

MPLSを用いた広域分散IXの実証実験

中川郁夫 (株) インテック・ネットコア
ikuo@inetcore.com

江崎 浩 東京大学
hiroshi@wide.ad.jp

菊池 豊 高知工科大学
yu@kikuken.org

永見健一 (株) インテック・ネットコア
nagami@inetcore.com

はじめに

本稿では、IX (Internet eXchange) に MPLS (Multi-Protocol Label Switching)¹⁾ 技術を適用した MPLS-IX アーキテクチャ²⁾ の技術解説を行うとともに、JGN (Japan Gigabit Network) 上で展開されている、広域分散 IX の実証実験についての報告を行う。

IX は自律的に運用されている複数のネットワーク同士を相互接続する仕組みである。IX には多数の ISP (Internet Service Provider) が接続し、お互いの経路情報の交換を行うことにより「ピアリング」、すなわち相互接続を行っている³⁾。現在のインターネットでは数百にも及ぶ IX が運用され⁴⁾、ISP 間のトラフィック交換を実現する上で、IX はきわめて重要な役割を果たしている。世界的な IX としては、PAIX や MAE、LINX などが有名である。国内では NSPIX2 や JPIX、JPNAP などの大規模な IX のほか、地域を中心とした TRIX、OKIX、TOYAMA-IX、など数々の地域 IX が構築され、運用されている。

著者らは次世代の IX のアーキテクチャとして MPLS-IX の研究・開発を進めている。MPLS-IX アーキテクチャでは、IX に接続する ISP 間に MPLS を用いて仮想的なパス (通信路) を確立し、その上で相互接続を行う。このため、ISP はデータリンク層に非依存な環境で相互接続を行うことができる。また、同アーキテクチャでは、仮想パスを実現するための制御ネットワークを IP ネットワークとして実現する。このことは、IX 全体のトポロジを柔軟に組むことを可能にし、広域分散環境での相互接続を実現する。

本稿では、前半で一般的な IX の技術、および著者が提案している MPLS-IX アーキテクチャについて述べ

る。ここでは、まず IX の仕組みについて述べ、既存の IX 技術の特徴について触れる。次に、MPLS-IX アーキテクチャについて、その実現技術や特徴について解説する。

また、後半では MPLS-IX アーキテクチャを用いた相互接続の実証実験について紹介する。著者らは次世代 IX 研究会を設立し、同研究会において MPLS-IX の研究開発と検証試験を行っている。同研究会では、MPLS-IX アーキテクチャを用いた相互接続の実証実験を JGN 上で行っている。本稿では、次世代 IX 研究会の概要について紹介するとともに、同研究会を中心とした MPLS-IX の実証実験についても報告を行う。特に、本稿では、JGN を用いて実施している実証実験について、その仕組みや接続構成について述べる。

IX- Internet eXchange

本章では、IX (Internet eXchange) の仕組みと技術について述べる。ここでは、まず、IX の特徴を明確にするため、プライベートピアリングと IX の仕組みについて述べ、その後、現在多くの IX で用いられている技術として、LAN (Local Area Network) 技術を用いるもの、および ATM (Asynchronous Transfer Mode) 技術を用いるものについて述べる。なお、本稿ではそれぞれの技術の特徴についてのみ述べる。詳細については参考文献 5) を参照されたい。

● プライベートピアリングと IX

ISP が相互接続を行う場合、何らかの手段で物理的な接続を行い、BGP4 (Border Gateway Protocol version 4)

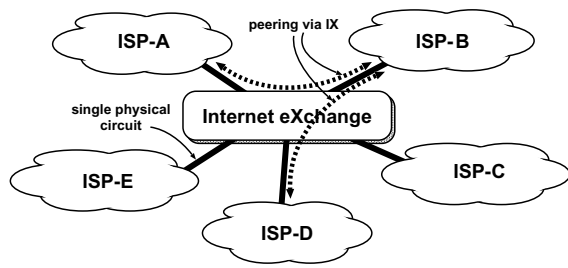


図-1 IXの仕組み

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。この際，相互接続に用いる物理的な回線の形態により，相互接続の方法はプライベートピアリング，およびIXに分類される。

プライベートピアリングでは，相互接続を行おうとする2つのISP間に専用の回線を準備して直接的に相互接続を行う。プライベートピアリングは他から独立した環境で2つのISP間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。反面，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ことになる。

プライベートピアリングにより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数は $O(N^2)$ にもおよび，明らかに拡張性に欠ける。

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」を提供し，各ISPはIXに接続するための回線を引き込む。IX内では各ISP間の相互接続を行うことができ，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-1)。このため，IX上での相互接続はパブリックピアリングとも呼ばれる。

IXを用いてISP間の相互接続を実現する場合，回線数は全体で $O(N)$ であり，この意味で，プライベートピアリングと比較して効率的で拡張性に優れている。また，1つのISPからみた場合，単一の物理回線で複数のISPと相互接続することによりトラフィックを集約できるため，回線コストを抑えることが可能である。

● LAN技術を用いたIX

現在，最も多くのIXではLANの技術を用いている。本稿では同技術を用いたIXを「LAN-IX」と呼ぶ。図-2はLAN-IXの基本的な仕組みを示したものである。

IXに接続を行うISPはそれぞれのルータ(境界ルータ)をLANメディアで接続する。最近では，IX側でギガビットイーサネットなどのLANスイッチを提供するケースが多い。LANスイッチは論理的に1つのLANセグメントとみなされ，IXに接続するすべてのルータに共通のサブネットとして機能する。各ISPのルータは同サブネット上でBGP4による経路情報の交換を行う。

● ATM技術を用いたIX

現在のインターネットでは，ATM技術を用いているIXも多い。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網から構成される。接続ISPはATMインタフェースを持ったルータを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。

ATM-IXでは，ISPのルータ間に確立されたPVCはポイントツーポイントの仮想的な回線としてみなすことができる。各ルータは仮想回線上でBGP4による経路制御を行うとともに，トラフィック交換を行う。

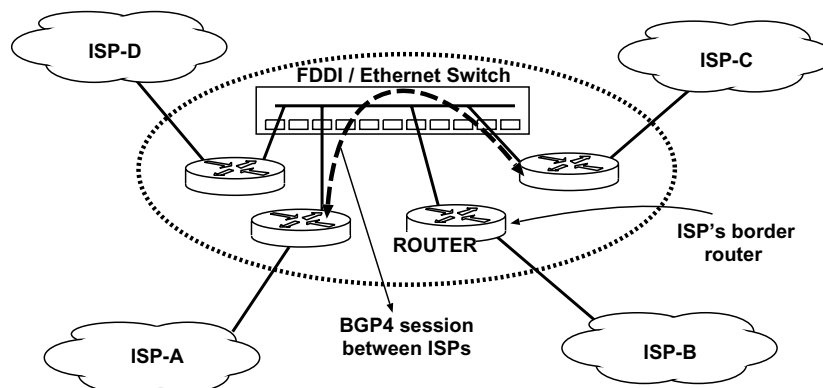


図-2 LAN-IXの実現例

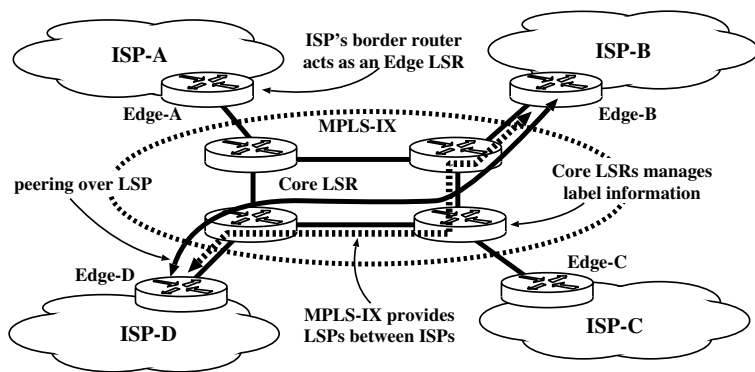


図-3 MPLS-IXの概要

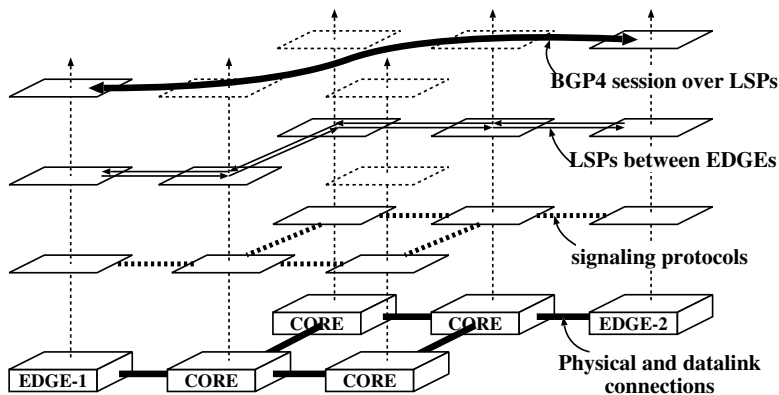


図-4 MPLS-IXのアーキテクチャ

用いることが特徴である。

図-3にMPLS-IXの概念図を示す。IX事業者はCore LSRと呼ばれるMPLS網の制御を行うルータでIXを構成する。IXに接続するISPはEdge LSRと呼ばれるMPLS網の境界に位置するルータをいずれかのCore LSRに接続する。各Edge LSRはMPLSの機能を用いて他のISPルータとの間にLSPを確立し、このLSPを用いてピアリング、すなわち相互接続を行う。

図-4は、トラフィック交換を行う際の、より具体的な仕組みについて表している。図のように、IXに接続するISPのルータ(Edge LSR)間では、MPLSのシグナリングプロトコルを用いてLSPを確立する。シグナリングプロトコルとしては、RSVP-TE (Resource reSerVation Protocol for Traffic Engineering)⁶⁾、およびLDP (Label Distribution Protocol)⁷⁾が利用できる。なお、シグナリングプロトコル、およびその動作に関する詳細については参考文献2)を参照されたい。

シグナリングプロトコルを用いて確立されたLSP上では、BGP4による経路情報の交換を行う。各ISPは受け取った経路情報に従い、LSPを通してデータトラフィックを他のISPに転送することによって、実データトラフィックの交換を実現する。

MPLS-IX

著者らはMPLS技術をIXに応用した、次世代のIXアーキテクチャ「MPLS-IX」を提案している²⁾。本章では、MPLS-IXの仕組み、および特徴について述べる。

● MPLS-IXの概念

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、従来のIP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網はLSR (Label Switching Router)と呼ばれるルータのネットワークとして構築される。LSR間にはデータを運ぶための仮想的なパスとしてLSP (Label Switched Path)を確立する。実データを網内で転送する際は、パケットにラベルが付与され、あらかじめ確立されたLSPに沿って転送される。

MPLS-IXでは、IXに接続するISPの境界ルータ間でLSPを確立することにより、LSP上で相互接続を実現する。一般に、MPLSは単一の管理ドメイン、すなわち1つのISPの内部で利用されることを前提としているが、MPLS-IXでは、複数の異なるISP間でMPLSの機能を

● MPLS-IXの特徴

MPLS技術をIXに応用する主な利点は2つある。1つは、MPLSを用いることにより、データリンク層に非依存な相互接続環境を実現できることである。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、異なるデータリンクメディアで接続するルータ間でも、LSP、すなわち、仮想的なパス上でデータを交換することができる。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。

現在のインターネットにおいて、ほとんどのIXで利用されている技術はLANやATMなどのデータリンク層に強く依存した仕組みであり、接続しようとするISPはデータリンクメディアを自由に選ぶことはできない。一方、MPLS-IXはデータリンク層に非依存であり、たとえば、現在事実上最大の通信速度を提供するOC-768 POSを用いることにより、IX上で40Gbpsの通信速度を実現することも可能である。

MPLS 技術を IX に応用するもう 1 つの利点は、広域分散環境での相互接続を実現できることである。MPLS-IX では制御ネットワーク、すなわち、Core LSR 間の接続や Core LSR と Edge LSR 間の接続において、通常の IP 技術を用いている。したがって、MPLS-IX はそのトポロジを柔軟に構成することが可能であり、広域分散環境への適用や、冗長構成の実現など、拡張性の面で特に優れている。

広域分散 IX の実証実験

著者らは、MPLS-IX アーキテクチャを用いた広域分散 IX 技術についての実証的な研究活動を行うことを目的として次世代 IX 研究会を設立した⁸⁾。同研究会では、MPLS-IX アーキテクチャを用いた相互接続技術の確立を目指すとともに、JGN 上で同技術を用いた広域分散 IX の実証実験を展開している。本章では次世代 IX 研究会の活動について紹介し、その実証実験の状況について報告する。

●次世代 IX 研究会

次世代 IX 研究会は 2001 年 10 月に設立された。同研究会は MPLS-IX アーキテクチャを用いた広域分散 IX 技術を確立するため、実証的な研究活動を推進している。同研究会への参加はオープンであり、通信事業者、プロバイダ、ルータベンダ、大学・研究機関などを中心に、2002 年 9 月現在で 68 組織が参加している。

現在、研究会では以下の 3 つのワーキンググループを設置し、それぞれのテーマに従って研究を進めている。

ルータ相互接続ワーキンググループ

ルータ相互接続ワーキンググループでは、MPLS-IX アーキテクチャを用いた相互接続をマルチベンダ環境で実現するため、MPLS ルータの相互接続性の研究、および検証試験を行っている。本ワーキンググループでは、MPLS-IX を実現するために MPLS ルータに求められる機能についての仕様書（ホワイトペーパー）を策定し、MPLS ルータベンダらとの意見交換を行うとともに、多数の MPLS ルータベンダの参加を受けて、マルチベンダ環境における、ルータの相互接続試験を実施している。

本ワーキンググループでは、以下のような相互接続試験を実施し、その結果を公開している。

- 第 1 回相互接続試験（2001 年 10 月 15 日～19 日）
TAO 大手町 IPv6 システム運用技術開発センター（東京）
10 社の MPLS ルータベンダが参加、
MPLS-IX アーキテクチャに必要な基本機能に関して

の検証を行った。

- 第 2 回相互接続試験（2002 年 1 月 28 日～2 月 1 日）
TAO 幕張ギガビットリサーチセンター（千葉）
12 社の MPLS ルータベンダが参加、
基本機能に加え、冗長性の確保やルータの性能などについても検証を行った。
- 第 3 回相互接続試験（2002 年 5 月 15 日～5 月 17 日）
（株）ネットマックス品質管理センター（東京）
12 社の MPLS ルータベンダが参加、
Core の機能試験、性能試験などを行った。
- Networld+Interop Tokyo 2002（2002 年 7 月 1 日～7 月 5 日）
幕張メッセ（千葉）
イベントにおける相互接続のショーケースを構築した。
- 第 4 回相互接続試験（2002 年 9 月 2 日～9 月 6 日）
TAO 北陸 IT 研究開発支援センター（石川）
10 社の MPLS ルータベンダが参加、
冗長性の試験、IPv6 による相互接続試験等を行った。

IX ユーザワーキンググループ

IX ユーザワーキンググループでは、MPLS-IX に接続する ISP、あるいは CSP（Content Service Provider）などの立場から、接続を行う際に必要になる機能や技術要件について検討し、テストベッド上で相互接続の実証実験を行っている。テストベッドは JGN を用いて広域分散環境に構築され、全国から 20 組織（2002 年 9 月現在）が接続している。実証実験では広域分散環境で複数の ISP が相互接続を行い、実際にトラフィックを交換することにより、MPLS-IX の機能検証を行っている。なお、本ワーキンググループで行っているテストベッド上での実証実験の詳細については後述する。

本ワーキンググループでは、MPLS-IX に接続するための設定、あるいは運用上のノウハウに関する情報共有も進めている。具体的には数種類の MPLS ルータの実装について、設定の仕方、あるいは運用方法をまとめ、研究会で発表の場を設けるとともに、Web 上でドキュメントを公開している。本ワーキンググループでの活動は教育（人材育成）の意味でも有効に機能している。

IX プロバイダワーキンググループ

IX プロバイダワーキンググループでは、IX を提供する事業者の立場から MPLS-IX の機能や技術について議論・検討を行っている。本ワーキンググループは、広域分散 IX を設計・構築・運用しようとする技術者らで構成される。次世代 IX 研究会のテストベッドの構築・運

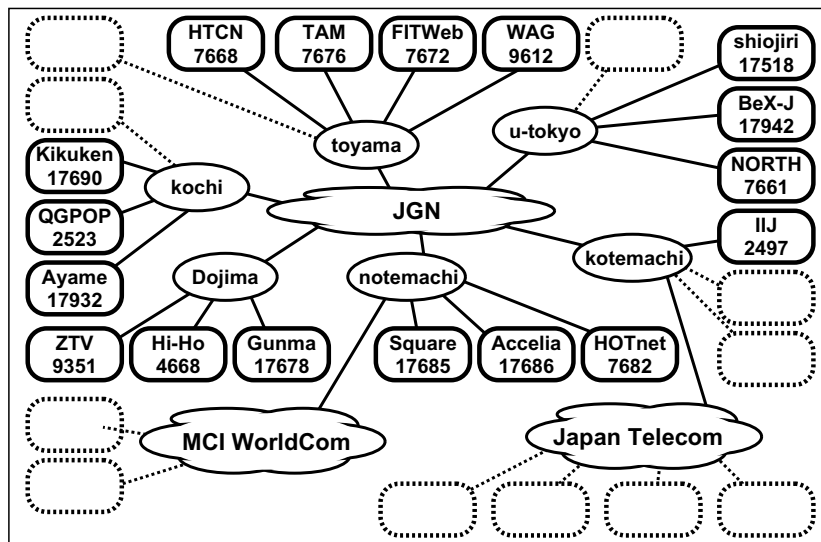


図-5 実証実験網のトポロジ

用を通して、MPLS-IX アーキテクチャによる相互接続環境を提供する際に求められる機能や運用技術の検討、あるいは運用上のノウハウの蓄積を行っている。

本ワーキンググループでは、MPLS-IX アーキテクチャによる相互接続環境を提供しようとする複数の事業者間で相互乗り入れする方法についても検討、実験を行っている。具体的には、次世代IX研究会で実験を行っているJGN上のテストベッドと、他のMPLS-IXの提供者が相互に接続し、複数のMPLS-IXを介してISPやCSPが相互接続するための仕組みについても実験を行っている。2002年9月現在では日本テレコム社、およびMCI WorldCom社がMPLS-IXによる相互接続環境を試験的に提供しており、次世代IX研究会のテストベッドと相互乗り入れを実施している。

●実証実験の概要

本節では、次世代IX研究会で構築・運用を行っているテストベッド、すなわち、MPLS-IX アーキテクチャを用いた広域分散IXの実証実験網について述べる。前述の通り、実証実験はJGN上に構築されており、JGN上の研究プロジェクトの1つである「地域間相互接続実験プロジェクト」に参加する組織を中心として相互接続の実験を行っている。

図-5にテストベッドのトポロジを示す。現在、広域分散IXのテストベッドは6つのCore LSRを持つ。Core LSRにはJuniper社製のM10もしくはM20を用いている。Core LSRは東京(u-tokyo, notemachi, kotemachi)、大阪(dojima)、富山(toyama)、高知(kochi)に分散配置され、その間をJGN上のATM PVCを用いて接続している。他の箱型は本テストベッドに接続しているISPの名前とAS (Autonomous System) 番号を示している。点

線の箱型は今後の予定である。本実証実験では、東京都内に東京大学、NTT 大手町ビル、KDDI 大手町ビルにそれぞれCore LSRを設置している。これは、実証的な相互接続実験を行うため、国内のISPが特に集中しているこれらの通信拠点に接続点を置くことが有効と判断したためである。

図は、他のMPLS-IX提供者との相互乗り入れについても示している。他MPLS-IX提供者との相互接続は、IXプロバイダワーキンググループの研究活動の一環として位置付けられる。前述の通り、現在、日本テレコム社とMCI WorldCom社が試験的にMPLS-IXによる相互接続環境を提供しており、次世代IX研究会のテストベッドと相互乗り入れを行っている。

●実証実験への参加方法

次世代IX研究会で運用を行っている本テストベッドは、実証実験への参加組織を随時募集している。ある組織が実証実験に参加する場合には、研究会に参加申込を行うとともに、これらのCore LSRのどれか1つに接続を行う必要がある。現在、Core LSRへの接続は以下の3つのうちいずれかの方法による。

- (a) JGNを用いて接続を行う。接続組織はJGNに接続し、いずれかのCore LSRまでATM PVCを確立する。
- (b) Core LSRに直接接続する。接続組織はルータを持ち込むか、もしくは回線の引き込むことにより、いずれかのCore LSRに対して直接接続を行う。なお、直接接続を行う場合には、ATM、POS、GbEなどを用いて接続することが可能である。
- (c) 他の事業者経由で接続する。次世代IX研究会のテストベッドと相互乗り入れを行っているMPLS-IX事

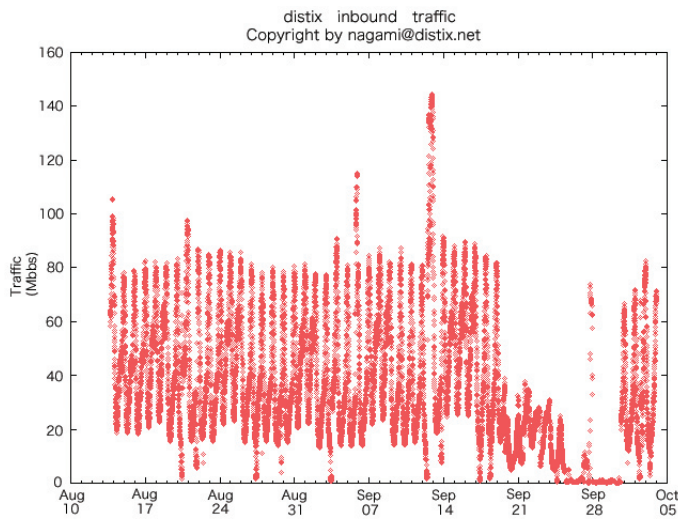


図-6 実証実験網のトラフィック状況

業者に接続することにより、実証実験に参加が可能である。

● 実証実験の状況

現在、次世代IX研究会の実証実験では20組織が相互接続に参加してトラフィック交換を行っている。実証実験には大学や研究機関のほか、ISP、CSPなどの民間企業も接続している。現在は、研究目的を主としているため、特定のアドレス空間での通信や、イベント用のトラフィックなど、実験用のトラフィック交換が中心である。

図-6は実証実験において交換されているトラフィックの合計値を示している。図は、直近の1カ月間について、テストベッドを介して交換されたトラフィックの合計を5分間隔で計測したものである。図に示される通り、ピーク時で100Mbps程度のトラフィック交換が行われていることが分かる。

● MPLS-IXの課題

次世代IX研究会では、これまで広域分散IXとしての基本的な機能の検証を中心に実証実験を行ってきた。これまでの実証実験では、相互接続実験を開始してから約6カ月間、計画的な停止を除いて、相互接続上の深刻な問題は発生していない。このことからMPLS-IXアーキテクチャを用いた広域分散IXの仕組みが非常に安定して動作していることが分かる。

一方、実証実験を通していくつかの細かな課題が指摘されている。以下では、これまでの実証実験で明らかになったMPLS-IXの課題についてまとめる。

(1) MPLSではパケットにラベルを付与するため、網内

のMTU (Maximum Transfer Unit) の設定に注意が必要である。一般に、各LSR間ではMPLSのラベル部分を除いてMTUが1500 Octets以上確保されるよう設定することが推奨される。

- (2) ルータの実装の差異に依存する問題の整理が必要である。多くのMPLSルータはMPLS-IXに必要な基本機能の実装を終えているが、一部のルータでは、たとえば、BGP4のTTL (Time To Live) の値が正しく指定できないケースなどがある。
- (3) MPLS網内での障害検出が難しい。MPLSではIPパケットの追跡が難しく、障害時にどこで問題が発生しているかを特定することが困難である。

おわりに

本稿では、広域分散IXを実現するMPLS-IXアーキテクチャに関する技術解説、および、JGN上で構築・運用されているMPLS-IXの実証実験の紹介を行った。MPLS-IXアーキテクチャを用いた広域分散IXは、新しいインターネットモデルを実現する可能性を秘めている。たとえば、コンテンツ事業者とアクセスプロバイダの直接接続や、地域間での広帯域アプリケーションの利用、大手ISPのトランジットサービスの地域での利用、商用IXの地方展開など、広域分散IX技術への期待は大きい。

次世代IX研究会での実証実験は、広域分散IXの実現という基本機能のフェーズから、今後は、IPv6などの次世代プロトコルへの適用、QoS (Quality of Service) などの品質制御技術の適用など、次のフェーズへ移って行く予定である。

謝辞 本研究の実施にあたって有益なご意見をいただいた林英輔教授、および全国のIX関係者に感謝いたします。また、本研究の一部は通信・放送機構の委託研究「委121-401」に基づいて実施いたしました。

- 1) Rosen, E., Viswanathan, A. and Callon, R.: Multiprotocol Label Switching Architecture, IETF Internet-Draft (Apr. 1999).
- 2) Nakagawa, I., Esaki, H. and Nagami, K.: A Next Generation IX Architecture using MPLS, SAINT2002, Nara (Jan 2002).
- 3) Huston, G.: Interconnection, Peering and Settlements, The Internet Protocol Journal, Vol.2, No.1 (Mar. 1999).
- 4) Manning, B.: Exchange Point Information, <http://www.ep.net/>
- 5) 中川 郁夫, 林 英輔, 高橋 徹, 江崎 浩: 次世代インターネットエクスチェンジの技術動向, 情報処理, Vol.42, No.7 (July 2001).
- 6) Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V. and Swallow, G.: RSVP-TE: Extensions to RSVP for LSP Tunnels, RFC3209 (Dec. 2001).
- 7) Andersson, L., Doolan, P., Feldman, N., Fredette, A. and Thomas, B.: LDP Specification, RFC3036 (Jan. 2001).
- 8) Next Generation IX Consortium: <http://www.distix.net/>
(平成14年10月2日受付)