

(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
2013年8月8日(08.08.2013)



(10) 国際公開番号
WO 2013/115177 A1

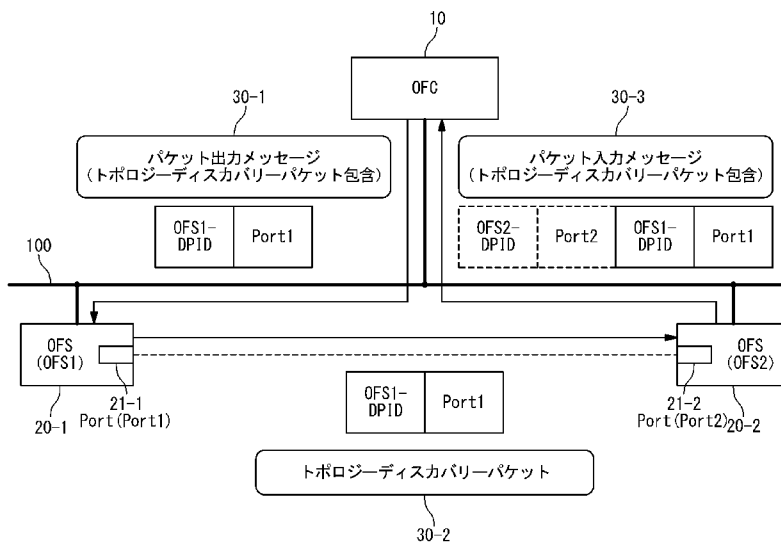
- (51) 国際特許分類:
H04L 12/70 (2013.01)
- (21) 国際出願番号: PCT/JP2013/051889
- (22) 国際出願日: 2013年1月29日(29.01.2013)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2012-016225 2012年1月30日(30.01.2012) JP
- (71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 Tokyo (JP).
- (72) 発明者; および
- (71) 出願人 (米国についてのみ): 高飛 (GAO Fei) [CN/JP]; 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP).
- (74) 代理人: 工藤 実 (KUDOH Minoru); 〒1400013 東京都品川区南大井六丁目24番10号カドヤビル6階 Tokyo (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, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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, SZ, TZ, UG, ZM, ZW), ユーラシア (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, 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: NETWORK SYSTEM AND TOPOLOGY MANAGEMENT METHOD

(54) 発明の名称: ネットワークシステム、及びトポロジー管理方法



- 30-1 Packet output message (containing topology discovery packet)
- 30-2 Topology discovery packet
- 30-3 Packet input message (containing topology discovery packet)

(57) Abstract: The purpose is to reduce the load on the switches on a secure channel network, while a controller is maintaining or updating the physical topology of the switches under conditions of high delay on the network between the switches, in an open network or the like. Specifically, the controller, which establishes flow entries defining actions and rules for uniformly controlling flows of packets for each of a plurality of switches, establishes for each of the plurality of switches a cyclic flow entry which is deleted in the event that cyclic packets sent and received among the plurality of switches have ceased to arrive. Then, in the event that a notification is received from the switches, to the effect that the cyclic flow entries have been deleted, a malfunction among the plurality of switches is detected.

(57) 要約: オープンフローネットワーク等において、コントローラがスイッチ間のネットワークが高遅延な状況でのスイッチの物理トポロジーを維持・更新する際のセキュア

チャンネルネットワーク上のスイッチの負荷を低減する。具体的には、複数のスイッチの各々に対して、パケットをフローとして一律に制御するためのルールとアクションとが定義されたフローエントリを設定するコントローラが、複数のスイッチの各々に対して、複数のスイッチ間で相互に送受信される巡回パケットが到着しなくなった場合に削除される巡回用フローエントリを設定する。そして、各スイッチから、巡回用フローエントリを削除した旨の通知を受信した場合、複数のスイッチ間の障害を検出する。

WO 2013/115177 A1

明 細 書

発明の名称： ネットワークシステム、及びトポロジー管理方法 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を構成するスイッチの物理トポロジーの管理方法に関する。

背景技術

[0002] [CD分離型ネットワークの説明]

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（データプレーン）を制御するCD（C：コントロールプレーン/D：データプレーン）分離型ネットワークが提案されている。

[0003] CD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（OpenFlow）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。オープンフロー技術の詳細については、非特許文献1（OpenFlow Switch Specification Version 1.1.0）に記載されている。なお、オープンフローネットワークは一例に過ぎない。

[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、オープンフローコントローラ（OFC：OpenFlow Controller）が、オープンフロースイッチ（OFS：OpenFlow Switch）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であるオープンフローメッセージ（OpenFlow Message）を用いてスイッチを制御するためのセキュアチャンネル（Secure Channel）により接続されている。

[0005] 以下、記載の簡略化のため、オープンフローコントローラ（OFC）を「

コントローラ（OFC）」と表記し、オープンフロースイッチ（OFS）を「スイッチ（OFS）」と表記する。また、コントローラとスイッチとの間で、セキュアチャンネルにより構成されたネットワークを、「セキュアチャンネルネットワーク」と呼ぶ。

[0006] オープンフローネットワークにおけるスイッチ（OFS）とは、オープンフローネットワークを構成し、コントローラ（OFC）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。

[0007] 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。一方、フレームは、「Ethernet（登録商標）」のPDUである。

[0008] 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。

[0009] 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（dst：Destination Address）、送信元アドレス（src：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、入力ポート（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。

[0010] 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」、「ヘッダを書き換える」といった動作を示す。例えば、スイッチ（OFS）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場合は、パケットを破棄する。或いは、スイッチ（OFS）は、フローエントリのアクション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。

[0011] スイッチ（OFS）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。

[0012] なお、オープンフローでは、コントローラ（OFC）は、スイッチ（OFS）側のフローエントリを把握・管理するため、スイッチ（OFS）側のフローエントリのコピー（複写）を保持している。例えば、コントローラ（OFC）は、スイッチ（OFS）のフローテーブルと同じフローテーブルを保持している。

[0013] [オープンフローネットワークにおけるトポロジー検出の現状]

オープンフローネットワークでは、コントローラ（OFC）は、LLDP（Link Layer Discovery Protocol）やOFDP（OpenFlow Discovery Protocol）等のトポロジーディスカバリープロトコル（Topology Discovery Protocol）を利用し、隣接スイッチ（OFS）間の接続情報を収集する。なお、OFDPは、LLDPを拡張したオープンフロー用のトポロジー検出プロトコルである。

[0014] また、コントローラ（OFC）は、収集した隣接スイッチ（OFS）間の接続情報に基づいて、オープンフローネットワーク全体のトポロジーを検出

する。オープンフローネットワーク全体のトポロジーの例として、スイッチ（OFS）間の接続情報や、接続用ポートの情報等が考えられる。

[0015] コントローラ（OFC）は、検出したスイッチ（OFS）の接続情報に基づいて、スイッチ（OFS）間の相互接続ポートや、各スイッチ（OFS）と通信端末機器（ホスト）との接続ポートに、適切なフローエントリを設定することで、オープンフローネットワーク内の通信を実現する。

[0016] 更に、運用中に、相互接続ポート間のリンクダウン以外の障害を検出するため、コントローラ（OFC）は、各スイッチ（OFS）に対して、定期的に、LLDPやOFDP等のトポロジーディスカバリーパケット（TDP：Topology Discovery Packet）を送信するように指示する。

[0017] 各スイッチ（OFS）は、ポート状態メッセージ（PSM：Port Status Message）をコントローラ（OFC）に送信する。

[0018] コントローラ（OFC）は、各スイッチ（OFS）から送信されたポート状態メッセージ（PSM）の通知を受信する。これにより、相互接続ポート間のリンクダウンの障害の発生時にはリンクダウンの障害を検出する。

[0019] 以上のように、コントローラ（OFC）は、検出したオープンフローネットワークトポロジーを維持・更新する。

[0020] [既存のオープンフローネットワークにおけるトポロジー検出・維持の手順]

図1を参照して、既存のオープンフローネットワークにおけるトポロジー検出・維持の手順について説明する。

[0021] ここでは、説明の簡略化のため、1台のコントローラ（OFC）と、2台のスイッチ（OFS）で構成されるオープンフローネットワークを例に説明する。

[0022] 既存のオープンフローネットワークは、コントローラ（OFC）10と、スイッチ（OFS）20-1と、スイッチ（OFS）20-2を含む。

[0023] コントローラ（OFC）10は、セキュアチャンネルネットワーク100

を介して、スイッチ（OFS）20-1及びスイッチ（OFS）20-2と接続されている。セキュアチャンネルネットワーク100には、コントローラ（OFC）とスイッチ（OFS）間のコントロール信号が流れる。コントロール信号の1つに、オープンフローメッセージがある。

[0024] コントローラ（OFC）10は、パケット出力メッセージ（POM：Packet-Output Message）30-1にトポロジーディスカバリーパケット（TDP）30-2を格納し、セキュアチャンネルネットワーク100を介して、パケット出力メッセージ（POM）30-1をスイッチ（OFS）20-1に送信する。なお、パケット出力メッセージ（POM）30-1は、オープンフローメッセージの1つである。

[0025] スイッチ（OFS）20-1は、接続ポート21-1を備える。スイッチ（OFS）20-2は、接続ポート21-2を備える。接続ポート21-1及び接続ポート21-2は、スイッチ（OFS）間の接続ポートである。接続ポート21-1及び接続ポート21-2は、スイッチ（OFS）20-1とスイッチ（OFS）20-2を接続する。

[0026] スイッチ（OFS）20-1は、パケット出力メッセージ（POM）30-1に格納されたトポロジーディスカバリーパケット（TDP）30-2を取得し、トポロジーディスカバリーパケット（TDP）30-2を接続ポート21-1に伝送する。

[0027] トポロジーディスカバリーパケット（TDP）30-2は、スイッチ（OFS）20-1の接続ポート21-1からスイッチ（OFS）20-2の接続ポート21-2に到達する。

[0028] スイッチ（OFS）20-2は、受信したトポロジーディスカバリーパケット（TDP）30-2をパケット入力メッセージ（PIM：Packet-Input Message）30-3に格納し、セキュアチャンネルネットワーク100を介して、パケット入力メッセージ（PIM）30-3をコントローラ（OFC）10に送信する。なお、パケット入力メッセージ（PIM）30-3は、オープンフローメッセージの1つである。

[0029] 以下に、具体的な処理について説明する。

[0030] [初期設定処理]

まず、トポロジー検出処理の開始前に行うべき初期設定処理について説明する。

[0031] コントローラ (OFC) 10は、各スイッチ (OFS) に、トポロジーディスカバリーパケット (TDP) 用のフローエントリを設定する。

[0032] コントローラ (OFC) 10は、このフローエントリの整合領域 (MF : Match Field) に、トポロジーディスカバリーパケット (TDP) 30-2がマッチする整合値 (MV : Match Value) を指定する。すなわち、整合領域 (MF) の整合値 (MV) がフローエントリのルールとなる。

[0033] コントローラ (OFC) 10は、このフローエントリのアクション領域 (AF : Action Field) に、アクション (Action) として、「コントローラ (OFC) 10に対して、トポロジーディスカバリーパケット (TDP) に関するパケット入力メッセージ (PIM) を送信する」と指定する。すなわち、アクション領域 (AF) のアクションがフローエントリのアクションとなる。

[0034] したがって、各スイッチ (OFS) は、受信したトポロジーディスカバリーパケット (TDP) 30-2が、上記のフローエントリにマッチした場合、コントローラ (OFC) 10に対して、トポロジーディスカバリーパケット (TDP) 30-2に関するパケット入力メッセージ (PIM) 30-3を送信することになる。

[0035] このとき、各スイッチ (OFS) は、このパケット入力メッセージ (PIM) 30-3のデータ領域 (DF : Data Field) に、トポロジーディスカバリーパケット (TDP) 30-2を格納する。

[0036] また、各スイッチ (OFS) は、このパケット入力メッセージ (PIM) 30-3に、当該スイッチ (OFS) の「DPID」と「Port」を格納する。

- [0037] 「DPID」(Delivery Point Identifier)は、スイッチ(OFS)の識別情報である。
- [0038] ここでは、スイッチ(OFS)20-1の「DPID」を、「OFS1-DPID」とする。また、スイッチ(OFS)20-2の「DPID」を、「OFS2-DPID」とする。
- [0039] 「Port」は、スイッチ(OFS)の接続ポートの識別情報である。
- [0040] ここでは、スイッチ(OFS)20-1の接続ポート21-1の「Port」を、「Port1」とする。また、スイッチ(OFS)20-2の接続ポート21-2の「Port」を、「Port2」とする。
- [0041] [トポロジー検出処理]
- 次に、初期設定処理の完了後に行われるトポロジー検出処理について説明する。
- [0042] コントローラ(OFC)10は、初期設定処理完了後、各スイッチ(OFS)のリンクアップしたポートに、セキュアチャンネルネットワーク経由で、オープンフローメッセージとして、トポロジーディスカバリーパケット(TDP)30-2に関するパケット出力メッセージ(POM)30-1を送信する。
- [0043] このとき、コントローラ(OFC)10は、送信するパケット出力メッセージ(POM)30-1のアクション領域(AF)に、アクションとして、「当該リンクアップしたポートからトポロジーディスカバリーパケット(TDP)を送信する」と指定する。
- [0044] また、コントローラ(OFC)10は、トポロジーディスカバリーパケット(TDP)30-2に、パケット出力メッセージ(POM)30-1の送信先であり、このトポロジーディスカバリーパケット(TDP)30-2の送信元となるスイッチ(OFS)の「DPID」と「Port」を指定する。
- [0045] 図1の例では、パケット出力メッセージ(POM)30-1の送信先であり、トポロジーディスカバリーパケット(TDP)30-2の送信元となる

スイッチ (OFS) は、スイッチ (OFS) 20-1である。

[0046] したがって、コントローラ (OFC) 10は、スイッチ (OFS) 20-1に対して、トポロジーディスカバリーパケット (TDP) 30-2に関するパケット出力メッセージ (POM) 30-1を送信する。

[0047] このとき、コントローラ (OFC) 10は、このパケット出力メッセージ (POM) 30-1のデータ領域 (DF) に、トポロジーディスカバリーパケット (TDP) 30-2を格納する。

[0048] また、コントローラ (OFC) 10は、このパケット出力メッセージ (POM) 30-1のアクション領域 (AF) に、アクションとして、「スイッチ (OFS) 20-1の接続ポート21-1からトポロジーディスカバリーパケット (TDP) 30-2を送信する」と指定する。

[0049] また、コントローラ (OFC) 10は、このトポロジーディスカバリーパケット (TDP) 30-2に、当該スイッチ (OFS) 20-1の「DPID」と「Port」として、「OFS1-DPID」と「Port1」を格納する。「OFS1-DPID」は、スイッチ (OFS) 20-1の識別情報である。このスイッチ (OFS) 20-1は、このトポロジーディスカバリーパケット (TDP) 30-2の送信元となるスイッチ (OFS) である。「Port1」は、接続ポート21-1の識別情報である。この接続ポート21-1は、トポロジーディスカバリーパケット (TDP) 30-2の送信ポートである。

[0050] スイッチ (OFS) 20-1は、コントローラ (OFC) 10から、上記のパケット出力メッセージ (POM) 30-1を受信する。

[0051] スイッチ (OFS) 20-1は、パケット出力メッセージ (POM) 30-1のデータ領域 (DF) に格納されたトポロジーディスカバリーパケット (TDP) 30-2を取得する。

[0052] スイッチ (OFS) 20-1は、このパケット出力メッセージ (POM) 30-1のアクション領域 (AF) に指定されたアクションに基づき、このトポロジーディスカバリーパケット (TDP) 30-2を接続ポート21-

1 から送信する。

[0053] 図1に示すように、スイッチ(OFS)20-1の接続ポート21-1は、スイッチ(OFS)20-2の接続ポート21-2と接続されている。したがって、スイッチ(OFS)20-1の接続ポート21-1から送信されたトポロジーディスカバリーパケット(TDP)30-2は、スイッチ(OFS)20-2の接続ポート21-2に到達する。

[0054] スイッチ(OFS)20-2は、接続ポート21-2で、トポロジーディスカバリーパケット(TDP)30-2を受信する。

[0055] スイッチ(OFS)20-2は、受信したトポロジーディスカバリーパケット(TDP)30-2が、初期設定処理で設定されたフローエントリにマッチした場合、コントローラ(OF C)10に対して、トポロジーディスカバリーパケット(TDP)30-2に関するパケット入力メッセージ(PIM)30-3を送信する。

[0056] このとき、スイッチ(OFS)20-2は、このパケット入力メッセージ(PIM)30-3に、当該スイッチ(OFS)20-2の「DPID」と「Port」として、「OFS2-DPID」と「Port2」を格納する。「OFS2-DPID」は、スイッチ(OFS)20-2の識別情報である。このスイッチ(OFS)20-2は、パケット入力メッセージ(PIM)30-3の送信元となるスイッチ(OFS)である。「Port2」は、接続ポート21-2の識別情報である。この接続ポート21-2は、トポロジーディスカバリーパケット(TDP)30-2の受信ポートである。

[0057] また、スイッチ(OFS)20-2は、このパケット入力メッセージ(PIM)30-3のデータ領域(DF)に、トポロジーディスカバリーパケット(TDP)30-2を格納する。

[0058] コントローラ(OF C)10は、上記のパケット入力メッセージ(PIM)30-3を受信する。

[0059] コントローラ(OF C)10は、このパケット入力メッセージ(PIM)30-3に格納された送信元のスイッチ(OFS)20-2の「OFS2-

DPID」と「Port 2」を取得する。

[0060] また、コントローラ（OFC）10は、このパケット入力メッセージ（PIM）30-3のデータ領域（DF）に格納されたトポロジーディスカバリーパケット（TDP）30-2を取得し、トポロジーディスカバリーパケット（TDP）30-2に格納されたスイッチ（OFS）20-1の「OFS 1-DPID」と「Port 1」を取得する。

[0061] これにより、コントローラ（OFC）10は、スイッチ（OFS）20-1の接続ポート21-1がスイッチ（OFS）20-2の接続ポート21-2と接続していることを検出する。

[0062] また、コントローラ（OFC）10は、上記と同様の仕組みで、逆方向にトポロジーディスカバリーパケット（TDP）30-2を流し、スイッチ（OFS）20-2の接続ポート21-2がスイッチ（OFS）20-1の接続ポート21-1と接続していることを検出する。

[0063] 更に、コントローラ（OFC）10は、運用中に、一定的なインターバル／リトライ回数で、上記のトポロジーディスカバリーパケット（TDP）30-2と、上記の仕組みを利用し、検出したトポロジーを維持・更新する。

[0064] なお、DPID、パケット入力メッセージ（PIM）30-3、パケット出力メッセージ（POM）30-1、フローエントリの仕様については、非特許文献1（OpenFlow Switch Specification Version 1.1.0）によって規定されている。

[0065] [検出したトポロジーの維持・更新における問題点]

上記の既存のオープンフローネットワークにおけるトポロジー検出・維持の手順は、オープンフローネットワークにおけるトポロジーの初期検出には有用であるが、検出したトポロジーを維持・更新する場合には、以下の（1）～（3）のような問題点がある。

[0066] （1）セキュアチャンネルネットワークへの負荷が高まる。

[0067] コントローラ（OFC）は、検出したトポロジーを維持・更新するために、各スイッチ（OFS）の全リンクアップしたポートに定期的にトポロジー

ディスカバリーパケット (TDP) をセキュアチャンネルネットワークに流す必要がある。

[0068] また、各スイッチ (OFS) も受信したトポロジーディスカバリーパケット (TDP) をセキュアチャンネルネットワーク経由でコントローラ (OFC) に流す必要がある。

[0069] 1つの双方向接続関係を維持・更新するため、一回のトポロジー走査で4個のトポロジーディスカバリーパケット (TDP) がセキュアチャンネルネットワーク上に流れる。

[0070] (2) 検出したトポロジーを維持・更新はコントローラ (OFC) 走査方法を利用するので、トポロジー変更があると、その変更を検出するために時間がかかる。

[0071] トポロジー変更を検出するタイミングは、運用中にコントローラ (OFC) がトポロジーディスカバリーパケット (TDP) を出すインターバル時間と、スイッチ (OFS) からのトポロジーディスカバリーパケット (TDP) のリプライタイムアウトに依存する。

[0072] インターバル時間とリプライタイムアウトが短ければ短いほど素早く検出できるが、その代わりに、セキュアチャンネルネットワーク上のスイッチ (OFS) への負荷が増大し、発生していないトポロジー変更を誤検出するリスクが増大する。

[0073] (3) 既存の多くのスイッチ (OFS) は、従来のレガシースイッチ (legacy switch) に、オープンフローに対応したファームウェア (firmware) を実装することで実現されている。

[0074] 例えば、スイッチ (OFS) は、オープンフローメッセージの処理において、「パケットがマッチしたフローエントリを検索し、マッチしたフローエントリに指定されたアクションに従ってパケットを処理する部分」以外の処理部分については、従来のレガシースイッチの汎用プロセッサを利用し、ソフトウェア (software) で実現している。

[0075] 汎用プロセッサのリソースは非常に限られているので、処理する必要があ

るオープンフローメッセージのロード（load：読込）が増大すると、計算リソース不足で、オープンフローメッセージの処理が遅延したり、オープンフローメッセージ自体が消失（ロスト）したりする可能性が大きい。

[0076] そのため、このようなスイッチ（OFS）を利用したオープンフローネットワークでは、上記の既存のオープンフローネットワークにおけるトポロジー検出・維持の手順を利用すると、コントローラ（OFC）は、発生していないトポロジー変更を誤検出し、システムを通信不能の状態に陥らせる可能性が非常に高い。

[0077] 関連する技術として、特許文献1（特開2003-143169号公報）にルーティングブリッジシステムが開示されている。このルーティングブリッジシステムでは、複数のノードをリング状に接続したリング状ネットワークが接続ノードを介して複数接続されたネットワークの、データリンクレイヤのアドレスを用いてデータの送信経路を決定する。各ノードは、所属するリング上の各ノード間のノード間スパニングツリーの作成を行うノードツリー作成部を備える。各接続ノードは、所属する各リング上の各ノード間のノード間スパニングツリーの作成を行うノードツリー作成部と、各リングを論理リンクに見立てた、各リング間のリング間スパニングツリーの作成を行うリング間ツリー作成部を備える。

[0078] また、特許文献2（特開2006-340361号公報）にホームネットワークのコネクショントポロジーの決定方法が開示されている。このホームネットワークのコネクショントポロジーの決定方法は、多重のハブ及びスイッチの周辺に配置されている複数のノードによって構成されているホームネットワークのコネクショントポロジーを決定するための方法である。この方法は、該複数のノードによってトポロジー決定メッセージをランダムなシーケンスにおいて1つずつ伝送し、該トポロジー決定メッセージの受信に基づいてノードの全てのコネクショントポロジーを1つずつ決定し、ノードの全てのローカルノードのリストを生成しかつ更新するステップを有している。トポロジー決定メッセージは、2つの連続するパケットを含んでいる。第1

の packets は、それをトポロジー決定 packet と識別する予め定めたペイロードを備えているブロードキャスト packet である。第 2 の packet は、同じコンテンツ及び非実在デスティネーション MAC アドレスを備えているユニキャスト packet である。

[0079] また、特許文献 3（特開 2008-172449 号公報）にネットワークの論理トポロジーを検出するトポロジー検出方法が開示されている。この関連技術では、通信装置は、ポート番号と送信元 MAC (Media Access Control) アドレスとを対応付けた転送情報を記録している。管理装置が、TTL (Time To Live) を 1 から順に 1 増分する毎に、該 TTL を含むトレース要求を通信装置へ送信する。次に、通信装置が、TTL = 1 のトレース要求に対して、転送情報を含むトレース応答を管理装置へ返信する。そして、管理装置が、受信したトレース応答から転送情報を取得し、ネットワークの装置トポロジーにおけるポートトポロジを導出する。

[0080] また、特許文献 4（特開 2009-111976 号公報）に無線メッシュ型ネットワークの構成及び最適化が開示されている。この関連技術では、ネットワーク設計ツールは、無線ネットワーク内でのノード及び機器の追加、除去、及び位置決めのための対話式グラフィックインターフェースと、閾値、ネットワークトポロジ選択、経路指定設定、及び無線メッシュ型ネットワーク内での通信経路及びスケジュールの生成及び最適化に関連するその他の構成パラメータを指定するための複数の対話式画面を含むメニューと、を提供する。ネットワーク設計ツールは、最適化規則の集合をユーザが入力するパラメータと共にネットワークモデルに自動的に適用して効率のよいネットワーク構成データを生成する。

先行技術文献

特許文献

[0081] 特許文献 1：特開 2003-143169 号公報

特許文献 2：特開 2006-340361 号公報

特許文献3：特開2008-172449号公報

特許文献4：特開2009-111976号公報

非特許文献

- [0082] 非特許文献1：“OpenFlow Switch Specification, Version 1.1.0 Implemented”, [online], February 28, 2011, インターネット (URL: <http://www.openflowswitch.org/documents/openflow-spec-v1.1.0.pdf>)

発明の概要

- [0083] 本発明の目的は、オープンフローネットワーク等において、コントローラがスイッチ間のネットワークが高遅延な状況でのスイッチの物理トポロジーの維持・更新を行うネットワークシステムを提供することである。
- [0084] 本発明に係るネットワークシステムは、パケットをフローとして一律に制御するためのルールとアクションとが定義されたフローエントリに従って、受信したパケットの処理を行う複数のスイッチと、複数のスイッチの各々に対して、フローエントリを設定を行うコントローラとを含む。コントローラは、複数のスイッチの各々に対して、複数のスイッチ間で相互に送受信される巡回パケットが到着しなくなった場合に削除される巡回用フローエントリを設定し、各スイッチから、巡回用フローエントリを削除した旨の通知を受信した場合、複数のスイッチ間の障害を検出する。
- [0085] 本発明に係るコントローラは、複数のスイッチの各々に対して、パケットをフローとして一律に制御するためのルールとアクションとが定義されたフローエントリを設定する機構と、複数のスイッチの各々に対して、複数のスイッチ間で相互に送受信される巡回パケットが到着しなくなった場合に削除される巡回用フローエントリを設定する機構と、各スイッチから、巡回用フローエントリを削除した旨の通知を受信した場合、複数のスイッチ間の障害を検出する機構とを備える。
- [0086] 本発明に係るトポロジー管理方法は、計算機により実施されるトポロジー

管理方法であって、コントローラとして、複数のスイッチの各々に対して、パケットをフローとして一律に制御するためのルールとアクションとが定義されたフローエントリを設定することと、複数のスイッチの各々に対して、複数のスイッチ間で相互に送受信される巡回パケットが到着しなくなった場合に削除される巡回用フローエントリを設定することと、各スイッチから、巡回用フローエントリを削除した旨の通知を受信した場合、複数のスイッチ間の障害を検出することを含む。

[0087] 本発明に係るプログラムは、上記のトポロジー管理方法における処理を、コントローラ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。

[0088] これにより、オープンフローネットワーク等において、トポロジーを維持・更新する際のセキュアチャンネルネットワーク上のスイッチの負荷を低減する。

図面の簡単な説明

[0089] [図1]既存のオープンフローネットワークにおけるトポロジー検出・維持の手順について説明するための図である。

[図2]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3]本発明に係るトポロジーキープアライブパケットの仕様について説明するための図である。

[図4]本発明に係るパケット出力メッセージについて説明するための図である。

[図5]本発明に係るトポロジーキープアライブフローエントリの仕様について説明するための図である。

[図6]本発明に係るフローエントリ変更メッセージの仕様について説明するための図である。

[図7]本発明に係るフローエントリ削除済メッセージの仕様について説明する。

[図8]本発明に係るネットワークシステムの実施例について説明するための図である。

発明を実施するための形態

[0090] 本発明は、CD分離型ネットワークを対象としている。ここでは、CD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。

[0091] <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。

[0092] 本実施形態では、検出したトポロジを維持・更新するために、「トポロジーキープアライブフローエントリ (TKAFE: Topology Keep Alive Flow Entry)」と、「トポロジーキープアライブパケット (TKAP: Topology Keep Alive Packet)」を規定して利用する。

[0093] トポロジーキープアライブパケット (TKAP) は、検出したトポロジを維持・更新するために予め規定された巡回パケットであり、トポロジーキープアライブフローエントリ (TKAFE) にマッチするパケットである。

[0094] 「トポロジーキープアライブフローエントリ (TKAFE)」や、「トポロジーキープアライブパケット (TKAP)」については後に詳述する。

[0095] [システム構成]

図2を参照して、本発明に係るネットワークシステムの構成例について説明する。

[0096] 本発明に係るネットワークシステムは、コントローラ (OFC) 10と、スイッチ (OFS) 20 (20-i、i=1~n: nは任意) を含む。

[0097] コントローラ (OFC) 10は、スイッチ (OFS) 20 (20-i、i=1~n) の各々 (以下、各スイッチ (OFS) 20) に対して、フローエントリを設定する。

[0098] コントローラ (OFC) 10は、セキュアチャンネルネットワーク100を介して、各スイッチ (OFS) 20と接続されている。セキュアチャネ

ルネットワーク100には、コントローラ（OFC）とスイッチ（OFS）間のコントロール信号が流れる。

[0099] 各スイッチ（OFS）20は、パケットを受信した際、コントローラ（OFC）10から設定されたフローエントリと受信パケットを比較し、受信パケットのヘッダ領域（HF：Header Field）の値が、フローエントリの整合領域（MF）の整合値（MV）とマッチした場合、このフローエントリのアクション領域（AF）に指定されたアクションを実行する。

[0100] [コントローラ（OFC）の構成]

次に、コントローラ（OFC）10の構成例について説明する。

[0101] コントローラ（OFC）10は、トポロジー管理部11と、トポロジー検出部12と、パケット梱包・振り分け処理部13と、パケット送信部14と、パケット受信部15を備える。

[0102] トポロジー管理部11は、トポロジー情報保存データベースを管理しており、コントローラ（OFC）10が検出したオープンフローネットワークのトポロジー情報をトポロジー情報保存データベースに保存する。また、トポロジー管理部11は、トポロジー検出部12からトポロジー情報の更新を指示された場合、指示の内容に従って、トポロジー情報保存データベースに保存されているトポロジー情報を更新する。

[0103] トポロジー検出部12は、既存のトポロジー検出機能を有し、トポロジーを検出・維持・更新する。すなわち、トポロジー検出部12は、従来通り、トポロジーディスカバリーパケット（TDP）を作成し、トポロジーを検出・維持・更新することができる。また、トポロジー検出部12は、検出したトポロジーを維持・更新するために、本実施形態において規定するトポロジーキープアライブフローエントリ（TKAFE）と、トポロジーキープアライブパケット（TKAP）を作成する。トポロジー検出部12は、作成した情報をパケット梱包・振り分け処理部13に渡す。更に、トポロジー検出部12は、パケット梱包・振り分け処理部13から受け取ったトポロジー変更通知メッセージを解析し、トポロジー管理部11に対し、トポロジー情報保

存データベースに保存されているトポロジー情報の更新を指示する。

[0104] パケット梱包・振り分け処理部13は、トポロジー検出部12から受け取った情報で実際のパケットを梱包し、パケット送信部14に渡す。例えば、パケット梱包・振り分け処理部13は、各スイッチ（OFS）20にトポロジーキープアライブフローエントリ（TKAFE）を追加（登録）するためのフローエントリ変更メッセージ（MFEM: Modify Flow Entry Message）を作成し、このフローエントリ変更メッセージ（MFEM）をパケット送信部14に渡す。或いは、パケット梱包・振り分け処理部13は、トポロジー検出部12から受け取ったトポロジーキープアライブパケット（TKAP）をパケット出力メッセージ（POM）に格納し、このパケット出力メッセージ（POM）をパケット送信部14に渡す。また、パケット梱包・振り分け処理部13は、パケット受信部15から受け取ったパケットを解析し、適切なモジュールに振り分ける。例えば、パケット梱包・振り分け処理部13は、パケット受信部15から受け取ったパケット入力メッセージ（PIM）を解析してトポロジー変更通知メッセージを取得／作成し、このトポロジー変更通知メッセージをトポロジー検出部12に渡す。

[0105] パケット送信部14は、セキュアチャンネルネットワーク100を介して、パケット梱包・振り分け処理部13から受け取ったパケットを各スイッチ（OFS）20に送信する。例えば、パケット送信部14は、パケット出力メッセージ（POM）を各スイッチ（OFS）20に送信する。

[0106] パケット受信部15は、セキュアチャンネルネットワーク100を介して、各スイッチ（OFS）20からパケットを受信し、受信パケットをパケット梱包・振り分け処理部13に渡す。例えば、パケット受信部15は、各スイッチ（OFS）20からパケット入力メッセージ（PIM）を受信し、このパケット入力メッセージ（PIM）をパケット梱包・振り分け処理部13に渡す。

[0107] [トポロジーキープアライブパケットの仕様]

図3を参照して、本発明に係るトポロジーキープアライブパケット（TKAP）の仕様について説明する。

[0108] トポロジーキープアライブパケット（TKAP）として、以下の条件を満たす任意のパケットを利用することが可能である。

[0109] （1）ブロードキャスト（BC）・マルチキャスト（MC）と一般的に使われているパケットタイプ以外の任意のタイプのパケット

[0110] （2）非特許文献1（OpenFlow Switch Specification Version 1. 1. 0）に規定されたフローエントリの整合領域（MF）、或いは、その一部を定義したパケット

[0111] 本実施形態では、パケットのオーバーヘッドを小さくするため、単純なMACヘッダフレーム（Ether Header Frame）を使用する。図3では、MACヘッダ（Ether Header）領域以外の領域の記載を省略する。

[0112] トポロジー検出部12は、MACヘッダ（Ether Header）領域のMAC送信元（Ether src）領域、MAC送信先（Ether dst）領域、及びフレームタイプ（Ether Type）領域の各々に、以下のような値を指定する。

[0113] トポロジー検出部12は、MAC送信元（Ether src）領域に、ワイルドカードの「*」を指定する。

[0114] トポロジー検出部12は、MAC送信先（Ether dst）領域に、「ブロードキャスト（BC）・マルチキャスト（MC）以外のMACアドレス」（Not BC/MC）を指定する。

[0115] トポロジー検出部12は、フレームタイプ（Ether Type）領域に、IANA（Internet Assigned Number Authority）に規定された「Experimental Type」の「0x1111」を指定する。

[0116] [パケット出力メッセージの仕様]

図4を参照して、本発明に係るパケット出力メッセージ（POM）の仕様

について説明する。

- [0117] パケット出力メッセージ (POM) は、ヘッダ領域 (HF) と、アクション領域 (AF) と、データ領域 (DF) を有する。
- [0118] ヘッダ領域 (HF) は、本メッセージがパケット出力メッセージ (POM) である旨を示す情報を格納するための領域である。
- [0119] アクション領域 (AF) は、データ領域 (DF) に格納されたパケットを送出する送信ポートを指定するための領域である。
- [0120] データ領域 (DF) は、トポロジーキープアライブパケット (TKAP) を格納するための領域である。
- [0121] ここでは、コントローラ (OFC) 10は、パケット出力メッセージ (POM) のヘッダ領域 (HF) に、「Type is Packet Out」を指定する。「Type is Packet Out」は、本メッセージがパケット出力メッセージ (POM) である旨を示す情報である。
- [0122] 各スイッチ (OFS) 20は、このパケット出力メッセージ (POM) を受信した場合、パケット出力メッセージ (POM) のデータ領域 (DF) に格納されたトポロジーキープアライブパケット (TKAP) を、アクション領域 (AF) に指定された送信ポートから送出する。
- [0123] [トポロジーキープアライブフローエントリの仕様]
図5を参照して、本発明に係るトポロジーキープアライブフローエントリ (TKAFE) の仕様について説明する。
- [0124] トポロジーキープアライブフローエントリ (TKAFE) は、整合領域 (MF) と、クッキー領域 (CF: Cookie Field) と、アクション領域 (AF) を有する。
- [0125] [整合領域]
まず、整合領域 (MF) について説明する。
- [0126] 整合領域 (MF) は、パケットのマッチングに使用されるエンタリー項目を示す領域である。エンタリー項目は、整合領域 (MF) のサブフィールドである。コントローラ (OFC) 10は、整合領域 (MF) に示されたエン

トリー項目毎に、整合値 (MV) を指定する。

[0127] なお、整合領域 (MF) のエントリー項目には、入力ポート (Ingress Port)、メタデータ (Meta data)、MAC送信元 (Ether src)、MAC送信先 (Ether dst)、フレームタイプ (Ether type)、仮想LAN識別子 (VLAN id)、仮想LAN優先度 (VLAN priority)、MPLSラベル (MPLS label)、MPLSトラフィッククラス (MPLS traffic class)、IP送信元 (IP src)、IP送信先 (IP dst)、IPプロトコル (IP Protocol)、IPサービスタイプ (IP ToS bits)、送信元ポート (TCP/UDP src Port)、及び送信先ポート (TCP/UDP dst Port) 等がある。

[0128] 入力ポート (Ingress Port) は、最初に本パケットを受信したポートを示す。メタデータ (Meta data) は、本パケットに関する情報を示す。MAC送信元 (Ether src)、MAC送信先 (Ether dst)、フレームタイプ (Ether type) は、それぞれ、送信元MACアドレス、送信先MACアドレス、上位層のプロトコルを識別するための16ビットの情報を示す。MPLS (Multi-Protocol Label Switching) は、ラベルスイッチング方式を用いたパケット転送技術である。MPLSラベル (MPLS label) は、IPヘッダの代わりに利用される固定長の識別標識を示す。MPLSトラフィッククラス (MPLS traffic class) は、異なるクラス又は優先順位の特定制と識別を行うための情報を示す。IP送信元 (IP src)、IP送信先 (IP dst)、IPプロトコル (IP Protocol)、IPサービスタイプ (IP ToS bits) は、それぞれ、送信元IPアドレス、送信先IPアドレス、IPプロトコル番号、TOSオクテット (IPパケットのヘッダ中に含まれる8ビットのオクテット部分) を示す。TOSオクテットは、通信のサービス品質 (QoS) を制御する目的で、処理の優先度を指定するために使う。送信元ポート (TCP/U

DP src Port)、及び送信先ポート(TCP/UDP dst Port)は、それぞれ、送信元ポート番号、送信先ポート番号を示す。

[0129] 本実施形態では、入力ポート(Ingress Port)とフレームタイプ(Ether type)のみ、マッチングに使用する。

[0130] コントローラ(OFC)10は、各スイッチ(OFS)20にトポロジーキープアライブフローエントリ(TKAFE)を設定する際、入力ポート(Ingress Port)とフレームタイプ(Ether type)以外の整合領域(MF)のエントリー項目の整合値(MV)に「ANY」を指定する。

[0131] 整合値(MV)が「ANY」である整合領域(MF)のエントリー項目は、任意の値(全ての値)にマッチするため、パケットのマッチングには使用されない。

[0132] これにより、各スイッチ(OFS)20は、受信パケットとトポロジーキープアライブフローエントリ(TKAFE)とを比較し、入力ポート(Ingress Port)及びフレームタイプ(Ether type)について、双方の値がマッチした場合、受信パケットとトポロジーキープアライブフローエントリ(TKAFE)がマッチしたと判断する。

[0133] ここでは、コントローラ(OFC)10は、入力ポート(Ingress Port)の整合値(MV)に、特定のポート(Specified Port)を指定する。この特定のポート(Specified Port)は、コントローラ(OFC)10が検出した各スイッチ(OFS)20のポートのうち、他の各スイッチ(OFS)20と繋がるポートを示す。

[0134] また、コントローラ(OFC)10は、フレームタイプ(Ether type)の整合値(MV)に「0x1111」を指定する。「0x1111」は、本実施形態におけるトポロジーキープアライブパケット(TKAP)のフレームタイプ(Ether type)を示す。

[0135] [クッキー領域]

次に、クッキー領域(CF)について説明する。

- [0136] クッキー領域 (C F) は、付加情報等のデータを一時保存しておくためのクッキー (C o o k i e) を示す領域である。クッキーには、クッキー値 (C V : C o o k i e V a l u e) が指定される。
- [0137] コントローラ (O F C) 10 は、クッキーのクッキー値 (C V) に、64 ビットの識別子を指定する。
- [0138] この64ビットの識別子は、トポロジーキープアライブフローエントリ (T K A F E) を設定したコントローラ (O F C) 10 が内部管理しているフローエントリを特定するための識別子である。オープンフローでは、コントローラ (O F C) 10 は、各スイッチ (O F S) 20 側のフローエントリを把握・管理するため、各スイッチ (O F S) 20 側のフローエントリのコピー (複写) を保持している。
- [0139] [アクション領域]
次に、アクション領域 (A F) について説明する。
- [0140] アクション領域 (A F) は、当該フローエントリに関するプロパティ項目 (p r o p e r t i e s) を示す領域である。プロパティ項目は、アクション領域 (A F) のサブフィールドである。コントローラ (O F C) 10 は、アクション領域 (A F) に示されたプロパティ項目毎に、特性を示す値 (特性値) を指定する。
- [0141] なお、アクション領域 (A F) のプロパティ項目には、優先度 (E n t r y P r i o r i t y) 、遊休時間 (I d l e t i m e) 、固定時間 (H a r d t i m e) 、フラグ (F l a g) 、アクション (A c t i o n) 等がある。
- [0142] コントローラ (O F C) 10 は、優先度 (E n t r y P r i o r i t y) に「0 x f f f f」を指定する。優先度 (E n t r y P r i o r i t y) は、フローエントリ自体の優先度を示す。「0 x f f f f」は、最も高い優先度 (最優先) を示す。すなわち、トポロジーキープアライブフローエントリ (T K A F E) は、最優先のフローエントリとなる。これにより、各スイッチ (O F S) 20 は、受信パケットに対して、トポロジーキープアライ

ブフローエントリ (TKAFE) を最優先に比較し、マッチした場合、トポロジーキープアライブフローエントリ (TKAFE) のアクションを実行する。

[0143] また、コントローラ (OFC) 10は、遊休時間 (Idle time) に「0x0003」を指定し、固定時間 (Hard time) に「0x0000」を指定する。時間の指定において、「0x0003」は「3秒」を示す。これにより、各スイッチ (OFS) 20は、トポロジーキープアライブパケット (TKAP) を受信してから、3秒以内に、次のトポロジーキープアライブパケット (TKAP) を受信しなかった場合、トポロジーキープアライブフローエントリ (TKAFE) がエクスパイア (expire: 期限満了、失効) したものと判断し、フローテーブルから当該トポロジーキープアライブフローエントリ (TKAFE) を削除 (エージアウト) する。

[0144] また、コントローラ (OFC) 10は、フラグ (Flag) に、「送信フロー削除フラグ (SFRF: Send Flow Removed Flag) = on」を指定する。すなわち、送信フロー削除フラグ (SFRF) を「on」(有効) にする。これにより、各スイッチ (OFS) 20は、エクスパイアしたトポロジーキープアライブフローエントリ (TKAFE) を削除した後、エクスパイアしたトポロジーキープアライブフローエントリ (TKAFE) を削除した旨を示すオープンフローメッセージをコントローラ (OFC) 10に送信する。本実施形態では、各スイッチ (OFS) 20は、エクスパイアしたトポロジーキープアライブフローエントリ (TKAFE) を削除した後、後述するフローエントリ削除済メッセージ (FERM: Flow Entry Removed Message) をコントローラ (OFC) 10に送信する。

[0145] また、コントローラ (OFC) 10は、アクション (Action) に、「マッチしたパケットを受信ポートに送り戻す動作」 (Send to import) を指定する。「マッチしたパケットを受信ポートに送り戻す動作」とは、接続ポートで受信したパケットが当該フローエントリにマッチ

した場合、当該パケットを、当該接続ポートに送り戻す動作を示す。これにより、各スイッチ（OFS）20は、トポロジーキープアライブフローエントリ（TKAFE）にマッチしたトポロジーキープアライブパケット（TKAP）を、受信ポートに送り戻す。したがって、トポロジーキープアライブパケット（TKAP）は、相互接続され、トポロジーキープアライブフローエントリ（TKAFE）が設定されたスイッチ（OFS）20間を巡回する。

[0146] [フローエントリ変更メッセージの仕様]

図6を参照して、コントローラ（OFC）10が各スイッチ（OFS）20にトポロジーキープアライブフローエントリ（TKAFE）を追加（登録）するためのフローエントリ変更メッセージ（MFEM）の仕様について説明する。

[0147] フローエントリ変更メッセージ（MFEM）は、ヘッダ領域（HF）と、整合領域（MF）と、クッキー領域（CF）と、アクション領域（AF）を有する。

[0148] ヘッダ領域（HF）は、本メッセージがフローエントリ変更メッセージ（MFEM）である旨を示す情報を格納するための領域である。

[0149] ここでは、コントローラ（OFC）10は、フローエントリ変更メッセージ（MFEM）のヘッダ領域（HF）に、「Type is Modify State」を指定する。「Type is Modify State」は、本メッセージがフローエントリ変更メッセージ（MFEM）である旨を示す情報である。

[0150] 整合領域（MF）、クッキー領域（CF）、及びアクション領域（AF）については、基本的に、図5で説明したトポロジーキープアライブフローエントリ（TKAFE）と同様である。

[0151] なお、アクション領域（AF）のプロパティ項目には、更に、命令（Command）が追加されている。

[0152] コントローラ（OFC）10は、新規のトポロジーキープアライブフロー

エントリ (TKAFE) を追加する場合、命令 (Command) に、「フローエントリの追加」 (Add Flow entry) を指定する。また、コントローラ (OFC) 10は、既存のトポロジーキープアライブフローエントリ (TKAFE) を修正する場合、命令 (Command) に、「フローエントリの変更」 (Modify Flow entry) を指定する。

- [0153] また、コントローラ (OFC) 10は、整合領域 (MF)、クッキー領域 (CF)、及びアクション領域 (AF) の各項目に、設定値を指定する。
- [0154] 各スイッチ (OFS) 20は、このフローエントリ変更メッセージ (MFEM) を受信した場合、アクション領域 (AF) の命令 (Command) を確認する。
- [0155] 各スイッチ (OFS) 20は、アクション領域 (AF) の命令 (Command) に「フローエントリの変更」 (Modify Flow entry) が指定されている場合、整合領域 (MF)、クッキー領域 (CF)、及びアクション領域 (AF) の各項目を参照し、該当するトポロジーキープアライブフローエントリ (TKAFE) が存在するか確認する。
- [0156] 各スイッチ (OFS) 20は、該当するトポロジーキープアライブフローエントリ (TKAFE) が存在しない場合、或いは、アクション領域 (AF) の命令 (Command) に「フローエントリの追加」 (Add Flow entry) が指定されている場合、このフローエントリ変更メッセージ (MFEM) の整合領域 (MF)、クッキー領域 (CF)、及びアクション領域 (AF) の各項目に指定された値に基づいて、新規のトポロジーキープアライブフローエントリ (TKAFE) を追加する。
- [0157] 各スイッチ (OFS) 20は、該当するトポロジーキープアライブフローエントリ (TKAFE) が存在する場合、このフローエントリ変更メッセージ (MFEM) の整合領域 (MF)、クッキー領域 (CF)、及びアクション領域 (AF) の各項目に指定された値を、トポロジーキープアライブフローエントリ (TKAFE) に反映する。すなわち、フローエントリ変更メッ

ページ (MFEM) の上記の各領域に指定された値を、トポロジーキープアライブフローエントリ (TKAFE) の該当する領域に上書きする。

[0158] これにより、コントローラ (OFC) 10は、作成/修正したトポロジーキープアライブフローエントリ (TKAFE) を、各スイッチ (OFS) 20側に設定することができる。

[0159] なお、コントローラ (OFC) 10が、既存のトポロジーキープアライブフローエントリ (TKAFE) を削除する場合には、命令 (Command) に、「フローエントリの削除」 (Delete Flow entry) を指定すれば良い。「フローエントリの削除」 (Delete Flow entry) の手順は、「フローエントリの変更」 (Modify Flow entry) の手順と同様である。

[0160] [フローエントリ削除済メッセージの仕様]

図7を参照して、各スイッチ (OFS) 20がエキスパイヤしたトポロジーキープアライブフローエントリ (TKAFE) を削除した後に、コントローラ (OFC) 10に送信するフローエントリ削除済メッセージ (FERM) の仕様について説明する。

[0161] ヘッダ領域 (HF) は、本メッセージがフローエントリ削除済メッセージ (FERM) である旨を示す情報を格納するための領域である。

[0162] ここでは、各スイッチ (OFS) 20は、フローエントリ削除済メッセージ (FERM) のヘッダ領域 (HF) に、「Type is Flow Removed」を指定する。「Type is Flow Removed」は、本メッセージがフローエントリ削除済メッセージ (FERM) である旨を示す情報である。

[0163] 整合領域 (MF)、クッキー領域 (CF)、及びアクション領域 (AF) については、基本的に、図6で説明したフローエントリ変更メッセージ (MFEM) と同様である。

[0164] 各スイッチ (OFS) 20は、命令 (Command) に、「フローエントリの削除」 (Delete Flow entry) を指定する。

- [0165] また、各スイッチ（OFS）20は、整合領域（MF）、クッキー領域（CF）、及びアクション領域（AF）の各項目に、削除対象となるトポロジーキープアライブフローエントリ（TKAFE）に指定されている値（既定値）を指定する。
- [0166] コントローラ（OFC）10は、このフローエントリ削除済メッセージ（FERM）を受信した場合、ヘッダ領域（HF）、又は、アクション領域（AF）の命令（Command）の少なくとも一方を確認する。
- [0167] コントローラ（OFC）10は、ヘッダ領域（HF）に「Type is Flow Removed」が指定されている場合、又は、アクション領域（AF）の命令（Command）に「フローエントリの削除」（Delete Flow entry）が指定されている場合、クッキー領域（CF）のクッキーに格納された64ビットの識別子を参照し、或いは、整合領域（MF）、クッキー領域（CF）、及びアクション領域（AF）の各項目を参照し、該当するトポロジーキープアライブフローエントリ（TKAFE）が存在するか確認する。
- [0168] コントローラ（OFC）10は、該当するトポロジーキープアライブフローエントリ（TKAFE）が存在しない場合、このフローエントリ削除済メッセージ（FERM）を無視・破棄する。
- [0169] コントローラ（OFC）10は、該当するトポロジーキープアライブフローエントリ（TKAFE）が存在する場合、そのトポロジーキープアライブフローエントリ（TKAFE）を削除する。
- [0170] これにより、各スイッチ（OFS）20は、エキスパイヤしたトポロジーキープアライブフローエントリ（TKAFE）の削除を、コントローラ（OFC）10側に反映することができる。
- [0171] <実施例>
図8を参照して、本発明に係るネットワークシステムの実施例について説明する。
- [0172] 本実施例では、1台のコントローラ（OFC）10と4台のスイッチ（O

F S) 20でオープンフローネットワークが構成されている。

[0173] コントローラ (OFC) 10は、セキュアチャンネルネットワーク100を介して、4台のスイッチ (OFS) 20の各々と接続されている。セキュアチャンネルネットワーク100には、コントローラ (OFC) とスイッチ (OFS) 間のコントロール信号が流れる。

[0174] 4台のスイッチ (OFS) 20を、それぞれ、スイッチ (OFS) 20-1、スイッチ (OFS) 20-2、スイッチ (OFS) 20-3、及びスイッチ (OFS) 20-4とする。

[0175] ここでは、スイッチ (OFS) 20-1の「DPID」を、「OFS1-DPID」とする。スイッチ (OFS) 20-2の「DPID」を、「OFS2-DPID」とする。スイッチ (OFS) 20-3の「DPID」を、「OFS3-DPID」とする。スイッチ (OFS) 20-4の「DPID」を、「OFS4-DPID」とする。

[0176] また、スイッチ (OFS) 20-1の接続ポートの1つである「Port 1」は、スイッチ (OFS) 20-2の接続ポートの1つである「Port 2」と接続されている。スイッチ (OFS) 20-2の接続ポートの1つである「Port 3」は、スイッチ (OFS) 20-3の接続ポートの1つである「Port 4」と接続されている。スイッチ (OFS) 20-3の接続ポートの1つである「Port 5」は、スイッチ (OFS) 20-4の接続ポートの1つである「Port 6」と接続されている。スイッチ (OFS) 20-4の接続ポートの1つである「Port 7」は、スイッチ (OFS) 20-1の接続ポートの1つである「Port 8」と接続されている。

[0177] [トポロジー初期検出]

まず、トポロジーを初期検出した際の動作について説明する。

[0178] コントローラ (OFC) 10のトポロジー検出部12は、既存のトポロジー検出機能を利用して、繋がっているスイッチ (OFS) 20の相互接続関係を収集し、オープンフローネットワークのトポロジーを検出する。

[0179] トポロジー検出部12は、検出したトポロジー情報を、以下のように、ト

ポロジ-管理部11のトポロジ-情報保存データベースに格納する。

- [0180] (Connection (OFS1-DPID: Port1 to
OFS2-DPID: Port2), KeepAlive Flow
Entry Cookie (0x0000000000000000))
- [0181] (Connection (OFS2-DPID: Port2 to
OFS1-DPID: Port1), KeepAlive Flow
Entry Cookie (0x0000000000000000))
- [0182] (Connection (OFS2-DPID: Port3 to
OFS3-DPID: Port4), KeepAlive Flow
Entry Cookie (0x0000000000000000))
- [0183] (Connection (OFS3-DPID: Port4 to
OFS2-DPID: Port3), KeepAlive Flow
Entry Cookie (0x0000000000000000))
- [0184] (Connection (OFS3-DPID: Port5 to
OFS4-DPID: Port6), KeepAlive Flow
Entry Cookie (0x0000000000000000))
- [0185] (Connection (OFS4-DPID: Port6 to
OFS3-DPID: Port5), KeepAlive Flow
Entry Cookie (0x0000000000000000))
- [0186] (Connection (OFS4-DPID: Port7 to
OFS1-DPID: Port8), KeepAlive Flow
Entry Cookie (0x0000000000000000))
- [0187] (Connection (OFS1-DPID: Port8 to
OFS4-DPID: Port7), KeepAlive Flow
Entry Cookie (0x0000000000000000))

[0188] [トポロジ-維持・更新]

次に、トポロジ-を維持・更新する際の動作について説明する。

[0189] トポロジ-検出部12は、トポロジ-を初期検出した後、トポロジ-管理

部 1 1 のトポロジー情報データベースに格納した各スイッチ (O F S) とポート (P o r t) の組み合わせを順次に取り出し、そのスイッチ (O F S) とポート (P o r t) の組み合わせに基づいて、トポロジーキープアライブフローエントリ (T K A F E) を作成する。

[0190] また、トポロジー検出部 1 2 は、トポロジーキープアライブフローエントリ (T K A F E) の作成と同時に、作成したトポロジーキープアライブフローエントリ (T K A F E) を識別するためのクッキーを生成する。

[0191] 本実施例では、トポロジー検出部 1 2 は、まず「O F S 1 - D P I D」の「P o r t 1」に基づいて、トポロジーキープアライブフローエントリ (T K A F E) を生成する。

[0192] トポロジー検出部 1 2 は、生成したトポロジーキープアライブフローエントリ (T K A F E) の整合領域 (M F) の各エントリー項目について、以下のような値を指定する。

[0193] トポロジー検出部 1 2 は、入力ポート (I n g r e s s P o r t) の整合値 (M V) に「P o r t 1」を指定する。

[0194] トポロジー検出部 1 2 は、フレームタイプ (E t h e r t y p e) の整合値 (M V) に「0 x 1 1 1 1」を指定する。

[0195] トポロジー検出部 1 2 は、優先度 (E n t r y P r i o r i t y) に「0 x f f f f」を指定する。

[0196] トポロジー検出部 1 2 は、遊休時間 (I d l e t i m e) に「0 x 0 0 0 3」を指定する。

[0197] トポロジー検出部 1 2 は、固定時間 (H a r d t i m e) に「0 x 0 0 0 0」を指定する。

[0198] トポロジー検出部 1 2 は、フラグ (F l a g) に、「送信フロー削除フラグ (S F R F) = o n」を指定する。すなわち、送信フロー削除フラグ (S F R F) を「o n」(有効)にする。

[0199] また、トポロジー検出部 1 2 は、上記の指定と同時に、クッキーを生成し、クッキーのクッキー値 (C V) に、フローエントリを特定するための 6 4

ビットの識別子を指定する。

- [0200] 本実施例では、クッキーのクッキー値 (CV) に、「OFS1-DPID」の「Port1」のトポロジーキープアライブフローエントリ (TKAFE) を識別するために生成した「0x0000000000000001」を指定する。
- [0201] トポロジー検出部12は、クッキーを生成した後、生成したトポロジーキープアライブフローエントリ (TKAFE)、クッキー値 (CV)、送信先のスイッチ (OFS) 20-1の「DPID」等の情報をパケット梱包・振り分け処理部13に渡し、新規のフローエントリを追加するためのフローエントリ変更メッセージ (MFEM) に梱包する旨の指示を出す。
- [0202] パケット梱包・振り分け処理部13は、トポロジー検出部12から受け取った情報及び指示に基づいて、図6で説明したフローエントリ変更メッセージ (MFEM) を作成する。また、パケット梱包・振り分け処理部13は、作成したフローエントリ変更メッセージ (MFEM) と、送信先スイッチ (OFS) のIPアドレスと、セキュアチャンネルポートのポート番号を、パケット送信部14に渡す。
- [0203] パケット送信部14は、パケット梱包・振り分け処理部13から受け取った情報に基づいて、フローエントリ変更メッセージ (MFEM) を、セキュアチャンネルポートから、送信先スイッチ (OFS) に送信する。
- [0204] 本実施例では、パケット送信部14は、フローエントリ変更メッセージ (MFEM) をスイッチ (OFS) 20-1に送信する。
- [0205] トポロジー検出部12は、非特許文献1 (OpenFlow Switch Specification Version 1.1.0) によって規定されている「Barrier Request/Reply」の同期の仕組みで、スイッチ (OFS) 20-1に対するトポロジーキープアライブフローエントリ (TKAFE) の追加が成功したことを知り、追加に成功したトポロジーキープアライブフローエントリ (TKAFE) のクッキー値 (CV) をトポロジー情報データベースに追加する。

- [0206] 本実施例では、トポロジー検出部12は、トポロジー管理部11のトポロジー情報データベースに、以下のようにクッキー値(CV)を追加する。
- [0207] (Connection (OFS1-DPID: Port1 to OFS2-DPID: Port2), KeepAlive Flow Entry Cookie (0x000000000000000001))
- [0208] また、本実施例では、トポロジー検出部12は、「OFS1-DPID」の「Port1」と相互接続している「OFS2-DPID」の「Port2」に対して、上記と同様な手順で、トポロジーキープアライブフローエントリ(TKAFE)を追加し、クッキー値(CV)をトポロジー情報データベースに設定する。
- [0209] トポロジー検出部12は、トポロジー管理部11のトポロジー情報データベースに、以下のようなクッキー値(CV)を追加する。
- [0210] (Connection (OFS2-DPID: Port2 to OFS1-DPID: Port1), KeepAlive Flow Entry Cookie (0x000000000000000002))
- [0211] トポロジー検出部12は、相互接続しているスイッチ(OFS)の両方に対して、トポロジーキープアライブフローエントリ(TKAFE)の追加に成功した後に、トポロジーキープアライブパケット(TKAP)と、図4で説明したパケット出力メッセージ(POM)を両方のスイッチ(OFS)に送信する旨の指示を、パケット梱包・振り分け処理部13に渡す。
- [0212] パケット梱包・振り分け処理部13は、トポロジー検出部12から受け取った情報及び指示に基づいて、トポロジーキープアライブパケット(TKAP)をパケット出力メッセージ(POM)のデータ領域(DF)に格納する。
- [0213] パケット出力メッセージ(POM)のアクション領域(AF)に、スイッチ(OFS)間の相互接続ポートからデータ領域(DF)のパケットを送信する動作を指定する。
- [0214] 本実施例では、パケット梱包・振り分け処理部13は、スイッチ(OFS

) 20-1に対して、アクション領域 (AF) に「Port 1からパケットを送信する動作」を指定したパケット出力メッセージ (POM) を作成する。また、スイッチ (OFS) 20-2に対して、アクション領域 (AF) に「Port 2からパケットを送信する動作」を指定したパケット出力メッセージ (POM) を作成する。

[0215] パケット梱包・振り分け処理部13は、これらのパケット出力メッセージ (POM) と、送信先スイッチ (OFS) のIPアドレスと、セキュアチャンネルポートのポート番号を、パケット送信部14に渡す。

[0216] パケット送信部14は、パケット梱包・振り分け処理部13から受け取った情報に基づいて、これらのパケット出力メッセージ (POM) を、それぞれ該当する送信先スイッチ (OFS) に送信する。

[0217] 本実施例では、パケット送信部14は、スイッチ (OFS) 20-1に対して、アクション領域 (AF) に「Port 1からパケットを送信する動作」を指定したパケット出力メッセージ (POM) を送信する。また、スイッチ (OFS) 20-2に対して、アクション領域 (AF) に「Port 2からパケットを送信する動作」を指定したパケット出力メッセージ (POM) を送信する。

[0218] スイッチ (OFS) 20-1は、受け取ったパケット出力メッセージ (POM) に従って、図3で説明したトポロジーキープアライブパケット (TKAP) を「Port 1」から送信する。

[0219] スイッチ (OFS) 20-2は、受け取ったパケット出力メッセージ (POM) に従って、図3で説明したトポロジーキープアライブパケット (TKAP) を「Port 2」から送信する。

[0220] また、トポロジーキープアライブパケット (TKAP) は、スイッチ (OFS) 20-1とスイッチ (OFS) 20-2の各々に設定されたトポロジーキープアライブフローエントリ (TKAFE) にマッチするので、障害が発生しない限り、トポロジーキープアライブパケット (TKAP) は、スイッチ (OFS) 20-1の「Port 1」とスイッチ (OFS) 20-2の

「Port 2」との間で往復する。

[0221] 以上の仕組みで、コントローラ（OFC）10は、検出した全部のスイッチ（OFS）とポート（Port）の組み合わせに対して、トポロジーキープアライブフローエントリ（TKAFE）の登録と、トポロジーキープアライブパケット（TKAP）の送信を実施する。

[0222] [障害発生時]

次に、スイッチ（OFS）間の相互接続ポートやネットワークに障害が発生した際の動作について説明する。

[0223] スイッチ（OFS）間の相互接続ポートやネットワークに障害が発生すると、トポロジーキープアライブパケット（TKAP）は、正しくスイッチ（OFS）間の相互接続ポートに到着できなくなる。すなわち、スイッチ（OFS）において、トポロジーキープアライブフローエントリ（TKAFE）にマッチするパケットが到着しなくなる。

[0224] スイッチ（OFS）は、トポロジーキープアライブフローエントリ（TKAFE）にマッチするパケットが到着しない時間が、トポロジーキープアライブフローエントリ（TKAFE）の遊休時間（Idle time）に設定した時間以上になると、トポロジーキープアライブフローエントリ（TKAFE）がエキスパイアしたと判断し、このトポロジーキープアライブフローエントリ（TKAFE）を削除する。

[0225] スイッチ（OFS）は、エキスパイアしたトポロジーキープアライブフローエントリ（TKAFE）を削除した後に、コントローラ（OFC）10に対して図7で説明したフローエントリ削除済メッセージ（FERM）を送信する。

[0226] コントローラ（OFC）10のトポロジー検出部12は、スイッチ（OFS）から送信されたフローエントリ削除済メッセージ（FERM）のクッキー値（CV）でトポロジー情報データベースに格納したトポロジー情報を検索することで、どのスイッチ（OFS）のどのポート（Port）のトポロジーキープアライブフローエントリ（TKAFE）がエキスパイアしたこと

を知り、該当のトポロジー情報をデータベースから削除する。

- [0227] 本実施例では、スイッチ（OFS）20-1の「Port 1」とスイッチ（OFS）20-2の「Port 2」（相互接続しているポート）の間に障害が発生した場合を例として説明する。
- [0228] 障害が発生すると、スイッチ（OFS）20-1とスイッチ（OFS）20-2の各々に、スイッチ（OFS）20-1の「Port 1」とスイッチ（OFS）20-2の「Port 2」との間のトポロジーキープアライブパケット（TKAP）が届かなくなる。
- [0229] スイッチ（OFS）20-1とスイッチ（OFS）20-2は、それぞれ、このトポロジーキープアライブフローエントリ（TKAFE）にマッチするトポロジーキープアライブパケット（TKAP）が届かなくなってから3秒が経つと、このトポロジーキープアライブフローエントリ（TKAFE）がエキスパイヤしたと判断し、このトポロジーキープアライブフローエントリ（TKAFE）を削除する。
- [0230] スイッチ（OFS）20-1とスイッチ（OFS）20-2は、それぞれ、トポロジーキープアライブフローエントリ（TKAFE）を削除した後、フローエントリ削除済メッセージ（FERM）をコントローラ（OFC）10に送信する。
- [0231] コントローラ（OFC）10は、スイッチ（OFS）20-1とスイッチ（OFS）20-2の各々から、それぞれフローエントリ削除済メッセージ（FERM）を受信する。
- [0232] コントローラ（OFC）10は、それぞれのフローエントリ削除済メッセージ（FERM）に格納されたクッキー値（CV）に基づいて、相互接続がなくなっていると判断し、相互接続情報をトポロジー情報データベースから削除することで、トポロジー変更を更新する。
- [0233] なお、スイッチ（OFS）20-1から受信したフローエントリ削除済メッセージ（FERM）に格納されたクッキー値（CV）は、「0x0000000000000001」である。

[0234] また、スイッチ（OFS）20-2から受信したフローエントリ削除済メッセージ（FERM）に格納されたクッキー値（CV）は、「0x0000000000000002」である。

[0235] 本実施例では、コントローラ（OFC）10は、以下の相互接続情報をトポロジー情報データベースから削除することで、トポロジー変更を更新する。

[0236] (Connection (OFS1-DPID: Port1 to
OFS2-DPID: Port2), KeepAlive Flow
Entry Cookie (0x0000000000000001))

[0237] (Connection (OFS2-DPID: Port2 to
OFS1-DPID: Port1), KeepAlive Flow
Entry Cookie (0x0000000000000002))

[0238] <第2実施形態>

以下に、本発明の第2実施形態について説明する。

[0239] 本実施形態では、トポロジーキープアライブフローエントリ（TKAFE）がエクスパイアした際に、即時に相互接続ポートに障害が発生したと判断せずに、LLDPパケットを利用して一定のリトライ回数でトポロジー検査を行う。

[0240] すなわち、二重の検査を行い、LLDPパケットを利用したトポロジー検査でも相互接続関係を確認できない場合に、スイッチ（OFS）間の相互接続ポートに障害が発生したと判断する。

[0241] <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。

[0242] <本発明の特徴>

次に、本発明の特徴について説明する。

[0243] 本発明に係るネットワークシステムは、例えばオープンフローネットワークにおいて、コントローラ（OFC）がスイッチ（OFS）間のネットワークが高遅延な状況でのスイッチの物理トポロジーの維持・更新を行うネット

ワークシステムに関する。

- [0244] コントローラ (OFC) は、スイッチ (OFS) 間の初期接続の完了後に、各スイッチ (OFS) の相互接続ポートに対して、アクション領域 (AF) に「`Send to in-port`」を指定し、遊休時間経過領域 (`Idle timeout Field`) に遊休時間を指定した巡回用のフローエントリを設定する。なお、トポロジーキープアライブフローエントリ (TKAFE) は、巡回用のフローエントリの1つである。
- [0245] コントローラ (OFC) は、パケット出力メッセージ (POM) を利用して、予め規定された巡回パケットを各スイッチ (OFS) の相互接続ポートに送る。なお、トポロジーキープアライブパケット (TKAP) は、巡回パケットの1つである。
- [0246] 巡回パケットは、アクション領域 (AF) に「`Send to in-port`」が指定されている巡回用のフローエントリにマッチし、各スイッチ (OFS) の相互接続ポート間で往復する。
- [0247] OPS間の相互接続ポートに障害発生時に、巡回パケットが各スイッチ (OFS) の相互接続ポートに届かなくなる。
- [0248] 各スイッチ (OFS) において、アクション領域 (AF) に「`Send to in-port`」が指定されている巡回用のフローエントリにマッチするパケットがなくなる。
- [0249] 各スイッチ (OFS) は、巡回用のフローエントリの遊休時間経過領域 (`Idle timeout Field`) に指定されている遊休時間の経過後に、当該巡回用のフローエントリを削除 (エージアウト) する。
- [0250] また、スイッチ (OFS) は、巡回用のフローエントリを削除 (エージアウト) したことを示すエージアウトメッセージを、コントローラ (OFC) へ通知する。
- [0251] コントローラ (OFC) は、スイッチ (OFS) から通知された巡回用のフローエントリのエージアウトメッセージに基づいて、該当スイッチ (OFS)

S) 間の相互接続ポートに障害が発生したと判断する。

[0252] なお、本発明に係るネットワークシステムでは、オープンフローネットワークにおいて、コントローラ（OFC）が検出したトポロジを維持し、トポロジ変更を検出するために、コントローラ（OFC）が頻繁的に走査パケットを流す仕組みを利用するのではなく、トポロジ変更をスイッチ（OFS）が積極的にコントローラ（OFC）に通知する仕組みを利用する。そのため、セキュアチャンネルネットワーク上のスイッチ（OFS）への負荷を軽減することができる。

[0253] また、本発明に係るネットワークシステムでは、トポロジキープアライブフローエントリ（TKAFE）の遊休時間（Idle time）と、エキスパイヤと、フローエントリ削除済メッセージ（FERM）の仕組みを利用する。そのため、セキュアチャンネルネットワーク上のスイッチ（OFS）に負荷を与えることなく、トポロジ変更の検出を素早くするように調整することができる。

[0254] また、本発明に係るネットワークシステムでは、セキュアチャンネルに流すトポロジ変更を通知するパケットが、1つのスイッチ（OFS）間の相互接続ポートに対して1つしか存在しない。そのため、セキュアチャンネルが高負荷の状態でも、トポロジ変更の検出が漏れる可能性を、既存のトポロジ検出技術よりも低くすることができる。

[0255] また、本発明に係るネットワークシステムでは、スイッチ（OFS）が、トポロジキープアライブパケット（TKAP）を汎用プロセッサ以外のLSI等で処理する。そのため、スイッチ（OFS）の汎用プロセッサのリスース不足が原因で、発生していない物理トポロジの変更を誤って検出してしまう可能性を軽減することができる。

[0256] 更に、本発明に係るネットワークシステムでは、オープンフロープロトコル（OpenFlow Protocol）に対応したスイッチ（OFS）を利用すれば、スイッチ（OFS）に新たな機能（ハードウェア実装、ソフトウェア実装）を追加する必要がなく、コントローラへのソフトウェア実装

だけで実現することができる。

[0257] <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。

[0258] コントローラ (OFC) 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。なお、コントローラ (OFC) は、端末やサーバに限らず、中継機器や周辺機器でも良い。また、コントローラ (OFC) は、計算機等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。

[0259] スイッチ (OFS) 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、物理マシン上に構築された仮想マシン (VM) により実現される仮想スイッチでも良い。

[0260] コントローラ (OFC) 及びスイッチ (OFS) 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。

[0261] 図示しないが、コントローラ (OFC) 及びスイッチ (OFS) 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、ネットワークとの通信に用いら

れるインターフェースによって実現される。

[0262] 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。

[0263] 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。また、バッファ (buffer) やレジスタ (register) 等でも良い。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。

[0264] 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、電子機器等に搭載される1チップマイコンが、上記のプロセッサ及び上記のメモリを備えている事例も考えられる。

[0265] 上記の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 (Ne

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。

[0266] 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。

[0267] なお、コントローラ (OFC) 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。

[0268] 但し、実際には、これらの例に限定されない。

[0269] <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。

[0270] なお、本出願は、日本出願番号2012-016225に基づく優先権を主張するものであり、日本出願番号2012-016225における開示内容は引用により本出願に組み込まれる。

請求の範囲

- [請求項1] パケットをフローとして一律に制御するためのルールとアクションとが定義されたフローエン트리に従って、受信したパケットの処理を行う複数のスイッチと、
- 前記複数のスイッチの各々に対して、前記複数のスイッチ間で相互に送受信される巡回パケットが到着しなくなった場合に削除される巡回用フローエントリを設定し、前記各スイッチから、前記巡回用フローエントリを削除した旨の通知を受信した場合、前記複数のスイッチ間の障害を検出するコントローラと
- 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、
- 前記コントローラは、
- 前記巡回用フローエントリに、パケットが入ってきたポートに当該パケットを転送する旨のアクションと、マッチするパケットが到着しなくなってからの有効期間を示す遊休時間と、を指定して前記各スイッチに設定する手段と、
- 前記巡回パケットを作成して前記各スイッチに送信する手段と、
- 前記各スイッチから、前記遊休時間が経過した巡回用フローエントリを削除した旨の通知を受信する手段と
- を具備する
- ネットワークシステム。
- [請求項3] 請求項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、
- 前記遊休時間が経過した巡回用フローエントリを削除した旨の通知を受信した際に、LLDPパケットを利用して一定のリトライ回数でトポロジー検査を行う手段と、
- 前記LLDPパケットを利用したトポロジー検査でも相互接続関係

を確認できない場合、前記複数のスイッチ間に障害が発生したと判断する手段と

を更に具備する

ネットワークシステム。

[請求項4]

複数のスイッチの各々に対して、パケットをフローとして一律に制御するためのルールとアクションとが定義されたフローエントリを設定する手段と、

前記複数のスイッチの各々に対して、前記複数のスイッチ間で相互に送受信される巡回パケットが到着しなくなった場合に削除される巡回用フローエントリを設定する手段と、

前記各スイッチから、前記巡回用フローエントリを削除した旨の通知を受信した場合、前記複数のスイッチ間の障害を検出する手段とを具備する

コントローラ。

[請求項5]

請求項4に記載のコントローラであって、

前記巡回用フローエントリに、パケットが入ってきたポートに当該パケットを転送する旨のアクションと、マッチするパケットが到着しなくなってからの有効期間を示す遊休時間と、を指定して前記各スイッチに設定する手段と、

前記巡回パケットを作成して前記各スイッチに送信する手段と、

前記各スイッチから、前記遊休時間が経過した巡回用フローエントリを削除した旨の通知を受信する手段と

を更に具備する

コントローラ。

[請求項6]

請求項5に記載のコントローラであって、

前記遊休時間が経過した巡回用フローエントリを削除した旨の通知を受信した際に、LLDPパケットを利用して一定のリトライ回数でトポロジー検査を行う手段と、

前記LLDPパケットを利用したトポロジー検査でも相互接続関係を確認できない場合、前記複数のスイッチ間に障害が発生したと判断する手段と

を更に具備する

コントローラ。

[請求項7]

計算機により実施されるトポロジー管理方法であって、

コントローラとして、複数のスイッチの各々に対して、パケットをフローとして一律に制御するためのルールとアクションとが定義されたフローエントリを設定することと、

前記複数のスイッチの各々に対して、前記複数のスイッチ間で相互に送受信される巡回パケットが到着しなくなった場合に削除される巡回用フローエントリを設定することと、

前記各スイッチから、前記巡回用フローエントリを削除した旨の通知を受信した場合、前記複数のスイッチ間の障害を検出することとを含む

トポロジー管理方法。

[請求項8]

請求項7に記載のトポロジー管理方法であって、

前記巡回用フローエントリに、パケットが入ってきたポートに当該パケットを転送する旨のアクションと、マッチするパケットが到着しなくなってからの有効期間を示す遊休時間と、を指定して前記各スイッチに設定することと、

前記巡回パケットを作成して前記各スイッチに送信することと、

前記各スイッチから、前記遊休時間が経過した巡回用フローエントリを削除した旨の通知を受信することと

を更に含む

トポロジー管理方法。

[請求項9]

請求項8に記載のトポロジー管理方法であって、

前記遊休時間が経過した巡回用フローエントリを削除した旨の通知

を受信した際に、LLDPパケットを利用して一定のリトライ回数でトポロジー検査を行うことと、

前記LLDPパケットを利用したトポロジー検査でも相互接続関係を確認できない場合、前記複数のスイッチ間に障害が発生したと判断することと

を更に含む

トポロジー管理方法。

[請求項10]

コントローラとして、複数のスイッチの各々に対して、パケットをフローとして一律に制御するためのルールとアクションとが定義されたフローエントリを設定するステップと、

前記複数のスイッチの各々に対して、前記複数のスイッチ間で相互に送受信される巡回パケットが到着しなくなった場合に削除される巡回用フローエントリを設定するステップと、

前記各スイッチから、前記巡回用フローエントリを削除した旨の通知を受信した場合、前記複数のスイッチ間の障害を検出するステップと

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。

[請求項11]

請求項10に記載の記憶媒体であって、

前記巡回用フローエントリに、パケットが入ってきたポートに当該パケットを転送する旨のアクションと、マッチするパケットが到着しなくなってからの有効期間を示す遊休時間と、を指定して前記各スイッチに設定するステップと、

前記巡回パケットを作成して前記各スイッチに送信するステップと

、

前記各スイッチから、前記遊休時間が経過した巡回用フローエントリを削除した旨の通知を受信するステップと

を更に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。

[請求項12]

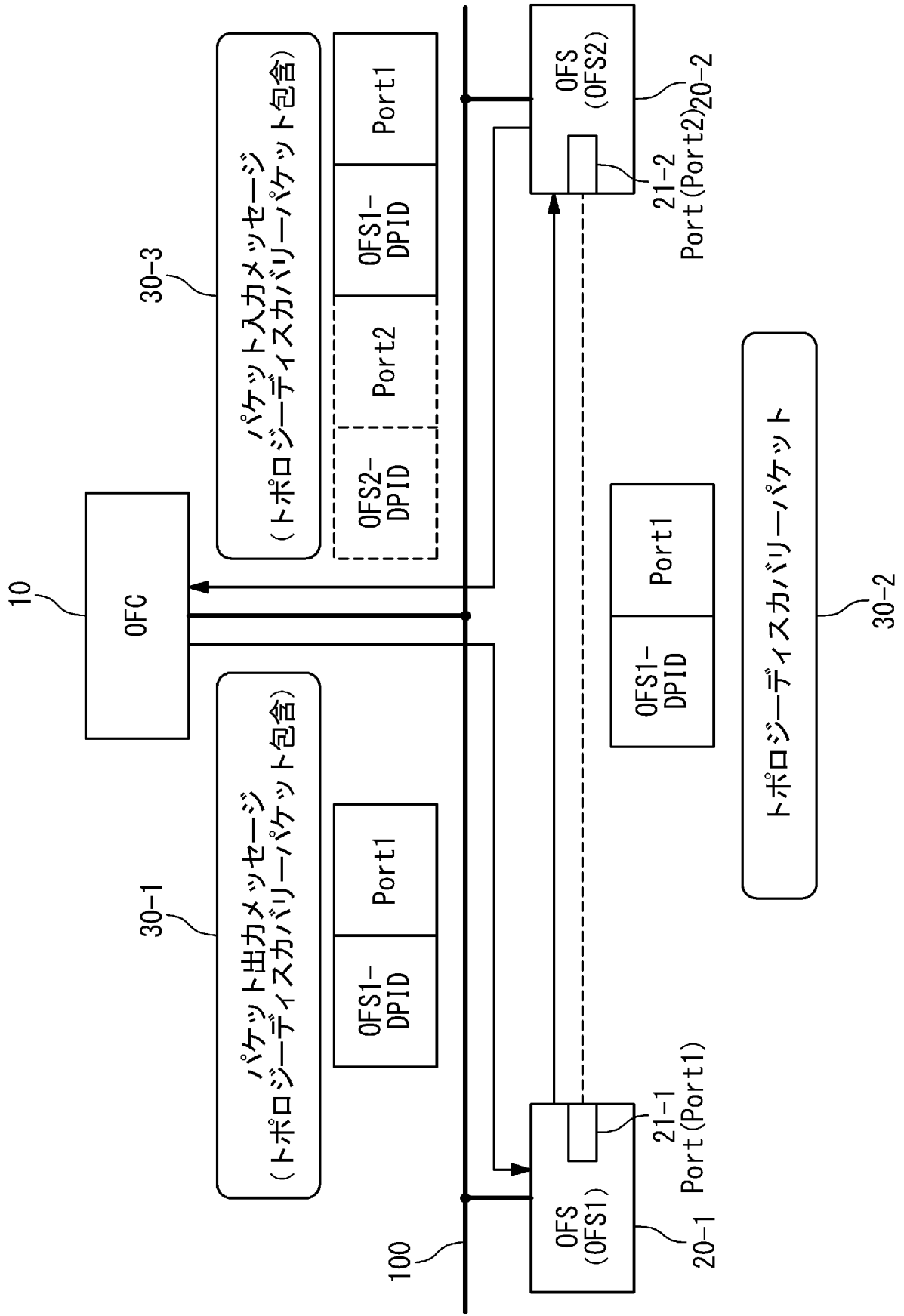
請求項11に記載の記憶媒体であって、

前記遊休時間が経過した巡回用フローエントリを削除した旨の通知を受信した際に、LLDPパケットを利用して一定のリトライ回数でトポロジー検査を行うステップと、

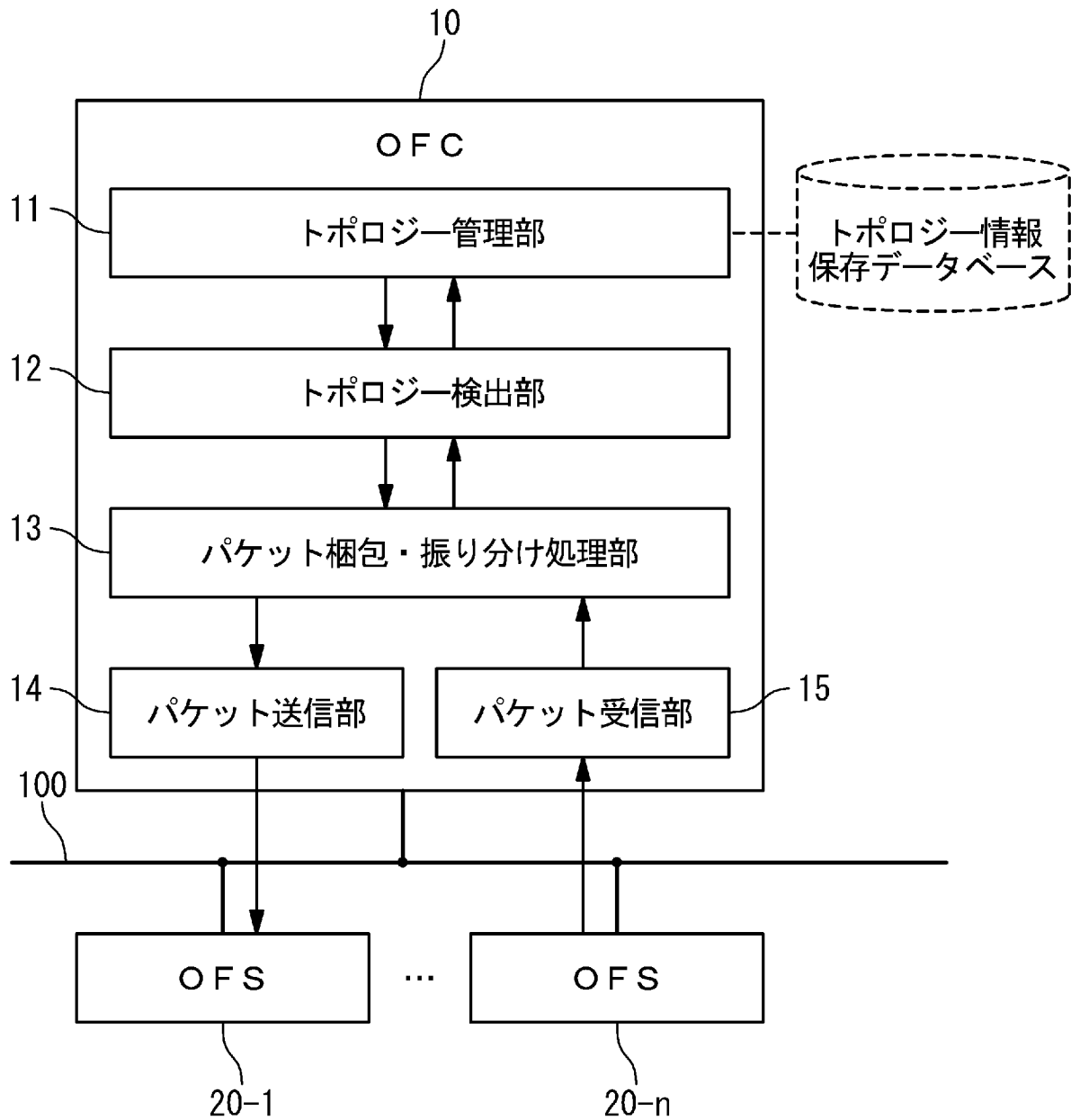
前記LLDPパケットを利用したトポロジー検査でも相互接続関係を確認できない場合、前記複数のスイッチ間に障害が発生したと判断するステップと

を更に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。

[図1]



[図2]



[図3]

トポロジーキープアライブパケット
(Topology KeepAlive Packet)

MACヘッダ (Ether Header)		
MAC送信元 (Ether src)	MAC送信先 (Ether dst)	フレームタイプ (Ether type)
*	Not BC/MC	0x0111

[図4]

パケット出カメッセージ (Packet-out Message)

ヘッダ領域 (Header Field)	アクション領域 (Action Field)	データ領域 (Data Field)		
Type is Packet Out	Send data field to the Specified port	MACヘッダ (Ether Header)		
		MAC送信元 (Ether src)	MAC送信先 (Ether dst)	フレームタイプ (Ether type)
		*	Not BC/MC	0x0111

トポロジーキープアラライブパケット (Topology KeepAlive Packet)

[図5]

トポロジキーベースドフローエントリ (Topology KeepAlive Flow Entry)

整合領域 (Match Field)	クッキー領域 (Cookie Field)	アクション領域 (Header Field)
※エントリ項目	クッキー	※プロパティ項目
※整合値 (Match Value)	Specified	※整合値
	unit64_t	

※エントリ項目	※整合値 (Match Value)
入力ポート (Ingress Port)	Specified Port
メタデータ (Meta data)	ANY
MAC送信元 (Ether src)	ANY
MAC送信先 (Ether dst)	ANY
フレームタイプ (Ether type)	0x1111
仮想LAN識別子 (VLAN id)	ANY
仮想LAN優先度 (VLAN priority)	ANY
MPLSラベル (MPLS label)	ANY
MPLSトラフィッククラス (MPLS traffic class)	ANY
IP送信元 (IP src)	ANY
IP送信先 (IP dst)	ANY
IPプロトコル (IP protocol)	ANY
IPサービスタイプ (IP ToS bits)	ANY
送信元ポート (TCP/UDP src Port)	ANY
送信先ポート (TCP/UDP dst Port)	ANY

※プロパティ項目	※特性値
優先度 (Entry Priority)	0xffff
遊休時間 (Idle time)	0x0003
固定時間 (Hard time)	0x0000
フラグ (Flag)	Send Flow Removed Flag=on
アクション (Action)	Send to in - port

[図6]

フローエントリ変更メッセージ (Modify Flow Entry Message)

ヘッダ領域 (Header Fild)	整合領域 (Match Fild)	クッキー領域 (Cookie Fild)	アクション領域 (Header Fild)
Type is Modify State	※エントリ項目	クッキー	※プロパティ項目
	※整合値 (Match Value)	Specified	※特性値

※エントリ項目	※整合値 (Match Value)
入力ポート (Ingress Port)	Specified Port
メタデータ (Meta data)	ANY
MAC送信元 (Ether src)	ANY
MAC送信先 (Ether dst)	ANY
フレームタイプ (Ether type)	0x1111
仮想LAN識別子 (VLAN id)	ANY
仮想LAN優先度 (VLAN priority)	ANY
MPLSラベル (MPLS label)	ANY
MPLSトラフィッククラス (MPLS traffic class)	ANY
IP送信元 (IP src)	ANY
IP送信先 (IP dst)	ANY
IPプロトコル (IP protocol)	ANY
IPサービスタイプ (IP ToS bits)	ANY
送信元ポート (TCP/UDP src Port)	ANY
送信先ポート (TCP/UDP dst Port)	ANY

※プロパティ項目	※特性値
命令 (Command)	Add/Modify Flow entry
優先度 (Entry Priority)	0xffff
遊休時間 (Idle time)	0x0003
固定時間 (Hard time)	0x0000
フラグ (Flag)	Send Flow Removed Flag=on
アクション (Action)	Send to in - port

[図7]

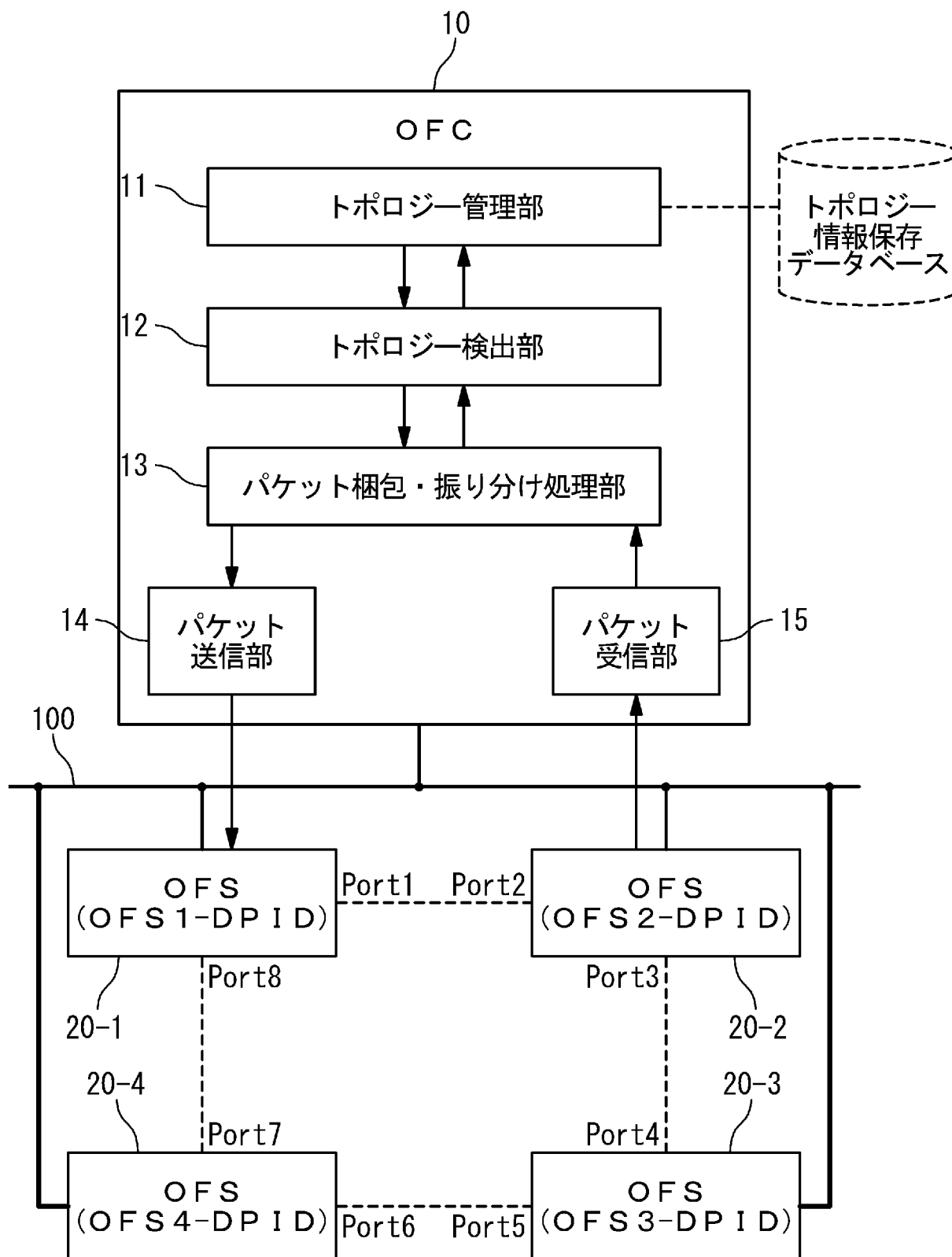
フローエントリ削除メッセージ (Flow Entry Removed Message)

ヘッダ領域 (Header Field)	整合領域 (Match Field)	クッキー領域 (Cookie Field)	アクション領域 (Header Field)
Type is Flow Removed	※エントリ項目	クッキー	※プロパティ項目
	※整合値 (Match Value)	Specified uinit64_t	※特性値

※エントリ項目	※整合値 (Match Value)	※特性値
入力ポート (Ingress Port)	Specified Port	Delete Flow entry
メタデータ (Meta data)	ANY	
MAC送信元 (Ether src)	ANY	0xffff
MAC送信先 (Ether dst)	ANY	
フレームタイプ (Ether type)	0x1111	0x0003
仮想LAN識別子 (VLAN id)	ANY	
仮想LAN優先度 (VLAN priority)	ANY	
MPLSラベル (MPLS label)	ANY	0x0000
MPLSトラフィッククラス (MPLS traffic class)	ANY	
IP送信元 (IP src)	ANY	Send Flow Removed Flag=on
IP送信先 (IP dst)	ANY	
IPプロトコル (IP protocol)	ANY	
IPサービスタイプ (IP ToS bits)	ANY	
送信元ポート (TCP/UDP src Port)	ANY	
送信先ポート (TCP/UDP dst Port)	ANY	Send to in - port

※プロパティ項目	※特性値
命令 (Command)	Delete Flow entry
優先度 (Entry Priority)	0xffff
遊休時間 (Idle time)	0x0003
固定時間 (Hard time)	0x0000
フラグ (Flag)	Send Flow Removed Flag=on
アクション (Action)	Send to in - port

[図8]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/051889

A. CLASSIFICATION OF SUBJECT MATTER

H04L12/70 (2013.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L12/00-12/28, 12/44-12/955

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2013
Kokai Jitsuyo Shinan Koho	1971-2013	Toroku Jitsuyo Shinan Koho	1994-2013

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	WO 2010/143607 A1 (NEC Corp.), 16 December 2010 (16.12.2010), paragraphs [0029] to [0036], [0045] to [0057]; fig. 1 to 3 & US 2012/0026891 A1	1, 2, 4, 5, 7, 8, 10, 11 3, 6, 9, 12
Y A	Hisatoshi OIKAWA et al., "Open Flow Network ni Okeru Keiro Joho no Tsuchi Shuho ni Kansuru Teian", Dai 73 Kai (Heisei 23 Nen) Zenkoku Taikai Koen Ronbunshu (3), Information Processing Society of Japan, 02 March 2011 (02.03.2011), pages 3-319 to 3-320	1, 2, 4, 5, 7, 8, 10, 11 3, 6, 9, 12
Y A	US 2011/0286324 A1 (Elisa BELLAGAMBA et al.), 24 November 2011 (24.11.2011), paragraph [0017]; fig. 1 & WO 2011/144495 A1	1, 2, 4, 5, 7, 8, 10, 11 3, 6, 9, 12

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"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)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search
06 March, 2013 (06.03.13)

Date of mailing of the international search report
19 March, 2013 (19.03.13)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

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

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. H04L12/00-12/28, 12/44-12/955

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	1971-2013年
日本国実用新案登録公報	1996-2013年
日本国登録実用新案公報	1994-2013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
Y A	WO 2010/143607 A1 (日本電気株式会社) 2010. 12. 16, 段落【0029】 - 【0036】、【0045】 - 【0057】 , 図 1-3 & US 2012/0026891 A1	1, 2, 4, 5, 7, 8, 10, 11 3, 6, 9, 12
Y A	及川永寿 他, 「OpenFlow ネットワークにおける経路情報の通知手 法に関する提案」, 第 73 回 (平成 23 年) 全国大会 講演論文集 (3) , 一般社団法人情報処理学会, 2011. 03. 02, pp. 3-319~3-320	1, 2, 4, 5, 7, 8, 10, 11 3, 6, 9, 12

C 欄の続きにも文献が列挙されている。

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	の日の後に公表された文献
「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	「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
「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	「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
「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)	「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
「O」口頭による開示、使用、展示等に言及する文献	「&」同一パテントファミリー文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願	

国際調査を完了した日
06. 03. 2013

国際調査報告の発送日
19. 03. 2013

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)	5 X	3 3 5 8
浦口 幸宏		
電話番号 03-3581-1101 内線 3596		

C (続き) . 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
Y A	US 2011/0286324 A1 (Elisa BELLAGAMBA et al.) 2011.11.24, paragraph [0017], Figure 1 & WO 2011/144495 A1	1, 2, 4, 5, 7, 8, 10, 11 3, 6, 9, 12