第 7 章　アドレス資源管理と経路情報の現状

内容

- インターネット経路制御
- BGP-4
- 最新経路情報
- アドレス資源管理と経路
- インターネットルーティングレジストリ

ほか

7. アドレス資源管理と経路情報の現状

本章では、本調査を実施する上で基礎となる情報を解説する。特に、IP アドレスの資源管理方法やインターネットの経路制御の本調査報告書執筆時点における現状についてまとめると共に、本調査で必要となる基礎知識について解説する。

本章では、まずインターネットにおける経路制御の考え方と経路制御方式について解説し、インターネット全体における経路制御を行うための経路制御プロトコルである BGP-4 について、簡単な方式の説明と動向について解説すると共に、現状の経路制御の実態を報告する。

次に、IP アドレス資源の管理方式について解説する。IP アドレスの資源管理と管理されている IP アドレスがインターネットでどのように経路制御されるかについての関連性について解説する。

次に、インターネット経路制御における問題点を整理し、その実態について報告する。

最後に、これら問題点に着眼し、インターネット経路制御を行う上で、本調査の核となる「インターネット・ルーティング・レジストリ」および「IP レジストリシステム」について解説する。

7.1. インターネット経路制御

本節では、インターネットの経路制御について、最初に基礎的な部分を解説し、次に経路制御を実施する上での単位となる経路制御単位(Routing Domain)について解説する。

最後に、会社単位などの組織内での経路制御方式とインターネット全体の経路制御方式の違いを明らかにし、そこで使われている実際の経路制御プロトコルについて解説する。

7.1.1. 経路制御基礎

インターネット上を流れる一連のデータは、パケットといわれるデータ単位に分割され、送信元から送信先にインターネット網上を利用して送信される。このとき、それぞれのパケットには、「送信元アドレス」と「送信先アドレス」が記録されており、インターネット網を構成する「ルータ」は、「送信先アドレス」よって送信先のデータ回線を選択し、目的の送信先端末までパケットを送り届ける。このときの「ルータ」は、電話回線における「交換機」に相当する装置と考えて良い。

この様子を図 7-1 に示す。

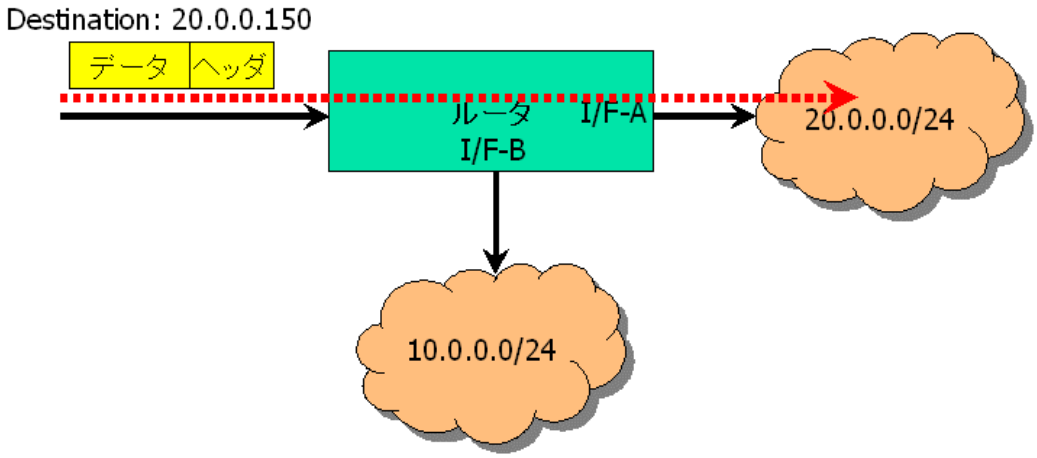


図 7-1　ルータの基本的パケット転送の様子

ルータでは、受信したパケットを、そのパケットの目的地に向かう最良の回線を選択し、その回線にパケットを送信する責務を負っている。この送信回線の選択には、「経路表」といわれる、送信先のリストを使って行われる。実際には、ルータは送信先アドレスの固まりである

プレフィックスとそのプレフィックスに該当するパケットを受信した際の最良の転送先ルータアドレスが書かれた「経路表(Routing Table)」と言われるデータベースを持ち、ルータではこのデータベースに従って経路の選択が行われる。これらの選択は通過するルータで何度も実施されて、パケットは送信先の端末まで送信される。この通過していくパケットの道筋を「経路」という。この様子を図 7-2 に示す。

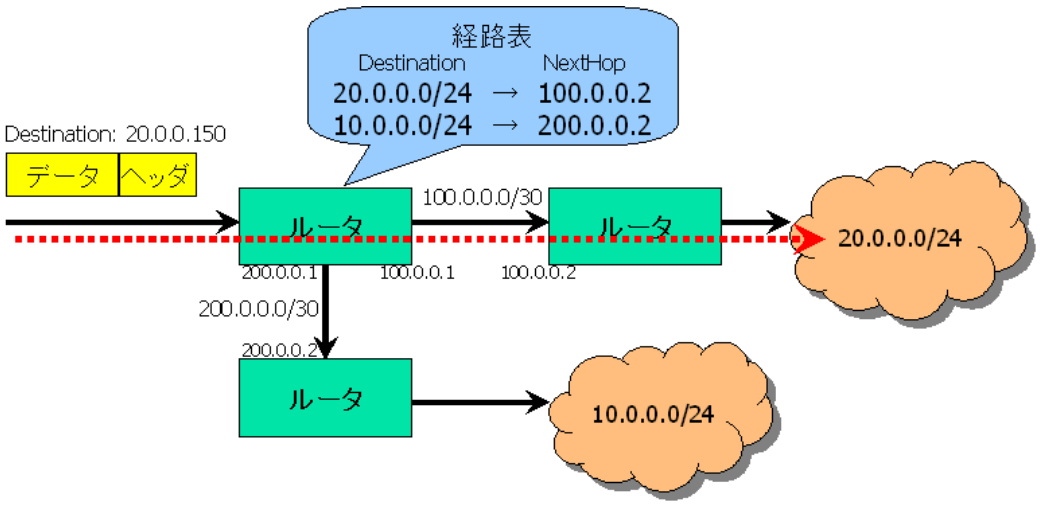


図 7-2 経路表によるパケット転送先の選択

以下に、経路制御の概念についてまとめる。

- 経路
  - パケットが送信元から送信先に向けて転送されてゆく「道筋（パス）」である。

- 経路表
  - 受信したパケットの送信先(Destination)アドレスに向けて、そのパケットを転送する次のルータのアドレスを示したデータベースである。

基本的な経路制御では、この経路表は人の手によって作成され、個々のルータに設定されることによって動作する。しかし、人の手による作業では、管理対象のネットワークが大規模になるにつれ、これらの設定が煩雑になるほか、ネットワークの機器や回線に問題が発生した場合に、人の手を介して対策を講じる必要があり、安定したネットワーク運用に問題が出ることが考えられる。このため、これらの経路表の作成を自動化し、ネットワーク障害が発生した場合にでも、図 7-3 に示すように、障害のポイントを自動的に迂回するような機能を持たせる必要がある。これらの機能を実現する手法として「動的経路制御プロトコル」を利用

する。一方、前者の人の手によって経路表を設定することを、固定的にルータに経路表を設定することから「静的経路制御」と呼んでいる。

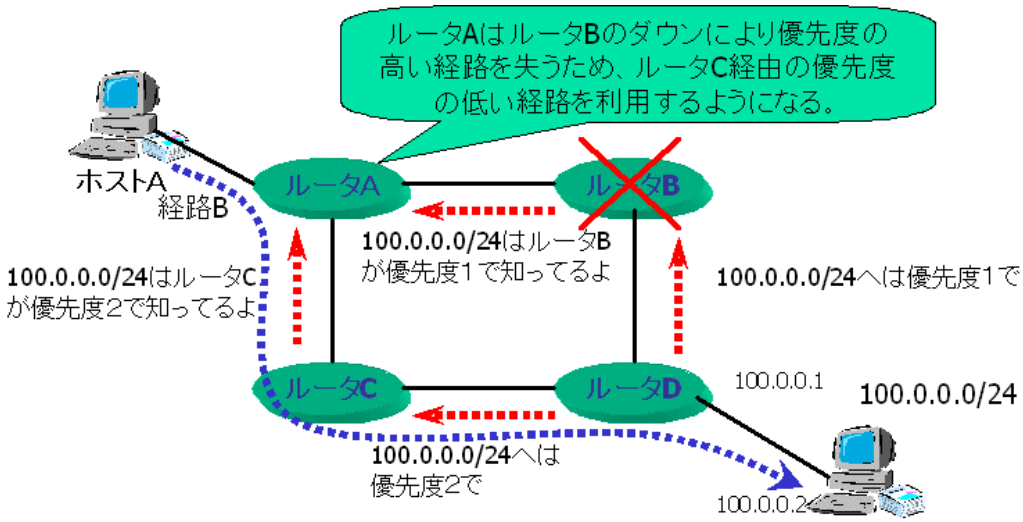


図 7-3　ルータダウンによる迂回の例

経路制御の一般的な特徴は、送信先のネットワークがどこに（どの回線）につながっているかという情報によって送信先を判断している。これは、動的経路制御においても基本的に同じである。具体的には、動的経路制御では、自分が接続しているネットワークの情報を他のルータに伝えるという手段によって、ネットワークの接続情報を構成していく。つまり、あるルータが接続しているネットワークの情報を他のルータに通知し、それをさらにその先のルータに伝搬することで実現するのである。これをデータの流れに置き換えると、データパケットが送信先に向かって送信されていくのに対し、ネットワークの情報である経路情報は、概念的に送信先から送信元に向かって情報が流れていくということになる。

動的経路制御を行う上で、インターネットを安定的に保つというさらに重要な機能を得ることができる。先ほども述べたように、ネットワーク障害の発生している箇所を迂回するという機能である。たとえば、送信元の端末から送信先の端末の間に、複数の経路が存在するような場合に動的経路制御を用いたとする。このような場合、動的経路制御では、これら複数の経路のうち、送信元と送信先の間を結ぶ最良の経路を選択し、その経路を使ってデータを送受信する。そして、万が一最良の経路に問題が発生した場合でも、もう一つある迂回の経路を用いてパケットを送受信するように経路を自動的に選択することが可能になる。

7.1.2. 経路制御ドメイン

経路制御の基本的な考え方は、「7.1.1 経路制御基礎」で解説した。しかし、ここで解説した動的経路制御を大規模なネットワークで実施するには、経路情報が大きくなり過ぎてルータがかえって不安定に可能性があるほか、管理するネットワークが複数の部署や会社にまたがることで、単一の方針によって管理できないなどいくつかの問題が生じてくる。

そこで、動的経路制御では、経路制御できる範囲を分割して管理することで、このような問題を避けるようにすることができる。

この経路制御できる範囲を「経路制御ドメイン(Routing Domain)」と呼ぶ。

また、この経路制御ドメインの範囲内のことを、「イントラ・ドメイン」と呼び、この経路制御ドメイン間を接続することを「インター・ドメイン」と呼ぶ。この様子を図 7-4 に示す。

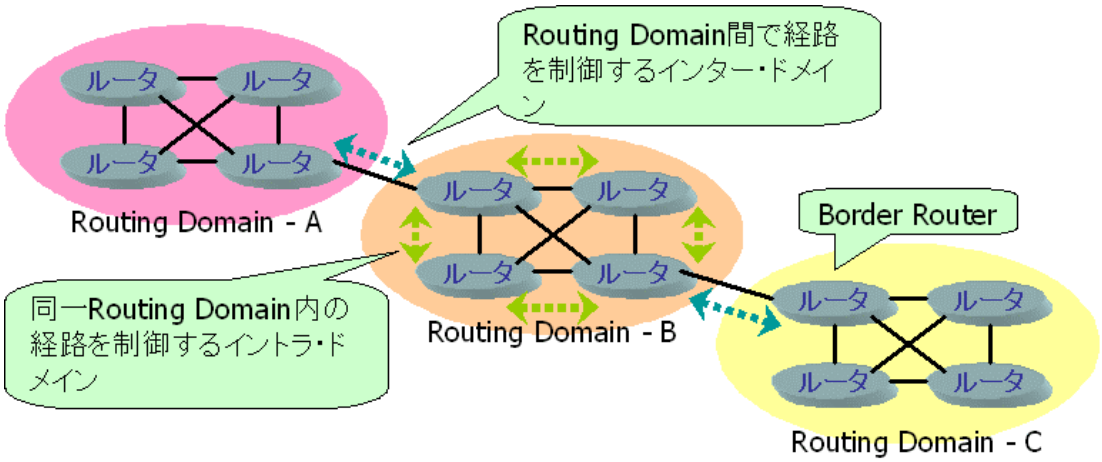


図 7-4　イントラ・ドメインとインター・ドメイン

動的経路制御では、イントラ・ドメインでは一般的に1つの経路制御プロトコルによって動的経路制御を実施する。一方、インター・ドメインでは、同一、または異なる経路制御プロトコル間で互いの経路制御ドメインの経路情報を交換しながら、イントラ・ドメインとインター・ドメインの間の連携を保つ。つまり、インター・ドメインの経路制御で交換される経路情報は、他の経路制御ドメインから受け取った経路情報も伝搬することもでき、これによって、複数の経路制御ドメインを跨いで経路情報を伝搬させることが可能になる。経路情報が複数の経路制御ドメインに行き渡ることができれば、パケットは、複数の経路制御ドメインを跨いで送信可能になると言える。

7.1.3. IGP と EGP

経路制御ドメインは、最小の経路制御の単位となるが、インターネットという膨大なネットワークで経路制御を行うことを考えると、さらに大きな単位での経路制御を行う必要がある。その経路制御単位が「自律システム（Autonomous System）」であり、一般的に「AS」と略して呼ばれものとなる。

ASは、会社やISPなど比較的大規模な単位で構成されるが、その単位毎に一定の経路制御のポリシ（方針）を持って運用されることが前提とされている。一方、AS の内部では、前節で述べたとおり、1つ以上の制御ドメインによって管理されることになる。

AS というような大きな単位間での経路制御プロトコルを「EGP(Exterior Gateway Protocol)」といい、AS内部で利用される経路制御プロトコルを「IGP(Interior Gateway Protocol)」という。この様子を図 7-5 に示す。

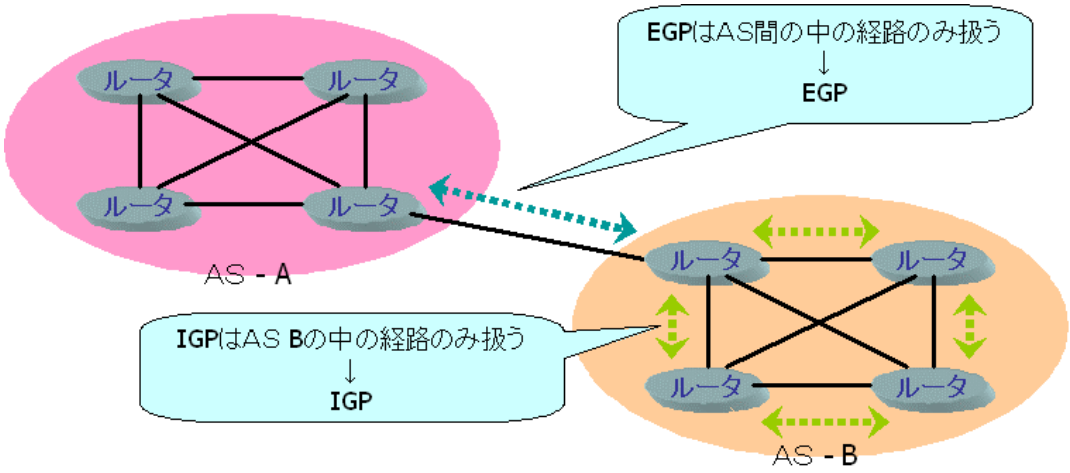


図 7-5 IGP と EGP の違い

IGP と EGP では、経路制御プロトコルの概念が大きく異なる。IGP では、会社や ISP などの組織内部で機敏な制御が可能となるようにサブネット単位での経路制御を基本としている。一方、EGP では、組織単位での経路制御を念頭に置くため、AS という単位で経路制御が行われる。つまり、EGP では、概念的に1つの AS が 1 つのルータのような扱いで経路制御が実施されている。以降の節において、IGP と EGP についてそれぞれ簡単に解説する。

7.1.4. IGP

IGP は、主に AS 内部で経路制御を行うための経路制御プロトコルである。IGP には、主に、RIP、OSPF、IS-IS などのプロトコルが多く使われている。小規模のネットワークでは、RIP が多く使われ、比較的大きなネットワークでは、OSPF がよく使われている。IS-IS は、OSPF と同様の考え方で作られているプロトコルだが、OSPF に比べ早期に作られ、米国などでは未だに多く利用されている。

本節では、以下に RIP と OSPF についてその特徴を整理する。

- RIP (Routing Information Protocol)

  経路制御プロトコルの中で最も単純なプロトコルであり、「距離ベクトル(Distance Vector)型」プロトコルとして、RFC1058[1]にまとめられている。

  初期バージョンの RIP では、アドレスクラス毎の経路制御しかできなかったため、RIP Version2.0 で VLSM(Variable Length Subnet Mask)に対応した。

  RIP での経路制御は、「メトリック（Metric）」といわれる送信先ネットワークまでのルータの数、つまり、送信先ネットワークまでの距離によって制御する。しかし、この Metric の値は最大で 16 までしか対応しておらず、16 台以上のルータを越えた経路制御はできない仕組みになっており、それ以上の大規模なネットワークでは利用できない。

- OSPF (Open Shortest Path First)

  現在 IGP として最もポピュラーなプロトコルであり、「リンク状態(Link State)型」プロトコルとして、RFC2328[2]にまとめられている。現在のバージョンは 2。

  経路制御は、コスト(Cost)という単位を用いて行い、回線毎にコストが設定され、通信端点間の総コストによって最良の経路を選択して、パケットの送信を行っている。

  OSPF では、大規模なネットワークを効率良く管理できるようにするために、複数の

---

[1] Routing Information Protocol(RFC1058)
http://www.ietf.org/rfc/rfc1058.txt

[2] OSPF Version 2(RFC2328)
http://www.ietf.org/rfc/rfc2328.txt

「エリア」という経路制御ドメインを設定できることが一つの特徴である。

7.1.5. EGP

EGP は先にも解説したように、主に AS 間での経路制御プロトコルとして用いられている。EGP は、現在のところ RFC4271[3]にまとめられている BGP-4(Border Gateway Protocol Version.4)が実質的に唯一の物として利用されている。

BGP-4 は、「パスベクタ(Path Vector)型」のプロトコルで、AS までのパス（経路）とその距離によって最良の経路を判断している。

BGP-4 は、インターネット全体という大規模なネットワークを効率よく経路制御できるように設計されている。このため、経路制御用のパラメータとして、内部で利用できるものや、AS 間で情報を伝搬して利用できる物など複数の経路判断基準を持ち合わせている。このようなことから、BGP-4 は高い拡張性と高いスケーラビリティを持ち合わせたプロトコルといえる。

一方で、柔軟過ぎる構造のためこのプロトコルの利用者とも言えるネットワーク運用者にゆだねられる部分も多く、現在の様な高度なセキュリティが要求されるネットワークにおいていくつかの問題点が出はじめている。

次節において、これらの問題点を正しく把握するために、BGP-4 の動作について解説する。

[3] A Border Gateway Protocol 4 (BGP-4) (RFC4271)
http://www.ietf.org/rfc/rfc4271.txt

7.2. BGP-4

本章では、本調査の主たる目的であるインターネット全体の経路情報に直接関係のある、インターネット全体の経路制御を実質的に行っている、BGP-4(Border Gateway Protocol Version 4)について解説する。

最初に BGP-4 の概要について解説し、BGP-4 というプロトコルがどのように作られているかについて解説する。

次に、このプロトコルを用いて、インターネットでは、どのような経路制御が行われているかについて解説する。

最後に、BGP-4 に関して IETF で行われている最新の議論、および BGP-4 に関連して採択されているプロトコルのうち現在のインターネットに大きな影響を持つものについてリストし、解説する。

7.2.1. BGP-4 とは

BGP-4 は、「7.1.5 EGP」の節でも述べたように、パスベクタ(Path Vector)型の経路制御プロトコルであり、主に AS 間での経路制御を行うためのプロトコルとして開発され、現在は RFC4271 としてまとめられている。

BGP-4 の最大の特徴は、経路を制御する単位を概念的に AS という単位で管理することにある。つまり、AS というネットワークの管理主体の方針に従って AS という単位で経路制御を行うことで、その AS の内部で行われている細かい経路制御などに一切関知することなくインターネット全体として経路制御を可能にしている。

このような BGP-4 には、いくつかのバージョンが存在し、現在は Version4 となっている。このプロトコルの前には Version3 が利用されていた。Version3 と Version4 の違いは、CIDR(Classless Inter-Domain Routing)をサポートしていることであり、それ以外の違いはほとんどない。

また、BGP-4 はインターネット全体という大きなネットワークを管理可能にするために、IGP の様な範囲の限られたネットワークを管理するプロトコルとは、以下の様な点が異なる。

- TCP を用いたプロトコル

　通常、経路制御プロトコルは、IP レイヤ、つまり IP アドレスを用いて通信を行うための情報を交換するプロトコルであるため、経路制御プロトコル自身は IP アドレスを用いた通信ではなく、その下位層の通信手段を用いて隣接するルータと直接情報を交換して経路情報のやりとりを行うことが多い。

　しかし、BGP では、隣接する AS の情報を送受信し、AS 間でどのような宛先のパケットを交換すべきかという情報を交換し、それに従ってパケットを転送することを目的としている。このため、IP レイヤとしては、隣接する AS と接続性を持っていることが前提であり、その上で互いの AS の経路情報を交換するため、IGP の様に下位層を用いた通信を行う必要がない。むしろ、BGP にとっては、隣接の AS の BGP ルータと信頼性のある通信セッションを持ち、常に互いの情報を交換できる環境を維持する必要がある。

　BGP では、これらの要件を満たすために IP レイヤ上の TCP によるコネクションによって通信を行うように設計されている。

- 差分のみの広告

　RIP などのように比較的小規模な経路制御プロトコルは、扱うネットワークの情報も少ないため一定時間毎に各ルータが持つ経路情報をすべて交換することで各ルータが持つ経路情報を最新の状態に維持している。

　しかし、経路制御プロトコルが大規模なネットワークに対応した場合、その経路制御プロトコルが取り扱う情報が膨大となり、すべての経路情報をいちいち交換していたのでは効率が良くない。そこで、BGP のような大規模なネットワークを扱うことのできる経路制御プロトコルは、通信相手のルータに対して更新された差分の経路情報を送付することで、全体の経路情報を最新の状態に維持するように設計されている。

- パスベクタ型の採用

　BGP は、パスベクタ型プロトコルであることは何度も述べてきた。パスベクタ型プロトコルの特徴は、どのような AS を経由して経路情報が到達したかを示す経路（パス）に主眼をおいて経路制御されるところにある。

パスに主眼を置くことで、経路情報が AS を経由してゆく段階で、同一の AS を何度も通過するような「ループ」という状態を検出することができ、このような不要な経路を排除することができる。さらに、経路情報の受信場所から、その経路情報の発信元までいくつの AS を経由してきたかという、概念的な距離も知ることができる。BGP-4 では、ここであげた情報も利用して最良の経路の選択をするすることが可能である。

このような、AS の連続によって表された経路を「AS パス」とよび、この AS パスに関連づけられた情報を「パス属性」とよび、BGP の経路制御に用いられる。

7.2.2. BGP-4 のプロトコル概要

BGP-4 は、TCP/IP 上で動作する経路制御プロトコルである。この経路制御プロトコルでは、前節でも述べたように BGP のコネクション(BGP セッション)を通じてお互いが持つ経路情報を、それぞれが望む属性(Attribute)を付与して相手側に広告(Announce)することによって経路情報を交換する。この様子を図 7-6 に示す。

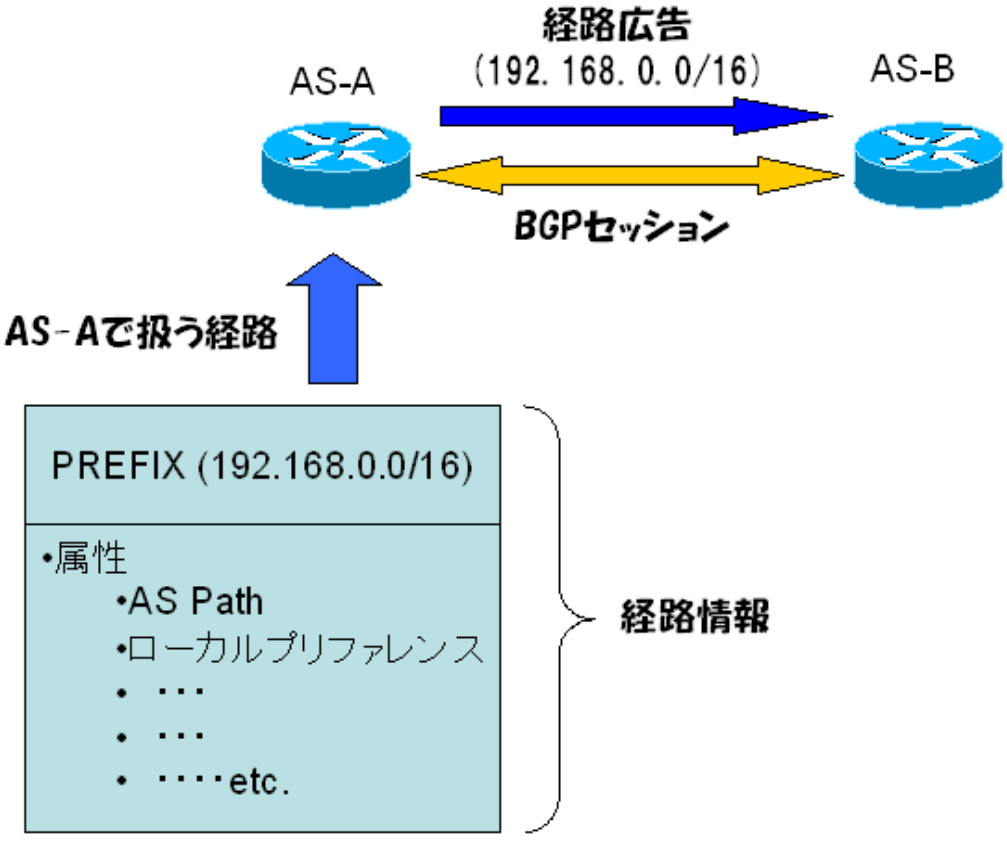


図 7-6 経路の構成とその広告

このため、BGP-4 ではお互いの状態を確実に把握するため、以下の 4 種類のメッセージを用いて互いの状態は管理している。

- OPEN

  待ち受け状態の BGP ルータに対してこのメッセージを送信することで、BGP の接続を開始する通知を行うと共に、互いの BGP ルータのサポートしている機能に関する情報を交換し、接続条件を互いのルータ間で交渉するために用いられるメッセージである。

- UPDATE

  経路情報の追加、削除、更新などの経路の更新情報を相手側に伝えるために用いるメッセージである。

- KEEPALIVE

  BGP-4 では、更新情報を接続相手に伝えるように設計されているため、経路の更新がなければ、メッセージの送信自体が発生しないことになる。このため、BGP セッションが確実に動作していることを確認するために、このメッセージを一定間隔毎に送付し、互いの BGP ルータが動作中であることを確認する設計となっている。

  この KEEPALIVE メッセージは、BGP ルータで持つ HOLD TIMER というタイマの三分の一の時間間隔で送信し、最短で 1 秒の間隔に設定可能である。HOLD TIMER は、KEEPALIVE メッセージの到着を最長で何秒間待つかを設定するタイマで、最低で 3 秒の値をとる。RFC4271 に推奨されているデフォルトの値は 180 秒であることから、KEEPALIVE メッセージの推奨送信間隔は、その三分の一の 60 秒となる。

  HOLD TIMER は、UPDATE メッセージ、または KEEPALIVE メッセージのいずれかを受信することで初期化されるが、HOLD TIMER が超過した場合は、その BGP セッションが正しく動作していない、または、どちらかの BGP ルータが正しく動作していないと判断し、その BGP セッションを破棄し、BGP 接続を切断する仕組みになっている。

- NOTIFICATION

　BGP-4 では、OPEN メッセージで正しくお互いの接続条件が満たされない場合や HOLD TIMER が超過した場合など、いくつかの状況において BGP セッションを維持できない、または確立できない場合がある。

　このような場合、BGP では、NOTIFICATION メッセージを相手側に送信し、エラー内容を通知すると共に、BGP セッションを切断するように設計されている。

BGP-4 では、これら4つのメッセージを使って BGP セッションを以下の 6 つの状態のいずれかにあることを管理している。

- IDLE

　BGP が動作していない状態を示している。

- ACTIVE

　BGP が動作している状態を示すが、BGP セッションの確立動作をしているわけではなく、BGP セッション接続先ルータとの TCP 接続受信待ち状態であることを示している。

- CONNECT

　OPEN メッセージの送受信前の TCP セッションの接続動作を実施中である状態を示している。

- OPEN SENT

　TCP セッションが確立し、OPEN メッセージを接続相手に送信した状態を示している。

- OPEN CONFIRM

　BGP の接続相手から OPEN メッセージを受信し、その動作条件を確認し、送信元の BGP ルータに対して確認の OPEN メッセージを送信し、その確認中の状態を示し

ている。

- ESTABLISHED

お互いの BGP の接続状態が確認でき、BGP セッションが正しく確立されている状態を示している。

BGP による経路情報の送受信は、この ESTABLISHED の状態に遷移した後に行われる。

BGP-4では、これらのメッセージを利用し状態の管理を行い、BGPセッションを正しく起動する。起動された BGP セッションは、ESTABLISHED 状態となり、経路情報の交換がUPDATE メッセージを用いて行われる。

これらの状態の遷移図を図 7-7 に示す。

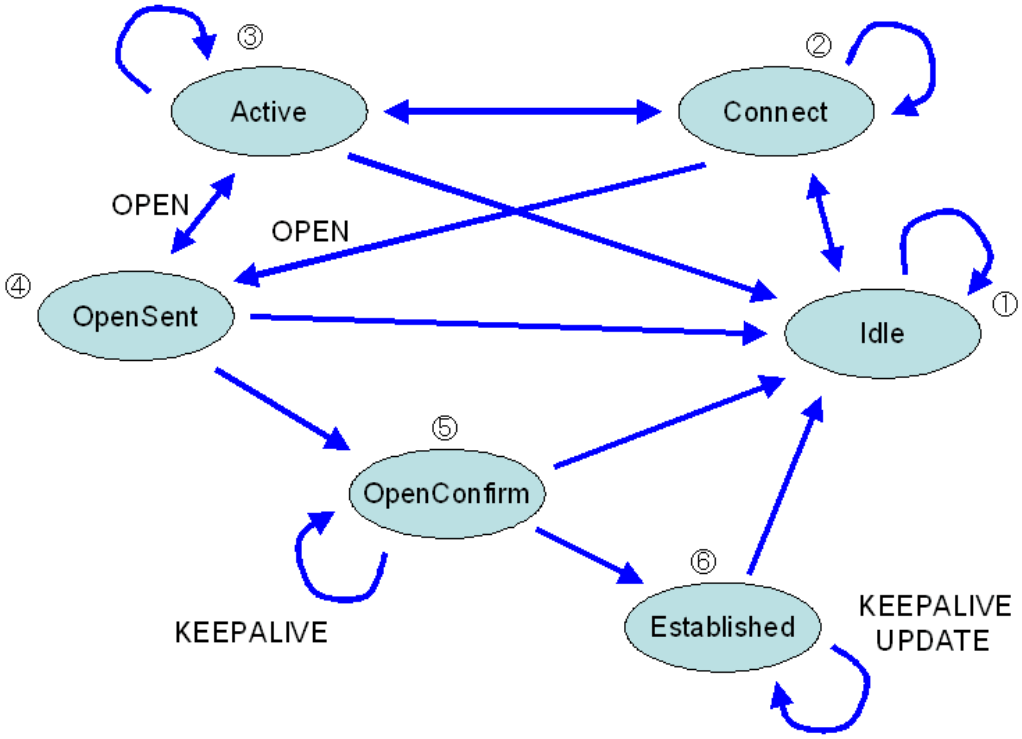


図 7-7 BGP の状態遷移図

UPDATE メッセージには、図 7-8 に示すように、パス属性、NLRI(Network Layer

Reachability Information)、および Withdrawn Routes の3種類の情報によって構成されている。

NLRI は、更新する経路の情報であり、送信元 BGP ルータが該当するネットワークに対する到達性情報を知っているということを受信側 BGP ルータに伝えてきている。この NLRI に対するパスの情報はパス属性によって与えられる。一方、Withdrawn Routes は、送信元 BGP ルータで該当する経路を失ったことを受信側 BGP ルータに伝える役割がある。

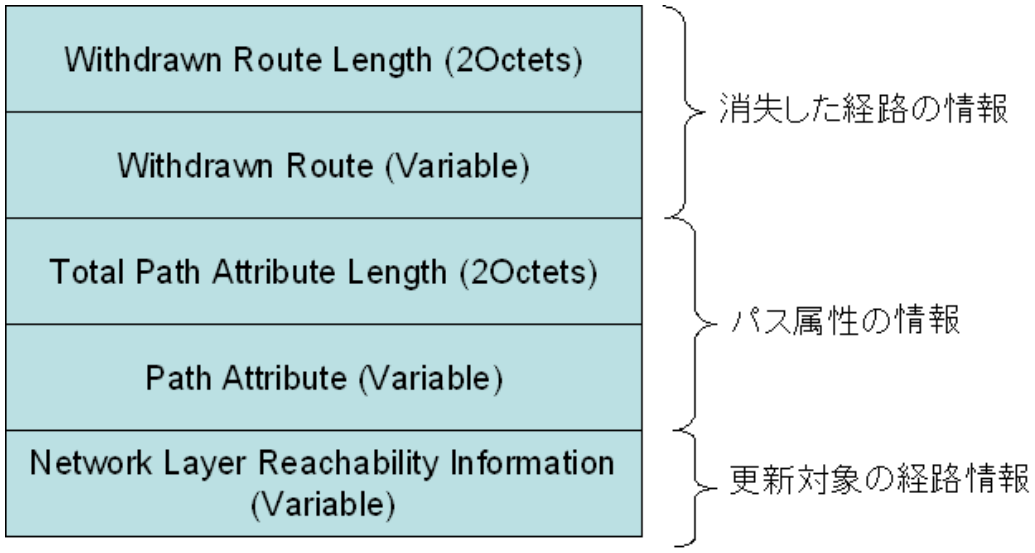


図 7-8 BGP Update メッセージのフォーマット

BGP-4 では、この UPDATE メッセージを用いて経路情報を交換し、最良の経路（パス）を選択している。次節に、BGP-4 による経路制御について解説する。

7.2.3. BGP-4 の経路制御方式

BGP は、AS 間で経路情報を交換するためのプロトコルであるが、AS の内部では他の AS から受け取った経路を自分の AS の中の他の BGP ルータに伝える必要もあるため、AS の内部でも BGP の接続を行うのが通例である。

このような AS 内部の BGP 接続を「IBGP（Internal BGP）接続」といい、通常の AS 間の接続を「EBGP(External BGP)接続」という。この様子を図 7-9 に示す。

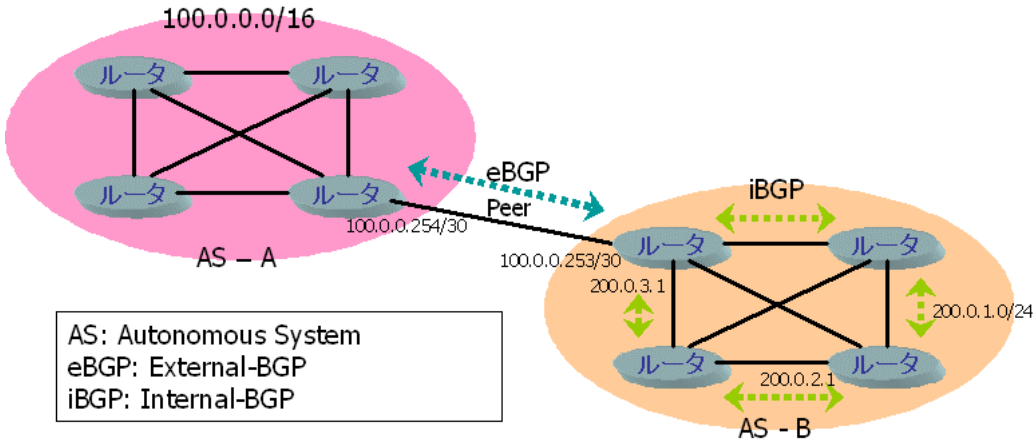


図 7-9 iBGP と eBGP

IBGP と EBGP では、プロトコル的な動作はほとんど変わらないが、唯一違う動作として、IBGP 接続から受信した経路情報を他の IBGP 接続へ転送しないという動作をする。このため、EBGP 接続から受信した経路は直接 BGP で受信した経路を必要とするルータに IBGP 接続で伝える必要がある。このため、通常では AS の内部にあるすべての BGP ルータ同士フルメッシュで IBGP 接続を設定しなくてはならない。

IBGP のフルメッシュによる接続によって AS 内部のすべての BGP ルータで、その AS から他の AS に向かう経路情報を共有することができる。しかし、複数の外部接続を持つ AS の場合、ある AS に向かう経路情報を複数の EBGP 接続から受信することがよくあるため、AS の内部には2つ以上の同一の AS に対する経路が存在することになる。このような場合、AS の内部でどちらの経路を選択するかを決める必要がある。

「7.1.5 EGP」でも述べたとおり、BGP では大規模なネットワークを柔軟に制御できるパス属性といわれる選択パラメータを持っている。以下に、BGP で標準的に利用できる代表的なパス属性について列挙する。

- Origin

  Origin は、何によってその経路が作られたかを示している。この属性の内容は基本的に変更せずに他のルータに伝搬する必要がある。

  Origin の値は、IGP、EGP、Incomplete のいずれかを採る。

- AS Path

  AS Path は、該当する経路情報が通過した AS の番号を列挙している。つまり、一番先頭に経路情報の発信元 AS が記録され、経路情報を伝搬した AS 番号が付加され、最後に自分の AS 番号が記録される形となる。

- Next Hop

  Next Hop は、該当する経路情報へ向かうパケットは、どの IP アドレスへ転送すればよいかを示している。

  EBGP から受信したばかりの経路情報の場合は、隣接するルータの IP アドレスが付与されるのが一般的である。しかし、IBGP 経由で受信した経路情報の場合、自分のルータに直接関係の無い、他のルータの IP アドレスである場合が多く、BGP ルータは、該当する経路情報でパケットの転送をする場合には、BGP の Next Hop アドレスの情報に従って IGP の経路情報を用いてパケットを転送しなくてはならない。

  このため、多くの場合、IBGP に経路情報を転送する場合はこの Next Hop 属性の値は変更されずに転送される。

- MED(Multi-Exit Discriminator)

  MED は、ある AS に対して直接的な接続を 2 つ以上持っている場合に、どの EBGP 接続を優先させるかという意志を接続先 AS に伝える為の属性として用いられる。

  最近の運用では、トラフィックの制御を緻密におこなうため MED などの外的要因によってトラフィックが変動しないように MED 属性を無視するような設定をしている場合が多いため、MED によってトラフィックの制御を行うような場合は、接続先の AS とあらかじめ調整が必要になるケースが多い。

- Local Preference

　Local Preference もパケットの転送先の優先順位を決めるためのパラメータである。主に AS 内部で利用するためのパラメータとして実装されている。

　たとえば、ある EBGP 接続で受信した経路情報を他の EBGP 接続でも受信していたとする。このとき、意図的にどちらかの接続の情報を優先させたい場合、つまり優先させた経路情報を受信している BGP ルータを利用してパケットを他の AS に転送したい場合に、このパラメータを利用して優先的に転送するように設定することができるのである。

　このパラメータの値は、経路情報とともに他のルータに転送してはいけない属性となっているため、基本的に他の AS には転送されない。

さて、BGP ではこれらのパス属性などを使って 1 つの転送先に向かう複数の経路情報からより良い経路情報を選択しなくてはならない。この選択基準は、ルータメーカによって若干の違いがあるが、多くの場合は同じ基準が採用されている。以下に、シスコシステムズ社製のルータで採用されている、経路選択基準を示す。なお、原文のまま引用しており、必要と思われる場合には、括弧内に解説を加えてある。（参照：インターネット・ルーティング・アーキテクチャー）

1) ネクストホップが到達できないときには、その経路は無視される。

2) 大きなウエイト(Weight)を持つパスを優先する。（Weight は、Cisco 独自のパラメータである。）

3) ウエイトが同じときには、大きなローカル優先度(Local Preference)を持つ経路を優先する。

4) 同じローカル優先度を持つときには、ローカルに生成された経路を優先する。

5) ローカル優先度が同じときには、短い AS パスを持つ経路を優先する。

6) AS パスの長さが同じときには、低い起源タイプ(Origin)を持つ経路を優先する。
(IGP < EGP < Incomplete の関係にある。つまり IGP が一番優先される。)

7) 起源タイプが同じときには、低い MED を持つ経路を優先する。
同じ MED を持つときには、次の方法で経路を優先する。外部（EBGP）は、コンフェデレーション外部経路よりも優先され、コンフェデレーション外部経路は内部(IBGP)よりも優先される。（コンフェデレーションは、BGP の拡張によって AS の内部をさらに複数の AS に分割することで、より大きなネットワークを管理しやすくするものである。次節にて解説する。）

8) 前記の条件がすべて同じときには、最も近い IGP 近隣ルータ経由で到達可能な経路を優先する。すなわち、その宛先に到達するのに AS 内で最短の内部パスを採る経路を優先する。

9) 内部パスが同じときには、BGP のルータ ID が決め手となる。最も低いルータ ID を持つ BGP ルータから来る経路が優先される。ルータ ID は通常、そのルータにおける最も大きな IP アドレスか、ループバック（仮想）アドレスである。ルータ ID は、実装に固有なものである。

このような優先される経路の選択は、AS の内外を問わずすべての BGP ルータで行われている。

一般的には、AS 内部では、先ほども説明したようにフルメッシュで IBGP 接続を行うため、優先される経路の情報は共有されており、AS の内部で優先される経路情報は複数あるうちの1つに選択されるが、IBGP 接続の接続先を操作したり、Local Preference の値を適宜操作したりするなどして、AS 内部の所々で選択する経路を変えることもでる。

BGP の経路制御に関する設定や運用が本報告書の主眼ではないため、詳しくは触れないが、BGP はここまでに説明したパス属性や BGP 接続を恣意的に操作するなどして非常にスケーラブルかつ柔軟な運用ができるよう設計されている。

7.2.4. 関連プロトコル

BGP に関する情報は、先にも述べたとおり RFC4271 にまとめられている。RFC4271 は、2006 年 1 月に発行された最新の RFC であり、その前には、RFC1771[4]として発行されていたものである。

[4] A Border Gateway Protocol 4 (BGP-4) (RFC1771)
http://www.ietf.org/rfc/rfc1771.txt

RFC1771 と RFC4271 の違いは、RFC4271 の Appendix. A にまとめられている。以下にその引用を日本語化したものを記載する。

Appendix. A RFC1771 との比較

RFC1771 との比較において、編集上の変更は数多くある。（ここに記載するには多すぎる。）

以下に技術的変更点を列挙する。

- TCP MD5[RFC2385 [5]]、BGP Route Reflectors[RFC2796 [6]]、BGP Confederations[RFC3065[7]]、BGP Route Refresh[RFC2918[8]]といったような機能の利用方法を反映し変更した。

- AGGREGATOR 属性の BGP Identifier の利用方法を明確にした。

- BGP ルータが Peer から受け付けるプレフィックスの上限を制限する手順

- AS 間のトラフィックエンジニアリングの為に AS_PATH 属性にある自 AS のもう一つの事象を BGP ルータの機能として含むようにした。

- NEXT_HOP タイプを明確にした。

- ATOMIC_AGGREGATE 属性の利用法を明確にした。

- 直接的な Next Hop 間の関係と NEXT_HOP パス属性で定義される Next Hop について記述した。

- Tie-Breaking 手順について明記した。

---

[5] Protection of BGP Session via the TCP MD5 Signature Option (RFC2385)
http://www.ietf.org/rfc/rfc2385.txt

[6] BGP Route Reflection – An Alternative to Full Mesh IBGP (RFC2796)
http://www.ietf.org/rfc/rfc2796.txt

[7] Autonomous System Confederations for BGP (RFC3065)
http://www.ietf.org/rfc/rfc3065.txt

[8] Route Refresh Capability for BGP-4 (RFC2918)
http://www.ietf.org/rfc/rfc2918.txt

- 経路広告の頻度について明記した。

- オプションパラメータ Type1(Authentication Information)については、推奨しないこととした。

- UPDATE メッセージエラーのサブコード 7 番 (AS Routing Loop) は、推奨しないこととした。

- OPEN メッセージエラーのサブコード 5 番 (Authentication Failure) は、推奨しないこととした。

- マーカーフィールドを認証の為に使うことを推奨しないこととした。

- TCP MD5 認証の実装を必須とした。

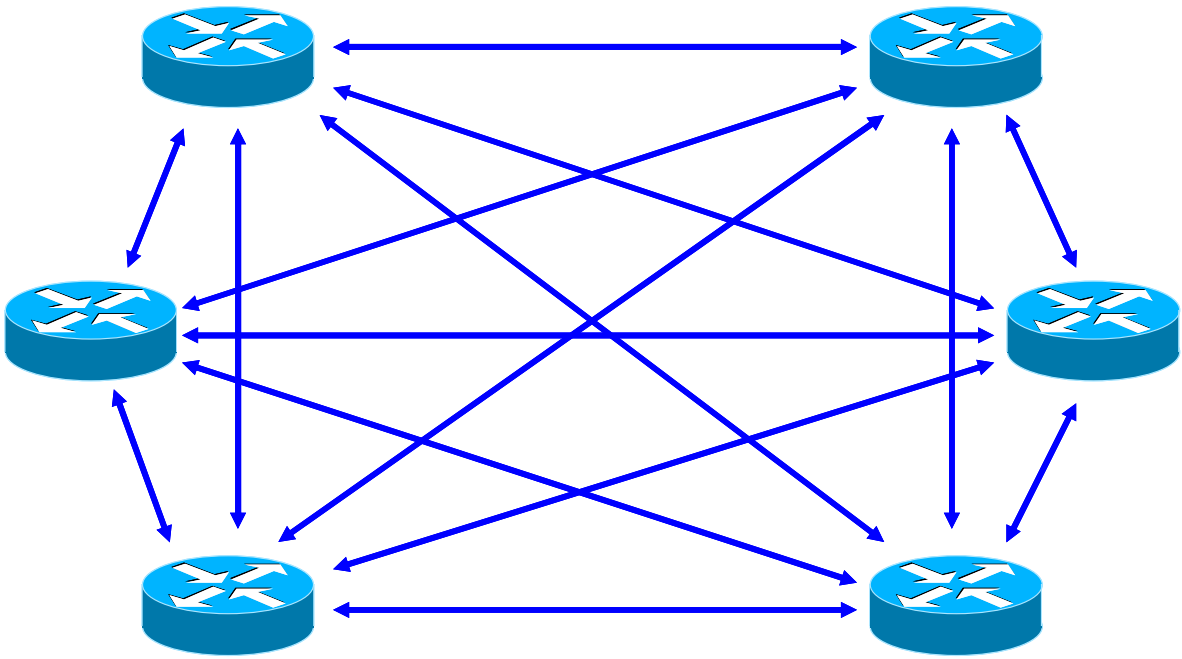
- BGP の FSM を明確にした。

この 2 つの RFC の間に大きな変更はなく、RFC1771 発行以降に発生したいくつかの修正などを反映したような作りになっているだけである。中でも、AS_Path 属性に関する修正が大きな物な修正の一つとしてあげられるが、これもシスコ社製ルータでは古くから実装されているもので、実際の運用現場では既にその機能が標準となっており、今回の RFC はこれに合わせた形になっている。

しかし、BGP はこれら基本的な機能だけでなく、変更点の第一項目で挙げられているような様々な拡張によって、より充実した機能を利用できるようになる。以下に、関連する機能、RFC のなかから重要な物を選択し、簡単な解説を加える。

■ BGP Route Reflection – An Alternative to Full Mesh IBGP (RFC2796)

IBGP 接続は、IBGP 接続経由で受信した経路は他の IBGP 接続に再広告しないため、EBGP で受信した経路は IBGP 接続を用いて AS 内のすべての BGP ルータに直接広告をしなくてはならない。このため、IBGP 接続は、フルメッシュで BGP 接続を行うこととなる。フルメッシュで IBGP 接続を行うと図 7-10 に示すように IBGP 接続が膨大な数となる。このため、大規模なネットワークで BGP を運用する場合には、ルータの性能などが問題となるケースがある。



BGPのフルメッシュのPeer数は、n×(n-1)の式に沿って増加する。
図の場合、n=6なので、6×(6-1)=30のPeer数となる。

図 7-10　フルメッシュによる IBGP 接続数の問題

このような問題を解消するために、ネットワーク内部をいくつかの「クラスタ」と言われるドメインを作り、その中に Route Reflector というそのクラスタのなかで経路制御を中核的に行うルータを設置し、他のルータは Route Reflector に対して IBGP を接続することで、IBGP 接続の数を減らすことが可能となる。

図 7-11 に Route Reflector を使った構成例を示す。

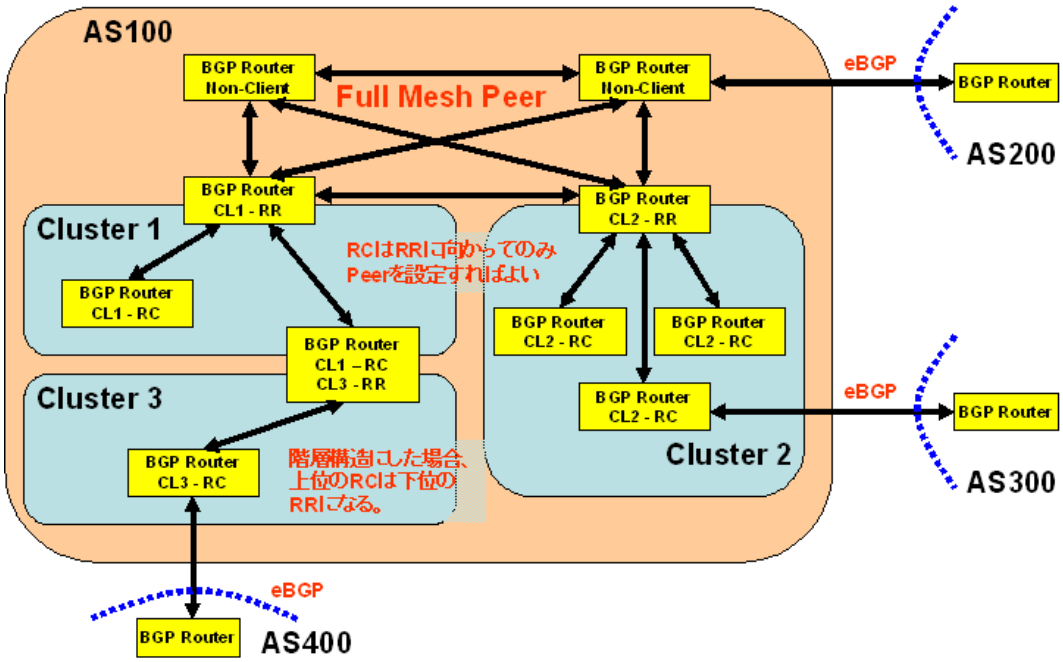


図 7-11　Route Reflector

Route Reflector となったルータは、IBGP で Route Reflector Client から接続されるが、通常の IBGP 接続とは異なり、一部の非転送なパス属性と共に Route Reflector Client から受信した経路を他の Route Reflector Client とクラスタに属さない IBGP ルータに経路情報を転送する処理を行う。

■ Autonomous System Confederations for BGP (RFC3065)

BGP Confederation という手法は、上記の Route Reflector よりも以前に考案された手法で結果的な効果として IBGP の数を減らすという結果を得ることができる。

BGP Confederation を利用した構成例を図 7-12 に示す。

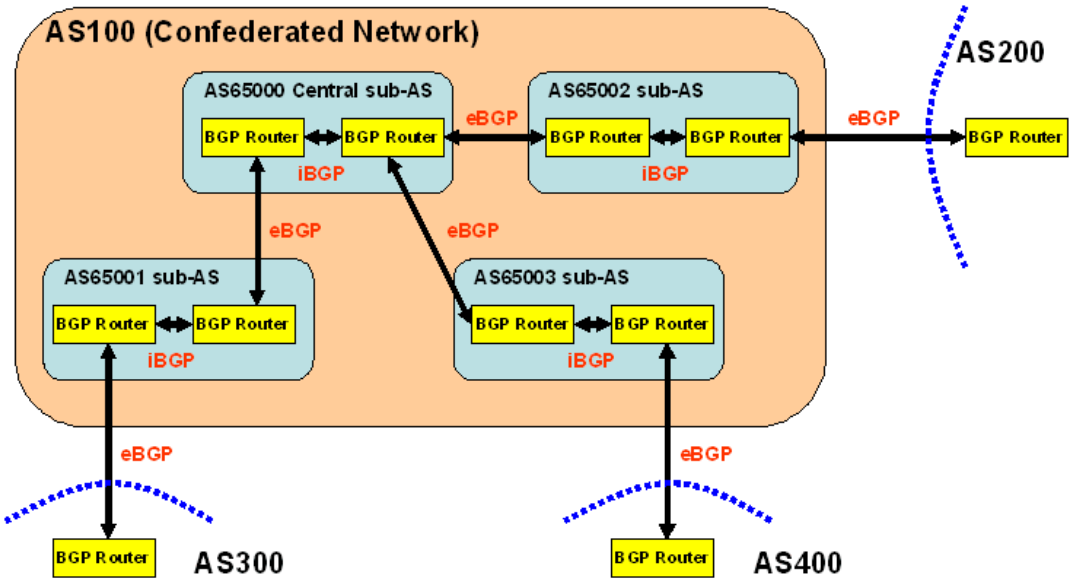


図 7-12　BGP Confederation

しかし、BGP Confederation の考え方は、Route Reflector とは異なり、AS 内部を複数の Sub-AS という AS に分割し、その Sub-AS のなかで独立した経路制御を実施することを推奨している。つまり、外部の AS から見れば通常の AS に見えるが、内部は、複数の AS があたかも小規模のインターネットのように BGP で接続されている状況といえる。これは、IBGP の接続数を減らすという単純な効果だけでなく、非常に大きな AS の場合は、各部署での経路制御ポリシを十分に反映させる方法として採用できるなどの効果もある。

■ Protection of BGP Sessions via the TCP MD5 Signature Option (RFC2385)

この RFC では、TCP の Extension を用いた BGP に対するセキュリティの拡張について記述されている。TCP のオプションでは、TCP セグメントに RFC1321[9]で規定される MD5 Digest を運ぶように設計されている。通常の TCP を利用した BGP セッションでは、何の暗号化もされていないことから、万が一の場合、BGP セッション自体を詐称されるなどの危険性があるが、この TCP MD5 認証方式を使うことで、その危険性から幾分か解消される。

■ Multiprotocol Extensions for BGP-4 (RFC2858[10])

この拡張は、BGP を IPv4 の経路制御だけではなく、ありとあらゆるネットワークアドレスや ID を転送する為の手段として利用できるようにした物である。

たとえば、IPv6 に対応した BGP は「BGP4+」と記されることが良くあるが、実体は、この Multiprotocol Extension を用い、アドレスファミリとして IPv6 を使っているに過ぎない。

このほか、この拡張は MPLS を用いて VPN を構築する際に、ラベルや MAC アドレスを転送するためにも利用することができる。

9 The MD5 Message-Digest Algorithm (RFC1321)
http://www.ietf.org/rfc/rfc1321.txt
10 Multiprotocol Extensions for BGP-4 (RFC2858)
http://www.ietf.org/rfc/rfc2858.txt

7.3. 最新経路情報

本節では、インターネットで実際に流れている経路情報について解説する。

最初に、インターネット全体で取り扱われている経路情報である「フルルート」の経路数の推移を紹介するとともに、フルルートの実態について解説する。以降、この情報を基本として、インターネットルーティングレジストリやインターネットレジストリからの割り振り状況との比較、1つの経路情報が複数の AS から広告される Multi-Origin Prefix の現状について、具体的な数字を挙げながら解説してゆく。

7.3.1. フルルートに関する状況

本節では、インターネットに流れている経路の実態について解説する。

なお、本節で利用する各種情報は特に断りが無い限り、株式会社インテック・ネットコアで取得したフルルートの情報を、経路情報を分析する独自のツールによって解析した 2006 年 1 月 26 日時点の結果を利用して紹介する。

図 7-13 にフルルートの経路数推移を示す。

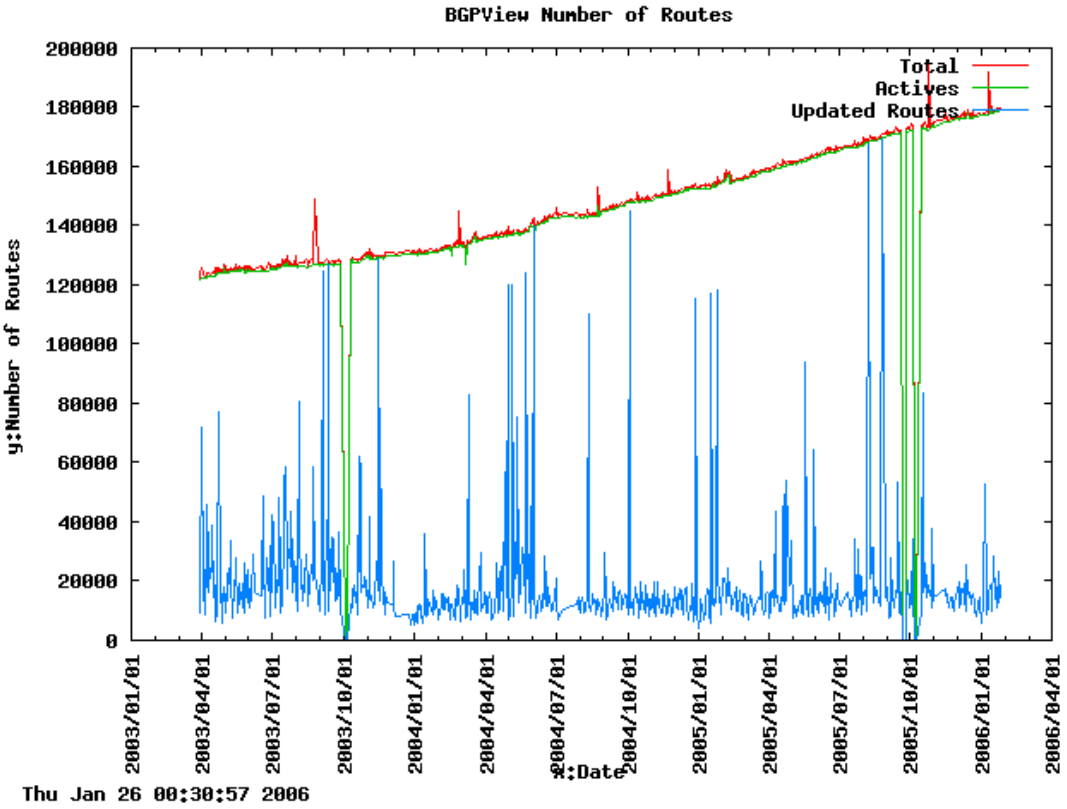


図 7-13 フルルートの経路数推移

グラフは、2003 年 4 月からの推移を示している。グラフでは、上部のなだらかな傾斜の線がフルルートの経路数を、下部の変化が激しい線がフルルートの中で 24 時間以内に更新が行われた経路数を示している。

フルルートの経路数は、グラフ作成当初の 2003 年 4 月において約 12 万経路程度だったが、グラフ作成時点の 2006 年 1 月の段階では約 18 万経路と 3 年間で約 6 万経路も増加していることがわかる。また、グラフの曲線から経路数は単調に増加傾向にあるのではなく、指数関数的に増加していることも推測できる。

一方、24 時間以内の更新される経路数の範囲は、測定開始当初からほとんど変化が無く、約 1 万から 2 万経路程度であることがわかる。

そこで、1日の経路更新数の推移を図 7-14 に示す。

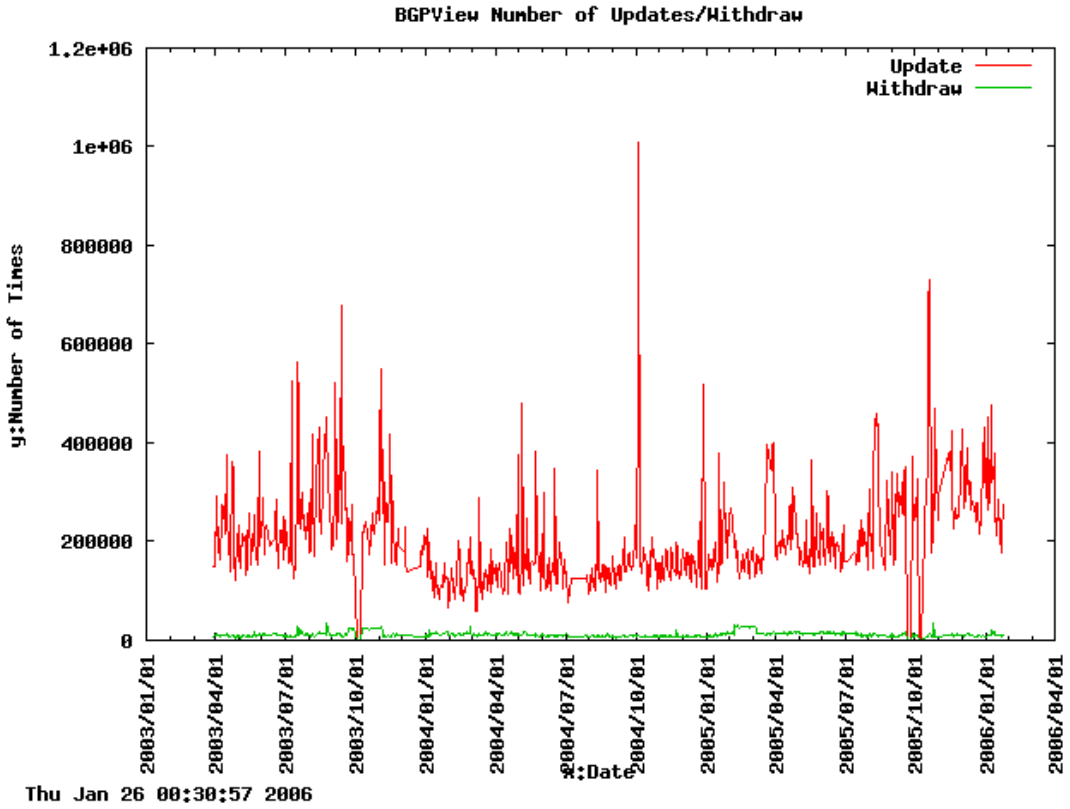


図 7-14 1 日の経路の更新回数の推移

この図では、経路更新数は 2003 年から現在に至るまで多少の変動はあるものの全体で約 2 万回の更新が行われ大きな変化が見られないことがわかる。

そこで、さらにこの更新される経路の内訳をプレフィックス長別に見たものを図 7-15 に示す。

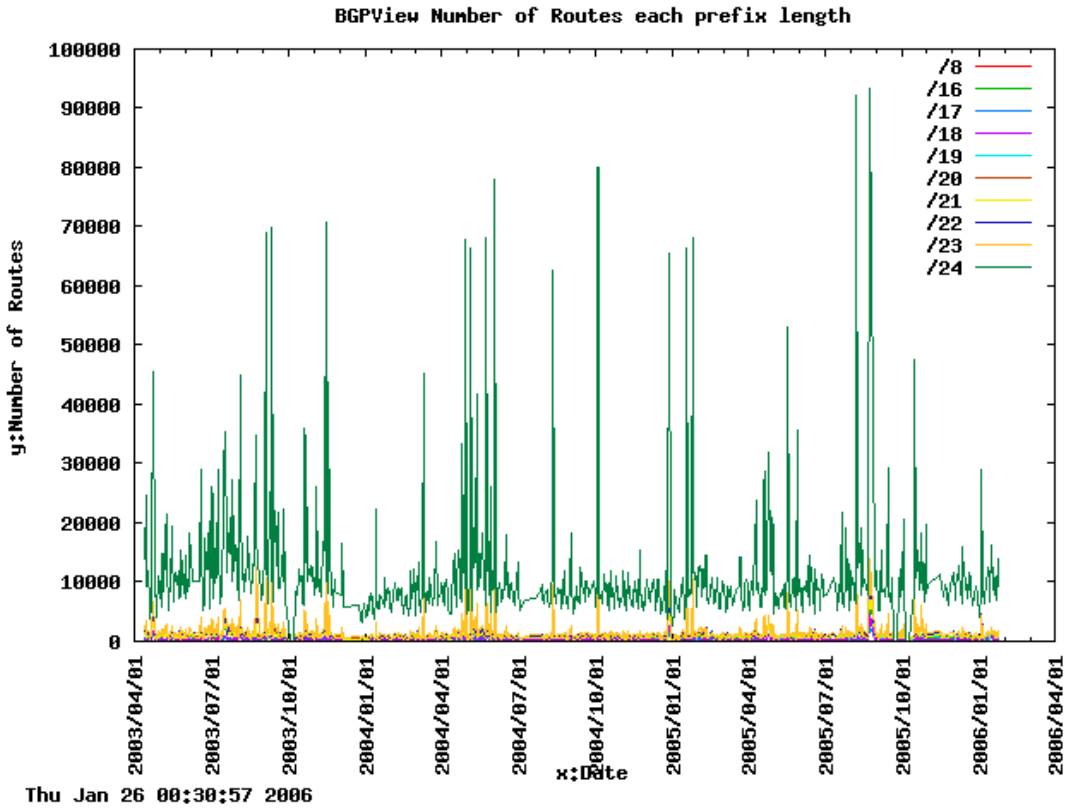


図 7-15 プレフィックス長別の経路更新数の推移

この図では、/24 の更新数を緑色で示しており、その数が他のプレフィックス長の更新回数よりも多く、約 1 万回から 1 万 5 千回で推移しており、その数に長期にわたってあまり変動が無いことがわかる。

つまり、経路数は指数関数的に増加傾向にあるが、そのなかで更新される経路のほぼ大半が/24 の経路であることがわかり、かつ、その更新回数は数年にわたって変化が無いことがわかる。

さて、一方でその経路を流す組織数を AS 単位で見た数、つまり Origin AS の数の推移を図 7-16 に示す。

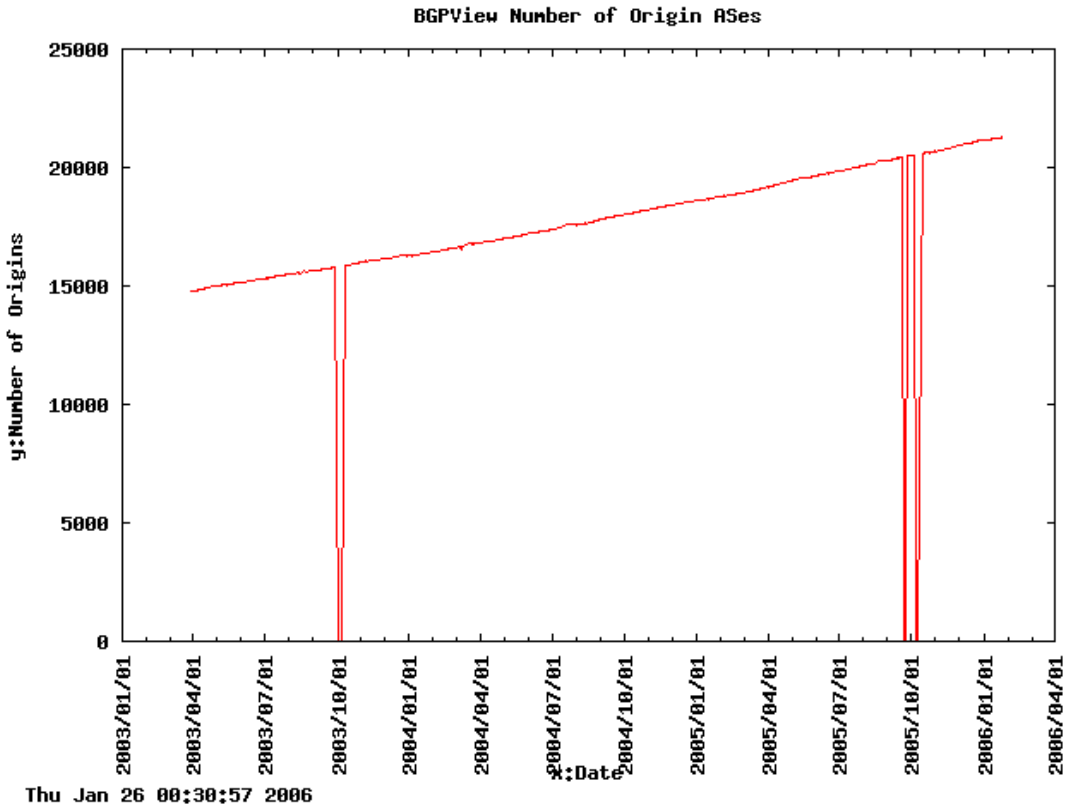


図 7-16 Origin AS 数の推移

図中に2箇所ほど 0 に落ち込んでいるところがあるが、これはデータの取得が失敗している部分であり無視できる。この図によると、2003 年の段階では約 15000 の AS が経路を広告していることがわかり、その数は 2006 年の段階で約 20000 にまで増加している。つまり、3年間で約 5000 の AS が新たにインターネットに接続し、経路の広告を開始したことになる。

この増加傾向と経路数を比較するために、一つの AS が広告する経路数の平均値の推移を図 7-17 に示す。

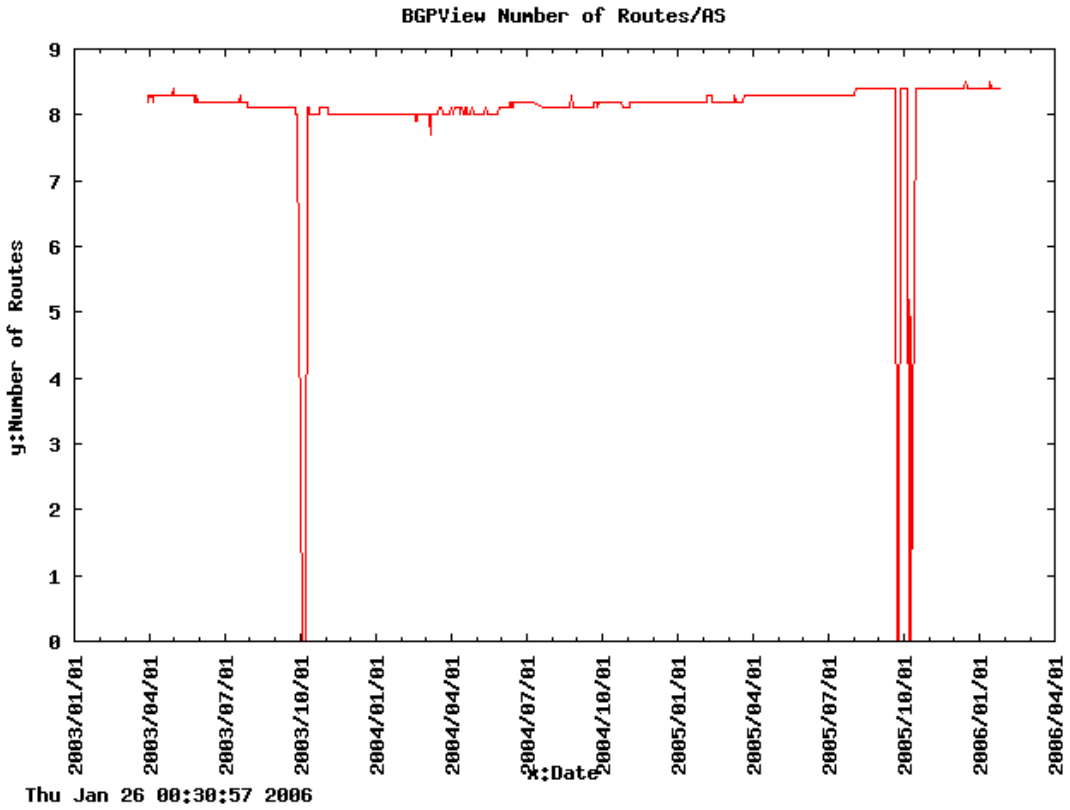


図 7-17 1AS あたりの経路広告数の平均値の推移

図のとおり経路数、Origin AS 数ともに増加傾向にはあるものの、1AS あたりが広告する経路数の推移は約 8 から 9 経路とその数に大きな変化が無いことがわかる。

これらのフルルートとその経路の内容が示す状況として、経路数や AS 数については、増加傾向にあるが、各 AS が取り扱う経路数については、2003 年以降大きな変化は見られないこと。そして、インターネット全体で日々更新される経路数についてもプレフィックス長が/24 のものが多く更新されるものの、更新数は 2003 年以降それほど大きく変化がなく、経路更新という観点から見たインターネットの安定性ついては、大きな変化が見られないことが理解できる。

7.3.2. パンチングホール経路の実態

前節では、流れている経路数そのものに注目してデータを見てきた。本節では、流れている経路の重なり具合を見ることにする。

あるネットワークまでの経路を表す経路情報は一つのネットワークに対して 1 つの経路情報があれば、基本的に問題がない。しかし、インターネットに流れるトラフィックをきめ細かく制御したり、接続事業者の運用上の方針などにより1つのネットワークに向かう経路を分割するなどして複数流したりすることがある。特に、192.168.0.0/16 という経路に 192.168.1.0/24 という重ね合った経路を流した場合に 192.168.1.0/24 は 192.168.0.0/16 のパンチングホールな経路ということができる。

経路制御上、このようなパンチングホール経路は正しく運用されている場合がほとんどだが、経路制御の特徴である Prefix 長[11]が長いほど優先させるという最長一致の原則を逆手にとり、Prefix 長の長い経路を故意にインターネットに流し、経路を横取りするケースもある。

ここでは、悪意のあるなしにかかわらずこの様なパンチングホールの経路がどれくらいあるのかを図 7-18 に示す。

11 本章では他の用語との関係をわかりやすくするため、「プリフィクス」を「Prefix」と表記する。

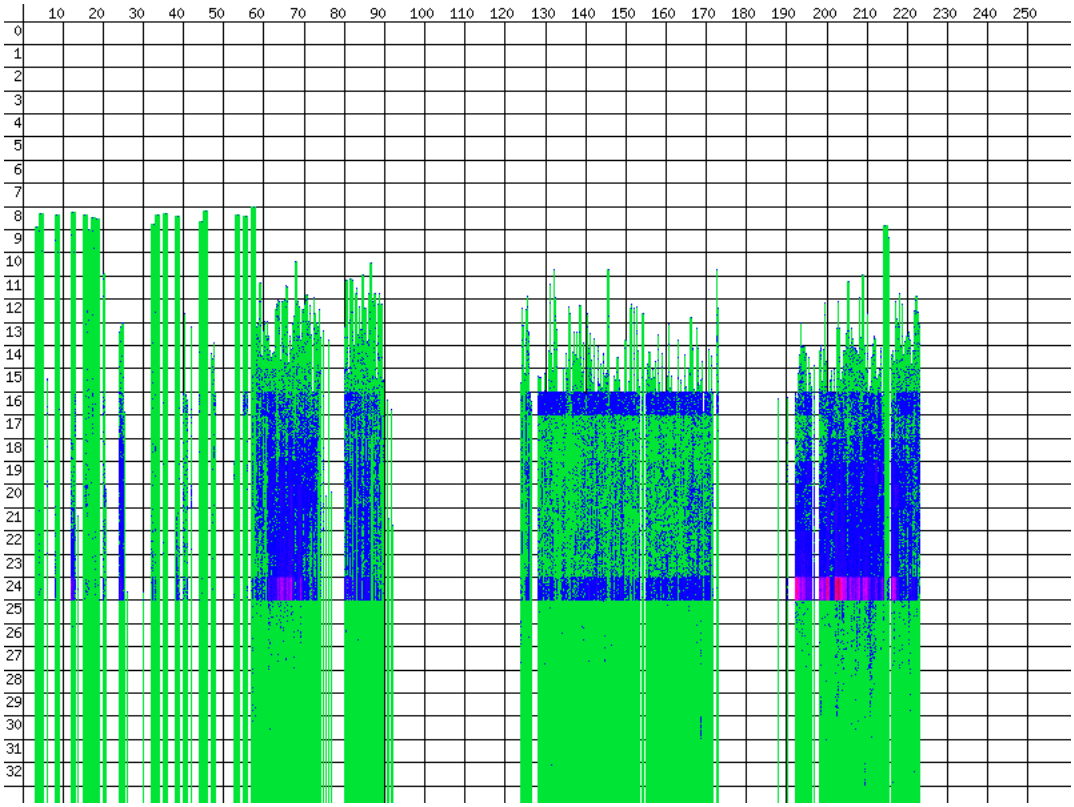


図 7-18　パンチングホール経路の状態

この図は、http://micho.mimora.com/routegraph/に紹介されている大久保氏が作成した経路を図化するツールを用いて描画している。

この図は、1つの経路が1つの点で描画され、横軸にIPアドレスの第一オクテットが表現され縦軸がカバーされるPrefix長となっている。また、1つの点は緑が基本で、複数の重なり合った経路があるほど、青や紫など色が濃くなってくように表現されている。

この図をみると 192/8 当たりに/24 の経路密度が多くなっていることが解る。このほか、青色のところが、ある経路に対するパンチングホールだと考えると、インターネットに流れている経路の広範囲にわたってパンチングホール経路が広告されていることが解る。

このようなパンチングホールが、意図的に正しく流されているのであれば全く問題はない。しかし、悪意を持って流された場合には問題となるが、この様なパンチングホール経路が正しいかどうかを判定することは通常非常に難しいと言える。

7.4. アドレス資源管理と経路

本節では、アドレス資源がどのように管理されているかについて解説する。

最初に、IP アドレスが同様な組織によってどういう形態で管理されているかを解説する。次に、その管理方針の考え方、その歴史について報告し、日本独自で持つ特殊方針の歴史についても言及する。

最後に、IP アドレスの管理と実際にその IP アドレスが経路情報という形でネットワーク運用の現場でどのように使われているか、そしてどのような乖離があるのかについて解説する。

7.4.1. アドレス資源管理の構造

IP アドレスは、IPv4 の場合、32 ビットの数値で表現可能な有限な資源として管理されている。インターネットの資源、つまり、インターネットで利用されるドメインや番号は、ICANN[12] (図 7-19) によって中心的に管理されており、その下部組織や管理を委譲された組織、そしてそれらの関係組織によってより細かい管理と運用がなされている。

特に番号関係については、ICANN の下部組織である IANA[13]にその管理がゆだねられている。

[12] Internet Corporation for Assigned Names and Numbers の略であり、民間の非営利団体です。(http://www.icann.org/)
[13] Internet Assigned Numbers Authority の略。

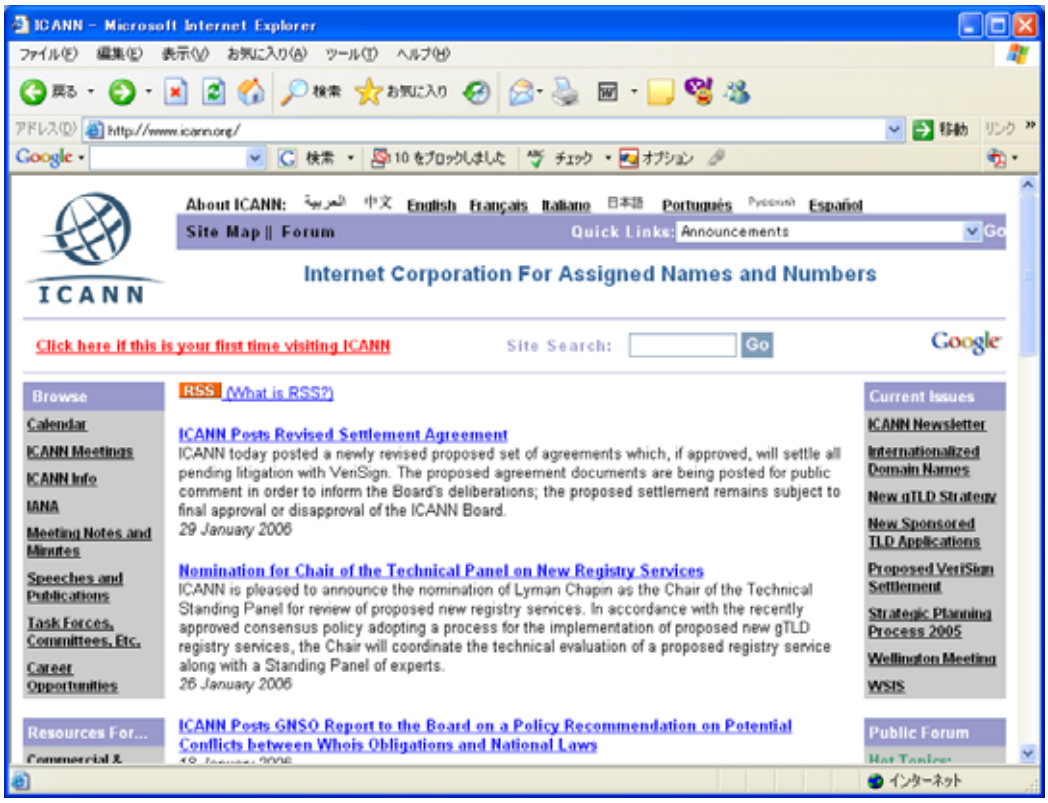


図 7-19 ICANN のホームページ

IANA では、/8 単位で IP アドレスの管理、つまり IP アドレスの上位 8 ビットだけを識別子として管理しており、それ以上の細かい管理は地域レジストリ(Regional Internet Registry : RIR)によって管理されている。

RIR は、管轄地域が分かれており、現在のところ ARIN(American Registry for Internet Numbers)、RIPE NCC(Réseaux IP Européens Network Coordination Center)、APNIC(Asia Pacific Network Information Center)、LACNIC(Latin American and Caribbean Network Information Center)、AfriNIC(African Network Information Center)の5つの RIR が存在する。

アドレス資源は、これらIANAとRIRを中心に階層的な構造を用いて管理しており、その階層図を図 7-20 に示す。

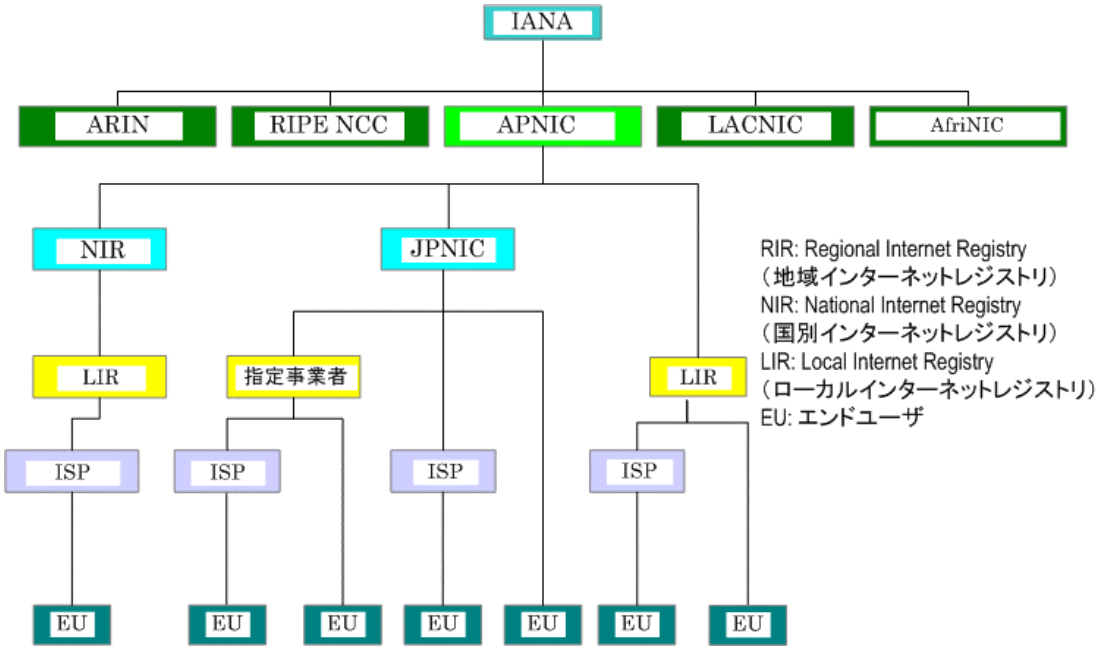


図 7-20　インターネットレジストリの階層構造

この図のように、IANA から RIR のアドレス資源管理は委譲され、そこから LIR と呼ばれる Local Internet Registry にさらに委譲され、最終的に ISP やエンドユーザーにアドレス資源が割り当てられ、世界中で唯一の IP アドレスとして利用することができるように管理されている。

ただし、APNIC および LACNIC においては、地域の中に多数の国が異なる文化を持ち成り立っている地域として、国別のインターネットレジストリである NIR(National Internet Registry)という、もう一つの中間階層が存在している。JPNIC もこの NIR に当たる。

NIR では、基本的に RIR の定めるルールに則った運用が求められているが、国毎に異なる特殊な事情を勘案し、より地域に根ざした方針を策定するための調整が行われている。

7.4.2. IP アドレス管理ポリシの考え方

前節では、アドレス資源管理機構の階層構造について述べた。これらの各 RIR では、アドレス資源を公平かつ効率的に分配するために、アドレス管理ポリシを策定している。

アドレス管理ポリシでは、アドレス資源を公平かつ効率的に分配すると同時に、各地域での特殊事情を勘案して策定されているが、それぞれの RIR が策定するアドレス管理ポリシの基本方針が異なると、地域格差が生まれたり、公平さが失われたりする可能性がある。そこで、IETF では、RFC2050[14]にて、このアドレス管理ポリシを作成するにあたって3つの目標を定めている。

以下に、ここで定める3つの目標について記載する。

(1) 節約

グローバルに一意なインターネットアドレス空間を、それらのアドレス空間を利用して運用されている ISP やエンドユーザーの必要に応じて、公平に分配すること。

アドレス空間をより長く利用できるようにするために、アドレス空間の先取りをさせないようにする。

(2) 経路制御適応性

グローバルに一意なインターネットアドレスを階層的に分配することで、それらのアドレスの経路制御的スケーラビリティが保たれる。インターネット内の経路制御が正しく動作するためには、このようなスケーラビリティが必要である。

ただし、IPv4 アドレスが割り当てられたからといって、経路制御に対する適応性が保証されるわけではない。

14 INTERNET REGISTRY IP ALLOCATION GUIDELINE (RFC2050)
http://www.ietf.org/rfc/rfc2050.txt

(3) 登録

> アドレス空間の割り振りや割り当てを記録する公的レジストリの提供。これは、アドレスの一意性を保証し、インターネットのトラブル対策のあらゆる場面で必要な情報を提供する。

また、この RFC ではこの目標に続いて以下の様にも記載されている。

> 上記の目標は、インターネットコミュニティ全体で満たされるべきものである。ただし、「節約」と「経路制御適応性」は対立しがちな目標であるという点に注意すべきである。さらに、上記の目標は時として個々のエンドユーザーまたはインターネットサービスプロバイダの利益に反する場合がある。したがって、個々のケースに応じて状況を注意深く分析し、判断して、適切な折衷案を見いだす必要がある。

このように、RFC2050 では、基本的な管理方針をガイドラインと示し、各 RIR がそれにそって地域ごとのアドレス管理ポリシを策定して運用できるようにしている。これにより、時代の変化によるニーズを吸収し、アドレスポリシを変更しながら長期にわたって共通の目標に向かって、アドレス資源の管理を行うことができるようになっている。

7.4.3. IP アドレス管理ポリシの変遷

インターネットが米国政府の強力な財政的援助を受けて発展していた 1990 年代初頭までは、NIC といえば世界で唯一の”The NIC”（SRI-NIC、あるいはその後身の nic.ddn.mil）を指していた。The NIC は、アドレスの取得やネームサーバーへの登録など、全世界からの申請を一手に引き受けて処理していた。

しかし、インターネットの急速な発展によって、この集中管理型の NIC 構造に変化が生じてきた。NIC に階層構造をもたせて、グローバルな The NIC から地域的な NIC への管理業務の分散化が図られるようになったからである。具体的には、従来 The NIC が各組織に対して直接割り当てていた IP アドレスをブロック化し、割当て業務を地域ごとの NIC に委任するようになった。この措置には、IP アドレスの地域的なまとまりを重視した CIDR という新しい経路制御技術の導入に対応する意味もあった。

1992 年 4 月には、ヨーロッパ地域を統括する RIPE（Réseaux IP Européens）の NCC（Network Coordination Center）が発足した。さらに 1993 年 4 月、The NIC は InterNIC として新たなスタートを切った。現在、InterNIC はグローバル NIC としての役割を引き継ぐととも

に、北米および周辺地域の NIC としても機能し、RIPE NCC などと協調しながらサービスを実施している。

このような状況のなかで、1993 年 1 月にホノルルで開催された APCCIRN（Asia-Pacific Coordinating Committee for International Research Networking：アジア・太平洋地域国際研究ネットワーク調整委員会）会議において、アジア・太平洋地域の NIC にあたる APNIC 設立に向けた調査・実験が提案された。APCCIRN の参加各国のあいだでも、このような地域的 NIC の発足を望む声が多かったため、APNIC の実験プロジェクトが開始されることになった。1993 年 8 月にサンフランシスコで開かれた APCCIRN 会議では、実験プロジェクトの期間を 1993 年 9 月～1994 年 6 月と定め、実験プロジェクトに関する報告を同会議でおこなうことが確認されている。

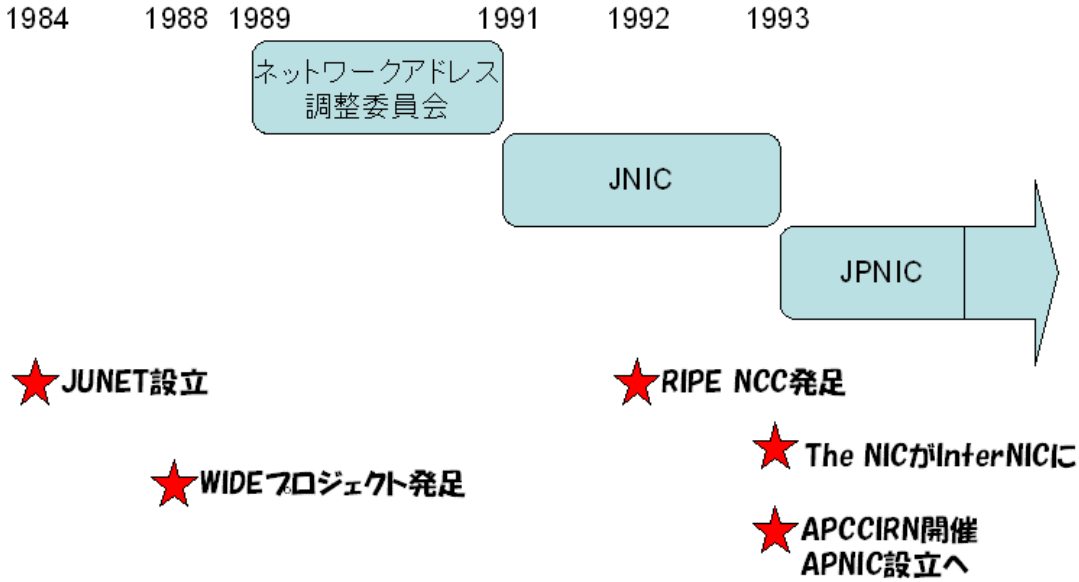


図 7-21 IR の歴史と出来事

APNIC は世界のレジストリの中でどのように位置づけられるかについては、図 7-20 に NIC の階層構造で示したとおりである。

地域インターネットレジストリの基盤が整備された当初、世界を北米、ヨーロッパ、アジア・太平洋を中心とする 3 つの地域に分け、それぞれを代表する地域 NIC（National Internet Registry）を設ける方向で議論が進められた。

この構想にもとづいて APNIC が正式に発足し、日本の国内 NIC（図 7-20 では National

Internet Registry) である JPNIC は、APNIC の下で相互に協力や支援をおこないながら機能していくことになったのである。そこで、JPNIC では、APNIC 実験プロジェクトに対し JPNIC 運営資金の 10%を上限として資金/資源を供与することとし、APNIC の発足に可能なかぎり協力する旨、合意され APNIC が設立されたのである。

その後、地域レジストリには、アフリカを担当する AfriNIC、ラテンアメリカを担当する LACNIC がさらに設立され、現在の 5RIR 体制として活動している。この様子を図 7-22 に示す。

図中にあるように、LACNIC は、主に ARIN の担当地域だったラテンアメリカ地域の管理を引き継ぐ形で新設され、同様に AfriNIC は、主に RIPE NCC が担当だったアフリカ地域の管理を引き継ぐ形で新設されたのである。図中では、ARIN が明確に ARIN と LACNIC に、RIPE NCC が RIPE NCC と AfriNIC に分割されたように表現されているが、実際には、多少の地域のずれがあることを付け加えておく。

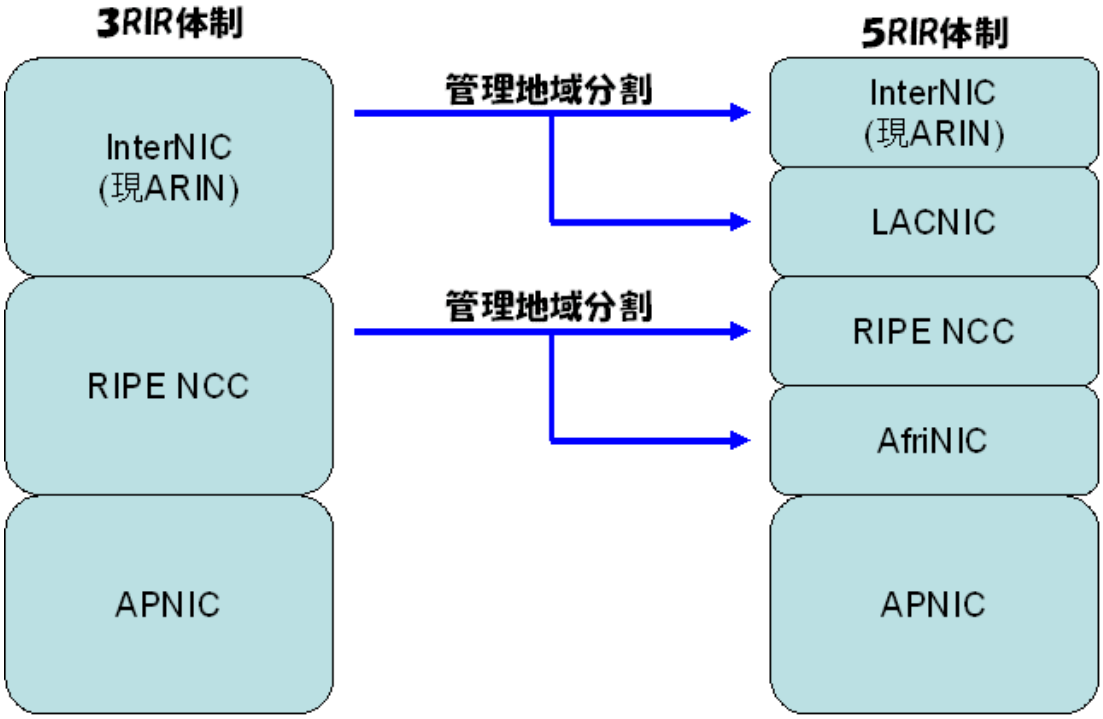


図 7-22　3RIR 体制から 5RIR 体制へ

一方、日本では、1984 年の JUNET 設立がその発端となる。JUNET は、東大、東工大、慶大を結んだ研究用のネットワークで、当時は UUCP をベースとして行われていた。その後、

1988 年に WIDE プロジェクトが開始され、日本でのインターネットの研究に加速がかかった。1989 年には、本格的にアドレスの管理を行うための「ネットワークアドレス調整委員会」が設置され、日本でのアドレスの管理が始まった。

その後、JNIC が 1991 年に発足し、アドレスの管理がネットワークアドレス調整委員会から引き継がれた。

さらに、インターネットの利用が加速され、1993 年には、JNIC は日本におけるサービスをより拡充するため、ネットワークプロジェクトを会員とする任意団体である JPNIC へとその姿を変え、その後社団法人化され、現在の形となっている。

図 7-21 に関連する出来事をまとめておく。

7.4.4. ローカルレジストリの特殊ルール

IP レジストリは、現在世界を5つの地域に分け、各地域に RIR を設置してアドレスの管理活動を行っている。そもそも、アドレスの割り振りなどにはインターネットの利用状況なども勘案した地域的な考慮が必要であり、それらの地域に依存する事柄は各地域レジストリの管理ポリシで吸収している。

特に、アジア太平洋地域の場合は、国数も多くまた経済状況もまちまちであり、APNIC における一つの運用ポリシで行うよりもさらに国別の考慮を行った国別レジストリが必要とされ、国によっては JPNIC の様な国別レジストリが APNIC のさらに下位層のレジストリとして活動をしている。

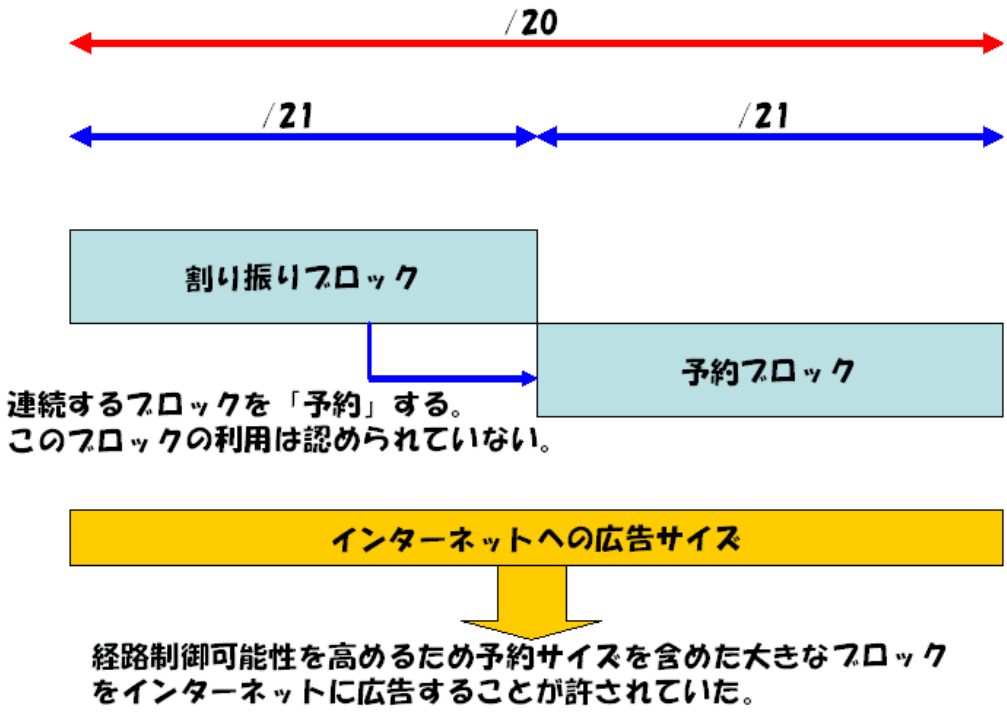


図 7-23　予約割り振り

これらのレジストリでは、RFC2050 や APNIC などの管理ポリシ、そしてインターネットの運用状況なども勘案しながら、地域に適切なローカルポリシの策定を行いる。日本の場合には、過去に「予約割り振り」というローカルルールが存在していた。予約割り振りについては図 7-23 に示しておく。

JPNIC が発足した当初は、まだ日本のインターネットは聡明期で、現在のように大規模な ISP が少なかった。さらに、レジストリからの最小割り振り単位は/20 などの比較的大きなものだったため、アドレスを割り振ってしまうとかなりの量が利用されずに死蔵されるのではないかとの懸念があった。さらに、当時のインターネットの経路制御では、細かい経路は経路フィルタで排除する等の措置も執られており、細かいアドレスを割り振り、経路制御性に問題が発生することも考えられた。そこで、JPNIC では、アドレスを割り振る際に、APNIC で定める最小割り振り単位のブロックを「予約」として扱い、その中から、/21 などの細かいアドレスブロックを割り振っていたのである。これが、予約割り振りが実施されていた経緯である。

その後、日本のインターネットも発展し、アドレスの需要も伸び、さらに、APNIC での最小割り振り単位もさらに小さくなったことから、この「予約割り振り」という制度は廃止されたのである。

7.4.5. 割り振り済みアドレスと経路情報

IP アドレスが ISP 等に割り振られ、それらすべてのアドレスが利用されるのが、アドレスを利用する最大効率である。しかし、経路制御性を考慮したり、ISP がサービスを続けてゆくためにあらかじめ取得するアドレスなどを考慮したりすると割り振ったすべてのアドレスを利用することは大変困難である。

本節では、割り振ったアドレスが実際どれくらいの割合で利用されているかを、インターネットに流れている経路から探ってゆく。

まずは、割り当て・割り振り済みブロックに対す経路情報カバー率について表 7-1 に示す。表中の「調査 prefix 数」は、レジストリから割り振られた「数」で、Prefix 長を考慮しないブロック数を表す。「(/24 にすると)」は、Prefix Length の長さが/24 のものとして換算した数を表す。

表 7-1 割り当て・割り振り済みブロックの経路広告カバー率

	PA		PI	
調査prefix数	1,937	100.00%	2,906	100.00%
(/24にすると)	134,140	100.00%	151,414	100.00%
exact matchのprefix数	1,060	54.72%	931	32.94%
(/24にすると)	100,508	74.93%	93,957	62.05%
広告されているprefixを/24にすると	130,446	97.25%	95,551	63.11%
広告されてないprefixを/24にすると	3,694	2.75%	55,863	36.89%

(JPNIC “経路情報の登録機構の検証”専門家チーム・メンバー 吉田友哉氏の資料より)

表の中では、カバー率を示すために PI と PA の二つに分けている。これは、インターネットレジストリが割り振り・割り当てを行っているアドレスには大きく、Provider Aggregatable(PA)と Provider Independent(PI)の 2 種類があるためである。グローバルにユニークな IP アドレスであるという点については基本的な特性に違いはないが、PA は ISP 等のさらに下位層のエンドユーザーに対してアドレスを割り当て、それらを集約 (Aggregate) する役割を担っているアドレスブロックであり、PI はそれら集約の役割を担わずに直接エンドユーザレベルに対して割り当てるためのブロックをさしている。

現在では、レジストリから割り振られるアドレスのほとんどが PA であるが、歴史的な理由や、特殊な用途のために PI が割り当てられているケースがある。特に、CIDR 以前に割り振られたアドレスを歴史的 PI とよび、特別に扱うケースもある。このため、PI は割り振られたネットワークでの実際のアドレス需要を適切に判断しなくて良い時代に割り振られているケースのものが多く、実際の利用率も低くなっていることが理解できる。

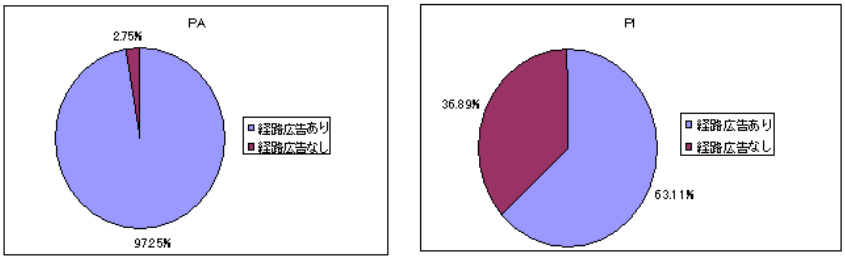
この PI、PA という前提を含めて表を再び解説する。

レジストリから割り振られるアドレスは、特に PA では ISP で集約可能な比較的大きなアドレスブロックになっている。理想的には、この割り振られたサイズそのままをインターネットに経路情報として広告することが望ましく、これにより経路数の爆発が抑制されることが期待できる。

このとき、すでに割り振ったブロック数は、PA で 1937、PI で 2906 となる。これに対し、割り振りブロックサイズがそのまま経路情報として流れているケースは、PA で 1060、PI で 931 であり、PA ですら 54%にしか及ばない。しかし、/24 換算で見た場合、PA で 75%、PI でも 62%とカバー率は高くなる。これは、多くの割り振りブロックが、割り振られたサイズそのままでインターネットに経路広告されているわけではなく、経路制御の都合などで、分割して広告されているという状況が推測できる。

PA と PI の比較についてさらに細かく対比したものを図 7-24 に示す。

- PA
  - ほぼ経路広告されている
  - 経路広告が無いアドレスは、現在広告準備中のものも含まれているので、経路広告ありの比率は実際にはこれ以上になるかもしれない
- PI
  - 6割程度しか広告されていない
  - Exact match で広告されているprefixの経路広告割合が62.05%で、全体の広告率（＝63.11%）に近い≒広告 or 未広告の結果になっている（後述の結果詳細を参照）



(JPNIC “経路情報の登録機構の検証”専門家チーム･メンバー　吉田友哉氏の資料より)

図 7-24　PA と PI の経路広告カバー率の比較

図にもあるように PA は、インターネットレジストリによってアドレスの利用効率を十分に考えられ運用されているブロックであるため、非常に利用効率が高いことが解る。一方、PI は利用効率が良くないことが解る。

注意したいのは、今回のこの評価はあくまでインターネットに流れている経路情報に対して、レジストリが割り振ったアドレスを重ね合わせたにすぎないということである。IP アドレスは、グローバルに一意性を保証して割り振り･割り当てを行うもので、本来、グローバルな経路表にそのアドレスを載せることが目的ではない。

実際には、インターネットに対する接続性を持たないが、多くの顧客を抱えるネットワークでは、プライベートアドレスの重複を避けるなどの目的で、一意性を確保する必要があり、レジストリからグローバルアドレスが割り振られているケースもある。このようなアドレスは、今回の調査のなかでは「未使用」に分類されてしまう。このため、PA、PIの利用率は、実際にはもう少し高くなると考えられることを付け加えておく。

7.4.6. レジストリによる割り振りと経路情報の溝

本章では、アドレスという資源の管理について述べ、それらのアドレスがインターネットの運用の現場に経路情報としてどのように扱われているかについて述べてきた。

本章で明らかにしたことは、アドレスの割り振りポリシは、時代背景と共に様々な変遷を遂げてきたこと、そして、それら割り振られたアドレスは、その時々で最大のパフォーマンスを出すようにアドレス割り振り・割り当てポリシを運用する現場で利用されてきたことである。

この割り振りポリシとその運用は、RFC2050 で扱われているように、効率的な割り振りと経路制御可能性という面でしばしば対立しがちである（図 7-25）。その背景を表しているのが 7.4.5 節で解説した割り振りブロックと経路情報の対比である。

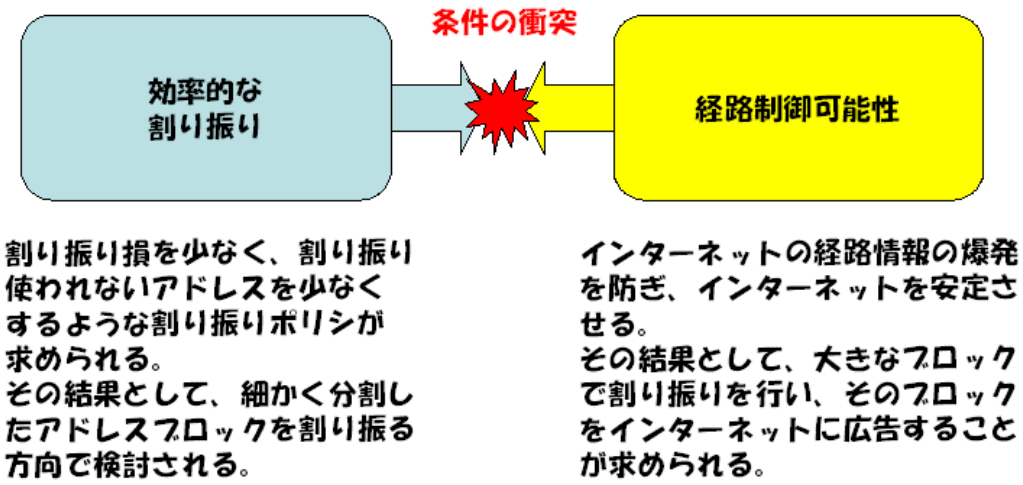


図 7-25　アドレスポリシの衝突する考え方

たとえば、レジストリのアドレス割り振りの運用現場では、割り振り効率を上げるために、/24 単位での割り振りが実際に行われていたのである。アドレスを割り振るという現場での運用は、この判断は非常に正しい判断だと言えるだろう、一方、経路制御と言う観点からは、/24 を 3 個と割り振られた場合には、/23+/24 という 2 つの経路として広告しなくてはならなくなるのである。逆転の発想で言えば、このような割り振りが行われていたために、割り振りサイズがそのままインターネットに広告できないという事態を引き起こしたのだともいえるのである。

現時点では、経路制御可能性とアドレスの割り振りの効率の両方が融和され、このようなことが少なくなってきているが、割り振り効率と経路制御可能性という2つの異なる目的を、同一のポリシのなかで運用することは困難といえ、異なるポリシの中で調和のとれた運用を行うような調整が今後も必要となっていくだろう。

7.5. インターネット経路制御の問題点

IP アドレスは、レジストリによって管理され適切な手続きによって ISP などの IP アドレス利用者に割り振られる。この手続きによってISPはグローバルユニークなIPアドレスを利用可能になり、このグローバルユニークな IP アドレスを利用してインターネット接続を行い様々なサービスが行われている。

割り振られた IP アドレスは、最初に BGP を利用して世界中に自分に割り振られたアドレスがインターネットに接続したことを広告する。BGP に対するアドレスの広告は、先にも述べたように経路広告といい、通常インターネットの接続サービスを提供する ISP 等に接続して、その ISP 経由で行われている。広告された経路は接続先 ISP を通じて世界中に伝搬しゆくのである。

しかし、この経路の伝搬には様々な問題がある。本節では、これらの問題点について解説し、現状での問題の解決方法について触れてゆく。

7.5.1. BGP 運用の根本的問題

先ほどの述べたように、インターネットに広告した経路は BGP を利用して世界中に広告される。これは、BGP を利用して世界中に点在する AS に伝搬してゆくことによって行われている。

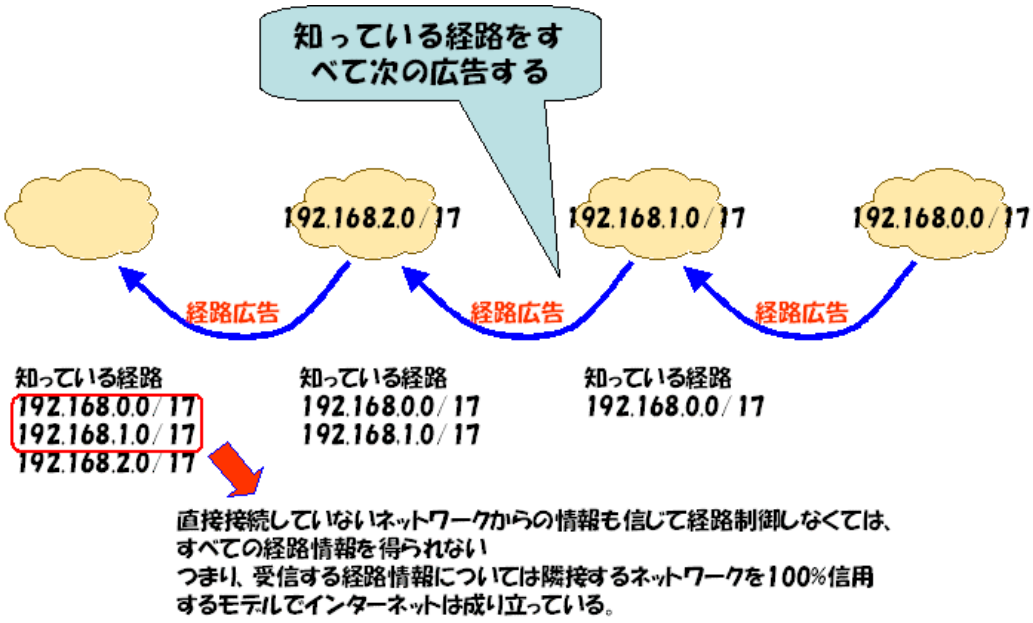


図 7-26　BGP の隣接ネットワーク信用モデル

そこで、いったん BGP での経路広告はどのように行われているかに戻って考えてみることにする。

BGP の経路広告は、異なる AS 間で BGP 接続を通じて、自分のネットワークが転送可能である転送先のアドレス情報を経路情報として、接続相手に広告するということである。一方、経路を受ける側は、自分が転送しようとしているパケットの転送先を他の AS から広告された経路情報によってのみ解決して転送してゆくことになる。この様子を図 7-26 に示す。

ここで問題となるのが、「他の AS から広告された経路情報のみによって解決」されるということである。BGP の世界では、この「他から広告された経路情報」をさらに他の AS に自分が転送可能なアドレス情報として広告するということを繰り返して経路情報が伝搬してゆくのである。これは、必ずしも自分が受信した経路情報が隣接する直近の AS、つまり自分の AS と直接接続している AS が広告した経路情報が必ず到達可能であるということを保証していないのである。

性善説に基づいて BGP の世界を考えた場合、つまり、経路の広告主 (Origin) は必ず正しい経路情報を広告し、経路を転送する途中の AS ではその情報に広告主に到達不可能となるような不正な変更を加えずに経路を転送するという場合には、現在の BGP のモデルでは何ら問題なく動作する。

しかし、これを性悪説で考えた場合には、いくつかの問題が生じてくる。考えられる状況を以下に列挙する。

(1) 広告主は、自分が割り振られている以外の経路情報を広告する。

先にも述べたように自分が経路情報として広告可能な経路情報は、自分がインターネットレジストリから直接割り振られたアドレス、もしくは、その管理権限があるアドレスに対する経路情報と考えるのが妥当である。

しかし、何らかの理由により、自分が管理すべき経路情報以外の情報をインターネットに流すことによって、未使用のアドレス空間を不正に利用したり、他の AS が利用中のアドレスを利用したりすることができるのである。

特に、他の AS が利用中のアドレスに対する経路情報が、関係のない AS から不正に広告された場合には、利用者が目的のネットワークにアクセスできなかったり、場合によっては犯罪に利用されたりする可能性もある。

この様子を図 7-27 に示す。

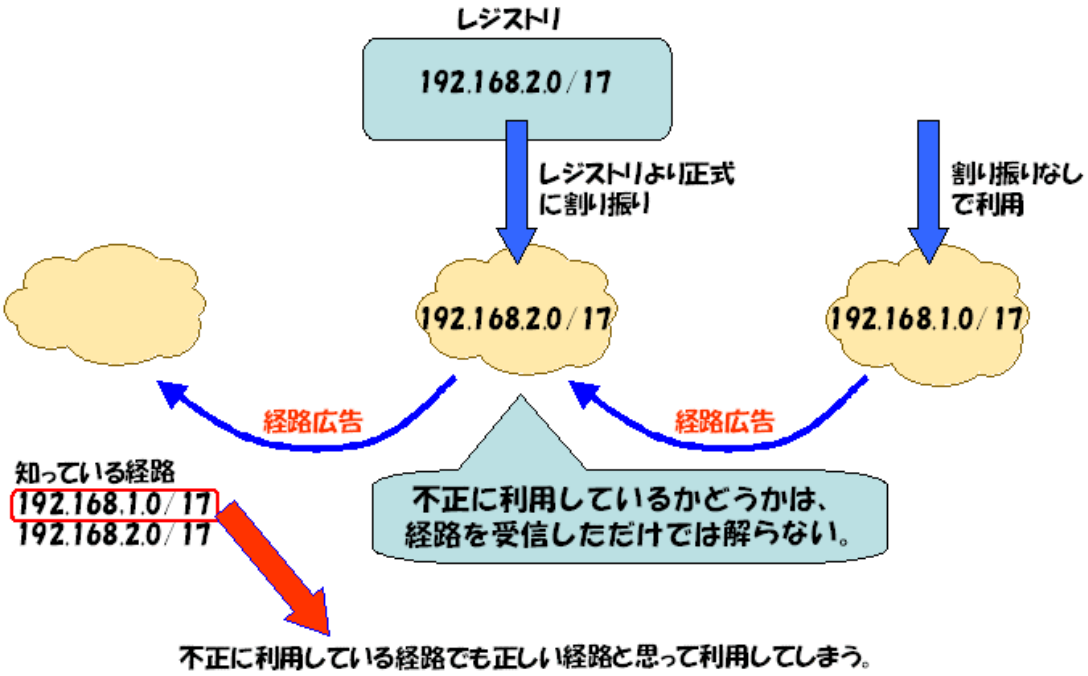


図 7-27　割り振られていないアドレスを利用する例

(2) 経路転送する途中の AS が、広告主を偽るような経路情報の改ざんをする。

BGP は、自分がパケットを転送可能なネットワークの情報を他の AS に経路情報という形で伝えるものあり、この繰り返しによって経路を世界中に転送している。つまり、ある AS に到達する自分の AS の経路情報は、複数の AS をまたがって転送されていることになる。

しかし、悪意を持ってこの転送を行うと、自分が受け取った経路情報を不正に改ざんして、他の AS に転送してしまうことが可能となる。この場合も(1)と同様な被害を考えることができる。

この事象に関する詳しい説明は、7.5.2.2 節で解説する。

(3) 意図的にインターネットを乱すためにある AS が経路情報を不正に流出する。

BGP 接続をしている AS は、自分を広告主としてしまえば事実上どのような経路情報もインターネットに広告可能である。しかし、ルータは制限のあるメモリ量、制限のある CPU 資源などを持った機器で構成され、不用意に大量の経路情報を送り込んだりすれば、これらの資源は使い尽くされ機器が停止してしまう可能性がある。

現在のインターネットの経路情報は、フルルートで約18万経路です。これに対して5万経路や10万経路という大量な経路情報を送り込めば、インターネットを構成するルータなどの機器資源は使い尽くされ一部で機器の停止などが考えられ、ひいてはインターネット全体の混乱を及ぼしかない。

この様子を図 7-28 に示す。

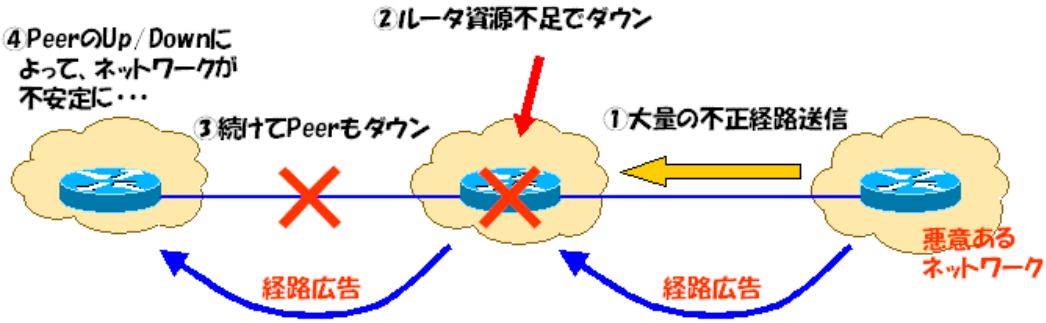


図 7-28　大量経路送信が原因でネットワークが不安定に

これらの例は、悪意こそないと考えられるが、運用上のミスなどによって酷似するケースのトラブルが過去に発生しており、今後深刻な事態に繋がる可能性がある。

特に、(1)や(2)で述べたような経路情報のなりすましや改ざんは、フィッシング詐欺などにつながりやすく今後のインターネットの安全性を考えるうえでは、非常に重要な問題である。

これらの問題の原点は、BGP は性善説に則り、接続先である隣接 AS からの情報を完全に信用するモデルで考案されたプロトコルであると言うところにある。

最近では、本節であげたような問題が指摘されているため、いくつかの問題回避策が考えられ実践されている。

以降に、悪意のあるなしにかかわらず、BGP を利用した経路伝搬の仕組みの中で運用上問題となる点について、より具体的に解説し、その後、回避手段などについて解説する。

7.5.2. 発生する可能性が高い問題

前節では、BGP が隣接 AS を信用するモデルで考案されたたためにいくつかの問題が発生する可能性があることを述べた。

本節では、この発生する問題のうち、比較的よく観測されるものと、悪意性の高い事例について取り上げ解説する。

7.5.2.1. 設定ミスによる影響

BGP は、隣接 AS から広告される経路情報を信用して経路制御を行う。これは、設定が正しく行われている場合、そして、悪意のない正しい情報が登録されている場合には、一切の問題が発生しない。

しかし、ルータなどの設定は基本的に人が行うものであること、そして、機器はしばしば誤動作を起こすものであることを考えると、ネットワークの運用者に悪意が全く無くても、正しくない経路情報がインターネットに流れ、一部のネットワークに問題が発生する、または、インターネット全体に及ぶ問題に発展する可能性すらある。

たとえば、ある内部経路が数万経路ほどある比較的大きなネットワークがあったとする。通常、IGP で扱う内部経路は、BGP のような経路制御プロトコルには転送しない。これは、IGP の様な細かい経路がインターネット全体にとって重要ではないことやインターネット全体の経路情報を減らすという意味もあるからである。

しかし、ルータの設定では、比較的簡単に IGP の経路情報を BGP の経路情報として広告するということが出来てしまうのである。(このような設定用のコマンドが用意されているのは、実際にそのような運用形態もあるからである。) 仮に、このルータの設定が間違えて実行されてしまった場合には、IGP に流れている数万経路がインターネットに流れてしまうのである。

現在のフルルートは約18万経路である。一方、ルータで管理可能な BGP の経路数は、ルータのメーカや性能に相当依存するが、通常25万経路程度のものもあり、これに IGP から8万経路も流れれば、合計が26万経路となり、ルータは BGP の経路情報が扱えなくなる。このとき、ルータは、経路情報を受け取るのをやめるか、リブートをするかのどちらかの動作をとることが多い。

リブートした場合は、そのルータが接続していた他の AS との接続が途絶え、再接続時に再度フルルートが交換され、場合によっては、ここでのトラブルと元となっている 26 万経路の一部も転送される可能性もある。この経路情報が他の隣接 AS に転送され、そこのルータの資源が足りなかった場合には、また同じような状況が発生し、トラブルはどんどん広がってゆくと言うことになる。

このトラブルは、数年前に実際に米国で発生している事例である。単純にルータの設定のミスだったことが解っているが、インターネットにとっては相当大きなトラブルだったと言って過言ではない。

このようにトラブルが連鎖してゆくことを「障害連鎖」といい、総務省の次世代 IP インフラ研究会[15]（図 7-29）の議論のなかでも取り上げられ、対策が必要であると議論されている。

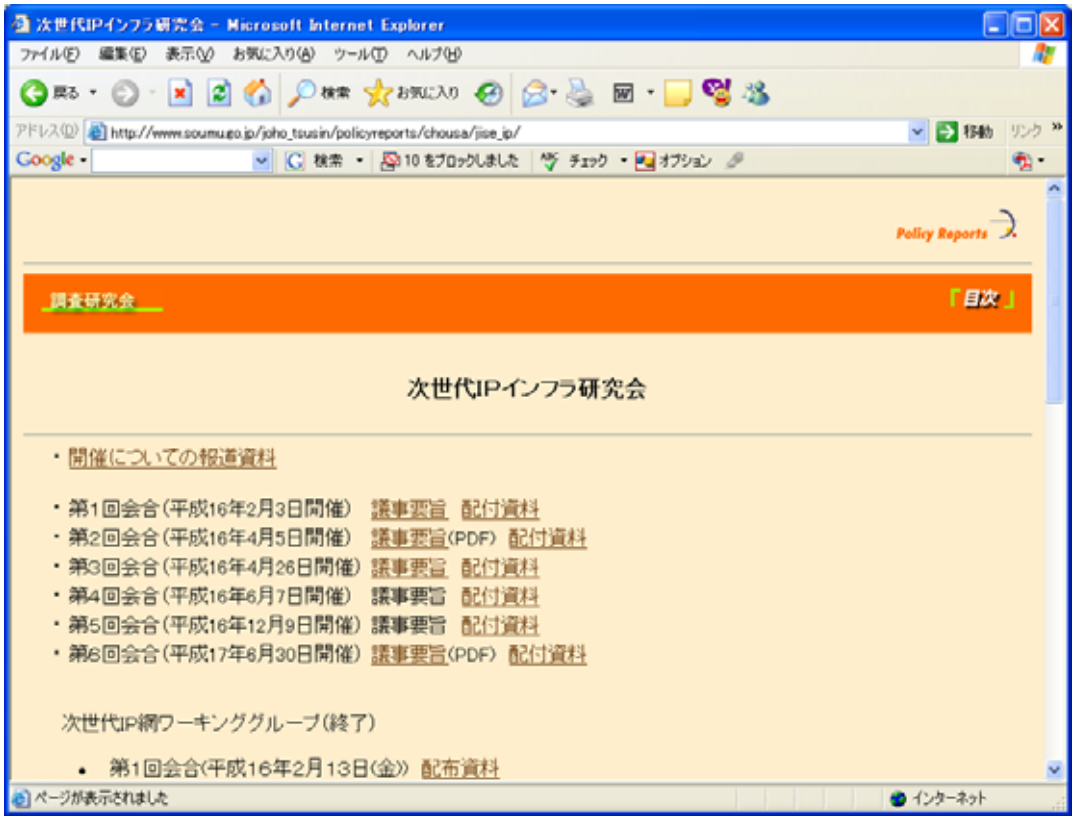


図 7-29　次世代 IP インフラ研究会のホームページ

[15] 次世代 IP インフラ研究会、次世代 IP 網ワーキンググループ（第 3 回会合）（平成 16 年 3 月 10 日開催）にて取り上げられた。
http://www.soumu.go.jp/joho_tsusin/policyreports/chousa/jise_ip/040310_3.html

この数万経路が流出するというトラブルは比較的極端な例ではあるが、細かいものでは、広告する経路自体の設定をミスしてしまうと言う例も以外と多く聞かれている。

たとえば、192.168.0.0/16 という経路を設定するところを、192.169.0.0/16 と設定してしまった場合などである。もちろん、ここであげた経路は単なる例だが、設定された例では、第二オクテットが1だけ違っている。広告主は、192.168.0.0/16 の割り振りを受けて、正しく広告しているつもりなのだが、実際には 192.169.0.0/16 がインターネットに広告されることになる。これにより、正しく 192.169.0.0/16 の割り振りを受けている AS は、インターネットの一部からのアクセスが不能になってしまうのである。

このように、経路情報の広告、そしてその伝搬には、単純な設定ミスで思いもよらぬトラブルを引き起こすことが解っている。このようなトラブルが無いように適切な回避策を講じておく必要がある。

7.5.2.2. 経路ハイジャック

もう一つの重要な問題として経路ハイジャックの問題がある。経路ハイジャックとは、現在利用中のネットワークに対する経路情報と同等、もしくは対象ネットワークの乗っ取りたい一部の経路情報をインターネットに広告し、正しく利用しているネットワークへの到達性を奪ったり、悪用するために正しいネットワークへのアクセスを横取りしたりすることを言いう。

乗っ取られた場合の状況として、正規な経路情報と同等もしくは、そのネットワークに含まれる経路情報がインターネットに広告されるため、それが悪意を持っているのか、単なる設定ミスであるのかについては見分けづらい。経路ハイジャックといった場合には、一般的に悪意を持ってこれらの行為を意図的に行っている場合を指す。

この経路ハイジャックは、BGP や経路制御方法の落とし穴をついて行われる。経路制御は、基本的に「Longest Match の原則」がある。経路情報は、Prefix と Prefix Length という2つの情報によって構成されるが、同一の Prefix が存在する場合には、Prefix Length がより長いもの(細かい経路)が優先されるという原則のことである。この Longest Match の原則が経路ハイジャックを行うため手法になるのである。

この様子を図 7-30 に示す。

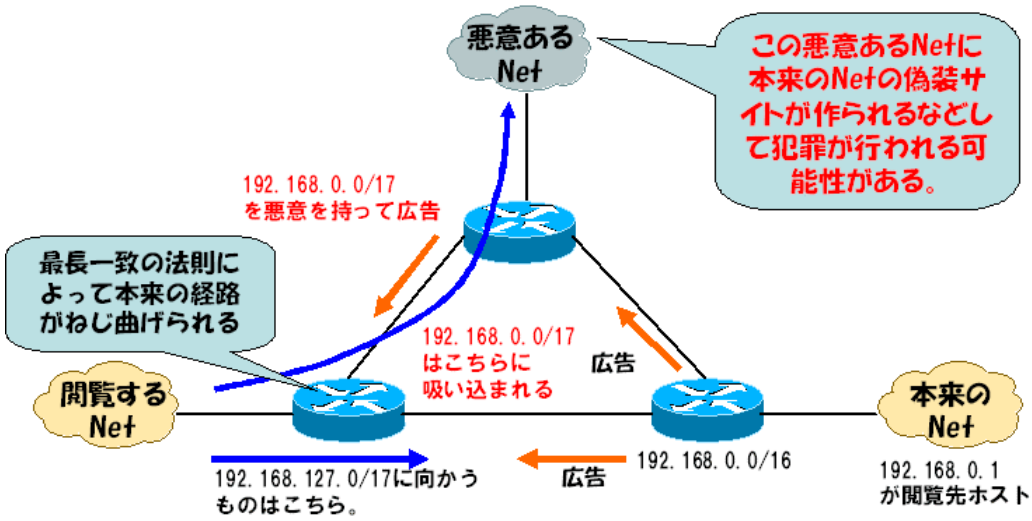


図 7-30　経路ハイジャックの手法

たとえば、AS-X が 192.168.0.0/16 という経路をインターネットに広告していたとする。このときの Prefix は 192.168.0.0 で Prefix Length は 16 である。これに対して悪意をもった AS-Z が、192.168.0.0./17 という経路をインターネットに広告した場合、Prefix は 192.168.0.0 で同じでだが、Prefix Length は 16 と 17 で、悪意を持った AS-Z の方がより長く、経路制御の特性上 AS-Z が優先すべき経路情報として扱われてしまうのである。AS-Z としては、この優先される経路情報を用いて、通常 AS-X に向かうトラフィックを横取りし、場合によってはフィッシング詐欺のためのサーバを立ちあげるなども考えられるのである。

このような例は、最近になって見られるようになってきた。自分のネットワークの経路情報を安全に伝える、また、不正な経路を受け取らないようにするための運用手段の確立、というものが必要になってきていると言える。

7.5.3. BGP 運用による問題回避手段と実情

本節では、前節までに述べた BGP が持つ問題について、現時点での運用上の対策について述べる。

7.5.3.1. 問題回避手段

前節で述べたような問題が発生する原因は、以下の 2 つの点から発生している。

- 広告元 AS が意図しない経路情報を広告している。

- 経路情報の受信時に広告元が意図している経路情報以外の経路情報を受け取っている。

逆に言えば、これらの意図しない経路情報を広告したり受信したりしなければ、経路情報は安定するということが言える。

多くのルータでは、送受信する経路情報を制限するための経路フィルタの機能が実装されており、これを利用して意図しない経路情報を遮断することができる。経路フィルタには、以下の様な種類がある。

(1) AS パスフィルタ

ASパスフィルタは、経路情報がどのようなASによって伝搬されてきたかを示しているBGPの経路情報の属性に含まれているAS Pathによって受信経路をフィルタするものである。

このフィルタでは、経由するASや、どのＡＳを通過してきたかという情報によって、その経路情報を受信するかどうかを決定する。たとえば、広告元のＡＳ番号が1で、その経路情報がＡＳ2と3の両方を経由してＡＳ4に到達しているような場合、ＡＳ4では、ＡＳ2経由の物だけを受信する、というような操作が可能になるのである。

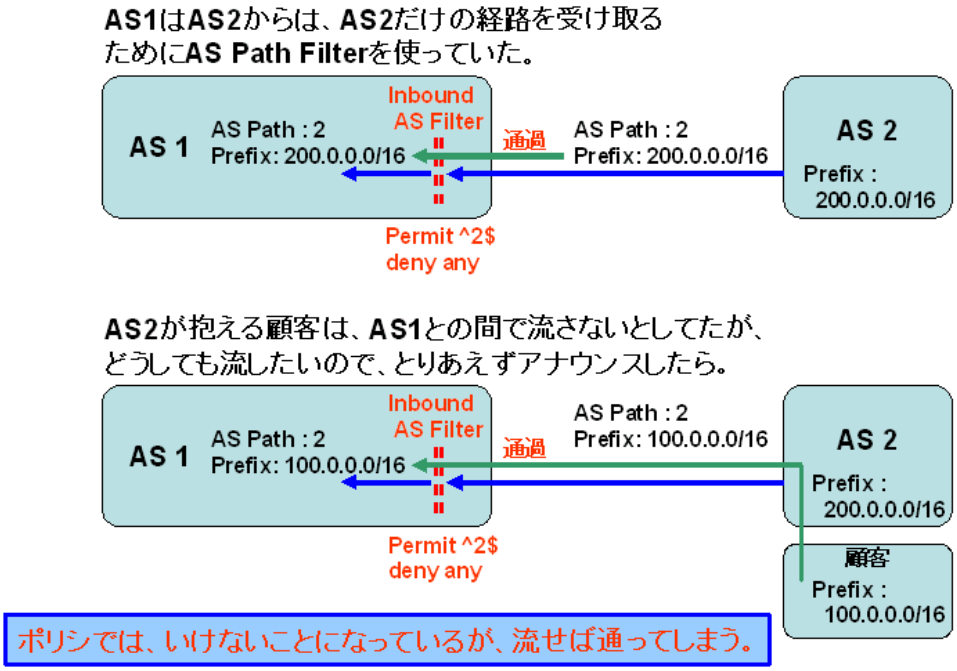


図 7-31　AS Path Filter の例

AS パスフィルタの動作の例について図 7-31 に示す。

一方、AS パスフィルタは、その AS 番号によってのみフィルタされるため、AS パスフィルタによって定義された条件だけを満たせば、実際に送られてくる Prefix の情報が不正な物であっても防げないという弱点もある。

(2) Prefix フィルタ

Prefix フィルタは、受信する経路の Prefix 情報に条件を設けることで経路情報をフィルタするものである。Prefix フィルタは、Prefix と Prefix Length の 2 つの情報によって 1 つの Prefix フィルタを構成する。また、Prefix Length は、多くの場合、設定した長さより「より長い」、「より短い」または「同じ」などの条件を付けることが可能で、Prefix に対して柔軟なフィルタを提供している場合が多い。

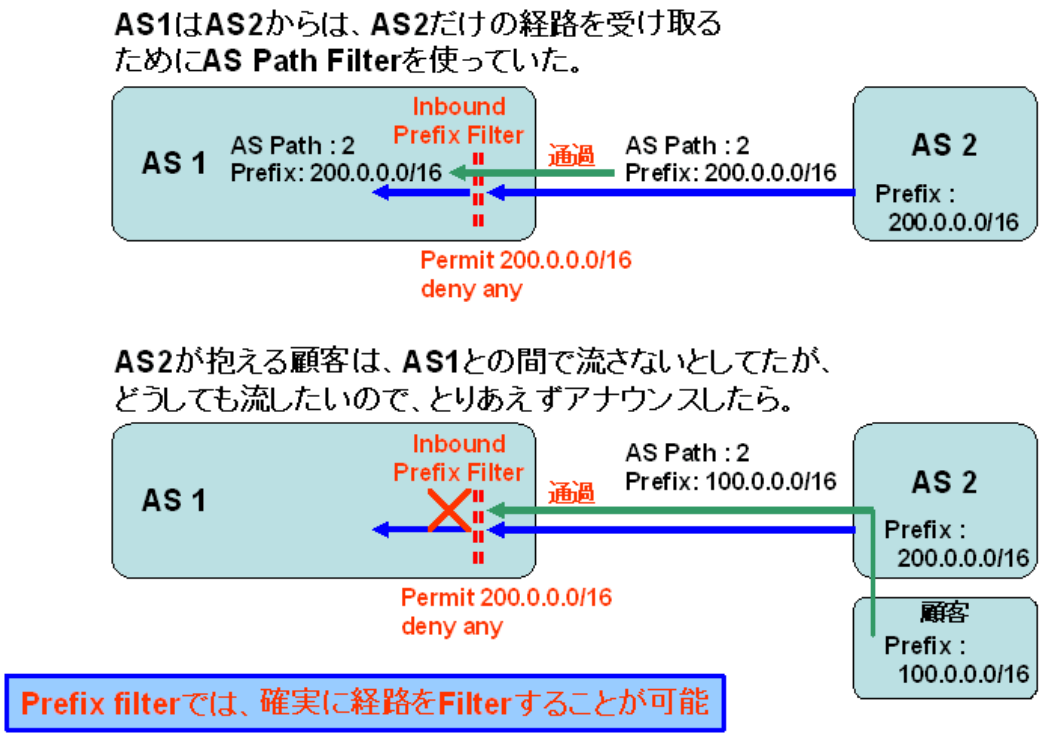


図 7-32 Prefix Filter の例

これは、経路情報として 192.168.0.0/16 とインターネットに広告されてしまっている場合でも、経路制御を行いやすくするために、192.168.0.0/17 と 192.168.128.0/17 の2つの経路情報に分割してインターネットに広告する場合などがあり、このような場合でも「/16 より長い物は受け付ける」と言うように設定することで、ある程度範囲を持ったフィルタの設定をすることが可能になるからである。

Prefix フィルタの動作の例について図 7-32 に示す。

一方、Prefix フィルタは、Prefix 情報をすべて網羅した設定を必要とするため、設定する情報が膨大になり、それによって管理も煩雑になり、設定ミスを引き起こす原因にもなりかねない。また、Prefix 情報が正しければ、経路情報を受け付けてしまうことから、意図しない AS から経路情報が広告されていても、問題なく受け取ってしまう可能性がある。

このように先に挙げた問題を解決する為の「経路フィルタ」という手段は、運用上の手法で対策は可能である。しかし、各フィルタの項目で挙げたように、どちらかのフィルタだけでは、

問題を完全に解決できず、これらのフィルタを複合的に利用することが不可欠である。

さらに、これらの情報を取得する手段についても問題がある。BGP の接続は、ISP から BGP の接続を購入するケース、IX 等に接続して他の ISP と相互接続するケースなどがある。IX に接続するケースは通常「対等ピア」と呼ばれ、お互いの AS の情報とその顧客の AS の情報を交換するにとどまる場合が多い。このため交換する経路情報が比較的少なく、メールなどで広告予定の経路情報を AS パスや Prefix の情報であらかじめ交換し、フィルタを設定しておくことで十分な対策が可能である。しかし、このようなことが可能なのは、実際には対等ピアの数が少なく、BGP の顧客も少ない場合である。これらの数が多くなってきた場合には、現実的にすべてに対してフィルタを設定することは難しくなり、何らかの自動的なフィルタ設定手段を講じるか、網羅的に対応可能なフィルタを考案して設定しておくなどの対応をするしかなく、対応策が機能としては存在していても現実的には設定が難しい状況と言える。

7.5.3.2. 現実的な対応例

前節では、フィルタの手法について述べた。このなかで、それらの手法を利用して細かく設定して、危険を回避することは、フィルタの設定量などの問題で現実的ではないと述べた。

そこで、本節では、現実的な範囲として実際に行われているフィルタの例について紹介する。

(1) IX での AS パスフィルタ

主に日本の IX で行われている。基本的に広告される Prefix に問題がある場合には完全に防ぎきれない AS パスフィルタであるが、IX で行う「対応ピア」は、通常 IX に接続しているもの同士である一定の合意を行い、場合によってはそれを書面で交換するなどの手続きを行って BGP 接続を行っている。また、IX で交換される経路情報は、隣接ASとその AS の顧客の情報を交換するというのが一般的で、交換する経路情報自体がASの管理の範疇であるため、そもそも安定した経路であると言える。そこで、IX での BGP 接続では、隣接 AS が管理している AS に限定することを目的に AS パスフィルタのみで対応しているケースが多いようである。

但し、IX にて AS パスフィルタを多用するケースは、日本における特徴的な運用方法で、米国などでは、そもそもフィルタをしないケースや、Prefix フィルタで行うことが多いようである。

(2) Prefix Length によるフィルタ

Prefix フィルタの一種になるが、特に Prefix Length に注目してフィルタを行う場合を指している。このモデルのフィルタには、より大きな Prefix をフィルタする場合とより細かい Prefix をフィルタするという二つのケースがある。

✧　より大きな Prefix をフィルタする

レジストリから ISP 等にアドレスを割り振る際に経路制御可能性を考慮し、「最小割り振りサイズ」というのが決められている。これは、経路情報が細かくなりすぎて経路数の爆発を防ぐために、ある程度大きな量を ISP に割り振ることで、インターネットに流れる経路情報の細分化を防ぐ為に決められているものである。

この最小割り振りサイズは、/8 単位で決められており、IANA のページや RIR のページで参照することができる。

この情報を利用すれば、意図的に分割して経路情報として流す場合を除けば、大幅に細かい経路情報としてインターネットに流れてくることは考えにくいため、この値を基準にフィルタをすることが可能となるのである。

✧　より細かい Prefix をフィルタする

BGP に流される経路情報は、先にも述べたようにレジストリからの最小割り振りサイズに大きく左右される。一方、この最小割り振りサイズは、CIDR が実施された後に考案された物で、それ以前の物は、クラスフルな割り当てが行われ、結果的にクラスフルな単位での経路情報がインターネットに広告されていたのである。

特にクラス C のアドレスは、/24 という経路情報として出現し、BGP の経路情報としては比較的細かい値となるが、この/24 に相当する経路情報は、今なお経路情報全体の約半数を占めているのである。

そこで、/24 という経路を一つの基準にして、インターネットに広告する経路情報としては、それ以上細かい経路は流さないようにという無言のルールが存在しており、このルールに従ってフィルタをしているケースが多いようである。

(3) その他最低限のフィルタ

最後にその他最低限実施すべきフィルタについて整理する。

このような最低限のフィルタは、JANOG（日本ネットワーク・オペレーターズ・グループ）で議論され、ドリーム・トレイン・インターネットの馬渡氏や株式会社インターネット総合研究所の平尾氏らの提案によって、JANOG の第 16 回ミーティングにて議論され、採択されている。[16]

16 本報告書執筆時点の 2006 年 2 月 9 日の時点では、JANOG のホームページにはまだ掲載されていないようである。詳しくは、議論を行った Inter-domain Routing Security Workshop のホームページから参照できる。(http://www.bugest.net/irs/)

7.6. インターネットルーティングレジストリ

本節では、インターネットルーティングレジストリ（以降、IRR）に関するさまざまな状況および実態について報告する。

最初に、IRR の歴史について解説し、IRR が開発された背景について報告する。次に、現在の国際、およびアジア太平洋での IRR の実態について報告する。次に、これら IRR が現時点で期待されている機能などについて解説し、さらに IRR の現状について報告する。最後に、これらの現状を踏まえ、実際の IRR が利用されている実態について報告する。

7.6.1. IRR の歴史

IRR を研究していた有名なプロジェクトの一つに米国の Merit、University of Southern California Information Sciences Institute(ISI)、Cisco Systems、University of Michigan とその関係者によって提供された Routing Arbiter Project(以下、RA プロジェクト)がある。

RA プロジェクトは、Route Server、Network Management System、Routing Arbiter Database(RADB)、 Routing Engineering Team の 4 つのプロジェクトで構成されていた。この RA プロジェクトは、NAP[17]での経路交換において、NAP 上でフルメッシュの BGP Peer をするのではなく、Route Server を使って、単一の Peer を Route Server に接続することで、BGP の経路情報の交換をスムーズに行うことを目的としていた。この Route Server には、データベースに蓄積された経路情報とポリシ情報を利用して、BGP Peer 間の経路制御を実装する機能があり、ここで利用するデータベースが RADB であり、IRR である。

RADB に蓄積されるポリシ情報は、RIPE-181 という形式で記述されていた。RIPE-181 形式は、ヨーロッパにある RIPE NCC によって開発されたもので、その原型は、1980 年後半に NSFNET のバックボーンルータを設定するために利用されていた PRDB(Policy Routing Database)である。PRDB は、この RADB と RIPE-181 の出現によって、1995 年までに置き換えられている。

[17] NAP(Network Access Point)：現在の IX(インターネットエクスチェンジ)に類似したもの

その後、IRR の記述言語である RIPE-181 はいくつかの拡張をされながら利用されてきたが、1999 年 6 月に RFC2622[18]として RPSL(Routing Policy Specification Language)が定義されると、Merit を初めとするいくつかの IRR サーバ提供者は、すぐさまこれに対応し、現在では、RPSL を利用することが標準となっているのである。

一方で、IRR サービスの状況も変化してきている。RA プロジェクトの発足に合わせる形で、NAP や ISP などでも IRR サービスを提供する動きがあった。1999 年までは、IRR サーバソフトウェアも十分なものがなく、実際にサービスを提供していたのは、RADB、RIPE、Cable & Wireless（旧 MCI）、ANS、CAnet の 5 か所のみだった。この中でも RADB はデータ量などから見ても世界最大のデータベースとなっていた。しかし、Merit は 1999 年後半に、RADB を有料化すると宣言し、同時に、IRRd という IRR サーバソフトウェアの提供を無償で開始すると発表したのである。

現在の IRR のホームページを図 7-33 に示す。

[18] Routing Policy Specification Language (RPSL) (RFC2622)
http://www.ietf.org/rfc/rfc2622.txt

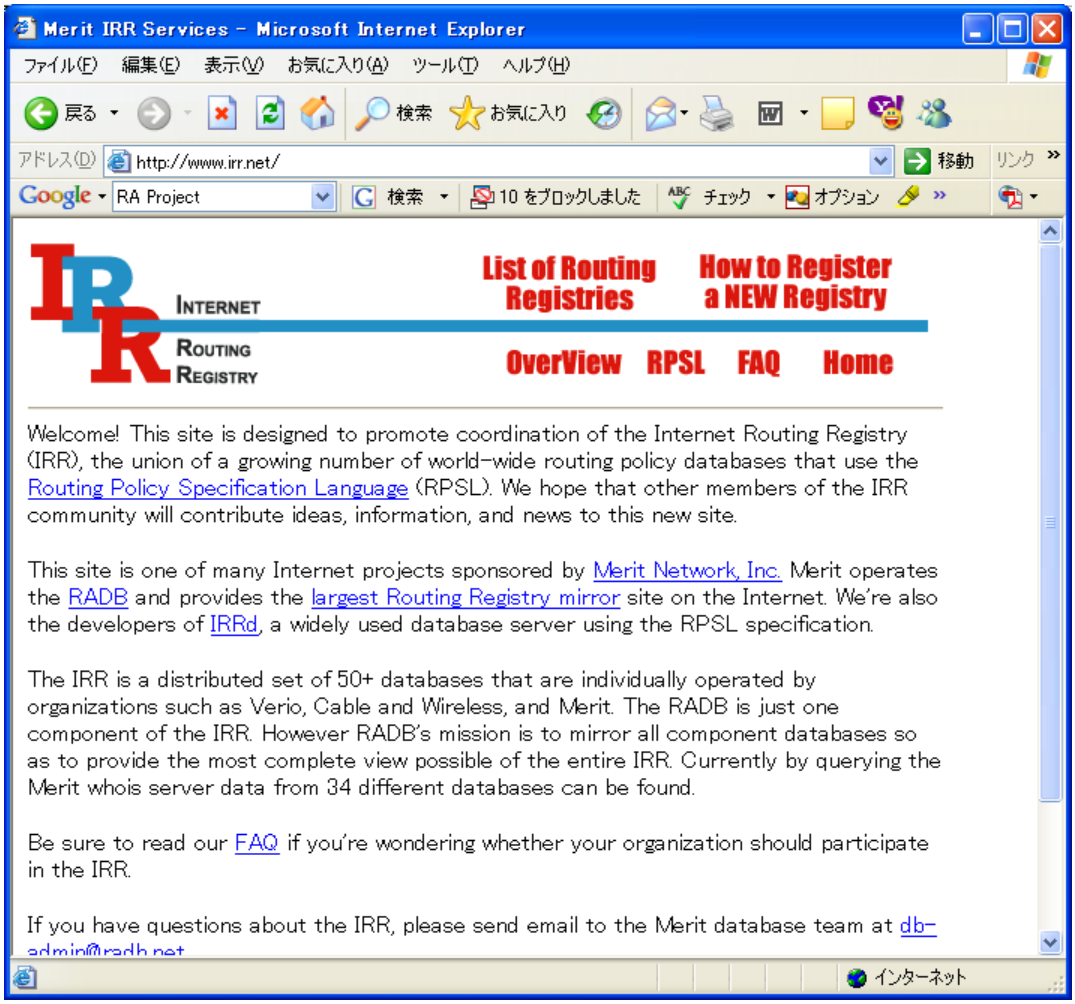


図 7-33 IRR のホームページ

IRRd は、IRR の情報を簡単に扱うことができ、かつ他の IRR とのミラーリングもサポートし、小規模向けには非常に強力なソフトウェアであった。このため、IRR を必要とする ISP などは、IRRd を用いて IRR サービスを独自に行い、そのデータを RADB とミラーするという手法を取り始めたのである。これにより、先にあげた 5 か所の IRR に多く集まっていたデータは徐々に分散化していったのである。

IRR のミラーリングの仕組みでは、ミラーリング先から他のミラーリング先へとデータを転送することは行わないため、今までのように RADB をミラーし、そのデータベースからルータへポリシを実装することが難しくなった。

この段階において、日本でも IRR の周辺事情に関する議論が盛り上がり、JPNIC が 2000 年に開催した IRR 研究会へとつながったのである。

現在、この活動は JPNIC が提供する IRR サービスである JPIRR に結びつき、試験サービスを行うに至った。JPIRR のホームページを図 7-34 に示す。

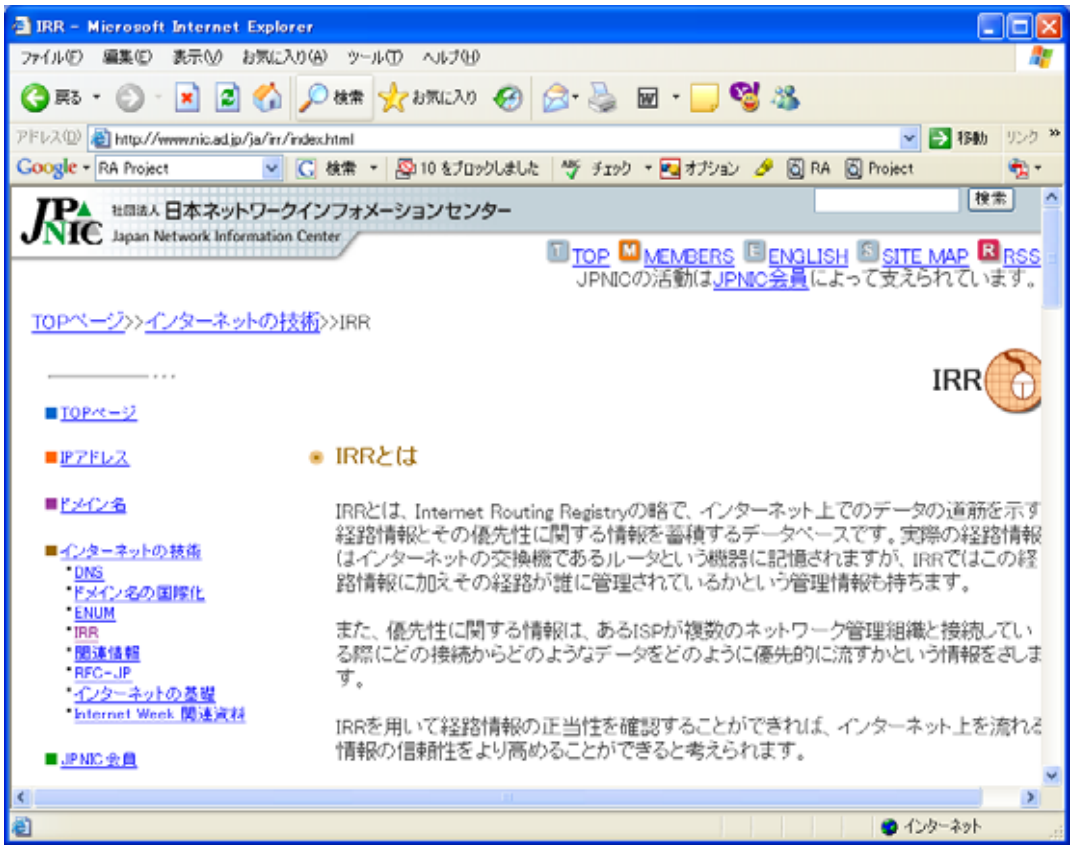


図 7-34 JPIRR のホームページ

7.6.2. 国際的な IRR の利用状況

国際的には、IRR のコミュニティは現在も Merit が運営する RADB を中心と考えて問題ない。Merit の有料化以降、Merit が配布する IRRd を用いて独自に IRR サーバを運営する ISP が数多くいるが、その多くは、RADB とミラーリングを実施しており、その数は、38[19]を越える。このため、RADB から直接検索を実施するのであれば、現在もなお、多くの IRR の情報を取得することが可能である。

[19] 2006 年 1 月 26 日時点で RADB のホームページ(http://www.radb.net/reports.html)にて公開されているミラーリングリストより。ただし、RIPE などが入っていないことから実際にも多数の IRR がミラーされていると思われる。

RADB のホームページを図 7-35 に示す。

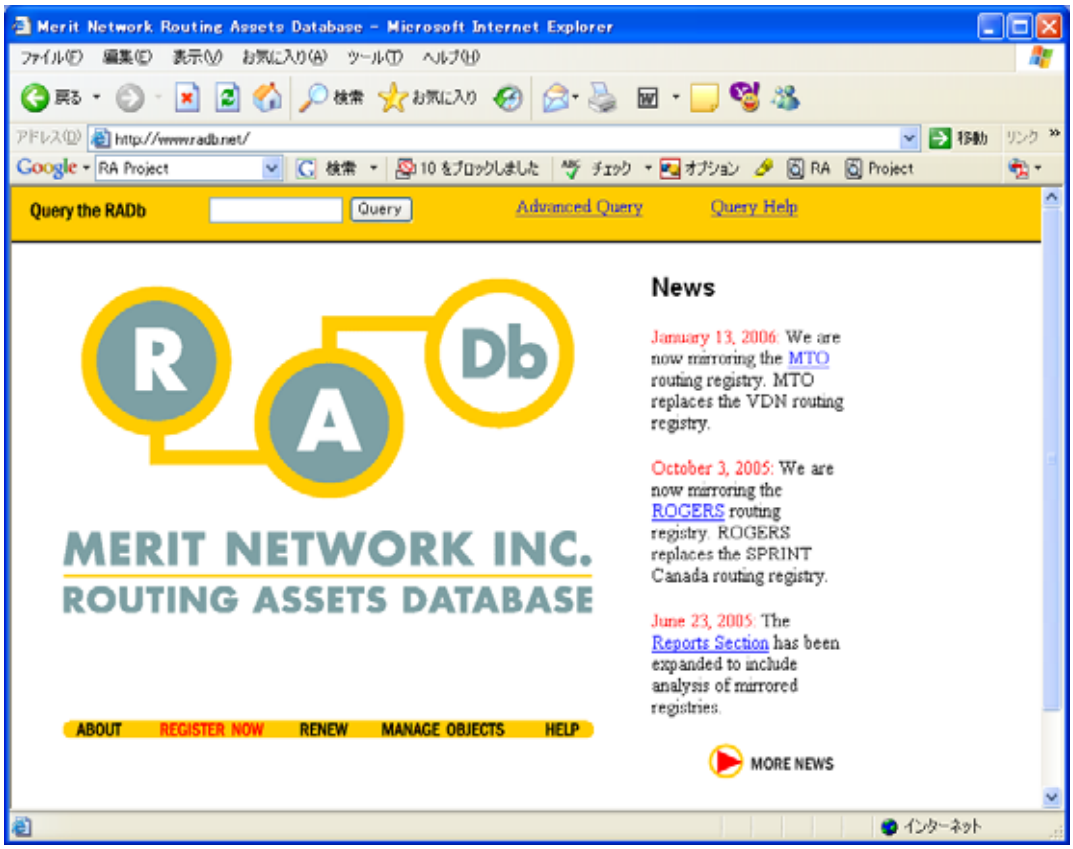


図 7-35 RADB のホームページ

一方で、インターネットレジストリで IRR の機能をサポートする動きもある。この動きは、米国のネットワーク運用グループである NANOG が開催するミーティングである NANOG22(2001 年 5 月開催)において、JPNIC IRR 企画策定専門家チームのメンバーが IRR とインターネットレジストリが持つ Whois データベースの連携によって IRR のデータがより信頼性の高いものになるという提案を行い、それに賛同する形で始まったものである。

この流れを受け、アジア太平洋地域を中心に活動する APNIC は、2002 年より RIPE NCC の Whois システムを採用し、同時に IRR のサービスも始めている。

ヨーロッパを中心に活動を行う RIPE NCC は、独自のデータベースを開発し利用している経緯もあり、古くから、IRR とインターネットレジストリが持つ Whois データベースが統合されたシステムで運用が行われてきたのである。

北米を中心に活動する ARIN(American Registry for Internet Numbers)では、Merit が大きな IRR である RADB を運営していることから、実験的に IRR を運営するにとどまり、積極的な運用は行っていない。

このように、IRR の共通のデータは、RIR(Regional Internet Registry:地域インターネットレジストリ)を中心にデータの統合化が進んできている[20]。しかし、各 ISP、特にトランジットサービスを行うような比較的大規模な ISP においては、現在もなお、独自に IRR サービスを実施しているところも少なくない。

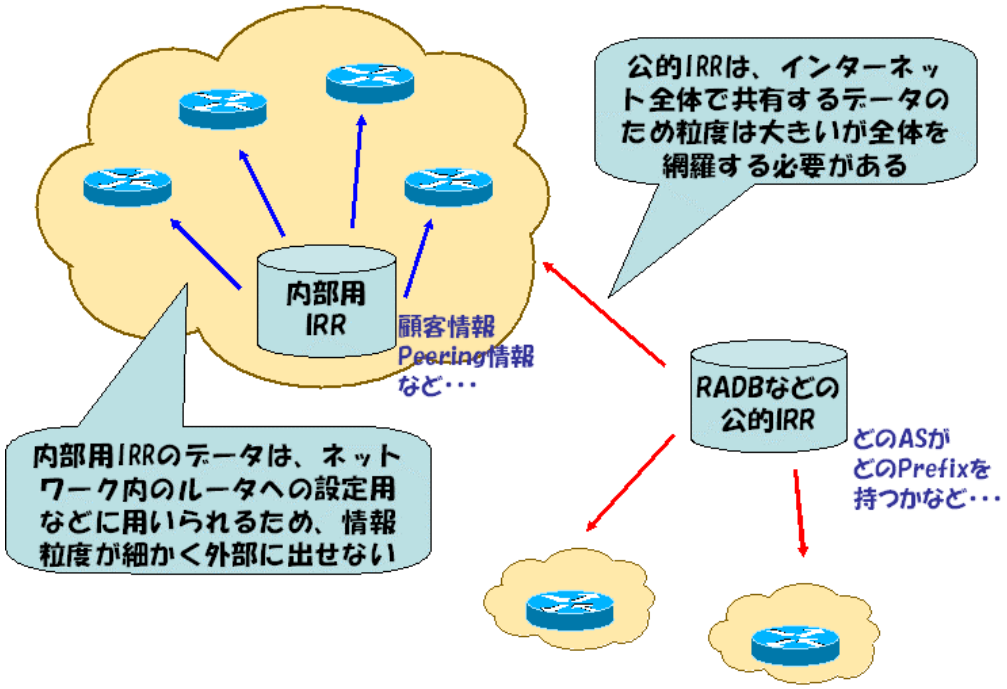


図 7-36 IRR の運用形態によって情報の特性が異なる

この 2 つの運用のされ方には、データの内容に大きな違いがあることがわかってきている。この様子を図 7-36 に示す。

先に説明した、RIR が運用する IRR データについては、より一般的な情報、つまり、AS 番号と Prefix、そしてそれらがグループ化された情報と運用者情報などが入り、経路の優先性情報や Peering 情報などの細かい情報はほとんど含まれていない。その一方で、各 ISP が運用するような IRR では、顧客の Peering 情報が扱われているケースも見受けられる。

[20] 北米では、いまもって RADB が中心的に利用されている。

このように ISP が独自に IRR を運用する背景には、BGP ルータの Prefix フィルタを自動生成するなど運用負荷の軽減や自 AS が管理すべき経路情報の安全な蓄積などという目的があるのである。

RIR が運用するような IRR でも、ごく一般的な情報はあるため Prefix フィルタ程度であれば自動生成が可能である。しかし、ISP が独自に運用する IRR などでは、これに加え、マルチホーム Peering の優先性情報などの細かい情報も扱うことで、より高度な運用負荷軽減を狙っているのが現実と言える。

7.6.3. アジア太平洋地域における IRR の利用状況

アジア太平洋地域における IRR の活動は APNIC を中心に行われている。先の第 22 回 NANOG ミーティングの動きや第 10 回 APNIC オープンポリシーミーティングでの JPNIC IRR 企画策定専門家チームからのインターネットレジストリによる IRR 実施の提案などを受けて、APNIC では 2001 年より IRR サービスを開始したのである。

この動きを経て、APNIC 管轄地域では、APNIC の IRR データベースが利用されるケースが増えてきたのである。

ISP 独自での IRR サービスの運用については、国際的な流れと変わることはないが、全体の数として APNIC 管轄地域では、大規模な ISP も少ないことからその数は多くない。

APNIC 管轄地域において比較的大きな動きがある地域は日本と韓国と言えるだろう。韓国は、先の APNIC の流れを受け、KRNIC(Korea Network Information Center)が中心に IRR サービスの検討プロジェクトが立ち上がっており、サービス開始に向けた検討が進んでいる。日本では、公共の IRR サービスとして現在 JPNIC が実施している JPIRR 試験サービスがあるが、民間・研究団体としては NTT や SINET なども IRR を運用していることがわかっている。

これら APNIC 管轄地域の IRR については、IRR の普及度などの面から、公共の IRR として運用されているものは APNIC の持つ IRR のみであり、JPNIC および KRNIC はいまだ実験段階、その他は ISP 独自となっている状況である。

7.6.4. IRR に期待される機能

ここまで IRR の歴史と世界・アジア地域での議論の状況について報告してきた。これらの

議論の中で、当初 IRR が特定の範囲、特に IX 接続業者間などの限られた範囲内でのみ利用されることを前提として進んできた物が、昨今では実際にインターネット上を流れてくる経路情報を確かめる手段としての価値が高まってきていることが解る。そこで本節では、これらの議論の中での IRR に登録される情報の種類と、それら情報がインターネット上に流れる経路情報の確認手段としての役割について、さらに詳しく報告する。

7.6.4.1. IRR で取り扱うデータの種別

IRR の利用形態はその利用者によって様々である。ある IRR では、特定のインターネットエクスチェンジポイント(IX)の経路交換ポリシを表現するケースや、特定の ISP の顧客の経路情報を管理するケースなどに使われていることが知られている。一方で、IRR のデータはインターネット全体の経路情報が蓄積されていると考えることができ、どの ISP がどのような Prefix や AS の塊で、経路をインターネット上に広告しているのかという、非常に一般的な情報の参照元としても利用されている。

これら 2 つの利用方法は、IRR に登録されるデータの粒度や参照元などが大きく異なるため、同じポリシでデータを取り扱うことは、Peering 情報などの守秘義務もあり、現実的ではない。JPNIC IRR 企画策定専門家チームが主催した会合などでも同様な議論が行われたほか、APNIC や RIPE NCC が主催するミーティングの経路制御関連の専門ワーキンググループ（Routing-WG）の議論でも同じような意見が寄せられている。

これらを整理すると、特定の部分で利用するデータとインターネット全体で利用するデータという 2 つの性質の異なる IRR データをプライベートデータとパブリックデータに分類することができる。

以下では、プライベートデータとパブリックデータの定義について明示する。

◯プライベートデータ

主に、特定の IX への参加者や特定の ISP の顧客向けに提供され、そのコミュニティに閉じて利用される IRR に登録されるデータを指す。これらのプライベートデータには、特定のコミュニティ内で閉じることが必要とされている情報が含まれていることが考えられ、他の IRR との情報の交換には、守秘義務協定など一定の制限を設けて実施するか、または一切の情報交換を実施しないなどの措置が必要である。

具体的には、特定の BGP Peer に対する Local Preference の値、AS-PATH や Prefix

の情報と、それらをどの BGP Peer にどういう優先度で広告するかという情報、およびどういう優先度で受信するかという情報が含まれていることが考えられる。これらの情報を使うことで、特定のコミュニティにおいてきめの細かい経路フィルタなどを自動的に生成することが可能になりうる。

○パブリックデータ

不特定多数からの参照を前提とし、インターネットの経路情報全体を参照できることを目的に運用される IRR に登録されるデータを指す。登録される情報は、インターネットレジストリから割り振られ、実際にインターネットに広告されている経路が登録されるもので、BGP 運用者がインターネットに対して広告する可能性のある経路とその Origin AS の情報などが登録されるほか、トラブル発生時のコンタクト先に関する情報の蓄積などに利用されることが期待されている。

パブリックデータは、インターネットに広告されている全ての経路情報が参照可能であるため、広告されている経路が正しいものであるかどうかを確認することができる点と、プライベートデータのような細かい優先性情報などは、付加的に登録可能、つまりオプションとしての情報であるという点が、重要なポイントとなっている。

もちろん、全ての経路情報に関連するデータが蓄積されるべきものであるため、フルルートに対する適切な Prefix フィルタが実施できるレベルの情報である必要がある。

これら 2 つの情報は、完全に分離して管理されるべきである。プライベートデータは、特定のコミュニティによって管理され、パブリックデータはインターネットレジストリのような公共的な機関によって管理されるべきである。しかし、これら 2 つのデータは時に相互に補完して形成されるべきである。例えば、特定のコミュニティにおいて蓄積されるプライベートデータは、そのコミュニティ内部で利用される情報のみが蓄積されたからである。それ以外の情報はパブリックデータとして蓄積されているデータによって補完されることで、必要な部分は細かく、それ以外の部分については粗く、全ての情報が参照できるようになる。

逆に、プライベートデータから細かい優先性情報などの守秘義務に抵触するような情報をそぎ落とし、パブリックデータに加工することで、パブリックデータを蓄積する IRR に流用することが可能になる。

現時点での IRR の仕組みには、このようなプライベートデータとパブリックデータの分類は存在していない。しかし、将来的に IRR の相互連携モデルが確立した時には、IRR の各オブ

ジェクト、さらには RPSL の各フィールド単位で、パブリック・プライベートなどの属性をもつことで、より信憑性の高い IRR 環境の構築が可能となるだろう。

7.6.4.2. 経路の台帳としての IRR

IRR で取り扱うデータは大きく分類して、「パブリックデータ」と「プライベートデータ」の 2 種類があることを、前節までに説明した。

プライベートデータには、経路の優先性情報などインターネットの運用上、重要な情報が含まれている。このような情報は、企業内や特定のコミュニティの中だけで利用され、外部に流出させないように運用される。その一方、パブリックデータは、プライベートデータをより一般的にし、特定のコミュニティ以外で参照可能な情報となる。

インターネットの運用上、優先性情報といった、より細かい情報を得ることで、運用の煩雑さが改善されることが期待できると考えられている。公共的な団体で実施する IRR では、特定のコミュニティに依存するプライベートデータではなく、広くデータを参照できるパブリックデータが求められている。このため、公的な IRR のサービスを実施する際の判断基準として、パブリックデータをインターネットの運用にどのように生かせるかという点が、一つのポイントとなると考えられる。

これまでの IRR の歴史の中で、この「パブリック」な IRR に近いものが RADB と言えるだろう。IRR が分散化する前の段階における RADB の利用方法を考察することで、パブリックデータの価値が見えてくることが考えられる。

IRR が分散化する前の段階では、RADB は CAnet や MCI をはじめとする他の IRR と比べて、非常に多くのデータを蓄積していた。この段階で、RADB のユーザは主に以下のような利用を行っていた。

1) 経路情報の Prefix や AS-PATH を利用した、フィルタリングの基礎情報の生成
2) 特定の Prefix に対するコンタクト情報の検索
3) 不審経路の確認

当時は、RADB と MCI の IRR を参照することでインターネットのほとんどの経路が参照できたことから、日本の ISP も RADB に自 AS の経路情報を登録しなければ、インターネット上で経路がフィルタされてしまうという懸念があった。

また、ISP の顧客が BGP によって接続する場合には、ISP が RADB への登録を代行してきたため、実質的に RADB を参照すればインターネットの経路情報はほとんど解決可能だと思われていたのである[21]。

もちろんこの段階では、米国の大手の ISP などに、不用意に Prefix や AS のフィルタがされてしまうことを避けるために登録されていたため、経路の優先性情報などの細かい情報はほとんど含まれていなかったのである。

このような環境が功を奏して、インターネットに接続されているネットワークの Prefix や AS に関する情報、コンタクト情報が容易に取得可能になっていたのである。また、このようなインターネットの経路情報が十分集まっているデータベースは、実際にルータで受けている経路の確認のためにも使われていた。

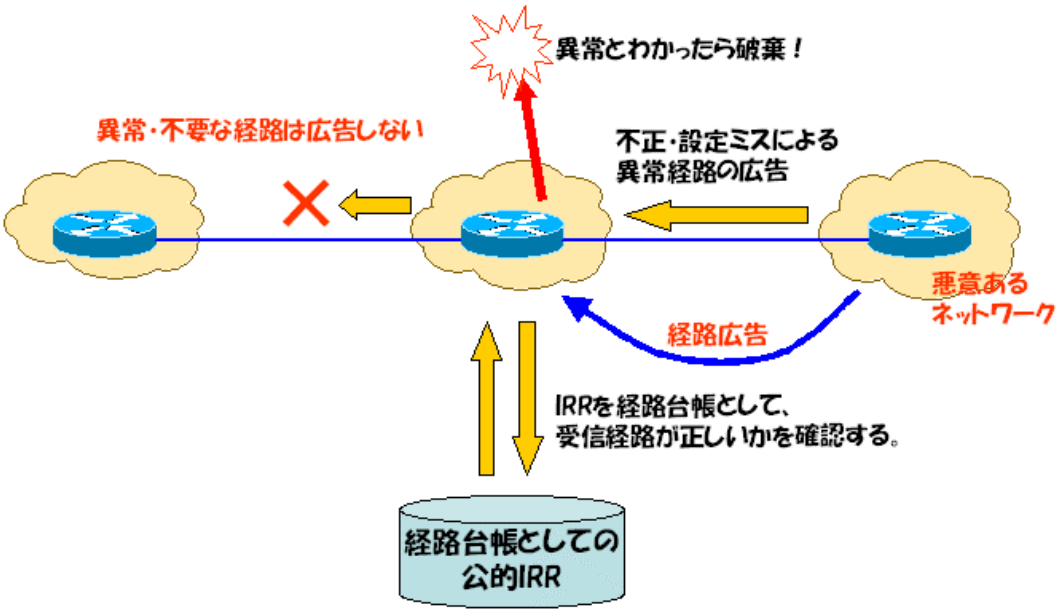


図 7-37　経路台帳として IRR の役割の例

インターネットのルータの運用は、一部自動化されている部分もあるが、多くの場合、人手によって行われている。これは同時にミスを引き起こす原因でもあり、ミスオペレーションなどによって、時として他の AS の Prefix や自分が割り当てられていない AS 番号の経路などをインターネットに流してしまうということがある。もちろん、このほかにも、悪意をもって他の ISP

21 実際は他の IRR も合わせて参照しなければ解決は不可能であった。

の経路を流し、インターネットを混乱させるということもある。

このような場合、RADB はその膨大なデータ量から、その経路は意図されてインターネットに広告されているのかを確認をするための台帳としての役割も果たしたのである。

このような IRR の利用例を図 7-37 に示す。

現時点での IRR のデータ、特に RADB などのパブリックデータは、IRR の分散化や長期間に渡ってデータが未更新となることなどによって、以前ほどの効果は得られていない。しかし、パブリックデータとして、インターネットの経路情報の台帳となるほどのデータが参照できるようになれば、運用上、非常に有用なデータベースとなることには間違いないであろう。

7.6.5. IRR の稼働実態

現在、世界中で稼働している IRR は大量に存在している。そのほとんどは ISP 内部の私的利用のものである。一方で、公的に利用されている IRR も複数ある。

以下に、私的･公的を問わず世界的に有名な IRR について列挙し、その特徴について解説する。

- RADB

  米国の Merit によって運営されている IRR である。7.6.1 項でも触れたように、RA プロジェクトの実験のために作られた IRR であるが、その後も世界的に利用され、現在では世界最多のオブジェクトが登録されている。

  また、Merit では、IRRd という簡単に IRR サーバを構築できるソフトウェアを配布しており、これを利用して多くの ISP が独自の IRR を構築し、さらにそれらは、RADB とミラーを行っているため、数多くの IRR とミラーを行っている。

  このため、RADB は IRR の代名詞として利用される場合が多く、IRR を用いて経路フィルタを実施する場合でも、現状では RADB に登録することが標準になっているのが現実である。

- RIPE-DB

RIPE によって運営されている IRR である。IRR のソフトウェアとして RIPE-DB という RIPE 独自に開発した IRR ソフトウェアを利用している。この IRR ソフトウェアは、RIPE がインターネットレジストリの WHOIS データベース用に開発したものを IRR もサポートするように拡張したもので、現在でも WHOIS データベースと IRR データベースが共存する形で利用されている。

また、RIPE-DB は IRR に経路情報広告ポリシを記述する言語である RPSL の原点ともなる RIPR-181 をいち早くサポートした IRR でもあり、Merit が開発した IRR ソフトウェアと同じ程度の歴史を持つ IRR といえる。

RIPE-DB は、レジストリが運営する IRR としては最初の IRR で、レジストリが利用する WHOIS データベース、そして IRR データベースの基本ソフトウェアとして広く利用されている。

- APNIC

APNIC によって運営されている IRR である。APNIC は、WHOIS データベースとして RIPE-DB を利用している。このため、IRR をサポートすることは、単純に IRR の機能を有効にするだけで利用できるため、APNIC ミーティングにて、APNIC による IRR の運営がリクエストされてから、IRR のサービスを始めるまでにそう時間がかからなかった。むしろ APNIC では、IRR で参照可能なデータベースを広範囲に広げるために、数多くの IRR とミラーを行うような努力を行い、その結果、レジストリが運用する公的な IRR としては、RADB に次いで多くのオブジェクトの検索が可能となっている。

このため、APNIC は現在、アジア太平洋地域の公的 IRR の中心的存在として運営されている。

- Verio

NTT/Verio が運営する IRR である。主に、顧客の経路情報を登録し NTT/Verio との接続ポイントで不正な経路が混入しないようにフィルタをしたり、確認したりするための参照先データベースとして構築された、私的な IRR である。

NTT/Verio では、この独自の IRR だけでなく、RADB ともミラーを行っており、RADB と NTT/Verio に登録されたデータを中心に経路のフィルタを作成し運用を行っている。

- SINET

  SINET[22]は、日本の学術情報ネットワークで、主に大学などに対してインターネットの接続サービスを行っている組織である。ただし、公式なサービスとしてIRRサービスは提供していないようである。

- JPIRR

  JPNIC が実験的に運営している IRR である。インターネットレジストリが運営する IRR によって IRR に登録する情報内容の正確性の向上と正しい情報が登録されることによって経路情報を確認するための台帳として IRR が機能出来る可能性について研究･開発を行っている。

  JPIRR では、主に JPNIC の IP アドレス管理指定事業者に対してサービスを行っているが、広く IRR の価値を理解してもらうことを目的に、現在のところ IP アドレス管理指定事業者に限らずサービスを提供している。

  2006 年中には、正式サービスへの展開を検討しており、より安定したサービスを展開予定である。

7.6.6. IRR 利用の実態

これら IRR に登録された情報は、インターネットに流れる経路の台帳として利用されることが期待されている。具体的な利用方法としては、IRR に登録されている情報を元にルータに設定する経路情報を生成したり、BGP で受信している経路を確認する手段として利用したりする。このほか、新しい利用方法としてルータの経路制御ソフトウェアに受信経路を IRR で確認する機構を実装し、経路情報の受信とほぼ同時に経路情報の真偽を確認するようなことも考えられている。

以下に、これらの IRR の利用手法の代表的なものと現在考えられている IRR の利用手法について解説する。

---

22 学術情報ネットワークは、国立情報研究所の運営するネットワークである。（http://www.sinet.ad.jp/）

(1) IRR を用いた経路フィルタの生成

IRR には、AS がインターネットに流す経路情報が登録されている。そこで、この情報を利用して自 AS が受信する経路情報をフィルタして、不正な経路情報を受信することを防ぐことができる。

フィルタの生成ツールとして IRRToolSet (図 7-38) が有名である。IRRToolSet は ISI[23] が開発した IRR を検索してその情報からルータに設定するフィルタなどを生成するツールである。フィルタを生成するだけでなく、IRR を利用する上で必要な文法チェックツールなどのパッケージである。

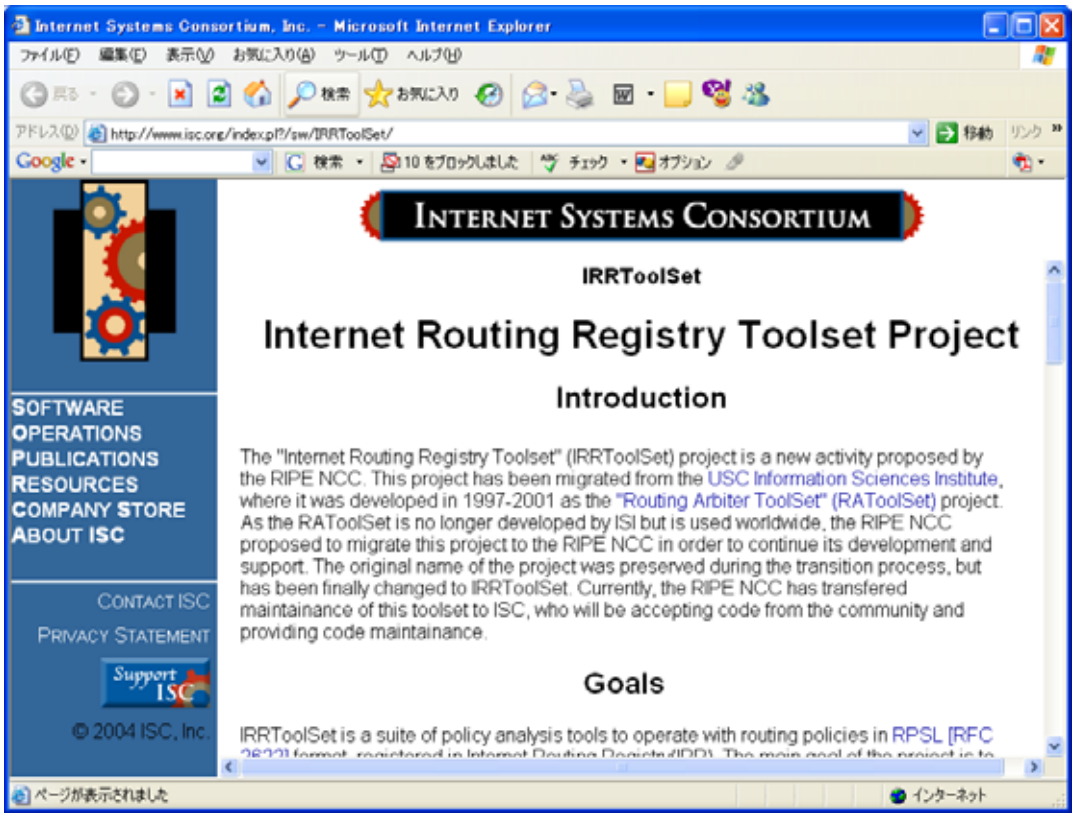


図 7-38 IRRToolSet のホームページ

一般的には、自 AS が接続する AS の情報から、その接続先 AS が接続相手に広告する情報を検索し、その情報から経路フィルタを生成する。この IRR を利用した機構が

[23] The University of Southern California Information Sciences Institute

無い場合に経路フィルタを生成するときには、接続先相手と広告する経路をメールなどで交換し、それをフィルタとして加工し利用することができる。この場合、接続相手の広告経路に追加･削除･変更が生じた場合に、その都度その更新を相手に通知し、フィルタを変更してもらう必要があり、比較的大きな AS の場合、接続 AS も多く、フィルタ更新作業だけでも相当な作業量となる可能性がある。IRR を用いてこれらの経路情報を交換することで、これらの情報のやりとりを簡素化できるだけでなく、フィルタの生成作業も簡素化することが出来るようになることが期待できる。

(2) IRR を用いた自動的な経路情報確認機構

先にも述べたように IRR には、ある AS が広告しようとしている経路情報が登録されている。この情報を利用すれば、経路を受信した時点で、その AS が意図的に広告しているかどうかを判断することができる。この特性を利用し、確認する機構をルータに実装した研究がある。

この研究は、2001 年に東京大学の長橋賢吾氏が「An Integrity Check for the Conflict Origin AS Prefixes in the Inter-domain Routing」という論文にまとめたもので、受信した経路を IRR に問い合わせ、その経路が登録されており、広告元 AS が正しいかなどのチェックを行う機構を経路制御ソフトウェア上に実装し、経路の受信段階で受信経路が広告元 AS によって意図的に広告されているかについて IRR を通して確認するための機構である。

論文発表段階では、IRR に登録されている情報がインターネット上に広告されている経路情報がすべて登録されていないことや、IRR に登録されている情報が正しくないなどの問題があり、有効性に疑問が呈されたが、正しい情報が登録されている IRR ができれば、経路情報のフィルタをダイナミックに行うことが出来る一つの解決策と言えるだろう。

7.6.7. IRR がもたらす影響

前節までに IRR の情報を元にフィルタを生成する手法などについて述べてきた。実際にこれらの手法を利用してフィルタを生成しているという確たる説明はどこの ISP も行っていない。

しかし、実際に IRR から経路情報を削除してしまった場合のトラブル、そして、複数の IRR が乱立し、連携が不十分な状態であるがために、経路制御上の問題が発生しているケース

がいくつか報告されている。

ある例では、AS に関連する情報を RADB に登録していたが、何らかの原因で RADB からオブジェクトが削除された。これにより、上流 AS が接続している他の AS の入り口で当該 AS の経路情報がフィルタで排除され、インターネットの広範囲にわたって接続性を失うトラブルが発生したのである。

そこで接続性の回復の手段として、当該 AS は日本の実験サービスである JPIRR にオブジェクトの登録をおこなった。しかし、これでも状況は改善できなかった。

これは、フィルタをしている AS が IRR オブジェクトの登録先 IRR が RADB または該当 AS が運営する私的 IRR のいずれかに登録しなくてはならないという制限があったためである。

非常に似た例が、JANOG16[24]において、NPO 法人北海道地域ネットワーク協議会の河合氏によって紹介されている。

IRR へのオブジェクト登録には、しばしばこのような誤解が生じている。IRR は、現在多く立ち上がっているが、それぞれの IRR に自身の IRR データベースへの登録を示すために、管理元の IRR の名称がすべてのオブジェクトに「Source」として記されている。IRR のミラーはこの「Source」毎に行われるため、「RADB にオブジェクトを登録」といった場合は、この「Source」が RADB になっていることが求められており、whois.radb.net を検索ホストとして検索が可能かどうかは問題ではないのである。図 7-39 にこの様子を示す。

[24] 日本ネットワーク・オペレーターズ・グループによる第 16 回目のミーティングで、2005 年 7 月 28 日～29 日に福岡で開催された。(http://www.janog.gr.jp/meeting/janog16/)

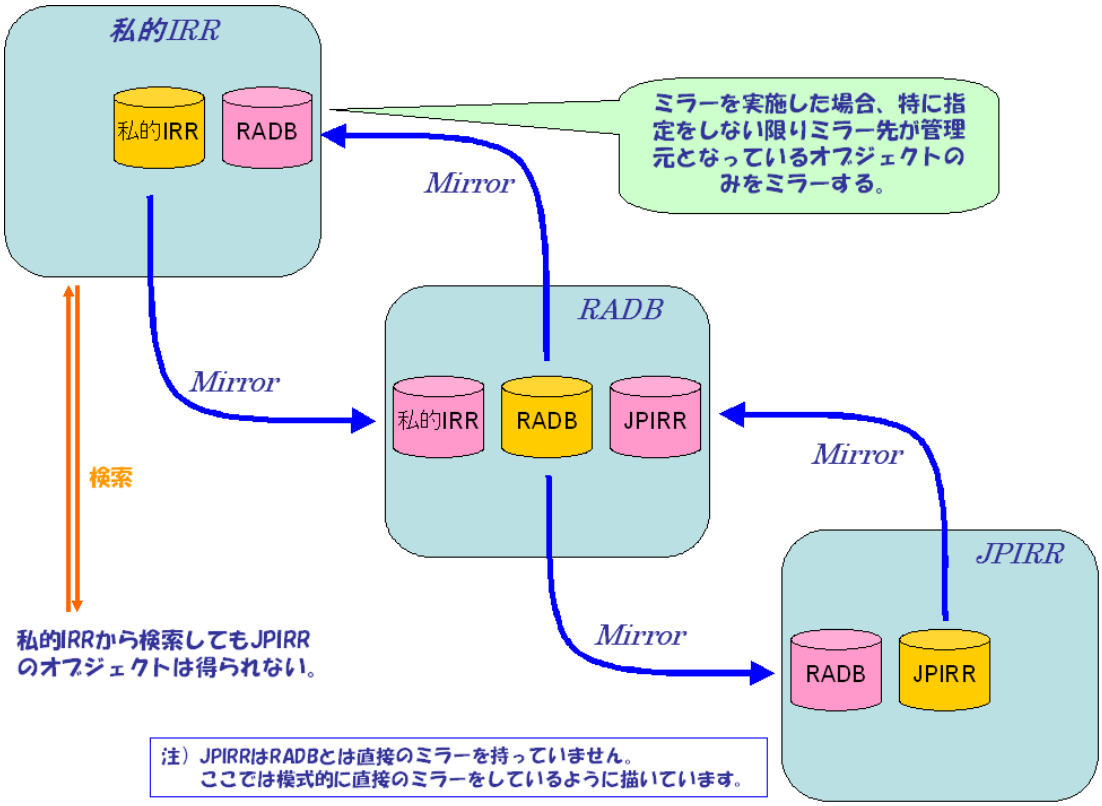


図 7-39 IRR のミラー構造

つまり、今回フィルタしていた AS は、自身の私的 IRR と RADB をミラーしており、私的 IRR で検索可能なオブジェクトを用いてフィルタを生成していたと考えられる。このような状況では、RADB で JPIRR のオブジェクトが検索できたとしても私的 IRR で検索ができないため問題が解決できないのである。

この問題を解決するためには、IRR システム全体としてミラーリングの手法などについて根本的に検討し直さなければならない。

最後に、この節で取り上げた問題を整理する。

(1) IRR への経路登録が無い場合、自 AS とは直接関係の無い AS において、経路がフィルタで排除される可能性があり、潜在的にインターネット全体に対して自 AS が広告した経路が正しく伝わらない可能性がある。

(2) RADB からオブジェクトの検索ができれば問題が解決するわけではなく、ミラーリングシステムを根本的に見直す必要がある。

7.7. IP レジストリシステム

本節では、JPNIC や APNIC などの IP レジストリ[25]が利用する IP レジストリシステムに関して解説を行う。

最初に、IP レジストリが運営する IP レジストリステムにはどのような役割があるのかを解説する。次に、IP レジストリシステムに登録される情報の特徴について解説する。

7.7.1. IP レジストリシステムの役割

IP レジストリは、7.4 節で解説したとおり IANA を頂点として、5つの地域レジストリを中心として、世界的に共通の IP アドレスの空間を管理し、IP アドレスを必要とする組織などに適切に割り当て、その一意性を保証する組織である。

IP レジストリシステムは、これらの IP レジストリが果たすべき役割を補助し、効率よく IP アドレスを管理するためのシステムである。

通常 IP レジストリといった場合、IP アドレスの管理のための階層構造に従い、ネットワーク接続事業などを行うローカルレジストリ（LIR）までを含むが、ここでは、この LIR を含まず、国別レジストリ、または地域レジストリにおける IP レジストリシステムに絞って解説を進めることとする。

この IP レジストリシステムには、IP アドレスを管理し、インターネットの世界を効率よく運用できることをサポートするために、以下の様な役割が実装されている。

(1) 管理すべきアドレス空間を把握する
(2) 利用中のアドレス空間を把握する
(3) アドレスを必要な人に必要なだけ効率よく割り当てる
(4) インターネット運用上必要な情報を公開する
(5) 公開すべき情報を最新に維持する

実際には、IP レジストリの職員が自ら作業するところも多くあるが、IP レジストリシステムとし

[25] ここでは IP レジストリと表記するが、前節までのインターネットレジストリと同義である。

て以上のような役割がある。以下に個々の役割について解説する。

(1) 管理すべきアドレス空間を把握する

IP レジストリが管理するアドレス空間は、IANA を頂点に管理するアドレス空間を必要に応じて地域レジストリに割り振りを行っている。また、APNIC 地域では、さらに国別レジストリが存在し、管理するアドレス空間はさらに国別レジストリへとその管理を委託している。

これら IP レジストリによるアドレス空間の管理では、IP アドレスの一意性を保つために適切に管理する必要があり、各レジストリがどのアドレス空間の管理が委譲されているのかをしっかり把握して管理しなくてはならないのである。

IP レジストリでは、このようなレジストリで管理しているアドレス空間を把握するような実装が必要なのである。

(2) 利用中のアドレス空間を把握する

IP レジストリでは、上記によって委譲されたアドレス空間をアドレスの利用者に対して、割り振り・割り当てを行う。通常、利用者をローカルレジストリとエンドユーザーに分けて考え、ローカルレジストリに対しては、ローカルレジストリがさらにエンドユーザーにアドレスの割り当てを行うことから、アドレスの「割り振り」と呼ぶ。

地域・国別レジストリでは、これら LIR やエンドユーザーに割り振り・割り当てたアドレス空間を管理し、二重に割り振り・割り当てを行わないようにしなくてはならない。

IP レジストリシステムでは、アドレス空間が二重に割り振り・割り当てを行わないような管理を行うための実装が必要なのである。

(3) アドレスを必要な人に必要なだけ効率よく割り当てる

IP アドレスは、インターネットを利用するための公共の資源として考えられる。このため、IP アドレスは、必要とする人に必要なだけ適切に割り当てる必要がある。適切な割り当てについては、先にも解説したように RFC2050 や各レジストリが定めているアドレスの割り振り・割り当てポリシに従って行われなくてはならない。

この割り振り・割り当てには、この「適切さ」を諮る仕組みがあり、IP アドレスの割り振り・割り当てルールが作成されており、このルールに従って実際お割り振り・割り当てが行われることになる。

通常、この割り振り・割り当ては、「申請」→「審査」→「割り振り・割り当て」の流れがあり、この申請の受理や、審査、審査結果の管理を補助するためのシステムが IP レジストリシステムに求められる。

(4) インターネット運用上必要な情報を公開する

RFC2050 でも記載されているとおり、割り振り・割り当てが行われたアドレス空間は、公開される必要がある。

広大なインターネットにおいて、あるネットワークが他のネットワークまでの到達性を確保するためには、複数のネットワークを跨いで接続されるが、多くの場合、直接接続しているネットワーク以外のネットワークの情報は知ることができない。しかし、インターネットの運用においては、離れたネットワークからの経路広告やパケット送受信が自分のネットワークに何らかの影響を及ぼす可能性があり、場合によっては、当該ネットワークに連絡をする必要が発生する。

このように、インターネットの安定運用を目指す上では、自ネットワークと直接接続性を持たないネットワークのコンタクト情報などの情報を知る必要性があり、IP レジストリシステムには、割り振り・割り当ての情報を元に、ネットワークに対するコンタクト情報など運用上必要な情報を適切に公開する責務がある。

IP レジストリステムには、このような割り振り・割り当てを行ったアドレス空間に対する運用上必要な情報を公開するための実装が必要なのである。

(5) 公開すべき情報を最新に維持する

上記によって公開される情報は、運用上最新のものでなくては役に立たない。IP レジストリシステムには、公開情報を最新の情報に維持するための仕組みが必要なのである。

IP レジストリシステムに登録されている情報で且つ運用上必要な情報は、通常

ネットワークの運用者へのコンタクト情報などである。多くのレジストリシステムでは、これらのコンタクト情報は、ネットワーク利用者自身によって変更が可能である。しかし、ネットワーク利用者自身によって更新可能であるということは、更新にあたって、更新者がネットワーク利用者自身であるかどうかを判断する必要であることを意味する。

そこで、IP レジストリシステムでは、これら情報を更新するものが、更新者として適切であるかどうかを判断するための仕組みが必要となるのである。

7.7.2. IP レジストリシステムに登録される情報

7.7.1 節では、IP レジストリシステムに必要な機能について解説した。これらの機能は、先にも述べたように IP アドレスの割り振り・割り当てを適切に行うために必要な機能である。

本節では、IP アドレスの割り振り・割り当てを適切に行うために IP レジストリシステムに登録されるべき情報について整理する。

IP レジストリシステムに登録されるべき基本的な情報は以下の 3 つである。

(1) 上位レジストリからの割り振り情報
(2) 下位レジストリへの割り振り情報
(3) アドレスの割り当て情報

上位レジストリからの割り振り情報の場合、JPNIC の場合は APNIC からの割り振り情報、APNIC の場合は IANA からの割り振り情報となる。

下位レジストリへの割り振り情報は、JPNIC の場合は LIR への割り振りに関する情報となり、割り振り先のコンタクト情報やすでに割り振りを行ったアドレス空間の管理、そして、割り振った時に適切に割り振ったかどうかを示すネットワークに関する情報となる。また、LIR では割り振られた空間をエンドユーザーに割り当てる作業を実際に行うため、この割り当てが実際にどの程度行われているか、つまり、割り振った空間のなかでの割り当てた量を管理しなくてはならない。

アドレスの割り当て情報は、どのアドレス空間をどの利用者が利用しているかを登録する。

これらを整理したものを表 7-2 に示す。

表 7-2 IP レジストリへの登録情報

上位レジストリからの割り振り情報	● 割り振り元レジストリ情報 ● 割り振りアドレス情報 ➢ 開始アドレス ➢ 終了アドレス
下位レジストリへの割り振り情報	● 割り振り先レジストリ情報 ● 割り振りアドレス情報 ➢ 開始アドレス ➢ 終了アドレス ● 割り振り済み空間情報 ➢ 割り振り時のネットワーク情報 ● 割り振り済み空間の利用情報
アドレスの割り当て情報	● 割当先情報 ➢ コンタクト情報など ● 割り当てアドレス情報 ➢ 開始アドレス ➢ 終了アドレス