

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5131651号
(P5131651)

(45) 発行日 平成25年1月30日(2013.1.30)

(24) 登録日 平成24年11月16日(2012.11.16)

(51) Int.Cl.

H04L 12/801 (2013.01)
H04L 12/911 (2013.01)

F 1

H04L 12/56 200Z

請求項の数 12 (全 17 頁)

(21) 出願番号 特願2011-543216 (P2011-543216)
 (86) (22) 出願日 平成22年11月18日 (2010.11.18)
 (86) 国際出願番号 PCT/JP2010/070527
 (87) 国際公開番号 WO2011/065268
 (87) 国際公開日 平成23年6月3日 (2011.6.3)
 審査請求日 平成24年5月9日 (2012.5.9)
 (31) 優先権主張番号 特願2009-269005 (P2009-269005)
 (32) 優先日 平成21年11月26日 (2009.11.26)
 (33) 優先権主張国 日本国 (JP)

(出願人による申告) 平成21年度 総務省「セキュアクラウドネットワーキング技術の研究開発」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

(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 100102864
 弁理士 工藤 実
 (72) 発明者 加藤 剛史
 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 玉木 宏治

最終頁に続く

(54) 【発明の名称】負荷分散システム、負荷分散方法、及びプログラム

(57) 【特許請求の範囲】

【請求項1】

ネットワークを構成するスイッチと、
 スイッチに対して経路を設定するコントローラと、
 1つのスイッチからの接続を複数のコントローラに通知し、スイッチからの問い合わせ
 メッセージを、マスタとなる1つのコントローラに転送するプロキシと
 を具備する

負荷分散システム。

【請求項2】

請求項1に記載の負荷分散システムであって、
 前記プロキシは、1つのスイッチからプロトコルのセキュアチャネル接続を受けると、
 マスタとなる1つのコントローラを接続先として決定し、前記マスタとなるコントローラ
 に対してセキュアチャネル接続を行い、前記1つのスイッチと前記マスタとなるコントローラとの間の接続を確立する

負荷分散システム。

【請求項3】

請求項1又は2に記載の負荷分散システムであって、
 前記プロキシは、スイッチの1つの接続セッションに対して、複数のコントローラからの
 経路情報登録メッセージを転送する

負荷分散システム。

10

20

【請求項 4】

請求項 1 乃至 3 のいずれか一項に記載の負荷分散システムであって、

前記プロキシは、取り扱い方法の不明なパケットを受信したスイッチからの問い合わせメッセージを、マスタとなる 1 つのコントローラに転送し、問い合わせメッセージの応答として、前記マスタとなるコントローラから経路情報登録メッセージを受けると、経路情報登録メッセージの転送先となるスイッチを決定し、前記転送先となる全てのスイッチに経路情報登録メッセージを転送する

負荷分散システム。

【請求項 5】

請求項 1 乃至 4 のいずれか一項に記載の負荷分散システムであって、

10

前記プロキシは、スイッチとコントローラとの対応関係を記憶し、スイッチとコントローラを監視し、どちらかに障害が発生したことを検出すると、スイッチとコントローラとの対応関係を変更する

負荷分散システム。

【請求項 6】

請求項 1 乃至 5 のいずれか一項に記載の負荷分散システムで使用されるプロキシ。

【請求項 7】

コントローラにおいて、ネットワークを構成するスイッチに対して経路を設定することと、

20

プロキシにおいて、1 つのスイッチからの接続を複数のコントローラに通知することと、

前記プロキシにおいて、スイッチからの問い合わせメッセージを、マスタとなる 1 つのコントローラに転送することと

を含む

負荷分散方法。

【請求項 8】

請求項 7 に記載の負荷分散方法であって、

前記プロキシにおいて、1 つのスイッチからプロトコルのセキュアチャネル接続を受けた際に、マスタとなる 1 つのコントローラを接続先として決定することと、

30

前記プロキシにおいて、前記マスタとなるコントローラに対してセキュアチャネル接続を行い、前記 1 つのスイッチと前記マスタとなるコントローラとの間の接続を確立することと

を更に含む

負荷分散方法。

【請求項 9】

請求項 7 又は 8 に記載の負荷分散方法であって、

前記プロキシにおいて、スイッチの 1 つの接続セッションに対して、複数のコントローラからの経路情報登録メッセージを転送すること

を更に含む

負荷分散方法。

40

【請求項 10】

請求項 7 乃至 9 のいずれか一項に記載の負荷分散方法であって、

前記プロキシにおいて、取り扱い方法の不明なパケットを受信したスイッチからの問い合わせメッセージを、マスタとなる 1 つのコントローラに転送することと、

前記プロキシにおいて、問い合わせメッセージの応答として、前記マスタとなるコントローラから経路情報登録メッセージを受けた際に、経路情報登録メッセージの転送先となるスイッチを決定することと、

前記プロキシで、前記転送先となる全てのスイッチに経路情報登録メッセージを転送することと

を更に含む

50

負荷分散方法。

【請求項 1 1】

請求項 7 乃至 10 のいずれか一項に記載の負荷分散方法であって、
前記プロキシで、スイッチとコントローラとの対応関係を保持することと、
前記プロキシで、スイッチとコントローラを監視し、どちらかに障害が発生したことを
検出すると、スイッチとコントローラとの対応関係を変更することと
を更に含む

負荷分散方法。

【請求項 1 2】

請求項 7 乃至 11 のいずれか一項に記載の負荷分散方法をプロキシに実行させるための
プログラムを格納する記憶媒体。 10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、負荷分散システムに関し、特にスイッチを監視・制御するコントローラが存在するネットワークにおける負荷分散システムに関する。

【背景技術】

【0002】

例えば、サーバ等のコントローラでネットワーク内のスイッチを監視・制御することとで
ネットワーク内を流れるデータフローを制御する技術は、オープン化されたネットワーク
技術の一つであり、大規模なネットワークの制御に適している。 20

【0003】

上記のような技術では、1つのコントローラが、ネットワークを制御するためにネット
ワークに属する全てのスイッチを管理下に置く必要があるため、大規模なネットワークと
なるにつれて、コントローラにデータフロー制御の負荷が集中する。更に、コントローラ
上では、ネットワーク監視ツール等の様々なアプリケーションが動作することになると考
えられるため、コントローラ自身の処理による負荷も増加すると考えられる。

【0004】

上記のような技術には、コントローラの負荷を制御する仕組みが無く、コントローラの
負荷を制御する仕組みを独自に実装すると、オープン化されたネットワーク技術を利用する
ことの優位性が失われる。 30

【0005】

また、上記のような技術では、1つのコントローラが全てのスイッチを制御するという
特徴から、スイッチの障害には柔軟に対応できる半面で、コントローラに障害が発生する
と全てのスイッチが制御できなくなるという弱点がある。

【0006】

しかし、複数のコントローラによりネットワークの管理を行うようとする場合、スイッ
チとコントローラのみの構成では、互換性や同期性の問題から、対応ソフトウェアやネッ
トワークの設計が複雑になる。

【0007】

関連する技術として、特開 2007-288711 号公報（特許文献 1）にゲートウェイ装置、設定コントローラ、ゲートウェイ装置の負荷分散方法及びプログラムが開示され
ている。この関連技術では、ゲートウェイ装置は、複数のネットワーク（NW）間でやり
取りされるパケットに対して、ゲートウェイコントローラ（G C）によって設定されたポ
リシーに従った処理を行うことにより、個々のネットワーク（NW）の運用ポリシーの差
異を吸収する機能を有している。このゲートウェイ装置は、設定コントローラと、2台の
分配ルータと、2台のスイッチングハブと、複数台のセッションボーダーコントローラ（
S B C）とから構成されている。

【0008】

また、特許第 3409726 号公報（特許文献 2）に転送先決定処理装置が開示されて 40

50

いる。この関連技術では、受信したIP (Internet Protocol) データグラムについて、フロー識別情報と、宛先IPアドレスを抽出すると、当該IPデータグラムについての転送先が、マルチパスとして設定されている場合において、フロー管理部は、マルチパス番号 (N) とフロー識別情報 (F) が入力されると、集約フローテーブルのみを参照し、転送パス (P) の決定を行う。

【0009】

また、特表2008-539643号公報（特許文献3）に通信ネットワーク内の複数のネットワーク要素間のセキュアな通信を確立するための方法が開示されている。この関連技術では、ゲートウェイとホストとの間に、セキュアチャネルSCが備えられる。加えて、他のセキュアチャネルSCが、アクセスコントローラとゲートウェイとの間に備えられる。この関連技術では、ゲートウェイを介してホストによって、セキュアなピアツーピア通信が確立される。

【先行技術文献】

【特許文献】

【0010】

【特許文献1】特開2007-288711号公報

【特許文献2】特許第3409726号公報

【特許文献3】特表2008-539643号公報

【非特許文献】

【0011】

【非特許文献1】“The OpenFlow Switch Consortium”
<http://www.openflowswitch.org/>

【非特許文献2】“OpenFlow Switch Specification Version 0.9.0 (Wire Protocol 0x98) July 20, 2009 Current Maintainer: Brandon Heller (brandonh@stanford.edu)” <http://www.openflowswitch.org/documents/openflow-spec-v0.9.0.pdf>

【発明の概要】

【0012】

サーバ等のコントローラでネットワーク内のスイッチにパケットの配達経路を動的に設定することでネットワーク内を流れるデータフローを制御するシステムにおいて、スイッチとコントローラとの間にプロトコルを中継するプロキシ (Proxy) を設置する。プロキシは、各スイッチからは単一のコントローラに見え、各コントローラからはネットワークの全てのスイッチに接続しているかのように振る舞うものである。

【0013】

本発明の負荷分散システムは、スイッチと、コントローラと、プロキシとを具備する。スイッチは、ネットワークを構成する。コントローラは、スイッチに対して経路を設定する。プロキシは、1つのスイッチからの接続を複数のコントローラに通知し、スイッチからの問い合わせメッセージを、マスタとなる1つのコントローラに転送する。

【0014】

本発明の負荷分散方法では、コントローラにおいて、ネットワークを構成するスイッチに対して経路を設定する。また、プロキシにおいて、1つのスイッチからの接続を複数のコントローラに通知する。また、プロキシにおいて、スイッチからの問い合わせメッセージを、マスタとなる1つのコントローラに転送する。

【0015】

本発明に係るプログラムは、ネットワークを構成するスイッチと、スイッチに対して経路を設定するコントローラとの間に設置されたプロキシにより実行されるプログラムである。このプログラムは、1つのスイッチからの接続を複数のコントローラに通知するステップと、スイッチからの問い合わせメッセージを、マスタとなる1つのコントローラに転

10

20

30

40

50

送するステップとをプロキシに実行させるためのプログラムである。なお、本発明に係るプログラムは、記憶装置や記憶媒体に格納することが可能である。

【0016】

サーバ等のコントローラでネットワーク内のスイッチにパケットの配送経路を動的に設定することでネットワーク内を流れるデータフローを制御するシステムにおいて、プロキシを導入することにより、独自に負荷分散機能を持たないスイッチとコントローラの組み合わせや、メーカーの違いにより負荷分散機能に互換性がないスイッチとコントローラの組み合わせにおいても、コントローラの負荷分散が可能となる。

【図面の簡単な説明】

【0017】

10

【図1】本発明の負荷分散システムの全体構成の例を示す図である。

【図2】本発明の第1実施形態におけるプロキシの構成例を示すブロック図である。

【図3】スイッチの起動時の動作（初期化）を示すフローチャートである。

【図4】初期化の概要を説明するための図である。

【図5】スイッチごとに決定されたマスタコントローラとの対応関係の例を示す図である。

【図6】経路制御の動作を示すフローチャートである。

【図7】フロー登録の概要を説明するための図である。

【図8】本発明の第2実施形態におけるプロキシの構成例を示すブロック図である。

【図9】障害発生後のスイッチとマスタコントローラとの対応関係の例を示す図である。

20

【発明を実施するための形態】

【0018】

本発明では、サーバ等のコントローラでネットワーク内のスイッチを監視・制御することでネットワーク内を流れるデータフローを制御する技術として、オープンフロー（OpenFlow）技術を例にして説明する。但し、実際には、オープンフロー技術に限定されない。

【0019】

オープンフロー技術とは、コントローラが、ルーティングポリシーとして自身に設定されたフロー定義情報（フロー：ルール+アクション）に従い、マルチレイヤ及びフロー単位の経路情報（フローテーブル）をスイッチに設定し、経路制御やノード制御を行う技術を示す。オープンフロー技術では、コントローラが、ネットワーク内のスイッチを監視し、通信状況に応じて、ネットワーク内のスイッチにパケットの配送経路を動的に設定する。これにより、経路制御機能がルータやスイッチから分離され、コントローラによる集中制御によって最適なルーティング、トラフィック管理が可能となる。オープンフロー技術が適用されるスイッチは、従来のルータやスイッチのようにパケットやフレームの単位ではなく、フロー単位で通信を取り扱う。

30

【0020】

フローテーブルとは、所定のマッチ条件（ルール）に適合するパケットに対して行うべき所定の処理（アクション）を定義したエントリ（entry）が登録されたテーブルである。ルールに適合するパケット群（パケット系列）をフローと呼ぶ。フローのルールは、パケットの各プロトコル階層のヘッダ領域に含まれるデスティネーションアドレス（Destination Address：宛先アドレス）、ソースアドレス（Source Address：送信元アドレス）、デスティネーションポート番号、ソースポート番号のいずれか又は全てを用いた様々な組み合わせにより定義され、区別可能である。なお、上記のアドレスには、MACアドレス（Media Access Control Address）やIPアドレス（Internet Protocol Address）を含むものとする。また、上記に加えて、入口ポート（Ingress Port）の情報も、フローのルールとして使用可能である。

40

【0021】

オープンフロー技術の詳細については、非特許文献1、非特許文献2にも記載されてい

50

る。

【0022】

<第1実施形態>

以下に、本発明の第1実施形態について添付図面を参照して説明する。

【0023】

[システム全体構成]

図1に示す通り、本発明の負荷分散システムは、オープンフロープロキシ(OpenFlow Proxy: OFPX)1と、オープンフローコントローラ(OpenFlow Controller: OFC)21、22と、オープンフロースイッチ(OpenFlow Switch: OFS)31～34を含む。

10

【0024】

オープンフロープロキシ(OFPX)1は、オープンフローコントローラ(OFC)21、22と、オープンフロースイッチ(OFS)31～34との間の通信を中継するプロキシである。オープンフロープロキシ(OFPX)1の例として、プロキシサーバ(proxy server)、ゲートウェイ(gateway)、ファイアウォール(firewall)、又はこれらに相当する計算機や中継装置を想定している。但し、実際には、これらの例に限定されない。

【0025】

オープンフローコントローラ(OFC)21、22は、オープンフロースイッチ(OFS)31～34を制御・監視し、オープンフロースイッチ(OFS)31～34に対して、パケットの配送経路を設定するサーバである。ここでは、オープンフロー技術を用いたフロースイッチング(Flowswitching)方式での設定を例に説明するが、実際には、送信先アドレス(Destination IP Address)ベースの静的ルーティング(Static Routing)方式や、MPLS(Multi Protocol Label Switching)ベースのパスルーティング(Path Routing)方式での設定でも良い。オープンフローコントローラ(OFC)21、22の例として、PC(パソコン)、シンクライアントサーバ、ワークステーション、メインフレーム、スーパーコンピュータ等の計算機を想定している。但し、実際には、これらの例に限定されない。

20

【0026】

オープンフロースイッチ(OFS)31～34は、ネットワークを構成し、受信したパケットを、設定された配送経路で配送するスイッチである。オープンフロースイッチ(OFS)31～34の例として、ネットワークスイッチ、マルチレイヤスイッチ(multi-layer switch)等が考えられる。マルチレイヤスイッチは、サポートするOSI参照モデルの層毎に、更に細かく分類されている。主な種別としては、ネットワーク層(第3層)のデータを読むレイヤ3スイッチ、トランスポート層(第4層)のデータを読むレイヤ4スイッチ、アプリケーション層(第7層)のデータを読むレイヤ7スイッチ(アプリケーションスイッチ)がある。オープンフロースイッチ(OFS)31～34は、少なくともレイヤ3スイッチの機能を備えているものとする。なお、オープンフローシステムでは、一般的なルータやスイッチングハブ等の中継装置をオープンフロースイッチ(OFS)として使用することもできる。但し、実際には、これらの例に限定されない。

30

【0027】

なお、図示しないが、オープンフロースイッチ(OFS)31～34の各々の配下には、サーバや各種のネットワーク対応機器が存在している場合がある。例えば、オープンフロースイッチ(OFS)31～34のそれぞれがサーバラック(Server Rack)に設置されている場合が考えられる。このような場合、オープンフロースイッチ(OFS)31～34の各々の配下のサーバは、論理構成において、仮想マシン(Virtual Machine: VM)や、仮想マシンモニタ(Virtual Machine Monitor: VMM)を備えている場合がある。上記のサーバや仮想マシンがオープ

40

50

ンフロースイッチ (OFS) 31～34を介してオープンフロープロキシ (OFPX) 1と通信する場合、直接的には、オープンフロースイッチ (OFS) 31～34がオープンフロープロキシ (OFPX) 1と通信することになる。

【0028】

[構成要素の詳細]

図2に示す通り、オープンフロープロキシ (OFPX) 1は、データ処理装置11と、記憶装置12と、ネットワーク処理装置13を備える。

【0029】

データ処理装置11は、問い合わせ処理部111と、フロー処理部112を備える。

【0030】

問い合わせ処理部111は、オープンフロープロキシ (OFPX) 1がオープンフロースイッチ (OFS) からの問い合わせメッセージを受信すると起動し、当該オープンフロースイッチ (OFS) のマスターOFCだけに当該オープンフロースイッチ (OFS) からの問い合わせメッセージを転送する。

【0031】

フロー処理部112は、オープンフロープロキシ (OFPX) 1がオープンフローコントローラ (OFC) から各OFS宛のフロー登録メッセージ (経路情報登録メッセージ) を受信すると起動し、フロー登録メッセージの宛先となるOFSに対して確立されたセキュアチャネルを使い、フロー登録メッセージを送信する。

【0032】

データ処理装置11の例として、マイクロプロセッサ (microprocessor) 、マイクロコントローラ、或いは、同様の機能を有するIC (Integrated Circuit: 半導体集積回路) 等が考えられる。但し、実際には、これらの例に限定されない。

【0033】

記憶装置12は、OFC記憶部121と、OFS記憶部122と、管理関係記憶部123を備える。

【0034】

OFC記憶部121は、全てのOFCのIPアドレスを記憶する。

【0035】

OFS記憶部122は、全てのOFSのIPアドレスを記憶する。

【0036】

管理関係記憶部123は、OFCが管理するOFSを記憶する。

【0037】

なお、IPアドレスは例示に過ぎない。実際には、ネットワーク上でOFCやOFSを特定可能な識別情報であれば良い。また、記憶装置12は、必要に応じて、データ処理装置11に所定の処理を実行させるためのプログラムを記憶する。

【0038】

記憶装置12の例として、RAM (Random Access Memory) 、ROM (Read Only Memory) 、EEPROM (Electrically Erasable and Programmable Read Only Memory) やフラッシュメモリ等の半導体記憶装置、HDD (Hard Disk Drive) やSSD (Solid State Drive) 等の補助記憶装置、又は、DVD (Digital Versatile Disk) やメモリカード等の記憶媒体 (メディア) 等が考えられる。但し、実際には、これらの例に限定されない。

【0039】

ネットワーク処理装置13は、ネットワークを介して、外部とデータの送受信を行う。ここでは、ネットワーク処理装置13は、オープンフロースイッチ (OFS) からの問い合わせメッセージを受信すると、問い合わせ処理部111を起動する。また、ネットワーク処理装置13は、オープンフローコントローラ (OFC) から各OFS宛のフロー登録

10

20

30

40

50

メッセージを受信すると、フロー処理部112を起動する。

【0040】

ネットワーク処理装置13の例として、NIC(Network Interface Card)等のネットワークアダプタや、アンテナ等の通信装置、接続口(コネクタ)等の通信ポート等が考えられる。また、ネットワークの例として、インターネット、LAN(Local Area Network)、無線LAN(Wireless LAN)、WAN(Wide Area Network)、バックボーン(Backbone)、ケーブルテレビ(CATV)回線、固定電話網、携帯電話網、WiMAX(IEEE 802.16a)、3G(3rd Generation)、専用線(leased line)、IrDA(Infra red Data Association)、Bluetooth(登録商標)、シリアル通信回線、データバス等が考えられる。但し、実際には、これらの例に限定されない。
10

【0041】

[動作の説明]

次に、本発明の負荷分散システムの動作について詳細に説明する。

【0042】

[前提条件]

本発明を実施する準備として、以下の条件を満たしている必要がある。

- 各OFSには、OFCのIPアドレスの代わりにオープンフロープロキシ(OFPX)1のIPアドレスを登録してあること。
20
- オープンフロープロキシ(OFPX)1には、予めオープンフローコントローラ(OFC)21とオープンフローコントローラ(OFC)22のIPアドレスをOFC記憶部に登録してあること。

【0043】

[スイッチの起動時の動作(初期化)]

最初に、図3を参照して、スイッチの起動時の動作を説明する。

【0044】

(1)ステップS101

オープンフロースイッチ(OFS)31が起動すると、オープンフロースイッチ(OFS)31は、予めOFCのアドレスとして記憶しているIPアドレスへオープンフローブロトコルのセキュアチャネル接続(Second Channel接続)を行う。ここでは、オープンフロースイッチ(OFS)31の接続先は、オープンフロープロキシ(OFPX)1となる。すなわち、オープンフロースイッチ(OFS)31は、オープンフロープロキシ(OFPX)1のIPアドレスを、OFCのアドレスとして記憶している。
30

【0045】

(2)ステップS102

オープンフロープロキシ(OFPX)1は、オープンフロースイッチ(OFS)31からのセキュアチャネル接続を受けると、OFS記憶部122にオープンフロースイッチ(OFS)31の情報(IPアドレス等)を記憶する。また、オープンフロープロキシ(OFPX)1は、OFC記憶部121に記憶しているOFCの中から、オープンフロースイッチ(OFS)31のマスタOFCとなるOFCを決定し、オープンフロースイッチ(OFS)31と決定されたマスタOFCとの対応関係を管理関係記憶部123に記憶する。ここでは、オープンフロースイッチ(OFS)31のマスタOFCとして、オープンフローコントローラ(OFC)21が選択されたこととする。
40

【0046】

(3)ステップS103

オープンフロープロキシ(OFPX)1は、オープンフローコントローラ(OFC)21とオープンフローコントローラ(OFC)22に対して、オープンフロースイッチ(OFS)31からの接続に見せかけたオープンフローブロトコルのセキュアチャネル接続(Second Channel接続)を行い、オープンフロースイッチ(OFS)31のオープンフロー
50

プロトコル接続を確立する。

【0047】

(4) ステップS104

同様にして、オープンフロープロキシ(OFPX)1は、全てのOFSのオープンフロープロトコル接続の確立を行う。すなわち、図4に示すように、オープンフロープロキシ(OFPX)1は、オープンフロースイッチ(OFS)31と同様に、オープンフロースイッチ(OFS)32、オープンフロースイッチ(OFS)33、オープンフロースイッチ(OFS)34のオープンフロープロトコル接続を確立する。ここでは、オープンフロープロキシ(OFPX)1は、オープンフローコントローラ(OFC)21とオープンフローコントローラ(OFC)22に対して、オープンフロースイッチ(OFS)32、オープンフロースイッチ(OFS)33、及びオープンフロースイッチ(OFS)34のそれぞれからの接続に見せかけたオープンフロープロトコルのセキュアチャネル接続を行う。

【0048】

(5) ステップS105

オープンフロープロキシ(OFPX)1は、全てのOFSのオープンフロープロトコル接続の確立が完了した後に、OFS記憶部122に、全てのOFSの情報(IPアドレス等)を記憶する。また、オープンフロープロキシ(OFPX)1は、OFC記憶部121に記憶しているOFCの中から、オープンフロースイッチ(OFS)32、オープンフロースイッチ(OFS)33、及びオープンフロースイッチ(OFS)34のそれぞれのマスタOFCとなるOFCを決定し、管理関係記憶部123に、OFSごとに決定されたマスタOFCとの対応関係を記憶する。

【0049】

ここでは、オープンフロープロキシ(OFPX)1は、管理関係記憶部123に、図5に示す対応関係の情報を記憶したものとする。すなわち、オープンフロープロキシ(OFPX)1は、オープンフロースイッチ(OFS)31及びオープンフロースイッチ(OFS)33のマスタOFCをオープンフローコントローラ(OFC)21とし、オープンフロースイッチ(OFS)32及びオープンフロースイッチ(OFS)34のマスタOFCをオープンフローコントローラ(OFC)22として、管理関係記憶部123に記憶する。

【0050】

[経路制御の動作]

次に、図6を参照して、経路制御の動作について説明する。

【0051】

(1) ステップS201

オープンフロースイッチ(OFS)31は、取り扱い方法の不明なパケットを受信した場合、オープンフロープロトコルに準拠したネットワーク経由でオープンフロープロキシ(OFPX)1へ、当該パケットの取り扱い方法を問い合わせるための問い合わせメッセージを送信する。なお、取り扱い方法の不明なパケットとは、初めて受信したパケット(first packet)のように、フローテーブルに登録されているエントリのいずれにも適合しない未登録のフローのパケットである。

【0052】

(2) ステップS202

オープンフロープロキシ(OFPX)1のネットワーク処理装置13は、オープンフロースイッチ(OFS)31からの問い合わせメッセージを受信すると、問い合わせ処理部111を起動する。問い合わせ処理部111は、管理関係記憶部123を参照し、オープンフロースイッチ(OFS)31のマスタOFCであるオープンフローコントローラ(OFC)21だけにオープンフロースイッチ(OFS)31からの問い合わせメッセージを転送する。

【0053】

10

20

30

40

50

(3) ステップS203

オープンフローコントローラ(OFC)21は、問い合わせメッセージを受けると、問い合わせを受けたパケットを配送するためのフローを確認する。ここでは、オープンフローコントローラ(OFC)21は、問い合わせを受けたパケットをオープンフロースイッチ(OFS)31 オープンフロースイッチ(OFS)33 オープンフロースイッチ(OFS)34の経路で配送するためのフローを登録する必要があると判断したとする。

【0054】

(4) ステップS204

オープンフローコントローラ(OFC)21は、図7に示すように、オープンフロープロキシ(OFPX)1に対して確立されたオープンフロースイッチ(OFS)31、オープンフロースイッチ(OFS)33、オープンフロースイッチ(OFS)34とのセキュアチャネル接続を使い、各OFSを宛先としたフロー登録メッセージを送信する。なお、実際には、オープンフローコントローラ(OFC)21は、オープンフロープロキシ(OFPX)1に対して、各OFSを宛先としたフロー登録メッセージを一括して送信するようにもしても良い。

10

【0055】

(5) ステップS205

オープンフロープロキシ(OFPX)1のネットワーク処理装置13は、オープンフローコントローラ(OFC)21から各OFS宛のフロー登録メッセージを受信すると、フロー処理部112を起動する。フロー処理部112は、フロー登録メッセージの宛先となるOFSに対して確立されたセキュアチャネルを使い、フロー登録メッセージを送信する。ここでは、フロー処理部112は、図7に示すように、オープンフロースイッチ(OFS)31、オープンフロースイッチ(OFS)33、及びオープンフロースイッチ(OFS)34のそれぞれに対して、フロー登録メッセージを送信する。

20

【0056】

(6) ステップS206

オープンフロースイッチ(OFS)31、オープンフロースイッチ(OFS)33、及びオープンフロースイッチ(OFS)34のそれぞれは、フロー登録メッセージを受信すると、フローを登録し、フローに基づいて、問い合わせたパケットと同じパターンのパケットを転送する。ここでは、オープンフロースイッチ(OFS)31は、問い合わせたパケットと同じパターンのパケットをオープンフロースイッチ(OFS)33に転送する。オープンフロースイッチ(OFS)33は、当該パケットを、オープンフロースイッチ(OFS)34に転送する。

30

【0057】

以降は同じパターンのパケットを各OFSが配送可能になる。

【0058】

同様にして、オープンフロースイッチ(OFS)32が不明パケットを受信した場合は、オープンフロースイッチ(OFS)32からの問い合わせメッセージがオープンフロープロキシ(OFPX)1によってオープンフローコントローラ(OFC)22に転送され、オープンフローコントローラ(OFC)22が必要に応じてフロー登録を行う。

40

【0059】

[セキュアチャネルのセッションの例]

次に、セキュアチャネルのセッションの例を説明する。

【0060】

ここでは、次のように記載を簡略化する。

「OFPX」は、オープンフロープロキシ(OFPX)1を示す。

「OFC」は、オープンフローコントローラ(OFC)21、22を示す。

「OFS」は、オープンフロースイッチ(OFS)31～34を示す。

【0061】

各OFSとOFPX間のセキュアチャネルにおいて、OFSからOFPXへ送信するパ

50

ケットのソースアドレス（送信側アドレス）はOFSのIPアドレスであり、デスティネーションアドレス（受信側アドレス）はOFPXのIPアドレスである。また、OFPXからOFSへ送信するパケットのソースアドレスはOFPXのIPアドレスであり、デスティネーションアドレスはOFSのIPアドレスである。

【0062】

OFPXからOFSへ送信するパケットは、OFCからOFSへ送信するパケットを中継するものであるが、OFSはOFPXとのセキュアチャネルを使用しているため、OFPXはOFCからOFSへ送信されるメッセージのソースアドレスをOFPXのIPアドレスとする必要がある。

【0063】

OFPXと各OFC間のセキュアチャネルにおいて、OFPXからOFCへ送信するパケットのソースアドレスはOFSのIPアドレスであり、デスティネーションアドレスはOFCのIPアドレスである。また、OFCからOFPXへ送信するパケットのソースアドレスはOFCのIPアドレスであり、デスティネーションアドレスはOFSのIPアドレスである。

【0064】

OFPXからOFCへ送信するパケットは、OFSとOFCの通信を中継するものであり、OFCがOFSからのメッセージであることを認識する必要があるため、ソースアドレスはOFSのアドレスとする必要がある。同様に、OFCからOFPXへ送信するパケットは、どのOFS宛のメッセージであるかをOFPXが認識する必要があるため、デスティネーションアドレスはOFSのアドレスとする必要がある。したがって、OFPXはOFCからOFSへの通信におけるゲートウェイとなる必要がある。

【0065】

【実施結果】

本実施形態では、フロー問い合わせ元のOFSごとに配送経路を選択するOFCを決定しており、OFCを負荷分散できる。

【0066】

一方で、各OFSとOFCはオープンフロープロトコルに沿った動作をしており、オープンフロープロキシ（OFPX）を介入させるための特別な処理は不要である。

【0067】

オープンフロープロキシ（OFPX）の処理は、各OFSからの問い合わせメッセージを、対応表を元にOFCへ転送することと、OFCからのメッセージをメッセージの宛先のOFSへ転送するという単純なものであることから、OFPXは安価なハードウェアで実現することができる。

【0068】

本発明によれば、OFS群を複数のOFCで制御可能になる。その理由は、プロキシの介入により、全てのOFSからは単一のOFCが存在するように見え、全てのOFCは全てのOFSとの接続が確立しているように見えるためである。

【0069】

【第2実施形態】

次に、本発明の第2実施形態について添付図面を参照して説明する。

本実施形態では、オープンフロープロキシ（OFPX）1のデータ処理装置11が、更に、生存確認処理部113を含む点を特徴とする。

【0070】

【システム全体構成】

負荷分散システムの全体構成については、図1に示す通りである。

【0071】

【構成要素の詳細】

図8に示す通り、第2実施形態のオープンフロープロキシ（OFPX）1は、データ処理装置11と、記憶装置12と、ネットワーク処理装置13を備える。

10

20

30

40

50

【0072】

記憶装置12とネットワーク処理装置13については、基本的に第1実施形態と同じである。

【0073】

第2実施形態のデータ処理装置11は、問い合わせ処理部111と、フロー処理部112と、生存確認処理部113を備える。

【0074】

問い合わせ処理部111とフロー処理部112については、基本的に第1実施形態と同じである。

【0075】

生存確認処理部113は、オープンフローコントローラ(OFC)21とオープンフローコントローラ(OFC)22を監視し、どちらかに障害が発生したことを検出する。

10

【0076】

ここで、管理関係記憶部123に図5の情報が記憶されている状態でオープンフローコントローラ(OFC)21に障害が発生したとする。生存確認処理部113は、オープンフローコントローラ(OFC)21の障害を検出すると、管理関係記憶部123において、マスタOFCがオープンフローコントローラ(OFC)21となっているエントリのマスタOFCを他のOFCへ変更する。この例では、生存確認処理部113は、オープンフロースイッチ(OFS)31及びオープンフロースイッチ(OFS)33のマスタOFCをオープンフローコントローラ(OFC)21からマスタオープンフローコントローラ(OFC)22へ変更する。この場合、管理関係記憶部123の内容は、図9に示す通りとなる。

20

【0077】

以降は、オープンフローコントローラ(OFC)21へ送られていたオープンフロースイッチ(OFS)31、オープンフロースイッチ(OFS)33の問い合わせメッセージは障害が発生していないオープンフローコントローラ(OFC)22へ送られることとなる。

【0078】

オープンフロープロキシ(OFPX)1は、オープンフローコントローラ(OFC)21の監視を継続し、オープンフローコントローラ(OFC)21の回復を検出したときは管理関係記憶部123を更新し、OFCの負荷分散を再開する。この例では、生存確認処理部113は、オープンフロースイッチ(OFS)31及びオープンフロースイッチ(OFS)33のマスタOFCをオープンフローコントローラ(OFC)22からマスタオープンフローコントローラ(OFC)21へ変更する。

30

【0079】

[実施結果]

第2実施形態では、OFCに障害が発生したときの切り替え動作が、管理関係記憶部に記憶されたOFSごとのマスタOFCの対応関係の更新のみで完了するため、短時間での切り替えが可能である。

【0080】

40

なお、上記の各実施形態は、組み合わせて実施することも可能である。

【0081】

<本発明の利用が考えられる分野>

本発明によれば、大規模ネットワークの性能改善と耐障害性を高める用途に適用できる。

【0082】

<まとめ>

以上のように、本発明の負荷分散システムにおいて、オープンフロープロキシ(OFPX)は、1つのオープンフロースイッチ(OFS)からのオープンフロープロトコル接続を複数のオープンフローコントローラ(OFC)に通知しつつ、オープンフロースイッチ

50

(OFS)からの問い合わせメッセージはマスタとなる1つのオープンフローコントローラ(OFC)のみに転送する。

【0083】

また、オープンフロープロキシ(OFPX)は、オープンフロースイッチ(OFS)の1つのオープンフロープロトコル接続セッションに対して、複数のオープンフローコントローラ(OFC)からのフロー登録メッセージを転送する。

【0084】

上記では、オープンフロー技術を例に説明しているが、無論、本発明は、オープンフロー技術以外の類似技術に対しても適用可能である。

【0085】

10

<付記>

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

【0086】

(付記1)

ネットワークを構成するスイッチと、スイッチに対して経路を設定するコントローラとの間に設置されたプロキシにより実行されるプログラムを格納する記憶媒体であって、

1つのスイッチからの接続を複数のコントローラに通知するステップと、

スイッチからの問い合わせメッセージを、マスタとなる1つのコントローラに転送するステップと

20

をプロキシに実行させるためのプログラムを格納する
記憶媒体。

【0087】

(付記2)

付記1に記載の記憶媒体であって、

1つのスイッチからプロトコルのセキュアチャネル接続を受けた際に、マスタとなる1つのコントローラを接続先として決定するステップと、

前記マスタとなるコントローラに対してセキュアチャネル接続を行い、前記1つのスイッチと前記マスタとなるコントローラとの間の接続を確立するステップと
を更にプロキシに実行させるためのプログラムを格納する

30

記憶媒体。

【0088】

(付記3)

付記1又は2に記載の記憶媒体であって、

スイッチの1つの接続セッションに対して、複数のコントローラからの経路情報登録メッセージを転送するステップ

を更にプロキシに実行させるためのプログラムを格納する
記憶媒体。

【0089】

40

(付記4)

付記1乃至3のいずれかに記載の記憶媒体であって、

取り扱い方法の不明なパケットを受信したスイッチからの問い合わせメッセージを、マスタとなる1つのコントローラに転送するステップと、

問い合わせメッセージの応答として、前記マスタとなるコントローラから経路情報登録メッセージを受けた際に、経路情報登録メッセージの転送先となるスイッチを決定するステップと、

前記転送先となる全てのスイッチに経路情報登録メッセージを転送するステップと
を更にプロキシに実行させるためのプログラムを格納する

記憶媒体。

【0090】

50

(付記 5)

付記 1 乃至 4 のいずれかに記載の記憶媒体であって、
 スイッチとコントローラとの対応関係を保持するステップと、
 スイッチとコントローラを監視し、どちらかに障害が発生したことを検出すると、スイ
 ッチとコントローラとの対応関係を変更するステップと
 を更にプロキシに実行させるためのプログラムを格納する
 記憶媒体。

【 0 0 9 1 】

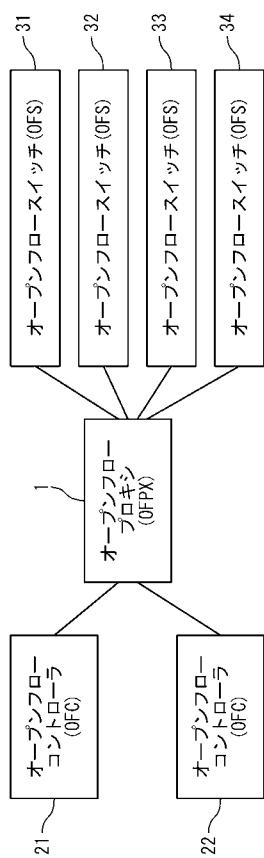
以上、本発明の実施形態を詳述してきたが、実際には、上記の実施形態に限られるもの
 ではなく、本発明の要旨を逸脱しない範囲の変更があっても本発明に含まれる。

10

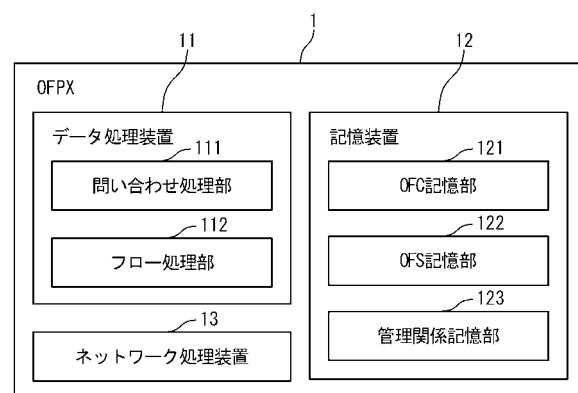
【 0 0 9 2 】

なお、本出願は、日本出願番号 2009-269005 に基づく優先権を主張するもの
 であり、日本出願番号 2009-269005 における開示内容は引用により本出願に組
 み込まれる。

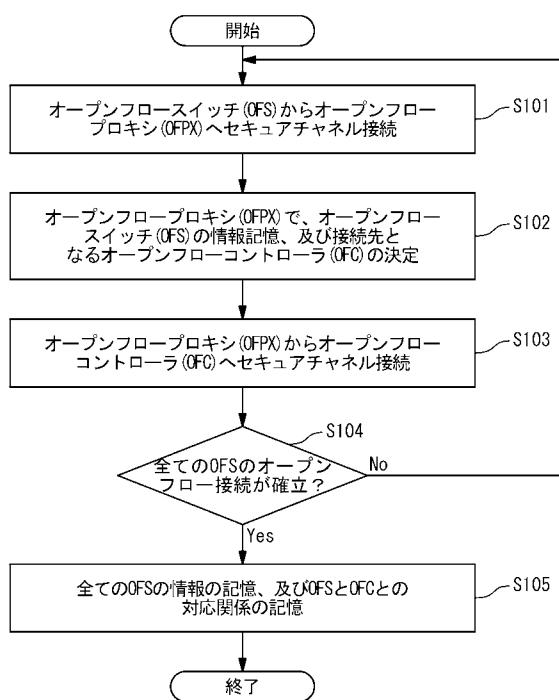
【 図 1 】



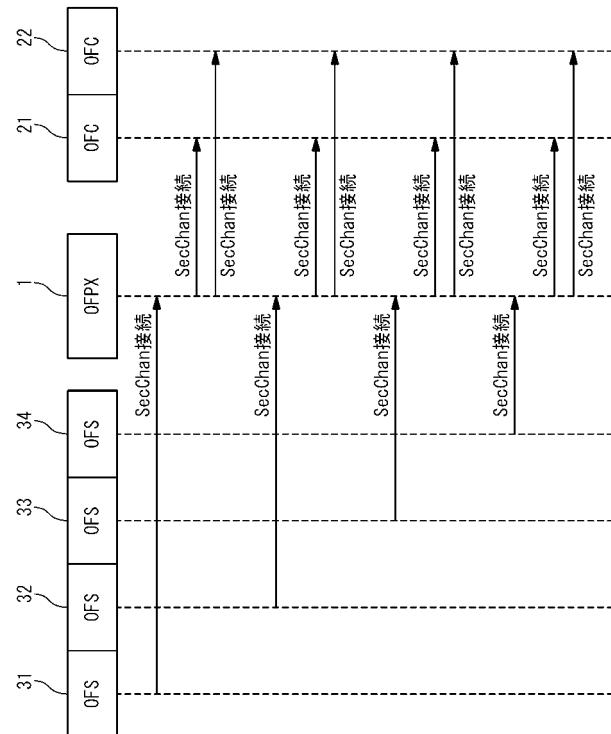
【 図 2 】



【図3】



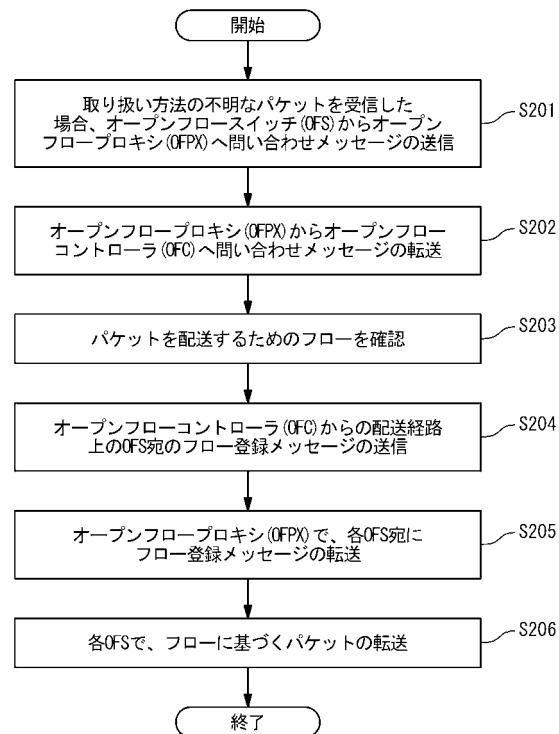
【 図 4 】



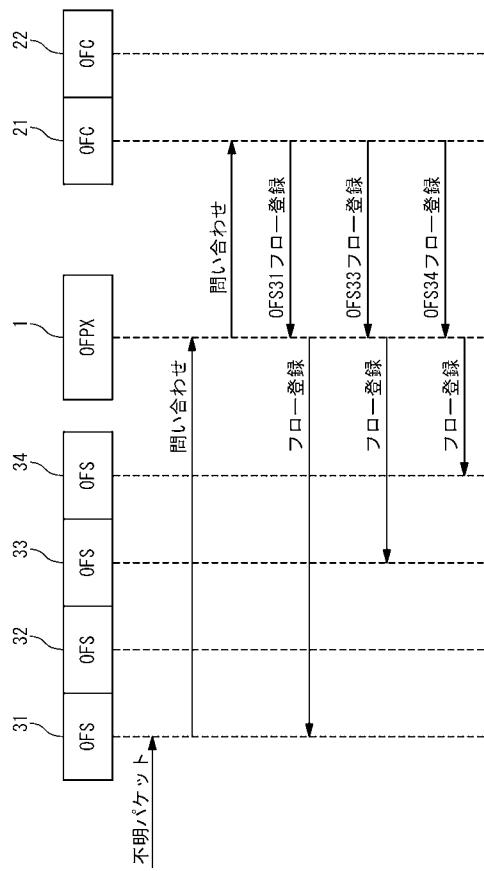
(5)

管理関係記憶部	
OF\$	マスタOF\$
OF\$ 31	OF\$ 21
OF\$ 32	OF\$ 22
OF\$ 33	OF\$ 21
OF\$ 34	OF\$ 22

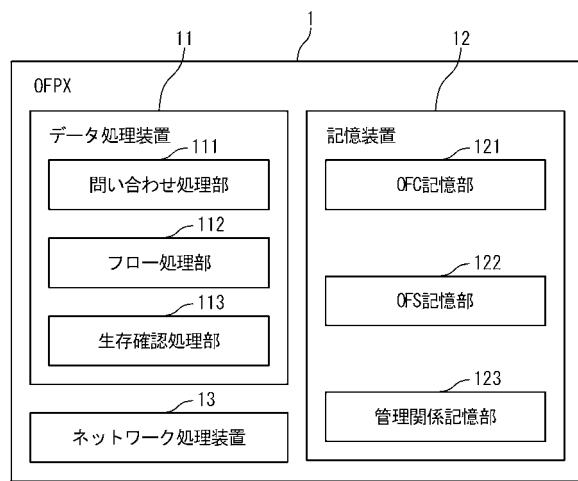
【 6 】



【図7】



【図8】



【図9】

障害発生後の管理関係記憶部

OF S	マスタOFC
OF S 31	OFC 22
OF S 32	OFC 22
OF S 33	OFC 22
OF S 34	OFC 22

フロントページの続き

(56)参考文献 特開2003-229889(JP, A)

Nick McKeown et al. , OpenFlow: Enabling Innovation in Campus Networks , <online> , 20

08年 3月14日 , <http://www.openflowswitch.org/documents/openflow-wp-latest.pdf>

小川 賢太郎 他 , オープンアーキテクチャルータにおける制御部冗長構成法 , 電子情報通信学会技術研究報告 (信学技報) NS2005-100 , 2005年10月13日

(58)調査した分野(Int.Cl. , DB名)

H04L 12/00-66