

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2016年2月4日(04.02.2016)



(10) 国際公開番号
WO 2016/017737 A1

- (51) 国際特許分類:
H04L 12/70 (2013.01) H04L 12/717 (2013.01)
- (21) 国際出願番号: PCT/JP2015/071611
- (22) 国際出願日: 2015年7月30日(30.07.2015)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2014-156562 2014年7月31日(31.07.2014) JP
- (71) 出願人: 日本電気株式会社(NEC CORPORATION)
[JP/JP]; 〒1088001 東京都港区芝五丁目7番1号
Tokyo (JP).
- (72) 発明者: 嶋田 望(SHIMADA, Nozomi); 〒1088001
東京都港区芝五丁目7番1号 日本電気株式
社内 Tokyo (JP).
- (74) 代理人: 加藤 朝道(KATO, Asamichi); 〒2220033
神奈川県横浜市港北区新横浜3丁目2番12
号加藤内外特許事務所内 Kanagawa (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保
護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA,

BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

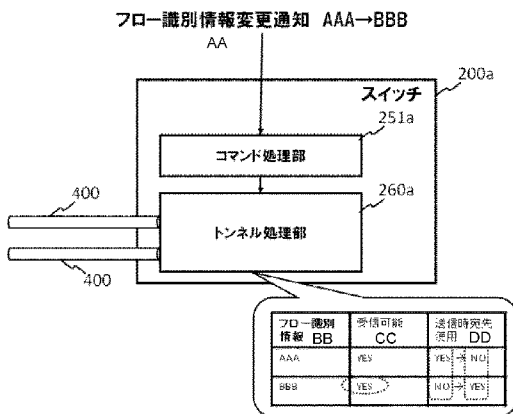
- (84) 指定国 (表示のない限り、全ての種類の広域保
護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユー
ロアジア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨー
ロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 国際調査報告 (条約第 21 条(3))

(54) Title: SWITCH, OVERLAY NETWORK SYSTEM, COMMUNICATION METHOD, AND PROGRAM

(54) 発明の名称: スイッチ、オーバーレイネットワークシステム、通信方法及びプログラム



200a Switch
251a Command processing unit
260a Tunnel processing unit
AA Flow identification information change notification
BB Flow identification information
CC Receiving possible?
DD Destination used for sending

(57) Abstract: The purpose of the present invention is to improve a function of changing tunnel setting information for a switch that functions as a tunnel endpoint of an overlay network system. In the present invention, a switch that functions as a tunnel endpoint of an overlay network system is equipped with: a tunnel processing unit that performs capsulization and decapsulization of frames sent to and received from the opposing tunnel endpoint; and a command processing unit that, upon receiving an instruction to change flow identification information for identifying a capsulized frame, instructs the tunnel processing unit to receive the frame with flow identification information for before and after the change, and thereafter instructs the tunnel processing unit to change the flow identification information to be set in a frame being sent.

(57) 要約: オーバーレイネットワークシステムのトンネルエンドポイントとして機能するスイッチのトンネル設定情報の変更機能を向上させる。オーバーレイネットワークシステムのトンネルエンドポイントとして機能するスイッチは、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネル処理部と、カプセル化されたフレームを識別するためのフロー識別情報の変更指示の受信を契機に、前

記トンネル処理部に対し、変更前後のフロー識別情報を持つフレームの受信を指示した後に、送信フレームに設定するフロー識別情報の変更を指示するコマンド処理部と、を備える。



WO 2016/017737 A1

明 細 書

発明の名称：

スイッチ、オーバーレイネットワークシステム、通信方法及びプログラム
技術分野

[0001] [関連出願についての記載]

本発明は、日本国特許出願：特願2014-156562号（2014年7月31日出願）に基づくものであり、同出願の全記載内容は引用をもって本書に組み込み記載されているものとする。

本発明は、スイッチ、オーバーレイネットワークシステム、通信方法及びプログラムに関し、特に、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネルエンドポイントとして機能するスイッチ、オーバーレイネットワークシステム、通信方法及びプログラムに関する。

背景技術

[0002] レイヤ3ネットワーク上に、仮想的なレイヤ2ネットワークを構築するタイプのオーバーレイネットワークシステムが知られている。このうちのVXLAN (Virtual eXtensible Local Area Network) は、UDPポート情報を追加ヘッダに埋め込み、これを用いて経路制御とロードバランシング機能を備えている。非特許文献1は、VXLANのドラフトである (Page 10、Outer UDP Headerの項参照)。

[0003] 特許文献1に、上記VXLAN等の大規模ネットワーク論理分割技術を用いて多数の論理的なネットワークを構築した構成におけるマルチキャストツリーの管理技術が開示されている。

[0004] 非特許文献2、3は、本発明の関連技術であるオープンフローのホワイトペーパーとオープンフローのスペックである。

先行技術文献

特許文献

[0005] 特許文献1：特開2014-7681号公報

非特許文献

[0006] 非特許文献1：M. Mahalingamほか7名、“VXLAN： A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Network”、[online]、[平成26（2014）年7月17日検索]、インターネット〈URL：<http://tools.ietf.org/pdf/draft-mahalingam-dutt-dcops-vxlan-02.pdf>〉

非特許文献2：Nick McKeownほか7名、“OpenFlow： Enabling Innovation in Campus Networks”、[online]、[平成26（2014）年7月17日検索]、インターネット〈URL：<http://archive.openflow.org/documents/openflow-wp-latest.pdf>〉

非特許文献3：“OpenFlow Switch Specification” Version 1.1.0 Implemented (Wire Protocol 0x02)、[online]、[平成26（2014）年7月17日検索]、インターネット〈URL：<http://archive.openflow.org/documents/openflow-spec-v1.1.0.pdf>〉

発明の概要

発明が解決しようとする課題

[0007] 以下の分析は、本発明によって与えられたものである。上記オーバーレイネットワークシステムにおいて、物理ネットワークがフローを識別するためのフロー識別情報（VXLANの場合、UDPポート番号）を動的に変更したいという要望がある。

[0008] しかしながら、それを実現するためには、オーバーレイネットワークシステムのトンネルエンドポイントとして機能する複数のスイッチに対して、ト

ンネル設定として、フロー識別情報の変更設定を行う必要がある。その際に、設定のタイミングがずれてしまうとその間はネットワークの通信断が発生してしまうという問題点がある。

[0009] 本発明は、オーバーレイネットワークシステムのトンネルエンドポイントとして機能するスイッチのトンネル設定の変更機能の向上に貢献できるスイッチ、オーバーレイネットワークシステム、通信方法及びプログラムを提供することを目的とする。

課題を解決するための手段

[0010] 第1の視点によれば、オーバーレイネットワークのトンネルエンドポイントとして、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネル処理部を備えるスイッチが提供される。このスイッチは、さらに、カプセル化されたフレームを識別するためのフロー識別情報の変更指示の受信を契機に、前記トンネル処理部に対し、変更前後のフロー識別情報を持つフレームの受信を指示した後に、送信フレームに設定するフロー識別情報の変更を指示するコマンド処理部と、を備える。

[0011] 第2の視点によれば、上記したスイッチを用いて構成されたオーバーレイネットワークシステムが提供される。

[0012] 第3の視点によれば、オーバーレイネットワークのトンネルエンドポイントとして、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネル処理部を備えたスイッチにおいて、カプセル化されたフレームを識別するためのフロー識別情報の変更指示の受信を契機に、前記トンネル処理部に対し、変更前後のフロー識別情報を持つフレームの受信を指示するステップと、前記指示の後に、所定の契機で、送信フレームに設定するフロー識別情報の変更を指示するステップと、を含む通信方法が提供される。本方法は、オーバーレイネットワークのトンネルエンドポイントとして機能するスイッチという特定の機械に結びつけられている。

[0013] 第4の視点によれば、オーバーレイネットワークのトンネルエンドポイント

トとして、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネル処理部を備えたスイッチに搭載されたコンピュータに、カプセル化されたフレームを識別するためのフロー識別情報の変更指示の受信を契機に、前記トンネル処理部に対し、変更前後のフロー識別情報を持つフレームの受信を指示する処理と、前記指示の後に、所定の契機で、送信フレームに設定するフロー識別情報の変更を指示する処理と、を実行させるプログラムが提供される。なお、このプログラムは、コンピュータが読み取り可能な（非トランジエントな）記憶媒体に記録することができる。即ち、本発明は、コンピュータプログラム製品として具現することも可能である。

発明の効果

[0014] 本発明によれば、オーバーレイネットワークシステムのトンネルエンドポイントとして機能するスイッチのトンネル設定情報の変更機能の向上に貢献することが可能となる。

図面の簡単な説明

[0015] [図1]本発明の一実施形態のスイッチの構成を示す図である。

[図2]本発明の第1の実施形態のオーバーレイネットワークシステムの構成を示す図である。

[図3]本発明の第1の実施形態の仮想スイッチの構成を示す図である。

[図4]本発明の第1の実施形態の仮想スイッチによりカプセル化後のフレームの構成を示す図である。

[図5]本発明の第1の実施形態の仮想スイッチ内のトンネル処理部の構成を示す図である。

[図6]本発明の第1の実施形態の動作を表したシーケンス図である。

[図7]本発明の第1の実施形態の仮想スイッチのUDPポート番号情報（初期状態）を示す図である。

[図8]本発明の第1の実施形態の仮想スイッチのUDPポート番号情報（UDPポート番号変更通知受信後）を示す図である。

[図9]本発明の第1の実施形態の仮想スイッチのUDPポート番号情報（送信UDPポート番号の変更通知受信後）を示す図である。

[図10]本発明の第1の実施形態の仮想スイッチのUDPポート番号情報（変更前UDPポート番号の解放通知受信後）を示す図である。

発明を実施するための形態

[0016] はじめに本発明の一実施形態の概要について図面を参照して説明する。なお、この概要に付記した図面参照符号は、理解を助けるための一例として各要素に便宜上付記したものであり、本発明を図示の態様に限定することを意図するものではない。

[0017] 本発明は、その一実施形態において、図1に示すように、オーバーレイネットワークのトンネルエンドポイントとして、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネル処理部260aと、コマンド処理部251aとを備えたスイッチ200aにて実現できる。

[0018] より具体的には、コマンド処理部251aは、カプセル化されたフレームを識別するためのフロー識別情報の変更指示（例：AAA→BBB）の受信を契機に、前記トンネル処理部260aに対し、変更前後のフロー識別情報を持つフレームの受信（フロー識別情報＝AAA及びBBBのフレームの受信）を指示する。そして、その後、所定の契機で、前記トンネル処理部260aに対し、送信フレームに設定するフロー識別情報の変更指示が行われる。なお、所定の契機としては、フロー識別情報の変更指示の受信から一定の期間の経過を条件としてもよいし、ユーザからの明示的な指示（送信フレームに設定するフロー識別情報の変更指示）がなされることを条件としてもよい。

[0019] 以上のように構成することにより、トンネル通信に使用するフロー識別情報の変更による通信断の発生を抑制することができる。その理由は、フロー識別情報の変更指示の受信を契機に、直ちに受信対象フレームのフロー識別情報を変更せず、まず、変更前後のフロー識別情報での受信を行わせ、その

後で、送信時に使用するフロー識別情報を変更することで、設定タイミングのずれを吸収するように構成したことにある。

[0020] [第1の実施形態]

続いて本発明の第1の実施形態について図面を参照して詳細に説明する。図2は、本発明の第1の実施形態のオーバーレイネットワークシステムの構成を示す図である。図2を参照すると、オーバーレイネットワークシステムのトンネルエンドポイントとして配置された複数の仮想スイッチ200A～200Cと、これら仮想スイッチ200を制御するフロー制御装置100と、これら仮想スイッチ200を利用して互いに通信する仮想マシン300A～300Eとを含む構成が示されている。仮想スイッチ200間は、後述するVXLANTunnel400AB-1～400BC-2が構築されている。以下、仮想スイッチ200A～200C、仮想マシン300A～300E、及びVXLANTunnel400AB-1～400BC-2を特に区別しない場合、それぞれ「仮想スイッチ200」、「仮想マシン300」、「VXLANTunnel400」と記す。

[0021] フロー制御装置100は、仮想スイッチ200間に、VXLANTunnel400をメッシュ状に構築するためのフロー管理をしており、1以上の仮想スイッチ200に対してフロー設定の指示を行う。これにより、フロー制御装置100は、仮想マシン300間の通信用ネットワークを実現する。なお、このようなフロー制御装置100は、非特許文献2、3のオープンフローコントローラにトンネル管理機能を追加することで実現できる。

[0022] 仮想スイッチ200は、仮想マシン300もしくは他の仮想スイッチからの入力フレームに対してフロー制御装置100の指示によって設定されたフローエントリの中から、入力フレームに適合するマッチ条件を持つフローエントリを選択して入力フレームの処理を行う。仮想スイッチ200は、選択したフローエントリによって入力フレームを他の仮想マシン300との接続用のVXLANTunnel400に出力し、基盤ネットワーク500を越えて他の仮想スイッチ200に出力したり、フロー制御装置100に出力したり

、ドロップしたりする。

[0023] 仮想マシン300は、他の仮想マシン300宛てのフレームを仮想スイッチ200に送信したり、他の仮想マシン300を送信元とするフレームを仮想スイッチ200から受信したりする。また、仮想マシン300は、仮想スイッチ200と直接接続している必要はなくレイヤ2スイッチ(L2SW)を介して接続していてもよいし、仮想マシン300がレイヤ3スイッチ(L3SW)やルータであり、その先に別の仮想マシン300が接続されている場合もある。

[0024] また、接続状態にある仮想スイッチ200と仮想マシン300は、サーバ仮想化技術を利用して、同一の仮想化サーバ上で動作する仮想エンティティであってもよい。

[0025] VXLANトンネル400は、仮想スイッチ200間で基盤ネットワーク500を介して通信を行うためのオーバーレイトンネルである。仮想ネットワーク識別子VNI(VXLAN Network Identifier)と通信を行う2台の仮想スイッチ200の組み合わせにつき1個存在する。図2の例では、3台の仮想スイッチ間にそれぞれ2つのVNIのVXLANトンネル400が構築されている。なお以下の説明において、仮想スイッチ200Aと仮想スイッチ200Bとの間のVNI=1のVXLANトンネル400を、VXLANトンネル400AB-1という形で表す。

[0026] 基盤ネットワーク500は、仮想スイッチ200間のフレーム転送を行うためのレイヤ3ネットワーク等である。本実施形態では、基盤ネットワーク500は、VXLANにてカプセル化された際に付与される宛先IPアドレスへの通信が可能であり、宛先にマルチキャストアドレスが設定されている場合には、全仮想スイッチ200への配信が可能であるものとする。

[0027] 図3は、第1の実施形態の仮想スイッチの構成を示す図である。図3を参照すると、フローエントリ検索部201と、フローエントリ記憶部202と、フローエントリ処理部203と、フロー処理部204と、フロー制御装置通信部205と、VXLAN処理部250とを備えた構成が示されている。

このうち、フローエントリ検索部201、フローエントリ記憶部202、フローエントリ処理部203、フロー処理部204、及びフロー制御装置通信部205が、オープンフロースイッチ機能部に相当する。

される。

[0028] VXLAN処理部250は、コマンド処理部251とトンネル処理部260-1、260-2（以降、トンネル処理部260-1、260-2を特に区別しない場合、「トンネル処理部260」と記す。）とを備えている。トンネル処理部260は、VXLANトンネル400から入力されたVXLANカプセル化されたフレームをフローエントリ検索部201に渡し、フロー処理部204の指示に従い、必要に応じVXLANデカプセル化して仮想マシン300に出力する。また、トンネル処理部260は、仮想マシン300から入力されたフレームをフローエントリ検索部201に渡して、フロー処理部204の指示に従い、必要に応じVXLANカプセル化し、VXLANトンネル400を介して対向する仮想スイッチ200へと転送する処理を行う。

[0029] コマンド処理部251はユーザからの操作を受け付け、トンネル処理部260に対して、UDPポート番号情報の変更コマンド等を出力する。

[0030] フローエントリ検索部201は、入力されたフレームからフローエントリを検索するためのフローエントリ検索条件情報を抽出し、フローエントリ検索条件情報を用いてフローエントリ記憶部203を検索する。フローエントリ検索部201は、前記検索の結果、マッチしたフローエントリのアクションと入力フレームをフロー処理部204に渡す。また、フローエントリ検索部201は、その際にマッチしたフローエントリのタイムアウト時間や統計情報などがあればこれらを更新する。

[0031] フローエントリ記憶部202は、フロー制御装置100より設定されたフローエントリテーブル等を用いて保持する。なお、フローエントリ記憶部202におけるフローエントリの異動は、フロー制御装置100に報告され、把握されている。従って、フロー制御装置100は、仮想スイッチ200の

ふるまいを把握し、制御可能となっている。

[0032] フローエントリ処理部203は、フロー制御装置通信部205を介してフロー制御装置100から来るフローエントリに関する追加・削除などの操作指示や参照指示に従い、フローエントリ記憶部202の更新等を行う。また、フローエントリ処理部203は、フローエントリ記憶部202を参照し、フローエントリでタイムアウトしたものに関しては削除し、フロー制御装置通信部205を介してフロー制御装置100に対して、該フローエントリが削除されたことを報告する。

[0033] フロー処理部204は、フローエントリ検索部201から、又は、フロー制御装置通信部205を介して、フロー制御装置100から渡される入力フレームとそのアクションに従ってフレームの値を変更したり、仮想マシン300にフレームを出力したり、他の仮想スイッチ200にフレームを出力する。また、必要に応じて、フロー処理部204は、VXL AN処理部250にVXL ANカプセル化を指示し、VXL ANトンネル400に出力したり、VXL ANトンネル400から入力されたフレームのVXL ANデカプセル化をしたり、フロー制御装置通信部205を介してフロー制御装置100にフレームを出力したり、ドロップしたりする。

[0034] 図4は、本発明の第1の実施形態の仮想スイッチによりVXL ANカプセル化されたフレームの構成を示す図である。VXL AN処理部250のトンネル処理部260は、VXL ANカプセル化する場合、オリジナルフレームに、Outer MAC、Outer IP、Outer UDP、VXL AN IDといった追加ヘッダにてカプセル化を行う。このうちのOuter UDPヘッダに後記するUDPポート情報番号601、602が設定され、フロー識別やこれを利用した負荷分散（ロードバランシング）に利用可能となっている。デカプセル化の場合には、上記カプセル化されたフレームから追加ヘッダを取り外す処理が行われる。

[0035] 図5は、第1の実施形態の仮想スイッチ200内のトンネル処理部260の構成を示す図である。図5を参照すると、VNI情報記憶部264と、ト

ネル解除部（デカプセル化部）265と、UDPポート情報記憶部266と、トンネル化部（カプセル化部）267とを備えた構成が示されている。

[0036] VN I情報記憶部264はトンネル処理部260が持つVXL ANの仮想ネットワーク識別子（VN I）を保持する。この仮想ネットワーク識別子（VN I）は、図4のVXL ANヘッダに格納される。

[0037] トンネル解除部（デカプセル化部）265は、VXL ANトンネル400から入力されたフレームのVXL ANフレームから追加ヘッダの除去等を行い、デカプセル化する。

[0038] UDPポート情報記憶部266は、トンネル処理部260がVXL ANトンネル400を介して送受信を行うためのUDPポート情報を保持する。UDPポート情報記憶部266に保持されるUDPポート番号は、コマンド処理部251から発行される変更コマンドにより変更される。なお、UDPポート情報記憶部266は、複数のトンネル処理部260で共有することが可能である。すなわち、複数のVN Iに対して1個のUDPポート番号を使用することが可能である。

[0039] トンネル化部（カプセル化部）267はVXL ANトンネル400へ出力する際に、図4に示すようにオリジナルフレームに追加ヘッダを付加してVXL ANカプセル化処理を行う。

[0040] コマンド処理部251はユーザからの操作を受け付け、トンネル処理部260に対して、UDPポート番号情報の変更コマンド（UDPポート番号変更通知）等を出力し、UDPポート情報記憶部266の内容等を書き換える動作を行う。

[0041] なお、本実施形態では、コマンド処理部251を設けて、UDPポート情報記憶部266を変更する構成を採用しているが、UDPポート番号変更のための特殊なフレーム／パケットを送受信することにより変更を行う方法も考えられる。

[0042] なお、図3、図5に示した仮想スイッチ200の各部（処理手段）は、仮想スイッチ200に搭載されるコンピュータに、そのハードウェアを用いて

、上記した各処理を実行させるコンピュータプログラムにより実現することもできる。

[0043] 続いて、本実施形態の動作について図面を参照して詳細に説明する。図6は、本発明の第1の実施形態の動作を表したシーケンス図である。

[0044] 初期状態として、図7に示すように、UDPポート情報記憶部266に、VLAN通信に用いるUDPポート番号として、#1が設定されているものとする。具体的には、図7で示すように、この時点では、UDPポート番号#1のみがフレーム受信可能、かつ、フレーム送信時の宛先UDPポート番号としてUDPポート番号#1が設定されている。

[0045] この状態で、図6に示すように、ユーザは、すべての仮想スイッチ200A~200Cの設定を切り替えるため、各仮想スイッチ200に対してUDPポート番号の変更を指示したものとする。前記指示を受けたコマンド処理部251は、トンネル処理部260に対して、UDPポート番号を#1から#2へ変更することを指示するUDPポート番号変更コマンド（第1の指示）を送信する（図6のS001）。

[0046] 前記UDPポート番号の変更コマンド（第1の指示）を受信したトンネル処理部260のUDPポート情報記憶部266の内容は、図8に示すとおり書き換えられる（図6のS002）。図8の例では、図7のエントリに、受信可能ポートとしてUDPポート番号#2を指定したエントリが追加されている。なお、この時点では、UDPポート番号#2は、送信時宛先に使用するUDPポート情報とはなっていない（送信時宛先使用フィールドがNO）。

[0047] 次に、ユーザは、仮想スイッチ200に対して送信時宛先ポートを新UDPポート番号#2に切替えるコマンドを入力する。前記指示を受けたコマンド処理部251は、トンネル処理部260に対して、送信時宛先に使用するUDPポート番号を#1から#2へ変更することを指示するUDPポート番号切替コマンド（第2の指示）を送信する（図6のS003）。

[0048] 前記UDPポート番号の切替コマンド（第2の指示）を受信したトンネル

処理部260のUDPポート情報記憶部266の送信時宛先使用フィールドが、図9に示すとおり書き換えられる(図6のS004)。図9の例では、図8の1番目のエントリの送信時宛先使用フィールドがYESからNOに書き換えられ、図8の2番目のエントリの送信時宛先使用フィールドがNOからYESに書き換えられている。この処理は、送信時宛先使用のUDPポート番号#1から#2に切り替える処理に相当する。なお、この時点においては、受信可能UDPポートとして、UDPポート番号#1、#2の双方が設定された状態が維持されている(受信可能ポートフィールドが共にYES)。

[0049] 次に、ユーザは、仮想スイッチ200に対して変更前UDPポート番号#1の開放を指示するコマンドを入力する。前記指示を受けたコマンド処理部251は、トンネル処理部260に対して、変更前UDPポート番号解放コマンド(第3の指示)を送信する(図6のS005)。

[0050] 前記変更前UDPポート番号解放コマンド(第3の指示)を受信したトンネル処理部260のUDPポート情報記憶部266の内容は、図10に示すとおり書き換えられる(図6のS006)。図10の例では、図9の1番目のエントリの受信可能UDPポートから、UDPポート番号#1が外されている(YES→NOに変更)。これにより、UDPポート番号#1が解放されたことになる。なお、図10の例では、各フィールドにNOを設定した状態で1番目のエントリを残しているが、このエントリは、再度、UDPポート#2から#1への変更時に再使用される。もちろん、その必要性が無い場合は、1番目のエントリ自体を削除してしまってもよい。

[0051] 以上説明したように、本実施形態によれば、システムを構成する各仮想スイッチに対してUDPポート番号の変更設定タイミングがずれてしまった場合でも、ネットワークの通信断の発生を抑制することができる。

[0052] 以上、本発明の各実施形態を説明したが、本発明は、上記した実施形態に限定されるものではなく、本発明の基本的技術的思想を逸脱しない範囲で、更なる変形・置換・調整を加えることができる。例えば、各図面に示したネ

ットワーク構成、各要素の構成、メッセージの表現形態は、本発明の理解を助けるための一例であり、これらの図面に示した構成に限定されるものではない。

[0053] 例えば、上記した第1の実施形態では、ユーザからUDPポート番号の変更コマンド（第1の指示）、前記UDPポート番号の切替コマンド（第2の指示）、変更前UDPポート番号解放コマンド（第3の指示）の発行指示がなされるものとして説明したが、前記UDPポート番号の切替と、変更前UDPポート番号解放は、所定期間の経過後に実行される自動処理とすることができる。また、変更前UDPポート番号解放が必要ない場合は、変更前UDPポート番号解放を省略することもできる。

[0054] また、ユーザに代わってフロー制御装置100が、UDPポート番号の変更コマンド（第1の指示）、前記UDPポート番号の切替コマンド（第2の指示）、変更前UDPポート番号解放コマンド（第3の指示）の発行指示をコマンド処理部251に送るようにしてもよい。この場合、フロー制御装置100が、仮想マシンに適用される通信ポリシー等を参照して、UDPポート番号の変更することが可能となる。

[0055] また例えば、上記した第1の実施形態では、VXLANで接続されることを想定して、仮想スイッチ200を例示して説明したが、本発明が適用されるスイッチは、仮想スイッチに限定されない。例えば、一部のスイッチ又は全部が物理スイッチであってもよい。

[0056] また例えば、上記した第1の実施形態では、トンネル処理部260は、VXLANトンネル400から入力されたVXLANカプセル化されたフレームをフローエントリ検索部201に渡し、フロー処理部204の指示に従い、必要に応じVXLANデカプセル化するものとして説明したが、トンネル処理部260が、VXLANデカプセル化してから、フレームをフローエントリ検索部201に渡す構成とすることもできる。同様に、仮想マシンから受信したフレームを送信する際も、入力されたフレームをVXLANカプセル化してから、フローエントリ検索部201に渡す構成とすることもできる。

。

[0057] 最後に、本発明の好ましい形態を要約する。

[第1の形態]

(上記第1の視点によるスイッチ参照)

[第2の形態]

第1の形態のスイッチにおいて、

前記コマンド処理部は、

第1の指示の受信を契機に、変更前後のフロー識別情報でのフレーム受信を開始するよう前記トンネル処理部に指示し、

第2の指示の受信を契機に、送信フレームへの変更後のフロー識別の格納を開始するよう前記トンネル処理部に指示するスイッチ。

[第3の形態]

第1又は第2の形態のスイッチにおいて、

前記コマンド処理部は、

第3の指示の受信を契機に、変更前のフロー識別情報でのフレーム受信を終了するよう前記トンネル処理部に指示するスイッチ。

[第4の形態]

第1から第3いずれかの形態のスイッチにおいて、

前記フロー識別情報は、UDP (User Datagram Protocol) ポート番号であるスイッチ。

[第5の形態]

第1から第4いずれかの形態のスイッチにおいて、

さらに、デカプセル化後の受信フレーム及びカプセル化前の送信フレームの処理を行うオープンフロースイッチ機能部を備えるスイッチ。

[第6の形態]

第5の形態のスイッチにおいて、

前記デカプセル化後の受信フレーム及びカプセル化前の送信フレームの処理に関する指示を受け取る制御装置から、前記フロー識別情報の変更指示を

受信するスイッチ。

[第7の形態]

(上記第2の視点によるオーバーレイネットワークシステム参照)

[第8の形態]

(上記第3の視点による通信方法参照)

[第9の形態]

(上記第4の視点によるプログラム参照)

なお、上記第7～第9の形態は、第1の形態と同様に、第2～第6の形態に展開することが可能である。

[0058] なお、上記の特許文献および非特許文献の各開示を、本書に引用をもって繰り込むものとする。本発明の全開示（請求の範囲を含む）の枠内において、さらにその基本的技術思想に基づいて、実施形態ないし実施例の変更・調整が可能である。また、本発明の開示の枠内において種々の開示要素（各請求項の各要素、各実施形態ないし実施例の各要素、各図面の各要素等を含む）の多様な組み合わせ、ないし選択が可能である。すなわち、本発明は、請求の範囲を含む全開示、技術的思想にしたがって当業者であればなし得るであろう各種変形、修正を含むことは勿論である。特に、本書に記載した数値範囲については、当該範囲内に含まれる任意の数値ないし小範囲が、別段の記載のない場合でも具体的に記載されているものと解釈されるべきである。

符号の説明

- [0059] 100 フロー制御装置
200、200A～200C 仮想スイッチ
200a スイッチ
201 フローエントリ検索部
202 フローエントリ記憶部
203 フローエントリ処理部
204 フロー処理部
205 フロー制御装置通信部

- 250 VXLAN処理部
- 251、251a コマンド処理部
- 260、260a、260-1、260-2 トンネル処理部
- 264 VNI情報記憶部
- 265 トンネル解除部（デカプセル化部）
- 266 UDPポート情報記憶部
- 267 トンネル化部（カプセル化部）
- 300、300A～300E 仮想マシン
- 400、400AB-1～400BC-2 VXLANトンネル
- 500 基盤ネットワーク
- 601、602 UDPポート番号情報

請求の範囲

- [請求項1] オーバーレイネットワークのトンネルエンドポイントとして、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネル処理部と、
- カプセル化されたフレームを識別するためのフロー識別情報の変更指示の受信を契機に、前記トンネル処理部に対し、変更前後のフロー識別情報を持つフレームの受信を指示した後に、送信フレームに設定するフロー識別情報の変更を指示するコマンド処理部と、
- を備えたことを特徴とするスイッチ。
- [請求項2] 前記コマンド処理部は、
- 第1の指示の受信を契機に、変更前後のフロー識別情報でのフレーム受信を開始するよう前記トンネル処理部に指示し、
- 第2の指示の受信を契機に、送信フレームへの変更後のフロー識別の格納を開始するよう前記トンネル処理部に指示する請求項1のスイッチ。
- [請求項3] 前記コマンド処理部は、
- 第3の指示の受信を契機に、変更前のフロー識別情報でのフレーム受信を終了するよう前記トンネル処理部に指示する請求項1又は2のスイッチ。
- [請求項4] 前記フロー識別情報は、UDP (User Datagram Protocol) ポート番号である請求項1から3いずれかのスイッチ。
- [請求項5] さらに、デカプセル化後の受信フレーム及びカプセル化前の送信フレームの処理を行うオープンフロースイッチ機能部を備える請求項1から4いずれかのスイッチ。
- [請求項6] 前記デカプセル化後の受信フレーム及びカプセル化前の送信フレームの処理に関する指示を受け取る制御装置から、前記フロー識別情報の変更指示を受信する請求項5のスイッチ。

[請求項7] オーバーレイネットワークのトンネルエンドポイントとして、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネル処理部と、

 カプセル化されたフレームを識別するためのフロー識別情報の変更指示の受信を契機に、前記トンネル処理部に対し、変更前後のフロー識別情報を持つフレームの受信を指示した後に、送信フレームに設定するフロー識別情報の変更を指示するコマンド処理部と、

 を備えたスイッチを用いて構成されたオーバーレイネットワークシステム。

[請求項8] オーバーレイネットワークのトンネルエンドポイントとして、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネル処理部を備えたスイッチにおいて、

 カプセル化されたフレームを識別するためのフロー識別情報の変更指示の受信を契機に、前記トンネル処理部に対し、変更前後のフロー識別情報を持つフレームの受信を指示するステップと、

 前記指示の後に、所定の契機で、送信フレームに設定するフロー識別情報の変更を指示するステップと、

 を含む通信方法。

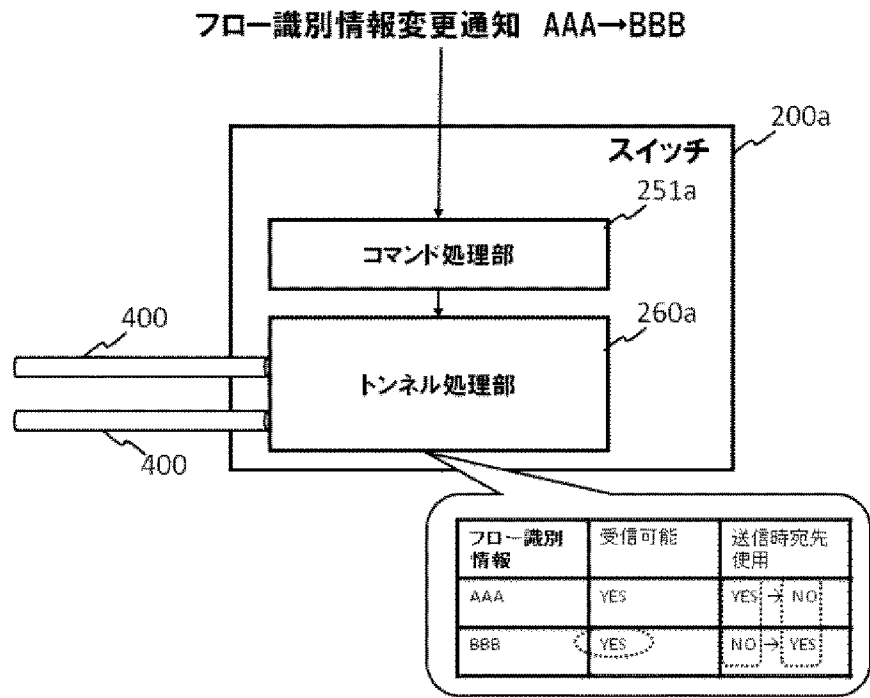
[請求項9] オーバーレイネットワークのトンネルエンドポイントとして、対向するトンネルエンドポイントと送受信するフレームのカプセル化とデカプセル化とを行うトンネル処理部を備えたスイッチに搭載されたコンピュータに、

 カプセル化されたフレームを識別するためのフロー識別情報の変更指示の受信を契機に、前記トンネル処理部に対し、変更前後のフロー識別情報を持つフレームの受信を指示する処理と、

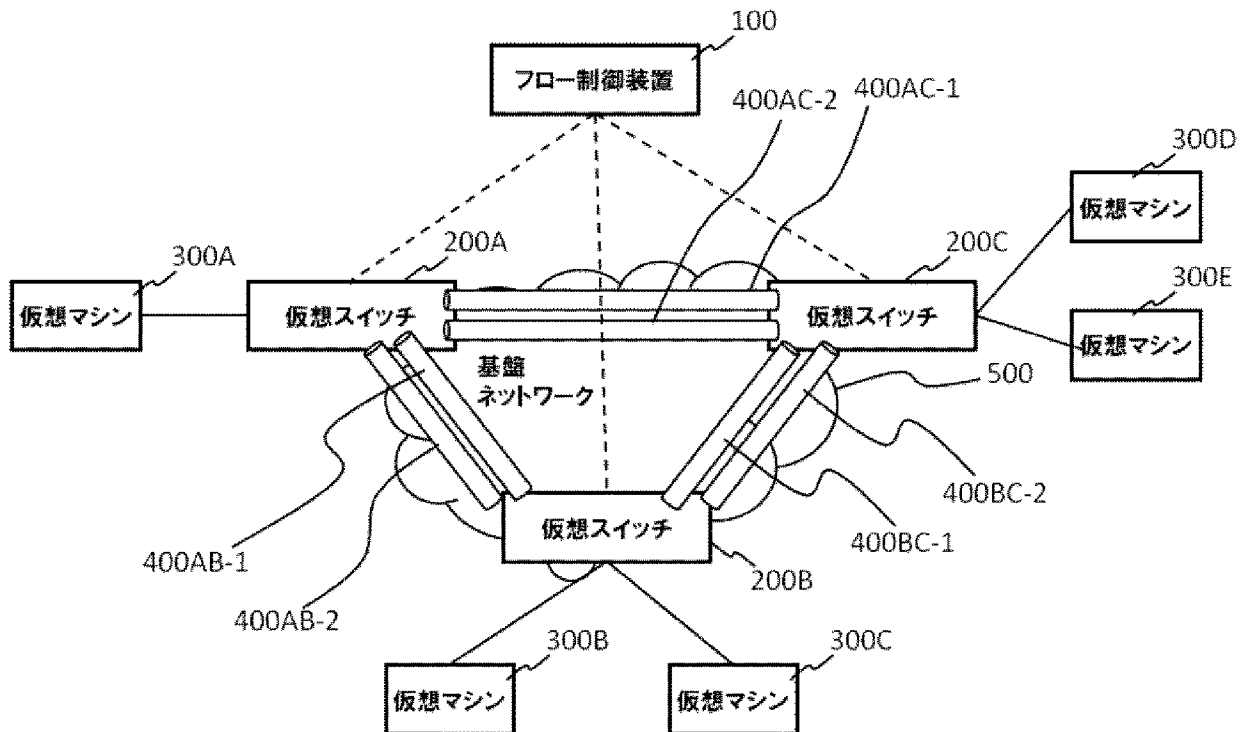
 前記指示の後に、所定の契機で、送信フレームに設定するフロー識別情報の変更を指示する処理と、

 を実行させるプログラム。

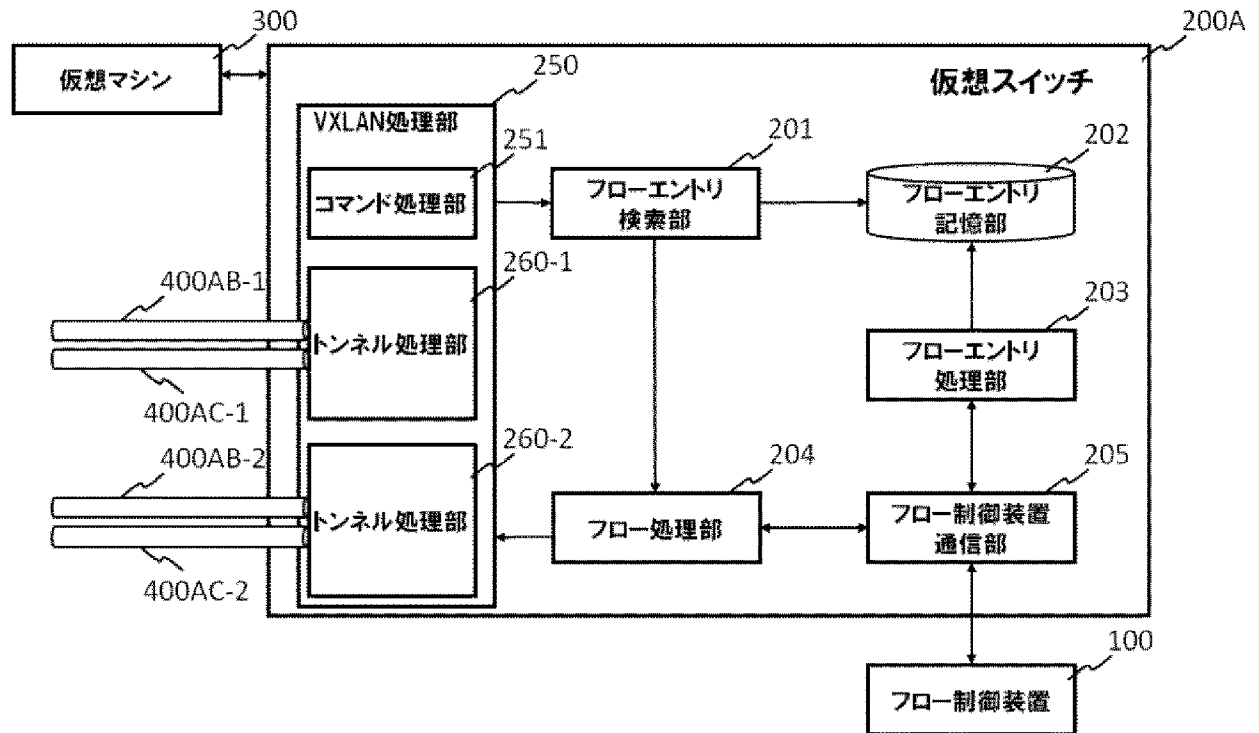
[図1]



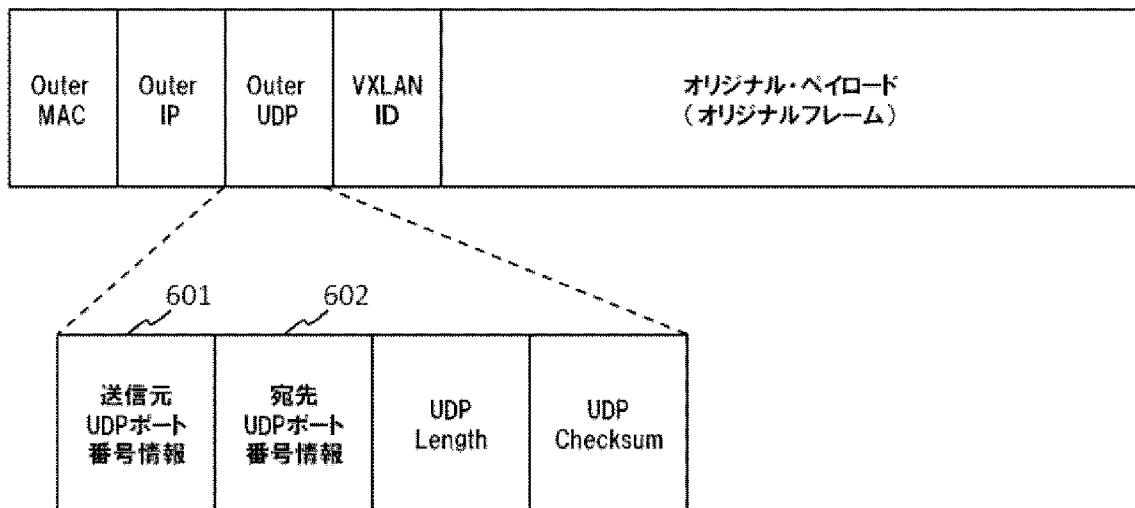
[図2]



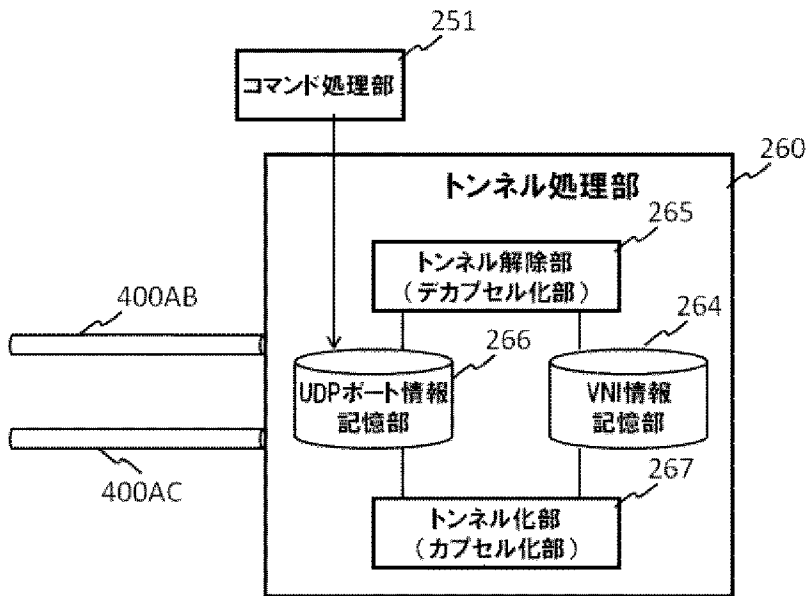
[図3]



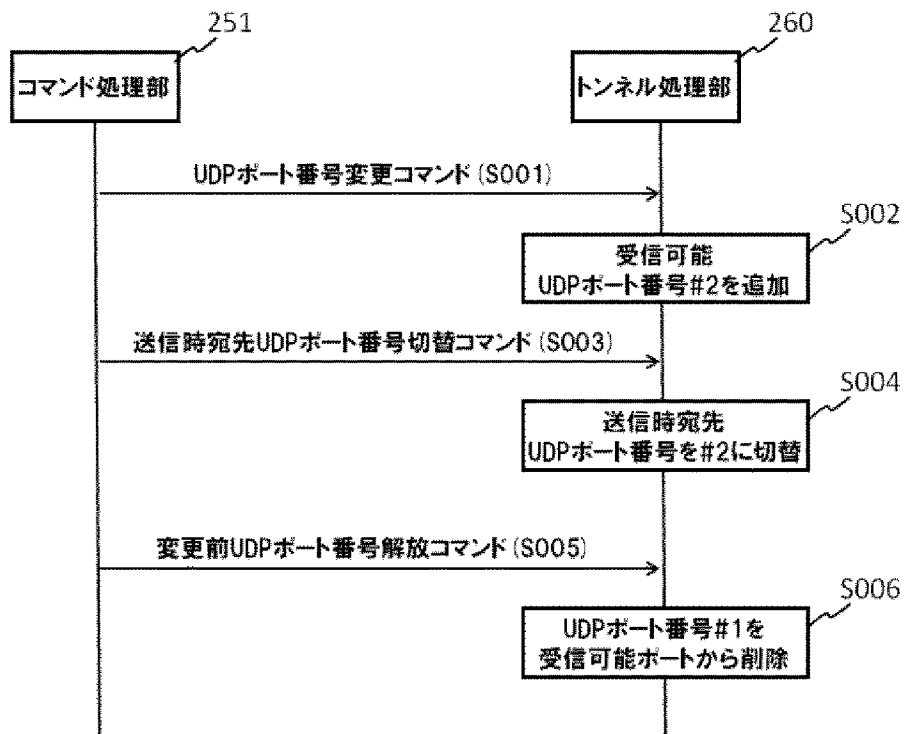
[図4]



[図5]



[図6]



[図7]

ポート番号	受信可能ポート	送信時宛先使用
#1	YES	YES

[図8]

ポート番号	受信可能ポート	送信時宛先使用
#1	YES	YES
#2	YES	NO

[図9]

ポート番号	受信可能ポート	送信時宛先使用
#1	YES	NO
#2	YES	YES

[図10]

ポート番号	受信可能ポート	送信時宛先使用
#1	NO	NO
#2	YES	YES

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2015/071611

<p>A. CLASSIFICATION OF SUBJECT MATTER <i>H04L12/70(2013.01) i, H04L12/717(2013.01) i</i></p> <p>According to International Patent Classification (IPC) or to both national classification and IPC</p>											
<p>B. FIELDS SEARCHED</p> <p>Minimum documentation searched (classification system followed by classification symbols) <i>H04L12/70, H04L12/717</i></p> <p>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched</p> <table style="width:100%; border:none;"> <tr> <td style="width:33%;"><i>Jitsuyo Shinan Koho</i></td> <td style="width:33%;"><i>1922-1996</i></td> <td style="width:33%;"><i>Jitsuyo Shinan Toroku Koho</i></td> <td style="width:33%;"><i>1996-2015</i></td> </tr> <tr> <td><i>Kokai Jitsuyo Shinan Koho</i></td> <td><i>1971-2015</i></td> <td><i>Toroku Jitsuyo Shinan Koho</i></td> <td><i>1994-2015</i></td> </tr> </table> <p>Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)</p>			<i>Jitsuyo Shinan Koho</i>	<i>1922-1996</i>	<i>Jitsuyo Shinan Toroku Koho</i>	<i>1996-2015</i>	<i>Kokai Jitsuyo Shinan Koho</i>	<i>1971-2015</i>	<i>Toroku Jitsuyo Shinan Koho</i>	<i>1994-2015</i>	
<i>Jitsuyo Shinan Koho</i>	<i>1922-1996</i>	<i>Jitsuyo Shinan Toroku Koho</i>	<i>1996-2015</i>								
<i>Kokai Jitsuyo Shinan Koho</i>	<i>1971-2015</i>	<i>Toroku Jitsuyo Shinan Koho</i>	<i>1994-2015</i>								
<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p> <table border="1" style="width:100%; border-collapse: collapse;"> <thead> <tr> <th style="width:10%;">Category*</th> <th style="width:70%;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="width:20%;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td align="center">A</td> <td>JP 2013-5110 A (NTT Communications Corp.), 07 January 2013 (07.01.2013), claim 1 (Family: none)</td> <td align="center">1-9</td> </tr> <tr> <td align="center">A</td> <td>JP 2013-38715 A (NTT Communications Corp.), 21 February 2013 (21.02.2013), paragraphs [0075] to [0082] (Family: none)</td> <td align="center">1-9</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	A	JP 2013-5110 A (NTT Communications Corp.), 07 January 2013 (07.01.2013), claim 1 (Family: none)	1-9	A	JP 2013-38715 A (NTT Communications Corp.), 21 February 2013 (21.02.2013), paragraphs [0075] to [0082] (Family: none)	1-9
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.									
A	JP 2013-5110 A (NTT Communications Corp.), 07 January 2013 (07.01.2013), claim 1 (Family: none)	1-9									
A	JP 2013-38715 A (NTT Communications Corp.), 21 February 2013 (21.02.2013), paragraphs [0075] to [0082] (Family: none)	1-9									
<p><input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.</p>											
<table style="width:100%; border:none;"> <tr> <td style="width:50%; vertical-align: top;"> <p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> </td> <td style="width:50%; vertical-align: top;"> <p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p> </td> </tr> </table>			<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>							
<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>										
<p>Date of the actual completion of the international search 02 October 2015 (02.10.15)</p>		<p>Date of mailing of the international search report 13 October 2015 (13.10.15)</p>									
<p>Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan</p>		<p>Authorized officer</p> <p>Telephone No.</p>									

A. 発明の属する分野の分類（国際特許分類（IPC）） Int.Cl. H04L12/70(2013.01)i, H04L12/717(2013.01)i		
B. 調査を行った分野 調査を行った最小限資料（国際特許分類（IPC）） Int.Cl. H04L12/70, H04L12/717		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2015年 日本国実用新案登録公報 1996-2015年 日本国登録実用新案公報 1994-2015年		
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2013-5110 A (エヌ・ティ・ティ・コミュニケーションズ株式会社) 2013.01.07, 請求項1 (ファミリーなし)	1-9
A	JP 2013-38715 A (エヌ・ティ・ティ・コミュニケーションズ株式会社) 2013.02.21, 【0075】-【0082】 (ファミリーなし)	1-9
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す） 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願日の後に公表された文献 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」 同一パテントファミリー文献		
国際調査を完了した日 02.10.2015	国際調査報告の発送日 13.10.2015	
国際調査機関の名称及びあて先 日本国特許庁 (ISA/J P) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 安藤 一道 電話番号 03-3581-1101 内線 3596	5X 3048