

第 XXXI 部  
IX の運用技術



## 第 31 部 IX の運用技術

### 第 1 章 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洗い出し、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する。具体的には、WIDE Project のサブプロジェクトである NSPIXP (Network Service Provider Internet exchange Point) 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う。

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、現在では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、本年は 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。

特に、Internet eXchange Point を物理的に配置した環境での ISP 間のトラフィック交換において、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、メディアおよび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ボリュームを考えた新規ネットワークポロジへの移行を終え、IPv6/IPv4 デュアルスタック運用、Root DNS や Teredo、6to4 などの、個々の組織のみではコスト的に運用メリットが存在しないものを取りまとめ、IX として公共的なサービスを提供することに重点をおいた実証実験を行っている。

### 第 2 章 本年の研究計画

本年の研究課題は以下の通りである。

#### 2.1 IX アーキテクチャに関する研究

##### (1) 広帯域化に対応し、安定性かつ冗長性を向上させた IX アーキテクチャの研究

2010 年は、2009 年に導入した DIX-IE 新アーキテクチャのスイッチ導入を受け継ぎ、新環境においてより安定した運用を提供するための運用管理に関する検証をすすめた。これを引き継ぎ、2010 年も、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構成に関する実証実験を引き続き実施している。また、NSPIXP-3 においても、2009 年に新アーキテクチャによる機材更新を行った。これによって、より大容量なトラフィックを交換することができるようになり、かつトラフィックの成分分析ができるようになった。この流れを引き継ぎ、2010 年においては成分分析や積極的な到達性監視による障害回避等の運用技術検証を進めた。

さらに、本年は DIX-IE と NSPIXP-3 の結合に関する試験も実施する予定である。これにより、NSPIXP-3 に存在する組織と DIX-IE に存在する組織が直接 BGP にて Peering を行うことが可能となる。この実証実験は、当初の計画には無かった項目であるが、その要望の高まりと、他の商用 IX との差別化のために計画された実証実験である。実証実験開始当初は、NSPIXP-3 と DIX-IE を結ぶ VLAN を別途用意し、希望する組織に対しては Tagged VLAN もしくは別ポートとして提供する。

##### (2) 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。平成22年9月現在、31組織がIPv6におけるBGP peeringを行っており、2009年に比べ、10組織ほど増加している。今後、さらにIPv6によるBGP peering参加組織は増加するものと思われる。また、研究目的にても述べたとおり、インターネットにおけるIXの果たす役割が変わりつつある。これにともない、IX上にて交換されるトラフィックの流量も成分も変化しつつある。この変化の傾向をとらえるために、sFlowを用いた成分分析を開始した。

成分分析を行うにあたっては、フリーソフトウェアならびに商用ソフトウェア、また独自開発のソフトウェアによるsFlowコレクターを運用し、それぞれによるトラフィック成分分析の結果を比較することで、よりIXの成分分析に適したsFlowコレクターのありかたに関する検証を開始した。

## 2.2 パブリックIXにおける各種サービスに関する研究

### (1) IPv6/DNSSEC対応RootDNSのIXにおけるサービス形態に関する研究

2009年は、DNSの世界に大きな変更が生じた。DNSSECの導入が本格的に開始され、RootDNSゾーンや逆引きの.arpaゾーンに対してDNSSECが導入され、運用が開始された。これにともない、RootDNSサーバへのトラフィック傾向に変化が現れている。

RootDNSサーバのような公的なサービスは、各ISPがそれぞれ独自にサービスを提供する類のものではなく、直接利益に結びつくサービスでもない。そのため、各ISPが独自に行うのではなく、公的な役割を持つIXにて提供される形態が望ましい。これは、ローカルピアリングが進行する傾向においては、より顕著に必要とされるサービスとなる。ローカルピアリングやトランジット接続では、そのトラフィック流量に応じて課金される形態が通常であり、顧客のみが使うのではない公的なサービスを提供するには適さない接続形態となっている。そのため、IXにおいてRootDNSサーバが提供されることは、インターネットの健全なサービス運用にとって重要であり、DIX-IEならびにNSPIX-3はDNSSEC時代におけるRootDNSの運用と

そのサービス状況に関する検証を引き続き行う。

また、2011年1月にはJPDNSのDNSSEC対応も表明されており、NSPIX-3にはJPDNSが存在している。そのため、2011年におけるJPDNSサーバDNSSEC対応後のトラフィックの変動に注目したい。

### (2) 6to4やTeredoといったIPv6移行に関するサービスの提供

前述の研究課題でも述べたとおり、IXの役割の変化に伴う公的なサービス提供がこれからのIX研究課題の主題となりつつある。その流れをうけて、IXにおいてIPv6移行サポートサービスを提供することを計画している。

2009年にも一部の移行サポートサービス提供を開始しており、6to4サービスは現在DIX-IEにおいて参加ISPに対して提供されている。6to4サービスに引き続き、Teredoサービス提供も予定しており、現在試行運用段階にある。2010年には、6to4サービスの提供とともに、Teredoサービスの本格的な提供が開始された。これによって、各ISPが個々に提供するのはコストに見合わない、もしくは効果が薄いIPv6移行サポートサービスをIXにて提供するための実証実験を行う。

本中間研究報告書では、これらの研究項目に関して中間成果報告を行う。

---

## 第3章 研究成果

---

本章では、各研究項目に関する成果を報告する。

### 3.1 IXアーキテクチャに関する研究

2010年では、2009年に引き続き、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3.1に平成22年9月時点でのDIX-IEならびにNSPIX-3の実証実験拠点を示す。

2010年10月から、研究費ならびにポート代金の改定が行われた。これによって、GbEと10GbEは実質の値下げとなった。FEに関しては、料金据え置きとなった。これは、NSPIX-3ならびにDIX-IE

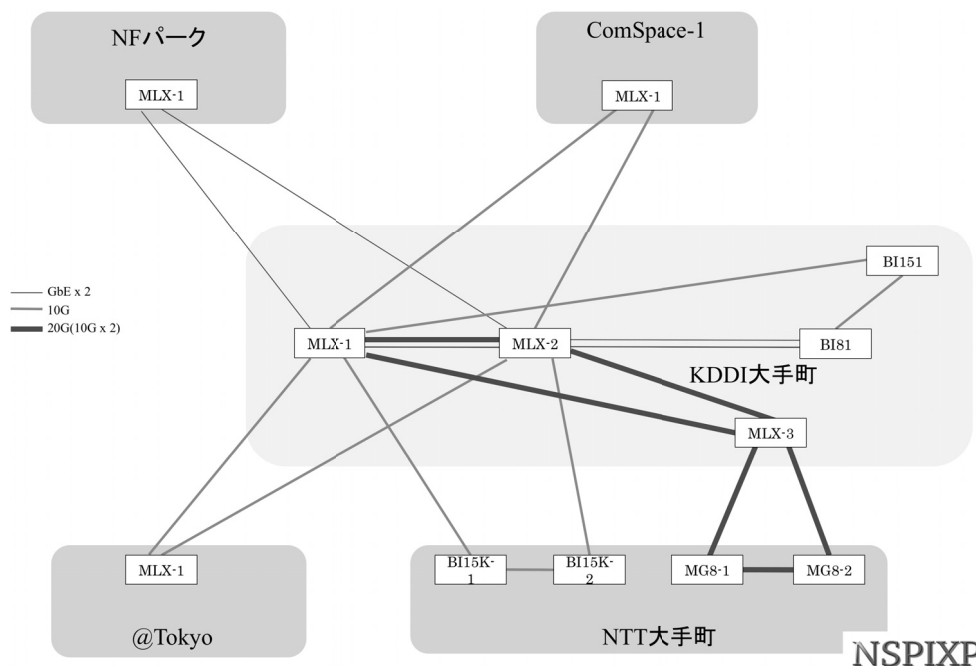


図 3.1. DIX-IE 拠点構成図

表 3.1. 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)
	K-Opti 湊町拠点 (WIDE)

表 3.2. DIX-IE における接続 ISP 数

接続メディア別	10 Gbps	9
	1 Gbps	41
	100 Mbps	20
接続拠点別	KDDI 大手町拠点	53
	NTT 大手町拠点	9
	NF パーク拠点	3
	ComSpace-1 拠点	4
	@Tokyo 拠点	1

において、基本的に FE による接続を終了する方向にあるからである。機材のコストやその需要の点から FE の接続サービスを維持することが困難となったためであり、NF 西大井拠点を除いて、その他の拠点では FE による接続サービスを終了する。図 3.1 に 2010 年 9 月における DIX-IE の拠点構成図を示す。

さらに、DIX-IE ならびに NSPIXP-3 の 2010 年 9 月現在における接続 ISP 数を表 3.2 と表 3.3 に示す。

次に、図 3.2 に 2010 年 6 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示す。

以上のデータより、本年の前半の傾向をまとめる。本年は、前年に比べ接続 ISP 数の減少ならびにトラフィック総量の減少が顕著に見取れる。これは、

表 3.3. NSPIXP-3 における接続 ISP 数

接続メディア別	1 Gbps	15
	100 Mbps	0
接続拠点別	NTT 堂島	14
	K-opti 湊町	1

ISP の統廃合による日本国内の ISP 数の減少、ならびにパブリック IX における BGP peering から、プライベート BGP peering への移行が原因となっている。この傾向は、総務省がまとめている「我が国のインターネットにおけるトラフィックの集計・試算<sup>1</sup>」にも顕著に現れている。

その一方で、IPv4 BGP peering は減少しているが、IPv6 peering は微増ながら増加している。これは、IPv4 のトラフィック交換は、大手 ISP が下流 ISP

1 [http://www.soumu.go.jp/menu\\_news/s-news/01kiban04\\_01000001.html](http://www.soumu.go.jp/menu_news/s-news/01kiban04_01000001.html)

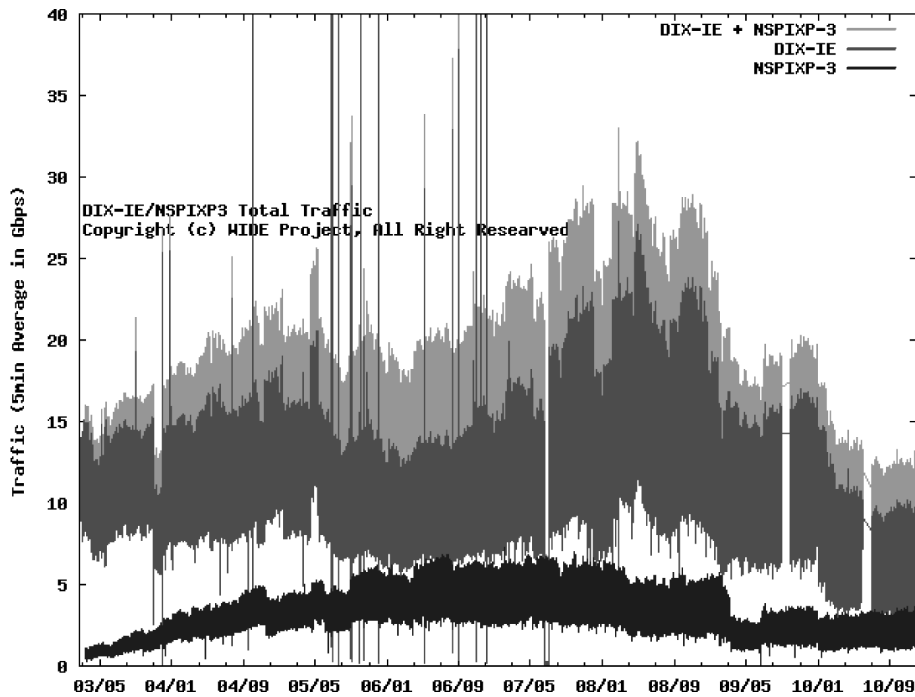


図 3.2. DIX-IE ならびに NSPIXP-3 におけるトラフィック総量の推移

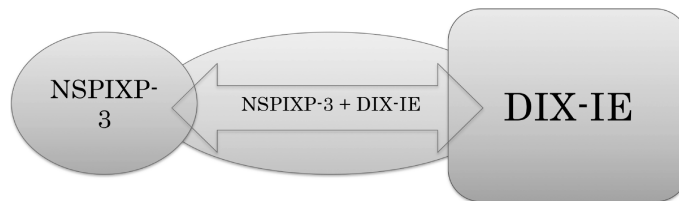


図 3.3. NSPIXP-3 と DIX-IE の結合実証実験イメージ

に対してトランジットを販売する形態が増えている一方、IPv6はまだ完全に商用ベースになっていないため、IXにおいてより多くのISPとBGP peeringすることが必要な段階であるためである。

また、本年前半は、sFlowを利用したトラフィック成分分析を開始した。NSPIXP-3において、sFlowを利用したトラフィック成分分析を行うための設定を機器に投入し、sFlowコレクターに対して情報提供を開始した。Flowサンプリングレートは8192にて情報収集を行っている。

現在はNSPIXP-3の管理目的のみにsFlowの分析結果を用いているが、今後傘下組織への提供も視野に入れて、アプリケーション作成を行っていく。傘下組織が、地組織のトラフィック成分情報ならびにNSPIXP-3全体のトラフィック成分情報を把握することで、IXにおけるトラフィック交換のありかたに関する議論を深めていきたい。

また、当初の研究計画には無かった実験項目であるが、DIX-IEとNSPIXP-3の連結に関する実験を計画している。これによって、NSPIXP-3参加組織とDIX-IE参加組織が直接BGP Peeringを行うことが可能となる。実証実験開始当初の形態としては、NSPIXP-3とDIX-IEを結ぶ別VLANを準備し、希望する組織にはTagged VLANもしくは別物理ポートでの提供を行う予定である。実験予定図を図3.3に示す。

### 3.2 パブリックIXにおける各種サービスに関する研究

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ている。

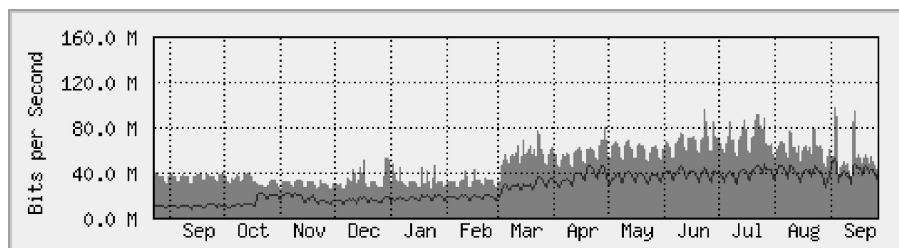


図 3.4. M Root DNS サーバのトラフィック総量

トラフィック総量が増えるに従って、ISP はトラフィックをさばくための設備投資費や人件費が増大する傾向にある。すると、資金面に強い ISP はその設備を増大し、顧客からの要望に応えることによって、さらに規模を拡大していく。その一方、あまり資金面に強くない ISP は、増え続けるトラフィック総量に対して設備投資を行うことができず、顧客はより良い ISP へと移動してしまう。そのため、現在 ISP の二極化が進んでおり、国内における ISP の数も減少傾向にある。

すると、大手 ISP がそのトラフィック交換のほとんどを担うようになり、中小 ISP は、大手 ISP からそのサービスを購入し、顧客に対してより安い価格で再販する形態をとりはじめている。そのため、パブリック IX における IPv4 BGP peering 数も減少している。パブリック IX における BGP peering では、基本的にトラフィック等価交換の原則で成り立っている。すなわち、お互いに交換し合うトラフィック総量に大きな差がないという前提で無償でのトラフィック交換が行われている。しかし、ISP の二極化が進むことでその前提がくずれ、中小 ISP は大手 ISP と、IX にて IPv4 BGP peering を行うことが困難となっている。その結果、中小 ISP はパブリック ISP から撤退し、大手 ISP からサービスを購入するという形態が増加している。

これは ISP の二極化であるとともに、インターネットに存在する公共的なサービスの冗長性や性能を低下させる結果にもつながる。Root DNS やその DNSSEC への対応、また IPv6 移行技術の提供といった公共的なサービスは、ISP 単体で行うことが難しい。これは、そのサービスが自身の顧客のためだけに存在しているものではなく、全世界のすべてのインターネットユーザに対するサービスとなり、ISP 単体でサービスを行うのは、その意図からもコスト見合いからも採算に合いにくいものであるからである。

このような公共的なサービスを提供するのはこれからの IX の役目であると考え、DIX-IE ならびに NSPIX-3 では積極的にこのようなサービス展開を行っている。一部の商用 IX においても、このような動きは出始めており、IPv6 移行サービスを展開する予定となっている。その一方で、IX への参加組織数が減少してしまうことにより、多くの ISP に対してそのサービスを直接提供することが困難になってしまう。当然、これらの公共サービスは大手 ISP を経由してユーザに提供されることとなるが、サービスの冗長性という観点から見ると、特定の手 ISP の接続性にそのサービスが依存することとなり、自律分散的なサービス展開が困難となる。

本中間報告では、以下の 2 つの公共サービスに関する報告を行う。

- (1) Root DNS サーバの DNSSEC 対応
  - (2) IPv6 移行サービス Teredo の本格運用開始
- (1) Root DNS サーバの DNSSEC 対応**

DIX-IE においては、Root DNS サーバを提供し続けている。本年前半では、この Root DNS サーバの DNSSEC 対応が行われた。これによるトラフィックデータの推移を図 3.4 に、PPS (Packet Per Second) 総量を図 3.5 に示す。

DNSSEC 対応後にトラフィックならびに PPS が増加しているのが見て取れる。これは、DNSSEC 対応によってユーザに返答されるデータ量が増加したため、トラフィック量の増加として表れている。また、従来は UDP で返答が行われていたものが、一部 TCP にて返答が行われていることも、トラフィック増の原因となっている。返答データが大きくなると、UDP の一つのパケットにすべてのデータを取めることが困難となり、その場合には TCP にてユーザへのデータ返答が行われる。TCP の場合、3 Way Handshake 等のコネクションのためのパケット転送も行われるため、PPS の増加という結果につながっている。



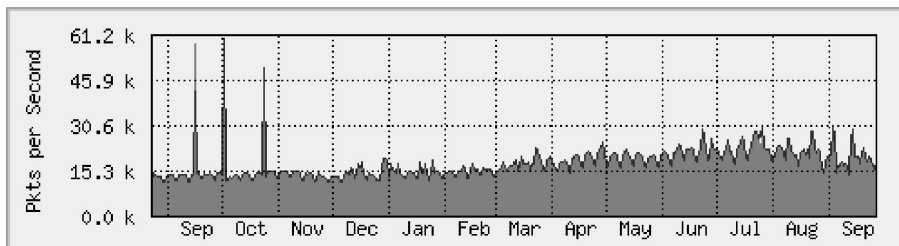


図 3.5. M Root DNS サーバへの PPS 総量

(2) IPv6 移行サービス Teredo の本格運用開始

本年前半は、DIX-IE において Teredo サービスの本運用が開始された。IPv6 移行技術それぞれの特徴を表 3.4 にまとめる。

また、Teredo にて利用される IPv6 アドレスのフォーマットを、表 3.5 に示す。

現在 DIX-IE においては、6to4 サービスが実運用されている。しかし、6to4 では多くのユーザが家庭内にて利用している NAT 技術とともに用いることができず、一部のユーザにしか IPv6 移行技術を提供できていない。一方、Teredo は NAT 内部からも IPv6 での通信が可能となるため、移行技術を利用できるユーザ数が増加する。そこで DIX-IE においても、Teredo サービスの実運用を開始する。

NAT 環境下における Teredo の動作概要を図 3.6 に示す。

Teredo の実サービス開始によって、より多くの

ISP ユーザに IPv6 移行技術を提供することが可能となった。これをうけ、IX における 6to4 ならびに Teredo を利用したトラフィックの総量も増加している。トラフィック総量増加の様子を図 3.7 に、Teredo 単体のトラフィック総量を図 3.8 に示す。

6to4 や Teredo といった公共的なサービスを展開することで、パブリック IX というものの存在意義を、従来の信頼性を有した大量トラフィック交換という役割から、新たな公的サービスを提供するための高度インフラとしての役割に変えていく方針である。2009 年、2010 年はこの方針に従い、公共的なサービスをより充実させる方向で実証実験を行ってきた。今後もさらにこの動きを加速させ、P2P のキャッシュサーバや著名 ftp サーバ等の誘致を行ない、IPv6 移行のためのアプリケーション層まで含めた公共的なサービスを展開していく方針である。

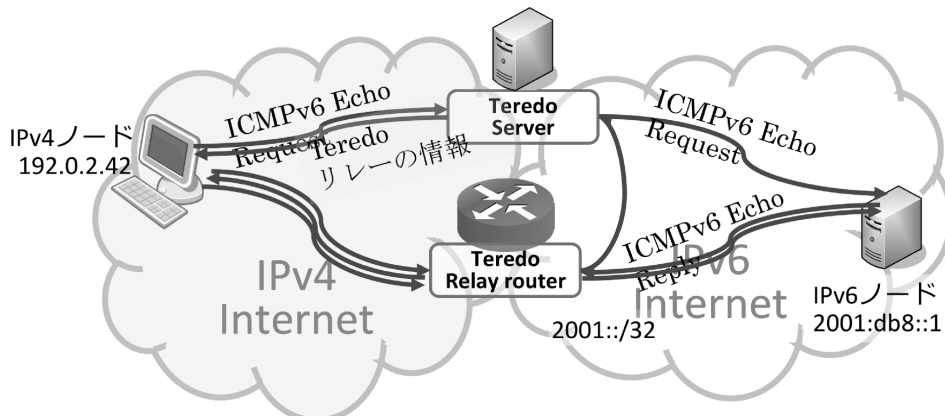
表 3.4. IPv6 移行技術のまとめ

	6to4	6rd	ISATAP	Teredo
IPv4 アドレス	Anycast	Unicast	Unicast	Unicast
IPv6 アドレス	Anycast	Unicast	Unicast	Anycast
NAT 越え	未対応	未対応	未対応	一部対応 (UDP/IPv4)
アドレス	/48	可変 (/56 等)	/128	/128

表 3.5. Teredo の IPv6 アドレスフォーマット

	0~31	32~63	64~79	80~95	96~127
長さ	32 bits	32 bits	16 bits	16 bits	32 bits
説明	Teredo プレフィックス	Teredo サーバ IPv4 アドレス	フラグ	UDP ポート	クライアント パブリック IPv4 アドレス
部分 復元	2001:0000	4136:e378	8000	63bf	3fff:ddd2
		65.54.227.120	cone NAT	40000	192.0.2.45





TeredoでNAT越えに利用する packets を「バブル」と呼ぶ  
通信が確立されるまで数秒程度かかる

図 3.6. NAT 環境下における Teredo の動作

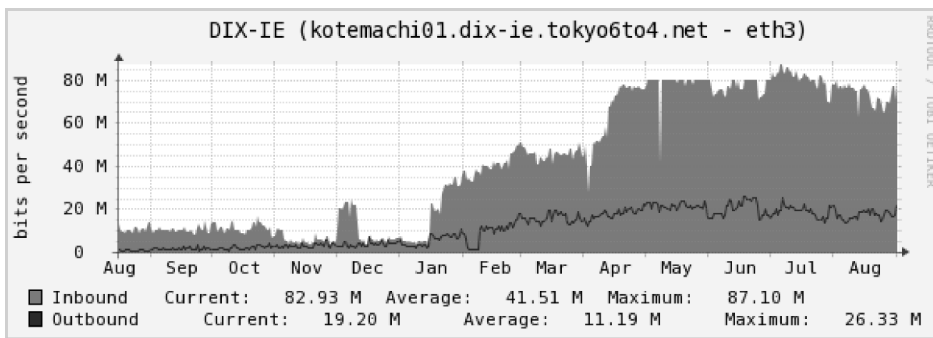


図 3.7. 6to4 + Teredo トラフィック総量

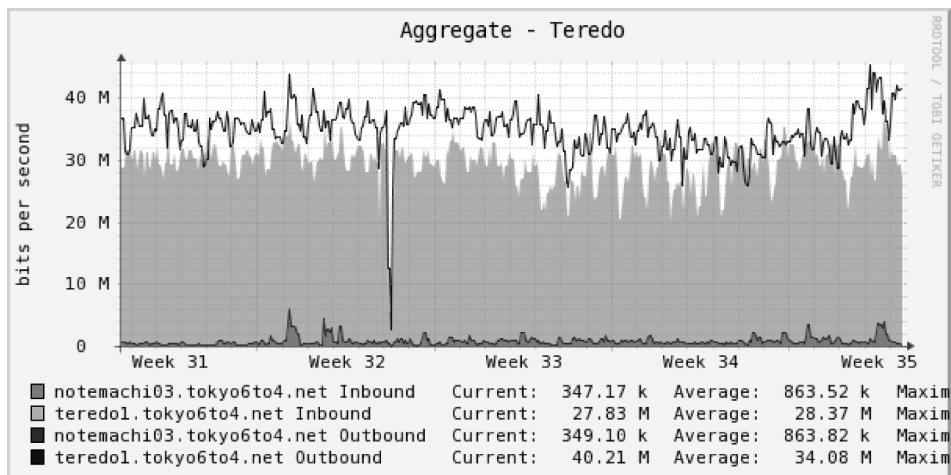


図 3.8. Teredo トラフィック総量

---

#### 第 4 章 おわりに

---

本報告書では、2010 年に行った実証実験に関する成果について述べた。2009 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、さらに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。具体的には、トラフィック成分分析による IPv6 移行支援や、Teredo サーバのさらなる増強、DIX-IE と NSPIX-3 の結合実験、EtherOAM による品質検証、アプリケーション層まで含め IPv6 移行支援技術の積極的な導入、また、クラウドを利用した仮想計算機提供サービス等との連携など、商用 IX とは異なった性格の実証実験を行う予定である。