

RPKI and IRR filtering at AMS-IX route servers

Stavros Konstantaras NOC Engineer

GRNOG7 2018



Personal introduction

- BSc in Information Technology (ATEL Thess.)
- MSc in System & Network Engineering from UvA
- 5 years of professional experience
 - Software engineer @ NEC Heidelberg, DE
 - Systems engineer @ UvA, NL
 - Network automation engineer
 @ NLnet Labs, NL
 - NOC engineer @ AMS-IX, NL



AMS-IX info

- A world leading independent Internet Exchange platform
 - Based in Amsterdam (14 PoPs)
 - 6 worldwide Exchange Points managed remotely (US, CW, HK, IN)
 - 5.6 Tbps peak traffic, 823 ASNs
 - Remote peering drives our growth
 - OTEGLOBE, CYTA, GEANT



Agenda

- Route Servers & Filtering – what
 - why
- AMS-IX implementation

 Architecture & Features
 How?
- Real-life examples/problems



Route Servers in IXPs

- Reduces the number of BGP connections per member/customer
- Manage only your most important peers, let the route server do the rest
- Send and receive routes from day one
- Use it as a backup

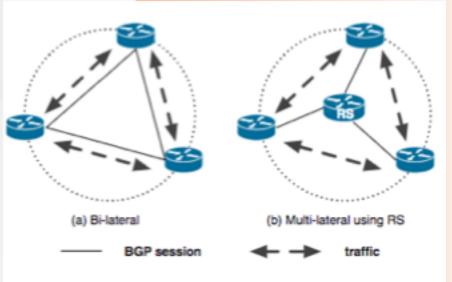


Figure 1: IXP peering options.

Figure retrieved from IMC238 (Richter et al): "Peering at Peerings: On the Role of IXP Route Servers", 2014



AMS-IX Route Servers

- 2 BIRD instances in high spec servers
- 775 IPv4 Peers & 636 IPv6 peers*
 - Prefixes received: 278.376 IPv4 II 46.233 IPv6
 - Prefixes Sent: **197.107** IPv4 II **31358** IPv6
 - Average Prefixes per peer: 386 IPv4 II 79 IPv6
- Neutral prefix handling
 - Local_pref = 100

*Updated 5/7/2018



Why filtering in IXPs?

- Prefixes received by IXPs have (usually)
 - higher local_pref and/or
 - shorter as_path
- RS are a vital part inside IXP's ecosystem
 - Small/medium customers peer only with RS to receive NLRI
 - Medium/Large customers use it as a secondary/backup source of NLRI
- Blocking Invalids in IXPs (IRRdb/RPKI)
 - keeps legitimate traffic inside the network
 - Reduces the impact of hijacking/miss-configuration to all members/customers that peer with RS.



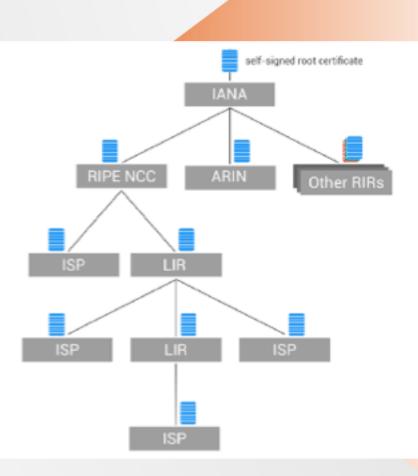
Internet Resource db

- Nice concept but...
 - Inconsistent information
 - No strings attached with routers and their BGP announcements
 - Fat fingers
 - BGP hijacking
 - RPSL and related tools

amsterdam Internet exchange

RPKI 1/2

- Provide a reliable infrastructure that can be used to secure BGP
 - Via Origin Validation
- Mirrors the existing resource allocation infrastructure
- Validate ownership of Internet resources
 - Resource Certificates (X.509)



RPKI 2/2



amsterdam Internet exchange

- Origin validation
 Valid
 - Invalid
 - Unknown
- Policies applied



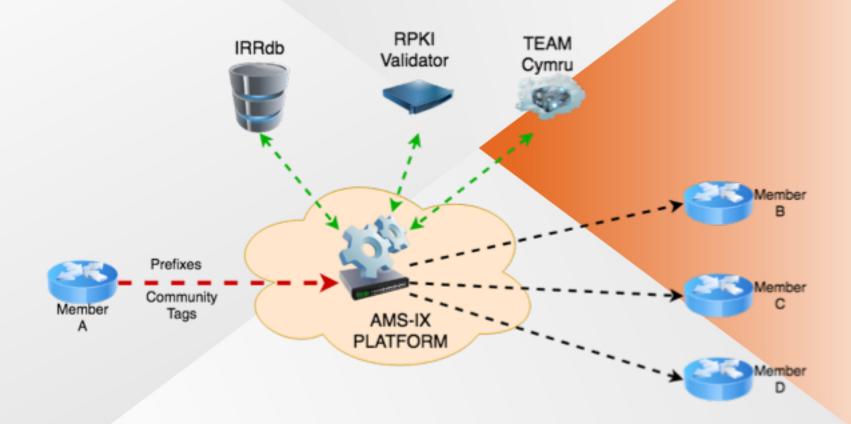
Features of AMS-IX RS

- Receive Prefixes / Propagate best paths
- Ensure peering rules are satisfied
- Perform IRR and RPKI based filtering
 The 3+1 filtering modes
- Perform community-based filtering
- Expose info to LG and notification system*





AMS-IX RS architecture





RS input explained

IRRdb

- Retrieve and Parse member's policy (import/export attributes)
- Retrieve and resolve
 AS-SETs

RPKI

 ROAs (via the RIPE validator which is connected to both RS)

Team Cymru

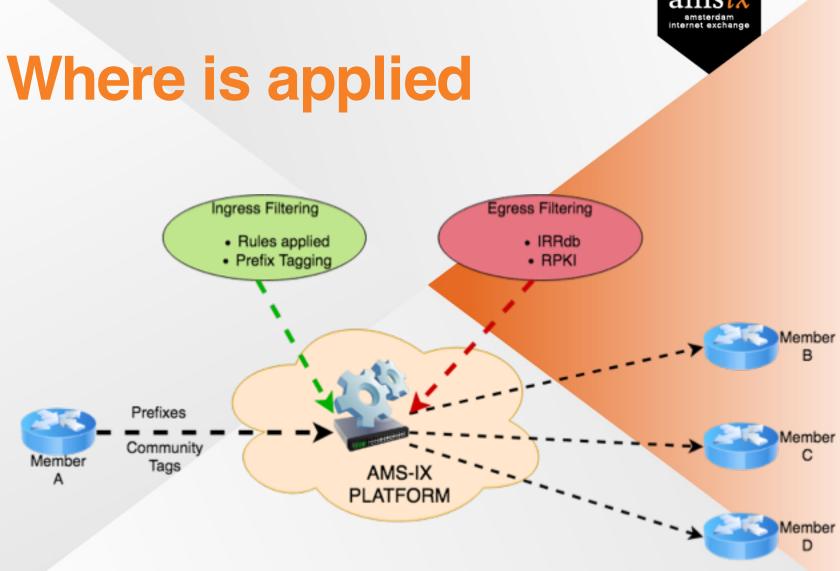
- Bogons
- Martians



Prefix filtering in AMS-IX

- Basic (ingress)
 The AMS-IX peering rules
- Extended (egress)
 The 3+1 peering modes







Peering rules (ingress)

- Not accepted prefixes:
 - Bogons & Martians
 - AMS-IX prefixes

- ACCESS DENIER
- Prefixes with AS path length > 64
- The first AS in AS path is **not** the customer one
- BGP next hop not belonging to the router advertising the prefix
- Discard AS Paths containing Private ASNs*



The 3+1 filtering modes (egress)

- "Filtering based on both IRRdb and RPKI data" (default)
- "Filtering based on IRRdb data"
- "Filtering based on RPKI data"
- o "No Filtering, Just tagging"



IRRdb Filtering 1/2

- Parse and resolve import/export records towards AS6777
 - Import-via/export-via are supported
- RS config is generated automatically based on IRRdb parser scripts
 - Info gathered from all major IRR DBs
 - We detect policy changes every hour



IRRdb Filtering 2/2

- What you receive is what you define
 You define your policy -> you instruct the RS
- Keep IRR objects up-to-date



RPKI Filtering

BGP Preview

This page provides a preview of the likely RPKI validity states your routers will associate with BOP announcements. This preview is based on:

- The RIPE NCC Route Collector information that was last updated 3 hours and 30 minutes ago.
- · BGP announcements that are seen by 5 or more peers.
- The validation rules defined in RFC 6483.
- . The validated ROAs found by this RPKI Validator after applying your filters and additional whitelist entries.

Please note that the BOP announcements your routers see may differ from the ones listed here.

Show 10	- entries										Search			
ASN							Prefix		0	Validity				0
1							41.78.36	0.94		2010/001				
1							41.78.37	0/24		-				
1							45.227.80	0/22		-				
1							91,200,82	0/22		240004				
1							91,210,36	1074		201000				
1							91,210,37	0.0/24		-				
1							91,210,38	10/24						
1							94.31.44	0/24		NULLE-ION				
1							154.86.10	8.0/22		210.000				
1							168.181.3	6.0.23		-				
Fest	Previous	1	2	3	4	5	Net	Let				Showing 1 to	10 of 788,34	1 ertiles

- BGP announcements are validated with RIPE's RPKI validator
- The prefixes that are being blocked are the ones with ROA status "INVALID"



Just tagging (per request)

- No filtering is applied to announced prefixes
 - But we still mark the received prefixes with the corresponding community tags:
 - ROA status: VALID (6777:65012)
 - ROA status: INVALID (6777:65022)
 - ROA status: UNKNOWN (6777:65023)
 - Present in AS's announced AS/AS-SET (6777:65011)
 - Not present in AS's announced AS/AS-SET (6777:65021)



Selectable peering modes

Change peering

Filtering based on IRRdb data Filtering based only on RPKI data Filtering based on both IRRdb and RPKI data

No prefix filtering, just tagging [NOT RECOMMENDED]

Change peering

GRNOG7 2018



BGP communities

- Manipulate prefix announcement via BGP community attributes:
 - Do not announce a prefix to a certain peer (0:peer-as)
 - Announce a prefix to a certain peer (6777:peer-as)
 - Do not announce a prefix to any peer (0:6777)
 - Announce a prefix to all peers (6777:6777)



Dynamic per-AS Prefix Limits

- Intended to prevent route leaks
- Dynamic limit is a necessity due to Tier 1 networks
 - Use IRRdb prefixes to calculate initial limit
 - For customers sending few prefixes limit=100
 - Maximum = 20.000



Tools used in implementation

External tools

- whois (to read member policy)
- bgpq3 (for resolving AS-SETs)
- RIPE validator (to validate announcements)
- Lots of internal tools
 - rs_configurator.pl
 - rs_prefixes_api



The road to MANRS

- Make "IRRdb+RPKI" filtering the default option
- Make "Just tagging" inaccessible
- Moved customers from "Just tagging" to the new default option
- Update relevant content in our website
- Inform our members and promote Routing Security



Upcoming features

- Policy Explorer
- E-mail notification system



Policy explorer

 Available at my.ams-ix.net (soon for users)

Route Server	filtering	and	policy	explorer
--------------	-----------	-----	--------	----------

lana and	mail share	Sec.	-	10.00	-	-	14 4 50	ы.
moon	policies	TOP 1	80	249	2028	-250	nna	51
		-						

Export policies for 80.249.208.50 (1103)

Detected export policy:

to AS6777 action community .= { 6777:6777 }; announce AS-SURFNET (OK)

Peering at rs1.ams-ix.net? 🖌

Announced (outgoing) Prefix	RCA valid?	IRRdb object present?
129.125.0.0/16	unknown	2 C
130.112.0.076	unknown	2 C
130.115.0.0/16	1 A A A A A A A A A A A A A A A A A A A	1
130.161.0.0/16	1	2 C
130.37.0.0/16	unknown	2 C
130.89.0.0/16	unknown	1
131.155.0.0/16	1	1
131.174.0.076	unknown	1
131.176.1.0/24	unknown	1
131.176.103.0/24	unknown	1
131.176.105.0/24	unknown	1
131.176.106.0/23	unknown	1
131.176.108.0/24	unknown	1
131.176.123.0/24	unknown	1
131.176.124.0/24	unknown	×
131.176.126.0/24	unknown	1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A



E-mail notifications

- Notify customers about:
 - Dropped prefixes in RS
 - IRRdb/RPKI invalids
 - Exceptions can be made upon request (whitelist prefixes)
- Can work as a "reminder" to keep IRR objects up-to-date



Real life example/problems

- Member A (old config)
- Member B (prefix hijack)



Member A example

- A big outage due to a BGP announcement to AMS-IX peering LAN (March 2011):
 - Containing the AMS-IX prefix (195.69.144.0/22)
 - ASN was not "6777"
 - The subnet mask was more specific



Member B example

- Classic prefix hijacking
 - Advertising 80.249.208.0/22 instead of /21
 - Announced by: ASXXXX
 - Upstream AS: ASYYYY
 - ASpath: YYYY XXXX
 - RPKI detected it successfully
 - "RPKI Status: ROA validation failed: Invalid Origin ASN, expected 1200"



Questions?

stavros.konstantaras@ams-ix.net