

いよいよ枯渇IPv4
～来るべきIPv6時代の実運用に備える～

株式会社まほろば工房
サイバー大学 准教授
代表取締役 近藤 邦昭

まほろば工房の実態(ごあいさつ)

■ 近藤邦昭とは

- ゲートボール発祥の地、北海道・芽室町生まれ(1970年)
- 神奈川工科大学 情報工学科 修了
- 独立系ソフトハウス勤務
 - 主に通信系のソフトウェア開発に従事
- DTIに入社
 - DTIの立ち上げ時に入社し、立ち上げ時期のネットワークを構築
 - JANOGの立ち上げ
- IIJに入社
 - NSPIXP(現DIXIE)に関する運用を中心に従事
 - NATSをIETFに提案(惨敗)
 - BGPViewの開発
 - JANOG会長就任、JPNIC IPアドレス検討委員長など
- Intec NetCoreに入社
 - ネットワーク運用ツールの開発に従事
 - その他、技術コンサルテーションなど
- 2006年5月個人事業として独立、まほろば工房設立
- 2007年3月まほろば工房の株式会社化
- 現在に至る。

近藤のIPv6関連活動

■ IPv6に対する近藤の基本理念

- 使って便利な所には、IPv6を使うべき
- IPv6の導入は、必要な時期にタイミングよく
 - 無駄なコストはなるべく出さないように
 - ビジネス的な優位性を出せるなら早めにやる
 - 事業継続性を鑑みて、遅れをとらないように
- 技術だけは、早めに理解しておくべき

こんなこと言っていると、近藤はIPv6は嫌いなんだと言われます。

■ そんな近藤の活動

- JPNIC番号資源利用状況調査研究専門家チーム チェア
 - 報告書「IPv4 アドレス枯渇に向けた提言」を提出
- JPNIC IPv4アドレス枯渇克服策検討ワーキンググループ メンバ
 - 報告書「IPv4アドレス在庫枯渇問題に関する検討報告書(第一次)」
- <http://www.nic.ad.jp/ja/ip/ipv4pool/> こちら辺にあります。

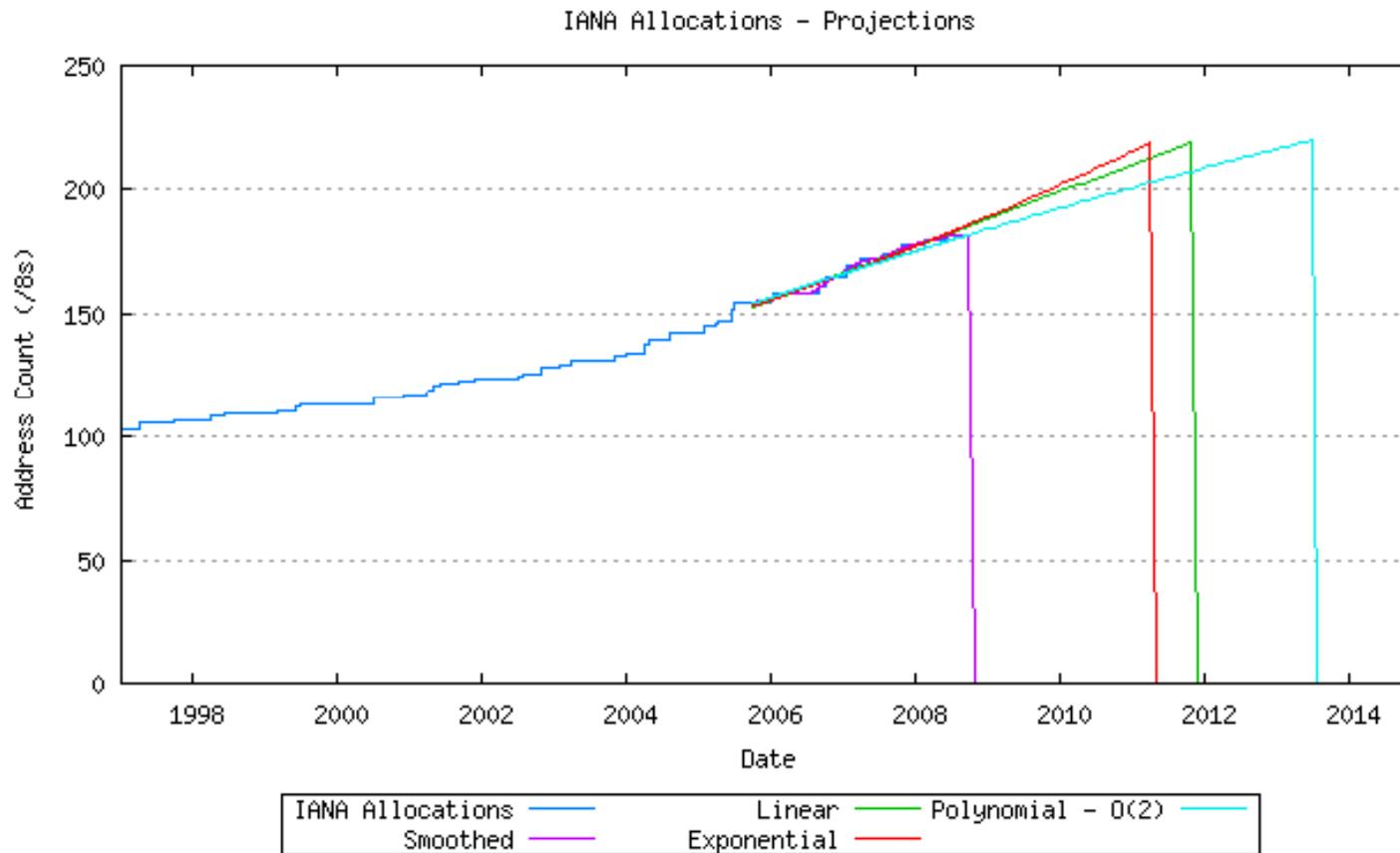
IPv4枯渇とIPv6に関する各種情報

- JPNIC
 - おそらく国内ではここが一番まとまっています。
 - <http://www.nic.ad.jp/ja/ip/ipv4pool/>
- IPv4アドレス枯渇対応タスク フォース
 - 2008年9月5日設立
 - インターネット/電気通信事業者、電気 通信機器ベンダー、その他関連する諸団体が連携
 - 国内13団体が協力
- APNIC/RIPE/ARIN/LACNIC/AfriNIC
 - RIRsでは、IPv4枯渇の状況を把握し、対策のためのポリシー策定などを行っている。
- 総務省
 - インターネットの円滑なIPv6移行に関する調査研究会
http://www.soumu.go.jp/joho_tsusin/policyreports/chousa/ipv6/
 - 「インターネットの円滑なIPv6移行に関する調査研究会」報告書(案)に関する意見募集
http://www.soumu.go.jp/s-news/2008/080411_3.html

APNICでのIPv4枯渇/IPv6の議論

- 2008年2月25日から29日に、台湾の台北で行われたAPNIC 25ミーティング
 - APRICOT/APOPSでのIPv4アドレス在庫枯渇関連報告
 - アドレスポリシー関連
 - [prop-050] IPv4アドレスの移管を認める提案
 - LIR間でアドレスの取引が行えるようにするもの
 - [prop-055] IANAからRIRへのグローバルアドレス分配の提案
 - IANAの在庫が/8x5個になった時点でRIRに1つずつ分配するもの
 - [prop-052] RIR間でIPv4アドレス在庫枯渇時期を調整する提案
 - アドレス在庫が少なくなった時に、RIR間でアドレスをやり取りして、枯渇時期を先延ばしにする提案
 - [prop-056] 在庫枯渇に向けたIPv4ソフトランディングの提案
 - アドレス枯渇の状況に応じてIPv4割り振り基準を厳しくしていく提案
 - [prop-058] LIR向け共有アドレスの新設
 - キャリアグレードNATに向けてプロバイダ用のプライベートアドレスを新設する提案
- なかなか、現場に影響のありそうな提案ばかりです。

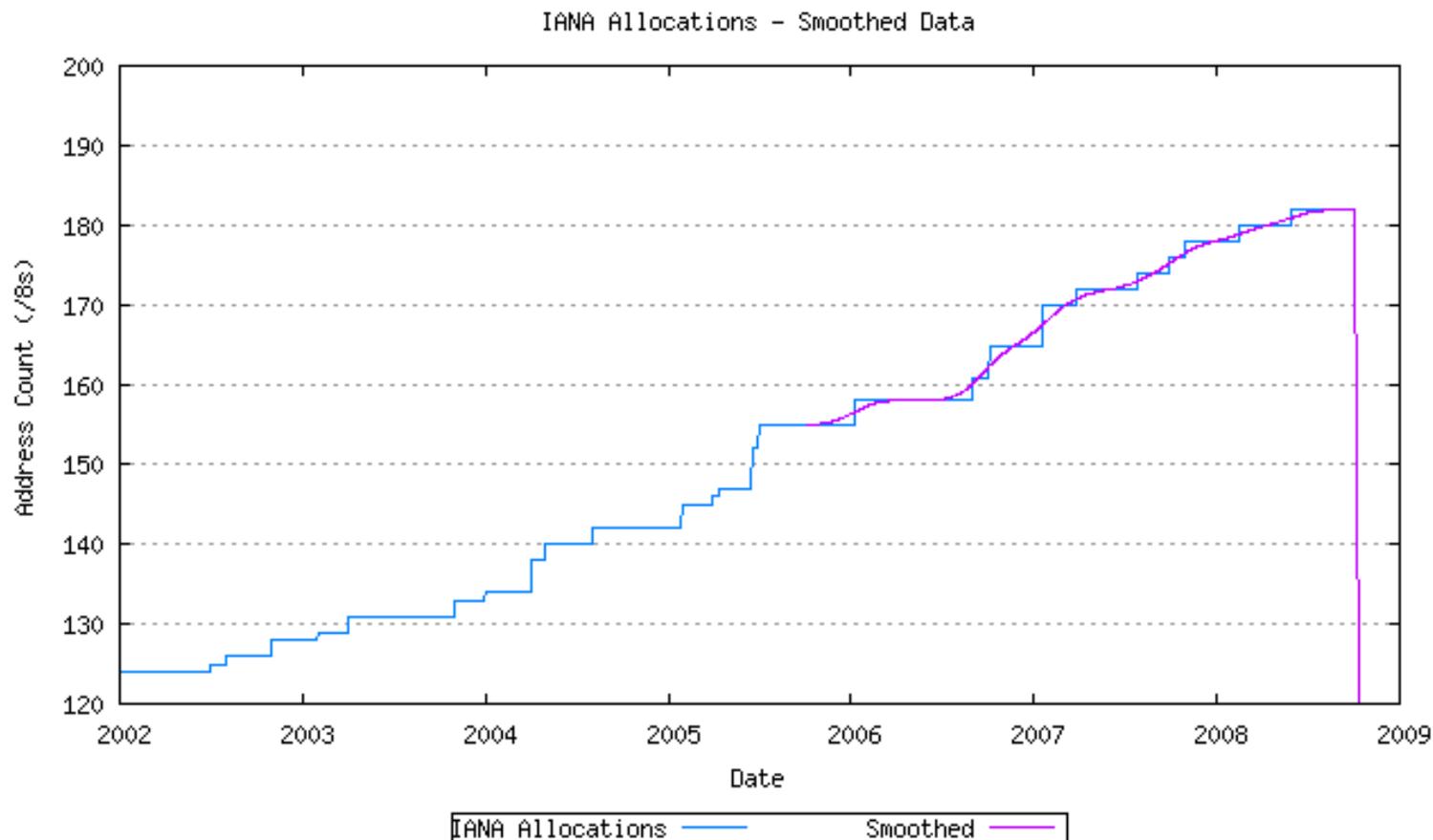
IPv4枯渇の状況



どうやら2011年というのは避けられないか？

出典: www.potaroo.net

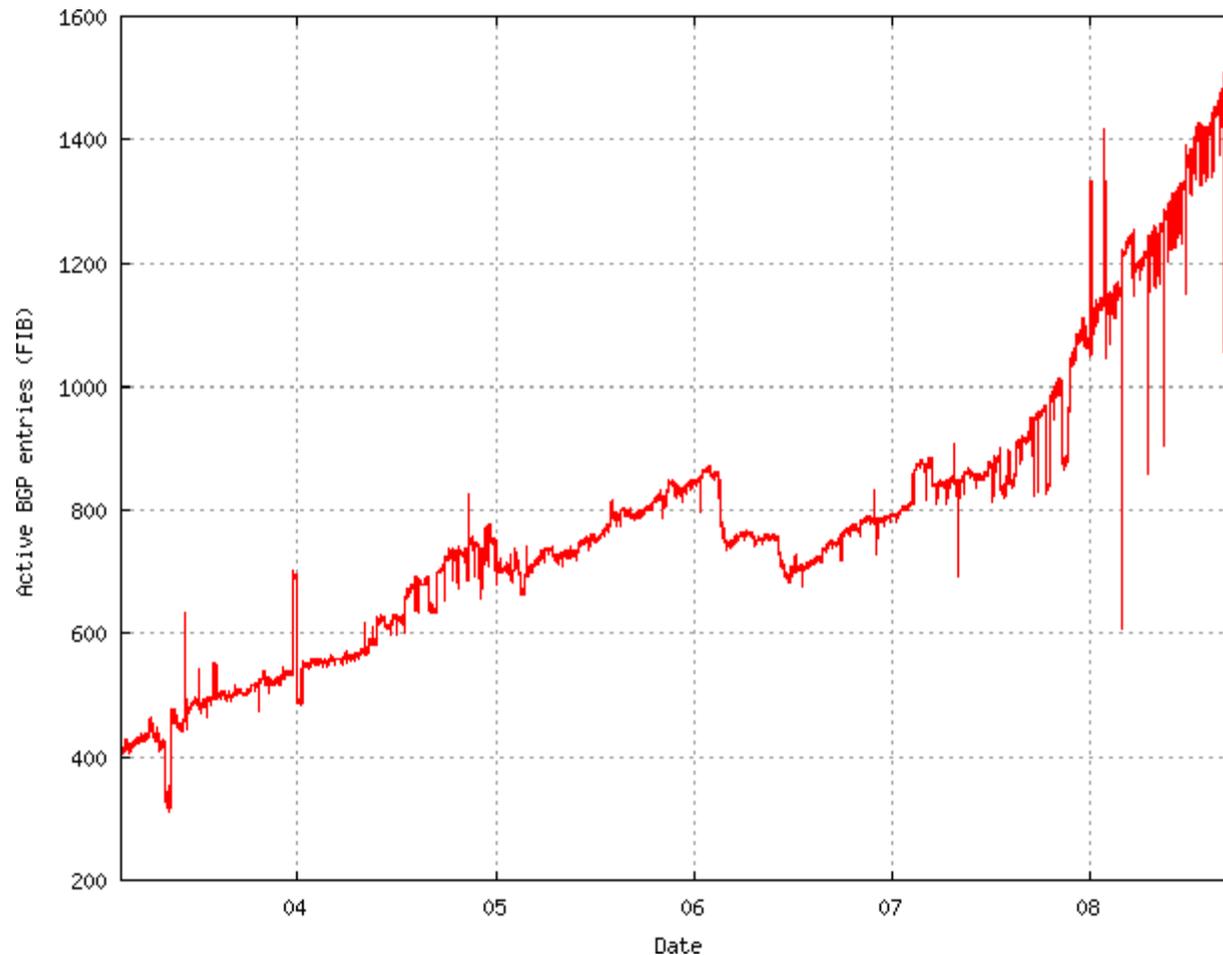
とはいえ傾向は少し変わったか？



なんとなく、消費の伸びが鈍化しているように見える

出典: www.potaroo.net

IPv6はどれくらい伸びているのか？



約1500Prefixes

伸びはかなり急激

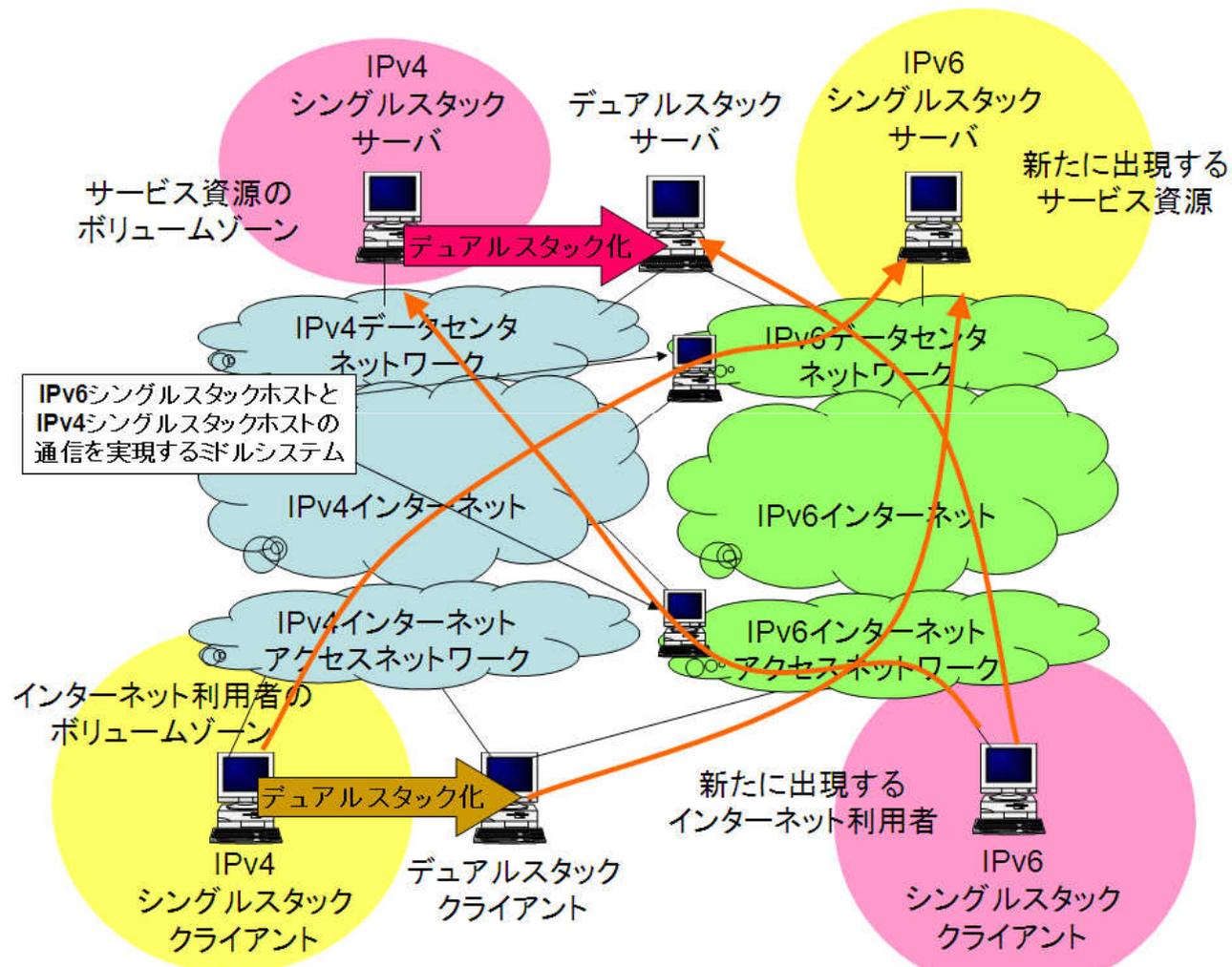
Aggregateしても
約1400Prefixes現存

出典: www.potaroo.net

IPv4とIPv6の関係

- では、IPv6が普及するとIPv4が減るのか？というのも含めて両者の関係を整理してみましょう。
 - IPv4とIPv6は、基本的に全く別のプロトコルである。
 - つまり、IPv4のインターネットとIPv6のインターネットは別物で、相互接続性は、原則存在しない。
 - え？じゃあ、トランスレーターって何者？
 - トランスレーターというのは、「プロトコル変換機」なので、一定の制限のもとIPv4とIPv6の間の相互接続を提供してくれる
 - もっと言うてしまうと、Application Level Gatewayと変わりありません。
 - さらに、アドレス帯域幅の関係もあり多くの制限が付きます。
 - じゃあ、IPv4でサービス続けられないってこと？
 - 端的に言うと、そういうことになるかもしれません。

IPv4とIPv6の関係を俯瞰してみよう



出典: 報告書「IPv4アドレス枯渇に向けた提言」

俯瞰図を整理すると

- IPv4とIPv6は、別々のネットワークとして構築される
 - まったく別の物理ネットワークかもしれないし、
 - デュアルスタックネットワークかもしれない。
- IPv6の規模が小さい場合(移行初期)
 - IPv4の必要なホストだけ、IPv6のデュアルスタック化
- IPv6の規模がIPv4より大きくなってきた場合(以降後記)
 - IPv6ホストの必要部分だけをIPv4のデュアルスタックするかもしれないし、
 - サービスが特定できれば、トランスレータを用意するかもしれない
- 接続サービスはどうするか？
 - ISPとしてトランスレータを入れるか？
 - 可能性はあるが、いずれにせよ完全なトランスレーションは無理
 - やるならユーザーのBBルータでトランスレーション？
 - キャリアグレードNATの可能性は？(次ページで・・・。)

キャリアグレードNATってどうなの？

- **キャリアグレードNATの真実**
 - ユーザーに多重NAT環境を提供することになります。
 - アプリケーションがますますつながらなくなります。
 - 自ISP内での折り返しは、NATBoxでの折り返しになります。
 - 対応していないNATBox多数。
 - 要参照: JANOG22「IPv4アドレス 販売終了のお知らせ～ISPによるNATで起きること～」
 - どっちにしろ、IPv4はそれなりの量が必要です。
 - 1つのIPで何人収容できますか？
 - 1つのユーザーで使っているセッション数ってはかったことありますか？
 - Ajaxバリバリだと、100セッション位使うユーザーがいます。
 - セッション制限かけると、アプリケーションに問題がでます。
 - 結構なコストがかかります。
 - ただでさえルータ高いのに、NATBoxに数千万払えますか？
 - トラフィックを抑制できれば安く済む可能性もあります。

IPv4が枯渇すると何が起きる？

- **単純ですが、IPv4の割り振り・割り当てが止まります。**
 - 他のISPからのIPv4買取は可能か？
 - そのようなポリシーは提案されていますが、策定は微妙な状態です
 - 仮に策定されたとしても、枯渇直後には相当な費用が予測されます
- **キャリアはどうするか？**
 - 方策1: キャリアグレードNATの導入
 - 方策2: IPv6接続サービスの開始
- **データセンターはどうするか？**
 - 方策1: キャリアグレードNATの導入？
 - あまり現実的ではないです。
 - 方策2: IPv6接続サービスの開始
- **ユーザーはどうするか？**
 - Vista使ってればそのままOKか？
 - NTTさん、JAIPAさんの議論がどうかなるでも異なるか？
 - いずれにせよあまり気にしない、というか気にさせない。
- **コンテンツサービス(EC業者を含む)はどうするか？**
 - ユーザ獲得を求めてIPv6対応業者へ流れる。

IPv4枯渇による厳しい現実

- **サーバーのアドレスは、NATの向こうに置けない**
 - Webサーバーなどは、現実的には、IPv4とIPv6の両方のアドレスが当面必要になる。そうしないと、ユーザーからのアクセスを受けられなくなる。
 - IPv4は、縮小傾向と考えれば、少なくともIPv6は必ず必要になる。
 - この時点で、IPv6サービスをしないという選択はない。
- **大手は、すべてIPv6 Ready**
 - 多くの大手では、IPv6サービス出してないじゃないか！
 - いえいえ、結構検証とかしっかりやって準備しています。
 - タイミングを見計らっているだけにしか、私には見えません。
- **IPv6の基本は、/32のProvider Aggregatable Addressです**
 - 一応、/48のPIアドレスも取得できます。
- **より進む、東京集中化**
 - FLETSのインターネット解放、NGNの開始など、東京集中型モデルを促進する要素は次々。

IPv4枯渇に備える

- とにかく準備しておきましょう。
 - IPv6アドレスをとりあえずもらいましょう。
 - 200顧客制限という条項は既にあります。
 - IPv6アドレスのアドレス維持料は、ほぼ無料です。
 - IPv4アドレスの維持料とIPv6の維持料を比較し、高い方がのみが課金されます。
 - <http://www.nic.ad.jp/ja/ip/member/ijiryo.html>
 - IPv6の運用経験を積みましょう
 - IPv6のネットワーク運用などは、どこを見ても経験豊かなところはありません。
 - 早めに、実験用ネットワークでも良いので敷設して、実験をしておきましょう。
 - 意外と落とし穴が多いです。
 - アドレス利用ポリシーも決めておく必要があります。
 - 対応は、コストがかからないところが順次していきましょう
 - ルータなどは、すでに対応しているものが多いはず。ファームウェアのアップグレードなどもしておきましょう。
 - サーバーは、ほとんどがデュアルスタックですので、影響の少ないところから徐々に対応しておきましょう。

IPv6ネットワーク管理の実態

- IPv4より簡単なアドレス管理
 - エンドネットワークには、/64が固定
 - このセグメントには、アドレス何個必要なんていう面倒な管理はいりません。
 - 何も考えず/64を割り当てちゃえばよいのです。
 - /48あれば、65536このエンドネットワークが収容できます。
 - PIでIPv6とっても十分な広さです。
- LIRからの割り当てでできないの？
 - 出来ます。が、BGPを使うという方法もあります。
 - これを機にBGP運用という考え方も良いかもしれません。
- 意外と面倒なIPv6の各種設定・運用
 - IPv6は、いろんな落とし穴があります。
 - CentOSで突然default gatewayがunreachableになる(バグ?)
 - 開発者によると既に治っているそうです。
 - なんか違うホストをつなぐと突然RAをはいて、Default Gatewayが奪われる
 - 気がつくと、インタフェースに謎のIPv6が割り当てられている
 - 実験用のホストでも接続した時に、そいつのIPv6 subnetでアドレスが勝手についてしまうなど・・・。

IPv6でBGPする

- BGPは大規模ネットワークだけのものではない
 - BGPの基本コンセプトは、「トラフィックエンジニアリング」
 - IPv6は、大きなアドレスブロック1つで制御する必要があるため、IPv4に比べてトラフィックエンジニアリングは、多少難しい
 - IXとBGP、地域内のトラフィック交換でトラフィックに対するコストを最小化する
- BGPにするリスク
 - BGP運用 = 独立ネットワーク
 - 「トラブルは自分で解決する」が原則
 - トラフィックエンジニアリングは上流に任せっきりにはできない
 - BGPに対する正しい知識が必要
 - リスクは回避できる
 - コミュニティやコンサルテーションによる補助をつけ、困ったことは相談できる環境を作る

BGPって難しいの？

- 「BGPって難しいんでしょ？」
 - 実はそんなに難しくありませんよ。
 - エッジネットワークの場合、2～3対地なのでバリバリの制御は必要なし
- BGPにするとお金かかるんでしょ？
 - AS番号とるのにJPNICに払う料金は、1万円
 - IPアドレスは、維持料がかかります。
 - IPv4の/22なら、年間26万円
 - 指定事業者になるなら、最初に26万円
 - 後はBGPルータを用意します。
 - 多分これが一番高いです。
- 技術的には？
 - 最初にPeer張ってしまえば、あとはよっぽどのことがない限り放置です。
 - ただ、回線などは、通常と同じようにちゃんと管理しましょう。

BGPの準備

- BGPに必要なもの
 - AS番号
 - PA/PIアドレス (IPv4/IPv6)
 - BGPルータ
 - BGP接続先
 - Full Transit, IX接続
- BGPを始めるための計画
 - 移行計画
 - 今のネットワークをポンとBGPに切り替えるわけではありません
 - どのように、BGP化していくかの計画を立て順序よく実施する必要があります

BGP接続先の選定

- Full Transit (Full route)を最低1か所から
 - BGPで運用する場合、デフォルトルートは存在しません。
 - インターネットの全経路をどこからか、「買う」必要があります。
 - おそらく、接続を買う場合、この部分が一番コスト高。
- トラフィック分散のためのBGP Peers
 - Full Transitは高いので、BGP Peerにトラフィックをより多く逃がし、Full Transitのコストを最小化します。
 - そのためには、トラフィック分析が必要です
 - NetFlow/sFlowなどで、トラフィック交換しているASを割り出します
 - 割り出したAS別トラフィックから、なるべく多いものをBGP Peerににがします。

BGP Peerの種類

■ Internet eXchange

- IXに接続すると、多くの仲間がそこにいます。
 - JPIX, JPNAP, Dixie
- 但し、すべてのASがPeerをしてくれるわけではありません。
 - 接続に先立って、Peerしてくれるところをあらかじめ調整しておく必要があります。
- 地方からのIX接続は、回線コストがばかにならない
 - IXのほとんどは東京・大阪
 - 基本的は、IXまで自前で接続を用意しないといけない
- 商品名で恐縮ですが・・・
 - Softbank TelecomのAssocioを經由してJPIXに接続できます。
 - 地方にいながらIXへのダイレクト接続を購入できる。
 - 但し、接続にMPLSの設定が必要。

■ Private Peer

- 地域の中で、直接トラフィックを交換する方法があります。
- 地域の中なので、ダークファイバーをうまく使うことで、コストは最少化できます。
- 複数の事業者が集まるなら、地域内専用のIXを共同で運用する方法もあります。

IPv6 BGP Operations

- IPv6におけるBGPとは
 - BGP4+がIPv6対応BGPとされています。
 - 実態は、BGP4そのままです。(RFC4771)
- では、どうやってIPv6を扱うの？
 - BGP4 Multiprotocol Extension Capabilityを使います。
 - つまり、BGP4のオプションでIPv6を扱うのです。
 - プロトコル的には、IPv4をしながらIPv6の経路情報交換も可能です。
- どんなプロトコルですか？
 - 基本的な情報は、IPv4と全く同じです。
 - 経路情報を扱う部分だけが違います。

Multiprotocol BGP Extension Capability

- RFC2283
- 目的
 - BGPは基本的にIPv4のみを扱うように設計されている
 - MPLS, IPv6などIP上で展開されるプロトコルの多様化によって、情報伝達ツールとしてのBGPに注目される
 - ニーズにこたえる形で、複数のプロトコルに関連する経路情報を伝達するために、マルチプロトコル対応の拡張が行われる

MP_BGP

- BGPのCapability Optionとして定義される
 - Multi-protocol Reachable NLRI
 - MP_REACH_NLRI : Type Code 14
 - Multi-protocol Unreachable NLRI
 - MP_UNREACH_NLRI: Type Code 15
 - 通常のBGPと同じようにNext_HopとNLRI(ネットワーク到達性情報:一般にPrefixとPrefix Lengthのペア)で構成される
 - このほか、プロトコルを特定するためにアドレスファミリ(AFI)とSubsequent AFIなどのフィールドもある
- BGP4の通常のNEXT_HOP属性はMandatoryであるが、MP_BGPの登場により、MP_BGP内のNEXT_HOPを正しく指定することによって、元来のNEXT_HOPを指定する必要はなくなっている。
- 現在では、IPv6経路の伝播やMPLSのラベル情報の伝播(RFC3107)などに利用されている

MP_BGP – パケットフォーマット

Address Family Identifier (2 octets)
Subsequent Address Family Identifier (1 octet)
Length of Next Hop Network Address (1 octet)
Network Address of Next Hop (variable)
Number of SNPAs (1 octet)
Length of first SNPA(1 octet)
First SNPA (variable)
Length of second SNPA (1 octet)
Second SNPA (variable)
...
Length of Last SNPA (1 octet)
Last SNPA (variable)
Network Layer Reachability Information (variable)

※RFC2283より
MP_REACH_NLRI フォーマット

IPv6の上流・・・。

- **明確なIPv6の上流はありません。**
 - いくつかのISPでは、IPv6のFull Routeを準備してくれていますが、IPv4と抱き合わせも多いです。

- **その他のIPv6 BGP経路を扱っているところ**
 - Distix
 - Associo経由で、MPLSでIX接続が可能です。
 - JPIX
 - IPv6のIXサービスを始めています。

- **IPv6のFull Route自体、明確に「コレ」というものはない。**
 - IPv6ネットワークが若干まだ流動的
 - そもそも物理ネットワークとしてIPv6を展開しているところは少ない
 - ほとんどは、tunnelで構成されている場合が多い

結論

- IPv6の準備は
 - 準備は、早々に初めて、自社としてどうアプローチすべきかは、検討しておかなくてはならない時期です。
 - サービスを開始するかどうかはタイミングを見ながらでよいでしょう。
 - ただし、先行投入の余裕があるならしておいたほうがよい
- IPv6 BGP
 - BGP Multihomingへ行くためのよいタイミングであることは間違いありません。
 - 検討しても損はないでしょう。
- IPv6は、もう目の前です。

The NEXT Network Management Solutions

S P E E D

Emergency notification system



The GOOD Collaboration Ware



<http://www.ate-mahoroba.jp/>