MYANMAR NETWORK OPERATORS GROUP

BGP Routing Security

Sai Nyan Lynn Swe
CCIE # 38501 (R&S , SP and DC) , CISSP
OPTIMITY Co Ltd

OPTIMITY

Agenda

01 Internet Today

02 Peering Security

03 Maximum BGP Prefix Tracking

04 Limiting AS Path Length

05 Filtering BGP route

06 Remotely Triggered Black Hole (RTBH)

07 Internet Routing Registry

The Internet Today (May 2024)

- Current IPv4 Internet Routing Table Statistics

BGP Routing Table Entries	951420
Prefixes after maximum aggregation	361839
Unique prefixes in Internet	462288
/24s announced	579462
ASNs in use	75837

  - (maximum aggregation is calculated by Origin AS)
  - (unique prefixes > max aggregation means that operators are announcing prefixes from their blocks without a covering aggregate)

The Internet Today (May 2024)

- Current IPv6 Internet Routing Table Statistics

BGP Routing Table Entries	194552
/48s announced	93851
ASNs in use	32692

MD5 keys on BGP peerings

- Use passwords on all BGP sessions
  - Not being paranoid, VERY necessary
  - It's a secret shared between you and your peer
  - If arriving packets don't have the correct MD5 hash, they are ignored
  - Helps defeat miscreants who wish to attack BGP sessions
- Powerful preventative tool, especially when combined with filters and GTSM

```
router bgp 100
 address-family ipv6
  neighbor 2001:db8::1 remote-as 200
  neighbor 2001:db8::1 description Peering with AS200
  neighbor 2001:db8::1 password 7 030752180500
!
```

BGP Maximum Prefix Tracking

- Allow configuration of the maximum number of prefixes a BGP router will receive from a peer
- Two level control:
  - Warning threshold: log warning message
  - Maximum: tear down the BGP peering, manual intervention required to restart

```
neighbor <x.x.x.x> maximum-prefix <max> [restart N] [<threshold>] [warning-only]
```

- Optional keywords:
  - `restart` will restart the BGP session after N minutes
  - <threshold> sets the warning level (default 75%)
  - `warning-only` only sends warnings

Limiting AS Path Length

- Some BGP implementations have problems with long AS_PATHS
  - Memory corruption
  - Memory fragmentation
- Even using AS_PATH prepends, it is not normal to see more than 20 ASes in a typical AS_PATH in the Internet today
  - The Internet is around 5 ASes deep on average
  - Largest AS_PATH is usually 16-20 ASNs

```
neighbor x.x.x.x maxas-limit 15
```

Limiting AS Path Length

- Some announcements have ridiculous lengths of AS-paths:

```
*> 3FFE:1600::/24     22 11537 145 12199 10318 10566 13193 1930 2200
   3425 293 5609 5430 13285 6939 14277 1849 33 15589 25336 6830 8002
   2042 7610 i
```

This example is an error in one IPv6 implementation

```
*>i193.105.15.0       2516 3257 50404 50404 50404 50404 50404 50404
   50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
   50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
   50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
   50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
   50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
   50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
   50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
   50404 50404 50404 50404 50404 50404 50404 i
```

This example shows 100 prepends (for no obvious reason)

- If your implementation supports it, limit the maximum AS-path length you will accept

Filtering BGP route

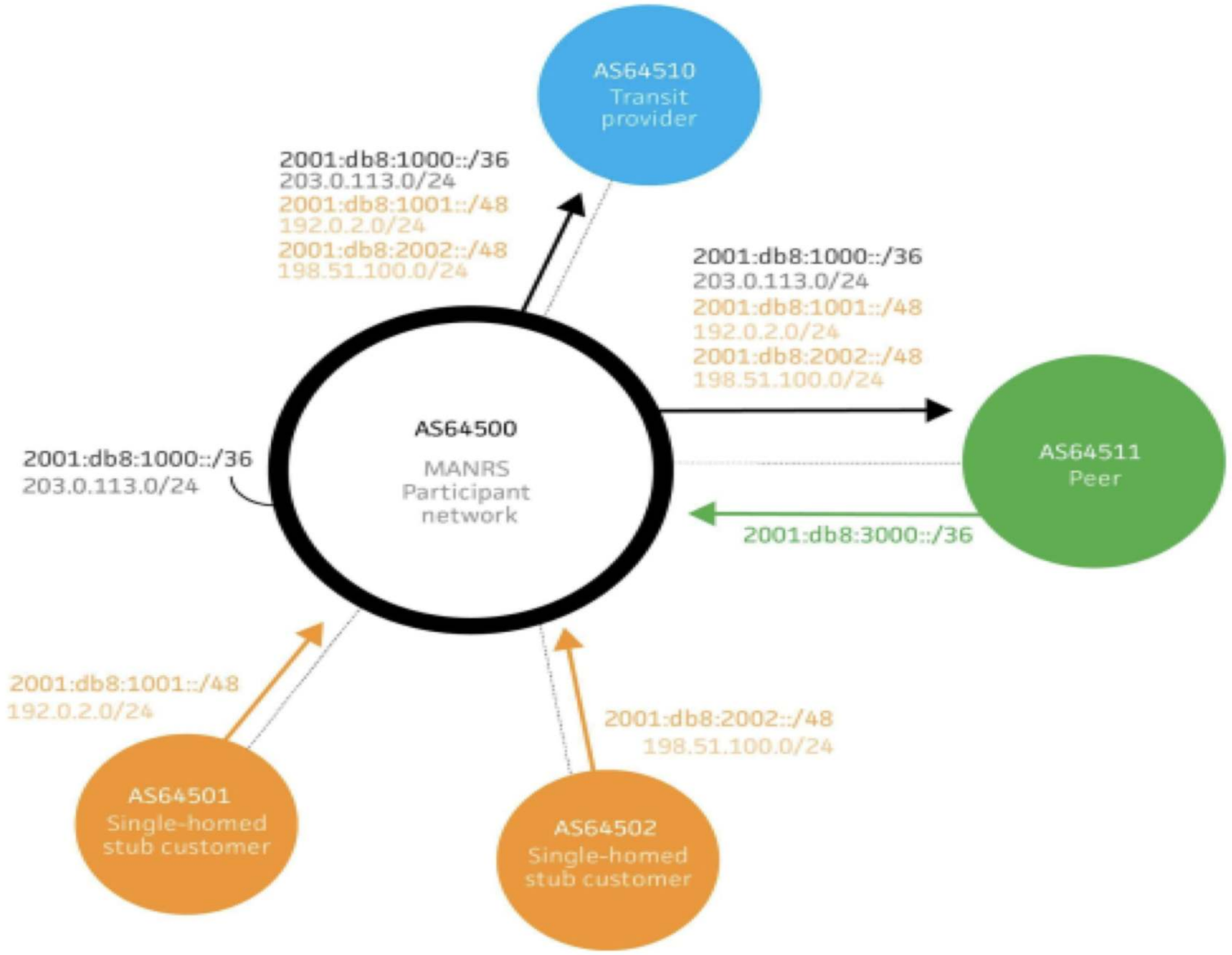
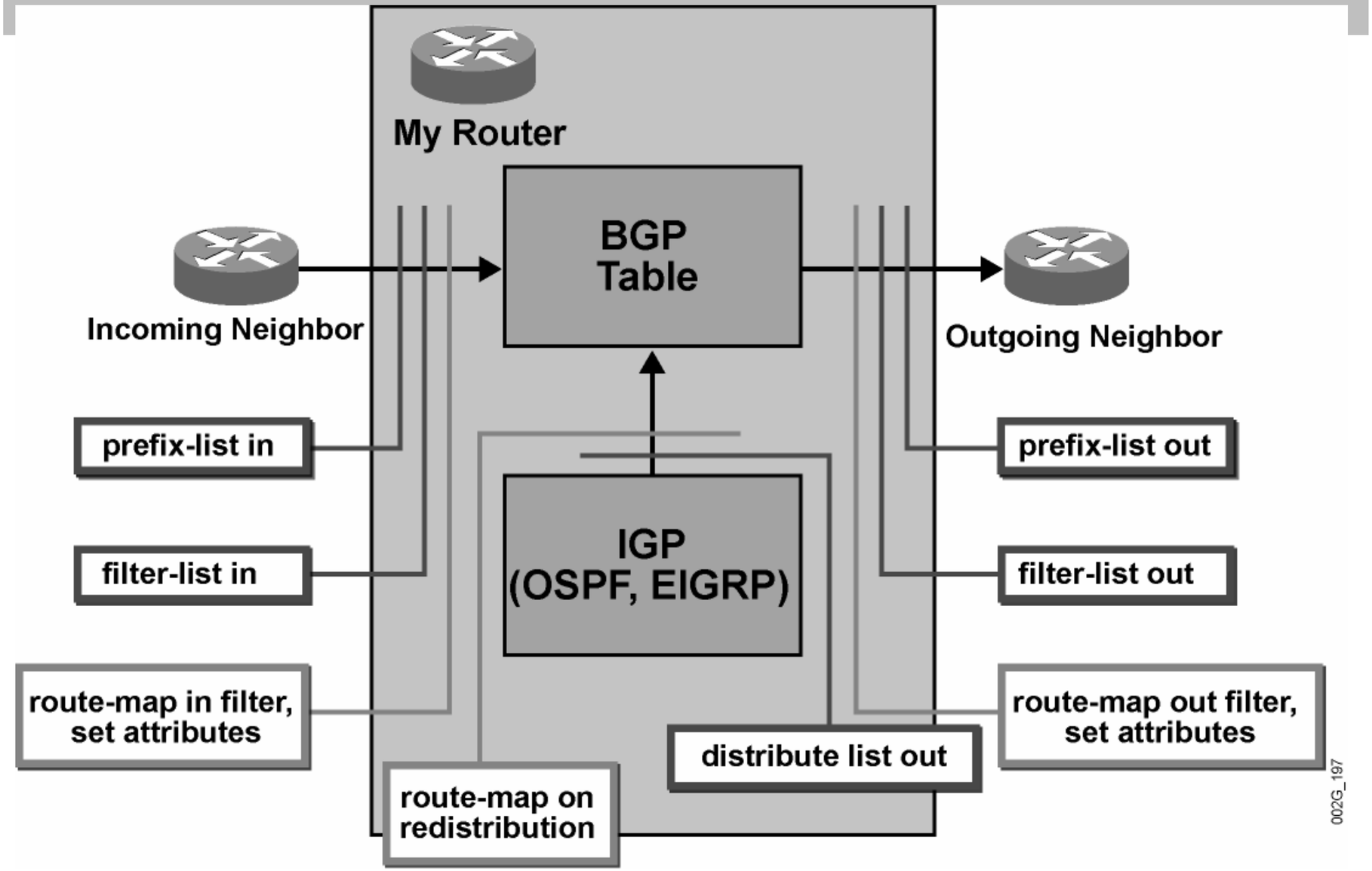


Fig 1. Simple network topology

Filtering BGP route



BGP Prefix Filtering

- Configuring BGP peering without using filters means:
  - All best paths on the local router are passed to the neighbour
  - All routes announced by the neighbour are received by the local router
  - Can have disastrous consequences
- Good practice is to ensure that each eBGP neighbour has inbound and outbound filter applied:

```
router bgp 64511
 neighbor 1.2.3.4 remote-as 64510
 neighbor 1.2.3.4 prefix-list as64510-in in
 neighbor 1.2.3.4 prefix-list as64510-out out
```

Receiving Prefixes from customer: Cisco IOS

- For Example:
  - Downstream has 100.69.0.0/20 block
  - Should only announce this to upstreams
  - Upstreams should only accept this from them
- Configuration on upstream

```
router bgp 100
 address-family ipv4
  neighbor 100.67.10.1 remote-as 101
  neighbor 100.67.10.1 prefix-list customer in
  neighbor 100.67.10.1 prefix-list default out
  neighbor 100.67.10.1 activate
!
ip prefix-list customer permit 100.69.0.0/20
!
ip prefix-list default permit 0.0.0.0/0
```

Receiving Prefixes: From Peers

- A peer is a Network Operator with whom you agree to exchange prefixes you originate into the Internet routing table
  - Prefixes you accept from a peer are only those they have indicated they will announce
  - Prefixes you announce to your peer are only those you have indicated you will announce

Receiving Prefixes: From Peers

- Agreeing what each will announce to the other:
  - Exchange of e-mail documentation as part of the peering agreement, and then ongoing updates

    OR
  - Use of the Internet Routing Registry and configuration tools such as:
    - IRRToolSet:
      https://github.com/irrtoolset/irrtoolset
    - bgpq4:
      https://github.com/bgp/bgpq4

Receiving Prefixes: From Upstream/Transit Provider

- Upstream/Transit Provider is a Network Operator who you pay to give you transit to the WHOLE Internet
- Receiving prefixes from them is not desirable unless really necessary
  - Traffic Engineering – see BGP Multihoming presentations
- Ask upstream/transit provider to either:
  - originate a default-route

    OR
  - announce one prefix you can use as default

Receiving Prefixes: From Upstream/Transit Provider

- Downstream Router Configuration

```
router bgp 100
 address-family ipv4
  network 100.66.0.0 mask 255.255.224.0
  neighbor 100.65.7.1 remote-as 101
  neighbor 100.65.7.1 prefix-list infilter in
  neighbor 100.65.7.1 prefix-list outfilter out
  neighbor 100.65.7.1 activate
!
ip prefix-list infilter permit 0.0.0.0/0
!
ip prefix-list outfilter permit 100.66.0.0/19
```

Receiving Prefixes from peer: Cisco IOS

- For Example:
  - Peer has 220.50.0.0/16, 61.237.64.0/18 and 81.250.128.0/17 address blocks
- Configuration on local router

```
router bgp 100
 address-family ipv4
  neighbor 100.67.10.1 remote-as 101
  neighbor 100.67.10.1 prefix-list my-peer in
  neighbor 100.67.10.1 prefix-list my-prefix out
  neighbor 100.67.10.1 activate
!
ip prefix-list my-peer permit 220.50.0.0/16
ip prefix-list my-peer permit 61.237.64.0/18
ip prefix-list my-peer permit 81.250.128.0/17
ip prefix-list my-peer deny 0.0.0.0/0 le 32
!
ip prefix-list my-prefix permit 100.67.16.0/20
```

Receiving Prefixes: From Upstream/Transit Provider

- If it is necessary to receive prefixes from any provider, care is required.
  - Don't accept default (unless you need it)
  - Don't accept your own prefixes
- Special use prefixes for IPv4 and IPv6:
  - http://www.rfc-editor.org/rfc/rfc6890.txt
- For IPv4:
  - Don't accept prefixes longer than /24 (?)
    - /24 was the historical class C
- For IPv6:
  - Don't accept prefixes longer than /48 (?)
    - /48 is the design minimum delegated to a site

Receiving Prefixes: From Upstream/Transit Provider

- Check Team Cymru's list of "bogons"
  - https://www.team-cymru.com/bogon-reference-http
- For IPv4 also consult:
  - https://www.rfc-editor.org/rfc/rfc6441.txt (BCP171)
- Bogon Route Server:
  - https://www.team-cymru.com/bogon-reference-bgp
  - Supplies a BGP feed (IPv4 and/or IPv6) of address blocks which should not appear in the BGP table

Receiving IPv4 Prefixes

```
router bgp 100
 network 101.10.0.0 mask 255.255.224.0
 neighbor 100.65.7.1 remote-as 101
 neighbor 100.65.7.1 prefix-list in-filter in
!
ip prefix-list in-filter deny 0.0.0.0/0              ! Default
ip prefix-list in-filter deny 0.0.0.0/8 le 32        ! RFC1122 local host
ip prefix-list in-filter deny 10.0.0.0/8 le 32       ! RFC1918
ip prefix-list in-filter deny 100.64.0.0/10 le 32    ! RFC6598 shared address
ip prefix-list in-filter deny 101.10.0.0/19 le 32    ! Local prefix
ip prefix-list in-filter deny 127.0.0.0/8 le 32      ! Loopback
ip prefix-list in-filter deny 169.254.0.0/16 le 32   ! Auto-config
ip prefix-list in-filter deny 172.16.0.0/12 le 32    ! RFC1918
ip prefix-list in-filter deny 192.0.0.0/24 le 32     ! RFC6598 IETF protocol
ip prefix-list in-filter deny 192.0.2.0/24 le 32     ! TEST1
ip prefix-list in-filter deny 192.168.0.0/16 le 32   ! RFC1918
ip prefix-list in-filter deny 198.18.0.0/15 le 32    ! Benchmarking
ip prefix-list in-filter deny 198.51.100.0/24 le 32  ! TEST2
ip prefix-list in-filter deny 203.0.113.0/24 le 32   ! TEST3
ip prefix-list in-filter deny 224.0.0.0/3 le 32      ! Multicast & Experimental
ip prefix-list in-filter deny 0.0.0.0/0 ge 25        ! Prefixes >/24
ip prefix-list in-filter permit 0.0.0.0/0 le 32
```

Receiving IPv6 Prefixes

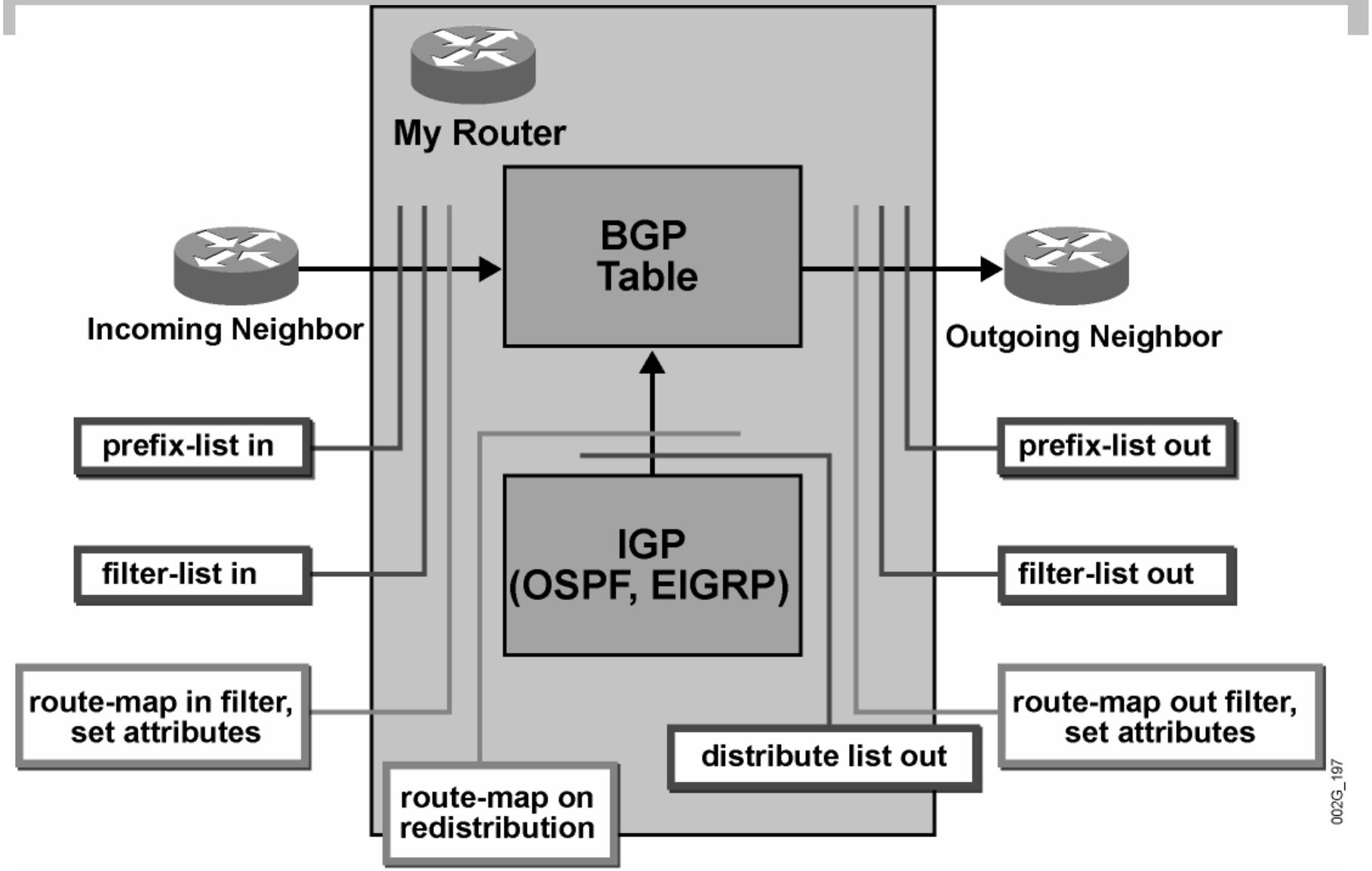
```
router bgp 100
 network 2020:3030::/32
 neighbor 2020:3030::1 remote-as 101
 neighbor 2020:3030::1 prefix-list v6in-filter in
!
ipv6 prefix-list v6in-filter permit 64:ff9b::/96          ! RFC6052 v4v6trans
ipv6 prefix-list v6in-filter deny 2001::/23 le 128        ! RFC2928 IETF prot
ipv6 prefix-list v6in-filter deny 2001:2::/48 le 128      ! Benchmarking
ipv6 prefix-list v6in-filter deny 2001:10::/28 le 128     ! ORCHID
ipv6 prefix-list v6in-filter deny 2001:db8::/32 le 128    ! Documentation
ipv6 prefix-list v6in-filter deny 2002::/16 le 128        ! Deny all 6to4
ipv6 prefix-list v6in-filter deny 2020:3030::/32 le 128   ! Local Prefix
ipv6 prefix-list v6in-filter deny 3ffe::/16 le 128        ! Formerly 6bone
ipv6 prefix-list v6in-filter permit 2000::/3 le 48        ! Global Unicast
ipv6 prefix-list v6in-filter deny ::/0 le 128
```

Note: These filters block Teredo (serious security risk) and 6to4 (deprecated by RFC7526)

Receiving Prefixes

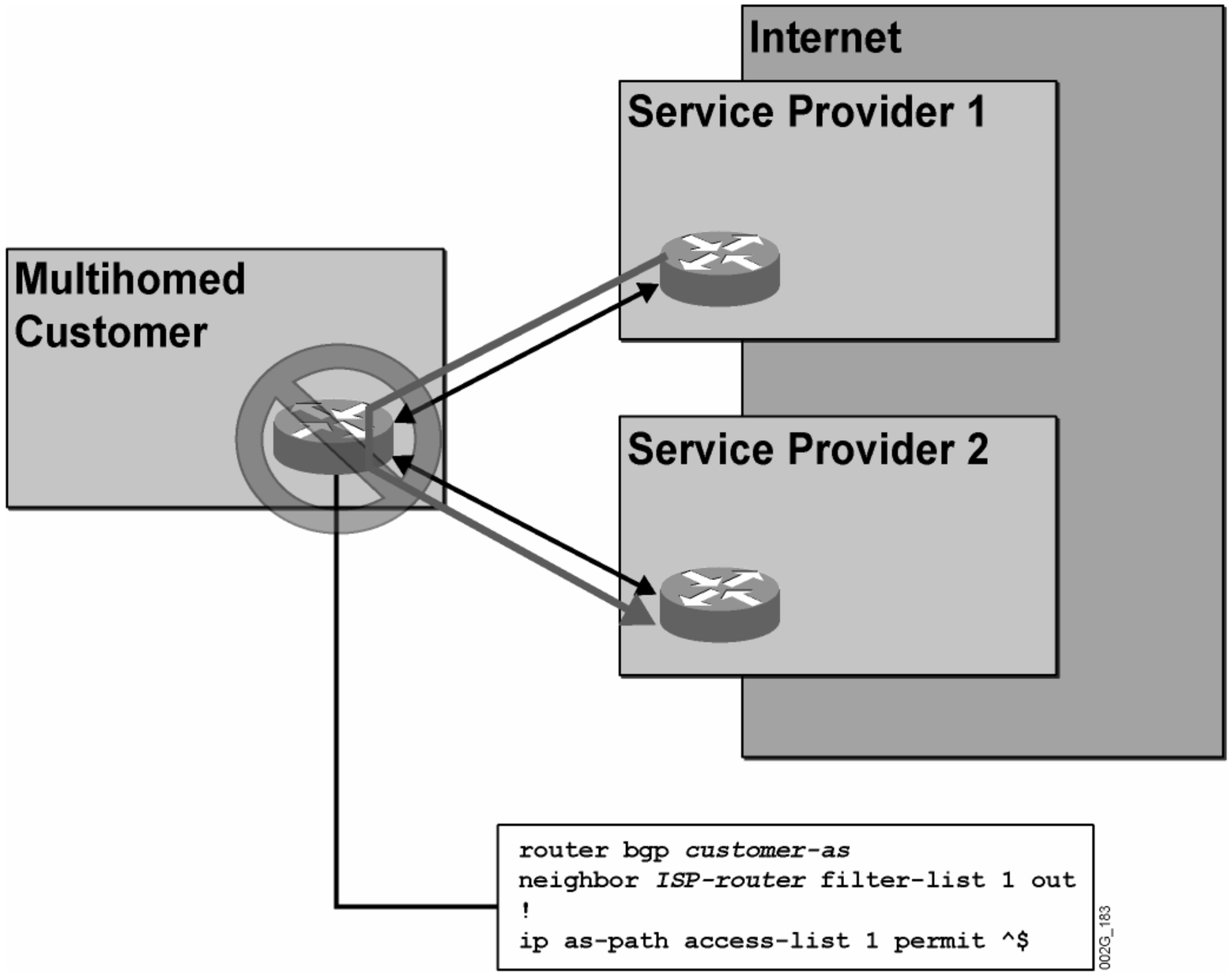
- Paying attention to prefixes received from customers, peers and transit providers assists with:
  - The integrity of the local network
  - The integrity of the Internet
- Responsibility of all Network Operators to be good Internet citizens

Filtering BGP route



AS-Path Filters

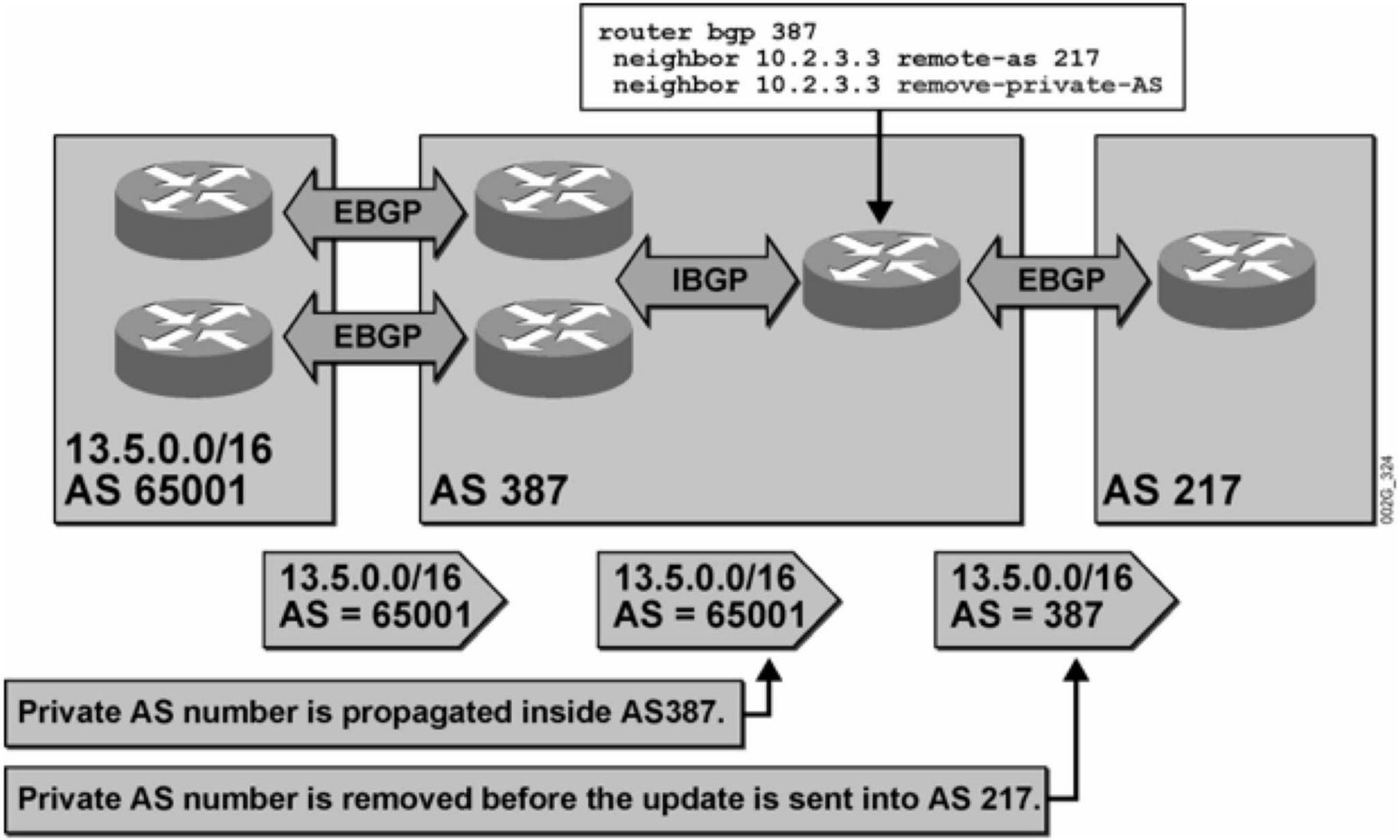
- AS-Path Filter Usages
  - Announce only local routes to the ISP—AS path needs to be empty
  - Select routes based on a specific AS number in the AS path
  - Accept routes for specific AS only from some BGP neighbours
- AS-Path Filter use regular expression


Internet
Service Provider 1
Multihomed
Customer
Service Provider 2
router bgp customer-as
neighbor ISP-router filter-list 1 out
!
ip as-path access-list 1 permit ^$
002G_183

Sample Regular Expressions

_100_	Going through AS 100
^100$	Directly connected to AS 100
_100$	Originated in AS 100
^100_.	Networks behind AS 100
^ [0-9]+$	AS paths one AS long
^([0-9]+)(_\1)*$	Prepending performed in neighbouring originating AS
^$	Networks originated in local AS
.*	Matches everything

Removing Private AS Numbers



Receiving Prefixes: Bogon ASNs?

- What about prefixes originated by bogon AS numbers?
  - Public ranges are 1-64495 (excluding 23456) and 131072-458751
    - IANA is distributing AS blocks to the RIRs from the latter range
  - All other ASNs are either for documentation, or for private use, or are unassigned
    - And any prefixes originating from those need to be dropped
    - Configuration error? Malicious intent?
- What would the AS_PATH filter look like?
  - Challenging with regular expression (as per IOS)
  - Easier with AS ranges (as per Bird or JunOS)

Filtering bogon ASNs – BIRD

- Here is a function showing how to filter bogon ASNs, as described previously:

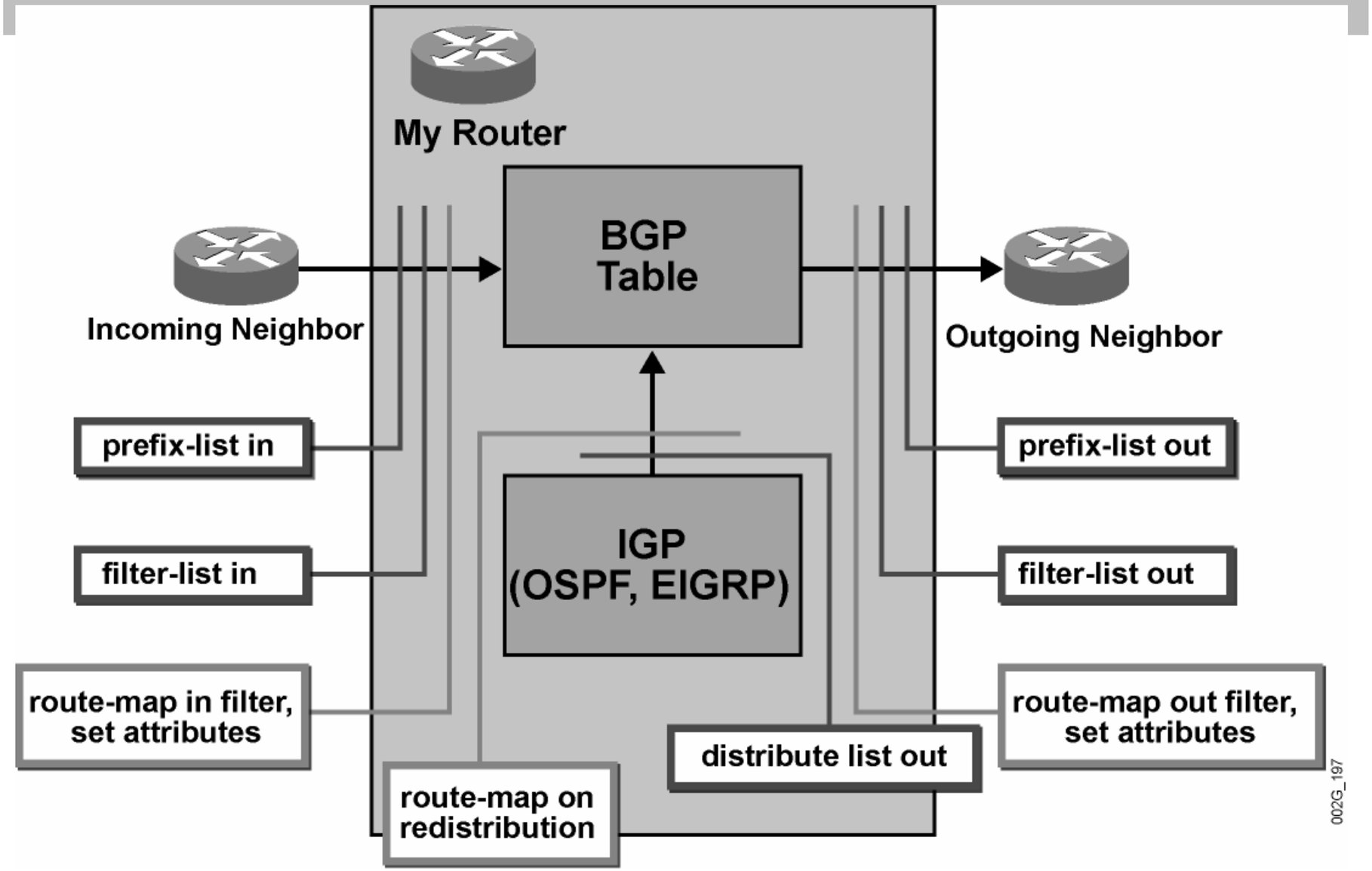
```
function as_path_contains_bogons()
int set invalid_asns;
{
    invalid_asns = [
        0,                      # Reserved
        23456,                  # Transition AS
        64496..64511,           # Documentation ASNs
        64512..65534,           # Private ASNs
        65535,                  # Reserved
        65536..65551,           # Documentation ASNs
        65552..131071,          # Reserved
        458752..4199999999,     # IANA Reserved
        4200000000..4294967294, # Private ASNs
        4294967295              # Reserved
    ];
    return bgp_path ~ invalid_asns;
}
```

Filtering bogon ASNs – FRR

- Here is an AS-PATH regexp showing how to filter bogon ASNs:

```
bgp as-path access-list Bogon_ASNs deny _0_
bgp as-path access-list Bogon_ASNs deny _23456_
bgp as-path access-list Bogon_ASNs deny _6449[6-9]_
bgp as-path access-list Bogon_ASNs deny _64[5-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _6[5-9][0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _[7-9][0-9][0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _1[0-2][0-9][0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _130[0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _1310[0-6][0-9]_
bgp as-path access-list Bogon_ASNs deny _13107[0-1]_
bgp as-path access-list Bogon_ASNs deny _45875[2-9]_
bgp as-path access-list Bogon_ASNs deny _4587[6-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _458[8-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _459[0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _4[6-9][0-9][0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _[5-9][0-9][0-9][0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _[0-9][0-9][0-9][0-9][0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _[0-9][0-9][0-9][0-9][0-9][0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _[0-9][0-9][0-9][0-9][0-9][0-9][0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs deny _[0-9][0-9][0-9][0-9][0-9][0-9][0-9][0-9][0-9][0-9]_
bgp as-path access-list Bogon_ASNs permit .*
```

Filtering BGP route



BGP and Route-maps

- Route-maps can set on:
  - Origin
  - BGP next-hop
  - Weight
  - BGP community
  - Local preference
  - MED

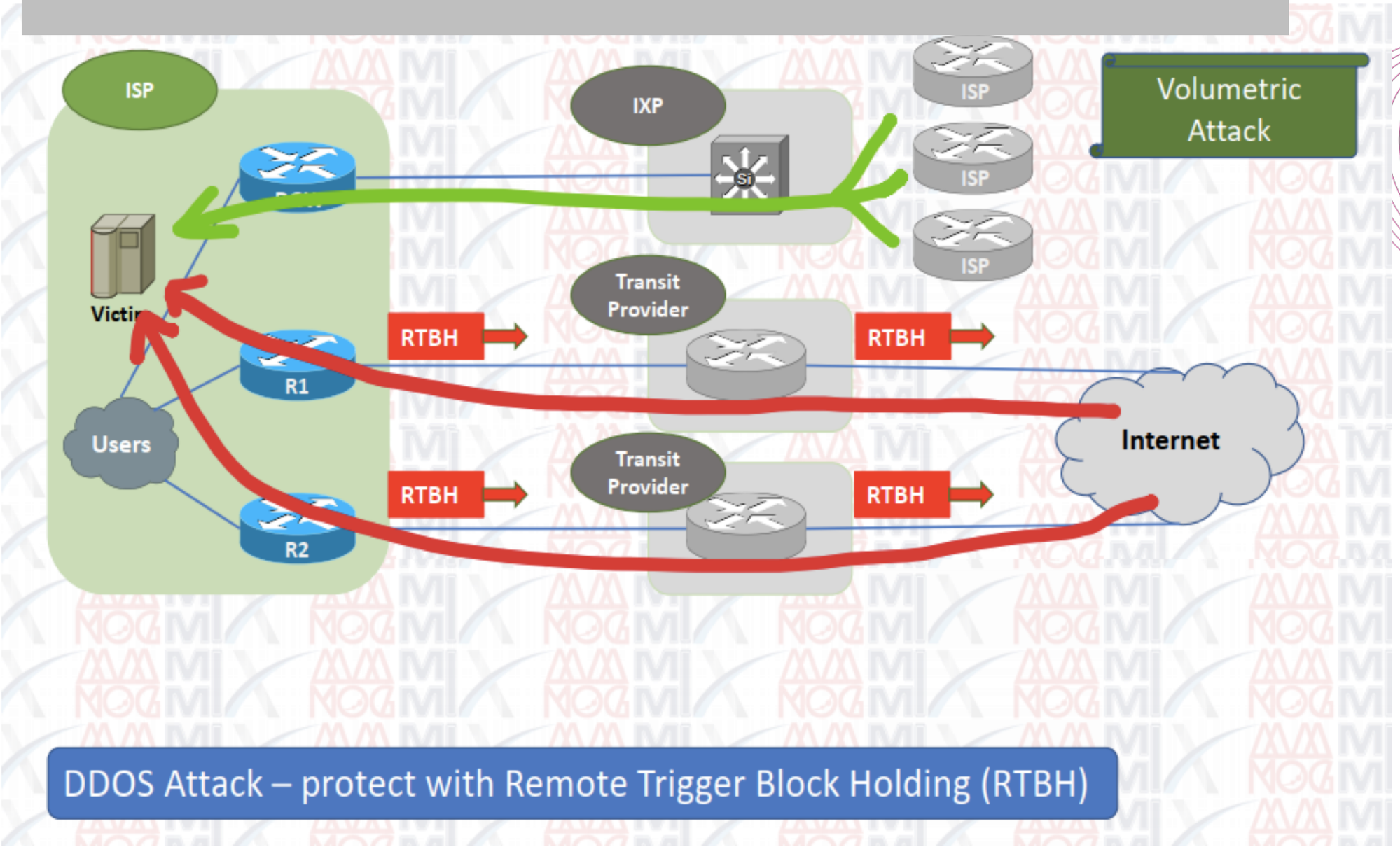
BGP Communities at MMIX

0:9654	Block Announcement of prefixes to all peers
0:(peer-as)	Block Announcement of prefixes to certain peer
9654:(peer-as)	Advertise to certain peer
9654:9654	Advertise of prefixes to all peers
9654:11344	Advertise to GGC
40027:40000	Advertise to Netflix
9654:20940	Advertise to Akamai
9654:54994	Advertise to Wangsu

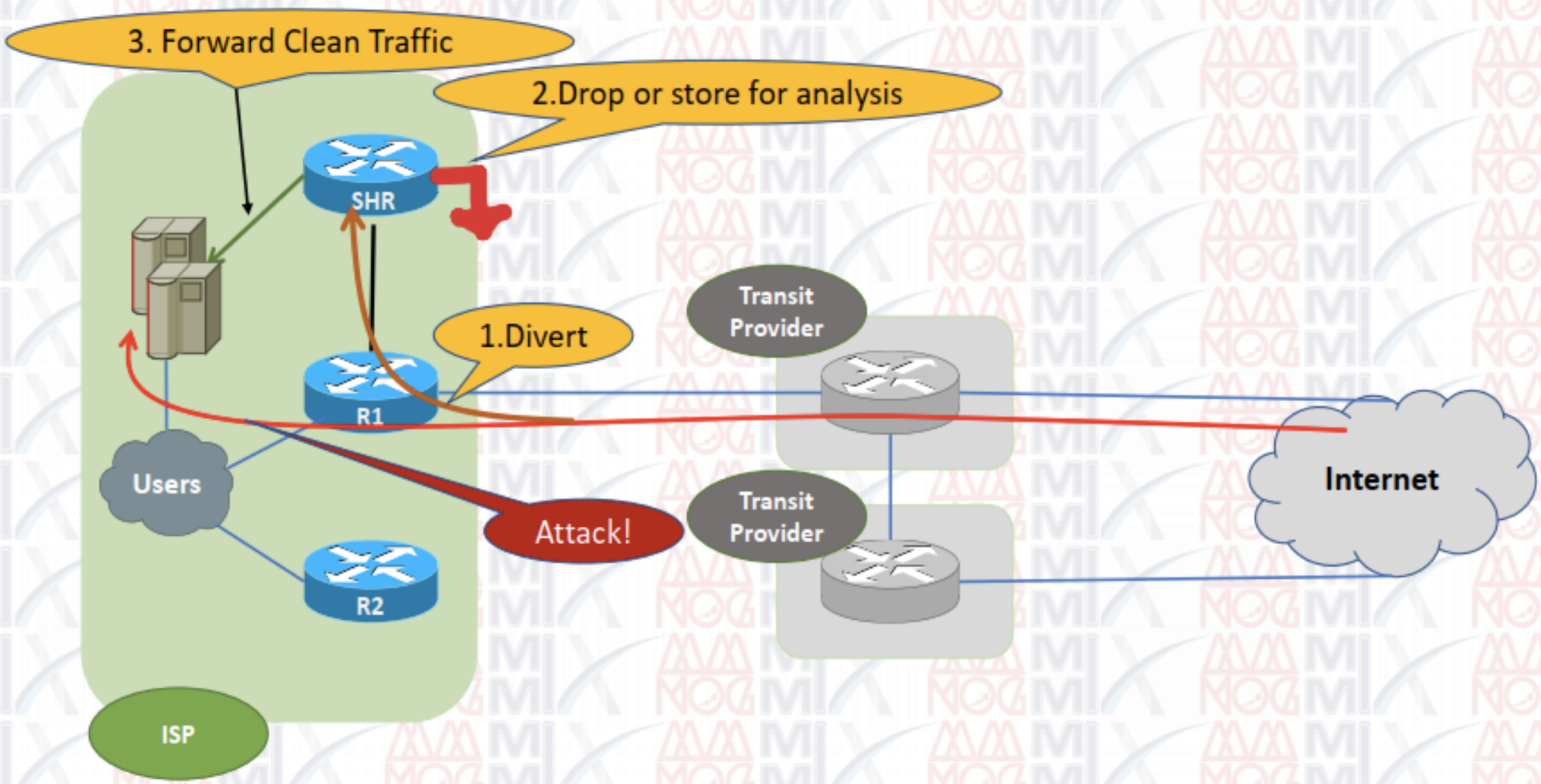
RTBH – How it works

- Network Operator deploys:
  - RTBH support across their entire backbone
    - Simply a null route for a specific next-hop address
    - (Router Null interfaces simply discard packets sent to them – negligible overhead in modern hardware)
  - A trigger router (usually in the NOC)
    - Talks iBGP with the rest of the backbone (typically as a client to route-reflectors in the core)
    - Used to trigger a blackhole route activity for any address under attack, as requested by a customer

Remotely Triggered Black Hole (RTBH)



Remotely Triggered Black Hole (RTBH)



DOS Attack – Use Sinkhole for clean pipe

Internet Routing Registry

- Many major transit providers and several content providers pay attention to what is contained in the Internet Routing Registry
  - There are many IRRs operating, the most commonly used being those hosted by the Regional Internet Registries, RADB, and some transit providers
- Best practice for any AS holder is to document their routing policy in the IRR
  - A route-object is the absolute minimum requirement

Internet Routing Registry

- IRR objects can be created via the database web-interfaces or submitted via email
- Policy language used to be known as RPSL
- Problems:
  - IRR contains a lot of outdated information
  - Network operators not following best practices
- Some network operators now using RPKI and ROAs to securely indicate the origin AS of their routes
  - Takes priority over IRR entries
  - RPKI and ROAs covered in other presentations

Internet Routing Registry

- Which IRR database to use?
  - Members of a Regional Internet Registry are recommended to use their RIR's Internet Routing Registry instance
    - Usually managed via the RIR's member portal giving easy access for creation and update of objects
    - Provided as part of the RIR's services to its members
  - Operators who do not belong to any RIR generally use:
    - Their upstream transit provider's Routing Registry (if provided)
    - The RADB
      - https://www.radb.net
      - Note: Placing objects in the RADB requires an annual subscription fee

Route Object: Purpose

- Documents which Autonomous System number is originating the route listed
- Required by many major transit providers
  - They build their customer and peer filter based on the route-objects listed in the IRR
  - Referring to at least the 5 RIR routing registries and the RADB
  - Some operators run their own Routing Registry
    - May require their customers to place a Route Object there (if not using the 5 RIR or RADB versions of the IRR)

Apnic => RADB

Create record at Apnic Portal.

Object Type

V4 prefix - ‘route’

V6 prefix – ‘route6’

```
route:          103.103.194.0/24
descr:          MMIX Net 1
origin:         AS137955
mnt-by:         MAIBT-MM-MMIX
last-modified:  2022-12-21T22:32:27Z
source:         APNIC
```

RADB

RADB

Anyone can update records.
Multiple wrong records.

Recommend

Let prefix owner update their owned records

```
route:      103.103.194.0/24
descr:      Proxy-registered route obje
origin:     AS9654
notify:     matthew.chan@hgc.com.hk
mnt-by:     MAINT-HGC-INTL
changed:    matthew.chan@hgc.com.hk 20220307
source:     RADB
```

MMIX

```
route:      103.103.194.0/24
descr:      CMI  (Customer Route)
origin:     AS132167
mnt-by:     MAINT-AS58453
changed:    qas_support@cmi.chinamobile.com 20200923
source:     RADB
```

Ooredoo

```
route:      103.103.194.0/24
descr:      CMI  (Customer Route)
origin:     AS9304
mnt-by:     MAINT-AS58453
changed:    qas_support@cmi.chinamobile.com 20200916
source:     RADB
```

HGC

```
route:          103.103.194.0/24
descr:          MMIX Net 1
origin:         AS137955
mnt-by:         MAIBT-MM-MMIX
last-modified:  2022-12-21T22:32:27Z
source:         APNIC
```

MMIX

AS Object: Purpose

- Documents peering policy with other Autonomous Systems
  - Lists network information
  - Lists contact information
  - Lists routes announced to neighbouring autonomous systems
  - Lists routes accepted from neighbouring autonomous systems
- Some operators pay close attention to what is contained in the AS Object
  - Some configure their border router BGP policy based on what is listed in the AS Object

AS Object: Example

```
aut-num:        AS17660
as-name:        DRUKNET-AS
descr:          DrukNet ISP, Bhutan Telecom, Thimphu
country:        BT
org:            ORG-BTL2-AP
import:         from AS6461    action pref=100;     accept ANY
export:         to AS6461      announce AS-DRUKNET-TRANSIT
import:         from AS2914    action pref=150;     accept ANY
export:         to AS2914      announce AS-DRUKNET-TRANSIT
<snip>
import:         from AS135666 action pref=250;    accept AS135666
export:         to AS135666    announce {0.0.0.0/0} AS-DRUKNET-TRANSIT
admin-c:        DNO1-AP
tech-c:         DNO1-AP
notify:         netops@bt.bt
mnt-irt:        IRT-BTTELECOM-BT
mnt-by:         APNIC-HM
mnt-lower:      MAINT-BT-DRUKNET
mnt-routes:     MAINT-BT-DRUKNET
last-modified:  2019-06-09T22:40:10Z
source:         APNIC
```

Examples of inbound and outbound policies – RPSL

AS-Set: Purpose

- The AS-Set is used by network operators to group AS numbers they provide transit for in an easier to manage form
  - Convenient for more complicated policy declarations
  - Used mostly by network operators who build their EBGP filters from their IRR entries
  - Commonly used at Internet Exchange Points to handle large numbers of peers

AS-Set: Example

```
as-set:          AS-DRUKNET-TRANSIT
descr:           DrukNet transit networks
members:         AS17660
members:         AS38004
members:         AS132232
members:         AS134715
members:         AS135666
members:         AS137925
members:         AS59219
members:         AS18024
members:         AS18025
members:         AS137994
admin-c:         DNO1-AP
tech-c:          DNO1-AP
notify:          netops@bt.bt
mnt-by:          MAINT-BT-DRUKNET
last-modified:   2019-01-15T08:51:21Z
source:          APNIC
```

Lists all the autonomous systems within the AS-DRUKNET-TRANSIT group

PeeringDB

Search here for a network, IX, or facility.

Advanced Search | Legacy Search

Register or

English (English)

Myanmar Internet Exchange Association Inc.

Also Known As	
Long Name	
Website	http://www.mm-ix.net
Address 1	Building 18, 2nd Floor, MICT Park,
Address 2	Hlaing University Campus, Hlaing Township
Floor	
Suite	
Location	Yangon, , 11051
Country Code	MM
Geocode	*Geocode data for this entity could not be obtained at this point. This is done automatically upon address field changes.*
Last Updated	2022-12-23T18:22:23Z
Notes	
Logo	

Networks

Filter

Name	ASN
MMIX IPT	137955
MMIX Mandalay Route Servers	9333
MMIX Yangon Route Servers	9654

Exchanges

Filter

Name	Country City
MMIX Mandalay	Mandalay
MMIX Yangon	Yangon

MMIX Mandalay

Peers	Connections	Open Peers	Total Speed	% with IPv6
12	12	12	382G	58

EXPORT

Organization	Myanmar Internet Exchange Association Inc.
Also Known As	MMIX
Long Name	Myanmar Internet Exchange
City	Mandalay
Country	MM
Continental Region	Asia Pacific
Media Type	Ethernet
Service Level	Not Disclosed
Terms	Not Disclosed
Last Updated	2023-05-30T13:23:01Z
Notes	

Contact Information

Company Website	https://www.mmix.net.mm
Traffic Stats Website	https://www.mmix.net.mm
Technical Email	noc@mm-ix.net
Technical Phone	+959881312340
Policy Email	admin@mm-ix.net
Policy Phone	
Sales Email	
Sales Phone	
Health Check	

Peers at this Exchange Point

Filter

Peer Name AŻ IPv4	ASN IPv6	Speed	Policy
103.116.193.25	2001:df3:1300:2::940:25		
ETPCL-AS-AP 103.116.193.45	134714	1G	Open
Horizon Sources Company Limited 103.116.193.18	149660	10G	Open
Kaopu Cloud HK 103.116.193.39	138915 2001:df3:1300:2::8915:39	10G	Open
meteversecloud 103.116.193.7	54994 2001:df3:1300:2::4994:7	20G	Open
MMIX IPT 103.116.193.3	137955 2001:df3:1300:2::7955:3	100G	Open
Myanmar Country Co. 103.116.193.28	134840	100G	Open
PCH AS3856 103.116.193.34	3856 2001:df3:1300:2::3856:34	10G	Open
PCH AS42 103.116.193.33	42 2001:df3:1300:2::42:33	10G	Open
RIPE NCC K-Root Operations 103.116.193.8	25152 2001:df3:1300:2::5152:8	1G	Open
Telcospeed Communication Co.,Ltd 103.116.193.37	139003	10G	Open
Zoom Plus 103.116.193.42	133433	10G	Open

Route Origin Authorisation (ROA)

- A digital object that contains a list of address prefixes and one AS number
- It is an authority created by a prefix holder to authorise an AS Number to originate one or more specific route advertisements
- Publish a ROA using your RIR member portal
  - Consult your RIR for how to use their member portal to publish your ROAs

Route Origin Authorisation

- A typical ROA would look like this:

Prefix	10.10.0.0/16
Max-Length	/18
Origin-AS	AS65534

- There can be more than one ROA per address block
  - Allows the operator to originate prefixes from more than one AS
  - Caters for changes in routing policy or prefix origin

- NB: Only create ROAs for the aggregate and the exact subnets expected in the routing table!! (See RFC9319)

Route Origin Validation

- Route Origin Validation means checking if the prefix received has a valid ROA
  - Valid ROA means that the prefix (and subnet) is being originated from the correct origin AS
  - See the "BGP Origin Validation" presentation for more in-depth content
- Implementing ROV means checking the validation database with what is learned from BGP peers:
  - Valid – allow; Invalid – drop; NotFound – allow (at lower preference?)
- Problem: how is this implemented in routers day?

Route Origin Validation

- The ideal would be to write directly to the active BGP table
- Some implementations use existing EBGP policy handling routines
  - ADJ-RIB-IN: table of all prefixes received prior to policy being applied
  - Route Refresh (RFC2918)

- Routers which maintain the ADJ-RIB-IN:
  - Apply the ROV policy to the stored received BGP table
  - Updates are applied "automatically" to the BGP table and therefore the FIB
  - No impact on any BGP peers (Route Refresh not needed)

Route Origin Validation

- Routers which do NOT maintain the ADJ-RIB-IN:
  - Apply the ROV policy by sending a Route Refresh to peers
  - When there are a large number of ROAs (November 2021 sees over 290k), and frequent changes or updates of ROAs:
    - Routers are sending frequent Route Refresh requests to peers (typically every few minutes)
    - Peers are being "bombarded" by Route Refresh requests: significant resource burden when they send the full or a large portion of the BGP table
    - Severe control plane CPU impact on the peer router (effectively a Denial of Service on the peer router)
  - As more and more ROAs are created and altered globally, this problem becomes significantly more serious!

Thank You!!!

https://sg.linkedin.com/in/sainyanlynnswe

https://www.facebook.com/AstraeaLwin