

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5565476号
(P5565476)

(45) 発行日 平成26年8月6日(2014.8.6)

(24) 登録日 平成26年6月27日(2014.6.27)

(51) Int.Cl.		F I
HO 4 L 12/749 (2013.01)		HO 4 L 12/749
HO 4 L 12/717 (2013.01)		HO 4 L 12/717

請求項の数 8 (全 17 頁)

(21) 出願番号	特願2012-557790 (P2012-557790)	(73) 特許権者	000004237
(86) (22) 出願日	平成23年12月8日(2011.12.8)		日本電気株式会社
(86) 国際出願番号	PCT/JP2011/078439		東京都港区芝五丁目7番1号
(87) 国際公開番号	W02012/111222	(74) 代理人	100102864
(87) 国際公開日	平成24年8月23日(2012.8.23)		弁理士 工藤 実
審査請求日	平成25年8月8日(2013.8.8)	(72) 発明者	森本 昌治
(31) 優先権主張番号	特願2011-31752 (P2011-31752)		東京都港区芝五丁目7番1号 日本電気株 式会社内
(32) 優先日	平成23年2月17日(2011.2.17)		
(33) 優先権主張国	日本国(JP)		
		審査官	松崎 孝大

最終頁に続く

(54) 【発明の名称】 ネットワークシステム、及びネットワークフロー追跡方法

(57) 【特許請求の範囲】

【請求項1】

設定されたフローテーブルのエントリの内容に従って受信パケットを処理する機能を持つスイッチと、

前記スイッチからパケットの問い合わせを受け、該パケットを一律に制御するためのルールと動作が定義されたエントリを、前記スイッチのフローテーブルに設定する機能を持つコントローラと、

前記スイッチ及び前記コントローラのうち少なくとも一方からパケットを受け取り、該パケットのヘッダと同じヘッダを使ってパケットをカプセル化する機能を持つカプセル化モジュールと

を含む

ネットワークシステム。

【請求項2】

請求項1に記載のネットワークシステムであって、

前記スイッチからパケットを受信した時に、受信パケットの先頭のヘッダ情報の変換を行って、他のスイッチに転送する機能を持つヘッダ変換装置を更に含み、

前記カプセル化モジュールは、前記カプセル化されたパケットが前記ヘッダ変換装置を経過した場合、変換前と変換後の2種類のヘッダが付与されたパケットを受け取り、変換前と変換後の2種類のヘッダの組を前記コントローラに通知する機能を持つ

10

20

ネットワークシステム。

【請求項 3】

請求項 1 又は 2 に記載のネットワークシステムであって、

前記カプセル化モジュールは、

前記スイッチが受信したパケットがカプセル化されたパケットかどうか確認する機能を持ち、

カプセル化されたパケットでない場合、該パケットのヘッダ情報を複製し、複製したヘッダ情報を該受信パケットに付与してカプセル化して前記スイッチに渡す機能を持ち、

カプセル化されたパケットである場合、該カプセル化されたパケットの先頭のヘッダ情報と本来のパケットのヘッダ情報との組をヘッダ変換情報として前記コントローラに通知する機能を持ち、

該カプセル化されたパケットを、前記先頭のヘッダ情報のみ付与されたパケットに変換して前記スイッチに渡す機能を持つ

ネットワークシステム。

【請求項 4】

カプセル化モジュールとしての機能を持つ計算機であって、

設定されたフローテーブルのエントリの内容に従って受信パケットを処理する機能を持つスイッチと、前記スイッチからパケットの問い合わせを受け、該パケットを一律に制御するためのルールと動作が定義されたエントリを、前記スイッチのフローテーブルに設定する機能を持つコントローラと、のうち少なくとも一方からパケットを受け取る手段と、

該パケットに対し、該パケットのヘッダと同じヘッダを使ってパケットをカプセル化する手段と

を具備する

計算機。

【請求項 5】

請求項 4 に記載の計算機であって、

前記カプセル化されたパケットが、受信パケットの先頭のヘッダ情報の変換を行って転送する機能を持つヘッダ変換装置を経過した場合、変換前と変換後の 2 種類のヘッダが付与されたパケットを受け取り、変換前と変換後の 2 種類のヘッダを参照する手段

を更に具備する

計算機。

【請求項 6】

請求項 4 又は 5 に記載の計算機であって、

受信パケットがカプセル化されたパケットかどうか確認する手段と、

前記受信パケットがカプセル化されたパケットでない場合、該パケットのヘッダ情報を複製する手段と、

複製したヘッダ情報を該受信パケットに付与してカプセル化する手段と、

前記受信パケットがカプセル化されたパケットである場合、該カプセル化されたパケットの先頭のヘッダ情報と本来のパケットのヘッダ情報との組を基にヘッダ変換情報を生成する手段と、

該カプセル化されたパケットを、前記先頭のヘッダ情報のみ付与されたパケットに変換する手段と

を具備する

計算機。

【請求項 7】

スイッチにおいて、設定されたフローテーブルのエントリの内容に従って受信パケットを処理することと、

コントローラにおいて、前記スイッチからパケットの問い合わせを受け、該パケットを一律に制御するためのルールと動作が定義されたエントリを、前記スイッチのフローテーブルに設定することと、

10

20

30

40

50

カプセル化モジュールにおいて、前記スイッチ及び前記コントローラのうち少なくとも一方からパケットを受け取り、該パケットのヘッダと同じヘッダを使ってパケットをカプセル化することとを含む

ネットワークフロー追跡方法。

【請求項 8】

カプセル化モジュールとしての機能を持つ計算機により実行されるプログラムを格納した記憶媒体であって、

設定されたフローテーブルのエントリの内容に従って受信パケットを処理する機能を持つスイッチと、前記スイッチからパケットの問い合わせを受けて該パケットを一律に制御するためのルールと動作が定義されたエントリを前記スイッチのフローテーブルに設定する機能を持つコントローラと、のうち少なくとも一方からパケットを受け取るステップと、

10

該パケットに対し、該パケットのヘッダと同じヘッダを使ってパケットをカプセル化するステップと

を計算機に実行させるためのプログラムを格納した

記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

20

本発明は、ネットワークシステムに関し、特にネットワークシステムにおけるネットワークフロー追跡方法に関する。

【背景技術】

【0002】

ネットワークシステムの制御方式の1つとして、外部の制御装置（コントロールプレーン）からノード装置（ユーザプレーン）を制御するCU（C：コントロールプレーン/U：ユーザプレーン）分離型ネットワークシステムが提案されている。

【0003】

CU分離型ネットワークシステムの一例として、コントローラからスイッチを制御してネットワークの経路制御を行うオープンフロー（OpenFlow）技術を利用したオープンフローネットワークシステムが挙げられる。オープンフロー技術の詳細については、非特許文献1に記載されている。なお、オープンフローネットワークシステムは一例に過ぎない。

30

【0004】

〔オープンフローネットワークシステムの説明〕

オープンフローネットワークシステムでは、OFC（OpenFlow Controller）等のコントローラが、OFS（OpenFlow Switch）等のスイッチのフローテーブルを操作することによりスイッチの挙動を制御する。コントローラとスイッチは、セキュアチャネル（Secure Channel）で接続されている。

【0005】

40

オープンフローネットワークシステムにおけるスイッチとは、オープンフローネットワークを形成し、コントローラの制御下にあるエッジスイッチ及びコアスイッチのことである。オープンフローネットワークにおける入力側エッジスイッチでのパケット（packet）の受信から出力側エッジスイッチでの送信までのパケットの一連の流れをフロー（Flow）と呼ぶ。

【0006】

パケットは、フレーム（frame）と読み替えても良い。パケットとフレームの違いは、プロトコルが扱うデータの単位（PDU：Protocol Data Unit）の違いに過ぎない。パケットは、「TCP/IP」（Transmission Control Protocol/Internet Protocol）のPDUである。

50

一方、フレームは、「Ethernet（登録商標）」のPDUである。

【0007】

フローテーブルとは、所定のマッチ条件（ルール）に適合するパケット（通信データ）に対して行うべき所定の動作（アクション）を定義したフローエントリ（Flow entry）が登録されたテーブルである。

【0008】

フローエントリのルールは、パケットの各プロトコル階層のヘッダ領域に含まれる宛先アドレス（Destination Address）、送信元アドレス（Source Address）、宛先ポート（Destination Port）、送信元ポート（Source Port）のいずれか又は全てを用いた様々な組み合わせにより定義され、区別可能である。なお、上記のアドレスには、MACアドレス（Media Access Control Address）やIPアドレス（Internet Protocol Address）を含むものとする。また、上記に加えて、入口ポート（Ingress Port）の情報も、フローエントリのルールとして使用可能である。また、フローエントリのルールとして、フローを示すパケットのヘッダ領域の値の一部（又は全部）を、正規表現やワイルドカード「*」等で表現したものを設定することもできる。

10

【0009】

フローエントリの実行は、「特定のポートに出力する」、「廃棄する」、「ヘッダを書き換える」といった動作である。例えば、スイッチは、フローエントリの実行に出力ポートの識別情報（出力ポート番号等）が示されていれば、これに該当するポートにパケットを出力し、出力ポートの識別情報が示されていない場合は、パケットを破棄する。或いは、スイッチは、フローエントリの実行にヘッダ情報が示されていれば、当該ヘッダ情報に基づいてパケットのヘッダを書き換える。

20

【0010】

オープンフローネットワークシステムにおけるスイッチは、フローエントリのルールに適合するパケット群（パケット系列）に対して、フローエントリの実行を実行する。

【0011】

オープンフローネットワークシステムのようなフローベースのネットワークでは、フローエントリの実行（所定のヘッダ条件等）にマッチするパケット群（パケット系列）をフロー（Flow）として扱う。フロー単位でトラフィックの監視や制御を行うことで、従来のネットワークよりも柔軟にネットワークを制御することができる。

30

【0012】

例えば、サーバとクライアント間のトラフィックをユーザ毎に制御したい場合、サーバとクライアントのIPアドレスを組み合わせ、エンド・ツー・エンドでフローを監視したり、トラフィック量を制御したりすることが可能となる。

【0013】

なお、現在のネットワークは非常に複雑な構成となっており、サーバやクライアントマシンの前段にはファイアウォール（firewall）やロードバランサ（負荷分散装置）といった様々な機能を持った機器が配置されていることが多く、これらの装置によってフローベースのネットワーク制御の利点が失われることがある。

40

【0014】

中でも、NAT（Network Address Translation）やNAPT（Network Address Port Translation）の機能を持つ装置は、パケットヘッダの変換を行う。例えば、NATの機能を持つ装置は、IPヘッダを書き換え、NAPTの機能を持つ装置は、IPヘッダ、及び、レイヤ4ヘッダをも書き換える。

【0015】

このような装置（以後、ヘッダ変換装置と称する）を経由した場合、パケットヘッダが

50

変換されてしまうため、経路前のフローと経路後のフローが別のフローになる。

【0016】

したがって、このようなヘッダ変換装置を中継した場合は、エンド・ツー・エンドのフロー毎の監視や制御ができない。

【0017】

このような課題を解決する手法として、例えば、ヘッダ変換装置が保持するヘッダ変換情報を利用する方法がある。

【0018】

具体的には、ヘッダ変換装置に問い合わせたアドレス変換テーブルを参照し、変換後のパケットヘッダの情報から変換前のパケットヘッダの情報を取得して、対応するフローを見つける、という方法が考えられる。

10

【0019】

しかしながら、こういった方法は、ヘッダ変換装置が装置の外部から問い合わせ可能なインターフェースを持っており、かつ、アドレス変換テーブルの情報を参照可能な場合のみ実現できる。もし、このような条件が揃っていなければ、ヘッダ変換装置を改造して条件を揃える必要がある。

【0020】

また、別の手法として、特許文献1(特開2005-210518号公報)では、IPトレースバックを行う装置である発信源追跡情報提供装置及び発信源追跡装置が開示されている。

20

【0021】

このIPトレースバック技術の代表例として、IETF(Internet Engineering Task Force)のICMPトレースバックワーキンググループが提唱するICMPトレースバック(Internet Control Message Protocol Traceback)がある。

【0022】

ICMPトレースバックは、途中経路上のルータ装置が一定確率で追跡対象のIPパケットを選択し、このIPパケットに対する追跡情報を生成し、この追跡情報をICMPメッセージでIPパケットの宛先に送信し、そして、宛先の装置がこの追跡情報を表示するものである。

30

【0023】

これと同様な手法を適用することで、分断されたフローの対応関係を取得できる可能性もある。

【0024】

しかし、このような手法を実現するには、ヘッダ変換装置にアドレス変換情報を外部に送信する仕組みが必要である。

【0025】

したがって、ヘッダ変換装置を改造せずを実現することは難しい。

【先行技術文献】

【特許文献】

40

【0026】

【特許文献1】特開2005-210518号公報

【非特許文献】

【0027】

【非特許文献1】“OpenFlow Switch Specification, Version 1.0.0”, [online], December 31, 2009, インターネット(URL: <http://www.openflowswitch.org/documents/openflow-spec-v1.0.0.pdf>)

【発明の概要】

【0028】

50

本発明の目的は、スイッチにて、パケットを、現在のヘッダと同じヘッダでカプセル化することにより、ネットワーク・アプライアンス通過後には、パケットに変換前と変換後の2種類のヘッダが付属している状態となるネットワークシステム及びネットワークフロー追跡方法を提供することである。なお、ヘッダは、レイヤ2からレイヤ4までの全てのヘッダを含む。具体的には、上記のヘッダは、「フローエントリのルールとなり得る情報の全部又は一部」と読み替える。

【0029】

本発明に係るネットワークシステムは、設定されたフローテーブルのエントリ（フローエントリ）の内容に従って受信パケットを処理する機能を持つスイッチと、該スイッチからパケットの問い合わせを受け、該パケットを一律に制御するためのルールと動作が定義されたエントリを、該スイッチのフローテーブルに設定する機能を持つコントローラと、該スイッチ及び該コントローラのうち少なくとも一方からパケットを受け取り、該パケットのヘッダと同じヘッダを使ってパケットをカプセル化する機能を持つカプセル化モジュールとを含む。

10

【0030】

本発明に係る計算機は、カプセル化モジュールとしての機能を持つ計算機であって、設定されたフローテーブルのエントリの内容に従って受信パケットを処理する機能を持つスイッチと、該スイッチからパケットの問い合わせを受け、該パケットを一律に制御するためのルールと動作が定義されたエントリを、該スイッチのフローテーブルに設定する機能を持つコントローラと、のうち少なくとも一方からパケットを受け取る装置と、該パケットに対し、該パケットのヘッダと同じヘッダを使ってパケットをカプセル化する装置とを具備する。

20

【0031】

本発明に係るネットワークフロー追跡方法では、スイッチにおいて、設定されたフローテーブルのエントリの内容に従って受信パケットを処理する。また、コントローラにおいて、該スイッチからパケットの問い合わせを受け、該パケットを一律に制御するためのルールと動作が定義されたエントリを、該スイッチのフローテーブルに設定する。また、カプセル化モジュールにおいて、該スイッチ及びコントローラのうち少なくとも一方からパケットを受け取り、該パケットのヘッダと同じヘッダを使ってパケットをカプセル化する。

30

【0032】

本発明に係るプログラムは、カプセル化モジュールとしての機能を持つ計算機により実行されるプログラムであって、設定されたフローテーブルのエントリの内容に従って受信パケットを処理する機能を持つスイッチと、該スイッチからパケットの問い合わせを受けて該パケットを一律に制御するためのルールと動作が定義されたエントリを該スイッチのフローテーブルに設定する機能を持つコントローラと、のうち少なくとも一方からパケットを受け取るステップと、該パケットに対し、該パケットのヘッダと同じヘッダを使ってパケットをカプセル化するステップとを計算機に実行させるためのプログラムである。すなわち、本発明に係るプログラムは、上記のネットワークフロー追跡方法における処理を、計算機に実行させるためのプログラムである。なお、本発明に係るプログラムは、記憶装置や記憶媒体に格納することが可能である。

40

【0033】

オープンフローネットワークシステムのようなフローベースのネットワークにおいて、NATや、NAPT等のパケットヘッダを変換するようなネットワーク・アプライアンスを経由した場合にも、その前後での2つのフローの対応関係が把握でき、エンド・ツー・エンドでのフローの追跡が可能となる。

【図面の簡単な説明】

【0034】

【図1】本発明に係るネットワークシステムの基本的な構成例、及び当該システムにおけるパケットの変化を説明するための図である。

50

【図2】ヘッダ変換装置の前段にあるスイッチに対するカプセル化モジュールの処理を説明するためのフローチャートである。

【図3】ヘッダ変換装置の後段にあるスイッチに対するカプセル化モジュールの処理を説明するためのフローチャートである。

【図4】本発明に係るネットワークシステムの概念及び実施例を説明するための図である。

【発明を実施するための形態】

【0035】

本発明は、CU分離型ネットワークシステムを対象としている。ここでは、CU分離型ネットワークシステムの1つであるオープンフローネットワークシステムを例に説明する。但し、実際には、オープンフローネットワークシステムに限定されない。

10

【0036】

<実施形態>

以下に、本発明の実施形態について添付図面を参照して説明する。

【0037】

[基本構成]

図1に示すように、本発明に係るネットワークシステムは、コントローラ10と、スイッチ20(20-i、i=1~n:nは台数)と、ヘッダ変換装置30を含む。

【0038】

[コントローラ]

コントローラ10は、ネットワークの接続状態を示すトポロジ情報(topology)を基に、スイッチ20(20-i、i=1~n)を検知した際、パケット転送経路を計算し、当該経路に関連するスイッチのフローテーブルにエントリ(フローエントリ)の登録を行う。

20

【0039】

[スイッチ]

スイッチ20(20-i、i=1~n)の各々は、受信したパケットを、自身のフローテーブルに登録されたエントリに従って転送する。

【0040】

ここでは、スイッチ20(20-i、i=1~n)の各々は、仮想スイッチであるものとする。仮想スイッチとは、物理マシン上で稼働する仮想マシンにより実現されるスイッチである。但し、実際には、スイッチ20(20-i、i=1~n)の各々は、本発明に係るスイッチとしての機能を実現するためのソフトウェアがインストールされた物理スイッチでも良い。

30

【0041】

また、スイッチ20(20-i、i=1~n)は、カプセル化モジュール21(21-i、i=1~n)と連携する。

【0042】

カプセル化モジュール21(21-i、i=1~n)は、IPパケットのカプセル化処理(encapsulation)を行う。

40

【0043】

なお、カプセル化モジュール21(21-i、i=1~n)は、スイッチ20(20-i、i=1~n)が稼働する物理マシンに内蔵されていても良いし、当該物理マシンに外部接続されていても良い。例えば、カプセル化モジュール21(21-i、i=1~n)は、当該物理マシンと通信可能な計算機上で稼働していても良い。

【0044】

[ヘッダ変換装置]

ヘッダ変換装置30は、所定のスイッチ20から受信したパケットのヘッダ情報の変換を行い、他のスイッチ20に転送する。

【0045】

50

[ハードウェアの例示]

以下に、本発明に係るネットワークシステムを実現するための具体的なハードウェアの例について説明する。

【 0046 】

コントローラ10の例として、PC(パソコン)、アプライアンス(appliance)、シンクライアント端末/サーバ、ワークステーション、メインフレーム、スーパーコンピュータ等の計算機を想定している。また、コントローラ10は、計算機に搭載される拡張ボードや、物理マシン上に構築された仮想マシン(Virtual Machine(VM))でも良い。

【 0047 】

仮想スイッチとしてスイッチ20(20-i、i=1~n)が動作する物理マシン、及びヘッダ変換装置30の例として、ネットワークスイッチ(network switch)、ルータ(router)、プロキシ(proxy)、ゲートウェイ(gateway)、ファイアウォール、ロードバランサ、基地局、アクセスポイント、或いは、複数の通信ポートを有する計算機等が考えられる。

【 0048 】

コントローラ10、スイッチ20(20-i、i=1~n)が動作する物理マシン、及びヘッダ変換装置30は、プログラムに基づいて駆動し所定の処理を実行するプロセッサと、当該プログラムや各種データを記憶するメモリと、ネットワークに接続するための通信用インターフェースによって実現される。

【 0049 】

上記のプロセッサの例として、CPU(Central Processing Unit)、ネットワークプロセッサ(NP:Network Processor)、マイクロプロセッサ(microprocessor)、マイクロコントローラ(microcontroller)、或いは、専用の機能を有する半導体集積回路(LSI:Large Scale Integration)等が考えられる。

【 0050 】

上記のメモリの例として、RAM(Random Access Memory)、ROM(Read Only Memory)、EEPROM(Electrically Erasable and Programmable Read Only Memory)やフラッシュメモリ等の半導体記憶装置、HDD(Hard Disk Drive)やSSD(Solid State Drive)等の補助記憶装置、又は、DVD(Digital Versatile Disk)等のリムーバブルディスクや、SDメモリカード(Secure Digital memory card)等の記憶媒体(メディア)等が考えられる。

【 0051 】

なお、上記のプロセッサ及び上記のメモリは、一体化していても良い。例えば、近年では、マイコン等の1チップ化が進んでいる。したがって、計算機等に搭載される1チップマイコンが、プロセッサ及びメモリを備えている事例が考えられる。

【 0052 】

上記の通信用インターフェースの例として、ネットワーク通信に対応した基板(マザーボード、I/Oボード)やチップ等の半導体集積回路、NIC(Network Interface Card)等のネットワークアダプタや同様の拡張カード、アンテナ等の通信装置、接続口(コネクタ)等の通信ポート等が考えられる。

【 0053 】

また、ネットワークの例として、インターネット、LAN(Local Area Network)、無線LAN(Wireless LAN)、WAN(Wide Area Network)、バックボーン(Backbone)、ケーブルテレビ(CATV)回線、固定電話網、携帯電話網、WiMAX(IEEE 802.16a)、3G(3rd Generation)、専用線(lease line)、IrDA(Infr

10

20

30

40

50

ared Data Association)、Bluetooth(登録商標)、シリアル通信回線、データバス等が考えられる。

【0054】

カプセル化モジュール21(21-i、i=1~n)の例として、スイッチ20(20-i、i=1~n)と同じ物理マシン上で稼働するソフトウェア又は仮想マシンを想定している。但し、実際には、カプセル化モジュール21(21-i、i=1~n)は、スイッチ20(20-i、i=1~n)の各々がアクセス可能な外部サーバ上で稼働するソフトウェア又は仮想マシンでも良い。また、カプセル化モジュール21(21-i、i=1~n)は、ソフトウェアに限らず、専用デバイス、物理マシンに搭載される拡張ボード又は周辺機器、或いはネットワーク上の中間装置(ミドルボックス)でも良い。

10

【0055】

但し、実際には、これらの例に限定されない。

【0056】

[コントローラの動作]

コントローラ10は、事前に、スイッチ20-1のカプセル化モジュール21-1に対して、「所定のフローに関するパケットのカプセル化を行う」ように設定しておく。

【0057】

具体的には、コントローラ10は、事前に、スイッチ20-1のカプセル化モジュール21-1に対して、「所定のフローに関するパケットのIPヘッダを複製(コピー)し、当該IPヘッダと同じヘッダでパケットをカプセル化し、カプセル化されたパケットを転送する」ように設定しておく。ここでは、IPヘッダを例に説明するが、IPヘッダはヘッダの一例に過ぎない。実際には、レイヤ3のIPヘッダに限らず、レイヤ2からレイヤ4までの全てのヘッダを対象としても良い。具体的には、「IPヘッダ」を「フローエントリのルールとなり得る情報の全部又は一部」と読み替える。

20

【0058】

また、コントローラ10は、事前に、スイッチ20-2のカプセル化モジュール21-2に対して、「変換前ヘッダと変換後ヘッダのペア(組)をヘッダ変換情報としてコントローラに送信すると共に、受信パケットのカプセル化を解除するため、受信パケットに変換後ヘッダがあれば除去し、変換前ヘッダがあれば変換後ヘッダに置き換える処理を行う」ように設定しておく。なお、「受信パケットから一旦ヘッダを全て除去した後、変換後ヘッダのみ付与する処理を行う」ように設定しても良い。また、「受信パケットに2段目のヘッダ(変換前ヘッダ)があれば除去する処理を行う」ように設定しても良い。

30

【0059】

[第1スイッチの動作]

次に、スイッチ20-1は、入力されたパケットをカプセル化モジュール21-1に渡す。ここでは、入力されたパケットの構成は、「IPヘッダ1-ペイロード」とする。

【0060】

カプセル化モジュール21-1は、コントローラ10から設定された通り、パケットのIPヘッダ1を複製し、IPヘッダ1と同じヘッダでパケットをカプセル化し、パケットを転送する。すなわち、転送されるパケットは、ペイロードにIPヘッダ1が二重(二段)に付与されている。したがって、スイッチ20-1から転送されるパケットの構成は、「IPヘッダ1-IPヘッダ1-ペイロード」となる。

40

【0061】

この場合、「IPヘッダ1-IPヘッダ1-ペイロード」のうち、先頭の「IPヘッダ1」の部分がカプセル化されたパケットのIPヘッダとなり、残りの「IPヘッダ1-ペイロード」の部分(本来のパケット)がカプセル化されたパケットのペイロードとなる。

【0062】

なお、パケットのカプセル化の処理については、例えば、RFC1701、RFC2784で提示されているGRE(Generic Routing Encapsulation)のような方式を用いれば良い。

50

【 0 0 6 3 】

[ヘッダ変換装置の動作]

次に、ヘッダ変換装置 3 0 は、入力されたパケットの IP ヘッダ 1 を参照し、アドレス変換処理を行い、IP ヘッダ 1 を IP ヘッダ 2 に変換する。

【 0 0 6 4 】

例えば、ヘッダ変換装置 3 0 は、入力されたパケットの IP ヘッダ 1 において、送信先 IP アドレスがプライベートアドレス「192.168.0.10」の場合、この IP ヘッダ 1 を、送信先 IP アドレスがグローバルアドレス「10.0.0.10」の IP ヘッダ 2 に変換する。

【 0 0 6 5 】

このようにして、ヘッダ変換装置 3 0 は、入力されたパケットの IP ヘッダ 1 を、異なる値を持つ IP ヘッダ 2 に変換し、パケットを転送する。ここでは、ヘッダ変換装置 3 0 は、二重（二段）に付与されている IP ヘッダ 1 のうち先頭のものを除去／変換して、代わりに IP ヘッダ 2 をペイロードに付与した後、パケットを転送する。したがって、転送されるパケットの構成は、「IP ヘッダ 2 - IP ヘッダ 1 - ペイロード」となる。

【 0 0 6 6 】

すなわち、ヘッダ変換装置 3 0 は、スイッチ 2 0 - 1 からカプセル化されたパケットを受信した際、「IP ヘッダ 1 - IP ヘッダ 1 - ペイロード」のうち、カプセル化されたパケットの IP ヘッダである先頭の「IP ヘッダ 1」の部分、「IP ヘッダ 2」に変換する。

【 0 0 6 7 】

この場合、「IP ヘッダ 2 - IP ヘッダ 1 - ペイロード」のうち、先頭の「IP ヘッダ 2」の部分が転送されるパケットの IP ヘッダとなり、残りの「IP ヘッダ 1 - ペイロード」の部分（本来のパケット）が転送されるパケットのペイロードとなる。

【 0 0 6 8 】

[第 2 スイッチの動作]

次に、スイッチ 2 0 - 2 は、入力されたパケットをカプセル化モジュール 2 1 - 2 に渡す。

【 0 0 6 9 】

カプセル化モジュール 2 1 - 2 は、コントローラ 1 0 から設定された通り、「変換前ヘッダと変換後ヘッダのペア（組）をヘッダ変換情報としてコントローラに送信すると共に、受信パケットに変換後ヘッダがあれば除去し、変換前ヘッダがあれば変換後ヘッダに置き換える処理を行う」ように設定しておく。或いは、「受信パケットから一旦ヘッダを全て除去した後、変換後ヘッダのみ付与する処理を行う」ように設定しても良い。又は、「受信パケットに 2 段目のヘッダ（変換前ヘッダ）があれば除去する処理を行う」ように設定しても良い。

【 0 0 7 0 】

ここでは、カプセル化モジュール 2 1 - 2 は、入力されたパケットがカプセル化されたパケットであるか確認し、カプセル化されたパケットであれば、「IP ヘッダ 2 - IP ヘッダ 1 - ペイロード」のうち、IP ヘッダ 2 と IP ヘッダ 1 を参照して、これらのペア（組）を基にヘッダ変換情報を作成し、コントローラ 1 0 に送信する。

【 0 0 7 1 】

上述の例であれば、ヘッダ変換情報は、「送信元 IP アドレスが「192.168.0.10」が「10.0.0.10」に変換された」ということを表す情報になる。

【 0 0 7 2 】

また、カプセル化モジュール 2 1 - 2 は、入力されたパケットについて、カプセル化を解除するため、「IP ヘッダ 2 - IP ヘッダ 1 - ペイロード」のうち、先頭の IP ヘッダ 2 を除去し、残りの「IP ヘッダ 1 - ペイロード」の部分（本来のパケット）のうち、「IP ヘッダ 1」を「IP ヘッダ 2」に置き換える。

【 0 0 7 3 】

10

20

30

40

50

このようにして、カプセル化モジュール 21 - 2 は、本来のパケットに対して、IPヘッダ 1 を除去して IPヘッダ 2 をペイロードに付与した後、パケットを転送する。したがって、転送されるパケットの構成は、「IPヘッダ 2 - ペイロード」となる。

【0074】

以上のように動作することで、ヘッダ変換装置 30 を一切改良・変更することなく、コントローラ 10 は、ヘッダ変換情報を取得することができ、その情報を用いてフローの追跡が可能となる。

【0075】

[応用例]

なお、スイッチ 20 (20 - i、i = 1 ~ n) における処理 (送信側での「カプセル化」、及び受信側での「カプセル化の解除」) は、コントローラ 10 自身が行っても良い。この場合、コントローラ 10 は、カプセル化モジュール 21 (21 - i、i = 1 ~ n) と連携する。また、コントローラ 10 と、カプセル化モジュール 21 (21 - i、i = 1 ~ n) が、同一の装置であっても良い。スイッチ 20 (20 - i、i = 1 ~ n) は、入力されたパケット全体をコントローラ 10 に送信し、応答として処理結果を受信する。例えば、図 1 において、コントローラ 10 は、カプセル化モジュール 21 - 1 及びカプセル化モジュール 21 - 2 と連携又は一体化しているものとする。スイッチ 20 - 1 は、「受信パケット」をコントローラ 10 に送信し、応答としてコントローラ 10 から「カプセル化されたパケット」を受信する。また、スイッチ 20 - 2 は、「カプセル化されたパケット」をコントローラ 10 に送信し、応答としてコントローラ 10 から「変換後ヘッダのみ付与されたパケット」を受信する。したがって、必要な場所で必要な処理を行うことができるならば、カプセル化モジュール 21 (21 - i、i = 1 ~ n) の配置 (所在) は、コントローラ 10 側及びスイッチ 20 (20 - i、i = 1 ~ n) 側のいずれでも良い。無論、スイッチ 20 (20 - i、i = 1 ~ n) とコントローラ 10 の両方で処理を行うことも可能である。例えば、通常のパケットはスイッチ 20 (20 - i、i = 1 ~ n) 上で処理を行い、重要フローのパケットはコントローラ 10 上で処理を行うようにしても良い。

【0076】

このとき、カプセル化モジュール 21 (21 - i、i = 1 ~ n) は、コントローラ 10 及びスイッチ 20 (20 - i、i = 1 ~ n) の各々と通信可能な計算機上で稼働しているも良い。

【0077】

また、ここでは、説明の簡略化のため、IPパケットをNATで変換する事例について説明しているが、MACフレームをMAT (MAC Address Translation) で変換する事例や、TCP/UDPパケットをNAPTで変換する事例についても、同様の方法で実施することができる。実際には、レイヤ 3 の IPヘッダに限らず、レイヤ 2 からレイヤ 4 までの全てのヘッダを変換の対象とすることができる。具体的には、本発明の説明における「IPヘッダ (又はヘッダ) 」を「フローエントリのルールとなり得る情報の全部又は一部」と読み替える。

【0078】

[第 1 のカプセル化モジュールの動作]

図 2 のフローチャートを用いて、カプセル化モジュール 21 - 1 の処理について説明する。

【0079】

(1) ステップ S 101

カプセル化モジュール 21 - 1 は、パケットを受信する。ここでは、受信パケットの構成は、「IPヘッダ 1 - ペイロード」とする。

【0080】

(2) ステップ S 102

次に、カプセル化モジュール 21 - 1 は、受信パケットがカプセル化対象のパケットであるかどうかを確認する。

【 0 0 8 1 】

(3) ステップ S 1 0 3

カプセル化対象のパケットである場合、カプセル化モジュール 2 1 - 1 は、IP ヘッダ 1 を参照し、IP ヘッダ 1 を複製して、同じ IP ヘッダ 1 でカプセル化を行う。この場合、パケットの構成は、「IP ヘッダ 1 - IP ヘッダ 1 - ペイロード」となる。

【 0 0 8 2 】

(4) ステップ S 1 0 4

カプセル化モジュール 2 1 - 1 は、スイッチ 2 0 - 1 を介してパケットを転送する。カプセル化対象のパケットである場合、パケットの構成は、「IP ヘッダ 1 - IP ヘッダ 1 - ペイロード」である。カプセル化対象のパケットでない場合、パケットの構成は、「IP ヘッダ 1 - ペイロード」である。

10

【 0 0 8 3 】

[ヘッダ変換装置の動作]

ヘッダ変換装置 3 0 は、転送されたパケットの IP ヘッダを変換する。カプセル化されたパケットである場合、ヘッダ変換装置 3 0 は、パケットの構成を、「IP ヘッダ 1 - IP ヘッダ 1 - ペイロード」から「IP ヘッダ 2 - IP ヘッダ 1 - ペイロード」に変換する。カプセル化されたパケットでない場合、ヘッダ変換装置 3 0 は、パケットの構成を、「IP ヘッダ 1 - ペイロード」から「IP ヘッダ 2 - ペイロード」に変換する。

【 0 0 8 4 】

ヘッダ変換装置 3 0 の処理については、本発明独自の処理ではなく、一般的なヘッダ変換であるため、説明を省略する。

20

【 0 0 8 5 】

[第 2 のカプセル化モジュールの動作]

次に、図 3 のフローチャートを用いて、カプセル化モジュール 2 1 - 2 の処理について説明する。

【 0 0 8 6 】

(1) ステップ S 2 0 1

カプセル化モジュール 2 1 - 2 は、パケットを受信する。

【 0 0 8 7 】

(2) ステップ S 2 0 2

次に、カプセル化モジュール 2 1 - 2 は、受信パケットがカプセル化されたパケットであるかどうかを確認する。カプセル化されたパケットである場合、パケットの構成は、「IP ヘッダ 2 - IP ヘッダ 1 - ペイロード」である。カプセル化されたパケットでない場合、パケットの構成は、「IP ヘッダ 2 - ペイロード」である。

30

【 0 0 8 8 】

(3) ステップ S 2 0 3

カプセル化されたパケットである場合、カプセル化モジュール 2 1 - 2 は、IP ヘッダ 2 と IP ヘッダ 1 の二つを参照し、ヘッダ変換情報を作成する。ヘッダ変換情報は、変換前ヘッダ「IP ヘッダ 1」と、変換後ヘッダ「IP ヘッダ 2」のペア(組)となる。

【 0 0 8 9 】

(4) ステップ S 2 0 4

次に、カプセル化モジュール 2 1 - 2 は、作成されたヘッダ変換情報をコントローラ 1 0 に送信する。

40

【 0 0 9 0 】

(5) ステップ S 2 0 5

次に、カプセル化モジュール 2 1 - 2 は、カプセル化されたパケット「IP ヘッダ 2 - IP ヘッダ 1 - ペイロード」から「IP ヘッダ 2」を除去し、残りの「IP ヘッダ 1 - ペイロード」の部分に対して、「IP ヘッダ 1」と「IP ヘッダ 2」を置き換える。このとき、カプセル化モジュール 2 1 - 2 は、最初に除去した「IP ヘッダ 2」を変換後ヘッダとして保持しておき、残りの「IP ヘッダ 1 - ペイロード」の部分から変換前ヘッダ「IP

50

Pヘッダ1」を除去し、変換後ヘッダ「IPヘッダ2」をペイロードに付け替える。すなわち、変換前ヘッダ「IPヘッダ1」を変換後ヘッダ「IPヘッダ2」に置換する。或いは、カプセル化されたパケット「IPヘッダ2 - IPヘッダ1 - ペイロード」から一旦全てのヘッダ「IPヘッダ2 - IPヘッダ1」を除去した後、変換後ヘッダ「IPヘッダ2」のみペイロードに付与する。又は、受信パケットに付与された2段目のヘッダ（変換前ヘッダ「IPヘッダ1」）を除去する処理を行う。これにより、パケットの構成は、「IPヘッダ2 - ペイロード」となる。

【0091】

(6)ステップS206

カプセル化モジュール21-2は、スイッチ20-2を介してパケットを転送する。パケットの構成は、「IPヘッダ2 - ペイロード」である。

10

【0092】

[補足]

なお、コントローラ10は、ヘッダ変換情報を一度取得できれば良く、取得後はパケットのカプセル化を停止するように処理しても良い。

【0093】

以上のように、カプセル化モジュール21-1とカプセル化モジュール21-2がそれぞれ、パケットのカプセル化、及びヘッダ変換情報の取得を行うことにより、コントローラは変換前と変換後のフローの対応関係を把握することができ、ヘッダ変換装置を経由した場合でもフローの追跡が可能となる。

20

【0094】

<本発明の要点>

本発明では、オープンフローにおけるスイッチの一つである「OpenVSwitch」等が実装しているパケットのカプセル化機能を利用して、ヘッダ変換装置の前後でフローを識別する手法を開示する。

【0095】

本発明では、図4に示すように、パケットのヘッダ変換装置の前段にあるスイッチは、パケットが持つヘッダと同じヘッダを複製してパケットをカプセル化する。

【0096】

ここでは、パケットPKT0がヘッダH0、ペイロードP0を持っていることを、「パケットPKT0 = (H0、P0)」と表すことにする。

30

【0097】

スイッチS0は、パケットPKT0を受信すると、ヘッダH0を参照して同じヘッダH0を複製し、パケットPKT0をカプセル化してパケットPKT1を作成する。すなわち、パケットPKT1はヘッダH0、ペイロード(H0、P0)を持つことになるため、「パケットPKT1 = (H0、(H0、P0))」となる。

【0098】

このパケットPKT1がヘッダ変換装置T0を経由した後のパケットをPKT2とすると、パケットPKT2のヘッダはヘッダH0からヘッダH1に変換されることになるため、「パケットPKT2 = (H1、(H0、P0))」となる。

40

【0099】

次に、スイッチS1は、パケットPKT2を受信すると、ヘッダH1と、ペイロード内のヘッダH0を見比べる(比較参照する)。これにより、スイッチS1は、ヘッダ変換装置T0によって元のヘッダH0がヘッダH1に変換されたことを認識する。

【0100】

スイッチS1は、パケットPKT2のカプセル化を解除し、パケットPKT2からヘッダH1を除去して、ペイロード内のヘッダH0をヘッダH1に置き換えてパケットPKT3を生成する。すなわち、スイッチS1は、「パケットPKT2 = (H1、(H0、P0))」を「パケットPKT3 = (H1、P0)」に置き換える。

【0101】

50

その後、スイッチ S 1 は、パケット P K T 3 を送信する。

【 0 1 0 2 】

このようにすることで、ヘッダ変換装置 T 0 に改造を加えることなく、ヘッダ変換装置の前後でフローの対応関係を取得することができ、フローの追跡が可能となる。

【 0 1 0 3 】

以上のように、本発明では、スイッチにて（又はスイッチから要求を受けたコントローラにて）、パケットをカプセル化する。その際に、現在のヘッダと同じヘッダでカプセル化することにより、ネットワーク・アプライアンス通過後には、パケットに変換前と変換後の 2 種類のヘッダが付属していることになる。これを利用して、フローの追跡が可能となる。

10

【 0 1 0 4 】

<まとめ>

本発明では、オープンフローネットワークシステムのようなフローベースのネットワークにおいて、現在のパケットヘッダと同じヘッダを使ってパケットをカプセル化することにより、N A T や、N A P T 等のパケットヘッダを変換するようなネットワーク・アプライアンスを経由した場合にも、その前後での 2 つのフローの対応関係が把握でき、エンド・ツー・エンドでのフローの追跡が可能となる。

【 0 1 0 5 】

また、本発明では、カプセル化によって付与されたパケットヘッダのみが変換され、その変換前のパケットヘッダと比較することで変換情報が取得できるため、パケットヘッダの変換を行うネットワーク・アプライアンス装置を改造することなく、パケットヘッダの変換情報の取得や、フローの追跡が可能となる。

20

【 0 1 0 6 】

<付記>

上記の実施形態の一部又は全部は、以下の付記のように記載することも可能である。但し、実際には、以下の記載例に限定されない。

【 0 1 0 7 】

(付記 1)

パケット受信時に、フローを構成する各パケットを一律に制御するためのルールと動作が定義されたエントリが設定されたフローテーブルを検索して、エントリに定義されたルールに適合する受信パケットに対して、エントリに定義された動作を行う機能を持つスイッチと、

30

スイッチからパケットの問い合わせを受け、当該パケットを一律に制御するためのルールと動作が定義されたエントリを、スイッチのフローテーブルに設定する機能を持つコントローラと、

パケット受信時に、受信パケットの先頭のヘッダ情報の変換を行って転送する機能を持つヘッダ変換装置と、

スイッチが受信したパケットがカプセル化されたパケットかどうか確認し、カプセル化されたパケットでない場合、該パケットのヘッダ情報を複製し、複製したヘッダ情報を該受信パケットに付与してカプセル化してスイッチに渡し、カプセル化されたパケットである場合、該カプセル化されたパケットの先頭のヘッダ情報を除去し、残りの部分にあるヘッダ情報を該除去したヘッダ情報と置き換えてスイッチに渡し、該残りの部分にあるヘッダ情報と該除去したヘッダ情報との組をヘッダ変換情報としてコントローラに通知する機能を持つカプセル化モジュールと

40

を含む

ネットワークシステム。

【 0 1 0 8 】

(付記 2)

付記 1 に記載のネットワークシステムであって、

カプセル化モジュールは、スイッチがヘッダ変換装置の前段にある第 1 スイッチである

50

場合、第1スイッチの受信パケットのヘッダ情報である第1ヘッダを複製し、第1ヘッダを該受信パケットに付与して、ペイロードに第1ヘッダが二重に付与されている「第1ヘッダ - 第1ヘッダ - ペイロード」という構成のパケットを、カプセル化されたパケットとして第1スイッチ経由で転送する機能を持ち、

ヘッダ変換装置は、第1スイッチからのパケットを受信した際、カプセル化されたパケットの先頭の第1ヘッダを第2ヘッダに変換して、「第2ヘッダ - 第1ヘッダ - ペイロード」という構成のパケットにして転送する機能を持ち、

カプセル化モジュールは、スイッチがヘッダ変換装置の後段にある第2スイッチである場合、第2スイッチの受信パケットのヘッダ情報を参照し、第1ヘッダと第2ヘッダとの組をヘッダ変換情報としてコントローラに通知し、該受信パケットの先頭の第2ヘッダを除去し、残りの部分にある第1ヘッダを第2ヘッダと置き換えて、「第2ヘッダ - ペイロード」という構成のパケットにした上で、第2スイッチ経由で転送する機能を持つ

ネットワークシステム。

【0109】

(付記3)

付記1に記載のネットワークシステムであって、

カプセル化モジュールは、スイッチ及びコントローラのうち少なくとも一方の側に設けられ、スイッチの側に設けられた場合、スイッチから直接パケットを受け取り、コントローラの側に設けられた場合、スイッチからコントローラ経由でパケットを受け取る機能を持つ

ネットワークシステム。

【0110】

(付記4)

カプセル化モジュールとしての機能を持つ計算機であって、

受信パケットがカプセル化されたパケットかどうか確認する手段と、

受信パケットがカプセル化されたパケットでない場合、該パケットのヘッダ情報を複製する手段と、

複製したヘッダ情報を該受信パケットに付与してカプセル化する手段と、

受信パケットがカプセル化されたパケットである場合、該カプセル化されたパケットの先頭のヘッダ情報を除去する手段と、

該カプセル化されたパケットの残りの部分にあるヘッダ情報を該除去したヘッダ情報と置き換える手段と、

該残りの部分にあるヘッダ情報と該除去したヘッダ情報との組を基にヘッダ変換情報を生成する手段と

を具備する

計算機。

【0111】

<備考>

以上、本発明の実施形態を詳述してきたが、実際には、上記の実施形態に限られるものではなく、本発明の要旨を逸脱しない範囲の変更があっても本発明に含まれる。

【0112】

なお、本出願は、日本出願番号2011-031752に基づく優先権を主張するものであり、日本出願番号2011-031752における開示内容は引用により本出願に組み込まれる。

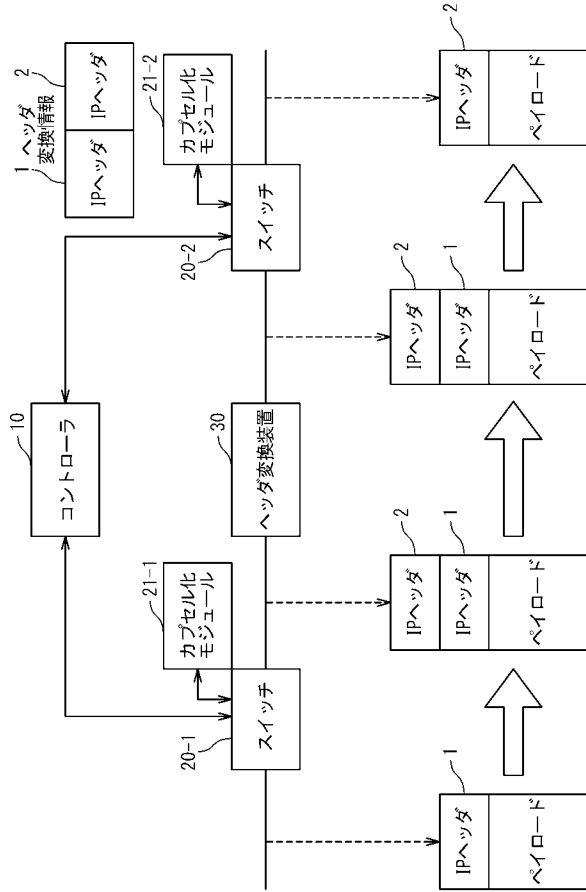
10

20

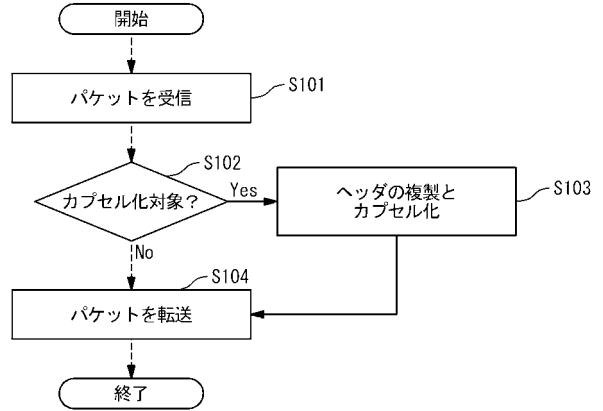
30

40

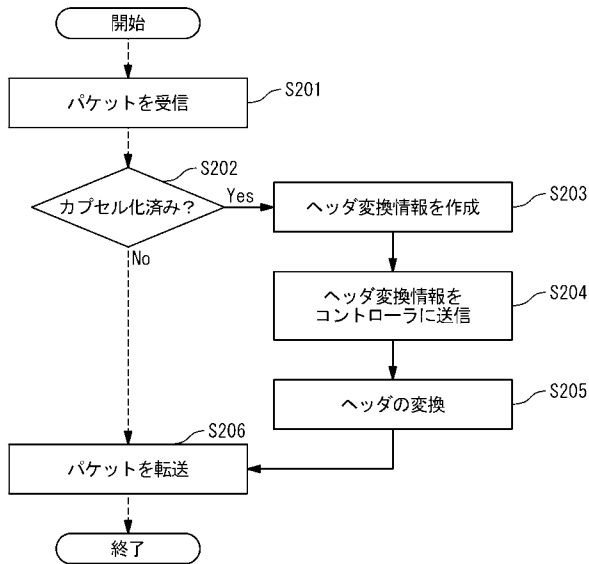
【図1】



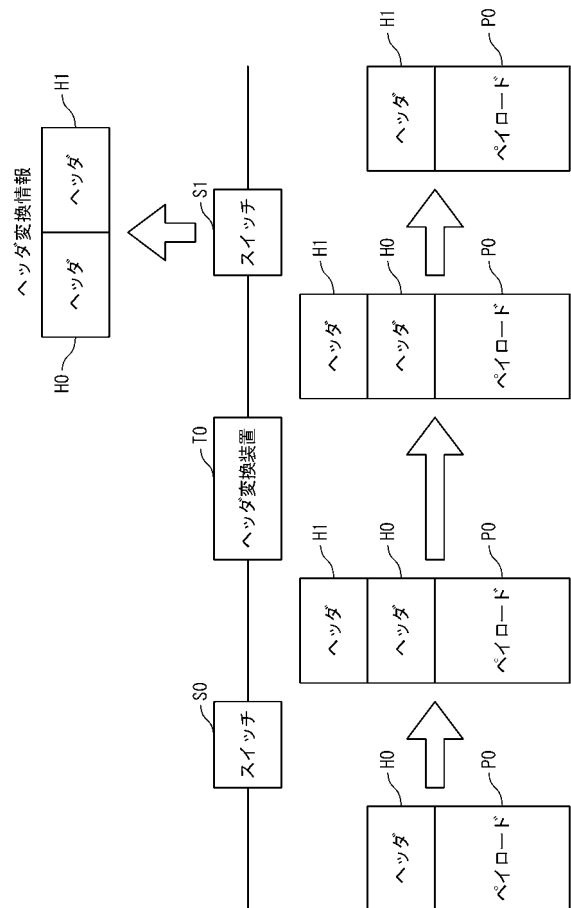
【図2】



【図3】



【図4】



フロントページの続き

(56)参考文献 特開2010-199669(JP,A)

横山 輝明、森島 直人、小川 晃通、宇夫 陽次郎、宇多 仁、江崎 浩、山口 英、NAT
ゲートウェイを通過するフローの網内での識別手法の提案、情報処理学会研究報告 Vol.2
001 No.59、社団法人情報処理学会、2001年 6月 8日、p.95-100

(58)調査した分野(Int.Cl., DB名)

H04L 12/749

H04L 12/717