RPKI and IRR filtering at AMS-IX route servers

Stavros Konstantaras
NOC Engineer

GRNOG7 2018

Personal introduction

- BSc in Information Technology (ATEI 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

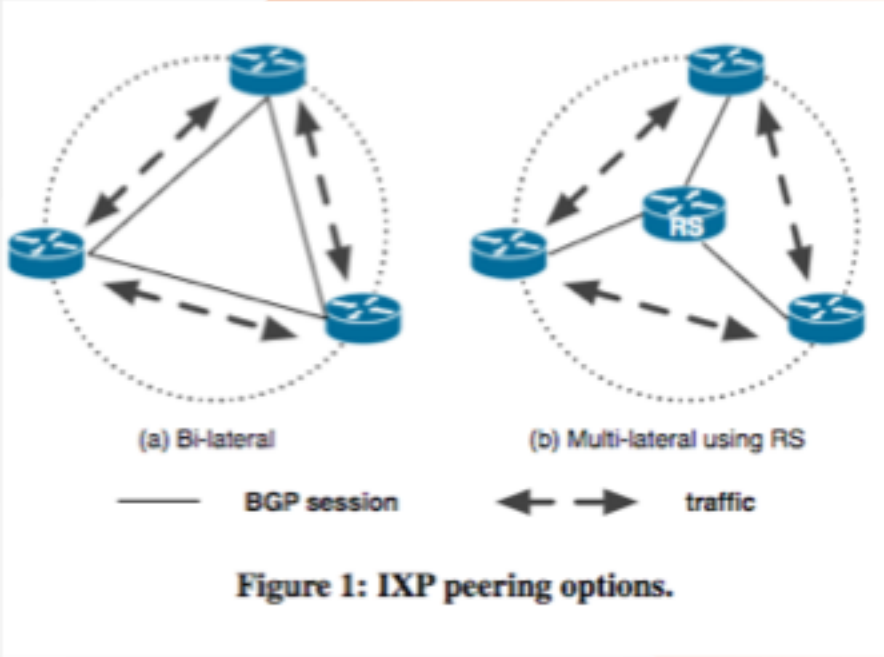


Figure 1: IXP peering options.

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?

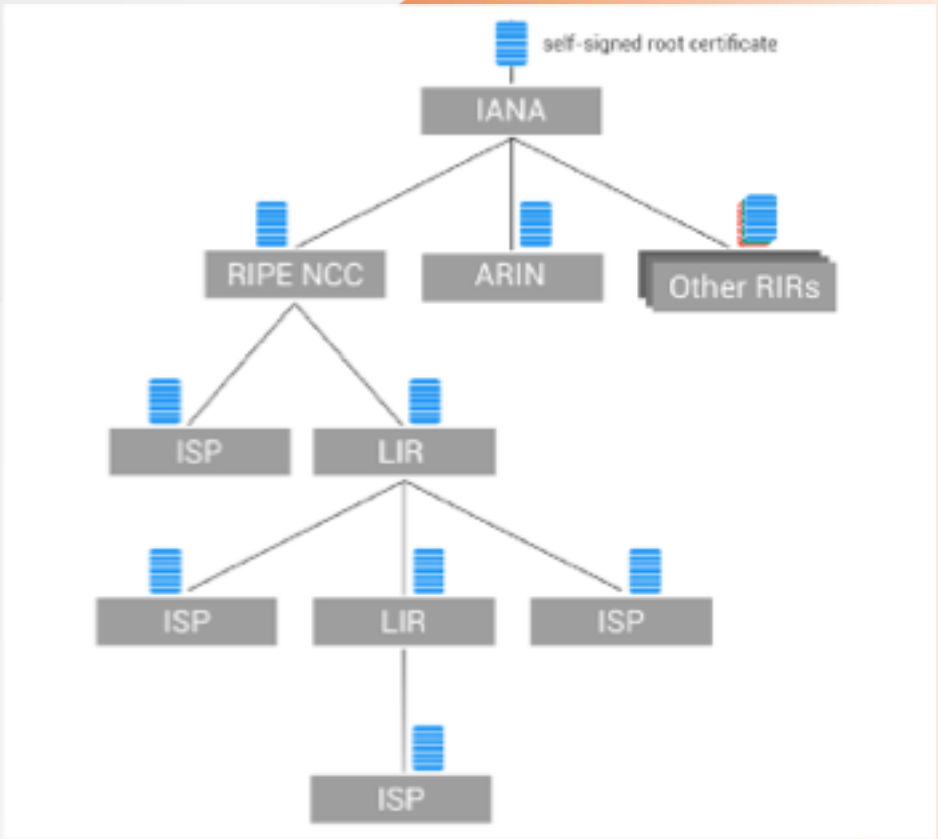
- Prefixes received by IXPs have (usually)
  - higher local_pref and/or
  - shorter as_path
- RS are a vital part inside IXP's ecosystem
  - Small/medium customers peer only with RS to receive NLRI
  - Medium/Large customers use it as a secondary/backup source of NLRI
- Blocking Invalids in IXPs (IRRdb/RPKI)
  - 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

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

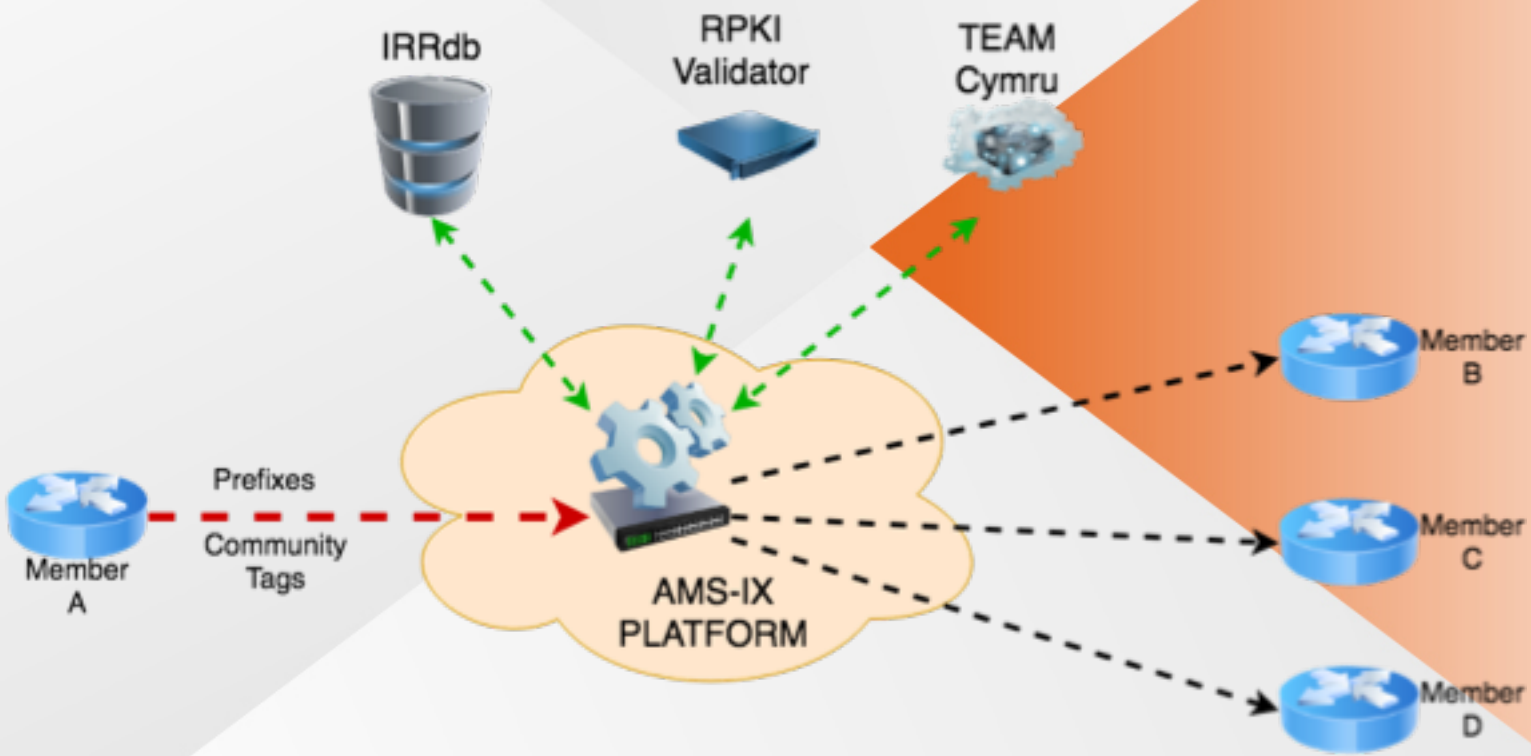
- Origin validation
  - Valid
  - Invalid
  - Unknown
- Policies applied

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*

*WiP

AMS-IX RS architecture



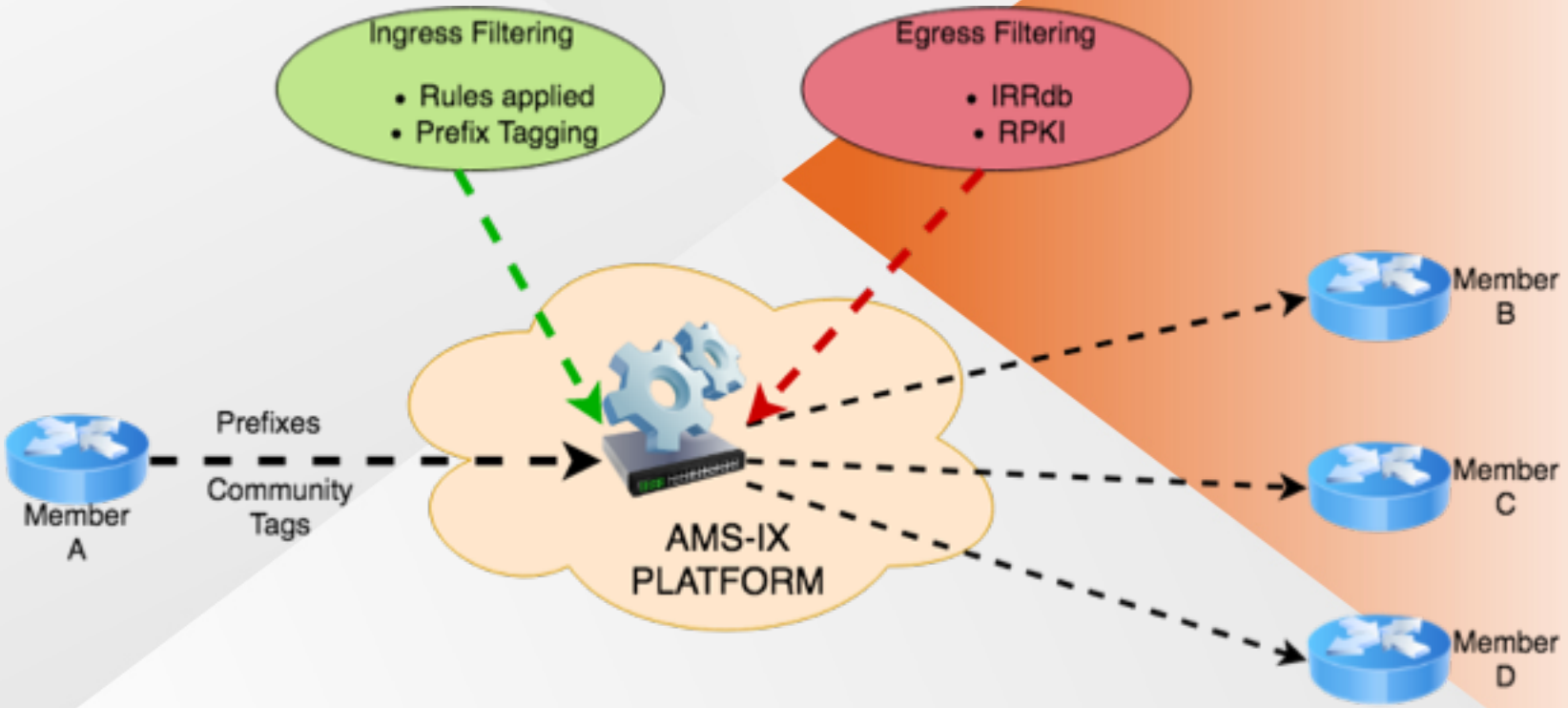
RS input explained

IRRdb	RPKI	Team Cymru
• Retrieve and Parse member's policy (import/export attributes) • Retrieve and resolve AS-SETs	• ROAs (via the RIPE validator which is connected to both RS)	• Bogons • Martians

Prefix filtering in AMS-IX

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

Where is applied



Peering rules (ingress)

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

*WiP

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*"
  - "No Filtering, *Just tagging*"

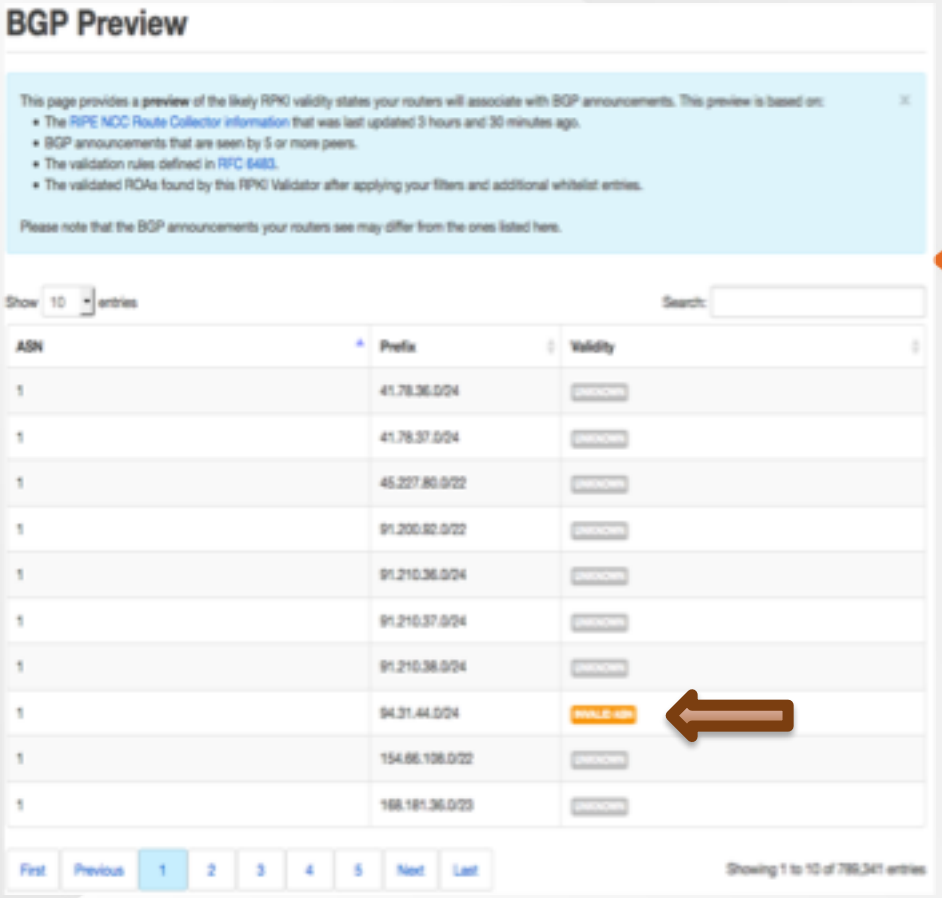
IRRdb Filtering 1/2

- 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



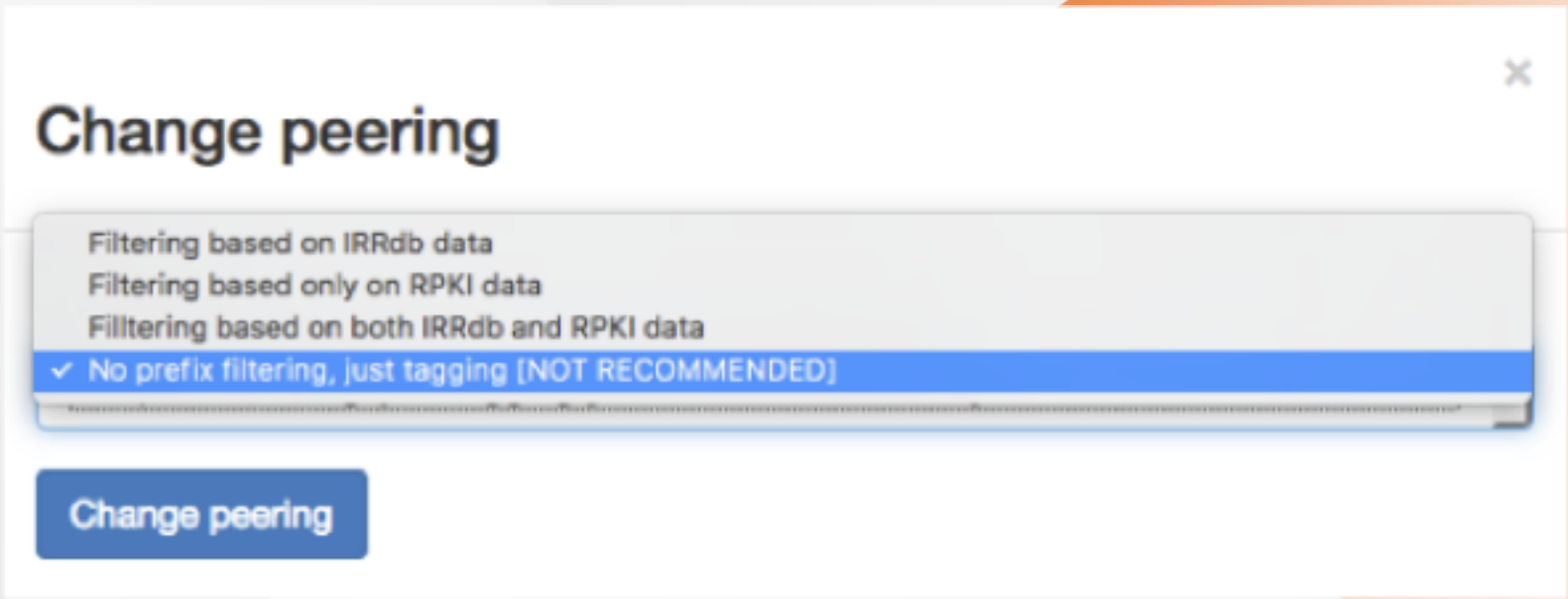
- BGP announcements are validated with RIPE's RPKI validator
- The prefixes that are being blocked are the ones with ROA status "**INVALID**"

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



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

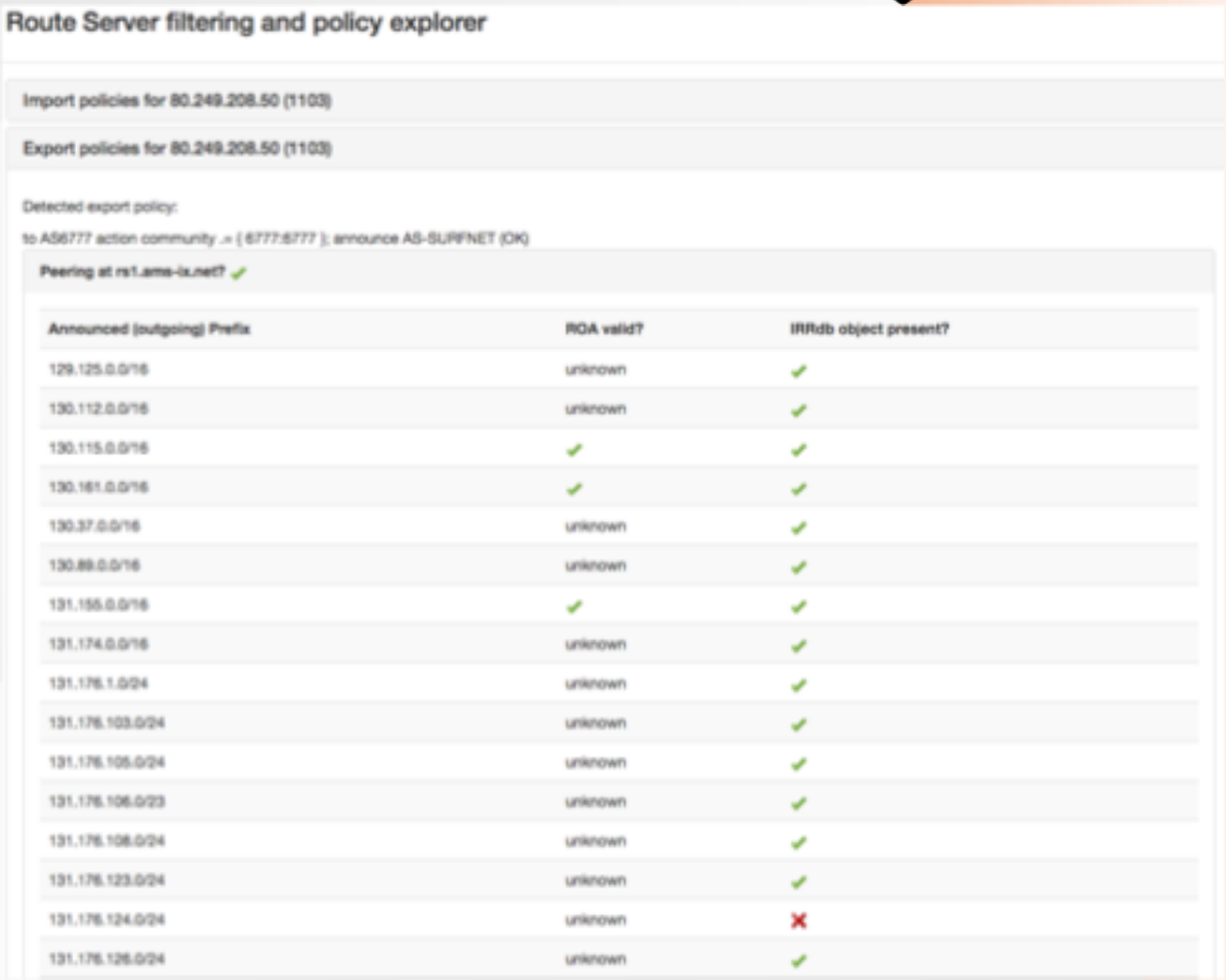
- Make “IRRdb+RPKI” filtering the default option
- Make “Just tagging” inaccessible
- Moved customers from “Just tagging” to the new default option
- Update relevant content in our website
- Inform our members and promote Routing Security

Upcoming features

- Policy Explorer
- E-mail notification system

Policy explorer

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



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