

## MPLS を用いた広域分散 IX の実現

中川 郁夫<sup>†1</sup> 江崎 浩<sup>†2</sup>  
 菊池 豊<sup>†3</sup> 永見 健一<sup>†4</sup>

著者らは、IX(Internet eXchange)にMPLS(Multi-Protocol Label Switching)を適用した次世代IXアーキテクチャMPLS-IXを提案してきた。IXは複数のISP(Internet Service Provider)間の相互接続を実現し、効率的にトラフィック交換を行うことを目的として運用されている。MPLS-IXでは、IXに接続するISP間にMPLSによる仮想的なパスを確立することにより、データリンク層に依存しない相互接続環境を提供する。また、MPLS-IXでは、柔軟なIXトポロジを構築することが可能である、などの特徴もある。

本稿では、MPLS-IXを用いて階層型IXを実現し、地域IXなどの既存のIXを階層型に接続することにより、広域分散IXを実現する手法について提案する。本稿で提案する手法は、既存のリソースを有効に活用し、既存のIXにおける相互接続に影響を与えることなく、広域分散環境での相互接続を可能にする。

また、著者らは通信放送機構の委託研究を受けてMPLS-IXアーキテクチャの研究を進めるとともに、次世代IX研究会を設立し、広域分散IXの実証実験を実施している。本稿では、同研究会の概要について紹介し、同研究会で実施している実証実験について報告する。

### Implementation of widely distributing IX using MPLS technology

IKUO NAKAGAWA,<sup>†1</sup> HIROSHI ESAKI,<sup>†2</sup> YUTAKA KIKUCHI<sup>†3</sup>  
 and KENICHI NAGAMI<sup>†4</sup>

Authors have proposed a next generation IX architecture, called **MPLS-IX**, in which we apply MPLS (Multi-Protocol Label Switching) technology to an IX. Many IXes exist in the current Internet, so that numerous ISPs interconnect and exchange traffic efficiently, each other. **MPLS-IX** establishes virtual paths between participating ISPs. **MPLS-IX** provides data-link medium independancy for interconnection, and enables an IX to be flexible in consiting IX topology.

In this paper, we propose an implementation of widely distributing IX. Our implementation has hierachical IX architecture based on **MPLS-IX** architecture. Deployment of our distributing IX is easy, because we connect existing IXes in hierachical model, and use existing resouces without any effect to current communications.

We are enforcing the research project of MPLS-IX which is funded by TAO (Telecommunications Advancement Organization of Japan) to develop and test several features of the MPLS-IX architecture. We established the 'Next Generation IX Consortium' based on the research project, and started the experimental IX testbed. We report the current status and major result of the consortium, as well.

#### 1. はじめに

本稿では、IX(Internet eXchange)にMPLS(Multi-Protocol Label Switching)技術を適用したMPLS-

IXアーキテクチャ<sup>24)</sup>を用いて、広域分散環境における階層型IXを実現する手法を提案するとともに、同技術を用いて展開している広域分散IXの実証実験の報告を行う。

IX(Internet eXchange)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IXが接続し、お互いの経路情報の交換を行うことにより、相互接続、すなわち「ピアリング」を行っている。

現在のインターネットでは数百にも及ぶIXが運用され<sup>2)</sup>、ISP間のトラフィック交換を実現するうえで、IX

†1 インテック・ウェブ・アンド・ゲノム・インフォマティクス(株)  
 INTEC Web and Genome Informatics Corporation

†2 東京大学  
 University of Tokyo

†3 高知工科大学  
 Kochi University of Technology

†4 東芝  
 Toshiba

は極めて重要な役割を果たしている。例えば PAIX<sup>3)</sup> や MAE<sup>4)</sup>, LINX<sup>5)</sup> など世界でも最大のトラフィック交換点として位置付けられている。また、国内では NSPIX2<sup>10)</sup> や JPIX<sup>11)</sup> が著明である。一方、世界的に地域 IX の構築を進める活動も進められている。日本でも TRIX<sup>14)</sup>, OKIX<sup>15)</sup>, TOYAMA-IX<sup>16)</sup>, BeX-J<sup>13)</sup> など、数々の地域 IX が構築され、運用されている。

一方、著者らは IX に MPLS (Multi-Protocol Label Switching) 技術を応用した、次世代 IX のアーキテクチャ MPLS-IX に関する研究を進めている<sup>24)</sup>。同技術は、IX に接続する ISP 間に MPLS を用いて仮想パスを確立し、その上で相互接続を行うため、データリンク層に非依存な IX を実現できる。また、MPLS-IX では、仮想的なパスを確立するための制御ネットワークはある種の IP ネットワークであり、IX 全体のトポロジを柔軟に組めることも大きな特徴である。

本稿では、MPLS-IX アーキテクチャを用いて、広域分散環境で階層型の相互接続環境を実現する手法について提案する。近年、インターネットでは次のような理由から、ISP 間の相互接続を広域分散環境で実現するための広域分散 IX の仕組みが求められている。

- コンテンツ提供者と地域のアクセスプロバイダを高品質な環境で相互接続する。
- 地域 IX に接続する大手 ISP が地域の ISP にトランジットサービスを提供する。
- 地域間において、高速・広帯域の通信を行う。
- 商用 IX の地方展開。

本稿で提案する手法では、例えば地域 IX など、既存の IX を階層型に接続し、広域分散環境に階層型の相互接続環境を実現する。同手法を用いることにより、既存のリソースを有効に活用しながら、容易に広域分散型の相互接続環境を実現することが可能になる。

また、著者らは通信放送機構の委託研究を受けて MPLS-IX アーキテクチャの研究を進めるとともに、次世代 IX 研究会を設立し、広域分散 IX の実証実験を実施している。本稿では、同研究会で実施している広域分散環境での相互接続試験の状況についても報告を行う。

なお、本稿では広域分散 IX を実現するための技術的な側面からの研究を対象としている。IX によるビジネスモデル、あるいは社会的な影響は本稿の対象外とする。

本稿では、第 2 節において IX の仕組みと技術について述べる。特に既存の IX 技術として LAN 技術を用いるもの、および ATM 技術を用いるものについて

述べる。

第 3 節では MPLS-IX アーキテクチャの概要について述べる。ここでは、MPLS-IX の基本的な仕組みと、広域分散環境へ適用する際に必要になる特徴について述べる。

第 4 節では MPLS-IX アーキテクチャを用いて、広域分散 IX を実現するための実現手法について提案する。本稿では、特に既存 IX を階層型に接続することを例に、現存の相互接続の仕組みを変えずに、シームレスに広域分散環境へ拡張するための手法について述べる。

第 5 節では、著者らが参加している次世代 IX 研究会での取り組みについて紹介し、同研究会で構築・運用を行っている広域分散 IX の実証実験について報告する。

## 2. IX - Internet eXchange

本節では、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る。また、既存の IX で用いられている技術として、LAN (Local Area Network) スイッチを用いるもの、および ATM (Asynchronous Transfer Mode) スイッチを用いるものについて述べる。

### 2.1 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) 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ふたつの 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ふたつの ISP 間のみ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。半面、ひとつの ISP は相互接続先の ISP ごとに専用の回線を準備する。

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  $O(N^2)$  にも及び、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

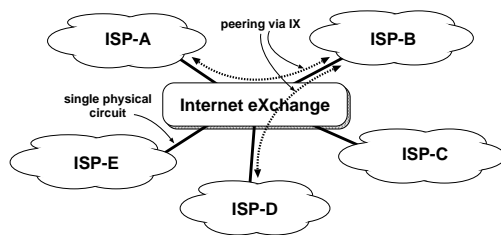


図 1 IX の概念図  
Fig. 1 IX model

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、ひとつ 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。

## 2.2 既存の IX 技術

既存の IX で利用されている技術は、LAN (Local Area Network) の技術を用いるものと、ATM (Asynchronous Transfer Mode) の技術を用いるものに分類される。詳細は参考文献<sup>22)</sup> に記述されているため、ここでは、それぞれの技術の特徴についてのみ述べる。本節で述べる技術は、後述の第 4 節で階層型の IX を実現する手法について述べる際に参照する。

### LAN-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。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 スイッチに接続する。LAN スイッチは論理的にひとつの 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-IX

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 インター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

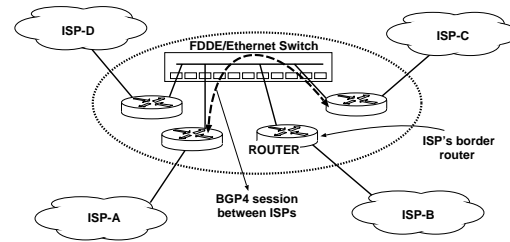


図 2 IX の例  
Fig. 2 Example of IX

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。

## 3. MPLS-IX

著者らは MPLS(Multi-Protocol Label Switching) 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<sup>24)</sup>。本節では、MPLS-IX の仕組み、および特徴について簡単に述べる。本稿で提案する広域分散 IX の構築手法は MPLS-IX に基づいている。

### 3.1 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、仮想的なパスを通して相互接続を実現する。一般に MPLS は単一の管理ドメイン、すなわちひとつの ISP の内部で利用されることが前提とされるが、MPLS-IX では、複数の異なる ISP 間で MPLS の機能を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(Label Switched Router) 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、同 LSP を用いてピア

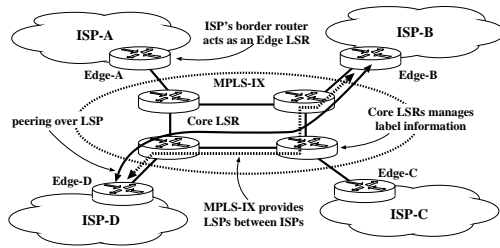


図 3 MPLS-IX の実現例  
Fig. 3 Example of MPLS-IX

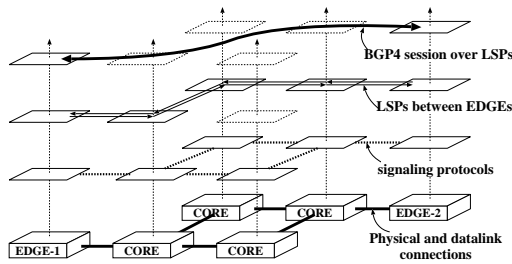


図 4 MPLS-IX の仕組み  
Fig. 4 Architecture of MPLS-IX

リング, すなわち相互接続を行う。

図 4 は, 相互接続を実現するまでの, 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。また, 同 LSP 上で 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したがい, LSP を通じてデータトラフィックを他の ISP に転送する。

### 3.2 MPLS-IX の特徴

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, 仮想的なパス (LSP) 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。

また, MPLS-IX では制御ネットワーク, すなわち Core LSR と Edge LSR を接続するためのネットワークを通常の IP ネットワークの技術を用いて構築可能である。

既存の IX 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 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 はデータリンク層に非依存であり, 例えば, 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。

## 4. 広域分散 IX の実現手法

本節では MPLS-IX アーキテクチャを用いて広域分散型の IX を実現する手法について述べる。以下では, まず, 広域分散 IX に対するニーズについて紹介する。次に, MPLS-IX アーキテクチャを用いて階層型 IX を実現する方法について述べ, 同技術を用いて階層的に既存の IX を接続することにより, 広域分散 IX を実現する手法について提案する。

### 4.1 広域分散 IX の必要性

国内のインターネットでは広域分散 IX の実現が強く求められている。これまで, 国内のインターネットは市場規模の差異に起因して, 東京一極集中型のトポロジを維持したまま成長してきた。このため地域間, もしくは地域と東京の通信は制約が多いことが一般的である。広域分散 IX の技術は, これらの遠距離の通信においても, 通信経路, 帯域, 通信遅延などの設計が容易で, より高速で安定した通信の実現を可能にする。

広域分散 IX 技術は, 次のような利用方法が期待されている。

- コンテンツプロバイダと各地アクセスプロバイダとの直接接続が可能になる。ゲーム<sup>18)</sup> やストリーム<sup>19)</sup> などのコンテンツを広域分散 IX を通じてエンドユーザに直接配信することは高品質なコンテンツの配信で有効である。現在も, 後述の広域分散 IX の実証実験を通して, コンテンツの直接配信に関する取り組みが進められている。
- 映像伝送, ファイル共有など地域間の通信の実現が可能になる。これまで, 国体映像や CATV の映像を地域間で交換する取り組みが実験的に行われている。<sup>21)23)</sup>
- 大手プロバイダが地域でトランジット (IP の接続性を提供する) サービスを提供することが可能になる。これまで, 外資系プロバイダは国内の回線を有していないため, 地域でトランジットサービスを提供することはできなかった。広域分散 IX は, 東京を拠点にするプロバイダが, 東京のユーザと同等に近いサービスを国内全域に提供可能にする。JANOG9 (Japan Network Operators Group) の会議では, 後述の広域分散 IX の実験網を介して AboveNet, ODN の 2 社がトランジット提供の実験を行った。
- 商用 IX の地方展開が可能になる。現在も JPIX<sup>11)</sup>, JPNAP<sup>12)</sup> などの商用 IX が大阪や名古屋に拠点

を増やそうとしている。これらの商用 IX はこれまで「点」の IX として、特定のビル内でのみサービスを提供してきたが、広域分散 IX を用いることにより容易に地方展開が可能になる。

#### 4.2 MPLS-IX による階層型 IX

MPLS-IX では制御ネットワークに階層構造を導入することにより、階層型 IX を構築することが可能である。本稿では、制御ネットワークを以下のふたつに分類する。

- リーフインターフェース
- IX バックボーン

リーフインターフェースは Edge LSR と接続するための、データリンク層のインターフェースを指す。各インターフェースは Core LSR で終端され、POS、ATM あるいは GbE など、データリンクメディアは任意である。各リーフインターフェースは独立したサブネットを構成する。

IX バックボーンは、Core LSR 間のネットワークとして構成する。IX バックボーンはリーフインターフェースの情報を含め、IX 全体の接続情報を管理する。Core LSR 間では OSPF や IS-IS などのリンクステート型の経路制御プロトコルで経路情報を交換し、IX 全体のトポロジ情報を更新する。すなわち、IX バックボーンは複数のリーフインターフェース間でパケットを交換するために経路制御を行う。

IX バックボーンは IP ネットワークの一種である。したがって、IX バックボーンを WAN(Wide Area Network) の技術を利用して広域分散環境で構築することにより、階層型 IX を広域分散環境に適用することができる。

#### 4.3 広域分散 IX の実現手法

以下では、MPLS-IX アーキテクチャを用いて広域分散 IX を実現する手法について述べる。ここでは、前述の階層型 IX の仕組みを利用し、既存の IX を階層的に接続することにより、広域分散環境での相互接続環境を実現する方法について述べる。

階層型 IX では、リーフインターフェースを用いて IX に接続する ISP のルータ(すなわち Edge LSR) と接続する際、次のいずれかの方法を取る。

- 直接接続する
- データリンク層のスイッチを介して接続する

前者は Core LSR と Edge LSR が物理的に直接接続されるケースである。この場合、ひとつのリーフインターフェースをひとつの Edge LSR が占有することになる。

後者は既存の IX を階層型に接続する場合に有用であ

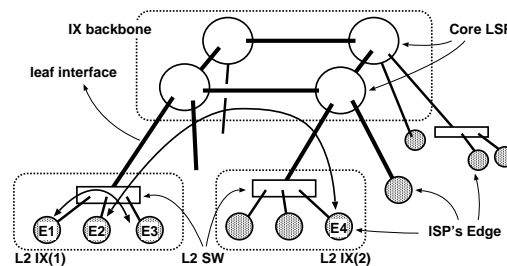


図 5 MPLS-IX による階層構造  
Fig. 5 Hierarchy of MPLS-IX

る。ここでは、既存の IX を LAN-IX もしくは ATM-IX とする。MPLS-IX はデータリンク層に非依存であるため、LAN-IX や ATM-IX 任意のデータリンク層の IX と階層的に接続できる。以下では、LAN-IX と ATM-IX をあわせて L2-IX と呼ぶ。

図 5 に階層型 IX を用いて構成される広域分散 IX を模式的に示している。白丸で表された C1~C4 は IX の Core LSR を表している。Core LSR とこれらの間を接続するネットワークが IX バックボーンを構成している。L2-IX と記述された箱型は既存の IX を表している。図に示すように、L2-IX を構成する L2-SW(イーサネットスイッチや ATM スイッチ) に MPLS-IX のリーフインターフェースを接続することにより、既存の IX を階層型に接続することが可能である。

本稿で示す広域分散 IX では、既存の IX と IX バックボーン提供者との責任分解点を定めることも容易である。図 5 では、リーフインターフェースによって既存の IX と接続しており、制御ネットワークの面でも同インターフェースで分離が可能である。したがって、L2-SW とリーフインターフェースの接続点を責任分解点とし、リーフインターフェースまでを広域分散 IX 提供者、すなわち IX バックボーン提供者の管理・運用範囲とし、L2-SW までを既存の IX 事業者の管理・運用範囲とするのが明解である。

#### 4.4 既存の IX との共存

本稿で提案する広域分散 IX の構築手法において、既存の IX 内での通信は MPLS-IX を用いる場合の通信と共存が可能である。すなわち、既存の IX に対して、広域分散 IX のリーフインターフェースを接続するだけで、既存の環境を維持したまま広域分散 IX 対応にすることができる。このことは、すでにあるリソースを有効に活かすことで広域分散対応へのコストや移行への手順を容易にするばかりではなく、トラフィックの集約効率を上げる意味でも重要な意味を持つ。

図 6 は既存の IX での通信と MPLS-IX を介した

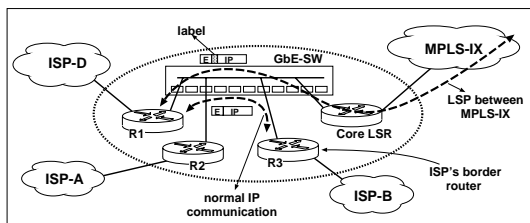


図 6 MPLS-IX による階層構造  
Fig. 6 Hierarchy of MPLS-IX

通信が共存可能なことを LAN-IX を例にして示している。既存の IX 内での通信は、イーサネットスイッチを介して、IP パケット (図で IP と記述されている箱) にイーサネットヘッダ (E と記述されている箱) が付与されて通信を行う。一方、既存の IX に接続する ISP のルータが MPLS-IX を介して通信を行う場合には、IP パケットとイーサネットヘッダの間にラベル情報が挿入されることにより通信を行う。L2-IX で利用されるスイッチでは、通常の IP による通信と、MPLS による通信は互いに干渉しないため、広域分散 IX と階層型に接続しても、既存の IX 上での通信に影響を与えることはない。

#### 4.5 他の手法との比較

以下では、MPLS-IX を用いて広域分散 IX を実現する場合と、それ以外の技術を用いて広域分散 IX を実現しようとする場合について、技術的な面からの比較を行う。

- ATM (Asynchronous Transfer Mode)  
ATM を用いることにより、広域分散環境で ISP 間の相互接続が可能である。ただし、ATM を用いて大規模な広域分散 IX を実現しようとした場合、PVC を設定する運用の手間が非常に大きいことが問題になる。また、他のデータリンク技術との相互接続性はないため、既存の IX の多くが利用している LAN-IX との接続はできない。
- 広域 LAN / 広域イーサネット  
近年、広域 LAN や広域イーサネットなどのサービスが提供され始めている。同技術を用いることにより、広域で LAN-IX を構築することが可能である。しかし、Ethernet によるサブネットを共有するため接続する ISP の数がサブネットの広さに依存する、など拡張性の面で問題がある。また、共通のサブネットを広域分散環境で共有することになり、既存の IX との共存が難しい。また、既存の IX との接続を行った場合でも、運用上、あるいは管理上の責任分解点が定義できない、などの問題がある。さらに、ATM-IX との接続性はないため、

ATM で IX を構築している場合には接続は不可能である。

- IP-VPN (RFC2547)  
RFC2547 は MPLS バックボーン上で VPN (Virtual Private Network) を実現する手段を定義している。同技術はすでに通信事業者がサービスとしても提供しており、企業などが社内網を構築する際に利用している。しかし、同技術は事業者の Core ルータにおいて、ユーザ経路をすべて保持管理する必要があるため、ISP などの多数の経路情報を有するネットワークの相互接続は現実的に不可能である。
- Optical-IX (Canarie)  
光スイッチの技術を応用した Optical-IX が提案されている。光スイッチ技術により高速広帯域による相互接続が可能である。しかし、ルータのインターフェースが DWDM (Dense Wavelength Division Multiplexer) に対応していることを想定しており、各組織が準備するルータが非常に高価なものになる。また、GMPLS や DWDM 対応のルータの実装は数年先といわれており、当面の実現技術としては不適切である。

## 5. 広域分散 IX の実証実験

著者らは、通信・放送機構の委託を受けて研究を進めている次世代広域分散 IX 技術をもとに、次世代 IX 研究会を設立した<sup>17)</sup>。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と、同技術を用いた広域分散 IX の実証実験を展開している。本節では次世代 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。

### 5.1 次世代 IX 研究会

次世代 IX 研究会は MPLS-IX アーキテクチャを用いて広域分散 IX を実証的に実験するための研究活動を行うことを目的として 2001 年 10 月に設立された。同研究会への参加はオープンであり、通信事業者、ISP、ルータベンダ、大学などを中心に、2002 年 3 月現在で 59 組織が参加している。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。

#### 5.1.1 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 を用いた相互接続を実現するため、マルチベンダ環境での MPLS ルータの相互接続性を検証する。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、MPLS ルータベンダらとの意見交換を行っている。

また、相互接続の要件・検証項目の検討を行うとともに、複数の MPLS ルータベンダの参加を受けて、ルータ相互接続試験を実施している。2001 年 10 月 15 日～19 日に行われた第 1 回相互接続試験では 10 社の MPLS ルータベンダが参加、MPLS-IX に必要な基本機能についての検証を行った。2002 年 1 月 28 日～2 月 1 日には第 2 回相互接続試験を実施、同試験では 12 社の MPLS ルータベンダが参加し、基本機能の検証のほか、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。

次世代 IX 研究会では、ウェブでこれらの相互接続試験の結果を公開している。また、本ワーキンググループでは、今後も継続的に相互接続試験を実施することを予定している。

#### 5.1.2 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、IX に接続する ISP、あるいは CSP(Content Service Provider)などの立場から、MPLS-IX に接続する際に必要になるの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(Japan Gigabit Network)を用いて広域分散環境に構築され、全国から 20 にも及ぶ組織が接続している。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、実証的に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、ウェブ上でドキュメントを公開している。これらの活動は教育(Education)の意味でも技術者の情報共有に有効に機能している。

#### 5.1.3 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・運用を通して、MPLS-IX による相互接続環境を提供する際に求められる機能や必要になる技術の関する検討、あるいは運用上のノウハウの蓄積を行っ

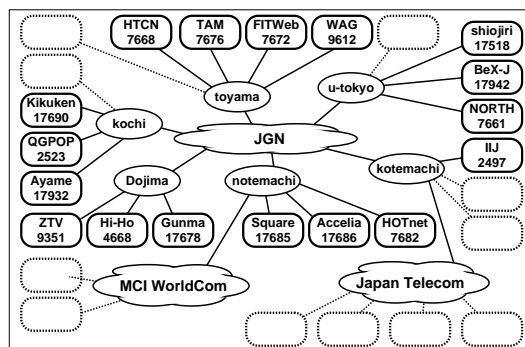


図 7 広域分散 IX のテストベッド  
Fig. 7 Testbed of distributed IX

ている。

本ワーキンググループでは、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テストベッドと他の MPLS-IX の提供者と協力し、複数の事業者を介して 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 3 月現在では Japan Telecom 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。

#### 5.2 実証実験の概要

本節では、次世代 IX 研究会で構築・運用を行っている広域分散 IX の実証実験について述べる。前述の通り、実証実験は JGN(Japan Gigabit Network)上に構築された広域分散環境のテストベッド上で行われている。図 7 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。

なお、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。

図では、IX プロバイダワーキンググループにおいて進めている、他の MPLS-IX 事業者との相互乗り入れ

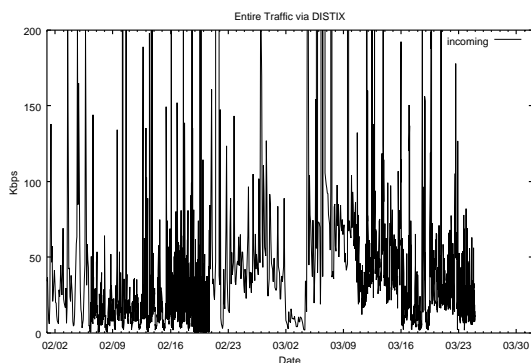


図 8 トラフィック状況 (1)

Fig. 8 Traffic over the testbed(1)

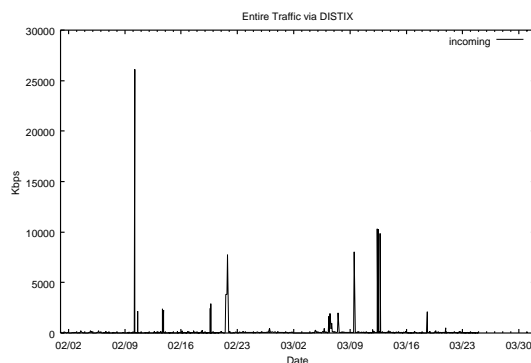


図 9 トラフィック状況 (2)

Fig. 9 Traffic over the testbed(2)

についても示している。前述の通り、現在 Japan Telecom 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。

### 5.3 実証実験への参加方法

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 のどれか一つに接続を行う必要がある。現在、Core LSR への接続の方法は以下の 3 つのうちいずれかの方法による。

- 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。
- 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。
- 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 事業者へ接続することにより、実証実験に参加が可能である。

### 5.4 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となっている。これは、MPLS-IX という新技術を用いて相互接続を行っているため、安定運用に至るまでの経過的な措置である。

図 8、図 9 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、5 分間隔でテストベッド上の 6 つの Core LSR で受け取ったトラフィックの合計を 2 か月間に渡って計測したものである。図 8 はトラフィックの推移を見やすくするため、Y 軸を 200Kbps で抑えている。機能検証を目的とした実験トラフィックが主であるためトラフィックの総和は必ずしも多くはないが、継続的なトラフィック交換が行われていることが分かる。

一方、図 9 は、おなじ期間のトラフィックを Y 軸を 30000Kbps(30Mbps) にして示したものである。図に示すように、広域分散 IX では瞬間的に 10~30Mbps 程度のトラフィックが交換されていることが分かる。なお、これらの瞬間的な高トラフィックの通信は、映像伝送などのイベント時のトラフィックである。次世代 IX 研究会では、実験目的のアプリケーションとして特定のストリームを定期的に流している。これらの実験から、本テストベッドにおいて、広域分散 IX により遠距離での相互接続環境上で高帯域アプリケーションを利用することが可能になったことが分かる。

### 5.5 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、多くの組織間で順調に相互接続ができていた。実証実験では相互接続実験を開始してから約 3 か月間、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。

- MPLS ではパケットにラベルを付与するため、通信経路上の MTU(Maximum Transfer Unit) の



設定に注意が必要である。実証実験でも、MTU の設定ミスにより、ストリームが見れないなどの問題が起こった。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必要である。

- 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 ルータベンダは MPLS-IX に必要な基本機能は実装を終えているが、一部のルータベンダでは、例えば、BGP4 の TTL(Time To Live) の値が指定できないなど、運用上の支障、もしくは混乱を招きやすい実装が残っているケースがある。次世代 IX 研究会では、これらのルータの実装上の問題についてまとめ、同時に各ベンダに対して対応を依頼している。

## 6. おわりに

本稿では、MPLS-IX アーキテクチャを用いて広域分散 IX を実現する手法について述べた。本稿で提案する手法は、既存の IX を階層的に接続することにより、既存のリソースを有効に利用しながら、スムーズに広域分散環境での相互接続環境へ移行することを可能にする。

また、著者らは次世代 IX 研究会において、広域分散 IX の実証実験を行っている。本稿では、これらの実証実験の概要について紹介するとともに、これまでの実験の状況についても報告した。

広域分散 IX は、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、あるいは商用 IX の地方展開など、さまざまな利用方法が期待されている。本研究は、これらの要求に対して、技術的な面から実現方法を示した。また、実証実験を通じて、その実用性についても示した。

今後は、広域分散 IX 技術において、IPv6 などの次世代プロトコルへの適用、あるいは QoS(Quality of Service) などの品質保証に関する技術についても研究を進めていく。

## 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、本研究は通信・放送機構の委託研究に基づいて実施しています。

## 参 考 文 献

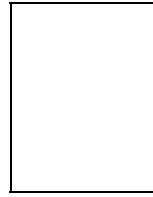
- 1) Geoff Huston: "Interconnection, Peering and Settlements", The Internet Protocol Journal, Vol 2, No 1, Mar. 1999.
- 2) Bill Manning: "Exchange Point Information", <http://www.ep.net/>
- 3) PAIX: "Palo Alto Internet eXchange", <http://www.paix.net/>
- 4) MCI WorldCom: "MAE Information", <http://www.mae.net/>
- 5) LINX: "LINX", <http://www.linx.net/>
- 6) Ameritech: "Chicago NAP", <http://nap.aads.net/main.html>
- 7) Telehouse, "NYIIX", <http://www.nyix.net/>
- 8) "Amsterdam IX", <http://www.ams-ix.net/>
- 9) "Star Tap", <http://www.startap.net/>
- 10) WIDE Project, "NSPIXP", <http://jungle.sfc.wide.ad.jp/NSPIXP/>
- 11) JPIX, "JaPan Internet eXchange", <http://www.jpix.ad.jp/>
- 12) JPNAP, "JaPan Network Access Point", <http://www.jpnap.ad.jp/>
- 13) BeX-J, "Business eXchange Japan", <http://www.bex-j.net/>
- 14) "Tohoku Regional IX", <http://www.tia.ad.jp/trix/>
- 15) "OKayama IX", <http://www.okix.or.jp/>
- 16) "Toyama Regional IX Consortium", <http://www.toyama-ix.net/>
- 17) "Next Generation IX Consortium", <http://www.distix.net/>
- 18) "PlayOnline", <http://www.playonline.com/>
- 19) "CRN Forum", <http://www.crnf.net/>
- 20) 中川 郁夫, 江崎 浩, 永見 健一: "ラベルスイッチを用いた分散 IX の設計", 分散システム/インターネット運用技術研究会 研究報告, 99-DSM-14, Jul 1999
- 21) 中川 郁夫, 林 英輔, 樋地 正浩, 八代 一浩, 菊池 豊, 西野 大: "ギガビットネットワークを用いた地域間相互接続の試み", 分散システム/インターネット運用技術研究会 研究報告, 99-DSM-15, Sep 1999
- 22) 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テンジの技術動向", 情報処理 (IPSJ Magazine), vol.42 no.7, Jul, 2001
- 23) Yutaka Kikuchi, Ikuo Nakagawa, Masahiro Hiji, Kazuhiro Yatsushiro, Dai Nishino, Eisuke Hayashi: "A trial for reconstructing the ground design of the Internet architecture in Japan", Proceedings of Second International Conference on Advances in Infrastructure for Electronic Business, Science and Education on the

Internet, Aug 2001

- 24) Ikuo Nakagawa, Hiroshi Esaki, Kenichi Nagami: "A Next Generation IX Architecture using MPLS", SAINT2002, Nara, Jan 2002
- 25) Y. Rekhter, T. Li: "A Border Gateway Protocol 4", IETF RFC1771, Mar. 1995
- 26) E. Rosen, A. Viswanathan, R. Callon: "Multi-protocol Label Switching Architecture", IETF Internet-Draft, April, 1999.
- 27) Peter Ashwood-Smith, Ayan Banerjee, et. al: "Generalized MPLS - Signaling Functional Description", IETF Internet-Draft, September, 2001.
- 28) Marc Blanchet, Florent Parent, Bill St-Arnaud: "Optical BGP: InterAS lightpath provisioning", IETF Internet-Draft, August, 2001.

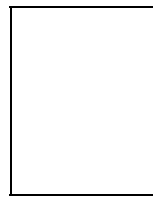
(平成 ? 年 ? 月 ? 日受付)

(平成 ? 年 ? 月 ? 日採録)



江崎 浩 (正会員)

1963年1月生。1987年九州大学工学研究科修士課程終了。同年(株)東芝入社。1998年東京大学情報基盤センター助教授。2001年東京大学大学院情報理工学系研究科, 助教授に就任。MPLS および IPv6 に関する研究開発に従事。WIDE プロジェクトボードメンバ。IPv6 普及・高度化推進協議会専務理事。電子情報通信学会会員。工学博士(東京大学, 1998)。



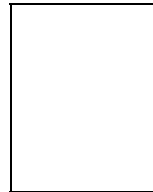
菊池 豊 (正会員)

1992年東京工業大学博士課程単位取得退学。同年より同大学情報工学科助手。1997年より高知工科大学情報システム工学科助教授。地域指向型のインターネットトラフィック交換の研究を行う。情報処理学会 DSM 研究会幹事。KPIX 実験研究協議会会長。博士(工学, 東京工業大学, 1994)。



中川 郁夫 (正会員)

1968年8月26日生。1991年東京工業大学理学部数学科卒業。1993年東京工業大学大学院総合理工学科学システム科学先攻修士課程終了。同年(株)インテック入社。同社研究所にてネットワーク管理, 大規模経路制御技術, 次世代インターネットに関する研究に従事。理学修士。情報処理学会会員。



永見 健一 (正会員)

1992年東京工業大学理工学研究科修士課程終了。同年(株)東芝入社。IETF MPLS WG で標準化活動を行い, CSR および MPLS に関する RFC を提出。現在, 東芝開発センターで MPLS および IPv6 の研究に従事。工学博士(東京工業大学, 2001)。