

ソース IP アドレスを考慮した経路制御システムの提案 ～地域 IX における経路制御問題の解決～

鍛治 武志†, 小林 和真†, 山口 英†, 尾家 祐二‡

† 奈良先端科学技術大学院大学

‡ 九州工業大学/奈良先端科学技術大学院大学

概要

インターネットが商用化され、利用者数が急激に増加するにしたがって、地方都市にインターネット接続サービスを提供する ISP(Internet Service Provider) の数も増加してきた。日本においては ISP が相互接続する IX(Internet eXchange) は東京や大阪に集中しているため、地方都市において IX(地域 IX) を構築しようという動きが急速に広まってきている。本稿では地域 IX の抱える問題の中でも経路制御に関する問題を明らかにし、その解決策として“ソース IP アドレスを考慮した経路制御システム”を提案する。提案するシステムは通信経路を決定する時点においてソース IP アドレスを考慮することにより、最適な経路選択を行うシステムである。また、提案した方式にもとづいて経路制御システム(STAR)を実装し、その有効性の評価を行った。

A proposal of routing system that depends on source IP address ～ Solution of routing issues on the regional IX ～

Takeshi Kajit†, Kazumasa Kobayashi†, Suguru Yamaguchi†, Yuji Oie†

†Nara Institute of Science and Technology

‡Kyusyu Institute of Technology/Nara Institute of Science and Technology

Abstract

Since the Internet is commercialized, the number of the Internet users is increasing drastically. ISPs(Internet Service Provider) that provide the Internet connectivity for the provinces also increase in its number simultaneously. As the Internet covers all over the Japan as its information infrastructure, many provinces provide information by means of it. ISPs connect to IX(Internet eXchange) to communicate with each other. But there are IXs only in Tokyo and Osaka in Japan. This is the reason why provinces wants their own IX in their regions individually. In this paper, we focus on the routing issues about regional IX and propose a routing system that depends on source IP address. We also implement the prototype of the routing system called STAR(Source address oriented Traffic Arrangement Router) and evaluate its performance in this paper.

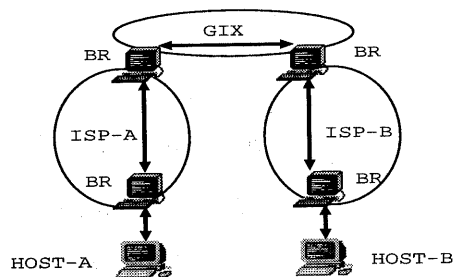
1 はじめに

WWWによる情報発信や電子メールなどの利便性が広く世間に認められるにしたがって、それらを利用するためにインターネットに接続する人の数が急激に増加している。それに伴いインターネット接続サービスを提供するISP(Internet Service Provider)の数も増加しており、インターネットを利用した情報インフラストラクチャの整備が各地で盛んに進められている。このような背景のもと、現在では地方自治体や地方公共団体、または地方の企業によるインターネットを利用した情報発信が数多く行なわれている。また、地域インターネット網の整備は情報化社会に向けた地域経済の産業振興策として各方面から期待が寄せられている。

一方、インターネットは自律系(Autonomous system, 以下AS)を基本単位としたネットワークの集合である。AS間の通信を行うために各ASを相互接続するポイントとしてIX(Internet eXchange)が存在する。現在運用中のIXには、WIDE ProjectによるNSPIXP[1]、商用IXのMEXやJPIXが挙げられるが、これらはGIX(Global Internet eXchange)と呼ばれており、東京や大阪という大都市に集中している。すなわち、現在のインターネットの設計は大都市を中心とした設計であり、このことは地方都市においてインターネットを利用する上で、通信遅延の発生、情報コンテンツの大都市への集中、他地域の災害の影響を受ける、地域内の通信路に関する決定権が地域にないなど様々な問題を抱える原因になっている[2]。

これらの問題はどの地方都市においても共通しており、地域内通信の効率化と安定化、地域に関する情報の地域からの発信、他の地域の障害に強いネットワーク構築、地域コミュニティの形成、通信路の決定権の取得などの目的を達成するために地方都市にもIXを構築しようという動きが広まっている。このようなIXはGIXに対してRIX(Regional Internet eXchange, 地域IX)と呼ばれ、既に岡山県[3]、山梨県、富山県、東海地域、東北地域などにおいて相互接続実験が行なわれている。

本稿では、地域IXが抱える経路制御に関する問題に着目し、その解決策である、“ソースIPアドレスを考慮した経路制御システム”を提案し、これに基づいた実装(STARシステム)について述べ、その有効性について考察する。



BR:Border Router

図 1: GIX を経由する経路

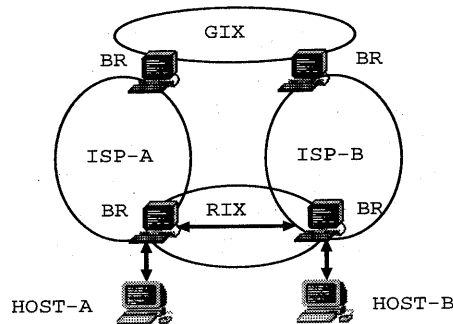


図 2: RIX で折り返す経路

2 地域IX(RIX:Regional Internet eXchange)

本章では、まず地域ネットワークにおける経路制御の重要性とその形態について述べる。また、現在のインターネットの技術的制約を説明し、地域IXにおいてGIXで取られている手法を適用できない理由を説明する。そして、様々な制約がある中で地域IXを実現するための方法として、“ソースIPアドレスを考慮した経路制御システム”を提案する。

2.1 地域ネットワークにおける経路制御の重要性

同じ地方都市にあるホスト間の通信に着目してみる。その際、それらが接続しているISPが互いに異なっている場合には、現在それらのホスト間の通信はGIXを経由して行われている(図1参照)。地方都市においてIXを作り、そのIXにおいてISPを相互接続すればその地域内での通信はその地域内で折り返すことができる(図2参照)。地域内の通信を地域

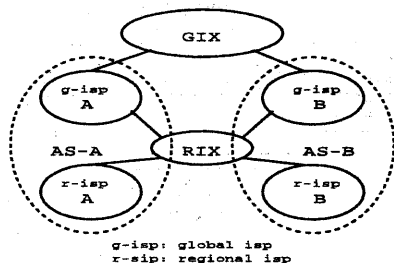


図 3: 地域 IX の形態

内に閉ることができれば、前章で述べた問題を解決することができる。このような理由から地方都市に IX を構築しようという動きが急速に広まっている。

2.2 地域 IX の形態

現在、相互接続実験を行なっている地域 IX には様々な形態が存在する [2]。その 1 つとして地域 IX 自体が外部との接続性を持ち、ISP の様な役割を果たす形態がある (図 3 参照)。この形態では地域 ISP (r-isp: regional isp) はそれぞれ大手 ISP (g-isp: global isp) と契約しインターネットとの接続性を確保する。地域 ISP は大手 ISP との接続を RIX を経由して行う。この場合、経路的には地域 ISP は大手 ISP の AS に含まれることになる。この形態の利点は、地域 ISP 間 (r-ispA, r-ispB) の通信を RIX を経由する最短経路で行うことができることである。地域内の経路情報は RIX 内部で閉じることができるので、大手プロバイダ間の経路交換を RIX で行う必要は特にない。この形態は岡山県が推進している岡山情報ハイウェイにおいて取り組まれている。

2.3 GIX における経路制御

現在 GIX においては、ISP 間で BGP4 により Peer を張り相互接続している。同一 ISP 間で複数の境界において Peer を張ると、複数の経路情報が交換されることになる。そうした場合、ISP は経路情報を集約し最適化してから使用する必要がある。しかし、この最適化を行なうためのコストは経路情報の数が増加するにつれ大幅に増加する。すなわち、この手法には拡張性がない。よって、RIX においても GIX と同様の手法によって相互接続することは、今後地域 IX が増加することを考慮すると現実的ではない。

また、図 3 の様な地域 IX の形態において、g-ispA と r-ispA、g-ispB と r-ispB がそれぞれ契約している場合、大手 ISP は契約した地域 ISP のパケットのみを伝送したいのであって、契約していない地域 ISP のパケットは伝送したくない。しかし、地域 IX において GIX と同様にすべての経路情報を交換すると、契約していない地域 ISP のパケットも伝送する可能性が出てくる。

このような問題を解決するために、地域 IX における新しい経路制御技術が必要である。

2.4 地域 IX の実現

地域内通信を地域内に閉じることを目的とした地域 IX を構築する上での技術的制約は以下の通りである。

- 制約 1 同一 AS 間での Peer の数が増加しない。
- 制約 2 AS 番号は有限であるため AS は取得しない。
- 制約 3 経路情報の伝送が可能である。
- 制約 4 ネットワーク管理のオーバーヘッドが少ない。
- 制約 5 高度なネットワーク管理技術を必要としない。
- 制約 6 地域 ISP ごとに大手 ISP の選択が可能である。

図 3 のような形態の地域 IX における経路制御の実現法として現在の技術で考えられる実現法を挙げる。[実現法 1] として地域 IX を 1 つの AS として運用する手法が考えられる。この手法では上記の制約 2 を満足できない。これは AS 番号は 2^{16} しかないため、今後地域 IX が増加することを考えると AS 番号の枯渇という問題が生じ拡張性がないためである。[実現法 2] として地域 IX がプライベート AS を使用する手法が考えられる。この場合 AS を取得しないため AS 番号の枯渇という問題は解決される。しかし、制約 3 を満足できない。すなわち、地域 IX の下流にさらに AS をつくった場合、プライベート AS ではトランジット AS になれないという問題がある。[実現法 3] として ISP 間で経路を静的に設定する手法が考えられる。この手法では制約 4 を満足できない。すなわち、経路を静的に設定すると他の ISP のネットワークトポロジが変更されるたびに、管理者は自分の管理するルータの経路の設定を行なわなければならない。ネットワーク管理者は他の ISP のた

めにルータの設定をし直す必要が生じるため、実際の運用環境においては現実的でない。[実現法 4]として IP トンネリングを使用する手法が考えられる。しかし、この手法では制約 5 を満足できない。なぜなら、IP トンネリングを行なう ISP の数が増加すると、ネットワークの構成が複雑になり管理が大変になるためである。さらに、トンネリングを行なう ISP 内部での経路制御も複雑になる可能性がある [4]。

以上、現在の技術を用いた [実現法 1]～[実現法 4]では制約 1～6 のすべてを満たすことはできない。上記すべての制約を満たすために、本研究では“ソース IP アドレスを考慮した経路制御システム”を提案する。この手法は、まず地域 IX より下流では IGP (Internal Gateway Protocol) により経路を交換するため、地域内の通信は地域内で閉じることができる。また、 $r\text{-ispA}$ は AS-A に、 $r\text{-ipsB}$ は AS-B に含まれるため AS の数は増やす必要はない (図 3 参照)。よって制約 1, 2 は問題にはならない。次に、 $r\text{-isp}$ の下流に別の AS を作ったとしても経路情報を伝送できるため制約 3 は満足できる。また、経路情報を動的に交換するため制約 4 は満足できる。現在あるシステムやプロトコルなどを全く変更する必要がなく、経路制御システムを置き換えるだけなので制約 5 も満足できる。最後に、提案する手法は経路制御システムにおいてソースアドレス集合を定義し、そのアドレス集合ごとに経路表を持つ。そしてソースアドレス集合ごとに専用の経路表を検索して、目的ホストまでの経路を選択するために制約 6 も満足できる。よって、提案する手法は制約 1～6 までのすべてを満たすことができる。

3 ソース IP アドレスを考慮した経路制御システム

本章では、パケット転送メカニズムについて述べ、提案する経路制御システムの実現方法とその効果について述べる。

3.1 パケット転送メカニズム

経路制御システムにおいてパケット転送時のメカニズムを図 4 に示す。経路制御システムは受け取ったパケットの宛先を判別し、自ホスト向けのパケットはそのままトランスポート層に渡す。他ホスト向けのパケットは関数 $ip_forward$ において経路表を探索

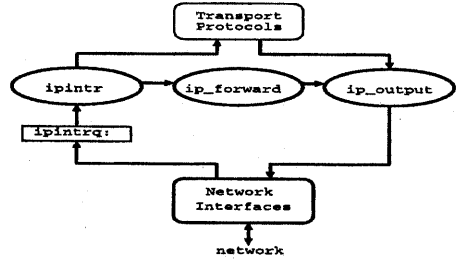


図 4: パケット転送メカニズム

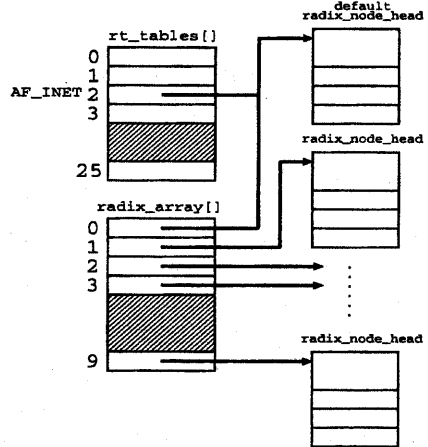


図 5: 複数の経路表

することによって次の宛先を決定する。

3.2 経路制御システムの実現

提案する経路制御システムでは、各 ISP に属するホスト群をアドレス集合として定義する。経路制御システムにおいては、1つのアドレス集合に1つの経路表という様に、アドレスファミリー AF_INET に対して複数の経路表を持つようにする (図 5 参照)。

そして、経路選択時、すなわち関数 $ip_forward$ においてパケットのソース IP アドレスを調べ、どのアドレス集合に属するホストから来たパケットかを判断する。その後、複数ある経路表の中から対応する経路表を選択し、選択した経路表から次のパケット転送先を決定する。この様に、アドレス集合ごとに経路表を選択する。

3.3 提案する経路制御システムにより得られる効果

- E1. 冗長な経路の短縮 地域内の通信は,GIX を經由することなく地域内で折り返すことができる.
- E2. 伝送回線の選択 回線には伝送速度や伝送コストの違う回線がある. 目的ホストまでの伝送回線を様々なポリシーにもとづいて選択できる.
- E3. ソースアドレス詐称の禁止 経路制御システムに予め登録していないアドレス集合から来たパケットは転送しない様に設定できる. よって, ソースアドレスを偽った通信を禁止できる.
- E4. 階層構造 地域 IX 内に, さらに狭い地域を対象とした地域 IX を作る事ができる. すなわち, 地域 IX を階層的に作り,E1~E3 までの効果を得ることができる.

4 経路制御システム (STAR)

既存の経路表はカーネル内にあるため, 経路表を複数もつ構成に変更するにはカーネル内に組み込む方式が適切であると考えられる. したがって, 本システム “Source address oriented Traffic Arrangement Router(STAR)” は FreeBSD ver2.2.1 のカーネル内に組み込む形で実装した.

4.1 カーネルへの組み込み

複数の経路表へのポインタを保持するために1次配列 (radix_array[]) を新たに設けた (図5参照). 経路表は最大10個まで持つことができるようになっているが, 想定するアドレス集合の数が変化すれば, カーネル内に定義されているマクロ (RADIX_MAX_NUM) を変化させることによって変更可能である.

現在の実装では, rt_tables[2] に保持している経路表へのポインタと, radix_array[0] に保持している経路表へのポインタは同じである. このポインタが指している経路表をデフォルトの経路表と呼ぶことにする. 現在の実装ではパケット転送時において, ソース IP アドレスが含まれるアドレス集合が定義されていない場合, デフォルトの経路表を検索し転送先を決定するようにしている.

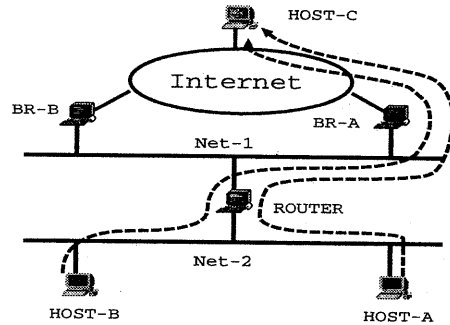


図 6: 通常のルータの場合

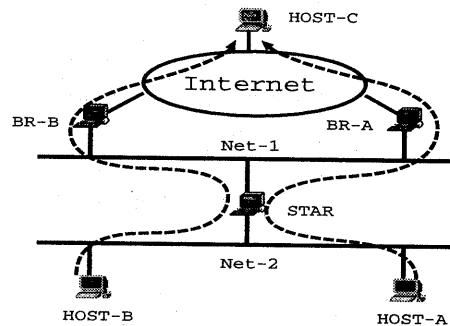


図 7: STAR システムの場合

4.2 route, snetstat, sifconfig コマンド

管理者が経路表に経路情報を追加あるいは削除するためには route コマンドが使用される. しかし, 既存の route コマンドでは複数ある経路表のどの経路表に対して経路情報を追加あるいは削除をするのかを指定できない. また, 新たな機能として経路表とアドレス集合の対応を指定する必要もあるため, sroute コマンドを用意した. また, 経路制御システムが保持している経路表を全て表示するために snetstat コマンドを, 経路表ごとにインターフェイスの設定ができるように sifconfig を用意した.

5 評価

5.1 実験環境

実験ネットワークとして Net-1 と Net-2 を使用し, ネットワーク間には ROUTER を設置する (図6参照). ROUTER のデフォルトルートは BR-A のため, HOST-A と HOST-B が HOST-C と通信する場合は BR-A を経由して行なわれる.

```

Radix tree:   No.1
Address group: HOST-A
Destination   Gateway   Flags
default       BR-A     UGSc

Radix tree:   No.2
Address group: HOST-B
Destination   Gateway   Flags
default       BR-B     UGSc
...

```

図 8: snetstat の出力例

Date Size	512B	1024B	1472B
FreeBSD	998	658	498
STAR	984	652	494

表 1: パケット転送能力比較 (単位:pps)

5.2 STAR システムの効果

HOST-A と HOST-B 専用の経路表を STAR システムに作成し (図 8 参照), デフォルトルートをそれぞれ BR-A, BR-B に設定する. すると, HOST-A の通信は BR-A を経由し, HOST-B の通信は BR-B を経由することが分かる (図 7 参照).

5.3 パケット転送能力

FreeBSD ver2.2.1 と, その上に実装した STAR システムとの間でパケット転送能力を比較した (表 1 参照). ただし, STAR システムの経路表の数は 10 個の場合である. この測定は Netperf[5] を使用し, 出来限り多くの UDP パケットをホスト間で送信し, STAR システムにおいて転送されるパケット数を計測した. また, データサイズは 512Byte, 1024Byte, 1472Byte の 3 つの場合について測定した. 表 1 からパケット転送能力はほとんど変化しないことが分かる.

5.4 安定性

STAR システムの安定性を実証するため研究室の有志数人に依頼し, このシステムを使用してもらっている. 現在, 運用実験を開始してから約 3 週間経つが, 目立ったトラブルもなく運用している.

6 今後の課題と考察

経路表へのポインタの探索は 1 次配列の線形探索のため, 経路表の数が増加すると経路表へのポインタを探索する時間は $O(n)$ で増加する. よって, 将来的には経路表へのポインタも radix tree に格納する予定である. この場合, 必要となる経路探索時間は経路表へのポインタを探す時間 $O(\log n)$ と実際に経路表で経路を探索する時間 $O(\log n)$ の和であり, $O(\log n)$ で抑えることができる.

また, 現在の実装ではデフォルトの経路表への経路情報の追加, 削除はルーティングデーモン (gated, routed) が, それ以外の経路表への経路情報の追加, 削除は管理者が行っている. 将来的には経路情報を動的に交換することによって, すべての経路表を自動的に更新できるようにする予定である.

7 まとめ

地域内の通信を地域内に閉じるために地域 IX を構築しようという動きが急速に広まってきている. しかし, 地域 IX において ISP を相互接続するためには様々な問題を解決する必要がある.

本稿では地域 IX における経路制御の問題に焦点を当て, 現在考えられている解決策を比較検討した. そして, 経路制御システムが経路表を複数持ち, パケット転送時にソース IP アドレスごとに使用する経路表を選択するシステムを提案した. 最後に, 提案にしたがって STAR システムの実装を行ない, その有効性の評価を行なった.

参考文献

- [1] WIDE Project: "WIDE/NSPIX Home Page", <http://xroads.sfc.wide.ad.jp/NSPIX>
- [2] 中川郁夫, 米田政明, 安宅彰隆: "国内における地域 IX の動向", 情報処理学会 分散システム運用技術 研究報告 No.7, 1997
- [3] OKIX: "岡山情報ハイウェイホームページ", <http://www.okix.or.jp/>
- [4] 今野幸典, 桶地正浩: "プライベートなインターネット エクスチェンジを実現する経路制御手法の提案", 情報処理学会 分散システム運用技術 研究報告 No.7, 1997
- [5] Hewlett-Packard Co.: "Netperf Home page", <http://www.cup.hp.com/netperf/NetperfPage.html>