

# **BGP Routing Security**

Training Course

RIPE NCC Learning & Development

# RIPE NCC Training Material



Please find your training material at:

https://www.ripe.net/training-material



# Schedule





09:00 - 09:30	Coffee and Tea
11:00 - 11:15	Break
13:00 - 14:00	Lunch
15:30 - 15:45	Break
17:30	End

## Introductions



- Names
- Experiences with BGP and routing security
- Goals



#### Overview



- The need for BGP Security
- Analyse BGP Threats and Attacks
- BGP Security Measures:
  - How to mitigate BGP threats
  - Protection of BGP sessions
  - LAB 1 Securing BGP Sessions
  - Implementing Route Filtering
  - LAB 2 Creating BGP Prefix Filters

- LAB 3 Filtering AS-Path/number of prefixes
- Registering in the IRR System
- LAB 4 Creating route(6) objects
- Implementing RPKI
- LAB 5.1 Creating ROAs
- LAB 5.2 BGP Origin Validation
- Next steps for BGP Security
- Best practices



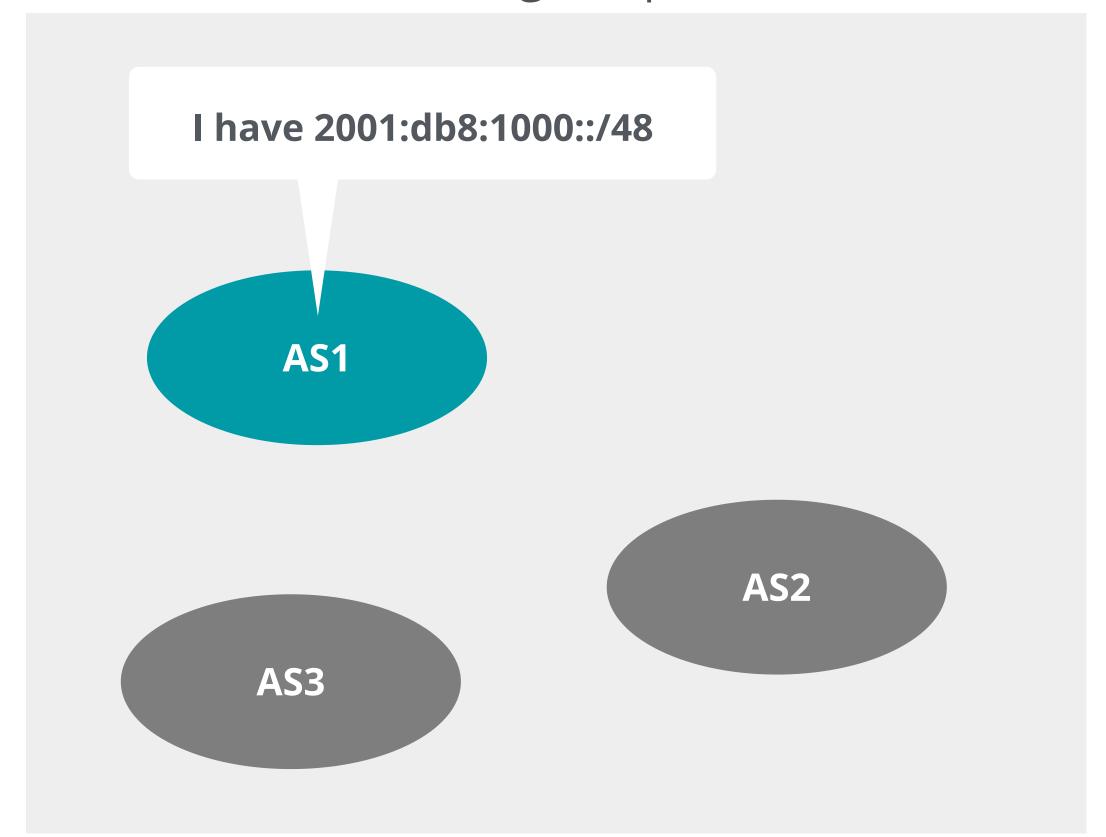
# The Need for BGP Security

Section 1



#### In theory:

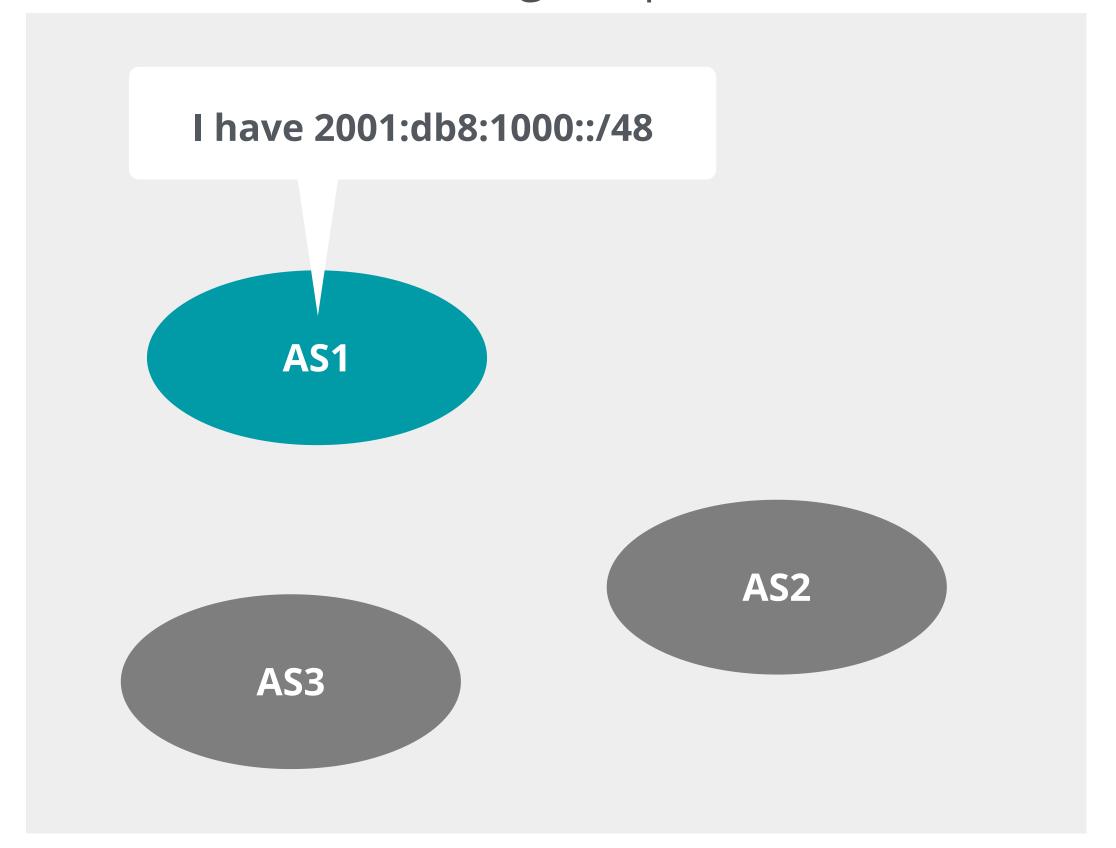
Only the legitimate resource holder should be announcing the prefix





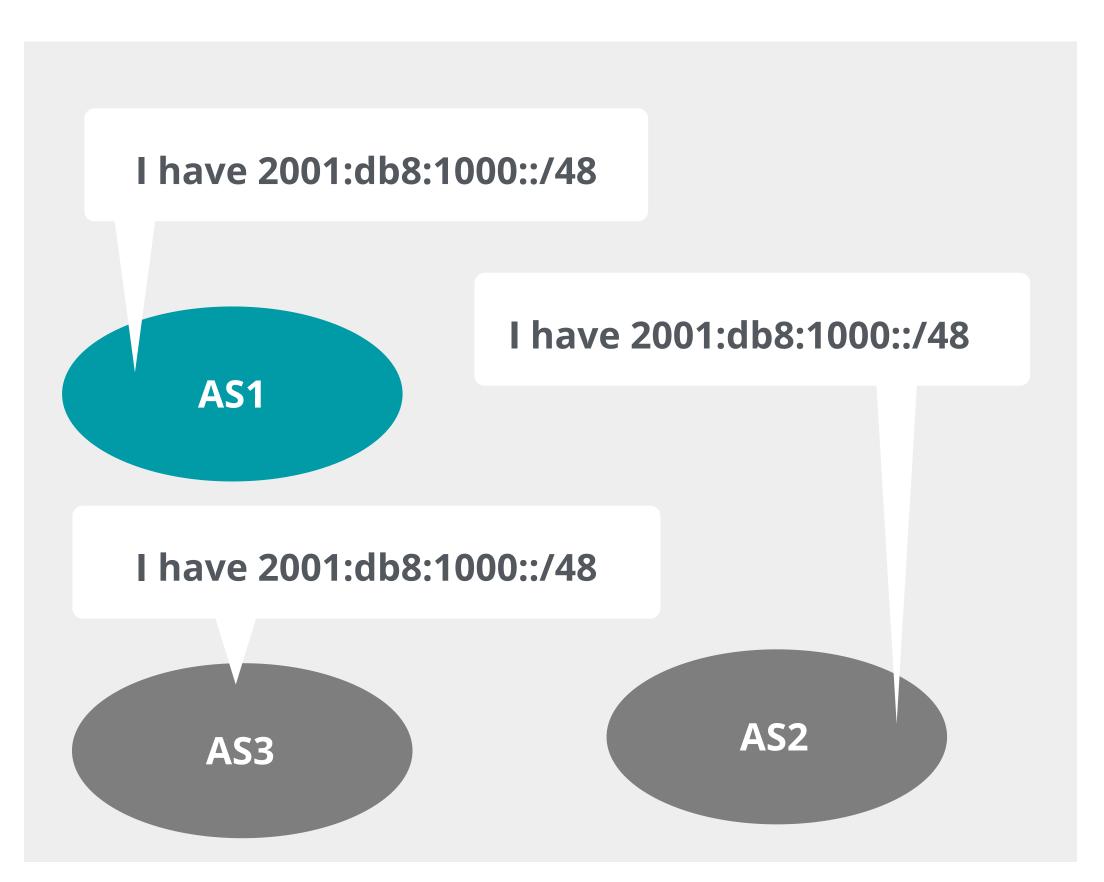
#### In theory:

Only the legitimate resource holder should be announcing the prefix



#### In practice:

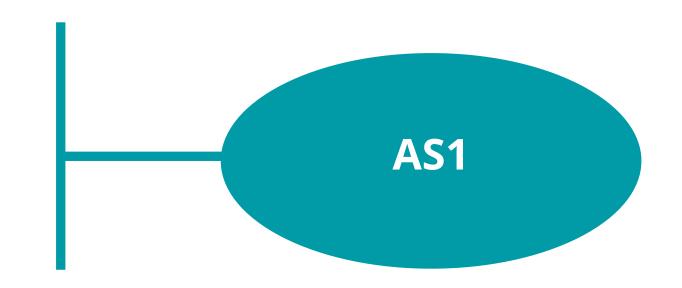
Any AS can announce any prefix!

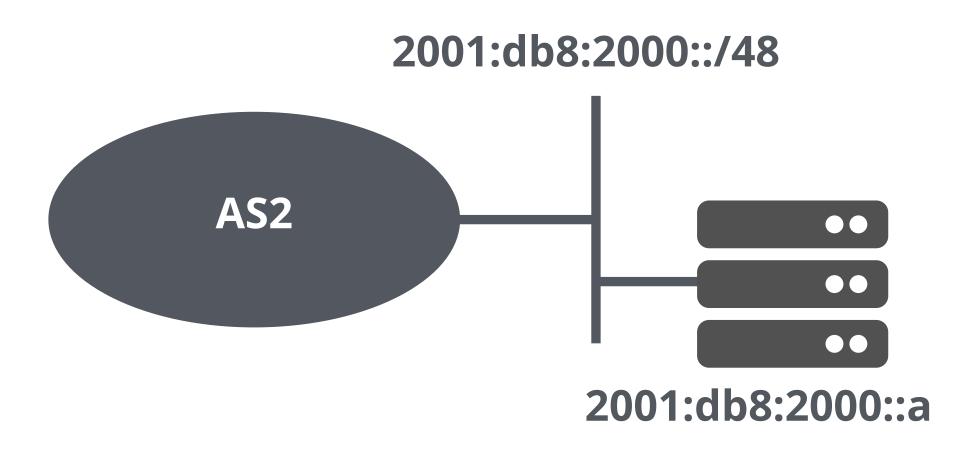




AS1 wants to access the server in AS2.

#### 2001:db8:1000::/48

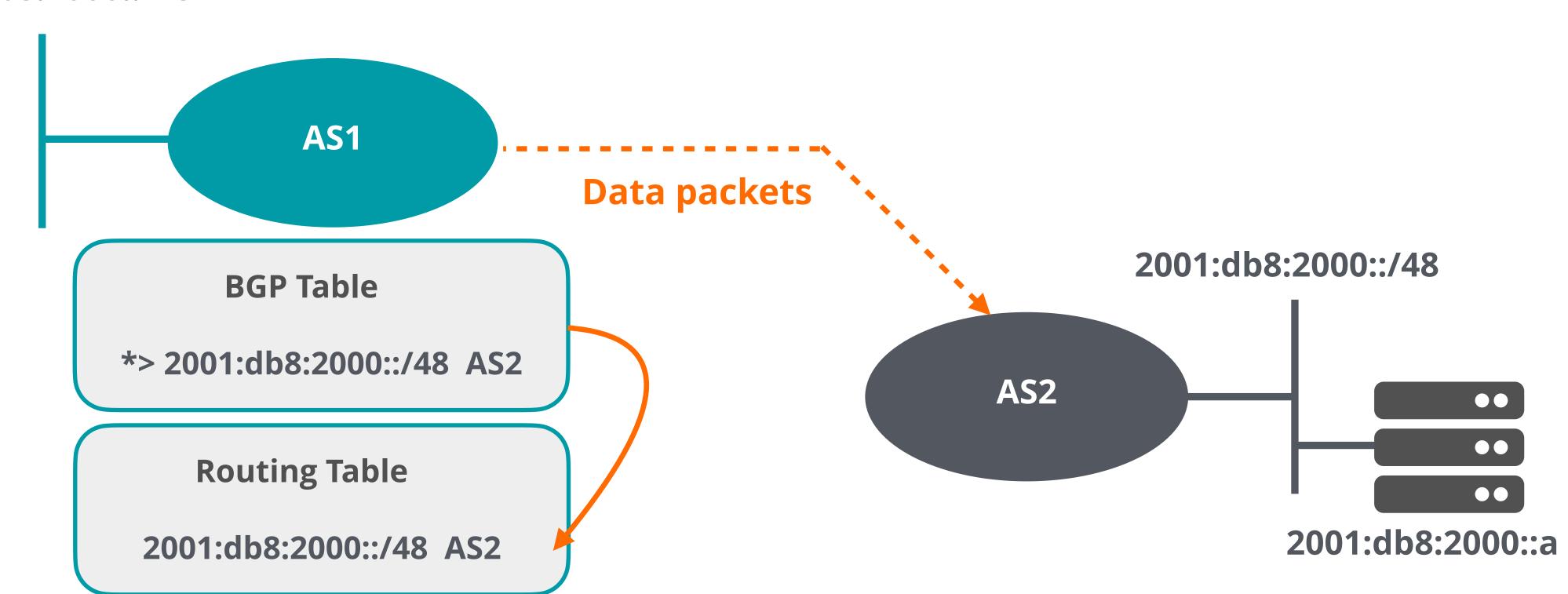






Data is forwarded based on routing table.

#### 2001:db8:1000::/48

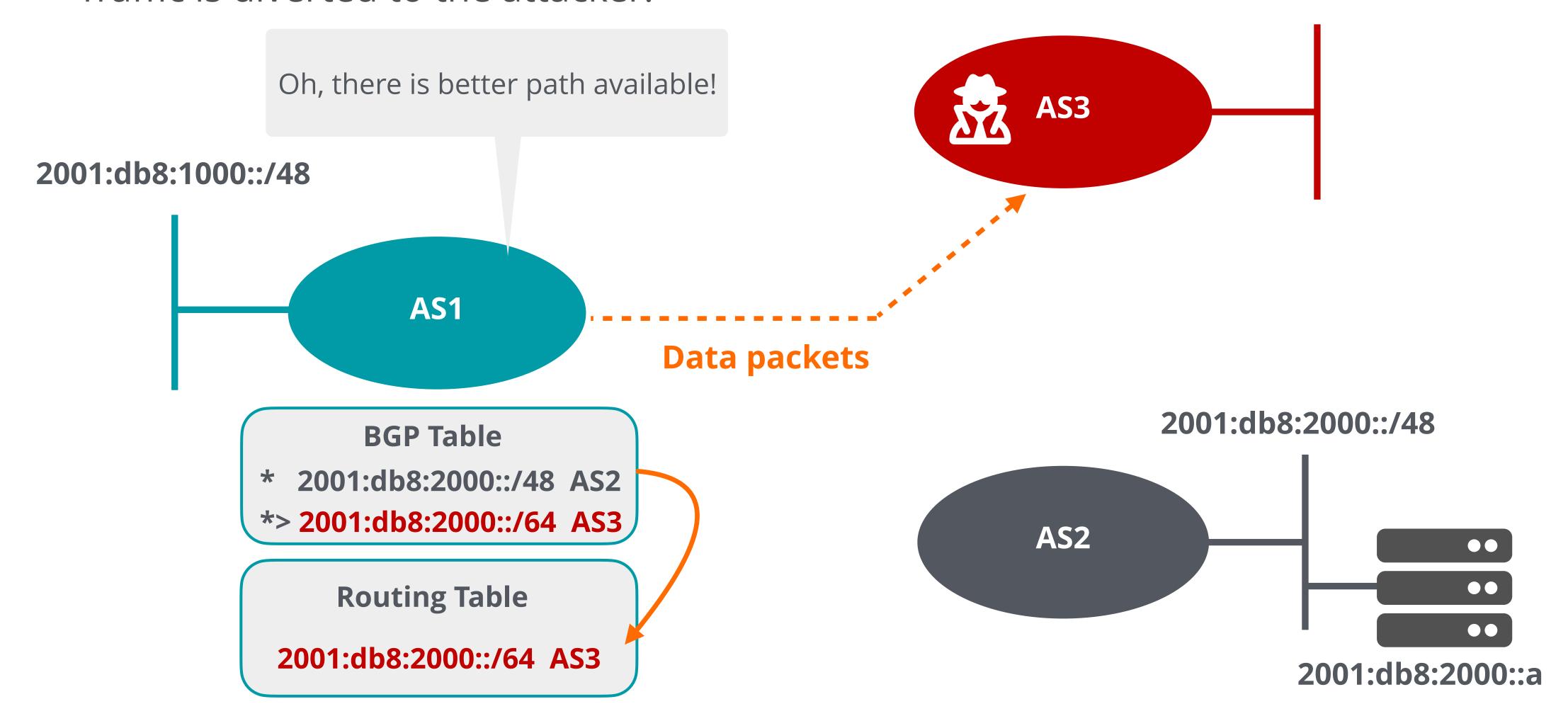




I have a more specific route! In case of a more specific announcement: **BGP Update** 2001:db8:2000::/64, AS3 2001:db8:1000::/48 AS1 2001:db8:2000::/48 **BGP Table** \* 2001:db8:2000::/48 AS2 AS2 \*> 2001:db8:2000::/64 AS3 2001:db8:2000::a

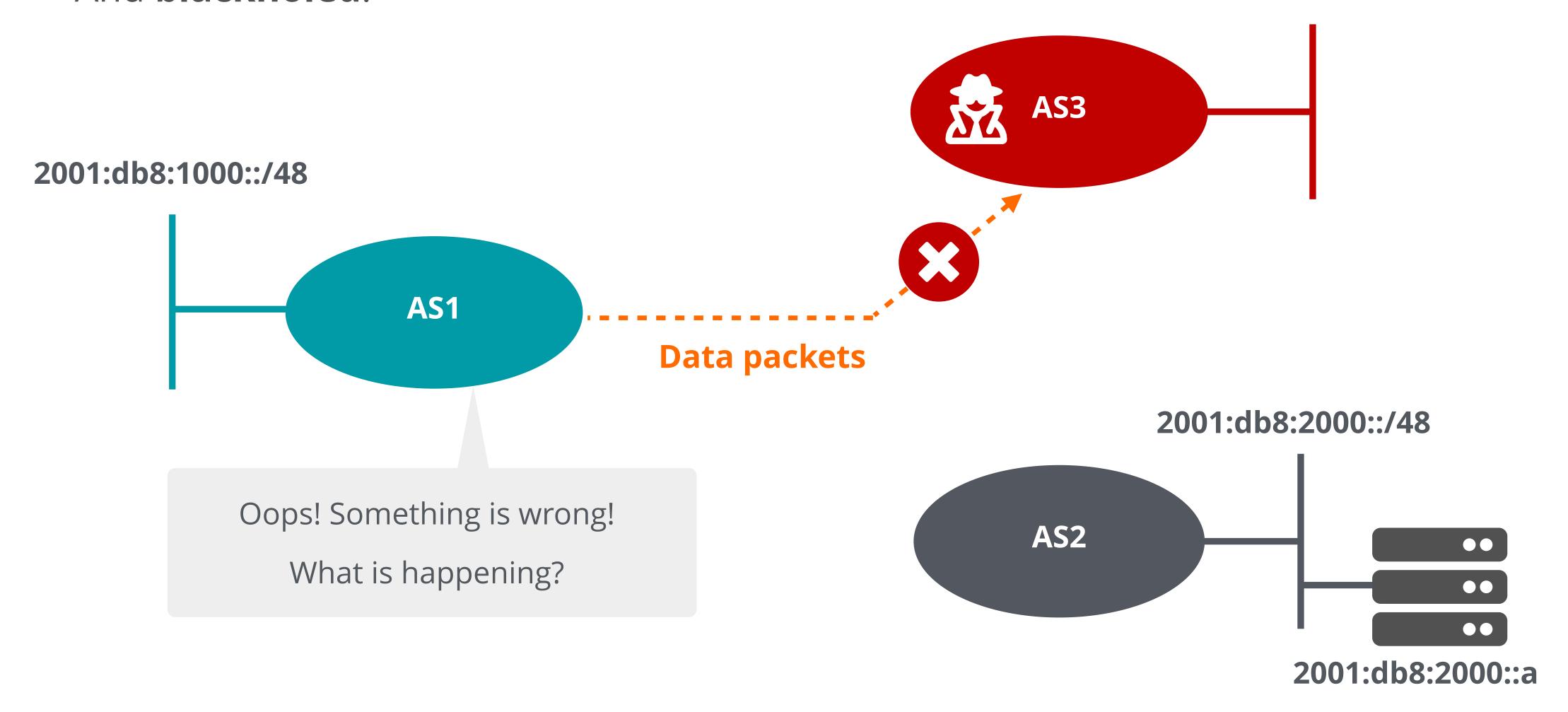


Traffic is diverted to the attacker:





#### And **blackholed**!





# Questions (2)



# Analyse BGP Threats and Attacks

Section 2



# Vulnerabilities of the BGP Protocol

Section 2.1

# **BGP Has Some Challenges**



- It is only based on trust, no built-in security
- No verification of the correctness of prefixes or AS paths



#### **BGP Has Three Main Vulnerabilities**



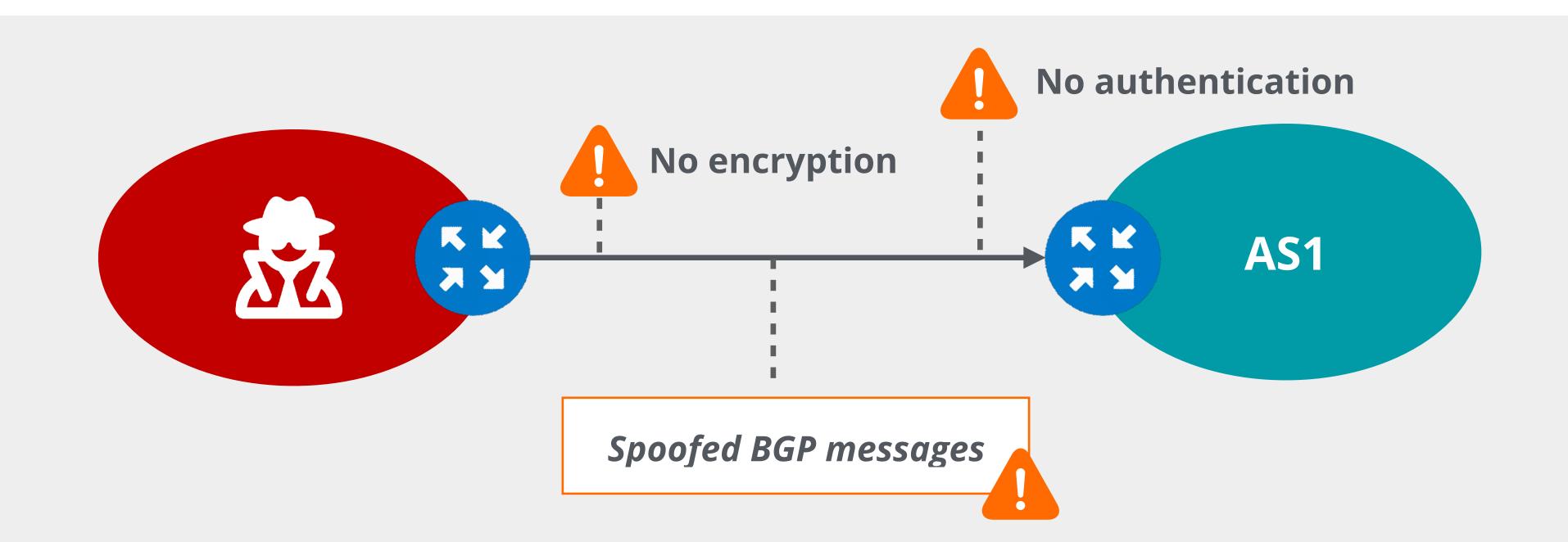
- No internal mechanism to protect the integrity and source authenticity of BGP messages, and no confidentiality
- No mechanism specified to validate the authority of an AS to announce a prefix

No mechanism to verify the authenticity of the attributes in a BGP update message

# No Encryption or Authentication



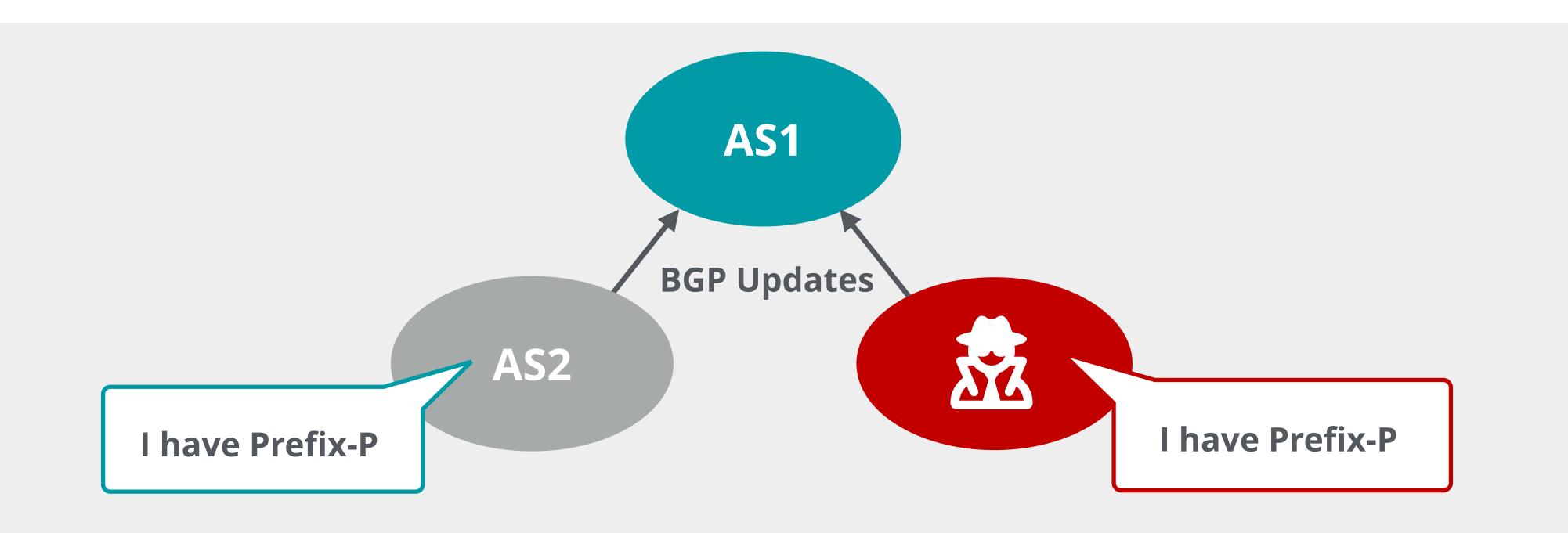
- BGP does not have a built-in authentication mechanism
- BGP provides no integrity or confidentiality
- BGP messages do not use a freshness service and can be replayed



# No Origin Validation



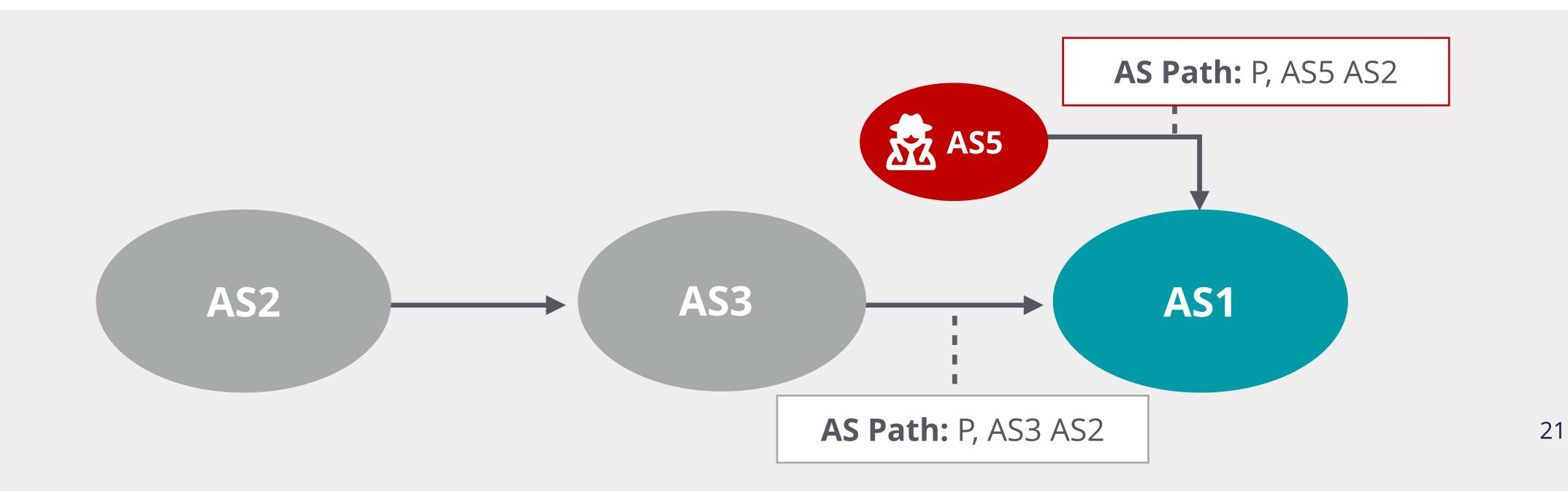
- BGP does not have a validity check for propagated routes
  - Any AS can announce any prefix



#### No Authentication of AS Path



- AS path attribute received in BGP update can not be validated
- Anyone can alter the path and prepend any ASN to the AS path



#### Due to these vulnerabilities





Attacks can be conducted by exploiting TCP or BGP messages



Any AS can announce any prefix



Any AS can prepend any ASN to the AS path

Fake routing information may disrupt Internet routing





# Causes of BGP Incidents

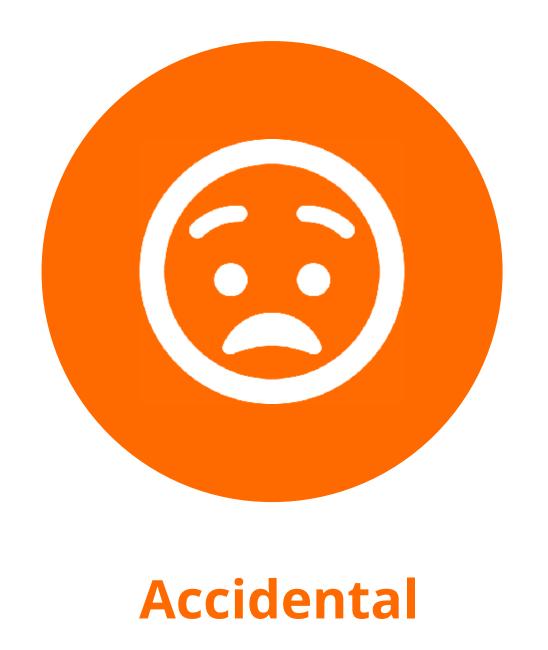
Section 2.2



#### Causes of BGP incidents can be divided into



Malicious



# Sometimes, it's Just Human Error



- Typo errors (fat finger)
  - May cause mis-origination
- Configuration errors
  - May cause mis-origination
  - AS path prepending mistake
- Simple mistakes may cause big problems!
  - BGP hijacks or route leaks



# But Sometimes They're Malicious!



ROUTE MANIPULATION ATTACKS	
TCP/IP PROTOCOL ATTACKS	
PROTOCOL MANIPULATION ATTACKS	
DENIAL-OF-SERVICE ATTACKS	

#### TCP/IP Protocol Attacks



- BGP uses TCP: vulnerable to TCP/IP based attacks
  - IP Spoofing
  - TCP Session Hijacking
  - SYN flooding attack



# IP Spoofing

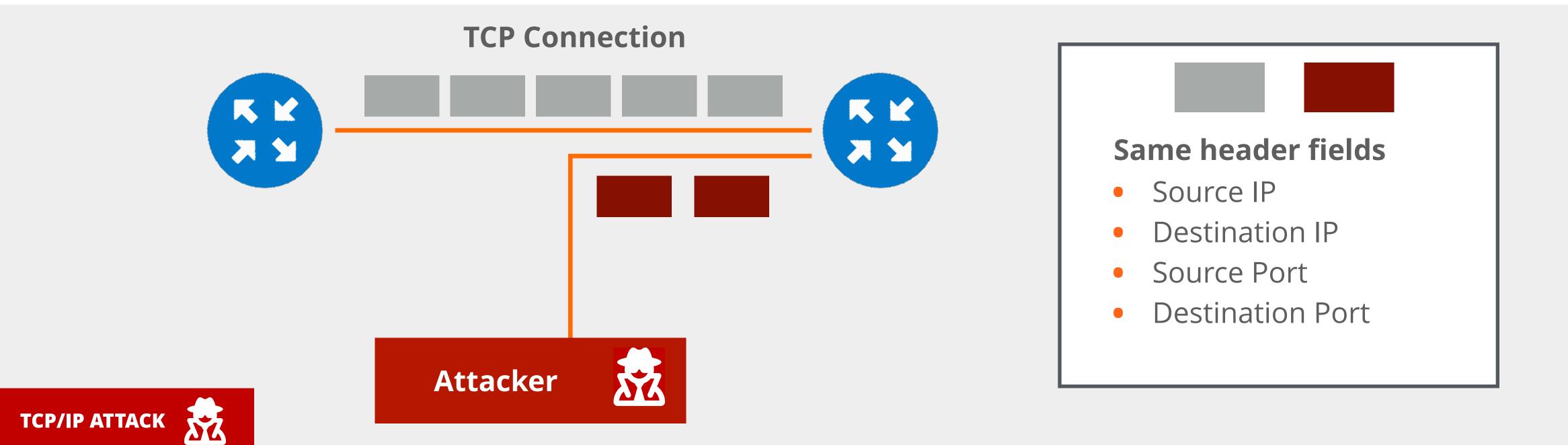


- An attacker could spoof an IP address of a BGP peer in order to:
  - Establish an unauthorised BGP session
  - Break an existing BGP session
  - Inject bogus routes or delete routes

# TCP Session Hijacking



Involves intrusion into an ongoing session



# TCP Session Hijacking



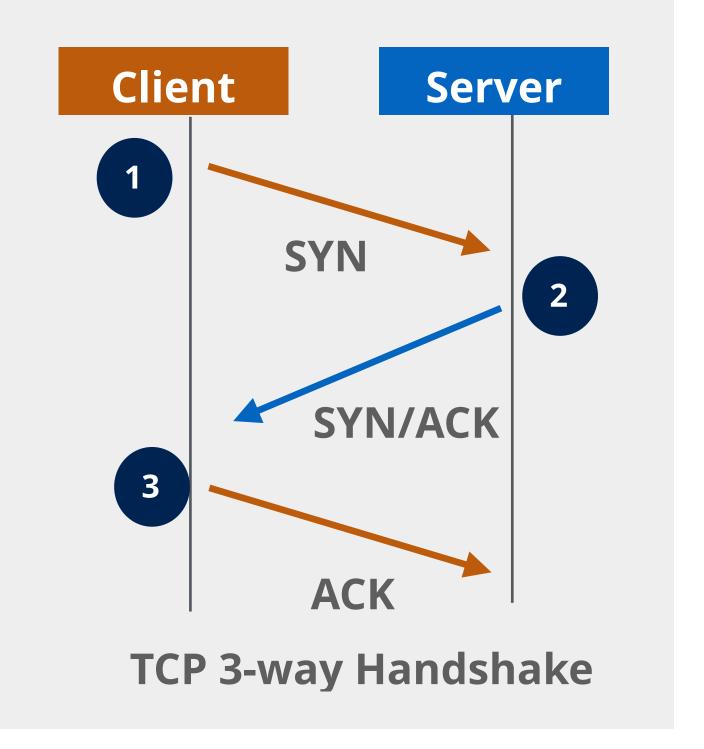
- Requires an attacker to discover the following:
  - src IP, dst IP, src port, dst port
  - sequence number of ongoing session
- In BGP, it could be used to
  - bring down the BGP session between peers (TCP RST)
  - inject false routes into BGP, delete or modify routes

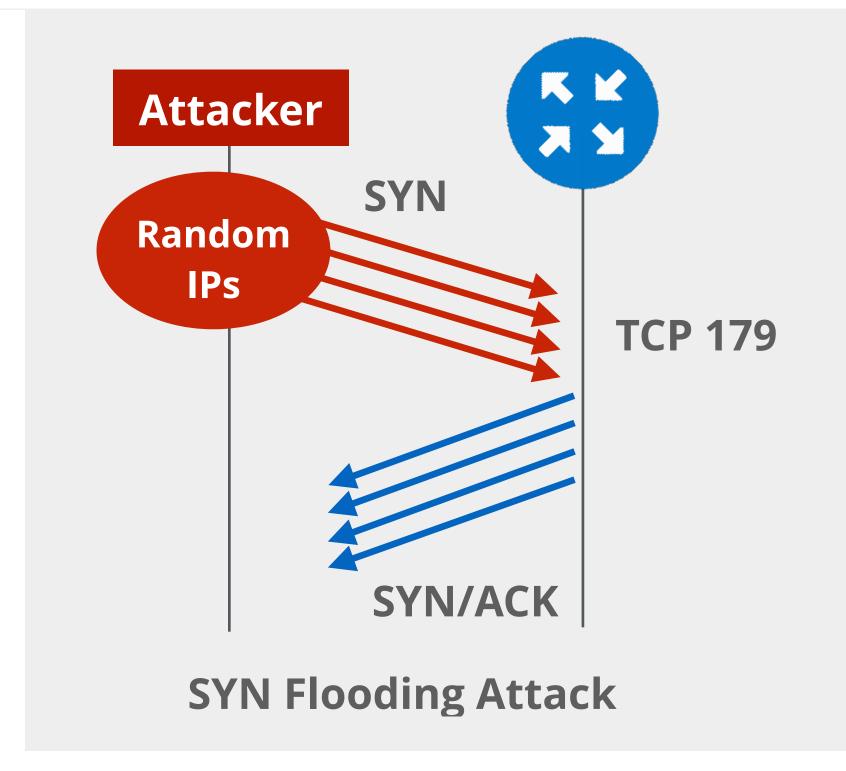


# SYN Flooding Attack



- A type of Denial-of-Service (DoS)
- Exploits the three-way hand shake process of a TCP connection
- Goal: Exhaust resources



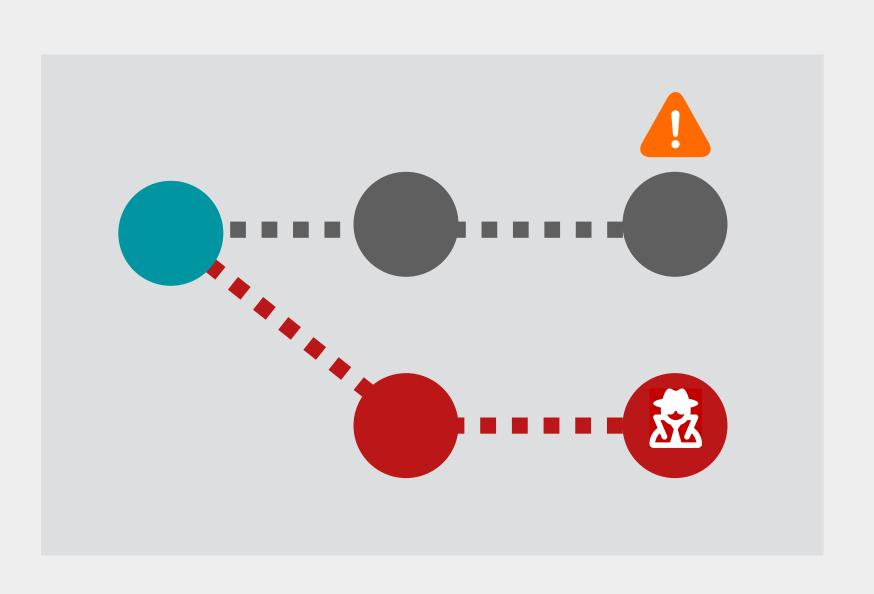




# **BGP Route Manipulation Attacks**



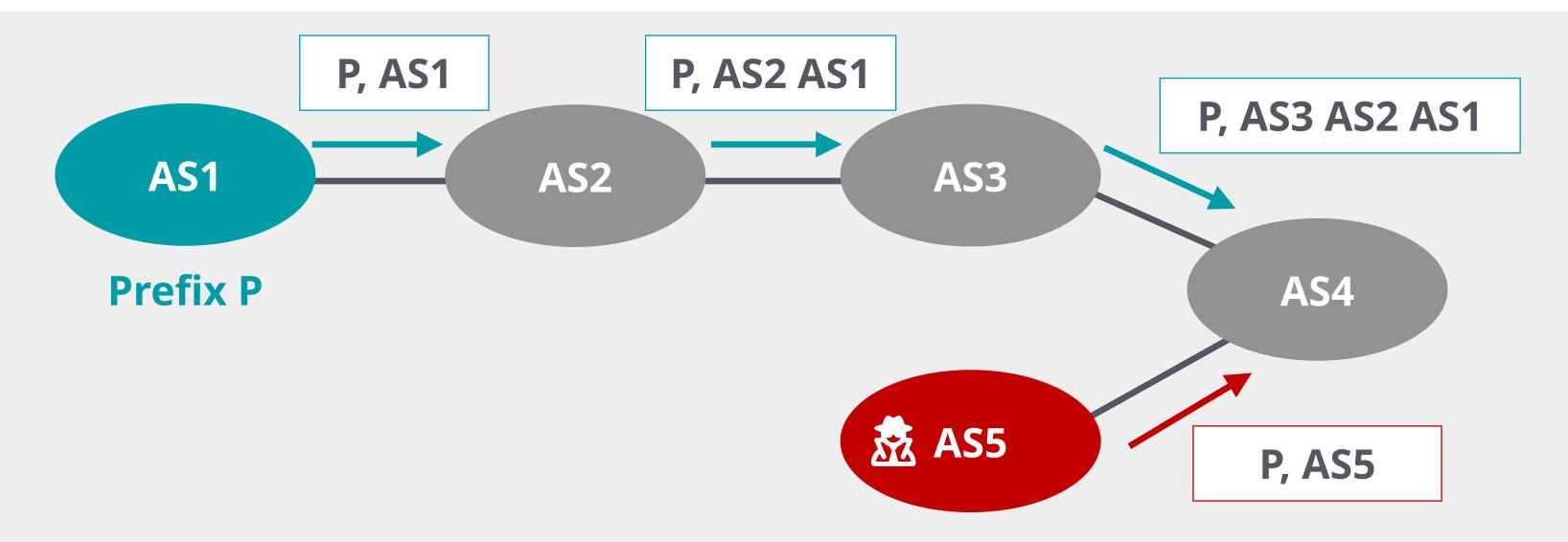
- Goal: Blackholing, eavesdropping or traffic analysis
- Attacker can:
  - Inject bogus routes into BGP tables
  - Reroute packets based on their intentions
  - Prevent traffic from reaching the intended destination
- Can be classified as:
  - BGP Origin Hijacks
  - BGP Path Hijacks
  - BGP Route Leaks
- Very common! Our focus on this course



# **BGP Origin Hijack**



- The hijacking AS:
  - Abuses mutual trust between ASes
  - Originates a prefix that it is not authorised to originate!
- Difficult to say whether it is an accident or an attack
- Traffic lost or received by attacker (eavesdrop)

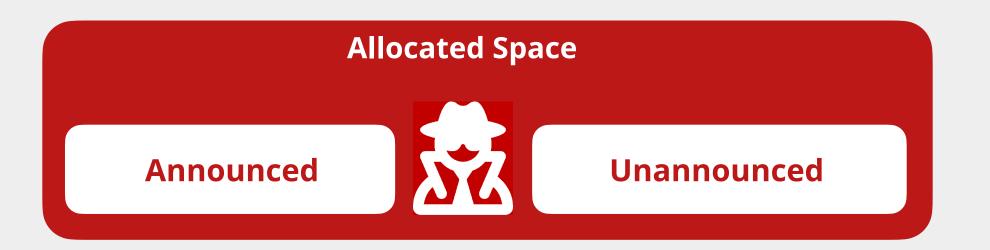




# Hijacks of Allocated Addresses



- Allocated address space could be:
  - Currently in use (announced prefixes)
  - Or unused IP space (unannounced prefixes)
- Unannounced prefixes are preferred by spammers
  - No operational impact
  - Potentially harms the reputation of the holder



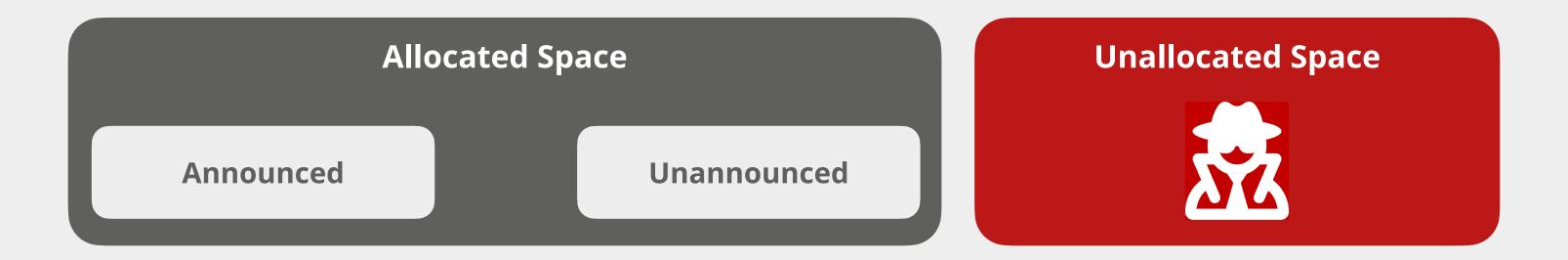
Unallocated Space



# Hijacks of Unallocated Addresses



- IP blocks not assigned by IANA or RIRs
- Effective if full bogon filters not applied
- No whois entries, no complaints!
- Again, a good choice for spammers

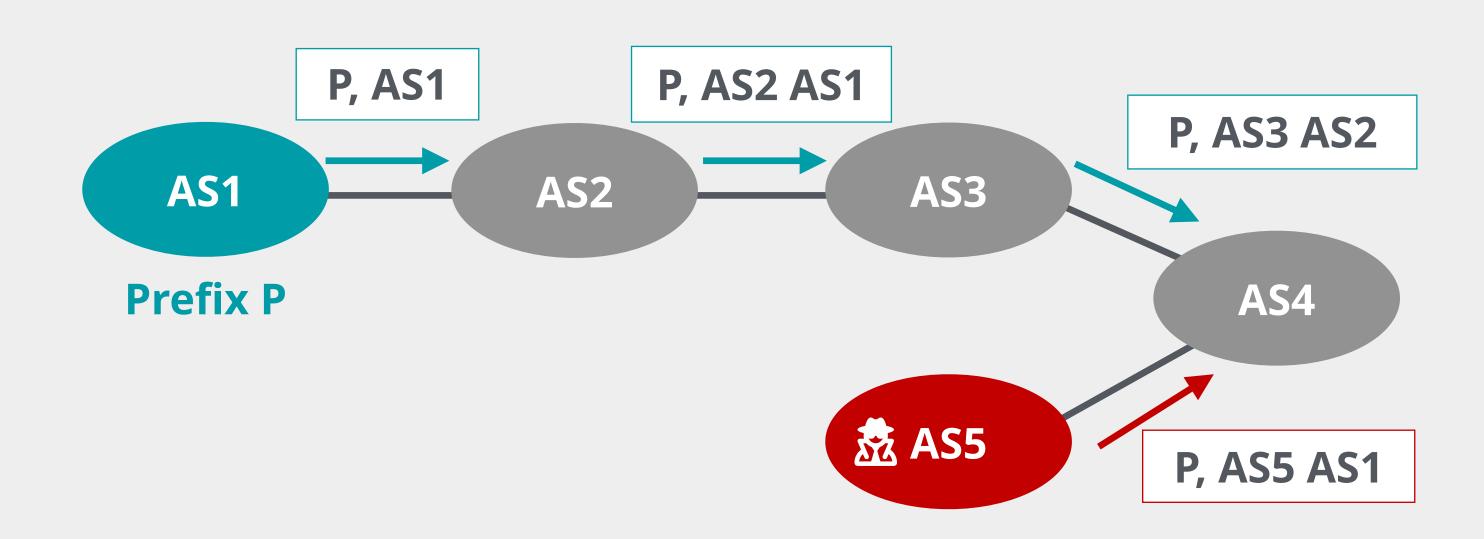




# BGP Path Hijack



- No verification of path attributes in received BGP updates
- Hijacker can modify the AS Path and redirect traffic
- Traffic lost or eavesdropped/modified (adds latency)

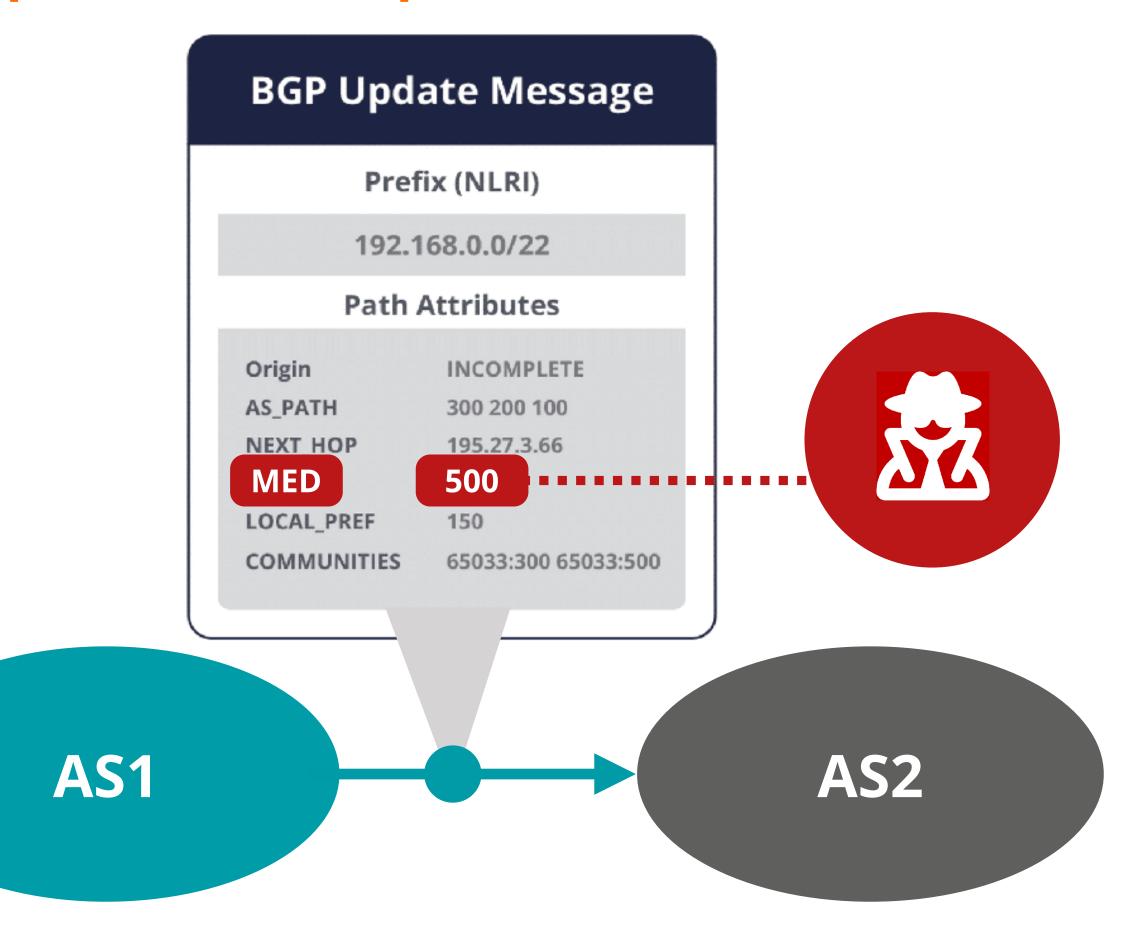




# **Protocol Manipulation Attacks**



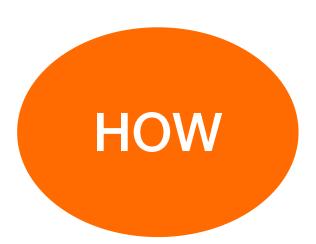
- Relatively new type of attack, no reports yet
- A malicious AS aims to manipulate properties of BGP protocol
- An attacker may:
  - Modify MED attribute
  - Exploit RFD/MRAI timer



# **Protocol Manipulation Attacks**



- Multi-exit-discriminator (MED)
  - A malicious AS may affect ASes' decisions by altering this attribute
- Route Flap Damping (RFD) / Minimum Route Advertisement Interval (MRAI) timers



A Malicious AS artificially withdraws and re-announces a route



ASes using RFD timer consider the route unstable and suppress it



ASes using MRAI timer delay the distribution of corresponding update message

# Denial of Service (DoS) Attacks



- An attacker can execute DoS attacks in several ways:
  - BGP session failure due to congestion
  - Deliberate route flapping
  - Explosion of routing table size
  - Blackholing traffic
  - TCP attacks (SYN flooding or TCP Reset)
- DDoS solutions are already available
  - Not specifically for BGP speakers



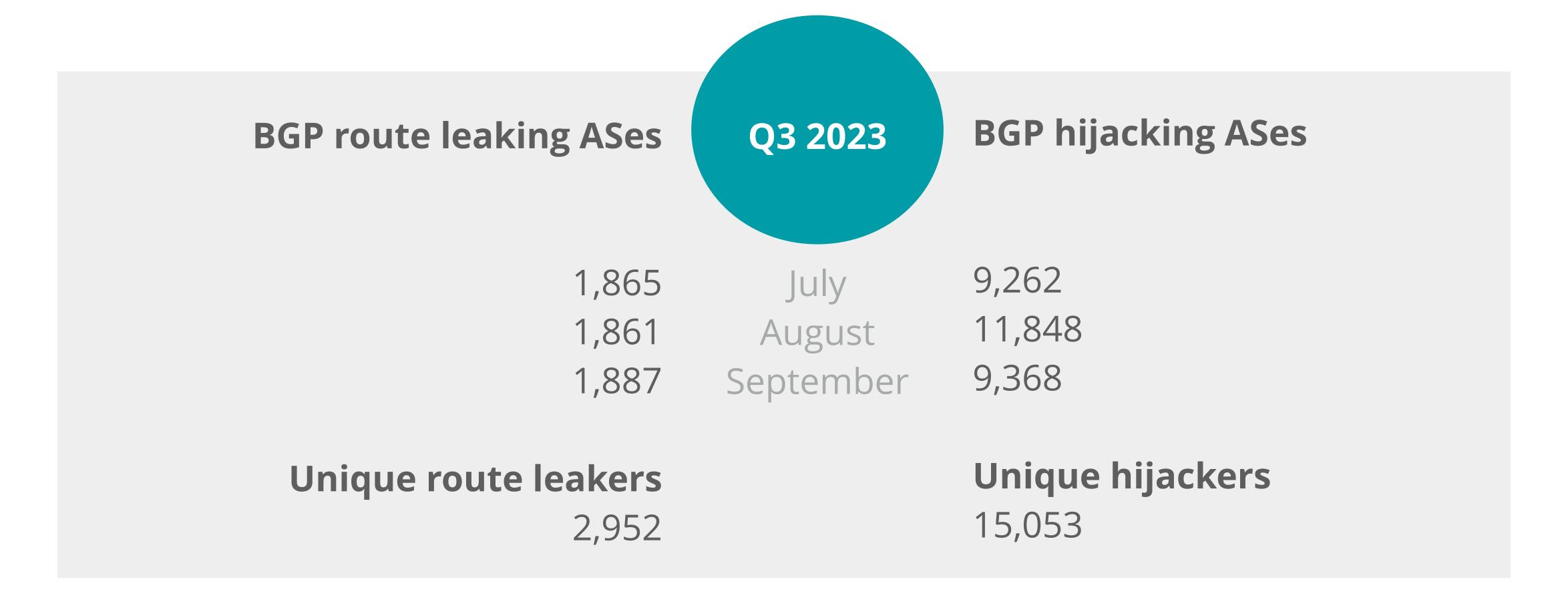
#### To Summarise



- BGP is vulnerable to mistakes and attacks
- Attackers could:
  - Inject bogus routes into the BGP table
  - Hijack a BGP session and break peer-to-peer connections
  - Initiate a DoS attack and exhaust victim's resources
  - Manipulate BGP and reroute packets
  - Intercept and eavesdrop
  - Blackhole the entire network, etc

#### BGP Incidents in Q3 2023





#### Source:

# April 2020: Akamai, Amazon

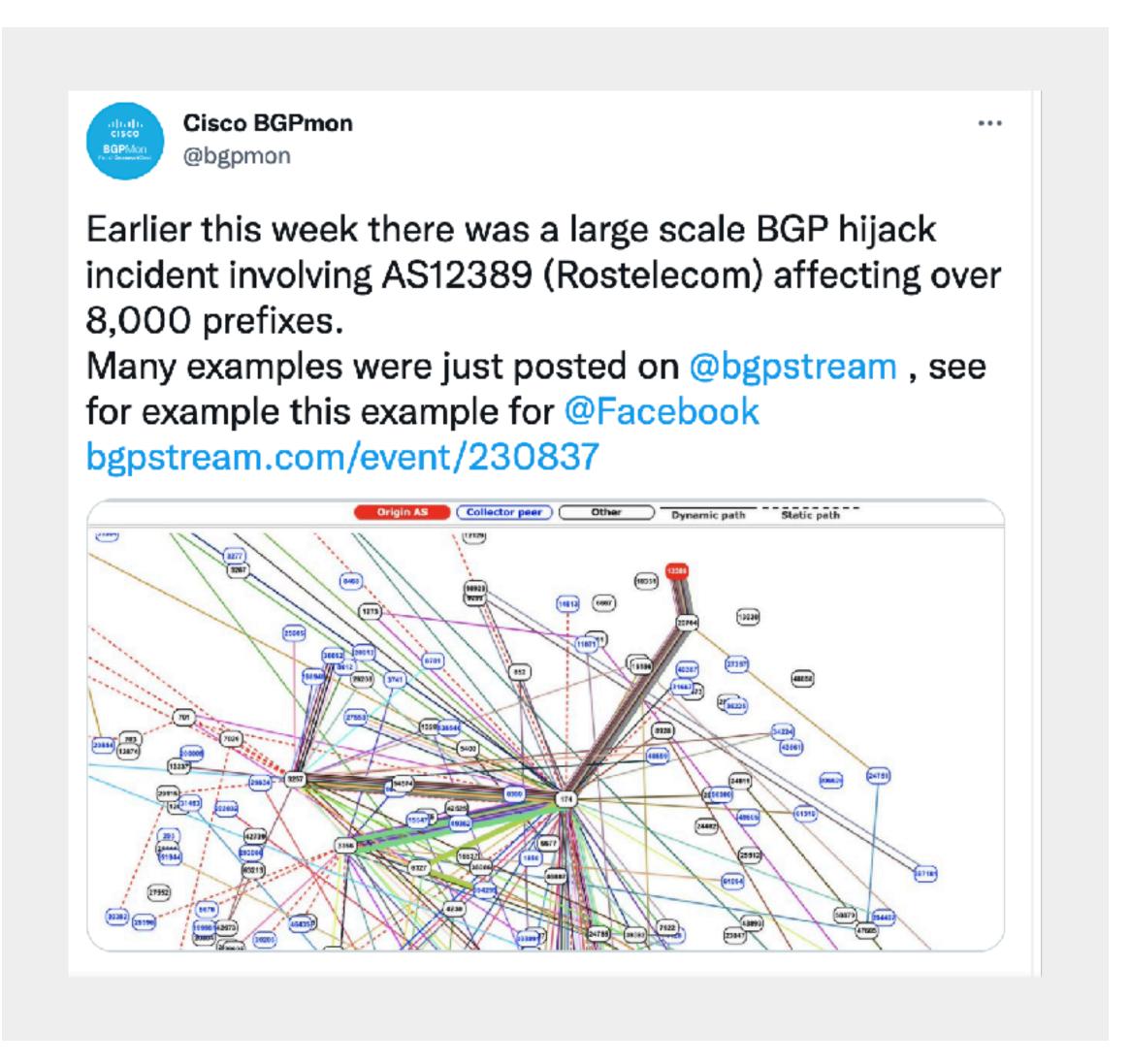


#### What happened?

- 8k+ routes hijacked by Rostelecom (AS12389)
- 200+ CDNs and cloud providers impacted
- Not known how much data leaked

#### Why?

- Malicious activity
- Lack of good filtering by upstream providers/peers





# Questions



Section 3



# How to Mitigate BGP Threats

Section 3.1

# How to Mitigate BGP Threats?



- To deal with this, you need to:
  - Secure message exchange between BGP speakers
  - Validate the routing information you receive



Some authentication and verification mechanisms should be in place



Prevent propagation of incorrect routing information

# How to Mitigate BGP Threats?



• It requires the following to be verified:



Does the BGP speaker belong to the AS that it claims?



Is the prefix originated by the legitimate holder or an AS that is authorised to originate it?

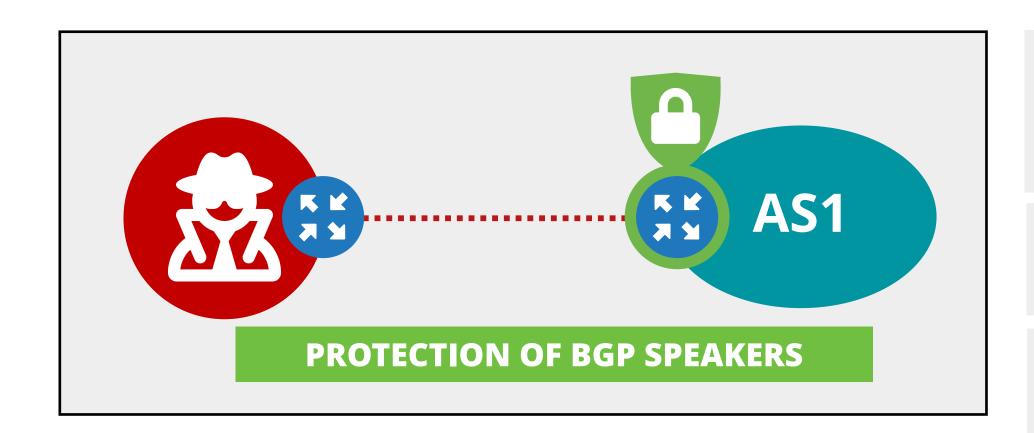


Does the AS path reflect the sequence of ASes that the BGP UPDATE packet has traversed?



Are the attributes in a BGP UPDATE message correct and have not been tampered with?





Only BGP peers to send packets to TCP 179: Control Plane Policing (CoPP), or ACLs (if CoPP not supported)

Limit accepted BGP traffic

uRPF to mitigate DoS/DDoS attacks

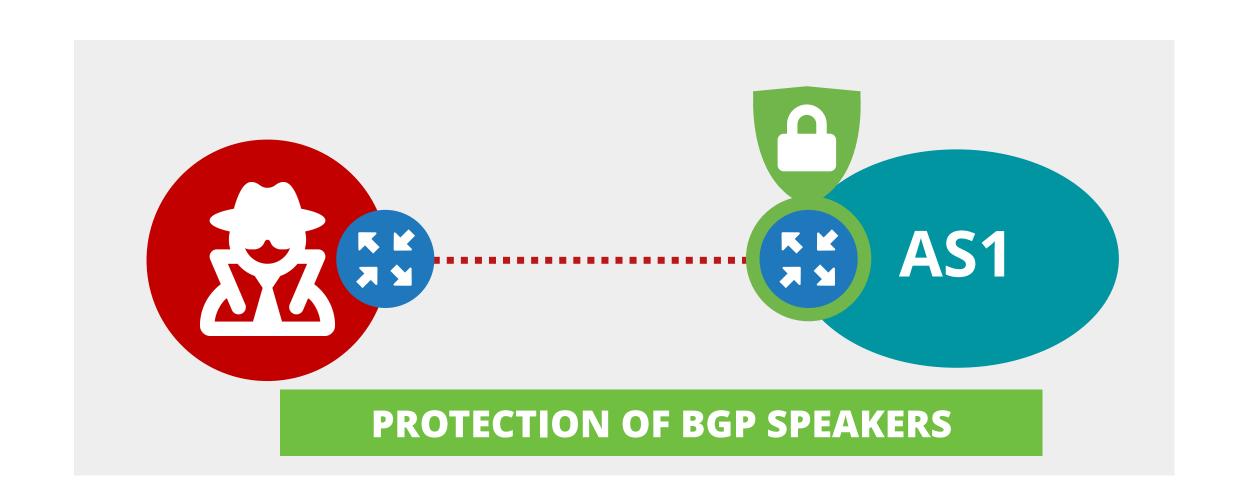


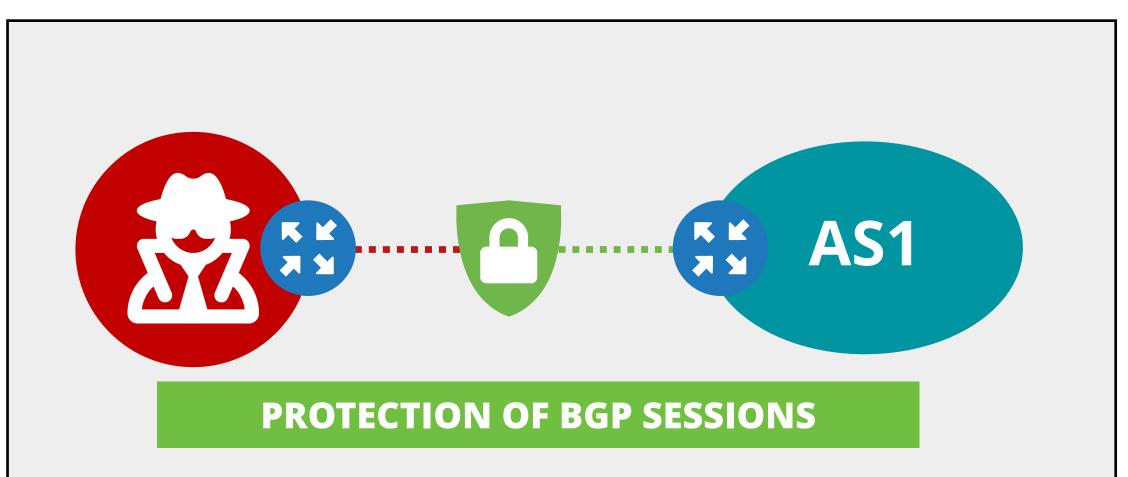


**BGP Security** 

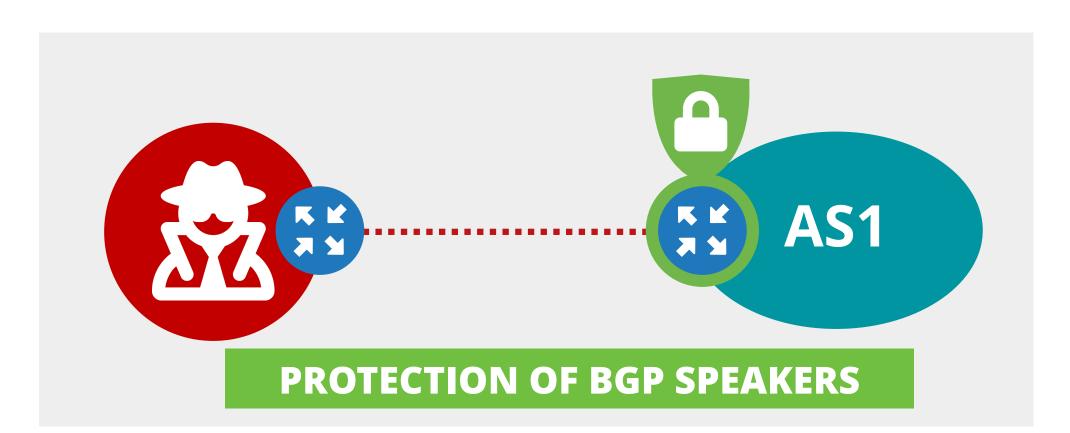
https://academy.ripe.net/bgp-security/



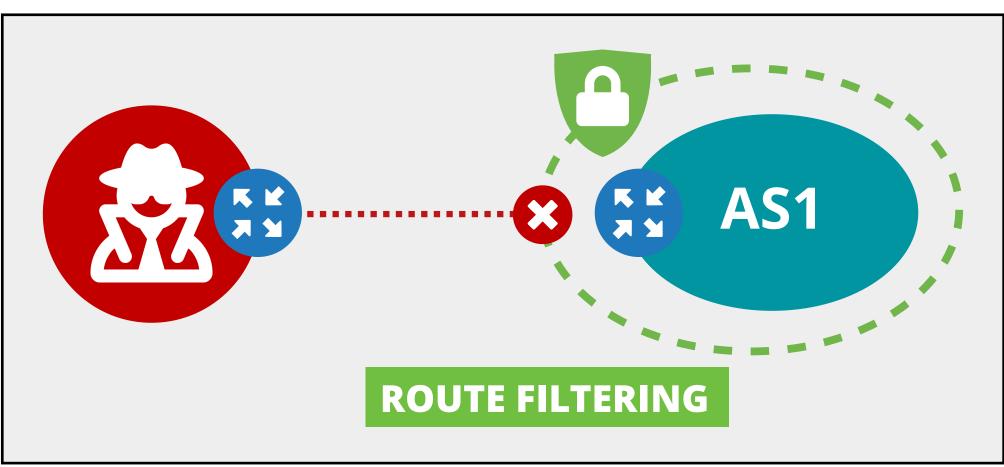




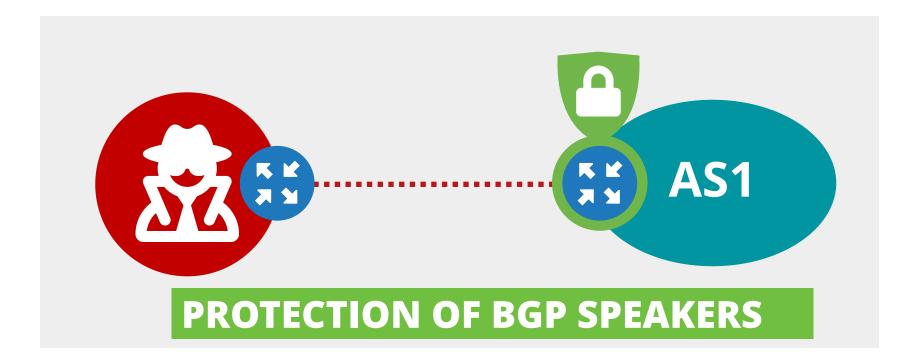




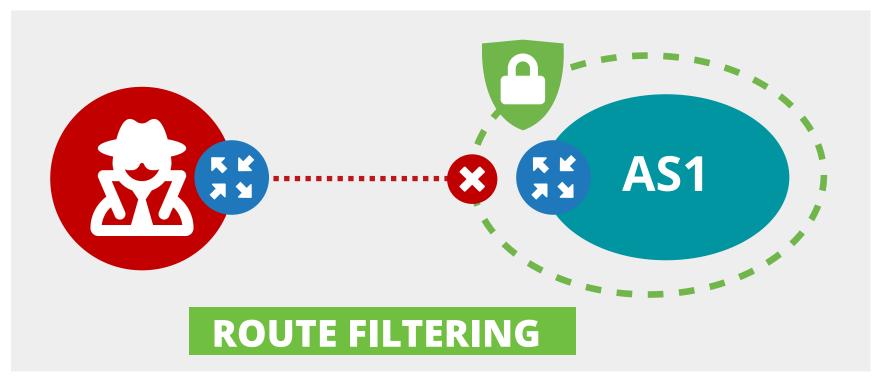


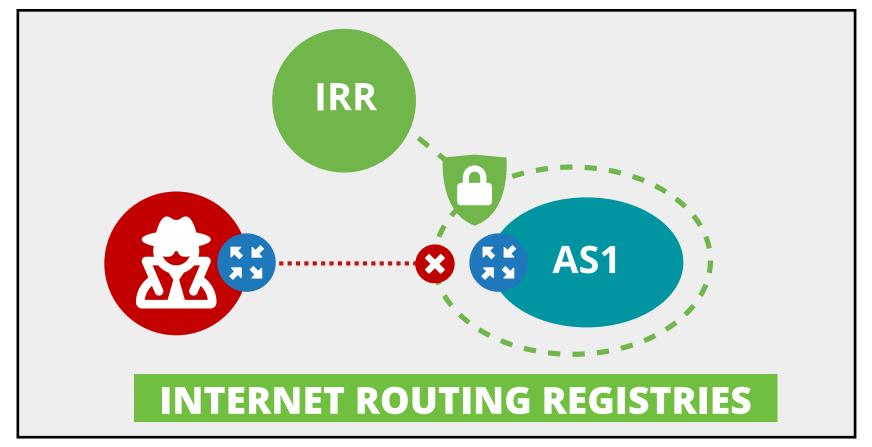




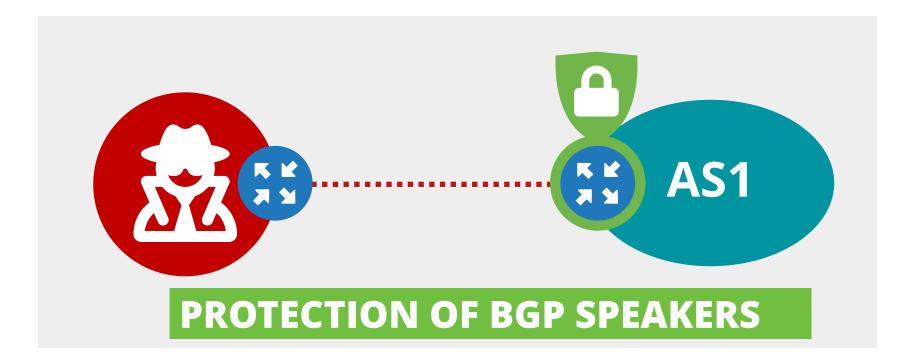




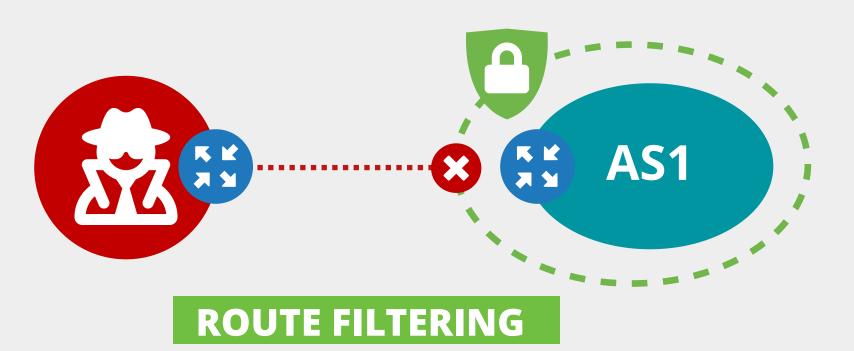


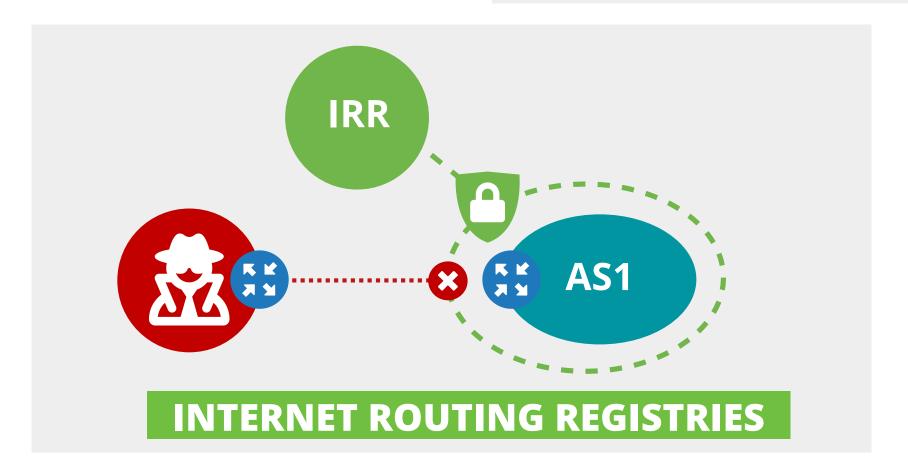


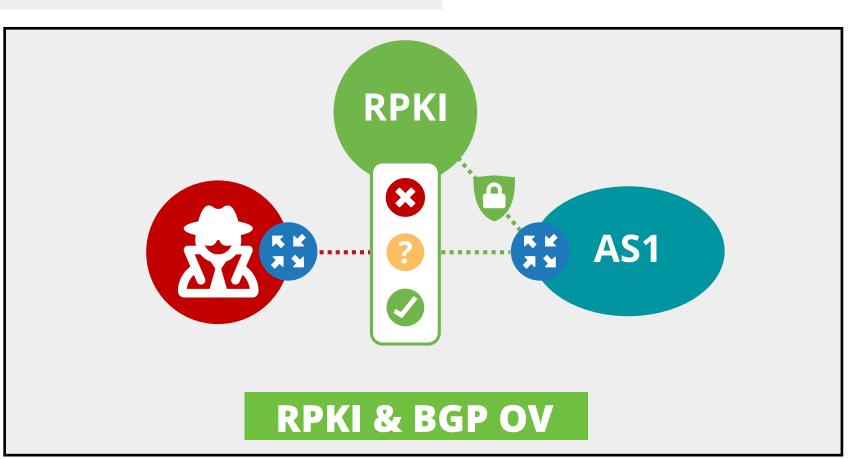














#### **Protection of BGP Sessions**

Section 3.2

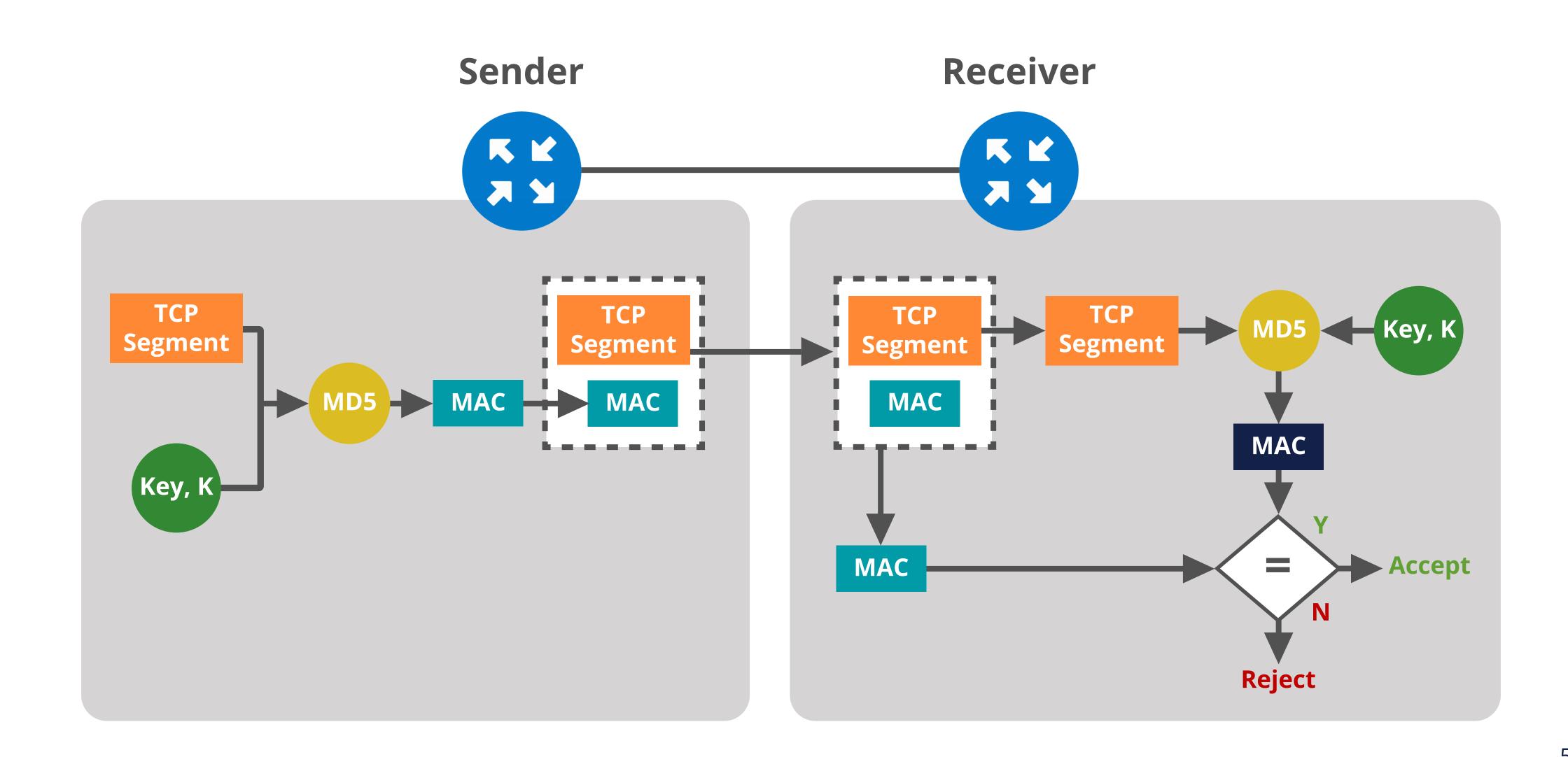
#### **BGP Session Protection**



- BGP sessions are subject to TCP/IP vulnerabilities
  - IP Spoofing, TCP session hijacking, SYN flooding
- Attacks against message integrity and confidentiality are possible
  - Man-in-the-middle and replay attacks
- We will see three solutions:
  - TCP MD5 and TCP-AO, to protect the BGP TCP session
  - BGP TTL Security (GTSM Generalised TTL Security Mechanism)

#### TCP MD5

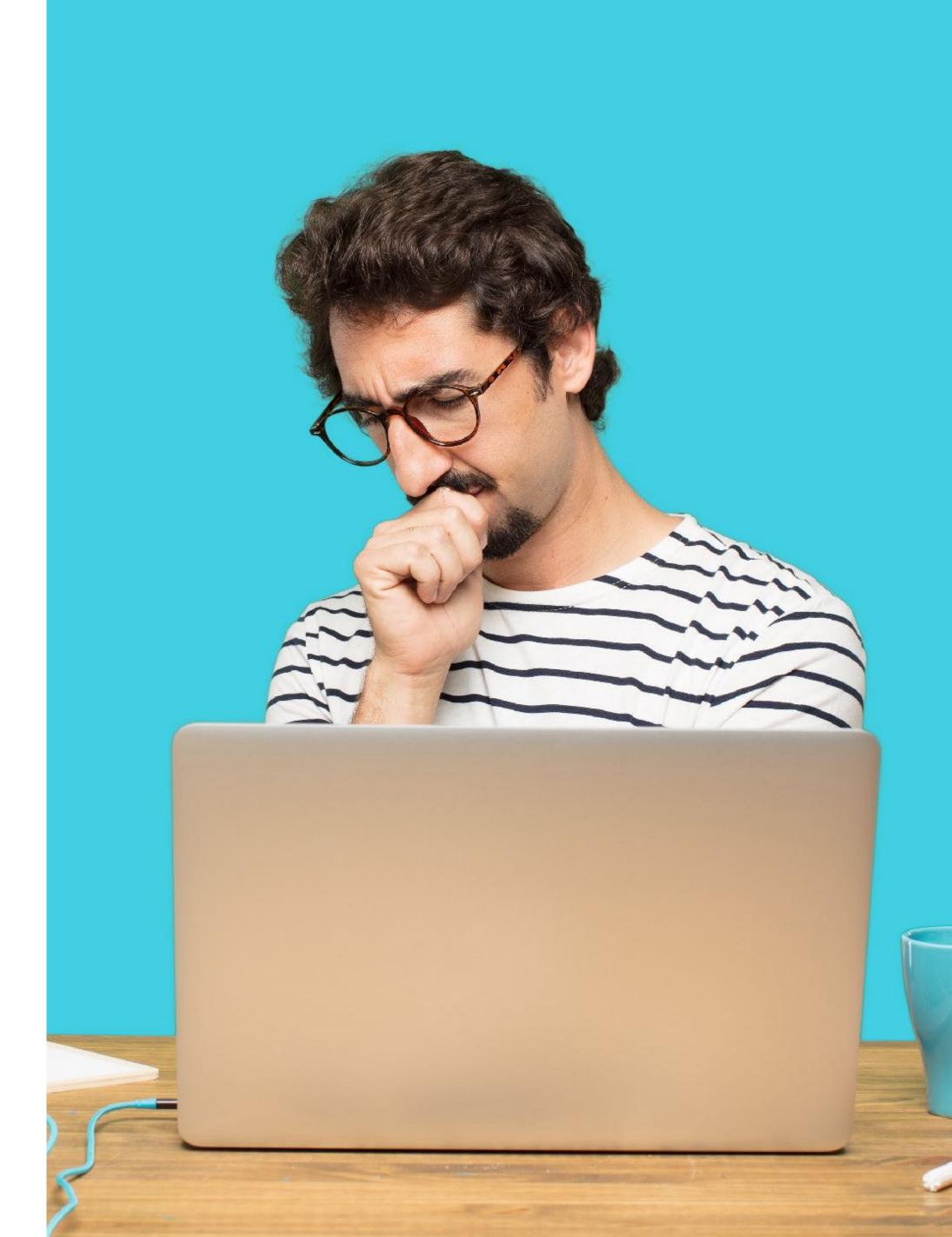




#### Limitations of TCP MD5

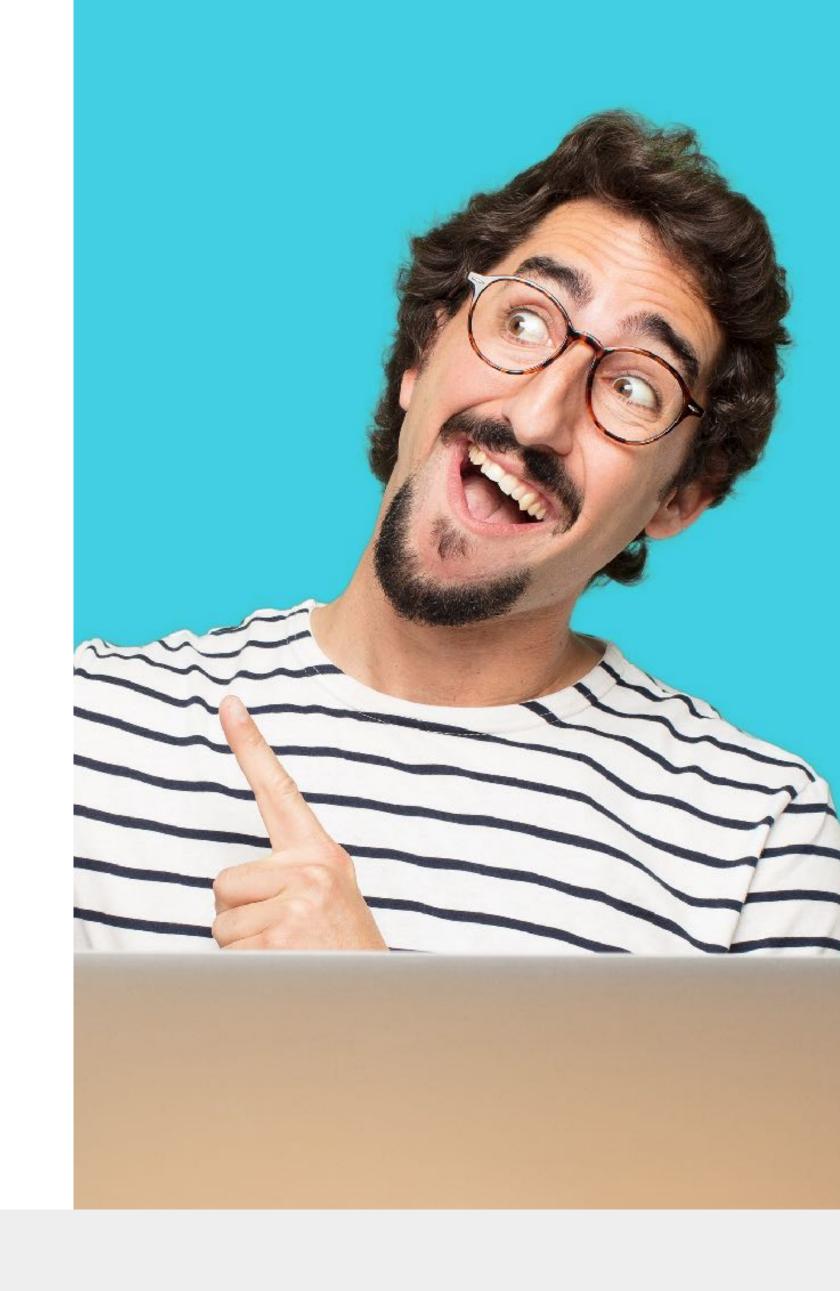
- Not a strong authentication mechanism
  - Supports only MD5 algorithm
- Doesn't allow dynamic key rollover
  - Changing pre-shared keys requires TCP session reset
  - Problem for long-lived sessions

MD5 has been obsoleted by TCP-AO, not recommended.



#### TCP-AO

- Enhances security and authenticity of TCP segments in BGP and LDP sessions
- Supports multiple stronger authentication algorithms
  - HMAC-SHA-1-96 and AES-128-CMAC-96
- Better key management and agility
  - Change keys without resetting TCP session
- Protects long-lived TCP sessions against replay attacks



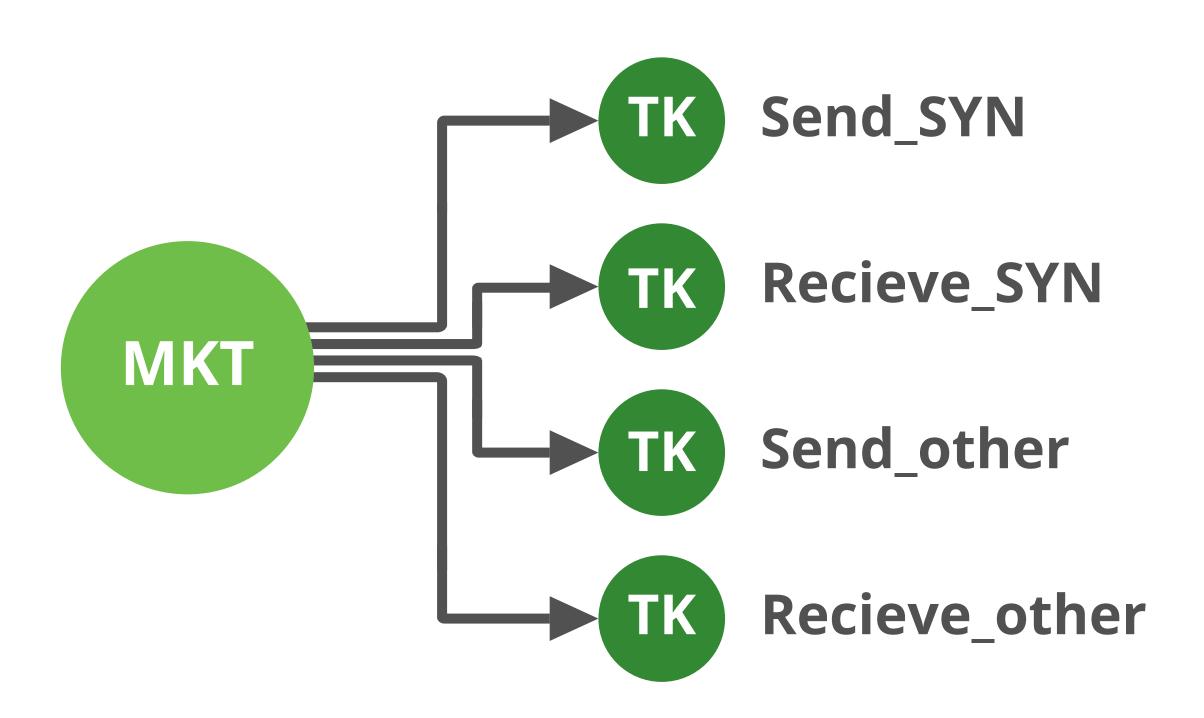
RFC 5925 - "The TCP Authentication Option"

RFC 5926 - "Cryptographic Algorithms for the TCP-AO"

#### TCP-AO



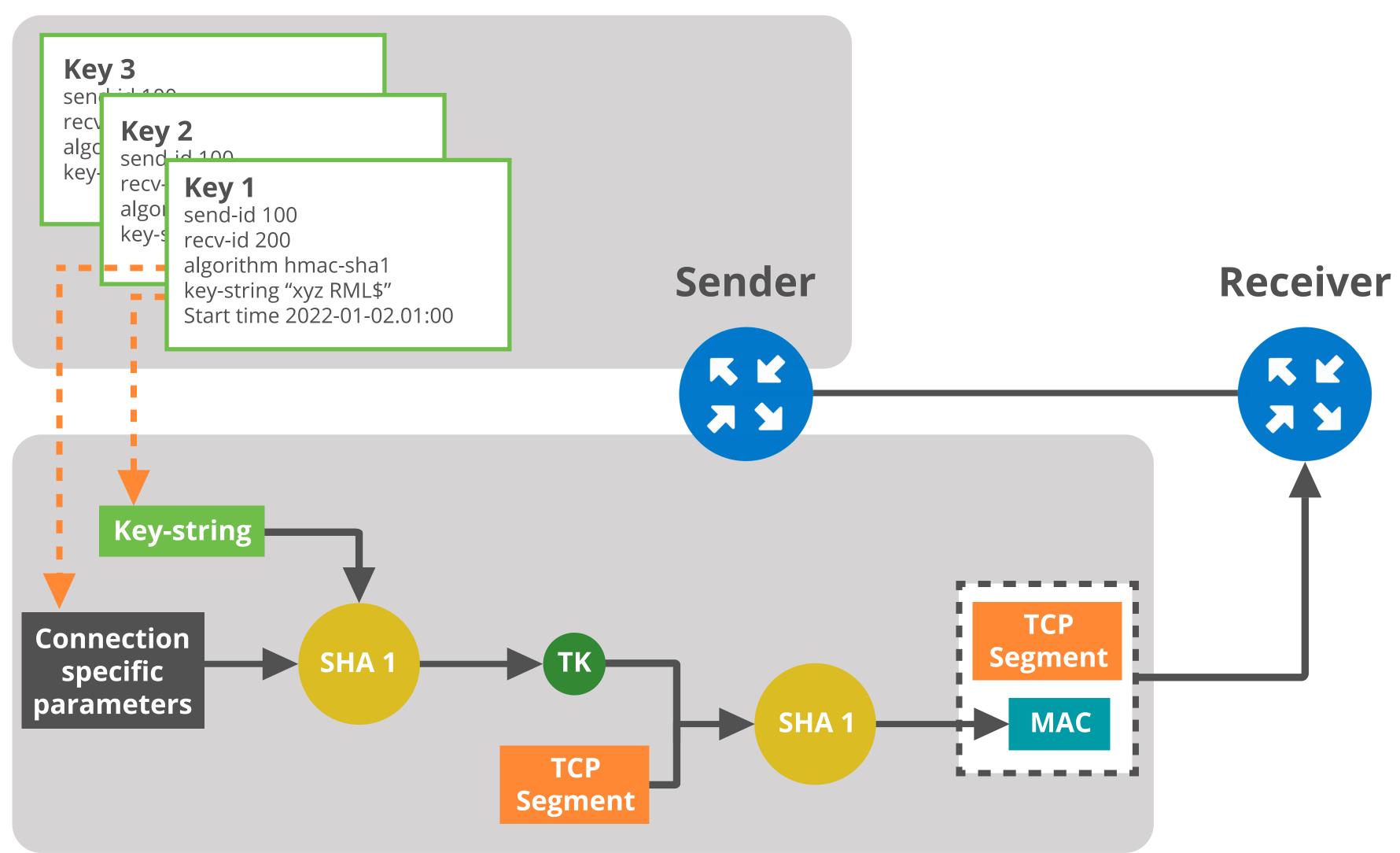
- Two sets of keys to authenticate incoming and outgoing segments:
  - Master Key Tuples (MKTs) (key-chain) and Traffic keys
- Four traffic keys are derived from each MKT



#### **How Does it Work?**

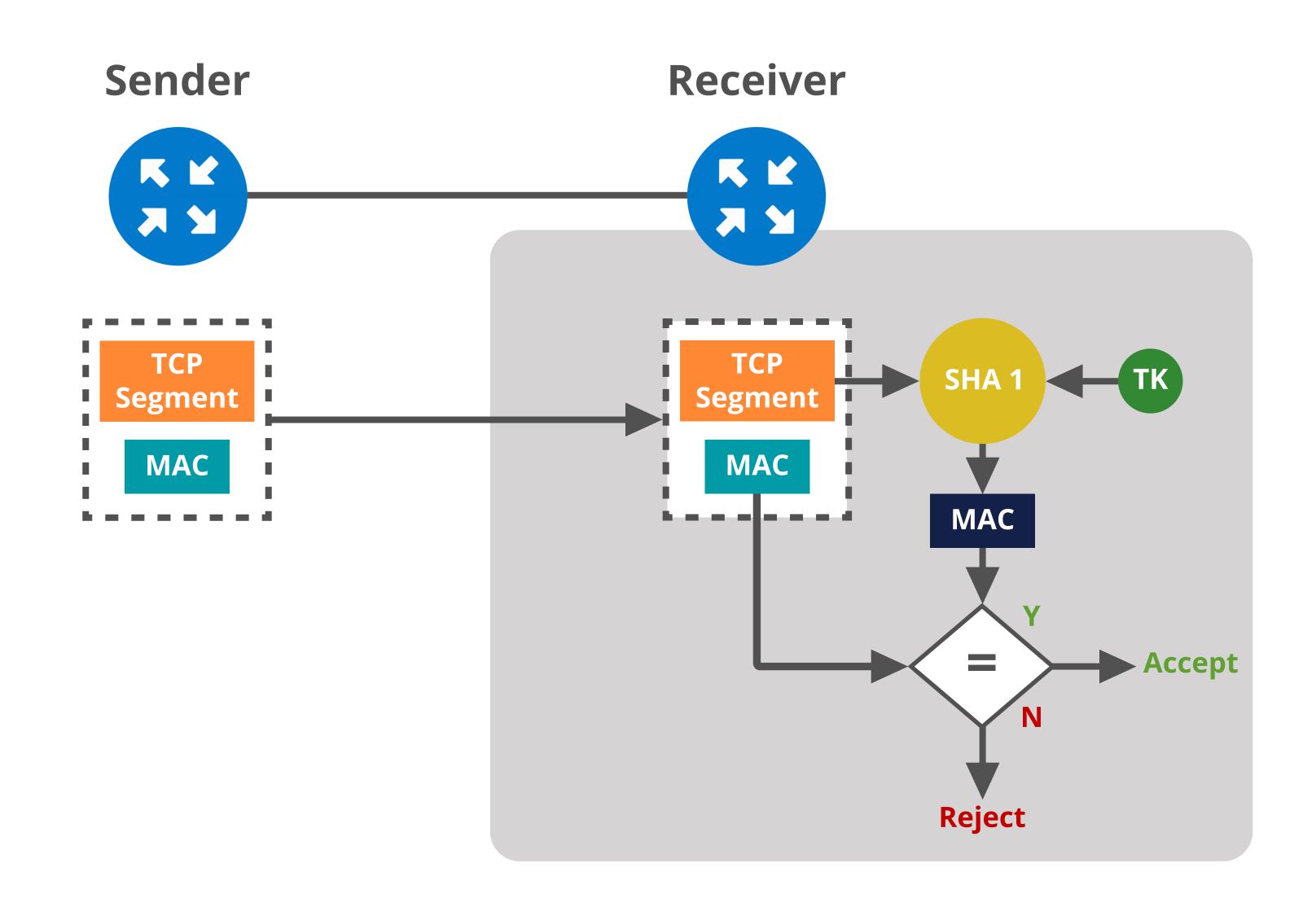


#### **MKTs**



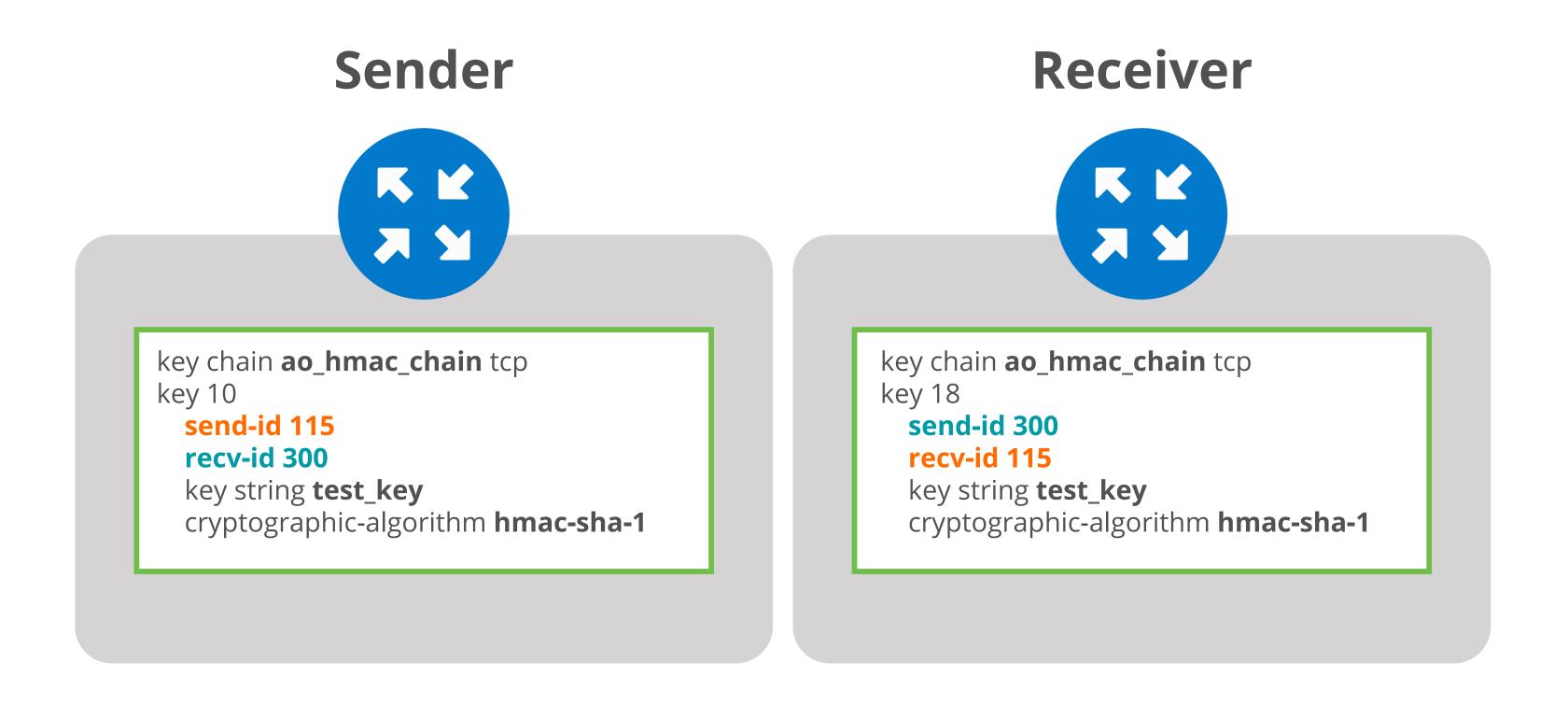
#### **How Does it Work?**





#### TCP AO Configuration





- The master-keys/key chains must be identical on both BGP peers
- Send and receive IDs must match
- Make sure the same MAC algorithm is used on both sides

#### TCP AO Configuration



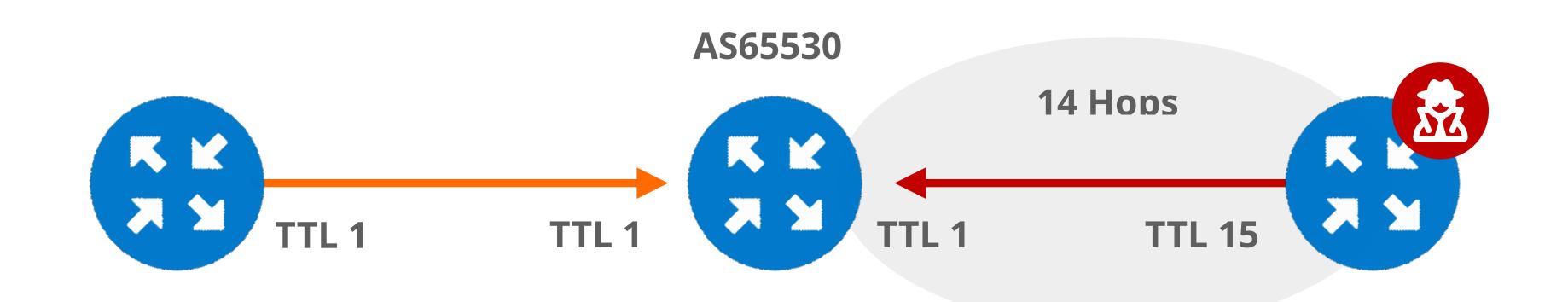
Last step is to apply it to BGP neighbour

```
(config)# router bgp 65530
(config-keychain-tcp)# neighbour <peer-IPv4/IPv6-address> ao <keychain-name>
[include-tcp-options]
```

# GTSM (TTL Security)



- TTL/Hop limit =1 by default for eBGP sessions
- Remote attacker may adjust TTL and send spoofed packets
  - May execute CPU utilisation-based attacks (DoS attacks)

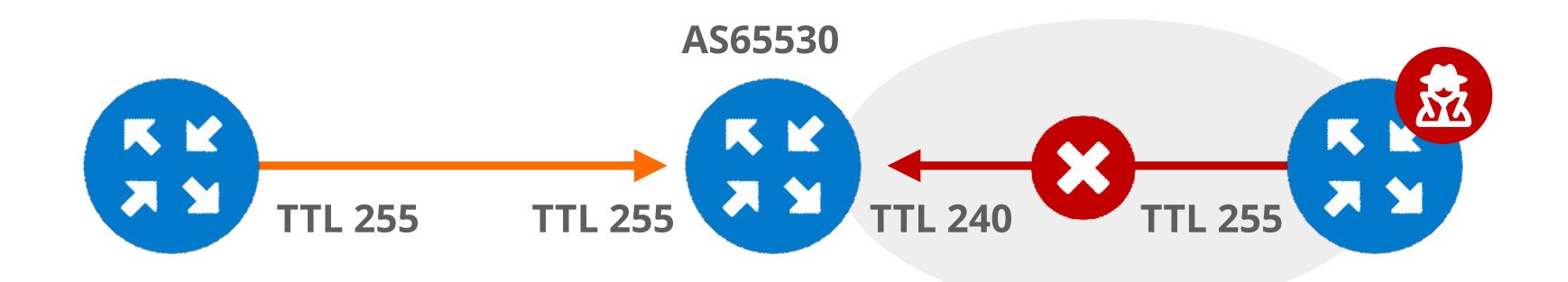


Attacker sends a large number of forged BGP packets

# GTSM (TTL Security)



- Should be implemented on directly connected eBGP peering
  - Send packets with TTL/Hop-limit 255. Discard packets if it is < 255.</li>
  - Configured on both ends of a BGP session
- Could be applied to multi-hop BGP peering, but not so effective



GTSM **enabled**, TTL of all BGP packets are set to 255

GTSM **enabled**, BGP packets with TTL less than 255 are dropped

Attacker sends a large number of forged BGP packets

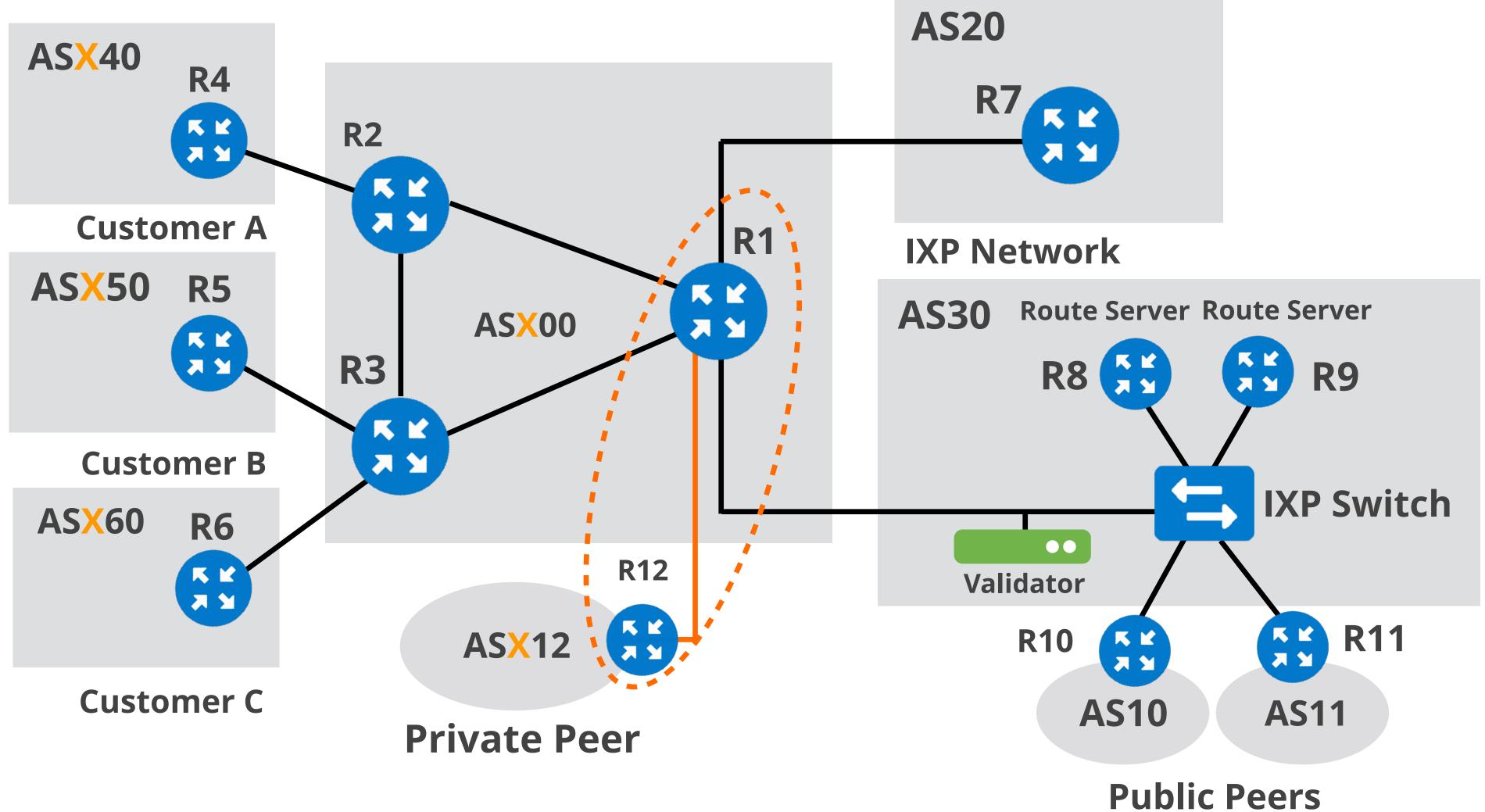


# Questions



#### **Transit Provider**





Your AS number AS X00 Your IPv6 allocation 2001:db8:X00::/48

# Lab 1 - Securing BGP Sessions



Description: Implement two techniques to protect BGP sessions

#### Goal:

Choose suitable and available security measures related to BGP sessions

• Time: 15 minutes

#### Tasks:

- 1.1 Configure MD5 authentication between two BGP routers
- 1.2 Configure GTSM (TTL Security) in addition to MD5





- What have you learned?
  - You have to check which features are available
  - You can combine protection techniques





# Implementing Route Filtering

Section 3.3

#### **Route Leaks**



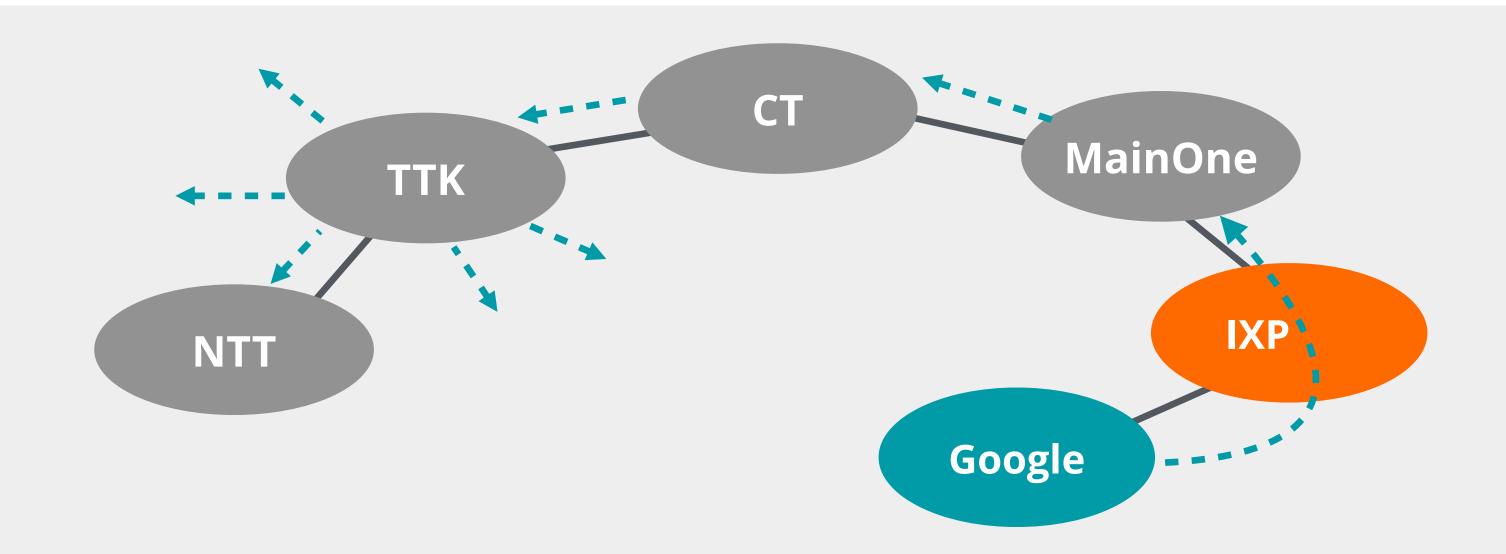
#### "The propagation of BGP announcements beyond their intended scope" [RFC7908]

- Illegitimate propagation of legitimate prefixes (not bogus routes)
- Result from human errors or misconfigurations
  - And/or improper or missing BGP route filters between BGP peers
- Leads to incorrect or suboptimal routing

# Google Prefix leak - November 2018



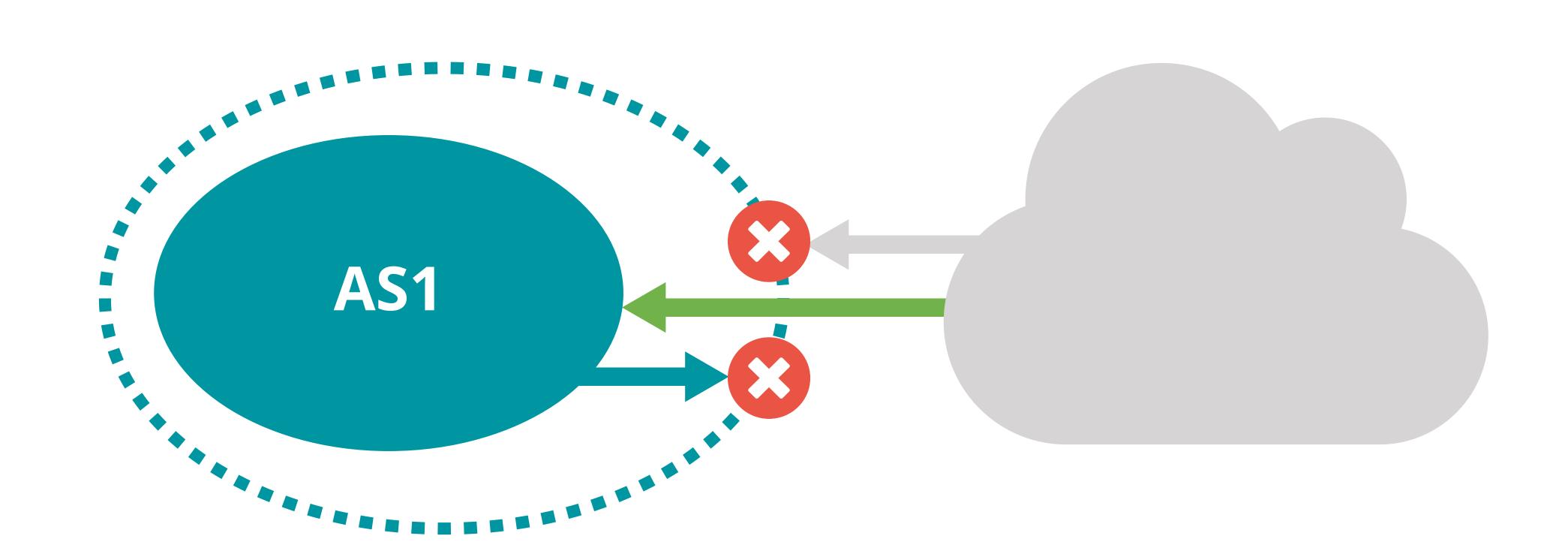
- What happened?
- MainOne leaked Google routes to CT and CT leaked them to other transits
  - Google services (G Suite and Google Search) affected by the leak
- Why?
  - Due to misconfigured filters



## **How to Prevent Route Leaks?**



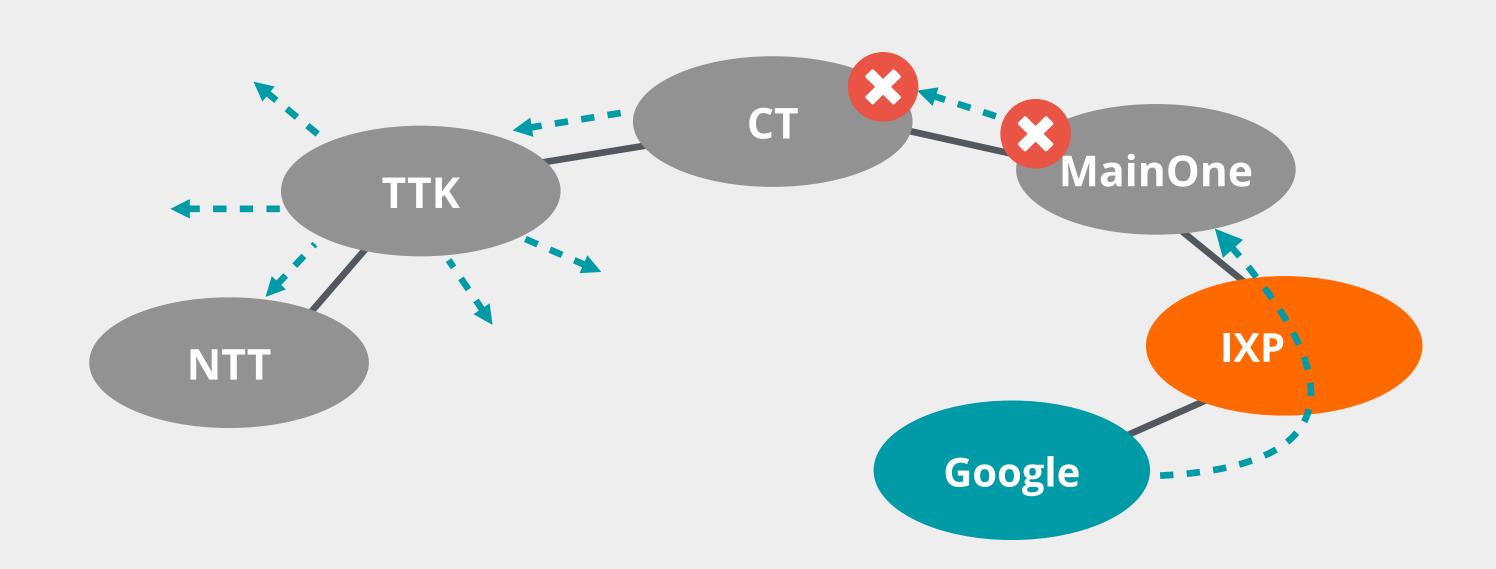
## Route filtering is the most powerful mechanism!



## Google Prefix leak - November 2018



- What's different with proper filters?
  - Google's prefix wouldn't reach CT
  - Proper outbound filters in MainOne, and/or
  - Proper inbound filters in CT



## What is BGP route filtering?



- The most basic **protection** mechanism against malicious or accidental BGP incidents:
  - Prevents route leaks
  - Mitigates the impact of BGP hijacks
- Technique used to control prefixes on the BGP peering
  - Which prefixes will you advertise to your peers?
  - Which prefixes will you accept into your network?

**Essential for routing security!** 



# Other Reasons for Filtering



### Business relationships

Customer-provider, peer-peer

### Technical reasons

Reduce memory utilisation, scalability

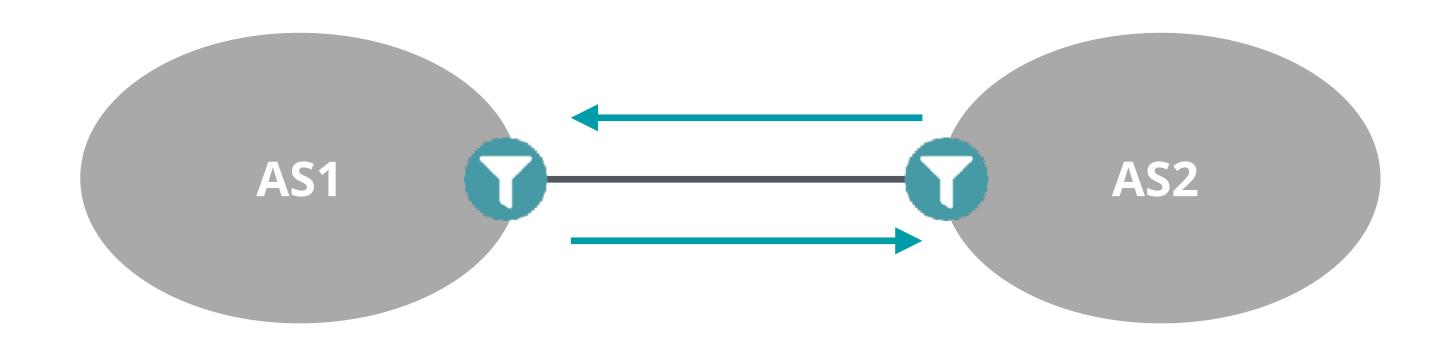
### Traffic engineering

Manipulate traffic flows and influence best path selection

## **BGP Filters (BGP Policies)**



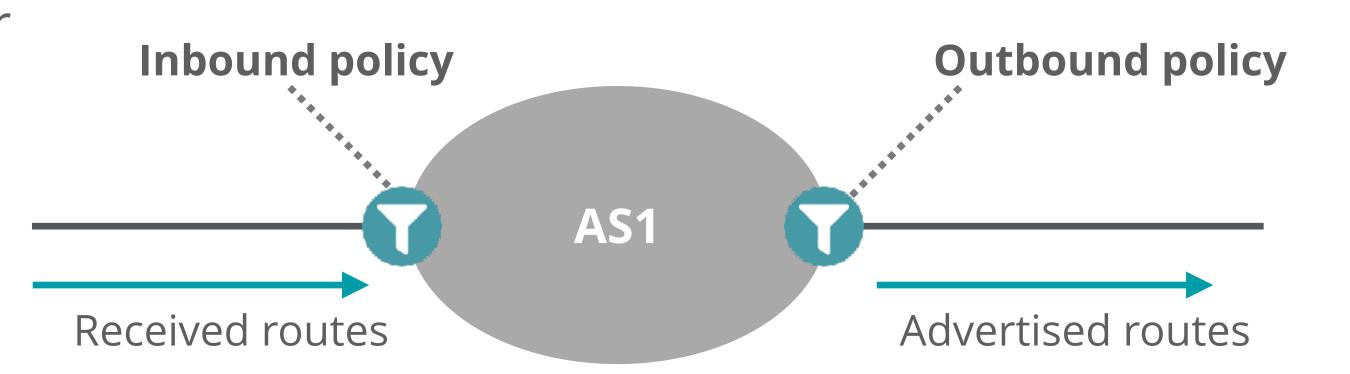
- Used to filter prefixes exchanged between BGP peers
- Describe BGP peers and routing relationships with them
- Filters can match on
  - IP prefixes
  - AS paths
  - Or any other BGP attributes (e.g. MED, BGP communities, etc)



## **BGP Filters (BGP Policies)**



- Inbound policy:
  - For **incoming** (received) routes
  - Detects configuration mistakes and attacks
- Should be applied on each eBGP peer
  - Both on ingress and egress
- Outbound policy:
  - For outgoing (advertised) routes
  - Limits propagation of routing information



# Filtering Principles



- Filter as close to the edge as possible
- Filter as precisely as possible
- Two filtering approaches:
  - Explicit Permit (permit then deny any)
  - Explicit Deny (deny then permit any)



## **Prefix List**



- Lists of routes you want to accept or announce
- You can create them manually or automatically with data from IRRs
- It can be done using scripts or tools:
  - Filtergen (Level3)
  - bgpq4
  - IRRToolSet
  - IRR Power Tools

## Which Routes Should be Filtered?

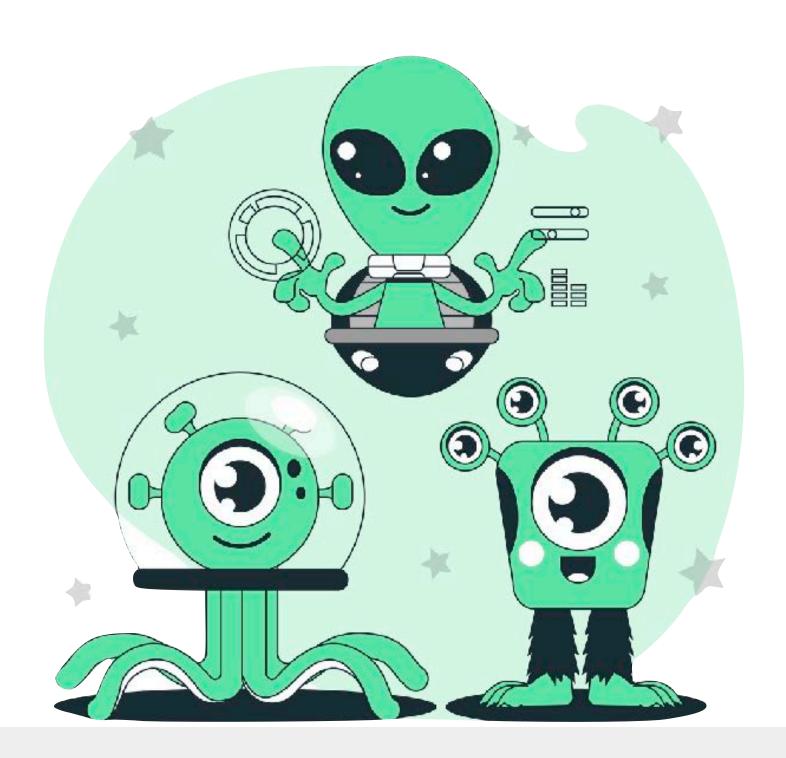


- Special-purpose prefixes (IPv4/IPv6) (Martians)
- Unallocated prefixes
- Routes that are too specific
- Prefixes belonging to the local AS
- IXP LAN prefixes
- The default route (0.0.0.0/0, ::/0)

# Special-purpose Prefixes



- Also known as Martians
  - RFC 1918 Private addresses
  - Reserved space (documentation, multicast, etc.)
- Not globally routable
  - Should be discarded on Internet BGP peering

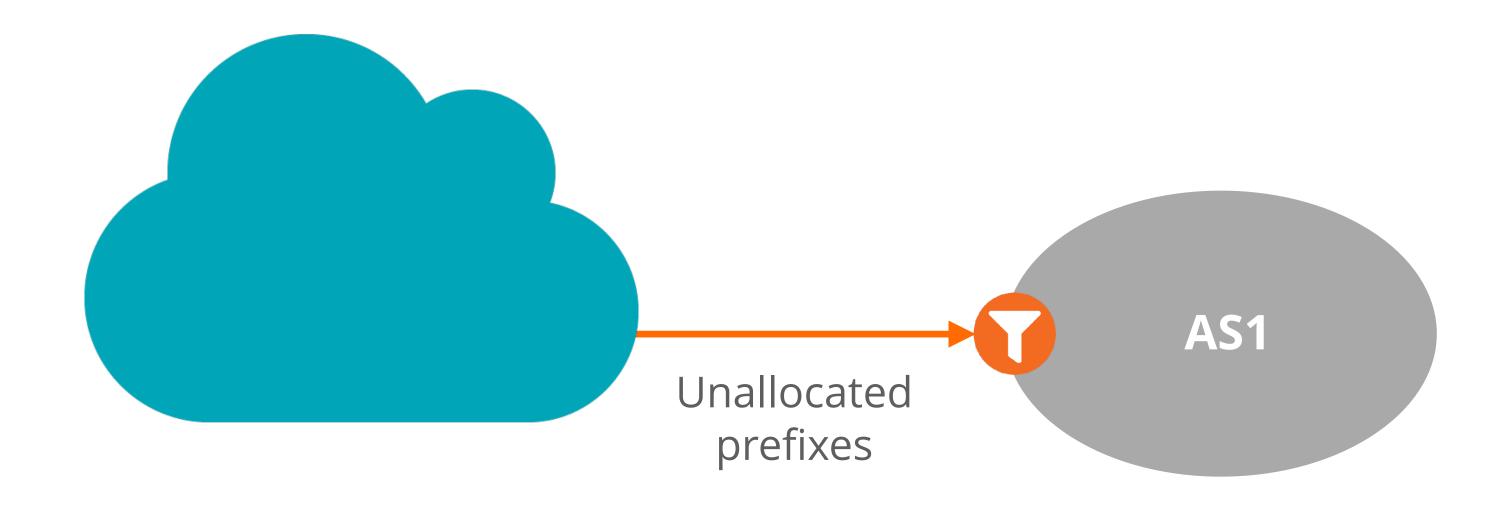


http://www.iana.org/assignments/iana-ipv4-special-registry http://www.iana.org/assignments/iana-ipv6-special-registry

## **Unallocated Prefixes**



- All unallocated prefixes should be filtered
  - Prefixes not yet allocated by IANA to RIRs (only for IPv6)
  - Prefixes allocated to an RIR but have not yet been distributed by an RIR to LIRs/End-users
- Filtering unallocated prefixes requires regular update



# Longest Accepted Prefixes



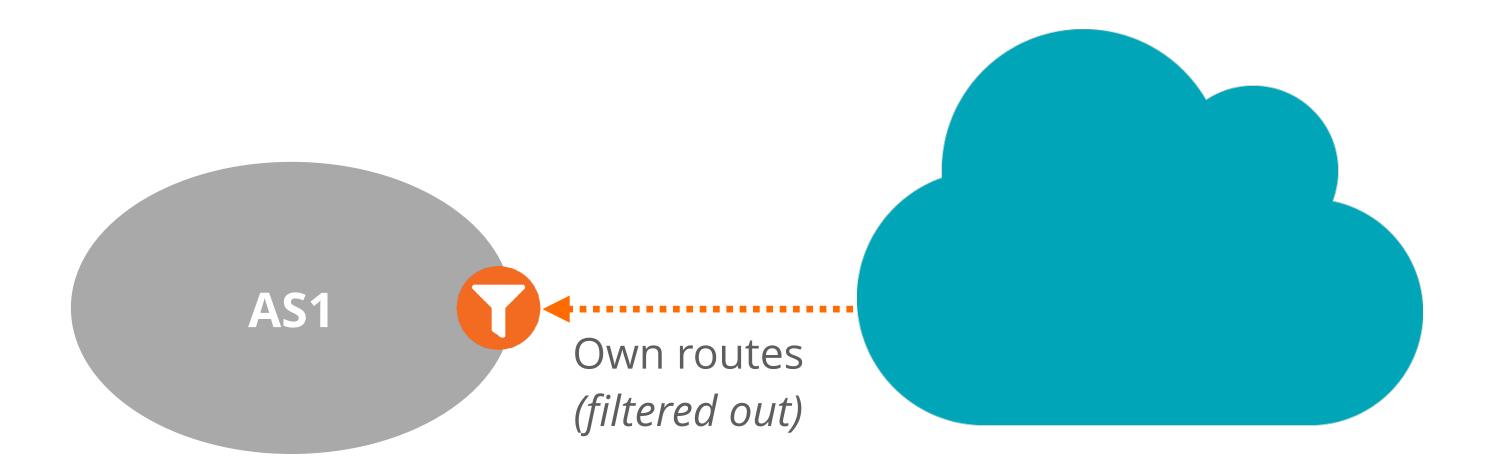
- Smaller prefixes should not be a part of global routing!
  - /24 for IPv4 (*RIPE-399*)
  - /48 for IPv6 (*RIPE-532*)
- Those prefixes are generally neither announced nor accepted on the Internet

ip prefix-list SMALL-V4 permit 0.0.0.0/0 le 24 ipv6 prefix-list SMALL-V6 permit 2000::/3 le 48

## Prefixes Belonging to the Local AS



- You should filter your own prefixes on all BGP peering
  - Prevents local traffic from leaking over an external peering
- Such filters can also be configured for downstream customers' prefixes
- In case of multi-homed customer, be careful not to break redundancy mechanism



## IXP LAN Prefixes

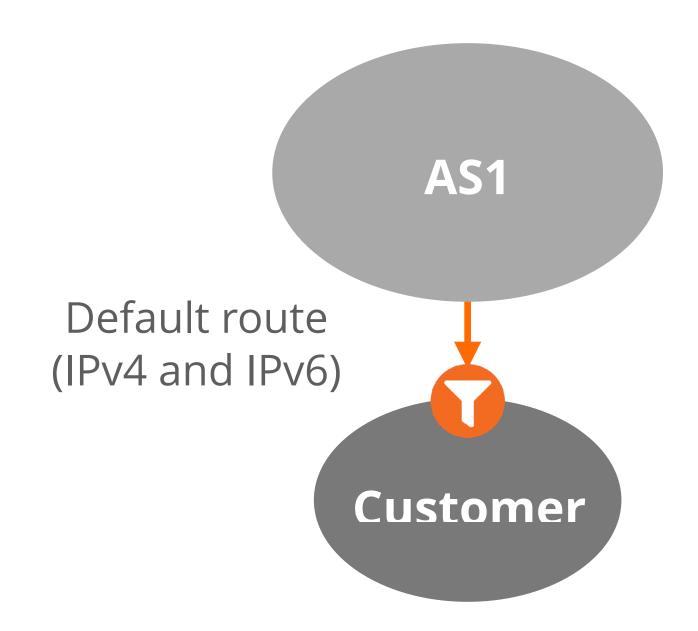


- IXP should originate its LAN prefix
  - Advertise it from its route server to all IXP members
- Do not accept IXP LAN prefix from any of your eBGP peers!
  - It may create a blackhole for connectivity to IXP LAN
- IXP prefix announcement should pass IRR-generated filters

## **Default Route**



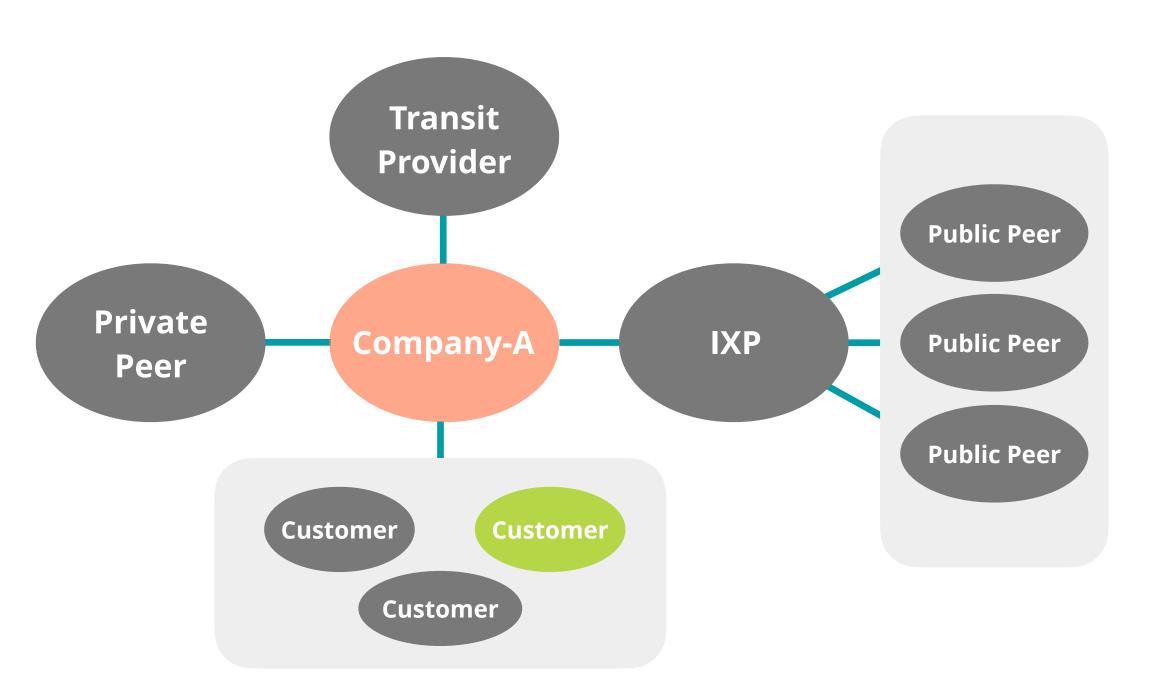
- 0.0.0.0/0 (IPv4) and ::/0 (IPv6)
- Advertised or accepted only in specific customer-provider peering relationships
  - E.g.customer with stub network
- Should be rejected unless a special peering agreement is in place



# Prefix Filtering Recommendations



- In full routing networks, some policies should be applied
  - On each BGP peer
  - For both received and advertised routes (inbound and outbound)
- Recommendations vary based on type of BGP peering relationships
  - Public and Private Peering
  - Transit Provider (Upstream)
  - Customer



# Prefix Filtering Recommendations



	With Public/Private Peers		With Transit Provider		With Customers		Leaf Customer Network	
	Inbound	Outbound	Inbound	Outbound	Inbound	Outbound	Inbound	Outbound
	Strict: Allow only IRR declared Loose: see below	Allow only own & customer's prefixes. Additionally: see below	Allow default only, or for Full routing table (FRT): see below	Allow only own & customer's prefixes. Additionally: see below	If known, allow only customer's prefixes. If not: see below	Allow default only, or for Full routing table (FRT): see below	Depends on agreement with upstream. If <b>default</b> only allow that. If <b>FRT</b> : see below	Only announce your own prefixes. Also filter:
Special Purpose Prefixes	X	X	X	X	X	X	X	X
Prefixes Not Allocated by IANA	X		X		X			
Too Specific Routes	X	X	X	X	X	X	X	X
Prefixes Belonging to the Local AS	X		X		X		X	
IXP LAN Prefixes	X	X	X	X	X			X
Default Route	X	X	Depends on needs / agreement	X	X	Depends on needs / agreement	Depends on needs / agreement	X



# Questions



# Lab 2 - Creating BGP Prefix Filters



• **Description**: Configure BGP prefix filters with different types of peers.

### Goals:

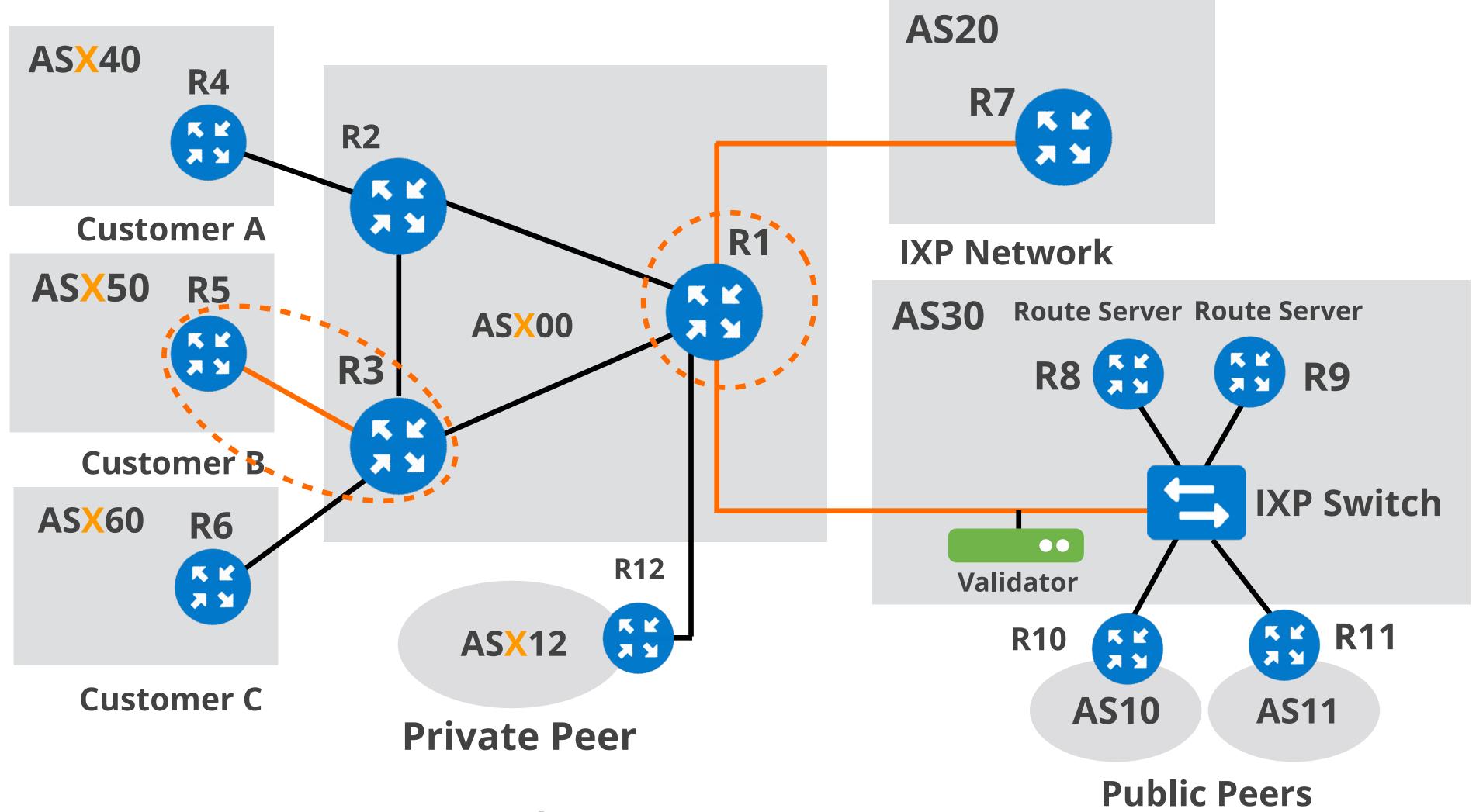
- Define BGP filter recommendations based on the routing relationships
- Choose the appropriate methods for implementing BGP filters
- Time: 40 minutes

### Tasks:

- 2.1 Configure prefix filters with Transit Providers and IXP Peers
- 2.2 Configure prefix filters with Customers
- (OPTIONAL) 2.3 Configure filters using communities with a private peer

## **Transit Provider**



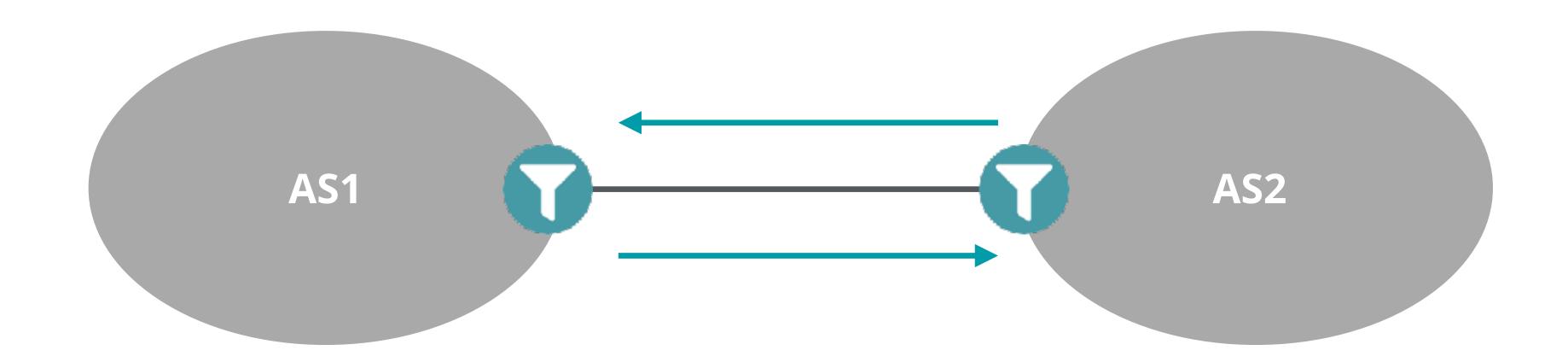


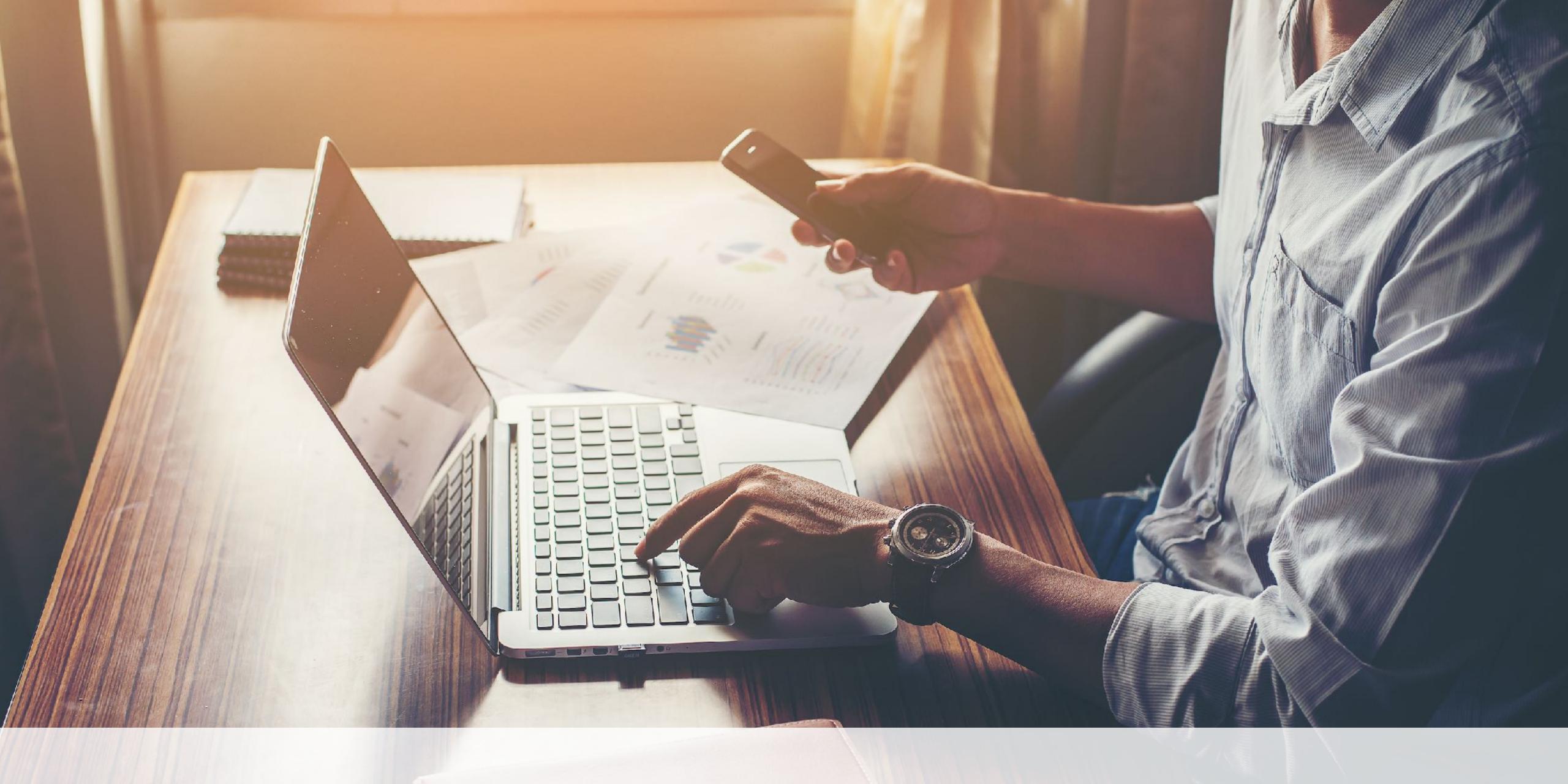
Your AS number AS X00 Your IPv6 allocation 2001:db8:X00::/48

## Lab 2 - Creating BGP Prefix Filters



- What have you learned?
  - Filtering rules are different for different types of peering relationships (Customers, Transit Providers, Peers)
  - Inbound and outbound filtering rules are different





After the Prefix Filtering, we continue with more filtering...





- Filters routes based on AS path
  - Permit or deny prefixes from certain ASes

```
router bgp 65564
network 10.0.0.0 mask 255.255.255.0
neighbor 172.16.1.1 remote—as 65563
neighbor 172.16.1.1 filter—list 1 out
neighbor 172.16.1.1 filter—list 2 in

ip as—path access—list 1 permit ^65564$
ip as—path access—list 2 permit ^65563$
```

# AS Path Filtering Recommendations



- From your customers accept only:
  - AS paths containing ASNs belonging to (or authorised to transit through) the customer

### Do not accept:

- Prefixes with private AS numbers in the AS path (unless from customers)
- Prefixes when the first AS number in the AS path is not the one of the peer's (unless towards a BGP route server)



## **AS Path Filtering Recommendations**



- Do not advertise:
  - Prefixes with a nonempty AS Path (unless you intend to provide transit for these prefixes)
  - Prefixes with upstream AS numbers in the AS Path to your peers (unless you intend to provide transit)
  - Private AS Paths (unless there is a special "private" arrangement with your peers)

### Do not override BGP's default behaviour

Do not accept your own AS in the AS-path



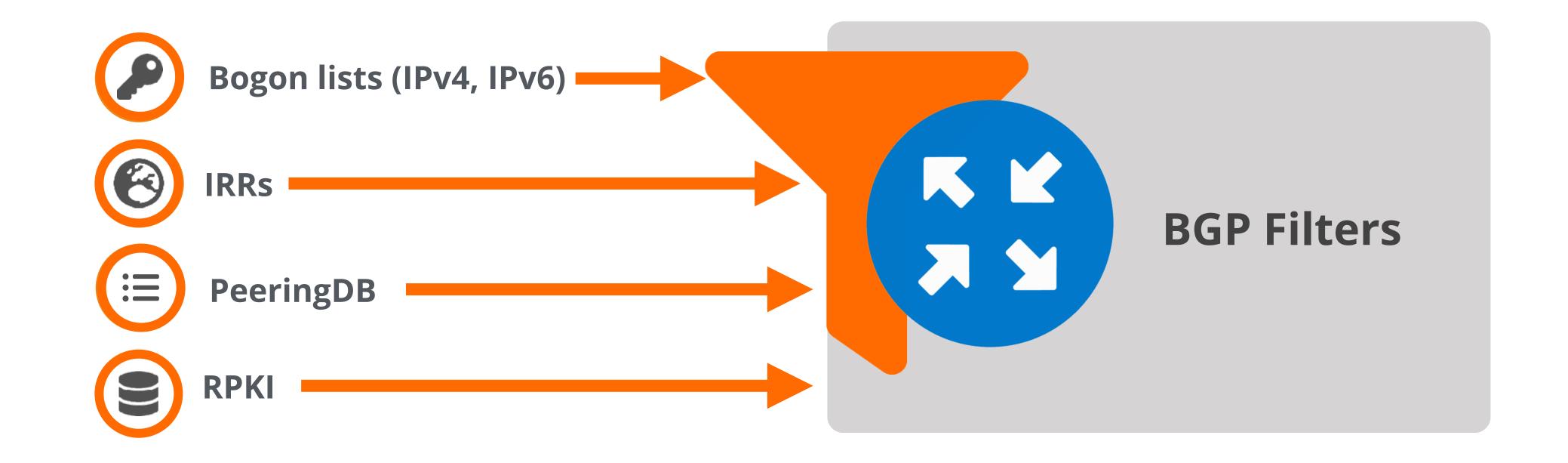
# **BOGON ASN filtering**



ASNs	Status	RFC
0	Reserved	RFC7607
23456	AS_TRANS	RFC6793
64496-64511	Reserved for use in docs and code	RFC5398
64512-65534	Reserved for Private Use	RFC6996
65535	Reserved	RFC 7300
65536-65551	Reserved for use in docs and code	RFC5398
65552-131071	Reserved	IANA
420000000 - 4294967294	Reserved for Private Use	RFC6996
4294967295	Reserved	RFC 7300

## Which Data Sources Can You Use for BGP Filters?





## **BOGON Lists**



- BOGONs are prefixes that should never appear in the Internet routing table
  - Martians (RFC1918 Private addresses + Reserved space)
  - IANA unallocated space
- Full BOGON should be filtered as well
  - BOGONs + prefixes unallocated by RIRs
- The BOGON and full BOGON lists are **not static**

## How to Get the List of (Full) BOGONs?



- Team Cymru provides lists of BOGONs and full BOGONs
- Offers variety of formats and methods
  - HTTP
  - BGP Peering (Bogon Route Server Project)
  - Routing Registries (RADB)
  - DNS

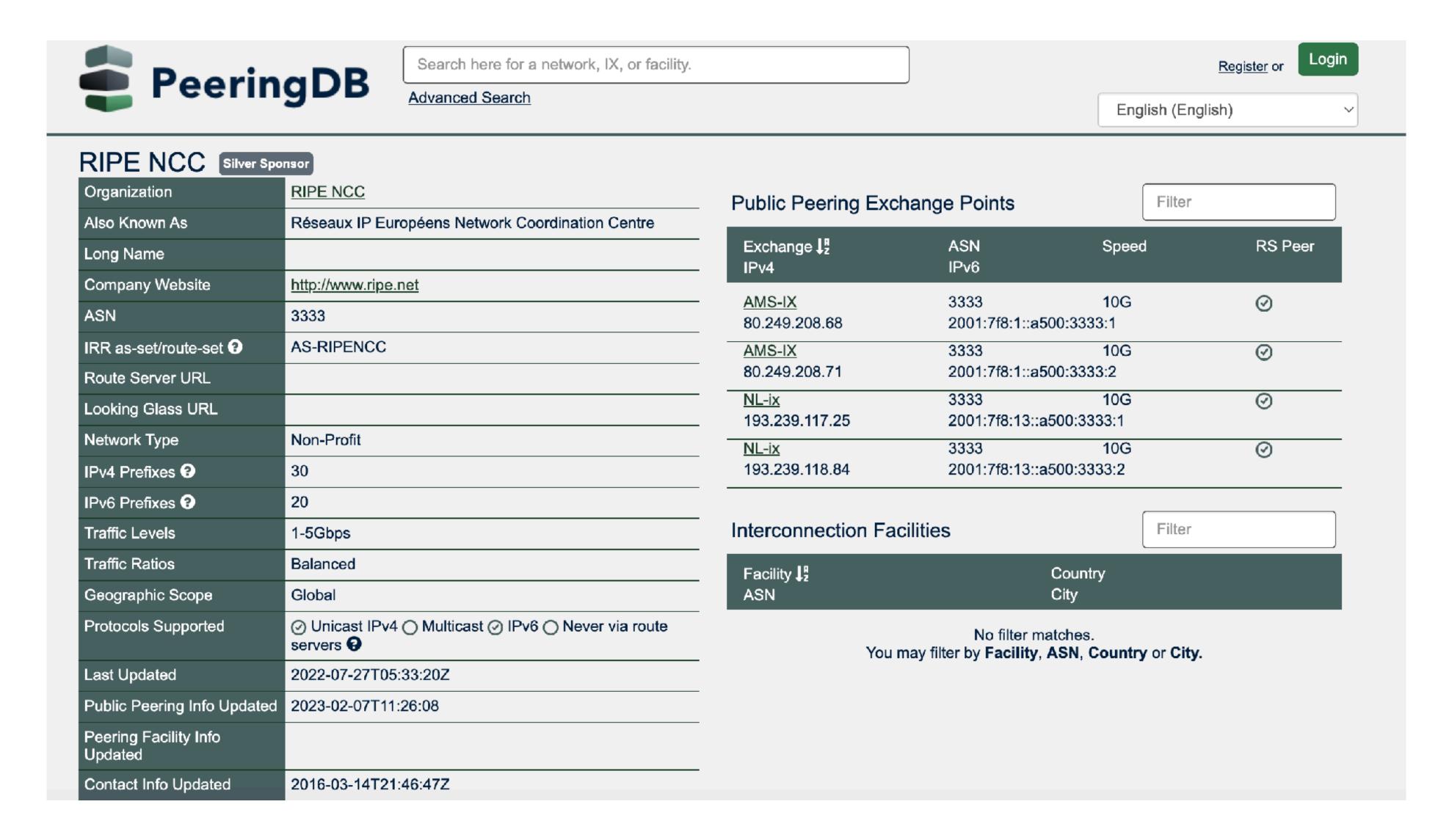
# PeeringDB



- Web-based public database for BGP peering
- Non-profit, community-driven initiative, run and promoted by volunteers
- First stop when making interconnection decisions
  - Default location for Internet peering data
  - Helps to decide where and whom to peer with
  - Provides contact info
  - Gives information about the peering policy

## PeeringDB





## Other Recommended Filtering



- Other methods to control BGP routes:
  - Max-prefix filtering

- BGP Route Flap Dampening
- Next-hop Filtering
- Optional BGP Community Scrubbing





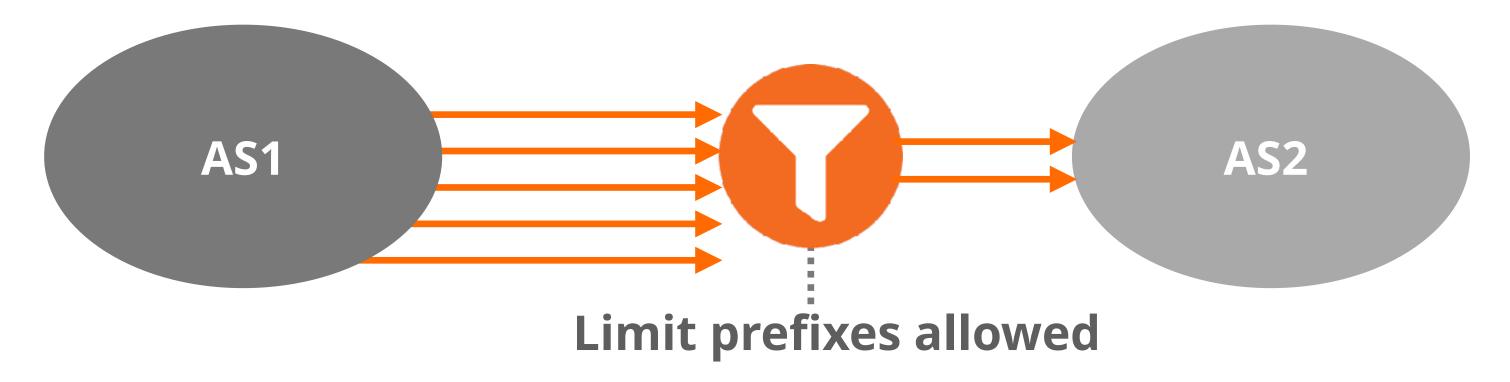
**BGP Security** 

https://academy.ripe.net/bgp-security/

## **Max-Prefix Filtering**



- From peers:
  - Set limit lower than the number of Internet routes
  - Different per peer based on expected number of routes
- From upstreams that provide full route:
  - Set limit higher than the number of Internet routes
  - Limit should be decided based on router's capacity
- Regularly review the limits









## Lab 3 - Filtering AS-Path/Number of Prefixes



Description: Configure AS-path filters and limit the number of accepted prefixes

#### Goals:

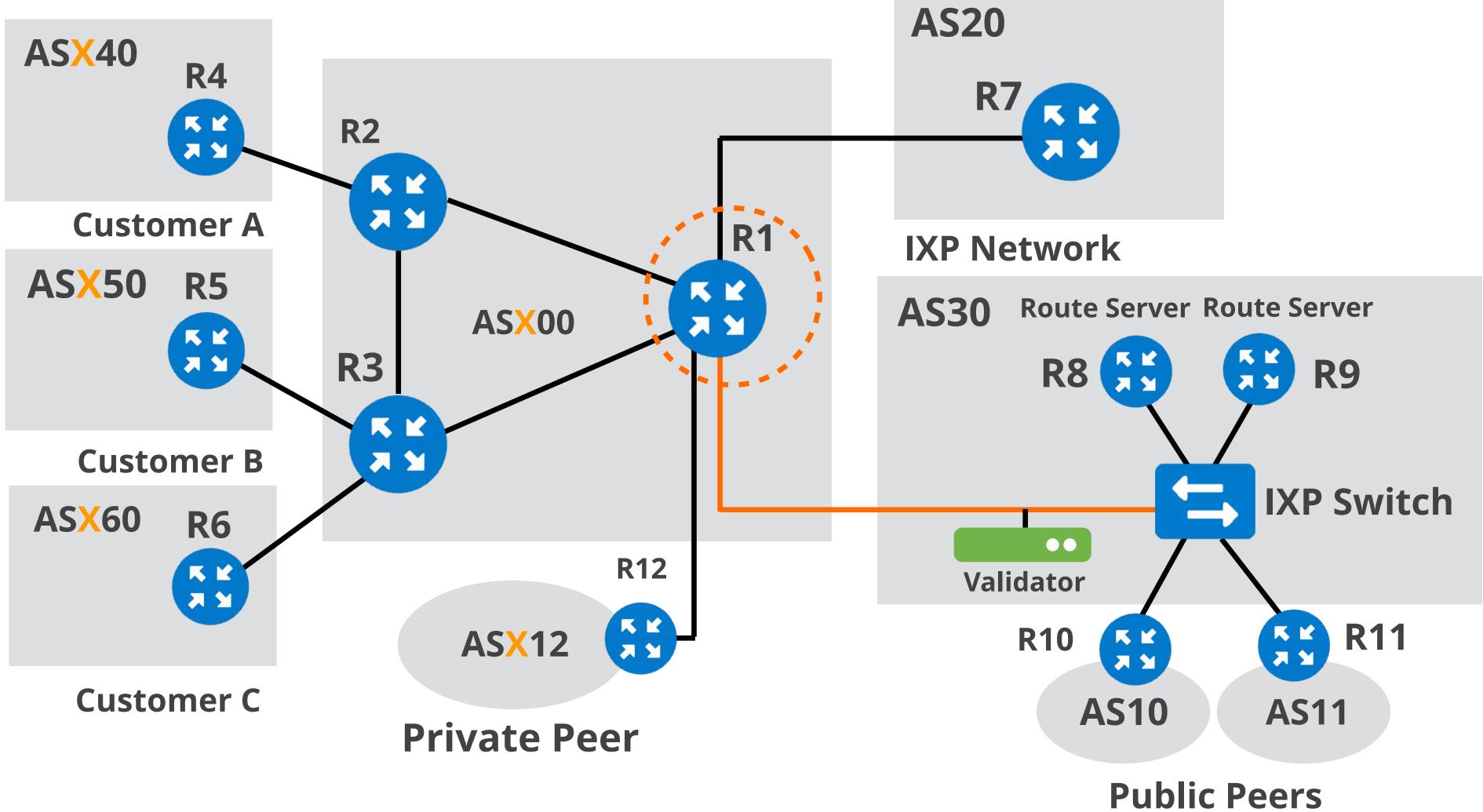
- Create consistent AS-path filter to secure a BGP network
- Create BGP prefix-filter to discard more specifics
- Time: 20 minutes

#### Tasks:

- 3.1 Configure BGP filters based on AS-path information
- 3.2 Configure a limit on the number of accepted BGP prefixes

#### **Transit Provider**





Your AS number AS X00 Your IPv6 allocation 2001:db8:X00::/48

#### Lab 3 - Filtering AS-Path/Number of Prefixes



#### What have you learned?

- You can filter routes based on the ASNs included in the AS-PATH
- Limiting the number of routes accepted in a BGP peering can avoid resource exhaustion
- Reaching the limit of accepted routes takes down the peering



## Registering in the IRR System

Section 3.4

## **IRR Support Routing Security**



- The Internet Routing Registry (IRR) composed by many databases:
  - RIPE NCC, APNIC, RADB, JPIRR, Level3, NTTCom, etc.
- Their information can be used to:
  - Improve stability and consistency of routing
  - Provide global view of routing policies
  - Automation of creating BGP filters
  - Network Troubleshooting

## Why Register Routing Information?





- Document your routing policy
  - Associate network prefixes with an origin AS



- Helps to filter unauthorised announcements
  - Mitigates route hijacks and denial-of-service



- Many transit providers and IXPs require it
  - They build their filters based on the Routing Registry

## The RIPE Routing Registry



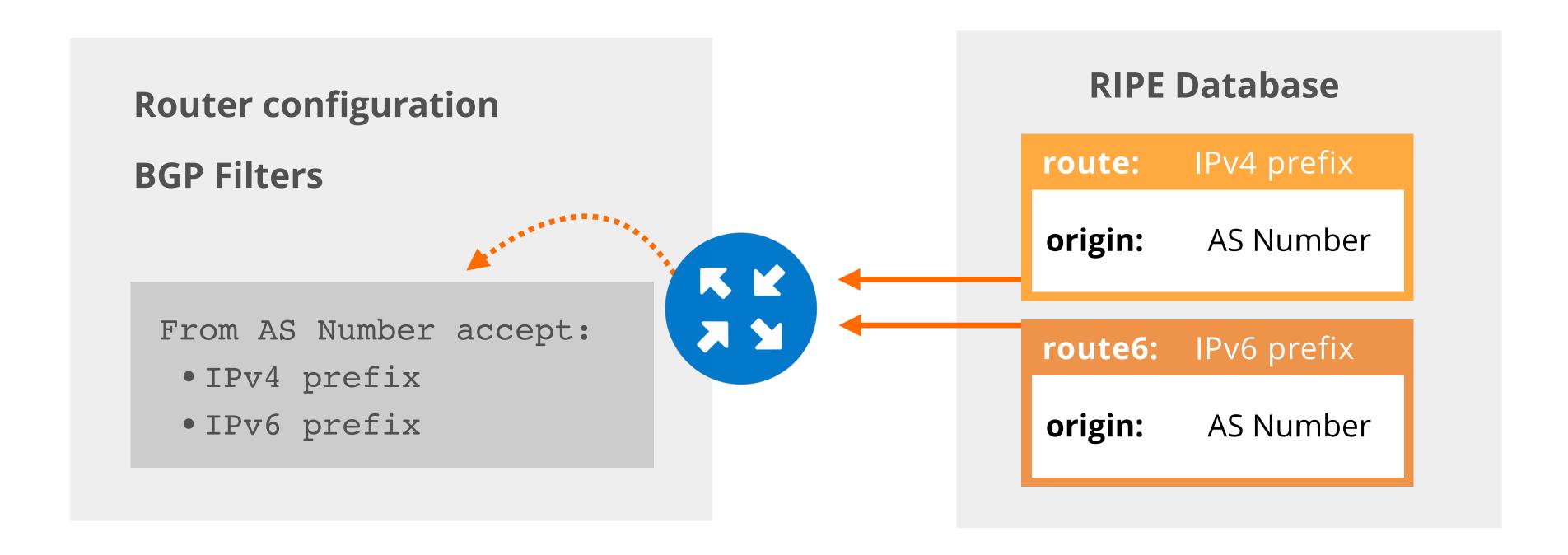
- A subset of the RIPE Database and part of the global IRR
- Used for registering routing policy information
- Includes several objects route aut-num inet-rtr route6 **Set Objects** peering-set filter-set as-set route-set rtr-set

**Source:** https://apps.db.ripe.net/docs/04.RPSL-Object-Types/





- Contains routing information for IPv4/IPv6 address space
- Specifies from which AS a certain prefix may be originated
- Used for creating BGP filters



#### **Authorisation Rules for Route(6)**



- You need permission from:
  - 1. inetnum or inet6num
  - 2. route or route6

Allocation

mnt-by: RIPE-NCC-HM-MNT
mnt-by: DEFAULT-LIR-MNT
mnt-routes: ANOTHER-MNT

route(6)

origin: AS65550
mnt-by: ANOTHER-MNT

<sup>\*</sup> mnt-routes delegates the creation of route(6) objects

#### Registering IP Routes



2002:ff30::/32 inet6num: inetnum: 10.30.0.0 - 10.30.3.255 mnt-by:TEST-NCC-HM-MNT mnt-by:TEST-NCC-HM-MNT mnt-by:SM30-MNT mnt-by:SM30-MNT route6: 2002:ff30::/32 10.30.0.0/22 route: origin: **AS65550** origin: **AS65550** mnt-by: **SM30-MNT** mnt-by: SM30-MNT

#### **Aut-num**



aut-num: AS64500

as-name: YOUR-AS-NAME

org: ORG-EE2-RIPE

import: from AS65550 accept ANY

export: to AS65550 announce AS64500

import: from AS64496 accept ANY

export: to AS64496 announce AS64500

admin-c: DV789-RIPE

tech-c: JS123-RIPE

status: ASSIGNED

mnt-by: RIPE-NCC-END-MNT

mnt-by: DEFAULT-LIR-MNT

source: RIPE

Registers who holds that AS Number

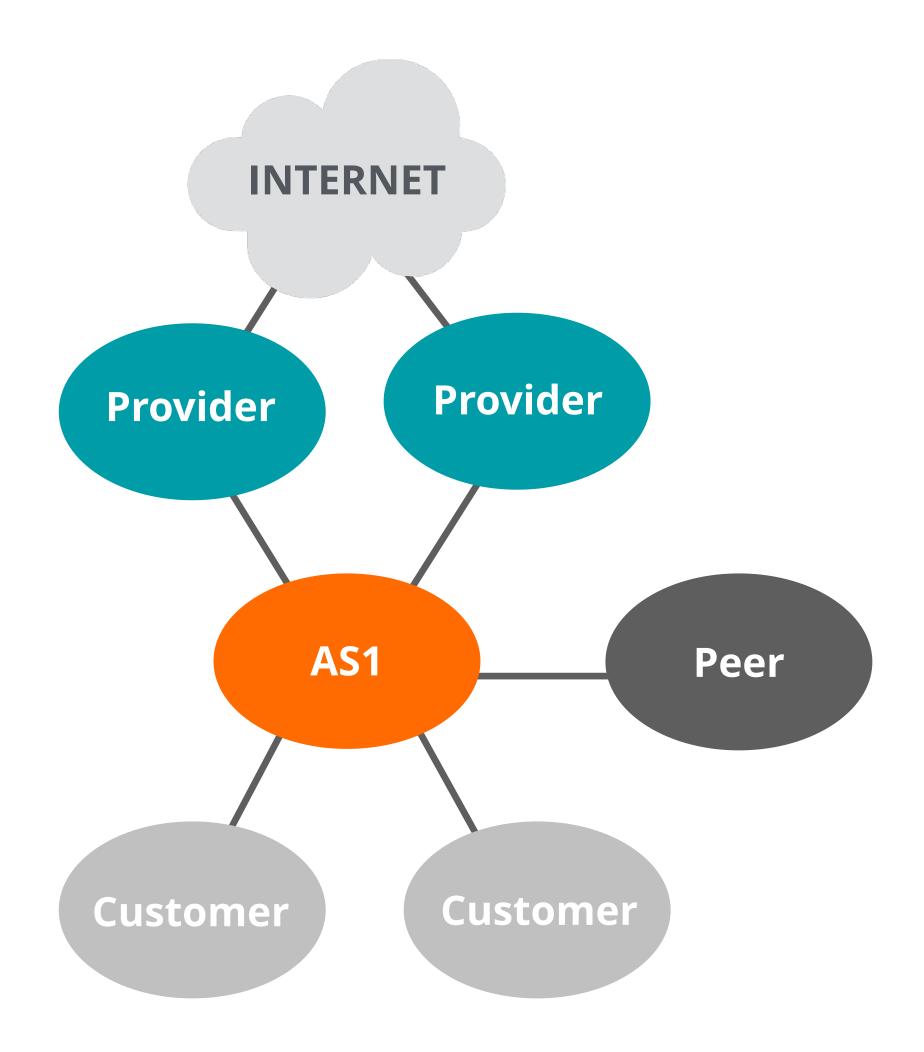
Defines the routing policy for an AS

Import - specifies which routes you accept

• Export - specifies which routes you announce

## **BGP Routing Policy**

- Who are your BGP peers? Which ASes
- What is your BGP relationship with them?
  - Customer, Provider, Peer
- What are your routing decisions?
  - Which prefixes to accept?
  - Which prefixes to announce?
  - Which prefixes will be preferred in case of multiple routes?



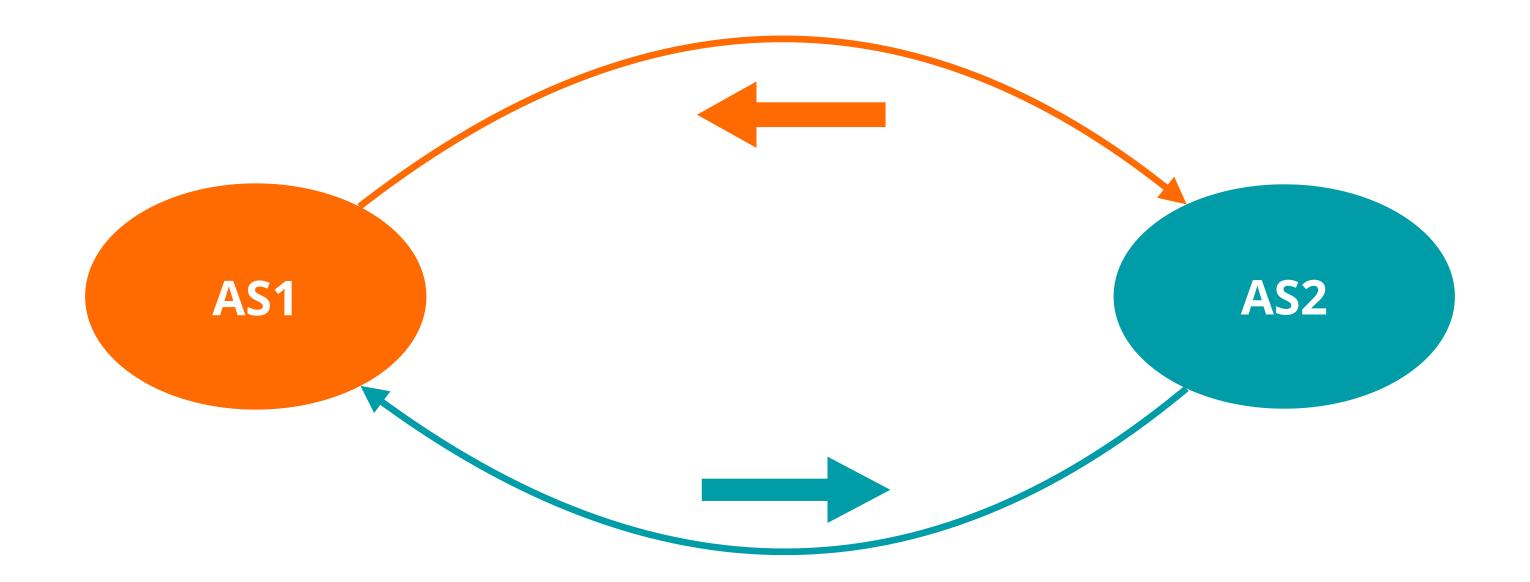
#### IRRs use RPSL Language



- RPSL Routing Policy Specification Language
- Allows network operators to specify their routing policies
  - Generic way to describe BGP configuration in the IRR
  - Not vendor-specific
- Originated from a RIPE Document (RIPE-181)
- Can be translated into router configuration

## Defining Routing Policy in RPSL





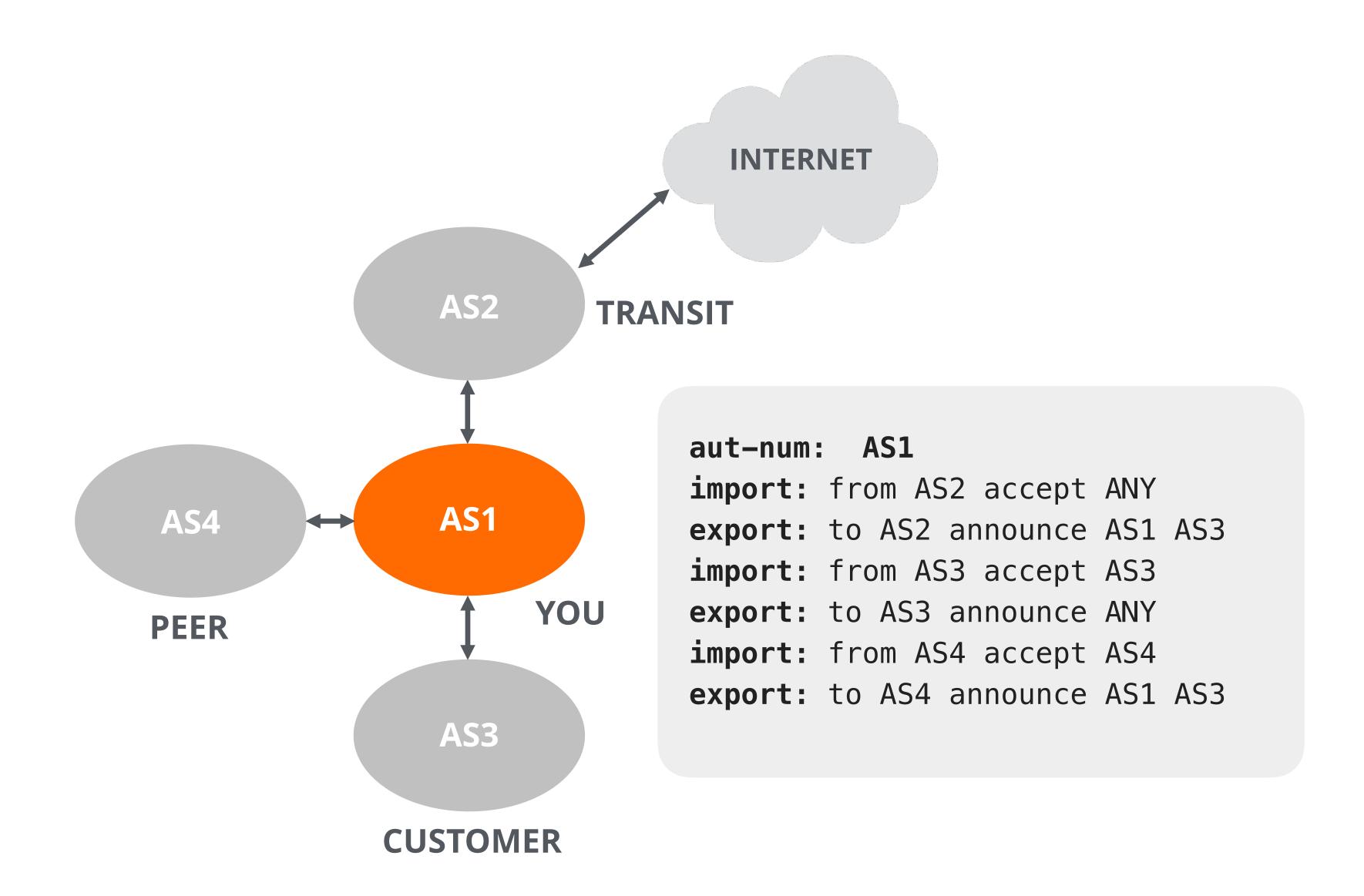
aut-num: AS1

import: from AS2 accept AS2

export: to AS2 announce AS1

#### Routing Policy Example





#### RPSLng



- **RPSL is older** than IPv6, the defaults are IPv4
- IPv6 was added later using a different syntax
- You have to specify that it's IPv6

aut-num: AS1

mp-import: afi ipv6.unicast from AS201 accept AS201
mp-export: afi ipv6.unicast to AS201 announce ANY

route-set: rs-customers members: 192.0.2.0/24

mp-members: 2001:db8:abcd::/48

#### Tools to check IRR status





11 records found for AS3333				
193.0.0.0/21	~	RIPE	$\odot$	✓ maxLength: 21
193.0.10.0/23	<b>✓</b>	RIPE	$\odot$	✓ maxLength: 23
193.0.12.0/23	✓	RIPE	$\odot$	✓ maxLength: 23
193.0.18.0/23	<b>✓</b>	RIPE	$\odot$	✓ maxLength: 23
193.0.20.0/23	<b>✓</b>	RIPE	$\odot$	✓ maxLength: 23
193.0.22.0/23	<b>✓</b>	RIPE	$\odot$	✓ maxLength: 23
193.230.194.0/24				✓ maxLength: 24
2001:610:240::/42				✓ maxLength: 42
2001:67c:2e8::/48	~	RIPE	$\odot$	✓ maxLength: 48



https://irrexplorer.nlnog.net/

https://stat.ripe.net

## Reality Check



- The IRR system has limitations
  - Conflicting data, no central authority, no holdership checks, not updated
- It is still widely used
- Improving IRR accuracy
  - Keep your IRR information up to date
  - Route filtering using IRRdv4 (validates against RPKI)
  - IRR databases should remove inconsistent records regularly







# Questions



## Lab 4 - Creating route(6) Objects



Description: Create a route6 object in the RIPE Test Database

#### Goals:

Register routing information in the RIPE Database

• Time: 15 minutes

#### Tasks:

- Create a RIPE NCC Access account (if you don't have one)
- Search for your IPv6 allocation and AS number
- Create a route6 object for your allocated IPv6 prefix

#### Lab 4 - Creating route(6) Objects



- What have you learned?
  - Which RIPE DB objects contain IPv6 routing information
  - How RIPE DB protects its objects and the accuracy of data

#### route(6)

origin: AS65550

mnt-by: ANOTHER-MNT

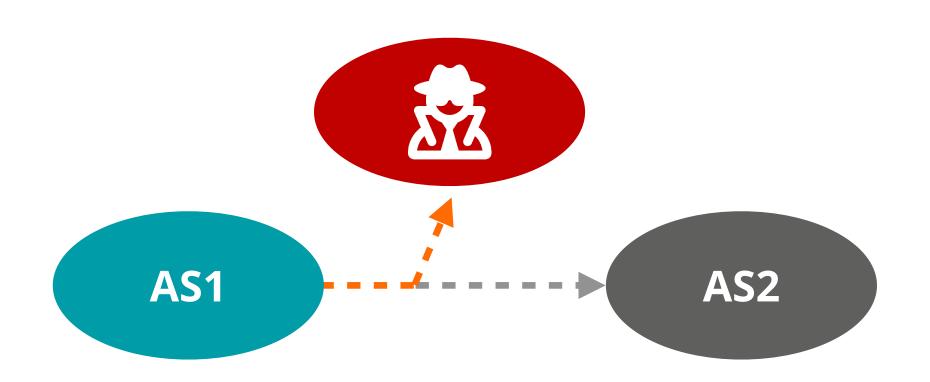


Section 3.5

#### **BGP Origin Hijacks**

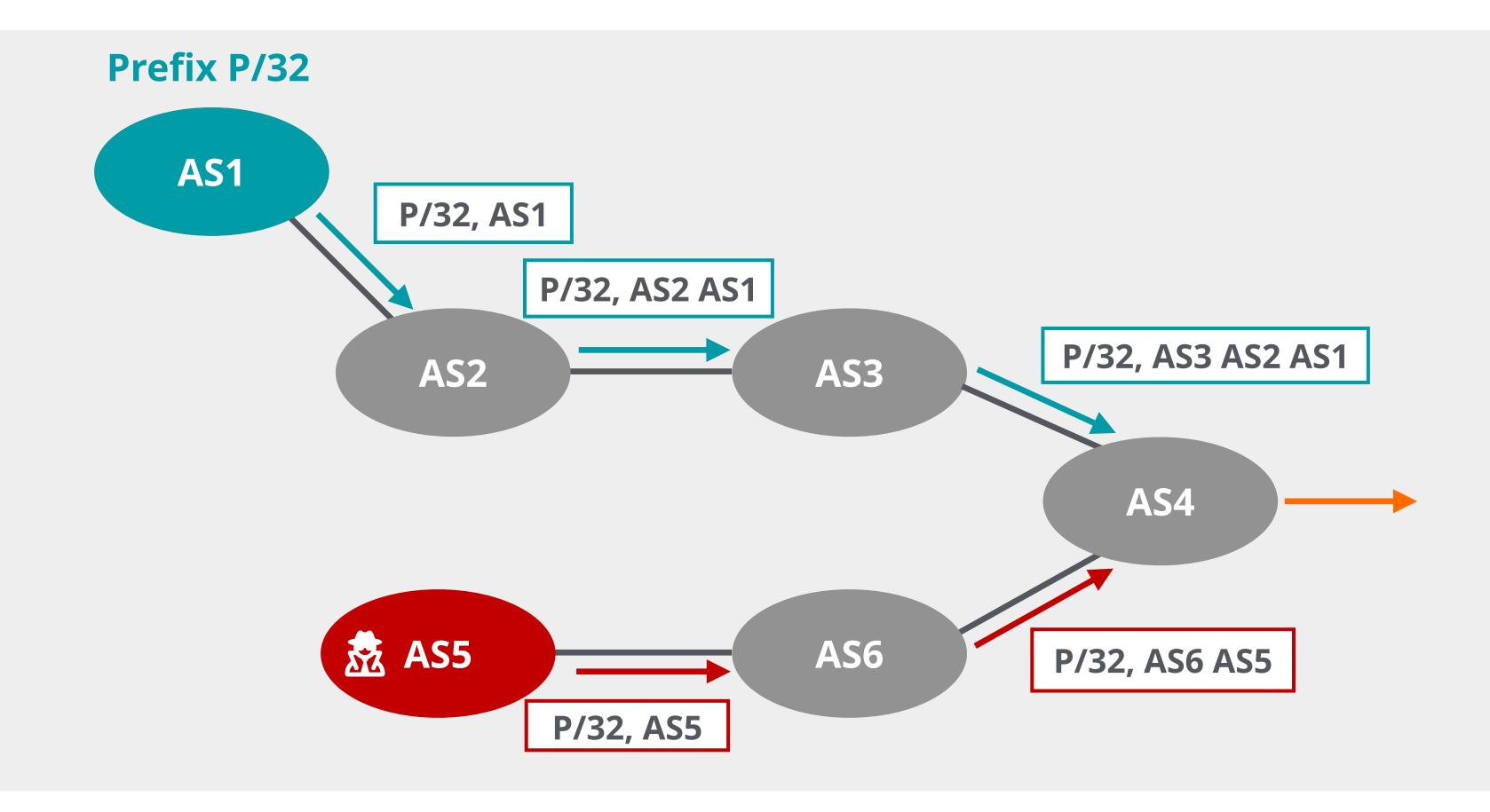


- An AS originates a prefix that is not authorised to originate
- Hijacker impersonates the legitimate holder
  - May hijack an allocated or unallocated address space
- It may announce the exact same prefix or more specifics
  - Prefix Hijack
  - Sub-prefix Hijack (De-aggregation hijack or subnet attack)



## Prefix Hijack



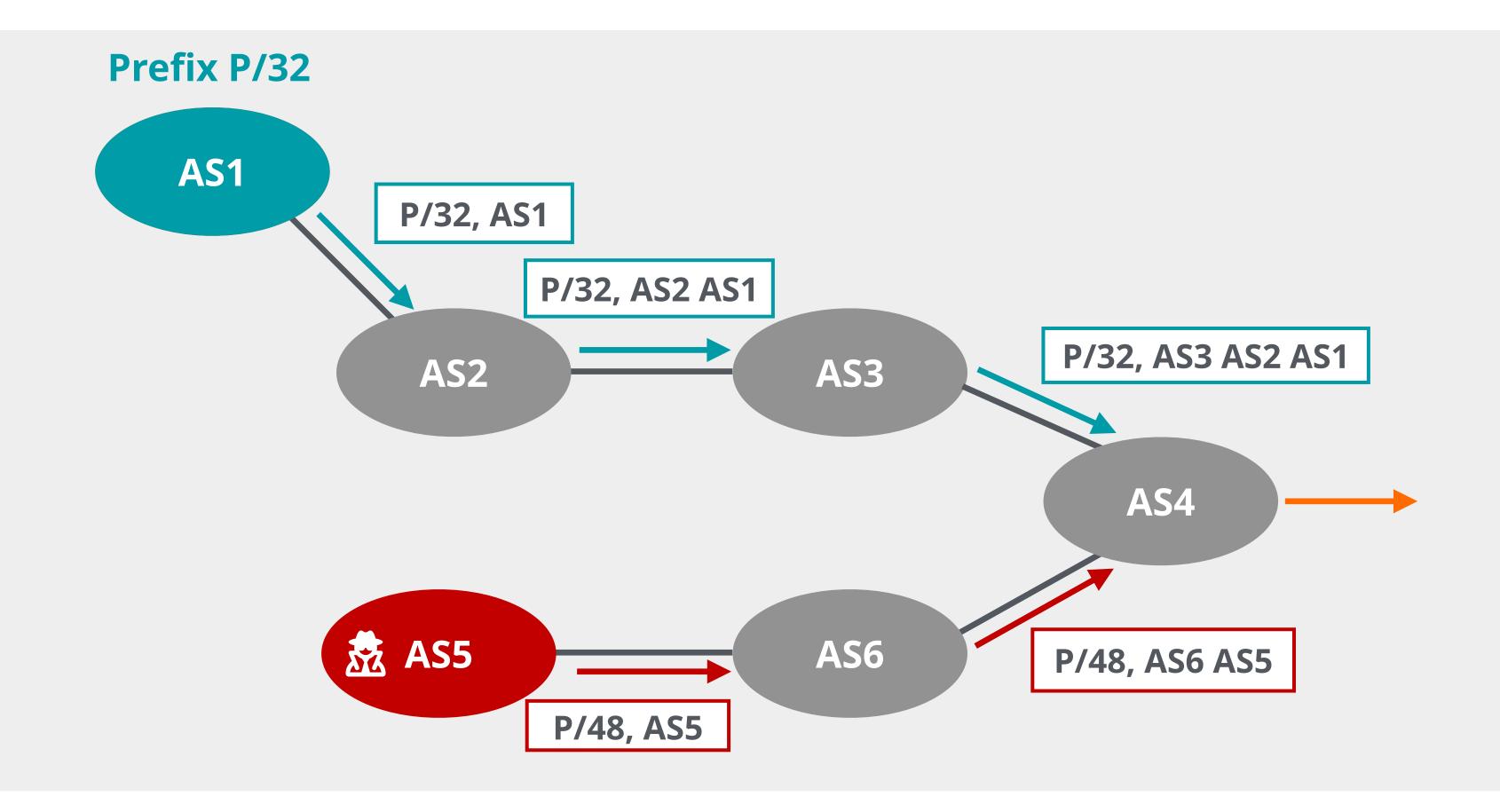


This is a **local hijack!** 

Only some networks are affected based on BGP path selection process

#### Sub-prefix Hijack (Subnet Attack)





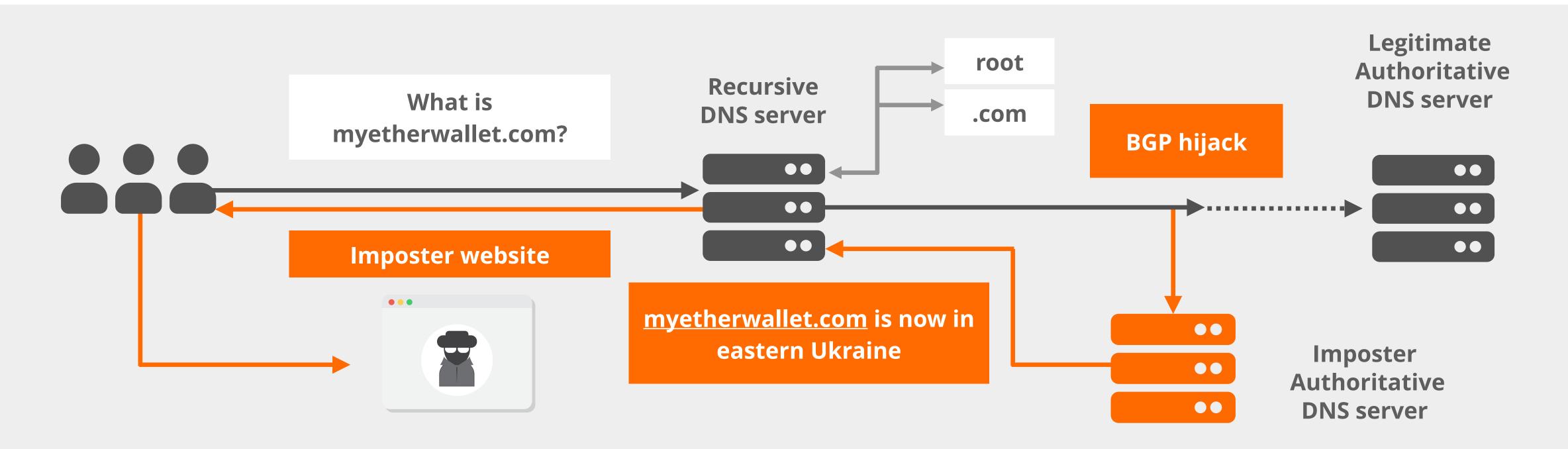
This is a **global hijack!** 

All traffic for more specific prefix will be forwarded to the hijacker's network

## April 2018: Amazon MyEtherWallet



- BGP hijack of Amazon DNS
- What happened?
- Why?
  - Attack to steal cryptocurrency



#### What is RPKI?

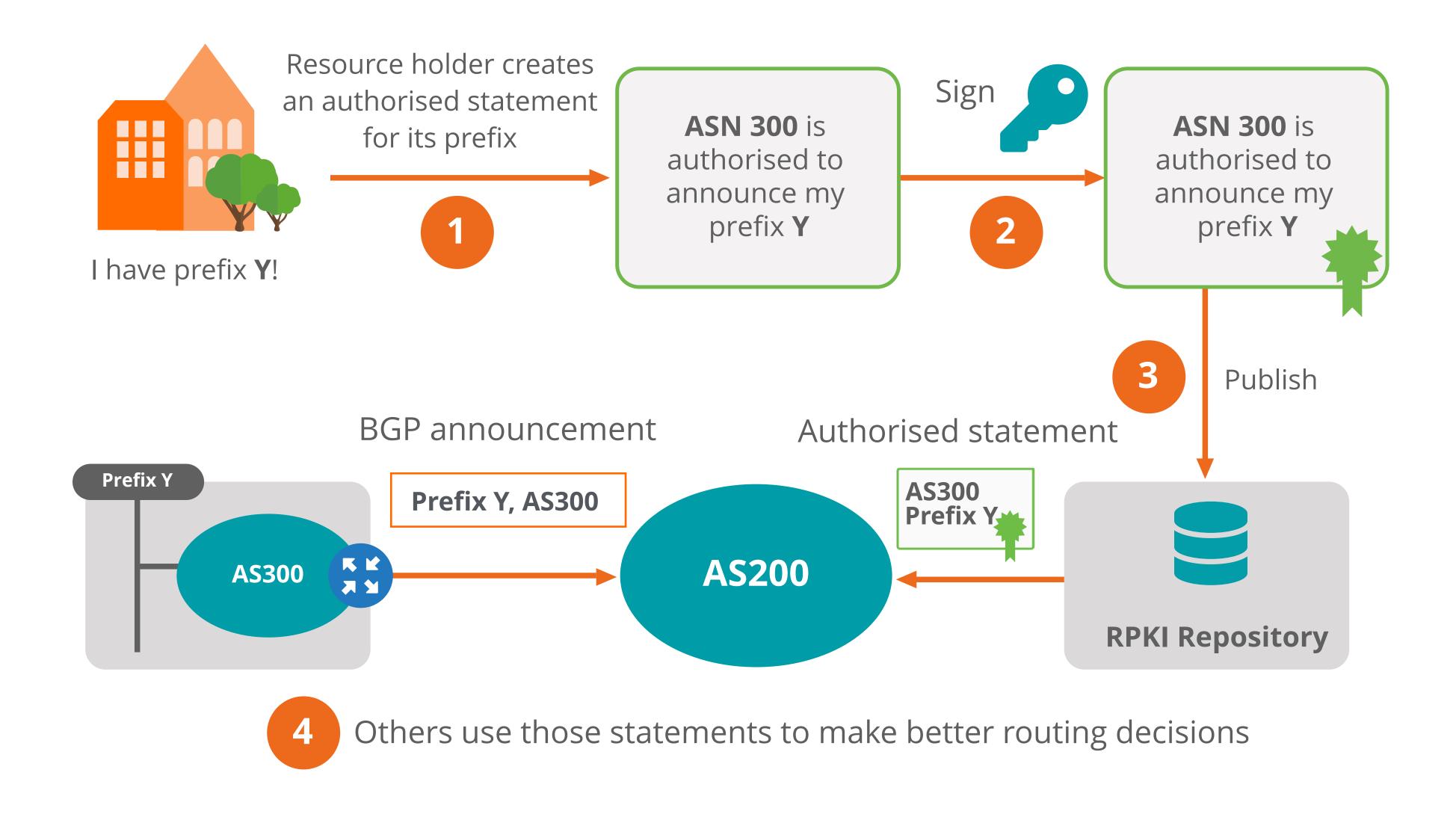


- A security framework for the Internet
- Verifies the association between resource holders and their resources
  - Attaches digital certificate to IP addresses and AS numbers
- Used to validate the origin of BGP announcements (BGP OV)
  - Is the originating ASN authorised to originate a particular prefix?
  - Helps to mitigate BGP Origin Hijacks and Route leaks



#### **How Does RPKI Work?**

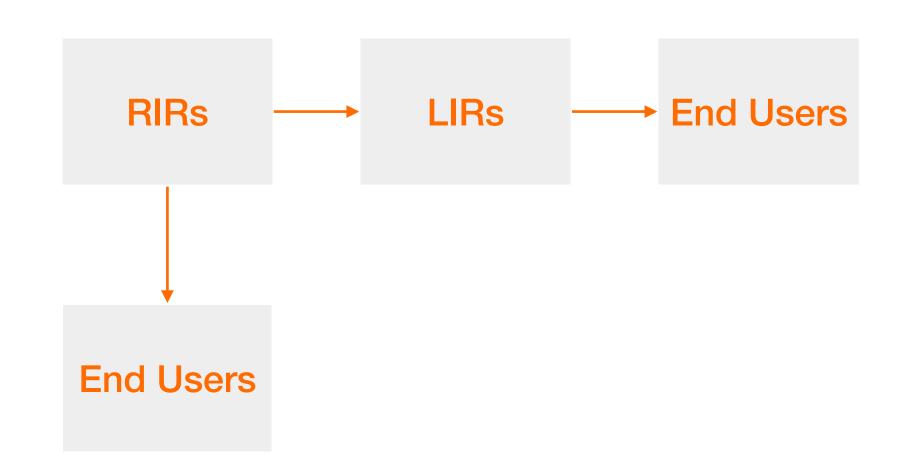


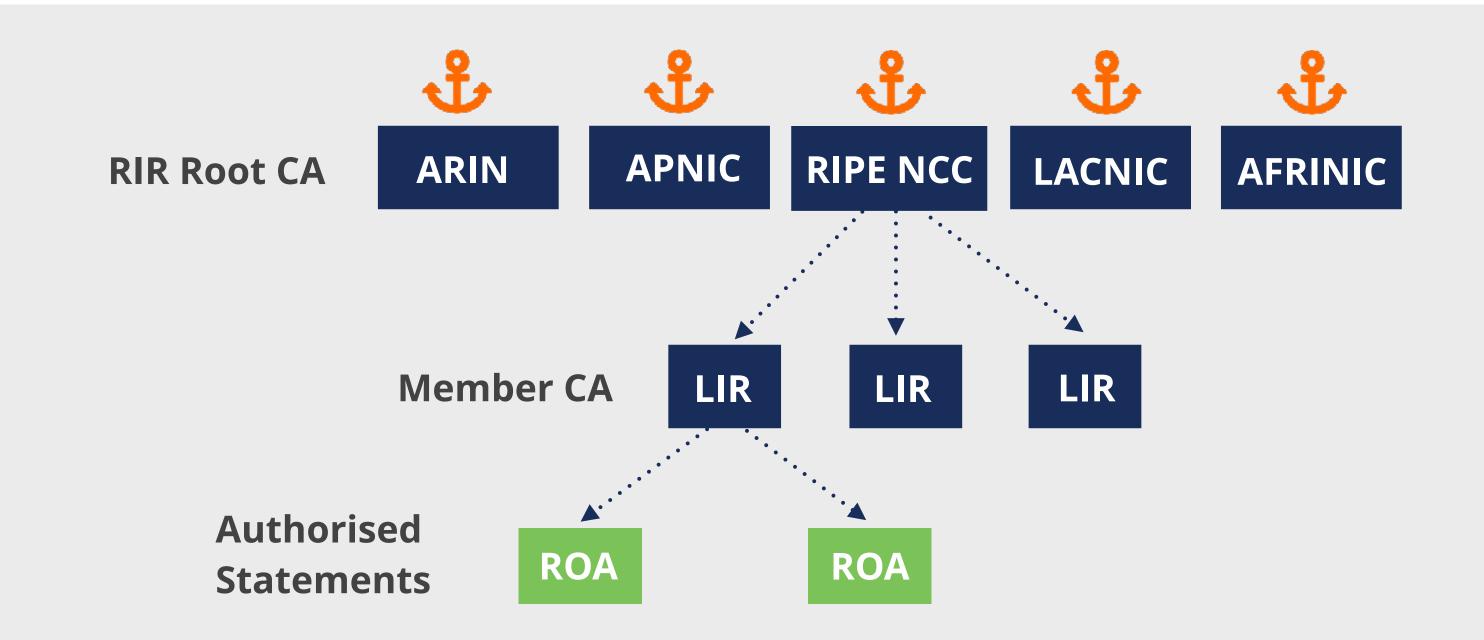


#### **Trust in RPKI**



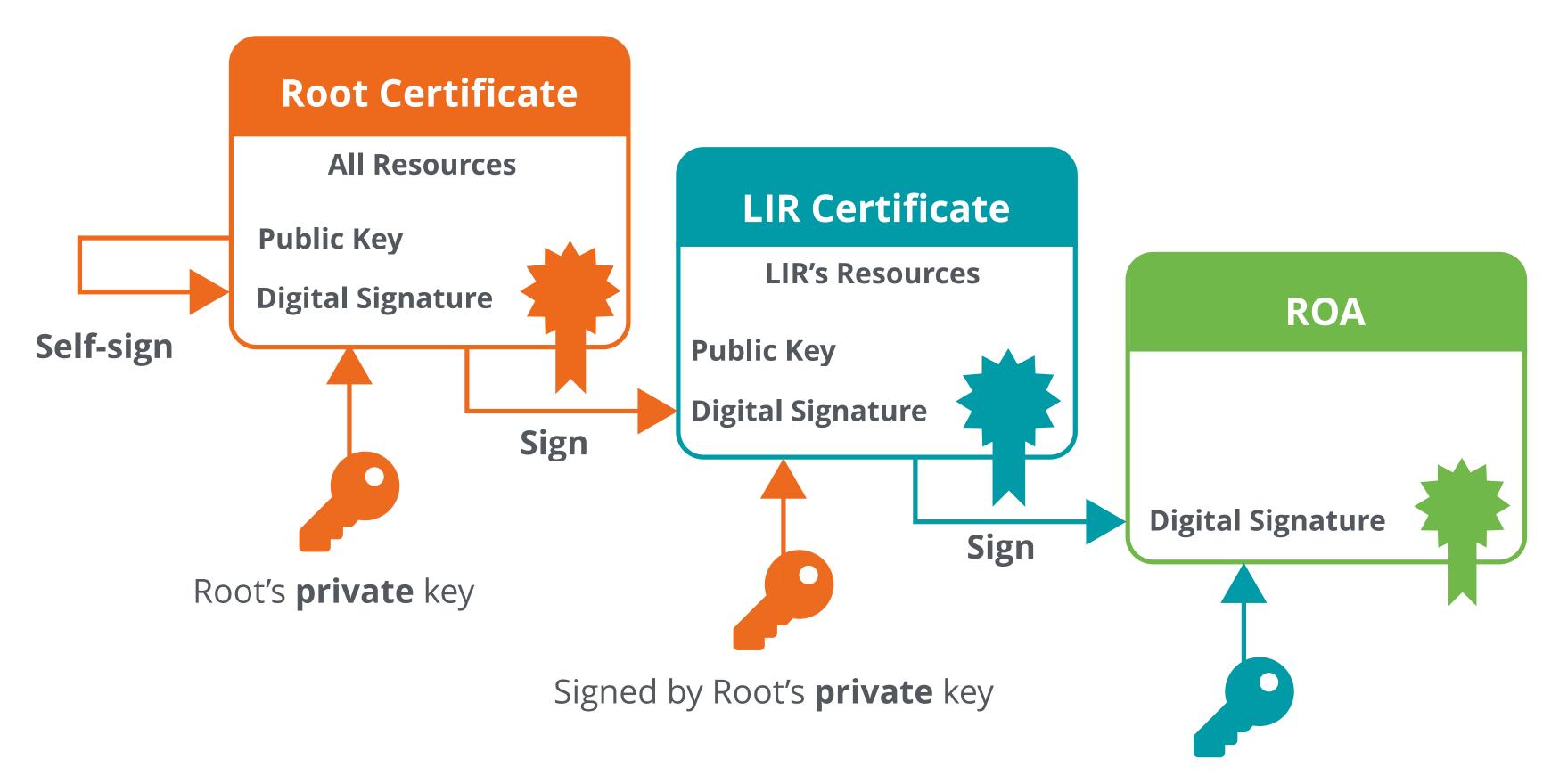
- RPKI relies on five RIRs as Trust Anchors
- Certificate structure follows the RIR hierarchy
- RIRs issue certificates to resource holders





#### **RPKI Chain of Trust**





Signed by LIR's **private** key

#### **Elements of RPKI**

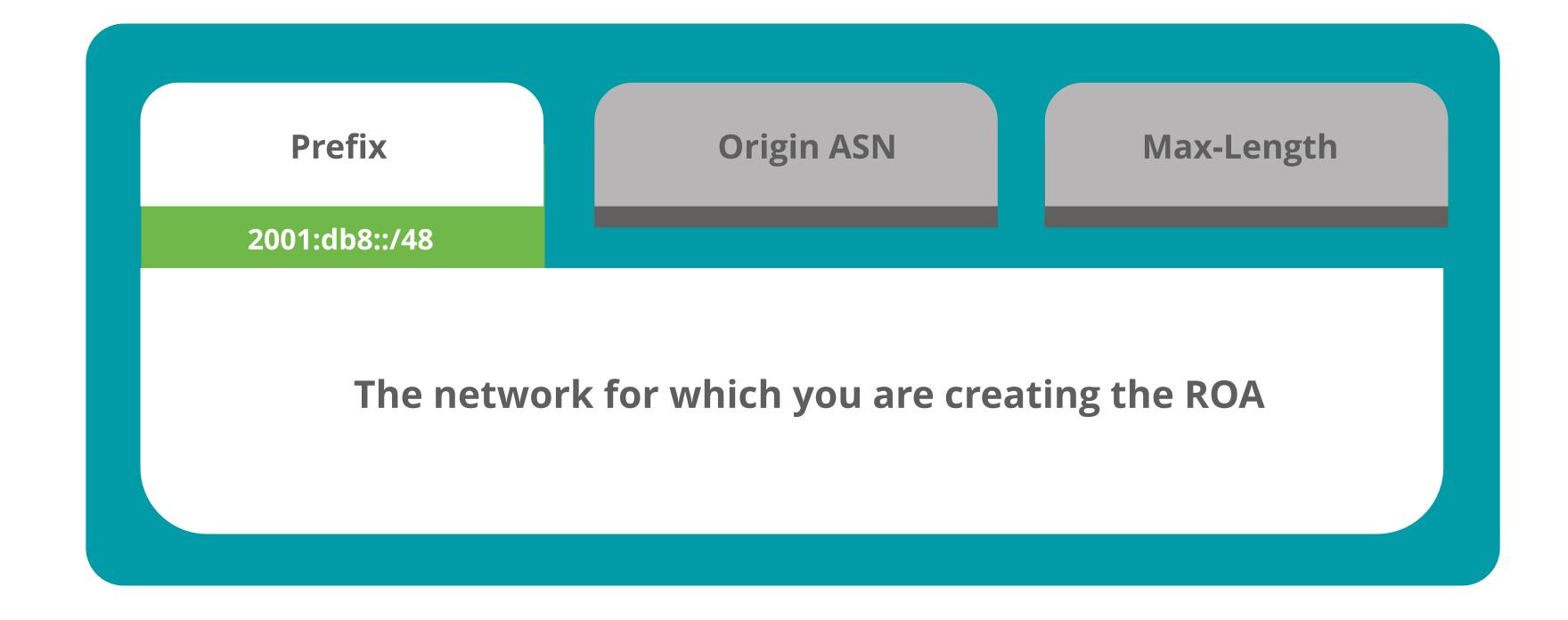


The RPKI system consists of two parts:



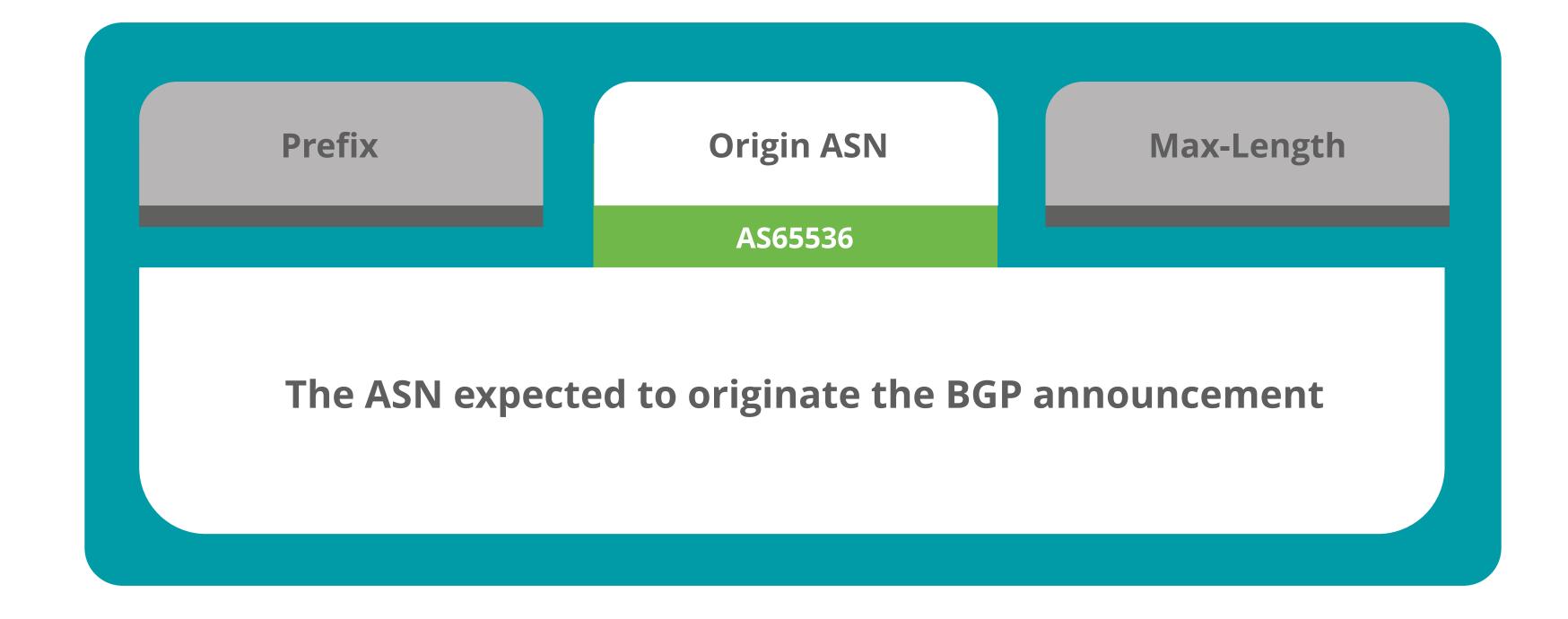
#### What is in a ROA?





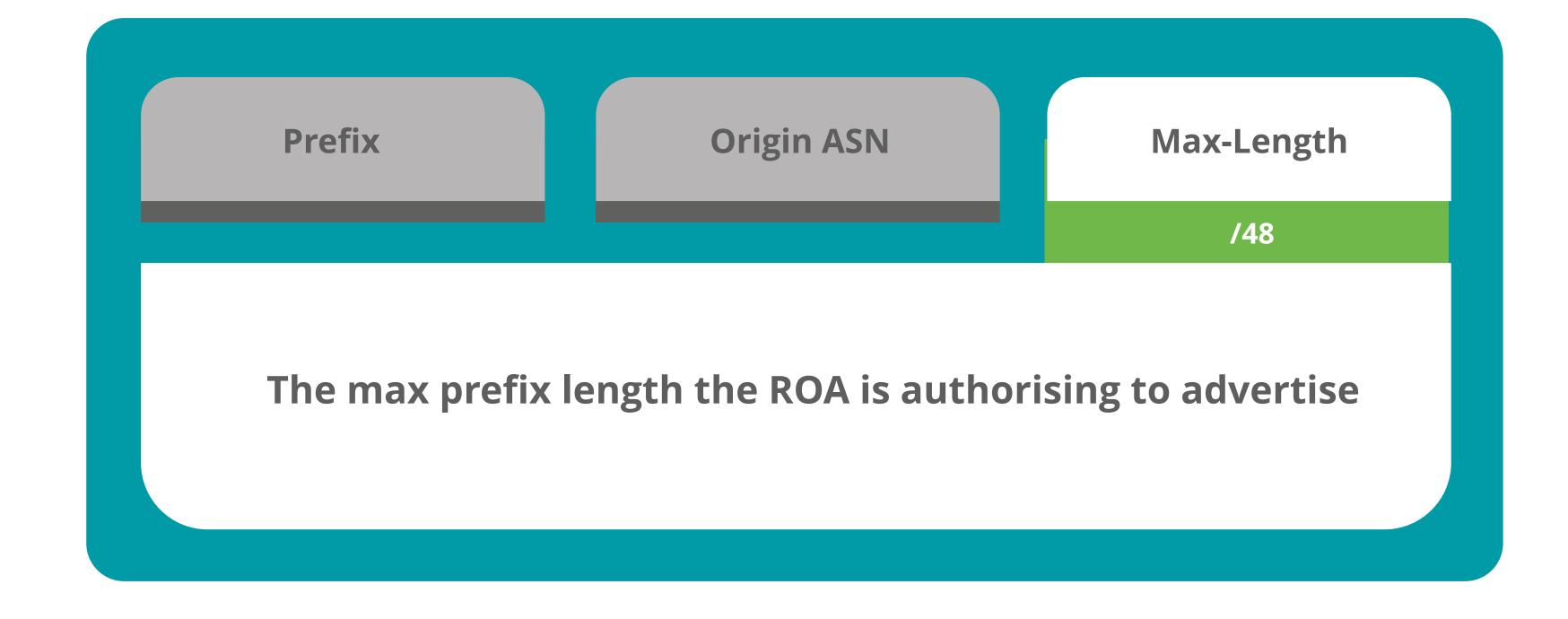
#### What is in a ROA?





#### What is in a ROA?





#### Max-Length

RIPE NCC (AS3333) has an IP address allocation

RIPE NCC creates this ROA

193.0.0.0/21



**Prefix** 

193.0.0.0/21

Max Length /22

**Origin ASN** 

**AS3333** 

# Max-Length

RIPE NCC (AS3333) has an IP address allocation

RIPE NCC creates this ROA

According to the ROA:

/21 /22

193.0.0.0/21

ROA	
Prefix	193.0.0.0/21
Max Length	/22
Origin ASN	AS3333

## Max-Length

RIPE NCC (AS3333) has an IP address allocation

RIPE NCC creates this ROA

According to the ROA:

/21 **/22 /22 /23 /23 /23 /23 /24 /24 /24 /24 /24 /24 /24 /24** 

193.0.0.0/21

**ROA** 193.0.0.0/21 Max Length /22

**AS3333** 

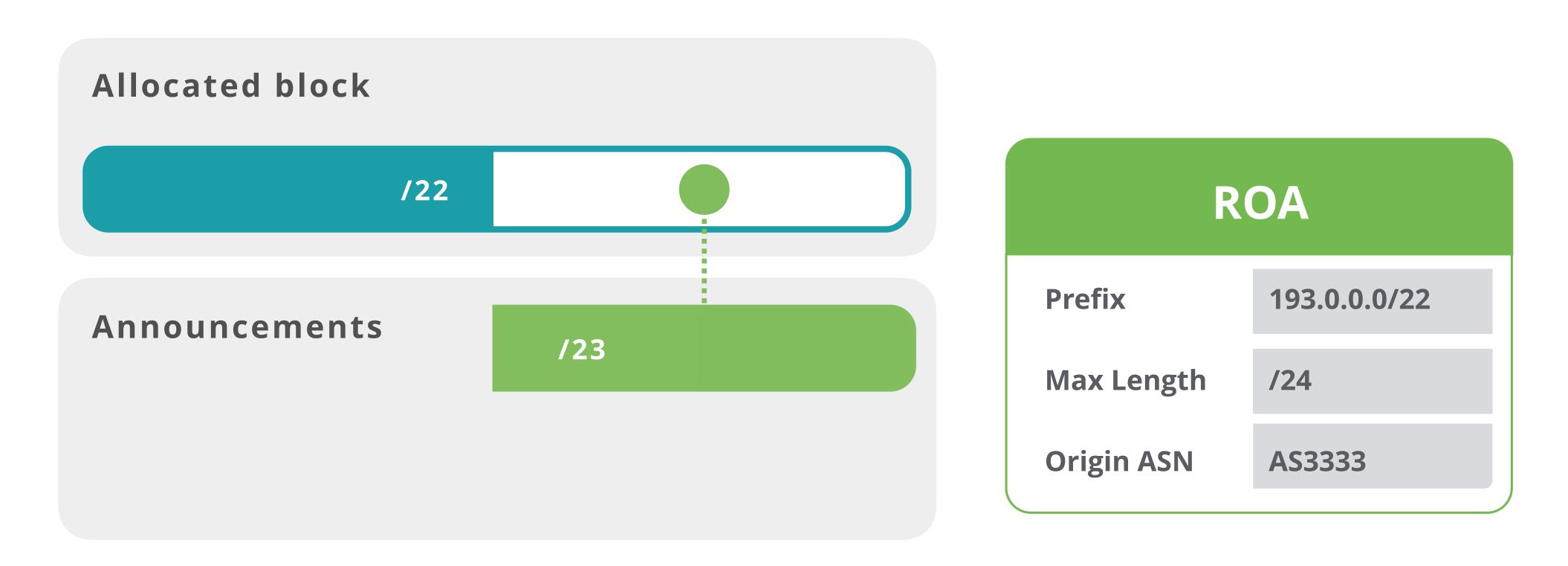
**Prefix** 

**Origin ASN** 

Any other more specific announcements are unauthorised by the ROA

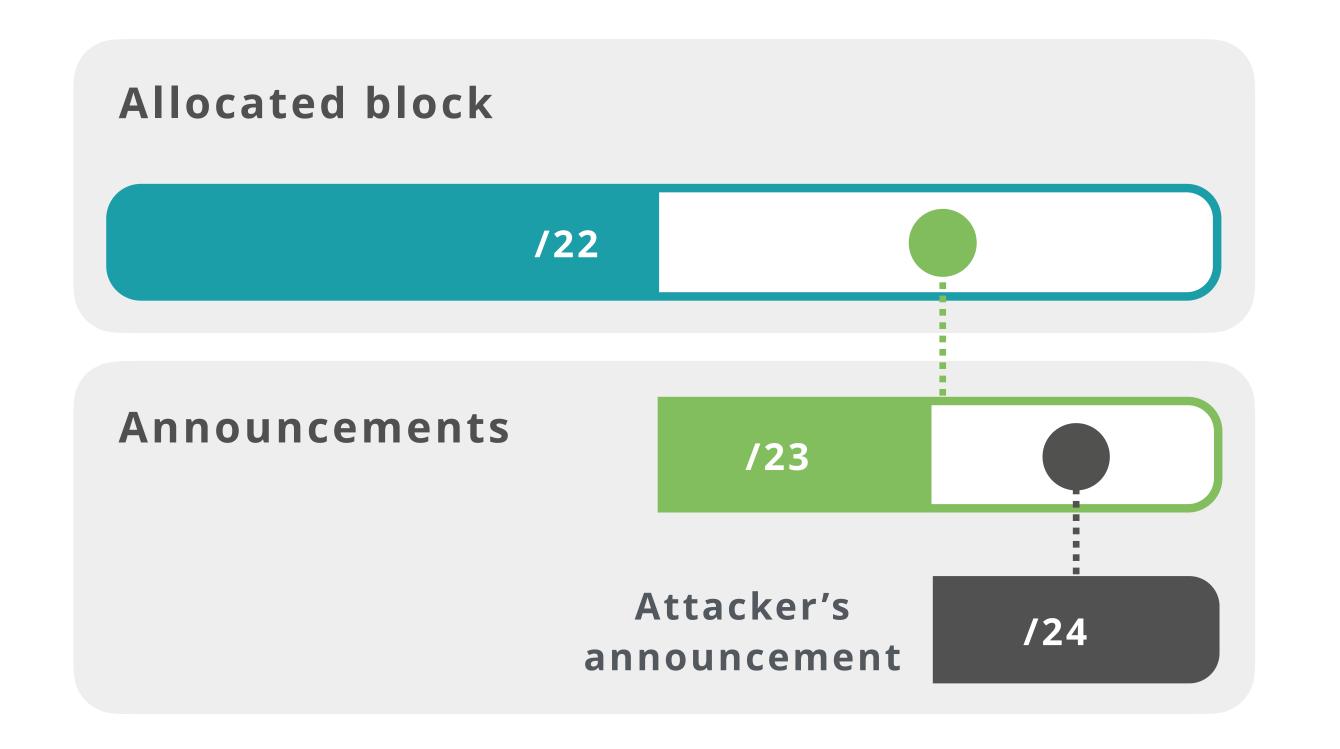


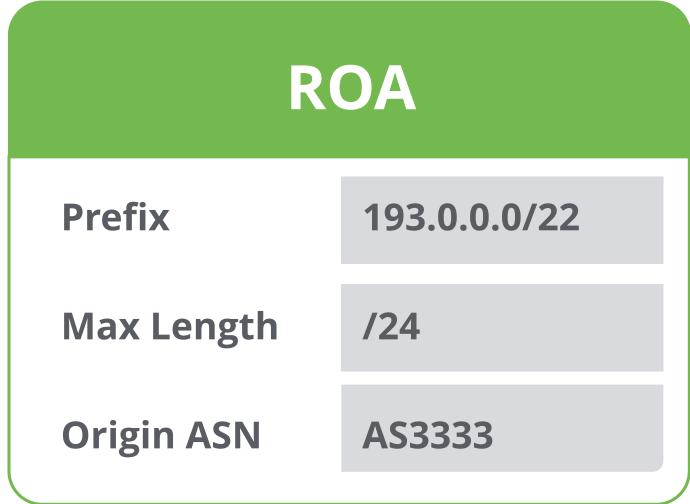
Case 1: You create a single ROA authorising the entire /22





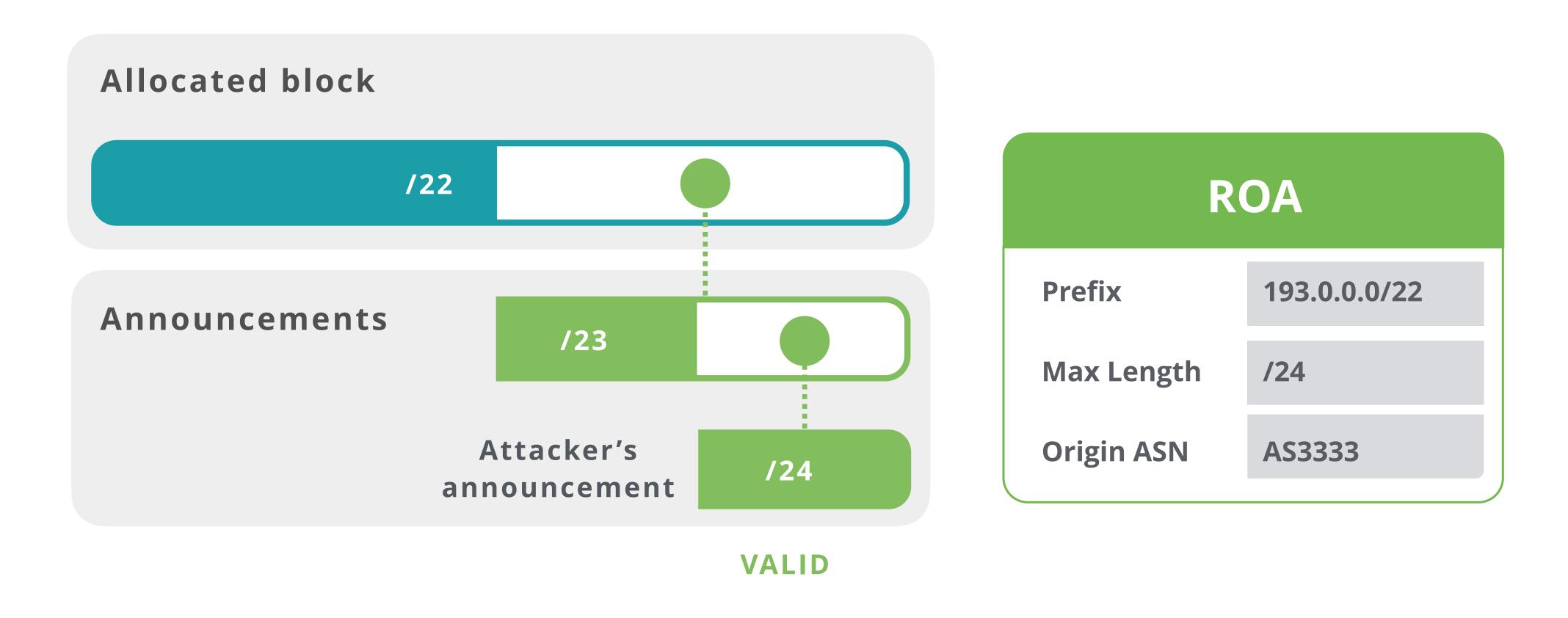
Case 1: You create a single ROA authorising the entire /22





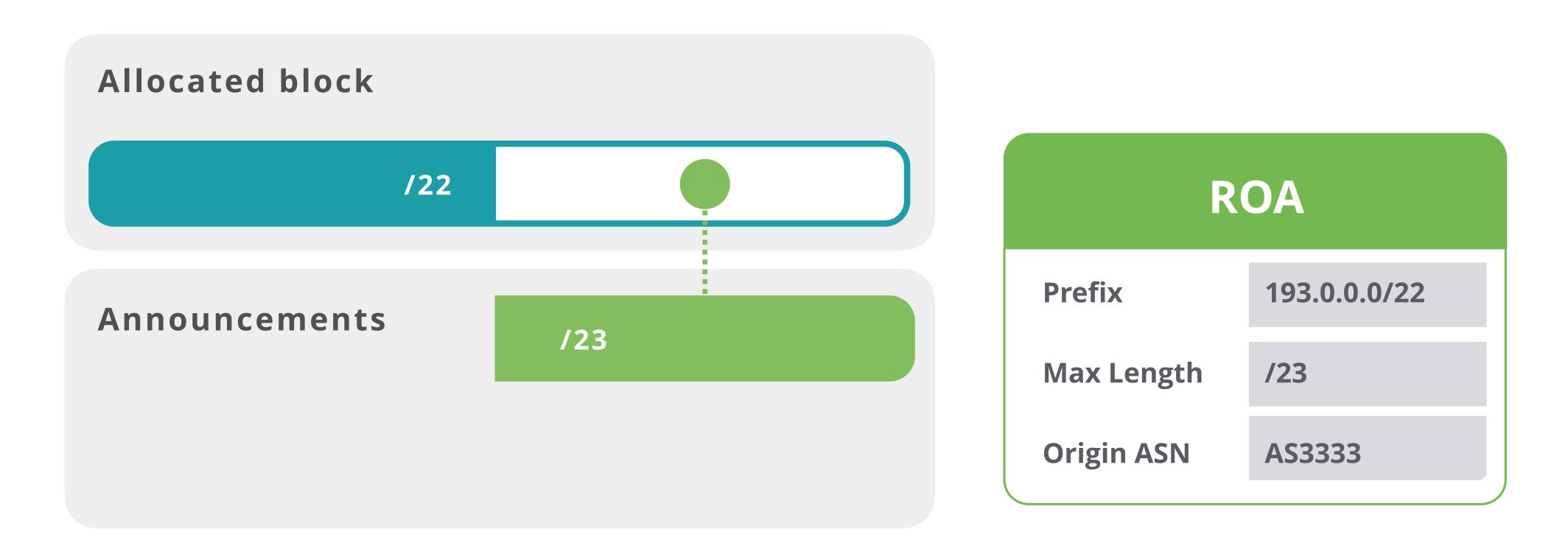


Case 1: You create a single ROA authorising the entire /22



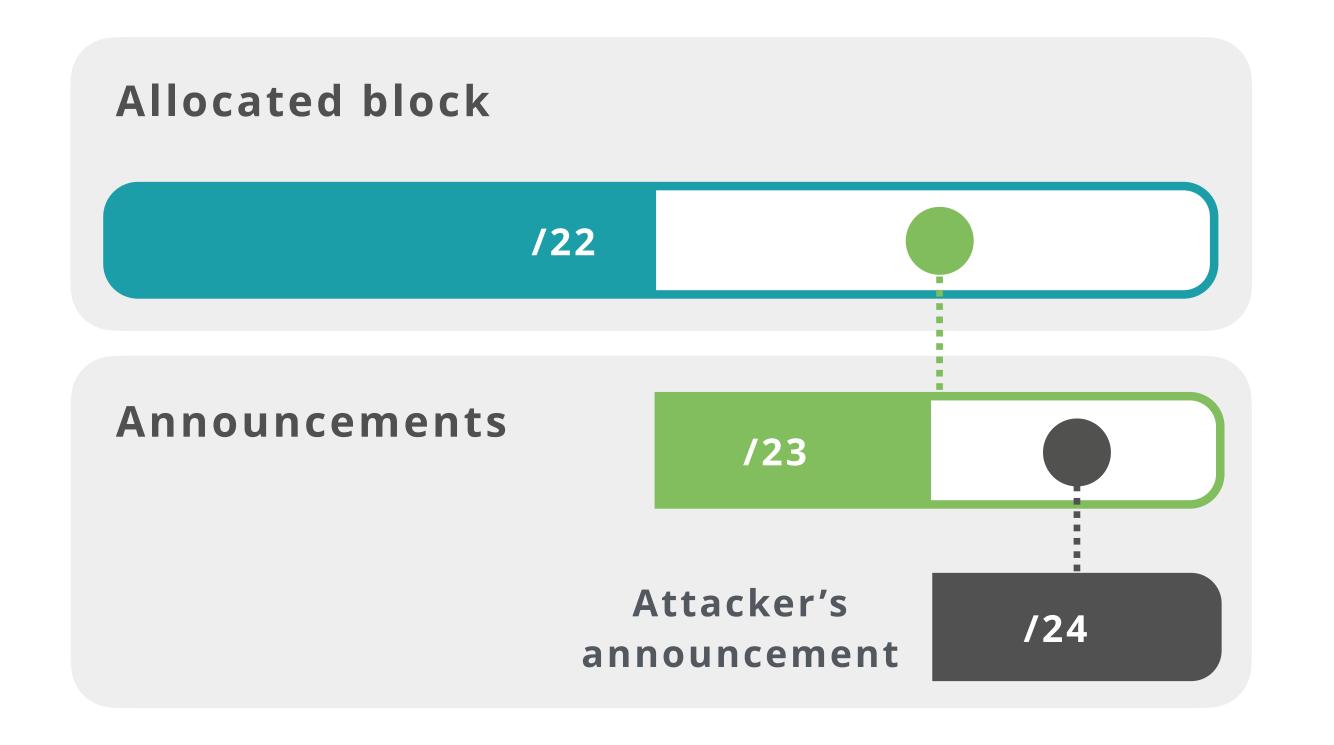


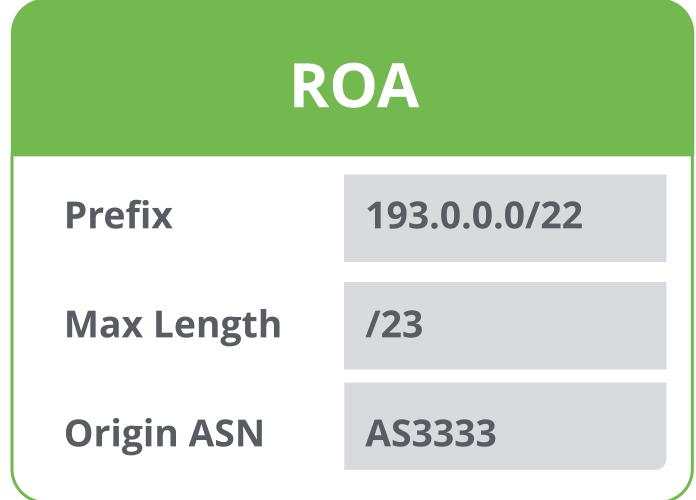
Case 2: You create ROAs only for your BGP announcement





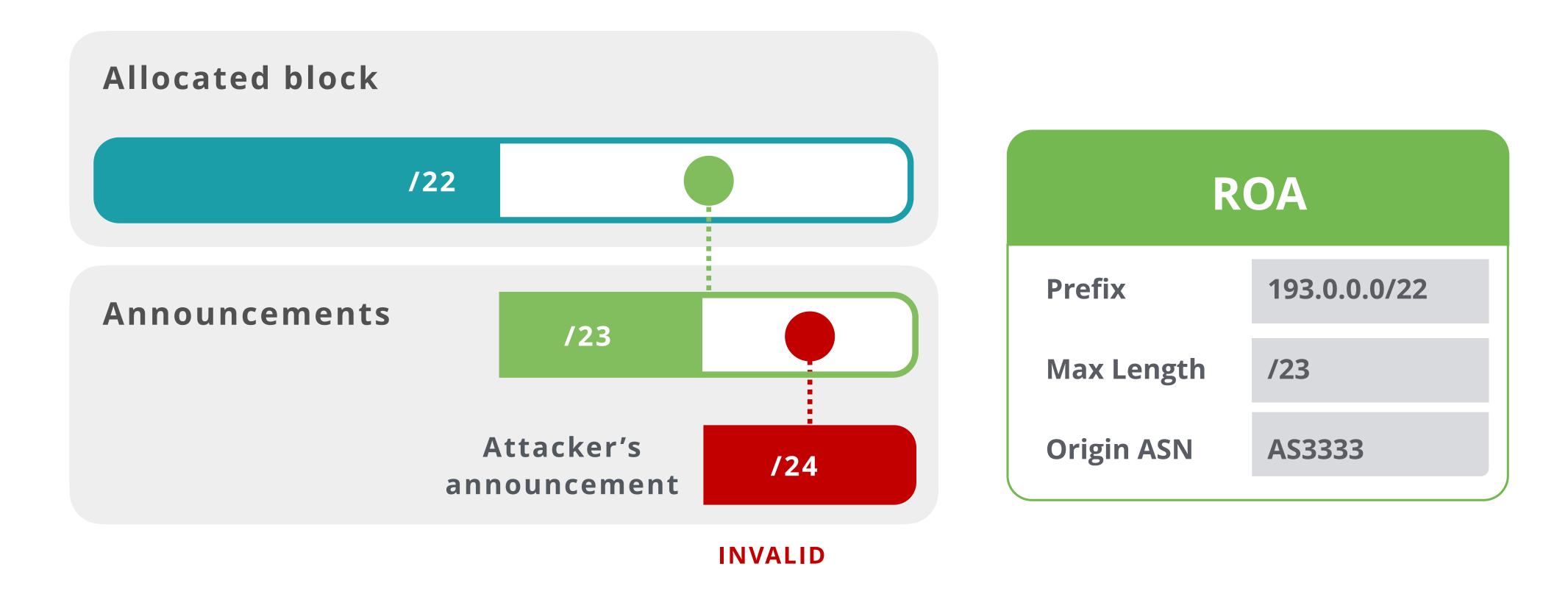
Case 2: You create ROAs only for your BGP announcement







Case 2: You create ROAs only for your BGP announcement

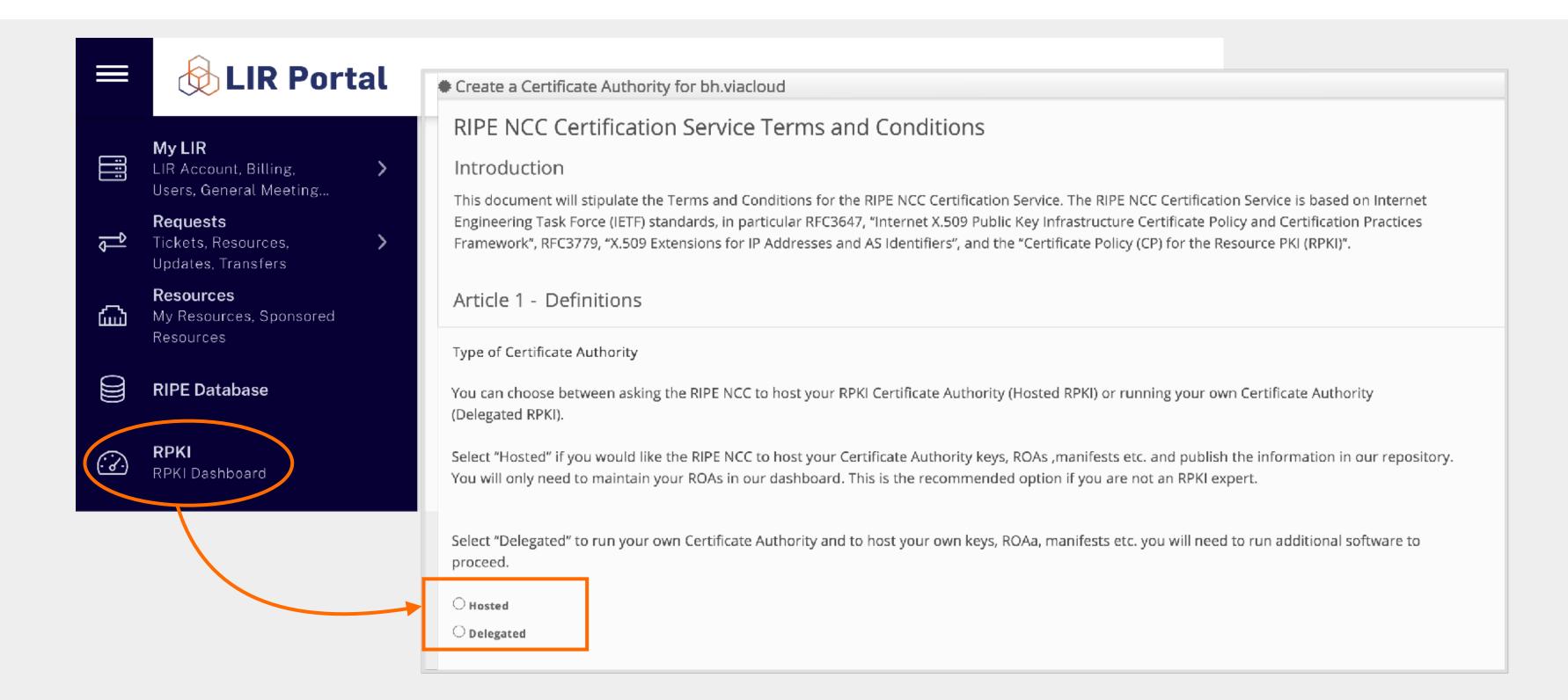


# How to create a ROA? Easy!



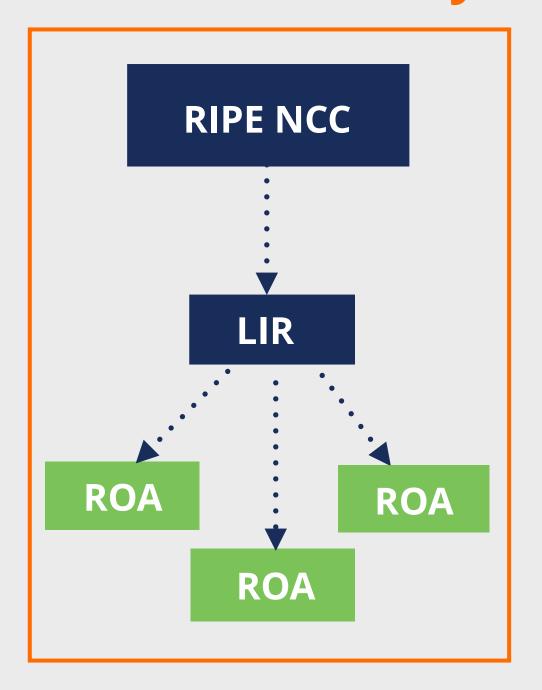
- Login to LIR Portal (<u>my.ripe.net</u>)
- Go to the RPKI Dashboard
- Choose which RPKI model to use





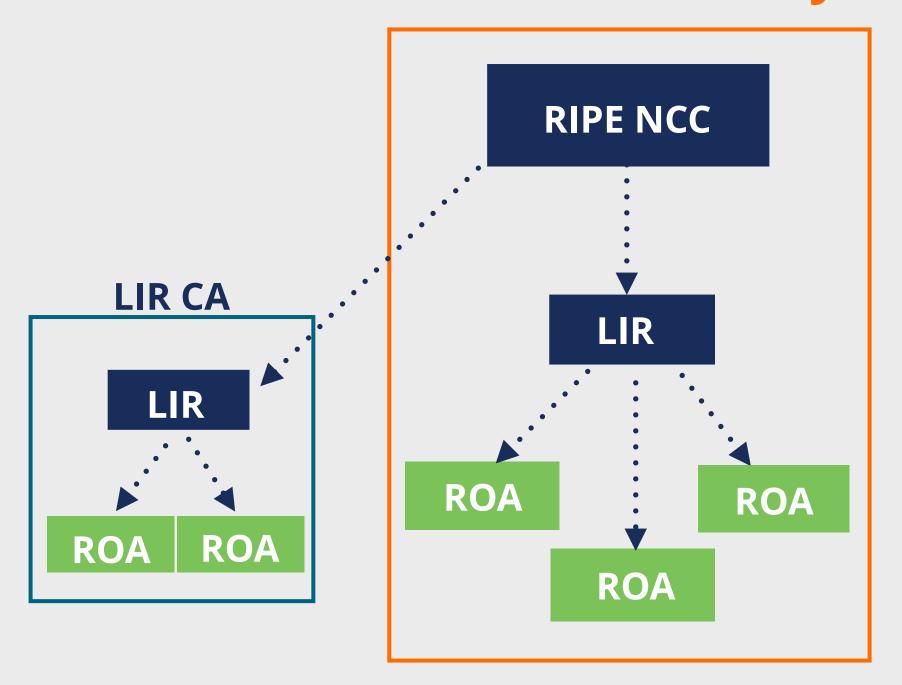
## **Hosted RPKI**

- ROAs are created and published using the RIR's member portal
- RIR hosts a CA for LIRs and signs all ROAs
- Automated signing and key rollovers
- Allows LIRs to focus on creating and publishing ROAs



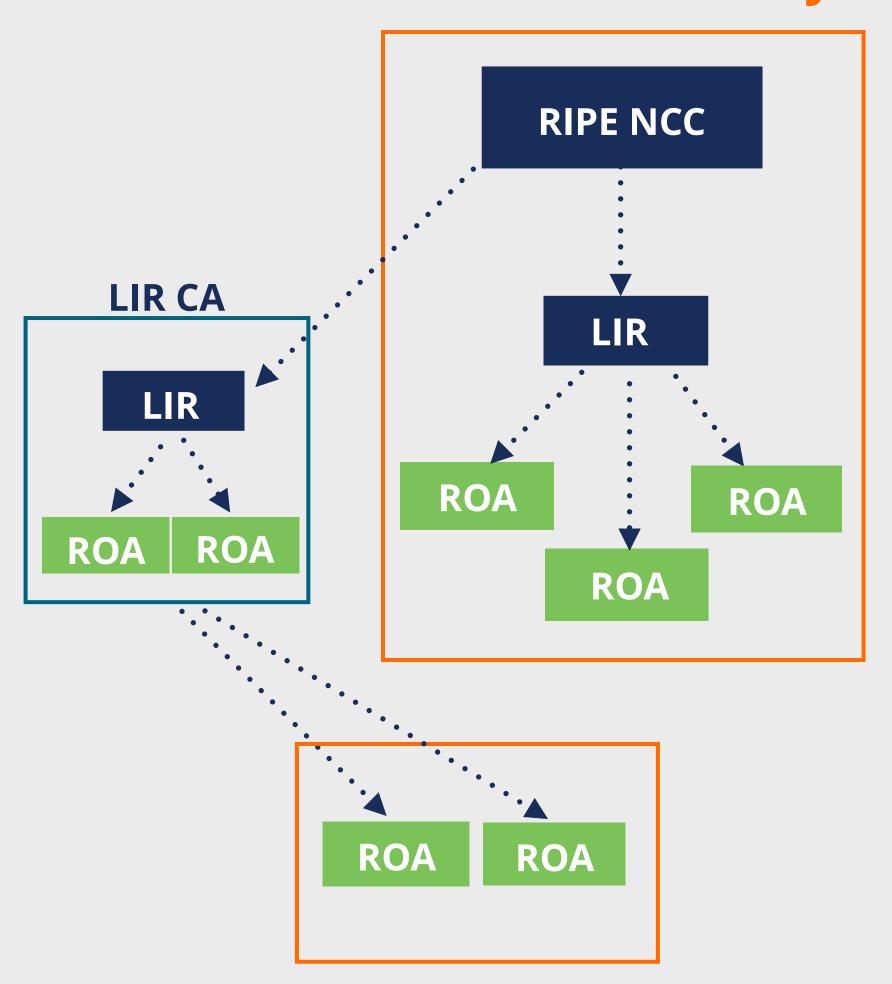
## Delegated RPKI

- Each LIR manages its part of the RPKI system:
  - Runs its own CA as a child of the RIR
  - Manages keys/key rollovers
  - Creates, signs and publishes ROAs
- Certificate Authority (CA) Software
  - Krill (NLnet Labs)
  - rpkid (Dragon Research Labs)



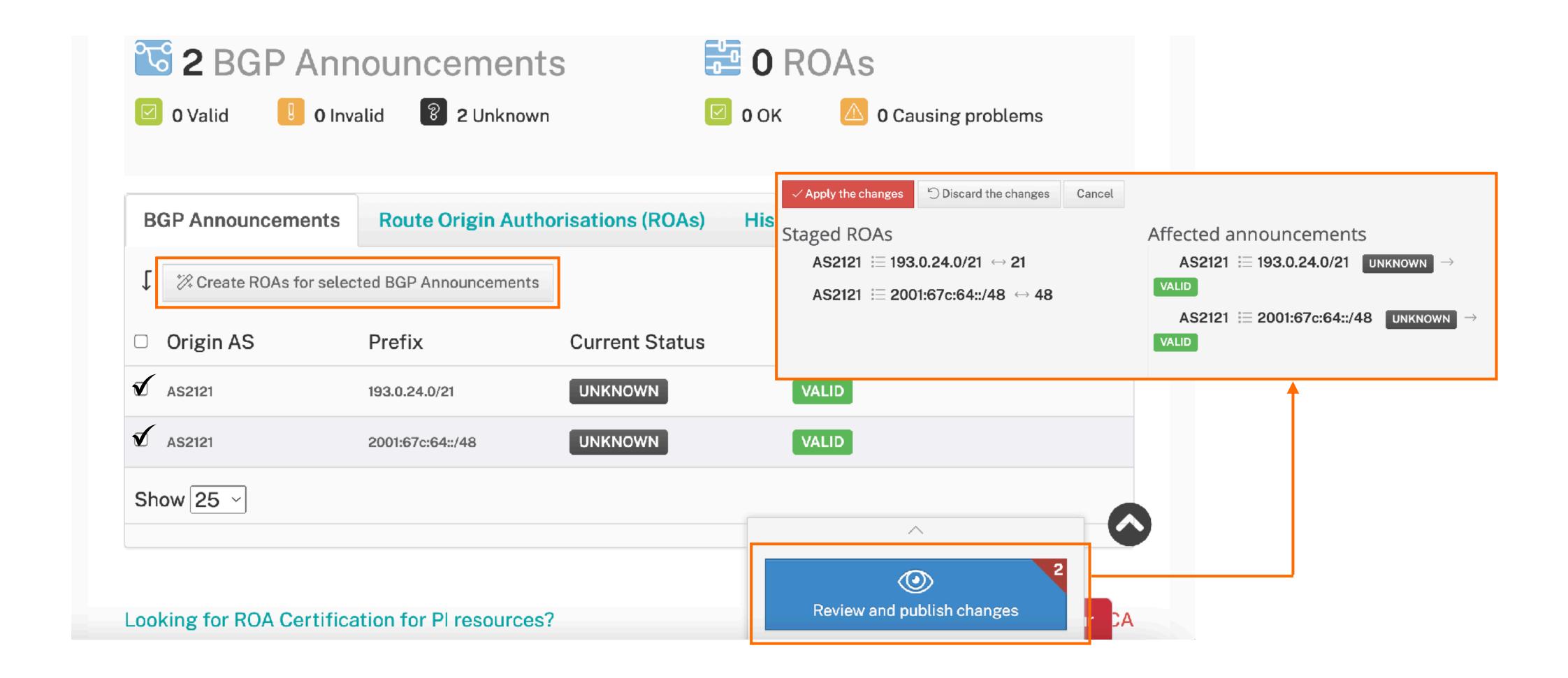
## **Hybrid RPKI**

- In-between hosted and delegated RPKI
- The LIR:
  - Runs its own CA as a child of the RIR
  - Manages keys/key rollovers and ROAs
  - Maintains key-pairs and objects and send them to RIR
  - RIR publishes ROAs in its repository
- Supported by APNIC, ARIN, RIPE NCC and NIRs
- AKA "Publication in parent" or "Publication as a service"



## RIPE NCC Hosted Solution





# Certifying Pl Resources



### Requested and managed by PI End User or by Sponsoring LIR

1. Complete the wizard successfully

Start the wizard to set up Resource Certification for PI and Legacy End User resources 🖸

- 2. Login to https://my.ripe.net and request a certificate
  - Sign in with your RIPE NCC Access account
- 3. Manage your ROAs



# Questions (2)



# Lab 5.1 RPKI - Creating ROAs



Description: Create (see demo) ROAs in the test RPKI Dashboard

#### Goals:

- Identify elements of the RPKI infrastructure
- Register routing information in the RPKI dashboard by creating a ROA
- Time: 15 minutes

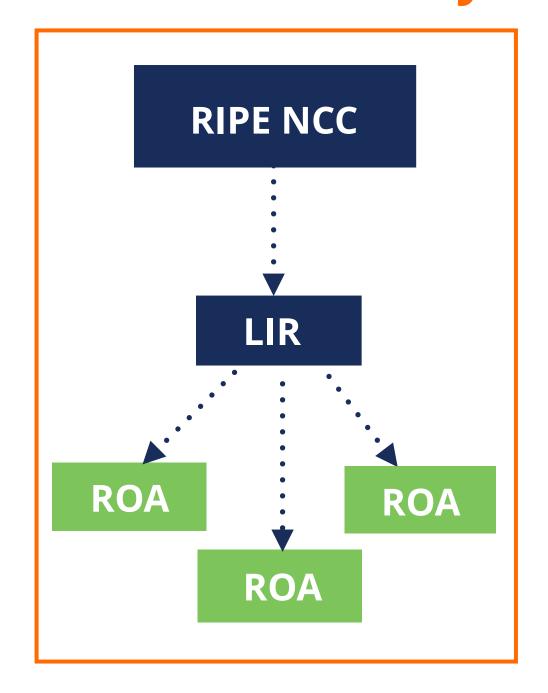
#### Tasks:

- 5.1 Check your BGP announcements
- 5.1 Create (or see a demo) ROAs in the test RPKI Dashboard
- 5.1 Create (or see a demo) a more specific ROA for one of your prefixes





- What have you learned?
  - How to create ROAs using the RIPE NCC Hosted RPKI service web interface



## **Elements of RPKI**



The RPKI system consists of two parts:



## **RPKI Validation**

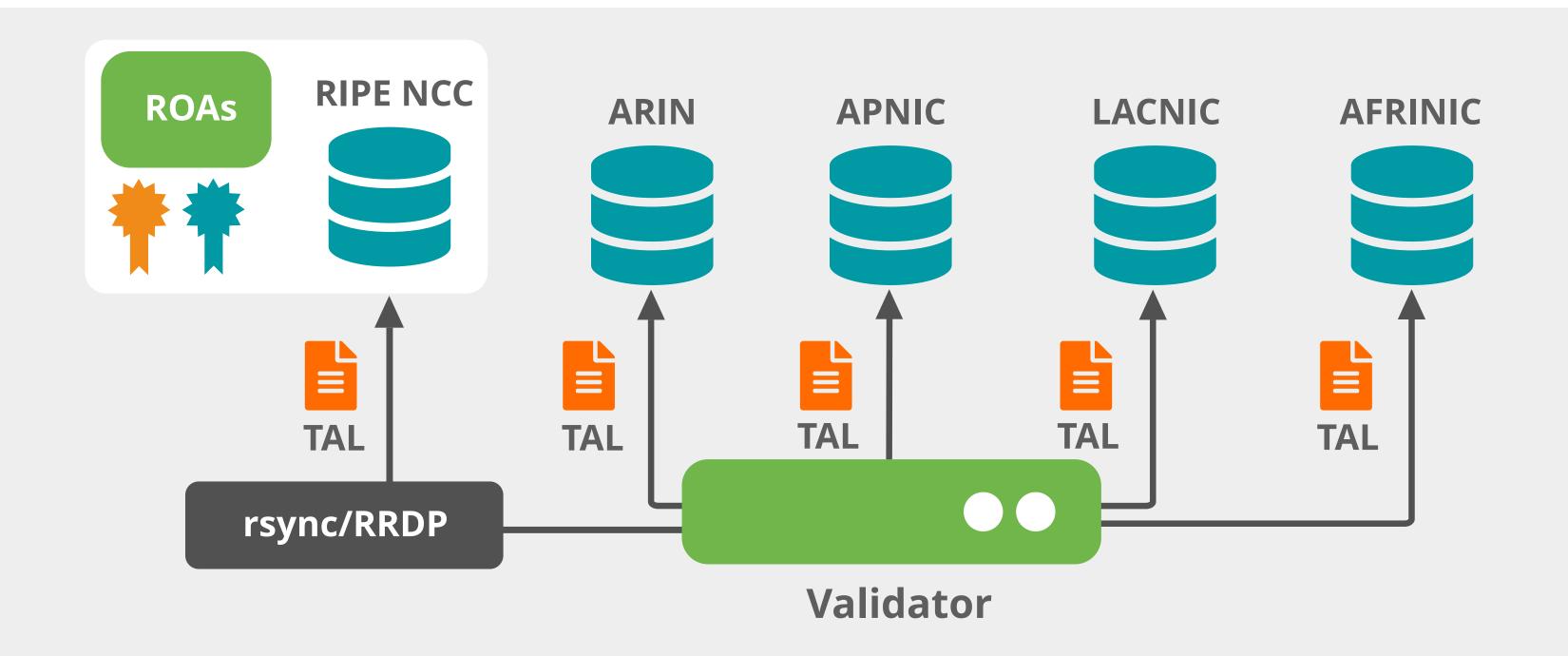


- Verifying the information provided by others
- First, validate the RPKI data
  - Install a validator software locally in your network
  - Verify holdership through a public key and certificate infrastructure
- Second, validate the origin of BGP announcements
  - Known as BGP Origin Validation (BGP OV) or Route Origin Validation (ROV)
  - This is done in a BGP router in your network

## **RPKI Validator**



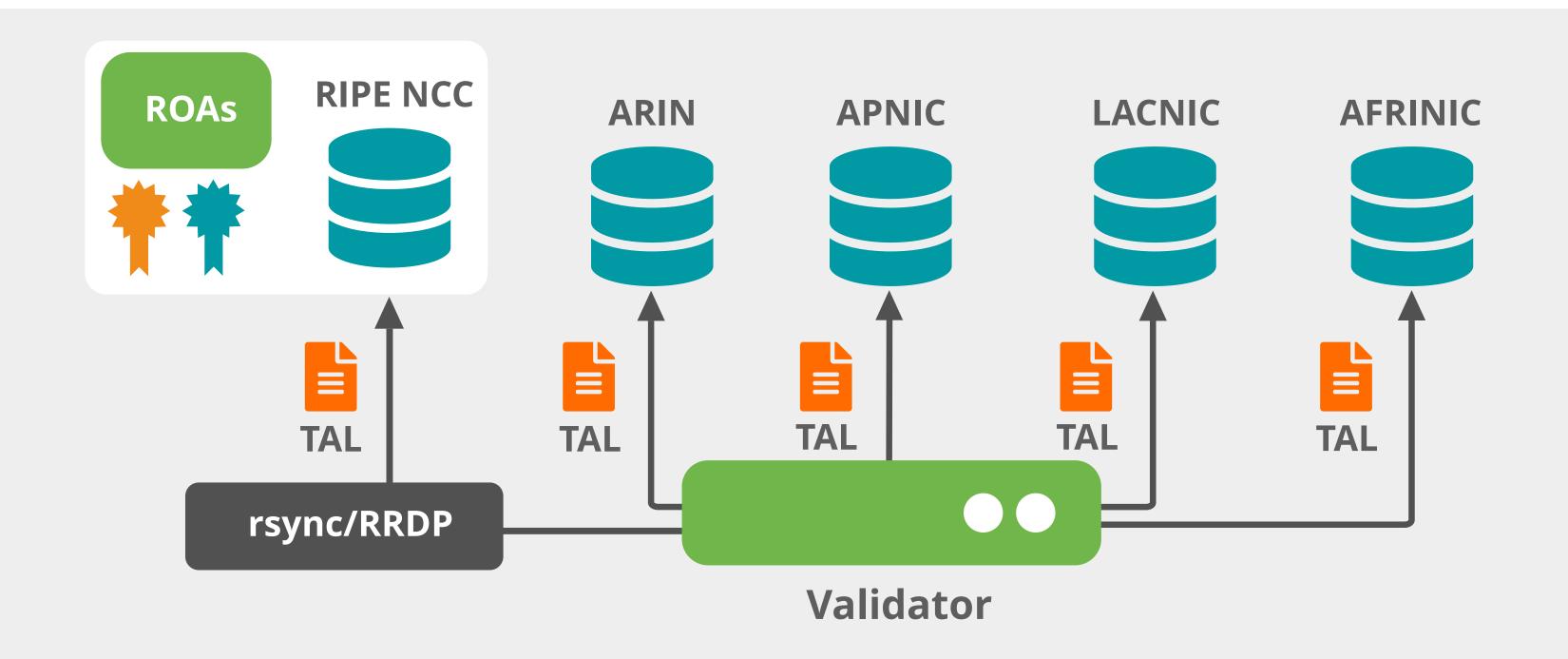
- Also known as Relying Party (RP) software
- Connects to RPKI repositories via **rsync** or **RRDP** protocol
- Uses information in TALs to connect to the repositories



## **RPKI Validator**

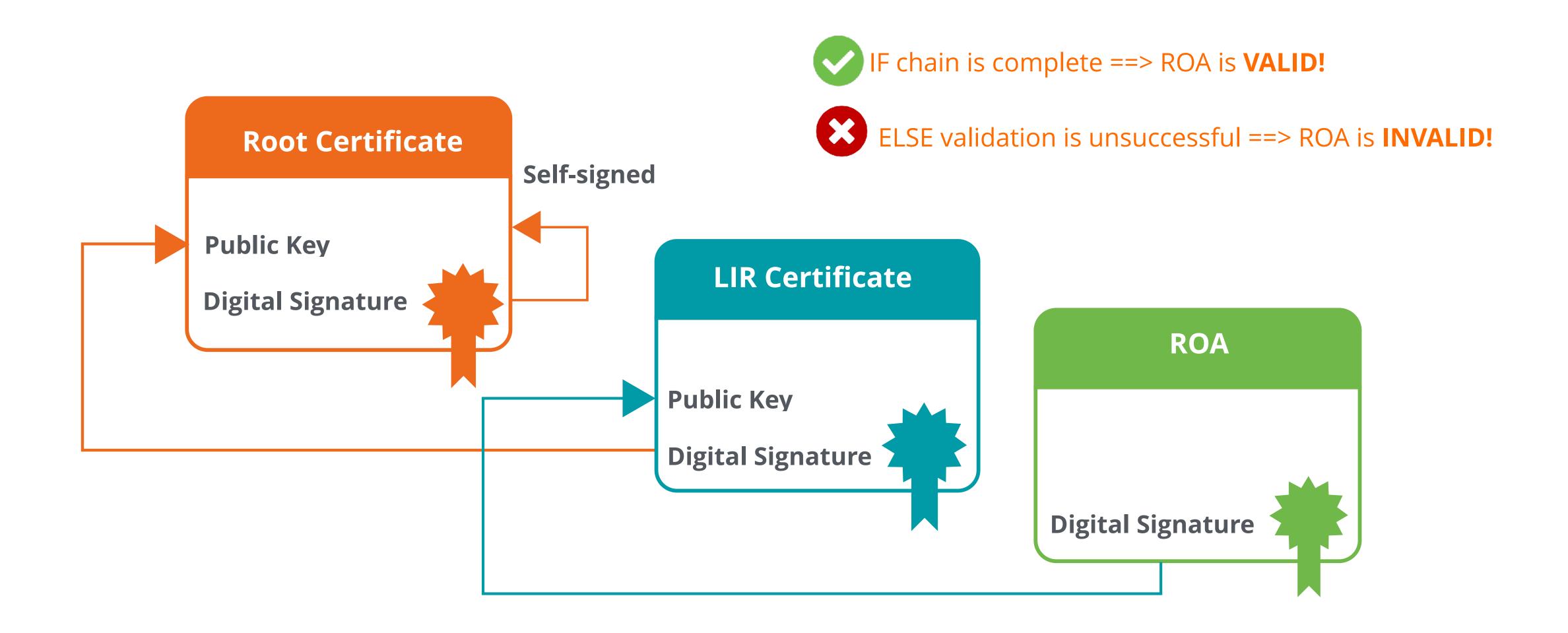


- Downloads ROAs from RPKI repositories
  - From RIRs and external repos
- Validates the chain of trust for all ROAs and associated CAs
  - Creates a local "validated cache" with all the valid ROAs



## **ROA Validation Process**





## **RPKI Validator Options**



- Routinator
  - Built by NLNet Labs
- FORT
  - Open source RPKI validator

- rpki-client
  - Integrated in OpenBsd

**Links for RPKI Validators:** 

https://github.com/NLnetLabs/routinator.git

https://github.com/NICMx/FORT-validator/

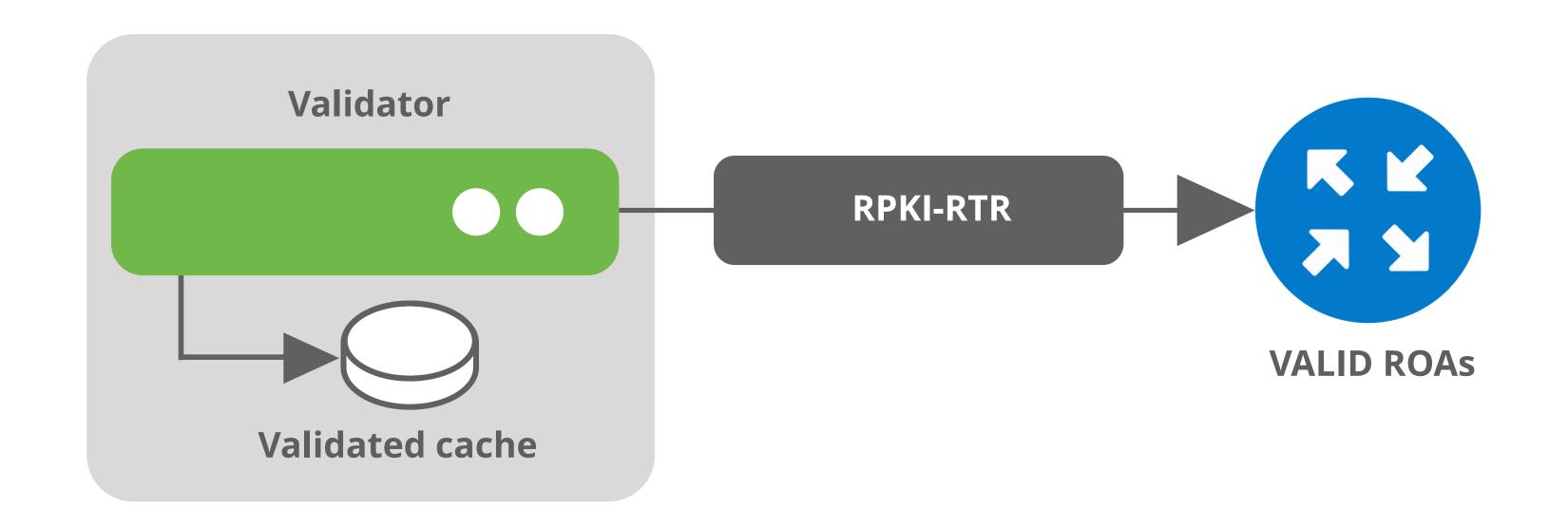
https://www.rpki-client.org/

**More Information:** 

https://rpki.readthedocs.io

## Valid ROAs are sent to the router





Router uses this information to make better routing decisions!

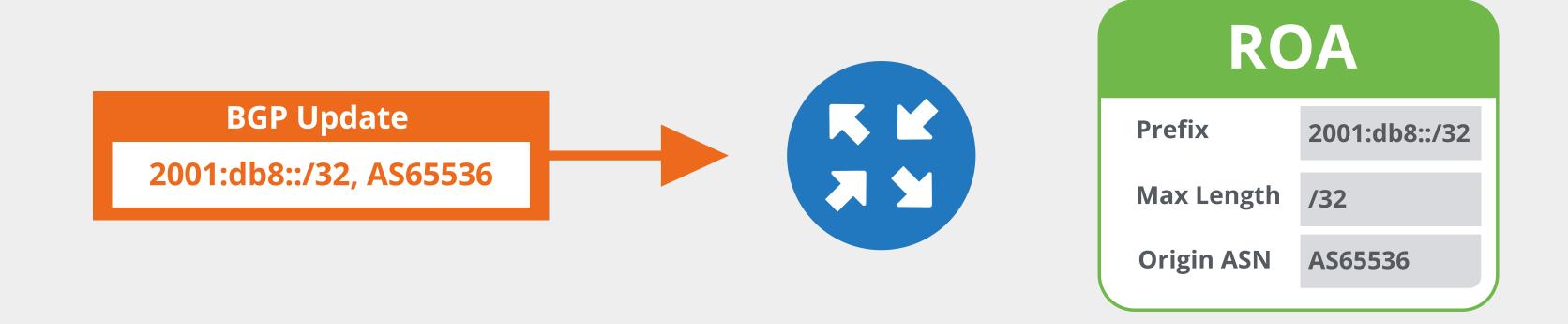




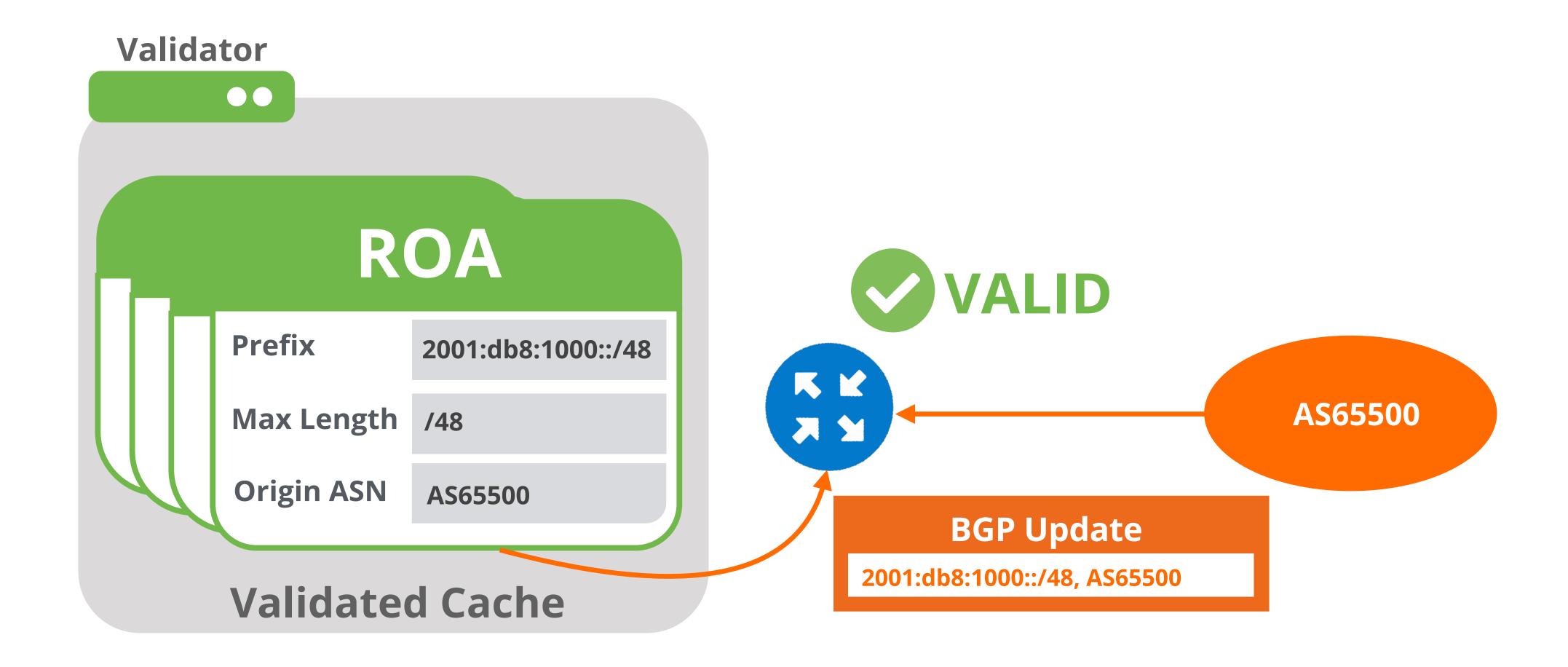




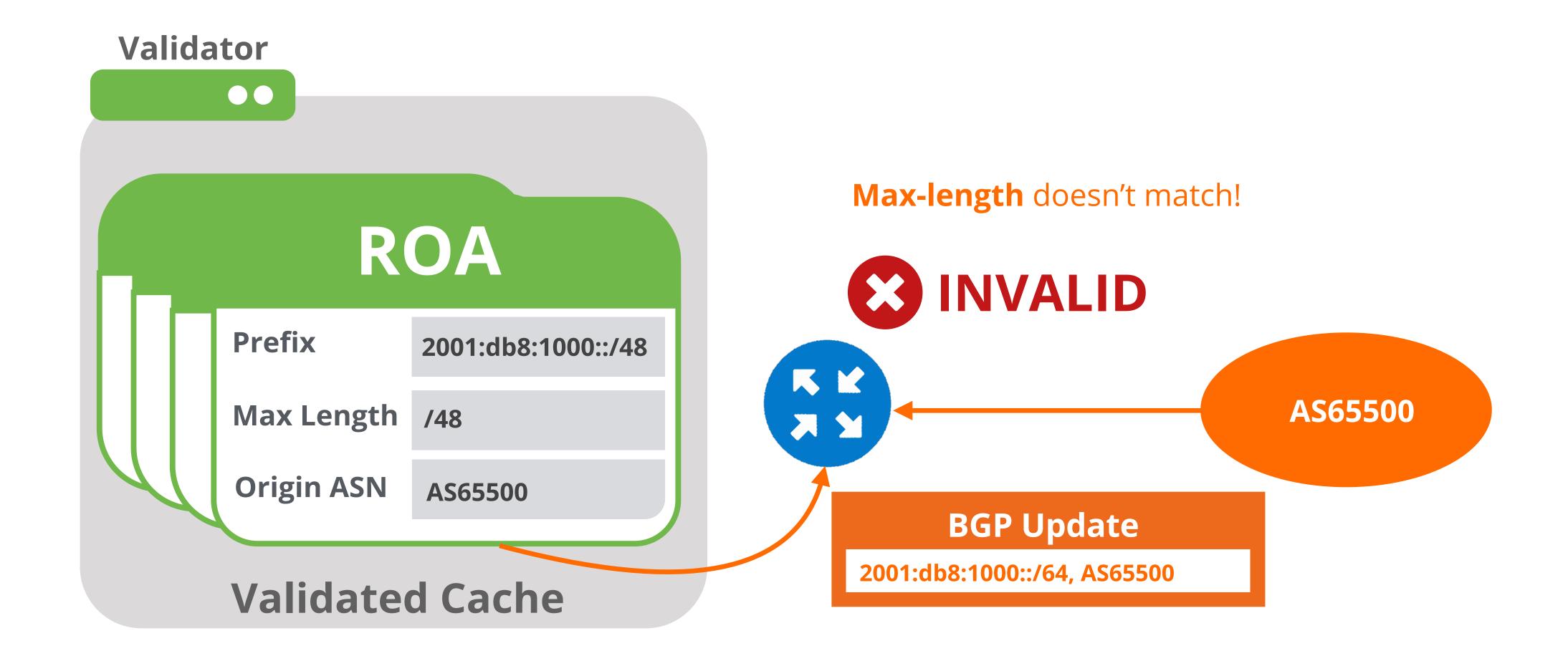
- RPKI based route filtering
- BGP announcements are compared against the valid ROAs
  - Origin ASN and max-length must match!
- Router decides the validation states of routes:
  - Valid, Invalid or Not-Found



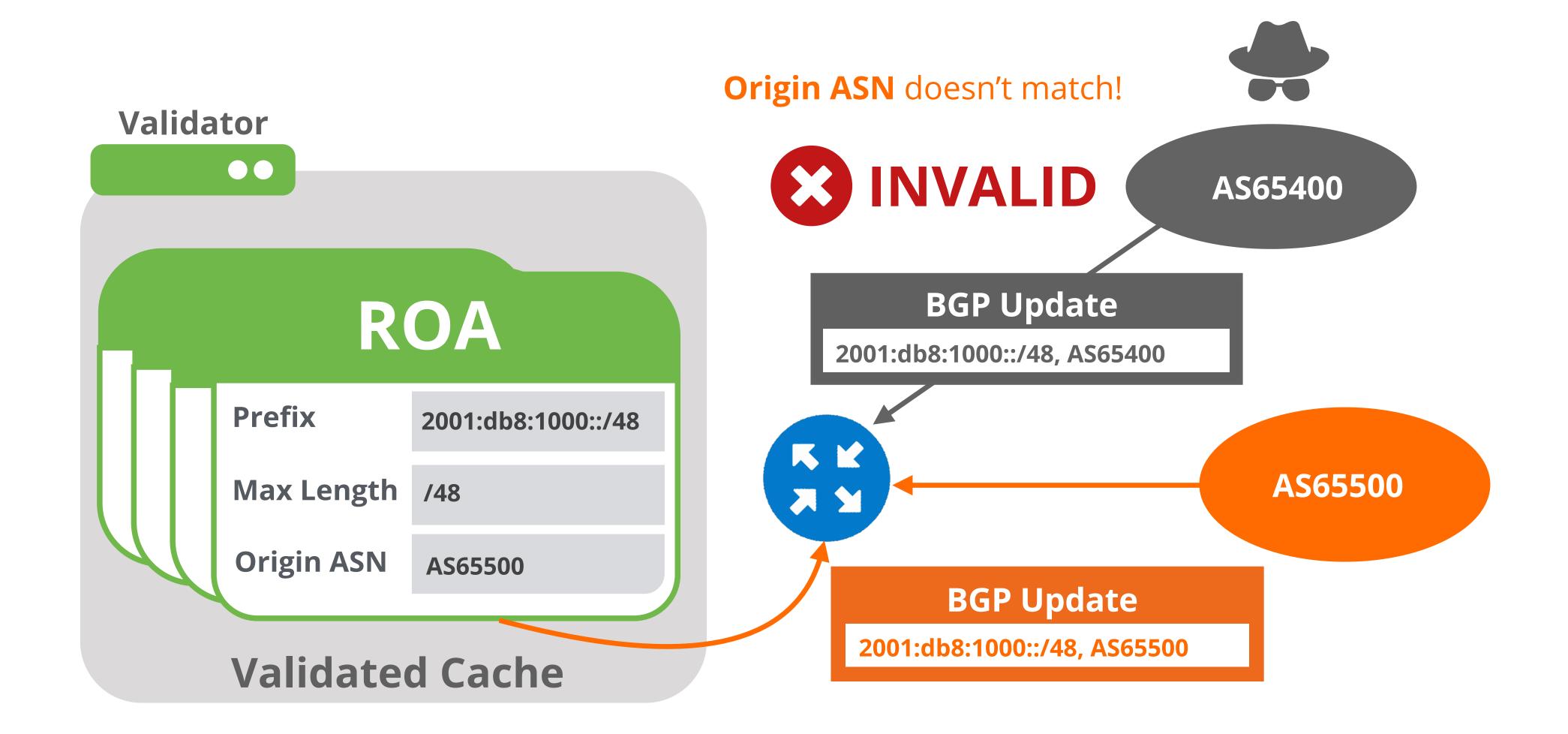




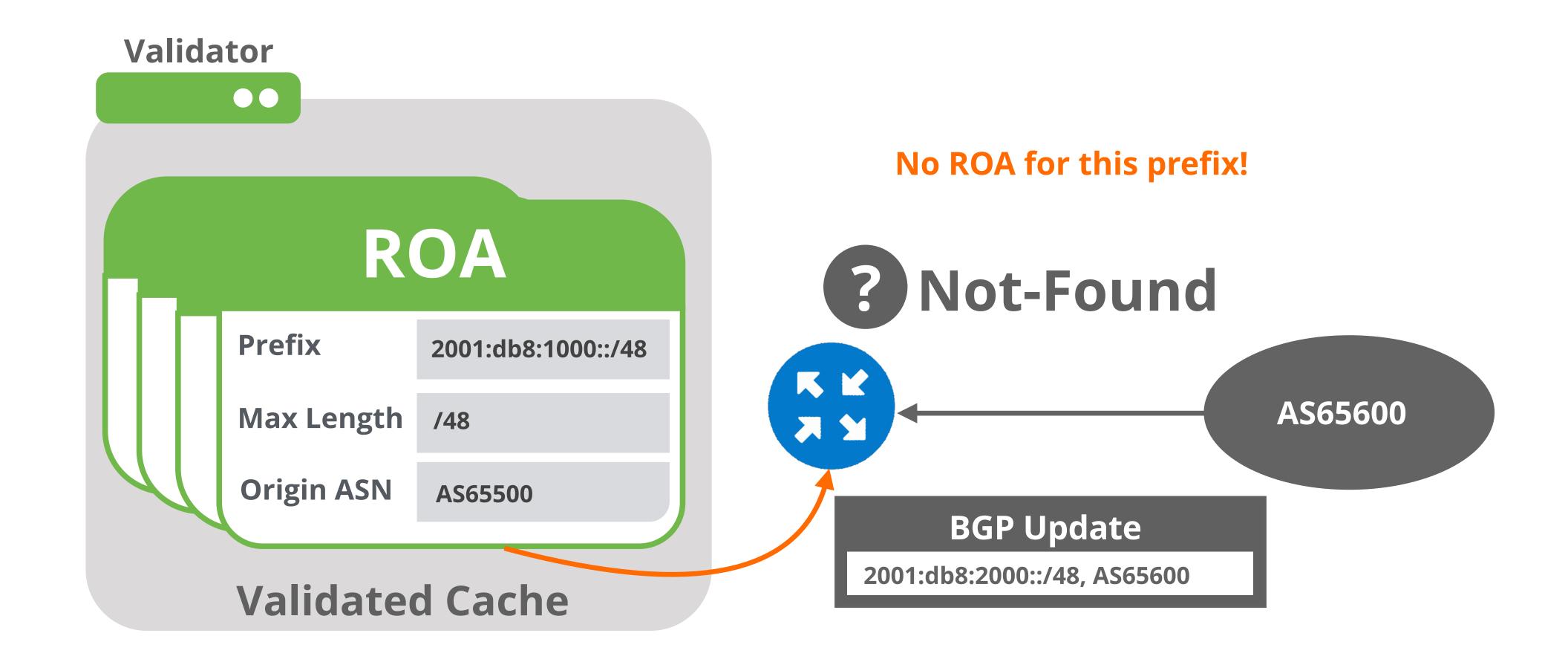






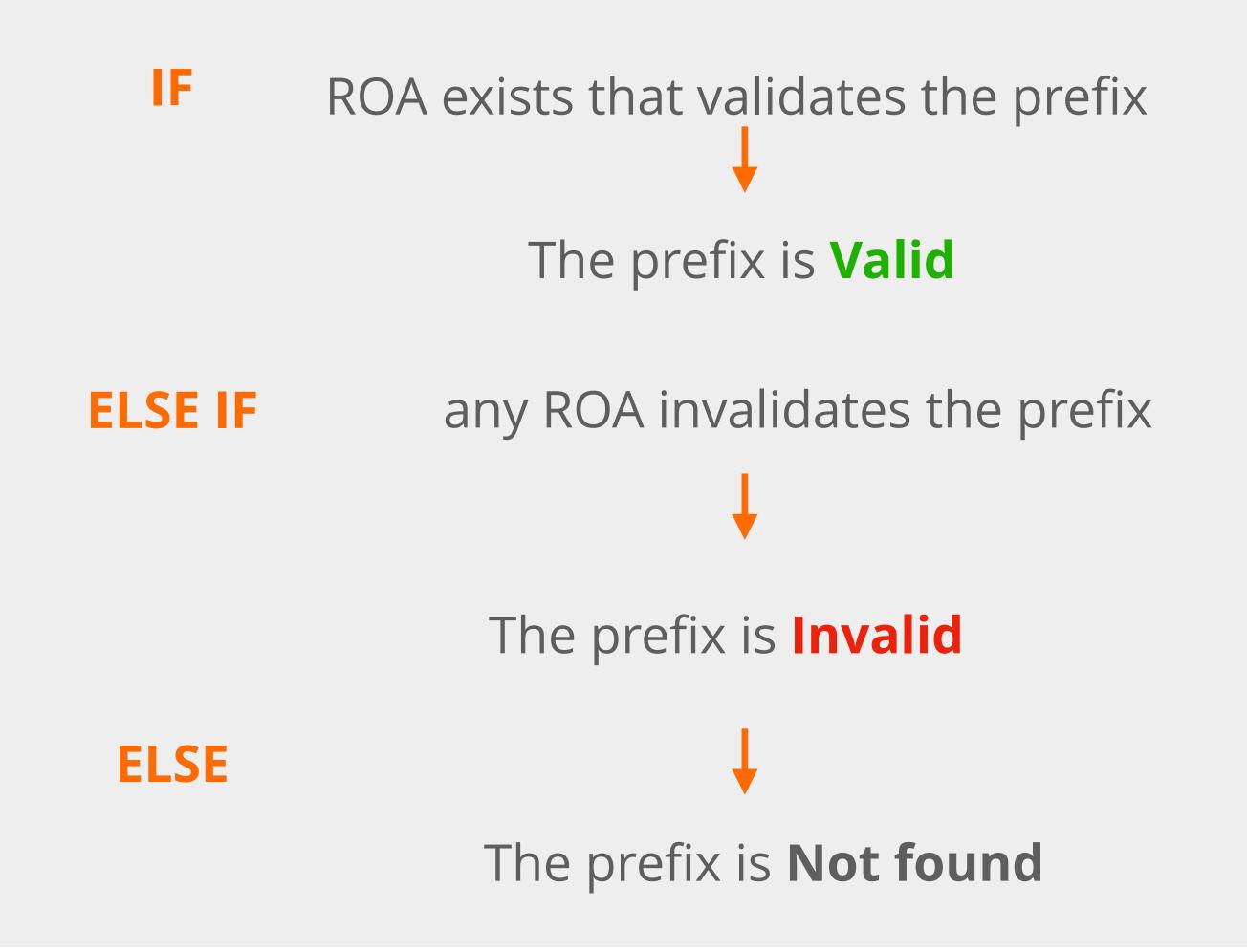






## The General Rule

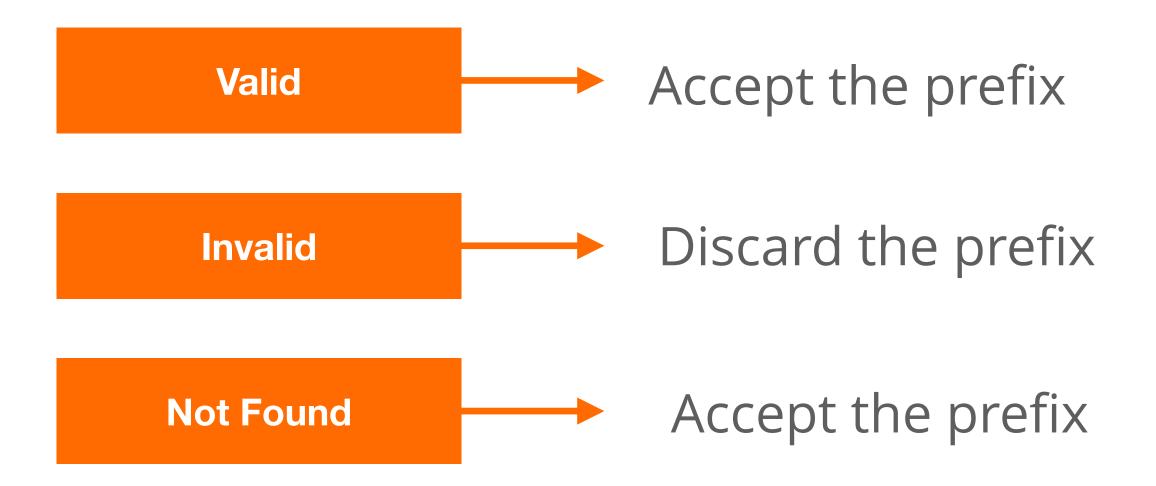




# After Validating



• You have to make a decision: Accept or Discard



Do not consider dropping prefixes with "Not-Found" RPKI validation state!

# Major Networks and RPKI Invalids

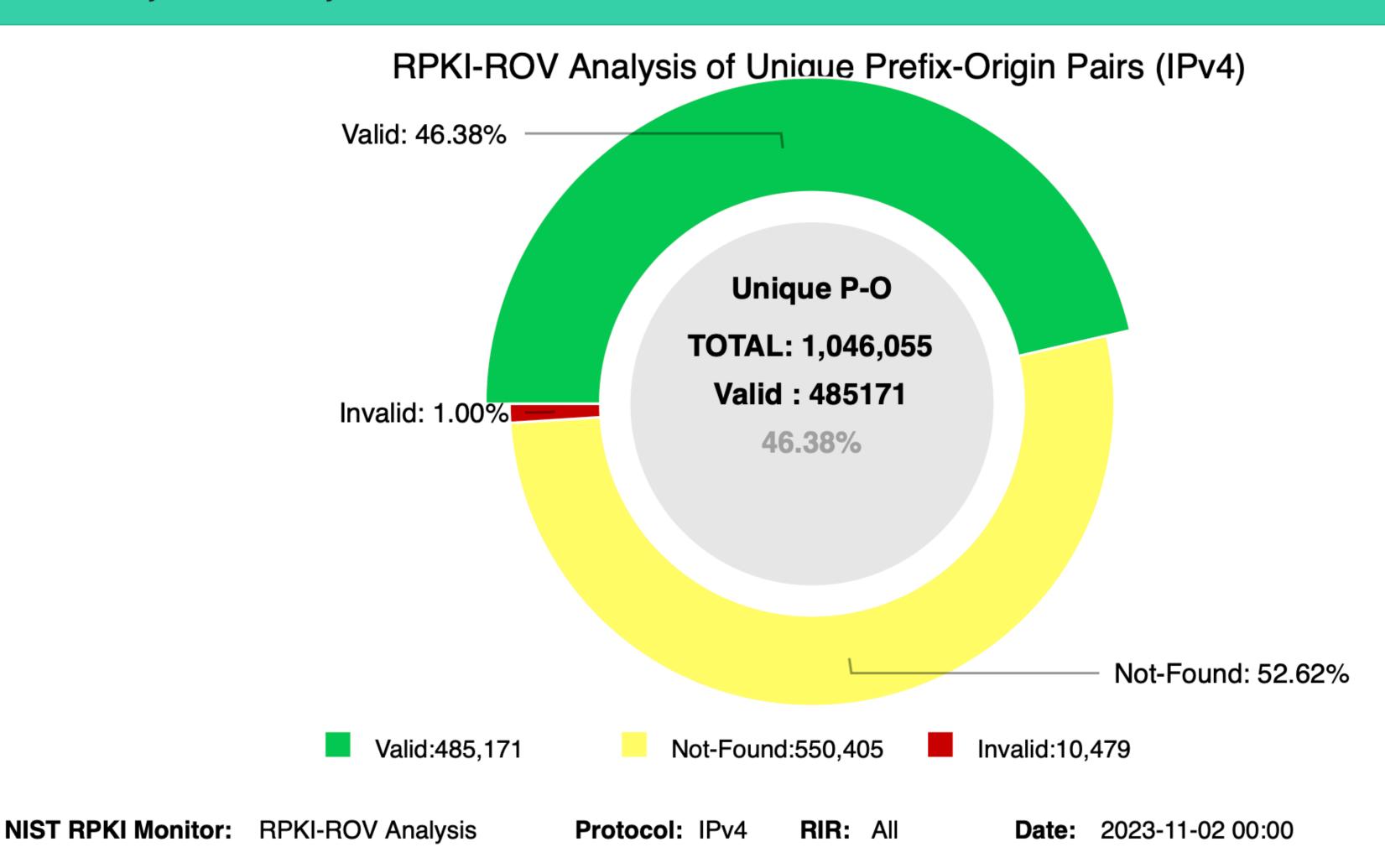


- Major networks are dropping invalids
  - Telia, AT&T, Cloudflare, Netflix, Swisscom, Cogent etc
- They follow a phased approach: First peers, then customers
  - Tag invalids on all peers, then on all customers
  - Drop invalids for all peers, then for all customers

# RPKI-ROV Analysis Globally (IPv4)



**RPKI-ROV Analysis: Global Analysis** 









# Lab 5.2 RPKI - BGP Origin Validation



 Description: Implement BGP Origin Validation and discard BGP announcements with "Invalid" RPKI status

#### Goals:

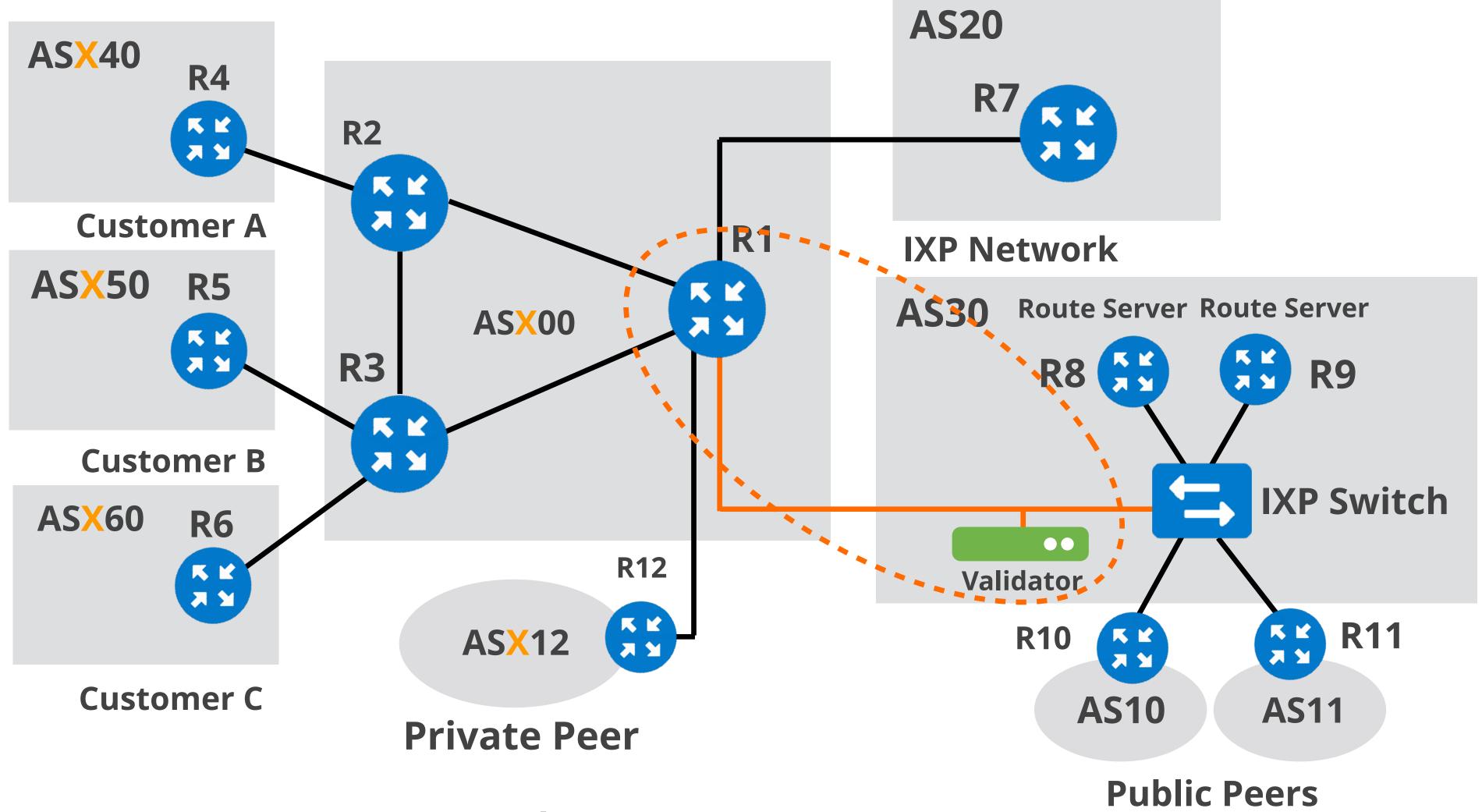
- Validate BGP announcements by using RPKI information (BGP OV)
- Use RPKI data to discard BGP Invalids
- Time: 20 minutes

#### Tasks:

- 5.2 Check valid ROAs on Routinator's GUI
- 5.2 Connect the Validator and the BGP router
- 5.2 Create a BGP Hijack

#### **Transit Provider**



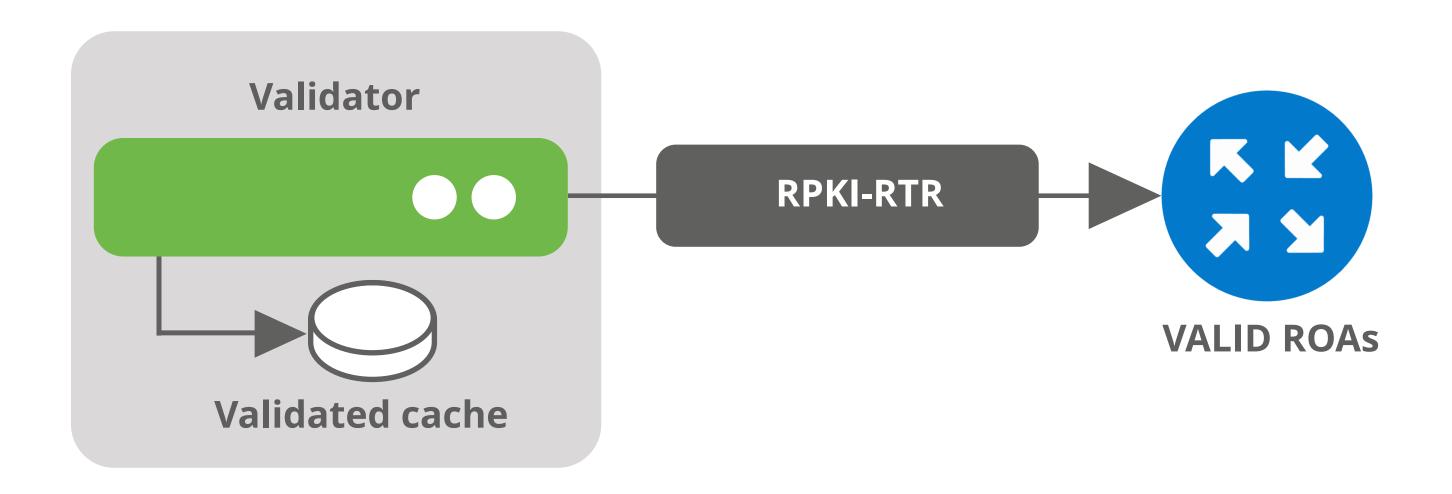


Your AS number AS X00
Your IPv6 allocation 2001:db8:X00::/48

# Lab 5.2 RPKI - BGP Origin Validation



- What have you learned?
  - How to use the GUI of Routinator validator software
  - The steps to follow to implement BGP OV in your network
  - How does RPKI OV works in a simple BGP Hijack scenario





# Next Steps for BGP Routing Security

Section 4

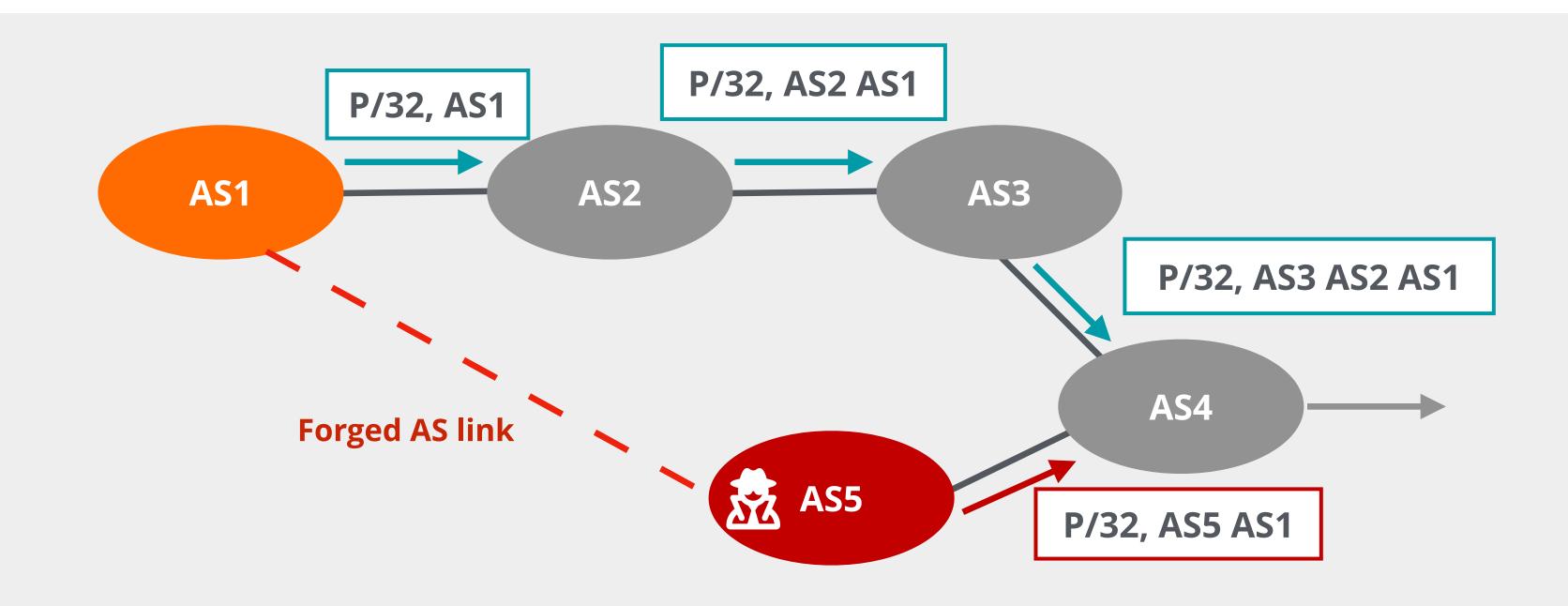
# What's Next for Routing Security?



# Fake Path with Correct Origin



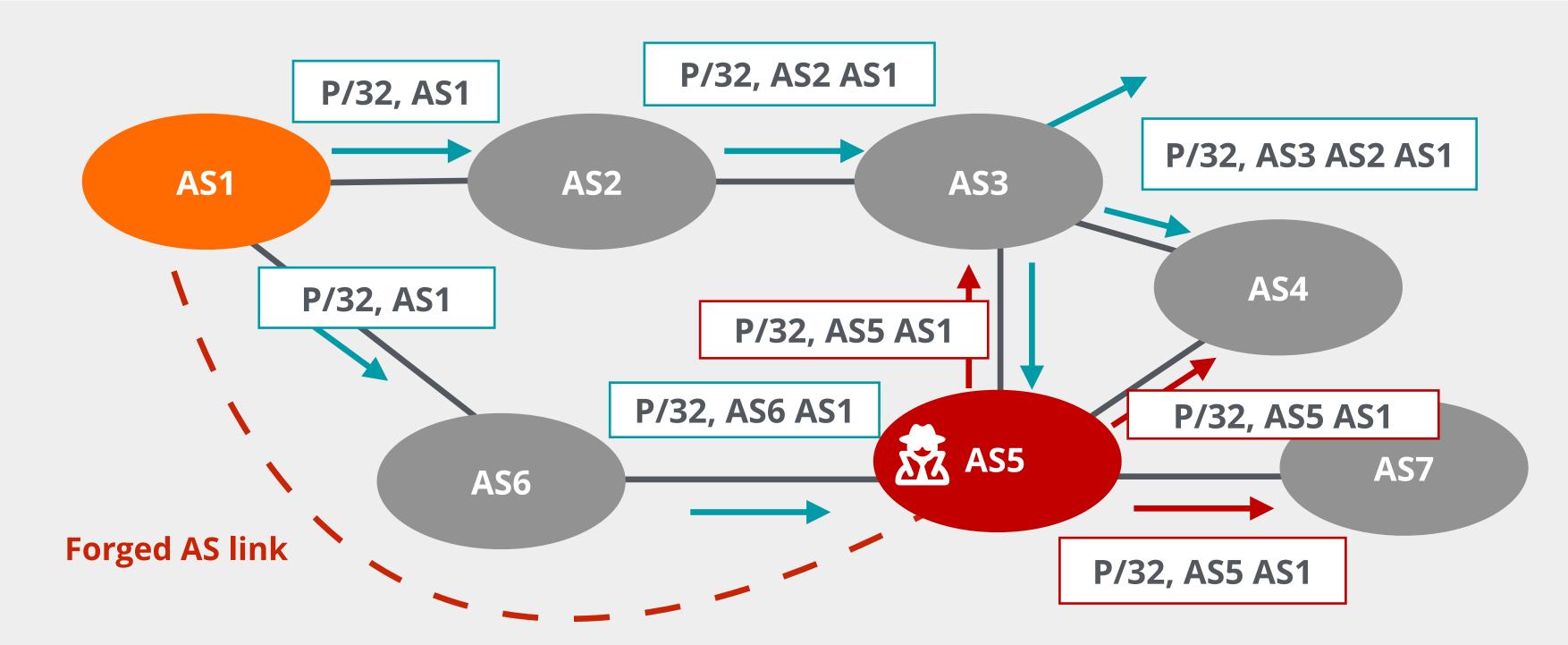
- The origin of the path does not change!
- The attacker:
  - Creates a forged AS link between two ASes
  - Reroutes the traffic to itself



# Modifying an Existing Path



- Neighbours of the attacker receive a false path
- The attacker can do either of these two things:
  - Analyse the traffic and then route to AS1
  - Drop the traffic to AS1



# What's Next for Routing Security?



- RPKI today focuses on Origin Validation!
- But RPKI OV can not detect path manipulations!
  - Origin AS remains intact in the altered AS Path
- Them, what to do?
- The solution is to validate the full BGP path
- Tentative solutions: BGPsec [RFC 8205] and ASPA

RPKI is a stepping stone to **Path validation**!

#### **BGPSec**



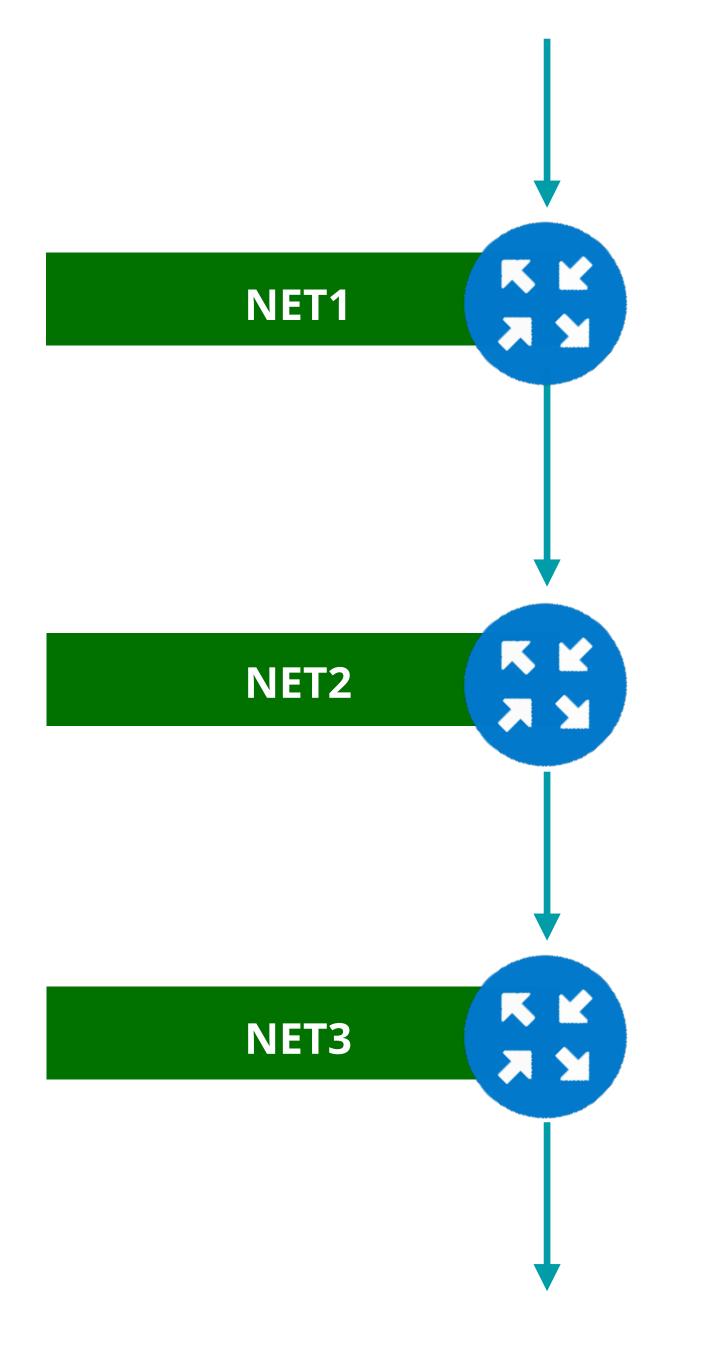
- Designed to supplement BGP Origin Validation
- Relies on the RPKI certificates
  - Router certificates are issued to routers within an AS
- Introduces a new BGP path attribute, BGPsec\_PATH
  - Optional, non-transitive attribute
  - Carries digitally signed AS path information
  - Support is negotiated between BGP speakers

## **BGP Operations**



- Routers sign the AS path in a BGP UPDATE message
- Each BGP UPDATE containing BGPsec\_PATH attribute:
  - Can advertise a single prefix only
  - Can only be **sent** to one AS at a time
- Routers verify the chain of trust of all of the signatures of the AS Path





Network: 192.168.0.0/16

AS Path: NET1, ...

**BGPSEC:** (key1, signature1)

Network: 192.168.0.0/16

AS Path: NET2, NET1, ...

**BGPSEC:** (key1, signature1)

(key2, signature2)

Network: 192.168.0.0/16

AS Path: NET3, NET2, NET1, ...

**BGPSEC:** (key1, signature1)

(key2, signature2)

(key3, signature3)

### **BGPSec Has Some Limitations...**



- Does not offer origin validation
- Does not prevent route leaks
- Expensive to run, requires more powerful routers
  - UPDATE messages are larger because of digital signatures
  - One UPDATE message is required for each prefix
  - BGP speakers need to perform cryptographic functions
- Does not support incremental deployment

That's why progress is very slow and no deployment yet!

#### **ASPA**



- Autonomous System Provider Authorisation
- Introduces a new digitally signed object, an ASPA
  - ASPA object defines upstreams for a defined AS
- ASPA proposes a lightweight solution for path validation
  - Leverages existing RPKI infrastructure
  - Does not require a new BGP attribute
  - Requires a database where ASPA objects could be queried
  - Verifies the sequence of ASes along the path

#### **How Does ASPA Work?**



- Customer AS creates an ASPA object and signs it
  - Authorises a set of **Provider ASes** to propagate its route announcements
- In the Validation process, receiving AS
- Verifies that if there is a cryptographically valid ASPA for that customer
  - Is provider AS authorised to propagate a given customer's route?
- 2 Verifies the AS path
  - Have routes been received from a customer, a provider, or from a route server?

If validation fails, then the route should be rejected!

#### More About ASPA



- ASPA helps to detect route leaks and hijacks
- Incremental deployment is possible
- Still in draft state (about to become an RFC)
- Already supported in a couple of validators
- Support in OpenBGPD and NIST BGP-SRx



# Best Practices

Section 5

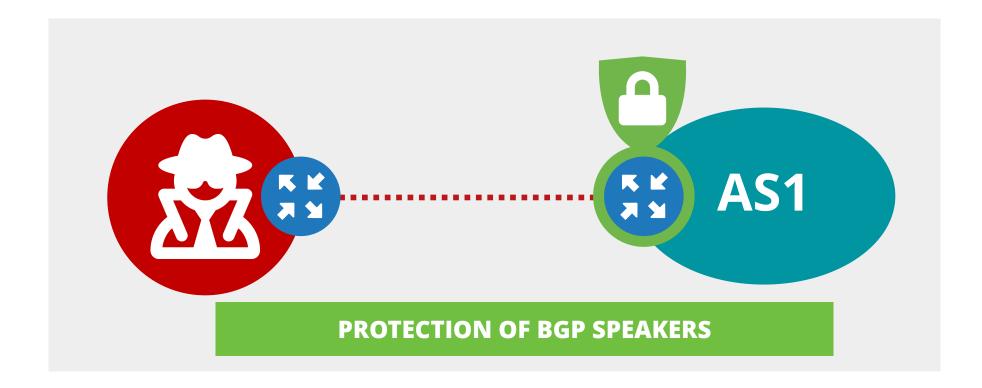
# For Secure Internet Routing



- Do not be the cause!
  - Announce the right prefixes to the right peers
- Do not distribute others' mistakes or attacks!
  - Validate the routing information you receive
- Do not be the victim!
  - Take all the measures you can to **protect** your network

## **BGP Security Measures**

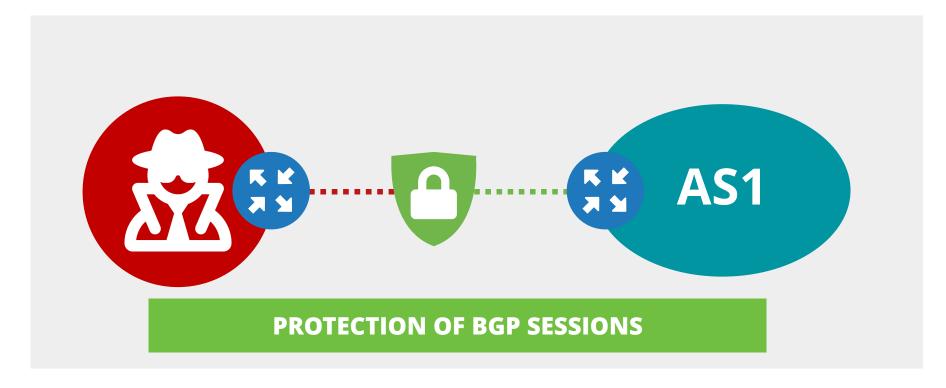




Only BGP peers to send packets to TCP 179: Control Plane Policing (CoPP) or ACLs (if CoPP not supported)

Limit accepted BGP traffic

uRPF to mitigate DoS/DDoS attacks



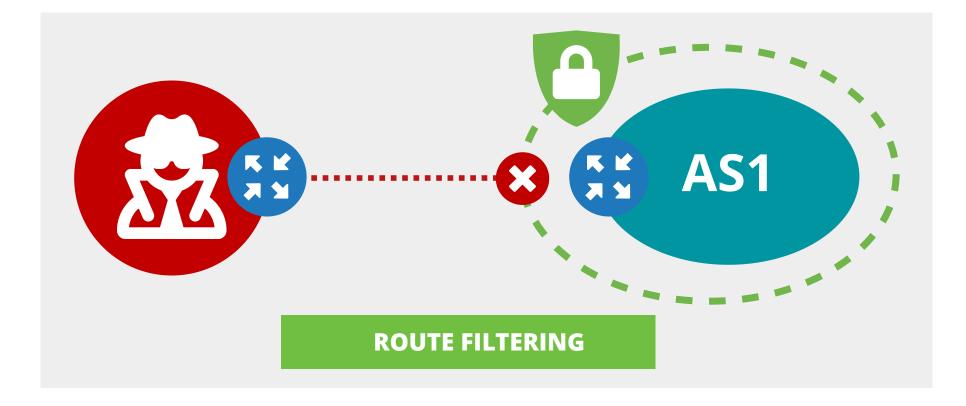
MD5



TCP-AO

BGP TTL Security (GTSM)





Inbound/outbound filters by type of peer



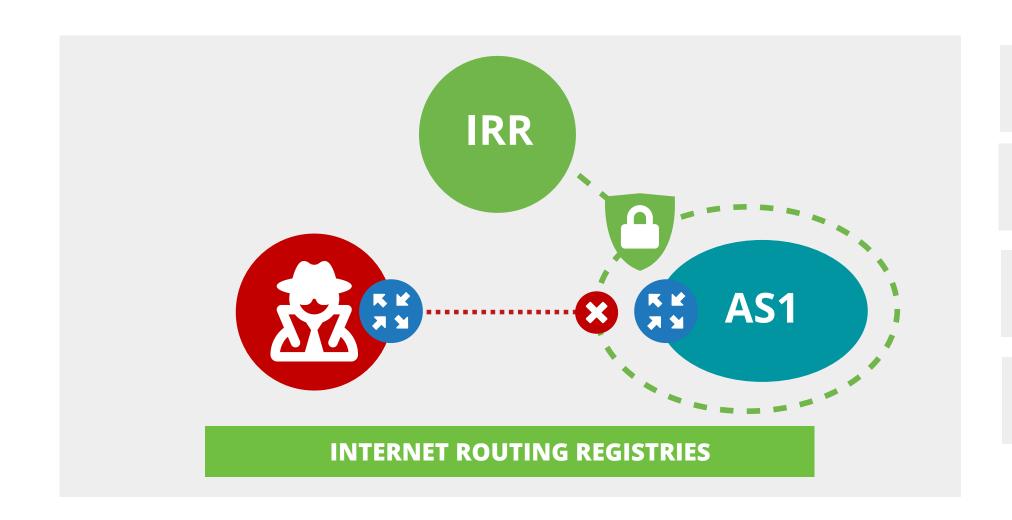
Filtering based on prefixes and/or AS Path



Manual or Automatic (IRRs) configuration

# **BGP Security Measures**





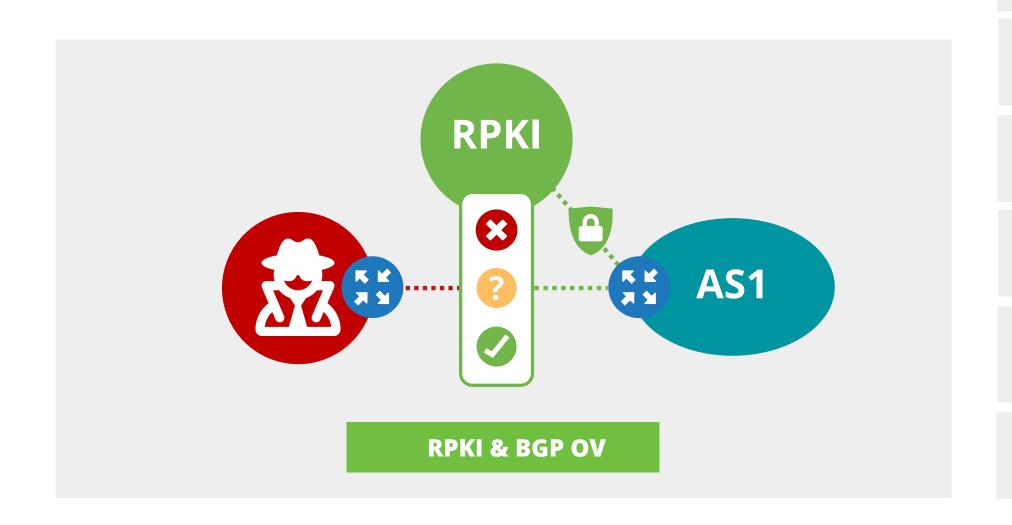
Create your route(6) objects in IRR



Update your IRR info regularly

Automate filters creation (IRR) (double check)

Document your Routing Policy



Create ROAs for your address space



max-length considerations (your most specific)

or, one ROA per prefix and more-specific

Implement BGP Origin Validation (OV)



Discard INVALID routes

Never discard NOT-FOUND routes

# **Check Your Routing**



- RIPEstat
  - https://stat.ripe.net
- IXP Country Jedi
  - https://jedi.ripe.net/latest/
- Bgpmon
  - https://routeviews.org
  - http://traceroute.org
- IRR Explorer
  - https://irrexplorer.nlnog.net

- RIPE Atlas
  - https://atlas.ripe.net
- NLNOG Ring
  - https://ring.nlnog.net
- HE BGP Toolkit
  - https://bgp.he.net
- BGP.tools
  - https://bgp.tools/



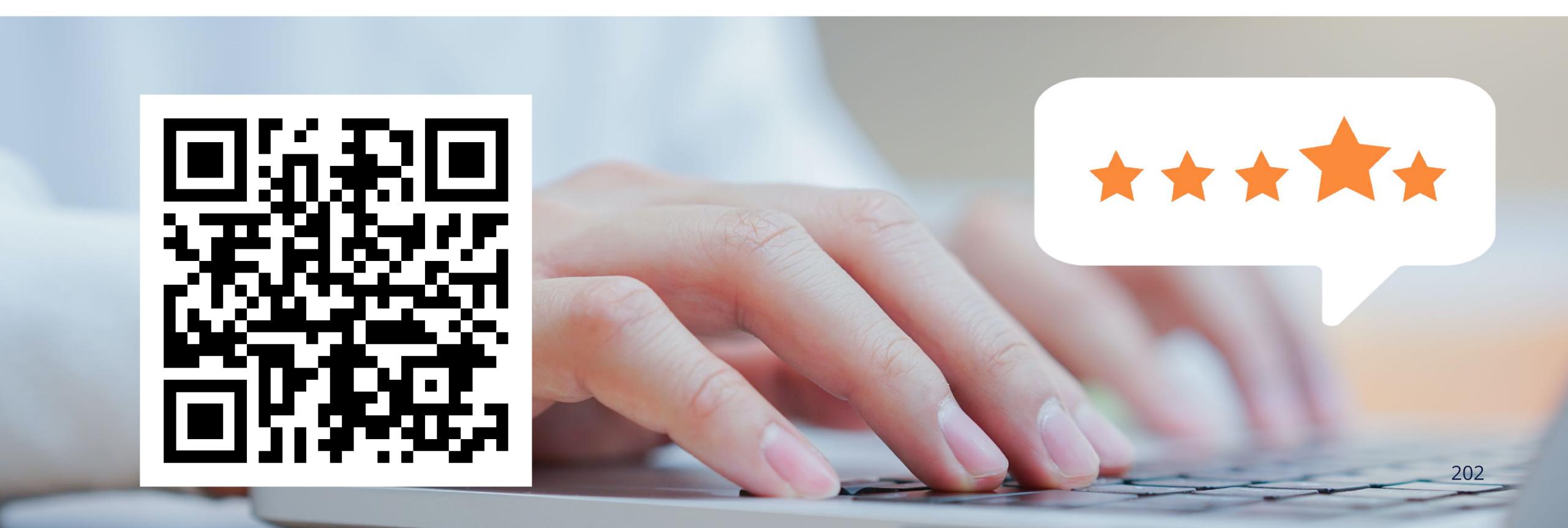
# Questions

# We want your feedback!



What did you think about this session? Take our survey at:

https://www.ripe.net/feedback/bgp/





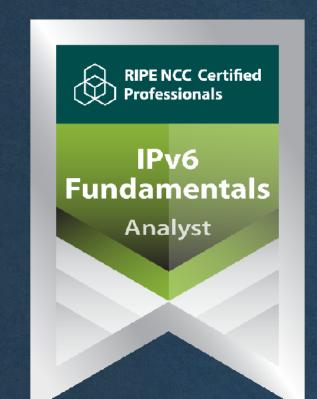
Learn something new today!

academy.ripe.net





# RIPE NCC Certified Professionals









https://getcertified.ripe.net/



## What's Next in BGP?



Webinars	Face-to-face	E-learning	Examinations	
Attend another webinar live Meet us at a location near you for a training session delivered in person.		Learn at your own pace at our online Academy.	Learnt everything you needed? Get certified!	
BGP Filtering (1 hr)		BGP Security (10 hrs)	BGP Security Associate	
Deploying RPKI (2 hrs)	DCD Douting Consulting			
Introduction to RPKI (1 hr)	*BGP Routing Security (6.5 hrs)			
Internet Routing Registry (1 hr)				
For more info click the link below		For more info click the link below	For more info click the link below	
learning.ripe.net		academy.ripe.net	getcertified.ripe.net	

Ënn	Соңы	An	Críoch	پایان	Ende	Y Diwedd
Vége	Endir		invezh			Koniec
Son	დასასრუ			վերջ	Кінець	Tin:
Lõpp	۸maia		הסוך	Tmiem	Liðugt	Finis
	Amaia	Loppi		Slutt		Kpaj
Kraj	Sfârşit	النهاية	Конец		Konec	Fund
Fine	Fin	Einde	Fí	Край	Beigas	Τέλος
Fim	Slut					Pabaiga

A STATE OF THE PARTY OF THE PAR

# Copyright Statement

The RIPE NCC Materials may be used for **private purposes**, **for public non-commercial purpose**, **for research**, **for educational or demonstration purposes**, or if the materials in question specifically state that use of the material is permissible, and provided the RIPE NCC Materials are not modified and are properly identified as RIPE NCC documents. Unless authorised by the RIPE NCC in writing, any use of the RIPE NCC Materials for advertising or marketing purposes is strictly forbidden and may be prosecuted. The RIPE NCC should be notified of any such activities or suspicions thereof.

[...]

#### Find the full copyright statement here:

https://www.ripe.net/about-us/legal/copyright-statement

