

# Large BGP Communities

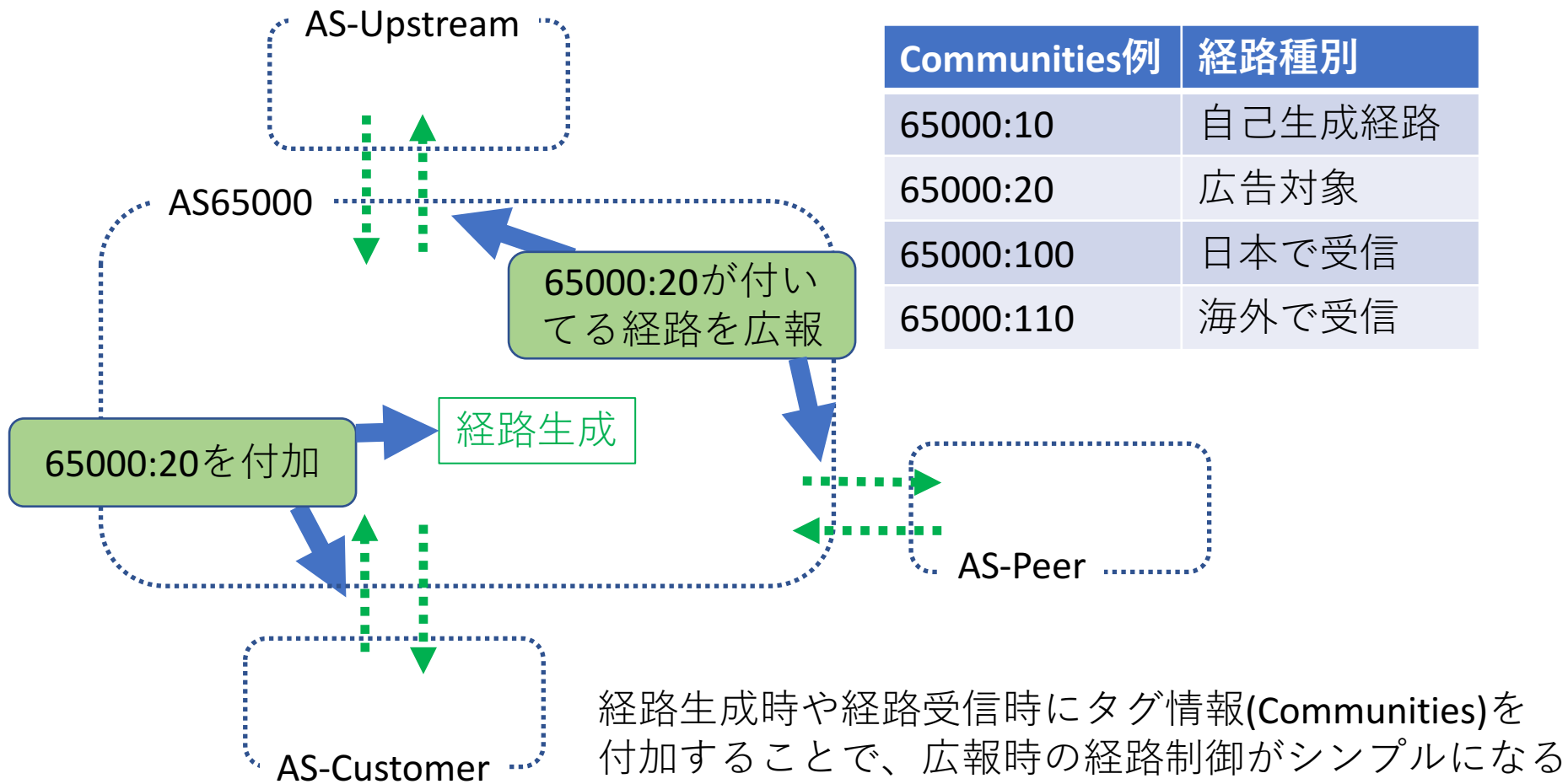
Matsuzaki 'maz' Yoshinobu

<maz@ij.ad.jp>

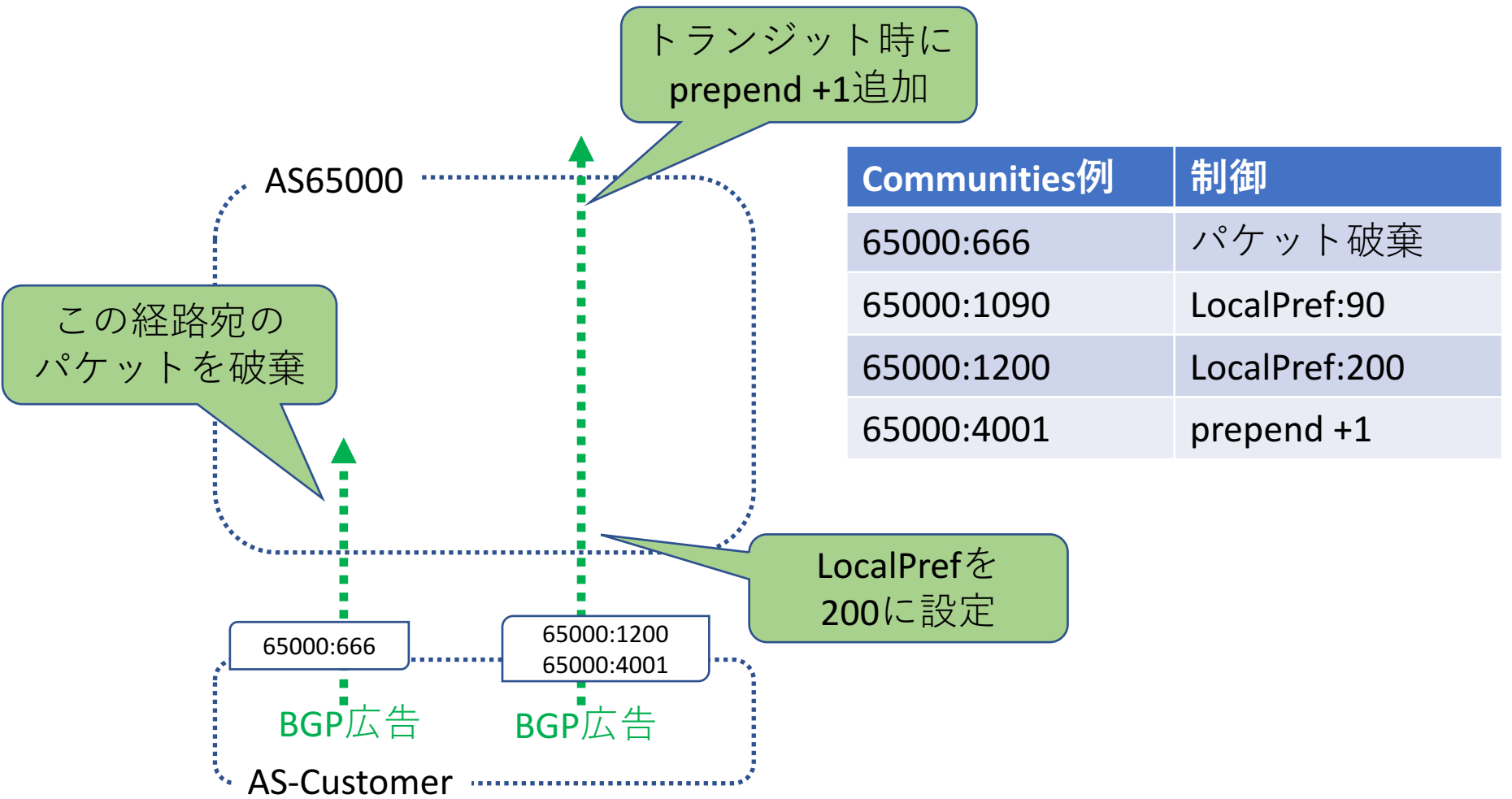
# BGP Communities [RFC1997]

- 経路広告に付加するタグみたいな属性値
  - 複数付加できる
  - 経路制御目的で広く利用されている
- **32bit長の数値で”16bit:16bit”な10進表記が主流**
  - 運用上、各ASでは最初の**16bit**に自AS番号を入れ、後ろの**16bit**を各種制御用に勝手に割り当てて利用
    - 例) 2497:1000 2497:3000 2497:110
  - **NO\_EXPORT (0xFFFFF01)**など、予め規定されているコミュニティもあるよ

# BGP Communities運用(内部例)



# BGP Communities運用(提供例)



# BGP Communitiesの問題

- 世の中、32bit AS番号な時代
  - 32bit AS番号は2007年ぐらいから分配開始
  - 現在では普通に割り当てされている状態
- 32bit AS番号が16bitフィールドに

入り  
きらない

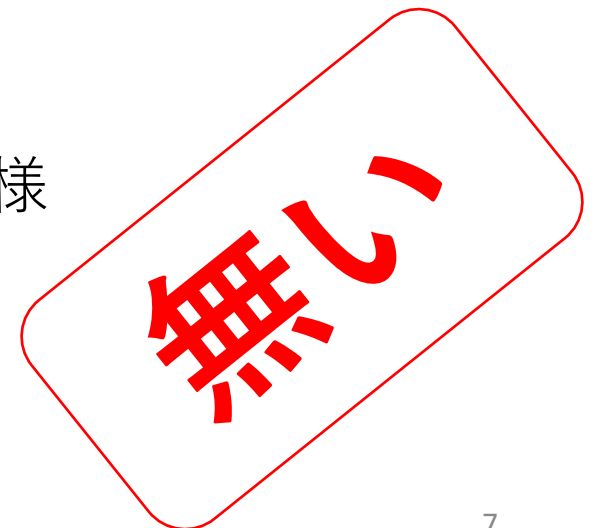
# BGP Communitiesの問題

問題の場合分け

- 他ASから既に提供されている制御を使うのはほぼ問題なし
  - 制御機能の提供ASや制御対象として興味のあるのが16bit ASなので、現状そんなに困っていない
- 自AS内のみの制御であればそんなに困らない
  - プライベートAS番号(64512-65534)で頑張るとか
  - それでも名前衝突する可能性はある
- 32bit AS番号なネットワークが、他ASにBGP Communitiesによる制御を提供する場合に困る
  - 何やっても名前衝突する可能性がある

# さて、使えそうな手段

- BGP Extended Communities [RFC4360]
  - L3VPNなどで利用されている属性値
  - RFC5668で32bit AS番号が使えるようになった
  - 複雑怪奇で主にVPN対応ルータでサポート
- I-D提案
  - draft-lange-flexible-bgp-communities
  - draft-ietf-idr-wide-bgp-communities
  - それぞれ色々な定義がある賢い仕様
  - コンセンサスには至らず



# Large BGP Communities



- 標準化作業中
  - [draft-ietf-idr-large-community](#)
  - IESG承認プロセスも通過したため、数ヶ月以内にRFCとして発行される予定
- 96bit長の数値
  - 単に大きくなったBGP Communities
  - 利用方法もほぼ一緒
- “32bit:32bit:32bit”の10進表記
  - <自AS番号>:<制御>:<対象AS番号>な利用を想定





# 2914での実装例

RFC 1997 (Current)	Large BGP Communities	Action
65400:peer-as	2914:65400:peer-as	Do not Advertise to peer-as in North America (NTT)
0:peer-as	6667:0:peer-as	Do not Announce to Route Server peer-as (AMS-IX)
65520:nnn	2914:65520:nnn	Lower Local Preference in Country nnn (NTT)
2914:410	2914:400:10	Route Received From a Peering Partner (NTT)
2914:420	2914:400:20	Route Received From a Customer (NTT)

# 既にテスト広報中

- Beacon Prefixes
  - 192.147.168.0/24
  - 2001:67c:208c::/48
- Large BGP Community: 15562:1:1
- みんなの運用しているルータでも見えます！
  - ただし、新しいPATH属性値なので・・・

# 手元でこんな感じ

```
IOS-XE15.4> >show ip bgp 192.147.168.0/24
```

```
:
```

```
unknown transitive attribute: flag 0xE0 type 0x20 length 0xC  
value 0000 3CCA 0000 0001 0000 0001
```

```
JUNOS14.2> show route 192.147.168.0/24 detail
```

```
:
```

```
Unrecognized Attributes: 15 bytes  
Attr flags e0 code 20: 00 00 3c ca 00 00 00 01 00 00 00 01
```

# looking glass でも

```
$ telnet route-views.routeviews.org
```

```
:
```

```
route-views>show ip bgp 192.147.168.0/24
```

```
:
```

```
Refresh Epoch 1
```

```
2497 2914 15562
```

```
202.232.0.2 from 202.232.0.2 (202.232.0.2)
```

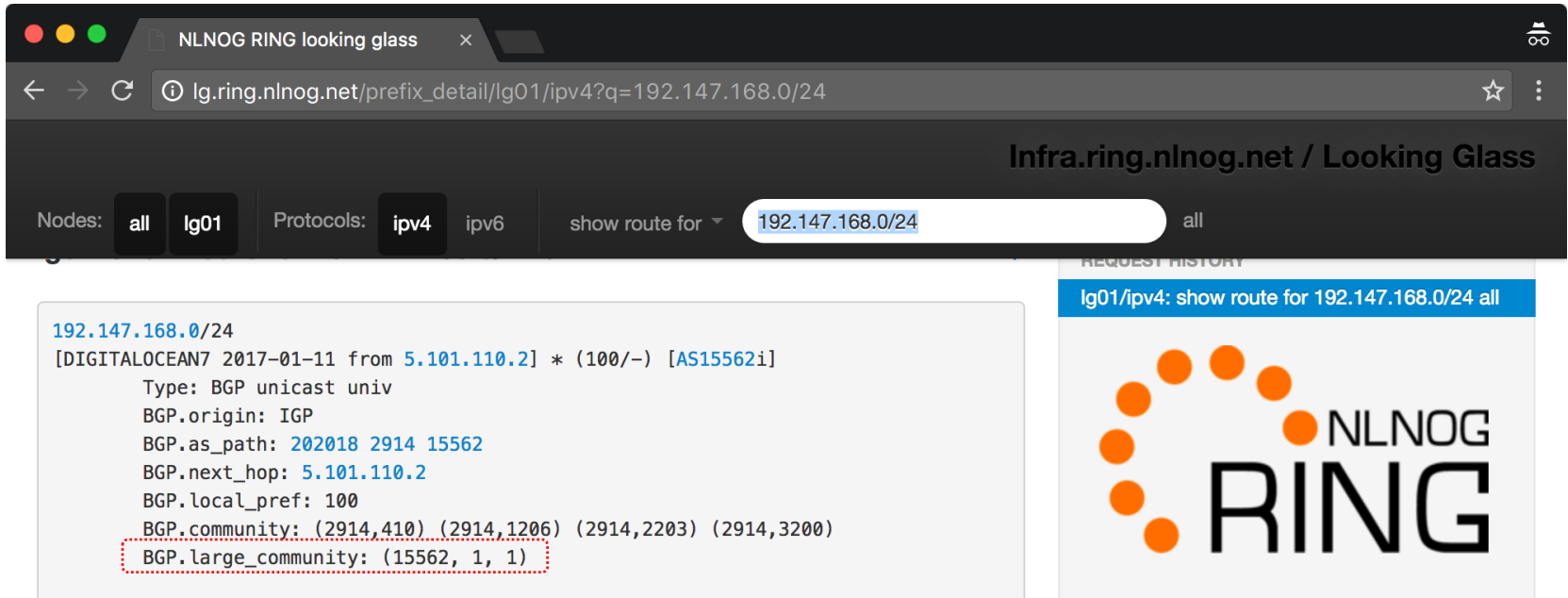
```
Origin IGP, localpref 100, valid, external
```

```
unknown transitive attribute: flag 0xE0 type 0x20 length 0xC
```

```
value 0000 3CCA 0000 0001 0000 0001
```

```
rx pathid: 0, tx pathid: 0
```

# Web UIなlooking glassでも



The screenshot shows a web browser window with the URL `lg.ring.nlnog.net/prefix_detail/lg01/ipv4?q=192.147.168.0/24`. The page title is "Infra.ring.nlnog.net / Looking Glass". The interface includes a search bar with the IP address `192.147.168.0/24` and buttons for "Nodes: all lg01" and "Protocols: ipv4 ipv6". Below the search bar, the BGP route details for `192.147.168.0/24` are displayed:

```
192.147.168.0/24
[DIGITALOCEAN7 2017-01-11 from 5.101.110.2] * (100/-) [AS15562i]
Type: BGP unicast univ
BGP.origin: IGP
BGP.as_path: 202018 2914 15562
BGP.next_hop: 5.101.110.2
BGP.local_pref: 100
BGP.community: (2914,410) (2914,1206) (2914,2203) (2914,3200)
BGP.large_community: (15562, 1, 1)
```

The `BGP.large_community: (15562, 1, 1)` line is highlighted with a red dashed box. To the right, there is a "REQUEST HISTORY" section with a button for "lg01/ipv4: show route for 192.147.168.0/24 all" and the NLNOG RING logo.

**BGP.large\_community: (15562, 1, 1)**

# BGP実装

Implementation	Software	Status	Details
Arista	<a href="#">EOS</a>	Planned	Feature Requested BUG169446
Cisco	<a href="#">IOS XR</a>	✓ Done!	Engineering Release
cz.nic	<a href="#">BIRD</a>	✓ Done!	BIRD 1.6.3 ( <a href="#">commit</a> )
ExaBGP	<a href="#">ExaBGP</a>	✓ Done!	<a href="#">PR482</a>
Juniper	<a href="#">Junos OS</a>	Planned	Second Half 2017
MikroTik	<a href="#">RouterOS</a>	Won't Implement Until RFC	Feature Requested 2016090522001073
Nokia	<a href="#">SR OS</a>	Planned	Third Quarter 2017
OpenBSD	<a href="#">OpenBGPD</a>	✓ Done!	OpenBSD 6.1 ( <a href="#">commit</a> )
OSRG	<a href="#">GoBGP</a>	✓ Done!	<a href="#">PR1094</a>
rtbrick	<a href="#">Fullstack</a>	Planned	<a href="#">February 2017</a>
Quagga	<a href="#">Quagga</a>	✓ Done!	Patch Provided for 1.1.0 <a href="#">875</a>
Ubiquiti Networks	<a href="#">EdgeOS</a>	Planned	<a href="#">Internal Enhancement Requested</a>
VyOS	<a href="#">VyOS</a>	Requested	Feature Requested <a href="#">T143</a>

最新情報は <http://largebgpcommunities.net/implementations/> にて

# ツール他

Implementation	Software	Status	Details
DE-CIX	<a href="#">pbgpp</a>	✓ Done!	<a href="#">PR16</a>
FreeBSD	tcpdump	✓ Done!	<a href="#">PR213423</a>
Marco d'Itri	<a href="#">zebra-dump-parser</a>	✓ Done!	<a href="#">PR3</a>
OpenBSD	tcpdump	✓ Done!	OpenBSD 6.1 ( <a href="#">patch</a> )
pmacct.net	<a href="#">pmacct</a>	✓ Done!	<a href="#">PR61</a>
RIPE NCC	<a href="#">bgpdump</a>	Planned	<a href="#">Issue 41</a>
tcpdump.org	<a href="#">tcpdump</a>	✓ Done!	<a href="#">PR543</a> ( <a href="#">commit</a> )
Yoshiyuki Yamauchi	<a href="#">mrtparse</a>	✓ Done!	<a href="#">PR13</a>
Wireshark	<a href="#">Dissector</a>	✓ Done!	18172 ( <a href="#">patch</a> )

最新情報は <http://largebgpcommunities.net/implementations/> にて



# やっつく事

- BGP関連実装
  - Large BGP Community対応
- 運用者
  - ベンダへのサポート依頼
  - 実装や活用の検討
  - ツール群の更新
  - 教育

# ポリシ例

Large BGP Communities (例)	意味
<AS>:0:<peer-as>	peer-asに広報しない
<AS>:n:<peer-as>	peer-asにprepend+n
<AS>:100:<local-pref>	AS受信時にlocal-prefを指定値に
<AS>:666:0	blackhole
<AS>:100n:0	全隣接にprepend+n
<AS>:101n:0	全peerにprepend+n

nは1~9の整数

# ちょっと話してみたい

- BGP Communityって使ってますか？
- BGP Large Communityってどう思いますか？
- どんな制御やりたいですか？
  - 隣接ASとの経路制御
  - 数ホップ先での制御
  - blackhole
- 足りないモノがありますか？

# 僕の懸念

- より多くの制御
  - 以前より細かく制御可能
  - 一般にBGPの制御ってのは見えにくい
- 昔より複雑なトラブルは発生しそう
- 悪用された場合に、検出が難しそう

# 第三者検証できる状態が大事

- 過去の経路情報の記録
  - Looking GlassやRoute Server
  - MRT RIB dumpに知らない属性値も含めて欲しい
- 蓄積情報の公開
  - 著名なのはRoute ViewsとRIPE RIS
  - 国内ではあまり無い
- 経路調査研究の促進

# Thanks

- Job Snijders
- Greg Hankins
- 最新情報は <http://LargeBGPCommunities.net/>