

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2016-85515
(P2016-85515A)

(43) 公開日 平成28年5月19日(2016.5.19)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 12/00 (2006.01)	G06F 12/00 571A	5B060
G06F 12/06 (2006.01)	G06F 12/00 570A	
	G06F 12/06 540E	

審査請求 有 請求項の数 11 O L (全 16 頁)

(21) 出願番号 特願2014-216196 (P2014-216196)
 (22) 出願日 平成26年10月23日 (2014.10.23)
 (11) 特許番号 特許第5911548号 (P5911548)
 (45) 特許公報発行日 平成28年4月27日 (2016.4.27)

(71) 出願人 390009531
 インターナショナル・ビジネス・マシーンズ・コーポレーション
 INTERNATIONAL BUSINESS MACHINES CORPORATION
 アメリカ合衆国10504 ニューヨーク州 アーモンク ニュー オーチャードロード
 New Orchard Road, Armonk, New York 10504, United States of America

(74) 代理人 100108501
 弁理士 上野 剛史

最終頁に続く

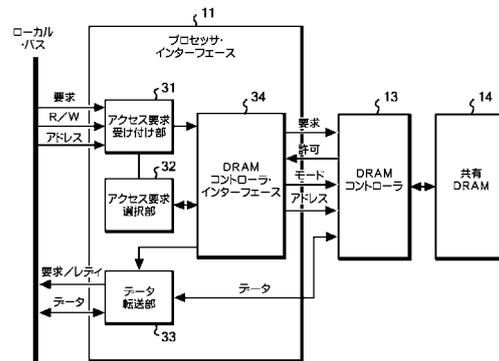
(54) 【発明の名称】 共有メモリへのアクセス要求をスケジューリングするための装置、方法およびコンピュータプログラム

(57) 【要約】

【課題】 共有メモリに対する複数のアクセス要求をうまくスケジューリングすることによって、ターンアラウンドタイムおよびバス利用効率を改善すること

【解決手段】 本発明に従うスケジューリング装置は、各要求元からのアクセス要求を受け付けるアクセス要求受け付け部と、アクセス要求受け付け部が受け付けたアクセス要求のうち、アクセス要求選択の基準となる最初のアクセス要求を選択し、この最初のアクセス要求に対して、BIモードで転送可能なアクセス要求を選択し、BIモードで転送可能なアクセス要求がない場合、または先行アクセス要求がBIモードもしくはCNモードの場合に、CNモードで転送可能なアクセス要求を選択する、アクセス要求選択部とを含み、アクセス要求選択部は、CNモードで転送可能なアクセス要求がない場合に、まだ選択されていないアクセス要求を対象に、最初のアクセス要求、BIモードで転送可能なアクセス要求、およびCNモードで転送可能なアクセス要求の選択を繰り返すように構成される。

【選択図】 図3



【特許請求の範囲】**【請求項 1】**

複数の要求元によって共有される共有メモリに対して、各要求元からのアクセス要求をスケジューリングするための装置であって、

前記各要求元からのアクセス要求を受け付けるアクセス要求受け付け部と、

前記アクセス要求受け付け部が受け付けたアクセス要求のうち、アクセス要求選択の基準となる最初のアクセス要求を選択し、前記最初のアクセス要求に対して、バンク・インターリーブ・モード（以下、BIモードという）で転送可能なアクセス要求を選択し、前記BIモードで転送可能なアクセス要求がない場合、または先行アクセス要求がBIモードもしくは連続リード/ライト・モード（以下、CNモードという）の場合に、CNモードで転送可能なアクセス要求を選択する、アクセス要求選択部と、

10

を含み、

前記アクセス要求選択部は、前記CNモードで転送可能なアクセス要求がない場合に、まだ選択されていないアクセス要求を対象に、前記最初のアクセス要求、前記BIモードで転送可能なアクセス要求、および前記CNモードで転送可能なアクセス要求の選択を繰り返す、スケジューリング装置。

【請求項 2】

前記アクセス要求受け付け部は、前記各要求元からのアクセス要求を到着順にFIFOバッファに保管し、

前記アクセス要求選択部は、前記FIFOバッファに保管されているアクセス要求のうち、最も待ち時間の長いアクセス要求を前記最初のアクセス要求として選択する、

20

請求項 1 に記載のスケジューリング装置。

【請求項 3】

前記アクセス要求選択部は、前記FIFOバッファに保管されているアクセス要求を待ち時間の長いものから順に走査することにより、前記BIモードまたは前記CNモードで転送可能なアクセス要求があるかどうかを判断する、請求項 2 に記載のスケジューリング装置。

【請求項 4】

前記スケジューリング装置に対してあらかじめ決められたアクセスタイムが割り当てられており、前記アクセス要求選択部は、前記アクセスタイムの制限内で、前記最初のアクセス要求、前記BIモードで転送可能なアクセス要求、および前記CNモードで転送可能なアクセス要求の選択を繰り返す、請求項 1 乃至 3 のいずれか 1 項に記載のスケジューリング装置。

30

【請求項 5】

前記アクセス要求選択部は、アクセス要求の選択のたびに、選択したアクセス要求のプロトコルオーバーヘッドを累算し、その累算値を前記アクセスタイムと比較することによって前記累算値が前記アクセスタイムの制限を超えたかどうかを判断する、請求項 4 に記載のスケジューリング装置。

【請求項 6】

複数の要求元によって共有されるメモリに対して、各要求元からのアクセス要求をスケジューリングするための方法であって、

40

前記各要求元からのアクセス要求を受け付けるステップと、

前記受け付けるステップで受け付けたアクセス要求のうち、アクセス要求選択の基準となる最初のアクセス要求を選択するステップと、

前記最初のアクセス要求に対して、BIモードで転送可能なアクセス要求を選択するステップと、

前記BIモードで転送可能なアクセス要求がない場合、または先行アクセス要求がBIモードもしくはCNモードの場合に、CNモードで転送可能なアクセス要求を選択するステップと、

を含み、

50

前記 B I モードで転送可能なアクセス要求または前記 C N モードで転送可能なアクセス要求がない場合に、まだ選択されていないアクセス要求を対象に、前記最初のアクセス要求を選択するステップ、前記 B I モードで転送可能なアクセス要求を選択するステップ、および前記 C N モードで転送可能なアクセス要求を選択するステップを繰り返す、スケジューリング方法。

【請求項 7】

前記アクセス要求を受け付けるステップは、前記各要求元からのアクセス要求を到着順に F I F O バッファに保管し、

前記最初のアクセス要求を選択するステップは、前記 F I F O バッファに保管されているアクセス要求のうち、最も待ち時間の長いアクセス要求を前記最初のアクセス要求として選択する、

10

請求項 6 に記載のスケジューリング方法。

【請求項 8】

前記 B I モードで転送可能なアクセス要求を選択するステップおよび前記 C N モードで転送可能なアクセス要求を選択するステップは、前記 F I F O バッファに保管されているアクセス要求を待ち時間の長いものから順に走査することにより、前記 B I モードで転送可能なアクセス要求があるかどうか、および前記 C N モードで転送可能なアクセス要求があるかどうかをそれぞれ判断する、請求項 7 に記載のスケジューリング方法。

【請求項 9】

当該スケジューリング方法を実行する装置に対してあらかじめ決められたアクセスタイムが割り当てられており、前記最初のアクセス要求を選択するステップ、前記 B I モードで転送可能なアクセス要求を選択するステップ、および前記 C N モードで転送可能なアクセス要求を選択するステップは、前記アクセスタイムの制限内で、前記最初のアクセス要求、前記 B I モードで転送可能なアクセス要求、および前記 C N モードで転送可能なアクセス要求の選択をそれぞれ繰り返す、請求項 6 乃至 8 のいずれか 1 項にスケジューリング方法。

20

【請求項 10】

前記最初のアクセス要求を選択するステップ、前記 B I モードで転送可能なアクセス要求を選択するステップ、および前記 C N モードで転送可能なアクセス要求を選択するステップは、アクセス要求の選択のたびに、選択したアクセス要求のプロトコルオーバーヘッドを累算し、その累算値を前記アクセスタイムと比較することによって前記累算値が前記アクセスタイムの制限を超えたかどうかを判断する、請求項 9 に記載のスケジューリング方法。

30

【請求項 11】

複数の要求元によって共有されるメモリに対して、各要求元からのアクセス要求をスケジューリングするためのコンピュータプログラムであって、請求項 6 乃至 10 のいずれか 1 項に記載のスケジューリング方法の各ステップをコンピュータに実行させるためのコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

40

【0001】

本発明は、共有メモリのアクセス制御に関し、特に、複数の要求元によって共有されるメモリに対して、各要求元からのアクセス要求をスケジューリングするための装置、方法およびコンピュータプログラムに関する。

【背景技術】

【0002】

コンピュータやストレージデバイスを含む様々な情報処理装置では、電力消費およびコストの削減が重要な課題になっている。例えば、最近のテープドライブは、プロセッサ (CPU) を含む複数の装置によって、D R A M のような外部メモリを共有する構成を採用している。メモリを共有すると、各装置が固有のメモリを持つ場合に比べて、メモリのチ

50

ップ数を減らすことができ、ひいては、電力消費およびコストの削減や、回路ボードの縮小に役立つ。

【0003】

しかしながら、共有メモリシステムを使用すると、プロセッサ等の要求元のメモリアクセスのターンアラウンドタイムが長くなって、パフォーマンスに悪影響を及ぼすことがある。パフォーマンスは、共有メモリシステムのバスの利用効率が悪い場合にも低下する。従って、共有メモリシステムにおいては、ターンアラウンドタイムの短縮と、バス利用率の改善が急務になっている。

【0004】

プロセッサからのアクセス要求が外部DRAMに送られる場合、アクセス要求のターンアラウンドタイムは、少なくともDRAMのプロトコルオーバーヘッド（アクセス対象のアドレスが有効化されてから、アクセスが終わって無効化されるまでの時間）に依存する。また、最初のアクセス要求が処理されているときに、第2のプロセッサが別のアクセス要求を出すと、第2のプロセッサは最初のアクセス要求の処理が終わるまで、自身のアクセス要求が処理されるのを待たなければならない。これは、第2のプロセッサのターンアラウンドタイムを増加させる。共有メモリシステムにおいてターンアラウンドタイムを改善するための手法として、バンクインターリーブモード（以下、BIモードという）および連続リード/ライトモード（以下、CNモードという）が知られている。

【0005】

BIモードでは、例えばアクティブコマンドによって、DRAMの複数のバンクを同時に開く、すなわち有効化することができる。コントローラは、異なるバンクアドレスを持つ複数のアクセス要求を、それらの有効化された複数のバンクに対してインターリーブ式に送り、それによりターンアラウンドタイムを短くすることができる。

【0006】

CNモードでは、コントローラは、前のアクセス要求で指定されたバンクアドレスおよびロウアドレスと同じバンクアドレスおよびロウアドレスを持つライトコマンドまたはリードコマンドを発行することによって、アクセスサイクルを連続させることができ、それによりプロトコルオーバーヘッドおよびターンアラウンドタイムの短縮が可能になる。

【0007】

BIモードおよびCNモードは、プロトコルオーバーヘッドの低減およびDRAMバス利用率の改善に寄与するが、アクセス要求のアドレスが、BIモードおよびCNモードの条件を満たさない場合は、ノーマルモードの転送が行われることになる。ノーマルモードでは、リードまたはライトのコマンド毎に、アドレスの有効化および無効化が実行される。

【0008】

DRAMとしてDDR3 SDRAMを用いた場合に、2つのリードコマンドを連続的に処理する場合のプロトコルオーバーヘッドの例を図1に示す。図1中の1文字のアルファベットは、下記非特許文献1に記載のコマンドを表しており、Aはアクティブコマンド、Rはリードコマンド、Pはプリチャージコマンドである。また、DQはデータ信号、DQSはデータストロブ信号を表している。

【0009】

(A) ノーマル転送

ノーマル転送では、リードコマンドRごとに、アクティブコマンドAでバンクアドレス及びロウアドレスの有効化を開始してから、プリチャージコマンドPでそれらの無効化を終了するまでに26クロックを要し、従って、2つのリードコマンドRを連続的に処理する場合のプロトコルオーバーヘッドは52クロックである。

【0010】

(B) BIモード

BIモードでは、2つのバンクからのリードが実行されるので、まずそれらをアクティブコマンドAで有効化した後2つのリードが連続的に実行され、最後にプリチャージコマ

10

20

30

40

50

ンドが1度だけ実行される。トータルのプロトコルオーバーヘッドは35クロックである。

【0011】

(c) ノーマル + CNモード

CNモードの転送は、先行コマンド（ここではノーマルモードのリードコマンド）と同じバンクアドレス及びロウアドレスに対して行われるので、アクティブコマンドAによるそれらのアドレスの有効化は最初だけでよく、プリチャージコマンドも最後に1度だけ実行すればよいので、プロトコルオーバーヘッドは最短の28クロックである。プロトコルオーバーヘッドだけを考えれば、CNモードがベストということになる。

【0012】

下記の特許文献1には、1回のバスサイクルで複数のアクセス要求を処理するようにしたメモリアクセス装置が記載されている。このメモリアクセス装置は、メモリをメインメモリとして使用する複数のCPUと、メモリをバッファとして使用する他の機能ブロックとに接続され、メモリへの複数のCPUからのアクセス転送を制御するCPUインターフェースと、メモリへのアクセス転送の調停を行うDRAMコントローラとを含んでいる。CPUインターフェースは、複数のCPUからのアクセス要求を待機させ、かつ各アクセスのアドレス、データ転送モード及びデータサイズを受取り保管し、DRAMコントローラにアクセス要求を知らせ、アクセス要求に対して許可信号を受けた場合に、許可信号に応じてDRAMコントローラに情報を送り、DRAMコントローラは、アクセス要求信号を受取り、アクセス調停に基づいて転送を許可されたCPUを指定して許可信号をCPUインターフェースに送るようになっている。

【0013】

また、下記の特許文献2には、共有メモリシステムを使用する複数の要求元の待ち時間を均等にするための手法が記載されており、それによれば、複数の要求元からのアクセス要求のうち、最も待ち時間の長い要求を選択して、その最長待ち時間要求を他のアクセス要求の後で共有メモリシステムに送り、最長待ち時間要求を出した要求元は、許可された最長待ち時間要求に続けて追加のアクセス要求を共有メモリシステムへ送ることができるようになっている。

【先行技術文献】

【特許文献】

【0014】

【特許文献1】米国特許出願公開第2014/0059286A1

【特許文献2】米国特許出願公開第2013/0179645A1

【非特許文献】

【0015】

【非特許文献1】DDR3 SDRAM STANDARD, JESD79-3F, July 2012 (<http://www.jedec.org/standards-documents/docs/jesd-79-3d>)

【発明の概要】

【発明が解決しようとする課題】

【0016】

共有メモリシステムのパフォーマンスの面からは、バス利用効率、すなわち、共有メモリに対するインターフェースの帯域幅（単位時間あたりの転送バイト数）を増大することが重要であり、そのためには、各インターフェースに割り当てられている許容可能なアクセスタイムの制限内で複数のアクセス要求のバースト転送サイズを最大化する必要がある。しかし、特許文献2では、転送の先頭でBIモードができない場合、共有メモリアクセスの要求元であるN台のプロセッサからのアクセス要求がノーマル転送されるため、オーバーヘッドが大きい。また、特許文献2では、1つの転送サイクルで転送可能なアクセス要求の数が制限されているが、上述のように、転送モードによってオーバーヘッドは異なるため、割り当て済みの許容可能なアクセスタイムに余裕がある場合にも転送が打ち切られることがある。従って、特許文献2に記載の手法はバースト転送サイズの最大化には向

10

20

30

40

50

いていない。

【 0 0 1 7 】

従って、本発明の目的は、共有メモリに対する複数のアクセス要求をうまくスケジューリングすることによって、ターンアラウンドタイムおよびバス利用効率を改善することにある。

【 課題を解決するための手段 】

【 0 0 1 8 】

本発明の第 1 の態様は、複数の要求元によって共有されるメモリに対して、各要求元からのアクセス要求をスケジューリングするための装置を提供し、該装置は、

前記各要求元からのアクセス要求を受け付けるアクセス要求受け付け部と、

前記アクセス要求受け付け部が受け付けたアクセス要求のうち、アクセス要求選択の基準となる最初のアクセス要求を選択し、前記最初のアクセス要求に対して、B I モードで転送可能なアクセス要求を選択し、前記 B I モードで転送可能なアクセス要求がない場合、または先行アクセス要求が B I モードもしくは C N モードの場合に、C N モードで転送可能なアクセス要求を選択する、アクセス要求選択部と、

を含み、

前記アクセス要求選択部は、前記 B I モードで転送可能なアクセス要求または前記 C N モードで転送可能なアクセス要求がない場合に、まだ選択されていないアクセス要求を対象に、前記最初のアクセス要求、前記 B I モードで転送可能なアクセス要求、および前記 C N モードで転送可能なアクセス要求の選択を繰り返すように構成される。

【 0 0 1 9 】

本発明の第 2 の態様は、複数の要求元によって共有されるメモリに対して、各要求元からのアクセス要求をスケジューリングするための方法を提供し、該方法は、

前記各要求元からのアクセス要求を受け付けるステップと、

前記受け付けるステップで受け付けたアクセス要求のうち、アクセス要求選択の基準となる最初のアクセス要求を選択するステップと、

前記最初のアクセス要求に対して、B I モードで転送可能なアクセス要求を選択するステップと、

前記 B I モードで転送可能なアクセス要求がない場合、または先行アクセス要求が B I モードもしくは C N モードの場合に、C N モードで転送可能なアクセス要求を選択するステップと、

を含み、

前記 B I モードで転送可能なアクセス要求または前記 C N モードで転送可能なアクセス要求がない場合に、まだ選択されていないアクセス要求を対象に、前記最初のアクセス要求を選択するステップ、前記 B I モードで転送可能なアクセス要求を選択するステップ、および前記 C N モードで転送可能なアクセス要求を選択するステップを繰り返すように構成される。

【 0 0 2 0 】

本発明の第 3 の態様は、複数の要求元によって共有されるメモリに対して、各要求元からのアクセス要求をスケジューリングするためのコンピュータプログラムを提供し、該コンピュータプログラムは、コンピュータに前記第 2 の態様に係る方法の各ステップを実行させる。

【 0 0 2 1 】

いずれの態様においても、アクセス要求選択の基準となる最初のアクセス要求は、F I F O バッファに保管されているアクセス要求のうち、最も待ち時間の長いアクセス要求であってもよい。

【 0 0 2 2 】

また、B I モードで転送可能なアクセス要求の選択、および C N モードで転送可能なアクセス要求の選択は、F I F O バッファに保管されているアクセス要求を待ち時間の長いものから順に走査することにより、B I モードで転送可能なアクセス要求があるかどうか

、およびCNモードで転送可能なアクセス要求があるかどうかをそれぞれ判断するようにしてもよい。

【0023】

また、スケジューリング装置に対してあらかじめ決められたアクセスタイムを割り当て、最初のアクセス要求、BIモードで転送可能なアクセス要求、およびCNモードで転送可能なアクセス要求を選択するときは、このアクセスタイムの制限内で、それぞれのアクセス要求の選択を繰り返すようにしてもよい。

【0024】

また、最初のアクセス要求、BIモードで転送可能なアクセス要求、およびCNモードで転送可能なアクセス要求を選択するときは、アクセス要求の選択のたびに、選択したアクセス要求のプロトコルオーバーヘッドを累算し、その累算値を割り当て済みのアクセスタイムと比較することによって累算値がアクセスタイムの制限を超えたかどうかを判断するようにしてもよい。

【図面の簡単な説明】

【0025】

【図1】2つのリードコマンドをノーマルモード、BIモードおよびCNモードで処理したときのプロトコルオーバーヘッドを示す図。

【図2】本発明に従うスケジューリング装置を含む共有メモリシステムの構成例を示すブロック図。

【図3】本発明に従うスケジューリング装置として機能するプロセッサインターフェースの詳細を示すブロック図。

【図4】プロセッサインターフェースのアクセス要求受け付け部に含まれる要求FIFOのエントリの一例を示すブロック図

【図5】要求FIFOの内容の具体例を示す図。

【図6】本発明に従うスケジューリング方法の実施形態を示すフローチャート。

【図7】図6の実施形態におけるアクセス要求を選択するステップの詳細を示すフローチャート。

【図8】図7の実施形態においてアクセス要求が選択されるたびに実行される動作を示すフローチャート。

【図9】DRAMコントローラへ送るために本発明の実施態様に従って作成されたアクセス要求の転送パッケージの一例を示す図。

【発明を実施するための形態】

【0026】

本発明に従うスケジューリング装置を含む共有メモリシステムの一構成例を図2に示す。図2において、本発明に従うスケジューリング装置として機能するのはプロセッサインターフェース11である。プロセッサインターフェース11には、それぞれのローカルバスを介して複数のプロセッサ12（プロセッサ1、プロセッサ2・・・プロセッサm）が接続されており、これらのプロセッサ12は共有DRAM14に対する要求元となる。プロセッサインターフェース11は、以下で説明するスケジューリングにより、各プロセッサ12からのアクセス要求を1つの転送パッケージにまとめて、DRAMコントローラ13に送る。DRAMコントローラ13は、この転送パッケージに含まれるアクセス要求（リードまたはライト）に従って共有DRAM14をアクセスする。

【0027】

プロセッサインターフェース11の詳細を図3に示す。図示のように、プロセッサインターフェース11は、各プロセッサからローカルバスを介してアクセス要求を受け付けるアクセス要求受け付け部31と、アクセス要求受け付け部31が受け付けたアクセス要求を所定の基準に従って選択し、1つの転送パッケージにまとめるアクセス要求選択部32と、プロセッサローカルバスとDRAMコントローラ13との間でリードデータまたはライトデータを転送するデータ転送部33と、DRAMコントローラ13へのインターフェースを提供するDRAMコントローラインターフェース34とを含んでいる。

【 0 0 2 8 】

アクセス要求受け付け部 3 1 は、ローカルバスの要求ライン、R / Wラインおよびアドレスラインに接続されており、アクセス要求を持つプロセッサが要求ラインを上げたことに応答して、リードまたはライトを示す R / Wライン上の信号と、アドレスライン上のアドレス信号、すなわち共有 D R A M 1 4 のバンクアドレス、ロウアドレスおよびコラムアドレスを保管する。本実施形態では、アクセス要求受け付け部 3 1 は F I F O バッファ（以下、要求 F I F O という）を含み、そのエントリの一例を図 4 に示す。

【 0 0 2 9 】

図 4 に示すように、要求 F I F O の各エントリは、アクセス要求を出したプロセッサを識別するプロセッサ I D と、上述のアドレス信号および R / W 信号とを含む。アクセス要求を出したプロセッサは、ローカルバスから識別することができる。すなわち、図 2 に示すように、各プロセッサ 1 2 には専用のローカルバスが接続されているので、アクセス要求受け付け部 3 1 は、要求ラインが上がったローカルバスにより、アクセス要求を出したプロセッサを識別することができる。アクセス要求受け付け部 3 1 は、識別したプロセッサの I D（図 2 に示すように、各プロセッサの番号 1、2・・・m でもよい）を要求 F I F O に保管する。アクセス要求受け付け部 3 1 は、プロセッサ 1 2 からのアクセス要求を受け付けると、そのことを D R A M コントローラインターフェース 3 4 に知らせる。

【 0 0 3 0 】

アクセス要求選択部 3 2 は、後で詳しく説明するように、D R A M コントローラ 1 3 によって許可されると、アクセス要求受け付け部 3 1 の要求 F I F O に保管されているアクセス要求から、所定の基準に従ってアクセス要求を選択し、それらのアクセス要求を含む転送パッケージを作成して、D R A M コントローラインターフェース 3 4 へ送る。

【 0 0 3 1 】

データ転送部 3 3 は、共有 D R A M 1 4 から読み出したリードデータ及び共有 D R A M 1 4 に書き込むライトデータを一時的に保管して、リードデータを要求元のプロセッサに転送し、ライトデータを D R A M コントローラ 1 3 へ転送する回路で、本実施形態においては、アクセス要求受け付け部 3 1 と同様に、F I F O バッファを使用して、それらのデータを保管する。図には示していないが、この F I F O バッファは、リードデータ用の F I F O バッファ（以下、リード F I F O という）およびライトデータ用の F I F O バッファ（以下、ライト F I F O という）を含む。

【 0 0 3 2 】

D R A M コントローラインターフェース 3 4 は、D R A M コントローラ 1 3 へのインターフェースを提供するもので、アクセス要求受け付け部 3 1 からアクセス要求の受け付けを通知されると、D R A M コントローラ 1 3 に要求信号を送り、D R A M コントローラ 1 3 からの許可信号を待つ。D R A M コントローラ 1 3 は、共有 D R A M 1 4 のアクセスが可能であれば、D R A M コントローラインターフェース 3 4 に許可信号を返す。D R A M コントローラインターフェース 3 4 は、許可信号を受け取ると、アクセス要求選択部 3 2 にアクセス要求の選択および転送パッケージの作成を開始させ、そしてアクセス要求選択部 3 2 からの転送パッケージに含まれる各アクセス要求に応じて、必要なモード信号およびアドレス信号を D R A M コントローラ 1 3 に送る。モード信号には、上述の R / W 信号の他に、B I モードまたは C N モードを示す信号も含まれる。D R A M コントローラ 1 3 は、これらの信号に応じて、共有 D R A M のアクセスを実行する。

【 0 0 3 3 】

図 3 に示す各構成要素は、アクセス要求選択部 3 2 を除き、上記の特許文献 1 に記載の対応する構成要素と同じものでもよい。

【 0 0 3 4 】

なお、図 2 および図 3 には示していないが、D R A M コントローラ 1 3 は、プロセッサインターフェース 1 1 の他にも、共有 D R A M 1 4 を使用する他の機能装置（例えば、テープドライブでは、サーボ回路、エラー訂正回路などが共有 D R A M を使用する）のインターフェースに接続することも可能である。その場合、D R A M コントローラ 1 3 は、特

10

20

30

40

50

許文献 1 にも記載されているように、各インターフェースからの要求を調停するためのアービタを含み、このアービタにより許可されたインターフェースに許可信号を送る。DRAMコントローラ 13 に複数のインターフェースが接続されている構成では、一般に、あらかじめ決められた許容可能なアクセスタイムが各インターフェースに割り当てられ、各インターフェースは、その制限内で複数のアクセス要求を時分割的に DRAM コントローラ 13 に送ることができる。

【0035】

次に、アクセス要求受け付け部 31 に含まれる要求 FIFO の具体的な内容を示す図 5 と、図 6 ~ 図 8 のフローチャートを参照しながら、本発明に従うスケジューリング装置、すなわちプロセッサインターフェース 11 の動作について説明する。なお、図 6 ~ 図 8 のフローチャートにおいては、「アクセス要求」を単に「要求」と表記している。

10

【0036】

図 5 に示す要求 FIFO は、各プロセッサ 12 からのアクセス要求に関する情報、すなわちプロセッサ ID、アドレスおよびリード/ライト (R/W) 信号を、アクセス要求の到着順に、上から下に向かって順に保管するよう構成されている。「プロセッサ ID」の欄に示されている $P_i(j)$ は、 i 番目 ($i = 1, 2, \dots, m$) のプロセッサ P_i からの j 番目 ($j = 1, 2, \dots, n$) のアクセス要求に係るプロセッサ ID である。なお、図 5 の要求 FIFO の場合は、これらの情報がアクセス要求の到着順に上から順に保管されるので、アクセス要求の順番を示す情報 (j) を保管しておかなくても、その順番を特定することは可能である。以下では、説明の便宜上、「プロセッサ ID」の欄に示されている $P_i(j)$ がアクセス要求も識別するものとする。

20

【0037】

「アドレス」の欄に示されている Bank はバンクアドレス、Row はロウアドレスを表している。実際には、共有 DRAM 14 はこれら 2 つのアドレスと、カラムアドレスとを用いてアクセスされるが、後述のように、本発明は、これら 3 つのアドレスのうち、バンクアドレスおよびロウアドレスだけを用いて、アクセス要求の転送をスケジューリングするので、図 5 ではカラムアドレスの表記を省略している。

【0038】

「R/W」の欄に示されている R は、アクセス要求がリードであることを表し、W は、アクセス要求がライトであることを表している。

30

【0039】

図 5 は、最初にプロセッサ 1 (P_1) が Bank 1、Row 10 に対してリードのアクセス要求を出したことを示す。アクセス要求受け付け部 31 は、図 6 のステップ S61 でこのアクセス要求 $P_1(1)$ を受け付け、そのことを DRAM コントローラインターフェース 34 に通知する。DRAM コントローラインターフェース 34 は、この通知にตอบสนองして、DRAM コントローラ 13 に要求信号を送り、許可されるのを待つ (ステップ S62)。DRAM コントローラ 13 は、要求信号を受け取ったとき、共有 DRAM がアクセス中でなければ、DRAM コントローラインターフェース 34 に許可信号を送るが、アクセス中の場合は、アクセスが終了するまで許可信号の返送を遅らせる。その間、アクセス要求受け付け部 31 の要求 FIFO には、後続のアクセス要求が到着順に保管され続ける (S62 からの「いいえ」のパス)。ここでは、DRAM コントローラ 13 によって許可されるまでに、プロセッサ m (P_m) の n 番目のアクセス要求 $P_m(n)$ までが要求 FIFO に保管されているものとする (図 5 参照)。DRAM コントローラ 13 から許可されると、DRAM コントローラインターフェース 34 は、それにตอบสนองして、アクセス要求選択部 32 にアクセス要求の選択を開始させる。

40

【0040】

アクセス要求選択部 32 は、要求 FIFO に保管されている複数のアクセス要求から、以下で説明する所定の基準に従って、アクセス要求を順に選択し、それらを含む転送パッケージを作成する (ステップ S63)。次に、アクセス要求選択部 32 は、選択した 1 つまたは複数のアクセス要求を含む転送パッケージを、DRAM コントローラインターフェ

50

ース34を介してDRAMコントローラ13に転送する(ステップS64)。最後に、データ転送部33は、アクセス要求がリードの場合には、共有DRAM14から読み出されたデータをDRAMコントローラ13から受け取って、リードFIFOに保管し、要求元のプロセッサにリードデータがあることをレディ信号で知らせ、リードデータを要求元プロセッサに転送する(ステップS65)。また、データ転送部33は、アクセス要求がライトの場合には、共有DRAM14に書き込むべきライトデータを要求元プロセッサに要求して、ライトFIFOに保管し、次いでそれをDRAMコントローラ14へ転送する(ステップS65)。

【0041】

上記の各ステップのうち、アクセス要求を選択するステップS63以外は、例えば特許文献1にも記載のように、よく知られているので、それらの詳細については省略する。

【0042】

アクセス要求選択部32が実行するステップS63の詳細を図7および図8に示す。

【0043】

アクセス要求選択部32は、まず、アクセス要求受け付け部31の要求FIFOに保管されている複数のアクセス要求から、あとのアクセス要求選択の基準となる最初のアクセス要求を選択する(ステップS71)。この最初のアクセス要求は、要求FIFOにおいて最も待ち時間が長いものが好ましい。その場合、図5の例では、アクセス要求選択部は、プロセッサ1からのリードのアクセス要求P1(1)を選択することになる。

【0044】

次に、プロセッサインターフェース11に割り当てられている許容可能なアクセスタイムの制限を超えたかどうかを判断するために、アクセス要求選択部32は、最初のアクセス要求P1(1)の選択に続いて、図8のステップS81に進み、プロトコルオーバーヘッドを累算する。最初のアクセス要求P1(1)は、図5の例ではリードであるので、そのプロトコルオーバーヘッドは、例えば26クロックである(図1参照)。次に、アクセス要求選択部32は、プロトコルオーバーヘッドの累算値を許容可能なアクセスタイムと比較することによって、その制限(例えば、1000クロック)を超えたかどうかを判断し(ステップS82)、超えていなければ選択したアクセス要求を転送パッケージに含めて(ステップS83)、次のステップ(今の場合は、図7のステップS72)に進む。最初のアクセス要求P1(1)の場合は、上の例では、累算値は制限を超えていないが、図8のフローは、後述のように、アクセス要求の選択のたびに実行されるので、複数のアクセス要求が選択されて、それらのプロトコルオーバーヘッドの累算値が制限を超えると、アクセス要求選択部32は、直前に選択したアクセス要求を廃棄して、アクセス要求選択プロセスを終了し(ステップS84)、図6のステップS64に進んで、作成した転送パッケージをDRAMコントローラインターフェース34に転送する。ステップS84で廃棄されたアクセス要求は要求FIFOには残っているので、図7のフローを再び実行するときは、選択の対象となる。

【0045】

アクセス要求選択部32は、次のステップS72で、要求FIFOのエントリを上から順に、すなわち、待ち時間が長いものから順に走査することにより、最初のアクセス要求に対してBIモードで転送可能なアクセス要求があるかどうかを判断する。図5の例では、最初のアクセス要求のバンクアドレスはBank1であり、これと異なるバンクアドレスを持つアクセス要求として、プロセッサ3(P3)からの2番目のアクセス要求P3(2)が見つかるので、アクセス要求選択部はこのアクセス要求を選択する(ステップS73)。もし見つからなければ、ステップS74に進む。

【0046】

前述のように、従来は、最初のアクセス要求に対してCNモードで転送可能なアクセス要求(図5の例では、同じプロセッサ1(P1)からの2番目のアクセス要求P1(2))が次に選択されていたが、本発明はBIモードを優先する。BIモードでは、少なくとも2つの異なった(Bank、Row)の組み合わせが使用されるため、それらの組み合

10

20

30

40

50

わせに対して、後続のアクセス要求をCNモードで転送できるようになるからである。

【0047】

アクセス要求選択部32は、BIモードで転送可能なアクセス要求の選択に続いて、図8のプロセスを前述のようにして実行し、累算値が制限を超えていなければ、選択したアクセス要求を転送パッケージに含めた後、ステップS72に戻って、ステップS73で前に選択したアクセス要求の次のアクセス要求(図5の例では、P2(1))のところから順に走査することにより、BIモードで転送可能なアクセス要求がまだあるかどうか、すなわちBank1およびBank8以外のバンクアドレスを持つアクセス要求があるかどうかを判断する。このように、ステップS72では、要求FIFOのエントリを上から順に、すなわち待ち時間が長いものから順に走査されるので、ステップS73で選択されるアクセス要求が複数の場合には、それらは待ち時間の長い順に選択されることになる。

10

【0048】

アクセス要求選択部32は、BIモードで転送可能なアクセス要求がなければ、ステップS74に進んで、要求FIFOにおいて選択済みのアクセス要求の次のアクセス要求から順に走査し、選択済みアクセス要求のバンクアドレスおよびロウアドレスの組み合わせを、まだ選択されていない他のアクセス要求のバンクアドレスおよびロウアドレスの組み合わせと比較することによって、CNモードで転送可能なアクセス要求があるかどうかを判断する。図5の例では、最初に選択したアクセス要求P1(1)は、バンクアドレスおよびロウアドレスの組み合わせが(Bank1、Row10)であり、これと同じ組み合わせを持つアクセス要求は、上から順に、すなわち待ち時間の長いものから順に、P1(2)~P1(n)、P2(2)、Pm(3)である。また、2番目に選択したアクセス要求P3(2)は、バンクアドレスおよびロウアドレスの組み合わせが(Bank8、Row6)であり、これと同じ組み合わせを持つアクセス要求はP3(3)である。このように、本発明によれば、CNモードに優先して、BIモードで転送可能なアクセス要求を選択しているのので、その後でCNモードで転送可能なアクセス要求を少なくとも2組選択することができる。

20

【0049】

アクセス要求選択部32は、要求FIFOにおいて選択済みのアクセス要求の次のアクセス要求から順に走査することによって、CNモードで転送可能なアクセス要求を見つけると(図5の例では、最初にP1(2)が見つかる)、それを選択し(ステップS75)、続いて図8のプロセスを実行する。図8のプロセスで累算値が制限を超えていなければ、選択したアクセス要求を転送パッケージに含めた後、ステップS74に戻って、CNモードで転送可能なアクセス要求の選択を繰り返す。ステップS74およびS75を繰り返すことにより、アクセス先のバンクが異なる2つのアクセス要求P1(1)およびP3(2)に対してCNモードで転送可能なアクセス要求が待ち時間の長い順に選択されることになる。

30

【0050】

上記の説明から明らかなように、CNモードで転送可能なアクセス要求が選択されるのは、BIモードで転送可能なアクセス要求がない場合(図7のステップS72の「いいえ」からステップS74を経由してステップS75に至るパス)、先行アクセス要求がBIモードの場合(ステップS73からステップS74を経由してステップS75に至るパス)、または先行アクセス要求がCNモードの場合(図8のステップS83から図7のステップS74を経由してステップS75に至るパス)である。

40

【0051】

アクセス要求選択部32は、CNモードで転送可能なアクセス要求がない場合は、ステップS71に戻って、まだ選択されていないアクセス要求を対象として、アクセス要求選択の基準となる次のアクセス要求(図5の例では、Pm(1))を選択し、上述の各ステップを繰り返す。

【0052】

図9は、このようにして選択された複数のアクセス要求からなる転送パッケージの例を

50

示す。図9の転送パッケージにおいて、アクセス要求P1(1)からアクセス要求P1(n)までは、図7のステップS74において「いいえ」のパスに出るまでに選択されたアクセス要求であり、アクセス要求Pm(1)からアクセス要求Pm(n)までは、ステップS74からステップS71に戻って開始される2回目の選択プロセスで選択されたアクセス要求である。アクセス要求Pm(1)の直後のアクセス要求P2(1)は、それまでに選択されなかったアクセス要求のうちで、アクセス要求Pm(1)に対してBIモードで転送可能なアクセス要求、すなわち、アクセス要求Pm(1)とは異なるバンクアドレスを持つアクセス要求であり、その後ろには、アクセス要求Pm(1)と同じバンクアドレスおよびロウアドレスを持つアクセス要求Pm(2)およびPm(n)が、CNモードで転送可能なアクセス要求として、待ち時間の長い順に選択されている。

10

【0053】

アクセス要求選択部32は、この転送パッケージをDRAMコントローラインターフェース34に送る。DRAMコントローラインターフェース34は、受け取った転送パッケージに含まれる各アクセス要求に応じて、モード信号(R/Wと、BIモードまたはCNモードを示す信号)およびアドレスをDRAMコントローラ13に順次供給し、アクセス要求がライトの場合は、データ転送部33からライトデータをDRAMコントローラ13に転送させ、アクセス要求がリードの場合は、DRAMコントローラ13から送られてきたリードデータをデータ転送部33に受け取らせる。

【0054】

DRAMコントローラ13は、DRAMコントローラインターフェース34からのモード信号に応じて適切なコマンドシーケンスを共有DRAMに送り、リードまたはライトを実行する。例えば、図9の例では、最初のリードアクセス要求P1(1)に対して次のリードアクセス要求P3(2)がBIモードになっているので、DRAMコントローラ13は、図1の(B)に示されるようなコマンドシーケンスを発行する。ただし、3番目のアクセス要求P1(2)からP1(n)まではCNモードになっているので、プリチャージコマンドPは、CNモードの最後のアクセス要求P1(n)まで発行されない。その次のPm(1)からは、前と同様のコマンドシーケンスが繰り返される。

20

【0055】

以上のように、本発明によれば、基準となる最初のアクセス要求が選択された後は、CNモードに優先して、BIモードで転送可能なアクセス要求が選択される。従って、その後、CNモードで転送可能なアクセス要求を選択する際には、バンクアドレスが異なる少なくとも2つのアクセス要求に対して連続的にアクセス要求を選択できるので、バースト転送サイズの最大化を図ることができる。ただし、最初にBIモードで転送可能なアクセス要求を選択すると、FIFOの順番外の選択になることがあるので、例えば、要求FIFOの対応するエントリに選択済みのタグをつけておけば、2回目以降のアクセス要求選択プロセスにおいて、当該アクセス要求が再度選択されるという不都合を回避することができる。

30

【0056】

以上、本発明の好ましい実施形態について説明してきたが、本発明は上述の実施形態に限定されるものではなく、本発明の範囲および趣旨から逸脱することなく様々な変更や修正を行うことが可能である。例えば、上述の実施形態では、あらかじめ決められた許容可能なアクセスタイムがプロセッサインターフェース11に割り当てられていたが、DRAMコントローラ13に接続されているのがプロセッサインターフェース11だけであれば、そのような割り当てを行わなくてもよい。その場合は、図8のフローは不要になるため、ステップS84に代わって、どのようなときに転送を打ち切るかをあらかじめ定義しておく必要があるが、それは任意に決めてよい。

40

【符号の説明】

【0057】

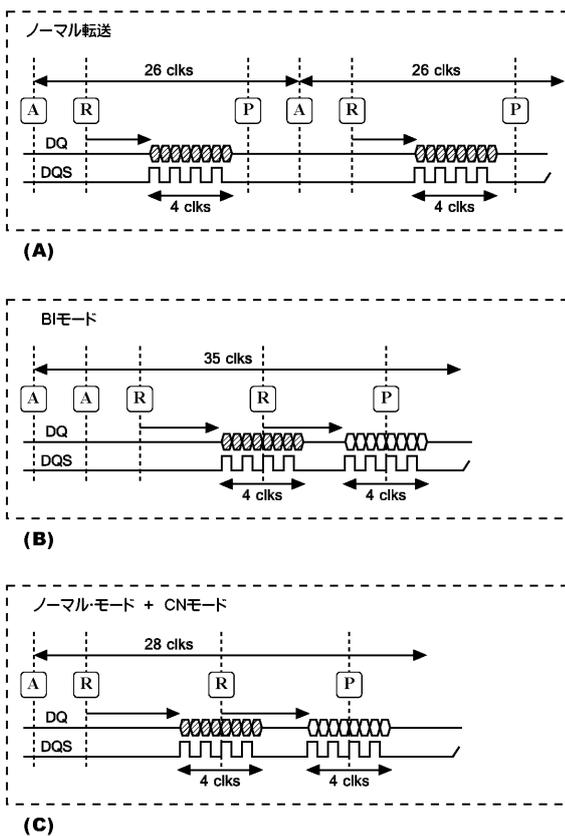
11：プロセッサインターフェース

12：プロセッサ

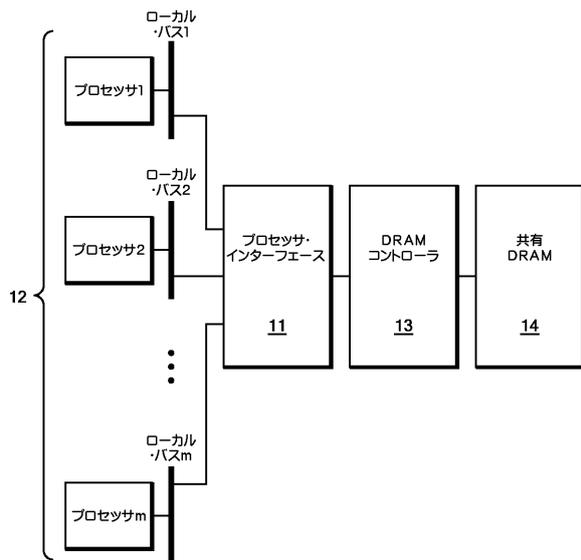
50

- 1 3 : D R A M コントローラ
- 1 4 : 共有 D R A M
- 3 1 : アクセス要求受け付け部
- 3 2 : アクセス要求選択部
- 3 3 : データ転送部
- 3 4 : D R A M コントローラインターフェース

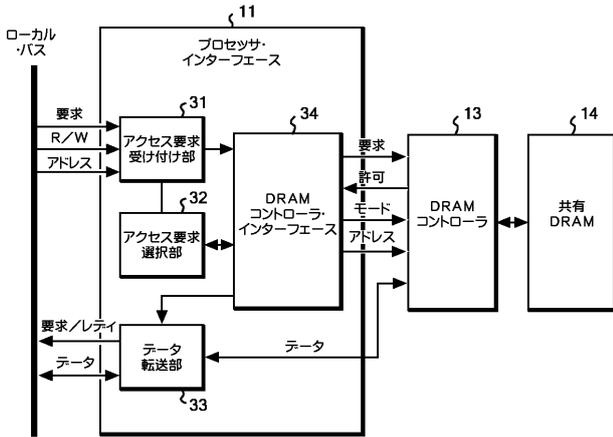
【 図 1 】



【 図 2 】



【 図 3 】



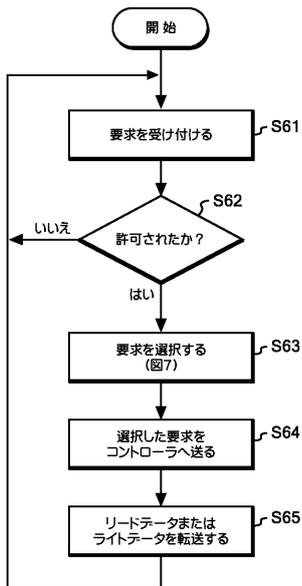
【 図 4 】

プロセッサID	アドレス	R/W
---------	------	-----

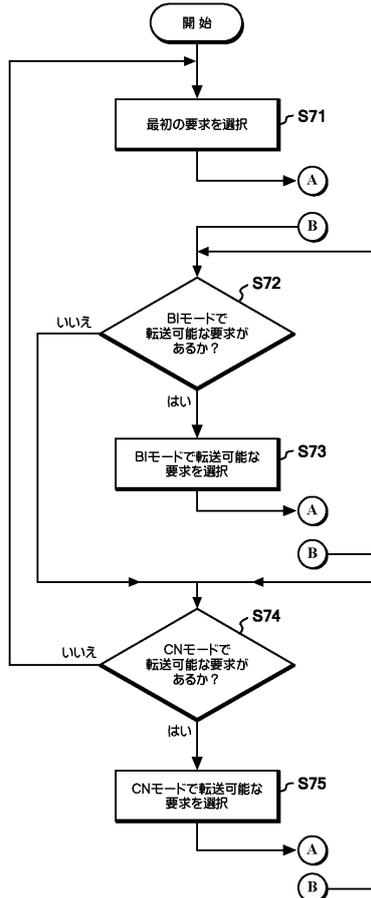
【 図 5 】

プロセッサID	アドレス	R/W
P1 (1)	Bank1, Row10	R
P1 (2)	Bank1, Row10	R
Pm (1)	Bank1, Row3	W
P3(1)	Bank1, Row7	R
P3 (2)	Bank8, Row6	W
P2 (1)	Bank8, Row10	W
P2 (2)	Bank1, Row10	R
P3 (3)	Bank8, Row6	R
Pm (2)	Bank1, Row3	W
Pm (3)	Bank1, Row10	R
P1 (3)	Bank1, Row10	R
⋮	⋮	⋮
P1 (n)	Bank1, Row10	R
⋮	⋮	⋮
Pm (n)	Bank1, Row3	W

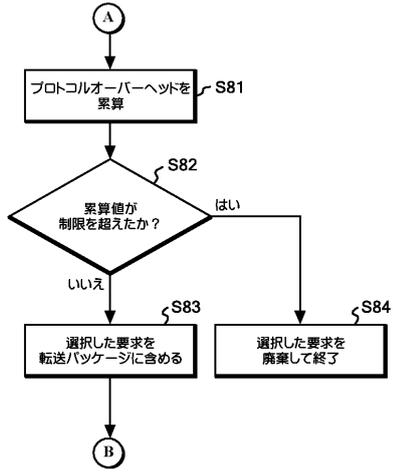
【 図 6 】



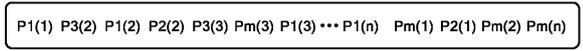
【 図 7 】



【 図 8 】



【 図 9 】



フロントページの続き

(74)代理人 100112690

弁理士 太佐 種一

(72)発明者 長原 里華

東京都江東区豊洲五丁目6番52号 NBF豊洲キャナルフロント 日本アイ・ビー・エム株式会社
東京基礎研究所内

(72)発明者 松尾 久人

東京都江東区豊洲五丁目6番52号 NBF豊洲キャナルフロント 日本アイ・ビー・エム株式会社
IBM東京ラボラトリー内

Fターム(参考) 5B060 CA08 CD03 CD11