

(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2012-146201

(P2012-146201A)

(43) 公開日 平成24年8月2日(2012. 8. 2)

(51) Int.Cl.

G06F 15/173 (2006.01)

F I

G06F 15/173 640C

テーマコード (参考)

5B045

審査請求 未請求 請求項の数 5 O L (全 15 頁)

(21) 出願番号 特願2011-5106 (P2011-5106)
 (22) 出願日 平成23年1月13日 (2011. 1. 13)

(71) 出願人 000003078
 株式会社東芝
 東京都港区芝浦一丁目1番1号
 (74) 代理人 100117787
 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 100152205
 弁理士 吉田 昌司

最終頁に続く

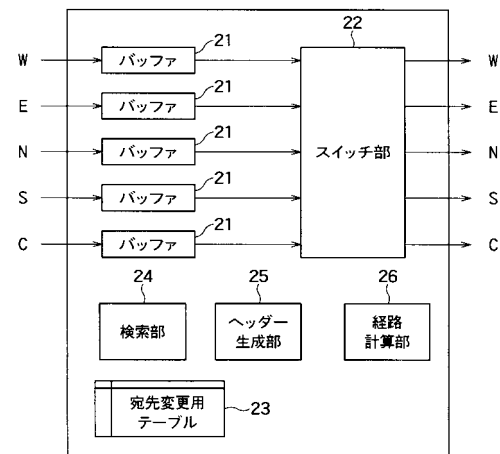
(54) 【発明の名称】 オンチップルータ及びそれを用いたマルチコアシステム

(57) 【要約】

【課題】コア間通信用のバッファを新たに設けることなく、コア間通信のレイテンシを低減可能なオンチップルータ及びそれを用いたマルチコアシステムを提供する。

【解決手段】実施形態に係るオンチップルータは、宛先変更用テーブルと、検索部と、経路計算部と、ヘッダ生成部とを備える。宛先変更用テーブルは、共有メモリ上の宛先アドレスと、宛先コアに接続されたルータの識別子とを対応させる。検索部は、書き込み要求パケットの宛先アドレスと同一のアドレスを有する有効なエントリが宛先変更用テーブルに存在するか否かを検索する。経路計算部は、検索の結果、有効なエントリが前記宛先変更用テーブルに存在すれば、そのエントリの識別子のルータを宛先として経路計算を行う。ヘッダ生成部は、計算されたルータの識別子を書き込み要求パケットのヘッダに格納し、かつ宛先を変更したことを示すルーテッドフラグを立てる。

【選択図】図4



【特許請求の範囲】**【請求項 1】**

パケットを受信する複数の入力ポートと、
前記パケットを送信する複数の出力ポートと、
前記入力ポートのそれぞれに設けられており、前記受信したパケットのデータを格納する複数のバッファと、
前記パケットが所望の出力ポートから送出されるように、パケットの出力先を切り替えるスイッチ部と、
送信元コアにより設定された共有メモリ上の宛先アドレスとルータの識別子とを対応させて格納する宛先変更用テーブルと、
受信した書き込み要求パケットに格納された共有メモリ上の宛先アドレスと同一のアドレスを有する有効なエントリが前記宛先変更用テーブルに存在するかどうかを検索する第 1 の検索を行う検索部と、
前記第 1 の検索の結果、前記宛先変更用テーブルに存在する有効なエントリのルータを宛先として経路計算を行う経路計算部と、
前記経路計算部により求められた経路を示すルータの識別子を、前記書き込み要求パケットに格納し、かつ宛先を変更したことを示すリルーテッドフラグを立てるヘッダー生成部と、を備えることを特徴とするオンチップルータ。

10

【請求項 2】

前記検索部は、前記宛先コアから受信した読み出し要求パケットに格納された共有メモリ上の宛先アドレスと同一のアドレスを有し、かつ、前記リルーテッドフラグが立てられた書き込み要求パケットが、前記バッファに蓄積されているかどうかを検索する第 2 の検索を行い、
前記ヘッダー生成部は、前記第 2 の検索の結果、該当する書き込み要求パケットが存在すれば、その書き込み要求パケットのヘッダーを書き換え、前記読み出し要求パケットに対する応答パケットのヘッダーを生成する、
ことを特徴とする請求項 1 に記載のオンチップルータ。

20

【請求項 3】

前記ヘッダー生成部は、前記書き込み要求パケットのリルーテッドフラグが立っている場合、1 サイクル経過する毎に、前記バッファに蓄積された前記書き込み要求パケットのライフサイクル値を減じ、
前記経路計算部は、前記ライフサイクル値が所定の値まで減少すると、前記書き込み要求パケットの共有メモリ上の宛先アドレスを宛先として経路計算を行い、
前記ヘッダー生成部は、その結果求められたルータ識別子を前記書き込み要求パケットに格納し、前記リルーテッドフラグを倒す、
ことを特徴とする請求項 1 または 2 のいずれかに記載のオンチップルータ。

30

【請求項 4】

転送先のルータのバッファに空きがない場合、前記経路計算部はパケットに格納された共有メモリ上の宛先アドレスを宛先として経路計算を行い、前記ヘッダー生成部は前記リルーテッドフラグを倒すことを特徴とする請求項 1 乃至 3 のいずれかに記載のオンチップルータ。

40

【請求項 5】

複数のプロセッサコアと、前記プロセッサコアに接続され、前記プロセッサコア間で送受信されるパケットを転送する複数のオンチップルータとを備えるマルチコアシステムであって、
前記オンチップルータは、
前記パケットを受信する複数の入力ポートと、
前記パケットを送信する複数の出力ポートと、
前記入力ポートのそれぞれに設けられており、前記受信したパケットのデータを格納する複数のバッファと、

50

前記パケットが所望の出力ポートから送出されるように、パケットの出力先を切り替えるスイッチ部と、

自身が接続された前記プロセッサコアにより設定された共有メモリ上の宛先アドレスと、ルータの識別子とを対応させて格納する宛先変更用テーブルと、

受信した書き込み要求パケットに格納された共有メモリ上の宛先アドレスと同一のアドレスを有する有効なエントリが前記宛先変更用テーブルに存在するかどうかを検索する第1の検索を行う検索部と、

前記第1の検索の結果、前記宛先変更用テーブルに存在する有効なエントリのルータを宛先として経路計算を行う経路計算部と、

前記経路計算部により求められた経路を示すルータの識別子を、前記書き込み要求パケットに格納し、かつ宛先を変更したことを示すリルーテッドフラグを立てるヘッダー生成部とを備える、

10

ことを特徴とするマルチコアシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明の実施形態は、オンチップルータ及びそれを用いたマルチコアシステムに関する。

【背景技術】

【0002】

20

共有メモリ型のマルチコアシステムでは、プロセッサコア（以下、「コア」）間の通信は共有メモリを介して行われる。まず、コアAが書き込み要求パケットをオンチップルータ（以下、「ルータ」）により共有メモリに送信する。そして、書き込み要求パケットに含まれるデータが共有メモリ上の所定アドレスに書き込まれる。その後、所定アドレスのデータを必要とするコアBがデータを読み出すための読出し要求パケットを共有メモリに送信し、コアAが書き込んだデータを取得する。

【0003】

このように共有メモリ型マルチコアシステムにおけるコア間通信では、共有メモリを介してデータをやりとりするため、一般にレイテンシが大きいという問題がある。さらに、コア間通信による共有メモリの一時的な使用が、共有メモリを圧迫するという問題もある。

30

【0004】

一方、いわゆるメッセージパッシング型と呼ばれるコア間通信方式では、共有メモリを介さずに、コアAからコアBにデータを直接送信する。この方式の場合、レイテンシを低減させることができるが、次の問題がある。コア間通信に用いる専用バッファをコアごとに設ける必要があるため、実装コストの増大、チップ面積の増加、消費電力の増大という問題が発生する。さらに、専用バッファには、他のコアからアクセス可能なようにグローバルアドレスを割り当てる必要があるため、限られたアドレス空間を圧迫するという問題もある。

【先行技術文献】

40

【特許文献】

【0005】

【特許文献1】特開2009-37273号公報

【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、コア間通信用のバッファを新たに設けることなく、コア間通信のレイテンシを低減可能なオンチップルータ及びそれを用いたマルチコアシステムを提供することである。

【課題を解決するための手段】

50

【 0 0 0 7 】

一実施形態に係るオンチップルータは、共有メモリ型のマルチコアシステムにおいてコア間通信を行うオンチップルータであって、バッファと、スイッチ部と、宛先変更用テーブルと、検索部と、経路計算部と、ヘッダー生成部と、を備える。

【 0 0 0 8 】

前記バッファは、各入力ポートに設けられており、前記入力ポートで受信したデータを格納する。前記スイッチ部は、受信したパケットが所望の出力ポートから送出されるように、パケットの出力先を切り替える。前記宛先変更用テーブルは、共有メモリ上の宛先アドレスと、宛先コアに直接接続されたルータの識別子とを対応させており、送信元コアにより設定される。前記検索部は、前記送信元コアから受信した書き込み要求パケットのヘッダーに格納された前記共有メモリ上の宛先アドレスと同一のアドレスを有する有効なエントリが前記宛先変更用テーブルに存在するかどうかを検索する第1の検索を行う。前記経路計算部は、前記第1の検索の結果、前記有効なエントリが前記宛先変更用テーブルに存在すれば、そのエントリの識別子のルータを宛先としてソースルーティング方式により経路計算を行う。前記ヘッダー生成部は、前記経路計算により求められたルータの識別子を、前記書き込み要求パケットのヘッダーのホップルータフィールドに格納し、かつ宛先を変更したことを示すリルーテッドフラグを立てる。

10

【 図面の簡単な説明 】

【 0 0 0 9 】

【 図 1 】一実施形態に係る共有メモリ型のマルチコアシステムの概略的な構成を示す図である。

20

【 図 2 】マルチコアシステムのコア間通信で用いられるパケットの構造を示す図である。

【 図 3 】ソースルーティング方式におけるパケットのヘッダーの詳細を示す図である。

【 図 4 】一実施形態に係るルータの構成を示す図である。

【 図 5 】宛先変更用テーブルの一例を示す図である。

【 図 6 】送信元コアと直接接続されたルータにより生成されたヘッダーの詳細を示す図である。

【 図 7 】第1の実施形態におけるパケットの転送経路の一例を示す図である。

【 図 8 】適応型ルーティング方式におけるパケットのヘッダーの詳細を示す図である。

【 図 9 】第3の実施形態に係るルータの構成を示す図である。

30

【 図 1 0 】第4の実施形態におけるパケットの転送経路の一例を示す図である。

【 図 1 1 】第5の実施形態に係る宛先変更用テーブルの一例を示す図である。

【 図 1 2 】第5の実施形態におけるパケットの転送経路の一例を示す図である。

【 発明を実施するための形態 】

【 0 0 1 0 】

以下、本発明に係る実施形態について図面を参照しながら説明する。なお、各図において同等の機能を有する構成要素には同一の符号を付し、同一符号の構成要素の詳細な説明は繰り返さない。

【 0 0 1 1 】

まず、各実施形態に共通するマルチコアシステムの構成およびコア間通信に用いられるパケットの概略的な構成について説明する。

40

【 0 0 1 2 】

図1は、本実施形態に係る共有メモリ型のマルチコアシステムの概略的な構成を示している。図1のマルチコアシステムは、16個のコア101～116と、20個のルータ201～220と、4個の共有メモリ301～304と、を含んでいる。

【 0 0 1 3 】

ルータ201～220は、格子状に配置され、メッシュ型のネットワークを構成している。ルータ201～216は対応するコア101～116と接続されており、ルータ217～220は対応する共有メモリ301～304と接続されている。例えば、コア101にはルータ201が接続されており、共有メモリ301にはルータ217が接続されてい

50

る。各コアは、メッシュ状に接続されたルータを介して、互いに接続されるとともに共有メモリと接続されている。

【 0 0 1 4 】

なお、図 1 においてルータを示す円内の数字は、ルータの識別子 (I D) を示している。即ち、ルータ 2 0 1 ~ 2 2 0 の識別子はそれぞれ、0 ~ 1 9 である。また、共有メモリ 3 0 1 ~ 3 0 4 については、キャッシュ機能を有する場合、メモリが用いられることもある。

【 0 0 1 5 】

次に、マルチコアシステムの通信で用いられるパケットの構造について、図 2 を用いて説明する。パケットは、宛先などに関する情報が格納されるヘッダーと、書き込みデータ又は読出しデータ等が格納されるボディとから構成される。ヘッダーおよびボディはフリット (F l i t) と呼ばれるデータ単位から構成されており、図 2 に示すパケットの例では、ヘッダーは 2 つのヘッダーフリット 0 及び 1 からなり、ボディは 5 つのボディフリット 0 ~ 4 からなる。

10

【 0 0 1 6 】

次にパケットの転送方式について説明する。パケットを転送する方式は、ソースルーティング (固定型ルーティング) 方式と適応型ルーティング方式がある。ソースルーティング方式では、送信元コアに接続されたルータが、宛先アドレスに応じてパケットの転送経路、即ちパケットの経由するルータを予め全て決める。よって、ソースルーティング方式の場合、転送経路が途中で変更されることはない。適応型ルーティング方式では、パケットを受信した各ルータは、宛先アドレスの他、ネットワークの混雑状況等を考慮して経路計算を行い、パケットの次の転送先を決定する。このように適応型ルーティングの場合、転送経路はパケットを受信したルータごとに決められる。

20

【 0 0 1 7 】

以下、第 1 乃至第 5 の実施形態について説明する。

(第 1 の実施形態)

第 1 の実施形態は、ソースルーティング方式を採った場合の実施形態である。第 1 の実施形態ではコア間通信は非同期で行われる。

【 0 0 1 8 】

まず、図 3 を用いて、ソースルーティング方式の場合のパケットのヘッダーの詳細を説明する。ヘッダーフリット 0 は、コマンドフィールド (C m d) と、リルーテッドフラグ (R e r o u t e d F l a g) と、A c k 要求フラグ (A c k R e q u i r e d F l a g) と、宛先アドレスフィールド (D e s t A d d r e s s) と、ソースコアフィールド (S r c C o r e) と、ライフサイクルフィールド (L i f e C y c l e) と、ホップルータフィールド (H o p R o u t e r 1) と、を有する。ヘッダーフリット 1 は、ホップルータフィールド (H o p R o u t e r 2 ~ 5) を有する。

30

【 0 0 1 9 】

コマンドフィールドはパケットの種類を示すフィールドである。例えば、書き込み要求パケットの場合、コマンドフィールドには “ w r i t e ” を意味する符号が格納され、読み出し要求パケットの場合、コマンドフィールドには “ r e a d ” を意味する符号が格納される。リルーテッドフラグは、パケットの宛先を変更したことを示すフラグである。

40

【 0 0 2 0 】

A c k 要求フラグは、発信元のコアが受領通知を要求するフラグである。第 1 の実施形態ではコア間通信は非同期であるため、A c k 要求フラグには “ 0 ” が格納される。宛先アドレスフィールドは、共有メモリ上の宛先アドレスを格納するフィールドである。ソースコアフィールドは、発信元のコアの識別子を格納するフィールドである。

【 0 0 2 1 】

ライフサイクルフィールドは、パケットの寿命を示すフィールドである。このフィールドに設定されるライフサイクル値は、ネットワークの特性等に依存する。例えば初期値として “ 2 0 ” が格納され、ライフサイクル値が所定の値 (例えば零) になるとタイムアウトとなる。

50

【 0 0 2 2 】

ホップルータフィールドは、パケットが経由するルータの識別子を格納するフィールドである。ホップルータフィールドの数は、ルータの中継段数に応じて変化する。発信元コアに接続されたルータは、パケットが宛先コア若しくは共有メモリに接続されたルータに達するまでに経由する全てのルータの識別子をヘッダーに格納する。このため、ルータの中継段数によっては、複数のヘッダーフリットが必要になる。

【 0 0 2 3 】

なお、マルチコアシステムに設けられた各ルータは、受信パケットのホップルータフィールドのうち最後尾のフィールドを参照することで、自分が宛先コアに接続されたルータであるか否かを認識することができる。つまり、最後尾のホップルータフィールドに格納された識別子が自身の識別子と一致すれば、自身が宛先コアに接続されたルータであると認識することができる。

10

【 0 0 2 4 】

図 4 は、第 1 の実施形態に係るルータの構成図の一例である。図 4 のルータは入力ポートおよび 5 つの出力ポートを 5 つずつ有する。ここでは、入力ポート及び出力ポートはそれぞれ、西方向 (W)、東方向 (E)、北方向 (N)、南方向 (S) およびコア方向 (C) と表現する。ただし、ルータの配置位置により入出力ポートの数は異なる。例えば、ルータ 2 0 1 は、コア方向、北方向および東方向について入出力ポートを有する。ルータ 2 0 2 は、コア方向、西方向、東方向及び北方向について入出力ポートを有する。

20

【 0 0 2 5 】

ルータは、各入力ポートに設けられたバッファ 2 1 と、スイッチ部 2 2 と、宛先変更用テーブル 2 3 と、検索部 2 4 と、ヘッダー生成部 2 5 と、経路計算部 2 6 とを有する。

【 0 0 2 6 】

バッファ 2 1 は、入力ポートで受信したデータを格納する。バッファ 2 1 の深さ (大きさ) は 1 F l i t 以上 (例えば 4 ~ 8 F l i t) である。スイッチ部 2 2 は、受信パケットが所望の出力ポートから送出されるように、パケットの出力先を切り替える。このスイッチ部 2 2 は経路計算部 2 6 により制御される。

【 0 0 2 7 】

宛先変更用テーブル 2 3 は、共有メモリ上の宛先アドレスと、宛先コアに接続されたルータの識別子とを対応させた表である。この宛先変更用テーブル 2 3 は送信元コアにより設定される。送信元コアは、自身の制御用レジスタにアクセスすると同様に、接続されたルータの宛先変更用テーブル 2 3 に書き込むことができる。

30

【 0 0 2 8 】

図 5 は宛先変更用テーブル 2 3 の一例を示している。図 5 の宛先変更用テーブル 2 3 は、データの有効性を示す V フラグ (Valid Flag) を格納する V と、共有メモリ上の宛先アドレスを格納する Dest. Address と、宛先ルータの識別子を格納する Dest Router ID とから構成される。V フラグ、宛先アドレスおよび宛先ルータ識別子のセットをエントリと呼ぶ。図 5 のエントリ 0 は V フラグが立ち、有効なエントリである。エントリ 0 において、宛先アドレスは 0 x 3 1 B 8 _ 9 B C 0 (以下、アドレス D A 1) であり、宛先ルータ識別子は “ 1 0 ” (ルータ 2 1 1) である。なお、図 5 ではエントリ数は 1 であるが、宛先変更用テーブル 2 3 は複数のエントリを有することができる。

40

【 0 0 2 9 】

次に、コア 1 0 1 を送信元コア、コア 1 1 1 を宛先コアとした場合における、第 1 の実施形態によるコア間通信の詳細について説明する。

まず、送信元のコア 1 0 1 は、書き込み要求パケットを送信する前に、自身と接続されたルータ 2 0 1 に対して設定を行う。具体的には、コア 1 0 1 は、宛先変更用テーブル 2 3 のエントリ 0 の宛先アドレス D A 1 を、宛先ルータ識別子として、ルータ 2 1 1 の識別子 “ 1 0 ” をそれぞれ書き込む。さらに、コア 1 0 1 は、エントリ 0 が有効であることを示す “ 1 ” を V フラグに書き込む。

【 0 0 3 0 】

50

宛先変更用テーブルを設定した後、コア 1 0 1 は、書き込み要求パケットを生成し、ルータ 2 0 1 へ送信する。書き込み要求パケットは、宛先アドレスフィールドに共有メモリ 3 0 1 のアドレス D A 1 が格納されている。

【 0 0 3 1 】

次に、ルータ 2 0 1 の検索部 2 4 は、コア 1 0 1 から受信した書き込み要求パケットのヘッダーに格納された宛先アドレスと同一のアドレスを有する有効なエントリが宛先変更用テーブル 2 3 に存在するかどうかを検索する。検索の結果、同一のアドレスを有する有効なエントリが存在する場合、ルータ 2 0 1 の経路計算部 2 6 は、エントリの識別子のルータを宛先として経路計算を行う。第 1 の実施形態の場合、ソースルーティングによる経路計算が行われる。

10

【 0 0 3 2 】

そして、ヘッダー生成部 2 5 は、求められた全てのルータの識別子を、パケットの経由する順にポップルータフィールドに格納するとともに、リルーテッドフラグに “ 1 ” を設定する。図 6 は、このようにして生成されたパケットのヘッダーの一例を示している。図 6 の各ホップルータフィールドには、経路計算により求められたルータの識別子 “ 1 ”、“ 2 ”、“ 6 ”、“ 1 0 ” がそれぞれ格納されている。

【 0 0 3 3 】

その後、経路計算部 2 6 からの制御を受けて、スイッチ部 2 2 はパケットの出力先を、ルータ 2 0 2 (I D = 1) の方向である E 方向のポートに切替える。これにより、宛先変更された書き込み要求パケットがルータ 2 0 1 からルータ 2 0 2 に送出される。

20

図 7 は、パケットの転送経路の一例を示す図である。書き込み要求パケットは、例えば図 7 の経路 A に示すように、ルータ 2 0 1、2 0 2、2 0 3 及び 2 0 7 を経由してルータ 2 1 1 に達する。

【 0 0 3 4 】

ここで、コア 1 1 1 は、アドレス D A 1 のデータを読み出す場合、ルータ 2 1 1 に対して読み出し要求パケットを送信する。読み出し要求パケットの宛先アドレスフィールドにはアドレス D A 1 が格納されている。

【 0 0 3 5 】

ルータ 2 1 1 はコア 1 1 1 から読み出し要求パケットを受信すると、検索部 2 4 により読み出し要求パケットの宛先アドレスと同一のアドレスを有し、かつ、リルーテッドフラグが “ 1 ” の書き込み要求パケットがバッファに蓄積されているかどうかを検索する。検索の結果、バッファ内に存在する場合、ルータ 2 1 1 はヘッダー生成部 2 5 により書き込み要求パケットのヘッダーを書き換え、読み出し要求パケットに対する応答パケットのヘッダーを生成し、コア 1 1 1 に送出される。したがって、宛先のコア 1 1 1 は、アドレス D A 1 のデータを迅速に読み出すことができる。なお、応答パケットを生成する際、バッファ内に蓄積されていた書き込み要求パケットは削除される。

30

【 0 0 3 6 】

次に、コア間通信が非同期であることに起因する問題と、それに対する対処方法について説明する。ここでは、コア 1 0 1 からの書き込み要求パケットがルータ 2 1 1 に到達した後に、宛先のコア 1 1 1 が別のアドレス D A 2 の読み出しを行った場合を考える。

40

【 0 0 3 7 】

ルータ 2 1 1 内のバッファ 2 1 に空きがない場合、アドレス D A 2 のデータを有する読み出し応答パケットは、途中のルータに留まり、いつまで経ってもルータ 2 1 1 に到達することができない。このためコア 1 1 1 の処理が停止するという不具合が生じる。

【 0 0 3 8 】

本実施形態では、ルータ 2 1 1 はライフサイクルを利用し、タイムアウトが発生した場合には、既に到達していた書き込み要求パケットを共有メモリに転送する。具体的には、ルータ 2 1 1 は、自身のバッファ内にリルーテッドフラグが “ 1 ” の書き込み要求パケットがある場合、ライフサイクル値を 1 サイクル経過するごとに 1 ずつ減らす。タイムアウトすると、ルータ 2 1 1 は、その書き込み要求パケットを共有メモリ 3 0 1 に転送する。

50

ルータ 2 1 1 は、ヘッダー生成部 2 5 により書き込み要求パケットのアドレス D A 1 を宛先として転送経路を計算し、求められたルータ識別子をホップルータフィールドに格納する。書き込み要求パケットのリルーテッドフラグは、“ 0 ” に設定する。これにより、書き込み要求パケットはルータ 2 1 1 から共有メモリ 3 0 1 に向けて送信され、例えば図 7 に示す経路 B に沿って転送される。

【 0 0 3 9 】

上記のようにしてルータ 2 1 1 に滞留した書き込み要求パケットを共有メモリに転送することで、ルータ 2 1 1 のバッファ 2 1 に空きを生じさせる。これにより、コア 1 1 1 による読出し要求に対する応答パケットは、ルータ 2 1 1 に到達することができる。その結果、コア 1 1 1 の処理が停止する事態を回避することができる。

10

【 0 0 4 0 】

第 1 の実施形態によれば、コア間通信用のバッファを新たに設けることなく、共有メモリを介さないコア間通信を実現することができる。このため、従来に比べてコア間通信のレイテンシを低減することができる。

【 0 0 4 1 】

また、第 1 の実施形態によれば、共有メモリがキャッシュ機能を有する場合、従来に比べてキャッシュラインの追い出しが低減することでヒット率を向上させることができ、その結果、マルチコアシステム全体の性能を向上させることができる。

【 0 0 4 2 】

(第 2 の実施形態)

20

第 2 の実施形態では、書き込み要求パケットの A c k 要求フラグを用い、同期をとってコア間通信を行う。宛先コアに接続されたルータは、A c k 要求フラグが立てられた書き込み要求パケットを受信すると、送信元コアにパケットの転送完了を通知する。そして、送信元コアは宛先コアにデータ書き込み完了通知を送信する。その後、宛先コアは読出し要求パケットを送信する。よって、同期したコア間通信を行うことができる。以下、詳細について説明する。

【 0 0 4 3 】

コア 1 0 1 を送信元コア、コア 1 1 1 を宛先コアとする。送信元コア 1 0 1 から書き込み要求パケットを受信したルータ 2 0 1 は、書き込み要求パケットの A c k 要求フラグに“ 1 ”を設定する。宛先コア 1 1 1 と接続されたルータ 2 1 1 は、コア 1 0 1 にパケットの転送が完了したことを通知する。

30

【 0 0 4 4 】

なお、この通知は、専用の信号線等により行ってもよいし、通常のデータ転送と区別可能な特殊なパケットを用いて行ってもよい。後者の場合は、特殊パケットがルータのバッファでデータ転送用パケットを追い越すことができるようにすることが好ましい。これにより、特殊パケットが通常のパケットに阻害されて送信元のコア 1 0 1 に到達できないために同期が完了しないという事態を回避することができる。

【 0 0 4 5 】

ルータ 2 1 1 からパケットの転送が完了した旨が通知されると、送信元コア 1 0 1 は、マルチコアシステムの提供する同期機構（メモリ上の共有変数や割り込み等）により、コア 1 1 1 へデータ書き込みの完了を通知する。コア 1 1 1 はその通知を受信すると、アドレス D A 1 に対する読出し要求パケットを送信する。それ以降の動作は第 1 の実施形態と同様であるので詳しい説明は省略する。

40

【 0 0 4 6 】

第 2 の実施形態によれば、宛先コアは送信元コアからデータ書き込みの完了通知を受けてから、読出し要求を行うため、宛先コアは送信元コアからの書き込みデータを確実に読み出すことができる。また、第 2 の実施形態によれば、従来に比べて、コア間通信のレイテンシを低減することができる。

【 0 0 4 7 】

(第 3 の実施形態)

50

第 3 の実施形態について説明する。まず、図 8 を用いて、適応型ルーティング方式の場合のパケットのヘッダーの構成について説明する。

【 0 0 4 8 】

ソースルーティング方式の場合との相違点は、ホップルータフィールドの代わりに、宛先ルータフィールド (Dest Router) が設けられていることである。宛先ルータフィールドには、宛先コアに接続されたルータの識別子が格納される。なお、従来の場合、宛先ルータフィールドは存在せず、宛先アドレスフィールドの情報を用いて経路計算が行われる。

【 0 0 4 9 】

図 9 は第 3 の実施形態に係るルータの構成を示している。図 9 のルータは、バッファ 2 1 と、スイッチ部 2 2 と、宛先変更用テーブル 2 3 と、検索部 2 4 と、ヘッダー生成部 2 5 と、経路計算部 2 6 とを有する。

【 0 0 5 0 】

さらに、第 3 の実施形態に係るルータは、隣接するルータのバッファの状態 (データの埋まり具合) を通知するフィードバック線 2 9 を有する。経路計算部 2 6 は、フィードバック線 2 9 からの情報を考慮して経路計算を行い、パケットの出力先を決定することができる。

【 0 0 5 1 】

次に、コア 1 0 1 を送信元コア、コア 1 1 1 を宛先コアとした場合における、第 3 の実施形態によるコア間通信の詳細について説明する。

まず、送信元のコア 1 0 1 は、書き込み要求パケットを送信する前に、ルータ 2 0 1 の宛先変更用テーブル 2 3 に対して設定を行う。即ち、コア 1 0 1 は宛先変更用テーブル 2 3 のエントリに、V フラグ “ 1 ”、アドレス D A 1 及びルータ識別子 “ 1 0 ” を書き込む。

【 0 0 5 2 】

次に、コア 1 0 1 は、宛先アドレスをアドレス D A 1 とした書き込み要求パケットをルータ 2 0 1 へ送信する。ルータ 2 0 1 の検索部 2 4 は、書き込み要求パケットの宛先アドレスと同一のアドレスを有する有効なエントリが宛先変更用テーブル 2 3 に存在するかどうかを検索する。検索の結果、同一のアドレスを有する有効なエントリが存在する場合、ヘッダー生成部 2 5 はそのエントリの宛先ルータ識別子を宛先ルータフィールドに格納するとともに、リルーテッドフラグに “ 1 ” を設定する。なお、この際、ヘッダー生成部 2 5 は、Ack 要求フラグについて、非同期の場合には “ 0 ” を設定し、同期の場合には “ 1 ” を設定する。

【 0 0 5 3 】

次に、ルータ 2 0 1 の経路計算部 2 6 は、格納された識別子のルータを宛先として経路計算を行い、次の転送先を決定する。その決定に従い、経路計算部 2 6 はスイッチ部 2 2 を制御する。これにより、書き込み要求パケットは、ルータ 2 0 1 から決定されたルータに向けて送信される。

【 0 0 5 4 】

パケットが転送された中継ルータの経路計算部 2 6 は、リルーテッドフラグが立っている場合には、宛先ルータフィールドに格納された識別子のルータを転送先として経路計算を行い、パケットの出力先を決定する。一方、リルーテッドフラグが立っていない場合には、経路計算部 2 6 は、宛先アドレスフィールドに格納されたアドレスを転送先として経路計算を行う。

【 0 0 5 5 】

コア 1 0 1 からの書き込み要求パケットは転送され、最終的に宛先ルータであるルータ 2 1 1 に到達する。それ以降の動作は、非同期の場合は第 1 の実施形態、同期の場合は第 2 の実施形態と同様であるので詳しい説明は省略する。

【 0 0 5 6 】

第 3 の実施形態によれば、第 1 の実施形態と同様、従来に比べてコア間通信のレイテン

10

20

30

40

50

シを低減することができるとともに、共有メモリがキャッシュ機能を有する場合には、従来に比べてマルチコアシステム全体の性能を向上させることができる。

【 0 0 5 7 】

また、第 3 の実施形態によれば、同期式のコア間通信を行う場合、第 2 の実施形態と同様、宛先コアは送信元コアからの書き込みデータを確実に読み出すことができるとともに、従来に比べて、コア間通信のレイテンシを低減することができる。

【 0 0 5 8 】

(第 4 の実施形態)

第 4 の実施形態は、ネットワークの混雑等により書き込み要求パケットが宛先コアに接続されたルータに到達できない場合の動作に関する。コア 1 0 1 からコア 1 1 1 に書き込み要求パケットを転送する場合を例にとって説明する。なお、ルーティング方式は、ソースルーティング方式でも適応型ルーティング方式でもよい。

【 0 0 5 9 】

コア 1 0 1 から送信された書き込み要求パケットは、ルータ 2 0 1 において宛先変更用テーブル 2 3 の設定に従って宛先をルータ 2 1 1 に変更され、経路計算を行って求められた出力ポートから送出され、ルータ 2 1 1 に転送される。

【 0 0 6 0 】

ここで、ルータ 2 1 1 への経路が混雑しているために、書き込み要求パケットが途中のルータ 2 0 7 (I D = 6) で滞留し、ルータ 2 1 1 に到達できない場合を考える。第 4 の実施形態では、ルータ 2 0 7 は、ルータ 2 1 1 に転送することを断念し、宛先を宛先アドレスの存在する共有メモリ 3 0 1 に変更する。具体的には、ルータ 2 0 7 の経路計算部 2 6 が、アドレス D A 1 を宛先として経路計算を行い、新たな転送経路を求める。また、ヘッダー生成部 2 5 は書き込み要求パケットのリルーテッドフラグを “ 0 ” にする。そして、書き込み要求パケットはルータ 2 0 7 から共有メモリ 3 0 1 に向けて転送される。図 1 0 の経路 E は、このときの転送経路の一例を示している。

【 0 0 6 1 】

その後、コア 1 1 1 が読み出し要求を行った場合、ルータ 2 1 1 のバッファには書き込み要求パケットは存在しない。このため、従来と同様に共有メモリから宛先アドレスのデータが読み出される。

【 0 0 6 2 】

次に、送信元コア 1 0 1 に接続されたルータ 2 0 1 がルータ 2 1 1 への経路が混雑していると判断した場合について説明する。この場合は、初めから共有メモリ 3 0 1 に向けて書き込み要求パケットを転送する。

【 0 0 6 3 】

ルータ 2 0 1 は、コア 1 0 1 からパケットを受信すると、宛先変更用テーブル 2 3 を参照し、宛先を変更すべきパケットであるかどうかを判断する。そして、宛先を変更すべきパケットである場合には、以下の処理を行う。

【 0 0 6 4 】

ルータ 2 0 1 がルータ 2 1 1 への経路が混雑していると判断した場合、ルータ 2 0 1 の経路計算部 2 6 は、アドレス D A 1 を宛先として経路計算を行う。また、ヘッダー生成部 2 5 は書き込み要求パケットのリルーテッドフラグを “ 0 ” に設定する。書き込み要求パケットは共有メモリ 3 0 1 に転送される。なお、ネットワークが混雑しているかどうかを判断する方法として、例えば、フィードバック線 2 9 からの情報を用いる。

【 0 0 6 5 】

第 4 の実施形態では、宛先ルータに到達することが困難なことが分かった段階で、ルータの動作を変更して、宛先コアへの転送を中止し、共有メモリに書き込み要求パケットを送信する。送信元コア及び宛先コアの動作を変更する必要はない。これにより、第 4 の実施形態によれば、ネットワークの混雑状況に応じて、共有メモリを介したコア間通信と、共有メモリを介さないコア間通信とを柔軟に切替えることができる。

【 0 0 6 6 】

10

20

30

40

50

(第5の実施形態)

第5の実施形態は、パケットのマルチキャスト転送を行う。第5の実施形態では、送信元コアからのデータを共有メモリに接続されたルータのバッファに蓄積し、宛先コアからの読み出し要求に応じてこのバッファから複数のコアに対してデータを送信する。これにより共有メモリへの書き込み・読み出しが発生しないことから、従来の共有メモリ型のコア間通信よりもレイテンシの小さいマルチキャストを実現させることができる。

【0067】

以下、送信元コア101から2つのコア111とコア114へデータを転送する場合について詳細に説明する。ここでは、同期の適応型ルーティング方式の場合について説明するが、第5の実施形態はソースルーティング方式や非同期の場合についても適用可能である。

10

【0068】

まず、送信元コア101は、書き込み要求パケットを送信する前に、ルータ201の宛先変更用テーブル23の設定を行う。例えば図11に示すように、コア101は、宛先変更用テーブル23のエントリに、Vフラグ“1”、アドレスDA1、及び共有メモリ301に接続されたルータの識別子“16”をそれぞれ書き込む。

【0069】

次に、コア101は、宛先アドレスをアドレスDA1とした書き込み要求パケットをルータ201へ送信する。ルータ201の検索部24は、書き込み要求パケットの宛先アドレスと同一のアドレスを有する有効なエントリが宛先変更用テーブル23に存在するかどうかを検索する。検索の結果、同一のアドレスを有する有効なエントリが存在する場合、ヘッダー生成部25は、そのエントリのルータ識別子“16”を宛先ルータフィールドに格納するとともに、リルーテッドフラグに“1”を設定する。この際、ヘッダー生成部25は、Ack要求フラグに“1”を設定する。

20

【0070】

次に、ルータ201の経路計算部26は、宛先ルータフィールドに格納された識別子のルータを宛先として経路計算を行い、次の転送先を決定する。その決定に従い、経路計算部26はスイッチ部22を制御する。これにより、書き込み要求パケットは、例えば図12の経路Hに沿って転送され、最終的にルータ217へ到達する。

【0071】

ルータ217の検索部は、宛先ルータフィールドに格納された識別子が自身のものと一致するパケットを検索し、そのようなパケットが存在すると、ルータ217はコア101に転送完了を通知する。

30

【0072】

コア101は、ルータ217から転送完了通知を受信すると、コア111及びコア114に対し、割り込みなどの手段により同期をかける。コア111及びコア114は、宛先をアドレスDA1とした読み出し要求パケットを送信する。これらの読み出し要求パケットは、共有メモリ301に向けて転送され、ルータ217へ到達する。

ルータ217の検索部は読み出し要求パケットを受信すると、自身のバッファ21の中からリルーテッドフラグが“1”であり、かつ宛先アドレスフィールドの宛先アドレスが読み出し要求パケットの宛先アドレスと一致する書き込み要求パケットを検索する。ルータ217のヘッダー生成部は、該当する書き込み要求パケットを書き換え、読み出し要求パケットに対する応答パケットのヘッダーを生成する。

40

【0073】

そして、生成した応答パケットを読み出し結果としてコア111及びコア114に送信する。なお、マルチキャストに対応するため、バッファに蓄積された書き込み要求パケットはそのまま保存しておき、マルチキャストが終了した後、共有メモリ301に書き込むとともにバッファから消去する。なお、マルチキャストの終了のタイミングとして、例えばライフサイクル値を用いたタイムアウトを用いることができる。図12の経路I及び経路Jは応答パケットの転送経路の一例を示している。

50

【 0 0 7 4 】

第5の実施形態によれば、共有メモリにデータの書き込み及び読み出しを行わない分のレイテンシを低減させることができる。

【 0 0 7 5 】

以上、本発明に係る実施形態について説明した。本実施形態ではメッシュ型のマルチコアネットワークを例にとったが、これに限らず、コア及び共有メモリにそれぞれ対応するルータが接続されていればよく、リング型やトラス型でもよい。

【 0 0 7 6 】

また、非同期のコア間通信を行う実施形態において、宛先コアにパケットを直接転送するとともに、共有メモリにもパケットを転送し、共有メモリの宛先アドレスにデータを書き込むようにしてもよい。このようにすることで、予め判明しているコア間通信についてはパケットを直接転送することでレイテンシを低減するとともに、その後他のコアが読み出し要求を行った場合には共有メモリからデータを読み出すことができる。

10

【 0 0 7 7 】

本発明のいくつかの実施形態を説明したが、これらの実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。これら実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれると同様に、特許請求の範囲に記載された発明とその均等の範囲に含まれるものである。

20

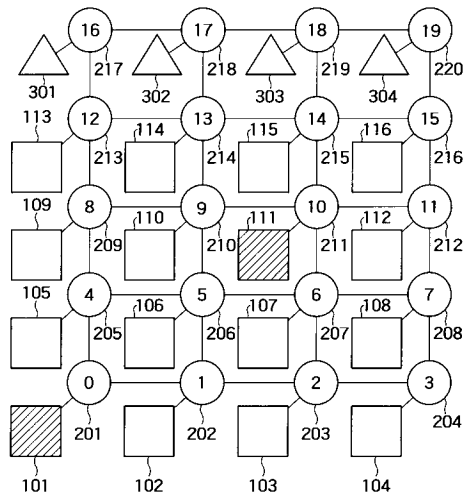
【 符号の説明 】

【 0 0 7 8 】

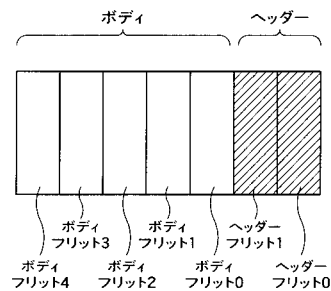
- 2 1 バッファ
- 2 2 スイッチ部
- 2 3 宛先変更用テーブル
- 2 4 検索部
- 2 5 ヘッダー生成部
- 2 6 経路計算部
- 2 9 フィードバック線
- 1 0 1 ~ 1 1 6 コア
- 2 0 1 ~ 2 2 0 ルータ
- 3 0 1 ~ 3 0 4 共有メモリ

30

【図 1】



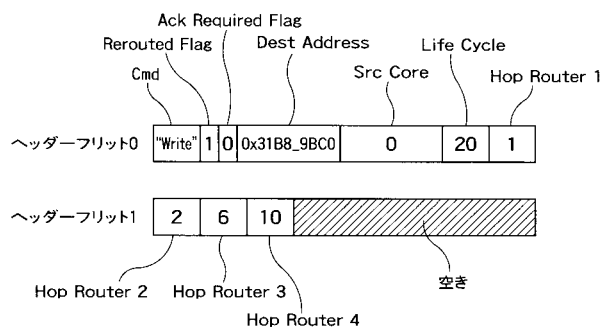
【図 2】



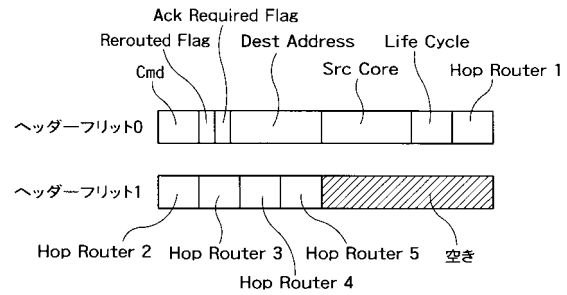
【図 5】

V	Dest. Address	Dest. Router ID
1	0x31B8_9BC0	10
0	don't care	don't care

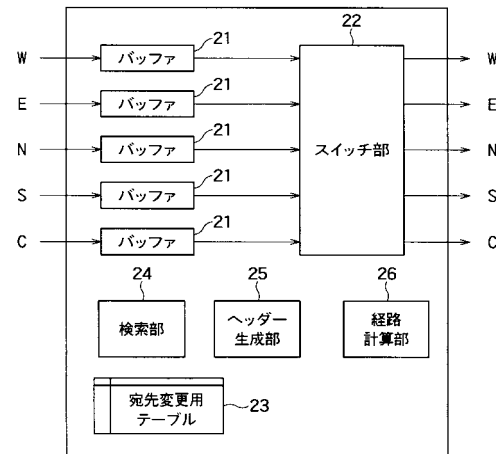
【図 6】



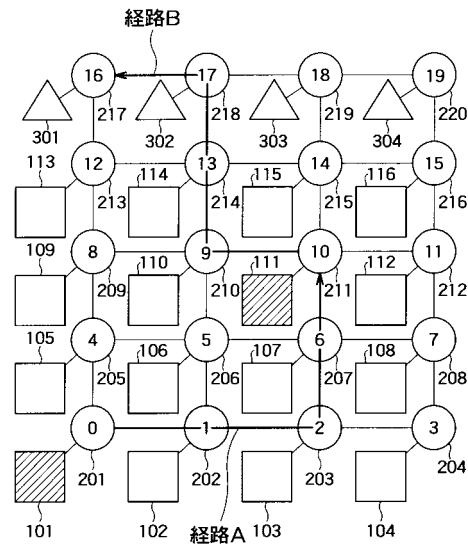
【図 3】



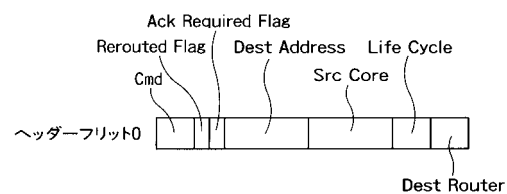
【図 4】



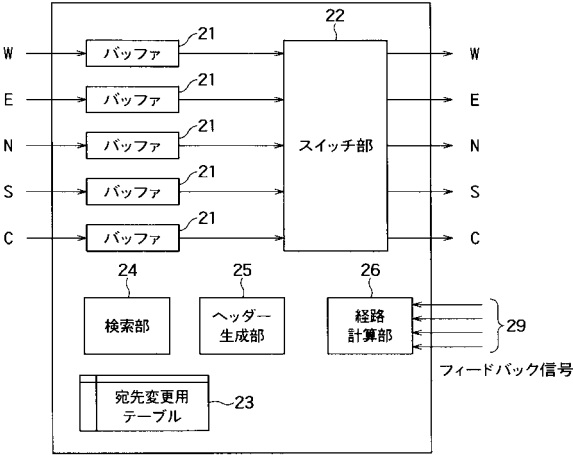
【図 7】



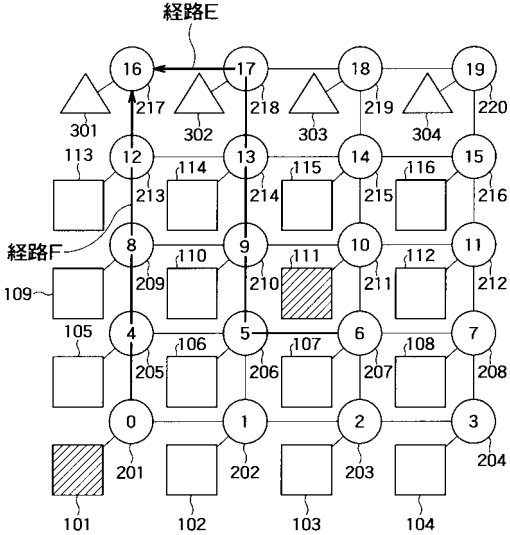
【図 8】



【 図 9 】



【 図 1 0 】

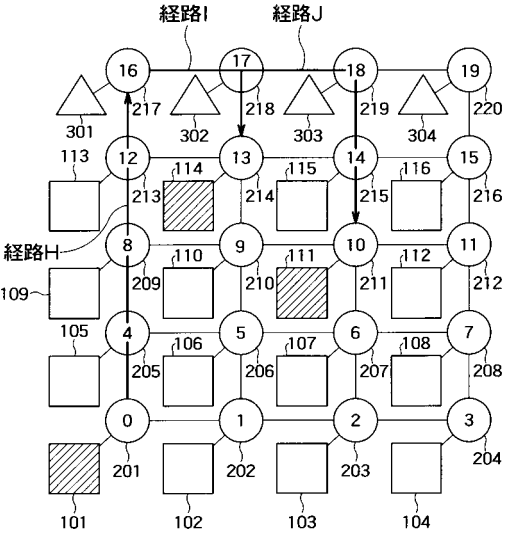


【 図 1 1 】

V	Dest. Address	Dest. Router ID
1	0x31B8_9BC0	16
0	don't care	don't care

エントリ0
エントリ1

【 図 1 2 】



フロントページの続き

(72)発明者 佐 野 徹
東京都港区芝浦一丁目 1 番 1 号 株式会社東芝内
F ターム(参考) 5B045 BB02 BB15 BB28 BB36