

# 色々なトラヒック制御の 利点と欠点

Matsuzaki 'maz' Yoshinobu

<maz@ij.ad.jp>

# トラヒック制御

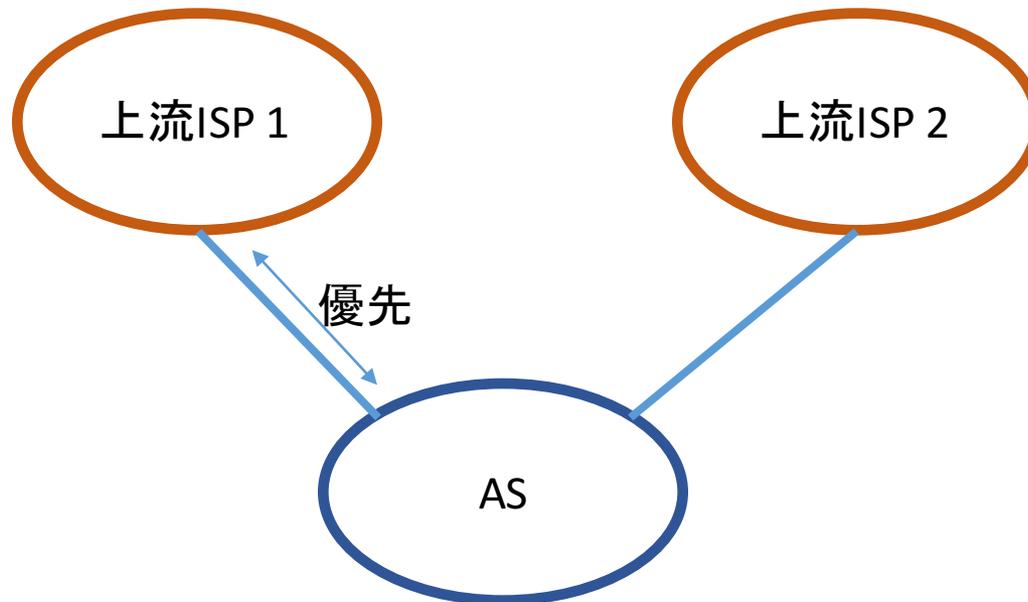
- 無限の帯域を持ってはいない
  - 限りある資源を有効/思ったように使いたい
- 物理法則を超えられない
  - 通信距離や通信経路にもよるけど遅延を短くしたい
- ユーザの利用が変化する
  - 変化に追従する必要がある
- いろいろな都合を鑑みる必要がある
  - 特定の回線を優先的に使ったり、特定の経路を曲げたりする必要がある

# BGP

- インターネットでは事実上ネットワーク間の標準的経路制御プロトコル
- ポリシーに基づいた経路制御が可能だとか言われているけど、細やかな制御には向いていない
  - ざっくりとした経路制御は可能

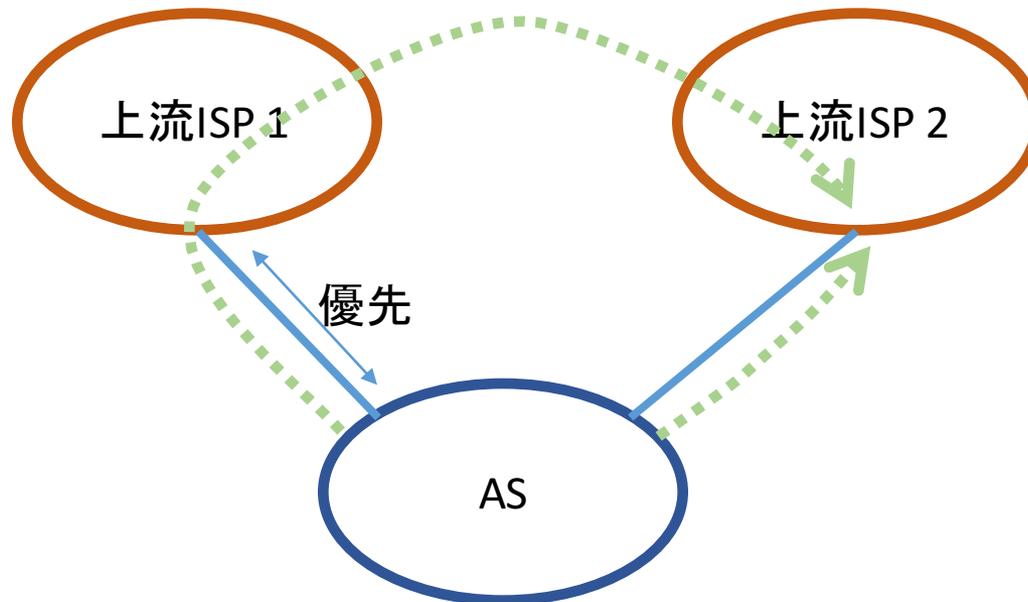
# 例えばこんな制御

- 通常時はISP1経由を優先(ISP2もISP1経由で通信)
- 障害時にはISP2に切り替え



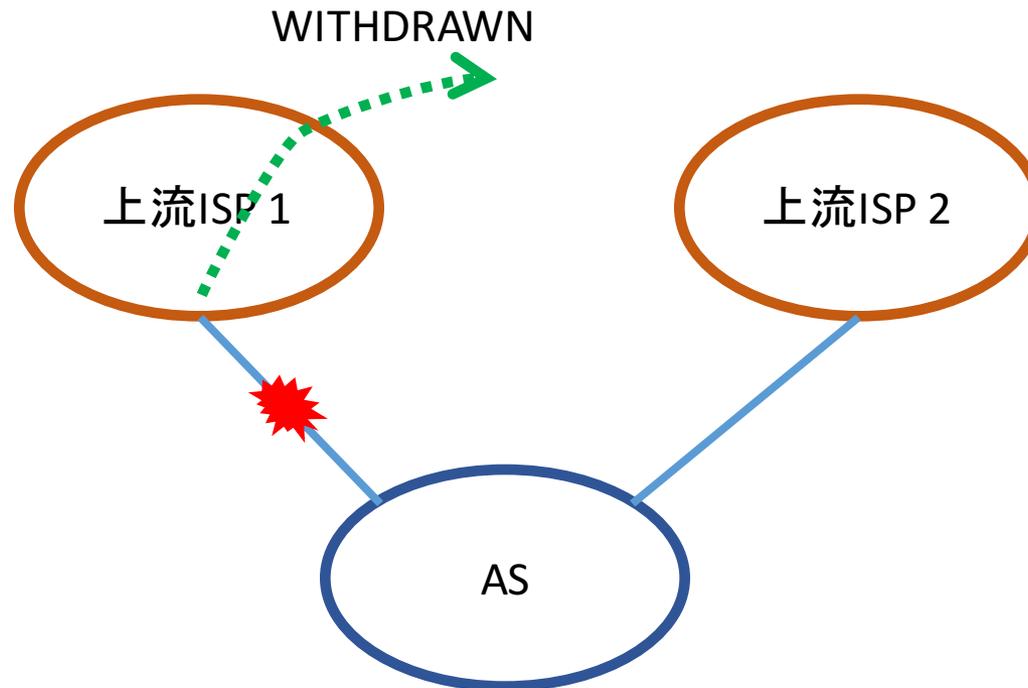
# 通常時の経路状態

- ISP1経由の経路広報がISP2まで到達
- ISP2でもほとんどのルータはISP1経由を優先



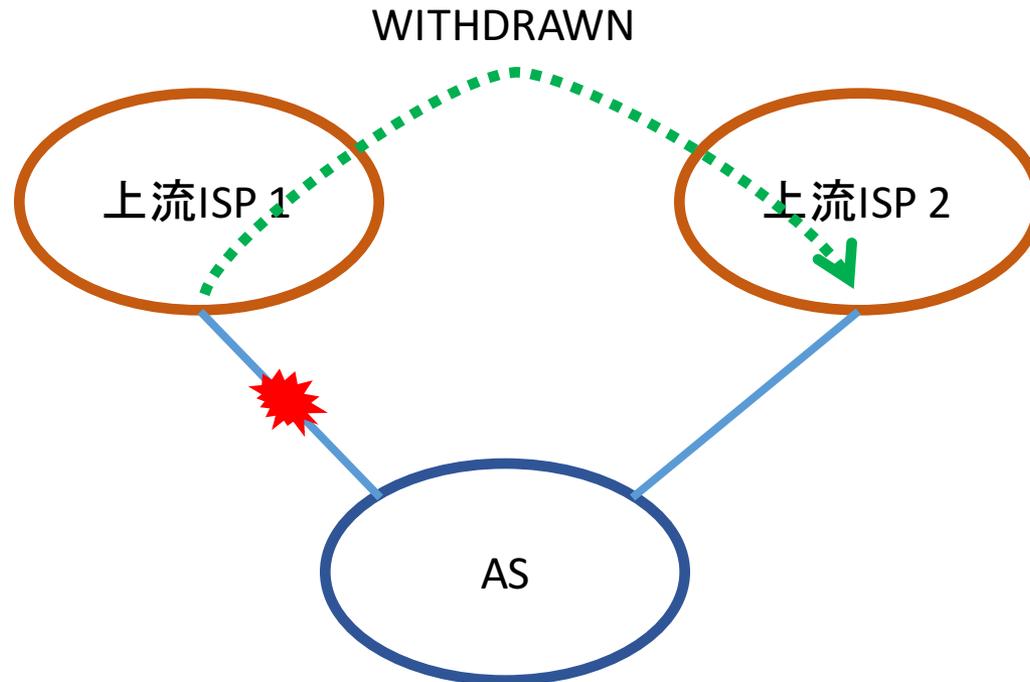
# 障害時の流れ

## 1. ISP1がWITHDRAWNを広報



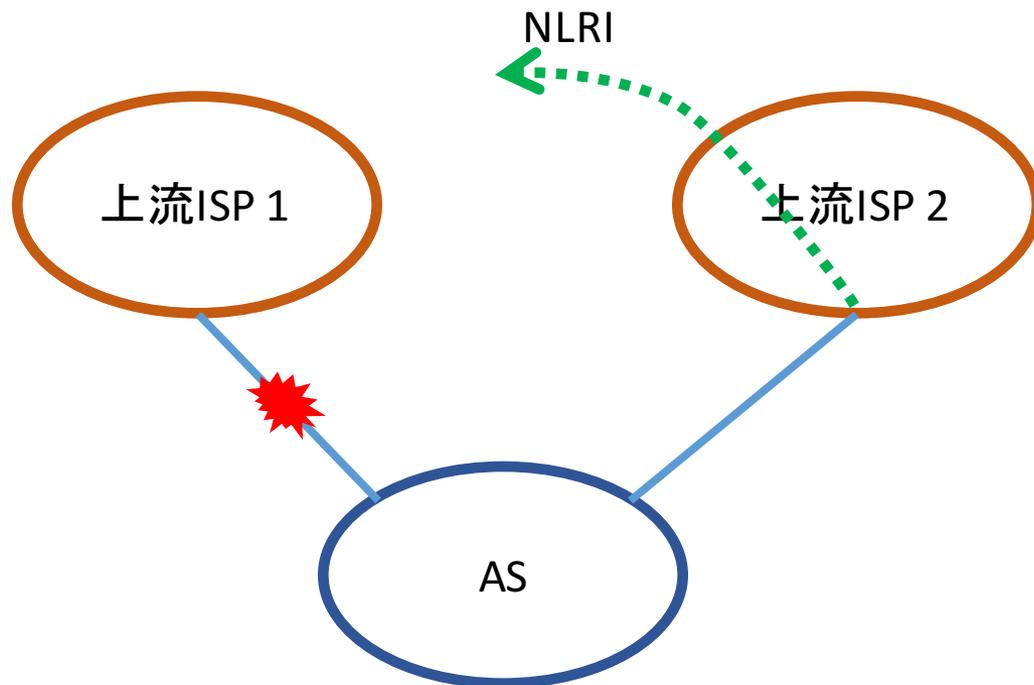
# 障害時の流れ

## 2. WITHDRAWNがISP2に到着



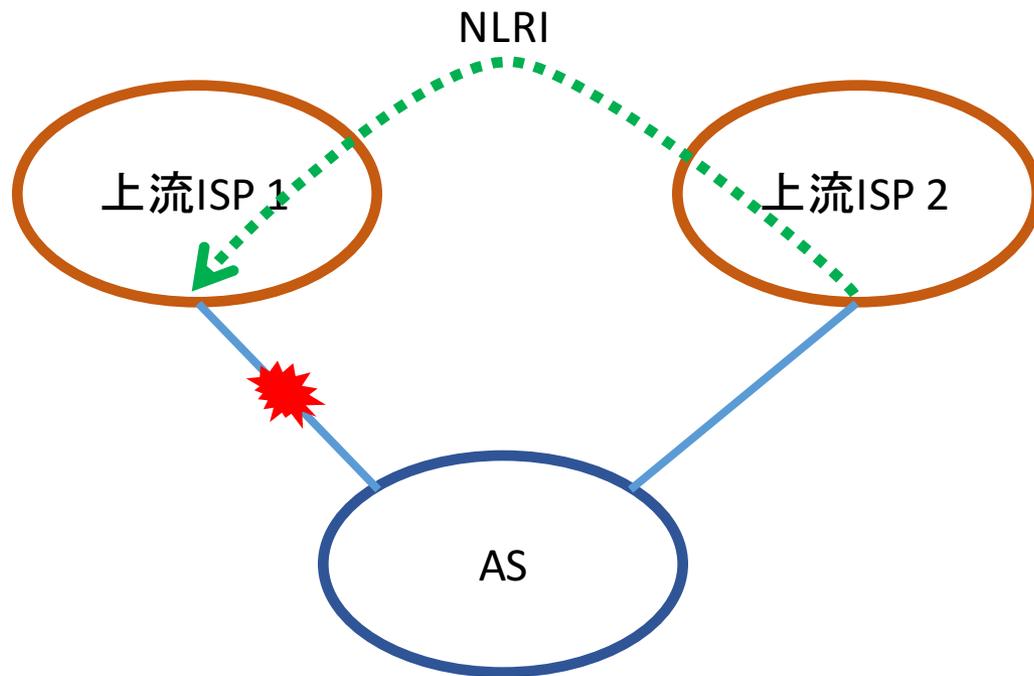
# 障害時の流れ

## 3. ISP2がNLRI(到達性情報)を広報



# 障害時の流れ

## 4. 経路広報がISP1にも到着

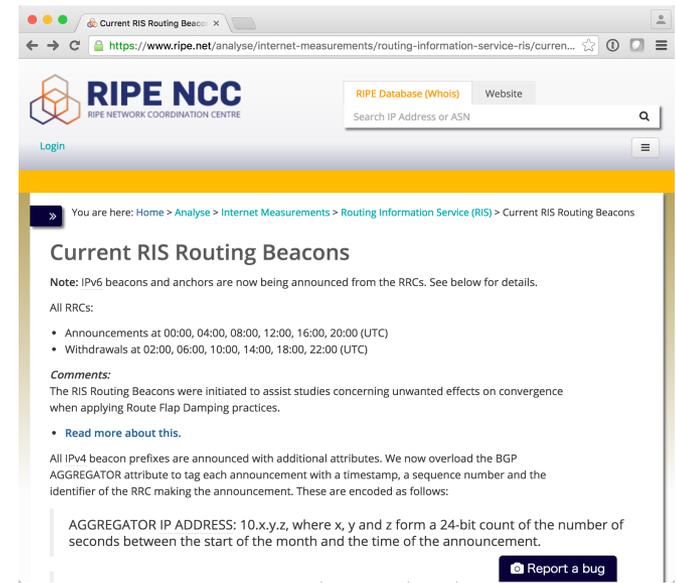
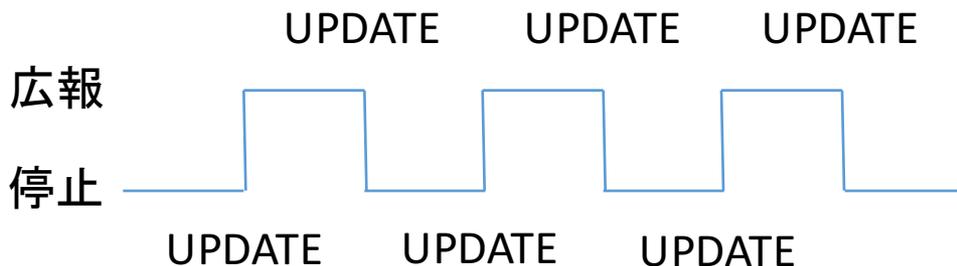


# 収束時間を見積もってみる

- ひとまずBGP UPDATEの収束時間
  - 本当はFIBにエントリが入って、みんなが通信できる状態にならないと意味ないけど、ここではざっくり概要を見ることに注目
- 収束の見極め
  - ざっくり300秒以内で最後に受信した関連BGP UPDATE (announce[経路広報]もしくはwithdrawn[経路削除])のタイミングで収束だと見なした

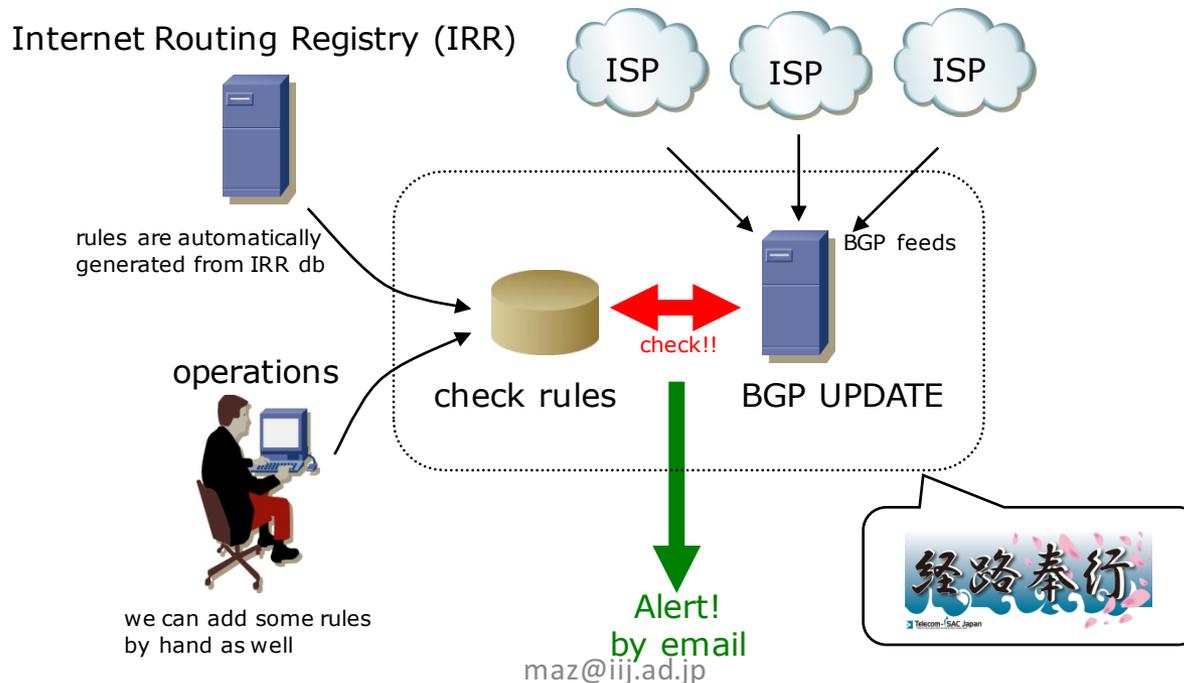
# BGP BEACONを見れば推定できる

- RIPE NCCのRIS Routing Beacons
  - <https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris/current-ris-routing-beacons>
- 世界各地からBGP UPDATE送出
- 2時間毎に広報/停止を繰り返し

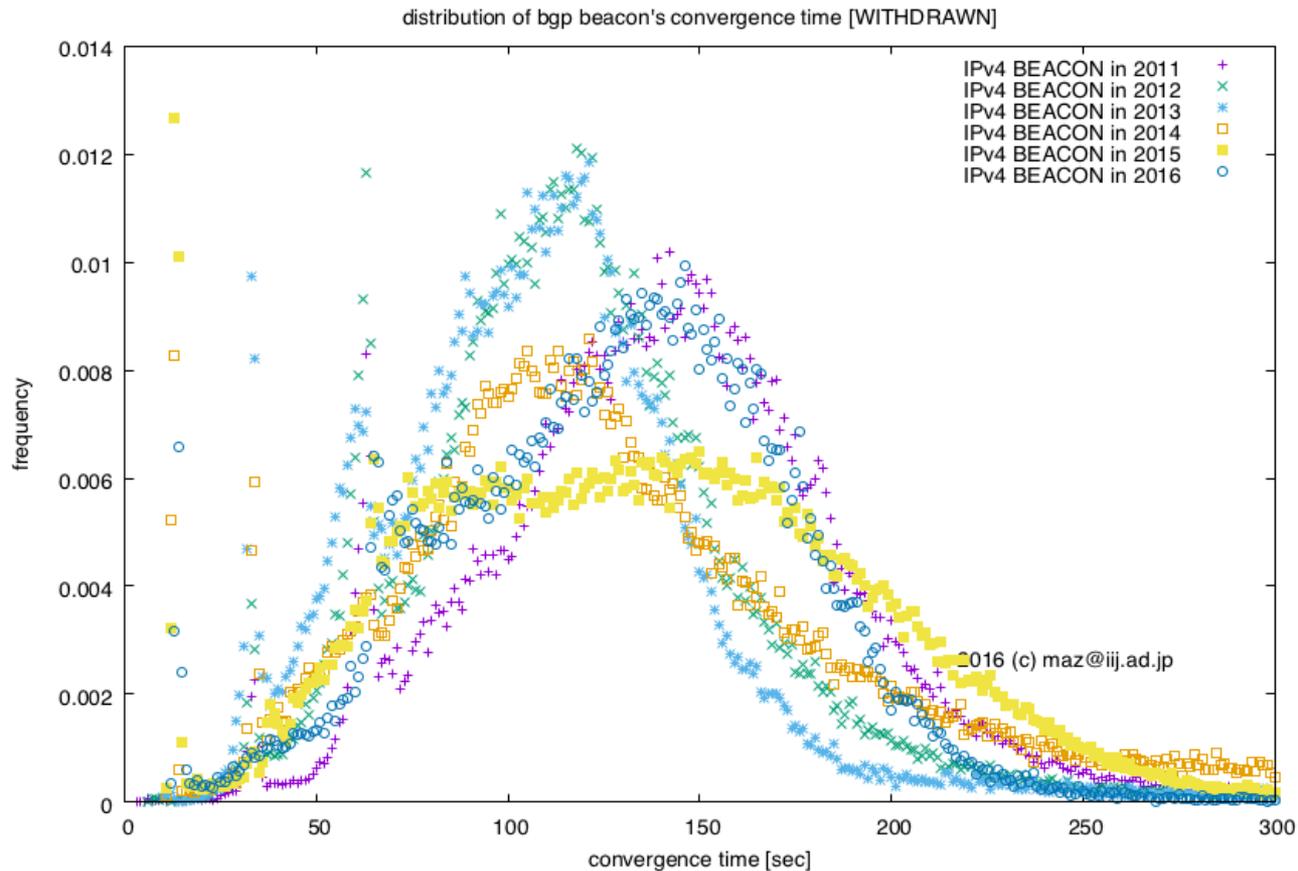


# 経路奉行にデータがあったよ

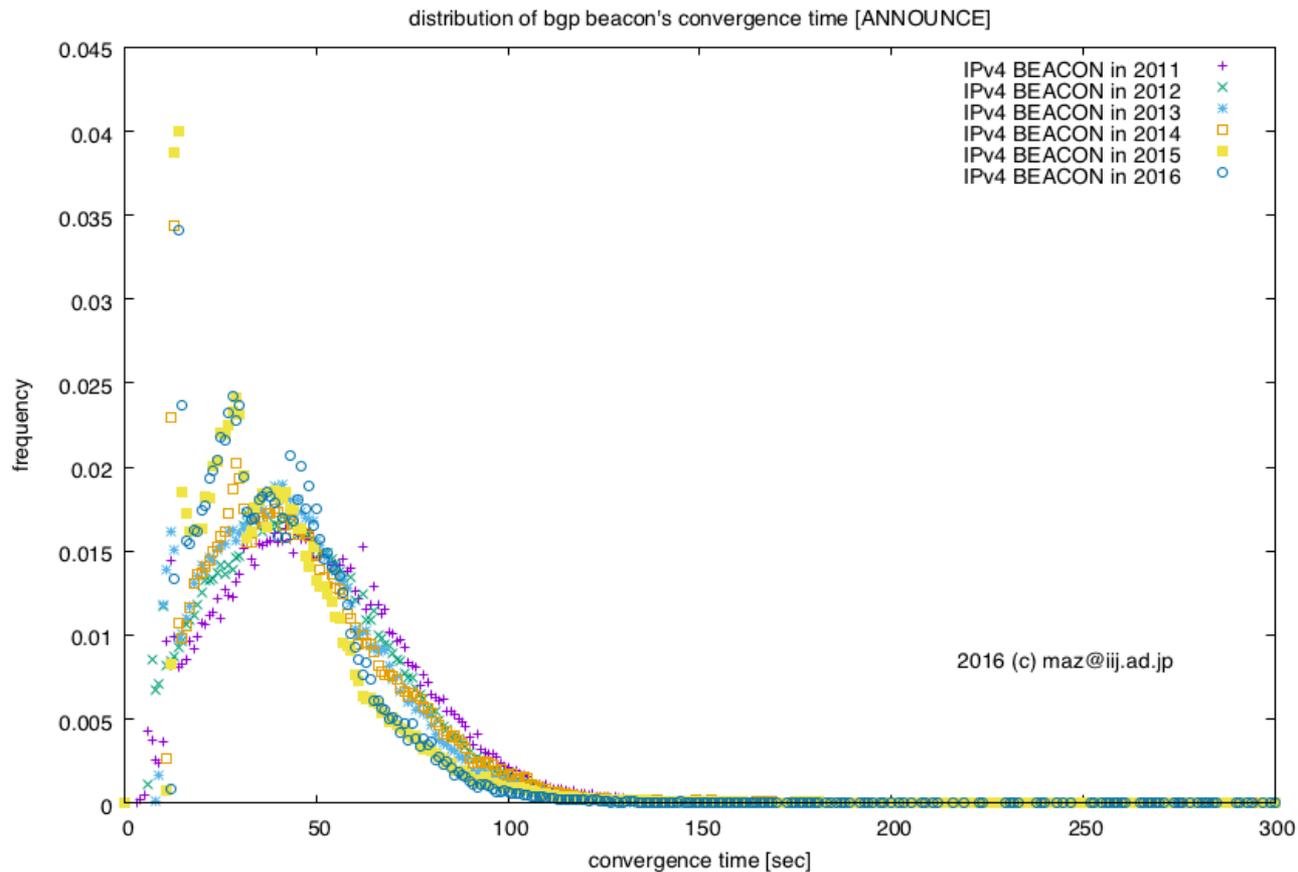
- 経路奉行は、Telecom-ISAC Japanの経路情報共有WG(BGPWG)によるプロジェクト
- 蓄積されたUPDATEから該当の約480万件を抽出



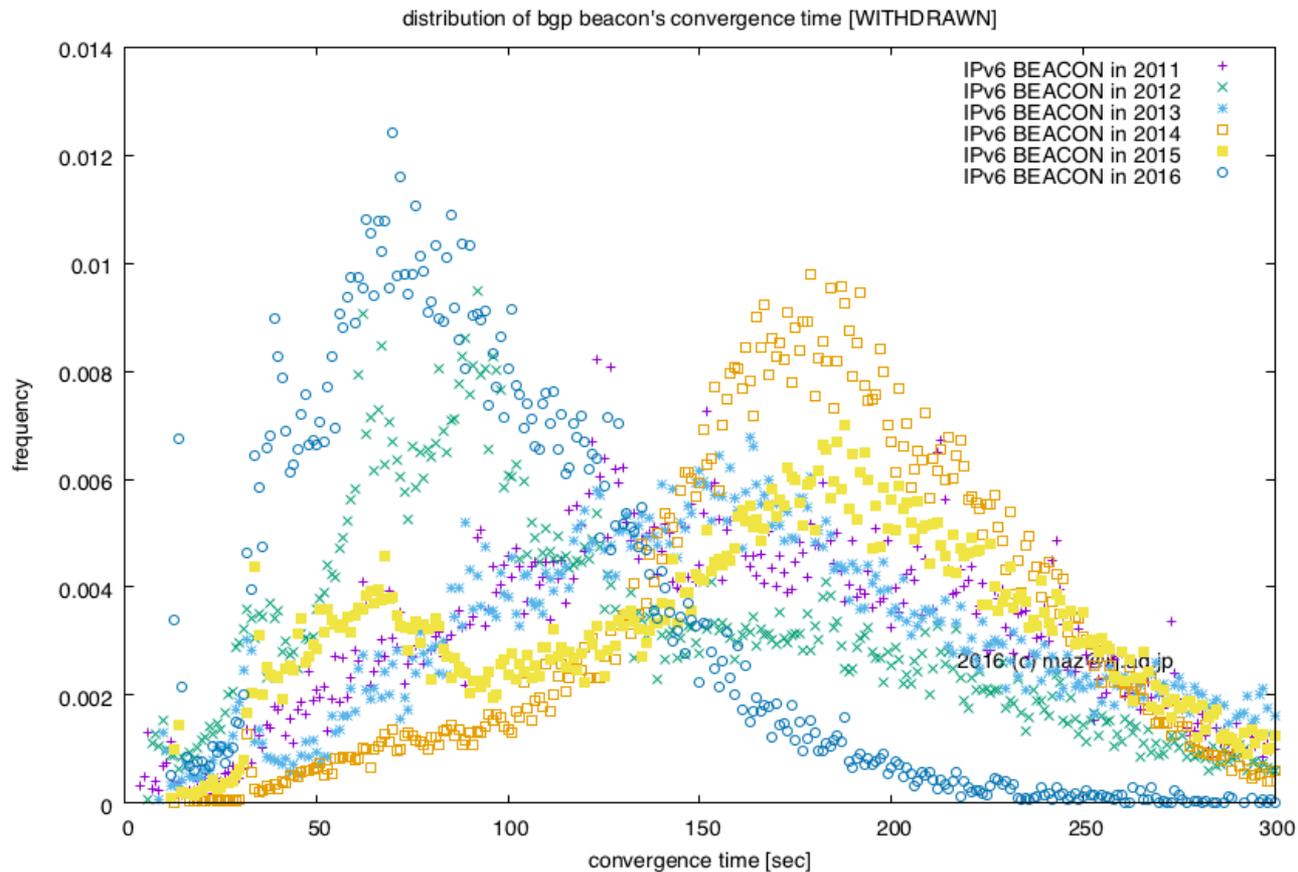
# 年次別 IPv4の経路削除収束時間



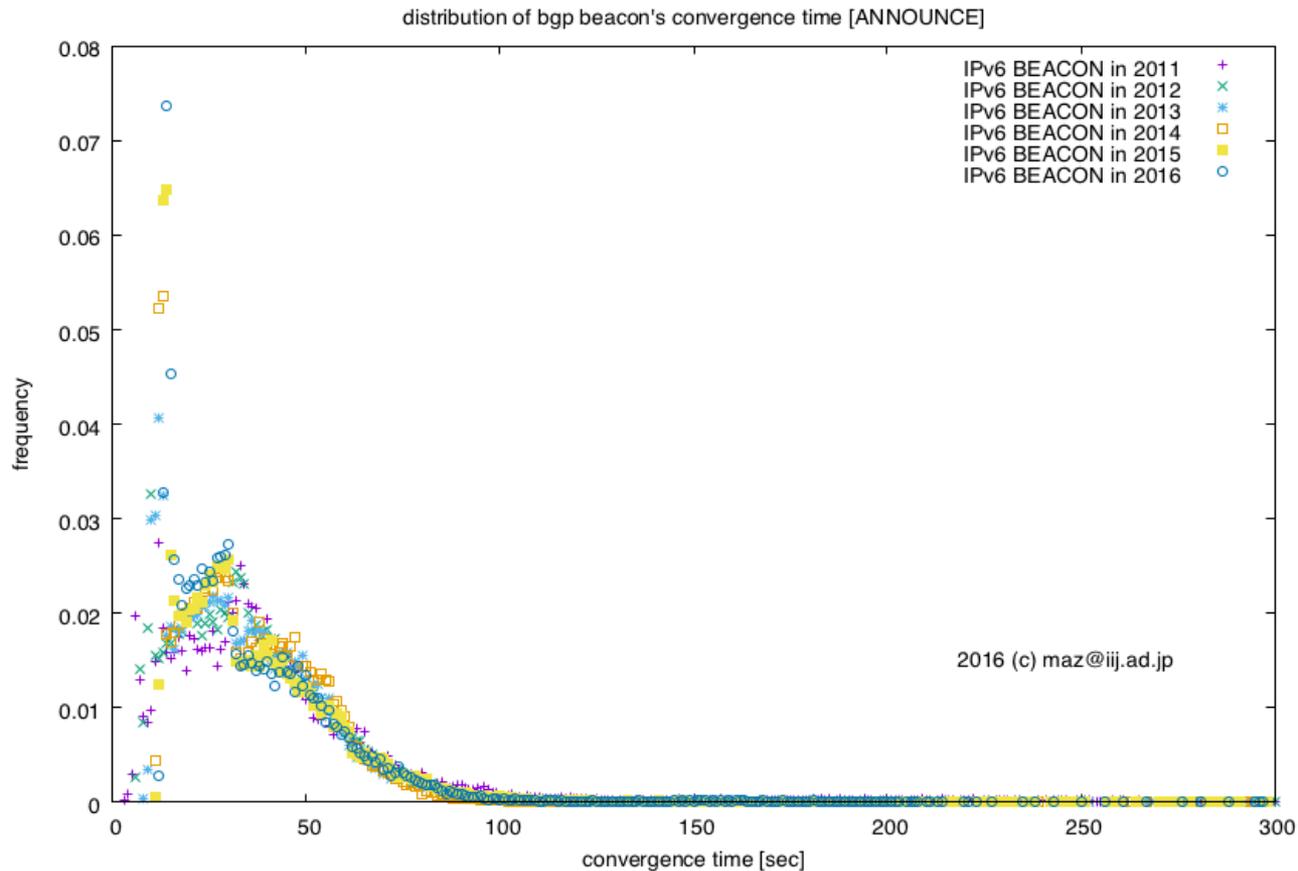
# 年次別 IPv4の経路広報収束時間



# 年次別 IPv6の経路削除収束時間



# 年次別 IPv6の経路広報収束時間

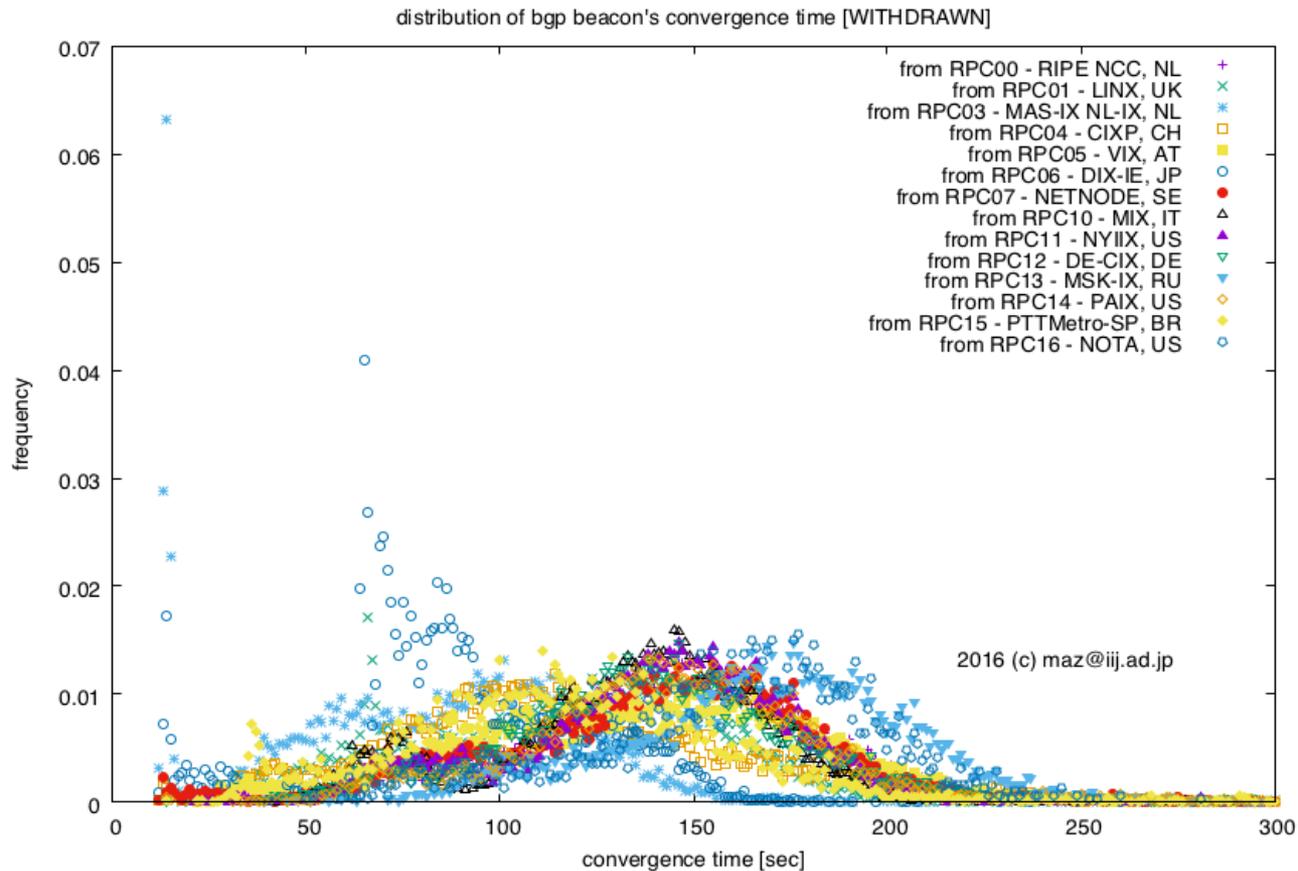


# 収束時間

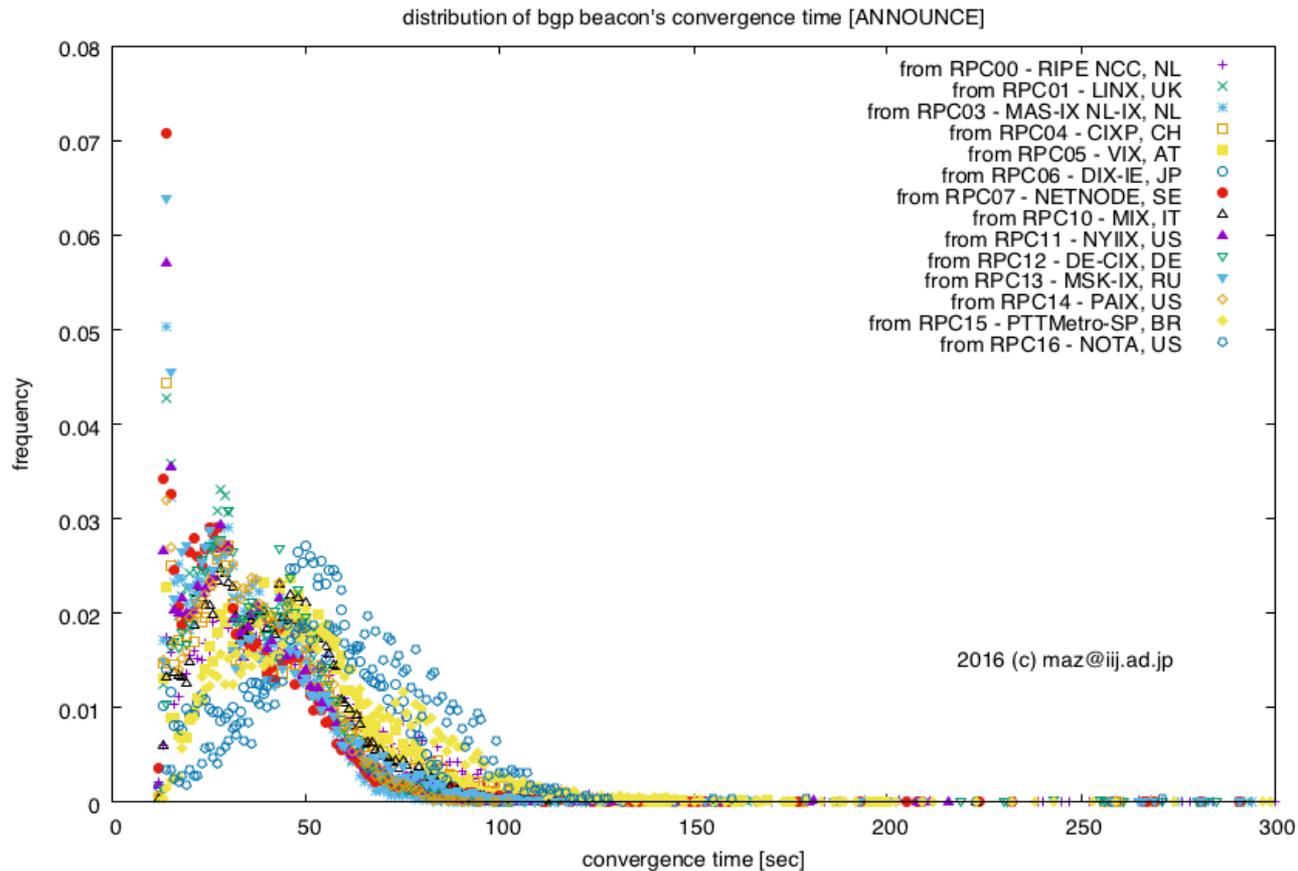
- 新規に経路広報するANNOUNCEより、経路削除するWITHDRAWNの方が収束に時間がかかる
  - 最適経路が削除されると、トランジット経由などの利用可能な経路が見え隠れし、すべての利用可能な経路が消されるまで時間がかかる
- 収束時間はUPDATEが経由するルータの実装や相互接続関係に依存する
  - MRAI
  - トランジット、ピア、カスタマー
- なので、細かな分散具合にはあまり意味がない

# 広報元ノード別

## IPv4 WITHDRAWN (2015+2016)

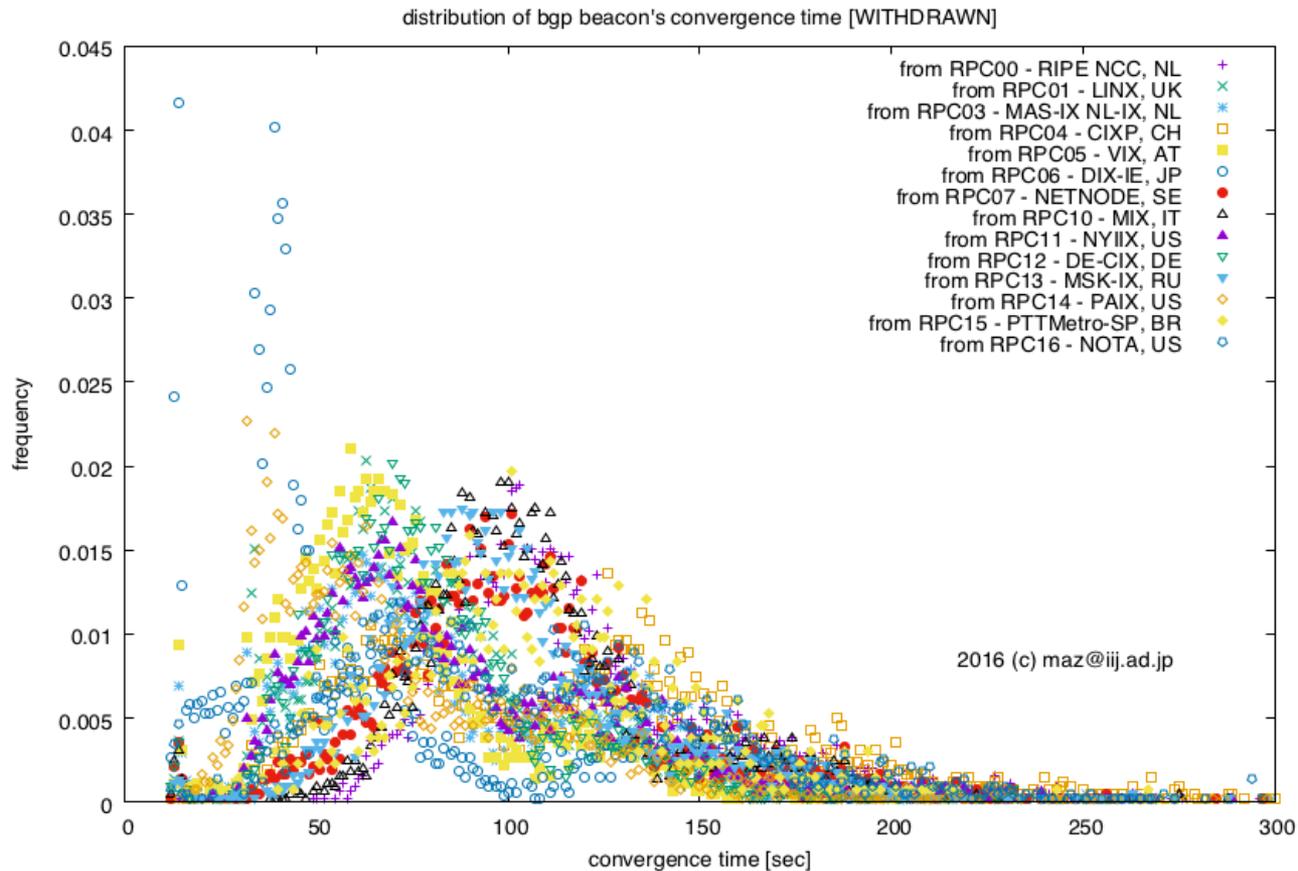


# 広報元ノード別 IPv4 ANNOUNCE (2015+2016)

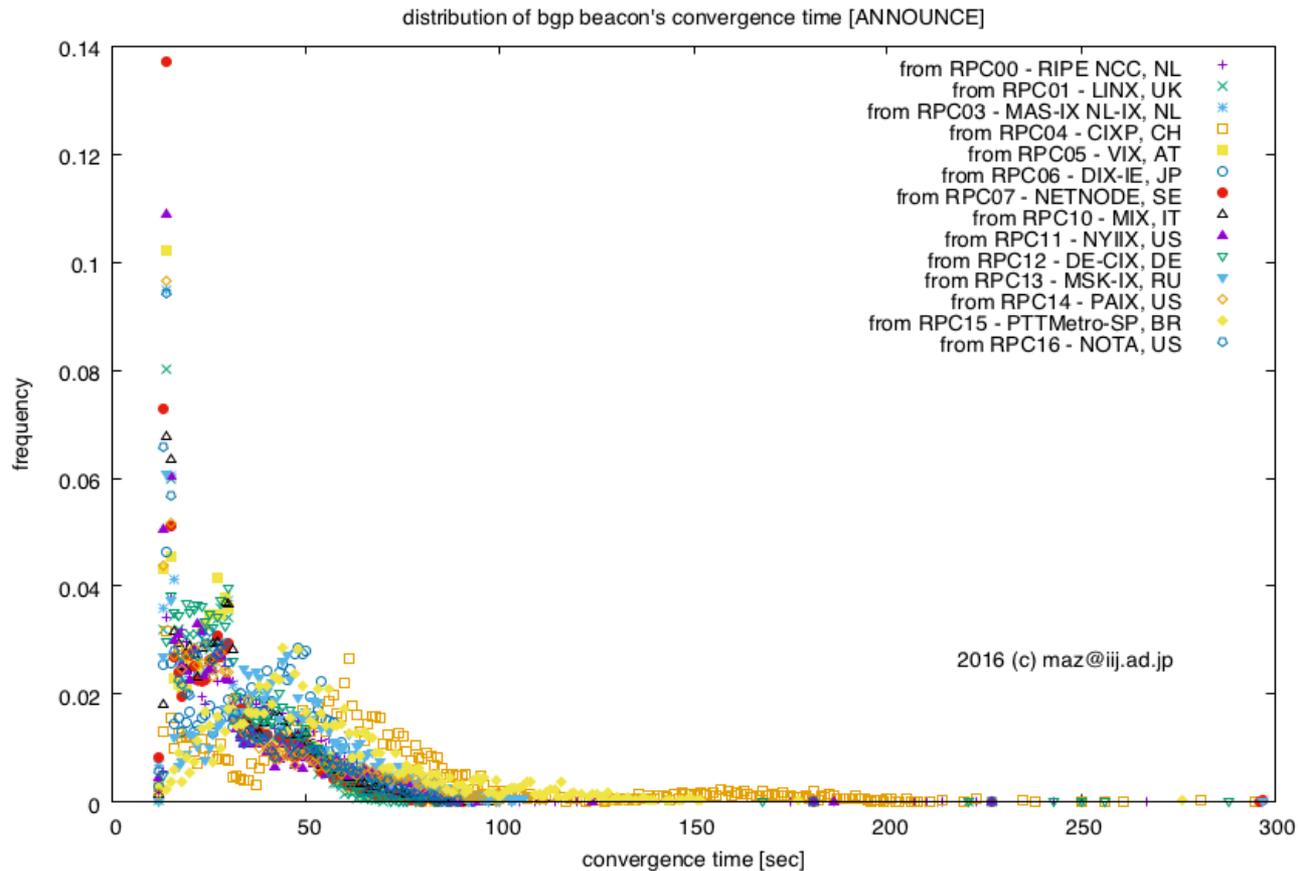


# 広報元ノード別

## IPv6 WITHDRAWN (2015+2016)



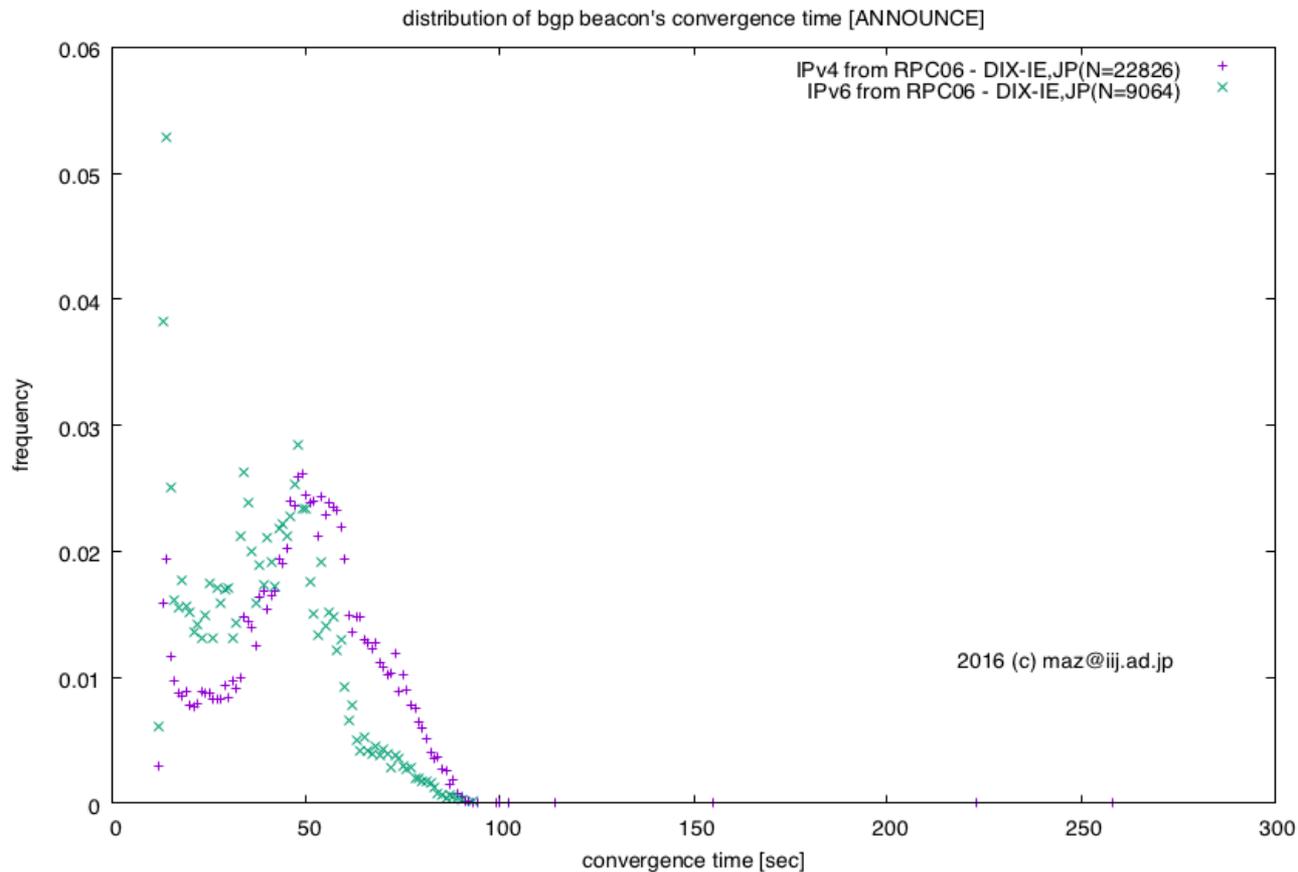
# 広報元ノード別 IPv6 ANNOUNCE (2015+2016)



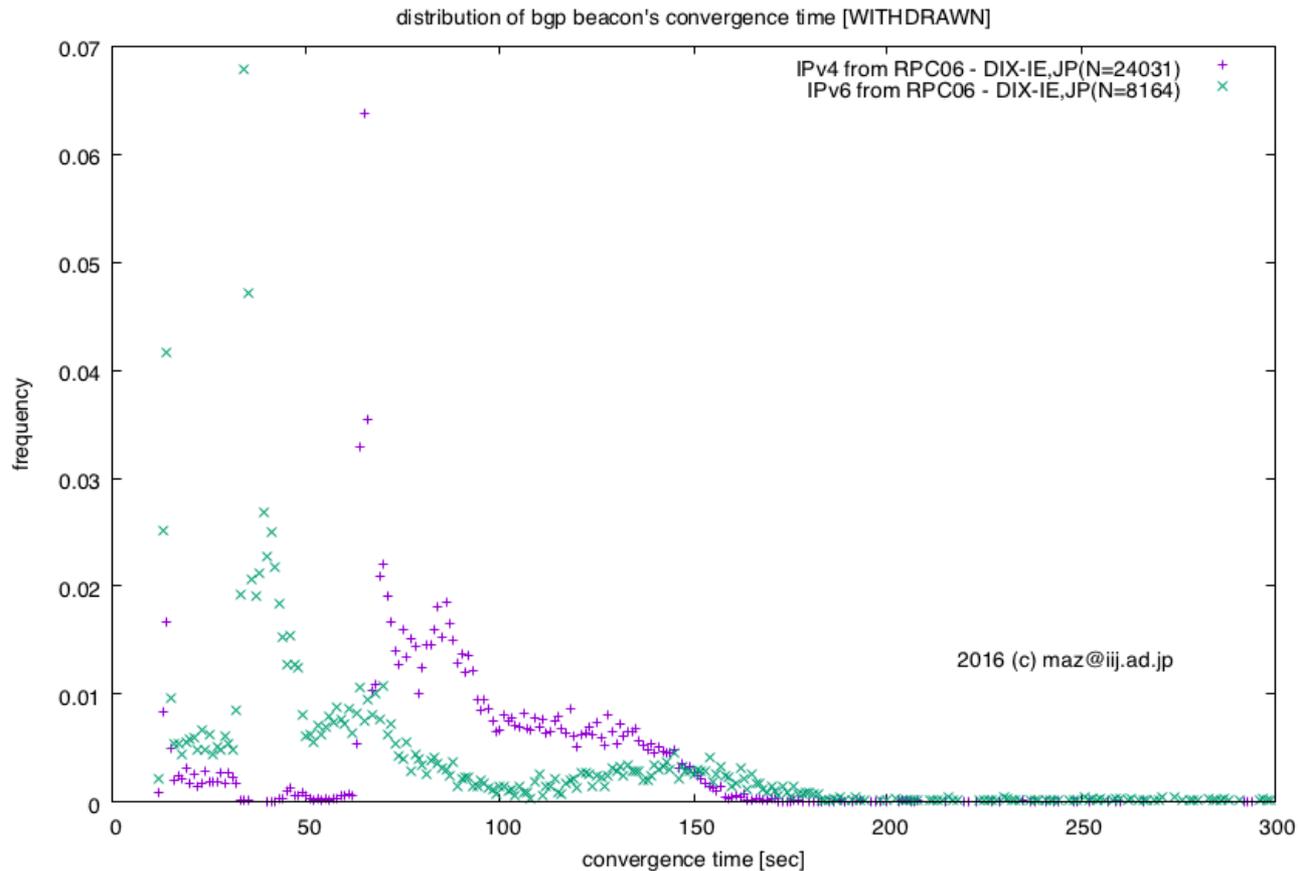
# 広報元別に見ると

- 遠方からの広報ほど収束に時間が掛かる
  - BGP UPDATEが届くまでに経由するAS/ルータが増える
  - WIDTDRAINでは削除されなければならない選択肢が増えるため、収束に時間が掛かる

# 東京ノードからの収束時間 ANNOUNCE (2016年分)



# 東京ノードからの収束時間 WITHDRAWN (2016年分)

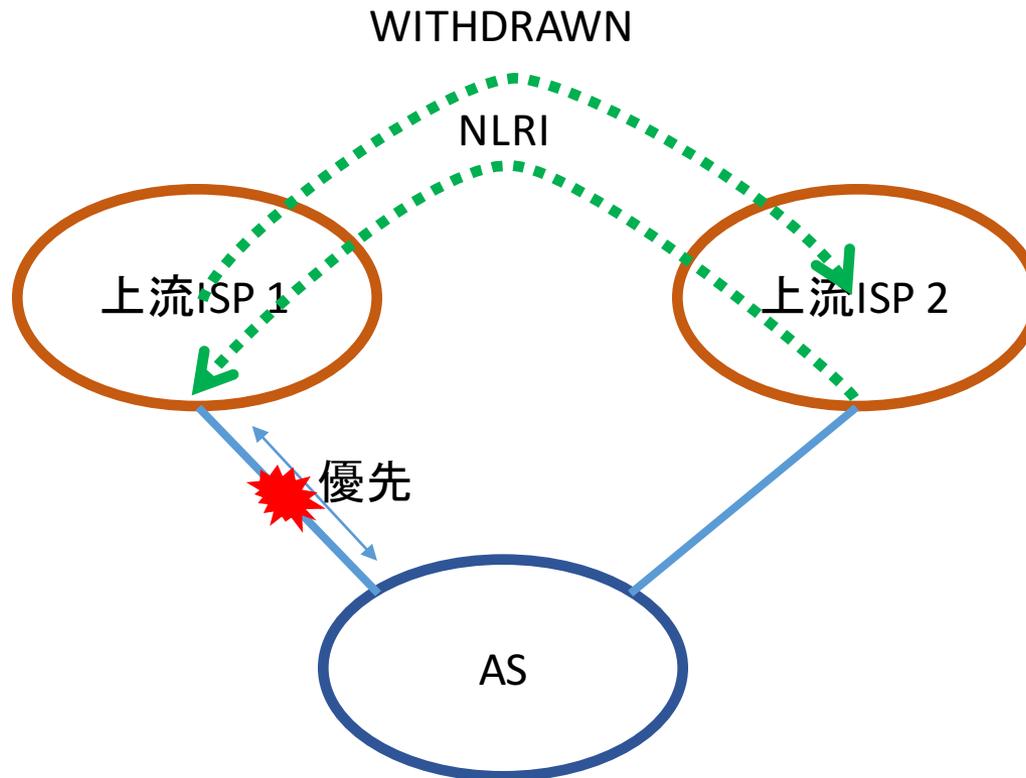


# 東京ノードからの経路収束時間

- 経路広報(announce)に**90秒**
- 経路削除(withdrawn)に**180秒**
  
- この値を参考に国内での経路切替時間を推定

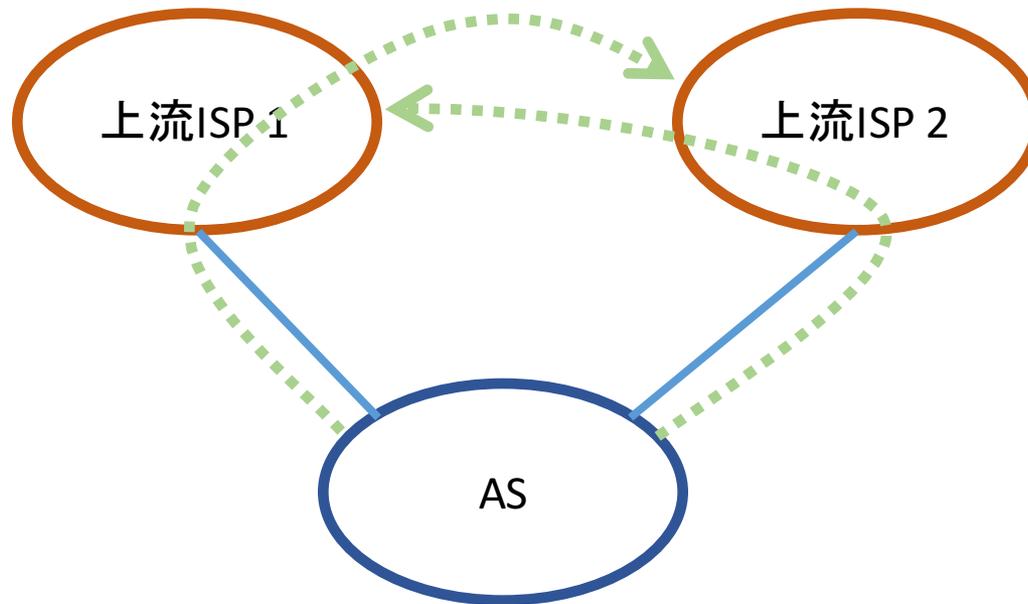
# 国内だと切り替わりに270秒

- withdrawn(180)秒+announce(90)秒



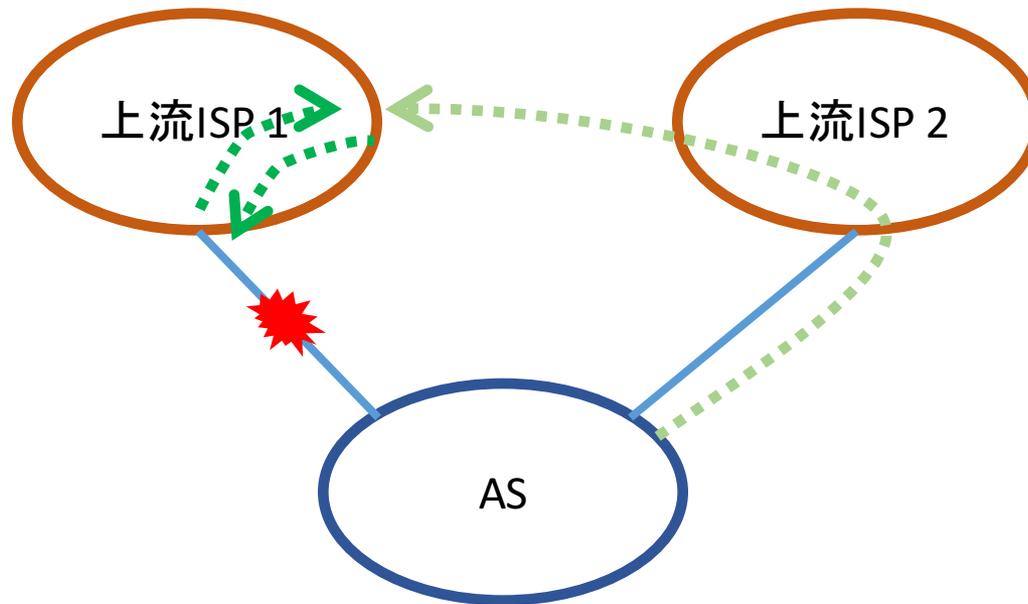
# 普通に経路広報している

- ISP1,2はそれぞれ直接接続した顧客経路を優先
- それぞれ最適経路として他のISPにも広報



# 経路障害時

- ISP1内で経路収束すれば良いため、比較的高速
  - 期待値としては数秒～数十秒

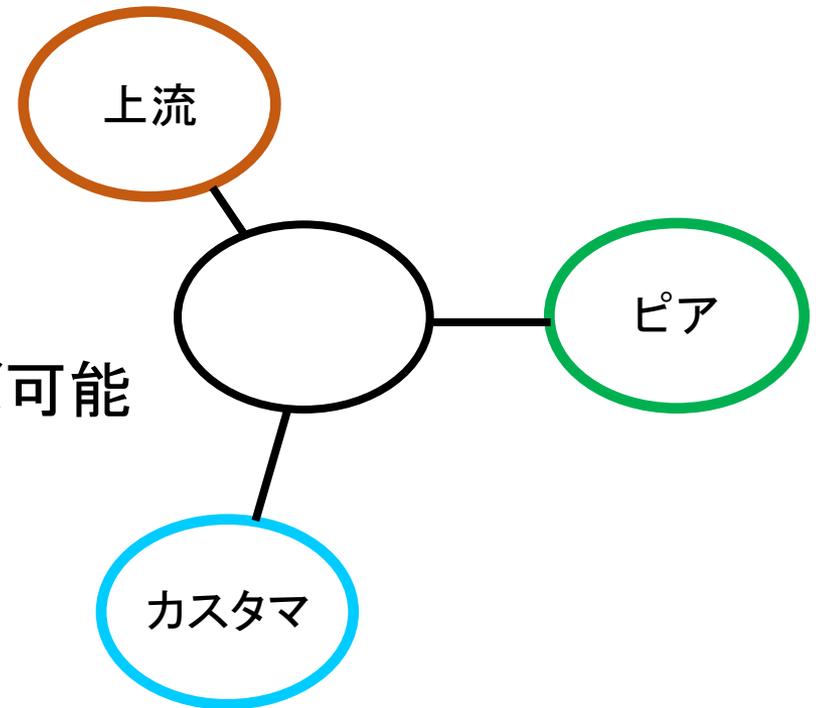


# BGPでの経路制御

- 細かい制御には向かない
  - 頑張ってもprefix単位
  - 流入方向のトラヒックを制御するのは、とても難しい
- 必要な帯域が確保できるまでの‘つなぎ’手段ぐらいに考えるのが無難
  - 需要に対して解離のある状態でネットワークを運用すると非効率になってくる
  - 基本的にあるがままに運用した方が期待通りに動く

# BGPでトラフィック制御できる範囲

- 対上流
  - 比較的簡単
  - お金を払ってるので応談
- 対ピア
  - あまりできない
  - お互いの利害が一致すれば可能
- 対カスタマ
  - 普通カスタマのポリシー優先
  - 事情があれば応談？



# ピア

- 他<sup>o</sup>のASと相互接続して直接トラフィック交換
  - トランジットコストを下げる
  - 品質の担保ができるようになる
- コストは接続形態による
  - 相互接続回線
  - IXP接続費用
  - ピア<sup>o</sup>収容ルータのポート費用

# ピアのモデル

- お互いに似たようなネットワークを想定
  - トラフィックパターン
  - ユーザ層、利用形態
  - トラフィック制御手法
  - 故に相互接続でお互いにネットワーク機能を拡張

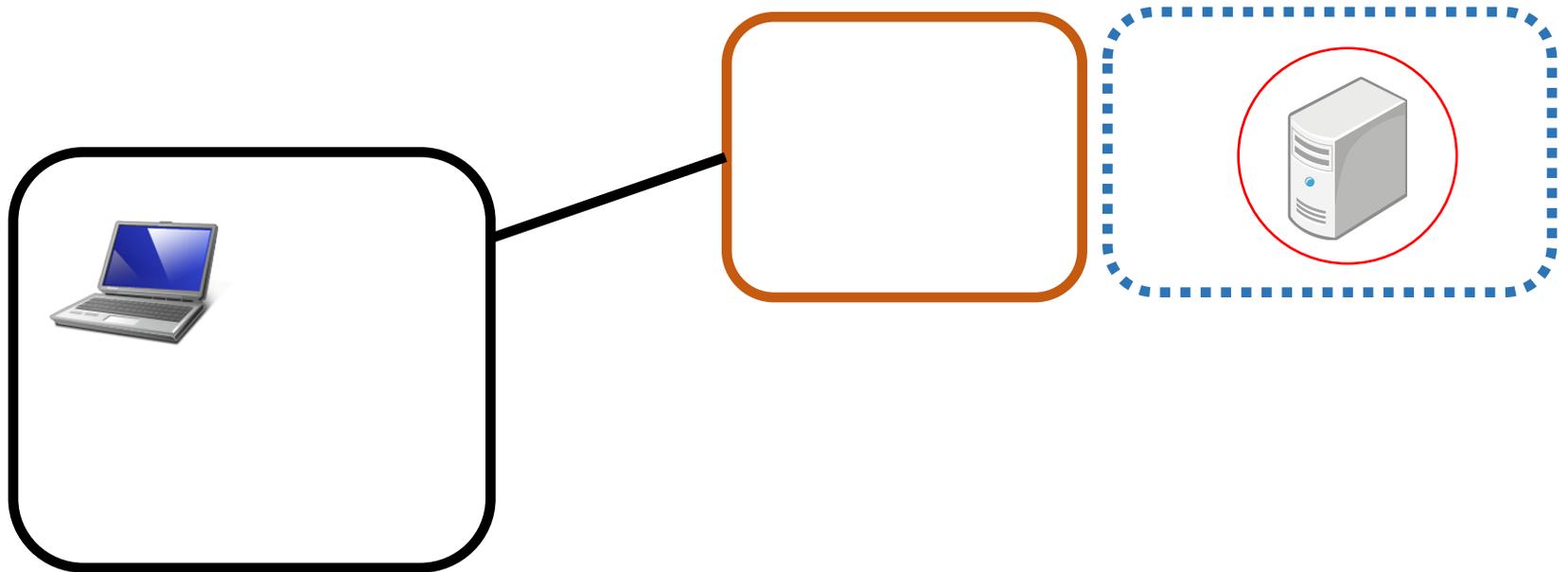
# 有償ピア

- メリットの不均衡感をお金で埋める
- 規模が小さい方がお金を払え
- 少なく運んだ方がお金を払え
- トラヒックを吐く方がお金を払え

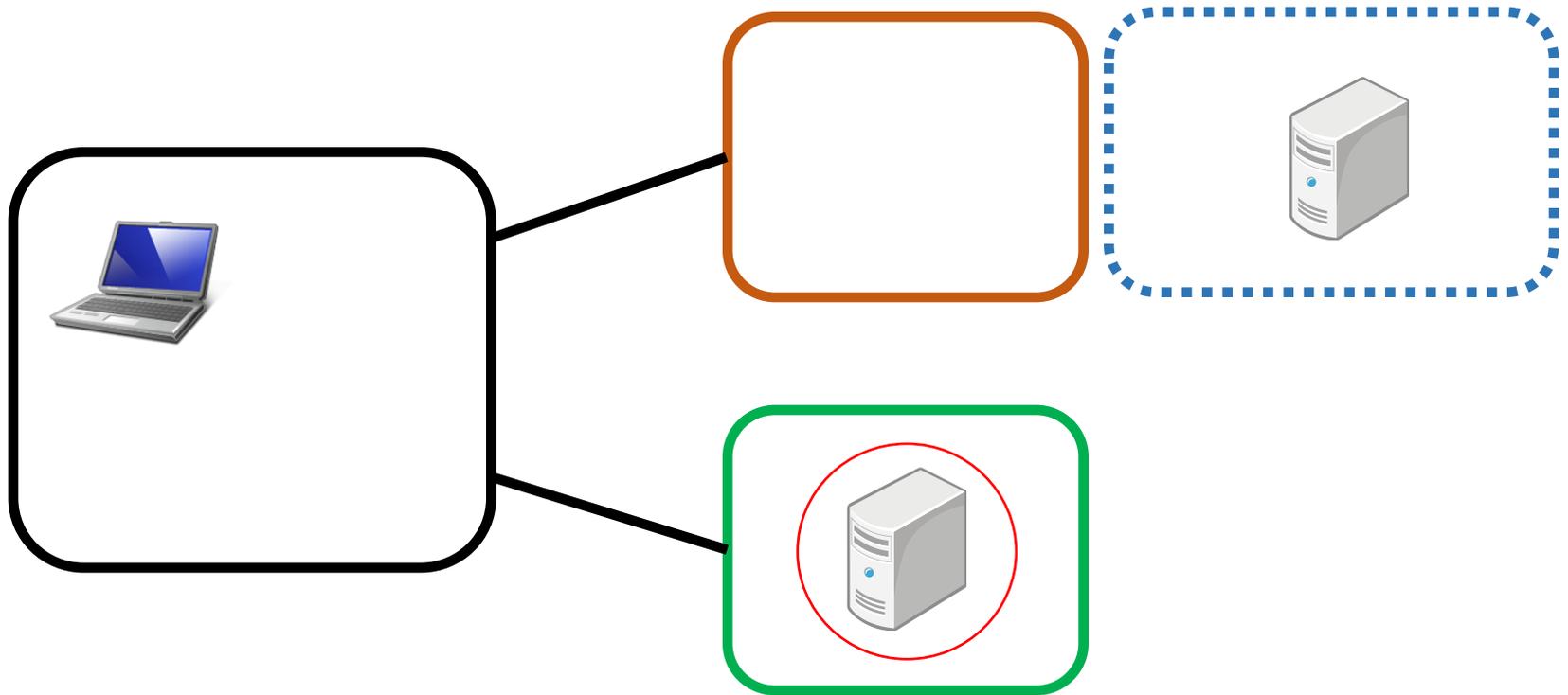
# CDN キャッシュ

- CDNが提供するキャッシュ箱を導入
  - トランジットコストを下げる
  - ユーザがコンテンツを近傍から受け取れるようになる
- コストは規模次第
  - キャッシュ箱向けラック代、電気代
  - 収容スイッチ
  - 収容ルータのポート費用
- トランジットよりは安い(はず?)

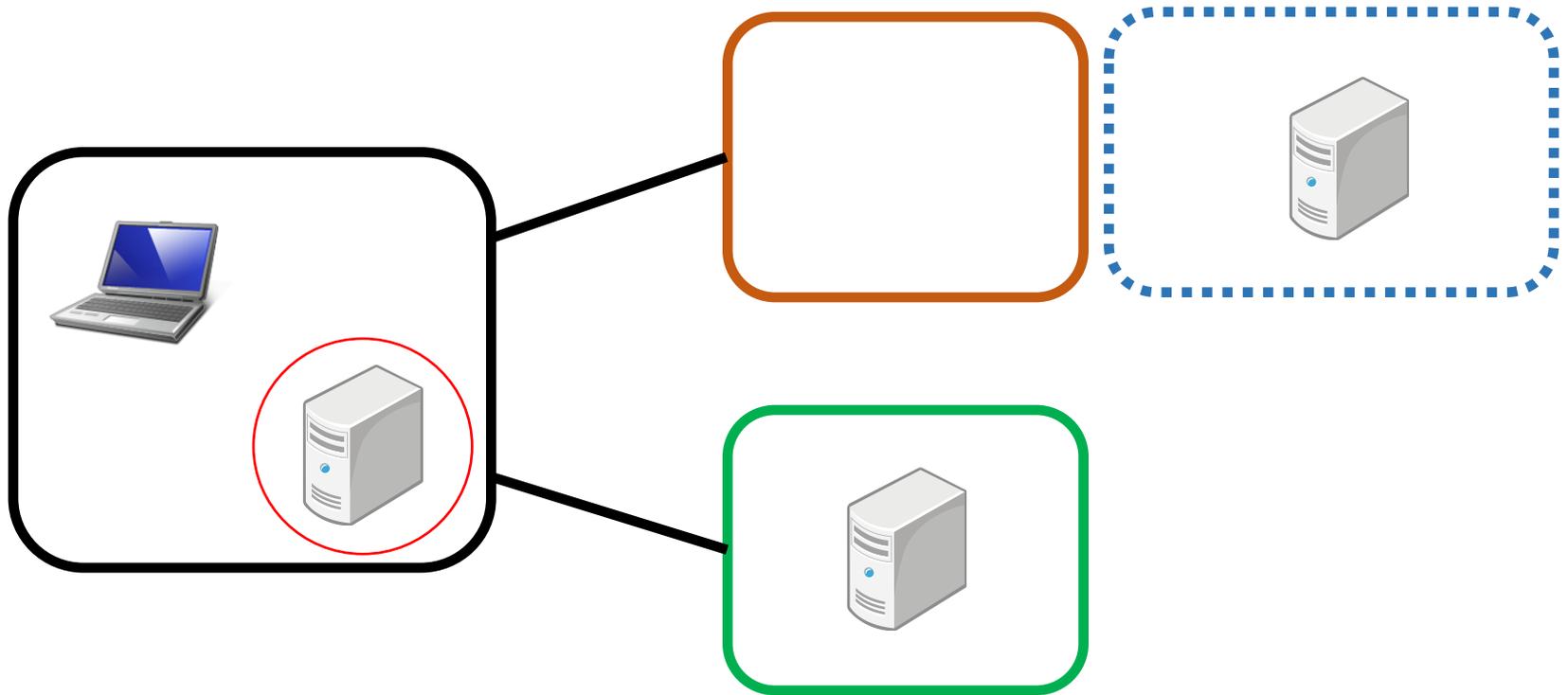
# Option 1: 上流から受信

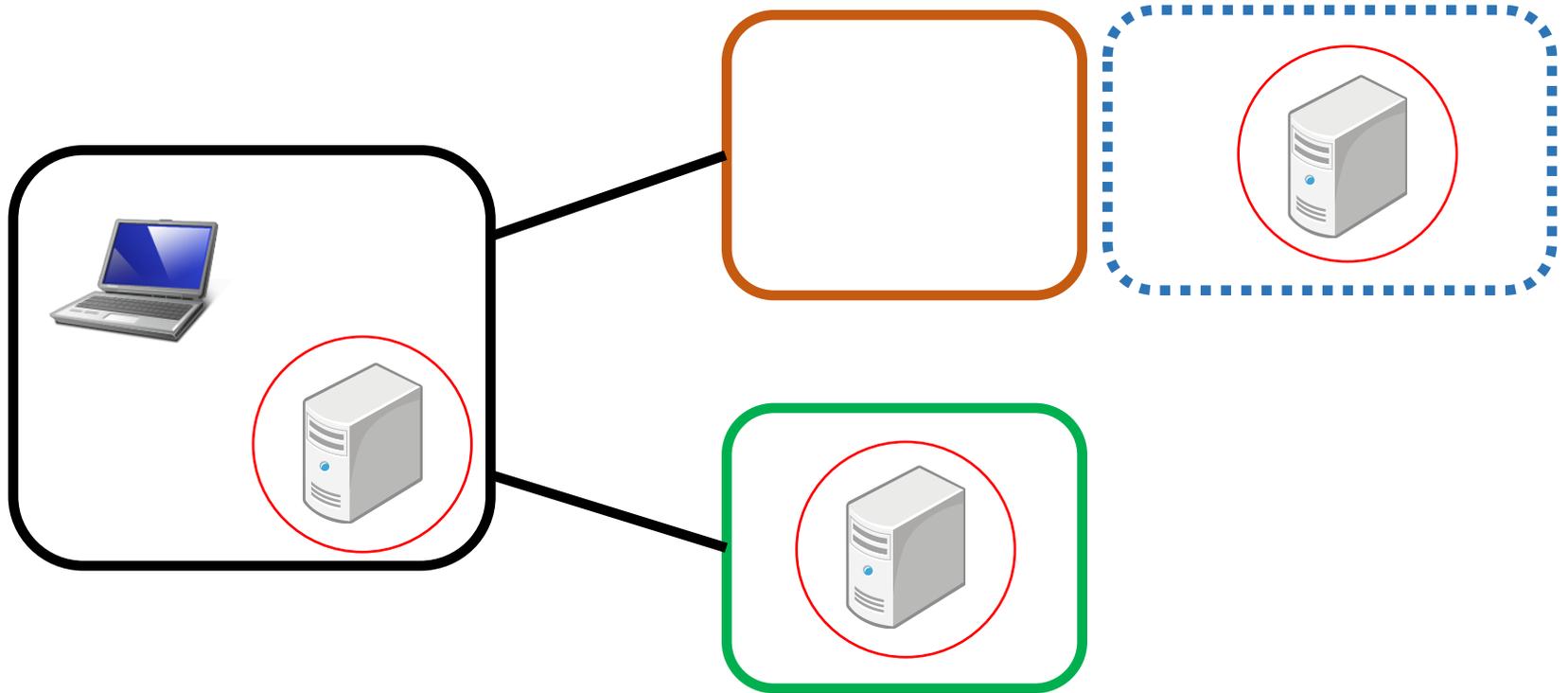


# Option 2: 直接ピア



# Option 3: キャッシュ箱の導入





# CDN

- CDNはキャッシュの配備によって、大きなトラフィックコントロール能力を獲得し、増強している
  - 通信の方端を担うため、ASをまたがったトラフィックの負荷分散など細やかな制御が可能
  - BGPなどでは到底実現できないトラフィックの制御
- 事業者のトラフィック制御能力が不均衡になる事態
  - ISPから見ると、制御をお任せしちゃっていて、手出しできない状態

# ピアとCDNキャッシュ

- とっても違う背景、アイデア
- ‘ピア’は対等な関係を前提にしてる
  - お互いにあまり恣意的なトラフィック制御できない前提
  - 暗黙に静的、あるいは緩やかな変動を期待している
- ‘キャッシュ’はCDNの機能増強
  - ISPがCDNのトラフィック効率配信に協力し、対価としてトランジットコストの削減が実現される

# ご意見聞いてみたい

- BGPで頑張りたい人いますか？
  - もっと早く収束させたいとか
  - 何ができるかとか、何を調べないといけないかとか
  - あと、なぜ頑張らなきゃいけないかとか
- ピアとCDNとキャッシュの整理
  - ピアとの付き合い方
  - CDNとの付き合い方
  - キャッシュの付き合い方

# まとめ

- BGPはもうちょっと頑張れる
  - 網内は比較的簡単
  - ASを超えて頑張ろうとすると協調が必要
  - チューニングや手法などはちょっと話していきたい
- BGPだけ見ているとトラフィック制御はできない
  - ピア、CDN、キャッシュ
- CDNは大きなトラフィックを送出
  - 交渉可能な相手がいることは良いこと ☺