

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5594552号
(P5594552)

(45) 発行日 平成26年9月24日(2014.9.24)

(24) 登録日 平成26年8月15日(2014.8.15)

(51) Int.Cl. F I
H O 4 L 12/717 (2013.01) H O 4 L 12/717

請求項の数 9 (全 19 頁)

(21) 出願番号	特願2014-92791 (P2014-92791)	(73) 特許権者	00004237
(22) 出願日	平成26年4月28日(2014.4.28)		日本電気株式会社
(62) 分割の表示	特願2012-552660 (P2012-552660) の分割		東京都港区芝五丁目7番1号
原出願日	平成23年12月27日(2011.12.27)	(74) 代理人	100102864
(65) 公開番号	特開2014-168286 (P2014-168286A)		弁理士 工藤 実
(43) 公開日	平成26年9月11日(2014.9.11)	(72) 発明者	高島 正徳
審査請求日	平成26年4月28日(2014.4.28)		東京都港区芝五丁目7番1号 日本電気株 式会社内
(31) 優先権主張番号	特願2011-5137 (P2011-5137)	(72) 発明者	加瀬 知博
(32) 優先日	平成23年1月13日(2011.1.13)		東京都港区芝五丁目7番1号 日本電気株 式会社内
(33) 優先権主張国	日本国(JP)		
早期審査対象出願		審査官	大石 博見

最終頁に続く

(54) 【発明の名称】 ネットワークシステム、及び経路制御方法

(57) 【特許請求の範囲】

【請求項1】

パケットが属する仮想ネットワークに対応する識別子に基づいてパケットを転送するための第一のエントリをコアスイッチに送信する第一の手段と、

パケットを識別するための識別ルールと、当該識別ルールで識別される受信パケットが属する仮想ネットワークに対応する前記識別子を当該受信パケットに付与する指示とを含む第二のエントリを、前記コアスイッチで構成されるネットワークの端点に設けられたエッジスイッチに対して予め送信可能な第二の手段と

を含むネットワークシステム。

【請求項2】

前記第二の手段は、前記識別子が付与された受信パケットを復元する指示を含むエントリを、前記エッジスイッチに対して予め送信する

ことを特徴とする請求項1のネットワークシステム。

【請求項3】

前記第二の手段は、前記エッジスイッチに仮想マシンが接続されたことに応じて、パケットを識別するための識別ルールと、当該識別ルールで識別される受信パケットが属する仮想ネットワークに対応する前記識別子を当該受信パケットに付与する指示とを含む第二のエントリを、前記エッジスイッチに対して予め送信する

ことを特徴とする請求項1または2のネットワークシステム。

【請求項4】

パケットが属する仮想ネットワークに対応する識別子に基づいてパケットを転送するための第一のエントリをコアスイッチに送信する第一の手段と、

パケットを識別するための識別ルールと、当該識別ルールで識別される受信パケットが属する仮想ネットワークに対応する前記識別子を当該受信パケットに付与する指示とを含む第二のエントリを、前記コアスイッチで構成されるネットワークの端点に設けられたエッジスイッチに対して予め送信可能な第二の手段と

を含むコントローラ。

【請求項 5】

前記第二の手段は、前記識別子が付与された受信パケットを復元する指示を含むエントリを、前記エッジスイッチに対して予め送信する

ことを特徴とする請求項 4 のコントローラ。

【請求項 6】

前記第二の手段は、前記エッジスイッチに仮想マシンが接続されたことに応じて、パケットを識別するための識別ルールと、当該識別ルールで識別される受信パケットが属する仮想ネットワークに対応する前記識別子を当該受信パケットに付与する指示とを含む第二のエントリを、前記エッジスイッチに対して予め送信する

ことを特徴とする請求項 4 または 5 のコントローラ。

【請求項 7】

パケットが属する仮想ネットワークに対応する識別子に基づいてパケットを転送するための第一のエントリをコアスイッチに送信し、

パケットを識別するための識別ルールと、当該識別ルールで識別される受信パケットが属する仮想ネットワークに対応する前記識別子を当該受信パケットに付与する指示とを含む第二のエントリを、前記コアスイッチで構成されるネットワークの端点に設けられたエッジスイッチに対して予め送信する

ことを特徴とする通信方法。

【請求項 8】

前記識別子が付与された受信パケットを復元する指示を含むエントリを、前記エッジスイッチに対して予め送信する

ことを特徴とする請求項 7 の通信方法。

【請求項 9】

前記エッジスイッチに仮想マシンが接続されたことに応じて、パケットを識別するための識別ルールと、当該識別ルールで識別される受信パケットが属する仮想ネットワークに対応する前記識別子を当該受信パケットに付与する指示とを含む第二のエントリを、前記エッジスイッチに対して予め送信する

ことを特徴とする請求項 7 または 8 の通信方法。

【発明の詳細な説明】

【技術分野】

【0001】

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の経路制御方法に関する。

【背景技術】

【0002】

外部のコントローラ（コントロールプレーン）からスイッチや端末等（ユーザプレーン）を制御する方式を C U（C：コントロールプレーン / U：ユーザプレーン）分離型アーキテクチャと呼ぶ。このような C U 分離型アーキテクチャに基づく構成のネットワークを C U 分離型ネットワークと呼ぶ。

【0003】

C U 分離型ネットワークの一例として、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（Open Flow）技術を利用したオープンフローネットワークが挙げられる。なお、オープンフローネットワークは一例に過ぎない。

10

20

30

40

50

【0004】

【オープンフローネットワークの説明】

オープンフローネットワークでは、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。

【0005】

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処理内容(アクション)を定義したエントリ(entry)が登録されたテーブルである。パケットは、フレームと読み替えても良い。ルールに適合するパケット群(パケット系列)をフロー(Flow)と呼ぶ。

10

【0006】

フローのルールは、パケットの各プロトコル階層のヘッダ領域(フィールド)に含まれる宛先アドレス(DA: Destination Address)、送信元アドレス(SA: Source Address)、宛先ポート(DP: Destination Port)、送信元ポート(SP: 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、フローのルールとして使用可能である。

20

【0007】

フローのアクションは、通常、所定の転送先へのパケット転送である。無論、フローのアクションとして、パケット破棄を指定することも可能である。

【0008】

オープンフローネットワークでは、通常、スイッチは、該当するエントリがないパケットを受信した場合、コントローラに対して、当該パケットについての問い合わせ(エントリ要求)を送信する。通常、スイッチは、当該パケットについての問い合わせとして、当該パケットをコントローラに転送する。

【0009】

オープンフローネットワークでは、通常、コントローラは、セキュアチャネル(Secure Channel)接続で、管理下にあるスイッチと接続されている。コントローラは、管理下にあるスイッチからパケットについての問い合わせを受けた場合、当該パケット群(フロー)の経路(パス)を計算し、この経路に基づき、当該スイッチのフローテーブルに、「当該パケット群(フロー)を所定の転送先へ転送する」という旨のエントリを登録する。このとき、コントローラは、当該エントリをフローテーブルに登録するための制御メッセージを、スイッチに対して送信する。

30

【0010】

オープンフロー技術の詳細については、非特許文献1、2に記載されている。

【先行技術文献】

【非特許文献】

40

【0011】

【非特許文献1】 “The OpenFlow Switch Consortium” <<http://www.openflowswitch.org/>>

【非特許文献2】 “OpenFlow Switch Specification Version 1.0.0 (Wire Protocol 0x01) December 31, 2009” <<http://www.openflowswitch.org/documents/openflow-spec-v1.0.0.pdf>>

【発明の概要】

【0012】

オープンフロー技術において、スイッチのフローテーブルにエントリを登録する方式は

50

、大きく「Proactive型」と、「Reactive型」の2つの方式に分けられる。

【0013】

「Proactive型」では、コントローラが「事前に」（データ通信が始まる前に）所定のケット群（フロー）の経路（パス）を計算し、スイッチのフローテーブルにエントリを登録する。すなわち、ここでいう「Proactive型」とは、コントローラが自発的に行う「事前のエントリ登録」を指す。

【0014】

「Reactive型」では、コントローラが「スイッチから1stパケット（該当エントリがない新規のケット）についての問い合わせを受けた際に」当該ケット群（フロー）の経路を計算し、スイッチのフローテーブルにエントリを登録する。すなわち、ここでいう「Reactive型」とは、実際のデータ通信時に、コントローラがスイッチからの問い合わせに応じて行う「リアルタイムのエントリ登録」を指す。

10

【0015】

オープンフローネットワークでは、基本的に、コントローラがスイッチから1stパケットについての問い合わせを受けた際に当該受信ケットに関するエントリを登録する「Reactive型」が中心となっている。

【0016】

しかし、実際のハードウェア（HW）では、フローテーブルの処理頻度を軽減し性能の問題を解決するためには、「Proactive型」が好適であると考えられる。例えば、大量の1stパケットがコントローラに到着しても処理し切れるようにするためには、Proactive型の方が好適であると考えられる。但し、実際には、完全なProactive型にするとエントリ数が膨大になると考えられるため、一部をReactive型にすることにより、エントリ数の制約から逃れるといったことも考えられる。

20

【0017】

また、「Proactive型」を用いれば、通信開始前にフローを定義できるため、Nimda等のウィルスによる大量フロー発生問題や、不明なケットによる不正アクセス等が回避可能になると考えられる。

【0018】

従って、オープンフローネットワークにおいて、「Proactive型」を実現するための具体的な方式が求められる。

30

【0019】

本発明に係るネットワークシステムは、複数のスイッチと、コントローラとを含む。複数のスイッチの各々は、自身のフローテーブルに登録されたエントリのルールに適合する受信ケットに対し、エントリに定義されたアクションに従った動作を行う。コントローラは、複数のスイッチによって形成されるネットワークの物理トポロジに応じて計算された経路に固有の識別子をルールとし、所定の出力ポートからの出力をアクションとしたエントリを、複数のスイッチ間において通信が開始される前に、各スイッチに登録する。

【0020】

本発明に係る経路制御方法では、複数のスイッチの各々において、自身のフローテーブルに登録されたエントリのルールに適合する受信ケットに対し、エントリに定義されたアクションに従った動作を行う。また、コントローラにおいて、複数のスイッチによって形成されるネットワークの物理トポロジに応じて計算された経路に固有の識別子をルールとし、所定の出力ポートからの出力をアクションとしたエントリを、複数のスイッチ間において通信が開始される前に、各スイッチに登録する。

40

【0021】

本発明に係るプログラムは、上記の経路制御方法におけるコントローラの動作を、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。

【0022】

50

これにより、オープンフローネットワークにおいて、「Proactive型」を実現し、ハードウェア（HW）の性能の問題を解決することができる。

【図面の簡単な説明】

【0023】

【図1】本発明に係るネットワークシステムの構成例を示す図である。

【図2】トポロジ検知の際の処理を説明するための図である。

【図3】ステーション検知（ARPリクエスト（要求）利用）の際の処理を説明するための図である。

【図4】ステーション検知（ARPリプライ（応答）利用）の際の処理を説明するための図である。

10

【図5】エン트리登録完了後のデータ通信の際の処理を説明するための図である。

【図6】コントローラへのパケットの問い合わせの際の処理を説明するための図である。

【発明を実施するための形態】

【0024】

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

【0025】

<第1実施形態>

以下に、本発明の第1実施形態について添付図面を参照して説明する。

20

【0026】

[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、スイッチ10（10-i、 $i = 1 \sim n$ ：nはスイッチ台数）と、コントローラ20を含む。

【0027】

スイッチ10（10-i、 $i = 1 \sim n$ ）、及びコントローラ20は、オープンフローネットワークを形成する。スイッチ10（10-i、 $i = 1 \sim n$ ）は、オープンフローネットワークにおけるノード（node）である。

【0028】

[スイッチ]

スイッチ10（10-i、 $i = 1 \sim n$ ）は、内部にフローテーブルを持ち、コントローラ20からフローテーブルに登録されたエントリに従って、パケットを転送する。

30

【0029】

[コントローラ]

コントローラ20は、トポロジ（topology：ネットワークの接続形態）検知を行い、ネットワークを構成するスイッチ10（10-i、 $i = 1 \sim n$ ）を検知し、フロー毎の経路（パス）を計算する。これにより、コントローラ20は、ネットワークを構成する全てのスイッチの識別情報（スイッチID、MACアドレス等）、及び各スイッチの接続形態を認識し、各スイッチの次段のスイッチを特定する。

【0030】

40

ここでは、コントローラ20は、通信が開始される前に、各スイッチのスイッチID（64ビット）と独自に定義したノードID（16ビット）とを1対1で対応付ける。但し、ビット数は一例に過ぎない。すなわち、コントローラ20は、各スイッチにノードIDを割り当てる。また、コントローラ20は、端末接続可能なエッジスイッチ間の経路を計算し、当該経路上の中継用スイッチであるコアスイッチ（Core）の各々のフローテーブルに、「受信パケットの宛先情報の領域の少なくとも一部に所定のノードIDが記載されていれば、当該受信パケットを（所定の出力ポートから）次段のスイッチに転送する」という旨のCoreエントリ（中継用エントリ）を登録する。すなわち、コアスイッチは、受信パケットの宛先情報の領域に記載されたノードIDをマッチ条件（ルール）として、転送の可否を判断する。無論、コントローラ20は、ノードIDに加えて、受信パケッ

50

トの宛先情報の領域に記載された他の情報もマッチ条件（ルール）として指定するようにしても良い。

【0031】

なお、実際には、コントローラ20は、コアスイッチ（Core）の各々のフローテーブルに、「（ノードIDに関わらず、）当該受信パケットを（所定の出力ポートから）次段のスイッチに転送する」という旨のCoreエントリを登録することも可能である。この場合、コアスイッチ（Core）は、無条件で、受信パケットを次段のスイッチに転送する。

【0032】

また、コントローラ20は、ステーション検知（端末検知）を行い、端末30（30-j、j=1~m:mは端末台数）を検知し、端末の宛先情報（MACアドレス等）、及び接続形態を認識し、端末とユーザIDとを1対1で対応付ける。すなわち、コントローラ20は、各端末にユーザIDを割り当てる。このとき、コントローラ20は、端末30（30-j、j=1~m）が接続されたエッジスイッチも検知する。

10

【0033】

また、コントローラ20は、通信が開始される前に、エッジスイッチのフローテーブルに、「受信パケットの宛先情報の領域の少なくとも一部に、当該エッジスイッチのノードIDと、配下の端末のユーザIDが記載されていれば、当該受信パケットの宛先情報を当該端末宛に戻し、当該受信パケットを当該端末に転送する」という旨のEgressエントリ（出力用エントリ）を登録する。

20

【0034】

ここで、配下の端末のユーザIDをマッチ条件（ルール）としたのは、配下の端末が複数存在する場合もあるからである。また、エッジスイッチのノードIDと端末のユーザIDとの組み合わせをマッチ条件（ルール）としているため、エッジスイッチ間で重複したユーザIDを使用することが可能になる。但し、同一のエッジスイッチの配下の端末の各々に対しては、重複したユーザIDを使用することはできない。

【0035】

また、コントローラ20は、入力側エッジスイッチ（Ingress）のフローテーブルに、「所定のパケットを受信した際、宛先情報を検索のキーとして使用し、受信パケットの宛先情報の領域の少なくとも一部に、出力側エッジスイッチ（Egress）のノードIDと、宛先となる端末のユーザIDとを記載して次段のスイッチに転送する」という旨のIngressエントリ（入力用エントリ）を登録する。ここで、上記の「所定のパケット」は、「所定のマッチ条件（ルール）に適合するパケット」と読み替えても良い。本発明では、入力側エッジスイッチ（Ingress）が最初にフローを特定するため、Ingressエントリには、通常のオープンフローと同様のパケットマッチングルールと、適合したパケットに対する上記のアクションが定義されることになる。

30

【0036】

[Ingressエントリ登録のタイミング]

なお、コントローラ20が入力側エッジスイッチのフローテーブルにエントリを登録するタイミングとしては、「通信が開始される前」（事前登録）と、「実際の通信時」（リアルタイム登録）の2つが考えられる。

40

【0037】

「通信が開始される前」（事前登録）の場合、コントローラ20は、通信が開始される前に、予め所定のパケットの転送先となる端末（宛先候補端末）を決定する。そして、コントローラ20は、通信が開始される前に、入力側エッジスイッチとなり得るエッジスイッチのフローテーブルに、「所定のパケットを受信した際、宛先情報を検索のキーとして使用し、受信パケットの宛先情報の領域の少なくとも一部に、当該宛先候補端末が接続されたエッジスイッチのノードIDと、当該宛先候補端末のユーザIDとを記載して次段のスイッチに転送する」という旨のIngressエントリを登録する。本実施形態では、この事例について説明する。

50

【0038】

「実際の通信時」(リアルタイム登録)の場合、コントローラ20は、送信元の端末からパケットを受信した入力側エッジスイッチから当該受信パケットについての問い合わせを受け取ると、当該受信パケット群(フロー)の経路を計算し、この経路に基づき、当該入力側エッジスイッチのフローテーブルに、「所定のパケットを受信した際、宛先情報を検索のキーとして使用し、受信パケットの宛先情報の領域の少なくとも一部に、当該宛先候補端末が接続されたエッジスイッチのノードIDと、当該宛先候補端末のユーザIDとを記載して次段のスイッチに転送する」という旨のIngressエントリを登録する。この事例については、第2実施形態にて説明する。

【0039】

[経路の指定]

更に、各スイッチの次段のスイッチが複数存在する場合(経路が複数の場合)、コントローラ20は、経路毎に冗長IDを定義する。次段のスイッチの各々はそれぞれの経路上に存在するため、次段のスイッチの各々と冗長IDとが対応付けられる。コントローラ20は、コアスイッチ(Core)のフローテーブルに、「受信パケットの宛先情報の領域の少なくとも一部に(所定のノードIDと)冗長IDが記載されていれば、当該受信パケットを当該冗長IDに対応する次段のスイッチに転送する」という旨のCoreエントリを登録する。また、コントローラ20は、入力側エッジスイッチのフローテーブルに、「所定のパケットを受信した際、宛先情報を検索のキーとして使用し、受信パケットの宛先情報の領域の少なくとも一部に、出力側エッジスイッチのノードIDと、冗長IDと、宛先となる端末のユーザIDとを記載して次段のスイッチに転送する」という旨のIngressエントリを登録する。冗長IDは、出力側エッジスイッチのノードIDの一部でも良い。例えば、ノードIDの領域の先頭又は末尾の数ビット(bit)を、冗長IDの領域として使用しても良い。

【0040】

[ハードウェアの例示]

スイッチ10(10-i、 $i = 1 \sim n$)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(firewall)、ロードバランサ(load balancer: 負荷分散装置)、帯域制御装置(packet shaper)、セキュリティ監視制御装置(SCADA: Supervisory Control And Data Acquisition)、ゲートキーパー(gatekeeper)、基地局(base station)、アクセスポイント(AP: Access Point)、通信衛星(CS: Communication Satellite)、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。また、スイッチ10(10-i、 $i = 1 \sim n$)は、物理マシン上に構築された仮想スイッチでも良い。

【0041】

コントローラ20及び端末30(30-j、 $j = 1 \sim m$)の各々の例として、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20及び端末30(30-j、 $j = 1 \sim m$)は、物理マシン上に構築された仮想マシン(VM)でも良い。

【0042】

なお、端末30(30-j、 $j = 1 \sim m$)の各々は、携帯電話機、スマートフォン、スマートブック、カーナビ(カーナビゲーションシステム)、携帯型ゲーム機、家庭用ゲーム機、携帯型音楽プレイヤー、ハンディターミナル、ガジェット(電子機器)、双方向テレビ、デジタルチューナー、デジタルレコーダー、情報家電(information home appliance)、OA(Office Automation)機器等でも良い。また、端末30(30-j、 $j = 1 \sim m$)の各々は、車両や船舶、航空機等の移動体に搭載されていても良い。

10

20

30

40

50

【0043】

図示しないが、スイッチ10(10-i、i=1~n)の各々、コントローラ20、及び端末30(30-j、j=1~m)の各々は、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。

【0044】

上記のプロセッサの例として、CPU(Central Processing Unit)、マイクロプロセッサ(microprocessor)、マイクロコントローラ、或いは、専用の機能を有する半導体集積回路(IC: Integrated Circuit)等が考えられる。

10

【0045】

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。

【0046】

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。従って、電子機器等に搭載される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。

20

【0047】

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。

【0048】

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(Infrared 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。

30

【0049】

スイッチ10(10-i、i=1~n)の各々、コントローラ20、及び端末30(30-j、j=1~m)の各々の内部の構成要素は、モジュール(module)、コンポーネント(component)、或いは専用デバイス、又はこれらの起動(呼出)プログラムでも良い。

40

【0050】

但し、実際には、これらの例に限定されない。

【0051】

[トポロジ検知]

図2を参照して、トポロジ検知の際の処理について説明する。

【0052】

コントローラ20は、LLDP(Link Layer Discovery Protocol)によって、ネットワークの物理トポロジを検知する。LLDPとは、制御フレームを送受信することにより定期的に隣接する装置情報を収集するためのプロトコルで

50

ある。

【0053】

事前に、管理者端末等から、スイッチ10(10-i、i=1~n)の各々に、内向き/外向きのコンフィグ(Configuration:設定情報)を設定する。或いは、コントローラ20が、セキュアチャネル(Secure Channel)接続で、管理下にあるスイッチ10(10-i、i=1~n)の各々に、内向き/外向きのコンフィグを投入するようにしても良い。

【0054】

「内向きのコンフィグ」とは、ネットワークの内側との通信用の設定情報である。「外向きのコンフィグ」とは、ネットワークの外側との通信用の設定情報である。

10

【0055】

スイッチ10(10-i、i=1~n)の各々は、内向き/外向きのコンフィグを、ポート(Port)のステータス情報(PortStat)として格納する。初期設定(default)では、スイッチ10(10-i、i=1~n)の各々は、内向きのコンフィグを、ポートのステータス情報(PortStat)として格納している。

【0056】

コントローラ20は、事前にスイッチ10(10-i、i=1~n)に内向き/外向きの設定があるため、トポロジを検知するスピードが速くなる。

【0057】

コントローラ20は、トポロジを検知して、スイッチ10(10-i、i=1~n)の各々が持つポートのそれぞれについて、ポートのステータス情報(PortStat)を収集して、スイッチ10(10-i、i=1~n)の各々が持つポートが内向きか外向きかを判断する。

20

【0058】

コントローラ20は、ポートのステータス情報において、内向きと明示的な設定のあるポートを「内向きポート」(Inside Port)と認識する。また、コントローラ20は、ポートのステータス情報において、外向きと明示的な設定のあるポートを「外向きポート」(Outside Port)と認識する。

【0059】

コントローラ20は、内向きポート、及び外向きと明示的な設定のないポート(未設定ポート等)に対して、LLDPの制御フレームを送信する。そして、コントローラ20は、LLDPの制御フレームに対する応答から、ネットワークの物理トポロジを検知し、トポロジ情報を作成する。

30

【0060】

このとき、コントローラ20は、管理下にあるスイッチ10(10-i、i=1~n)の各々のスイッチIDを取得し、各スイッチ10(10-i、i=1~n)のスイッチIDと、ノードIDとを対応付ける。このとき、ノードIDを対応付けるのは、外向きポートが存在するスイッチ(エッジスイッチ)のみでも良い。ここでは、スイッチIDとして、「DPID:#1~#6」を示している。なお、実際には、「DPID:#1~#6」をそのままノードIDとして使用することも可能である。

40

【0061】

また、ノードIDは、サブノードIDと、冗長IDを含むものとする。サブノードIDは、スイッチを特定するためのノードIDの実質的な本体である。サブノードIDは、単体でスイッチを特定することが可能な識別情報でも良い。或いは、サブノードIDは、冗長IDと組み合わせることで、スイッチを一意に特定するためのノードIDを形成するものでも良い。冗長IDは、経路を特定するための識別情報である。各スイッチ10(10-i、i=1~n)は、この冗長IDに基づいて、次段のスイッチに転送するためのポートを決定し、当該ポートから受信パケットを送出するようにしても良い。なお、実際には、冗長IDは、ノードIDとの関連性が保たれており、冗長IDとして特定可能であれば、別の領域に格納されていても良い。

50

【0062】

コントローラ20は、外向きポートが存在するスイッチ（エッジスイッチ）間の経路を計算し、当該経路上にあるコアスイッチ（Core）のフローテーブルに、「受信パケットの宛先情報の領域の少なくとも一部に所定のノードID（外向きポートが存在するエッジスイッチのノードID）が記載されていれば、当該受信パケットを当該経路上の次段のスイッチに転送する」という旨のCoreエントリを登録する。すなわち、コアスイッチは、受信パケットの宛先情報の領域に記載されたノードIDをマッチ条件（ルール）として、転送の可否を判断する。無論、コントローラ20は、ノードIDに加えて、受信パケットの宛先情報の領域に記載された他の情報（VTNID、ユーザID等）もマッチ条件（ルール）として指定するようにしても良い。

10

【0063】

なお、実際には、コントローラ20は、コアスイッチ（Core）の各々のフローテーブルに、「（ノードIDに関わらず、）当該受信パケットを（所定の出力ポートから）次段のスイッチに転送する」という旨のCoreエントリを登録することも可能である。この場合、コアスイッチ（Core）は、無条件で、受信パケットを次段のスイッチに転送する。受信パケットの転送の可否については、入力側エッジスイッチ（Ingress）や出力側エッジスイッチ（Egress）で判断することになる。

【0064】

ここでは、コントローラ20は、外向きポートが存在する全てのスイッチ間の経路を計算し、当該経路上にあるコアスイッチ（Core）のフローテーブルに、上記のようにCoreエントリを登録する。

20

【0065】

[エントリの構成例]

以下に、エントリの構成例について説明する。

エントリは、「Port」、「DA」（Destination Address）、「SA」（Source Address）、「OPort」（Output Port）、「Mod」（Modify）等のデータ格納領域を含む。

【0066】

「Port」は、受信パケットの入力ポートを示す情報の格納領域である。「DA」は、受信パケットの宛先情報の格納領域である。「SA」は、受信パケットの送信元情報の格納領域である。「OPort」は、受信パケットの出力ポートを示す情報の格納領域である。「Mod」は、受信パケットに対して行う処理を定義した情報の格納領域である。

30

【0067】

「Port」、「DA」、及び「SA」は、マッチ条件（ルール）に該当する。また、「OPort」及び「Mod」は、処理内容（アクション）に該当する。

【0068】

「DA」に格納される「ID群」とは、「ノードID」、「VTNID」、「ユーザID」等の情報である。「ノードID」は、出力側エッジスイッチとなるスイッチ（外向きポートが存在するノード）を特定するための識別情報の格納領域である。「VTNID」は、VTN（Virtual Tenant Network）等の、外向きポートが存在するスイッチ間の経路を通過するパケット群（フロー）が属するVN（Virtual Network）の識別情報の格納領域である。「ユーザID」は、宛先となる端末（外向きポートが存在するスイッチに接続された/接続され得る端末）を特定するためのユーザIDの識別情報である。端末とユーザIDとの対応付けは、以下に説明する「ステーション検知」において行う。

40

【0069】

[ステーション検知]

図3、図4を参照して、ステーション検知の際の処理について説明する。

【0070】

コントローラ20は、端末がアドレス解決のために送信するARP（Address

50

Resolution Protocol)の制御フレームを利用して、ステーション検知を行う。

【0071】

なお、ARPの制御フレームは一例に過ぎない。例えば、DHCP(Dynamic Host Configuration Protocol)等の制御フレームでも良い。また、実際には、制御フレームに限定されない。

【0072】

ここでは、端末30-1を「端末A」とし、端末30-2を「端末B」とする。

【0073】

(1) ARP_Req(ARPリクエスト(Request:要求))を利用

10

図3に示すように、端末Aから端末Bに対して通信する際に、端末BのMACアドレスがわからず、端末BのIPアドレスだけがわかっている場合、端末Aは、端末Bのアドレスを解決するためのARP_Req(ARPリクエスト)をブロードキャストで送信する。

【0074】

端末Aが接続されたエッジスイッチ10-1は、セキュアチャネル(Secure Channel)接続で、コントローラ20宛にARP_Req(ARPリクエスト)を転送する。このとき、コントローラ20は、ARPプロキシ(ARP Proxy)として動作する。

【0075】

20

コントローラ20は、端末Aが接続されたエッジスイッチ10-1からARP_Req(ARPリクエスト)を受信した際、ARP_Req(ARPリクエスト)の送信元情報から端末AのMACアドレス(及びIPアドレス)を取得し、端末AにユーザIDを割り当てる。すなわち、端末AのMACアドレス(及びIPアドレス)と、ユーザIDとを対応付ける。

【0076】

コントローラ20は、端末Aが接続されたエッジスイッチ10-1のフローテーブルに、「受信パケットの宛先情報の領域の少なくとも一部に(当該エッジスイッチのノードIDと)配下の端末AのユーザIDが記載されていれば、当該受信パケットの宛先情報を端末AのMACアドレス宛に戻し、当該受信パケットを端末AのMACアドレス宛に転送する」という旨のEgressエントリを登録する。

30

【0077】

コントローラ20は、ARPプロキシ(ARP Proxy)として、ターゲット(Target)となる端末Bのアドレス解決をするため、セキュアチャネル(Secure Channel)接続で、管理下にあるスイッチ10(10-i、i=1~n)の各々に対して、ARP_Req(ARPリクエスト)をブロードキャストで送信する。このとき、ARP_Req(ARPリクエスト)の送信元MACアドレスは、端末AのMACアドレスである。

【0078】

端末Bが接続されたエッジスイッチ10-6は、ブロードキャストで送信されたARP_Req(ARPリクエスト)を端末Bに転送する。

40

【0079】

ここでは、説明の簡略化のため、端末Bのみを宛先端末として説明しているが、他の宛先端末が接続されたエッジスイッチについても、上記と同様の処理を行う。

【0080】

(2) ARP_Rep(ARプリプライ(Reply:応答))を利用

図4に示すように、端末Bは、ARP_Req(ARPリクエスト)に対する応答として、端末A宛にARP_Rep(ARプリプライ)を送信する。

【0081】

端末Bが接続されたエッジスイッチ10-6は、セキュアチャネル(Secure C

50

hannel)接続で、コントローラ20宛にARP__Rep(ARPリプライ)を転送する。このとき、コントローラ20は、ARPプロキシ(ARP Proxy)として動作する。

【0082】

コントローラ20は、端末Bが接続されたエッジスイッチ10-6からARP__Rep(ARPリプライ)を受信した際、ARP__Rep(ARPリプライ)の送信元情報から端末BのMACアドレス(及びIPアドレス)を取得し、端末BにユーザIDを割り当てる。すなわち、端末BのMACアドレス(及びIPアドレス)と、ユーザIDとを対応付ける。

【0083】

コントローラ20は、端末Bが接続されたエッジスイッチ10-6のフローテーブルに、「受信パケットの宛先情報の領域の少なくとも一部に(当該エッジスイッチのノードIDと)配下の端末BのユーザIDが記載されていれば、当該受信パケットの宛先情報を端末BのMACアドレス宛に戻し、当該受信パケットを端末BのMACアドレス宛に転送する」という旨のEgressエントリを登録する。

【0084】

この時点で、コントローラ20は、端末Aと端末Bとの間の通信が可能であると判断し、端末Aが接続されたエッジスイッチ10-1のフローテーブルに、「端末B宛のパケットを受信した場合、当該受信パケットの宛先情報の領域の少なくとも一部に、ID群(端末Bが接続されたエッジスイッチのノードIDと、当該フローのVTNIDと、端末BのユーザID)を記載して次段のスイッチに転送する」という旨のIngressエントリを登録することも可能である。このエントリ登録方式は「Proactive型」(事前にIngressエントリを登録する方式)である。

【0085】

コントローラ20は、ARPプロキシ(ARP Proxy)として、端末AにARP__Rep(ARPリプライ)を転送するため、セキュアチャネル(Secure Channel)接続で、端末Aが接続されたエッジスイッチ10-1に対して、ARP__Rep(ARPリプライ)を送信する。このとき、ARP__Rep(ARPリプライ)の送信元MACアドレスは、端末BのMACアドレスである。

【0086】

端末Aが接続されたエッジスイッチ10-1は、コントローラ20から受信したARP__Rep(ARPリプライ)を端末Aに転送する。

【0087】

端末Aは、ARP__Req(ARPリクエスト)に対する応答として受信したARP__Rep(ARPリプライ)から、端末BのMACアドレスを取得する。

【0088】

[エントリ登録完了後のデータ通信]

図5を参照して、エントリ登録完了後のスイッチ間でのIPパケット等の通信の際の処理について説明する。

【0089】

この時点では、スイッチ間の全てのスイッチに対して、必要なエントリ登録は完了しているものとする。すなわち、コントローラ20の作業は完了している。

【0090】

まず、端末Aは、宛先となる端末BのMACアドレス(及びIPアドレス)を、パケットの宛先情報の領域に記載し、端末B宛のパケットを送信する。

【0091】

端末Aが接続されたエッジスイッチ10-1は、入力ポート1で端末B宛のパケットを受信すると、自身のフローテーブルに当該受信パケットと適合するエントリが登録されているか確認する。

【0092】

10

20

30

40

50

端末Aが接続されたエッジスイッチ10-1は、自身のフローテーブルに、「端末B宛の packets を受信した場合、当該受信パケットの宛先情報の領域の少なくとも一部に、ID群（端末Bが接続されたエッジスイッチ10-6のノードIDと、当該フローのVTN IDと、端末BのユーザID）を記載して次段のスイッチに転送する」という旨のIngress エントリが登録されているため、受信パケットの宛先情報の領域の少なくとも一部に、当該ID群を記載して次段のスイッチに転送する。ここでは、エッジスイッチ10-1は、受信パケットの宛先情報に記載されている端末BのMACアドレスをID群に変更（宛先情報にID群を上書き）し、当該変更後のパケット（以下、ID化パケット）を、出力ポート2に転送し、出力ポート2から次段のスイッチ10-2に転送する。

【0093】

エッジスイッチ10-2は、入力ポート3でID化パケットを受信すると、自身のフローテーブルに当該ID化パケットと適合するエントリが登録されているか確認する。

【0094】

エッジスイッチ10-2は、自身のフローテーブルに、「受信パケットの宛先情報の領域の少なくとも一部にエッジスイッチ10-6のノードIDが記載されていれば、当該受信パケットを（所定の出力ポートから）次段のスイッチに転送する」という旨のCore エントリが登録されているため、当該ID化パケットを、出力ポート4に転送し、出力ポート4から次段のスイッチ10-3に転送する。なお、エッジスイッチ10-2は、出力ポート4と出力ポート5のいずれに転送するかは、当該Core エントリの「OPort」の値を基に決定する。

【0095】

このとき、エッジスイッチ10-2は、当該ID化パケットに冗長IDが含まれている場合、冗長IDの値を検索キーとして、該当するCore エントリを検索する。

【0096】

エッジスイッチ10-3は、入力ポート10でID化パケットを受信すると、自身のフローテーブルに当該ID化パケットと適合するエントリが登録されているか確認する。

【0097】

エッジスイッチ10-3は、自身のフローテーブルに、「受信パケットの宛先情報の領域の少なくとも一部にエッジスイッチ10-6のノードIDが記載されていれば、当該受信パケットを（所定の出力ポートから）次段のスイッチに転送する」という旨のCore エントリが登録されているため、当該ID化パケットを、出力ポート12に転送し、出力ポート12から次段のスイッチ10-6に転送する。

【0098】

エッジスイッチ10-6は、入力ポート13でID化パケットを受信すると、自身のフローテーブルに当該ID化パケットと適合するエントリが登録されているか確認する。

【0099】

端末Bが接続されたエッジスイッチ10-6は、自身のフローテーブルに、「受信パケットの宛先情報の領域の少なくとも一部に（当該エッジスイッチ10-6のノードIDと）配下の端末BのユーザIDが記載されていれば、当該受信パケットの宛先情報を端末BのMACアドレス宛に戻し、当該受信パケットを端末BのMACアドレス宛に転送する」という旨のEgress エントリが登録されているため、当該ID化パケットの宛先情報の領域に記載されているID群を、端末BのMACアドレス宛に変更（宛先情報に端末BのMACアドレスを上書き）し、当該変更後のパケット（元の受信パケット）を、出力ポート14に転送し、出力ポート14から端末Bに転送する。

【0100】

なお、上記では、宛先情報の領域として、「MACアドレス領域」を例に説明しているが、実際には「IPアドレス領域」でも良い。すなわち、「宛先アドレスの領域」であれば良い。

【0101】

< 第2実施形態 >

10

20

30

40

50

以下に、本発明の第2実施形態について説明する。

本実施形態では、ステーション検知の際ではなく、スイッチ間でのIPパケット等の通信を開始した際に、入力側エッジスイッチのフローテーブルに「Ingressエントリ」を登録する事例について説明する。

【0102】

ここでは、図3に示すステーション検知(ARPリクエスト利用)により、エッジスイッチに「Egressエントリ」を登録する処理のみ実施した状態、或いは、図4に示すステーション検知(ARプリプライ利用)の際に、エッジスイッチに「Ingressエントリ」を登録する処理を実施していない状態を想定している。

【0103】

[パケットの問い合わせ]

図6を参照して、データ通信開始時のパケットの問い合わせの際の処理について説明する。

【0104】

コントローラ20は、セキュアチャネル(Secure Channel)接続で、管理下にあるスイッチから受信パケットについての問い合わせを受けた場合に、応答としてIngressエントリを登録する。

【0105】

まず、端末Aは、宛先となる端末BのMACアドレス(及びIPアドレス)を、パケットの宛先情報の領域に記載し、端末B宛のパケットを送信する。

【0106】

端末Aが接続されたエッジスイッチ10-1は、端末B宛のパケットを受信すると、自身のフローテーブルに当該パケットと適合するエントリが登録されているか確認する。この時点では、端末Aが接続されたエッジスイッチ10-1には、当該パケットと適合するエントリ(Ingressエントリ)は登録されていないため、端末Aが接続されたエッジスイッチ10-1は、セキュアチャネル(Secure Channel)接続で、コントローラ20宛に端末B宛のパケットを転送する。

【0107】

コントローラ20は、端末B宛のパケットの宛先情報から端末BのMACアドレス(及びIPアドレス)を取得し、自身が保持しているトポロジ情報及びステーション情報から、端末BのユーザIDと、端末Bが接続されたエッジスイッチ10-6のノードIDを特定する。

【0108】

コントローラ20は、端末Aが接続されたエッジスイッチ10-1のフローテーブルに、「端末B宛のパケットを受信した場合、当該受信パケットの宛先情報の領域の少なくとも一部に、ID群(端末Bが接続されたエッジスイッチのノードIDと、当該フローのVTNIDと、端末BのユーザID)を記載して次段のスイッチに転送する」という旨のIngressエントリを登録する。このエントリ登録方式は「Reactive型」(最初のパケット到着時にIngressエントリを登録する方式)である。

【0109】

<各実施形態の関係>

なお、上記の各実施形態は、組み合わせて実施することも可能である。例えば、本発明に係るネットワークシステムにおいて、第1実施形態に対応するスイッチと、第2実施形態に対応するスイッチが混在していても良い。また、複数のVTNが存在する場合、VTN毎に実施形態を別々にしても良い。

【0110】

<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。

【0111】

10

20

30

40

50

[付記 1 : システム構成]

それぞれが、自身のフローテーブルに登録されたエントリのルールに適合する受信パケットに対し、前記エントリに定義されたアクションに従った動作を行う複数のスイッチと、

前記複数のスイッチによって形成されるネットワークの物理トポロジに応じて計算された経路に固有の識別子をルールとし、所定の出力ポートからの出力をアクションとしたエントリを、複数のスイッチ間において通信が開始される前に、前記各スイッチに登録するコントローラとを含む

ネットワークシステム。

10

【 0 1 1 2 】

[付記 2 : Core エントリ登録]

付記 1 に記載のネットワークシステムであって、

前記コントローラは、

通信開始前のトポロジ検知の際、前記各スイッチに内向き及び外向きの設定情報を投入し、前記各スイッチのポート情報を収集するポート情報収集部と、

前記各スイッチのポート情報を基に、内向きポートと外向きポートとを特定し、エッジスイッチとコアスイッチとを特定するスイッチ特定部と、

内向きポート、及び外向きと明示的な設定のないポートに対して、隣接スイッチ情報を収集するための制御フレームを送信する制御フレーム送出部と、

20

前記制御フレームに対する応答から、ネットワークの物理トポロジを検知し、トポロジ情報を作成するトポロジ情報作成部と、

前記各スイッチに、前記固有の識別子としてノード ID を対応付けるノード ID 割り当て処理部と、

外向きポートが存在するスイッチ間の経路を計算し、コアスイッチのフローテーブルに、受信パケットの宛先情報の領域にエッジスイッチのノード ID が記載されている場合は次段のスイッチに当該受信パケットを転送する旨の中継用エントリを登録する中継用エントリ登録部と

を具備する

ネットワークシステム。

30

【 0 1 1 3 】

[付記 3 : 送信側の Egress エントリ登録]

付記 2 に記載のネットワークシステムであって、

前記コントローラは、

通信開始前のステーション検知の際、送信元端末からの ARP 要求を基に、前記送信元端末の宛先情報を特定する送信元端末特定部と、

前記送信元端末に、前記固有の識別子としてユーザ ID を対応付ける送信元ユーザ ID 割り当て処理部と、

前記送信元端末が接続されたエッジスイッチのフローテーブルに、受信パケットの宛先情報の領域に当該エッジスイッチのノード ID 及び前記送信元端末のユーザ ID が記載されている場合は、当該受信パケットの宛先情報の領域に前記送信元端末の宛先情報を記載し、前記送信元端末に当該受信パケットを転送する旨の出力用エントリを登録する送信元端末側出力用エントリ登録部と

40

を更に具備する

ネットワークシステム。

【 0 1 1 4 】

[付記 4 : 受信側の Egress エントリ登録]

付記 3 に記載のネットワークシステムであって、

前記コントローラは、

通信開始前のステーション検知の際、宛先端末からの ARP 応答を基に、前記宛先端末

50

の宛先情報を特定する宛先端末特定部と、

前記宛先端末に、前記固有の識別子としてユーザIDを対応付ける宛先ユーザID割り当て処理部と、

前記宛先端末が接続されたエッジスイッチのフローテーブルに、受信パケットの宛先情報の領域に当該エッジスイッチのノードID及び前記宛先端末のユーザIDが記載されている場合は、当該受信パケットの宛先情報の領域に前記宛先端末の宛先情報を記載し、前記宛先端末に当該受信パケットを転送する旨の出力用エントリを登録する宛先端末側出力用エントリ登録部と

を更に具備する

ネットワークシステム。

10

【0115】

[付記5：送信側のIngressエントリ登録(Proactive型)]

付記4に記載のネットワークシステムであって、

前記コントローラは、

通信開始前のステーション検知の際、前記宛先端末からのARP応答を基に、前記送信元端末と前記宛先端末との間の通信が可能であると判断する通信可否判断部と、

前記送信元端末が接続されたエッジスイッチのフローテーブルに、前記宛先端末宛のパケットを受信した場合、当該受信パケットの宛先情報の領域の少なくとも一部に、前記宛先端末が接続されたエッジスイッチのノードID及び前記宛先端末のユーザIDを記載して次段のスイッチに転送する旨の入力用エントリを登録する送信元端末側入力用エントリ登録部と

20

を更に具備する

ネットワークシステム。

【0116】

[付記6：送信側のIngressエントリ登録(Reactive型)]

付記4に記載のネットワークシステムであって、

前記コントローラは、

通信が開始された際、前記送信元端末が接続されたエッジスイッチからのパケットの問い合わせを基に、前記送信元端末と前記宛先端末との間の通信が可能であると判断する通信可否判断部と、

30

前記送信元端末が接続されたエッジスイッチのフローテーブルに、前記宛先端末宛のパケットを受信した場合、当該受信パケットの宛先情報の領域の少なくとも一部に、前記宛先端末が接続されたエッジスイッチのノードIDと前記宛先端末のユーザIDを記載して次段のスイッチに転送する旨の入力用エントリを登録する送信元端末側入力用エントリ登録部と

を更に具備する

ネットワークシステム。

【0117】

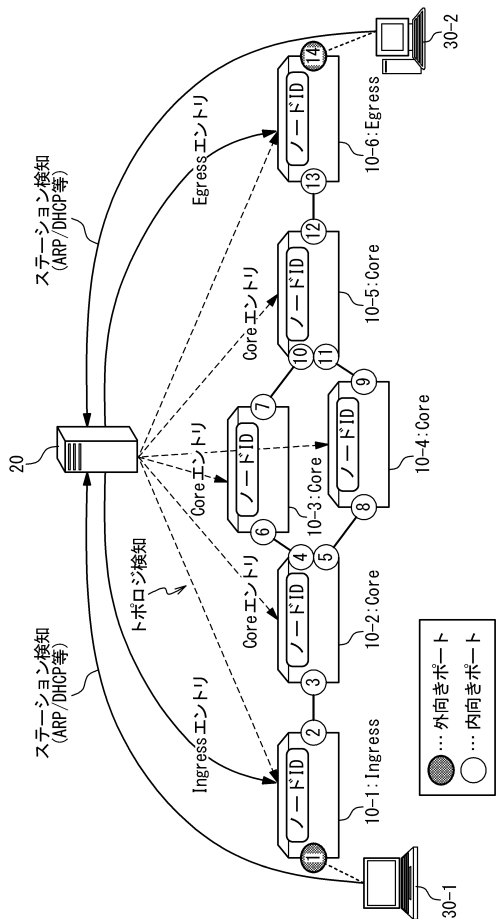
以上、本発明の実施形態を詳述してきたが、実際には、上記の実施形態に限られるものではなく、本発明の要旨を逸脱しない範囲の変更があっても本発明に含まれる。

40

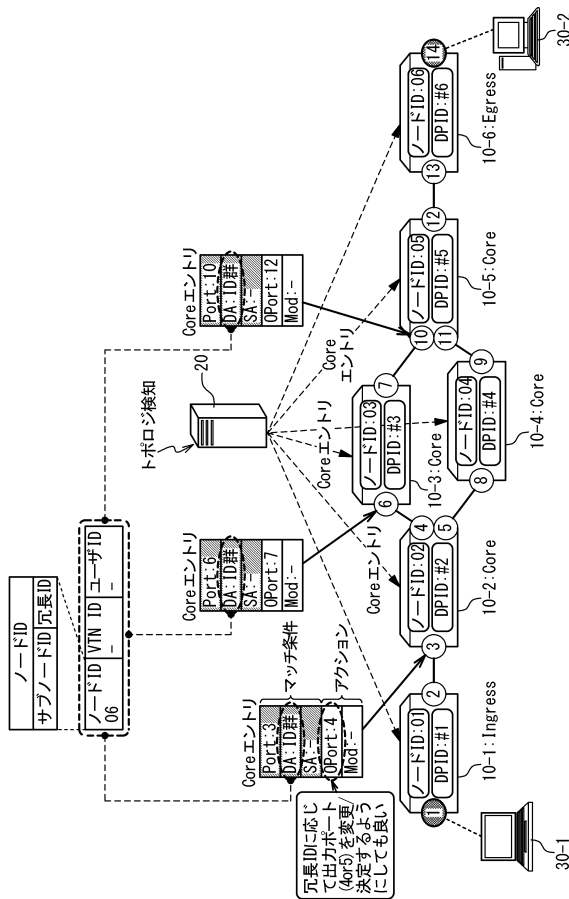
【0118】

なお、本出願は、日本出願番号2011-005137に基づく優先権を主張するものであり、日本出願番号2011-005137における開示内容は引用により本出願に組み込まれる。

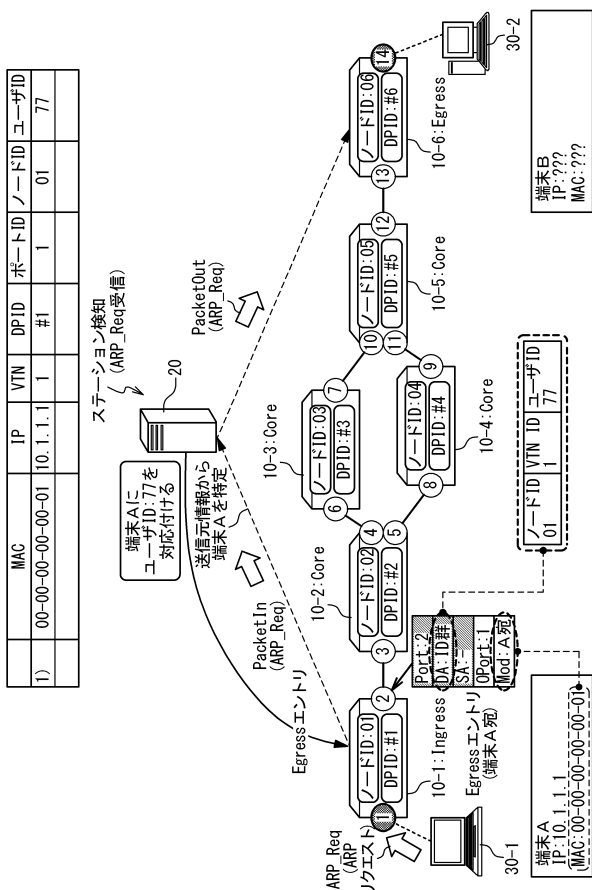
【図 1】



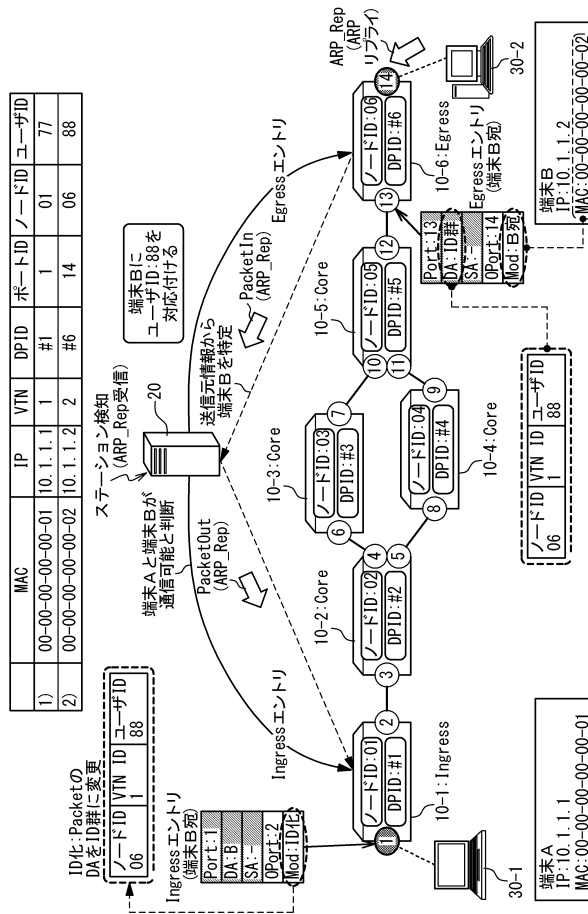
【図 2】



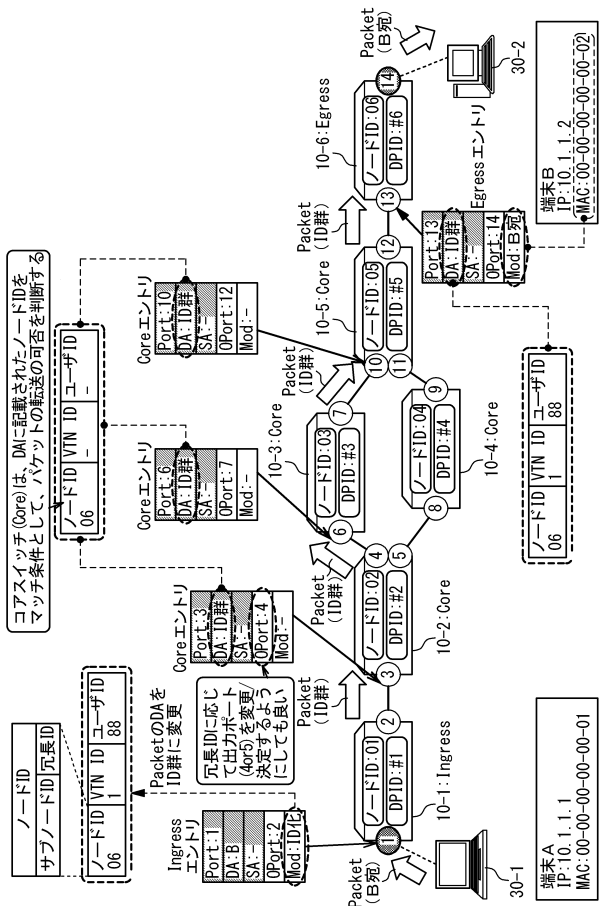
【図 3】



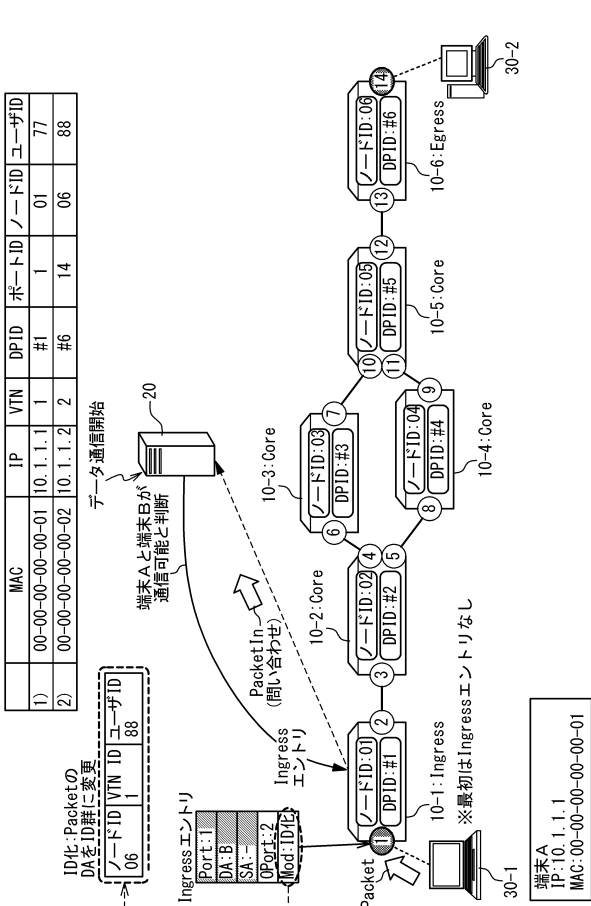
【図 4】



【図5】



【図6】



フロントページの続き

(56)参考文献 特開2009-77219(JP,A)
特開2008-187686(JP,A)

(58)調査した分野(Int.Cl., DB名)
H04L 12/717