

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5854048号
(P5854048)

(45) 発行日 平成28年2月9日(2016.2.9)

(24) 登録日 平成27年12月18日(2015.12.18)

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

請求項の数 15 (全 20 頁)

<p>(21) 出願番号 特願2013-534082 (P2013-534082) (86) (22) 出願日 平成23年9月21日 (2011.9.21) (65) 公表番号 特表2014-504048 (P2014-504048A) (43) 公表日 平成26年2月13日 (2014.2.13) (86) 国際出願番号 PCT/JP2011/005303 (87) 国際公開番号 W02012/101689 (87) 国際公開日 平成24年8月2日 (2012.8.2) 審査請求日 平成26年8月6日 (2014.8.6) (31) 優先権主張番号 特願2011-16323 (P2011-16323) (32) 優先日 平成23年1月28日 (2011.1.28) (33) 優先権主張国 日本国 (JP)</p> <p>(出願人による申告) 平成22年度 総務省、「グリーンネットワーク基盤技術の研究開発」委託研究、産業技術力強化法第19条の適用を受ける特許出願</p>	<p>(73) 特許権者 000004237 日本電気株式会社 東京都港区芝五丁目7番1号 (74) 代理人 100080816 弁理士 加藤 朝道 (72) 発明者 千葉 靖伸 東京都港区芝五丁目7番1号 日本電気株式会社内 審査官 菊地 陽一</p>
---	--

最終頁に続く

(54) 【発明の名称】 通信システム、転送ノード、制御装置、通信制御方法およびプログラム

(57) 【特許請求の範囲】

【請求項1】

パケットの転送経路を設定し、該転送経路に対応する処理規則を決定する制御装置と、前記制御装置から送信された処理規則により受信パケットを処理する処理部と、受信パケットへの処理に関する情報を表す照合情報を、受信パケットに付加して転送する照合情報処理部と、を含む転送ノードと、

前記受信パケットに付加された照合情報が、予め計算された照合情報と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認する確認部と、を含む通信システム。

【請求項2】

前記照合情報処理部は、受信パケットに前記照合情報が含まれる場合、該照合情報と受信パケットへの処理に関する情報とに基づく照合情報を、該受信パケットに付加し、

前記受信パケットに付加された照合情報が、前記転送経路に対応する予め計算された照合情報と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認する請求項1の通信システム。

【請求項3】

前記制御装置は、複数の転送ノードのうちの所定の転送ノードに、前記予め計算した照合情報を通知し、

前記所定の転送ノードが、前記通知された照合情報を用いて、前記転送ノードと前記制御装置との同期を確認する請求項1または2の通信システム。

【請求項 4】

複数の転送ノードのうちの所定の転送ノードが、前記制御装置に前記照合情報を送信し

、
前記制御装置が、前記転送ノードと前記制御装置との同期を確認する請求項 1 または 2 の通信システム。

【請求項 5】

前記確認部を備える同期確認装置を含む請求項 1 または 2 の通信システム。

【請求項 6】

前記転送ノードは、前記照合情報を生成することが可能である請求項 1 から 5 いずれか一の通信システム。

10

【請求項 7】

前記照合情報は、前記受信パケットに対応する処理規則、前記転送ノードの識別情報、前記転送経路に対応するリンク情報の少なくとも 1 つに関連する情報を表す請求項 1 から 6 いずれか一の通信システム。

【請求項 8】

制御装置から送信された処理規則により受信パケットを処理する処理部と、
受信パケットへの処理に関する情報を表す照合情報を、受信パケットに付加して転送する照合情報処理部と、

前記受信パケットに付加された照合情報が、予め計算された照合情報と一致するか否かにより、前記制御装置との同期を確認する確認部と、

20

を含む転送ノード。

【請求項 9】

前記照合情報処理部は、受信パケットに前記照合情報が含まれる場合、該照合情報と受信パケットへの処理に関する情報とに基づく照合情報を、受信パケットに付加する請求項 8 の転送ノード。

【請求項 10】

前記照合情報を生成することが可能である請求項 8 または 9 の転送ノード。

【請求項 11】

前記照合情報は、前記受信パケットに対応する処理規則、前記転送ノードの識別情報、前記制御装置によって設定された前記受信パケットの転送経路に対応するリンク情報の少なくとも 1 つに対応する情報を表す請求項 8 から 10 いずれか一の転送ノード。

30

【請求項 12】

前記制御装置から、前記予め計算された照合情報を受信する請求項 8 から 11 のいずれか一の転送ノード。

【請求項 13】

パケットの転送経路を設定し、該転送経路に対応する処理規則を、該処理規則によりパケットを処理する転送ノードに送信する手段と、

前記処理規則により受信パケットを処理した転送ノードから、受信パケットへの処理に関する情報を表す照合情報を受信する手段と、

転送ノードから受信した前記照合情報が、予め計算された照合情報と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認する確認部と、

40

を含む制御装置。

【請求項 14】

制御装置から設定された処理規則に従ってパケット処理を行なう転送ノードが、受信パケットへの処理に関する情報を表す照合情報を生成し、前記受信パケットに付加するステップと、

前記受信パケットに付加された照合情報が、予め計算された照合情報と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認するステップと、を含む通信制御方法。

【請求項 15】

50

制御装置から設定された処理規則に従ってパケット処理を行なう転送ノードに搭載されたコンピュータに実行させるプログラムであって、

受信パケットへの処理に関する情報を表す照合情報を、受信パケットに付加する処理と

前記受信パケットに付加された照合情報が、予め計算された照合情報と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認する処理と、

を実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願についての記載)

本発明は、日本国特許出願：特願2011-016323号(2011年1月28日出願)の優先権主張に基づくものであり、同出願の全記載内容は引用をもって本書に組み込み記載されているものとする。

本発明は、通信システム、転送ノード、制御装置、通信制御方法およびプログラムに関し、特に、ネットワークに配置された転送ノードを集中制御する制御装置を有する通信システム、転送ノード、制御装置、通信制御方法およびプログラムに関する。

【背景技術】

【0002】

制御装置が転送ノードを集中制御することにより通信を実現する通信システムにおいては、転送ノードと制御装置とが同期していることが求められる。制御装置と転送ノードとが同期していないと、制御装置が転送ノードに送信したパケットの転送方法に関する指示と、転送ノードによるパケットの転送処理との間に齟齬が生じ、制御装置が意図しないパケット転送が行われてしまうからである。

【0003】

上記のように制御装置が転送ノードを集中制御する通信システムとしては、オープンフロー(OpenFlow)という技術が知られている(特許文献1、非特許文献1、2参照)。オープンフローは、通信をエンドツーエンドのフローとして捉え、フロー単位で経路制御、障害回復、負荷分散、最適化を行うものである。非特許文献2に仕様化されているオープンフロースイッチは、制御装置と位置付けられるオープンフローコントローラとの通信用のセキュアチャネルを備え、オープンフローコントローラから適宜追加または書き換え指示されるフローテーブルに従って動作する。フローテーブルには、フロー毎に、パケットヘッダと照合するマッチングルール(ヘッダフィールド)と、フロー統計情報(Counters)と、処理内容を定義したアクション(Actions)と、の組が定義される(図15参照)。

【0004】

例えば、オープンフロースイッチは、パケットを受信すると、フローテーブルから、受信パケットのヘッダ情報に適合するマッチングルール(図15のヘッダフィールド参照)を持つエントリを検索する。検索の結果、受信パケットに適合するエントリが見つかった場合、オープンフロースイッチは、フロー統計情報(カウンタ)を更新するとともに、受信パケットに対して、当該エントリのアクションフィールドに記述された処理内容(指定ポートからのパケット送信、フラグディング、廃棄(drop)等)を実施する。一方、前記検索の結果、受信パケットに適合するエントリが見つからなかった場合、オープンフロースイッチは、セキュアチャネルを介して、オープンフローコントローラに対して受信パケットを転送し、受信パケットの送信元・送信先に基づいたパケットの経路の決定を依頼し、これを実現するフローエントリを受け取ってフローテーブルを更新する。このように、オープンフロースイッチは、フローテーブルに格納されたエントリを処理規則として用いてパケット転送を行っている。

【0005】

しかしながら、上記特許文献1、非特許文献1、2においては、オープンフローコント

10

20

30

40

50

ローラとオープンフロースイッチとの間の同期の確認方法について、具体的な検討はなされていない。

【0006】

その他、特許文献2に、携帯型のデバイスが保持するデータと、該携帯型のデバイスが接続するデータベースに記憶されたデータとの同期を確認する技術が開示されている。同公報によれば、携帯型のデバイスは、携帯型のデバイスが保持するデータについてハッシュを生成し、該ハッシュを同期化サーバに送信し、該データの同期確認を要求する。同期化サーバは、同期確認を要求されたデータについて、ハッシュを生成する。同期化サーバは、携帯型のデバイスから送信されたハッシュと、自身が生成したハッシュとを比較する。同期化サーバは、比較結果に基づいて、データの同期を確認する。同期確認の対象となるデータからハッシュを生成し、該ハッシュに基づいて同期確認を実施することで、同期確認に要する送信時間や帯域幅を削減できると、記載されている。

10

【先行技術文献】

【特許文献】

【0007】

【特許文献1】国際公開第2008/095010号

【特許文献2】特表2009-510566号公報

【非特許文献】

【0008】

【非特許文献1】Nick McKeownほか7名、“OpenFlow: Enabling Innovation in Campus Networks”、[online]、[平成22年12月22日検索]、インターネット URL: <http://www.openflowswitch.org/documents/openflow-wp-latest.pdf>

20

【非特許文献2】“OpenFlow Switch Specification” Version 1.0.0. (Wire Protocol 0x01) [平成22年12月22日検索]、インターネット URL: <http://www.openflowswitch.org/documents/openflow-spec-v1.0.0.pdf>

【発明の概要】

30

【発明が解決しようとする課題】

【0009】

以下の分析は、本発明によって与えられたものである。上述のように、特許文献1、非特許文献1、2には、Stats Request/Replyというフローテーブル、もしくは、フローエントリの内容を応答する制御メッセージが用意されている程度であり（非特許文献2の「5.3.5 Read State Messages」参照）、転送ノードと制御装置間の効率の良い同期確認方法は記載されていない。

【0010】

また、特許文献2の技術は、個々のデータベースを代表するハッシュ（データベースハッシュ）の比較の結果、ミスマッチが検出された場合、携帯型のデバイスが保持する個々のデータ毎に同期確認（記録ハッシュの比較）を実行している。

40

【0011】

制御装置が転送ノードを集中制御することにより通信を実現する通信システムにおいては、非特許文献2のStats Request/Replyのような方法や、特許文献2のような方法で同期確認を実行すると、制御装置と転送ノードとの間に同期確認のためのネットワークトラフィックが発生し、制御装置および転送ノードに掛かる負荷がより増大してしまうという問題点がある。さらには、制御対象の転送ノードの数が増大するに従って、制御装置の負荷が爆発的に増大してしまう点も考慮する必要がある。

【0012】

本発明は、上記した事情に鑑みてなされたものであって、その目的とするところは制御

50

装置が転送ノードを集中制御することにより通信を実現する通信システムにおける、転送ノードと制御装置間の効率の良い同期確認方法を提供することにある。

【課題を解決するための手段】

【0013】

本発明の第1の視点によれば、パケットの転送経路を設定し、該転送経路に対応する処理規則を決定する制御装置と、前記制御装置から送信された処理規則により受信パケットを処理する処理部と、受信パケットへの処理に関する情報を表す照合情報を、受信パケットに付加して転送する照合情報処理部と、を含む転送ノードと、前記受信パケットに付加された照合情報が、予め計算された照合情報と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認する確認部と、を含む通信システムが提供される。

10

【0014】

本発明の第2の視点によれば、制御装置から送信された処理規則により受信パケットを処理する処理部と、受信パケットへの処理に関する情報を表す照合情報を、受信パケットに付加して転送する照合情報処理部と、前記受信パケットに付加された照合情報が、予め計算された照合情報と一致するか否かにより、前記制御装置との同期を確認する確認部と、を含む転送ノードが提供される。

【0015】

本発明の第3の視点によれば、上記した転送ノードに、前記照合情報を用いた処理規則の同期状態の確認動作を行わせる制御装置、または、上記した転送ノードから受信した前記照合情報を用いて転送ノードとの処理規則の同期状態の確認動作を行う制御装置が提供される。特に、前記制御装置は、パケットの転送経路を設定し、該転送経路に対応する処理規則を、該処理規則によりパケットを処理する転送ノードに送信する手段と、前記処理規則により受信パケットを処理した転送ノードから、受信パケットへの処理に関する情報を表す照合情報を受信する手段と、転送ノードから受信した前記照合情報が、予め計算された照合情報と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認する確認部と、を備えた構成とすることができる。

20

【0016】

本発明の第4の視点によれば、制御装置から設定された処理規則に従ってパケット処理を行なう転送ノードが、受信パケットへの処理に関する情報を表す照合情報を生成し、前記受信パケットに付加するステップと、前記受信パケットに付加された照合情報が、予め計算された照合情報と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認するステップと、を含む通信制御方法が提供される。本方法は、転送ノードおよび転送ノードを集中制御する制御装置という、特定の機械に結びつけられている。

30

【0017】

本発明の第5の視点によれば、制御装置から設定された処理規則に従ってパケット処理を行なう転送ノードに搭載されたコンピュータに実行させるプログラムであって、受信パケットへの処理に関する情報を表す照合情報を、受信パケットに付加する処理と、前記受信パケットに付加された照合情報が、予め計算された照合情報と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認する処理と、を実行させるプログラムが提供される。なお、このプログラムは、コンピュータが読み取り可能な記憶媒体に記録することができる。即ち、本発明は、コンピュータプログラム製品として具現することも可能である。また、このプログラムは、前記した転送ノード、制御装置、又はノ及び、通信システムまたは、これらの一部によって実現される技術的機能において物理的かつ電子工学的な役割を果たす。

40

【発明の効果】

【0018】

本発明によれば、制御装置が転送ノードを集中制御することにより通信を実現する通信システムにおける、転送ノードと制御装置間の同期確認を効率良く実施することが可能になる。

【図面の簡単な説明】

50

【 0 0 1 9 】

【図 1】本発明の概要を説明する図である。

【図 2】本発明の概要を説明する別の図である。

【図 3】本発明の第 1 の実施形態の構成を示す図である。

【図 4】本発明の第 1 の実施形態の制御装置および転送ノードの詳細構成を示すブロック図である。

【図 5】本発明の第 1 の実施形態の転送ノードの処理規則記憶部に保持されるテーブルの一例である。

【図 6】照合情報を含んだ（付加した）パケットの一例である。

【図 7】図 6 の照合情報格納ヘッダの構成例を示す図である。

10

【図 8】本発明の第 1 の実施形態の制御装置の処理規則記憶部に保持されるテーブルの一例である。

【図 9】本発明の第 1 の実施形態の制御装置の処理規則の設定要求受付時の動作を説明するための流れ図である。

【図 10】本発明の第 1 の実施形態の転送ノードのユーザパケット受信時の動作を説明するための流れ図である。

【図 11】本発明の第 1 の実施形態の制御装置における同期異常通知動作を説明するための流れ図である。

【図 12】本発明の第 1 の実施形態の動作シーケンスの一例を示す図である。

【図 13】図 10 のシーケンスに対応する動作を表した図である。

20

【図 14】照合情報がパケットヘッダに埋め込まれたパケットの一例である。

【図 15】非特許文献 2 に記載のフローエントリの構成を表した図である。

【発明を実施するための形態】

【 0 0 2 0 】

はじめに、本発明の概要について説明する。図 1 に示すように、本発明は、複数の転送ノードに、所定のパケット転送経路に従ったパケット処理を行わせる処理規則を設定する制御装置 20 と、前記制御装置から設定された処理規則に従って受信パケットを処理する転送ノード 10 と、を含む構成にて実現できる。なお、この概要に付記した図面参照符号は、理解を助けるための一例として各要素に便宜上付記したものであり、本発明を図示の態様に限定することを意図するものではない。

30

【 0 0 2 1 】

前記転送ノード 10（図 1 の転送ノード # 1、転送ノード # 2 参照）は、受信パケットを転送する際に、前記受信パケットに、受信パケットへの処理に関する情報を表す照合情報（図 1 の「C1」、「C2」参照）を付加する照合情報処理部 113 を備える。所定の転送ノード 10（例えば、図 1 の転送ノード # 2）の照合情報処理部 113 は、自身が付加した照合情報「C2」が、予め計算された内容「c2」と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認する。もちろん、所定の転送ノード 10（例えば、図 1 の転送ノード # 2）の照合情報処理部 113 が、受信したパケットに含まれる照合情報「C1」が予め計算された内容「c1」と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認するようにしてもよい。

40

【 0 0 2 2 】

転送ノード 10 が、転送ノード 10 と制御装置 20 との同期を確認するために用いる照合情報を、パケット中に含めて転送する。よって、転送ノード 10 と制御装置 20 との間に発生する同期確認のためのランザクションが削減される。その結果、実フローを利用して簡便に、転送ノード 10 と制御装置 20 間の処理規則の同期状態を確認することが可能となる。

【 0 0 2 3 】

上記照合情報を用いた同期確認は、図 1 のように転送ノード 10 側で行う形態のほか、制御装置 20 側で行うことも可能である。例えば、図 2 に示すように、所定の転送ノード 10（例えば、図 1 の転送ノード # 2）の照合情報処理部 113 が、自身が付加した照合

50

情報「C2」を制御装置20に送信し、制御装置20に設けられた同期確認部25Aが予め計算された内容「c2」と一致するか否かにより、前記転送ノードと前記制御装置との同期を確認する構成も採用可能である。

【0024】

ここで、上記した照合情報は、転送ノード10の識別子、MAC(Media Access Control)アドレス、転送ノード10のインタフェースに紐付けられているリンクIDなどの個々の転送ノードを区別できるような情報を用いることができる。また、これらの情報に代えて、個々のノードが当該パケットの処理に用いた処理規則、もしくは、その一部を用いることも可能である。また、ユーザパケットに照合情報を付加する形態としては、図1、図2に示すようにユーザパケットのヘッダ部分に埋め込む形態のほか、ユーザパケットに追加ヘッダとして付加する形態など(図6、図7、図14参照)、種々の形態を採用することができる。

10

【0025】

[第1の実施形態]

続いて、本発明の第1の実施形態について図面を参照して詳細に説明する。図3は、本発明の第1の実施形態に係る通信システムの構成を示す図である。図3を参照すると、3つの転送ノード10と、制御装置20と、転送ノード10を経由して通信するホスト(Host(A)、(B))が示されている。なお、図3の例では、3つの転送ノード10と、制御装置20と、2つのホスト(Host(A)、Host(B))を示しているが、それぞれの数は、あくまで例示であり、それぞれ任意の数とすることができる。

20

【0026】

図4は、転送ノード10および制御装置20の詳細構成を表したブロック図である。図4を参照すると、転送ノード10は、転送処理部11と、制御装置20と通信する制御装置通信部12と、処理規則記憶部14を管理する処理規則管理部13と、パケットバッファ15と、を備えて構成される。なお、転送ノード10は、必ずしもパケットバッファ15を備えていなくともよい。

【0027】

図5は、処理規則記憶部14に保持される処理規則格納用テーブルの一例である。Priorityフィールドは、各処理規則の優先度を表わしている。また、マッチングルールおよびアクション(Actions)は、図15のオープンフロースイッチのフローエントリのマッチングルール(照合規則)およびアクション(Actions)に相当する。照合情報フィールドは、各処理規則中の特定のフィールドの値(例えば、マッチングルールおよびアクション(Actions))を引数として所定の関数(Message Digest Algorithm 5(MD5)などに代表されるハッシュ関数等)を用いて、当該転送ノードが持つ処理規則を象徴する値(以下、「照合情報」という。)が格納される。図5の例では、照合情報フィールドに値が入っているが、照合情報フィールドは空値であってもよい。制御装置20から設定された処理規則の照合情報フィールドに値が入っていない場合、後記する照合情報を用いた同期確認は省略される。なお、図5に示したテーブルの形式は例示である。本発明における処理規則格納用テーブルは、オープンフローの技術に基づくものに限定されるものではない。

30

40

【0028】

なお、図5の例では、フロー統計情報(Stats)フィールドを省略しているが、フロー統計情報(Stats)を追加することも可能である(但し、フロー統計情報(Stats)は、パケット受信の都度変化するものであるため、照合情報の作成には使用しない)。従って、図5の処理規則格納用テーブルは、非特許文献2のオープンフロースイッチのフローエントリに、Priorityフィールドと照合情報フィールドとを追加したものとなっている。

【0029】

転送処理部11は、処理規則検索部111と、アクション実行部112とを含んで構成される。処理規則検索部111は、上記のような構成の処理規則記憶部14から、受信パ

50

ケットに適合するマッチングルール（照合規則）を持つ処理規則の処理内容（アクション）をアクション実行部 112 に出力する。アクション実行部 112 は、前記処理規則検索部 111 から出力された処理内容（アクション）を実行する。処理規則検索部 111 は、さらに、照合情報処理部 113 を備えている。

【0030】

照合情報処理部 113 は、受信したユーザパケットに照合情報が含まれる場合、この照合情報と、前記自装置を処理したことを示す情報とを用いて、受信パケットに新しい照合情報を付加する。

【0031】

照合情報処理部 113 が付加する照合情報は、例えば、次式 [数 1]、[数 2] により算出することができる。

【0032】

【数 1】

$$C_n = f(C_{n-1}, FlowEntry_n)$$

【数 2】

$$C_n = \{h_n, h_{n-1}, h_{n-2}, h_{n-3}, \dots, h_0\}$$

$$h_n = f(FlowEntry_n)$$

ここで、 C_n は、パケット転送経路上の n 番目の転送ノード # n が生成する照合情報であり、 C_{n-1} は、受信パケットに含まれている $n-1$ 番目の転送ノード # $n-1$ が生成した照合情報である。また、 $FlowEntry_n$ は、パケット転送経路上の n 番目の転送ノード # n が、当該ユーザパケットに対応するとして処理規則記憶部 14 から検索した処理規則中の特定のフィールドの値（例えば、マッチングルールおよびアクション（Actions））を示す。また、関数 $f(x)$ としては、MD5 などのハッシュ関数を用いることができる。このようなハッシュ関数を用いることで、照合情報 C_n は、ユーザパケットが経験した処理規則（個々の転送ノードにおいて適用されてきた処理規則）の組み合わせを一意に表わした値となる。また、上記マッチングルールおよびアクション（Actions）に代えて、非特許文献 2 に記載されている Flow Cookie 等を用いることもできる。

【0033】

また、照合情報処理部 113 は、処理規則検索部 111 により検索された処理規則の照合規則フィールドに、図 5 に示すように値が入っている場合、上記のようにして算出された照合情報 C_n と、制御装置 20 から設定された処理規則中の照合規則フィールドとを比較する。前記比較の結果、不一致が検出された場合、照合情報処理部 113 は、制御装置 20 に対して、同期異常（同期外れ）を報告する。

【0034】

図 6 は、転送ノード 10 にて作成される、照合情報を付加したパケットの構成例を示す図である。図 6 の例では、照合情報を付加したパケット 32 は、照合情報を含んだ照合情報格納ヘッダ 33 をユーザパケット 31 の先頭に付加した構成となっている。

【0035】

図 7 は、上記照合情報格納ヘッダ 33 の構成例を示す図である。図 7 の例では、照合情報格納ヘッダ 33 は、MAC 宛先アドレス（MAC DA）、MAC 送信元アドレス（MAC SA）、上位プロトコルタイプ（Ether Type）、トータルヘッダ長（Total Length）に、照合情報 C_n を付加した構成となっている。

【0036】

上記のような転送ノード 10 は、非特許文献 2 のオープンフロースイッチのフローテーブルに上記照合情報フィールドを追加し、さらに、照合情報処理部 113 を追加した構成にて実現することが可能である。また、上記のような転送ノード 10 の照合情報処理部 1

10

20

30

40

50

13を含む転送処理部11に相当する機能は、転送ノード10を構成するコンピュータに実行させるプログラムによって実現することもできる。

【0037】

再度、図4を参照して、制御装置20の構成を説明する。制御装置20は、転送ノード10との通信を行うノード通信部21と、制御メッセージ処理部22と、ノード通信部21を介して収集された転送ノード10の接続関係に基づいてネットワークトポロジ情報を構築するトポロジ管理部23と、トポロジ管理部23にて構築されたネットワークトポロジ情報に基づいてパケットの転送経路および該転送経路上の転送ノード10に実行させるアクションを求める経路・アクション計算部24と、転送ノード10への処理規則の設定や、転送ノード10からの処理規則のタイムアウト通知に応じて、処理規則記憶部28の内容が各転送ノード10の処理規則記憶部14の内容と一致するよう管理する処理規則管理部25と、照合情報生成部27と、処理規則記憶部28とを備えて構成される。

10

【0038】

さらに、制御メッセージ処理部22は、転送ノード10から受信した制御メッセージを解析して、必要な処理を行うメッセージ解析・処理部221と、転送ノード10に送信するメッセージを生成するメッセージ生成部222とを備えて構成される。

【0039】

照合情報生成部27は、上記処理規則管理部25からの要求に基づいて、同期確認を行わせる転送ノードに設定する処理規則の照合情報フィールドに設定する照合情報を生成する手段である。照合情報の具体的な算出方法は、上記した照合情報処理部113が採用する算出方法と同一のものを用いることとすればよい。

20

【0040】

図8は、制御装置20の処理規則記憶部28に保持される処理規則格納用テーブルの一例である。図8に示すとおり、制御装置20の処理規則記憶部28には、制御装置20の制御対象の転送ノード10の処理規則記憶部14に保持される処理規則がまとめて保持される。例えば、図8の上から1～3番目のエントリは、マッチングルールがAであるフローについて、DPID(ノード10の識別子; Datapath Identifier)が1～3のノードにそれぞれ設定される処理規則を示している(上から4～6番目のエントリも同様)。なお、図8の例では、マッチングルールをキーにして処理規則が格納されているが、DPID等をキーにして処理規則を格納する構成とすることも可能である。

30

【0041】

DPIDが1であるノード(例えば、図3のノード#1)は、上記上から1番目のエントリのとおり、「A」というマッチングルールに適合するパケットを受信した場合に、そのアクションフィールド(Actions)に従い、9番ポートから出力する処理を実行する。また、DPIDが2であるノード(例えば、図3のノード#2)は、DPIDが1であるノード(例えば、図3のノード#1)から「A」というマッチングルールに適合するパケットを受信すると、6番ポートから出力する処理を実行する。また、DPIDが3であるノード(例えば、図3のノード#3)は、DPIDが2であるノード(例えば、図3のノード#2)から「A」というマッチングルールに適合するパケットを受信すると、1番ポートから出力する処理を実行する。以上により、「A」というマッチングルールに適合するパケットを指定した経路にて転送させる制御が実現される。

40

【0042】

また、図8の例では、DPIDが1であるノード(例えば、図3の転送ノード#1)は、照合フィールドが空値であるので、照合情報のパケットへの付加を行う。DPIDが2、3であるノード(例えば、図3の転送ノード#2、#3)は、照合フィールドに値が入っているので、照合情報のパケットへの付加に加えて、前記生成した照合情報が処理規則中の照合フィールドと一致しているか否かの確認を行う。以上のように、各転送ノード10によって、パケット転送処理とともに照合情報の再付加が行われ、さらに、照合フィールドに値が設定されている場合には、当該転送ノードにより、前記付加した照合情報によ

50

る同期確認処理が行われる。

【 0 0 4 3 】

なお、図 4 の構成では、処理規則記憶部 2 8 を制御装置 2 0 の内部に設けているが、別途外部サーバ等に処理規則記憶部 2 8 を設ける構成も採用可能である。

【 0 0 4 4 】

また、上記のような制御装置 2 0 は、非特許文献 1、2 のオープンフローコントローラをベースに、上記のような照合情報の生成機能を追加することで実現することが可能である。但し、制御装置 2 0 は、オープンフローに基づくことに限定されるものではない。また、上記のような制御装置 2 0 の処理規則管理部 2 5、照合情報生成部 2 7 に相当する機能は、制御装置 2 0 を構成するコンピュータに実行させるプログラムによって実現することもできる。

10

【 0 0 4 5 】

続いて、上記した制御装置 2 0 および転送ノード 1 0 の動作について説明する。図 9 は、上記制御装置 2 0 の処理規則の設定要求受付時の動作を説明するための流れ図である。図 9 を参照すると、制御装置 2 0 は、転送ノード 1 0 から処理規則の設定要求；図 1 0 のステップ S 1 1 0 参照）を受けると（ステップ S 0 0 1）、トポロジ管理部 2 3 にて構築されたネットワークトポロジ情報を取得し、パケットの転送経路を計算する（ステップ S 0 0 2）。

【 0 0 4 6 】

前記パケットの転送経路の計算の結果、経路を作成できない、経路上のノードが故障している等の理由により、転送できないと判定した場合（ステップ S 0 0 3 の No）を除き、制御装置 2 0 は、前記計算された転送経路に対応するアクションを計算する（ステップ S 0 0 4）。続いて、制御装置 2 0 は、当該処理規則の設定に係るパケットが属するフローを特定するためのマッチングルールを作成し、当該転送経路上の各転送ノード 1 0 に適用する処理規則（照合情報なし）を生成する（ステップ S 0 0 5）。さらに、制御装置 2 0 は、当該経路上の各転送ノード 1 0 のうち、同期確認処理を行わせる転送ノードの処理規則の照合情報フィールドに設定する照合情報を生成する（ステップ S 0 0 6）。

20

【 0 0 4 7 】

処理規則および照合情報の生成が完了すると、制御装置 2 0 は、処理規則の設定メッセージを生成し（ステップ S 0 0 7）、当該転送経路上の転送ノード 1 0 に対して、処理規則設定メッセージを送信する（ステップ S 0 0 8）。

30

【 0 0 4 8 】

その後、転送ノード 1 0 がパケットをバッファしていない場合（ステップ S 0 0 9 の No）、制御装置 2 0 は、パケットの送信指示（Packet - Out）を行う（ステップ S 0 1 0）。このパケットの送信指示は、送信すべきパケット（ステップ S 0 0 1 で受信したパケット）と、当該パケットに対し実行すべきアクション（照合情報の付与と指定ポートから出力）とを指示すること、あるいは、送信すべきパケット（ステップ S 0 0 1 で受信したパケット）と、当該パケットに対し実行すべきアクション（処理規則記憶部 1 4 の検索）を指示することによって行われる。なお、転送ノード 1 0 がパケットをバッファしている場合（ステップ S 0 0 9 の Yes）、後に説明するように、転送ノード 1 0 は自身でパケットを送信できるため、制御装置 2 0 側の処理は省略される。

40

【 0 0 4 9 】

図 1 0 は、転送ノード 1 0 の動作を表したフローチャートである。図 1 0 を参照すると、転送ノード 1 0 は、ホストや他の転送ノード 1 0 からパケットを受信すると（ステップ S 1 0 1）、処理規則記憶部 1 4 から、受信パケットに適合するマッチングルールを持つ処理規則を検索する（ステップ S 1 0 2）。

【 0 0 5 0 】

ステップ S 1 0 2 にて、処理規則を抽出できた場合（ステップ S 1 0 3 の Yes）、転送ノード 1 0 は、照合情報 C_n を作成する（ステップ S 1 0 4）。なお、前記受信パケットに照合情報 C_{n-1} が含まれている場合には、前記抽出した処理規則と、その照合情報

50

C_{n-1} とを用いて、新しい照合情報 C_n を作成する。

【0051】

次に、転送ノード10は、前記生成した照合情報 C_n を用いて同期確認処理を行うか否かを確認する(ステップS105)。前記同期確認処理を行うか否かは、上述のとおり、ステップS102で抽出された処理規則の照合情報フィールドに照合情報が含まれているか否かで判断することができる。

【0052】

ここで、ステップS102で抽出された処理規則の照合情報フィールドに照合情報が含まれていない場合(ステップS105のNo)、転送ノード10は、同期確認処理を省略し、前記抽出した処理規則に記述されたアクションを実行する(ステップS108)。

10

【0053】

一方、ステップS102で抽出された処理規則の照合情報フィールドに照合情報が含まれている場合(ステップS105のYes)、転送ノード10は、ステップS104で作成した照合情報 C_n とステップS102で抽出された処理規則の照合情報フィールドに格納された照合情報とを比較することにより、同期確認を行う(ステップS106)。ステップS106にて、照合情報が一致していることが確認できた場合(ステップS106のYes)、前記抽出した処理規則に記述されたアクションを実行する(ステップS108)。

【0054】

一方、ステップS106にて、照合情報の不一致が検出された場合(ステップS106のNo)、転送ノード10は、制御装置20に対して、同期異常が発生したことを通知する(ステップS107)。

20

【0055】

一方、ステップS103で受信パケットに対応する処理規則が見つからなかった場合(ステップS103のNo)、転送ノード10は、制御装置20に対し、受信パケットに対応する処理規則の作成および送信要求を行う(ステップS110)。

【0056】

ステップS110では、転送ノード10は、受信パケットをパケットバッファ15に保存するとともに、制御装置20に、受信パケットを送信し、処理規則の作成を要求する。その後は、制御装置20にて、図9に示す手順に従って、照合情報を含んだ処理規則の作成処理が行われる。

30

【0057】

制御装置20から、処理規則の設定メッセージを受信すると、転送ノード10は、自身の処理規則記憶部14に、処理規則の設定メッセージに従い処理規則を格納する(ステップS111)。

【0058】

次に、転送ノード10は、パケットバッファ15に受信パケットが保存されているか否かを確認し(ステップS112)、保存されていれば(ステップS112のYes)、当該受信パケットを読み出し(ステップS113)、前記設定した処理規則(フローエントリ)に定められた処理内容(アクション;照合情報の付加と受信パケットの指定ポートからの出力)を実行する(ステップS108)。これにより、受信パケットが次ホップの転送ノードに転送される。

40

【0059】

一方、転送ノード10がパケットバッファ15を備えていない場合など、転送ノード10が受信パケットを保存していない場合(ステップS112のNo)、制御装置20からのパケット送信指示(Packet-Out)が受信される(ステップS114)。

【0060】

前記パケット送信指示(Packet-Out)を受信した転送ノード10は、パケットバッファ15を備えている場合にはパケットバッファ15にパケットが保存されているか否かを確認し(ステップS115)、保存されていれば(ステップS115のYes)

50

、当該パケットを読み出し（ステップS116）、パケット送信指示（Packet - Out）とともに受信した処理内容（アクション；ここでは、照合情報の付加と指定ポートからの出力、または、処理規則記憶部14の検索）を実行する（ステップS108）。また、パケットを保存していない場合（ステップS115のNo）、転送ノード10は、パケット送信指示（Packet - Out）とともに受信したパケットについて、パケット送信指示（Packet - Out）とともに受信した処理内容（アクション；ここでは、照合情報の付与と指定ポートからの出力、または、処理規則記憶部14の検索）を実行する（ステップS108）。これにより、受信パケットが次ホップの転送ノードに転送される。

【0061】

図11は、上記ステップS107において、転送ノード10が制御装置20に同期異常を通知した場合の制御装置20の動作を表したフローチャートである。図11に示したとおり、制御装置20は、転送ノード10から同期異常を受信すると（ステップS201）、その内容を同期異常ログとして記録するとともに、所定のネットワーク管理者の端末等に通知する動作を行う（ステップS202）。

【0062】

続いて、上記転送ノード10および制御装置20の一連動作を図12のシーケンス図および図13の参考図を用いてより具体的に説明する。以下の説明では、制御装置20により、転送ノード#1、#2の処理規則記憶部14には、照合情報の無い処理規則が設定され、転送ノード#3の処理規則記憶部14に照合情報を含む処理規則が設定されているものとして説明する。

【0063】

図12、図13に示すように、Host(A)が、転送ノード#1に対し、Host(B)宛てのパケットを送信すると（図12、図13のST1）、転送ノード#1は、処理規則記憶部14から、当該パケットに対応する処理規則を抽出し（図12、図13のST2）、その内容を用いて照合情報C1を作成し、作成した照合情報C1を受信パケットに付加して転送ノード#2に送信する（図12、図13のST3）。

【0064】

転送ノード#1からパケットを受信した転送ノード#2は、処理規則記憶部14から、受信パケットに対応する処理規則を抽出し（図12、図13のST4）、その内容と受信パケットに含まれる照合情報C1を用いて照合情報C2を作成し、作成した照合情報C2を受信パケットに付加して転送ノード#3に送信する（図12、図13のST5）。

【0065】

転送ノード#2からパケットを受信した転送ノード#3は、同様に、処理規則記憶部14から、受信パケットに対応する処理規則を抽出し（図12、図13のST6）、その内容と受信パケットに含まれる照合情報C2を用いて照合情報C3を作成する。さらに、転送ノード#3は、前記抽出した処理規則に照合情報C3が含まれているため、作成した照合情報C3と前記抽出した処理規則に含まれている照合情報C3とが一致するか否かを確認する。

【0066】

本実施形態では、当該パケットが通過してきた転送ノードにおいて適用された処理規則群から、照合情報C3を生成しているため、転送ノード#1～#3に設定された処理規則が制御装置20が保持している処理規則と同期していない場合、転送ノード#3が作成した照合情報C3は、処理規則に含まれている照合情報C3と不一致となる。この場合、転送ノード#3は、上記した図10のステップS107で説明したように、制御装置20に対し、同期異常を通知する（図12、図13のST7-1）。

【0067】

一方、転送ノード#3が作成した照合情報C3が、処理規則に含まれている照合情報C3と一致する場合、転送ノード#1～#3に設定された処理規則が制御装置20が保持している処理規則と同期していると判断される。この場合、転送ノード#3は、転送ノード

10

20

30

40

50

2 から受信したパケットから照合情報 C 2 を削除した上で、ホスト (B) に転送する (図 1 2、図 1 3 の S T 7 - 2)。

【 0 0 6 8 】

以上のように、本実施形態によれば、転送ノード 1 0 と制御装置 2 0 とが直接処理規則の送受信を行わなくとも、転送ノード 1 0 と制御装置 2 0 間の処理規則の同期状態を確認することが可能となる。

【 0 0 6 9 】

さらに、本実施形態では、上記 [数 1]、[数 2] に示したように、処理規則を用いて照合情報を生成するようにしているため、例えば、経路途中の転送ノードで本来意図しない処理規則に適合してパケットの転送が行われた場合にも、その誤りを検出することが可能となっている。

【 0 0 7 0 】

[第 2 の実施形態]

続いて、本発明の第 2 の実施形態について図面を参照して詳細に説明する。上記本発明の第 1 の実施形態では、照合情報格納ヘッダによりパケットに照合情報を付加するものとして説明したが、追加ヘッダを用いずに照合情報をパケットに含めることもできる。以下、追加ヘッダを用いずに照合情報をパケットに含めるようにした第 2 の実施形態について説明する。なお、第 2 の実施形態の転送ノード 1 0 および制御装置 2 0 の基本的な構成は、上記第 1 の実施形態と同等であるので、以下、相違点を中心に説明を行う。

【 0 0 7 1 】

図 1 4 は、本実施形態の転送ノード 1 0 が生成するパケットを説明するための図である。符号 3 1 は、H o s t (A)、(B) 等から送信されたユーザパケットを表わしている。符号 3 2 a は、照合情報が埋め込まれたパケットの構成を示している。図 1 4 の例では、M A C D A フィールドに、図 1 4 の下段に示す 4 8 ビットの情報を書き込んだ構成となっている。なお、図 1 4 の例では、照合情報を含むことを示す 8 ビットの固定データ、関数 f_1 (最終ホップのノードの D P I D , オリジナルの M A C D A)、称号情報 C_n を書き込むものとしている。

【 0 0 7 2 】

本実施形態によれば、上記した第 1 の実施形態と比較して、先頭と最終ホップの転送ノード 1 0 において、オリジナルの M A C D A を、縮退・復元する処理が追加されることになるが、追加ヘッダを用いないため、オーバーヘッドを少なくすることが可能となる。なお、図 1 4 に示したパケット構成は、あくまで一例であり、照合情報の格納に用いるフィールドや、照合情報の内容、オリジナルのデータを縮退するための関数等は適宜変更することができる。

【 0 0 7 3 】

以上のように、本発明は、既存パケットの特定のフィールドに照合情報を含める形態でも実現することが可能である。

【 0 0 7 4 】

[第 3 の実施形態]

次に、本発明の第 3 の実施形態について説明する。上記実施形態では、処理規則を用いて照合情報を生成する例を説明した。第 3 の実施形態では、照合情報の他の生成例を説明する。

【 0 0 7 5 】

照合情報は、転送ノード 1 0 の識別子 (例えば、転送ノード 1 0 に付与された I D や M A C アドレス等) を用いて生成されてもよい。転送ノード 1 0 の識別子 (D P I D) を用いて照合情報を算出することとした場合、次式 [数 3]、[数 4] により照合情報を算出することができる。

【 0 0 7 6 】

10

20

30

40

【数3】

$$C_n = f(C_{n-1}, DPID_n)$$

【数4】

$$C_n = \{h_n, h_{n-1}, h_{n-2}, h_{n-3}, \dots, h_0\}$$

$$h_n = f(DPID_n)$$

【0077】

ここで、 C_n は、パケット転送経路上の n 番目の転送ノード# n が生成する照合情報であり、 C_{n-1} は、受信パケットに含まれている $n-1$ 番目の転送ノード# $n-1$ が生成した照合情報である。また、 $DPID_n$ は、パケット転送経路上の n 番目の転送ノード# n の識別子である。また、関数 $f(x)$ としては、MD5などのハッシュ関数を用いることができる。このようなハッシュ関数を用いることで、照合情報 C_n は、ユーザパケットが転送過程で経由した転送ノードの組み合わせを一意に表わした値となる。

10

【0078】

照合情報は、処理規則記憶部14に格納されているテーブルのID(テーブルID)を用いて生成してもよい。転送ノード n のテーブルIDを【数3】、【数4】に代入して算出することができる。

【0079】

また、照合情報は、転送ノード間のリンクのID(リンクID)を用いて生成してもよい。転送ノード n がパケットを転送するリンクのIDをリンクID n とすると、リンクID n を【数3】、【数4】に代入することで、照合情報を生成することができる。

20

【0080】

本実施形態で説明した、照合情報の生成例により、転送ノード10によるパケット転送が、制御装置20が決定したパケット転送経路通りに実行されたかを確認できる。つまり、第1の実施形態及び第2の実施形態とは異なる粒度で、制御装置20と転送ノード10との同期を確認することができる。なお、転送ノード10と制御装置20との同期確認の粒度が、動的に変更されてもよい。例えば、制御装置20は、処理規則を用いて生成していた照合情報を、本実施形態で説明した転送ノード10の識別子等を用いて生成するように、動的に変更してもよい。このような構成により、システム管理者が、必要に応じて同期確認検査の粒度を動的に切り替えることが可能になる。

30

【0081】

以上、本発明の実施形態を説明したが、本発明は、上記した実施形態に限定されるものではなく、本発明の基本的技術的思想を逸脱しない範囲で、更なる変形・置換・調整を加えることができる。例えば、転送ノード10が、制御装置20からの指示に基づいて、上記した照合情報の生成の可否、もしくは、生成方法を変更するようにしてもよい。また、制御装置20が、予め作成されたスケジュールに基づいて、照合情報の生成を指示した上で、照合情報が設定された処理規則を設定することで、種々の検査、サンプリングを行うことも可能となる。

40

【0082】

また、上記した各実施形態では、転送ノード10側で照合情報による同期確認を行うものとして説明したが、冒頭の概要に述べたように、任意の位置の転送ノード10が制御装置20に照合情報を送信し、制御装置20側で同期確認を行うようにすることも可能である。上記照合情報の送信は、該当する転送ノードの処理規則に、照合情報の送信を行わせるアクションを追加することで容易に実現することが可能である。

【0083】

また、制御装置20が、上記照合情報による同期確認を行う主体(転送ノード10/制御装置20)を切り換えることができるようにしてもよい。

【0084】

50

また、上記した第2の実施形態では、最終ホップのノードが追加ヘッダの除去やMAC DAフィールドを書き戻すものとして説明したが、ホスト側でこれらの処理を行う構成も採用可能である。

【0085】

また、上記した各実施形態で説明した検証情報として、非特許文献2のFlow Cookieなどのコントローラがフローエントリを一意に特定するために生成する情報を用いることも可能である。

【0086】

なお、上記の特許文献および非特許文献の各開示を、本書に引用をもって繰り込むものとする。本発明の全開示（請求の範囲を含む）の枠内において、さらにその基本的技術思想に基づいて、実施形態ないし実施例の変更・調整が可能である。また、本発明の請求の範囲（クレーム）の枠内において、種々の開示要素の多様な組み合わせないし選択が可能である。

10

【符号の説明】

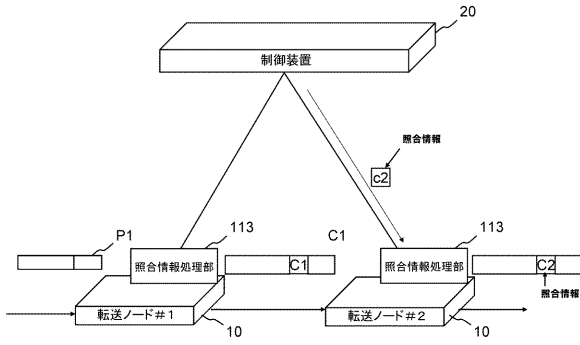
【0087】

- 10 転送ノード
- 11 転送処理部
- 12 制御装置通信部
- 13 処理規則管理部
- 14 処理規則記憶部
- 15 パケットバッファ
- 20 制御装置
- 21 ノード通信部
- 22 制御メッセージ処理部
- 23 トポロジ管理部
- 24 経路・アクション計算部
- 25 処理規則管理部
- 25A 同期確認部
- 27 照合情報生成部
- 28 処理規則記憶部
- 31 ユーザパケット
- 32、32a 照合情報付きパケット
- 33 照合情報格納ヘッダ
- 111 処理規則検索部
- 112 アクション実行部
- 113 照合情報処理部
- 221 メッセージ解析・処理部
- 222 メッセージ生成部

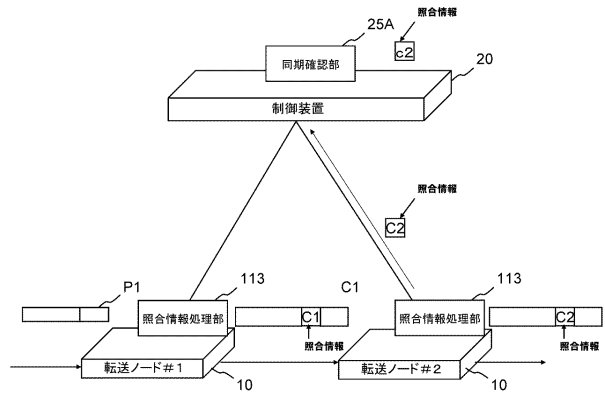
20

30

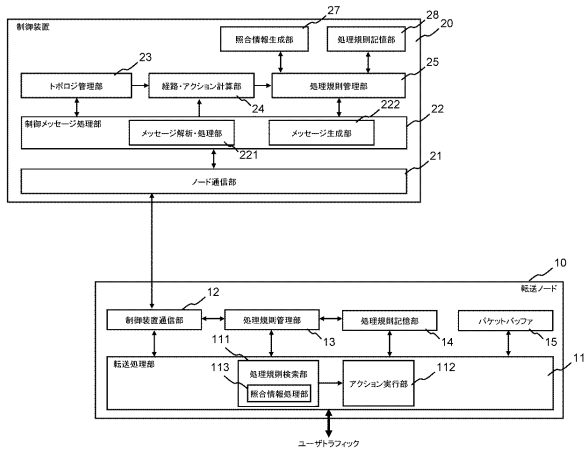
【図1】



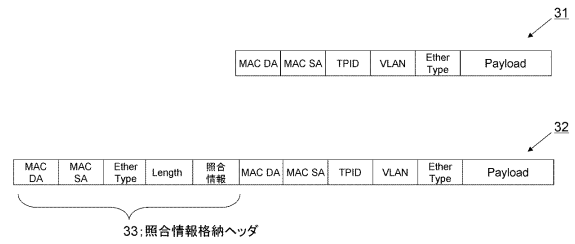
【図2】



【図4】



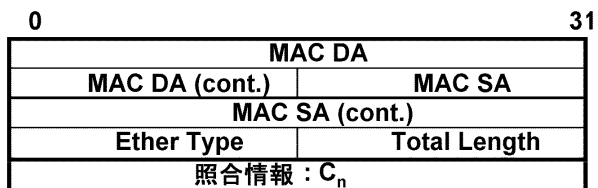
【図6】



【図5】

Priority	マッチングルール	Actions	照合情報
65535	A	Output to Port #9	0x0ceb
:	:	:	:
2048	B	Output to Port #7	0xff01
:	:	:	:
32	X	Output to Port #2	0xa310
:	:	:	:

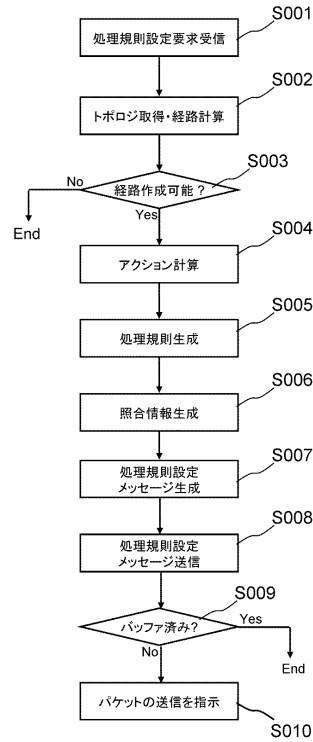
【図7】



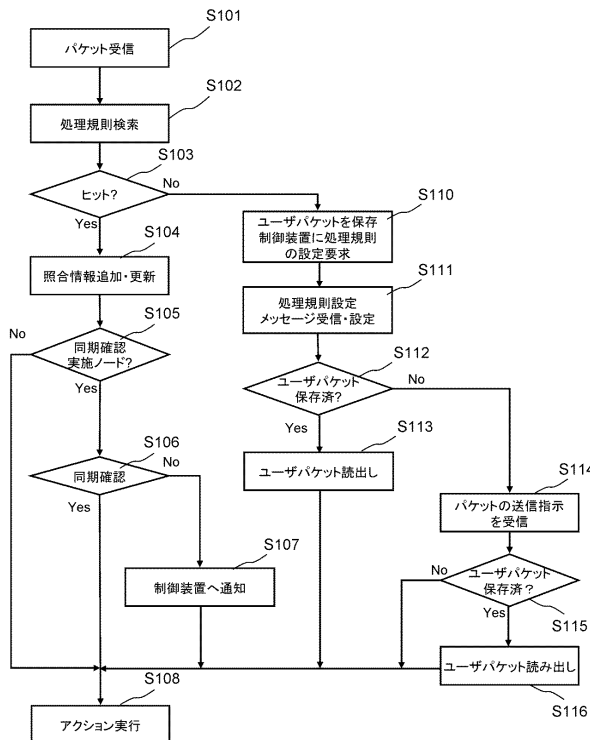
【図8】

DPID	Priority	マッチングルール	Actions	照合情報
1	65535	A	Output to Port #9	N/A
2	65535	A	Output to Port #6	0xa92b
3	65535	A	Output to Port #1	0x0ceb
1	4096	B	Output to Port # 7	N/A
2	2048	B	Output to Port #4	0x349a
3	4096	B	Output to Port #3	0xff01
1	32	X	Output to Port #2	N/A
2	32	X	Output to Port #5	0x945d

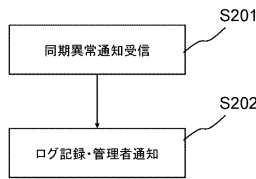
【図9】



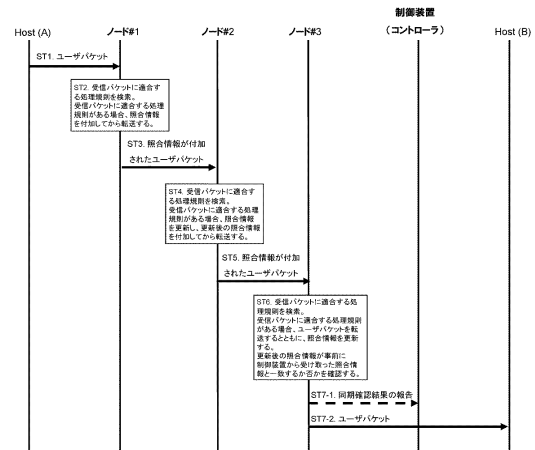
【図10】



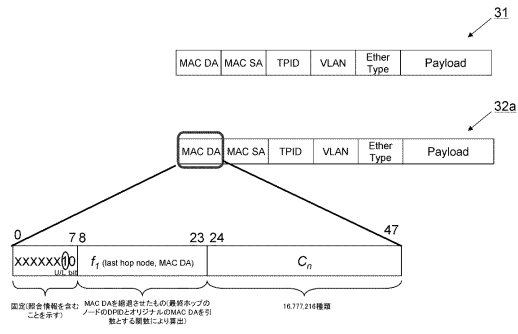
【図11】



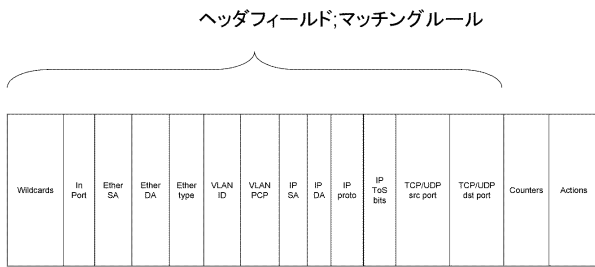
【図12】



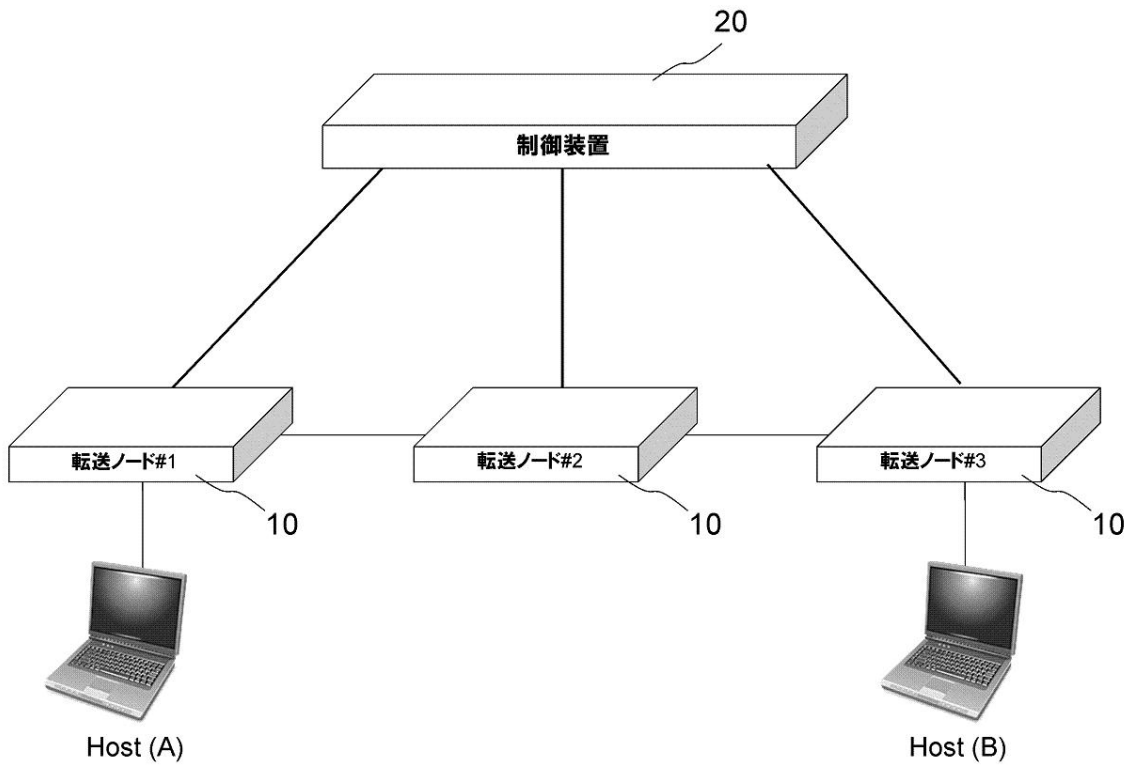
【図14】



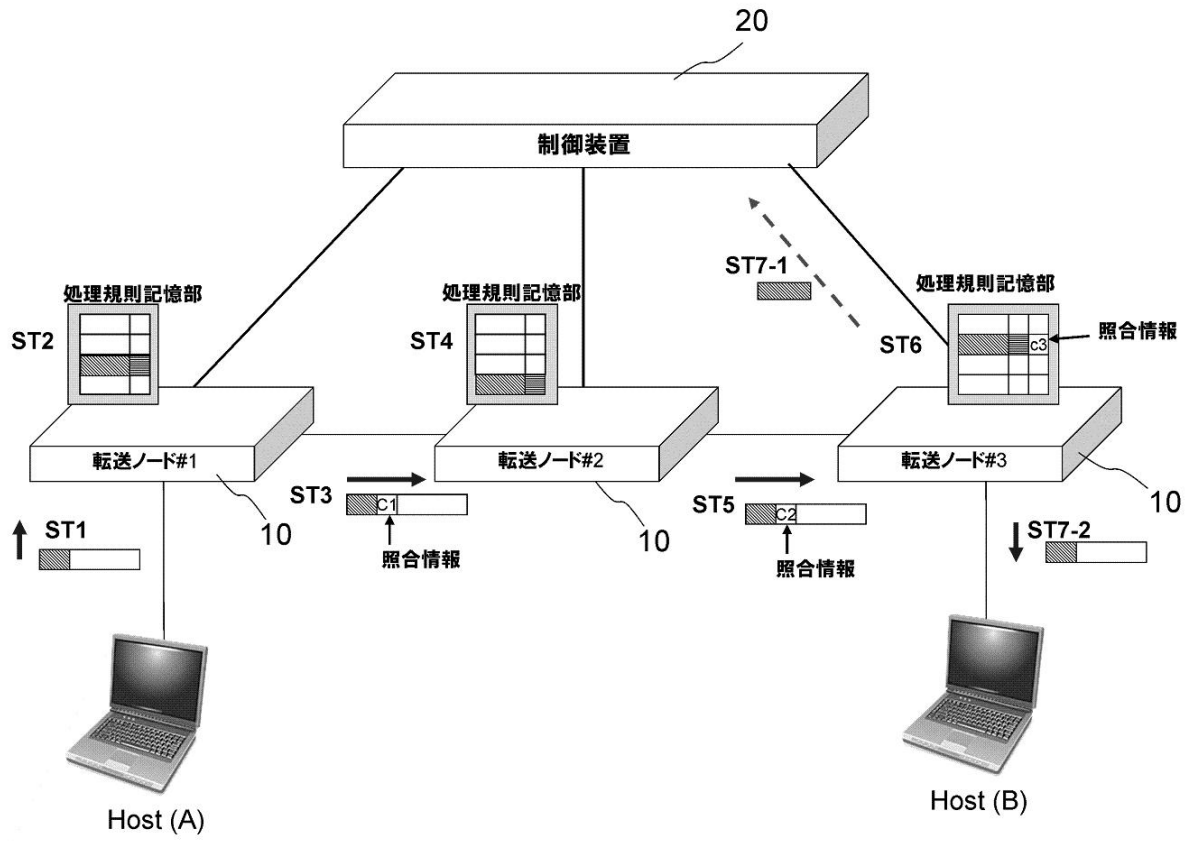
【図15】



【図3】



【図13】



フロントページの続き

(56)参考文献 特開平09 - 098189 (JP, A)

Nick McKeown et.al., OpenFlow: Enabling Innovation in Campus Networks, ACM SIGCOMM Computer Communication Review, 2008年 4月, vol.38, No.2, pp.69-74

(58)調査した分野(Int.Cl., DB名)

H04L 12/717