BGP security

19 april 2018

Copenhagen

Agenda

14:30 Welcome and registration

15:00 Presentation

17:00 Questions

17:30 Beer & Burgers 🍻 & 🍔

Who are we?

Lucas
Senior network engineer @ NL-ix
in ISP business
since 2013

Martin
Consultant
in ISP business
since 1995

Iljitsch
Author
(was) in ISP business
since 1995

Topics

- Some BGP history
- BGP security:
  - security of the BGP TCP sessions
  - security of prefixes in BGP
- BGP flowspec
- Use BGP for more security-related stuff
- Conclusion: what do you use when?
- The NL-ix secure routeserver

Inter-domain routing timeline

- GGP: ~1979
- EGP: 1982
- BGP 1: 1989
- BGP 2: 1990
- BGP 3: 1991
- BGP 4: 1994
- OSI IDRP: never

BGP-4

- Up to BGP-3: only class A, B, C networks in inter-domain routing
- New in BGP-4: classless inter-domain routing (CIDR)
  - where we got /-notation 🙃
- BGP-4 is very easy to extend
- So when IPv6 came along, *multiprotocol extensions* were added so IPv6 (and other protocols) can be carried in BGP-4
- Many other new features and modifications have been incorporated into BGP-4 over the years

BGP security

- **BGP security has two parts:**
  1. ***Are BGP messages exchanged between the right BGP speakers unmodified?***
- **And:**
  2. ***Are they saying the right things?***
     - 2a. ***only the owner of a prefix gets to to advertise (originate) that prefix***
     - 2a. ***the advertisement may only be distributed in accordance with the wishes of the originator of the prefix***

The right speakers

- What if someone intercepts the cable between two BGP routers and starts to impersonate router A to router B?

- What if someone sends spoofed TCP reset packets that kill the TCP session BGP runs over?

Security of the BGP TCP sessions

- Solution: RFC 2385 (1998): *Protection of BGP Sessions via the TCP MD5 Signature Option*
  - this is where the crypto guys cringe: yes, in 2018 we still use MD5, and not even as HMAC
- RFC 5925 (2010), *The TCP Authentication Option*, solves this
  - but only if you use it...
- MD5 hash calculation is light weight, so in theory 👍
- But attacker can send garbage packets, then router says 😰

GTSM

- RFC 5082 (2007): *The Generalized TTL Security Mechanism*
  - normally, the TTL field in BGP packets = 1
    - to detect extra hops in the BGP path
    - but attacker 2 hops away sets TTL = 3, receiver sees 1
  - GTSM: sender sets TTL = 255, receiver checks TTL == 255
    - attacker 2 hops away can't set TTL = 257
    - so protects agains spoofed packets one or more hops away
  - no crypto, so router CPU says 👍

Security of prefixes in BGP

- The Youtube/Pakistan incident (2008)
- China Telecom (2010)
- The Enzu/Spotifiy route leak (2015)
- Unused Russian AS (2017)

- The list goes on...

Security of prefixes in BGP (2)

- Filter your customers. <u>Always</u>. ☠️
- Can't filter transit providers, they send you all prefixes 😎
- Filtering peers:
  - manually: unworkable, too many new prefixes without notice 😥
  - based on routing registry: 😀 → 😕 → 😀
  - based on RPKI: 🤔
- Further security: BGPsec: 🤯

History of BGP security

- Around 2000, BBN (builders or the original ARPANET!) proposed S-BGP (Secure BGP)
  - S-BGP adds certificates to BGP and signs every update
- Shortly after that, Cisco proposed soBGP (Secure Origin BGP)
  - soBGP provides similar security with less overhead
- In 2003, the "National Strategy to Secure Cyberspace" identified IP(v4), the DNS and BGP as key protocols that needed "security and resilience" improvements

History of BGP security

So lots of fun when I (Iljitsch) went to the an IETF meeting for the first time in Atlanta in 2002!

Securing information in BGP – how?

- First:
  - only the owner of a prefix gets to to advertise (originate) that prefix
- Second:
  - the advertisement may only be distributed in accordance with the wishes of the originator of the prefix

S-BGP → RPKI

- Making sure only the owner (holder) of a prefix can originate a prefix:
  - Resource Public Key Infrastructure
    - RFC 6480 (2012)
  - these resources are IP addresses and AS numbers
  - RPKI is not part of BGP, it works "out of band"

RPKI

- RFC 3779 (2004) adds extensions to X.509 certificates for IPv4 and IPv6 prefixes and AS numbers
- RIRs (= the RIPE NCC in Europe) give out these certificates
  - certificates contain no identifiable information
  - the RIRs use a self-signed root certificate
- The address space holder generates a Route Origination Authorization (ROA)
  - contains expiry date and maximum prefix length
  - signed with address certificate's private key

RPKI (2)

- ROAs are uploaded to public repositories
- Everyone who uses RPKI downloads all the certs and ROAs
- Server crunches all the signatures and generates a filter based on the ROAs
- Routers can then download a copy of this filter
  - so the filter isn't part of the router's configuration

Routing table != bank account

- Your bank says:

  Today we operate at 90% capacity. So if you transfer € 1000,- the receiver will get € 900,-.

- You say:

  That sucks, I'll wait until you're at 100%.

- Your ISP says:

  Today we operate at 90% capacity. So if you transfer 1000 packets, the receiver will get 900 of them.

- You say:

  That sucks, but it's better than nothing so I'll take it.

Connectivity trumps security?

- In other words:
  - we'll take *secure* connectivity over *insecure* connectivity
  - but we'll take *insecure* connectivity over *no* connectivity

- So what's the appropriate action when RPKI doesn't validate?
  - filter?
  - lower local pref?

RPKI on the router

```
router bgp 65000
 address-family ipv4 unicast
 neighbor 10.0.102.1 route-map rtmap-PEX1-3 in
 bgp bestpath prefix-validate allow-invalid
!
route-map rtmap-PEX1-3 permit 10
 match rpki invalid
 set local-preference 50
!
route-map rtmap-PEX1-3 permit 20
 match rpki not-found
 set local-preference 100
!
route-map rtmap-PEX1-3 permit 30
 match rpki valid
 set local-preference 200
!
route-map rtmap-PEX1-3 permit 40
```

https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_bgp/command/irg-cr-book/bgp-m1.html#wp3677719851

BGPsec

- So RKPI solves:
  - only the owner of a prefix gets to to advertise (originate) a prefix
- But (path to) the origin AS can still be faked
- Which is solved by BGPsec:
  - an advertisement may only be distributed in accordance with the wishes of the originator of the prefix

BGPsec (2)

- RFC 8205 (2017)
- Very similar to S-BGP
  - but with the parts covered by RPKI removed
- BGPsec capability is negotiated between BGP routers
- BGPsec only works when there is an unbroken path of BGPsec-capable routers from the origin
- Can be deployed incrementally
  - but not guidance on how, exactly...

BGPsec (3)

- BGPsec_Path attribute replaces the AS_PATH attribute
  - includes the AS number of the router a BGP update is sent to
  - and a cryptographic signature
  - with "Subject Key Identifier" pointing to the RPKI certificate of the router that created the signature
- Receivers of an update check the BGPsec_Path and the signatures

BGPsec performance

- Must be a separate update for every prefix
- Must be a separate update for every neighbor
- Updates get larger due to signatures
- S-BGP performance and deployment paper (2000):
  - CPU busy with signing/checking updates: 140 minutes per day
- 750k prefixes with 4 hops each: 3 million signatures to check after establishing a BGP session
  - ECDSA 256 bit signature check ~= 4 msec (Intel i5 @ 2.5 GHz)
  - 3 hours to process the IPv4 BGP table at startup! 😱 😱 😱

Peering LAN

- After all this theory, a story from the trenches!
- In 2003, the AMS-IX needed to go from a /24 to a /23 for the peering LAN
- But: someone typed <prefix>/24 rather than <prefix>/23
- And advertised <prefix>/24 to their peers
- The peers started sending the BGP packets not to their neighbors directly, but through the /24
- BGP neighbors were no longer directly connected, so sessions went down

Peering LAN (2)

- This all happened during a RIPE meeting, lots of fun! 😅
- Moral of the story:
  - never accept any BGP advertisements of the peering LAN prefixes for the internet exchanges you're connected to
- Then, in 2014, they needed to go from a /22 to a /21
- Guess what happened...

Conclusion: what do you use when?

- Use GTSM when you can!
- Prefer TCP-AO over the MD5 password!
- Unfortunately, only work if both ends enable them manually 😟
- In general, MD5 password is a good idea
  - but if there aren't any good filters, attackers can send fake BGP packets with invalid MD5 checksums
  - this is a denial-of-service on your router's CPU
  - if that's a risk and the BGP session doesn't have many prefixes anyway and/or with BGP graceful restart, maybe skip the MD5

The Interconnect Exchange

BGP flowspec

- What is it?
  - RFC 5575 (2009) *Dissemination of Flow Specification Rules*
  - RFC 7674 (2015) *Clarification of the Flowspec Redirect Extended Community*
- A firewall filter distributed via BGP in real time

BGP flowspec (2)

- What can you filter on?
  - IPv4 or IPv6 destination address
  - IPv4 or IPv6 source address
  - IPv4 protocol or IPv6 last next header
  - TCP/UDP source or destination port
  - ICMP type/code
  - TCP flags
  - packet length
  - DiffServ code point
  - fragmentation bits

BGP flowspec (3)

- What can you do with it?
  - advanced remote triggered black hole
  - drop trafic
  - shape trafic
  - redirect-VRF
  - redirect-marking (DiffServ)
  - redirect IP next hop
  - all on specific sources, destinations, ports

BGP flowspec (4)

- Why don't we speak it between us?
  - Several large ISPs use it internally, so why can't we cry for help?
  - RFCs defined what and how it works not how you restrict it and control what you signal your neighbors

BGP flowspec (5)

- Use case: Arbor Networks
  - Border routers looks at traffic and send samples to Peakflow
  - Peakflow looks at flow data and looks for anormalies
  - Peakflow injects flowspec prefixes with actions back into routers

Use BGP for more security-related stuff: mail filtering

- Traditional ways to filter mail:
  - download lists
  - DNS lookups
- Realtime distribution via BGP:
  - good and bad guys is added to central route servers
  - whitelists and black lists with communities
  - users peer with routeservers and update firewall filters

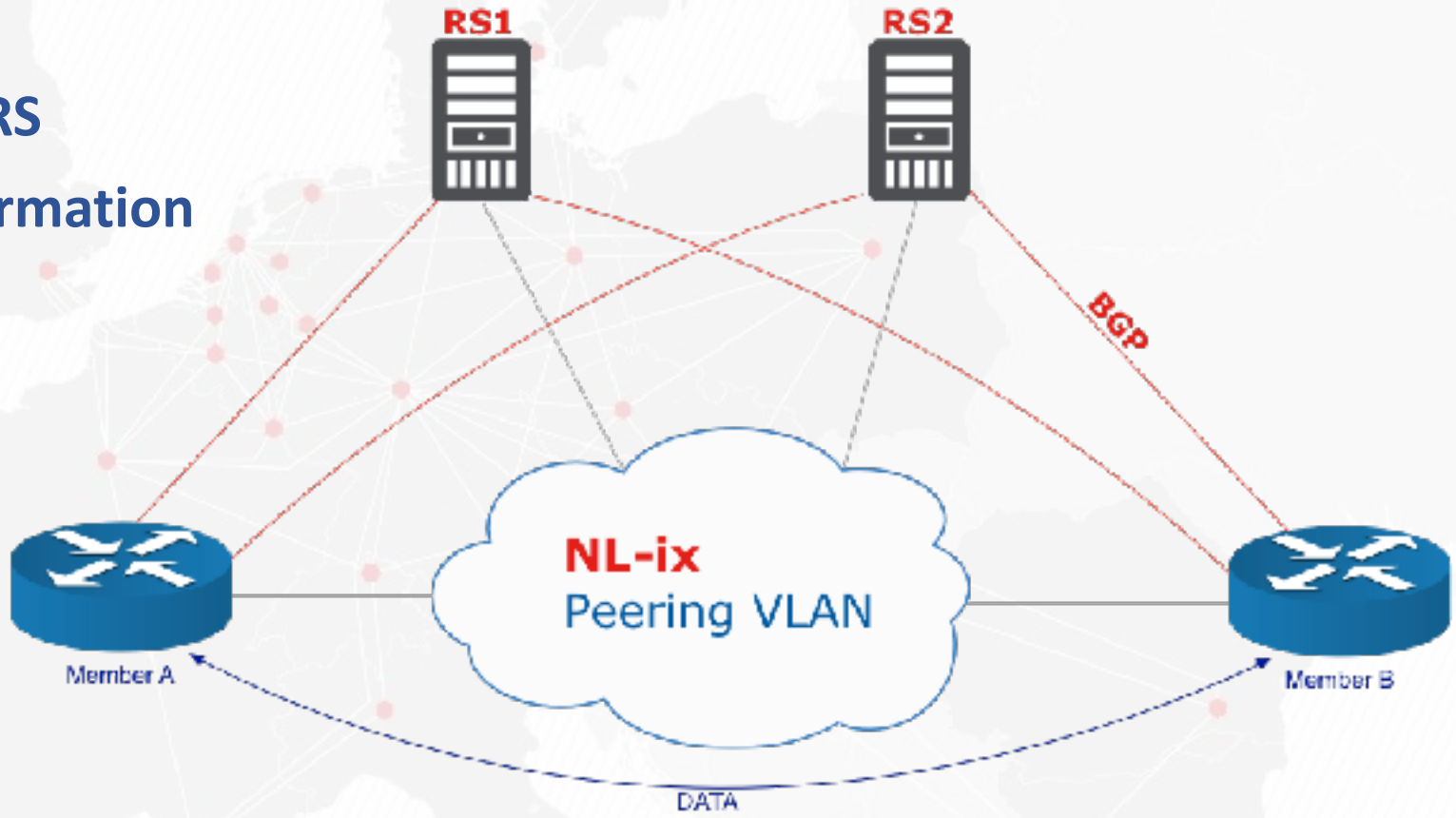
Use BGP for more security-related stuff: firewalling

- Run BGP on your average server:
  - open/close firewall ports
  - redirecting traffic
  - shaping
- Use routeservers to distribute prefixes
  - use communities to signal actions
  - use route alternative route tables to keep prefixes locally
- OpenBSD can do this natively, on other systems you will need some script magic

NLix
The Interconnect Exchange

What are route servers?

- Not a router
- No data going through the RS
- Used to aggregate BGP-information



Main benefits

- Less BGP-sessions to configure
- Quick and easy way to get lots (50k+) prefixes
- Less time spent making peering arrangements
- Automatically filter prefixes (RPKI and ROA)

So you're a peer

- You're on the RS
- Someone makes a boo-boo

```
lucasb@jointtransit-nikhef> show route receive-protocol bgp 213.207.9.124

inet.0: 702211 destinations, 6297272 routes (699195 active, 34 holddown, 680690 hidden)
  Prefix   Nexthop      MED     Lclpref    AS path
* 8.8.8.0/24          213.207.9.124                    20562 I
```

Preventing route leaks

- By default the route server filters on:
  - RFC-bogons
  - Fullbogons
  - IRRDB
  - ROA (RPKI)
  - DROP-List (Spamhaus)

BOGON prefixes

- RFC-bogons
  - Bogons are defined as martians and netblocks that have not been allocated to a regional internet registry (RIR)
    - martians: addresses set aside for special uses in RFCs / by IANA
- Fullbogons
  - Fullbogons are a larger set which also includes IP space that has been allocated to an RIR, but not assigned by that RIR to an actual ISP or other end-user
  - This provides a much more granular and enumerative view of IP space that should not appear on the Internet
- Updated every day!

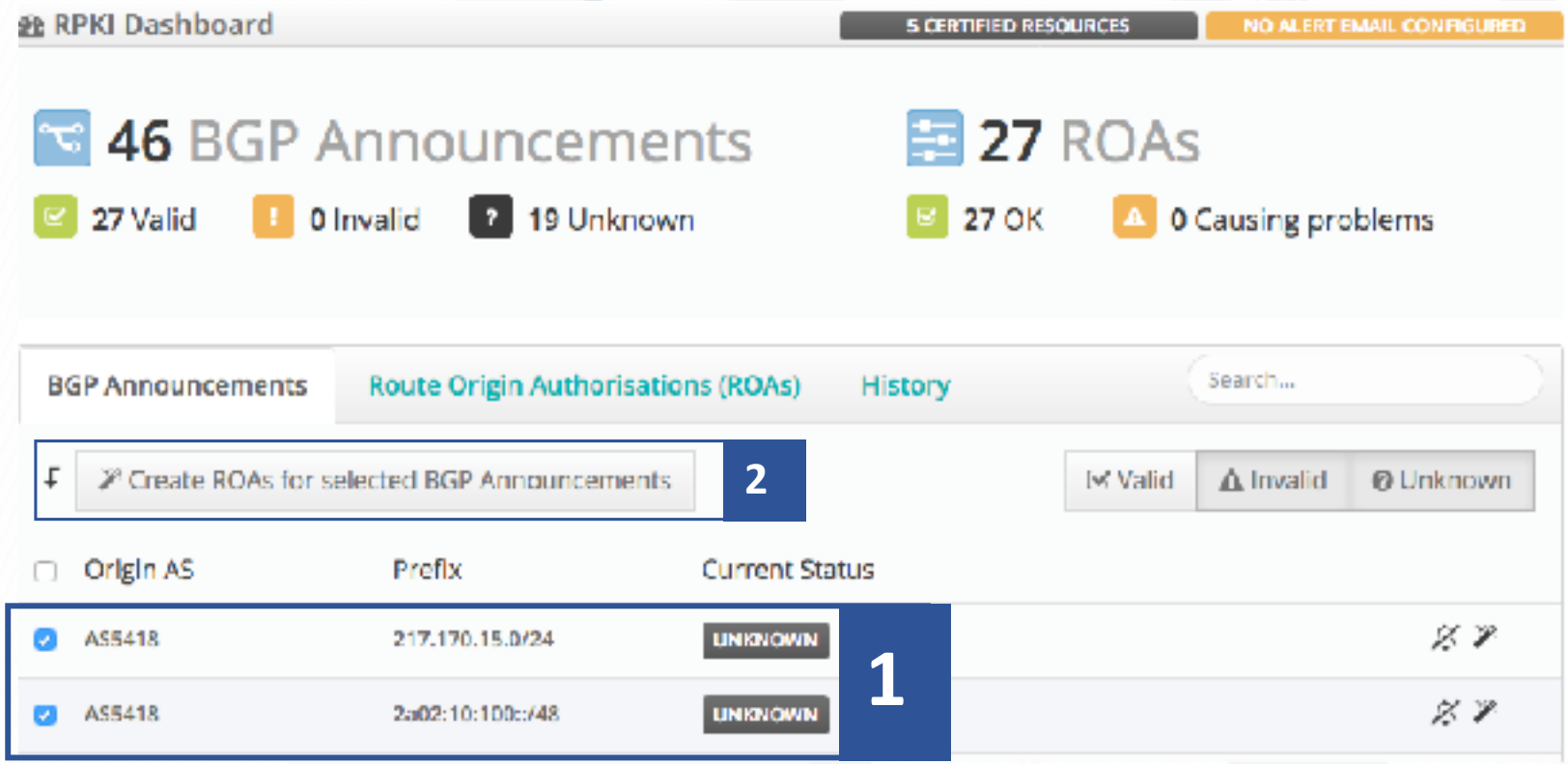
IRRDB and ROA

- IRRDB
  - AS-SET and ROUTE-OBJECTS
  - ASN and prefix validation

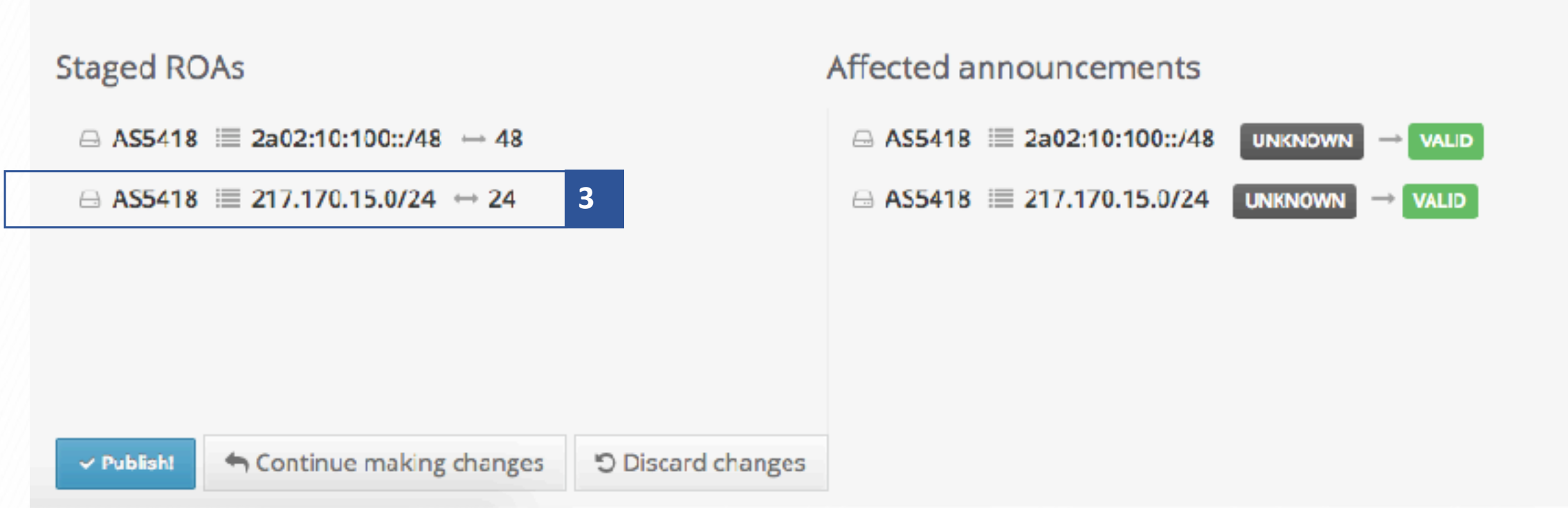
- Resource Public Key Infrastructure (RPKI)
  - It allows us to verify whether an AS is authorized to announce a specific prefix.
  - The main building blocks in the RPKI infrastructure are trust-anchors, ROA's and validators.
    - Trust-anchors used today are the RIR's (RIPE, APNIC, etc)
  - ROA is made by the maintainer of the AS.

So how do I (Lucas) do it?

- **At the RIPE RPKI-Dashboard!**



So how do I (Lucas) do it?



So how do I (Lucas) do it?

46 BGP Announcements

27 Valid | 0 Invalid | 19 Unknown

27 ROAs

27 OK | 0 Causing problems

BGP Announcements | Route Origin Authorisations (ROAs) | History | Search...

Discard Changes | Delete ROAs | Causing Problems | Not Causing Problems | + New ROA

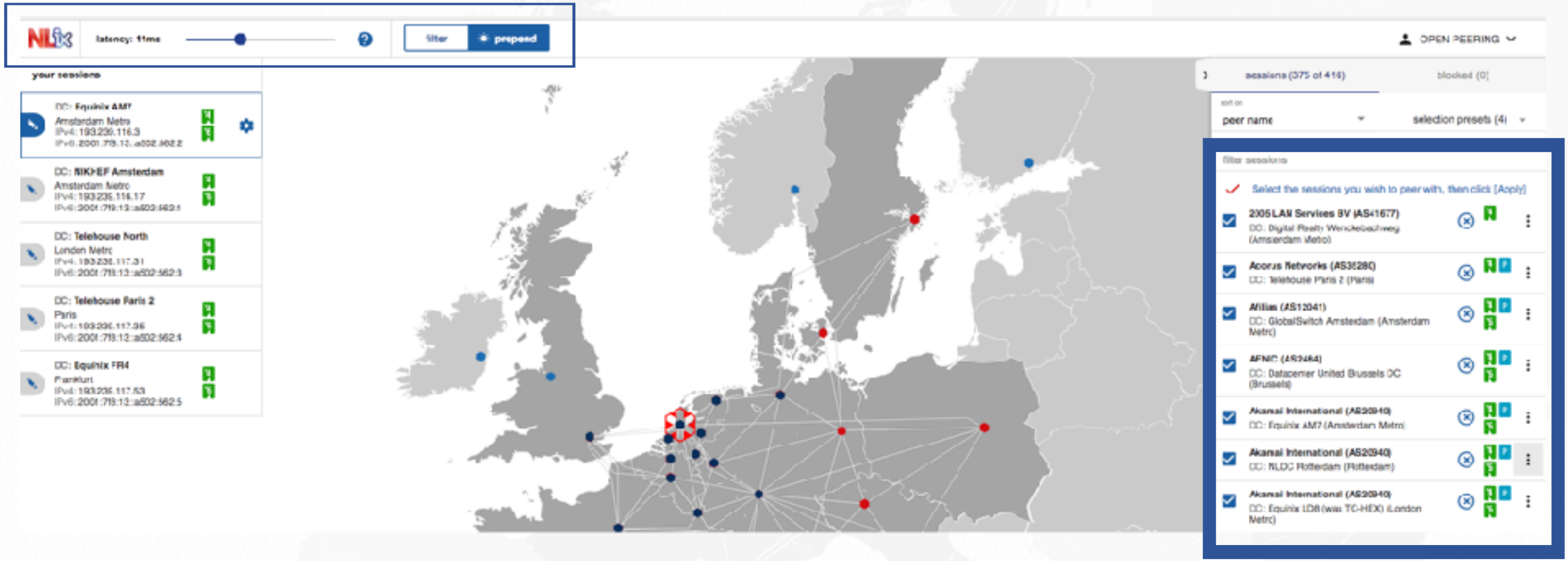
AS number	Prefix	Most specific length allowed	Affects
AS Number	Prefix	Max length	
AS20562	82.150.151.0/24	24	1
AS24785	213.207.3.0/24	24	1
AS24785	213.207.16.0/24	24	1
AS24785	213.207.0.0/24	24	1

DROP LIST

- From Spamhaus
- Advisory "drop all traffic" lists:
  - Consisting of netblocks that are "hijacked”
  - Used by professional spam or cyber-crime operations and botnets

But wait.. There’s more!

Latency based



So how do I do it?

available on route server	yes / no
show available capacity	yes / no
port capacity	10 Gb/s
maximum load	90 %

filtering

RFC Bogons	always activated
Fullbogons	(recommended) powered by Team Cymru
IRRDB	(recommended)
ROA	(recommended)
DROP List	powered by Spamhaus

CANCEL APPLY

How do I (you) get started?

- Sign up to be an NL-ix customer! 😇
  - *take advantage of the DK 1G and 10G promo!* 🤑
- You receive your login credentials for the My NL-ix Portal 🤖
- Log in to the dashboard 🤩
- Click on the right buttons! 😘

The End

- Questions?
- We'll be here afterwards to answer more questions one-on-one!