

# Euro-IX Route Server Workshop

Barry O'Donovan  
NAMEX Offices, Rome, Italy  
March 13th, 2024

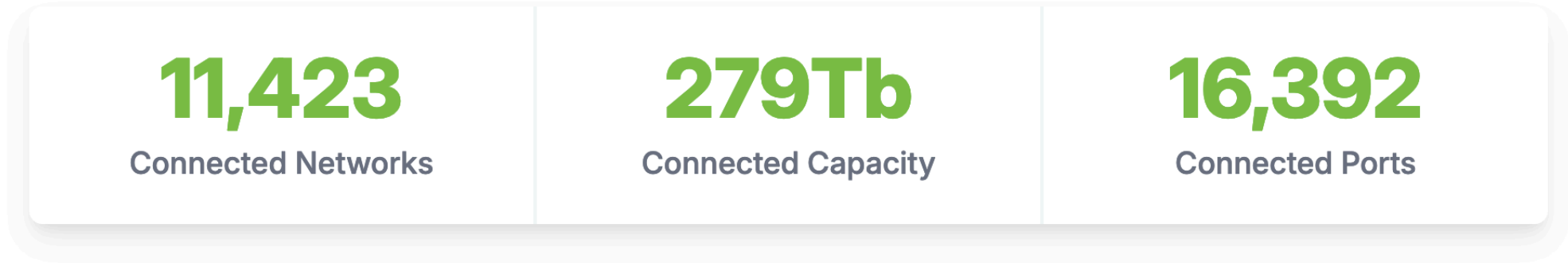
# INEX

- Peering point for the island of Ireland, member owned, cost recovery basis, founded in 1996.
- ~110 members (~95% of eyeballs)
- New peak @ 1.27 Tbps last week
- Dual infrastructure, 7+2 PoPs, DF
- Opened INEX Cork in 2016
- Home of IXP Manager
- $3 \times 2 \times (1+1+1+2) = 30$  Bird instances



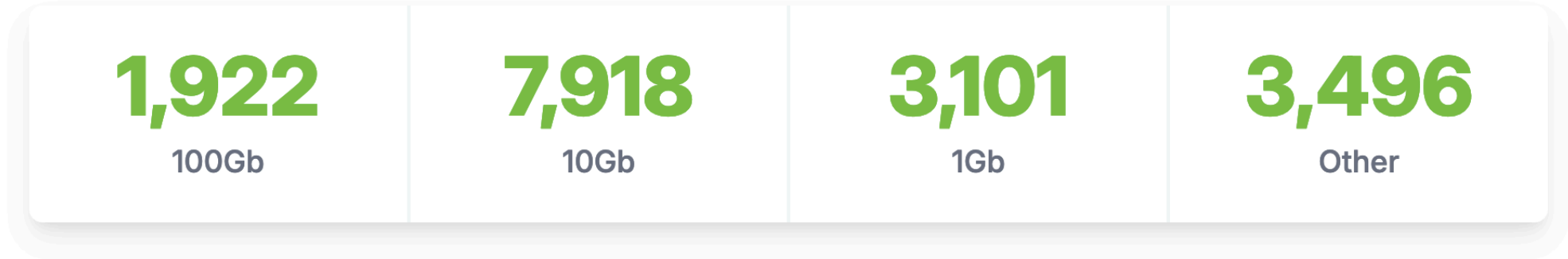
# IXP Manager





There are at least **230** IXPs Powering Peering using IXP Manager around the world to connect **11,423** networks of which **6,724** are unique. The edge peering capacity of these networks is **279Tb** over **16,392** connected ports. We aggregate traffic statistics from 135 IXP Manager platforms which show us that the peak traffic exchanged is **74.4 Tbps**.

Port Breakdown



# IXP Manager - Most Connected Networks

NETWORK	ASN	# IXPS
Packet Clearing House	3856	80
Packet Clearing House AS42	42	78
Hurricane Electric	6939	78
Cloudflare	13335	78
Akamai Technologies	20940	56
Microsoft	8075	46
Google LLC	15169	46
Facebook Inc	32934	46
Amazon.com	16509	45
DNS-OARC-112	112	41

# IXP Manager - New Developer

- Feb 2024 - second attempt at hiring
- “PHP Developer - Help Build the Internet”
  - $\geq$  3 years professional PHP development
  - Hybrid but requires some office attendance
  - Desirable: Laravel; ‘full-stack’; FOSS experience
- Job accepted last week & started this week
- Fully funded via the “IXP Manager Sponsorship Program”
  - <https://www.ixpmanager.org/sponsorship/terms>
  - Funding on hand to cover the next year
  - Reviewing other possible sponsorship streams

# Route Servers & Filtering



## [inex-tech] Route server system now in beta

Nick Hilliard

Fri Nov 23 12:20:17 GMT 2007

Following the announcement at the last INEX members meeting that we were looking into running a route server system, we are now pleased to announce that we now have a route server system which is in stable beta.

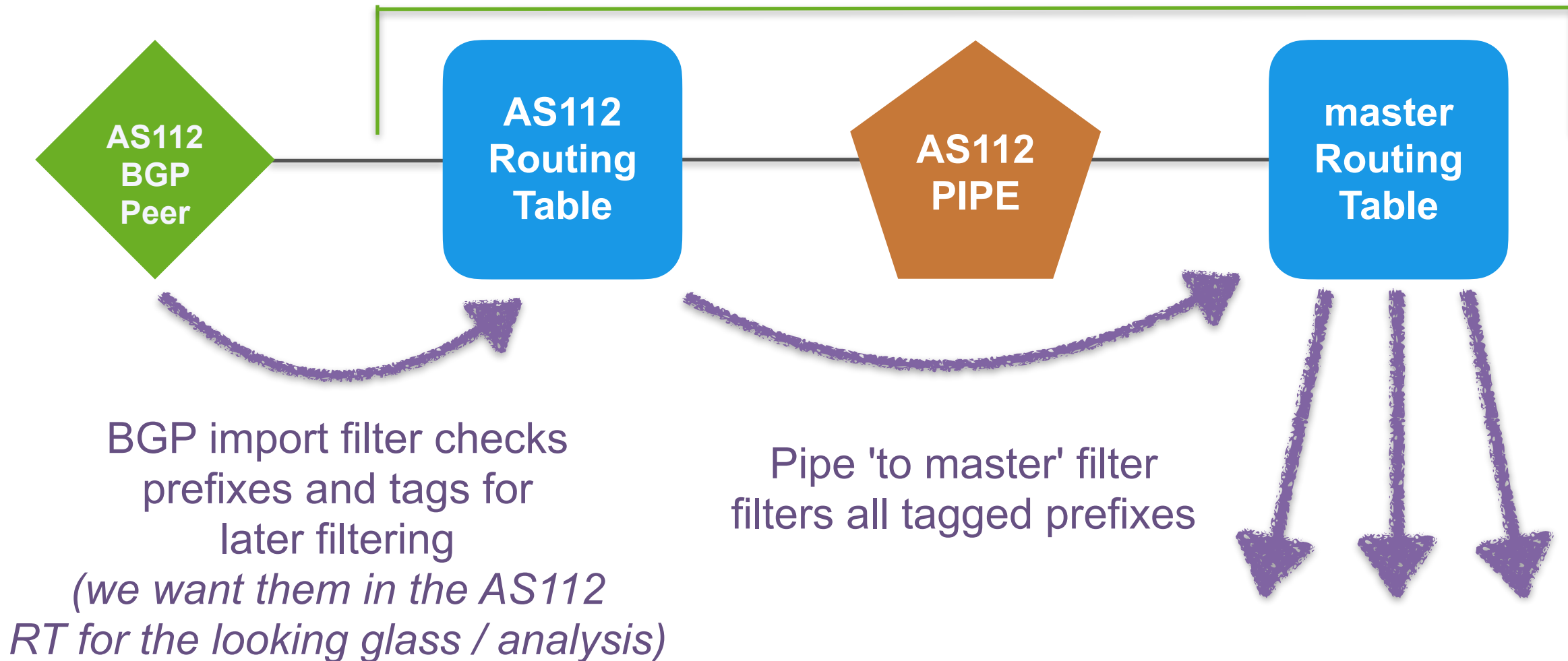
As a brief summary, the route server system offers the following advantages:

- dramatically reduces the number of BGP sessions required to peer with other INEX members
- **strict route filtering on inbound announcements means that only prefixes registered at RIPE by exchange members will be visible**
- dual-hosted system offers high reliability
- **community based filtering** allows route server users to control which INEX members their prefixes are sent to

# 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 *(not implemented: PeeringDB never\_via\_ixp)*
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

# IXP Manager v5 Bird Topology - Import From Member



# Communities

```
define IXP_LC_FILTERED_PREFIX_LEN_TOO_LONG      = ( routeserverasn, 1101, 1 );
define IXP_LC_FILTERED_PREFIX_LEN_TOO_SHORT    = ( routeserverasn, 1101, 2 );
define IXP_LC_FILTERED_BOGON                   = ( routeserverasn, 1101, 3 );
define IXP_LC_FILTERED_BOGON_ASN              = ( routeserverasn, 1101, 4 );
define IXP_LC_FILTERED_AS_PATH_TOO_LONG       = ( routeserverasn, 1101, 5 );
define IXP_LC_FILTERED_AS_PATH_TOO_SHORT     = ( routeserverasn, 1101, 6 );
define IXP_LC_FILTERED_FIRST_AS_NOT_PEER_AS   = ( routeserverasn, 1101, 7 );
define IXP_LC_FILTERED_NEXT_HOP_NOT_PEER_IP   = ( routeserverasn, 1101, 8 );
define IXP_LC_FILTERED_IRRDB_PREFIX_FILTERED  = ( routeserverasn, 1101, 9 );
define IXP_LC_FILTERED_IRRDB_ORIGIN_AS_FILTERED = ( routeserverasn, 1101, 10 );
define IXP_LC_FILTERED_PREFIX_NOT_IN_ORIGIN_AS = ( routeserverasn, 1101, 11 );

define IXP_LC_FILTERED_RPKI_UNKNOWN           = ( routeserverasn, 1101, 12 );
define IXP_LC_FILTERED_RPKI_INVALID          = ( routeserverasn, 1101, 13 );
define IXP_LC_FILTERED_TRANSIT_FREE_ASN      = ( routeserverasn, 1101, 14 );
define IXP_LC_FILTERED_TOO_MANY_COMMUNITIES  = ( routeserverasn, 1101, 15 );
```

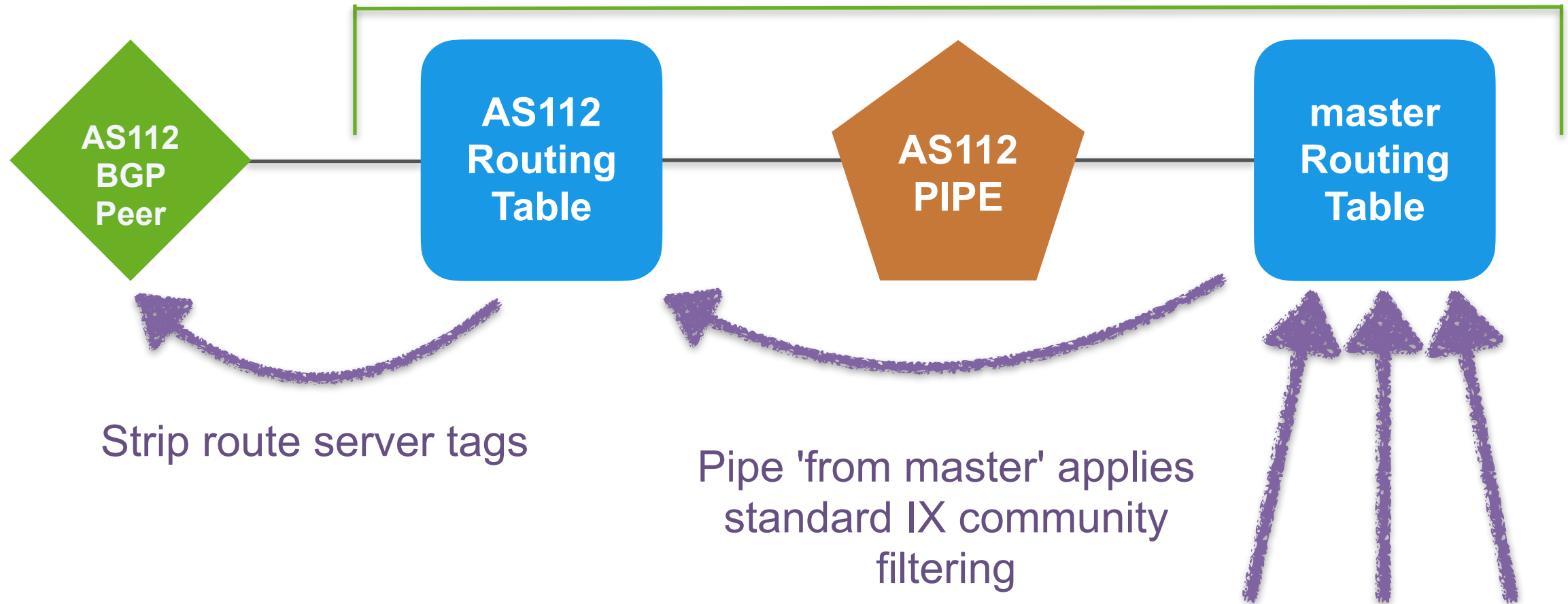
# Communities

```
define IXP_LC_INFO_RPKI_VALID          = ( routeserverasn, 1000, 1 );
define IXP_LC_INFO_RPKI_UNKNOWN        = ( routeserverasn, 1000, 2 );
define IXP_LC_INFO_RPKI_NOT_CHECKED   = ( routeserverasn, 1000, 3 );

define IXP_LC_INFO_IRRDB_VALID         = ( routeserverasn, 1001, 1 );
define IXP_LC_INFO_IRRDB_NOT_CHECKED  = ( routeserverasn, 1001, 2 );
define IXP_LC_INFO_IRRDB_MORE_SPECIFIC = ( routeserverasn, 1001, 3 );

define IXP_LC_INFO_IRRDB_FILTERED_LOOSE = ( routeserverasn, 1001, 1000 );
define IXP_LC_INFO_IRRDB_FILTERED_STRICT = ( routeserverasn, 1001, 1001 );
define IXP_LC_INFO_IRRDB_PREFIX_EMPTY  = ( routeserverasn, 1001, 1002 );

define IXP_LC_INFO_SAME_AS_NEXT_HOP = ( routeserverasn, 1001, 1200 );
```

**NEW ROUTE SERVERS****IXP Manager v5 Bird Topology - Export To Member**

# The Need for Route Server Filtering

- You are essentially “outsourcing” your routing policy
  - Usually fine - most networks peer openly at an IXP
- “I want to peer with everyone except X”
  - Route servers need some knobs to allow this
- Mostly standardised BGP community schema provides this
  - <https://github.com/euro-ix/rs-workshop-july-2017/wiki/Route-Server-BGP-Community-usage>
- Certainly standardised at over 200 exchanges using IXP Manager

# (Large) Community Based Filtering

Action	Community
Prevent announcement of a prefix to a peer	43760:0:peer-as
Announce a route to a certain peer	43760:1:peer-as
Prevent announcement of a prefix to all peers	43760:0:0
Announce a prefix to all peers ( <i>default</i> )	43760:1:0

NAMESPACE : ACTION : TARGET



# Extra Filtering Options Large Communities

Action	Community
Prepend to peer AS once	43760:101:peer-as
Prepend to peer AS twice	43760:102:peer-as
Prepend to peer AS thrice	43760:103:peer-as

peer-as == 0 ==> ALL PEERS

NAMESPACE : ACTION : TARGET

# Community Based Filtering in Practice

- Difficult at both ends of the network-size scale:
  - Small networks rarely touch their border routers
  - Large networks need cumbersome change control procedures
- Very complicated in a pinch
  - Community filtering is only half the story!
    - Still need to filter the routes you learn from the route servers
- DDoS events of Q2 2021
  - INEX Operations implemented route server filtering on an emergency basis for a number of members.

# UI Based Filtering in IXP Manager







# NEW: UI Based Filtering in IXP Manager

- Move the complexity from member router to route server (RS)
- Mechanism is unchanged - just “where it happens” moves:
  - RS tags your routes in ingress rather than you doing it on egress
  - RS filters routes to be advertised to you on egress rather than you on ingress
- Solve for 90% of use-cases

# UI Based Filtering in IXP Manager

## Route Server Filtering for AS112 Reverse DNS

✓ Your filters are in sync with our production configuration.

Peer	LAN	Protocol	Advertised Prefix	Advertise Action	Received Prefix	Receive Action	Enabled	Order	Actions
<a href="#">AwkwardNet</a>	All	IPv6	*	Do Not Advertise	*	Do Not Receive (Drop)	Yes	1	     

# UI Based Filtering in IXP Manager

! Your filters are not in sync with our production configuration. You can continue editing or:

Revert

Commit

## Staged Rules (Deploy via Commit above)

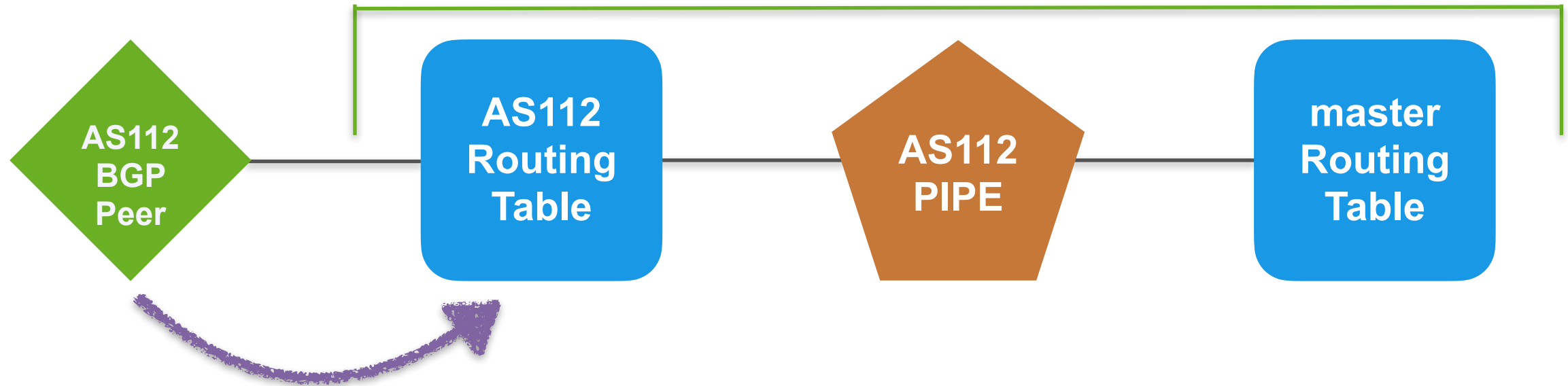
Peer	LAN	Protocol	Advertised Prefix	Advertise Action	Received Prefix	Receive Action	Enabled	Order	Actions
<a href="#">AwkwardNet</a>	All	IPv4	*	Do Not Advertise	*	Do Not Receive (Drop)	Yes	1	     

## Rules in Production

Peer	LAN	Protocol	Advertised Prefix	Advertise Action	Received Prefix	Receive Action	Enabled	Order
<a href="#">AwkwardNet</a>	All	IPv6	*	Do Not Advertise	*	Do Not Receive (Drop)	Yes	1

## NEW ROUTE SERVERS

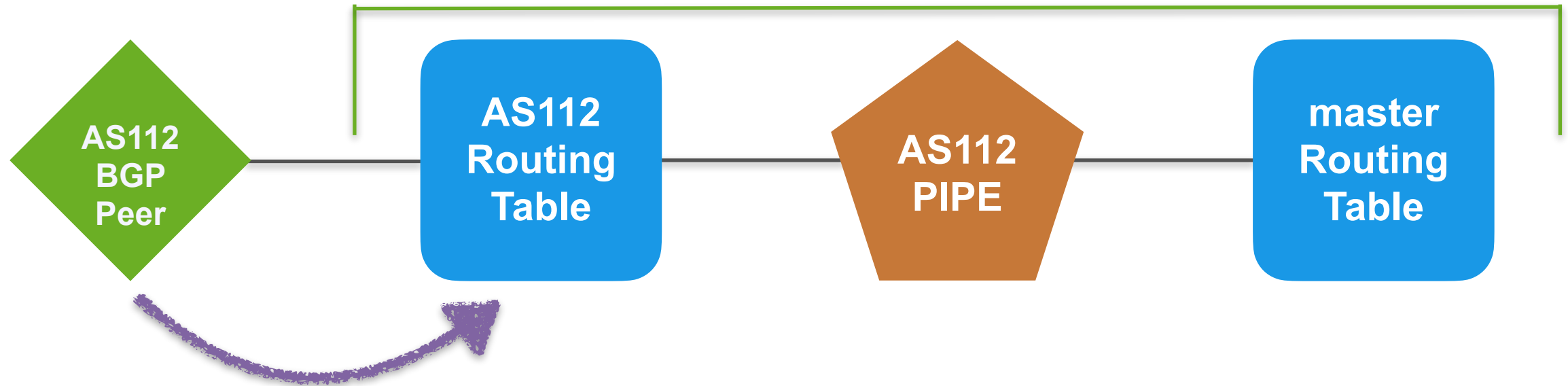
# IXP Manager v5 Bird Topology - Import From Member



Tagging moves from member  
router egress to route  
server ingress

## NEW ROUTE SERVERS

# IXP Manager v5 Bird Topology - Export To Member



Filtering moves from member ingress  
to route server egress



A globe of the Earth is shown from a high-angle perspective, centered on the Atlantic Ocean. Overlaid on the globe is a complex network of glowing orange lines that represent server connections or data routes. The lines are most dense in North America and Europe, with many lines radiating from central hubs. The globe's colors are muted, with blues for the oceans and greens/browns for the continents. The overall aesthetic is technical and digital.

# Route Server Resilience

# Background

- Route servers considered a critical production service
  - Deployed in pairs on each LAN
- Run on Dell hypervisors treated as production network appliances
  - No other function
  - Dual PSU, hardware RAID, iDRAC
  - Deployed in different PoPs
  - Full array of monitoring / checks

# Previous Configuration Update Method

- Via script ~4 times per day, offset against its resilient route server
- Sanity check - don't install a rs2 config on rs1 ('rs1-lan1-ipv4')
- Check retval for curl api call for config
- Check downloaded file exists and is non-zero size
- Check there are member BGP config stanzas in file (grep)
- Use Bird to parse the config file and check retval
- Backup old config file and replace with new
- Is Bird running? Either start or reconfigure as appropriate
- If reconfiguring and it fails, revert to old configuration file
- API call to IXP Manager to signal update complete

# New Configuration Update Method

- On demand and no longer offset. Script augmented:
- IXP Manager database updated so that all routers now have named pairs
- Script now obtains a lock before starting the update process
  - A locked router prevents its pair from updating
- Additionally, old config and new config are diff'd
  - Bird only reloaded if configuration has actually changed
- API signal to mark update as complete also releases lock
- I.e. A failed update run will prevent the paired router from updating
- Stress tested on the route collectors

Handle	Name	Vlan	Pair	Peering IP	ASN	Last Updated	Actions
rs1-cork-ipv4	RS1 - Cork - IPv4	INEX Cork	rs2-cork-ipv4	185.1.69.8	43760	5 minutes ago	
rs1-cork-ipv6	RS1 - Cork - IPv6	INEX Cork	rs2-cork-ipv6	2001:7f8:18:210::8	43760	3 minutes ago	
rs1-lan1-ipv4	RS1 - LAN1 - IPv4	INEX LAN1	rs2-lan1-ipv4	185.6.36.8	43760	36 minutes ago	
rs1-lan1-ipv6	RS1 - LAN1 - IPv6	INEX LAN1	rs2-lan1-ipv6	2001:7f8:18::8	43760	36 minutes ago	
rs1-lan2-ipv4	RS1 - LAN2 - IPv4	INEX LAN2	rs2-lan2-ipv4	194.88.240.8	43760	36 minutes ago	
rs1-lan2-ipv6	RS1 - LAN2 - IPv6	INEX LAN2	rs2-lan2-ipv6	2001:7f8:18:12::8	43760	36 minutes ago	

# GitHub Actions for RS Testing

🏠 Summary

Jobs

✅ ci-ex-dusk

Run details

🕒 Usage

📄 Workflow file

## ci-ex-dusk

succeeded 3 weeks ago in 1m 8s

🔽 Setup php server

🔽 ✅ Execute tests (Unit and Feature tests) via PHPUnit

```

1 ▶ Run vendor/bin/phpunit --testsuite 'Docstore Test Suite,IXP Manager Test Suite'
6 PHPUnit 9.6.15 by Sebastian Bergmann and contributors.
7
8 ..... 63 / 277 ( 22%)
9 ..... 126 / 277 ( 45%)
10 ..... 189 / 277 ( 68%)
11 ..... 252 / 277 ( 90%)
12 ..... 277 / 277 (100%)
13
14 Time: 00:30.474, Memory: 131.50 MB
15
16 OK (277 tests, 890 assertions)
  
```



**What's Next?**

# Development Plans

- Release the new features!
- IRRDB update could crash causing unprocessed members to go stale
- Iterative / per-member config generation rather than “monolithic”
- Event-based triggers rather than cron-like
- OpenBGPd (‘Equiv class’ support: config gen; monitoring & alerting; looking glass; prefixes filtered reporting; ...)
- rfc8950 - Advertising IPv4 Network Layer Reachability Information (NLRI) with an IPv6 Next Hop
- Birdwatcher / others + other telemetry support
- BFD support





# Thank you



**INEX**

INTERCONNECTING NETWORKS  
AND PEOPLE FOR OVER 25 YEARS