

第 3 部

商用インターネット相互接続実験

第 1 章

はじめに

本章では、WIDE プロジェクトが行っている商用インターネットの相互接続に関する実験 NSPIXP について述べる。NSPIXP 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実証的な手法を用いて研究を行なっているプロジェクトであり、現在 20 の商用ネットワークと、JPNIC、APNIC、EXPO、そして WIDE インターネットが WNOC-TYO 内に設置された拠点で相互接続され、商用インターネットが相互に接続された場合の問題点の洗い出し、そして解決方法の研究を行っている。

1.1 背景

日本国内における商用ネットワークサービスプロバイダは、1993 年に 2 社がサービスを開始してから、この 3 年間で急速にその数は増加し、1996 年現在で大小合わせて 200 社を越えるプロバイダがサービスを行っている。

インターネットは、独自に運用されているネットワークが相互に接続され、全体が構成されている。このため、何らかの方法でネットワーク同士が相互に接続されなければ、エンド-エンドのサービスを提供することができない。これは、今までの電話等にみられる通信サービスのように、大手通信事業社が日本全国にサービスを展開し、利用者にとっては、エンド-エンドのサービスが一社によって提供されるような形態とは異なり、どちらかといえば、鉄道網における利用形態に似ている。鉄道網の利用者は、複数の事業者が提供しているサービスを、駅で乗り換えながら目的地まで辿りつく。インターネットも複数のプロバイダが提供しているネットワークを渡り歩くことによって、エンド-エンドのサービスが提供されているのである。

インターネットにおけるネットワーク相互接続の方法は、世界中のインターネットでいろいろな試みが行われてきた。

アメリカでは、NSFNET があったころには、NSFNET を中心に、地域ネットワークが NSFNET に接続され、各地域ネットワークに組織のネットワークが接続される形態が試みされた。また、1995 年からは、NAP と呼ばれるネットワークを相互に接続するためのポイントと NAP 間を相互接続するためのバックボーン、そしてネットワークサービスプロバイダ（以下 NSP と呼ぶ）が NAP へ接続するといった形態が試みられている。この 2

つの試みは、共に全米各地を相互に接続するためのバックボーンがまず最初であり、そして長距離系バックボーンに地域または組織のネットワークが接続されるといった、バックボーン主体の相互接続形態である。

これに対して、CIX プロジェクトや MAE-EAST/WEST プロジェクト、そして SWAB プロジェクトなどは、各 NSP をどのようにして相互に接続するかを検討しているプロジェクトである。この2つの大きな違いは、前者は、バックボーン主体の相互接続であるのに対して、後者は、NSP を相互に接続するためのポイントの構成方法に主眼がある点である。

ネットワークを相互に接続する場合、相互接続に介在するネットワークが存在すると、介在したネットワークの運用ポリシーに影響を受けてしまう。従って、商用インターネットを相互に接続する場合、介在するネットワークが存在しない、相互接続ポイントが必要となってくる。これは、CIX プロジェクトが始まった経緯を見ることから理解することができる。CIX プロジェクトは、当時 NSFNET がインターネットのバックボーンとして機能していた時代、多くのネットワークがこの NSFNET を経由して相互接続されていた。しかし、NSFNET が国の科学技術の振興に対する資金を用いて運用されていたため、その利用目的が研究目的といった利用に制限されていた。このような制限が無い相互接続を行うための相互接続ポイントとして CIX が設立された。

1993 年から日本国内でも商用インターネットのサービスが開始された。これに伴い WIDE プロジェクトでは、国内の NSP の相互接続が整然と行われ、また技術的な視点に立って相互接続の問題を研究するためのプロジェクトを発足させた。これが、NSPIXP である。NSPIXP とは、Network Service Provider Internet eXchange Project の略である。NSPIXP は、その研究方法として、実証的な研究を行うため、実際に NSP を相互に接続するためのポイントを用意し、このポイントを用いて研究を行うことにした。このポイントのことを、NSPIXP-1(Network Service Provider Internet eXchange Point 1) と呼ぶ。現在この NSPIXP-1 は、WNOC-TYO(WIDE Network Operation Center - TOKYO) 内に設置されている。

第 2 章

NSPIXP-1

2.1 Phase I

NSPIXP は、1994 年 4 月から Phase I がスタートし、1995 年 3 月に Phase I が終了した。Phase I では、実証的な研究基盤の構築と、相互接続における問題点の洗い出しが主な研究課題であった。

2.1.1 独立な相互接続ポイントの構築

NSP を相互に接続するためのポイントは、AUP 的に各 NSP から独立でなければならない。WIDE プロジェクトが行っている NSPIXP プロジェクトであっても、WIDE インターネットの AUP で運用されるポイントであってはならない。従って、NSPIXP-1 の設置場所としては、WNOC-TYO 内としたが、NSPIXP-1 は、WIDE インターネットとは、独立なセグメントとして構成し、このセグメントは、WIDE の AUP で制約される範囲外とした (図 2.1)。

2.1.2 レイヤ 3 とレイヤ 2 相互接続モデル

ネットワーク相互接続ポイントのモデルには、レイヤ 3 相互接続モデルとレイヤ 2 相互接続モデルがある。2 つのモデルを図 2.2 に示す。

この 2 つのモデルは、NSP 間を相互接続するポイントのレイヤを示している。レイヤ 3 モデルでは、NSP 間をレイヤ 3 すなわち IP ルータで相互接続をするのに対して、レイヤ 2 モデルでは、NSP 間をレイヤ 2、すなわちデータリンクレベルで相互接続する。たとえば CIX は、レイヤ 3 モデルで相互接続を行っており、MAE-EAST/WEST は、レイヤ 2 モデルである。

この 2 つのモデルの簡単な比較を表 2.1 に示す。

レイヤ 3 モデルにおける相互接続ポイントは、ルータとなる。従って、経路制御的には、1 つの AS (Autonomous System) となるり、何らかの AUP が存在することになる。これに対して、レイヤ 2 モデルにおける相互接続ポイントは、データリンク層なので、その上を流れているパケットの IP レベルでの処理は行わない。単なる伝送媒体を共有しているポイ

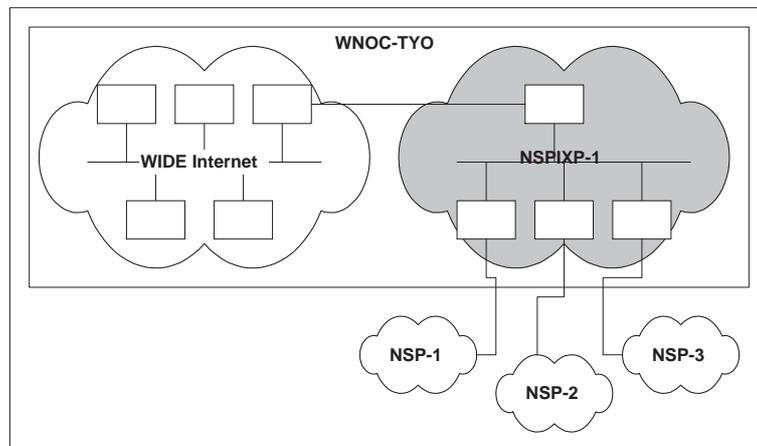


図 2.1: WNNOC-TYO 内に設置された NSPIXP-1

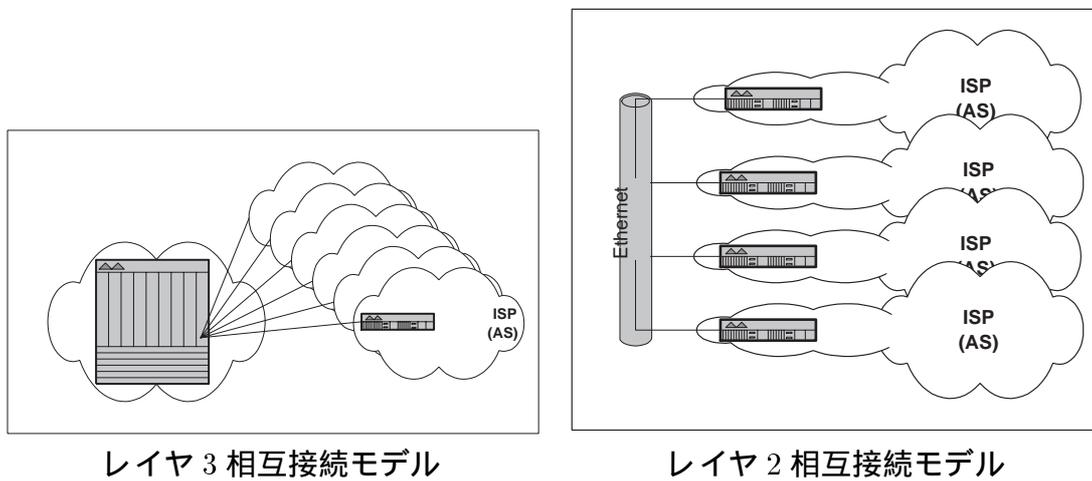


図 2.2: 2つの相互接続モデル

ントがあるだけとなり、IP レベルでのポリシーは、各 NSP が管理運用しているルータ間で処理される。従って相互接続ポイントとしては、ポリシーに加担しない。

ポリシーの面からはレイヤ 2 相互接続モデルは、非常に単純であるが、実際の運用面から見てみると、各ルータの運用管理は非常に複雑となる。レイヤ 3 モデルでは、新しい NSP が接続される場合、単に中心となるルータと新しく設置されたルータ間で新たな BGP パスを張り、経路情報を交換すれば、ほぼ自動的に各 NSP のルータに対しては、新しい経路情報が伝播される。しかし、レイヤ 2 モデルの場合、新しい NSP が接続されると、すべての NSP のルータがこの新しいルータとの間で BGP パスを設定する必要がある。経路制御的な運用の視点からは、レイヤ 3 相互接続モデルでは、1 対 N の運用管理となり、N の増減に対する作業量は、 $O(N)$ となるのに対して、レイヤ 2 相互接続モデルでは、N 対 N の運用管理となり、N の増減に対する作業量は、 $O(N^2)$ となる。

NSPIXP では、当初各 NSP の技術レベルのアンバランスなどを考慮し、レイヤ 3 相互接続モデルに従った相互接続を行っていた。しかし、レイヤ 3 接続の場合、各 NSP は、NSPIXP の運用に対して従属的な運用になり、今後の NSP の複雑な相互接続に対する発展性の無いことなどの理由から、1994 年 10 月よりレイヤ 2 相互接続モデルに従った相互接続に変更した。

2.2 Phase II

1995 年 4 月から 1997 年 3 月を Phase II として、引き続き研究を続けている。NSPIXP 1 は、現在 20 の NSP、JPNIC と APNIC そして、EXPO および WIDE インターネットが相互に接続されている。ルータとしては、ほとんどが CISCO2500 シリーズを用い、Etherswitch を用いてデータリンクレベルでの相互接続を行っている。各ルータは、各 NSP の手により運用されている。各 NSP の AS 番号と各ネットワークへの接続帯域を表 2.2 に示す。

2.2.1 ルーティングレジストリ, ルートサーバ

Phase II では、レイヤ 2 相互接続モデルにおける最大の問題である経路制御の運用技術が主な研究課題である。レイヤ 2 相互接続モデルにおける運用作業量は、前にも述べたように $O(N^2)$ となる。現在相互接続を行っている NSP の数は 20 であり、現在もこの数は増えつつある。すでに各ルータでの BGP パスの設定は、非常に複雑なものとなってきているが、今後の規模の増加に対し、現状の方法では安定した運用は不可能に近い。そこで、Phase II では、主にルーティングレジストリ・ルートサーバの研究を行っている。

	レイヤ 3 相互接続モデ	レイヤ 2 相互接続モデル
接続媒体	ルータ	データリンク
IX の AUP	IP レベルでのポリシーが存在する	IP レベルでのポリシーは存在しない
BGP パス	1 対 N	N 対 N

表 2.1: 2 つの相互接続モデルの比較

IP Addr on IX	AS #	Net	DSU	Router
202.249.3.33	2500	WIDE	N.A.	C4500
202.249.3.34	2497	IIJ	TTNET/T1	C2501
202.249.3.35	2915	SPIN	TTNET/T1	C2501
202.249.3.36	2510	WEB	TTNET/T1	C4000
202.249.3.37	2518	MESH	NTT/T1	IP45/620
202.249.3.38	2521	TOKYONET	NTT/T1	C2501
202.249.3.39	2519	NIS	NTT/T1	C2501
202.249.3.40	2527	SINFONY	NTT/T1	C2501
202.249.3.41	2514	SPHERE	NTT/T1	C2501
202.249.3.42	4197	GOL	NTT/T1	C2501
202.249.3.43	2526	SPACE	NTT/768	C4000
202.249.3.44	4681	MAJIC	NTT/T1	C2501
202.249.3.45	4685	ASAHI-NET	NTT/T1	C2514
202.249.3.46	2515	JPNIC	NTT/512	C2514
202.249.3.47	4686	BEKKOAME	NTT/T1	C2501
202.249.3.48	4689	GSL	Ttnet/T1	C7010
202.249.3.49	4691	DTI	NTT/T1	C2501
202.249.3.50	4690	EXPO	NTT/T3cell	C7010
202.249.3.51	4673	InterVia	NTT/T1	C4500
202.249.3.52	4677	PTOP	NTT/T1	C2501
202.249.3.53	2554	PSI	NTT/768	C2503
202.249.3.54	4608	APNIC	NTT/T1	C2501
202.249.3.55	2524	ASN-RIM	NTT/T1	C2501
202.249.3.56	4672	ASN-IBMNet	Ttnet/T1	C2501

表 2.2: NSPIXP-1 との接続ネットワーク ('95 11 月 13 日現在)

第 3 章

NSPIXP-2/3

NSPIXP-1 では、実験上の性格から、国内の経路のみを扱うと同時に、共同研究者はすべて経路情報の 1 次的な相互交換を行うことなど、相互接続に関する制限を設けて共同研究を行ってきた。このような制限によって、IX の基本的な実験環境が整備されるとともに、国内のインターネットにおける基本的な相互接続環境を提供することができた。しかし、インターネットの利用者の増加やアプリケーションの多様化に伴い、より運用に即した技術の研究開発を行う必要が生じてきている。そこで、1996 年に新に 2 つの相互接続実験ポイントを新設し研究を行うことを提案する。

3.1 実験計画

新に設置する相互接続実験ポイントでは、

1. 経路情報を交換するための BGP の Peering は、各共同研究者間の合意により設定する。
2. 交換する経路情報は、特に国内の経路のみといった制限は設けず、Peering した共同研究者間で合意した経路情報の交換を行う。

3.2 相互接続実験ポイント 予定場所

1. NSPIXP-2: 東京 大手町 KDD ビル
2. NSPIXP-3: 大阪 (検討中)

ネットワーク相互接続ポイントの実験場所として、この 2 ポイントを選んだ理由は、

1. この 2 実験ポイントとも通信事業者が提供するハウジングサービスを利用し、電源・空調設備が完備され、運用上セキュアな場所である。
2. ネットワークの相互接続のために必要な機器を収容するための十分なスペースが確保できる。

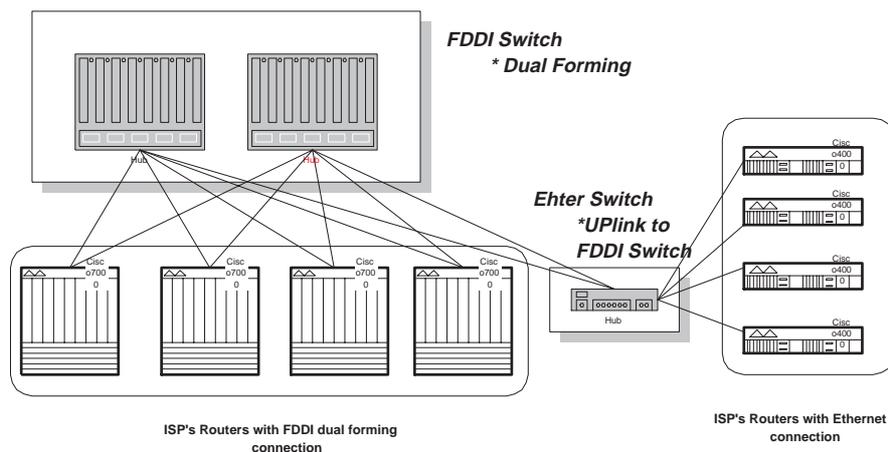


図 3.1: NSPIXP-2 の構成

3. 相互接続ポイントを地理的に分散することにより、安定した相互接続を可能にする。
 などといった理由による。

3.3 相互接続アーキテクチャ

レイヤ2 コネクティビティは、FDDI と Ethernet を提供し、可能な限り 2 重系による信頼性向上を図る様設計する。基本的な構成を図 3.1に示す。

3.4 NSPIXP-1 との関係

NSPIXP は、商用インターネットの相互接続に関する研究を行っていく研究プロジェクトであるが、実質的には、国内のインターネットの良好な相互接続環境も提供している。NSPIXP-1 では、各 ISP の一次的な経路情報を相互に交換するといった制約を設けているため、このような良好なコネクティビティを提供することが可能であった。しかし、NSPIXP-2/3 における経路制御には、このような制約を設けないため、NSPIXP-2/3 だけの接続を許すと、今まで培ってきた国内の良好なコネクティビティを不安定なものにする可能性がある。そこで暫定的 1997 年 3 月末日までは、NSPIXP-2/3 へ接続を行う場合には、NSPIXP-1 への参加も必要条件とすることとした。

第 4 章

おわりに

NSPIXP は、現在 Phase II として研究を行なっている。本稿で述べたような技術的な課題に関しては、技術のサーベや各種ソフトウェアの開発を現在行なっている。また、国内の通信事業者間の相互接続に対する行政のあり方や国際的な相互接続関連についても、共同研究者各位と議論を続けている。

現在 NSPIXP は、多くの商用インターネットサービスプロバイダの方々と共同研究を行ない、少なくとも Phase II 完了までには、多くの成果を期待できると確信している。また、将来のインターネットの相互接続のあり方についても、議論を行なって行きたいと考えている。

