

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5398552号

(P5398552)

(45) 発行日 平成26年1月29日(2014. 1. 29)

(24) 登録日 平成25年11月1日(2013. 11. 1)

(51) Int.Cl.

F I

G O 6 F 15/173 (2006.01)

G O 6 F 15/173 6 4 O C

請求項の数 39 (全 18 頁)

(21) 出願番号	特願2009-553208 (P2009-553208)	(73) 特許権者	509257307
(86) (22) 出願日	平成20年3月13日 (2008. 3. 13)		エックスモス リミテッド
(65) 公表番号	特表2010-521731 (P2010-521731A)		X M O S L T D
(43) 公表日	平成22年6月24日 (2010. 6. 24)		イギリス, ビーエス1 4 ビービー ブリ
(86) 国際出願番号	PCT/GB2008/000880		ストル, キング ストリート, ヴェンチャ
(87) 国際公開番号	W02008/110806		ーズ ハウス (番地なし)
(87) 国際公開日	平成20年9月18日 (2008. 9. 18)	(74) 代理人	110000202
審査請求日	平成23年2月21日 (2011. 2. 21)		新樹グローバル・アイピー特許業務法人
(31) 優先権主張番号	11/717, 621	(72) 発明者	メイ, デービット
(32) 優先日	平成19年3月14日 (2007. 3. 14)		イギリス, ビーエス8 2 イージェイ ブ
(33) 優先権主張国	米国 (US)		リストル, クリフトン, イートン クレセ
			ント 9
		審査官	篠塚 隆
			最終頁に続く

(54) 【発明の名称】 メッセージルーティング構造

(57) 【特許請求の範囲】

【請求項 1】

各ノードがアレイ中において識別されるローカルノードアドレスをそれぞれ備え、各ローカルノードアドレスは、最上位から最下位までのアドレス重要性の順を有する複数の要素を備え、前記各ノードは、

ローカルノードアドレスの要素を、それぞれ、他の1つのノードに向かうルーティング方向にマッピングするマッピング手段と、

送信先ノードを識別する送信先ノードアドレスを有するメッセージを受信するスイッチング手段であって、前記スイッチング手段は：最上位の不一致要素を識別することで、ローカルノードアドレスを送信先ノードアドレスと比較する手段と；ローカルノードアドレスが送信先ノードアドレスと一致しない場合に、マッピング手段により前記ローカルノードアドレスの最上位の不一致要素にマッピングされた方向に、他のノードにメッセージをルーティングする手段とを備える、スイッチング手段と、

を備えるプロセッサノードのアレイ。

【請求項 2】

前記要素のそれぞれは1ビットである、請求項1に記載のアレイ。

【請求項 3】

前記マッピング手段はルーティングルックアップテーブルを備える、請求項1または2に記載のアレイ。

【請求項 4】

前記ルーティングルックアップテーブルはソフトウェアによりプログラマブルである、請求項 3 に記載のアレイ。

【請求項 5】

前記マッピング手段はルーティングアルゴリズムを備える、請求項 1 または 2 に記載のアレイ。

【請求項 6】

各ノードは、少なくとも 1 つのローカルプロセッサを備え、前記スイッチング手段は、前記ローカルノードアドレスが前記送信先ノードと一致する場合に、前記ローカルプロセッサの 1 つに前記メッセージをルーティングするように構成される、請求項 1 に記載のアレイ。

10

【請求項 7】

前記ローカルノードアドレスを前記送信先ノードアドレスと比較する手段は、1 回に 1 ビット実行され、前記最上位の不一致ビットに遭遇した場合に停止される、請求項 2 に記載のアレイ。

【請求項 8】

前記送信先ノードアドレスは、前記スイッチング手段により受信されるメッセージの最初の部分である、請求項 1 に記載のアレイ。

【請求項 9】

前記送信先ノードアドレスは最上位ビットから最下位ビットの順で前記スイッチング手段により受信される、請求項 2 に記載のアレイ。

20

【請求項 10】

前記アレイは少なくとも 2 次元である、請求項 1 に記載のアレイ。

【請求項 11】

前記アレイは少なくとも 3 次元である、請求項 10 に記載のアレイ。

【請求項 12】

前記アレイは少なくとも 4 次元である、請求項 11 に記載のアレイ。

【請求項 13】

前記スイッチング手段は、一致を検出した場合に、1 つのプロセッサのメッセージをルーティングする前に、前記送信先ノードアドレスを廃棄するように構成される、請求項 1 に記載のアレイ。

30

【請求項 14】

少なくとも 1 つのノードは、複数のプロセッサで構成され、前記メッセージは、ノード中の送信先プロセッサを識別する送信先プロセッサアドレスをさらに備え、前記スイッチング手段は前記送信先ノードアドレスが前記ローカルノードアドレスと一致する場合に前記送信先プロセッサに前記メッセージをルーティングするように構成される、請求項 1 に記載のアレイ。

【請求項 15】

前記スイッチング手段は、前記メッセージを前記送信先プロセッサにルーティングする前に前記送信先プロセッサアドレスを廃棄するように構成される、請求項 14 に記載のアレイ。

40

【請求項 16】

少なくとも 1 つのノードにおける少なくとも 1 つのプロセッサは、複数の I/O チャンネルを備え、前記メッセージは、前記プロセッサ中の送信先チャンネルを識別する送信先チャンネルアドレスをさらに備え、前記スイッチング手段は前記送信先ノードアドレスが前記ローカルノードアドレスと一致する場合に前記メッセージを前記送信先チャンネルにルーティングするように構成される、請求項 1 に記載のアレイ。

【請求項 17】

前記スイッチング手段は、前記メッセージを前記送信先チャンネルにルーティングする前に、前記送信先チャンネルアドレスを廃棄するように構成される、請求項 16 に記載のアレイ。

50

【請求項 18】

前記スイッチング手段は、送信先ノードアドレスの1またはそれ以上のビットが前記ローカルノードアドレスと一致すると、前記送信先ノードアドレスからこれらビットを廃棄するように構成される、請求項2に記載のアレイ。

【請求項 19】

プロセッサノードのアレイ中のメッセージをルーティングする方法であって、各ノードはそれぞれアレイ内において識別するためのローカルノードアドレスを有し、各ローカルノードアドレスは最上位から最下位までのアドレス重要性の順を有する複数の要素を備え、前記方法は：

送信先ノードアドレスを備えるメッセージを1つのノードにおいて受信し；

10

最上位の不一致要素を識別するために前記ローカルノードアドレスを前記送信先ノードアドレスと比較し；

前記ローカルノードアドレスが前記送信先ノードアドレスと一致しない場合に、前記ローカルノードアドレスの各要素を、それぞれ、他の1つのノードに向かうルーティング方向にマッピングするマッピング手段を調査し、前記マッピング手段により前記ローカルノードアドレスの最上位の不一致要素にマッピングされた方向にある他のノードに前記メッセージをルーティングする、方法。

【請求項 20】

前記要素のそれぞれは1ビットである、請求項19に記載の方法。

【請求項 21】

20

前記マッピング手段はルーティングルックアップテーブルを備える、請求項19または20に記載の方法。

【請求項 22】

前記ルーティングルックアップテーブルはソフトウェアによりプログラマブルである、請求項21に記載の方法。

【請求項 23】

前記マッピング手段はルーティングアルゴリズムでなる、請求項19または20に記載の方法。

【請求項 24】

各ノードは少なくとも1つのローカルプロセッサを備え、前記方法は、前記ローカルノードアドレスが前記送信先ノードアドレスと一致する場合に前記ローカルプロセッサの1つに前記メッセージをルーティングすることをさらに含む、請求項19に記載の方法。

30

【請求項 25】

前記比較するステップは：1回で1ビットの比較処理を実行し、最上位の不一致ビットに遭遇した場合に比較処理を停止する処理を含む、請求項20に記載の方法。

【請求項 26】

前記送信先ノードアドレスは、前記ノードによって受信されるメッセージの最初の部分である、請求項19に記載の方法。

【請求項 27】

前記送信先ノードアドレスは最上位ビットから最下位ビットの順に前記ノードによって受信される、請求項20に記載の方法。

40

【請求項 28】

前記アレイは少なくとも2次元である、請求項19に記載の方法。

【請求項 29】

前記アレイは少なくとも3次元である、請求項28に記載の方法。

【請求項 30】

前記アレイは少なくとも4次元である、請求項29に記載の方法。

【請求項 31】

一致を検出された場合に前記メッセージを1つのプロセッサにルーティングする前に、前記送信先ノードアドレスを廃棄することをさらに備える、請求項19に記載の方法。

50

【請求項 3 2】

少なくとも 1 つのノードは複数のプロセッサを含み、前記メッセージはノード内において送信先プロセッサを識別するための送信先プロセッサアドレスをさらに備え、前記方法は、前記送信先ノードアドレスが前記ローカルノードアドレスに一致している場合に、前記送信先プロセッサに前記メッセージをルーティングすることをさらに含む、請求項 1 9 に記載の方法。

【請求項 3 3】

前記メッセージを前記送信先プロセッサにルーティングする前に前記送信先プロセッサアドレスを廃棄することをさらに含む、請求項 3 2 に記載の方法。

【請求項 3 4】

少なくとも 1 つのノードの少なくとも 1 つのプロセッサは、複数の I/O チャンネルを含み、前記メッセージは前記プロセッサ中における送信先チャンネルを識別するための送信先チャンネルアドレスをさらに含む、前記方法は、前記送信先ノードアドレスが前記ローカルノードアドレスと一致する場合に前記送信先チャンネルに前記メッセージをルーティングすることを含む、請求項 1 9 に記載の方法。

【請求項 3 5】

前記メッセージを前記送信先チャンネルにルーティングする前に前記送信先チャンネルアドレスを廃棄することをさらに含む、請求項 3 4 に記載の方法。

【請求項 3 6】

前記送信先ノードアドレスの 1 または複数のビットが前記ローカルノードアドレスと一致する場合に前記送信先ノードアドレスからこれらビットを廃棄する、請求項 2 0 に記載の方法。

【請求項 3 7】

n を整数とするとき、 n 次元プロセッサアレイのノード間のメッセージをルーティングする方法であって、前記方法は：

アレイ中の次元の数 n を決定し；

L をノードの数とするとき、次元 $k = 1 \sim n$ のそれぞれにおいてアレイの最大範囲 $L(k)$ を決定し；

各次元 $k = 1 \sim n$ に沿って、バイナリで連続的に番号が付された座標を定義し、この座標は少なくとも $L(k)$ の長さをカバーしており、そのようなノードは各 n 次元中のそれぞれの座標により記述される；

各次元 $k = 1 \sim n$ に沿って各ノードをアドレスするために、 $2^{m(k)} \geq L(k)$ となるような $m(k)$ ビットのアドレス部分を割り当て、それぞれの $m(k)$ ビットアドレス部分中の各次元 k に沿ったノード座標を格納し；

アレイ中の各ノードの各アドレス部分に対して、値 0 のビットに、それぞれ、他の 1 つのノードに向かう第 1 方向をマッピングし、値 1 のビットに、それぞれ、別の他の 1 つのノードに向かう第 2 方向をマッピングし；

複数の前記アドレス及び前記方向を参照することによってメッセージをルーティングする、方法。

【請求項 3 8】

n を整数とするとき、 n 次元プロセッサアレイのノードにアドレスを割り当てる処理をコンピュータに実行させるためのプログラムであって、前記プログラムは、

アレイ中の次元の数 n を決定するステップ；

L をノードの数とするとき、次元 $k = 1 \sim n$ のそれぞれにおいてアレイの最大範囲 $L(k)$ を決定するステップ；

次元 $k = 1 \sim n$ のそれぞれに沿って、バイナリで連続的に番号が付された座標を定義し、この座標は少なくとも $L(k)$ の長さをカバーしており、そのようなノードは各 n 次元中のそれぞれの座標により記述される；

各次元 $k = 1 \sim n$ に沿って各ノードをアドレスするために、 $2^{m(k)} \geq L(k)$ とな

10

20

30

40

50

るような $m(k)$ ビットのアドレス部分を割り当て、それぞれの $m(k)$ ビットアドレス部分中の各次元 k に沿って、ノード座標を格納するステップ；

アレイ中の各ノードの各アドレス部分に対して、値0のビットに、それぞれ、他の1つのノードに向かう第1方向をマッピングし、値1のビットに、それぞれ、別の他の1つのノードに向かう第2方向をマッピングするステップ
を前記コンピュータに実行させるためのプログラム。

【請求項39】

プロセッサノードのアレイであって、各ノードはそれぞれアレイ中において識別するためのローカルノードアドレスを有し、各ローカルノードアドレスは最上位から最下位のアドレス重要性の順を有する複数の要素を備え、各ノードは、

前記ローカルノードアドレスの各要素に、それぞれ、他の1つのノードに向かうルーティング方向をマッピングするマッピング装置と、

送信先ノードを識別するための送信先ノードアドレスを有するメッセージを受信するための配置されるスイッチと、
を備え、前記スイッチは、

最上位の不一致要素を識別することにより、前記ローカルノードアドレスを前記送信先ノードアドレスと比較するための比較部と、

前記ローカルノードアドレスが前記送信先ノードアドレスと一致しなかった場合に、前記マッピング装置により前記ローカルノードアドレスの最上位の不一致要素にマッピングされた方向にある、他のノードに前記メッセージをルーティングするルーティング部と、
を備えるアレイ。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、プロセッサ間のメッセージをルーティングすることに関し、特に、プロセッサアレイのノード間のメッセージをルーティングする構造またはプロトコルに関する。

【背景技術】

【0002】

図1に示すように、集積回路1はプロセッサタイル2のアレイとして構成することができ、各タイルはそれぞれプロセッサ4、メモリ6、及び通信手段8を含んでいる。タイル2は、チップ1全体にデータ及び制御情報を送信する相互接続10を介して接続されている。チップ上のタイルのいくつかは、外部接続12を介して外部装置に結合されている。これに代えて、あるいは付加的に、他のチップ上のプロセッサが1つの回路を形成するようにともに配列されるように構成できる。

【0003】

このような構成の利点は、設計段階においてより小さいモジュール化された複数のプロセッサをともにリンクすることにより、製造業者が、特定のコスト、消費電力、地域目標などのような、問題のアプリケーションまたは装置のために、その処理とメモリ必要条件に従って特定のパフォーマンスを持つチップまたは回路を容易に作成することができる。

【0004】

しかしながら、また、アレイ全体にメッセージをルーティングすることには困難性を有している。メッセージは遅延するおそれがあり、最も効率的な経路によってルーティングされないかもしれず、また、デッドロックしてしまうおそれもある。さらに、大規模アレイにおいてルーティングプロトコルは複雑であり動作が重くなるおそれもある。

【発明の概要】

【0005】

本願の第1の観点では、プロセッサノードのアレイが提供され、各ノードはそれぞれ前記アレイ中において識別されるローカルノードアドレスを備え、各ローカルノードアドレスは最上位から最下位までのアドレス重要性の順を備える複数の要素を備え、各ノードは：ローカルノードアドレスの各要素にそれぞれのルーティング方向をマッピングするよう

に構成されるマッピング手段と；送信先ノードを識別する送信先ノードアドレスを備えるメッセージを受信するために配置されたスイッチとを備え、前記スイッチは：最上位の不一致要素を識別することによって、ローカルノードアドレスを送信先ノードアドレスと比較する比較手段と；前記ローカルノードアドレスが前記送信先ノードアドレスと一致しない場合に、前記マッピング手段によりローカルノードアドレスの最上位の不一致要素にマッピングされた方向にある、他のノードにメッセージをルーティングする手段とを備える。

【0006】

必須ではないが、前記要素のそれぞれは1つのビットであることが好ましい。

【0007】

驚くことに、発明者は、ノードアドレスのビットにルーティング方向をマッピングすることによって、n次元プロセッサアレイの全体にメッセージをルーティングするための効率的で、素早く、容易な方法で、デッドロックすることのないプロトコルの提供が可能であることを発見した。さらなる利点として、どのような次元数をもつどのようなサイズのアレイにもこの同様の簡単な構造が作用する。

【0008】

実施例において、スイッチは、前記比較の処理を1回に1ビット実行し、最上位の不一致ビットに遭遇した場合に比較処理を停止するように構成することができる。送信先ノードアドレスは、スイッチが受信するメッセージの最初の部分とすることができる。送信先ノードアドレスは、最上位の不一致ビットから最下位のビットの順にスイッチに受信されるように構成できる。

【0009】

これらの実施例では、多くのノードにおいて、スイッチが他のノードに転送する前にメッセージの最初のいくつかのビットを読み出すことだけを必要とすることから、特に有利である。この技術では、アドレスはメッセージの残りの部分が到着する前に得ることができることから、それによって格別に速いワームホールルーティングを容易にする。これから、スイッチは、最小の遅延でメッセージのルーティングを即座に設定することが可能である。

【0010】

マッピング手段は、ルーティングルックアップテーブルを含むように構成できる。このルーティングルックアップテーブルは、ソフトウェアによりプログラム可能に構成できる。マッピング手段はルーティングアルゴリズムを含む構成とすることができる。

【0011】

ノードは、少なくとも2次元、少なくとも3次元、または少なくとも4次元のアレイに配置することができる。

【0012】

各ノードは、少なくとも1つのローカルプロセッサ、及びローカルノードアドレスが送信先ノードアドレスに一致する場合に、ローカルプロセッサの1つにメッセージをルーティングするように構成されるスイッチを含むように構成できる。スイッチは、プロセッサにメッセージをルーティングする前に一旦一致が検出された送信先ノードアドレスを廃棄するように構成できる。

【0013】

少なくとも1つのノードは、複数のプロセッサを含み、前記メッセージは前記ノード中の送信先プロセッサを識別する送信先プロセッサアドレスを含み、前記スイッチは、送信先ノードアドレスがローカルノードアドレスに一致する場合に、送信先プロセッサにメッセージをルーティングするように構成される。スイッチは、送信先プロセッサにメッセージをルーティングする前に、送信先プロセッサアドレスを廃棄するように構成することができる。

【0014】

少なくとも1つのノードにおける少なくとも1つのプロセッサは、複数のI/Oチャンネル

10

20

30

40

50

を含み、前記メッセージはプロセッサ中に送信先チャネルを識別する送信先チャネルアドレスをさらに含み、前記スイッチは送信先ノードアドレスがローカルノードアドレスに一致する場合に、送信先チャネルにメッセージをルーティングするように構成することができる。

【0015】

前記スイッチは、送信先ノードアドレスのビットがローカルノードアドレスと一致すると、送信先ノードアドレスから1またはそれ以上のビットを廃棄するように構成することができる。たとえば、メッセージの位置がその次元内の特定の座標に絞り込めるように、特定次元に関連するノードアドレス部分が廃棄されるように構成できる。

【0016】

本発明の他の観点によれば、プロセッサノードのアレイ内におけるメッセージのルーティング方法を提供するものであって、各ノードは、アレイ内においてそれを識別するためのローカルノードアドレスをそれぞれ備えており、各ローカルノードアドレスは最上位から最下位までのアドレス重要性の順を備える複数の要素を備え、前記方法は：ノードにおいて、送信先ノードアドレスを含むメッセージを受信し；最上位の不一致要素を識別するためにローカルノードアドレスを送信先ノードアドレスと比較し；ローカルノードアドレスが送信先ノードアドレスに一致しなかった場合に、それぞれのルーティング方向にローカルノードアドレスの各要素をマッピングするマッピング手段を調査し、前記マッピング手段によりローカルノードアドレスの最上位の不一致要素にマッピングされた方向にある、他のノードにメッセージをルーティングする。

【0017】

本発明の他の観点によれば、 n を整数とするとき、 n 次元プロセッサアレイのノード間においてメッセージをルーティングする方法を提供するものであって、この方法は：アレイ中の次元 n の数を決定し； L をノードの数とするとき、次元 $k = 1 \sim n$ のそれぞれにおいてアレイの最大範囲 $L(k)$ を決定し；各次元 $k = 1 \sim n$ に沿って、バイナリで連続的に番号が付された座標を定義し、この座標は少なくとも $L(k)$ の長さをカバーしており、そのようなノードは各 n 次元中のそれぞれの座標により記述される；各次元 $k = 1 \sim n$ に沿って各ノードをアドレスするために、 $m(k)$ ビットのアドレス部分を割り当て、それぞれの $m(k)$ ビットアドレス部分中の各次元 k に沿って、ノード座標を格納し；アレイ中の各ノードの各アドレス部分に対して、値0のビットにそれぞれ第1方向をマッピングし、値1のビットにそれぞれ第2方向をマッピングし、複数の前記アドレス及び前記方向を参照することによってメッセージをルーティングする。

【0018】

本発明の他の観点によれば、 n を整数とするとき、 n 次元プロセッサアレイのノードにアドレスを割り当てるコンピュータプログラム製品を提供するものであって、コンピュータプログラムは以下のステップを実行するコードを含み： L をノードの数とするとき、アレイ中の次元 n の数を特定し、次元 $k = 1 \sim n$ のそれぞれにおいてアレイの最大範囲 $L(k)$ を決定し、次元 $k = 1 \sim n$ のそれぞれに沿って、バイナリで連続的に番号が付された座標を定義し、この座標は少なくとも $L(k)$ の長さをカバーしており、そのようなノードは各 n 次元中のそれぞれの座標により記述される；各次元 $k = 1 \sim n$ に沿って各ノードをアドレスするために、 $m(k)$ ビットのアドレス部分を割り当て、それぞれの $m(k)$ ビットアドレス部分中の各次元 k に沿って、ノード座標を格納し；アレイ中の各ノードの各アドレス部分に対して、値0のビットにそれぞれ第1方向をマッピングし、値1のビットにそれぞれ第2方向をマッピングする。

【0019】

本発明をより理解し、どのようにして上の構成が効果を導出できるのかについて、実施例と対応する図面を用いて説明する。

【図面の簡単な説明】

【0020】

【図1】図1はプロセッサアレイの一例を示す。

【図2】図2はプロセッサノードの概略的な代表例である。

【図3】図3はルーティングに適したメッセージフォーマットを示す。

【図4】図4は1次元プロセッサアレイを示す

【図5】図5は2次元4×4プロセッサアレイを示す。

【図6】図6は2次元3×3プロセッサアレイを示す。

【図7】図7は3次元3×3×2プロセッサアレイを示す。

【図8】図8は4次元2×2×2×2プロセッサアレイを示す。

【図9】図9はルーティング方法を表すフローチャートである。

【図10】図10はアドレスとルーティング方向の割り当て方法を表すフローチャートである。

10

【発明を実施するための形態】

【0021】

図2は、本発明の典型的な実施例によるプロセッサタイル2'を概略的に示す。このような複数のタイルは結合された機構またはアレイと一緒に配置され、アレイ中の各タイル2'はシステムスイッチ16及びアレイ中の他のタイルにそれぞれリンクする複数の2方向システムリンク18を備えている。システムリンク18のいくつかまたは全ては、それぞれ他のプロセッサにそれぞれの方向でリンクし、および／またはいくつかのシステムリンク18は同一のプロセッサに同一の方向でリンクすることで、システムリンク18のグループがより大きい容量で1つのリンクとして扱うことができる。システムリンク18は、オンチップまたはインターチップであってもよい。各システムスイッチ16はアレイ中のノードとして定義される。

20

【0022】

この例において、タイル2'は4つのプロセッサ4を含んでいる。各プロセッサ4は2方向チャンネルリンク22及びプロセッサスイッチ14を介してアクセス可能な多数の異なるI/Oチャンネルを備えている。プロセッサスイッチは、1または複数の2方向プロセッサリンク20によってシステムスイッチに接続される。タイル2'は、後述するように、システムスイッチ16によってアクセス可能なルーティングルックアップテーブルを備えている。タイル2'は、RAM及びROM（図示せず）のようなメモリおよび外部リンクをさらに備えている。

【0023】

30

実行中において、システムスイッチ16はローカルプロセッサの1つ（たとえば、同一ノードのプロセッサ）から、または異なるノードのプロセッサから、入ってくるメッセージ26を受信する。各ノードはアレイ中において各ノードを識別するためのアドレスが割り当てられ、各プロセッサはそれぞれのノード中の各プロセッサを識別するためのアドレスが割り当てられ、各チャンネルはそれぞれのプロセッサ中における各チャンネルを識別するためのアドレスが割り当てられている。アドレスは、多くの場合個々のビットであるアドレスデータの構成要素量子で構成される。図3に概略的に図示するように、メッセージは、送信先ノードアドレスを特定する送信先ノード部分28、送信先プロセッサアドレスを特定する送信先プロセッサ部分30、および送信先チャンネルアドレスを特定する送信先チャンネル部分32を備えるヘッダ27を含む。また、メッセージはペイロード34を含む。好ましい実施例では、ヘッダは2バイト（16ビット）の長さであり、8ビットの送信先ノード部分、4ビットの送信先プロセッサ部分、4ビットの送信先チャンネル部分で構成される。このことは、256ノード、4096プロセッサ、65536チャンネル、各ノードは16プロセッサを有し、各プロセッサが16チャンネルを有することを許容する。

40

【0024】

システムスイッチ16は、ローカルノードアドレス（たとえば、自身のノードアドレス）を、メッセージの送信先ノードアドレスと比較する。これらが一致した場合、システムスイッチはメッセージの送信先プロセッサアドレスを調査し、対応するプロセッサスイッチ14にこれをルーティングする。プロセッサスイッチ14は、その後、メッセージの送信先チャンネルアドレスを調査し、対応するチャンネルにこれをルーティングする。なお、好ましく

50

は受信ソフトウェアがヘッダを再度見る必要がないように、ヘッダ27はメッセージが最終的にプロセッサ4に送信される前に廃棄されるように構成できる。メッセージがプロセッサスイッチ14にルーティングされる前にノードアドレス部分28及びプロセッサアドレス部分が廃棄され、チャンネルアドレス部分32は、メッセージがプロセッサスイッチ14にルーティングされた後であって、メッセージがプロセッサ4にルーティングされる前に廃棄されるように、ヘッダの各部分が廃棄される。

【0025】

しかしながら、一致しなかった場合、システムスイッチ16は、他のノードを送信先とするメッセージを受信し、これをどの方向に転送するか、たとえば、どのシステムリンク18がメッセージを再送信するか、どのノードが次にこれを受信するか、を決定する必要がある。アレイは、数十、数百、数千のノードで構成され、このためメッセージが送信先に到達するまでに多数のノード間を中継される必要があるかもしれないことに留意してほしい。これから、アレイ全体を通して最も効率的なルートを見つけることは、簡単な仕事ではない。

【0026】

本発明によれば、ルックアップテーブル24を調査することにより達成される。ルックアップテーブル24は、ローカルノードアドレスのビットに対して表にされた多数の可能なルーティング方向を備え、1つの方向がそれぞれ各ビットにマッピングされている。ノードのローカルアドレスとメッセージの送信先アドレスの間の比較に基づいて、システムスイッチ16は不一致ビットを識別し、そのビットにどの方向がマッピングされているかを特定するためにルックアップテーブル24を参照する。

【0027】

ルックアップテーブルは、顧客がモジュラーファッションで多くのプロセッサを互いに接続し、必要とされるルーティングをプログラムすることが可能となるように、ソフトウェアプログラマブルであることが好ましい。これに代えて、アレイが供給者によって予め設計されている場合には、ルックアップテーブルをアレイのために固定的にプログラムされたものに代えることができる。

【0028】

ルックアップテーブルの一例を図2に示すが、実施例においては、各ノードにこのようなルックアップテーブルが複数提供されるように構成できる。たとえば、システムリンク毎に1つのルックアップテーブルを備えるように構成できる。このことは、ルックアップテーブルリソースのためにメッセージが競合することを防止し、複数のメッセージが同時にルーティングされることを許容することから、メッセージのルーティングを有利に促進する。

【0029】

特に好ましい実施例では、システムスイッチ16は、たとえば、一度に1ビットのように、ビット対ビットでローカルノードアドレスを送信先ノードアドレスと比較し、最も先に不一致ビットが発見されるとすぐに比較を停止する。このことは、多くのノードにおいてシステムスイッチ16は、他のノードにこれを転送する前にメッセージの送信先アドレスの小さい部分を読み込むことだけを必要とすることから、特にルーティングを速めることとなる。スイッチは、1つのビットを読み込むだけでよい場合もしばしばある。好ましくは、システムスイッチ16は、最上位ビットを比較することで開始し、次のビット、そしてその次と進み、不一致を発見するまで行う。好ましい実施例では、ルーティング方法がしばしばメッセージの最初の数ビットを読み込むだけで決定でき、また最初のビットのみを読み込むことでさえ決定できるように、送信先ノードアドレスの最上位ビットはヘッダ27のまさしく最初に置かれる。

【0030】

この技術は、メッセージの残りの部分が到着する前にアドレスを得ることができることから、格別に速い「ワームホールルーティング」を容易にする。これから、システムスイッチ16は最小の遅延でメッセージのルーティングを即座に設定することが可能である。

【0031】

ルーティングルックアップテーブルの使用は、本発明の1実施例に関する簡単な1次元レイを概略的に表す図4の例を参照してさらに詳細に説明される。

【0032】

この例では、各タイル2'は単一のプロセッサ4のみを含み、プロセッサアドレスは実質的にノードアドレスと同じである。また、各タイルは、ルーティングルックアップテーブル24およびシステムスイッチ16を備えている。さらに、レイが1次元であることから、各タイル2'は最高でも2つのシステムリンク18と接続される。簡潔化のために、プロセッサスイッチ14、プロセッサリンク20及びチャネルリンク22の詳細は含まない。

【0033】

8個のノードに適用するために、各ノードは3ビットアドレスに割り当てられる。したがって、各ルックアップテーブル24はノードアドレスの各アドレスに対して1つ、合計3つのエントリを有している。各エントリは、アップ(U)及びダウン(D)の2つの可能なルーティング方向のうちの1つの特定するように構成できる。ルックアップテーブルは次のように表される：

ノード	ノードアドレス	ルックアップテーブル
N 1	0 0 0	D D D
N 2	0 0 1	D D U
N 3	0 1 0	D U D
N 4	0 1 1	D U U
N 5	1 0 0	U D D
N 6	1 0 1	U D U
N 7	1 1 0	U U D
N 8	1 1 1	U U U

この例のように、最上位ビットはレイの特定の半分を識別する。所定の半分の中で、2番目の上位ビットは、レイの特定の4分の1を識別し；所定の4分の1の中で、最下位ビットは特定のノードまで狭められる。これから、メッセージが誤った半分の中から始まった場合には、関連するシステムスイッチ16は、送信先とローカルアドレスの最上位ビット間の単一の比較だけで発見することができる。ルックアップテーブル24を参照すると、システムスイッチが最上位ビットにどの方向をマッピングするかを特定し、これから、正しい半分に向けてこれを移動するためにメッセージをどの方向にルーティングするかを特定する。メッセージが正しい半分であれば、関連するシステムスイッチ16は少なくとも2つの比較を実行する必要がある、1つは最上位ビット間の比較であり、1つは2番目に上位のビット間の比較である。2番目に上位のビット間の比較は、メッセージが誤った4分の1にあるか否かをシステムスイッチ16に教え、そうであるならば、正しい4分の1に向けてこれを移動するためにメッセージをどの方向にルーティングする必要があるかを教える。メッセージが正しい4分の1にルーティングされるのであれば、最下位ビットを用いて最終的な送信先ノードを見つけるために同様の原理が適用される。

【0034】

説明を目的として、ノードN7からノードN2にメッセージをルーティングする場合を考える。ノードN7、N6及びN5のそれぞれにおいて、システムスイッチは最上位ビットのみを比較する。なぜなら、それらが一致しておらず、ルックアップテーブル24を参照して、メッセージは上方にルーティングされるべきと決定されるからである。したがって、ノードN7-N5において、受信したメッセージの単一のビットのみを調査することで速やかにメッセージをルーティングすることができる。ノードN4及びN3において、システムスイッチ16は、送信先ノードN2に到着するまで、メッセージをさらに上方に向けてルーティングする必要があると決定するために、2番目の上位ビットを比較することに頼らなければならない。

【0035】

なお、もちろん、最上位から最下位ビットの順をどのように名付けるかは、通常、慣例の問題だけである。しかしながら、この発明の目的のために、「重要性」の意味はノードのアレイ中における位置とリンクしており、これから最上位ビットはアレイを2つの部分に分割し、次の上位ビットはこれら部分のそれぞれを2つのサブ部分にサブ分割する、以下同様である。すなわち、最上位ビットから最下位ビットまでのアドレスビットの重要度の順が、アレイのより狭い空間的領域に対応する。重要性の順は、メッセージのアドレス部分中にどのビットが配列されるかの順を意味するものではなく、メモリやその他のものにどのように配列されるかの順でもない。このような要因は発明の要点ではない（とはいえ、上述したように、好ましい実施例では、最上位ビットがヘッダ27の最初に位置している）。本文を単純にする目的で、最上位から最下位までのビットは左から右に読むようにする。

10

【0036】

この構造の速さと単純さは、より複雑な複数次元のアレイにおいて特に有利である。図5は、 4×4 の2次元アレイに拡張した本願発明の一例を概略的に図示したものである。便宜上、各タイル2'は、それぞれのシステムリンク18とそれぞれのノードアドレス0000~1111のみで示している。

【0037】

各ノードアドレスは、2つの異なる部分を含むことを考慮することができる：左-右方向の座標を特定する部分、上-下方向のアドレスを特定する座標。この例では、最上位2ビットはアレイの左から右に00,01,10,11が割り当てられ、最下位2ビットはアレイの上から下に00,01,10,11が割り当てられている。ルーティングルックアップテーブルはノードアドレスに対して、2つの付加的方向である左（L）及び右（R）をマッピングしている。各ノードのルックアップテーブルは以下の通りである：

20

ノード	ノードアドレス	ルックアップテーブル
N 1	0 0 0 0	R R D D
N 2	0 0 0 1	R R D U
N 3	0 0 1 0	R R U D
N 4	0 0 1 1	R R U U
N 5	0 1 0 0	R L D D
N 6	0 1 0 1	R L D U
N 7	0 1 1 0	R L U D
N 8	0 1 1 1	R L U U
N 9	1 0 0 0	L R D D
N 1 0	1 0 0 1	L R D U
N 1 1	1 0 1 0	L R U D
N 1 2	1 0 1 1	L R U U
N 1 3	1 1 0 0	L L D D
N 1 4	1 1 0 1	L L D U
N 1 5	1 1 1 0	L L U D
N 1 6	1 1 1 1	L L U U

30

40

この場合、システムスイッチ16は最初にメッセージを受信し、その送信先ノードアドレスの最上位ビットをローカルノードアドレスの最上位ビットと比較する。最上位が一致しているか否かを特定することにより、メッセージが左-右方向におけるアレイの正しい半分にあるか否か、例えば、それが最も左の2つの縦列に位置するか、あるいは最も右の2つの縦列に位置するかをスイッチ16に教える。もし、そうでない場合、ルックアップテーブルは、最上位ビットにマッピングされるルーティング方向によって、正しい半分に向けてメッセージを転送するためにどの方向にルーティングするかをスイッチに教える。正しい半分であれば、2番目の上位ビットが正しい縦列に絞込む。正しい縦列であれば、上-下方向において、最下位ビットについて同様の工程を適用する。

50

【0038】

どの次元のアドレスビットを最上位ビットおよび最下位ビットとするかは用語上の問題だけである。つまり、発明は、メッセージのための正しい横列を見出すために最初に上一下方向を特定するビットを読み、その後続いて、正しい縦列に絞り込むために左一右方向におけるペアを読み込むことを同等とすることができる。この適用の目的のために、最上位の次元は、単に使用される最初のアドレス部分のそれとして考えられ、最下位の次元は使用される最後の部分のそれである。異なる次元の部分からのビットは、インターリーブ形式で使用されるか、および／または配列することができる。これら及びバリエーションは、基本的な原理を害するものではなく、実質的に、最上位から最下位までのアドレスビットの重要性の順はアレイの空間的領域をより狭める順を意味する。

10

【0039】

図6は、 3×3 の2次元アレイの例を概略的に示す。再度便宜上、各タイル2'はシステムリンク18とそれぞれのノードアドレス0000~1010のみを示す。この例は、不規則なアレイまたは都合のよいサイズである 2^n （ n は整数であるとき）のサイズを有していないアレイにおけるルーティングを提供するために、本発明がどのように用いられるかを示す。説明を単純化することを目的として、余分なノードアドレスである架空の位置2''を示している。 3×3 アレイのためのルックアップテーブルは以下の通りである：

ノード	ノードアドレス	ルックアップテーブル
N 1	0 0 0 0	R R D D
N 2	0 0 0 1	R R D U
N 3	0 0 1 0	R R U x
N 4	0 1 0 0	R L D D
N 5	0 1 0 1	R L D U
N 6	0 1 1 0	R L U x
N 7	1 0 0 0	L x D D
N 8	1 0 0 1	L x D U
N 9	1 0 1 0	L x U x

20

存在するノード2'のアドレス及びルックアップテーブルは、図5の 4×4 アレイ中の対応するノードと、余分なノードアドレスのためのそれを除いて同一であり、ルックアップテーブル中のいくつかのエントリは、それらが調査されることができる状況にないことから、実質的に「気にしない」(x)となっている。これから、どのようにして 3×3 アレイを 4×4 グリッドから実質的に「切り抜く」ことができるかを見せることができる。同様の原理は、不規則な形状のアレイのアドレスに使用することが可能である。

30

【0040】

図7は、 $3 \times 3 \times 2$ の3次元アレイに拡張した本発明の一例を示す。これは、前方(F)及び後方(B)の2つの新たな方向を生成する。後方一前方方向においてノード位置を特定するためにノードアドレスに余分なビットを追加することで、発明の原理は全く同一である。ルックアップテーブルは以下の通りである：

ノード	ノードアドレス	ルックアップテーブル
N 1	0 0 0 0 0	R R D D B
N 2	0 0 0 1 0	R R D U B
N 3	0 0 1 0 0	R R U x B
N 4	0 1 0 0 0	R L D D B
N 5	0 1 0 1 0	R L D U B
N 6	0 1 1 0 0	R L U x B
N 7	1 0 0 0 0	L x D D B
N 8	1 0 0 1 0	L x D U B
N 9	1 0 1 0 0	L x U x B
N 10	0 0 0 0 1	R R D D F

40

50

N 1 1	0 0 0 1 1	R R D U F
N 1 2	0 0 1 0 1	R R U x F
N 1 3	0 1 0 0 1	R L D D F
N 1 4	0 1 0 1 1	R L D U F
N 1 5	0 1 1 0 1	R L U x F
N 1 6	1 0 0 0 1	L x D D F
N 1 7	1 0 0 1 1	L x D U F
N 1 8	1 0 1 0 1	L x U x F

この場合、図 5, 6 に関連して記載したように、メッセージは、まず正しい縦列及び横列にルーティングされる。次に、システムスイッチ16は、ローカルアドレス及び送信先アドレスの最下位ビットを比較することに頼る。このことは、メッセージが正しい平面（たとえば、前方または後方）にあるか否かをスイッチ16に教える。もしそうでなければ、スイッチ16は、正しい平面を得るためにメッセージをどの方向にルーティングするかを決定するためにルーティングテーブルを参照する。

【0041】

発明の1つの実施例によれば、メッセージの位置がこの次元内の特定の座標に絞り込むように、特定の次元に関連するノードアドレス部分が廃棄されるように構成できる。図7のアレイを例として示すように、正しい縦列にメッセージがルーティングされてしまうと、第1及び第2の最上位ビットは必要がなくなり廃棄されるように構成できる。同様に、メッセージが正しい横列にルーティングされると、第3及び第4の最上位ビットは必要がなくなり廃棄されるように構成できる。

【0042】

メッセージがルーティングされると個々のビットすら廃棄されるように構成でき、たとえば、一致した最上位ビットを廃棄し、その後一意した第2上位ビットを廃棄し、以下同様にすることができる。

【0043】

要求されるビット比較の回数を減少させることから、部分または個々のビットを送信先アドレスから廃棄することはルーティング速度を改善する。しかし、このような部分またはビットを廃棄することを許容するために、短くされたアドレスに基づくメッセージのルーティングのために、ある方法を実行しなければならない。これを達成する1つの方法は、所定のリンクのルックアップテーブルから取り去られたビットを有する各システムリンク18のために分離されたルックアップテーブル24を提供することである。たとえば、図7の例を再度参照すると、メッセージがまず左-右方向にルーティングされると、上-下方向及び前方-後方方向におけるリンク18が原アドレスの最初の2つの最上位ビットを比較するためのいずれのビットにも必要ではない。同様に、このメッセージは続いて上-下方向にルーティングされると、前方-後方リンク18のルックアップテーブルは単一ビットを含むことだけが必要となる。このことは、ルックアップテーブルのいくつかのサイズを減少させ、リソースを節約する。

【0044】

この発明は4またはそれ以上の次元に拡張することができる。もちろん、アレイは3つの空間的次元に依然として存在し—この「次元」という語はここでは単に「自由度」を意味する。ここまでアレイは4またはそれ以上の次元をもち、アレイ中の少なくとも1つのノードは4つまたはそれ以上の他のノードとリンクしている。これは4またはそれ以上のメッセージをルーティングすることが可能な方向があることを意味している。

【0045】

図8は、正八胞体として知られる、 $2 \times 2 \times 2 \times 2$ の4次元のハイパーキューブに拡張した本発明の一例を示している。各ノードN1~N16は4つの他のノードに接続され、いずれのノードもメッセージをルーティングする際に4つの自由度を持っている。

【0046】

図7の例のように、左ー右、上ー下及び前ー後の3つの自由度を有している。図8の例は、メッセージのルーティングのために4番目の自由度を追加するための付加的リンク18を提供する。任意の用語の問題として、2つの新たに作成した方向をここではイエロー（Y）及びグリーン（G）とし、イエローはグリーンの反対方向とする。各ノードにおけるルーティングテーブルは以下の通りである：

ノード	ノードアドレス	ルックアップテーブル
1	0 0 0 0	R D B Y
2	0 0 0 1	R D B G
3	0 0 1 0	R D F Y
4	0 0 1 1	R D F G
5	0 1 0 0	R U B Y
6	0 1 0 1	R U B G
7	0 1 1 0	R U F Y
8	0 1 1 1	R U F G
9	1 0 0 0	L D B Y
10	1 0 0 1	L D B G
11	1 0 1 0	L D F Y
12	1 0 1 1	L D F G
13	1 1 0 0	L U B Y
14	1 1 0 1	L U B G
15	1 1 1 0	L U F Y
16	1 1 1 1	L U F G

10

20

図9は、本発明に係る方法を示すフローチャートである。この方法は、ステップ100において、1つのノードのシステムスイッチ16がメッセージの送信先ノードアドレスの第1ビットを受信した場合に開始する。記述のために、いずれの時間においても考慮されるアドレスビットは、ここではビット番号*i*としている。ステップ102において、送信先ノードアドレスの*i*番目のビットはローカルノードアドレスの*i*番目のビットと比較される。ステップ104において、送信先アドレスとローカルアドレスの*i*番目のビットが一致したか否かに基づいてこの方法は分岐する。それらが一致しない場合には、この方法は、*i*番目のビットに対してマッピングする方向を決定するためにルックアップテーブルを調査するステップ106に移行する。ステップ108において、メッセージはその方向にルーティングされる。送信先とローカルアドレスの*i*番目のビットが一致した場合には、この方法はノードアドレスが到達する最後のビットであるか否かを特定するステップ110に移行する。そうであれば、メッセージはその送信先に到着し、この方法は、メッセージがローカルプロセッサにルーティングされるステップ112に分岐する。最後のビットが到達してない場合には、この方法は次のアドレスビットを考慮するステップ114に移行し、ステップ102からの方法を繰り返す。

30

【0047】

上の記述から、ノードにアドレスを付け、対応するルックアップテーブルに割り当てる好ましい方法としてパターンがまた出現する。これは、図10のフローチャートを参照して説明する。ステップ200において、アレイにおける次元の数*n*を決定する。ステップ202において、*L*をノードの数とする場合に、アレイの最大範囲*L* (*k*) は各次元*k* = 1 ~ *n* に対して決定される。ステップ204において、座標は次元*k* = 1 ~ *n* のそれぞれにおいて定義される。その座標はバイナリで連続しており、少なくとも*L* (*k*) の長さをカバーしなければならない。各ノードの位置は、*n*次元のそれぞれにおけるそれぞれの座標によって記述されることができる。ステップ206において、*m* (*k*) ビットのアドレス部分は、各次元*k* = 1 ~ *n* に沿って各ノードに対して制約 $2^{m(k)} \geq L(k)$ が割り当てられ、次元*k* に沿ったノードの座標はそれぞれの*m* (*k*) ビットのアドレス部分に格納されるように構成できる。ステップ208において、ノードアドレスビットに方向をマッピングするこ

40

50

とにより、各ノードのためのルックアップテーブルが引き出される。アレイ中の各ノードの各アドレス部分に対して、これは、値0のビットにそれぞれの第1方向にマッピングし（例えば、右、下、後方、イエロー）、値1のビットにそれぞれ第2方向（例えば、左、上、前方、グリーン）をマッピングすることを意味する。

【0048】

たとえば、図7を再び参照すると、アドレスとこれに対応するルックアップテーブルは、以下のように割り当てられる。まず、アレイは3次元であり、 $3 \times 3 \times 2$ ノードの最大に拡張されている。 $2^1 < 3 < 2^2$ であることから、3つのノードの方は2ビットアドレス部分を必要とする。2つのノードの方は1ビットアドレス部分を必要とするだけである。これから、アレイはノードにアドレスを付与するために5ビットアドレスを用いるものであって、左-右方向においてノードにアドレスを付与するために2つの最上位ビットを用い、上-下方向においてノードにアドレスを付与するために3番目及び4番目の上位ビットを用い、前方-後方方向においては最下位ビットを用いる。ルックアップテーブルを引き出すために、2つの最上位ビットにおける右及び左、3番目及び4番目の上位ビットにおける下及び上、最下位ビットにおける後方及び前方に対してそれぞれ0及び1がマッピングされる。

【0049】

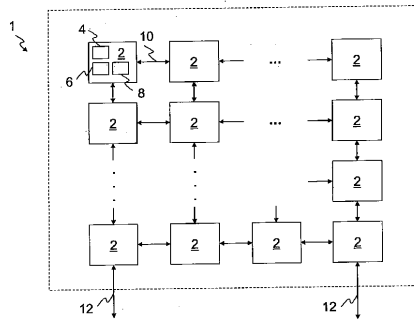
上記実施例は単なる例証として記載したものであることはいうまでもない。他の実施例として、他のアドレスパターンは当業者にとって自明である。さらに、最上位から最下位ビットの順は、メッセージが平面の次に横列、その次に縦列の順でルーティングされるか、横列の次に平面、その次に縦列の順にルーティングされるかなどのように変更することが可能である。さらに、ルーティングルックアップテーブルは、ルーティングアルゴリズムのようなより好ましくないオプションによって置き換えることも可能である。たとえば、扱いやすい幾何学形状を形成しない不規則アレイやまたツリータイプ構造を有するアレイを含む、異なるサイズ、異なる形状、異なる次元数のアレイの例に対する他の構成は、当業者により実行可能である。本発明の観点は記述した実施例によって限定されるものではなく、以下のクレームによってのみ限定される。

10

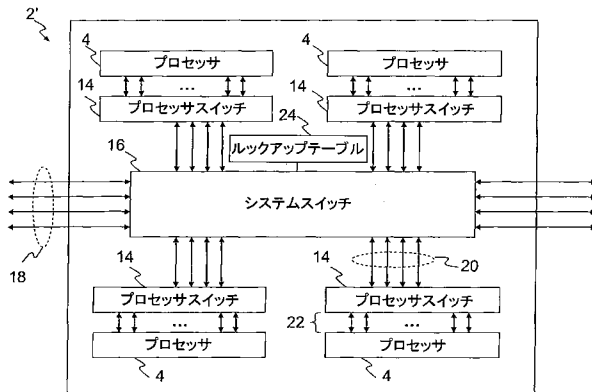
20

【図 1】

Figure 1

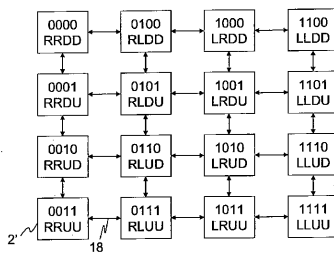


【図 2】



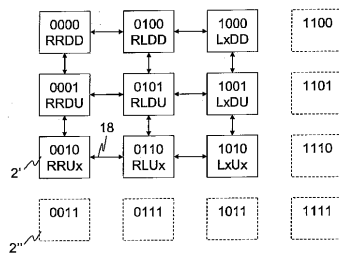
【図 5】

Figure 5

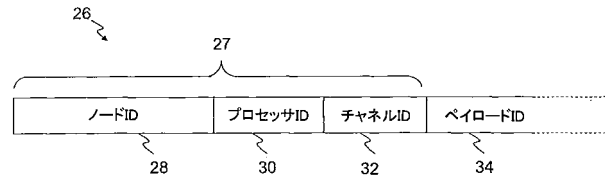


【図 6】

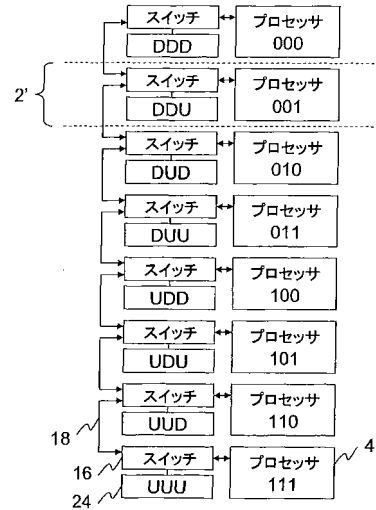
Figure 6



【図 3】

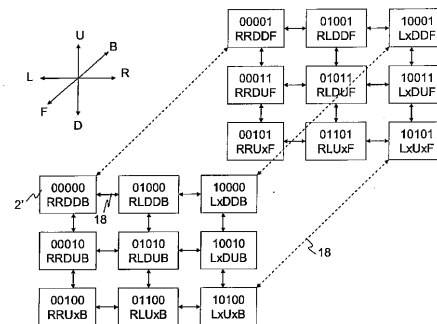


【図 4】



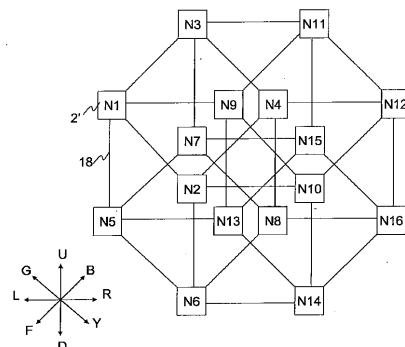
【図 7】

Figure 7

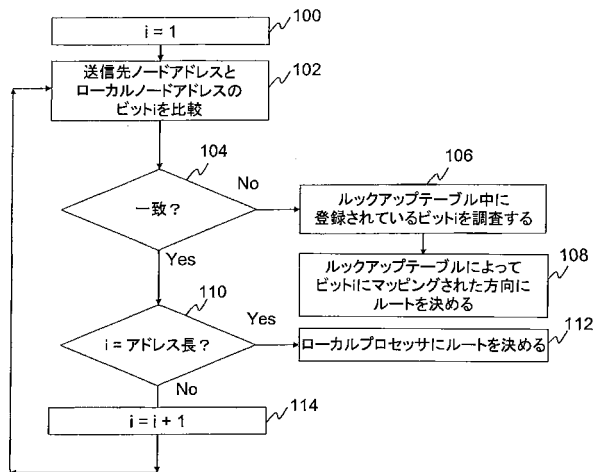


【図 8】

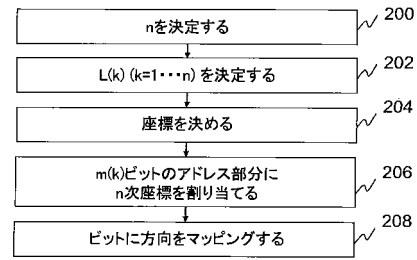
Figure 8



【図 9】



【図 10】



フロントページの続き

(56)参考文献 特開昭63-124162 (JP, A)

(58)調査した分野(Int.Cl., DB名)
G06F15/173