(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3564343号
(P3564343)
(45)発行日　平成16年9月8日(2004.9.8)　(24)登録日　平成16年6月11日(2004.6.11)

(51)Int.Cl.[7]　F I
G06F 12/08　G06F 12/08　517C
G06F 12/08　501D

請求項の数 3　(全 12 頁)

(21)出願番号　特願平11-334029
(22)出願日　平成11年11月25日(1999.11.25)
(65)公開番号　特開2001-154914(P2001-154914A)
(43)公開日　平成13年6月8日(2001.6.8)
審査請求日　平成12年10月11日(2000.10.11)

前置審査

(73)特許権者　000168285
エヌイーシーコンピュータテクノ株式会社
山梨県甲府市大津町１０８８－３
(74)代理人　100109313
弁理士　机　昌彦
(74)代理人　100085268
弁理士　河合　信明
(74)代理人　100111637
弁理士　谷澤　靖久
(72)発明者　矢崎　大輔
山梨県甲府市大津町１０８８－３
甲府日本電気株式会社内

審査官　相崎　裕恒

最終頁に続く

(54)【発明の名称】キャッシュバイパス時のデータ転送装置と方法

(57)【特許請求の範囲】

【請求項１】

主記憶装置とキャッシュメモリへのデータ転送巾が、キャッシュメモリと汎用レジスタへのデータ巾より大きい演算処理装置において、中央処理装置からのメモリアクセス命令に対するデータがバッファメモリ上のキャッシュメモリに存在しないときのキャッシュ制御において、メモリコントローラでの読み出しデータに対応するページテーブル中のキャッシュバイパス情報が登録しないことを示しているページ属性のバイパスビットがオンの場合、メモリコントローラを介して主記憶装置へのリクエストを目的データだけアクセスするように縮退させる手段を備え、目的データのみを主記憶装置からロードして、汎用レジスタに目的データを格納し、汎用レジスタ書き込みで主記憶装置からのリプライデータ輻輳が発生しないように主記憶装置に対し、タイマーにより予め決められた間隔でリクエストを発行するリクエスト制御回路を備え、汎用レジスタに目的データを格納させることを特徴とするキャッシュバイパス時のデータ転送装置。

【請求項２】

主記憶装置とキャッシュメモリへのデータ転送巾が、キャッシュメモリと汎用レジスタへのデータ巾より大きい演算処理装置において、中央処理装置からのメモリアクセス命令に対するデータがバッファメモリ上のキャッシュメモリに存在しないときのキャッシュ制御において、メモリコントローラでの読み出しデータに対応するページテーブル中のキャッシュバイパス情報が登録しないことを示しているページ属性のバイパスビットがオンの場合、メモリコントローラを介して主記憶装置へのリクエストを目的データだけアクセスす

るように縮退させ、目的データのみを主記憶装置からロードして、汎用レジスタに目的データを格納し、汎用レジスタ書き込みで主記憶装置からのリプライデータ輻輳が発生しないように主記憶装置に対し、タイマーにより予め決められた間隔でリクエストを発行し、汎用レジスタに目的データを格納させることを特徴とするキャッシュバイパス時のデータ転送方法。

【請求項３】

主記憶装置とキャッシュメモリへのデータ転送巾が、キャッシュメモリと汎用レジスタへのデータ巾より大きい演算処理装置において、中央処理装置からのメモリアクセス命令に対するデータがバッファメモリ上のキャッシュメモリに存在しないときのキャッシュ制御において、メモリコントローラでの読み出しデータに対応するページテーブル中のキャッシュバイパス情報が登録しないことを示しているページ属性のバイパスビットがオンの場合、メモリコントローラを介して主記憶装置へのリクエストを目的データだけアクセスするように縮退させ、目的データのみを主記憶装置からロードして、汎用レジスタに目的データを格納し、汎用レジスタ書き込みで主記憶装置からのリプライデータ輻輳が発生しないように主記憶装置に対し、タイマーにより予め決められた間隔でリクエストを発行し、汎用レジスタに目的データを格納させる処理をコンピュータに実行させるためのプログラムを記録したことを特徴とする記録媒体。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明はキャッシュバイパス時のデータ転送装置と方法に関し、特にキャッシュバイパス情報がオンのデータを主記憶装置に参照しに行く際、ブロックリードから目的のデータのみのリードにすることと、書き込み輻輳が発生しないように、リクエストの発行を制御するキャッシュバイパス時のデータ転送装置と方法に関する。

【０００２】

【従来の技術】

従来のキャッシュバイパス時のデータ転送方式は、本発明の実施の形態の図面を参照して説明する。

【０００３】

図１は本発明のキャッシュバイパス時のデータ転送装置と方法の一実施の形態を示すブロック図であり、図２はキャッシュバイパス時のデータ転送装置と方法のバッファメモリを示す構成図で、また図３はキャッシュバイパス時のデータ転送装置と方法のメモリコントローラを示す構成図である。

【０００４】

中央処理装置（１）より発行されたメモリアクセス命令に対応するデータがバッファメモリ（２）内のキャッシュメモリ（２７）に格納されていないキャッシュミスの場合、主記憶装置（４）からデータを読み出す処理として、論理アドレス（３１）から物理アドレス（３６）にアドレス変換し、同時に物理アドレス（３６）の管理サイズ毎に付随するページ属性をチェックした後、主記憶装置（４）からデータ供給を受けるデータ転送方式を持つ処理装置において、前記中央処理装置（１）からのメモリアクセス命令に対するデータが前記バッファメモリ（２）上のキャッシュメモリ（２７）に存在しない場合、メモリコントローラ（３）を介して（データ局所性原理により）目的データを含むブロックデータを主記憶装置（４）からロードする処理を行っているが、メモリコントローラ（３）でのページ属性のバイパスｂｉｔがオンの場合、目的データを含むブロックデータをロードするが、主記憶装置（４）から読み出されたデータをキャッシュメモリ（２７）に保持しないので、目的データを含むブロックデータを主記憶装置（４）から読み出しても、目的データ以外使用されずにデータ破棄していた。主記憶装置（４）の読み出しサイクルは、一般にマシンサイクルより中央処理装置の動作が遅いので殆ど使用されないブロックデータのロード処理に主記憶装置（４）が占有され、主記憶装置（４）の処理速度低下を招いていた。

【０００５】
中央処理装置（１）より発行されたメモリアクセス命令に対応するデータがバッファメモリ（２）内のキャッシュメモリ（２７）に格納されていないキャッシュミスの場合、主記憶装置（４）からデータを読み出す処理として、論理アドレス（３１）から物理アドレス（３６）にアドレス変換し、同時に物理アドレス（３６）の管理サイズ毎に付随するページ属性をチェックした後、主記憶装置（４）からデータ供給を受けるデータ転送方式を持つ処理装置で、かつ、主記憶装置（４）とキャッシュメモリ（２７）へのデータ転送巾が、キャッシュメモリ（２７）と汎用レジスタへのデータ巾より大きい処理装置において、主記憶装置（４）から複数個のデータを読み出して汎用レジスタ（１０）に格納する命令が中央処理装置（１）で実行され、命令に対するデータが前記バッファメモリ（２）上のキャッシュメモリ（２７）に存在せず、前記命令により主記憶装置（４）をアクセスするアドレスに対応するメモリコントローラ（３）のページ属性のバイパスｂｉｔがオンの場合、主記憶装置（４）からキャッシュメモリ（２７）へのデータ転送巾に較べて、キャッシュメモリ（２７）から汎用レジスタ（１０）のデータ転送巾の能力が小さいので、主記憶装置（４）から読み出された複数のデータがバッファメモリ（２）から汎用レジスタ（１０）への書き込みパスで書き込み輻輳が発生するため、複数データを主記憶装置（４）から読み出す場合、バッファメモリ（２）内のキャッシュメモリ（２７）もしくはキャッシュメモリ近傍のバッファにデータを保持して、バッファメモリ（２）から汎用レジスタ（１０）への書き込みで発生する書き込み輻輳を吸収しているためハードウェア量が増加する問題があった。

【０００６】
【発明が解決しようとする課題】
上述した従来のキャッシュバイパス時のデータ転送方式は、
第１の問題点は、目的データと使用されずにデータ破棄される目的データ以外の全てのデータを含むブロックデータを主記憶装置から読み出ため殆ど使用されないブロックデータのロード処理に主記憶装置が占有され、主記憶装置の処理速度低下を招いていた。

【０００７】
その理由は、主記憶装置の読み出しサイクルは、一般にマシンサイクルより中央処理装置の動作が遅いので殆ど使用されないブロックデータのロード処理に主記憶装置が占有されるためである。

【０００８】
第２の問題点は、複数データを主記憶装置から読み出す場合、メモリのハードウェア量が増加していた。

【０００９】
その理由は、複数データを主記憶装置から読み出す場合、主記憶装置からキャッシュメモリへのデータ転送巾に較べて、キャッシュメモリから汎用レジスタのデータ転送巾の能力が小さいので、バッファメモリ内のキャッシュメモリもしくはキャッシュメモリ近傍のバッファにデータを保持して、バッファメモリから汎用レジスタへの書き込みで発生する書き込み輻輳を吸収しているためである。

【００１０】
本発明の目的は、キャッシュバイパス情報がオンのデータを主記憶装置に参照しに行く際、ブロックリードから目的のデータのみのリードにすることにより、後続命令を早く行えるようにし、性能向上をはかると共に、メモリの負荷を少なくすることができるデータ転送方式と、汎用レジスタで書き込み輻輳が発生しないように、リクエストの発行を制御するリクエスト制御回路を設けることにより、汎用レジスタで書き込み輻輳の発生を防ぎ、従来書き込み輻輳が発生した際、キャッシュメモリ近傍のバッファに一時的に保存していたハードウェア量の削減ができるキャッシュバイパス時のデータ転送装置と方法を提供することにある。

【００１１】
【課題を解決するための手段】

本発明のキャッシュバイパス時のデータ転送装置と方法は、主記憶装置とキャッシュメモリへのデータ転送巾が、キャッシュメモリと汎用レジスタへのデータ巾より大きい演算処理装置において、中央処理装置からのメモリアクセス命令に対するデータがバッファメモリ上のキャッシュメモリに存在しないときのキャッシュ制御において、メモリコントローラでの読み出しデータに対応するページテーブル中のキャッシュバイパス情報が登録しないことを示しているページ属性のバイパスビットがオンの場合、メモリコントローラを介して主記憶装置へのリクエストを目的データだけアクセスするように縮退させる手段を備え、目的データのみを主記憶装置からロードして、汎用レジスタに目的データを格納し、汎用レジスタ書き込みで主記憶装置からのリプライデータ輻輳が発生しないように主記憶装置に対し、タイマーにより予め決められた間隔でリクエストを発行するリクエスト制御回路を備え、汎用レジスタに目的データを格納させる。

【００１２】

【発明の実施の形態】

次に、本発明の実施の形態について図面を参照して説明する。

【００１３】

図１は本発明のキャッシュバイパス時のデータ転送装置と方法の一実施の形態を示すブロック図であり、図２はキャッシュバイパス時のデータ転送装置と方法のバッファメモリを示す構成図で、また図３はキャッシュバイパス時のデータ転送装置と方法のメモリコントローラを示す構成図である。

【００１４】

図１、図２、および図３を参照すると、本発明の第１の実施の形態は、複数の演算処理装置（５）が主記憶装置（４）に接続されるマルチプロセッサ構成において、各演算処理装置（５）は中央処理装置（１）、バッファメモリ（２）、メモリコントローラ（３）から構成されている。中央処理装置（１）は、データ処理を管理する処理装置である。バッファメモリ（２）は、主記憶装置（４）の写しであるキャッシュメモリ（２７）と、キャッシュメモリ（２７）に対応するアドレスを保持するタグメモリ（２４）と、キャッシュメモリ（２７）に目的データが存在するかのチェックを行う比較回路（２５）および制御部（２２）等により構成され、メモリコントローラ（３）は、論理アドレス（３１）から物理アドレス（３６）へのアドレス変換を行うアドレス変換部と、主記憶装置（４）へのリクエスト制御を管理するリクエスト生成部（３８）およびアドレス生成を行うアドレス生成部（３７）より構成される。

【００１５】

バッファメモリ（２）において中央処理装置（１）から主記憶リードリクエスト（１０１）とアドレス（１０２）を受信すると、バッファメモリ（２）内の制御部（２２）とアドレスバッファ（２３）にセットし、アドレスバッファ（２３）にセットされたアドレス（２０３）でキャッシュメモリ（２７）に目的データ（１１０）が存在するかチェックを行う。

【００１６】

具体的には、アドレスバッファ（２３）の下位ｂｉｔ（２０２）でタグメモリ（２４）をサーチし、タグメモリ（２４）の該当するデータ（２０４）とアドレスバッファ（２３）の上位ｂｉｔ（２０３）との比較を比較回路（２５）で行い、一致した（ＨＩＴ）か一致しない（ＭＩＳＳ）の情報（２０５）を制御部（２２）へ通知する。一致した（ＨＩＴ）ら目的データ（１１０）がキャッシュメモリ（２７）に存在するので、制御部（２２）の指示（２０９）によりキャッシュメモリ（２７）から読み出す処理を行う。不一致（ＭＩＳＳ）なら、目的データ（１１０）がキャッシュメモリ（２７）に存在しないので、制御部（２２）の指示によりメモリコントローラ（３）を経由して主記憶装置（４）に対してリクエスト（１０３）を発行する。

【００１７】

一般に、主記憶リードリクエスト（１０１）は目的データ（１１０）に隣接したデータが必要になる可能性が高い（局所性原理）。このため、目的データ（１１０）に隣接データ

も含むブロックデータを読み出す処理を行い、読み出したブロックデータをキャッシュメモリ（２７）に登録する処理を行っている。主記憶装置（４）からのリプライデータ（１０８）は通常、目的データ（１１０）を含むブロックデータリプライであり、主記憶装置（４）から複数のリプライデータ（１０８）がバッファメモリ（２）に送付される。

【００１８】

バッファメモリ（２）は、メモリコントローラ（３）から送付されたキャッシュバイパス情報（１０５）を受信し、キャッシュバイパス情報（１０５）がオンの場合、主記憶装置（４）からはブロックデータではなく目的データ（１０８）しかバッファメモリ（２）に戻らないので、目的データリプライタグ（１０９）の受信により制御部（２２）の動作が完了するように動作させる。

【００１９】

ページング方式を用いて主記憶装置（４）にアクセスする情報処理装置では、アドレス変換が必要であり本発明ではメモリコントローラ（３）で行う。アドレス変換は、中央処理装置（１）から送付される論理アドレス（３１）中のページデータ（上位ｂｉｔ）（３０２）でメモリコントローラ（３）内アドレス変換バッファ（３３）をサーチし、アドレス変換バッファ（３３）の該当するページテーブル（３３－ａ）から、主記憶装置（４）にアクセスするページアドレス（３０４）を読み出し、中央処理装置（１）からの論理アドレス（３１）のページ内データ（下位ｂｉｔ）（３５）と上記ページデータ（３４）を加えて、物理アドレス（３６）を生成する。

【００２０】

バッファメモリ（２）より送付されたリクエスト（１０３）は目的データ（１１０）のみをアクセスするリクエスト（１０３）であるため、目的データ（１１０）を含むブロックデータを主記憶装置（４）から読み出すため、主記憶装置（４）をアクセスする個数分の複数アクセスリクエスト（１０７）をメモリコントローラ（３）内リクエスト生成部（３８）で生成し、主記憶装置（４）にリクエスト（１０７）送付を行っている。

【００２１】

同様にアドレス生成部（３７）で生成された物理アドレス（３６）もアドレス生成部（３７）で主記憶アクセス毎に物理アドレス（１０６）が生成され、リクエスト（１０７）と対で主記憶装置（４）に送付される。

【００２２】

メモリコントローラ（３）ではアドレス変換と同時にページアドレス毎の管理情報もアドレス変換バッファ（３３）から読み出され、その中のキャッシュバイパス情報（１０５）はメモリコントローラ（３）内のリクエスト生成部（３８）とバッファメモリ（２）へ送付される。

【００２３】

複数のリクエスト（１０７）を生成するリクエスト生成部（３８）には、アドレス変換バッファ（３３）内のページテーブル（３３－ａ）にあるキャッシュバイパス情報（１０５）が報告される。キャッシュバイパス情報（１０５）がオンである場合、主記憶装置（４）から読み出されたデータ（１０８）は、キャッシュメモリ（２７）には登録しないため、目的データ（１１０）を含むブロックデータを読み出す処理を行っても、目的データ（１１０）以外のデータは利用されずに破棄されていた。そこで、リクエスト生成部（３８）ではキャッシュバイパス情報（１０５）がオンである場合、目的データ（１１０）のみをアクセスするリクエスト（１０７）に縮退動作させる。

【００２４】

次に、本発明の第１の実施の形態の動作について詳細に説明する。

【００２５】

図１において、メモリロード命令を処理する場合、中央処理装置（１）は、バッファメモリ（２）に対し、主記憶リードリクエスト（１０１）とアドレス（１０２）を発行する。バッファメモリ（２）は、中央処理装置（１）より主記憶リードリクエスト（１０１）とアドレス（１０２）をアドレスを制御部（２２）とアドレスバッファ（２３）に受信する

。制御部（２２）は、リクエストが主記憶リードリクエストと判断したので、キャッシュメモリ（２７）に目的データが格納されているかチェックを行う。具体的には、アドレスバッファ（２３）の下位ｂｉｔで信号線２０２を介しタグメモリ（２４）をサーチし、信号線２０４を介しタグメモリ（２４）の該当データと信号線２０３を介しアドレスバッファ（２３）の上位ｂｉｔを比較回路（２５）で比較し、キャッシュヒットか、キャッシュミスかを区別し、キャッシュミスと判断した場合、制御部（２２）よりメモリコントローラ（３）へ信号線１０３を介し前記リクエストと信号線１０４を介し前記アドレスを発信する。

【００２６】

メモリコントローラ（３）は、バッファメモリ（２）より信号線１０３を介しリクエスト（３０）と信号線１０４を介し論理アドレス（３１）を受信すると、論理アドレス（３１）の上位アドレスにより信号線３０２を介しアドレス変換バッファ（３３）の該当するページテーブル（３３－ａ）からデータパス３０４を介しデータを読み出し、信号線３０５を介し読み出したデータ（３４）に前記論理アドレス（３１）の変位（３２）と、前記リクエスト（３０）とキャッシュバイパス情報（信号線１０５）をそれぞれ信号線３０１と信号線１０５を介して読み出された信号を加えて物理アドレス（３６）を生成する。

【００２７】

アドレス変換バッファ（３３）の該当するデータを信号線３０４を介して読み出すとの同時に、該当するテーブルにはページ情報も格納されており、キャッシュメモリ（２７）にデータを格納しないキャッシュバイパス情報（信号線１０５）も含まれている。

【００２８】

主記憶装置（４）は一度に、決められたデータ巾の処理しか行わないので、リクエスト生成部（３８）では、ブロックデータリクエストから、主記憶装置（４）が処理出来るデータ巾に分割した複数のリクエストを生成して、主記憶装置（４）に信号線１０７を介して発信する。そのため、キャッシュバイパス情報（信号線１０５）がオンの場合、目的データのみをロードするリクエスト命令の縮退をする回路をメモリコントローラ（３）は有している。複数のリクエスト生成時、リクエスト管理のため、タグ情報生成も同時に行う。

【００２９】

目的データを含むリクエストに最若番のタグを付与るるようにタグ生成を行い、主記憶装置（４）に信号線１０７を介して送付する。

【００３０】

主記憶装置（４）は、メモリコントローラ（３）より信号線１０７を介して受信したリクエストの処理を行い、リクエストがメモリロードの場合、信号線１０６を介して受信したアドレスで示されるメモリよりデータを読み出してデータパス１０８を介して返送し、メモリコントローラ（３）からの信号線１０７を介して受信したタグを付与して信号線１０９を介してバッファメモリ（２）に返送する。

【００３１】

前記リプライタグは、バッファメモリ（２）内の信号線１０９を介して制御部（２２）へ送られ、メモリコントローラ（３）より受信したバイパス情報（信号線１０５）と合わせて主記憶装置（４）からのリプライデータ制御を行う。

【００３２】

バッファメモリ（２）でのタグ管理は、目的データを含むリクエストに最若番のタグを付与したるものとして動作する。主記憶アクセスは通常、目的データを含むブロックデータリプライであり、主記憶装置（４）から複数のリプライデータがデータパス１０８を介してバッファメモリ（２）に送付される。

【００３３】

モリコントローラ（３）から送付されたキャッシュバイパス情報（信号線１０５）がオンの場合、主記憶装置（４）から目的データしかバッファメモリ（２）にデータパス１０８を介して戻らないので目的データリプライを受信しただけで、制御部（２２）の動作が完了するように動作させる。

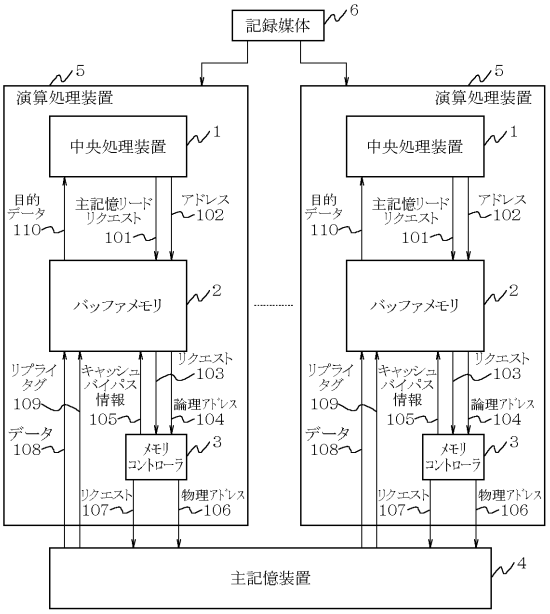
【００３４】
次に本発明の第２の実施の形態について説明すると、主記憶装置（４）とキャッシュメモリ（２７）へのデータ転送巾が、キャッシュメモリ（２７）と汎用レジスタ（１０）へのデータ転送巾より大きい多数の演算処理装置（５）に主記憶装置（４）が接続されるマルチプロセッサ構成において、各演算処理装置（５）は中央処理装置（１）、バッファメモリ（２）、メモリコントローラ（３）から構成されている。
【００３５】
中央処理装置（１）は、データ処理を管理する処理装置、バッファメモリ（２）は、主記憶装置の写しであるキャッシュメモリ（２７）と、キャッシュメモリ（２７）に対応するアドレスを保持するタグメモリ（２４）と、キャッシュメモリ（２７）に目的データ（１１０）が存在するかのチェックを行う比較回路（２５）および制御部（２２）等により構成され、メモリコントローラ（３）は、論理アドレス（３１）から物理アドレス（３６）へのアドレス変換を行うアドレス変換部と、主記憶装置へのアクセス制御を管理するリクエスト生成部（３８）およびリクエストと対をなすアドレス生成部（３７）より構成される。
【００３６】
バッファメモリ（２）の構成で、キャッシュメモリ（２７）は、レジスタ（２６）を挟んで読み出し速度が低速であるが、大容量の情報が格納できる主記憶装置（４）と、レジスタ（２８）を挟んで読み出しを高速に処理出来る中央処理装置（１）内の汎用レジスタ（１０）に接続されており、情報供給量のネックを解消するために主記憶装置（４）の情報の写しを持つ役割を果たしている。
【００３７】
中央処理装置（１）から主記憶リードリクエスト（１０１）とアドレス（１０２）を受信すると、第１の実施の形態で示したようにバッファメモリ（２）内のアドレスバッファ（２３）に格納されたアドレス（２０３）に対応すえるデータがキャッシュメモリ（２７）に存在するかチェックを行い、存在していなければ制御部（２２）の指示によりメモリコントローラ（３）を経由して主記憶装置（４）に対してデータ読み出しリクエスト（１０３）を行う。
【００３８】
主記憶装置（４）から複数個のデータ（１０８）を読み出して汎用レジスタ（１０）に格納する命令の場合も、複数読み出しリクエスト（１０３）として、メモリコントローラ（３）に発行される。
【００３９】
メモリコントローラ（３）の構成で、メモリコントローラ（３）では、実施例１と同様に、アドレス変換と同時にページアドレス毎の管理情報もアドレス変換バッファ（３３）から読み出され、その中のキャッシュバイパス情報（１０５）は メモリコントローラ（３）内のリクエスト生成部（３８）とバッファメモリ（２）へ送付される。主記憶装置（４）に対してデータ読み出しリクエスト（１０３）を受信したメモリコントローラ（３）は、第１の実施の形態で示したように、バッファメモリ（２）から受信した目的データ（１１０）のみを読み出すリクエストをブロックデータ読み出しリクエストとして扱い、主記憶装置（４）が処理出来るデータ巾に分割した複数のリクエスト（１０７）を生成して、主記憶装置（４）に発行する。
【００４０】
主記憶装置（４）からバッファメモリ（２）へのリプライデータ処理は、データ転送巾の広いデータパス１０８を使用して転送するので、ブロックデータリプライの複数データ（１０８）が同時に転送される。主記憶装置（４）からバッファメモリ（２）内のキャッシュメモリ（２７）間のデータ転送巾は、キャッシュメモリ（２７）から中央処理装置（１）内の汎用レジスタ（１０）間のデータ転送巾より大きいため、汎用レジスタ（１０）へ複数のデータ（１０８）を書き込む動作の場合、データ転送巾の小さいキャッシュメモリ（２７）汎用レジスタ（１０）間で書き込み輻輳が発生する。

【００４１】
従来技術では、書き込み輻輳を解消するため、キャッシュメモリ（２７）近傍のレジスタに一時的にデータを格納していたが、この方法ではハードウェア量が増大するという問題を抱えていた。そこで、複数のリクエスト（１０７）を生成するリクエスト生成部（３８）において、複数読み出しリクエスト（１０７）で、リクエスト（１０７）に対応するキャッシュバイパス情報（１０５）がオンである場合、汎用レジスタ（１０）への書き込み輻輳が発生する可能性があるため、書き込み輻輳の発生しない間隔でリクエスト（１０７）発行を行う回路を設けることで、書き込み輻輳の発生を押さえる。
【００４２】
次に、本発明の第２の実施の形態の動作について詳細に説明する。
【００４３】
図１において、メモリロード命令の場合、中央処理装置（１）は、バッファメモリ（２）に対し、信号線１０１を介し主記憶リードリクエストと信号線１０２を介しアドレスを発行する。バッファメモリ（２）は、中央処理装置（１）より信号線１０１を介し主記憶リードリクエストと信号線１０２を介しアドレスを制御部（２２）とアドレスバッファ（２３）に受信する。制御部（２２）は、リクエストが主記憶リードリクエストと判断したので、信号線２０９を介しキャッシュメモリ（２７）に目的データが格納されているかチェックを行う。
【００４４】
第１の実施の形態の動作説明のようにしてキャッシュヒットか、キャッシュミスかを区別し、キャッシュミスと判断した場合、制御部（２２）よりメモリコントローラ（３）へ信号線１０３を介し前記リクエストと信号線１０４を介し前記アドレスを発信する。
【００４５】
メモリコントローラ（３）は、バッファメモリ（２）より信号線１０３を介しリクエストと信号線１０４を介しアドレスを受信すると、第１の実施の形態の動作説明のようにして物理アドレス（３６）を生成する。主記憶装置（４）は一度に決められたデータ巾の処理しか行えないので、リクエスト生成部（３８）では、信号線３０１を介して受信されるブロックデータリクエストから、主記憶装置（４）が処理出来るデータ巾に分割した複数のリクエストを生成して、信号線１０４を介し主記憶装置（４）に発信する。
【００４６】
データパス１０８を介して送信される主記憶装置（４）からレジスタ（２６）を通してデータパス２０６を介しバッファメモリ（２）内のキャッシュメモリ（２７）間のデータ転送巾に比べ、データパス１１０を介して送信されるバッファメモリ（２）内キャッシュメモリ（２７）と汎用レジスタ（１０）へデータを転送するレジスタ（２８）と中央処理装置（１）内の汎用レジスタ（１０）間のデータ転送能力が小さいために、書き込み輻輳が発生するので、複数読み出しリクエストで、リクエストに対応するキャッシュバイパス情報（信号線１０５）がオンである場合、書き込み輻輳の発生しない間隔でリクエスト発行を行うように動作させ、書き込み輻輳の発生を押さえる。書き込み輻輳の発生しない間隔でのリクエスト発行を行うように動作させる一つの方法としてはタイマーによる予じめ決められた間隔でのリクエスト発行が実現できる。
【００４７】
以上はに図１に示すような複数の演算処理装置で主記憶装置を共有するキャッシュバイパス時のデータ転送装置と方法で説明したが、演算処理装置と主記憶装置が一対一で接続されているキャッシュバイパス時のデータ転送装置と方法においても同様に適応することができる。
【００４８】
尚以上のような処理プログラムを記録した記録媒体（６）を有しコンピュータに実行させることもできる。
【００４９】
【発明の効果】

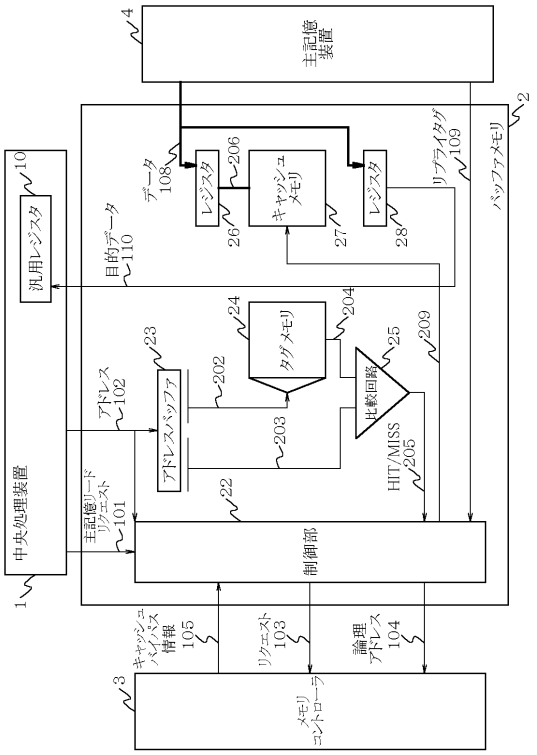
以上説明したように、本発明のキャッシュバイパス時のデータ転送装置と方法は、
第１の効果は、目的のデータのみをアクセスするリクエストにすることにより目的のデータのみを汎用レジスタにリードし、中央処理装置にロードすることで、後続命令を早く行えるようにし性能向上をはかると共に、メモリの負荷を少なくすることが出来る。
【００５０】
その理由は、キャッシュバイパス情報がオンである場合のメモリロード命令について、ブロックデータを読み出す処理を行うのではなく、メモリコントローラ内のリクエスト生成部にてメモリアクセス命令を縮退させ、目的のデータのみをアクセスするリクエストにすることにより目的のデータのみを汎用レジスタにリードし、中央処理装置にロードしているからである。
【００５１】
第２の効果は、キャッシュバイパス情報がオンである場合のメモリロード命令について、主記憶装置からバッファメモリ間のデータ転送巾に比べ、バッファメモリと中央処理装置間のデータ転送能力が小さいため発生する中央処理装置内汎用レジスタへのデータ書き込み輻輳を吸収するために、一時的にデータを格納するキャッシュメモリ近傍のレジスタを設ける必要がなくなりハードウェア量の増大を防ぐことが出来る。
【００５２】
その理由は、汎用レジスタで書き込み輻輳が発生しないように、メモリコントローラ内にあるリクエスト生成部で主記憶装置へのリクエスト発行間隔を制御するリクエスト制御回路を設けることで、書き込み輻輳の発生を押さえるからである。
【図面の簡単な説明】
【図１】本発明のキャッシュバイパス時のデータ転送装置と方法の一実施の形態を示すブロック図である。
【図２】本発明のキャッシュバイパス時のデータ転送装置と方法のバッファメモリを示す構成図である。
【図３】本発明のキャッシュバイパス時のデータ転送装置と方法のメモリコントローラを示す構成図である。
【符号の説明】
１　　中央処理装置
２　　バッファメモリ
３　　メモリコントローラ
４　　主記憶装置
５　　演算処理装置
６　　記録媒体
１０　　汎用レジスタ
２２　　制御部
２３　　アドレスバッファ
２４　　タグメモリ
２５　　比較回路
２６，２８　　レジスタ
２７　　キャッシュメモリ
３０　　リクエスト
３１　　論理アドレス
３２　　変位
３３　　アドレス変換バッファ
３３－ａ　　ページテーブル
３４　　ページデータ
３５　　ページ内データ
３６　　物理アドレス
３７　　アドレス生成部

３８　　リクエスト生成部

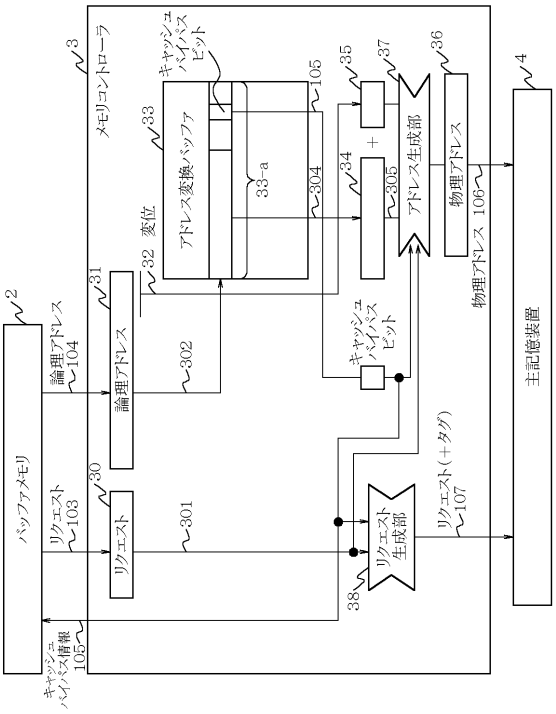
【図１】



【図２】



【図３】


バッファメモリ
2
論理アドレス
104
リクエスト
103
キャッシュバイパス情報
105
3
メモリコントローラ
31
論理アドレス
302
30
リクエスト
301
32
変位
33
アドレス変換バッファ
33-a
キャッシュバイパスビット
105
304
34
35
+
305
37
アドレス生成部
36
物理アドレス
106
物理アドレス
キャッシュバイパスビット
38
リクエスト生成部
リクエスト(+タグ)
107
4
主記憶装置

フロントページの続き

(56)参考文献 特開平０１－１０８６５１（ＪＰ，Ａ）
特開平０９－３１９６５４（ＪＰ，Ａ）
特開平０２－１１４３４２（ＪＰ，Ａ）
特開平０６－３４８５９３（ＪＰ，Ａ）
特開昭６０－１５３５５４（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.[7]，ＤＢ名)
G06F 12/08
G06F 12/00