

誘導経路情報を用いたインターネットパス制御アーキテクチャ

Internet Routing Path Control Architecture using Inviting Routing Information Advertisement

東京大学 大学院 情報理工学系研究科

江崎 浩 (Hiroshi ESAKI)

<Hiroshi@wide.ad.jp>

あらまし

インターネットシステムにおける経路制御は、ますます複雑化するインターネットの構造と要求に対応しなければならない。しかし、現在の経路制御システムはグローバルスケールで、宛先アドレスをルートとしたスパンニングツリー型のパスを形成するために、必ずしも自由なパスの制御(トラフィックエンジニアリング)やIngress側での迅速なトラフィックの制御を実現することが容易ではない。本論文では、エニキャストの概念を拡張した誘導経路情報の広告を用いたインターネットパス制御アーキテクチャ(IRIDES; Invitation Routing Information aDvertisement for path Engineering System)の提案を行っている。誘導経路情報を用いて、送信ノードにより近いポイント(誘導ポイント)において、トラフィックエンジニアリング可能なルーティング空間(誘導空間)にIPパケットを誘導する。誘導空間におけるトラフィックエンジニアリングは、任意の手法を用いて実現することができる。提案アーキテクチャは、既存の経路制御システムとの共存が可能であり、十分なスケーラビリティを持つ。さらに、提案アーキテクチャをもとに、いくつかの具体的なアプリケーション(e.g., マルチエージェント型マルチホームNEMO/MIP)が実現されることを示す。

Abstract

Routing system in the Internet must come up with the complex technical requirements and with complex network structure and topology. The existing Internet routing system is based on the IP packet forwarding using only the destination IP address. Routing protocol can be realized to establish the (dynamic) spanning tree, whose root is the destination node/network and the end leaves are the source nodes/networks. Therefore, it would not be easy to achieve a flexible path control for the traffic engineering purpose or rapid traffic control near the ingress point to the Internet from the source node. In this paper, we proposed an new routing architecture, called IRIDES(Invitation Routing Information aDvertisement for path Engineering System), that extend the anycast architecture. In IRIDES, the IP packet is invited to a nearest entry point to the invited space, where can apply any traffic engineering technique in it. The proposed architecture has well scalability and deployment capability by the coexistence with the legacy Internet routing system. Also, we show some practical applications based on the proposed architecture, using some particular technical instances.

キーワード インターネット, トラフィックエンジニアリング, 誘導経路情報, ルーティング, 経路制御, IRIDES

(注意) 本論文と同様の内容の論文は、電子情報通信学会 和文論文誌B に 投稿されている(2005年1月)。

1. はじめに

TCP/IPを基盤技術とするインターネットシステムは、エンド・ツー・エンドアーキテクチャを基本原理として、1980年代後半から驚異的な成長を遂げてきた。その成長は、インターネットに接続されるノードの数とネットワークの数のみならず、その利用者の多様度の増大に関しても著しいものがある。インターネットの成長とともに、インターネットシステムにおいて、その根本的な機能を提供するルーティング制御技術(経路制御技術)も、変遷を遂げてきた。インターネットが運用を開始した当初は、ノードのインターフェースを識別するためのIPアドレス長は4ビットであり、今日広く利用されている動的経路制御は適用されず、すべて、Static Routingで運用することができたし、ノードのIPアドレスの解決は、DNSのようなディレクトリサービスを利用することなく、Staticなテーブルを各計算機のファイルに書き込む形態であった。やがて、インターネットシステムの成長とともに、IPアドレスは、8ビットに拡張され、その後、32ビット[1]に拡張されて、今日、128ビットのアドレス長(IPv6)[2]へと拡張された。その過程で、グローバルなIPアドレスと計算機の論理的な名称(FQDN; Fully Qualified Domain Name)のマッピング情報を提供するDNSシステム[3]が導入された。

一方、経路制御システムは、当初のStatic Routing方式のみで運用されたが、やがて、ネットワークの状況に応じて、到達可能な最適の経路を自動的に決定して、IPパケットの転送経路の制御と管理を行う、動的経路制御が導入されるようになった。現在でも、Static Routing(あるいはDefault Routing)は、実際のネットワーク運用において、広く適用されているが、経路の冗長性の確保や宛先ネットワークへの接続性の確保を実現するために、インターネットのコア部では、動的経路制御が適用されているのが一般的である。動的経路制御として、インターネットにおいて、現在広く適用されているプロトコルは、IGP(Interior Gateway Protocol)としてはRIP[4]およびOSPF[5]/ISIS[6]であり、EGP(Exterior Gateway Protocol)としてはBGP[7]である。これらのプロトコルは、それぞれ固有の経路決定アルゴリズムを採用しているが、根本的には、すべて共通のアーキテクチャを適用することで、大規模化(スケラビリティ)への対応を行っている。これは、宛先IPアドレスのみを用いた経路表の作成/管理と、それに基づいたIPパケットの転送処理である。これは、宛先ノードをルート(頂点)とする、スパニングツリーをグローバルに形成することによって、経路制御の複雑度をオーダー $O(n)$ に押さえ込んでいる。ただし、ここで、 n は、経路制御ドメインに存在するノード/ネットワークの数を表す。

ますます複雑化および多様化するインターネットの構造と要求に対して、インターネットシステムにおける経路制御は、適切に対応することが要求されている。しかし、現在の経路制御システムは、上述の通り、グローバルスケールで宛先IPアドレスをルートとしたスパニングツリー型のパスを形成するために、必ずしも自由なパスの制御(トラフィックエンジニアリング)や、DOS攻撃に対するIngress側の迅速な制御など、現在のインターネットシステムが必要としている機能を実現することが容易ではない。これまで、インターネットにおける経路制御システム全体の抽象化を行い問題点の整理が行われず、解決すべき技術課題に対応するための具体的な手法やアーキテクチャの提案と実装および適用は、ある意味、アドホックに行われてきた。その結果、本質的には、同じ手法やアプローチを用いているのも関わらず、個別の技術エレメントに整合性がなく、それぞれ、個別の技術が新たに提案され適用されてきた。

そこで、本論文では、改めて、現在のインターネットで適用されている経路制御アーキテクチャの整理と抽象化を行い、その問題点と新たな要求条件を整理し、その一般的な解決手法を整理し抽象化した形で提案することを目的としている。具体的には、エニキャストの概念を拡張した誘導経路情報の広告を用いたインターネットパス制御アーキテクチャ(IRIDES; Invitation Routing Information advertisement for path Engineering System)の提案を行っている。誘導経路情報を用いて、送信ノードにより近いポイント(誘導ポイント)において、トラフィックエンジニアリング可能なルーティング空間(誘導空間)にIPパケットを誘導する。誘導空間におけるトラフィックエンジニアリングは任意の手法を用いて実現することができる。提案アーキテクチャは、既存の経路制御システムとの共存が可能であり、十分なスケラビリティを持つ。さらに、提案アーキテクチャをもとに、いくつかの具体的なアプリケーション(e.g., マルチエージェント型マルチホームNEMO/MIP)が実現されることも示す。

以下が本論文の構成である。第2章では、現在広く利用されているインターネット経路制御技術の抽象化とその問題点ならびに要求条件を明らかにする。第3章では、IRIDESアーキテクチャの具体的な提案を行う。第4章では、IRIDESアーキテクチャを適用した、個別の経路制御の技術課題の解決手法の具体例を示す。第5章では、本論文の結論を述べる。

2. インターネット経路制御技術の問題点

2.1 既存の経路制御技術の分析

本節では、本論文で提案するアーキテクチャに関連する2つの既存経路制御技術の特徴を整理する。

2.1.1 Destination-Based Routing

インターネットにおいて、現在広く適用されている経路制御プロトコルは、IGP(Interior Gateway Protocol)としてはRIPとOSPF/ISISであり、EGP(Exterior Gateway Protocol)としてはBGPである。RIPはDistance Vector方式、OSPFおよびISISはLink State方式、さらにBGPはPath Vector方式といった、それぞれ異なる固有の経路決定方式を用いている。一般的に、(大規模な)AS内部での経路制御ではOSPFおよびISISが、一方、AS間での経路制御ではBGPが適用されている。これらの経路制御プロトコルは、異なる経路決定方式を用いているが、根本的には、宛先IPアドレスのみを用いた経路表の作成/管理と、それに基づいたIPパケットの転送処理を行っている。すなわち、一見異なる経路制御決定方式を用いているように見えるが、すべて共通のアーキテクチャを適用することで、大規模化(スケーラビリティ)への対応を行っているといえよう。具体的には、宛先ノード(あるいはネットワーク)をルート(頂点)とする、スパニングツリー(Spanning Tree)を、与えられたルーティングドメインにおいて形成することによって、経路制御の複雑度をオーダー $O(n)$ に押さえ込んでいる。Inter-Domainルーティングを司るBGPにおけるDefault-Freeな運用形態では、いわゆるフルルート(Full Routes)を隣接する(Peeringしている)BGPルータと相互交換することで、基本的にはグローバルスケールのスパニングツリーを管理している。一方、Intra-Domainを司るOSPFやISISでは、そのルーティングドメイン内部のネットワークにおけるスパニングツリーの管理を行っている。

スパニングツリーは、閉じたトポロジー部分を持たない開放型のトポロジー構成を持つ。スパニングツリーが閉じた部分を持たないので、経路制御プロトコルに要求されるループフリー機能の提供が実現されていると考えることが可能である。すなわち、ルーティングドメイン上のすべてのオブジェクト(=ノードあるいはネットワーク)から、注目する宛先オブジェクト(=ノードあるいはネットワーク)、すなわちツリーのRootノードへの、ループ(閉じた部分)のないパスを提供することができている。このスパニングツリーは、すべてのリーフ(Leaf)ノード、すなわち、すべての送信元オブジェクトで共有されたスパニングツリーとなっている。送信元オブジェクトごとのスパニングツリーを形成管理することで、ルータが管理すべき経路情報の量のオーダーを $O(n^2)$ から $O(n)$ に削減することに成功しているわけである。これが、インターネット経路制御システムにおけるDestination-Based Routingであり、すべての経路制御プロトコルが、この原理を共通に適用している。

2.1.2 経路情報の集約化

もともと各組織は、個別のIPアドレス空間(=Portable IP Address Space)を持ち、このアドレス情報を経路制御情報としてインターネットに対して広告していた。しかし、インターネットの大規模化に伴い、このような個別のIPアドレス空間を各組織がインターネットに広告すると、ルータは、経路テーブルのエントリーとして各組織ごとにエントリーを生成/管理しなければならなくなってしまう。そこで、この問題を解決するために、アドレスの集約化(Address Aggregation)が適用されるようになった。これは、インターネットの成長に伴う経路表の増大に対応するための運用方法である。アドレスの集約化を実現するために、IPアドレス空間は、各組織が独立に取得管理するものではなく、集約化を行うネットワーク単位ごとに行い、そのネットワーク単位に属する組織は、そこからIPアドレスの部分空間を割り当ててもらった形態が一般的になってきた。集約化を行うネットワーク単位は、一般的には、ISP(Internet Service Provider)やNIR(National Internet Registry)となり、最上位は3つのRIR(Regional Internet Registry)となる。したがって、結果的に、IPアドレス空間に関して、階層的なツリー構造が構築されていると見ることができる。なお、経路制御におけるIPアドレスの集約化によって、このようなIPアドレス空間の階層構造とIPパケットの転送経路との間には強い相関関係が構築されることになる。

2.1.3 経路制御システムへの要求事項

インターネットシステムにおける経路制御は、ますます複雑化および多様化するインターネットの構造と要求に対して、適切に対応できることが要求されている。前節における議論の通り、現在の経路制御システムは、以下の2つの技術的な特徴(=制約)を持っている。

- ・ルーティングドメイン内において、宛先オブジェクト(=宛先ノードや宛先ネットワーク)をルートとした、

すべての送信元オブジェクトで共有されたスパニングツリーパスを生成管理している。

- ・グローバルスケールに、階層的なIPアドレス空間の割り当てが行われている。

インターネットの経路制御システムは、この2つの技術的な制約を持つために、さまざまな課題が顕在化し、これまで、それぞれの課題に対してアドホックな対応がとられてきた。これらの課題は、以下の3点に集約化することができる。

1. 共有されない転送パスの提供

宛先オブジェクトごとに管理されるスパニングツリー上のパスに依存しない IPパケットの転送パスの提供は、トラフィックエンジニアリングに必要な重要な要素技術の一つである。現在の経路制御システムでは、異なる送信元オブジェクトから、同一の宛先オブジェクトに向けて転送されるIPパケットは、共通のパス(スパニングツリー)を共用しなければならなくなってしまう。宛先オブジェクトへの複数のIPパケット転送パスの提供は、シングルISP内でのサービス品質提供の冗長度の向上に貢献するだけでなく、複数のISPが関与するマルチホーム環境の提供における通信品質の向上にも貢献する。

2. プラグアンドプレイ環境の提供

IPアドレス空間の階層化と集約化に伴い、オブジェクト(ノードあるいはネットワーク)の、物理的あるいは論理的な移動時には、該当するオブジェクトシステムにおいては、IPアドレスのリナンバリングなど、さまざまな、ネットワーク関連の設定変更を行う必要がある。オブジェクトの論理的な移動は、物理的な移動時には、ほぼ、必然的に発生するが、物理的な移動が伴わない場合でも、契約先のISPを変更する場合などに発生する。契約先のISPを変更する場合には、一般的には、そのオブジェクト(ノードあるいはネットワーク)が使用するIPアドレスが変更になってしまい、それに伴い、ネットワーク関連の設定を多数変更しなくてはならない。すなわち、本論文で議論するプラグアンドプレイは、設定フリーの環境を提供する自動構成認識技術ではなく、設定"変更"フリーの環境を提供するものを意味する。

3. Ingress側での迅速な対応を実現するフレームワークの提供。

Destination-Based Routingを適用した経路制御システムにおいては、送信元オブジェクトに近い地点での迅速な対処や制御が、一般的に難しい。特に、DDoS(Distributed Denial of Service)攻撃への対処のためのIPパケットフィルタリングなど、基本的には、現在は、Egress側に近い地点(すなわち受信オブジェクト側)で対処している。Ingress側での迅速な対応が実現すれば、ネットワーク内に、本質的には不要なトラフィックの流入を削減できるなどの効用も期待できるであろう。

3. 提案アーキテクチャ; IRIDES

3.1 IRIDESアーキテクチャ

Destination-based Routingアーキテクチャに伴う不十分な転送パス制御とアドレスの集約化に伴う階層化が原因で発生する不十分なプラグアンドプレイ機能の2つの問題を解決するために、誘導経路情報の広告を用いたIRIDES(Invitation Routing Information aDvertisement for path Engineering System)アーキテクチャの提案を行う。本アーキテクチャでは、通常のレイヤ3転送空間とは別に、トラフィックエンジニアリング可能なルーティング空間(誘導空間)を定義し、誘導空間と通常のレイヤ3転送空間との間には複数の誘導ポイント(IIR; Ingress IRIDES Router)ルータを設置する。なお、誘導空間は、物理的に独立したネットワーク空間でもよいし、通常のレイヤ2やレイヤ3の基盤上に構築される仮想的なオーバーレイネットワークのどちらでも構わない。すなわち、独立したネットワーク空間を持つ構造にするか、それともオーバーレイ型の仮想的な誘導空間を構築するかは、アーキテクチャの具体的な実装方法の選択肢である。

図1に、IRIDESシステムにおける通常のレイヤ3空間の概念図を示した。図では、2つのIIR(IIR1およびIIR2)が、インターネット上に分散配置されている。IIR1とIIR2は、まったく同一の経路情報をインターネットに対して広告する(エニキャスト[8])。送信元ノードから転送されたIPパケットは、経路制御プロトコル的に、より近いIIR(誘導ポイント)に転送(誘導)され、IIRを通して誘導空間に誘導転送される。

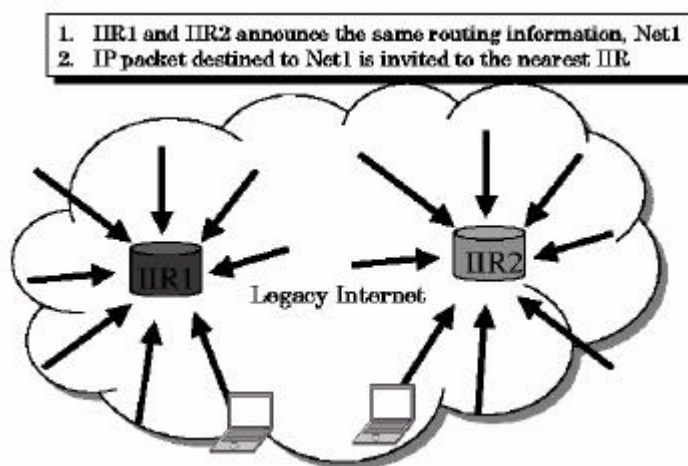


図1. 通常のレイヤ空間でのIRIDESの構成概念図
Fig.1 IRIDES Configuration in Legacy Internet Space

図2に、誘導空間(Traffic Engineering Plane)上に存在する宛先ホスト(DH1)に、3つの異なるホスト(SH1, SH2, SH3)からIPパケットが転送される場合を示した。送信側から送信された同一宛先IPアドレスを持つIPパケットは、経路制御的に、もっとも近接する誘導ポイント(IIR)に誘導される。IIRを通じて誘導空間に転送されたIPパケットは、誘導空間で適用されている経路制御方式に従って、宛先ホスト(DH1)に転送される。IIRでは、受信したIPパケット内の宛先IPアドレス情報(+その他のIPパケットフロー識別子情報)から、誘導空間での経路制御で使用されるIPパケットの転送のための識別子を選択し、適切なデータフォーマットに変換し、これを誘導空間に転送する。なお、誘導空間で使用される経路制御のための識別子としては、MPLSのラベルやVLANタグ、IPTunnel用IPアドレス、あるいはIPv6アドレスなど、誘導空間ネットワークの運用者が、適切な識別子を自由に選択可能とすることができる。

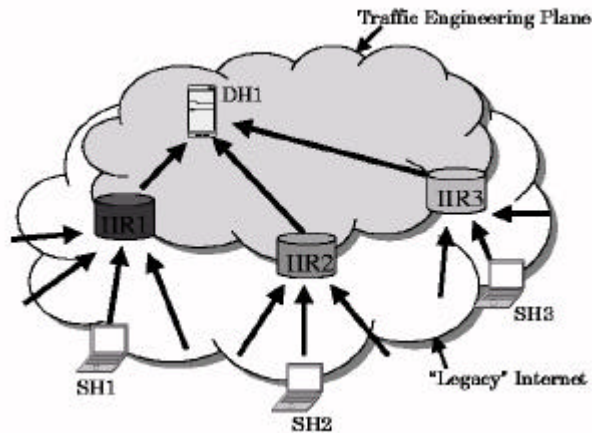


図2. IRIDESにおけるIPパケットの転送概念図
 Fig.2 IP Packet Transmission in IRIDES System}

さらに、図3には、いったん誘導空間に転送されたIPパケットが、EIR(Egress IRIDES Router)を経由して、再び、通常のレイヤ3空間上の宛先ネットワークに転送される場合を示した。宛先ネットワークが、トランジットを行わないスタブ型のネットワークであれば、誘導空間をバイパス経路として、誘導空間が接続するルーティングドメイン上で自由な転送パスを定義利用することを可能とすることができる。

図4に、IIRからEIR、あるいは宛先ホストに対して、複数の経路を提供する場合の構成例を示した。図では、2つの誘導ポイント(IIR)から、それぞれ、2つの独立なデータ転送パスを提供しており、合計で4本の転送パスが提供されている。IIRにおいては、どのようなポリシーで、2つのパスを利用するかを判断を、ローカル判断で行うことができる。

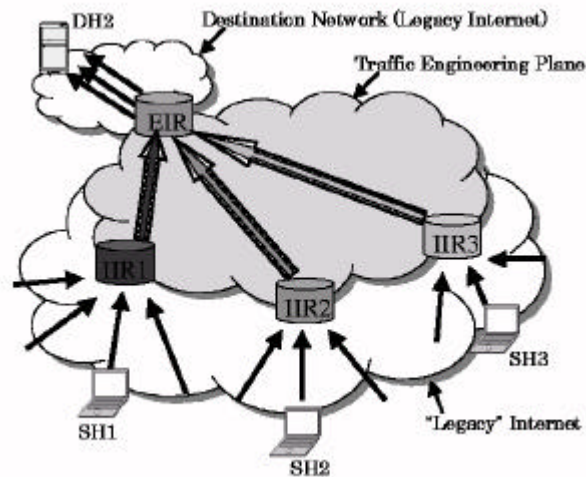


図3. 誘導空間をバイパス経路とする構成例
 Fig.3 IRIDES Configuration for Bypass Transmission Network

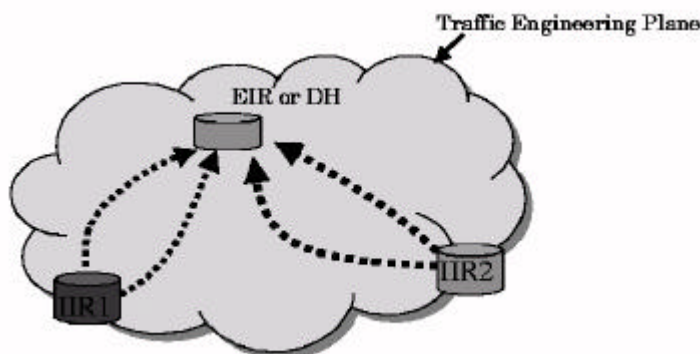


図4. 誘導空間におけるマルチパスの提供
Fig.4 Multi-Paths Data Transmission in Traffic Engineering Plane

上述した IRIDESアーキテクチャは、以下に示したような技術的な特長を持つ。

1. エニキャスト技術を用いた誘導ポイント (IIR) の分散配置

- (a) スケーラビリティ
IIRの負荷分散が、(拡張された) エニキャスト技術を用いた経路制御によって実現されるとともに、宛先オブジェクトへの転送パスを分散させることができる。IIRは、原理的には、グローバルインターネット上に分散配置することも可能である。
- (b) 転送経路の最適化
送信元オブジェクトに最も近接するIIRを経由してIPパケットの転送を行うことができる。その結果、より適切な経路の利用を可能にすることができる。
- (c) Ingress側での制御の可能性
DDoSへの対応、輻輳制御、あるいは、ポリシー制御を、IIRにおいて実行することが可能となる。

2. 既存の IPインフラストラクチャーとの共存

- (a) エッジ部を既存のIPインフラストラクチャー、コア部をトラフィックエンジニアリング可能な誘導空間として構築することで、すべての既存ホストおよび既存ネットワークとの共存を可能にする。
- (b) 誘導空間で使用するIPパケットのトランスポート技術は、特定の技術に固定されず、任意の転送方法を適用可能にすることができる。
- (c) 図2および図3に示したように、既存のIPインフラと誘導空間をまたがったネットワーク上で、宛先オブジェクトからすべての送信元オブジェクトへのスパンニングツリーが形成される。このスパンニングツリーは、既存の経路制御システムで形成されるスパンニングツリーよりも、共有部分を小さくすることが可能となる。
- (d) 誘導空間(トラフィックエンジニアリングプレーン)でのIPパケットの転送パスは、既存の経路制御をもとに計算される転送パスとは独立に定義することができるので、既存の経路制御システムとの共存が容易に実現可能である。
- (e) 誘導空間におけるネットワークポロジの定義は、既存のIPインフラのものとは独立に定義可能であり、不均一なネットワーク基盤でも、同一の機能を持ったノードからなる均一な仮想ネットを定義することができる。現実のネットワークにおいては、仮に、単一ベンダーで構成されたネットワークでも、各ノードごとに、提供機能に関する不均一性が存在する場合が一般的である。

3. ポータブルアドレス環境の提供

エンドツーエンドでのIPパケット通信のためのIPアドレスと、誘導空間におけるトラフィックエンジニアリングのためのアドレス(識別子)の分離により、エンドオブジェクト(ノードあるいはネットワーク)は、結果的に、ポータブルなIPアドレス空間を獲得することができるようになる。ポータブルアドレスの提供により、モバイルオブジェクトへの対応とプラグアンドプレイ環境(設定変更フリーという意味での)の提供を可能とする。

IRIDESアーキテクチャでは、現在のインターネットシステムにおける、エンドノード(正確にはインターフ

エース)の識別を行うためのグローバルな識別子と、そこへのIPパケットの配送に必要な経路制御に必要な識別子を同一のものとし、この識別子のみを用いて転送パスの決定を行っているという、現在のインターネットの本質的な識別子管理と利用のアーキテクチャへの改善を提案している。すなわち、少なくとも、誘導空間においては、エンドノード識別子と転送経路決定のための識別子を分離することで、多様な、転送パスのエンジニアリングを可能としている。

3.2 IRIDES プロトコルアーキテクチャ

IRIDESシステムのプロトコルアーキテクチャは、有線と無線の区別、および、Stable/Nomadic/Mobile/Adhocの区別を行わず、プロトコルアーキテクチャとしては、個別の要素技術を選択/定義可能なメタ言語の構造に似たフレームワークのみの定義を行うアーキテクチャとして設計する。これにより、同一のプロトコルアーキテクチャ、およびこれを実現するソフトウェアアーキテクチャにおいて、個別技術を取替え可能な構造にすることを可能にする。

IRIDESシステムでは、以下の3つのフレームワークを提供する具体的なプロトコル設計を、今後、進める予定である。

- (1) 誘導空間(トラフィックエンジニアリングプレーン)におけるデータ転送方式/プロトコルを選択可能にするAPI(Application Interface)。
- (2) EIRおよびIIRの間でのデータベース同期プロトコル。 具体的に必要な作業は、iBGPにおける作業とほぼ同一であるので、iBGPベースのプロトコルアーキテクチャが現実的であろう[9]。また、大規模化(IIRやEIRの数の増大への対処)は、RR(Route Reflector)と同様の手法を導入することによって、必要となるデータ同期の量を、 $O(n^2)$ から $O(n)$ に削減させることが可能である。
- (3) ポリシー登録プロトコル。
IIRにおける制御ポリシーは、送信元オブジェクトおよび受信オブジェクト、あるいはネットワーク運用者が指定/登録可能にすることが望ましい。ポリシーの矛盾、設定に関する設定者の認証と権限の確認を行うプロトコルもあわせて必要になると考える。

3.3 問題点と今後の課題

IRIDESアーキテクチャの問題点と課題としては、以下の点が挙げられる。

3.3.1 問題点

- (1) ルーティンググループへの対処:
既存の経路制御プロトコルが決定するIPパケットの転送経路とは異なるパスを利用するので、本質的に、ルーティンググループの発生可能性がある。この問題は、類似の機能とサービスを提供するMPLSにおいて、既に、議論され、ループ検出のためのプロトコルの提案と標準化が完了している[10]。IRIDESシステムにおいても、同様のループ検出プロトコルを定義する必要があるかは、実運用の形態などを鑑み決定されるべきであろう。
- (2) エンドエンドアーキテクチャ:
IIRおよびEIRにおいては、透明なレイヤ3以上の処理を行い、エンドエンドアーキテクチャの特長を喪失していないかの検討を行う必要がある。
- (3) 中継ノードでの処理:
IIRおよびEIRで必要となる機能は、現在のルータで行われているものとほぼ同程度のものであり、IPパケットの転送パフォーマンス上の問題は、発生しないと考えられる。
- (4) エンドエンド到達性の保障:
IIRおよびEIRでは、アーキテクチャ上、NATアーキテクチャと同様に、IPアドレスの変換も可能とすることができる。これによる、エンドエンドでの到達性の喪失の可能性が発生する可能性がある。プロトコルの設計とシステムの運用においては、この点に関する十分な考慮と検討が必要であると考えられる。

3.3.2 今後の課題

- (1) 宛先のアドレス解決と表現手法
誘導空間でのEIRおよび宛先ホストへのデータ転送に必要な識別子(=アドレス)の解決手法と、その表現手法の正規化、ならびに標準化を行う必要がある。
- (2) トラフィックエンジニアリングツールの汎用化

ルーティングループの検出、転送経路の把握とエンジニアリングツールの確立など、現在のインターネットシステムでは、確立していないツール群の汎用化と整備を行う必要がある。

(3) 個別プロトコルの設計

上述した3つのプロトコルコンポーネントフレームワークの具体的な設計定義を行わなければならない。

4. IRIDESアーキテクチャの適用例

本章では、IRIDESアーキテクチャのアプリケーション(適用例)という観点から、いくつかの既存のアーキテクチャを考察する。本論文では、IRIDESアーキテクチャとして、エッジ部が既存インターネットシステムでコア部がトラフィックエンジニアリング可能なネットワークとしているが、実際には、これらは、基本的には双対の関係にすることが可能である。すなわち、上述のアーキテクチャとは逆の構成で、コア部が既存のインターネットシステムで、エッジ部が、新しいネットワーク(例えば、IPv6やプライベートネットワーク)とすることも可能である。

4.1 MIPとNEMO

文献[11]-[13]で提案されているマルチホームアーキテクチャでは、複数のHA(Home Agent)をインターネット上に分散配置し、複数のIPアドレスを持つMH(Mobile Host)あるいはMR(Mobile Router)に対して、HAから複数のパケット転送パスを設定することで、マルチホーム環境を、MHおよびMRに接続された移動ネットワークに提供するものである。複数のHAは、すべて同一の経路情報をインターネットに広告し(誘導経路情報=エニキャスト)することによって、インターネットからの誘導ポイント(IIR)へのアクセスを地理的に分散することができる。また、[14]で議論されているように、HA間では、iBGPに似たHA-HAプロトコルを用いてHAが持つデータベースの同期が行われる。

4.2 SHAKE

SHAKE(SHARed multiple paths protocol for cluster network Environment)[15]は、Mobile IPを拡張して、複数のExternalへの接続性を持つノードが、広帯域なInternalリンク(e.g., WiFiやBlueTooth)を利用しながら協力しあい、より、広帯域なExternal通信を実現するアーキテクチャである。このアーキテクチャは、図3に示した構成と同一のものと捕らえることができる。SHAKEでは、HAがIIR(誘導ポイント)に対応する。

4.3 MPLS

MPLSのエッジルータが、IIR(誘導ポイント)と考えればよい。ただし、IIRは誘導経路情報の広告を行わず、通常のレイヤ3の転送経路上で、IIR(=MPLSエッジルータ)が、IPパケットを強制的にトラップして、誘導空間に導入するアーキテクチャであると、解釈することができる。誘導空間におけるIPパケットフローの識別と転送経路の決定のための識別子は、データプレーンとコントロールプレーンに別れていて、これらのマッピングが、各ルータでとられている。データプレーンでは、MPLSラベルが、両方に共通の識別子となる。コントロールプレーンでは、IPパケットフローの識別はFEC(Forward Equivalent Class)で行われ、経路決定のための識別子は宛先のIPアドレスとなる。

4.4 NAT

エッジがPrivateアドレスを用いたネットワーク、コアがグローバルIPアドレスを用いたネットワークという構成である[16]。NATルータは、IIRにあたる。図3に示したEIRが存在する場合も、さまざまな技術を用いて実現することが可能である。一般的に、IIR(NATルータ)は、NATセグメントに対しては、グローバルには重複のある(=同一の)Default Routeの情報を広告することになる。IIRでは、IPアドレスの変換テーブルが生成される。誘導空間でのIPパケットフローの識別子はグローバルな宛先ホストのIPアドレスかEIR(=NATルータ)の誘導空間側のグローバルなIPアドレスとなる。

4.5 6to4, ISATAP

6to4[17]は、IPv4のネットワーク上で、IPv6のネットワークセグメントを相互接続しルーティングを行うアーキテクチャである。IPv6のネットワークセグメントには、IPv6のアドレスプレフィックスとともに、中継ルータのIPv4アドレスが割り当てられる。IPv4ネットワークがコア部(誘導空間)、IPv6ネットワークがエッジ部となり、6to4ルータがIIRおよびEIRに対応することになる。誘導空間でのフローの識別子は宛先ホストのIPv6

のアドレス(プレフィックス)であり、経路制御決定のための識別子は6-to-4ルータ(=EIR)のIPv4アドレスとなる。

ISATAP[18]は、IPv6アドレスフォーマットの下位32ビット中に、IPv4のアドレスを埋め込むアーキテクチャである。6-to-4と同様に、IPv4ネットがコア部(誘導空間)、IPv6ネットがエッジ部となる。誘導空間でのフローの識別子と経路制御決定のための識別子は、ともに、宛先ホストの(宛先IPv6アドレス中の)IPv4のアドレス部分である。NATやMPLSと同様に、IIRからの誘導経路の広告は行われない。

4.6 マルチキャスト

現在のインターネットは、ユニキャストサービスを基本としており、必ずしもマルチキャストサービスを本格的に提供はしていないし、ネットワーク内で使用されているルータにおけるマルチキャスト機能のサポートの状況は、同一ベンダーが提供するルータですら、残念ながら、一様ではない状況である。また、現在最も広く利用されているPIM (Protocol Independent Multicast)は、RPF(Reverse Path Forwarding)[19]のメカニズムを用いたマルチキャストパケットの転送を行っており、マルチキャストパケットの転送ツリーは、ユニキャストの転送経路から一意に決められてしまう。これら2つの問題を、IRIDESアーキテクチャを適用することで解決することができる可能性がある。PIMにおけるRP(Rendezvous Point)をIIRとし、誘導空間のネットワークポロジをマルチキャスト機能を持った適切なノードのみから構成されるネットワークにすることで、不均一機能をもつノードから構成されるネットワークにおいても、ユニキャストのパケット転送経路に独立なマルチキャストサービスが提供可能となる。

5. まとめ

現在の経路制御システムは、グローバルスケールでの、宛先アドレスをルートとしたスパンニングツリー型のパスをもとにIPパケットの転送を行っているために、必ずしも自由なパスの制御を実現することが容易ではない。本論文では、ユニキャストの概念を拡張した誘導経路情報の広告を用いたインターネットパス制御アーキテクチャ(IRIDES; Invitation Routing Information advertisement for path Engineering System)の提案を行った。提案アーキテクチャは、既存の経路制御システムとの共存が可能であり、十分なスケラビリティを持つ。さらに、提案アーキテクチャをもとに、いくつかの具体的なアプリケーション(e.g., マルチエージェント型マルチホームNEMO/MIP)が実現されることを示した。

今後は、誘導空間における汎用的な宛先のアドレス解決と表現手法と、トラフィックエンジニアリングツールの汎用化、さらに、3つのプロトコルコンポーネントの設計を行う予定である。

謝辞

本論文で議論したアーキテクチャの提案に関しては、個別に問題提起された経路制御の問題点を解決するための解決法に関する議論を、多くの方々が行う過程で、アーキテクチャが抽象化されていった。特に、インテックネットコア社の中川郁夫氏/永見健一氏、Cornell大学Paul Francis氏、東京大学吉田薫氏との建設的で具体的な議論に感謝する。

参考文献

- [1] Jon Postel, "Internet Protocol DARPA Internet Program Protocol Specification", IETF RFC791, 1981年9月.
- [2] S.Deering, R.Hinden, "Internet Protocol, Version 6 (IPv6) Specification", IETF RFC1883, 1998年7月.
- [3] P.Mockapetris, "Domain Names - Concepts and Facilities", IETF RFC882, 1983年11月.
- [4] G.Malkin, "RIP Version 2", IETF RFC2453, 1998年11月.
- [5] J.Moy, "OSPF Version 2", IETF RFC2328, 1998年4月.
- [6] D.Oran, "OSI IS-IS Inter-Domain Routing Protocol", IETF RFC1142, 1990年2月.
[7] Y.Rehker, T.Li, "A Border Gateway Protocol 4 (BGP-4)", IETF RFC1771, 1995年3月.
- [8] C.Partridge, T.Mendez, W.Milliken, "Host Anycasting Service", IETF RFC1546, 1993年11月.
- [9] T.Bates, R.Chandra, "BGP Route Reflection An alternative to full mesh IBGP", IETF RFC1966, 1996年6月.
- [10] Y.Ohba, Y.Katsube, E.Rosen, P.Doolan, "MPLS Loop Prevention Mechanism", IETF RFC3063, 2001年2月.

- [11] H.Esaki, "Multi-Homing and Multi-Path Architecture Using Mobile IP and NEMO Framework", Proceedings of IEEE/IPSJ SAINT2004, 2004年1月.
- [12] S.Uda, N.Ogashiwa, K.Nagami, K.Kondo, I.Nakagawa, Y.Shinoda, H.Esaki, "Design and Implementation of Overlaying Multi-Homing Architecture", Proceedings of IEEE/IPSJ SAINT2004 Workshop on IPv6 Technology and Deployment, pp.178-185, 2004年1月.
- [13] K.Nagami, S.Uda, N.Ogashiwa, R.Wakikawa, H.Esaki, H.Ohnishi, "Multi-homing for small scale fixed network Using Mobile IP and NEMO", (Work-in-Progress), IETF Internet-Draft, draft-nagami-mip6-nemo-multihome-fixed-network-01.txt, 2004年7月.
- [14] R. Wakikawa, et al, "Inter Home Agents Protocol (HAHA)", (Work-in-Progress), IETF Internet-Draft, draft-wakikawa-mip6-nemo-haha-01.txt, 2004年2月.
- [15] 伊藤、小山、太田、石原、"Mobile IPを用いた通信回線共有方式の提案"、情報処理学会 DICO2003 論文集、pp.97-100, 2003年6月.
- [16] K.Egevang, P.Francis, "The IP Network Address Translator (NAT)", IETF RFC1631, 1994年5月.
- [17] B. Carpenter, "Connection of IPv6 Domains via IPv4 Clouds", IETF RFC3056, 2001年2月
- [18] R.Gilligan, E.Nordmark, "Transition Mechanisms for IPv6 Hosts and Routers", IETF RFC2893, 2000年8月.
- [19] D.Estrin, D.Farinacci, et.al, "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification", IETF RFC2362, 1998年6月.