Telia Carrier

ROUTE FILTERING WITH RPKI

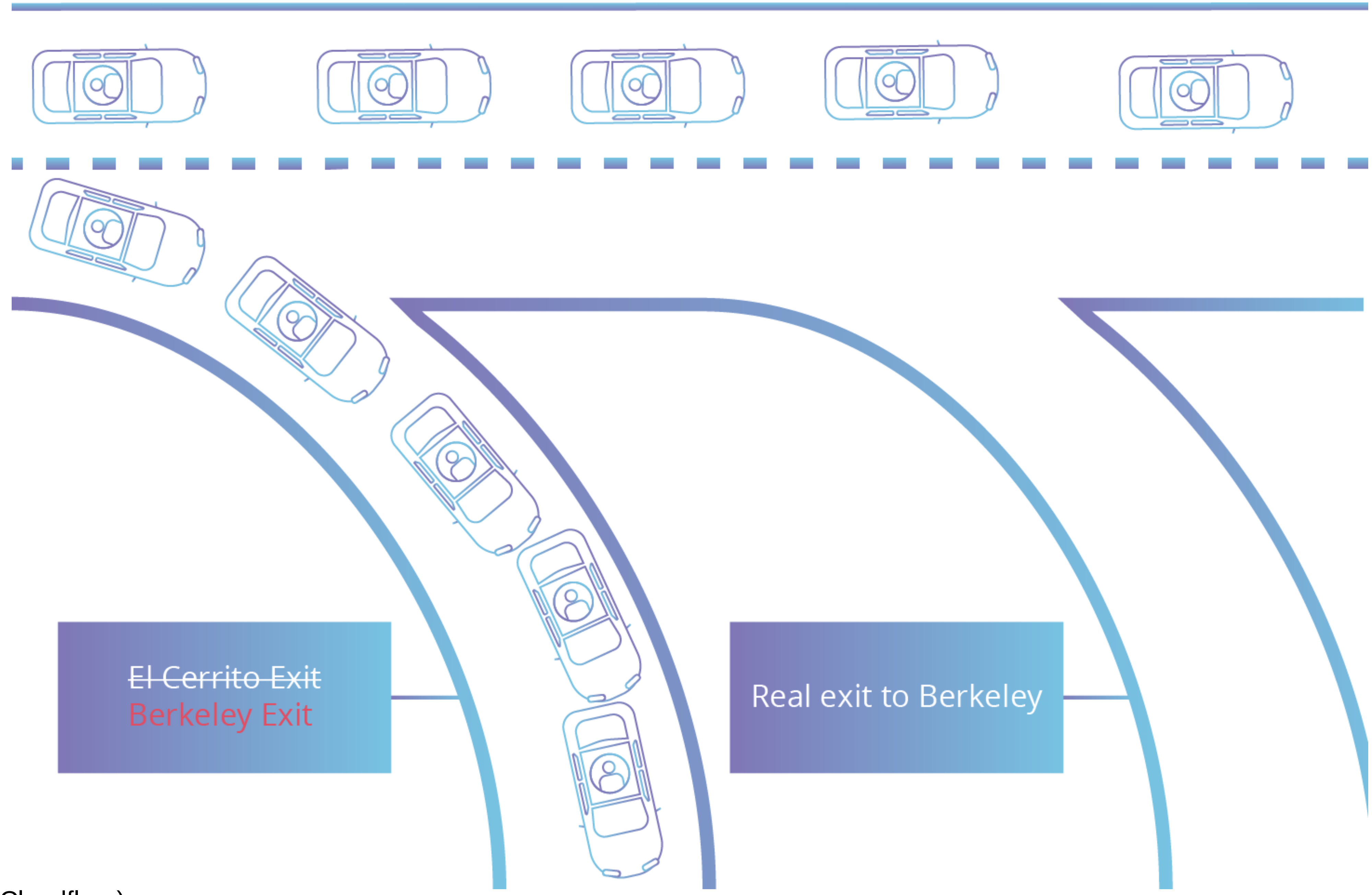
Carl Fredrik Lagerfeldt
Global Peering Manager

CONTENTS

- What is RPKI and what problem is it trying to solve?
- What we’ve done at Telia Carrier
- Reactions from customers & pitfalls to avoid

WHAT IS A BGP HIJACK OR ROUTE LEAK?

- BGP is too trusting in nature
- BGP neighbours are able to ‘advertise’ routes to anywhere



~~El Cerrito Exit~~
Berkeley Exit

Real exit to Berkeley

(image credit: Cloudflare)

WHAT IS A BGP HIJACK OR ROUTE LEAK?

BGP Hijacks

Illegitimate advertisement of foreign address or AS number space.

This can be intentional or unintentional announcement

BGP Route Leaks

Illegitimate announcement of a route received from a peer/upstream to another peer/upstream.

EXAMPLES

- **1997** - AS7007 mistakenly (re)announces 72K+ routes (becomes the poster-child for route filtering).
- **2008** - ISP in Pakistan accidentally announces IP routes for YouTube by blackholing the video service internally to their network.
- **2017** - Russian ISP leaks 36 prefixes for payments services owned by Mastercard, Visa, and major banks.
- **2018** - BGP hijack of Amazon DNS to steal crypto currency.
- **2019** - AS21217 leaked 70K+ routes to China Telecom in Frankfurt, redirecting a lot of European networks through their network.
- **2019** - AS33154 leaked more specific routes generated by a *“BGP Optimizer”* to Verizon, impacting major parts of the Internet.

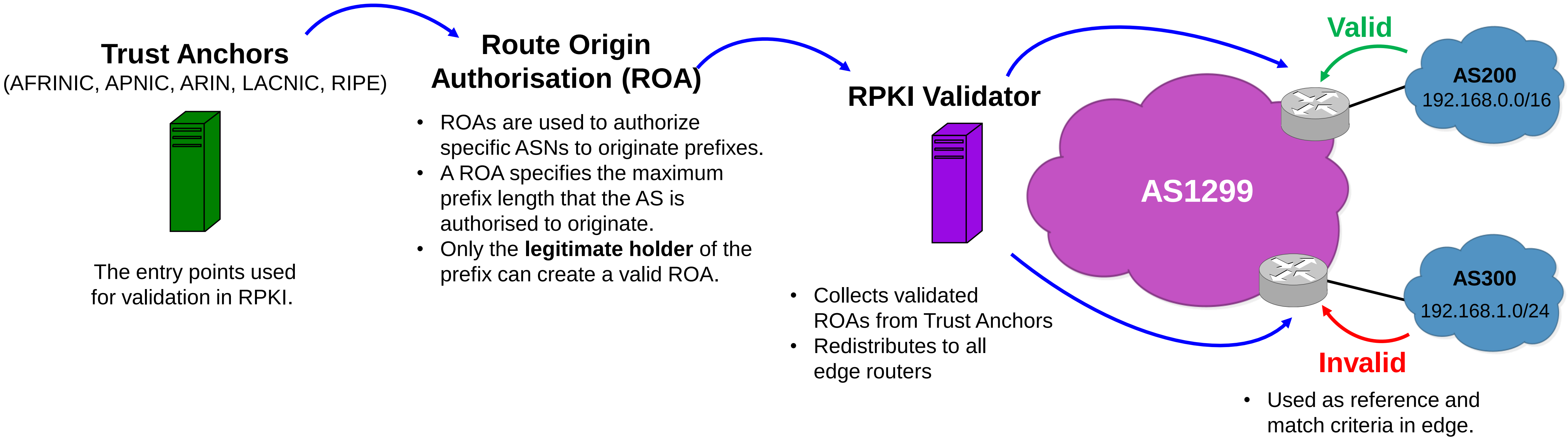
RPKI?

WHAT IS RPKI?

- Resource Public Key Infrastructure
- Route Validation for BGP Announcements
- Helps prevent ‘BGP Hijacks’ and ‘Route Leaks’

RESOURCE PUBLIC KEY INFRASTRUCTURE (RPKI)

The Resource Certification (RPKI) system allows Local Internet Registries (LIRs) to request a digital certificate listing the Internet number resources they hold. It offers validatable proof of holdership of a resource's registration by a Regional Internet Registry (RIR).



Trust Anchors
(AFRINIC, APNIC, ARIN, LACNIC, RIPE)

The entry points used for validation in RPKI.

Route Origin Authorisation (ROA)

- ROAs are used to authorize specific ASNs to originate prefixes.
- A ROA specifies the maximum prefix length that the AS is authorised to originate.
- Only the **legitimate holder** of the prefix can create a valid ROA.

RPKI Validator

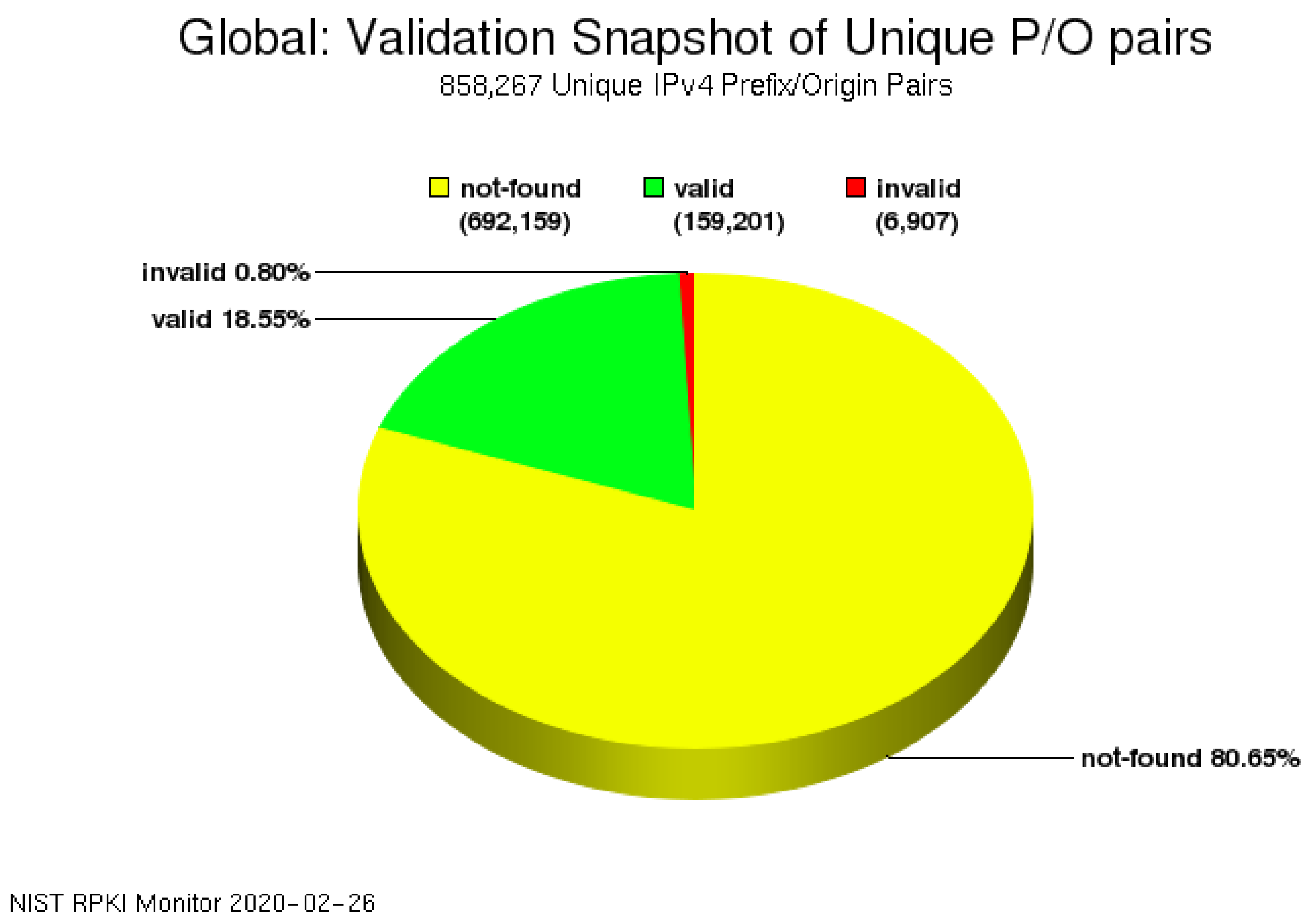
- Collects validated ROAs from Trust Anchors
- Redistributes to all edge routers

- Used as reference and match criteria in edge.

RPKI VALIDATION STATES

STATE	DESCRIPTION	RECOMMENDED ACTION
Valid	A matching ROA exists, all criteria matches	None, all good
Unknown	No ROA is registered for the prefix	Register ROAs for your IP space
Invalid	A matching ROA exists for the prefix, origin-AS and/or mask-length is not matching record	Possible hijack or route leak, need to withdraw & re-register ROA
Unverified	Validation error, origin-validation is skipped	Check your RPKI infrastructure/validator

CURRENT STATE OF THE DFZ



Global: Validation Snapshot of Unique P/O pairs

858,267 Unique IPv4 Prefix/Origin Pairs

NIST RPKI Monitor 2020-02-26

https://rpki-monitor.antd.nist.gov/

AS1299

SOME PERSPECTIVE

2,100+

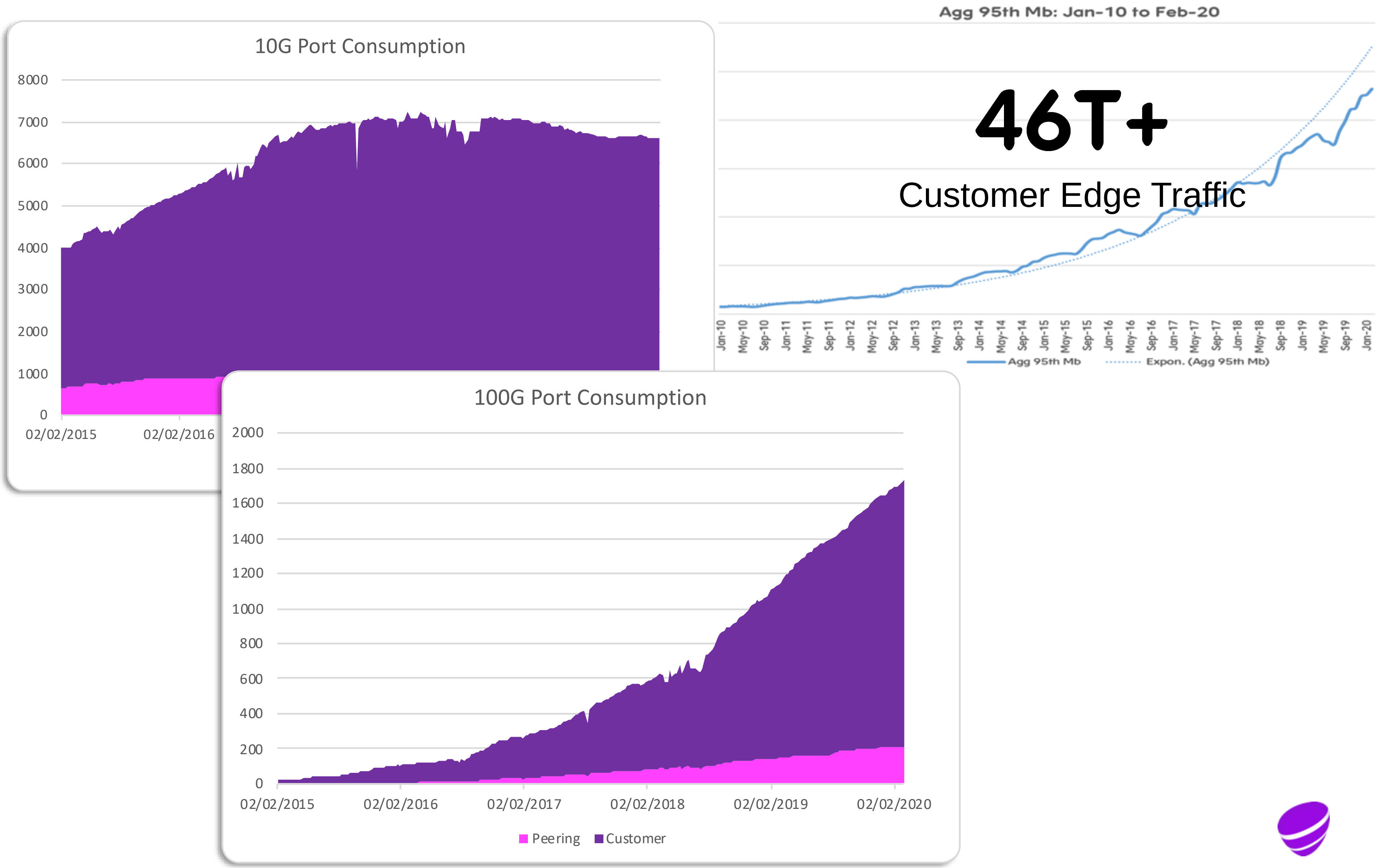
Directly Connected Customer ASNs

200+

Number of Edge devices

11,000+

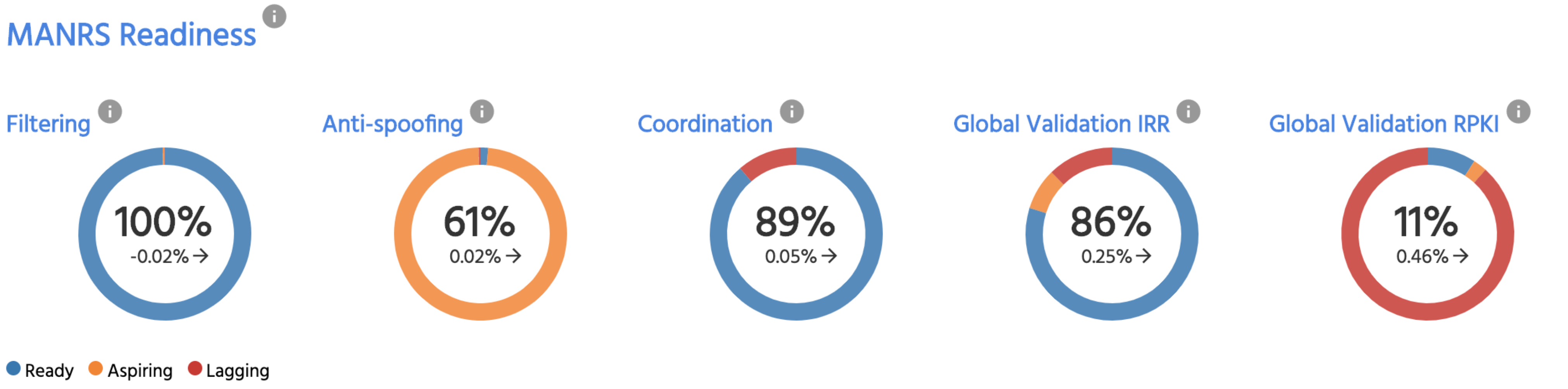
Number of ‘Connections’



46T+

Customer Edge Traffic

MANRS & OBSERVATORY



MANRS Readiness

Filtering
100%
-0.02% →

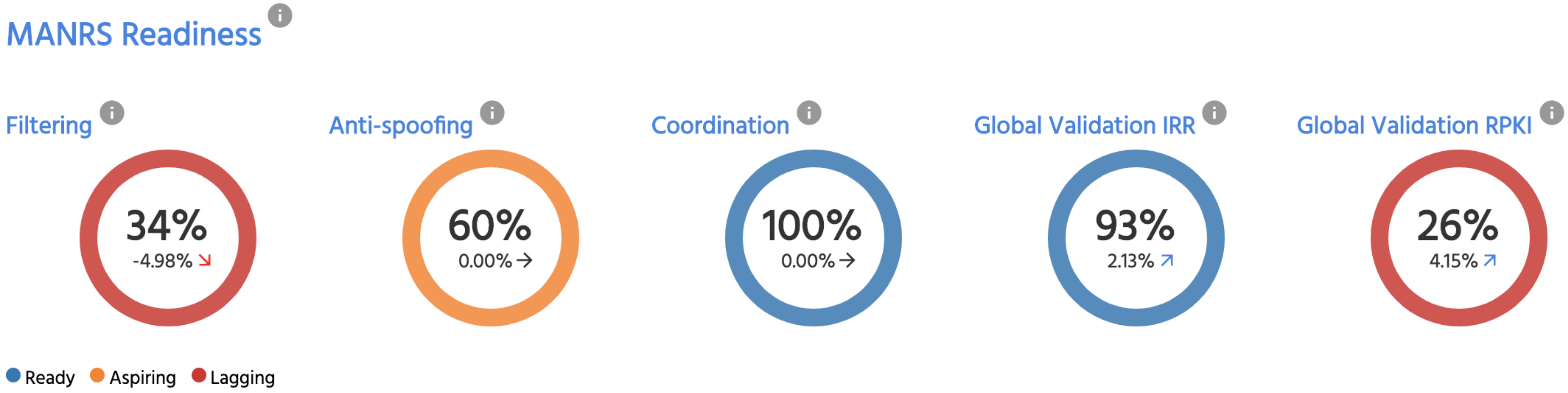
Anti-spoofing
61%
0.02% →

Coordination
89%
0.05% →

Global Validation IRR
86%
0.25% →

Global Validation RPKI
11%
0.46% →

● Ready ● Aspiring ● Lagging



MANRS Readiness

Filtering
34%
-4.98% ↘

Anti-spoofing
60%
0.00% →

Coordination
100%
0.00% →

Global Validation IRR
93%
2.13% ↗

Global Validation RPKI
26%
4.15% ↗

● Ready ● Aspiring ● Lagging

WHY AND WHAT STARTED IT?

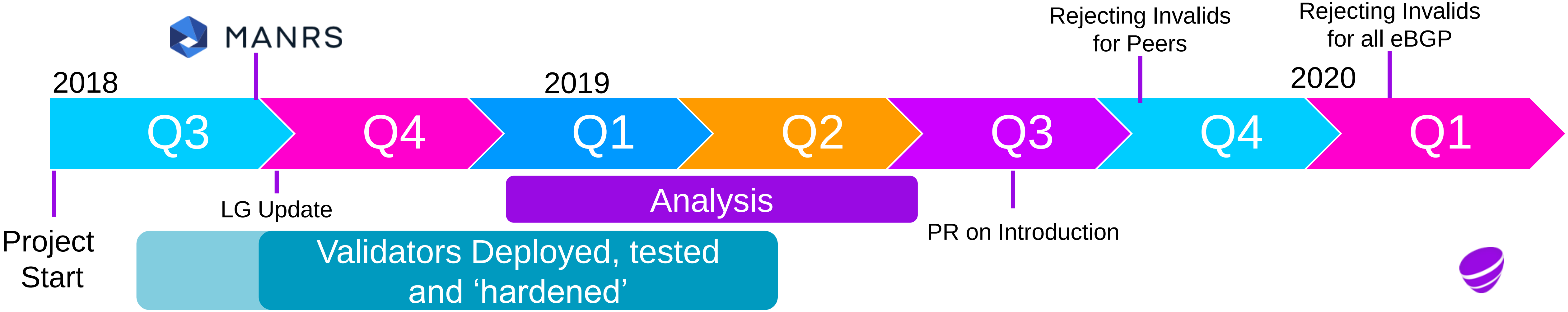
Aim : Further secure our BGP Routing. To be good custodians, not to be mentioned in post-mortem blogs

Considerations;

- Let’s not break anything
- Stability / Reliability
- Will it break anything?
  - RTBH was a concern
- How ‘bad’ are our customers/peers currently?

Plan;

- Testing
- Start with validating – no action
- Analysis
- Implement Rejects
  - Customer & Peers at same time
  - Controlled Introduction
  - Enabled by default for new connections



MANRS

2018

2019

2020

Q3 Q4 Q1 Q2 Q3 Q4 Q1

Rejecting Invalids for Peers

Rejecting Invalids for all eBGP

Project Start

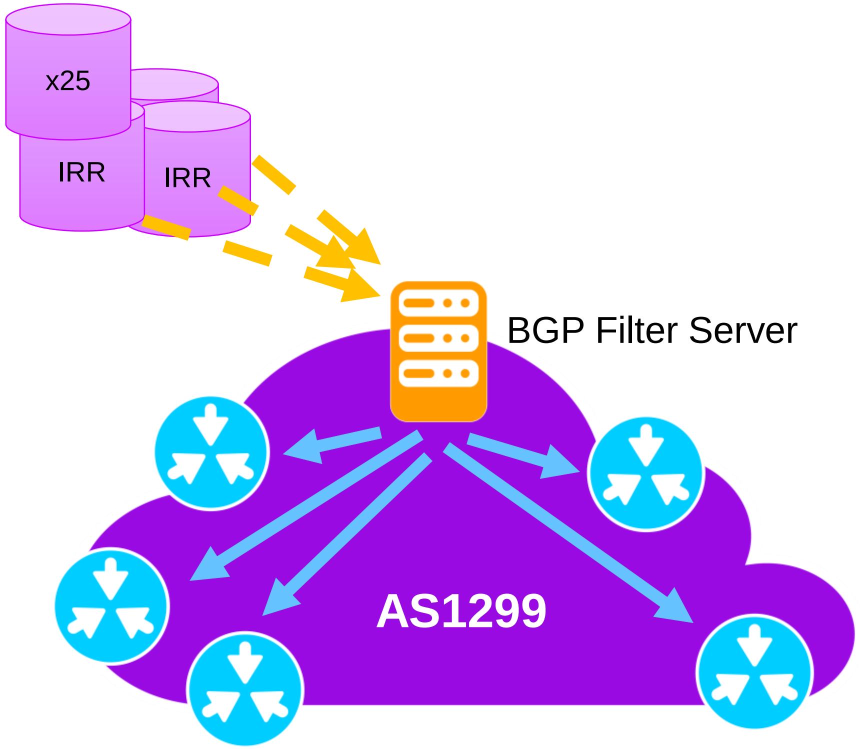
LG Update

Analysis

PR on Introduction

Validators Deployed, tested and ‘hardened’

EXISTING BGP FILTERS

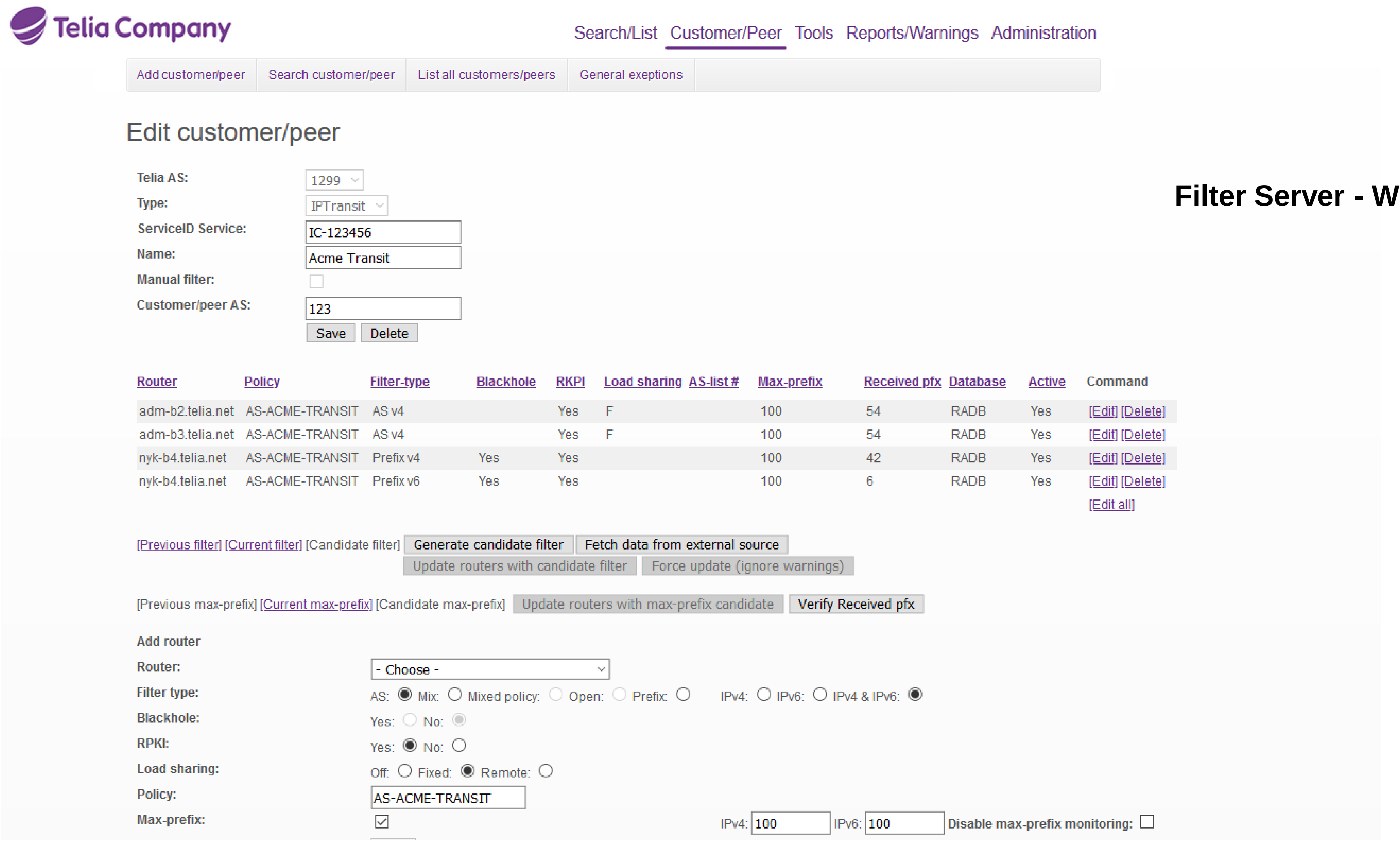


BGP Filter Server

- Born 2005
- Pulls from 25 IRR DB's and generates filters
  - RPKI ROA's also used as a source
- Pushes to all edges twice a day
- Manages all AS- and prefix-filters (v4 & v6)
  - Manipulation of entries
- Central point for:
  - Max-prefix control and monitoring
  - RTBH enabling
  - iBGP loadsharing
  - RPKI activation, exceptions

25 IRR Databases queried – in order: RADB, AFRINIC, RIPE, RIPE-NONAUTH, BELL, APNIC, NTTCOM, ALTDB, PANIX, RISQ, NESTEGG, LEVEL3, REACH, AOLTW, OPENFACE, ARIN, OTTIX, EASYNET, JPIRR, HOST, RGNET, ROGERS, BBOI, TC and CANARIE

Filter Server - Web GUI



Edit customer/peer

Telia AS:	1299
Type:	IPTransit
ServiceID Service:	IC-123456
Name:	Acme Transit
Manual filter:	□
Customer/peer AS:	123

Save Delete

Router	Policy	Filter-type	Blackhole	RKPI	Load sharing	AS-list #	Max-prefix	Received pfx	Database	Active	Command
adm-b2.telia.net	AS-ACME-TRANSIT	AS v4		Yes	F		100	54	RADB	Yes	[Edit] [Delete]
adm-b3.telia.net	AS-ACME-TRANSIT	AS v4		Yes	F		100	54	RADB	Yes	[Edit] [Delete]
nyk-b4.telia.net	AS-ACME-TRANSIT	Prefix v4	Yes	Yes			100	42	RADB	Yes	[Edit] [Delete]
nyk-b4.telia.net	AS-ACME-TRANSIT	Prefix v6	Yes	Yes			100	6	RADB	Yes	[Edit] [Delete]
											[Edit all]

[Previous filter] [Current filter] [Candidate filter] Generate candidate filter Fetch data from external source
Update routers with candidate filter Force update (ignore warnings)

[Previous max-prefix] [Current max-prefix] [Candidate max-prefix] Update routers with max-prefix candidate Verify Received pfx

Add router

Router:	- Choose -
Filter type:	AS: ● Mix: ○ Mixed policy: ○ Open: ○ Prefix: ○ IPv4: ○ IPv6: ○ IPv4 & IPv6: ●
Blackhole:	Yes: ○ No: ●
RPKI:	Yes: ● No: ○
Load sharing:	Off: ○ Fixed: ● Remote: ○
Policy:	AS-ACME-TRANSIT
Max-prefix:	☑ IPv4: 100 IPv6: 100 Disable max-prefix monitoring: □

TOPOLOGY AND TESTING

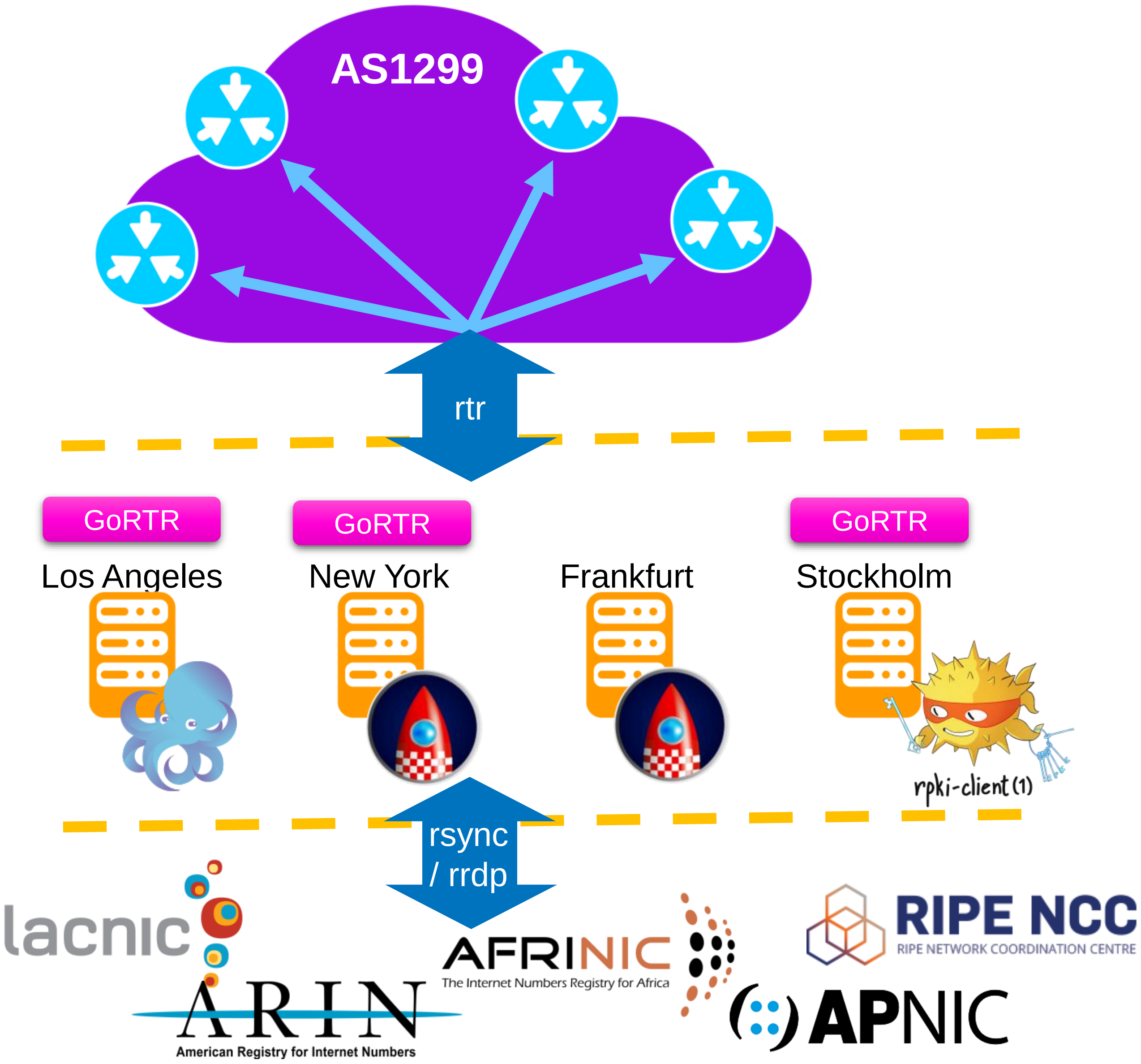
Validators used:

- Routinator
- OctoRPKI / GoRTR
- OpenBSD rpki-client
- *RIPE's RPKI Validator v3*

RIR TA update frequency:

- Routinator: every 60 min
  - Frankfurt On the hour
  - New York 15 min past the hour
- OctoRPKI / Los Angeles: every 20 min
- rpki-client / Stockholm: every 15 min

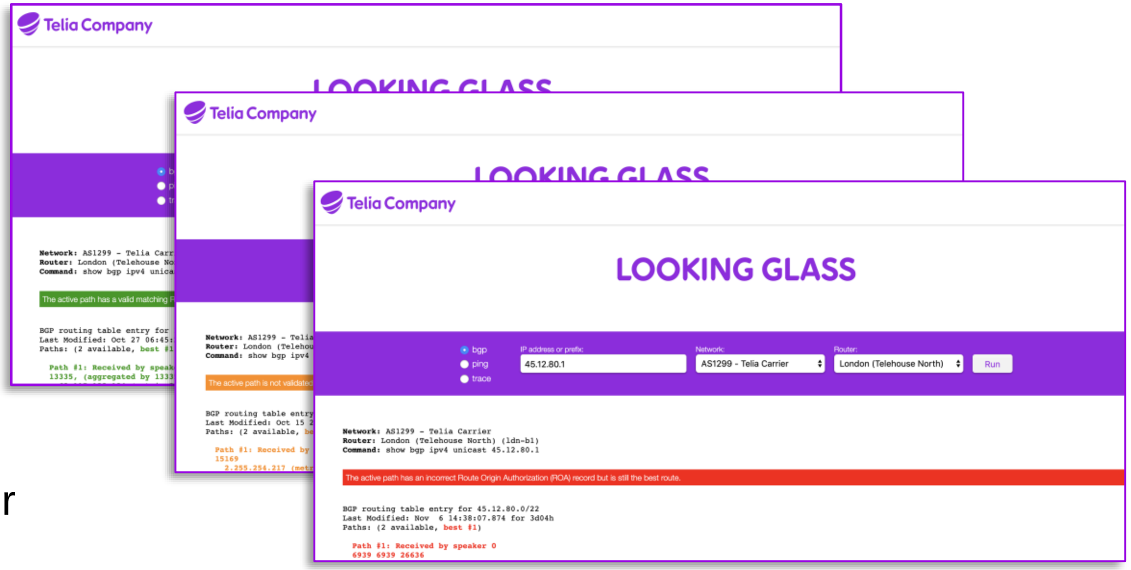
Each edge device has four validator sessions



REPORTING AND ANALYSIS

BGP COMMUNITY	RPKI STATUS
1299:430	Valid
1299:431	Unknown
1299:432	Invalid

Looking Glass updated Oct 2018 with RPKI status



```
Network: AS1299 - Telia Carr
Router: London (Telehouse No
Command: show bgp ipv4 unica
```

The active path has a valid matching R

```
BGP routing table entry for
Last Modified: Oct 27 06:45:
Paths: (2 available, best #1

  Path #1: Received by speak
  13335, (aggregated by 1333
```

```
Network: AS1299 - Telia
Router: London (Telehou
Command: show bgp ipv4
```

The active path is not validated

```
BGP routing table entry
Last Modified: Oct 15 2
Paths: (2 available, be

  Path #1: Received by
  15169
    2.255.254.217 (metr
```

● bgp
○ ping
○ trace

IP address or prefix: 45.12.80.1

Network: AS1299 - Telia Carrier

Router: London (Telehouse North)

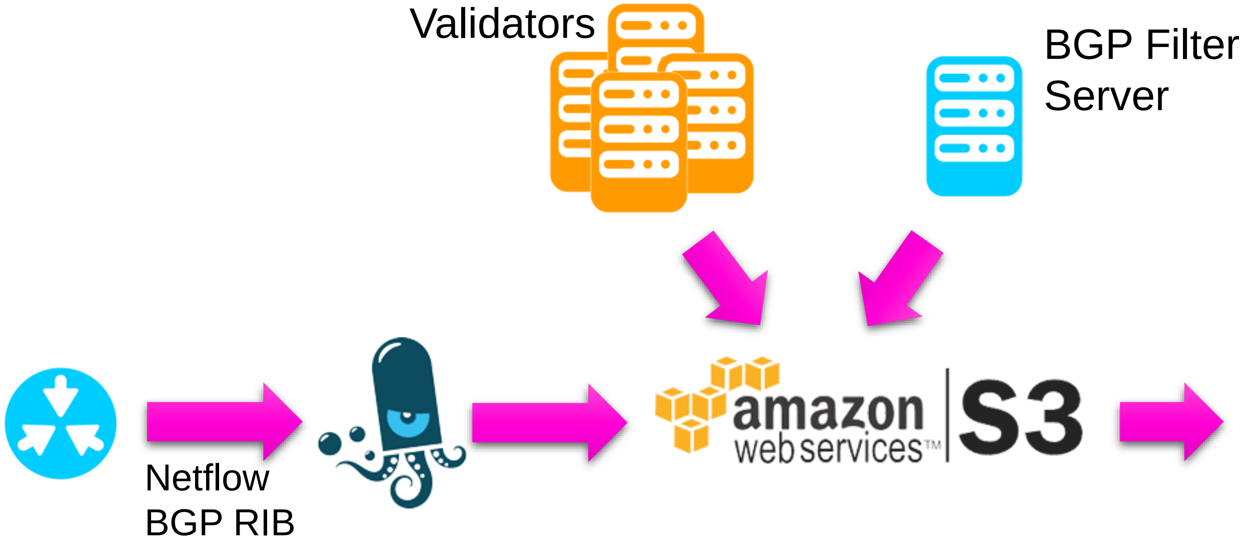
Run

```
Network: AS1299 - Telia Carrier
Router: London (Telehouse North) (ldn-b1)
Command: show bgp ipv4 unicast 45.12.80.1
```

The active path has an incorrect Route Origin Authorization (ROA) record but is still the best route.

```
BGP routing table entry for 45.12.80.0/22
Last Modified: Nov  6 14:38:07.874 for 3d04h
Paths: (2 available, best #1)

  Path #1: Received by speaker 0
  6939 6939 26636
```



Validators

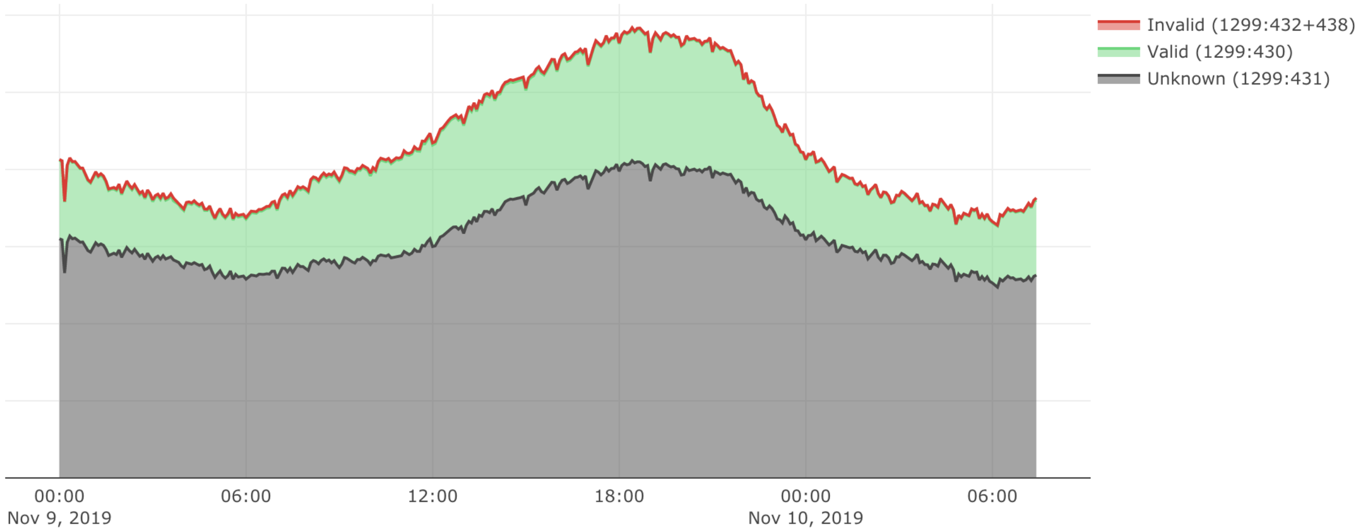
BGP Filter Server

Netflow
BGP RIB

REPORTING AND ANALYSIS

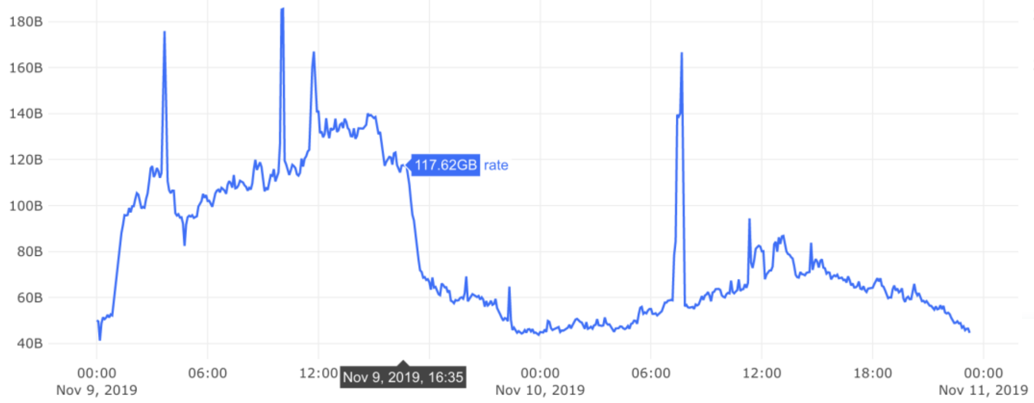
Traffic volumes by RPKI status;

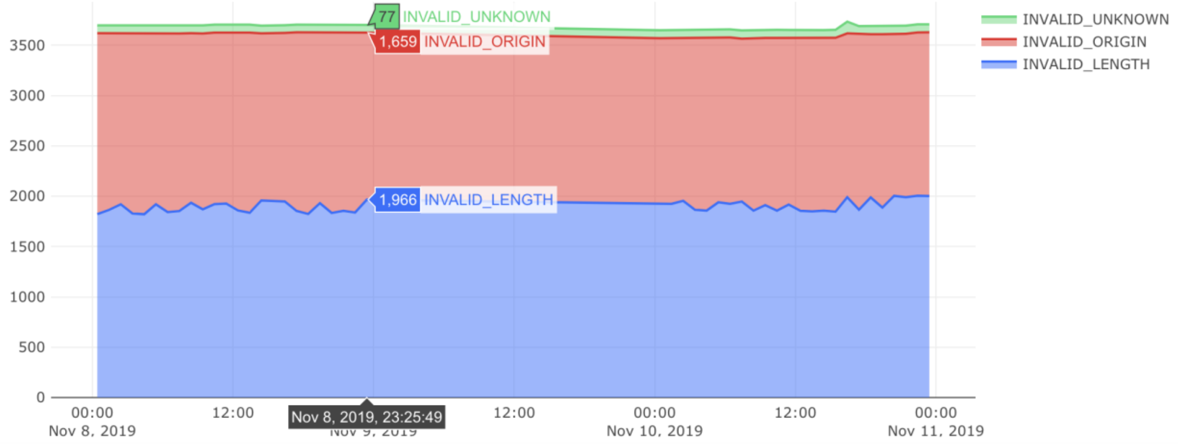
- By Adjacent AS
- By Router
- By Region
- etc.



Detailed Traffic;

- 5min data samples





Reasons for Invalid

NOTE TO OUR CUSTOMERS

Further to Telia Carrier's recent announcements regarding RPKI implementations (see: Sep 2019 Press Release and BGP Routing Security Policy), we have taken another step forward in the journey and are, as of Feb 4th 2020 Rejecting Invalids on all BGP sessions.

We take our responsibility as a Tier1 provider seriously – and this step aligns to that. While it's probably fair to say that not everyone is ready for this, and some undesired impact will occur, we felt the time is right to take this next step. This honours the work done by IP resource owners who have registered their ROAs, and further reduces the risks of malicious and accidental BGP Hijacks.

If you have an issue with BGP connectivity, and you suspect this is related to RPKI Invalids – the first step is to check the prefix in question on the Telia Carrier Looking glass here : https://lg.telia.net/.
If you get no result found - then check the ROA status using one of the publicly available RPKI Validators (eg https://rpki-validator.ripe.net/roas or https://rpki.cloudflare.com/). If the prefix is not valid then this is most likely the reason for connectivity loss.

Resource owner will need to fix the relevant ROA to ensure that the mask-length or Origin is correct. Once corrected with the relevant RIR; propagation and acceptance by AS1299 Routers can take up to an hour, but typically a lot less.

See here for more details about correcting ROAs : https://rpki.readthedocs.io/en/latest/about/help.html#what-can-i-do-about-my-route-having-an-invalid-state

Thank you for your understanding and support in this positive step for Internet Security.

Jorg Dekker
IP Services and DDoS Product Manager

CUSTOMER THOUGHTS

- A content / Mega Content / CDN customer
  - Super happy(!)
  - We’re seeing RPKI as a requirement in new RFPs
- Smaller Tier1 or Tier2-3
  - Might be concerned
  - ROAs are not in their control – could cause issues
    - Few cases reported to our CSC
- End Networks / Eyeballs
  - Should be happy
  - Work to get ROAs registered or fixed

PITFALLS

RPKI & RTBH

Blackholing is often triggered by announcing a host route (/32 or /128) with a BGP community.

Example: 1299:666

Problem:

If you have a ROA with max-length set, the blackholing request is now RPKI invalid.

Is this a problem?

RPKI & RTBH

”Solution” 1:

Don’t reject RPKI invalids for RTBH requests that are trusted.

Solution 2 (complicated):

Re-engineer how RTBH validation is done in the network.

RTBH RE-ENGINEERED

Old method:

RTBH requests are validated in router against customer specific prefix-sets.

New method:

Check if the next hop of the covering route is the same as the blackhole request.

- Cannot be done on routers today, external validation server based solution required
- (see talk from IETF 104 by Job Snijders)

PITFALLS

- RPKI Validators
  - They are memory hogs...
  - crontab?

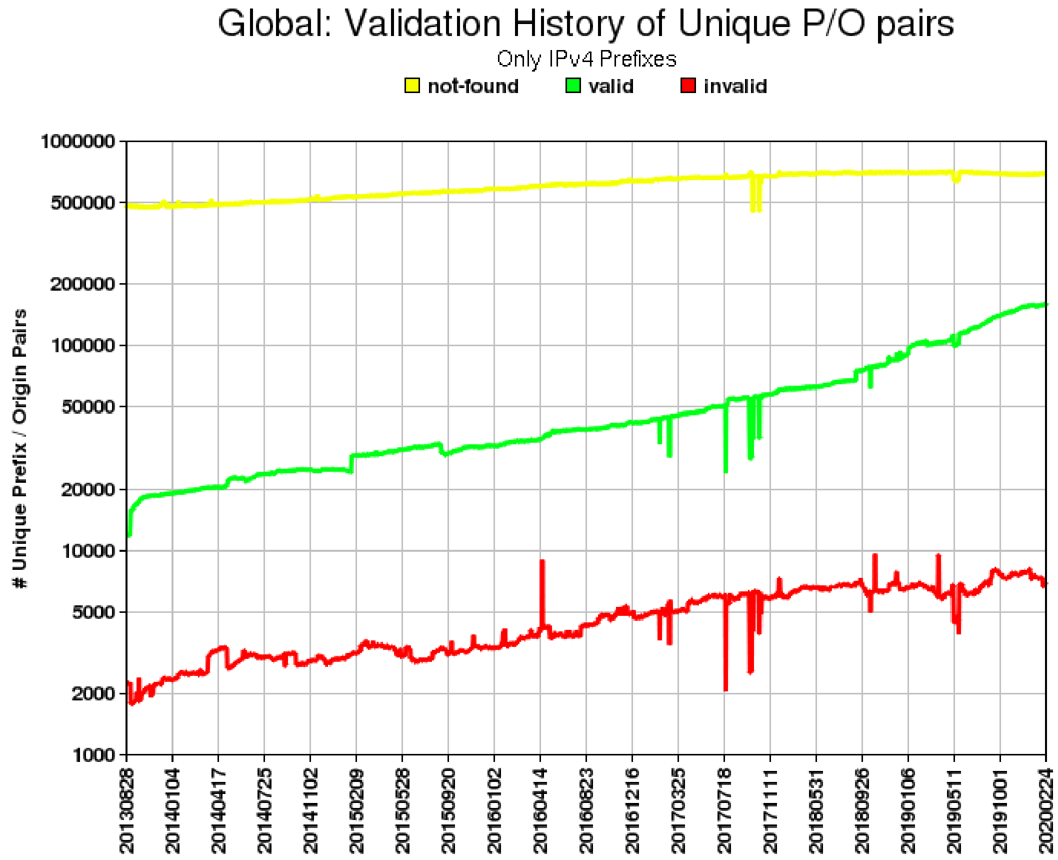
- Routers
  - Use loopback for RTR
  - Be aware of issues with route damping
  - Hidden route-refresh?
  - Weird bugs, lack of knowledge by TAC

OPERATIONAL PITFALLS

- Does your NOC have server knowledge?
  - “NetDevOps”
  - Keep your IT department far away from the servers
- Monitor RTR-feeds in routers, create alarms
- Stay away from using whitelist functionality in validators
  - Use validation exception in BGP policies instead

MORE INFORMATION

- **Good global tracker**
  - https://rpki-monitor.antd.nist.gov/
  - Shows uptake in ROA registration
  - And % of Invalid / Valid / Unknown
- **Excellent general resource**
  - https://rpki.readthedocs.io/en/latest/
  - Good one to send to customers
- **RIRs**
  - ARIN
  - RIPE
  - AFRINIC
  - APNIC
  - LACNIC



NIST RPKI Monitor 2020-02-25

OUR “SOFT” AIMS

- We want to be seen to be taking our responsibility seriously
- Leading by example
- Not rushing and breaking things, but not lagging and being complacent

Telia Carrier

THANK YOU!

Carl Fredrik Lagerfeldt
cf@telia.net