

Echigo-IXの経路制御と BGP no-export community

- ENOG47 Meeting -

2017/10/27

かねこやすゆき

<yasuyuki.kaneko@global-netcore.jp>

Echigo-IXとは

<http://www.echigo-ix.jp/>



Echigo-IX

[トップページ](#) [プロジェクト概要](#) [ニュース・お知らせ](#) [関連リンク](#)

Echigo-IX プロジェクト

メニュー

- [Echigo-IX について](#)
- [Echigo-IX の特徴](#)
- [Echigo-IX 参加組織](#)
- [Echigo-IX 活動履歴](#)
- [Echigo-IX 連絡先](#)

Echigo-IX プロジェクト

Echigo-IX は、新潟県内におけるインターネット通信品質の向上、インターネット利活用の促進を目的とした地域IX (Internet eXchange) プロジェクトです。

Echigo-IX は、[ENOG \(Echigo Network Operators' Group\)](#) の活動の一環として運営しており、この活動を通じて、新潟県における地域コミュニティの活性化とビジネス創出に寄与することを目指します。



Echigo-IXとは

- コミュニティ運営型 地域IX
- 現在は6つのASが接続



AS9601



AS18070



AS23631



AS23641



AS38652



AS131913

- Akamai CDNキャッシュの共同利用も

コミュニティ運営型とは

- みんなの「持ち寄り」で構築・運用

サーバ提供するよ！

IPアドレス割り当て
ます！



Echigo-IX

AS番号は当社負担
で！

ルータなら任せろ！

DCは当社でどうぞ！
スイッチも出すよ！

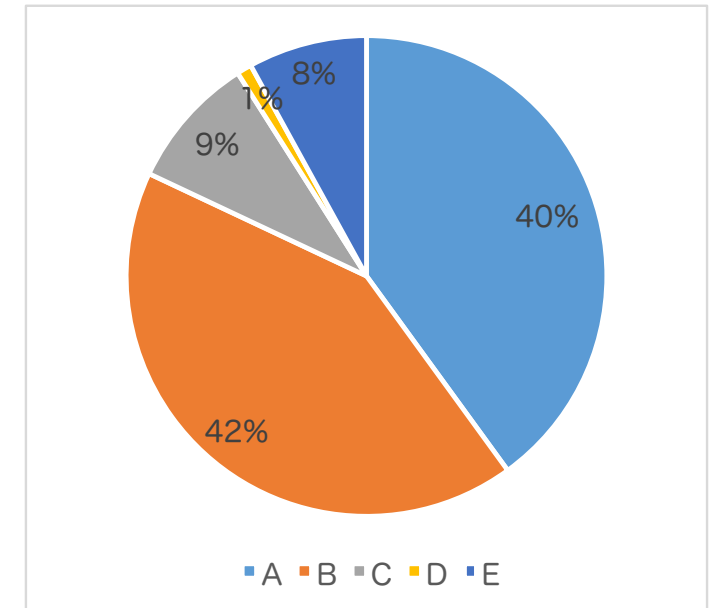
費用負担ルール

- 「持ち寄り」 設備の維持経費は各々負担
- 「運用」 はみんなでやる
- 「トランジット」 は流量見合いで各社提供
- 「現ナマ」 のやり取りはしない

トランジット提供のルール

- 各ASの流量占有率を計測し、一年間の負担月数をざっくり決定

• AS A =	40%	→	5ヶ月
• AS B =	42%	→	5ヶ月
• AS C =	9%	→	1ヶ月
• AS D =	1%	→	0ヶ月
• AS E =	8%	→	1ヶ月



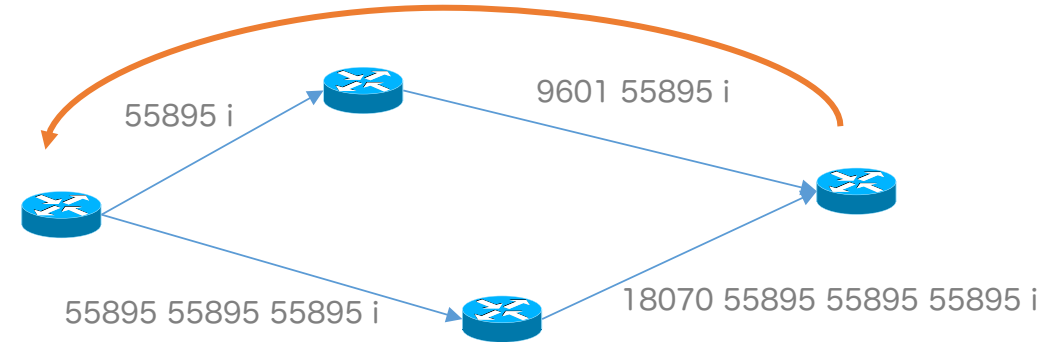
というわけで

- 定期的にトランジット提供元を変更するオペレーションが必要！
- このオペレーションはIX側の設定変更のみで実現したい
 - 各社のルータ設定はいちいちいじりたくない

従来のやり方は

- AS Path Prepend

- 各ASは常にフルルートをIXに提供する
- IXルータは当月のトランジットAS以外のASに対しては、prependした経路を広告する
- 結果、トランジットAS経由が最短経路として選択されることを期待する



課題

- 完全に制御できない
- というか、むしろ全然効かない?!?!

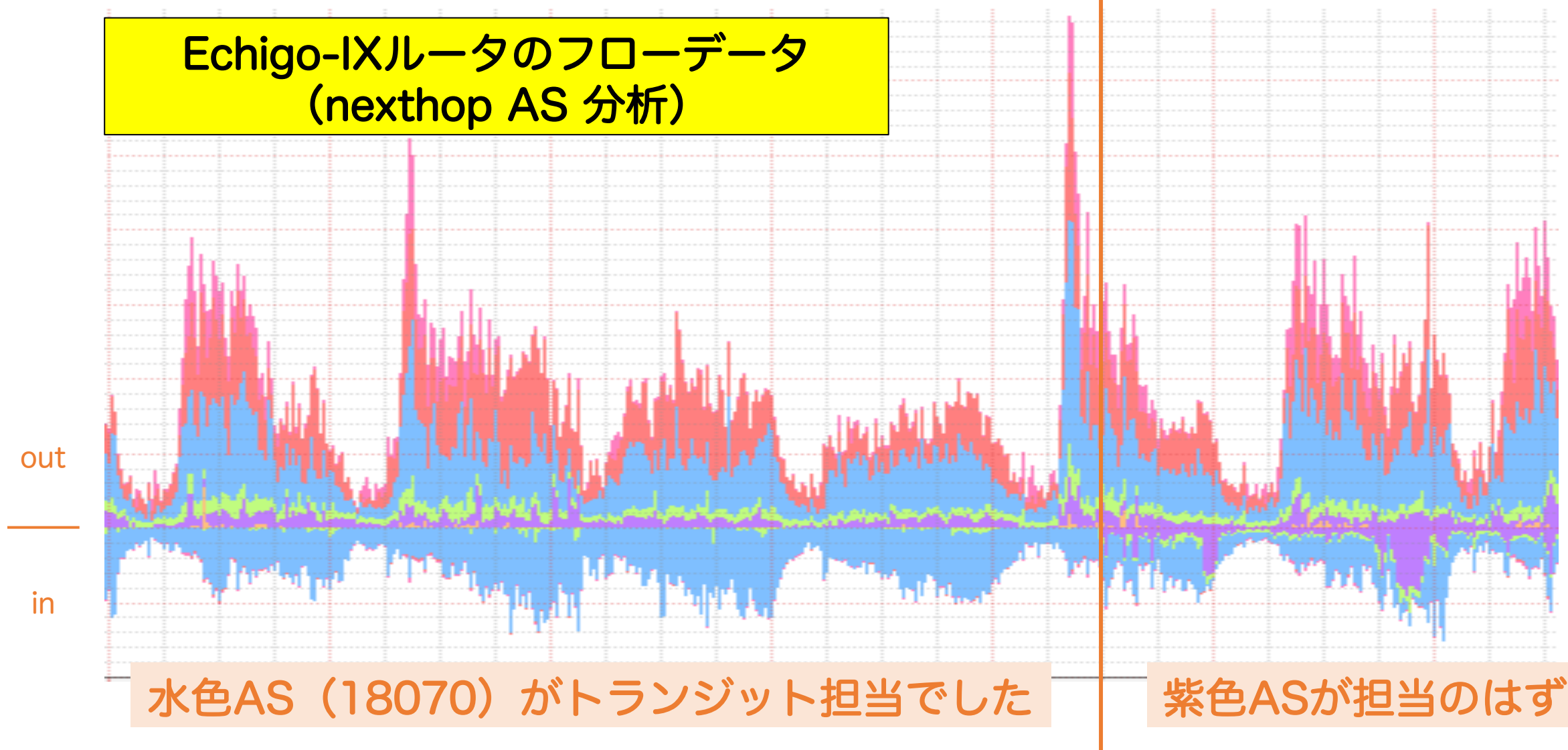
- なぜ



実際のトラフィック状況

設定変更

Echigo-IXルータのフローデータ
(nexthop AS 分析)



どこから流入している？

AS18070ルータのフローデータ
(Origin AS 分析)

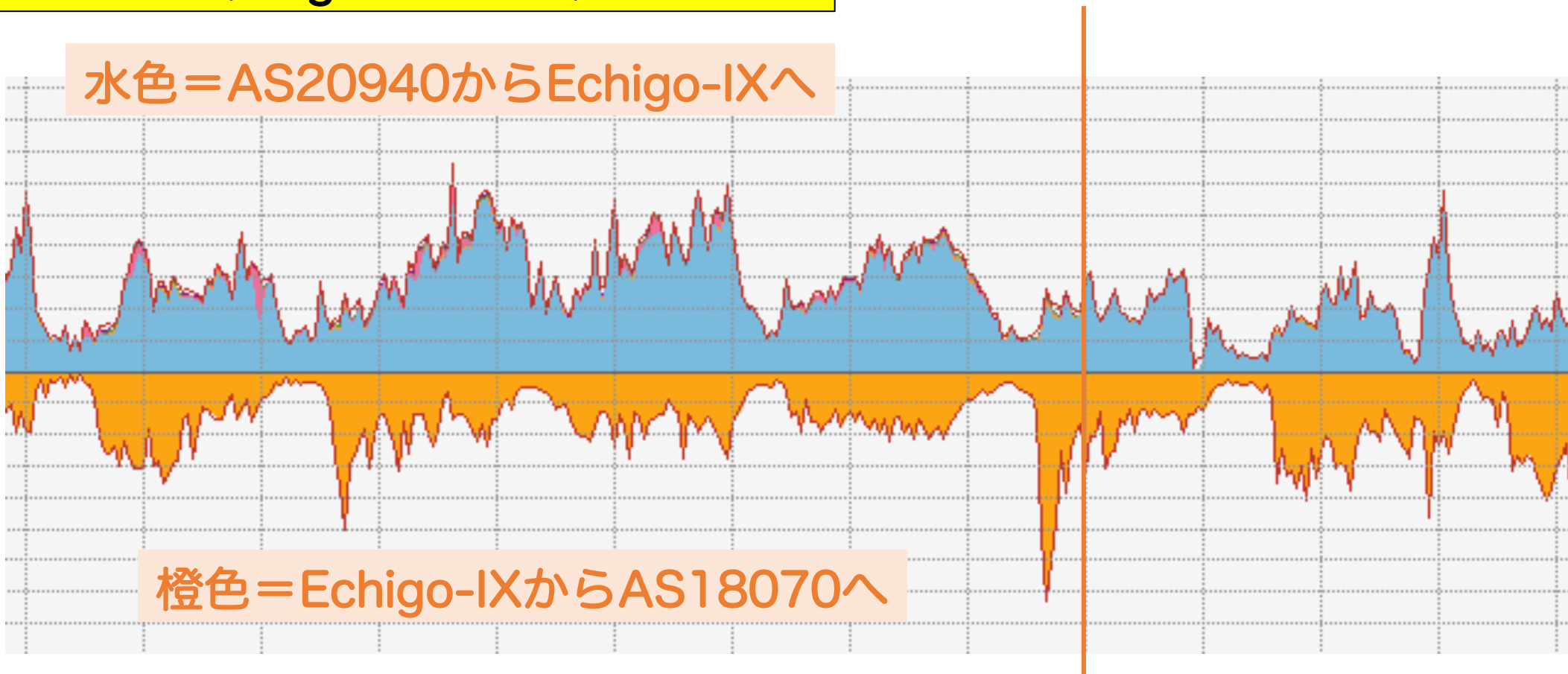
設定変更

水色 = AS20940からEchigo-IXへ

out

in

橙色 = Echigo-IXからAS18070へ



ここから見えること

- Echigo-IXのinboundトラフィックのほとんどは、AS20940からの通信っぽい
 - AS20940 = Akamai International B.V.
- AS Path Prependの効果は、極めて限定的である

ナカノヒトに直接聞いてみた



金子 康行

ホーム 20+



AS55895 (Echigo-IX) のトランジットを調整しているんですが、AS20940からAS55895へ流れる通信が、AS18070経由になってるのが直せなくて、、、そちらから見てどういう風に見えてますかね？

path prependでその月優先するAS以外は遠ざけてるんですけど

それを今日やったんですが、どうにもAS18070経由の通信がなくなってますね・・・（汗

ナカノヒトに直接聞いてみた



金子 康行

ホーム 20+



うち、prepend は苦手なんです...



経路の有無しか実際見ていないので



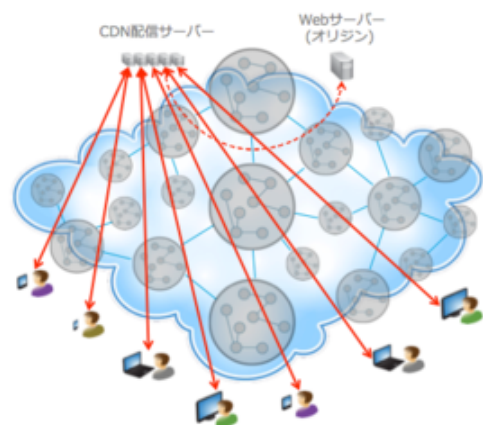
—人人人人人人人人人人人人人人—
> 経路の有無しか実際見ていない <
—Y^Y^Y^Y^Y^Y^Y^Y^Y^Y^Y^Y^Y^Y^Y^Y—



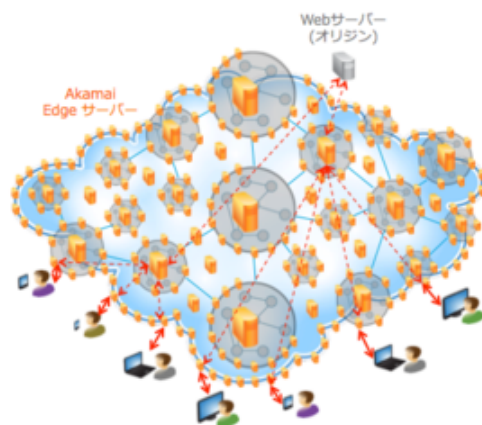
Akamaiの仕組み

『超分散アーキテクチャ』

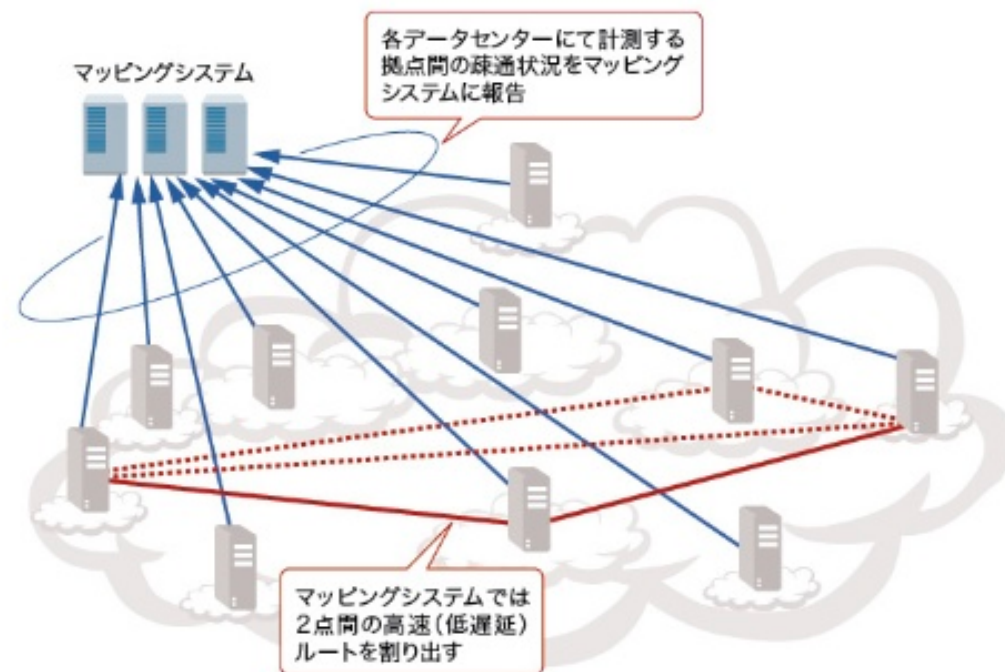
一般的なCDN
(集中型/限定分散型アーキテクチャ)



Akamai Intelligent Platform
(超分散型アーキテクチャ)



『ダイナミックマッピング』



想像するに

- 独自の仕組みで品質を計測している
- 経路情報ではなく計測に基づく最適選択
- AS Path長では制御不能！
- 経路そのものをなくすしかない



やりたいこと、おさらい

- 対外的には経路広告をなくしたい
- 直結するIXメンバーには経路を広告したい
- IXルータの設定変更だけで管理実現したい
- で、結論
 - BGP no-export communityの利用

そもそも、BGP communityとは

- 経路に対する 「タグづけ」 を行うもの
 - Community属性それ自体はBGPの最適経路選択に関与しない
 - Communityは経路(prefix)に対する付加的な情報として存在
 - 受信側が属性の変更やフィルタ設定などにより経路制御を行う判断条件としてCommunityを利用することができる

RFC1997での定義

Well-known Communities

The following communities have global significance and their operations shall be implemented in any community-attribute-aware BGP speaker.

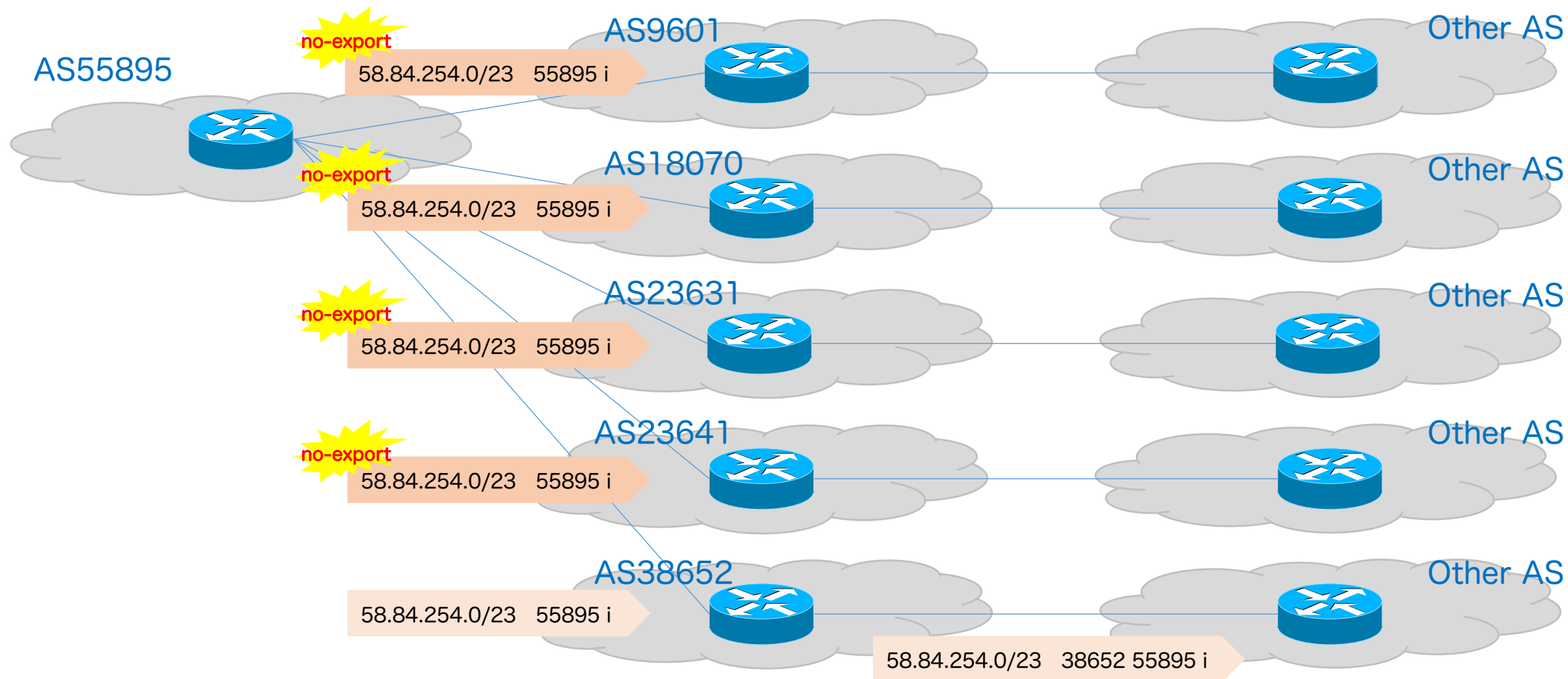
NO_EXPORT (0xFFFFFFFF01)

All routes received carrying a communities attribute containing this value MUST NOT be advertised outside a BGP confederation boundary (a stand-alone autonomous system that is not part of a confederation should be considered a confederation itself).

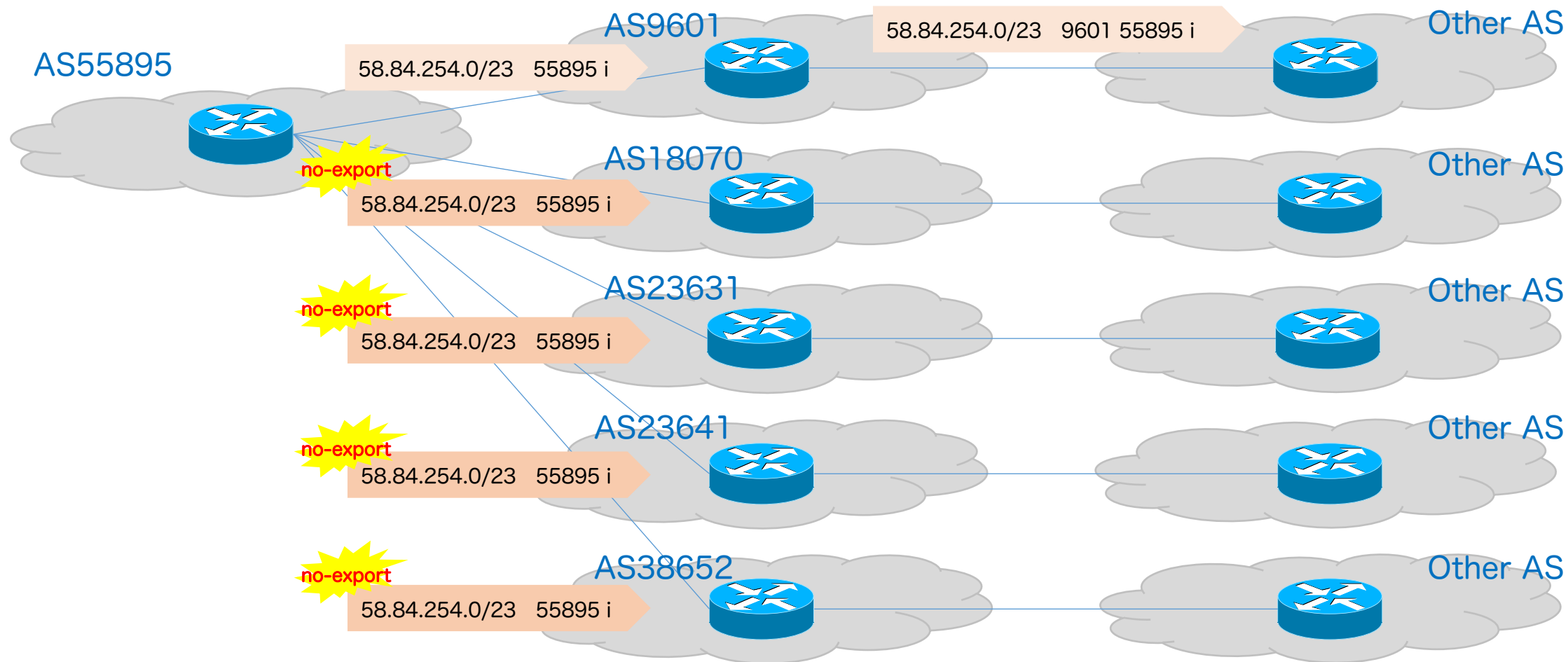
no-export communityとは

- 「他ASには経路情報を広告しないでね」
- という意図を持つ、定義済みコミュニティ

実現イメージ



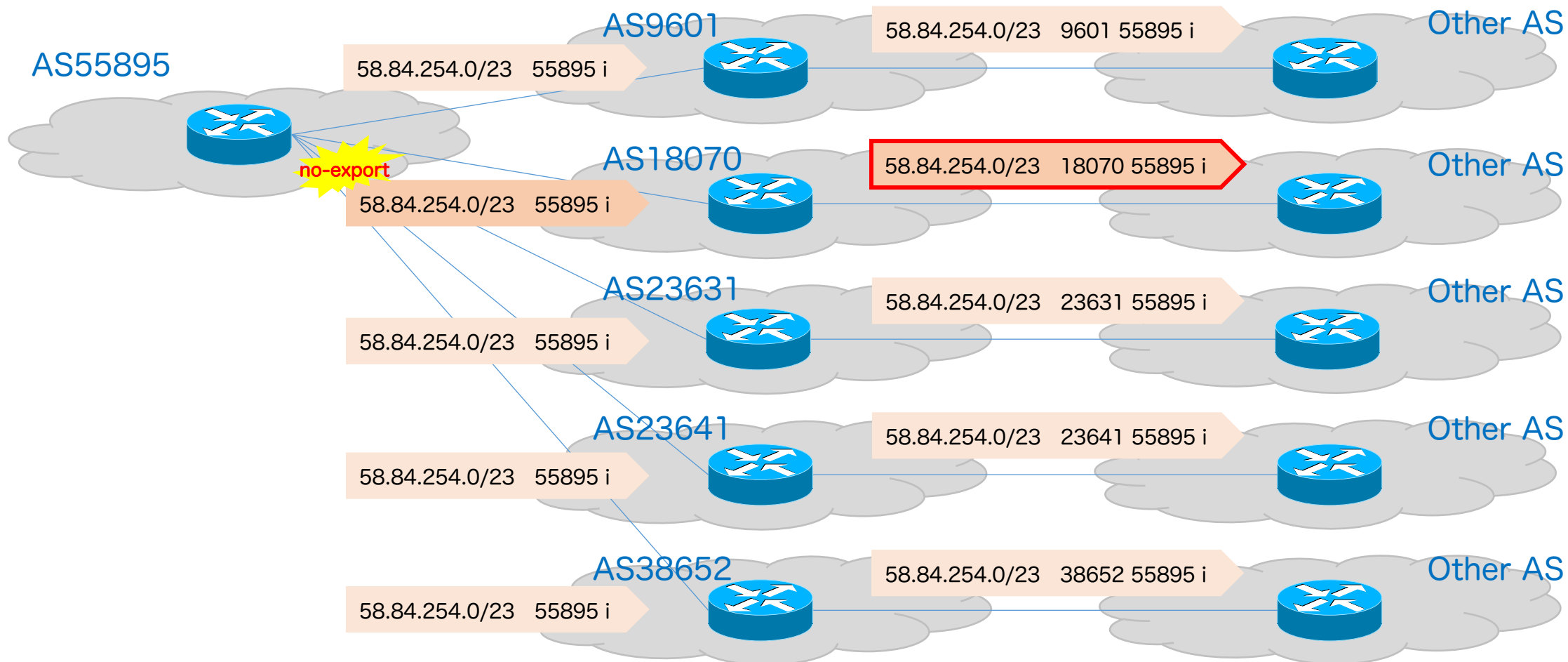
実現イメージ



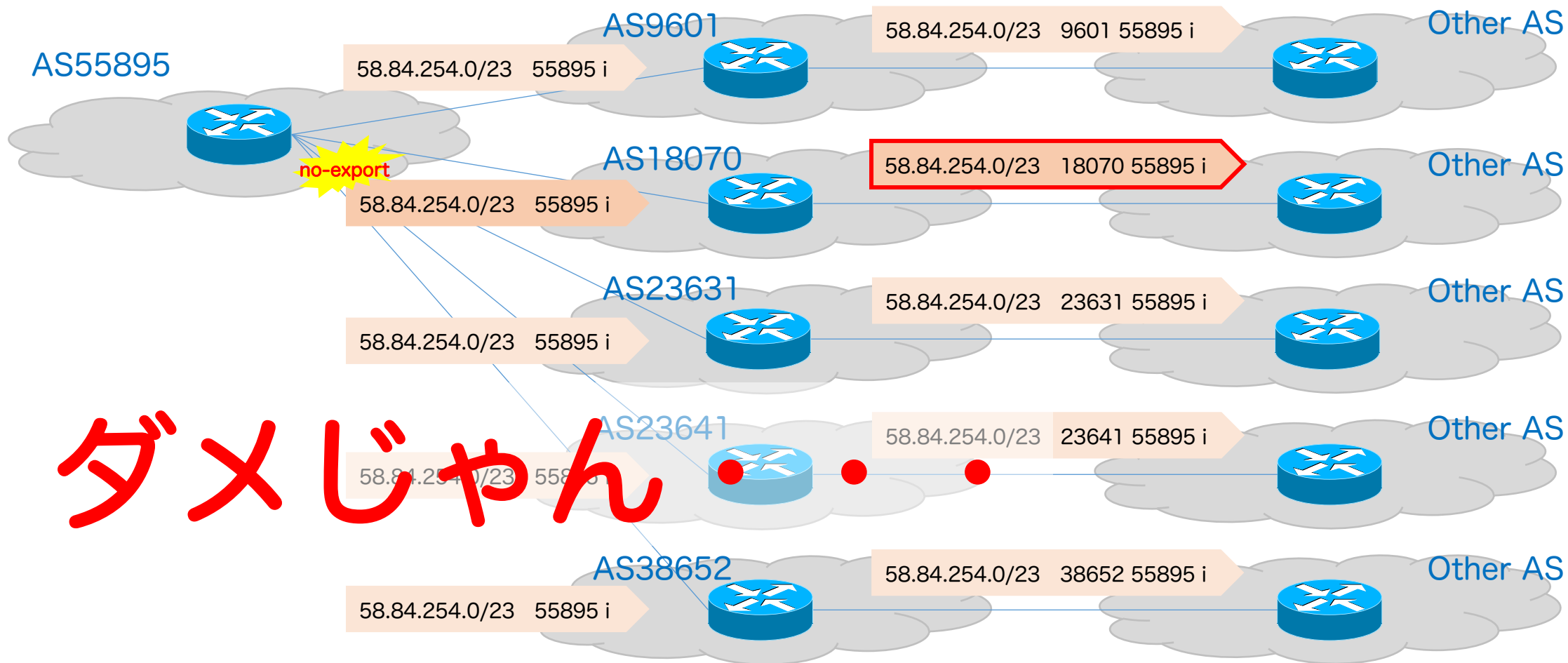
さっそくテストしてみた

- とりあえず、AS18070向けにだけ no-export をつけてみたよ

AS18070向けでテストしてみた

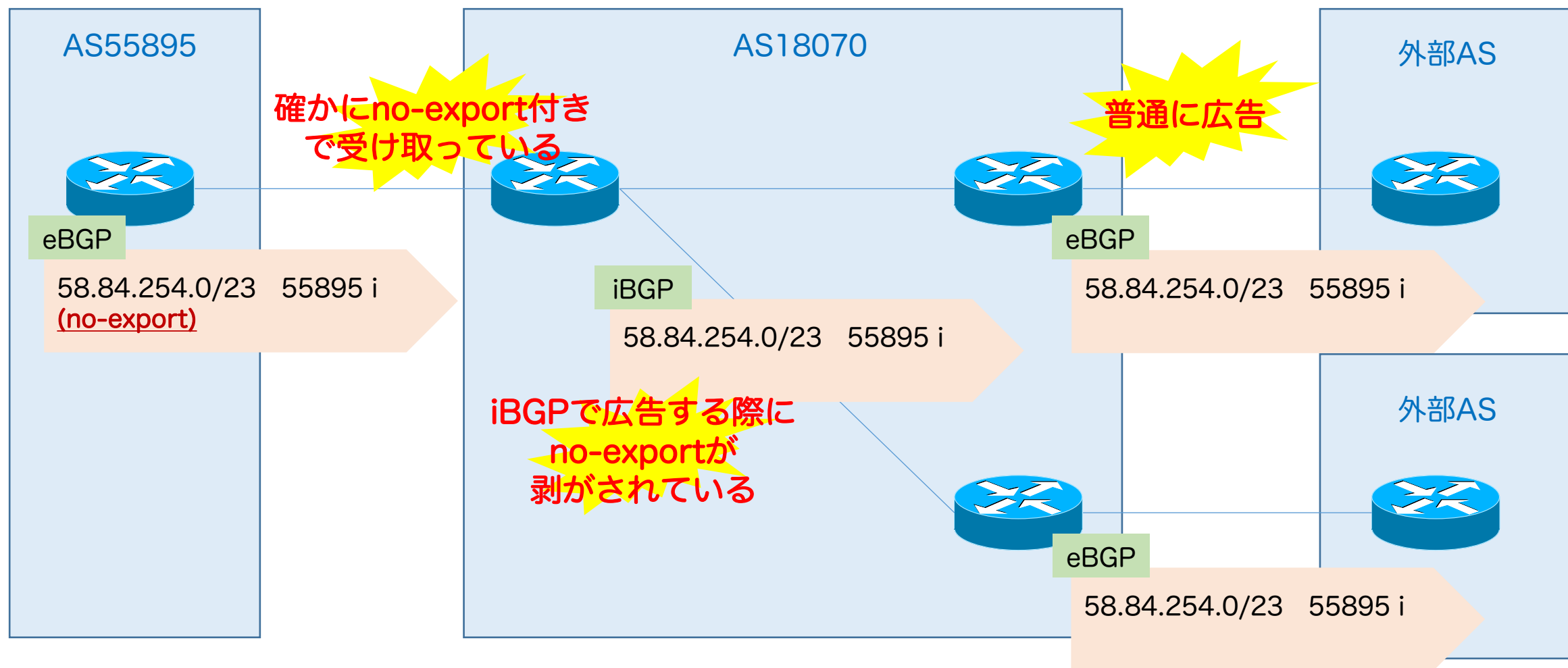


AS18070向けでテストしてみた



ダメじゃん

ダメだった原因を探ると

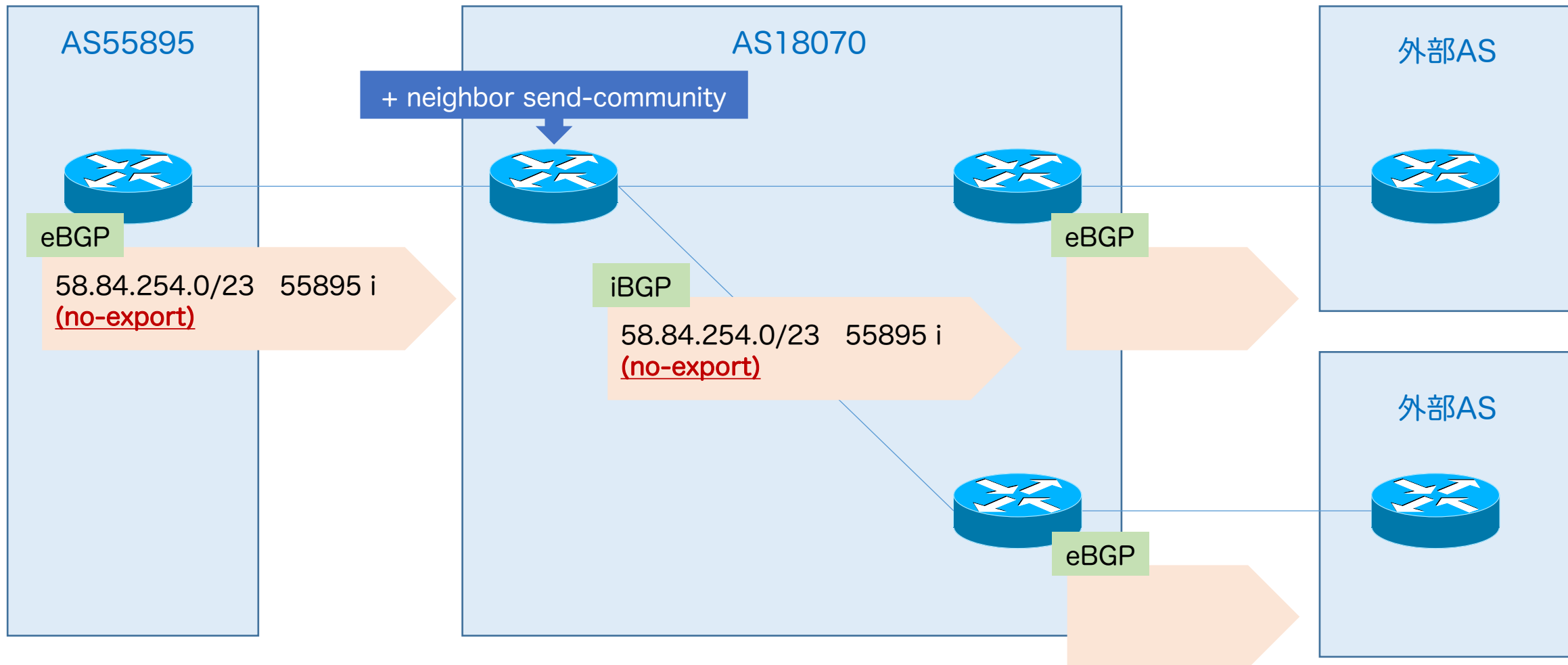


理由は簡単

- Cisco IOSでは以下の設定が必要
 - neighbor *ip-address* send-community
- その設定入れてなかったわ・・・



うまくいきました。。。。



IXルータの設定 (Vyatta)

- route-mapに一行追加するだけ
 - Neighborが5つ、それぞれv4/v6があるので、合計で10行追加

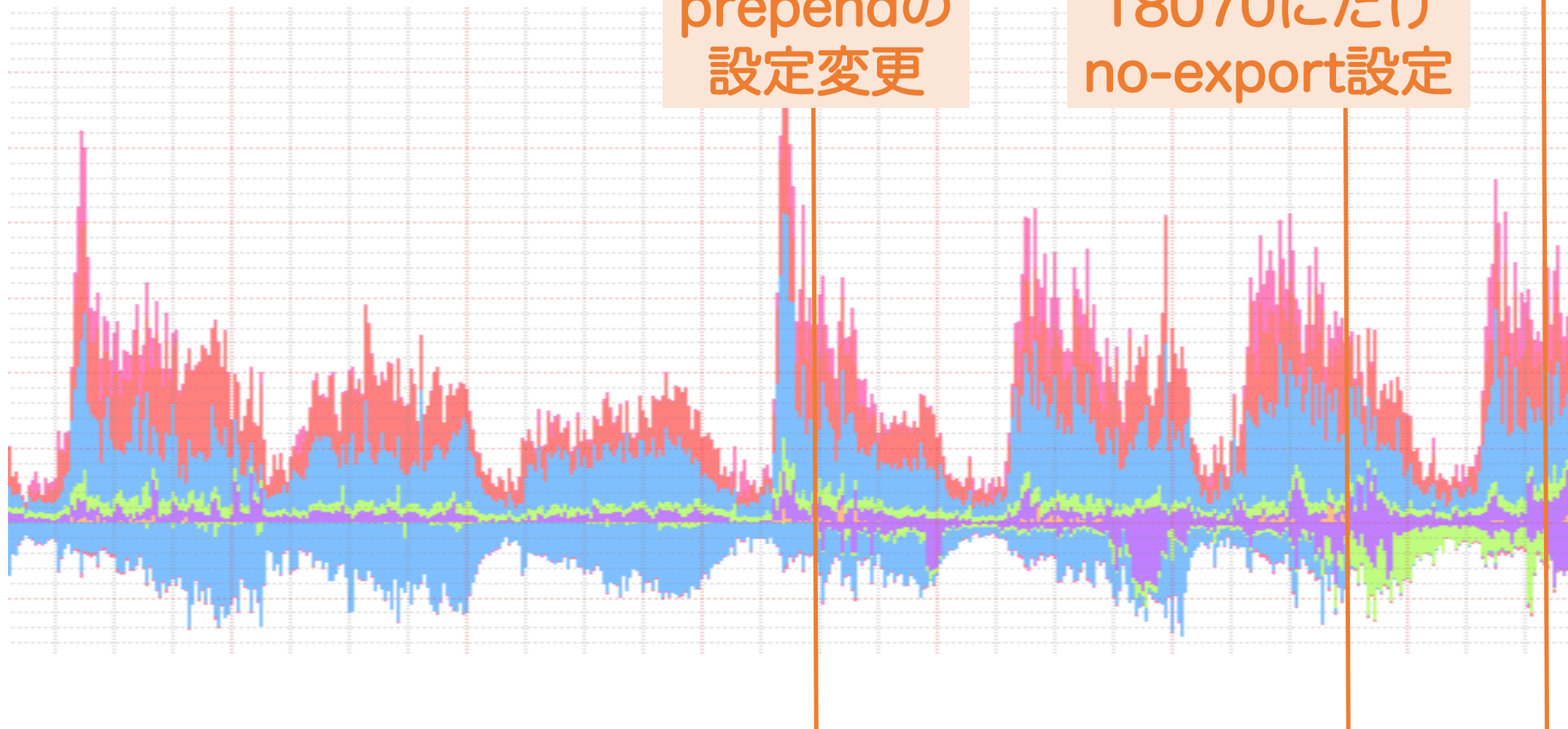
```
set protocols bgp 55895 neighbor 58.84.254.17 remote-as '18070'  
set protocols bgp 55895 neighbor 58.84.254.17 route-map export 'rm-18070-export-v4'  
...  
  
set policy route-map rm-18070-export-v4 rule 10 action 'permit'  
set policy route-map rm-18070-export-v4 rule 10 match as-path 'apl-my-origin'  
set policy route-map rm-18070-export-v4 rule 10 set as-path-prepend '55895 55895 55895 55895'  
set policy route-map rm-18070-export-v4 rule 10 set community 'no-export'  
set policy route-map rm-18070-export-v4 rule 990 action 'deny'
```

うまくいったようです

紫以外の全てに
no-export設定

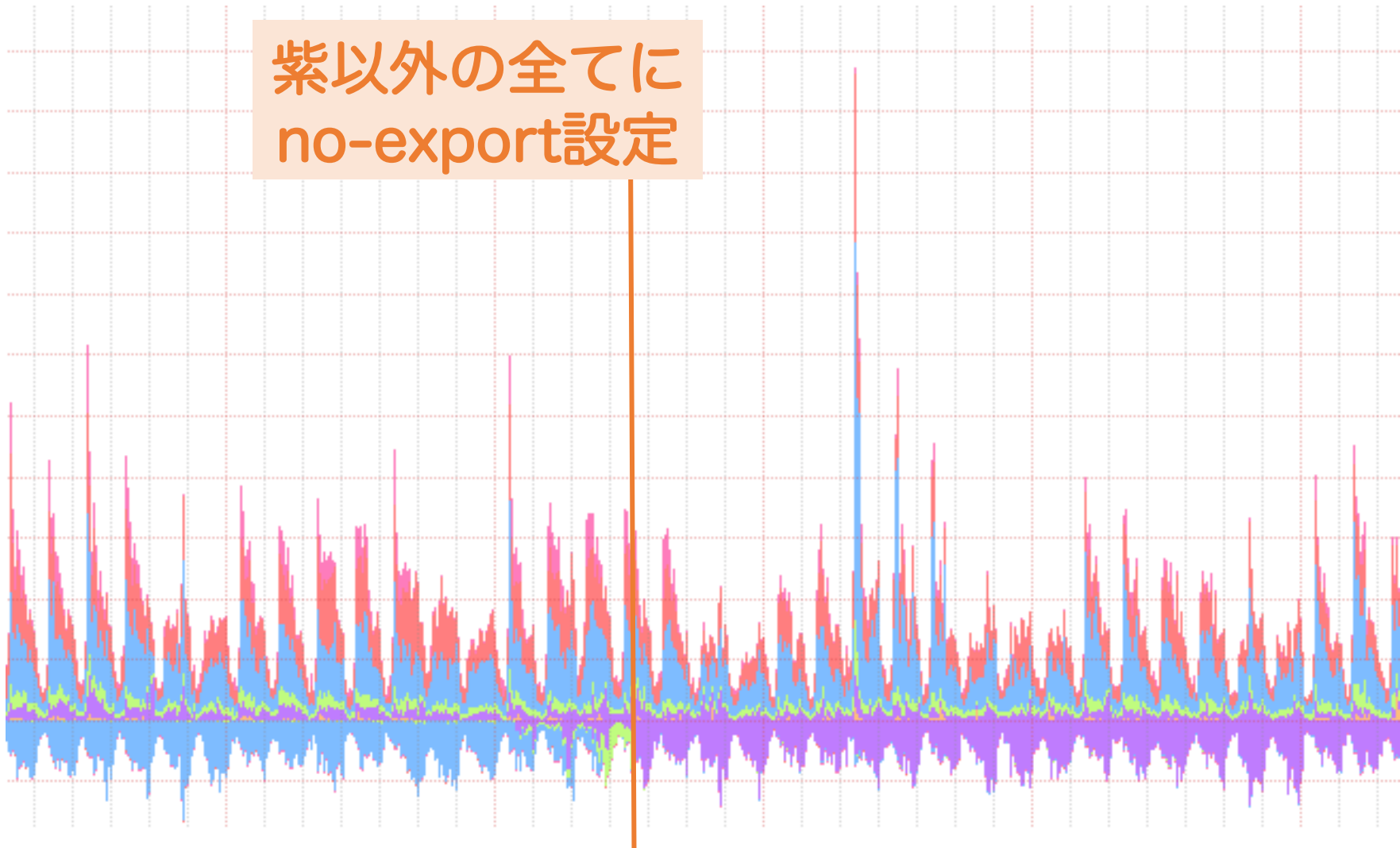
prependの
設定変更

18070にだけ
no-export設定

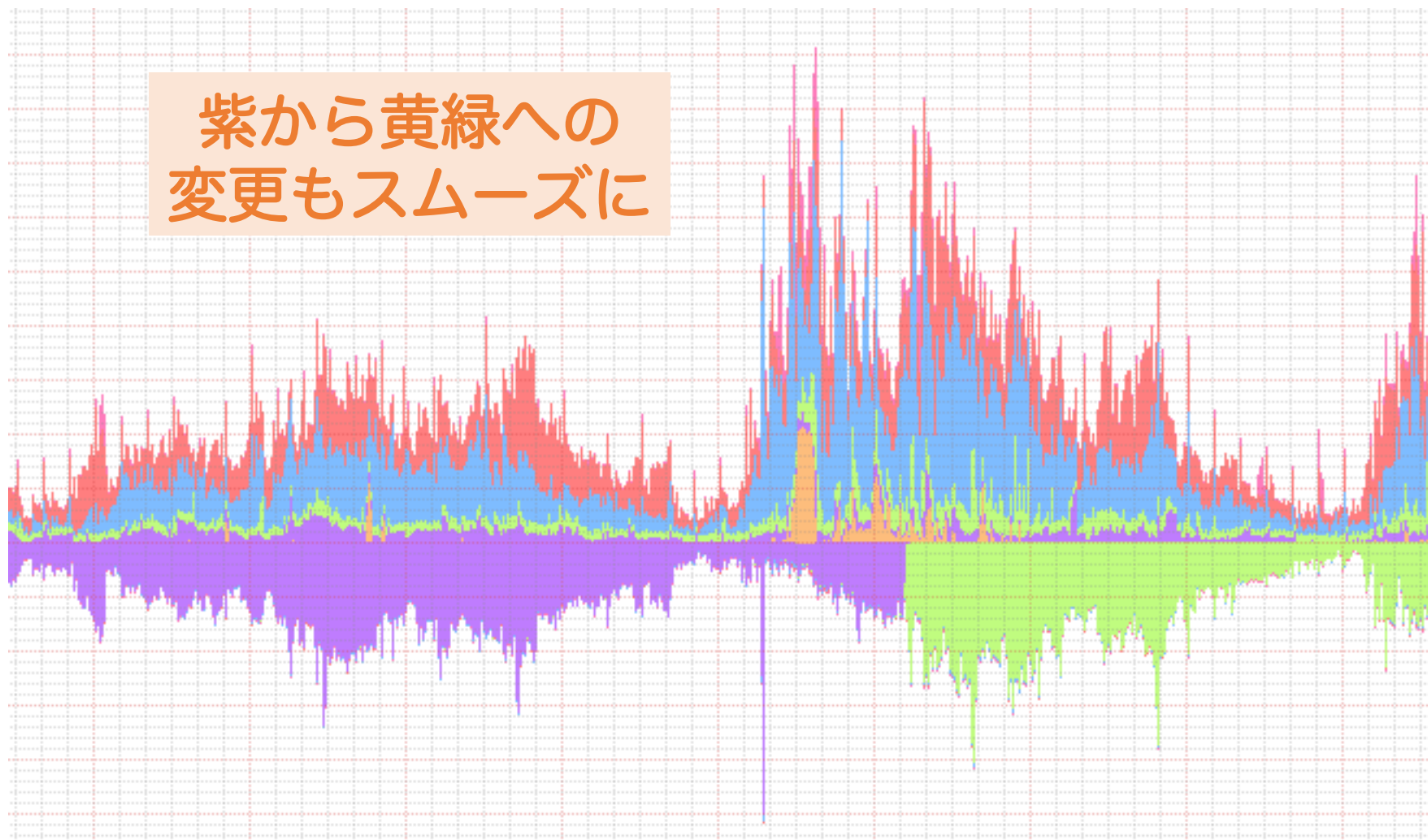


その後もうまくいってます

紫以外の全てに
no-export設定



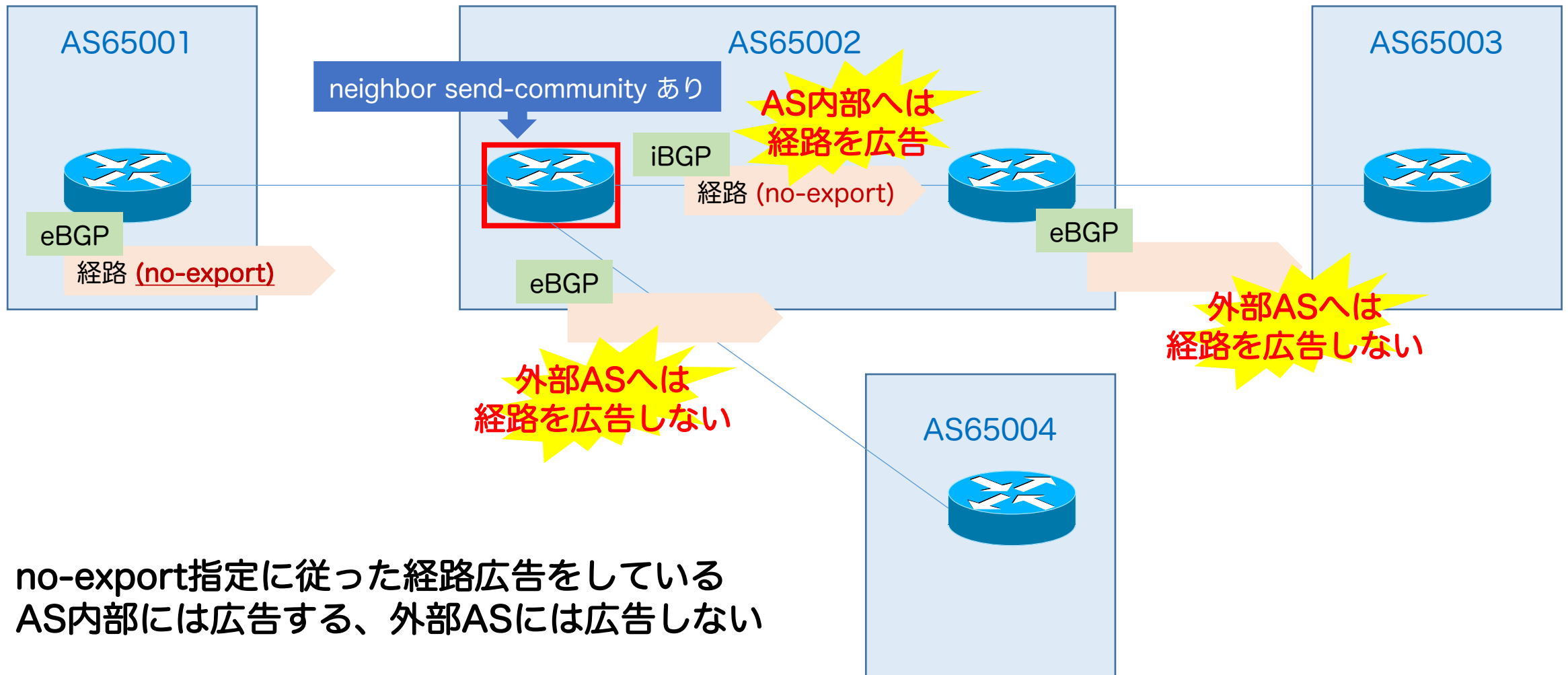
さらにその後も、うまくいってます



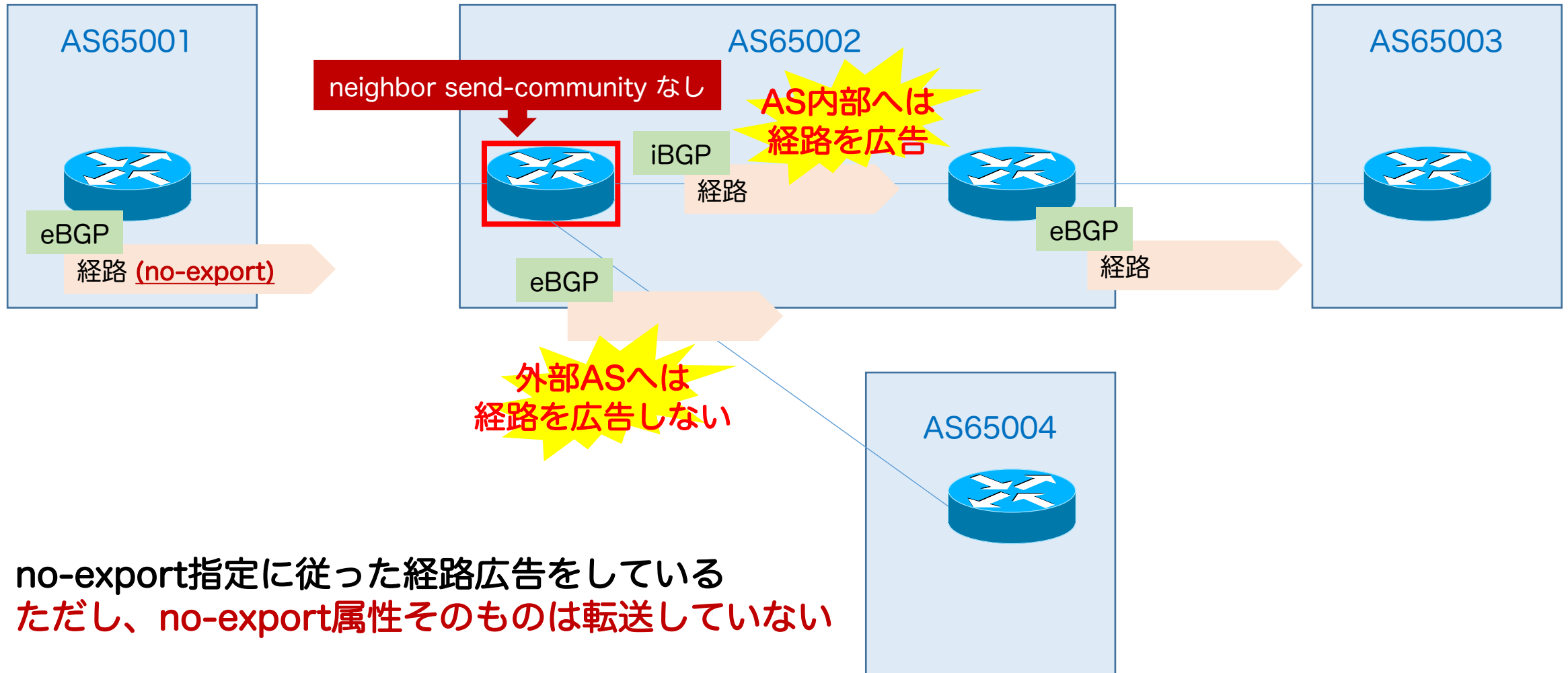
ちなみに

- ルータ自身は、well-known communityのついた経路の取り扱いは正しく判断する
- neighbor send-communityは、あくまでもコミュニティ属性を隣接ルータに転送するかしないかの設定であって、経路広告の判断とは無関係
- 次ページで図解

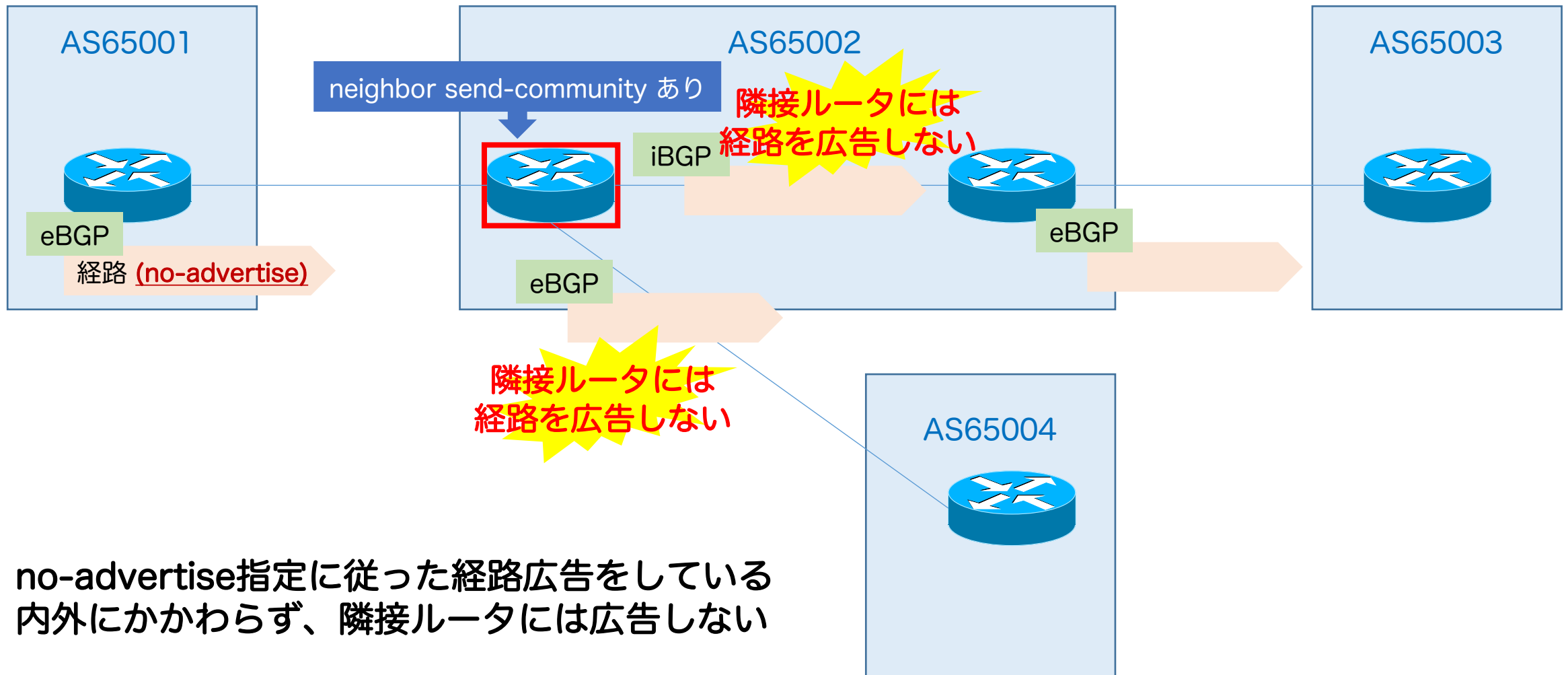
no-export経路に対するルータの動作



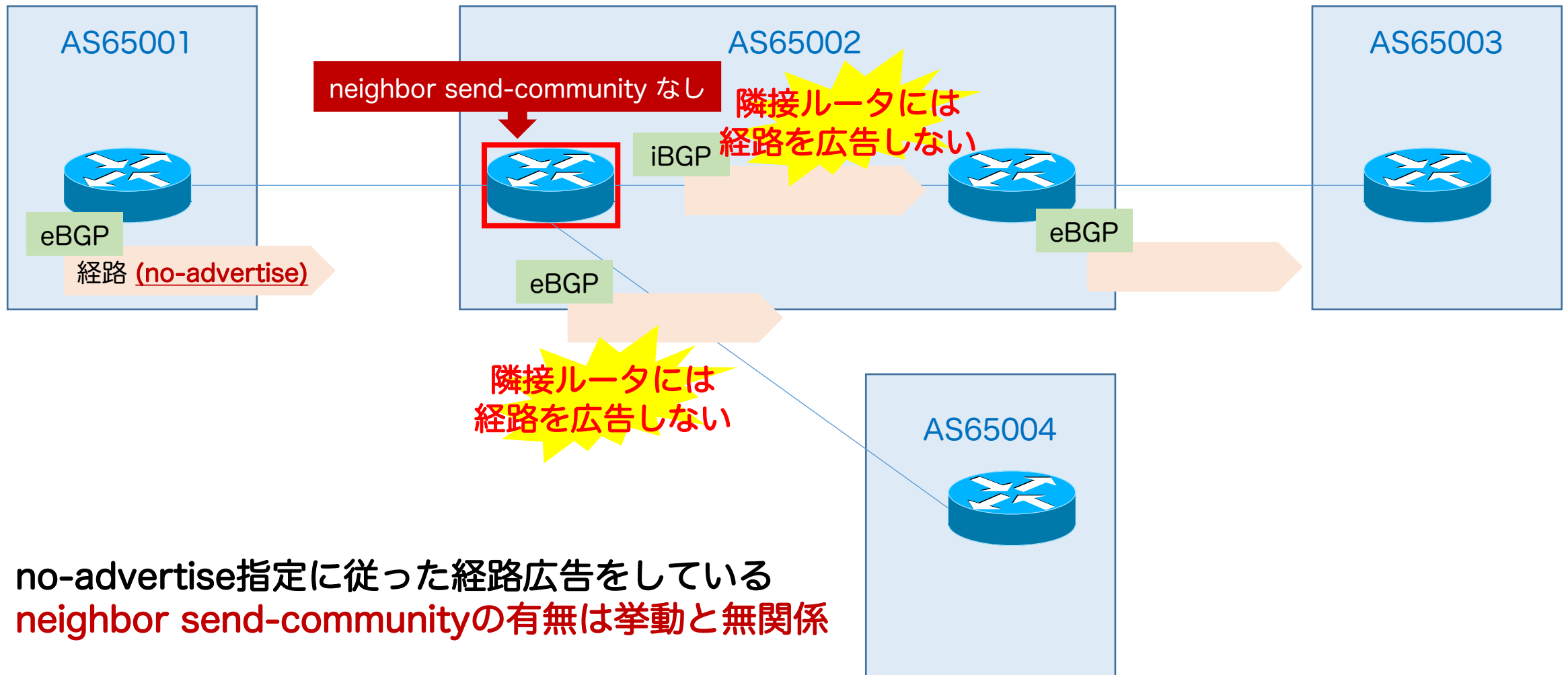
no-export経路に対するルータの動作



no-advertiseの場合の挙動



no-advertiseの場合の挙動



IOS以外のルータでは？

- VyOSの場合

- neighbor send-communityのようなコマンドは特になさそう
- 検証したところ、デフォルト設定でno-exportはiBGP転送された

- JUNOSの場合

 本発表時に会場から、JUNOSでは特にそのようなコマンドはない、
というコメントをいただきました

- ~~neighbor send-community~~というコマンドがあるっぽい？
- vSRXで検証しようとしたけど、サーバリソースが足らず・・・
- ということで実際の状況は未確認です m(_ _)m

まとめ

- AS Path Prependでの制御はやっぱり限界がある
 - 必ずしもASパス長で決まるわけではない
 - Akamaiじゃなくても、local-prefとか相手のポリシー次第だから
- no-exportコミュニティを使って希望の運用ができた
 - IX側の設定変更のみでトランジットをきっちり変更できる
 - 冗長性はないけどね！
- 少なくともIOSでは、コミュニティ転送はデフォルトではない
 - 使う場合はneighbor send-communityを入れましょう

参考までに

- 小島さんのこの資料、素晴らしいですよ！
- <https://speakerdeck.com/codeout/bgp-communityfalse-ji-ben-she-ji>
 - BGP Communityの基本設計



ついでに宣伝

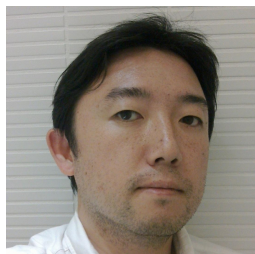


- 2017年12月1日（金） IP Meeting 2017@浅草橋

パネルディスカッション

今そこにある「グローバル」インターネット
～ 私たちは繋がっている ～

- 豪華パネリスト



私がモデレータ
やります
見に来てね！！