



# INEX's Shiny New Route Servers

INEX Members' Meeting, Dublin, March 2019

Barry O'Donovan

@ComePeerWithMe / @barryo79



RPKI

## IRRDB vs. RPKI ROAs

```
route6:          2001:db8::/32
descr:           Example IPv6 route object
origin:          AS65500
created:         2006-07-12T16:11:58Z
last-modified:  2011-02-22T15:58:03Z
source:          SOME-IRRDB
```

```
route:           192.0.2.0/24
descr:           Example IPv4 route object
origin:          AS65500
created:         2004-12-06T11:43:57Z
last-modified:  2016-11-16T22:19:51Z
source:          SOME-IRRDB
```

**RPKI**

## ROAs - Route Origin Authorisations

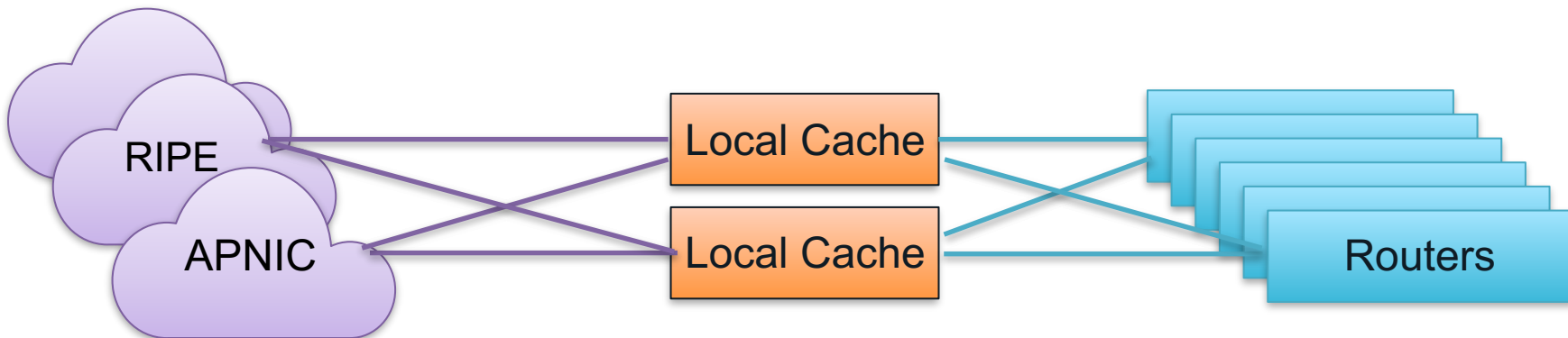
- A cryptographically secure replacement for route[6] objects
- Adds maximum prefix length
- Yields route origin triplets that have been validated

```
( Origin AS, Prefix, Max Length )  
( AS65500, 2001:db8::/32, /48 )  
( AS65501, 192.0.2.0/24, /24 )
```

RPKI

## Validating BGP Routing with RPKI-RTR

- A cache server (*validator*) does the cryptographic heavy lifting
- Routers receive and maintain the set of ROAs via RPKI-RTR from the cache
- RPKI gives three validation results: VALID, INVALID, UNKNOWN



RPKI

## Creating ROAs

- Use the RIPE LIR Portal - it's really easy
  - Multi-user support
  - Two-factor authentication
  - BGP route collector suggestions
  - Auto-renew ROAs



Home > ... > LIR Portal

You are editing

Internet Neutral Exchange Ass... ▾

My LIR >

**Resources** ▾

My Resources

Request Resources

Request Transfer

IPv4 Transfer Listing Service

RPKI Dashboard





RIPE Database >



BGP Announcements **Route Origin Authorisations (ROAs)** History

↓   Valid  Invalid  Unknown

<input type="checkbox"/>	Origin AS	Prefix	Current Status	
<input type="checkbox"/>	AS65500	192.0.2.0/24	<b>UNKNOWN</b>	 

Show  of 8 items

**RPKI**

## ROAs on the INEX Route Collector

```
bird> show route
  filter {
    if bgp_large_community ~ [( 2128, 1000, 1 )] then accept;
  }
table master4 count
```

**17710** of 244895 routes for **144834** networks in table master4

**=> 12.2% of routes at INEX have a ROA**



**RPKI**

## ROAs on the INEX Route Collector

```
bird> show route
  filter {
    if bgp_large_community ~ [( 2128, 1101, 13 )] then accept;
  }
table master4 count
```

535 of 244890 routes for **144825** networks in table master4

**=> 0.2% of routes at INEX have an invalid ROA**

\* 466 of these are via HE, 53 via Virgin Media; leaving only 16 for the other 98 route collector sessions

RPKI

## ROAs on the INEX Route Collector

AS6939	(1)	HE	11,346
AS6830	(2)	Virgin Media	3,799
AS8220	(3)	Colt	414
AS21320	(4)	GEANT	362
AS16509	(5)	AWS	307
AS13237	(6)	euNetworks	211
AS43531	(7)	IX Reach	88

AS15830	(11)	Equinix	53/322
AS5466	(13)	eir	26/77
AS207044	(14)	enet	22/53
AS31122	(15)	Viatel	20/64
AS34245	(15)	Magnet	20/22
AS39122	(16)	Blacknight	20/53
AS1213	(17)	HEAnet	12/23
AS2110	(17)	BT Ireland	12/95

RPKI

## ROAs on the INEX Route Collector [27/03/2019, 56 ASNs]

11651	6939	12	56911	2	62129
3882	6830	12	1213	2	61145
515	8220	10	8075	2	44384
377	21320	9	42	2	43192
307	16509	8	7713	2	41678
248	13237	8	51677	2	41073
91	43531	7	42473	2	39093
90	702	7	25441	2	31641
89	5400	5	44451	2	200562
74	15169	5	35226	2	199346
53	15830	5	15533	2	197853
31	22822	5	13335	2	15612
27	31122	4	39449	1	39319
26	5466	4	200005	1	3856
22	207044	4	199256	1	30900
21	39122	3	61194	1	203754
21	34245	3	60277	1	201607
20	14537	3	50326	1	12388
19	2110	3	32934		



**NEW ROUTE SERVERS**

## **INEX's Shiny New Route Servers**

## NEW ROUTE SERVERS

# Route Server Refresh at INEX & IXP Manager

- RPKI just one element
- Upgrade configuration from Bird v1.6 to Bird v2.0
- Complete rewrite of filtering workflow
  - Large communities used extensively within the route server
- Upgrade Bird's Eye<sup>1</sup> for Bird v2 BGP
- Overhaul IXP Manager looking glass

1. A secure micro service for querying Bird - <https://github.com/inex/birdseye>

## NEW ROUTE SERVERS


# Bird v1 to v2 Changes

- RPKI-RTR supported
- Collapsed separate daemons for IPv4 and IPv6 into a single daemon
  - master route table becomes master4 / master6
  - new protocol blocks: `ipv4 { ... } / ipv6 { ... }`
- Other very minor configuration changes

## NEW ROUTE SERVERS

# Bird v1 to v2 Changes

```
protocol bgp pb_as112_vli249_ipv4 {  
    description "AS112";  
    local as routerasn;  
    source address 192.0.2.8;  
    neighbor 192.0.2.6 as 112;  
    import all;  
    export none;  
    table master;  
}
```



```
protocol bgp pb_as112_vli249_ipv4 {  
    description "AS112";  
    local as routerasn;  
    source address 192.0.2.8;  
  
    neighbor 192.0.2.6 as 112;  
    ipv4 {  
        import all;  
        export none;  
        table master4;  
    };  
}
```

## NEW ROUTE SERVERS

# Standard IX Route Server Community Filters

Description	Community	Large Community
Prevent announcement of a prefix to a certain peer	0:peer-as	43760:0:peer-as
Announce a prefix to a certain peer	43760:peer-as	43760:1:peer-as
Prevent announcement of a prefix to all peers	0:43760	43760:0:0
Announce a prefix to a all peers	43760:43760	43760:1:0

Path prepends now available: <https://www.inex.ie/technical/route-servers/>



43760:1101:\* are filtered

## NEW ROUTE SERVERS

# Route Server BGP Community Usage

Description	Large Community
RPKI Valid	43760:1000:1
RPKI Unknown	43760:1000:2
IRRDB Valid	43760:1001:1
...	...

Description	Large Community
Bogon Prefix	43760:1101:3
IRRDB Invalid	43760:1101:9
RPKI Invalid	43760:1101:13
...	...

## NEW ROUTE SERVERS

# IXP Manager v5 Route Server Filtering

1. Small prefixes (default is  $> /24$  /  $/48$  for ipv4 / ipv6)
2. Martians / bogons
3. Ensure at least 1 ASN and  $\leq 64$  ASNs in path
4. Ensure peer AS is the same as first AS in the prefix's AS path
5. Prevent next-hop hijacking
6. Filter known transit networks
7. Ensure origin AS is in set of ASNs from member AS-SET
8. RPKI:
  - Valid -> accept
  - Invalid -> drop
9. RPKI Unknown -> revert to standard IRRDB prefix filtering

## NEW ROUTE SERVERS

# IXP Manager v5 Route Server Filtering

1. Small prefixes (default is  $> /24$  /  $/48$  for ipv4 / ipv6)
2. Martians / bogons
3. Ensure at least 1 ASN and  $\leq 64$  ASNs in path
4. Ensure peer AS is the same as first AS in the prefix's AS path
5. Prevent next-hop hijacking
- 6. Filter known transit networks**
7. Ensure origin AS is in set of ASNs from member AS-SET
8. RPKI:
  - Valid -> accept
  - Invalid -> drop
9. RPKI Unknown -> revert to standard IRRDB prefix filtering

## NEW ROUTE SERVERS

# Filter Known Transit Networks

These do not peer at IX's and they aren't typically customers of IX participants

```

14  define TRANSIT_ASNS = [ 174,           # Cogent
15                          209,           # Qwest (HE carries this on IXPs IPv6 (Jul 12 2018))
16                          701,           # UUNET
17                          702,           # UUNET
18                          1239,          # Sprint
19                          1299,          # Telia
20                          2914,          # NTT Communications
21                          3257,          # GTT Backbone
22                          3320,          # Deutsche Telekom AG (DTAG)
23                          3356,          # Level3
24                          3549,          # Level3
25                          3561,          # Savvis / CenturyLink
26                          4134,          # Chinanet
27                          5511,          # Orange opentransit
28                          6453,          # Tata Communications
29                          6461,          # Zayo Bandwidth
30                          6762,          # Seabone / Telecom Italia
31                          7018 ];          # AT&T

```

## NEW ROUTE SERVERS

# IXP Manager v5 Route Server Filtering

1. Small prefixes (default is  $> /24$  /  $/48$  for ipv4 / ipv6)
2. Martians / bogons
3. Ensure at least 1 ASN and  $\leq 64$  ASNs in path
4. Ensure peer AS is the same as first AS in the prefix's AS path
5. Prevent next-hop hijacking
6. Filter known transit networks
- 7. Ensure origin AS is in set of ASNs from member AS-SET**
8. RPKI:
  - Valid -> accept
  - Invalid -> drop
9. RPKI Unknown -> revert to standard IRRDB prefix filtering

## NEW ROUTE SERVERS

# Ensure Origin AS is Member's AS-SET

```
as-set:          AS-HEANET
descr:          Autonomous Systems routed by HEAnet
members:        AS1213, AS2128, AS112, AS42310, AS2850, AS-IEDR
remarks:        Group ASs routed by HEAnet together
mnt-by:         HEANET-NOC
source:         RIPE
```

No ability to create AS sets in RPKI

draft-ietf-grow-rpki-as-cones will resolve this

This is a regression over static IRRDB filtering

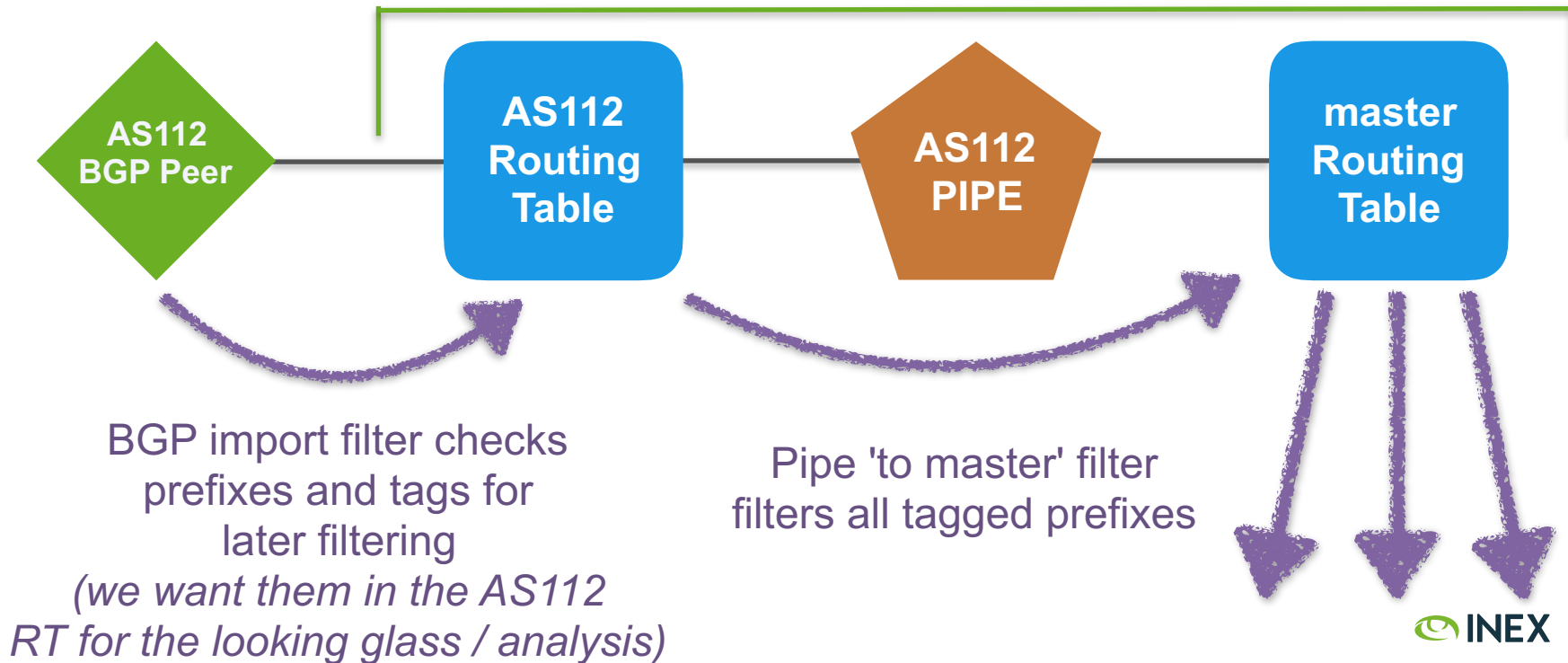
## NEW ROUTE SERVERS

# IXP Manager v5 Route Server Filtering

1. Small prefixes (default is  $> /24$  /  $/48$  for ipv4 / ipv6)
2. Martians / bogons
3. Ensure at least 1 ASN and  $\leq 64$  ASNs in path
4. Ensure peer AS is the same as first AS in the prefix's AS path
5. Prevent next-hop hijacking
6. Filter known transit networks
7. Ensure origin AS is in set of ASNs from member AS-SET
8. RPKI:
  - Valid -> accept
  - Invalid -> drop
9. RPKI Unknown -> revert to standard IRRDB prefix filtering

## NEW ROUTE SERVERS

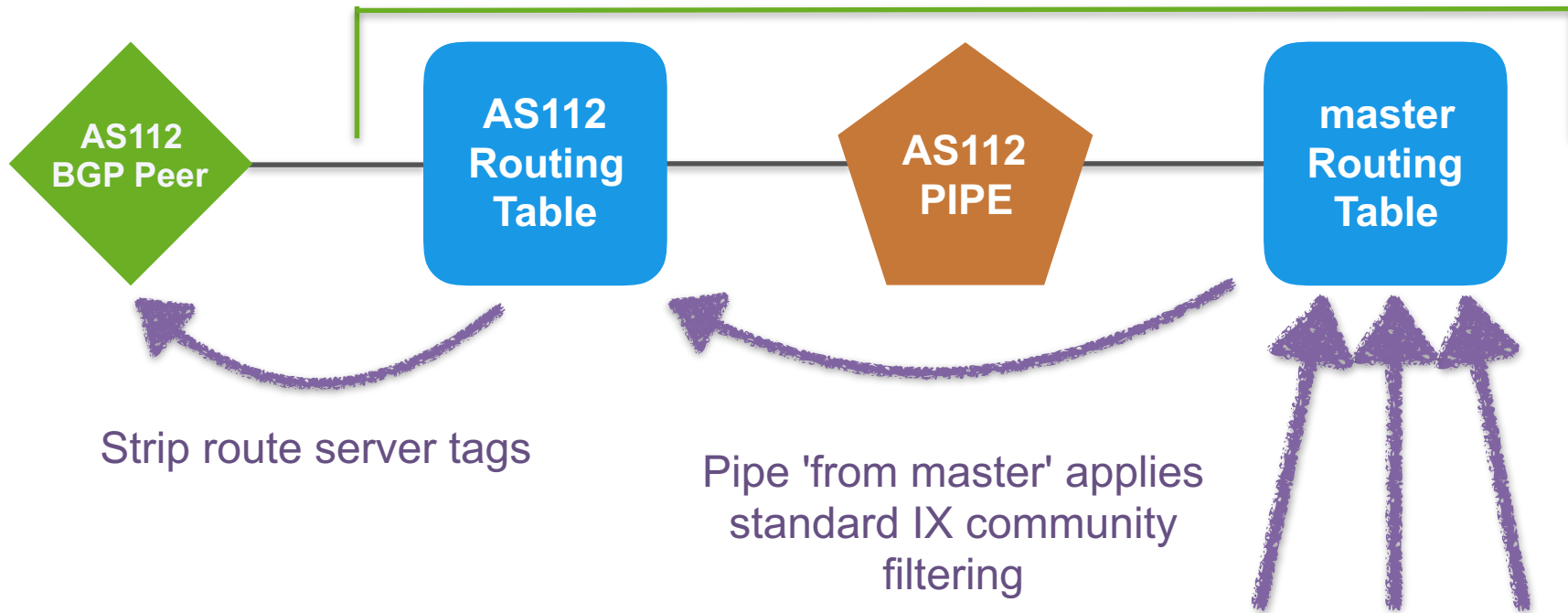
## IXP Manager v5 Bird Topology - Import From Member





## NEW ROUTE SERVERS

## IXP Manager v5 Bird Topology - Export To Member





RPKI @ INEX

## RPKI Implementation Notes

## Validator Software - RIPE NCC RPKI Validator 3

- RIPE NCC RPKI Validator 3 released in 2018
  - <https://github.com/RIPE-NCC/rpki-validator-3>
- Dramatically reduces installation complexity
- Modest VM requirements, runs on standard OS distributions
- Requirement to download ARIN TAL separately

```
$ wget https://ftp.ripe.net/tools/rpki/validator3/rc/generic/rpki-validator-latest-dist.tar.gz
$ tar xzf rpki-validator-latest-dist.tar.gz
$ ./rpki-validator-3.0-x/rpki-validator-3.sh
$ open http://localhost:8080
```

```
$ wget https://ftp.ripe.net/tools/rpki/validator3/rc/generic/rpki-rtr-server-latest-dist.tar.gz
$ tar xzf rpki-rtr-server-latest-dist.tar.gz
$ ./rpki-rtr-server/rpki-rtr-server-3.sh
```

## Validator Software - Routinator 3000

- Routinator 3000 by NLnet Labs
  - <https://github.com/NLnetLabs/routinator>
- First impressions: low overhead, installation simplicity, stable, "just works"
- Requirement to download ARIN TAL separately

```
$ curl https://sh.rustup.rs -sSf | sh
$ source ~/.cargo/env
$ cargo install routinator
$ routinator rtrd -al 127.0.0.1:3323
```

## Validator Software - RPKI-RTR and Bird

```
roa4 table t_roa;

protocol rpki rpki1 {

    roa4 { table t_roa; };

    remote "192.0.2.67" port 3323;

    retry keep 90;
    refresh keep 900;
    expire keep 172800;
}
```

RPKI @ INEX

## Validator Software - RPKI-RTR and Bird

```
# RPKI check
rpki_result = roa_check( t_roa, net, bgp_path.last_nonaggregated );

if( rpki_result = ROA_INVALID ) then {
    ...
}

# or ROA_VALID / ROA_UNKNOWN
```

# Implementation Process at INEX

- INEX has two route servers and a route collector per LAN
- Upgrade route collector to Bird v2 + RPKI first
  - identify members who peer on the route server with RPKI invalid prefixes
  - found 4 members of ~80 with issues
    - 1 x more specific advertised than ROA allowed for
    - 1 x origin AS not matching ROA
    - 1 x member still advertising transferred space, new owners had ROAs
    - 1 x member created ROA for upstream peer-as rather than origin-as
  - members alerted to this on a "FYI basis" (i.e. non-blocking for INEX)
- Route server #1 completed Feb 7th
- Route server #2 completed Feb 14th

## Implementation Process at INEX

- Outside of the four members with issues, no other member issues
- No issues to date with Bird v2
- Some issues with RIPE's validator (crashing, disk space)
- No issues with Routinator 3000
- There's a lot in this (Bird v2, route collector vs server, large community tagging and filtering, RPKI vs IRRDB, etc.)



## Looking Glass INEX Cork - Route Collector - IPv4

INEX Cork - Route Collector - IPv4   

This is the public looking glass. Uncached results and additional routers available when logged in.

Bird v2 2.0.3 | API: 1.2.0 | Router ID: 185.1.69.126 | Uptime: 11 days. | Last Reconfigure: 2019-02-16 15:12:02 | JSON: [\[status\]](#) [\[bgp\]](#)

Search: 

Neighbor	Description	ASN	Table	PfxLimit	State/PfxRcd	PfxExp	Actions
185.1.69.6	AS112 - AS112 Reverse DNS	112	<a href="#">master4</a>		2	0	<a href="#">Details</a>
185.1.69.24	AS714 - Apple Distribution International	714	<a href="#">master4</a>		596	0	<a href="#">Details</a>
185.1.69.26	AS714 - Apple Distribution International	714	<a href="#">master4</a>		597	0	<a href="#">Details</a>
185.1.69.11	AS1213 - HEAnet	1213	<a href="#">master4</a>		23	0	<a href="#">Details</a>
185.1.69.12	AS5466 - Eir	5466	<a href="#">master4</a>		77	0	<a href="#">Details</a>
185.1.69.17	AS15405 - East Cork Broadband	15405	<a href="#">master4</a>		5	0	<a href="#">Details</a>
185.1.69.14	AS16171 - Strencom	16171	<a href="#">master4</a>		4	0	<a href="#">Details</a>
185.1.69.16	AS20940 - Akamai Technologies	20940	<a href="#">master4</a>		1	0	<a href="#">Details</a>
185.1.69.23	AS25152 - RIPE NCC k-root server	25152	<a href="#">master4</a>		1	0	<a href="#">Details</a>
185.1.69.10	AS31122 - Viatel	31122	<a href="#">master4</a>		90	0	<a href="#">Details</a>
185.1.69.19	AS41736 - Nova Telecom	41736	<a href="#">master4</a>		3	0	<a href="#">Details</a>
185.1.69.21	AS42090 - Rapid Broadband	42090	<a href="#">master4</a>		6	0	<a href="#">Details</a>

Network	Next Hop	Metric	Communities?	AS Path	
<a href="#">104.132.227.0/24</a>	185.1.69.12	P 100	1 LC: 2	5466 41264	<a href="#">Details</a>
<a href="#">109.125.0.0/18</a>	185.1.69.12	P 100	1 LC: 2	5466 15751	<a href="#">Details</a>
<a href="#">132.189.78.0/24</a>	185.1.69.12	P 100	1 LC: 3 ⚠	5466 8116	<a href="#">Details</a>
<a href="#">132.189.79.0/24</a>	185.1.69.12	P 100	1 LC: 3 ⚠	5466 8116	<a href="#">Details</a>
<a href="#">132.237.132.0/24</a>	185.1.69.12	P 100	1 LC: 2	5466 30614	<a href="#">Details</a>
<a href="#">132.237.167.0/24</a>	185.1.69.12	P 100	1 LC: 2	5466 30614	<a href="#">Details</a>
<a href="#">134.191.192.0/24</a>	185.1.69.12	P 100	1 LC: 2	5466 4983	<a href="#">Details</a>
<a href="#">134.191.216.0/22</a>	185.1.69.12	P 100	1 LC: 2	5466 4983 4983 4983 4983 4983 4983 4983 4983 4983	<a href="#">Details</a>
<a href="#">134.191.220.0/23</a>	185.1.69.12	P 100	1 LC: 2	5466 4983 4983 4983 4983 4983 4983 4983 4983 4983	<a href="#">Details</a>
<a href="#">134.191.240.0/22</a>	185.1.69.12	P 100	1 LC: 3 ⚠	5466 4983	<a href="#">Details</a>
<a href="#">134.191.244.0/24</a>	185.1.69.12	P 100	1 LC: 3 ⚠	5466 4983	<a href="#">Details</a>
<a href="#">134.191.246.0/23</a>	185.1.69.12	P 100	1 LC: 2	5466 4983	<a href="#">Details</a>
<a href="#">135.74.153.0/24</a>	185.1.69.12	P 100	1 LC: 3 ⚠	5466 18676	<a href="#">Details</a>
<a href="#">146.214.64.0/23</a>	185.1.69.12	P 100	1 LC: 3 ⚠	5466 42213	<a href="#">Details</a>

### Route Details - 132.189.78.0/24 as received from protocol pb\_as5466\_vli223\_ipv4

Network	132.189.78.0/24	
Gateway	185.1.69.12	PRIMARY
From Protocol	pb_as5466_vli223_ipv4	
Age	2019-02-12 09:12:03	
Metric	100	
Type	BGP univ	
BGP :: AS Path	5466 8116	
BGP :: Local Pref	100	
BGP :: Communities	5466:20	
BGP :: Large Communities	2128:1000:2	RPKI UNKNOWN
	2128:1101:9	IRRDB PREFIX FILTERED
	2128:1001:1001	IRRDB FILTERED STRICT

Close



Background interface showing a list of IP networks on the left and a 'Details' button for each row on the right. The selected row is 132.189.78.0/24.

RPKI @ INEX

# New Route Server Filtered Prefixes Tool

Your INEX - IXP Manager Dashboard



The screenshot shows the top portion of the INEX IXP Manager dashboard. At the top is a navigation bar with several menu items: 'Overview', 'Details', 'Ports', 'Cross Connects', 'Filtered Prefixes' (which is highlighted with a green arrow), 'Peering Manager', 'Statistics', and 'Peer to Peer Traffic'. Below the navigation bar, the dashboard is divided into two main sections. On the left is a section titled 'Aggregate Traffic Statistics' with a search icon. On the right is a section titled 'Recent Members' with a sub-header 'Our five most recent members are listed below. Have you arranged peering with them yet?'.

# Route Server Filtered Prefixes

**Bad news!** We found 9 prefix(es) that are currently being filtered.

These are listed below with the reason for the filtering and the route server where filtering has been applied.

Prefix	Filtered Because	Filtered On Router(s)
87.232.5.0/24	IRRDB PREFIX FILTERED	rs1-lan1-ipv4 rs2-lan1-ipv4
87.232.128.0/21	RPKI INVALID	rs1-lan1-ipv4 rs2-lan1-ipv4
87.232.64.0/18	NEXT HOP NOT PEER IP	rs1-lan1-ipv4 rs2-lan1-ipv4
87.232.32.0/19	RPKI INVALID	rs1-lan1-ipv4 rs2-lan1-ipv4
91.197.36.0/22	TRANSIT FREE ASN	rs1-lan1-ipv4 rs2-lan1-ipv4

THANK YOU

# Any Questions?