

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3807582号
(P3807582)

(45) 発行日 平成18年8月9日(2006.8.9)

(24) 登録日 平成18年5月26日(2006.5.26)

(51) Int. Cl.	F I
G06F 12/08 (2006.01)	G06F 12/08 501C
G06F 12/06 (2006.01)	G06F 12/08 501D
	G06F 12/08 519Z
	G06F 12/08 503B
	G06F 12/08 575
請求項の数 11 (全 16 頁) 最終頁に続く	

(21) 出願番号	特願平11-39746	(73) 特許権者	503121103
(22) 出願日	平成11年2月18日(1999.2.18)		株式会社ルネサステクノロジ
(65) 公開番号	特開2000-242559(P2000-242559A)		東京都千代田区丸の内二丁目4番1号
(43) 公開日	平成12年9月8日(2000.9.8)	(74) 代理人	100081938
審査請求日	平成15年3月26日(2003.3.26)		弁理士 徳若 光政
		(74) 代理人	100075096
			弁理士 作田 康夫
		(72) 発明者	菅野 雄介
			東京都国分寺市東恋ヶ窪一丁目280番地
			株式会社日立製作所中央研究
			所内
		(72) 発明者	水野 弘之
			東京都国分寺市東恋ヶ窪一丁目280番地
			株式会社日立製作所中央研究
			所内
			最終頁に続く

(54) 【発明の名称】 情報処理装置及び半導体装置

(57) 【特許請求の範囲】

【請求項1】

中央演算部と、
 ライトバック形式のキャッシュメモリと、
 複数のメモリバンクと、データプリロードレジスタと、リードデータ線と、ライトデータ線とを有するメインメモリと、
 入力アドレスから前記メインメモリのアクセスのためのバンクアドレス、ロウアドレス、カラムアドレスを形成するアドレス変換回路と、
 前記中央演算部と前記キャッシュメモリと前記メインメモリ間を結ぶデータ線とを備え、
 前記アドレス変換回路は、前記入力アドレスのINDEXビットの一部のビットを固定した値として前記入力アドレスのTAGビットの一部のビットを変化させた場合に互いに異なる前記バンクアドレス及び同一のカラムアドレスとロウアドレスを出力するものであり、
 前記中央演算部からのリクエストアドレスによる前記キャッシュメモリへのアクセスがミスした場合には、
 前記キャッシュメモリは、リクエストアドレスとINDEXビットが同じでTAGビットの異なるライトアドレスを出力し、
 前記アドレス変換回路は、前記リクエストアドレスに対応するバンクアドレスであるリクエストバンクアドレスを出力し、かつ、前記ライトアドレスに対応するバンクアド

レスであるライトバンクアドレスを出力し、

前記キャッシュメモリからのライトデータは、前記データ線から前記プリロードレジスタへ格納された後、前記ライトデータ線を経由して前記ライトバンクアドレス、前記ロウアドレス、前記カラムアドレスにより指定される箇所へ格納され、

前記リクエストバンクアドレス、前記ロウアドレス、前記カラムアドレスにより指定される箇所のリードデータは、前記リードデータ線を介して前記各バンクから前記データ線へ出力されるものであって、前記リクエストバンクアドレスにより指定されるバンクと、前記ライトバンクアドレスにより指定されるバンクとは同一サイクルでアクセスが開始されることを特徴とする情報処理装置。

【請求項 2】

請求項 1 において、

前記 INDEX ビットの一部のビットのビット数、前記 TAG ビットの一部のビットのビット数、及び前記バンクアドレスのビット数は互いに等しいことを特徴とする情報処理装置。

【請求項 3】

請求項 1 において、

前記複数のメモリバンクの数は、2 の N 乗個 ($N = 1, 2, 3, \dots$) であり、

前記 INDEX ビットの一部のビットは、前記 INDEX ビットの下位の N ビットであり、

前記 TAG ビットの一部のビットは、前記 TAG ビットの下位の N ビットであることを特徴とする情報処理装置。

【請求項 4】

請求項 1 において、

前記キャッシュメモリは、前記 INDEX ビットで表される数を持つ複数のエントリーブブロックを含み、

前記複数のエントリーブブロックのそれぞれは、前記リクエストアドレスに含まれる前記 TAG ビットと比較するためのエントリーブ TAG アドレスを記憶するためのアドレス領域と前記メインメモリのデータが複写すべきデータ領域とを有することを特徴とする情報処理装置。

【請求項 5】

請求項 1 において、

前記アドレス変換回路は、前記リクエストアドレスの INDEX ビットの一部のビットと前記リクエストアドレスの TAG ビットの一部のビットとを受けて、前記リクエストバンクアドレスを示すバンクアドレスビットを出力する論理回路を含むものであり、

前記キャッシュメモリは、前記キャッシュメモリに含まれる一つのブロック中で書き換えられたデータをそれに対応するアドレスを持つメインメモリに書き戻す動作であるライトバック動作を行うための 前記ライトアドレスを示すライトバックアドレスビットを出力可能とするものであり、

前記アドレス変換回路は、前記ライトバックアドレスビットの一部のビットと前記 INDEX ビットの一部のビットとを受けて、前記ライトバンクアドレスを示すライトバック用バンクアドレスビットを出力する第 2 論理回路を更に有し、

前記ライトバック動作を行う場合に、前記ライトバック用バンクアドレスビットは、書き込みが行うべき前記複数のメモリバンクの一つを指定することを特徴とする情報処理装置。

【請求項 6】

請求項 5 において、

前記第 2 論理回路は、前記論理回路と同一の論理演算をする回路であることを特徴とする情報処理装置。

【請求項 7】

請求項 5 において、

10

20

30

40

50

前記情報処理装置は、前記アドレス変換回路から前記メインメモリに前記バンクアドレスビットを伝達するための第1アドレスバスと前記アドレス変換回路から前記メインメモリに前記ライトバック用バンクアドレスビットを伝達するための第2アドレスバスとを更に有し、

前記メインメモリは、前記バンクアドレスビットによって選択される前記複数のメモリバンクの一つからの読み出し動作と、前記ライトバック用バンクアドレスビットによって選択される前記複数のメモリバンクの他の一つへの書き込み動作とが、並列して実行される期間を有することを特徴とする情報処理装置。

【請求項8】

請求項1から7のいずれかにおいて、

前記キャッシュメモリは、ダイレクトマップ方式又はセットアソシアティブ方式とされることを特徴とする情報処理装置。

【請求項9】

請求項1から8のいずれかにおいて、

前記中央演算部と、前記キャッシュメモリとは第1半導体チップ上に形成され、前記アドレス変換回路は第2半導体チップ上に形成され、前記メインメモリは第3半導体チップ上に形成されたメモリチップを含むことを特徴とする情報処理装置。

【請求項10】

請求項1から8のいずれかにおいて、

前記中央演算部と、前記キャッシュメモリと、前記アドレス変換回路は第1半導体チップ上にモノリシックに形成され、前記メインメモリは第2半導体チップ上に形成されたメモリチップを含むことを特徴とする情報処理装置。

【請求項11】

請求項1から10のいずれかにおいて、

前記メインメモリはダイナミック形メモリであることを特徴とする情報処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

CPUとメモリを含む情報処理システムにおいて、特に、キャッシュメモリを用いる場合の主記憶装置(メインメモリ)に対するアドレスの割付に関する。

【0002】

【従来の技術】

この明細書で参照される文献のリストは以下の通りであり、文献の参照は文献番号をもってすることとする。[文献1]:日経マイクロデバイス 1998年2月号、pp.134-141、[文献2]:David A. Patterson and John L Hennessy, "Computer Architecture A Quantitative Approach Second Edition", Morgan Kaufmann Publisher Inc., (1996), pp.375-384。

【0003】

[文献1]には、DRAMへのアドレスマッピング例がpp.141に開示されている。アドレスマッピングとは、アドレス信号とDRAMチップ内で選択される場所との対応づけのことである。この例では1チップ4バンクの64Mbit DRAMを2チップ用いている。この例で64ビット単位の連続するアドレスへのアクセスは、同文献の図C(a)に示すようにアドレスビットの下位から列、行、デバイス、バンクの順にアドレスを割り当てると、第一チップと第二チップの2つのメモリバンクへ隙間なく16MBのデータを記憶できる。同文献の図C(b)はアドレスビットを下位から列、バンク、デバイス、行の順で割り当てるとの例である。これにより8個のバンクに連続するアクセスが分散するようにデータを記憶することが記載される。

【0004】

【発明が解決しようとする課題】

本願に先立ち本願発明者は、現実の情報処理システムが中央演算部(以後CPUとよぶ)と

10

20

30

40

50

キャッシュメモリ及びメインメモリとからなり、それらの相互関係を考慮したアドレスマッピングについて検討した。その結果、CPUの発行するアドレスは、キャッシュメモリによってメインメモリへ行われるものが選別されるため、キャッシュメモリとの関係を考慮したメインメモリへのアドレス変換（アドレスマッピング）の必要性があることを見いだした。

【0005】

図2は、本願に先だって検討したキャッシュメモリのアドレス管理を示す図である。この図は、物理アドレスをキャッシュメモリが管理する領域に分けたものであり、[文献2]のpp.378よりの引用である。キャッシュメモリは、物理アドレスを大きくブロックオフセットとブロックアドレスの2つに分割して扱う。ブロックオフセットはキャッシュメモリが保持するデータを扱う最小単位であり、ラインサイズとも呼ばれる。ブロックアドレスはブロックオフセット単位のアドレスである。また、ある種のキャッシュメモリは、ブロックアドレスをさらに二分割して管理する方式をとる。これは、ダイレクトマップ方式、セットアソシアティブ方式と呼ばれるキャッシュメモリの場合である。これらの方式では、下位をINDEX、上位をTAGと呼ぶ。

【0006】

図3はキャッシュメモリがアドレスとデータを管理する様子を示したものである。この図は、[文献2]のpp.381からの引用である。ここでCMはキャッシュメモリであり、容量が8-KBでラインサイズが32Bのダイレクトマップのキャッシュメモリを示している。まず、CPUからメモリシステムへの要求アクセスにおいて、キャッシュメモリへリードアクセスがヒットした場合について説明する。CPUからのリクエストアドレス(以後、リクエストアドレスとする)はアドレス線にてキャッシュメモリへ伝達される(図中 1)。その後、リクエストアドレスの一部であるINDEX情報を元に、キャッシュメモリにエンターされている番号が決定される。ここではダイレクトマップ形式のキャッシュメモリを例にしているため、INDEXアドレスの総数(この例では2の8乗の256)とキャッシュメモリのエンター数が一致している。そのためこの対応は1対1対応できる(図中 2)。アドレスのINDEX情報からキャッシュメモリのエンター番号が選択された後、そのエンター番号で示されるエンターに格納されているTAGと、リクエストアドレスのTAGを比較する(図中 3)。これはキャッシュメモリのエンターが有効である場合のみ行う(有効ビットValidにて確認する)。リクエストアドレスとキャッシュにエンターされているTAGが一致している場合は、ブロックオフセットのアドレスを利用して、4:1のマルチプレクサで所望の8バイトのデータをCPUへ伝達させる(図中 4)。一方、リクエストアドレスとキャッシュに保持されているデータのアドレスが一致しなかった場合はメインメモリへアクセスすることになる。

【0007】

次にメモリシステムへの書き込みについて述べる。ここでは、ライトバック形式(後述)のキャッシュメモリにおいて、キャッシュメモリへライトアクセスがヒットした場合について説明する。CPUから書き込み要求が発生すると、前述したリードアクセスと同様、キャッシュメモリへリクエストアドレスが伝達される。その後、前述した手続きをおこない、キャッシュメモリに保持されたデータのアドレスと一致するかが判定される。ここで、リクエストアドレスのTAGがキャッシュメモリ内に保持されているTAGと一致していれば、キャッシュメモリに保持しているデータを変更し、メインメモリのデータとは一致していないことを示すダーティービットを立てる(不図示)。CPUはキャッシュメモリの内容のみ更新し処理を続行することが可能なので高速アクセスができる。

【0008】

さて、このようにキャッシュメモリの内容が変化している状態で、次のアクセスがキャッシュミスである場合の処理について述べる。この場合、キャッシュメモリの内容を新たに要求された内容に置き換える必要が生じる。これは、参照の局所性(新しく必要となったデータは近い将来ふたたび必要とされる確率が高いという性質)を利用しているためである。このキャッシュ内容の置き換え動作を以後ライトバックと呼び、この形式をとるキャ

10

20

30

40

50

ッシュメモリをライトバック型キャッシュメモリと呼ぶ。キャッシュメモリはアドレスの一部であるINDEX部でデータを管理するので、ここで置き換えられるアドレスはリクエストアドレスとINDEX部が同じでTAG部が異なる性質がある（ブロックオフセット分はそっくり置き換えられる。これをキャッシュのリプレースと呼ぶ）。

【0009】

このようなキャッシュメモリの動作を考慮すると、メインメモリへのアクセスには、参照の局所性とライトバック時のINDEXが同じでTAGが異なるアドレス間のアクセスの2種類の性質を考慮する必要がある。[文献1]の例では、アクセスの空間的局所性に依存した前者のタイプのアクセスに対しては、アクセスが異なるバンクへ分散するため高速アクセスが可能であるが、後者のライトバック時のアクセスに対しては考慮されていない。即ちライトバックにより同一バンクの異なるワード線へのアクセス（バンク競合状態）が生ずると高速アクセスがこんなんとなる。そのため、ライトバックが頻繁に発生するアプリケーション（プログラム）を実行する際に問題となる。

10

【0010】

そこで本発明の目的の一つは、キャッシュメモリを有する情報処理装置の主記憶装置（メインメモリ）にDRAMを用いる場合に、バンク競合状態を避けるために、アクセスの局所性とライトバック時のアクセスの規則性を取り入れて、これらのアクセスを高確率でDRAMの異なるバンクに割り当てることである。さらには、本願の他の目的はライトバック処理後の次のリードアクセス高速化することである。

【0011】

【課題を解決するための手段】

本願発明の代表的な手段を示せば以下の通りである。CPUからのリクエストアドレスをDRAMの各バンクに割り当てる際に、局所的なアクセスとライトバック時のアクセスを異なるバンクへ割り当てられるように、DRAMのバンクアドレスをINDEX部とTAG部の演算によってDRAMのバンクアドレスを生成することである。より具体的には、CPUの発するリクエストアドレスのINDEX部とTAG部から演算によってDRAMのバンクアドレスを生成するアドレス変換回路も設ける。INDEX部とTAG部からDRAMのバンクアドレスを生成するための典型的な演算は加算である。

20

【0012】

【発明の実施の形態】

<実施例1>

図1は本願の情報処理装置の基本概念を示す構成図である。これは、キャッシュメモリCMとCPU装置と、アドレスマッピング回路AMCを含むメモリコントローラMCONと、1以上のDRAMを含むメインメモリMMで構成される。CPUからのリクエストアドレスは、ADBUSにてCMおよびMCONへ伝達される。CMへ伝達されたリクエストアドレスはAMCにてDRAMアクセスのためのバンクアドレス、ロウアドレス、カラムアドレスに変換された後、ADBUSにてMMへ伝達される。ここではアドレス変換の方法を説明するため小規模なシステムを用いて説明する。そのため、キャッシュメモリの容量が128B、ブロックオフセットが4ビット（16B）、のダイレクトマップ型キャッシュメモリを想定して説明する。この例ではINDEXが3ビット、TAGが23ビットである。

30

40

【0013】

図4は本発明で用いられるメインメモリMMの構成例である。ここでは容量が512B（アドレスとしては9ビット必要）、構成バンクが2バンク（アドレスは1ビット必要）のDRAMでMMを構成すると考える。ここで、基本的なメモリアレイは、複数のワード線WLとビット線BLの交点に設けられたダイナミック型メモリセルMC（1個のトランスファーMOSFETと1個のキャパシタを持つ）を複数持ち、センスアンプSAはWLで選択されたメモリセルのデータを増幅する機能を持つ。SAにて増幅されたデータはBL上に保持され、切り替えスイッチYSによってグローバルビット線GBLにつながれてCPUへ伝達される。

【0014】

図5はアドレス変換の一実施例である。ここではリクエストアドレスRQADRがアドレス変

50

換回路（アドレスマッピング回路）AMCにてDRAMアクセスのアドレスDRADRへ変換される様子を説明する。AMCは論理回路LCを含む。リクエストアドレスRQADRの上に記した数字は、アドレス構成ビットの下位からの番号である。LINEはキャッシュメモリのブロックオフセットを示す。ここで後の便宜のため、このアドレスをDRAM容量を考慮して別の割り当て方を定義する。D-LINEはDRAMのワード線単位のブロックオフセットと定義し、D-TAGはTAGの下位からDRAM容量（ここでは9ビット）までと定義し、D-INDEXは、D-TAGと同数ビットで構成されるINDEXの一部と定義する。この例では、D-LINEはアドレスの下位からINDEXの下位まで割り当てている。これは、連続するアドレスがより多く同一のワード線へ集まるように割り当てているため、高速ページモード等を用いることで高速アクセスが可能である。また場合によってはD-LINEをLINEとINDEXの上位で構成することも可能である。このよ

10

【0015】

次に、AMCで変換をおこなうアドレスについて説明する。ここでは参照の局所性に起因するアクセス（局所アクセス）を同一ワード線へ集め、同一ワード線へ集めきれない局所アクセスとライトバック時のアクセスを異なるバンクへ割り当てる例を考える。これは、DRAMは同一のワード線上のデータには最初のアクセス以外は高速に行え、また、同一ワード線上にないデータでも、異なるバンクのワード線上のデータであれば、バンクを独立動作させることによって高速アクセスがおこなえるからである。したがって、ここで変換をおこなうアドレスは、D-INDEXとD-TAGである。これらのアクセスをできる限り異なるバンクへ

20

割り当てるアドレス変換方式は、排他論理和をD-INDEXの一部とD-TAGの一部に対しておこなうことで実現する。ここではDRAMのバンク数が2であることを考慮して、D-INDEXとD-TAGの下位1ビットを排他論理和で演算した結果をDRAMのバンクアドレスRBAに割り当て、D-TAGの上位1ビットとD-INDEXの2ビットをロウアドレスRRAへ割り当て、D-LINEの上位1ビットをカラムアドレスRCAへ割り当てる。このような変換を行う論理回路は排他論理和が唯一ではなく、たとえば加算器をもちいても達成できる。加算器を用いる場合は、

30

【0016】

図6はこのようなアドレス変換を行ってバンクアドレスを割り当てた結果である。横軸はD-INDEXの下位一ビットD-INDEX_L1（2進表記）であり、縦軸はD-TAGの下位一ビット（2進表記）D-TAG_L1である。図中の各柵内の数字はバンクアドレスである。本発明のマッピングをおこなうと、D-INDEXのみ変化する連続アクセスの一部に対しても、また同時に、INDEXが同じでTAGが異なるライトバック時のアクセスの一部についても異なるバンクに割り当てることが可能となる。そのため、空間的局所性を有するアクセスもライトバック時のアクセスも高い確率で異なるバンクへ割り当てられるので、高速アクセスが可能となる。

【0017】

<実施例2>

上述のようにINDEXが同じでTAGが異なるアドレスは、高い確率でバンク競合が起こらないようにDRAMの異なるバンクへ割り付けることが可能である。このアドレス割り付けを拡張すると、ライトバック時のリードとライトを完全に別バンクに割り当てることが可能となるので、ライトバック時のリードアクセスとライトアクセスを同一のサイクルで行うことが可能となる。以下でその実施例について述べる。

40

【0018】

図7は、ライトバック時（キャッシュメモリをミスした際にキャッシュ内容の置き換えが発生するアクセス）に発生するリードアクセスとライトアクセスを完全に別バンクに割り当て、それによってライトバック時のリードとライトを同時に行う発明の一実施例である

50

。これは、CPUと、キャッシュメモリCMと、DRAMを含むメインメモリMMと、MMを制御するメモリコントローラMCONと、MCONに含まれるアドレス変換回路AMCによって構成される。

【0019】

CPUからのリクエストアドレスは、ADBUSAにてCMに伝達されるとともにMCONへも伝達される。MCONへ伝達されたリクエストアドレスはAMCへ伝達され、DRAMのバンクアドレスRBA、ロウアドレスRRA、カラムアドレスRCAに変換される。また、ライトバック時のライトアドレスのTAGはCMよりWBADBUSBにてMCONへ伝達され、AMCにてDRAMのライトバック時のライト用バンクアドレスWBAに変換される。AMCで変換されたこれらのアドレスはADBUSBにてMMへ伝達される。なお詳しくは後述するが、このADBUSBはリクエストアドレスのバンクアドレスRBAを伝達アドレス線RBALと、ライトバック時のライトアドレスのバンクアドレスWBAを伝達するアドレス線WBALと、ロウアドレスRRA、カラムアドレスRCAを伝達するアドレス線RCLとで構成される。CPUからCMとMCONへの制御はCTBUSAにて行われる。MMへの制御は、MCONからCTBUSBにて行われる。CPUとCMおよびMMで構成されるメモリシステムへのデータ授受は、DABUSにて行われる。

10

【0020】

図8(A)はメインメモリの一実施例である。ここで考えるCMは容量が128B、ブロックオフセット(LINE)が4ビット(16B)、のダイレクトマップ型とし、MMは容量が512B(アドレスとしては9ビット必要)とする。この場合、D-TAGが2ビットとなるので、INDEXが同じでTAGが異なるアクセスを別バンクへ割り当てるためには、

バンク数 = 2のD-TAG乗 = 2の2乗 = 4

20

とすればよい。各バンクはRA0からRA3で示される4本のワード線で構成される。このDRAMの構成は、ライトバック時のリードアクセスとライトアクセスが完全に別バンクに割り当てられることが特徴である。なお、ワード線の上に記した数字は、後程説明する加算によってバンクアドレスを変換する場合のD-INDEXとD-TAG部分のアドレスである。

【0021】

ここで、各部の動作を説明する。RBAL、WBAL、RCLで構成されるADBUSBにて入力されたDRAMアクセスのアドレスはアドレスバッファADBUFで受け、その後ADBUSBにて各バンクへ伝達される。CTBUSBにて入力された制御信号は、制御部CTRへ入力された後、CTRからDRAM内の各部を制御する。DRAMへのデータ入出力はDABUSを用いておこない、DRAM内のI/OバッファIOBUFへ結線される。この実施例では、ライトバック時のリードとライトを同時に行うが、DABUSの構成線数を従来と同じとすることを特徴とする。したがって、このDABUSはリードとライトで共用する。また、ライトバック時のリードとライトを同時に行うため、本発明ではDRAM内部でのリードとライトのデータの伝達経路を独立化する。まずリードデータはリードデータ線RDLにて各バンクから直接このIOBUFへ伝達し、ライトデータはライトデータ線WDL1にてIOBUFからデータプリロードレジスタDPLREGへ伝達した後、ライトデータ線WDL2にて各バンクへ伝達する。このようにDPLREGを設置することにより、詳しくは後述する1ポートで擬似的な2ポートアクセスが可能となる。

30

【0022】

図8(B)はこのような構成のDRAMの1バンクの構成例である。記号の意味は図4と同じであり、ここではその構成数のみが異なる。

40

【0023】

図9はリクエストアドレスとライトバック時のライトアドレスのTAGアドレスをDRAMアクセスのアドレスへ変換するアドレス変換回路AMCを含むメモリコントローラMCONの一実施例である。リクエストアドレスRQADRは、AMCによってDRAMアクセスのアドレスDADR1に変換され、ライトバック時のアドレスWBADRはAMCによってDRAMアクセスのアドレスDADR2に変換される。また、外部からの制御線はCTBUSAにてMCON内の制御回路CTR1へ伝達され、DRAMアクセスの制御信号を生成し、DRAMへCTBUSBにて伝達する。

【0024】

ここで、AMCによるアドレス変換について説明する。まず、D-TAGとD-INDEXとD-LINEの一部がMCONへ伝達されMCON内部のAMCへ伝達される。AMCへ入力されたD-TAGとD-INDEXはAMC

50

内部のLC1へ伝達される。LC1は2つの排他論理和回路XOR1、XOR2で構成される。LC1へはD-INDEXの下位1ビットがLC1のノード1 N1へ、D-TAGの下位1ビットがLC1のノード2 N2へ、D-INDEXの上位1ビットがLC1のノード3 N3へ、D-TAGの上位1ビットがLC1のノード4 N4へ伝達される。LC1内では、N1とN2が排他論理和回路XOR2に入力され、演算結果がノード5 N5へ出力される。また、N3、N4が排他論理和回路XOR1に入力され、演算結果がノード6 N6へ出力される。N5はバンクアドレスRBAの下位1ビットに割り当てられ、N6はRBAの上位1ビットに割り当てられる。また、D-INDEXの下位1ビットはロウアドレスRRAの下位1ビットへ、D-INDEXの上位1ビットはロウアドレスRRAの上位1ビットへ割り当てられる。ここではキャッシュメモリへのデータ転送(キャッシュフィル)を32ビットづつ4回おこなう例を考
10

【0025】

次に、ライトバック時のライトアドレスWBADRの変換について説明する。AMCへ伝達されたWBADRのD-TAGとRQADRのD-INDEXとがLC2へ伝達される。LC2は2つの排他論理和回路XOR3、XOR4で構成される。ここでも、LC2へはD-INDEXの下位1ビットがLC1のノード7 N7へ、D-TAGの下位1ビットがLC1のノード8 N8へ、D-INDEXの上位1ビットがLC1のノード9 N9へ、D-TAGの上位1ビットがLC1のノード10 N10へ伝達される。N7とN8は排他論理和回路XOR4に入力され、演算結果がノード11 N11へ出力される。またN9、N10が排他論理和回路XOR3に入力され、演算結果がノード12 N12へ出力される。N11はWBAの下位1ビットに割り当てられ、N12はWBAの上位1ビットに割り当てられる。ライトバック時のライトアドレスのロウアドレスとカラムアドレスは、リクエストアドレスのINDEXとLINEと等しいので、リクエストア
20

【0026】

ここで得られたDRAMアクセスのアドレスRBA、WBA、RRA、RCAはタイミングコントローラTCRへ伝達され、その後RBAL、WBAL、RCLで構成されるADBUSBにてDRAMへ伝達される。DRAMへのアドレスの投入タイミングはこのTCRにて制御される。

【0027】

図10はリクエストアドレスとライトバック時のライトアドレスのTAGアドレスをDRAMアクセスのアドレスへ変換する図9のアドレス変換回路AMC内の論理回路LC1及びLC2の別の実施例である。ここでは、LC1、LC2に加算器を用いる例を示し、その構成と機能について説明する。LC1は1つの論理積回路AND1と3つの排他論理和回路XOR1、XOR2、XOR5にて構成
30

【0028】

図11はAMCによるバンクアドレスの変換結果である。この図で縦軸はD-TAG(2進表記)であり、横軸はD-INDEX(2進表記)であり、各桁目内はバンクアドレス(10進表記)である。(A)はAMCの論理回路LC1、LC2に排他論理和を用いた図9の場合のバンクアドレス変換結果であり、(B)はAMCの論理回路LC1、LC2に図10の加算器を用いた場合のバンクアドレス変換結果である。いずれの論理回路を用いても(A)、(B)に示すように、バンクアドレスは横軸方向も縦軸方向も全て異なる数字が割り当てられている。即ち、アドレス変換に必要な論理回路を一般化すれば、INDEXビットの一部のビットとTAGビットの一部のビットを受けてバンクアドレスビットを出力するものであって、INDEXビットの一部のビットを固定した値としてTAGビットの一部のビットを変化させた場合に互いに異なる前記バンクアドレスビットを形成し、かつTAGビットの一部のビットを固定した値としてINDEXビットの一部のビットを変化させた場合に互いに異なるバンクアドレスビットを出力するものである。このようにメインメモリへ行われる空間的局所性の高いアクセスを別バンクへ割り当て、同時に、ライトバック時のリードとライトのアクセスを完全に別バンクに割り当てることができる。
40

【0029】

図12は図8のDRAMに対して、ライトバック（書き戻しと読み出し）の後に読み出しを行う時のアクセスプロトコルを示すものである。図12（B）は本発明の擬似的な2ポートアクセスの動作を示す実施例である。ここで擬似的な2ポートアクセスとは、一つのメモリバンクに対する書き込みと他の一つのメモリバンクからの読み出しを時間的に並列に行うメモリアccessをさす。図12（A）は、比較のための擬似的2ポートアクセスを行わないでライトバックを行う例をしめしている。（A）、（B）ともに最上段はDRAMの動作クロック（CLK）を示しクロックの立ち上がりエッジに通し番号を付けた。次段はDRAMに入力されるアドレス（Address）を示し、3段目はCPU-DRAM間のデータバスの使用状況（Data）を示す。ここでは、ライトバック動作としてDRAMのバンク1にリードアクセスが、バンク2にライトアクセスが生じ、その後、バンク3にリードアクセスが発生した場合について記した。また、最下段は各バンクの使用状況を示している。なお、#の付いた数字は、アクセスの順番を表わしている。

10

【0030】

まず図12（A）のアクセス方法について述べる。DRAMへのリードアクセス要求が発生すると、所望のリードデータが格納されているバンクのワード線を立ち上げてメモリセル内のデータをセンスアンプで増幅する動作（バンクアクティブ動作）がおこなわれる。これは、バンクアドレスRBAとロウアドレスRRAをDRAM側へ伝達し、同時に、バンクアクティブコマンドACT#1が投入されることで行われる。所望のリードデータがセンスアンプにて増幅されると（バンクアクティブ状態）、カラムアドレスRCAとリードコマンドR#1をDRAMへ投入することによってセンスアンプに増幅されたデータの一部がデータバスに読み出される。その後のライトは、ライトのバンクアドレスWBAとロウアドレスRRAとバンクアクティブコマンドACT#2を投入することでバンクアクティブ動作を行い、バンクアクティブ状態になった後にライトコマンドW#2とライトデータを入力することでおこなう。ここではバンクアクティブコマンドを投入してからリード及びライトコマンドを投入するまで2クロックかかるとし、リードコマンド投入からデータ出力まで2クロック、ライトコマンド投入からライトデータ入力まで0クロックとした。さらにその後のバンク3に発生したリードアクセスはバンク1でのリードアクセスと同様に処理される。但し、ライトデータのバーストアクセスを妨げないために、最初のバンクアクティブコマンドACT#1投入後10クロック後に3番目のバンクアクティブコマンドACT#3を投入する。この時リードコマンドR#3はその後2クロック後に投入される。このように（A）では最初のバンクアクティブコマンドを投入してから2回目のリードデータが出力完了するまで17クロックかかる。

20

30

【0031】

一方、図12（B）に示す本発明のアクセス方法は、ライトバックが発生するとバンク1のリードアクセスを開始すると同時にバンク2のライトアクセスを開始する。このアクセスによりDRAMへ送られるコマンドとアドレスは、ライトバックコマンドWB#1とリードバンクアドレスRBAとライトバンクアドレスWBAとロウアドレスRRAである。ここでロウアドレスRRAがリードとライトで同一なのは、ロウアドレスがINDEXの一部であり、ライトバック時のリードとライトのアドレス間にはINDEXが同じでTAGが異なる性質があることを利用したものである。また同時に、データバスにはライトデータの転送が開始される。これはリードデータがバンクアクティブコマンドを投入後4クロック後に出力されることを考慮して、ライトデータをその待ち時間に入力するためである。ライトバックコマンドを投入後2クロック後にカラムコマンドC#1とカラムアドレスRCAを投入する。このカラムアドレスもリードとライトで同一である。それは、カラムアドレスがINDEXの一部とLINEの一部で構成されており、ライトバック時は、INDEXが同じでTAGが異なる同一LINEが置き換えられるからである。このようにすると、ライトバック時のリードとライトアクセスを同時おこなうことができる。このとき、ライトデータはライトバンクがアクティブとなる前に入力されるため、データプリロードレジスタDPLREGで一旦このライトデータを受ける。このライトデータは、ライトバンクが活性化した後に、カラムコマンドC#1とカラムアドレスRCAによって、DPLREGから該当バンクへ転送される。一方、バンク2ではこのC#1とRCAによってリードデータが出力される。このように、バンク1とバンク2にてライトバック時のリ

40

50

ードとライトが処理できるので、同一サイクルでライトバック時のリード動作とライト動作が終了できる。その後バンク3に発生するリードアクセスは、(A)のリードアクセスと同様の手続きでおこなわれる。この場合、バンク3へのリードアクセスが(A)に比べて早く開始できるため、ライトバックが発生してからバンク3のリードアクセスが終了するまで11サイクルで済む。これは(A)に比べて6サイクル高速化することになる。

【0032】

本願発明の疑似2ポートアクセスは、アドレスバスの最小限の追加で2つのメモリバンクへの並列アクセスできる点に特徴がある。「疑似」でない通常の2ポートアクセスとは、アドレスバスもデータバスも2組を使用することを意味する。即ち、書き込みと読み出しを時系列で進めではなく並列に行うには、通常は2倍のバスが必要となるが、ハードウェアが増加する点で好ましいものではない。本願の疑似2ポートとアクセスでは、並列アクセスを可能とするためにバンクアドレスだけは2組必要とするが、他のアドレスバスとデータバスを追加する必要がない点が利点である。即ち、ライトバック時には、バンクアドレスだけが異なり他のロウアドレスとコラムアドレスは同じ場所がアクセスされるため、バンクアドレス以外のアドレスバスの追加を必要としない。

10

【0033】

また、DRAMバンクを活性化してデータを読み出すには所定の時間がかかってしまう。この時間を利用して、ライトバックにおいてデータバスは時分割的に使い、データバスの増加を防止している。即ち、最初キャッシュからの書き戻しデータがデータバスによりDRAMに伝達されDRAMのプリロードレジスタDPLREGに取り込まれる。その後で他のメモリバンクからのデータがデータバスによりキャッシュに伝達される。従ってデータバスを時分割的に利用するために、DRAMには、バンクが活性化して書き込みが可能になる前に書き込みデータを取り込んでおくためのプリロードレジスタDPLREGを設けておけば良い。

20

【0034】

なお、バースト長を複数設定できるシンクロナスDRAM(SDRAM)のようなプロトコルでは、連続するデータの長さ(バースト長)を調節できる。このような方式では、バースト長分のデータが出力される前に別のリードコマンドを投入すると、以前のデータのバースト転送が遮断されてしまう。そこでキャッシュメモリの内容を置き換える目的のライトバック時は、必要となる転送データ長が決まっているため、バースト転送を遮断することなく次のリードコマンドを投入できるようにすれば、ライトバック時のリードデータが出力完了した後に直に次のリードデータが出力開始できるようになる。

30

【0035】

<実施例3>

図13は、図7の情報処理装置に対して本発明を現在用いられている典型的な容量を有するキャッシュメモリとDRAMを用いた場合に適用した他の実施例である。ここでは、キャッシュメモリとして、容量が512KB、ブロックオフセット(LINE)が32Bであるダイレクトマップ形式のライトバック方式を考える。この場合、LINEは5ビット[0:4]、INDEXは14ビット[5:18]、残りがTAG[19:31]となる。一方、メインメモリの総容量は128MBとし、容量が8MBのマルチバンク構成のDRAMで構成することを考える。まず、この128MBが1チップDRAMであると仮に考えれば、先に定義したD-TAGの概念を用いることができ、D-TAGは8ビットとなる。これにより、本システムで必要な独立バンク数は256と決まる。しかし、ここでは1チップDRAMの容量は8MBであり、このような大容量に対応していないため、16チップを用いてメインメモリを構成する。その結果、1つのDRAMチップのバンク数は16バンク、1バンクあたりのワード線数は、256本、一本のワード線下のメモリセル数は16384個とすれば実現される。

40

【0036】

ここでメモリコントローラMC内でのAMCによるアドレス変換について説明する。ここでは加算演算によりDRAMのバンク及びチップ番号を決定する方法を述べる。まずリクエストアドレスRQADRの変換について述べる。ここで加算演算するアドレスの部分はD-TAG[19:26]

50

とD-INDEX[5:12]である。このD-TAGとD-INDEXを加算器ADDER1へ入力することによって、その出力結果の下位4ビットをバンクアドレスRBA、上位4ビットをチップセレクトCS1へ変換する。このうちRBAはタイミングコントローラTCRへ伝達され、CSはRQADRの[27:31]と共にチップセレクト回路CSCへ伝達される。また、ロウアドレスRRAはD-INDEXを、カラムアドレスはRQADRの[2:3]と[13:18]を割り当てることで、DRAMへのアドレス変換が完了する。

【0037】

次にライトバック時のライトアドレスの変換について述べる。ライトバック時のライトアドレスは、キャッシュメモリからTAGのみMCONへ伝達し、これをライトバック時のライトアドレスWBADRとする。ここで加算演算するアドレスの部分はWBADRのD-TAG[19:26]とRQADRのD-INDEXである。これらを加算器ADDER2へ入力することで、その出力結果の下位4ビットをバンクアドレスWBAに上位4ビットをチップセレクトCS2へ変換する。このうちCS2はCSCへ伝達され、WBAはTCRへ伝達される。なお、ライトバック時のライトアクセスのロウアドレスとカラムアドレスは、アドレスのINDEXとLINEに相当するため、それぞれRRA、RCAに等しい。

【0038】

タイミングコントローラTCRは、上記AMCで変換されたRBA、WBA、RRA、RCAをDRAMへ発行するタイミングを計測する。ライトバック時には、TCRは第一のタイミングでRBA、WBA、RRAを発行し、第二のタイミングでRCAを発行する。このように、本発明のライトバック時の擬似2ポートメモリは、メインメモリが1チップに収まらない大容量の場合でもライトバック時のライトバンクアドレス線とチップセレクト信号線を増設するだけで実現され、情報の高速処理が可能となる。

【0039】

ところで、完全にライトバック時のリードとライトが同時処理されなくてもよいと考えられる場合には、独立バンク数を少なくして、最大限バンク競合が起こらないように設定できる。この例は、図5の場合のようにD-TAGとD-INDEXの一部を用いてバンクアドレスを演算すればよい。

【0040】

最後に、実施例1～3に示してきた情報処理装置の現実的な適用形態について説明する。図14(A)はCPUとキャッシュメモリCMとメモリコントローラMCONとメインメモリMCが同一の半導体基盤上にモノリシックに形成される実施例である。

【0041】

図14(B)はCPUとCMとMCONとが第1の半導体チップ上にモノリシックに形成され、MMが第2の半導体チップ上に形成される実施例である。図14(C)はCPUとCMとが同一の半導体チップ上にモノリシックに形成され、MCONが第2の半導体チップ上に形成され、MMが第3半導体チップ上に形成される実施例である。図14(D)はCPUが第1の半導体チップに形成され、CMが第2の半導体チップに形成され、MCONが第3の半導体チップに形成され、MMが第4の半導体チップ上に形成される実施例である。

【0042】

図14(C)～(D)のような実施例では、本願発明に特徴的なアドレス変換回路は、メモリコントローラMCON半導体チップ内に搭載されることで実施される。また図14(B)～(D)のように複数の半導体チップをからなる場合は、各半導体チップは各々レジン等で保護されたICパッケージに封止され、典型的にはガラスエポキシ等でできた配線基板上に半田付けされる。また複数の半導体チップがセラミック等でできた配線基板にボンディングワイヤ又は半田バンプで接続されるマルチチップモジュールの形態を取ってもよい。

【0043】

【発明の効果】

キャッシュメモリを有する計算機システムにおけるメインメモリアクセスは、参照の局所性と呼ばれるINDEXが変化するアクセスとライトバック時のINDEXが同じでTAGが異なるア

クセスがある。本発明はこれらのアクセスを別バンクへ割り当てることができるため、高速アクセスが可能となる。また、ライトバック時のリードアクセスとライトアクセスを完全に別バンクへ割り当てられれば、1つのポートを利用した擬似的な2ポートアクセスが可能となり、ライトバックのアクセスとその後のリードアクセスの高速化が実現される。

【図面の簡単な説明】

【図1】本発明の第1の実施例による情報処理装置を示す図。

【図2】キャッシュメモリの管理アドレスの例。

【図3】キャッシュメモリを含む情報処理装置の構成例。

【図4】図1の実施例に用いるメインメモリMMの構成例。

【図5】図1の実施例に用いるアドレス変換回路の一実施例。

【図6】図5のアドレス変換回路によるバンクアドレスの変換結果。

【図7】本発明の第2の実施例による情報処理装置を示す図。

【図8】図7の実施例に用いるメインメモリMMの構成例。

【図9】図7の実施例に用いるアドレス変換回路の一例。

【図10】図9の論理回路の他の構成例。

【図11】排他論理和(A)及び加算(B)によるバンクアドレスの変換結果。

【図12】図7の実施例のアクセスシーケンス。

【図13】図7の情報処理装置の他の適用例。

【図14】本発明の情報処理装置の具体的な実現方法を示す図。

【符号の説明】

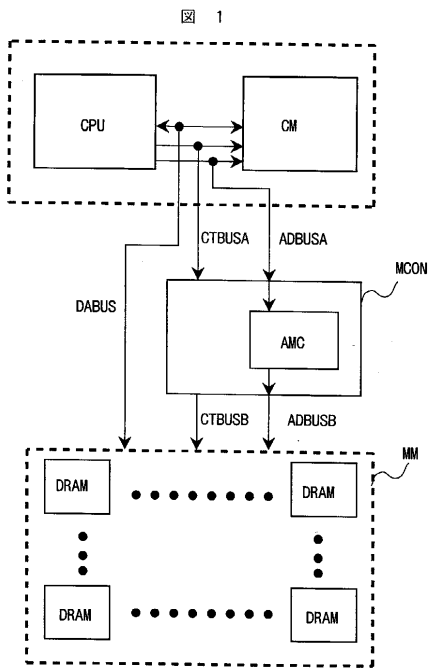
CPU...中央演算部、CM...キャッシュメモリ、MCON...メモリコントローラ、AMC...アドレスマッピング回路(アドレス変換回路)、MM...メインメモリ、DRAM...ダイナミック型ランダムアクセスメモリ、TAG(Tag)...アドレスのTAG部、INDEX(Index)...アドレスのINDEX部、D-TAG...DRAMの構成を考慮したTAGの一部、D-INDEX...DRAMの構成を考慮したINDEXの一部、SA...センスアンプ、MC...メモリセル、BL...ビット線、WL...ワード線、GBL...グローバルビット線、ADBUF...アドレスバッファ、DPLREG...データプリロードレジスタ、IOBUF...データI/Oバッファ、ADBUSA...アドレス線A、ADBUSB...アドレス線B、CTBUSA...制御線A、CTBUSB...制御線B、DABUS...データ線、RBA...リクエストバンクアドレス、RRA...リクエストロウアドレス、RCA...リクエストコラムアドレス、WBA...ライトバック時のライトバンクアドレス。

10

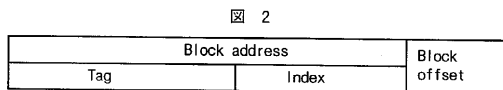
20

30

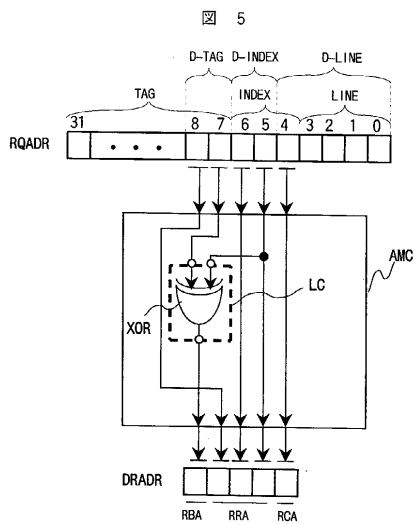
【 図 1 】



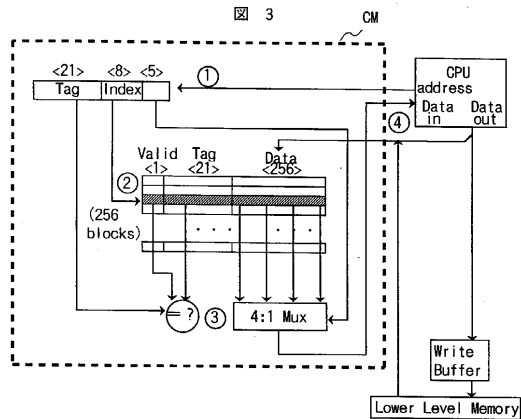
【 図 2 】



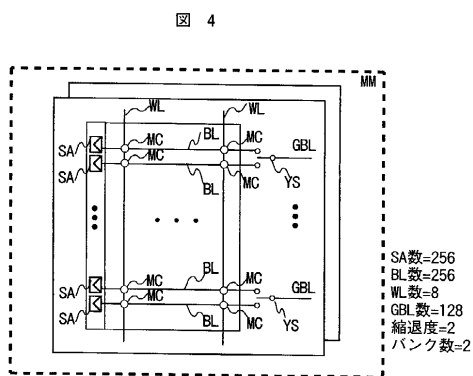
【 図 5 】



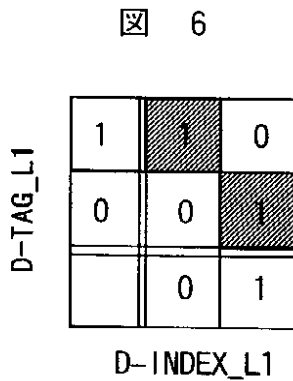
【 図 3 】



【 図 4 】



【 図 6 】



【 図 7 】

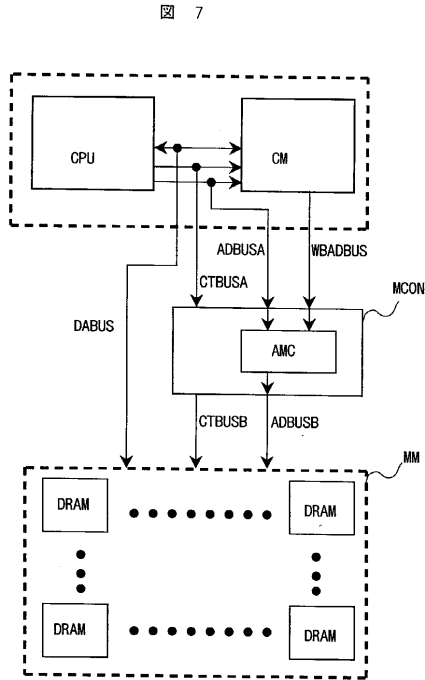


図 7

【 図 8 】

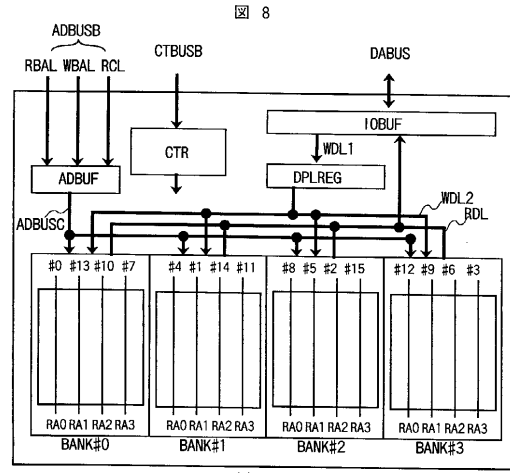
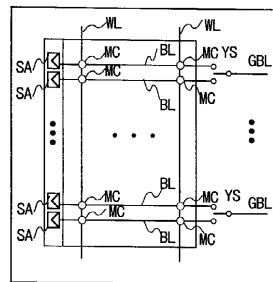


図 8

(A)



(B)

SA数=256
BL数=256
WL数=4
GBL数=128
縮退度=2

【 図 9 】

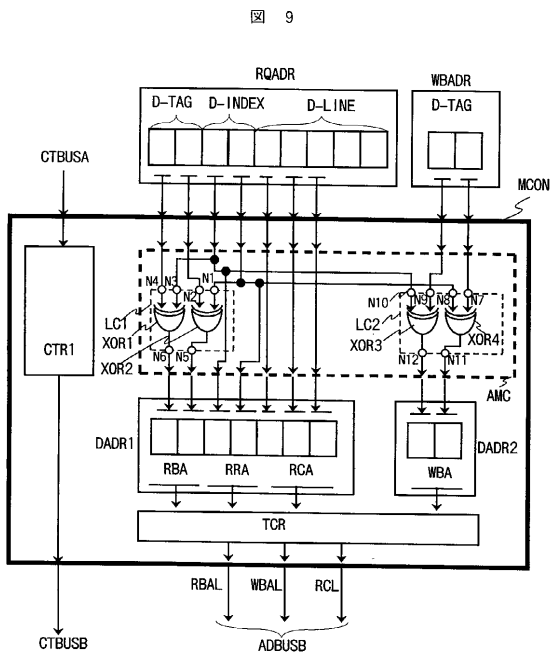


図 9

【 図 10 】

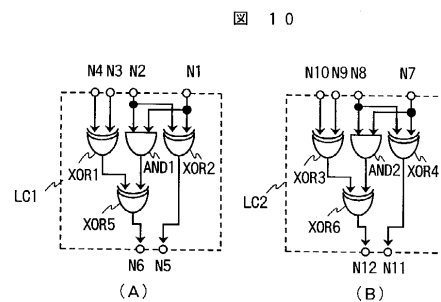


図 10

(A)

(B)

【 1 1 】

11

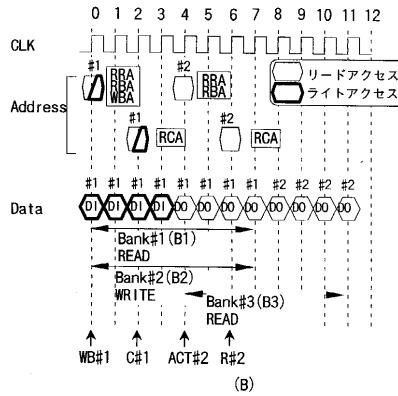
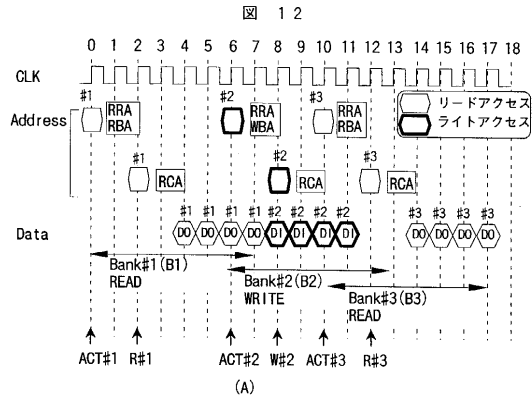
D-TAG	11	3	2	1	0
	10	2	3	0	1
	01	1	0	3	2
	00	0	1	2	3
		00	01	10	11

D-INDEX (A)

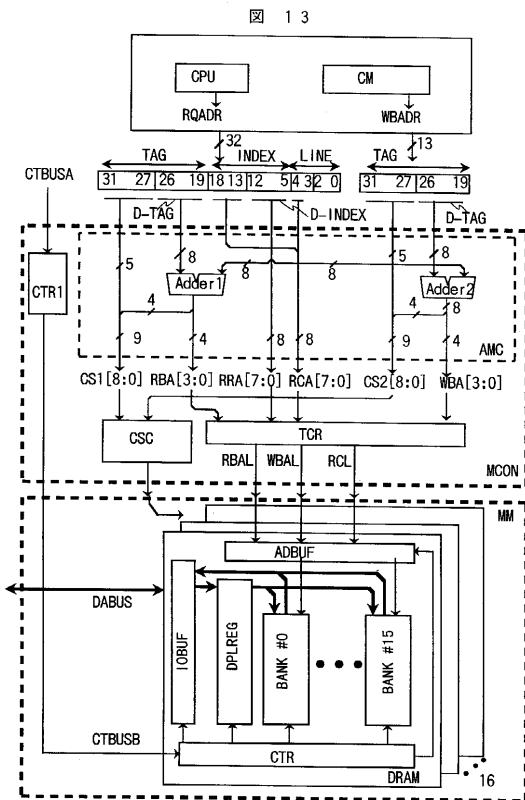
D-TAG	11	3	0	1	2
	10	2	3	0	1
	01	1	2	3	0
	00	0	1	2	3
		00	01	10	11

D-INDEX (B)

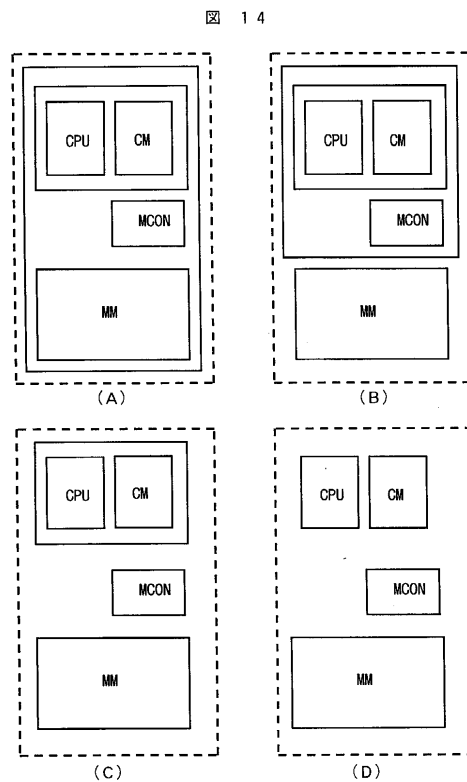
【 1 2 】



【 1 3 】



【 1 4 】



フロントページの続き

(51) Int.Cl. F I
G 0 6 F 12/06 5 4 0 C
G 0 6 F 12/06 5 2 5 B

(72) 発明者 渡部 隆夫
東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内

審査官 清木 泰

(56) 参考文献 特開昭56-114185(JP,A)
特開平06-095962(JP,A)
特開平11-265315(JP,A)
特開平09-288614(JP,A)
特開平04-233050(JP,A)
特開平10-031621(JP,A)
特開平07-175724(JP,A)
特開平04-069750(JP,A)
特開昭58-220287(JP,A)
特開昭56-019575(JP,A)
特開昭51-113429(JP,A)
特開2000-090005(JP,A)
特開2000-029782(JP,A)
特開2000-021160(JP,A)
特開2000-194596(JP,A)
米国特許第06070227(US,A)
Tomas Rokicki, Minimizing the Page Close Penalty: Indexing Memory Banks Revisited, HP Labs 1997 Technical Reports, Hewlett-Packard Development Company, 1997年, HPL-97-18R1

(58) 調査した分野(Int.Cl., DB名)

G06F12/08-12/12

G06F12/06