

# 大規模SDNネットワークを実現する OpenFlowコントローラアーキテクチャ

千葉 靖伸 須堯 一志

## 要旨

OpenFlowのような集中制御型のネットワークアーキテクチャを大規模ネットワークに適用する際、制御機能を提供するコントローラのスケーラビリティと信頼性をいかに担保するかが課題となります。本稿では、クライアント・サーバシステムに広く適用されているデザインパターンをネットワーク制御に応用した、新たなコントローラアーキテクチャを提案します。更に、当該アーキテクチャを、データセンターにおいて仮想レイヤ2ネットワークの提供を行うOpenFlowコントローラへ適用することにより、大規模SDNネットワークに適用可能なOpenFlowコントローラを実現可能なことを示します。



OpenFlow／OpenFlowコントローラ／スケーラビリティ／ロードバランサパターン／三階層アーキテクチャ

## 1. はじめに

Software-Defined Networking (SDN) を構成するための要素技術であるOpenFlow<sup>1)</sup>では、制御機能を提供するOpenFlowコントローラをソフトウェアによって実装することにより、OpenFlowスイッチによって構成されたネットワークが提供する機能を柔軟に定義できます。

OpenFlowを実用ネットワークに適用する際には種々の課題があり、特に、大規模ネットワークに適用する際の代表的な課題として、以下の点が挙げられます（図1）。

### （1）コントローラ性能とスケーラビリティ

- 1) 制御可能なスイッチ数に対するスケーラビリティ
- 2) スイッチによって発生し、コントローラへ通知される非同期イベントの処理性能
- 3) コントローラが能動的に行うスイッチ制御処理性能

### （2）信頼性と耐障害性

- 1) コントローラが単一障害点とならないコントローラ・システム設計が必要
- 2) コントローラとスイッチ間の制御チャネルの可用性確保と障害時にも運用継続可能なシステム設計が必要

本稿では上記課題のうち、主に性能とスケーラビリティに注目しその解決方法を示します。具体的には、大規模デー

タセンターにおける顧客に対する仮想レイヤ2ネットワークの提供を想定し、クライアント・サーバシステムに広く適用されているデザインパターンをOpenFlowに応用することにより、課題解決を試みました。以下、その解決方法、及び、実装に基づく性能評価結果について報告します。

## 2. 提案手法

### 2.1 前提条件

OpenFlowコントローラにより提供されるネットワーク機能、及び、当該ネットワーク機能を実現するためのコントローラの動作は、ソフトウェア実装により自由に定義できます。

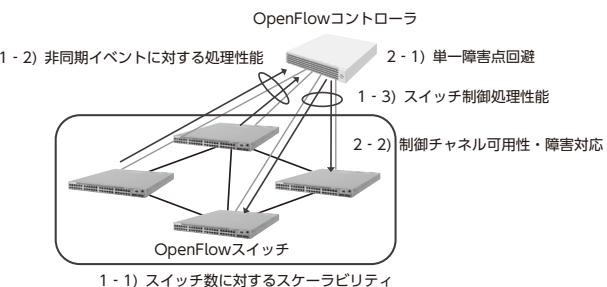


図1 OpenFlowを大規模ネットワークへ適用する際の課題

そのため、想定されるコントローラの動作を一般化し、全てのコントローラに対して適用可能なアーキテクチャを決定することは不可能です。そこで、本検討では、大規模データセンターにおけるマルチテナント環境において、仮想レイヤ2ネットワークの提供を行うコントローラを想定し、コントローラ動作を以下のように想定しました。

- (1) スイッチの制御は、主にコントローラ外から与えられた設定・条件に基づきコントローラが能動的・主体的に行う
- (2) スイッチが発生させる非同期イベントをトリガーとしたスイッチの制御は、(1)と比較して稀である
- (3) スイッチの制御はスイッチに対する独立性が高く、複数のスイッチに対して依存関係を伴って一連の制御を行うことは稀である

また、検討において以下の点で定性的・定量的評価を行いました。

#### (1) 性能・スケーラビリティ

前述の想定されるコントローラ動作に対して高い性能・スケーラビリティを確保できることを最低限の判断基準としました。

#### (2) 設計・実装容易性

システムの複雑度が低く容易にソフトウェア実装が可能であり、かつ、既存ソフトウェア資産を活用できることを重視しました。

#### (3) 運用容易性

ソフトウェア実装の容易性に加え、システムの運用性も重視しました。また、既存の運用ノウハウ・システムを再利用可能か否かという点も考慮しました。

#### (4) 信頼性と耐障害性

本検討の主眼は、性能とスケーラビリティの確保ですが、コントローラの信頼性と耐障害性についても考慮しました。

## 2.2 アプローチ

検討の結果、クライアント・サーバシステムのスケールアウトのためのデザインパターンである「ロードバランサパターン」と、クライアント・サーバアーキテクチャの1つである「三階層アーキテクチャ」を応用したOpenFlowコントローラアーキテクチャを設計しました(図2)。

ロードバランサパターンでは、ネットワーク規模に応じて、ワーカ数を増減させることにより、求められるコントローラ性能を得ることができます。また、設計・実装容易性が高く、

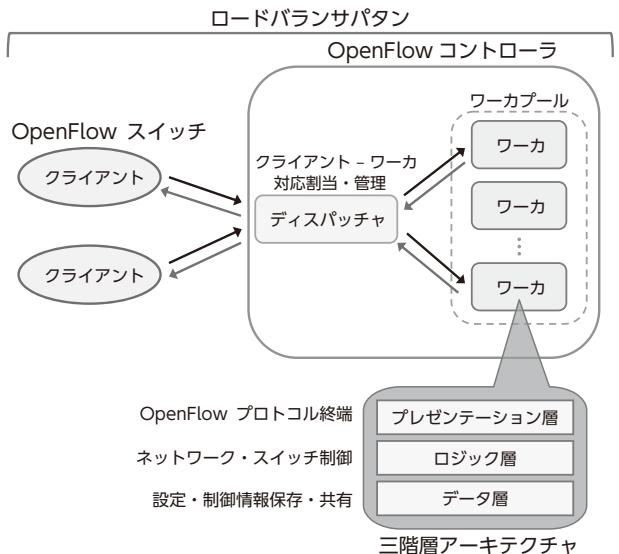


図2 提案アーキテクチャ

また、Webシステムなどで多用されていることから、その設計・運用ノウハウやソフトウェアを再利用できます。ディスパッチャは、OpenFlowスイッチからの制御チャネル確立要求を受け付け、事前に設定された基準に基づき、その接続をワーカに割り当てます。ワーカは、ディスパッチャの割り当てに従い、制御チャネルを確立します。そして、クライアントに相当するOpenFlowスイッチからの非同期イベントを処理し、必要に応じて応答を返します。

ワーカは、三階層アーキテクチャに類似したアーキテクチャにより構成し、OpenFlowプロトコルを終端するプレゼンテーション層、OpenFlowプロトコルを用いたネットワーク・スイッチ制御処理を実行するロジック層、制御情報の保存・複数ワーカ間での共有を行うデータ層から構成しました。他のワーカとの状態共有は、全てデータ層を介して行われます。データ層では、ワーカ間の状態共有に加え、ワーカの動作を決定する外部システムから与えられる設定・制御情報を保持できるようにしました。

通常のロードバランサパターン・三階層アーキテクチャと本検討のアーキテクチャの異なる点は、ワーカが能動的にスイッチの制御を行う場合があるという点です。これについては、スイッチとワーカ間の対応関係が確立された後、スイッチに対応するワーカがスイッチに対する能動的な制御を実行することとしました。ワーカはクライアントに相当するOpenFlowスイッチからの非同期イベントの発生有無にかか

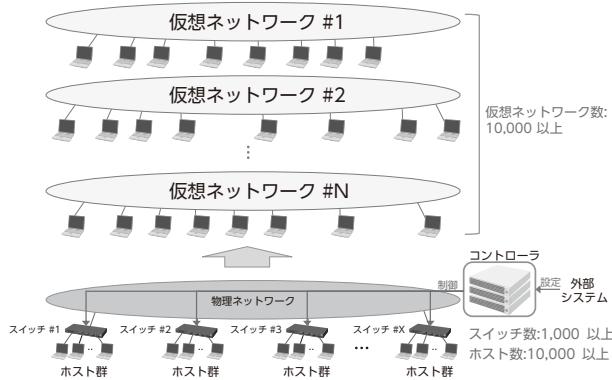


図3 コントローラ・ネットワーク要件

わらず、データ層に保存されたスイッチ設定・制御情報を確認し、必要に応じてスイッチに対する制御を実行します。

### 3. 実装

提案アーキテクチャの有効性を評価するため、提案アーキテクチャに基づくOpenFlowコントローラの実装・評価を行いました。

#### 3.1 機能・性能要件

コントローラの機能・性能、及び、管理・制御対象ネットワークの要件は以下のとおりとしました（図3）。

##### （1）機能要件

- 1) VXLAN<sup>2)</sup>により構成される仮想レイヤ2ネットワークの提供・管理を行う
- 2) OpenFlowスイッチのポートと仮想ネットワーク間の対応を管理する
- 3) OpenFlowスイッチのポートと当該ポートに接続されたホストのMACアドレスの対応を管理する
- 4) 上記の操作を行うためのインターフェースをRESTスタイルで提供する

##### （2）性能要件

- 1) 1,000台を超えるOpenFlowスイッチを管理可能であること
- 2) 10,000を超える仮想ネットワークを生成・管理可能であること
- 3) 10,000を超えるエンドホストを仮想ネットワークに接続可能であること

#### 3.2 実装

実装したコントローラの構成を図4に示します。

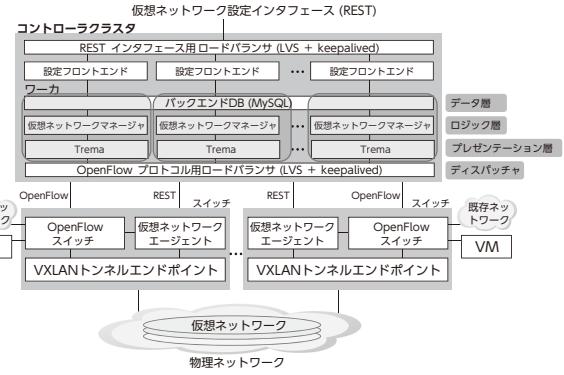


図4 コントローラ構成

ディスパッチャは、Linux Virtual Server (LVS)<sup>3)</sup>のL4ロードバランス機能により実装しました。加えてLVSの設定及びワーカの監視のため、keepalived<sup>4)</sup>を利用しました。ワーカは、OpenFlowコントローラ・スイッチ開発フレームワークであるTrema<sup>5)</sup>、新規開発した仮想ネットワークマネージャ、MySQL<sup>6)</sup>によるバックエンドデータベースにより構成しました。本実装では、仮想ネットワークの設定を行うためのインターフェースを提供する必要があるため、バックエンドデータベースに設定を保存するための機能を設定フロントエンドとして実装しました。

### 4. 評価

#### （1）スイッチ数に対するスケーラビリティ

1つのワーカに対して約410台のスイッチを割り当て可能であり、割り当てられた全スイッチに対する制御を実行可能なことを確認しました。また、ワーカ数を増加させることでシステムとして制御可能なスイッチ数がリニアに増加することを確認しました。

#### （2）仮想ネットワーク数に対するスケーラビリティ

1,024スイッチ、各スイッチに接続されるホスト数（VM数）128、という条件において16,000を超える仮想ネットワークを実際に生成・管理可能であることを確認しました。

#### （3）仮想ネットワーク設定性能・スケーラビリティ

仮想ネットワーク数に依存せず、仮想ネットワーク生成・設定にかかる時間が一定であることを確認しました（図5）。

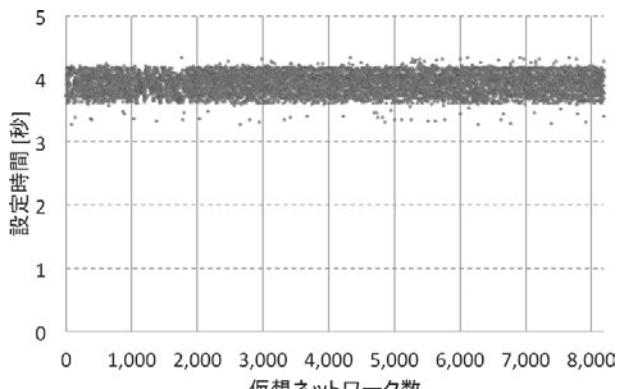


図5 仮想ネットワーク設定性能

### 執筆者プロフィール

千葉 靖伸

情報・ナレッジ研究所  
主任

須堯 一志

情報・ナレッジ研究所  
主任研究員

### 関連URL

NEC BIGLOBE Virtualizes the BIGLOBE Data Center  
Using State-of-the-Art OpenFlow Technology  
<http://www.biglobe.co.jp/en/press/2013/0422.html>

## 5.まとめ

大規模 SDN ネットワークを実現する OpenFlow コントローラーアーキテクチャ、及び、その実装例を示しました。本件の成果は、NEC ビッグローブにおいて商用サービスの提供に活用されています。今後は、異なる要件・機能を持つコントローラの性能向上・スケーラビリティ確保について検討を進め、幅広い領域への SDN 適用を目指します。

\* OpenFlow は、Open Networking Foundation の商標または登録商標です。

\* Linux は、Linus Torvalds の米国及びその他の国における登録商標です。

\* MySQL は、Oracle Corporation またはその関連会社の商標です。

### 参考文献

- 1) OpenFlow - Open Networking Foundation  
<https://www.opennetworking.org/sdn-resources/onf-specifications/openflow>
- 2) VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks  
<http://tools.ietf.org/id/draft-mahalingam-dutt-dcops-vxlan-06.txt>
- 3) The Linux Virtual Server Project  
<http://www.linuxvirtualserver.org/>
- 4) Keepalived for Linux  
<http://www.keepalived.org/>
- 5) Trema Full-Stack OpenFlow Framework in Ruby and C  
<http://trema.github.io/trema/>
- 6) MySQL :: The world's most popular open source database  
<http://www.mysql.com/>

# NEC 技報のご案内

NEC 技報の論文をご覧いただきありがとうございます。  
ご興味がありましたら、関連する他の論文もご一読ください。

NEC 技報 WEB サイトはこちら

NEC 技報(日本語)

NEC Technical Journal(英語)

## Vol.66 No.2 ICTシステムを高度化するSDN特集

ICTシステムを高度化するSDN特集によせて  
SDNがもたらすICTシステムの高度化とIT・ネットワーク市場の変化  
NECのSDNへの取り組みとNEC SDN Solutions  
SDN実用化に向けた標準化

### ◇ 特集論文

#### NEC Enterprise SDN Solutions

WANの利用、運用を効率化する拠点・データセンター接続最適化ソリューション  
安全で柔軟なネットワークアクセスを提供する「アクセス認証ソリューション」

#### NEC Data Center SDN Solutions

仮想環境の効率化を実現するIaaS運用自動化ソリューション

#### NEC SDN Solutionsを支える最新技術

SDNコントローラ作成のシンプル化を実現するネットワーク抽象化モデル  
Wi-Fiの利便性向上を実現するスマートデバイス通信制御技術  
大規模SDNネットワークを実現するOpenFlowコントローラーアーキテクチャ  
ヘテロジニアス網統合制御基盤を実現するマルチレイヤ抽象化技術  
運用省力化を実現するIP-VPN向けOpenFlowコントローラ



### 導入事例

乱立する部門LAN、移動する検査機器 医療現場のネットワークをOpenFlowで改革  
事業拡大を見据えデータセンターにSDNを導入 サービスのスピード、信頼性、他社優位性を向上

### ◇ 普通論文

iPASOLINK All Outdoor Radio (AOR) 装置の開発  
iPASOLINKシリーズ及び超多値変調技術の開発  
10Gbps伝送を実現する超大容量無線伝送技術  
メタマテリアルを用いた電磁ノイズ抑制技術とその実用化

### ◇ NEC Information

C&Cユーザーフォーラム & iEXPO2013  
人と地球上にやさしい情報社会へ～インフラで、未来をささえる～  
NEC講演  
展示会報告

### NEWS

2013年度C&C賞表彰式典開催

## Vol.66 No.2 (2014年2月)

特集TOP