

大規模データセンターの管理自動化をサポートするSDN対応製品 UNIVERGE PFシリーズ

早野 慎一郎 宮本 武 山下 敦也

要旨

SDN対応製品であるUNIVERGE PFシリーズについて、大規模ITシステム基盤向けの強化内容や採用技術・適用メリットを紹介します。UNIVERGE PFシリーズは、大規模化対応や運用性の向上を行う機能強化を行うとともに、OpenFlowに加えオーバーレイ方式の対応と、ネットワークの統合運用管理機能を提供しており、NEC Cloud IaaSにも採用されました。本稿では、製品の構成や機能、連携方法の説明、仮想ネットワークの分割制御・統合運用について紹介します。



SDN／OpenFlow／仮想ネットワーク／仮想化／オーバーレイ／クラウド／システム基盤

1. はじめに

仮想化技術の進展によって、業務システムを支えるITシステム基盤の集約や、その運用効率化が一般的になってきました。またクラウドサービスの発展や機能の拡充に伴い、クラウドサービス事業者が提供するITシステム基盤サービスを利用するケースも格段に増えてきました。これらのITシステム移行・集約トレンドに伴い、企業向けやクラウドサービス向けに構築される仮想化されたITシステム基盤には、大規模化や運用自動化などのニーズが急速に高まっています。コンピュータ、ストレージと並び、ITシステム基盤の重要な要素であるネットワークにおいても、仮想化や大規模・運用自動化への対応がいまや必須条件となっていました。

このような市場ニーズのなか、SDN (Software-Defined Networking) は、ITシステム基盤の大規模化や運用自動化に加え、そのITシステム全体の柔軟性や経営要求への迅速な対応を高めるアーキテクチャとして、大きな期待が寄せられています¹⁾。

2011年に世界に先駆けて製品化されたNECのSDN対応製品であるUNIVERGE PFシリーズは、大規模化や運用性を強化するとともに、OpenFlow²⁾だけでなく複数の既存ネットワーク技術に対応し、統合運用管理機能を新たに備

いました。本稿では、クラウド基盤サービス「NEC Cloud IaaS」にも採用されている、UNIVERGE PFシリーズの大規模仮想ITシステム基盤向けの強化ポイントと、それを支える技術や適用メリットについて紹介します。

2. UNIVERGE PFシリーズの概要

クラウドサービスで利用される仮想ITシステム基盤では、スイッチ台数が1,000台を超えるような大規模なネットワークとなるケースがあり、従来のネットワーク技術では拡張性の制限や構成変更時の設定の複雑さが課題でした。

こうした課題を解決するため、UNIVERGE PFシリーズは、階層的なアーキテクチャを採用しています。またOpenFlowだけでなく、オーバーレイ方式のネットワークも採用可能とすることにより、拡張性が高く柔軟なネットワークを実現します。

図1に示すようにUNIVERGE PF6800（以下、PF6800）は、OpenFlowに対応したコントローラであり、ハードウェアのUNIVERGE PFシリーズのスイッチ製品群（以下、PFS）を制御します。またUNIVERGE PF6700（以下、PF6700）は、VXLAN³⁾技術を採用したオーバーレイ方式によるネットワークの設定・管理を行います。オーバーレイ方式は、異

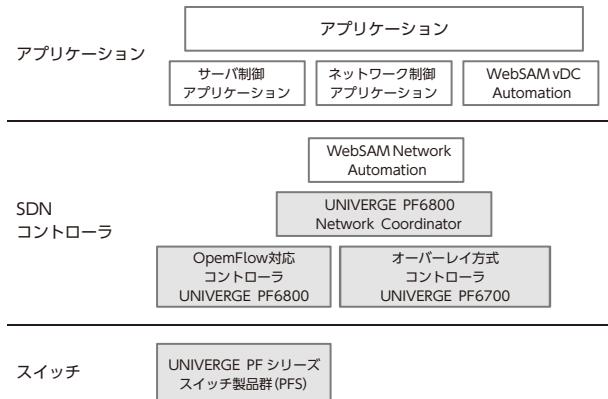


図1 UNIVERGE PFシリーズの位置付け

なるサーバの仮想スイッチ間をVXLAN技術などによるトンネルで接続することにより、仮想ネットワークの柔軟な構築を可能にします。

UNIVERGE PF6800 Network Coordinator (UNC) は、PF6800及びPF6700を統合的に管理・制御します。システム管理製品であるWebSAM vDC Automationは、WebSAM Network Automationと連携してITシステム基盤のユーザーからのリクエストに応じて、ネットワークリソースをUNCに要求し、UNCがPF6800及びPF6700を制御してリソースを提供します。こうしたネットワーク設計・設定は従来、人手でされてきましたが、ユーザーからのリソース要求を迅速にネットワーク設定へ反映させるためには、自動化する必要がありました。

3. コントローラの階層化

3.1 分割制御の実現

OpenFlowは集中制御方式であるため、図2に示すように、スイッチの台数が多くなるとコントローラの性能不足や制御ネットワークの帯域不足が課題となります。これを解決するために、図3のようにUNIVERGE PFシリーズでは、複数のPF6800が分割されたドメインごとにネットワークを制御し、UNCが統合的に仮想ネットワークの設定・管理を行なう分割処理方式を実現しています。1式のUNCで最大40クラスタのPF6800を管理でき、これにより管理対象のスイッチが1,000台を超えるような大規模ネットワークの管理が可能となりました。

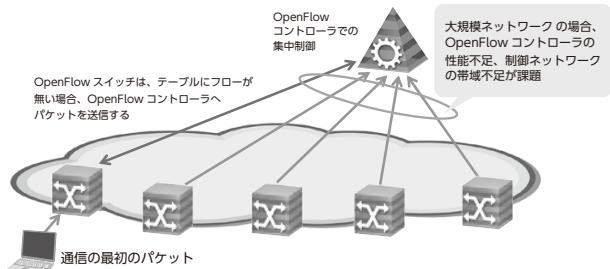


図2 OpenFlowの大規模ネットワークでの課題

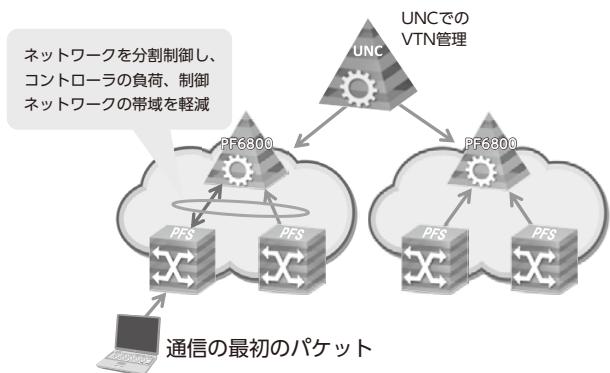


図3 UNIVERGE PFシリーズによる大規模ネットワークの実現

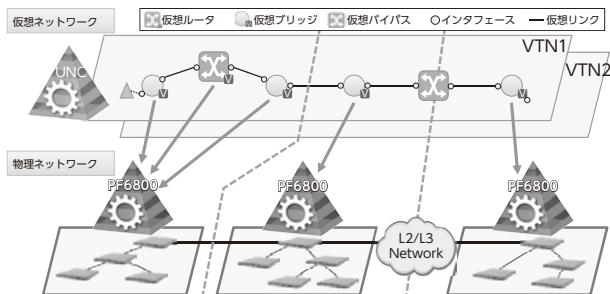


図4 UNCによる統合的な仮想ネットワークの管理

3.2 仮想ネットワークの統合管理

UNIVERGE PFシリーズでは、SDNで実現する仮想ネットワークをVTN (Virtual Tenant Network)⁴⁾と呼びます。VTNではそれぞれのVTNが論理的に分離されており、利用するIPアドレス空間をVTNごとに管理しています。このため、IPアドレスの重複を気にせずに複数のVTNを1つの物理ネットワークに構築することができます。

ユーザーはUNCでVTNを作成することができ、図4のよ

うに複数のPF6800にまたがったVTNを作成することも可能です。更に物理ネットワークのトポロジーが可視化され、VTNとあわせて物理/仮想の両面から直観的にネットワークの状態が把握可能となるため、大規模ネットワークを効率的に運用管理できます。このように、コントローラの階層化は、統合運用管理可能なネットワークの規模を大幅に拡大し、大規模ITシステム基盤の運用管理効率化に大きく貢献します。

なお、UNCはVTNに関してOSS(Open Source Software)のSDNコントローラであるOpenDaylight⁵⁾ソフトウェアに準じたAPI(Application Programming Interface)を持ちます。ユーザーがSDNコントローラやSDNアプリケーションを選定する際、OSSに準じたAPIを持つことは重要な要素となります。

4. オーバーレイ方式への対応

4.1 オーバーレイ方式の実現

オーバーレイ方式とは、異なるサーバの仮想スイッチ間をVXLAN技術などで接続して、仮想ネットワークを構成するものです。これによりサーバ間の物理ネットワークを変更することなく、かつ物理ネットワークの制限を超えて仮想ネットワークを柔軟に拡張することが可能となり、大規模なITシステム基盤を効率的に運用できます。例えば、従来課題であったVLAN IDの制限を超えて4,096個以上の仮想ネットワークを実現でき、異なる物理サーバ間で仮想サーバのマイグレーションが可能になります。

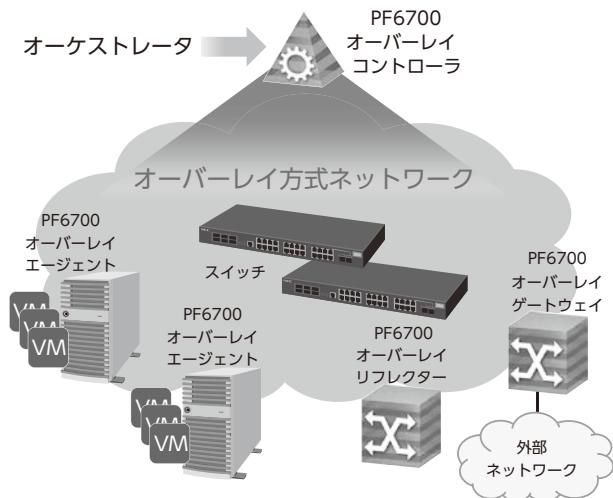
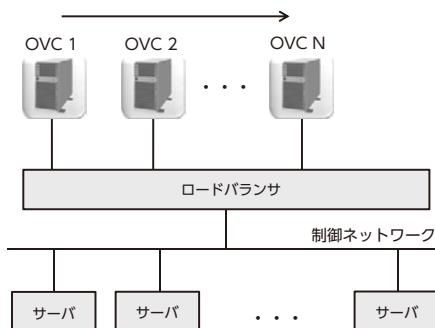


表 PF6700を構成するコンポーネントと役割

ソフトウェア名	役割
PF6700 オーバーレイ コントローラ	すべてのオーバーレイゲートウェイ、リフレクター、エージェントを制御して、オーバーレイ方式ネットワークを管理
PF6700 オーバーレイ ゲートウェイ	オーバーレイ方式ネットワークと外部ネットワークを接続し、相互にまたがる通信を中継
PF6700 オーバーレイ リフレクター	宛先不明なデータ、ブロードキャストやマルチキャストのデータを仮想サーバ(VM)やPF6700 オーバーレイゲートウェイに転送
PF6700 オーバーレイ エージェント	PF6700 オーバーレイコントローラから受けた要求メッセージをサーバ内の仮想スイッチに設定



PF6700は、オーバーレイ方式を採用した新製品であり、図5及び表に示すような複数のコンポーネントにより構成されています。

4.2 大規模システムに対応したスケーラビリティ

従来、大規模なシステムでオーバーレイ方式を利用すると、宛先不明なデータやブロードキャスト/マルチキャストデータが、ネットワーク全体の負荷となる課題がありました。UNIVERGE PF6700オーバーレイリフレクター（以下、OVR）が、宛先不明なデータやブロードキャスト/マルチキャストデータなどをネットワーク全体に転送することなく、サーバやPF6700オーバーレイゲートウェイ（以下、OVG）に適切に転送することで、大規模システムでの性能劣化を防ぎます。

またPF6700全体を制御・管理するPF6700オーバーレイ

コントローラ（以下、OVC）、OVR、OVGは、大規模データセンターを考慮し、柔軟にスケールアウトできるシステム構成としています。

・ OVCのスケールアウト

ロードバランサ配下にOVCを配置することにより、サーバの増加に応じて、OVCをスケールアウトさせることができます（図6）。

・ OVRのスケールアウト

OVCと同じように、ロードバランサ配下にOVRを配置することにより、スケールアウトを実現できます。また、OVRをグループ化し、そのグループ単位でロードバランサを配置することができるため、ロードバランサの負荷分散も実現可能となっています（図7）。

・ OVGのスケールアウト

複数のアクティブ/スタンバイのOVGペアをグループとして管理し、外部ネットワークとやりとりする仮想ネットワーク単位にOVGグループを割り当てることにより、負荷分散を実現しています。また、グループ内においてOVGのペアに重みづけを行い、その重みづけに応じた負荷分散を行います（図8）。

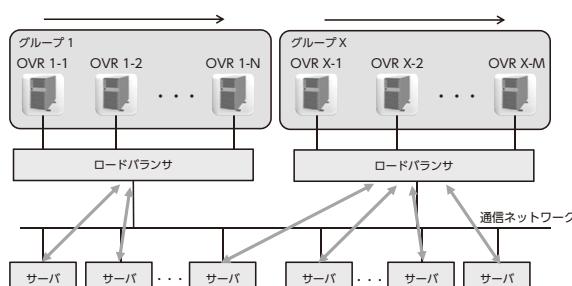


図7 OVRのスケーラビリティ

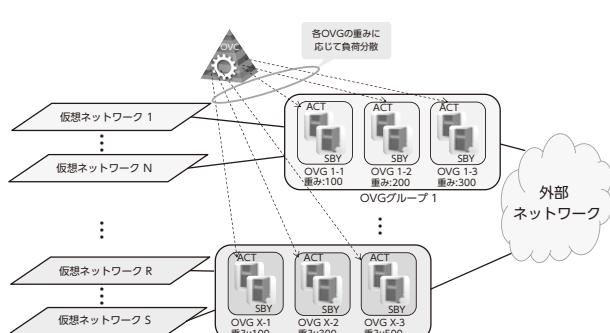


図8 OVGのスケーラビリティ

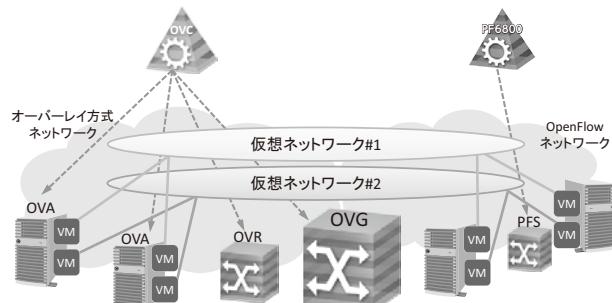


図9 PF6700の適用事例

4.3 適用事例

NEC Cloud IaaSでは、「ハイアベイラビリティ（HA）」と「スタンダード（STD）」の2つのサービスがあります。

HAサービスでは、VTNによりネットワーク仮想化を実現するPF6800を利用しています。

また、STDサービスでは、VXLANオーバーレイ方式によりネットワーク拡張性を実現するPF6700を利用しています。

仮想ネットワークがHAサービスとSTDサービス両方を跨ぐ場合は、図9のようにOVGにてVXLAN変換を行います。

5. スイッチ製品の大容量化

第3章1節に説明したように、OpenFlow技術を使う大規模ネットワークでは、コントローラの性能が課題となります。この対策として、先に述べたコントローラの階層化に加え、OpenFlowコントローラの負荷を軽減させるスイッチUNIVERGE PF5459（以下、PF5459）（写真）を開発しました。本章では、この製品に採用した負荷軽減手法（以下、MAC Forwarding手法）を紹介します。

5.1 MAC Forwarding手法

弊社の従来のOpenFlowのフローテーブルは、送信元/宛先MACアドレス・VLAN IDなどの組み合わせを利用してきました。MAC Forwarding手法は、これに従来技術のMAC（Media Access Control）学習テーブルを前段に加えた多段テーブルをOpenFlowの仕様で実現したものです。

このMAC Forwarding手法は、図10に示すように2つのテーブルを活用します。Dynamic Mac Flow tableとは、従来のレイヤ2スイッチ（L2スイッチ）が持つMAC学習テーブルと同等の動作をします。またStandard Flow tableは、弊



写真 UNIVERGE PF5459の外観

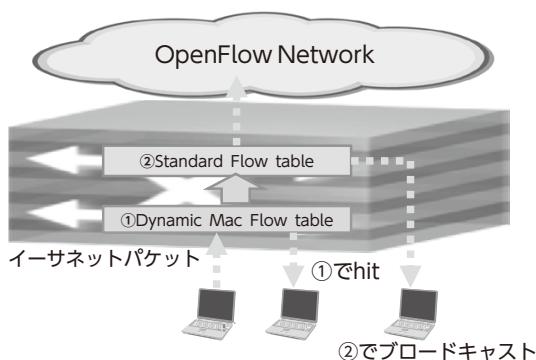


図10 テーブル構造とスイッチ動作

社の従来のOpenFlowフローテーブルと同等な動作をします。

MAC Forwarding手法は、入力されたパケットに対して、まずDynamic Mac Flow tableを検索します。宛先MACアドレスが学習済みであった場合は、L2スイッチと同等なパケット転送を行い、未学習の宛先MACアドレスであった場合は、Standard Flow tableを検索し、パケットを転送します。

このようにパケットの宛先を学習する機能があるため、従来のOpenFlowの動作において未知のパケットが到達するたびに発生していた、コントローラへの問い合わせが発生しません。つまり、コントローラの負荷を大幅に軽減することができます。

また、スイッチにおいても、高速・大容量化が難しいStandard Flow tableをDynamic Mac Flow tableで補完することで、従来のL2スイッチと同様に大量の端末を収容することが可能となります。これにより、OpenFlow技術に対応した多ポート(48ポート)10Gbpsスイッチを開発することができました。

5.2 VLAN設定の一元管理

PF5459では、SDNにより複数スイッチのVLAN設定を一元管理できるようにしました。従来のスイッチでは、構成

を変更するごとに、各スイッチに最大4,095個のVLAN設定を行う必要がありました。その運用工数を大幅に削減できます。

6. むすび

以上述べたように、SDN対応製品の階層化や階層ごとの大規模ネットワーク対応により、拡張性が高く柔軟なネットワークを実現し、かつ大規模システムの運用自動化、管理効率化を行うことができるようになりました。今後、複数拠点の統合運用などの機能強化を進めていく予定です。

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

* OpenDaylightは、OpenDaylight Project, Inc.の商標または登録商標です。

* Ethernetは、富士ゼロックス株式会社の登録商標です。

参考文献

- 1) NEC SDN Solutions
<http://jpn.nec.com/sdn/index.html>
- 2) OpenFlow -Open Networking Foundation
<https://www.opennetworking.org/ja/>
- 3) VXLAN
<https://tools.ietf.org/html/rfc7348>
- 4) VTN
[https://wiki.opendaylight.org/view/OpenDaylight_Virtual_Tenant_Network_\(VTN\):Overview](https://wiki.opendaylight.org/view/OpenDaylight_Virtual_Tenant_Network_(VTN):Overview)
- 5) OpenDaylight
<http://www.opendaylight.org/>

執筆者プロフィール

早野 慎一郎

ソリューションプラットフォーム統括本部
シニアエキスパート

宮本 武

コンバージドネットワーク事業部
シニアエキスパート

山下 敦也

コンバージドネットワーク事業部
シニアエキスパート

関連URL

UNIVERGE PFシリーズ製品情報

<http://jpn.nec.com/univerge/pflow/>

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.67 No.2 ICTシステムを担うこれからのクラウド基盤特集

ICTシステムを担うこれからのクラウド基盤特集によせて
NECのクラウド基盤への取り組み

◇ 特集論文

NEC C&Cクラウド基盤 NEC Cloud IaaSのサービス

マルチ環境統合を実現するポータルサービス
多用途環境に対応するハイブリッド型サーバーサービス
多様なネットワーク環境を提供するネットワークサービス
内部統制手法を活用した堅牢なセキュリティサービス
クラウド基盤を支えるデータセンターサービス

NEC C&Cクラウド基盤を支える製品、最新技術

運用の自動化によりトータルコストを最適化する「WebSAM vDC Automation」
運用自動化により効率的な管理を実現する統合運用管理基盤
データセンターのTCO削減に貢献するマイクロモジュラーサーバー及び相変化冷却機構
クラウド環境に適した高信頼基盤を提供するiStorage M5000
データ保存に最適な、優れた圧縮効率と高速性を両立するiStorage HSシリーズ
大規模データセンターの管理自動化をサポートするSDN対応製品 UNIVERGE PFシリーズ
省電力を実現する相変化冷却技術・熱輸送技術

NEC C&Cクラウド基盤の将来技術

低コスト・省電力・低フットプリントを実現するアクセラレータ活用技術
スケールアップにより多種多様なコンピューティングを実現するResource Disaggregated Platform
クラウド環境を対象にしたモデルベース設計支援技術
モデルベースでのサイジングと構成管理によりクラウド上のSIを効率化するクラウド型SI
ビッグデータ分析とクラウド～異常を見抜くインパリヤント分析技術～

導入事例

クラウドで遠隔監視保守システムの安定稼働を実現 全国約1,100基のタワーパーキングの安全を支える
ビジネスの中核を担うシステムをNEC Cloud IaaSへ移行 NECのトータルサポート力を評価
クラウド基盤サービスでグループのIT環境を共通化 ITガバナンスのさらなる強化を目指す

◇ NEC Information

C&Cユーザーフォーラム &iExpo2014

Orchestrating a brighter world 世界の想いを、未来へつなげる。

基調講演

展示会報告

NEWS

2014年度C&C賞表彰式典開催



Vol.67 No.2
(2015年3月)

特集TOP