

今更ながらRFC 5549で IX接続するぞ

unnumberedの話じゃないよ

株式会社ミクシィ 開発本部 吉野純平

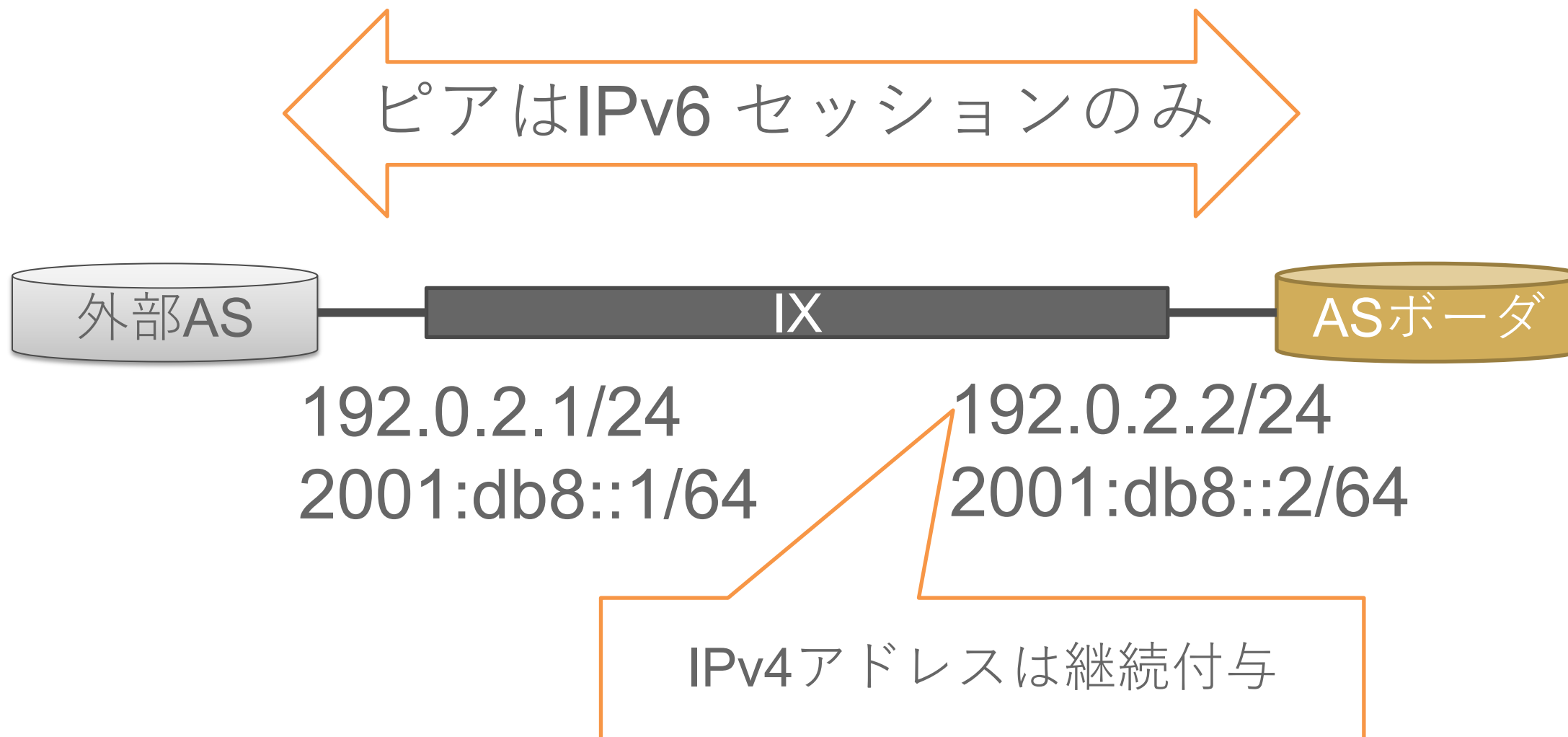
自己紹介

- 吉野純平 株式会社ミクシィ 開発本部インフラ室
- 最近
 - ネットワークに触る機会も減りましたorz
 - アナログ回路を勉強中
 - ローカル5G商用運用はじめました
 - 引き続きSMPTE 2110してます
 - 先週子供3号が生まれました

- IX利用者が増えていると思います
- IXのIPv4リナンバは発生していくだろう
 - 手間がかかりますね
 - 広くするにも限りもある
- IXのIPv6のリナンバは多分来ないですよね？
- IX上でのIPv4のピアを全廃しませんか？

みんなでやってみませんか？





○みなさまご自分で検証の上ご利用ください



ごめんなさい。
JANOG49までに
実網では、
できませんでした
た。



- RFC5549の話をします
 - (8950が出ましたが)
- Unnumberedの話をしません



過去にも色々なところで話されています

○ IRS26

- RFC5549-v4 NLRI with v6 NH -(ユースケースと実装状況) 土屋さん
- <http://irs.ietf.to/wiki.cgi?action=ATTACH&page=IRS26&file=rfc5549%2Dshtsuchi%2D01%2Epdf>

○ NANOG 56

- “RFC5549 Advertising IPv4 NLRI with an IPv6 Next Hop” Ariën Vijn
- <https://archive.nanog.org/meetings/nanog56/presentations/Monday/mon.general.vijn.20.pdf>

今日の発表のポイント

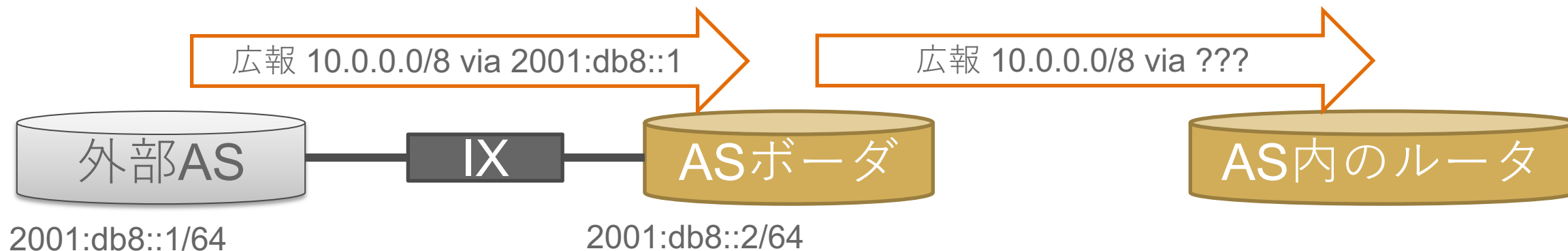


- 論点 1 : 対外接続でのポイント
- 論点 2 : AS内でのIXでの受信経路の広報に関するポイント

論点1：RFC5549に限らず検討

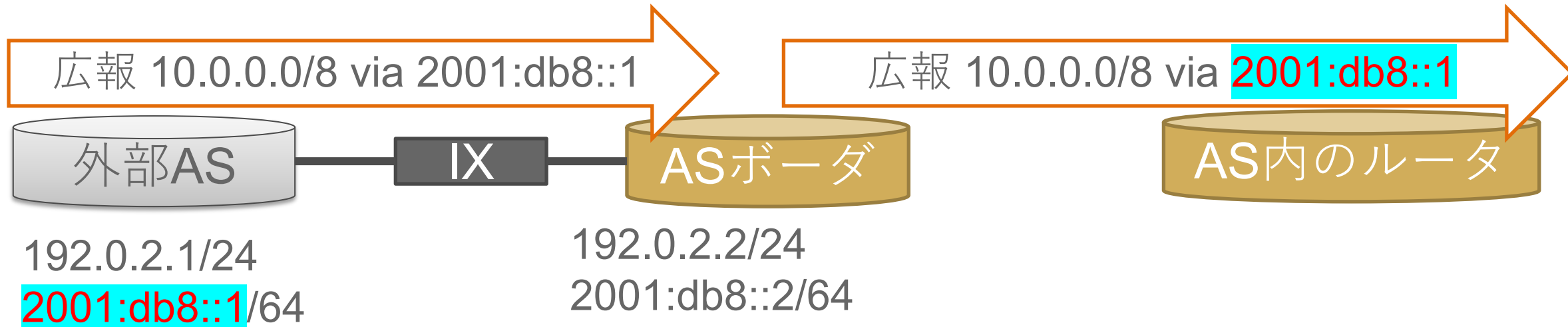
- BGPセッションをIPv6オンリーにする
- Extended Nexthop Capabilityを使う
 - BGP Session: IPv6
 - IPv4 NLRI with **IPv6 Nexthop**
 - IPv6 NLRI with IPv6 Nexthop
- Extended Nexthop Capabilityを使わない
 - BGP Session: IPv6
 - IPv4 NLRI with **IPv4 Nexthop**
 - IPv6 NLRI with IPv6 Nexthop

論点2：網内広報



- Extended Nexthop Capabilityを使う場合はリナンバケア不要
- 課題は自身の網内への広報方法
- iBGPでExtended Nexthop Capablityを使わないneighborへの広報
 - IPv4のNLRIはNexthopをIPv4にしないと広報できない
 - Nexthop Selfしている場合は考慮しなくていい
 - (たぶんみなさん使ってないですね?)

既存運用がIX経路をNexthop Selfしない場合の案1



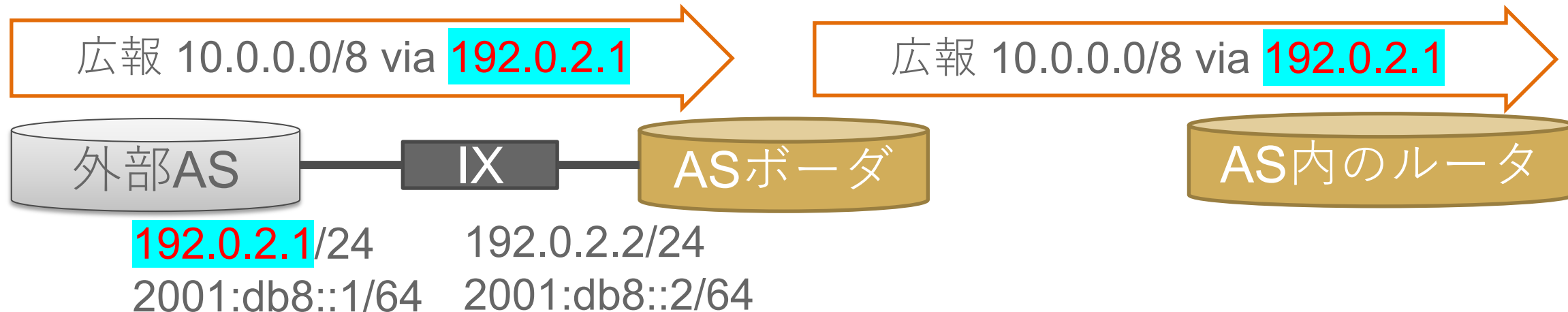
- IPv6でIXセグメントが網内になんらか広報されている場合
- iBGPにもExtended Nexthop Capablityを入れる
- 変更範囲が広いのが難点

既存運用がIX経路をNexthop Selfしない場合の案2



- NexthopをIXのIPv4セグメントに書き換えて広報する
- 一発でやれそうな1行の便利コンフィグがなさそう
- ラボでやってみた例
 - eBGPで受けるときにIXセグメントごとのcommunityを付与
 - iBGP広報時にnexthopにIXのIPv4セグメントのIPを書き換え
 - ちょっとわかりづらいなあというのが本音

既存運用がIX経路をNexthop Selfしない場合の案3

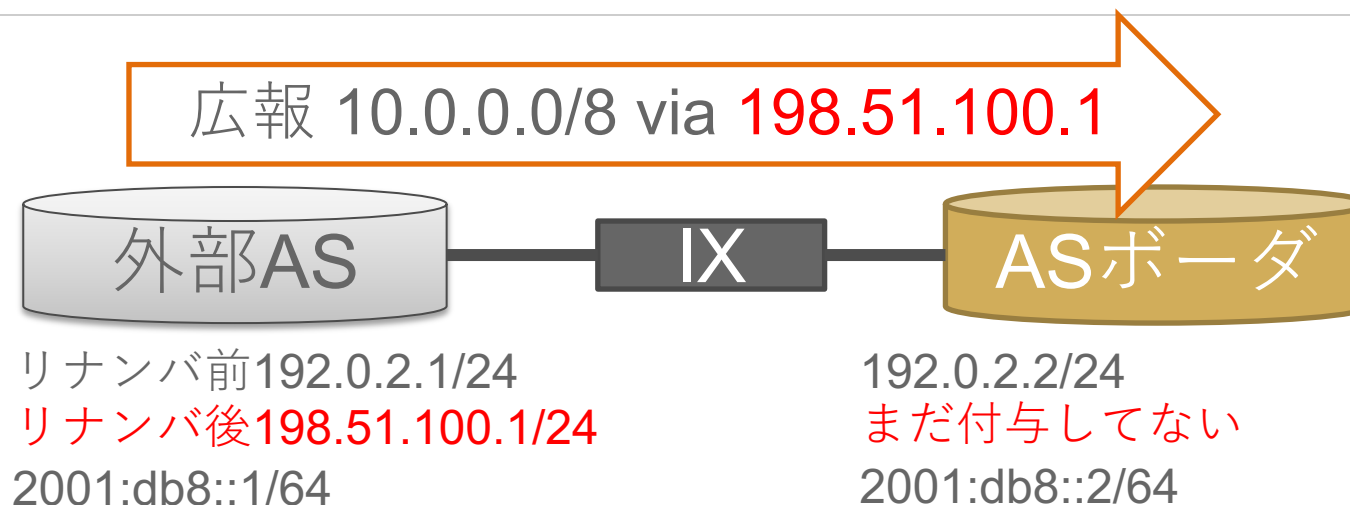


○Extended Nexthop Encoding活用をしないパターン

○IPv6のセッションでIPv4 NLRI with IPv4 Nexthopを送ることは可能

- Juniperなら `neighbor 2001:db8::2 family inet unicast local-ipv4-address 198.51.100.1`
- Aristaなら `neighbor 2001:db8::2 local-v4-addr 198.51.100.1`

Extended Nexthop Capabilityを活用しないデメリット



- 右側の機器でリナンバ後のアドレス付与前に広報すると・・・
 - 変なルートを回る？
 - invalidな経路になる？
- 機種によっては第3者のIPv4を広報することもできる??
 - ミスオペで悲しいことになる

ここから運用上必要な要素の確認

tracerouteは？

- IPv4が付与されていればそれが使われる
- 付与されていない場合、ループバックアドレスが使われる
 - 機種によるかも？

受信経路数制限はAFI/SAFIごとにできるか？

○ Juniper : できた

- set bgp group hoge family inet unicast accepted-prefix-limit xxxx
- set bgp group hoge family inet unicast extended-nexthop
- set bgp group hoge family inet6 unicast accepted-prefix-limit xxxx

○ Arista : できた

- neighbor 2001:db8::10 peer group hoge
- address-family ipv4
 - neighbor hoge activate
 - neighbor hoge next-hop address-family ipv6 originate
 - neighbor 2001:db8::10 maximum-accepted-routes 1000
- address-family ipv6
 - neighbor hoge activate
 - neighbor 2001:db8::10 maximum-accepted-routes 1000

経路フィルタは？

- v4とv6のルールを両方書く
- Juniperなら並べたり、Aristaならsub-route-mapを使ったり。
- Arista のサンプルイメージ
 - route-map hoge-export permit 10
 - match ip address prefix-list hoge-ipv4
 - route-map hoge-export permit 20
 - match ipv6 address prefix-list hoge-ipv6
 - BGP配下で, neighbor hoge route-map hoge-export out

経路数の監視は？

- AFI/SAFIごとに監視できそう
 - 経路数はMIB定義がありそう

フロー解析は？

○ 現在検証中

- nexthopの情報がでるところいけるか？
- あれ？ うまく行く気がしない . . .
- netflow v5 たぶんむり
- netflow v9 テンプレートが増えそうだけど動くかも
- sflow router拡張がv6 nexthop対応しているだろうか？
- 代わりにv4何かを埋めておいてもらえるといいのかなあ？
- 記録用に使うv4アドレスを通知できないかなあ？
- あれ、コレクタ側は . . .

まとめ

- Extended Nexthop Encodingは良さそう。flowは追試
- 一緒にやってみませんか？
- 将来的にネットワーク作るならこんな感じか？
 - BGPセッションはIPv6オンリー
 - IGPはOSPFv3かISIS？
 - SRを考えるならISIS？
 - IPv4はコントロールプレーンから消えて行く？

質疑応答

- 懸念あれば教えてください！
- IX事業者のみなさん
 - そもそもダメって方いらっしゃいます？？？
 - ルートサーバでサポートなどどうでしょう？？？
- IX利用者のみなさん
 - 対応相談可能ですか？
 - 特にNexthop Selfされてない方どうでしょう？
- 機材ベンダーのみなさん
 - 何か懸念ありますか？

○ sflow出力

- Aristaは出力できた。Gateway拡張部分に出る（事後確認）
- IPv4がなまじ残ると殺せなくて大変なのでIPv6オンリーの方が良いのでは？
- AFI/SAFIのどちらかでprefix limit超えるとどうなるか要確認
- RFC5549の解釈のはなし
 - 該当AFI/SAFIはExtended Nexthop Capabilityで指定されたものでしか送ってはないと解釈すべきか？



We are hiring!!