

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6576006号
(P6576006)

(45) 発行日 令和1年9月18日 (2019.9.18)

(24) 登録日 令和1年8月30日 (2019.8.30)

(51) Int. Cl.

F I

H04L 12/717 (2013.01)

H04L 12/717

請求項の数 20 (全 28 頁)

(21) 出願番号 特願2016-552529 (P2016-552529)
 (86) (22) 出願日 平成27年2月19日 (2015.2.19)
 (65) 公表番号 特表2017-506462 (P2017-506462A)
 (43) 公表日 平成29年3月2日 (2017.3.2)
 (86) 国際出願番号 PCT/US2015/016658
 (87) 国際公開番号 WO2015/127107
 (87) 国際公開日 平成27年8月27日 (2015.8.27)
 審査請求日 平成30年2月19日 (2018.2.19)
 (31) 優先権主張番号 14/188,027
 (32) 優先日 平成26年2月24日 (2014.2.24)
 (33) 優先権主張国・地域又は機関
 米国 (US)

(73) 特許権者 508140877
 レベル スリー コミュニケーションズ、
 エルエルシー
 アメリカ合衆国、コロラド州 80021
 , ブルームフィールド、エルドラド ブー
 ルバード 1025
 (74) 代理人 110000877
 龍華国際特許業務法人
 (72) 発明者 セラ、ウィリアム、トーマス
 アメリカ合衆国、コロラド州 80021
 , ブルームフィールド、エルドラド ブー
 ルバード 1025 レベル スリー コ
 ミュニケーションズ、エルエルシー内

最終頁に続く

(54) 【発明の名称】 分離した制御デバイスおよび転送デバイスを備えるネットワークでの制御デバイス検出

(57) 【特許請求の範囲】

【請求項 1】

複数の転送デバイスからルーティング制御デバイスへの複数の経路を検出する方法であって、

(a) 前記複数の転送デバイスのうち、転送デバイスのデータリンク層ポート上で、前記ルーティング制御デバイスからの制御パケットを受信する段階であって、前記データリンク層ポートは、前記転送デバイスを他の転送デバイスにリンクを介して接続し、前記制御パケットは、パケットを前記他の転送デバイスから前記ルーティング制御デバイスへとルーティングするための推定時間を示すパフォーマンスメトリックを含む、段階と、

(b) (ア) 前記制御パケットを受信した前記データリンク層ポート、および、(イ) 前記制御パケットにおける前記パフォーマンスメトリックに少なくとも部分的に基づいて、前記ルーティング制御デバイスに到達させるべく、前記転送デバイスの、MACアドレスによって識別されるどのデータリンク層ポート上でデータを転送するか識別する段階と

(c) 前記転送デバイスでの新規データストリームの受信に応答して、(b)で識別された前記データリンク層ポート上で、メッセージを前記ルーティング制御デバイスへ送信する段階であって、前記メッセージは前記ルーティング制御デバイスに対し、前記複数の転送デバイスを通る、宛先へのパスを決定することと、決定された前記パスに沿って前記新規データストリームを転送するように前記複数の転送デバイスのサブセットを構成することとを命令する、送信する段階と、

10

20

(d) 前記ルーティング制御デバイスの複数の構成コマンドに従って、前記新規データストリームを転送する段階と
を備える方法。

【請求項 2】

前記送信する段階 (c) は、前記ルーティング制御デバイスを通して、前記新規データストリームの初期パケットを転送する段階を含む、請求項 1 に記載の方法。

【請求項 3】

(e) 段階 (a) で前記制御パケットを受信した前記データリンク層ポートと異なる、前記転送デバイスの残りのデータリンク層ポート上で、前記制御パケットを前記転送デバイスから他の転送デバイスへと転送する段階を更に含む、請求項 1 または 2 に記載の方法。

10

【請求項 4】

(f) 前記転送デバイスで、前記複数の転送デバイスのうち、他の転送デバイスへとラベルを送信し、前記ラベルは、前記転送デバイスで受信されるパケットに付加されている場合に、前記転送デバイスに対し、(b) で識別された前記データリンク層ポート上で前記ルーティング制御デバイスへ前記パケットを転送するように指示する段階を更に含む、請求項 3 に記載の方法。

【請求項 5】

前記制御パケットがシリアル番号を含み、
前記方法は、

20

(f) 前記転送デバイスが、他の制御パケットを以前に転送したかどうかを判定する段階であって、前記他の制御パケットは、前記ルーティング制御デバイスからのものであり、前記シリアル番号を有し、パケットを前記転送デバイスから前記ルーティング制御デバイスへとルーティングする他の推定時間であって、前記制御パケット内の前記推定時間より少ない推定時間を有する、判定する段階を更に含む、

前記転送デバイスが、(f) において、前記他の制御パケットを以前に転送していないと判定された場合に、前記転送する段階 (e) の段階を行う、

請求項 3 または 4 に記載の方法。

【請求項 6】

(e) (b) で識別された前記データリンク層ポート上で他のメッセージを前記ルーティング制御デバイスへと送信する段階であって、前記他のメッセージは、前記転送デバイスと、前記複数の転送デバイスのうち、前記転送デバイスが接続されている複数の他の転送デバイスとを識別する、送信する段階

30

を更に含む、

前記ルーティング制御デバイスは前記他のメッセージを使用してネットワークのトポロジを表すデータベースを構築し、かつ、前記データベースを使用して前記パスを決定する、

請求項 1 から 5 のいずれか一項に記載の方法。

【請求項 7】

前記制御パケットは、前記ルーティング制御デバイスの識別子を含み、

40

(e) 前記転送デバイスのデータリンク層ポート上で、他のルーティング制御デバイスの識別子を含む他の制御パケットを受信する段階と、

(f) 前記他の制御パケットを受信した前記データリンク層ポートに少なくとも部分的に基づいて、前記転送デバイスのどのデータリンク層ポート上でデータを転送し、前記他のルーティング制御デバイスへ到達させるか識別する段階と、

(g) 新規データストリームの前記転送デバイスでの受信に応答して、前記パスが前記複数の転送デバイスを通して構成されるよう命令する前記メッセージを、前記ルーティング制御デバイスおよび前記他のルーティング制御デバイスのいずれが送信するか決定する段階と

を更に含む、

50

請求項 1 から 6 のいずれか一項に記載の方法。

【請求項 8】

ルーティング制御デバイスへの経路を検出するための転送デバイスであって、

(i) 前記転送デバイスのデータリンク層ポート上で前記ルーティング制御デバイスからの制御パケットを受信し、前記データリンク層ポートは、前記転送デバイスを他の転送デバイスにリンクを介して接続し、前記制御パケットは、パケットを前記他の転送デバイスから前記ルーティング制御デバイスへとルーティングするための推定時間を示すパフォーマンスメトリックを含み、(i i) (ア) 前記制御パケットを受信した前記データリンク層ポート、および、(イ) 前記制御パケットにおける前記パフォーマンスメトリックに少なくとも部分的に基づいて、前記ルーティング制御デバイスに到達させるべく、前記転送デバイスの、MAC アドレスによって識別されるどのデータリンク層ポート上でデータを転送するか識別するように構成されたスイッチ検出モジュールと、

新規データストリームの前記転送デバイスでの受信にตอบสนองして、識別された前記データリンク層ポート上で前記ルーティング制御デバイスへとメッセージを送信するように構成されたパス要求モジュールであって、前記メッセージは前記ルーティング制御デバイスに対し、(i) 複数の前記転送デバイスを通る、宛先へのパスを決定することと、(i i) 決定された前記パスに沿って前記新規データストリームを転送するように、複数の前記転送デバイスのサブセットを構成することとを命令する、パス要求モジュールと、

前記ルーティング制御デバイスの複数の構成コマンドに従って、前記転送デバイスが前記新規データストリームを転送することを可能にするべくデータ入力された、前記転送デバイス内のルーティングテーブルと

を備える転送デバイス。

【請求項 9】

前記パス要求モジュールは更に、前記ルーティング制御デバイスを通して前記新規データストリームの初期パケットを転送するよう構成される、請求項 8 に記載の転送デバイス。

【請求項 10】

前記スイッチ検出モジュールは、前記制御パケットを受信した前記データリンク層ポートとは異なる前記転送デバイスの残りのデータリンク層ポート上で、前記転送デバイスから他の転送デバイスへと前記制御パケットを転送するよう構成され、任意で、

前記スイッチ検出モジュールは、前記複数の転送デバイスの中の、他の転送デバイスへとラベルを送信するよう構成され、

前記ラベルは、前記転送デバイスで受信されたパケットに付加されている場合に、前記転送デバイスに対し、前記スイッチ検出モジュールによって識別された前記データリンク層ポート上で、前記ルーティング制御デバイスへ前記パケットを転送するように指示する、

請求項 8 または 9 に記載の転送デバイス。

【請求項 11】

前記制御パケットはシリアル番号を含み、前記転送デバイスは更に、パフォーマンスメトリックモジュールを備え、

前記パフォーマンスメトリックモジュールは、前記転送デバイスが、他の制御パケットを以前に転送したかどうか判定するよう構成され、前記他の制御パケットは、前記ルーティング制御デバイスからのものであり、前記シリアル番号を含み、前記転送デバイスから前記ルーティング制御デバイスへとパケットをルーティングする他のパフォーマンスメトリックであって、前記制御パケット内のパフォーマンスメトリックより良い他のパフォーマンスメトリックを含み、

前記パス要求モジュールが、前記他の制御パケットを以前に転送しなかったと判定した場合に、前記転送を行う、

請求項 10 に記載の転送デバイス。

【請求項 12】

前記スイッチ検出モジュールは更に、

識別された前記データリンク層ポート上で前記ルーティング制御デバイスへ他のメッセージを送信するよう構成され、前記他のメッセージは、前記転送デバイスと、前記転送デバイスが接続されている前記複数の転送デバイスのうち複数の他の転送デバイスとを識別し、

前記ルーティング制御デバイスは、ネットワークのトポロジを表すデータベースを構築するべく前記他のメッセージを使用し、前記パスを決定するべく前記データベースを使用するよう構成される、

請求項 10 または 11 に記載の転送デバイス。

【請求項 13】

前記制御パケットは、前記ルーティング制御デバイスの識別子を含み、前記スイッチ検出モジュールは更に、

前記転送デバイスのデータリンク層ポート上で、他のルーティング制御デバイスの識別子を含む、他の制御パケットを受信し、

前記他の制御パケットを受信した前記データリンク層ポートに少なくとも部分的に基づいて、前記他のルーティング制御デバイスに到達するべく、前記転送デバイスのどのデータリンク層ポート上でデータを転送するか識別し、

新規データストリームの前記転送デバイス上での受信にตอบสนองして、前記パスが前記複数の転送デバイスを通して構成されるように命令する前記メッセージを、前記ルーティング制御デバイス、および前記他のルーティング制御デバイスのいずれに送信するか決定するよう構成される

請求項 10 から 12 のいずれか一項に記載の転送デバイス。

【請求項 14】

複数の転送デバイスからルーティング制御デバイスへの経路を検出するための方法を少なくとも1つのマシンに実行させるプログラムであって、前記方法は、

(a) 前記複数の転送デバイスのうち、転送デバイスのデータリンク層ポート上で、前記ルーティング制御デバイスからの制御パケットを受信する段階であって、前記データリンク層ポートは、前記転送デバイスを他の転送デバイスにリンクを介して接続し、前記制御パケットは、パケットを前記他の転送デバイスから前記ルーティング制御デバイスへとルーティングするための推定時間を示すパフォーマンスメトリックを含む、段階と、

(b) (ア) 前記制御パケットを受信した前記データリンク層ポート、および、(イ) 前記制御パケットにおける前記パフォーマンスメトリックに少なくとも部分的に基づいて、前記ルーティング制御デバイスへ到達させるべく、前記転送デバイスの、MACアドレスによって識別されるどのデータリンク層ポート上でデータを転送するか識別する段階と、

(c) 新規データストリームの前記転送デバイスでの受信にตอบสนองして、(b)で識別された前記データリンク層ポート上で、メッセージを前記ルーティング制御デバイスへ送信する段階であって、前記メッセージは前記ルーティング制御デバイスに対し、前記複数の転送デバイスを通る、宛先へのパスを決定することと、決定された前記パスに沿って前記新規データストリームを転送するように前記複数の転送デバイスのサブセットを構成することとを命令する、送信する段階と、

(d) 前記ルーティング制御デバイスの複数の構成コマンドに従って、前記新規データストリームを転送する段階と

を備える、プログラム。

【請求項 15】

前記送信する段階(c)は、前記新規データストリームの初期パケットを前記ルーティング制御デバイスを通して転送する段階を含む、請求項 14 に記載のプログラム。

【請求項 16】

前記方法は、(e) (a)で前記制御パケットが受信された前記データリンク層ポートと異なる、前記転送デバイスの残りのデータリンク層ポート上で、前記制御パケットを前

10

20

30

40

50

記転送デバイスから他の転送デバイスへと転送する段階を更に含み、任意で、

前記方法は、

(f) 前記転送デバイスで、前記複数の転送デバイス中の他の転送デバイスにラベルを送信する段階を更に備え、前記ラベルは、前記転送デバイスで受信されたパケットに付加されている場合に、前記転送デバイスに対し、(b)で識別された前記データリンク層ポート上で、前記ルーティング制御デバイスへ前記パケットを転送するよう指示する、請求項14または15に記載のプログラム。

【請求項17】

前記制御パケットはシリアル番号を含み、前記方法は、

(f) 前記転送デバイスが、以前に転送された他の制御パケットを有するかどうか判定する段階であって、前記他の制御パケットは、前記ルーティング制御デバイスからのものであり、前記シリアル番号を有し、パケットを前記転送デバイスから前記ルーティング制御デバイスへとルーティングする他の推定時間であって、前記制御パケット内の前記推定時間より少ない推定時間を有する、判定する段階

を更に備え、

前記転送デバイスが、(f)において、前記他の制御パケットを以前に転送していないと判定された場合に、前記転送する段階(e)の転送段階を行う、

請求項16に記載のプログラム。

【請求項18】

前記方法は、(e)(b)で識別された前記データリンク層ポート上で前記ルーティング制御デバイスへ他のメッセージを送信する段階であって、前記他のメッセージは、前記転送デバイスと、前記転送デバイスが接続されている、前記複数の転送デバイスのうち複数の他の転送デバイスとを識別する段階を更に備え、

前記ルーティング制御デバイスは、ネットワークのトポロジを表すデータベースを構築するべく前記他のメッセージを使用し、前記パスを決定するべく前記データベースを使用する、

請求項14から17のいずれか一項に記載のプログラム。

【請求項19】

前記制御パケットは、前記ルーティング制御デバイスの識別子を含み、

前記方法は、

(e) 前記転送デバイスのデータリンク層ポート上で、他のルーティング制御デバイスの識別子を含む他の制御パケットを受信する段階と、

(f) 前記転送デバイスの受信した前記データリンク層ポートに少なくとも部分的に基づいて、前記他のルーティング制御デバイスへ到達するべく、前記転送デバイスのどのデータリンク層ポート上でデータを転送するか識別する段階と、

(g) 新規データストリームの前記転送デバイスでの受信にตอบสนองして、前記複数の転送デバイスを通る前記パスが構成されるように命令する前記メッセージを、前記ルーティング制御デバイス、および、前記他のルーティング制御デバイスのいずれに送信するか決定する段階と

を更に含む、請求項14から18のいずれか一項に記載のプログラム。

【請求項20】

複数の転送デバイスのネットワークにおいて、少なくとも、転送デバイスのデータリンク層ポートに接続されている、ルーティング制御デバイスであって、

前記ルーティング制御デバイスからの制御パケットをデータリンク層ポートから送信することによって、前記複数の転送デバイスの前記ネットワークのトポロジを判定するよう構成されたコントローラ検出モジュールであって、前記制御パケットは、パケットを各転送デバイスから前記ルーティング制御デバイスへとルーティングするための推定時間を示すパフォーマンスメトリックを含む、コントローラ検出モジュールと、

前記ネットワークのトポロジを保存するよう構成されたトポロジデータベースと、

前記ルーティング制御デバイスから、前記ネットワーク内の各転送デバイスへのパスを

10

20

30

40

50

確立し、

前記ネットワーク内の前記転送デバイスのそれぞれから前記ルーティング制御デバイスへのパスを確立し、

転送元デバイスからの接続要求をデータリンク層ポート上で受けると、前記転送元デバイスから転送先デバイスへの接続パスを確立する

よう構成されたパス決定モジュールと、

前記接続パスが確立されている間、前記転送元デバイスから前記ルーティング制御デバイス、および、前記ルーティング制御デバイスから前記転送先デバイスまでの前記パスを通して、前記接続要求に対して少なくとも1つの初期パケットを送るよう構成されたデータ転送モジュールと

10

を備える、ルーティング制御デバイス。

【発明の詳細な説明】

【技術分野】

【0001】

本分野は概して、ネットワークのルーティングに関する。

【背景技術】

【0002】

通信ネットワークは、例えば、2つの地理的に離れた場所の間でデータを転送することを可能にするネットワーク接続を提供し得る。ネットワーク接続は、ルータなどの通信デバイスを接続する、複数のリンクに及び得る。通信デバイスを通してリンクがどのように相互接続されるかに応じて、ネットワークは異なるトポロジを有し得る。特定のネットワークトポロジにおいて、送信元と宛先との間で、複数の経路が利用可能な場合がある。現在の容量および利用状況に応じて、いくつかの経路は、他の経路より好ましいことがある得る。

20

【0003】

従来のルーティングのアルゴリズムは、各ルータが有する、その隣接するリンクおよびデバイスからのローカル情報を利用して、データをルーティングし得る。ルータはそのような情報をルーティングテーブルに保持する。そして、ルータは受信パケットの宛先アドレスに基づいて、そのルーティングテーブルを使用して、受信パケットを特定の隣接するデバイスへ転送し得る。ルーティングテーブルを構築するべく、各ルータは、ボーダゲートプロトコル(BGP)などのプロトコルを使用して、ローカルの隣接するルータと、ルーティングおよび到達可能性の情報を交換する。このように、各ルータは、パケットを転送し、かつ、制御機能を実行して自身のルーティングテーブルを更新する。

30

【0004】

ローカル情報の使用は、いくつかの状況では望ましいが、常に効率的にデータをルーティングできるとは限らない。より効率的にデータをルーティングするべく、ソフトウェア定義ネットワーク(SDN)と呼ばれる他の技術は、制御機能と転送機能とを別個のデバイスに分離する。制御デバイスは、ネットワークトポロジの全体的な知識を使用して、個別のデータフローに対し、転送デバイスのネットワークを通るパスを決定する。このように、ルーティング制御デバイスは、例えば、遅延を最小化するか帯域幅を最大化するパスを、ネットワークを通して確立し得る。

40

【発明の概要】

【0005】

一実施形態において、コンピュータ実装方法は、ネットワーク転送デバイスからルーティング制御デバイスへの経路を検出する。この検出方法は、転送デバイスのデータリンク層ポート上でルーティング制御デバイスからの制御パケットを受信する段階であって、データリンク層ポートは、転送デバイスを他の転送デバイスにリンクを介して接続し、制御パケットは、パケットを他の転送デバイスからルーティング制御デバイスへとルーティングするための推定時間を示すパフォーマンスメトリックを含む、段階を含む。この方法では次に、少なくとも部分的には、(ア)制御パケットを受信したデータリンク層ポート、

50

および、(イ)制御パケットにおけるパフォーマンスメトリックに基づいて、ルーティング制御デバイスに到達するべく、転送デバイス上の、MACアドレスによって識別されるデータのデータリンク層ポートでデータを転送するか識別する。この方法では、新規データストリームの転送デバイスでの受信に応答して、識別されたデータリンク層ポート上で、コマンドをルーティング制御デバイスへ送信する。このコマンドは、ルーティング制御デバイスに対し、転送デバイスを通る、宛先までのパスを決定することと、決定されたパスに沿って新規データストリームを転送するように転送デバイスを構成することとを命令する。この方法では次に、ルーティング制御デバイスの構成コマンドに従って、新規データストリームを転送する。

【0006】

10

システムおよびコンピュータプログラム製品の実施形態も開示される。

【0007】

本発明の更なる実施形態、特徴、および利点、ならびに、各種実施形態の構造及び動作が、複数の添付図面を参照して以下に詳細に説明される。

【図面の簡単な説明】

【0008】

本明細書に組み込まれ、本明細書の一部を形成する複数の添付図面は本開示を図示し、説明と合わせて、本開示の複数の原理を説明し、当業者が本開示を作成及び使用することを可能にするのに更に役立つ。

【0009】

20

【図1A】制御パケットを転送デバイスへ送信するルーティング制御デバイスを図示するダイアグラムである。

【0010】

【図1B】ネットワーク全体にわたって、制御パケットをフラディングする、転送デバイスを図示するダイアグラムである。

【0011】

【図1C】ネットワークトポロジを反映した情報を制御デバイスが収集する方法を図示するダイアグラムである。

【0012】

【図2】制御デバイスから転送デバイスへのパスを見つけるための方法を図示するフローチャートである。

30

【0013】

【図3】各転送デバイスから制御デバイスへのパスをセットアップする方法を図示するフローチャートである。

【0014】

【図4A】転送デバイスから制御デバイスへのパスを使用するネットワーク接続を図示するダイアグラムである。

【0015】

【図4B】高速パスをセットアップするルーティング制御デバイスを図示するダイアグラムである。

40

【0016】

【図5】複数の制御デバイスを有するネットワークを図示するダイアグラムである。

【0017】

【図6】制御デバイスおよび転送デバイスのモジュールを図示するダイアグラムである。

【0018】

要素が最初に現れる図面は、概して、対応する参照番号で左端の1または複数の数字で示される。複数の図面において、同様の参照番号は、同一または機能的に類似の要素を示し得る。

【発明を実施するための形態】

【0019】

50

上述のように、SDNルーティング技術は、ネットワークトポロジの全体的な知識を使用することで、効率的にパケットをルーティングする。これらの技術は、互いに分離したルーティング制御デバイスおよび転送デバイスを使用する。転送デバイスがネットワークを通して新規データフローを受信すると、転送デバイスは、そのデータフローの新規パスを決定するべく、ルーティング制御デバイスに連絡する。ルーティング制御デバイスは、ネットワークの転送デバイスを通る新規パスを決定し、それに応じて、転送デバイスを構成する。

【0020】

機能するには、ルーティング制御デバイスおよび転送デバイスは、ネットワーク情報で構成される必要がある。具体的には、相互接続された転送デバイスを通る新規パスを作成するように要求するには、転送デバイスは、ルーティング制御デバイスにデータをルーティングする方法を知っている必要がある。そして、ルーティング制御デバイスは、ネットワークを通るパスを決定するには、ネットワークトポロジを知っている必要がある。この情報を手動で構成することは、長い時間を要し、エラーが発生し得る。

10

【0021】

本明細書で開示される複数の実施形態は、トポロジ情報が自動的に検出されることを可能にする。一実施形態において、ルーティング制御デバイスは、転送デバイスに接続され、それに制御パケットを送信する。制御パケットは、制御デバイスのアドレスを、その送信元アドレスとして、アドレスフィールド内に含む。また、制御デバイスは、シリアル番号を各制御パケットに付加する。

20

【0022】

転送デバイスは、例えば、ネットワークが当該転送デバイスからルーティング制御デバイスへと情報を送信するのにかかり得る時間を示すパフォーマンス情報を付加する。次に、転送デバイスは、他の転送デバイスに接続されている、その他のポート上で、制御パケットをブロードキャスト、またはフラディングする。そして、他の転送デバイスが、より良いパフォーマンス情報を含むパケットを以前に再ブロードキャストしていない場合、当該他の転送デバイスは、パフォーマンス情報を更新し、パケットを再ブロードキャストする。

【0023】

転送デバイスは、制御パケットを受信すると、最高のパフォーマンス情報を有するパケットを制御パケットがどのポートで受信したかを記憶する。転送デバイスは、このポートが、ルーティング制御デバイスへの最速経路上であることが分かる。そして、各転送デバイスは、ラベルスイッチパスをその経路上で設定するように、ルーティング制御デバイスに要求し得る。

30

【0024】

パスが確立されると、転送デバイスは、近隣のネットワークトポロジと、ネットワーク状態とに関する情報を、制御デバイスに送信する。制御デバイスは、すべての転送デバイスからの、そのような情報を使用することで、全体的なネットワークトポロジと、ネットワーク状態とのデータベースを作成できる。このトポロジデータベースによって、制御デバイスは、ネットワーク内で転送元デバイスと転送先デバイスとの間のパスを確立することが可能になる。

40

【0025】

コントローラに対してこのパスが確立されると、コントローラは、それ以降いつでも、より完全なトポロジおよびパフォーマンス情報を利用し得る他の方法を使用することで、転送デバイスとコントローラとの間の通信により良いパスを決定し得る。この、より良いパスが決定された場合、既存のパスは、コントローラによって置き換えられ得る。この場合、コントローラは、特定の新規パスを再度シグナリングし、次に、既存のパスを削除する。

【0026】

後述の詳細な説明は、5つの項に分けられる。第1項は、図1A 図1Cに関連し、制

50

御パケットを各転送デバイスへ送信し、トポロジおよびネットワーク状態情報を転送デバイスから収集する、ネットワーク内の制御デバイスを説明する。第2項は、図2 図3に関連し、制御デバイスと各転送デバイスとの間の双方向接続のセットアップを説明する。第3項は、図4 A 図4 Bに関連し、大量のパケットをバッファリングするようにエッジルータへ要求することなく、制御デバイスを使用することで、データを確立し、転送することを説明する。第4項は、図5に関連し、ネットワーク内における複数の制御デバイスの使用を説明する。最後の第5項は、図6に関連し、制御デバイスシステム、および、そのモジュール、ならびに、転送デバイスシステム、および、そのモジュールを説明する。

【制御デバイスおよび制御パケット】

【0027】

10

図1 Aは、通信ネットワークを図示するダイアグラム100である。通信ネットワークは、ローカルエリアネットワーク(LAN)、メトロポリタンエリアネットワーク(MAN)、または、広域ネットワーク(WAN)であり得る。通信ネットワークは、ポイントツーポイント、または、マルチポイントツーマルチポイントのネットワークプロトコルを利用し得る。使用されるネットワークアクセスプロトコルは、例えば、マルチプロトコルラベルスイッチング(MPLS)、イーサネット(登録商標)、非同期転送モード(ATM)、ハイレベルデータリンク制御(HDLC)、または、パケットリレーを含み得る。

【0028】

通信ネットワークは、転送デバイス104、106、108、110および112など、リンクによって相互接続された、複数の転送デバイスを含む。転送デバイスは、パケットを転送するデバイスであり、データリンク層(OSI第2層)、および、ネットワーク層(OSI第3層)のデバイスを含む。

20

【0029】

また、通信ネットワークは、ルーティング制御デバイス102を含む。ルーティング制御デバイス102は、例えば転送デバイス104など、少なくとも1つの転送デバイスに接続され得る。ルーティング制御デバイス102は、ネットワーク内の他の転送デバイスから、地理的に遠く離れた場所にあり得る。

【0030】

例として、ユーザは、例えば転送デバイス106などの転送元デバイスから、例えば転送デバイス110などの転送先デバイスへとデータを送信し得る。データは、パケットに分割されるストリームであり得、各パケットは、転送デバイス110、または、他の下流のデバイスを、その宛先として指定し得る。

30

【0031】

ルーティング制御デバイス102は、ネットワーク接続を確立するためのインテリジェントなルーティングを提供する。これを行うには、制御デバイス102は、トポロジ、ならびに、ネットワークリンクおよびネットワークデバイスの状態に関する知識が必要である。ルーティング制御デバイス102は、転送デバイス104に直接接続する。制御デバイスがネットワークの知識を獲得するべく、制御デバイス102は、制御パケット120を転送デバイス104へ送信する。

【0032】

40

制御パケット120は、当該パケットが制御パケットであることを示すものと、メディアアクセス制御アドレスなどのルーティング制御デバイス102の識別子と、パケットのシリアル番号とを含み得る。一実施形態において、制御デバイス102は、一定の時間間隔で、新規制御パケットを送信し得る。制御デバイス102が新規パケットを送信するたびに、制御デバイス102は、シリアル番号をインクリメントし、異なる番号を有する新規制御パケットを送り出す。

【0033】

説明のための例において、制御パケット120は、以下の情報を含み得る。

タイプフラグ：C(制御パケット)

制御デバイスのMACアドレス：01:23:45:67:89:ab

50

制御パケットのシリアル番号：1

【0034】

図1Bで図示するように、ルーティング制御デバイス102が制御パケット120を隣の転送デバイス104へ送信すると、ネットワーク転送デバイスは、そのパケットをネットワーク全体にフラディングする。

【0035】

図1Bは、制御パケットをネットワーク全体にフラディングする、ネットワーク転送デバイスを図示するダイアグラムである。制御パケット120を受信後、転送デバイス104は、制御パケットを受信したポートを除く、そのすべてのポートで、制御パケットを送り出し得る。

10

【0036】

上述した通り、ルーティング制御デバイス102が制御パケット120を生成する場合、制御デバイスは、そのアドレスとシリアル番号とを制御パケットに含め得る。転送デバイス104が制御パケットを受信すると、それを転送する前に、転送デバイス104は、2つの修正を行い得る。第1に、転送デバイスは、それ自身の識別子（そのMACアドレスなど）を制御パケットに追加し得る。このアドレスは、制御パケットが通過するすべての転送デバイスを記録するように、制御パケットのアドレスフィールドに追加される。第2に、転送デバイス104は、制御パケット内のパフォーマンスメトリックを修正し得る。パフォーマンスメトリックには、例えば、制御パケットを制御デバイス102から転送デバイス104へ送信する際の遅延が含まれ得る。この遅延情報は、例えば、リンク層検出プロトコル（LLDP）のパケット交換を使用して、転送デバイス104によって収集され得る。トラフィック量の変化にネットワークが反応する際の一時的な変化を防止するべく、遅延は移動平均として決定されるか、さもなければ平準化され得る。

20

【0037】

転送デバイス104は、これら2つの修正を加え、制御パケット142、144および146として、他のポートで制御パケットを転送する。この時点では、3つの制御パケット142、144および146はすべて、同じ情報を含み得る。引き続き上の例で、転送デバイス104のMACアドレスが02：02：02：02：02：bcであり、制御デバイス102と転送デバイス104との間の遅延が5ナノ秒であると仮定すると、制御パケット142、144および146は、以下の情報を含み得る。

30

タイプフラグ：C（制御パケット）

制御デバイスのMACアドレス：01：23：45：67：89：ab

制御パケットのシリアル番号：1

仲介転送デバイスのアドレス：02：02：02：02：02：bc

パフォーマンスメトリック：5ナノ秒

【0038】

転送デバイス104は、制御パケット144をリンク114上で転送デバイス106に転送し、制御パケット142をリンク116上で転送デバイス108に転送し、制御パケット146をリンク118上で転送デバイス110に転送する。

40

【0039】

転送デバイス106は、制御パケット144を受信すると、自身のアドレスをアドレスフィールドに追加し、制御パケットのパフォーマンスメトリックを更新することで、制御パケット148を生成する。例えば、転送デバイス106は、遅延を追加し、データを転送デバイス106から転送デバイス104へ送信し得る。引き続き上の例で、転送デバイス106のMACアドレスが03：03：03：03：03：cdであり、コントローラ102と転送デバイス104との間の遅延が2ナノ秒であると仮定すると、制御パケット148は、以下の情報を含み得る。

タイプフラグ：C（制御パケット）

制御デバイスのMACアドレス：01：23：45：67：89：ab

制御パケットのシリアル番号：1

50

仲介転送デバイスのアドレス：02：02：02：02：02：bc、03：03：03：03：cd

パフォーマンスメトリック：7ナノ秒

【0040】

転送デバイス106は、パケット148をリンク122上で転送デバイス108へ送信する。

【0041】

転送デバイス108は、制御パケット142および148を受信する。転送デバイス108は、これら2つの制御パケットの各々に含まれる遅延情報を使用することで、制御デバイス102への遅延がより少ないパスを確立することを選択できる。転送デバイス108は、制御パケット内のアドレス系列フィールドを使用することで、制御デバイスに到達するには、どの隣接する転送デバイスと通信する必要があるか知ることができる。

【0042】

例えば、転送デバイス108は、リンク122上でデータを転送デバイス106へ送信するための時間は3ナノ秒であり、リンク116上でデータを転送デバイス104へ送信するための時間は9ナノ秒であることを知り得る。この時間を、受信したパケット148および142内のパフォーマンスメトリックに追加することで、転送デバイス108は、転送デバイス106を介した制御デバイス102への伝送には10ナノ秒かかるのに対し、転送デバイス104を介した伝送には、14ナノ秒かかるかと判定できる。そのため、転送デバイス108は、デバイス104ではなく、転送デバイス106を介して、制御デバイス102へデータをルーティングすることを選択し、より速いパスを取り得る。

【0043】

また、転送デバイス108は、制御パケット142および148を再転送し得る。転送デバイス108が、より遅いパスを示すパケット（この場合、パケット142）を受信する前に、より速いパスを示すパケット（この場合、パケット148）を受信した場合、転送デバイス108は、より遅いパケットを転送しなくてよい。なぜなら、転送デバイス108は、制御デバイス102へと向かう如何なるデータも、そのデータが転送デバイス108に由来するものか、または、デバイス112などの他の転送デバイスに由来するものに関係なく、転送デバイス106を通る、より速いパスを取るべきであることを知っているからである。

【0044】

制御デバイス102への最適パスは転送デバイス106を通るパスであることを決定すると、転送デバイス108は、転送デバイス106との双方向パスを確立し得る。一実施形態において、この双方向パスを確立するべく、転送デバイス108は、制御デバイスへ到達するには、特定のポート（当該転送デバイスが制御パケット144を受信したポート）で特定のラベル付きのパケットを送信するというルールを、転送テーブルとも呼ばれ得る、そのルーティングテーブルに設定する。この情報は、転送デバイス106に通信され、その結果、両方の転送デバイスが同じルールに従う。したがって、転送デバイス106が、転送デバイス108から、設定ラベルとともにパケットを設定ポートで受信した場合、転送デバイス106は、このパケットは制御デバイス102宛てであることを知る。同様に、転送デバイス106は、制御デバイス102宛てのトラフィックを、転送デバイス104を通して転送することを知る。そのため、転送デバイス106は、制御デバイス102へとデータをルーティングするべく、転送デバイス104と共にラベルを設定する。また、制御デバイス102から転送デバイス104を通して転送デバイス106へとデータをルーティングするべく、同様の方法でラベルが設定され得る。

【0045】

このように、制御デバイスからの制御パケットを分析し、再転送することによって、各転送デバイスは、制御デバイスと通信する方法を検出し、制御デバイスへのラベルスイッチパスを設定できる。図4Aおよび図4Bに関して後述するように、転送デバイスが新規データストリームを受信すると、制御デバイスがネットワークを通るパケットのためのパ

10

20

30

40

50

スを設定するまで、転送デバイスは新規データストリームからのパケットを制御デバイスへ転送し得る。

【 0 0 4 6 】

実施形態では、ラベルスイッチパスを使用してデータを制御デバイスへ転送するだけでなく、図 1 C で図示されているように、ラベルスイッチパスを使用して、制御デバイスがネットワークのトポロジ情報を検出することも助ける。

【 0 0 4 7 】

図 1 C は、ネットワークトポロジを反映している情報を収集する制御デバイス 1 0 2 を図示するダイアグラム 1 7 0 である。

【 0 0 4 8 】

各転送デバイスは、何の転送デバイスに接続されているか、および、周囲のリンクおよび転送デバイスの状態はどうかを示す情報をルーティング制御デバイスへ送信する。周囲のリンクおよび転送デバイスの状態には、例えば、トラフィックおよび輻輳情報、レイテンシ、パケット損失率、ジッタなどが含まれ得る。この情報は、例えば、リンク層検出プロトコル (L L D P) を使用して収集され得る。一実施形態において、転送デバイスは、この情報を定期的に送信する。そして、転送デバイスは、転送デバイスと制御デバイスとの間で確立されたパスを使用して、この情報を送信する。

【 0 0 4 9 】

図 1 C で示されている例では、制御デバイス 1 0 2 への転送デバイス 1 0 8 用のパスが確立されている。このパスには、転送デバイス 1 0 8 、 1 0 6 および 1 0 4 が含まれる。転送デバイス 1 0 8 は、その周囲のトポロジを含む情報を制御デバイス 1 0 2 へ送信する。例では、転送デバイス 1 0 8 は、この情報をパケット 1 7 2 に入れてルーティング制御デバイス 1 0 2 へと送信する。

【 0 0 5 0 】

ルーティング制御デバイス 1 0 2 は、転送デバイス 1 0 8 を含む転送デバイスのすべてから受信するトポロジ情報を使用して、転送デバイスのネットワークのトポロジを反映するデータベース 1 7 4 を作成する。 [制御デバイスとの間で両方向のパスを確立する]

【 0 0 5 1 】

図 2 は、制御デバイスから転送デバイスへのパスを発見するための方法 2 0 0 を図示するフローチャートである。

【 0 0 5 2 】

段階 2 0 2 で、制御デバイスは転送デバイスのネットワークのうち 1 つの転送デバイスに物理的に接続する。段階 2 0 4 で、制御デバイスは、制御 (または検出) パケット、例えば、リンク層検出プロトコル (L L D P) パケットを、制御デバイスが接続されている転送デバイスへ送信する。

【 0 0 5 3 】

段階 2 0 6 で、パケットを受信した転送デバイスは、パケットが制御デバイスに由来するものかどうか検証する。例において、転送デバイスは、例えば、特定のフラグが立てられているかどうかをチェックすることによって、パケットが制御デバイスからのものであることを検証できる。

【 0 0 5 4 】

制御パケットは、ネットワークリンクまたはネットワークデバイスの状態を示すパフォーマンスメトリックを含む。例えば、メトリックはリンクのレイテンシを示し得る。したがって、この例では、メトリックが低いことは、レイテンシがより低いこと、したがって、パフォーマンスがより良いことを示す。また、メトリックは、レイテンシ、パケット損失、ジッタなどの複数の状態の関数であり得る。

【 0 0 5 5 】

制御パケットを受信する各転送デバイスは、パフォーマンスメトリックを更新する。更新されたメトリックは、例えば、制御パケットが前の、すなわち上流の転送デバイスから、当該転送デバイスへ到達するための追加のレイテンシを示す。メトリックの更新後、転

10

20

30

40

50

送デバイスは、制御パケットを受信したポートを除く、そのすべてのポート上で、制御パケットをフラッディングする。

【 0 0 5 6 】

転送デバイスは、過去に受信したものと同一制御パケットを再度受信し得る。これが起きるのは、制御パケットがネットワークを通してフラッディングされるからである。転送デバイスは、制御デバイスが各制御パケットに付加する固有のシリアル番号によって、繰り返しの制御パケットを認識できる。

【 0 0 5 7 】

段階 2 0 8 で、転送デバイスは、パケットのパフォーマンスメトリックが、制御デバイスから受信した以前のメトリックより良いかどうかをチェックする。転送デバイスが受信したすべての制御パケットの中で、当該メトリックが最高のパフォーマンスメトリックである場合、段階 2 1 0 で、転送デバイスは、制御デバイスへ到達するための最適な経路としてパケットの受信ポートを保存する。

10

【 0 0 5 8 】

段階 2 1 2 で、制御パケットがネットワークを通して無限に循環することを回避するべく、転送デバイスはシリアル番号をチェックして、制御パケットが以前に転送されたかどうか判定する。制御パケットは固有のシリアル番号を、生成する各制御パケットへ追加するので、転送デバイスは制御パケットのシリアル番号を使用して、制御パケットを以前に転送したかどうか判定できる。または、転送デバイスが、例えば、より大きいシリアル番号を有する、より最近の制御パケットを以前に転送した場合、転送デバイスはシリアル番号を使用して、現在の制御パケットが古いかどうかを判定する。パケットが無限にネットワークを通して循環することを回避するべく、制御パケットの少なくとも一部のハッシュまたはチェックサムなど、多くの他の機構が使用され得る。

20

【 0 0 5 9 】

段階 2 1 4 で、転送デバイスは、パケットが以前に転送されたかどうか、および、最高のパフォーマンスを示すメトリックを含んでいるかどうかを判定する。制御パケットが最高のパフォーマンスメトリックを示さないか、または、当該転送デバイスによって以前に転送されていた場合、転送デバイスはパケットを破棄し、プロセスは終了する。

【 0 0 6 0 】

制御パケットが最高のパフォーマンスメトリックを含み、かつ、以前に転送されていない場合、プロセス 2 0 0 は段階 2 1 6 でパスを設定する。段階 2 1 6 は、後述の図 3 の説明で詳細に説明される。

30

【 0 0 6 1 】

次に段階 2 1 8 で、転送デバイスは、上流の転送デバイスと現在の転送デバイスとを接続しているリンクのパフォーマンス、および、2つの転送デバイス自体のパフォーマンスを反映するように、パフォーマンスメトリックを更新する。次に段階 2 2 0 で、受信した転送デバイスは、制御パケットを受信したポートを除く、そのすべてのポート上でパケットをフラッディングする。このように、更新された制御パケットは、他の転送デバイスへ送信される。

【 0 0 6 2 】

40

他の転送デバイスは、段階 2 0 6 からプロセスを繰り返し、同じ方法でパケットを処理する。

【 0 0 6 3 】

図 3 は、転送デバイス 1 0 4 と 1 0 6 との間のパスをセットアップするための方法 2 1 6 を図示するフローチャートである。例として、制御パケットが転送デバイス 1 0 4 から転送デバイス 1 0 6 へと移動した場合、転送デバイス 1 0 4 は 1 0 6 にとって、上流の転送デバイスである。

【 0 0 6 4 】

図 1 A、図 1 B、図 2 に関連して上述に引き続き、転送デバイス 1 0 6 は、上流の転送デバイス 1 0 4 から新規制御パケットを受信し、新規制御パケットは、これまでで最高の

50

パフォーマンス特性を有すると仮定する。この理由から、図2の段階216でパスが設定される。

【0065】

段階302で、転送デバイス106は、1つ上流の転送デバイスに対し、制御デバイスへのパスの要求を送信する。段階306で、転送デバイス104はパスを確立するためのラベルを決定する。段階308で、転送デバイス104は、このラベルを転送デバイス106へ送信する。段階304で、転送デバイス106はそれに応じて、そのルーティングテーブル内でこのラベルルールを確立する。

【0066】

段階310で、転送デバイス104は制御デバイスへのパスを要求する。パスを作成するべく、転送デバイス104は、他の上流の転送デバイスからラベルを要求し、段階302、306、308および310を繰り返し得る。段階312で、転送デバイス104は、ラベルルールをそのルーティングテーブル内に確立する。

【0067】

例えば、転送デバイス104は、段階302で転送デバイス106からの要求を受信後、ラベルL1を二者間のパスに選択する。このラベルを転送デバイス106へ送信後、転送デバイス106は、制御デバイス宛てのデータパケットすべてに、このラベルを付加する。

【0068】

この例では、転送デバイス104は、転送デバイス106へ接続されているポート上でラベルL1が付いたデータパケットを受信すると、そのパケットを制御デバイスへ転送する。制御デバイスへパケットを転送するべく、転送デバイス104は、特定のラベルとともに特定のポート上で転送することを(段階312でラベルルールを確立したことで)知り得る。このように、転送デバイスから制御デバイスへのパスが確立される。

【0069】

また、制御デバイスから転送デバイスへの逆方向のパスも、同様の方法で確立され得る。この場合、各転送デバイスは、そのルーティングテーブル内でラベルを設定し得る。ルーティングテーブル内のエントリには、データをルーティングするための関連するポートが含まれ得る。

【0070】

そのようなパスがネットワーク内の転送デバイスの中で確立された後、ネットワーク内では、制御デバイスと各転送デバイスとの間で、両方向のパスが存在する。図3の例では、転送デバイス106と、転送デバイス104と、制御デバイスとの間にパスが確立された後、転送デバイス106はこのパスを使用して、制御デバイスにデータを送信する。逆の場合も同様であり、制御デバイスは、このパスを使用して転送デバイス106へデータを送信する。

[新規データストリーム用の高速パスを確立する]

【0071】

図4Aは、サーバ444へアドレス指定されたネットワークユーザ442からのデータストリームを図示するダイアグラム400を示す。データストリームには、パケット402、404、406、408、410および412が含まれる。データストリームの第1パケット(パケット402)が転送デバイス106へ到達すると、それは制御デバイス102へルーティングされる。制御デバイス102は、このパケットが新規接続に属していると判定し、制御デバイス102を流れることがない、接続のための高速パスを確立するプロセスを開始する。

【0072】

例において、パケット402の後、パケット404および406も送信される。転送デバイス106は、パケット404および406をバッファリングするか、または、第1パケット402に対して行ったことと同様に、これらのパケットを制御デバイス102へと転送する。

10

20

30

40

50

【 0 0 7 3 】

制御デバイスは第 1 パケットを受信した後、転送デバイス 1 0 4 へパケットを送信する。転送デバイス 1 0 4 は、パケット内に含まれる宛先アドレス、および、図 2 および図 3 の段階 2 1 6 で設定されたパスを使用して、対応するラベルとともにパケットを対応するポートへと転送する。

【 0 0 7 4 】

次に、制御デバイスは送信元から宛先へのパスを計算する。パスにはネットワークリンクおよびルータが含まれるが、制御デバイスは含まれない。例えば、制御デバイスは、M P L S プロトコルを使用して、データストリーム用のラベルスイッチパス (L S P) を作成する。パスを決定するべく、制御デバイスは、現在または過去の帯域幅の利用状況を考慮し得る。例えば、制御デバイスは、前の数週間 (または他の周期的な期間) の間のネットワーク利用状況を考慮し得る。また、制御デバイスは、複数のパス上で負荷分散を可能にするべく、複数のパスを接続のために作成し得る。

10

【 0 0 7 5 】

また、制御デバイスは、各種パス全体における、レイテンシ、ジッタ、パケット損失、もしくは、他の何らかのパフォーマンスメトリック、ユーザのサービス水準合意、または、転送されるデータのタイプを考慮し得る。例えば、ブロードキャストビデオデータは、大きな帯域幅を必要とするが、レイテンシは比較的重要でないことがあり得る。他方、ボイスオーバー IP (V o I P) データは、そこまで大きな帯域幅を必要としないが、レイテンシはより重要であることがあり得る。ブロードキャストビデオデータの場合、制御デバイスは高帯域幅、高レイテンシパスを選択し、V o I P データの場合、制御デバイスは低帯域幅、低レイテンシパスを選択し得る。

20

【 0 0 7 6 】

他の実施形態において、制御デバイスは、特定のサーバを通してデータをルーティングし得る。例えば、特定のタイプを有するデータ、または、特定の宛先に宛てられたデータは、途中で特定のスクラビングサーバによって、スクラビングを受ける必要があり得る。スクラビングサーバは、受信データを監視するべく、または、データ上で他の分析を実施するべく、悪意のある内容がないかデータをスキャンするのに使用され得る。その実施形態において、制御デバイスは、データが特定のスクラビングサーバ、またはサーバのセットを通過するように決定し得る。

30

【 0 0 7 7 】

高速パスを決定した後、制御デバイスは、パスに沿った各転送デバイスのルーティングテーブルを更新する。データストリームに関して、複数のパスが計算された場合、制御デバイスはすべてのパスに沿ったルーティングテーブルを作成する。一実施形態において、更新済みルーティングテーブルは転送デバイスに対し、送信元 / 宛先アドレスおよび送信元 / 宛先ポートの特定の組合せを有するパケットをどのように転送するか命令し得る。代替的な実施形態において、データストリームは、ラベルによって識別され、更新済みルーティングテーブルは、ラベルが付いたパケットをどのように転送するか指示し得る。

【 0 0 7 8 】

制御デバイスは、更新済みルーティングテーブルをネットワークルータへ送信する。制御デバイスは、段階 2 1 6 ですべての転送デバイスに対して確立されたパスを使用して、ルーティングテーブルを送信し得る。ルータがどのように構成されるかについての例は、図 4 B で図示されている。

40

【 0 0 7 9 】

図 4 B は、ユーザ 4 4 2 とサーバ 4 4 4 との間でパスを確立するべく、制御デバイス 1 0 2 がどのようにネットワーク転送デバイスを構成するか図示するダイアグラム 4 2 0 を示す。

【 0 0 8 0 】

図 4 B で、制御デバイス 1 0 2 は、ユーザ 4 4 2 とサーバ 4 4 4 との間のデータストリームが、リンク 1 2 0、1 2 2、および 1 2 4、ならびに、転送デバイス 1 0 6、1 0 8

50

、 1 1 2、および 1 1 0 を含むパスを通して進むことを決定する。転送デバイスを構成するべく、制御デバイス 1 0 2 は、構成コマンド 4 2 2、4 2 4、4 2 6 および 4 2 8 を使用して、更新済みルーティングテーブルを送信する。これらの構成コマンドは、ユーザ 4 4 2 と宛先 4 4 4 とを接続する高速パス上にあるすべての転送デバイスを構成する。具体的には、コマンド 4 2 2 は転送デバイス 1 0 6 に対し、データストリーム内のパケットをリンク 1 2 0 へ転送するよう命令し、コマンド 4 2 4 は転送デバイス 1 0 8 に対し、データストリーム内のパケットをリンク 1 2 2 へ転送するよう命令し、コマンド 4 2 6 は転送デバイス 1 1 2 に対し、データストリーム内のパケットをリンク 1 2 4 へ転送するよう命令し、コマンド 4 2 8 は転送デバイス 1 1 0 に対し、データストリーム内のパケットをリンク 1 2 4 へ転送するよう命令する。

10

【 0 0 8 1 】

一実施形態において、パスが完全に確立されるまで、パケットが制御デバイスを通して流れ続けることを確実にするべく、パスに沿ったルータが、出口ポイント（転送デバイス 1 0 4 ）から入口ポイント（転送デバイス 1 0 6 ）まで構成される。第 1 に、コマンド 4 2 8 は、転送デバイス 1 1 0 を構成する。第 2 に、コマンド 4 2 6 は、転送デバイス 1 1 2 を構成する。第 3 に、コマンド 4 2 4 は、転送デバイス 1 0 8 を構成する。第 4 に、コマンド 4 2 2 は、転送デバイス 1 0 6 を構成する。

【 0 0 8 2 】

図 4 A に戻って、パスが確立されるまで、ユーザ 4 4 2 は、データストリームのパケット（パケット 4 0 2、4 0 4、4 0 6 ）の送信を継続する。上述のように、転送デバイス 1 0 6 は、パス上でこれらのパケットを制御デバイス 1 0 2 へ向かわせ続ける。

20

【 0 0 8 3 】

制御デバイス 1 0 2 は、これらのパケットを受信すると、図 2 の段階 2 1 6 であらかじめ決定された既定のパスを使用して、パケットを宛先へとルーティングする。このように、ネットワークを通るパスが確立されている間（図 4 B で図示）、パケットは宛先へルーティングされ続けるので、初期パケットをエッジルータにバッファリングする必要性が回避される。

【 0 0 8 4 】

パスが設定されると（例えば、転送デバイスが新規ルーティングテーブルで構成されると）、制御デバイスによって設定された高速パスに沿ってデータが流れる。この時点で、データは制御デバイスを流れなくなるので、転送デバイスを通るより速い速度で、より低いエンドツーエンドのレイテンシで流れる。

30

[複数の制御デバイス]

【 0 0 8 5 】

図 5 は、複数の制御デバイスを有するネットワークを図示するダイアグラムである。図 5 で示されている例示的な実施形態において、制御デバイス 1 0 2 が転送デバイス 1 0 4 に接続され、制御デバイス 5 0 4 が転送デバイス 1 1 0 に接続され、制御デバイス 5 0 6 が転送デバイス 5 0 6 に接続される。

【 0 0 8 6 】

例示的な実施形態において、各制御デバイス 1 0 2、5 0 4 または 5 0 6 は、制御パケットを定期的に、または、様々な時間間隔で、各々に直接接続された転送デバイスへ送信する。そして、図 2 で図示されているのと同じプロセスに従って、各制御デバイスは、ネットワーク内の各転送デバイスに対し、双方向パスを確立する。

40

【 0 0 8 7 】

図 5 の例示的な実施形態において、各転送デバイスは、その隣接する転送デバイスのトポロジを示す情報、および、ネットワーク状態情報をネットワーク内の各制御デバイスへ送信する。この実施形態において、転送デバイスは、プロセス 2 0 0 の段階 2 1 6 で確立された双方向パスを使用して、この情報を送信する。一実施形態において、転送デバイスは、この情報を制御デバイスに対し、定期的に、または、様々な時間間隔で送信する。

【 0 0 8 8 】

50

冗長性の目的で、単一の物理ネットワークが、複数のトポロジデータベースと共に、複数の制御デバイス上にミラーリングされ得る。これらの制御デバイスは各々、自身とネットワーク内の各ルータとの間で確立されたパスを有する。1つの制御デバイスへの既定パスに障害が発生した場合、トラフィックは異なる制御デバイスへのパスを選択し得る。

【0089】

例示的な実施形態において、各転送デバイスは、データを転送する際、および、宛先への高速パスを確立する際、1度に1台の制御デバイスを選択する。この実施形態において、転送デバイスは選択アルゴリズムに基づいて、制御デバイスのいずれか1つを選択できる。選択アルゴリズムには、例えば、ラウンドロビンアルゴリズムがあり得る。また、転送デバイスは、制御デバイスを選択するアルゴリズムで、制御デバイスが1度にセットアップしようとしているネットワーク負荷、または、新規接続の数も考慮し得る。加えて、転送デバイスは、レイテンシがもっとも低いコントローラに対し、宛先への高速パスを確立するように依頼し得る。ラウンドロビンのアプローチは、最低レイテンシのアプローチと併せて使用され得る。例えば、デバイスは、コントローラへの複数のパスにわたって、トラフィックのバランスをとることを選択し得る。例えば、おそらく、もっとも近い5への丸め処理をメトリックに行い、その結果、3、4、5、6、7、のパフォーマンス値を5の値と同じように扱う。転送デバイスは、高速パスを作成するコントローラを選択するべく、もっとも低い丸め処理済みメトリックを有するコントローラの間でラウンドロビンアルゴリズムを採用し得る。これは、ネットワークが実世界の状況に対応する際に、パフォーマンス値のわずかな修正の影響を制限するのに役立つ。

【0090】

一実施形態において、ネットワーク内の制御デバイスは、互いに同期する必要はない。制御デバイスに障害が発生した場合、制御デバイスから制御パケットを一定期間受信しなかった後で、転送デバイスは、その制御デバイスへのパスを維持するためのラベルを除去する。転送デバイスは、ラウンドロビンなどの同じ選択アルゴリズムによって、続けて他の制御デバイスを使用する。

【0091】

複数の制御デバイスが存在することで、ネットワークのレジリエンスが大幅に向上する。制御デバイスに障害が発生した場合、または、制御デバイスに接続しているリンクに障害が発生した場合、転送デバイスは、ネットワーク内にある他の制御デバイスを続けて使用する。また、複数の制御デバイスは、ネットワーク内の負荷のバランスをとることに役立つ。新規接続要求の数がより少ない制御デバイスをインテリジェントに選択することで、または、レイテンシがもっとも低いパスなどの最適パスを有する制御デバイスを選択することで、転送デバイスは、新規接続を要求する場合に、ネットワークの輻輳区間を回避できる。

[制御デバイスおよび転送デバイスモジュール]

【0092】

図6は、制御デバイス102、およびそのモジュール、ならびに、転送デバイス104、およびそのモジュールを図示するダイアグラム600を示している。

【0093】

具体的には、制御デバイス102は、コントローラ検出モジュール610、パス決定モジュール608、データ転送モジュール612、および、トポロジデータベース606を含む。各種モジュールは、図2の方法に関して上述したように機能し得る。

【0094】

コントローラ検出モジュール610は、転送デバイスのネットワークのトポロジを判定するように構成される。コントローラ検出モジュール610は、制御デバイスに直接接続されている転送デバイスへ制御パケットを送信する。制御パケットは、制御デバイスのアドレス、および、シリアル番号を含む。アドレスとシリアル番号の組合せを使用することで、各制御パケットはネットワーク全体で一意に識別され得る。

【0095】

制御パケットがネットワークを通過して移動するにつれて、制御パケットを受信する各転送デバイスは、自身のアドレスを制御パケットに追加する。また、各転送デバイスは、各制御パケットにネットワークの特性を示すメトリックを更新する。転送デバイスは、制御パケットを受信すると、制御パケットから、制御パケットが転送デバイスに到着すべく移動してきたパスが分かる。また、プロセス 200 で説明されているように、転送デバイスは、パフォーマンスメトリック情報を使用して、複数の転送パケットが移動した複数のパスの中で、どのパスが制御デバイスへの最適パスであるか分かる。したがって、転送デバイスは、制御デバイスへの最適パスを動的に更新できる。

【0096】

転送デバイスへのパスが設定された後で、コントローラ検出モジュール 610 は、ネットワーク内の転送デバイスから、トポロジおよびネットワーク状態情報を受信する。コントローラ検出モジュール 610 は、この情報、および、ネットワーク内の他の転送デバイスからの同様の情報を使用して、トポロジデータベース 606 を作成する。

【0097】

トポロジデータベース 606 は、ネットワーク内の転送デバイスの検出されたトポロジを保存する。トポロジデータベース 606 は、転送デバイスおよびリンクが共にどのように接続されているかに関する現在のトポロジを反映する。また、トポロジデータベース 606 は、ネットワークの現在の状態を示す情報を保存する。ネットワークの状態には、ネットワークリンクまたはデバイス上のレイテンシ、パケット損失、輻輳、またはジッタが含まれ得る。

【0098】

パス決定モジュール 608 は、新規ネットワーク接続のためのパスを確立する。転送元デバイスから接続の要求があると、パス決定モジュール 608 は、転送元デバイスから転送先デバイスへの接続パスを決定する。パス決定モジュール 608 は、パスを決定するのにトポロジデータベース 606 を使用する。トポロジデータベース 606 に保存されたトポロジおよびネットワーク状態情報を使用することで、パス決定モジュール 608 は、転送元デバイスから宛先への最適パスを発見できる。最適パスには、例えば、最短パス、または、遅延がもっとも少ないパス、または、パケット損失がもっとも少ないパスがあり得る。

【0099】

接続のためのパスを決定後、パス決定モジュール 608 はまた、決定されたパス上にある各転送デバイスに対し、構成コマンドを送信する。図 4B の例示的な実施形態で図示されているように、パス決定モジュールは、構成コマンド 422、424、426 および 428 を作成し、送信する。構成コマンドは、対応する転送デバイスを構成する。構成された転送デバイスは、パス決定モジュール 608 によって決定された接続用パス上の接続に属するパケットをルーティングする。

【0100】

データ転送モジュール 612 は、パス決定モジュールが接続用のパスを決定している間に、要求された接続に関する初期パケットを送り、構成コマンドを送信することによって転送デバイスを構成する。図 4A で図示されているように、初期パケット 402、404、406、408 および 410 は、転送元デバイスから制御デバイスへのパスを通過して、制御モジュールへ転送される。接続パスが確立されている間、データ転送モジュール 612 は、制御デバイスから転送先デバイスへのパスを通して初期パケットを転送する。

【0101】

図 6 で示されている例示的な実施形態において、転送デバイス 104 は、パス要求モジュール 622、スイッチ検出モジュール 624、ルーティングテーブル 626、パフォーマンスメトリックモジュール 628 を含む。

【0102】

転送デバイスで新規データ接続の要求を受信した後、パス要求モジュール 622 は、メッセージをルーティング制御デバイスへ送信し、接続パスの確立を要求する。パス要求モ

10

20

30

40

50

ジュール 6 2 2 は、プロセス 2 0 0 の段階 2 1 6 で設定された制御デバイスへのパスを使用して、接続パスを確立するための要求を送信する。パス要求は、ルーティング制御デバイスに対し、転送元デバイスから転送先デバイスへのパスを決定するように命令する。パス決定モジュール 6 0 8 は次に、接続パスを決定し、決定されたパスに沿って新規データストリームを転送するように転送デバイスを構成する。

【 0 1 0 3 】

スイッチ検出モジュール 6 2 4 は、制御デバイスのコントローラ検出モジュール 6 1 0 によって送信された制御パケットを受信する。スイッチ検出モジュール 6 2 4 は、方法 2 0 0 の段階 2 1 6 で示されているように、転送デバイスから制御デバイスへのパスを決定する。一実施形態において、スイッチ検出モジュール 6 2 4 は、制御パケット内の情報、および、制御パケットを受信するポートに基づいて、この決定を下す。

10

【 0 1 0 4 】

転送デバイスから制御デバイスへのパスを確立するべく、スイッチ検出モジュール 6 2 4 は、制御パケットを受信したポートに部分的に基づいて、ルーティング制御デバイスに到達するには、転送デバイス上のどのポート上でデータを転送するか識別する。

【 0 1 0 5 】

一実施形態において、隣接するデバイスから到着する制御パケット内のパフォーマンスメトリックが、制御デバイスから最高のパフォーマンスを有する転送デバイスへのパスを示す場合、スイッチ検出モジュール 6 2 4 は、方法 3 0 0 の段階 3 0 6 で示されているように、制御デバイスへルーティングするためのラベルを決定する。

20

【 0 1 0 6 】

また、スイッチ検出モジュール 6 2 4 は、プロセス 2 0 0 の段階 2 1 4 で示されているように、受信した制御パケットが最高のパフォーマンスのパスを示し、かつ、転送デバイスの他のポート上でまだ転送されていないかどうか判定した後で、制御パケットを受信したポートを除くそのポート上で、制御パケットをフラッディングする。

【 0 1 0 7 】

転送デバイスのルーティングテーブル 6 2 6 には、ネットワーク内でデータパケットをどのように転送するかに関する情報が含まれている。構成コマンド 4 2 2、4 2 4、4 2 6 および 4 2 8 は、パケットをどのように転送するかに関する情報で、転送デバイス内のルーティングテーブルを更新する。また、制御デバイス 1 0 2 へのパスをセットアップするためのラベルを受信後、スイッチ検出モジュール 6 2 4 は、決定されたラベルに基づいて、ルーティングテーブル 6 2 6 内に転送ルールを確立する。

30

【 0 1 0 8 】

パフォーマンスメトリックモジュール 6 2 8 は、受信制御パケットのパフォーマンスメトリックのフィールドを更新する。受信制御パケットには、制御デバイスから前の転送デバイスまでのパスのネットワークパフォーマンスメトリックを含むフィールドが含まれる。パフォーマンスメトリックモジュールは、現在の転送デバイスまでのパスのパフォーマンスを反映するように、パフォーマンスメトリックのフィールドを更新する。

【 0 1 0 9 】

また、パフォーマンスメトリックモジュール 6 2 8 は、転送デバイスが、より良いパフォーマンスメトリックを有する他の制御パケットを以前に転送したかどうか判定する。パフォーマンスメトリックモジュール 6 2 8 は、スイッチ検出モジュール 6 2 4 に対し、より良いメトリックを有する、以前の同じ制御パケットが、転送デバイスポート上でフラッディングされたかどうか通知する。

40

[結論]

【 0 1 1 0 】

本明細書で使用されるユーザという用語は、ネットワーク接続サービスを利用する企業の従業員など、ネットワーク接続サービスの顧客、および、サービス提供者自体のネットワーク管理者の両方を包含し得る。また、ユーザは異なる企業、または組織に所属し得る。

50

【 0 1 1 1 】

トポロジデータベース 6 0 6 は、永続的なメモリを含む、任意の保存タイプの構造化メモリであってよい。例では、各データベースは、関係データベース、またはファイルシステムとして実装され得る。

【 0 1 1 2 】

図 6 の複数のデバイスおよびモジュールのそれぞれは、ハードウェア、ソフトウェア、ファームウェア、またはそれらの組合せで実装され得る。

【 0 1 1 3 】

図 6 の複数のデバイス及びモジュールのそれぞれは、同じまたは異なる複数のコンピューティングデバイス上に実装され得る。そのようなコンピューティングデバイスは、限定されないが、パーソナルコンピュータ、携帯電話などのモバイルデバイス、ワークステーション、組み込みシステム、ゲームコンソール、テレビ、セットトップボックス、または、他の何らかのコンピューティングデバイスを含み得る。更に、コンピューティングデバイスは、限定されないが、命令を実行し格納するための、プロセッサと、非一時的メモリを含むメモリとを有するデバイスを含み得る。メモリは、データ及びプログラム命令を有形に具体化し得る。ソフトウェアは、1または複数のアプリケーションとオペレーティングシステムとを含み得る。ハードウェアは、限定されないが、プロセッサ、メモリ、およびグラフィカル・ユーザ・インタフェース・ディスプレイを含み得る。また、コンピューティングデバイスは、複数のプロセッサ、および複数の共有または別個のメモリ構成要素も有し得る。例えば、コンピューティングデバイスは、クラスタ化された又は分散された、コンピューティング環境またはサーバファームの一部または全体であってよい。

【 0 1 1 4 】

「(a)」、「(b)」、「(i)」、「(i i)」などの識別子は、異なる要素または段階のために用いられることがある。これらの識別子は明確にするために用いられ、必ずしも要素または段階の順序を示すものではない。

【 0 1 1 5 】

本発明は、複数の特定機能、および、それらの関係の実装を図示する複数の機能的構成単位を用いて上述された。これらの機能的構成単位の境界は、説明の利便性のために、本明細書で任意に画定された。複数の特定機能およびそれらの関係が適切に実行される限り、代替の境界が画定され得る。

【 0 1 1 6 】

複数の具体的な実施形態についての前述の説明は、本発明の全般的性質を十分に明らかにしているので、当技術分野の技術内の知識を応用することで、過度の実験なしに、本発明の一般的概念から逸脱することなく、その他の人々がそのような複数の具体的な実施形態を様々な用途のために容易に修正し得る、および/または容易に改作し得る。したがって、そのような改作物および修正物は、本明細書に提供された教示および指導に基づいて、開示された複数の実施形態の均等物の趣旨内および範囲内であるよう意図されている。本明細書の語句または用語は、説明を目的とするものであって限定するものではなく、これにより本明細書の語句または用語は、教示および指導を考慮して当業者により解釈されることになることが理解されるべきである。

【 0 1 1 7 】

本発明の広さおよび範囲は、上述された複数の例示的实施形態のいずれによっても限定されるべきではなく、次の特許請求の範囲及びその均等物のみに基づいて定義されるべきである。

本願によれば、以下の各項目もまた開示される。

[項目 1]

複数の転送デバイスからルーティング制御デバイスへの複数の経路を検出する方法であって、

(a) 前記複数の転送デバイスのうち、転送デバイスのポート上で、制御パケットを受信する段階と、

10

20

30

40

50

(b) 前記制御パケットを受信した前記ポートに少なくとも部分的に基づいて、前記ルーティング制御デバイスに到達させるべく、前記転送デバイスのどのポート上でデータを転送するか識別する段階と、

(c) 前記転送デバイスでの新規データストリームの受信に応答して、(b)で識別された前記ポート上で、メッセージを前記ルーティング制御デバイスへ送信する段階であって、前記メッセージは前記ルーティング制御デバイスに対し、前記複数の転送デバイスを通る、宛先へのパスを決定することと、決定された前記パスに沿って前記新規データストリームを転送するように前記複数の転送デバイスのサブセットを構成することとを命令する、送信する段階と、

(d) 前記ルーティング制御デバイスの複数の構成コマンドに従って、前記新規データストリームを転送する段階と
を備える方法。

[項目 2]

前記制御パケットは、前記ルーティング制御デバイスからのものであり、
前記ポートは、前記転送デバイスのデータリンク層ポートであり、前記転送デバイスを他の転送デバイスにリンクを介して接続し、

前記制御パケットは、パケットを前記他の転送デバイスから前記ルーティング制御デバイスへとルーティングするための推定時間を示すパフォーマンスメトリックを含み、

前記識別することは、前記パフォーマンスメトリックおよび前記制御パケットを受信した前記データリンク層ポートに少なくとも部分的に基づいており、

前記識別することは、前記転送デバイス上の、MACアドレスによって識別されるどのデータリンク層ポートが、前記ルーティング制御デバイスへ到達するべくデータを転送するかを識別する

項目 1 に記載の方法。

[項目 3]

前記送信する段階(c)は、前記ルーティング制御デバイスを通して、前記新規データストリームの初期パケットを転送する段階を含む、項目 1 または 2 に記載の方法。

[項目 4]

(e) 段階(a)で前記制御パケットを受信した前記ポートと異なる、前記転送デバイスの残りのポート上で、前記制御パケットを前記転送デバイスから他の転送デバイスへと転送する段階を更に含む、項目 1 から 3 のいずれか一項に記載の方法。

[項目 5]

(f) 前記転送デバイスで、前記複数の転送デバイスのうち、他の転送デバイスへとラベルを送信し、前記ラベルは、前記転送デバイスで受信されるパケットに付加されている場合に、前記転送デバイスに対し、(b)で識別された前記ポート上で前記ルーティング制御デバイスへ前記パケットを転送するように指示する段階を更に含む、項目 4 に記載の方法。

[項目 6]

前記制御パケットが、パケットを前記転送デバイスから前記ルーティング制御デバイスへとルーティングするためのシリアル番号および好ましくは推定時間を含み、
前記方法は、

(f) 前記転送デバイスが、他の制御パケットを以前に転送したかどうかを判定する段階であって、前記他の制御パケットは、前記ルーティング制御デバイスからのものであり、前記シリアル番号を有し、パケットを前記転送デバイスから前記ルーティング制御デバイスへとルーティングする他の推定時間であって、前記制御パケット内の前記推定時間より少ない推定時間を有する、判定する段階を更に含む、

前記転送デバイスが、(f)において、前記他の制御パケットを以前に転送していないと判定された場合に、前記転送する段階(e)の段階を行う、

項目 4 または 5 に記載の方法。

[項目 7]

10

20

30

40

50

(e) (b) で識別された前記ポート上でメッセージを前記ルーティング制御デバイスへと送信する段階であって、前記メッセージは、前記転送デバイスと、前記複数の転送デバイスのうち、前記転送デバイスが接続されている複数の他の転送デバイスとを識別する、送信する段階

を更に含み、

前記ルーティング制御デバイスは前記メッセージを使用してネットワークのトポロジを表すデータベースを構築し、かつ、前記データベースを使用して前記パスを決定する、

項目 1 から 6 のいずれか一項に記載の方法。

[項目 8]

前記制御パケットは、前記ルーティング制御デバイスの識別子を含み、

(e) 前記転送デバイスのポート上で、他のルーティング制御デバイスの識別子を含む他の制御パケットを受信する段階と、

(f) 前記他の制御パケットを受信した前記ポートに少なくとも部分的に基づいて、前記転送デバイスのどのポート上でデータを転送し、前記他のルーティング制御デバイスへ到達させるか識別する段階と、

(g) 新規データストリームの前記転送デバイスでの受信にตอบสนองして、前記パスが前記複数の転送デバイスを通して構成されるよう命令する前記メッセージを、前記ルーティング制御デバイスおよび前記他のルーティング制御デバイスのいずれが送信するか決定する段階と

を更に含む、

項目 1 から 7 のいずれか一項に記載の方法。

[項目 9]

ルーティング制御デバイスへの経路を検出するための転送デバイスであって、

(i) 前記転送デバイスのポート上で制御パケットを受信し、(i i) 前記制御パケットを受信した前記ポートに少なくとも部分的に基づいて、前記ルーティング制御デバイスに到達させるべく、前記転送デバイスのどのポート上でデータを転送するか識別するように構成されたスイッチ検出モジュールと、

新規データストリームの前記転送デバイスでの受信にตอบสนองして、識別された前記ポート上で前記ルーティング制御デバイスへとメッセージを送信するように構成されたパス要求モジュールであって、前記メッセージは前記ルーティング制御デバイスに対し、(i) 複数の前記転送デバイスを通る、宛先へのパスを決定することと、(i i) 決定された前記パスに沿って前記新規データストリームを転送するように、複数の前記転送デバイスのサブセットを構成することとを命令する、パス要求モジュールと、

前記ルーティング制御デバイスの複数の構成コマンドに従って、前記転送デバイスが前記新規データストリームを転送することを可能にするべくデータ入力された、前記転送デバイス内のルーティングテーブルと

を備える転送デバイス。

[項目 10]

前記ポートは、前記転送デバイスのデータリンク層ポートであり、前記転送デバイスを他の転送デバイスにリンクを介して接続し、

前記制御パケットは、パケットを前記他の転送デバイスから前記ルーティング制御デバイスへとルーティングするための推定時間を示すパフォーマンスメトリックを含み、

前記識別することは、前記パフォーマンスメトリックおよび前記制御パケットを受信した前記データリンク層ポートに少なくとも部分的に基づいており、

前記識別することは、前記転送デバイス上で、MAC アドレスによって識別されるどのデータリンク層ポートが、前記ルーティング制御デバイスへ到達するべくデータを転送するかを識別する

項目 9 に記載の転送デバイス。

[項目 11]

前記パス要求モジュールは更に、前記ルーティング制御デバイスを通して前記新規デー

10

20

30

40

50

タストリーム of の初期パケットを転送するよう構成される、項目 9 または 10 に記載の転送デバイス。

[項目 12]

前記スイッチ検出モジュールは、前記制御パケットを受信した前記ポートとは異なる前記転送デバイスの残りのポート上で、前記転送デバイスから他の転送デバイスへと前記制御パケットを転送するよう構成され、任意で、

前記スイッチ検出モジュールは、前記複数の転送デバイスの中の、他の転送デバイスへとラベルを送信するよう構成され、

前記ラベルは、前記転送デバイスで受信されたパケットに付加されている場合に、前記転送デバイスに対し、前記スイッチ検出モジュールによって識別された前記ポート上で、前記ルーティング制御デバイスへ前記パケットを転送するように指示する、

項目 9 から 11 のいずれか一項に記載の転送デバイス。

[項目 13]

前記制御パケットは、前記転送デバイスから前記ルーティング制御デバイスへとパケットをルーティングするためのシリアル番号および好ましくは推定時間を含み、前記転送デバイスは更に、パフォーマンスメトリックモジュールを備え、

前記パフォーマンスメトリックモジュールは、前記転送デバイスが、他の制御パケットを以前に転送したかどうか判定するよう構成され、前記他の制御パケットは、前記ルーティング制御デバイスからのものであり、前記シリアル番号を含み、前記転送デバイスから前記ルーティング制御デバイスへとパケットをルーティングする他のパフォーマンスメトリックであって、前記制御パケット内のパフォーマンスメトリックより良い他のパフォーマンスメトリックを含み、

前記パス要求モジュールが、前記他の制御パケットを以前に転送しなかったと判定した場合に、前記転送を行う、

項目 12 に記載の転送デバイス。

[項目 14]

前記スイッチ検出モジュールは更に、

識別された前記ポート上で前記ルーティング制御デバイスへメッセージを送信するよう構成され、前記メッセージは、前記転送デバイスと、前記転送デバイスが接続されている前記複数の転送デバイスのうち複数の他の転送デバイスとを識別し、

前記ルーティング制御デバイスは、ネットワークのトポロジを表すデータベースを構築するべく前記メッセージを使用し、前記パスを決定するべく前記データベースを使用するよう構成される、

項目 12 または 13 に記載の転送デバイス。

[項目 15]

前記制御パケットは、前記ルーティング制御デバイスの識別子を含み、前記スイッチ検出モジュールは更に、

前記転送デバイスのポート上で、他のルーティング制御デバイスの識別子を含む、他の制御パケットを受信し、

前記他の制御パケットを受信した前記ポートに少なくとも部分的に基づいて、前記他のルーティング制御デバイスに到達するべく、前記転送デバイスのどのポート上でデータを転送するか識別し、

新規データストリームの前記転送デバイス上での受信に応答して、前記パスが前記複数の転送デバイスを通して構成されるように命令するメッセージを、前記ルーティング制御デバイス、および前記他のルーティング制御デバイスのいずれに送信するか決定するよう構成される

項目 12 から 14 のいずれか一項に記載の転送デバイス。

[項目 16]

複数の転送デバイスからルーティング制御デバイスへの経路を検出するための方法を少なくとも 1 つのマシンに実行させるプログラムであって、前記方法は、

(a) 前記複数の転送デバイスのうち、転送デバイスのポート上で、制御パケットを受信する段階と、

(b) 前記制御パケットを受信した前記ポートに少なくとも部分的に基づいて、前記ルーティング制御デバイスへ到達させるべく、前記転送デバイスのどのポート上でデータを転送するか識別する段階と、

(c) 新規データストリームの前記転送デバイスでの受信に応答して、(b) で識別された前記ポート上で、メッセージを前記ルーティング制御デバイスへ送信する段階であって、前記メッセージは前記ルーティング制御デバイスに対し、前記複数の転送デバイスを通る、宛先へのパスを決定することと、決定された前記パスに沿って前記新規データストリームを転送するように前記複数の転送デバイスのサブセットを構成することとを命令する、送信する段階と、

(d) 前記ルーティング制御デバイスの複数の構成コマンドに従って、前記新規データストリームを転送する段階と

を備える、プログラム。

[項目 1 7]

前記ポートは、前記転送デバイスのデータリンク層ポートであり、前記転送デバイスを他の転送デバイスにリンクを介して接続し、

前記制御パケットは、パケットを前記他の転送デバイスから前記ルーティング制御デバイスへとルーティングするための推定時間を示すパフォーマンスメトリックを含み、

前記識別することは、前記制御パケットおよび前記パフォーマンスメトリックを受信した前記データリンク層ポートに少なくとも部分的に基づいており、

前記識別することは、前記転送デバイス上で、MAC アドレスによって識別されるどのデータリンク層ポートが、前記ルーティング制御デバイスへ到達するべくデータを転送するかを識別する

項目 1 6 に記載のプログラム。

[項目 1 8]

前記送信する段階 (c) は、前記新規データストリームの初期パケットを前記ルーティング制御デバイスを通して転送する段階を含む、項目 1 6 または 1 7 に記載のプログラム

。

[項目 1 9]

前記方法は、(e) (a) で前記制御パケットが受信された前記ポートと異なる、前記転送デバイスの残りのポート上で、前記制御パケットを前記転送デバイスから他の転送デバイスへと転送する段階を更に含み、任意で、

前記方法は、

(f) 前記転送デバイスで、前記複数の転送デバイス中の他の転送デバイスにラベルを送信する段階を更に備え、前記ラベルは、前記転送デバイスで受信されたパケットに付加されている場合に、前記転送デバイスに対し、(b) で識別された前記ポート上で、前記ルーティング制御デバイスへ前記パケットを転送するよう指示する、項目 1 6 から 1 8 のいずれか一項に記載のプログラム。

[項目 2 0]

前記制御パケットは、前記転送デバイスから、前記ルーティング制御デバイスへと、パケットをルーティングするためのシリアル番号および好ましくは推定時間を含み、前記方法は、

(f) 前記転送デバイスが、以前に転送された他の制御パケットを有するかどうか判定する段階であって、前記他の制御パケットは、前記ルーティング制御デバイスからのものであり、前記シリアル番号を有し、パケットを前記転送デバイスから前記ルーティング制御デバイスへとルーティングする他の推定時間であって、前記制御パケット内の前記推定時間より少ない推定時間を有する、判定する段階

を更に備え、

前記転送デバイスが、(f) において、前記他の制御パケットを以前に転送していない

10

20

30

40

50

と判定された場合に、前記転送する段階（e）の転送段階を行う、
項目19に記載のプログラム。

[項目21]

前記方法は、（e）（b）で識別された前記ポート上で前記ルーティング制御デバイスへメッセージを送信する段階であって、前記メッセージは、前記転送デバイスと、前記転送デバイスが接続されている、前記複数の転送デバイスのうち複数の他の転送デバイスとを識別する段階を更に備え、

前記ルーティング制御デバイスは、ネットワークのトポロジを表すデータベースを構築するべく前記メッセージを使用し、前記パスを決定するべく前記データベースを使用する、

10

項目16から20のいずれか一項に記載のプログラム。

[項目22]

前記制御パケットは、前記ルーティング制御デバイスの識別子を含み、

前記方法は、

（e）前記転送デバイスのポート上で、他のルーティング制御デバイスの識別子を含む他の制御パケットを受信する段階と、

（f）前記転送デバイスの受信した前記ポートに少なくとも部分的に基づいて、前記他のルーティング制御デバイスへ到達するべく、前記転送デバイスのどのポート上でデータを転送するか識別する段階と、

（g）新規データストリームの前記転送デバイスでの受信に応答して、前記複数の転送デバイスを通る前記パスが構成されるように命令する前記メッセージを、前記ルーティング制御デバイス、および、前記他のルーティング制御デバイスのいずれに送信するか決定する段階と

20

を更に含む、項目16から21のいずれか一項に記載のプログラム。

[項目23]

複数の転送デバイスのネットワークにおいて、少なくとも、転送デバイスのポートに接続されている、ルーティング制御デバイスであって、

前記複数の転送デバイスの前記ネットワークのトポロジを判定するよう構成されたコントローラ検出モジュールと、

前記ネットワークのトポロジを保存するよう構成されたトポロジデータベースと、

30

前記ルーティング制御デバイスから、前記ネットワーク内の各転送デバイスへのパスを確立し、

前記ネットワーク内の前記転送デバイスのそれぞれから前記ルーティング制御デバイスへのパスを確立し、

転送元デバイスからの接続要求を受けると、前記転送元デバイスから転送先デバイスへの接続パスを確立する

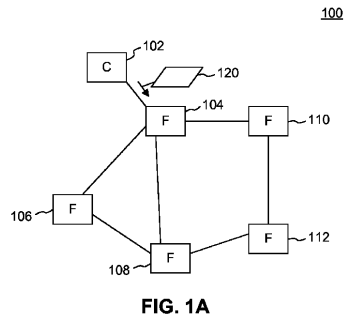
よう構成されたパス決定モジュールと、

前記接続パスが確立されている間、前記転送元デバイスから前記ルーティング制御デバイス、および、前記ルーティング制御デバイスから前記転送先デバイスまでの前記パスを通して、前記接続要求に対して少なくとも1つの初期パケットを送るよう構成されたデータ転送モジュールと

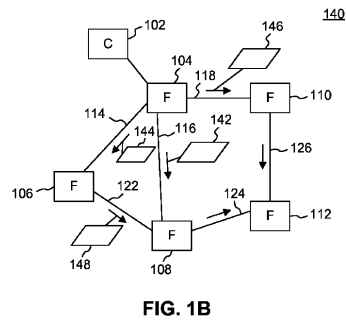
40

を備える、ルーティング制御デバイス。

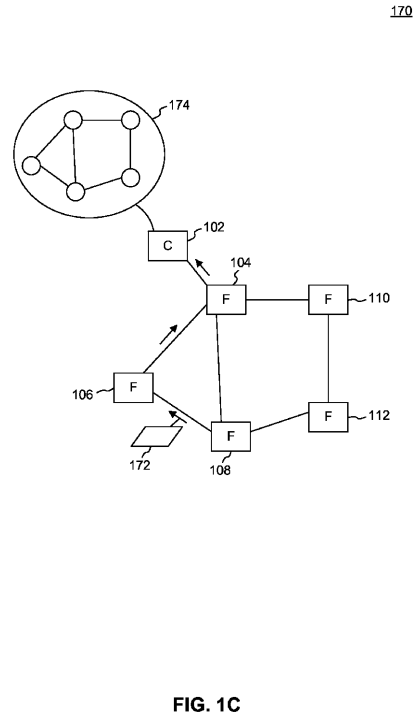
【図 1 A】



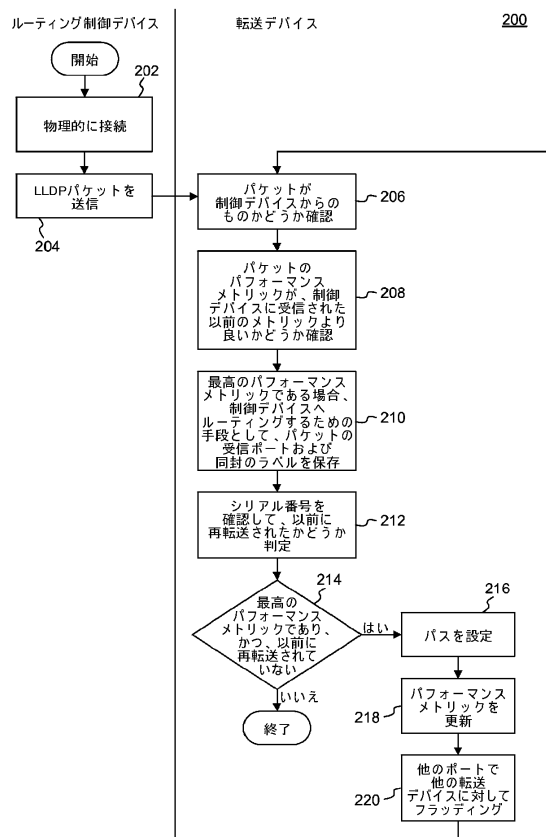
【図 1 B】



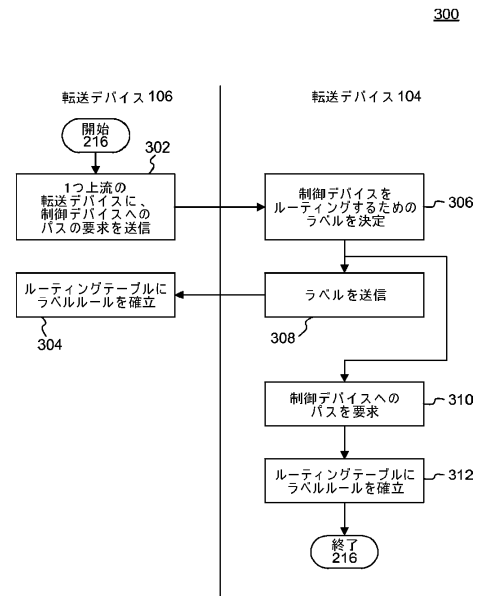
【図 1 C】



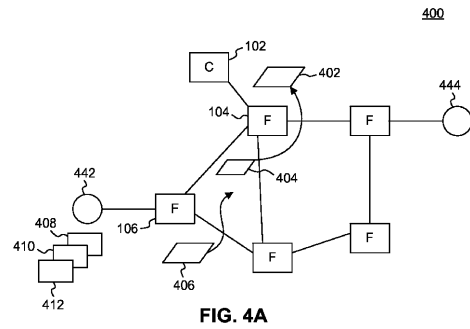
【図 2】



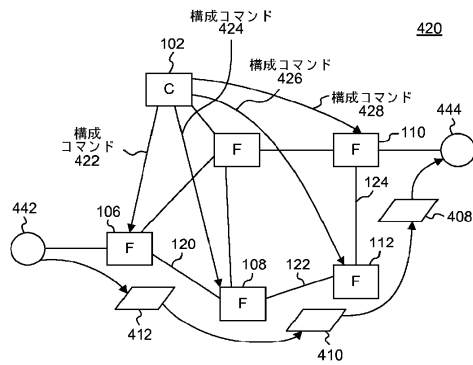
【図 3】



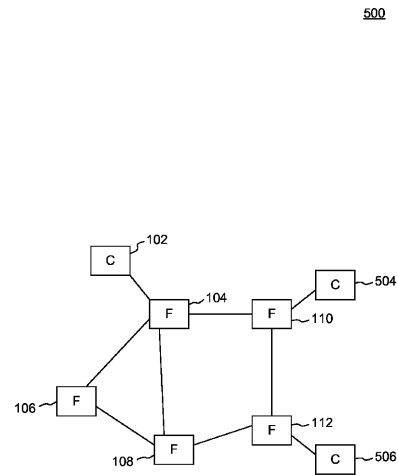
【図 4 A】



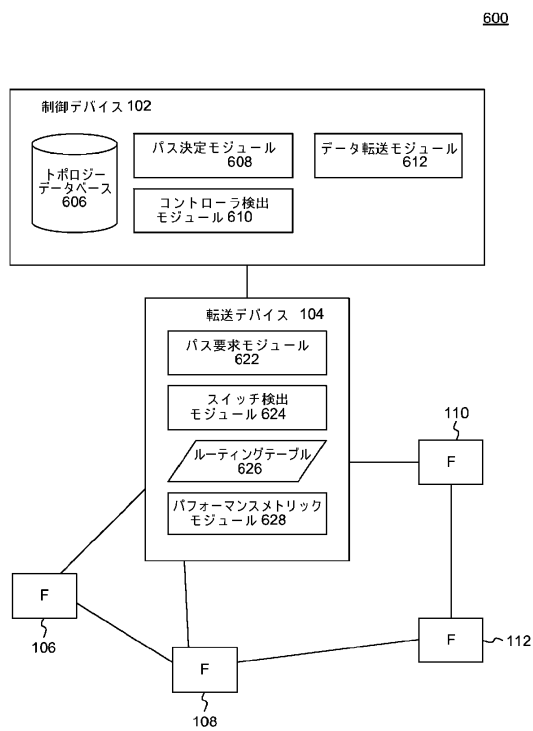
【図 4 B】



【図 5】



【図 6】



フロントページの続き

(72)発明者 セラ、ジェイムス、マイケル
アメリカ合衆国，コロラド州 80021，ブルームフィールド，エルドラド プールバード 1
025 レベル スリー コミュニケーションズ，エルエルシー内

審査官 松崎 孝大

(56)参考文献 米国特許出願公開第2008/0212585 (US, A1)
国際公開第2008/095010 (WO, A1)
国際公開第2011/108205 (WO, A1)
小出 俊夫、下西 英之，OpenFlowネットワークにおける制御ネットワークの構築自動化に関する一検討，電子情報通信学会技術研究報告，社団法人電子情報通信学会，2010年
2月25日，Vol.109 No.448，p.19-24

(58)調査した分野(Int.Cl.，DB名)
H04L 12/717