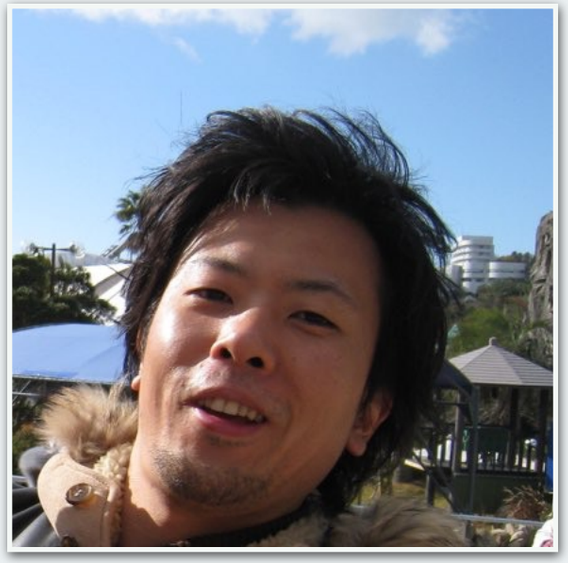
2017年ルーティングアップデートと情報収集手段



小島 慎太郎

codeout

http://about.me/codeout

2004 ISP ntt.net

2009 IX JPNAP

2014 フリーランスの AS手伝い

本日のゴール

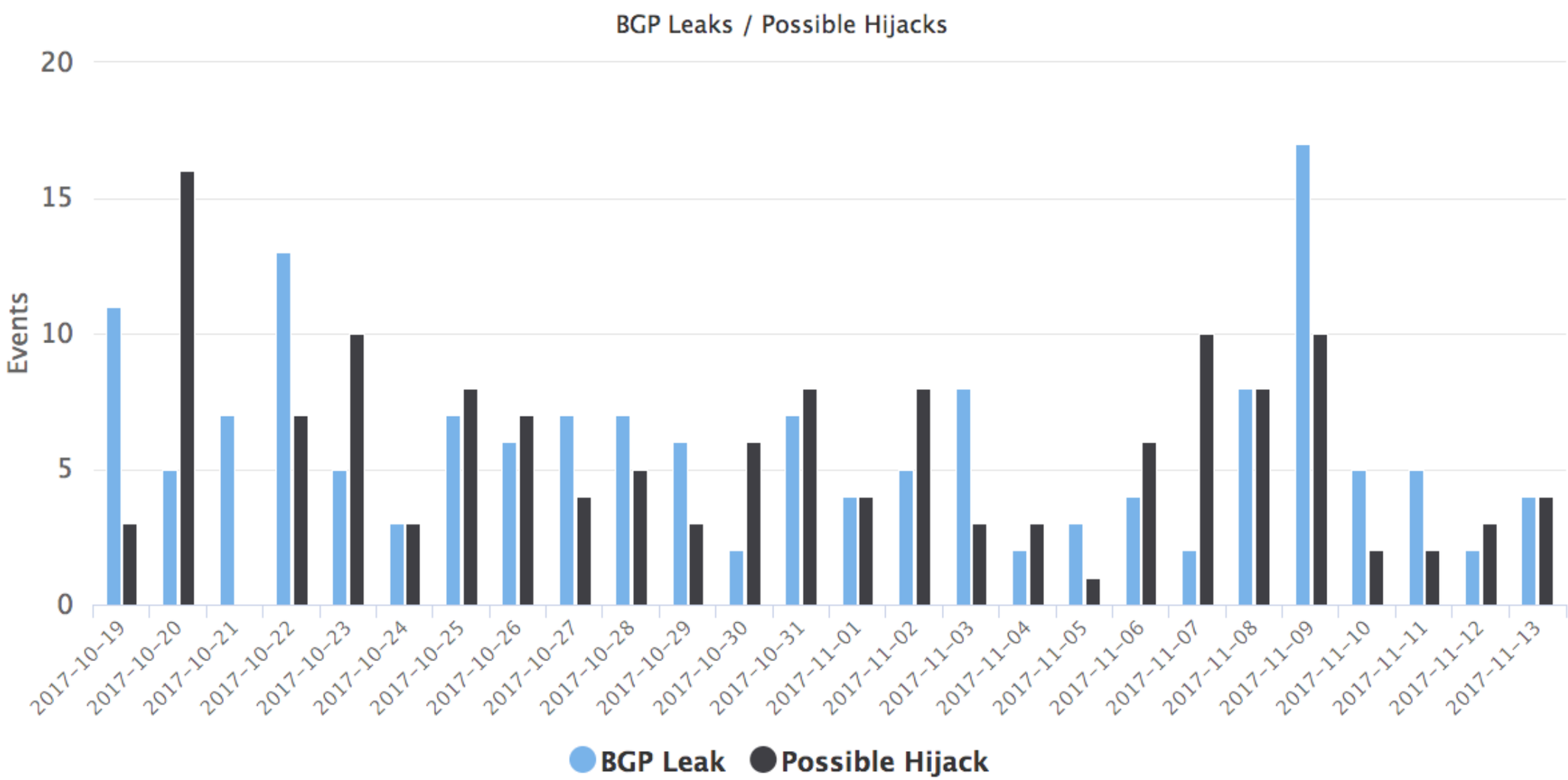
- 経路障害について、2017年のトレンドを知る
- 経路障害について、自社に影響がなくても調査できる。その方法を学ぶ

2017年

気になる経路障害

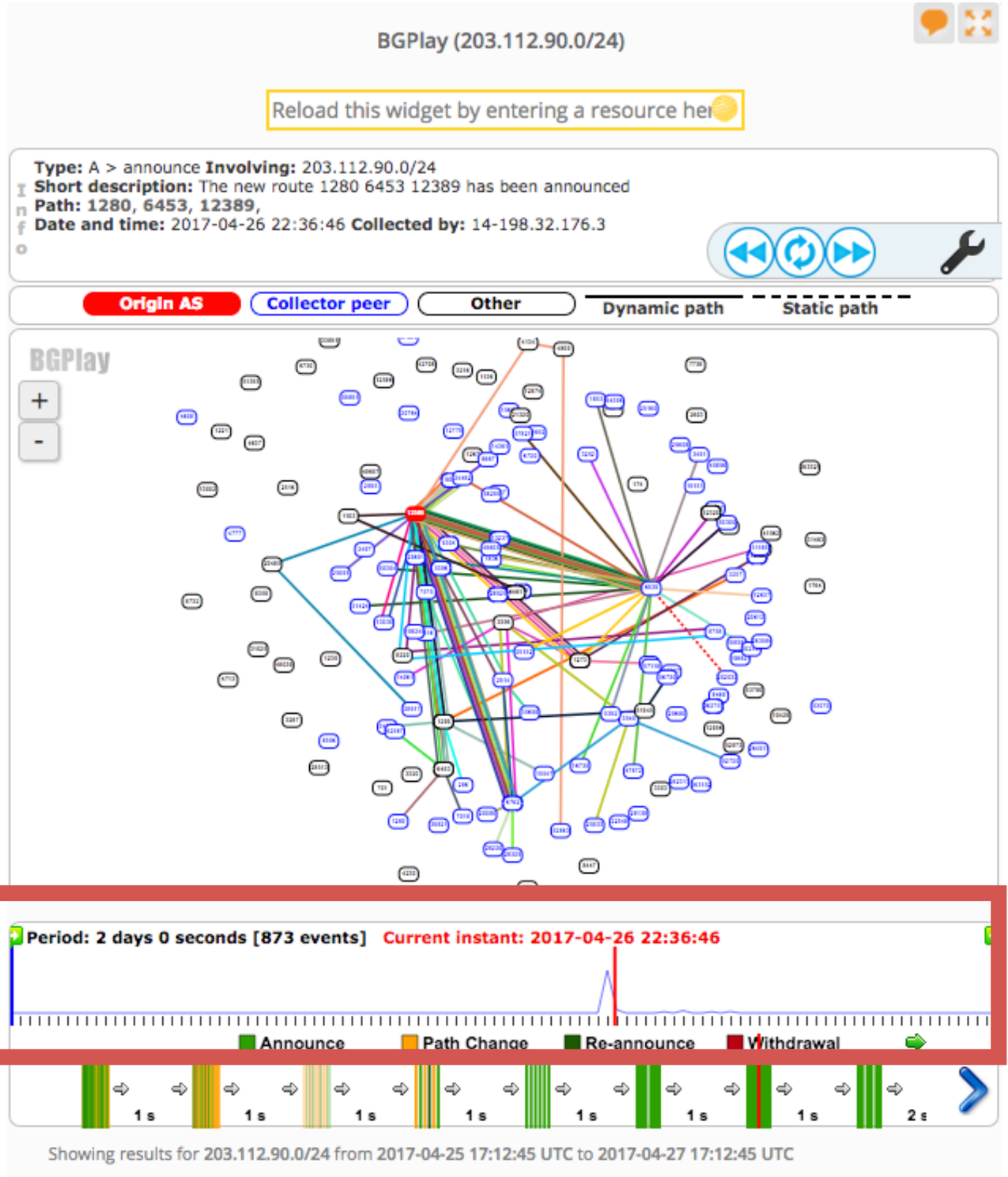
毎日毎日、なにか起こっています

- https://bgpstream.com/ のデータ
- 経路リーク、ハイジャックは 平均 10件/日



4/26 Rostelecom(AS12389) が金融系サービスの経路をハイジャック

- MasterCard、Visa、銀行などの経路がハイジャックされた
- 50 経路、37 ASN
- 通常時は存在しない経路
- ほとんどが /24
- 7分間

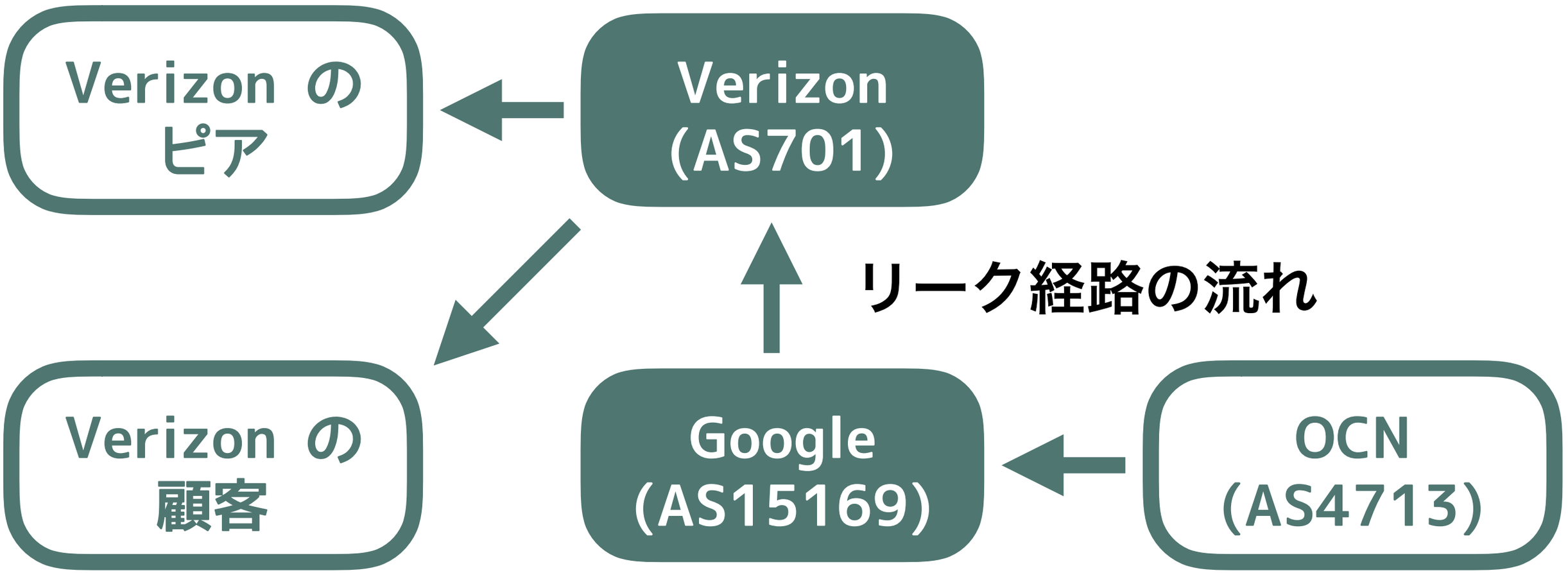


https://stat.ripe.net/widget/bgplay#w.resource=203.112.90.0/24&w.ignoreReannouncements=false&w.starttime=1493140365&w.endtime=1493313165&w.rrcs=0,1,2,5,6,7,10,11,13,14,15,16,18,20&w.instant=null&w.type=bgp

4/26 Rostelecom(AS12389) が金融系サービスの経路をハイジャック

- 悪意によるものかは不明だが、経路はグローバルに伝搬
  - 悪意があったとすれば、これはオペミス
  - グローバルに検知されることは本意ではないはず
- 当然暗号化されているが、その暗号は安全か？
  - 脆弱性、弱いCipher Suite
- ふだん存在しない経路であり、本来の Origin ASはバラバラ
- Rostelecom所有の、ちがうASNからも同時に出ていた
  - BGPMON「悪意があったという根拠は出てない。ミスっぽい」
  - https://bgpmon.net/bgpstream-and-the-curious-case-of-as12389/
- 7分。あらかじめ対策が必要

8/25 Google(AS15169) が OCN(AS4713)などの経路をリーク



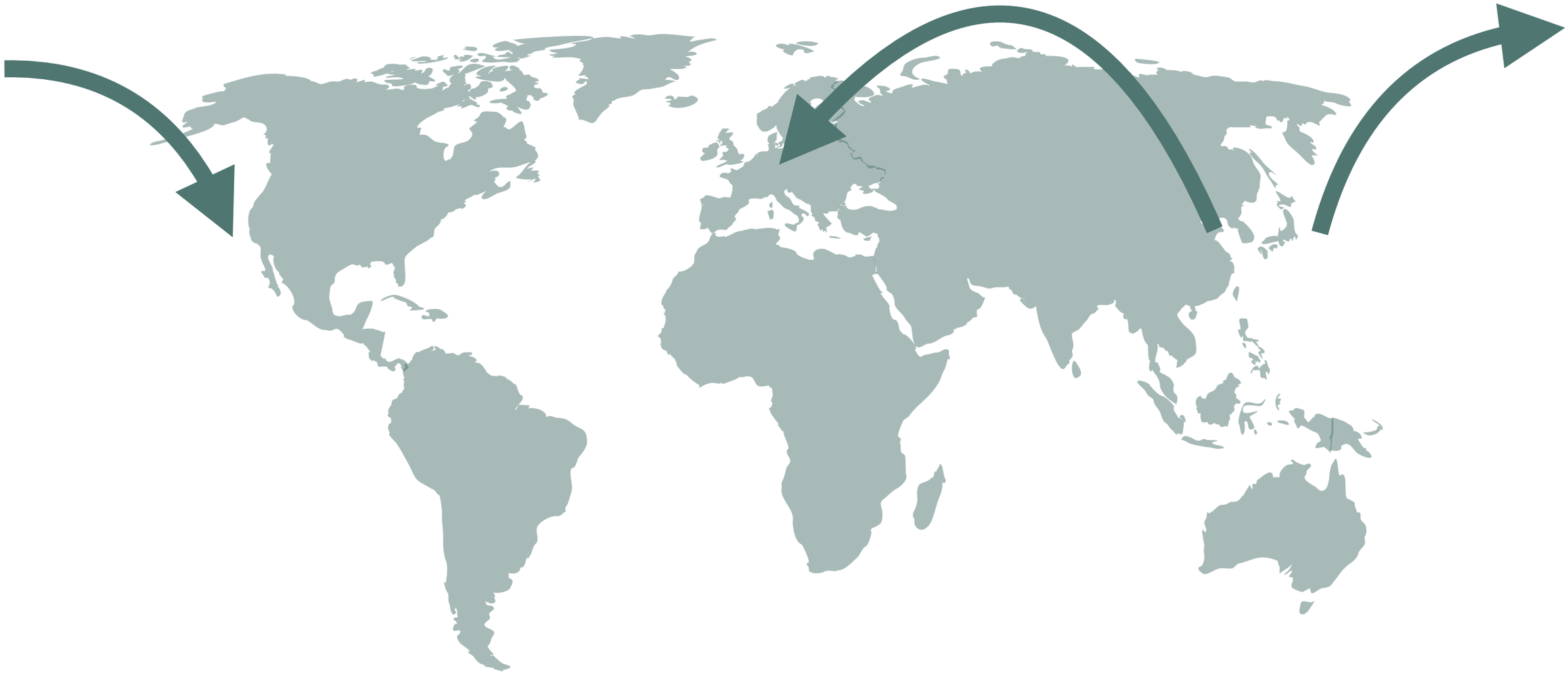
通常時は存在しない、more speific 経路

8/25 Google(AS15169) が OCN(AS4713)などの経路をリーク

- 92,000 経路、6,300 Origin ASN
  - ベストパスが変わった → 遅延 + パケロス
  - RIB / FIB があふれた
- 通常時は存在しない、more specific 経路
  - ほとんどが /24
  - 12分間
  - 報道によれば「ネットワークの誤設定」
  - 日本は大騒ぎ。影響ありました？

これが、日本 / 北米 / 南米 / EU で観測された。
経路リーク。ハイジャックではない。Origin AS は正しい

8/25 Google(AS15169) が OCN(AS4713)などの経路をリーク



- ふだん流れないところにトラフィック流入 → 輻輳
- 経路の受け取り方によって海外まわりに → 遅延
- リークした経路、Reachable ではあった模様

11/6 Level3(AS3356) が Comcastなどの経路をリーク

リーク経路の流れ

- NTTCOM (AS2914)
- TELEFONICA (AS12956)
- TATA (AS6453)

- BACOM (AS577)
- COMCAST (AS33491)
- COMCAST (AS7725)
- ...

通常時は存在しない、more speific 経路

11/6 Level3(AS3356) が Comcastなどの経路をリーク

リーク経路の流れ

- 遅延 + パケロス
- リークした経路、観測点によりReachable ではあった模様
- ふだん流れないところにトラフィック流入 → 輻輳

11/6 Level3(AS3356) が Comcastなどの経路をリーク

- 8,000 経路、56 Origin ASN
  - ベストパスが変わった → 経路上のどこかで遅延 + パケロス
- 通常時は存在しない、more specific 経路
- /14 ~ /24
- 97分間
- 報道によれば「ネットワークの誤設定」
- 日本ではさほど騒がれなかった

11/6 Level3(AS3356) が Comcastなどの経路をリーク

日本で騒がれなかったのはなぜか？

- 8/25 の経路リークと比べて注目度が低い
- Level3 のリーク経路は、日本でも観測されている
  - リーク経路の量が少なかった
    - 92000 vs. 8000
    - RIB / FIB サイズ ギリギリ運用しすぎ疑惑
  - リーク経路との通信がもともとなかった

2017年、気になる経路障害

- 大規模事業者のオペミスによる事故が目立つ
- 局所的にトラフィックエンジニアリングなどに使っているmore specific のリーク。普段グローバルには存在しない

- 2016年は
  - 悪意ある、局所的なハイジャックがトレンドだった
  - 一時的にIXにつないで経路を出す、とか
  - ソーシャルハッキング的
  - 現在そのトレンドは過ぎて、攻撃が減ったのか？
    攻撃はあるが、うまく防いでいるのか？

2017年トレンドまとめ

- 大規模事業者の 悪意のないミス によって影響を受けるケース
- 大きな事業者であるほどインパクト大

自社や顧客を守るために、

- どのような影響を受けそうかを知り、
- 対策を打ちたい
- 特に RIB / FIB サイズ要注意

本日のゴール

✅ 経路障害について、2017年のトレンドを知る

- 経路障害について、自社に影響がなくても調査できる。その方法を学ぶ

後半はこのあたりの話をします

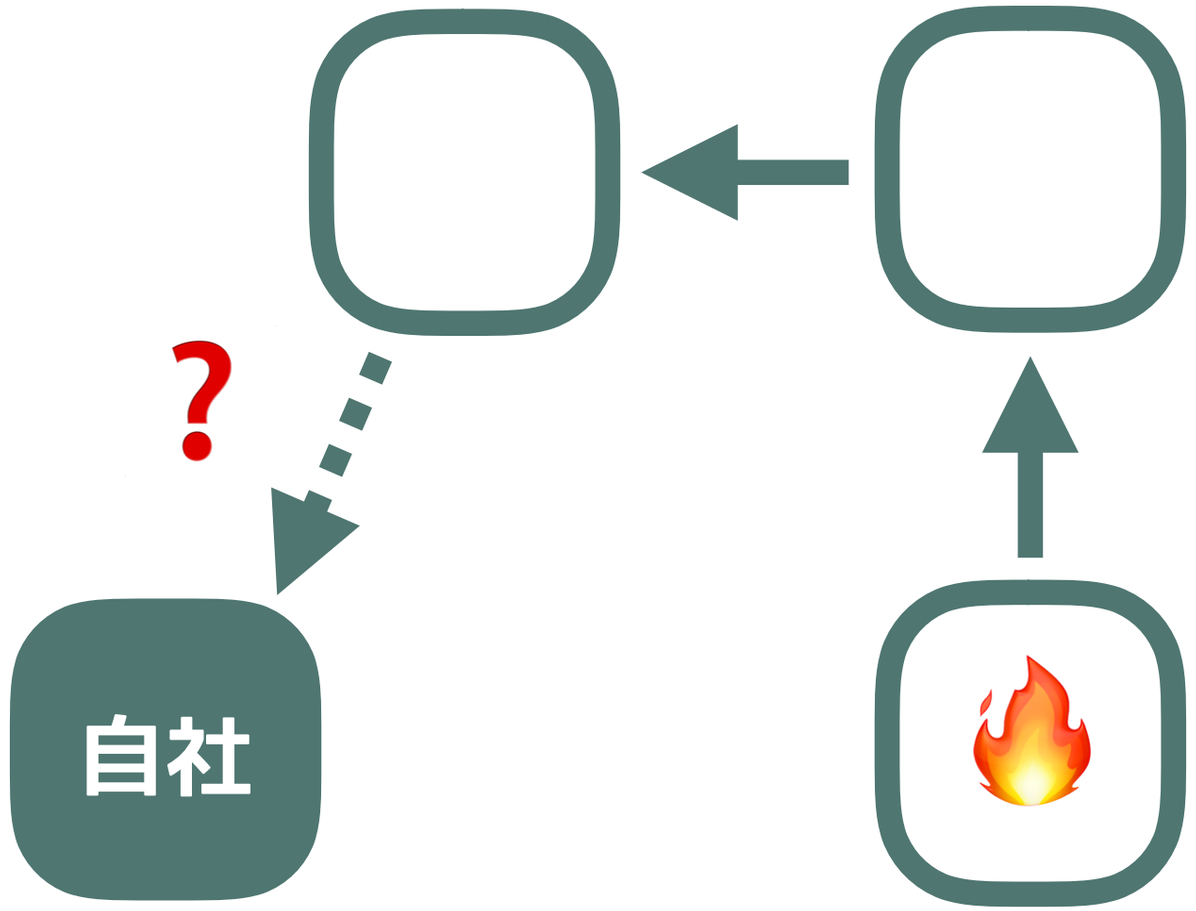
- 影響のあった経路障害は、調査して対策検討できる
  - 「リスクを容認して対策しない」も含めて

- 影響のなかった経路障害は？
  次も影響がないと言えるか？
- 「データがないのでわかりません」からの脱却

近くで起こった経路障害

- インパクトを受けやすいが、調査もしやすい
- 自社で観測できることが多い

遠くで起こった経路障害



8/25、11/6 ふたつのリーク、BGPセッション自体はともにUS にあったと思われる

❌ 遠いなら大丈夫

❌ 今回大丈夫だったから次回も大丈夫

自社網で検知できなかった
経路障害を調査する

- 経路障害のしくみ
- 影響範囲

を調査し、推測することは可能です

ケーススタディ

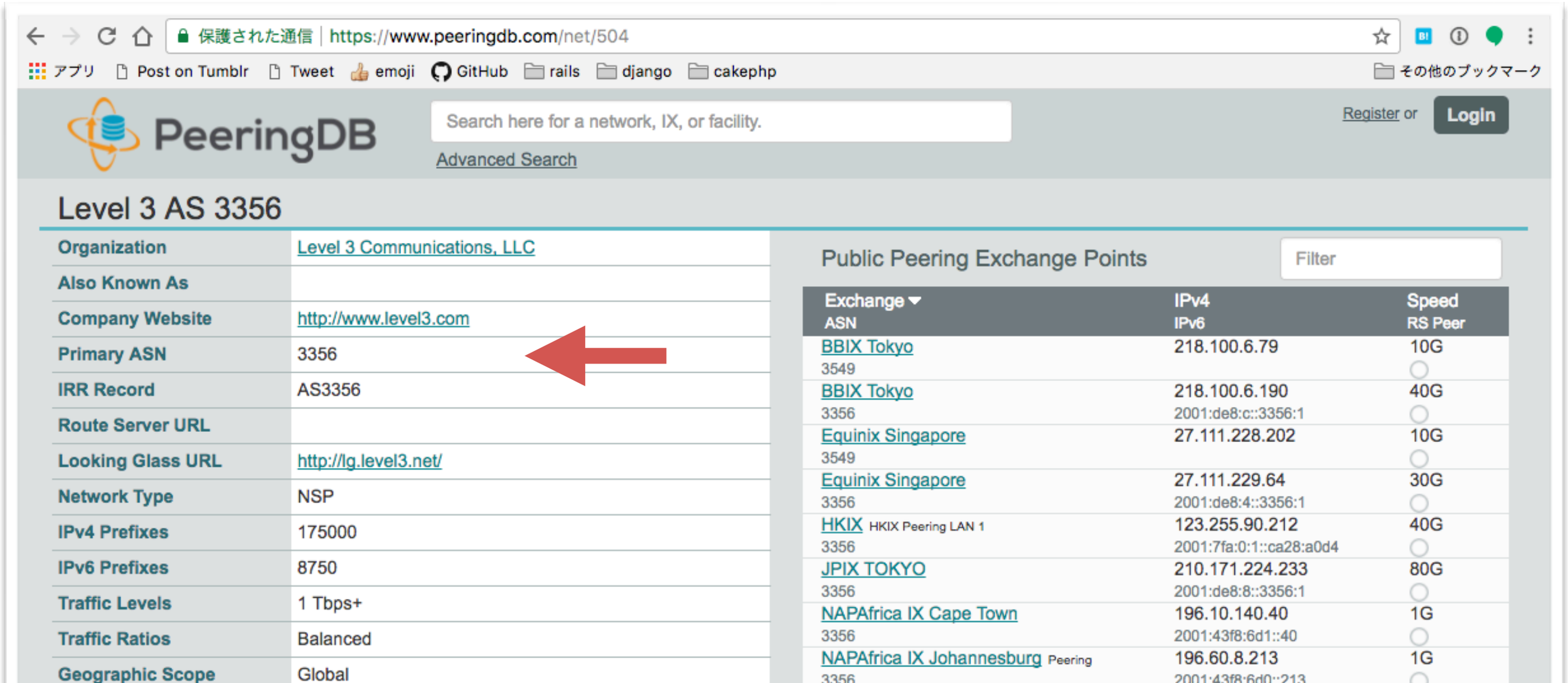
- 11/6 Level3 の経路リークを題材に

11/6 に Level3 が
経路リークしたっぽい

https://techcrunch.com/2017/11/06/comcast-xfinity-slow-outage/

11/6 に Level3 が経路リークした

- Level3 ?
- ASN = 3356 らしい
- AS3549 のほうは 旧Global Crossing

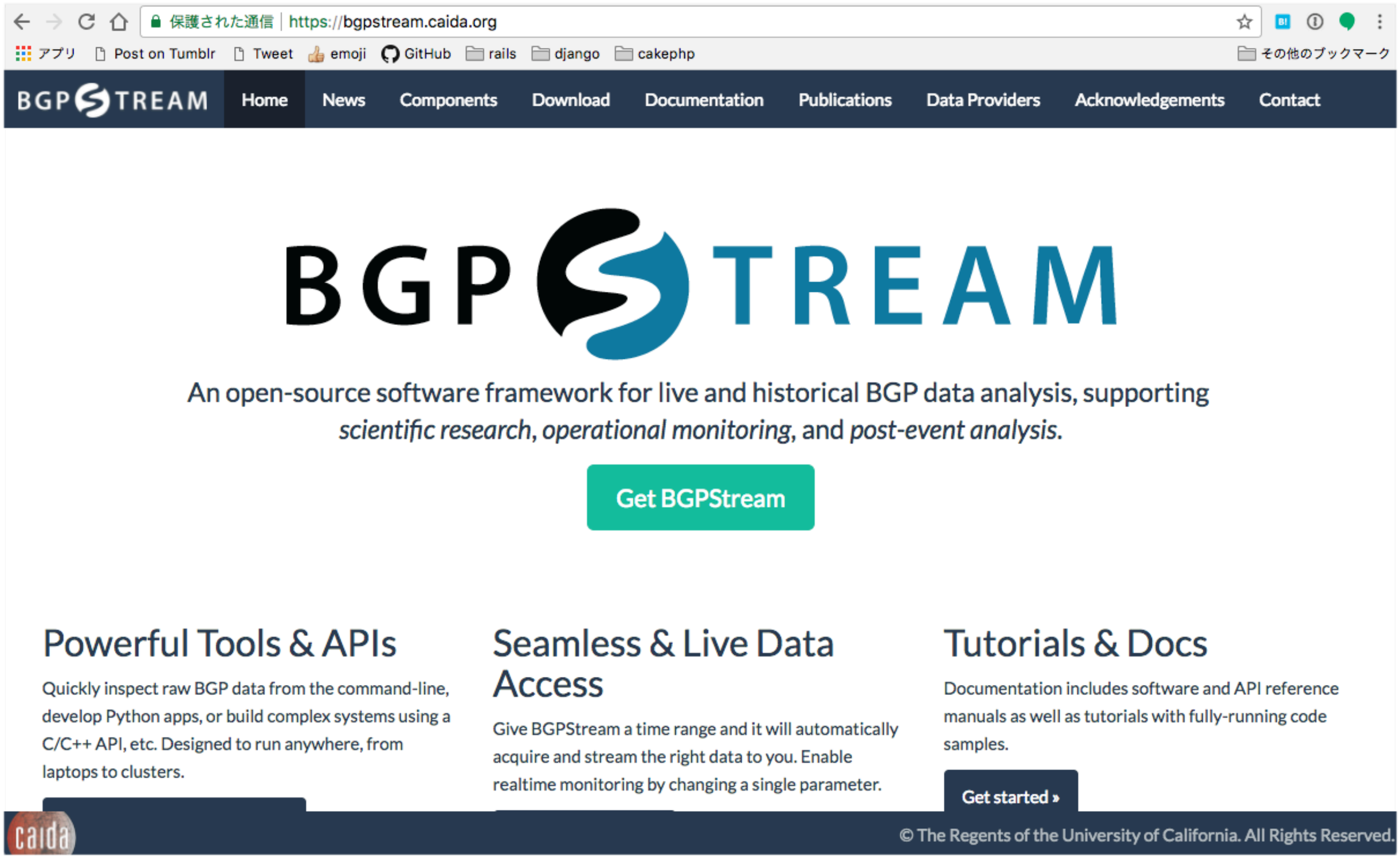


わかったこと

✅ Level3 = AS3356

次は、Level3の広告経路 を調べる

11/6 の Level3 広告経路を調べる



BGP Stream

- https://bgpstream.caida.org/data
- 様々なBGP Update / RIB アーカイブを串刺し検索するAPI + API クライアント
- CAIDA のプロジェクト。bgpdump の代替
- データソース
  - Route Views
  - RIPE NCC
  - OpenBMP
  - BGPmon

BGP Stream

```
iw2017 $ bgpreader -m -c route-views.eqix -w 1509926400,1509926700 -t
updates

BGP4MP|1509926485|A|2001:504::2:0:1:9151:1|19151|2804:35a0::/32|19151 6939
16735 263097 263650 266297|IGP|2001:504::2:0:1:9151:1|0|27|19151:1000
19151:1500 19151:61007 19151:65040|NAG||
BGP4MP|1509926485|A|206.126.236.47|19151|146.16.0.0/22|19151 2914 209 721
27064 5953|IGP|206.126.236.47|0|0|2914:420 2914:1001 2914:2000 2914:3000
19151:1000 19151:1500 19151:61006 19151:65050 65504:209|AG|65037
144.104.74.0|
BGP4MP|1509926485|W|206.126.236.120|41095|146.16.184.0/24
BGP4MP|1509926485|A|2001:504:0:2::3257:1|3257|2001:7fb:fe15::/48|3257 6453
29075 12654|IGP|2001:504:0:2::3257:1|0|0|3257:8111 3257:30118 3257:50002
3257:51100 3257:51102|NAG|64845 10.6.151.128|
BGP4MP|1509926485|W|206.126.236.25|6079|146.16.184.0/24
BGP4MP|1509926485|W|206.126.236.25|6079|195.64.138.0/23
BGP4MP|1509926485|W|206.126.236.25|6079|177.38.238.0/24
BGP4MP|1509926485|W|206.126.236.25|6079|131.108.171.0/24
BGP4MP|1509926485|A|2001:504:0:2::6762:1|6762|2001:7fb:fe15::/48|6762 29075
12654|IGP|2001:504:0:2::6762:1|0|100|6762:1 6762:92 6762:13300|NAG|64845
10.6.151.128|
BGP4MP|1509926485|W|2001:504:0:2::5769:1|5769|2407:6c00:8::/48
```

11/6 を指定し、bgpdump 同等の情報が取れる

わかったこと

- ✅ Level3 = AS3356
- ✅ 広告経路を取ってこれそう

データ量がまだ多い。時刻を絞る

データ量がまだ多い
時刻を絞る

こういう感じの
グラフが欲しい

11/6 Level3の
広告経路数の変動をみたい

- BGP Update アーカイブ中のどのピアを選択するか
- とりあえずRIB を眺める
- US に近そうなEquinix Ashburn を選択
- AS_PATH 上で "3356" の左隣を一覧する

```
iw2017 $ wget http://archive.routeviews.org/route-views.eqix/bgpdata/
2017.11/RIBS/rib.20171106.0000.bz2
iw2017 $ bgpdump -M rib.20171106.0000.bz2 | grep ' 3356 ' > 3356.txt
iw2017 $ cut -d \| -f 7 bgpstream/3356.txt | sed 's/ 3356 .*//g; s/.* //g' |
sort -nu
174
209
297
577
1299
2914
3257
```

11/6 Level3の広告経路数の変動をみたい

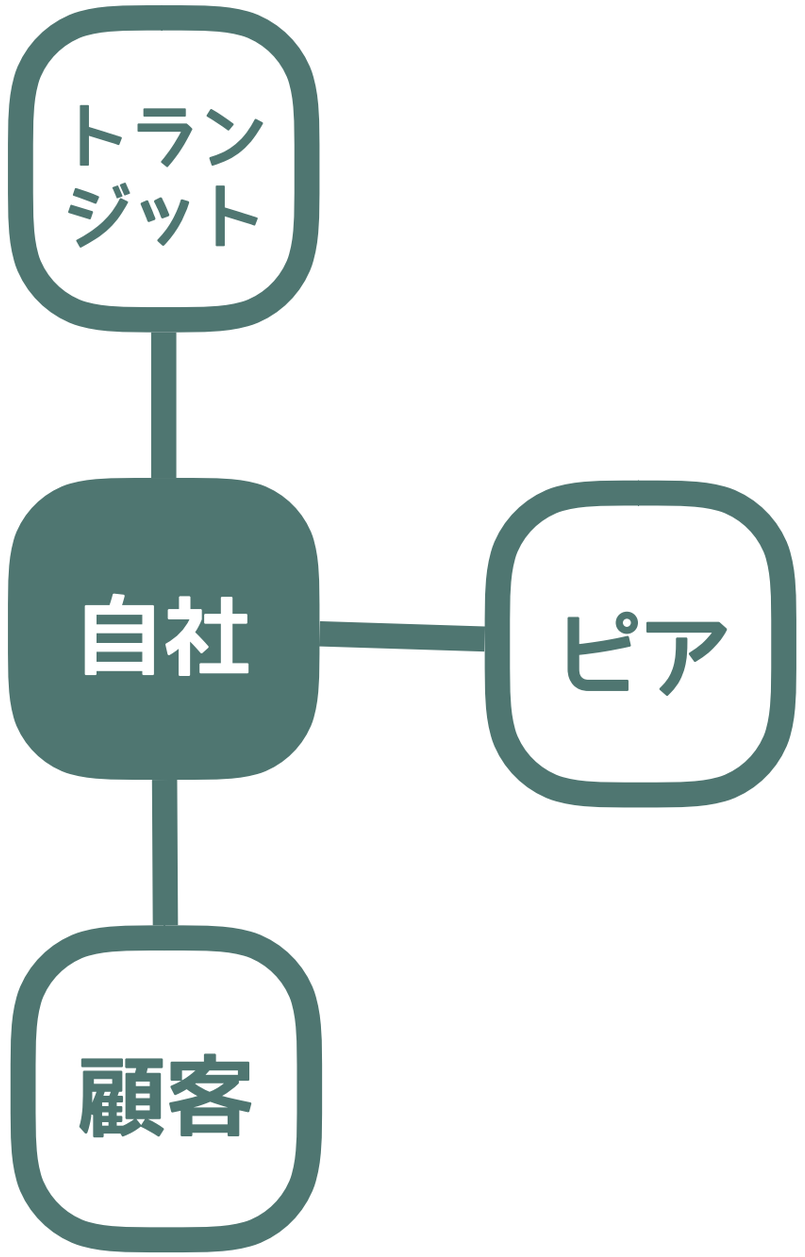
- かつて(?) Tier1 と呼ばれたNTTCOM(AS2914)がいる
- Level3 はTier1 であり、経路リークを探すならTier1とのピアがよさそう

```
iw2017 $ wget http://archive.routeviews.org/route-views.eqix/bgpdata/2017.11/RIBS/rib.20171106.0000.bz2
iw2017 $ bgpdump -M rib.20171106.0000.bz2 | grep ' 3356 ' > 3356.txt
iw2017 $ cut -d \| -f 7 bgpstream/3356.txt | sed 's/ 3356 .*//g; s/.* //g' | sort -nu
174
209
297
577
1299
2914
3257
```

おさらい

❌ リークが疑われるもの

	トランジットに広告	ピアに広告	顧客に広告
トランジット経路を	❌	❌	
ピア経路を	❌	❌	
顧客経路を			



Tier1 という条件で考えられる範囲

11/6 Level3の
広告経路数の変動をみたい

- AS_PATH "2914 3356" が含まれていれば
  なんでもOK
- 適当に "206.126.236.47" に注目する

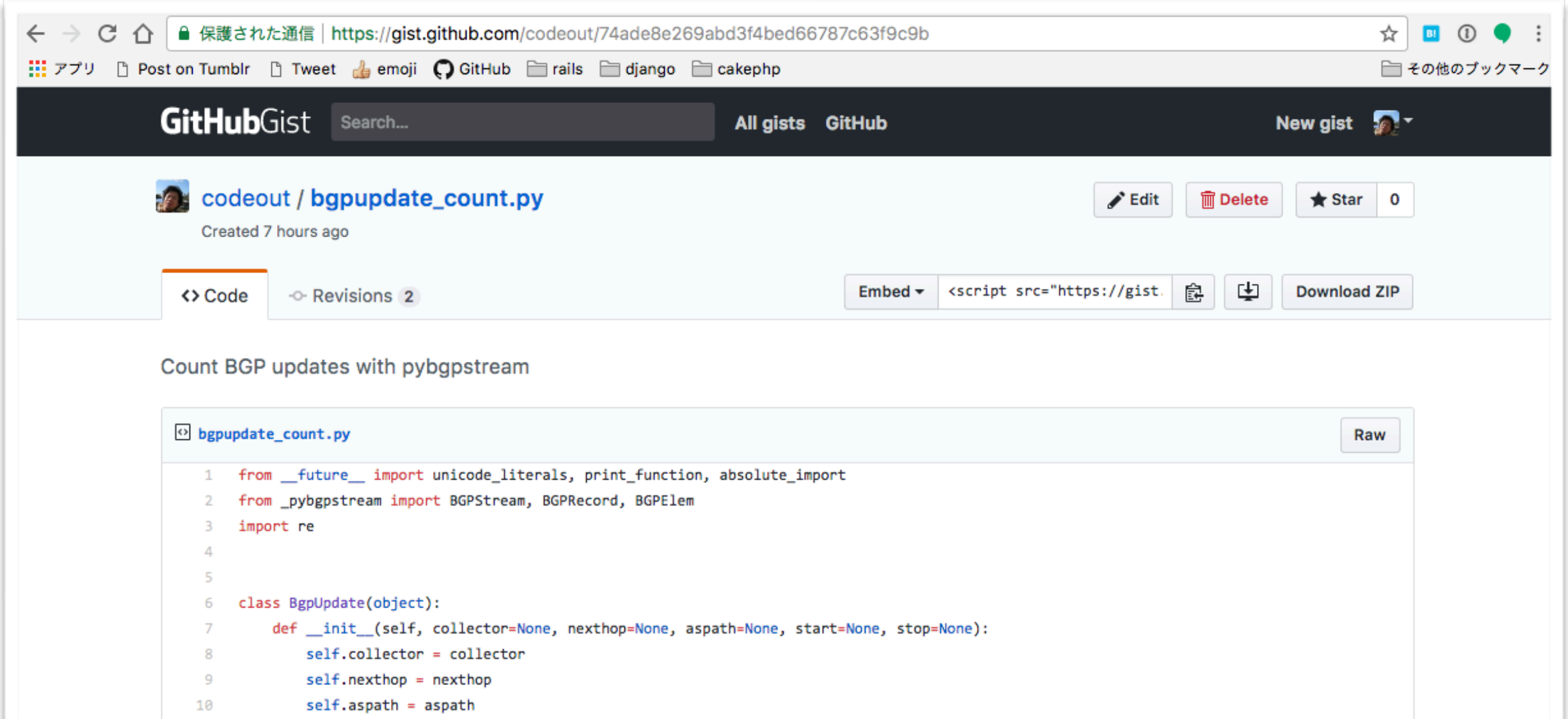
```
iw2017$ grep ' 2914 3356 ' 3356.txt|head
TABLE_DUMP2|11/06/17 00:00:00|B|206.126.236.47|19151|1.119.192.0/21|19151
2914 3356 4837 4808|IGP
TABLE_DUMP2|11/06/17 00:00:00|B|206.126.236.47|19151|1.119.200.0/22|19151
2914 3356 4837 4808|IGP
TABLE_DUMP2|11/06/17 00:00:00|B|206.126.236.10|4589|2.18.64.0/24|4589 2914
3356 6057|IGP
TABLE_DUMP2|11/06/17 00:00:00|B|206.126.236.47|19151|2.18.64.0/24|19151 2914
3356 6057|IGP
TABLE_DUMP2|11/06/17 00:00:00|B|206.126.236.10|4589|2.19.251.0/24|4589 2914
3356 6057|IGP
```

11/6 Level3の
広告経路数の変動をみたい

- 検索条件が確定
  - コレクター "route-views.eqix" のアーカイブ中から、
- neighbor "206.126.236.47" から受信し記録した経路を検索し、
- AS_PATH "2914 3356" を含むBGP Update 数を数える

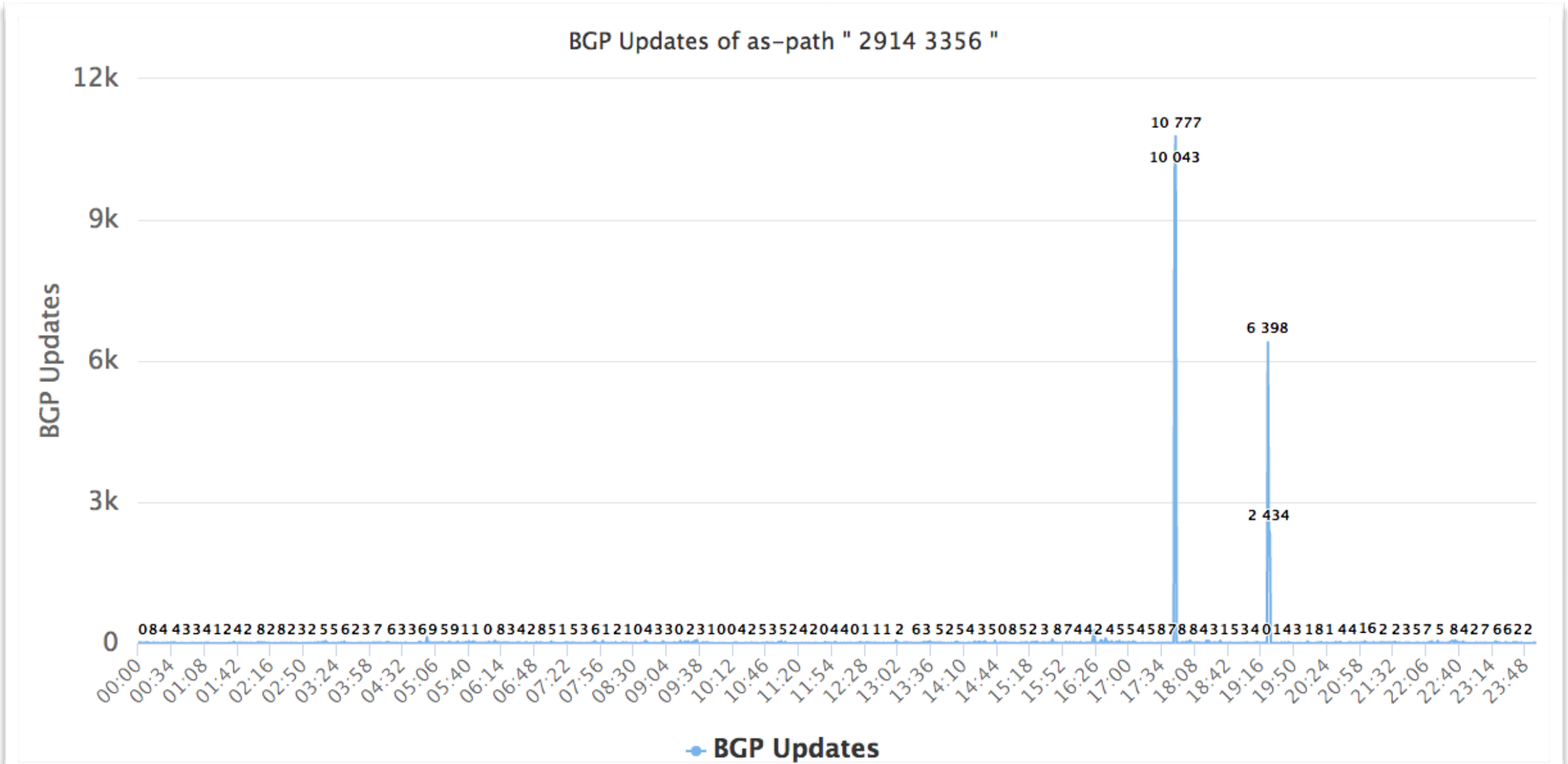
pybgpstream

- 複雑な条件検索、カウント処理が必要
- bgpreader(CLIクライアント) はそこまでやってくれないので、python binding を使う
- https://gist.github.com/codeout/74ade8e269abd3f4bed66787c63f9c9b



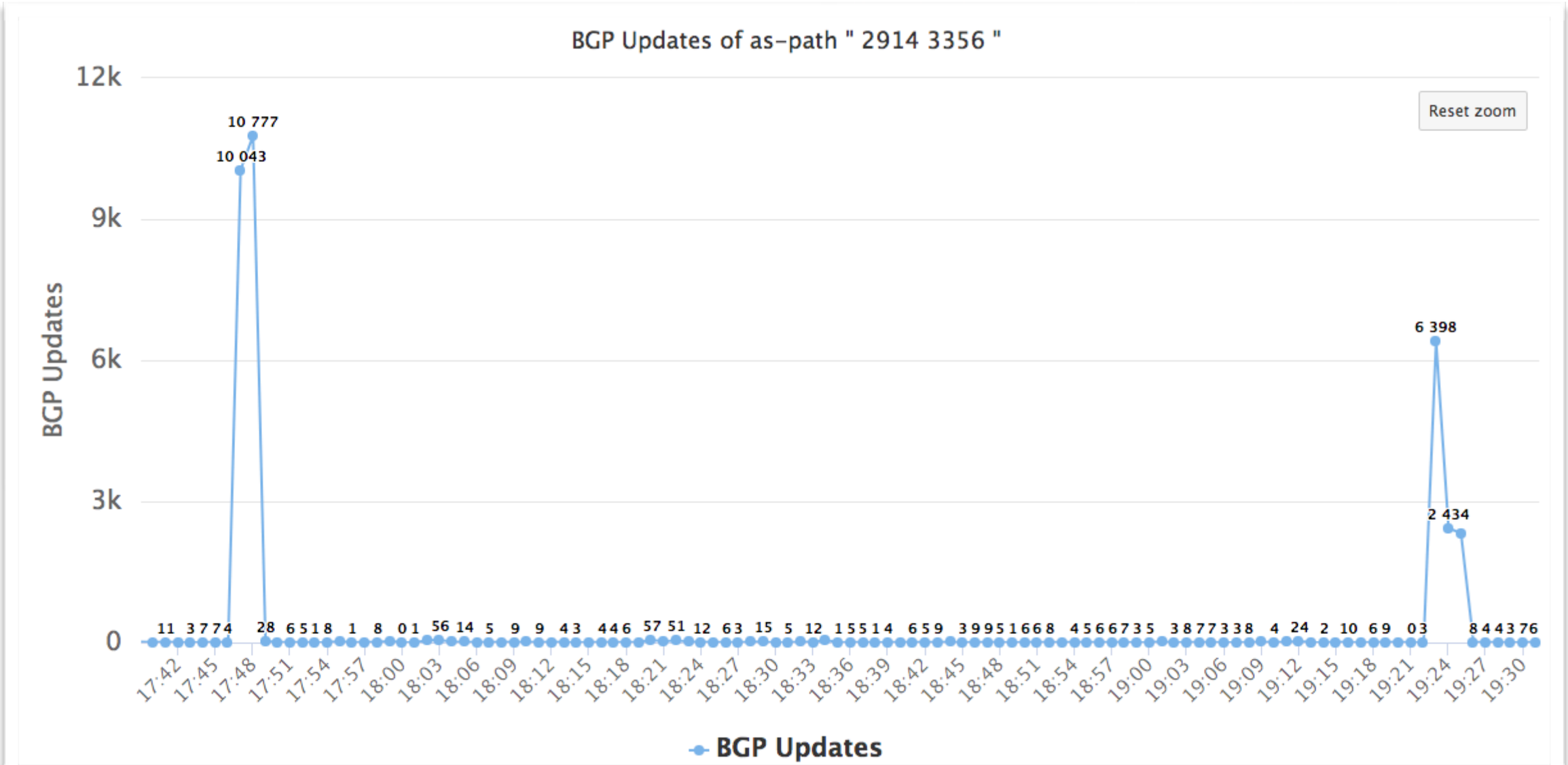
広告経路数の変動、すごく特徴がある

- 実行し、好きなチャートライブラリで描画



ズーム

- 17:47~17:48(UTC)、19:23~19:25(UTC) がおかしい



わかったこと

- ✅ Level3 = AS3356
- ✅ 広告経路を取ってこれそう
- ✅ 多数のBGP Update を送っている時刻がある

リーク経路かどうかは まだわからない。
経路の中身を深掘りする

17:47~17:48 を深掘りする

- 該当の時間帯に観測したBGP Update で、AS_PATH ".* 2914 3356 .*" かつ
- 普段は存在しない経路を探す
- python ではスループットが足りないし、クエリを書きにくい

データベースでやる。

mask 演算ができる postgresql で

BGP Update / RIB アーカイブを データベースに入れる

- 複雑なクエリを実行したい
- 例: Prefix長ごとのBGP Update 数

```
route_leak=# SELECT masklen(prefix), count(*) FROM updates GROUP BY masklen
ORDER BY count DESC;
 masklen | count
---------+--------
      24 | 133430
      22 |  38445
      23 |  22782
      48 |  15206
      21 |  13182
      20 |  12611
      19 |  12209
...
```

BGP Update / RIB アーカイブを データベースに入れる

- https://gist.github.com/codeout/b777f6e1ebb45c908dc96df92a1aa9e2

GitHubGist Search... All gists GitHub New gist

codeout / route_views.rb

Last active 6 hours ago

Edit Delete Star 0

Code Revisions 5

Embed <script src="https://gist. Download ZIP

Download MRT archive from Route Views Project to load into PostgreSQL

route_views.rb Raw

```
1  require 'open-uri'
2  require 'pg'
3  require 'uri'
4
5  module RouteViews
6    DIR = 'mrt'
7    BASE_URL = 'http://archive.routeviews.org'
8    SERVERS = {
9      'oregon2' => '',
10     'oregon3' => 'route-views3',
11     'oregon4' => 'route-views4',
12     'eqix' => 'route-views.eqix',
13     'isc' => 'route-views.isc',
```

BGP Update / RIB アーカイブを データベースに入れる

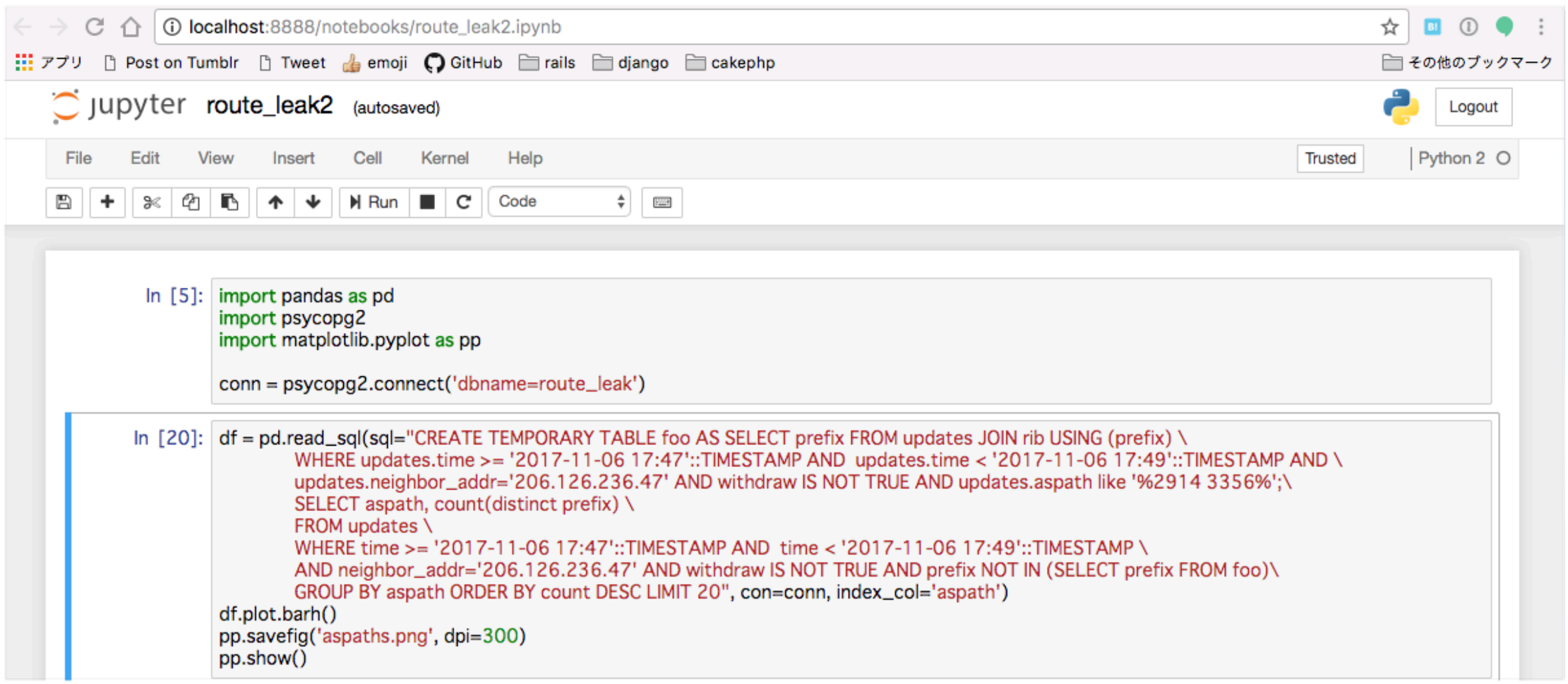
- コマンド例

```
# DB 作成
iw2017 $ createdb -E UTF8 -T template0 route_leak
mkdir -p mrt/eqix
iw2017 $ ruby route_views.rb migrate route_leak

# アーカイブダウンロード
iw2017 $ wget http://archive.routeviews.org/route-views.eqix/bgpdata/
2017.11/RIBS/rib.20171106.1600.bz2 -P mrt/eqix
iw2017 $ wget http://archive.routeviews.org/route-views.eqix/bgpdata/
2017.11/UPDATES/updates.20171106.1745.bz2 -P mrt/eqix

# アーカイブを読んでDBに入れる
iw2017 $ ruby route_views.rb update load route_leak
iw2017 $ ruby route_views.rb rib load route_leak
```

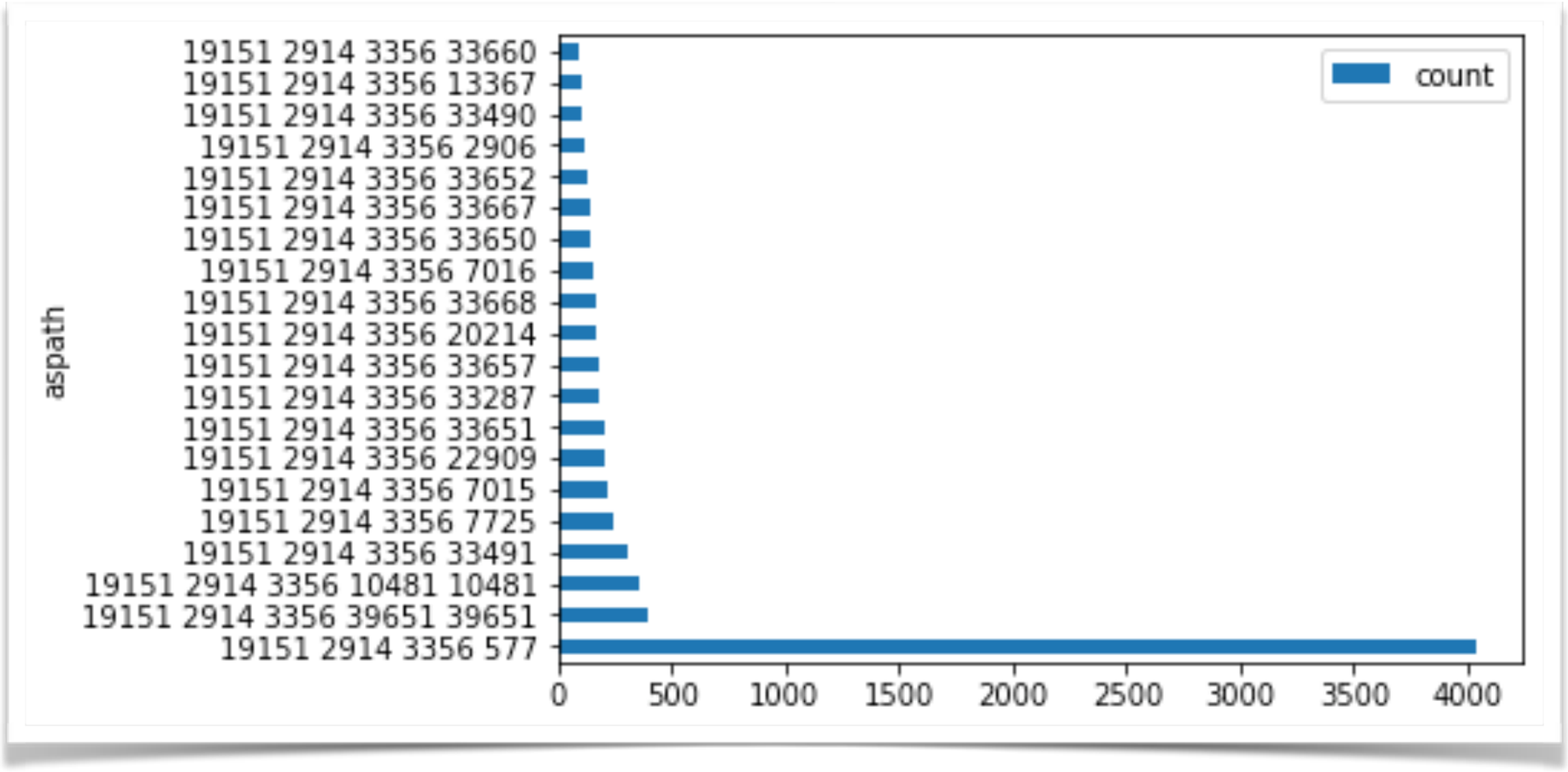
検索する

- jupyter-notebook が便利
- https://www.dropbox.com/s/p5umcql8lw5adgu/route_leak.ipynb?dl=0

```
iw2017 $ pip install jupyter pandas matplotlib psycopg2
iw2017 $ jupyter-notebook
```



AS_PATH ごとのUpdate 数



- 3356 の右隣は、577、39651、10481、33491などなど
- 普段は存在しない経路であるが、リークかどうかはまだわからない

わかったこと

✅ Level3 = AS3356

✅ 広告経路を取ってこれそう

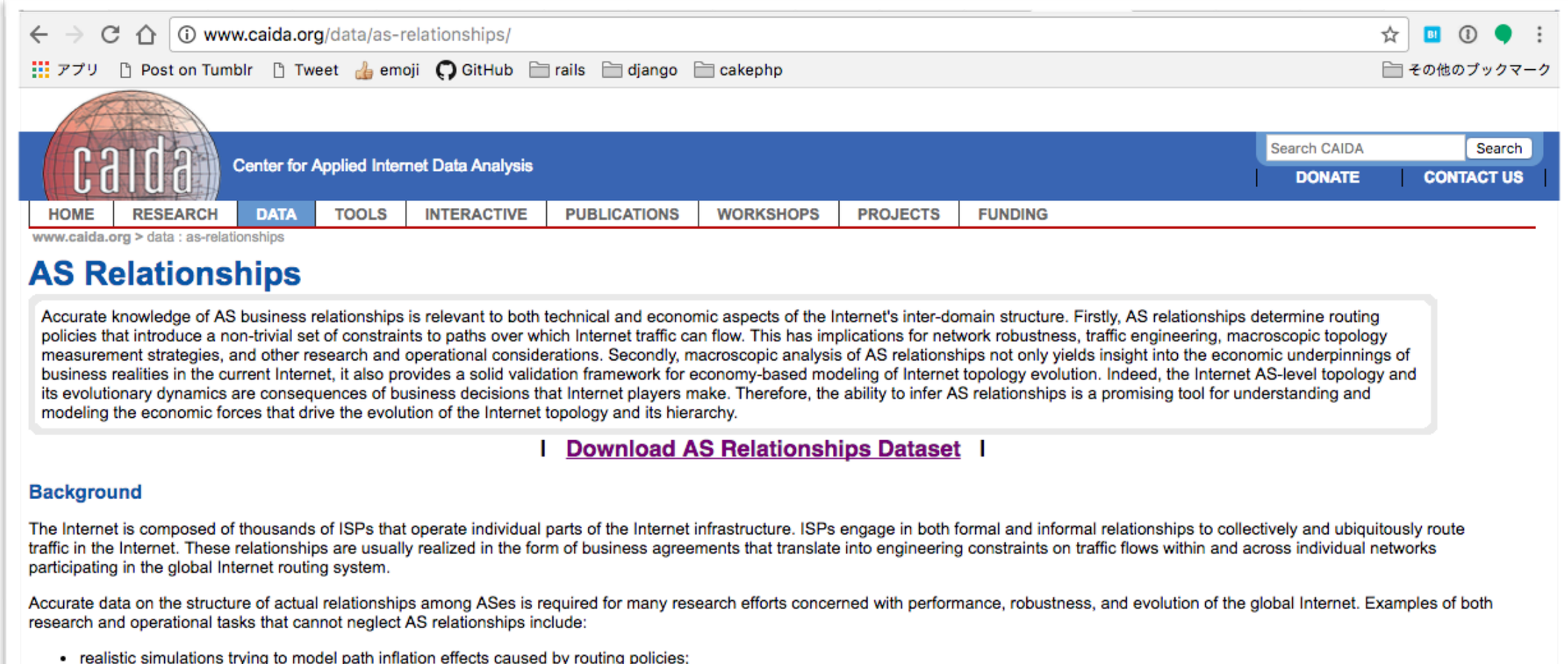
✅ 多数のBGP Update を送っている時刻がある

✅ 普段は存在しない経路を探し、
(Level3から見た) 隣接AS一覧を得た

リーク経路かどうかは まだわからない。
隣接AS と Level3 の関係を推測する

隣接AS と Level3 の関係は？

- CAIDA AS Relationships
- http://www.caida.org/data/as-relationships/
- 精度に問題はあるものの、トランジット / ピア判別に使える



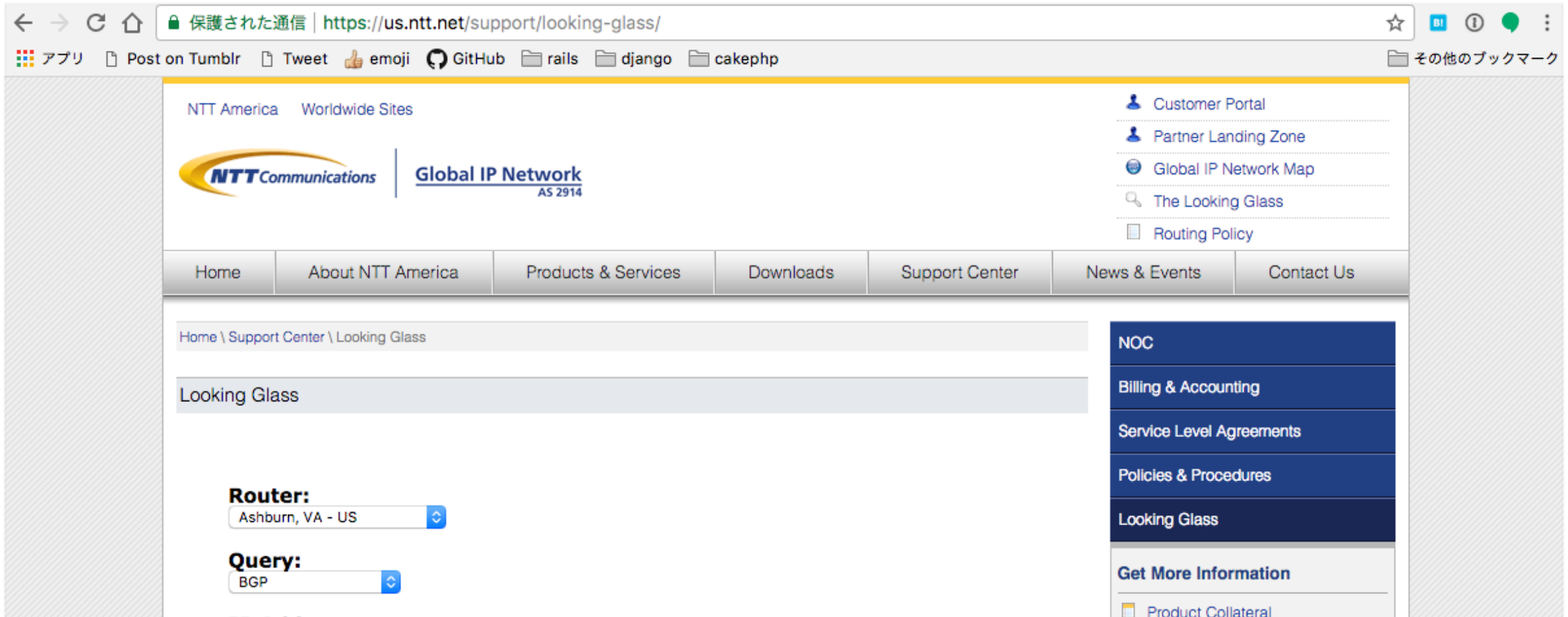
CAIDA AS Relationsihps

```
iw2017 $ grep "577|3356" ~/darwin/doc/archive/as-rel2.171001.txt
577|3356|0|bgp
```

3カラム目	意味
0	AS577 と AS3356 はピア と思われる
-1	AS577 は AS3356 に トランジット提供していると思われる

隣接AS と
Level3 の関係は？

- Looking Glass
- ここではAS2914 のものを使う
- https://us.ntt.net/support/looking-glass/
- 普段、同じAS_PATH の経路がなければLevel3 と該当AS はピアっぽい



これらのASと Level3の関係は？

3356 右隣	AS Relationsihp	Looking Glass
577	ピア	ピア
39651	トランジット	トランジット
10481	トランジット	トランジット
33491	ピア	ピア
7725	ピア	ピア
7015	ピア	ピア
22909	ピア	ピア
33651	ピア	ピア
33287	ピア	ピア
33657	ピア	ピア
20214	ピア	ピア

リークと考えてよさそうな経路

わかったこと

- ✅ Level3 = AS3356
- ✅ 広告経路を取ってこれそう
- ✅ 多数のBGP Update を送っている時刻がある
- ✅ 普段は存在しない経路を探し、
  (Level3から見た) 隣接AS一覧を得た
- ✅ リークとおぼしきAS_PATH を得た

リークの影響範囲は？

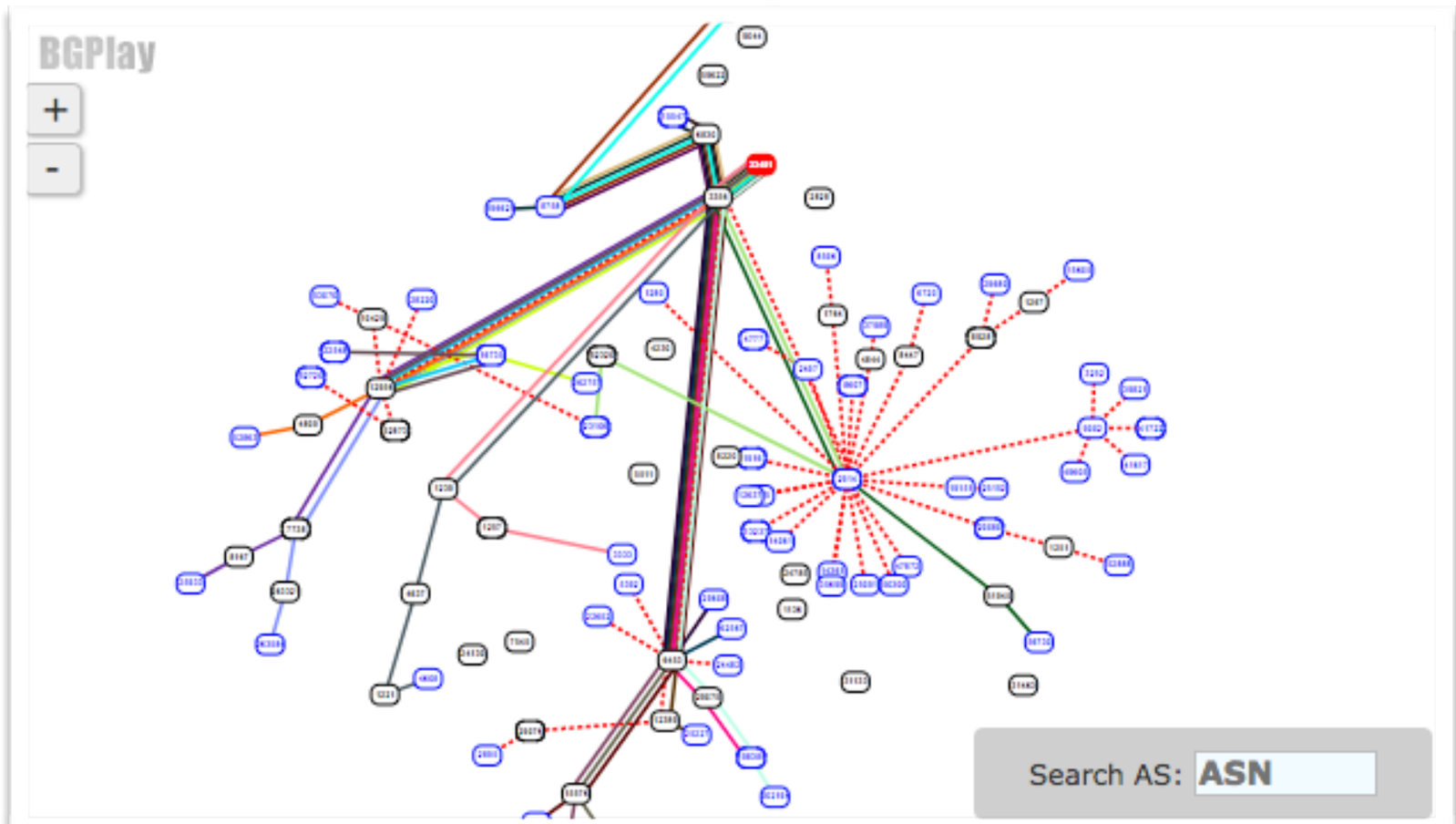
NTTCOM(AS2914) 以外のリーク先は？

- RIPE BGPlay
- https://stat.ripe.net/widget/bgplay#w.resource=23.25.0.0/18&w.ignoreReannouncements=false&w.starttime=1509926400&w.endtime=1510012800&w.rrcs=0,1,2,5,6,7,10,11,13,14,15,16,18,20&w.instant=null&w.type=bgp

NTTCOM(AS2914) 以外のリーク先は？

- RIPE BGPlay
- 適当な経路をひとつえらぶと、TELEFONICA (AS12956)、TATA (AS6453)が見える。

→ 影響の大きさが想像できる 💦



わかったこと

- ✅ Level3 = AS3356
- ✅ 広告経路を取ってこれそう
- ✅ 多数のBGP Update を送っている時刻がある
- ✅ 普段は存在しない経路を探し、
  (Level3から見た) 隣接AS一覧を得た
- ✅ リークとおぼしきAS_PATH を得た
- ✅ Level3がどのASにリークしたか

ここまでのまとめ

「 11/6 に Level3 が経路リークした 」くらいの情報から、何が起こったかはだいたい調査可能

- 時刻
- 経路アトリビュート
- 仲介したAS
- 通信状態を調べるのは困難

8/25 のリークも同様に調査できる

- IRS27 BGP経路問題発生時の行動を考えよう
  - https://speakerdeck.com/codeout/bgpjing-lu-wen-ti-fa-sheng-shi-falsexing-dong-wokao-eyou-as-nakutemoda-zhang-fu-da

遠くで起こった経路障害について、
直接影響なくても何が起こったかを推測することができる

同じ障害は起こる。次は大丈夫か？

- 影響なかった？ どういう原理で？
  - リーク経路が流入しなかった。止めてくれたのは誰？
  - 流入したが耐えた。閾値はどれくらい？
- 影響あった場合のインパクトは容認可能？

リスク評価することが重要

まとめ

- 経路障害について、2017年のトレンドをふりかえった
- 経路障害について、自社に影響がなくても調査できる。その方法を紹介した

Questions ?