

研究用 Internet バックボーンの運用について (省際研究情報ネットワークを例にして)

鈴木亮一¹ 福田晴元¹ 三上博英¹

ryoichi@slab.ntt.jp, harumoto@slab.ntt.jp, mikami@slab.ntt.jp

NTT ソフトウェア研究所

研究用インターネットバックボーンである省際研究情報ネットワーク (IMnet) の運営に関して報告する。IMnet は、独立した運営を行なっている研究試験機関の LAN(WAN) を統合し他のインターネットバックボーンと接続を行なっている。本稿では、IMnet の運営を通して得られたインターネットバックボーン維持の問題点について、経路制御の観点から、他ネットワークとの接続、一般組織との接続、バックボーン自身について問題を指摘した。また、今後のインターネットに必要な技術について考察した。

A case study of Internet backbone Operation (Inter-Ministry Research Information Network)

We describe the case study of the academic Internet backbone, IMnet(Inter-Ministry research information network). IMnet connects many Japanese national institutes, laboratories, and related organizations to Internet for research. We show problems with connecting to other ISPs, subscribers, and the IMnet backbone at the routing point of view. We also discuss about future technology for the Internet.

1 はじめに

インターネットはここ数年で大きく発展した。日本においても、数多くの学術プロバイダ、地域プロバイダ、商用プロバイダが生まれ、相互接続を行なっている。プロバイダの数が増えるにつれ、相互接続における技術的複雑さは増す。各プロバイダは自分のポリシー (Acceptable Use Policy:AUP) を持ち、それに従って接続されるためである。

また、プロバイダに接続される組織内ネットワークにも、AUP もしくはそれに近い運用ポリシーがある場合がある。また、組織内ネットワークの事情から、プロバイダに複雑な要求が上がる場合がある。

このような状況で、インターネットプロバイダの経路制御に関する要求は、日々複雑になっている。本稿で

は、そのような複雑な経路制御に関する問題について、考察する。

なお、本稿では、ISP はインターネットに対して複数の物理接続点を持っているものを前提とする。

2 省際研究情報ネットワーク

省際研究情報ネットワーク (Inter-Ministry research information Network:IMnet)^{[1][2]} は、科学技術庁の科学技術振興調整費によって、平成6年より実験を開始したインターネットサービスプロバイダ (ISP) である。これは、国公立研究試験機関ならびに、支援機関を結ぶ実験ネットワークとして設立された物である。バックボーンの運営は、KDD 研究所を中心とする経路制御分科会、

¹Ryoichi SUZUKI, Harumoto Fukuda, and Hirohide MIKAMI
NTT Software Laboratories
Midori-cho 3-9-11, Musashino, Tokyo.180 Japan

NTT ソフトウェア研究所を中心とする整備・運用・管理分科会の二つの分科会のもと、行なわれている。

平成6年11月より実験を開始し、そのバックボーンを使って経路制御、運営管理の実験を行なうとともに、多くの国立研究試験機関をつなぐインターネットサービスプロバイダとして運営されている。

現在IMnetに接続されているのは、110ドメインとなっている。これらは、IMnetの実験に参加するという位置付けで接続しており、参加組織と呼ばれている。

3 IMnet のトポロジー

IMnet のトポロジーを図1に示す。

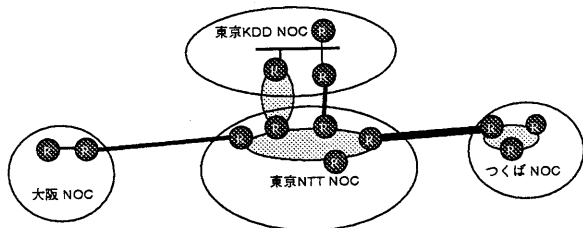


図1 IMnet のバックボーン

IMnet は、東京、つくば、大阪に NOC がおかれている。

東京とつくば間は 45Mbps の専用線接続、東京と大阪は 3Mbps 専用線接続となっている。

東京は、他の ISP との接続行なっている KDD NOC と、一般組織との接続を行なっている NTT NOC がある。

KDD NOC 内は Ether Switch、NTT NOC 内は FDDI を使用している。KDD NOC と NTT NOC の間は 3Mbps の専用線と 100Mbps の実験用光ファイバで接続されている。

つくば NOC 内は FDDI を使用している。

大阪 NOC 内は Ethernet を使用している。

各参加組織は、事情に応じて 64Kbps から 150Mbps までの速度で最寄りの NOC に接続されている。

このようなルータ配置構成は、

- 時間経過

当初の計画と異なった参加組織の増加や、回線速度の変更によるルータ負荷の分散を図ったため

- 予算

国家予算のため計画通りの支出が求められ、予定外の事態に対応が困難

などの制限による。

3.1 IMnet における問題

3.1.1 回線速度のアンバランス

東京つくば間、東京大阪間のバックボーンアンバランスにより、トラフィックがより細い線によって律速されている。

また、参加組織は細い線によって接続されてきた場合、バックボーンからのトラフィックを処理しきれない場合がある。

3.1.2 トポロジーのアンバランス

東京を中心としたスター型のため、経路的冗長性が無い。

高速度の専用線敷設にともなったルータ増強が困難なため、小数のルータをパフォーマンスに応じて配置変えをおこなった。その場合には、大規模な回線の別ルータへの収容変えが伴うこととなった。また、このような場合でもバックボーンを停止することをなるべく避けるため、2回線同時に接続されている時期もあることなどから、ルータの役割が時系列で変化し、ルータ相互の結合を階層化が保てなくなっている。

このため、経路制御に制限がかけられている。

4 経路制御

今日、インターネットに対して複数の接続をもつ ISP は AS 番号 (Autonomous System 番号、自立システム番号) をもち、BGP4 (Border Gateway Protocol version 4)^[3] と呼ばれる EGP (Exterior Gateway Protocol) を使用して接続される。

IMnet が運用を開始した時点では、国内の多くの学術 ISP は RIP (Routing Information Protocol) を使って

接続されていた。しかし、RIPの限界から、ASの概念を使用し整理された経路制御が必要とされ、RIPからBGP4へのプロトコル変更が行なわれる時期とほぼ同時に、IMnetはインターネットの一部となった。

BGP4の使用は、以下のメリットをもたらした。

- ISP境界の明確化

ASの導入により、ISP内部の経路制御と、ISP外部との経路制御が論理的に分割できるようになり、ISP内部での経路制御が容易となった。

以前は、日本全体としてネットワーク・マップを作成し、それよりRIPのメトリック値を調整することで接続されていた。そのため、ローカルなメトリック値の調整が大きな影響を及ぼす可能性が高かった。

- Classlessルーティング

未割当IP address空間の減少のため、従来Class A/B/Cとわけられていたネットワークの単位を、より実情にあった単位で割当が始まった。

また、インターネットの全経路数が増え、またそれにもなって経路のふらつきなども激しくなった。そのため連続したネットワークアドレスを集約することで、見かけの経路数の削減、小さい経路のふらつきの隠蔽を図った。

これらに対応するためのルーティングがClasslessルーティングである。

- 経路情報の詳細化

古いEGPは、ツリートポロジーを前提としていたが、BGPは経路情報をリレーしてきたASの情報がメトリックとしてありループフリーとなっている。またそのメトリックを利用して、より細かい経路制御が可能となっている。

IMnetでは、当初よりBGP4を使用し、JOIN、SINET、TISN、WIDEなどの国内プロバイダ、NSI、Sprint、KREOnetなどの海外との接続を行なっている。

そのため、AS境界ルータは各NOCに最低一つは配置されAS間ルーティングを担当している。内部的に

はそのルータ間はiBGPでメッシュを張ってルーティングされている。

ASのIGP(Interior Gateway Protocol)として、OSPF⁴を使用し、すべてのルータがIMnet内の経路情報を交換している。

これは、

- Classlessルーティングを行なうIGP

当時利用可能なのは、Cisco社のEIGRPとOSPFのみであった。

- RFCになっていた。

- BGPとのインタラクションが定義されていた。

より採用が決定された。

IMnet内Classlessルーティングを行なったのは、参加組織から一つのネットワークアドレスをサブネットワークに分割し、それぞれ別のNOCに接続するという要求があったためである。

4.1 階層的経路制御

基本的にISPでの経路制御は、EGPに関する問題と、IGPに関する問題がある。

- EGP

1. 外部接続をどのように使い分けるか
2. 他ASの通過バケットを許すか

- IGP

1. 内部のどの線をどのように使うか。
2. 参加組織との接続用プロトコル

これらの問題は常に、独立して現れるのではなく、お互いに関係しあっている。

しかし、これらの問題を論理的に分割し解決を容易にするため、階層化の手法を取る。具体的には、

1. インターネットの全経路をもつルータ群
2. 自ASの経路をもつルータ群
3. 直接接続した組織の経路情報をもつルータ群

の3階層にわけろ。それぞれ、

1. EGP の問題 1, EGP の問題 2
2. IGP の問題 1, EGP の問題 2
3. IGP の問題 1, IGP の問題 2

に分割する。これによって問題を分割、システムとしての管理を用意することを目指す。

しかし、IGP と EGP の問題の完全な分割は困難である。ルータとリンクに余裕を持たせることが可能な場合には、冗長な経路になることもあるが、ある程度分割可能である。

また、参加組織との接続は、一般的な IGP もしくは静的な設定を利用する。これは参加組織側の事情により、いくつかの選択肢を用意すべきである。

このような場合、IGP を複数の種類使用するため、IGP 同士のインタラクションも設計項目に入れる必要がある。

4.2 IMnet における問題

4.2.1 Classless ルーティングに非対応

IMnet の運営が始まった当時は、参加組織などに、classless ルーティングに対応していない機器が多く残っていた(現在でも一部残っているところもある)。

また、一部の参加組織は中間に別のネットワークが入って間接的に IMnet 接続されているところがあり、その中間ネットワークが classless ルーティングをサポートしていない場合もあった。

そのため、東京とつくばに分かれている研究所が同じネットワーク番号を使う場合に問題があり、その研究所が一時内部リンクをもつことで一時的な解決図ったなどがあった(静的経路を利用するなど他の解決法もあるが、マジックナンバーが入るため管理が複雑になる)。

しかし、classless な経路制御を行なう OSPF の設定に比べ、RIP の設定は単純であるため、Class A/B/C の単位での経路制御を希望するような参加組織には RIP を接続プロトコルを希望するところが少なくない。

4.2.2 階層化の崩れ

運営を継続するにつれ、接続組織が増えてきた。そのため、大阪、つくばの NOC の AS 境界ルータにも参加組織が接続されることができた。

そのため、参加組織が収容ルータによって、経路が大幅に異なる場合が出てきている。

例えば、WIDE とは東京ならびに大阪で接続されているが、大阪の WIDE との接続に使われているルータに収容された参加組織からの WIDE へのパケットは、大阪からでることになり、それ以外のルータに収容された組織のパケットは東京から WIDE にでるといふことが起こった。

5 マルチホーム接続

ISP が増加とともに、様々な事情からマルチホーム接続を希望する参加組織が増えてきた。マルチホーム接続とは、一つのネットワークが複数の ISP と同時に接続を持つことである。

マルチホーム接続は、

- トポロジーが複雑になる。
- 使いわけをはっきりと決定する必要がある。

複数の ISP に対して、それぞれをどのような位置付け(例えば、主に使う ISP とバックアップの ISP、国内向けの ISP と海外向けの ISP など。もっと複雑な場合もある)を決める必要がある。ルータの設定は、この決定に基づいて行なわれる。

しかし、技術的制限から、この決定が必ずしも希望通りになるとは限らない。

例えば、ISP A には 1.5Mbps、バックアップ用に ISP B に 128Kbps で接続をしたと仮定する。バックアップがうまく働くためには普段は ISP から ISP B から経路情報を広報してもらう必要がある。その結果、ISP B から入ってくるパケットが出てくる。このように速い ISP A をうまく使えない可能性が出てくる場合がある。

- トラブルの切りわけが難しい。

インターネットは、Hop-by-Hop のパケットフローディングを行なっているため、行きと帰りの経

路が異なる。トポロジーが複雑になればパケットが通る可能性がある経路が増え、トラブル時には切り訳が難しくなる。

- 経路制御プロトコルが、マルチホーム接続を考慮していない

現在の BGP は、マルチホーム接続を扱う機能が無い。そのため、マルチホームを行なうネットワークに対して、多くのプロバイダは特別の設定を行なう。

このような特別な設定は、設定情報の維持管理のうえで問題となる場合が多い。

などの理由から、単一の ISP に接続されている場合に比べて管理が困難になる。また、マルチホーム接続された経路は不安定になりやすいという報告^[5]もある。そのため、マルチホーム接続にはより高い技術と、自組織と接続された ISP、さらにはその接続された ISP 同士が密に連絡を取り合っている必要がある。

以下にトラブルの例を示す。

5.1 事例 1:

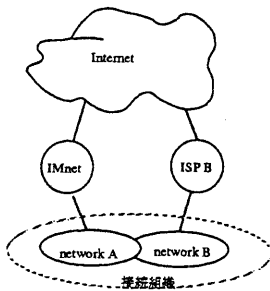


図 2 マルチホームの例

図 2 のように接続された組織から、通信ができないとの報告があった。

原因は、network B の経路情報を IMnet も ISP B も他 AS に広報していたことによる。

インターネット上、遠くにあるネットワークが network B への経路として IMnet と ISP B のどちらかを選ぶのは一般的に困難である。

その組織は外部向けのルータに firewall の設定を施しており、それぞれ直接接続指定のネットワークへのパ

ケット以外は filter out していた。今回は、network B のパケットが IMnet を経由したため、パケットが廃棄されていたためであった。

今まで ISP B 経由だった経路がなくなかはずみで、IMnet 経由になったため、原因がすぐには分からなかった。

接続組織は、正しく自分の希望を伝え、各 ISP に設定を要求する必要がある。

5.2 事例 2:

IMnet では自バックボーン接続データベースをもとに、各ルータの経路制御表の sanity check を定期的に行なっている。

ある日突然、ある経路が日に何回となく up-down を繰り返すため、BGP の広報すべき経路情報から外れてしまった。しかし、通信出来ないという報告はなかった。

後日、そのネットワークが他の ISP とマルチホームを既に行なっているとの報告を受けた。

6 今後の技術

以上、IMnet の経路制御とその問題点について述べてきた。

以下に、今後 ISP を運営するにあたって必要となるであろう技術についての考察を行なう。

6.1 変更に柔軟に対応できる設定

- 高速なルータについて

“高速なルータ”には以下の意味がある。

- 高速なスイッチングの出来るルータ

今後、ますますリンクの速度が上がる事が予想される。そのリンクが複数接続された場合に対応出来るバス、CPU、ソフトウェアを兼ね備えていること。

- 高速なルーティングの出来るルータ

経路情報も増え、ネットワークトポロジーも複雑化するにつれて経路のふらつきも出てくる。5 万以上の経路情報を余裕を持って扱えること。

余裕をもったルータがなければ、ルータの配置に関しての問題が複雑になり、初期設計に変更を加える場合に大きな障害となる。

● 柔軟な設定の出来るルータ

設定は、かならず大きな変更が必要になる。これは ISP 自身の問題よりも、他 ISP との関係、接続組織との関係の問題の方が多。

そのため、条件の変更にともなった設定変更に対応出来るように設計するとともに、ルータ自身が設定変更に対応出来る設計である必要がある。

以下にその要件を挙げる。

－ 強力な設定環境

設定確認、変更の際にはエラーが予想される。それを可能限りに防ぐためには、強力なエディタ、デバッガ、さらには、実行前に確認するシステムなどが必要である。

実際には、ネットワークを止めずに、または止める範囲を最小限にして設定変更が行なわれるので、シミュレーション環境、テスト環境があり、その結果を実ネットワークと正確に反映されるシステムが必要である。

－ 設定構文の構造化

gated などでは設定構文が構造化されているが、専用ルータでは構造化されていない(あるいは少ない)。同じ文を各インターフェースに設定する時などには、問題がおきることがある。

－ 設定の共有化

多くのルータは設定を共有しておくべき部分を持っている。たとえば、filter 用のリストなどがその例である。数十台のルータに同じ変更を加えるのは効率が悪い。共通部分は 1 箇所に集中管理される必要がある。

－ 安全な設定変更

一部のルータでは、入力した直後にそれが有効となる場合がある。これが必要な場合と危険な場合がある。設定変更の場面に依じて、

変更が有効になるタイミングを選択出来るようにするべきである。

設定の維持管理を容易にするために、ルータ機能のモジュール化と標準化を支援する機能が必要である。

6.2 プロバイダと接続組織の連絡

管理運営体制も重要であることが、マルチホームのトラブルの例からも挙げられる。前出のように、問題はますます複雑化し、日々状況が変化している。トラブル解決には、正確で最新の情報が必要である。

そのため、プロバイダ間、プロバイダと接続組織間の連絡体制を整える必要がある。またそのチャンネルを通して、プロバイダや接続組織の計画工事/事故などを伝えることで、問題が大きくなるのを防ぐことが出来る。

また、現在の技術で何が容易で何が困難かを整理しておく必要もある。

7 まとめ

今後は、今まで得られた知見を元に、

- 頑強で現実に適したルーティング・プロトコル
マルチホームなども考慮に入れた経路制御技術
 - ネットワークの自動診断の研究
ネットワークの変更・事故を監視するシステム
- の研究を行なう予定である。

参考文献

- [1] 福田・鈴木・三上, “省際研究情報ネットワークにおける接続形態”, 第 51 回情報処理学会全国大会, 1-101
- [2] 鈴木・福田・三上, “省際研究情報ネットワークの構成について”, 第 51 回情報処理学会全国大会, 1-103
- [3] Y. Rekhter, T. Li, “A Border Gateway Protocol 4 (BGP-4)”, rfc1771
- [4] J. Moy, “OSPF Version 2”, rfc1583
- [5] 上水流由香, “安定性を重視したマルチホーム接続について”, 第 51 回情報処理学会全国大会, 1-143