University of Rome Tor Vergata
ICT and Internet Engineering

Network and System Defense

Alessandro Pellegrini, Angelo Tulumello

A.A. 2023/2024

Lecture 10: BGP Security

Angelo Tulumello

Slides by Marco Bonola

sources:
[1] Sriram, Kotikalapudi, and Doug Montgomery. "Resilient Interdomain Traffic Exchange." NIST Special Publication 800 (2019): 189. Available online @ https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-189.pdf
[2] Kuhn, Rick, Kotikalapudi Sriram, and Doug Montgomery. "Border gateway protocol security." NIST Special Publication 800 (2007): 54. - Withdrawn NIST Technical Series Publication

Intro

- BGP security is still a big issue
  - see some famous attacks reported in slide 20-21
- BGB routing ***Control Plane*** threats
  - Border Gateway Protocol (BGP) prefix hijacking,
  - Route leaks
  - Other forms of misrouting resulting in denial-of-service (DoS)
  - Unwanted data traffic detours
  - Performance degradation
- ***Data Plane*** attacks (we'll cover this in a separate class)
  - Large-scale distributed DoS (DDoS) attacks on servers using spoofed internet protocol (IP) addresses
  - Reflection amplification
- Threats related to the ***underlying protocols*** (IP/TCP)
  - which are nowadays considered more as "historical vulnerabilities" - still interesting...
- [1] provides technical guidance and recommendations for technologies that facilitate ***Resilient Interdomain Traffic Exchange (RITE)***
  - [1] superseded [2], which was used anyway for this class as it covers transport layer security related to BGP
  - [1] represents both a good survey and a good summary as well

Further Reading: BGP Security Best Practice

NSA 2018 report: https://apps.nsa.gov/iad/library/reports/a-guide-to-border-gateway-protocol-bgp-best-practices.cfm

Table of Contents

IP/TCP vulnerabilities related to BGP and relative countermeasures

problema da affrontare

Peer Spoofing and TCP Resets

- ***BGP messages can be spoofed*** in order to look like a message from a valid peer and insert false information into a BGP routing tables
- IP addresses can often be found using the ICMP traceroute function, so BGP implementations should include countermeasures against this attack
- A special case of peer spoofing, called a ***reset attack***, involves inserting TCP RESET messages into an ongoing session between two BGP peers
- When a reset is received, the target router drops the BGP session and both peers withdraw routes previously learned from each other
  - thus disrupting network connectivity until recovery, which may take several minutes to hours, depending on the number of BGP peers affected

Peer Spoofing and TCP Resets Countermeasures

Table 3-1. Peer Spoofing Countermeasures

Method	Reference or RFC	Strength	Cost	Notes
Strong sequence number randomization	CERT Advisory CA-2001-09 [5]	M	L	Varies with the underlying operating system See Section 4.3
TTL Hack	RFC 3682	M	L	Simple configuration option; not effective against machines one hop away See Section 4.4
MD5 Signature option	RFC 2385	H	M	Widely available option; may be significant administrative cost See Section 4.5

TCP Resets Using ICMP

- ***ICMP*** can be used to produce ***session resets***
- Older IETF specifications do not require checking sequence numbers of received ICMP messages
- Only require knowledge of the victim's IP address and port number

Table 3-2. TCP Reset Countermeasures

Method	Reference or RFC	Strength	Cost	Notes
TCP sequence number checking	[18][53]	M	L	Varies with the underlying operating System. Included on Linux, FreeBSD, OpenBSD.
TTL Hack	RFC 3682	M	L	Simple configuration option; not effective against machines one hop away See Section 4.4
Router access control	[18]	H	M	Block packets of ICMP Type 3 codes 2, 3, and 4 See also NISCC Vulnerability Advisory ICMP – 532967 [53]
IPsec authentication	[45]	H	M	Widely available; may be significant administrative cost See Section 4.6.

Session Hijacking

è difficile da performare, è complessa la rete internet. La non applicapibiltà di questo attacco è dovuto al fatto che tutti gli AS sono pubblici ma l'ingresso agli AS dall'esterno è consentito solo da un gateway, questo aumenta la sicurezza.

- ***Session hijacking involves intrusion into an ongoing BGP session***
- Requires similar (a bit more) information needed to accomplish the reset attack
- Hijacking attack may be designed to achieve more than simply bringing down a session between BGP peers
  - For example, the objective may be to change routes used by the peer, in order to facilitate eavesdropping, blackholing, or traffic analysis

Table 3-3. Session Hijacking Countermeasures

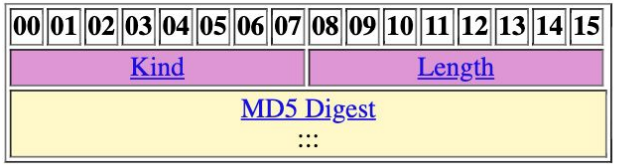
Method	Reference or RFC	Strength	Cost	Notes
Strong sequence number randomization	CERT Advisory CA-2001-09 [5]	M	L	Varies with the underlying operating system See Section 4.3
TTL Hack	RFC 3682	M	L	Simple configuration option; not effective against machines one hop away
MD5 Signature option	RFC 2385	H	M	Widely available option; may be significant administrative cost See Section 4.5
IPsec	RFC 4301, plus many related RFCs (RFCs 4302-4309)	H	H	See Section 4.6

TCP MD5 Option

contromisura è l'autenticazione del tcp header

- The MD5 hash algorithm can be used to protect BGP sessions by creating a keyed hash for TCP message authentication (***RFC 2385***)
- Every segment sent on a TCP connection to be protected against spoofing will contain the ***16-byte MD5 digest*** produced by applying the MD5 algorithm to these items in the following order:
  - TCP pseudo-header (in the order: source IP address, destination IP address, zero-padded protocol number, and segment length)
  - TCP header, excluding options, and assuming a csum of zero
  - TCP segment data (if any)
  - independently-specified key or password, known to both TCP and presumably connection-specific

TCP Option 19:



Kind. 8 bits. Set to 15.

Length. 8 bits. Set to 18.

MD5 Digest. 16 bytes.

TCP Authentication Option

calcola message authentication code

- Even though you find many documents referring to the use of TCP MD5 signature to protect BGP sessions, this option (RFC 2385) has been obsoleted
- ***RFC 5925*** specifies the ***TCP Authentication Option*** (TCP-AO)
- It supports strong Message Authentication Codes (MACs) algorithms
  - From RFC 5926: HMAC-SHA-1-96 and AES-128-CMAC-96
- It provides better key generation/coordination
  - Different traffic keys derived from the Master Key Tuples
  - Different MKT can be associated to different TCP parameters
  - MKT can change event in the same connection (long lived flow protection)
- It does not specify how to negotiate the MKT
- H-MAC input: Sequence Number Extension (SNE), IP pseudo header, TCP header (with TCP AO set to zero), TCP payload

time to live

TTL Hack as an alternative to TCP authentication

- ***The TTL is an 8-bit field in each IP packet*** that prevents packets from circulating endlessly in the Internet

  tra me e google.com ci sono ad esempio 10 hops = router nell'internet

  - At each network node, the TTL is decremented by one, and is discarded when it is reduced to zero without reaching its destination

    utile nel caso di loop
- It is not unusual for 20 or more hops to be required before a packet is finally received, so a packet that starts with a value lower than this has a high probability of being discarded before it reaches its intended destination
  - ***With BGP, however, peers are normally adjacent***, thus only one hop should be required for a packet sent in a BGP message
  - eBGP peer are always adjacent
  - iBGP not necessarily (but in many cases the AS has an MPLS backbone)
- The ***Generalized TTL Security Mechanism (TTL Hack)*** sets the TTL to 255 on outgoing packets
  - Adjacent peers should see incoming packets with TTL = 255
  - When implementing the TTL hack, it is also possible to set an expected incoming value
    - 255 for adjacent peers, 254 for peers @ 1 hop, 253 @ 2 hop, etc..

parliamo di Hack poichè non è un vero e proprio metodo.

Vulnerability: Route Flapping

- ***Route flapping refers to repetitive changes to the BGP routing table,*** often several times a minute.
  - A “route flap” occurs when a route is withdrawn and then re-advertised
- ***High-rate route flapping can cause a serious problem for routers***, because every flap causes route changes or withdrawals that propagate through the network of ASes
  - If route flaps happen fast enough – e.g., 30 to 50 times per second – the router becomes overloaded, eventually preventing convergence on valid routes
- The potential impact for Internet users is a slowdown in message delivery, and in some cases packets may not be received at all
- In other words, route flapping can result in a denial of service, either accidental or from an intentional attack

Countermeasure: Route Flap Damping

- ***Route flap damping*** is a method of reducing route flaps by implementing an algorithm that ignores the router sending flapping updates for a configurable period of time
- Each time a flapping event occurs, peer routers ***add a penalty value*** to a total for the flapping router
- The penalty decays exponentially over time
  - If route flaps persist often enough, the total exceeds a configurable cutoff threshold
  - If no further flaps are seen, the penalty will reach a reuse threshold
- While this mechanism helps to reduce instability caused by spontaneous faults in the network, it can be misused by an attacker
  - Because of the potential problems described above, and the fact that faster processors in routers have obviated some of the original need for RFD, network administrators should be cautious about enabling RFD

What about IPsec for BGP authentication?

The use of IPsec to protect BGP is considered also by the IETF in RFC 7454 “BGP Operations and Security” (2015)

<< IPsec could also be used for session protection. At the time of publication, there is not enough experience of the impact of using IPsec for BGP peerings, and further analysis is required to define guidelines >>

In the past somebody thought that for several reasons BGP and IPsec don’t get along well in real networks (IETF proceeding from 2006 https://www.ietf.org/proceedings/66/slides/saag-2.pdf)

- ❏ Key distribution with static SAs
- ❏ IKE rekeying
- ❏ IKEv1 DoS vulnerability
- ❏ Cost of public key operation

What really happens is that IPsec is not really used for protecting BGP

How do we proceed from here..

- [2] was withdrawn because, IMHO, the vulnerabilities described are
  - either pretty unlikely (e.g. ISP already implements anti spoofing mechanisms)
  - more theoretical than practical (e.g. are we really able to guess the correct TCP sequence number and break into an already established TCP connection?)
  - solved by some widely adopted countermeasures (e.g. TTL hacking or TCP auth implemented by almost all vendors)
- The real problems todays are others (DDoS and BGP Hijacking)
  - [2] partially addresses these aspects..
  - … [1] is much more focused on these topics!

BGP control plane vulnerabilities

1. Prefix Hijacking and Announcement of Unallocated Addresses

A BGP prefix hijack occurs when an AS ***accidentally*** or ***maliciously*** originates a prefix that it is not authorized (by the prefix owner) to originate

If an AS is authorized to originate/announce a prefix by the prefix owner, then such a route origination/announcement is called ***legitimate***

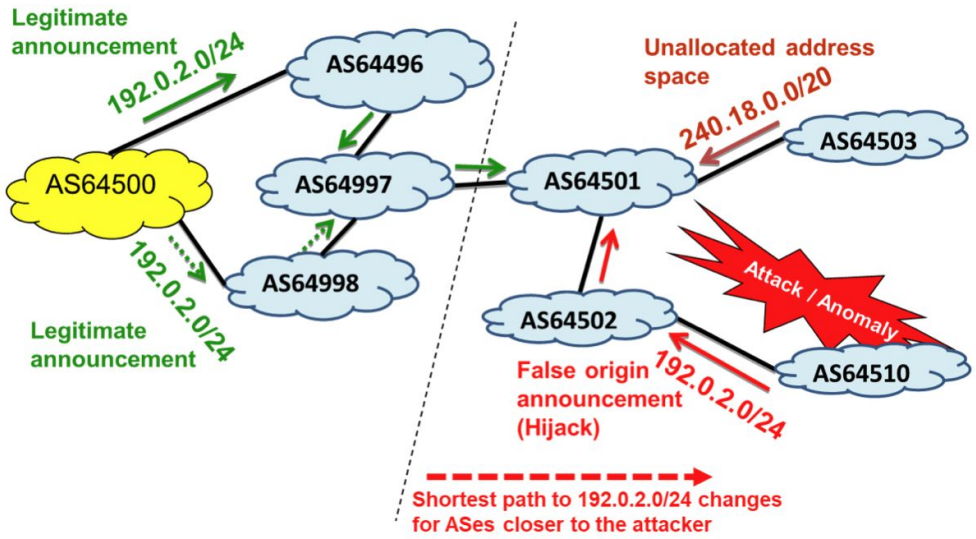


Fig 1 BGP Hijacking Example

AS64510 invia lo stesso IP, c'è un problema non è legittimo è lo stesso di AS64500. C'è una parte della rete che ha accettato l'announcement fatto in modo malevolo (sono più vicini all'AS).

Prefix hijack, sub-prefix hijack and prefix squatting

- In ***Figure 1***, prefix 192.0.2.0/24 is legitimately originated by ***AS64500***, but ***AS64510*** falsely originates it
- The path to the prefix via the false origin AS ***will be shorter for a subset of the ASes on the internet*** (it will be installed in their routing tables)
  - ASes for which AS64510 is closer would choose the false announcement
  - data traffic in those ASes destined for the network 192.0.2/24 will be misrouted to AS64510
- ***Things are even worse!***
- IP route selection always prefer the most specific rule (i.e. longest prefix)
  - When an AS falsely announces a more-specific prefix (than a prefix announced by an authorized AS), the ***longer, unauthorized prefix will be widely accepted***
  - in Figure 1 unauthorized origination of reserved address space 240.18.0.0/20 (240.0.0.0/8 is reserved for future use)
- Similarly, an AS may also falsely originate allocated but currently unused address space. This is called ***prefix squatting***

altra variante, c'è prefisso non posseduto/usato da altri, e lo usiamo noi

BGP Hijack Consequences (1)

- ***The consequences of such adverse actions can be serious and include***
  - denial-of-service, eavesdropping, misdirection to imposter servers (to steal login credentials or inject malware), defeat of IP reputation systems to launch spam email
- There have been numerous incidents involving prefix hijacks in recent years
  - ***December 2017:*** Eighty high-traffic prefixes usually used by ***Google, Apple, Facebook, Microsoft, TwitchTV and Riot Games*** were hijacked by an unknown Russian Autonomous System (AS) simply known as DV-LINK-AS (AS39523). User information such as email addresses, passwords, usernames and other login details were suspected to be compromised
  - ***April 2018:*** Approximately 1300 IP addresses belonging to ***Amazon Web Services*** were hijacked by eNet (or a customer of theirs), an ISP in Columbus, Ohio. Several partners, such as Hurricane Electric routed traffic through the hijacked addresses, exacerbating the issue. ***The attacker was suspected to be after cryptocurrency, stealing a total of about $150,000 from MyEtherWallet users***
  - ***November 2018***: China Telecom was suspected of hijacking a total of 180 prefixes, affecting a vast scope of ***Google*** services, including a massive denial of service to GSuite and Google Search. Regardless of whether intention was involved, ***valuable Google traffic data fell into the hands of the attackers***

https://uniroma2-my.sharepoint.com/:u:/g/personal/michele_tosi_students_uniroma2_eu/EcLx7L1rM7pBn83V5oZTLXYBMdROXt1aZSevC3WBdM1y0Q?e=YhNCQl

BGP Hijack Consequences (2)

- ***May 2019***: Taiwan fell victim to an unknown Brazilian attacker using two prefixes for advertising purposes that belonged to The ***Taiwan Network Information Center***, a non-profit organisation officially funded by the Taiwanese Directorate General Telecommunications of the Ministry of Transportation and Communication. ***The attack lasted three and a half minutes where public data was vulnerable***
- ***June 2019***: A ***Swiss data centre hosting company accidentally leaked over 70 000 routes from its internal routing table to China Telecom***. Instead of ignoring the BGP leak, ***China Telecom re-announced these routes*** as its own and declared itself as the shortest way to reach the network of the Swiss data centre operator and other nearby European telecommunication companies and ISPs. Some of the most impacted European networks included Swisscom (AS3303) of Switzerland, KPN (AS1130) of Holland, and Bouygues Telecom (AS5410) and Numericable-SFR (AS21502) of France. ***This particular incident was severe, lasting over two hours. Users of the affected networks suffered slow connections and denial of service to some servers.***
- ***April 2020***: A massive BGP hijack involving over ***8800 prefixes affected companies such as Akamai, Amazon and Alibaba*** on April 1, 2020. Initiated by a Rostelecom user, the attack caused service disruptions throughout the world. It is currently unknown how much data was leaked or for what purposes, but it generally acknowledged that stricter network filtering by Rostelecom could have prevented the attack

2. BGP UPDATE modification (1)

- BGP messages carry a sequence of AS numbers that indicates the “path” of interconnected networks over which data will flow (*we know this...*)
- This ***AS_PATH*** data is often used to implement routing policies that reflect the business agreements and peering policies that have been negotiated between networks
- ***BGP AS_PATH modification***
  - A malicious AS which receives a BGP update may illegitimately remove some of the preceding ASes in the AS_PATH attribute of the update to make the ***path length seem shorter***
  - ASes upstream can be deceived to believe that the path to the advertised prefix via the adversary AS is shorter
  - the adversary AS may seek to illegitimately increase its revenue from its customers, or may be able to eavesdrop on traffic that would otherwise not transit through their AS

2. BGP UPDATE modification (2)

- ***BGP Prefix modification***

in questo caso viene modificato il prefisso in uno più specifico, nel caso di hijacked viene inviato un indirizzo falso

  - An adversary AS replaces a prefix in a received update with a more-specific prefix (subsumed by the prefix) and then forwards the update to neighbors
  - ***Kapela-Pilosov attack***
    - only the prefix is replaced by a more-specific prefix, but the AS path is not altered
    - this means that ASes on the internet would widely accept and use the adversary AS's advertisement for the more-specific prefix
  - The exceptions are the ASes that are in the AS path from the adversary to the prefix
    - Standard BGP loop detection
  - The adversary would be able to force almost all traffic for the more-specific prefix to be routed via their AS
  - They can eavesdrop on the data (destined for the more-specific prefix) while channeling it back to the legitimate destination to avoid detection

non è un problema di BGP o dei suoi protocolli, bensì tra le policies tra due AS che non vengono rispettate (agreements, etc..)

3. Route Leaks: definitions of peering relations

BGP peer non rispetta le policies con un altro AS

- BGP peering policies often specify limits on what routing announcements will be accepted by each party
- Policies reflect customer, transit provider, and/or lateral peer business relationship between networks
- Definitions of peering relations (useful to understand what follows):
  - ***Transit provider:*** typically provides service to connect its customer(s) to the global internet
  - ***Customer AS:*** single-homed to one transit provider or multi-homed to more than one transit providers
  - ***Stub customer:*** an AS that has no customer ASes or lateral peer ASes of its own
  - ***Leaf customer:*** is a stub customer that is single-homed to one transit provider (and no any other ASes)
  - ***Peering relationships: provider-to-customer (P2C)***, ***customer-to-provider (C2P)***, ***peer-to-peer (p2p)***
    - ***Public Peering***
    - ***Private (Bilateral) Peering***
    - ***Upstream (Transit Peering)***
    - ***Downstream (Customer Peering)***
  - ***Customer cone of AS A:*** AS A plus all the ASes that can be reached from A following only P2C links
  - ***Customer cone prefixes***: the union of the prefixes received from all directly connected customers and the prefixes originated by the AS itself

Se Google ha nuovo prefisso, il customer deve mantenere l'informazione e e non deve annunciarla ad altri peers.
E' il router che propaga l'annuncio che deve farlo. Prima va al customer, poi all'ISP NON va direttamente da ISP1 a ISP2.

3. Route Leaks: basics and a simple example

- ***A route leak is the propagation of routing announcements beyond their intended scope***
  - i.e. in violation of the intended policies of the receiver, the sender, and/or one of the ASes along the preceding AS path [RFC 7908]
- In Figure 2 AS3 “leaks” the update from AS1 to AS2, which in turn propagate the update to its peers
- A stub or customer AS should never be traversed between 2 transit ASes
- stub or customer ASes do not pass BGP routing information received from one transit provider to another

In general, ISPs prefer customer routes over those from others

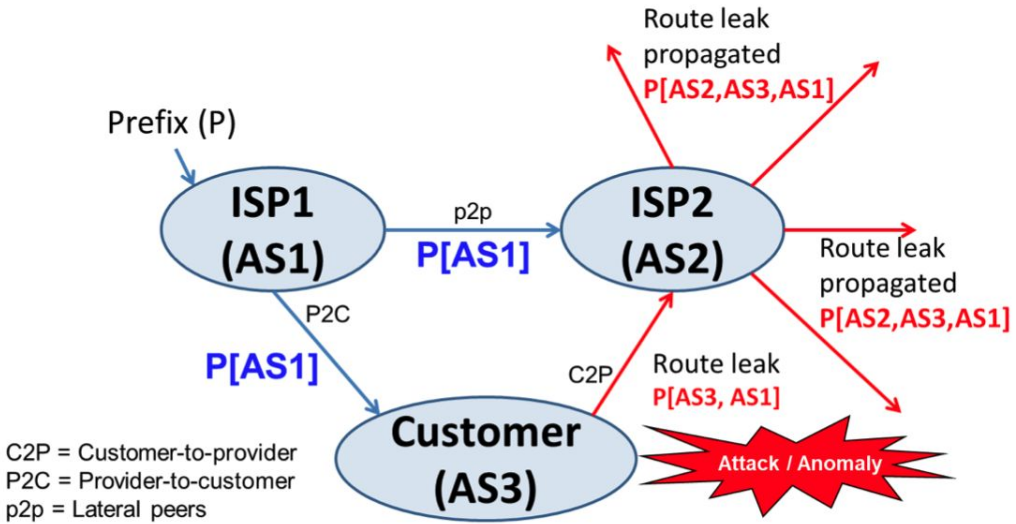
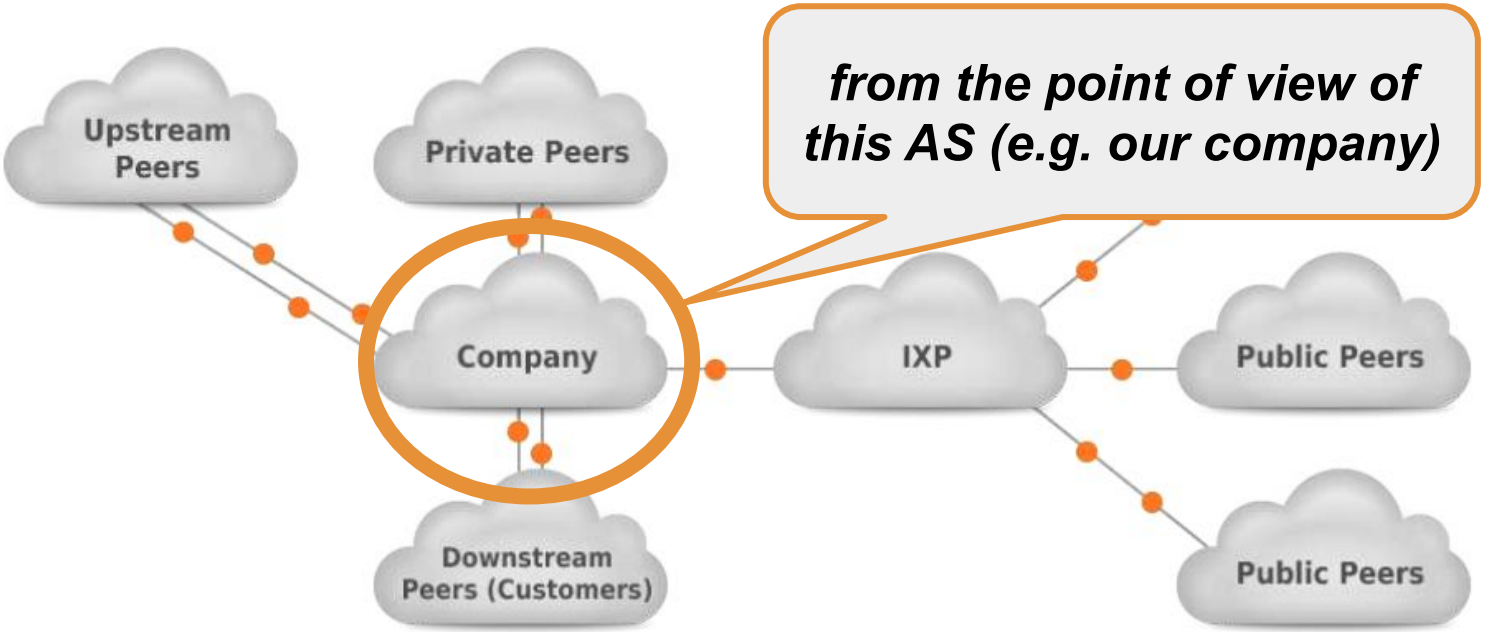


Fig 2 Illustration of the basic notion of a route leak

intended scope: sono un ISP e ricevo Google network prefix e lo do al cliente per raggiungerlo. Il customer. Il customer dovrebbe mantenere questa informazione senza condividerla. Il problema si ha se la condivide ad esempio ad un'altra ISP. Questa ISP può preferire il route tramite il customer e propaga questo link [AS3, AS1].

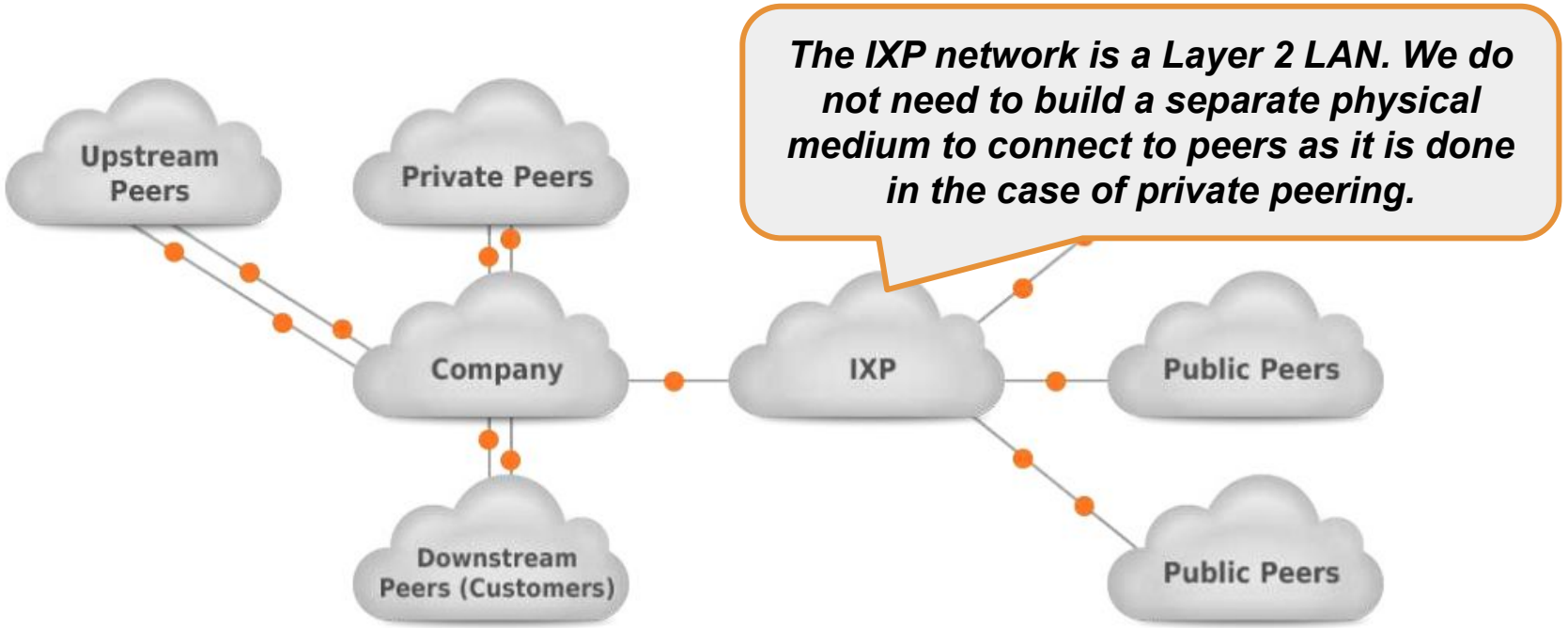
More about peering relationships

Siamo l'azienda cerchiata ad esempio un ISP.



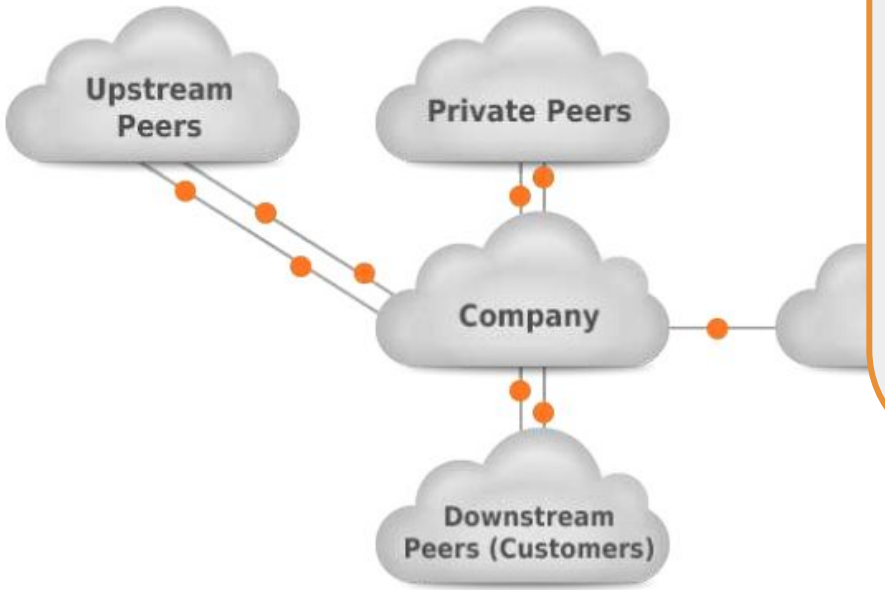
More about peering relationships

svariati AS sono collegati nel network, e scambiano BPG announcements (routes). In italia c'è Namex, con sede a San Lorenzo.



IXP: Intenet Exachange Point, aree di rete in cui diversi AS possono collegarsi per scambiare BGP announcement, prefix

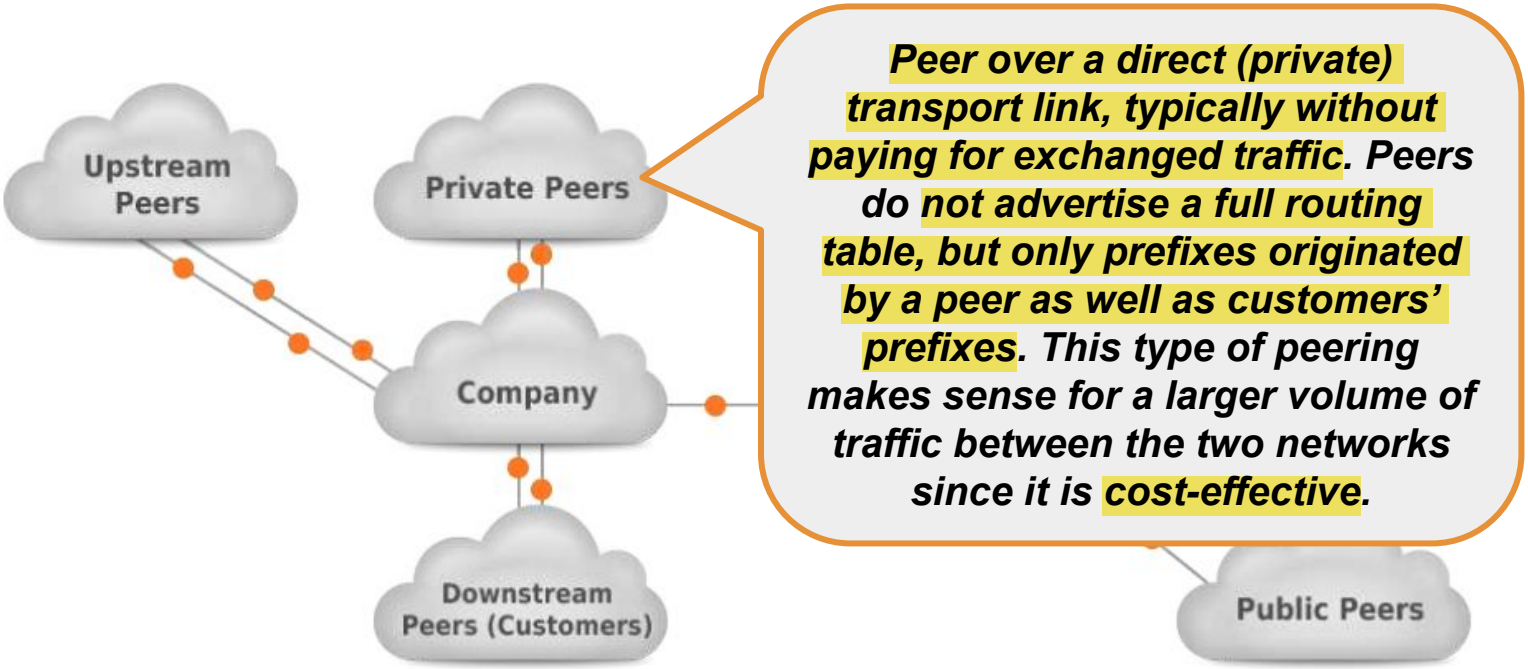
More about peering relationships



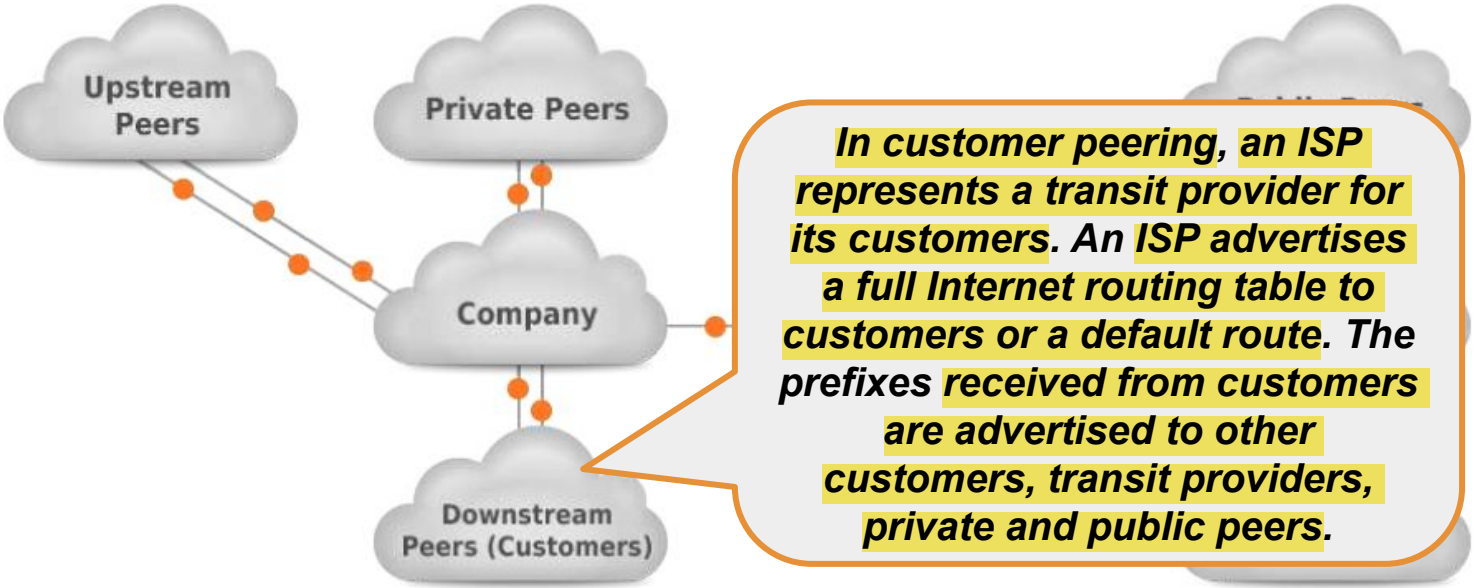
Public peering is a type of relationship where two ISPs exchange prefixes and traffic via a single public IXP.
Typically, each ISP participating in IXP brings its own border router that connects with the Ethernet port to the IXP LAN. The WAN port is used to connect back to the ISP network. The ISP's routers peer with each other (if no route server is used) using eBGP and can freely exchange own prefixes and IP traffic (without fees).

Non c'è relazione customer-provider, ma solo "laterali", e quindi parliamo di public peers.

More about peering relationships



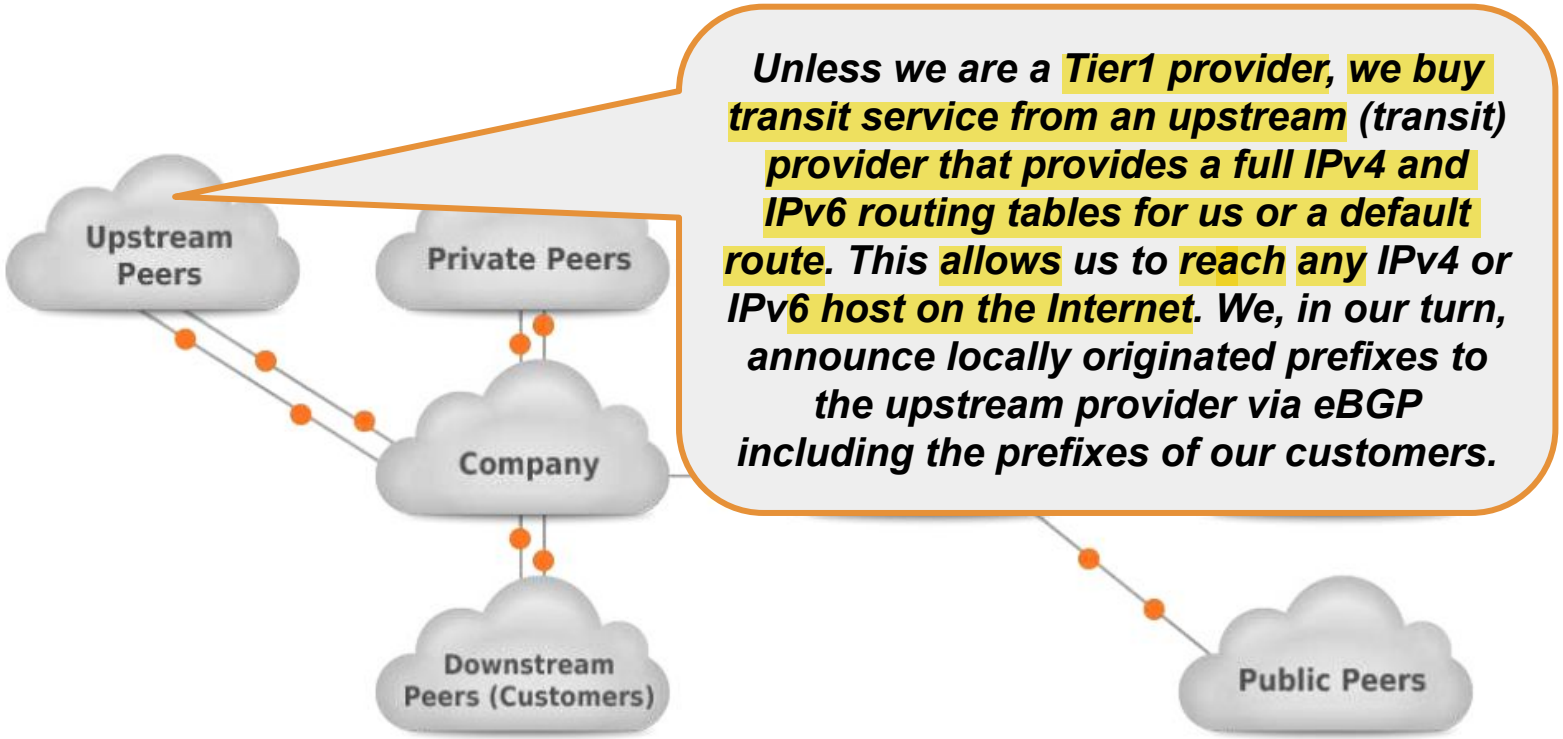
More about peering relationships



piccole compagnie ISP che fanno da tramite.

queste piccole ISP si appoggiano a questi Upstream Peers.

More about peering relationships



3. Route Leaks: well known incidents and consequences

Generalmente nascono da errate/mancate configurazioni, quasi mai attacchi.

- MainOne (a Nigerian ISP) leak of Google prefixes, which caused an outage of Google services for over an hour in ***November 2018***
- Dodo-Telstra incident in ***March 2012***, which caused an outage of internet services nationwide in Australia
- Massive Telekom Malaysia route leaks in ***June 2015***, which Level3, in turn, accepted and propagated
- Consequences of a route leak
  - ***Redirection of traffic through an unintended path***, which may enable eavesdropping or malicious traffic analysis
  - When a large number of routes is leaked simultaneously, the ***offending AS is often overwhelmed by the resulting unexpected data traffic*** and drops much of the traffic that it receives. This causes blackholing and denial-of-service for the affected prefixes.
- Route leaks can be accidental or malicious but most often arise from accidental misconfigurations.

RPKI and BGP Origin Validation

Resource public key infrastructure per le internet resources

Registration of Route Objects in Internet Routing Registries

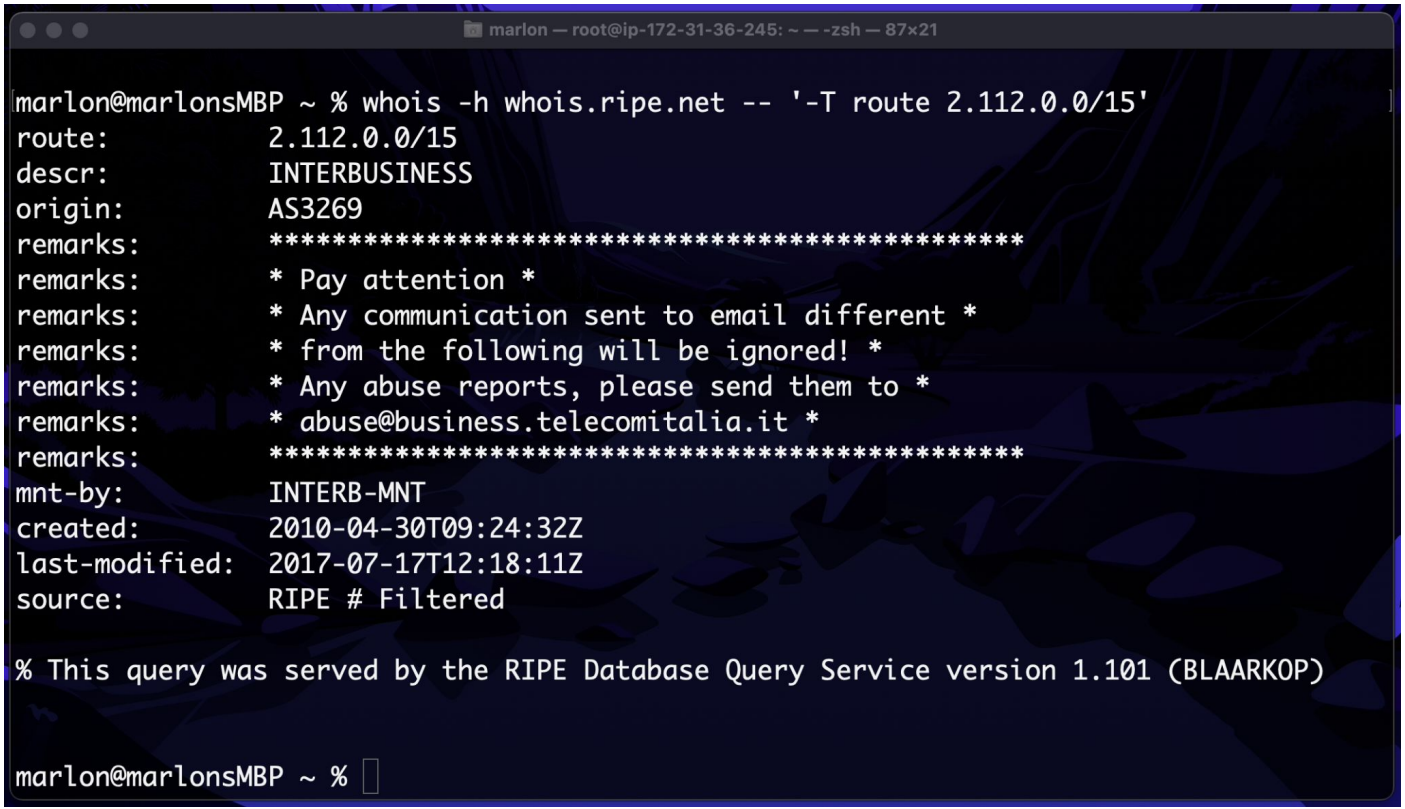
- Declarative data about internet resource allocations and routing policies have traditionally been available from ***regional internet registries (RIRs)*** and ***internet routing registries (IRRs)***
  - The ***RIR*** data are maintained regionally by ***ARIN*** in North America, ***RIPE*** in Europe, ***LACNIC*** in Latin America, ***APNIC*** in Asia-Pacific, and ***AfriNIC*** in Africa
  - The ***IRRs*** are maintained by the ***RIRs*** as well as some major ***ISPs***
  - Merit's Routing Assets Database (***RADb***) and other similar entities provide a collective routing information base consisting of registered (at their site) as well as mirrored (from the IRRs) data
- ***The route objects available in the IRRs provide routing information declared by network operators***
  - the route objects contain information regarding the origination of prefixes (i.e., the association between prefixes and the ASes which may originate them)
  - The completeness, correctness, freshness, and consistency of the data derived from these sources vary widely, and the data is not always reliable

Il problema di questo sistema centralizzato è che richiede update periodici, download delle informazioni nell'IRR

per evitare BGP hijack dovrei avere questa informazione su ogni BGP router

Querying the RIPE DB with whois (AS3269 = Telecom Italia)

se riceviamo bgp hijack dobbiamo vedere queste info e fare un check con l'annuncio.



Registration of Route Objects in Internet Routing Registries

- Network operators often obtain route object information from the IRRs and/or RADb for creating prefix filters in their BGP routers
- Efforts are encouraged to create complete and accurate IRR data in line with the current operational reality…
- … but even greater efforts should be devoted to creating ***route origin authorizations (ROAs)*** because ***RPKI*** provides a stronger authentication and validation framework for network operators than IRR (see next slides)

Security Recommendation: route objects corresponding to the BGP routes originating from an AS should be registered and actively maintained in an appropriate RIR's IRR. Enterprises should ensure that appropriate IRR information exists for all IP address space used directly and outsourced

Certification of Resources in Resource Public Key Infrastructure (1)

- ***Resource Public Key Infrastructure (RPKI)*** is a standards-based approach for providing cryptographically secured registries of internet resources and routing authorizations (see RFC 6480 and RFC 6482)
- To better understand the role of RPKI let's review the hierarchical approach for IPv4/IPv6 address and AS number resource allocations
  - The Internet Assigned Numbers Authority (***IANA***) allocates resources to the regional internet registries (***RIRs***)
  - The ***RIRs*** sub-allocate resources to ***ISPs*** and ***enterprises***
  - The ***ISPs*** may further suballocate to ***other ISPs*** and ***enterprises***
  - (In some regions) ***RIRs*** sub-allocate to ***local internet registries*** (***LIRs***) which in turn suballocate to ***ISPs*** and ***enterprises***
    - this is the case in Europe. RIPE delegates resource allocation to a number of LIRs
      - LIRs include ISPs (e.g. Telecom Italia) and enterprises (e.g. Aruba)

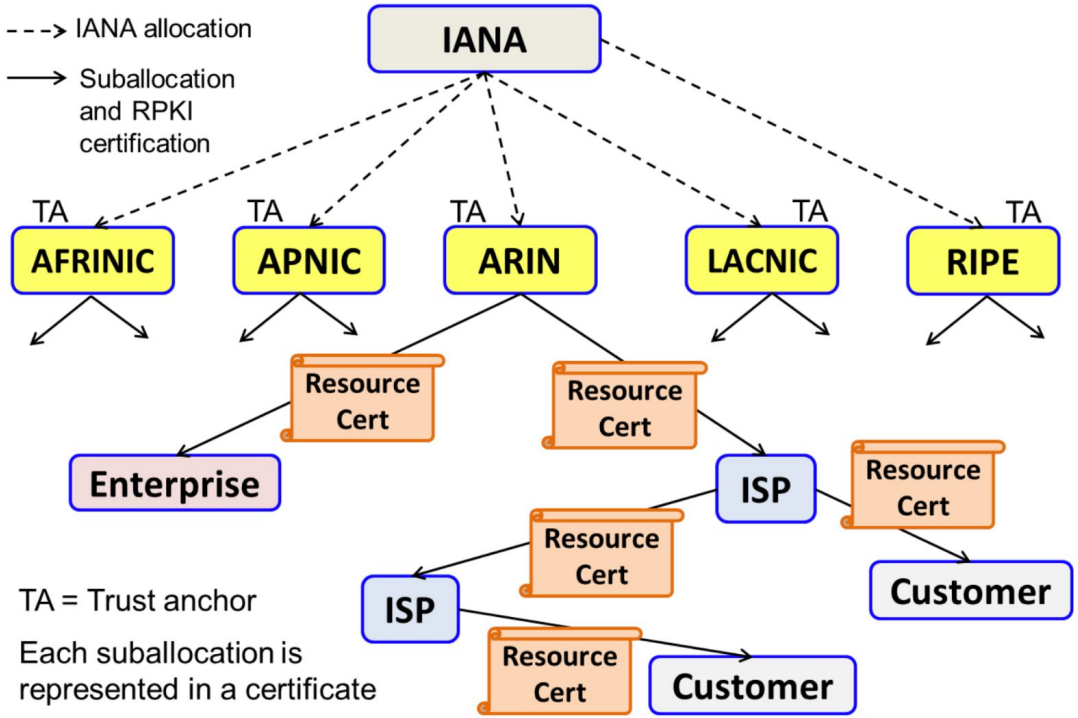
Certification of Resources in Resource Public Key Infrastructure (2)

- ***RPKI*** is a ***global certificate authority*** (CA) and ***registry service*** offered by all regional internet registries (RIRs)
- The RPKI certification chain follows the same allocation hierarchy as for the resource allocation (see the figure in the next slide)
- Each of the five RIRs (AFRINIC, APNIC, ARIN, LACNIC, and RIPE) maintains an independent trust anchor for RPKI certification services in its respective region
  - ***IOW they are 5 independent root CAs***

Resource allocation and certificate chain in RPKI

Note: for the other RIRs the approach is the same (a LIR may be included in the chain)

sono indipendent CA, possono anche delegare qualcuno per firmare risorse di customer.



Certification of Resources in Resource Public Key Infrastructure (3)

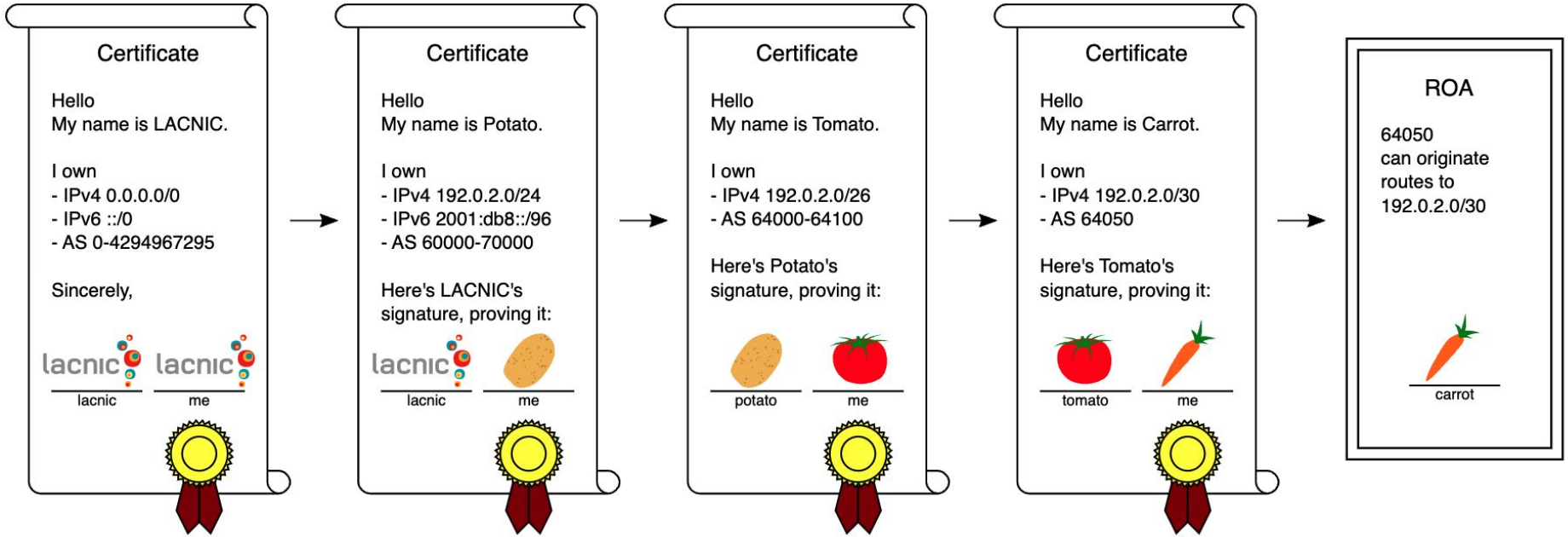
stesso standard di https

- RPKI is based on the X.509 standard with RFC 3779 extensions that describe special certificate profiles for internet number resources (prefixes and AS numbers) [RFC 5280, RFC 6487, RFC 3779].
- The ***RIRs issue resource certificates*** (i.e., certificate authority (CA) certificates) to ISPs and enterprises with registered number resource allocations and assignments.
- There are two models of resource certification
  - ***hosted***: the RIR keeps and manages keys and performs RPKI operations on their servers
  - ***delegated***: a resource holder (an ISP or enterprise) receives a CA certificate from their RIR, hosts their own certificate authority, and performs RPKI operations
    - e.g., signs route origin authorizations, issues subordinate resource certificates to their customers

BGP Origin Validation (BGP-OV) (1)

- Once an address prefix owner obtains ***a CA certificate***, they can generate an ***end-entity (EE) certificate*** and use the private key associated with the EE certificate ***to digitally sign a route origin authorization (ROA)***
- ***ROA functions:***
  - It declares a specific AS as an authorized originator of announcements for the prefix
  - It specifies one or more prefixes (optionally a maxlength per prefix) and a single AS number
    - If a maxlength is specified for a prefix in the ROA, then any more-specific (i.e., longer) prefixes (subsumed under the prefix) with a length not exceeding the maxlength are permitted to be originated from the specified AS
    - In the absence of an explicit maxlength for a prefix, the maxlength is equal to the length of the prefix itself
  - If the resource owner has a resource certificate listing multiple prefixes, they can create one ROA in which some or all those prefixes are listed.
    - Alternatively, they can create one ROA per prefix
  - ROAs can also be created (and signed) by an ISP (transit provider) on behalf of its customer

RPKI certificate chain: high level example

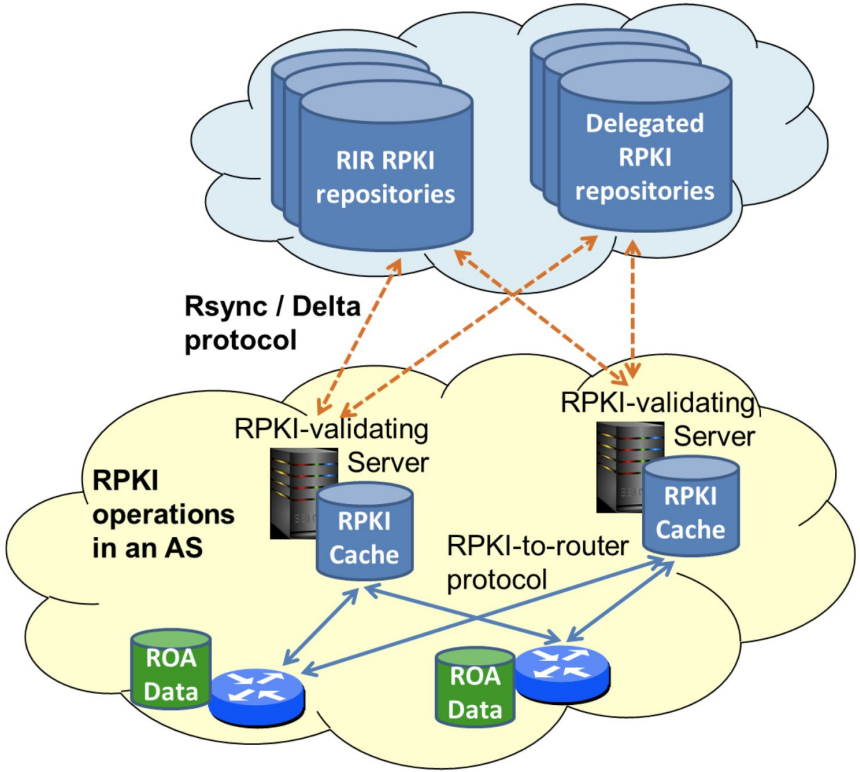


RPKI data retrieval, caching, and propagation to routers

- Once created, RPKI data is used throughout the internet by relying parties (RPs), such as ***RPKI-validating servers***
- RPs can access RPKI data from the ***repositories*** using either the rsync protocol or the ***RPKI Repository Delta Protocol*** (***RRDP***, AKA “delta protocol”) (RFC8182)
- A BGP router typically accesses the required ROA data from one or more RPKI cache servers that are maintained by its AS
- The ***RPKI-to-router (RTR)*** protocol is used for communication between the RPKI cache server and the router (RFC 8210) viene prodotta lista prefissi ed AS.

BGP router non fa la validazione, dal server il BGP router scarica la lista di prefissi e AS numbers da un cache server

RPKI data retrieval, caching, and propagation to routers



nell'AS ci sono alcuni RPKI VS che fanno la validazione dei prefissi che scaricano e nelle cache salvano i dati necessari. Questi dati vengono aggiornati periodicamente.

How routers use ROAs as prevention for hijacking and leaking

nelle data structures mandate a BGP abbiamo:

- A BGP router receives a validated list of ***(prefix, maxlength, origin AS)*** tuples (derived from valid ROAs) from one or more RPKI cache servers
- The router makes use of this list with the ***BGP origin validation (BGP-OV) process*** to determine the validation state of an advertised route (RFC 6811)
  - A route has a ***Valid*** origin if the ***(prefix, origin AS)*** pair in the advertised route can be corroborated with the list
  - A route is considered ***Invalid*** if there is a mismatch with the list (i.e., AS number does not match, or the prefix length exceeds maxlength)
  - A route is deemed ***NotFound*** if the prefix announced is not covered by any prefix in the white list (i.e., there is no ROA that contains a prefix that equals or subsumes the announced prefix)
  - When an AS_SET (unordered AS_PATH sequence) is present in a BGP update, it is not possible to clearly determine the origin AS from the AS_PATH

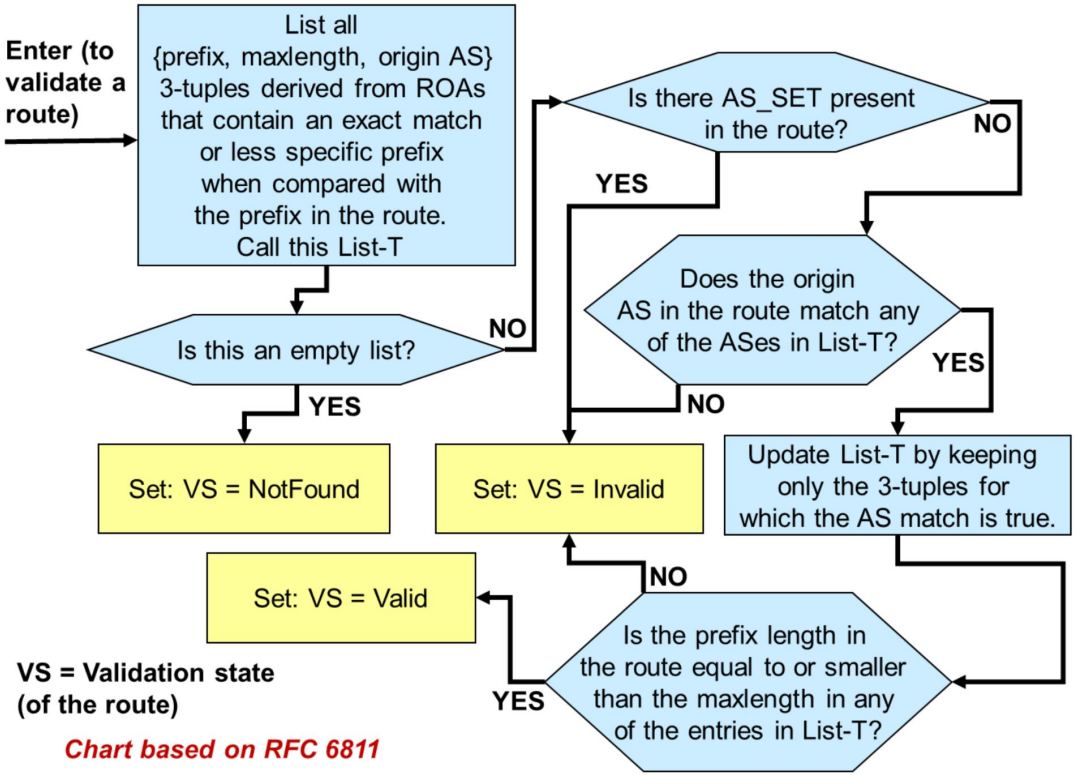
viene fatto il detect in questo modo di un hijack attack

quando viene ricevuto un announcement bisogna vedere le ROA information

nelle RFC abbiamo best practices, in cui suggeriscono che se riceviamo AS_SET, dobbiamo buttarlo.

Algorithm for origin validation (based on RFC 6811)

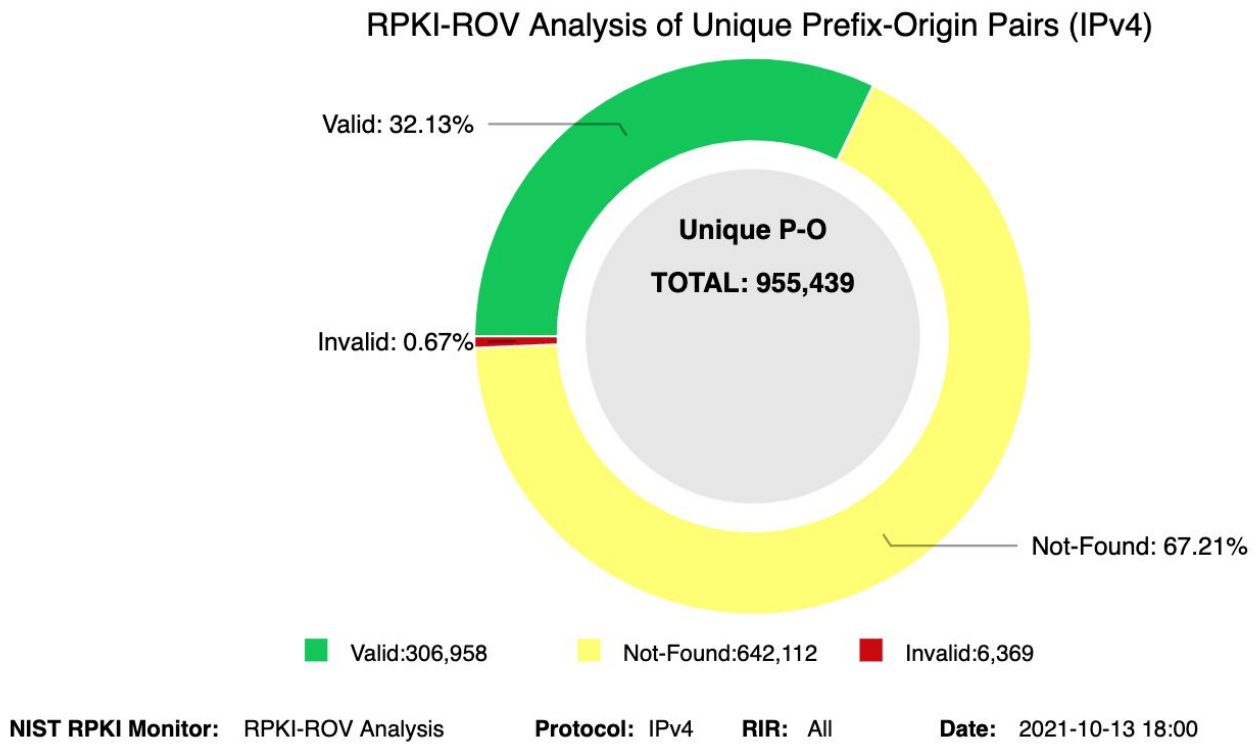
per mitigare hijacking



RPKI BGP OV current adoption status

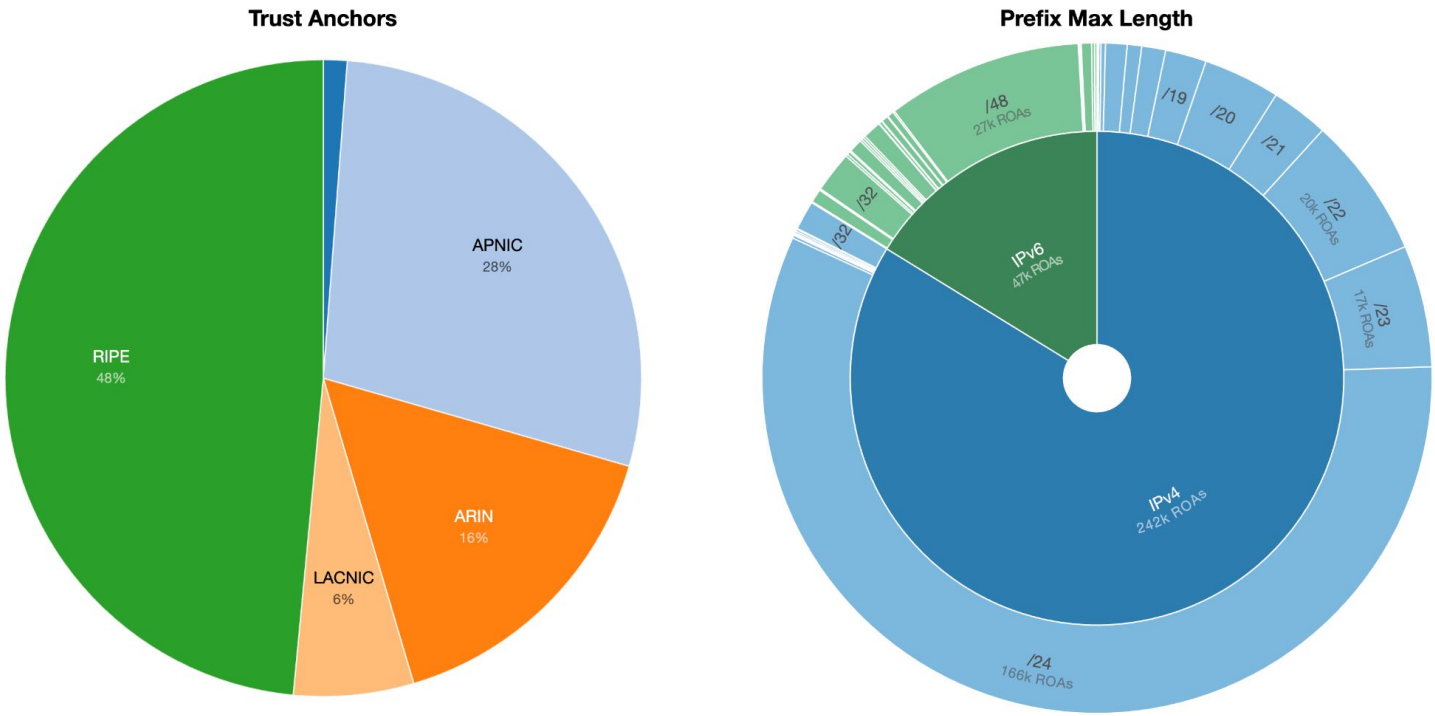
- There are several implementations of RPKI-based BGP OV in both hardware and software-based router platforms [Juniper] [Cisco] [Patel] [Scudder] [NIST-SRx] [Parsons2] [goBGP] [RTRlib].
- Deployment guidance and configuration guidance for many of these implementations are available from several sources
- Although BGP-OV is already implemented in commercial BGP routers, the activation and ubiquitous use of RPKI and BGP-OV in BGP routers require motivation and commitment on the part of network operators.
- Check the following link for a continuous monitoring of the adoption status: https://rpki-monitor.antd.nist.gov/

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



Cloudflare RPKI portal https://rpki.cloudflare.com/

Found **288,698** ROAs in the global RPKI system.

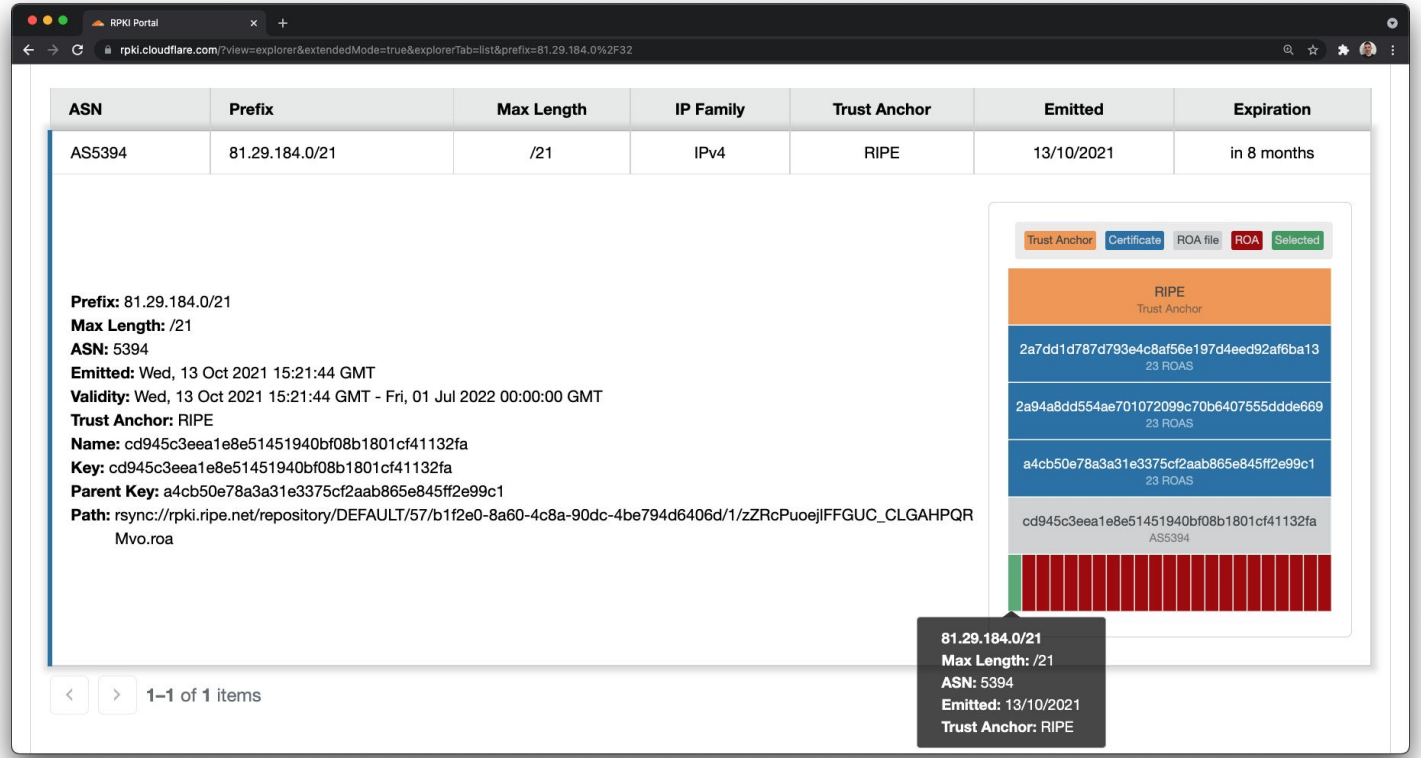


What should the operators do in this partial deployment status?

- In partial/incremental deployment state of the RPKI, the permissible {prefix, origin ASN} pairs for performing BGP-OV should be generated by taking the union of such data obtained from ROAs, IRR data, and customer contracts
- BGP-OV results should be incorporated into local policy decisions to select BGP best path
- How BGP-OV results are used in path selection is strictly a local policy decision for each network operator
- Typical policy choices include:
  - ***Tag-Only*** – BGP-OV results are only used to tag/log data about BGP routes for diagnostic purposes
  - ***Prefer-Valid*** – Use local preference settings to give priority to valid routes. Note that this is only a tie-breaking preference among routes with the exact same prefix
  - ***Drop-Invalid*** – Use local policy to ignore invalid routes in the BGP decision process

A deeper look at a real ROA certification chain

Online ROA validator from https://rpki.cloudflare.com/

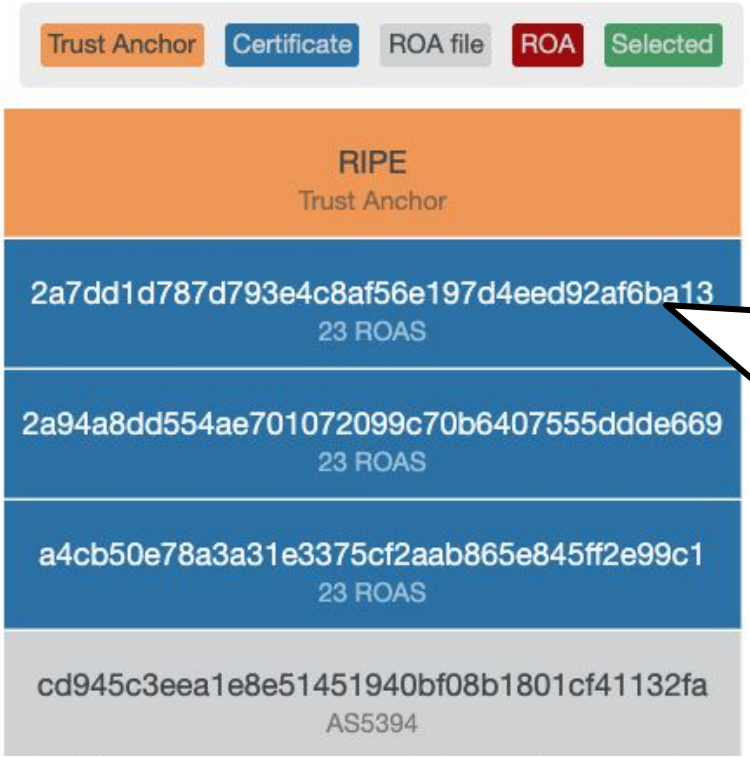


RIPE Trust Anchor

uguale per tutte IRR

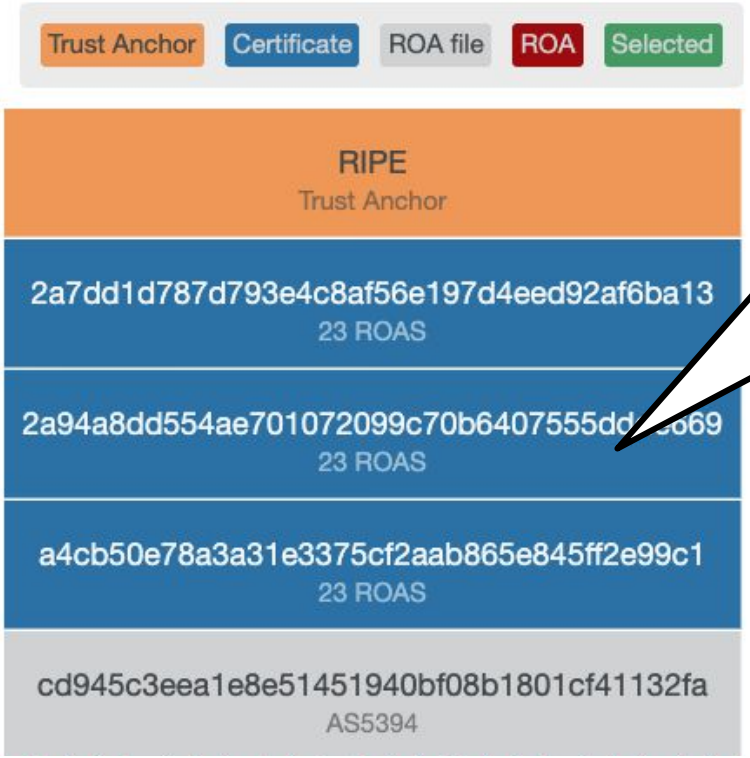
ASNs: 0-4294967295
IPs: 0.0.0.0/0, ::/0
Validity: Tue, 28 Nov 2017 14:39:55 GMT - Sun, 28 Nov 2117 14:39:55 GMT
Trust Anchor: RIPE
Name: ripe-ncc-ta
Key: e8552b1fd6d1a4f7e404c6d8e5680d1ebc163fc3
Parent Key: -
Path: rsync://rpki.ripe.net/ta/ripe-ncc-ta.cer
Subject Information Access (SIA): rsync://rpki.ripe.net/repository/ripe-ncc-ta.mft
https://rrdp.ripe.net/notification.xml
rsync://rpki.ripe.net/repository/

"url" in cui scarichiamo certificati relativi a RIPE



"All Resources" intermediate certificate

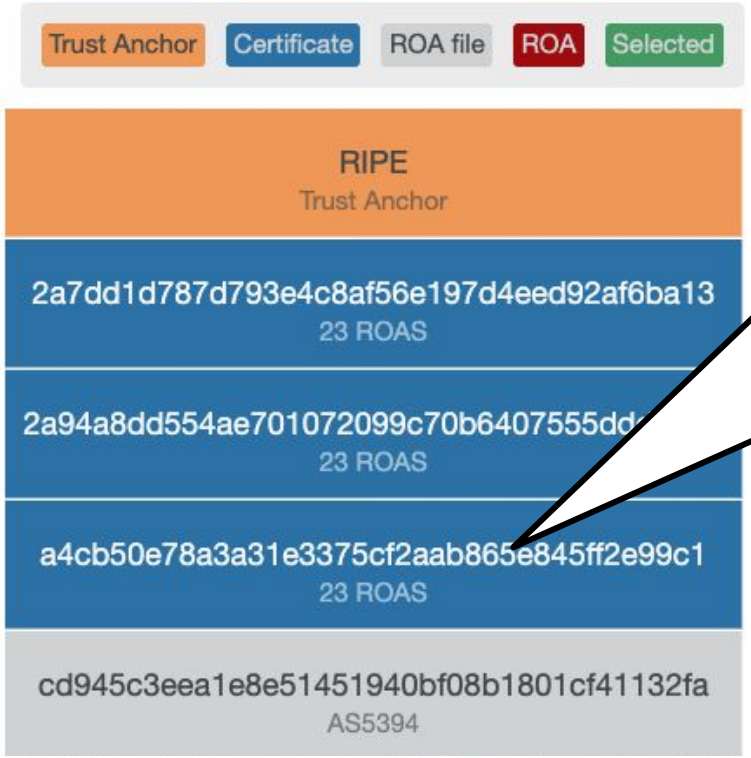
ASNs: 0-4294967295
IPs: 0.0.0.0/0, ::/0
Validity: Tue, 28 Sep 2021 10:31:33 GMT - Fri, 01 Jul 2022 00:00:00 GMT
Trust Anchor: RIPE
Name: 2a7dd1d787d793e4c8af56e197d4eed92af6ba13
Key: 2a7dd1d787d793e4c8af56e197d4eed92af6ba13
Parent Key: e8552b1fd6d1a4f7e404c6d8e5680d1ebc163fc3
Path: rsync://rpki.ripe.net/repository/2a7dd1d787d793e4c8af56e197d4eed92af6ba13.cer
Subject Information Access (SIA): rsync://rpki.ripe.net/repository/aca/
rsync://rpki.ripe.net/repository/aca/Kn3R14fXk-TIr1bhl9Tu2Sr2uhM.mft
https://rrdp.ripe.net/notification.xml



Resource Owning Certificate

ASNs	IPs	Expiration
7	1.178.224.0/19	
28	1.179.112.0/20	
137	2.0.0.0/8	
224	5.0.0.0-5.28.31.255	
248-251	5.28.40.0-5.45.35.255	
261	5.45.40.0-5.254.127.255	in 8 months
286	5.254.160.0-5.255.255.255	
288	13.116.0.0-13.123.255.255	
294	13.140.0.0/14	
375	13.168.0.0-13.183.255.255	
and 240 more...	and 2808 more...	

certificato per la specifica entity AS che possiede la risorsa
(nell'esempio è per UNIDATA ed ECITY NETKA, due compagnie fuse)



End Entity Certificate

ASNs: 5394, 16035
IPs: 77.39.160.0/19, 77.39.224.0/19, 81.29.180.0-81.29.191.255, 185.152.156.0/22, 194.79.192.0/19, 194.183.0.0/19, 195.94.128.0/18, 195.250.224.0/19, 213.233.0.0/18, 217.72.96.0/20, 2a02:688::/32
Validity: Fri, 01 Jan 2021 04:47:35 GMT - Fri, 01 Jul 2022 00:00:00 GMT
Trust Anchor: RIPE
Name: a4cb50e78a3a31e3375cf2aab865e845ff2e99c1
Key: a4cb50e78a3a31e3375cf2aab865e845ff2e99c1
Parent Key: 2a94a8dd554ae701072099c70b6407555ddde669
Path: rsync://rpki.ripe.net/repository/DEFAULT/pMtQ54o6MeM3XPKquGXoRf8umcE.cer
Subject Information Access (SIA): rsync://rpki.ripe.net/repository/DEFAULT/57/b1f2e0-8a60-4c8a-90dc-4be794d6406d/1/
rsync://rpki.ripe.net/repository/DEFAULT/57/b1f2e0-8a60-4c8a-90dc-4be794d6406d/1/pMtQ54o6MeM3XPKquGXoRf8umcE.mft
https://rrdp.ripe.net/notification.xml

ASN 5349: UNIDATA Unidata
ASN 16035: ECITY NETKA

Why the same EE? << In December 2002, the extraordinary shareholders' meeting of eCity S.r.l. resolved to transform the company into a joint stock company and changed its name to "Unidata S.p.A." >>

ROA file containing 23 ROAS

ASN	Prefix	Max Length	IP Family	Trust Anchor	Emitted	Expiration
AS5394	81.29.184.0/21	/21	IPv4	RIPE	13/10/2021	in 8 months
AS5394	81.29.180.0/22	/22	IPv4	RIPE	13/10/2021	in 8 months
AS5394	195.94.152.0/24	/24	IPv4	RIPE	13/10/2021	in 8 months
AS5394	194.183.0.0/19	/19	IPv4	RIPE	13/10/2021	in 8 months
AS5394	77.39.224.0/20	/20	IPv4	RIPE	13/10/2021	in 8 months
AS5394	77.39.224.0/19	/19	IPv4	RIPE	13/10/2021	in 8 months

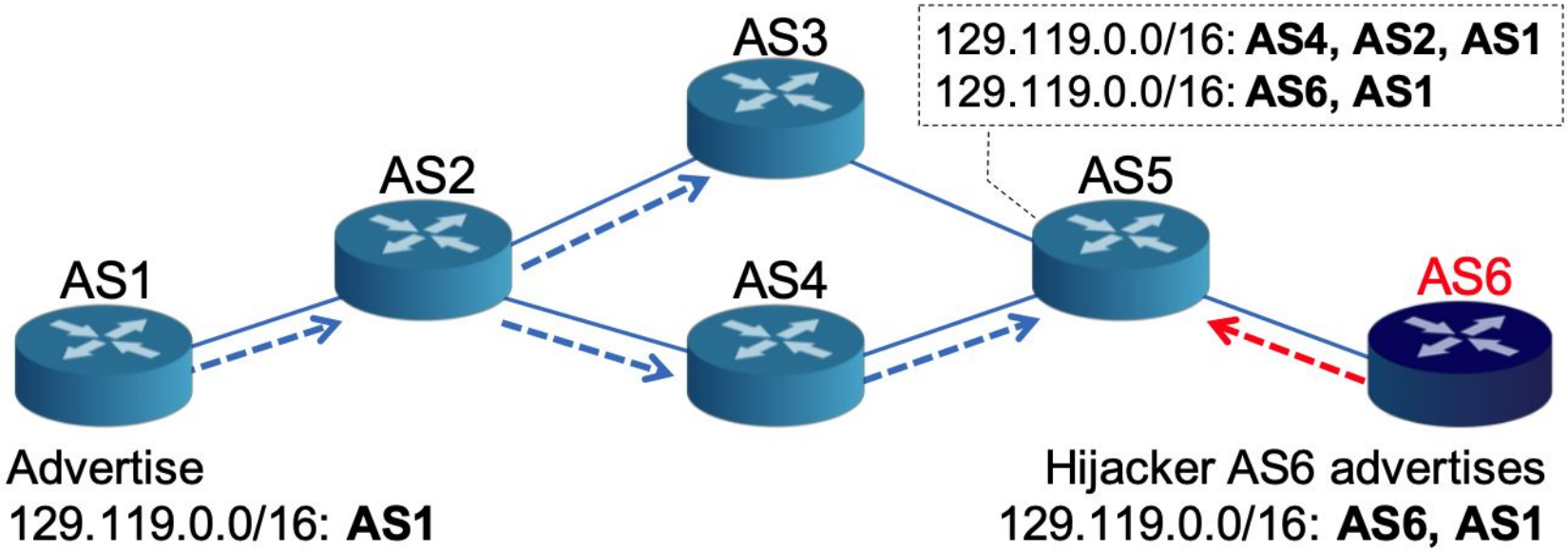
...

Forged-Origin Hijacks and BGP path validation

Forged-Origin Hijacks

- With ROA-based origin validation an AS can not announce a prefix unless it has a valid ROA
- However, a purposeful malicious hijacker can forge the origin AS of any update by prepending the number of an AS found in a ROA for the target prefix onto their own unauthorized BGP announcement
- ***BGP origin validation is necessary but, by itself, is insufficient for fully securing the prefix and AS path in BGP announcements!*** Il problema nasce quindi sul PATH-AS
- For greater impact, in conjunction with forging the origin, the attacker may replace the prefix in the route with a more-specific prefix (subsumed under the announced prefix) that has a length not exceeding the maxlength in the ROA

Forged-Origin Hijacks



BGP origin validation non fa check path, solo corrispondenza AS number - network.
Sembra un update (c'è virtual link tra AS6 e AS1) ma è un origination di AS6.

BGP Path Validation

attualmente non usato dagli operatori

- BGP-PV is a mechanisms for protecting BGP announcement against prefix modifications and forged-origin attacks
- It is specified in the ***BGPSec*** protocol (RFC 8025)

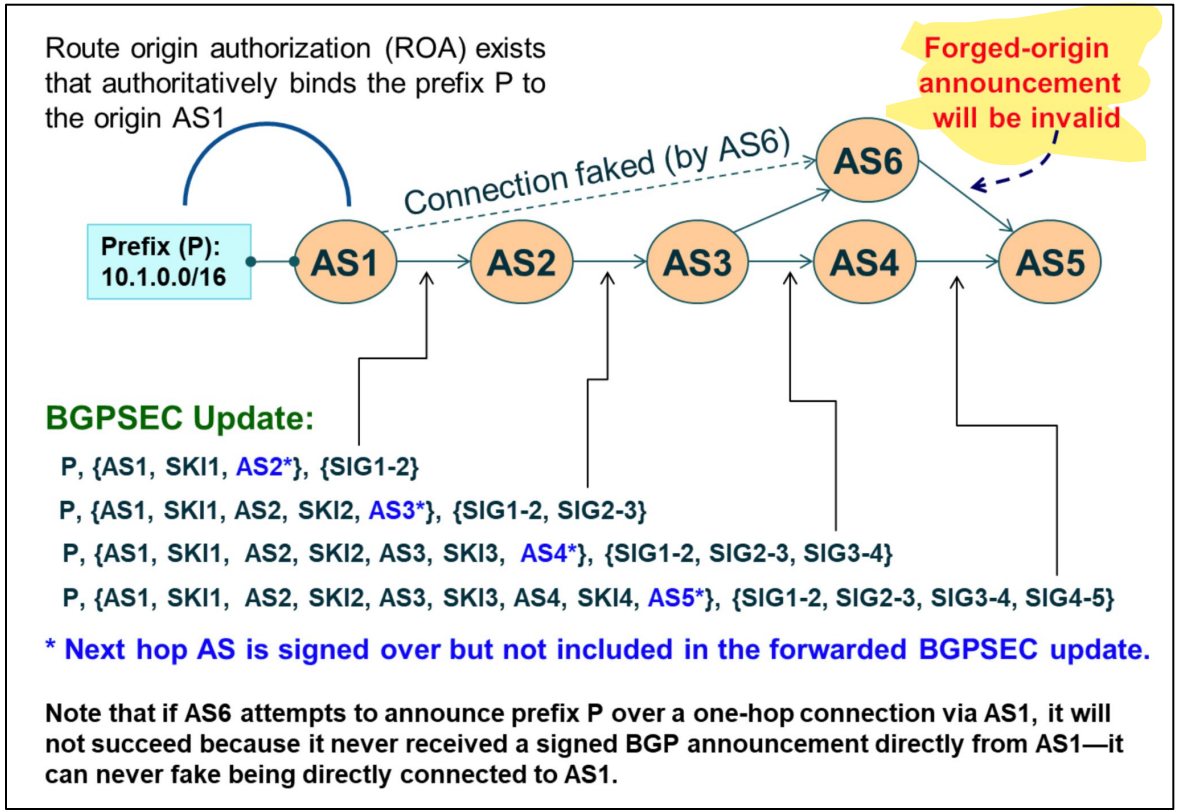
prof passa alla slide dopo con figura)

  - each AS implementing BGP-PV has an RPKI resource certificate for their ASN
  - each router in the path has certificate and a private key to sign the BGP updates
  - the certificates for all BGP-PV routers are retrieved by all participating ASes
  - the public keys of all BGP-PV routers are expected to be available at each BGP-PV router
  - each AS uses the private key to sign the BGP update and includes in the data to be signed the next AS supposed to receive the update
  - the update includes the ***subject key identifier*** (SKI) for the public key of each AS in the path
  - each AS will receive multiple signatures to be verified
  - if all signatures verify correctly and the origin validation check also passes, the BGP update is valid

AS6 vuole eseguire forged hijhack su AS1.
L'idea è di legare crittograficamente l'announce con una catena di segnature, in cui ciascuno applica la propria segnatura.

BGP Path Validation

ovviamente c'è
private e public key,
meccanismi di validazione..
dentro l'annuncio abbiamo:
- AS1
- SKI1 (public key)
- AS2* (next hop)
sappiamo ovviamente a chi inviarlo!
E' stato segnato ma non incluso nella signature.
- SIG1-2 è la signature fatta con private key.
con cui applica la signature.
Fatto da tutti gli AS nel path



Prefix Filtering

Prefix Filtering: intro

Gli operatori che hanno accettato il fake route di Twitter non hanno applicato Prefix Filtering, ovvero una lista di prefissi attesi.
Se ricevo twitter prefix da AS ignoto, non dovrei accettarlo

- ***BGP prefix filtering*** (also known as route filtering) is the most basic mechanism for protecting BGP routers from accidental or malicious disruption
- Prefixes expected in a peering (e.g., customer) relationship are accepted, and prefixes not expected, including bogons and unallocated, are rejected
- ***Inbound*** and ***outbound*** prefix filtering should be both implemented
- Route filters are typically specified using a syntax similar to that used for access control lists (we'll see this in a lab at the end of this lecture)
- Types of prefix filters: ***(i) Unallocated Prefixes***; ***(ii) Special Purpose Prefixes***; ***(iii) Prefixes Owned by an AS***; ***(iv) Prefixes that Exceed a Specificity Limit***; ***(v) Default Route***; ***(vi) IXP LAN Prefixe***s

(es: rifiuto tutto ciò che è \20)

Simple example (found on google image)

```
router bgp 100
 network 105.7.0.0 mask 255.255.0.0
 neighbor 102.10.1.1 remote-as 110
 neighbor 102.10.1.1 prefix-list AS110-IN in
 neighbor 102.10.1.1 prefix-list AS110-OUT out
!
ip prefix-list AS110-IN deny 218.10.0.0/16
ip prefix-list AS110-IN permit 0.0.0.0/0 le 32
ip prefix-list AS110-OUT permit 105.7.0.0/16
ip prefix-list AS110-OUT deny 0.0.0.0/0 le 32
```

qui mettiamo route che non partecipano in RPKI, poichè altrimenti sono espressi lì dentro!

1. Unallocated Prefixes and Special Purpose Prefixes

- The Internet Assigned Numbers Authority (IANA) allocates address space to RIRs.
- All the IPv4 address space (or prefixes), except for some reserved for future use, have been allocated by IANA
- The IPv6 address space is much larger than that of IPv4, and, understandably, the bulk of it is unallocated.
- It is a good practice to accept only those IPv6 prefix advertisements that have been allocated by the IANA è facile vedere che, senza prefix filter di questo tipo, accettare un numero elevato di indirizzi!
- Network operators should ensure that the IPv6 prefix filters are updated regularly
  - In the absence of such regular updating processes, it is better not to configure filters based on allocated prefixes
  - If prefix resource owners regularly register AS0 ROAs for allocated (but possibly currently unused) prefixes, then those ROAs could be a complementary source for the update of prefix filters
- Moreover, IANA maintains registries for special-purpose IPv4 and IPv6 addresses, these should be filtered

1. Unallocated Prefixes and Special Purpose Prefixes

Security Recommendations

1. IPv6 routes should be filtered to permit only allocated IPv6 prefixes. Network operators should update IPv6 prefix filters regularly to include any newly allocated prefixes.

2. Prefixes that are marked "False" in column "Global" [IANA-v4-sp] [IANA-v6-sp] are forbidden from routing in the global internet and should be rejected if received from an external BGP (eBGP) peer.

- [IANA-v4-sp] https://www.iana.org/assignments/iana-ipv4-special-registry/iana-ipv4-special-registry.xhtml
  [IANA-v6-sp] https://www.iana.org/assignments/iana-ipv6-special-registry/iana-ipv6-special-registry.xhtml

2. Prefixes Owned by an AS

- An AS may originate one or multiple prefixes
- In the inbound direction, the AS should (in most cases) reject routes for the prefixes (subnets) it originates if received from any of its eBGP peers (transit provider, customer, or lateral peer).
  - In general, the data traffic destined for these prefixes should stay local and should not be leaked over external peering.
- However, if the AS operator is uncertain whether a prefix they originate is single-homed or multi-homed, then the AS should accept the prefix advertisement from an eBGP peer (and assign a lower local preference value) so that the desired redundancy is maintained.

2. Prefixes Owned by an AS

- An AS may originate one or multiple prefixes
- In the inbound direction, the AS should (in most cases) reject routes for the prefixes (subnets) it originates if received from any of its eBGP peers (transit so that the desired redundancy is maintained.

(se non abbiamo network in posti geografici diversi)

Security Recommendation: for single-homed prefixes (subnets) that are owned and originated by an AS, any routes for those prefixes received at that AS from eBGP peers should be rejected.

3. Prefixes that Exceed a Specificity Limit

- Normally, ISPs neither announce nor accept routes for prefixes that are more specific than a certain level of specificity.
  - For example, maximum acceptable prefix lengths are mentioned in existing practices as /24 for IPv4 and /48 for IPv6.
  - The level of specificity that is acceptable is decided by each AS operator and communicated with peers.
  - In instances when Flowspec [RFC5575] [RFC5575bis] is used between adjacent ASes for DDoS mitigation, the two ASes may mutually agree to accept longer prefix lengths (e.g., a /32 for IPv4) but only for certain pre-agreed prefixes.
  - That is, the announced more-specific prefix must be contained within a pre-agreed prefix
- Some operators may choose to reject prefix announcements that are less-specific than /8 and /11 for IPv4 and IPv6, respectively

3. Prefixes that Exceed a Specificity Limit

- Normally, ISPs neither announce nor accept routes for prefixes that are more specific than a certain level of specificity.
  - For example, maximum acceptable prefix lengths are mentioned in existing practices as /24 for IPv4 and /48 for IPv6

Security Recommendation: it is recommended that an eBGP router should set the specificity limit for each eBGP peer and reject prefixes that exceed the specificity limit on a per-peer basis

- Some operators may choose to reject prefix announcements that are less-specific than /8 and /11 for IPv4 and IPv6, respectively

4. Default Route

- A route for the prefix 0.0.0.0/0 is known as the default route in IPv4, and a route for ::/0 is known as the default route in IPv6.
- The default route is advertised or accepted only in specific customer-provider peering relations.
  - For example, a transit provider and a customer that is a stub or leaf network may make this arrangement between them whereby the customer accepts the default route from the provider instead of the full routing table.
  - In general, filtering the default route is recommended except in situations where a special peering agreement exists.

4. Default Route

- A route for the prefix 0.0.0.0/0 is known as the default route in IPv4, and a route for ::/0 is known as the default route in IPv6.
- The default route is advertised or accepted only in specific customer-provider peering relations.

Security Recommendation: the default route (0.0.0.0/0 in IPv4 and ::/0 in IPv6) should be rejected except when a special peering agreement exists that permits accepting it.

IXP LAN Prefixes

- Typically, there is a need for the clients at an internet exchange point (IXP) to have knowledge of the IP prefix used for the IXP LAN which facilitates peering between the clients.

  - See [RFC7454] for more details on this topic.

IXP LAN Prefixes

- Typically, there is a need for the clients at an internet exchange point (IXP) to have knowledge of the IP prefix used for the IXP LAN which facilitates peering between the clients.

Security Recommendation: An internet exchange point (IXP) should announce — from its route server to all of its member ASes — its LAN prefix or its entire prefix, which would be the same as or less specific than its LAN prefix. Each IXP member AS should, in turn, accept this prefix and reject any more-specific prefixes (of the IXP announced prefix) from any of its eBGP peers.

non accettiamo, tranne original announce, altri specific prefix rispetto quello ricevuto da IXP da altri eBGP

customer diverso da TIER-1

Prefix Filtering for Peers of Different Types

- The inbound and outbound prefix filtering recommendations vary based on the type of peering relationship that exists between networks: lateral peer, transit provider, customer, or leaf customer
- A number of publicly available documents (including [1]) give a list of detailed recommendations for each type of peer relationship
- For example, prefix filtering performed in a Leaf Customer Network:
  - A leaf customer may request only the default route from its transit provider. In this case, only the default route should be accepted and nothing else.
  - If the leaf customer requires the full routing table from the transit provider, then it should apply the following inbound prefix filters: Unallocated prefixes, Special-purpose prefixes, Prefixes that the AS (i.e., leaf customer) originates, Prefixes that exceed a specificity limit, Default route

Role of RPKI in Prefix Filtering

La validità o meno di RPKI è crittograficamente sicura

- An ISP can retrieve (from RPKI registries) all available route origin authorizations (ROAs) corresponding to autonomous systems (ASes) that are known to belong in their customer cone.
- From the available ROAs, it is possible to determine the prefixes that can be originated from the ASes in the customer cone.
- As the RPKI registries become mature with increasing adoption, the prefix lists derived from ROAs will become useful for prefix filtering.
- Even in the early stages of RPKI adoption, the prefix lists (from ROAs) can help cross-check and/or augment the prefix filter lists that an ISP constructs by other means.

Role of RPKI in Prefix Filtering

- An ISP can retrieve (from RPKI registries) all available route origin authorizations (ROAs) corresponding to autonomous systems (ASes) that are known to belong in their customer cone.

Security Recommendation: the ROA data (available from RPKI registries) should be used to construct and/or augment prefix filter lists for customer interfaces

- Even in the early stages of RPKI adoption, the prefix lists (from ROAs) can help cross-check and/or augment the prefix filter lists that an ISP constructs by other means.

Route Leak Solution

Intra-AS route leaking mitigation

ignoriamo AS per ciò che entra ed esce, ad esempio con prefix non riannunciamo routes

- Many operators currently use an intra-AS solution, which is done by tagging BGP updates from **ingress to egress** (within the AS) using a BGP large community
- The BGP large community does not propagate in eBGP
- Each BGP update is **tagged** on ingress to indicate that it was received in eBGP from a customer, lateral peer, or transit provider
- Further, a route that originated within the AS is tagged to indicate the same
- At the egress point, the sending router applies an egress policy that makes use of the tagging
  - Routes that are received from a customer are **allowed on the egress** to be forwarded to any type of peer (e.g., customer, lateral peer, or transit provider).
  - Routes received from a lateral peer or transit provider are **forwarded only to customers** (i.e., they are not allowed to be forwarded to a lateral peer or transit provider).
- Further reading https://tools.ietf.org/id/draft-ietf-grow-route-leak-detection-mitigation-04.html

Inter-AS route leaking mitigation

- The second type of inter-AS solution is intended to work in eBGP across AS hops.
- With the inter-AS solution, the focus shifts to detection and mitigation in case a route leak has already occurred and started to propagate.
- If a leak indeed propagates out of an AS, then the peer AS or any AS along the subsequent AS path should be able to detect and stop it.
- For robustness of the internet routing infrastructure, inter-AS route leak detection and mitigation capabilities will also need to be implemented in addition to the intra-AS prevention capability

vediamo un leak quando questi è già stato propagato.
Ci sono tag basati su prefix filtering. Quando un route leak esce, un AS lo identifica e lo stoppa.

A better (crypto) solution to Route Leaks: ASPA

- The mitigations discussed before are *best practices* for operators
  - We can make use of the RPKI to detect and mitigate Route Leaks
- **ASPA**: *Autonomous System Provider Authorization*
  - Basic idea: use the RPKI to certify the **relationships** between ASes
  - For each AS, this means having a list of ***authorized*** providers
    - cryptographically verifiable!
  - When a BGP peer receives an announcement, it can verify that the AS_PATH couples of ASNs have the *"right"* relationship
    - In this way, we have a solution both for Route Leaks and for BGP Forged Origin hijacks!

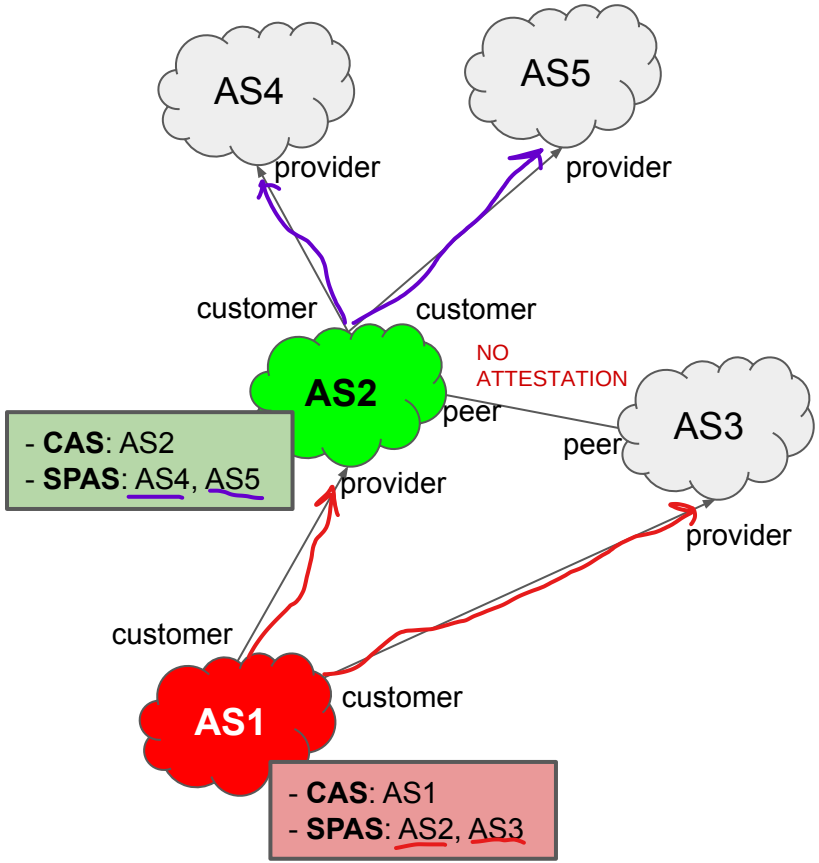
i providers sono AS che, per la loro compagnia, fornisco servizi di transit.
Ogni AS è un customer, e c'è lista di providers. Quando BGP router con ASPA attiva, riceve announcement con AS path, per ogni AS nel path noi sapremo le relazioni tra loro, e possiamo parlare di "valley" (che vediamo nella slide DOWNSTREAM PEERS)

ASPA resource content

- The content of an ASPA has the following information:
  - Identification of the Customer AS (**CAS**) → customerASID field
    - it is called customer, but could also be a provider
  - A Set of *(authorized)* Providers AS (**SPAS**) → providers field
- It is encoded in an RPKI signed object

(Io sono tutti)

https://datatracker.ietf.org/doc/draft-ietf-sidrops-aspa-profile/

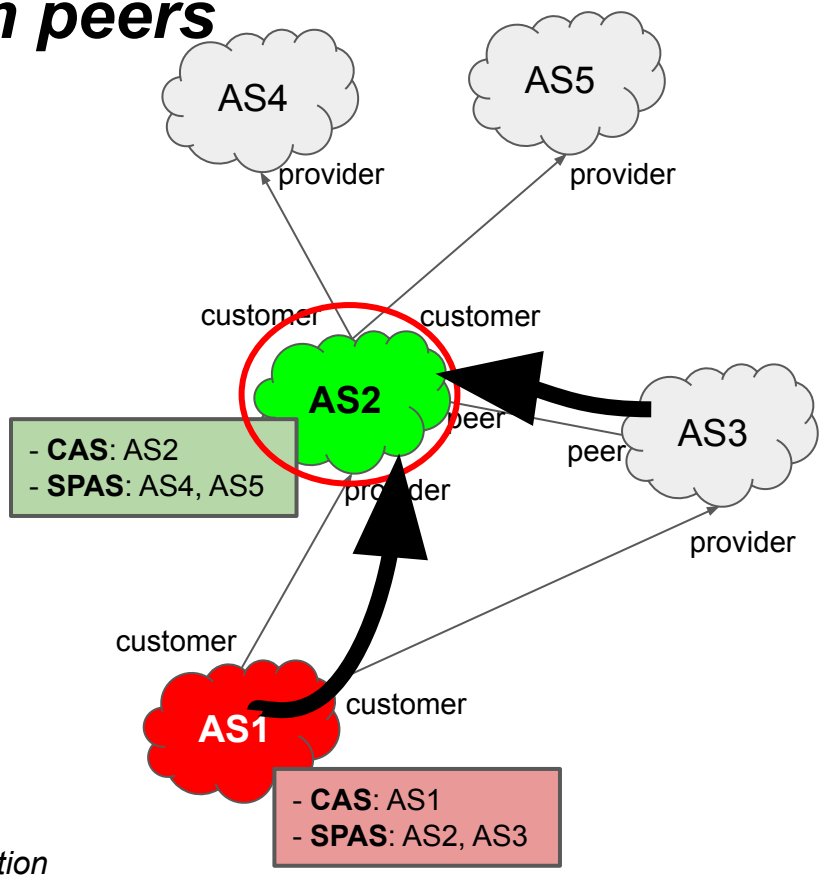


ASPA verification: upstream peers

Basic principle: let { AS[N], AS[N-1], … , AS[2], AS[1] } be the AS_PATH sequence

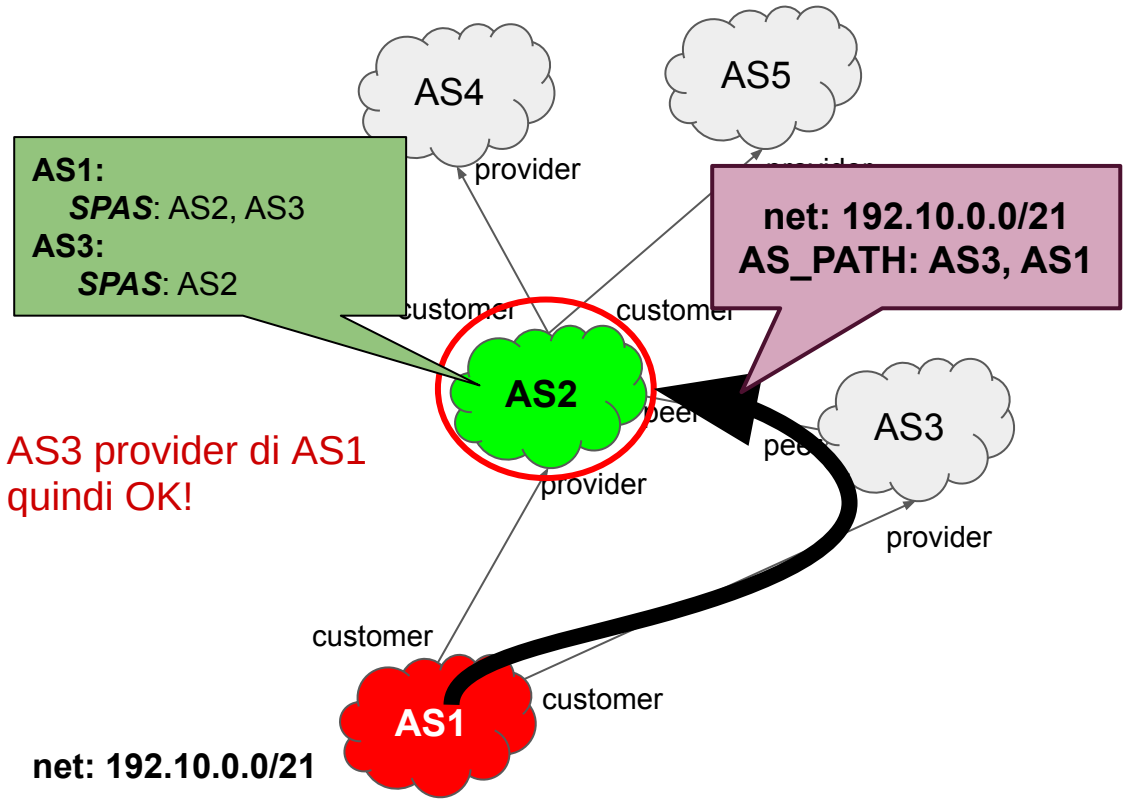
1. If N = 1 → **valid**
2. For N >= 2, let *i* be 2 <= i <= N:
   a. if *hop*(AS[i-1], AS[i]) outcome is "Not Provider" → **invalid**
   b. if *hop*(AS[i-1], AS[i]) outcome is "No Attestation" → **unknown**
   c. if *hop*(AS[i-1], AS[i]) outcome is "Provider" → **valid**

true se AS[i-1] è provider di AS[i], false altrimenti

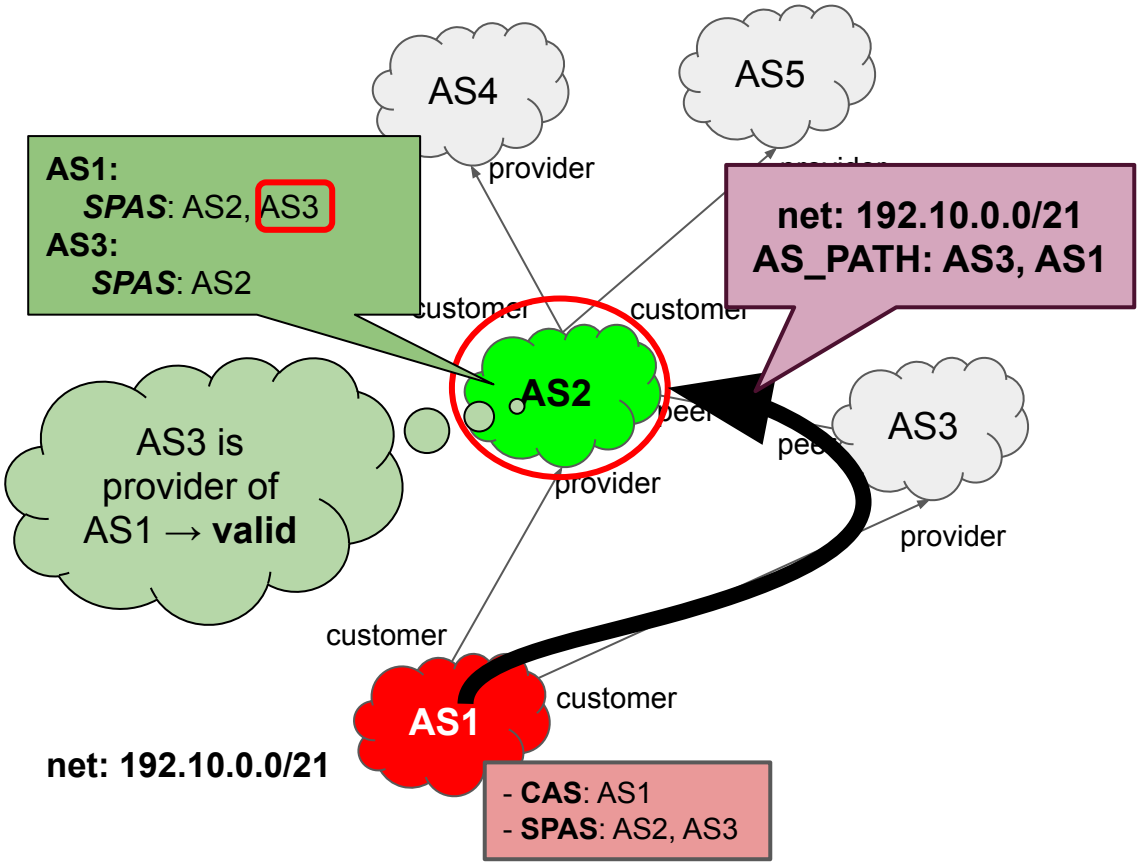


https://datatracker.ietf.org/doc/html/draft-ietf-sidrops-aspa-verification

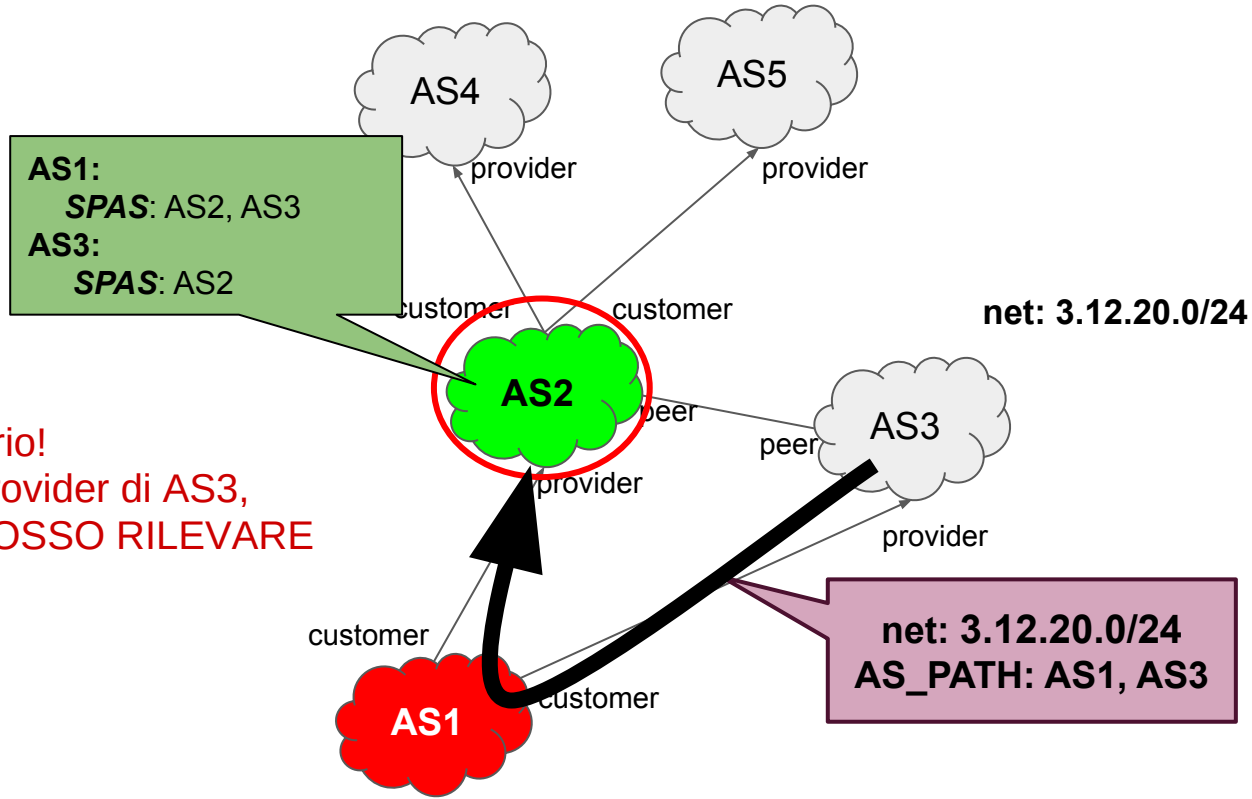
ASPA verification: upstream peers



ASPA verification: upstream peers

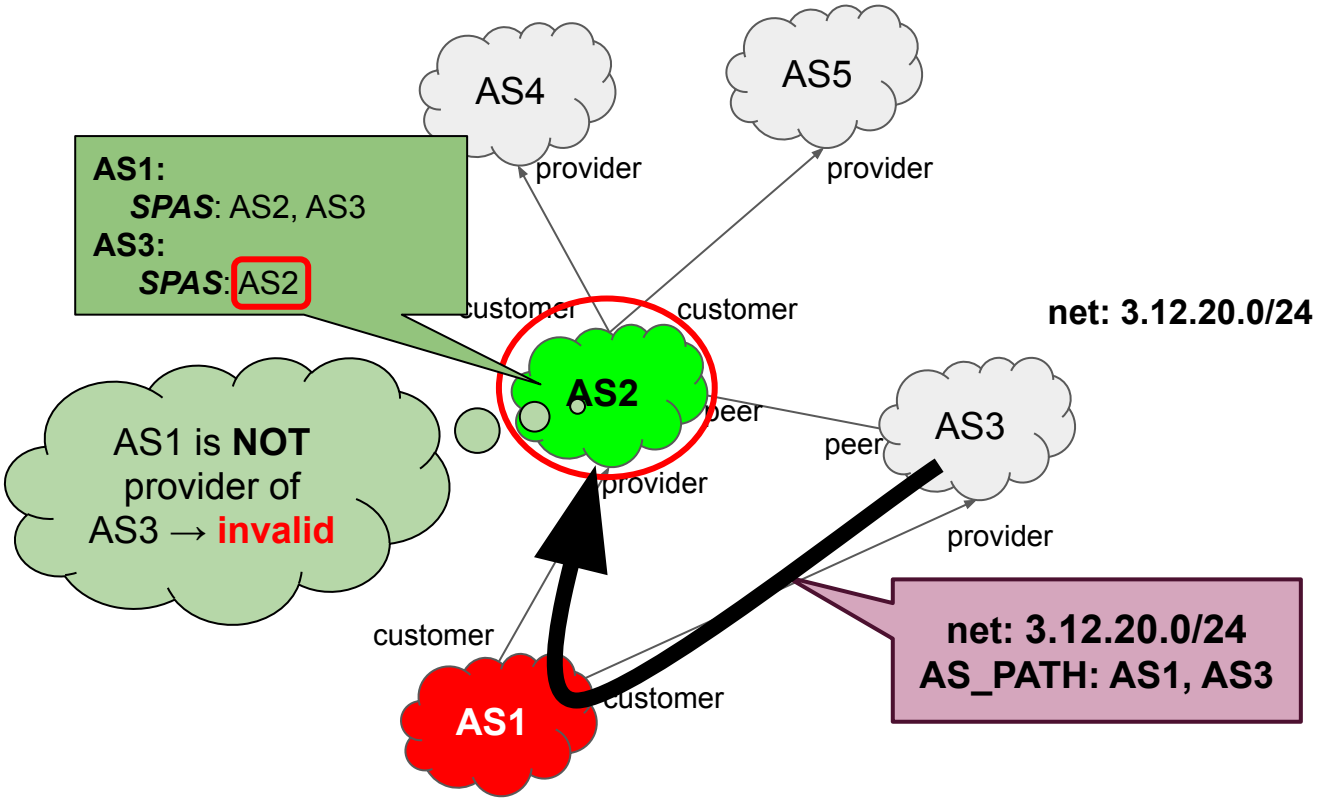


ASPA verification: upstream peers



Qui è il contrario!
AS1 NON è provider di AS3, quindi NON POSSO RILEVARE LEAK

ASPA verification: upstream peers

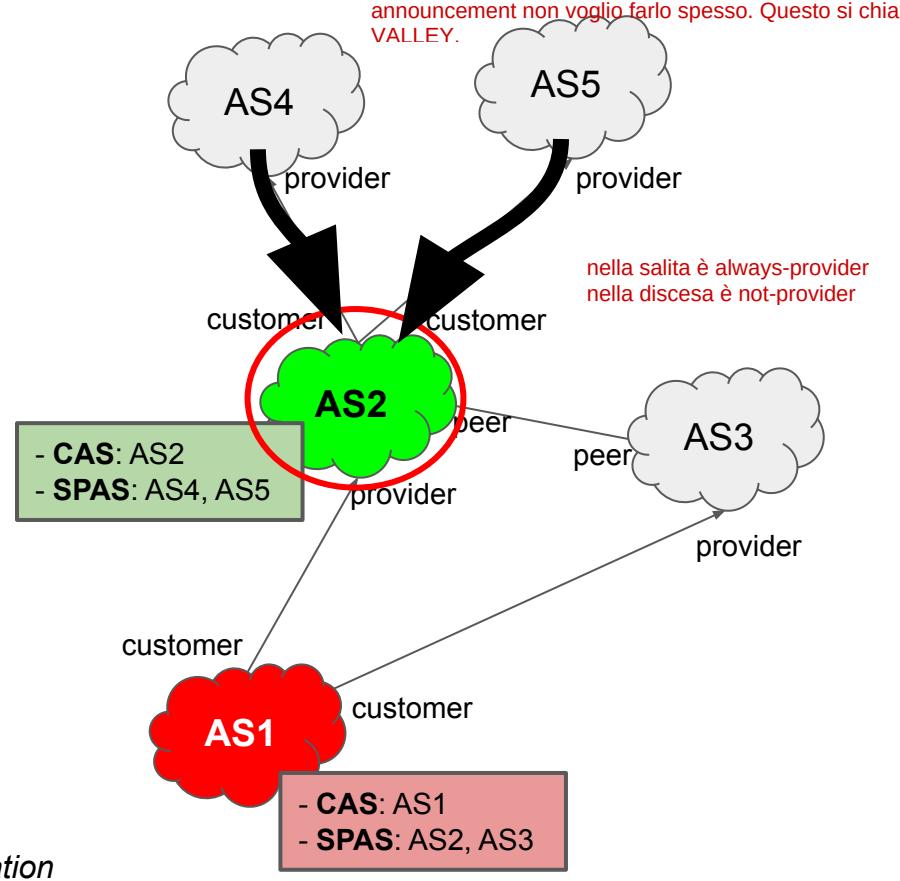


ASPA verification: downstream peers

immaginiamo di avere 2 customer, ne arriva un terzo che esporta un address. Ci sono i vari livelli fino a TIER-1, che propaga annuncio ad altri TIER, per poi "riscendere" di gerarchia. Quindi SALE e SCENDE, ma nello stesso BGP announcement non voglio farlo spesso. Questo si chiama VALLEY.

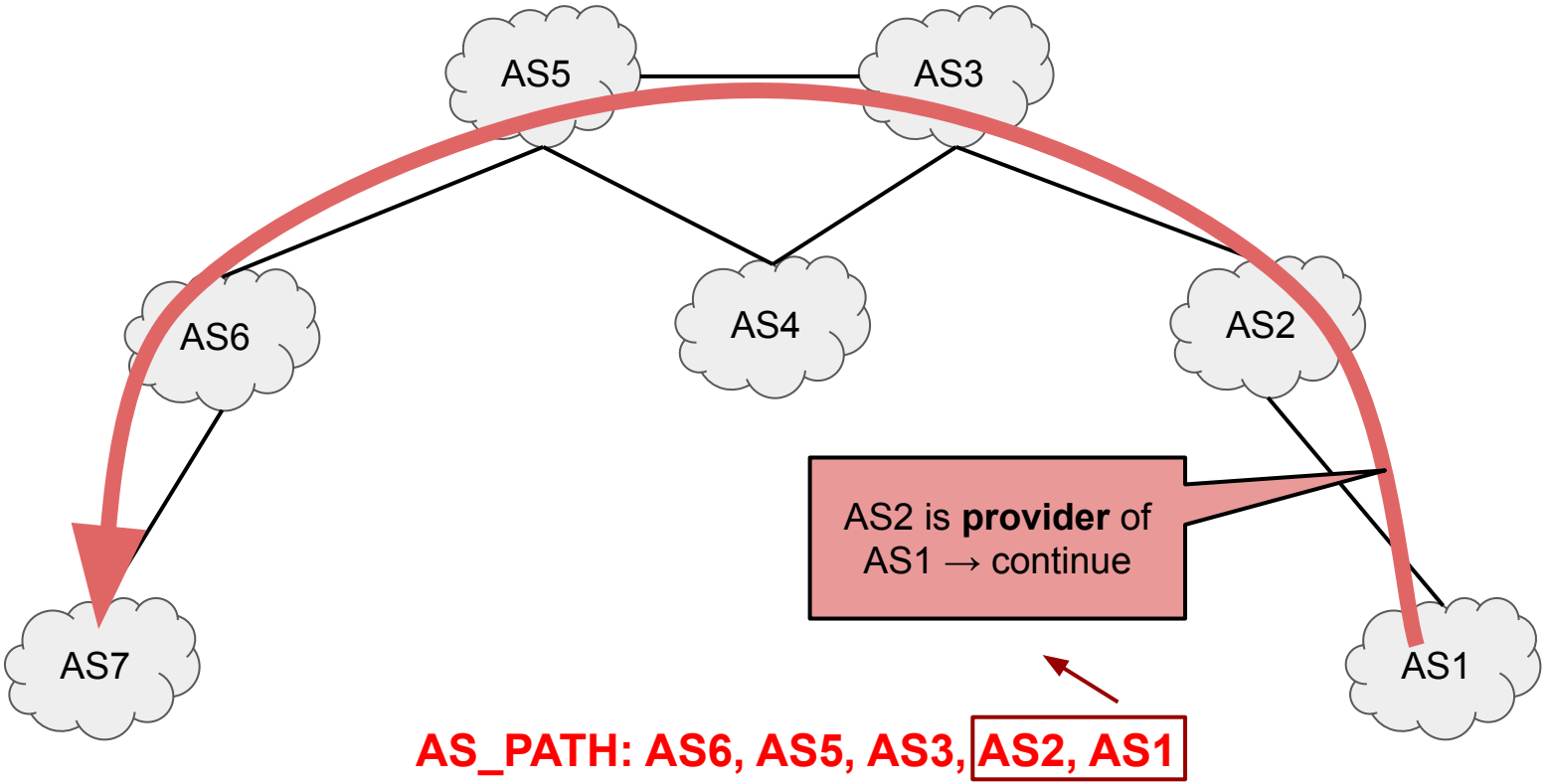
Basic principle: let { AS[N], AS[N-1], … , AS[2], AS[1] } be the AS_PATH sequence, with **N>=3**

1. let u, v with u <= v
   a. if (AS[u-1], AS[u]) outcome is “Not Provider” **and** (AS[v+1], AS[v]) outcome is “Not Provider” → **invalid**
2. Up-ramp: determine the highest K such that (AS[K-1], AS[K]) = “Provider”
3. Down-ramp: determine the highest L such that (AS[L+1], AS[L]) = “Provider”
4. If L-K <= 1 → **valid** else **unknown**

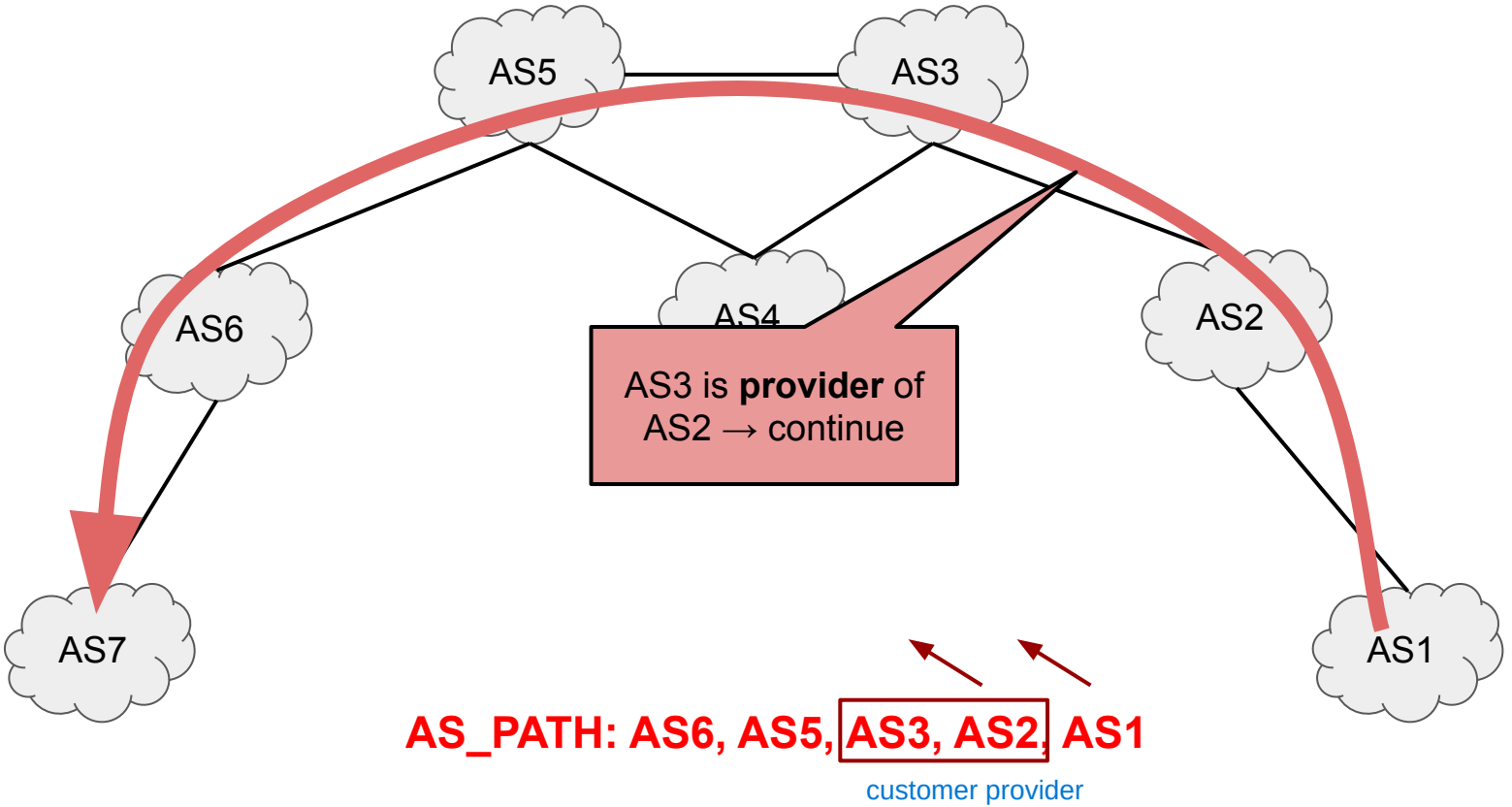


https://datatracker.ietf.org/doc/html/draft-ietf-sidrops-aspa-verification

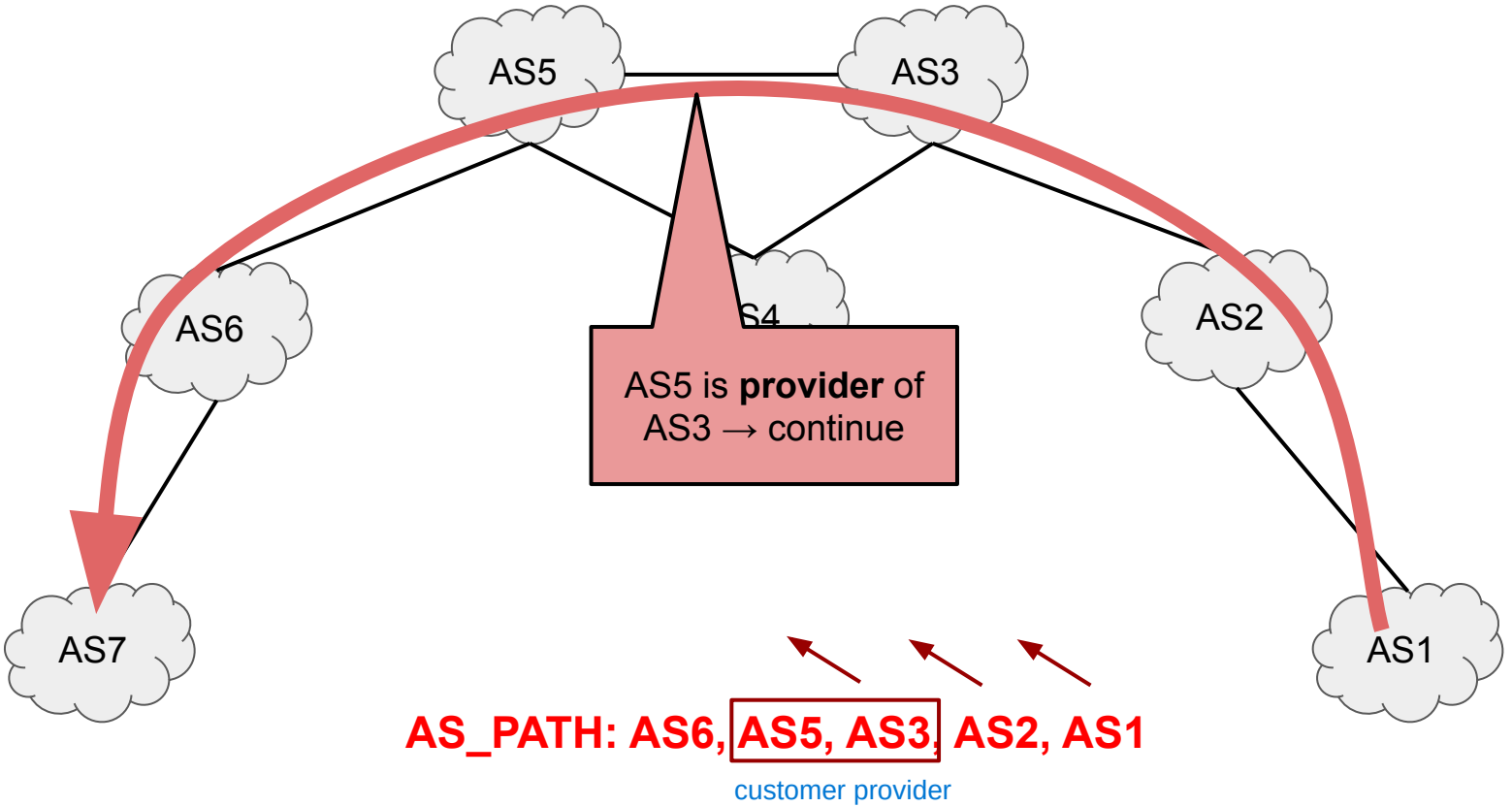
ASPA verification: downstream peers



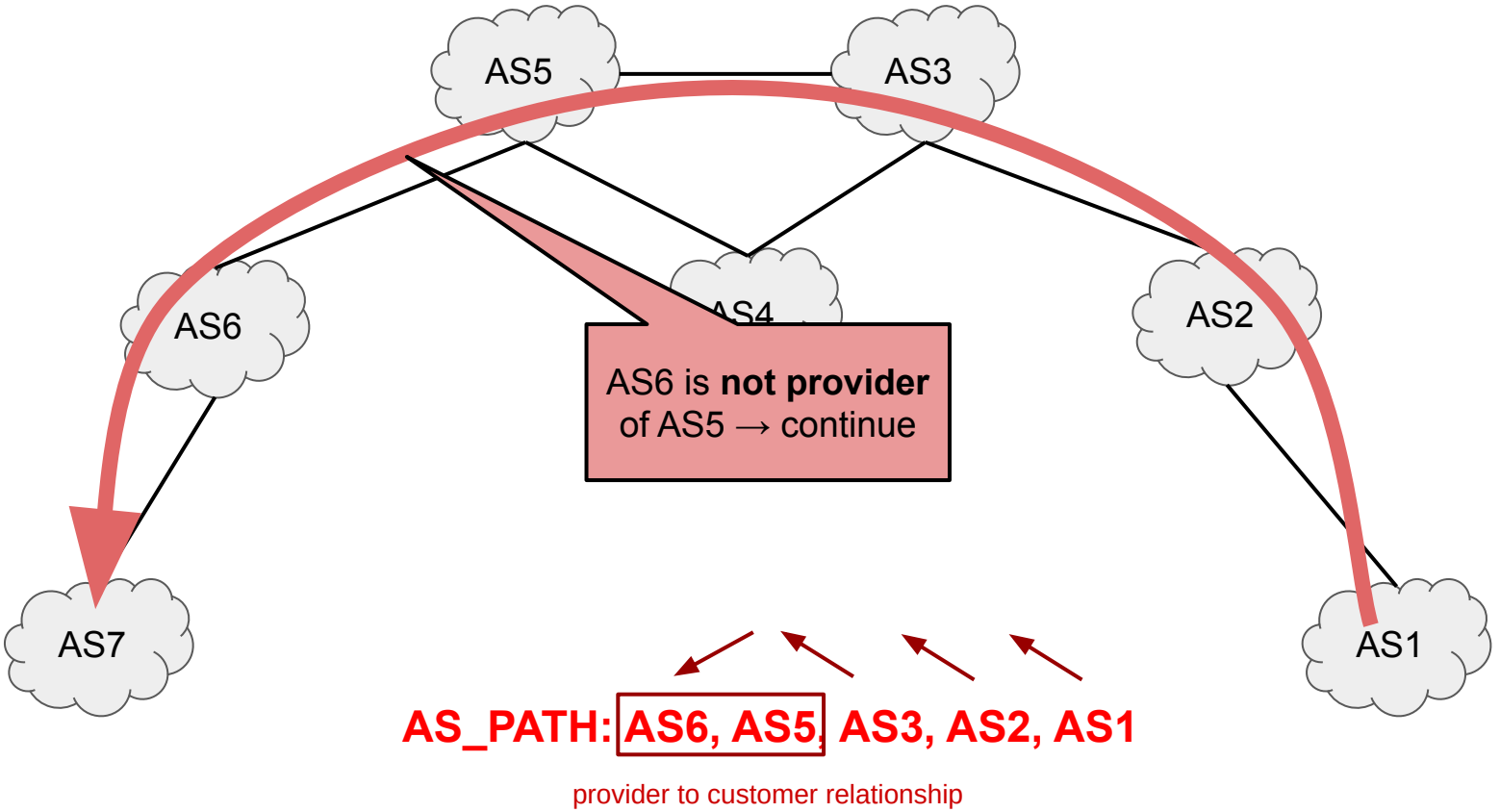
ASPA verification: downstream peers



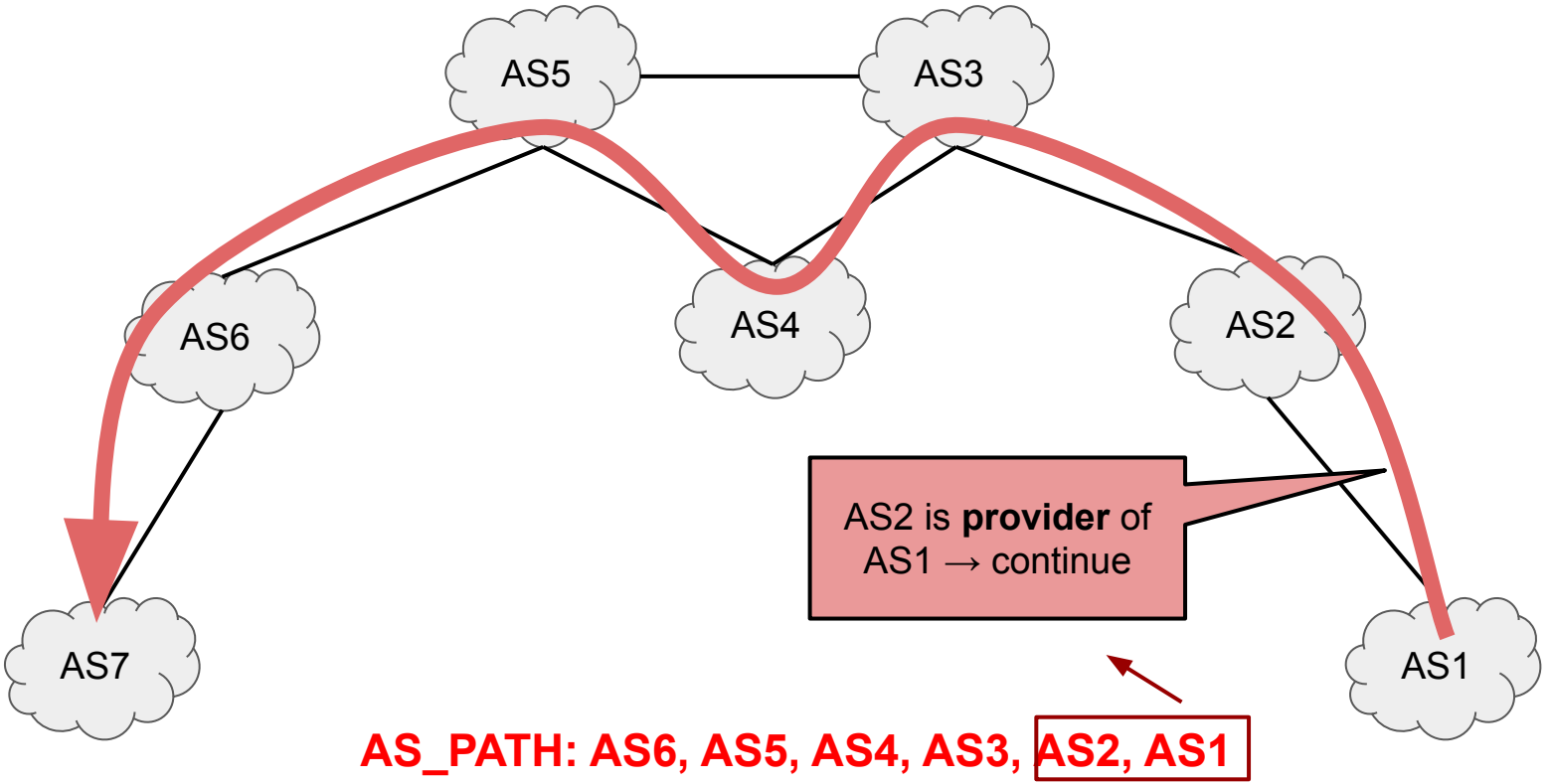
ASPA verification: downstream peers



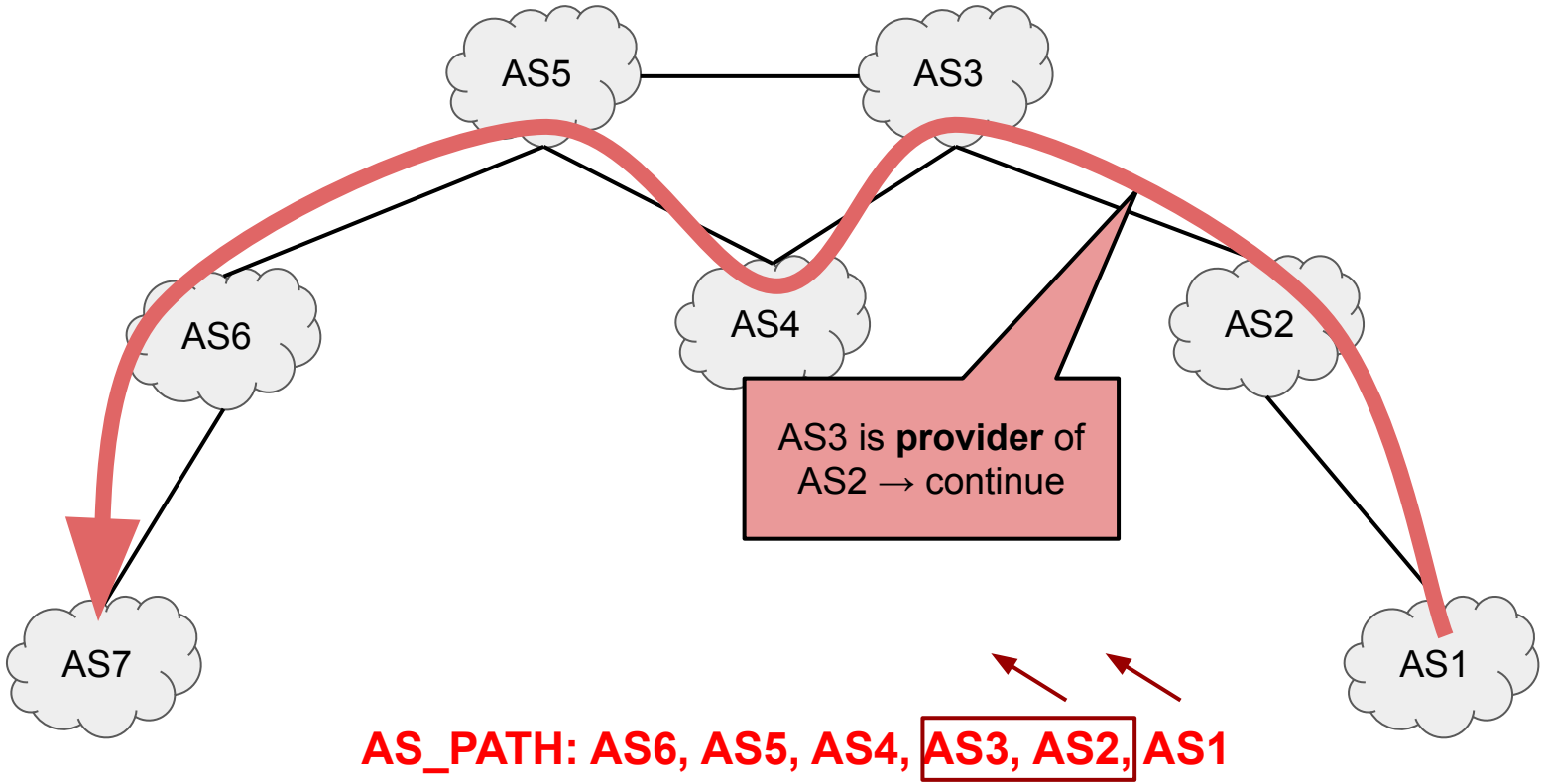
ASPA verification: downstream peers



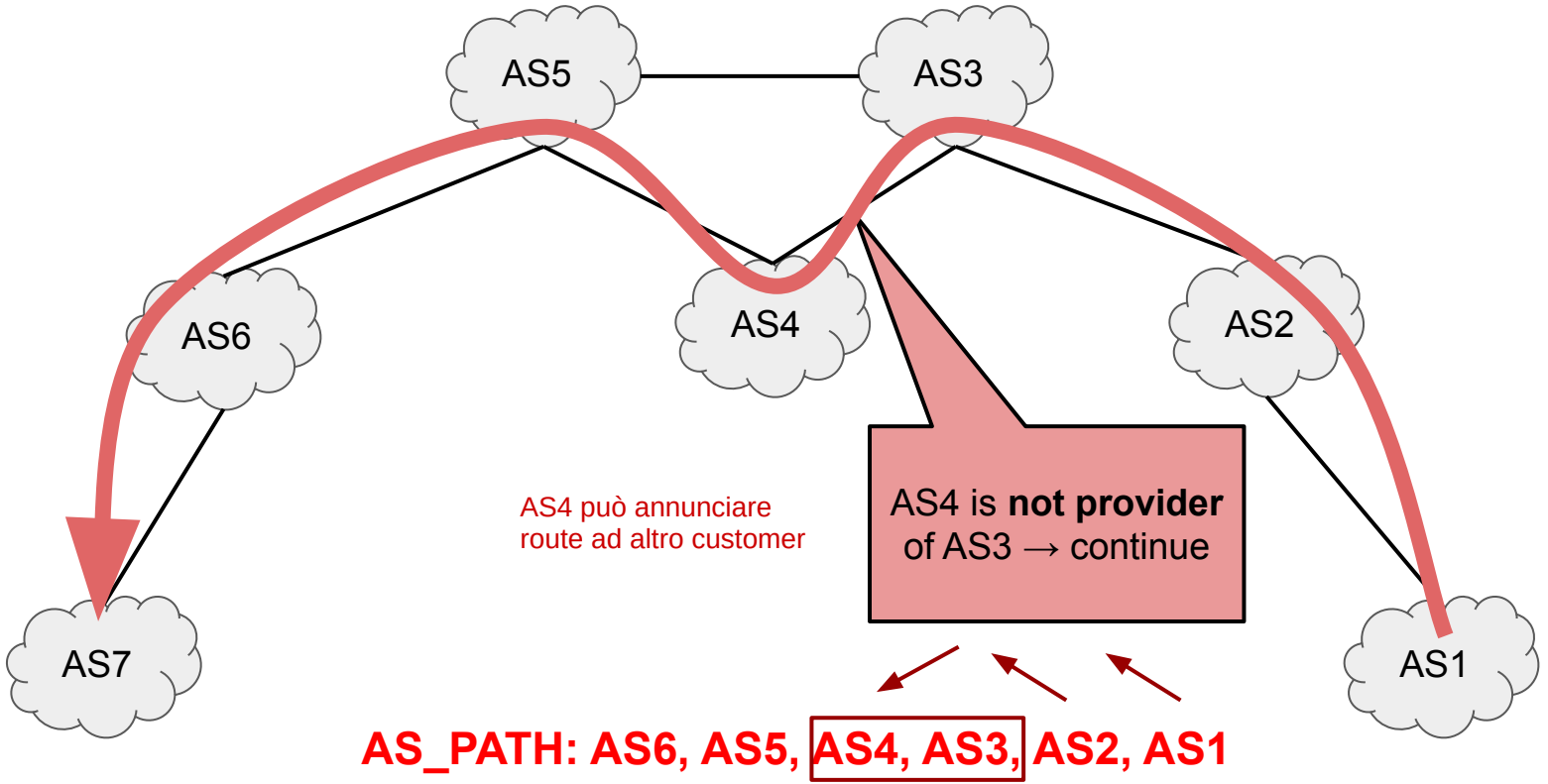
ASPA verification: downstream peers



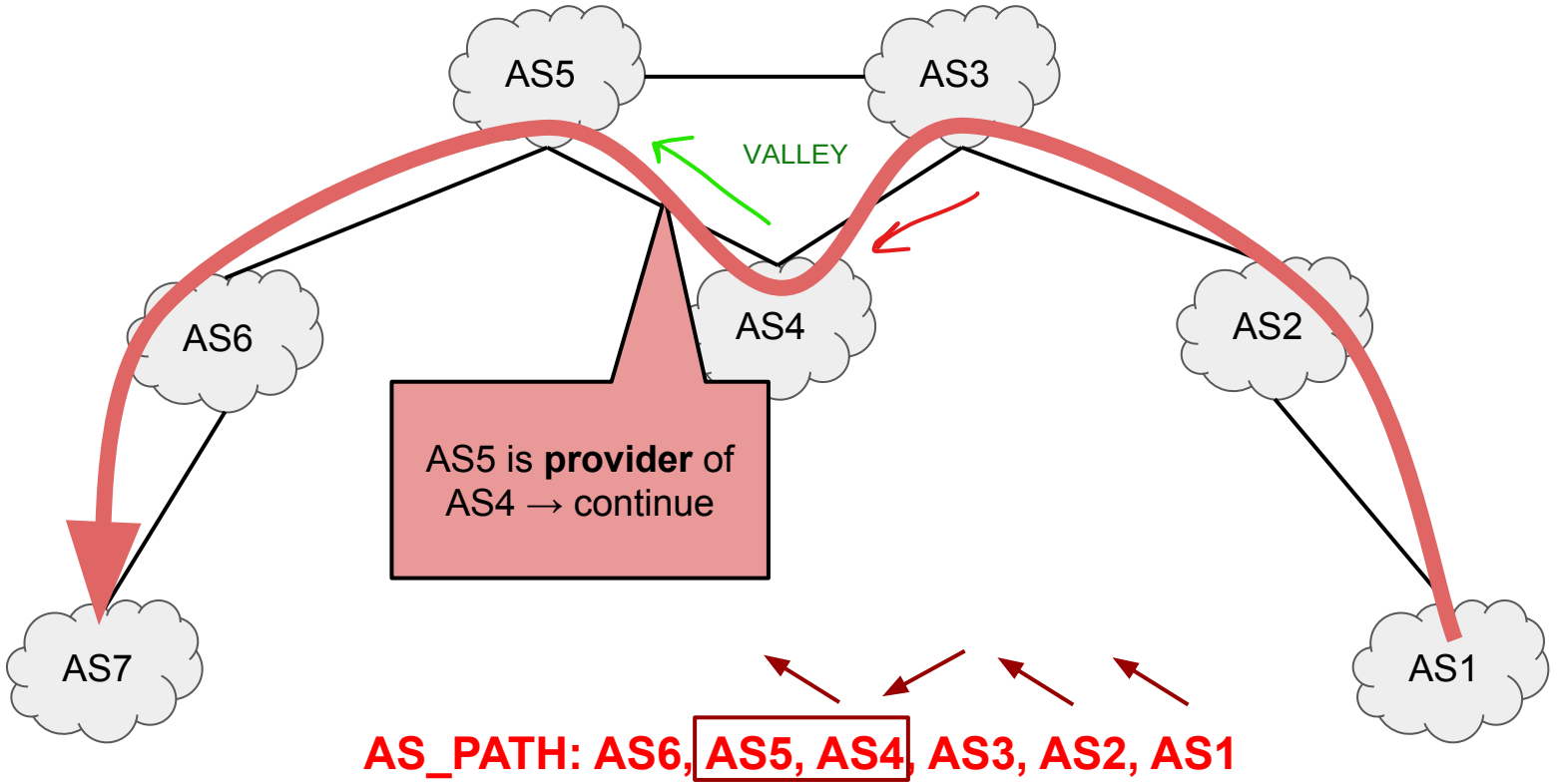
ASPA verification: downstream peers



ASPA verification: downstream peers

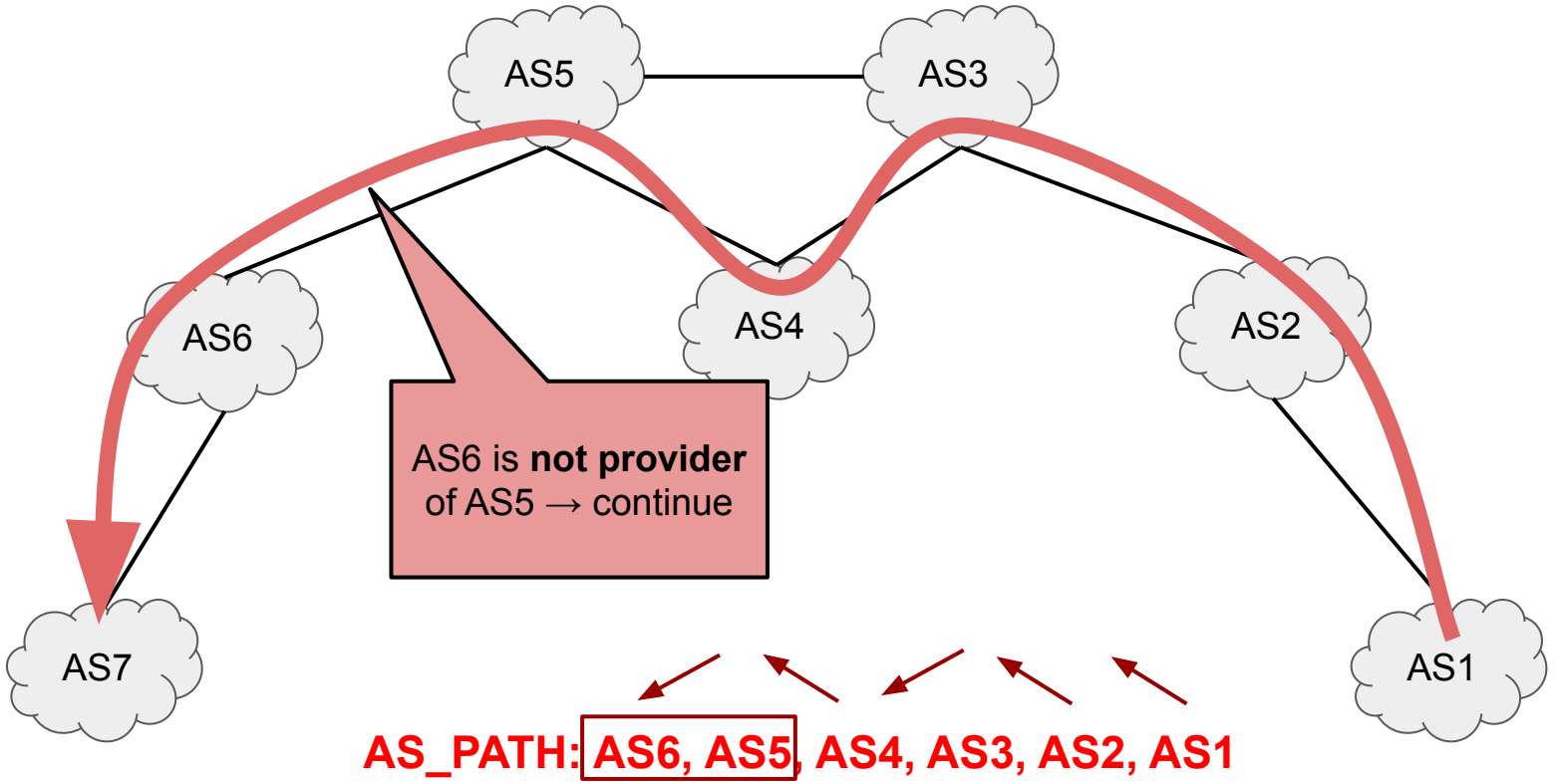


ASPA verification: downstream peers



identificata tramite questo meccanismo, perchè per ogni AS ho lista dei provider.

ASPA verification: downstream peers



RPKI implementations, LABs and tutorials

- NIST BGP (S)ecure (R)outing E(x)tension Software Suite (with demos) https://github.com/usnistgov/NIST-BGP-SRx
- FRRouting Project, a free and open source Internet routing protocol suite for Linux and Unix platforms: https://frrouting.org/
- Krill, an RPKI daemon, featuring a Certificate Authority (CA) and publication server, written in Rust: https://github.com/NLnetLabs/krill
- Online training lab on the APNIC virtual lab platform https://academy.apnic.net/en/virtual-labs?labId=87395
- Offline virtual lab based on 4 LXD containers and one Docker container https://github.com/eololab/rpki-lab https://eolo.it/nic/download/20190927/2-vergani.pdf
- How to Install an RPKI Validator https://labs.ripe.net/author/tashi_phuntsho_3/how-to-install-an-rpki-validator/
- Yet another one: https://blog.apnic.net/2021/08/11/rpki-mylab-environment/
- https://www.nsg.ee.ethz.ch/fileadmin/user_upload/theses/SA-2021-11.pdf
- **Last year students' project: https://github.com/ThetaRangers/KatharaRPKI**

Nella veloce demo del laboratorio, viene eseguita una versione SENZA RPKI ed una seconda con RPKI

Prima shell:
kathara lstart con tmux
inizializziamo la topology
customer1
kathara connect customer1r1
vogliamo fare trace route con customer3
passando per ISP1, poi IXP, e grazie a router3

traceroute 150.0.0.3

kathara connect customer2r1 su altra shell
nella cartella shared c'è attack.sh
in cui annunciamo network 150.0.0.0/28 e prefix list
avviamo con bash attack.sh

in pratica risponde customer2

Lab: RPKI with Kathara and BGP Hijacking

kathara lstart sempre con tmux
kathara customer1r1
traceroute 150.0.0.3
lab.conf è la configurazione sui link

kathara connect customer2r1
bash shared/attack.sh

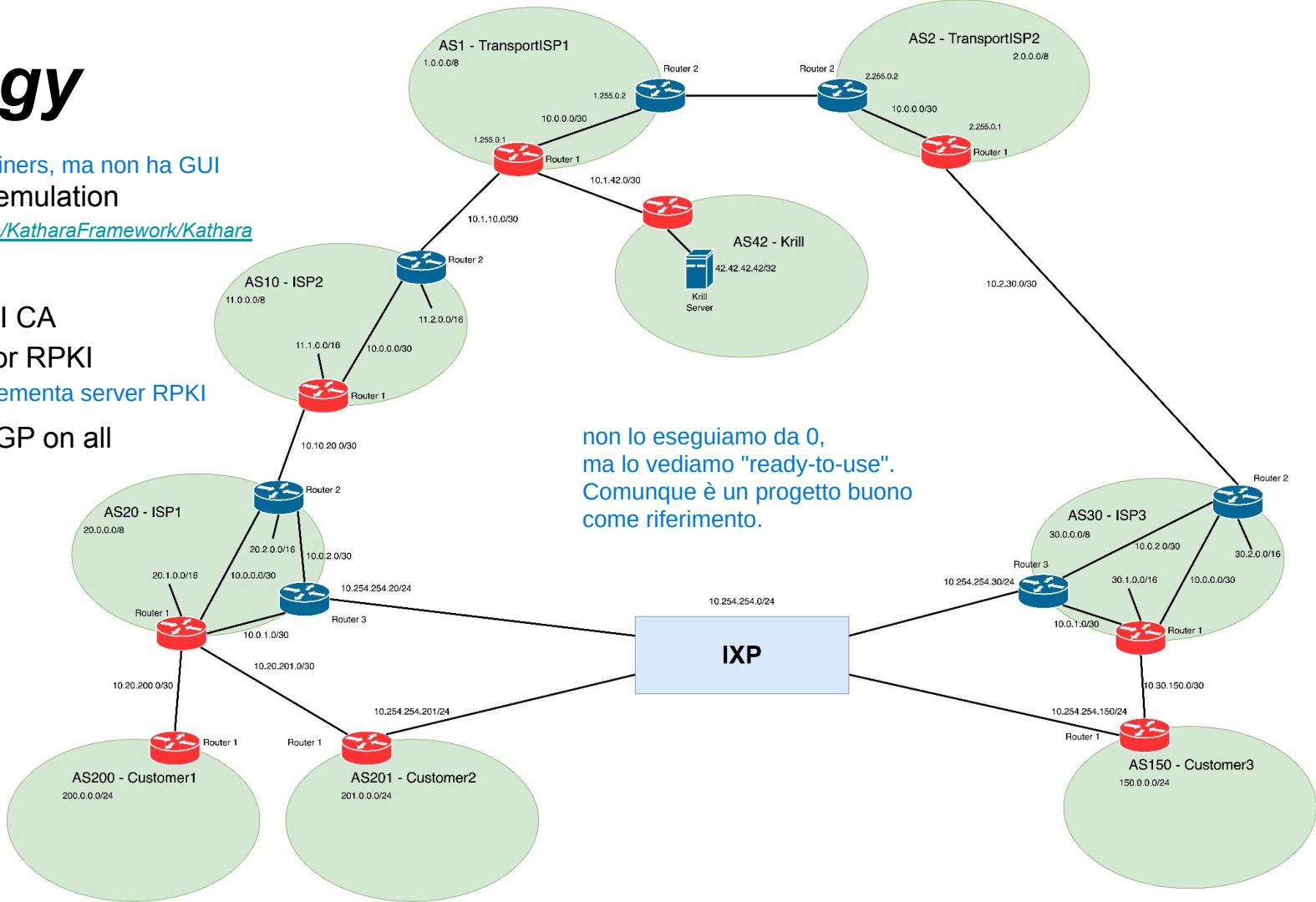
https://github.com/ThetaRangers/KatharaRPKI

Topology

basata su containers, ma non ha GUI

- Kathara for emulation
  - https://github.com/KatharaFramework/Kathara

- Krill for RPKI CA
- Routinator for RPKI routers implementa server RPKI
- FRR runs BGP on all routers



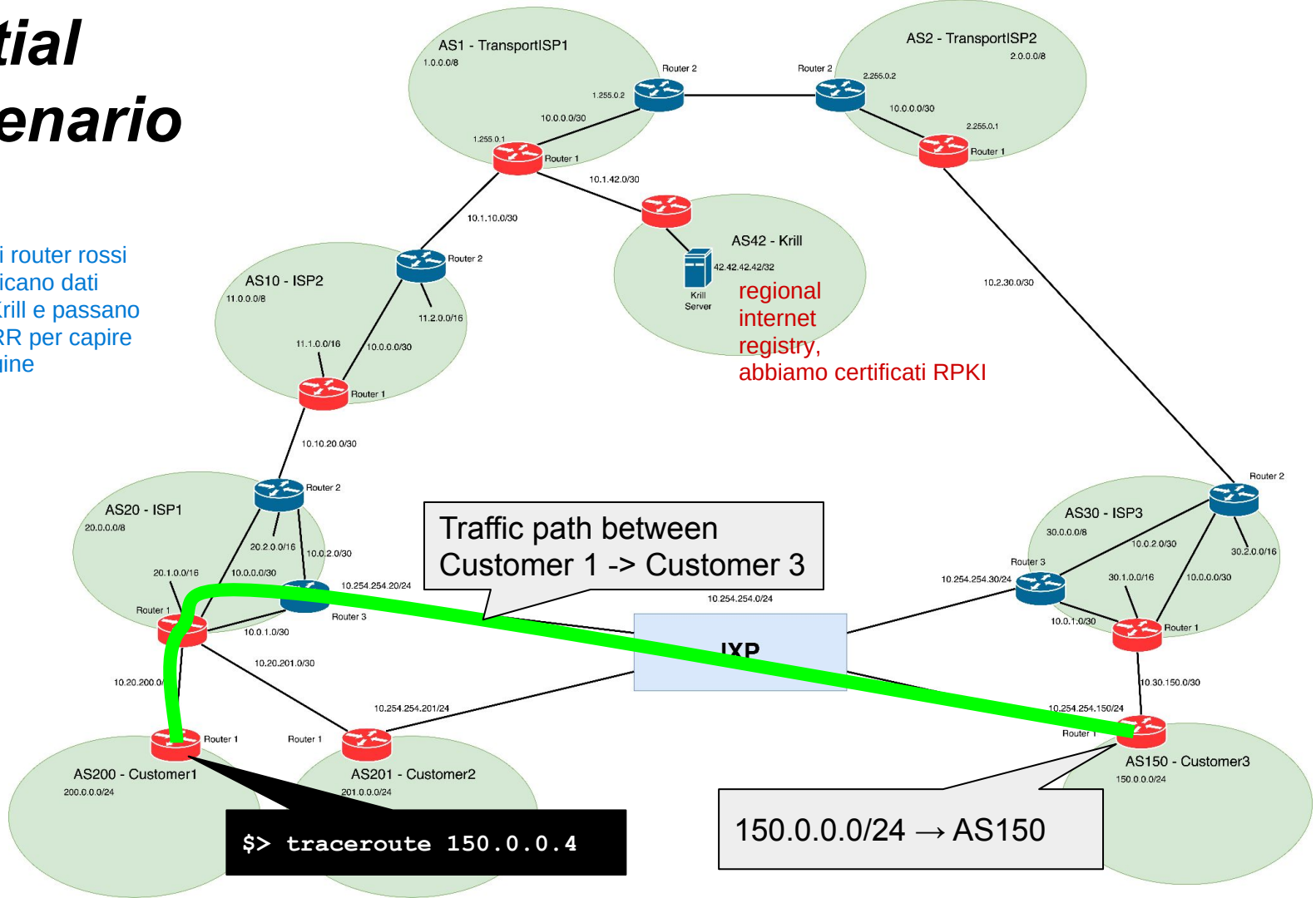
non lo eseguiamo da 0,
ma lo vediamo "ready-to-use".
Comunque è un progetto buono
come riferimento.

Demonstration

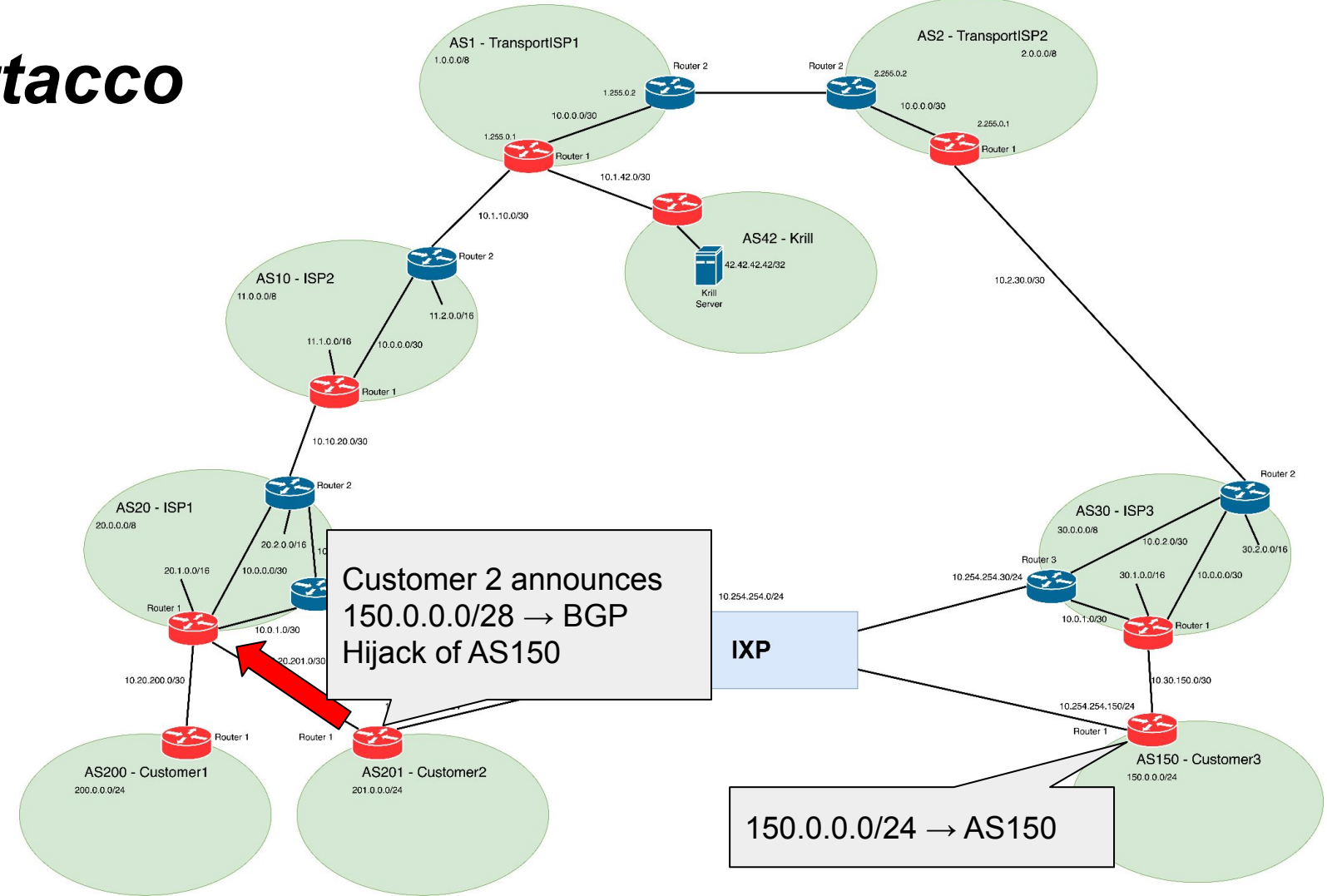
- Step 1:
  - Run the lab without RPKI
- Step 2:
  - From customer 1 run a traceroute to customer 3 network, to verify the path
- Step 3:
  - Customer 2 executes a BGP Subprefix Hijacking attack sending a more specific prefix for customer 3 network
  - Verify that the path of the packet is changed → successful attack
- Step 4:
  - Run again the lab with RPKI enabled
- Step 5:
  - Replicate the attack and verify that this time is not successful

Initial Scenario

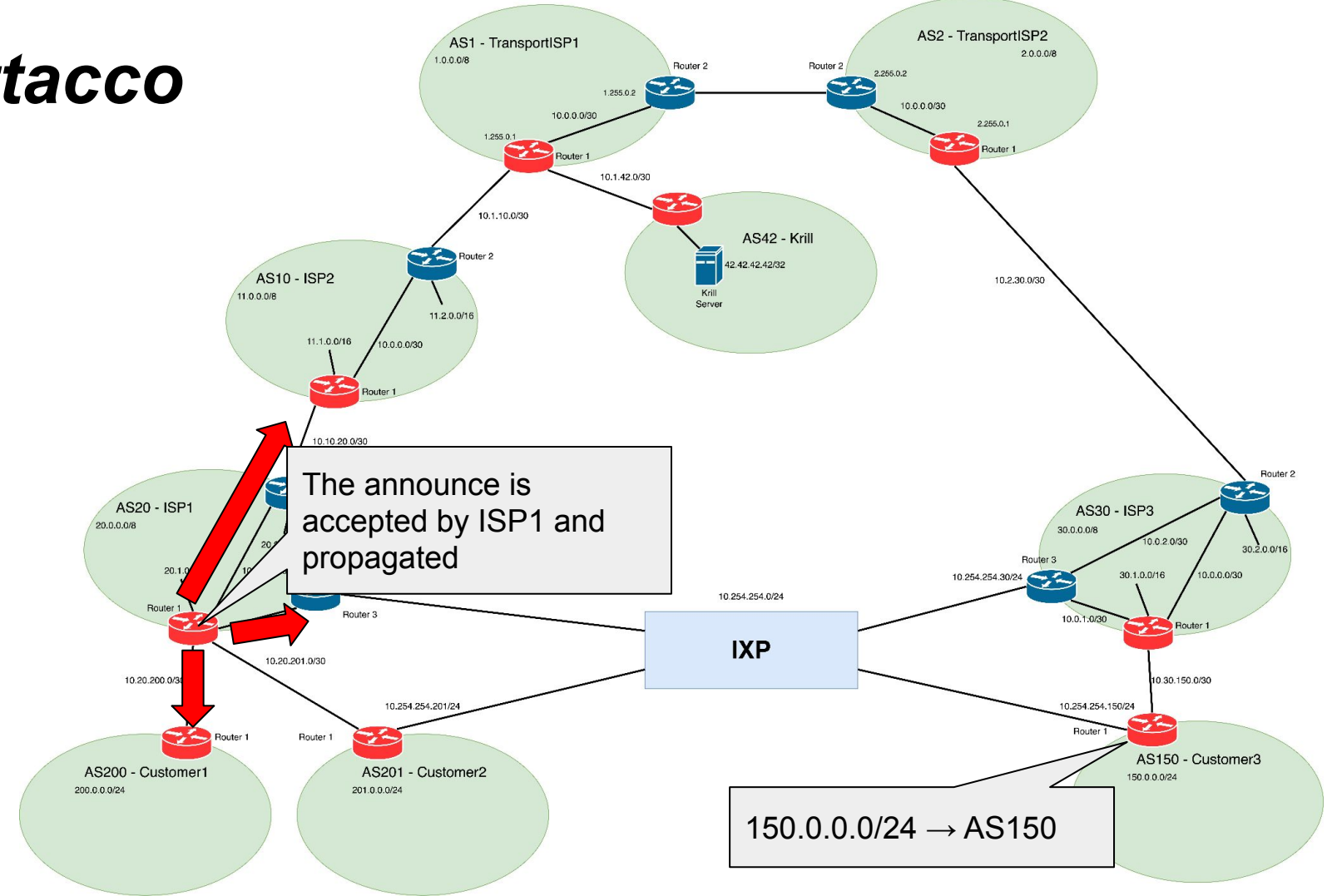
tutti i router rossi scaricano dati da Krill e passano a FRR per capire l'origine



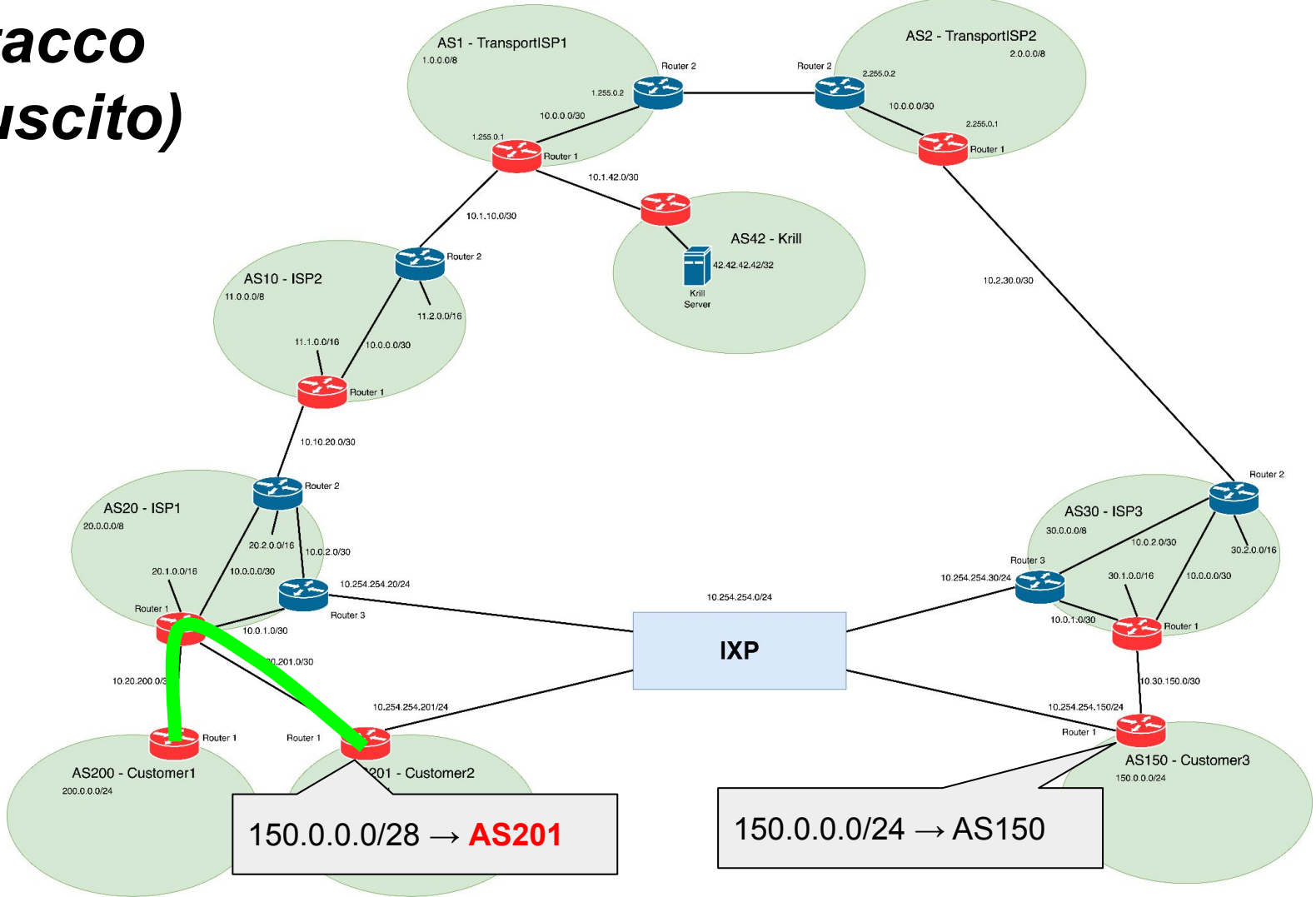
Attacco


AS1 - TransportISP1
1.0.0.0/8
Router 2
1.255.0.2
10.0.0.0/30
1.255.0.1
Router 1
AS2 - TransportISP2
2.0.0.0/8
Router 2
2.255.0.2
10.0.0.0/30
2.255.0.1
Router 1
10.1.42.0/30
10.1.10.0/30
AS42 - Krill
42.42.42.42/32
Krill Server
10.2.30.0/30
AS10 - ISP2
11.0.0.0/8
Router 2
11.2.0.0/16
11.1.0.0/16
10.0.0.0/30
Router 1
10.10.20.0/30
AS20 - ISP1
20.0.0.0/8
Router 2
20.2.0.0/16
20.1.0.0/16
10.0.0.0/30
Router 1
10.0.1.0/30
20.201.0/30
10.20.200.0/30
Router 1
AS200 - Customer1
200.0.0.0/24
Router 1
AS201 - Customer2
201.0.0.0/24
Customer 2 announces
150.0.0.0/28 → BGP
Hijack of AS150
10.254.254.0/24
IXP
Router 2
AS30 - ISP3
30.0.0.0/8
10.0.2.0/30
30.2.0.0/16
Router 3
10.254.254.30/24
30.1.0.0/16
10.0.0.0/30
10.0.1.0/30
Router 1
10.30.150.0/30
10.254.254.150/24
Router 1
AS150 - Customer3
150.0.0.0/24
150.0.0.0/24 → AS150

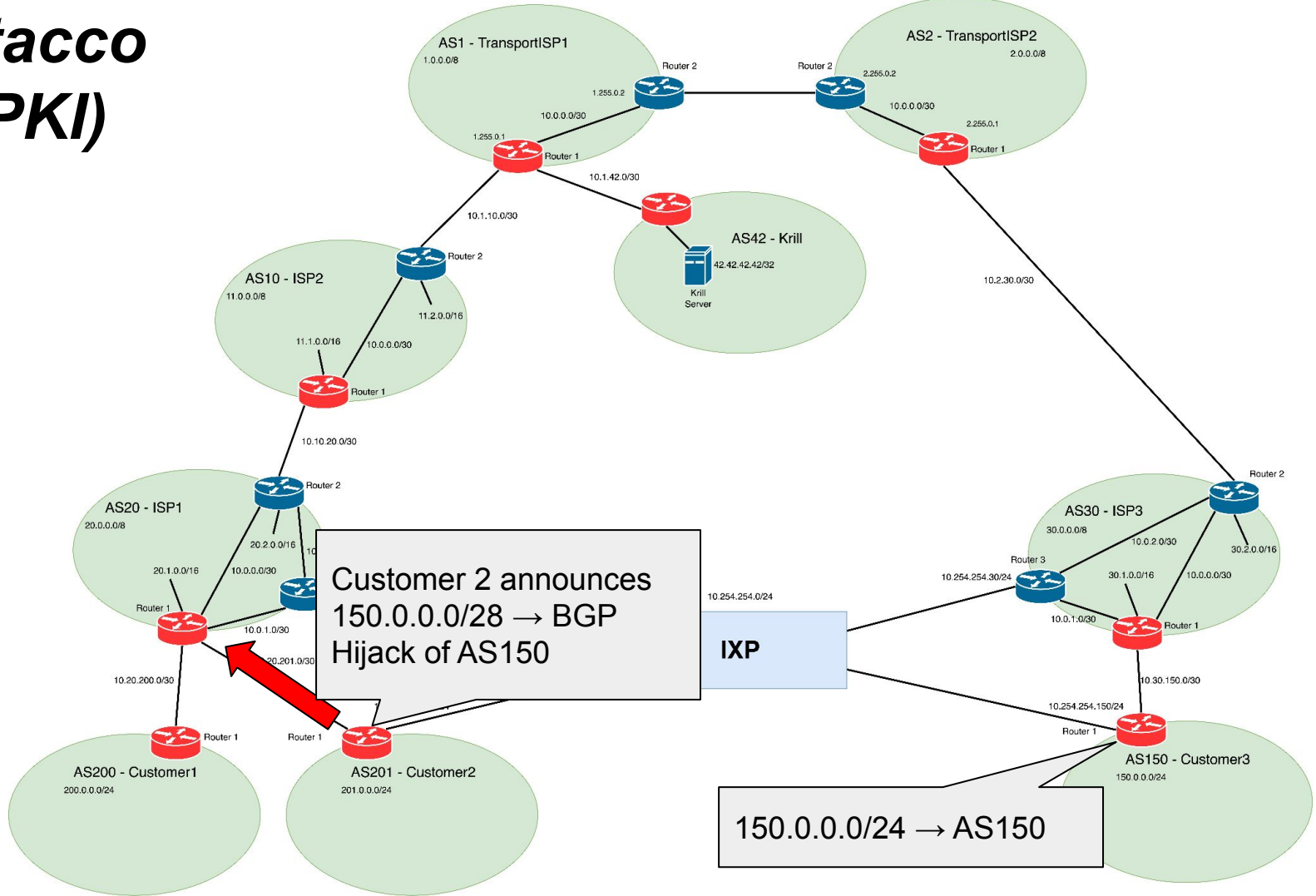
Attacco


AS1 - TransportISP1
1.0.0.0/8
Router 2
1.255.0.2
10.0.0.0/30
1.255.0.1
Router 1
AS2 - TransportISP2
2.0.0.0/8
Router 2
2.255.0.2
10.0.0.0/30
2.255.0.1
Router 1
10.1.42.0/30
10.1.10.0/30
AS42 - Krill
42.42.42.42/32
Krill Server
AS10 - ISP2
11.0.0.0/8
Router 2
11.2.0.0/16
11.1.0.0/16
10.0.0.0/30
Router 1
10.2.30.0/30
10.10.20.0/30
The announce is accepted by ISP1 and propagated
AS20 - ISP1
20.0.0.0/8
Router 1
Router 3
10.254.254.0/24
IXP
AS30 - ISP3
30.0.0.0/8
Router 2
10.0.2.0/30
30.2.0.0/16
Router 3
10.254.254.30/24
30.1.0.0/16
10.0.0.0/30
10.0.1.0/30
Router 1
10.20.201.0/30
10.30.150.0/30
10.254.254.201/24
10.254.254.150/24
Router 1
Router 1
Router 1
AS200 - Customer1
200.0.0.0/24
AS201 - Customer2
201.0.0.0/24
AS150 - Customer3
150.0.0.0/24
150.0.0.0/24 → AS150

Attacco (riuscito)



Attacco (RPKI)



Attacco (RPKI)