

(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
2012 年 9 月 13 日(13.09.2012)



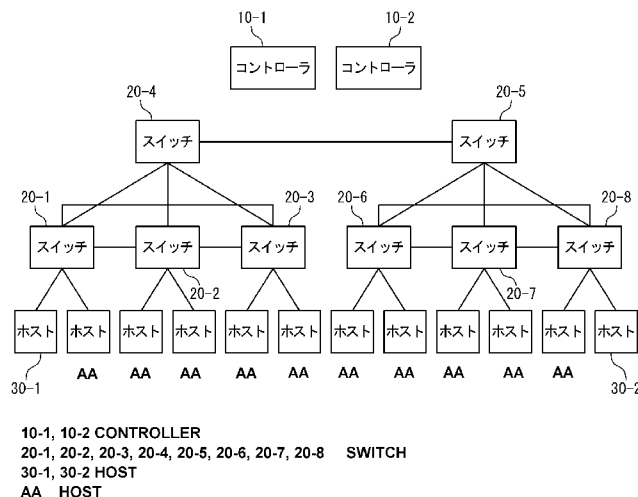
(10) 国際公開番号
WO 2012/121114 A1

- (51) 国際特許分類:
H04L 12/56 (2006.01)
- (21) 国際出願番号: PCT/JP2012/055272
- (22) 国際出願日: 2012 年 3 月 1 日(01.03.2012)
- (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 2011-048143 2011 年 3 月 4 日(04.03.2011) 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
- (75) 発明者/出願人 (米国についてのみ): 勝倉 辰之助 (KATSUKURA Shinnosuke) [JP/JP]; 〒1088001 東京都港区芝五丁目 7 番 1 号 Tokyo (JP).
- (74) 代理人: 工藤 実 (KUDOH Minoru); 〒1400013 東京都品川区南大井六丁目 2 4 番 1 0 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, 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, 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, MD, RU, TJ, TM), ユーロッパ (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, NETWORK DEVICE, AND PATH CONTROL METHOD

(54) 発明の名称: ネットワークシステム、ネットワーク機器、及び経路制御方法

[図1]



(57) Abstract: In the present invention, in a network constituted by network devices from which control functions are separated to external controllers, as typified by an openflow network, a path which is not adopted can be deleted when the controllers are multiplexed/redundant. Specifically, a network device sets a flow entry in which a rule and an operation for uniformly controlling packets as a flow are defined, in a flow table, on the basis of the control from each of a plurality of controllers. At this time, a path set by a controller which has completed the setting of the flow entries regarding all the network devices on the path, earliest among the plurality of controllers, is adopted, and paths set by other controllers are not adopted, and packets are transferred in accordance with the flow entry of the adopted path.

(57) 要約:

[続葉有]

WO 2012/121114 A1

オープンフローに代表される制御機能を外部のコントローラに分離したネットワーク機器で構成されるネットワークにおいて、コントローラを多重化／冗長化した際の不採用経路の削除を実現する。具体的には、ネットワーク機器は、複数のコントローラの各々からの制御に基づいて、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動作が定義されたフローエントリをフローテーブルに設定する。このとき、該複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、該採用された経路のフローエントリに従って、パケットを転送する。

明 細 書

発明の名称：

ネットワークシステム、ネットワーク機器、及び経路制御方法

技術分野

[0001] 本発明は、ネットワークシステムに関し、特にネットワーク通信の経路制御を行うコントローラに関する。

背景技術

[0002] ネットワークシステムの制御方式の1つとして、外部の制御装置（コントロールプレーン）からノード装置（ユーザプレーン）を制御するCU（C：コントロールプレーン／U：ユーザプレーン）分離型ネットワークシステムが提案されている。

[0003] CU分離型ネットワークシステムの一例として、コントローラからスイッチを制御してネットワークの経路制御を行うオープンフロー（OpenFlow）技術を利用したオープンフローネットワークシステムが挙げられる。オープンフロー技術の詳細については、非特許文献1に記載されている。なお、オープンフローネットワークシステムは一例に過ぎない。

[0004] [オープンフローネットワークシステムの説明]

オープンフローネットワークシステムでは、OFC（OpenFlow Controller）等のコントローラが、OFS（OpenFlow Switch）等のスイッチ内のフローテーブル（Flow table）を操作することによりスイッチの挙動を制御する。

[0005] オープンフローネットワークシステムにおけるスイッチとは、オープンフローネットワークを形成し、コントローラの制御下にあるエッジスイッチ及びコアスイッチのことである。オープンフローネットワークにおける入力側エッジスイッチでのパケット（packet）の受信から出力側エッジスイッチでの送信までのパケットの一連の流れをフロー（Flow）と呼ぶ。

[0006] パケットは、フレーム（frame）と読み替えても良い。パケットとフ

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。

[0007] フローテーブルとは、所定のマッチ条件（ルール）に適合するパケット（packet）に対して行うべき動作（アクション）を定義したフローエントリ（Flow entry）が登録されたテーブルである。

[0008] フローエントリのルールは、パケットの各プロトコル階層のヘッダ領域に含まれる宛先アドレス（Destination Address）、送信元アドレス（Source Address）、宛先ポート（Destination Port）、送信元ポート（Source Port）のいずれか又は全てを用いた様々な組み合わせにより定義され、区別可能である。なお、上記のアドレスには、MACアドレス（Media Access Control Address）やIPアドレス（Internet Protocol Address）を含むものとする。また、上記に加えて、入口ポート（Ingress Port）の情報も、フローエントリのルールとして使用可能である。また、フローエントリのルールとして、フローを示すパケットのヘッダ領域の値の一部（又は全部）を、正規表現やワイルドカード「*」等で表現したものを設定することもできる。

[0009] 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ければ、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。

[0010] オープンフローネットワークシステムにおけるスイッチは、フローエント

り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。

[0011] オープンフローネットワークシステムのような、「パケット転送の機能」と「経路制御の機能」をネットワーク機器とコントローラとに分離したネットワークにおいて、パケットの経路を計算し設定を行うコントローラに障害が発生した場合、パケットの転送処理を行うネットワーク機器は、パケットを転送することができなくなる。或いは、他のコントローラに再度、経路計算を依頼しなければならなくなるため、処理速度が非常に遅くなる。

[0012] このため、制御機能を担当するコントローラの耐障害性を考慮すると、コントローラの多重化が考えられる。

[0013] また、耐障害性や処理速度の観点から、複数のコントローラが同期をとることなく独立してパケットの経路計算とネットワーク機器への設定を行うことが考えられる。

[0014] しかし、可用性と処理速度の観点から、複数のコントローラが同期をとることなく独立して経路を計算し、ネットワーク機器に設定を行う場合、同期を取らないためにネットワーク機器への設定指示が競合することによって、ネットワーク機器の設定情報に不整合が生じる可能性がある。その結果、ネットワーク上に閉路（ループ）が形成され、送信先にパケットが届かなくなる可能性がある。

[0015] 関連する技術として、特許文献1（特開2006-020054号公報）にネットワーク接続装置、経路情報配布プログラム及び経路情報配布方法が開示されている。この関連技術では、子ルータに割り当てられたプレフィックスと割り当て先の子ルータの情報と割り当て管理表とに基づいて経路情報更新通知メッセージを生成する。このメッセージは経路情報の更新（追加、削除）可能性を伝えるものである。既にプレフィックスを割り当て済のルータに対して、経路情報更新通知を送信する。

[0016] また、特許文献2（特開2007-096912号公報）にネットワーク中継器が開示されている。この関連技術では、ネットワークプロセッサ（N

P) より経路情報変更パケットを受け取り、そのイベントを解釈する。イベントの解釈は該経路情報変更パケットのヘッダ情報等により行う。イベント情報が追加であった場合には、経路情報追加処理を実施し、変更であった場合には、経路情報変更処理を、削除であった場合には、経路情報削除処理をそれぞれ実施する。ネットワークプロセッサ (NP) からの経路情報変更パケット全てについて処理が終わったら、ネットワークプロセッサ (NP) へのリクエスト送出处理を起動するためのタイマーをセットする。

先行技術文献

特許文献

[0017] 特許文献1：特開2006-020054号公報

特許文献2：特開2007-096912号公報

非特許文献

[0018] 非特許文献1：“OpenFlow Switch Specification, Version 1.0.0”, [online], December 31, 2009, インターネット (URL: <http://www.openflowswitch.org/documents/openflow-spec-v1.0.0.pdf>)

発明の概要

[0019] 上記の問題を解決するために、複数のコントローラ間で同期を取ることなく独立して経路を計算し、ネットワーク機器に設定を行う場合において、パケットが送信元から送信先へ転送されることを保証する手法が求められている。

[0020] このような手法の1つとして、ネットワーク機器がパケットに情報を付加し、経路上のネットワーク機器がその情報を読み取ることで、正しい経路にパケットを転送する手法が考えられる。

[0021] しかし、ネットワーク機器がパケットの情報を変更する場合、パケットへの処理のオーバーヘッドや、情報付加による転送処理のオーバーヘッドの問題が

発生する可能性もある。

[0022] そこで、本発明では、パケットが送信元から送信先へ転送されることを保証しつつ、ネットワーク機器がパケット自体に対する変更を行わない手法を提案する。

[0023] 本発明に係るネットワークシステムは、複数のネットワーク機器と、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動作が定義されたフローエントリを、経路上のネットワーク機器のフローテーブルに設定する複数のコントローラとを含む。該経路上のネットワーク機器は、該複数のコントローラの中で最も早く該経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、該採用された経路のフローエントリに従って、パケットを転送する。

[0024] 本発明に係るネットワーク機器は、複数のコントローラの各々からの制御に基づいて、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動作が定義されたフローエントリをフローテーブルに設定する装置と、該複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリを設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、該採用された経路のフローエントリに従って、パケットを転送する装置とを具備する。

[0025] 本発明に係る経路制御方法は、ネットワーク機器により実施される経路制御方法であって、複数のコントローラの各々からの制御に基づいて、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動作が定義されたフローエントリをフローテーブルに設定することと、該複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリを設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、該採用された経路のフローエントリに従って、パケットを転送することとを含む。

[0026] 本発明に係るプログラムは、複数のコントローラの各々からの制御に基づ

いて、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動作が定義されたフローエントリをフローテーブルに設定するステップと、該複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、該採用された経路のフローエントリに従って、パケットを転送するステップとをネットワーク機器に実行させるためのプログラムである。

[0027] すなわち、本発明に係るプログラムは、上記の経路制御方法における処理を、ネットワーク機器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。

[0028] コントローラを多重化／冗長化した際の不採用経路の削除が可能になる。

図面の簡単な説明

[0029] [図1]本発明に係るネットワークシステムの構成例を示す図である。

[図2]該ネットワークシステムにおけるコントローラの構成例を示す図である。

[図3]該ネットワークシステムにおけるスイッチの構成例を示す図である。

[図4]該ネットワークシステムにおけるフローエントリの構成例を示す図である。

[図5]該ネットワークシステムにおける経路削除パケットの構成例を示す図である。

[図6] 1 s t パケット受信時の動作を説明するための図である。

[図7]複数のコントローラが設定した経路について説明するための図である。

[図8]採用された経路の順について説明するための図である。

[図9]不採用経路に経路削除パケットが流れる状況について説明するための図である。

[図10]不採用経路の削除処理全体の手順を説明するための図である。

[図11]不採用経路の設定前に、不採用経路の削除処理が開始された場合について説明するための図である。

[図12]本発明の実施例を説明するための図である。

発明を実施するための形態

[0030] 本発明は、C U分離型ネットワークシステムを対象としている。ここでは、C U分離型ネットワークシステムの1つであるオープンフローネットワークシステムを例に説明する。但し、実際には、オープンフローネットワークシステムに限定されない。

[0031] <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。

[0032] [システム構成]

図1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。

[0033] 本発明に係るネットワークシステムは、コントローラ10（10-i、i = 1 ~ x : xは台数）と、スイッチ20（20-j、j = 1 ~ y : yは台数）と、ホスト30（30-k、k = 1 ~ z : zは台数）を含む。

[0034] コントローラ10（10-i、i = 1 ~ x）の各々は、ネットワーク機器におけるパケットの転送経路等を制御する制御装置である。

[0035] スwitch 20（20-j、j = 1 ~ y）の各々は、パケットを転送するネットワーク機器である。

[0036] ホスト30（30-k、k = 1 ~ z）の各々は、パケットを送受信する端末である。

[0037] [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。

[0038] コントローラ10（10-i、i = 1 ~ x）の例として、P C（パソコン）、アプライアンス（a p p l i a n c e）、シンク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。

[0039] スwitch 20（20-j、j = 1 ~ y）の例として、ネットワークスイッ

チ (network switch) を想定している。他にも、ルータ (router)、プロキシ (proxy)、ゲートウェイ (gateway)、ファイアウォール、ロードバランサ、帯域制御装置、セキュリティ監視制御装置、基地局、アクセスポイント、或いは、複数の通信ポートを有する計算機等が考えられる。

[0040] ホスト30 (30-k、k=1~z) の例として、上記の計算機を想定している。他にも、携帯電話機、カーナビ (カーナビゲーションシステム)、携帯型ゲーム機、家庭用ゲーム機、携帯型音楽プレーヤー、ハンディターミナル、ガジェット (電子機器)、双方向テレビ、デジタルチューナー、デジタルレコーダー、情報家電 (information home appliance)、OA (Office Automation) 機器等が考えられる。ホスト30 (30-k、k=1~z) 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。

[0041] なお、ホスト30 (30-k、k=1~z) は、オープンフロー非対応のネットワークにおけるスイッチ等の中継装置でも良い。すなわち、スイッチ20 (20-j、j=1~y) により形成されるネットワークが、オープンフロー対応のネットワーク (オープンフローネットワークシステム) であれば良い。ホスト30 (30-k、k=1~z) のいずれかに接続されるスイッチ20 (20-j、j=1~y) は、オープンフローネットワークシステムにおけるエッジスイッチに相当するため、配下の装置が端末でも外部ネットワークの中継装置でも同様に処理できる。

[0042] また、コントローラ10 (10-i、i=1~x)、スイッチ20 (20-j、j=1~y)、及びホスト30 (30-k、k=1~z) 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 (Virtual Machine (VM)) でも良い。

[0043] 図示しないが、コントローラ10 (10-i、i=1~x)、スイッチ20 (20-j、j=1~y)、及びホスト30 (30-k、k=1~z) 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、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フェースによって実現される。

[0044] 上記のプロセッサの例として、CPU (Central Processing Unit)、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。

[0045] 上記のメモリの例として、RAM (Random Access Memory)、ROM (Read Only Memory)、EEPROM (Electrically Erasable and Programmable Read Only Memory) やフラッシュメモリ等の半導体記憶装置、HDD (Hard Disk Drive) やSSD (Solid State Drive) 等の補助記憶装置、又は、DVD (Digital Versatile Disk) 等のリムーバブルディスクや、SDメモリカード (Secure Digital memory card) 等の記憶媒体 (メディア) 等が考えられる。また、レジスタ (register) でも良い。

[0046] なお、上記のプロセッサ及び上記のメモリは、一体化していても良い。例えば、近年では、マイコン等の1チップ化が進んでいる。従って、上記の計算機等に搭載される1チップマイコンが、プロセッサ及びメモリを備えている事例が考えられる。

[0047] 上記のインタフェースの例として、ネットワーク通信に対応した基板 (マザーボード、I/Oボード) やチップ等の半導体集積回路、NIC (Network Interface Card) 等のネットワークアダプタや同様の拡張カード、アンテナ等の通信装置、接続口 (コネクタ) 等の通信ポート等が考えられる。

[0048] また、ネットワークの例として、インターネット、LAN (Local

Area Network)、無線LAN (Wireless LAN)、WAN (Wide Area Network)、バックボーン (Back bone)、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。

[0049] 但し、実際には、これらの例に限定されない。

[0050] [コントローラの詳細]

図2を参照して、コントローラ10 (10-i、 $i = 1 \sim x$) の構成例について説明する。

[0051] コントローラ10 (10-i、 $i = 1 \sim x$) の各々は、経路計算部11と、ネットワーク管理部12と、パケット処理部13を備える。

[0052] 経路計算部11は、ネットワーク管理部12からスイッチ20 (20-j、 $j = 1 \sim y$) のトポロジー情報を受け取り、パケットの経路を計算し、スイッチ20 (20-j、 $j = 1 \sim y$) に設定を行う。

[0053] ネットワーク管理部12は、スイッチ20 (20-j、 $j = 1 \sim y$) のトポロジー情報や、経路計算部11で計算した経路の管理を行う。ここでは、ネットワーク管理部12は、トポロジー情報や計算した経路を基に、オープンフロープロトコルに準拠した制御メッセージにより、スイッチ20 (20-j、 $j = 1 \sim y$) の挙動を制御する。

[0054] パケット処理部13は、スイッチ20 (20-j、 $j = 1 \sim y$) から転送されたパケットの受信、パケットの情報解析、経路計算部11へ解析結果の送信を行う。

[0055] なお、経路計算部11、ネットワーク管理部12、及びパケット処理部13は、モジュール (module) 又はコンポーネント (component) でも良い。

[0056] コントローラ10 (10-i、 $i = 1 \sim x$) は、複数のスイッチ20 (2

0-j、j=1~y)をネットワーク管理部12によって管理することができ、スイッチ20(20-j、j=1~y)と専用線、若しくは通常のネットワークのセキュアチャネル(Secure Channel)で接続されている。

[0057] [スイッチの詳細]

図3を参照して、スイッチ20(20-j、j=1~y)の構成例について説明する。

[0058] スwitch20(20-j、j=1~y)の各々は、フローテーブル21と、フローテーブル管理部22と、パケット処理部23を備える。

[0059] フローテーブル21は、経路情報となるフローエントリを管理する。フローテーブル21は、記憶装置や記憶媒体に格納することができる。

[0060] フローテーブル管理部22は、コントローラ10(10-i、i=1~x)からの制御メッセージに応じて、フローテーブル21のフローエントリを追加・更新・削除する。フローテーブル管理部22は、フローテーブル21のフローエントリの内容を確認することもできる。

[0061] パケット処理部23は、パケットを処理する。すなわち、パケット処理部23は、スイッチ20(20-j、j=1~y)内部において、パケットに関する処理全般を行う。

[0062] パケット処理部23は、パケット受信部231と、パケット転送部232と、パケット破棄部233と、経路削除パケット生成部234と、経路削除パケット解析部235を備える。

[0063] パケット受信部231は、パケットを受信し解析する。パケット受信部231は、パケット解析として、パケットのヘッダ情報を参照し、フローテーブル21に該パケットのヘッダ情報と適合するフローエントリがあるか否かを確認する。適合するフローエントリがある場合、該フローエントリの内容が経路上の次段のノードに転送する旨であれば、パケット受信部231は、パケット転送部232にパケットを渡し、経路上の次段のノードに転送する。また、該フローエントリの内容がパケットを破棄(廃棄)する旨であれば

、パケット受信部231は、パケット破棄部233にパケットを渡し、パケットを破棄する。更に、適合するフローエントリがない場合、パケット受信部231は、パケット転送部232にパケットを渡し、コントローラ10（10-i、 $i = 1 \sim x$ ）に該パケットの経路を問い合わせる。

[0064] パケット転送部232は、該スイッチに隣接する他のスイッチ20（20-j、 $j = 1 \sim y$ ）、コントローラ10（10-i、 $i = 1 \sim x$ ）、ホスト30（30-k、 $k = 1 \sim z$ ）等にパケットを転送する。パケット転送部232は、出力ポート毎に設けられ、受け取ったパケットを該出力ポートから出力する。

[0065] パケット破棄部233は、受け取ったパケットを破棄する。なお、パケット破棄部233は、通常パケットのみならず、後述する経路削除パケットも破棄することができる。

[0066] 経路削除パケット生成部234は、不採用経路を削除するための経路削除パケットを生成する。

[0067] 経路削除パケット解析部235は、受信した経路削除パケットを解析する。経路削除パケット解析部235は、パケット解析として、経路削除パケットのヘッダ情報を参照し、フローテーブル21に経路削除パケットのヘッダ情報と適合する不採用経路のフローエントリがあるか否かを確認する。適合するフローエントリがある場合、該フローエントリの内容が経路上の次段のノードに転送する旨であれば、パケット受信部231は、パケット転送部232に経路削除パケットを渡し、経路上の次段のノードに転送する。また、該フローエントリの内容が経路削除パケットを破棄する旨であれば、パケット受信部231は、パケット破棄部233に経路削除パケットを渡し、経路削除パケットを破棄する。更に、適合するフローエントリがない場合、経路削除パケット解析部235は、経路削除パケットを一時的に保持し、所定の時間の経過後に、パケット破棄部233に経路削除パケットを渡し、経路削除パケットを破棄する。経路削除パケット解析部235は、パケット受信部231と一体化していても良い。

[0068] なお、フローテーブル管理部 22 や、パケット処理部 23（パケット受信部 231、パケット転送部 232、パケット破棄部 233、経路削除パケット生成部 234、及び経路削除パケット解析部 235）は、モジュール（module）又はコンポーネント（component）でも良い。

[0069] [フローテーブルの詳細]

フローテーブル 21 は、フローエントリ 211 を保持する。すなわち、フローテーブル 21 は、フローエントリ 211 の集合である。フローエントリ 211 は、各フローに対応する経路情報である。

[0070] [フローエントリの詳細]

図 4 を参照して、フローエントリ 211 の構成例について説明する。

[0071] フローエントリ 211 は、フローヘッダ（Flow header）2111 と、アクション（Action）2112 と、経路 ID 2113 と、コントローラ ID 2114 を含む。

[0072] フローヘッダ 2111 は、送信元 IP アドレスや送信先 IP アドレス等のフローの基本情報を記述するための領域（フィールド）である。すなわち、フローヘッダ 2111 は、フローエントリ 211 のルールを記述するための領域である。

[0073] アクション 2112 は、フローの処理方法を記述するための領域である。すなわち、アクション 2112 は、フローエントリ 211 のアクションを記述するための領域である。

[0074] 経路 ID 2113 は、経路毎に一意な値（経路の識別子）を記述するための領域である。

[0075] コントローラ ID 2114 は、フローエントリ 211 を設定したコントローラ 10（10-i、 $i = 1 \sim x$ ）の識別子を記述するための領域である。

[0076] [経路削除パケットの詳細]

図 5 を参照して、本発明で定義する経路削除パケットについて説明する。

[0077] 経路削除パケット 50 は、スイッチ 20（20-j、 $j = 1 \sim y$ ）のフローテーブル 21 を構成するフロー毎のフローエントリ 211 を削除するため

の制御メッセージのパケットである。

[0078] 経路削除パケット50は、フローヘッダ51と、経路ID52と、コントローラID53を含む。

[0079] フローヘッダ51は、図4に示したフローヘッダ2111と同様、送信元IPアドレスや送信先IPアドレス等のフローの基本情報を記述するための領域である。

[0080] 経路ID52は、図4に示した経路ID2113と同様、経路毎に一意な値（経路の識別子）を記述するための領域である。

[0081] コントローラID53は、図4に示したコントローラID2114と同様、フローエントリ211を設定したコントローラ10（10-i、 $i=1\sim x$ ）の識別子を記述するための領域である。

[0082] [本発明における経路制御処理]

図6を参照して、図1のような構成で、以下の（1）から（6）の順に、1stパケット（First Packet）受信時の動作について説明する。

[0083] （1）ステップS1

スイッチ20（20-j、 $j=1\sim y$ ）は、フローテーブル21に情報のないパケット（1stパケット）が届いた場合、複数のコントローラ10（10-i、 $i=1\sim x$ ）に1stパケットを転送する。すなわち、スイッチ20（20-j、 $j=1\sim y$ ）は、受信パケットのヘッダ情報と適合するフローテーブル21のフローエントリ211がない場合、該パケットを1stパケットと判断して、複数のコントローラ10（10-i、 $i=1\sim x$ ）に該1stパケットを転送する。

[0084] （2）ステップS2

コントローラ10（10-i、 $i=1\sim x$ ）は、1stパケットを受け取った場合、経路を計算し、経路上のスイッチ20（20-j、 $j=1\sim y$ ）に設定する。

[0085] （3）ステップS3

コントローラ 10 (10-i、i=1~x) は、スイッチ 20 (20-j、j=1~y) の設定を終えた場合、1st パケットの送信元であるスイッチ 20 (20-j、j=1~y) に応答する。

[0086] (4) ステップ S4

スイッチ 20 (20-j、j=1~y) は、最も早く応答したコントローラ 10 (10-i、i=1~x) ではないコントローラ 10 (10-i、i=1~x) に不採用通知を出す。

[0087] (5) ステップ S5

スイッチ 20 (20-j、j=1~y) は、不採用となったコントローラ 10 (10-i、i=1~x) の設定したフローエントリ 211 の削除処理を行う。

[0088] (6) ステップ S6

スイッチ 20 (20-j、j=1~y) は、採用された経路のフローエントリ 211 に従って、パケットを処理する。

[0089] [ステップ S1 の詳細]

スイッチ 20 (20-j、j=1~y) は、フローテーブル 21 に情報のないパケット (1st パケット) が届いた場合、複数のコントローラ 10 (10-i、i=1~x) に 1st パケットを転送する。

[0090] 例えば、スイッチ 20-1 がホスト 30-1 からパケットを受け取るとき、パケット処理部 23 は、通常パケットか経路削除パケット 50 かの判定を行う。

[0091] パケットが通常パケットであるとき、パケット処理部 23 は、そのパケットのヘッダ情報が、フローテーブル 21 に設定されているフローエントリ 211 のフローヘッダ 2111 と適合するか確認する。

[0092] 適合する場合、パケット処理部 23 は、該当フローエントリ 211 のアクション 2112 に従ってパケットの処理を行う。

[0093] 適合しなかった場合、パケット処理部 23 は、そのパケット (1st パケット) をスイッチ 20-1 の管理を行うコントローラ 10-1 とコントローラ

ラ 10-2 に転送する。

[0094] [ステップ S 2 の詳細]

コントローラ 10 (10-i、 $i=1\sim x$) は、1st パケットを受け取った場合、経路を計算し、経路上のスイッチ 20 (20-j、 $j=1\sim y$) に設定する。

[0095] コントローラ 10-1 とコントローラ 10-2 のそれぞれにおいて、パケット処理部 13 が 1st パケットを受け取った際、経路計算部 11 は、ネットワーク管理部 12 のトポロジー情報と、パケット処理部 13 の解析結果を基に、そのパケット (1st パケット) が通るべき経路を計算する。

[0096] ここで、図 7 のように、コントローラ 10-1 が計算した経路を経路 100-1、コントローラ 10-2 が計算した経路を経路 100-2 とする。

[0097] 経路 100-1 は、「スイッチ 20-1」→「スイッチ 20-4」→「スイッチ 20-5」→「スイッチ 20-6」→「スイッチ 20-7」→「スイッチ 20-8」の順となる。

[0098] 経路 100-2 は、「スイッチ 20-1」→「スイッチ 20-2」→「スイッチ 20-4」→「スイッチ 20-5」→「スイッチ 20-7」→「スイッチ 20-6」→「スイッチ 20-8」の順となる。

[0099] 経路の計算終了後、コントローラ 10-1 とコントローラ 10-2 のそれぞれにおいて、ネットワーク管理部 12 は、計算して得られた経路に基づくフローエントリ 211 を、経路上の各スイッチ 20 (20-j、 $j=1\sim y$) に設定する。

[0100] スイッチ 20 (20-j、 $j=1\sim y$) へのフローエントリ 211 の設定については、経路上の送信元ホスト 30-1 に最も近いスイッチ 20-1 から順に行う。

[0101] 例えば、経路 100-1 の場合、コントローラ 10-1 は、図 8 の (1) から (6) の順に、スイッチ 20 (20-j、 $j=1\sim y$) の設定を行う。この順序は、通常のオープンフローの設定順序とは異なる。

[0102] 経路上の送信元ホスト 30-1 に近いスイッチ 20-1 から順に設定を行

う場合、スイッチ20-1は、設定された瞬間にフローエントリ211が有効となり、そのフローエントリ211のアクション2112に従って受信パケットを処理してしまうことになる。そのため、例えば、経路上の転送先であるスイッチ20-4の設定が終了する前に、スイッチ20-4がスイッチ20-1からパケットを受信した場合、スイッチ20-4は、該パケットを1stパケットとして、再度コントローラ10(10-i、i=1~x)に転送してしまう可能性がある。

[0103] 従って、オープンフローにおいては経路上の送信先ホスト30-2に近いスイッチ20(20-j、j=1~y)から順に設定を行う。

[0104] 本発明では、経路上の送信元ホスト30-1に最も近いスイッチ20-1の設定のみ仮設定(仮登録)とすることで、全てのスイッチ20(20-j、j=1~y)の設定が終了する前にパケットがフローエントリ211に従ってされることを防ぐ。

[0105] すなわち、経路上の全てのスイッチ20(20-j、j=1~y)の設定が完了しない限り、パケットが送信元ホスト30-1から送信先ホスト30-2に転送されることはない。

[0106] 経路100-1の場合、コントローラ10-1は、経路100-1上の送信元ホスト30-1に最も近いスイッチ20-1にフローエントリ211を仮設定した後、以降のスイッチ20(20-j、j=1~y)は仮設定することなくフローテーブル21にフローエントリ211を設定する。

[0107] 例えば、スイッチ20(20-j、j=1~y)が保持するフローテーブル21とは別に、仮のフローエントリ211を保持することのできる記憶領域を持つことで、そこに仮設定するフローエントリ211を保持しても良い。別の方法では、フローエントリ211に仮設定なのか、本登録なのかを判断できる領域を持つことでも実現可能である。

[0108] [ステップS3の詳細]

コントローラ10(10-i、i=1~x)は、スイッチ20(20-j、j=1~y)の設定を終えた場合、1stパケットの送信元であるスイッ

チ 20 (20-j、j=1~y) に応答する。

[0109] コントローラ 10 (10-i、i=1~x) は、経路上の全てのスイッチ 20 (20-j、j=1~y) の設定が完了すると、スイッチ 20-1 に応答として設定完了通知を送信する。

[0110] このとき、コントローラ 10 (10-i、i=1~x) は、スイッチ 20-1 に仮設定したフローエントリ 211 をフローテーブル 21 に登録する。

[0111] スwitch 20-1 は、最も早く応答したコントローラ 10 (10-i、i=1~x) が設定した経路を、パケットが通る経路として採用する。

[0112] 仮に、コントローラ 10-1、コントローラ 10-2 が同時に応答した場合、スイッチ 20-1 は、採用された経路が 1 つとなるようにどちらかを選択する。

[0113] [ステップ S4 の詳細]

スイッチ 20 (20-j、j=1~y) は、最も早く応答したコントローラ 10 (10-i、i=1~x) ではないコントローラ 10 (10-i、i=1~x) に不採用通知を出す。

[0114] 採用された経路は、コントローラ 10-1 が設定した経路 100-1 であると仮定する。

[0115] コントローラ 10-1 が最も早く応答したため、スイッチ 20-1 は、不採用となったコントローラ 10-2 に経路不採用を通知する。

[0116] これは、不採用が決まった時点で設定途中のコントローラ 10-2 に経路の設定を中止させるために行う。

[0117] そのため、コントローラ 10-2 は、スイッチ 20-1 から不採用通知を受け取ると、コントローラ 10-2 による経路上の全てのスイッチ 20 (20-j、j=1~y) の設定が完了していなくとも、残りの設定を行わない。

[0118] [ステップ S5 の詳細]

スイッチ 20 (20-j、j=1~y) は、不採用となったコントローラ 10 (10-i、i=1~x) の設定したフローエントリ 211 の削除処理

を行う。

[0119] ここでは、スイッチ20-1は、コントローラ10-2の設定した不採用経路100-2のフローエントリ211の削除を行う。

[0120] 不採用経路100-2の削除には、経路削除パケット50を用いて行う。
スイッチ20-1は、不採用経路100-2上のスイッチに向けて、経路削除パケット50を送信する。

[0121] 図9に示すように、不採用経路100-2に、経路削除パケット50が流れる。

[0122] [不採用経路の削除処理]

図10を参照して、不採用経路の削除処理全体の手順について説明する。

[0123] (1) ステップS501

スイッチ20-1において、パケット処理部23は、経路削除パケット50に、コントローラ10-2の設定した不採用経路100-2の情報を設定する。ここでは、パケット処理部23は、不採用経路100-2のフローエントリ211のフローヘッダ2111、経路ID2113、及びコントローラ10-2を示すコントローラID2114を基に、経路削除パケット50を作成する。すなわち、経路削除パケット50のフローヘッダ51、経路ID52、及びコントローラID53は、それぞれ上記の不採用経路100-2のフローヘッダ2111、経路ID2113、及びコントローラID2114と等しい。

[0124] (2) ステップS502

スイッチ20-1において、パケット処理部23は、不採用経路100-2のフローエントリ211のアクション2112の内容が送信先ホスト30-2に転送する内容か確認する。すなわち、経路の末端のスイッチであるか確認する。

[0125] (3) ステップS503

スイッチ20-1において、パケット処理部23は、フローエントリ211のアクション2112の内容が送信先ホスト30-2に転送する内容では

ない場合、フローエントリ 211 のアクション 2112 に従って経路削除パケット 50 を処理する。ここでは、スイッチ 20-1 において、パケット処理部 23 は、フローエントリ 211 のアクション 2112 に従って、経路削除パケット 50 を、次段のスイッチ 20-2 に転送する。

[0126] (4) ステップ S504

不採用経路上にある各スイッチ 20 (20-j、j=2~y) において、フローテーブル管理部 22 は、経路削除パケット 50 を次段のスイッチ 20 (20-j、j=2~y) に転送した後、仮設定又は正式に設定されている不採用経路 100-2 のフローエントリ 211 を削除する。

[0127] (5) ステップ S505

不採用経路上にある各スイッチ 20 (20-j、j=2~y) において、パケット処理部 23 は、経路削除パケット 50 を受信した際、経路削除パケット 50 のフローヘッダ 51、経路 ID 52、及びコントローラ ID 53 を基に、フローテーブル 21 に登録されているフローエントリ 211 の中に、経路削除パケット 50 のフローヘッダ 51、経路 ID 52、コントローラ ID 53 と適合するフローエントリ 211 があるか確認する。

[0128] (6) ステップ S506

不採用経路上にある各スイッチ 20 (20-j、j=2~y) において、パケット処理部 23 は、経路削除パケット 50 のフローヘッダ 51、経路 ID 52、コントローラ ID 53 と適合するフローエントリ 211 がある場合、該フローエントリ 211 のアクション 2112 の内容が送信先ホスト 30-2 に転送する内容か確認する。すなわち、経路の末端のスイッチであるか確認する。

[0129] (7) ステップ S507

不採用経路上にある各スイッチ 20 (20-j、j=2~y) において、パケット処理部 23 は、フローエントリ 211 のアクション 2112 の内容が送信先ホスト 30-2 に転送する内容ではない場合、フローエントリ 211 のアクション 2112 に従って経路削除パケット 50 を処理する。ここで

は、各スイッチ20（20-j、j=2~y）において、パケット処理部23は、フローエントリ211のアクション2112に従って、経路削除パケット50を、次段のスイッチ20（20-j、j=2~y）に転送する。

[0130] （8）ステップS508

不採用経路上にある各スイッチ20（20-j、j=2~y）において、フローテーブル管理部22は、経路削除パケット50を次段のスイッチ20（20-j、j=2~y）に転送した後、仮設定又は正式に設定されている不採用経路100-2のフローエントリ211を削除する。

[0131] （9）ステップS509

不採用経路上にある各スイッチ20（20-j、j=1~y）において、パケット処理部23は、フローエントリ211のアクション2112の内容が送信先ホスト30-2に転送する場合、すなわち、スイッチ20（20-j、j=1~y）が経路の最後のスイッチ20-8の場合、転送処理を行わず、経路削除パケット50のフローヘッダ51、経路ID52、コントローラID53と適合するフローエントリ211を削除し、経路削除パケット50を破棄する。なお、1stパケットを受信したスイッチ20-1が経路の最後のスイッチである場合、スイッチ20-1は、経路削除パケット50を生成せずに、不採用経路のフローエントリ211を削除しても良い。例えば、不採用経路がスイッチ20-1の配下のホスト間の通信経路である場合や、スイッチ20-1の配下のホスト30（30-k、k=1~z）の各々が、オープンフロー非対応のネットワークにおけるルータ等の中継装置である場合が考えられる。

[0132] （10）ステップS510

また、不採用経路上にあるスイッチ20（20-j、j=2~y）は、フローテーブル21に登録されているフローエントリ211の中に、経路削除パケット50のフローヘッダ51、経路ID52、コントローラID53と適合するフローエントリ211がない場合、受信した経路削除パケット50を一時的に保持する。

[0133] (11) ステップS511

スイッチ20 (20-j、j=2~y) は、コントローラ10 (10-i、i=1~x) からフローエントリ211の追加設定があるか確認する。

[0134] (12) ステップS512

スイッチ20 (20-j、j=2~y) は、コントローラ10 (10-i、i=1~x) からフローエントリ211の追加設定がある場合、追加設定されたフローエントリ211のフローヘッダ2111、経路ID2113、及びコントローラID2114が、一時的に保持されている経路削除パケット50のフローヘッダ51、経路ID52、及びコントローラID53と適合するか確認する。

[0135] (13) ステップS513

フローエントリ211と経路削除パケット50が適合する場合、スイッチ20 (20-j、j=2~y) は、そのフローエントリ211の登録を破棄する。

[0136] (14) ステップS514

スイッチ20 (20-j、j=2~y) は、経路削除パケット50を一時的に保持してから所定の時間が経過したか確認する。

[0137] (15) ステップS515

スイッチ20 (20-j、j=2~y) は、経路削除パケット50を一時的に保持してから所定の時間が経過した場合、保持している経路削除パケット50を破棄する。

[0138] [ステップS510以降の詳細]

例えば、図11のように、経路100-2が不採用となったコントローラ10 (10-i、i=1~x) がスイッチ20-5までしか設定を完了させることができなかった状態で、経路の削除処理が始まってしまった場合等が考えられる。

[0139] この場合、スイッチ20-7は、不採用となった経路100-2のフローエントリ211が設定されていないため、経路削除パケット50を処理でき

ない。

[0140] このとき、スイッチ 20-7 は、受信した経路削除パケット 50 を一時的に保持する。

[0141] 経路削除パケット 50 を一時的に保持する理由は、フローエントリ 211 の登録処理が遅れて行われる可能性があるからである。

[0142] 不採用経路の削除処理が始まった時点で、経路が不採用となったコントローラ 10-2 は、不採用通知を受け取っているため、スイッチ 20-5 より先にある不採用経路の残りのスイッチ（「スイッチ 20-7」→「スイッチ 20-6」→「スイッチ 20-8」）に対してはフローエントリ 211 の設定を行わない。

[0143] しかし、ネットワークの遅延等の理由から、不採用経路の削除処理後にフローエントリ 211 が設定される可能性もあるので、スイッチ 20-7 は、経路が不採用となったコントローラ 10-2 が遅れてフローエントリ 211 を設定してきた場合のために、スイッチ 20-5 から受信した経路削除パケット 50 を保持しておく。

[0144] コントローラ 10-2 からフローエントリ 211 が遅れて設定された場合、スイッチ 20-7 は、そのフローエントリ 211 の登録を破棄する。

[0145] なお、コントローラ 10-2 からフローエントリ 211 が遅れて設定された場合、スイッチ 20-7 は、念のために、そのフローエントリ 211 に従って、次段のスイッチ 20-6 に経路削除パケット 50 を送信するようにしても良い。スイッチ 20-6 も、経路削除パケット 50 を受信した場合、上記のスイッチ 20-7 と同様の処理を行うものとする。

[0146] 一定時間、保持する経路削除パケット 50 の情報と適合するフローエントリ 211 が登録されない場合、スイッチ 20-7 は、登録を破棄すべき不採用経路のフローエントリはないと判断し、経路削除パケット 50 を破棄する。

[0147] [ステップ S6 の詳細]

スイッチ 20 (20-j、j=1~y) は、採用された経路のフローエン

トリ 2 1 1 に従って、パケットを処理する。

[0148] ここでは、スイッチ 2 0 - 1 は、最も早く応答したコントローラ 1 0 - 1 が仮設定したフローエントリ 2 1 1 をフローテーブル 2 1 に設定する。

[0149] この時点で、スイッチ 2 0 - 1 は、1 s t パケットの情報と一致するフローエントリ 2 1 1 が設定されるため、1 s t パケットを該当フローエントリ 2 1 1 のアクション 2 1 1 2 に従って処理する。

[0150] 仮設定のフローエントリ 2 1 1 をフローテーブル 2 1 に設定するタイミングは、ステップ S 5 における「不採用経路の削除処理」とほぼ同時である。

[0151] スwitch 2 0 - 1 は、経路削除パケット 5 0 をスイッチ 2 0 - 4 に転送した後に、仮設定のフローエントリ 2 1 1 をフローテーブル 2 1 に設定する。

[0152] 仮設定のフローエントリ 2 1 1 がフローテーブル 2 1 に設定された後、スイッチ 2 0 - 1 は、受信したパケットを、フローエントリ 2 1 1 に従って、採用された経路上のスイッチ 2 0 - 4 に転送する。

[0153] 以降、採用された経路上のスイッチ 2 0 (2 0 - j、j = 1 ~ y) の各々は、受信したパケットを、フローエントリ 2 1 1 に従って、採用された経路上の次段のノードに転送し、パケットをホスト 3 0 - 2 まで転送する。

[0154] ここで、不採用経路の削除処理が終了していないスイッチ 2 0 (2 0 - j、j = 1 ~ y) がパケットを受け取る可能性がある。

[0155] その場合、該スイッチ 2 0 (2 0 - j、j = 1 ~ y) は、どのフローエントリ 2 1 1 が経路 1 0 0 - 1 か判断できないため、不採用経路の削除処理の終了を待つ。

[0156] ここで、スイッチ 2 0 (2 0 - j、j = 1 ~ y) は、次の 2 つの場合から、不採用経路の削除処理が終了したことを判断する。

[0157] (1) パケットのフローヘッダの情報に適合するフローヘッダ 2 1 1 1 を持つフローエントリ 2 1 1 が 1 つしかない場合、スイッチ 2 0 (2 0 - j、j = 1 ~ y) は、採用されたコントローラ 1 0 (1 0 - i、i = 1 ~ x) から設定されたフローエントリ 2 1 1 のみが登録されている状態であると判断し、不採用経路の削除処理が終了したことを判断する。

[0158] (2) パケットのフローヘッダの情報に適合するフローエントリ 211 は複数存在するが、これらのフローエントリ 211 のコントローラ ID 2113 及び経路 ID 2113 が全て同一のものである場合、スイッチ 20 (20-j、j = 1 ~ y) は、これらのフローエントリ 211 全てが、採用されたコントローラ 10 (10-i、i = 1 ~ x) から登録されている状態であると判断し、不採用経路の削除処理が終了したことを判断する。

[0159] スイッチ 20 (20-j、j = 1 ~ y) は、不採用経路の削除処理の終了を一定時間判断できない場合、すなわち、ネットワークの遅延等の原因によって、一定時間待っても経路削除パケット 50 が届かず、不採用経路が削除されない場合、その時点で 1st パケットをコントローラ 10 (10-i、i = 1 ~ x) に転送し、経路の再設定を行う。

[0160] [実施例]

図 12 を参照して、本発明の実施例について説明する。

[0161] ここでは、コントローラ 10 (10-i、i = 1 ~ x) の例として、コントローラ 1 及びコントローラ 2 を示す。また、スイッチ 20 (20-j、j = 1 ~ y) の例として、スイッチ 1 ~ スイッチ 4 を示す。また、ホスト 30 (30-k、k = 1 ~ z) の例として、ホスト 1 及びホスト 2 を示す。ホスト 1 は、送信元ホストであり、ホスト 2 は送信先ホストである。

[0162] (1) スイッチ 1 は、ホスト 1 からパケットを受信する。

[0163] (2) スイッチ 1 は、自身が保持するフローエントリに一致しないパケットである場合、該パケットを 1st パケットとして、自身を制御する複数のコントローラに転送する。ここでは、スイッチ 1 は、1st パケットを、コントローラ 1 及びコントローラ 2 に転送する。

[0164] (3) コントローラ 1 及びコントローラ 2 は、スイッチから 1st パケットを転送された場合、経路を計算し、経路上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。このとき、コントローラ 1 及びコントローラ 2 は、1st パケットを受信したスイッチ 1 から順に、経路上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。但し、1st パケットを受信したス

スイッチ 1 に設定したフローエントリには、仮設定であることを示す情報（フラグ等）を付与しておく。すなわち、この時点では、1 s t パケットを受信したスイッチ 1 に設定したフローエントリは、正式に設定したフローエントリではなく、仮設定のフローエントリである。仮設定であることを示す情報とは、該フローエントリが未だ無効である旨を示す情報である。1 s t パケットを受信したスイッチ 1 は、仮設定のフローエントリに従って、パケットを転送することはしない。なお、仮設定であることを示す情報を格納する領域は、フローエントリ内に新たに設けても良いし、フローエントリ内の既存の領域の一部を利用しても良い。

[0165] (4) コントローラ 1 及びコントローラ 2 は、経路上のスイッチの設定が全て完了した場合、設定が完了した旨の通知（設定完了通知）を送信し、1 s t パケットを受信したスイッチ 1 に設定したフローエントリから仮設定を示す情報を取り除く。

[0166] (5) 1 s t パケットを受信したスイッチ 1 は、最も早く（最先に）全てのスイッチの設定が完了したコントローラ 1 の設定した経路をパケットの通信路に採用し、経路を採用したコントローラ 1 以外のコントローラ 2 に不採用通知を送る。コントローラ 2 は、不採用通知を受け取った際に、経路上のスイッチの設定が途中の場合、以降の設定を中止する。ここでは、コントローラ 2 は、経路上のスイッチ 4 の設定前に、不採用通知を受け取り、経路上のスイッチ 4 の設定を中止する。

[0167] (6) 1 s t パケットを受信したスイッチ 1 は、採用されなかった経路を削除するために、経路削除パケットを生成し、経路削除パケットを、採用されなかった経路のフローエントリに従って、隣接するスイッチ 3 に転送し、採用されなかった経路のフローエントリを削除する。スイッチ 3 を始め、採用されなかった経路上のスイッチは、経路削除パケットを受信した際、経路削除パケットのヘッダ情報と適合するフローエントリに従って、隣接するスイッチ 4 に転送した後、該フローエントリを削除する。スイッチ 4 は、経路削除パケットのヘッダ情報と適合するフローエントリが設定されていないた

め、経路削除パケットを一時的に保持し、所定の時間が経過した後に、経路削除パケットを破棄する。なお、スイッチ4は、経路削除パケットのヘッダ情報と適合するフローエントリが設定されていた場合、該フローエントリの内容から、採用されなかった経路上の最後のスイッチであると判明するため、隣接するホスト2には経路削除パケットを送信せず、経路削除パケットのヘッダ情報と適合するフローエントリを削除し、経路削除パケットを破棄する。これらの処理の詳細については、図10に示した不採用経路の削除処理と同様である。

[0168] (7) 1stパケットを受信したスイッチ1は、1stパケット以降のパケットを、採用された経路のフローエントリに従って、隣接するスイッチ2に転送する。

[0169] (8) スwitch2は、1stパケット以降のパケットを、フローエントリに従って、隣接するスイッチ4に転送する。

[0170] (9) 採用された経路上の最後のスイッチ4は、1stパケット以降のパケットを、採用された経路のフローエントリに従って、ホスト2に転送する。

[0171] [補足]

本発明の実施例では、1stパケットを受信したスイッチが2台のコントローラによって制御されているが、3台以上のコントローラによる制御も実現可能である。すなわち、図示された2台のコントローラは、「経路を採用したコントローラ」と、「経路を採用されなかったコントローラ」を示すものであり、個々のコントローラの台数は任意である。

[0172] <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。

[0173] オープンフローにおいて、スイッチのフローテーブルにフローエントリを登録する方式は、大きく「Proactive型」と、「Reactive型」の2つの方式に分けられる。

[0174] 「Proactive型」では、コントローラが「事前に（データ通信が

始まる前に)」所定の packets 群 (フロー) の経路 (パス) を計算し、スイッチのフローテーブルにフローエントリを登録する。すなわち、ここでいう「Proactive型」とは、コントローラが自発的に行う「事前のフローエントリ登録」を指す。

[0175] 「Reactive型」では、コントローラが「スイッチから 1st packets (該当フローエントリがない新規の packets) についての問い合わせを受けた際に」該 packets 群 (フロー) の経路を計算し、スイッチのフローテーブルにフローエントリを登録する。すなわち、ここでいう「Reactive型」とは、実際のデータ通信時に、コントローラがスイッチからの問い合わせに応じて行う「リアルタイムのフローエントリ登録」を指す。

[0176] オープンフローネットワークシステムでは、基本的に、コントローラがスイッチから 1st packets についての問い合わせを受けた際に該受信 packets に関するフローエントリを登録する「Reactive型」が中心となっている。

[0177] しかし、実際のハードウェア (HW) では、フローテーブルの処理頻度を軽減し性能の問題を解決するためには、「Proactive型」が好適であると考えられる。例えば、大量の 1st packets がコントローラに到着しても処理し切れるようにするためには、「Proactive型」の方が好適であると考えられる。但し、実際には、完全な「Proactive型」にするとフローエントリ数が膨大になると考えられるため、一部を「Reactive型」にすることにより、フローエントリ数の制約から逃れるといったことも考えられる。

[0178] また、「Proactive型」を用いれば、通信開始前にフローを定義できるため、Nimda等のウィルスによる大量フロー発生問題や、不明な packets による不正アクセス等が回避可能になると考えられる。

[0179] 上記の第1実施形態における説明では、「Reactive型」を前提に説明してきたが、実際には、「Proactive型」のオープンフローネットワークシステムにおいて、本発明を実施することも可能である。

[0180] すなわち、複数のコントローラの各々が、スイッチからのパケットの問い合わせを受けることなく、事前に（データ通信が始まる前に）、到着が予想される所定のパケット群の経路を計算し、経路上のスイッチのフローテーブルにフローエントリを登録するようにしても良い。スイッチにおける不採用経路の削除処理等については、上記の第1実施形態と同様である。

[0181] <まとめ>

以上のように、本発明は、オープンフローに代表される制御機能を外部のコントローラに分離したネットワーク機器で構成されるネットワークにおいて、コントローラを多重化／冗長化した際の不採用経路の削除を実現する。

[0182] 本発明では、パケット転送と経路制御の機能が分離されたネットワーク機器で構成されるネットワークを対象とする。

[0183] 本発明では、複数のコントローラが独立して動作し、コントローラ間の通信を伴わずに全てのネットワーク機器に対してフローエントリを設定できる。

[0184] 本発明では、経路は、複数のコントローラの中で最も早く設定を完了させたコントローラの設定した経路である。

[0185] 本発明では、ネットワーク機器が、通信に使用されない経路のフローエントリを削除する経路削除パケットを生成する。

[0186] 本発明では、ネットワーク機器が、経路削除パケットを、通信に使用されない経路上の隣接するネットワーク機器に転送する。

[0187] 本発明では、ネットワーク機器は、経路削除パケットを受信した際、経路削除パケットを解析し、解析結果を基に、自身に設定されたフローエントリのうち、経路削除パケットのヘッダ情報と適合するフローエントリを削除する。

[0188] 本発明では、ネットワーク機器が、経路削除パケットに適合するフローエントリがない場合、経路削除パケットを一時的に保持する。

[0189] 本発明では、ネットワーク機器が、コントローラによって登録されるフローエントリと、一時的に保持している経路削除パケットのヘッダ情報とが適

合する場合、その登録を破棄する。

[0190] 本発明により、複数のコントローラ間で同期を取ることなく、通信路を設定可能となる。また、不採用経路を削除することで、ネットワーク上に閉路が作成されることがなく、送信先へ正しく転送されないという障害が発生しない。また、ネットワーク機器がパケットに対して情報の付加や削除といった処理を行わない。

[0191] <付記>

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

[0192] (付記1)

複数のコントローラの各々からの制御に基づいて、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動作が定義されたフローエントリをフローテーブルに設定する装置と、

複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、採用された経路のフローエントリに従って、パケットを転送する装置と

を具備する

ネットワーク機器。

[0193] (付記2)

付記1に記載のネットワーク機器であって、

複数のコントローラの各々から、経路上のネットワーク機器全ての設定を完了させた旨の設定完了通知を受信する装置と、

設定完了通知に応じて、不採用とした経路のフローエントリと適合する経路削除パケットを作成する装置と、

外部から経路削除パケットを受信した場合、経路削除パケットと適合するフローエントリの有無を確認する装置と、

不採用とした経路のフローエントリに従って、経路削除パケットを、不採

用とした経路上の隣接するネットワーク機器に転送する装置と、
不採用とした経路のフローエントリを削除する装置と
を更に具備する
ネットワーク機器。

[0194] (付記3)

付記2に記載のネットワーク機器であって、
経路削除パケットと適合するフローエントリの内容を確認する装置と、
確認の結果、隣接する送信先ホストに転送する内容であれば、該フローエントリを削除し、経路削除パケットを破棄する装置と
を更に具備する
ネットワーク機器。

[0195] (付記4)

付記3に記載のネットワーク機器であって、
経路削除パケットと適合するフローエントリがない場合、経路削除パケットを一時的に保持する装置と
複数のコントローラのいずれかからフローエントリに設定が行われた際、
該フローエントリが経路削除パケットと適合すれば、該フローエントリの設定を破棄する装置と、
保持してから所定の時間の経過後に、経路削除パケットを破棄する装置と
を更に具備する
ネットワーク機器。

[0196] (付記5)

ネットワーク機器により実施される経路制御方法であって、
複数のコントローラの各々からの制御に基づいて、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動作が定義されたフローエントリをフローテーブルに設定することと、
複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他の

コントローラの設定した経路を不採用とし、採用された経路のフローエントリに従って、パケットを転送することと

を含む

経路制御方法。

[0197] (付記 6)

付記 5 に記載の経路制御方法であって、

複数のコントローラの各々から、経路上のネットワーク機器全ての設定を完了させた旨の設定完了通知を受信することと、

設定完了通知に応じて、不採用とした経路のフローエントリと適合する経路削除パケットを作成することと、

外部から経路削除パケットを受信した場合、経路削除パケットと適合するフローエントリの有無を確認することと、

不採用とした経路のフローエントリに従って、経路削除パケットを、不採用とした経路上の隣接する経路制御方法に転送することと、

不採用とした経路のフローエントリを削除することと

を更に含む

経路制御方法。

[0198] (付記 7)

付記 6 に記載の経路制御方法であって、

経路削除パケットと適合するフローエントリの内容を確認することと、

確認の結果、隣接する送信先ホストに転送する内容であれば、該フローエントリを削除し、経路削除パケットを破棄することと

を更に含む

経路制御方法。

[0199] (付記 8)

付記 7 に記載の経路制御方法であって、

経路削除パケットと適合するフローエントリがない場合、経路削除パケットを一時的に保持することと

複数のコントローラのいずれかからフローエントリに設定が行われた際、
該フローエントリが経路削除パケットと適合すれば、該フローエントリの設定を破棄することと、

保持してから所定の時間の経過後に、経路削除パケットを破棄することと
を更に含む

経路制御方法。

[0200] (付記 9)

複数のコントローラの各々からの制御に基づいて、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動作が定義されたフローエントリをフローテーブルに設定するステップと、

複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、採用された経路のフローエントリに従って、パケットを転送するステップと
をネットワーク機器に実行させるための

プログラム。

[0201] (付記 10)

付記 9 に記載のプログラムであって、

複数のコントローラの各々から、経路上のネットワーク機器全ての設定を完了させた旨の設定完了通知を受信するステップと、

設定完了通知に応じて、不採用とした経路のフローエントリと適合する経路削除パケットを作成するステップと、

外部から経路削除パケットを受信した場合、経路削除パケットと適合するフローエントリの有無を確認するステップと、

不採用とした経路のフローエントリに従って、経路削除パケットを、不採用とした経路上の隣接するプログラムに転送するステップと、

不採用とした経路のフローエントリを削除するステップと
を更にネットワーク機器に実行させるための

プログラム。

[0202] (付記 1 1)

付記 1 0 に記載のプログラムであって、
経路削除パケットと適合するフローエントリの内容を確認するステップと、

確認の結果、隣接する送信先ホストに転送する内容であれば、該フローエントリを削除し、経路削除パケットを破棄するステップと
を更にネットワーク機器に実行させるための
プログラム。

[0203] (付記 1 2)

付記 1 1 に記載のプログラムであって、
経路削除パケットと適合するフローエントリがない場合、経路削除パケットを一時的に保持するステップと

複数のコントローラのいずれかからフローエントリに設定が行われた際、
該フローエントリが経路削除パケットと適合すれば、該フローエントリの設定を破棄するステップと、

保持してから所定の時間の経過後に、経路削除パケットを破棄するステップと
を更にネットワーク機器に実行させるための
プログラム。

[0204] <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。

[0205] なお、本出願は、日本出願番号 2 0 1 1 - 0 4 8 1 4 3 に基づく優先権を主張するものであり、日本出願番号 2 0 1 1 - 0 4 8 1 4 3 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]

複数のネットワーク機器と、

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動作が定義されたフローエントリを、前記経路上のネットワーク機器のフローテーブルに設定する複数のコントローラと
を含み、

前記経路上のネットワーク機器は、前記複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、前記採用された経路のフローエントリに従って、パケットを転送する

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、

前記複数のコントローラの各々は、前記経路上のネットワーク機器全ての設定を完了させた際に、前記パケットの問い合わせを行ったネットワーク機器に対して、設定完了通知を送信し、

前記パケットの問い合わせを行ったネットワーク機器は、前記設定完了通知に応じて、前記不採用とした経路のフローエントリと適合する経路削除パケットを作成し、前記不採用とした経路のフローエントリに従って、前記経路削除パケットを、前記不採用とした経路上の隣接するネットワーク機器に転送し、前記不採用とした経路のフローエントリを削除し、

前記経路上のネットワーク機器は、前記経路削除パケットを受信すると、前記経路削除パケットと適合するフローエントリの有無を確認し、適合するフローエントリがある場合、該フローエントリに従って、前記経路削除パケットを、前記不採用とした経路上の隣接するネットワーク機器に転送し、該フローエントリを削除する

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、
前記経路上のネットワーク機器は、前記経路削除パケットと適合するフローエントリの内容を確認し、確認の結果、隣接する送信先ホストに転送する内容であれば、該フローエントリを削除し、前記経路削除パケットを破棄する
ネットワークシステム。
- [請求項4] 請求項3に記載のネットワークシステムであって、
前記経路上のネットワーク機器は、前記経路削除パケットと適合するフローエントリがない場合、前記経路削除パケットを一時的に保持し、前記複数のコントローラのいずれかからフローエントリに設定が行われた際、該フローエントリが前記経路削除パケットと適合すれば、該フローエントリを設定を破棄し、保持してから所定の時間の経過後に、前記経路削除パケットを破棄する
ネットワークシステム。
- [請求項5] 複数のコントローラの各々からの制御に基づいて、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動作が定義されたフローエントリをフローテーブルに設定する手段と、
前記複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、前記採用された経路のフローエントリに従って、パケットを転送する手段とを具備する
ネットワーク機器。
- [請求項6] 請求項5に記載のネットワーク機器であって、
前記複数のコントローラの各々から、前記経路上のネットワーク機器全ての設定を完了させた旨の設定完了通知を受信する手段と、
前記設定完了通知に応じて、前記不採用とした経路のフローエントリと適合する経路削除パケットを作成する手段と、

外部から前記経路削除パケットを受信した場合、前記経路削除パケットと適合するフローエントリの有無を確認する手段と、

前記不採用とした経路のフローエントリに従って、前記経路削除パケットを、前記不採用とした経路上の隣接するネットワーク機器に転送する手段と、

前記不採用とした経路のフローエントリを削除する手段と
を更に具備する

ネットワーク機器。

[請求項7]

請求項6に記載のネットワーク機器であって、

前記経路削除パケットと適合するフローエントリの内容を確認する手段と、

確認の結果、隣接する送信先ホストに転送する内容であれば、該フローエントリを削除し、前記経路削除パケットを破棄する手段と
を更に具備する

ネットワーク機器。

[請求項8]

請求項7に記載のネットワーク機器であって、

前記経路削除パケットと適合するフローエントリがない場合、前記経路削除パケットを一時的に保持する手段と

前記複数のコントローラのいずれかからフローエントリに設定が行われた際、該フローエントリが前記経路削除パケットと適合すれば、該フローエントリの設定を破棄する手段と、

保持してから所定の時間の経過後に、前記経路削除パケットを破棄する手段と

を更に具備する

ネットワーク機器。

[請求項9]

ネットワーク機器により実施される経路制御方法であって、

複数のコントローラの各々からの制御に基づいて、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動作が定義されたフローエン

トリをフローテーブルに設定することと、

前記複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、前記採用された経路のフローエントリに従って、パケットを転送することを含む

経路制御方法。

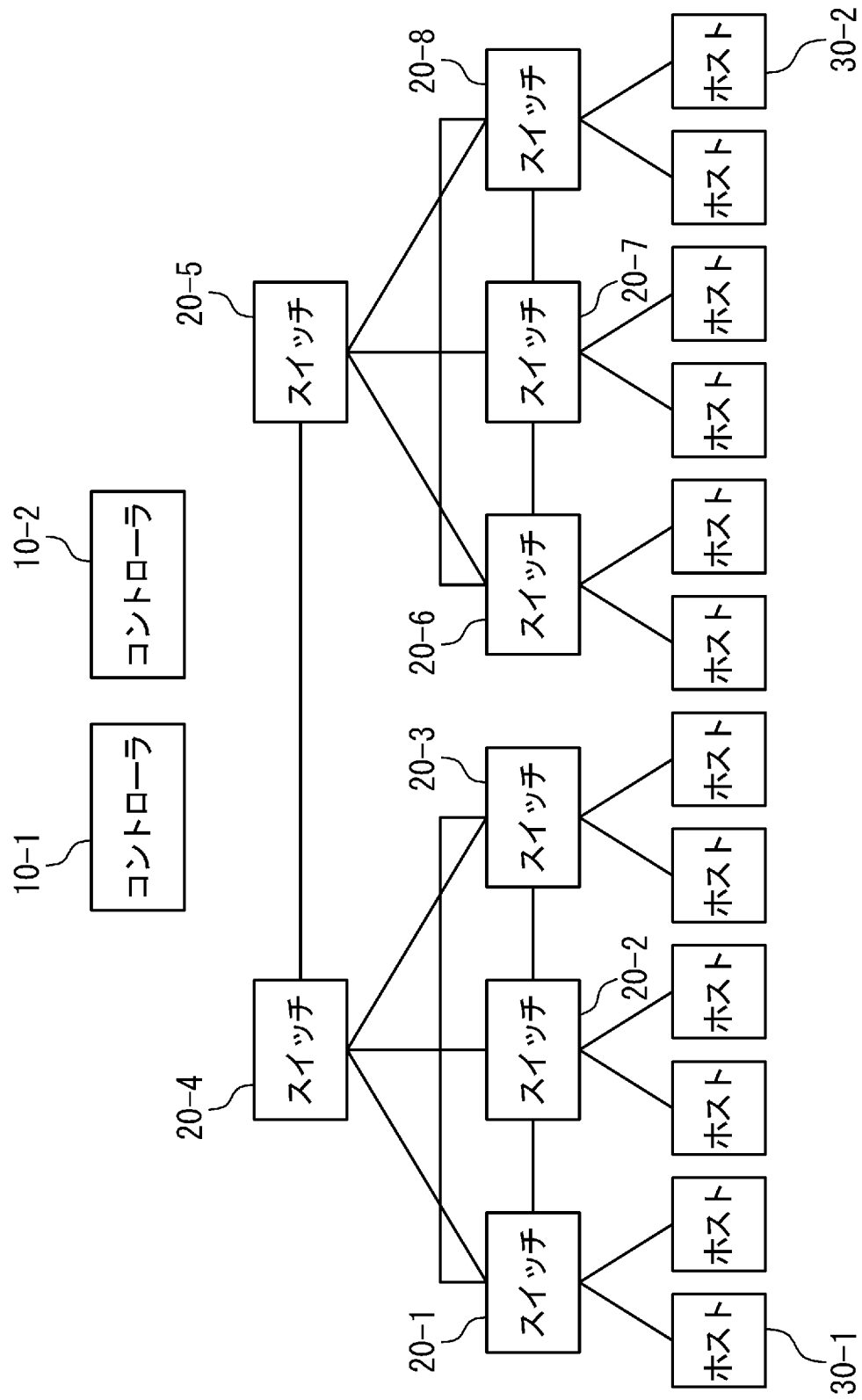
[請求項10]

複数のコントローラの各々からの制御に基づいて、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動作が定義されたフローエントリをフローテーブルに設定するステップと、

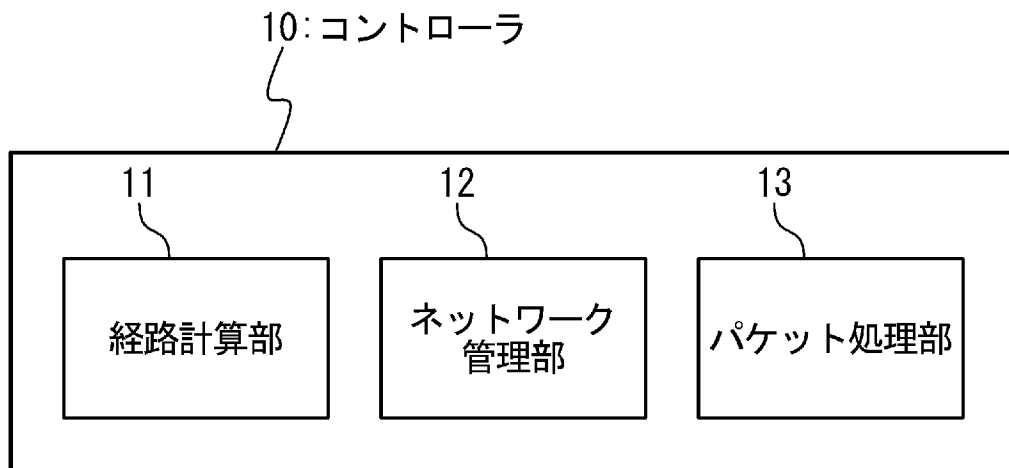
前記複数のコントローラの中で最も早く経路上のネットワーク機器全てのフローエントリの設定を完了させたコントローラの設定した経路を採用し、他のコントローラの設定した経路を不採用とし、前記採用された経路のフローエントリに従って、パケットを転送するステップと

をネットワーク機器に実行させるためのプログラムを格納した記憶媒体。

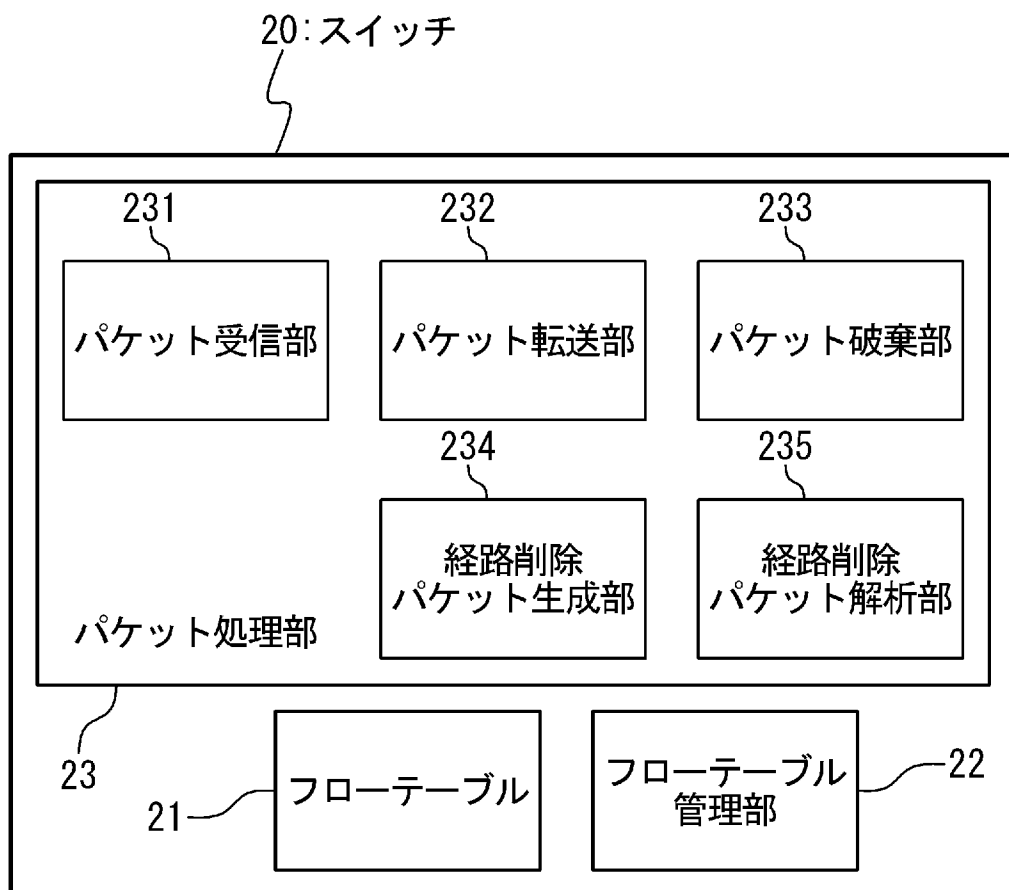
[図1]



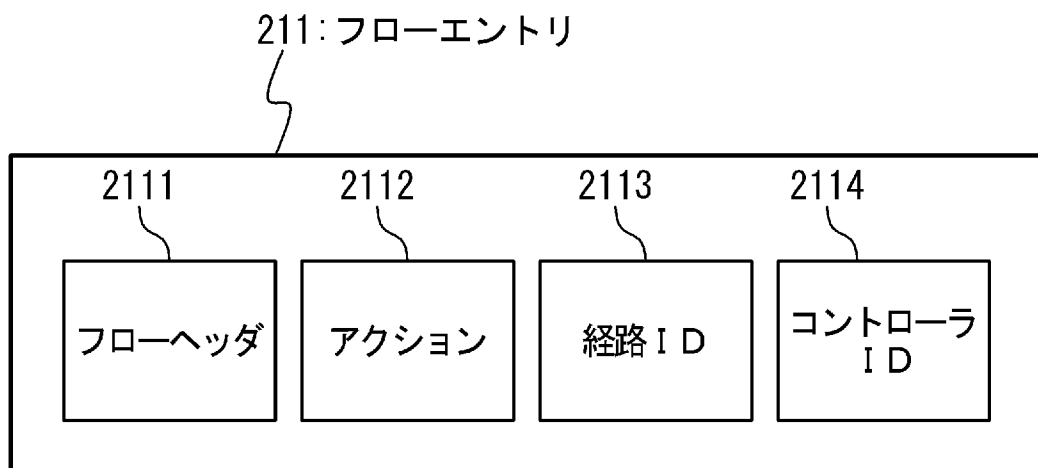
[図2]



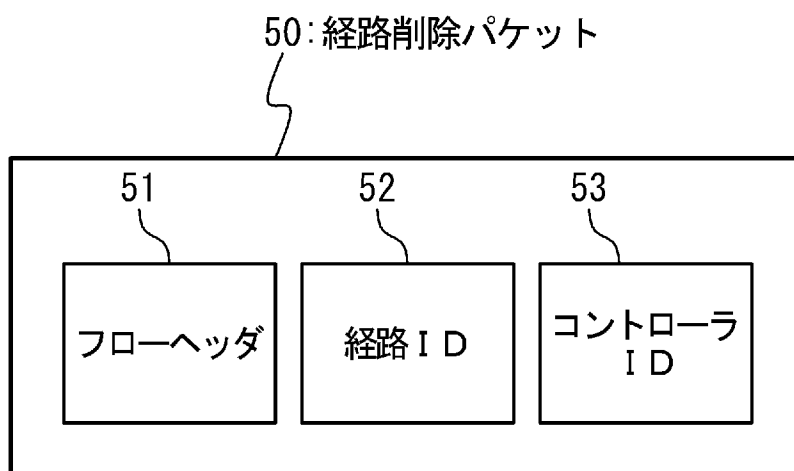
[図3]



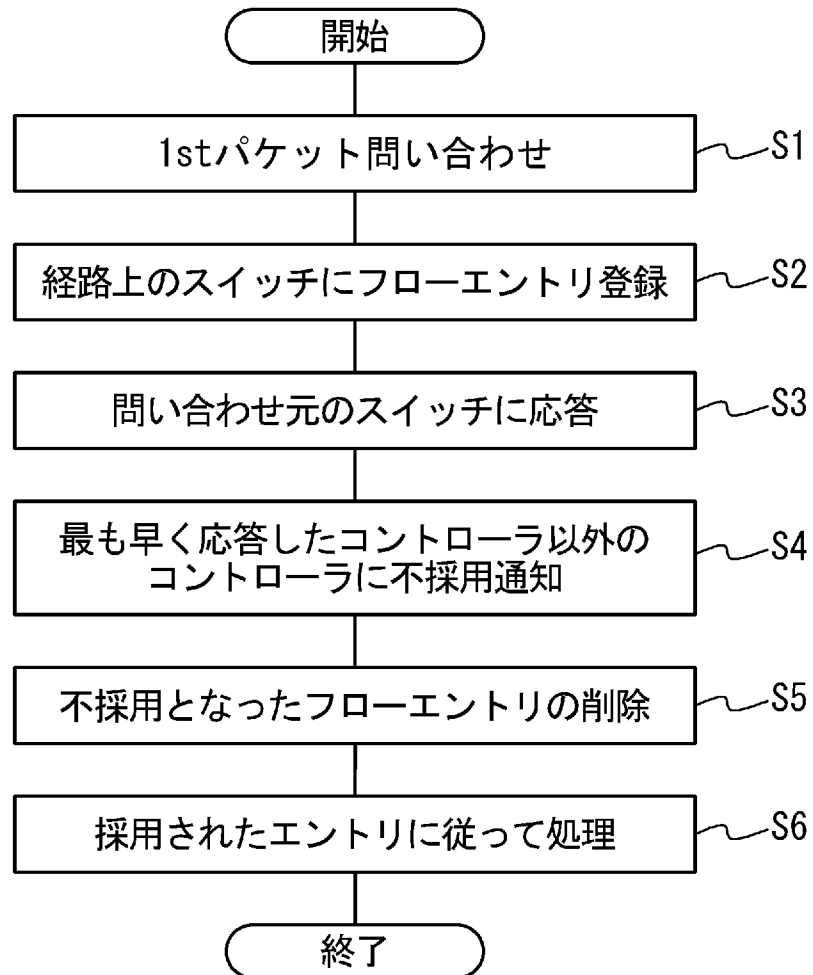
[図4]



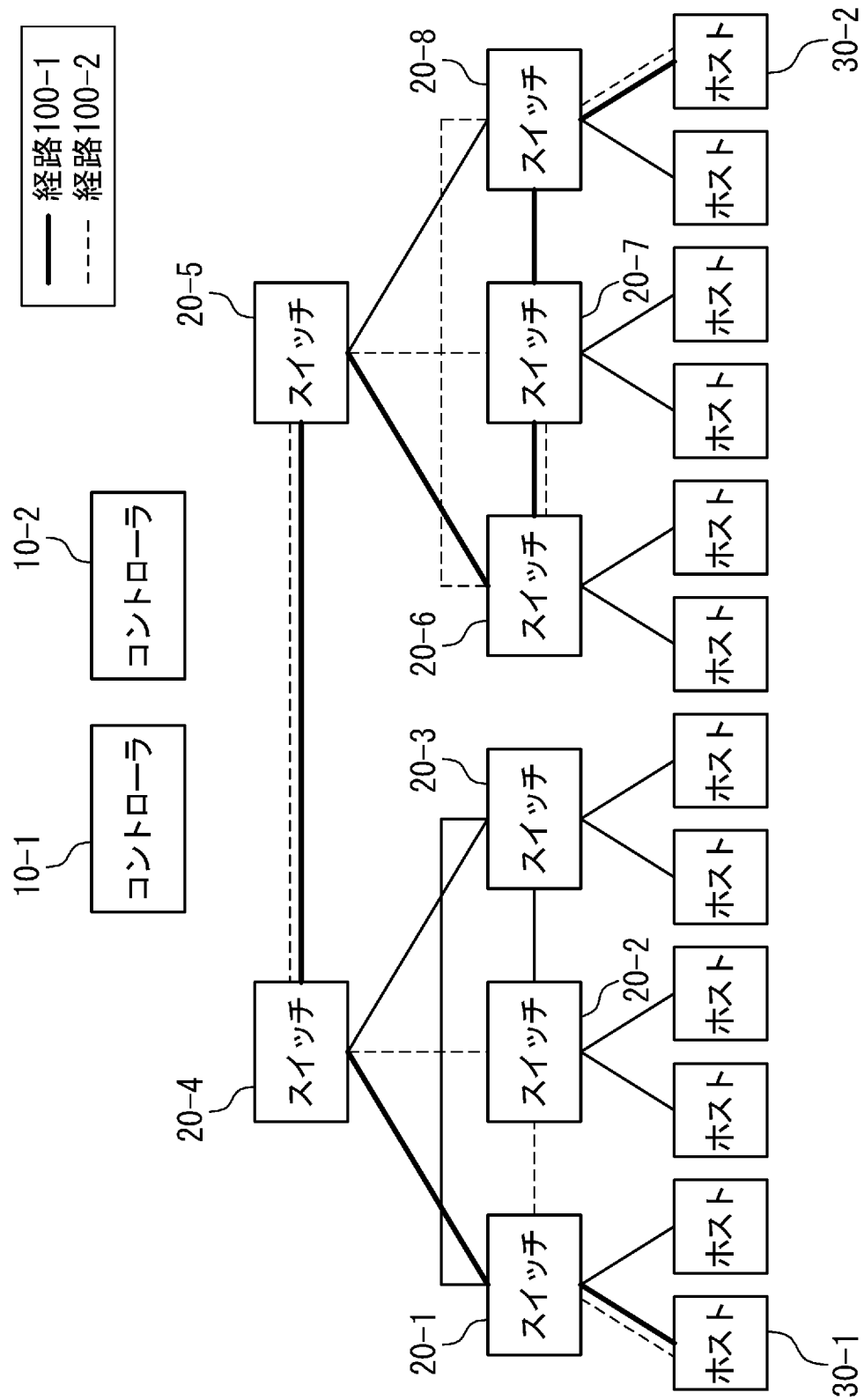
[図5]



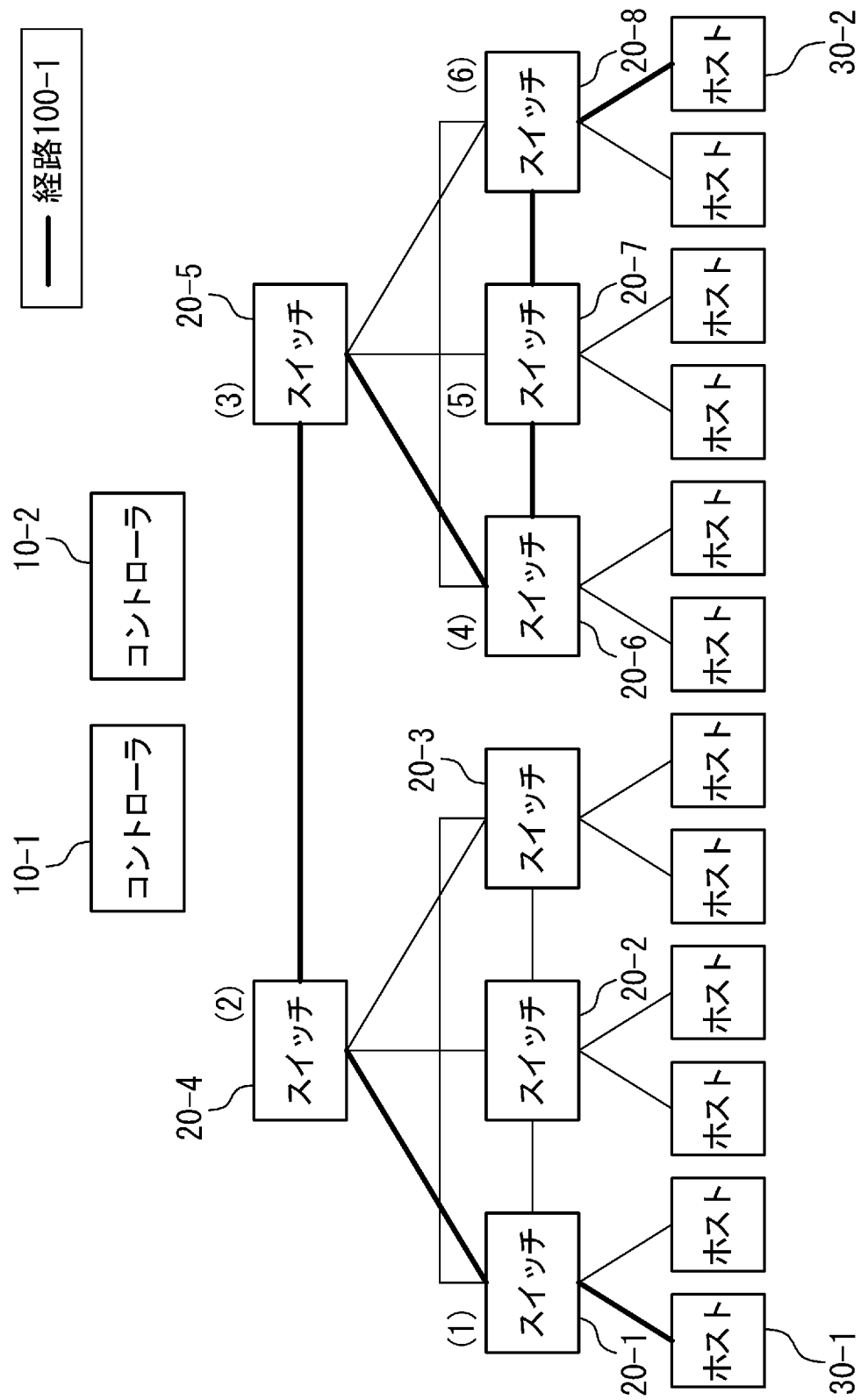
[図6]



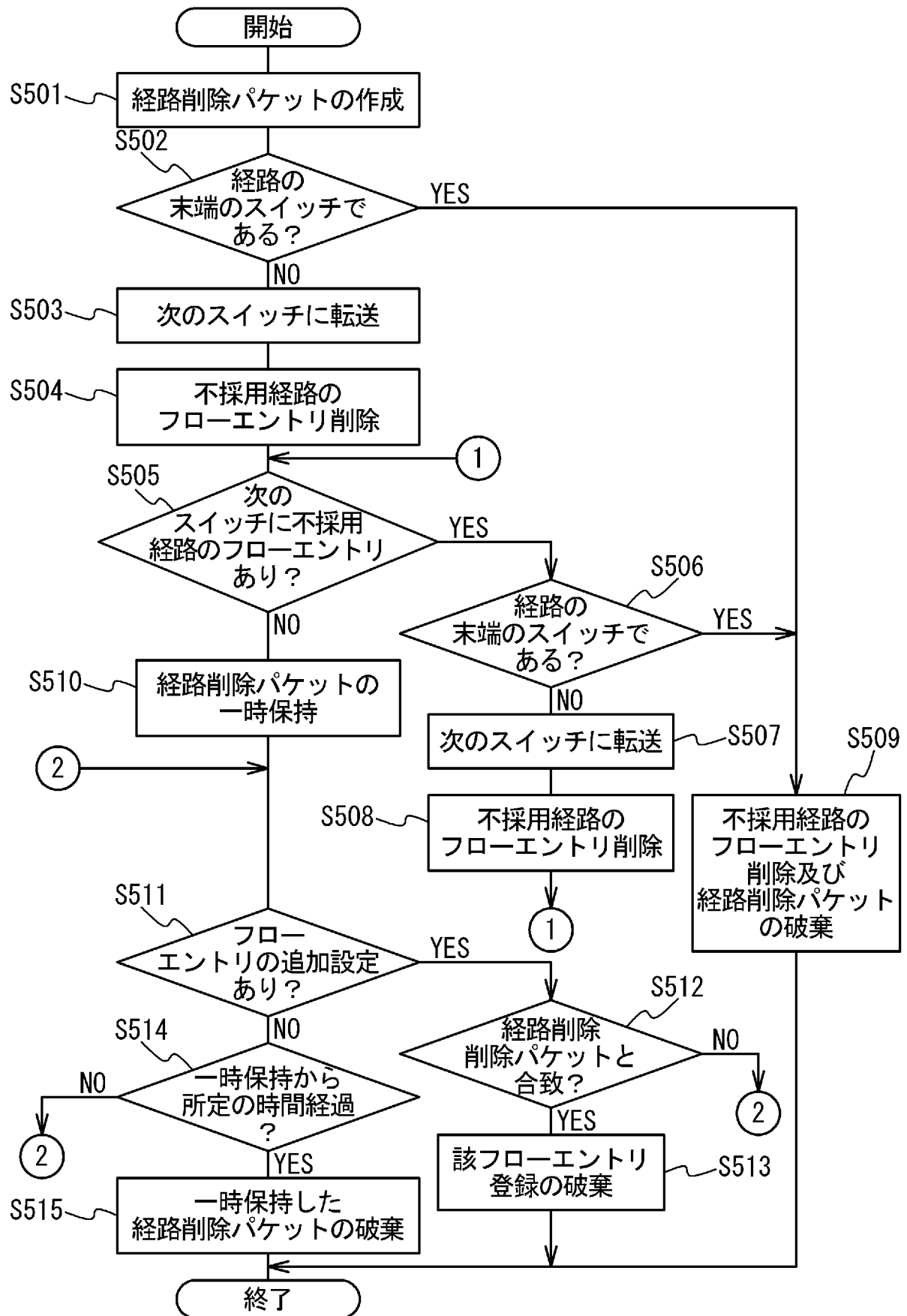
[図7]



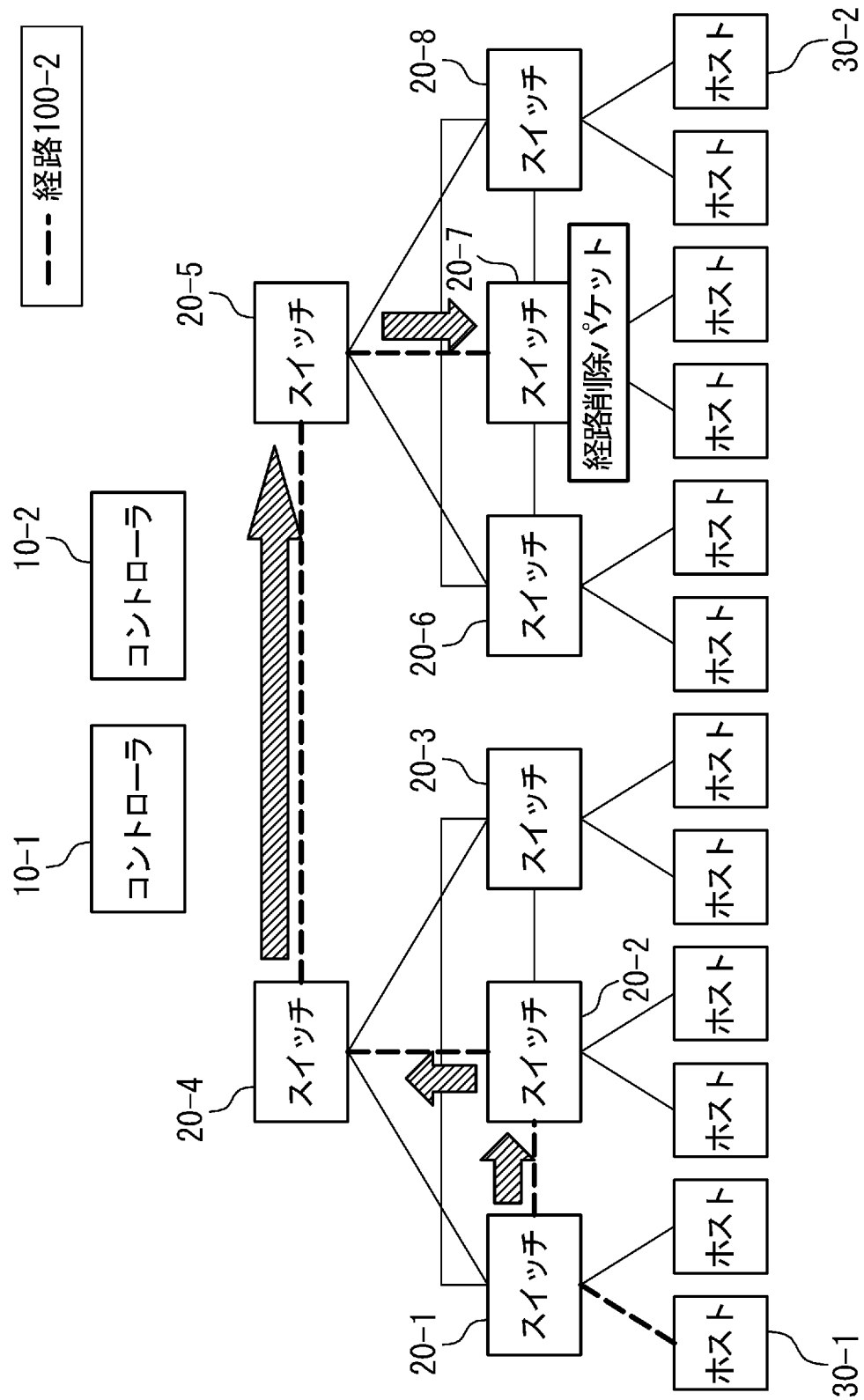
[図8]



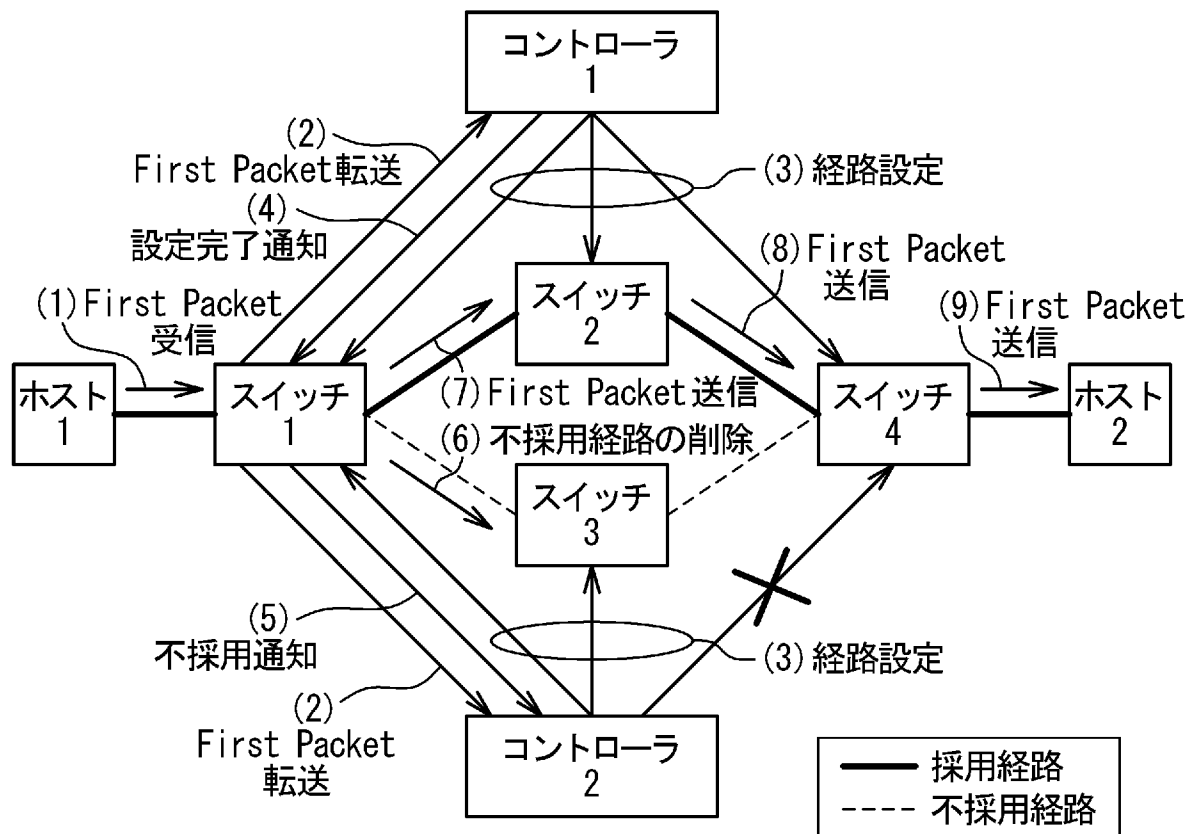
[図10]



[図11]



[図12]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/055272

A. CLASSIFICATION OF SUBJECT MATTER

H04L12/56(2006.01) 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)

H04L12/56

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	1996-2012
Kokai Jitsuyo Shinan Koho	1971-2012	Toroku Jitsuyo Shinan Koho	1994-2012

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.
A	OpenFlow Switch Specification, Version 1.1.0 Implemented, [online], The OpenFlow Switch Consortium, 2011.02.28, [retrieval date 23 March 2012 (23.03.2012)], Internet <URL: http://www.openflow.org/documents/openflow-spec-v1.1.0.pdf >	1-10
A	WO 2010/103909 A1 (NEC Corp.), 23 February 2010 (23.02.2010), paragraphs [0001] to [0026] (Family: none)	1-10



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
23 March, 2012 (23.03.12)

Date of mailing of the international search report
03 April, 2012 (03.04.12)

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（I P C））

Int.Cl. H04L12/56(2006.01)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（I P C））

Int.Cl. H04L12/56

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	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 2 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 2 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 2 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	OpenFlow Switch Specification, Version 1.1.0 Implemented, [online], The OpenFlow Switch Consortium, 2011.02.28, [平成 2 4 年 3 月 2 3 日 検 索], インターネット <URL: http://www.openflow.org/documents/openflow-spec-v1.1.0.pdf >	1-10
A	W0 2010/103909 A1（日本電気株式会社）2010.02.23, 段落 0 0 0 1 - 0 0 2 6（ファミリーなし）	1-10

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

「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

の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

2 3 . 0 3 . 2 0 1 2

国際調査報告の発送日

0 3 . 0 4 . 2 0 1 2

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号 1 0 0 - 8 9 1 5

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）

松崎 孝大

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 6

5 X

4 0 6 0