

# ルーティングを語ろう

Yoshida 'tomo' Tomoya

<yoshida@ocn.ad.jp>

Matsuzaki 'maz' Yoshinobu

<maz@iij.ad.jp>

# はじめに

- インターネットの運用は協調と連携が必要
- 経験や知識は共有してこそ、みんなの幸せ
  - 日本から色々と提案をしていきたい
- 皆様のご意見を聞きたい！

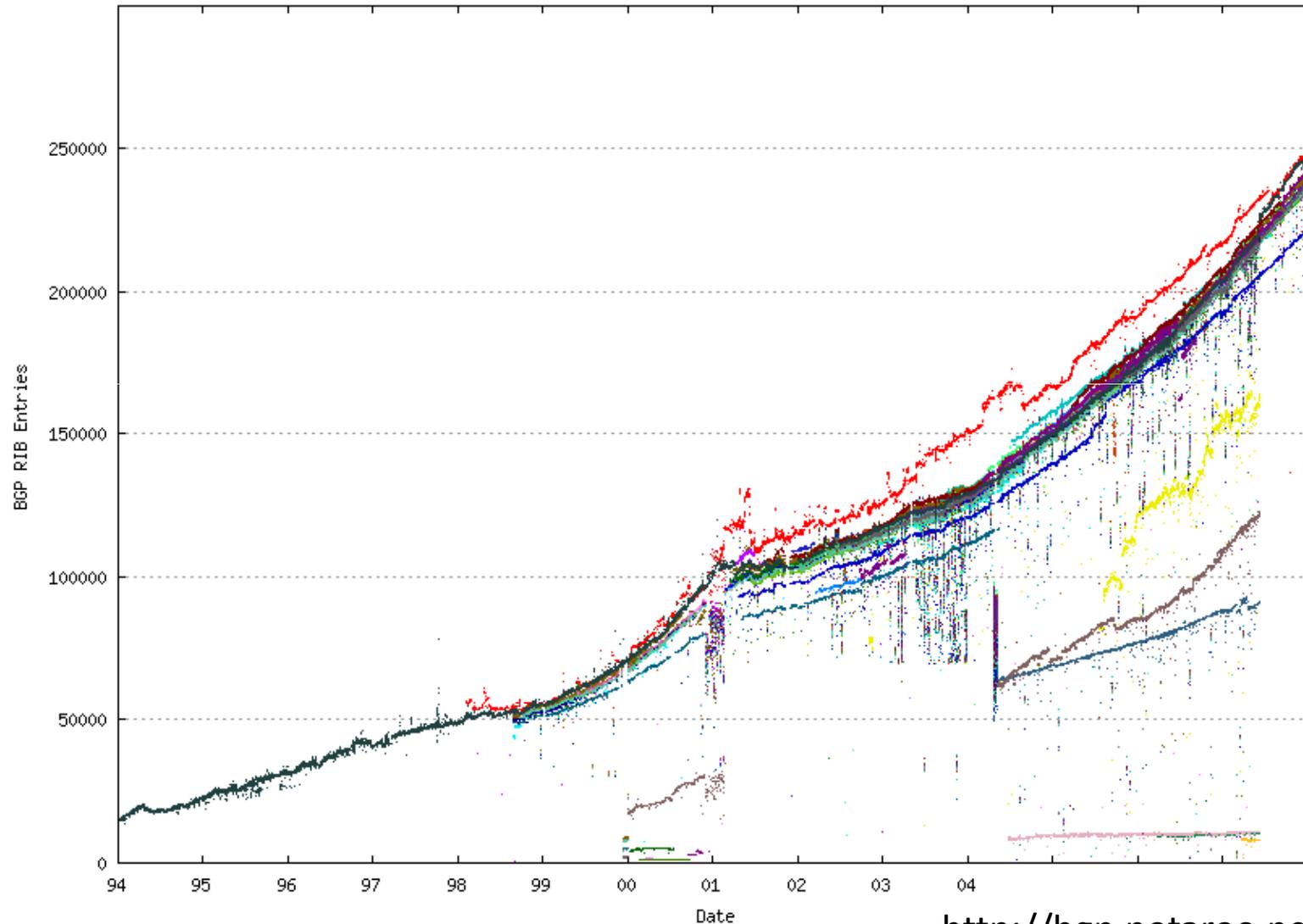
# 最新動向

1. 経路数の増大
2. 4-octet AS
3. IPv4枯渇関連

# 経路数の増大

- 待ったなしで伸び続ける
- IPv4枯渇期周辺には、40万経路に到達かも
- メモリも使うし、コンバージェンスにも影響
- 最近では256kの壁にぶち当たることが懸念されている
  - 過去は192Kや224Kなど、何度か壁があった
- 不必要に多いのなら良いことはない

# IPv4経路数の推移

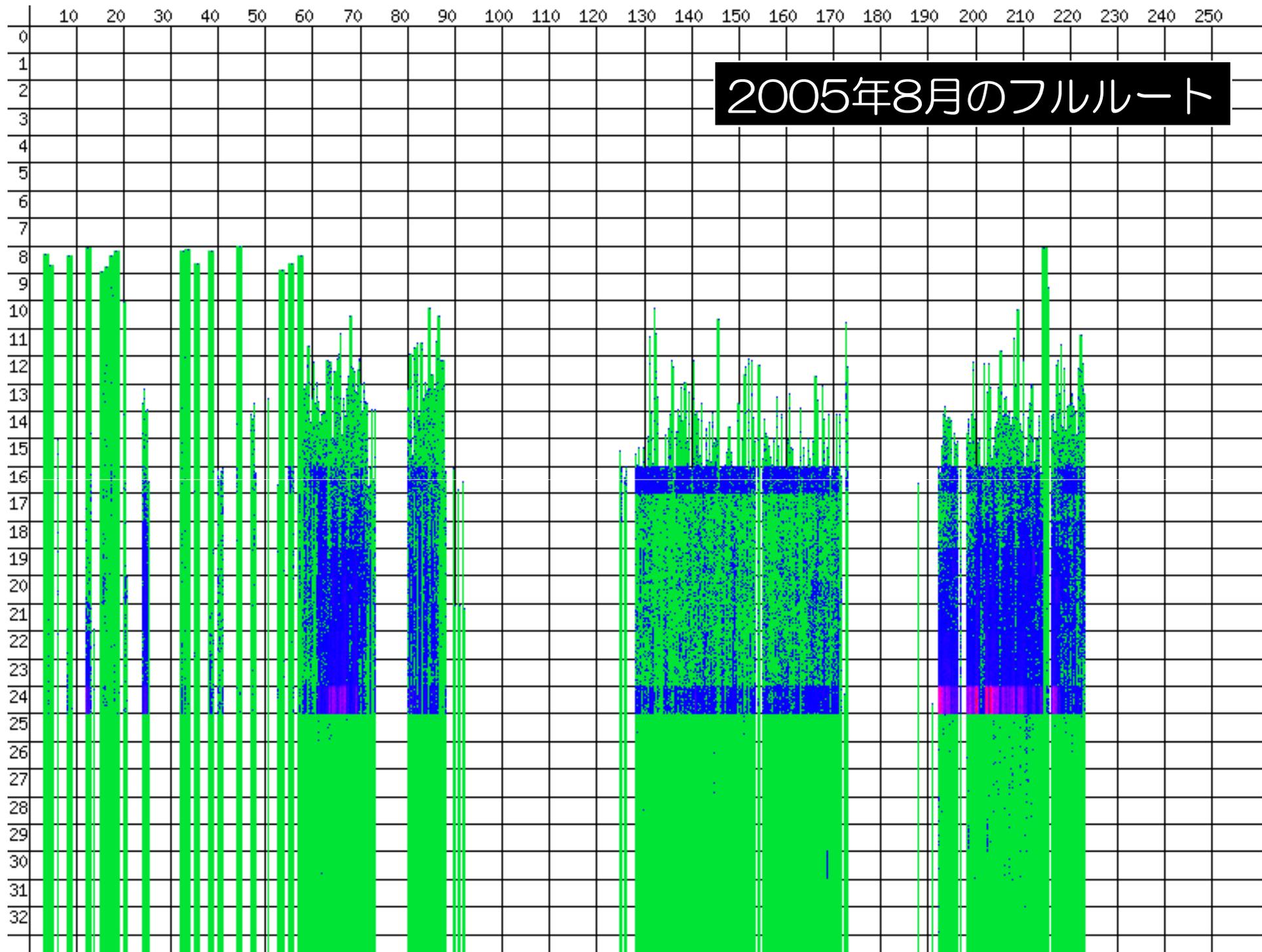


2008/01/24

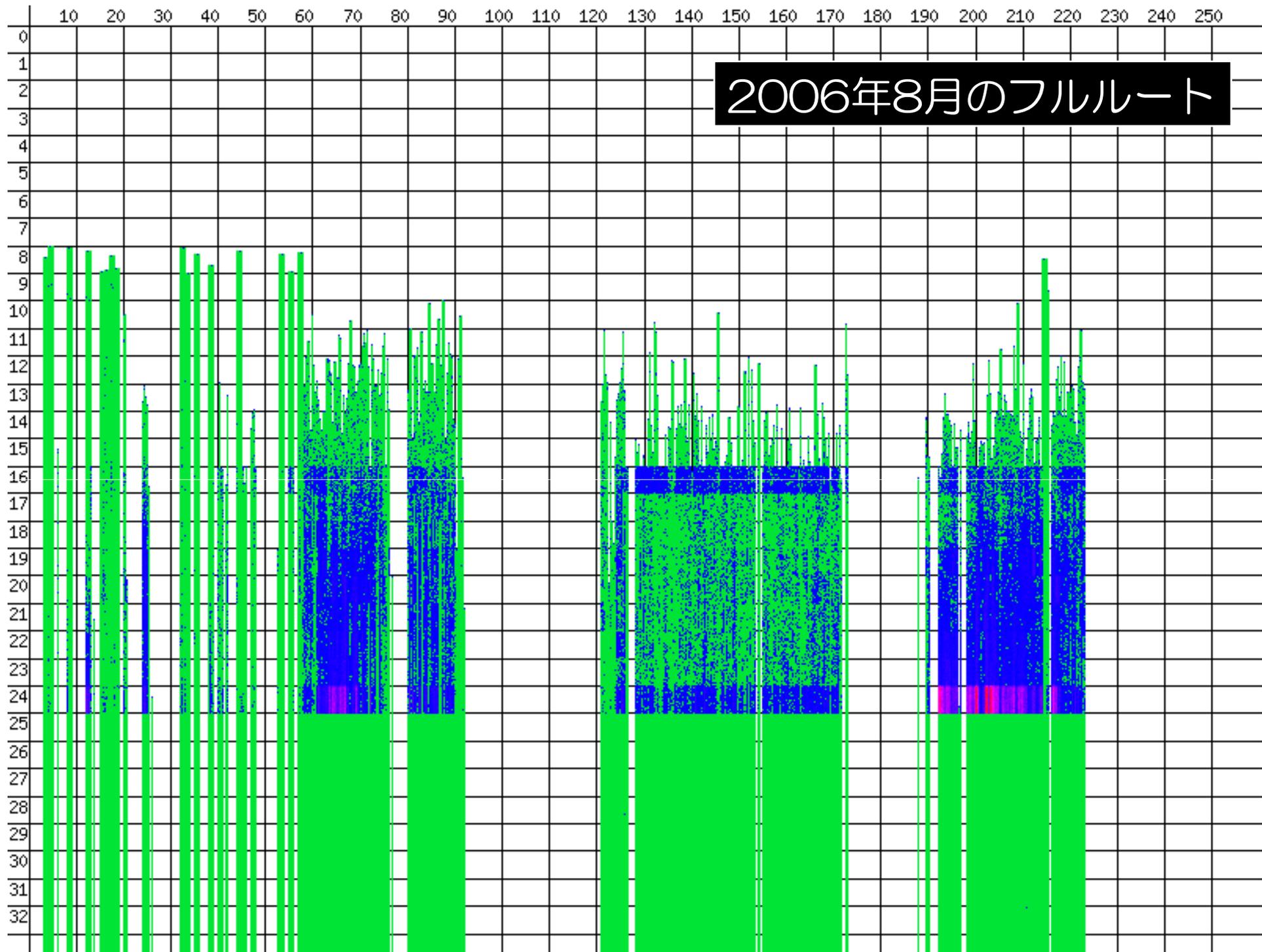
Date  
tomo&maz

<http://bgp.potaroo.net/>

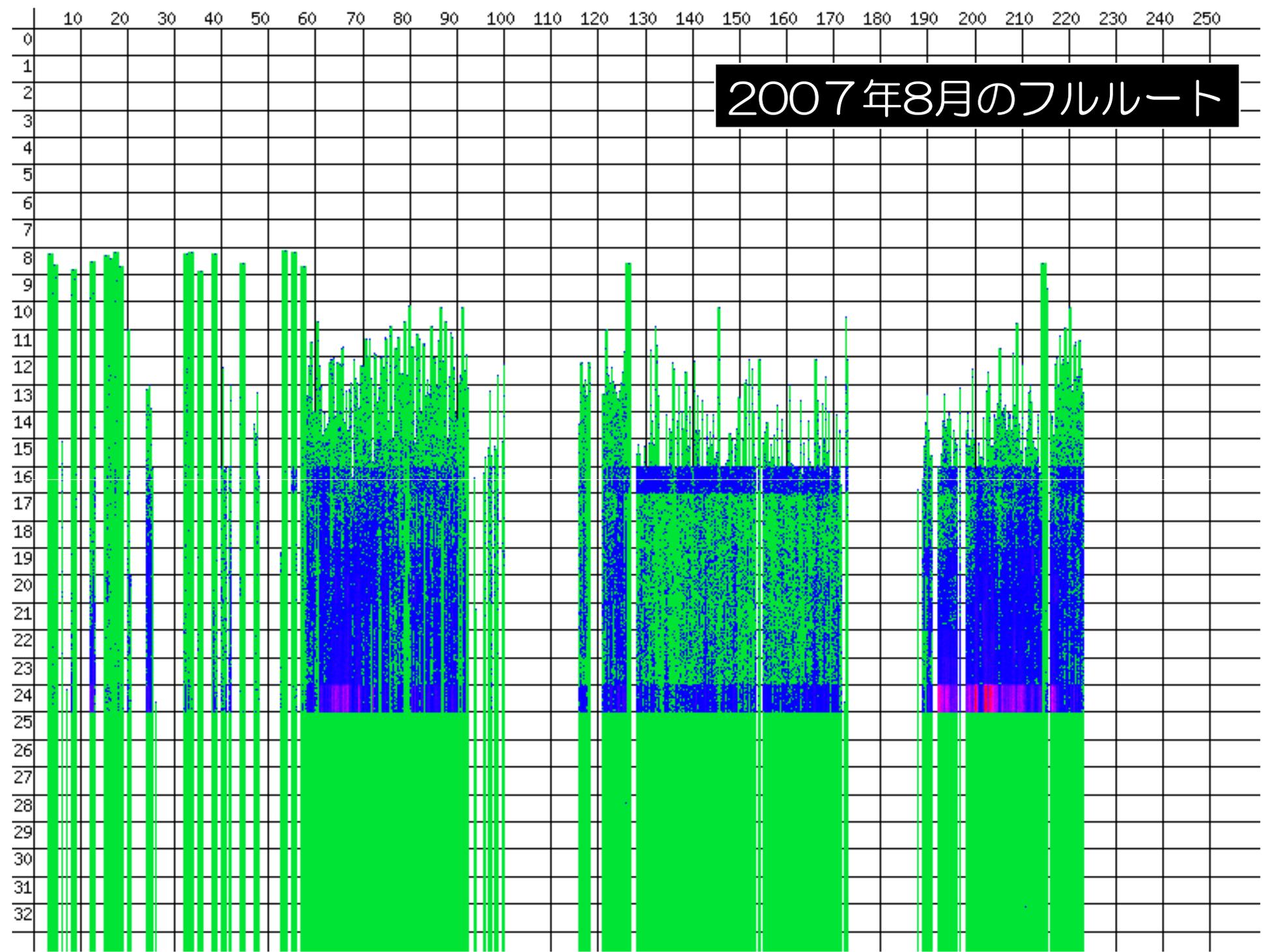
# 2005年8月のフルルート



# 2006年8月のフルルート



# 2007年8月のフルルート



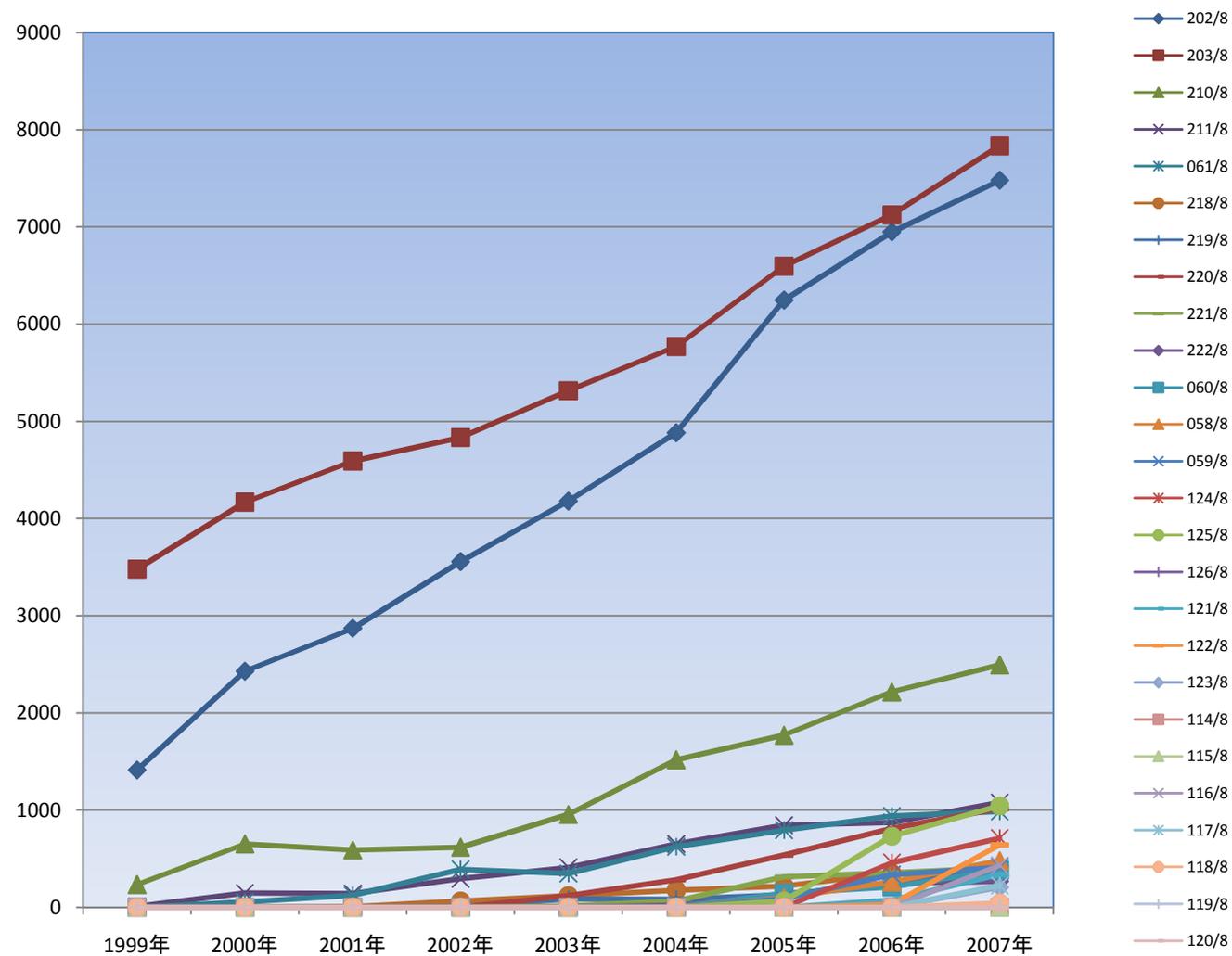
# 経路増大の原因

- トラフィックエンジニアリング
- パンチングホール
  - 新規ブロックの代替かもしれないけど、、
  - 細かい経路は増加する傾向に
- 特にアジア地域がひどい
- アグリゲーションすれば、約15万ぐらいに
  - 当面はメモリの枯渇や装置の仕様上の上限値を超えないで済むこと間違いなし
- IPv6時代を考えると、最初からある程度コンセンサスを作りたい

# /8毎の/24の数

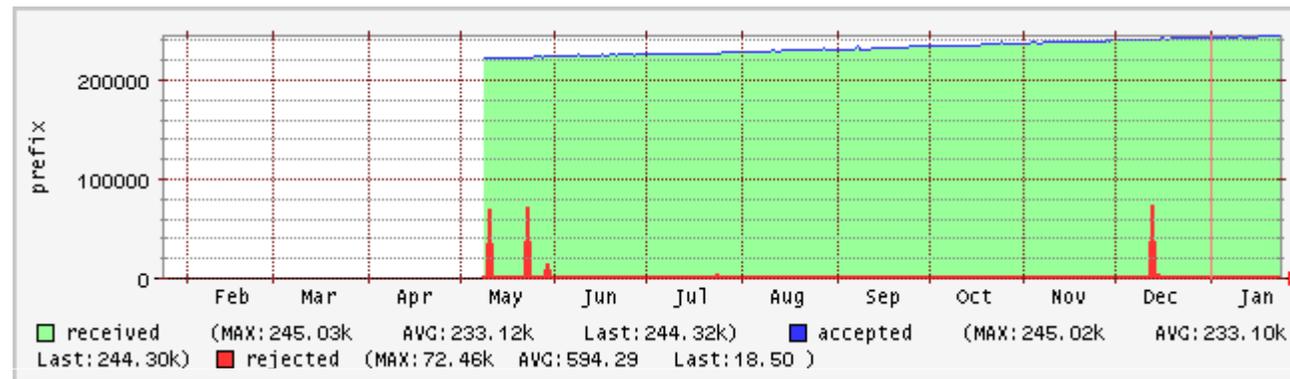
4:9	8:89	12:1830	15:19	16:3	17:6
18:13	20:35	24:994	25:1	32:63	33:3
38:360	40:64	41:390	44:2	47:8	50:1
52:3	55:3	56:3	57:21	58:479	59:413
60:384	61:987	62:939	63:1980	64:3299	65:2373
66:3592	67:909	68:679	69:2016	70:466	71:106
72:1437	73:7	74:760	75:264	76:153	77:420
78:272	79:11	80:876	81:793	82:579	83:350
84:617	85:824	86:358	87:597	88:283	89:1018
90:39	91:741	92:5	96:2	97:6	98:11
116:430	117:224	118:44	121:338	122:642	123:207
124:711	125:1043	128:309	129:195	130:142	131:369
132:56	133:9	134:194	135:34	136:192	137:82
138:152	139:74	140:516	141:146	142:373	143:258
144:323	145:48	146:326	147:125	148:486	149:184
150:169	151:137	152:95	153:121	154:11	155:261
156:164	157:224	158:160	159:197	160:155	161:101
162:190	163:184	164:497	165:405	166:290	167:309
168:513	169:130	170:420	171:29	172:5	189:116
190:1200	192:5725	193:4036	194:3202	195:2446	196:927
198:3804	199:3227	200:5250	201:1252	202:7477	203:7832
204:3981	205:2001	206:2421	207:2781	208:2955	209:3596
210:2492	211:1079	212:1308	213:1536	214:419	215:44
216:4095	217:1189	218:320	219:399	220:1016	221:407
222:266					

# アジア地域の/24の推移

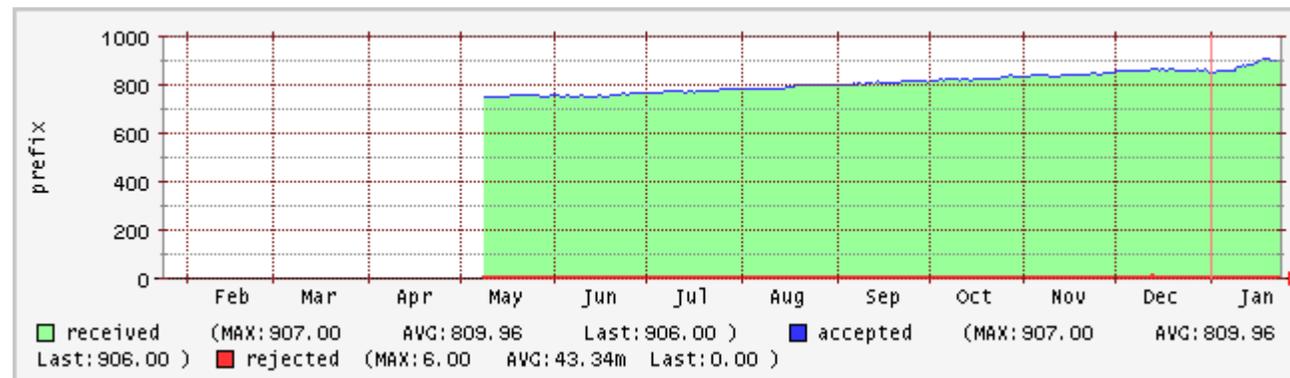


# 内部で見えてる経路数

- IPv4



- IPv6



# ルータのメモリ

- Control plane
  - RIBの経路数
  - route reflectorなどで経路数を減らす？
- Forwarding plane
  - FIBの経路数
  - multipathとかで増える傾向に
  - full routeをあきらめる？
- 今のところ、メモリの増強で対応！

# メモリ枯渇

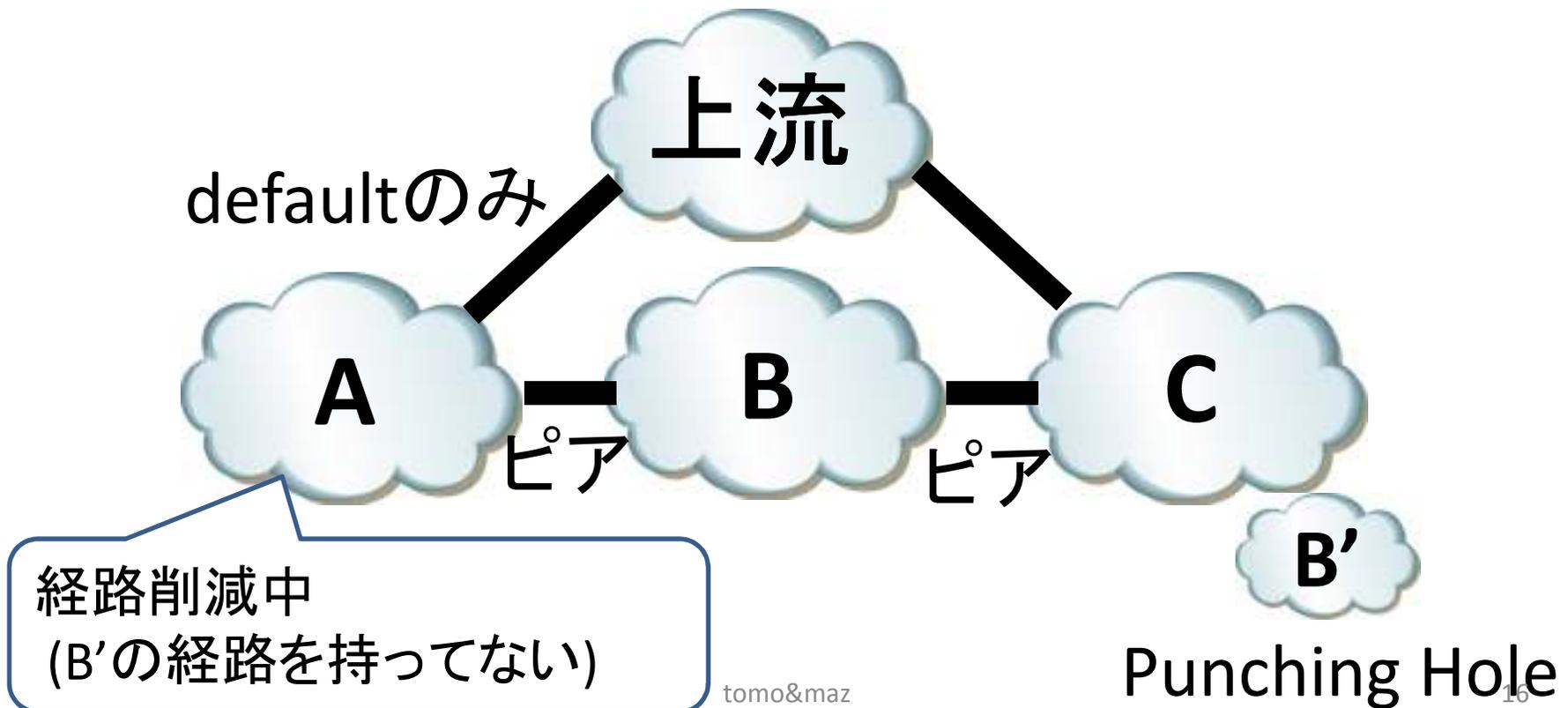
- 真っ向勝負な解決策＝メモリ追加
  - 機種によっては、メモリ搭載量の制限
  - これまで通り動く
- 背に腹は代えられない解決策＝経路削減
  - 適当なルールで経路を削減
    - ex. 上流ASからはdefaultのみ広報してもらう
    - ex. 細かい経路を受け取らない
  - 副作用が怖い

# ad-hocな対応で問題が・・・

- 経路削減
  - default運用すると宛先がroutableじゃないパケットも運ぶ
  - BGP的な最適経路を選べない
  - uRPF loose modeとか使ってる場合は要注意
    - 気にせず落としちゃうかも

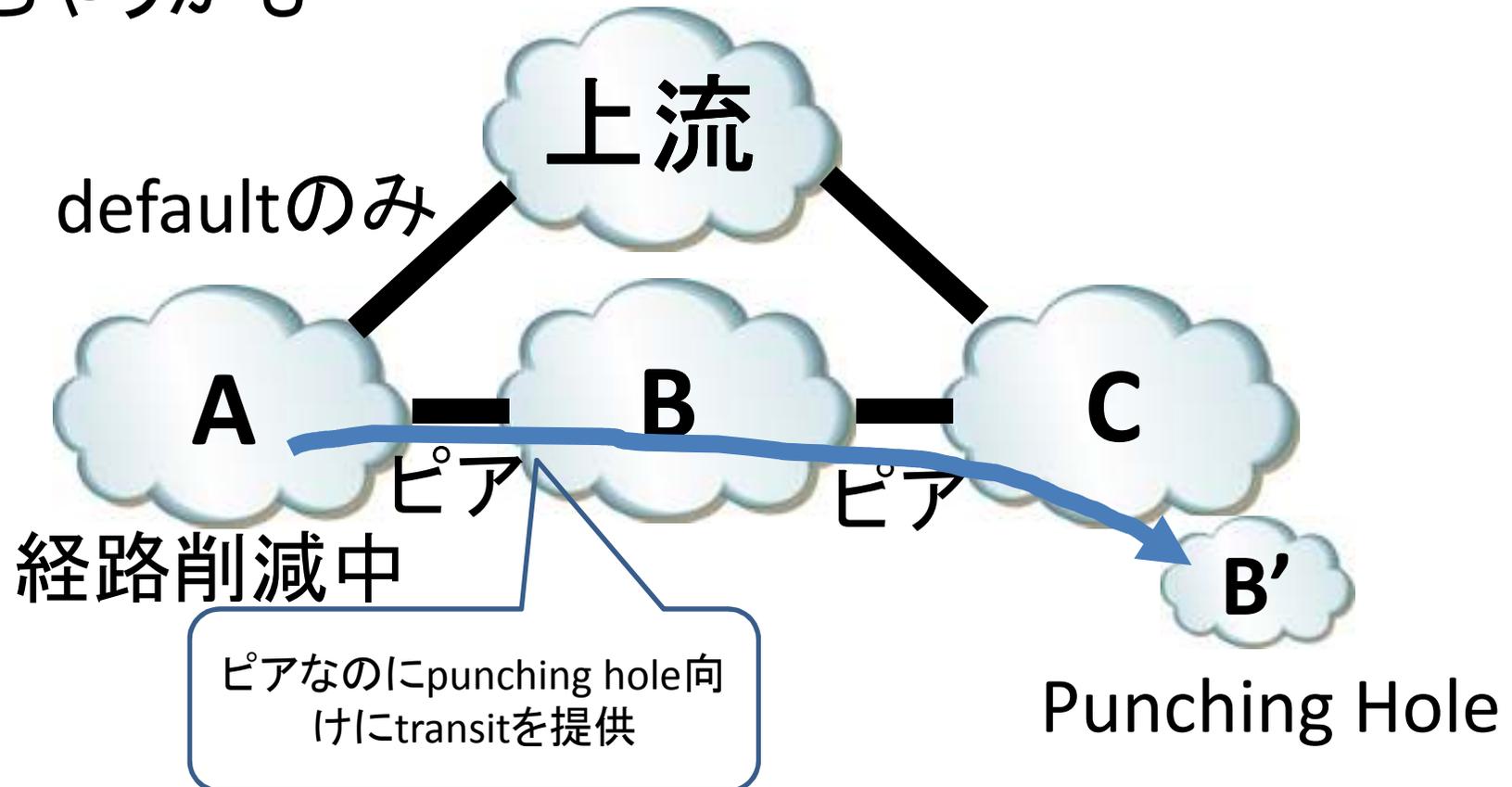
# punching holeと経路削減

- Bのprefixをpunching holeして別ISPから広報
- Aは経路削減中で、上流からdefaultのみ受信



# punching holeと経路削減の副作用

- どこかで意図しない経路制御状態が生まれちゃうかも・・・



# どうするのがいい？

- punching hole
  - やめといた方が安全
  - どうしてもやりたいと言われたら、元となるPAのISPとpunching holeを広報するISPを集めて話し合っておくべき (ex. routing, タダ乗り)
- 経路削減
  - 最適経路を選べない副作用を十分理解する
  - どうしてもコストが問題ならBGP運用をあきらめて、上流に運用を任せた方が安定するかも

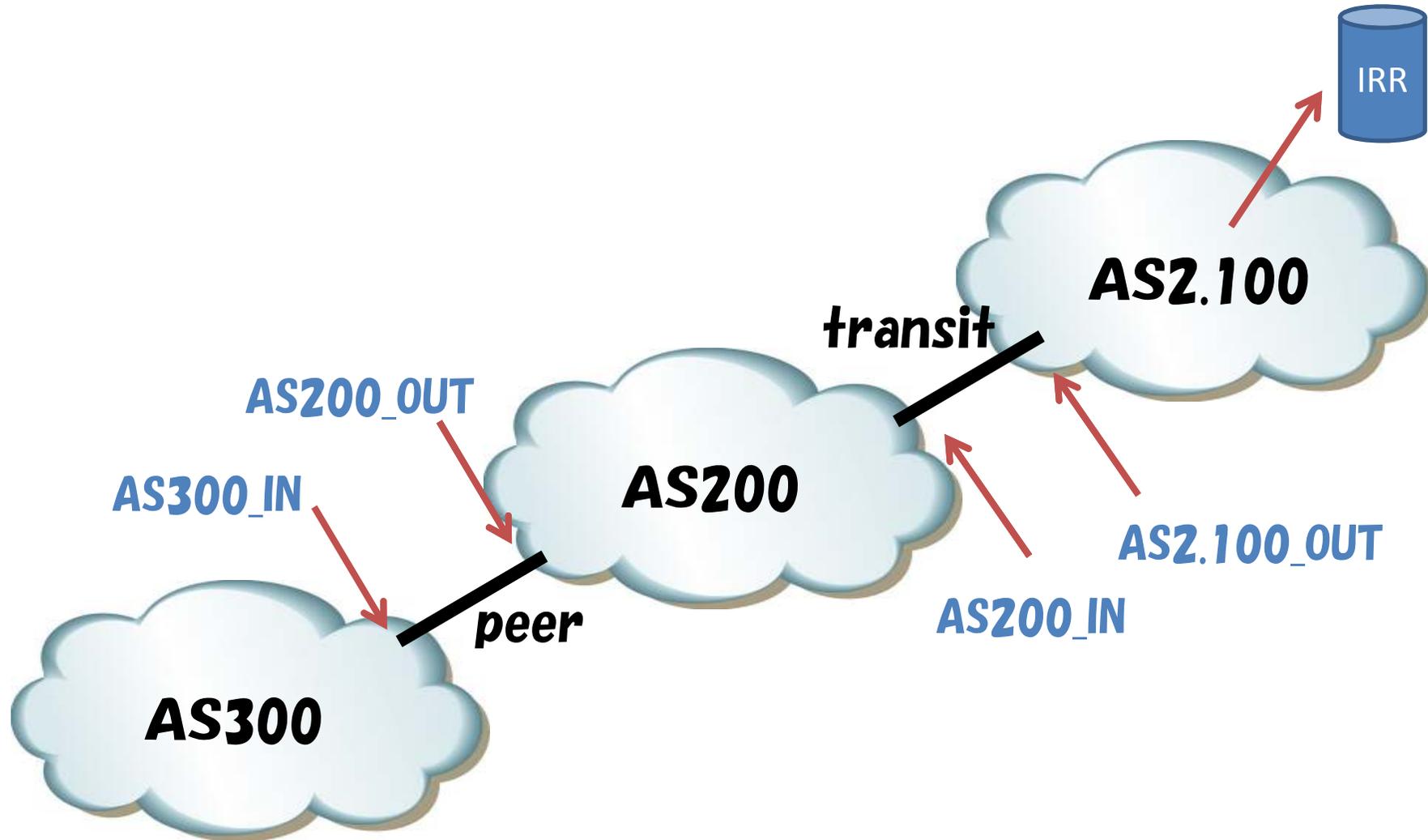
# 対応

- 細かい経路によるトラヒック制御は今後、経路数増加に伴って、意図しない状況を作り出すかも
- やっぱりメモリ追加で対抗するしかない
  - ちゃんとした運用は身を助ける
  - 不必要な経路を減らす努力もしましょう

# 4octet-AS

- 最近4octet-ASのAS\_PATH UPDATEが日本国内でも流通しはじめた
  - 自分が2octetの場合どう対処すべきか？
  - 自分が4octetの場合どう対処すべきか？
  - Transit ASではどう対処したらよいか
- 標準的な対応方法や注意事項は皆で共有して、問題をあきらかにしていく
- この営みがとても大事
- 前回のIRS15でも議題の1つに

# 登場人物と確認ポイント



# 1. IRRに登録

- AS2.100で登録
  - Route、AS-Set、AS
- AS23456でも登録?
  - Route、AS-Set
  - ASは4octetのみで登録か？

移行段階では、Transit AS 等がFilterを作成す  
際や、一般に経路確認時に両ASで見える必要  
がありそう

# IRRへの登録例1

```
route: 196.1.15.0/24
descr: Afrinic 32bit ASN test c/o Internet Solutions (Pty) Ltd (South Africa)
origin: AS23456
mnt-by: MAINT-AS3741
changed: nishal@is.co.za 20070914
source: RADB
```

```
route: 196.1.15.0/24
descr: AFRINIC 4-btyle ASN testing prefix
origin: AS5.1
mnt-by: AAP-MNT
changed: aalain@trstech.net 20070917
source: RIPE
```

# IRRへの登録例2

```
route: 59.106.254.0/24
descr: SAKURA-D
      SAKURA Internet
origin: AS23456
notify: peer@sakura.ad.jp
mnt-by: MAINT-AS7684
changed: w-yoshioka@sakura.ad.jp 20071226
source: JPIRR
```

```
route: 59.106.254.0/24
descr: SAKURA-D
      SAKURA Internet
origin: AS23456
notify: peer@sakura.ad.jp
mnt-by: MAINT-AS7684
changed: w-yoshioka@sakura.ad.jp 20071226
source: RADB
```

# AS2.0の対応方法

	IRRが4octet未対応	IRRが4octet対応済 (RIPE, APNIC)
2octetの人に対して	<p>① 問題: 4octetが記述できない</p> <p>↓</p> <p>AS23456で記述するが、 本当のoriginがわからない</p> <p>対処案</p> <ol style="list-style-type: none"><li>1. description等に4octet-ASを書く</li><li>2. attributeを新規に定義する</li></ol>	<p>③ 問題: AS23456を記述しない</p> <p>↓</p> <p>2octetの人にAS23456と認識されない</p> <p>対処案</p> <ol style="list-style-type: none"><li>1. ツールで変換対応</li><li>2. description等にAS23456を記述する</li><li>3. Route ObjectをAS23456でも作成</li></ol>
4octetの人に対して	<p>② 問題: 4octetが記述できない</p> <p>対処案</p> <ol style="list-style-type: none"><li>1. description等に4octet-ASを書く + ツールで4octet-ASを参照</li></ol>	<p>④ 問題: 基本的には無し</p> <p>1. ツールが4octet-AS対応済みならOK</p>

# 4octet AS PATH Update

- こんなAS-PATHのアップデートだけではNG
  - $^(200_)+(23456_)+\$$
- きちんと4octet-ASを伝えないといけない
  - $^(200_)+(2.100_)+\$$
- “.”表記にも気をつけないといけない
  - そもそも”.”を使わない箱もある
    - $^(200_)+(131172_)+\$$ 
      - $65536*2+100=131172$
- OSのバージョンアップ時には要フィルタ変更

# IPv4枯渇関連

- LISP
  - LocatorとIDを分離して経路制御しよう
  - 複雑さが増す欠点もあり、未だに議論は継続中
- 240/4
  - IETF Reserved (旧classE空間)を有効利用
    - IABの一部の人が率先して提案している
  - 現状使えない機械は多く、ハードルは高い
    - Windows, UNIX, ルータ

# インターネットとルーティング

1. インターネットのとらえ方
2. IPv6 routing

# インターネットのとらえ方

- ルーティングという切り口で考えてみよう
- 人/状況によって注目するところが異なる
  - 到達性
  - 回線の利用効率
  - 最適化
  - 制御ポリシ

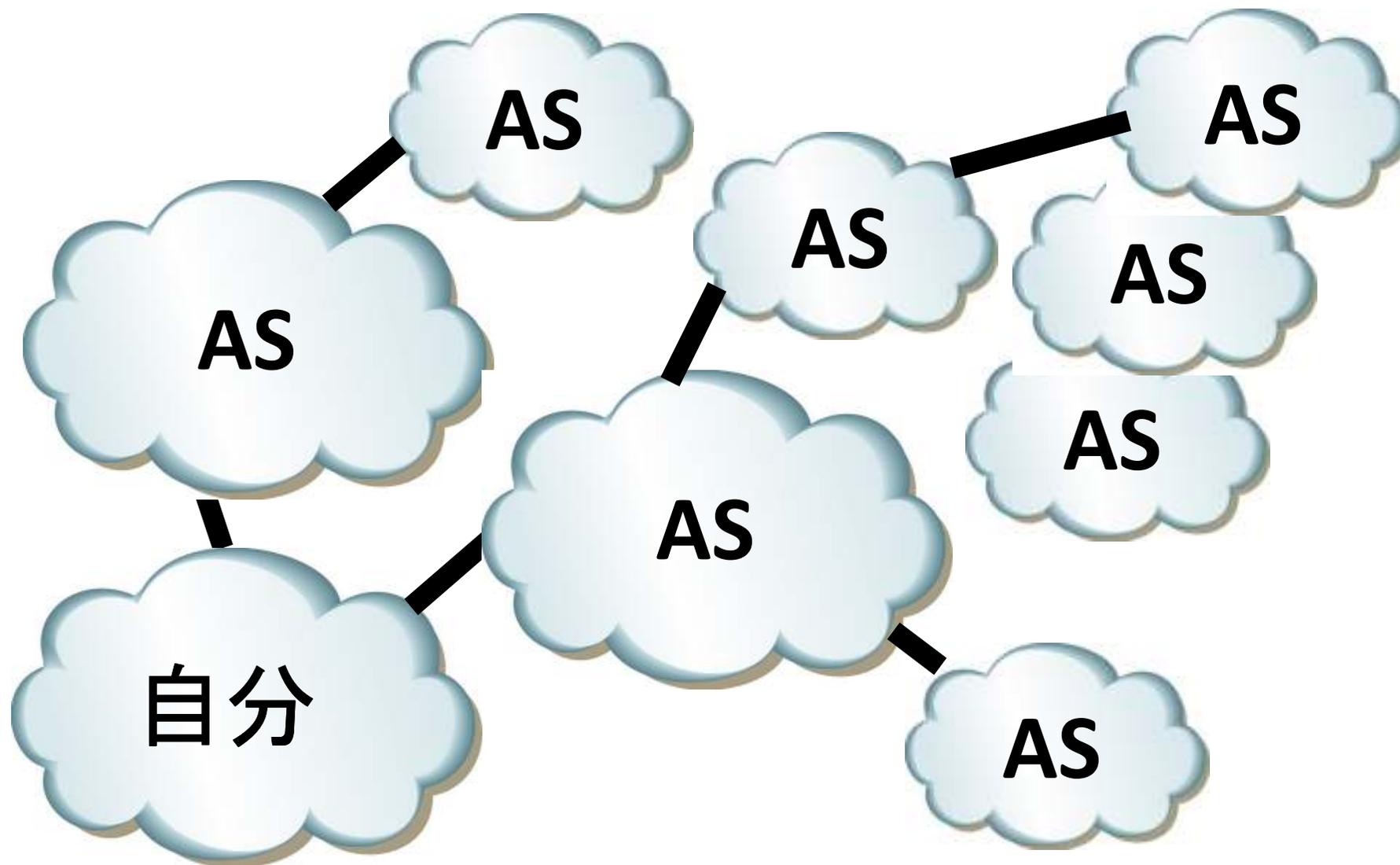
# 1.もよん派



# 1.もよん派

- インターネットはもやもやした雲だと思ってる
  - パケットをよろしく相手まで届けてくれる
  - 使えてればその構成にはあんまり興味がない
- 近隣ネットワークとの制御が主な目的
  - 接続回線の利用効率が心配
    - うまく負荷分散できればいい
  - 実はfull route運用じゃなくてdefaultでもいいかも

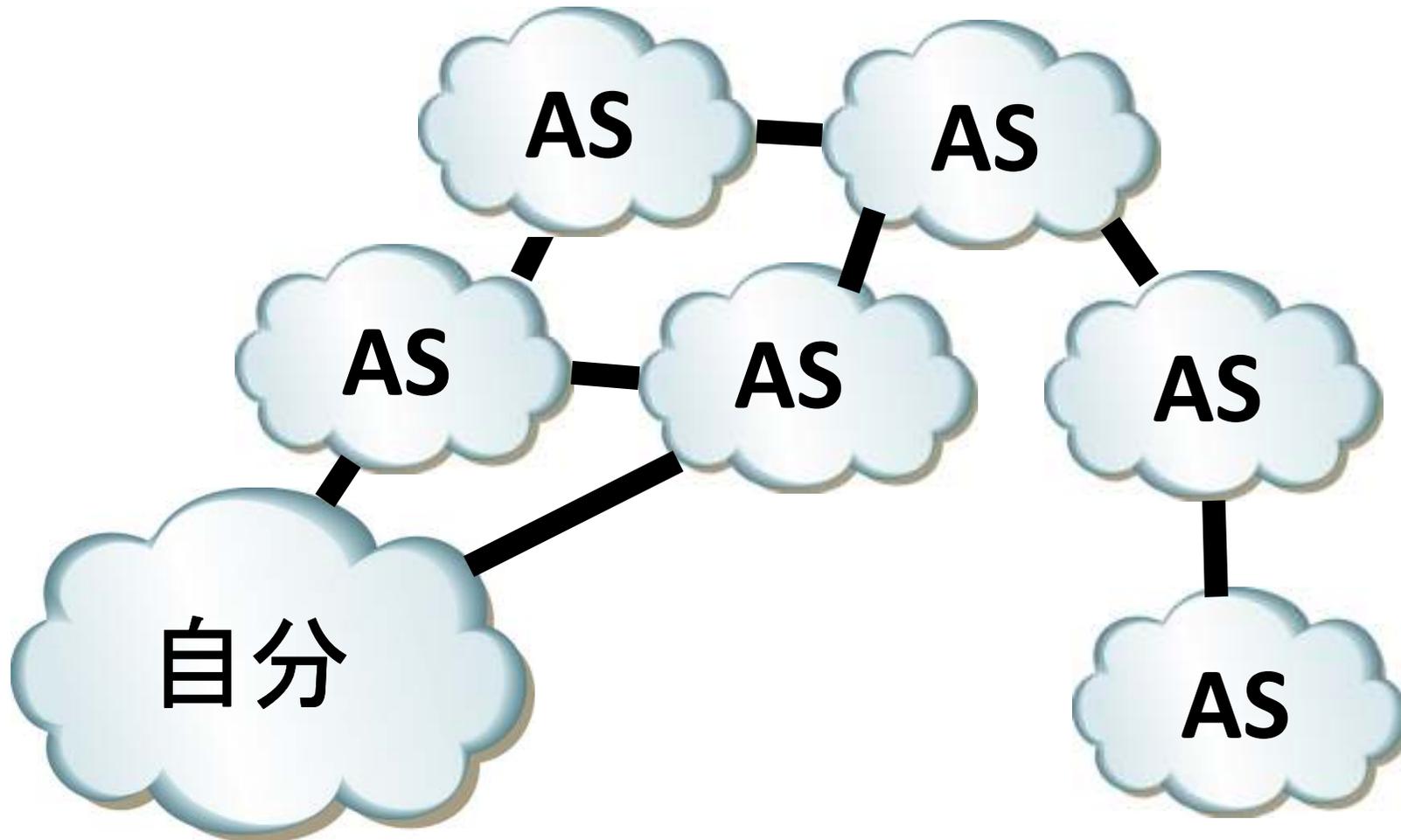
## 2. ASの集まりだよ派



## 2. ASの集まりだよ派

- ASがいっぱい繋がってるのがインターネット
  - ISP相互接続マップとか大好きだった
  - 通信の際は必要に応じて幾つかのASを経由するけど、遠いところはやっぱり闇の中
  - AS hop数が近ければ良いと思う
- 主なトラヒック要因となるAS/prefixをチェック
  - 頑張って制御
  - netflowは便利！

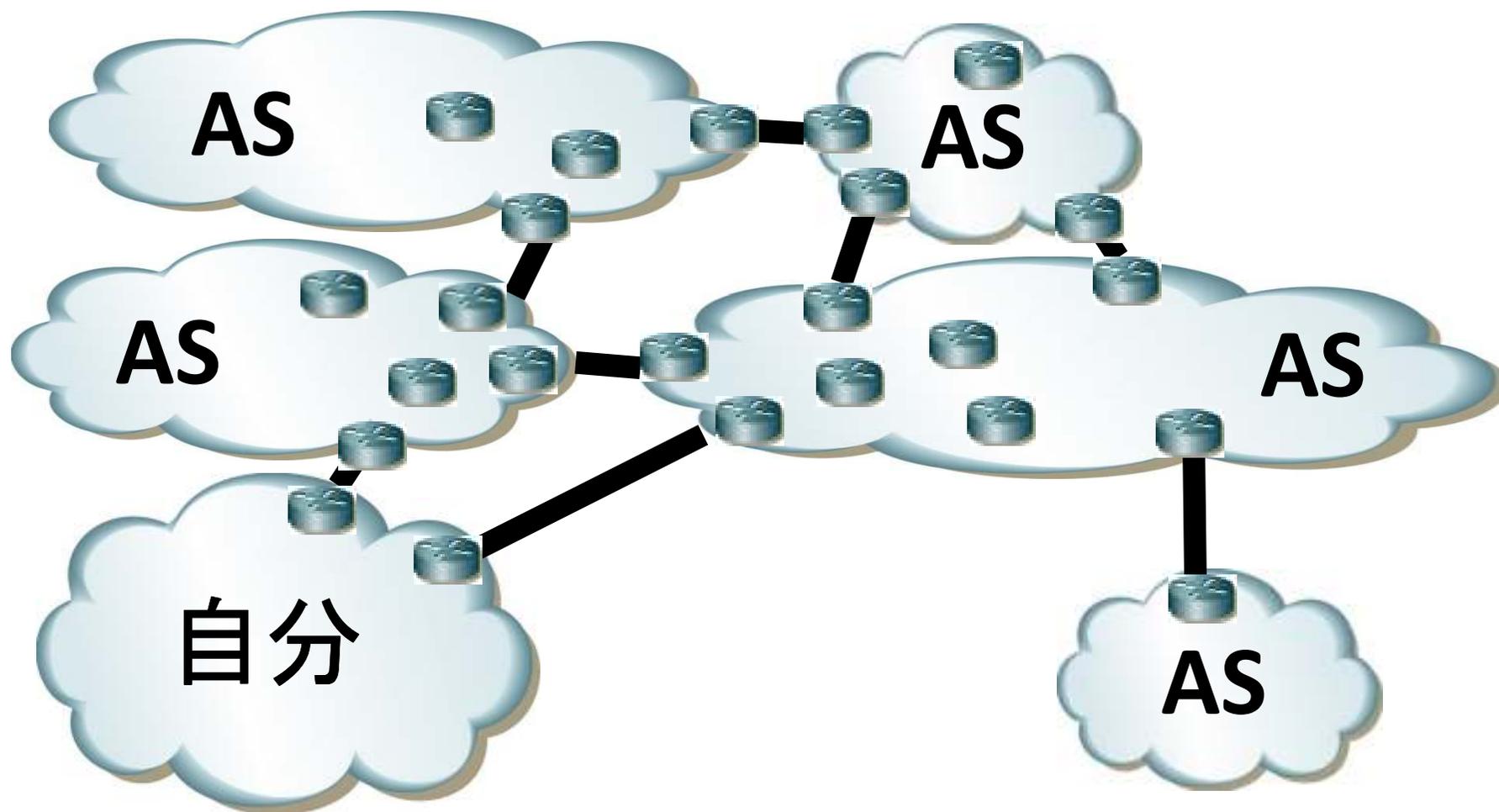
### 3.構造派



# 3.構造派

- ASが階層構造を持ちながら繋がってるよ
  - 上流、ピア、カスタマ
  - あるASのM単価と流量を考えるようになる
- AS毎の特徴をチェック
  - AS Pathを見て顧客関係がわかる
  - このASがtransitを提供するはずがないから、この経路は間違ってるはずとか言い出す
  - 上流の選び方などが気になる

## 4. 経路制御の集合だよ派



## 4. 経路制御の集合だよ派

- 各ルータにポリシーが実装されて動いてるよ
  - AS内部は単一ポリシーではないかも
    - ASの中でも地域、場所によって制御が違う
    - regional peerとかconfederationとか
- みんなの制御ポリシーを推測する
  - looking glass/netflow/traceroute等を駆使して、それぞれの経路制御ポリシーを推測する
  - 流れるはずのないtrafficからポリシーの実装ミスを見つける

# 詳細なところまで考える時

- 細かいことまで気にしなきゃいけない場合
  - 経路障害、意図しない経路状態
  - 設計時、構成変更時
- 困ったときに状況を知れる手段は僕らの保険
  - いざって時に身を助けるし品質も良くなる。
  - わかんない時はjanog@とかで聞いてもいい

# 難しい経路制御は不幸を招く

- 状況を分かりにくくする
- 問題発生時に取れる手段が限られる
- 構成変更時に失敗する
  - 忘れる、間違う、バグを踏む

# なぜシンプルなのがいいのか

- ポリシが素直な方が見通しが良い
  - トラブル時、設計時に間違わない
  - 新人教育、引き継ぎが楽
- 変更、変化に無理なく対応できる
  - IPv6, 4octet AS
  - 使用機種の変更

# お勧めしない経路制御の例

- トラブルの原因になりうる
  - source address based routing
- 世界に影響がある
  - more specific route
  - multiple origin
  - punching hole

# source address based routing

- かなり嫌。ふつつ採用しない
- そういう手段を思いつく人がいるみたい
- 世の中はdestination address basedなので、導入には相当注意が必要
  - routing loopなどを容易に引き起こす
  - 閉域網などで、もうどうしようもなく代替手段がない場合に、駄目だ駄目だ、こんなんやりたくないんだけど他に手段が無いと泣きながら使うぐらい

# more specific route

- お勧めしない
- 主にトラヒックの吸い込み目的で広報
  - トラヒック制御
  - 経路hijackへの対抗
- 経路数の増加を引き起こす
  - 世界中のルータのメモリを消費
  - トラヒック制御目的に利用し始めると、なかなかやめられなくなっちゃう
  - 必要な帯域を必要なところに確保した方が素直

# multiple origin

- お勧めしない
- 複数ASでprefixを生成して広報
  - bgpを運用せずにお手軽にマルチホーム
- 経路hijackと見分けがつきにくい
  - IRRの登録情報でも権威ある広報元がわかりにくい
- 回線を冗長化したいなら、同じASに複数回線もった方が高速に迂回できる
- 回線を目的別に使いたいとか、あるASに依存しない冗長性が必要ならASを取った方が素直

# punching hole

- かなりお勧めしない
- リナンバ無しでISPを変更する
- multiple originとmore specific routeの悪いところを併せもつ
  - 経路hijackと見分けがつかない
  - 経路数の増加を引き起す
- 今後、IPv4の新規割り当てが困難になった際に増加するかもしれないので、注意が必要

# ここまでのまとめ

- 今後、経路数は増え、制御は複雑になる
  - IPv4枯渇
  - トラヒック制御
  - 経路削減するISPの出現
- 人によって困るポイントが違うかも
  - トラヒック制御で困る人
  - 経路増加で困る人
  - 到達性で困る人

# IPv6 routing

- ちゃんと動く製品は出てきた
  - 高い箱、最近の箱だと性能も良い
  - まだちょっと足りない機能はあるけど
- まだまだ相互接続が足りません
  - ISP間
  - ccTLD DNS
  - root-servers.net ?!

おわり