バックボーンアドレス分離とセキュリティの考察

さくらインターネット(株)
技術部 ネットワークチーム
大久保 修一 ohkubo@sakura.ad.jp

本日のAgenda

- ルータ宛てDoSアタックの危険性
- ルータへ設定するIPアドレスの見直し
- 到達性のない空間へのリナンバ
- 到達性のないIPアドレスの注意点、検討事項
- まとめ、議論

ルータ宛てDoSアタックの危険性

- ルータに設定しているIPアドレスは、Tracerouteすると外部からわかる。
- ルータはパケットを転送するのは速いが、自身へ向かってくるトラフィック(DoSアタックなど)には弱い。
- ルータ宛てDoSアタックの例
  - 大量のパケット(UDPフラッドなど)
  - telnet,ssh,BGPへのsynアタック

直接的なアタック

  - TTL0になるパケット
  - 未解決ARPな宛先へのパケット
  - 宛先がランダムなパケットの中継

間接的にルータに負荷を与える。ルータのアーキテクチャに依存。

ルータ宛てDoSアタックの危険性

- DoSアタックを受けた時の症状
  - CPU負荷上昇
  - telnetできない
  - BGPピアダウン
  - VRRPの状態がフラップ(Master←→Backup)
  - LACPが切断される
  - OSPFのneighborダウン、LSAの不伝播

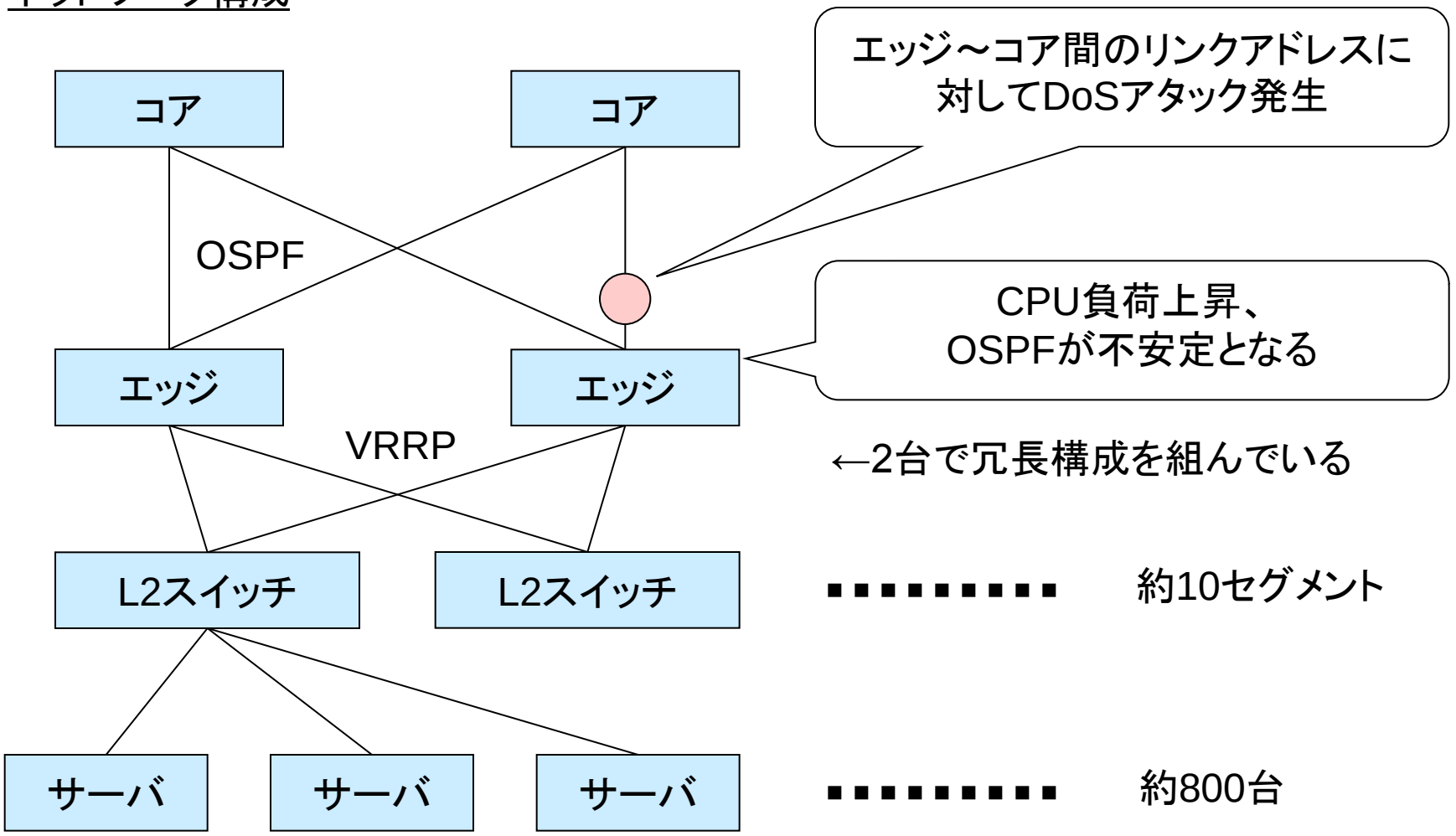
    →実際に発生しました（涙

ルータ宛てDoSアタックの事例

- 障害情報
  http://www.sakura.ad.jp/news/archives/20080904-001.news
- 障害発生日時
  2008年09月04日 00時36分 - 01時46分
- 影響範囲
  「さくらウェブ」と「さくらの専用サーバ」
- 障害内容
  通信が不安定となる

実際になにが起きていたか？

ネットワーク構成



実際になにが起きていたか？

sFlowによるサンプルを、tcpdumpで解析した結果

01:03:05.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:05.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:05.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:05.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:05.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:05.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:05.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:06.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:06.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:06.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1
01:03:06.000000 IP xx.xx.xx.xx.4620 > xx.xx.xx.xx.80: UDP, length 1

宛先がルータのリンクアドレス

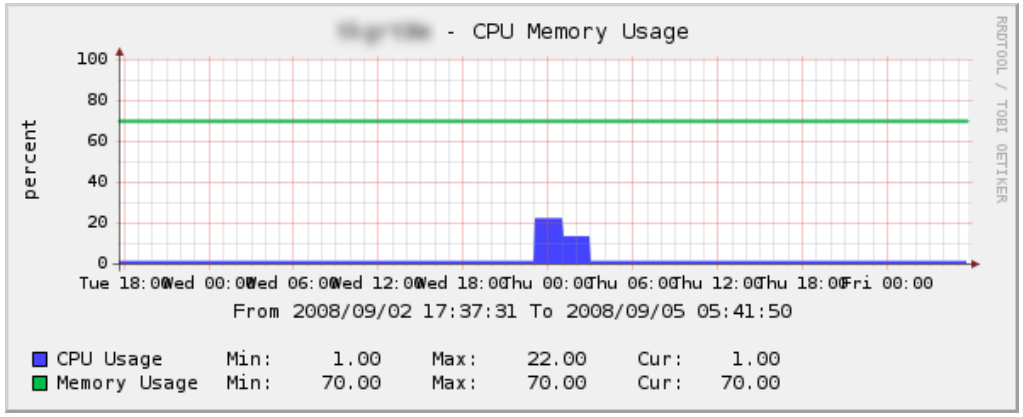
実際になにが起きていたか？

ルータのログ

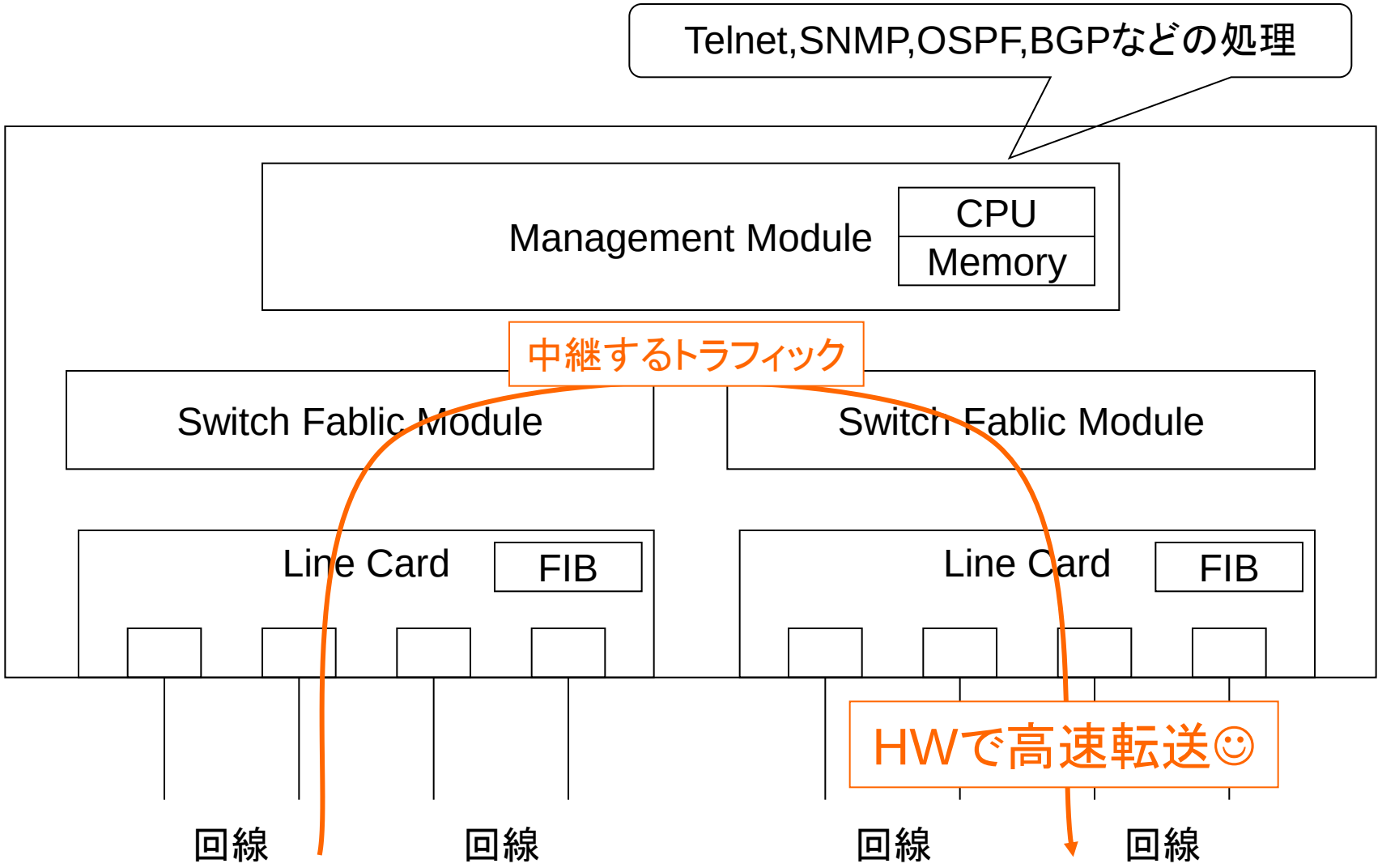
Sep 4 01:33:44 xxxxxxx-grt1b.bb.sakura.ad.jp OSPF: intf retransmit, rid xx.xx.xx.xx, intf addr xx.xx.xx.xx, nbr rid xx.xx.xx.xx, pkt type is link state request, LSA type 1, LSA id xx.xx.xx.xx, LSA rid xx.xx.xx.xx
Sep 4 01:33:49 xxxxxxx-grt1b.bb.sakura.ad.jp OSPF: intf retransmit, rid xx.xx.xx.xx, intf addr xx.xx.xx.xx, nbr rid xx.xx.xx.xx, pkt type is link state request, LSA type 1, LSA id xx.xx.xx.xx, LSA rid xx.xx.xx.xx
・・・省略・・・

OSPF LSAの再送が頻発し、
LSAが正常に伝わらない状態。

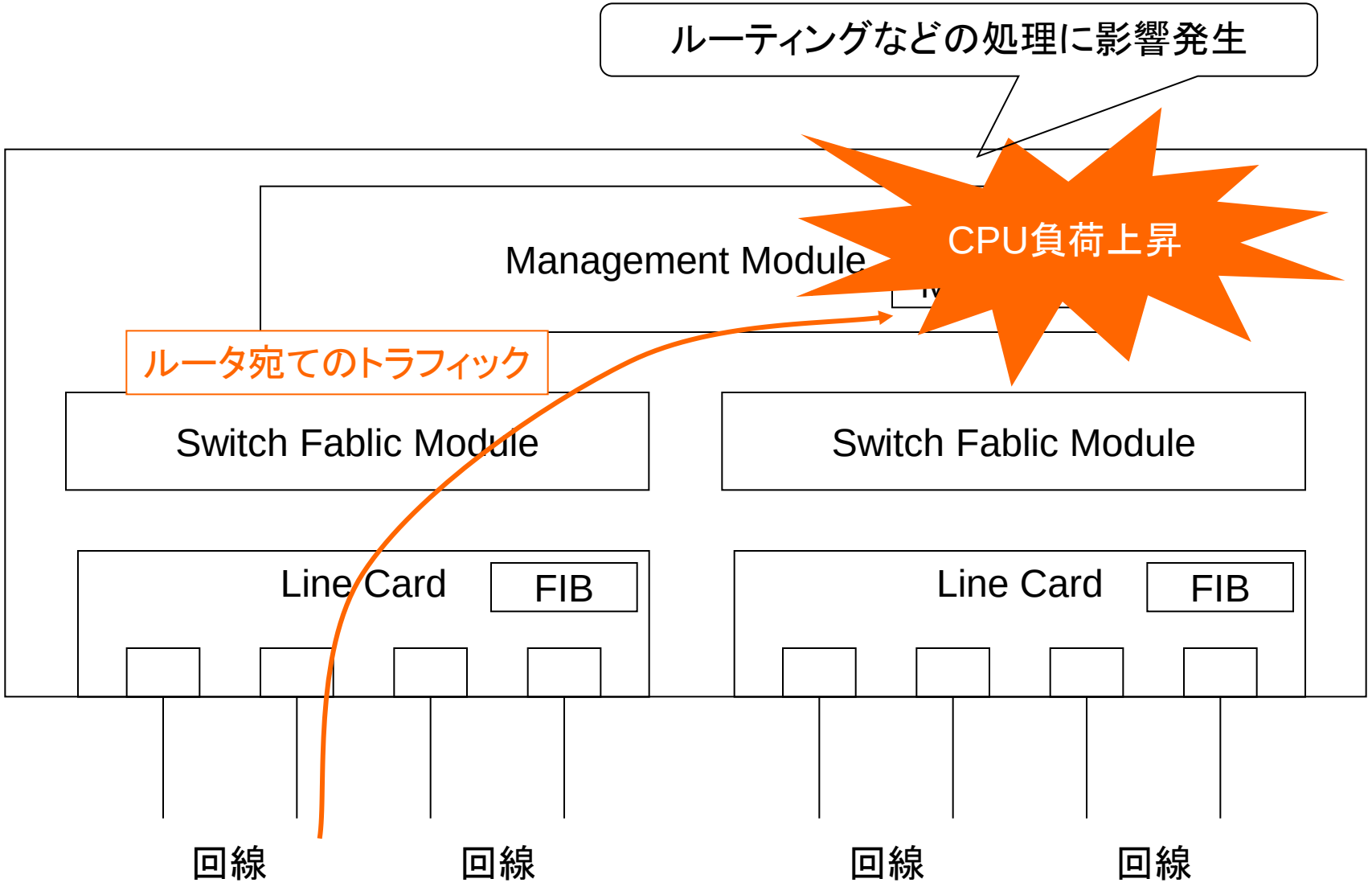
CPU負荷



一般的なルータのアーキテクチャ



一般的なルータのアーキテクチャ



ルータ宛てのDoSアタックの対策は?

- DoSアタックに強いルータを買う
  - 最近のルータは、自身宛てDoSアタックに強い製品もあります。
  - ただ、DoSアタックに弱い古いルータが残ってると。。。
- ACLを書いて頑張る
  - 自身宛のパケットにフィルタ設定
  - 機種によっては、自身宛のACLが簡単にかけない
  - ACLがCPU処理な機種

    ラインカードでパケット落とせないので、結局CPU負荷上昇が発生。
- ルータに外部ASから到達性のないIPアドレスを使う
  - そもそもルータが狙われることがなくなる
  - 今回はこれを検討

到達性のないIPアドレスを使っても大丈夫？

- そもそも、ルータにIPアドレスを設定する目的は？
  - BGPのピアを張るため
  - OSPFのLSIDを決めるため
  - next-hop解決のため
  - マネジメント(telnet,snmp,ntpなど)するため
  - etc…
- 外部ASからの(外部ASへの)到達性は必要なさそう。

到達性のないIPアドレス

- プライベート(RFC1918)アドレス
  - 逆引きできない。
    - tracerouteした時に、逆引き設定してると結構便利だったりする。
  - お客さんがtracerouteした時に見え方が良くない。
    - なぜプライベートアドレスが見えるのか、問い合わせがきそう。
    - 特にBGPのトランジットサービスをしていると。
- 外部に経路を広告しないグローバルアドレス
  - 空いている空間があれば、すぐにでも実施可能。
  - 上記プライベートアドレスを利用した問題は解決。
- 特殊用途用プロバイダ非依存アドレスを取得？
  - 要件(IX,DNS,小規模マルチホーム)にはマッチしなさそう。

到達性のないIPアドレスが使われている例(1)

- JPNAP東京II(218.100.45.0/24)
  - tracerouteの結果

    % traceroute -n xx.xx.xx.xx

    traceroute to xx.xx.xx.xx (xx.xx.xx.xx), 64 hops max, 60 byte packets

    1 xx.xx.xx.xx 0.789 ms 0.447 ms 0.48

    2 xx.xx.xx.xx 0.477 ms 0.482 ms 0.486 ms

    3 218.100.45.xx 0.986 ms 0.479 ms 0.483 ms

    4 xx.xx.xx.xx 0.981 s

    5 xx.xx.xx.xx 0.982 ms 0.982 ms 0.986 ms

    ・・省略・・

    一部会場のみ

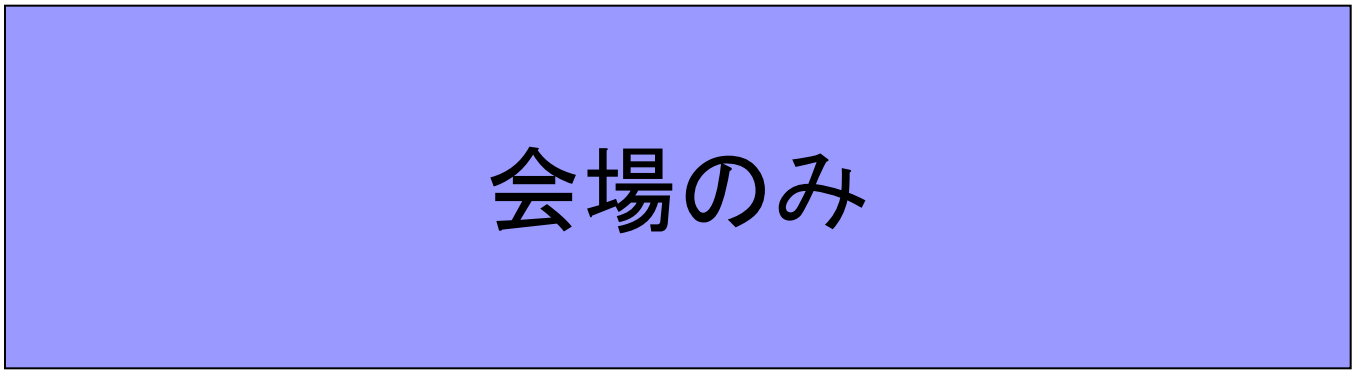
    JPNAP東京IIのIPアドレス
  - BGPテーブルの検索結果

    #show ip bgp 218.100.45.0/24

    BGP4 : None of the BGP4 routes match the display condition
  - インターネットに広報されていない。

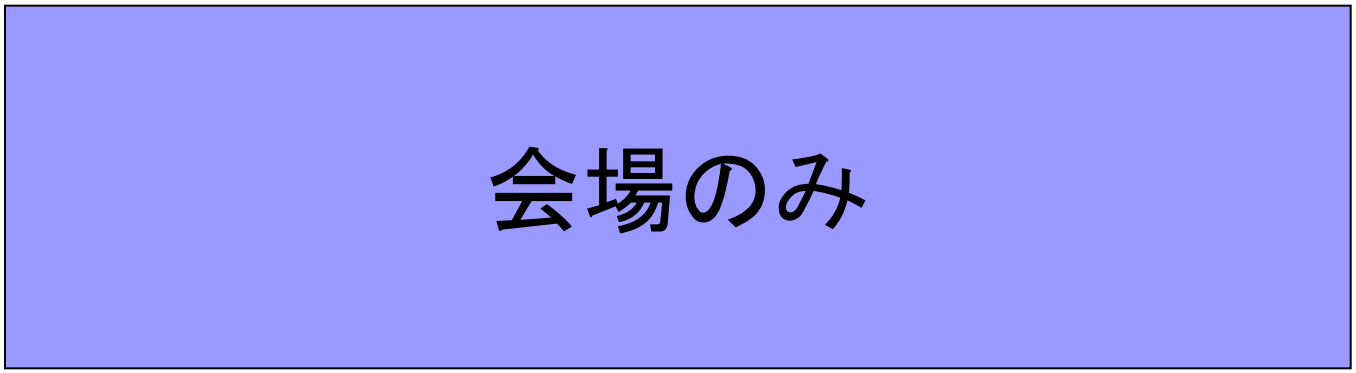
到達性のないIPアドレスが使われている例(2)

- AS17676さんのバックボーン



到達性のないIPアドレスが使われている例(3)

- AS2510さんのバックボーン



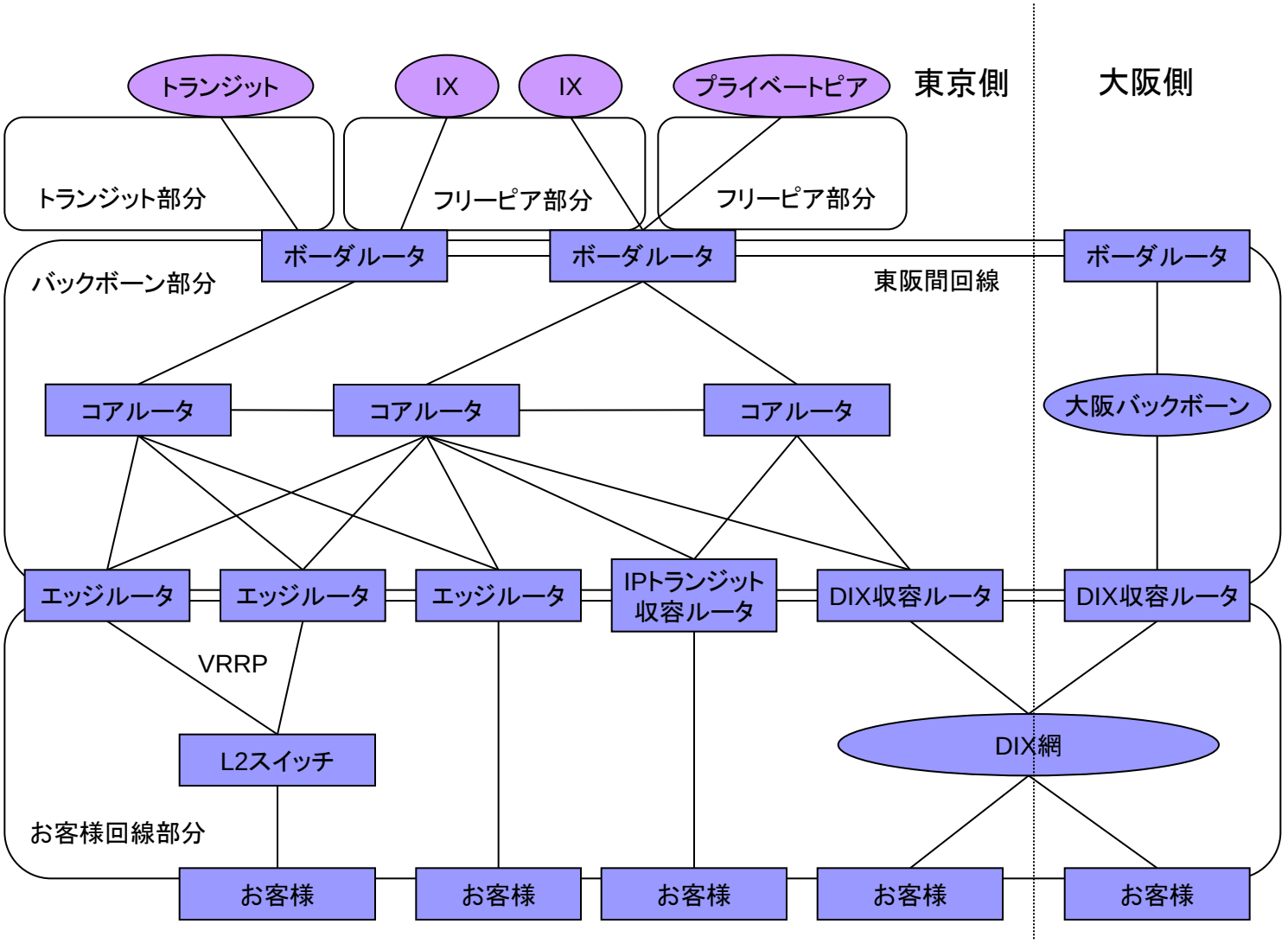
ということで、、、

うちでもやってみました。

バックボーン不到達化計画

- 不到達IPアドレス設定ポリシー決め
  - バックボーンのどの部分を不到達空間にするか？
- 不到達IPアドレスの確保
  - 必要な不到達空間のIPアドレスのサイズをカウント
  - どのCIDRブロックのどのあたりから確保するか？
- 不到達IPアドレス空間を作る
  - 不到達空間の経路情報を、外部に広報しないようにする。
- 新規に設置するルータは、不到達IPアドレスを設定
- 既存のルータはリナンバする
  - リナンバメンテナンスのスケジューリングとアナウンス
  - リナンバ作業の実施

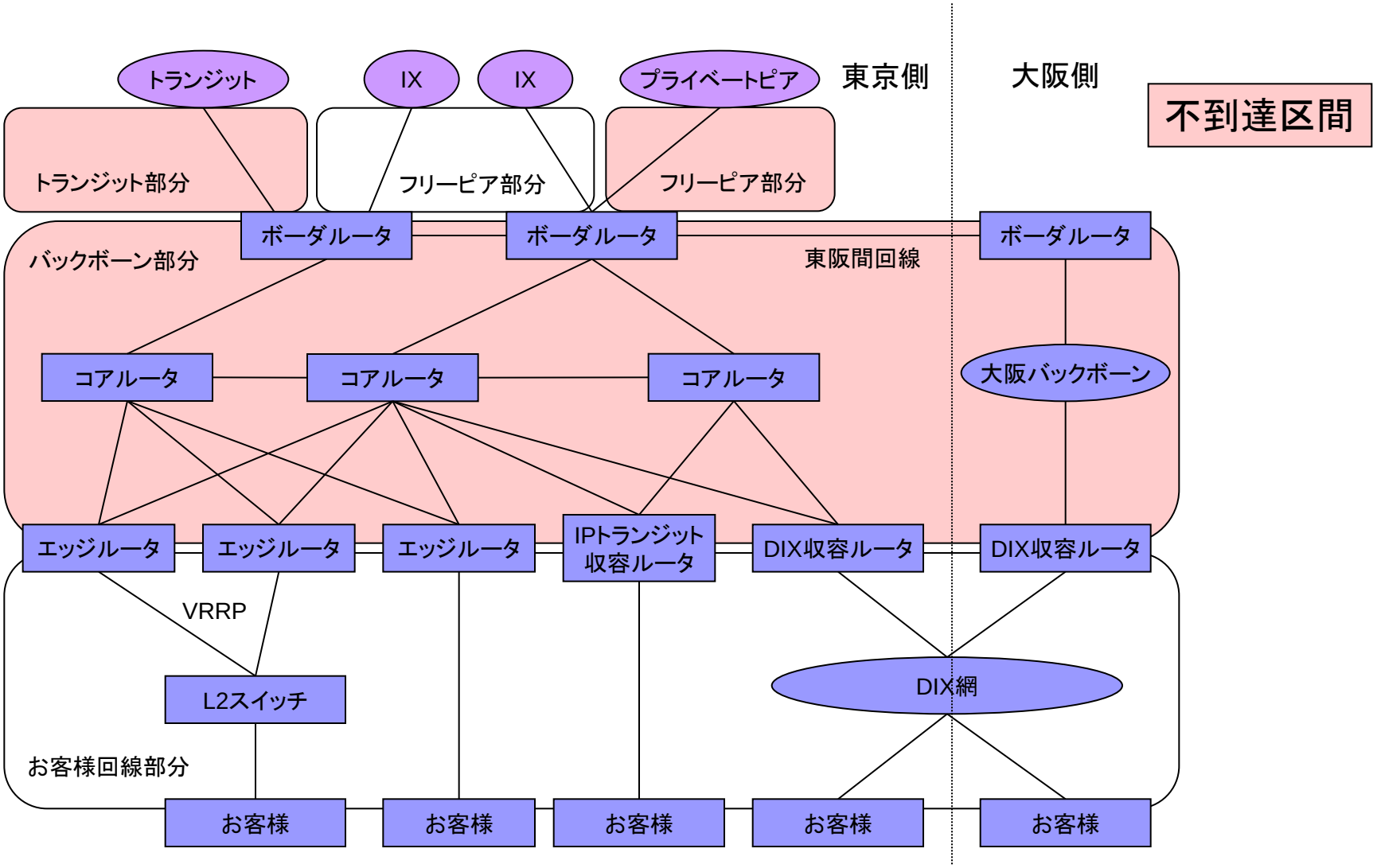
バックボーンの概要


トランジット
IX
IX
プライベートピア
東京側
大阪側
トランジット部分
フリーピア部分
フリーピア部分
ボーダルータ
ボーダルータ
ボーダルータ
バックボーン部分
東阪間回線
コアルータ
コアルータ
コアルータ
大阪バックボーン
エッジルータ
エッジルータ
エッジルータ
IPトランジット
収容ルータ
DIX収容ルータ
DIX収容ルータ
VRRP
L2スイッチ
DIX網
お客様回線部分
お客様
お客様
お客様
お客様
お客様

不到達IPアドレス設定ポリシー

- 以下のIPアドレスは、不到達のものを設定
  - 対象機器
    - コアルータ、ボーダルータ、エッジルータ
  - 対象のインターフェイス
    - ループバックアドレス (=router-id)
    - ルータ間のリンクアドレス(/30)
    - トランジットやプライベートピアのリンクアドレス(/30)
- 以下のIPアドレスは対象外
  - お客様向けインターフェイスのリンクアドレス
    - 基本的にはルータ間リンクでも、到達性を確保しておいた方が無難。
    - 希望するお客様のみ対応する。
  - IX向けインターフェイスのリンクアドレス
    - IX事業者側で指定されているので変更不可

不到達IPアドレス設定ポリシー



IPアドレスの確保

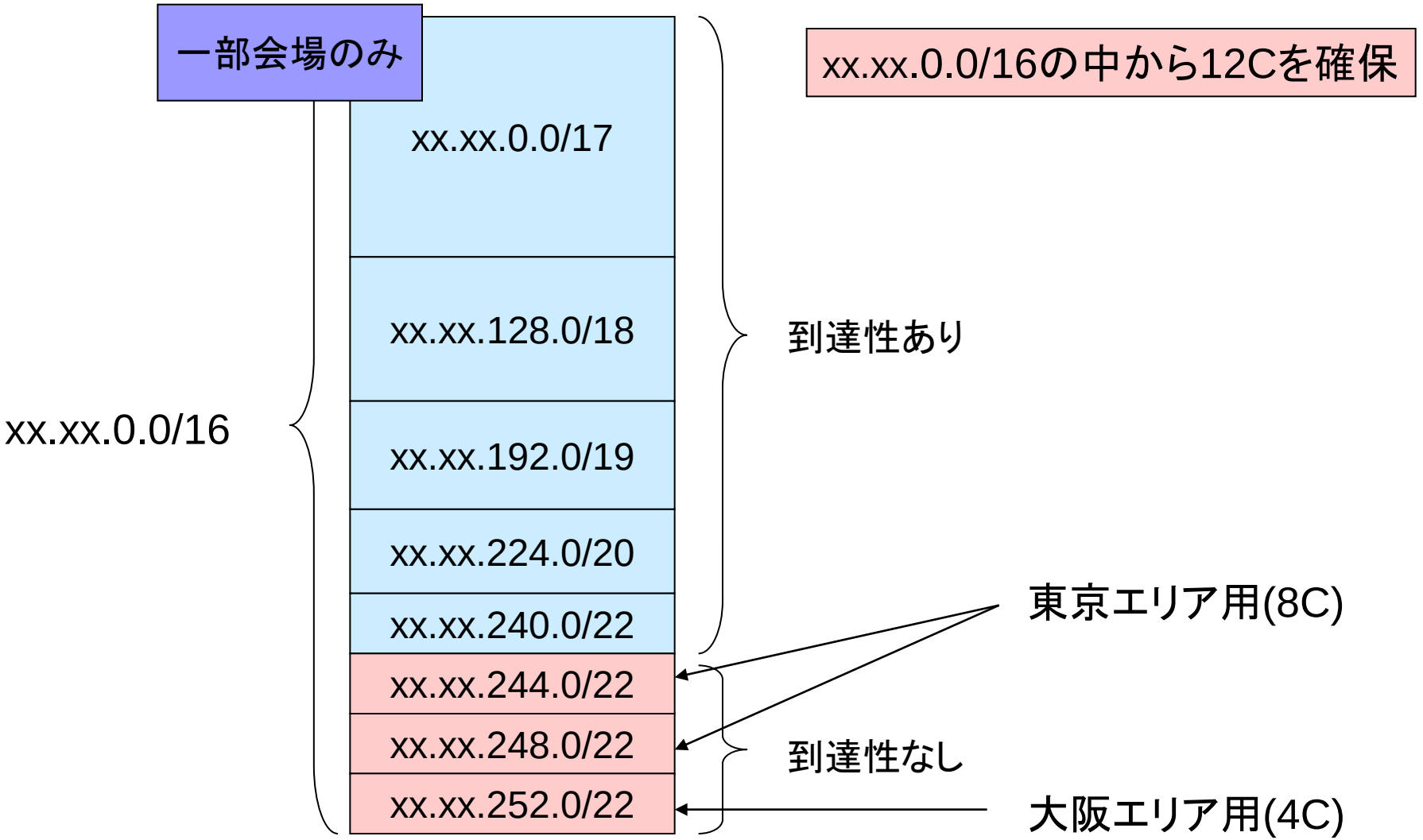
- ループバックアドレス、ルータ間リンクアドレスに必要なIPアドレスの個数を、ルータの台数、ルータ間リンクの数から計算。

	東京エリア	大阪エリア
必要なIPアドレスの個数	970個	643個
確保する空間	8C(=2048個)	4C(=1024個)

※ 東京、大阪、合計12C分を確保。

※ ルータ間リンクアドレスは/30で設定するため、4個必要。

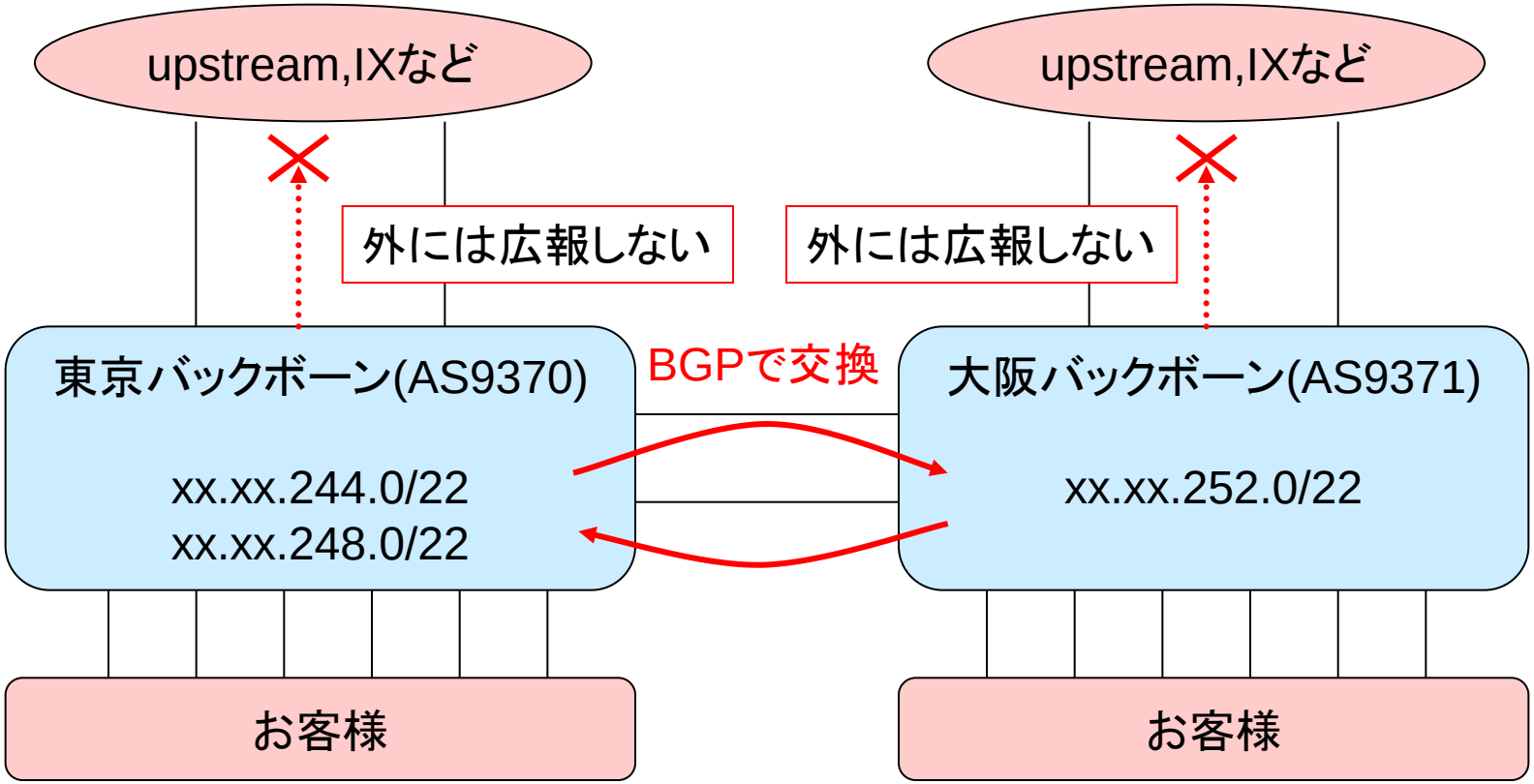
到達性のないIPアドレス空間を作る



到達性のない空間を作る

- 新規に以下経路を広報(2008/3/12実施)
  - xx.xx.0.0/17
  - xx.xx.128.0/18
  - xx.xx.192.0/19
  - xx.xx.224.0/20
  - xx.xx.240.0/22
  - ※ あらかじめIRRへの登録や、ピア先ISPへの連絡を行う(2008/2/28実施)
- その後、以下の広報を停止(2008/3/25実施)
  - xx.xx.0.0/16
- 結果、以下の到達性のない空間が生まれる
  - xx.xx.244.0/22
  - xx.xx.248.0/22
  - xx.xx.252.0/22

東京、大阪AS間での経路共有

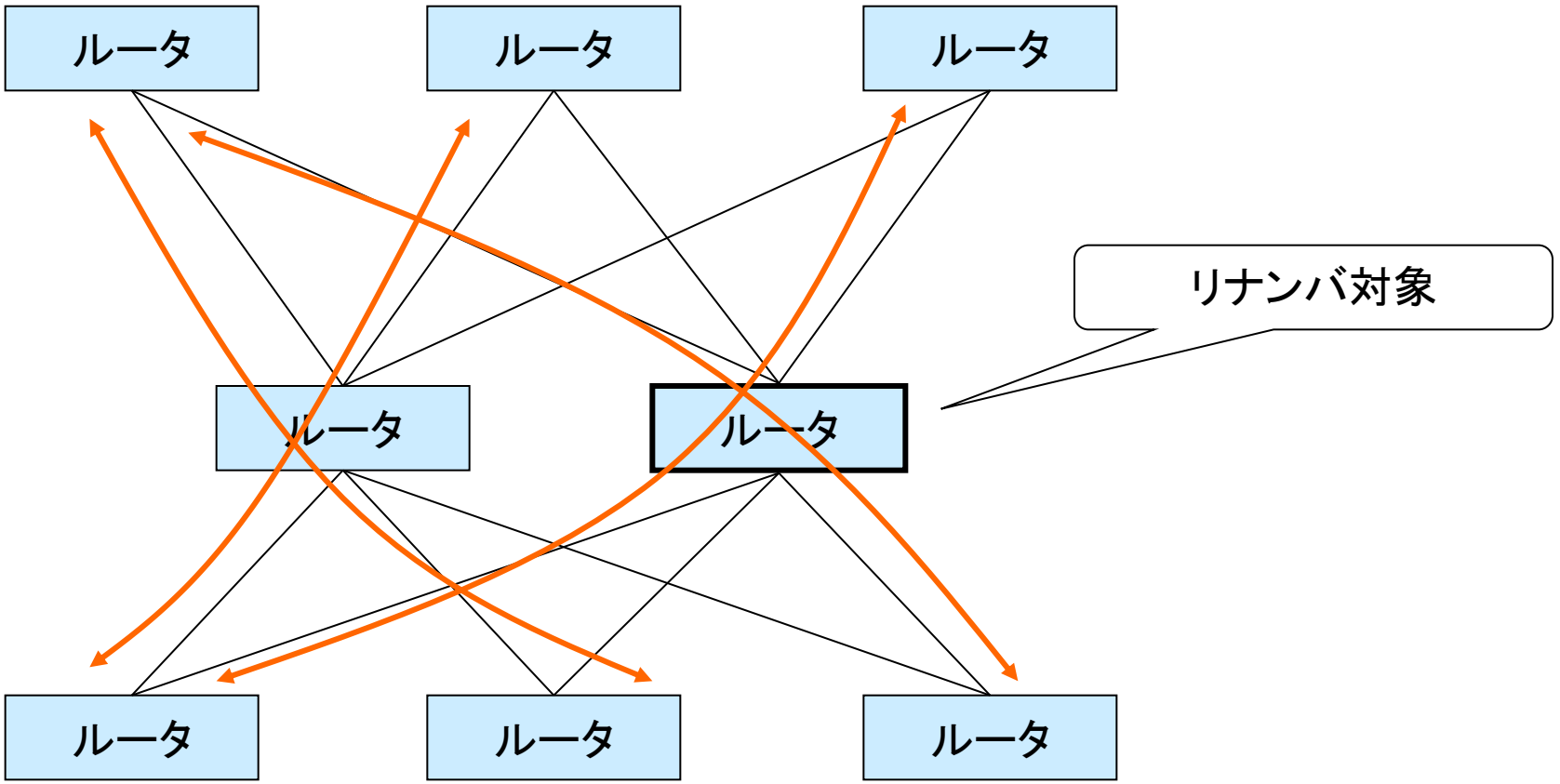


ルータマネジメント利便性向上のため、
東京、大阪のAS間では、BGPで経路の交換を行う。

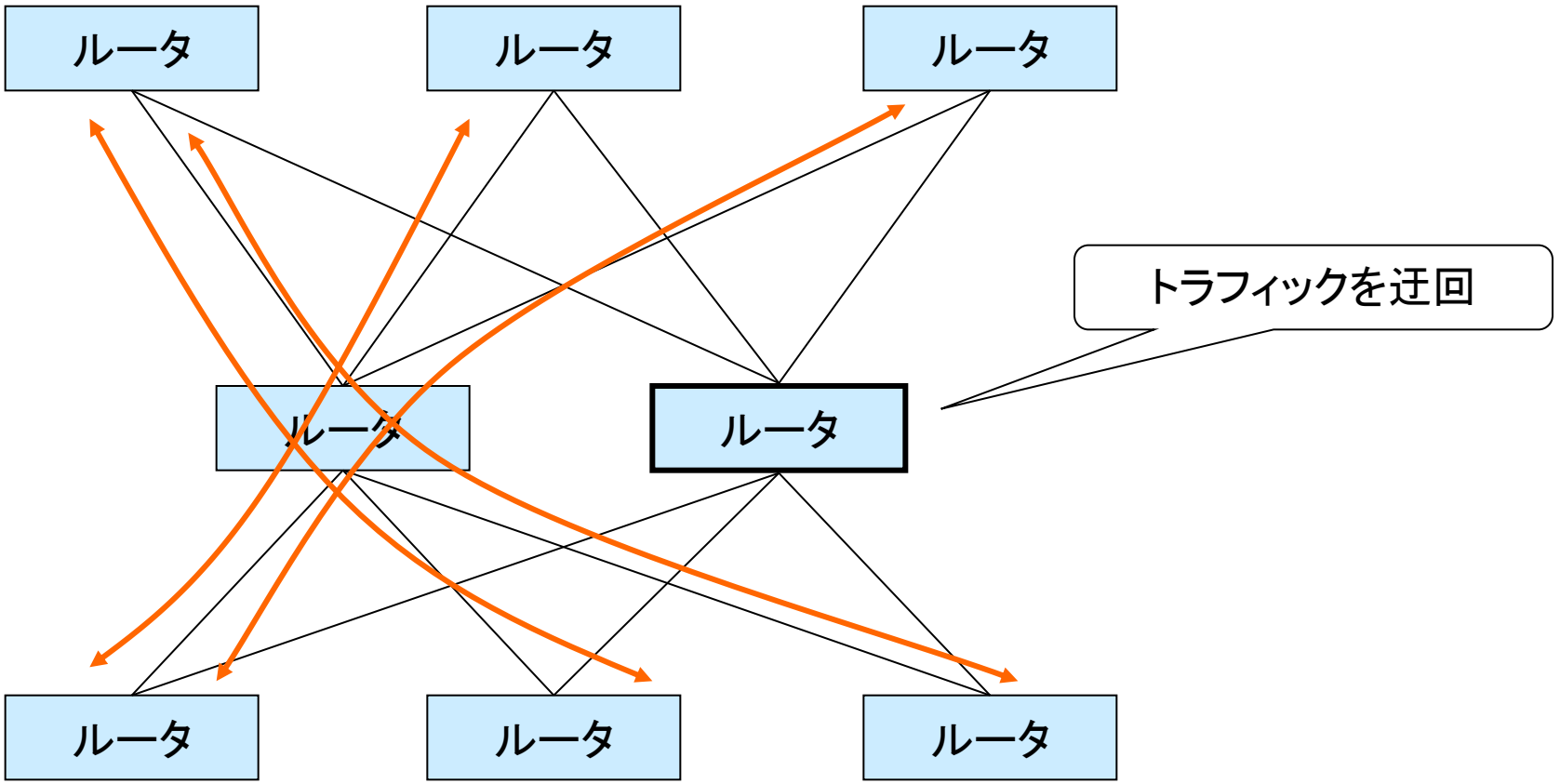
リナンバの方針

- 到達性のあるIPアドレスを設定している既存のルータは、到達性のないIPアドレスへリナンバを行う。
- 基本はお客様のトラフィックを止めずに実施。
  - router-id(=loopbackアドレス)を変更する際に、OSPF neighborがダウンするので、要注意。
- コアルータ、ボーダルータはバージョンアップなどのメンテナンスと併せて行う。
  - トラフィックを迂回させれるので、問題なし。

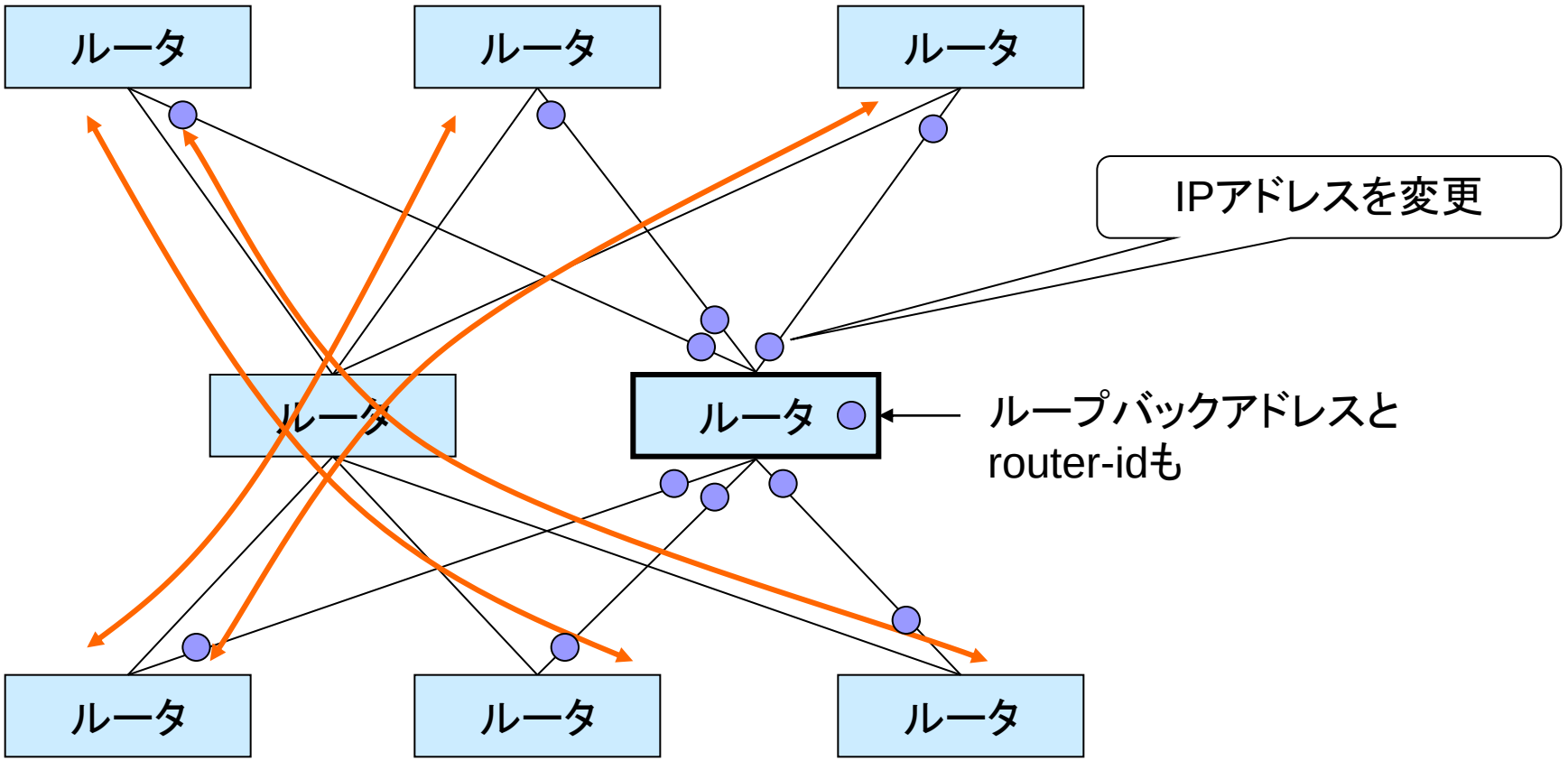
コア、ボーダルータのリナンバ



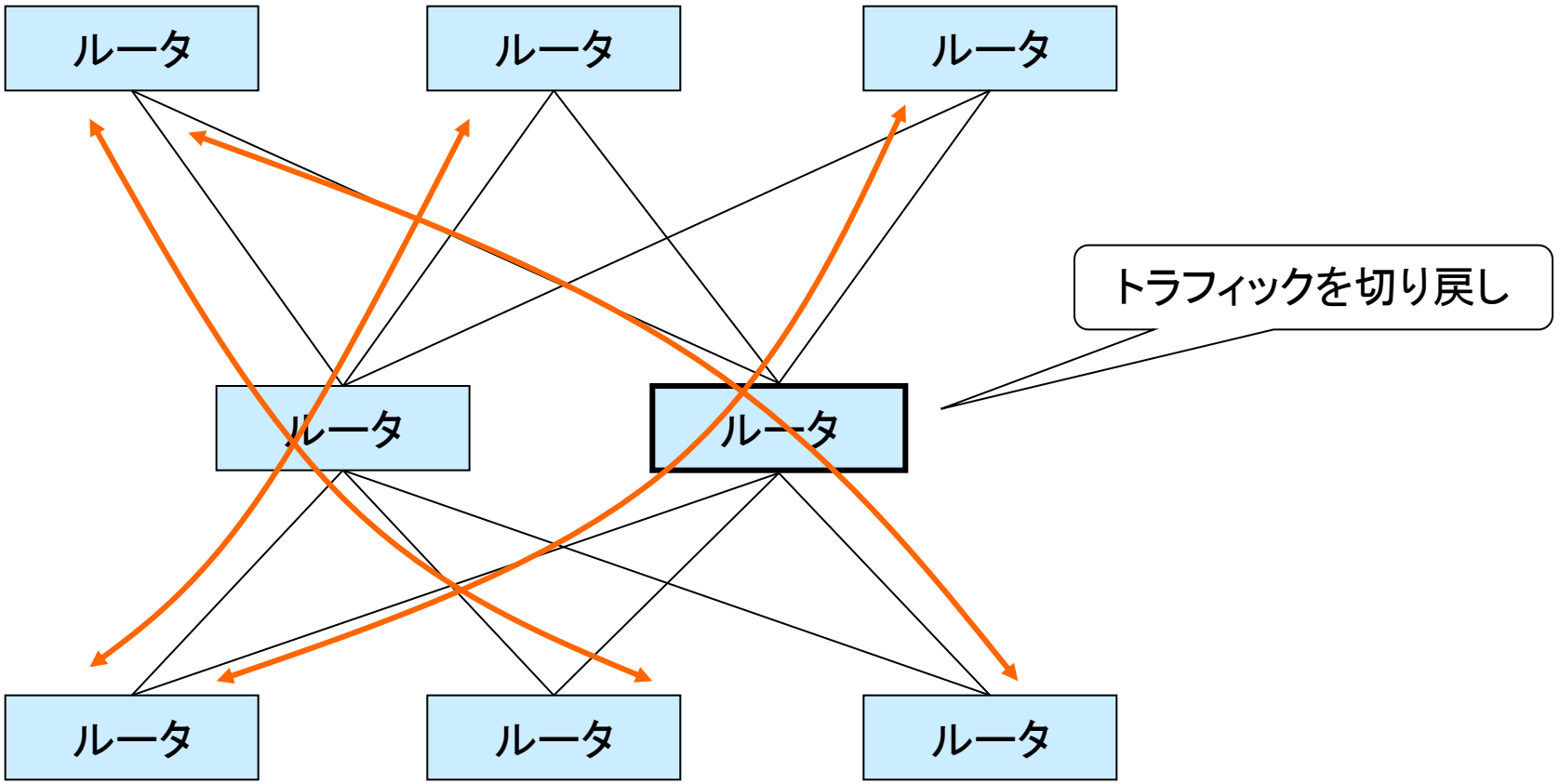
コア、ボーダルータのリナンバ



コア、ボーダルータのリナンバ



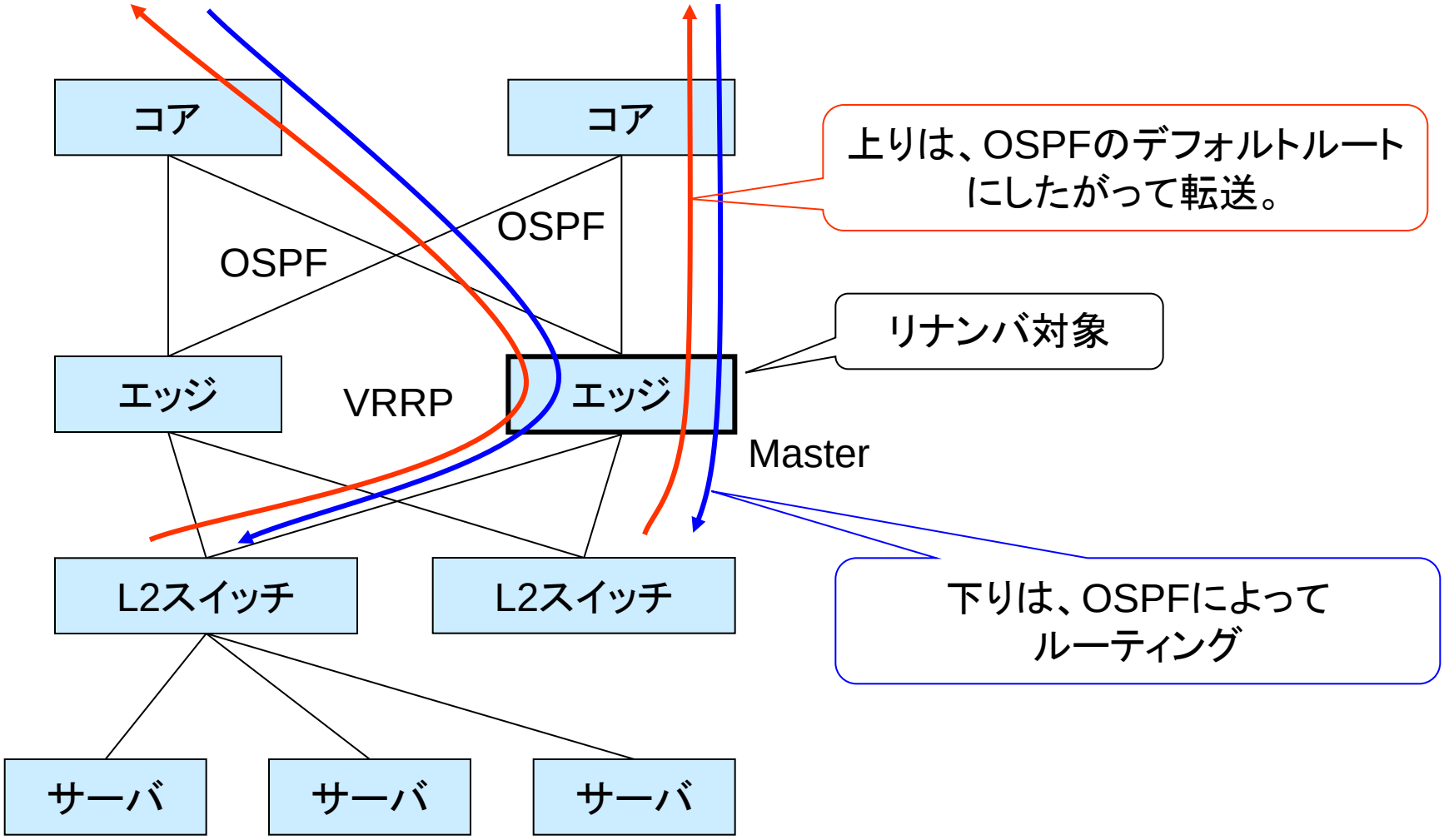
コア、ボーダルータのリナンバ



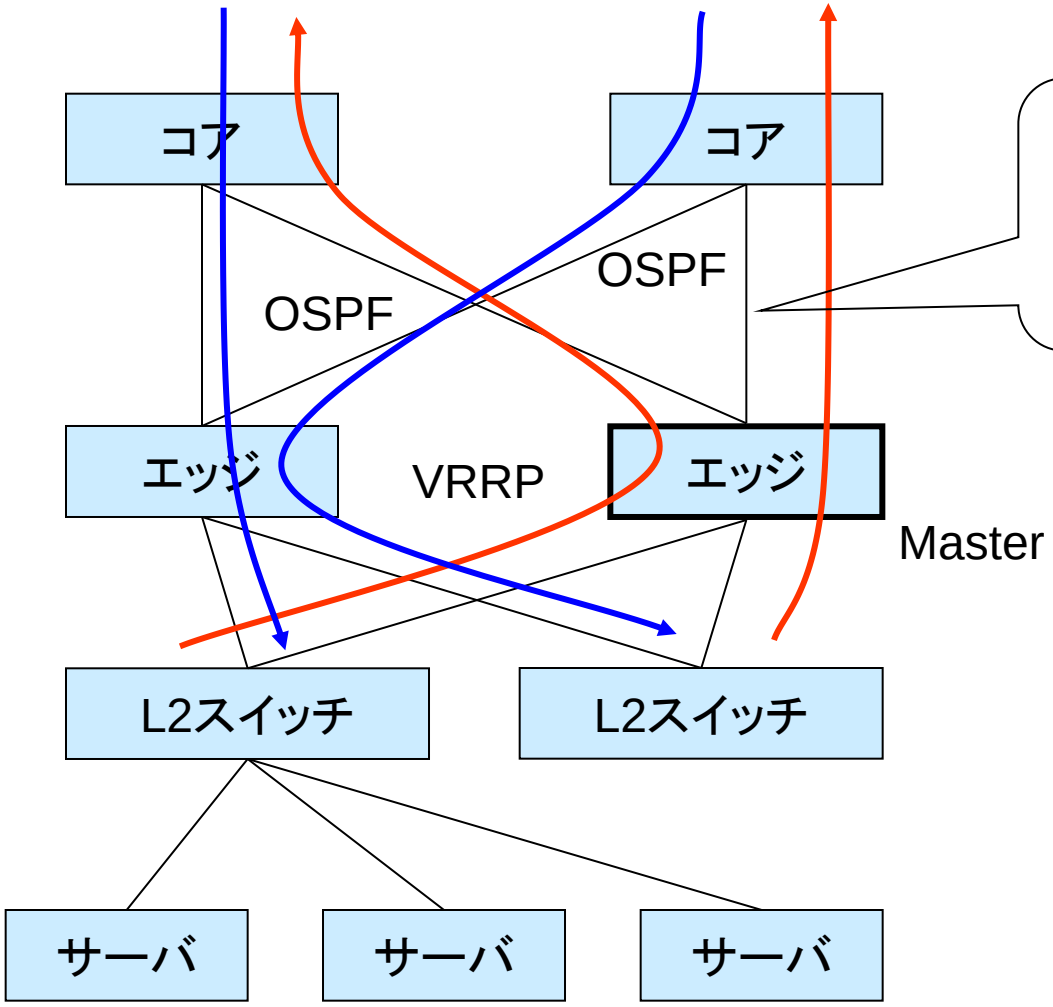
エッジルータのリナンバ

- エッジルータ
  - VRRPにて冗長化している場合は片系に迂回させてから。
  - 直収の場合は、OSPF neighborダウン時に通信が停止するので、工夫が必要。

VRRP冗長化エッジルータの場合

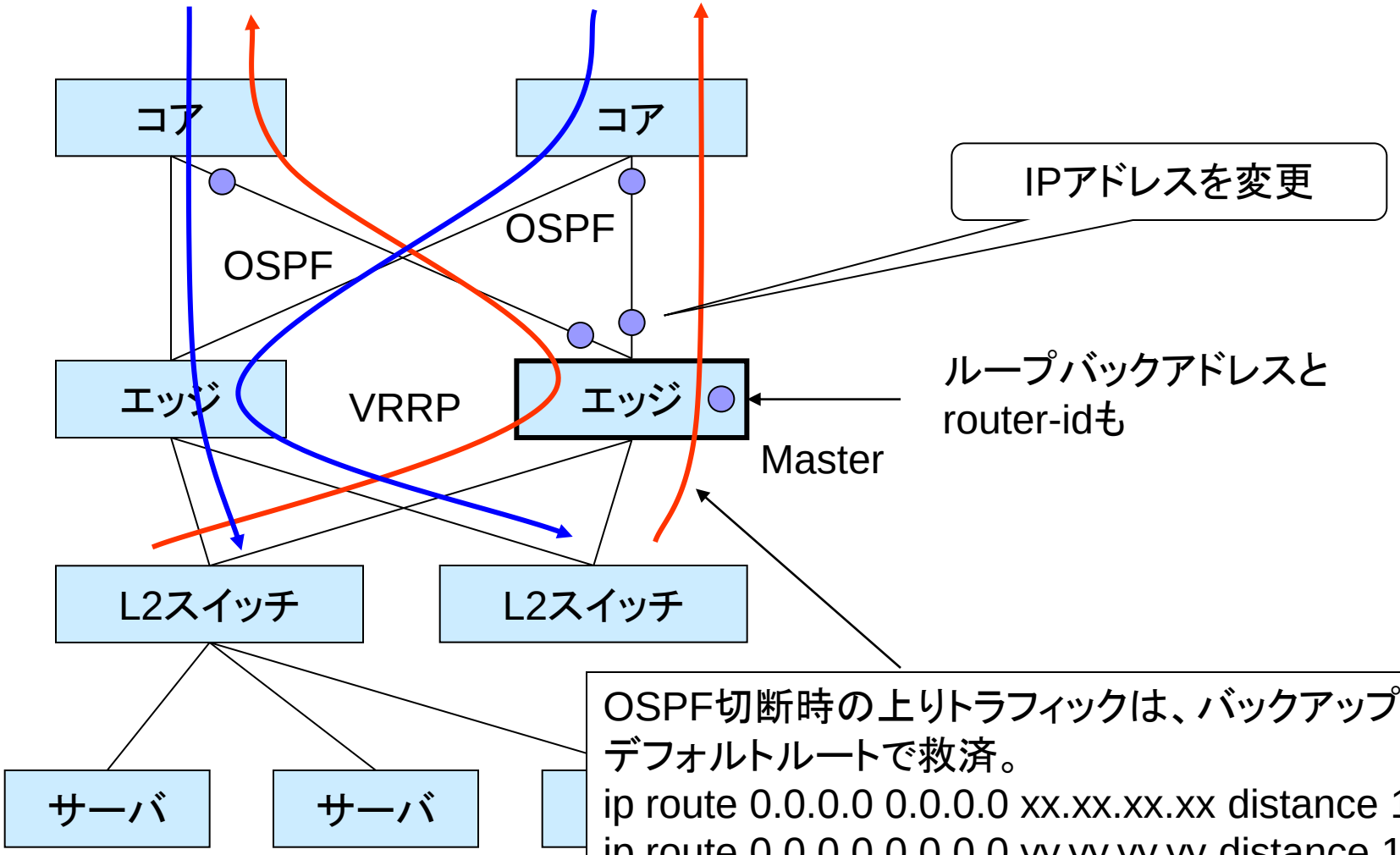


VRRP冗長化エッジルータの場合

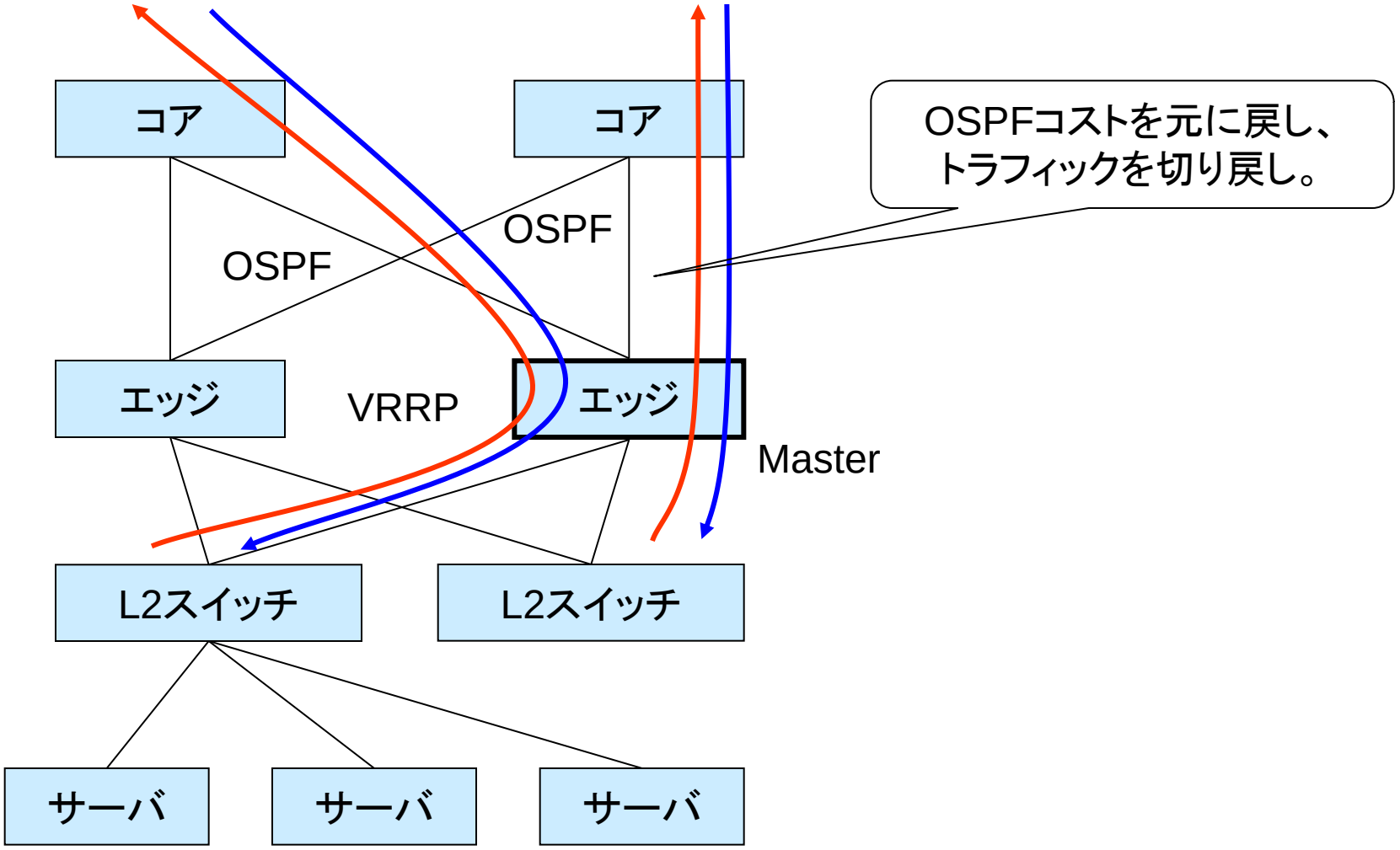


コアルータとのリンクのOSPFコストを上げ、下りトラフィックのみ対向に迂回。
VRRP Masterはそのまま。

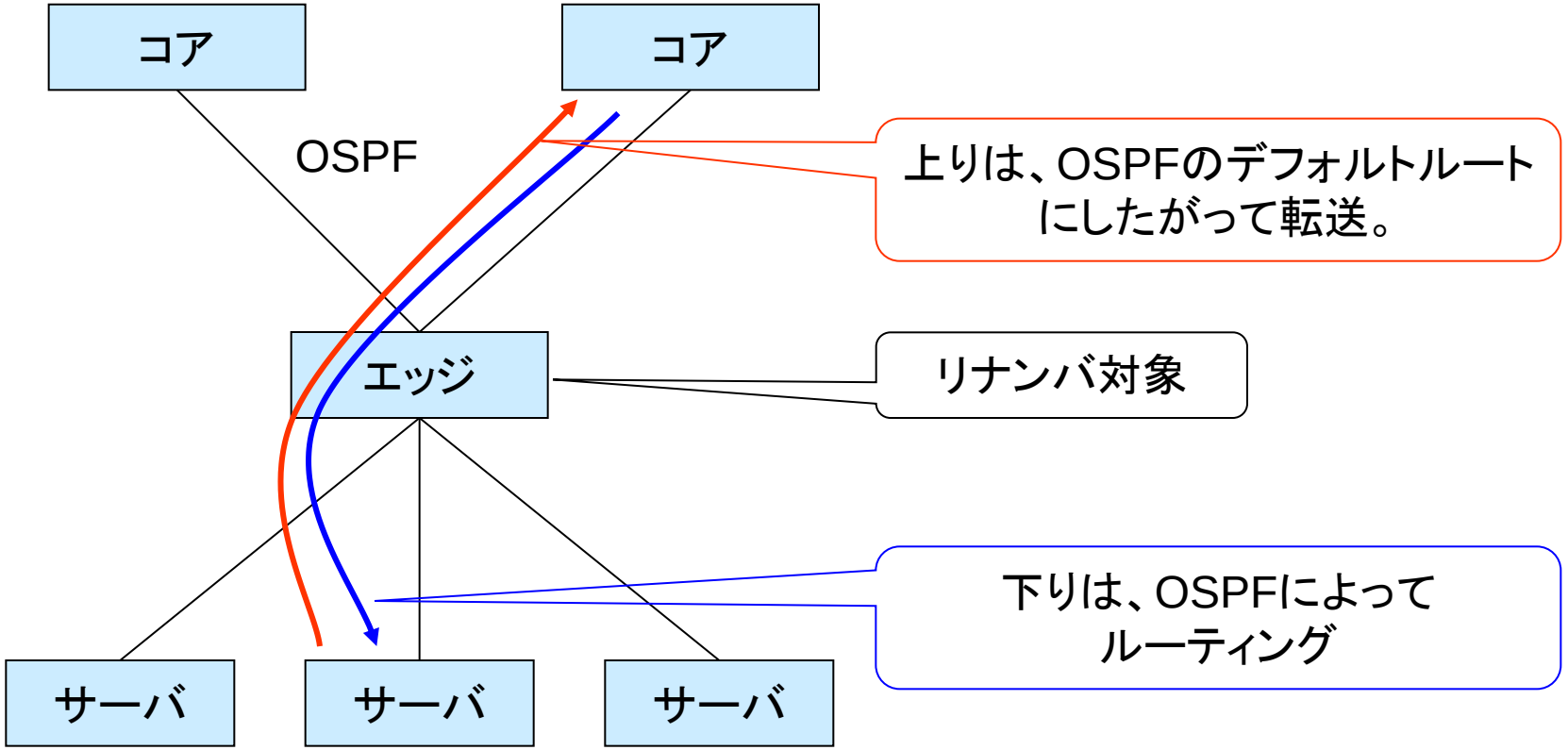
VRRP冗長化エッジルータの場合



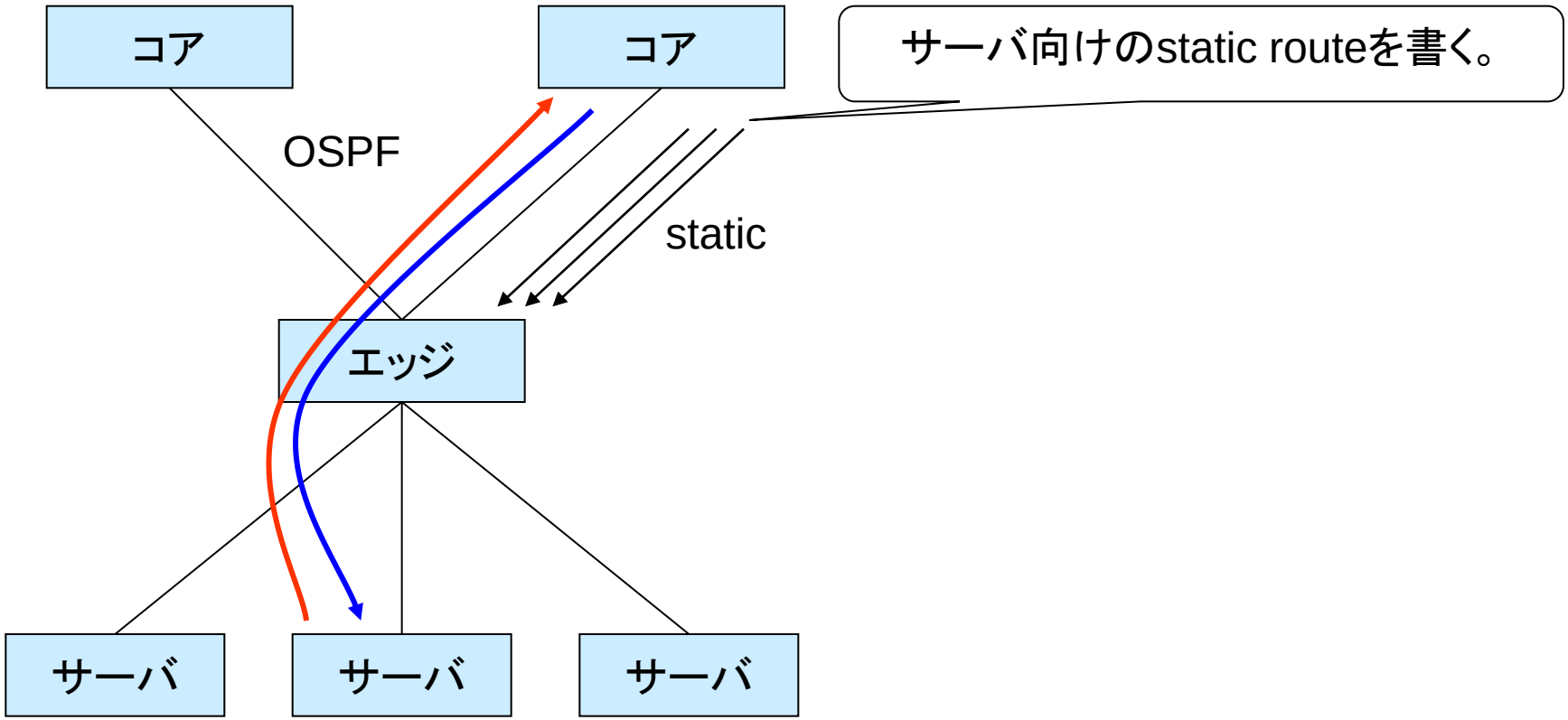
VRRP冗長化エッジルータの場合



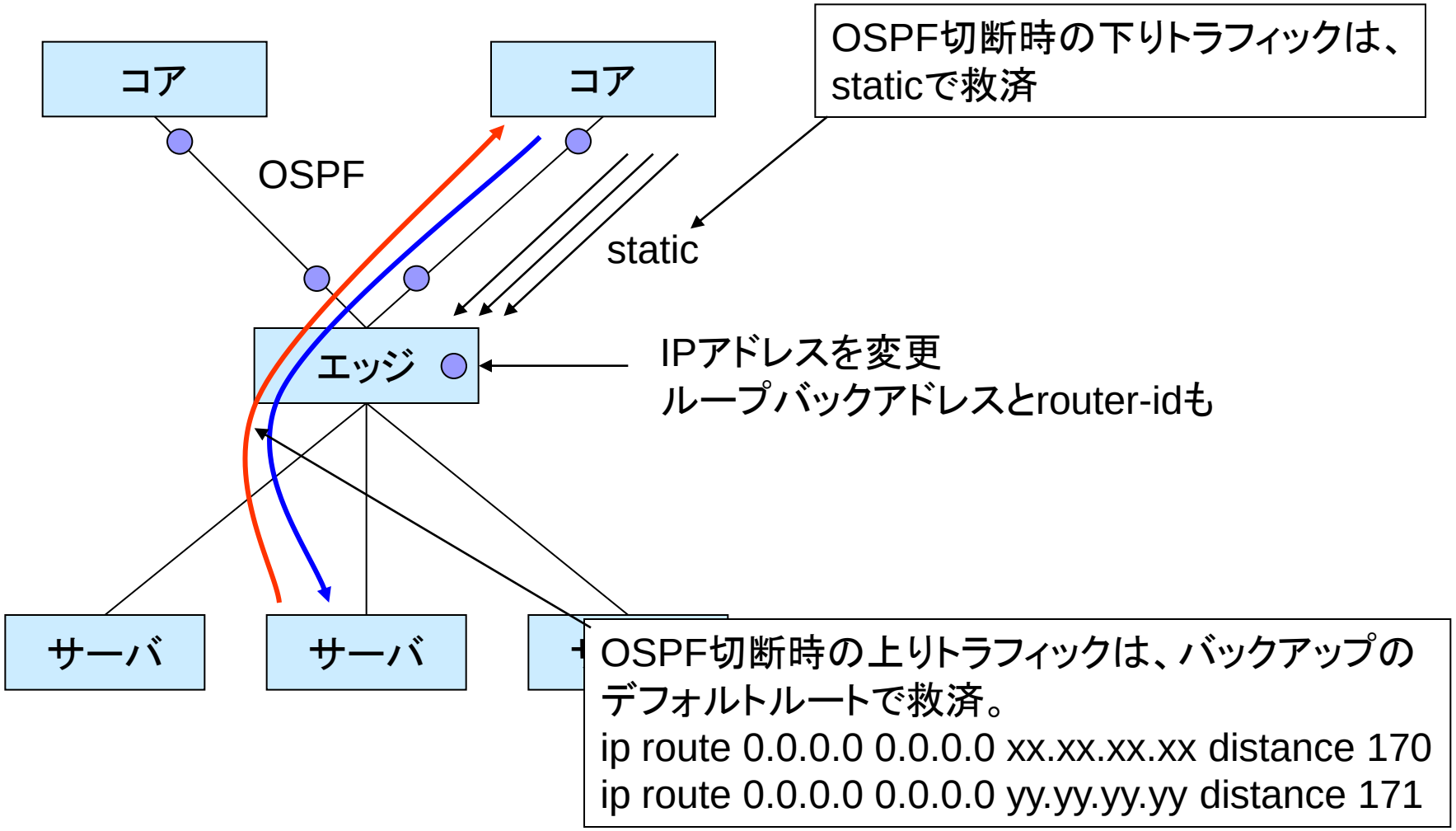
直収エッジルータの場合



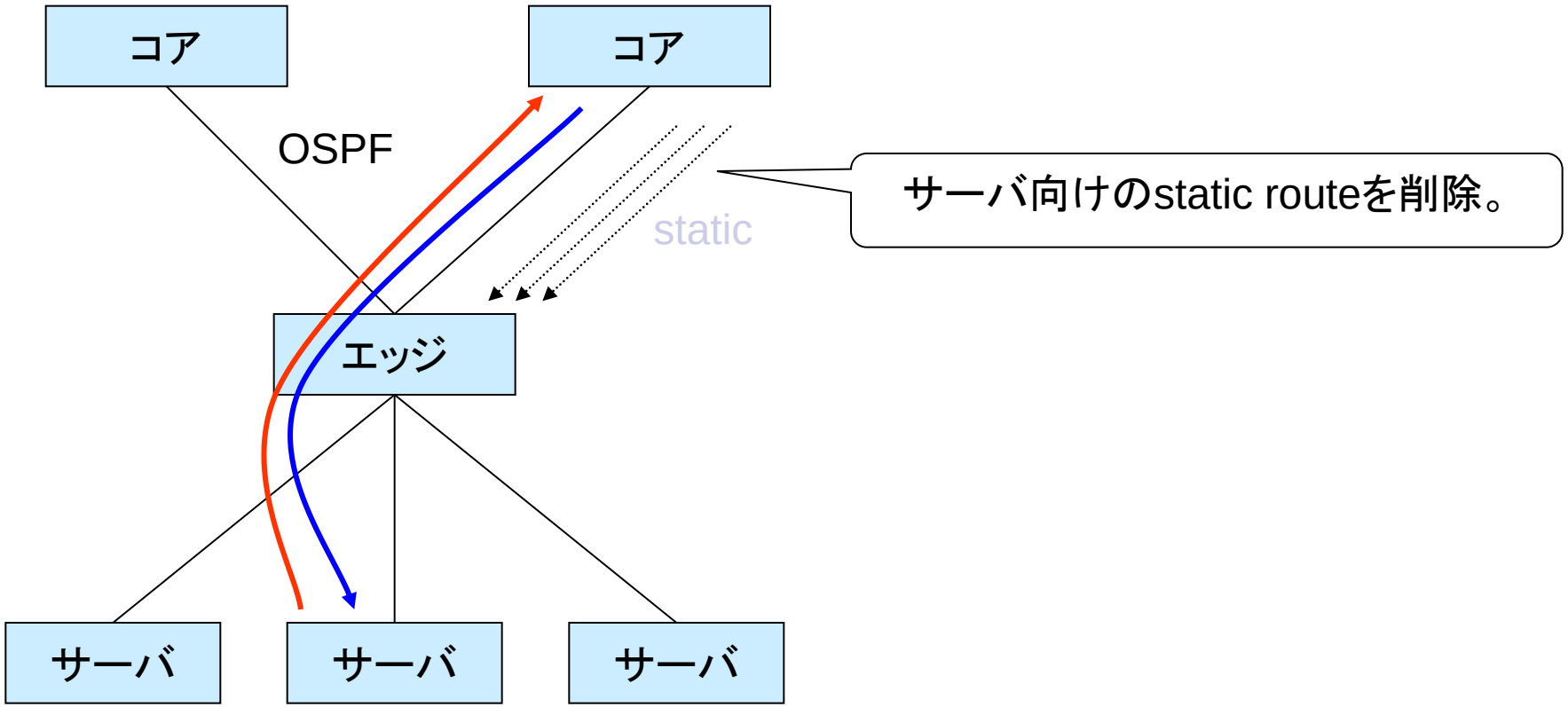
直収エッジルータの場合



直収エッジルータの場合



直収エッジルータの場合

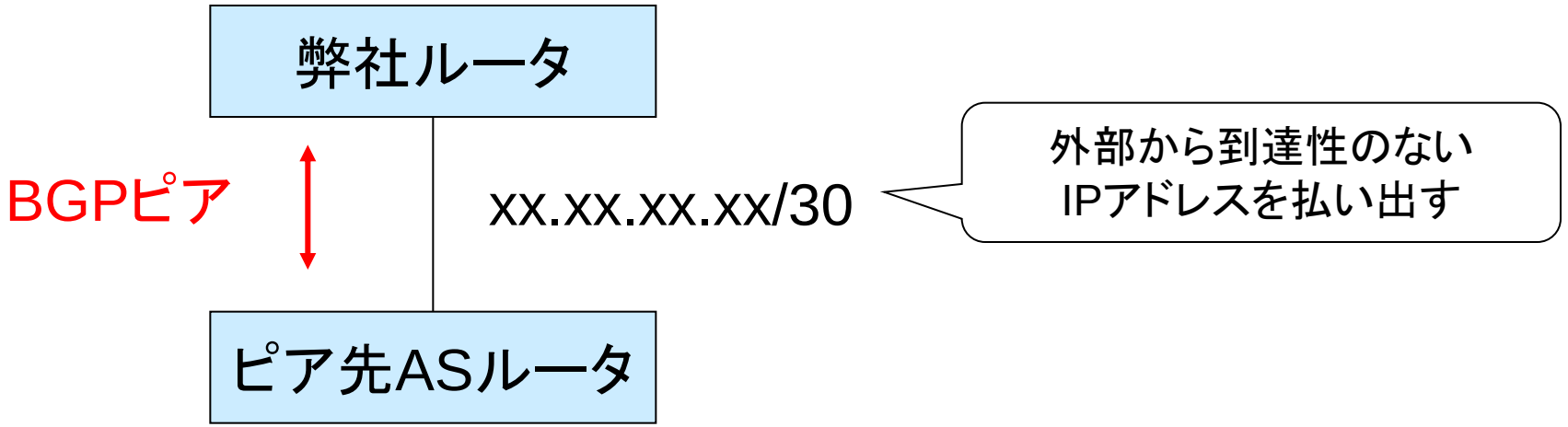


リナンバのスケジュール

- 2008/3～2008/9
  - 新規に設置するルータは、到達性のないアドレスを利用
  - ルータリプレイスの際には、ついでにリナンバを実施
  - エッジルータリナンバ変更のマニュアル作成
- 2008/10～2009/9
  - 毎週4台ずつ、ルータのリナンバメンテを実施
  - 2009/5/20現在、153台中、127台が到達性のないIPアドレスへ移行済み
  - 2009/9までには全てのルータのリナンバが完了する見込み

外部AS向けインターフェイスのリナンバ

- プライベートピア
  - 弊社から到達性のないIPアドレスを払い出し。
  - 既存ピアについては、今後リナンバ予定。
- BGPカスタマ
  - 2008/9より、お客様向けリンクアドレスとして、到達性のないアドレスを提供開始(選択可能)。



外部AS向けインターフェイスのリナンバ

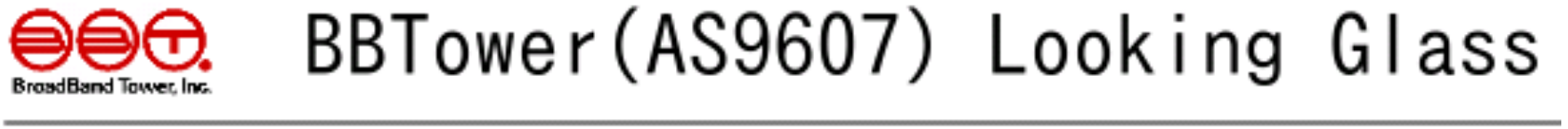
- 上位ISP(トランジット)
  - 到達性のないIPアドレスの払い出し
  - もしくは、弊社からリンクアドレスの指定
  - ISPさんによって対応が異なる。

成果のほどは？

- 原理的にルータ宛てDoSアタックの危険性は減った
- 一部リナンバできないところへのDoSアタックの可能性は残っている
  - 別途ACLの設定で防御
- tftpサーバやsyslogサーバへのアクセス制御ルールがシンプルになった。
  - ipfwの設定例

```
## Router Interface
${fwcmd} add pass ip from xx.xx.244.0/22 to any     # Tokyo Router Interface
${fwcmd} add pass ip from xx.xx.248.0/22 to any     # Tokyo Router Interface
${fwcmd} add pass ip from xx.xx.252.0/22 to any     # Osaka Router Interface
```

実際の例



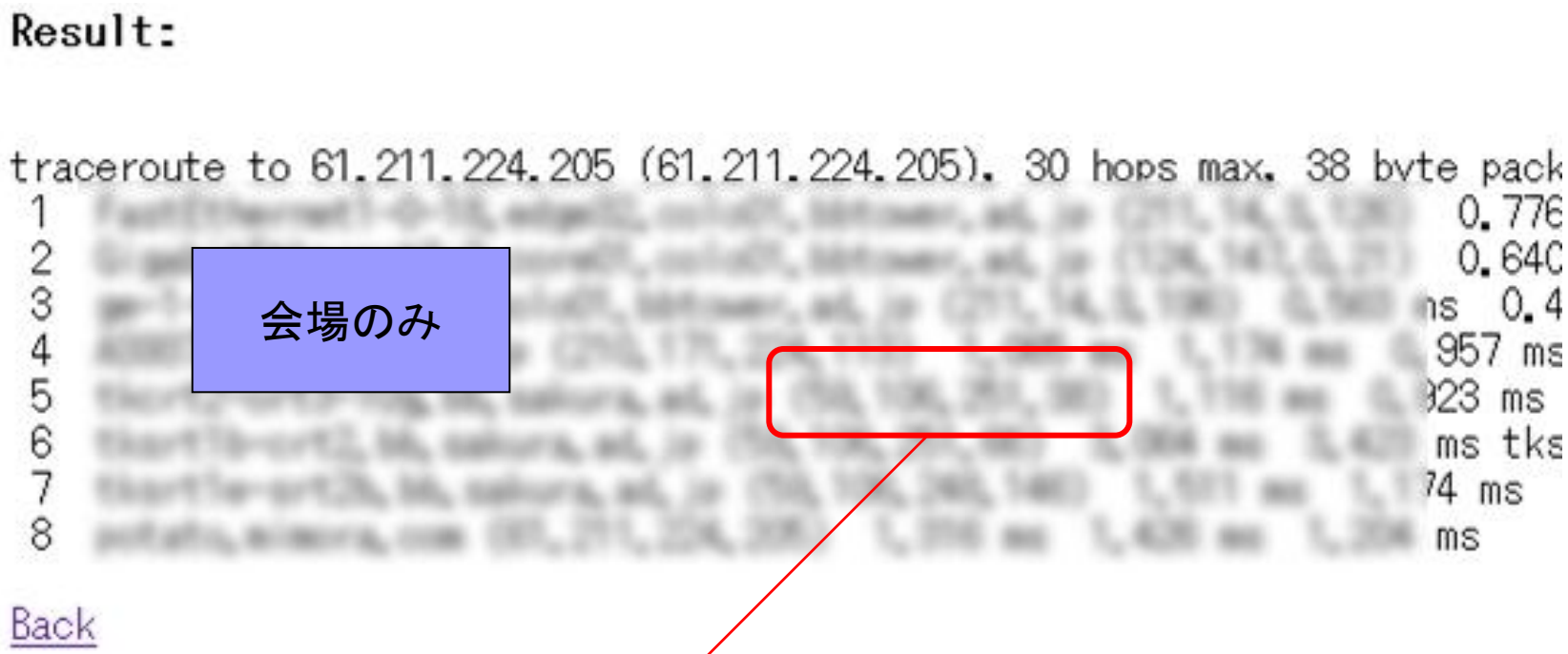
- ○ ping
- ◉ traceroute
- ○ show ip bgp (prefix is allowed as argument)
- ○ show ip bgp regexp
- ○ show ip bgp summary
- ○ ping6
- ○ traceroute6
- ○ show bgp ipv6 (prefix is allowed as argument)
- ○ show bgp ipv6 regexp
- ○ show bgp ipv6 summary

Argument: 61.211.224.205

SUBMIT RESET

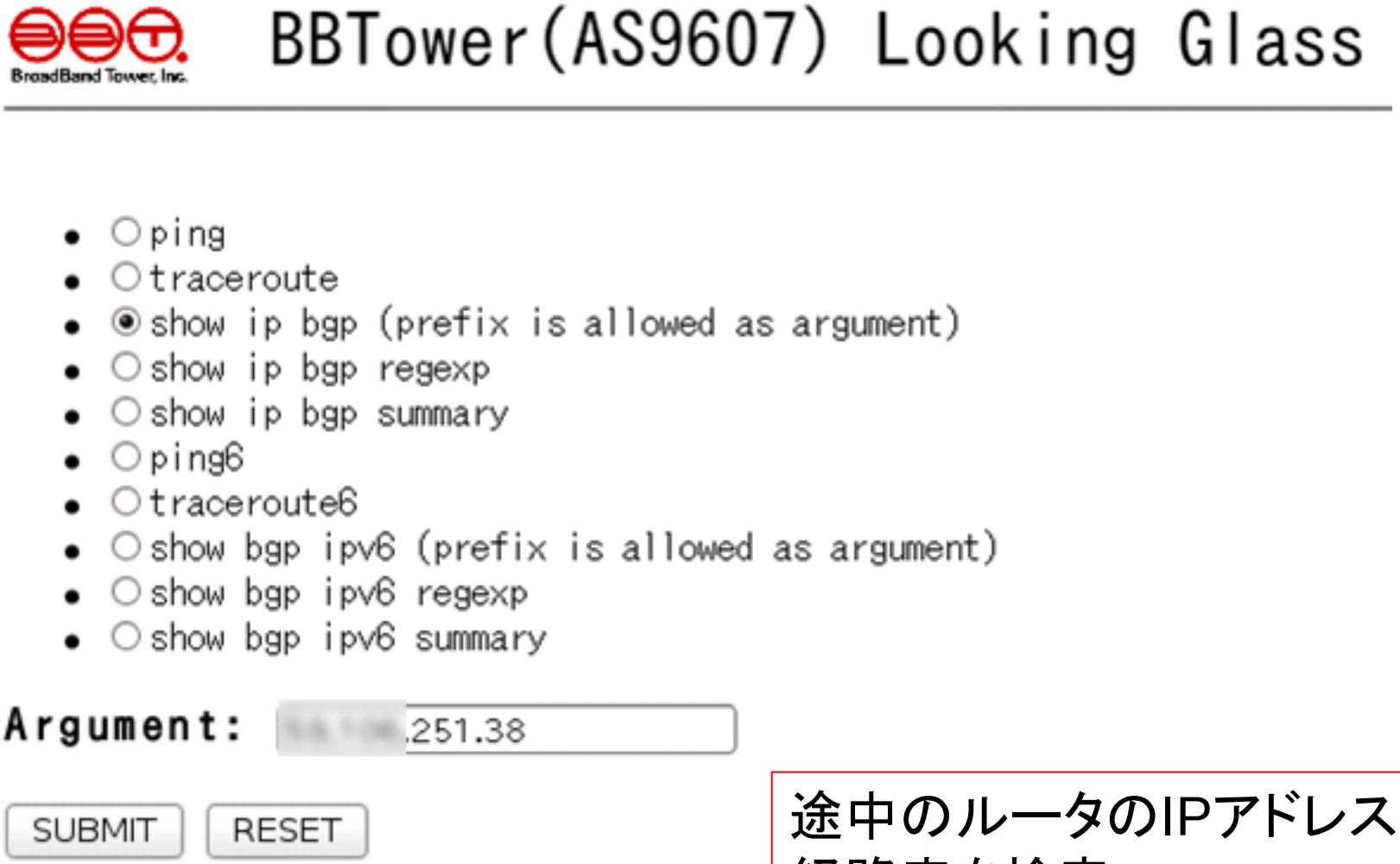
弊社AS内に存在する
サーバ宛てにtraceroute

実際の例



xx.xx.251.38
到達性なしのIPアドレスも見える

実際の例



途中のルータのIPアドレスの経路表を検索

実際の例

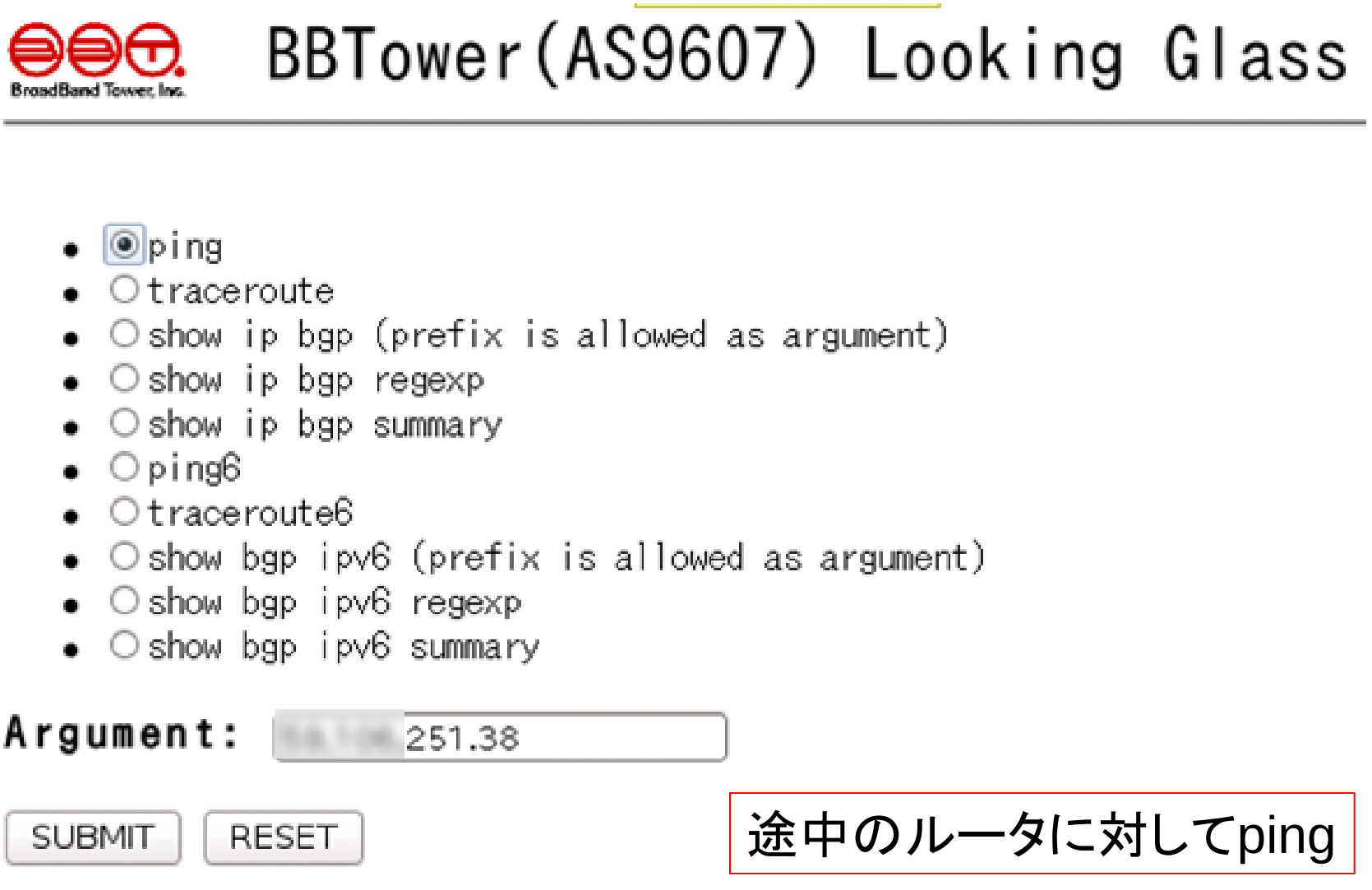
Result:

% Network not in table

Back

BGPテーブルには載っていない。

実際の例



途中のルータに対してping

実際の例

Result:

PING [redacted] 251.38 ([redacted].251.38) 56(84) bytes of data.

--- [redacted] 251.38 ping statistics ---
5 packets transmitted, 0 received, 100% packet loss, time 4009ms

Back

当然ながら到達しない。

到達性のないIPアドレスの注意点

- syslog,ntp,snmpサーバはIGPで到達できるように
  - 例：マルチフィードさんのntpサーバが使えない。
- 外部ネットワークからルータにログインできない。
  - 外部ネットワークから直接ログインする機会はない？
- ルータが発するICMPパケットの到達性に注意
  - ここ数年、uRPFがインターネットの多くのASで実装されるようになってきた。
  - 広報されていないアドレスがSrcとなったパケットは、uRPFによってDropされる。

uRPFとICMP

- 過去にも議論がありました。
  - 参照：IRS16
    石田慶樹さん
    「ICMPv6とパケットフィルタの微妙な関係」
- Dropすると困りそうなICMPパケット
  - ICMP Time exceeded
    - Tracerouteの途中で該当hopに＊が表示される。
  - ICMP Fragmentation Needed
    - Path MTU Discoveryで必須のパケット
  - ICMP (Network|Host) Unreachable
    - そもそも返さないように設定していて困らない？

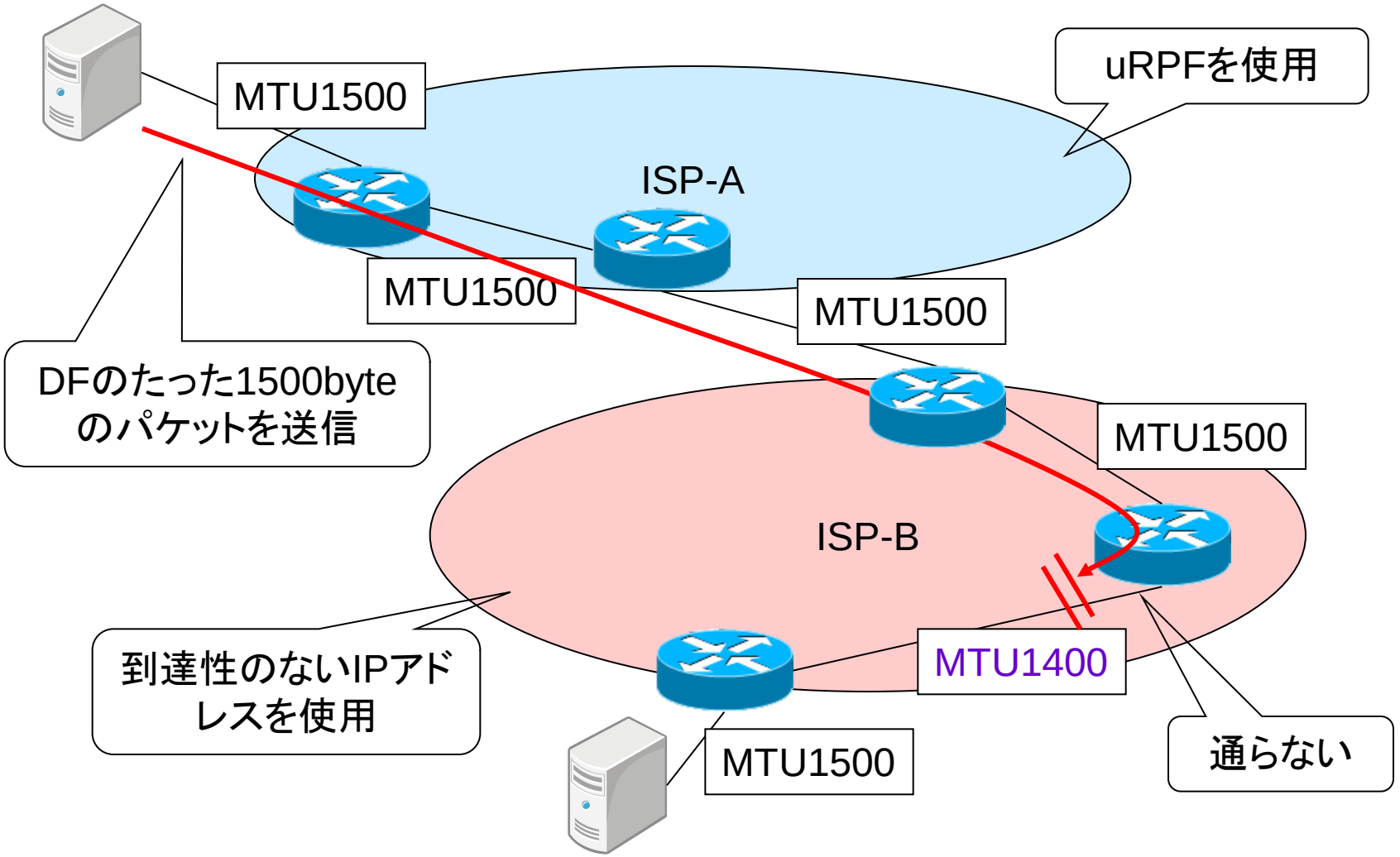
uRPFによってICMPパケットがDropしている例

abovenetのlooking glassからtraceroute。
ICMP Time exceededパケットがDropしている。

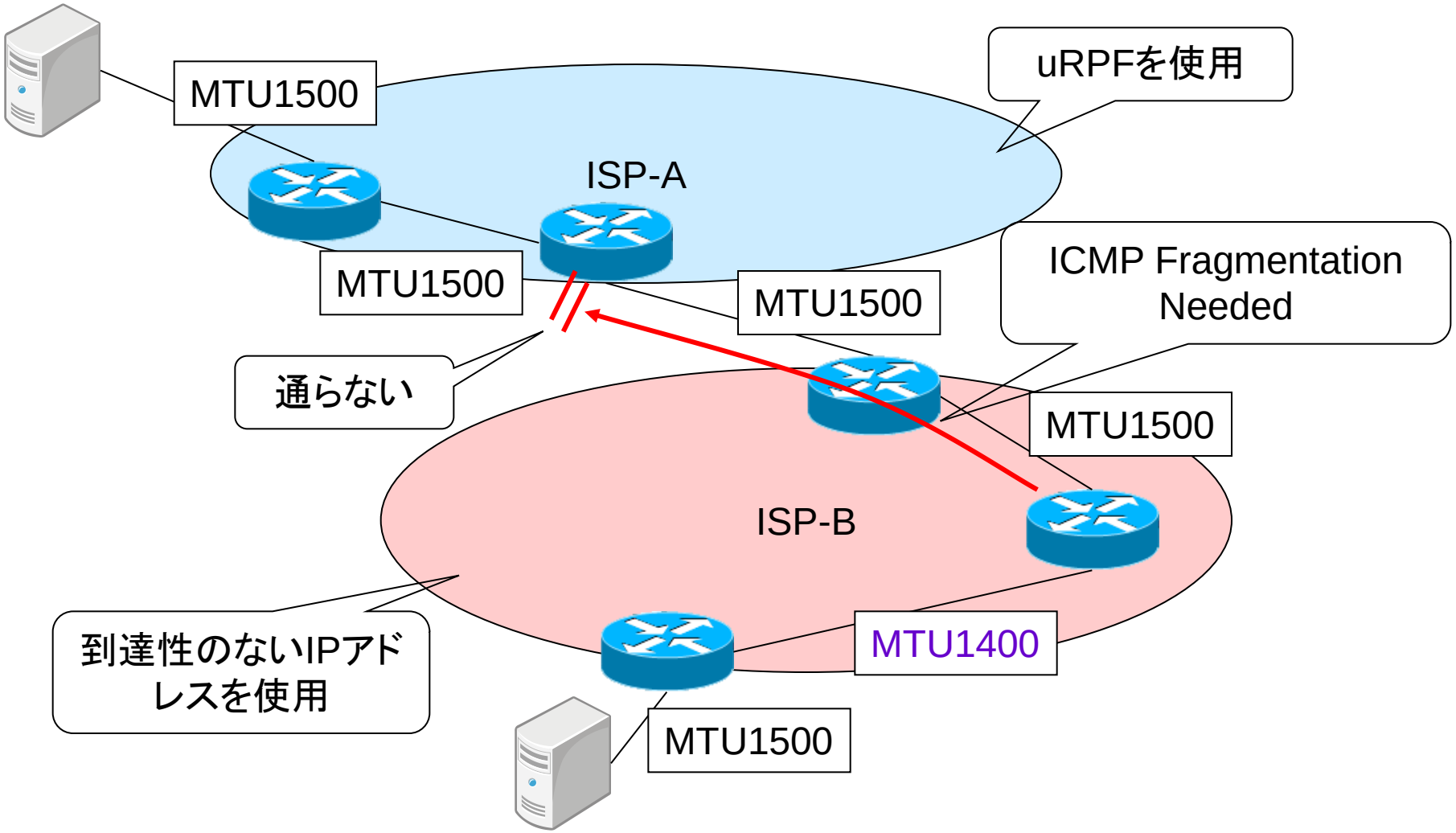
```
Router: mpr1.ams1.nl.above.net
Command: traceroute 59.106.1.133


traceroute to 59.106.1.133 (59.106.1.133), 30 hops max, 40 byte packet
 1  ten-gige-1-1.mpr1.ams2.nl.above.net (64.125.26.73)  1.202 ms  1.48
 2  ten-gige-2-2.mpr2.ams2.nl.above.net (64.125.26.70)  16.408 ms  46.
 3  pni-verio.ams2.nl.above.net (82.98.247.10)  1.128 ms  1.222 ms  1.
 4  ae-1.r22.amstnl02.nl.bb.gin.ntt.net (129.250.4.221)  23.103 ms  1.
 5  as-0.a21.tokyjp01.jp.ra.gin.ntt.net (129.250.17.61)  235.407 ms  2
 6  xe-2-1.a14.tokyjp01.jp.ra.gin.ntt.net (61.120.145.190)  244.340 ms
 7  xe-1-3.a14.tokyjp01.jp.ra.gin.ntt.net (61.120.145.170)  235.362 ms
 8  * * *
 9  * * *
10  * * *
11  report5.sakura.ad.jp (59.106.1.133)  232.362 ms  214.302 ms  236.0
```

弊社バックボーン区間

Path MTU Discoveryが機能しない例



Path MTU Discoveryが機能しない例



uRPFへの対策

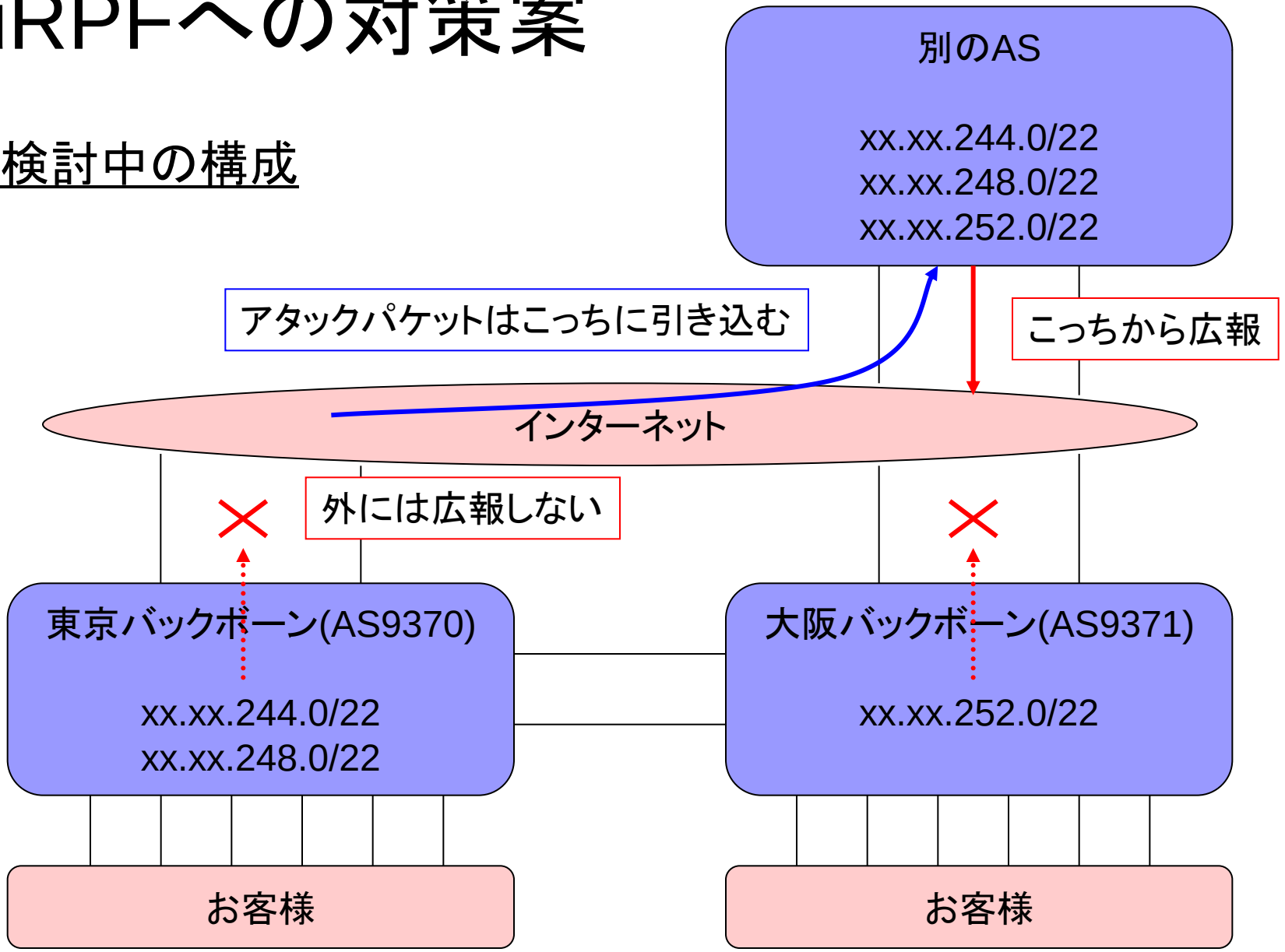
- 到達性のないアドレスを利用しているルータのリンクのMTUを、全て統一しておく

  →現在、IP MTU 1500で統一しているため、実害はない。
- 松崎さんのjanog@janogへの投稿より:

  インターネット上を流れるパケットのソースアドレスは経路情報に乗ってるのがお勧め。エラー通知のためICMPのソースアドレスになりうるアドレスは広報しとくとトラブルが少ないと思います。
- 別のところから、その空間の経路を広報してはどうか？（トラフィックはそっちに引き込む）

uRPFへの対策案

検討中の構成



議論

- ルータに設定するIPアドレスはどうするべきか？
  - 到達性のあるものを使うべき。
  - 到達性のないものを使ってもよい。
- uRPFへの対策はどうするべきか？
- 上位ISPさんへお願い
  - リンクアドレスは、インターネットから到達性のないものを使わせてください(OKの回答をいただけたところは、今のところありません。。。)
- IX事業者さんへのお願い
  - IXセグメントは、インターネットから到達性のないものをご用意いただけるとありがたいです。

議論

- 会場のISPさんに聞きたいこと
  - ルータが狙われたケースはありますか?
  - どんなアタックが多いですか?
  - ルータに設定するIPアドレスはどのようにしてますか?
  - ルータ導入時にDoSアタック耐性の検証はしていますか?
- ルータメーカーさんへ質問
  - ルータ宛てDoSアタックの耐性を上げるために工夫している仕組みはありますか？
  - uRPFに対応できる実装はありますか？

    例えば、ルータが返すICMPパケットのソースIPアドレスを指定できたりなど。