Network Service Provider
Internet eXchange Point
WIDE Project

高度情報インフラストラクチャの構築に関する研究

# NSPIXP WG 実証実験活動報告

2013 年研究報告書

2013 年 12 月 30 日

WIDE Project NSPIXP WG

# 1. はじめに

本研究では、商用インターネットを相互に接続する場合の問題点を明確にし、それを解決するための技術や手法の研究開発ならびに実証実験を行うことを目的とした実証実験を行う。また、近年成長し続ける動画系のインターネットトラフィックに対して、トラフィック輻輳を防ぎ、ユーザへの応答性を保つためのトラフィックエンジニアリングや、大規模災害等の障害にも対応できるための強固なインターネットバックボーンの形成に関する実証実験を行うことを目的とする。さらに、Software Defined Network (SDN) の IX への適用を検討することで、より柔軟な構成やトラフィック制御を行うことができる、次世代 IX のモデルに関する研究を行う。

具体的には、WIDE Project のサブプロジェクトである、 Network Service Provider Internet exchange Point (NSPIXP) プロジェクトにて行われている、DIX-IE ならびに NSPIXP-3 の運用を通じて、新技術の研究開発や実証実験を行った。これにより、インターネットがより信頼性を有した高度情報インフラストラクチャとして機能するために必要となる機能の検証や開発、ならびにその実証実験を行った。

NSPIXP プロジェクトは、1994 年の NSPIXP-1 運用開始、1996 年の NSPIXP-2 運用開始、1997 年の NSPIXP-3 運用開始を経て、東京に分散配置した DIX-IE、大阪に分散配置した NSPIXP-3、IPv6 に特化した NSPIXP-6 の運用を基盤とした実証実験を行ってきた。2008 年 6 月に NSPIXP-6 の運用を終了し、現在は DIX-IE、NSPIXP-3 におけるプロダクション品質の IPv6/IPv4 デュアルスタック運用に取り組んでいる。さらに 2012 年は、DIX-IE と NSPIXP-3 を連結した、広域 IX である NSPIXP-23 の運用も開始した。本年度は、さらに柔軟なトラフィック制御を可能とすべく、SDN を取り入れた SDN-IX の構築に関する検討と準備をすすめた。

本報告書では、第 2 節にて本年度の研究計画を再掲し、その研究計画に基づいた本年度の成果を第 3 節にて報告する。最後に、第 4 節にて本年度後半の活動方針について述べる。

## 2. 本年度の研究計画

NSPIXP Project では特に近年、高信頼性および高効率性を考慮した上での分散ネットワークアーキテクチャに着目し、トラフィック制御技術の実証および展開の検討と議論をおこなっている。現在、今後のトラフィック動向をふまえた IX の利用方法に関する議論と、次世代 IX に向けた取り組みである SDN-IX の実現手法に関する議論を行っている。さらに、東阪を接続した広域 IX である NSPIXP-23 では、トラフィックエンジニアリングに重心をおいた実証実験を行なっている。関東と関西にまたがった広域 IX を利用することで、より強固な、拠点単位の災害に対応することのできるインターネットバックボーンと IX のアーキテクチャに関する実証実験に取り組んでいる。

年次初頭の研究計画書にて示した通り、本年度の研究課題は以下の 5 項目である。

### (1) 拠点障害にも対応できる IX ならびにサービスアーキテクチャの研究

2011 年から 2012 年は、トラブルを極力低減し、万が一の障害発生時においても自動的に回復することのできるような IX アーキテクチャに関する設計と検証を進めてきた。この成果として、2012 年には、DIX-IE と NSPIXP-3 を相互接続する形での NSPIXP-23 を構築し、サービス開始を行った。これをふまえ、2013 年は、より障害に強い IX アーキテクチャとして、東京と大阪にまたがる広域 IX の構成のみならず、それを利用したサービスの展開手法についても、検討と検証を行うことを目指す。

NSPIXP-23 を実現するにあたっては、それぞれのネットワークをいきなり相互接続するのではなく、DIX-IE と NSPIXP3 をまたがる新たなネットワークを別途構築し、DIX-IE と NSPIXP3 の両拠点に 802.1Q tagged VLAN によって提供することを行った。これは、いきなり DIX-IE と NSPIXP3 のネットワークを相互接続した場合に、同じブロードキャストドメインに属してしまうため、東京と大阪の間にてどの程度の広域ブロードキャストパケットが発生してしまうか予想が困難であるため、まずは新規に別ネットワークを構築して、DIX-IE と NSPIXP3 を接続する形態にて開始した。この運用にあたっても、広域となるがゆえにその障害や品質低下をより即時性をもって検知する技術が必要となった

ため、EtherOAM を用いた IX の監視と品質保証に関する実験を行った。

2013 年はこの流れを受け、EtherOAM を活用した CC(Continuity Check)と LT(Link Trace)による広域かつ多段スイッチで連結された IX の監視手法の確立に取り組む。NSPIXP-23 は、その構成上多段スイッチ構成にて構築されているため、何か障害が発生した場合には、どの地点の回線、もしくは機器にて障害が発生したのかを突き止める必要がある。従来の手法であれば、スイッチのリンク状態の監視や、スイッチ間に ping 用の端末を設置することでの ping 監視を行うことで、障害箇所の特定を行なっていた。一方、EtherOAM を利用した場合には、CC や LT を用いることで、追加の機材や監視のためのネットワーク設定を必要とすることなく、多段スイッチ構成においても障害点を特定することが可能となる。そのため、広域 IX を構成する場合の多段構成での監視モデルの確立に関する研究を行う。

また、ISP が両拠点に接続点を有することによって、東京エリアの一部の拠点に災害や障害が発生したとしても、自動的に他の拠点のバックアップ接続ポイントを利用してトラフィック交換を継続して行うことを目指す。これは 2012 年から行なっている研究課題であり、いくつかの ISP に参加してもらうことで、検証実験を行なっている。また、これに加え 2013 年は、サービスを提供するサーバ環境を、仮想化技術を用いてまるごと高速マイグレーションさせるための研究も行う。図 1 にその概念図を示す。

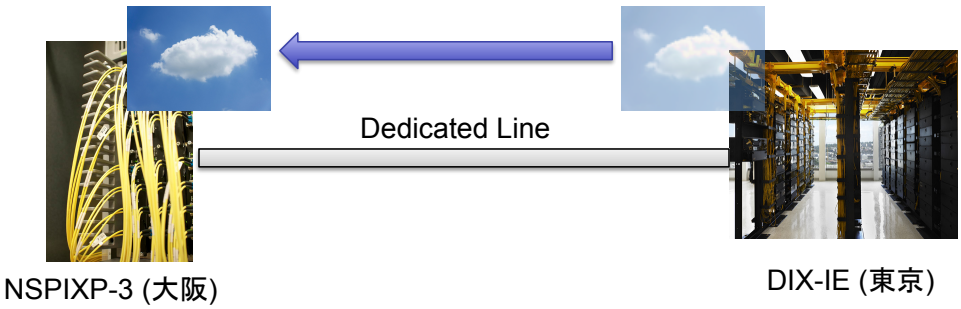



図 1：広域サービスマイグレーション

近年の仮想化技術では、一台の物理サーバの中に仮想サーバを複数起動し、多重化して利用する傾向にある。この際、サーバの仮想化を行うソフトウェア技術をハイパーバイザと呼ぶが、本研究では仮想サーバ単体ではなく、ハイパーバイザごとすべてマイグレーションすることを目的とした実証実験を行う。このためには、ハイパーバイザ技術自体の抽象化や仮想化、ストレージの移動技術といった要素技術が必要となる。2013 年はこれらの研究課題にも取り組み、

検証を行うことを目指す。

### (2) IPv4/IPv6 トラフィック成分分析に関する研究

IPv4 から IPv6 への移行が進むにつれて、IX における IPv6 での BGP peering も増加している。これは IPv4 アドレス枯渇にともなう IPv6 への移行が本格化してきたことを反映していると考えられる。

DIX-IE, NSPIXP-3 では、sflow を利用したトラフィック成分分析を 2011 年から開始した。NSPIXP-23 も同様に sflow によるトラフィック収集を行なっている。NSPIXP-3 では、すべての拠点、ならびに参加組織を対象として sflow データの収集を開始し、IPv6 のトラフィックに関しても、その成分分析を行なっている。DIX-IE は 2013 年 2 月時点において、KDDI 拠点、NTT-C 拠点、MIND 拠点において sflow によるデータ収集を行なっている。2013 年は、これをさらにすすめ、DIX-IE の前拠点におけるデータ集取を行い、より詳細な IPv4/IPv6 トラフィック成分分析を行うことを目指す。

### (3) 公共サービスならびにキャッシュアーキテクチャの展開に関する研究

ISP 同士のトラフィック交換が、パブリックピア接続からプライベートピア接続に移行するにあたって、公共的なサービスを提供するためのインフラが問題となっている。例えば、
Root DNS や ccTLD DNS、また逆引きゾーンを受け持つ APNIC 管理 DNS に代表されるような公共的なサービスである。これらの公共的な DNS サービスは、効率よく多組織に対して提供することが求められており、ある特定のユーザに対する利益ではなく、インターネットを利用するすべてのユーザの利益のために提供されているサービスである。したがって、これらサービスに接続性を提供するためには、プライベートピア接続よりパブリックピア接続の方が適している。今まで DIX-IE ならびに NSPIXP3 では、積極的にこのような公共インフラとしての性格をもつ組織の誘致を行なってきた。2012 年もその流れを引きつぎ、商用 IX とは性格の異なったサービスを提供できるよう、公共的なサービスの誘致を行なっていく。

また、近年 IX でのトラフィック交換からプライベートピアリングでのトラフィック交換が主流になるにつれて、海外トラフィックによる帯域の圧迫が問題

となっている。これは、海外動画サイトやダウンロードサイトから日本に流入するトラフィックが増大するにつれて、プライベートピアリングやトランジットの帯域のほとんどをこれらのトラフィックが占めるようになり、下位 ISP が上位 ISP に支払う課金が増大するという問題である。この問題に対応すべく、Google 等の大手コンテンツプロバイダーは独自に ISP に対してキャッシュ機能を有した装置を設置することを進めているが、Google が設置したキャッシュ装置は Google に関連するトラフィックのみしかキャッシュすることができず、全体としてのトラフィックを低減することはできない。そこで、IX においてキャッシュ装置を設置することで、IX 経由で海外からの流入トラフィックを低減することができるような運用モデルの研究を行う。これによって、下位 ISP に発生する大きな課金を防ぐことができ、かつ限られた海外回線の帯域を有効に利用することが可能となる。IX におけるキャッシュ有効利用モデルについて 2013 年も引き続き検討と検証を進める。

### (4) EtherOAM の導入

IX におけるサービス品質の向上と定常監視技術として、EtherOAM をどのように利用することができるか、その運用モデルを含めた検証を開始する。

IX アーキテクチャが広域になるにしたがって、その品質監視が従来の技術では難しくなってくる。そこで、EtherOAM を利用した品質管理と監視アーキテクチャを検討し、懸賞を行う。これは、研究課題(1)にも関連するものであり、より信頼性のおけるサービス展開や、効率的な迂回路の確立のために必要な要素技術となる。

EtherOAM 技術を利用した監視においては、MEP をどの範囲に設定するかが大きな問題となる。MEP を設定する地点によって、監視できる機器と監視できる技術(CC,LB,LT)が決定され、監視対象も自ずと決定される。したがって、EtherOAM 導入時には、MEP の綿密な設計を行うことが必要不可欠であり、MEP 設定を適切に行うことが運用モデルの定義につながる。そのため、IX に適した MEP 設定を形成することは、EtherOAM を利用した IX 監視モデルを定義することにつながる。

2012 年は、NSPIXP-23 への EtherOAM の導入検証を行った。この結果を用いて、2013 年は、DIX-IE 全体への EtherOAM 導入をすすめる。

### (5) Software Defined IX の設計検討

現在の IX は、Layer-2 もしくは Layer-3 において BGP peering を行い、経路を交換することでトラフィック交換を行うことが前提となっている。この際のトラフィック交換の粒度は、あくまで BGP で交換できるプレフィクスが単位となり、IPv4 の場合は通常 /24 と呼ばれる 256 個の IPv4 アドレスを単位とした制御となる。また、IPv6 の場合は通常/48 と呼ばれる、2^76 個の IPv6 アドレスを単位とした制御となる。

さらに、トラフィック制御に用いる指標は、あくまでも IP アドレスであり、それ以外の例えば URI であるとか、ポート番号であるとかいった情報を指標とした経路制御を行うことはできない。これは、インターネットを情報インフラストラクチャとして見た場合には、単純かつ冗長性を確保するために十分な仕組みであるが、より高度なトラフィック制御を行おうと思った場合には、機能が不足する。そのため、BGP を用いたトラフィック交換制御は限界があり、より増加し続ける動画トラフィックを高度に制御するためには、より柔軟なトラフィック制御手法が求められる。そこで SDN を用いた IX、すなわち Software Defined IX を構築し、様々な指標を使い、かつより詳細な粒度でのトラフィック交換制御ができる IX を構築することを目指す。2013 年は、この要求事項と要素技術の検討と検証を行う。

## 3. 研究成果

本節では、(1)～(5)の各研究項目に関する、研究成果を報告する。

### (1) 拠点障害にも対応できる IX ならびにサービスアーキテクチャの研究

2013 年では、前年度に引き続き、広帯域化によるトラフィック増加に対応するための、IX アーキテクチャの研究と運用を通じたその実証を行った。表 1 に 2013 年 12 月時点での、DIX-IE ならびに NSPIXP-3 の実証実験拠点を示す。

表 1：DIX-IE / NSPIXP-3 実証実験拠点一覧

| | |
|---|---|
| DIX-IE | KDDI 大手町拠点（WIDE） |
| | NTT 大手町拠点（NTT Communications） |
| | NF 西大井拠点（MIND） |
| | ComSpace-1 拠点（Vectant） |
| | @Tokyo 拠点（@Tokyo） |
| NSPIXP-3 | NTT 堂島拠点（WIDE） |

2013 年は、特に大きな構成変更を伴う作業は発生しなかった。2013 年 12 月次点での、DIX-IE ならびに NSPIXP-23 の構成図を図 1 に示す。

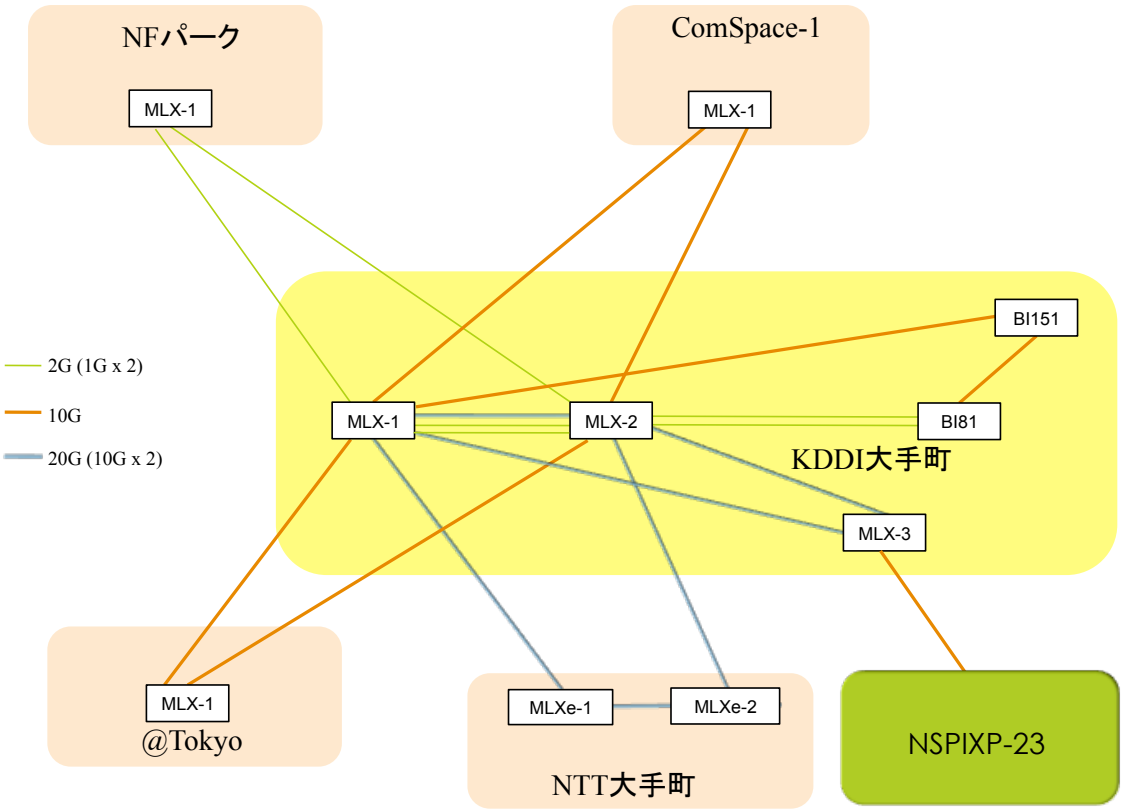
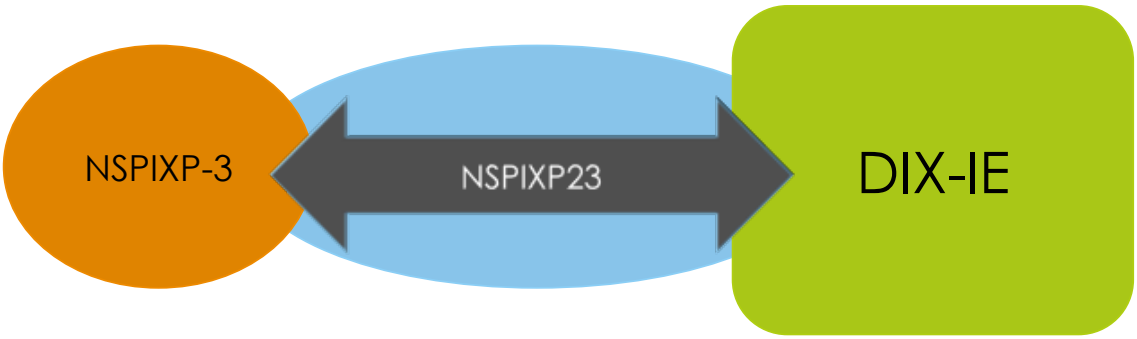



図 1：DIX-IE ならびに NSPIXP-23 構成図

また、NSPIXP-23 の内部構成を、図 2 に示す。




| | | |
|---|---|---|
| DIX-IE | : | VLAN 2 (202.249.2.0/24, 2001:200:0:fe00::/64) |
| NSPIXP3 | : | VLAN 3 (202.249.38.0/26, 2001:200:0:fe10::/64) |
| NSPIXP23 | : | VLAN 3000 (202.249.38.64/27, 2001:200:0:fe20::/64) |

**図 2：NSPIXP-23 の内部構成**

次に、2013 年 12 月時点における、DIX-IE ならびに NSPIXP-3 にて交換されたトラフィック総量の推移を図 3 に示す。

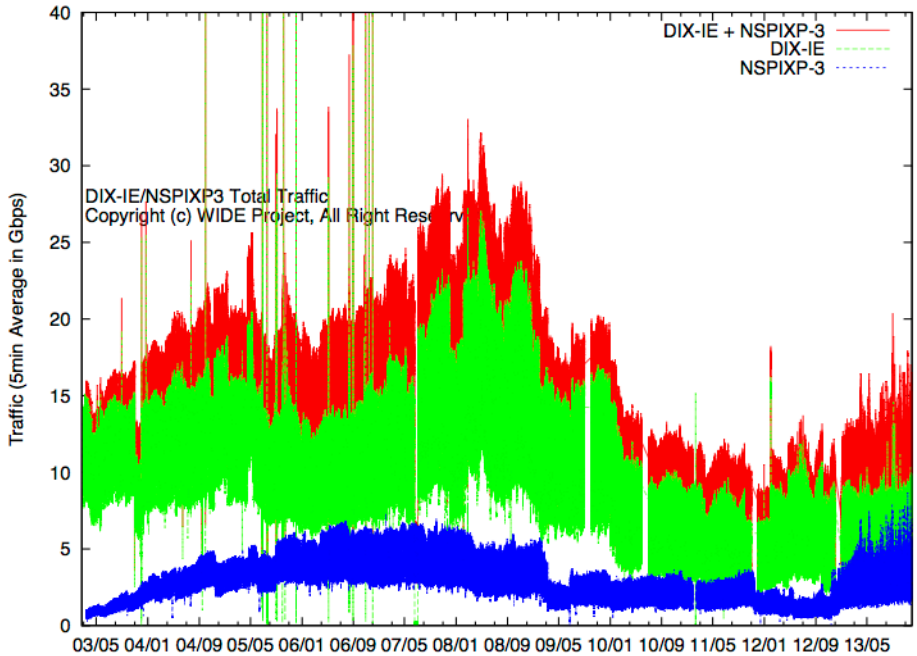



**図 3：DIX-IE ならびに NSPIXP-3 におけるトラフィック総量の推移**

本年度は、前年度から比較すると、トラフィック総量が少量であるが増加に

転じていることが見て取れる。この傾向は、前年度の報告書においても紹介した、総務省「我が国のインターネットにおけるトラヒックの集計・試算[1]」の 2013 年 5 月版においても同様の傾向が示されており、プライベートピアへの移行が一段落し、また IX へのトラフィック流入が多少なりとも増加している傾向にあるととれる。また、NSPIXP-23 によって、東西のトラフィック交換が活発に行われるようになったことも影響していると考えられる。トラフィック増加については、今年度後半において詳細に分析する。

また、2013 年は、特に明記すべき計画メンテナンスや障害は発生しなかった。アナウンスとして一点、緊急連絡受付電話番号の変更があった。新番号は 2013 年 7 月 1 日 0:00 より有効となった。

## (2) IPv4/IPv6 トラフィック成分分析に関する研究

2012 年度の報告書において報告した通り、DIX-IE ならびに NSPIXP-3 ともに、sflow を用いたトラフィック成分分析を開始している。DIX-IE においては全拠点のコアスイッチと参加者収容収容スイッチの全インタフェースにおいて sflow を有効にしている。NSPIXP-3 においても、NTT 堂島拠点のスイッチにおいて、sflow を有効にしている。すなわち、DIX-IE ならびに NSPIXP-3 上にて交換されるトラフィックすべてを成分分析することが可能となっている。

sflow を用いたトラフィック成分分析の解析例として、ある拠点の一部の収容ポートにおける、2013 年 9 月 20 日から 2013 年 9 月 30 日の間の IPv4/IPv6 トラフィック動向を図 4 に示す。グラフにおいて緑色にて示される部分が IPv6 トラフィックの帯域である。総トラフィック量の約 1 割を IPv6 トラフィックが占めており、IPv6 への以降が確実に進行していることを示す証拠となる。

[1] http://www.soumu.go.jp/menu_news/s-news/01kiban04_02000061.html

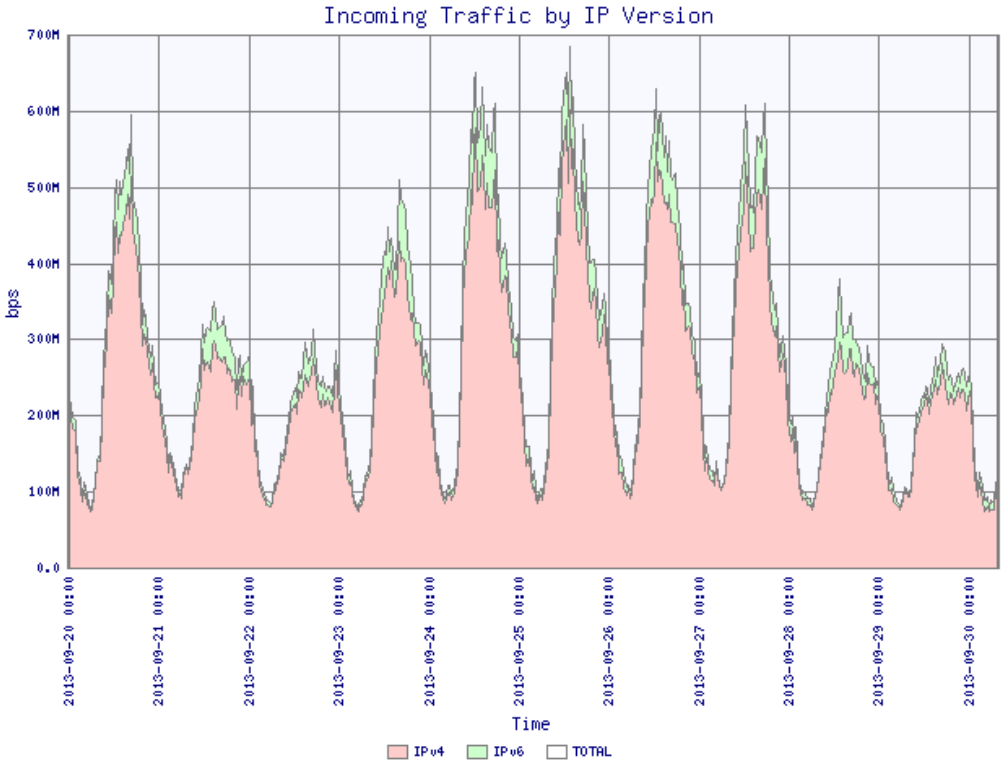



図 4： ある拠点における IPv4/IPv6 トラフィック割合

### (3) 公共サービスならびにキャッシュアーキテクチャの展開に関する研究

DIX-IE ならびに NSPIXP-3 においては、いくつかの公共的なサービスを提供する実証実験を行っている。IX にて公共的なサービスを提供することは、日本ならびに世界のインターネットの健全な運営のためにこれからの IX が担う必要な役割であると考える。

主要なトラフィックがプライベートピアによるトラフィック交換に移行しつつあるが、やはり公共的なサービスを提供するのは、IX の役目であると考え、DIX-IE ならびに NSPIXP-3 では、積極的に公共的なサービス展開を行ってきた。その例として、 Root DNS や ccTLD DNS、ならびに 6to4 や Teredo といったサービスがあげられる。

例として、

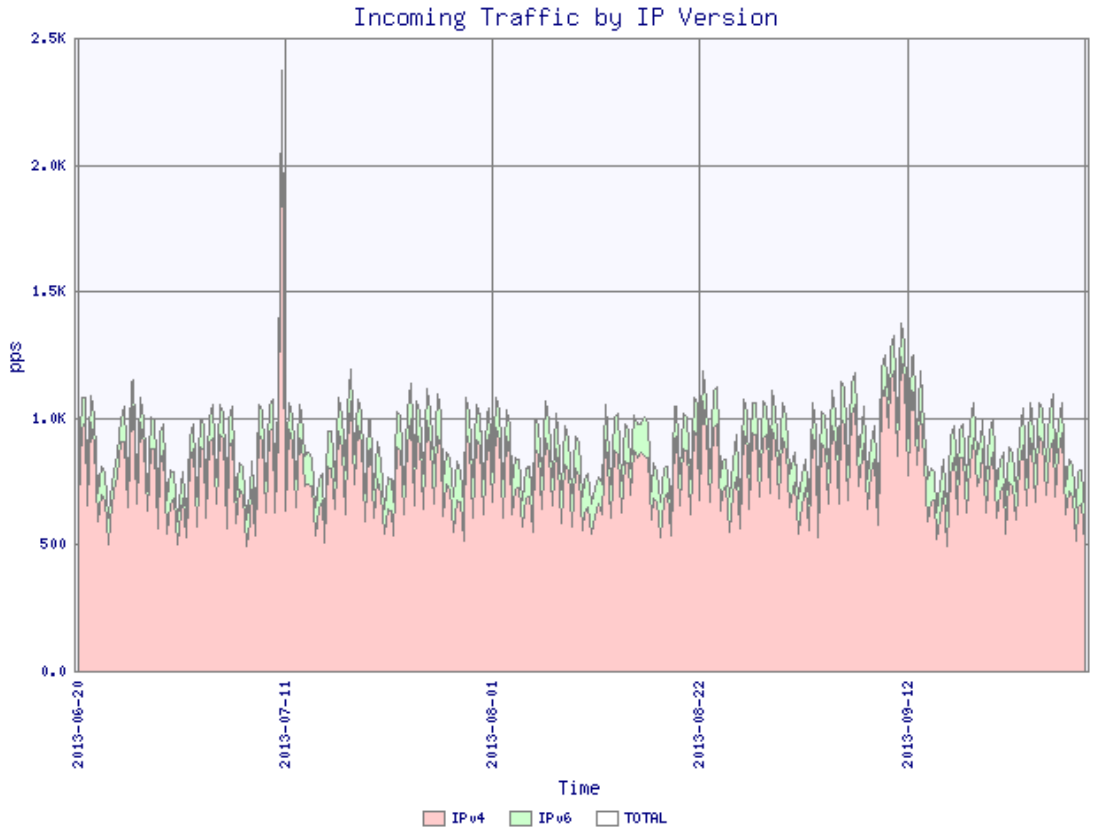



図 5 に DIX-IE における Root DNS のトラフィック動向を示す。DIX-IE から見た、Root DNS サーバからのトラフィックを pps にて示したグラフである。Root DNS はインターネットにおける重要なインフラのひとつであり、サービスを健全に提供することが命題となっている。このサービスに貢献するために、DIX-IE ではトラフィック閾値による監視を提供し、一定時間にある一定以上のトラフィックが発生した場合には、管理者にアラートを送付する仕組みを提供した。この仕組により、例えば図中の 7 月 10 前後に発生した Root DNS サーバへの大量問い合わせに関しても、正確にアラートを送付することが可能であった。

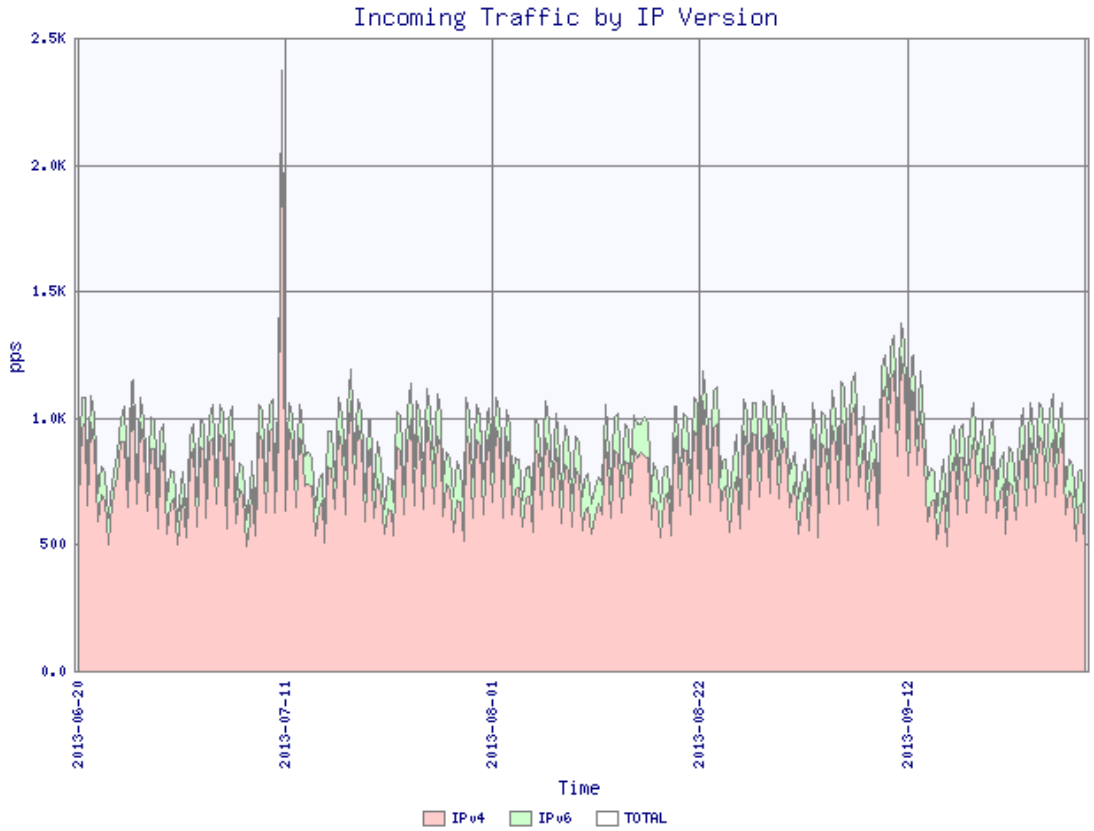



図 5：DIX-IE における Root DNS のトラフィック動向

## (4) EtherOAM の導入

本年度前半においては、EtherOAM 導入に関する進展は無かった。本年度後半において、まずは NSPIXP-23 の東西間での導入に向け、検証と検討を進め、導入を計画する所存である。

## (5) Software Defined IX の設計検討

SDN-IX の設計に関して、より多くの意見を求めるため、次の 2 つの会合を開催した。

- SDN-IX 導入に向けたブレインストーミング in IXP meeting
  (2013 年 6 月 13 日＠幕張メッセ)
- SDN-IX パネルセッション
  (2013 年 9 月 19 日＠恵比寿ガーデンホール)

これらの会合を通じ、ISP やコンテンツ事業者が持つ SDN-IX のイメージな

らびに SDN-IX の機能といったものを洗い出すことができた。

IX は、AS 間を相互に繋ぐ中間に位置しており、IX でフィ ルタリングができれば、サイバー攻撃を AS に流入する手前で止めることが可能となる。これにより、各 AS の対外接続の帯域が不要なトラフィックに浪費されることを防ぐことができる。しか、既存の IX に設置された L2 スイッチでは、AS 運用者からの要望に合わせて IX 内でフィルタリングを行ったり、トラフィック毎の転送制御を行ったりすることができない。そこで SDN 技術を用いると、外部のプログラムから動的に転送制御を行うことができ、かつ既存の経路制御プロトコルと異なる独自の転送制御を行なうことができる。

これらの利点を生かし、NSPIXP Project では、次の 4 機能を提供する SDN-IX を構築することを目指すこととした。

- コンテンツ種別による転送先制御
- 悪意のあるトラフィックフィルタリング
- 悪意のあるトラフィックの緩和
- 外部の脅威情報共有機構との連携

これにより、従来の IX では困難であった、コンテンツに応じたトラフィック制御や、より無駄のないトラフィック交換が可能になると考える。例えば、図 6 に示すように、トラフィックの内容に応じて囮となるサーバにトラフィックを誘導し、フィルタリングを行なうことで、正しいトラフィックのみをピアリング相手に渡すことが可能となる。

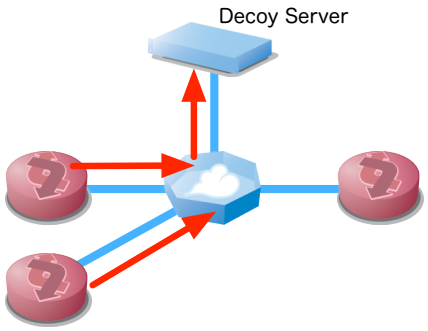



**図 6：DDoS トラフィックの緩和**

さらに、SDN-IX を提供する機器は、従来の機器に比べてかなり安価な機器を使って実現する予定である。これにより、ポート単価を抑えることができ、10Gbps や 40Gbps のポートを、現在の 1Gbps ポートの単価程度にて提供することが可能となる予定である。これにより、他の IX のバックアップ用途として

の IX に用いることができるようになり、非常時のトラフィック交換に用いることができる IX を実現することができる。

現在、SDN-IX の実現に向けて設計や機能を検討中であり、本年度中のテスト運用開始を目指している。

## 4. おわりに

本報告書では、2013 年に行った、実証実験に関する成果報告について述べた。2012 年に引き続き、これからの ISP やエンドユーザに求められる高度情報インフラストラクチャとしてのパブリック IX サービスのありかたというものを念頭に、より強固なインターネットバックボーンとサービスを実現するための、高度な運用技術の研究開発ならびに実証実験を行っていく方針である。

特に SDN-IX に関しては、本年度中の実現を目指し研究を進める。SDN-IX として用いる機材は、ファームウェアレベルで柔軟なソフトウェア改造が行える機材、もしくは OpenFlow に適している機材を用いる予定であり、現在機種選定を進めている。さらに、SDN-IX においては従来の IX に無い機能を提供するのみならず、DIX-IE や他の IX のバックアップ IX として機能すべく、10G ポートならびに 40G ポートのみの提供とし、10G ポートならびに 40G ポートを安価に提供する予定である。その代わり、用いる機材やソフトウェアの信頼性は、従来の DIX-IE 機材より研究側に傾向したものとなる予定である。SDN-IX の新機能やバックアップ IX としての機能を理解し、積極的に参加してくれる参加組織を募集する。

NSPIXP Project では、従来の IX とは違うより柔軟な IX の構築を目指すことで、商用 IX とは異なった性格の実証実験を推進していく所存である。