

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5561366号
(P5561366)

(45) 発行日 平成26年7月30日(2014.7.30)

(24) 登録日 平成26年6月20日(2014.6.20)

(51) Int.Cl. F I
HO 4 L 12/717 (2013.01) HO 4 L 12/717
HO 4 L 12/741 (2013.01) HO 4 L 12/741

請求項の数 7 (全 29 頁)

| | | | |
|---------------|------------------------------|-----------|--------------------------|
| (21) 出願番号 | 特願2012-532901 (P2012-532901) | (73) 特許権者 | 000004237 |
| (86) (22) 出願日 | 平成23年7月21日(2011.7.21) | | 日本電気株式会社 |
| (86) 国際出願番号 | PCT/JP2011/066524 | | 東京都港区芝五丁目7番1号 |
| (87) 国際公開番号 | W02012/032864 | (74) 代理人 | 100102864 |
| (87) 国際公開日 | 平成24年3月15日(2012.3.15) | | 弁理士 工藤 実 |
| 審査請求日 | 平成25年3月6日(2013.3.6) | (72) 発明者 | 鈴木 洋司 |
| (31) 優先権主張番号 | 特願2010-200690 (P2010-200690) | | 東京都港区芝五丁目7番1号 日本電気株式会社社内 |
| (32) 優先日 | 平成22年9月8日(2010.9.8) | (72) 発明者 | 高島 正徳 |
| (33) 優先権主張国 | 日本国(JP) | | 東京都港区芝五丁目7番1号 日本電気株式会社社内 |
| | | (72) 発明者 | 久保田 一志 |
| | | | 東京都港区芝五丁目7番1号 日本電気株式会社社内 |

最終頁に続く

(54) 【発明の名称】 スイッチシステム、スイッチ制御方法、及び記憶媒体

(57) 【特許請求の範囲】

【請求項1】

所定のケットに対する処理を定義した複数のテーブルを、各テーブルに定義された条件及び処理内容に応じて論理的に統合してオープンフローテーブルを構築し、前記オープンフローテーブルを参照して、受信ケットに対する処理内容を決定するオープンフロー機能部と、

前記決定された処理内容に従って、前記受信ケットに対する処理を実行するアクション機能部と

を具備し、

前記オープンフロー機能部は、

前記受信ケットについて、前記複数のテーブルのうち、L2テーブル、L3テーブル、及びその他テーブルを検索し、該当するエントリを検索結果として出力するL2/L3/その他テーブル検索機能部と、

前記受信ケットについて、前記複数のテーブルのうち、TCAM(Ternary Content Addressable Memory)を検索し、該当するエントリを検索結果として出力するTCAM検索機能部と、

TCAM、L2テーブル、L3テーブル、及びその他テーブルの間の優先順位を定義し、L2/L3/その他テーブル検索機能部及びTCAM検索機能部から検索結果を受け取り、前記優先順位に応じて、採用するエントリを決定し、前記受信ケットに対する処理内容を決定するオープンフロー処理部と

を具備する

スイッチシステム。

【請求項 2】

請求項 1 に記載のスイッチシステムであって、

前記オープンフロー機能部は、

前記オープンフローテーブルにおける優先順位の一定範囲ずつを各テーブルに割り当てる手段と、

コントローラから前記オープンフローテーブルの制御が行われる際、前記コントローラから指定された優先順位の範囲内の値を基に、使用するテーブルを判別する手段と

を具備する

スイッチシステム。

【請求項 3】

請求項 1 に記載のスイッチシステムであって、

前記オープンフロー機能部は、

前記オープンフローテーブルを構成する各テーブルにテーブル ID を設定する手段と、コントローラから前記オープンフローテーブルの制御が行われる際、前記コントローラから指定されたテーブル ID を基に、使用するテーブルを判別する手段と

を具備する

スイッチシステム。

【請求項 4】

オープンフロースイッチにおいて実施されるスイッチ制御方法であって、

所定のバケットに対する処理を定義した複数のテーブルを、各テーブルに定義された条件及び処理内容に応じて論理的に統合してオープンフローテーブルを構築し、前記オープンフローテーブルを参照して、受信バケットに対する処理内容を決定することと、

前記決定された処理内容に従って、前記受信バケットに対する処理を実行することと、

前記受信バケットについて、前記複数のテーブルのうち、L2 テーブル、L3 テーブル、及びその他テーブルを検索し、該当するエントリを検索結果とすることと、

前記受信バケットについて、前記複数のテーブルのうち、TCAM (Ternary Content Addressable Memory) を検索し、該当するエントリを検索結果とすることと、

TCAM、L2 テーブル、L3 テーブル、及びその他テーブルの間の優先順位を定義し、前記優先順位に応じて、各テーブルの検索結果の中から採用するエントリを決定し、前記受信バケットに対する処理内容を決定することと

を含む

スイッチ制御方法。

【請求項 5】

請求項 4 に記載のスイッチ制御方法であって、

前記オープンフローテーブルにおける優先順位の一定範囲ずつを各テーブルに割り当てることと、

コントローラから前記オープンフローテーブルの制御が行われる際、前記コントローラから指定された優先順位の範囲内の値を基に、使用するテーブルを判別することと

を更に含む

スイッチ制御方法。

【請求項 6】

請求項 4 に記載のスイッチ制御方法であって、

前記オープンフローテーブルを構成する各テーブルにテーブル ID を設定することと、コントローラから前記オープンフローテーブルの制御が行われる際、前記コントローラから指定されたテーブル ID を基に、使用するテーブルを判別することと

を更に含む

スイッチ制御方法。

10

20

30

40

50

【請求項 7】

所定の packets に対する処理を定義した複数のテーブルを、各テーブルに定義された条件及び処理内容に応じて論理的に統合してオープンフローテーブルを構築し、前記オープンフローテーブルを参照して、受信 packets に対する処理内容を決定するステップと、

前記決定された処理内容に従って、前記受信 packets に対する処理を実行するステップと、

前記受信 packets について、前記複数のテーブルのうち、L2 テーブル、L3 テーブル、及びその他テーブルを検索し、該当するエントリを検索結果とするステップと、

前記受信 packets について、前記複数のテーブルのうち、TCAM (Ternary Content Addressable Memory) を検索し、該当するエントリ

10

を検索結果とするステップと、
TCAM、L2 テーブル、L3 テーブル、及びその他テーブルの間の優先順位を定義し、前記優先順位に応じて、各テーブルの検索結果の中から採用するエントリを決定し、前記受信 packets に対する処理内容を決定するステップと

をスイッチとして使用される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】**

本発明は、スイッチシステムに関し、特に個々のスイッチが複数のテーブルを有するスイッチシステムに関する。

20

【背景技術】**【0002】**

ネットワークシステムにおける通信経路の制御については、近年、通信機器の制御プロトコルであるオープンフロー (OpenFlow) 技術を使用した経路制御方式が研究されている。オープンフロー技術による経路制御が行われるネットワークを、オープンフローネットワークと呼ぶ。

【0003】

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) 等のスイッチのオープンフローテーブル (OpenFlow Table) 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の間は、コントローラがオープンフロープロトコルに準拠した制御メッセージを用いてスイッチを制御するためのセキュアチャネル (Secure Channel) により接続されている。

30

【0004】

オープンフローネットワーク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 packets の受信から出力側エッジスイッチでの送信までの packets の一連の流れをフロー (Flow) と呼ぶ。

【0005】

40

オープン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 packets (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) を定義したフローエントリ (Flow entry) が登録されたテーブルである。

【0006】

フローエントリのルールは、packets の各プロトコル階層のヘッダー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合せにより定義され、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 Prot

50

ocol 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。

【0007】

フローエントリのアクションは、「特定のポートに出力する」、「廃棄する」、「ヘッダーを書き換える」といった動作を示す。例えば、スイッチは、フローエントリのアクション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ならば、パケットを破棄する。或いは、スイッチは、フローエントリのアクションにヘッダー情報が示されていれば、当該ヘッダー情報に基づいてパケットのヘッダを書き換える。

【0008】

オープンフローネットワーク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、フローエントリのアクションを実行する。

【0009】

オープンフロースイッチの詳細については、非特許文献1、2に記載されている。

【0010】

ネットワーク上の大量のフローを制御するには、大容量のオープンフローテーブルを有する必要がある。現状では、オープンフローテーブルに用いられるTCAM(Ternary Content Addressable Memory)はあまり大容量ではないため、必要十分な容量が確保できているとは言えない。また、オープンフローテーブルに使用されているスイッチのテーブル(主にTCAM)の容量自体を増やすことは難しい。

【0011】

上記の課題の解決方法の1つとして、外部TCAMを使用するという方法があるが、これにはコストがかかる。また、10G多ポート(データ伝送速度10Gビット/秒に対応した複数のポートを有するネットワーク機器)のような高速転送の機器では、外部TCAMの使用自体が不可という状況がある。少なくとも現状では、10Gベースのスイッチで動作可能な外部TCAMは存在しない。

【先行技術文献】

【非特許文献】

【0012】

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

【非特許文献2】“OpenFlow Switch Specification Version 0.9.0(Wire Protocol 0x98) July 20, 2009 Current Maintainer: Brandon Heller(brandonh@stanford.edu)”
<<http://www.openflowswitch.org/documents/openflow-spec-v0.9.0.pdf>>

【発明の概要】

【0013】

本発明の目的は、既存のリソースであるスイッチの各テーブルを用いてオープンフローテーブルのエントリ数拡張を実現することにある。

【0014】

本発明に係るスイッチシステムは、所定のパケットに対する処理を定義した複数のテーブルを、各テーブルに定義された条件及び処理内容に応じて論理的に統合してオープンフローテーブルを構築し、オープンフローテーブルを参照して、受信パケットに対する処理内容を決定するオープンフロー機能部と、決定された処理内容に従って、受信パケットに対する処理を実行するアクション機能部とを具備する。

【0015】

本発明に係るスイッチ制御方法は、オープンフロースイッチにおいて実施されるスイッ

10

20

30

40

50

チ制御方法であって、所定の packets に対する処理を定義した複数のテーブルを、各テーブルに定義された条件及び処理内容に応じて論理的に統合してオープンフローテーブルを構築し、オープンフローテーブルを参照して、受信 packets に対する処理内容を決定することと、決定された処理内容に従って、受信 packets に対する処理を実行することを含む。

【0016】

本発明に係るプログラムは、上記のスイッチ制御方法における処理を、スイッチとして使用される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。

【0017】

これにより、コントローラからスイッチ内部の複数のテーブルを、1つの大容量のオープンフローテーブルとして用いることを可能とする。

【図面の簡単な説明】

【0018】

【図1】本発明に係るスイッチシステムの構成例を示す概念図である。

【図2】本発明の第1実施形態に係るオープンフロー機能部の詳細を示す概念図である。

【図3】本発明の第1実施形態に係るオープンフロー処理部の詳細を示す概念図である。

【図4】オープンフローテーブル制御の概要を示す図である。

【図5】オープンフローテーブル制御の第1方式の例の詳細を示す図である。

【図6】オープンフローテーブル制御の第2方式の例の詳細を示す図である。

【図7A】パケット流入時のスイッチシステムの第1の動作例を示す図である。

【図7B】パケット流入時のスイッチシステムの第1の動作例を示す図である。

【図8A】パケット流入時のスイッチシステムの第2の動作例を示す図である。

【図8B】パケット流入時のスイッチシステムの第2の動作例を示す図である。

【図9】オープンフローテーブル制御の第1方式の具体例を説明するための図である。

【図10】オープンフローテーブル制御の第2方式の具体例を説明するための図である。

【図11】本発明の第2実施形態に係るスイッチシステムの実施例1の詳細を示す概念図である。

【図12】本発明の第2実施形態に係るスイッチシステムの実施例2の詳細を示す概念図である。

【図13】本発明の第2実施形態に係るスイッチシステムの実施例3の詳細を示す概念図である。

【発明を実施するための形態】

【0019】

< 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。

【0020】

[システム構成]

図1に示すように、本発明に係るスイッチシステムは、コントローラ101と、スイッチ102を含む。

【0021】

コントローラ101は、オープンフロープロトコルに準拠した処理でスイッチ102を制御する。

【0022】

[スイッチの構成]

スイッチ102は、プロトコル制御部103と、入力ポート104と、オープンフロー機能部105と、レガシー(Legacy)機能部108と、アクション(Action)機能部111と、出力ポート112を備える。

【0023】

10

20

30

40

50

プロトコル制御部 103 は、コントローラ 101 がオープンフロープロトコルに準拠した処理でスイッチ 102 を制御するための通信を行う際に、コントローラ 101 とスイッチ 102 の間でプロトコル制御を行う。プロトコル制御部 103 は、スイッチ 102 の内部に限らず、スイッチ 102 の前段でも良い。

【0024】

入力ポート 104 は、パケット入力用インタフェースである。入力ポート 104 は、オープンフロー有効ポートと、オープンフロー無効ポートを有する。オープンフロー有効ポートとは、オープンフロープロトコルに対応した入力ポートであり、オープンフロー無効ポートとは、オープンフロープロトコルに未対応の入力ポートである。

【0025】

オープンフロー機能部 105 は、オープンフロー有効ポートから入力したパケットを処理する。

【0026】

オープンフロー機能部 105 は、オープンフローテーブル管理部 106 と、オープンフロー処理解決部 (OpenFlow Action Resolver) 107 を備える。

【0027】

オープンフローテーブル管理部 106 は、スイッチ 102 が使用するオープンフローテーブルを保持する。オープンフローテーブルには、オープンフロープロトコルに準拠したパケットに対するアクション (オープンフロー処理のアクション) が定義されている。

【0028】

オープンフロー処理解決部 107 は、オープンフローテーブル管理部 106 の検索 (Lookup) 結果を基に、オープンフロー処理のアクションを決定する。

【0029】

レガシー機能部 108 は、オープンフロー無効ポートから入力したパケットを処理する。

【0030】

レガシー機能部 108 は、レガシーテーブル (Legacy Table) 管理部 109 と、レガシー処理解決部 (Legacy Action Resolver) 110 を備える。

【0031】

レガシーテーブル管理部 109 は、スイッチ 102 が使用するレガシーテーブルである。レガシーテーブル管理部 109 は、オープンフロープロトコルに準拠していないパケット (通常のパケット等) に対するアクション (レガシー処理のアクション) を定義している。

【0032】

レガシー処理解決部 110 は、レガシーテーブル管理部 109 の検索結果を基に、該当するレガシー処理のアクションを決定する。レガシー処理については、通常のスイッチ機能を用いる。

【0033】

アクション機能部 111 では、オープンフロー機能部 105、もしくはレガシー機能部 108 で決定されたアクションを実行する。

【0034】

出力ポート 112 は、パケット出力用インタフェースである。

【0035】

[オープンフロー処理とレガシー処理の違い]

オープンフロー処理では、パケットの経路制御を外部のコントローラ経由で行う。コントローラは、ネットワーク全体を見て最適な経路を選択する。一方、レガシー処理では、通常のスイッチやルータで行われているように、自律分散で経路制御を行う。通常のスイッチやルータは、自身の近傍に関する情報からネットワーク状況を判断して最適経路を選

10

20

30

40

50

択する。

【 0 0 3 6 】

オープンフロー処理では、最大 1 2 種類の情報の組合せでパケット識別を行うことができる。一方、レガシー処理では、L 2 ネットワークであれば宛先 M A C アドレス、L 3 ネットワークであれば宛先 I P アドレスというように、パケット識別に使用できる情報が少ない。このため、細かなフロー制御を行うことが難しい。例えば、レガシー処理では、同じ宛先 I P アドレスでも、送信元 T C P ポート番号が異なれば、別のフローと判断して別の経路を選択する。

【 0 0 3 7 】

[スイッチシステムの全体動作]

以下に、図 1 のスイッチシステムの全体動作について説明する。

10

【 0 0 3 8 】

[パケット入力]

スイッチ 1 0 2 に新規パケットが流入した際、スイッチ 1 0 2 は、入力ポート 1 0 4 で、このパケットを受け取る。

【 0 0 3 9 】

スイッチ 1 0 2 は、パケットを受け取った入力ポート 1 0 4 がオープンフロー有効ポートか否か確認する。例えば、スイッチ 1 0 2 は、スイッチ 1 0 2 自体又は入力ポート 1 0 4 の設定情報 (c o n f i g) 等を参照して、入力ポート 1 0 4 がオープンフロー有効ポートか否か確認する。

20

【 0 0 4 0 】

[パケット入力からオープンフロー処理への移行]

スイッチ 1 0 2 は、入力ポートがオープンフロー有効ポートであった場合、パケットを、入力ポート 1 0 4 からオープンフロー機能部 1 0 5 に渡す。

【 0 0 4 1 】

[オープンフロー処理]

オープンフロー機能部 1 0 5 は、渡されたパケットについて、スイッチ 1 0 2 の複数のテーブルを保持するオープンフローテーブル管理部 1 0 6 において検索処理を行う。

【 0 0 4 2 】

次に、オープンフロー機能部 1 0 5 は、オープンフロー処理解決部 1 0 7 において、検索結果、及び各テーブルの優先順位 (P r i o r i t y) からパケットのアクションを決定する。なお、優先順位は、優先度と読み替えても良い。

30

【 0 0 4 3 】

[オープンフロー処理からアクション実行への移行]

オープンフロー機能部 1 0 5 は、決定されたアクションが「 P a c k e t - I N 」 (コントローラに対する該当パケットのアクション問い合わせ) であった場合 (例えば、該当するフローエントリが存在せずアクションが決定できなかった場合) は、プロトコル制御部 1 0 3 を通じてコントローラ 1 0 1 に問い合わせ (該当パケットの転送等) を行う。最初は、レガシー処理の対象となるパケット以外の全てのパケットのアクションを、無条件で「 P a c k e t - I N 」としておいても良い。オープンフロー機能部 1 0 5 は、問い合わせに対する応答として、「 P a c k e t - O U T 」 (コントローラからのアクション問い合わせ結果) を受け取り、その内容を該当パケットのアクションとして決定し、オープンフローテーブル管理部 1 0 6 が保持するテーブルに登録する。以降、オープンフロー機能部 1 0 5 は、オープンフロー処理解決部 1 0 7 において、このパケットと同じルールに該当するパケットに対して、このアクションを決定することになる。

40

【 0 0 4 4 】

オープンフロー機能部 1 0 5 は、決定されたアクションに基づいて、該当パケットをアクション機能部 1 1 1 に渡す。すなわち、該当パケットの処理の主体が、オープンフロー機能部 1 0 5 からアクション機能部 1 1 1 に移行する。

【 0 0 4 5 】

50

[オープンフロー処理からレガシー処理への移行]

オープンフロー機能部105は、決定されたアクションが「NORMAL」（レガシー機能部108を用いたパケット処理）であった場合は、該当パケットをレガシー機能部108に渡す。すなわち、該当パケットの処理の主体が、オープンフロー機能部105からレガシー機能部108に移行する。

【0046】

[パケット入力からレガシー処理への移行]

また、スイッチ102は、入力ポート104がオープンフロー無効ポートであった場合、もしくは当該パケットに対して先に（以前に）オープンフロー機能部105で決定されたアクションが「NORMAL」であった場合、パケットを、入力ポート104からレガシー機能部108に渡す。すなわち、該当パケットの処理の主体が、入力ポート104からレガシー機能部108に移行する。

10

【0047】

[レガシー処理]

レガシー機能部108は、渡されたパケットについて、スイッチ102の複数のテーブルから構成されるレガシーテーブル管理部109において検索処理を行う。

【0048】

[レガシー処理からアクション実行への移行]

次に、レガシー機能部108は、レガシー処理解決部110において、検索結果、及び各テーブルの優先順位（Priority）からパケットのアクションを決定する。すなわち、該当パケットの処理の主体が、レガシー機能部108からアクション機能部111に移行する。

20

【0049】

なお、レガシー機能部108～レガシー処理解決部110でのレガシー処理については、通常のスイッチ機能を用いるため、本発明では詳細に触れない。

【0050】

[アクション実行]

アクション機能部111は、該当パケットに対して、オープンフロー機能部105、もしくはレガシー機能部108で決定されたアクションを実行する。

【0051】

オープンフロー機能部105で決定されたアクションの例として、ヘッダー（Header）情報の書き換え、指定された出力ポートからのパケット出力、パケット破棄等が考えられる。レガシー機能部108で決定されたアクションの例として、ルーティング（routing）等によるパケット転送等が考えられる。但し、実際には、これらの例に限定されない。

30

【0052】

最後に、アクション機能部111は、実行されるアクションに「パケット出力」が含まれている場合は、その内容に従って、適切な出力ポート112から該当パケットを出力する。

【0053】

[コントローラによるオープンフローテーブルの制御]

なお、コントローラ101は、プロトコル制御部103を通じて、スイッチ102のオープンフローテーブル管理部106の制御を行うことが可能である。ここで、「オープンフローテーブル管理部106の制御」とは、オープンフローテーブルに対するフローエントリの登録/変更/削除/一括削除等である。

40

【0054】

オープンフローテーブル管理部106がオープンフローテーブルとして保持するスイッチ102の各テーブルは、必ずしもオープンフロー仕様（OpenFlow Spec）で規定されている全ての操作に対応している訳ではない。

【0055】

50

このため、コントローラ 101 は、オープンフローテーブル管理部 106 がオープンフローテーブルとして保持する各テーブルで実現可能な機能（設定可能なアクション）を踏まえて制御を行う必要がある。

【0056】

[オープンフロー機能部の詳細]

図2は、本発明のオープンフロー機能部105の詳細について説明するための図である。

【0057】

オープンフロー機能部105、オープンフローテーブル管理部106、オープンフロー処理解決部107、アクション機能部111は、図1に表されている機構・機能と同一のものである。

10

【0058】

オープンフロー機能部105は、オープンフローテーブル管理部106と、オープンフロー処理解決部107を備える。

【0059】

オープンフローテーブル管理部106は、テーブル群113と、検索(Lookup)機能部114を備える。

【0060】

テーブル群113は、オープンフローテーブルを構成するテーブル群である。

【0061】

検索機能部114は、テーブル群113を用いて入力パケットの検索を行う。

20

【0062】

検索機能部114は、L2/L3/その他テーブル(OF)検索機能部115と、TCAM(OF)検索機能部116を備える。

【0063】

なお、「OF」とは、「OpenFlow」の略語である。

【0064】

L2/L3/その他テーブル(OF)検索機能部115は、入力パケットについて、L2テーブル(OF)、L3テーブル(OF)、及びその他テーブル(OF)を参照し、該当するエントリがないか検索する。その他テーブル(OF)の一例として、マルチキャストルーティングテーブル等が挙げられる。すなわち、L2/L3/その他テーブル(OF)検索機能部115は、入力パケットについて、プロトコル単位でテーブルの検索を実行する。

30

【0065】

TCAM(OF)検索機能部116は、入力パケットについて、TCAM(OF)を参照し、該当するエントリがないか検索する。すなわち、TCAM(OF)検索機能部116は、入力パケットについて、TCAMの検索を実行する。

【0066】

[オープンフロー機能部の動作]

以下に、図2のオープンフロー機能部105の動作について説明する。

40

【0067】

入力ポート104は、オープンフロー有効ポートから入力されたパケットを、オープンフロー機能部105に渡す。

【0068】

オープンフロー機能部105は、渡されたパケットについて、スイッチ102の複数のテーブルから構成されるオープンフローテーブル管理部106において検索処理を行う。

【0069】

このとき、オープンフローテーブル管理部106の検索機能部114は、オープンフローテーブルを構成するテーブル群113に登録されたエントリ情報に基づいて検索を行う。

50

【0070】

具体的には、検索機能部114において、まず、L2/L3/その他テーブル(OF)検索機能部115が検索を行い、次いで、TCAM(OF)検索機能部116が検索を行う。

【0071】

検索機能部114は、この検索結果を、オープンフロー処理解決部107に渡す。

【0072】

オープンフロー処理解決部107は、検索結果、及び各テーブルの優先順位(Priority)からパケットのアクションを決定する。

【0073】

[ハードウェアの例示]

以下に、本発明に係るスイッチ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。

【0074】

コントローラ101の例として、PC(パソコン)、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101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(VM:Virtual Machine)でも良い。

【0075】

スイッチ102の例として、L3スイッチ(layer 3 switch)、L4スイッチ(layer 4 switch)、L7スイッチ/アプリケーションスイッチ(layer 7 switch)、或いは、マルチレイヤスイッチ(multi-layer switch)等のネットワークスイッチ(network switch)を想定している。他にも、スイッチ102の例として、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/セキュリティ監視制御装置(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。

【0076】

プロトコル制御部103、オープンフロー機能部105、オープンフローテーブル管理部106、オープンフロー処理解決部107、レガシー機能部108、レガシーテーブル管理部109、レガシー処理解決部110、及びアクション機能部111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によって実現される。

【0077】

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、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。

【0078】

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리카ード(Secure Digital memory card)等の記憶媒

10

20

30

40

50

体（メディア）等が考えられる。また、バッファ（buffer）やレジスタ（register）でも良い。或いは、DAS（Direct Attached Storage）、FC-SAN（Fibre Channel - Storage Area Network）、NAS（Network Attached Storage）、IP-SAN（IP - Storage Area Network）等を用いたストレージ装置でも良い。

【0079】

また、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。

10

【0080】

或いは、プロトコル制御部103、オープンフロー機能部105、オープンフローテーブル管理部106、オープンフロー処理解決部107、レガシー機能部108、レガシーテーブル管理部109、レガシー処理解決部110、及びアクション機能部111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（VM）でも良い。

【0081】

入力ポート104及び出力ポート112の例として、ネットワーク通信に対応した基板（マザーボードやI/Oボード）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。

20

【0082】

また、入力ポート104及び出力ポート112が利用する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

【0083】

なお、プロトコル制御部103、オープンフロー機能部105、オープンフローテーブル管理部106、オープンフロー処理解決部107、レガシー機能部108、レガシーテーブル管理部109、レガシー処理解決部110、及びアクション機能部111の各々は、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。

【0084】

但し、実際には、これらの例に限定されない。

【0085】

[オープンフロー処理解決部の詳細]

40

図3は、本発明のオープンフロー処理解決部107の詳細について説明するための図である。

【0086】

オープンフロー処理解決部107、TCAM（OF）検索機能部116は、図2に表されている機構・機能と同一のものである。

【0087】

オープンフロー処理解決部107は、TCAM（OF）のエントリマッピングを調整することにより、TCAM（OF）検索機能部116の一部として実現される。

【0088】

このため、実質的には、オープンフロー処理解決部107と、TCAM（OF）検索機

50

能部 116 は、1つの機能ブロック (TCAM (OF) Lookup & Open Flow Action Resolver) となる。この機能ブロックは、テーブル間優先順位 117 と、TCAM (OF) 内部のエントリ 118 を有する。

【0089】

テーブル間優先順位 117 は、期待するアクション優先順位 (Priority) を示す。TCAM (OF) 内部のエントリ 118 は、優先順位 (Priority) に対応する TCAM (OF) 内のエントリマッピングを示す。

【0090】

TCAM (OF) 内部のエントリ 118 は、TCAM (OF) 検索用エントリ群 119 と、L2 テーブル (OF) 検索結果参照用エントリ 120 と、L3 テーブル (OF) 検索結果参照用エントリ 121 と、その他テーブル (OF) 検索結果参照用エントリ 122 と、Miss-hit 用エントリ 123 を含む。

10

【0091】

TCAM (OF) 検索用エントリ群 119 は、TCAM (OF) 検索機能部 116 における TCAM (OF) 検索を実現するためのエントリの集合である。L2 テーブル (OF) 検索結果参照用エントリ 120 は、L2 テーブル (OF) の検索結果を参照するためのエントリである。L3 テーブル (OF) 検索結果参照用エントリ 121 は、L3 テーブル (OF) の検索結果を参照するためのエントリである。その他テーブル (OF) 検索結果参照用エントリ 122 は、その他テーブル (OF) の検索結果を参照するためのエントリである。Miss-hit 用エントリ 123 は、いずれのエントリにもヒット (hit) しなかったパケットを「Miss-hit (ヒットせず)」として取り扱うためのエントリである。すなわち、これらは上記のそれぞれのアクションを定義したエントリである。

20

【0092】

[オープンフロー処理解決部の動作]

以下に、図 3 のオープンフロー処理解決部 107 の動作について説明する。

【0093】

オープンフローテーブル内の検索順序として、最後に、TCAM (OF) 検索機能部 116 が検索を行う。なお、TCAM (OF) 検索機能部 116 での検索の際に、L2 / L3 / その他テーブルの検索結果を参照することが可能なスイッチでは、TCAM (OF) のエントリマッピングを調整することにより、オープンフロー処理解決部 107 を実現することが可能である。

30

【0094】

例えば、テーブル間優先順位 117 のような期待するテーブルごとのアクション優先順位 (Priority) を、TCAM (OF) 内部のエントリ 118 のようにマッピングすることにより対応することができる。

【0095】

図 3 の例の場合は、アクション優先順位 (Priority) 117 は、優先順位 (Priority) の高い順から、「全部対応 (TCAM (OF) 相当)」、「L2 対応 (L2 テーブル (OF) 相当)」、「L3 対応 (L3 テーブル (OF) 相当)」、「その他対応 (その他テーブル相当)」となっている。

40

【0096】

オープンフロー処理解決部 107 は、これを TCAM (OF) 内のエントリマッピングとして、TCAM (OF) の検索優先順位 (Priority) の高い順から、「TCAM (OF) 検索機能実現用エントリ群 119」、「L2 テーブル (OF) 検索結果参照用エントリ 120」、「L3 テーブル (OF) 検索結果参照用エントリ 121」、「その他テーブル (OF) 検索結果参照用エントリ 122」、「Miss-hit 用エントリ 123」の順に並べれば良い。

【0097】

[TCAM (OF) 検索機能実現用エントリ群]

TCAM (OF) 検索機能実現用エントリ群 119 は、検索機能の TCAM (OF) 検

50

索機能部 116 を実現するためのエン트리群である。オープンフロー処理解決部 107 は、入力パケットが TCAM (OF) 検索機能実現用エン트리群 119 のいずれかにヒットした場合、そのエントリのアクションを、当該パケットに対するアクションとして選択する。

【0098】

[L2テーブル(OF)検索結果参照用エン트리]

L2テーブル(OF)検索結果参照用エン트리120は、事前に行われたL2テーブル(OF)の検索の結果を踏まえ、入力パケットに対応するエントリがL2テーブル(OF)上に存在した場合にヒットするエントリである。

【0099】

例えば、L2テーブル(OF)検索時に入力パケットに対応するエントリがあった場合は、あるフラグ「X=1」となり、TCAM(OF)上のL2テーブル(OF)検索結果参照用エン트리120では、このフラグ「X=1」の場合に同エントリがヒットする。

【0100】

オープンフロー処理解決部107は、L2テーブル(OF)検索結果参照用エン트리120がヒットした場合、L2テーブル(OF)の該当エントリのアクションを、当該パケットに対するアクションとして選択する。

【0101】

[L3テーブル(OF)検索結果参照用エン트리]

L3テーブル(OF)検索結果参照用エン트리121は、事前に行われたL3テーブル(OF)の検索の結果を踏まえ、入力パケットに対応するエントリがL3テーブル(OF)上に存在した場合にヒットするエントリである。

【0102】

例えば、L3テーブル(OF)検索時に入力パケットに対応するエントリがあった場合は、あるフラグ「Y=1」となり、TCAM(OF)上のL3テーブル(OF)検索結果参照用エン트리121では、このフラグ「Y=1」の場合に同エントリがヒットする。

【0103】

オープンフロー処理解決部107は、L3テーブル(OF)検索結果参照用エン트리121がヒットした場合、L3テーブル(OF)の該当エントリのアクションを、当該パケットに対するアクションとして選択する。

【0104】

[その他テーブル(OF)検索結果参照用エン트리]

その他テーブル(OF)検索結果参照用エン트리122は、事前に行われたその他テーブル(OF)の検索の結果を踏まえ、入力パケットに対応するエントリがその他テーブル(OF)上に存在した場合にヒットするエントリである。

【0105】

例えば、その他テーブル(OF)検索時に入力パケットに対応するエントリがあった場合は、あるフラグ「Z=1」となり、TCAM(OF)上のその他テーブル(OF)検索結果参照用エン트리122では、このフラグ「Z=1」の場合に同エントリがヒットする。

【0106】

オープンフロー処理解決部107は、その他テーブル(OF)検索結果参照用エン트리122がヒットした場合、その他テーブル(OF)の該当エントリのアクションを、当該パケットに対するアクションとして選択する。

【0107】

[Miss-hit用エン트리]

Miss-hit用エン트리123は、TCAM(OF)のいずれのエントリにも入力パケットがヒットしなかった場合にヒットするエントリである。

【0108】

ここでは、Miss-hit用エン트리123は、全てのパターンのパケットがヒット

10

20

30

40

50

するエントリである。オープンフロー処理部 107 は、入力パケットが TCAM (OF) のいずれのエントリにもヒットせず、Miss-hit 用エントリ 123 にのみヒットした場合、オープンフローの設定に応じて、「Packet-IN」(コントローラに対する該当パケットのアクション問い合わせ)、もしくは「NORMAL」(レガシー機能部を用いたパケット処理)を、当該パケットに対するアクションとして選択する。

【0109】

なお、各テーブルのアクション優先順位 (Priority) を変えたい場合は、TCAM (OF) 内の該当エントリの順番を入れ替えれば良い。

【0110】

[オープンフローテーブル制御の概要]

図4は、本発明におけるコントローラからのオープンフローテーブル制御の概要図である。コントローラ101、スイッチ102、プロトコル制御部103、テーブル群113は、図1、図2に表されている機構・機能と同一のものである。

【0111】

スイッチ102は、TCAM124と、L2テーブル125と、L3テーブル126と、その他テーブル127を含む。

【0112】

TCAM124は、TCAM (OF) と、TCAM (Legacy) を含む。L2テーブル125は、L2テーブル (OF) と、L2テーブル (Legacy) を含む。L3テーブル126は、L3テーブル (OF) と、L3テーブル (Legacy) を含む。その他テーブル127は、その他テーブル (OF) と、その他テーブル (Legacy) を含む。

【0113】

TCAM (OF)、L2テーブル (OF)、L3テーブル (OF)、その他テーブル (OF) は、オープンフローテーブルを構成する。

【0114】

TCAM (Legacy)、L2テーブル (Legacy)、L3テーブル (Legacy)、その他テーブル (Legacy) は、レガシーテーブルを構成する。

【0115】

スイッチ102上のTCAM124、L2テーブル125、L3テーブル126、及びその他テーブル127は、通常の場合、物理的には1つのテーブルである。

【0116】

本発明におけるスイッチ102では、この物理的に1つのテーブル (TCAM124、L2テーブル125、L3テーブル126、及びその他テーブル127) を、オープンフローテーブルを構成するテーブル群113と、レガシーテーブルを構成するテーブル群128に論理的に分ける機能を有する。すなわち、スイッチ102は、物理的に1つのテーブル (TCAM124、L2テーブル125、L3テーブル126、及びその他テーブル127) を、各テーブルに定義された条件及び処理内容に応じて論理的に統合してオープンフローテーブル (テーブル群113) とレガシーテーブル (テーブル群128) を構築する。

【0117】

オープンフローテーブルを構成するテーブル群113は、TCAM (OF)、L2テーブル (OF)、L3テーブル (OF)、その他テーブル (OF) を含む。

【0118】

レガシーテーブルを構成するテーブル群128は、TCAM (Legacy)、L2テーブル (Legacy)、L3テーブル (Legacy)、その他テーブル (Legacy) を含む。

【0119】

[オープンフローテーブル制御の概要]

以下に、図4のコントローラからのオープンフローテーブル制御の概要について説明す

10

20

30

40

50

る。

【0120】

コントローラ101は、プロトコル制御部103を通じて、スイッチ102のオープンフローテーブルの制御を行うことが可能である。

【0121】

スイッチ上のTCAM124、L2テーブル125、L3テーブル126、その他テーブル127は、通常の場合、物理的には各々1つのテーブルである。

【0122】

本発明におけるスイッチでは、これらのテーブルのリソースの一部をオープンフロー用に切り出して使用し、オープンフローテーブルを構成するテーブル群113とレガシーテーブルを構成するテーブル群128を論理的に構築する機能を有する。すなわち、スイッチ102は、TCAM124、L2テーブル125、L3テーブル126、及びその他テーブル127を基に、論理的なオープンフローテーブル(テーブル群113)とレガシーテーブル(テーブル群128)を構築する。

10

【0123】

オープンフローテーブルを構成する各テーブルは、それぞれ実現できるオープンフロー機能が異なる。

【0124】

このため、コントローラ101は、「1:各テーブルで実現可能なオープンフロー機能はどのようなものか」、「2:制御するエントリはオープンフローテーブルを構成するどのテーブルのものか」を踏まえた上でオープンフローテーブル制御を行う必要がある。

20

【0125】

「1」については、「コントローラ側:予め各テーブルで実現可能な機能を入力(Input)しておいて、それ以外の制御を行おうとした場合にエラーを返す機構を持たせる」、「スイッチ側:コントローラからの制御命令について、対象となるテーブルが制御に対応する機能を有していない場合にエラーを返す機構を持たせる」といった対応がある。

【0126】

「2」については、「オープンフローテーブルにおける優先順位(Priority)(0-64k)の一定範囲ずつを各テーブルに割り当て、コントローラから制御を行う際はこの優先順位(Priority)の範囲を基に、使用するテーブルを判別する」、「オープンフローテーブルを構成する各テーブルにIDを持たせ、コントローラから制御を行う際はこのIDを基に、使用するテーブルを判別する」といった対応がある。

30

【0127】

[オープンフローテーブル制御の詳細(1)]

図5は、本発明におけるコントローラからのオープンフローテーブル制御の第1方式の例の詳細について説明するための図である。ここでは、優先順位(Priority)範囲でテーブル指定する場合について説明する。

【0128】

コントローラ101、スイッチ102、オープンフローテーブル管理部106、テーブル群113は、図1、図2に表されている機構・機能と同一のものである。

40

【0129】

オープンフロー機能部105のオープンフローテーブル管理部106は、オープンフローテーブルを構成するテーブル群113の各々のテーブルに、優先順位(Priority)範囲を割り当てる。

【0130】

優先順位(Priority)範囲でテーブル指定する場合、各々のテーブルの優先順位(Priority)範囲は重複してはならない。また、優先順位(Priority)範囲の合計は、オープンフローで規定されている優先順位(Priority)範囲を越えてはならない。

【0131】

50

コントローラ101は、各テーブルに割り当てられた優先順位(Priority)範囲内の値でテーブルを指定する。

【0132】

オープンフロー機能部105のオープンフローテーブル管理部106は、コントローラ101から指定された優先順位(Priority)範囲内の値を基に、使用するテーブルを判別する。

【0133】

[オープンフローテーブル制御の詳細(2)]

図6は、本発明におけるコントローラからのオープンフローテーブル制御の第2方式の例の詳細について説明するための図である。ここでは、テーブルIDでテーブル指定する場合について説明する。

10

【0134】

コントローラ101、スイッチ102、オープンフローテーブル管理部106、テーブル群113は、図1、図2に表されている機構・機能と同一のものである。

【0135】

オープンフロー機能部105のオープンフローテーブル管理部106は、オープンフローテーブルを構成するテーブル群113に、各々テーブルID129を割り当てる。

【0136】

テーブルIDでテーブル指定する場合、各々のテーブルIDは重複してはならない。一方、各テーブルで規定される優先順位(Priority)範囲は、重複しても問題ない。各テーブルはテーブルID129により別のテーブルとして識別され、各々のテーブルの優先順位(Priority)はオープンフロー処理部107により決定されるためである。

20

【0137】

コントローラ101は、各テーブルに割り当てられたテーブルID129でテーブルを指定する。

【0138】

オープンフロー機能部105のオープンフローテーブル管理部106は、コントローラ101から指定されたテーブルID129を基に、使用するテーブルを判別する。

【0139】

[パケット流入時の処理の動作例(1)]

図7A、図7Bは、本発明におけるパケット流入時のスイッチシステムの第1の動作例である。である。簡単のため、動作に関連する機能ブロックのみ記載している。ステップS101~ステップS107は、パケット流入からの一連の動作の流れを示す。

30

【0140】

(1)ステップS101

コントローラ101は、予めスイッチ102に対して、{Match条件:宛先IP=AA、Action:Drop}というTCAM(OFF)エントリー(1)の登録を行う。スイッチ102は、テーブル群113のTCAM(OFF)に、TCAM(OFF)エントリー(1)の登録を行う。

40

【0141】

なお、「Match条件」は、L1~L4までの任意のヘッダ情報の組合せにより定義される。また、「Action」は、Match条件に適合したパケットに対する中継/破棄、ヘッダ情報の書き換え等のアクションを定義したものである。

【0142】

(2)ステップS102

コントローラ101は、予めスイッチ102に対して、{Match条件:宛先IP=AA、Action:Port1から出力}というL3テーブル(OFF)エントリー(1)と、{Match条件:宛先IP=BB、Action:Port2から出力}というL3テーブル(OFF)エントリー(2)の登録を行う。スイッチ102は、テーブル群1

50

13のL3テーブル(OF)に、L3テーブル(OF)エントリ(1)及びL3テーブル(OF)エントリ(2)を登録する。

【0143】

(3)ステップS103

入力ポート104は、パケット流入時に、オープンフロー有効ポートに対して宛先IP = AAのパケットが流入した場合、オープンフロー機能部105に送信する。

【0144】

(4)ステップS104

オープンフロー機能部105は、オープンフロー有効ポートに対して宛先IP = AAのパケットが流入した場合、当該パケットを処理する。

10

【0145】

(5)ステップS105

まず、オープンフロー機能部105において、L2/L3/その他テーブル(OF)検索機能部115は、流入したパケットについて、L2/L3/その他テーブル(OF)での検索を行う。この検索において、当該パケットは、L3テーブル(OF)エントリ(1)にヒットする。

【0146】

(6)ステップS106

L2/L3/その他テーブル(OF)検索機能部115は、この検索結果を、オープンフロー処理解決部107に通知する。図7Bでは、オープンフロー処理解決部107と、TCAM(OF)検索機能部116は、1つの機能ブロック(TCAM(OF) Lookup & OpenFlow Action Resolver)として示す。

20

【0147】

(7)ステップS107

TCAM(OF)検索機能部116は、流入したパケットについて、TCAM(OF)での検索を行う。この検索において、当該パケットは、TCAM(OF)エントリ(1)が最初にヒットする。この時点でTCAM(OF)の検索処置は終了する。TCAM(OF)検索機能部116は、この検索結果を、オープンフロー処理解決部107に通知する。

【0148】

(8)ステップS108

オープンフロー処理解決部107は、L2/L3/その他テーブル(OF)検索機能部115とTCAM(OF)検索機能部116の各々から検索結果を受け取り、テーブル間の優先順位に従って、流入したパケットに対するアクションを決定する。ここでは、TCAM(OF)の優先順位が一番高い。そのため、オープンフロー処理解決部107は、TCAM(OF)エントリ(1)のアクション(Drop)を、流入したパケットに対するアクションとして決定し、決定したアクション(Drop)を、アクション機能部111に通知する。

30

【0149】

(9)ステップS109

アクション機能部111は、オープンフロー処理解決部107で決定されたアクション(Drop)を実行する。ここでは、アクション機能部111は、アクションが「Drop」であるため、パケットの出力を行わない。アクション機能部111は、流入したパケット及びこれと同一フローに属する以降のパケットを破棄する。

40

【0150】

[パケット流入時の処理の動作例(2)]

図8A、図8Bは、本発明におけるパケット流入時のスイッチシステムの第2の動作例である。簡単のため、動作に関連する機能ブロックのみ記載している。ステップS201~ステップS210は、パケット流入からの一連の動作の流れを示す。

【0151】

50

(1) ステップ S201

コントローラ101は、予めスイッチ102に対して、{ Match条件：宛先IP = AA、Action：Drop } というTCAM(OF) エントリー(1)の登録を行う。スイッチ102は、テーブル群113のTCAM(OF)に、TCAM(OF) エントリー(1)の登録を行う。

【0152】

(2) ステップ S202

コントローラ101は、予めスイッチ102に対して、{ Match条件：宛先IP = AA、Action：Port1から出力 } というL3テーブル(OF) エントリー(1)と、{ Match条件：宛先IP = BB、Action：Port2から出力 } というL3テーブル(OF) エントリー(2)の登録を行う。スイッチ102は、テーブル群113のL3テーブル(OF)に、L3テーブル(OF) エントリ(1)及びL3テーブル(OF) エントリー(2)を登録する。

10

【0153】

(3) ステップ S203

入力ポート104は、パケット流入時に、オープンフロー有効ポートに対して宛先IP = BBのパケットが流入した場合、オープンフロー機能部105に送信する。

【0154】

(4) ステップ S204

オープンフロー機能部105は、オープンフロー有効ポートに対して宛先IP = BBのパケットが流入した場合、当該パケットを処理する。

20

【0155】

(5) ステップ S205

まず、オープンフロー機能部105において、L2/L3/その他テーブル(OF) 検索機能部115は、当該パケットについて、L2/L3/その他テーブル(OF)での検索を行う。この検索において、当該パケットは、L3テーブル(OF) エントリー(2)にヒットする。

【0156】

(6) ステップ S206

L2/L3/その他テーブル(OF) 検索機能部115は、この検索結果を、オープンフロー処理解決部107に通知する。図8Bでは、オープンフロー処理解決部107と、TCAM(OF) 検索機能部116は、1つの機能ブロック(TCAM(OF) Lookup & OpenFlow Action Resolver)として示す。

30

【0157】

(7) ステップ S207

TCAM(OF) 検索機能部116は、流入したパケットについて、TCAM(OF)での検索を行う。しかし、流入したパケットは、宛先IP = BBのパケットであるため、TCAM(OF)にはヒットするエントリがない。この時点でTCAM(OF)の検索処置は終了する。TCAM(OF) 検索機能部116は、この検索結果を、オープンフロー処理解決部107に通知する。この時点でTCAM(OF)の検索処置は終了する。TCAM(OF) 検索機能部116は、この検索結果を、オープンフロー処理解決部107に通知する。

40

【0158】

(8) ステップ S208

オープンフロー処理解決部107は、L2/L3/その他テーブル(OF) 検索機能部115とTCAM(OF) 検索機能部116の各々から検索結果を受け取り、テーブル間の優先順位に従って、流入したパケットに対するアクションを決定する。ここでは、対象となるエントリがL3テーブル(OF) エントリー(2)しか存在しない。そのため、オープンフロー処理解決部107は、L3テーブル(OF) エントリー(2)のアクション(Port2から出力)を、流入したパケットに対するアクションとして決定し、決定し

50

たアクション (P o r t 2 から出力) を、アクション機能部 1 1 1 に通知する。

【 0 1 5 9 】

(9) ステップ S 2 0 9

アクション機能部 1 1 1 は、オープンフロー処理解決部 1 0 7 で決定されたアクション (P o r t 2 から出力) を実行する。

【 0 1 6 0 】

(1 0) ステップ S 2 1 0

アクション機能部 1 1 1 は、アクション機能部 1 1 1 は、流入したパケット及びこれと同一フローに属する以降のパケットを、 P o r t 2 を持つ出力ポート 1 1 2 に出力する。

【 0 1 6 1 】

(1 1) ステップ S 2 1 1

出力ポート 1 1 2 は、アクション機能部 1 1 1 から出力されたパケットを、 P o r t 2 から出力する。

【 0 1 6 2 】

(1 2) ステップ S 2 1 2

P o r t 2 から出力されたパケットは、ネットワーク上に流出し、宛先 I P = B B に向けて伝送される。

【 0 1 6 3 】

なお、図 8 A、図 8 B において、L 2 / L 3 / その他テーブル (O F) 検索機能部 1 1 5 は、流入したパケットについて、複数のテーブルにヒットするエントリがある場合は、ヒットした全てのエントリを検索結果としてオープンフロー処理解決部 1 0 7 に通知する。オープンフロー処理解決部 1 0 7 は、優先順位が一番高い T C A M (O F) にヒットするエントリが存在しない場合、L 2 / L 3 / その他テーブル (O F) 検索機能部 1 1 5 の検索結果の中で優先順位が一番高いテーブルにヒットするエントリを採用する。

【 0 1 6 4 】

また、図 8 A、図 8 B において、L 2 / L 3 / その他テーブル (O F) 検索機能部 1 1 5 は、流入したパケットについて、複数のテーブルにヒットするエントリがある場合は、テーブル間優先順位 1 1 7 に従い、これらのテーブルのうち優先順位が一番高いテーブルにヒットするエントリを検索結果としてオープンフロー処理解決部 1 0 7 に通知するようにしても良い。

【 0 1 6 5 】

[オープンフローテーブル制御例 (1)]

図 9 は、優先順位 (P r i o r i t y) 範囲でテーブル指定する場合におけるコントローラからのオープンフローテーブル制御の第 1 方式の例である。

【 0 1 6 6 】

登録情報 1 3 2 は、コントローラ 1 0 1 からスイッチ 1 0 2 のオープンフローテーブル管理部 1 0 6 への登録情報である。登録結果 1 3 3 は、オープンフローテーブルを構成するテーブル群 1 1 3 への登録結果である。

【 0 1 6 7 】

オープンフローテーブルを構成するテーブル群 1 1 3 を 1 つのテーブルに見立てた場合、スイッチ 1 0 2 は、優先順位 (P r i o r i t y) により期待するテーブルへの制御を行う。

【 0 1 6 8 】

(1) ステップ S 3 0 1

コントローラ 1 0 1 は、スイッチ 1 0 2 に対して、{ 優先順位 (P r i o r i t y) : 5 0 0 0 1、M a t c h 条件 : X X X X、A c t i o n : Y Y Y Y } というエントリの登録を行う。

【 0 1 6 9 】

(2) ステップ S 3 0 2

スイッチ 1 0 2 は、コントローラ 1 0 1 からエントリが登録された場合、このエントリ

10

20

30

40

50

の優先順位 (Priority) に該当するテーブルである L 2 テーブル (OF) を選択し、この L 2 テーブル (OF) に該当エントリを登録する。

【 0 1 7 0 】

[オープンフローテーブル制御例 (2)]

図 1 0 は、テーブル ID でテーブル指定する場合におけるコントローラからのオープンフローテーブル制御の第 2 方式の例である。

【 0 1 7 1 】

登録情報 1 3 4 は、コントローラからのオープンフローテーブルへの登録情報である。登録結果 1 3 5 は、オープンフローテーブルを構成するテーブル群への登録結果である。

【 0 1 7 2 】

オープンフローテーブルを構成するテーブル群の各テーブルを、それぞれ別のオープンフローテーブルに見立てた場合、スイッチ 1 0 2 は、テーブル ID により期待するテーブルへの制御を行う。

【 0 1 7 3 】

(1) ステップ S 4 0 1

コントローラ 1 0 1 は、スイッチ 1 0 2 に対して、{ テーブル ID : # 2、優先順位 (Priority) : 1、Match 条件 : X X X X、Action : Y Y Y Y } というエントリの登録を行う。

【 0 1 7 4 】

(2) ステップ S 4 0 2

スイッチ 1 0 2 は、コントローラ 1 0 1 からエントリが登録された場合、このエントリのテーブル ID に該当するテーブルである L 2 テーブル (OF) を選択し、この L 2 テーブル (OF) に該当エントリを登録する。

【 0 1 7 5 】

[第 1 実施形態の特徴]

本実施形態によれば、オープンフロー処理部により複数のテーブル間のリソースを統合し、かつ、各テーブルの優先順位 (Priority) を比較してアクション (Action) 解決を行うことができる。

【 0 1 7 6 】

従って、スイッチの複数のテーブルのリソースを用いて大容量のオープンフローテーブルを構築することができる。これにより、スイッチにより大量のフローを制御することが可能となる。

【 0 1 7 7 】

また、本実施形態によれば、オープンフローテーブルを 1 つの大容量のオープンフローテーブルとして使用する場合における「優先順位 (Priority) 範囲」や、オープンフローテーブルを構成するテーブル群の各テーブルを複数の異なるオープンフローテーブルとして使用する場合における「テーブル ID」を用いて、オープンフローテーブルを構成する各テーブルを識別することができる。

【 0 1 7 8 】

従って、スイッチの複数のテーブルから構成されるオープンフローテーブルを 1 つの大容量のオープンフローテーブル、もしくは複数の異なるオープンフローテーブルとして使用することができる。これにより、複数のテーブルから構成されるオープンフローテーブルを柔軟に制御することが可能となる。

【 0 1 7 9 】

< 第 2 実施形態 >

以下に、本発明の第 2 実施形態について添付図面を参照して説明する。

第 2 実施形態は、T C A M (OF) 検索機能部にオープンフロー処理部を含めない場合の、オープンフロー処理部の実施例である。

【 0 1 8 0 】

[実施例 1]

10

20

30

40

50

図 1 1 は、「 T C A M (O F) 検索機能部 1 1 6 」 「 L 2 / L 3 その他テーブル (O F) 検索機能部 1 1 5 」の順に検索が実施される場合のオープンフロー機能部 1 0 5 の詳細について説明するための図である。

【 0 1 8 1 】

オープンフロー機能部 1 0 5、オープンフローテーブル管理部 1 0 6、オープンフロー処理解決部 1 0 7、アクション機能部 1 1 1、テーブル群 1 1 3、検索機能部 1 1 4、L 2 / L 3 / その他テーブル (O F) 検索機能部 1 1 5、T C A M (O F) 検索機能部 1 1 6 は、図 2 に表されている機構・機能と同一のものである。

【 0 1 8 2 】

図 1 1 を参照して、「 T C A M (O F) 検索機能部 1 1 6 」 「 L 2 / L 3 その他テーブル (O F) 検索機能部 1 1 5 」の順に検索が実施される場合のオープンフロー機能部 1 0 5 の動作について説明する。

【 0 1 8 3 】

図 1 1 は、オープンフローテーブル管理部 1 0 6 の検索機能部 1 1 4 において、T C A M (O F) 検索機能部 1 1 6 が最後に検索を行わない事例である。

【 0 1 8 4 】

図 1 1 では、「 T C A M (O F) 検索機能部 1 1 6 」 「 L 2 / L 3 その他テーブル (O F) 検索機能部 1 1 5 」の順に検索が実施されるため、T C A M (O F) 検索機能部 1 1 6 にオープンフロー処理解決部 1 0 7 を含めることができない。

【 0 1 8 5 】

[実施例 2]

図 1 2 は、L 2 / L 3 その他テーブル (O F) 検索機能部 1 1 5 と T C A M (O F) 検索機能部 1 1 6 の検索が同時に実施される場合のオープンフロー機能部 1 0 5 の詳細について説明するための図である。

【 0 1 8 6 】

オープンフロー機能部 1 0 5、オープンフローテーブル管理部 1 0 6、オープンフロー処理解決部 1 0 7、アクション機能部 1 1 1、テーブル群 1 1 3、検索機能部 1 1 4、L 2 / L 3 / その他テーブル (O F) 検索機能部 1 1 5、T C A M (O F) 検索機能部 1 1 6 は、図 2 に表されている機構・機能と同一のものである。

【 0 1 8 7 】

図 1 2 を参照して、L 2 / L 3 その他テーブル (O F) 検索機能部 1 1 5 と T C A M (O F) 検索機能部 1 1 6 の検索が同時に実施される場合のオープンフロー機能部 1 0 5 の動作について説明する。

【 0 1 8 8 】

図 1 2 は、オープンフローテーブル管理部 1 0 6 の検索機能部 1 1 4 において、T C A M (O F) 検索機能部 1 1 6 が最後に行われない事例である。

【 0 1 8 9 】

L 2 / L 3 その他テーブル (O F) 検索機能部 1 1 5 と T C A M (O F) 検索機能部 1 1 6 の検索が同時に実施されるため、T C A M (O F) 検索機能部 1 1 6 にオープンフロー処理解決部 1 0 7 を含めることができない。

【 0 1 9 0 】

[実施例 3]

図 1 3 は、T C A M (O F) 検索機能部 1 1 6 にオープンフロー処理解決部 1 0 7 を含められない場合におけるオープンフロー処理解決部 1 0 7 の実施例である。

【 0 1 9 1 】

オープンフロー処理解決部 1 0 7、アクション機能部 1 1 1 は、図 2 に表されている機構・機能と同一のものである。

【 0 1 9 2 】

オープンフロー処理解決部 1 0 7 は、検索受信部 (L o o k u p R e c e i v e r) 1 3 0 と、処理解決部 (A c t i o n R e s o l v e r) 1 3 1 を備える。

10

20

30

40

50

【0193】

検索受信部 (Lookup Receiver) 130 は、各テーブルの検索結果を受け取る。処理解決部 (Action Resolver) 131 は、各テーブルの検索結果を基に、オープンフロー処理を決定する。

【0194】

図13を参照して、TCAM (OF) 検索機能部 116 にオープンフロー処理解決部 107 を含まれない場合におけるオープンフロー処理解決部 107 の動作について説明する。

【0195】

TCAM (OF) 検索機能部 116 にオープンフロー処理解決部 107 を含めることができない場合、スイッチ 102 に、図13で示されるようなオープンフロー処理解決部 107 を搭載する。

10

【0196】

オープンフロー処理解決部 107 において、検索受信部 (Lookup Receiver) 130 は、各テーブルの検索結果を受け取る。処理解決部 (Action Resolver) 131 は、各テーブルの検索結果、及び予め設定されたテーブル優先順位 (Priority) によりオープンフローのアクションを決定し、決定されたアクションをアクション機能部 111 に通知する。

【0197】

[他の実施例]

20

なお、上記の各実施形態において、スイッチ 102 はオープンフロースイッチであるものとして説明しているが、オープンフロースイッチは例示に過ぎない。実際には、オープンフロースイッチに限定されるものではなく、オープンフロースイッチと同様の機構・機能を有するスイッチに対しても、本発明を適用することは可能である。

【0198】

<本発明の特徴>

本発明は、複数のテーブルの集約によるオープンフローテーブルの拡張方法に関する。

【0199】

本発明は、スイッチの単一テーブル (主に TCAM) で実現されるオープンフローテーブルについて、スイッチの複数のテーブルを用いてオープンフローテーブルを構築することによりオープンフローテーブルフローエントリ数の拡張を実現するものである。

30

【0200】

なお、オープンフローでは、これまで同一の NW 機器 (ルータ/スイッチ等) に実装されていた転送機能、制御機能を分離し、転送機能は NW 機器に残し、制御機能は外出しのコントローラに任せる。コントローラは、遠隔からオープンフロープロトコルを用いて NW 機器のオープンフローテーブルを操作することにより NW 装置の挙動を制御する。オープンフローテーブルは、{Match 条件、Action、統計情報} の 3 種の情報からなるフローエントリ群から構成される。オープンフローでは、この適合 (Match) 条件から制御対象とするフローを定義し、このフローを単位としてアクション (Action) や統計情報の取得を行うことができる。

40

【0201】

オープンフローの {Match 条件、Action、統計情報} の概要は、以下の通りである。

【0202】

[Match 条件]

「Ingress port (入力ポート)」/「Src MAC (送信元 MAC アドレス)」/「Dst MAC (宛先 MAC アドレス)」/「Ether type (タイプ・フィールド)」/「VLAN ID (仮想 LAN 識別情報)」/「VLAN Priority (仮想 LAN 優先順位)」/「Src IP (送信元 IP アドレス)」/「Dst IP (宛先 IP アドレス)」/「IP protocol (IP プロトコル番号)

50

」/「IP ToS (upper 6 bit)」/「Src Port (送信元ポート番号)」/「Dst Port (宛先ポート番号)」

【0203】

[Action]

「Forward (ある物理ポートから出力)」/「All (入力ポート以外全てから出力)」/「Controller (コントローラ向けに出力)」/「Local (自装置のローカルスタック宛に出力)」/「テーブル (オープンフローテーブルの内容に従って出力)」/「In_port (入力ポートから出力)」/「Normal (レガシーテーブルの内容を用いて出力)」/「Flood (入力ポート、及びSpanning Treeのブロックポート以外全てから出力)」/「Drop (パケットを破棄)」/「Modify - Field (パケットのヘッダ情報を書き換え)」

10

【0204】

例えば、「Modify - Field」の場合は、「VLAN ID」、「Vlan 優先順位 (Priority)」、「Src MAC」、「Dst MAC」、「Src IP」、「Dst IP」、「IP ToS」、「Src Port」、「Dst Port」を書き換え可能である。

【0205】

[統計情報]

「テーブル単位」、「フロー単位」、「物理ポート単位」、「キュー (Queue) 単位」での各種統計情報。

20

【0206】

本発明は、スイッチのテーブル (主にTCAM) の容量自体を増やすことなく、スイッチの複数のテーブルを用いてオープンフローテーブルを構築することにより、装置としてのオープンフローテーブル合計フローエントリ数の拡張を実現するものである。すなわち、スイッチ内部で複数のテーブルを使用して、コントローラからこれを大容量オープンフローテーブルとして用いることを可能とした。

【0207】

具体的には、各テーブルの機能 (Match条件 / Action) 差分を吸収してオープンフローテーブルとしての統合を実現した。なお、スイッチの複数のテーブルは、それぞれ本来の用途 (例えば、L2テーブルであればL2中継、L3テーブルであればL3中継) があるため、全てのリソースを使用するのではなく、その一部を切り出す形で使用する。

30

【0208】

以上のように、本発明の特徴は、「複数のテーブルのMatch条件 / Action差分を吸収して、オープンフローテーブルとして統合すること」と、「その際の具体的なアクション決定方法を提供すること」である。

【0209】

本発明では、スイッチの複数のテーブルリソースの一部をオープンフローテーブルとして使用する。

【0210】

また、本発明では、スイッチの各テーブルを、実現可能な機能 (Match条件 / Action) に応じて「機能制限付きのオープンフローテーブルリソース」として取り扱っている。

40

【0211】

また、本発明では、オープンフロー処理部により、各テーブルリソースの機能差分 (Match条件 / Action) を吸収して、オープンフローテーブルリソースとして統合している。

【0212】

また、本発明では、オープンフロー処理部において、TCAMを含むテーブル間の優先順位 (Priority) によりアクション決定を実現している。

50

【 0 2 1 3 】

なお、TCAM (OF) 検索が最後に行われるスイッチにおいては、オープンフロー処理部をTCAM (OF) 検索機能部に含める形で実現している。

【 0 2 1 4 】

更に、本発明では、複数のテーブルから構成されるオープンフローテーブルを、コントローラから柔軟に制御する。

【 0 2 1 5 】

また、本発明では、複数のテーブルを1つのオープンフローテーブルとして使用する場合、優先順位 (Priority) の範囲でのテーブル識別を実現している。

【 0 2 1 6 】

また、本発明では、複数テーブルを個別のオープンフローテーブルとして使用する場合、テーブルIDでのテーブル識別を実現している。

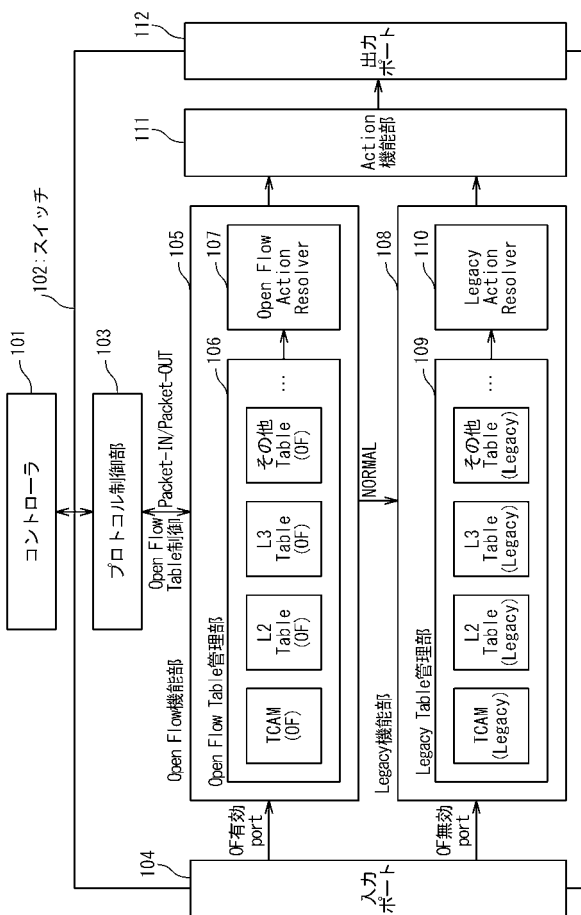
【 0 2 1 7 】

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。

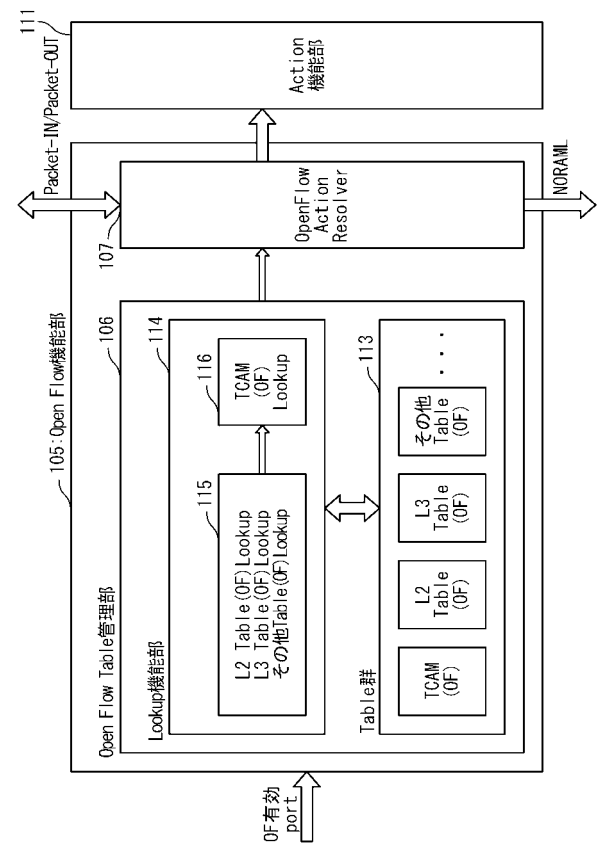
【 0 2 1 8 】

なお、本出願は、日本出願番号2010-200690に基づく優先権を主張するものであり、日本出願番号2010-200690における開示内容は引用により本出願に組み込まれる。

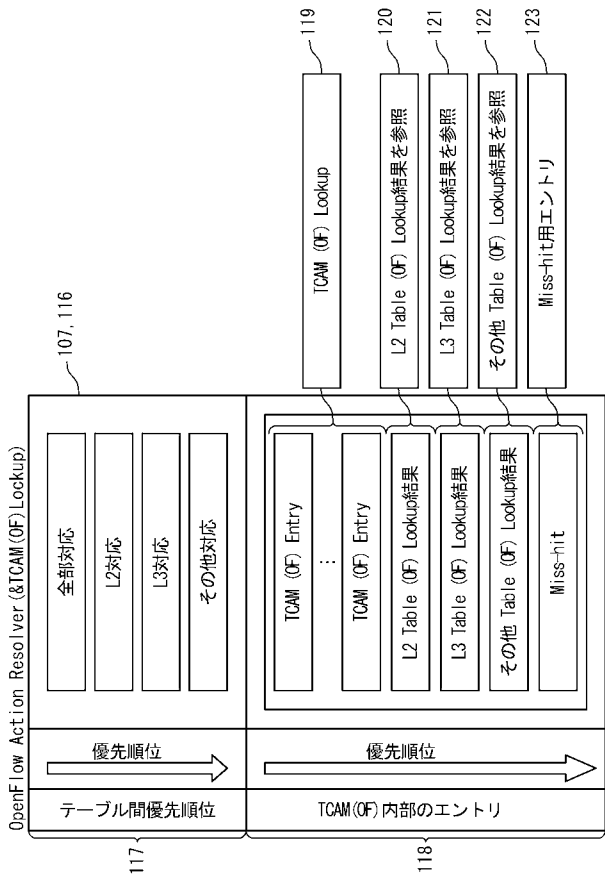
【 図 1 】



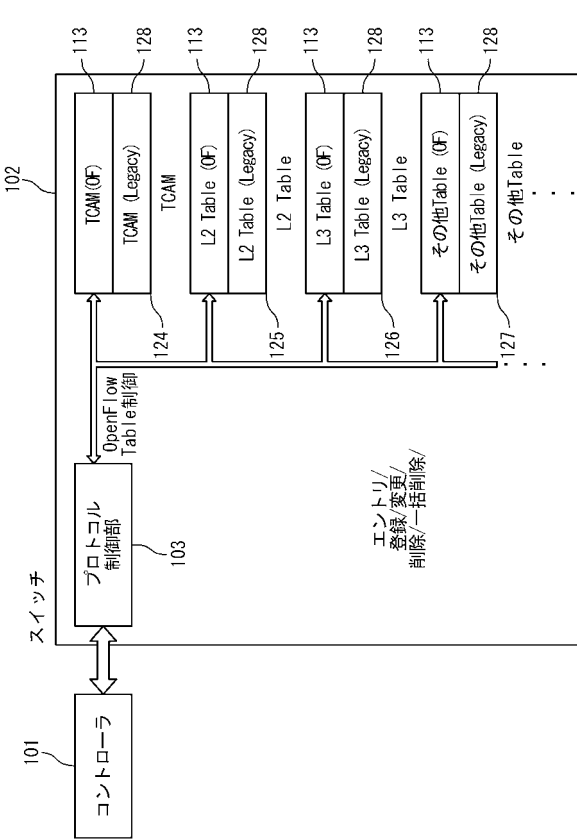
【 図 2 】



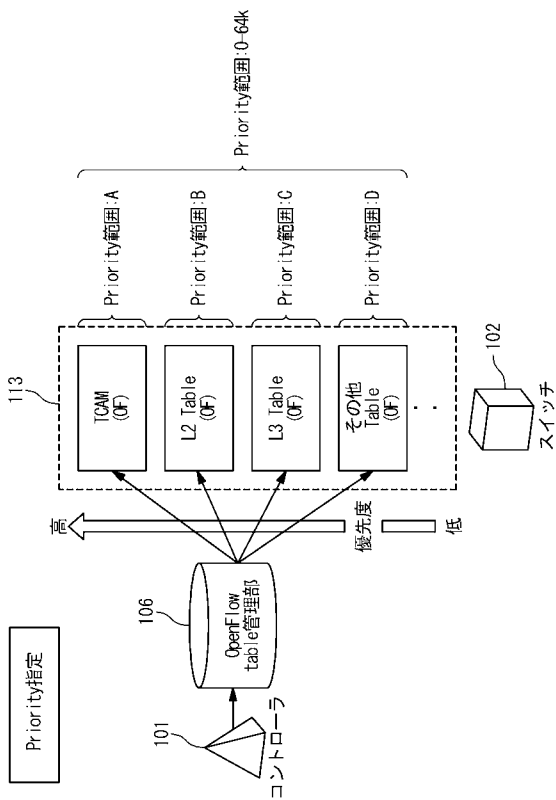
【図3】



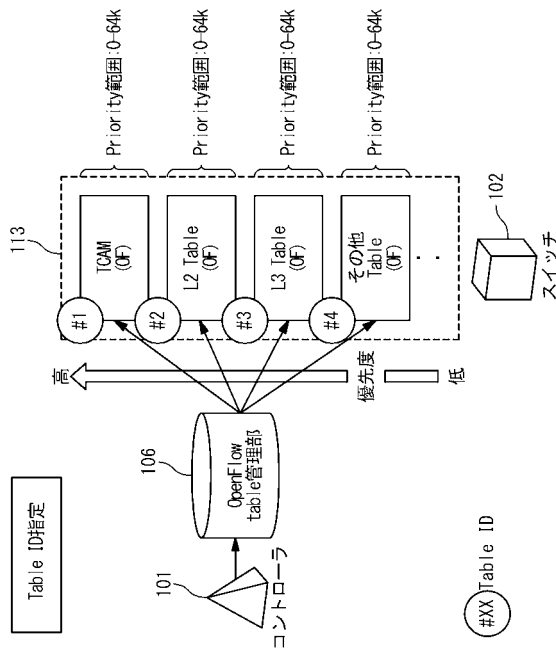
【図4】



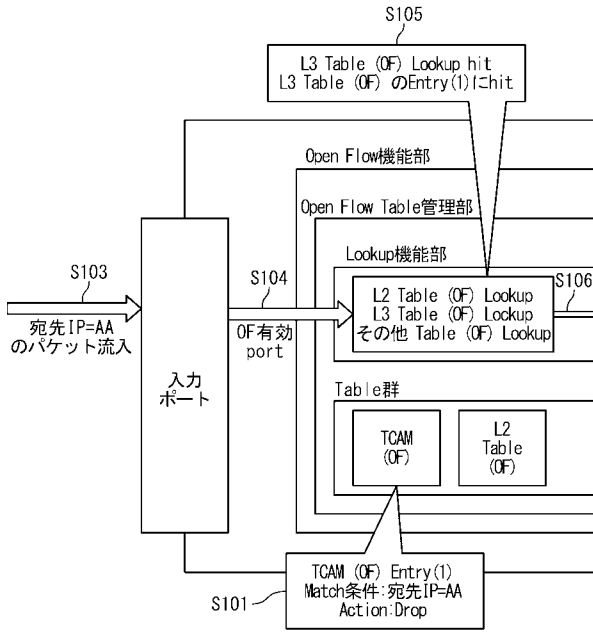
【図5】



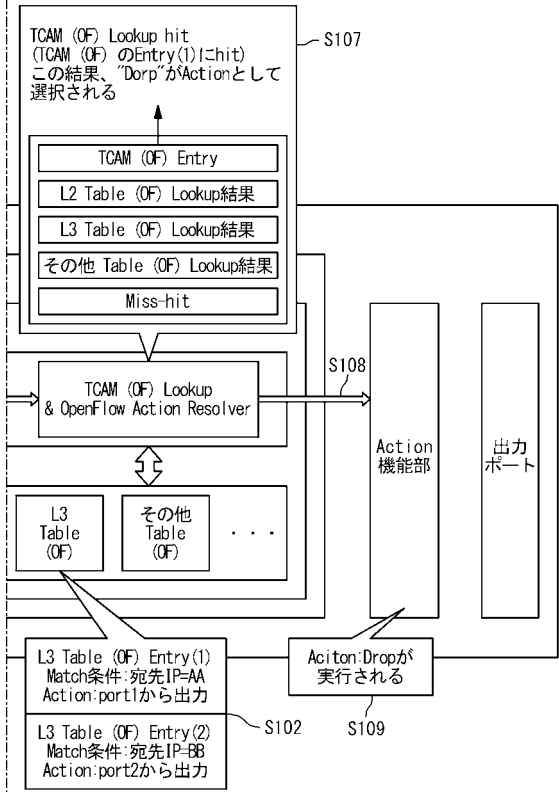
【図6】



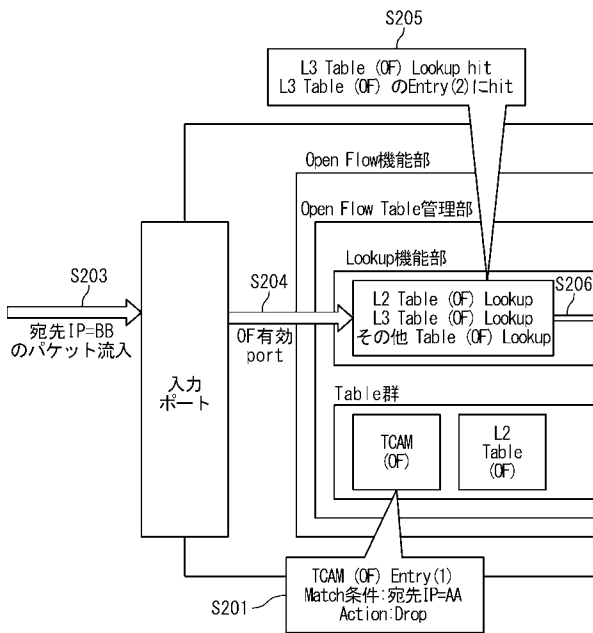
【図7A】



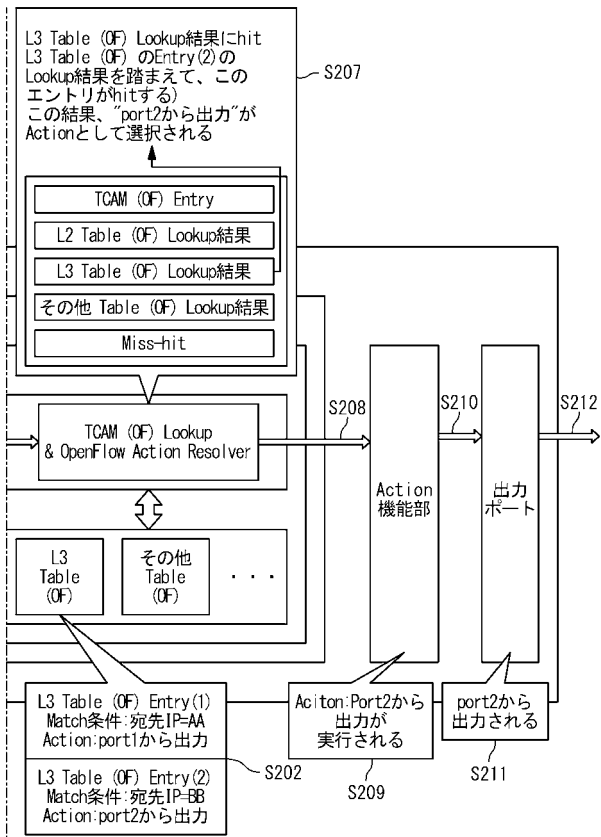
【図7B】



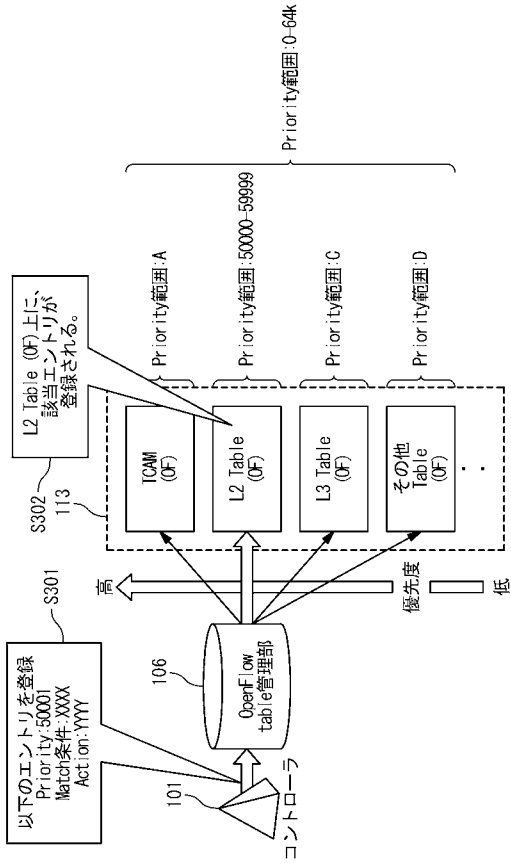
【図8A】



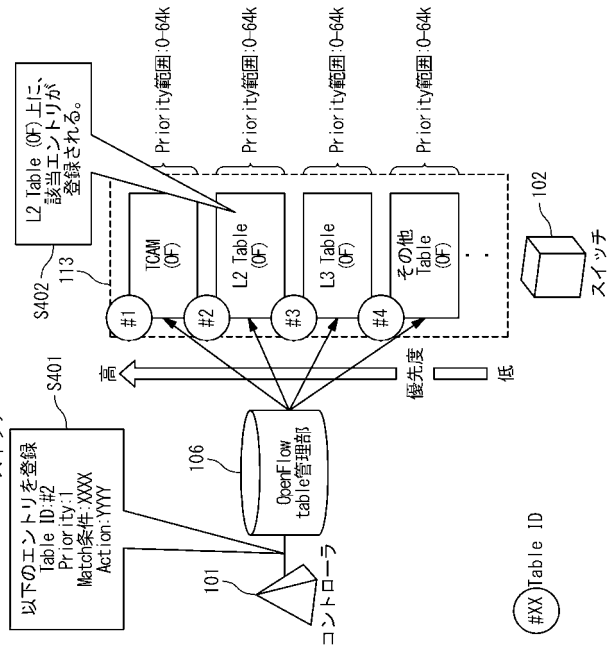
【図8B】



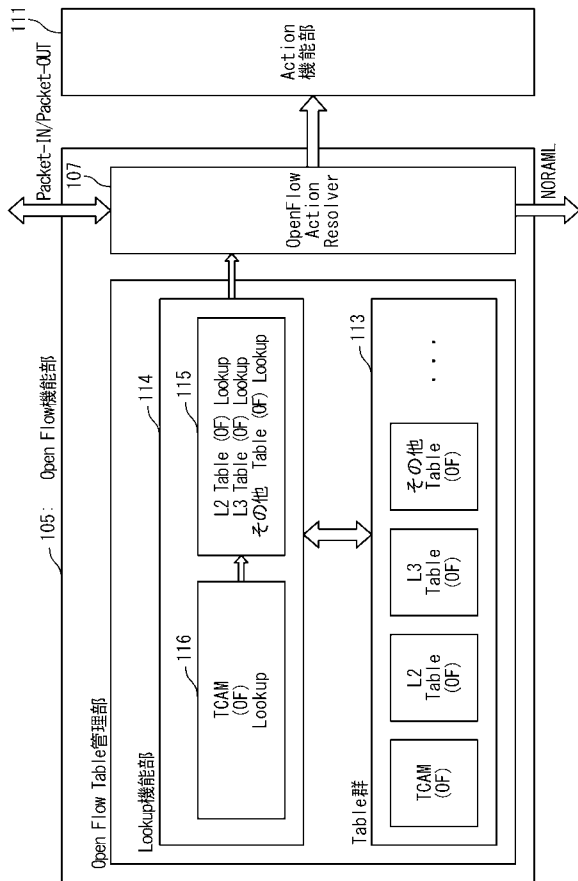
【図 9】



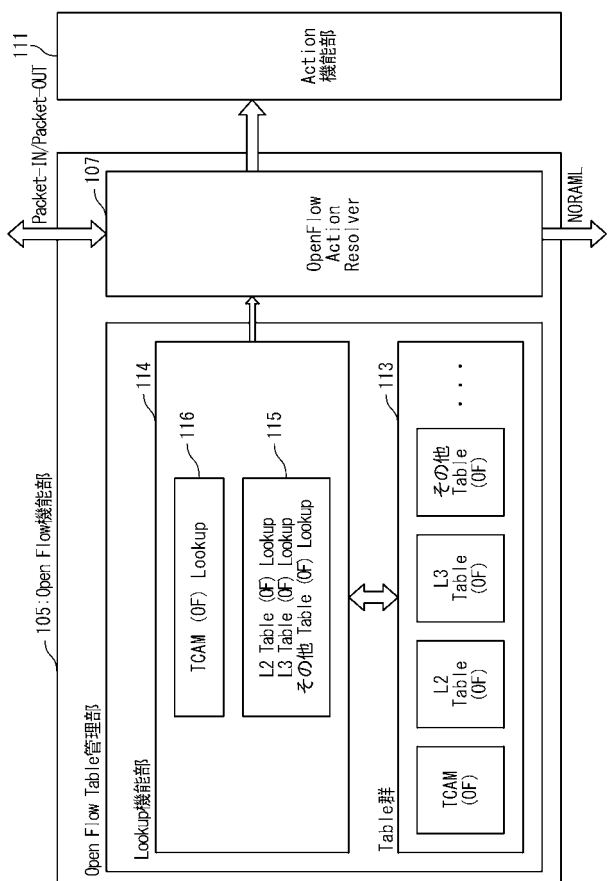
【図 10】



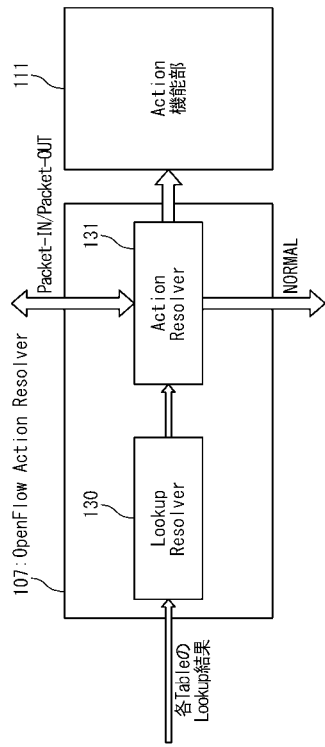
【図 11】



【図 12】



【 図 13 】



フロントページの続き

- (72)発明者 伊澤 徹
東京都港区芝五丁目7番1号 日本電気株式会社内
- (72)発明者 林 将志
東京都港区芝五丁目7番1号 日本電気株式会社内

審査官 松崎 孝大

- (56)参考文献 特開2003-304293(JP,A)
特開2005-252954(JP,A)
特開2006-101343(JP,A)

- (58)調査した分野(Int.Cl., DB名)
- | | |
|------|--------|
| H04L | 12/717 |
| H04L | 12/741 |