

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5644948号
(P5644948)

(45) 発行日 平成26年12月24日(2014.12.24)

(24) 登録日 平成26年11月14日(2014.11.14)

(51) Int.Cl. F I
HO 4 L 12/717 (2013.01) HO 4 L 12/717
HO 4 L 12/761 (2013.01) HO 4 L 12/761

請求項の数 10 (全 18 頁)

<p>(21) 出願番号 特願2013-528074 (P2013-528074) (86) (22) 出願日 平成24年8月10日 (2012.8.10) (86) 国際出願番号 PCT/JP2012/070429 (87) 国際公開番号 W02013/022082 (87) 国際公開日 平成25年2月14日 (2013.2.14) 審査請求日 平成26年2月7日 (2014.2.7) (31) 優先権主張番号 特願2011-176226 (P2011-176226) (32) 優先日 平成23年8月11日 (2011.8.11) (33) 優先権主張国 日本国(JP)</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 大嶽 貴寛 東京都港区芝五丁目7番1号 日本電気株式会社内 審査官 松崎 孝大</p>
---	---

最終頁に続く

(54) 【発明の名称】 パケット転送システム、制御装置、パケット転送方法およびプログラム

(57) 【特許請求の範囲】

【請求項1】

互いに接続された第1の中継装置群と、
 複数のポートを備えて、前記第1の中継装置とそれぞれ接続された第2の中継装置と、
 前記第1、第2の中継装置間の物理リンクを用いて、仮想的な論理リンクとなるトランクを構成する制御装置と、を含み、
 前記制御装置は、
 前記第1の中継装置側の前記トランクの構成ポートの中から代表ポートを決定し、
 前記第1の中継装置群に含まれる中継装置が所定の制御対象パケットを受信した場合、
 前記代表ポートを有する中継装置を介して前記受信した制御対象パケットを送信するよう
 、前記第1の中継装置群を制御すること、
 を特徴とするパケット転送システム。

【請求項2】

前記所定の制御対象パケットは、ブロードキャストパケットまたはマルチキャストパケットである請求項1のパケット転送システム。

【請求項3】

前記各第1の中継装置は、少なくとも、
 前記第2の中継装置と接続されたポートであり、前記代表ポートの候補となるメンバポートと、
 他の第1の中継装置と接続されたスタックリンクポートと、

前記第 1、第 2 の中継装置以外の装置と接続された出力ポートと、を備える請求項 1 または 2 のパケット転送システム。

【請求項 4】

前記制御装置は、

前記第 1 の中継装置のポート情報を含むトランクの構成情報を記憶するトランク情報記憶部と、

前記トランク情報記憶部を参照して前記代表ポートを選択するとともに、前記第 1 の中継装置を制御するトランク制御コマンド生成部と、

を備える請求項 1 から 3 いずれか一のパケット転送システム。

【請求項 5】

前記トランク制御コマンド生成部は、前記第 1 の中継装置間のスタックリンクの障害状況に応じて、前記第 1 の中継装置間の前記代表ポートを有する中継装置までのパケット転送経路を変更する請求項 4 のパケット転送システム。

【請求項 6】

前記第 1 の中継装置が 3 台以上あり、互いにリング状に接続されている請求項 1 から 5 いずれか一のパケット転送システム。

【請求項 7】

前記各第 1 の中継装置は、

パケットの受信ポートと、前記所定の制御対象パケットを特定するための照合条件と、出力ポートと、を対応付けたエントリを格納する転送テーブルを保持し、

前記制御装置は、前記各第 1 の中継装置の転送テーブルを書き換えることにより、前記各第 1 の中継装置を制御する請求項 1 から 6 いずれか一のパケット転送システム。

【請求項 8】

互いに接続された第 1 の中継装置群と、

複数のポートを備えて、前記第 1 の中継装置とそれぞれ接続された第 2 の中継装置と、に接続され、

前記第 1、第 2 の中継装置間の物理リンクを用いて、仮想的な論理リンクとなるトランクを構成する制御装置であって、

前記第 1 の中継装置側の前記トランクの構成ポートの中から代表ポートを決定し、

前記第 1 の中継装置群に含まれる中継装置が所定の制御対象パケットを受信した場合、前記代表ポートを有する中継装置を介して前記受信した制御対象パケットを送信するよう、前記第 1 の中継装置群を制御する制御装置。

【請求項 9】

互いに接続された第 1 の中継装置群と、

複数のポートを備えて、前記第 1 の中継装置とそれぞれ接続された第 2 の中継装置と、に接続され、

前記第 1、第 2 の中継装置間の物理リンクを用いて、仮想的な論理リンクとなるトランクを構成する制御装置が、

前記第 1 の中継装置側の前記トランクの構成ポートの中から代表ポートを決定するステップと、

前記第 1 の中継装置群に含まれる中継装置が所定の制御対象パケットを受信した場合、前記代表ポートを有する中継装置を介して前記受信した制御対象パケットを送信するよう、前記第 1 の中継装置群を制御するステップと、を含むパケット転送方法。

【請求項 10】

互いに接続された第 1 の中継装置群と、

複数のポートを備えて、前記第 1 の中継装置とそれぞれ接続された第 2 の中継装置と、に接続され、

前記第 1、第 2 の中継装置間の物理リンクを用いて、仮想的な論理リンクとなるトランクを構成する制御装置を構成するコンピュータに実行させるプログラムであって、

前記第 1 の中継装置側の前記トランクの構成ポートの中から代表ポートを決定する処理

10

20

30

40

50

と、

前記第1の中継装置群に含まれる中継装置が所定の制御対象パケットを受信した場合、前記代表ポートを有する中継装置を介して前記受信した制御対象パケットを送信するように、前記第1の中継装置群を制御する処理と、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2011-176226号(2011年8月11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。

10

【0002】

本発明は、パケット転送システム、制御装置、パケット転送方法およびプログラムに関し、特に、複数の物理リンクを用いて仮想的な論理リンクとなるトランクが構成されているパケット転送システム、制御装置、パケット転送方法およびプログラムに関する。

【背景技術】

【0003】

近年、オープンフロー(OpenFlow)という技術が提案されている(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、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インストラクション(Instructions)と、の組が定義される(図16参照)。

20

【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16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、当該エントリのインストラ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、フラディング、廃棄等)を実施する。一方、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エントリ設定の要求、即ち、受信パケットの処理内容の決定の要求を送信する。オープンフロースイッチは、処理内容が定められた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う。

30

【0005】

上記特許文献1や非特許文献1、2に記載されているネットワークにおいて、ブロードキャストやマルチキャストを実現するには、個々のオープンフロースイッチに、複数のポートからパケットを転送させるフローエントリを設定することになる。

40

【先行技術文献】

【特許文献】

【0006】

【特許文献1】国際公開第2008/095010号

【非特許文献】

【0007】

【非特許文献1】Nick McKeownほか7名、“OpenFlow: Enab

50

ling Innovation in Campus Networks”、[online]、[平成23(2011)年7月26日検索]、インターネット URL: <http://www.openflow.org/documents/openflow-wp-latest.pdf>

【非特許文献2】“OpenFlow Switch Specification” Version 1.1.0 Implemented (Wire Protocol 0x02) [平成23(2011)年7月26日検索]、インターネット URL: <http://www.openflow.org/documents/openflow-spec-v1.1.0.pdf>

【発明の概要】

【発明が解決しようとする課題】

【0008】

以下の分析は、本発明によって与えられたものである。上記特許文献1や非特許文献1、2に記載されているネットワークに、オープンフロースイッチが追加されるなど、ネットワークポロジに事後的な変更が生じることがある。個々のオープンフロースイッチは、自装置のフローテーブルに受信パケットに適合するフローエントリがあれば、オープンフローコントローラに、フローエントリの設定要求を行わない。このため、前記ネットワークポロジの変更によって、新規に追加されたオープンフロースイッチや接続に変更が加えられたオープンフロースイッチに、ブロードキャストパケットが転送されなくなってしまうことが起こりうる。

【0009】

そこで、ブロードキャストパケットの転送に関しては、ネットワークポロジの変化時にブロードキャスト配信経路を計算し、各経路上のオープンフロースイッチに対して、ブロードキャストパケット用のフローエントリを設定しておく方が望ましい。

【0010】

ところで、上記特許文献1や非特許文献1、2に記載されている技術を用いれば、並列するリンクを持ついくつかのオープンフロースイッチに適切なフローエントリ(例えば、特定のパケットを同一の宛先に転送する。)を設定することで、これらリンクを束ね、仮想的な論理リンクとなるトランクを構成することができる。

【0011】

しかしながら、上記トランクの一端の装置がオープンフローコントローラの制御対象外の通信装置である場合、各オープンフロースイッチは、それぞれ保持するフローエントリに従ってブロードキャストパケットを転送してしまうため、特定の宛先に同一パケットが複数送信されてしまうことが起こりうる。また仮に、前記ブロードキャストパケット用のフローエントリが設定されていない場合においても、その他のフローエントリにより意図しない経路でブロードキャストパケットが転送されてしまうことが起こりうる。

【0012】

本発明の目的とするところは、制御装置の制御対象外の外部中継装置との間にトランクが配置されている構成において、前記トランクを構成する中継装置(上記のオープンフロースイッチに相当。)に、ブロードキャストやマルチキャスト等の所定の制御対象パケットを適切に転送させることのできる構成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、互いに接続された第1の中継装置群と、複数のポートを備えて、前記第1の中継装置とそれぞれ接続された第2の中継装置と、前記第1、第2の中継装置間の物理リンクを用いて、仮想的な論理リンクとなるトランクを構成する制御装置と、を含み、前記制御装置は、前記第1の中継装置側の前記トランクの構成ポートの中から代表ポートを決定し、前記第1の中継装置群に含まれる中継装置が所定の制御対象パケットを受信した場合、前記代表ポートを有する中継装置を介して前記受信した制御対象パケットを送信するよう、前記第1の中継装置群を制御するパケット転送システムが提供される。

【0014】

10

20

30

40

50

本発明の第2の視点によれば、互いに接続された第1の中継装置群と、複数のポートを備えて、前記第1の中継装置とそれぞれ接続された第2の中継装置と、に接続され、前記第1、第2の中継装置間の物理リンクを用いて、仮想的な論理リンクとなるトランクを構成する制御装置であって、前記第1の中継装置側の前記トランクの構成ポートの中から代表ポートを決定し、前記第1の中継装置群に含まれる中継装置が所定の制御対象パケットを受信した場合、前記代表ポートを有する中継装置を介して前記受信した制御対象パケットを送信するよう、前記第1の中継装置群を制御する制御装置が提供される。

【0015】

本発明の第3の視点によれば、互いに接続された第1の中継装置群と、複数のポートを備えて、前記第1の中継装置とそれぞれ接続された第2の中継装置と、に接続され、前記第1、第2の中継装置間の物理リンクを用いて、仮想的な論理リンクとなるトランクを構成する制御装置が、前記第1の中継装置側の前記トランクの構成ポートの中から代表ポートを決定するステップと、前記第1の中継装置群に含まれる中継装置が所定の制御対象パケットを受信した場合、前記代表ポートを有する中継装置を介して前記受信した制御対象パケットを送信するよう、前記第1の中継装置群を制御するステップと、を含むパケット転送方法が提供される。本方法は、前記第1の中継装置群を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。

【0016】

本発明の第4の視点によれば、互いに接続された第1の中継装置群と、複数のポートを備えて、前記第1の中継装置とそれぞれ接続された第2の中継装置と、に接続され、前記第1、第2の中継装置間の物理リンクを用いて、仮想的な論理リンクとなるトランクを構成する制御装置を構成するコンピュータに実行させるプログラムであって、前記第1の中継装置側の前記トランクの構成ポートの中から代表ポートを決定する処理と、前記第1の中継装置群に含まれる中継装置が所定の制御対象パケットを受信した場合、前記代表ポートを有する中継装置を介して前記受信した制御対象パケットを送信するよう、前記第1の中継装置群を制御する処理と、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。

【発明の効果】

【0017】

本発明によれば、制御装置の制御対象外の外部中継装置との間にトランクが配置されている構成において、前記トランクを構成する中継装置に、ブロードキャストやマルチキャスト等の所定の制御対象パケットを適切に転送させることが可能となる。

【図面の簡単な説明】

【0018】

【図1】本発明の第1の実施形態の構成を表した図である。

【図2】本発明の第1の実施形態の制御装置の詳細構成を表したブロック図である。

【図3】本発明の第1の実施形態の制御装置のトランク情報記憶部に保持されるトランク構成情報の例である。

【図4】代表ポートに決定されたポート（メンバポート）を持つ中継装置の転送テーブルの内容を示す図である。

【図5】代表ポートに決定されなかったポート（メンバポート）を持つ中継装置の転送テーブルの内容を示す図である。

【図6】中継装置1101 - 1103間のリンクに障害が生じた場合の中継装置1102の転送テーブルの内容を示す図である。

【図7】中継装置1101 - 1103間のリンクに障害が生じた場合の中継装置1103の転送テーブルの内容を示す図である。

【図8】ブロードキャストパケットの転送経路の一例を示す図である。

【図9】外部中継装置から中継装置1102にブロードキャストパケットが転送された場合の転送経路を示す図である。

10

20

30

40

50

【図10】外部中継装置から中継装置1103にブロードキャストパケットが転送された場合の転送経路を示す図である。

【図11】スタックリンクに障害が生じた場合の転送経路を示す図である。

【図12】スタックリンクに障害が生じた場合の転送経路を示す別の図である。

【図13】本発明の第2の実施形態の構成を表した図である。

【図14】本発明の第3の実施形態の構成を表した図である。

【図15】本発明の第4の実施形態の構成を表した図である。

【図16】非特許文献2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】

【0019】

はじめに本発明の一実施形態の概要について図面を参照して説明する。以下、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。

【0020】

本発明は、その一実施形態において、図1に示すように、互いに接続された第1の中継装置群1101～1103と、複数のポートを備えて、前記第1の中継装置1101～1103とそれぞれ接続された第2の中継装置（外部中継装置）1106と、に接続され、前記第1、第2の中継装置間の物理リンクを用いて、仮想的な論理リンクとなるトランクを構成する制御装置1100を含む構成にて実現できる。

【0021】

例えば、図1の外部中継装置1106と中継装置1101、1102、1103との物理リンクを用いてトランクが構成されているものとする。ここで、制御装置1100は、前記トランクを構成するポートのうち中継装置1101のポートを代表ポートに決定したものとす。制御装置1100は、中継装置1101が、制御対象パケットを受信した場合、所定の経路に従って転送するよう制御する（図8参照）。

【0022】

また、制御装置1100は、第1の中継装置1102、1103が前記所定の制御対象パケットを受信した場合、中継装置1101を介して前記受信した制御対象パケットを送信するよう、前記第1の中継装置1102、1103を制御する（図9、図10参照）。

【0023】

このようにすることで、トランクを流れてくる制御対象パケットが、意図しない他の中継装置から重複して転送されてしまったり、ループを形成しないように制御することができる。

【0024】

[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1は、本発明の第1の実施形態の構成を表した図である。図1を参照すると、外部コマンドで受信パケットの転送等のパケット処理を制御可能な複数の中継装置（中継装置1101～1105）と、外部中継装置1106と、端末1107、1108と、前記中継装置1101～1105を制御する制御装置1100と、を含むパケット転送システムが示されている。

【0025】

中継装置1101～1105は、破線で示す制御チャネルを介して制御装置1100と接続された通信装置である。中継装置1101～1105は、制御装置1100からの制御に従って、外部中継装置1106に接続された端末1107と、中継装置1105に接続された端末1108間の通信を実現する。図1では、データ用の通信路を実線で示し、制御用の通信路を破線で示しているが、これら2つが混在していてもよい。

【0026】

本実施形態では、上記中継装置として、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を用いるものとする。具体的には、中継装置1101～1105は、図16に

10

20

30

40

50

示すようなマッチングルールと処理内容 (I n s t r u c t i o n s) とを対応付けたフローエントリを保持し、受信パケットに適合するフローエントリに従ってパケットの転送等を実施する。なお、上記オープンフロースイッチ以外の装置を中継装置として用いることもできる。例えば、T e l n e t 経由のC L I (コマンドラインインタフェース) にて、任意のパケットのマッチング条件を指定してパケットの経路を設定できるような装置を用いることも可能である。

【 0 0 2 7 】

外部中継装置 1 1 0 6 は、複数の物理ポートで L i n k A g g r e g a t i o n (I E E E 8 0 2 . 3 a d 参照。以下、「L A G」ともいう。) を構成する機能を有する。外部中継装置 1 1 0 6 は、L A G を構成する物理ポートで中継装置 1 1 0 1、中継装置 1 1 0 2、中継装置 1 1 0 3 と接続されている。

10

【 0 0 2 8 】

端末 1 1 0 7 は、外部中継装置 1 1 0 6 に接続されたホスト端末等である。また、端末 1 1 0 8 は、中継装置 1 1 0 5 に接続されたホスト端末等である。

【 0 0 2 9 】

また、本実施形態では、中継装置 1 1 0 1 ~ 1 1 0 5 のうち、外部中継装置 1 1 0 6 と接続された中継装置 1 1 0 1 ~ 1 1 0 3 が、それぞれ他の中継装置と リング 状に接続されている。

【 0 0 3 0 】

このほか、中継装置 1 1 0 4 は、中継装置 1 1 0 1、1 1 0 2、1 1 0 5 と接続され、中継装置 1 1 0 5 は、中継装置 1 1 0 1、1 1 0 3、1 1 0 4 と接続されている。

20

【 0 0 3 1 】

なお、以上のようなネットワークポロジは、本発明を簡単に説明するための一例として示したものであり、種々の変更を加えることができる。例えば、図 1 に示した中継装置、外部中継装置、端末の数等は、特に限定されない。

【 0 0 3 2 】

続いて、図 1 の制御装置 1 1 0 0 について詳細に説明する。図 2 は、本発明の第 1 の実施形態の制御装置 1 1 0 0 の詳細構成を表したブロック図である。図 2 を参照すると、中継装置通信部 1 0 1 と、トポロジ取得部 1 0 2 と、トポロジテーブル 1 0 3 と、ユニキャスト経路探索部 1 0 4 と、ユニキャスト用経路制御コマンド生成部 1 0 5 と、ブロードキャスト経路探索部 (B C 経路探索部) 1 0 6 と、ブロードキャスト経路コマンド生成部 (B C 用経路制御コマンド生成部) 1 0 7 と、トランク情報入力部 1 0 8 と、トランク情報記憶部 1 0 9 と、トランク制御コマンド生成部 1 1 0 とを備えた構成が示されている。

30

【 0 0 3 3 】

中継装置通信部 1 0 1 は、トポロジ取得部 1 0 2、ユニキャスト用経路制御コマンド生成部 1 0 5、B C 用経路制御コマンド生成部 1 0 7 またはトランク制御コマンド生成部 1 1 0 にて生成された制御用コマンドを中継装置 1 1 0 1 ~ 1 1 0 5 に送信し、中継装置 1 1 0 1 ~ 1 1 0 5 からの応答をトポロジ取得部 1 0 2、ユニキャスト用経路制御コマンド生成部 1 0 5、B C 用経路制御コマンド生成部 1 0 7 またはトランク制御コマンド生成部 1 1 0 に入力する動作を行う。

40

【 0 0 3 4 】

より具体的には、中継装置通信部 1 0 1 は、中継装置 1 1 0 1 ~ 1 1 0 5 との制御用セッションの確立、制御用コマンドの送信・受信を行う。制御用コマンドとしては、非特許文献 2 に記載されているオープンプロトコルを用いることができる。また、T e l n e t 経由のC L I や、S N M P (S i m p l e N e t w o r k M a n a g e m e n t P r o t o c o l) を用いて通信を行ってもよい。

【 0 0 3 5 】

トポロジ取得部 1 0 2 は、中継装置通信部 1 0 1 を介して中継装置 1 1 0 1 ~ 1 1 0 5 と通信し中継装置 1 1 0 1 ~ 1 1 0 5 間の接続関係を示すトポロジ情報を取得し、トポロジテーブル 1 0 3 に登録する。前記トポロジの取得方法としては、中継装置間で隣接する

50

中継装置のインタフェース情報等を収集させ、吸い上げる方法がある。このようなプロトコルとして、Link Layer Discovery Protocolが代表的である。その他、制御装置1100から、特定の中継装置の特定ポートから当該中継装置のIDとポート番号を含むパケットを出力するように制御を行い、対向となる中継装置から該パケットを受信することで、中継装置間の認識を行う方法も採用できる。前記制御装置からのパケット出力とパケット受信には、非特許文献2のPacket-out、Packet-inメッセージを用いることができる。もちろん、ネットワークトポロジを事前に設定しておく方法も採用できる。

【0036】

ユニキャスト経路探索部104は、トポロジテーブル103を参照し、端末と接続している間の中継装置間の経路を計算する。ユニキャスト経路探索部104における経路の計算方法としては、最短経路木(ダイクストラ法が代表的)を用いることができる。これら経路は、単一とは限らず、通信単位毎に、異なる経路が計算されるようにしてもよい。また、事前に、すべての中継装置を起点または終点とした計算を行っておき、必要に応じてこれら予め計算した経路を用いることとしてもよい。

10

【0037】

ユニキャスト用経路制御コマンド生成部105は、ユニキャスト経路探索部104の計算した経路上の中継装置に対して、制御用コマンドを送信し、経路を設定する。制御用コマンドとしては、例えば、中継装置が保持するフローエントリを書き換える非特許文献2のFLOW Table Modificationメッセージを用いることができる。例えば、図1の端末1108のIP(Internet Protocol)アドレス、MAC(Media Access Control)アドレス、TCP(Transmission Control Protocol)/UDP(User Datagram Protocol)のポート等をマッチングルールとし、経路上の次の中継装置と接続されたポートから当該パケットを出力させる処理内容(Instructions)と定めたフローエントリを設定することで、端末1108宛てのパケットの経路制御が実現される。

20

【0038】

BC経路探索部106は、ブロードキャスト用に、端末と接続している中継装置から、少なくとも当該端末と接続している他のすべての中継装置へ配信できるような全域木(スパンニングツリー)経路を計算する。BC経路探索部106は、ネットワークで一つの全域木経路を計算してもよいし、複数の全域木経路を計算してもよい。複数の全域木経路を計算する場合には、端末と接続している中継装置毎に異なる全域木経路を計算してもよい。全域木の計算方法としては、最小全域木(プリム法、クラスカル法が代表的)を用いることができる。また、その計算の際には、ある中継装置が端末と接続しておらず末端である場合には除外することを繰り返して、端末と接続している中継装置間に存在しない中継装置を除外するようにしてもよい。さらに、BC経路探索部106は、トランク制御コマンド生成部110から代表ポートの選択結果を受信している場合、前記代表ポートが経路に含まれるよう経路の再計算を行う。

30

【0039】

BC用経路制御コマンド生成部107は、ブロードキャスト用配信経路制御処理とブロードキャストパケット許可処理とを行う。ブロードキャスト用配信経路制御処理は、ブロードキャスト用の経路を構築する処理であり、上述のユニキャスト用の経路と同様に中継装置にフローエントリを設定する。なお、このとき、ブロードキャスト用のフローエントリは、前記ユニキャスト用のフローエントリより低い優先度が設定される。これにより、ユニキャスト用のフローエントリにマッチングしないパケットであって、ブロードキャスト用のフローエントリにマッチングするパケットが、ブロードキャストの対象となる。

40

【0040】

ブロードキャストパケット許可処理としては、端末からのブロードキャストパケットのブロードキャスト用の配信経路での転送を許可する処理、パケットの任意のフィールドに

50

ブロードキャストパケットであることを示す情報を書き込むよう中継装置を制御する処理が行われる。例えば、端末からの入口にある中継装置に、送信MACアドレスの先頭ビット(I/Gビット)を1に書き換えるフローエントリを設定することで実現できる。

【0041】

また、別の方法としては、ユニキャスト用経路制御コマンド生成部105およびBC用経路制御コマンド生成部107のうち一つもしくは両方が、端末からの入口にある中継装置に、ユニキャスト経路による配信、ブロードキャスト経路による配信であることを区別(Unicast/BC配信フラグ)できるように任意のフィールドを変更させるようにしてもよい。また、この場合において、宛先側端末への出口にある中継装置で、前記任意のフィールドを元に戻すようにしてもよい。変更フィールドとしては、IPのTosフィールド、VLAN Priorityフィールド等が考えられる。

10

【0042】

また別の方法としては、端末の入り口にある中継装置に、任意のアドレスを縮退して、この縮退したアドレスに、Unicast/BC配信フラグを入れるようにしてもよい。例えば、送信先MACアドレスを縮退して、送信先MACアドレスに縮退したMACアドレスと前記区別フラグを入れることが挙げられる。なお、この場合、宛先側端末への出口にある中継装置にて、縮退したアドレスから元のアドレスを復元する処理が必要となる。

【0043】

また別の方法としては、ユニキャストとマルチキャストに別のVLAN IDを割り当ててもよい。例えば、VLAN IDを、別途割り当てた専用のVLAN IDに変換して、送信先MACアドレスと元のVLAN IDを縮退したIDを生成し、送信先MACをこの縮退IDとすることが挙げられる。

20

【0044】

これら任意のフィールドの変更により、ユニキャスト経路とブロードキャスト経路による配信の区別ができるため、ブロードキャスト用の配信経路制御処理で、宛先不明のユニキャストパケットもブロードキャスト用配信経路で配信できるようになる。

【0045】

トランク情報入力部108は、CLIやGUI(グラフィカルユーザインターフェース)などにより中継装置のトランク構成情報の入力を受け付ける。

【0046】

トランク情報記憶部109は、トランク情報入力部108で入力されたトランク構成情報を記憶する。

30

【0047】

図3は、トランク情報記憶部に保持されるトランク構成情報の一例である。図3の例では、トランクを構成する中継装置について、外部中継装置1106と接続する(中継装置の)ポート(以下、「メンバポート」という。)であるメンバポート情報と、トランク構成時のパケット回り込みを防ぐパケット回り込み抑制処理のためにトランクを構成する中継装置間を接続するポートであるスタックリンクポート情報とを指定した内容となっている。さらに、図3の例では、スタックリンクポートとして、1つの中継装置にEAST、WESTの2つのポートが指定されている。EASTスタックリンクポートにはトランクを構成する別の中継装置のWESTスタックリンクポートを接続される。このようにしてトランクを構成するスタックリンクポート間のリンクはリングを構成するように接続される。

40

【0048】

例えば、図3によると、中継装置1101のポート#1、中継装置1102のポート#4および中継装置1103のポート#3がメンバポートとなっている。これは、図1に示すように、中継装置1101のポート#1、中継装置1102のポート#4および中継装置1103のポート#3が外部中継装置1106と接続され、トランクを構成していることを示している。また、中継装置1101のポート#2は、EASTスタックリンクポートとなっている。このEASTスタックリンクポートは、図1に示すように、他の中継装

50

置 1 1 0 3 の W E S T スタックリンクポートであるポート # 2 と接続されている。また、中継装置 1 1 0 1 のポート # 3 は、W E S T スタックリンクポートとして他の中継装置 1 1 0 2 の E A S T スタックリンクポートであるポート # 5 と接続されている。同様に、中継装置 1 1 0 2 のポート # 6 は、W E S T スタックリンクポートとして他の中継装置 1 1 0 3 の E A S T スタックリンクポートであるポート # 1 と接続され、リングが構成されている。

【 0 0 4 9 】

なお、図 3 では、1 組のトランク構成情報が示されているが、ネットワーク内に複数のトランクが構成されている場合には（図 1 3 ~ 図 1 5 参照）、複数組のトランク構成情報を記憶することが可能である。

10

【 0 0 5 0 】

トランク制御コマンド生成部 1 1 0 は、代表ポート選出処理と、スタックリンク制御処理とを実行する。代表ポート選出処理では、トランク情報記憶部 1 0 9 で保持するメンバポートの中から代表ポートを 1 つ選出する処理が行われる。トランク制御コマンド生成部 1 1 0 は、前記メンバポートから代表ポートを選出する際に、トポロジテーブル 1 0 3 を参照して、候補となる各メンバポートが外部中継装置 1 1 0 6 と通信可能となっていることを確認する。

【 0 0 5 1 】

この代表ポートの選出結果は B C 経路探索部 1 0 6 にも通知される。B C 経路探索部 1 0 6 は、ブロードキャスト用の配信経路の再計算を行う。代表ポートの選出処理は、トランク情報入力部 1 0 8 にトランク構成情報が入力された場合や、代表ポートが障害となりトポロジテーブル 1 0 3 から代表ポートの物理トポロジの変化が通知された場合などに実行される。

20

【 0 0 5 2 】

スタックリンク制御処理は、中継装置の代表ポートの有無やスタックリンクポートの障害に基づいて、中継装置通信部 1 0 1 を介して、トランクを構成する中継装置 1 1 0 1、1 1 0 2、1 1 0 3 に対し、メンバポートおよびスタックリンクポートから受信したブロードキャストパケットの転送先を指示するフローエントリを設定する。

【 0 0 5 3 】

具体的には、選出された代表ポートでブロードキャストパケットが受信された場合、当該ブロードキャストパケットの送信は許可される。一方、代表ポート以外のメンバポートで受信された場合、当該ブロードキャストパケットの送信は禁止され、スタックリンクポートからの転送や廃棄が行われる。これにより、トランクを構成するポートを持つ中継装置から複数のブロードキャストパケットが出力されることを抑制できる。

30

【 0 0 5 4 】

図 4 は、代表ポートとして選出されたポートを有する中継装置（例えば、中継装置 1 1 0 1）に保持される転送テーブルの内容（マッチングルールは受信ポートを除き省略。）の例である。図 4 に示されたように、代表ポートとして選出されたポートを有する中継装置では、ブロードキャストパケットを受信した場合、その受信ポートが代表ポート、スタックリンクポートのいずれの場合も、当該パケットをブロードキャスト転送経路の次ホップの中継装置に転送するフローエントリが設定される。

40

【 0 0 5 5 】

図 5 は、メンバポートが代表ポートとして選出されなかった中継装置（例えば、中継装置 1 1 0 2、1 1 0 3）に保持される転送テーブルの内容（マッチングルールは受信ポートを除き省略。）の例である。図 5 の例では、メンバポートまたは W E S T スタックリンクポートからブロードキャストパケットを受信した場合、E A S T スタックリンクポートから転送するフローエントリと、E A S T スタックリンクポートからブロードキャストパケットを受信した場合、W E S T スタックリンクポートから転送するフローエントリとが設定されている。

【 0 0 5 6 】

50

上記のように、メンバポートが代表ポートとして選出されなかった中継装置においては、スタックリンクポートを経由して隣接する中継装置へブロードキャストパケットの転送が行われる。そして、ブロードキャストパケットが、代表ポートを有する中継装置に到達すると、当該中継装置から、ブロードキャスト用の経路に従ってパケット転送が行われる。

【 0 0 5 7 】

また、トランク制御コマンド生成部 1 1 0 によるスタックリンク制御処理では、スタックリンクポートの障害等も考慮される。図 6 は、メンバポートが代表ポートとして選出されなかった中継装置（例えば、中継装置 1 1 0 2、1 1 0 3）において、WESTスタックリンクポートに障害が発生している場合の転送テーブルの内容（マッチングルールは受信ポートを除き省略。）の例である。図 6 の例では、メンバポートで受信したパケットをEASTスタックリンクポートから転送するフローエントリと、EASTスタックリンクポートまたはWESTスタックリンクポートで受信したパケットを廃棄するフローエントリとが設定されている。

10

【 0 0 5 8 】

図 7 は、メンバポートが代表ポートとして選出されなかった中継装置（例えば、中継装置 1 1 0 2、1 1 0 3）において、EASTスタックリンクポートに障害が発生している場合の転送テーブルの内容（マッチングルールは受信ポートを除き省略。）の別の一例である。図 7 の例では、メンバポートまたはWESTスタックリンクポートで受信したパケットを、WESTスタックリンクポートから転送するフローエントリと、EASTスタックリンクポートで受信したパケットを廃棄するフローエントリとが設定されている。

20

【 0 0 5 9 】

図 6、図 7 のようなフローエントリを設定することにより、リングを構成する中継装置 1 1 0 1 ~ 1 1 0 3 のスタックリンクポートで 1 重障害が発生した場合、EASTのスタックリンクポートからWESTのスタックリンクポートに転送されるパケットは、EASTのスタックリンクポートで障害が発生している中継装置で折り返されて代表ポートが存在する中継装置まで転送される。WESTのスタックリンクポートからEASTのスタックリンクポートに転送されるパケットは、WESTのスタックリンクポートで障害が発生している中継装置で廃棄される。このため、パケットがループすることを防ぐことができる。

30

【 0 0 6 0 】

なお、上記した制御装置 1 1 0 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、BC経路探索部 1 0 6、BC用経路制御コマンド生成部 1 0 7、トランク情報入力部 1 0 8、トランク情報記憶部 1 0 9 およびトランク制御コマンド生成部 1 1 0 を追加して構成することができる。

【 0 0 6 1 】

また、図 2 に示した制御装置 1 1 0 0 の各部（処理手段）は、制御装置 1 1 0 0 を構成するコンピュータに、そのハードウェアを用いて、上記した各処理を実行させるコンピュータプログラムにより実現することもできる。

【 0 0 6 2 】

続いて、本実施形態の動作について図面を参照して詳細に説明する。以下の説明では、図 8 に示すように、外部中継装置 1 1 0 6 経由で端末 1 1 0 7 から端末 1 1 0 8 宛てに送信されたブロードキャストパケットについて、中継装置 1 1 0 1、中継装置 1 1 0 4、中継装置 1 1 0 5 経由で転送する経路が計算されているものとする（図 8 の実線参照）。また、代表ポート選出処理の結果、中継装置 1 1 0 1 のメンバポートが、代表ポートとして選出されているものとする。

40

【 0 0 6 3 】

図 8 の端末 1 1 0 7 がブロードキャストパケットを外部中継装置 1 1 0 6 に送信すると、外部中継装置 1 1 0 6 は、LAGを構成するポートのうち中継装置 1 1 0 1 の代表ポートにパケットを転送する。

50

【 0 0 6 4 】

中継装置 1 1 0 1 は、図 4 に示すフローエントリに従って、ブロードキャスト転送経路の次ホップの中継装置である中継装置 1 1 0 4 に接続されているポートからブロードキャストパケットを転送する。前記中継装置 1 1 0 4 は、制御装置 1 1 0 0 から設定されたブロードキャスト転送経路を実現するフローエントリに従って、中継装置 1 1 0 5 にパケットを転送する。中継装置 1 1 0 5 は、同様に、制御装置 1 1 0 0 から設定されたブロードキャスト転送経路を実現するフローエントリに従って、端末 1 1 0 8 にパケットを転送する。この結果、図 8 の矢線に示されたとおりにブロードキャストパケットが転送される。

【 0 0 6 5 】

図 9 は、端末 1 1 0 7 からブロードキャストパケットを受信した外部中継装置 1 1 0 6 が、LAG を構成するポートのうち中継装置 1 1 0 2 のメンバポートにパケットを転送した場合の転送経路を示す図である。

【 0 0 6 6 】

この場合、中継装置 1 1 0 2 は、図 5 に示すフローエントリに従って、自装置（中継装置 1 1 0 2 ）の E A S T スタックリンクポート（図 9 のポート # 5 ）からパケットを転送する。前記パケットを受信した中継装置 1 1 0 1 は、図 4 に示すフローエントリに従って、ブロードキャスト転送経路の次ホップの中継装置である中継装置 1 1 0 4 に接続されているポートからブロードキャストパケットを転送する。その後は、図 8 と同様に、ブロードキャスト用の経路に従ってブロードキャストパケットが転送される。

【 0 0 6 7 】

図 1 0 は、端末 1 1 0 7 からブロードキャストパケットを受信した外部中継装置 1 1 0 6 が、LAG を構成するポートのうち中継装置 1 1 0 3 のメンバポートにパケットを転送した場合の転送経路を示す図である。

【 0 0 6 8 】

この場合、中継装置 1 1 0 3 は、図 5 に示すフローエントリに従って、自装置（中継装置 1 1 0 3 ）の E A S T スタックリンクポート（図 1 0 のポート # 1 ）からパケットを転送する。前記パケットを受信した中継装置 1 1 0 2 は、図 5 に示すフローエントリに従って、自装置（中継装置 1 1 0 2 ）の E A S T スタックリンクポート（図 1 0 のポート # 5 ）からパケットを転送する。前記パケットを受信した中継装置 1 1 0 1 は、図 4 に示すフローエントリに従って、ブロードキャスト転送経路の次ホップの中継装置である中継装置 1 1 0 4 に接続されているポートからブロードキャストパケットを転送する。その後は、図 8 と同様に、ブロードキャスト用の経路に従ってブロードキャストパケットが転送される。

【 0 0 6 9 】

図 1 1 は、中継装置 1 1 0 1 と中継装置 1 1 0 2 間のスタックリンクに障害が発生した場合において、端末 1 1 0 7 からブロードキャストパケットを受信した外部中継装置 1 1 0 6 が、LAG を構成するポートのうち中継装置 1 1 0 3 のメンバポートにパケットを転送した場合の転送経路を示す図である。

【 0 0 7 0 】

この場合、中継装置 1 1 0 3 は、図 5 に示すフローエントリに従って、自装置（中継装置 1 1 0 3 ）の E A S T スタックリンクポート（図 1 1 のポート # 1 ）からパケットを転送する。前記パケットを受信した中継装置 1 1 0 2 では、図 1 1 に示すように、E A S T スタックリンクポート（図 1 1 のポート # 5 ）に障害が生じているため、図 7 に示すフローエントリが設定されている。従って、中継装置 1 1 0 2 では、図 7 に示すフローエントリに従って、自装置（中継装置 1 1 0 2 ）の W E S T スタックリンクポート（図 1 1 のポート # 6 ）からパケットを折り返し転送する。前記パケットを受信した中継装置 1 1 0 3 は、図 5 に示すフローエントリに従って、自装置（中継装置 1 1 0 3 ）の W E S T スタックリンクポート（図 1 1 のポート # 2 ）からパケットを転送する。前記パケットを受信した中継装置 1 1 0 1 は、図 4 に示すフローエントリに従って、ブロードキャスト転送経路の次ホップの中継装置である中継装置 1 1 0 4 に接続されているポートからブロードキャ

10

20

30

40

50

ストパケットを転送する。その後は、図8と同様に、ブロードキャスト用の経路に従ってブロードキャストパケットが転送される。

【0071】

図12は、中継装置1101と中継装置1102間のスタックリンクに障害が発生した場合において、端末1107からブロードキャストパケットを受信した外部中継装置1106が、LAGを構成するポートのうち中継装置1102のメンバポートにパケットを転送した場合の転送経路を示す図である。

【0072】

中継装置1102では、図12に示すように、EASTスタックリンクポート(図12のポート#5)に障害が生じているため、図7に示すフローエントリが設定されている。従って、中継装置1102では、図7に示すフローエントリに従って、自装置(中継装置1102)のWESTスタックリンクポート(図11のポート#6)からパケットを転送する。前記パケットを受信した中継装置1103は、図5に示すフローエントリに従って、自装置(中継装置1103)のWESTスタックリンクポート(図12のポート#2)からパケットを転送する。前記パケットを受信した中継装置1101は、図4に示すフローエントリに従って、ブロードキャスト転送経路の次ホップの中継装置である中継装置1104に接続されているポートからブロードキャストパケットを転送する。その後は、図8と同様に、ブロードキャスト用の経路に従ってブロードキャストパケットが転送される。

【0073】

以上のように本実施形態によれば、図8～図10を用いて示したように、スタックリンクポートを用いた転送を行い、選出した代表ポートを有する中継装置以外から、ブロードキャストパケット等の制御対象パケットが出力されないように制御される。

【0074】

また、本実施形態によれば、図11～図12を用いて示したように、スタックリンクに障害が生じた場合でもループすることなくパケットを転送することができるになっている。

【0075】

以上本発明の第1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られず、種々の変形実施形態を採ることができる。以下、複数のトランクを有するいくつかの変形実施形態を説明する。

【0076】

[第2の実施形態]

図13は、本発明の第2の実施形態の構成を表した図である。本発明の第1の実施形態との相違点は、中継装置1201～1203と外部中継装置1209との間のトランクと、中継装置1205～1207と外部中継装置1210との間のトランクとの2つのトランクが構成されている点である。

【0077】

このような構成においても、トランク情報記憶部109のトランク構成情報を参照して、上記のように、中継装置1201～1203のメンバポートと、中継装置1205～1207のメンバポートとの中からそれぞれ1つの代表ポートを選出し、スタックリンクを介して各代表ポートからパケットが転送されるように制御する。これにより、選出した代表ポートを有する中継装置以外から、ブロードキャストパケット等の制御対象パケットが出力されないように制御することが可能となる。

【0078】

[第3の実施形態]

図14は、本発明の第3の実施形態の構成を表した図である。本発明の第1の実施形態との相違点は、中継装置1101～1103の別のポートを用いて、外部中継装置1106とは別のトランクを構成する外部中継装置1106aが追加されている点である。

【0079】

10

20

30

40

50

このような構成においても、トランク情報記憶部 109 のトランク構成情報を参照して、上記のように、中継装置 1101 ~ 1103 の外部中継装置 1106a と接続された第 2 のメンバポートから代表ポートを選出し、スタックリンクを介して代表ポートから、外部中継装置 1106a 経由の packets が転送されるように制御する。これにより、選出した代表ポートを有する中継装置以外から、ブロードキャスト packets 等の制御対象 packets が出力されないように制御することが可能となる。

【0080】

[第4の実施形態]

図 15 は、本発明の第 4 の実施形態の構成を表した図である。本発明の第 1 の実施形態との相違点は、中継装置 1101 ~ 1103 と外部中継装置 1106 との間に、第 2 のトランクが構成されている点である。なお、中継装置 1101 ~ 1103 の第 2 のトランクを構成するメンバポートは、第 1 のトランクを構成するメンバポートと同一のポートであってもよい。

10

【0081】

このような構成においても、トランク情報記憶部 109 のトランク構成情報を参照して、上記のように、中継装置 1101 ~ 1103 の第 2 のトランクを構成するメンバポートから代表ポートを選出し、スタックリンクを介して当該代表ポートから、外部中継装置 1106 経由の packets が転送されるように制御する。これにより、選出した代表ポートを有する中継装置以外から、ブロードキャスト packets 等の制御対象 packets が出力されないように制御することが可能となる。

20

【0082】

以上、本発明の各実施形態を説明したが、本発明は、上記した実施形態に限定されるものではなく、本発明の基本的技術的思想を逸脱しない範囲で、更なる変形・置換・調整を加えることができる。例えば、上記した各実施形態では、外部中継装置は、中継装置と端末との間に配置されているものとし説明したが、中継装置 1101 ~ 1105 の間に配置されていてもよい。

【0083】

また、上記した各実施形態では、ブロードキャスト packets の転送制御を行う例を挙げて説明したが、ブロードキャスト packets 以外の packets を制御対象 packets とすることもできる。例えば、マルチキャスト packets 等の外部中継装置から複数の中継装置に packets が転送される可能性のある packets に適用することができる。

30

【0084】

また、上記した各実施形態では、図中に実線と破線で示したように、データ用の通信路と、制御用の通信路とをそれぞれ設けられているものとして説明したが、これらは混在してもよい。

【0085】

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の枠内において種々の開示要素（各請求項の各要素、各実施例の各要素、各図面の各要素等を含む）の多様な組み合わせないし選択が可能である。すなわち、本発明は、請求の範囲を含む全開示、技術的思想にしたがって当業者であればなし得るであろう各種変形、修正を含むことは勿論である。

40

【符号の説明】

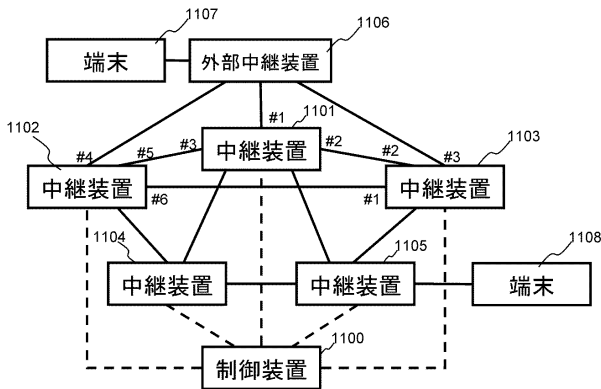
【0086】

- 101 中継装置通信部
- 102 トポロジ取得部
- 103 トポロジテーブル
- 104 ユニキャスト経路探索部
- 105 ユニキャスト用経路制御コマンド生成部

50

- 106 ブロードキャスト経路探索部 (BC経路探索部)
- 107 ブロードキャスト経路コマンド生成部 (BC用経路制御コマンド生成部)
- 108 トランク情報入力部
- 109 トランク情報記憶部
- 110 トランク制御コマンド生成部
- 1100 制御装置
- 1101 ~ 1105、1201 ~ 1208 中継装置
- 1106、1106a、1209、1210 外部中継装置
- 1107、1108、1211 ~ 1213 端末

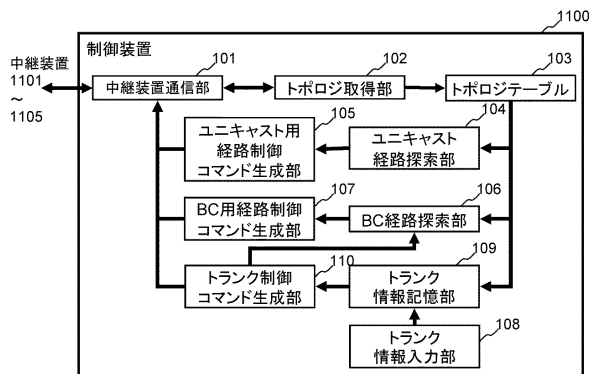
【図1】



【図3】

中継装置	メンバポート	EAST スタックリンク ポート	WEST スタックリンク ポート
1101	#1	#2	#3
1102	#4	#5	#6
1103	#3	#1	#2

【図2】



【図4】

受信ポート	処理内容(Instructions)
代表ポート	BC転送経路の次ホップの中継装置と接続しているポートから転送
EAST スタックリンクポート	BC転送経路の次ホップの中継装置と接続しているポートから転送
WEST スタックリンクポート	BC転送経路の次ホップの中継装置と接続しているポートから転送

【図5】

受信ポート	処理内容(Instructions)
メンバポート	EASTスタックリンクポートから転送
EAST スタックリンクポート	WESTスタックリンクポートから転送
WEST スタックリンクポート	EASTスタックリンクポートから転送

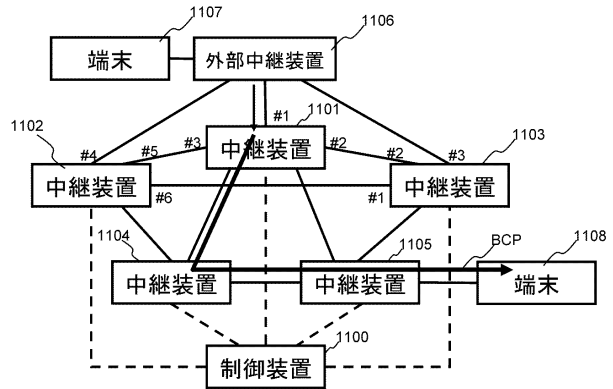
【図7】

受信ポート	処理内容(Instructions)
メンバポート	WESTスタックリンクポートから転送
EAST スタックリンクポート	廃棄
WEST スタックリンクポート	WESTスタックリンクポートから転送

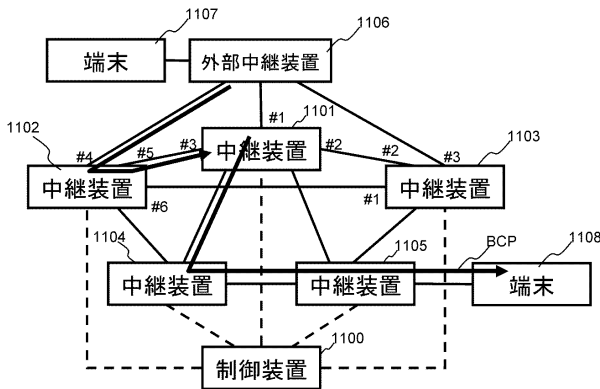
【図6】

受信ポート	処理内容(Instructions)
メンバポート	EASTスタックリンクポートから転送
EAST スタックリンクポート	廃棄
WEST スタックリンクポート	廃棄

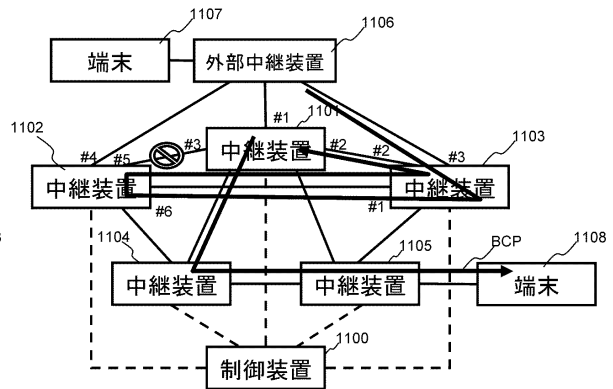
【図8】



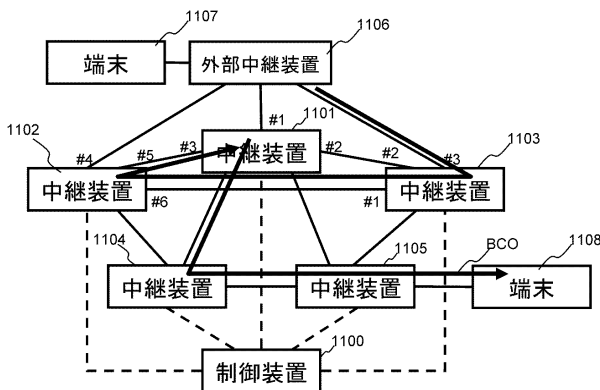
【図9】



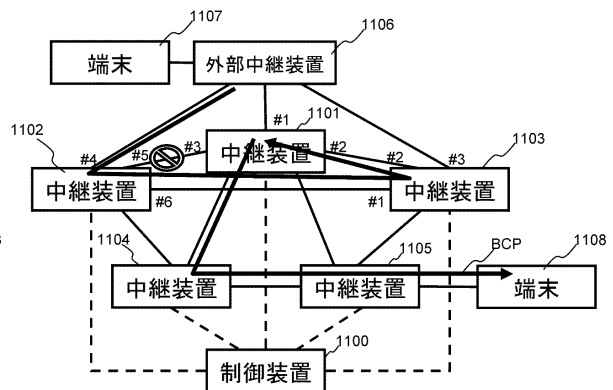
【図11】



【図10】



【図12】



フロントページの続き

(56)参考文献 特開2011-103592(JP, A)

ファウエイスリーコムが提案する次の5年を見越したネットワーク機器選び, NETWORK
MAGAZINE 第10巻 第6号, 株式会社アスキー, 2005年 6月 1日, p.2
0 - 23

根本 浩之, サーバーからネットワーク, ストレージまで ブームを迎えた「仮想化」って何だ
ろう?, 日経NETWORK 第101号, 日経BP社, 2008年 8月28日, p.052
- 059

(58)調査した分野(Int.Cl., DB名)

H04L 12/717

H04L 12/761