(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5654142号
(P5654142)
(45)発行日　平成27年1月14日(2015.1.14)　(24)登録日　平成26年11月28日(2014.11.28)

(51)Int.Cl.　F I
H04L 12/717 (2013.01)　H04L 12/717

請求項の数 12　(全 37 頁)

(21)出願番号　特願2013-544766(P2013-544766)
(86)(22)出願日　平成23年12月15日(2011.12.15)
(65)公表番号　特表2014-506409(P2014-506409A)
(43)公表日　平成26年3月13日(2014.3.13)
(86)国際出願番号　PCT/US2011/065077
(87)国際公開番号　WO2012/082988
(87)国際公開日　平成24年6月21日(2012.6.21)
審査請求日　平成25年8月12日(2013.8.12)
(31)優先権主張番号　12/971,924
(32)優先日　平成22年12月17日(2010.12.17)
(33)優先権主張国　米国(US)

(73)特許権者　513146011
ビッグ　スウィッチ　ネットワークス　インコーポレイテッド
アメリカ合衆国　カリフォルニア州　９４０４１　マウンテン　ヴィュー　ウェスト　イヴリン　アベニュー　１００　スイート　１１０
(74)代理人　100092093
弁理士　辻居　幸一
(74)代理人　100082005
弁理士　熊倉　禎男
(74)代理人　100067013
弁理士　大塚　文昭
(74)代理人　100086771
弁理士　西島　孝喜

最終頁に続く

(54)【発明の名称】ネットワーク・スイッチを構成するための方法

(57)【特許請求の範囲】
【請求項１】
コントローラ・サーバを用いて、ネットワーク内のネットワーク・スイッチにフロー・テーブル・エントリを与える方法であって、前記ネットワーク・スイッチの各々は、パケット・フィールドを前記フロー・テーブル・エントリのフィールドと比較することによりパケットを処理し、前記方法は、
前記コントローラ・サーバによって、前記ネットワーク・スイッチの幾つかをエッジ・ネットワーク・スイッチとしてカテゴリ化し、前記ネットワーク・スイッチの幾つかを非エッジ・スイッチとしてカテゴリ化するステップと、
前記コントローラ・サーバによって、第１のフロー・テーブル・エントリのセットを、エッジ・ネットワーク・スイッチとしてカテゴリ化された前記ネットワーク・スイッチに分散させ、かつ前記第１のフロー・テーブル・エントリのセットとは異なる第２のフロー・テーブル・エントリのセットを、非エッジ・スイッチとしてカテゴリ化された前記ネットワーク・スイッチに分散させるステップと、
を含み、
前記第１及び第２のフロー・テーブル・エントリを分散させるステップは、完全なフィールドのみを有するフロー・テーブル・エントリを前記エッジ・スイッチに分散させて、少なくとも幾つかのワイルドカード指定されたフィールドを有するフロー・テーブル・エントリを前記非エッジ・スイッチに分散させるステップを含むことを特徴とする方法。
【請求項２】

前記第１及び第２のフロー・テーブル・エントリを分散させるステップは、前記フロー・テーブル・エントリを、ネットワーク接続上で前記コントローラ・サーバから、前記ネットワーク・スイッチ上の対応するコントローラ・クライアントに分散させるステップを含み、

前記第１及び第２のフロー・テーブル・エントリを分散させるステップは、前記コントロール・サーバにおけるネットワーク・プロトコル・スタックを用いて、前記ネットワーク接続上で、前記コントローラ・クライアントにおける対応するネットワーク・プロトコル・スタックと通信するステップを含むことを特徴とする、請求項１に記載の方法。

【請求項３】

前記フロー・テーブル・エントリの各々は、送信元インターネット・プロトコル（ＩＰ）アドレス・フィールドを含み、前記エッジ・スイッチについての前記フロー・テーブル・エントリの各々の前記送信元ＩＰアドレス・フィールドにはワイルドカードがなく、前記非エッジ・スイッチについての前記フロー・テーブル・エントリの各々の前記送信元ＩＰアドレス・フィールドは、少なくとも幾つかのワイルドカード指定を含むことを特徴とする、請求項２に記載の方法。

【請求項４】

前記ネットワーク・スイッチはポートを含み、前記フロー・テーブル・エントリは、前記ネットワーク・スイッチがパケットを前記ポートのどれに転送すべきかを指定するアクション・フィールドを含むことを特徴とする、請求項１に記載の方法。

【請求項５】

デフォルトのインターネット・ゲートウェイによりインターネットに結合されたネットワークを動作させるための方法であって、前記ネットワークはエンド・ホストを前記デフォルト・インターネット・ゲートウェイに結合するネットワーク・スイッチを含み、前記ネットワーク・スイッチはコントローラによって制御される、前記方法は、

前記コントローラによって、ワイルドカード指定されていない物理ポートのエントリを含むフロー・テーブル・エントリを前記ネットワーク・スイッチ内のエッジ・スイッチに与えるステップと、

前記コントローラによって、少なくとも所与のフロー・テーブル・エントリを前記ネットワーク・スイッチ内の非エッジ・スイッチに与えるステップと、

を含み、前記非エッジ・スイッチは、前記デフォルトのインターネット・ゲートウェイに接続されており、前記所与のフロー・テーブル・エントリは、ワイルドカード指定された物理ポート・フィールドを含み、かつ、部分的にワイルドカード指定されて、前記非エッジ・スイッチに、パケットを前記デフォルトのインターネット・ゲートウェイに転送するように指示する宛先インターネット・プロトコルアドレス・フィールドを有することを特徴とする方法。

【請求項６】

前記ネットワーク・スイッチの各々は、前記コントローラと通信するコントローラ・クライアントを含み、前記コントローラはコントローラ・サーバを含み、前記フロー・テーブル・エントリを前記エッジ・スイッチに与えるステップは、前記コントローラ・サーバ及び前記コントローラ・クライアントにおけるネットワーク・プロトコル・スタックを用いて、ネットワーク接続上で前記フロー・テーブル・エントリを伝達することを含み、前記方法はさらに、

前記ネットワーク・スイッチの各々において、受け取ったパケット・フィールドを、そのネットワーク・スイッチに与えられた前記フロー・テーブル・エントリのフィールドと比較するステップをさらに含むことを特徴とする、請求項５に記載の方法。

【請求項７】

コントローラを用いて、ネットワーク内のネットワーク・スイッチにフロー・テーブル・エントリを与える方法であって、前記ネットワーク・スイッチの各々は、パケット・フィールドを前記フロー・テーブル・エントリのフィールドと比較し、一致が検出された場合には対応するアクションを取ることによってパケットを処理し、エンド・ホストは前記

ネットワーク・スイッチに接続されており、前記方法は、
前記コントローラによって、前記ネットワーク・スイッチの第１のものが、前記エンド・ホストに接続された入力・出力ポートを有するエッジ・スイッチであり、前記ネットワーク・スイッチの第２のものが前記ネットワーク・スイッチの前記第１のものに接続された入力・出力ポートを有する集約スイッチであると判断するステップと、
前記ネットワーク・スイッチの第１のものがエッジ・スイッチであり、前記ネットワーク・スイッチの前記第２のものが集約スイッチであるとの判断に応答して、前記コントローラを用いて、前記第１のネットワーク・スイッチに第１のフロー・テーブル・エントリを与え、前記第２のネットワーク・スイッチに第２のフロー・テーブル・エントリを与えるステップと、
を含み、前記第１のフロー・テーブル・エントリは、ワイルドカード指定のない物理ポート・フィールドを含み、前記第２のフロー・テーブル・エントリはワイルドカード指定を有する物理ポート・フィールドを含むことを特徴とする方法。
【請求項８】
前記第１のネットワーク・スイッチに前記第１のフロー・テーブル・エントリを与えるステップは、前記第１のネットワーク・スイッチに、ワイルドカード指定のないインターネット・プロトコル送信元アドレス・フィールドを含むフロー・テーブル・エントリを与えることを含み、
前記第２のネットワーク・スイッチに前記第２のフロー・テーブル・エントリを与えるステップは、前記第２のネットワーク・スイッチに、ワイルドカード指定されたインターネット・プロトコル送信元アドレス・フィールドを含むフロー・テーブル・エントリを与えるステップを含み、
前記第１のネットワーク・スイッチはコントローラ・クライアントを含み、前記第２のネットワーク・スイッチはコントローラ・クライアントを含み、前記コントローラは、ネットワーク・プロトコル・スタックを用いて、ネットワーク接続上で前記コントローラ・クライアントと通信することを特徴とする、請求項７に記載の方法。
【請求項９】
前記ネットワーク・スイッチの第３のものが、エンド・ホストに接続され、かつ、前記第２のスイッチに結合され、前記方法は、
前記第３のネットワーク・スイッチに、ワイルドカード指定されたインターネット・プロトコル送信元アドレス・フィールドと、ワイルドカード指定のない物理ポート・フィールドを含む第３のフロー・テーブル・エントリを与えるステップを含み、
前記第１、第２、及び第３のネットワーク・スイッチの前記フロー・テーブル・エントリに応答して、パケットを、前記第１、第２、及び第３のネットワーク・スイッチを通じて、前記第１のネットワーク・スイッチに接続された第１のエンド・ホストから、前記第３のネットワーク・スイッチに接続された第２のエンド・ホストに転送するステップをさらに含むことを特徴とする請求項８に記載の方法。
【請求項１０】
前記第１のネットワーク・スイッチに前記第１のフロー・テーブル・エントリを与えるステップは、前記第１のネットワーク・スイッチに、ワイルドカード指定のないイーサネット送信元アドレス・フィールドを含むフロー・テーブル・エントリを与えるステップを含むことを特徴とする、請求項７に記載の方法。
【請求項１１】
前記第１のネットワーク・スイッチに前記第１のフロー・テーブル・エントリを与えるステップは、前記第１のネットワーク・スイッチに、ワイルドカード指定のない仮想ローカル・エリア・ネットワーク・タグを含むフロー・テーブル・エントリを与えるステップを含むことを特徴とする、請求項７に記載の方法。
【請求項１２】
前記第１のネットワーク・スイッチに前記第１のフロー・テーブル・エントリを与えるステップは、前記第１のネットワーク・スイッチに、ワイルドカード指定のないインター

ネット・プロトコル送信元アドレス、ワイルドカード指定のないイーサネット送信元アドレス、及びワイルドカード指定のない仮想ローカル・エリア・ネットワーク・タグからなる群から選択されるワイルドカード指定のないアドレスを含むフロー・テーブル・エントリを与えるステップを含み、

前記第２のネットワーク・スイッチに前記第２のフロー・テーブル・エントリを与えるステップは、前記第２のネットワーク・スイッチに、所与のアドレスを含むフロー・テーブル・エントリを与えるステップと、パケット・フィールドが前記所与のアドレスと一致することを検出したことに応答して、前記第２のネットワーク・スイッチに前記第２のネットワーク・スイッチが取るアクションを与えるステップを含むことを特徴とする、請求項７に記載の方法。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、通信ネットワークに関し、より特定的には、通信ネットワークにおけるスイッチの構成に関する。

本出願は、全体が引用により本明細書に組み入れられる２０１０年１２月１７日に出願された米国特許出願第１２／９７１，９２４号に基づく優先権を主張するものである。

【背景技術】

【０００２】

インターネット及びインターネットに接続されたローカル・データ・ネットワークのようなパケット・ベースのネットワークは、ネットワーク・スイッチを含む。ネットワーク・スイッチは、パケット送信元（ｐａｃｋｅｔ　ｓｏｕｒｃｅ）からパケット宛先（ｐａｃｋｅｔ　ｄｅｓｔｉｎａｔｉｏｎ）へパケットを転送するために用いられる。

【０００３】

１つのベンダーのスイッチを別のベンダーの機器を用いて制御するのは困難であるか又は不可能である場合がある。これは、１つのベンダーのスイッチ機器が、別のベンダーのスイッチ機器とは異なるオペレーティング・システム及び制御手順の組を用いることがあるためである。異なるタイプのスイッチ・プラットフォームの制御に関連した問題に対処するために、クロスプラットフォーム（ｃｒｏｓｓ－ｐｌａｔｆｏｒｍ）のプロトコルが開発された。これらのプロトコルは、他の場合には互換性のないスイッチの集中制御を可能にする。

【０００４】

クロスプラットフォームのコントローラ・クライアントを、ネットワーク内のスイッチ上に含ませることができる。コントローラ・クライアントは、ネットワーク経路上で対応するコントローラ・サーバと通信することができる。コントローラ・クライアントは様々なスイッチ・ハードウェア上に実装できるので、単一のコントローラが、他の場合には互換性のないことがあるスイッチ機器を制御することが可能である。

【０００５】

コントローラ・クライアントが実装されたネットワーク・スイッチの各々は、そのスイッチによってパケットがどのように転送されるかを指定するエントリを有するフロー・テーブルを含むことができる。気を付けなければ、このタイプの構成を実装するために必要なフロー・テーブル・エントリの数が、ネットワーク内のスイッチの一部の能力を超えることがある。このタイプのクロスプラットフォームのネットワーク・スイッチに基づく隔離されたネットワーク・ドメイン間でトラフィックを伝達する際にも問題が生じ得る。

【発明の概要】

【発明が解決しようとする課題】

【０００６】

従って、ネットワーク・スイッチを動作させるための改良された構成を提供できることが望ましい。

【課題を解決するための手段】

【０００７】
フロー・テーブルを用いて、ネットワーク・スイッチを構成することができる。フロー・テーブル・エントリは、ヘッダ・フィールドと、関連したアクションとを含むことができる。パケットがネットワーク・スイッチにより受信されると、ネットワーク・スイッチは、パケット・フィールドをフロー・テーブル・エントリのフィールドと比較することができる。ネットワーク・スイッチは、一致が検出された場合に適切なアクションを取ることができる。例えば、ネットワーク・スイッチは、パケットを適切なスイッチ・ポートに転送することができる。

【０００８】
コントローラ・サーバを用いて、ネットワーク・スイッチを制御することができる。ネットワーク・スイッチの各々は、コントローラ・クライアントを含むことができる。コントローラ・サーバ及びコントローラ・クライアントは、ネットワーク・プロトコル・スタックを用いて、ネットワーク接続上で通信することができる。例えば、コントローラ・サーバは、ネットワーク・スイッチに、所望のパケット処理動作を行うように指示するフロー・テーブル・エントリを、コントローラ・クライアントに分散させることができる。

【０００９】
コントローラ・サーバは、ネットワークのトポロジーを判断し、かつ、ネットワーク・スイッチの容量及び他のネットワーク・スイッチの能力についての情報を集めることができる。この情報に基づいて、コントローラ・サーバは、スイッチに、パケットを、ネットワークを通じて所望の経路に沿って転送するように指示する、ネットワーク・スイッチについてのフロー・テーブルを生成することができる。ネットワーク・コアに近いスイッチについてのフロー・テーブル・エントリには、ネットワーク・エッジに近いスイッチより多くのワイルドカード指定（ｗｉｌｄｃａｒｄｉｎｇ）を与え、ネットワーク・コアに近いスイッチの容量に追いつかなくなることを回避することができる。ネットワーク・エッジにおけるスイッチについてのフロー・テーブル・エントリが、ネットワーク・コアに近いスイッチより少ないワイルドカード指定を有することを保証することによって、ワイルドカード指定の存在下で、パケット転送機能を保つことができる。

【００１０】
幾つかのネットワークにおいては、コア・スイッチのようなスイッチが、コントローラ・サーバが生成したフロー・テーブル・エントリによって構成されていないローカル・コア・ネットワークを形成することができる。ローカル・コア・ネットワークは、フロー・テーブルを用いて構成されるスイッチのネットワークにおけるドメインの間に置くことができる。従って、ローカル・コア・ネットワークは、ネットワーク・ドメインを互いに隔離することができる。

【００１１】
このタイプのネットワークにおいては、ローカル・コア・ネットワークを通じて、トラフィックを、１つの隔離ドメインから他方の隔離ドメインに通すことができる。トラフィックは、パケットをカプセル化する１つのドメインにおけるカプセル化（ｅｎｃａｐｓｕｌａｔｉｏｎ）エンジンと、パケットをカプセル化解除する他方のドメインにおける対応するカプセル化解除（ｄｅｅｎｃａｐｓｕｌａｔｉｏｎ）エンジンとを用いて通すことができる。カプセル化エンジン及びカプセル化解除エンジンを実装するのに用いられるネットワーク・スイッチは、隔離されたネットワーク・ドメイン内の潜在的に任意の位置におけるネットワーク・スイッチ上に配置することができる。コントローラ・サーバは、これらスイッチの位置を発見し、ネットワーク・スイッチに、カプセル化エンジン及びカプセル化解除エンジンを含む適切な経路に沿ってトラフィックを転送するように指示するフロー・テーブル・エントリを生成することができる。

【００１２】
本発明のさらなる特徴、その性質及び種々の利点は、添付図面及び以下の好ましい実施形態の詳細な説明からより明らかになるであろう。

【図面の簡単な説明】

【００１３】
【図１】本発明の実施形態による、コントローラ及びパケット転送システムを含む例証となるネットワークの図である。
【図２】本発明の実施形態による、パケット処理エンジンを実行するマイクロプロセッサ・ベースの機器を用いて、パケット転送システムをどのように実装できるかを示す図である。
【図３】本発明の実施形態による、仮想スイッチを用いてパケット転送システムの機能をどのように実行できるかを示す図である。
【図４】本発明の実施形態による、パケット転送システムが制御ユニット及び関連したスイッチング集積回路を含む、パケット転送システム及び関連したコントローラの図である。
【図５】本発明の実施形態による、パケット転送システムがマスター・コントローラ及びスレーブ・コントローラを有し、かつ、コントロール・サーバを遠隔コンピューティング機器上又はパケット転送システムにおけるライン・カード上に実装できる、ネットワークの図である。
【図６】本発明の実施形態による、ネットワーク接続上で通信するコントローラ・サーバ及びコントローラ・クライアントの図である。
【図７Ａ】本発明の実施形態による、パケット処理システムにより用いることができるタイプの例証となるフロー・テーブルの図である。
【図７Ｂ】本発明の実施形態による、フロー・テーブルのフロー・テーブル・エントリに基づいて実行することができる３つの例証となるタイプのパケット転送を示す、パケット処理システムにより用いることができるタイプの例証となるフロー・テーブルの図である。
【図７Ｃ】本発明の実施形態による、特定アドレスを有するパケットがスイッチ内の第３の物理ポートに転送される、例証となるフロー・テーブルの図である。
【図７Ｄ】本発明の実施形態による、特定アドレスを有するパケットがスイッチ内の第４の物理ポートに転送される、例証となるフロー・テーブルの図である。
【図８】本発明の実施形態による、パケット処理システムにおいてパケットを処理する際に必要とされる例証となるステップのフローチャートである。
【図９】本発明の実施形態による、コントローラが、複数のネットワーク・スイッチをどのように制御できるかを示すネットワークの図である。
【図１０】本発明の実施形態による、ネットワークの異なる部分を通して、スイッチをどのように分散できるかを示す例証となるネットワークの図である。
【図１１】本発明の実施形態による、様々な量のワイルドカード指定を有する例証となるフロー・テーブル・エントリのテーブルである。
【図１２】本発明の実施形態による、送信元パケット・エンド・ホストと宛先パケット・エンド・ホストとの間の経路に沿って、データをどのように伝達できるかを示す例証となるネットワークの図である。
【図１３】図１２の経路に沿ってパケットを転送する際に用いることができる従来のフロー・テーブル・エントリのテーブルである。
【図１４】図１３の経路に沿ってパケットを転送する際に用いることができる従来のフロー・テーブル・エントリの別のテーブルである。
【図１５】本発明の実施形態による、図１２の経路に沿ってパケットを転送する際に用いることができるフロー・テーブル・エントリの例証となるテーブルである。
【図１６】本発明の実施形態による、図１２の経路に沿ってパケットを転送する際に用いることができるフロー・テーブル・エントリの別の例証となるテーブルである。
【図１７】本発明の実施形態による、図１５及び図１６に示されるタイプのフロー・テーブル・エントリを用いて、図１２に示されるタイプのネットワークを動作させる際に用いることができる例証となるステップのフローチャートである。
【図１８】本発明の実施形態による、ネットワークにおけるスイッチのトポロジー及びス

イッチ容量を判断する際に必要とされる例証となるステップのフローチャートである。
【図１９】本発明の実施形態による、デフォルト・ゲートウェイを用いてインターネットに結合されたネットワークの図である。
【図２０】本発明の実施形態による、図１９に示されるタイプのネットワークにおいてパケットを転送する際に用いることができる例証となるエントリを含むフロー・テーブルである。
【図２１】本発明の実施形態による、隔離されたネットワーク・ドメインに接続するように、ローカル・コア・ネットワークを通るトンネルを形成する際に用いることができる、パケット・カプセル化エンジン及びパケット・カプセル化解除エンジンを含むネットワークの図である。
【図２２】本発明の実施形態による、図２１に示されるタイプのネットワークを動作させる際に必要とされる例証となるステップのフローチャートである。
【図２３】本発明の実施形態による、ネットワークにおいて用いることができる、例証となるエッジ・スイッチのフロー・テーブル・エントリ及び非エッジ・スイッチのフロー・テーブル・エントリの図である。
【発明を実施するための形態】
【００１４】
インターネット、並びにインターネットに結合されたローカル及び地域ネットワークのようなネットワークは、パケット・ベースのスイッチに依存する。本明細書ではネットワーク・スイッチ、パケット処理システム、又はパケット転送システムと呼ばれることがあるこれらのスイッチは、アドレス情報に基づいてパケットを転送することができる。このように、パケット送信元により伝送されるデータ・パケットをパケット宛先に配信することができる。ネットワーク用語では、パケットの送信元（ｓｏｕｒｃｅ）及び宛先（ｄｅｓｔｉｎａｔｉｏｎ）は、エンド・ホストと呼ばれることがある。エンド・ホストの例は、パーソナル・コンピュータ、サーバ、及び他のコンピューティング機器である。
【００１５】
ネットワーク・スイッチの能力は、比較的小型のイーサネット・スイッチ及び無線アクセス・ポイントから、複数のライン・カード、冗長電源、及びスーパバイザ能力を含む大型のラック・ベースのシステムまでの範囲に及ぶ。イーサネット・スイッチは、ネットワークのエッジ付近で用いられることがあるので、エッジ・スイッチ又はトップ・オブ・ラック・スイッチと呼ばれることがある。より大型のラック・ベースのシステムは、多くの場合、ネットワーク・コア位置で用いられ、かつ、ルータ、コア・ルータ、又はコア・スイッチと呼ばれることがある。幾つかのネットワーク環境においては、コア・スイッチとエッジ・スイッチとの間にあるネットワーク・スイッチは、集約（ａｇｇｒｅｇａｔｉｏｎ）スイッチ又は分散スイッチと呼ばれる。集約スイッチ及びコア・スイッチは、まとめて非エッジ・スイッチと呼ばれることもある。
【００１６】
ネットワークが、複数のベンダーからの機器を含むのは珍しいことではない。一例として、大学又は会社構内のためのネットワークが、１つのベンダーからのコア・スイッチと、別のベンダーからのエッジ・スイッチと、さらに別のベンダーからの集約スイッチとを含むことがある。異なるベンダーからのネットワーク・スイッチを相互接続して、パケット転送ネットワークを形成することができるが、それらのオペレーティング・システムと制御プロトコルとの間の非互換性のために、集中方式で管理するのは困難であり得る。
【００１７】
これらの潜在的な非互換性は、共通のクロスプラットフォーム制御モジュール（本明細書ではコントローラ・クライアントと呼ばれることがある）を各々のネットワーク・スイッチに組み込むことによって克服することができる。集中型クロスプラットフォーム・コントローラ・サーバは、それぞれのネットワーク・リンク上で、制御クライアントの各々と対話することができる。クロスプラットフォーム・コントローラ・サーバ、及び対応するコントローラ・クライアントを用いることにより、潜在的に異種のネットワーク・スイ

ッチ機器の中央管理が可能になる。

【００１８】

本明細書では一例として説明されることがある１つの例証となる構成の場合、集中制御は、図１のコントローラ・サーバ１８のような１つ又はそれ以上のコントローラ・サーバにより提供される。コントローラ・サーバ１８は、スタンド・アロン・コンピュータ上、コンピュータのクラスタ上、複数の位置の間に分散された一組のコンピュータ上、ネットワーク・スイッチ内に埋め込まれたハードウェア上、又は他の適切なコンピューティング機器１２上に実装することができる。コントローラ・サーバ１０は、単一のコンピュータ上の単一のプロセスとして実行することができ、又は、冗長性のために幾つかのホストにわたって分散させることができる。分散型構成の使用は、予期しないネットワーク分割（例えば、２つの構内間のネットワーク・リンクが中断される状況）に対して、ネットワーク１０に弾力性を与える助けになり得る。

【００１９】

分散型コントローラ構成において、コントローラ・ノードは、コントローラ内のプロトコルを用いて、情報を交換することができる。例えば、新しいエンド・ホストが、第１のコントローラ・ノードにのみ接続されたネットワーク・ハードウェア（例えば、スイッチ）に接続される場合、その第１のコントローラ・ノードは、コントローラ内のプロトコルを用いて、他のコントローラ・ノードに、新しいエンド・ホストの存在を知らせることができる。所望であれば、スイッチ又は他のネットワーク・コンポーネントを複数のコントローラ・ノードに接続することができる。単一のコントローラ・サーバを用いて、関連したスイッチのネットワークを制御する構成が、本明細書において一例として説明されることがある。

【００２０】

図１のコントローラ・サーバ１８は、ネットワーク１０のトポロジーについての情報を収集することができる。例えば、コントローラ・サーバ１８は、ネットワークを通じて、リンク層検出プロトコル（Ｌｉｎｋ　Ｌａｙｅｒ　Ｄｉｓｃｏｖｅｒｙ　Ｐｒｏｔｏｃｏｌ：ＬＬＤＰ）プローブ・パケットを送信して、ネットワーク１０のトポロジーを見いだすことができる。コントローラ・サーバ１８は、ネットワーク・トポロジーに関する情報、及びネットワーク機器の能力に関する情報を用いて、ネットワークを通って流れるパケットのための適切な経路を判断することができる。ひとたび適切な経路が識別されると、コントローラ・サーバ１８は、対応する設定データをネットワーク１０内のハードウェアに送信し、パケットが所望の通りにネットワークを通って流れることを保証することができる。これらのようなネットワーク構成動作は、システム設定動作中バックグラウンドで連続的に、又は、新たに伝送されたデータ・パケット（すなわち、既存の経路が確立されなかったパケット）の出現に応答して、実行することができる。

【００２１】

コントローラ・サーバ１８を用いて、ネットワーク構成規則２０を実装することができる。規則２０は、種々のネットワーク・エンティティがどのサービスを利用可能であるかを指定することができる。一例として、規則２０は、ネットワーク２０におけるどのユーザ（又はどのタイプのユーザ）が特定のサーバにアクセスできるかを指定することができる。規則２０は、例えば、コンピューティング機器１２のデータベース内に維持することができる。

【００２２】

コントローラ・サーバ１８、及びそれぞれのネットワーク・スイッチ１４におけるコントローラ・クライアント３０は、ネットワーク・プロトコル・スタックを用いて、ネットワーク・リンク１６上で通信することができる。

【００２３】

各々のスイッチ（パケット転送システム）１４は、入力・出力ポート３４を有することができる。ケーブルを用いて、機器の部品をポート３４に接続することができる。例えば、パーソナル・コンピュータ、ウェブ・サーバ、及び他のコンピューティング機器などの

エンド・ホストを、ポート３４にプラグ接続することができる。ポート３４を用いて、スイッチ１４の１つを他のスイッチ１４に接続することもできる。

【００２４】

ポート３４の１つからポート３４の別のものにパケットを転送する際、及び着信パケットに対して他の適切なアクションを行う際に、パケット処理回路３２を用いることができる。パケット処理回路３２は、専用高速スイッチ回路のような１つ又はそれ以上の集積回路を用いて実装することができ、かつ、ハードウェアのデータ経路として働くことができる。所望であれば、ソフトウェアのデータ経路を実装する際に、制御ユニット２４上で実行されているパケット処理ソフトウェア２６を用いることができる。

【００２５】

制御ユニット２４は、制御ソフトウェアを格納及び実行するための処理及びメモリ回路（例えば、１つ又はそれ以上のマイクロプロセッサ、メモリ・チップ、及び他の制御回路）を含むことができる。例えば、制御ユニット２４は、パケット処理ソフトウェア２６のようなソフトウェアを格納及び実行することができ、フロー・テーブル２８を格納することができ、コントローラ・クライアント３０の動作をサポートするために用いることができる。

【００２６】

コントローラ・クライアント３０及びコントローラ・サーバ１８は、ＯｐｅｎＦｌｏｗプロトコル（例えば、ＯｐｅｎＦｌｏｗスイッチ仕様バージョン１．０．０を参照されたい）などのネットワーク・スイッチのプロトコルに準拠することができる。コントローラ・クライアント３０の中の１つ又はそれ以上のクライアントは、他のプロトコル（例えば、簡易ネットワーク管理プロトコル）に準拠することもできる。ＯｐｅｎＦｌｏｗプロトコル又は他の適切なプロトコルを用いて、コントローラ・サーバ１８は、コントローラ・クライアント３０に、スイッチ１４が、入力・出力ポート３４からの着信パケットをどのように処理するかを決定するデータを与えることができる。

【００２７】

１つの適切な構成の場合、コントローラ・サーバ１８からのフロー・テーブル・データを、フロー・テーブル２８のようなフロー・テーブル内に格納することができる。スイッチ１４を構成する際に、フロー・テーブル２８のエントリを用いることができる（例えば、パケット処理回路３２及び／又はパケット処理ソフトウェア２６の機能）。典型的なシナリオにおいては、フロー・テーブル２８は、フロー・テーブル・エントリのキャッシュ記憶装置として働き、これらのフロー・テーブル・エントリの対応するバージョンが、パケット処理回路３２の回路が維持する設定内に埋め込まれる。しかしながら、これは例証にすぎない。フロー・テーブル２８は、スイッチ１４内のフロー・テーブル・エントリのための排他的な記憶装置として働くことができ、又は、パケット処理回路３２内のフロー・テーブルの記憶装置リソースを優先して、省くことができる。一般に、フロー・テーブル・エントリは、いずれかの適切なデータ構造体（例えば、１つ又はそれ以上のテーブル、リスト等）を用いて格納することができる。明確にするために、フロー・テーブル２８のデータ（制御ユニット２４のデータベース内に維持されたものでも、又はパケット処理回路３２の構成内に埋め込まれたものでも）は、本明細書では、フロー・テーブル・エントリを形成するもの（例えば、フロー・テーブル２８における行）と呼ばれる。

【００２８】

所望であれば、スイッチ１４は、制御ソフトウェアを実行し、パケット処理回路３２を省いた、図２の汎用処理プラットフォームを用いて実装することができる。このタイプの構成が図２に示される。図２の例証となる構成に示されるように、コンピューティング機器１２上のコントローラ・サーバ１８は、ネットワーク・リンク１６上で、スイッチ（パケット転送システム）１４上のコントローラ・クライアント３０と通信することができる。コントローラ・サーバ１８は、例えば、フロー・テーブル２８内に維持されるフロー・テーブル・エントリをコントローラ・クライアント３０に伝送することができる。パケット処理ソフトウェア４０は、ネットワーク・インターフェース３８を用いて、パケット（

例えば、ポート３４を用いて送受信されるパケット）を転送し、他の方法で処理することができる。ネットワーク・インターフェース３８は、スイッチ１４内のシステム・ボード（一例として）にプラグ接続された１つ又はそれ以上のネットワーク・インターフェースを用いて実装することができる。

【００２９】

別の例示的なタイプのネットワーク・スイッチが、図３に示される。図３の例においては、コンピューティング機器４２は、仮想マシン４４を実装するために用いられる。コンピューティング機器４２は、例えば、１つ又はそれ以上のコンピュータに基づくサーバとすることができ、仮想マシン４４は、ウェブ・サーバ又は他のオンライン・サービスを実装するために用いることができる。典型的なシナリオにおいては、仮想マシンサービスを購入した顧客に、仮想マシン４４の番号が割り当てられることがある。これらの仮想マシンが互いに通信できることを保証するために、コンピューティング機器４２のリソースの一部を用いて、ネットワーク・スイッチ１４（例えば、パケット処理ソフトウェア４０のようなソフトウェア、フロー・テーブル２８、及びコントローラ・クライアント３０に基づいたパケット処理システム）を実装する。仮想スイッチと呼ばれることもあるスイッチ１４は、それぞれの仮想マシン４４間でパケットを転送できる１つのタイプのパケット転送システムを形成する。

【００３０】

図１のネットワーク・スイッチ１４のようなネットワーク・スイッチは、１つ又はそれ以上の高速スイッチング集積回路（「スイッチＩＣ」）に結合された制御回路を用いて実装することができる。このタイプの構成が、図４に示される。図４に示されるように、コンピューティング機器１２上のコントローラ・サーバ１８は、経路１６を介してネットワーク・スイッチ１４と通信することができる。スイッチ１４は、処理回路２４と、スイッチＩＣ３２－１．．．スイッチＩＣ３２－Ｎのような１つ又はそれ以上の関連したスイッチＩＣ３２とを含むことができる。制御回路２４は、例えば、マイクロプロセッサ及びメモリに基づくことができる。スイッチＩＣ３２－１．．．ＩＣ３２－Ｎは、高速でパケット処理タスクを処理することができる専用スイッチング回路とすることができる。一例として、制御回路２４は、５００ＭＨｚのマイクロプロセッサに基づくことができ、スイッチＩＣ３２－１．．．ＩＣ３２－Ｎは、入力・出力ポート３４の４８からのデータを処理することができ、その各々は、（一例として）１－１０Ｇｂｐｓの関連したデータ転送速度を有する。

【００３１】

図１のネットワーク・スイッチ１４を実装する際に用いることができる別の例証としてのスイッチ・アーキテクチャが、図５に示される。図５の例において、スイッチ（パケット転送システム）１４は、プロセッサ２４－１のようなマスター・プロセッサと、スレーブ・プロセッサ２４－２のような１つ又はそれ以上の関連したスレーブ・プロセッサとを含むことができる。スイッチＩＣ３２及びプロセッサ２４－２のようなスレーブ・プロセッサは、ライン・カード４８のようなライン・カード上に実装することができる。ライン・カード５０のような１つ又はそれ以上のライン・カードは、処理回路（例えば、マイクロプロセッサ及びメモリ）を含むことができる。ライン・カード４８及び５０は、バックプレーン５２を用いて相互接続することができる。

【００３２】

図５に示されるタイプの構成の場合、コントローラ・サーバは、ライン・カードの処理リソースを用いて実装することができる。例えば、コントローラ・サーバは、図５のコントローラ・サーバ１８－Ｂにより示されるように、ライン・カード５０上に実装することができる。所望であれば、コントローラ・サーバは、コンピューティング機器１２上に実装することができる（例えば、図５のコントローラ・サーバ１８－Ａのように）。コントローラ・サーバ１８－Ａ又はコントローラ・サーバ１８－Ｂは、プロセッサ２４－１及び／又は２４－２のようなプロセッサを用いて実装されるコントローラ・クライアント３０と通信することができる。コントローラ・サーバ１８－Ａとコントローラ・クライアント

との間の通信は、ネットワーク接続１６上で行われ得る。コントローラ・サーバ１８－Ｂとコントローラ・クライアントとの間の通信は、バックプレーン５２上で（例えば、ＴＣＰ／ＩＰのようなプロトコルを用いたネットワーク接続上で）行われ得る。

【００３３】

図６に示されるように、コントローラ・サーバ１８及びコントローラ・クライアント３０は、ネットワーク・プロトコル・スタック５８及びネットワーク・プロトコル・スタック６０のようなネットワーク・プロトコル・スタックを用いて、ネットワーク経路６６上で通信することができる。スタック５８及び６０は、例えば、Ｌｉｎｕｘ　ＴＣＰ／ＩＰスタック又はＶｘＷｏｒｋｓオペレーティング・システムにおけるＴＣＰ／ＩＰスタックとすることができる（例として）。経路６６は、例えば、スイッチ１４と外部機器との間のネットワーク接続をサポートする経路（例えば、図１のネットワーク経路１６）とすることができ、又は図５に示されるような、スイッチ１４のバックプレーン５２におけるネットワーク接続をサポートする経路とすることができる。経路６６が経路１６のようなネットワーク経路である構成が、一例として本明細書で説明されることもある。

【００３４】

制御プロトコル・スタック５６は、ネットワーク・プロトコル・スタック５８と制御ソフトウェア５４との間のインターフェースとして働く。制御プロトコル・スタック６２は、ネットワーク・プロトコル・スタック６０と制御ソフトウェア６４との間のインターフェースとして働く。動作中、コントローラ・サーバ１８がコントローラ・クライアント３０と通信するとき、制御プロトコル・スタック５６は、制御プロトコル・メッセージ（例えば、ポートを作動させ又は特定のフロー・テーブル・エントリをフロー・テーブル２８にインストールするための制御メッセージ）を生成し、構文解析する。図６に示されるタイプの構成を用いることにより、コントローラ・サーバ１８とコントローラ・クライアント３０との間のリンク上に、ネットワーク接続が形成される。コントローラ・サーバ１８とコントローラ・クライアント３０は、インターネット・プロトコル（ＩＰ）ネットワーク接続上で、伝送制御プロトコル（Ｔｒａｎｓｍｉｓｓｉｏｎ　Ｃｏｎｔｒｏｌ　Ｐｒｏｔｏｃｏｌ、ＴＣＰ）又はユーザ・データグラム・プロトコル（Ｕｓｅｒ　Ｄａｔａｇｒａｍ　Ｐｒｏｔｏｃｏｌ、ＵＤＰ）を用いて通信することができる。ネットワーク接続上でコントローラ・サーバ１８とコントローラ・クライアント３０との間で通信するときに用いることができる制御プロトコルの例として、ＳＮＭＰ及びＯｐｅｎＦｌｏｗプロトコル・スタック・バージョン１．０．０が挙げられる（例として）。

【００３５】

フロー・テーブル２８は、複数のフィールド（ヘッダ・フィールドと呼ばれることがある）を有するフロー・テーブル・エントリ（例えば、テーブルにおける行）を含む。スイッチ１４により受信されたパケット・フィールドを、フロー・テーブルのフィールドと比較することができる。各々のフロー・テーブル・エントリは、関連したアクションを有することができる。パケット・フィールドと、フロー・テーブル・エントリのフィールドとの間に一致がある場合、そのフロー・テーブル・エントリについての対応するアクションをとることができる。

【００３６】

例証となるフロー・テーブルが、図７に示される。図７Ａに示されるように、テーブル２８は、フロー・テーブル・エントリ（行）６８を有することができる。各々のフロー・テーブル・エントリは、ヘッダ７０、アクション７２、及び統計データ７４と関連付けることができる。ヘッダ７０は各々、複数のヘッダ・フィールド７６を含むことができる。各々のフロー・テーブル・エントリにおけるアクションは、パケット・フィールドと、そのフロー・テーブル・エントリのヘッダの対応するフィールドとの間に一致が検出されたときに、スイッチ１４がパケットに対してどのアクションを行うべきかを示す。スイッチ１４は、統計データ（カウンタ値）をフロー・テーブル２８の統計データ部分内に維持することができ、この統計データには、スイッチ１４の性能についての情報を得ることが望まれる場合、コントローラ・サーバ１８により照会することができる。

【００３７】
ヘッダ７０のヘッダ・フィールド（及び、各着信パケットの対応するフィールド）は、以下のフィールド、すなわち入力ポート（すなわち、パケットが受信される、スイッチ１４の物理ポートのアイデンティティ）、イーサネット送信元アドレス、イーサネット宛先アドレス、イーサネットのタイプ、仮想ローカル・エリア・ネットワーク（ＶＬＡＮ）ｉｄ、ＶＬＡＮ優先順位、ＩＰ送信元アドレス、ＩＰ宛先アドレス、ＩＰプロトコル、ＩＰＴｏＳ（サービスのタイプ）ビット、トランスポート送信元ポート／インターネット制御メッセージ・プロトコル（ＩＣＭＰ）タイプ（送信元ＴＣＰポートと呼ばれることもある）、及びトランスポート宛先ポート／ＩＣＭＰコード（宛先ＴＣＰポートと呼ばれることがある）を含むことができる。所望であれば、他のフィールドを用いることができる。
【００３８】
各々のフロー・テーブル・エントリ（フロー・エントリ）は、スイッチが一致するパケットをどのように処理するかを命令する、ゼロ又はそれ以上のアクションと関連付けられる。転送アクションが存在しない場合、パケットはドロップされることが好ましい。パケット・フィールドと、フロー・テーブル・エントリのヘッダ・フィールドとの間に一致が検出された場合、スイッチ１４が取り得るアクションは、以下のアクション、すなわち、転送（例えば、着信インターフェースを含まない全てのインターフェース上でパケットを送信するためのＡＬＬ、パケットをカプセル化し、コントローラ・サーバに送信するためのＣＯＮＴＲＯＬＬＥＲ、パケットをスイッチのローカル・ネットワーキング・スタックに送信するためのＬＯＣＡＬ、フロー・テーブル２８におけるアクションを行うためのＴＡＢＬＥ、パケットを入力ポートの外に送信するためのＩＮ＿ＰＯＲＴ、例えば、伝統的なレベル２、ＶＬＡＮ、及びレベル３の処理を用いて、スイッチによりサポートされるデフォルト転送経路によってパケットを処理するためのＮＯＲＭＡＬ、及び着信インターフェースを含まない最小スパニング・ツリーに沿ってパケットをフラッドさせるためのＦＬＯＯＤ）を含むことができる。スイッチ１４が取ることができる付加的なアクションは、ポートに取り付けられたキューを通してパケットを転送するためのエンキュー・アクション、及びドロップ・アクション（例えば、指定されたアクションなしに、フロー・テーブル・エントリと一致するパケットをドロップするための）を含む。（フィールド修正（Ｍｏｄｉｆｙ－ｆｉｅｌｄ）アクションをスイッチ１４によりサポートすることもできる。取り得るフィールド修正アクションの例は、ＶＬＡＮ ＩＤの設定（Ｓｅｔ ＶＬＡＮ ＩＤ）、ＶＬＡＮ優先順位の設定（Ｓｅｔ ＶＬＡＮ ｐｒｉｏｒｉｔｙ）、ＶＬＡＮヘッダの除去（Ｓｔｒｉｐ ＶＬＡＮ ｈｅａｄｅｒ）、イーサネット送信元アドレスＭＡＣ（媒体アクセス制御）アドレスの修正（Ｍｏｄｉｆｙ Ｅｔｈｅｒｎｅｔ ｓｏｕｒｃｅ ＭＡＣ ａｄｄｒｅｓｓ）の修正、イーサネット宛先ＭＡＣアドレスの修正（Ｍｏｄｉｆｙ Ｅｔｈｅｒｎｅｔ ｄｅｓｔｉｎａｔｉｏｎ ＭＡＣ ａｄｄｒｅｓｓ）、ＩＰｖ４送信元アドレスの修正（Ｍｏｄｉｆｙ ＩＰｖ４ ｓｏｕｒｃｅ ａｄｄｒｅｓｓ、Ｍｏｄｉｆｙ ＩＰｖ４ Ｔｏｓビット、Ｍｏｄｉｆｙ ＩＰｖ４ ＴｏＳビットの修正（Ｍｏｄｉｆｙ ＩＰｖ４ ＴｏＳ ｂｉｔｓ）、トランスポート宛先ポートの修正（Ｍｏｄｉｆｙ ｔｒａｎｓｐｏｒｔ ｄｅｓｔｉｎａｔｉｏｎ ｐｏｒｔ）を含む。
【００３９】
図７Ｂは、３つのフロー・テーブル・エントリを有する例証となるフロー・テーブルである。エントリは、ワイルドカード（例えば、「＊」記号）を有するフィールドを含む。特定のフィールド内にワイルドカードが存在する場合、着信パケット内のフィールドの特定の値に関係なく、全ての着信パケットは、フィールドに対する「一致」を形成すると考えられる。
【００４０】
図７Ｂのテーブルの第１行のエントリは、フロー・テーブル・エントリが動作しているスイッチが、イーサネットのスイッチングを行うように指示する。特に、一致するイーサネット宛先アドレスを有する着信パケットは、ポート３に転送される。
【００４１】

図７Ｂのテーブルの第２行のエントリは、インターネットの経路指定を行うために（すなわち、パケットはその宛先ＩＰアドレスに基づいて転送される）、スイッチをどのように構成するかを示す。

【００４２】

図７Ｂのテーブルの第３行は、ファイヤウォール処理を行うのに、どのようにスイッチを構成できるかを示すエントリを含む。宛先ＩＰポート値が８０であるパケットを受信すると、そのパケットはドロップされる（すなわち、スイッチは、ポート８０のトラフィックをブロックするファイヤウォールとして働くように構成される）。

【００４３】

図７Ｂに示されるタイプのフロー・テーブル・エントリは、システムの設定動作中にコントローラ・サーバ１８によりスイッチ１４にロードすることができ、又は、コントローラ・サーバ１８におけるスイッチ１４からのパケットの受信及び処理に応答して、リアルタイムでコントローラ・サーバ１８からスイッチ１４に与えることができる。多数のスイッチ１４を有するネットワークにおいては、ネットワークを通して経路を形成するために、各スイッチに、適切なフロー・テーブル・エントリを与えることができる。

【００４４】

一例として、それぞれのエンド・ホスト間に直列に接続された第１及び第２のスイッチを含むネットワークを考える。トラフィックをエンド・ホストの第１のものからエンド・ホストの第２のものに送信する場合、第１及び第２のスイッチを通じてトラフィックを経路指定することが望ましい。第２のスイッチが第１のスイッチのポート３に接続されている場合、第２のエンド・ホストが第２のスイッチのポート５に接続されている場合、及び第２のエンド・ホストの宛先ＩＰアドレスが１７２．１２．３．４である場合、コントローラ・サーバ１８は、第１のスイッチに、図７Ｃのフロー・テーブル・エントリを与え、かつ、第２のスイッチに、図７Ｄのフロー・テーブル・エントリを与えることができる。宛先ＩＰアドレス１７２．１２．３．４を有するパケットが、第１のスイッチにおいて受信された場合、これらは、図７Ｃのテーブルにおける「ポート３に転送」アクションに従って、第２のスイッチに転送される。これらのパケットが第２のスイッチにおいて受信された場合、図７Ｄにおける「ポート５に転送」アクションに従って、第２のスイッチのポート５に接続された第２のエンド・ホストに転送される。

【００４５】

入力・出力ポート３４上で受信されたパケットを処理する際にスイッチ１４によって行うことができる例証となるステップが、図８に示される。ステップ７８において、スイッチ１４は、そのポートの１つ（例えば、図１の入力・出力ポート３４の１つ）においてパケットを受信する。

【００４６】

ステップ８０において、スイッチ１４は、受信したパケット・フィールドを、そのスイッチのフロー・テーブル２８におけるフロー・テーブル・エントリのフィールドと比較して、一致があるかどうかを判断する。フロー・テーブル・エントリにおける幾つかのフィールドは、完全な値（すなわち、完全なアドレス）を含むことができる。他のフィールドは、ワイルドカードを含むことができる（すなわち、「＊」の「ｄｏｎ’ｔ　ｃａｒｅ」ワイルドカード文字でマーク付けされたフィールド）。さらに他のフィールドは、部分的に完全なエントリ（すなわち、部分的にワイルドカード指定された部分アドレス）を有することができる。幾つかのフィールドは、範囲を用いることができ（例えば、ＴＣＰのポート番号を１から４０９６までの間の値に制限することにより）、事実上、この範囲を用いて、１つのタイプの部分的ワイルドカード指定を実施することができる。受信したパケットとフロー・テーブル・エントリとの比較を行う際、スイッチ１４は、フロー・テーブル・エントリにおける各フィールドが、いずれのワイルドカード指定もない完全な値、ワイルドカード指定を有する部分的な値、又はワイルドカード文字（すなわち、完全にワイルドカード指定されたフィールド）を含むかどうか考慮することができる。

【００４７】

ステップ８０の動作中、パケット・フィールドとフロー・テーブル・エントリの対応するフィールドとの間に一致がないと判断された場合、スイッチ１４は、リンク１６上でパケットをコントローラ・サーバ１８に送信することができる（ステップ８４）。

【００４８】

ステップ８０の動作中、パケットとフロー・テーブル・エントリとの間に一致があると判断された場合、スイッチ１４は、そのフロー・テーブル・エントリと関連付けられたアクションを実行し、かつ、そのフロー・テーブル・エントリの統計フィールドにおけるカウンタ値を更新することができる（ステップ８２）。次いで、線８６で示されるように、処理は、ステップ７８にループバックすることができるので、スイッチ１４により別のパケットを処理することができる。

【００４９】

図９は、複数の関連したネットワーク接続１６を用いて、コントローラ・サーバ１８が複数のスイッチ１４をどのように制御できるかを示す例証となるネットワークの図である。図９に示される例証となるネットワークにおいて、第１のエンド・ホスト（図９の左側のエンド・ホスト８８）は、第２のエンド・ホスト（図９の右側のエンド・ホスト８８）と通信する。エンド・ホスト８８は、コンピュータ（例えば、パーソナル・コンピュータ）、サーバ、コンピュータのクラスタ、セットトップ・ボックス、手持ち式機器、又はいずれかの他のコンピューティング機器とすることができる。エンド・ホスト８８の間の通信の一部の間、第１のエンド・ホストはパケット送信元として働くことができ、第２のエンド・ホストはパケット宛先として働くことができる。あるときには、役割を逆にすることができ、第２のエンド・ホストがパケット送信元として働き、一方、第１のエンド・ホストがパケット宛先として働くことができる。

【００５０】

ネットワークを通してパケットが正しく転送されることを保証するために、コントローラ１８は、図９に示されるスイッチの各々に、適当なフロー・テーブル・エントリを与えることができる。１つの適切な構成において、コントローラ・サーバ１８は、着信パケットとそのフロー・テーブル・エントリとの間に一致を検出しなかったスイッチからコントローラ・サーバ１８に送信されたパケットの受信に応答して、スイッチ１４に、フロー・テーブル・エントリを与えることができる。コントローラ・サーバ１８がパケットを受信すると、コントローラ・サーバ１８は、スイッチ１４のフロー・テーブル２８についての適切なエントリを判断するのに、ネットワーク構成規則２０（図１）、パケットからの情報、ネットワーク・トポロジー情報、及び他の情報を用いることができる。次に、コントローラ・サーバ１８は、フロー・テーブル・エントリをスイッチ１４に与えて、ネットワークを通してパケットを転送するように、スイッチを構成することができる。別の適切な構成の場合、コントローラ・サーバ１８は、設定動作中に、フロー・テーブル２８をスイッチ２８に与える。

【００５１】

コントローラ・サーバ１８が、スイッチからパケットを受信する前に又はその受信に応じてリアルタイムで、スイッチ１４にフロー・テーブル・エントリを与えるかどうかに関係なく、ひとたび各スイッチ１４にフロー・テーブル・エントリが与えられると、フロー・テーブル・エントリは、スイッチ１４が、ネットワークを通して満足のいく経路に沿ってパケットを転送することを保証する。

【００５２】

スイッチ１４のリソースに過負荷をかけないように注意しなければならない。通常、各スイッチ１４のフロー・テーブルの容量は制限されている。スイッチの容量は、例えば、１０，０００を超えないフロー・テーブル・エントリを処理するように、そのスイッチを制限することができる。この制限を超過するのを回避するために、異なるネットワーク位置におけるスイッチに、異なるレベルの特定性（ｓｐｅｃｉｆｉｃｉｔｙ）を有する規則を実装するフロー・テーブル・エントリを与えることができる。

【００５３】

一例として、図１０のネットワーク１０を考える。図１０のネットワーク１０において、エンド・ホスト８８は、ネットワーク１０のエッジ部分１０Ｅにおけるエッジ・スイッチ１４Ｅ、ネットワーク１０の集約・分散部分１０ＡＤにおける集約スイッチ１４ＡＤ、及びネットワーク１０のコア部分１０Ｃにおけるコア・スイッチ１４Ｃ（その１つ又はそれ以上がインターネットに結合される）を通して通信する。エンド・ホスト８８の各々と関連したアクティブなユーザが存在し得る。典型的なエッジ・スイッチ１４Ｅは、約５０の異なるエンド・ホスト８８に接続することができる（一例として）。典型的には、集約スイッチ１４ＡＤは、約２０のエッジ・スイッチ１４Ｅに接続することができる（一例として）。このアーキテクチャの結果として、集約スイッチ１０ＡＤは、約１０００のエンド・ホスト８８からのトラフィックを処理する必要があり得る。

【００５４】

アクティブなユーザは、ウェブ・ブラウザ、及び多数のネットワーク接続をもたらす他のアプリケーションを用いることができるので、集約スイッチ１４ＡＤが、１０，０００乃至２０，０００又はそれ以上のネットワーク接続を処理するように要求されることもある。集約スイッチ１４ＡＤのハードウェアが、最大１０，０００のフロー・テーブル・エントリしか処理できない場合、ネットワーク１０が過負荷になる可能性があるので、幾つかの所望のネットワーク接続が、ユーザに利用不能になる。

【００５５】

この潜在的な問題の回避を確実にするために、コントローラ・サーバ１８は、ネットワーク１０の異なる部分におけるスイッチに、異なる特定性のフロー・テーブル・エントリ（一致規則）を与えることができる。例えば、エッジ・スイッチ１４Ｅに、集約スイッチ１４ＡＤよりも制限的なエントリを与えることができ、及び／又は、集約スイッチ１４ＡＤに、コア・スイッチ１４Ｃよりも制限的なエントリを与えることができる。このタイプの構成により、所望のレベルのパケット処理を全体的に維持することを可能にする一方で、他の場合にはその容量に追いつかなくなることがある、集約スイッチ１４ＡＤ及び１４Ｃのようなスイッチにかかる負担を軽減することができる。

【００５６】

図１１は、より多くのワイルドカード指定を組み込むことにより、例証となるヘッダ・フィールド（すなわち、ＩＰ宛先アドレス・フィールド）と関連した一致規則をどのようにより制限的でなくできるかを示す。図１１の例において、ＩＰ宛先アドレス・フィールドに用いることができる最も制限的な値は、完全なＩＰアドレス「２．２．２．２」である。フロー・テーブル・エントリがこの値を有する場合、このフィールドに一致するパケットのみが、同一のＩＰ宛先アドレス（すなわち、２．２．２．２）を含むものとなる。フロー・テーブル・エントリに用いることができる、僅かにより制限的でないＩＰ宛先アドレス・フィールドは、「２．２．２．*」である。最後のＩＰアドレス位置にワイルドカードを用いることにより、２．２．２．１、２．２．２．２、２．２．２．３等のようなアドレスは全て２．２．２．*エントリと一致するので、着信パケットにおけるＩＰ宛先アドレス・フィールドに一致するための要件が緩和される。より多くのワイルドカード指定を組み込むことにより、さらにより制限的でない一致規則がもたらされる。例えば、図１１における「２．２．*」エントリは、２．２．０．０、２．２．０．１、．．．、２．２．１．０、２．２．１．１、．．．のＩＰ宛先アドレスと一致する。フロー・テーブル・エントリ・フィールドが２．*の値を有する場合、アドレスにおける最後の３つの位置の値に関係なく、「２」で始まる全てのＩＰ宛先アドレスが一致する。フィールド全体がワイルドカード指定された場合（すなわち、図１１のＩＰ宛先アドレス・フィールドが「*」の値を有する場合）、全ての着信パケットが、フロー・テーブルのＩＰ宛先アドレス・フィールドに一致するとみなされるＩＰ宛先アドレス・フィールドを含む（すなわち、「*」はあらゆるアドレス値に一致するとみなされる）。

【００５７】

図１１の例が実証するように、より多くワイルドカード指定されたフロー・テーブルの値は、より制限的でない一致規則に対応し、一方、より少なくワイルドカード指定された

フロー・テーブルの値は、より特定の一致基準に対応する。所定のネットワークにおけるパケット処理のために特定の機能を実装する場合には、ネットワークにおけるスイッチ１４の少なくとも一部が詳細なフロー・テーブル・エントリ、及びそれに対応する制限的な一致規則を用いることが望ましい。例えば、ネットワーク・エッジにおける又はその近くのスイッチが、詳細なフロー・テーブル・エントリを処理できる場合がある。しかしながら、ネットワークのコアのより近くでは、図１０に関連して説明されたトラフィックにおける集中のせいで、そのような詳細なフロー・テーブル・エントリが、スイッチに追いつかなくなる傾向がある。従って、ネットワーク・コアにより近いスイッチに対して、より制限的でない規則を用いることが有利であり得る。一例として、「２．２．２．２」のＩＰ宛先アドレス・フィールドを、エッジ・スイッチ１４Ｅにおけるフロー・テーブルに用いることができ、一方、コア・スイッチ１４Ｃにおけるフロー・テーブル・エントリに対して、*のＩＰ宛先アドレス・フィールドを用いることができる。コア・スイッチにおける「*」のエントリは、多数の可能な宛先ＩＰアドレスに適用できるため、あまり制限的でない「*」値を使用することは、コア・スイッチ１４Ｃに必要とされる異なるフロー・テーブル・エントリ数を最小にする助けとなる。

【００５８】

一例として、図１２の例証となるネットワークを考える。図１２の例においては、２つのエンド・ホスト８８がある。エンド・ホストＥＨＡはパケット送信元として働くことができ、エンド・ホストＥＨＢはパケット宛先として働くことができる。スイッチ１４は、エッジ・スイッチ１４Ｅと、スイッチ１４Ｃのようなコア・スイッチとを含む。設定動作中、又はリアルタイムのネットワーク構成動作中、コントローラ・サーバ１８は、スイッチＳＷ　Ｅ１、ＳＷ　Ｃ１及びＳＷ　Ｅ２に、スイッチＳＷ　Ｅ１、ＳＷ　Ｃ１及びＳＷ　Ｅ２を通じて、ネットワーク経路上で、エンド・ホストＥＨＡからエンド・ホストＥＨＢにパケットを転送するのに適したフロー・テーブル・エントリを与えることができる。図１２のスイッチと関連した物理的な入力・出力ポートが、スイッチを接続するリンクに隣接して示される。例えば、スイッチＳＷ　Ｅ１に隣接するラベル「１」は、エンド・ホストＥＨＡがスイッチＥ１のポート１に接続されることを示す。スイッチＳＷ　Ｅ１に隣接するラベル２及びスイッチＳＷ　Ｃ１に隣接するラベル「３」は、スイッチＥ１のポート２がスイッチＳＷ　Ｃ１のポート３に接続されることなどを示す。

【００５９】

１つの従来の手法を用いる場合、完全な一致規則が、ワイルドカードのないフロー・テーブル・エントリの形で、ネットワーク内のスイッチに与えられる。エントリは、スイッチがどのように着信パケットを転送するかを指定する。図１３は、このタイプの構成を用いて、図１２のネットワークのスイッチを与えることができるフロー・テーブル・エントリを示す。図１３に示されるように、フロー・テーブル・エントリＥ１は、スイッチＳＷ　Ｅ１に、着信パケットをポート１からポート２に転送するように指示する。フロー・テーブル・エントリＣ１は、スイッチＳＷ　Ｃ１に、着信パケットをスイッチＳＷ　Ｃ１のポート３からポート４に転送するように指示する。スイッチＳＷ　Ｅ２は、フロー・テーブル・エントリＥ２に従って、着信パケットをスイッチＳＷ　Ｅ２のポート５からポート６に転送する。従って、図１３のフロー・テーブル・エントリは、図１２のネットワーク１０において、パケットをエンド・ホストＥＨＡからＥＨＢに流れさせる。このタイプの手法は、フロー・テーブル・エントリにおいてワイルドカードを使用することを必要とせず、このことは、フロー・テーブル・エントリがスイッチの容量に追いつかなくなる状況をもたらし得る。例えば、スイッチＣ１のようなスイッチには、多数のエッジ・スイッチ及びエンド・ホストについてのフロー・テーブル・エントリを与える必要があり得る。スイッチＳＷ　Ｃ１内に格納する必要があるフロー・テーブル・エントリの数が、スイッチＳＷＣ１のハードウェア能力を上回ることがある。

【００６０】

この問題に対処する従来の構成が、図１４に示される。図１４の構成では、フロー・テーブル・エントリの選択されたフィールド内に、ワイルドカード指定が用いられる。例え

ば、図１４の各フロー・テーブル・エントリの物理ポートのフィールドが、ワイルドカードを有する。これにより、各エンド・ホストが取り付けられる物理ポートの識別情報に関係なく、スイッチに結合されたエンド・ホストの全てからのパケットが、同じく処理されることが可能になる。物理ポートの情報に基づいてパケットを異なるように処理する試みは行われないため（及び、図１４のテーブルにおける他のワイルドカード指定情報のため）、必要とされるフロー・テーブル・エントリの数は低減する。

【００６１】

図１４の従来の手法は、低減した数のフロー・テーブル・エントリを用いて、パケットをエンド・ホストＥＨＡからエンド・ホストＥＨＢに転送するのを可能にするが、セキュリティが危険にさらされる。特に、スイッチは物理ポートの情報に基づいてトラフィックをブロックできないため、図１４の手法は、攻撃者が物理ポートをスプーフィング（ｓｐｏｏｆ）するのを可能にする。

【００６２】

フロー・テーブル・エントリからネットワークのスイッチ（例えば、集約スイッチ及びコア・スイッチ）に課される負担を減らしながら、所望のレベルのネットワーク性能を維持するために（例えば、セキュリティを維持するために）用いることができるスキームが、図１５のフロー・テーブル・エントリにより示される。フロー・テーブル・エントリＥ１’、Ｃ１’及びＥ２’をネットワークの異なる部分にあるスイッチに与えることができる。例えば、エントリＥ１’及びＥ２’を、それぞれ図１２のスイッチＳＷ　Ｅ１及びＳＷ　Ｅ２に用いることができ、エントリＣ１’をスイッチ１４Ｃに用いることができる。フロー・テーブル・エントリＥ１’、Ｃ１’及びＥ２’は、異なるレベルの制限性を有することができる。特に、ネットワークのコアにより近いスイッチ（例えば、図１２のコア・スイッチＳＷ　Ｃ１についてのエントリＣ１’）のための一致規則は、より多くのワイルドカード指定を有することができ、かつ、ネットワークのコアからより遠くにあるスイッチ（例えば、エッジ・スイッチ１４Ｅについてのエントリ Ｅ１’及びＥ２’）のための一致規則より制限的でないものとすることができる。

【００６３】

例えば、スイッチＳＷ　Ｅ１に、着信パケットをどのように転送するかを指示するのに用いることができる、エントリＥ１’のようなフロー・テーブル・エントリは、物理ポートの入力フィールド及びＩＰ送信元アドレスのフィールド内に特定の情報を含むことができ、一方、集約スイッチ又はコア・スイッチＳＷ　Ｃ１に、着信パケットをどのように転送するかを指示するのに用いることができる、エントリＣ１’のようなフロー・テーブル・エントリは、少なくとも幾つかのワイルドカード指定を含むことができる。図１５の例の場合、エントリＥ１’（スイッチＳＷ　Ｅ１により用いられる）及びエントリＥ２’（スイッチＳＷ　Ｅ２により用いられる）の両方とも、完全なフィールドのみを含み、ワイルドカード指定を有するフィールドは含んでおらず、一方、エントリＣ１’は、完全にワイルドカード指定された複数のフィールド（すなわち、物理入力ポート、ＩＰ送信元アドレス、宛先ＴＣＰポート、及び送信元ＴＣＰポート）を含む。

【００６４】

Ｃ１’エントリのようなエントリは、Ｅ１’及びＥ２’エントリのようなエントリより多くのワイルドカード指定を含むので、Ｃ１’エントリのようなエントリを用いることは、コア・スイッチＳＷＣ１（本例において）が維持するフロー・テーブル・エントリの数を減らす助けとなる。これにより、スイッチＳＷ　Ｃ１へのフロー・テーブルの負担が減り、スイッチ１４に追いつかなくなるのを防ぐ助けとなる。同時に、エッジ・スイッチＳＷ　Ｅ１及びＳＷ　Ｅ２（本例において）に対応するフロー・テーブル・エントリの完全性のために、セキュリティが保護される。エントリＥ１’及びＥ２’は、物理ポートの情報を保持し、かつ、コア・スイッチＣ１にアクセスすることができないので、物理ポートのスプーフィングは可能でない（エンド・ホストをコア・スイッチＣ１に物理的に直接接続できないと仮定して）。

【００６５】

コントローラ・サーバ１８は、図１５に示されるタイプのテーブルについてのフロー・テーブル・エントリを選択的に分散させることができる。例えば、コントローラ・サーバ１８は、フロー・テーブル・エントリＥ１’のようなフロー・テーブル・エントリを、スイッチＳＷ　Ｃ１のようなネットワーク・コアにより近いスイッチではなく、スイッチＥ１のようなエッジ・スイッチに分散させ、かつ、フロー・テーブル・エントリＣ１’のようなフロー・テーブル・エントリを、エッジ・スイッチではなく、スイッチＳＷ　Ｃ１のようなコア・スイッチ又はコアに近いスイッチに分散させることができる。

【００６６】

図１６は、エッジ・スイッチのフロー・テーブル・エントリがワイルドカード指定を含むスキームについてのフロー・テーブル・エントリを示す。例えば、図１２のスイッチＳＷ　Ｅ１についてのエントリＥ１’’により示されるように、宛先ＴＣＰポート・フィールドをワイルドカード指定し、かつ、送信元ＴＣＰポート・フィールドをワイルドカード指定することができる。図１２のスイッチＳＷ　Ｅ２についてのエントリＥ２’’により示されるように、ＩＰ送信元アドレス・フィール及び送信元ＴＣＰポートをワイルドカード指定することができる。パケットにおける送信元ＴＣＰポート情報は、ＩＰリンクを確立する通常のプロセスの一部としてランダムに割り当てられ（戻りトラフィックが適切なエンド・ホスト処理に向けられることを保証するために）、かつ、通常有用なセキュリティ情報は含まないため、送信元ＴＣＰポートにワイルドカード指定を使用することにより、セキュリティを低減させずに、フロー・テーブル・エントリを低減させることが可能になる。Ｅ１’’エントリにおける宛先ＴＣＰポート情報のワイルドカード指定は、スイッチＳＷ　Ｅ１が、エンド・ホストＥＨＢにおける権限のないＴＣＰポートに向かうその関連したエンド・ホストＥＨＡからのトラフィックをブロックすることができないので、セキュリティ上の問題をもたらす可能性がある。それにもかかわらず、宛先ＴＣＰポート・フィールドは、フロー・テーブル・エントリＥ２’’において２２の完全な（ワイルドカード指定のない）エントリで満たされているので、スイッチＳＷ　Ｅ２は、この所望のトラフィック・ブロック動作を行うことができる。

【００６７】

従って、図１６のフロー・テーブル・エントリにおいて幾つかのワイルドカード指定が存在していても、パケット転送機能の全体的な損失はない。エントリＥ１’’のワイルドカード指定により失われたどの機能も、フロー・テーブル・エントリＥ２’’のより完全なフィールドの機能によって埋め合わせられるので、図１５の例において実行された同じパケット処理機能が、図１６の例において実行される。図１５の例におけるように、エントリＣ１’’のようなコア・スイッチのフロー・テーブル・エントリは、スイッチＳＷ　Ｃ１により維持されなければならないフロー・テーブル・エントリの総数を減らすために、ワイルドカード指定（例えば、ネットワーク・エッジにより近いスイッチについてのフロー・テーブル・エントリより多いワイルドカード指定）を含むことができる。

【００６８】

図１７は、ネットワークの異なる位置において異なるレベルの特定性を有するフロー・テーブル・エントリ（すなわち、ネットワークのエッジにおける又はこれに近いスイッチに対してはより制限的であり、ネットワークのコア又はその近くであまり制限的ではないフロー・テーブル・エントリ）を含むネットワークを動作させる際に必要とされる例証となるステップのフローチャートである。

【００６９】

ステップ９０の動作中、コントローラ・サーバ１８はネットワーク機器を識別し、ネットワークのトポロジーを判断することができる。例えば、ステップ９２の動作中、コントローラ・サーバ１８は、各スイッチ１４の能力を判断することができる。ステップ９４の動作中、コントローラ・サーバ１８は、ネットワークのレイアウトについての情報（例えば、どのスイッチ及びエンド・ホストが、スイッチにおける入力・出力ポートの各々に接続されているか等）を得ることができる。スイッチ能力について収集することができる情報は、各スイッチにおける最大公称フロー・テーブル容量（例えば、各スイッチにおいて

処理することができるフロー・テーブル・エントリの公称最大数）、フロー・テーブル・エントリを処理するためのスイッチの実際の現在の容量（すなわち、スイッチのフロー・テーブル内に現在存在する新しいフロー・テーブル・エントリについての空いている行の数）、各スイッチが行うことができるアクションのタイプ等を含む。所望であれば、ネットワーク内のエンド・ホストの能力についての情報を収集することができる。エンド・ホストの能力について収集することができる情報は、どのタイプの処理がサポートされるか、及びどの接続規則がそれらのプロセスと関連付けられるか（例えば、エンド・ホスト番号Ｘは、いずれかのエンド・ホストがポート８０を用いて接続するのを可能にするウェブ・サーバである）情報を含む。ネットワーク・トポロジー情報は、どのスイッチ・ポートが互いに接続されているか、幾つのエンド・ホストが各スイッチに取り付けられるか、幾つの他のスイッチが各スイッチに接続されるか、及びエンド・ホストが取り付けられたポートのアイデンティティについての情報を含むことができる。ネットワークのトポロジーを判断するために、コントローラ・サーバ１８は、リンク層検出プロトコル（ＬＬＤＰ）パケットのようなプローブ・パケットをネットワーク全体にわたって送信することができる。スイッチ及び他のネットワーク・コンポーネントは、コントローラ・サーバにより照会されたときに、それらの能力についての情報を戻すことができる。ステップ９０の動作は、ネットワーク１０の動作中連続的に行うことができる。

【００７０】

ステップ９４の動作中にネットワークのトポロジーを判断する際、コントローラ・サーバ１８は、スイッチ１４を、主にネットワーク・エッジ１０Ｅ、集約（集約・分散）ネットワーク部分１０ＡＤ、又はネットワーク・コア１０Ｃと関連付けられたものとしてカテゴリ化することができる（例えば、図１０を参照されたい）（例えば、スイッチ１４を、エッジ・ネットワーク・スイッチとして又は非エッジ・スイッチとしてカテゴリ化する）。測定基準を各々のスイッチに適用して、スイッチがエッジ・スイッチであるか（例えば、スイッチが多数のエンド・ホストに接続されている場合）、又は非エッジ・スイッチであるか（例えば、スイッチがどのエンド・ホストにも接続されていない及び／又は監視ホストにだけ接続されている場合）を判断することができる。測定基準を各々の非エッジ・スイッチに適用して、非エッジ・スイッチが集約スイッチであるか（例えば、スイッチが多数のエッジ・スイッチに接続されている場合）、又はコア・スイッチであるか（例えば、集約スイッチ又はコア・スイッチに接続されており、エッジ・スイッチにほとんど又は全く接続されていないスイッチ）を判断することができる。

【００７１】

１つの例証となる測定基準では、スイッチが１つ又はそれ以上のエンド・ホスト（例えば、多数のエンド・ホスト）に接続されている場合、スイッチをエッジ・スイッチとしてカテゴリ化することができ、スイッチがエンド・ホストに接続されておらず、及び／又は１つのホスト（又は、場合によっては１つより多いホスト）に排他的に又は主として監視目的のために接続されている場合、非エッジ・スイッチとしてカテゴリ化することができる。スイッチ１４をカテゴリ化する際にコントローラ・サーバ１８により用いることができる別の例証としての測定基準では、第１のスイッチが第２のスイッチより多くのエンド・ホストに接続されている場合、第１のスイッチは、第２のスイッチよりエッジ様のものとしてカテゴリ化することができる。第１のスイッチがエンド・ホストより多くの取り付けられたスイッチを有し、かつ、第２のスイッチがエンド・ホストより少ない取り付けられたスイッチを有する場合、第１の非エッジ・スイッチは、第２の非エッジ・スイッチよりコア様（非エッジ様）であるとみなすことができる。所望であれば、スイッチをカテゴリ化するのに、他の測定基準を用いることができる。これらは、単に説明に役立つ実例にすぎない。ひとたび判断されると、スイッチのカテゴリは、ネットワーク構成中に適切なフロー・テーブル・エントリを分散させる際に用いることができる。

【００７２】

ステップ９６の動作中、コントローラ・サーバ１８は、パケット送信元（例えば、図１２のエンド・ホストＥＨＡのようなエンド・ホスト８８の１つ）からパケット宛先（例え

ば、図１２のエンド・ホストＥＨＢのようなエンド・ホスト８８の１つ）に送信されたパケットに対して、ネットワーク１０を通る適切な経路を判断することができる。コントローラ・サーバ１８は、ステップ９０の動作中に収集した情報のような情報を用いて、パケットについての適切な経路を判断することができる。経路は、ネットワーク設定動作中に識別することができ、又は、コントローラ・サーバ１８においてスイッチ１４の１つからのパケット（例えば、パケットを受信したが、パケットに対する一致を生じたフロー・テーブル・エントリは含まなかったスイッチによりコントローラ・サーバ１８に送信されたパケット）の受信に応答して、リアルタイムで判断することができる。

【００７３】

ステップ９８の動作中、コントローラ・サーバ１８は、ステップ９６の動作中に識別された経路及びネットワーク構成規則２０（図１）を用いて、完全なフロー・テーブル・エントリ（すなわち、図１３に示されるタイプのエントリ）を生成することができる。ネットワーク構成規則２０は、どのエンド・ホストがどのサービスにアクセスできるか（所望であれば、集合エンド・ホストを用いて）の規則を含むことができる。これらのネットワーク構成規則は、コントローラ・サーバ１８が生成するフロー・テーブル・エントリのセットにおいて具体化することができる。例えば、特定のエンド・ホストが特定のサービスにアクセスできない場合、このタイプの権限のないアクセスを防ぐように（例えば、ポートをブロックすること等により）フロー・テーブル・エントリを構築することができる。ステップ９８において生成されたフロー・テーブル・エントリは、好ましくは完全なフィールド（ワイルドカード指定のないフィールド）を含み、従って、ネットワークについての完全なパケット転送機能を維持する。

【００７４】

図１０及び図１２に関連して説明されたように、スイッチ１４におけるフロー・テーブル・エントリの全てに対して、完全な（ワイルドカード指定のない）フィールドを用いると、スイッチ１４に、具体的にはネットワークのエッジから離れたところに位置するスイッチに負担がかかることがある。これらのスイッチが過負荷にならないことを保証するために、コントローラ・サーバ１８は、ステップ１００の動作中、ステップ９８のフロー・テーブル・エントリの圧縮バージョンを生成することができる。図１５及び図１６に関連して説明されたように、これらのフロー・テーブル・エントリは、必要とされるフロー・テーブル・エントリ数を低減させるためにワイルドカード指定を含んでいる。インテリジェントなワイルドカード割り当てを用いて、スイッチ１４の所望のパケット転送機能が、ステップ９８で生成されたフロー・テーブル・エントリに対して保護されることを保証することができる。例えば、エッジ・スイッチが物理ポート情報を維持することを保証することにより（すなわち、非エッジ・スイッチのみについての物理ポート情報をワイルドカード指定することにより）、物理ポートのスプーフィングを防ぐためのネットワーク・スイッチの能力を保護することができる。別の例として、ワイルドカード指定されていない適切なＴＣＰポートのフィールドがエッジ・スイッチのエントリに保持されることを保証することによって、ＴＣＰポートのブロックを実装するためのスイッチの能力を保持することができる。

【００７５】

ステップ９８の動作を行う際、サーバ・コントローラ１８は、スイッチに、ネットワーク内のそれらの位置に合わせたフロー・テーブル・エントリが与えられることを保証することができる。ネットワーク・エッジにおける又はその近くのネットワーク・スイッチ（例えば、エッジ・スイッチ）のために、より制限的なフロー・テーブル・エントリを用いることができ、一方、ネットワーク・コアのより近くにあるスイッチ（例えば、非エッジ・スイッチ）のために、より制限的でないフロー・テーブルのエントリを用いることができる。スイッチの位置、公称スイッチ容量、実際のスイッチ容量等のような要因に基づいて、フロー・テーブル・エントリをスイッチに与えることができる。

【００７６】

ネットワークにおけるコア・スイッチが、フロー・テーブルを含み、かつ、コントロー

ラ１８により調整可能である場合、コア・スイッチには、集約スイッチについてのフロー・テーブル・エントリと同じだけ制限的な又はそれより制限的ではないフロー・テーブル・エントリを与えることができる。幾つかのネットワークにおいて、コア・スイッチは、コントローラ・サーバ１８とは独立して動作するコントローラにより制御することができ、かつ、コントローラ・サーバ１８と互換性がないものとすることができる。このタイプの状況において、コア・スイッチは、それぞれのコントローラを用いて構成できるので、コントローラ・サーバ１８は、コア・スイッチに、フロー・テーブル・エントリを与える必要がない。

【００７７】

典型的なシナリオにおいて、コントローラ・サーバ１８は、エッジ・スイッチ１４Ｅに、ワイルドカード指定されるフィールドがほとんどないか又は全くない、完全な又はほぼ完全なフロー・テーブル・エントリを与えることができる。集約スイッチ１４ＡＤのような非エッジ・スイッチには、より制限的でないフロー・テーブル・エントリを与えることができる。例えば、集約スイッチ１４ＡＤには、その唯一の完全なフィールドが宛先ＩＰフィールドであり、その他のフィールドは完全な又は部分的なワイルドカードを含むフロー・テーブル・エントリを与えることができる。コア・スイッチ１４Ｃがコントローラ１８によって制御されない場合、コア・スイッチ１４Ｃのような非エッジ・スイッチには、コントローラ１８からフロー・テーブル・エントリを与える必要はない。しかしながら、コア・スイッチ１４Ｃがコントローラ１８により制御される場合、コントローラ１８は、部分的にワイルドカード指定された宛先ＩＰアドレスを除いて、完全にワイルドカード指定されたフロー・テーブル・エントリをコア・スイッチ１４Ｃに与えることができる。一例として、コア・スイッチについてのフロー・テーブル・エントリは、宛先ＩＰアドレス・フィールドを除いて、全てのフィールドにおいてワイルドカードを有することができる。コア・スイッチについてのフロー・テーブル・エントリにおける宛先ＩＰアドレス・フィールドには、「１７１．６４．１２３．*」のような部分的にワイルドカード指定された値を与えることができ（所望のサブネットにアドレス指定されたパケットと一致するように）、この部分的にワイルドカード指定された宛先ＩＰアドレス・フィールドに対応するアクションは、「ポート３に送信する」とすることができる。

【００７８】

ひとたびステップ１００のフロー・テーブル・エントリが生成されると、コントローラ・サーバ１８は、これらのフロー・テーブル・エントリを適切なスイッチ１４に分散させることができる。スイッチ１４がこのように構成される場合、パケットは、パケットの送信元とパケット宛先との間でネットワーク１０を通じて流れることができる。

【００７９】

所望であれば、ステップ９８の動作をステップ１０の動作と結合させることができる（すなわち、ステップ９８の完全なフロー・テーブル・エントリを計算する中間ステップを行うことなく、経路及びネットワーク構成規則から、選択的なワイルドカード指定を含む圧縮されたフロー・テーブル・エントリを直接生成することができる。）

【００８０】

図１８は、図１７のステップ９０の動作を行う際（すなわち、ネットワーク１０のトポロジー及びそのスイッチ１４の能力を判断するとき）に必要とし得る動作を示す。ステップ１０４の動作中、コントローラ・サーバ１８は、適切な測定基準（例えば、取り付けられたエンド・ホストの数、取り付けられたスイッチの数等に基づいた測定基準）を用いて、ネットワーク１０内のそれらの位置に応じてスイッチをカテゴリ化することができる。例えば、１つ又はそれ以上のエンド・ホスト８８に接続されたスイッチをエッジ・スイッチとしてカテゴリ化することができる。他のスイッチ１４は、ネットワークのトポロジー内の位置、及び／又は、取り付けられたスイッチの数等などの要因に基づいて、ネットワーク・コア１０Ｃ、又は、ネットワーク１０ＡＤの集約・分散部分に属するとカテゴリ化することができる。

【００８１】

ステップ１０６及び１０８の動作中、コントローラ・サーバ１８は、ネットワーク接続１６上で、ネットワーク１０内の個々のスイッチ１４にクエリを発行することができる。例えば、コントローラ・サーバ１８は、ステップ１０６の動作中、その公称容量についてスイッチ１４に照会することができる。スイッチ１４は、その公称容量（すなわち、いずれの既存のエントリもない場合、スイッチが処理することができるフロー・テーブル・エントリの理論上の最大数）についての情報で応答することができる。実際の容量の情報について照会された場合（ステップ１０８）、スイッチ１４は、コントローラ・サーバ１８に、その実際の（現在の）容量についての情報（すなわち、スイッチの能力を超過することなく、スイッチにロードすることができる付加的なフロー・テーブル・エントリの数）を与えることができる。ステップ９０の動作中に収集される情報は、ネットワーク１０におけるスイッチについての適切なフロー・テーブル・エントリを生成する際に（例えば、フロー・テーブル・エントリによりスイッチに追いつかなくならないように、図１７のステップ９６、９８及び１００の動作中、スイッチ１４についてのフロー・テーブル・エントリをどのように生成するかを判断する際に）用いることができる。

【００８２】

図１９の例証となるネットワーク１０において、エンド・ホストＥＨＣ及びＥＨＤは、スイッチＳＷ　Ｅ３及びＳＷ　Ｅ４のようなエッジ・スイッチ１４Ｅ、及び、コア・スイッチＳＷ　Ｃ２のようなコア・スイッチ１４Ｃを介して、インターネット１１０に結合される。コア・スイッチＳＷ　Ｃ２は、デフォルトのインターネット・ゲートウェイＤＧによりインターネットに結合することができる。本例において、エンド・ホスト機器１１２のようなインターネット機器は、３．３．３．３の関連したＩＰアドレスを有することができる。エンド・ホストＥＨＣは、１．１．１．１のＩＰアドレスを有することができ、かつ、ＴＣＰポート２２と関連付けることができる。エンド・ホストＥＨＤは、２．２．２．２のＩＰアドレスを有することができ、かつ、ＴＣＰポート２２と関連付けることができる。

【００８３】

図１９の例証となるネットワークのようなネットワークにおいて、コントローラ・サーバ１８は、ネットワークにおいて使用中のサブネットの知識を有し得る。この情報に基づいて、コントローラ・サーバ１８は、ネットワーク内に「３」で始まるＩＰアドレスを含むエンド・ホストがない（すなわち、ネットワーク１０内に、３．*の有効なＩＰ宛先アドレスがない）と結論付けることができる。このことは、コントローラ・サーバ１８が、３．*のＩＰアドレスを有するトラフィックをデフォルトのゲートウェイＤＧに転送するフロー・テーブルを構築することを可能にする。ネットワーク１０内に、３．*の宛先アドレスがないので、この転送タスクを行う際に、フロー・テーブル・エントリ内の他のフィールドの値を必要としない。

【００８４】

図２０は、図１９のネットワーク１０においてこのタイプのパケット転送スキームを実装するために用いることができる例証となるフロー・テーブルである。スイッチＳＷ　Ｅ３により用いることができるフロー・テーブル・エントリＥ３により示されるように、エンド・ホストＥＨＣから宛先ＩＰアドレス３．３．３．３へのパケットは、スイッチＳＷ　Ｅ３のポート２に転送される（送信元アドレス・ポート５における２．２．２．２を有する、転送ポートと同じタイプのエントリを、エンド・ホストＥＨＤからのパケットのために用いることができる）。スイッチＳＷ　Ｃ２により用いることができるフロー・テーブルのエントリＣ２により示されるように、ＩＰ宛先アドレスが「３．*」に一致する全てのパケットがポート７に、従って、ゲートウェイＤＧに転送される。フロー・テーブル・エントリＣ２における他のフィールドをワイルドカード指定し、必要とされるフロー・テーブル・エントリの数を最小にすることができる。

【００８５】

ネットワーク１０のための別の可能なトポロジーが、図２１に示される。このタイプの構成では、ネットワークのドメインＡ及びＢが、コントローラ・サーバ１８からフロー・

テーブルをロードすることができるスイッチ１４にポピュレートされる。技術的な制限のため又は許可がないために、介在するローカル・コア１１４におけるネットワーク・スイッチは、コントローラ・サーバ１８からフロー・テーブル・エントリがロードされず、コントローラ・サーバ１８からフロー・テーブル・エントリをロードすることができない。

【００８６】

コントローラ・サーバ１８により与えられるフロー・テーブル・エントリの制御下でドメインＡ及びＢが機能するのを可能にするために、ネットワーク・トンネル（例えば、仮想ローカル・エリア・ネットワーク・トンネル）を、ローカル・コア・ネットワーク１１４を通して確立することができる。このトンネルは、ネットワーク１０の１つのドメイン（例えば、ドメインＡ）においてカプセル化エンジンを用いて他方のドメインに向かうデータ・トラフィックをカプセル化し、ネットワーク１０の他方のドメイン（例えば、ドメインＢ）においてカプセル化解除（ｄｅｅｎｃａｐｓｕｌａｔｉｏｎ）エンジンを用いてカプセル化解除し、従って、カプセル化されたデータを復旧することにより、形成することができる。ネットワーク１０におけるカプセル化エンジン及びカプセル化解除エンジンは、Ｇｅｎｅｒｉｃ Ｒｏｕｔｉｎｇ Ｅｎｃａｐｓｕｌａｔｉｏｎ（ＧＲＥ）、マルチプロトコル・ラベル・スイッチング（Ｍｕｌｔｉｐｒｏｔｏｃｏｌ Ｌａｂｅｌ Ｓｗｉｔｃｈｉｎｇ、ＭＰＬＳ）、仮想ローカル・エリア・ネットワーク（ＶＬＡＮ）カプセル化技術、ネットワーク・トンネルのデータをカプセル化するための他の技術等といったカプセル化方法を用いることができる。

【００８７】

スイッチ１４の利用可能なハードウェア及びソフトウェア能力に起因して、特定のスイッチ１４が、カプセル化エンジン及びカプセル化解除エンジンを実行するのにより適していることがある。例えば、ドメインＡにおけるスイッチ１４－１は、カプセル化エンジンＥＥを実行するのに適切であり、ドメインＢにおけるスイッチ１４－２は、カプセル化解除エンジンＤＥを実行するのに適切であり得る。カプセル化エンジンＥＥ及びカプセル化解除エンジンＤＥが実装されるスイッチは、ローカル・コア１１４に直接接続される必要はない。図２１に示されるように、例えば、１つ又はそれ以上の他のスイッチ１４（例えば、カプセル化エンジン又はカプセル化解除エンジンを有さないスイッチ）を、スイッチ１４－１とローカル・コア１１４との間に置くことができ、スイッチ１４の１つ又はそれ以上（例えば、カプセル化エンジン又はカプセル化解除エンジンを有さないスイッチ）をスイッチ１４－２とローカル・コア１１４との間に置くことができる。

【００８８】

トラフィックがローカル・コア１１４を通って（すなわち、コントローラ・サーバからのフロー・テーブル・エントリにより制御されていないネットワークの部分を通って）進むことができるのを確実にするために、コントローラ・サーバ１８は、パケットをネットワーク１０の各ドメイン内に転送するフロー・テーブル・エントリを生成し、トラフィックがエンジンＥＥ及びＤＥにより適切にカプセル化及びカプセル化解除されるようにすることができる。例えば、トラフィックがドメインＡのエンド・ホストＥＨ１によりドメインＢのエンド・ホストＥＨ２に送信される場合、コントローラ・サーバ１８は、パケットを、経路１１６に沿ってスイッチを通ってＥＨ１からスイッチ１４－１上のカプセル化エンジンＥＥに転送し、カプセル化エンジンＥＥによりカプセル化されたパケットを、経路１１８に沿ってローカル・コア１１４を通ってスイッチ１４－２上のカプセル化解除エンジンＤＥに転送し、かつ、カプセル化解除エンジンＤＥによりカプセル化解除されたパケットを、経路１２０を介してエンド・ホスト８８に転送するフロー・テーブル・エントリを生成することができる。パケットが取る経路は、特定のスイッチを複数回通ることができる。例えば、パケットは、経路１１６に沿って流れているときに最初に、経路１１８に沿って流れているとき二回目に、スイッチ１４－３を通過することができ、かつ、経路１１８に沿って流れているとき最初に、経路１２０に沿って流れているときに二回目にスイッチ１４－４を通過することができる（本例において）。

【００８９】

図２１のネットワーク１０のスイッチのためのフロー・テーブル・エントリを生成するのに必要とされる例証となるステップのフローチャートが、図２２に示される。ステップ１２２の動作中、コントローラ・サーバ１８は、ネットワーク１０のトポロジーについての情報（例えば、ドメインＡ及びＢの機器の位置、間に置かれたローカル・コア１１４又はコントローラ・サーバ１８により制御されていない他の機器の性質）、フロー・テーブル・エントリの能力及び各スイッチ１４の他の能力（例えば、コア１１４を通してネットワーク・トンネルを形成するためのあらゆるカプセル化エンジンＥＥ及びカプセル化解除エンジンの位置）、並びに他のネットワーク情報を収集することができる。

【００９０】

ステップ１２２の動作中に収集された情報に基づいて、コントローラ・サーバ１８は、ステップ１２４において、カプセル化エンジンＥＥ、ローカル・コア１１４、及びカプセル化解除エンジンＤＥを通して、エンド・ホストＥＨ１からエンド・ホストＥＨ２にトラフィックを指向させる、図２１の経路１１６、１１８及び１２０のような経路を識別することができる。次に、コントローラ・サーバ１８は、これに応じて、ステップ１２６において、適切なフロー・テーブル・エントリを有するスイッチ１４をロードすることによって、ネットワークを構成することができる。ネットワーク１０の動作中、エンド・ホストＥＨ１からのトラフィックは、ローカル・コア１１４のネットワーク・スイッチを通り抜け、エンド・ホストＥＨ２に到達する。

【００９１】

より制限的なフロー・テーブル・エントリを、非エッジ・スイッチではなく、エッジ・スイッチに対して用いることにより、ネットワーク１０において安全な通信を保証することができる。例えば、エッジ・スイッチには、特定の物理ポート上のホストが特定のアドレス（例えば、ＩＰ送信元アドレス、イーサネット送信元アドレス、ＶＬＡＮタグ、又はイーサネット・アドレス＋ＶＬＡＮタグといったこれらと他のフィールドとの組み合わせ）を用いることを要求するフロー・テーブル・エントリを与えることができる。図２３の例は、エッジ・スイッチについてのフロー・テーブル・エントリが、関連したアドレス情報に加えて、特定の物理ポート情報（例えば、ポート＝４）をどのように含むことができるかを示す。このタイプの構成の場合、エッジ・スイッチ上のフロー・テーブルは、物理ポートのフィールド及び関連したアドレス・フィールドのためのワイルドカードをもたない。エッジ・スイッチのフロー・テーブル・エントリが指定する転送アクション又は他のアクションは、物理ポート及びパケットのアドレスの両方がフロー・テーブル・エントリにより確立された基準を満たす場合にのみ実行される。

【００９２】

非エッジ・スイッチ（例えば、集約スイッチ）には、より制限的でないフロー・テーブル・エントリを与えることができる。例えば、図２３の例証となる非エッジ・スイッチのフロー・テーブル・エントリにより示されるように、集約スイッチには、物理ポートがワイルドカード指定され、かつ、転送を決定するためにアドレス・フィールド情報だけが使用されるフロー・テーブル・エントリを与えることができる。エッジ・スイッチは、悪意のあるホストからの攻撃を防ぐフロー・テーブル・エントリを含むので、集約スイッチは、エッジ・スイッチからの着信パケットにおけるアドレス情報が信頼できると考えることができる。

【００９３】

例えば、エッジ・スイッチの１つに接続された悪意あるエンド・ホストが、別のエンド・ホストになりすまそうとして、ＩＰ送信元アドレスを偽造しようとする場合、悪意あるエンド・ホストが接続されるエッジ・スイッチは、悪意あるエンド・ホストの物理ポートに対して適切な一致を検出しない。悪意あるホストが接続されたエッジ・スイッチにおけるフロー・テーブル・エントリは、物理ポート番号情報及びアドレス情報の両方を含む。アドレスが悪意あるエンド・ホストによって成功裏に偽造された場合でも、悪意あるエンド・ホストからのいずれのパケットも、悪意あるエンド・ホストと関連した物理ポート情報を含み、なりすまされたエンド・ホストの正しい物理ポート情報は含まない。悪意ある

エンド・ホストの物理ポートは、エッジ・スイッチのフロー・テーブル・エントリにおける要求される物理ポートと一致しないため、エッジ・スイッチは、パケットを悪意あるエンド・ホストから集約スイッチに転送せず、なりすましの試みは失敗する。

【００９４】

１つの実施形態によれば、コントローラ・サーバを用いて、ネットワーク内のネットワーク・スイッチにフロー・テーブル・エントリを与える方法が提供される。各々のネットワーク・スイッチは、パケット・フィールドをフロー・テーブル・エントリのフィールドと比較することにより、パケットを処理する。この方法は、コントローラ・サーバによって、スイッチの幾つかをエッジ・ネットワーク・スイッチとしてカテゴリ化し、コントローラ・サーバによってネットワーク・スイッチを非エッジ・スイッチとしてカテゴリ化することを含む。この方法は、コントローラ・サーバによって、異なるフロー・テーブル・エントリを、非エッジ・スイッチとしてカテゴリ化されたネットワーク・スイッチではなく、エッジ・ネットワーク・スイッチとしてカテゴリ化されたネットワーク・スイッチに分散させることを含む。

【００９５】

別の実施形態によれば、フロー・テーブル・エントリを分散させることは、フロー・テーブル・エントリを、ネットワーク接続上でコントローラ・サーバから、ネットワーク・スイッチ上の対応するコントローラ・クライアントに分散させることを含む。

【００９６】

別の実施形態によれば、フロー・テーブル・エントリを分散させることは、コントローラ・サーバにおけるネットワーク・プロトコル・スタックを用いて、ネットワーク接続上で、コントローラ・クライアントにおける対応するネットワーク・プロトコル・スタックと通信することを含む。

【００９７】

別の実施形態によれば、フロー・テーブル・エントリを分散させることは、完全なフィールドのみを有するフロー・テーブル・エントリをエッジ・スイッチに分散させて、少なくとも幾つかのワイルドカード指定されたフィールドを有するフロー・テーブル・エントリを非エッジ・スイッチに分散させることを含む。

【００９８】

別の実施形態によれば、各々のフロー・テーブル・エントリは、送信元インターネット・プロトコル（ＩＰ）アドレス・フィールドを含み、エッジ・スイッチについての各々のフロー・テーブル・エントリの送信元ＩＰアドレス・フィールドにはワイルドカードがなく、非エッジ・スイッチについての各々のフロー・テーブル・エントリの送信元ＩＰアドレス・フィールドは、少なくとも幾つかのワイルドカード指定を含む。

【００９９】

別の実施形態によれば、ネットワーク・スイッチはポートを含み、フロー・テーブル・エントリは、ネットワーク・スイッチがパケットをポートのどれに転送すべきかを指定するアクション・フィールドを含む。

【０１００】

１つの実施形態によれば、第１の組のエンド・ホスト及びネットワーク・スイッチと関連した第１のネットワーク・ドメインと、第２の組のエンド・ホスト及びネットワーク・スイッチと関連した第２のネットワーク・ドメインと、第１のネットワーク・ドメイン及び第２のネットワーク・ドメイン内のネットワーク・スイッチに対してフロー・テーブル・エントリを与えるコントローラ・サーバと、ローカル・コア・ネットワークとを有するネットワークを動作させる方法が提供される。各々のネットワーク・スイッチは、パケット・フィールドをフロー・テーブル・エントリのフィールドと比較することによって、パケットを処理する。カプセル化エンジンが第１のネットワーク・ドメイン内のネットワーク・スイッチの１つの上に実装され、カプセル化解除エンジンが第２のネットワーク・ドメイン内のネットワーク・スイッチの１つの上に実装される。この方法は、制御サーバによって、第１のドメイン及び第２のドメイン内のネットワーク・スイッチに、パケットを

第１のネットワーク・ドメイン内の第１のエンド・ホストからカプセル化エンジンに転送し、カプセル化されたパケットを、ローカル・コア・ネットワークを通じてカプセル化エンジンからカプセル化解除エンジンに転送し、かつ、パケットをカプセル化解除エンジンから第２のドメイン内の第２のエンド・ホストに転送するように指示する、第１のドメイン内のネットワーク・スイッチ及び第２のドメイン内のネットワーク・スイッチについてのフロー・テーブル・エントリを生成することを含む。

【０１０１】

別の実施形態によれば、ローカル・コア・ネットワークは、コントローラ・サーバからのフロー・テーブル・エントリによって制御されず、第１のネットワーク・ドメイン内のネットワーク・スイッチの少なくとも所与のものが、カプセル化エンジンが実装されるネットワーク・スイッチとローカル・コア・ネットワークとの間に置かれ、フロー・テーブル・エントリを生成することは、ネットワーク・スイッチに、カプセル化されたパケットを、ネットワーク・スイッチの所与のものを通じて転送するように指示するフロー・テーブル・エントリを生成することを含む。

【０１０２】

別の実施形態によれば、ローカル・コアは、コントローラ・サーバからのフロー・テーブル・エントリをロードできないネットワーク・スイッチを含み、この方法は、コントローラ・サーバからのフロー・テーブル・エントリをネットワーク・スイッチにおける対応するコントローラ・クライアントに与えることをさらに含む。

【０１０３】

別の実施形態によれば、フロー・テーブル・エントリを与えることは、コントローラ・サーバ及びコントローラ・クライアントにおけるネットワーク・プロトコル・スタックを用いて、ネットワーク接続上で前記フロー・テーブル・エントリを伝達することを含む。

【０１０４】

別の実施形態によれば、前記ネットワーク・プロトコル・スタックを用いることは、伝送制御プロトコル（ＴＣＰ）／インターネット・プロトコル（ＩＰ）スタックを用いて、フロー・テーブル・エントリを伝達することを含む。

【０１０５】

別の実施形態によれば、フロー・テーブル・エントリは、ヘッダ・フィールド及びアクション・フィールドを含み、フロー・テーブル・エントリを生成することは、ネットワーク・スイッチに、パケットをネットワーク・スイッチの少なくとも所与のものを通じて２度転送するように指示するフロー・テーブル・エントリを生成することを含む。

【０１０６】

別の実施形態によれば、この方法は、カプセル化エンジンによって、マルチプロトコル・ラベル・スイッチング（Ｍｕｌｔｉｐｒｏｔｏｃｏｌ　Ｌａｂｅｌ　Ｓｗｉｔｃｈｉｎｇ）を用いて、カプセル化されたパケットを生成することをさらに含む。

【０１０７】

別の実施形態によれば、この方法は、総称ルーティング・カプセル化（Ｇｅｎｅｒｉｃ　Ｒｏｕｔｉｎｇ　Ｅｎｃａｐｓｕｌａｔｉｏｎ）を用いて、カプセル化されたパケットを生成することをさらに含む。

【０１０８】

１つの実施形態によれば、デフォルトのインターネット・ゲートウェイによりインターネットに結合されたネットワークを動作させるための方法が提供される。ネットワークはエンド・ホストをデフォルトのインターネット・ゲートウェイに結合するネットワーク・スイッチを含む。この方法は、ワイルドカード指定されていない物理ポートのエントリを含むフロー・テーブル・エントリをネットワーク・スイッチ内のエッジ・スイッチに与えることと、少なくとも所与のフロー・テーブル・エントリをネットワーク・スイッチ内の非エッジ・スイッチに与えることとを含む。非エッジ・スイッチは、デフォルトのインターネット・ゲートウェイに接続されており、所与のフロー・テーブル・エントリは、ワイルドカード指定された物理ポート・フィールドを含み、かつ、部分的にワイルドカード指

定されて、非エッジ・スイッチに、パケットをデフォルトのインターネット・ゲートウェイに転送するように指示する宛先インターネット・プロトコル・アドレス・フィールドを有する。

【０１０９】

別の実施形態によれば、ネットワーク・スイッチの各々は、コントローラ・サーバと通信するコントローラ・クライアントを含み、フロー・テーブル・エントリをエッジ・スイッチに与えることは、コントローラ・サーバ及びコントローラ・クライアントにおけるネットワーク・プロトコル・スタックを用いて、ネットワーク接続上で前記フロー・テーブル・エントリを伝達することを含む。

【０１１０】

別の実施形態によれば、この方法は、各々のネットワーク・スイッチにおいて、受け取ったパケット・フィールドを、そのネットワーク・スイッチに与えられたフロー・テーブル・エントリのフィールドと比較することを含む。

【０１１１】

１つの実施形態によれば、コントローラ・サーバを用いて、ネットワーク内のネットワーク・スイッチにフロー・テーブル・エントリを与える方法が提供され、ここでは、各々のネットワーク・スイッチは、パケット・フィールドをフロー・テーブル・エントリのフィールドと比較し、一致が検出された場合には対応するアクションを取ることによってパケットを処理し、エンド・ホストはネットワーク・スイッチに接続される。この方法は、コントローラ・サーバによって、ネットワーク・スイッチの第１のものが、エンド・ホストに接続された入力・出力ポートを有するエッジ・スイッチであり、ネットワーク・スイッチの第２のものがネットワーク・スイッチの第１のものに接続された入力・出力ポートを有する集約スイッチであると判断することを含む。この方法は、ネットワーク・スイッチの第１のものがエッジ・スイッチであり、ネットワーク・スイッチの第２のものが集約スイッチであるとの判断に応答して、コントローラ・サーバを用いて、第１のネットワーク・スイッチに第１のフロー・テーブル・エントリを与え、第２のネットワーク・スイッチに第２のフロー・テーブル・エントリを与えることをさらに含み、第１のフロー・テーブル・エントリは、ワイルドカード指定のない物理ポート・フィールドを含み、第２のフロー・テーブル・エントリはワイルドカード指定を有する物理ポート・フィールドを含む。

【０１１２】

別の実施形態によれば、第１のネットワーク・スイッチに第１のフロー・テーブル・エントリを与えることは、第１のネットワーク・スイッチに、ワイルドカード指定のないインターネット・プロトコル送信元アドレス・フィールドを含むフロー・テーブル・エントリを与えることを含む。

【０１１３】

別の実施形態によれば、第２のネットワーク・スイッチに第２のフロー・テーブル・エントリを与えることは、第２のネットワーク・スイッチに、ワイルドカード指定されたインターネット・プロトコル送信元アドレス・フィールドを含むフロー・テーブル・エントリを与えることを含む。

【０１１４】

別の実施形態によれば、第１のネットワーク・スイッチはコントローラ・クライアントを含み、第２のネットワーク・スイッチはコントローラ・クライアントを含み、コントローラ・サーバは、ネットワーク・プロトコル・スタックを用いて、ネットワーク接続上でコントローラ・クライアントと通信する。

【０１１５】

別の実施形態によれば、ネットワーク・スイッチの第３のものが、エンド・ホストに接続され、かつ、第２のスイッチに結合され、この方法は、第３のネットワーク・スイッチに、ワイルドカード指定されたインターネット・プロトコル送信元アドレス・フィールドと、ワイルドカード指定のない物理ポート・フィールドとを含む第３のフロー・テーブル

・エントリを与えることを含む。
【０１１６】
別の実施形態によれば、この方法は、第１、第２、及び第３のネットワーク・スイッチのフロー・テーブル・エントリに応答して、パケットを、第１、第２、及び第３のネットワーク・スイッチを通じて、第１のネットワーク・スイッチに接続された第１のエンド・ホストから、第３のネットワーク・スイッチに接続された第２のエンド・ホストに転送することを含む。
【０１１７】
別の実施形態によれば、第１のネットワーク・スイッチに第１のフロー・テーブル・エントリを与えることは、第１のネットワーク・スイッチに、ワイルドカード指定のないイーサネット送信元アドレス・フィールドを含むフロー・テーブル・エントリを与えることを含む。
【０１１８】
別の実施形態によれば、第１のネットワーク・スイッチに第１のフロー・テーブル・エントリを与えることは、第１のネットワーク・スイッチに、ワイルドカード指定のない仮想ローカル・エリア・ネットワーク・タグを含むフロー・テーブル・エントリを与えることを含む。
【０１１９】
別の実施形態によれば、第１のネットワーク・スイッチに第１のフロー・テーブル・エントリを与えることは、第１のネットワーク・スイッチに、ワイルドカード指定のないインターネット・プロトコル送信元アドレス、ワイルドカード指定のないイーサネット送信元アドレス、及びワイルドカード指定のない仮想ローカル・エリア・ネットワーク・タグからなる群から選択されるワイルドカード指定のないアドレスを含むフロー・テーブル・エントリを与えることを含む。本実施形態において、第２のネットワーク・スイッチに第２のフロー・テーブル・エントリを与えることは、第２のネットワーク・スイッチに所与のアドレスを含むフロー・テーブル・エントリを与えることと、パケット・フィールドが所与のアドレスと一致することを検出したことに応答して、第２のネットワーク・スイッチに、第２のネットワーク・スイッチが取るアクションを与えることを含む。
【０１２０】
上記は、本発明の原理を例証するものにすぎず、当業者であれば、本発明の範囲及び趣旨から逸脱することなく、様々な修正を行うことができる。
【符号の説明】
【０１２１】
１０：ネットワーク
１０Ｅ：エッジ部分
１０ＡＤ：集約・分散部分
１０Ｃ：コア部分
１２、４２：コンピューティング機器
１４：ネットワーク・スイッチ
１４Ｅ：エッジ・スイッチ
１４ＡＤ：集約スイッチ
１４Ｃ：コア・スイッチ
１６：ネットワーク・リンク
１８：コントローラ・サーバ
２０：ネットワーク構成規則
２２：ＴＣＰポート
２４：制御ユニット
２４－１：マスター・プロセッサ
２４－２：スレーブ・プロセッサ
２６：パケット処理ソフトウェア

２８：フロー・テーブル
３０：コントローラ・クライアント
３２：パケット処理回路
３４：入力・出力ポート
３８：ネットワーク・インターフェース
４０：パケット処理ソフトウェア
４４：仮想マシン
４８、５０：ライン・カード
５２：バックプレーン
５４、６４：制御ソフトウェア
５６、６２：制御プロトコル・スタック
５８、６０：ネットワーク・プロトコル・スタック
６６：ネットワーク経路
６８：フロー・テーブル・エントリ
７０：ヘッダ
７２：アクション
７４：統計データ
７６：ヘッダ・フィールド
８８：エンド・ホスト
１１０：インターネット
１１２：エンド・ホスト機器
１１４：ローカル・コア
１１６、１１８、１２０：経路
ＥＨ：エンド・ホスト
ＤＧ：デフォルトのインターネット・ゲートウェイ
ＥＥ：カプセル化エンジン
ＤＥ：カプセル化解除エンジン

【図１】

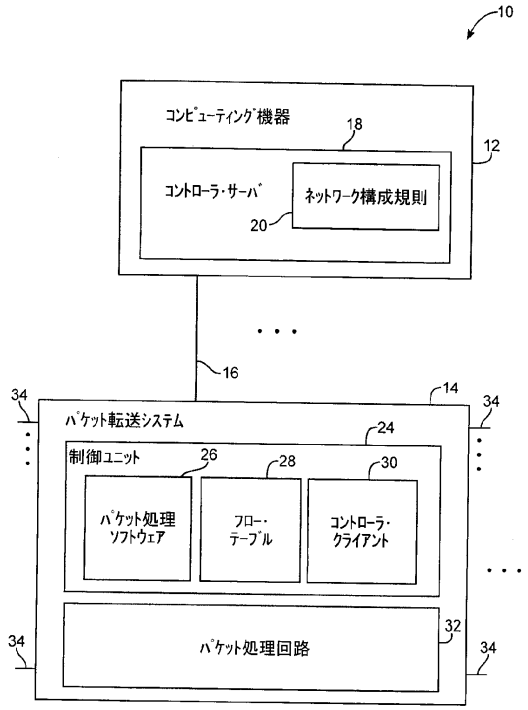

FIG. 1

【図２】

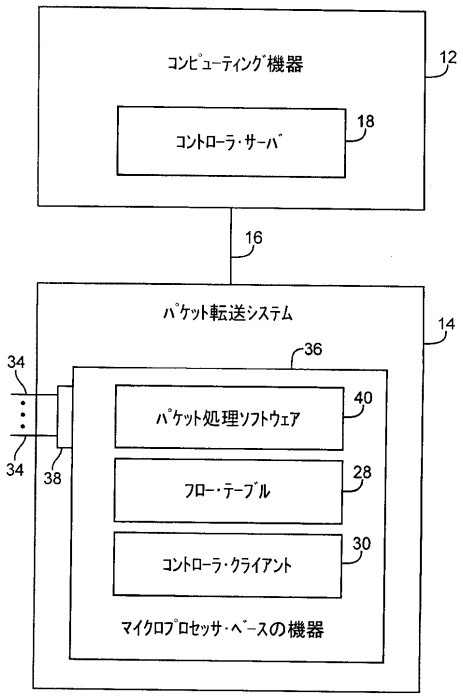

FIG. 2

【図３】

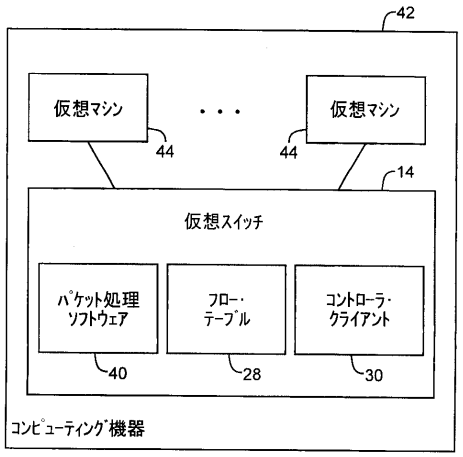

FIG. 3

【図４】

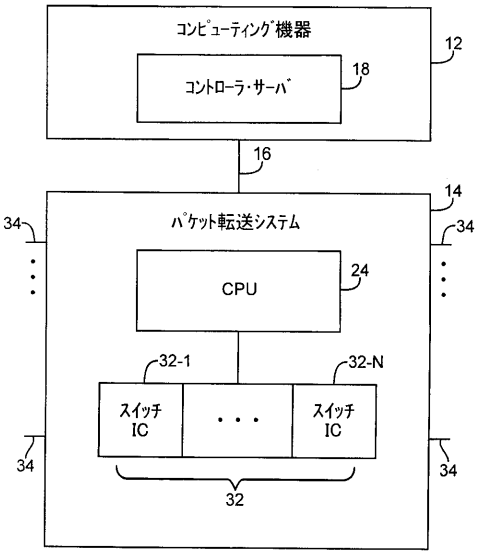

FIG. 4

【図５】

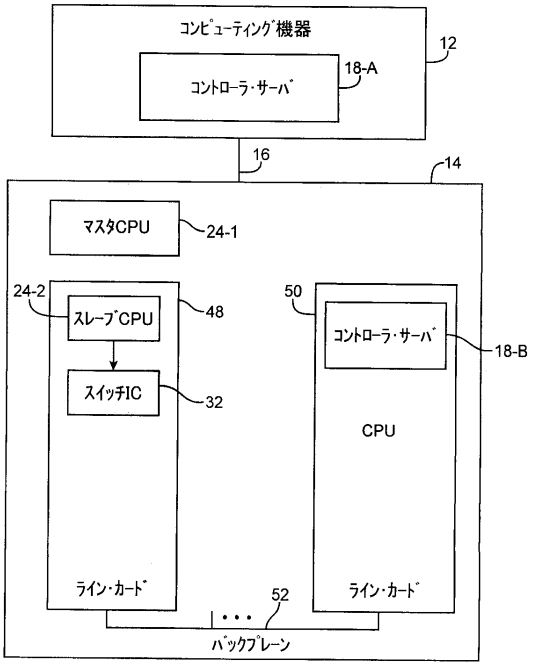

FIG. 5

【図６】

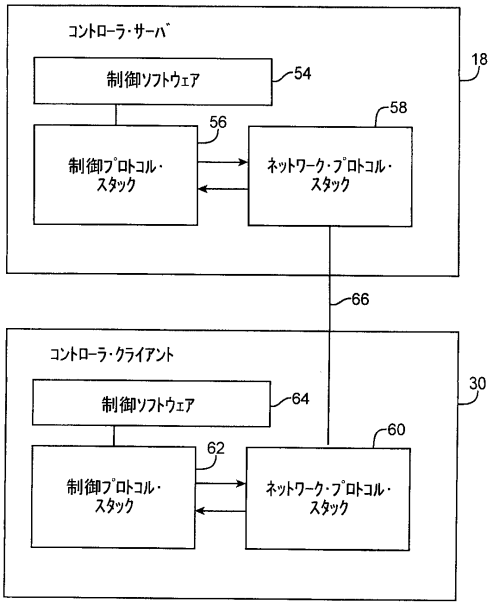

FIG. 6

【図７Ａ】

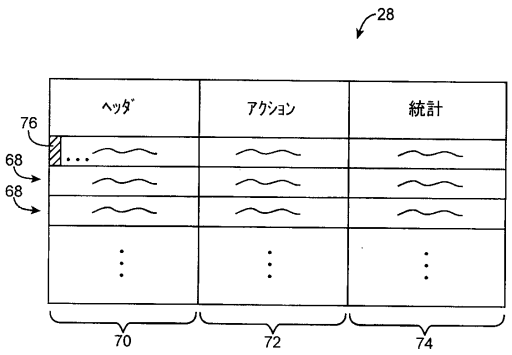

FIG. 7A

【図７Ｂ】

アクション	ポート3に送信	ポート4に送信	ドロップ
宛先TCPポート	*	*	80
送信元TCPポート	*	*	*
宛先IPアドレス	*	172.12.3.4	*
送信元IPアドレス	*	*	*
宛先イーサネット・アドレス	00:1FAB	*	*
送信元イーサネット・アドレス	*	*	*
物理入力ポート	*	*	*

FIG. 7B

【図７Ｃ】

...	宛先IPアドレス	...	アクション
...	172.12.3.4	...	ポート3に送信

FIG. 7C

【図７Ｄ】

...	宛先IPアドレス	...	アクション
...	172.12.3.4	...	ポート5に送信

FIG. 7D

【図８】

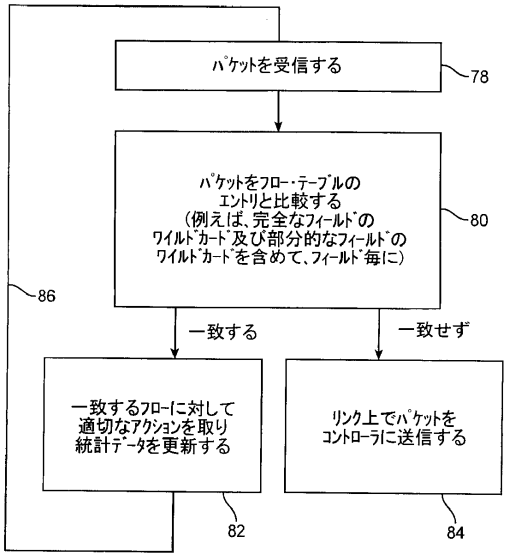

FIG. 8

【図９】

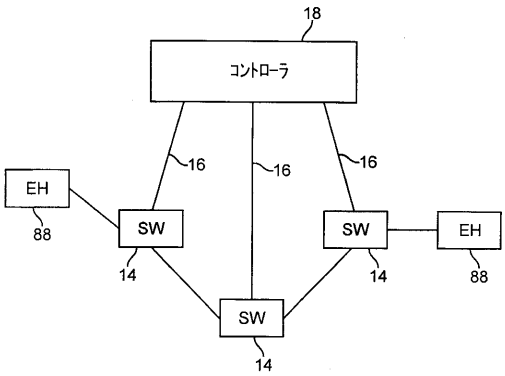

FIG. 9

【図１０】

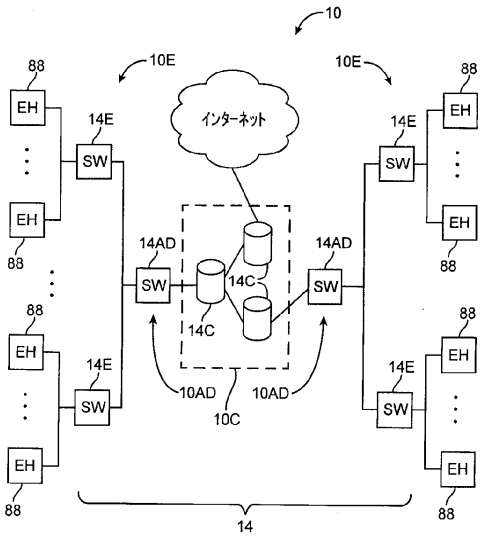

FIG. 10

【図１１】

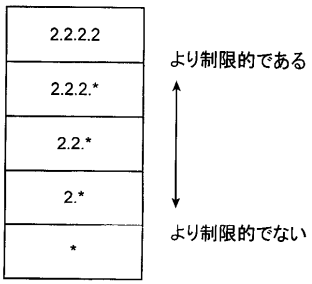

FIG. 11

【図１２】

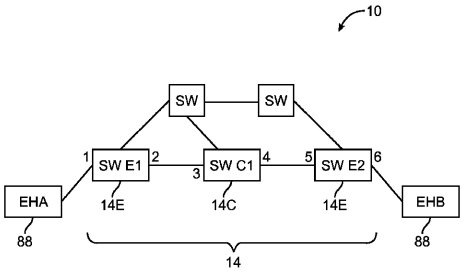

FIG. 12

【図１３】

物理的INﾎﾟｰﾄ	IP送信元ｱﾄﾞﾚｽ	IP宛先ｱﾄﾞﾚｽ	宛先TCPﾎﾟｰﾄ	送信元TCPﾎﾟｰﾄ	ｱｸｼｮﾝ
1	1.1.1.1	2.2.2.2	22	7777	ﾎﾟｰﾄ2に送信
・・・					
3	1.1.1.1	2.2.2.2	22	7777	ﾎﾟｰﾄ4に送信
・・・					
5	1.1.1.1	2.2.2.2	22	7777	ﾎﾟｰﾄ6に送信

E1 C1 E2

（従来技術）

FIG. 13

【図１４】

物理的INﾎﾟｰﾄ	IP送信元ｱﾄﾞﾚｽ	IP宛先ｱﾄﾞﾚｽ	宛先TCPﾎﾟｰﾄ	送信元TCPﾎﾟｰﾄ	ｱｸｼｮﾝ
*	*	2.2.2.2	*	*	ﾎﾟｰﾄ2に送信
・・・					
*	*	2.2.2.2	*	*	ﾎﾟｰﾄ4に送信
・・・					
*	*	2.2.2.2	*	*	ﾎﾟｰﾄ6へ送信

E1' C1' E2'

（従来技術）

FIG. 14

【図１５】

物理的INﾎﾟｰﾄ	IP送信元ｱﾄﾞﾚｽ	IP宛先ｱﾄﾞﾚｽ	宛先TCPﾎﾟｰﾄ	送信元TCPﾎﾟｰﾄ	ｱｸｼｮﾝ
1	1.1.1.1	2.2.2.2	22	7777	ﾎﾟｰﾄ2に送信
・・・					
*	*	2.2.2.2	*	*	ﾎﾟｰﾄ4に送信
・・・					
5	1.1.1.1	2.2.2.2	22	7777	ﾎﾟｰﾄ6へ送信

E1' C1' E2'

FIG. 15

【図１６】

物理的INﾎﾟｰﾄ	IP送信元ｱﾄﾞﾚｽ	IP宛先ｱﾄﾞﾚｽ	宛先TCPﾎﾟｰﾄ	送信元TCPﾎﾟｰﾄ	ｱｸｼｮﾝ
1	1.1.1.1	2.2.2.2	*	*	ﾎﾟｰﾄ2へ送信
・・・					
*	*	2.2.2.2	*	*	ﾎﾟｰﾄ4へ送信
・・・					
5	*	2.2.2.2	22	*	ﾎﾟｰﾄ6へ送信

E1'' C1'' E2''

FIG. 16

【図１７】

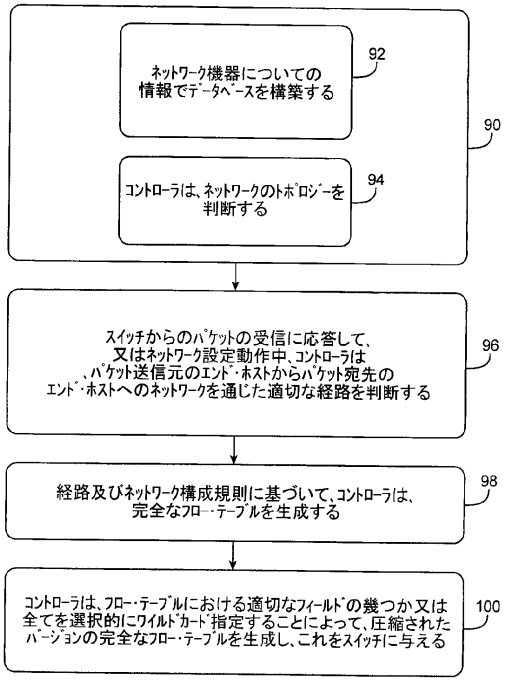

FIG. 17

【図１８】

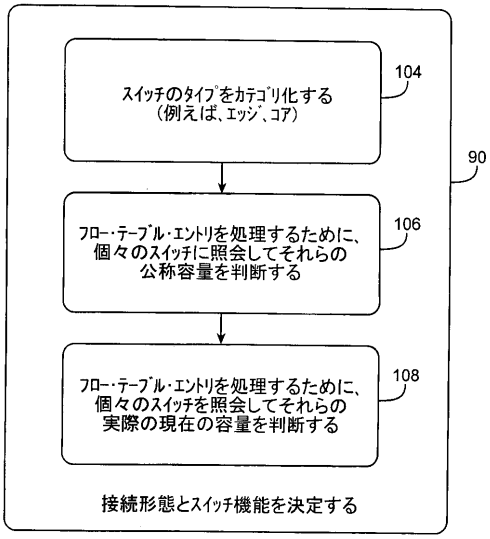

FIG. 18

【図１９】

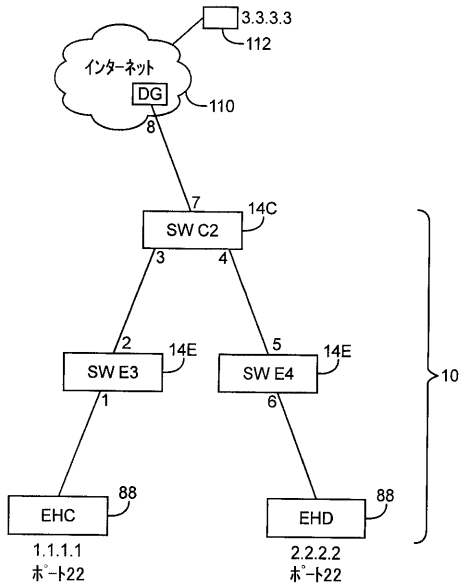

FIG. 19

【図２０】

物理的INポート	IP送信元アドレス	IP宛先アドレス	宛先TCPポート	送信元TCPポート	アクション
1	1.1.1.1	3.3.3.3	22	7777	ポート2に送信
・・・					
*	*	3.*	*	*	ポート7に送信
・・・					

E3 C2

FIG. 20

【図２１】

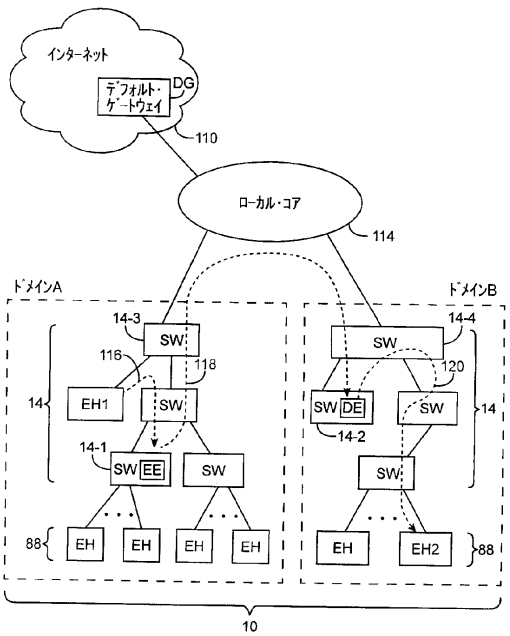

FIG. 21

【図２２】

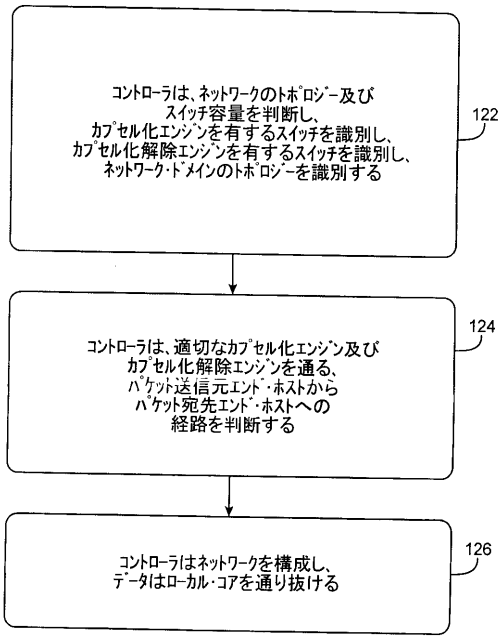

FIG. 22

【図２３】

エッジ

ポート4	アドレス(例えば、送信元IPアドレス、イーサネット送信元アドレス、VLANタグ、アドレスとイーサネット・アドレス及びVLANタグのような他のフィールドとの組み合わせ等)	対応するアクション

非エッジ

ポート=*	アドレス	対応するアクション

FIG. 23

フロントページの続き

(74)代理人 100109070
弁理士 須田 洋之
(74)代理人 100109335
弁理士 上杉 浩
(72)発明者 アッペンツェラー グイド
アメリカ合衆国 カリフォルニア州 ９４３０１－２３２９ パロ アルト エル カミノ リアル ８５５ スイート ２６０

審査官 上田 翔太

(56)参考文献 千葉靖伸, 他, フローベースネットワークにおけるフローエントリ削減手法の提案とＯｐｅｎＦｌｏｗネットワークへの適用, 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, ２０１０年 2月２５日, ネットワークシステム 109(448), p.7-12
LingYun Lu, 他, OpenFlow control for cooperating AQM scheme, Signal Processing (ICSP), 2010 IEEE 10th International Conference on, ２０１０年１０月２４日, p.2560-2563
Brandon Heller, 他, ElasticTree: saving energy in data center networks, NSDI'10 Proceedings of the 7th USENIX conference on Networked systems design and implementation, ２０１０年 4月２８日, p.1-16

(58)調査した分野(Int.Cl., ＤＢ名)
Ｈ０４Ｌ １２／７１７