

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号
特許第5118731号
(P5118731)

(45) 発行日 平成25年1月16日(2013. 1. 16)

(24) 登録日 平成24年10月26日(2012. 10. 26)

(51) Int.Cl.

F I

G O 6 F 12/08 (2006.01)

G O 6 F 12/08 5 1 7 C

G O 6 F 12/08 5 0 9 Z

請求項の数 6 (全 17 頁)

(21) 出願番号	特願2010-180807 (P2010-180807)	(73) 特許権者	000003078
(22) 出願日	平成22年8月12日 (2010. 8. 12)		株式会社東芝
(65) 公開番号	特開2012-42994 (P2012-42994A)		東京都港区芝浦一丁目1番1号
(43) 公開日	平成24年3月1日 (2012. 3. 1)	(74) 代理人	100117787
審査請求日	平成24年8月9日 (2012. 8. 9)		弁理士 勝沼 宏仁
		(74) 代理人	100082991
			弁理士 佐藤 泰和
		(74) 代理人	100103263
			弁理士 川崎 康
		(74) 代理人	100107582
			弁理士 関根 毅
		(74) 代理人	100118843
			弁理士 赤岡 明
		(74) 代理人	100146123
			弁理士 木本 大介

最終頁に続く

(54) 【発明の名称】 キャッシュユニット及びプロセッシングシステム

(57) 【特許請求の範囲】

【請求項 1】

クリティカルワードファーストに対応しない外部バスを介して接続された主記憶装置から前記クリティカルワードファーストに対応する内部バスに接続された第1ラインサイズを有する1次キャッシュにデータを転送するキャッシュユニットにおいて、

前記第1ラインサイズ以上の第2ラインサイズを有する2次キャッシュと、前記1次キャッシュの第1リフィル要求のバースト転送の先頭アドレスが前記2次キャッシュに存在していない場合に、前記第1リフィル要求を、前記バースト転送の先頭アドレスが処理対象となる処理データのアドレスであり、バースト長が前記第1ラインサイズ未満である第2リフィル要求に変換する要求変換部と、を備えるキャッシュ制御部と、

前記要求変換部により変換された第2リフィル要求を前記主記憶装置に転送し、前記第2リフィル要求に対応する処理データを前記主記憶装置から受け取り、受け取った処理データを前記1次キャッシュに転送する非キャッシュ制御部と、を備えることを特徴とするキャッシュユニット。

【請求項 2】

前記処理データのアドレスが所定の閾値アドレス以上である場合に、前記要求変換部をオンにする切換部をさらに備える、

請求項1記載のキャッシュユニット。

【請求項 3】

前記切換部は、前記第1リフィル要求が発行された時刻からの経過時間が所定時間以上

となる場合、又は前記第1リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットである場合に、前記閾値アドレスをアドレス順が小さい方向にシフトする、
請求項2記載のキャッシュユニット。

【請求項4】

データを変換するデータ変換部をさらに備え、

前記キャッシュ制御部は、前記処理データを含む第2ラインサイズのデータを前記主記憶装置に要求し、

前記データ変換部は、前記キャッシュ制御部により要求された前記処理データを含む第2ラインサイズのデータを、前記1次キャッシュのリフィルを実行するときに前記処理データの後ろに配置すべきリフィルデータに変換し、変換したリフィルデータを前記プロセッシングユニットに転送する、

請求項1乃至3の何れか1項記載のキャッシュユニット。

【請求項5】

前記非キャッシュ制御部は、前記2次キャッシュのリフィル要求を前記キャッシュ制御部に許可する、

請求項1乃至4の何れか1項記載のキャッシュユニット。

【請求項6】

主記憶装置と、

第1ラインサイズを有する1次キャッシュと、

前記主記憶装置から前記1次キャッシュにデータを転送するキャッシュユニットと、

前記1次キャッシュと前記キャッシュユニットとを接続するクリティカルワードファーストに対応しない外部バスと、

前記主記憶装置と前記キャッシュユニットとを接続するクリティカルワードファーストに対応する内部バスと、を備え、

前記キャッシュユニットは、

前記第1ラインサイズ以上の第2ラインサイズを有する2次キャッシュと、前記1次キャッシュの第1リフィル要求のバースト転送の先頭アドレスが前記2次キャッシュに存在していない場合に、前記第1リフィル要求を、前記バースト転送の先頭アドレスが処理対象となる処理データのアドレスであり、バースト長が前記第1ラインサイズ未満である第2リフィル要求に変換する要求変換部と、を備えるキャッシュ制御部と、

前記要求変換部により変換された第2リフィル要求を前記主記憶装置に転送し、前記第2リフィル要求に対応する処理データを前記主記憶装置から受け取り、受け取った処理データを前記1次キャッシュに転送する非キャッシュ制御部と、
を備えることを特徴とするプロセッシングシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明の実施形態は、キャッシュユニット及びプロセッシングシステムに関する。

【背景技術】

【0002】

従来、クリティカルワードファースト（以下、「CWF（Critical Word First）」という）対応バスとCWF非対応バスとが混在するシステムオンチップ（以下、「SoC（System on Chip）」という）などのチップセットにおいて、CWF対応バスに接続されたプロセッシングユニットが要求するデータがCWF非対応バスに接続されたDRAM（Dynamic Random Access Memory）等の主記憶装置に格納されている場合には、主記憶装置からプロセッシングユニットにCWF非対応バスを介してデータが転送される。

【0003】

しかしながら、プロセッシングユニットが処理しようとしているデータ（以下、「クリティカルワード」という）は、主記憶装置から転送されたデータの任意の位置に含まれて

10

20

30

40

50

いる。この場合には、プロセッシングユニットは、クリティカルワードを受け取るまでストール状態を維持する。その結果、プロセッシングユニットの処理効率が低下する。

【0004】

従来、CWF非対応バスに接続された主記憶装置からCWF対応バスに接続されたプロセッシングユニットにデータが転送される場合のプロセッシングユニットの処理効率を改善することが望まれている。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開2008-510246号公報

10

【発明の概要】

【発明が解決しようとする課題】

【0006】

本発明が解決しようとする課題は、CWF非対応バスに接続されたユニットからCWF対応バスに接続されたユニットにデータが転送される場合のCWF対応バスに接続されたユニットの処理効率を改善することである。

【課題を解決するための手段】

【0007】

本実施形態に係るキャッシュユニットは、クリティカルワードファーストに対応しない外部バスを介して接続された主記憶装置からクリティカルワードファーストに対応する内部バスに接続された第1ラインサイズを有する1次キャッシュにデータを転送するモジュールである。キャッシュユニットは、キャッシュ制御部と、非キャッシュ制御部と、を備える。キャッシュ制御部は、第1ラインサイズ以上の第2ラインサイズを有する2次キャッシュと、1次キャッシュの第1リフィル要求のバースト転送の先頭アドレスが2次キャッシュに存在していない場合に、第1リフィル要求を、バースト転送の先頭アドレスが処理対象となる処理データのアドレスであり、バースト長が第1ラインサイズ未満である第2リフィル要求に変換する要求変換部と、を備える。非キャッシュ制御部は、要求変換部により変換された第2リフィル要求を主記憶装置に転送し、第2リフィル要求に対応する処理データを主記憶装置から受け取り、受け取った処理データを1次キャッシュに転送するモジュールである。

20

30

【図面の簡単な説明】

【0008】

【図1】第1実施形態に係るプロセッシングシステム1の構成を示すブロック図。

【図2】キャッシュユニット10の構成を示すブロック図。

【図3】プロセッシングシステム1のデータ処理の手順を示すフローチャート。

【図4】図3のデータ処理において取り扱われるデータの構造を示す概略図。

【図5】図3のデータ処理におけるキャッシュユニット10内のデータフローを示すブロック図。

【図6】第2実施形態に係るキャッシュユニット10の構成を示すブロック図。

【図7】切換部146の切換処理（第1例）の手順を示すフローチャート。

40

【図8】S701を説明するための概略図。

【図9】切換部146の切換処理（第2例）の手順を示すフローチャート。

【図10】図9の切換処理において用いられる要求テーブルの概略図。

【図11】切換部146の切換処理（第3例）の手順を示すフローチャート。

【図12】図11の切換処理において用いられる優先テーブルの概略図。

【図13】切換部146の切換処理（第4例）の手順を示すフローチャート。

【図14】シフトステップ（S1303）を説明するための概略図。

【図15】切換部146の切換処理（第5例）の手順を示すフローチャート。

【発明を実施するための形態】

【0009】

50

本実施形態について図面を参照して説明する。

(第1実施形態)

第1実施形態について説明する。

第1実施形態に係るプロセッシングシステムの構成について説明する。図1は、第1実施形態に係るプロセッシングシステム1の構成を示すブロック図である。

【0010】

プロセッシングシステム1は、データ処理を実行するシステムである。プロセッシングシステム1は、複数のプロセッシングユニットPU0～PU3と、キャッシュユニット10と、主記憶装置20と、内部バス30と、外部バス40と、を備える。

【0011】

各プロセッシングユニットPU0～PU3は、1次キャッシュL1を備える。各プロセッシングユニットPU0～PU3は、処理対象となるデータ（以下、「処理データ」という）にデータ処理を実行するモジュールである。1次キャッシュL1は、複数ワードのデータを記憶可能な第1ラインサイズを有するメモリである。例えば、第1ラインサイズは8ワード（64バイト）である。

【0012】

キャッシュユニット10は、CWFに対応しない外部バス40を介して接続された主記憶装置20からCWFに対応する内部バス30に接続された第1ラインサイズを有する1次キャッシュL1にデータを転送するモジュールである。

【0013】

主記憶装置20は、データを記憶可能な記憶媒体である。例えば、主記憶装置20はDRAMである。

内部バス30は、CWFに対応するデータバスである。

外部バス40は、インクリメンタル（以下、「INCR」という）バーストに対応するデータバスである。外部バス40は、CWFに対応していない。

【0014】

なお、図1には、4つのプロセッシングユニットPU0～PU3が1次キャッシュL1を備えるプロセッシングシステム1が図示されるが、プロセッシングユニットの数は幾つでも良いし、プロセッシングユニットは1次キャッシュL1を備えていなくても良い。

【0015】

図2は、キャッシュユニット10の構成を示すブロック図である。

キャッシュユニット10は、内部バスインタフェース12と、キャッシュ制御部14と、非キャッシュ制御部16と、外部バスインタフェース18と、を備える。

【0016】

内部バスインタフェース12は、内部バス30の転送方式であるCWFに従ってデータを転送するモジュールである。内部バスインタフェース12は、キャッシュ制御部14、及び非キャッシュ制御部16と、内部バス30との間に設けられる。

【0017】

キャッシュ制御部14は、2次キャッシュL2と、要求変換部142と、データ変換部144と、を備える。キャッシュ制御部14は、2次キャッシュL2へのアクセスを制御するとともに、2次キャッシュL2のリフィルを実行するモジュールである。

【0018】

2次キャッシュL2は、複数ワードのデータを記憶可能な第2ラインサイズを有するメモリである。第2ラインサイズは、第1ラインサイズ以上である。例えば、第2ラインサイズは32ワード（256バイト）である。

要求変換部142は、第1リフィル要求を第2リフィル要求に変換するモジュールである。

データ変換部144は、データを変換するモジュールである。

【0019】

非キャッシュ制御部16は、主記憶装置20を含む2次キャッシュL2以外のモジュール

10

20

30

40

50

ルへのアクセスを制御するモジュールである。

外部バスインタフェース18は、キャッシュ制御部14、非キャッシュ制御部16と外部バス40との間に設けられる。

【0020】

プロセッシングシステム1の動作について説明する。以下、プロセッシングユニットPU0が、1次キャッシュL1に記憶されていない処理データにデータ処理を実行する例について説明する。

図3は、プロセッシングシステム1のデータ処理の手順を示すフローチャートである。図4は、図3のデータ処理において取り扱われるデータの構造を示す概略図である。図5は、図3のデータ処理におけるキャッシュユニット10内のデータフローを示すブロック図である。

10

【0021】

<第1リフィル要求発行ステップ(S301)> プロセッシングユニットPU0が、1次キャッシュL1のリフィルを要求するための第1リフィル要求(図4(1))を発行し、発行した第1リフィル要求を内部バスインタフェース12に送信する(図5(1))。図4(1)に示すように、第1リフィル要求では、要求内容を示すコマンドは読み出し(Rd)であり、バースト長は第1ラインサイズ(8ワード)であり、バースト転送の先頭アドレスは処理データのアドレス(0x1000_00F0)であり、バーストパターンは内部バス30の転送方式(CWF)である。すなわち、図4(1)の第1リフィル要求は、アドレス“0x1000_00F0”のデータ、アドレス“0x1000_00F8”のデータ、アドレス“0x1000_00C0”のデータ、アドレス“0x1000_00C8”のデータ、アドレス“0x1000_00D0”のデータ、アドレス“0x1000_00D8”のデータ、アドレス“0x1000_00E0”のデータ、及びアドレス“0x1000_00E8”のデータをCWFで順番に転送するための要求である。

20

【0022】

<S302> 内部バスインタフェース12が、第1リフィル要求のアドレスに基づいて、キャッシュ制御部14へのアクセス(以下、「キャッシュアクセス」という)か、非キャッシュ制御部16へのアクセス(以下、「非キャッシュアクセス」という)かを判定する。例えば、内部バスインタフェース12は、キャッシュアクセスのメモリ空間アドレスのリストと第1リフィル要求のアドレスとを比較し、第1リフィル要求のアドレスがそのリストに含まれる場合にキャッシュアクセスと判定し、含まれない場合に非キャッシュアクセスと判定する。キャッシュアクセスと判定された場合には(S302-YES)、第1リフィル要求が要求変換部142へ送信され(図5(2))、S303が実行される。非キャッシュアクセスと判定された場合には(S302-NO)、第1リフィル要求が非キャッシュ制御部16へ送信され、非キャッシュアクセスステップ(S321)が実行される。

30

【0023】

<S303> 要求変換部142が、2次キャッシュL2に処理データが記憶されているか否かを判定する。すなわち、要求変換部142は、第1リフィル要求のバースト転送の先頭アドレスが2次キャッシュL2に存在しているか否かを判定する。例えば、内部バスインタフェース12は、2次キャッシュL2に記憶されたデータのアドレスリストと第1リフィル要求のアドレスとを比較する。第1リフィル要求のアドレスがアドレスリストに含まれる(すなわち、第1リフィル要求のバースト転送の先頭アドレスが2次キャッシュL2に存在している)場合には、2次キャッシュL2に処理データが記憶されていると判定し、第1リフィル要求のアドレスがアドレスリストに含まれない(すなわち、第1リフィル要求のバースト転送の先頭アドレスが2次キャッシュL2に存在していない)場合には、2次キャッシュL2に処理データが記憶されていないと判定する。2次キャッシュL2に処理データが記憶されていると判定された場合には(S303-YES)、1次キャッシュリフィルステップ(S312)が実行される。2次キャッシュL2に処理データが記憶されていないと判定された場合には(S303-NO)、要求変換ステップ(S304)が実行される。

40

50

【0024】

＜要求変換ステップ（S304）＞ 要求変換部142が、第1リフィル要求を第2リフィル要求（図4（2））に変換し、変換した第2リフィル要求を非キャッシュ制御部16に送信する（図5（3））。より具体的には、要求変換部142は、第1リフィル要求のバースト転送の先頭アドレスが2次キャッシュL2に存在していない場合に、第1リフィル要求のバースト長を短縮する（例えば、1ワードに変更する）ことにより、第1リフィル要求を第2リフィル要求に変換する。第2リフィル要求では、バースト転送の先頭アドレスが処理対象となる処理データのアドレスであり、バースト長が第1ラインサイズ未満である。図4（2）に示すように、第2リフィル要求では、要求内容は読み出し（Rd）であり、バースト転送の先頭アドレスは処理データのアドレス（0x1000_00F0）であり、バースト長は1ワードであり、バーストパターンは外部バス40の転送方式（INCR）である。すなわち、図4（2）の第2リフィル要求は、アドレス“0x1000_00F0”のデータをINCRで転送するための要求である。

10

【0025】

＜第2リフィル要求転送ステップ（S305）＞ 非キャッシュ制御部16が、第2リフィル要求を外部バスインタフェース18に転送する（図5（4））。次いで、外部バスインタフェース18が、外部バス40を介して、第2リフィル要求を主記憶装置20に転送する（図5（5））。すなわち、非キャッシュ制御部16は、外部バスインタフェース18を介して、要求変換部142により変換された第2リフィル要求を主記憶装置20に転送する。

20

【0026】

＜許可ステップ（S306）＞ 非キャッシュ制御部16が、所定のタイミングで、2次キャッシュL2のリフィル要求を許可するための許可通知をキャッシュ制御部14に送信する（図5（6））。例えば、許可通知は、第2リフィル要求が外部バス40により受理されたことを示すフラグ等の1ビット情報である。例えば、所定のタイミングは、外部バス40が第2リフィル要求を受理したタイミング、又は要求変換ステップ（S304）の後にキャッシュユニット10が発行するコマンドが第2リフィル要求を追い越さないことが確定したタイミングである。これにより、1次キャッシュ向けのリフィルデータが処理データより先にプロセッシングユニットPU0に転送されることを確実に防ぐことができる。

30

【0027】

＜2次キャッシュリフィル要求転送ステップ（S307）＞ キャッシュ制御部14が、許可通知を受け取った後、2次キャッシュL2のリフィル要求（図4（3））を発行し、発行した2次キャッシュL2のリフィル要求を外部バスインタフェース18に送信する（図5（7））。次いで、外部バスインタフェース18が、2次キャッシュL2のリフィル要求を主記憶装置20に転送する（図5（8））。すなわち、キャッシュ制御部14は、外部バスインタフェース18を介して、処理データを含む第2ラインサイズのデータを主記憶装置20に要求する。

【0028】

＜処理データ転送ステップ（S308）＞ 外部バスインタフェース18が、外部バス40を介して、第2リフィル要求に対応する1ワードの処理データ（図4（4））を主記憶装置20から受け取る（図5（9））。次いで、外部バスインタフェース18が、処理データを非キャッシュ制御部16に転送する（図5（10））。次いで、非キャッシュ制御部16が、処理データを内部バスインタフェース12に転送する（図5（11））。次いで、内部バスインタフェース12が処理データをプロセッシングユニットPU0に転送する（図5（12））。すなわち、非キャッシュ制御部16は、外部バスインタフェース18を介して主記憶装置20から受け取った処理データを、内部バスインタフェース12を介して1次キャッシュL1に転送する。これにより、プロセッシングユニットPU0は、1ワードの処理データのみを先に受け取ることができる。換言すると、プロセッシングユニットPU0は、バーストパターンがCWFである第1リフィル要求に対応する処理デ

40

50

ータを、CWFに対応しない外部バス40を介して接続された主記憶装置20から受け取る。その後、プロセッシングユニットPU0は、ストール状態を解除し、処理データにデータ処理を実行する。

【0029】

<2次キャッシュリフィルステップ(S309)> 外部バスインタフェース18が、外部バス40を介して、2次キャッシュL2のリフィル要求に対応する32ワードのデータ(図4(5))を主記憶装置20から受け取る(図5(13))。図4(5)に示すように、2次キャッシュL2のリフィル要求に対応するデータは、24ワードの先頭データと、処理データの前方に位置する6ワードの第1リフィルパートと、1ワードの処理データと、処理データの後方に位置する1ワードの第2リフィルパートと、を備える。次いで、外部バスインタフェース18が、2次キャッシュL2のリフィル要求に対応するデータをキャッシュ制御部14に転送する(図5(14))。次いで、キャッシュ制御部14が、2次キャッシュL2のリフィル要求に対応するデータに基づいて2次キャッシュL2のリフィルを実行する。

10

【0030】

<データ変換ステップ(S310)> データ変換部144が、2次キャッシュL2のリフィル要求に対応するデータを7ワードのリフィルデータ(図4(6))に変換し、変換したリフィルデータを非キャッシュ制御部16に送信する(図5(15))。図4(6)に示すように、7ワードのリフィルデータは、1ワードの第2リフィルパートと、第2リフィルパートの後方に位置する6ワードの第1リフィルパートと、を備える。1ワードの第2リフィルパートは、図4(5)の2次キャッシュL2のリフィル要求に対応するデータにおいて処理データの後方に位置していたデータである。6ワードの第1リフィルパートは、図4(5)の2次キャッシュL2のリフィル要求に対応するデータにおいて処理データの前方に位置していたデータである。

20

【0031】

例えば、データ変換部144は、2次キャッシュL2のリフィル要求に対応するデータにおいて1ワードの処理データの前後に合計7ワードのリフィルパートが存在する場合(図4(5)の場合)には、処理データの前方に位置するデータ(第1リフィルパート)の位置を処理データの後方に位置するデータ(第2リフィルパート)の後方に変更することにより、2次キャッシュL2のリフィル要求に対応するデータを7ワードの1次キャッシュL1向けのリフィルデータに変換する。例えば、データ変換部144は、処理データが2次キャッシュL2のリフィル要求に対応するデータの最後尾に位置する場合には、処理データの前方に位置する7ワードのデータ(リフィルパート)の位置を処理データの後方に変更することにより、2次キャッシュL2のリフィル要求に対応するデータを7ワードのリフィルデータに変換する。例えば、データ変換部144は、処理データの後方に7ワードのデータ(リフィルパート)が存在する場合には、処理データの後方に位置する7ワードのデータ(リフィルパート)を抽出することにより、2次キャッシュL2のリフィル要求に対応するデータを7ワードのリフィルデータに変換する。

30

【0032】

<リフィルデータ転送ステップ(S311)> 非キャッシュ制御部16が、リフィルデータを内部バスインタフェース12に転送する(図5(16))。次いで、内部バスインタフェース12が、内部バス30を介して、リフィルデータをプロセッシングユニットPU0に転送する(図5(17))。すなわち、データ変換部144は、キャッシュ制御部14により要求された処理データを含む第2ラインサイズのことを、1次キャッシュL1のリフィルを実行するときに処理データの後方に配置すべきリフィルデータに変換し、非キャッシュ制御部16及び内部バスインタフェース12を介して、変換したリフィルデータをプロセッシングユニットPU0に転送する。なお、2次キャッシュL2のリフィルを実行する前にデータのバイパス転送が可能な場合には、データ変換ステップ(S310)及びリフィルデータ転送ステップ(S311)の組み合わせは、2次キャッシュリフィルステップ(S309)の前に実行されても良いし、2次キャッシュリフィルステップ

40

50

(S 3 0 9) と並列に実行されても良い。

【0033】

<1次キャッシュリフィルステップ(S 3 1 2)> プロセッシングユニットPU0は、1ワードの処理データ及び7ワードのリフィルデータから構成される8ワードのデータセット(図4(7))に基づいて1次キャッシュL1のリフィルを実行する。換言すると、プロセッシングユニットPU0は、CWFに対応しない外部バス40を介して接続された主記憶装置20から受け取ったCWFに対応するデータ構造のデータに基づいて、1次キャッシュL1のリフィルを実行する。1次キャッシュリフィルステップ(S 3 1 2)の後、図3のデータ処理が終了する。

【0034】

<非キャッシュアクセスステップ(S 3 2 1)> 非キャッシュ制御部16が、2次キャッシュL2以外のモジュール(例えば、主記憶装置20)へのアクセスを制御する。非キャッシュアクセスステップ(S 3 2 1)の後、図3のデータ処理が終了する。

【0035】

第1実施形態によれば、CWF非対応バス(外部バス40)に接続されたユニット(主記憶装置20)からCWF対応バス(内部バス30)に接続されたユニット(プロセッシングユニットPU0~PU3)にデータが転送される場合に、CWF対応バスに接続されたユニットの処理効率を改善することができる。

【0036】

(第2実施形態)

第2実施形態について説明する。第2実施形態は、切換部を備えるキャッシュ制御部の例である。なお、第2実施形態に係るプロセッシングシステムの構成は、第1実施形態と同様であり、第1実施形態と同様の説明については省略する。

【0037】

図6は、第2実施形態に係るキャッシュユニット10の構成を示すブロック図である。

キャッシュユニット10は、内部バスインタフェース12と、キャッシュ制御部14と、非キャッシュ制御部16と、外部バスインタフェース18と、を備える。内部バスインタフェース12、非キャッシュ制御部16、及び外部バスインタフェース18は、第1実施形態と同様である。

【0038】

キャッシュ制御部14は、2次キャッシュL2と、要求変換部142と、データ変換部144と、切換部146と、を備える。キャッシュ制御部14は、2次キャッシュL2へのアクセスを制御するとともに、2次キャッシュL2のリフィルを実行するモジュールである。2次キャッシュL2、要求変換部142、及びデータ変換部144は、第1実施形態と同様である。

【0039】

切換部146は、要求変換部142のオン及びオフを切り換えるモジュールである。

切換部146の動作の第1例について説明する。図7は、切換部146の切換処理(第1例)の手順を示すフローチャートである。図8は、S701を説明するための概略図である。

図7の切替処理は、S302においてキャッシュアクセスと判定された場合に実行される。

【0040】

<S701> 切換部146が、第2ラインサイズに相当する32ワードのアドレス空間における第1リフィル要求のアドレスと所定の閾値アドレスとを比較する。例えば、閾値アドレスは、ユーザにより設定される任意の値であって、切換部146のレジスタ(図示せず)に格納されている。第1リフィル要求のアドレスが閾値アドレス以上(図8(1))である場合には(S701-YES)、S702が実行される。第1リフィル要求のアドレスが閾値アドレスより小さい(図8(2))(すなわち、処理データのアドレスが閾値アドレスより前のアドレスである)場合には(S701-NO)、S711が実行さ

10

20

30

40

50

れる。

【0041】

<S702> 切換部146が、要求変換部142をオンにする。すなわち、切換部146は、処理データのアドレスが所定の閾値アドレス以上である場合に、要求変換部142をオンにする。これにより、要求変換部142は、要求変換ステップ(S304)を実行可能になる。S702の後に、図7の切換処理が終了する。

【0042】

<S711> 切換部146が、要求変換部142をオフにする。これにより、要求変換部142は、要求変換ステップ(S304)を実行不可能になる。この場合には、第2リフィル要求転送ステップ(S305)～リフィルデータ転送ステップ(S311)は実行されず、一般的な2次キャッシュL2のリフィル及び1次キャッシュL1のリフィルが実行される。S711の後に、図7の切換処理が終了する。

【0043】

切換部146の動作の第1例では、処理データのアドレスが閾値アドレス以上であるか否かに基づいて要求変換部142のオン及びオフが切り換えられるので、キャッシュユニット10の消費電力を低減するとともに、要求変換部142のオン及びオフを切り換える条件をユーザが任意に設定することができる。

【0044】

切換部146の動作の第2例について説明する。図9は、切換部146の切換処理(第2例)の手順を示すフローチャートである。図10は、図9の切換処理において用いられる要求テーブルの概略図である。

図9の切替処理は、S302においてキャッシュアクセスと判定された場合に実行される。

【0045】

<タイムスタンプ記録ステップ(S901)> 切換部146が、第1リフィル要求を発行したプロセッシングユニット(以下、「発行元プロセッシングユニット」という)の識別情報(以下、「PUID(Processing Unit Identification)」という)と、第1リフィル要求が内部バスインタフェース12により受け付けられた時刻(以下、「タイムスタンプ」という)と、を要求テーブルに記録する。PUIDは、第1リフィル要求に含まれる情報である。タイムスタンプは、内部バスインタフェース12により発行される情報である。要求テーブルは、切換部146のレジスタに格納されている。図10は、プロセッシングユニットPU0が発行した第1リフィル要求がバスインタフェース12により受け付けられた時刻がT1であり、プロセッシングユニットPU3が発行した第1リフィル要求がバスインタフェース12により受け付けられた時刻がT2であり、プロセッシングユニットPU1が発行した第1リフィル要求がバスインタフェース12により受け付けられた時刻がT3であり、プロセッシングユニットPU2が発行した第1リフィル要求がバスインタフェース12により受け付けられた時刻がT4である、ことを示している。

【0046】

<S902> 切換部146が、要求テーブルに記録されたタイムスタンプと閾値時間とを比較する。例えば、閾値時間は、ユーザにより設定される任意の値であって、切換部146のレジスタに格納されている。要求テーブルのタイムスタンプが閾値時間以上である(すなわち、内部バスインタフェース12が第1リフィル要求を受け付けてから閾値時間以上の時間が経過した)場合には(S902-YES)、S903が実行される。要求テーブルのタイムスタンプが閾値時間より小さい(すなわち、内部バスインタフェース12が第1リフィル要求を受け付けてから閾値時間以上の時間が経過していない)場合には(S902-NO)、S911が実行される。

【0047】

<S903> S702と同様である。すなわち、切換部146は、第1リフィル要求が発行された時刻からの経過時間が所定時間以上となる場合に、要求変換部142をオンにする。これにより、要求変換部142は、要求変換ステップ(S304)を実行可能に

なる。S 9 0 3の後に、図 9 の切換処理が終了する。

< S 9 1 1 > S 7 1 1と同様である。S 9 1 1の後に、図 9 の切換処理が終了する。

【 0 0 4 8 】

切換部 1 4 6 の動作の第 2 例では、第 1 リフィル要求が発行された時刻からの経過時間に基づいて要求変換部 1 4 2 のオン及びオフが切り換えられるので、キャッシュユニット 1 0 の消費電力を低減するとともに、プロセッシングユニットのストール時間を一定時間以内に抑えることができる。

【 0 0 4 9 】

切換部 1 4 6 の動作の第 3 例について説明する。図 1 1 は、切換部 1 4 6 の切換処理（第 3 例）の手順を示すフローチャートである。図 1 2 は、図 1 1 の切換処理において用いられる優先テーブルの概略図である。

10

図 1 1 の切替処理は、S 3 0 2 においてキャッシュアクセスと判定された場合に実行される。

【 0 0 5 0 】

< S 1 1 0 1 > 切換部 1 4 6 が、第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットであるか否かを判定する。例えば、第 1 リフィル要求の発行元プロセッシングユニットの P U I D が優先テーブルに含まれている場合には、第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットであると判定される。第 1 リフィル要求の発行元プロセッシングユニットの P U I D が優先テーブルに含まれていない場合には、第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットでないと判定される。図 1 2 は、プロセッシングユニット P U 0 が優先プロセッシングユニットであることを示している。例えば、優先テーブルは、ユーザにより任意に設定される情報であって、切換部 1 4 6 のレジスタに格納されている。第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットであると判定された場合には（S 1 1 0 1 - Y E S）、S 1 1 0 2 が実行される。第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットでないと判定された場合には（S 1 1 0 1 - N O）、S 1 1 1 1 が実行される。

20

【 0 0 5 1 】

< S 1 1 0 2 > S 7 0 2 と同様である。すなわち、切換部 1 4 6 は、第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットである場合に、要求変換部 1 4 2 をオンにする。図 1 2 の場合には、第 1 リフィル要求の発行元プロセッシングユニットがプロセッシングユニット P U 0 である場合に、切換部 1 4 6 は要求変換部 1 4 2 をオンにする。これにより、要求変換部 1 4 2 は、要求変換ステップ（S 3 0 4）を実行可能になる。S 1 1 0 2 の後に、図 1 1 の切換処理が終了する。

30

< S 1 1 1 1 > S 7 1 1 と同様である。S 1 1 1 1 の後に、図 1 1 の切換処理が終了する。

【 0 0 5 2 】

切換部 1 4 6 の動作の第 3 例では、第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットであるか否かに基づいて要求変換部 1 4 2 のオン及びオフが切り換えられるので、キャッシュユニット 1 0 の消費電力を低減するとともに、優先度の高いプロセッシングユニットのストール時間を短縮することができる。

40

【 0 0 5 3 】

切換部 1 4 6 の動作の第 4 例について説明する。図 1 3 は、切換部 1 4 6 の切換処理（第 4 例）の手順を示すフローチャートである。図 1 4 は、図 1 3 のシフトステップ（S 1 3 0 3）を説明するための概略図である。

図 1 3 の切替処理は、S 3 0 2 においてキャッシュアクセスと判定された場合に実行される。

【 0 0 5 4 】

< タイムスタンプ記録ステップ（S 1 3 0 1） > タイムスタンプ記録ステップ（S 9 0 1）と同様である。

50

【0055】

＜S1302＞ 切換部146が、要求テーブルに記録されたタイムスタンプと閾値時間とを比較する。例えば、閾値時間は、ユーザにより任意に設定される情報であって、切換部146のレジスタに格納されている。要求テーブルのタイムスタンプが閾値時間以上である（すなわち、内部バスインタフェース12が第1リフィル要求を受け付けてから閾値時間以上の時間が経過した）場合には（S1302－YES）、シフトステップ（S1303）が実行される。要求テーブルのタイムスタンプが閾値時間より小さい（すなわち、内部バスインタフェース12が第1リフィル要求を受け付けてから閾値時間以上の時間が経過していない）場合には（S1302－NO）、S1304が実行される。

【0056】

＜シフトステップ（S1303）＞ 切換部146が、閾値アドレスをアドレス順が小さい方にシフトする。すなわち、切換部146は、第1リフィル要求が発行された時刻からの経過時間が所定時間以上となる場合に、第1リフィル要求のアドレスより大きな閾値アドレスをアドレス順が小さい方向にシフトする（図14（1）、（2））。これにより、第1リフィル要求のアドレスが閾値アドレス以上となる（図14（3））。

【0057】

＜S1304＞ S701と同様である。

＜S1305＞ S702と同様である。これにより、第1リフィル要求のアドレスが閾値アドレスより小さい場合であっても、要求変換部142は、要求変換ステップ（S304）を実行可能になる。S1305の後に、図13の切換処理が終了する。

＜S1311＞ S711と同様である。S1311の後に、図13の切換処理が終了する。

【0058】

切換部146の動作の第4例では、処理データのアドレスが閾値アドレス以上であるか否かに基づいて要求変換部142のオン及びオフが切り換えられる際に、第1リフィル要求が発行された時刻からの経過時間が所定時間を超えると要求変換部142のオン及びオフを切り換えるための条件が緩和されるので、キャッシュユニット10の消費電力を低減するとともに、ストール時間が一定時間を超えたプロセッシングユニットに対して要求変換部142がオンになり易くすることができる。

【0059】

切換部146の動作の第5例について説明する。図15は、切換部146の切換処理（第5例）の手順を示すフローチャートである。

図15の切替処理は、S302においてキャッシュアクセスと判定された場合に実行される。

【0060】

＜S1501＞ S1101と同様である。第1リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットであると判定された場合には（S1501－YES）、シフトステップ（S1502）が実行される。第1リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットでないと判定された場合には（S1501－NO）、S1503が実行される。

【0061】

＜シフトステップ（S1502）＞ シフトステップ（S1303）と同様である。すなわち、切替部146は、第1リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットである場合に、第1リフィル要求のアドレスより大きな閾値アドレス（図14（1））をアドレス順が小さい方向にシフトする（図14（2））。これにより、第1リフィル要求のアドレスが閾値アドレス以上となる（図14（3））。

【0062】

＜S1504＞ S701と同様である。これにより、第1リフィル要求のアドレスが閾値アドレスより小さい場合であっても、要求変換部142は、要求変換ステップ（S304）を実行可能になる。S1504の後に、図15の切換処理が終了する。

< S 1 5 1 1 > S 7 1 1と同様である。S 1 5 1 1の後に、図 1 5の切換処理が終了する。

【 0 0 6 3 】

切換部 1 4 6の動作の第 5 例では、処理データのアドレスが閾値アドレス以上であるか否かに基づいて要求変換部 1 4 2のオン及びオフが切り換えられる際に、プロセッシングユニットが優先プロセッシングユニットであれば要求変換部 1 4 2のオン及びオフを切り換えるための条件が緩和されるので、キャッシュユニット 1 0の消費電力を低減するとともに、優先度の高いプロセッシングユニットに対して要求変換部 1 4 2がオンになり易くすることができる。

【 0 0 6 4 】

切換部 1 4 6の動作の第 1 例～第 3 例は、それぞれ、組み合わせられても良い。

例えば、切換部 1 4 6の動作の第 1 例と第 2 例とを組み合わせた場合には、第 1 リフィル要求のアドレスが閾値アドレスより小さかったとしても、要求テーブルに記録された時刻 T i が閾値時刻以上であれば、要求変換部 1 4 2がオンになる。これにより、ユーザが設定した要求変換部 1 4 2のオン及びオフを切り換える条件にかかわらず、プロセッシングユニットのストール時間を一定時間以内に抑えることができる。

【 0 0 6 5 】

例えば、切換部 1 4 6の動作の第 2 例と第 3 例とを組み合わせた場合には、第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットであるときには要求テーブルに記録された時刻 T iにかかわらず要求変換部 1 4 2がオンにされ、第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットでないときには要求テーブルに記録された時刻 T i が閾値時刻以上になった後に要求変換部 1 4 2がオンになる。これにより、第 1 リフィル要求が発行された時刻からの経過時間にかかわらず、優先度の高いプロセッシングユニットのストール時間を短縮することができる。

【 0 0 6 6 】

例えば、切換部 1 4 6の動作の第 3 例と第 1 例とを組み合わせた場合には、第 1 リフィル要求のアドレスが閾値アドレスより小さかったとしても、第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットであるときには、要求変換部 1 4 2がオンになる。これにより、ユーザが設定した要求変換部 1 4 2のオン及びオフを切り換える条件にかかわらず、優先度の高いプロセッシングユニットのストール時間を短縮することができる。

【 0 0 6 7 】

切換部 1 4 6の動作の第 4 例及び第 5 例は、組み合わせられても良い。例えば、タイムスタンプが閾値時間以上であること及び第 1 リフィル要求の発行元プロセッシングユニットが優先プロセッシングユニットであることの両方を満たすと、シフトステップ (S 1 3 0 3) が実行される。これにより、キャッシュユニット 1 0の消費電力をさらに低減するとともに、優先プロセッシングユニットのストール時間が一定時間を超えたアクセス (すなわち、緊急度の高いアクセス) を優先的に高速化することができる。

【 0 0 6 8 】

なお、本実施形態では、データ変換部 1 4 4がキャッシュ制御部 1 4に設けられる例について説明したが、データ変換部 1 4 4は、外部バスインタフェース 1 8又は非キャッシュ制御部 1 6に設けられても良い。

【 0 0 6 9 】

また、本実施形態では、データ変換ステップ (S 3 1 0) において、データ変換部 1 4 4がリフィルデータを非キャッシュ制御部 1 6に送信し、リフィルデータ転送ステップ (S 3 1 1) において、非キャッシュ制御部 1 6が、リフィルデータをプロセッシングユニット P U 0に転送する例について説明したが、本実施形態の範囲はこれに限られるものではない。例えば、本実施形態は、データ変換ステップ (S 3 1 0) において、データ変換部 1 4 4がリフィルデータを保持し、リフィルデータ転送ステップ (S 3 1 1) において、データ変換部 1 4 4がリフィルデータをプロセッシングユニット P U 0に送信する場合

10

20

30

40

50

についても適用可能である。

【0070】

なお、本発明は、上述した実施形態に限定されるものではなく、その要旨を逸脱しない範囲で構成要素を変形して具体化される。また、上述した実施形態に開示されている複数の構成要素の適宜な組み合わせにより、種々の発明が形成可能である。例えば、上述した実施形態に示される全構成要素から幾つかの構成要素を削除してもよい。さらに、異なる実施形態にわたる構成要素を適宜組み合わせてもよい。

【符号の説明】

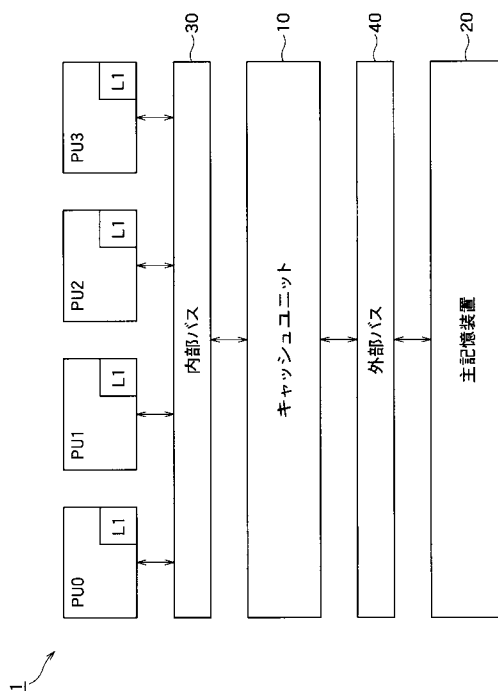
【0071】

- 1 プロセッシングシステム
- 10 キャッシュユニット
- 12 内部バス
- 14 キャッシュ制御部
- 142 要求変換部
- 144 データ変換部
- 146 切換部
- 16 非キャッシュ制御部
- 18 外部バスインタフェース
- 20 主記憶装置
- 30 内部バス
- 40 外部バス
- L1 1次キャッシュ
- L2 2次キャッシュ

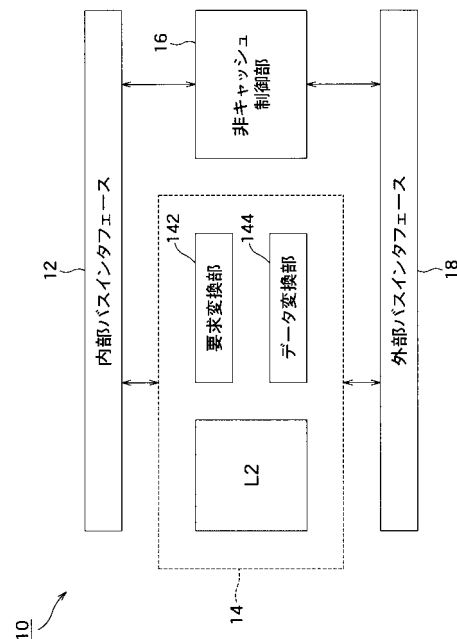
10

20

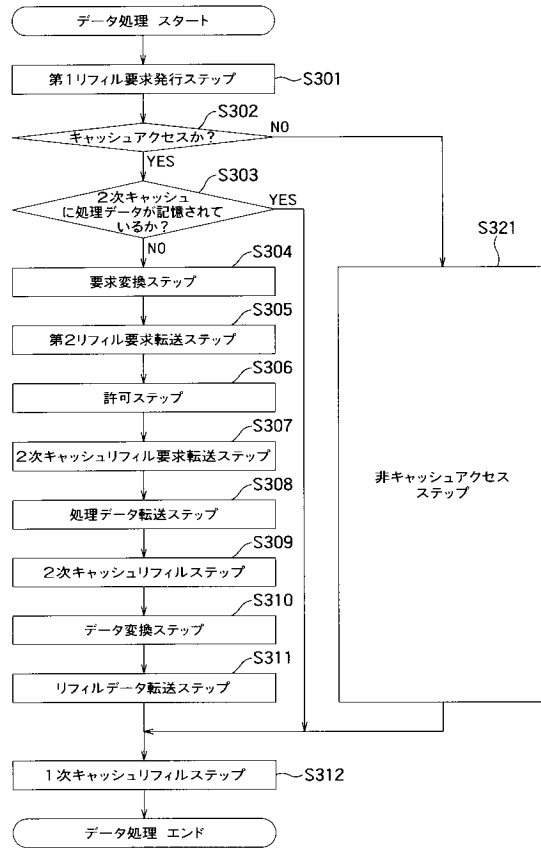
【図1】



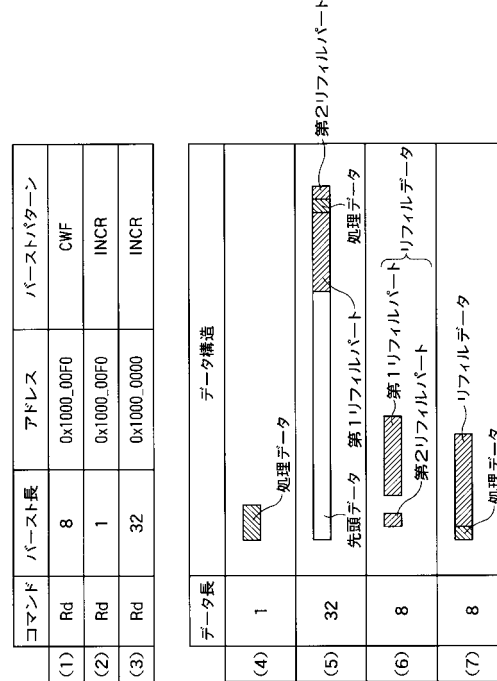
【図2】



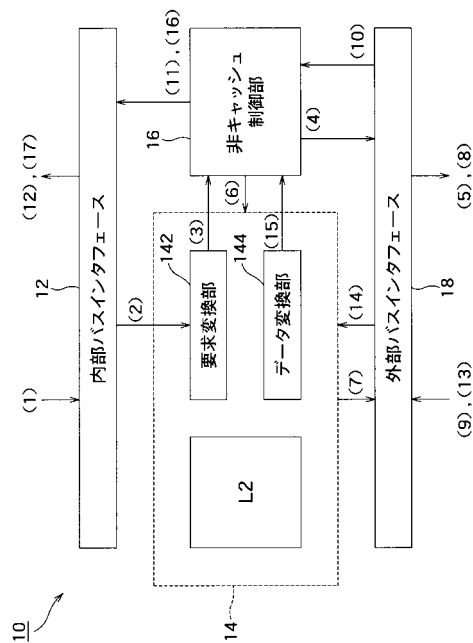
【図 3】



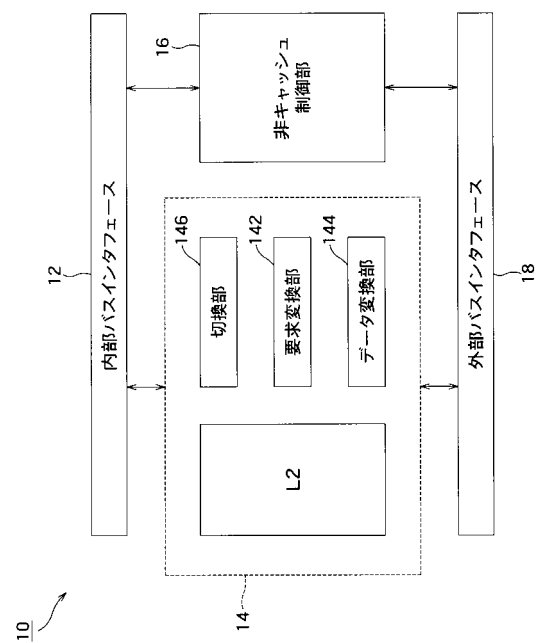
【図 4】



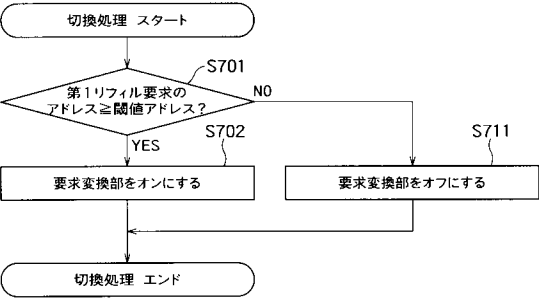
【図 5】



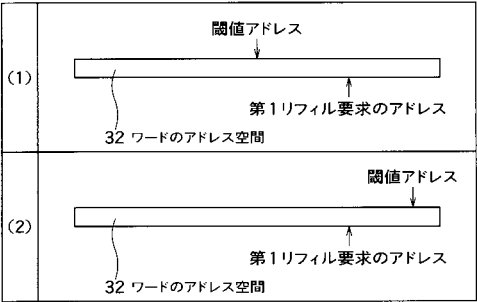
【図 6】



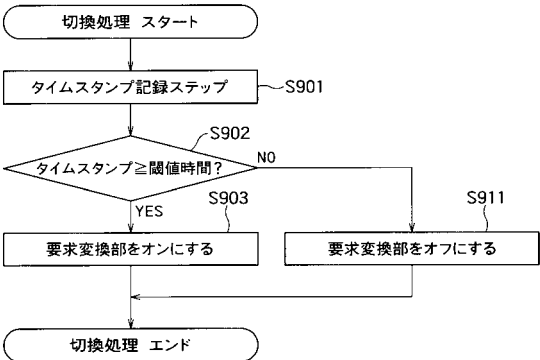
【図 7】



【図 8】



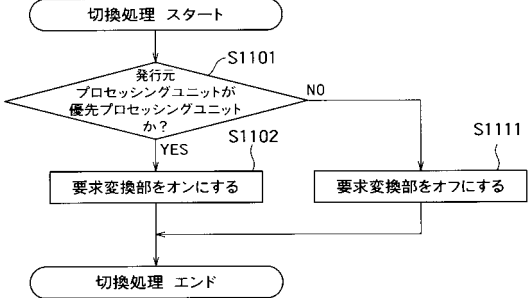
【図 9】



【図 10】

PUID	タイムスタンプ
0	T1
3	T2
1	T3
2	T4

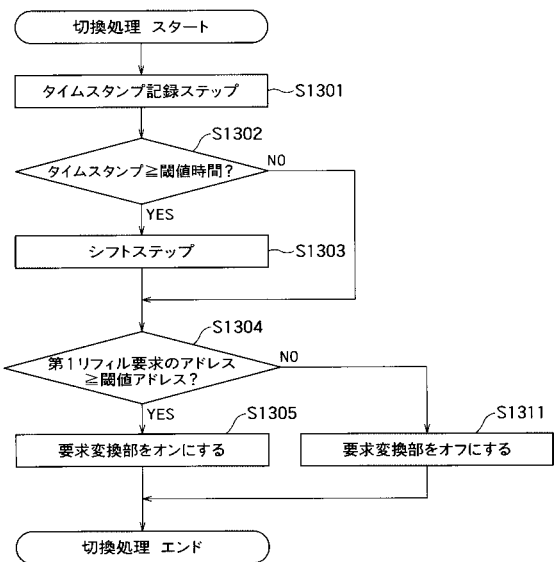
【図 11】



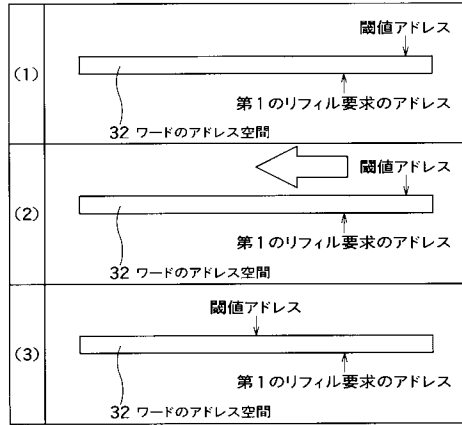
【図 12】

PUID
PU0

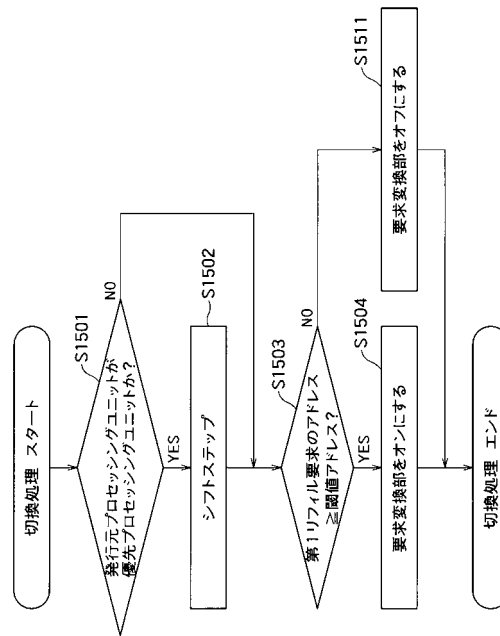
【図 13】



【図 14】



【図 15】



フロントページの続き

(72)発明者 細 田 宗一郎
東京都港区芝浦一丁目1番1号 株式会社東芝内

審査官 桜井 茂行

(56)参考文献 特表2008-510246 (JP, A)
特開2010-140140 (JP, A)
特開2009-217714 (JP, A)

(58)調査した分野(Int.Cl., DB名)
G06F 12/08-12/12
G06F 12/00-12/06
G06F 13/16-13/18