次世代のオープンネットワーク基盤としての「SDN」

仮想ネットワークをソフトウェアで自由に構成・制御する技術「SDN」の実現に必要な技術要件と今後の可能性、ならびに株式会社ストラトスフィアのストラトスフィアSDNプラットフォームの特長について解説します。

3.1 はじめに

SDN(Software Defined Network)[*1]とは、ユーザがプログラムによりネットワークを自由に制御しようというコンセプトです。IIJグループでは、本年4月に株式会社アクセスと共同出資の合弁会社、株式会社ストラトスフィアを設立し、SDNを実現するための基盤ソフトウェアの開発に取り組んでいます。本稿では、SDNに関する技術要件について簡単に説明し、これらの技術要件が実現されることによる次世代のクラウドシステムの発展の可能性について解説します。

3.2 クラウド環境における仮想化への要求とネットワークの課題

3.2.1 クラウドへの要求と既存ネットワーク技術の限界

クラウドサービスを利用するユーザと、サービス提供を行うオペレータの、両方の視点から考えた場合、クラウドシステムには下記の要件が求められていると考えられます。

- システム構築の柔軟性・拡張性・迅速性の確保
- システム全体の稼働率向上
- 標準化された機器構成と運用体制の構築
- システムの信頼性の確保

一方、現状のクラウドシステムで用いられるVLAN(Virtual LAN)というネットワーク仮想化技術には、下記の制約があります。

- VLAN IDが4,094通りしか取れない
- ルータ経由で接続されているサーバクラスタ間をまたがるL2セグメントが作れない

現在は、サーバクラスタ内のネットワークはL2スイッチで構成され、ネットワークの他の部分とはルータを経由してL3で接続されるのが一般的です。クラスタ内部ではVLANを用いてユーザごとに仮想L2セグメントが構成されています。しかし、こういった環境では、上記の制約により、1つのクラスタ内で構成可能なL2セグメントの数は、VLAN IDの上限である、4,094に制限されます。VLAN数が上限に達してしまったクラスタでは、新規のユーザを収容できなくなりますので、クラスタ内のサーバの稼働率をなかなか上げられません。また、あるユーザのL2セグメントは、クラスタを越えて延ばすことができませんので、同一L2セグメント内のVM(Virtual Machine)の数を増やしたいといった要求に応えられないケースも発生します。

このような制約を持つシステムでは、サービス品質の定義に合わせて、クラスタを構成するサーバの台数、サーバごとに収容するVMの数やスペックなど異なるポリシーを設定し、構成やポリシーの異なる複数のクラスタを運用することになります。すると、運用体制の複雑化やコスト増、そして、複数の異なる運用フローによる全体の信頼性低下を招くことにもなりかねません。

3.2.2 ネットワーク仮想化からSDNへ

これらの既存ネットワーク技術の制約を超え、クラウドサービスの普及により発生した新たな要件に応え、クラウドシステムを進化させるためには、従来のVLANに替わるネットワーク仮想化の技術が必要となります。

3.2.1での考察から、こうした新たなネットワーク仮想化技術は、L2セグメントを物理的なネットワークの境界を越えて展開できるもの、かつ、VLANの制約を越えて多数のユーザネットワークを収容できるものでなければならないことが分かります。

*1 Open Networking Foundation, “Software-Defined Networking: The New Norm for Networks” ONF White Paper, Apr. 2012

また、これからのクラウドシステムでは、多数のサーバや、その上で稼働する膨大な数のVM、そしてそれらを結ぶ複雑なネットワークの構築や管理を行いながら、持続的で安定した運用の実現も必要とされています。そのためにはシステムやネットワーク管理のかなりの部分を自動化し、運用をスケーラブルなものにする必要があります。このような要件を満たすために考えられた概念がSDNです。SDNでは、従来のVLANの限界やL2境界を超えて柔軟に仮想ネットワークを構築することと、仮想ネットワークも含めたネットワーク全体の構成管理や運用を自動化することが、ソフトウェアにより実現できます。そして、それらの機能を他のシステムから利用可能にするためのAPIを定義し、ユーザに提供することで、ユーザが柔軟にネットワークインフラを利用しながら、クラウドインフラ上に新しいサービスを構築することが可能になります(図-1)。

SDNの実現のために必要な技術項目を、下記のようにまとめてみました。

1. 1つの物理ネットワーク上で、多数の論理的なネットワークを互いに独立して動作させる仮想化技術
2. 仮想化技術を用いて構成される論理ネットワークの物理ネットワークへの収容管理・構成管理を行う技術、及び物理ネットワークのリソース共用の管理技術を実現するためのネットワークモデリングとAPI整備
3. 物理ネットワーク上に多重展開された論理ネットワーク全体を動作させるために、ネットワーク機器の設定や動作を遠隔から集中制御する技術

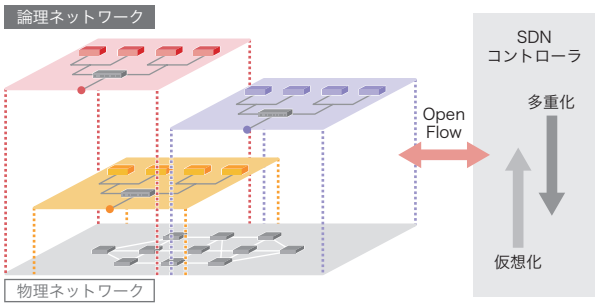

図-1 SDNの概念図

1.や3.に関しては、スタンフォード大学のClean Slate Program産学共同研究を源流とするOpenFlow[*2]というプロトコルがあります。OpenFlowでは、FlowRuleと呼ばれる、フロー単位でパケット転送動作を細かく規定するルールを遠隔からネットワーク機器に設定することで、従来とは異なるパケット転送動作を実現し、ネットワーク仮想化を実現します。

一方、VXLAN[*3]、NVGRE[*4]、STT[*5]のような仮想L2ネットワークをIPネットワーク経由で展開するための新たなプロトコルも提案されています。これらのプロトコルでは、サーバ間にIPトンネルを設定し、それらのサーバ上にあるVM同士を同一の仮想L2ネットワークに接続することができます。VLANの制限は、VNI(VXLAN Network Identifier)やVSID(Virtual Subnet ID)などのより大きなIDフィールドを新たに導入することで拡張し、論理ネットワークの区別を行います。

ネットワーク仮想化の実現には、2つの方法があります(図-2)。1つは、ネットワーク機器ごとにFlowRuleを設定し、フロー単位で仮想ネットワークを実現する方式で、Hop-by-Hop型と呼びます。また、VXLANやNVGREのようなIPトンネリングを用いて仮想ネットワークを実現する方式をEdge Overlay型と呼びます。

Hop-by-Hop型は、各ネットワークリンクに流れるトラフィックをフロー単位できめ細かに制御できる反面、すべてのネットワークノードにFlowRuleを設定し、全体として一貫性のあるパケット転送を行う必要があります。

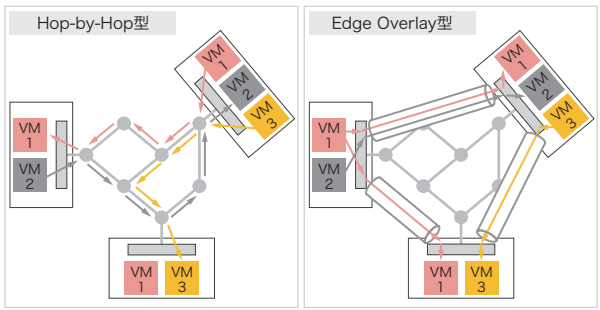

図-2 Hop-by-Hop型とEdge Overlay型

*2 Nick MacKeown, et al. "OpenFlow: Enabling Innovation in Campus Networks", Mar. 2008
*3 M. Mahalingam, et al. "VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", Internet-Drafts, Aug. 2012
*4 M. Sridharan, et al. "NVGRE: Network Virtualization using Generic Routing Encapsulation", Internet-Drafts, Jul. 2012
*5 B. Davie, et al. "A Stateless Transport Tunneling Protocol for NetworkVirtualization(STT)", Internet-Drafts, Sep. 2012

一方、Edge Overlay型は、各サーバ上のVMが割当てられているユーザと、VNIやVSIDとの対応を管理しさえすれば、IPトンネルにより柔軟に仮想ネットワークを展開することができる反面、サーバ間のIPネットワーク上でのきめ細かなパケット転送制御を行うことはできません。一般的にはEdge Overlay型のほうが、既存のIPネットワークに手を入れずに利用できるので、ネットワーク機器をすべてOpenFlow対応の機器に入れ替えOpenFlowによる制御を行わなければならないHop-by-Hop型よりも導入が簡単だと考えられています。Edge Overlay型であっても、IPトンネルを終端するサーバにソフトウェアスイッチを設置して、ソフトウェアスイッチの動作をOpenFlowで制御することも可能です。Hop-by-Hop型とEdge Overlay型は、それぞれの特徴を組み合わせながら、ハイブリッドに活用することを考えるのが良いでしょう。

3.3 ストラトスフィアSDNプラットフォーム(SSP)

ここで株式会社ストラトスフィアの製品の特長について解説します。

仮想ネットワークの実装方式としては、Edge Overlay型を採用しています。最初は導入の簡単なEdge Overlay型を採用し、将来的にはHop-by-Hop型に対応する機能拡張や、MPLS(Multi-Protocol Label Switching)やPBB(Provider Backbone Bridge)といったバックボーンプロトコルとの相互接続、そして、SDNを使ったWANへの拡張に取り組み、広域なネットワーク全体の仮想化を実現します。

オーバレイのプロトコルは、VXLANとSTTに対応します。VXLANのVNIは24ビットの識別子ですので、約1,600万通りの仮想L2セグメントをサポートできます。

SSPでは、従来利用されてきた様々なネットワーク技術に新しい技術を融合させることで、既存の環境からシームレスに移行可能なSDNソリューションを実現しています。図-3に示すのは、SSPでの、VXLANを用いたEdge Overlay型の構成イメージです。Compute Nodeというのは仮想化されたサーバのことで、このサーバのハイパーバイザー内に仮想スイッチを設置し、この仮想スイッチ間でVXLANのトンネルを設定します。このとき、仮想スイッチ上ではVLANを用いてL2セグメントの識別を行います。仮想スイッチからVXLANのトンネルに載せるときに、VLAN IDをVNIに変換します。このような構成を取ることで、Compute Nodeで保持しなければならないFlowRuleの数を格段に減らすことができ、OpenFlowによるパケット処理のオーバヘッドを最小限に抑えることが可能になります。SSPのやり方では、1台のCompute Node内で4,094通り以上のVLANは設定できないのですが、実運用上は4,094のVLANを使い尽くすことは暫く起こらないと考えます。それよりも、現状VLANで運用しているシステムが大部分であることを考えると、このような構成にした方が導入時のストレスが少ないと考えられます。

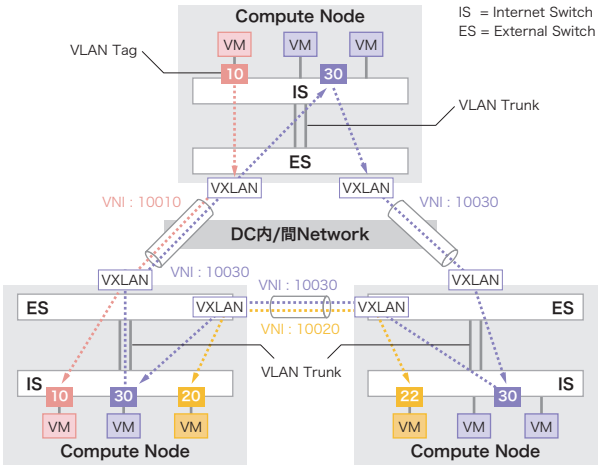

図-3 SSPのトンネル構成概念図(ビルトインモード)

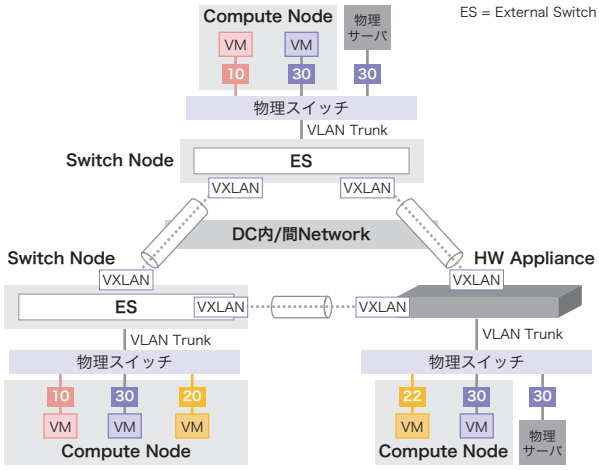

図-4 セパレートモード

また、SSPは構成要素がモジュラー構成(図-4)になっていますので、Compute NodeのHyperVisor上で動いていたExternal Switchを(ビルトインモード)、必要に応じて別のサーバで動かして(セパレートモード)、Compute Nodeでの仮想スイッチの処理負荷を軽減させたりすることが可能ですし、場合によってはハードウェアアプライアンスを利用したハードウェアによるオーバレイの処理を行うことも可能です。

3.4 ネットワークモデリング化とAPI化

3.2.2で述べた技術項目2.を実現するためには、SDNで扱う論理ネットワークや物理ネットワークの抽象モデルが必要となります。このようなモデルから、クラウドユーザの用いる論理ネットワークやそれらを多重収容する物理ネットワークを定義することで、SDNシステム上で、論理ネットワークの物理ネットワーク上での最適配置計算やリソース共有の制御などを取り扱うことが可能になります[*6*7]。更に、これらのネットワークモデルを操作するためのメソッドをAPIとして定義することで、仮想ネットワークを設定・管理・制御する機能をサービスとして外部のアプリケーションに提供することができます。

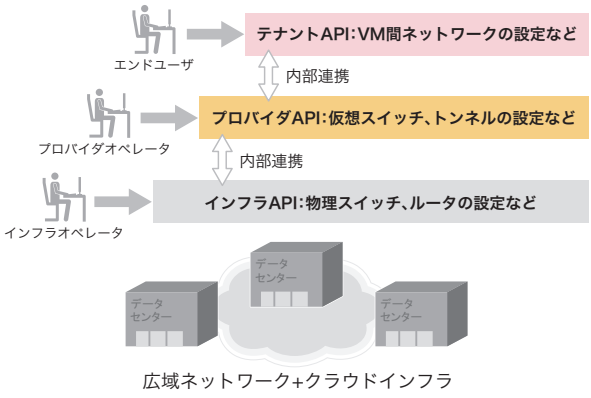

図-5 クラウドインフラのユーザへの開放

SSPでは、APIを3階層に分けて定義しています(図-5)。最上層は、エンドユーザのVM同士を繋ぐ仮想L2セグメントなどの操作を行うレイヤーであり、最下層はインフラオペレータのための、物理ホストやネットワーク機器を操作するためのレイヤーです。インフラの一部を分割し、ユーザレベルの仮想ネットワークを多重収容するための中間層は、インフラオペレータからリソースをまとめて借りて、その中からエンドユーザにリセールするようなサービスプロバイダが必要とする各種操作を行うレイヤーです。この中間層のプロバイダ向けAPIにより、インフラの一部がユーザに解放され、プロバイダユーザはこのAPIを用いてそれぞれの目的に応じた仮想ネットワークサービスを構築し、エンドユーザに提供することが可能となります。

3.5 結論

SDNの基盤の整備とAPIの整備により、サービスプロバイダやエンドユーザは、論理的なシステムが物理的にどこにどう配置されているかなどの詳細を気にすることなく、APIを通じてクラウドインフラ上の抽象化されたリソースを自由に利用することが可能になる一方、インフラオペレータはサービスプロバイダやエンドユーザが展開している仮想ネットワークが、自社インフラのどこにどう展開されて動作しているかを、インフラAPIを通じて把握し、必要に応じて再配置し、最適化できるようになります。このような環境を実現するクラウド基盤を整えることで、クラウドインフラの利用効率を最大化しながら、サービスプロバイダやエンドユーザの自由度をも最大化する、次世代のオープンネットワークシステムが実現するのです。

執筆者:

浅羽 登志也(あさば としや)
株式会社IIJイノベーションインスティテュート 代表取締役社長。株式会社ストラトスフィア 代表取締役社長。1992年、IIJの設立と共に入社し、バックボーンの構築、経路制御、国内外ISPとの相互接続などに従事。1999年より取締役、2004年より取締役副社長として技術開発部門を統括。2008年6月に株式会社IIJイノベーションインスティテュートを設立、同代表取締役社長に就任。2012年4月に株式会社ストラトスフィアを設立、同代表取締役社長に就任。

*6 T. Koponen, et al., "Onix: A Distributed Cotrol Platform for Large-scale Production Networks" Operating Systems Design and Implementation (OSDI), Oct. 2010
*7 浅羽, 「SDNの可能性と次世代クラウドの展望」電子情報通信学会インターネットアーキテクチャ研究会, 2012年9月