DNS Anycast と Routing に関する考察

Interdomain Routing Security (IRS) Workshop 6
2005年10月7日

石田慶樹＠文京区在住
森下泰宏＠JPRS
吉田友哉＠OCN

背景

- DNS Anycastに関するID
  - Operation of Anycast Services
    - draft-ietf-grow-anycast-01.txt (2005/07/18)
  - BGP Anycast Node for Authoritative Name Server Requirements
    - draft-morishita-dnsop-anycast-node-requirements-01.txt (2005/07/18)
- Routing に関してはあまり考察されていない
  - この場にいる関係者にInputがあることが重要
- さらに何らかのCollaboration Workも考えられる
  - I-Dの更新

内容

1. uRPFとは(復習)
2. DNS Anycastとは(復習)
3. DNS AnycastとuRPF (復習+α)
4. DNS AnycastとIRR

1. uRPF とは？(復習)

uRPF とは

- 経路情報を利用した Ingress Filter の手法
  - unicast Reverse Path Forwarding
- 利点
  - 静的な設定の整合性を保つ必要がない
  - 経路情報の変化に応じて動的に対応可能
- 欠点
  - 経路制御変更の際に細心の注意が必要

uRPF とは(続き)

- RFC3704(BCP84)
  - Ingress Filtering for Multi-homed Networks
- RPF(Reverse Path Forwarding)
  - Loose Reverse Path Forwarding w/o Default
  - Strict Reverse Path Forwarding
  - Feasible Reverse Path Forwarding

uRPF とは(続き)

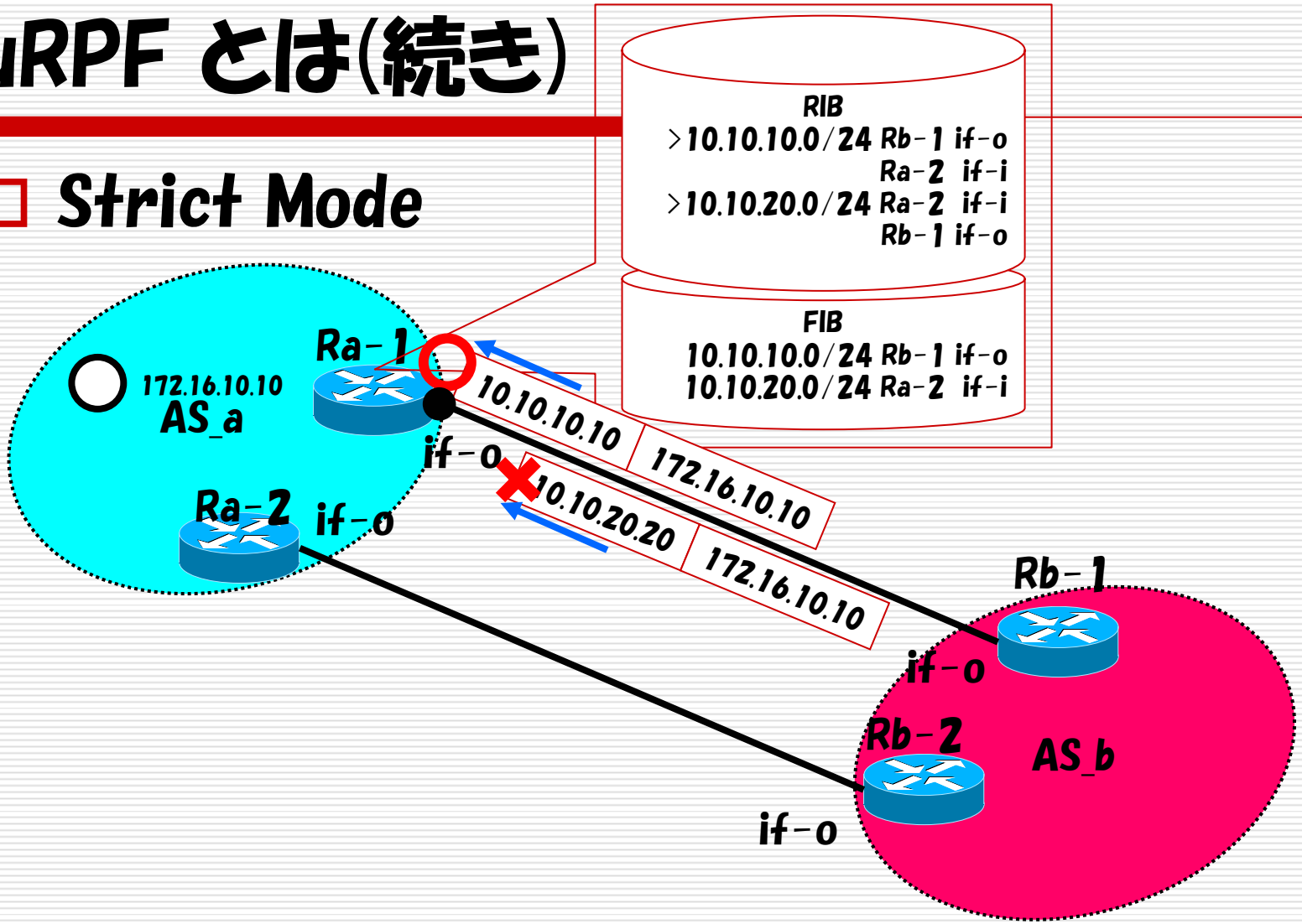
- Loose Reverse Path Forwarding
  - パケットのソースアドレスがルーティングテーブルにあるかどうかのみを確認
  - 厳密に言うと Reverse Path Forwarding ではない
  - Default経路の処理をどうするかでさらに扱いが分かれる

uRPF とは(続き)

- Strict Reverse Path Forwarding
  - パケットのソースアドレスについてFIBを参照
  - パケットを受け取ったインタフェースがforwardするべきインタフェースなら、そのパケットは通過可能
  - Strict Reverse Path Forwarding ではパスの対称性があることを前提としているが、それはありえない
  - ただしこの問題は運用技術的に解決可能である(とされている)

uRPF とは(続き)

- Strict Mode

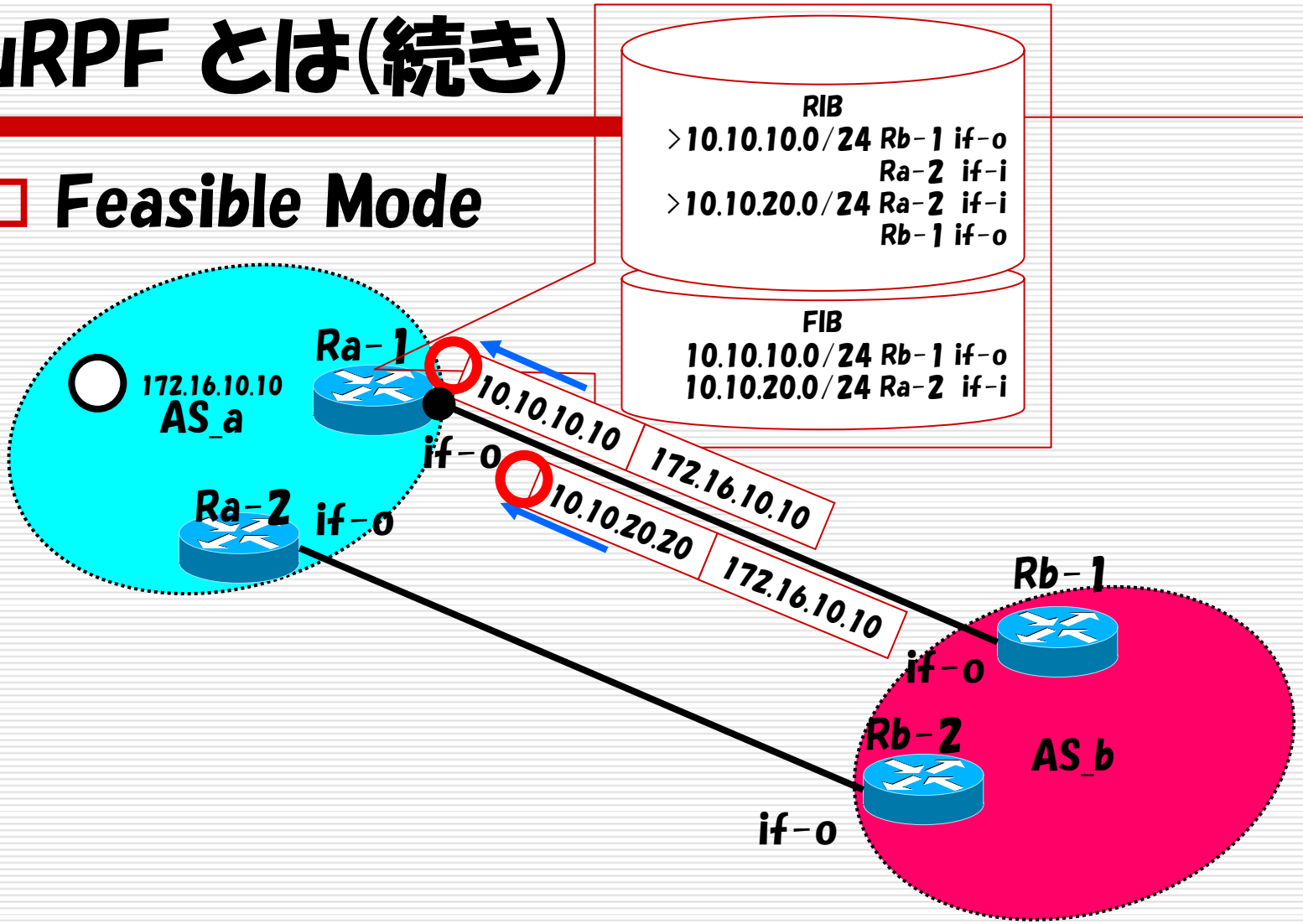


uRPF とは(続き)

- Feasible Reverse Path Forwarding
  - パケットのソースアドレスについてFIBではなくRIBを参照する
  - パケットを受け取ったインタフェースが経路的にbestでなくとも、代替経路として利用される可能性のあるインタフェースなら、パケットは通過可能
  - パスが非対称でも経路がアナウンスされていれば大丈夫(のはず)

uRPF とは(続き)

- Feasible Mode



Feasible Case Example 1

- ISP-Aと複数箇所でPeerしている場合
  （赤の部分にfeasible_RPFを設定）

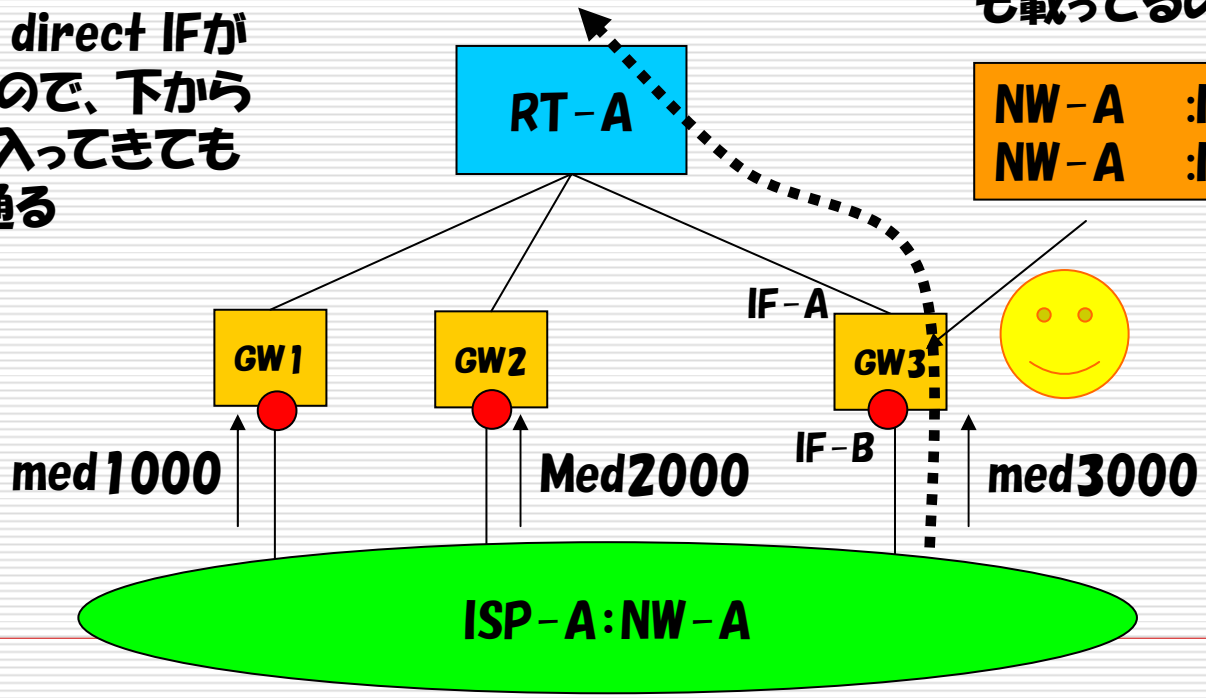
FIBだと以下

NW-A	:IF-A

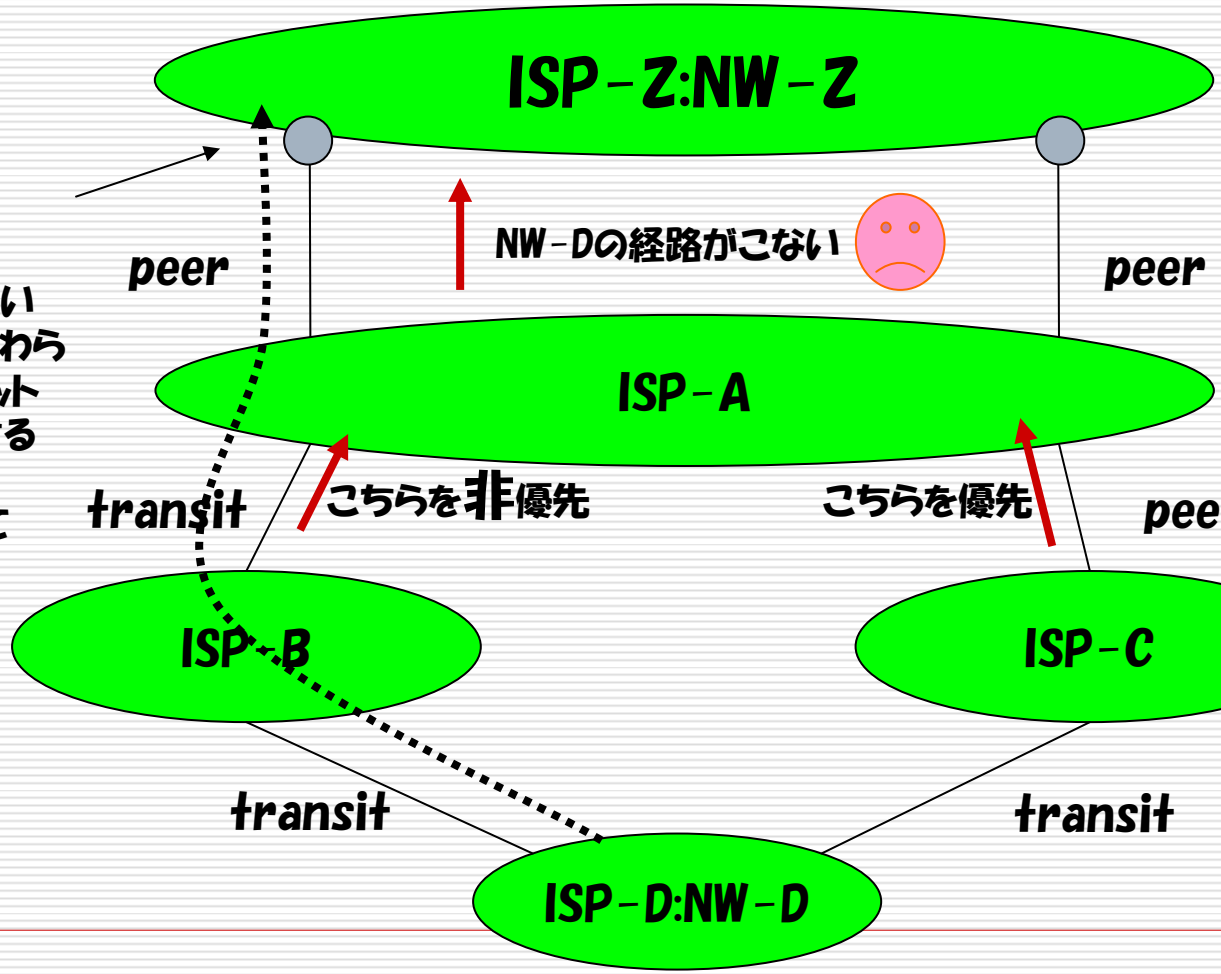
だけど、RIBにはIF-Bも載ってるので大丈夫

NW-A	:IF-A
NW-A	:IF-B

Med=3000 のGW3では、RIB上だと direct IFが載っているので、下からパケットが入ってきても問題なく通る



Feasible Case Example 2

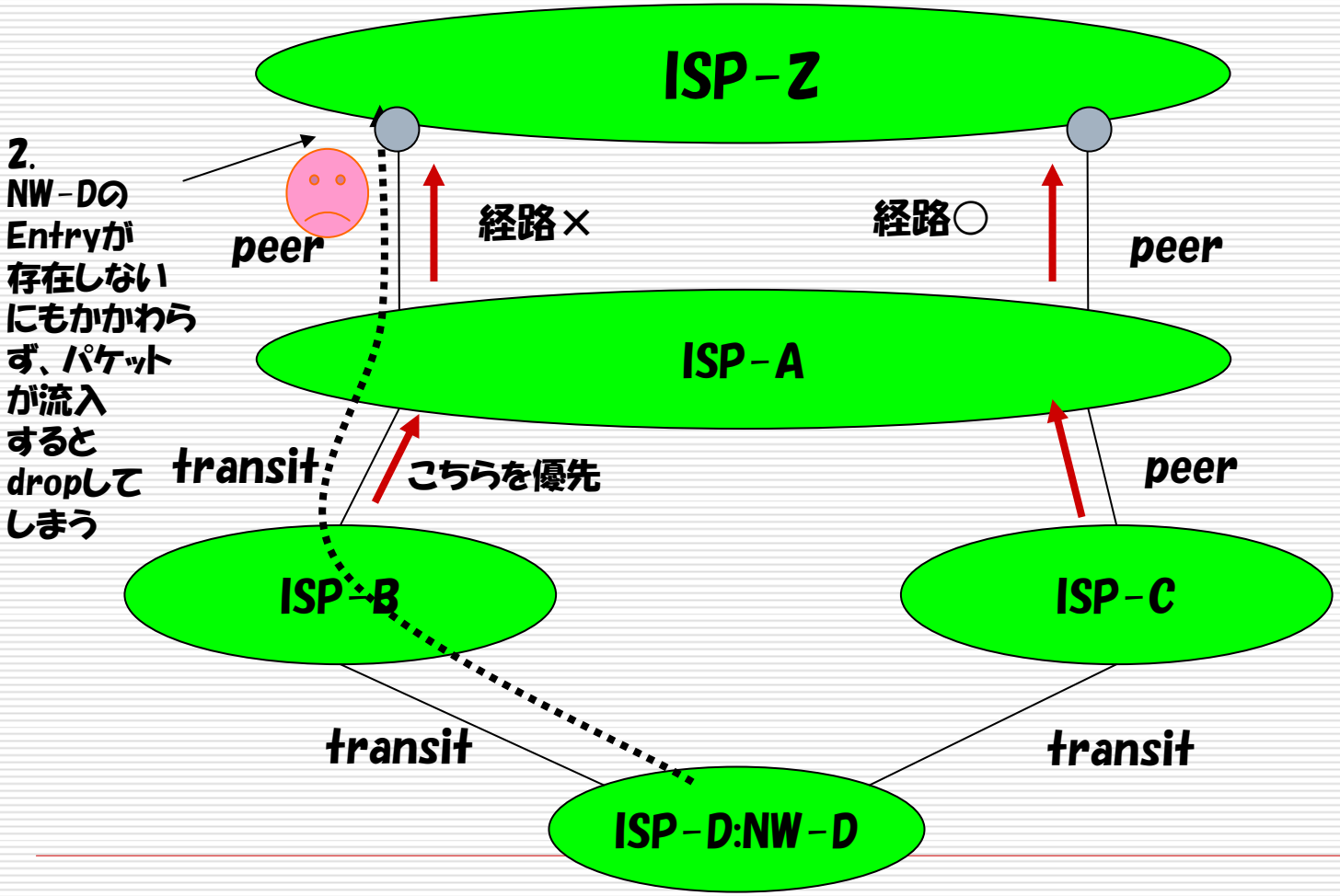


3.
NW-Dの
Entryが
存在しない
にもかかわら
ず、パケット
が流入する
すると、
Dropして
しまう

1.
ISP-AがNW-D
の経路をISP-C
経由を優先した場合、
NW:Dの経路がISP-Z
に伝播されない

2.
が、ISP-DはISP-Z
の経路:NW-ZをISP-B
及びISP-Cより受信
しているので、点線
矢印のようなパケット
の流れが発生する

Feasible Case Example 3



1.
ISP-Aの片方の
回線からのみ
NW-Dの経路が
広告されてこない
場合

Prefixを分けて
経路制御をして
いる場合(一部の
経路だけを広告)
あるいは、設定ミス
などでたまたま
該当の経路が
伝播していない
など

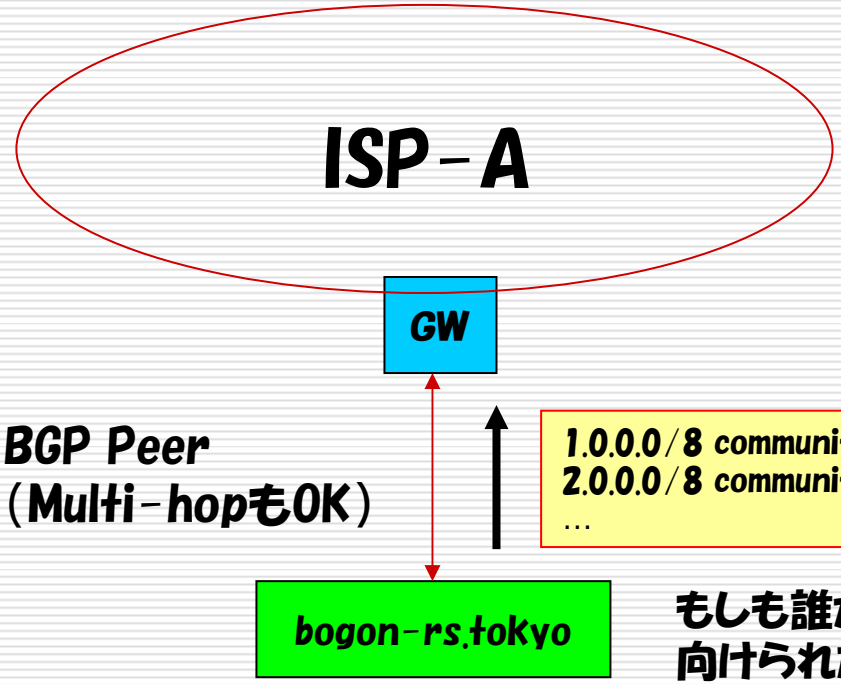
Feasible 考察

- 動くことは動く
- 使いかたに依存しそう
  - BGP Multipathを片方のASのみがやっている場合や、TEでprefix分けをしている場合、相手と非対称ルーティングを行っていると、変に経路が伝播しない場合、あるいは、RIBのupdateのタイミングが遅くなってしまったほうが、相手からのpacketをdropしてしまう可能性大
- けっこう微妙かもしれない
  - 幾つかのベンダの方々が実装したほうがいいかききにこられる今日この頃

本件とは直接関係ないですが・・・

Bogon route server in Tokyo coming soon ! Please peer !

- Bogonな最新の経路情報をBGPで配信するProject
  - http://www.cymru.com/BGP/bogon-rs.html.jis



GWで経路を受信する際に
match community=65333:888
set next-hop 192.0.2.1
などを設定し、
ip route 192.0.2.1 255.255.255.255 null0
などにしておくと、パケットを廃棄可能

もしも誰かにDefaultでbogonルート宛てにpacketを向けられた場合には、null0にたたきおとすことが出来る

2. DNS anycast とは？

DNS Anycast とは

- RFC3258
  - Distributing Authoritative Name Servers via Shared Unicast Addresses
- 13個(しかない)ルートネームサーバ
  - 負荷分散
  - 耐障害性

  のための技術
- DNSの特性を利用
  - UDPで1パケット
  - 必ずQueryに対してのAnswerとなる
- 地理的 and/or ネットワーク的に分散した複数の箇所から同一のIPアドレスブロックをアナウンスする
- IGP AnycastとBGP Anycastに大別される

DNS Anycast とは(続き)

- 利点
  - サーバ処理能力の増強
  - DDoS発生時の影響の制約
- 欠点
  - 従来の方法では監視が困難
- どれくらいDeployしているか？
  - http://www.root-servers.org/
  - 様々な理由から積極的に進行してしまっている
    - 地理的、技術的、政治的等
  - 13のルートサーバのIPアドレスに対して105箇所(2005年10月現在)
- 参考（森下が昔書いた記事）
  - 【レポート】急速に進むDNSルートサーバーのAnycast化
    - ルートサーバー間の国際協調が今後の課題に～RIPE Meetingから
  - http://internet.watch.impress.co.jp/cda/special/2003/10/03/6333.html

DNS Anycast とは(続き)

Server	Operator	Locations	IP Addr	Home ASN
A	VeriSign Naming and Directory Services	Dulles VA	198.41.0.4	19836
B	Information Sciences Institute	Marina Del Rey CA	192.228.79.201/2001:478:65::53	tba
C	Cogent Communications	Herndon VA; Los Angeles;New York City; Chicago	192.33.4.12	2149
D	University of Maryland	College Park MD	128.8.10.90	27
E	NASA Ames Research Center, Inc.	Mountain View CA	192.203.230.10	297
F	Internet Systems Consortium	Ottawa; Palo Alto; San Jose CA; New York City; San Francisco; Madrid; Hong Kong; Los Angeles; Rome; Auckland; Sao Paulo; Beijing; Seoul; Moscow; Taipei; Dubai; Paris; Singapore; Brisbane;Toronto; Monterrey; Lisbon; Johannesburg; Tel Aviv; Jakarta; Munich; Osaka; Prague; Amsterdam; Barcelona; Nairobi	192.5.5.241/2001:500::1035	3557
G	U.S. DOD Network Information Center	Vienna VA	192.112.36.4	568
H	U.S. Army Research Lab	Aberdeen MD	128.63.2.53/2001:500:1::803f:235	13
I	Autonomica/NORDUnet	Stockholm; Helsinki; Milan; London; Geneva; Amsterdam;Oslo; Bangkok; Hong Kong; Brussels; Frankfurt; Ankara; Bucharest; Chicago; Washington DC; Tokyo; Kuala Lumpur; Palo Alto; Jakarta; Wellington; Johannesburg; Perth; San Francisco; New York; Singapore; Miami; Ashburn (US); Mumbai	192.36.148.17	29216
J	VeriSign Naming and Directory Services	Dulles VA (4 locations); Mountain View CA; Seattle WA;Atlanta GA; Los Angeles CA; Miami FL; Sunnyvale CA; Amsterdam; Stockholm; London; Tokyo; Seoul; Singapore	192.58.128.30	26415
K	Reseaux IP Europeens -Network Coordination Centre	London (UK); Amsterdam (NL); Frankfurt (DE); Athens (GR); Doha (QA); Milan (IT); Reykjavik (IS); Helsinki (FI); Geneva (CH); Poznan (PL); Budapest (HU); Abu Dhabi (AE); Tokyo (JP); Brisbane (AU); Miami (US)	193.0.14.129/2001:7fd::1	25152
L	Internet Corporation for Assigned Names and Numbers	Los Angeles	198.32.64.12	20144
M	WIDE Project	Tokyo; Seoul (KR); Paris (FR)	202.12.27.33/2001:dc3::35	7500

http://www.root-servers.org/ より

DNS Anycast とは(続き)

- DNS Anycast に関する2つのI-D
  - Operation of Anycast Services
    - draft-ietf-grow-anycast-01.txt
      IP Anycastを用いた分散サービスを構築する際の留意点や推奨点を、IGPとBGPの双方について体系的にまとめようとしているもの。このI-Dはサービス対象をDNSのみに限定していないが、参照例や実装例としてDNSにおけるケースが多く登場する。著者はISCでF.root-servers.netを管理運用しているJoe Abley氏。
  - BGP Anycast Node for Authoritative Name Server Requirements
    - draft-morishita-dnsop-anycast-node-requirements-01.txt
      権威DNSサーバにIP Anycastを導入する場合に必要となる、具体的な要求事項・前提条件についてまとめようとしているもの。既存ドキュメント(RFC 2182, 2870, 3258, draft-ietf-grow-anycast-...)での規定事項を実現する、あるいはAnycastの導入によってそれらがおびやかされることがないようにするために必要な事項・条件について考察しようとしている。著者はJPRS佐藤・松浦・森下。

DNS Anycast とは(続き)

- draft-ietf-grow-anycast-01.txt
  - 4.4 Routing Considerations 9
    - 4.4.1 Signalling Service Availability 9
    - 4.4.2 Covering Prefix 10
    - 4.4.3 Equal-Cost Paths 10
    - 4.4.4 Route Dampening 11
    - 4.4.5 Reverse Path Forwarding Checks 12
    - 4.4.6 Propagation Scope 12
    - 4.4.7 Other Peoples' Networks 13
    - 4.4.8 Aggregation Risks 14

DNS Anycast とは(続き)

- draft-ietf-grow-anycast-01.txt
  - 4.4.5 Reverse Path Forwarding Checks
    Reverse Path Forwarding (RPF) checks, first described in [RFC2267], are commonly deployed as part of ingress interface packet filters on routers in the Internet in order to deny packets whose source addresses are spoofed (see also RFC 2827 [RFC2827]). Deployed implementations of RPF make several modes of operation available (e.g. "loose" and "strict").

    Some modes of RPF can cause non-spoofed packets to be denied when they originate from multi-homed site, since selected paths might legitimately not correspond with the ingress interface of non-spoofed packets from the multi-homed site. This issue is discussed in [RFC3704].

    A collection of anycast nodes deployed across the Internet is largely indistinguishable from a distributed, multi-homed site to the routing system, and hence this risk also exists for anycast nodes, even if individual nodes are not multi-homed. Care should be taken to ensure that each anycast node is treated as a multi-homed network, and that the corresponding recommendations in [RFC3704] with respect to RPF checks are heeded.

DNS Anycast とは(続き)

- draft-morishita-dnsop-anycast-node-requirements-01.txt
  - RoutingへのRequirementに関する言及は今のところなし
    - これから書く予定(コメント求む)
  - Routingというワードは6回だけ出現

3. DNS Anycast と uRPF

DNS Anycast と uRPF

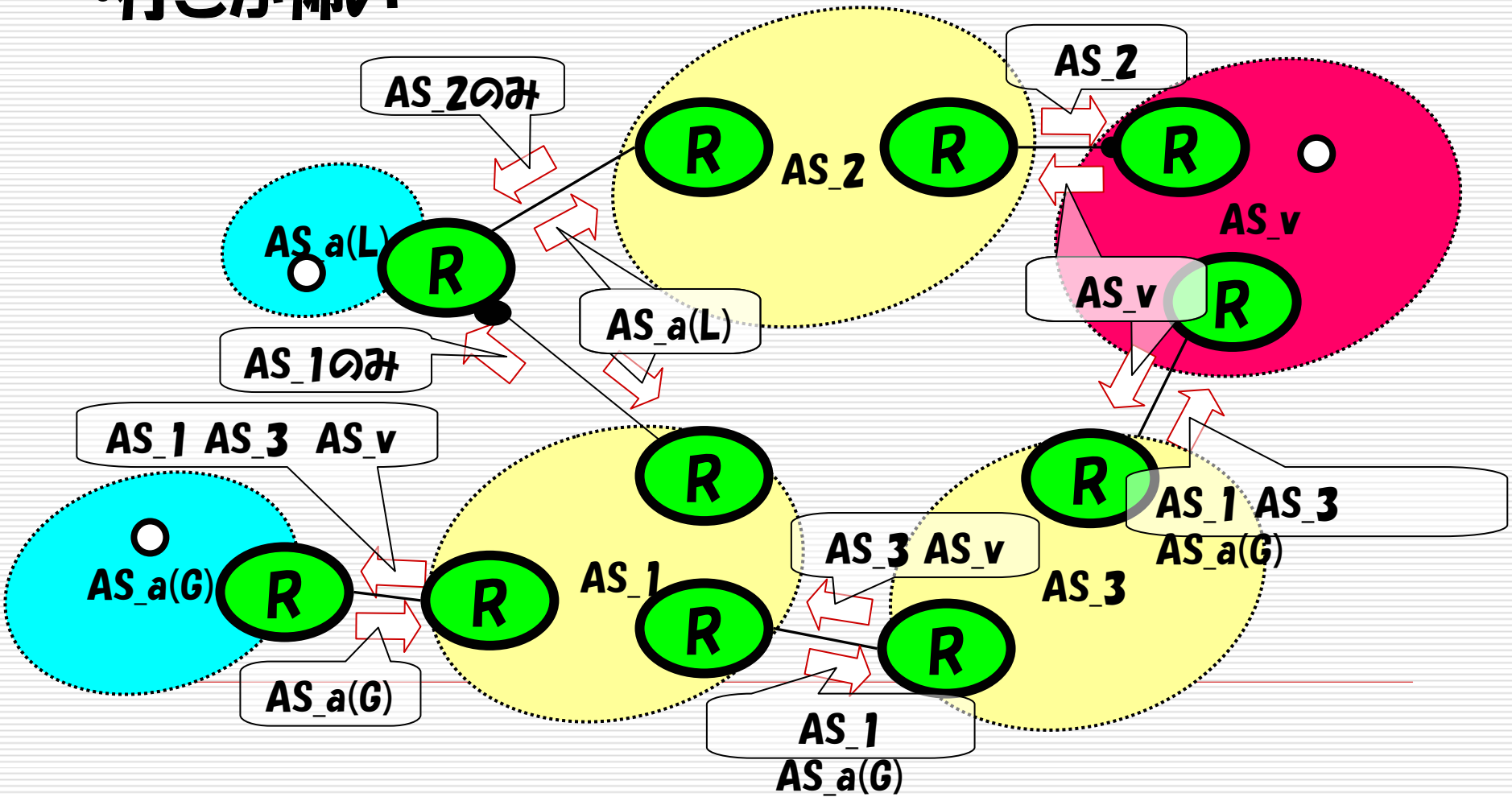
- 現在のrootサーバにおけるDNS Anycastの実装
  - Global Nodeが少数
  - Local Nodeが多数
  - No-ExportのCommunityをつけている場合がある
  - しかしLocal Nodeも部分的にはトランジットされてしまっている

DNS Anycast と uRPF

- 何が問題か？
  - 従来のインターネットにおける基本原則である「IPアドレスにより通信相手を特定した1対1通信」が、根本から崩れる
  - 基本的にはRoutingの非対称性であるが。。。
  - No-Export Communityの利用
- 異常発生時に大問題⇒問題解決が非常にやりにくい
  - 関係者が複数存在している
    - DNS Anycastのグローバルな管理者
    - DNS Anycastのローカルサイトのオペレータ
    - DNS Anycastとピアしているオペレータ
    - 異常を受けた当事者(とばっちり)
  - 当事者には問題の原因もよくわからない
  - 途中の経路上にRoutingとDNS Anycastの両方を理解したオペレータが必要

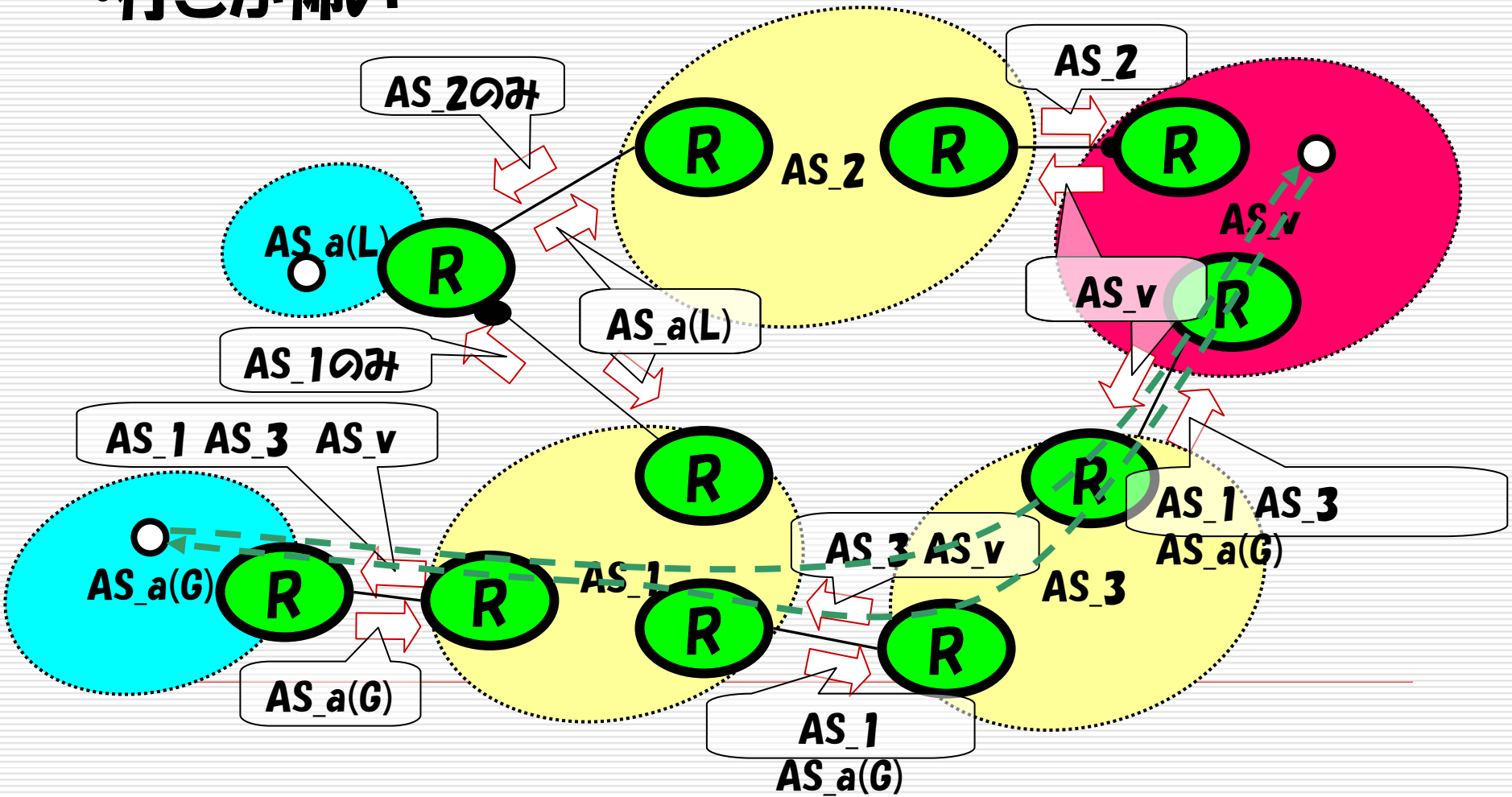
uRPF vs BGP Anycast 問題(その1)

- 行きが怖い



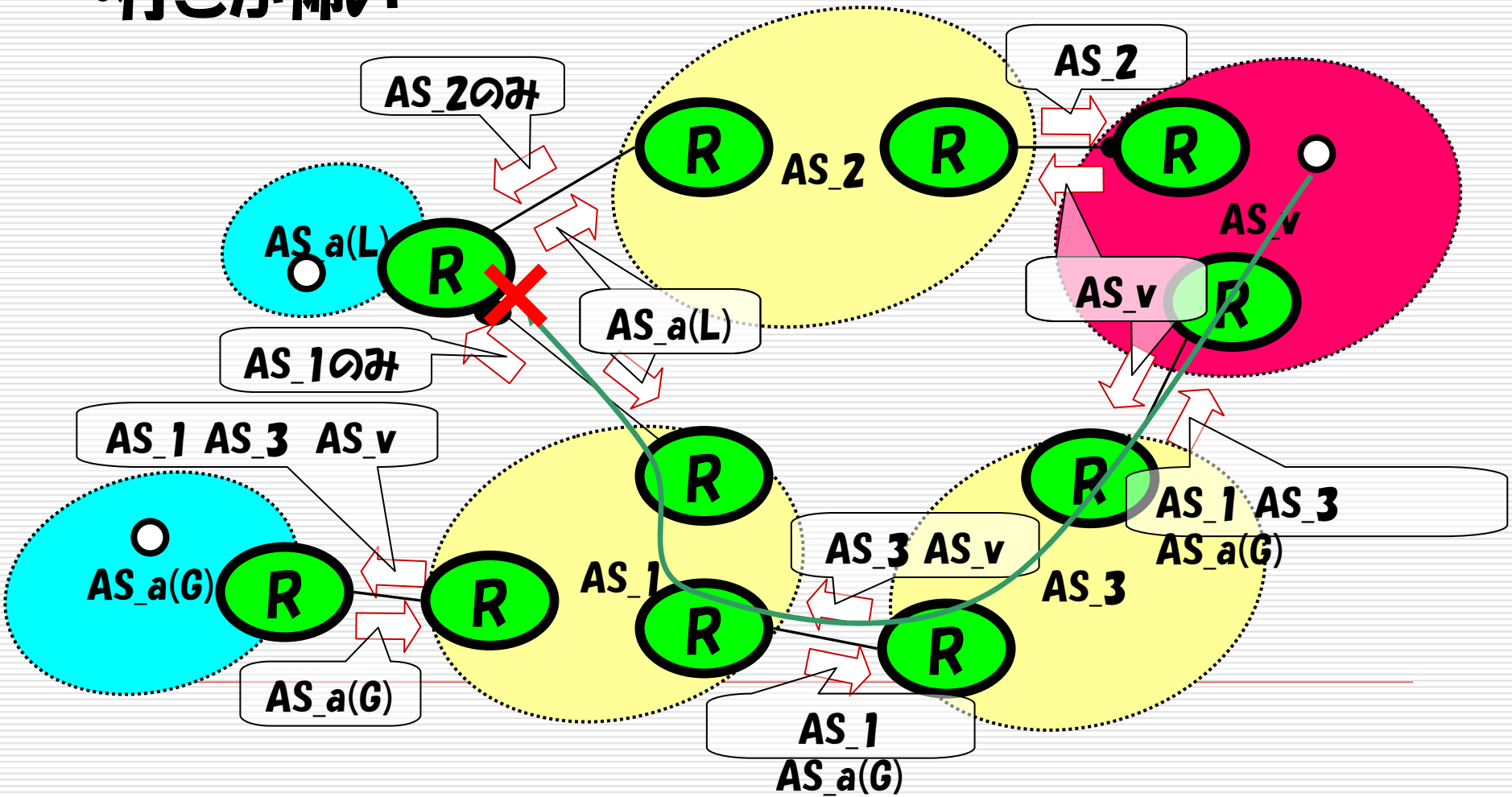
uRPF vs BGP Anycast 問題(その1)

- 行きが怖い



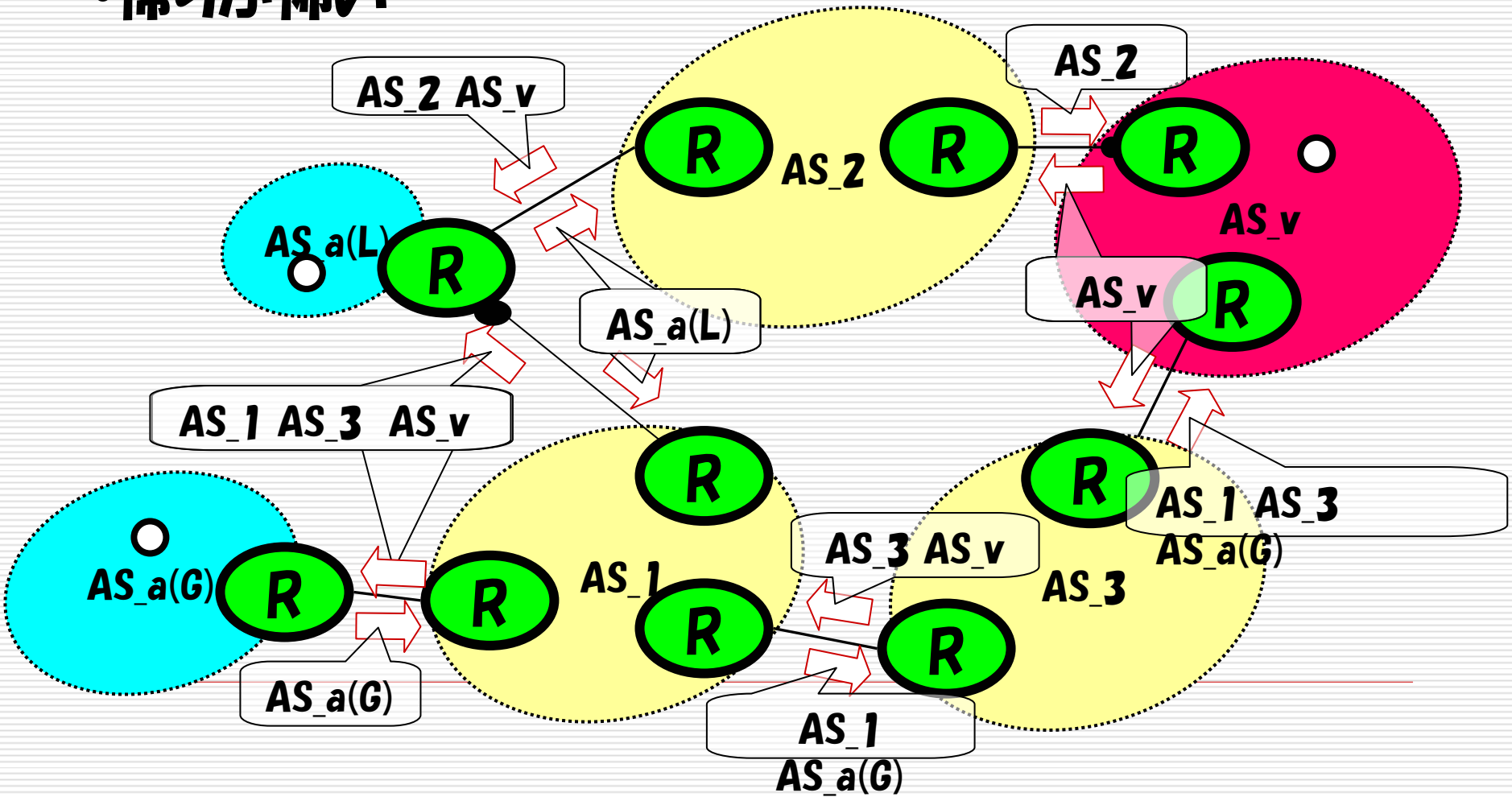
uRPF vs BGP Anycast 問題(その1)

- 行きが怖い



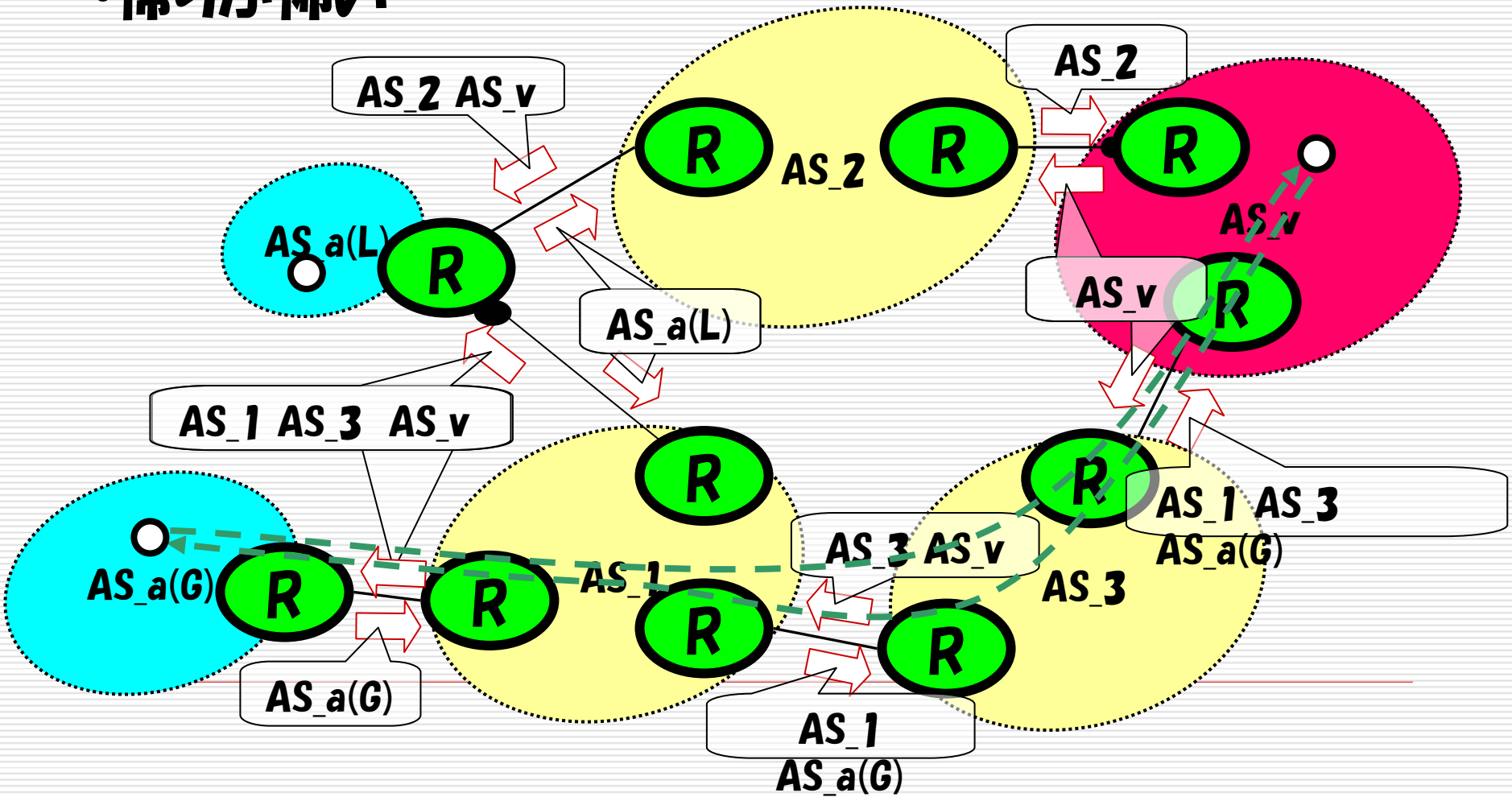
uRPF vs BGP Anycast 問題(その2)

- 帰りが怖い



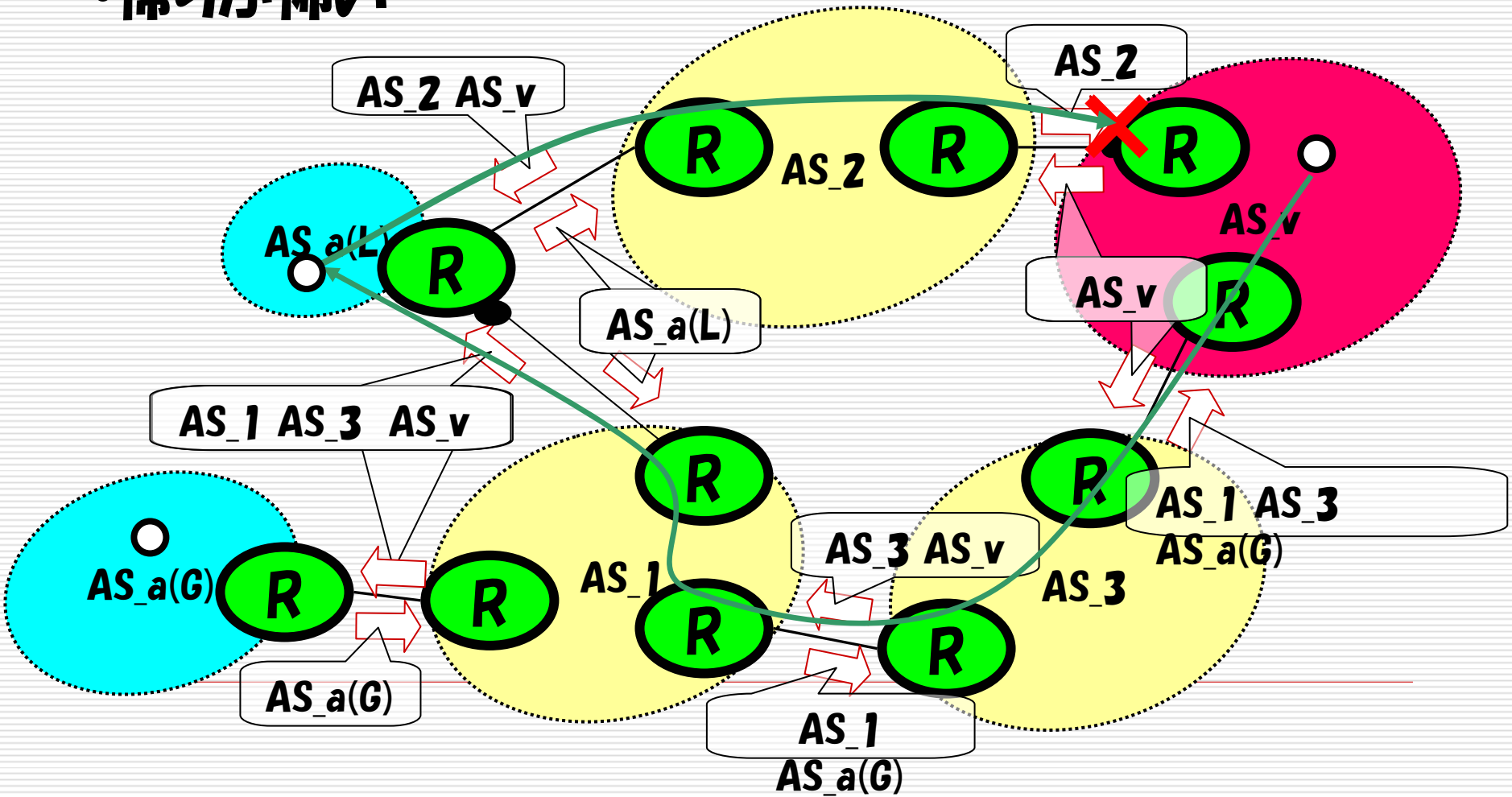
uRPF vs BGP Anycast 問題(その2)

- 帰りが怖い



uRPF vs BGP Anycast 問題(その2)

- 帰りが怖い

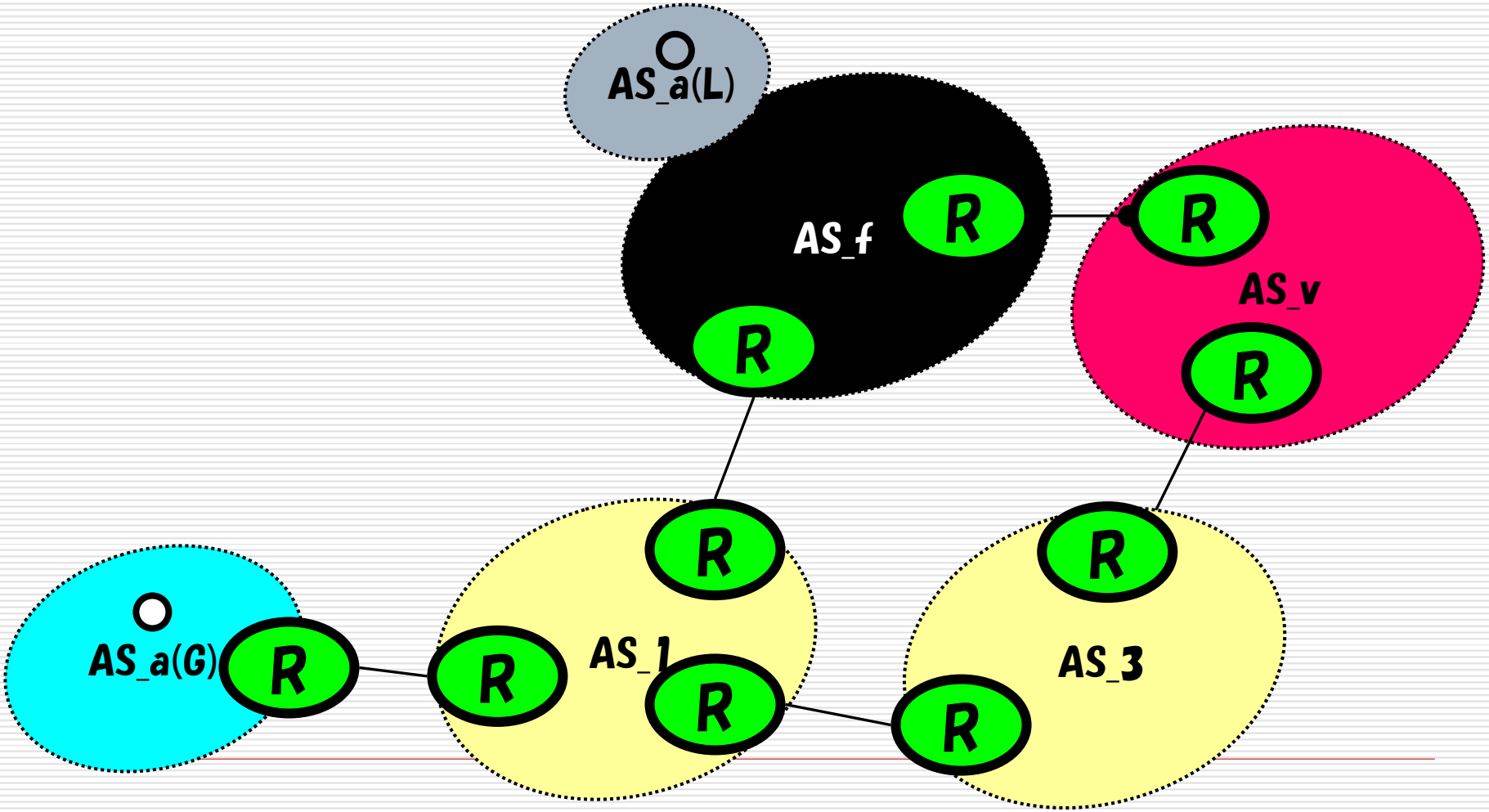


4. DNS Anycast と IRR

DNS Anycastを騙る

- 懸念事項！
  - DNS Anycast のAS番号/IPアドレスを騙られた時に、それをいちはやく検出できるか？
  - DNS Anycastの経路がどこから聞こえてくるのが正しいか確認できるか？
- あくまで可能性ですが。。。
  - AS番号とIPアドレスブロックは既知
  - 悪者がトランジットしているISPのふりをすれば。。。
  - 目的は？
    ⇒もちろんPharming, Phishing

DNS Anycastを騙る



DNS AnycastとIRR

- 経路の検証にIRRを使う手があるが
  - 騙られたDNS Anycastに対して利用可能か？

```
% whois -h whois.ra.net. as3557
aut-num:     AS3557
as-name:     ISC-CALIFORNIA
descr:       Internet Systems Consortium, Inc.
admin-c:     PV860-RIPE
tech-c:      JA856-RIPE
export:      to AS3557:AS-FLN announce AS3557 AND {192.5.5.0/24}
export:      to AS-ANY announce AS3557:AS-ISC
mp-export:    afi ipv4 to AS-ANY announce AS3557:AS-ISC
mp-export:    afi ipv4 to AS3557:AS-FLN announce AS3557 AND {192.5.5.0/24}
mp-export:    afi ipv6 to AS-ANY announce AS3557:AS-ISC6
mp-export:    afi ipv6 to AS3557:AS-FLN announce AS3557 AND {2001:500::/48}
remarks:     ----------------------------------------
remarks:     Contacts per RFC2142:
remarks:
remarks:     Abuse/UCE: abuse@isc.org
remarks:     Security: noc@isc.org
remarks:
remarks:     Peering information can be found at
remarks:     http://www.isc.org/peering
remarks:     ----------------------------------------
notify:      noc^chat@isc.org
mnt-by:      MAINT-ISC
changed:     jabley@isc.org 20050124
source:      RIPE
```

DNS AnycastとIRR

- その他のASはピア相手をすべてを尽くさず
- IRRの専門家の方教えてください！
  - 悪意を持ったアナウンスに対しては？
  - つまるところセキュアなBGPが来る必要がある？