

第1部

特集1 データセンターネットワークイノベーション

関谷 勇司、宇多 仁

第1章 背景

近年、データセンターにおけるネットワーク技術が注目を浴びており、従来のレイヤリングモデルとは異なった、様々な技術が登場している。その理由として、クラウドやビッグデータ処理の普及があげられる。クラウドやビッグデータを処理するにあたっては、多くの計算資源やストレージ資源を必要とするため、多くの場合データセンターに機材が設置され、運用される。そして、それらの機材を連結し、ひとつのサービスとして形成するためにネットワークが利用される。そのため、データセンター内部のネットワークは、多くの機材を接続し、多くのトラフィックを低遅延にて処理することが求められるネットワークとなる。さらに、単一の障害にサービスが影響されないという冗長性や、コスト面から、できる限り接続リンクを効率的に利用することが求められるという、従来のインターネットの思想を受け継ぎつつも、新たな要求を含んだネットワークとなっている。そこで、WIDE Projectでは、このデータセンターネットワークの現状を正確に把握し、求められる要件とそれに応えるあるべき技術の姿を探るため、WIDE Project夏合宿ならびにWIDE Project秋合宿において、特別セッションを開催し、集中的な議論を行った。本報告書はその議論の中身と、議論から導き出されたこれからのデータセンターネットワークのあるべき姿について報告する。

第2章 WIDE Project夏期合宿における議論

WIDE Projectでは、2013年8月4日から2013年8月6日の期間、函館大沼プリンスホテルにおいて、WIDE Project夏期合宿を開催した。この合宿は、毎年集中的に議論を行うテーマをしばり、有志メンバーを募って開催される合宿である。今年は、WIDE Project Boardメンバー 9人、WIDE Projectメンバー 4人、ゲストスピーカー 6人の計20人が参加して開催された。

本年の夏期合宿のテーマは「データセンターネットワーク」であり、データセンターネットワークの現状と、今求められている技術に関して集中的な議論を行うために開催された。ゲストスピーカーとして、ヤフー株式会社、さくらインターネット株式会社、NTTコミュニケーションズ株式会社、NECビッグロブ株式会社、ウインドリバー株式会社の方にご参加頂き、各組織にとってのデータセンターネットワークの現状をご報告して頂いた。夏期合宿のプログラムを表2.1に示す。

2.1 IETFレビュー

IETFレビューは、IETFにおいて標準化が進行中の技術の中で、データセンターネットワークの技術に関わりがあると思われるワーキンググループを抜き出し、ここ一年の標準化動向をレビューするためのセッションである。参加者の中から、BoardメンバーとWIDE Projectメンバーが分担し、調査結果を報告した。レビュー対象となったワーキンググループ一覧を、表2.2に示す。

本報告では、まず IETFレビューの中から、特にデータセンターネットワークに関連する技術を抜粋し、その技術動向とレビューにおける議論の結果について報告する。

trillワーキンググループレビューからは、trillに関する標準化動向の報告があり、trillという技術に関しての仕様の確認と、その利用方法に関する議論が行われた。L2ネットワークに経路制御を導入したものであり、リンクの冗

長性と利用効率の向上をねらったものである。LISPワーキンググループレビューでは、LISPの標準化動向と、従来のIdentifierとLocatorの分離という目的のみならず、動的なVPN制御といった用途への利用がいくつか画策され

表2.1 WIDE Project夏期合宿プログラム

2013/08/04	14:00-16:00	IETF Review 1	
	16:00-16:15	Break	
	16:15-18:00	IETF Review 2	
	18:00-20:00	夕食	
	20:00-23:00	基礎知識の共有セッション	(1) データセンターネットワーク要素技術
			(2) SDN とクラウドデータセンター
			(3) データセンターネットワークに関する研究動向
2013/08/05	9:00-12:00	ゲストセッション 1	データセンターとクラウド
			「石狩データセンターとさくらのクラウドの運用について」(さくらインターネット 須藤様)
			「パブリッククラウドサービス Cloudn サービスの現状と課題」(NTT コミュニケーションズ 安達様)
			「BIGLOBE クラウドの現状と課題」(NEC ビッグロブ 津山様)
	12:00-13:30	昼食	
	13:30-15:30	ゲストセッション 2	サービスを支える技術 「実サービス向上のための要素技術」 (ウインドリバー 志方様) 「Yahoo! Japan の DCNW」(ヤフー 高澤様)
	15:45-18:30	ディスカッション 1 (グループワーク)	クラウドサービスとデータセンター データセンターは必要なのか なぜデータセンターがサービスを提供するのか クラウドが提供するサービスとは
	18:30-20:30	夕食	
	20:30-23:00	ディスカッション 2 (グループワーク)	データセンターネットワークはインターネットの特異点なのか データセンターによってネットワークは変わったのか データセンター内部と外部 データセンター間ネットワーク
2013/08/06	9:00-10:30	ディスカッション 3	これからのデータセンターモデル あるべき姿は 足りない技術は 解決すべき課題は
	10:45-12:00	WrapUp	まとめと今後の方針

ている、という現状が確認された。

また、l2vpn、l3vpnワーキンググループレビューでは、MPLSやVPLSを利用したトラフィックエンジニアリングの動向や、新技術としてのEVPN(Ethernet VPN)、MVPN(Multicast VPN)の動向に関して報告が行われた。VPN構成のシグナリングプロトコルとして、BGPやPIM-SMを利用する提案が多く出されており、特にBGPを用いたシグナリングはMPLSのシグナリングプロトコルとしてデファクトスタンダードの地位を得ていることが確認された。さらに、nvo3ワーキンググループレビューでは、オーバーレイネットワークの一般的な概念化についての報告があり、特定の技術仕様を標準化するのではなく、オーバーレイネットワークというものの構成要素とその概念を定義するワーキンググループであることが共有された。nvo3ワーキンググループでは、主にL3ネットワークの上にオーバーレイネットワークを構築することが概念化されており、VXLANやNVGREといったオーバーレイネットワーク構築技術をもとに、制御技術としてLISPを用いるといった提案が行われている。

表2.2 IETF レビュー対象ワーキンググループ一覧

Internet Area	lisp	eden
	trill	motonori
	softwires	osamu
Operations and Management Area	eman	hiroshi
	grow	kato
Routing Area	forces	shigeya
	i2rs	k_akashi
	idr	zin
	l2vpn	sekiya
	l3vpn	sekiya
	nvo3	upa
	pwe3	shinoda
Transport Area	mptcp	jo2lxq
	nfsv4	onoe
	storm	sunao
IRTF	sdnrg	sekiya

これらのIETFワーキンググループレビューから、データセンターネットワークの構築には、従来のL2ネットワークを用いて構築される技術と、L3の上に独自のL2もしくはL3ネットワークを構築する、オーバーレイネットワーク技術が用いられるのが主流である、ということが確認された。

2.2 基礎知識の共有セッション

次に、IETFワーキンググループレビューで確認された要素技術に関してさらに理解を深め、実際の運用に関するイメージを把握するために、基礎知識の共有を行うためのセッションが開催された。

■ データセンターネットワーク要素技術

データセンターネットワークで利用されている/されることが期待されている各種技術について、さらなる理解を深めるための技術解説とレビューを行った。具体的に、以下の技術に関して詳細な紹介と議論が行われた。

- VXLAN(draftmahalingamduddtcopsvxlan04)
- TRILL(RFC6325,6326,6327)
- SPB(IEEE802.1aq)
- DCB
- Priority Flow Control(PFC) IEEE802.1Qbb
- Enhanced Transmission Selection(ETS) IEEE802.1Qaz
- Data Center Bridging eXchange(DCBX) IEEE802.1Qaz
- Congestion Notification(CN) IEEE802.1Qau
- EVB(IEEE802.1Qbg)
- draftlapukhovbgproutinglargedc05

DCBやEVBでは、データセンターネットワークにおいて、より低遅延で高速なネットワーク通信を実現するための仕組みや工夫に関して、理解を深めた。ストレージネットワークに利用されるプライオリティ付けの仕組みや、大規模データセンターにおいて、L3とBGPを利用したデータセンターネットワークの構築手法に関する理解を深めた。

■ SDNとクラウドデータセンター

データセンターネットワークでは、その管理や運用においてSDN(Software Defined Networking)技術を利用する

ことが試みられている。そのため、SDN技術はデータセンターネットワークを議論するにあたって外すことのできない技術要素であり、その正しい理解と議論を行うための、基礎知識の共有を行った。具体的には、以下の技術に関して、詳細なレビューと議論を行った。

- OpenStack
- OpenDaylight
- OpenFlow
- Northbound / Southbound API

特に、SDNの代表的な技術となっているOpenFlowに関しては、特に詳細なレビューと議論が行われた。参加したメンバーの中にも、OpenFlowは集中管理にてフローを制御する、程度の技術知識しか持っていない者もいたため、その仕様と適用できる範囲、実運用のイメージに関して活発な議論がかわされた。

■ データセンターネットワークに関する研究動向

最後に、データセンターネットワークに関する最新の研究動向についてレビューを行った。ここ2年ほどの、SIGCOMMやINFOCOMといった著名な国際カンファレンスにおける、データセンター関連論文をとりあげ、研究テーマとしてどのようなテーマや技術が注目されているのか、レビューと知識の共有を行った。データセンター関連論文では、どのようにデータセンター内部のリンク利用効率を上げるかといった論文が多く見受けられ、データセンター内部のトポロジ構造に関する論文もいくつか見受けられた。

2.3 ゲストスピーカーセッション

ゲストスピーカーセッションでは、各企業から参加して頂いた方に、それぞれの企業内におけるデータセンターネットワークの現状と課題を報告して頂いた。

まず、さくらインターネット株式会社の須藤様より、さくらインターネットの石狩データセンターにおけるデータセンターネットワーク、ならびにデータセンター管理の現状に関して報告を頂いた。電力や空調に関する工夫、また社内におけるサービスの立案から実現までのプロセスをご紹介頂き、SDNを利用したネットワーク攻撃防御

に関する事例をご紹介頂いた。

次に、NTTコミュニケーションズ株式会社安達様より、NTTコミュニケーションズがサービス展開しているCloudnの詳細をもとに、パブリッククラウド運用の現状と課題に関してご報告を頂いた。Cloudnは、IaaS(Infrastructure as a Service)型のクラウドサービスであり、パブリッククラウドと呼ばれる、多くのユーザに対してサービスを提供する形態のクラウドである。その中で課題となっているのは、VLAN数の制限に起因するデータセンターネットワークの構成の不自由さであり、仮想ルータの信頼性や冗長性問題であるとの報告があった。VLAN数の制限に起因する問題は、多くのデータセンター管理者から述べられている問題であり、特にIaaSパブリッククラウドの場合は、ユーザがクラウド内部で利用するネットワーク数だけでも莫大な数になるため、既存のVLANの4096という識別子数ではあつという間にネットワーク数が不足することが既知の問題となっている。Cloudnでは、VXLANを導入することでその問題の解決に取り組んでいるとの報告があった。やはり、VXLANはVLAN数の制限を解決するための技術として注目されているが、その一方で、もともとの仕様ではオーバーレイ網の制御のためにL3マルチキャストを用いており、データセンター内部ネットワークにてL3マルチキャストネットワークを構築する必要がある。そのため、本来のVXLANの仕様通りではないが、単なるトンネリングプロトコルとしてVXLANを用いるという手法が用いられることが多くなっている。この手法は、ユニキャストVXLANと呼ばれており、データセンターネットワーク管理者の間で、VLAN数の不足問題を解決するという観点において注目されている技術であるとの現状が共有された。

次に、NECビッグロブ津山様より、BIGLOBEクラウドの現状と課題についてご報告を頂いた。BIGLOBEクラウドでは、サービス展開の即時性を重視し、SDN技術を用いたサービス展開を開始しているとの報告がなされた。OpenFlowを使ったパスの設定や、VXLANを用いたハイパーバイザー間のパス設定、ならびにそれらのクラウドコントローラとの連携について報告して頂いた。ビッグロブでは、SDNを用いたクラウドサービスを展開しており、その現状と課題を共有し、議論を行った。また、

BIGLOBEクラウドにおいても、さらなるネットワークサービスの仮想化を目指している旨が報告され、今後の課題としてL3ルータの仮想化やデータセンター間連携の実現があげられた。

さらに、ウインドリバー株式会社の小宮山様より、実サービス向上のための要素技術に関するご報告を頂いた。NFV(Network Functions Virtualization)の一環として、ネットワークのパケット処理やプロトコル処理を一部オフロードし、ハードウェア処理することによる高速化をねらった製品に関してのご報告を頂いた。今後のネットワーク仮想化やSDN化によるソフトウェア処理と、ハードウェアを用いた高速化をどう両立していくかに関する一つの方向性についてご報告頂き、議論を行った。

最後に、ヤフー株式会社の高澤様より、ヤフージャパンのデータセンターネットワークに関してのご報告を頂いた。ヤフージャパンのデータセンターネットワークは、その多くの部分がレガシーな技術を用いて構築されており、また独自のツール等を用いて管理の効率化と共通化がなされている、との報告を頂いた。SDNや仮想化はまだ本格的には用いられておらず、ハードウェアの共通化やロードバランサーの駆使などによって、大規模なサイト運用を行なっているとの報告がなされた。一方で、IPv6対応は積極的に行われており、IPv4/IPv6デュアルスタック対応は完了しているとのことであり、データセンター間の連携は積極的には行っていない旨が報告された。

これらゲストスピーカーの方々からの現状報告をうけ、これからのデータセンターネットワークの課題を洗い出し、参加者間で共有することができた。

2.4 データセンターネットワークに関する議論

IETFにおけるデータセンターネットワーク関連技術の標準化動向や関連技術の調査、ならびにゲストスピーカーからのデータセンターネットワークの現状と課題の報告を受け、参加者全員にてデータセンターネットワークの現状に関する評価と、それら技術のあるべき姿に関する議論を行った。

まず、データセンターの成り立ちについて確認を行い、

議論を行った。従来は、電話会社の局舎に機材をコロケーションすることで形成されていたインターネットが、インターネットの拡大とともにサービスのための計算資源や記憶装置を設置する需要が高まり、必然的にデータセンターが生まれた。その後、クラウドやビッグデータ処理の台頭によってデータセンターは確固たる地位を得て現在に至る、という歴史が確認された。特に、近年のクラウドの台頭により、自社内に機材を置くのではなく、データセンターに自社機材を設置するオフプレミス型のプライベートクラウドの利用や、クラウド事業者がデータセンターに設置した機材を利用してサービスを展開している、パブリッククラウドの利用が増えているため、よりデータセンターの利用と重要性が高まっていることが確認された。この場合、現在のクラウド利用にて需要が高まっているのが、オフプレミスへの機材移行である。その過程としてオンプレミス環境と、オフプレミス環境のプライベートクラウド環境をVPNや専用線で接続し、ハイブリッドな運用やシームレスな移行を行いたいという需要が高まっているとの報告がなされた。これは、オフプレミス環境のプライベートクラウドにおいてもユーザにある程度のネットワーク構成の自由度を与えることであり、VPNや専用線によってプライベートクラウドとユーザ、もしくはプライベートクラウド同士を接続する技術が必要とされていることであると結論づけた。さらに、ユーザにネットワーク構成の自由度を与えるにあたっては、先に述べたデータセンターネットワーク内部のVLAN数の制約が存在するため、VLAN以外のVXLAN等の技術を用いてVLANとVXLANの相互変換を行う必要があることも確認された。

一方で、これらの技術が本当に問題を解決するための正しい手法なのか、という問題提起もなされた。VPNやオーバレイを用いて、ユーザ単位のネットワークを構築することは可能であるが、集中管理や冗長性の無い技術になりがちであり、従来のインターネットが有する特性を損なっているとの指摘もなされた。

最終的に、データセンターネットワークにおいては、やはり多くのL2ネットワークを構成することが求められる一方、従来のL2を構築する技術であるVLANやSTPでは、広域かつ多数のL2ネットワークを管理しにくい、と

というのが根本にある問題ではないか、という議論がなされた。この解法としては、そもそもL2管理の問題点となる、ブロードキャストやマルチキャストといった、シェアドメディアを想定したプロトコルを利用しないようEthernetプロトコルを改造するであるとか、はじめからデータセンター間をインターコネクトできるような仕様を決定しておくであるとか、もっと運用管理をシンプルに行い、かつ相互接続するための技術を作るのではないか、という議論が活発になされた。

2.5 WIDE Project夏期合宿のまとめ

WIDE Project夏期合宿では、データセンターネットワークの現状と課題に関して集中的な議論を行った。現在、データセンターネットワークの構築に用いられている要素技術を調査し、その最新動向を把握するとともに、実際の事業者の方々から運用の現状と課題を伺うことで、より実践的なデータセンターネットワークの議論を行なうことを目指した。ゲストスピーカーの方々も加わって頂き、白熱した議論が展開された結果、これからのデータセンターネットワークは、やはりL2 over L3のオーバーレイネットワークも含んだL2技術の改善と、データセンターとユーザ間、もしくはデータセンター間のインターコネクトに関する技術の重要性がより高まっていくであろうという結論を得た。この結論に基づき、WIDE Projectでのさらなる議論と、cloudワーキンググループを中心とした活動を行っていくことが確認された。

第3章 WIDE Project秋期合宿における議論

WIDE Project秋期合宿は、2013年9月10日から2013年9月13日の日程において、信州松代ロイヤルホテルにて開催された。その中のボードプレナリセッションにおいて、夏期合宿の報告と、より多くのWIDEメンバーを交えた議論が行われた。

まず、夏期合宿にて行われた報告と議論のまとめを報告し、参加者全員での知識と議論過程の共有を行った。夏期合宿にて導き出されたデータセンターネットワークにおける課題としては、

- ブロードキャストとマルチキャストの無いネット

ワーク構築

- インターコネクトできるアーキテクチャ
- 識別子のグローバルな管理
- 遅延の克服

が報告された。

参加したメンバーからは、VXLANの利用やオープンな技術を用いたデータセンター間の接続に関する意見が出され、仮想化の上に仮想化を重ねるような方式ではなく、シンプルにインターコネクトを実現する技術が必要である、との認識が共有された。その一方で、遠隔地同士の接続において、ネットワークが遅延を超えられないのは事実であり、ストレージとネットワークの中間層において、遅延を吸収するための技術が必要であるとの認識も共有された。これら課題を、学生は卒論や修論、企業の方はサービスの要素技術として生かせるような議論と活動の方向性が、これからのWIDE Projectに必要なものであるとの認識が共有された。

第4章 WIDE Project報告会におけるパネルセッション

2013年10月9日に、慶應義塾大学三田キャンパスにおいて、WIDE Project拡大報告会が開催された。その中で初めての試みとして、パネルディスカッションを開催した。これは、WIDE Projectがデータセンターネットワークの議論と課題の解決に注力していることを示すために、あえて活動報告とは別の形で、特別パネルセッション「データセンターネットワーク ～今ある未来とまだ無い未来～」と題して活動報告させて頂いた。このパネルセッションは、NTTコミュニケーションズ株式会社 安達様、さくらインターネット株式会社 田中様、アラクスネットワークス株式会社 新様にご登壇頂き、WIDE Projectボードメンバーからは関谷と宇多が参加した。

データセンターネットワークは、「ネットワーク」ではあるが「インターネット」ではないという認識の元に、それでもインターネット内に存在するネットワークであれば、閉じた技術でもいいのか、また相互接続できなくても良いのか、といった観点を中心として、データセンターネットワークの今見えている未来とまだ見えていない未

来に関してパネルディスカッションを行った。

さくらインターネット株式会社の田中様からは、L2 over L3のオーバーレイネットワークの必要性和、シンプルかつ高信頼な技術を求める意見が述べられた。現在話題となっているOpenFlowやSDNは、技術的に高機能であるが、複雑であり信頼性が乏しい、といった指摘がなされた。オーバーレイネットワークをより柔軟に、かつシンプルに管理できる技術を求めているとの結論であった。

次に、NTTコミュニケーションズの安達様からは、VLAN数の制限に関する課題や、その解決案としてのVXLANの利用について意見が述べられ、クラウドコントローラとの連携のためのAPI標準化が必要であると結論が述べられた。

さらに、アラクスラネットワークス株式会社の新様からは、新たな技術の実装までに必要とする時間とコストによる、製品化までのタイムラグ問題や、標準化に関わる人間が正しい場所にて発言していないといった問題点が述べられた。

最後に、WIDE Projectボードメンバの宇多から、データセンターネットワークのL2に対する要求の高さや、その管理の要件、集中管理の是非に関する意見が述べられた。WIDE Projectがすべき活動として、軽いフットワークで実証実験を行いつつ、オープンな技術による解決策を探るべき、との意見が述べられた。

報告会にてパネルディスカッションを行なうという初めての試みであったが、報告会参加者からも、複数の意見が出され、議論に加わって頂くことができた。今後もWIDE Projectは、データセンターネットワークの課題解決に関して、活動を継続していく旨を報告させて頂いた。

第5章 データセンターネットワークのあるべき姿

以上にて、WIDE Project夏期合宿、秋期合宿、ならびに拡大報告会の報告を述べた。データセンターネットワークの課題解決は、便利だから、コストが安いから、といった観点のみでそれぞれが独自の仕様で独自の解決策を導入すると、次のステップへの大きな足枷になると考えられる。その例が、データセンターとユーザ、もしくはデータセンター間のインターコネクトであり、複雑化による信頼性の欠如である。

データセンターネットワークに求められる「ネットワーク」とは、従来の「インターネット」とは異なる性質であるということを認識する一方で、データセンターネットワークもインターネットに接続されるネットワークである、ということを忘れないよう、設計と構築を行っていく必要がある。

その例として、夏期合宿での結論である、

- ブロードキャストとマルチキャストの無いネットワーク構築

- インターコネクトできるアーキテクチャ

という2大要件があげられる。これらは、従来のインターネットと異なる性質の「ネットワーク」であるが、オープンかつ相互接続可能な、シンプルなネットワークを目指す必要があるという事例になっている。

今後もWIDE Projectは、WIDEクラウドという実証実験の場を有しており、この利点を活かしたデータセンターネットワーク課題解決のための要素技術の研究と実証実験を行っていく所存である。