

第 XXXIII 部

IX の運用技術

第 33 部 IX の運用技術

第 1 章 はじめに

NSPIXP (Network Service Provider Internet eXchange Project) は、商用インターネットを相互に接続する場合の問題点について、実証的な手法で研究を進めるプロジェクトである。ここでは、NSPIXP WG が 2004 年に行った活動内容や、NSPIXP の現状について記す。

第 2 章 DIX-IE (Distributed IX In EDO)

DIX-IE が稼働してほぼ 1 年が経過した。2004 年度には、10 Gigabit Ethernet での接続希望増にともない新たな Switch である Foundry 社製 BigIron

MG8 を KDDI (WIDE Project) および NTT Communications 拠点に設置し、KDDI 拠点では、既存の BigIron15000 および BigIron8000 からの Gigabit Ethernet、10 Gigabit Ethernet ユーザの移行、NTT Communications 拠点では、新規の 10 Gigabit Ethernet 接続を収容することとした。

MCI 拠点は 2004 年度の接続復帰を試みたが、BPDU その他の問題により、現在も接続を停止している状態である。

現在、DIX-IE は以下の 6 拠点で構成されている。

- KDDI (WIDE Project) 大手町
- MCI 新川
- MIND 西大井
- NTT Communications 大手町
- @Tokyo 豊洲
- AboveNet 日本橋

上記の拠点は、10 Gigabit Ethernet や Gigabit Ethernet のリンクアグリゲーションを用いることにより、4~20*2 Gbps の帯域で Core Switch と接続されている。現在の構成図を図 2.1 に示す。

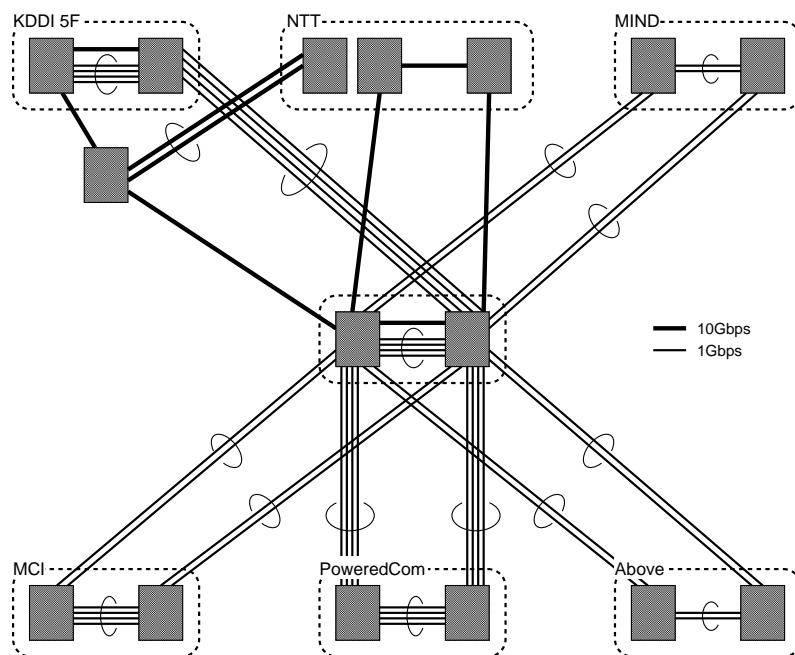


図 2.1. 現在の DIX-IE の構成図

2.1 接続組織数

2004 年 9 月 7 日現在、DIX-IE に接続している組織数は 73 である。また、1 組織がトラフィック増加への対応や冗長性確保のために、複数ポートで接続するケースも珍しくない。そのため、接続ポート数は接続組織数を大きく上回る 103 ポートにのぼる。これらをメディアの種別ごとに整理すると、以下のようになる。

- 10 Gigabit Ethernet 3
- Gigabit Ethernet 72
- Fast Ethernet 26

トラフィック増加への対応に、10 Gigabit Ethernet での接続を検討している組織が徐々に現れはじめており、上記のようにすでに接続している組織も現れはじめています。また、接続ポート数を分散拠点ごとに整理したものは以下の通りである。

- KDDI (WIDE Project) 65
- MCI (拠点停止中) 1
- MIND 5
- NTT Communications 24
- @Tokyo 2
- AboveNet 8

この 1 年で 10 Gigabit Ethernet の接続要求が増加するとともに IPv6 もしくは IPv4/IPv6 Dual での接続が 8 port 以上となった。

2.2 運用の履歴

冒頭で述べたように、2003 年 3 月より DIX-IE は実運用に入ったが、運用開始後も引き続き新規技術の deploy field としての役目を担い、数々の firmware を試しているが、それにともないいくつかの問題も発生している。2004 年 9 月には、MG8 の導入にも関連し、不定期に Core からの BPDUs が特定拠点の Switch に届かなくなるという問題が発生しており、当該 Switch 間での packet dump を含め、問題解決のための取り組みを実施している。KDDI 拠点での旧 Switch から MG8 への移行においては、いくつかの組織において、当該組織側のハードウェア故障により問題が発生したが、ハードウェアの交換によりこれらの問題を解決している。

2.2.1 運用問い合わせ窓口

運用問い合わせ先として新たに、24 時間の電話応

対窓口を開設し、共同研究組織からの問い合わせのみを 24 時間受け付けることとした。共同研究組織からの問い合わせに限定するため、各組織からは運用担当者のメンバリストを提出して頂き、そのメンバからの問い合わせのみを受け付ける。運用問い合わせ先では、連絡を受けたことに対して、現状を確認し、DIX-IE 運用チームへのエスカレーションだけでなく、関連するマシンの監視についても行うこととした。

2.2.2 拠点間リンクでの障害

2004 年 1 月 5 日、Core-MIND 拠点間のリンクにおいてインタフェースが Down/Up する事象が発生したが、これは当該リンクを収容する伝送装置が reboot したことに起因することであった。

Link Aggregation の問題から Core-Above 拠点間リンクでは IEEE 802.3ad による Link Aggregation を外し Gigabit Ethernet でのリンクにより対応した。

2004 年 1 月 17 日、2004 年 2 月 17 日、2004 年 3 月 17 日、Core-MIND 拠点間リンクにおいて Topology Change が発生したが、これは当該リンクにおいて IEEE 802.3ad の bug に関連することがわかり firmware を入れ替えることにより対応した。

継続して、何らかの問題により連続して 2 つ以上 BPDUs を正しく受け取れない場合に起こる、MsgAgeExpiry が Core-MIND 拠点間にて発生した。Core 側の収容ポート変更を行うことで、問題を解決した。

2004 年 5 月 28 日、Core-NTT Communications 拠点間リンクで Down/Up が発生した。これは XENPAK に関連する光レベルの振れが原因と思われる、接続面の清掃後問題は収まった。

2004 年 6 月 14 日、Core 内リンクにおいて Topology Change が多発し、全拠点に対して影響を及ぼす事となった。原因としては Core Switch 間の冗長リンクに使用しているインタフェースに問題があり、それを外す事で対処を行った。

2004 年 12 月 29 日、Core-NTT Communications 拠点間リンクで Topology Change が多発する状態になった。Core 側 XENPAK のハードウェア故障が原因であったため、XENPAK の交換を行うことにより、問題を解決した。

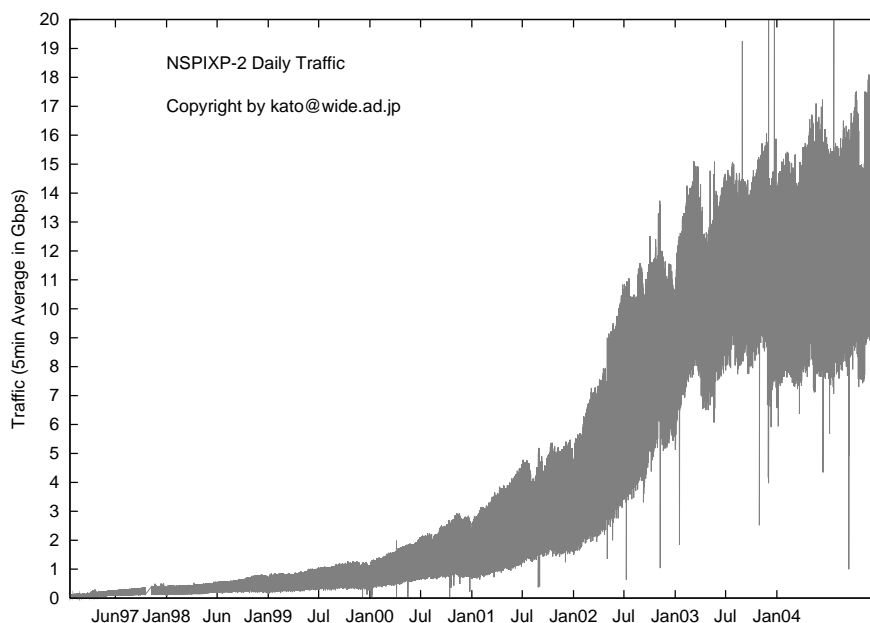


図 2.2. 現在までの総トラフィックの推移

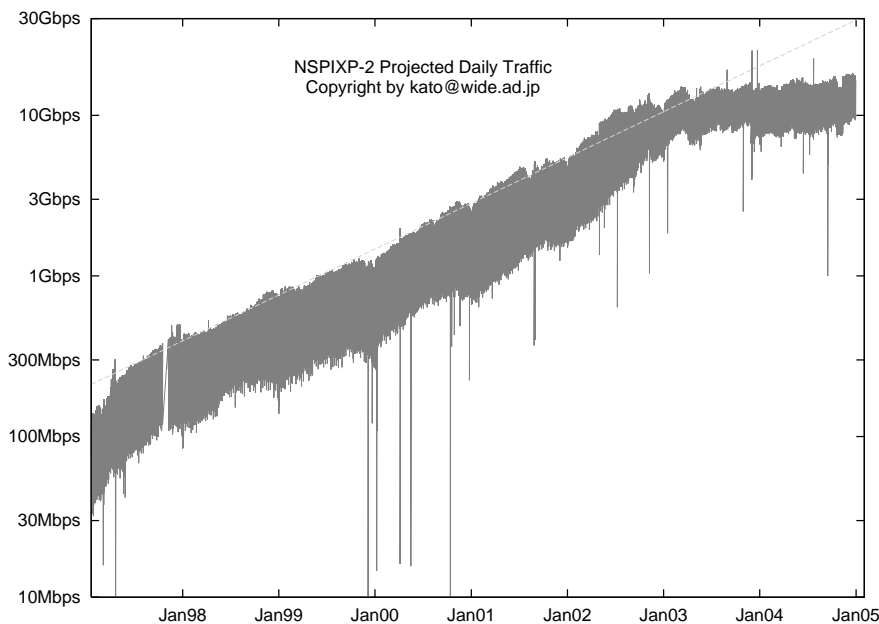


図 2.3. 現在までの総トラフィックの推移 (ログスケール)

2.2.3 ファームウェア更新

上記の IEEE 802.3ad bug 問題へ対応するために、2004 年 3 月 22 日、KDDI (WIDE Project) 拠点、MIND 拠点、Above 拠点の Switch において、ファームの更新作業を行った。@Tokyo 拠点の firmware bug のため、2004 年 3 月 31 日に@Tokyo 拠点の Switch において、ファームの更新作業を行ったが、入れ替えたファームにマネージメント機能の bug が発見さ

れたため、切り戻しを行った。

2.2.4 未解決の問題

ファームの変更、Core 側での収容ポート変更後も、不定期に発生している BPDU を連続して受信できていないために Topology Change してしまう問題に関して、再度 packet dump を含む調査を実施中である。

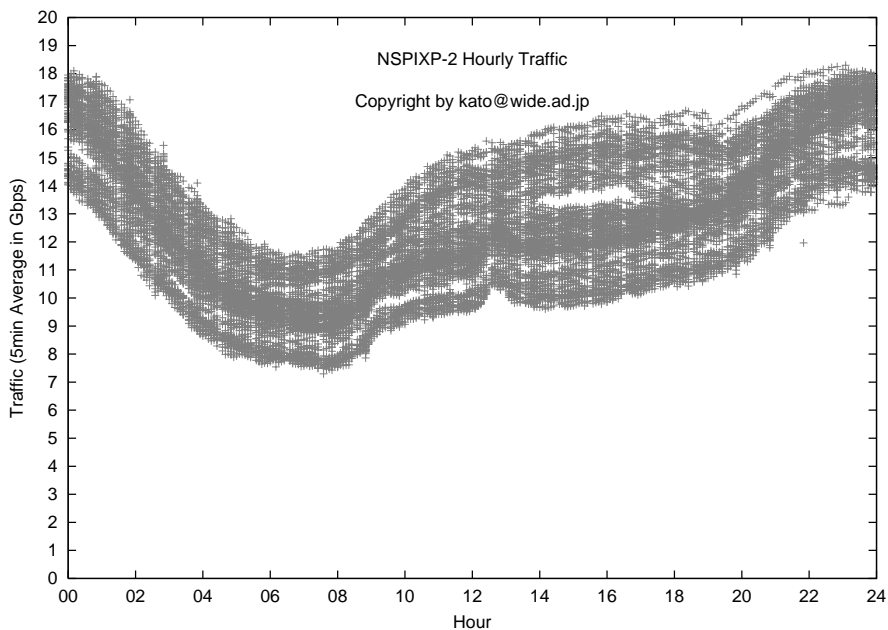


図 2.4. 1 日のトラフィックの推移

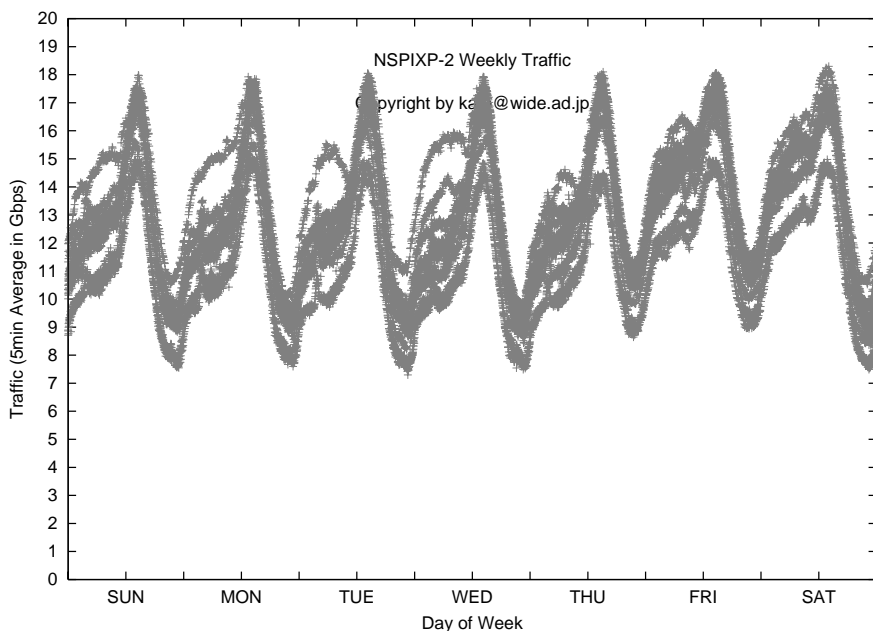


図 2.5. 1 週間のトラフィックの推移

2.3 トラフィックの推移

本節では、DIX-IE で交換されているトラフィックに関して述べる。

図 2.2 のグラフは、NSPIXP-2 の稼働時からの総トラフィック量の推移を示したものである。ここで総トラフィックとは、ISP を接続しているポートの流入トラフィックの総和である。2004 年 4 月後半、

11 月にトラフィックの減少傾向が見られる。

図 2.3 は、図 2.2 の Y 軸のみをログスケールに変換したものである。破線は

$$Traffic = b * \exp(a * year)$$

で示される近似であり、*a* は 0.657 である。この図の通り、DIX-IE で交換されているトラフィックはこの近似線に非常によく沿った形での増加傾向を見

せている。2003年度まで、この増加傾向はおおよそ毎年2倍であったが、ブロードバンドアクセス環境の普及が一段落したこと、ISPの商用IXの接続やISP間のdirect peerが広帯域化していることもあり、DIX-IEにおけるトラフィックの増加傾向は多少抑えられてきていることがわかる。

図2.4は、2003年11月～2004年1月のトラフィックを1日単位で示したものである。ISDNがインターネットへ接続するための主な手段であった数年前は、NTTのテレホーダイサービスの影響で、午後11時にトラフィックの急激な増加が観測された。現在では、ケーブルテレビやADSL、あるいはFTTHのような、いわゆるブロードバンド技術がインターネットへの主な接続手段となり、その結果として過去にみられたような増加はみられなくなった。

また、平日は午後1時前後に若干のトラフィックの増加が見られる。これは、昼食後の休みにホームページを閲覧するなどの習慣によるものであると思われる。

図2.5は、図2.4のデータを1週間単位で示したものである。インターネットの黎明期には週末に大きな減少がみられたが、現在ではそのような減少は観測されず、平日と週末でほぼ同様のトラフィックがみられる。

第3章 NSPIXP-3

NSPIXP-3は、1996年に関西で初のIXとして設置された。また、設置当初より分散IXを意識しており、2拠点の構成で運用が開始された。現在は拠点が増強され、以下の3拠点となっている。

- NTT 堂島
- C&W IDC 福島
- OMP 湊町

上記の3拠点はすべて相互に接続する構成になっており、トライアングルを形成している。なお、拠点間の接続は以下の通りである。

- 堂島～福島 10 Gigabit Ethernet
- 福島～湊町 Gigabit Ethernet
- 湊町～堂島 Gigabit Ethernet

このように、拠点間は10Gbpsあるいは1Gbps

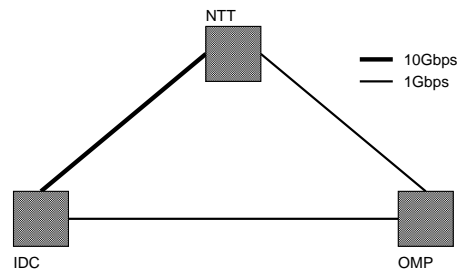


図3.1. 現在のNSPIXP-3の構成図

での接続となっている。また、トライアングルはIEEE802.1w Spanning-Treeによりループを回避しており、定常状態では湊町～堂島拠点間でブロックしている。現在の構成図を図3.1に示す。

3.1 接続組織数

2005年1月8日現在、NSPIXP-3に接続している組織数は23である。また、DIX-IEと同様に、トラフィック増加への対応や冗長性確保のために複数ポートで接続している組織もあり、接続ポート数は26となっている。メディア種別ごとの接続数は以下の通りである。

- Gigabit Ethernet 19
- Fast Ethernet 7

また、接続ポート数を拠点ごとに整理したものは以下の通りである。

- NTT 20
- C&W IDC 4
- OMP 2

3.2 トラフィックの推移

本節では、NSPIXP-3で交換されているトラフィックに関して述べる。

図3.2のグラフは、NSPIXP-3で交換されている総トラフィック量の推移を示したものである。2001年7月～2002年10月までの取得データに関しては、縮退して保存する手法をとっていたため、1日平均の最大値のみとなっている。また、2002年10月～2003年3月までは、NSPIXP-3の構成変更にもなって技術的に正しい値が取得できていないため、表示していない。

この図にみるように、全体としてNSPIXP-3でほぼ単調にトラフィックが増加する傾向に変化はない。しかし、ログスケールでみたときにその傾きが大きく減少していることがわかる。図3.3のグラフ

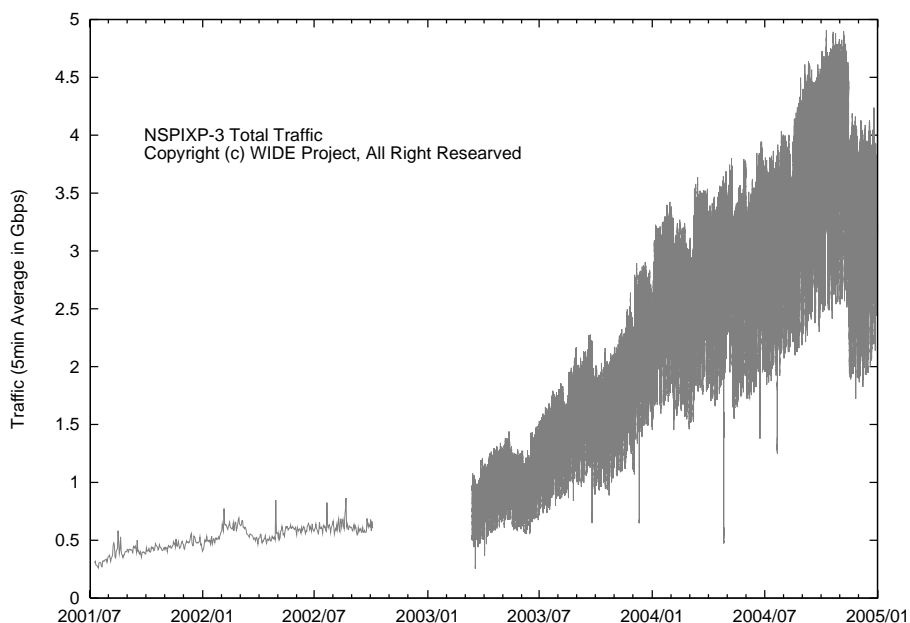


図 3.2. 現在までの総トラフィックの推移

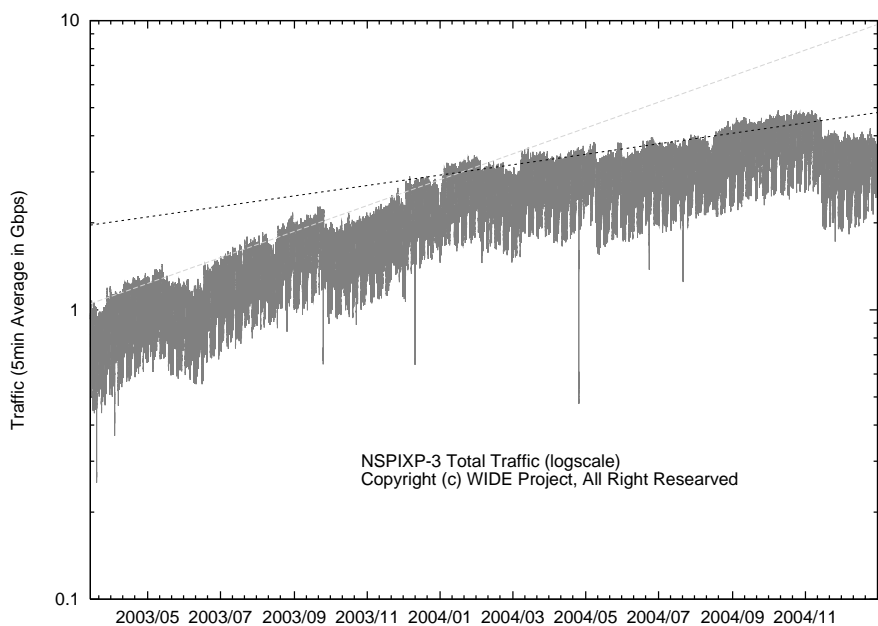


図 3.3. 現在までの総トラフィックの推移 (ログスケール)

は、図 3.2 の Y 軸のみをログスケールに変換したものである。これは、昨年度みられた大手 ISP のトラフィック交換点の分散化がさらに進み、プライベート・ピアの活用が多くなったためであると推測している。

図 3.3 中の 2 本の破線は、

$$Traffic = b * \exp(a * epoch)$$

で示される近似である。ここで a は、傾きの大きい直線に対して $3.91007e - 08$ 、傾きの小さい直線に対して $1.59964e - 08$ 、 $epoch$ は 1970 年 1 月 1 日からの経過秒数である。

図 3.4 から図 3.6 のグラフは、2003 年 10 月～2004 年 12 月のトラフィックをそれぞれ日単位、週単位、および、月単位で示したものである。

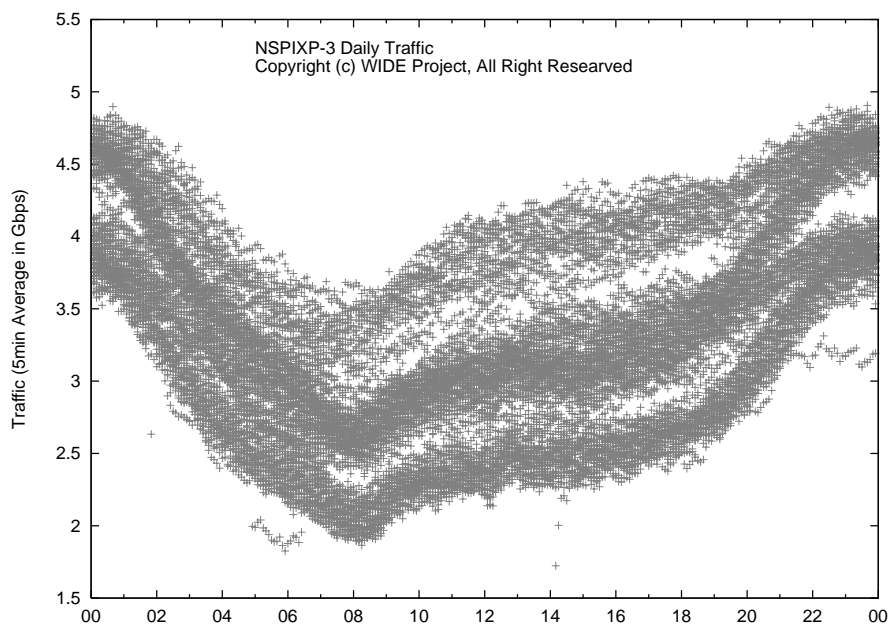


図 3.4. 1日のトラフィックの推移

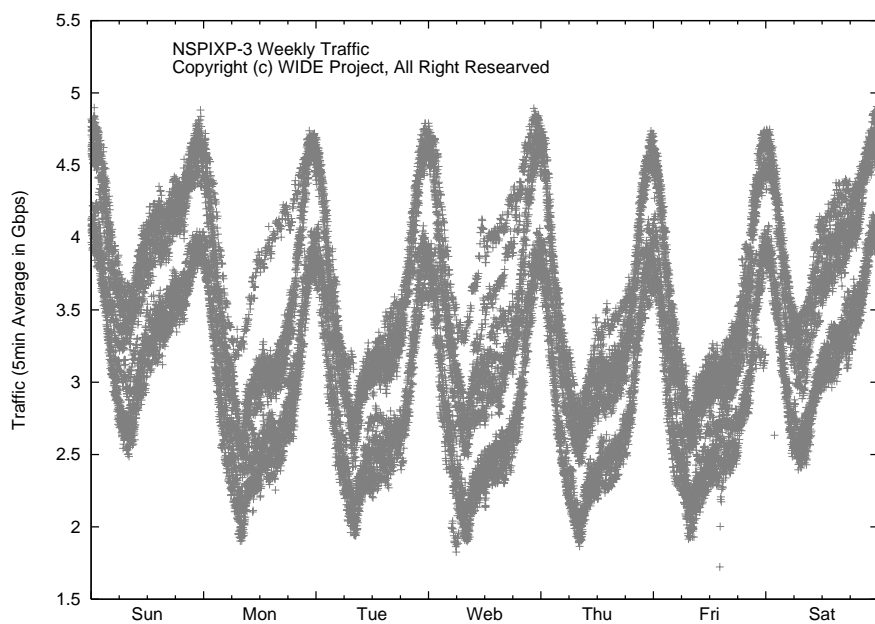


図 3.5. 1週間のトラフィックの推移

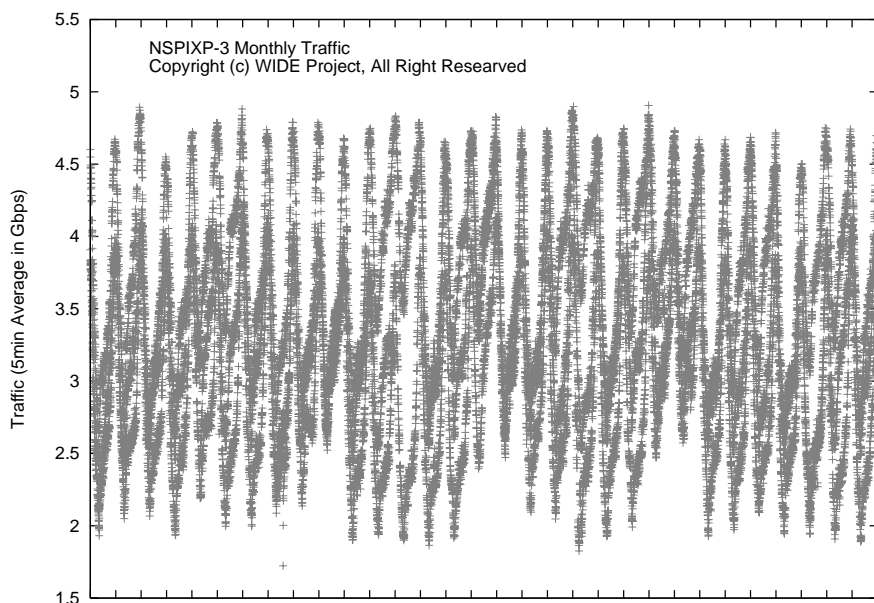


図 3.6. 1ヶ月のトラフィックの推移

第 4 章 関連イベント

NSPIXP では、一時的に接続性を必要とするイベントなどに協力している。ここでは、毎年恒例になっている、NetWorld+Interop への接続性の提供に関して報告する。

毎年、6月の終わりから7月のはじめにかけて、千葉県幕張メッセにおいて Networld+Interop (以下、N+I) と呼ばれるイベントが開催されている。このイベントは、インターネットに関する技術の展覧会であり、各種のベンダが新しい技術を実装した機材の展示やデモンストレーションを行う。会場内には、出展社の機材を収容するためのネットワークが敷設される。このネットワークは、標準化される前の技術などを駆使したものであり、ここで利用された技術が次の年の主流になっているなど、インターネット技術の最先端を結集したものとなっている。

N+I では、各出展社や関係施設に対してインターネットへの接続性を提供する、ShowNet と呼ばれるネットワークが構築される。ShowNet は独自のアドレスブロック及び AS 番号を持ち、全世界で開催される展示会とともに移動するイベントネットワーク

である。東京開催時には、東京の主要な IX に接続して各 ISP との間に peer/transit を行う。NSPIXP は ShowNet の対外接続において中心的役割を果たしている。

『NetWorld+Interop 2004 東京』は、2004年6月28日から7月2日にかけて幕張メッセで開催された。NSPIXP との接続は、大手町に位置する NTT Communications 拠点で行い、DIX-IE および NSPIXP-6 と接続した。DIX-IE に対しては、ShowNet のルータからの 10 Gigabit Ethernet LR による接続、および Gigabit Ethernet LX による接続の 2 種類のメディアを利用した。また、DIX-IE における ShowNet への peer/transit 組織数は 47 を数えた。一方、NSPIXP-6 に対しては Fast Ethernet を利用し、peer/transit 組織数は 22 であった。

第 5 章 他の IX 運用者との情報交換

日本の主要な IX には、NSPIXP/JPIX/JPNAP がある。これらは一部で競合関係にある反面、障害発生時にトラフィックを相互に逃がすなどの協力が必須である。このようなことから、上記の主要 IX 間で情報交換を行うための連絡会である IX-exchanges

を開催している。2004年2月2日に第3回会合が開催され、近年のインターネットの急速な普及、アクセス回線の高速化および低価格化によるトラフィックの急増により、IX間の情報交換だけでなくISPを含めて連携する必要があるという観点から、ISP3社を招いて議論が行われた。今年度は、関係者の都合がつかず、来年度に会合の開催が持ち越され、引続きトラフィックの分散を考慮したアーキテクチャおよび効率的な運用技術にポイントをおき、ISPのトラフィック動向や接続状況の情報交換を行い、各IXにおいてどのようなトラフィック分散をするべきか、どのような構成でIXを形成するかなどの議論を行うことが予定されている。

