Interop Tokyo 2013 ShowNet テクニカルレポート

Interop Tokyo 2013 / 2014 NOC チームメンバー

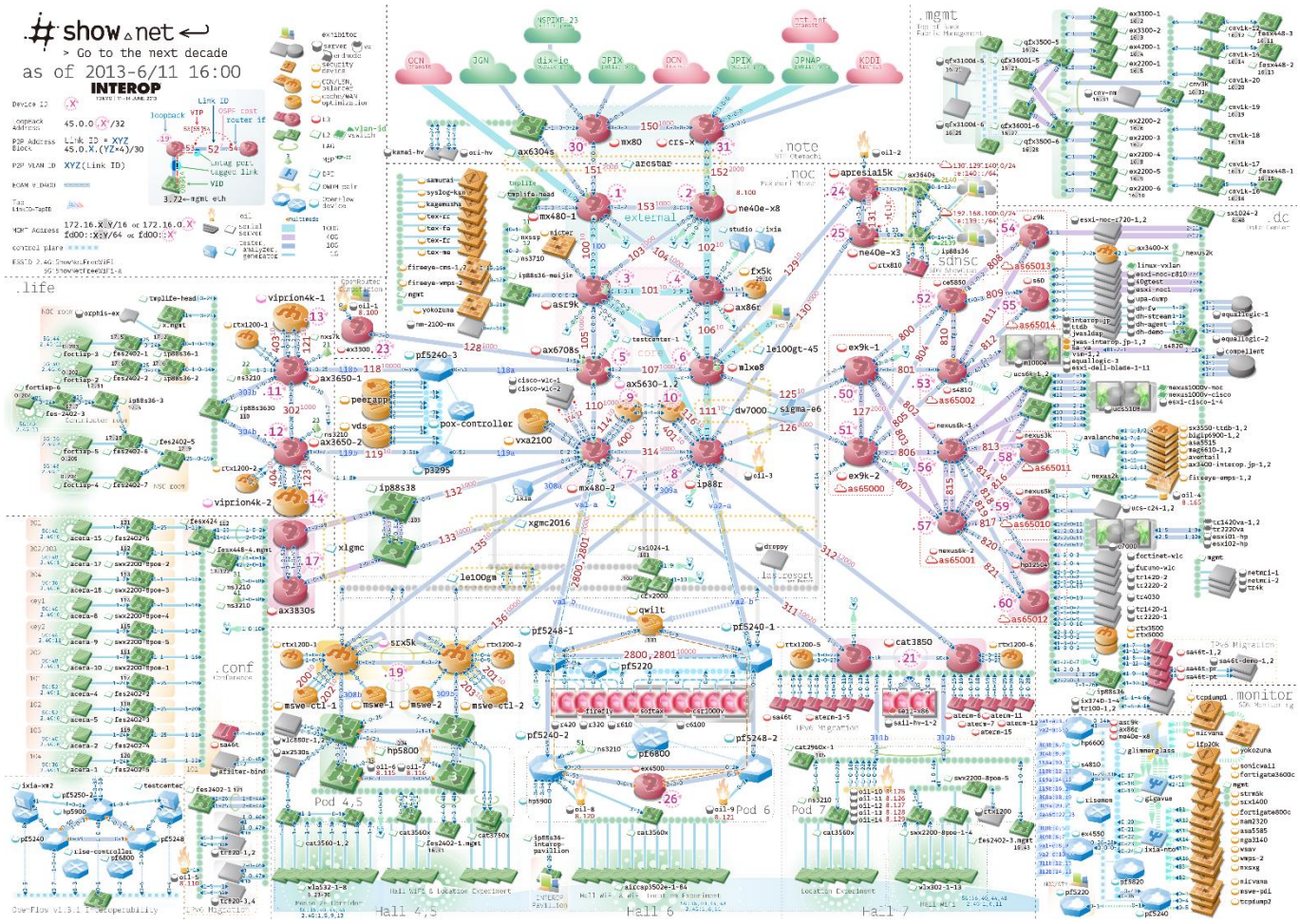
1 はじめに

Interop Tokyo ShowNet は 2013 年の開催にて 20 回目の実施となりました。これまで、多くの最新技術の相互接続検証やネットワーク設計、運用におけるチャレンジを ShowNet NOC チームメンバーでは行ってきました。しかしながら、その相互接続検証の結果や技術上のノウハウが広く世の中に普及しているかといえば、そうとはいえないのが現状です。この状況をかんがみ、Interop Tokyo 2013 ShowNet から、これまで公開してきた プレゼンテーションフォーマットの報告書以外に、電子情報通信学会などでの学術研究界での発表や、技術文書として ShowNet にて取り組んだ相互接続検証や技術チャレンジとその運用結果を報告しようと考えております。その取り組みの中の一環として、本技術報告書では Interop Tokyo 2013 ShowNet NOC チームメンバーにより執筆・発表された、学術研究界での発表内容をテクニカルレポートとしてまとめ、ShowNet での相互接続検証結果や技術チャレンジの内容を広く公開するものです。このテクニカルレポートが学術での研究や産業界での実用化などのきっかけになれば幸いです。

Interop Tokyo 2013

ShowNet NOC チーム一同

2 INTEROP 2013 SHOWNET 概要



上記の図は Interop 2013 ShowNet のネットワークトポロジー図です．Interop 2013 ShowNet では，つぎのようなチャレンジを行いました．

- 100G 対外線＋100G バックボーン（図中上段）
- One armed 構成の 44/444/46/64 トランスレータ（図中央，図右中央，図下中央ホール 4 など）
- IPv4/IPv6 共存・移行技術による出展者収容（図下中央，ホール 7）
- V6 only バックボーンと 464 移行技術を用いた会議棟ネットワーク構築（図中央左）
- OpenFlow と virtual appliance router の NFV による出展者収容（図下中央，ホール 6）
- BGP によるデータセンターバックボーン（図中央右 .dc）
- インラインキャッシュ（図中央）
- イーサファブリック，TRILL による管理ネットワーク構築（図上右.mgmt）

- OpenFlow 1.3.1 相互接続とアクセスコーナー（図左下）
- 会場すべてでの無線 LAN アクセス提供トライアルと電波観測・監視運用技術の検証（図下）
- DNS64/NAT64 を用いた IPv6 Only WiFi アクセス提供実験（図下）
- WiFi によるロケーションサービス提供実験（図下）
- OpenFlow を用いた観測網構築（図右下）

など，さまざまなデザインとトライアルを実施しました．Interop 2013 ShowNet のデザインはデザイン担当 NOC 河口大樹さん曰く「A3 で描ける限界が来た！これ以上無理！」と言わしめたほどの密度になっています.

このカッティングエッジネットワークは，NOC チームだけで出来たものではなく，多数のコントリビュータの皆様からの機器提供，各コントリビュータのコントリビュータスペシャリストの皆様とのディスカッションを通してデザインを進め，ホットステージ期間中に STM（ShowNet Team Member）や CTM（Contributor Team Member）などボランティア参加の皆様との構築と検証を通して，やっと完成したネットワークです．ホットステージはその名のとおり，まさに「鉄火場」です．皆さん睡眠時間を削って構築，チューニング，検証を行ってくださいました（楽しいのか，はまっているのかはわかりませんが，「明日もあるから早く帰ってください」とお願いしてもなかなか帰ってくれないのです（涙））． おかげさまで，2013 は「近年まれに見るほどの安定運用」でした.

さまざまな皆様のご協力の結果, いろいろな相互接続検証結果や運用知見を得ることが出来ました. NDA の関係上公開できない内容もあるのですべてを皆様にお伝えすることは出来ませんが，公開できる相互接続検証結果や運用知見を次の章から順次説明して行きたいと思います.

3 BGP を用いたデータセンターネットワークの構築

本章では、2013 年 9 月 6 日に開催された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（IA 研究会）にて発表した論文に修正を加えたものを掲載する。本内容は Interop 2013 Tokyo ShowNet のデータセンターネットワークを構築する上で BGP を用いた耐障害性の高い、スケーラブルなデータセンターネットワークの構築手法に取り組んだ結果をまとめたものである。

BGP を用いたデータセンターネットワークの構築
－Interop Tokyo 2013 における ShowNet data center の取り組み－

奥澤 智子† 関谷 勇司‡

†株式会社データホテル 〒160-0022 東京都新宿区新宿 6-27-30
‡東京大学 情報基盤センター 〒113-8658 東京都文京区弥生 2-11-6
E-mail: †okuzawa@datahotel.co.jp, ‡sekiya@nc.u-tokyo.ac.jp

あらまし 多くのサービスがインターネット上で展開され，データセンター内で管理運用されるサーバ台数は日増しに増大している．しかしながら，そのサーバを収容するネットワークは，レイヤ 2 ネットワークで構成されている事例が多く，拡張性と耐障害性に課題を抱えていると言える．そこで本論文では，経路制御プロトコルとして BGP のみを用いたレイヤ 3 データセンターネットワークを設計，及び構築した事例を述べ，レイヤ 2 ネットワークの影響範囲を狭めたネットワーク構成における，拡張性と耐障害性を考察する．

キーワード データセンター, BGP, 拡張性, 耐障害性

Bulding a data center network using BGP as the routing protocol
－The challenge of ShowNet data center in Interop Tokyo 2013－

Tomoko OKUZAWA† Yuji SEKIYA‡

†DataHotel Co., Ltd. 6-27-30 Shinjuku, Shinjuku-ku, Tokyo, 160-0022 Japan
‡Information Technology Center, The University Tokyo 2-11-16 Yayoi, Bunkyo-ku, Tokyo, 113-8658 Japan
E-mail: †okuzawa@datahotel.co.jp, ‡sekiya@nc.u-tokyo.ac.jp

Abstract Many services are deployed on the Internet, the number of servers that are supported in data center is increasing day by day. However, most of the network that accommodates servers used layer 2 network protocol has a problem in stability and scalability. In this paper, the designing and building data center network using BGP as the only routing protocol are described. Furthermore, stability and scalability in Layer 3 only design are discussed.

Keyword data center, BGP, network stability, network scalability

3.1 はじめに

データセンターサービスを用いて Web システムを構築するユーザは，レイヤ 2 ネットワーク構成を好む傾向にある．その理由としては，ロードバランサの接続方式における制限事項があげられる．

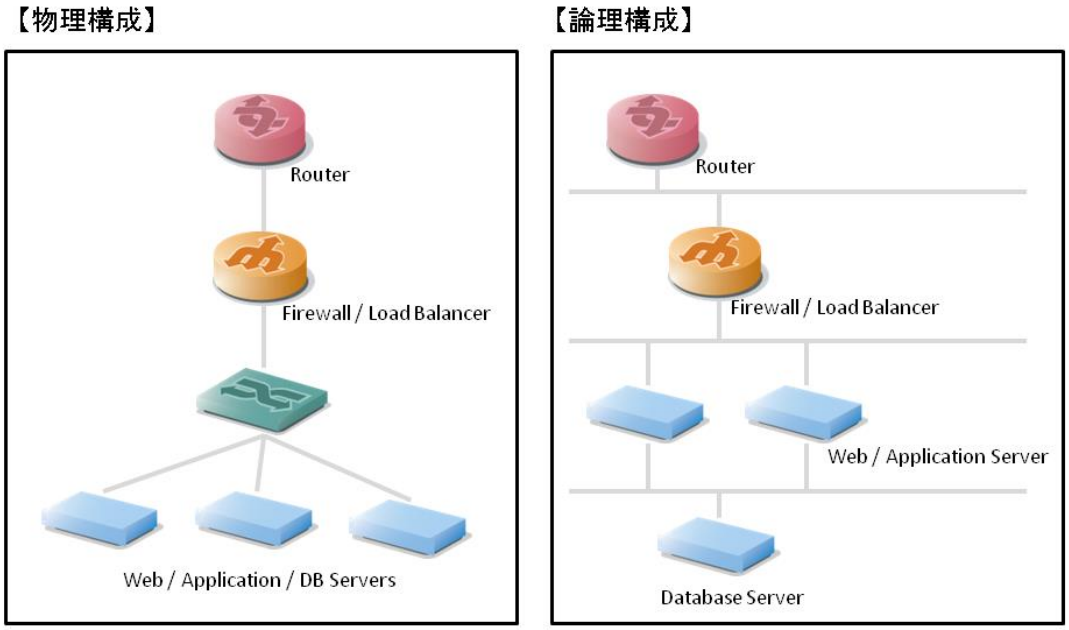


図 1 一般的な 2 層 Web/DB システム

図 1 に示す通り，現在の一般的な Web システムにおいては，冗長性を確保する為，ロードバランサを導入する事例が多い．ロードバランサを NAT (Network Address Translation) 方式や，レイヤ 2 DSR (Direct Server Return) 方式で用いた場合，サーバとロードバランサが同じレイヤ 2 ネットワークに収容されていなければならない制限がある．よって，必然的にレイヤ 2 ネットワークで構成する事を選択するのではあるが，一方では，サーバ管理者がサーバを追加する際，ネットワーク管理者に頼らずに拡張できるという，管理運用上の側面もあると考える．

システム規模が数十台のサーバで構成される場合は，レイヤ 2 ネットワークで構成する事は常である．しかしながら，その規模が数万台のサーバで構成され，トラフィック量も数十 Gbps を超えるサービスシステムに成長した場合，レイヤ 2 ネットワーク上で，ブロードキャストストームに代表されるネットワーク障害が発生すると，その影響は大きい．

故にここ数年，レイヤ 2 ネットワークにおける耐障害性という欠点を補う技術として TRILL[1]が，STP に代わり用いられる構成が増えてきているが，現状ではまだ，ベンダ独自の仕様で拡張実装されている機器が多く，構成上のベンダダイバーシティが困難な為，サービス事業者の設計ポリシーによっては導入されにくい側面もある．

大規模システムになるほど，レイヤ 2 ネットワークにおける障害が起きた場合の影響は大きい．そして，その影響を限定する為にネットワークの細分化を試みると，分割単位として用いられる VLAN 番号の制限数に達し，それ以上の拡張が見込めなくなる．よって現状では，大規模システムを支えるネットワークには，耐障害性と拡張性において課題を抱えていると言える．

そこで本論文では，この課題に対して試みた，レイヤ 3 データセンターネットワークの設計，及び構築事例を述べ，レイヤ 2 ネットワークの影響範囲を狭めたネットワーク構成における，耐障害性と拡張性を考察する．

3.2 レイヤ 3 データセンターネットワークの構築

本節では，大規模システムにおける BGP[2]を用いたデータセンターネットワークを提案する IETF Internet-Draft “Use of BGP for routing in large-scale data centers” [3] を，Interop Tokyo 2013 ShowNet において設計，及び構築を行った事例を述べる．

ShowNet とは，Interop Tokyo が開催されるイベント会場に構築するライブデモンストレーションネットワークの総称であり，開催期間中は実際に，出展者や来場者に対してネットワークコネクティビティを提供するという機能も有している．2013 年の ShowNet データセンターは，前述の IETF Internet-Draft を基に，ToR（Top of Rack）装置が並ぶアクセス層，ToR 装置を集約するアグリゲーション層，及びコア層の 3 階層で構成し，そのネットワーク装置間は全て，BGP によって経路制御を行う設計とした．ここで構築したネットワークの構成概要を，図 2 に示す．

現在の一般的なデータセンターネットワークは，アグリゲーション層以下はレイヤ 2 ネットワークで構成される事例が多い．一方，本構成においては，レイヤ 2 ネットワークの範囲はサーバと ToR 装置間のポイント・トゥ・ポイントネットワークのみに留めた．ToR 装置までレイヤ 3 ネットワークで構成した点が特徴である．

BGP の管理単位となる AS（Autonomous System）番号は，プライベート AS 番号を用い，アクセス層においては，各 ToR 装置毎に番号を付与し，区分した．

また，各装置における，BGP による広告経路は，図 2

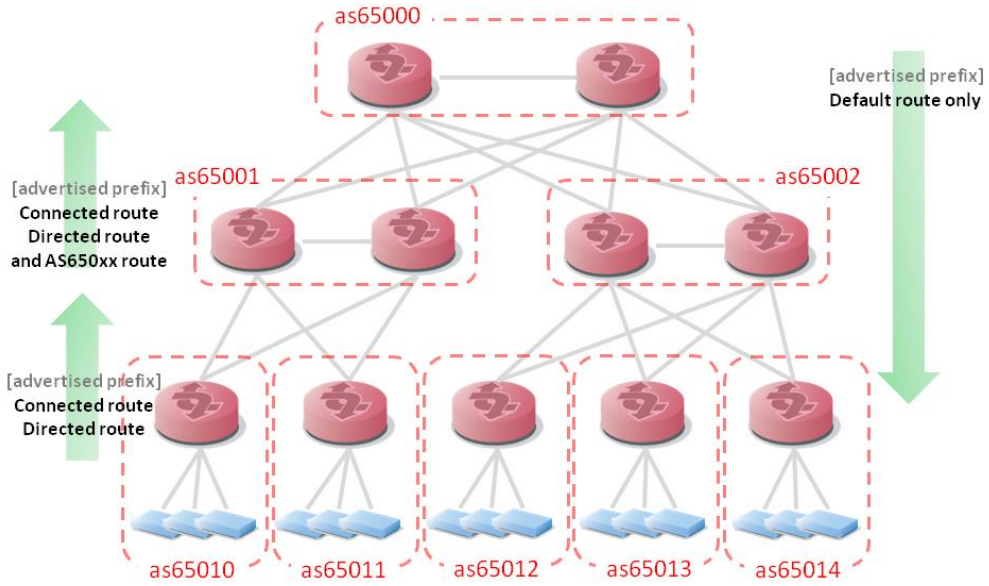


図 2　ShowNet データセンターネットワーク概要

に示す通り，下流の AS に対してはデフォルト経路のみを広告し，上流の AS に対しては，各装置が保持する直接接続経路，及び静的経路を BGP に再配信し，下流 AS から受信した経路と併せて広告するように設計した．直接接続経路，及び静的経路の再配信は，ToR 装置にサーバを接続したと同時に対象経路を広告させ，サーバに対してすぐにネットワークコネクティビティを提供するという目的の為である．以上の内容にて，ShowNet 内のデモンストレーション用サーバを収容する為のデータセンターネットワークを設計し，構築を実施した．

次節では，構築を通じて明らかになった，本構成における設計上の留意点を述べる．

3.3 構築上の留意点

BGP で経路制御を実施しているネットワークと，その他の経路制御プロトコルで制御しているネットワークを接続する際は，その経路交換方法によっては，経路ブラックホールや，意図しない経路が選択される場合があるので設計時に予め留意しておく必要がある．例として，図 3 を示す．

データセンター内の BGP 経路を，境界に位置する装置から他ネットワークの装置へ，OSPF にて経路の再配信を実施した場合，R3 のルータにおいては，R5 から受信した BGP 経路，及び R1 から受信した

OSPF 経路を保持する事となる．多くのベンダにおいて，その装置内の初期経路優先順位は，BGP 経路よりも OSPF 経路が高く設定されている為，R3 において，対象経路宛ての通信は R1 に転送される．結果，R6 から R5 宛ての通信が，R3 に転送された場合，R6→R3→R1→R2→R4→R5 という経路を辿り，データセンター内の通信が，データセンター外を経由するという意図しない経路選択をしてしまう事象が発生する．

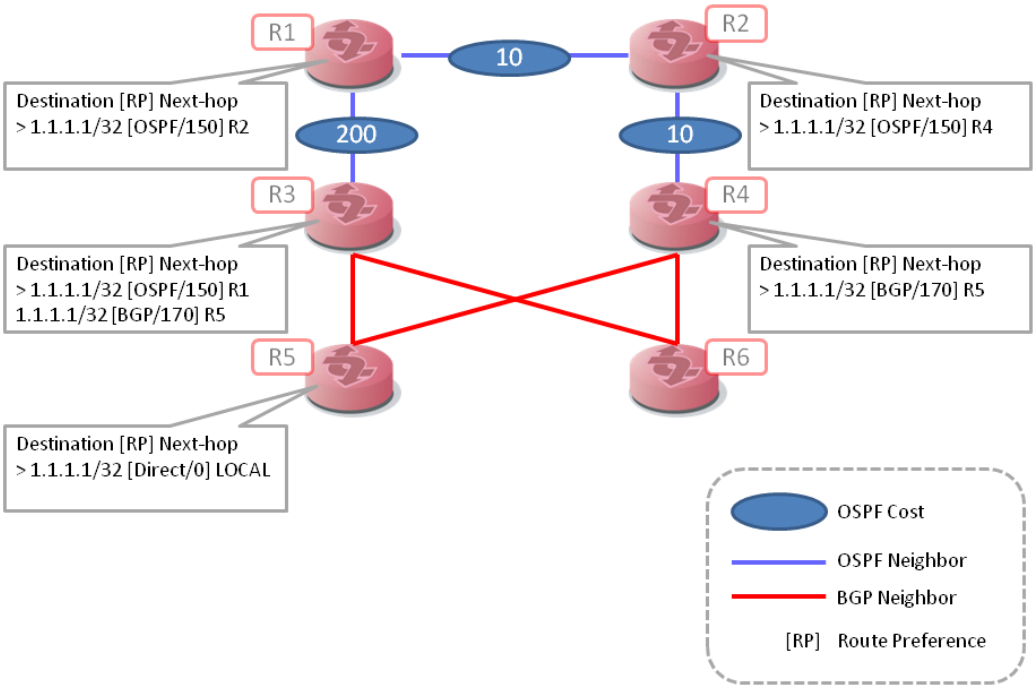


図 3 意図しない経路選択例

このような事象を回避する為には，他プロトコルに再配信する際，その境界装置内の経路優先順位において，BGP 経路が該当するプロトコル経路よりも高くなるよう，優先値の調整を実施した方が望ましいと言える．

3.4 レイヤ 3 データセンターネットワークの利点と課題

本節では，耐障害性と拡張性に関する本構成の利点と，ShowNet におけるレイヤ 3 データセンターネットワークの構築，及び運用を通じて考察した，課題について述べる．

耐障害性と拡張性については，レイヤ 3 プロトコルがレイヤ 2 プロトコルに比べ，元々優位を持つ性質である．

殊に耐障害性に関しては，障害時における影響範囲を局所化できる事が，本構成の一番の利点である．運用上，人為的ミスを含めたレイヤ 2 ネットワークにおけるストーム障害は少なくない．また，ストーム障害が一旦発生すると，その対応にかかる人的負担は非常に大きい．本ネットワークの構成は，レイヤ 2 ネットワークの局所化により耐障害性の向上が見込まれる為，それに伴い運用に係る負担も軽減するであろうと考える．

また，拡張性については，全てのネットワーク装置間は BGP を用いた経路制御を実施している為，サービスの成長に応じて, AS 単位で，容易に規模を拡張する事ができる．この点は，拡張性における一番の利点である．

さらに本構成においては，エニーキャスト[4]によるサーバの拠点間負荷分散, RTBH (Remote Triggered Black Hole Filtering) [5]による DoS 攻撃対処など, BGP に元々備わる経路制御の仕組みのみを用いて，データセンターサービスの運営上，利便的な機能を拡充する事も可能である．この点は，機能の拡張性という側面から見る，拡張性の向上であると言える．

一方，構築，及び運用を通じて明らかになった課題もある．ひとつは，障害時の収束時間である．障害時に求める復旧までに要する時間は利用者，及びサービスの性質によって異なるが，数ミリ秒の収束時間を期待する場合は, BGP による経路収束時間はそぐわない．

また，現在のサーバ仮想化技術の普及を考えるに, VM の移動に用いるレイヤ 2 ネットワークの必要性は無視できない．この点については, VXLAN[6]や NVGRE[7]などのオーバーレイ技術と組み合わせて，本ネットワークの利点を生かしながら，利用者の要求に応じる事ができるネットワークにする為，更なる設計の検討が必要になる点であろう．

3.5 おわりに

本論文では, Interop Tokyo における ShowNet データセンターにおいて挑戦した，耐障害性と拡張性の向上を目的とする, BGP を用いたレイヤ 3 データセンターネットワークの構築事例を述べた．

ShowNet はデモンストレーションネットワークであるが，本構成を実際の商用サービスに適用すると想定した場合，その費用対効果については，各々の事業者において十分な検討がなされるであろう. BGP がサポートされた機器は，非常に高額な印象があり，設備投資上敬遠される可能性があるが，今回 ShowNet の構築で用いた ToR 装置は全て, ToR 用途で実際に販売されている機器である．また, BGP は WAN で使用するプロトコルのイメージが高い為，運用者の慣れを懸念し，運用上敬遠される可能性もあるが，今回用いているプロトコルは BGP のみである為，一方では平易な運用が期待できる．

よって今後，システムが大規模化し，レイヤ 2 ネットワークでの拡張性に限界を感じ，ネットワーク更改を余儀なくされた場合，本レイヤ 3 データセンターネットワーク構成が，検討の土台にあがる事を期待したい．

文　献

[1] R. Perlman, D. Eastlake 3rd, D. DuttS. Gai, A. Ghanwani, “Routing Bridges (RBridges): Base Protocol Specification”, Internet Engineering Task Force, Request for Comments 6325, July 2011.

[2] Y. Rekhter, Ed., T. Li, Ed., S. Hares, Ed., “A Border Gateway Protocol 4 (BGP-4)”, Internet Engineering Task Force, Request for Comments 4271, January 2006.

[3] P. Lapukhov, A. Premji, J. Mitchell, Ed., “Use of BGP for routing in large-scale data centers”, Internet Engineering Task Force, Internet-Draft, draft-lapukhov-bgp-routing-large-dc-05, July 2013.

[4] J. Abley, K. Lindqvist, “Operation of Anycast Services”, Internet Engineering Task Force, Request for Comments 4786, December 2006.

[5] W. Kumari, D. McPherson, “Remote Triggered Black Hole Filtering with Unicast Reverse Path Forwarding (uRPF)”, Internet Engineering Task Force, Request for Comments 5635, August 2009.

[6] M. Mahalingam, D. Dutt, K. Duda, P. Agarwal, L. Kreeger, T. Sridhar, M. Bursell, C. Wright, “VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks”, Internet Engineering Task Force, Internet-Draft, draft-mahalingam-dutt-dcops-vxlan-04.txt, May 2013.

[7] M. Sridharan, A. Greenberg, Y. Wang, P. Garg, N. Venkataramiah, K. Duda, I. Ganga, G. Lin, M. Pearson, P. Thaler, C. Tumuluri, “NVGRE: Network Virtualization using Generic Routing Encapsulation”, Internet Engineering Task Force, Internet-Draft, draft-sridharan-virtualization-nvgre-03.txt, August 2013.

4 ETHER OAM/CFM の相互接続の課題とこれから

本章では Interop Tokyo 2013 で実施した Ether OAM /CFM の相互接続検証の結果と、相互接続検証および運用での課題をまとめたものを報告する。なお、本章の内容は、2013 年 9 月 6 日に開催された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（IA 研究会）にて発表した論文に修正を加えたものを掲載する。

Ethernet OAM／CFM の相互接続の課題とこれから
－Interop Tokyo 2013 ShowNet Ethernet OAM の取り組み－

長谷川 幹人† 関谷 勇司‡

†エスアイアイ・ネットワーク・システムズ（株） 〒261-8507 千葉県千葉市美浜区中瀬 1-8
‡東京大学 情報基盤センター 〒113-8658 東京都文京区弥生 2-11-16
E-mail: †mikito.hasegawa@sii.co.jp, ‡sekiya@nc.u-tokyo.ac.jp

あらまし Interop Tokyo／ShowNet では、2009～2013 の 5 年に渡り、Ethernet OAM(EOAM)の相互接続性の確認や、実運用化に向けた取り組みを行っている。そこで本論文では、これまでの EOAM／Connectivity Fault Management(CFM)で取り組んだ結果を元に、CFM の相互接続の現状と課題を述べる。その上で、これからのイーサネットレイヤの運用監視をよくするために何が必要なのか、どのような解決方法が考えられるのか考察したい。

キーワード Ethernet OAM, Connectivity Fault Management(CFM), ネットワーク監視,

And future challenges and current status of the interconnection of Ethernet OAM/CFM
－Efforts ShowNet Ethernet OAM in 2013 Interop Tokyo－

Mikito HASEGAWA† Yuji SEKIYA‡

† SII Network Systems Inc. 8, Nakase 1-chome, Mihama-ku, Chiba-shi, Chiba, 261-8507 Japan
‡ Information Technology Center, The University of Tokyo 2-11-16 Yayoi, Bunkyo-ku, Tokyo, 113-8658 Japan
E-mail: †mikito.hasegawa@sii.co.jp, ‡sekiya@nc.u-tokyo.ac.jp

Abstract At Interop Tokyo / ShowNet, over 5 years of 2009-2013, I have done check and interoperability of Ethernet OAM (EOAM) , efforts to operationalize real. In this paper, based on the result of efforts by EOAM / Connectivity Fault Management (CFM) so far, we describe the current status and issues of mutual connection of CFM. On top of that, you want to consider what's necessary in order to improve the operational monitoring of Ethernet layer in the future, what solutions or be considered.

Keyword Ethernet OAM, Connectivity Fault Management(CFM), Network Management,

4.1 はじめに

Interop Tokyo／ShowNet では、2009～2013 の 5 年に渡り、Ethernet OAM(EOAM)の相互接続性の確認や、実運用化に向けた取り組みを行っている。そこで本論文では、これまでの EOAM／Connectivity Fault Management(CFM)で取り組んだ結果を元に、CFM の相互接続の現状と課題を述べる。その上で、これからのイーサネットレイヤの運用監視をよくするために何が必要なのか、どのような解決方法が考えられるのか考察したい。

4.2 ETHERNET OAM 概要

本節では、Ethernet OAM について概要を述べる。本プロトコルは、Ethernet Operation, Administration, and Maintenance(OAM)という名称の通り、「イーサネット上で利用可能な保守・管理・運用機能」を定義しているものである。ただ一言に、Ethernet OAM(以降 EOAM)と言っても複数の規格が存在する。IEEE では 802.3ah(Clause 57)Link OAM[1]と、802.1ag Connectivity Fault Management(CFM)[2]の 2 つの規格があり、ITU-T では Y.1731 OAM Functions and Mechanisms for Ethernet -Based Networks[3]が、TTC では JT-Y1731 OAM functions and mechanisms for Ethernet based network[4]が、MEF では MEF16 Ethernet Local Management Interface (E-LMI)[5]がそれぞれ定義されている。各規格はそれぞれ適用範囲が異なっており、Link OAM は機器間の物理回線や通信状態の管理を、CFM は VLAN 単位の障害監視を、Y.1731 はキャリア利用を意識した CFM のアップグレード版を、JT-Y1731 は Y.1731 TTC 版を、最後の MEP16 は UNI 間で情報をやり取りするインタフェース仕様という適用範囲になっている。そのため、それぞれの適用されるレイヤも異なっている。レイヤイメージを図 1 に示す。

図 1：EOAM レイヤイメージ

Link OAM はトランスポートレイヤを、E-LMI はサービスレイヤを、CFM、Y.1731、JT-Y1731 はトランスポートレイヤ～ネットワークレイヤ～サービスレイヤという適用となっている。

4.3 CFM 概要

本節では、CFM の概要について述べる。(詳細は[2]を参照のこと) 本プロトコルは、前述の通り「VLAN 単位の障害監視を主目的としたプロトコル」である。本プロトコルでは Continuity Check(CC)、LoopBack(LB)、LinkTrace(LT)という 3 つの機能が定義されている。CC は、CFM の主となる機能であり、監視ポイント間のフレーム定期送信による疎通性確認を行う機能で、MEP(Maintenance association End Point)間のイーサネットフレーム疎通性の定期監視（複数の送信間隔が定義）に利用する。LB は、監視ポイント間の疎通確認（IP Ping に相当）を行う機能で、MEP や MIP(Maintenance association Intermediate Point)間のイーサネットフレーム疎通性のオンデマンド確認に利用する。LT は、監視ポイント間の経路確認（IP Traceroute に相当）を行う機能で、MEP-MIP-MEP のイーサネットフレームの経路状態のオンデマンド確認に利用する。

次に CFM の管理ネットワーク階層化の概念について説明する。階層化イメージを図 2 に示す。

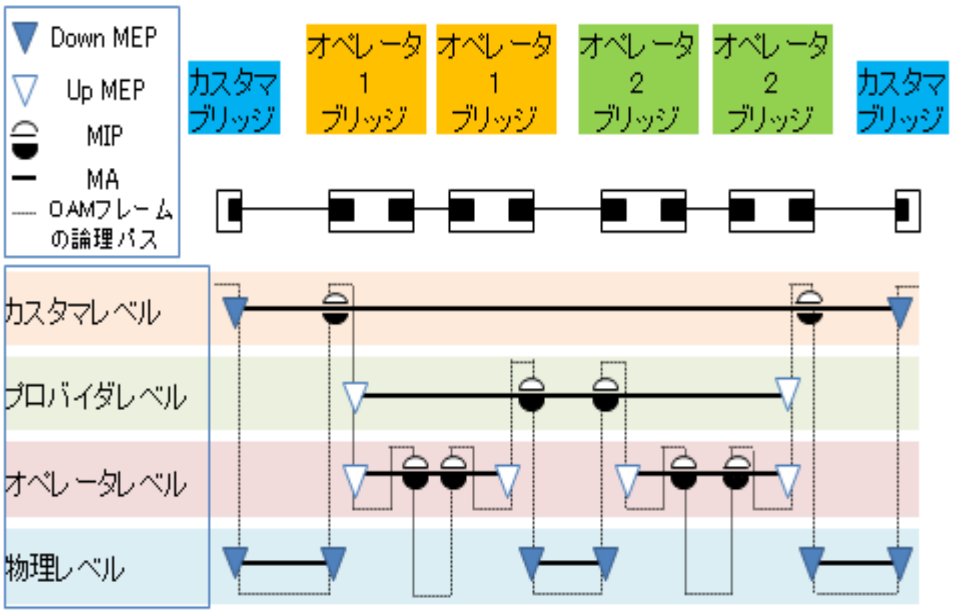


図 2：CFM 階層化イメージ

CFM では、管理対象のグループとして Maintenance Domain(MD)と Maintenance Association(MA)を定義している。MD は同じ管理者によって接続性が管理されているネットワークという位置づけであり、

1 つの MD に最大 8 つの MD レベルが設定できるようになっている。このレベルを用いて 1 つの管理ネットワーク内でのサービスの階層化を提供する。階層化の概念として重要な点は 3 つ。1 つ目は、同じレベルの MD は隣接することがあっても絶対に重ならない。2 つ目は、自分より高い MD レベルの EOAM フレームは透過する。3 つ目は、自分より低い MD レベルの EOAM フレームは廃棄する、の 3 点である。各 MD レベルの適用例として、MD レベル 0～2 が回線事業者、3～4 がサービス事業者、5～7 がサービスユーザー、とされている。MA は同じ MD 内で、実際に接続性を管理する対象(MEP)の集まりという位置づけであり、管理する対象は物理ポートまたは VLAN ごととしている。

1. ShowNet CFM Challenge

本節では、ShowNet における CFM のこれまでの取り組みについて述べる（本稿では、EOAM Challenge として同じく取り組んだ Y.1731 と Link OAM については概要の表 1 のみの掲載とさせて頂きたい）。

	区分	2009	2010	2011	2012	2013
Y.1731	相互接続	—	CC,LB,LT,DM	CC,LB,LT,DM	—	—
Link OAM	相互接続	—	—	—	Discovery,Link Monitoring,Remote Failure Indication	Remote Loopback
	運用監視	—	—	—	—	Discovery,Link Monitoring,Remote Failure Indication

表 1：ShowNet EOAM Challenge(Y.1731、Link OAM)

ShowNet における CFM の取り組みは 2009 年に遡る。これまでの取り組み概要を表 2 に示す。

	区分	2009	2010	2011	2012	2013
CFM	相互接続	CC,LB,LT	CC,LB,LT	LT（監視ツールから各装置）	LT（各装置間）	LT（各装置間）
	運用監視	ー	CC専用GUI	CC,LB専用GUI	CC,LB専用GUI	CC,LB専用GUI
	他プロトコルとの併用	ー	STP上での利	STP上での利用	VPLS上での利用	OpenFlow上での利用

表 2：ShowNet CFM Challenge

CFM の取り組みとして掲げた 2009 年のテーマは『ShowNet における EOAM 相互接続元年の年』。コントリビュータは全 11 社（計 14 機種）が参加し、主に CC、LB、LT の相互接続性の確認を行った。ネットワークトポロジーはコの字型（ループ構成ではない）とし、1VLAN(EOAM 専用)・1 つの MD・1 階層の MD レベル、CCM の送信間隔は 1 分とした。

2010 年のテーマは『相互接続中心のチャレンジ』。コントリビュータは全 12 社（計 22 機種）で、3 つのチャレンジに取り組んだ。1 つ目の取り組みは、2009 年に引き続き CC、LB、LT の相互接続性の確認とした。構成を MD レベル 3 階層とし、その上で各機能の動作を確認した。2 つ目は、運用に即した利用へのチャレンジ。CC を活用し ShowNet ネットワーク全体の死活監視を行った。死活監視の見える化は、Interop 専用にコントリビューション頂いたツールで実施した。3 つ目は、L2 冗長プロトコルとの親和性の確認。CC と L2 冗長プロトコル(STP)の親和性の確認を数台のスイッチ上で実施することとした。ネットワークトポロジーはコの字型（ループ構成ではない）とし、相互接続用とネットワーク監視用に EOAM 専用 VLAN を 1 つずつ構築。CCM の送信間隔は 1 秒とした。

2011 年のテーマは『構築・運用ツールの活用にチャレンジ』。コントリビュータは全 12 社（計 21 機種）で、2 つのチャレンジに取り組んだ。1 つ目の取り組みは、主に LT の相互接続性の確認とした。構成を MD レベル 2 階層とし、その上で監視装置から各装置との相互接続性を確認した。（各装置間の確認は時間の兼ね合いから実施できず）2 つ目は、2010 年に引き続き、運用に即した利用へのチャレンジとした。MD レベル 2 階層の構成で、ShowNet 全体で CC による死活監視を実施（CFM 未対応機器は脇に CPE を配置）。並行して EOAM 監視装置より LB の死活監視を定期的に行った。監視の見える化は Interop 専用にコントリビューション頂いた 2 つの異なるツールで各々実施した。ネットワークトポ

ロジーはループ構成とし、STP を併用。L3 ShowNet ネットワーク構成に近似の構成とした。CCM の送信間隔は 1 秒とした。

2012 年のテーマは『構築・運用ツールとしての活用と Link Trace の相互接続』。コントリビュータは全 10 社（計 16 機種）で、2 つのチャレンジに取り組んだ。1 つ目の取り組みは、2011 年に続き、主に LT の相互接続性の確認とした。構成を MD レベル 1 階層とし、その上で各装置間の相互接続性を確認した。2 つ目は、前年に続き運用に即した利用へのチャレンジとした。VPLS 上での利用も積極的に行った。ネットワークトポロジーはコの字型(ループ構成ではない)とし、1VLAN (EOAM 専用)・1 ドメイン・1 階層のドメインレベルで構成。CCM の送信間隔は 1 秒とした。

2013 年のテーマは前年に続き『構築・運用ツールとしての活用と Link Trace の相互接続』。コントリビュータは全 7 社（計 10 機種）で、2 つのチャレンジに取り組んだ。1 つ目の取り組みは、前年に続き主に LT の相互接続性の確認とした。構成を MD レベル 1 階層とし、その上で各装置間の相互接続性を確認した。2 つ目は、前年に続き運用に即した利用へのチャレンジとした。OpenFlow 上での利用も積極的に行った。ネットワークトポロジーはコの字型(ループ構成ではない)とし、1VLAN (EOAM 専用)・1 ドメイン・1 階層のドメインレベルで構成。CCM の送信間隔は 1 秒とした。

4.4 CFM 相互接続の現状と課題

本節では、ShowNet でこれまでの実績を元に、CFM の相互接続の現状と課題について述べる。（本節で述べる内容は、あくまで ShowNet において確認した結果であり内容を保証するものではないことを前提としたい。）まず、CFM の各機能の相互接続の現状について述べる。CC は、これまでの相互接続性確認において特に課題は見受けられず、実運用ツールとしての利用が可能な状況と思われる。ただ、LB と LT は、一部の機器の実装において課題と思われる点が幾つか確認されている。そこで、LB と LT の実装課題と思われる点について、それぞれ例をあげて説明する。まず、LB の実装課題と思われる点について 1 件の例をあげる。動作概要を図 3 に示す。

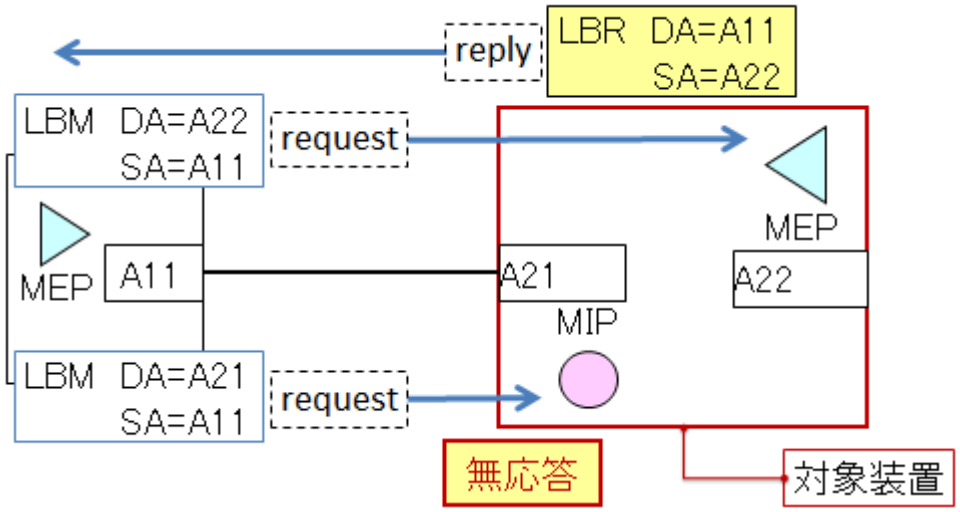


図 3：LB 実装課題例

ここでの課題は、対象装置に MEP と MIP が同じ VLAN(domain)に存在する場合に、対象装置の MIP の LB 応答がないことである。そのため、本実装の機器が中継機器の場合、LB による疎通性確認ができないということになる。

次に LT の実装課題と思われる点について 2 件の例をあげる。1 件目の動作概要を図 4 に示す。

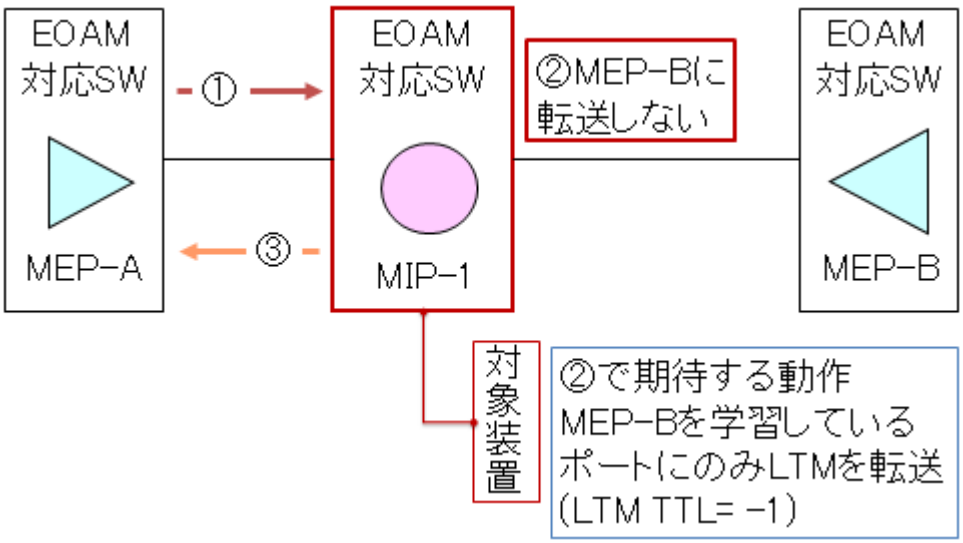


図 4：LT 実装課題例（1 件目）

図 4 の LT の動作の流れを①から順に説明する。①MEP-A から MEP-B へ LTM を送信する。LTM TTL=N をセットして規格に定義された Multicast DA で送信する。②対象装置の MIP-1 が LTM を

MEP-B に転送しない。③MIP-1 は MEP-A に LTR を応答する。MEP-A は MIP-1 の LTR を受信する。となる。

ここでの課題は、対象装置に MIP が設定されている場合に、MEP-A の LTM を転送しないことである。そのため、本実装の機器が中継機器の場合、LT による経路状態確認ができないということになる。

次に 2 件目の動作概要を図 5-1 に示す。

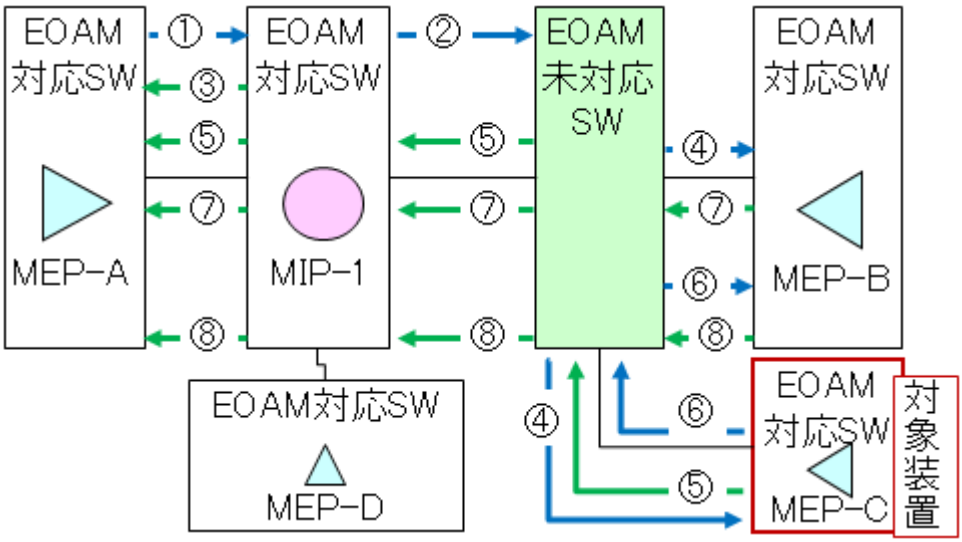


図 5-1：LT 実装課題例（2 件目）

図 5-1 の LT の動作の流れを①から順に説明する。①MEP-A から MEP-B へ LTM を送信する。②MIP-1 は MEP-B を学習しているポートにのみ LTM を転送する。③MIP-1 は MEP-A に LTR を応答する（①に対する応答)。 ④EOAM 未対応スイッチは LTM が Multicast DA であるためフラッディングする。⑤対象装置の MEP-C が自分宛ではない LTM に LTR を応答する(④に対する応答)。⑥対象装置の MEP-C が MEP-B に LTM を送信する。⑦MEP-B は MEP-A に LTR を応答する(④に対する応答)。⑧MEP-B は MEP-C から受信した LTM(SA MAC = MEP-A)の応答として MEP-A に LTR を応答する(⑥に対する応答)。となる。

ここでの課題は、対象装置に MEP が設定されていて自分宛以外の LTM を受信する場合に、MEP-A に MEP-B から 2 つの LTR が応答されることである。そのため、本実装の機器が自分宛以外の LTM を受信する構成の場合、LT による経路状態確認ができないということになる。

TTL の減算状態を見ると、本課題がより分かり易い。図 5-2 に TTL 減算の流れを示す。MEP-A の LTM TTL が 64 の場合、MIP-1 から TTL63 の LTR が、MEP-B から TTL62 の LTR が MEP-A に対し応答される。MEP-B までの経路状態確認はここで終わるべきである。ただ、本来不要な LTR として、MEP-C

から TTL62 の LTR が、MEP-B から TTL61 の LTR が、同じタイミングで MEP-A に対し応答されるため、実際の経路状態とは異なる LTR が MEP-A で見えてしまうのである。

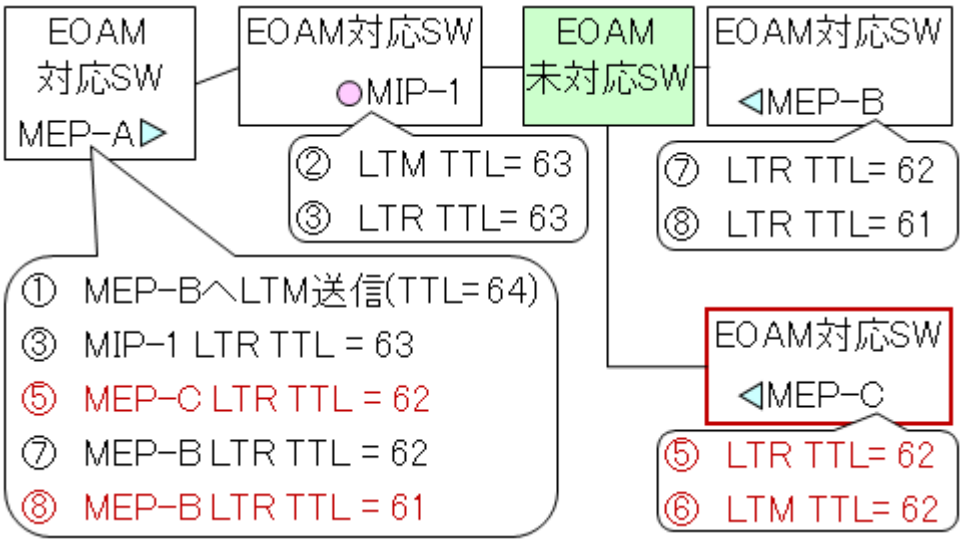


図 5-2：TTL 減算の流れ

4.5 CFM のこれから

本節では、これまでの ShowNet における CFM の取り組みを受けて、これからの CFM 相互接続で取り組むべきと考えていることを述べる。その上で、これからのイーサネットレイヤの運用監視をよくするために何が必要なのか、どのような解決方法が考えられるのか考察したい。まず現状確認されている LB と LT の幾つかの実装課題については、各々の当該機器メーカにフィードバックを行っており、その感触から、これらの課題が解決される可能性は高いと考えている。仮にこれらの課題が解決された場合、これまで主に CC による死活監視を中心に行ってきたイーサネットレイヤのネットワーク監視は、CFM の 3 つの機能をフルに活用することで、ICMP による IP レイヤのネットワーク監視とほぼ同等のことができるようになるため、これまで以上に Ethernet OAM の実環境利用を検討するユーザは増えると思われる。このようなユーザが CFM 相互接続の知見を得られる場所として、今後も ShowNet における CFM の取り組みは続けていくべきと考えている。

次に、これからのイーサネットレイヤの運用監視をよくするために何が必要なのか、どのような解決方法が考えられるのか考察したい。まず「イーサネットレイヤを運用監視する」ためのプロトコルとしては、現状の Ethernet OAM は必要十分な機能を有すると考えている。なぜなら物理レイヤを監視する Link OAM や VLAN 単位に監視する CFM や Y.1731 など、必要に応じた機能が定義されているからである。但し、プロトコル利用の仕方がまだまだ浸透していないという課題もある。次に「運用監視するためのツール」であるが、ここが現状で最も課題と考える点である。なぜなら市販品のある IP レイヤの

監視ツールとは異なり、特注品を選択する以外に選択肢がないからである。特注品となる理由の一つに、Ethernet OAM プロトコル自体の難しさもあると考えている。なぜなら難しいが故にツール化する難易度が上がり、モチベーションが下がるからである。そのための解決策としては、遠回りかもしれないが時間をかけて Ethernet OAM の利用価値が理解されるような取り組み（Ethernet OAM プロトコルの利用シーンの公開や、トレーニング資料の作成・配布など）を通して、少しずつ浸透させていくことだと考えている。

2. おわりに

本論文では、これまでの EOAM／Connectivity Fault Management(CFM)で取り組んだ結果を元に、CFM の相互接続の現状と課題を述べ、その上で、これからのイーサネットレイヤの運用監視をよくするために何が必要なのか、どのような解決方法が考えられるのか考察した。ShowNet では引き続き Ethernet OAM の利活用に取り組むことで、これからのイーサネットレイヤ品質向上の一端を担えればと考えている。

文　献

[8] http://www.ieee802.org/21/doctree/2006_Meeting_Docs/2006-11_meeting_docs/802.3ah-2004.pdf
[9] http://standards.ieee.org/getieee802/download/802.1ag-2007.pdf
[10] https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-Y.1731-201007-S!Amd1!PDF-E&type=items
[11] http://www.ttc.or.jp/jp/document_list/pdf/j/STD/JT-Y1731v1.pdf
[12] http://metroethernetforum.org/Assets/Technical_Specifications/PDF/MEF16.pdf

5 IPv4/IPv6 共存技術の現状と SHOWNET での相互接続検証結果について

IPv4/IPv6 共存技術とは、IPv6 only となったバックボーン上を介して、顧客ネットワーク側に IPv4/IPv6 のデュアルスタック環境を届ける技術の総称である。本章では、Interop Tokyo 2013 ShowNet で実施した IPv4/IPv6 共存技術の相互接続検証と、サービス利用の結果に関して報告を行う。本章の内容は 2013 年 9 月 6 日に開催された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（IA 研究会）にて発表した論文に修正を加えたものを掲載する。

IPv4/IPv6 共存技術の現状と ShowNet での相互接続検証結果について

大久保 修一† 櫨山 寛章†† 真野 桐郎††† 渡邊 茂††††

† さくらインターネット株式会社 研究所 〒160-0023 東京都新宿区西新宿 4-33-4

†† 奈良先端科学技術大学院大学

††† A10 ネットワークス株式会社

E-mail: †ohkubo@sakura.ad.jp,††hiroa-ha@is.naist.jp,
†††kmano@a10networks.com,††††shigeru@interop-tokyo.net

あらまし 現在 IETF で標準化が進められている各種 IPv4/IPv6 共存技術の適用領域の考察と、Interop2013 ShowNet にて実施した相互接続検証の結果について報告する。

キーワード IPv4/IPv6 共存技術,相互接続

Current status of IPv4/IPv6 co-existent technologies and interoperability test in ShowNet

Shuichi Ohkubo† Hiroaki Hazeyama†† Kiriro Mano††† Shigeru Watanabe††††

† Research Center, SAKURA Internet Inc. 4-33-4 Nishi shinjuku, Shinjuku-ku, Tokyo, 160-0023 Japan

†† Nara Institute of Science and Technology

††† A10 Networks, K.K.

E-mail: †ohkubo@sakura.ad.jp, ††hiroa-ha@is.naist.jp,
†††kmano@a10networks.com,††††shigeru@interop-tokyo.net

Abstract Applicability of IPv4/IPv6 co-existent technologies and report the result of interoperability test in Interop2013 ShowNet.

Keyword IPv4/IPv6 co-existent technology,interoperability

5.1 はじめに

今後インターネットの IPv6 普及が進むにつれ、キャリア等が新規に構築するバックボーンは IPv6 を前提とした設計になると考えられる。一方、インターネット上の IPv4 ユーザは今後数十年に渡って残り続けると予測されており、キャリアはユーザに対して引き続き IPv4 サービスの提供を継続する必要がある。

そのような IPv6 を前提としたネットワーク環境で、IPv4 サービスの提供を可能とする各種 IPv4/IPv6 共存技術の標準化が IETF にて進められている。本稿では、各方式の特徴と適用領域について考察し、Interop2013 ShowNet にて行った相互接続検証の概要と結果について報告する。

5.2 IPv4/IPv6 共存技術の必要性

2011 年 4 月、APNIC(アジア太平洋地域のレジストリ)において IPv4 アドレスの在庫が枯渇し、それ以降 World IPv6 Day や World IPv6 Launch といったイベントを経て、国内の組織においても IPv6 対応が進んでいる。これらの商用サービスで運用されているネットワークは、IPv4、IPv6 のデュアルスタック構成となっているケースが多い。

しかし、シングルスタックに比べデュアルスタック構成は IPv4 と IPv6 二面のバックボーン運用が必要となり、ロギングやアクセス制御、ルーティング設定やトポロジ管理といった点で、運用コストが倍増してしまうことが課題となっている。

そこで、バックボーンを IPv6 シングルスタックで運用しながらも IPv4 サービスの提供を可能とする IPv4/IPv6 共存技術を導入することで、運用コストの削減が期待できる。

5.3 各種 IPv4/IPv6 共存技術の特徴と適用領域

現在標準化が進められている(もしくは RFC 化が完了している)主要な IPv4/IPv6 共存技術を表 1 に示す。

表 1 主要な IPv4/IPv6 共存技術

名称	NAT の場所	パケットフォーマット
DS-Lite [1]	キャリア側	カプセル化
464XLAT [2]	キャリア側	トランスレーション
MAP-E [3]	ユーザ側	カプセル化
SA46T [4]	N/A	カプセル化
NAT64/DNS64[5][6]	キャリア側	トランスレーション

これらは優劣を付けられるものではなく、それぞれ異なった特徴を持つため、導入する事業者の特性にマッチしたものを選択することが望ましい。各共存技術は、大きく以下の 2 種類に分類することができる。

A) NAT をキャリア側で行うもの
B) NAT をユーザ側で行うもの

A)については、膨大な NAT セッションをキャリア側で運用しなければならないため、大規模な NAT 処理が可能な CGN(Carrier Grade Network Address Translation)機器の導入が必要となる。また、セッションログを保存するための大容量のストレージが必要となるため、スケーラビリティに乏しいアーキテクチャである。その反面、通信不良や不正通信などのインシデントが発生した際の追跡が容易であることや、ユーザに割り当てるポート数を柔軟に制御することで IPv4 アドレス資源を有効活用しやすいといったメリットが存在する。したがって、中小規模の ISP 等での導入に向いているといえる。

一方 B)は、アドレスマッピングルールをあらかじめ配布しておくことで、ユーザ自身が割り当てられた IPv6 アドレスから使用可能な IPv4 グローバルアドレスとポート番号の範囲を導き、ユーザ側に設置される CPE にてその範囲に NAT を行う。したがって、キャリア側設備は単純にアドレスマッピングを行うシンプルな機器でルーティング可能なため、スケーラビリティを確保しやすい。ユーザが使用するポート番号の範囲のみを管理すれば良く、膨大な NAT セッションを扱う CGN も、ログを保存するストレージも必要ない。ただし、あらかじめアドレスマッチングルールを事前に決め打ちしな

ければならず、後で変更することも困難なため、IPv4、IPv6 共に大きなアドレス空間を割り当てることが可能な大規模 ISP での導入に適していると考えられる。

5.4 INTEROP2013 SHOWNET での実装

ShowNet では、過去にも各種 IPv4/IPv6 共存技術の提供を行なってきた。しかしながら、アクセスコーナーでの提供を中心とした、あくまでもトライアル的な位置づけであった。Interop2013 では、これまで積み上げられたノウハウをもとにした商用品質での提供を目標とし、出展社向けに提供する回線への IPv4 over IPv6 技術の導入を行った。IPv4 over IPv6 技術で利用した Interop2013 ShowNet 参加企業からの提供機器と用途を表 2 に示す。

表 2　Interop2013 ShowNet にて使用した機器

提供社名	製品名	用途
シスコシステムズ	ASR9000	MAP-E BR
インターネットイニシアティブ	SEIL/x86	MAP-E CE、DS-Lite B4
古河電工	FX5000	MAP-E BR
A10 ネットワークス	AX5630	DS-Lite AFTR、464XLAT PLAT
セイコーソリューションズ	iX-3740	464XLAT PLAT、NAT64
ジュニパーネットワークス	SRX5600	DS-Lite AFTR、464XLAT PLAT
NEC アクセステクニカ	Aterm	DS-Lite B4、464XLAT CLAT
Infoblox	Trinzic820/vNIOS	DNS64、DHCP、DHCPv6
富士通	SA46T/SA46T-PR/SA46T-PT	

464XLAT、DS-Lite については、キャリア側機器(PLAT,AFTR)とユーザ側機器(CLAT,B4)に関して複数のメーカから提供を受けたため、全ての組み合わせを実現できるよう実装を行った。

MAP-E については、ASR9000 と FX5000 の 2 台の BR を IP Anycast を用いた負荷分散、冗長構成とした。MAP-E ではキャリア側設備をステートレスにすることができるため、一般的なルーティングの技術を用いて負荷分散、冗長構成を実現可能である。なお、MAP-E パラメータは表 3 の通りとした。

表 3 MAP-E パラメータ

パラメータ	値
Rule IPv6 prefix	2001:3e8:6000::/56
Rule IPv4 prefix	130.129.255.92/30
End-user IPv6 prefix	2001:3e8:6000:00xx::/64
EA bits	8bit(64-56)
Port-Set ID	6bit(8-2)
PSID offset	4
ports/user	960ports
BR Address	2001:3e8:0:2900::1/128

5.5 相互接続検証結果

相互接続性の検証は、参加企業から提供された各機器のあらゆる組み合わせパターンにおいて、下記項目の確認を行った。

- MTU の正常性
  設定した MTU サイズまでのパケットで通信が行えるかを確認。
- TCP や UDP パケットの疎通

TCP についてはSYN パケット中の MSS 値が、正しく CE で調整されているかについても確認。

- ICMP パケットの疎通
  Echo Request/Reply、Time Exceeded、Fragment Needed、Host Unreachable の 4 種別にて確認。
- フラグメントパケットの疎通
  分割された IP パケットをユーザ側機器、キャリア側機器、共に正しくハンドリングできるかを確認。
- 主要な Web サイトの閲覧
  ユーザが利用すると考えられる主要な Web サイトをブラウザで閲覧し、遅延やコンテンツの欠落（歯抜け）が発生しないかを確認。
- P2P 系アプリケーションの動作
  Skype などのクライアント端末間で直接通信するアプリケーションは、CE の NAT の実装に強く依存するため、音声やビデオが疎通するかを確認。
- 異ドラフトによる実装(MAP-E のみ)
  後ほど詳細を説明するが、MAP-E は参照ドラフトによって 2 種類の実装があるため両方で確認。

これらの結果、以下に示すいくつかの設定調整が必要であった。しかし、本会期に入ってからのトラブルは全くなく、ユーザからの不具合報告もなかった。結果として、相互接続性の問題はなく、非常に良好に動作したといえる。

A) MTU の問題

一般的な PPPoE を使用するようなアクセス回線も含め、トンネルを使用する方式では常につきまとう問題ではあるが、今回実装した IPv4 over IPv6 トンネリングにおいても同様の問題が発生した。

まず、IPv4 インターネットにおける Path MTU Discovery Blackhole による通信不良を避けるため、CPE にて IPv4/TCP Syn パケットの MSS 値の調整を行うような設定を行った。具体的な値は、以下の式によって導かれる。

TCP MSS 値 = IPv6 MTU - 40(IPv6 ヘッダ長) - 20(IPv4 ヘッダ長) - 20(TCP ヘッダ長)

なお、今回使用した MAP-E 機器の中には、IPv6 ネットワークの Path MTU Discovery Blackhole も避けるため、IPv6 MTU が 1280Bytes 固定設定となっている機器が存在した。そのため、MAP-E ドメインについてはこの値で統一した。

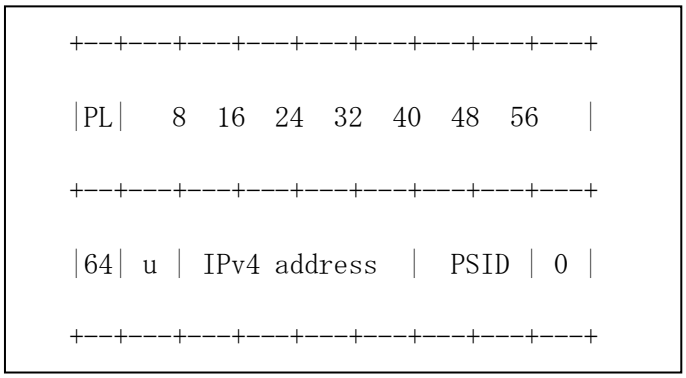
B) フラグメントの扱いが機器によって異なる

上記 MTU の問題とも関連するが、サイズの大きい IPv4 パケットを IPv6 でカプセル化する際に、IPv6 フラグメントを用いて転送する機器が存在した。その際、対向機器側でフラグメントされた ICMP パケットをドロップするような実装があり、ICMP パケットのみ通過できないようなケースが存在した。

5.6 MAP-E 最新ドラフトでの相互接続検証

MAP-E は、本稿執筆時点で IETF draft となっているが、改版の度に数回の仕様改定が行われており、2013 年 8 月時点でのリビジョンは -07 である。リビジョン間で最も大きな変更点として、draft-ietf-softwire-map -03 から-04 へのアップデートにおいて、カプセル化した IPv6 パケットに付与する IPv6 アドレスのインターフェイス ID のフォーマットが変更になった。変更点を図 1 に示す。

draft-ietf-softwire-map-03 における Interface ID



draft-ietf-softwire-map-04 における Interface ID

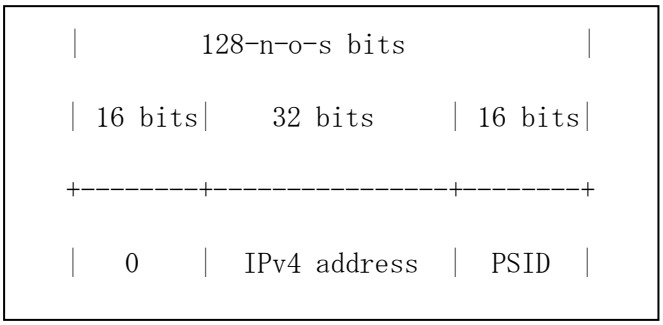


図 1: MAP-E インターフェイス ID の差異

多くのメーカは古いフォーマット(-03)で既に実装を進めている状況であったため、本会期中に提供する仕様は-03 のフォーマットに統一した。

しかし、新しいフォーマット(-04)をサポートしている機器もあったため、対応機器のメーカに協力を仰ぎ、ホットステージ期間中に相互接続検証を行った。あるメーカにおいて、ICMP パケットの新フォーマットによるカプセル化が未対応であったが、TCP/UDP パケットについては、問題なく接続できた。図 2 に示すパケットダンプにて、新フォーマットでのアドレスマッピング(下線部)が行われていることがわかる。

```
19:12:49.571434 2001:3e8:6000:c0:0:8281:ff5f:0 >2001:3e8:0:2900::1: IP 130.129.255.95.45089

> 133.242.53.30.80: S 1937843878:1937843878(0) win 14600 <mss 1420,sackOK,timestamp 344073592

0,nop,wscale 6>
```

図 2 : MAP-E 新フォーマットにおけるパケット

5.7 まとめ

本稿では、IPv4/IPv6 共存技術の特徴とその適用領域を考察し、Interop2013 ShowNet での相互接続検証結果について述べた。相互接続を行うにあたっては設定調整が必要な

部分が確認され注意を要するが、Interop 本会期中、トラブルなく提供を行えたことから、各種 IPv4/IPv6 共存技術は既に実用化段階に入っているといえる。

謝 辞

本稿執筆においてご協力いただいた、Interop 2013 ShowNet コントリビュータ各社、Interop 2013 ShowNet NOC チーム、STM / CTM、Interop 2013 ShowNet 事務局、他、Interop 2013 ShowNet の運営に関わって頂いた皆様に厚く感謝する。

文 献

[1] A. Durand and R. Droms and J. Woodyatt and Y. Lee, Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion, RFC 6333 (Proposed Standard), Internet Engineering Task Force, Aug., 2011, http://tools.ietf.org/html/rfc6333

[2] M. Mawatari and M. Kawashima and C. Byrne, 464XLAT: Combination of Stateful and Stateless Translation, RFC 6877 (Informational), Internet Engineering Task Force, Apr., 2013 http://tools.ietf.org/html/rfc6877

[3] O. Troan, W. Dec, X. Li, C. Bao, S. Matsushima, T. Murakami, Mapping of Address and Port with Encapsulation (MAP), IETF Softwire Working Group Internet Draft, May, 2013, (version 07) http://tools.ietf.org/html/draft-ietf-softwire-map

[4] N.Matsuhira, Stateless Automatic IPv4 over IPv6 Encapsulation / Decapsulation Technology: Specification, Internet Draft, July, 2013, http://tools.ietf.org/html/draft-matsuhira-sa46t-spec

[5] M. Bagnulo and P. Matthews and I. van Beijnum , Stateful NAT64: Network Address and Protocol Translation, RFC 6146 (Proposed Standard), Apr., 2011, http://tools.ietf.org/html/rfc6146

[6] M. Bagnulo and A. Sullivan and P. Matthews and I. van Beijnum , DNS64: DNS Extensions for Network Address Translation, RFC 6147 (Proposed Standard), Apr., 2011, http://tools.ietf.org/html/rfc6147

6 INTEROP 2014 SHOWNET に向けて

さて，本技術報告書もこの章で終わりとなりますが，ここで Interop 2014 ShowNet に関して簡単な紹介をさせていただきます.

Interop 2013 ShowNet は，20 回記念ということもあり（？），かなりいろいろなトライアルを実施しました. かなりたくさんの課題があるため来場者の方にはフォーカスがぶれて見えたかもしれませんが, それだけインターネットを構築する技術は多種多様で，常に進化し続けていると言えます．また，Interop 2013 ShowNet では現在の NOC チーム，STM，CTM の体制で何とか期間内に構築出来る最大数の構成でした．そのため，十分な検証が出来たとは言えず，課題が見えた相互接続検証や運用もさまざま浮かび上がってきました．Interop 2013 ShowNet NOC チームは反省会の際に「はたして，インターネットは 10 年後も生き残っているのだろうか？現状のままだと生き残っていないんじゃないか？または，とても快適とはいえないネットワークライフになっているんじゃないか？」という思いを持ちました.

そこで Interop 2014 ShowNet では，まず Interop 2013 ShowNet NOC チームメンバー，NOC アドバイザリーメンバー，Interop 2014 ShowNet NOC チームメンバーに「<インターネット未来予想図 Worst Case（現実） と Best Case （希望）>」というタイトルにてアンケートをとりました.

アンケートの内容ですが，

1) 研究や運用や一般ユーザとして，普段かかわっているネットワークやネットワーク構築技術に関して，3 年後はこうなっているという予想図を下記の 3 ケースで大胆に予想してください.

- ワーストケース（こうなっていたらいやだという構成）
- 現実路線（ビジネスとしての硬い構成や方向性）
- ベストケース（自分が欲しいと思っている技術革新・標準規格の構成）

2）2015 年 Apricot/APNIC 日本開催や，IETF 94th Yokohama までに普及しておいて欲しい標準規格やインターネットサービス，解決しておいて欲しい諸問題を教えてください.

3）2020 年東京オリンピックまでに実現していて欲しい, 標準技術やインターネットサービスを教えてください.

4）30th Interop Tokyo（2023 年）に，「こうなってたらいいなぁ」という，漠然とした夢を教えてください.

アンケート結果の詳しい内容は 2014 年版の「ゲンバブログ」に譲りますが，悲観的な内容もあれば，新しい技術を取り込んだかっとんだ未来予測など，さまざまな「現実の課題」と「夢」が浮かび上がりました.

この未来予想図を踏まえ，Interop 2014 ShowNet NOC チームメンバーでは 3 年間かけてインターネットとインターネットを支える技術を再構築していくことをグランドテーマとして設定しました．Interop 2014 ShowNet のテーマは「Scratch & Rebuild the Internet：phase 1 - tough core, soft edge, for future apps -」です．まだ見ぬ未来のアプリケーションを支えるために, 何があっても壊れないタフなバックボーンコアネットワークをどのように構築し, さまざまなユーザやアプリケーションの要望や要求に柔軟にこたえるエッジネットワークをどのように構築すればよいか？，いい意味で「枯れた」技術のよい部分を継承し，「時代の要望を先取りして答えようとする」技術・標準規格を商用サービススペックまでこなれたものにどうやれば出来るか？，また，そのときの効率的な運用方法はどのようなものなのか？という課題に取り組んで，
ShowNet のネットワークデザインを進めています.

読者の皆様は，ぜひ，幕張の会場に足を運んでいただき，Interop 2014 ShowNet NOC チームのベスト・プラクティスをその目で確かめていただければ，幸いです.

文責

Interop 2014 ShowNet NOC チームメンバー

NOC ジェネラリスト

櫨山 寛章，関谷 勇司，宇多 仁