

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

**特許第3739798号
(P3739798)**

(45) 発行日 平成18年1月25日(2006.1.25)

(24) 登録日 平成17年11月11日(2005.11.11)

(51) Int. Cl.

H04L 12/56 (2006.01)

F I

H04L 12/56 400Z

請求項の数 15 (全 18 頁)

(21) 出願番号	特願平9-530304	(73) 特許権者	富士通株式会社
(86) (22) 出願日	平成9年2月21日(1997.2.21)		神奈川県川崎市中原区上小田中4丁目1番1号
(65) 公表番号	特表平11-504494	(74) 代理人	弁理士 石田 敬
(43) 公表日	平成11年4月20日(1999.4.20)		弁理士 土屋 繁
(86) 国際出願番号	PCT/US1997/002631	(74) 代理人	弁理士 戸田 利雄
(87) 国際公開番号	W01997/031458		弁理士 西山 雅也
(87) 国際公開日	平成9年8月28日(1997.8.28)		
審査請求日	平成16年2月20日(2004.2.20)		
(31) 優先権主張番号	08/605,676		
(32) 優先日	平成8年2月22日(1996.2.22)		
(33) 優先権主張国	米国(US)		

最終頁に続く

(54) 【発明の名称】 ダイナミックなネットワークポロジープローブのシステム及び方法

(57) 【特許請求の範囲】

【請求項1】

ネットワークのネットワークポロジープローブ情報をダイナミックに決定する方法であって、該ネットワークは宛先要素に結合された第1のノードと第1の要素とを含む複数のネットワーク要素を含んでおり、

(a) 該第1のノードから該宛先要素への第1のルートを特定する第1のルーチング情報と、該ネットワーク要素の該第1の要素から該第1のノードへの第2のルートを特定する第2のルーチング情報とを含むピンフレームを発生するステップと、

(b) 該第1のルートを介して該宛先要素へ該第1のピンフレームを伝送するステップと、

(c) 該宛先要素の第1のポートを通して該宛先要素でピンフレームを受信するステップと、

(d) 該ピンフレームに応答してエコーフレームを発生するステップであって、該エコーフレームは第1の接続情報と第3のルーチング情報と宛先要素識別子とを含み、該第1の接続情報は該宛先要素に隣接するネットワーク要素を特定し、該第3のルーチング情報は該第2のルーチング情報と該第1のポートの識別子とを含むものと、

(e) 前記の第3のルートを介して該宛先要素から該第1のノードへ該エコーフレームを伝送するステップと、

(f) 該第1のノードで該エコーフレームを受信するステップと、

(g) 該宛先要素識別子と該第1の接続情報とを含むことによって該ネットワークポロ

ジー情報を更新するステップと、
をそなえる方法。

【請求項 2】

該宛先要素は第 2 のノード及びルータのうちの 1 つである、請求項 1 に記載の方法。

【請求項 3】

次のステップすなわち、

(h) ネットワーク内の該ネットワーク要素のすべてに対し各ステップ (a) ~ (g) を
繰返すステップを更にそなえる、請求項 1 に記載の方法。

【請求項 4】

該ネットワークトポロジ内の各ループを特定するために、同一の宛先要素識別子を有す
るネットワーク要素を特定するステップを更にそなえる、請求項 3 に記載の方法。

【請求項 5】

次のステップすなわち、

(i) ネットワークを通して複数のルートを有するルータテーブルを決定するステップで
あって、該第 1 のノードから該ネットワーク要素の 1 つに至る各ルートは該ネットワー
クトポロジ情報にもとづいているものを更にそなえる、請求項 1 に記載の方法。

【請求項 6】

各ステップ (a) ~ (i) は、ネットワークが動作している間実行される、請求項 5 に記
載の方法。

【請求項 7】

各ステップ (a) ~ (i) は、ネットワークの初期化の間実行される、請求項 5 に記載の
方法。

【請求項 8】

ステップ (g) は不正に機能しているネットワーク要素と不正に機能しているネットワー
クリンクとを特定し、該ネットワークリンクは該ネットワーク要素を結合するものである
、請求項 1 に記載の方法。

【請求項 9】

ネットワークのネットワークトポロジ情報を決定するシステムであって、該ネットワー
クは宛先要素に結合された第 1 のノードと第 1 の要素とを含む複数のネットワーク要素を
含んでおり、

ピンフレームを発生するために該第 1 のノードに位置付けられたピンフレーム発生器であ
って、該ピンフレームは該第 1 のノードから該宛先要素への第 1 のルートを特定する第 1
のルーティング情報と、該ネットワーク要素の該第 1 の要素から該第 1 のノードへの第 2 の
ルートを特定する第 2 のルーティング情報とを含むものと、

該ピンフレーム発生器に結合され、該第 1 のルートを介して該宛先要素へ第 1 の該ピンフ
レームを送信するための第 1 のノードトランスミッタと、

該ピンフレームに回答してエコーフレームを発生するために該宛先要素内に位置付けられ
たエコーフレーム発生器であって、該エコーフレームは第 1 の接続情報と第 3 のルーチ
ング情報と宛先要素識別子とを含み、該第 1 の接続情報は該宛先要素に隣接するネットワ
ーク要素を特定し、該第 3 のルーティング情報は該第 2 のルーティング情報と該第 1 のポートの
識別子とを含むものと、

該宛先要素の第 1 のポートを通してピンフレームを受信し、該ピンフレームを該エコーフ
レーム発生器にトランスバレントに伝送するために、該宛先要素内に位置付けられ、該
第 1 のピンフレームを受信するために配置され、該エコーフレーム発生器に結合されたピ
ンフレームレシーバと、

前記の第 3 のルートを介して該宛先要素から該第 1 のノードへ該エコーフレームを送信
するために、該エコーフレーム発生器に結合されたエコーフレームレシーバと、

該宛先要素識別子と該第 1 の接続情報とを付加することによって該ネットワークトポロ
ジ情報を更新するために、該ピンフレームの識別子に結合され、該エコーフレームを受信
するために配置されたトポロジ更新ユニットと、

10

20

30

40

50

をそなえたシステム。

【請求項 1 0】

該エコフレームレシーバに結合され、該ネットワークトポロジ内の各ループを特定するために、同一の識別子を有するネットワーク要素を特定するためのトポロジーループ識別子を更にそなえる、請求項 9 に記載のシステム。

【請求項 1 1】

該トポロジ更新ユニットに結合され、該更新されたネットワークトポロジ情報にもとづいて、該第 1 のノードから各ネットワーク要素まで該ネットワークを通るルートを含むルーチングテーブルを発生するためのルーチングテーブル発生器を更にそなえる、請求項 9 に記載のシステム。

10

【請求項 1 2】

該宛先要素は第 2 のノード及びルータのうちの 1 つである、請求項 9 に記載のシステム。

【請求項 1 3】

次のステップすなわち、

(j) ネットワーク内の該ネットワーク要素の第 2 のものに対して各ステップ (a) ~ (g) を繰返すステップを更にそなえ、前記の第 2 のネットワーク要素は該第 1 の接続情報内で特定される、請求項 1 に記載の方法。

【請求項 1 4】

該第 1 のルーチング情報は以前に受信された接続情報から生成される、請求項 13 に記載の方法。

20

【請求項 1 5】

該ピンフレーム発生器は該第 1 の接続情報内で特定された第 2 の宛先ノードにアドレス指定された第 2 のピンフレームを発生する、請求項 9 に記載のシステム。

【発明の詳細な説明】

関連出願の相互参照

本願発明の主題は、下記に掲げる出願の主題と関連している。

出願番号____、弁護士用ドケット番号 2268、“非同期パケット交換”の名称で Thomas M. Wicki, Patrick J. Helland, Takeshi Shimizu, Wolf-Dietrich Weber および Winfried W. Wilcke によって 1996 年 2 月 22 日に出願、

出願番号____、弁護士用ドケット番号 2270、“低い待ち時間、高いクロック周波数プレシオ非同期パケットベースクロスバー・スイッチング・チップ・システムおよび方法”の名称で、Thomas M. Wicki, Jeffrey D. Larson, Albert Mu および Raghu Sastry によって 1996 年 2 月 22 日に出願、

30

出願番号____、弁護士用ドケット番号 2271、“パケットスイッチングネットワーク内のルーチングデバイス出力アクセス用調整方法および装置”の名称で、Jeffrey D. Larson, Albert Mu および Thomas M. Wicki によって 1996 年 2 月 22 日に出願、

出願番号____、弁護士用ドケット番号 2272、“電圧の揺れを少なくし、かつ、内部のブロック化データパスを生じさせないクロスバースイッチおよびその方法”の名称で、Albert Mu および Jeffrey D. Larson によって 1996 年 2 月 22 日に出願、

出願番号____、弁護士用ドケット番号 2274、“フロー制御プロトコル・システムおよび方法”の名称で Thomas M. Wicki, Patrick J. Helland, Jeffrey D. Larson, Albert Mu, Raghu Sastry および Richard L. Schober, Jr. によって 1996 年 2 月 22 日に出願、

40

出願番号____、弁護士用ドケット番号 2275、“相互接続の障害検出およびその位置特定方法および装置”の名称で、Raghu Sastry, Jeffrey D. Larson, Albert Mu, John R. Slice, Richard L. Schober, Jr. および Thomas M. Wicki によって 1996 年 2 月 22 日に出願、

出願番号____、弁護士用ドケット番号 2277、“多重ワード通信におけるエラー検出方法および装置”の名称で、Thomas M. Wicki, Patrick J. Helland および Takeshi Shimizu によって 1996 年 2 月 22 日に出願、

出願番号____、弁護士用ドケット番号 2278、“正のソース帰還をそなえたクロック動

50

作されるセンス増幅器”の名称で、Albert Muによって1996年2月22日に出願、参考として、上記の出願の全てを本願発明の全体に亘って取り入れている。

発明の背景

1. 発明の分野

本発明は、一般にネットワーク解析の分野、より具体的に言うと、コンピュータブートプロセス中又はオンラインオペレーション中に発信元ノードルーチングパケット交換網のトポロジをダイナミックに決定するためのシステム及び方法に関する。

2. 背景技術の説明

ネットワークの端子（ノード）間で情報を効率良くかつ高い信頼性で転送するためにコンピュータネットワークにより使用される技術は数多く存在する。このような1つの技術はパケット交換である。パケット交換においては、送信ノードは受信ノードに対しフレームを送る。メッセージは数多くの可変的サイズの部分に分割される。これらの部分はパケットと呼ばれる。各々のパケットにはデータ部分、パケットヘッダそして往々にしてエラー検出情報例えばパリティが含まれる。データ部分は、ネットワーク内のより高位のレイヤ例えばOSI基準モデル内で規定されているようなアプリケーションレイヤ、プレゼンテーションレイヤ及びセッションレイヤからのその他のプロトコル情報と共に送るべきメッセージ内の情報を内含している。パケットヘッダは、なかでもパケットシーケンス内のパケットの場所に関する情報を内含している。

ネットワークのノード間でパケットを搬送するために付加的な情報が必要とされる。この付加的情報はフレームヘッダ内に記憶される。1つのフレームヘッダが複数のパケットに付加され、パケットとフレームヘッダの組合せはフレームと呼ばれる。各々のネットワークは1つのフレームのサイズを制限し、従ってパケットがあまりに大きすぎて単一のフレーム内にはめ込むことができない場合、パケットは2つ以上のフレームに分割されることになる。パケット内に含まれた情報は、そのフレームのための究極的な宛先ノードを特定する情報を内含している。

発信元ノードと宛先ノードはネットワーク環境内で直接接続されていないことが多い。各ノードは、リンク又は信号回線によって少なくとも1つのルータに接続される。ネットワークは多くのルータを含むことができる。発信元ノードから宛先ノードまで1つのフレームを転送するためには、フレームは標準的に複数のリンク及びルータの中を横断する。ネットワークを通してのフレームの経路はフレームのルートと呼ばれる。ネットワーク内でフレームをルーチングするのに使用できる技術は数多く存在する。いくつかの従来のルーチング技術には、フラッドルーチング、ランダムルーチング、ディレクトリルーチング及び適応ディレクトリルーチングがある。フラッドルーチング技術は、発信元ノードと宛先ノードの間でネットワーク内の考えられる全ての経路に亘ってフレームを伝送する。フラディング技術は、有効な何らかの経路が存在する場合にフレームがうまく送受信されることを保証する。しかしながらネットワーク上のトラフィックは、各フレームが何度も送られることから著しく増大する。ランダムルーチング技術では、フレームは発信元ノードからルータまでランダムに送られ、宛先がそのフレームを受信するまでプロセスは反復する。各フレームはネットワークのまわりを「さまよい」、結局宛先ノードに達する。ルータから特定の信号回線を選択する確率は、トラフィックライン容量又はその他のネットワーク条件に基づいて偏倚させられる。ランダムルーチングに付随する問題点は、フレームが発信元ノードから宛先ノードまで直接的経路をとり得ないためにネットワークが多数のルータを有する場合にシステムの待ち時間が著しく高くなるという点にある。

第3のルーチング技術は、ディレクトリルーチング又は発信元ノードルーチングと呼ばれる。ディレクトリルーチングにおいては、各々の発信元ノードは、特定の宛先ノードのために選ばれるネットワーク内の経路を表示するルーチングテーブルを内含する。このルーチングテーブルは、「オフライン」で開発されダイナミックに修正できない。しかしながらメモリの制約条件のため各々の宛先ノードについて1つの経路しか含まれずその経路上の1つのリンクが破損していることが必要とされる場合、この破損経路上で走行するフレームはいずれもうまく伝送されないだろう。4番目のルーチング技術は適応ディレクトリ

10

20

30

40

50

ルーチングと呼ばれる。適応ディレクトリルーチングにおいては、各ルータは、隣接するルータ又はノードのためのネットワーク内の最良のルートを表示するメモリ内のルーチングテーブルをもつ。ルーチングテーブル内のエントリは実時間で修正できる。従来の適応ルーチング技術がもつ1つの問題点は、ルータが例えばフレーム遅延といった単数又は複数のネットワークパラメータに基づいてフレームのための最も効率の良いルートを決定するという点にある。ルータはフレームをダイナミックにルーチングするため、各ルータの待ち時間はディレクトリルーチング技術における各ルータの待ち時間よりも大きい。フレームのルータは発信元ノードによって完全に決定されないことから、発信元ノードルーチングネットワークによって適応ディレクトリルーチング技術が利用されることはない。

発信元ノードルーチングネットワークの利点は、ルータがルーチング命令を作り出す必要がないということすなわちルータは発信元ノードによってフレームをどこにルーチングすべきかの命令を受けているということにある。従って、このようなルータのコストはダイナミックルーチングを実行するルータのコストよりも低い。発信元ノードルーチングネットワークの第2の利点は、これらのネットワークの待ち時間が縮減されるということにある。すなわち、各ルータを通してフレームをルーチングするのに必要とされる時間は例えばフラッドルーチング又はランダムルーチング技術の場合のようにダイナミックルーチングを利用するネットワークに比べると減少している。発信元ノードルーチングネットワーク内での待ち時間は減少するがこれはルータがネットワークを通してのフレームの経路をダイナミックに決定する必要がなく、その代りに受信するルータを決定するのにルータはフレームヘッダを単に読みとるだけであり、それはルーチング情報をダイナミックに決定するよりも迅速な作業であるからである。

ネットワークを通してのフレームのルートは、フレームが発信元ノードルーチングネットワーク内で生成された時点で決定されることから各ノードは、ネットワークを通してのその他の全てのノードへの少なくとも1つの経路を特定するルーチングテーブルを有する。上述の通り従来の発信元ノードルーチングシステムにおいては、このようなルーチングテーブルは予め決定されており静的である。すなわちルーチングテーブルは、システムがオフラインであるときに構築される。さらに、ルーチングテーブルに対する修正は、ノードがオフラインのときにも発生する。いくつかの発信元ノードルーチングネットワーク内では、任意のノード内のルーチングテーブルは、ネットワーク全体が作動中でないときのみ修正される。

ルーチングテーブルの修正は、リンクが非作動状態であり、もう1つのリンク、ルータ又はノードが付加されているか又は削除されている場合又は重大なトラフィック状況下ではネットワークの特定の部分上のフレームトラフィックのフローを低減させるためにシステムブート時間中に発生し得る。場合によっては、1つのノードとルータの間又は2つのルータの間のネットワークリンクが非作動状態となり（故障）、従ってこの故障したリンクを横断するフレームトラフィックは中断される。ネットワークを通してのフレーム経路は予め決定され静的であることから、この故障したリンクを通してルーチングされる全てのフレームはその宛先に到達せず、失われたものとみなされる。ルーチングテーブル内に記憶されたノードAからノードBまでのルートのみがフレームが破損したリンク内を通過することを要求しているようなルーチングテーブルをノードAが有している場合、ネットワーク内でその他のルートが利用できたとしてもいかなるフレームもノードAからうまく送られず又ノードBでもうまく受信されない。同様にしてノード、ルータ又はリンクがネットワークに付加された場合、ルーチングテーブルは、これらの付加的なネットワークエレメントを説明するため修正されなくてはならない。ルータが非作動状態となったならば、そのときこのルータを通してルーチングされた全てのトラフィックは失われることになる。上述の状況下では、例えば、ルーチングテーブルはネットワークトポロジー内の変更、例えば付加的なルータ、ノード又はリンク又は非作動状態のルータ、ノード又はリンクを反映するために修正されなくてはならない。しかしながら上述のように、いくつかの従来の発信元ノードルーチングネットワークでは、ルーチングテーブルが存在するノード又はネットワーク全体が作動していない時点でのみルーチングテーブルを修正することが

10

20

30

40

50

できる。

ルーチングテーブルを修正する目的でノードを非作動状態にすることは、ネットワーク性能及びノード性能の点からみて高くつくことである。ノードが大量のフレームトラフィックを取扱っている場合には、フレームのうちの小さい割合のものだけが非作動状態のリンクの影響を受け、ノードをオフに切換えることによってノードへ及びノードから流れるトラフィックの全てを停止することが可能である。その上、非作動状態のリンクを通してルーチングできる全てのノードは同様に、そのルーチングテーブルを修正するため作動を停止しなくてはならない。

適応ディレクトリルーチング技術においてさえ、発信元ノード内のルーチングテーブルを修正する能力は、ネットワーク性能の低下及びネットワーク待ち時間の増大を対価として

10

もたらされる。必要とされるのは、ネットワークが作動を停止したり又はネットワークの性能を著しく低減させる必要なく、発信元ノードルーチングネットワークのトポロジを効率良くかつダイナミックに決定するためのシステム及び方法である。この方法は高価なハードウェアを要することなくネットワーク性能に対する影響を最小限におさえなくてはならない。

発明の概要

本発明は、ネットワークの性能への影響を最小にして、かつ実行するために高価なハードウェアを必要とすることなく、発信元ノードルーチングネットワークのトポロジをダイナミックに決定するためのシステム及び方法である。本発明では、発信元ノードはピンフレーム (ping frame) を発生する。該ピンフレームは、データフレームからピンフレーム

20

を区別するために該ピンフレームのヘッダに特別のコードを含む。発信元ノードは該発信元ノードに結合される第1のルータに該ピンフレームを伝送する。該第1のルータは該フレームをピンフレームとしてトランスペアレントに特定し、該ピンフレームを内部制御フレームハンドラにルーチングする。該制御フレームハンドラは該発信元ノードに返送されるエコーフレームを生成する。該第1のルータは該ピンフレームが受信されるポートを特定し、エコーフレーム識別子とともに該エコーフレームのヘッダ内にこの情報を配置する。該エコーフレームの本体は、第1ルータのIDコード、該ピンフレームが受信されるポートのID及び接続情報すなわちネットワーク要素が該第1のルータの各ポートに接続されているか否かについての情報を含んでいる。

発信元ノードはエコーフレームを受信し、該受信されたエコーフレーム内の接続情報にも

30

とづいて、ピンフレームが送られなかったルータとノードとを特定する。他のノード又はルータ例えば第2のルータを選択した後、該発信元ノードは第2のピンフレームを発生する。該第2のピンフレームは第1のピンフレームと同様の態様で発生され動作する。しかし、該第2のピンフレームは異なる宛先を有しているので、該第2のピンフレームは第1のピンフレームとは異なる情報を含む。このピンフレームは、ルータによってなされるのに必要なルーチングの決定を最小限にするために、フレームヘッダ内にルーチング情報を含み、フレーム本体内にリターンルーチング情報を含む。第2のピンフレームは第1のルータを介して第2のルータへ送られる。該第2のルータはそのフレームがピンフレームであることをトランスペアレントに決定し、該フレームをその制御フレームハンドラに送る。該制御フレームハンドラは該ピンフレームヘッダを削除し、該ピンフレームの本体内の

40

ルーチング情報からエコーフレームヘッダを生成する。第2のルータは該ピンフレームが受信されるポートを特定し、この情報をエコーフレーム識別子とともにエコーフレームのヘッダ内に配置する。エコーフレームの本体は第2のルータIDコード、ピンフレームを受信したこのルータのポート番号、及び接続情報を含む。次いで第2のルータはエコーフレームを発信元ノードへ返送する。

50

ことができ、また高速のトラフィック期間中に実行されてもシステムの待ち時間に最小限の増加をもたらすにすぎない。その理由は各ピンフレームが小さくて、宛先ルータ又はノードの制御フレームハンドラにトランスペアレントに送られるからである。

【図面の簡単な説明】

図1は、本発明の好ましい実施形態に従ったノード及びルータを含む発信元ノードルーチングネットワークトポロジーの一例を示す図である。

図2は、本発明の好ましい実施形態に従った発信元ノードルーチングネットワーク内のルータのより詳細な図である。

図3は、本発明の好ましい実施形態に従ったクロスポイントマトリクスのより詳細な図である。

10

図4は、本発明の好ましい実施形態に従った制御フレームハンドラのより詳細な図である。

図5は、本発明の好ましい実施形態のトポロジー探索技術の流れ図である。

図6は、本発明の好ましい実施形態に従ったピンフレーム (ping frame) 生成技術の流れ図である。

図7は、本発明の好ましい実施形態に従ったエコフレーム生成技術の流れ図である。

図8は、本発明の好ましい実施形態に従ったトポロジー情報更新技術の流れ図である。

図9(a) - 9(g)は、本発明の好ましい実施形態に従ったトポロジー探索技術の例である。

図10は、本発明の好ましい実施形態に従った発信元ノードルーチングネットワーク内のノードのより詳細な図である。

20

好ましい実施形態の詳細な説明

本発明の好ましい実施形態についてここで図面を参照しながら説明する。なお図中、同じ参照番号は、同一の又は機能的に類似の要素を表わしている。同様に、図中、各々の参照番号の最も左の数字は、その参照番号が最初に使用されている図に対応している。

図1は、本発明の好ましい実施形態に従ったノード及びルータを含む発信元ノードルーチングネットワークトポロジーの例である。図1中のネットワークは、8つのノード102、すなわちノードA - H、102A - H及び7つのルータ104A - Fを含んでいる。各々のノード102及び各々のルータ104は、トポロジー探索技術においてルータ及びトポロジーループを特定するために発信元ノードによって用いられる一意のIDコードを有する。例えば、第1のルータ104AのためのIDコードは20である。ノード及びルータは、合わせてネットワークエレメントを成し、リンク106により接続されている。各リンク106は、全二重通信信号回線である。すなわち、各リンクは、相対する方向で2つの独立した信号を同時に搬送することができる。つまり1つの信号は発信元ノード102Aからルータ104Aにより受信され、1つの信号は同時にルータ104Aにより発信元ノード102Aに伝送される。従って各々のリンク106は、図2及び図10に示されこれらの図を参照しながら以下で記述されるように、2つの半二重通信信号回線として表わすことができる。以下で記述する通り本発明は、ネットワークのトポロジーを各ノード102がダイナミックに決定できるようにするシステム及び方法である。

30

図2は、本発明の好ましい実施形態に従った発信元ノードルーチングネットワーク内のルータのより詳細な例示である。図2は、6つのポート202A - Fを含んでいる。変形実施形態においては、ルータは異なる数のポート、例えば2つのポート又は30のポートを内含している。各ポート、例えばポート1、202Aは、2重通信信号回線106に結合されている。各々のポート202は同様にアービトレーションユニット204にも接続されている。アービトレーションユニット204については、以上でクロスリファレンスされ本書に全体が参考として内含されているJeffrey D.Larson、Albert Mu及びThomas M.Wickiにより1996年2月22日に提出された「パケットスイッチングネットワーク内のルーチングデバイス出力アクセス用調整方法および装置」という題の同時係属特許出願を参考にしながら、以下でより詳しく説明する。各ポートは、バッファ信号回線210を介してクロスバースイッチ206に接続されている。クロスバースイッチ206のオペレーションについては以下で図3を参考

40

50

にしながらより詳しく記述されており、以上でクロスリファレンスされその全体が本書に参考として内含されているAlbert Mu及びJeffrey D.Larsonにより1996年2月22日に提出された「電圧の揺れを少なくし、かつ、内部のブロック化データパスを生じさせないクロスバースイッチおよびその方法」という題の同時係属特許出願の中でやや修正された形で記述されている。好ましい実施形態においては、各々のバッファ信号回線210Aは、以下に記述する各々のバッファをクロスバースイッチ206と接続するための6つの接続信号回線を含んでいる。ルータ104のオペレーションについては以下で図3～9を参考にしながらより詳細に記述されている。

図10は、本発明の好ましい実施形態に従った発信元ノードルーチングネットワーク内の例えばノードA 102Aといったノードをより詳しく示している。図10には、トポロジー探索プロセスに關与するノードA 102Aの部分が含まれる。ノードは、以上で記述された2重通信信号回線106Aに結合されている。フレームがポート1 1006により受信される。ポート1 1006のオペレーションは、以下で記述するポート202のオペレーションに類似している。ポート1 1006は、受信したフレームが以下で記述するようにデータフレーム、ピンフレーム又はエコフレームのいずれであるかを決定する。フレームがピンフレームであるならば、そのフレームは制御フレームハンドラ310に送られる。制御フレームハンドラ310は、以下でより詳細に記述されている。フレームがデータフレームである場合には、そのフレームはデータユニット1002に送られる。フレームがエコフレームである場合、フレームはピンユニット1004に送られる。ピンユニット1004はエコフレーム内に含まれた情報を読み取りこの情報に基づいてルーチングテーブルを更新する。さらに、ピンユニット1004はピンフレームを生成する。制御フレームハンドラ310及びピンユニット1004は、ハードウェア又はソフトウェアのいずれででも実施できる。ピンユニット1004のオペレーションについては以下でさらに詳しく記述する。

図3は、本発明の好ましい実施形態に従ったルータ104及びクロスポイントマトリクス206のさらに詳しい図である。ルータ104の各々のポート202はポート入力端302、例えばポート1入力端302Aを内含する。各々のポート入力端302は、例えば、ポート入力端302で受信されたフレームを記憶する6つのバッファ（図示せず）を内含する。各バッファは、クロスバースイッチ206に結合される。図3を簡略化する目的で、各々のポート302をクロスバースイッチ206に結合するバッファ信号回線210のみが示されている。上述の通り、各ポートは6つのバッファを有し、各々のバッファはバッファ信号回線210を介してクロスバースイッチ206に接続される。従って、6つのポート202の各々は、合計36のバッファ信号回線210について6つの接続を含む。同様にして、36のバッファ信号回線210の各々は、7つのクロスポイント回線306に結合される。図3を簡略にする目的で、各ポート202について7つのクロスポイント回路306だけが示されている。

ルータ104は、7つのアービトレーションユニット304を内含する。これらのアービトレーションユニットのうちの6つの304A～Fはクロスバースイッチ出力信号回線312A～Fを介して出力信号回線212A～Fに結合されている。7番目のアービトレーションユニットは、クロスバースイッチ出力信号回線312Gを介して制御フレームハンドラ310に結合される。さらに第7のアービトレーションユニット304Gは制御回線308を介して制御フレームハンドラ310に直接結合される。クロスバースイッチ出力信号回線312A～Gは、クロスポイント回路306を介して各々の出力信号回線212に接続される。

ここでルータ108のオペレーションの全体的概観を記す。ルータ104のポート202、例えばポート1 202Aの中でフレームが受信される。このフレームは、フレームヘッダの中の情報に基づいて受信したフレームがデータフレームであるかピンフレームであるかを見極めるデータシンクロナイザ（図示せず）によって受信される。受信したフレームがデータフレームである場合、このデータフレームは、ポート1入力端302A内のバッファの中に記憶される。アービトレーションユニット304は、各バッファ内のデータがポート202を通してもう1つのノード又はバッファまで伝送され得るか否かを決定する。ポート1 202Aにてデータフレームが受信された場合、データフレームが送られるべき出力ポートはアービトレーションユニット304によって決定される。データフレームを出力ポート上で送

10

20

30

40

50

ることができるということをアービトレーションユニット304が表示した時点で、アービトレーションユニットはデータフレームをもつバッファをクロスバースイッチ206を用いて適切な出力ポートに結合する。クロスバースイッチ206のオペレーションのより詳細な説明については、先に引用したAlbert Mu及びJeffrey D.Larsonにより1996年2月22日に提出された「電圧の揺れを少なくし、かつ、内部のブロック化データパスを生じさせないクロスバースイッチおよびその方法」という題の特許出願の中に提供されている。アービトレーションユニット304のオペレーションのより詳しい記述については、先に引用したJeffrey D.Larson、Albert Mu及びThomas M.Wickiにより1996年2月22日に提出された「パケットスイッチングネットワーク内のルーチングデバイス出力アクセス用調整方法および装置」という題の特許出願の中で提供されている。

10

受信したフレームがピンフレームである場合、そのピンフレームはバッファ内に記憶され、7番目のアービトレーションユニット304Gがクロスポイント回路306を介して制御フレームハンドラ310にピンフレームを接続する。従って、クロスバースイッチ208は、データフレームと同じ要領で交換されることから、ピンフレームをトランスペアレントに取扱う。しかしながら、出力ポートに交換される代りに、ピンフレームは制御フレームハンドラ310に交換され、クロスバースイッチ208は同時に、適切な出力信号回線212へとその他のバッファ内に記憶されたその他のフレームを交換することができる。ピンフレームは、制御フレームハンドラ310により受信される。制御フレームハンドラ310は、ルータ情報、受信ポート情報及び接続情報を含むエコフレームを作り出す。このとき、制御フレームハンドラ310はエコフレームを発信元ノードに送り返す。制御フレームハンドラ310のオペレーション及びピンフレームを受信しエコフレームを送信するときのルータのオペレーションについては、以下で図4-9を参照しながら記述する。

20

図4は、本発明の好ましい実施形態に従った制御フレームハンドラ310のさらに詳しい図である。制御フレームハンドラ310は、ルータ情報モジュール402、接続モジュール404、エコフレームヘッダレジスタ406及びエコフレーム発生器408を含む。ルータ情報モジュール402は、ルータIDコードを内含する。接続モジュール404は、各ポート202がいかに関係づけられるかすなわち、ポート202がもう1つのポート又はノードに接続されるか又はポートが接続されていないか（開放）を記述する情報を含んでいる。エコフレームヘッダレジスタ406は、エコフレームヘッダ情報を記憶する。以下で記述する通り、エコフレームヘッダレジスタ406は、ピンフレーム本体内の最初の8.5バイトの情報を、つまりピンフレームヘッダ後の最初の8.5バイトを内含している。エコフレーム発生器408はピンフレームに回答してエコフレームを生成する。エコフレームヘッダレジスタ406はクロスバースイッチ信号回線312Gを介してピンフレームを受信し、ルータ情報モジュール402、接続モジュール404及びエコフレームヘッダレジスタ406からの情報を受信する。さらに、エコフレーム発生器408は、受信ポート識別子を含む7番目のアービトレーションユニット304Gから制御信号308を受信する。受信ポート識別子は、ピンフレームが受信されるポートを特定する。7番目のアービトレーションユニット304Gは受信ポート識別子をエコフレーム発生器408に伝送し、この発生器はこの情報をエコフレームヘッダ内に入れ、かくしてフレームはピンフレームと同じルートを用いて発信元ノードへと伝送し戻される。さらに、受信ポート識別子はエコフレーム内に記憶され、受信ポート識別子は、ピンフレーム本体内に記憶されているようなエコノードのための将来のルーチング情報を決定する際に、発信元ノードによって利用される。エコフレーム発生器のオペレーションについては、図5~9を参照してより詳しく説明する。

30

40

図5は、本発明の好ましい実施形態のトポロジ探究技術の流れ図である。この技術については、図1及び図9(a)~9(g)を参考にしながら記述する。図9(a)~9(g)は、本発明の好ましい実施形態に従ったトポロジ探究技術の例である。本発明は、例えば発信元ノード102Aといった発信元ノード内のルーチングテーブルがシステムのブート時間で初期化される必要があるか又は正確でなく修正を要するような状況下で使用され得る。上述のようにルーチングテーブルの修正は、リンクが作動状態にない場合、もう1つのリンク、ルータ又はノードがネットワークに付加されるか又は削除される場合又は、

50

例えばトラフィックフローが高いために単数又は複数のリンクが伝送の遅延をひき起こしている場合に起こり得る。前述の通り、ノードとルータの間、又は2つのルータの間のネットワークリンクは時として非作動状態となり、従ってこの故障したリンクを横断してのフレームトラフィックは中断される。発信元ノードルーチングネットワーク内のフレーム経路は予め定められ静的であることから、この故障したリンクを通してルーチングされる全てのフレームは、その宛先に到達せず、失われたものとみなされる。発信元ノード102Aが、発信元ノード102Aから例えばノード102Fといった宛先ノードへの1つのルートのみがルーチングテーブル内に記憶されているようなルーチングテーブルを有し、このルート上のリンクが故障している場合、発信元ノード102Aから宛先ノード102Fまでいかなるトラフィックもうまく伝送されない。しかしながら、伝送の成功を結果としてもたらすような代替的ルートを使用できる可能性もある。このような状況下では、発信元ノードルーチングネットワークを通しての代替的ルートを特定するようにルーチングテーブルを修正しなくてはならない。同様に、ネットワークにノード、ルータ又はリンクが付加された場合、ルーチングテーブルはこれらの付加的ネットワークエレメントを説明するように修正されなくてはならない。ルータが非作動状態になった場合、このルータ内にルーチングされた全てのトラフィックは失われることになる。このような状況下では、ネットワークトポロジの変更を反映するようにルーチングテーブルを修正しなくてはならない。

本発明の1つの実施形態においては、トポロジ探索システム及び方法は、物理的に分散型であるものの論理的には共有型であるメモリのアーキテクチャをもつシステム内の発信元ノードルーチングネットワークの中で使用される。従って、1つのノード例えば処理ノードは、異なる場所に物理的に位置特定され異なるノードに結合されている記憶場所に対するアクセスを要求することができる。要求を行なっているノードすなわち発信元ノードはメモリの場所を特定し、ネットワークは迅速にデータを検索する。1つのリンクが非作動状態にある場合、データ要求は不成功に終わる。このようなシステムにおいては、従来のシステムにおいて必要とされるように単数又は複数のノード内のルーチングテーブルを修正するため又は例えばルータ104がルーチング決定を下すシステムの場合のようにネットワークの待ち時間を増大させる技術を用いてルーチングテーブルを修正するためにシステム全体をオフに切替えるのは費用が高くつくことである。本発明は、システムの実行及び待ち時間に著しく干渉することなくネットワークのトポロジをダイナミックに決定しルーチングテーブルを更新するためのシステム及び方法である。

本発明のトポロジ探索手順は、システムブート時点で予め定められた数の連続的に伝送されたフレームを「失なう」こと、及び予め定められた時間の後に手順を実行することを含め、数多くの方法でトリガーされ得る。図9(a)を参照すると、最初のピンフレームを伝送する前に発信元ノード102Aにわかっているネットワークトポロジが示されている。すなわち、発信元ノード102Aは、それが1つの回路エレメント例えばノード又はルータに接続されているもののこの要素を記述する特定のいかなる情報も有していないということを知っている。発信元ノードはピンフレームを生成する(502)ことによってトポロジ探索プロセスを開始する。ピンフレーム生成技術は、図6及び図9(b)を参考にし以下でさらに詳しく記述されている。

図6は、本発明の好ましい実施形態に従ったピンフレーム生成技術の流れ図である。発信元ノード102Aは、宛先エレメントを決定する(602)。図9(b)に示されているように第1のピンフレーム902Aを生成するとき、残りのネットワークエレメントに対する単一の接続のみを有する発信元ノード102Aが、それに接続された唯一の要素を選択することになる。発信元ノード102Aはピンフレームヘッダ904Aを生成する(604)。ピンフレームヘッダ904Aは、ルーチング情報及びピンフレームIDコードを含む。好ましい実施形態においては、ピンフレームIDコードは、値7の3ビット2進表示すなわち111であり、フレームヘッダの3つの最上位ビットの中に位置づけされているが、これらのビットの正確な場所は、変形実施形態において異なる可能性がある。好ましい実施形態においては、最初の3つのビットは受信ルータがフレームを送る出力ポートを特定する。これら3つのビットは、フレームを適切な出力ポートに接続するべくアービトレーションユニット304を特

10

20

30

40

50

定する。例えば3つのビットが1の2進表示を表わしている場合、第1のアービトレーションユニット304Aは、出力ポート1(図3中の「01」)を用いて出力されるべき特定のバッファの中にデータが記憶されるという通知を受ける。

ルーチング情報が7の2進表示である場合、これら3つのビットは7番目のアービトレーションユニット304Gを特定する。7番目のアービトレーションユニット304Gは、以下で記述するように、適切なクロスポイント回路306を活動化させることによりクロスバースイッチ206にバッファを制御フレームハンドラ310へと接続させる。ゼロの3ビット2進表示が、受信要素に対し、受信要素がノードである場合には受信したフレームがその宛先に到達したことを、又要素がルータである場合にはルーチングテーブル内のエントリ不良又は伝送エラーといったようなエラー条件が発生したことを、通知する。従って、図9(b)では第1のピンフレームヘッダ904Aが示され、111に等しい3つの最上位ビット及び000に等しい次の3つの最上位ビットを有する。発信元ノード102Aは、以下で記述されているようにエコフレームのためのルーチング情報を含むフレーム本体906Aを生成する(606)。しかしながら、第1のエコフレームはそれが受信されるのと同じポートを介して直接発信元ノードに伝送されることから、第1のピンフレームのためにこのようなルーチング情報は全く必要とされない。従って、第1のピンフレーム本体906Aはいかなるルーチング情報も内含していない。第1のピンフレームを生成した(502)後、発信元ノードはピンフレームを伝送する(504)。

ピンフレームは宛先エレメント例えばルータ104Aによって受信される(506)。上述のとおり、ピンフレームヘッダ904Aは、ピンフレームが受信される7番目のアービトレーションユニット304Gを通知するルータ104Aのポート3入力といったポート入力によって読み取られる。ポート3入力は、フレームヘッダ904Aの中の3つの最上位ビットすなわち111の値によりピンフレームとしてフレームを特定する。7番目のアービトレーションユニット304Gは、クロスポイント回路306Cに対し7番目のクロスバースイッチ出力信号回線312Gを接続するよう命令し、入力ポート302Cに対しピンフレーム902Aを伝送するように命令する。その結果として、ピンフレーム902Aは、制御フレームハンドラ310によって受信される。制御フレームハンドラ310は、第1のピンフレームに応答して第1のエコフレームを生成する(510)。エコフレーム生成技術は、以下で、図7及び図9(b)-(c)を参考にして記述される。

図7は、本発明の好ましい実施形態に従ったエコフレーム生成技術の流れ図である。制御フレームハンドラ310は、アービトレーションユニット304Gから制御信号回線308を介して、ピンフレームを受信した入力ポート302を受信する。ルータ情報モジュール402は、ルータの一意IDコードを含む。例えば第1のルータ104AのためのルータIDコードは20である。制御フレームハンドラ310は、ピンフレームからピンフレームヘッダを削除し(702)、エコフレームヘッダレジスタ406の中に第1のピンフレーム本体906Aの第1のラインを記憶する。好ましい実施形態では、エコフレームヘッダは8.5バイトである。制御フレームハンドラ310は、接続モジュール404からルータの接続情報を検索する。接続情報は、各ポートの状態すなわちネットワーク要素がそれに接続されているか否かを特定する。例えば、最初のルータ104Aについては、ポート1, 2, 3及び5がネットワーク要素に接続され、ポート4及び6は接続されていない。制御フレームハンドラ310は又、制御信号回線308上で7番目のアービトレーションユニット304Gからの制御信号も受信する。制御信号を受信した後、エコフレーム発生器408はエコフレーム908Aを作り出す(704)。エコフレームヘッダ910Aの3つの最上位ビットは、第1のピンフレームを受信したポートすなわちポート3又は011バイナリに等しい。最初のルータ104Aは第3のポートを介して発信元ノード102Aに直接連結されていることから、次の3つの最上位ビットは000であり、エコフレームがその宛先に達したことを表わしている。エコフレームヘッダ910は同様に、データフレームとは異なりエコフレームとしてフレームを特定する。例えばエコフレームヘッダの最下位ビット内にあるフラグ識別子をも含んでいる。このときエコフレーム発生器はエコフレーム本体912Aを作り出す(706)。このエコフレーム本体は、ルータIDコード例えば20、受信ポート識別子例えばポート3及び接続情報

10

20

30

40

50

を含む。エコフレーム発生器408は、エコフレームヘッダ910A内で特定されるように、ポート3を介して発信元ノードに対し第1のエコフレーム908Aを送信する(512)。第1のエコフレームは、発信元ノード102Aによって受信され(514)、発信元ノード102Aはそのネットワークトポロジ情報を更新する(516)。

図8は、本発明の好ましい実施形態に従ったトポロジ情報更新技術の流れ図である。発信元ノード102Aは、あらゆるトポロジーループを特定する(802)。この特定は、ネットワーク要素IDコードを比較することによって達成される。次に発信元ノードは、ルータIDコード及び接続情報を用いてネットワークトポロジ情報を修正する(804)。図9(d)は、第1のエコフレームを受信した後に発信元ノード102Aが知っているネットワークトポロジを例示する。ノードAは、それが20というIDコードをもつ第1のルータ104Aに対してこの第1のルータ104Aのポート3を介して接続されていることを知っている。第1のルータ104Aのポート1、2及び5はその他のネットワークエレメントに接続され、第1のルータ104Aのポート4及び6は接続されておらず、ポート3は発信元ノード102Aに接続されている。発信元ノード102Aはトポロジ情報を検査し、ネットワークの全てのネットワーク要素が探査されたか否かを決定する(518)。この例においては、第1のルータ104Aのポート1、2及び5に接続された要素は探査されていなかった。従って、これらの未探査のエレメントの1つ、例えば第1のルータ104Aのポート1に接続されたネットワークエレメントが選択され、ステップ502-518は、以下に記す通りに反復される。

全てのネットワークエレメントが探査されていなかったことから、発信元ノード102Aは第2のピンフレームを生成する(502)。上述のとおり、第2のピンフレームを生成するための技術は、第1のピンフレームを生成するための技術と同じである。第1のピンフレームと第2のピンフレームの間の差には、ピンフレームヘッダ内のピンフレームルーチング情報及びピンフレーム本体内のエコフレームルーチング情報が含まれる。第2のピンフレームを生成するための技術は、図6及び図9(e)を参照して以下に記述される。発信元ノード102Aは、更新された接続トポロジ情報に基づいてルーチング情報を決定する。発信元ノード102Aから宛先エレメントまでのピンフレームのルートは、発信元ノードから第1のルータ104A、そして次に第1のルータ104Aのポート1を通して、この第1のルータ104Aに接続された宛先エレメントまでである。発信元ノード102Aは、このルーチング情報に基づいて第2のピンフレームヘッダ904Bを生成する(604)。具体的に言うと、第2のピンフレームヘッダ904Bの3つの最上位ビットは、第2のピンフレーム902Bが送られるときに通る第1のルータ104A内のポートすなわちポート1に対応する。従って、これら3つのビットは、1の2進表示、すなわち001である。次の3つの最上位ビット、すなわち前の3つのビットが第1のルータ104Aによって削除された後で第2のピンフレームヘッダ904Bの中に残ることになる3つの最上位ビットは、上述のとおり第2のネットワークエレメントがこれらのビットを読みとり受信されたフレームがピンフレームであることを決定することになることから、7の2進表示すなわち111である。発信元ノード102Aは第2のピンフレーム本体906Bを生成する(606)。上述の通り、第2のピンフレーム本体906Bの第1のラインには、エコフレームのためのルーチング情報が含まれる。エコフレームルーチング情報は、宛先ノード内で必要な論理を最小限にするため、宛先ノードの代りに発信元ノード102Aによって決定される。このラインの最初の3つのビットは、第2のエレメント内に第2のピンフレーム902Bが受信されるときに通るポートとなる。このポートは同様に、第2のエレメントから第1のルータ104Aまで第2のエコフレームを送信することになる。発信元ノード102Aは、まだこのポート情報を有しておらず、従ってこれらの3つのビットは「ドントケア」ビットであり「×」により表わされる。しかしながら、次の3つの最上位ビットは、第1のルータ104Aによって受信された時点でエコフレームがとることになるルートを表わす。エコフレームは、第1のルータ104Aの第1のポートによって受信され、第3のポートを通して発信元ノード102Aまで伝送される。従って、次の3つのビットは3の2進表示すなわち011を表わす。第1のルータ104Aがポート3を通して第2のエコフレームをルーチングした後、このエコ

10

20

30

40

50

フレームは、発信元ノード102Aによって受信されることになる。上述の通り、ゼロの3ビット表示すなわち000が発信元ノードによって解釈されてフレームがその宛先に到達したことを意味する。従って第2のピンフレーム902Bは図9(e)に示された値に等しい。第2のピンフレーム902Bを生成した後、発信元ノード102Aは、この第2のピンフレーム902Bを第1のルータ104Aに伝送する(504)。

第1のルータ104Aは第2のピンフレーム902Bを受信し、第2のピンフレームヘッダ904Bを読み取る。第2のピンフレームヘッダ904Bの最初の3つのビットは、第2のピンフレーム902Bがそれを通して送られるべきポートを特定する。上述の通りこれら3つのビットは001に等しい。従って、第1のルータ104Aはピンフレームヘッダ904Bからこれら3つのビットを削除し、残りのルーチングビットについて3つのビットシフトオペレーションを実行し、第1のルータ104Aのポート1に接続されたネットワークエレメントに対し修正された第2のピンフレーム902Bをルーチングする。このネットワークエレメントは第2のルータ104Dである。第2のルータ104Dは、修正された第2のピンフレーム902Bを受信し(506)、第2のピンフレームヘッダ904Bを読み取る。第1の3つのビットはこのとき111に等しい。上述の通り、これらのビットは111に等しいことから、第2のピンフレームは第2のルータ104D内で制御フレームハンドラ310にルーチングされる。第2のルータ104Dは、上述の技術を用いて第2のエコフレーム908Bを生成する。特定的には、第2のルータ104D内の7番目のアービトレーションユニット304Gから制御信号を受信した後、エコフレーム発生器408は第2のピンフレームヘッダ904Bを削除し(702)、第2のエコフレームヘッダ910Bを作り出す。第2のエコフレームヘッダの3つの最上位ビットは、第2のエコフレーム908Bが最初送られるときに通ることになるポートすなわちポート5又は2進101を表わす。3つのビットフィールドの次の2つのセットは、上述の発信元ノード102Aによってセットされるように001及び000に等しい。第2のエコフレームヘッダ910Bは、受理されたフレームがデータフレームと違ってエコフレームであることを発信元ノード102Aに示すフラグを同様に内含する。第2のルータ104Dは同様に第2のエコフレーム本体912Bをも作り出す(706)。第2のエコフレーム本体912Bは、第2のルータ104DのIDコード例えば23及び入力ポート識別子すなわちポート5を内含する。第2のエコフレーム本体912Bは同様に第2のルータ104Dのための接続情報も内含している。特定的に言うと、接続情報は、第2のルータ104Dの6つのポート全てが使用中であることを表わす。

第2のルータ104Dは、第2のルータ104Dのポート5を通して第1のルータ104Aに対し第2のエコフレーム908Bを伝送する(512)。第1のルータ104Aは上述の通り、第2のエコフレーム908Bを受信し、エコフレームヘッダ910Bを読み取って、第2のエコフレーム908Bをルーチングすべき次のポートを特定し、第2のエコフレームヘッダ910Bの第1の3ビットを削除し、残りのルータビットをシフトする。第2のルータ104Dは第2のエコフレーム908Bをエコフレームとして特定しない。その代り第2のエコフレーム908Bは、あたかもデータフレームであるかのごとく、第1のルータ104Aにより処理される。従って第1のルータ104Aは、データフレームと異なるやり方で第2のピンフレーム902B又は第2のエコフレーム908Bを取り扱う必要がない。この特長は、これらの機能を実行するのに付加ロジックを必要とするネットワークに比べた場合、ハードウェアの複雑性を減少させ、ネットワーク待ち時間を減少させる。第1のルータ104Aは、第2のエコフレームヘッダ910Bの中で表示されているように第1のルータ104Aのポート3を介して発信元ノード102Aに第2のエコフレーム908Bを伝送する。発信元ノード102Aは第2のエコフレーム908Bを受信し(514)、トポロジー情報を更新する(516)。上述の通り、トポロジー情報を更新するステップ(516)については、図8にさらに詳しく記述されている。発信元ノードはあらゆるトポロジーループを特定する(802)。この例では、いかなるトポロジーループも特定されなかった。トポロジーループが存在する場合、発信元ノードは、ネットワークエレメントIDコードを比較することによってこのループを特定する。2つのIDコードが一致する場合、発信元ノード102Aは、下記の通りこれら2つの要素を等しいと定義する。

10

20

30

40

50

このとき、発信元ノード102Aは、ルータIDコード及び接続情報を用いてネットワークポロジ情報を修正する。図9(g)は、第2のエコフレーム908Bを受信した後発信元ノード102Aが知っているネットワークポロジ情報を例示する。発信元ノード102Aは、20に等しいIDコードをもつ第1のルータ104Aに接続される。第1のルータ104Aは、ポート2及び5を介して未知のネットワークエレメントに接続され、ポート4及び6においては接続されず、ポート3を介して発信元ノード102Aに接続され、ポート5を介して第2のルータ104Dに接続されている。第2のルータ104Dは23に等しいIDコードを有し、ポート1, 2, 3, 4及び6で5つの未知のネットワークエレメントに接続され、ポート5を介して第1のルータ104Aに接続されている。発信元ノード102Aは、未知のネットワークエレメント914Aが未知のネットワークエレメント914Bと同じであることを理解していない。発信元ノードは、第1のルータ104Aのポート2からの経路及び第2のルータ104Dのポート4からの経路を探索しこれら2つの未知のエレメント914A, 914BのIDコードが同じであることを決定した後、これら2つのエレメントが同じであることを決定することになる。同様にして、第2のエコフレーム908Bを受信した後、受信元ノードは未知のエレメント914C及び914Dが同じであることを知らない。この決定は、発信元ノード102Aが上述のとおりループを特定した時点で発生することになる。

このプロセスは、全てのポートが検査されるまで続く。リンク又はエレメントが非作動状態になった場合、このリンク又はエレメントは、ピンフレームに回答してエコフレームが受信されないために特定されることになる。かかるリンク又はエレメントは、更新されたトポロジの中には内含されない。このとき、発信元ノード102Aは、更新されたネットワークポロジに基づいてルーティングテーブルを生成する(520)。更新されたルーティングテーブルは例えば非作動状態のリンクを避けるためにネットワークエレメント間に代替的経路を提供することができ、或いは又、全ての潜在的ループを回避しフレームトラフィックでリンク又はルータを過剰ロードする可能性を最小限におさえ、ネットワークデッドロックを回避するような最適な解決法を達成するべく、このルーティングテーブルをしらみつぶしに計算することが可能である。ルーティングテーブルを更新するための技術のいくつかの例は、本書にその全体が参考として内含されているR.D.RosnerのPacket Switching, Tomorrow's Communications Today(1982)の中で示されている。

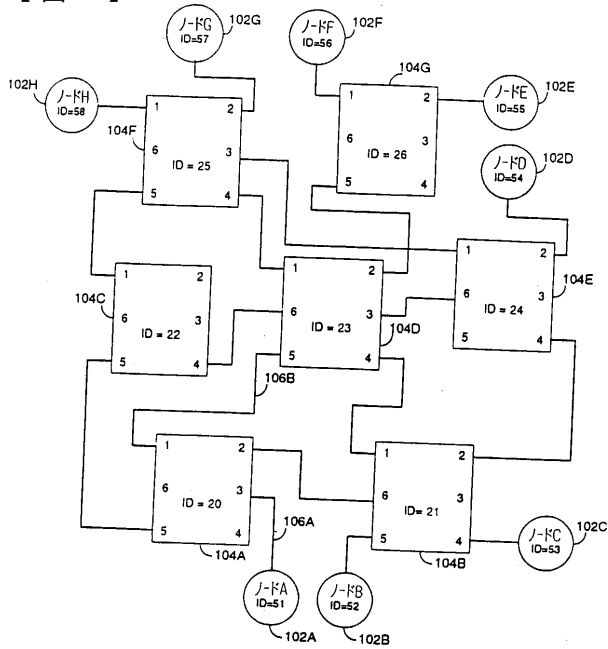
上述の例は、ルータによってピンフレームが受信されエコフレームが生成される場合を記述している。宛先エレメントが他のノードである場合でも、ピンフレームを受信しエコフレームを生成、送信するための技術は同じである。

10

20

30

【図 1】



100

Figure 1

【図 2】

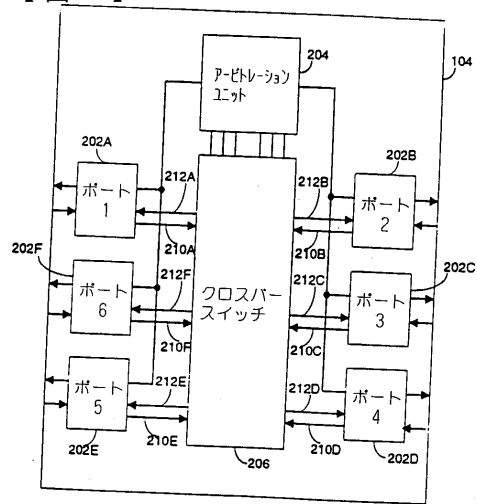


Figure 2

【図 3】

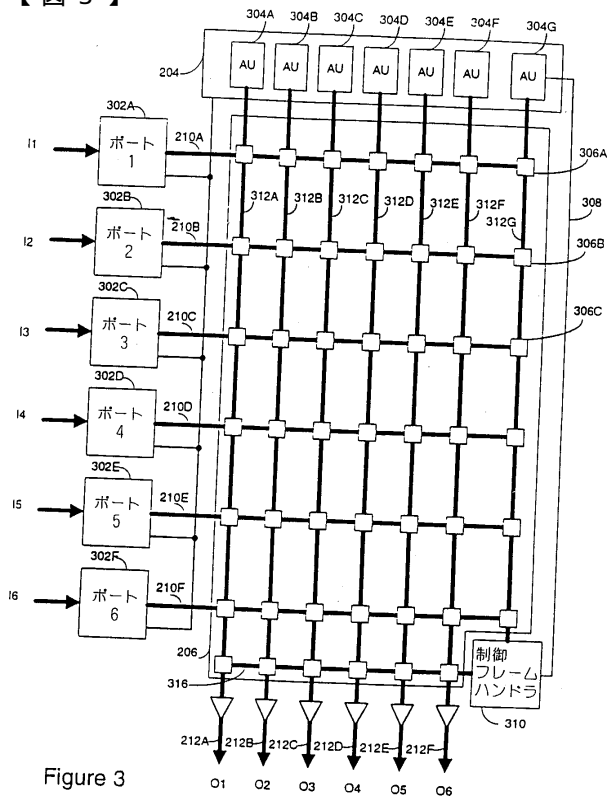


Figure 3

【図 4】

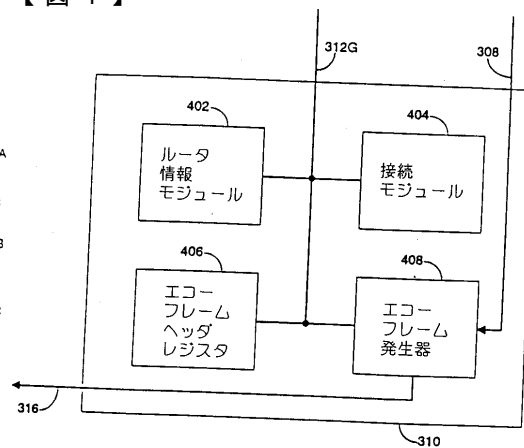


Figure 4

【図 5】

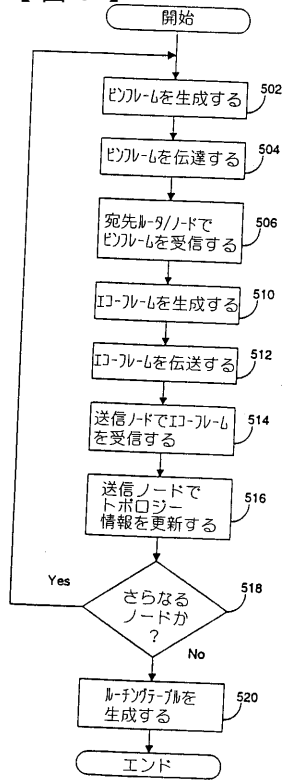
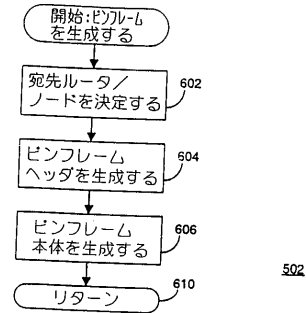


Figure 5

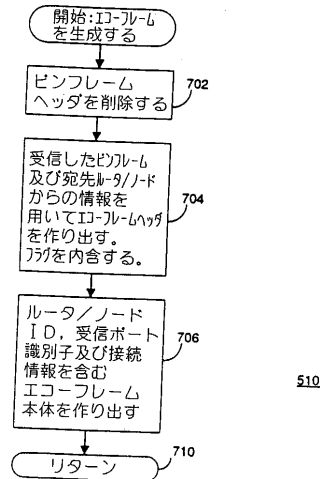
【図 6】

Figure 6



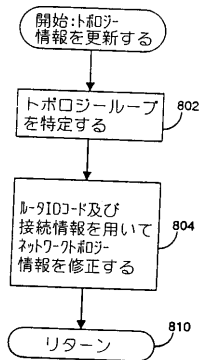
【図 7】

Figure 7



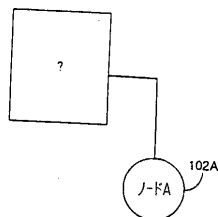
【図 8】

Figure 8



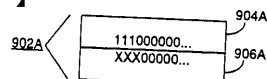
【図 9 (a)】

Figure 9(a)



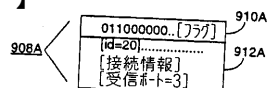
【図 9 (b)】

Figure 9(b)



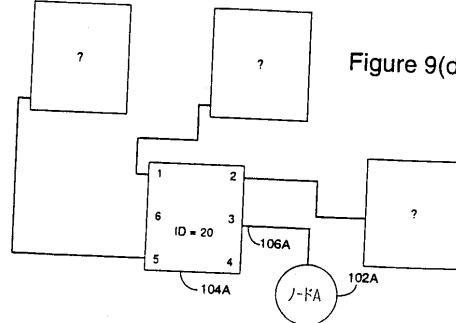
【図 9 (c)】

Figure 9(c)



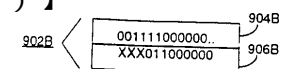
【図 9 (d)】

Figure 9(d)



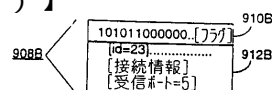
【図 9 (e)】

Figure 9(e)



【図 9 (f)】

Figure 9(f)



【図 9 (g)】

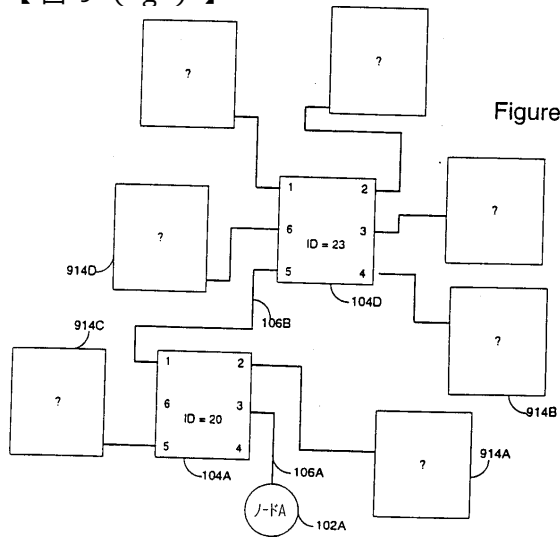


Figure 9(g)

【図 1 0】

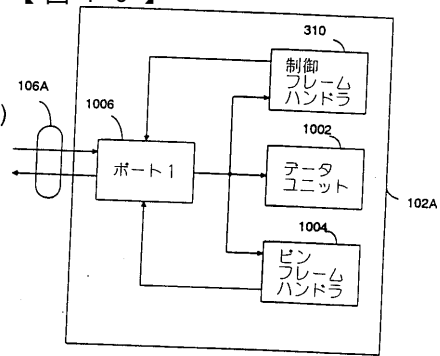


Figure 10

フロントページの続き

- (72)発明者 ウィッキー, トーマス エム.
アメリカ合衆国, カリフォルニア 94306, パロ アルト, ベンチュラ アベニュー 275,
アパートメント 7
- (72)発明者 ヘルランド, パトリック ジェイ.
アメリカ合衆国, ワシントン 98053, レドモンド, ノース イースト トウエンティーフィ
フス ウェイ 22401
- (72)発明者 ウェバー, ウォルフ-ディートリッヒ
アメリカ合衆国, カリフォルニア 94020, ラ ホンダ, ボックス 411, スター ルート
2
- (72)発明者 ウィッケ, ウィンフリード ダブリュ.
アメリカ合衆国, カリフォルニア 95120, サン ホセ, コッパー ピーク レーン 128
5

審査官 石井 研一

- (56)参考文献 特開平08-032597(JP, A)
米国特許第5138615(US, A)
米国特許第5485578(US, A)
特開平09-036873(JP, A)
特開平09-139736(JP, A)
特開平09-162870(JP, A)
特開平02-156753(JP, A)

(58)調査した分野(Int.Cl., DB名)

H04L 12/56
H04L 12/26
H04L 12/28
H04L 12/46
H04L 12/66