行期間を含むことを特徴とする請求項11に記載の高帯域メモリシステム。

【発明の詳細な説明】**【技術分野】****【0001】**

10

本発明は、高帯域メモリ（HBM）システムに関し、より詳細には、ホスト装置とインターフェースするためのメモリ動作において、決定論的及び非決定論的処理を含む高帯域メモリシステムに関する。

【背景技術】**【0002】**

ディープ神経網（Deep Neural Net）のような新しい応用プログラムは、膨大な計算及びメモリ機能を使用して他のデータセットを学習し、高い正確度で学習する。また、高性能コンピューティング（HPC）、グラフィックアルゴリズムなどのような応用プログラムがデータ及び演算集約的に変わることによって、エネルギーの効率性と低遅延（latency）が重要な特性となる。プロセッシングインメモリ（Processing In Memory : PIM）は、DRAMロジックダイ（logic dies）の複雑な作業をスケジューリングして低電力技術プロセスで追加コンピューティング機能を提供し、データのあるところにより近くPIMを配置することによって、このような問題を解決する。

20

【先行技術文献】**【特許文献】****【0003】**

30

【特許文献1】米国特許第6622235号明細書

【特許文献2】米国特許第8321627号明細書

【特許文献3】米国特許第9201777号明細書

【特許文献4】米国特許第9633748号明細書

【特許文献5】米国特許出願公開第2008/0162855号明細書

【特許文献6】米国特許出願公開第2016/0210174号明細書

【特許文献7】米国特許出願公開第2016/0379686号明細書

【特許文献8】米国特許出願公開第2017/0220499号明細書

【非特許文献】**【0004】**

【非特許文献1】PATTNAIK, Ashutosh et al., 「Scheduling Techniques for GPU Architectures with Processing-In-Memory Capabilities」, PACT '16, September 11-15, 2016, Haifa, Israel

40

【発明の概要】**【発明が解決しようとする課題】****【0005】**

本発明は、上記従来の問題点に鑑みてなされたものであって、本発明の目的は、ホスト装置とインターフェースするためのメモリ動作において、決定論的及び非決定論的処理を含む高帯域メモリシステムを提供することにある。

【課題を解決するための手段】**【0006】**

50

上記目的を達成するためになされた本発明の一態様による高帯域メモリ(H B M)システムは、高帯域メモリ(H B M)装置と、ホスト装置に連結された第1インターフェース及び前記 H B M 装置に連結された第2インターフェースを含み、前記第1インターフェースを介して前記ホスト装置から第1命令語を受信し、前記受信された第1命令語を、前記第2インターフェースを介して前記 H B M 装置に伝送される第1P I M (P r o c e s s i n g - I n - M e m o r y) 命令語に変換する論理回路と、を備え、前記第1命令語が前記ホスト装置から受信された時点と前記 H B M システムが前記ホスト装置から他の命令語を受信する準備ができた時点との間の時間は、決定論的である。

【 0 0 0 7 】

前記第1P I M 命令語は、完了のための決定論的遅延を有し得る。前記論理回路は、前記第1インターフェースを介して前記ホスト装置から第2命令語を受信し、前記受信された第2命令語を、前記第2インターフェースを介して前記 H B M 装置に伝送される第2P I M 命令語に変換し、前記第2P I M 命令語は、完了のための非決定論的遅延を有する。前記論理回路は、前記ホスト装置から受信された前記第1命令語に応答して前記 H B M 装置を制御し、前記 H B M 装置のチャンネルから少なくとも一つの選択されたバンクをプリチャージし得る。前記第2命令語が前記ホスト装置から前記論理回路によって受信された時点と前記 H B M システムが前記ホスト装置から他の命令語を受信する準備ができた時点との間の時間は、決定論的である。

【 0 0 0 8 】

上記目的を達成するためになされた本発明の他の態様による高帯域メモリ(H B M)システムは、高帯域メモリ(H B M)装置と、ホスト装置に連結された命令語 / アドレスバス及びデータバスを含む第1インターフェース、前記 H B M 装置に連結された第2インターフェース、及び前記ホスト装置に連結されたトランザクションバス(t r a n s a c t i o n b u s)を含み、前記第1インターフェースを介して前記ホスト装置から一つ以上の命令語を受信し、受信されたそれぞれの命令語を、前記第2インターフェースを介して前記 H B M 装置に伝送される少なくとも一つの対応するP I M (P r o c e s s i n g - I n - M e m o r y) 命令語に変換する論理回路と、を備え、前記論理回路は、前記ホスト装置から前記 H B M 装置内におけるP I M 動作のための第1命令語を受信し、前記受信された第1命令語を前記 H B M 装置に伝送される第1P I M 命令語に変換し、前記ホスト装置から前記第1命令語に続く第2命令語を更に受信し、前記第2命令語は、前記第2命令語が前記ホストから受信された時点と前記 H B M システムが前記ホスト装置から他の命令語を受信する準備ができた時点との間の時間に関する、決定論的期間及び非決定論的期間を含む時間推定情報を要求する。

【 0 0 0 9 】

前記第1P I M 命令語は、完了のための非決定論的遅延を有し得る。前記論理回路は、前記第1P I M 命令語の変換が完了した場合、前記トランザクションバスを介して前記ホスト装置に指示を伝送し得る。前記第1命令語に対応する第1命令語パケットは、前記ホスト装置からデータバスを介して前記論理回路で受信され、前記第1命令語が前記ホスト装置から受信された時点と前記 H B M システムが前記ホスト装置から他の命令語を受信する準備ができた時点との間の時間は、非決定論的であり得る。

【 発明の効果 】

【 0 0 1 0 】

本発明によると、ホスト装置とインターフェースするためのメモリ動作において、決定論的及び非決定論的処理を可能にする高帯域メモリシステム及びそのプロトコルを具現することができる。

【 図面の簡単な説明 】

【 0 0 1 1 】

【 図 1 】一実施形態による H B M + システムの一例の一部の断面図である。

【 図 2 】一実施形態による 1 段階 H B M + プロトコル命令語の一例のタイミング図である。

【図3】一実施形態によるPIM_CMD命令語のコンテンツの一例の配置を示す図である。

【図4】一実施形態による2段階HBM+プロトコル命令語の一例の概略的タイミング図である。

【図5】一実施形態によるPIM動作がHBM+装置内の单一のアドレスに向けられるか又は同一の行に向けられる1段階HBM+プロトコルPIM命令語の一例のタイミング図である。

【図6】一実施形態によるPIM動作が同一のチャンネルのバンク内で行われる1段階HBM+プロトコルPIM命令語の一例のタイミング図である。

【図7】一実施形態によるPIM動作が異なるチャンネルを介して行われる1段階HBM+プロトコルPIM命令語の一例のタイミング図である。 10

【図8】一実施形態によるPIM命令語が完了する時点でHBM+スタックが時間の推定値を提供する2段階HBM+プロトコルPIM命令語の一例のタイミング図である。

【図9】一実施形態によるPIM命令語が完了する時点でHBM+スタックがクレジット基盤指示を提供する2段階HBM+プロトコルPIM命令語の一例のタイミング図である。

【図10】一実施形態によるPIM命令語が完了する時点でHBM+スタックがフィードバック基盤指示を提供する2段階HBM+プロトコルPIM命令語の一例のタイミング図である。

【発明を実施するための形態】

20

【0012】

以下、本発明を実施するための形態の具体例を、図面を参照しながら詳細に説明する。発明の詳細な説明において、本発明の完全な理解を提供するために多数の特定の細部事項を説明する。しかし、当業者は開示された態様がこのような特定の細部事項無しで実施され得ることを理解するはずである。他の例において、公知された方法、手順、構成要素、及び回路は、ここに開示する主題を曖昧にしないために詳細に説明しない。

【0013】

本明細書の全体に亘って「一実施形態」又は「実施形態」は、本実施形態に関連して説明する特定の特徴、構造、又は特性が本明細書に開示する少なくとも一つの実施形態に含まれるということを意味する。従って、本発明の全体に亘って「一実施形態において」又は「実施形態において」又は「一実施形態によって」(又は類似の意味を有する他の文句)という表現は、全てが必ずしも同一の実施形態を示すものではない。なお、特定のフィーチャー(feature)、構造、又は特性は、一つ以上の実施形態において任意の適した方式に連結される。これに関連して、本明細書で使用するように、「例示的な」という単語は、「例示、実例、又は例示を提供する」ということを意味する。「例示的な」ものと本明細書に記載する任意の実施形態は、必ずしも他の実施形態よりも好ましく、且つ有利なものと解釈してはならない。更に、論議の文脈によって、単数は対応する複数の形態を含むことがあり、複数の用語は相応する単数の形態を含むことがある。本明細書に図示し、論議する多様な図面(構成要素も含む)は、単に例示的な目的のものであり、一定の比率で描かれるものではない。同様に、多様な波形及びタイミング図を、単に例示的な目的のために示す。例えば、一部の要素の寸法は、明確性のために他の要素に比べて誇張され得る。更に、適切なものとして考慮される場合、参照符号は、相応するか又は類似の要素を示すために図面の間で繰り返される。 30

【0014】

本明細書で使用する用語は、単に特定の実施形態を説明するためのものであり、特許請求の範囲を限定するものではない。本明細書で使用する単数の形態「a」、「a n」、及び「t h e」は、文脈上別に指示しない限り、複数の形態を含むものとする。本明細書で使用する「～含む」又は「含む～」という用語は、明示する特徴、整数、段階、動作、要素、構成要素、又は構成要素の存在を示すが、存在を排除しないということが理解されるはずであり、或いは一つ以上の他の特徴、整数、段階、動作、構成要素、又はグループの 40

50

追加を含む。本明細書で使用する「第一」、「第二」などの用語は、前に明示した名詞のラベルとして使用され、明示的に定義されない限り、全ての類型の順序（例えば、空間的、時間的、論理的など）を暗示しない。なお、同一且つ類似の機能を有する部品（parts）、コンポーネント（components）、ブロック、回路、ユニット、又はモジュールを示すために二つ以上の図面に亘って同一の参照番号を使用する。しかし、このような使用方法は、説明の簡素化及び論議の容易さのためだけに使用される。このような構成要素又はユニットの構成若しくは構造的細部事項が全ての実施形態に亘って同一であるということを意味しないか、或いは共通して参照する部品／モジュールが本明細書に開示する特徴の実施形態の教示を具現する唯一の方法であることを意味しない。

【0015】

10

別に定義しない限り、本明細書で使用する全ての用語（技術用語及び科学用語を含む）は、本発明の技術分野における当業者が一般的に理解するものと同一の意味を有する。なお、一般的に使用される辞書で定義された用語と同一の用語は、関連技術の脈絡における意味に一致する意味を有するものと解釈すべきであり、理想的又は過度に形式的な意味として解釈されない限り、明示的にここに定義される。

【0016】

本発明は、高帯域PIM（本明細書では、PIM（Processing - In - Memory）機能を支援する高帯域メモリ（HBM）を「HBM+」若しくは単に「HBM」と呼称する）システムのための準同期インターフェースプロトコルに関するものである。即ち、本発明は、HBM+システムとホスト装置との間のインターフェースとして、決定論的（deterministic）遅延（latency）及び非決定論的（non-deterministic）遅延の両方を含むのであり、従って、準同期式（quasi-synchronous）インターフェースプロトコルである。本明細書に開示する準同期プロトコルを用いる通信は、一つ以上のクロック信号に同期化されるが、PIM動作のような特定の動作に関連する特定の遅延は、完了のための決定論的遅延又は非決定論的遅延を有する。決定論的遅延を決定するPIM動作は1段階HBM+プロトコルを使用して遂行されるが、非決定論的遅延を提供するPIM動作は2段階HBM+プロトコルを使用して遂行される。

20

【0017】

30

図1は、一実施形態によるHBM+システム100の一例の一部の断面図である。HBM+システム100は、ホスト装置101、インターポーザー（interposer）102、及びHBM+スタック（stack）103を含む。HBM+スタック103は、ロジックダイ（die）（又は論理回路）104、及び複数のHBM+メモリ装置（105a～105d）を含む。ホスト装置101及びHBM+スタック103は、インターポーザー102の上部表面に固定される。

【0018】

30

一実施形態において、ホスト装置101は、中央処理装置（CPU）、グラフィック処理装置（GPU）、グラフィック加速器、又はFPGA（Field Programmable Gate Array）であるが、これに限定されるものではない。

【0019】

40

HBM+メモリ装置105は、チャンネル当たり16個のバンクを有する二つのチャンネルに分割される。また、一つ以上のHBM+メモリ装置（105a～105d）は、従来の判読及び記録動作のようなPIM機能及び規則的なデータ貯蔵機能を含む。HBM+メモリ装置に提供されるPIM機能は、ホスト装置101によってより効率的に提供される。ロジックダイ104は、HBM+メモリデバイス105内のPIM機能を制御するためのロジック機能を含む。HBM+内部バス111は、ロジックダイ104をHBM+メモリ装置105の各々に接続させる。HBM+内部バス111は、複数のアドレスライン、複数の命令語（コマンド）ライン、複数のデータライン、及び一つ以上の他の信号線を含む。4個のHBM+メモリ装置105をHBM+スタック103に示しているが、任意の数のHBM+メモリ装置105がHBM+スタック103を形成することはよく理解さ

50

れるはずである。HBM+システム100の一部のみを図1に示しているが、インターポーラー102の上部表面に取り付けられる追加的なHBM+スタック103が有り得ることを理解すべきである。

【0020】

インターポーラー102は、基板106の上部表面に固定される。基板106は、他の装置(図示せず)に電気接続を提供するために使用される端子107を含む。インターポーラー102は、ホスト装置101及びHBM+スタック103に対する構造的基盤を提供する以外に、ホスト装置101とHBM+スタック103との間の電気的連結を提供する。本実施形態において、インターポーラー102は、ホスト装置101とHBM+スタック103との間に電気的に連結される命令語(コマンド)/アドレス(CA)バス108、データバス109、及びトランザクションバス110を含む。インターポーラー102は、図示していない追加的な電気的連結を含む。

【0021】

CAバス108、データバス109、及びトランザクションバス110のそれぞれは、複数のライン及び/又はビットを含む。本実施形態において、トランザクションバス110は、トランザクション応答信号(RSP_R)及びトランザクションエラー信号(RSP_E)を含む。CAバス108、データバス109、及びトランザクションバス110は、HBM+システム100内において同期式方式で動作する。即ち、CAバス108、データバス109、及びトランザクションバス110は、一つ以上のクロック信号に同期して動作する。

10

【0022】

本実施形態において、HBM+システム100は、例えば命令語又はトランザクションの発行と応答出力との間のタイミングで、ホスト装置101とHBM+スタック103との間で命令語及びデータを通信するための1段階HBM+プロトコルを含み、正規のデータ貯蔵機能及び/又はPIM機能に基づくか、或いはHBM+スタック103が他の命令語又はトランザクションの準備をする時点は決定論的である。また、HBM+システム110は、例えば命令語又はトランザクションの発行と応答出力との間のタイミングで、ホスト装置101とHBM+スタック103との間で命令語及びデータを通信するための2段階HBM+プロトコルを含み、正規のデータ貯蔵機能及び/又はPIM機能に基づくか、或いはHBM+スタック103が他の命令語又はトランザクションの準備をする時点は非決定論的である。本明細書で使用する「準同期インターフェース(quasi-synchronous interface)」という用語は、決定論的通信及び非決定論的通信又はトランザクションの両方を経て通過するインターフェースを意味する。

20

【0023】

1段階HBM+プロトコルは、一般的にホスト装置101が結果を待たない比較的に規則的なデータ貯蔵及び簡単なPIM動作のためのものである。即ち、命令語の発行と応答出力との間の時間又は遅延は決定論的であるから、ホスト装置101は、命令語の発行と応答出力との間に他の動作を遂行する。代案として、ホスト装置101は、決定論的遅延の終端で、正規のトランザクションを再開するようにスケジューリングされる。一般的に簡単なPIM動作には、一つ又は二つのソース及び/又は対象アドレスが含まれ、ストライド媒介変数(parameter)は含まれない。

30

【0024】

2段階HBM+プロトコルは、一般的にホスト装置101が結果を要求する多数のソース及び/又は目的地アドレス、マトリックス動作、ストライドを含む動作のような、しかしこれに限定されないより複雑なPIM動作のためのものである。命令語の発行と応答出力との間の時間又は遅延は非決定論的であるから、ホスト装置101は、命令語の発行と応答出力との間、或いは命令語の発行とHBM+スタック103が他の命令語又はトランザクションを受信する準備ができた時点との間に他の動作を遂行する。後述の表2は、HBM+PIM命令語のカテゴリ及びタイミング推定の例を示す。

40

【0025】

50

図2は、一実施形態による1段階HBM+プロトコル命令語の一例のタイミング図200である。タイミング図200は、縮尺通りではなく、単に例示的なPIM命令語に対する1段階(one-step)HBM+プロトコルの相対的なシーケンスを描写するものであることを理解すべきである。タイミング図200は、一般的に特定のバスの単方向性又は両方向性を示すために異なるバスに関連する矢印を含む。特に、左側から右側に向かう矢印は、ホスト装置101からHBM+スタック103に伝送される通信、トランザクション、又は信号を示す。一方、右側から左側に向かう矢印は、HBM+スタック103からホスト装置101に伝送される通信、トランザクション、又は信号を示す。本明細書に開示する全てのタイミング図において、HBM+内部バス111を介して発生するものとして示したトランザクション及び動作は、HBM+スタック103の外部の装置には見えない。 10

【0026】

図2の201で、ホスト装置101はCAバス108を介してPIM_CHRG命令語を発行する。PIM_CHRG命令語は、発行されるPIM_CMDに対するバンク及びチャンネル識別情報を含む。ロジックダイ104は、PIM_CHRG命令語に応答して一つ以上の選択されたチャンネルの関連バンクをプリチャージし、一つ以上の選択されたチャンネルの関連バンクを固定(lock)して、ホスト装置101及び後続のPIM_CMDに対する関連バンクの所有権を保証する。PIM_CHRG命令語の直後に、ホスト装置101は、202でCAバス108を介してPIM_WR命令語を発行する。表1は、PIM_CHRG命令語及びPIM_WR命令語に対する例示的な物理的信号パラメータを示す。 20

【0027】

【表1】

Function	CKE_0		R[0]	R[1]	R[2]	R[3]	R[4]	R[5]
	Previous Cycle	Current Cycle						
PIM_CHRG	H	H	H	L	H	X	X	X
PIM_WR	H	H	L	L	L	X	X	X

【0028】

30

表1には、現在のJEDDEC HBM標準に基づく一つのピンの識別例を上段に示す。「H」は高い論理電圧レベルを表し、「L」は低い論理電圧レベルを表す。「X」は「相関なし(don't care)」の論理電圧レベルを表す。

【0029】

図3は、一実施形態によるPIM_CMD命令語300のコンテンツの一例の配置を示す図である。PIM_CMDは、実際のPIM命令語を含み、CAバス108を介して伝送されるものとは対照的に、データバス109を介して伝送される。ロジックダイ104は、PIM_CMDをデコードし、PIM_CMD命令に含まれるPIM動作を完了するために、HBM+スタック103内部の命令語を発行する。本実施形態において、全てのPIM命令語がPIM_CMDの全てのフィールドを使用するものではないが、PIM_CMDは固定された大きさ又は長さを有する。 40

【0030】

図3に示すように、PIM_CMD命令語300は、IDフィールド301、動作フィールド302、ストライドフィールド303、フラッグフィールド304、範囲フィールド305、値フィールド306、番号フィールド307、ソースアドレス1フィールド308、ソースアドレス2フィールド309、目的地アドレス1フィールド310、目的地アドレス2フィールド311を含む。本実施形態において、IDフィールド301は3ビットを含み、動作フィールド302は3ビットを含む。ストライドフィールド303は16ビットを含む。フラッグフィールド304は2ビットを含み、範囲フィールド305は32ビットを含む。値フィールド306は64ビットを含む。番号フィールド307は8 50

ビットを含む。ソースアドレス1フィールド308は32ビットを含み、ソースアドレス2フィールド309は32ビットを含む。目的地アドレス1フィールド310は32ビットを含み、目的地アドレス2フィールド311は32ビットを含む。その他、追加フィールド、フィールド配列、及びフィールドの大きさが追加され得る。

【0031】

再び図2を参照すると、202でCAバス108を介して伝送されたPIM_WR命令語は、203でデータバス109を介してPIM_CMDを伝送するために使用される。一実施形態において、データバス109に示すPIM_CMD命令語は、CAバス108上に示すPIM_WR命令語に後続する。他の実施形態において、データバス109上に示すPIM_CMD命令語は、CAバス108上に示すPIM_WR命令と少なくとも部分的に同時に有る。

10

【0032】

PIM_CMDに応答して、ロジックダイ104は、PIM_CMDに含まれるPIM動作に対するソースとなるデータを判読するために、HBM+内部バス111を介して204でソースリード命令語SRC_RDを伝送する。この例において、PIM動作(OOP)は205で遂行される。PIM動作の結果は、206でロジックダイ104により伝送されたDST_WR命令語を使用して目的地アドレスに記録される。207で、PIM動作は完了し、HBM+スタック103は、ホスト装置101から追加の正規動作、PIM動作、及び/又は他のトランザクションを受信する準備をする。

20

【0033】

図2に示した1段階HBM+プロトコルのタイミングは決定論的であり、従って、202でPIM_WR命令語が発行された時点と、HBM+スタック103がホスト装置101から他の命令語及び/又はトランザクションを受信する準備ができた(即ち、正規のトランザクションを再開する)時点との間の時間を t_{PIM_WR} で図2に示している。即ち、タイミング t_{PIM_WR} は、HBM+スタック103内の同一のチャンネル/バンクにホスト装置101によって新たな命令語又はトランザクションが発行される前の最小時間と見なされる。1段階HBM+プロトコルのタイミングが決定論的属性であることを示すために、図2に示した多様な命令語、トランザクション、及び応答の間の矢印を実線矢印で示している。なお、後述の図5及び図7は、それぞれ異なる決定論的タイミングシリオに対するタイミング図を示す。

30

【0034】

図4は、一実施形態による2段階HBM+プロトコル命令語の一例の概略的タイミング図400である。タイミング図400は、縮尺通りではなく、例示的なPIM命令語に対する2段階HBM+プロトコルの相対的なシーケンスを描写するものであることを理解すべきである。図2のタイミング図200と同様に、図4に示すタイミング図400は、一般的に特定のバスの単方向性又は両方向性を示すために異なるバスに関連する矢印を含む。

【0035】

描画スペースを考慮して図4には示していないが、ホスト装置101は、CAバス108を介してPIM_CHRG命令語を送り出す。1段階HBM+プロトコルと同様に、PIM_CHRG命令語はバンクを含み、PIM_CMDに対するチャンネル識別情報が直ぐに発行される。PIM_GHRG命令語に応答して、ロジックダイ104は、一つ以上の選択されたチャンネルの関連バンクをプリチャージし、HBM+スタック103の全体を固定して、ホスト装置101に対するHBM+スタック103の所有権を保証する。PIM_CHRG命令語(図示せず)の直後に、401で、ホスト装置101はCAバス108を介してPIM_WR命令語を発行する。402で、ホスト装置101はデータバス109上にPIM_CMDを発行する。

40

【0036】

PIM_CMDに応答して、ロジックダイ104は、403でソースリード命令語(SRC_RD)をHBM+内部バス111を介して伝送し、PIM_CMDに示す動作のた

50

めのソースデータを判読する。PIM動作(O P)は404で遂行される。PIM動作の結果は、405でDST_W R命令語を使用して目的地アドレスに記録される。図4に示す例示的なタイミングに対し、406で、ロジックダイ104は、トランザクションバス110を介して判読準備命令語(RD_R D Y)をホスト装置101に伝送する。407で、ホスト装置101は、CAバス108を介してリード命令語(RD)を伝送する。408で、ロジックダイ104は、リードデータ(RD_Data)を、データバス109を介して出力し、409で、ホスト装置101は、正規の動作及びトランザクションをもたらす。一部の2段階HBM+プロトコルPIM命令語に対し、判読準備命令語(RD_R D Y)をホスト装置101に伝送する代わりに、ホスト装置101が中間で他の作業をスケジュールするために、ロジックダイ104は、PIM動作がいつ完了するかに対する指示をホスト装置101に伝送する。そのような指示を伝送するロジックダイ104のいくつかの例示的なシナリオは、後述の図8及び図10で説明する。

【0037】

1段階HBM+プロトコルとは対照的に、2段階HBM+プロトコルは、PIM動作(O P)が一般的に2段階HBM+プロトコルに対してより複雑でありPIM動作に応じて一定ではない時間有るために、非決定論的タイミング特性を有する。例えば、 100×100 行列で遂行される転置行列のPIM演算は、PIM転置演算が両方の場合で同一の演算であっても、 10×10 行列に対する同一の転置行列のPIM演算よりも長くかかる。2段階HBM+動作に関連するタイミングが一般的に非決定論的であることを示すために、図4の命令語と応答との間に示した矢印中の一部を太い点線の矢印で示している。例えば、401におけるPIM_W R命令語と406におけるRD_R D Y応答との間の矢印と、406におけるRD_R D Y応答と407におけるRD命令語との間の矢印は、2段階HBM+プロトコルの非決定論的性質を示すために太い点線矢印で示す。また、非決定論的特性を、404におけるPIM_O Pと405におけるDST_W R応答との間の太い点線矢印で示す。

【0038】

【表2】

#	Category	Function Examples	Projected tIOP (cycles)	Type of Operation
1.	Data atomicity	Read-modify-write	1	Register copy
		Test and set	2	Register compare + copy
		Compare and Swap (CAS)	2	Register compare + copy
		Increment	1	Register ALU
2.	Data copying	Mem-copy	1	Register copy
		Mem-set	1	Register copy
3.	Data reshaping	Transpose, Pack/unpack, Swap	Protocol specified	Multiple copies
4.	Data reduction	Popcount, accumulation, bitwise operations, sum, min, dot-product	Protocol specified	Multiple copies, ALU
5.	Special functions	Map function, hash, pattern match	Protocol specified	Multiple copies, ALU

10

20

30

40

50

【0039】

ホスト装置101の観点において、例えば1段階HBM+プロトコルの決定論的特性は、ホスト基盤命令語スケジューリングプロトコルを提供するものと見なされる

【0040】

図5は、一実施形態によるPIM動作がHBM+装置内の単一のアドレスに向けられるか又は同一の行に向けられる1段階HBM+プロトコルPIM命令語の一例のタイミング図500である。タイミング図500は、縮尺通りではなく、単一のアドレスに向けられるか又はHBM+メモリ装置内の同一の行に向けられる例示的なPIM動作に対する1段階HBM+プロトコルの相対的なシーケンスを描写するものであることを理解すべきである。

10

【0041】

501で、ホスト装置101は、CAバス108を介してPIM_CHRG命令語を発行する。PIM_CHRG命令語は、直ぐに発行されるPIM_CMDに対するバンク及びチャンネル識別情報を含む。PIM_CHRG命令語に応答して、ロジックダイ104は、一つ以上の選択されたチャンネルの関連バンクをプリチャージし、PIM動作中、一つ以上の選択されたチャンネルの関連バンクを固定して、ホスト装置101に対する所有権を保証する。PIM_CHRG命令語の直後に、ホスト装置101は、502でCAバス108を介してPIM_WR命令語を発行する。503で、ホスト装置101はデータバス109を介してPIM_CMDを伝送する。この例では、PIM_CMDがHBM+メモリ装置105内のメモリ位置の値を増加させる命令であると見なす。PIM_WR命令語とPIM_CMD命令語との間の時間遅延は、 t_{WL} であり、これはPIM_WR命令語を送信してPIM_CMD命令語を送信する間に待機しなければならないライト時間の遅延である。PIM_CMD命令語は、DDR(Double Data Rate)HBMメモリ装置に対して $t_{BL/2}$ のバースト長さの遅延を要する。

20

【0042】

ロジックダイ104がPIM_CMDをデコーディングして504でHBM+内部バス111を介してHBM+スタック103にACT命令語を発行するまで t_{PD} の伝播遅延が存在する。選択された行に対する活性化遅延は t_{RC} である。505で、ロジックダイ104は、選択されたソースデータをロードするためにリード命令語(RD)を発行する。経路520に沿って選択されたデータをリードするまでの遅延は t_{RL} である。506で、選択されたソースデータは、 $t_{BL/2}$ の遅延で判読される。507で、PIM動作(IOP)は t_{IOP} の遅延で遂行される。この例で、PIM動作(IOP)は、相対的に複雑でないPIM動作であるHBM+メモリ装置105内のメモリ位置の値を増加させる。

30

【0043】

図5の530で示す並列経路において、ロジックダイ104は、507におけるPIM命令語(IOP)の終端でデータが再びメモリに記録されるため、 t_{RTW} の遅延後、508でライト命令語(WR)を発行する。ライト命令語(WR)に関連する遅延は t_{WL} である。本実施形態において、ロジックダイ104は、509で結果データを記録(WR_Data)するためのタイミングが正確であるように、508でライト命令語(WR)を伝送するための適切な時間を決定する。

40

【0044】

509で、ライト命令語(WR)に応答したPIM動作(IOP)の結果は、 $t_{BL/2}$ のバースト長さの遅延でメモリに再び記録(WR_Data)される。再びメモリに記録された結果の復旧時間遅延は t_{WR} である。510で、ロジックダイ104は結果が記録された行に対するプリチャージ命令語(PRE)を発行し、511で、ホスト装置101が追加のトランザクション及び/又は命令語をHBM+スタッ�103に発行する前に t_{RP} のプリチャージ遅延が続く。

【0045】

従って、PIM動作がHBM+メモリ装置105内の単一のアドレスに向けられるか又

50

は同一の行に向けられるこのシナリオにおいて、502でPIM_WR命令語が発行された時点と、511でホスト装置101からHBM+スタック103が他の命令語及び/又はトランザクションを受信する準備ができた時との間の時間(t_{PIM_WR})は、下記の数学式1の通りに決定される。

【0046】

【数1】

$$t_{PIM_WR} = t_{WL} + t_{BL/2} + t_{PD} + t_{RCD} + \max((t_{RL} + t_{BL/2} + t_{IOP}), (t_{RTW} + t_{WL})) + t_{BL/2} + t_{WR} + t_{RP}$$

【0047】

10

ここで、図5に示した経路(520、530)の間の最大遅延は、 t_{PIM_WR} を定義するために使用される。経路520の遅延は($t_{RL} + t_{BL/2} + t_{IOP}$)である。一方、経路530の遅延は($t_{RTW} + t_{WL}$)である。

【0048】

ホスト基盤命令語スケジューリングプロトコルを提供するものと見なされるもう一つの1段階HBM+プロトコルシナリオを図6に示す。

【0049】

図6は、一実施形態によるPIM動作が同一のチャンネルのバンク内で行われる1段階HBM+プロトコルPIM命令語の一例のタイミング図600である。特に、タイミング図600は、一定の縮尺ではなく、単に同一のチャンネルのバンクに向けられる例示的なPIM動作に対する1段階HBM+プロトコルの相対的なシーケンスを示すものであることを理解すべきである。

20

【0050】

601で、ホスト装置101は、CAバス108を介してPIM_CHRG命令語を発行する。PIM_CHRG命令語は、直ぐに発行されるPIM_CMDに対するバンク及びチャネル識別情報を含む。PIM_CHRG命令語に応答して、ロジックダイ104は、一つ以上の選択されたチャンネルの関連バンクをプリチャージし、PIM動作中、一つ以上の選択されたチャンネルの関連バンクを固定して、ホスト装置101に対する所有権を保証する。PIM_CHRG命令語の直後に、ホスト装置101は、602でCAバス108を介してPIM_WR命令語を発行する。603で、ホスト装置101はデータバス109を介してPIM_CMDを伝送する。この例では、PIM_CMDが、メモリ位置Bを、位置A及びBがHBM+メモリ装置105内の同一のチャンネル内のバンク内にあるメモリ位置Aと同一になるように設定する命令語であるものと見なす。PIM_WR命令語とPIM_CMD命令語との間のライト時間遅延は t_{WL} である。PIM_CMD命令語は、DDR HBMメモリ装置に対して $t_{BL/2}$ のバースト遅延を要する。

30

【0051】

603におけるPIM_CMDに応答して、ロジックダイ104は、604でHBM+内部バス111を介してHBM+スタック103に活性化命令語(ACT1)を伝送して第1ソースデータアドレス(即ち、位置A)を活性化させる。ロジックダイ104がPIM_CMDをデコーディングして604でACT1命令語を発行するまでの伝播遅延は t_{PD} である。

40

【0052】

第1並列経路620において、ロジックダイ104は、605で第2ソースデータアドレス(即ち、位置B)を活性化するために活性化命令語(ACT2)を発行する。ACT1命令語の発行とACT2命令語の発行との間の遅延は、 t_{RRD} 又は t_{RC} である。PIM動作が、二つの互いに異なるバンクの間で進行する場合、遅延は(一般的に) t_{RRD} になる(ソースと目的地アドレスが同一のバンクグループ内にある二つの互いに異なるバンクの間にある場合に遅延は t_{RRDL} になるが、ソースと目的地アドレスが他のバンクにある二つの異なるバンクにある場合、遅延は t_{RRDF} である。)。PIM動作が同一のバンク内にある場合、遅延は t_{RC} になる。この並列経路620において、ロジック

50

ダイ 1 0 4 が、 6 0 6 でライト命令語 (W R 2) を発行する前に t_{RCD} の遅延が存在し、 またライト命令語 (W R 2) に後続する t_{WL} の遅延が存在する。

【 0 0 5 3 】

第 2 並列経路 6 3 0 において、ロジックダイ 1 0 4 は、活性化命令語 (A C T 1) に応答して、 6 0 7 でリード命令語 (R D 1) を発行する。活性化命令語 (A C T 1) の後及びリード命令語 (R D 1) の前に t_{RCD} の遅延がある。リード命令語 (R D 1) が発行された時間と 6 0 8 のリードデータ (R D _ D a t a) の動作との間には t_{RL} の遅延がある。データは、 6 0 8 で $t_{BL/2}$ のバースト長さの遅延で読み出される (R D _ D a t a)。 6 0 9 で P I M 動作 (I O P) は t_{IOP} の遅延で遂行される。

【 0 0 5 4 】

P I M 動作 (I O P) で生成されたデータを記録するために、 6 1 0 でロジックダイ 1 0 4 がプリチャージ命令語 (P R E 1) を発行するためのリード命令語 (R D 1) 以降、 $t_{RTW} - t_{RCD}$ の遅延が存在する。 6 1 1 で、 6 0 6 におけるライト命令語 (W R 2) に応答して、 P I M 動作 (I O P) の結果は $t_{BL/2}$ の遅延でメモリに再び記録 (W R _ D a t a) される。再びメモリに記録された結果の復旧時間は t_{WR} である。 6 1 2 で、ロジックダイ 1 0 4 は、結果を復旧するために記録された行に対するプリチャージ命令語 (P R E 2) を発行し、 6 1 3 で、ホスト装置 1 0 1 が H B M + スタック 1 0 3 に追加のトランザクション及び / 又は命令語を発行する前に t_{RP} の遅延が続く。

【 0 0 5 5 】

従って、 P I M 動作が同一のチャンネル内のバンクに向けられるこのようなシナリオにおいて、 6 0 2 で P I M _ W R 命令語が発行された時点と、 H B M + スタック 1 0 3 が 6 1 3 でホスト装置 1 0 1 から他の命令語を受信する準備ができた時点との間の時間 (t_{PIMWWR}) は、下記の数学式 2 の通りに決定される。

【 0 0 5 6 】

【 数 2 】

$$t_{PIMWWR} = t_{WL} + t_{BL/2} + t_{PD} + \max((t_{RCD} + t_{RL} + t_{BL/2} + t_{IOP}), (t_{RRD} + t_{RCD} + t_{WL})) + t_{BL/2} + t_{WL} + t_{RP}$$

【 0 0 5 7 】

ここで、図 6 に示した経路 (6 2 0 、 6 3 0) の間の最大遅延は、 t_{PIMWWR} を定義するために使用される。経路 6 2 0 の遅延は ($t_{RCD} + t_{BL/2} + t_{IOP}$) である。一方、経路 6 3 0 の遅延は ($t_{RRD} + t_{RCD} + t_{WL}$) である。

【 0 0 5 8 】

ホスト基盤命令語スケジューリングプロトコルを提供するものと見なされるもう一つの 1 段階 H B M + プロトコルシナリオを図 7 に示す。

【 0 0 5 9 】

図 7 は、一実施形態による P I M 動作が異なるチャンネルを介して行われる 1 段階 H B M + プロトコル P I M 命令語の一例のタイミング図 7 0 0 である。特に、タイミング図 7 0 0 は、縮尺通りではなく、異なるチャンネルに亘る例示的な P I M 動作に対する 1 段階 H B M + プロトコルの相対的なシーケンスを示すためのものであることを理解すべきである。

【 0 0 6 0 】

7 0 1 で、ホスト装置 1 0 1 は、 C A バス 1 0 8 を介して P I M _ C H R G 命令語を発行する。 P I M _ C H R G 命令語は、直ぐに発行される P I M _ C M D に対するバンク及びチャネル識別情報を含む。 P I M _ C H R G 命令語に応答して、ロジックダイ 1 0 4 は、一つ以上の選択されたチャンネルの関連バンクをプリチャージし、 P I M 動作中、一つ以上の選択されたチャンネルの関連バンクを固定して、ホスト装置 1 0 1 に対する関連バンクの所有権を保証する。 P I M _ C H R G 命令語の直後に、ホスト装置 1 0 1 は、 7 0 2 で C A バス 1 0 8 を介して P I M _ W R 命令語を発行する。 7 0 3 で、ホスト装置 1 0 1 はデータバス 1 0 9 を介して P I M _ C M D を伝送する。 P I M _ W R 命令語と P I M

10

20

30

40

50

—CMD命令語との間の時間遅延は t_{WL} であり、これはPIM_WR命令語を送りPIM_CMD命令語を送る間に待機しなければならない時間である。PIM_CD命令語は、DDR_HBMメモリ装置に対して $t_{BL/2}$ のバースト長さの遅延を要する。

【0061】

703におけるPIM_CMDに応答して、ロジックダイ104は、704でHBM+内部バス111を介してHBM+スタック103に活性化命令語(ACT1)を伝送する。704で、ロジックダイ104がPIM_CMDをデコーディングしてACT1命令語を発行するまでの時間の遅延は t_{PD} である。遅延 t_{RCD} 後に、ロジックダイ104は705でリード命令語(RD1)を発行する。第1並列経路720において、706で $t_{BL/2}$ のバースト長さの遅延でデータが判読(RD_Data)される前に t_{RL} の遅延が存在する。707で、PIM動作(IOP)は t_{IOP} の遅延で遂行される。PIM動作(IOP)で生成されたデータを記録するためのロジックダイ104が708でプリチャージ命令語(PRE1)を発行するために、705のリード命令語(RD1)の以降に($t_{RTW} - t_{RCD}$)遅延が存在する。

【0062】

第2並列経路730において、ロジックダイ104は、709で活性化命令語(ACT2)を発行する。PIM動作が異なるチャンネルに亘って遂行される状況において、ロジックダイ104が活性化命令語(ACT2)を発行する時点に制約はない。ロジックダイ104が、710でライト命令語(WR2)を発行する前に t_{RCD} の遅延が存在する。711で、データが記録されるまでに(WR_Data)ライト命令語(WR2)が発行される時間の間に t_{WL} の遅延が存在する。711で、データは $t_{BL/2}$ のバースト長さの遅延で記録(WR_Data)される。ロジックダイ104が、712で復旧のために結果を記録した行に対してプリチャージ命令語(PRE2)を発行する前に t_{WR} の遅延があり、713で、ホスト装置101が追加トランザクション及び/又は命令語をHBM+スタック103に発行する前に t_{RP} の遅延が続く。

【0063】

従って、PIM動作が、他のチャンネルに亘るこのようなシナリオにおいて、702でPIM_WR命令語が発行された時点と、HBM+スタック103が713でホスト装置101から他の命令語を受信する準備ができた時点との間の時間(t_{PIM_WR})は、下記の数学式3に示される。

【0064】

【数3】

$$t_{PIM_WR} = t_{WL} + t_{BL/2} + t_{PD} + t_{RCD} + t_{RL} + t_{BL/2} + t_{IOP} + t_{BL/2} + t_{WR} + t_{RP}$$

【0065】

なお、後述する図8～図10は、PIM動作が、図5～図7で考慮された1段階HBM+プロトコル命令語よりもより複雑な2段階HBM+プロトコルPIM命令語に対するタイミング図を示し、結果的に完了のための全体の非決定論的タイミングを有する。従って、一部の2段階HBM+プロトコルPIM命令語は、クライアント基盤命令語スケジューリングプロトコルを提供するものと見なされ、HBM+スタックは、PIM命令語が2段階HBM+プロトコルPIM命令語のうちの一部の非決定論的特徴を説明するために完了する時点を通知する。これに関して、ロジックダイ104は、複雑なPIM命令語が完了する時間をホスト装置101に通知する。

【0066】

一実施形態において、有効時間の指示は時間の推定を含む。他の実施形態において、有効時間の指示はクレジット基盤の値を含む。更に他の実施形態において、有効時間の指示は、ホスト装置にPIM動作が完了したか否かを知らせるために、HBM+スタックをいつポーリングするかに対する指示が与えられる再試行基盤のフィードバックを含む。ホスト装置101に提供されるPIM命令語が完了する時間に対する指示は、自動ログの内容

10

20

30

40

50

、履歴統計情報、計算された推定値、進行中のトラフィック、最大PIM作業範囲などに基づくか又はこれらから予測されるが、本発明は、これらに限定されるものではない。

【0067】

後述する図8～図10のタイミング図に示す2段階HBM+プロトコルPIM命令語は、図1に示したトランザクションバス110を使用しない。その代わりに、図8～図10に示す2段階HBM+プロトコルPIM命令語は、CAバス108、データバス109、及びHBM+内部バス111を利用して、PIM命令語が完了する有効時間の表示をホスト装置101に提供する。

【0068】

図8は、一実施形態によるPIM命令語が完了する時点でHBM+スタック103が推定値を提供する2段階HBM+プロトコルPIM命令語の一例のタイミング図800である。タイミング図800は、縮尺通りではなく、単にHBM+スタックが例示的なPIM動作が完了する時点で推定値を提供する場合の2段階HBM+プロトコルの相対的なシーケンスを描写するものであることを理解すべきである。

【0069】

801で、ホスト装置101は、CAバス108を介してPIM_CHRG命令語を発行する。PIM_CHRG命令語は、直ぐに発行されるPIM_CMDに対するバンク及びチャンネル識別情報を含む。PIM_CHRG命令語に応答して、ロジックダイ104は、一つ以上の選択されたチャンネル内の関連バンクをプリチャージし、PIM動作中、HBM+スタック103を固定して、ホスト装置101に対するHBM+スタックの所有権を保証する。PIM_CHRG命令語の直後に、ホスト装置101は、802でCAバス108を介してPIM_WR命令語を発行する。ホスト装置101は、803でデータバス109を介してPIM_CMDを伝送する。PIM_WR命令語とPIM_CMD命令語との間の遅延は t_{WL} である。PIM_CMD命令語は、DDR_HBMメモリ装置に対して $t_{BL/2}$ バースト長さの遅延を要する。

【0070】

803におけるPIM_CMDに応答して、ロジックダイ104は、804でHBM+スタック103にHBM+内部バス111を介して活性化命令語(ACT)を伝送する。804でロジックダイ104がPIM_CMDをデコーディングしてACT命令語を発行する時間の遅延は t_{PD} である。遅延 t_{RC_D} の後に、ロジックダイ104は805でリード命令語(RD)を発行する。806で $t_{BL/2}$ のバースト長さの遅延でデータが判読(RD_Data)される前に t_{RL} の遅延が存在する。807で、PIM動作(IOP)は t_{IOP} の遅延で遂行される。しかし、PIM演算(IOP)が複雑であるため、PIM演算(IOP)に関連する遅延は非決定論的である。

【0071】

また、ロジックダイ104は、リード命令語(RD)が805で発行された後、PIM動作(IOP)が807で完了する前に808で t_{RTW} の遅延でライト命令語(WR)を発行する。PIM動作(IOP)の結果は、808におけるライト命令語(WR)の発行後に t_{WL} の遅延を有して809でメモリに記録(WR_Data)される。810で、ロジックダイ104は t_{WR} の遅延を有してプリチャージ命令語(PRE)を発行する。ホスト装置101が、811でHBM+スタック103に追加のトランザクション及び/又は命令語を発行する前に t_{RP} の遅延が続く。

【0072】

タイミング図800に示したトランザクションの多くは決定論的な側面であるが、トランザクション全体の全体的なタイミングは非決定論的である。807におけるPIM動作(IOP)の非決定論的遅延を説明するために、ホスト装置101は、812でCAバス108を介してPIM_RD命令語を発行する。ロジックダイ104が、813におけるPIM_EST応答で応答する前に t_{RL} の遅延がある。この例で、PIM_EST応答は、807におけるPIM動作(IOP)が完了する時間を示す時間推定を含む。一実施形態において、時間の推定は時間の単位である。他の実施形態において、時間の推定はク

10

20

30

40

50

ロックサイクルの単位からなる。

【0073】

従って、812でPIM_WR命令語が発行された時点と、811でHBM+スタック103がホスト装置101から他の命令語を受信する準備ができた時点との間の有効時間 $t_{PIM_WR(effective)}$ は、下記の数学式4に示される。

【0074】

【数4】

$$t_{PIM_WR(effective)} = t_{PIM_WR} + t_{PIM_EST}$$

10

【0075】

ここで、 t_{PIM_WR} はPIM演算が完了するのにかかる時間の決定論的部分を示し、 t_{PIM_EST} はPIM演算が完了するまでの時間の非決定論的部分の時間推定を示す。

【0076】

図9は、一実施形態によるPIM命令語が完了する時点でHBM+スタック103がクレジット基盤の表示を提供する2段階HBM+プロトコルPIM命令語の一例のタイミング図900である。タイミング図900は、縮尺通りではなく、単にHBM+スタックが例示的なPIM動作が完了する点でクレジット基盤の表示を提供する場合の2段階HBM+プロトコルの相対的なシーケンスを描写するものであることを理解すべきである。

20

【0077】

901で、ホスト装置101は、CAバス108を介してPIM_CHRG命令語を発行する。PIM_CHRG命令語は、直ぐに発行されるPIM_CMDに対するバンク及びチャンネル識別情報を含む。PIM_CHRG命令語に応答して、ロジックダイ104は、一つ以上の選択されたチャンネル内の関連バンクをプリチャージし、PIM動作中、HBM+スタック103を固定して、ホスト装置101に対するHBM+スタックの所有権を保証する。PIM_CHRG命令語の直後に、ホスト装置101は、902でCAバス108を介してPIM_WR命令語を発行する。ホスト装置101は、903でデータバス109を介してPIM_CMDを伝送する。PIM_WR命令語とPIM_CMD命令語との間の遅延は t_{WL} である。PIM_CMD命令語は、DDR_HBMメモリ装置に対して $t_{BL/2}$ のバースト長さの遅延を要する。

30

【0078】

903におけるPIM_CMDに応答して、ロジックダイ104は、904でHBM+スタック103にHBM+内部バス111を介して活性化命令語(ACT)を伝送する。904でロジックダイ104がPIM_CMDをデコーディングしてACT命令語を発行する時間の遅延は t_{PD} である。遅延 t_{RCD} の後に、ロジックダイ104は905でリード命令語(RD)を発行する。906で $t_{BL/2}$ のバースト長さの遅延でデータが判読(RD_Data)される前に t_{RL} の遅延が存在する。907で、PIM動作(IOP)は t_{IOP} の遅延で遂行される。しかし、PIM演算(IOP)が複雑であるため、PIM演算(OP)に関連する遅延は非決定論的である。

40

【0079】

また、ロジックダイ104は、リード命令語(RD)が905で発行された後、PIM動作(IOP)が907で完了する前に908で t_{RTW} の遅延でライト命令語(WR)を発行する。PIM動作(IOP)の結果は、908におけるライト命令語(WR)の発行後に t_{WL} の遅延を有して909でメモリに記録(WR_Data)される。910で、ロジックダイ104は、 t_{WR} の遅延を有してプリチャージ命令語(PRE)を発行する。ホスト装置101が、911でHBM+スタック103に追加のトランザクション及び/又は命令語を発行する前に t_{RP} の遅延が続く。

【0080】

タイミング図900に示したトランザクションの多くは決定論的な側面であるが、トラ

50

ンザクション全体の全体的なタイミングは非決定論的である。907におけるPIM動作(IOP)の非決定論的遅延を説明するために、ホスト装置101は、912でCAバス108を介してPIM_RD命令語を発行する。ロジックダイ104が、913におけるPIM_CRED応答で応答する前に t_{RL} の遅延がある。この例で、PIM_CRED応答は、ホスト装置101がスロットリング(throttling)メカニズムとして使用する多数のクレジット(credit)に関する情報を含む。例えば、PIM_CRED応答が、ホスト装置101が0より大きな整数のクレジットを有することを示す場合、ホスト装置101は、クレジットが残らなくなるまで、HBM+スタック103に命令語及び/又はトランザクションを発行し続ける。

【0081】

10

従って、912でPIM_RD命令語が発行された時点と、911でHBM+スタック103がホスト装置101から他の命令語を受信する準備ができた時点との間の有効時間 $t_{PIM_WR(effective)}$ は、下記の数学式5に示される。

【0082】

【数5】

$$t_{PIM_WR(effective)} = t_{PIM_WR} + t_{PIM_CRED}$$

【0083】

20

ここで、 t_{PIM_WR} はPIM動作が完了するのにかかる時間の決定論的部分を示し、 t_{PIM_CRED} は0よりも大きな整数のクレジットを示し、ホスト装置101がこれ以上残ったクレジットがなくなるまで、ホスト装置101は、HBM+スタック103に命令語及び/又はトランザクションを発行し続ける。

【0084】

20

図10は、一実施形態によるPIM命令語が完了する時点でHBM+スタック103がフィードバック基盤表示を提供する2段階HBM+プロトコルPIM命令語の一例のタイミング図1000である。タイミング図1000は、縮尺通りではなく、単にHBM+スタックが例示的なPIM動作が完了する時点でフィードバック基盤表示を提供する場合の2段階HBM+プロトコルの相対的なシーケンスを描写するものであることを理解すべきである。

30

【0085】

1001で、ホスト装置101は、CAバス108を介してPIM_CHRG命令語を発行する。PIM_CHRG命令語は、直ぐに発行されるPIM_CMDに対するバンク及びチャンネル識別情報を含む。PIM_CHRG命令語に応答して、ロジックダイ104は、一つ以上の選択されたチャンネル内の関連バンクをプリチャージし、PIM動作中、HBM+スタック103を固定して、ホスト装置101に対するHBM+スタックの所有権を保証する。PIM_CHRG命令語の直後に、ホスト装置101は、1002でCAバス108を介してPIM_WR命令語を発行する。ホスト装置101は、1003でデータバス109を介してPIM_CMDを伝送する。PIM_WR命令語とPIM_CMD命令語との間の遅延は t_{WL} である。PIM_CMD命令語は、DDR_HBMメモリ装置に対して $t_{BL/2}$ のバースト長さの遅延を要する。

40

【0086】

1003におけるPIM_CMDに応答して、ロジックダイ104は、1004でHBM+スタック103にHBM+内部バス111を介して活性化命令語(ACT)を伝送する。1004で、ロジックダイ104がPIM_CMDをデコーディングしてACT命令語を発行する時間の遅延は t_{PD} である。遅延 t_{RC_D} の後に、ロジックダイ104は1005でリード命令語(RD)を発行する。1006で $t_{BL/2}$ のバースト長さの遅延でデータが判読(RD_Data)される前に t_{RL} の遅延が存在する。1007で、PIM動作(IOP)は t_{IOP} の遅延で遂行される。しかし、PIM演算(IOP)が複雑であるため、PIM演算(OP)に関連する遅延は非決定論的である。

50

【0087】

また、ロジックダイ104は、リード命令語(RD)が1005で発行された後、PIM動作(IOP)が1007で完了する前に1008で t_{RTW} の遅延でライト命令語(WR)を発行する。PIM動作(IOP)の結果は、1008におけるライト命令語(WR)の発行後に t_{WL} の遅延を有して1009でメモリに記録(WR_Data)される。1010で、ロジックダイ104は、 t_{WR} の遅延を有してプリチャージ命令語(PRE)を発行する。ホスト装置101が、1011でHBM+スタック103に追加のトランザクション及び/又は命令語を発行する前に t_{RP} の遅延が続く。

【0088】

タイミング図1000に示したトランザクションの多くは決定論的な側面であるが、トランザクション全体の全体的なタイミングは非決定論的である。1007におけるPIM動作(IOP)の非決定論的遅延を説明するために、ホスト装置101は、1012でCバス108を介してPIM_RD命令語を発行する。ロジックダイ104が、1013におけるPIM_FDBK応答で応答する前に t_{RL} の遅延がある。この例で、PIM_FDBK応答は、PIM動作が完了したか否かを判別するために、ホスト装置101がHBM+スタック103をポーリングする前の時間周期に関する情報を含む。ホスト装置101は、フィードバック情報を使用してHBM+スタック103をポーリングするために戻る前に他の動作をスケジューリングして遂行する。

【0089】

従って、1012でPIM_RD命令語が発行された時点と、1011でHBM+スタック103がホスト装置101から他の命令語を受信する準備ができた時点との間の有効時間 $t_{PIM_WR(effective)}$ は、下記の数学式6に示される。

【0090】

【数6】

$$t_{PIM_WR(effective)} = t_{PIM_WR} + t_{PIM_FBDBK}$$

【0091】

ここで、 t_{PIM_WR} はPIM動作が完了するのにかかる時間の決定論的部分を示し、 t_{PIM_FBDBK} はPIM動作が完了したか否かを決定するためにホスト装置101がHBM+スタック103をポーリングするまでの時間周期に関する情報を示す。

【0092】

以上、本発明の実施形態について図面を参照しながら詳細に説明したが、本発明は、上述の実施形態に限定されるものではなく、本発明の技術的範囲から逸脱しない範囲内で多様に変更実施することが可能である。

【産業上の利用可能性】

【0093】

本発明は、ホスト装置とインターフェースするためのメモリにおいて決定論的処理及び非決定論的処理を可能にするシステムに有用である。

【符号の説明】

【0094】

- 100 HBM+システム
- 101 ホスト装置
- 102 インターポーザー
- 103 HBM+スタック
- 104 ロジックダイ(論理回路)
- 105a ~ 105d HBM+メモリ装置
- 106 基板
- 107 端子
- 108 C A(命令語/アドレス)バス

10

20

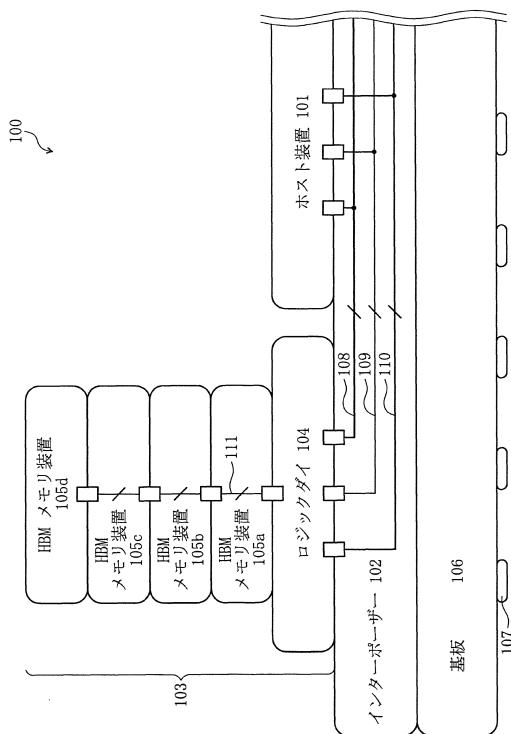
30

40

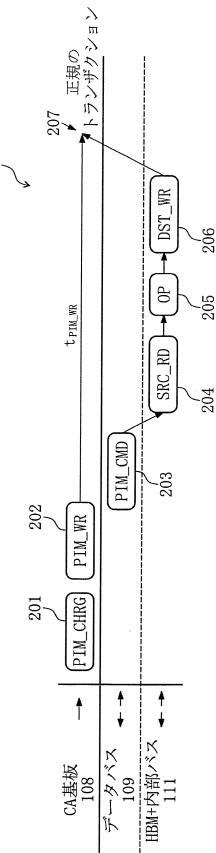
50

1 0 9	データバス	
1 1 0	トランザクションバス	
1 1 1	HBM + 内部バス	
2 0 0、4 0 0、6 0 0、7 0 0、8 0 0、9 0 0、1 0 0 0		タイミング図
3 0 0	P I M _ C M D 命令語	
3 0 1	I D フィールド	
3 0 2	動作フィールド	
3 0 3	ストライドフィールド	
3 0 4	フラッグフィールド	
3 0 5	範囲フィールド	10
3 0 6	値フィールド	
3 0 7	番号フィールド	
3 0 8	ソースアドレス 1 フィールド	
3 0 9	ソースアドレス 2 フィールド	
3 1 0	目的地アドレス 1 フィールド	
3 1 1	目的地アドレス 2 フィールド	

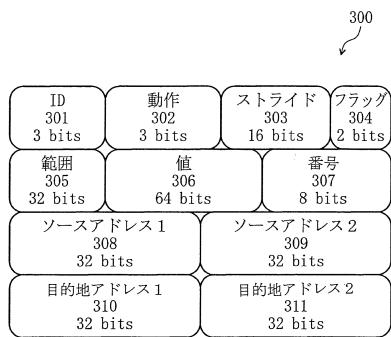
【図 1】



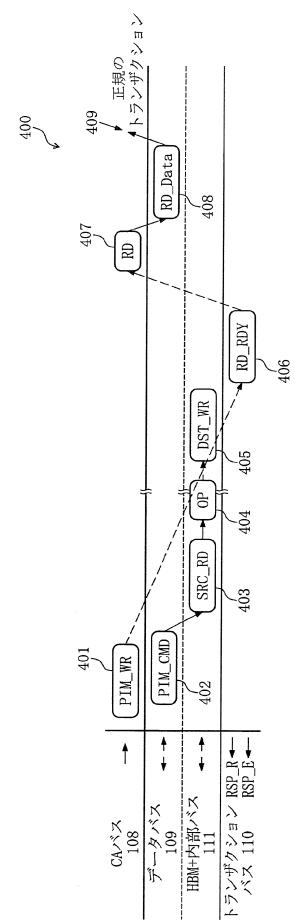
【図 2】



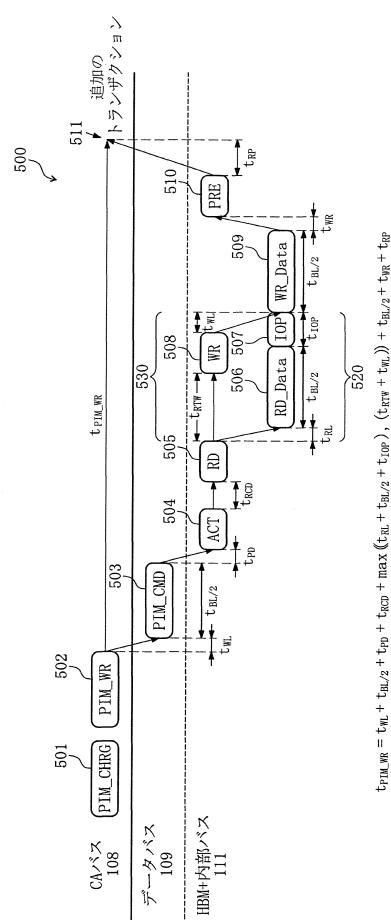
〔 叴 3 〕



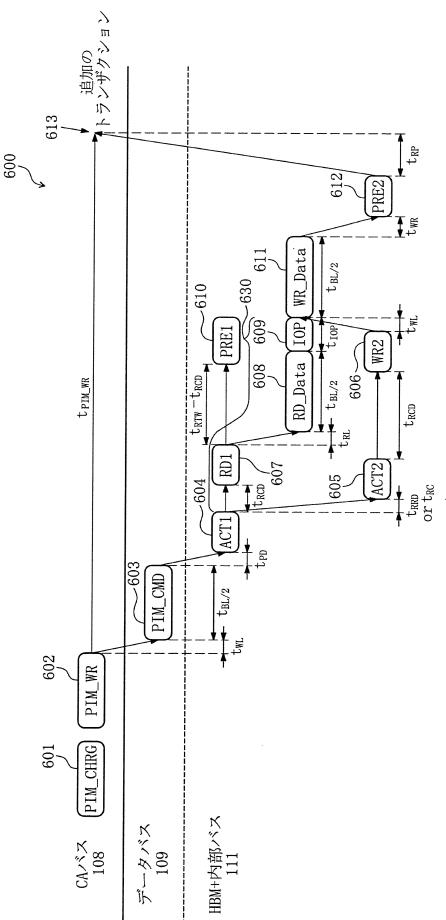
【 図 4 】



【 义 5 】

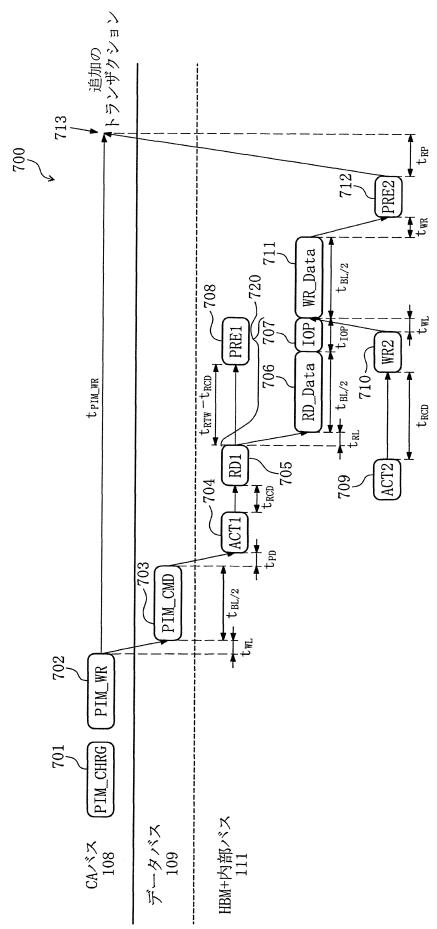


【図6】

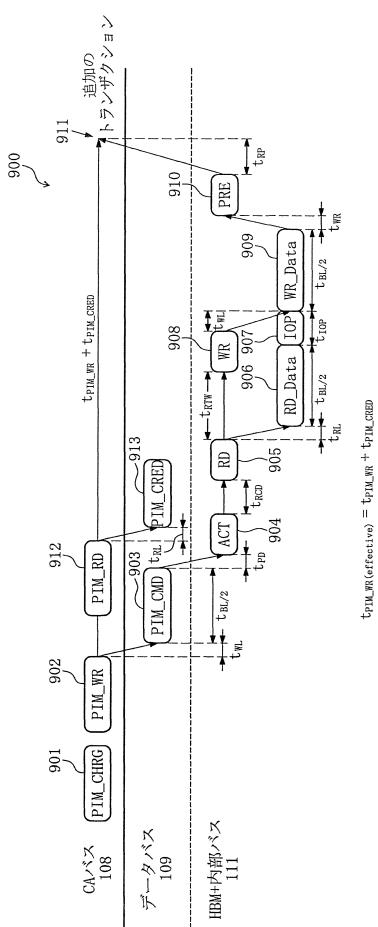


$$t_{PIM_WR} = t_{WL} + t_{BL/2} + t_{PD} + \max(t_{RCD} + t_{RL} + t_{BL/2} + t_{WL}), (t_{RCD} + t_{RCD} + t_{WL}) + t_{BL/2} + t_{WR} + t_{RP}$$

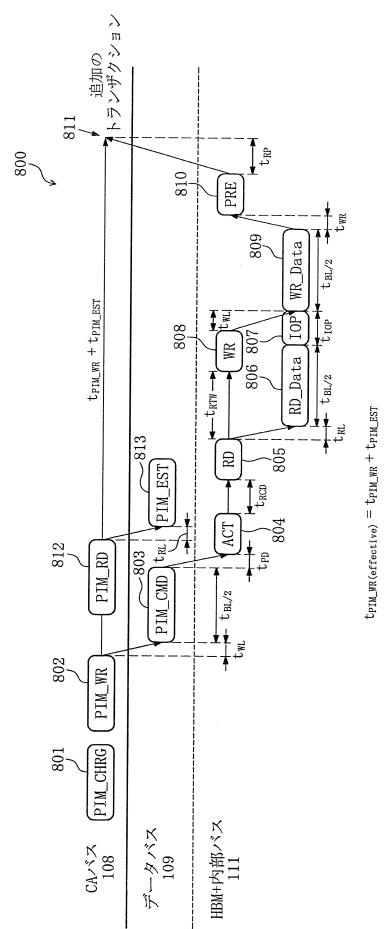
【図7】



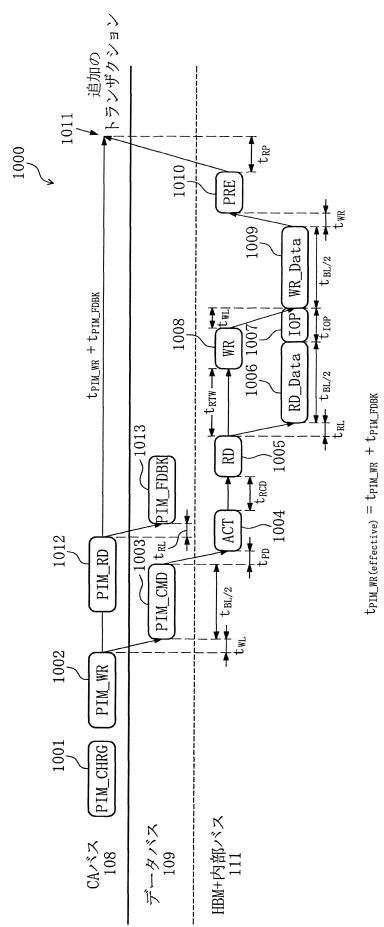
〔 図 9 〕



【 义 8 】



【図 1 0



$$t_{PIM_WR} = t_{WL} + t_{BL}/2 + t_{PP} + t_{RCD} + t_{RL} + t_{BL}/2 + t_{IOP} + t_{BL}/2 + t_{VR} + t_{RCD}$$

$$t_{PIM_WR(effective)} = t_{PIM_WR} + t_{PIM_EST}$$

$$t_{PIM_WR(effective)} = t_{PIM_WR} + t_{PIM_FDBK}$$

フロントページの続き

(51)Int.Cl. F I
G 1 1 C 5/06 2 0 0

(31)優先権主張番号 62/558,741
(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 15/821,686
(32)優先日 平成29年11月22日(2017.11.22)
(33)優先権主張国・地域又は機関
米国(US)

早期審査対象出願

(72)発明者 鄭 宏 忠
アメリカ合衆国, 95032, カリフォルニア州, ロス ガトス, カールトン アベニュー
- # 6, 120
(72)発明者 ブレナン, ロバート
アメリカ合衆国, 95054, カリフォルニア州, サンタ クララ, ミル クリーク レ
ーン # 201, 586

審査官 後藤 彰

(56)参考文献 特開2011-81732(JP, A)
特開2010-80802(JP, A)

(58)調査した分野(Int.Cl., DB名)

G 0 6 F 1 2 / 0 0
G 0 6 F 1 2 / 0 6
G 1 1 C 5 / 0 4
G 1 1 C 5 / 0 6