

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6127907号
(P6127907)

(45) 発行日 平成29年5月17日(2017.5.17)

(24) 登録日 平成29年4月21日(2017.4.21)

(51) Int.Cl.

F 1

G06F 12/08 (2016.01)

G06F 12/08 12/08 517B

G06F 12/10 (2016.01)

G06F 12/08 519E

G06F 12/08 523C

G06F 12/10 501Z

G06F 12/08 509F

請求項の数 7 (全 14 頁)

(21) 出願番号	特願2013-220675 (P2013-220675)	(73) 特許権者 000005223
(22) 出願日	平成25年10月23日 (2013.10.23)	富士通株式会社
(65) 公開番号	特開2014-112360 (P2014-112360A)	神奈川県川崎市中原区上小田中4丁目1番
(43) 公開日	平成26年6月19日 (2014.6.19)	1号
審査請求日	平成28年7月5日 (2016.7.5)	(74) 代理人 100105142
(31) 優先権主張番号	特願2012-248661 (P2012-248661)	弁理士 下田 憲次
(32) 優先日	平成24年11月12日 (2012.11.12)	(72) 発明者 三浦 隆司
(33) 優先権主張国	日本国 (JP)	神奈川県川崎市中原区上小田中4丁目1番 1号 富士通株式会社内

審査官 酒井 恭信

最終頁に続く

(54) 【発明の名称】演算処理装置及び演算処理装置の制御方法

(57) 【特許請求の範囲】

【請求項 1】

キャッシュメモリと、

前記キャッシュメモリを制御する第1のコントローラと、

前記キャッシュメモリを介さずにアクセスされるノンキャッシュ空間が割り当てられた第2のコントローラと、

を含み、前記第1のコントローラは、前記ノンキャッシュ空間への第1及び第2のアクセスリクエストの処理順序が追い越し可能であり且つ前記第1及び第2のアクセスリクエストのアクセス先が同一である条件が満足される場合、前記第2のコントローラへ先に発行した前記第1のアクセスリクエストに対する前記第2のコントローラからの完了通知を待たずに前記第2のアクセスリクエストを前記第2のコントローラに発行し、前記条件が満たされない場合、前記第2のコントローラへ先に発行した前記第1のアクセスリクエストに対する前記第2のコントローラからの完了通知を待ってから前記第2のアクセスリクエストを前記第2のコントローラに発行することを特徴とする演算処理装置。

【請求項 2】

前記第2のコントローラは複数のコントローラを含み、前記第1及び第2のアクセスリクエストが前記複数のコントローラのうちの1つの同一のコントローラにアクセスする場合に、前記第1のコントローラは、前記第1及び第2のアクセスリクエストのアクセス先が同一であると判定することを特徴とする請求項1記載の演算処理装置。

【請求項 3】

10

20

前記第1のコントローラは、1次キャッシュコントローラと2次キャッシュコントローラとを含み、

前記1次キャッシュコントローラは、前記ノンキャッシュ空間への前記第1及び第2のアクセスリクエストの処理順序が追い越し可能であると判定した場合、前記2次キャッシュコントローラへ先に発行した前記第1のアクセスリクエストに対する前記2次キャッシュコントローラからの応答を待ってから前記第2のアクセスリクエストを前記2次キャッシュコントローラに発行し、前記ノンキャッシュ空間への前記第1及び第2のアクセスリクエストの処理順序が追い越し可能でないと判定した場合、前記2次キャッシュコントローラへ先に発行した前記第1のアクセスリクエストに対する前記第2のコントローラからの完了通知を待ってから前記第2のアクセスリクエストを前記2次キャッシュコントローラに発行し。10

前記2次キャッシュコントローラは、前記第1のアクセスリクエストを前記第2のコントローラに発行すると、前記完了通知を待つことなく前記応答を前記1次キャッシュコントローラに送信する

ことを特徴とする請求項1又は2記載の演算処理装置。

【請求項4】

前記第1のコントローラは、前記ノンキャッシュ空間への複数のアクセスリクエストの処理順序が追い越し可能であるが前記複数のアクセスリクエストのアクセス先が同一でない場合、前記第2のコントローラへ先に発行した全てのアクセスリクエストに対する前記第2のコントローラからの完了通知を待ってから、次のアクセスリクエストを前記第2のコントローラに発行することを特徴とする請求項1乃至3何れか一項記載の演算処理装置。20

【請求項5】

論理アドレスを物理アドレスに変換するために用いるTLBを更に含み、

前記第1のコントローラは、前記TLBに含まれる情報に基づいて、前記ノンキャッシュ空間への前記第1及び第2のアクセスリクエストの処理順序が追い越し可能であるか否かを判定することを特徴とする請求項1乃至4何れか一項記載の演算処理装置。

【請求項6】

キャッシュメモリと、前記キャッシュメモリを制御する第1のコントローラと、前記キャッシュメモリを介さずにアクセスされるノンキャッシュ空間が割り当てられた第2のコントローラとを含む演算処理装置において、30

前記ノンキャッシュ空間への第1及び第2のアクセスリクエストの処理順序が追い越し可能である第1の条件が満たされるか否かを判定し、

前記第1及び第2のアクセスリクエストのアクセス先が同一である第2の条件が満たされるか否かを判定し、

前記第1の条件及び前記第2の条件の両方が満たされる場合、前記第2のコントローラへ先に発行した前記第1のアクセスリクエストに対する前記第2のコントローラからの完了通知を待たずに、前記第2のアクセスリクエストを前記第1のコントローラから前記第2のコントローラに発行し、40

前記第1の条件及び前記第2の条件の少なくとも一方が満たされない場合、前記第2のコントローラへ先に発行した前記第1のアクセスリクエストに対する前記第2のコントローラからの完了通知を待ってから、前記第2のアクセスリクエストを前記第1のコントローラから前記第2のコントローラに発行することを特徴とする演算処理装置の制御方法。

【請求項7】

前記第2のコントローラは複数のコントローラを含み、前記第1及び第2のアクセスリクエストが前記複数のコントローラのうちの1つの同一のコントローラにアクセスする場合に、前記第1のコントローラは、前記第2の条件が満たされると判定することを特徴とする請求項6記載の演算処理装置の制御方法。

【発明の詳細な説明】**【技術分野】****【0001】**

本願開示は、演算処理装置及び演算処理装置の制御方法に関する。

【背景技術】**【0002】**

CPUが備えるコマンドとして、ノンキャッシュ空間へのアクセス命令がある。ノンキャッシュ空間とは、キャッシュメモリを介在せずにアクセスされるメモリ空間をいう。ノンキャッシュ空間へのアクセス命令は、キャッシュメモリを介さないで、アクセス対象のデバイスに割り当てられているメモリ空間にアクセスするための命令である。ノンキャッシュ命令によるノンキャッシュ空間へのアクセスは、ノンキャッシュブル空間として定義されたアドレス空間への読み書きとして定義される。10

【0003】

ノンキャッシュ空間へのアクセスの動作は、レジスタの読み書きをしたり、I/O装置への動作を指示したりするものである。例えば、メモリコントローラへのノンキャッシュリクエストにより、メモリコントローラ内にあるレジスタにアクセスできる。また例えば、PCIe (Peripheral Component Interconnect Express) コントローラへのノンキャッシュリクエストにより、PCIeコントローラ内のレジスタ、又は、例えばPCIeカード等の外部装置内のレジスタにアクセスすることができる。また例えば、CPUインターフェースコントローラへのノンキャッシュリクエストにより、他のCPUに繋がっているメモリコントローラやPCIeコントローラ等の装置へアクセスすることができる。20

【0004】

デバイスのドライバからの割り込み処理として、当該デバイスに対してノンキャッシュライト動作を複数回実行し、その後同期をとるためにノンキャッシュリード動作を1回実行するアクセスパターンがよくある。従って、ノンキャッシュリクエストを連續発行する際に効率的に発行できることが望ましい。

【0005】

従来のノンキャッシュ制御において、命令コントローラから発行されたリクエストを受けた1次キャッシュコントローラは、TLB (Translation Lookaside Buffer) を用いて、アクセス先の仮想アドレスを物理アドレスへと変換する。この物理アドレス内のNCビット(キャッシュブル空間かノンキャッシュブル空間かを表すビット)がノンキャッシュ空間を示す場合、1次キャッシュコントローラは、ノンキャッシュリクエストを2次キャッシュコントローラへ発行する。2次キャッシュコントローラは、このノンキャッシュリクエストを、リクエスト宛先のシステムコントローラ(メモリコントローラ、PCIeコントローラ、CPUインターフェースコントローラ等)に発行する。30

【0006】

命令コントローラから発行される次のリクエストがノンキャッシュリクエストの場合、このリクエストは1次キャッシュコントローラにおいて待機する。システムコントローラにおいて最初のリクエストの処理が完了すると、当該システムコントローラは、2次キャッシュコントローラに完了通知を発行する。完了通知を受け取った2次キャッシュコントローラは、1次キャッシュコントローラに完了通知を通知する。完了通知を受け取った1次キャッシュコントローラは、待機させておいた次のノンキャッシュリクエストを2次キャッシュコントローラに発行することができる。40

上述の説明のように従来技術では、デバイスからの完了通知を待たなければ、1次キャッシュコントローラを内蔵するCPUコアから、2次キャッシュコントローラに対して、次のノンキャッシュリクエストを発行することができない。しかしながら、CPUコアからデバイスの完了通知を待たずに、連続でノンキャッシュリクエストを発行できる場合があるにも関わらず、一律にデバイスの完了通知を待つ方式としたのでは、ノンキャッシュリクエストの処理効率が悪くなってしまう。

【先行技術文献】

【特許文献】**【0007】**

【特許文献1】特開2007-172609号公報

【発明の概要】**【発明が解決しようとする課題】****【0008】**

以上を鑑みると、効率よくノンキャッシュリクエストを発行できる演算処理装置が望まれる。

【課題を解決するための手段】**【0009】**

演算処理装置は、キャッシュメモリと、前記キャッシュメモリを制御する第1のコントローラと、前記キャッシュメモリを介さずにアクセスされるノンキャッシュ空間が割り当てられた第2のコントローラを含み、前記第1のコントローラは、前記ノンキャッシュ空間への第1及び第2のアクセスリクエストの処理順序が追い越し可能であり且つ前記第1及び第2のアクセスリクエストのアクセス先が同一である条件が満足される場合、前記第2のコントローラへ先に発行した前記第1のアクセスリクエストに対する前記第2のコントローラからの完了通知を待たずに前記第2のアクセスリクエストを前記第2のコントローラに発行し、前記条件が満たされない場合、前記第2のコントローラへ先に発行した前記第1のアクセスリクエストに対する前記第2のコントローラからの完了通知を待ってから前記第2のアクセスリクエストを前記第2のコントローラに発行することを特徴とする。

10

【発明の効果】**【0010】**

少なくとも1つの実施例によれば、効率よくノンキャッシュリクエストを発行できる演算処理装置が提供される。

【図面の簡単な説明】**【0011】**

【図1】演算処理装置及び周辺デバイス等を含む演算処理システムの構成の一例を示す図である。

【図2】I U - R E Q のフォーマットの一例を示す図である。

【図3】N C - R E Q のフォーマットの一例を示す図である。

30

【図4】アクセスリクエスト発行処理の流れを示すフローチャートである。

【図5】アクセスリクエスト発行処理の動作の一例を示す図である。

【図6】アクセスリクエスト発行処理の動作の別の例を示す図である。

【図7】アクセスリクエスト発行処理の動作の更に別の例を示す図である。

【図8】T L B の T T E のフォーマットを示す図である。

【図9】アクセスリクエスト発行処理の回路構成を示す図である。

【発明を実施するための形態】**【0012】**

以下に、本発明の実施例を添付の図面を用いて詳細に説明する。

【0013】

図1は、演算処理装置及び周辺デバイス等を含む演算処理システムの構成の一例を示す図である。図1の演算処理システムは、演算処理装置であるC P U 1 0、メインメモリ1 1、外部装置1 2、及び他C P U 1 3を含む。C P U 1 0は、複数のC P Uコア2 1 - 1乃至2 1 - n、2次キャッシュコントローラ2 2、システムコントローラ2 3、及び2次キャッシュメモリ2 4を含む。複数のC P Uコア2 1 - 1乃至2 1 - nの各々は同一の構成を有しており、C P Uコア2 1 - 1に代表して示されるように、演算器3 0、命令コントローラ3 1、1次キャッシュメモリ3 2、1次キャッシュコントローラ3 3、及びT L B3 4を含む。また、2次キャッシュコントローラ内には、各C P Uコアに対応するアドレス判断部3 8 - 1乃至3 8 - nを含む。システムコントローラ2 3は、メモリコントローラ3 5、P C I e コントローラ3 6、及びC P Uインターフェースコントローラ3 7を含

40

50

む。

【0014】

なお図1において、各ボックスで示される各機能ブロックと他の機能ブロックとの境界は、基本的には機能的な境界を示すものであり、物理的な位置の分離、電気的な信号の分離、制御論理的な分離等に対応するとは限らない。各機能ブロックは、他のブロックと物理的にある程度分離された1つのハードウェアモジュールであってもよいし、或いは他のブロックと物理的に一体となつたハードウェアモジュール中の1つの機能を示したものであってもよい。

【0015】

CPUコア21-1乃至21-nは、2次キャッシュメモリ24を共有し、2次キャッシュコントローラ22を介して2次キャッシュメモリ24にアクセスする。またCPUコア21-1乃至21-nは、2次キャッシュコントローラ22を介してシステムコントローラ23の各コントローラにアクセスする。メモリコントローラ35は、外部メモリであるメインメモリ11を制御するコントローラである。PCIeコントローラ36は、PCIeカード等の外部装置12を制御するコントローラである。CPUインターフェースコントローラ37は、CPU10と同様の構成及び機能を有する他CPU13との情報のやり取りを制御する。システムコントローラ23のこれらのコントローラには、キャッシュメモリを介さずにアクセスされるノンキャッシュ空間が割り当てられている。

10

【0016】

命令コントローラ31は、1次キャッシュメモリ32からフェッチした命令をデコードし、デコード結果に応じて演算器30による演算命令の実行を制御する。また命令コントローラ31は、ロード命令やストア命令等のアクセスリクエスト(IU-REQ: Instruction Unit-Request)を1次キャッシュコントローラ33に発行し、1次キャッシュメモリ32に対するデータのロードやデータのストアを実行する。

20

【0017】

命令コントローラ31から発行されたアクセスリクエスト(IU-REQ)を受けた1次キャッシュコントローラ33は、TLB34を参照して、アクセスリクエスト中の仮想アドレスを物理アドレスへと変換する。図8よりTLB34のTTE(Translation Table Entry)には、アクセス空間にサイドエフェクトがあるかないかを示すEビット50(TTEe)と物理ページ番号を示すPA51がある。アクセスリクエスト中のアクセス先のアドレスが、インオーダー処理を要件とするデバイスのメモリ空間に対応する場合には、TLB34を参照することによりTTEe=1が得られる。またアクセスリクエスト中のアクセス先のアドレスが、アクセスリクエストの処理順序が追い越し可能であるデバイスのメモリ空間に対応する場合には、TLB34を参照することによりTTEe=0が得られる。

30

【0018】

またTLB34を参照して仮想アドレスとPA51から得られた物理アドレスには、当該アドレスがキャッシュブル空間かノンキャッシュブル空間かを示すビットであるNCビットが含まれる。1次キャッシュコントローラ33は、NCビットがノンキャッシュ空間を示す場合、ノンキャッシュリクエスト(NC-REQ: NonCache-Request)を、2次キャッシュコントローラ22へ発行する。1次キャッシュコントローラ33は、NCビットがキャッシュ空間を示す場合、1次キャッシュメモリ32に対するアクセスを実行する。1次キャッシュメモリ32と2次キャッシュメモリ24とは階層構造となっており、1次キャッシュメモリ32においてアクセスがヒットしない場合、2次キャッシュコントローラ22を介して2次キャッシュメモリ24へのアクセスが実行される。1次キャッシュコントローラ33と2次キャッシュコントローラ22とは、キャッシュメモリ(1次キャッシュメモリ32及び2次キャッシュメモリ24)を制御するコントローラである。

40

【0019】

図2は、IU-REQのフォーマットの一例を示す図である。図2に示されるように、命令コントローラ31から1次キャッシュコントローラ33に発行するアクセスリクエス

50

トである I U - R E Q は、命令コード 4 1 と仮想アドレス 4 2 を含む。命令コード（オペコード）4 1 は、当該命令の種類を示すコードであり、例えばストア命令（書き込み命令）であることを示すコードやロード命令（読み出し命令）であることを示すコードである。仮想アドレス 4 2 は、ストア命令やロード命令等においてアクセスする先を示す仮想アドレスである。

【0020】

図 3 は、N C - R E Q のフォーマットの一例を示す図である。図 3 に示されるように、1 次キャッシュコントローラ 3 3 から 2 次キャッシュコントローラ 2 2 へ発行するアクセスリクエストである N C - R E Q は、ノンキャッシュ命令コード（オペコード）4 3 と物理アドレス 4 4 と C O R E - I D 4 9 を含む。ノンキャッシュ命令コード 4 3 は、当該命令の種類を示すコードであり、例えば当該命令が例えばノンキャッシュストア命令（書き込み命令）であることを示すコードや、ノンキャッシュロード命令（読み出し命令）であることを示すコードである。物理アドレス 4 4 は、N C ビット 4 5 、C P U 識別子である C P U - I D 4 6 、コントローラ識別子である C T L - I D 4 7 、及びアドレス 4 8 を含む。また、C O R E - I D 4 9 は C P U コア識別子である。

10

【0021】

N C ビット 4 5 は、前述のように、アクセス先のアドレス（アドレス 4 8 が指し示す先）がキャッシュブル空間かノンキャッシュブル空間かを示すビットである。また C P U - I D 4 6 は、当該 N C - R E Q がアクセスする先の C P U を識別する識別子である。例えば N C - R E Q が、図 1 に示す C P U 1 0 内の P C I e コントローラ 3 6 にアクセスする場合、C P U - I D 4 6 は、図 1 に示す当該 C P U 1 0 を識別する識別子となる。また N C - R E Q が、図 1 に示す他 C P U 1 3 の P C I e コントローラにアクセスする場合、C P U - I D 4 6 は、図 1 に示す他 C P U 1 3 を識別する識別子となる。C T L - I D 4 7 は、アクセス先のコントローラを識別する識別子である。例えば N C - R E Q が、図 1 に示す C P U 1 0 内の P C I e コントローラ 3 6 にアクセスする場合、C T L - I D 4 7 は、図 1 に示す P C I e コントローラ 3 6 を識別する識別子となる。アドレス 4 8 は、メモリ空間中のアクセス先の物理アドレスである。例えば N C - R E Q が、図 1 に示す C P U 1 0 内の P C I e コントローラ 3 6 にアクセスする場合、アドレス 4 8 は、P C I e コントローラ 3 6 に割り当てられたメモリ空間中の特定のアドレスとなる。C O R E - I D 4 9 は各 C P U コア 2 1 - 1 乃至 2 1 - n からの N C - R E Q から作成され、N C - R E Q のリクエスト元である C P U コアを識別する識別子となる。

20

【0022】

図 4 は、アクセスリクエスト発行処理の流れを示すフローチャートである。図 4 を参照して、アクセスリクエスト発行処理について説明する。

【0023】

ステップ S 1 において、命令コントローラ 3 1 がリクエスト（I U - R E Q ）を発行する。ステップ S 2 において、1 次キャッシュコントローラ 3 3 が、リクエスト（I U - R E Q ）を受け付け、リクエスト中の仮想アドレスに基づき T L B 3 4 を参照することにより、仮想アドレスに対応する物理アドレスを得る。1 次キャッシュコントローラ 3 3 は更に、物理アドレス中の N C ビットを確認することにより、当該リクエストがノンキャッシュ空間へのアクセスであるか否かを判断する。当該リクエストがキャッシュ空間へのアクセスである場合には、1 次キャッシュメモリ 3 2 や 2 次キャッシュメモリ 2 4 に対する通常のキャッシュアクセス制御が実行される。当該リクエストがノンキャッシュ空間へのアクセスである場合には、ステップ S 3 以降の処理が実行される。

40

【0024】

ステップ S 3 で、1 次キャッシュコントローラ 3 3 は、T L B 3 4 を参照することにより得られた T T E e の値が 0 であるか否かを判定する。T T E e = 0 でない場合（T T E e = 1 の場合）、アクセス先のメモリ空間がインオーダー処理を要件とするデバイスであることを意味する。この場合、ステップ S 4 で、1 次キャッシュコントローラ 3 3 は、当該リクエストの直前のリクエストに対するアクセス先のデバイス又はコントローラからの

50

完了通知（リクエストされた処理の実行の完了を知らせる通知）を既に受領しているか否かを判定する。既に完了通知を受領していれば、1次キャッシュコントローラ33は、2次キャッシュコントローラ22にリクエスト（N C - R E Q）を発行する（ステップS8）。完了通知を受領していない場合、ステップS5で、1次キャッシュコントローラ33は、デバイスからの完了通知が到来するまで待機する。完了通知が到来すると、1次キャッシュコントローラ33は、2次キャッシュコントローラ22にリクエスト（N C - R E Q）を発行する（ステップS8）。

【0025】

このように1次キャッシュコントローラ33は、ノンキャッシュ空間へのアクセスリクエストの処理順序が追い越し可能でないと判定した場合、2次キャッシュコントローラ22へ先に発行したアクセスリクエストに対する完了通知を待つ。この完了通知は、アクセス先のデバイスからシステムコントローラ23を介して、又はシステムコントローラ23から到来する。1次キャッシュコントローラ33は、先に発行したアクセスリクエストに対する完了通知を待ってから（即ち完了通知が到来すると）、現在処理中のアクセスリクエストを2次キャッシュコントローラ22に発行する。

【0026】

ステップS3の判定結果がT T E e = 0である場合、アクセス先のメモリ空間がアクセスリクエストの処理順序が追い越し可能であるデバイスに対応することを意味する。この場合、ステップS6で、直前のリクエストについて2次キャッシュコントローラ22からの応答（N C - T K N）があるか否かを確認する。この応答（N C - T K N）は、2次キャッシュコントローラ22が、アクセスリクエストをシステムコントローラ23に発行すると、上記の完了通知を待つことなく1次キャッシュコントローラ33に送信するものである。

【0027】

応答を受領していない場合（ステップS6でN o）、ステップS7で、1次キャッシュコントローラ33は応答が到来するまで待機する。応答を受領すると（ステップS6でY e s）、ステップS8で、1次キャッシュコントローラ33は、2次キャッシュコントローラ22にリクエスト（N C - R E Q）を発行する。

【0028】

このように1次キャッシュコントローラ33は、ノンキャッシュ空間へのアクセスリクエストの処理順序が追い越し可能であると判定した場合、2次キャッシュコントローラ22へ先に発行したアクセスリクエストに対する応答を待つ。この応答は、2次キャッシュコントローラ22が、アクセスリクエストをシステムコントローラ23に発行すると、完了通知を待つことなく1次キャッシュコントローラ33に送信するものである。1次キャッシュコントローラ33は、先に発行したアクセスリクエストに対する応答を待ってから（即ち応答が到来すると）、現在処理中のアクセスリクエストを2次キャッシュコントローラ22に発行する。

【0029】

ステップS9で、2次キャッシュコントローラ22は、1次キャッシュコントローラ33からのアクセスリクエストを受け付ける。ステップS10で、2次キャッシュコントローラ22は、システムコントローラ23へのリクエストの発行数を確認する。ここでリクエストの発行数とは、2次キャッシュコントローラ22からシステムコントローラ23へ発行したリクエストであって、その完了通知がシステムコントローラ23から届いていないリクエストの数である。

【0030】

ステップS10での判定結果が発行数 = 0を示す場合、ステップS11で、2次キャッシュコントローラ22はシステムコントローラ23にアクセスリクエストを発行し、且つ、1次キャッシュコントローラ33に応答N C - T K Nを送信する。2次キャッシュコントローラ22は更に、発行したリクエストのアドレスを保持し（記憶し）、発行数のカウントを1増加する。

10

20

30

40

50

【 0 0 3 1 】

ステップ S 1 0 での判定結果が発行数が 0 より大きく且つ $f_u l_1$ 未満であることを示す場合、ステップ S 1 2 で、2 次キャッシュコントローラ 2 2 は、直前のアクセスリクエストの記憶してあるアドレスと現在処理中のアクセスリクエストのアドレスとを比較する。この比較はアクセス先（「宛先」）を確認するために行われる。

【 0 0 3 2 】

直前と現在との 2 つのアクセスリクエストのアクセス先が同一である条件が満される場合（ステップ S 1 2 で「同一宛先」）、ステップ S 1 1 で、2 次キャッシュコントローラ 2 2 は、システムコントローラ 2 3 にアクセスリクエストを発行する。この際、2 次キャッシュコントローラ 2 2 は、システムコントローラ 2 3 へ先に発行したアクセスリクエストに対するシステムコントローラ 2 3 からの完了通知を待たずに、現在処理中のアクセスリクエストをシステムコントローラ 2 3 に発行する。なおシステムコントローラ 2 3 は複数のコントローラ（35, 36, 37）を含んでいる。従って、2 次キャッシュコントローラ 2 2 がシステムコントローラ 2 3 にアクセスリクエストを発行する場合には、N C - R E Q の物理アドレス（図 3 参照）の C T L - I D 4 7 が示す 1 つのコントローラにアクセスリクエストが発行されることになる。また、2 つのアクセスリクエストがこれら複数のコントローラのうちの 1 つの同一のコントローラにアクセスする場合に、2 次キャッシュコントローラ 2 2 は、2 つのアクセスリクエストのアクセス先が同一であると判定する。より具体的には、図 3 に示す物理アドレス 4 4 のうち、上位ビット側の C P U - I D 4 6、及び C T L - I D 4 7 が 2 つのアクセスリクエストにおいて一致している場合に、2 つのアクセスリクエストのアクセス先が同一であると判定する。下位ビット側のアドレス 4 8 が 2 つのアクセスリクエストにおいて異なっていても、2 つのアクセスリクエストのアクセス先が同一であると判定する。

10

【 0 0 3 3 】

直前と現在との 2 つのアクセスリクエストのアクセス先が同一である条件が満されない場合（ステップ S 1 2 で「別宛先」）、ステップ S 1 3 で、2 次キャッシュコントローラ 2 2 は、発行数が 0 になるまで待機する。発行数は、システムコントローラ 2 3 へ既に発行したアクセスリクエストに対するシステムコントローラ 2 3 からの完了通知（リクエストされた処理実行が完了したことを知らせる通知：N C - E N D）が到来するたびに、1 減少する。発行数が 0 になると、2 次キャッシュコントローラ 2 2 は、現在処理中のアクセスリクエストをシステムコントローラ 2 3 に発行する（ステップ S 1 1）。このようにして、アクセス先が同一である条件が満されない場合、2 次キャッシュコントローラ 2 2 は、システムコントローラ 2 3 へ先に発行したアクセスリクエストに対するシステムコントローラ 2 3 からの完了通知を待機する。2 次キャッシュコントローラ 2 2 は、先に発行したアクセスリクエストに対するシステムコントローラ 2 3 からの完了通知を待ってから、現在処理中のアクセスリクエストをシステムコントローラ 2 3 に発行する。この際、システムコントローラ 2 3 へ先に発行したアクセスリクエストが複数有る場合には、全てのアクセスリクエストに対するシステムコントローラ 2 3 からの完了通知を待ってから、現在処理中のアクセスリクエストをシステムコントローラ 2 3 に発行する。

20

【 0 0 3 4 】

30

ステップ S 1 0 での判定結果が発行数 = $f_u l_1$ を示す場合、ステップ S 1 3 で、2 次キャッシュコントローラ 2 2 は、発行数 < $f_u l_1$ になるまで待機する。ここで $f_u l_1$ とは、2 次キャッシュコントローラ 2 2 が受け付けることのできるリクエストの数であり、例えば、受け付けたリクエストを保持する 2 次キャッシュコントローラ 2 2 に内蔵のバッファの容量に依存する。

40

【 0 0 3 5 】

ステップ S 1 4 で、システムコントローラ 2 3（メモリコントローラ 3 5、P C I e コントローラ 3 6、又は C P U インタフェースコントローラ 3 7）がリクエストを受け付け、必要に応じてデバイス側にリクエストを発行する。即ち、メインメモリ 1 1 や、外部装置 1 2 や、他 C P U 1 3 等にリクエストを発行する。なお、例えば P C I e コントローラ

50

36がリクエストを受け付けた場合、当該リクエストがPCIeコントローラ36内部のレジスタにアクセスするリクエストである場合には、外部装置12にリクエストを発行することはない。一方、当該リクエストが外部装置12にアクセスするリクエストである場合には、PCIeコントローラ36から外部装置12にリクエストが発行される。

【0036】

ステップS15で、デバイス側或いはシステムコントローラ23によりリクエスト処理が完了すると、処理完了を示す完了通知(NC-END)がシステムコントローラ23から2次キャッシュコントローラ22に送信される。前述のように、完了通知(NC-END)が到来するたびに、2次キャッシュコントローラ22側では、発行数を1減少する処理を行う。

10

【0037】

図5は、アクセスリクエスト発行処理の動作の一例を示す図である。まず命令コントローラ31から1次キャッシュコントローラ33にリクエストIU REQ1が発行される。このリクエストIU REQ1に対応するTTEeが0である場合、ノンキャッシュリクエストNC REQ1が、1次キャッシュコントローラ33から2次キャッシュコントローラ22へと発行される。ここで、次のリクエストIU REQ2及びIU REQ3が命令コントローラ31から発行され、対応するTTEeが0である場合、対応するノンキャッシュリクエストNC REQ2及びNC REQ3は、1次キャッシュコントローラ33で待機する。

【0038】

20

1次キャッシュコントローラ33からノンキャッシュリクエストNC REQ1を受け取った2次キャッシュコントローラ22は、初期状態での発行数が0であるので、直ちにシステムコントローラ23にノンキャッシュリクエストNC REQ1を発行する。この際、発行数は1増加される(ステップS51)。またNC REQ1の発行とともに、又は並行して、又は直前若しくは直後に、2次キャッシュコントローラ22は、NC REQ1の発行を示す通知として、1次キャッシュコントローラ33に応答NC TKN1を発行する。NC TKN1を受け取った1次キャッシュコントローラ33は、次のノンキャッシュリクエストNC REQ2を2次キャッシュコントローラ22に発行する。

【0039】

30

NC REQ2を受け取った2次キャッシュコントローラ22は、NC REQ1のアクセス先とNC REQ2のアクセス先とを比較して、同一アクセス先(同一宛先)ならば、直ちにシステムコントローラ23にNC REQ2を発行する。この際、発行数は1増加される(ステップS52)。またNC REQ2の発行とともに、又は並行して、又は直前若しくは直後に、2次キャッシュコントローラ22は、NC REQ2の発行を示す通知として、1次キャッシュコントローラ33に応答NC TKN2を発行する。NC TKN2を受け取った1次キャッシュコントローラ33は、次のノンキャッシュリクエストNC REQ3を2次キャッシュコントローラ22に発行する。

【0040】

30

NC REQ3を受け取った2次キャッシュコントローラ22は、NC REQ2のアクセス先とNC REQ3のアクセス先とを比較して、同一アクセス先(同一宛先)ならば、直ちにシステムコントローラ23にNC REQ3を発行する。この際、発行数は1増加される(ステップS53)。またNC REQ3の発行とともに、又は並行して、又は直前若しくは直後に、2次キャッシュコントローラ22は、NC REQ3の発行を示す通知として、1次キャッシュコントローラ33に応答NC TKN3を発行する。

40

【0041】

リクエスト完了通知NC END1及びNC END2がシステムコントローラ23から2次キャッシュコントローラ22に送信されると、2次キャッシュコントローラ22では各完了通知毎に発行数を1減算する(ステップS54、S56)。

図6は、アクセスリクエスト発行処理の動作の別の一例を示す図である。まず命令コントローラ31から1次キャッシュコントローラ33にリクエストIU REQ1が発行さ

50

れる。このリクエストIU - REQ 1に対応するTTEeが0である場合、ノンキャッシュリクエストNC - REQ 1が、1次キャッシュコントローラ33から2次キャッシュコントローラ22へと発行される。ここで、次のリクエストIU - REQ 2及びIU - REQ 3が命令コントローラ31から発行され、対応するTTEeが0である場合、対応するノンキャッシュリクエストNC - REQ 2及びNC - REQ 3は、1次キャッシュコントローラ33で待機する。

【0042】

1次キャッシュコントローラ33からノンキャッシュリクエストNC - REQ 1を受け取った2次キャッシュコントローラ22は、初期状態での発行数が0であるので、直ちにシステムコントローラ23にノンキャッシュリクエストNC - REQ 1を発行する。この際、発行数は1増加される(ステップS61)。またNC - REQ 1の発行とともに、又は並行して、又は直前若しくは直後に、2次キャッシュコントローラ22は、NC - REQ 1の発行を示す通知として、1次キャッシュコントローラ33に応答NC - TKN 1を発行する。NC - TKN 1を受け取った1次キャッシュコントローラ33は、次のノンキャッシュリクエストNC - REQ 2を2次キャッシュコントローラ22に発行する。

【0043】

NC - REQ 2を受け取った2次キャッシュコントローラ22は、NC - REQ 1のアクセス先とNC - REQ 2のアクセス先とを比較する。アクセス先が異なる(別宛先)であるので(ステップS62)、2次キャッシュコントローラ22は、先にシステムコントローラ23へ発行したNC - REQ 1に対する完了通知NC - END 1が到来するまでNC - REQ 2を待機させる。NC - REQ 1に対する完了通知NC - END 1がシステムコントローラ23から到来すると、2次キャッシュコントローラ22は、発行数を1減算する(ステップS63)。これにより発行数が0となるので、2次キャッシュコントローラ22は、システムコントローラ23にNC - REQ 2を発行する。これに応じて、発行数は1増加される(ステップS63)。またNC - REQ 2の発行とともに、又は並行して、又は直前若しくは直後に、2次キャッシュコントローラ22は、NC - REQ 2の発行を示す通知として、1次キャッシュコントローラ33に応答NC - TKN 2を発行する。NC - TKN 2を受け取った1次キャッシュコントローラ33は、次のノンキャッシュリクエストNC - REQ 3を2次キャッシュコントローラ22に発行する。

【0044】

NC - REQ 3を受け取った2次キャッシュコントローラ22は、NC - REQ 2のアクセス先とNC - REQ 3のアクセス先とを比較する。アクセス先が異なる(別宛先)であるので(ステップS64)、2次キャッシュコントローラ22は、先にシステムコントローラ23へ発行したNC - REQ 2に対する完了通知NC - END 2が到来するまでNC - REQ 3を待機させる。

【0045】

リクエスト完了通知NC - END 2がシステムコントローラ23から2次キャッシュコントローラ22に送信されると、2次キャッシュコントローラ22では発行数を1減算する(ステップS65)。

【0046】

図7は、アクセスリクエスト発行処理の動作の更に別の一例を示す図である。図7に示す例では、命令コントローラ31、1次キャッシュコントローラ33、及び2次キャッシュコントローラ22の動作は、図5に示す動作と同じである。図7においては、図5の場合と異なり、システムコントローラ23又はデバイスによるリクエスト処理が、NC - REQ 1よりもNC - REQ 2に対して先に完了している。その結果、NC - REQ 2に対する完了通知NC - END 2が、NC - REQ 1に対する完了通知NC - END 1よりも先に、システムコントローラ23から2次キャッシュコントローラ22に発行されている。NC - REQ 1及びNC - REQ 2は、TTEeが0であり、アクセス先のメモリ空間がアクセスリクエストの処理順序が追い越し可能であるデバイスに対応している。従って、図7に示す動作例の場合のように、後にリクエストが発行されたNC - END 2が、先

10

20

30

40

50

にリクエストが発行された N C - R E Q 1 を追い越し、先に処理が完了してしまってよい。このようにアクセスリクエストの処理順序が追い越し可能である場合、後に発行されたアクセスリクエストが、先に発行されたアクセスリクエストよりも、先に処理が開始されて先に処理が完了したり、後に処理が開始されて先に処理が完了したりしてよい。即ち、アクセスリクエストの処理順序が追い越し可能である場合、処理順序がリクエスト発行順序と異なってもよく、処理順序がリクエスト発行順序と異なったことによる結果として、処理結果に支障が生じることがない。

【 0 0 4 7 】

図 9 はアクセスリクエスト発行処理の回路構成を示す図である。複数の C P U コア 2 1 - 1 乃至 2 1 - n からそれぞれ 2 次キャッシュコントローラ 2 2 に N C - R E Q が発行される。2 次キャッシュコントローラ 2 2 には各 C P U コア 2 1 - 1 乃至 2 1 - n に対応したアドレス判断部 3 8 - 1 乃至 3 8 - n を有しており、それぞれのアドレス判断部 3 8 - 1 乃至 3 8 - n で宛先確認を行う。アドレス判断部 3 8 - 1 乃至 3 8 - n では、C P U コア 2 1 - 1 乃至 2 1 - n から N C - R E Q が来た場合、そのリクエストのアドレスを保持しておく。宛先確認は、現在処理中の N C - R E Q と保持していた直前の N C - R E Q の C P U - I D 4 6 、及び C T L - I D 4 7 を比較して同一宛先かどうかを判定する。システムコントローラ 2 3 ヘリクエスト発行後もアドレスは保持しておく。選択部 3 9 では発行可能となった C P U コア 2 1 - 1 乃至 2 1 - n からのリクエストを任意に選択してシステムコントローラ 2 3 へと N C - R E Q を発行する。また、システムコントローラ 2 3 からのリクエスト完了通知を受けた 2 次キャッシュコントローラ 2 2 は、応答部 4 0 で完了通知に付随された C O R E - I D を確認して、リクエスト発行元の C P U コアへと N C - E N D を返す。
10

【 0 0 4 8 】

以上、本発明を実施例に基づいて説明したが、本発明は上記実施例に限定されるものではなく、特許請求の範囲に記載の範囲内で様々な変形が可能である。

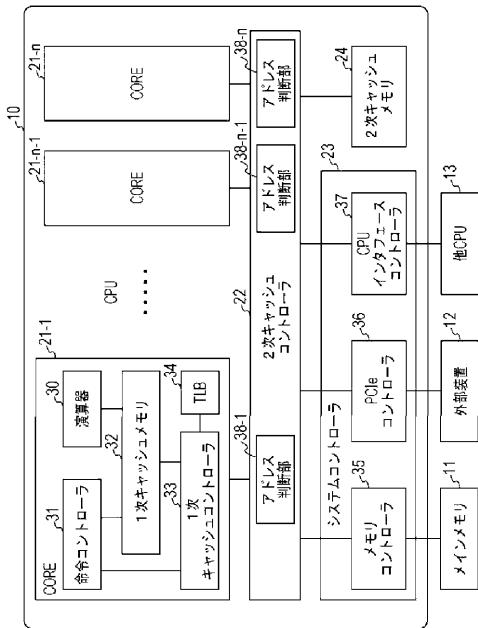
【 符号の説明 】

【 0 0 4 9 】

- 1 0 C P U
 - 1 1 メインメモリ
 - 1 2 外部装置
 - 1 3 他 C P U
 - 2 1 - 1 乃至 2 1 - n C P U コア
 - 2 2 2 次キャッシュコントローラ
 - 2 3 システムコントローラ
 - 2 4 2 次キャッシュメモリ
 - 3 0 演算器
 - 3 1 命令コントローラ
 - 3 2 1 次キャッシュメモリ
 - 3 3 1 次キャッシュコントローラ
 - 3 4 T L B
 - 3 5 メモリコントローラ
 - 3 6 P C I e コントローラ
 - 3 7 C P U インタフェースコントローラ
 - 3 8 - 1 乃至 3 8 - n アドレス判断部
 - 3 9 選択部
 - 4 0 応答部
-
- 30
-
- 40

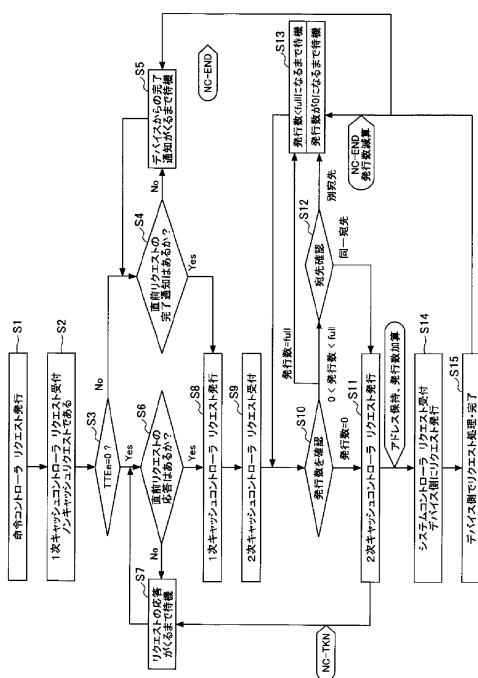
【図1】

演算処理装置及び周辺デバイス等を含む
演算処理システムの構成の一例を示す図



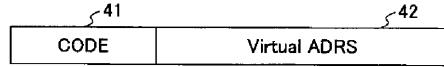
【図4】

アクセスリクエスト発行処理の流れを示すフローチャート



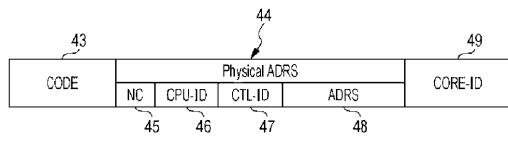
【図2】

IU-REQのフォーマットの一例を示す図



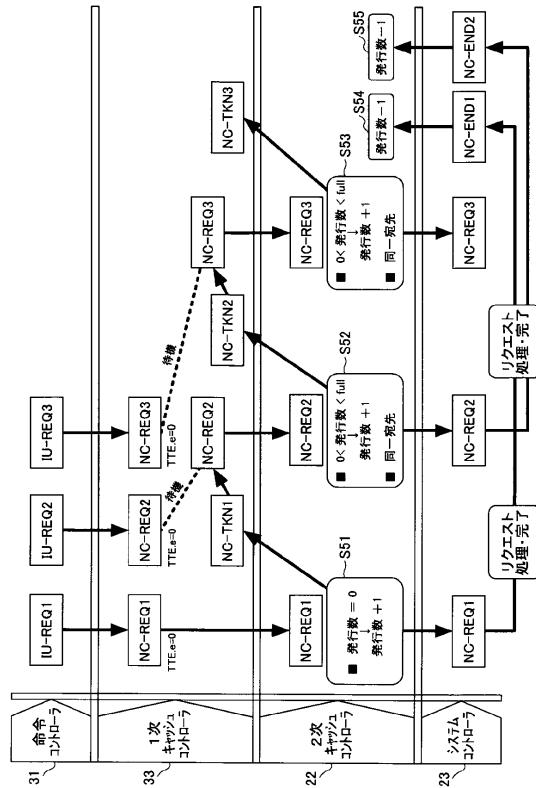
【図3】

NC-REQのフォーマットの一例を示す図



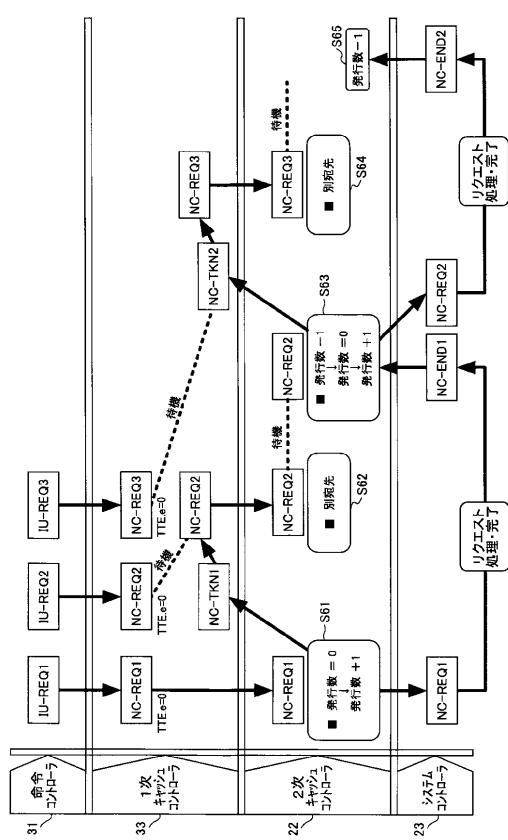
【図5】

アクセスリクエスト発行処理の動作の一例を示す図



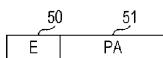
【 四 6 】

アクセスリクエスト発行処理の動作の一例を示す図



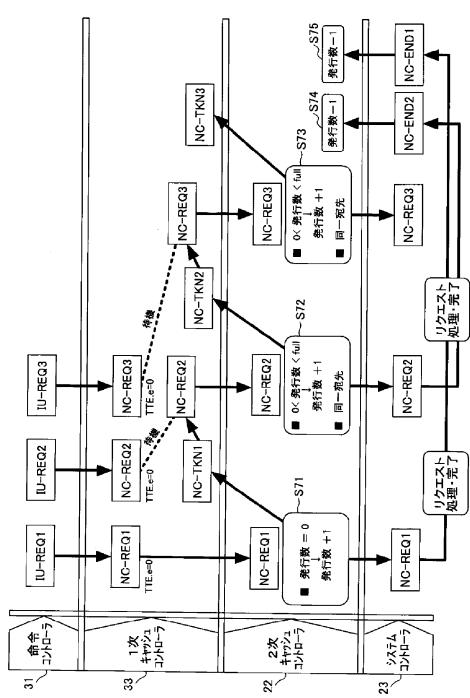
【 四 8 】

TLBのTTTのフォーマットを示す図



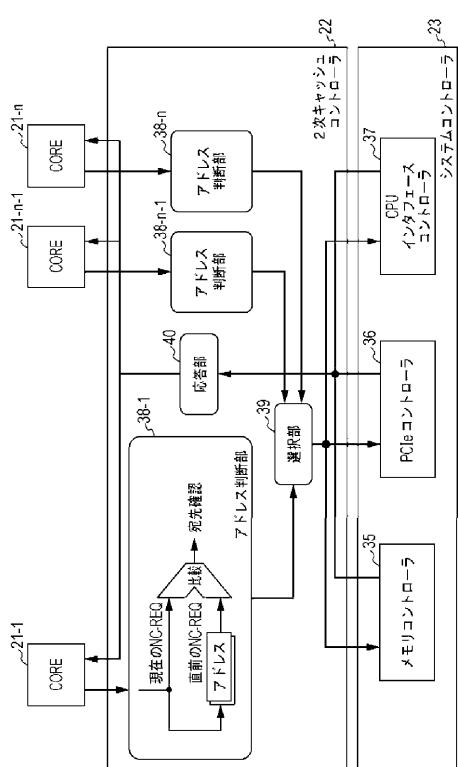
【図7】

アクセスリクエスト発行処理の動作の更に別の一例を示す図



【図9】

アクセスリクエスト発行処理の回路構成を示す図



フロントページの続き

(56)参考文献 国際公開第2010/032433 (WO, A1)

特開2001-117859 (JP, A)

米国特許第6697899 (US, B1)

特開平05-265916 (JP, A)

特開2001-154914 (JP, A)

米国特許第5751996 (US, A)

米国特許第6014737 (US, A)

(58)調査した分野(Int.Cl., DB名)

G 06 F 12/08 - 12/10