

次世代 インターネット エクスチェンジの 技術動向

INTEC W&G

中川郁夫 ikuo@intec.co.jp

麗澤大学

林 英輔 hayashi@reitaku-u.ac.jp

RIIS

高橋 徹 toru@riis.ad.jp

東京大学

江崎 浩 hiroschi@wide.ad.jp

はじめに

近年のインターネットでは、プロバイダ (ISP: Internet Service Provider) 間の相互接続を効率的に行うためにインターネットエクスチェンジ (IX: Internet eXchange) を構築・運用することが増えている。インターネットでは多数のISPが非常に複雑に相互接続を行っている。IXはISP間の相互接続を実現する1つの技術であり、ISP間のトラフィック交換を効率的に行うことを可能とする。現在では世界に数百ものIXが存在し⁶⁾、IXはISP間の膨大なトラフィック交換を支える重要な通信基盤の役割を果たして

いる。

本稿では、インターネットの基盤技術の1つとして、現在と次世代のIXの実現・運用技術に関する解説を行う。前半では相互接続技術としてのIXの概要と現状のIX技術について解説し、その問題点について触れる。また、後半では仮想的なネットワークによる相互接続環境、データリンク層に非依存な相互接続技術、広域分散型のIX、および光スイッチ技術を用いたIXなど、次世代IXのアーキテクチャと要素技術に関する最新動向を解説する。

IXの機能と構成

本稿では、まず最初にIXの概要について解説する。インターネットでは多数のISPが複雑に相互接続 (ピアリング) を行っている。相互接続を実現する手段の1つは、相互接続を行おうとする2つのISP間で単独で回線を準備しトラフィックの交換を行う方法で、これはプライベートピアリングと呼ばれる。これに対してIXを用いた相互接続では、多数のISPが集まる「場所」としてIXを構築し、多数のISP間での相互接続 (パブリックピアリング) を実現する。

以下では、プライベートピアリングとIXを比較しながらIXの機能、および構成について解説する。また、後述のIXの実現技術を方向づけるIXのポリシモデルについても述べる。

プライベートピアリング

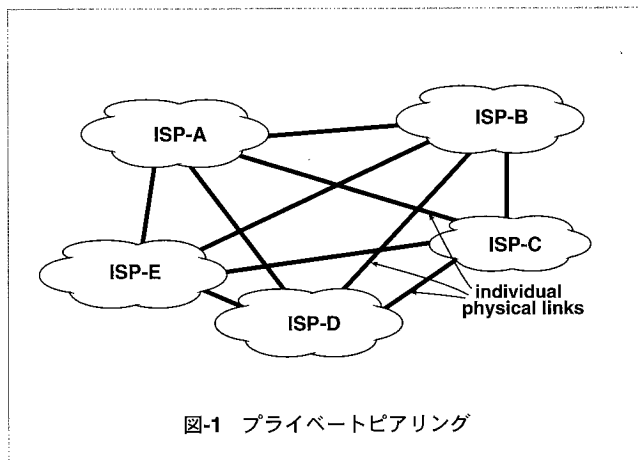
相互接続を行おうとする2つのISP間に専用回線を準備し直接的に相互接続を行う仕組みをプライベートピアリングと呼ぶ。プライベートピアリングは完全に他から独立した環境で2つのISP間だけの相互接続を行うため、物理的な構成やトラフィック制御などの面で自由度が高い。

図-1に複数のISPがプライベートピアリングにより相互接続を行っている例を示す。プライベートピアリングでは1つのISPは相互接続先のISPごとに専用の回線を準備する。本手法によりISP間の完全なメッシュ状の相互接続環境を実現するためには、ISPの数をNとして、全体で $N \times (N-1) / 2$ 本の回線を準備する必要がある。すなわち全体の回線コストは $O(N^2)$ にも及ぶ。インターネットではISPの数は数十万ともいわれており、プライベートピアリングでISP間の相互接続を展開することは拡張性に欠ける。

IXの基本モデル

IXはプライベートピアリングに比較して効率的にISP間の相互接続環境を実現する。IXは相互接続の「場所」を提供し、各ISPが専用の回線をその場所、すなわちIXに引き込むことにより全体の回線コストを抑える。一方、IX内では各ISP間の相互接続を行うことができるため、機能的に前述のプライベートピアリングによる完全メッシュの相互接続と同等の環境を実現できる。

図-2にIXによる相互接続環境の例を示す。図のように、1つのISPは単一の回線をIXに引き込むことにより他のすべてのISPとの相互接続が可能である。回線コストは全体でも $O(N)$ に抑えられることになり、この意味でIXは効率的であり拡張性に優れている。ISPにとっては、相互接続用の回線を1本にまとめることによりトラフィック集約が可能であるという意味でもIXが効果的であることが多い。複数のトラフィックを集約することは、全体の回線利用率を向上させ、回線コストの削減につながる。



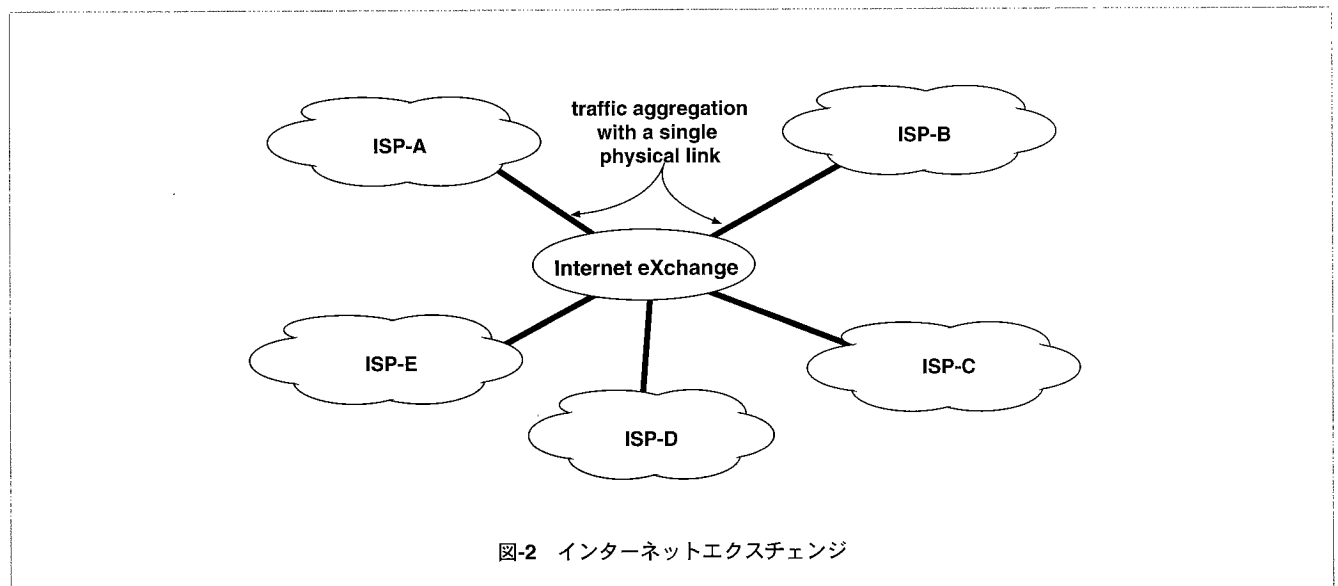
IXのポリシモデル

IXにおけるポリシモデルは、IXの実現方法に直結するIXの重要な技術要件である。インターネットでは、ISP間の相互接続におけるトラフィック制御はコストやビジネス構造などの面からもきわめて重要な意味を持つ¹⁾。IXはISP間の相互接続環境を提供しており、ISP間の相互接続ポリシの面で重要な役割を果たす。

インターネットにおける相互接続では、ISPは経路情報にポリシを反映させることによりトラフィックを制御する。現在のインターネットではISP間の経路制御はBGP4 (Border Gateway Protocol version 4)²⁾で行う。BGP4では授受する経路情報を柔軟に操作することが可能であり、ISPは授受する経路情報を適切に制御することによりポリシをトラフィックに反映させる。たとえば、あるISPから受け取るトラフィックは、そのISPにBGP4で広告する経路情報により操作し、同様に送り出すトラフィックはBGP4で受け取る経路情報を操作することにより制御する。このようなトラフィック制御、およびポリシ制御の仕組みはIXでの相互接続でも同様である。

接続ISPのポリシを各ISPごとに個別に制御することを可能にするため、現在のIXはバイラテラル (bilateral) と呼ばれるポリシモデルを採用している。バイラテラルの相互接続環境では、接続ISPはピアリングを行うISPごとに完全に独立に経路を制御することが可能であり、IXはISP間の経路制御について何ら関与しない。

図-3はバイラテラルによるポリシ制御の例を表している。図はIX上で3つの相互接続が行われていることを示している。また、USER-XはISP-CとISP-Dの2つのISPに接続を行っており、ISP-C、ISP-DはそれぞれUSER-Xの経路をIX上で広告している。すなわち、IXにおいてはUSER-Xの経路は異なる複数の経路が存在する。仮にIXが経路制御



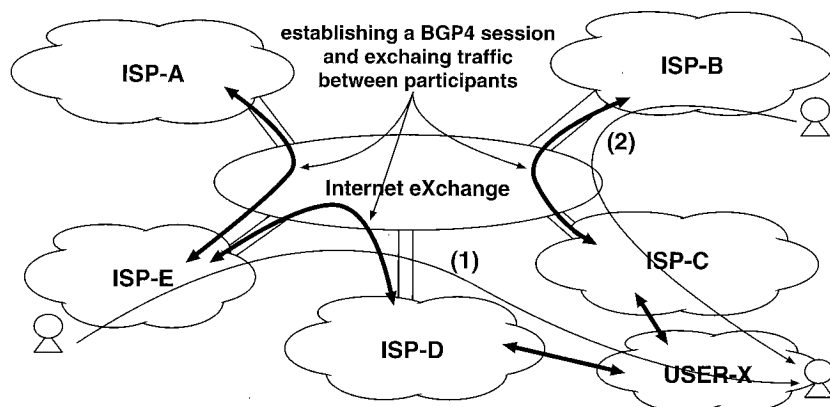


図-3 バイラテラルポリシモデル

に参与する場合はIX内で経路選択が行われ、IXを経由する通信ではUSER-Xへの経路はISP-C経由、もしくはISP-D経由のいずれかに決定される。一方バイラテラルによるポリシモデルでは、図に示すようにIXに接続する各ISPごとに通信経路を選択することが可能であり、たとえばISP-EのユーザはISP-D経由で、またISP-BのユーザはISP-C経由でUSER-Xと通信することが可能である。

バイラテラルの相互接続環境を実現するため、一般にIXではISP間にデータリンク層の接続性を提供する。IXをルータ等のネットワーク層で動作する機器で構築すると、前述の例のように同一のアドレス空間に対する複数の独立した経路情報がIX上で交換された場合に単一の経路選択を強要され、複数の経路を独立に扱うことができなくなる。そのため、現在の技術ではIXをデータリンク層の機器で構成することが必要になる。

現在のIX技術

現在インターネットで実際に運用されているIXは主として2つのIX技術に分類される。1つはLAN (Local Area Network) の技術を用いたもの、もう一方はATM (Asynchronous Transfer Mode) の技術を用いたものである。ここでは、これらのIX技術について基本的な仕組みを紹介し、それぞれの問題点について触れる。

LAN技術によるIX

現在、多くのIXではイーサネットなどのLANの技術を用いて相互接続環境を実現している。これらのIXはLANスイッチ等のLAN機器を準備し、接続ISPはそれぞれの

ルータを同LANスイッチに接続することにより、LAN上で他のISPのルータとの接続性を確保する。本稿ではこれらのIXを「LAN-IX」と呼ぶ。PAIX⁷⁾、NYIIX⁹⁾、AMS-IX¹⁰⁾、NSPIXP2¹²⁾、JPIX¹³⁾等多くのIXが本手法により相互接続環境を提供している。

図-4はLAN-IXの基本的な仕組みを示したものである。LANスイッチは複数の機器により構成されることも多いが、論理的には1つのLANセグメントとみなすことができる。IXに接続を行うISPはそれぞれのルータ(境界ルータ)を準備し、提供されるLANインタフェースでLANスイッチに接続する。

LAN-IXでは、相互接続を行うLANセグメントはすべての接続ISPの共通サブネットとして機能する。共通サブネットはある特定のネットワークアドレスを持ち、IX運用者は接続ISPに対してサブネット内のIPアドレスを割り当てる。各ISPは割り当てられたアドレスを用いて隣接ルータとBGP4による経路情報の交換を行い、トラフィックを交換する。

LAN-IXの問題点

前述のように共通のサブネットを用いて相互接続を行う場合、データリンク層の設定や隣接ルータとの相互接続の設定は非常に簡単であり、接続ISPにとっては容易に相互接続が可能になる。一方で、LAN-IXは以下のような問題点も抱えている。

- スイッチングスピード

現在のインターネットでは大規模なISPのバックボーン

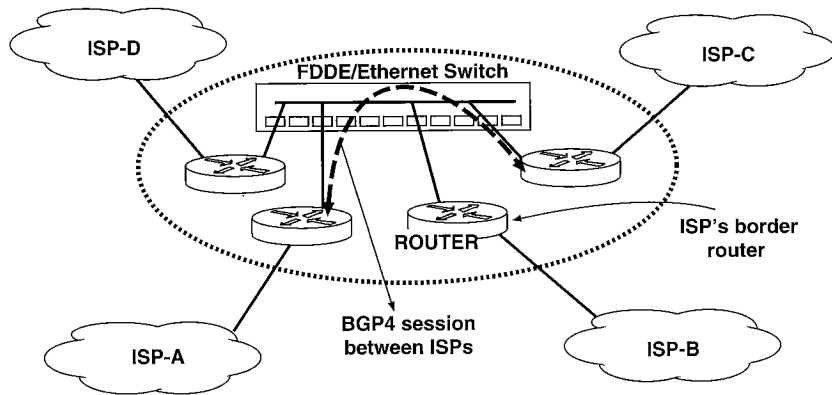


図-4 LAN技術を用いたIX

ンは10Gbps (OC-192) で構成される。一方、LAN技術ではギガビットイーサネットを用いても1Gbps程度までしかスループットはなく大規模ISP間のトラフィックの交換には不十分である。

●セキュリティ

LAN-IXではすべてのルータが共通のサブネットに接続するため、相互接続をしていないルータからもパケットを送出される可能性がある。たとえば、IXに接続するあるISPのルータがBGP4に無関係に他のルータを次ルータ (next-hop) として設定することによりトラフィックを意図的に流すことが可能である。これは thirdparty next-hop と呼ばれ、LAN-IXの大きな問題の1つとされている。

●ルータの増設

LANスイッチに接続するためのファイバあるいはケーブルの物理長が制限される。そのため、接続ISPはIXと物理的に同じ場所にルータを増設することが多い。境界ルータの増設は運用コストの増大につながり、結果的に1つのISPが多くのIXに接続することが難しくなる。

●拡張性

LAN-IXでは共通サブネットの規模により拡張性が制限される。また、ネットワークアドレスの拡張はアドレスやネットマスクの変更を伴うため、相互接続の一時的な停止など、品質の問題にもつながる。

ATM技術によるIX

現在のインターネットではATM技術を用いたIXも多く運用されている。ATM技術を用いたIXはATMスイッチ、もしくはATMスイッチからなるATM網を構築し、接続ISPはATMインタフェースを持ったルータでIXに接続する。IXは接続ISPのルータ間にPVC (Permanent Virtual Circuit) と呼

ばれる仮想的な回線を設定し、ISP間の相互接続を実現する。本稿では本技術によるIXを「ATM-IX」と呼ぶ。MAE-ATM⁸⁾ や StarTap¹¹⁾ などは同手法によるIXを構築している。

図-5はATM-IXの実現例を示している。この図ではISP-A～ISP-DがIXに接続している。いずれのISPもATM対応のルータを用いてIXにATMレベルで接続を行う。図ではISP-AとISP-Cの相互接続の様子を示している。IXはこれらのルータの間にPVCを設定し、ISP-AとISP-Cのそれぞれのルータ間で仮想的な回線を利用できるようにしている。ここで、両端ISPは仮想的な回線をポイントツーポイントの論理的に独立な回線としてみなすことができ、同回線上でBGP4による経路情報の交換、およびトラフィックの交換を行う。

ATM-IXの問題点

TCP/IPの通信をATM上で行うことは転送速度やオーバーヘッドなどのいくつかの問題点があるとされている。ATM-IXでも同様にATMを用いることによる問題点が指摘されている。以下では、ATM-IXでの問題点について述べる。

●オーバーヘッド

ATM上でTCP/IPの通信を行う場合、パケットを分割するためのオーバーヘッドが大きな問題になる。ATMではパケットを48オクテットのデータと5オクテットのヘッダから構成される固定長のセル (cell) に分割するため、最低でも9.4%の無駄が生じる。特に高速な通信ではオーバーヘッドは15%にもなることがあり、回線の利用率の面で

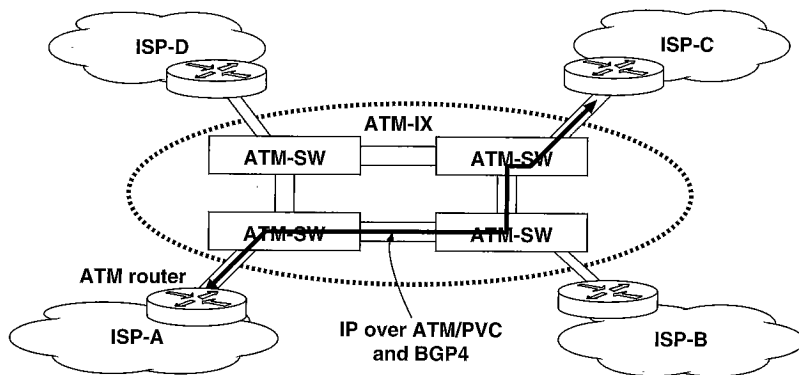


図-5 ATM技術を用いたIX

大きな問題になる。

- スイッチングスピード

ATMではセル単位で通信を行うため、転送表の検索やパケットの分割、再構成の処理能力などに起因してスイッチング速度、通信速度が問題になる。

- 運用コストおよび拡張性

ATM-IXではPVCの設定に伴う運用コスト、および拡張性の低下が問題とされる。

ATM-IXにおけるISP間のPVCの設定は、一般に完全な自動化はできていない。ISPの数をNとした場合のPVCの数は最大で $O(N^2)$ に及び、特にATM-IXが複数のATMスイッチなどで構成される場合は運用コスト、および拡張性が問題になる。

次世代IXの技術

前述のような現在のIX技術における問題点を解決し、ISP間のさらに大規模なトラフィック交換を実現するために、各地で次世代IX技術の研究開発が展開されている。次世代IX技術では大容量のトラフィック交換を実現するための処理能力の向上や、仮想的なプライベートピアリングの実現、広域分散型のIXなどさまざまな新機能についても研究されている。以下では、次世代IX技術の最新動向について解説する。

VLAN技術を用いたIX

前述のLAN-IXを提供するIX事業者の一部では、既存の仕組みを拡張し新しい機能を提供するための仕組みとしてVLAN (Virtual Local Area Network) を利用することを検討している。VLANはLAN上に仮想的に複数の独立したサブ

ネットを実現する技術で、1つの物理インタフェースで複数のネットワークを同時に利用することができる。たとえば、最近の多くのLAN-IXはギガビットイーサネットスイッチで構成しているが、これらのIXではVLAN技術を導入することにより、既存の構成を保ったまま容易に仮想的なネットワークを増やすことが可能である。以下ではVLAN技術を用いたIXを「VLAN-IX」と呼ぶ。

図-6にVLAN-IXの概念図を示す。図のようにVLAN-IXの物理構成はLAN-IXと同じになる。VLAN-IXではIXを構成するLANスイッチでVLANにより仮想的に独立なサブネットを設定することにより、たとえば(1)や(2)のようにISP間に個々の独立した相互接続環境を提供する。

前述のようにLAN-IXでは第三者によるトラフィックの送受信が可能であるなど、セキュリティ上の問題が指摘されているが、VLAN-IXでは相互接続ごとに仮想的に独立したサブネットを設定することにより仮想的なプライベートピアリングを可能にし、安全な相互接続ができる。

MPLSを応用したIX

一部の研究者の間では、MPLS (Multi-Protocol Label Switching)³⁾ 技術をIXに応用することにより、データリンク層に非依存な相互接続環境を実現しようとする試みが行われている。MPLSはIPパケットに固定長のラベルを付加することにより網の内部で柔軟なトラフィック制御を行うための技術である。MPLSはデータリンク層、およびネットワーク層に非依存に設計されており、MPLS網を介して、異なるデータリンクメディアで接続するルータ間に仮想的なパスを確立することができる。MPLS技術をIXに応用することによりデータリンク層に非依存な相互接続環境の実現を可能にする。以下ではMPLS技術を

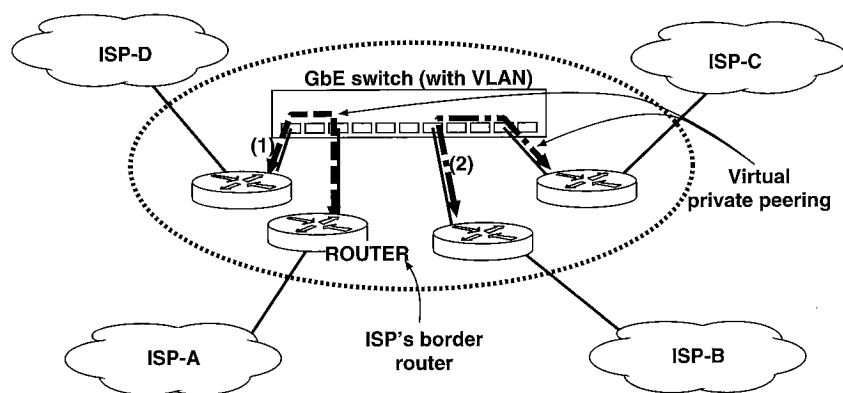


図-6 VLAN技術を応用したIX

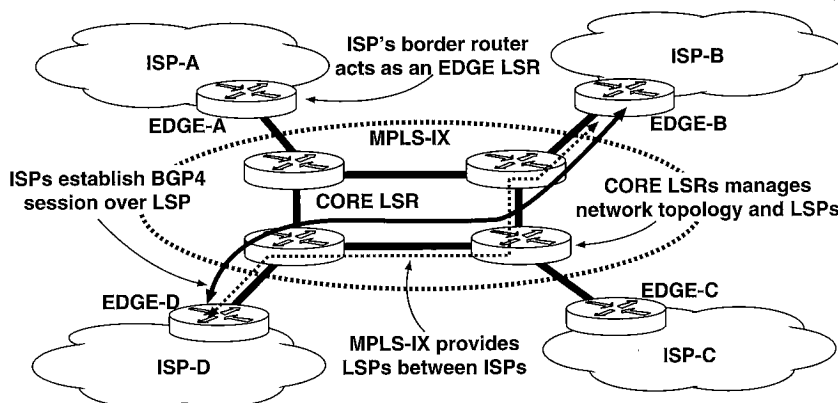


図-7 MPLSによる相互接続網

応用したIXを「MPLS-IX」と呼ぶ。

図-7にMPLS-IXの実現例を示す。IX事業者はコアLSR (Label Switched Router) と呼ばれるMPLS網の制御を行うルーターでIXを構成する。接続ISPはエッジLSRと呼ばれるMPLS網の境界に位置するルーターをいずれかのコアLSRに接続する。図-7ではISP-BとISP-Dの各ルーター間に仮想パスを確立し、同パス上でBGP4による経路情報の交換、およびトラフィックの交換を実現している。

MPLS技術をIXに応用する最大の利点はデータリンク層に非依存な相互接続環境を実現することである。既存のIX技術はLAN-IX、ATM-IXともにデータリンク層に強く依存した仕組みであり、接続ISPはデータリンクメディアを選ぶことはできない。一方、MPLSはデータリンク層に非依存であり、MPLS-IXではイーサネット、ATM、POS (Packet over Sonet) など任意のデータリンク層を用いて相互接続が可能である。たとえば、現在事実上最大の通信

速度を提供するOC-192 POSを用いることによりIX上で10Gbpsの通信速度を実現することも可能である。

また、MPLS-IXでは広域分散環境において相互接続環境を実現することも可能である。前述のとおり、MPLS-IXではコアLSRの網としてIXを構成する。コアLSR間は長距離回線などを用いることにより広域分散配置が可能であり、また、接続ISPについても、エッジLSRが接続を行うデータリンクメディアを選択することにより、IXと物理的に同じ場所にルーターを増設することなく、柔軟なIX接続が可能になる。

光スイッチを用いたIX

長期的な視点からは、さらに大容量のトラフィック交換を可能にするため光スイッチを用いたIX技術の研究が進められている。光ファイバ上で直接IP通信を行うIPO (IP over Optical) 技術や、DWDM (Dense Wavelength Division Mul-

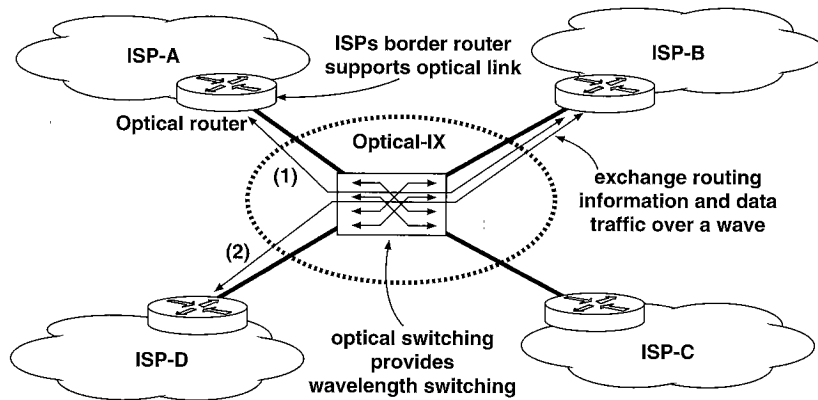


図-8 光スイッチによるIX

tiplexer)などの光多重技術を組み合わせることにより現在の1,000倍もの通信速度を実現することが期待されている。同技術をIXに応用したものが「Optical-IX」で、同技術ではISP間の相互接続における膨大なトラフィック交換を実現する。

図-8はOptical-IXの概念を示している。Optical-IXは光スイッチ等の装置で構成され、接続ISPは光多重、および光通信技術に対応したルータで光スイッチに接続を行う。光スイッチとルータ間は光ファイバで直結されており、光多重技術により通信ごとに波長(wavelength)を使いわけ、Optical-IXではGMPLS (Generalized MPLS)⁴⁾あるいはOBGP (Optical BGP)⁵⁾と呼ばれるプロトコルを用いてISPのルータ間に波長を割り当てることにより相互接続環境を提供する。すなわち、相互接続を行うISPのルータ間に必要に応じて波長を提供し、波長上でIPの通信路を実現することによりBGP4による経路制御、およびトラフィック交換を可能にする。

Optical-IXの最大の特徴は、IXを光スイッチおよび光ファイバ網として構成することにより、膨大なトラフィックを交換可能にすることである。Optical-IXでは、光多重や波長技術により現在の数百倍から数千倍の通信速度を実現することが可能といわれている。

むすび

本稿では技術的な観点から現状、および次世代のIX技術について解説を行った。現状のIX技術としてはLAN技術を用いたIX、およびATM技術を用いたIXを紹介し、それぞれの仕組み、および問題点について触れた。また、次

世代のIX技術としてVLANを用いたIX、MPLS技術を応用したIX、光スイッチによるIXについて、仕組みおよび特徴について解説した。

現在、インターネットでは数百ともいわれるIXが構築・運用され、ISP間の相互接続を実現する場として、IXは非常に重要な役割を果たすようになってきた。このような背景を考慮すると、今後のIX技術、IXモデルはインターネットの発展を支えるためにきわめて重要な課題であるといえる。たとえば、インターネットバックボーンが80波長から構成され、各波長が10Gbps (OC-192)の通信速度を持つ時代を考えたときに、IXでのトラフィック交換の性能、処理能力がそれに耐えることができるものでなければならない。一方で、インターネットが今後の最も重要な通信基盤であることを考えると、通信速度の問題ばかりではなく、ISP間の通信における信頼性や安定性、安全性など、IXにおける通信の品質を確保するための技術などもきわめて重要な課題である。

謝辞 本稿の執筆にあたってご協力をいただいた世界中のIXの技術者、および研究者の皆様に感謝します。

参考文献

- 1) Huston, G.: Interconnection, Peering and Settlements, The Internet Protocol Journal, Vol.2, No.1 (Mar. 1999).
- 2) Rekhter, Y. and Li, T.: A Border Gateway Protocol 4, IETF RFC1771 (Mar. 1995).
- 3) Rosen, E., Viswanathan, A. and Callon, R.: Multiprotocol Label Switching Architecture, IETF Internet-Draft (Apr. 1999).
- 4) Ashwood-Smith, P. and Banerjee, A. et al.: Generalized MPLS - Signaling Functional Description, IETF Internet-Draft (Sep. 2001).
- 5) Blanchet, M., Parent, F. and St-Arnaud, B.: Optical BGP: InterAS Lightpath Provisioning, IETF Internet-Draft (Aug. 2001).
- 6) Manning, B.: Exchange Point Information, <http://www.ep.net/>
- 7) PAIX: Palo Alto Internet eXchange, <http://www.paix.net/>
- 8) MCI WorldCom: MAE Information, <http://www.mae.net/>
- 9) Telehouse: NYIIX, <http://www.nyix.net/>
- 10) Amsterdam IX: Amsterdam IX, <http://www.ams-ix.net/>
- 11) Star Tap: Star Tap, <http://www.startap.net/>
- 12) WIDE Project: NSPIX, <http://jungle.sfc.wide.ad.jp/NSPIX/>
- 13) JPPIX: JaPan Internet eXchange, <http://www.jpix.ad.jp/>

(平成13年5月23日受付)

