

MPLSを用いた分散IXアーキテクチャの設計

永見 健一 中川 郁夫 菊地 豊 江崎 浩
東芝 INTEC W&G 高知工科大学 東京大学

概要

ISP(Internet Service Provider)間の相互接続を効率的に行う技術としてIX(Internet eXchange)が使われている。従来のIX(Internet eXchange)技術の典型的な例はイーサネットなどで構成されるLAN型IXとATMを用いたATM型IXである。これら2つは、いくつかの問題点がある。例えば、ATMではOC-48(2.4Gbps)が限界といわれており、また、イーサネットでは、ギガビットイーサネットが現在の主流であるため、これ以上の高速なデータリンクでIXを用いてISP間を接続することはできない。しかし、OC-192など高速なデータリンクが存在して来ているため、これに見合う高速なデータリンクでISP間を接続する技術が必要がある。本稿では、様々なデータリンクメディアを利用できるMPLSの利点を生かしたIXアーキテクチャを提案し、このMPLS IXに必要なルータの機能および設定項目を整理する。

1 はじめに

インターネットは、多数のISP(Internet Service Provider)が存在し、そのISP間を相互接続することによって成り立っている。ISP間の相互接続を効率的に行う技術としてIX(Internet eXchange)が使われている。現在では、世界に数百ものIXが存在し[1]、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。

従来のIX(Internet eXchange)技術の典型的な例はイーサネットなどで構成されるLAN型IXとATMを用いたATM型IXである。これら2つは、いくつかの問題点がある。1つの例として、ATMではOC-48(2.4Gbps)が限界といわれており、また、イーサネットでは、ギガビットイーサネットが現在の主流であるため、これ以上の高速なデータリンクでIXを用いてISP間を接続することはできない。しかし、OC-192など高速なデータリンクが存在して来ているため、これに見合う高速なデータリンクでISP間を接続する必要がある。

本論文では、これら典型的なIX技術の問題点

を解決する新しいIX技術であるMPLSを用いたIX技術(MPLS-IX)を提案する。このMPLS-IXは、主に以下の特長を持つ。

- 様々なデータリンクメディアを使用してIXに接続することができ、異なるデータリンクメディアで接続しているISP間でピアリングができる
- 広域に適したデータリンクメディア(POS等)を利用することで、IXを広域に分散させることができる
- 高速なデータリンクメディア(POSやDWDM等)を使用することで、その時代にあった高速なIXを実現することができる

まず始めに、MPLS IXに対する要求を列挙し、必要な機能を整理する。次に、MPLS IXのアーキテクチャの基本概念を説明する。最後に、MPLS IXに必要なルータの機能を整理する。

2 MPLSを用いたIX技術

この章では、我々が提案しているMPLSを用いた新しいIX技術(MPLS-IX)を説明する。

2.1 MPLS-IXの概要

MPLS-IXのアーキテクチャは、図1に示すように、MPLSコアルータを持つIX提供者とMPLSエッジルータを持つISPが存在する。従来のIXと同様にISPの境界ルータ同士がIXを通してBGPのピアリングを行う。従来のIXとの大きな違いは、ISPの境界ルータ同士が同一のサブネットに属さず、MPLS-IX内に複数のルータが存在していることである。以下では、図1を用いて、MPLS-IXの概要を説明する。

MPLS-IXは、1つ以上のMPLSコアルータで構成される。以下の説明のためにMPLSコアルータを「IXエッジルータ」と「IXコアルータ」という用語で区別することにする。IXエッジルータは、MPLS-IXの境界ルータであり、MPLSエッジルータであるISPの境界ルータと接続するルータである。IXコアルータは、MPLS-IXのルータ(すなわち、MPLSコアルータ)でありIXエッジルータではないルータである。IXエッジルータには、ISPの境界ルータであるMPLSエッジルータが接続する。IXエッジルータとMPLSエッジルータ間のデータリンクは、MPLSがサポートされているデータリンクメディアであればどんなデータリンクメディアでも構わない。理論的には全てのデータリンクメディアでMPLSが利用可能であるので、どんなデータリンクメディアも使うことができる。ISP-AとMPLS-IX間のデータリンクメディアとISP-BとMPLS-IX間のデータリンクメディアは、異なっても構わない。また、MPLS-IX内部も任意のデータリンクメディアを使用することができる。この図では、簡単のために2つのISPしか書いていないが一般的にMPLS-IXには、複数のISPが接続することができる。以下の説明では、図中の2つのISPのMPLSエッジルータ間のピアリングのみを説明するが、複数のISPが接続しているときも同様にピアリングを行うこ

とができる。

上記の様な接続状態でISP-AのMPLSエッジルータとISP-BのMPLSエッジルータの間でピアリングを行う。MPLSエッジルータ間は、ラベル配布プロトコルを用いてLSPを設定する。LSPが設定されるとそのLSPを通してBGPによる経路情報交換を行う。BGPによって経路情報が交換されるとそれにしたがって、データトラフィックがLSP上を流れる。

ここで重要なのは、LSPで運ばれるパケットは、専用線やトンネルと同じように途中のMPLSコアルータではIPヘッダの宛先アドレスではなく、ラベル情報を参照して転送されるため、MPLS-IXのルータ(MPLSコアルータ)においてBGPで交換される経路を持つ必要がない。またLSPは、MPLSエッジルータ間で設定できるため、MPLS-IXは介在せずに、ISP同士でパケット転送のポリシーを決めることができる。これにより現在のIXの基本ポリシーであるパイラテラルの実現が可能になる。

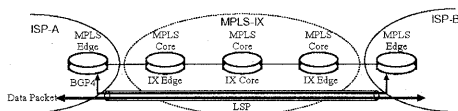


図1: MPLS-IX

2.2 MPLS-IXの動作手順

この節では、MPLS-IXの実現方法を説明する。

図2は、MPLS-IXに接続する2つのISP間にLSPを設定し、BGP4による経路情報交換を行う手順を示している。MPLS-IXでは、実際のデータを流すためのLSPを設定するために以下の手順を実行する。

1. ルータ間の物理接続を行う

2. MPLS エッジルータ間で LSP を設定するための経路情報を持つ
3. MPLS エッジルータおよび MPLS コアルータで MPLS シグナリングプロトコルを動作させる
4. MPLS エッジルータ間で LSP を設定する
5. MPLS エッジルータ間の BGP4 で経路情報交換を行う

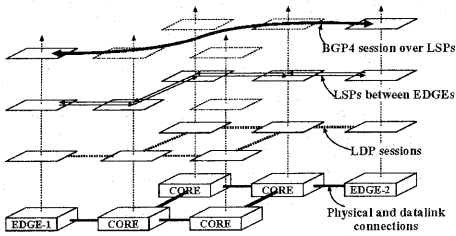


図 2: MPLS-IX アーキテクチャ

まずはじめに、MPLS コアルータ間および MPLS コアルータと MPLS エッジルータをあるデータリンクメディアで接続する必要がある。このデータリンクメディアは、MPLS が動作するものであればどのようなデータリンクメディアでも良い。理論的には、どんなデータリンクメディアも選ぶことができる。例えば、MPLS コアルータ間には、POS やギガビットイーサネットでも接続することもできるし、MPLS コアルータと MPLS エッジルータ間を ATM で接続することもできる。それぞれのルータ間は任意のデータリンクで接続することができる。

次に、MPLS エッジルータ間で LSP を設定するために必要な IP 経路情報を持つ必要がある。MPLS コアルータ間では、OSPF や IS-IS などの動的経路制御プロトコルを動作させ、IP 経路情報交換を行う。MPLS エッジルータと MPLS コアルータ間では、基本的にスタティック経路設定を行う。これは、MPLS コアルータをもつ IX 提供者と MPLS エッジルータを持つ ISP の管理者が

異なるため、運用上それぞれの IGP 経路情報交換を避けたいためである。MPLS エッジルータでは、LSP を設定する相手の MPLS エッジルータへのスタティック経路を設定し、MPLS コアルータでは、自分が接続している MPLS エッジルータへの経路を OSPF や IS-IS によって MPLS コアルータに配布する。

次に、それぞれのルータで MPLS 機能を動作させ、MPLS のシグナリングプロトコルを起動する。現在の MPLS のシグナリングプロトコルは大きく分けて 3 つ存在する。LDP[3]、RSVP[4]、CR-LDP である。これらのシグナリングプロトコルにより、MPLS エッジルータ間に LSP を設定する。これらのシグナリングプロトコルで設定される LSP は片方向なので、両方向のデータ交換ができるように 2 つの LSP を設定する。図 2 では、Edge-1 から Edge-2 への LSP と Edge-2 から Edge-1 への LSP の 2 つを設定することになる。

MPLS エッジルータ間に LSP を設定した後、その LSP を通じて BGP4 で経路情報交換を行う。

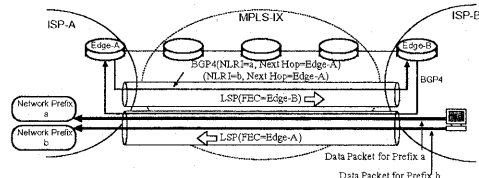


図 3: BGP による経路交換と LSP 上のデータ通信

図 3 で BGP による経路情報交換とその後のデータトラフィックの流れを説明する。ISP-A の MPLS エッジルータである Edge-A と ISP-B の MPLS エッジルータである Edge-B 間でピアリングを行う。この図では簡単のために、Edge-A から Edge-B へ BGP による経路情報通知を示している。Edge-A と Edge-B の間には、LSP が 2 本設定されている。1 本は、Edge-A から Edge-B へのデータ通信用であり、FEC は Edge-B が設定されている。もう 1 本は Edge-B から Edge-A へのデータ通信用であり、FEC は Edge-A が設定されている。現状の MPLS での LSP は、片方向通信にのみ使えるの

で、ある MPLS エッジルータ間では両方向通信のために 2 本の LSP が必要であるために、このような状況となる。

ISP-A には、2 つのネットワークプレフィックスがあり、それぞれ「a」と「b」である。Edge-A は、Edge-B に BGP で「ネットワークプレフィックス a のネクストホップが Edge-A」であること「ネットワークプレフィックス b のネクストホップが Edge-A」であることを通知する。この BGP パケットは、LSP (FEC=Edge-B) を通して送信される。この BGP パケットを受信した Edge-B は、ネットワークプレフィックス a および b のパケットを LSP (FEC=Edge-A) で送信できるように設定する。これは、ネットワークプレフィックス a および b のネクストホップが Edge-A であるため、Edge-A へパケットを配送できる LSP を通してデータパケットが転送される。

このように、MPLS エッジルータ間で片方向 1 つずつの LSP を設定し、BGP パケットとデータパケットの両方をその LSP で転送する。ネットワークプレフィックス毎に LSP を 1 本設定する方法もあるが、IX の様なフルルートを持つルータでプレフィックス毎に 1 本の LSP を設定すると莫大な LSP が必要となるため、ネクストホップに対して 1 本の LSP を設定することにより、MPLS-IX 中の LSP 数を少なく保つことができる。

2.3 MPLS-IX の特長

前節で説明した MPLS 技術を使った MPLS-IX は、以下のような特長を持つ。

- 様々なデータリンクメディアで IX に接続可能
従来の IX では、1 つのデータリンクメディアを使用している。そこで、IX に接続するためには、指定されたデータリンクメディアを使う必要があった。しかし、MPLS は、ATM、POS、GbE など様々なデータリンクメディアを扱うことができる特長を持つ。このため、MPLS を利用した MPLS-IX では、様々

なデータリンクメディアによって IX に接続することができる。

- 異なるデータリンクメディアで接続している ISP 同士でピアリング可能
MPLS では、異なるデータリンクメディア間で LSP を作成することができる。この LSP を用いて BGP およびデータパケットを交換することにより、異なるデータリンクメディアで接続している ISP 同士で BGP によるピアリングが可能になる。
- 従来の IX との接続性がある (マイグレーションが容易)
MPLS-IX では、異なるデータリンクメディアで接続することが可能である。このため、従来の IX と MPLS-IX を接続し、従来 IX に接続している ISP が MPLS 対応ルータを準備すれば、MPLS-IX に接続している ISP と従来の IX に接続している ISP 間でピアリングを行うことが可能である。
- 高速なデータリンクでの接続が可能
MPLS-IX では、様々なデータリンクメディアを使えるために、POS やギガビットイーサネット等の高速なデータリンクメディアを使うことが可能である。また、IETF MPLS WG では、高速なデータリンクメディアである DWDM 等に MPLS を適用する方式が検討されているため、今後の高速性が期待できる。
- 広域分散が可能
MPLS-IX では、データリンクメディアを選ばないため、ATM や POS のように広域で使うように設計されたデータリンクメディアを使うことができる。このデータリンクメディアを MPLS コアルータ間で使用することにより、MPLS-IX のバックボーンを全世界に拡張することも可能である。また、ISP と IX の接続を広域で使えるデータリンクメディアにすることで、IX に接続する ISP のルータを IX が存在するコロケーションに置く必要がなくなる。

2.4 MPLS IX に対する要求条件

この節では、MPLS IX に必要な要求条件をまとめる。MPLS IX では、図 4 の様に MPLS IX の提供者とその MPLS-IX を使用する利用者 (例えば、ISP) に分類される。MPLS IX 提供者とその利用者は、それぞれ別のネットワーク管理ポリシーで運用されることが一般的であるため、MPLS IX 提供者の境界ルータと MPLS IX 利用者の境界ルータとの間では、できるだけお互いに影響を与えないような構成にする必要がある。また、運用管理を容易にする必要がある。上記の要求条件は、以下のようにまとめられる。

- MPLS IX 提供者と利用者の間では共通の IGP (Interior Gateway Protocol) を使わない。

MPLS IX 提供者と利用者の間で共通の IGP を利用すると、相手のネットワーク内の経路変更が自分のネットワーク中に伝わり、相手の経路変更の影響を受け、ルータの経路計算の負荷が大きくなったり、経路トラブルを発生する可能性がある。そこで、MPLS IX 提供者と利用者の間で共通の IGP を使わないようにすることで、上記の問題を回避する。図 4 では、MPLS IX 提供者内ネットワークの経路制御を OSPF とし、MPLS IX 提供者の境界ルータと (図 4 Core-B, Core-D) と利用者の境界ルータ (図 4 Edge-A, Edge-E) では、スタティック経路を使うことにより、同一の IGP を使わない例を示している。

- MPLS IX の境界ルータで LSP のフィルタを設定できるようにする

LDP を MPLS のシグナリングとして用いる MPLS IX 利用者は、自分のアドレスを FEC としてラベル割当メッセージを MPLS 提供者の境界ルータに送信する。ところが、設定ミスや悪意のある利用者からは、利用者のアドレス以外を FEC としてラベル割当メッセージが送信される可能性がある。もし、このメッセージを MPLS IX 提供者が受信してしまうと、正常ではない LSP が設定されることにな

り、パケットが期待していないところに転送される可能性がある。このような MPLS IX 利用者の設定ミスによる影響を、他の利用者に与えないようにするため、自分のアドレス以外の FEC を受理しないように MPLS IX 提供者の境界ルータに LSP のフィルタを設定する必要がある。

- LSP 毎のトラフィック統計情報を取得可能にする。

IX の利用者である ISP 等は、どの ISP 間でトラフィックが交換されているのかを知ることによりネットワーク設計を容易にしたり、課金のための情報を得ることができる。MPLS IX では、相手先毎に LSP を設定するため、この LSP 毎にトラフィック情報を取得すれば、この目的を達成できる。そこで、LSP 毎にトラフィック情報を取得する必要がある。

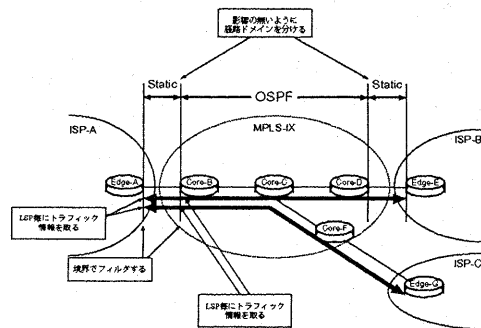


図 4: MPLS-IX アーキテクチャ

3 MPLS IX に必要なルータの機能

この章では、前章で説明した、MPLS IX のアーキテクチャと MPLS IX の要求条件を満たすための MPLS IX で使うルータに必要な機能を整理する。

3.1 コアルータの機能

この節では、MPLS IX 提供者のルータで必要な機能を示す。以下に必要な機能を示す。

- エッジルータへの経路を MPLS IX 内部の IGP(例えば OSPF) で MPLS IX 内部のルータに伝搬できること。

RSVP の場合は、エッジルータとコアルータ間のネットワークの経路情報あるいはエッジルータのループバックインターフェイスを伝搬できれば良い。図 5 では、Edge-A のループバックアドレスを A、Edge-A のインターフェイスアドレスを X.A とする。ここで、X は IP アドレスのネットワーク部を示し、A は IP アドレスのホスト部を示すとする。この場合には、Core-B は、MPLS IX 中にネットワークプレフィックス X と Edge-A のループバックインターフェイス A を伝搬できれば良い。

シグナリングプロトコルに LDP を利用する場合は、エッジルータのインターフェイスアドレスかエッジルータのループバックアドレスを伝搬する必要がある。例えば、図 5 では、Edge-A のループバックアドレス A と Edge-A のインターフェイスアドレス X.A を伝搬させないといけない。

RSVP と LDP とも LSP の FEC は、エッジルータのループバックアドレスかインターフェイスアドレスである。しかし、LDP の場合は、プロトコル仕様上、FEC と全く同じ経路情報をコアルータで持つ必要があるため、FEC をインターフェイスアドレスとする場合にインターフェイスアドレスと全く同じ経路情報を持つ必要があるため、インターフェイスアドレスを経路プロトコルで広告する必要がある。RSVP の場合は、宛先まで到達できる経路があれば良いので、インターフェイスアドレスではなく、そのインターフェイスが属しているネットワークプレフィックスを広告すれば十分である。

- LSP のフィルタができること。

MPLS のシグナリングプロトコルである LDP や RSVP を MPLS エッジルータから受信したときに拒否することができるようなフィルタを設定できる必要がある。例えば、特定の FEC のみを受理するように設定したり、RSVP Path メッセージの送信 IP アドレスにより拒否できるように設定する。

- LSP 毎の統計情報が取れること。
MPLS IX の入口および出口の境界ルータで LSP にながれるトラフィック統計情報を取得できる必要がある。これは、課金やパケットロスを監視するため等を使う。

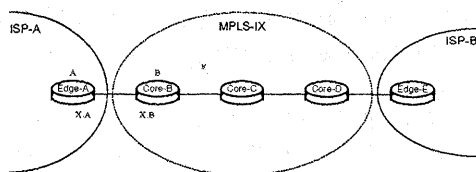


図 5: MPLS-IX ネットワーク構成

3.2 エッジルータの機能

この節では、MPLS IX 利用者であるエッジルータで必要な機能を示す。以下に必要な機能を示す。

- スタティック経路で設定した相手エッジルータへのアドレスに対して LSP を設定できること

要求条件で示したように、エッジルータとコアルータの間では、同一 IGP を動かさない。そこで、エッジルータでは、スタティック経路等で相手先エッジルータへの LSP を設定できる必要がある。一般的な MPLS の使い方では、エッジルータとコアルータ間で OSPF を用いることが想定されている。MPLS IX での使い方は一般の方法と異なるので、注意が必要がある。

- BGP パケットが LSP に流れること
BGP のピアアドレスは、上記で設定した LSP

の FEC アドレスである。この BGP のパケットを LSP に流す必要がある。これは、LSP が切断されたときに、LSP を通るデータパケットを迂回させるために必要である。この機能が実現されない場合は、BGP による経路情報は交換できるが、実際のデータパケットを交換する LSP が切れているため、データパケットがすべて廃棄されてしまうことになる。

- BGP で受信した経路にマッチするデータパケットが LSP に流れること
この機能は、MPLS IX の基本動作として必要なものである。BGP で受信した経路にマッチするデータパケットを LSP に流すことにより、コアルータで IP を見る必要がなくなるので、コアルータにフルルートを持つ必要がなくなり、エッジルータ間でバイラテラルポリシーでパケット交換ができる。
- IP TTL を MPLS TTL にコピーしない設定ができること
BGP パケットが MPLS IX 外部を経由して相手エッジルータに到達しないように IP TTL を 1 として送出する。通常、エッジルータでは、IP TTL を MPLS TTL にコピーして送信するが、この動作を行うと、コアルータでパケットが廃棄されてしまう。そこで、IP TTL を 1 としながら、コアルータでパケットを廃棄せずに転送できるように、エッジルータで挿入する MPLS TTL を IP TTL からコピーせずに設定値 (例えば、最大値の 255) で送信できるように設定できる必要がある。また、出口のエッジルータでも通常の動作で MPLS TTL を IP TTL にコピーするようになっているが、これも同様に MPLS TTL を IP TTL にコピーしないように設定できる必要がある。
- LSP のフィルタができること。
MPLS のシグナリングプロトコルである LDP や RSVP をコアルータから受信したときに拒否することができるようなフィル

タを設定できる必要がある。例えば、特定の FEC のみを受信するように設定したり、RSVP Path メッセージの送信 IP アドレスにより拒否できるように設定する。

- LSP 毎の統計情報が取れること。PHP 無のラベル割当ができること。
エッジルータでの LSP 毎の統計情報をとれる必要がある。送信側エッジルータでは、必ずラベルが割り当てられているので、そのラベルのトラフィック情報がとれば良いが、受信側エッジルータでは、PHP [2] を用いるとラベルがエッジルータの前段コアルータで削除されてしまうため、受信側エッジルータでトラフィック統計情報をとることができない。そこで、LSP 毎の統計情報をとるためにも PHP 無のラベル割当ができる必要がある。

4 まとめ

本論文では、MPLS を用いた新しい IX アーキテクチャ (MPLS IX) の概要を説明し、基本的な要求条件を整理した。これをもとに、MPLS IX を実現するために必要なコアルータおよびエッジルータの基本機能を整理した。

今後は、冗長性やロードバランスを考慮した場合や複数の MPLS IX 同士をつなぐ場合について検討するとともに、実ルータを用いて機能評価およびパフォーマンス評価を行う予定である。

参考文献

- [1] Bill Manning, "Exchange Point Information", <http://www.ep.net/>
- [2] E. Rosen, et al., "Multiprotocol Label Switching Architecture", IETF RFC 3031, Jan. 2001.
- [3] L. Andersson, et al., "LDP Specification", IETF RFC 3036, Jan. 2001.

- [4] D. Awduche, et al., "RSVP-TE: Extensions to RSVP for LSP Tunnels", IETF Internet Draft, draft-ietf-mpls-rsvp-lsp-tunnel-09.txt, Feb. 2002.