

第XIX部

IXの運用技術

第19部

IXの運用技術

第1章 はじめに

この章では IX 分科会の成果について報告する。NSPIXP では、すでに 2001 年 8 月 25 日に NSPIXP-2 において、データリンク技術として Gigabit Ethernet への移行を完了している。スイッチ間のトラフィック増加にともない、NSPIXP-2 においては bi151(Foundry 社製 BigIron15000) と bi81(Foundry 社製 BigIron8000) のリンクを、NSPIXP-3 においてはリンクアグリゲーション、ラビッドスパンニングツリーのサポートなどから、スイッチ機器の更改を行った。

第2章 他 IX 組織との交流

2002 年 12 月 4 日に東京大学情報基盤センターにおいて、国内主要 IX である JPIX および JPNAP とともに、IX 連絡会を開催し、各 IX 間の状況、障害、問題などをお互い情報交換する仕組み作りを開始した。

第3章 NSPIXP-2

この章では NSPIXP-2 における活動内容に関して説明を行う。

3.1 CableTron 社製 SSR(Smart Switch Router) の撤去

1999 年 12 月にデータリンクの Gigabit Ethernet 化のために、導入された CableTron 社製 SSR から bi151(Foundry 社製 BigIron15000) への収容機器変

更が、2002 年 4 月 22 日に行われた。データリンク技術の FDDI から Gigabit Ethernet 変更のために、設置され約 2 年半の役目を果たした SSR は、ポート収容密度や Ethernet トランク/冗長化技術対応などの理由の元、同日電源を落し撤去された。

3.2 スイッチ間リンクの増強

2002 年 4 月末時点において、4 GE であった bi151 (Foundry 社製 BigIron15000) と bi81(Foundry 社製 BigIron8000) 間のリンクが、スイッチ間トラフィックの増加により一部のリンクにおいて問題を起こし始めている事が観測できた。そのため、2002 年 5 月 2 日にスイッチ間リンクを 4 GE から 8 GE へ変更を行った。

3.3 bi81 のマネージメントモジュール変更

ラビッドスパンニングツリー (IEEE802.1w) や 10 GE サポートのために、2002 年 5 月 21 日に bi81 のマネージメントモジュールを高性能な M4 に変更を行った。

3.4 bi151 の GbE モジュール交換

bi151 で確認された ASIC 問題 (ポート 7/7、7/8) のために、bi151 slot7 にある GbE モジュールを 2002 年 6 月 6 日に交換を行った。

3.5 bi151 障害

2002 年 7 月 9 日 12 時 57 分ころより、bi151 において BGP 断および CPU ロードが 99% になる現象が発生した。調査の結果、ある ISP のポート障害に関連して診断中に誤って loop を形成されてしまったためであることが判明し、13 時 40 分ころに復旧を確認した。約 40 分間にわたり、bi151 に収容された ISP は BGP が不安定になり、ニュースや新聞等に取り上げられた。このことから、再度接続 ISP に対して、loopback/multicast packet/broadcast packet 等接続時の注意点について確認し、再発の防止を行った。

3.6 スイッチ間リンク 10 GE 化

スイッチ間トラフィックの単一 MAC アドレス GE

超などの問題やリンクアグリゲーション (802.3ad) 構成問題から、2002 年 11 月 8 日に bi151 と bi81 間のリンクの 10 GE 化を行った。両スイッチ間には 10 GE のバックアップとして、4 GE リンクを残し、STP を用いて切替える設定になっている。

3.7 分散 IX 収容ポートの増設

すでに試験運用を開始している分散 IX との接続に必要な 10 GE のモジュールを 2003 年 1 月 8 日に増設を行った。

3.8 NSPIXP-2 における収容組織

2002 年 1 月末現在、NSPIXP-2 に接続している組織 (複数接続を持つものも 1 つ) 数は 70 である。また、NSPIXP-2 上で IPv6 の経路/交換を行っている組織数は 15 である。

第 4 章 NSPIXP-3

この章では NSPIXP-3 における活動内容に関して説明を行う。

4.1 NSPIXP-3 機器変更

これまで Cisco 社製 Catalyst6509 で構成されていた NSPIXP-3 を拠点間リンク増強、IEEE802.3ad Link Aggregation、IEEE802.1w Rapid Spanning Tree、IPv6 対応にするため、Foundry 社製 BigIron8000 に機器変更を行った。Catalyst6509 を NTT 堂島、IDC 福島、OMP 湊町に設置し各拠点間を 1000Base-ZX で接続、IEEE802.1d にて冗長化構成をとっていたものを、10 GE-LR(NTT 堂島-IDC 福島)、IEEE802.1w にするために BigIron8000 に機器変更を行った。IEEE802.1w での冗長化構成により、障害時の切り替わり 50ms、復帰 2s になった。2002 年 10 月 4 日から順次別リングとして BigIron8000 を各拠点に設置し、NTT 堂島において既存リングと接続、各拠点において順次 BigIron への収容変更を行った。2003 年 2 月現在、NTT 堂島に接続する WIDE プロジェクト (場所変更中のため) を除き全て、BigIron の収容変更された。

第 5 章 分散 NSPIXP-2

この章では分散 IX(分散 NSPIXP-2:Tokyo-IX) における活動内容に関して説明を行う。

5.1 NSPIXP-2 の分散化

NSPIXP-2 の分散化に関しては、2000 年度の報告書でその方向性を示した。また昨年度の報告書において実験内容に関して示した。分散拠点間のデータリンクとして、IEEE802.3ad を用いる事で、複数の Gigabit Ethernet を束ね (トランキング)、また一部の拠点との間では、10 GE(IEEE802.3ae) の LAN-PHY である 10 GE-LR(1300 nm/Serial) を拠点間のリンクとして使用している。現状の分散 NSPIXP-2 の構成を 5.1 図に示す。

5.2 冗長化のための IEEE802.1w

2002 年 10 月以前は、冗長化構成のために IEEE802.1w を簡易実装した Foundry 社仕様の Fast Spanning Tree を使用していたが、IEEE802.1w を実装できたので、2002 年 10 月 8 日以降は IEEE802.1w(Rapid Spanning Tree) を使用している。IEEE802.1w を実装していない異機種に対しては IEEE802.1d を使用し冗長化を確保している。迅速に再計算を行うとともに、IEEE802.3ad の組上がり時間を考慮するために、Forward-delay 9、Hello-time 1、Max-age 6、Priority CORE から帯域順としている。

5.3 CORE 間冗長化

KDDI 大手町に設置した分散 IX 用 CORE スイッチ間のリンクは当初 10 GE だけであったが、10 GE ポートの障害を考慮し、別途 4 GE によるバックアップを作成した。

5.4 過去の問題点

過去の問題点として、IEEE802.3ad、IEEE802.3ae、IEEE802.1w など、標準化途中もしくは実装され始めたばかりの規格においての相互接続性があげられる。また、特定の WDM(Wavelength Division Mul-

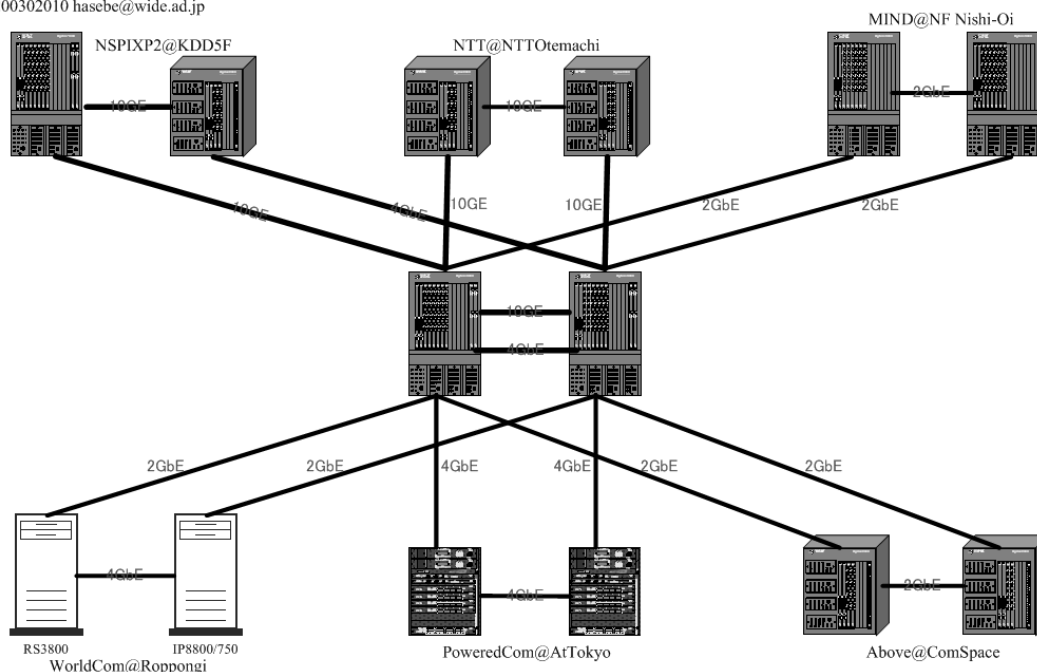
Tokyo-IX Configuration
ver.200302010 hasebe@wide.ad.jp

図 5.1. 分散 NSPIXP-2 の構成

tiplexing) 機器との相性問題なども発見された。

5.5 試験運用

IEEE802.3ad 回りの、標準化および実装問題で 1 年近く運用開始が延期されているが、現在安定が確認できたため、試験運用を開始している。2002 年 1 月末時点で 8 組織が、分散 IX 環境において接続/交換を行っている。

第 6 章 Traffic on NSPIXP-2

The following picture illustrates the traffic exchanged at NSPIXP-2, one of the major Internet exchange points in Japan. The statistics is obtained by polling `ifInOctets` mib variable every 15 minutes. (every 30 minutes before April 21, 1997, and every 5 minutes after Jan 14, 2000). The figures below are not automatically updated, however, they are updated once a month.

Following graph indicates the growth of aggre-

gated inbound traffic at NSPIXP-2 from its scratch.

All of the Digital Gigaswitches was retired by Aug 25, thus graph for Gigaswitch portion does not indicate any traffic. We also retired a couple of Cabletron's SSR8600 on Apr 22. Then we have a Foundry BigIron8000, and a Foundry BigIron15000 connected together.

The aggregated traffic plotted using logscale (y-axis only) is shown below. The greenline is a approximation by $Traffic = b * \exp(a * year)$, where a is 0.657. Don't you think it fits very well to the observed peak aggregated inbound traffic at NSPIXP2? Assuming the growth trend continues, the project peak traffic will reach 5 Gbps in November 2001, 7 Gbps in May 2002, 10 Gbps in December 2002, up to 20 Gbps in December 2003 (roughly to say it doubles in each year).

If we plot the same data for last 3 months with x-axis folded as 24hour, we get the following. You can see that the peak of traffic is at after-lunch and mid-night. Latter is thanks to the NTT's "telehodai" program, in which the contracted sub-

W I D E P R O J E C T 2 0 0 2

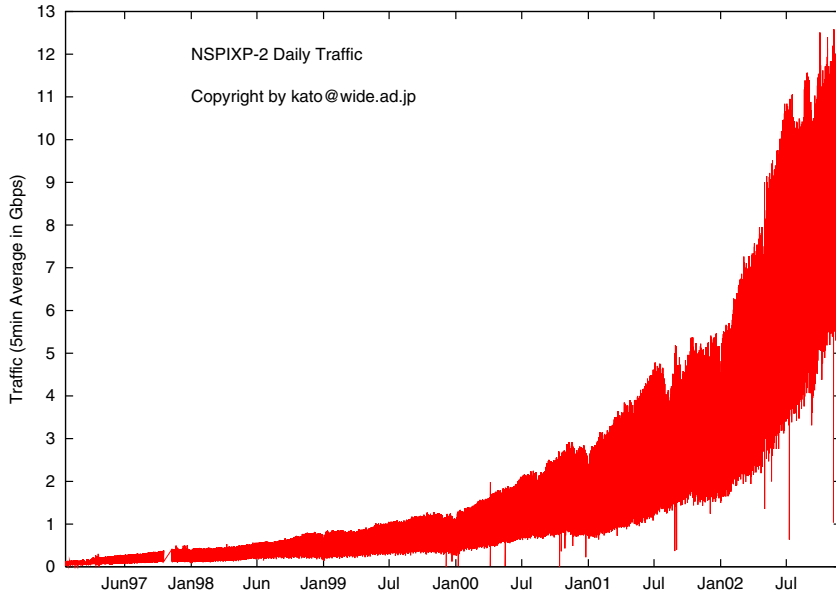


Fig. 6.1.

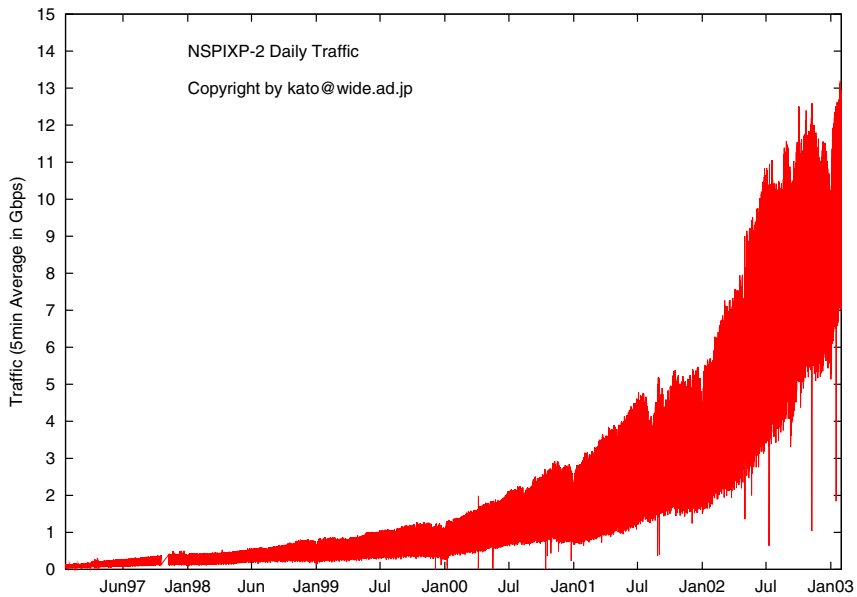


Fig. 6.2.

scribers are not charged for dialup calls to specified numbers (a user can specify up to 2 numbers) between 23:00 and 8:00 localtime. In 2002, broadband services including Cable TV based service and ADSL based service have been widely deployed in Japan. Those services provide not only large bandwidth but also at a flat-rate so people are able to access the Internet at any time.

The figure blow plots the hourly traffic averaged over a month for 6 months. There is no discontinuous point at 11pm due to the “tele-hodai” program. The reason for this is that “tele-hodai” program users had decreased, because a lot of users had changed their program from “tele-hodai” to flat-rate ADSL services. The traffic growth at 11pm is stable at about 300 Mbps regardless of the to-

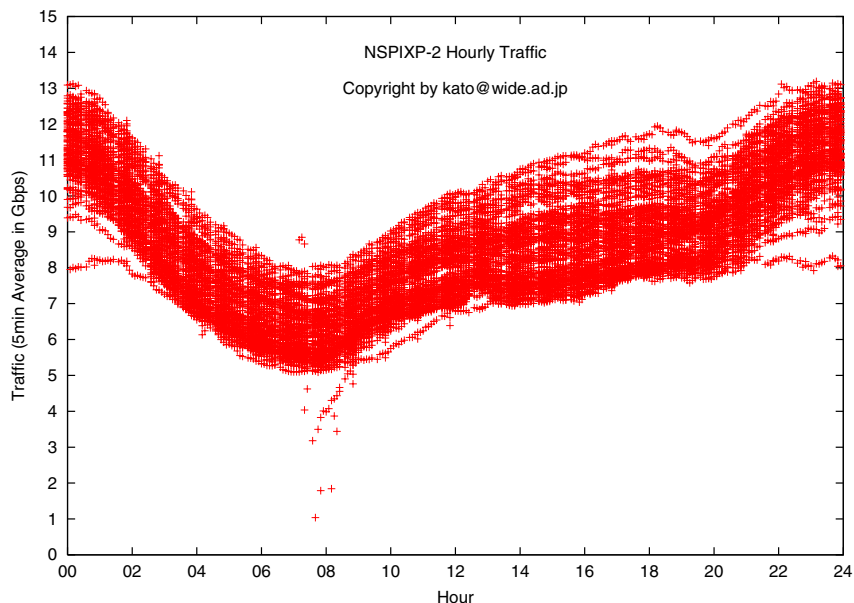


Fig. 6.3.

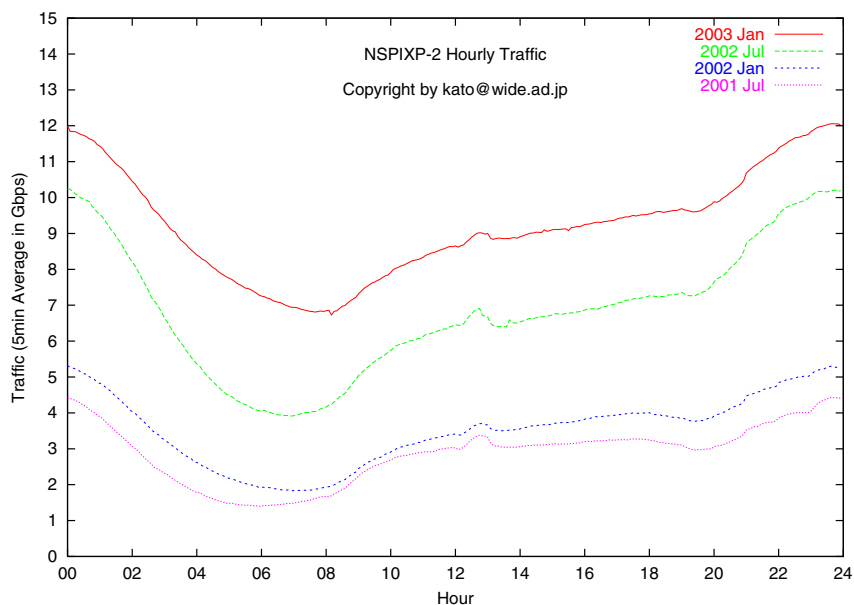


Fig. 6.4.

tal traffic. It is considered that the traffic growth after 8pm dues to the Internet access from home, and mainly dues to these flat-rate services. Generally, The traffic change during a day is equalized. This is caused by deploying flat-rate ADSL services.

第 7 章 イベント

この章では、IX(NSPIXP) の役割を示すために、NetWorld+Interop 2002 東京、第 54 回 IETF 横浜および Global IPv6 Summit in Japan に関して報告する。

7.1 NetWorld+Interop 2002 東京

NetWorld+Interop 2002 東京 は 2002 年 7 月 3 日から 7 月 5 日にかけて幕張メッセで開催され、各出展者および関係施設に提供されるネットワークとして、ShowNet がある。ShowNet は独自のアドレスブロックおよび AS を持ち、全世界をショーとともに回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、各 ISP との間に peer/transit を行う。NSPIXP は ShowNet の対外接続において、中心的役目を果たしており、NTT 大手町ビルにおいて NSPIXP-2 と NSPIXP-6 に接続を行った。

- NSPIXP-2
ShowNet の 2 台のルータから、NSPIXP-2 に対して、10 GE-LR、1 GE-LX の 2 種類のメディアを利用し、接続を行った。単一ユーザ (ISP) としては、世界発の 10 GE を用いた IX の接続を提供する事となった。NSPIXP-2 における ShowNet への peer/transit 組織数は 26 を数えた。
- NSPIXP-6
ShowNet の 2 台のルータから、NSPIXP-2 に対して、100Base-Tx を利用し、接続を行った。NSPIXP-6 における ShowNet への peer/transit 組織数は 21 を数えた。

7.2 第 54 回 IETF 横浜

第 54 回目の IETF は、7 月 14 日から 7 月 19 日にかけて、パシフィコ横浜で開催された。富士通がスポンサーを務め、WIDE プロジェクトがホストを行った。

NetWorld+Interop と同様に、第 54 回 IETF 横浜のネットワークは、来場者および WG の発表のために使用するため、JPNIC に臨時のアドレスブロック

および AS を申請し、独立した ISP として NSPIXP および東京での主要な IX に接続を行った。

NSPIXP は第 54 回 IETF 横浜ネットワークの対外接続において、中心的役目を果たしており、KDDI 大手町ビルにおいて NSPIXP-2 と NSPIXP-6 に接続を行った。

- NSPIXP-2
ShowNet の 2 台のルータから、NSPIXP-2 に対して、10 GE-LR、1 GE-LX の 2 種類のメディアを利用し、接続を行った。NSPIXP-2 における ShowNet への peer/transit 組織数は 39 を数えた。
- NSPIXP-6
ShowNet の 2 台のルータから、NSPIXP-2 に対して、100Base-Tx を利用し、接続を行った。NSPIXP-6 における ShowNet への peer/transit 組織数は 19 を数えた。

7.3 Global IPv6 Summit

2002 年 12 月に、Global IPv6 Summit in Japan をパシフィコ横浜で開催した。実行委員会が中心となり、IPv6 マルチキャストを用いて IPv6 インターネットに向けて放送を行った。

7.4 IPv6 マルチキャスト

IIJ、NTT コミュニケーションズ、NEC、WIDE プロジェクトが NSPIXP-6 を使い、KDDI、フリービット、早稲田大学、マイクロソフト、日立製作所、IPv6 普及・高度化推進協議会の協力の元、本会場の映像と音声 を 500 kbps のストリームを配信した。