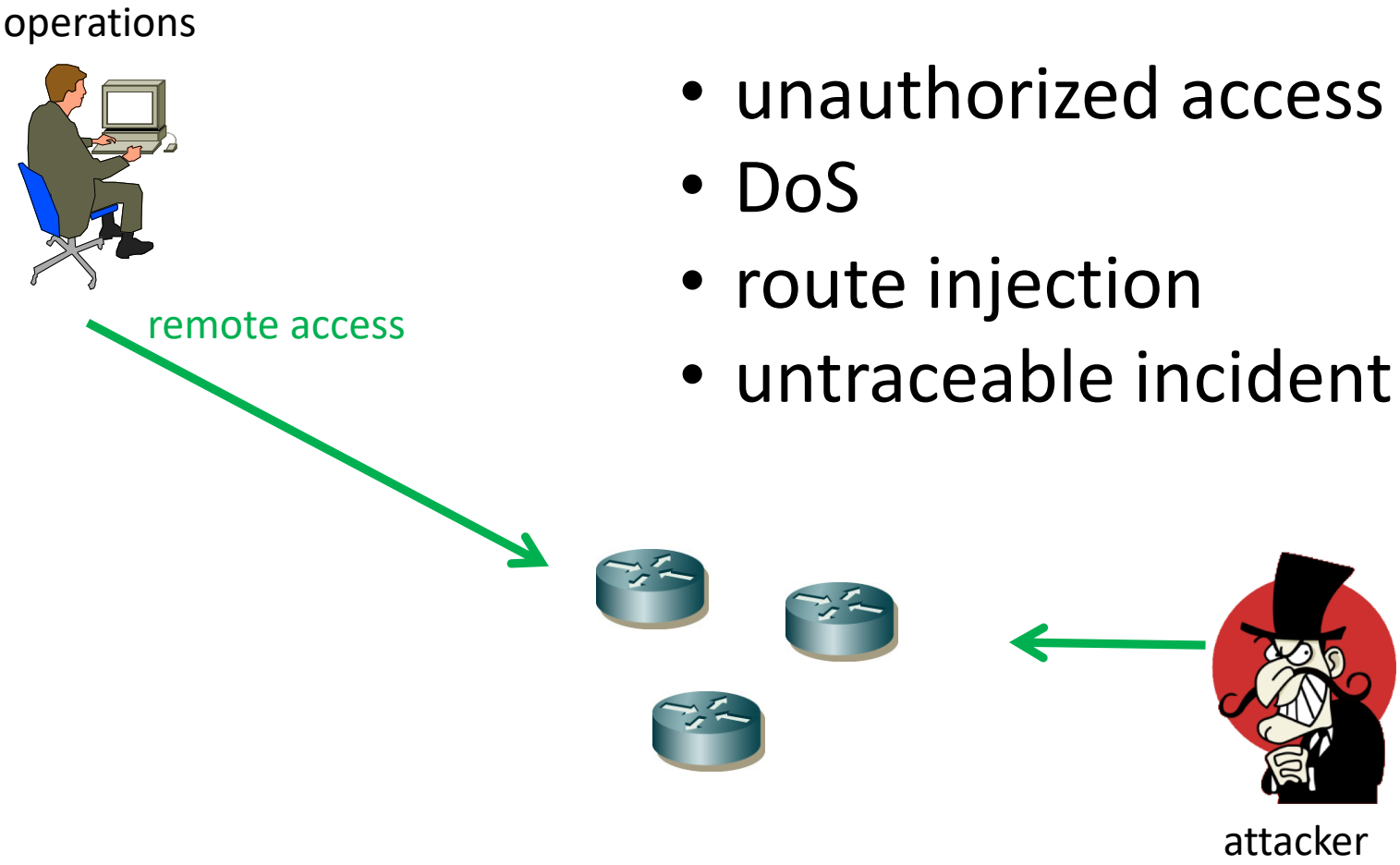
Securing network infrastructure

Matsuzaki ‘maz’ Yoshinobu
<maz@iij.ad.jp>

Our Goals

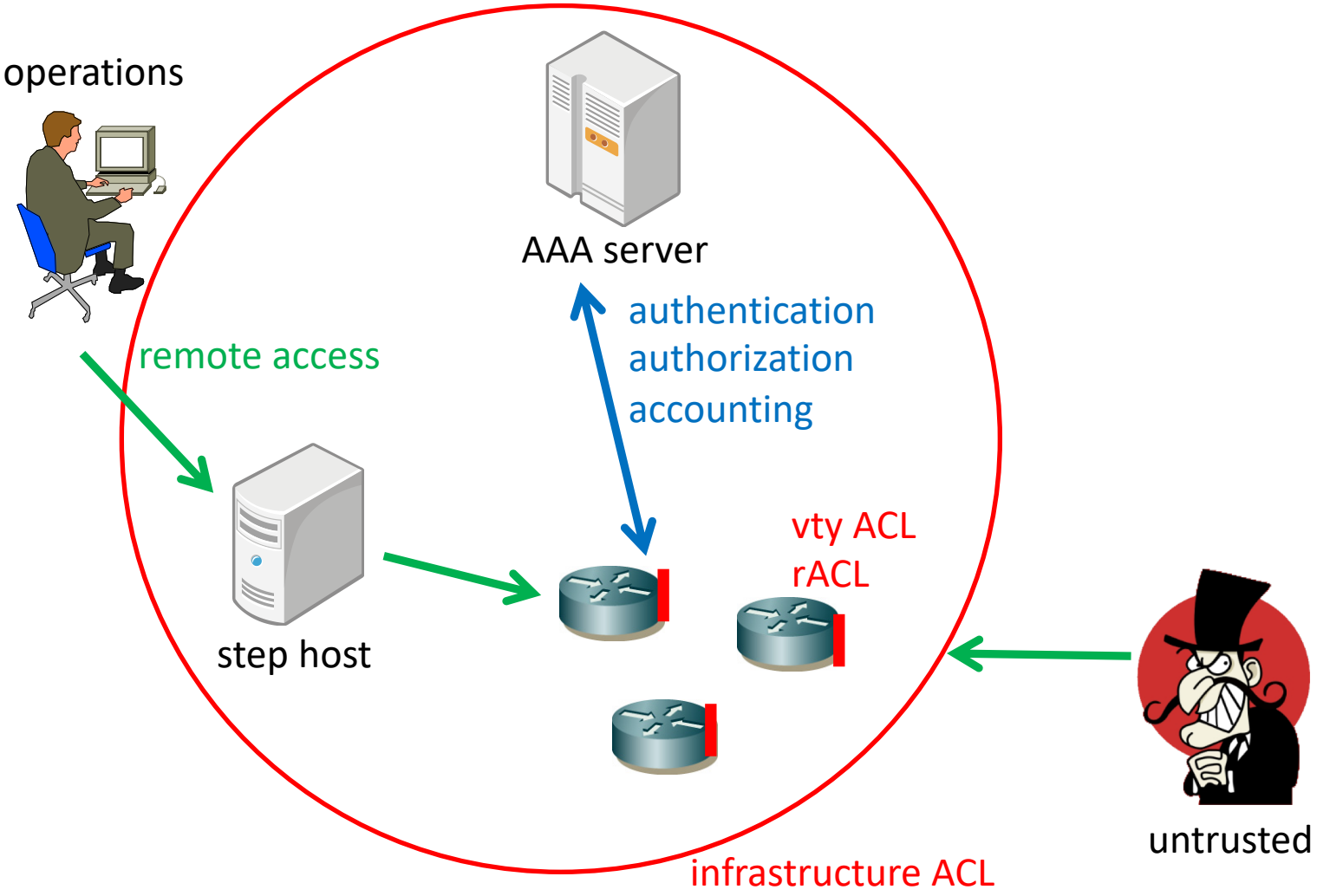
- Ensuring Network Availability
- Controlling Routing Policy
- Protecting Information
- Preventing Misuse
- Mitigating Attacks
- Responding to Incidents
- etc.

Risks



- unauthorized access
- DoS
- route injection
- untraceable incident

protecting devices



AAA server and remote access

- Authentication, Authorization, Accounting
  - tacacs, radius
- each operators has own login account
  - You can set privileges per tasks of the operator
- logging at AAA servers
  - where (device)
  - who (login account)
  - what (command)

Remote Access to Devices

- in-band access
  - vty, snmp, ntp, etc...
  - IP reachability is required
  - useful for daily operations
- out-of-band access
  - serial console
  - workable without IP reachability
  - useful for restoration

Access Control for in-band access

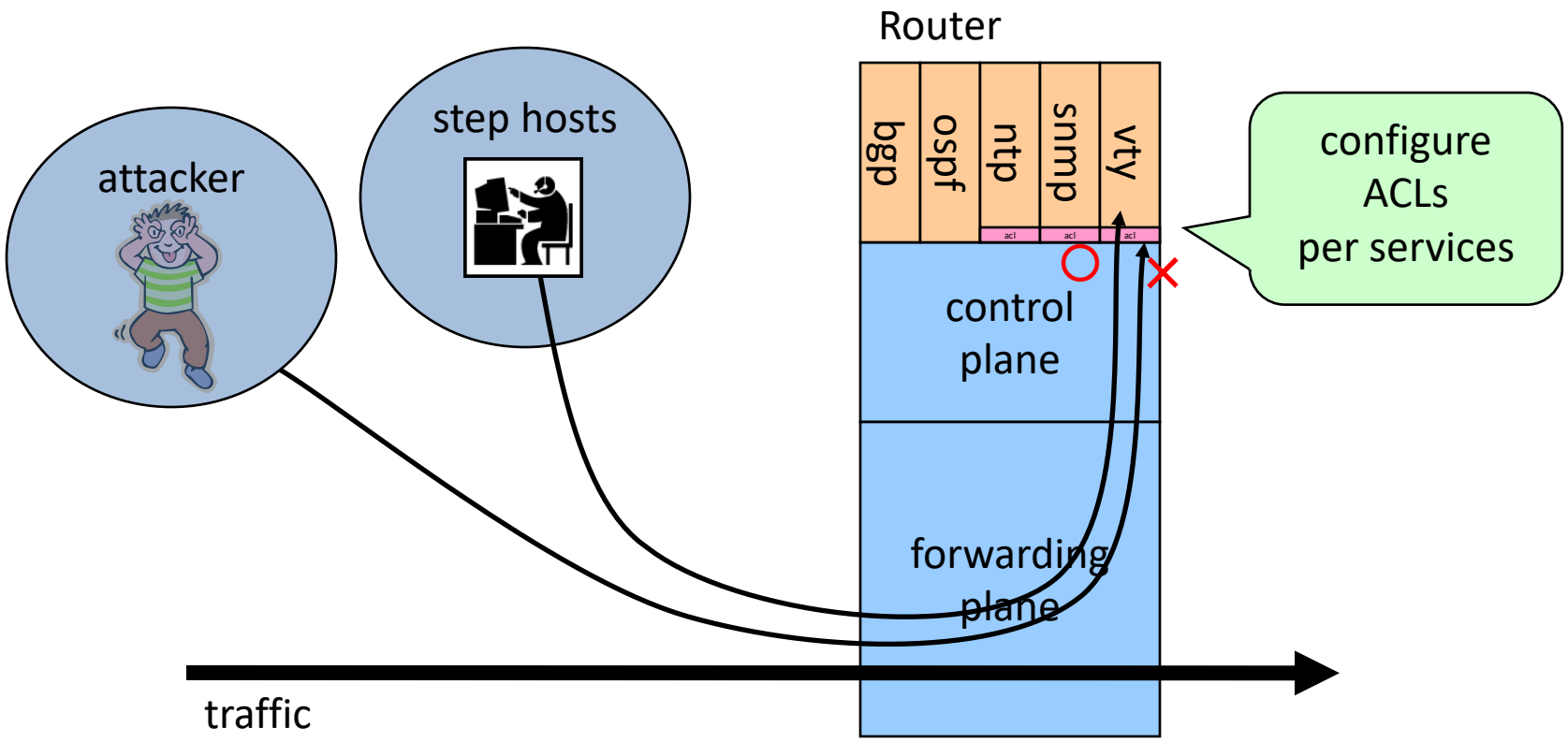
- operations need to access remote devices
  - to manage the devices
- packet filtering on vty, snmp and etc
  - to protect devices from unauthorized access
  - allow access from trusted network only
    - source IP address based filtering

step hosts

- are placed on a trusted network
- useful to enforce more restricted control
- Easier management of OPS user account

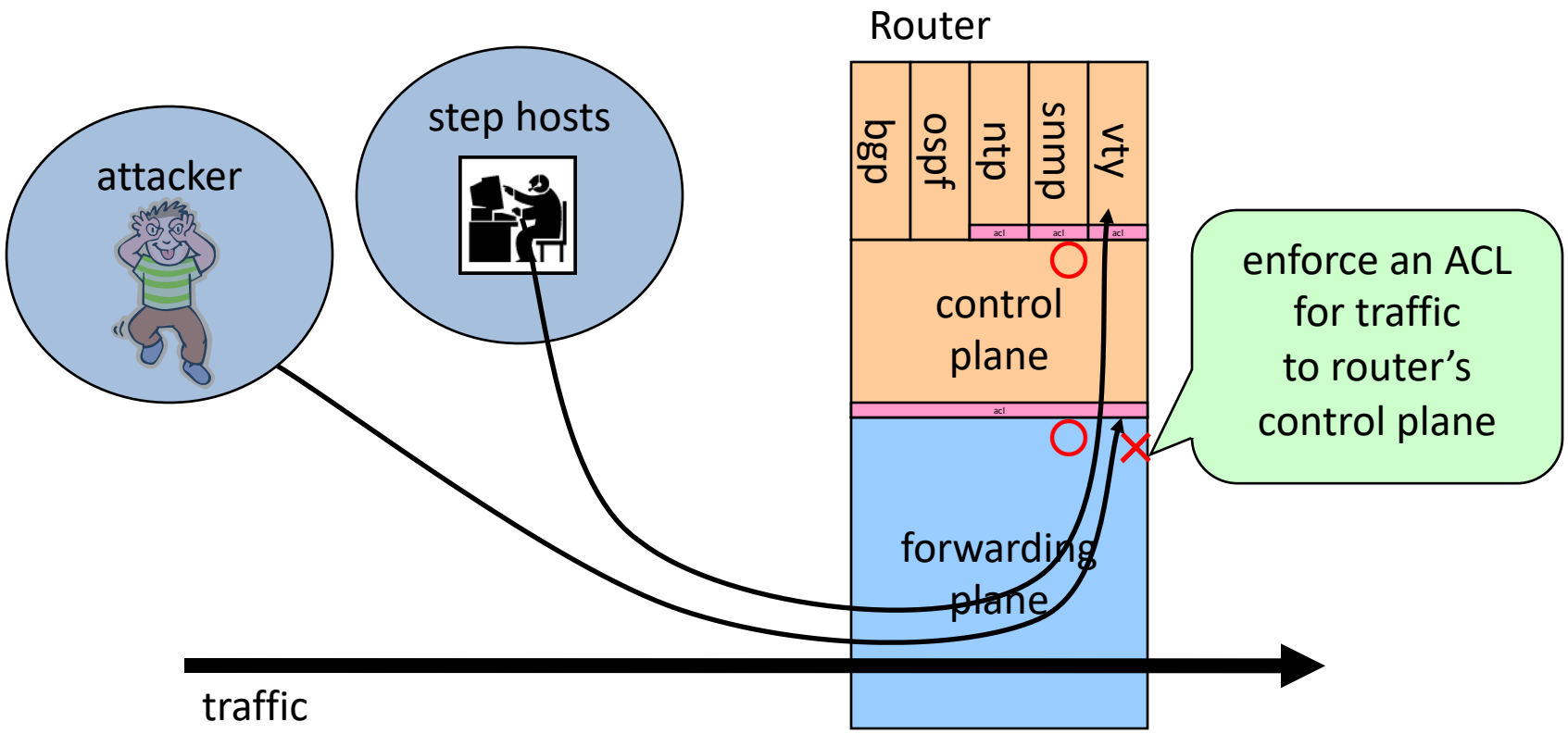
- logging on step hosts
  - typescript of a VTY session
  - login/logout

access control per services



Received/Router ACL (rACL)

access control against control plane

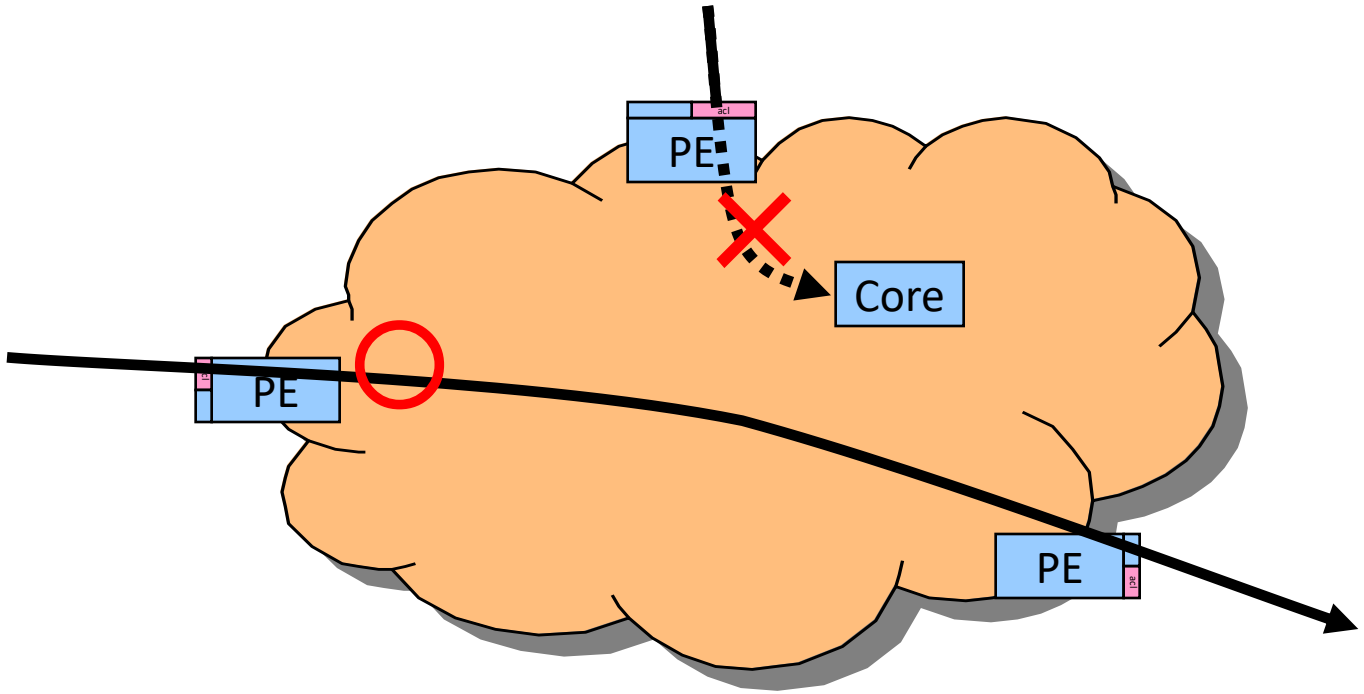


infrastructure ACL

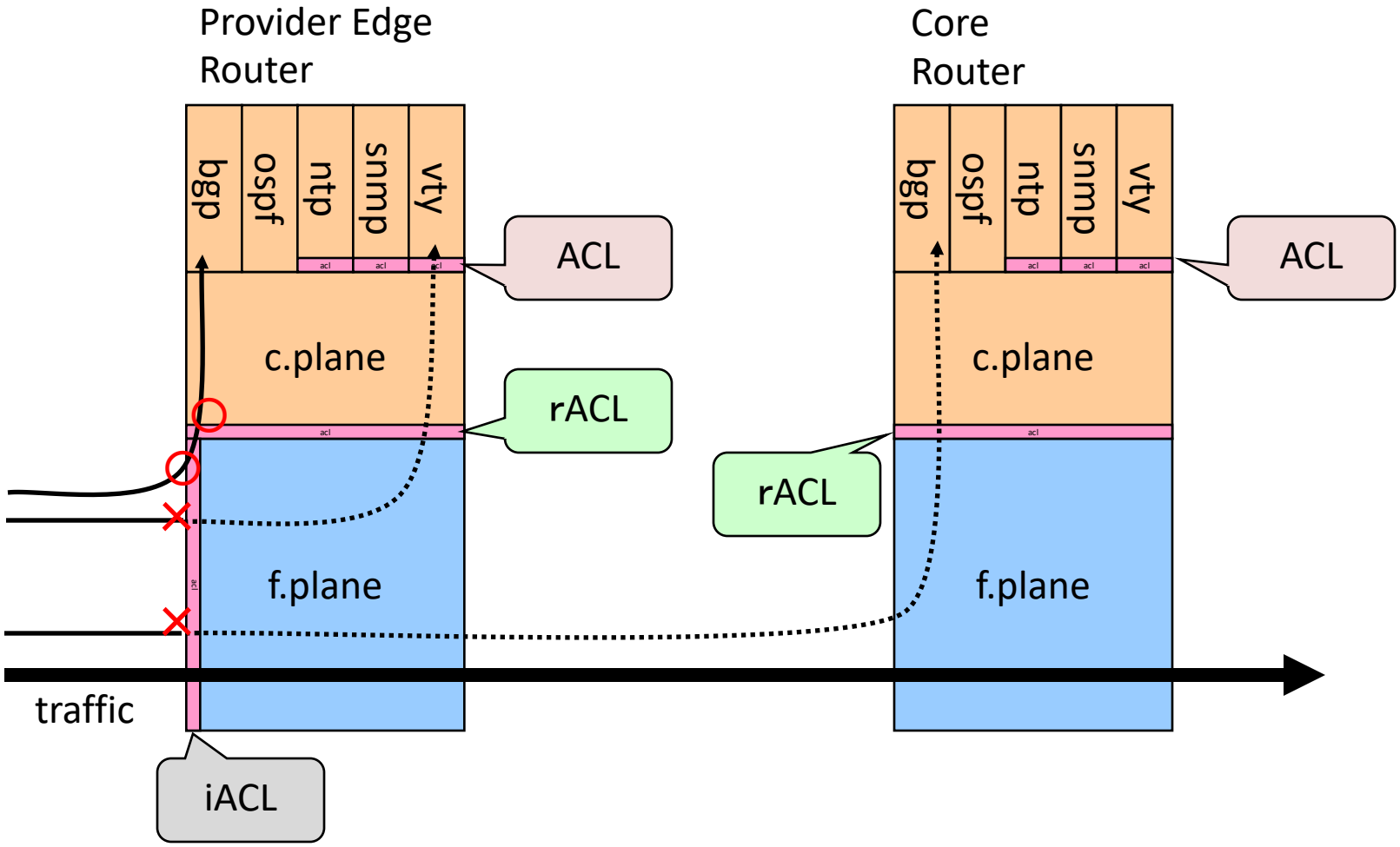
- to protect our management traffic
  - not too much
  - ping, traceroute to our devices should be workable
- deny packets from INFRA to INFRA on edge
  - INFRA: routers, step hosts and so on
    - these ip range should be stayed inside

Infrastructure ACL (iACL)

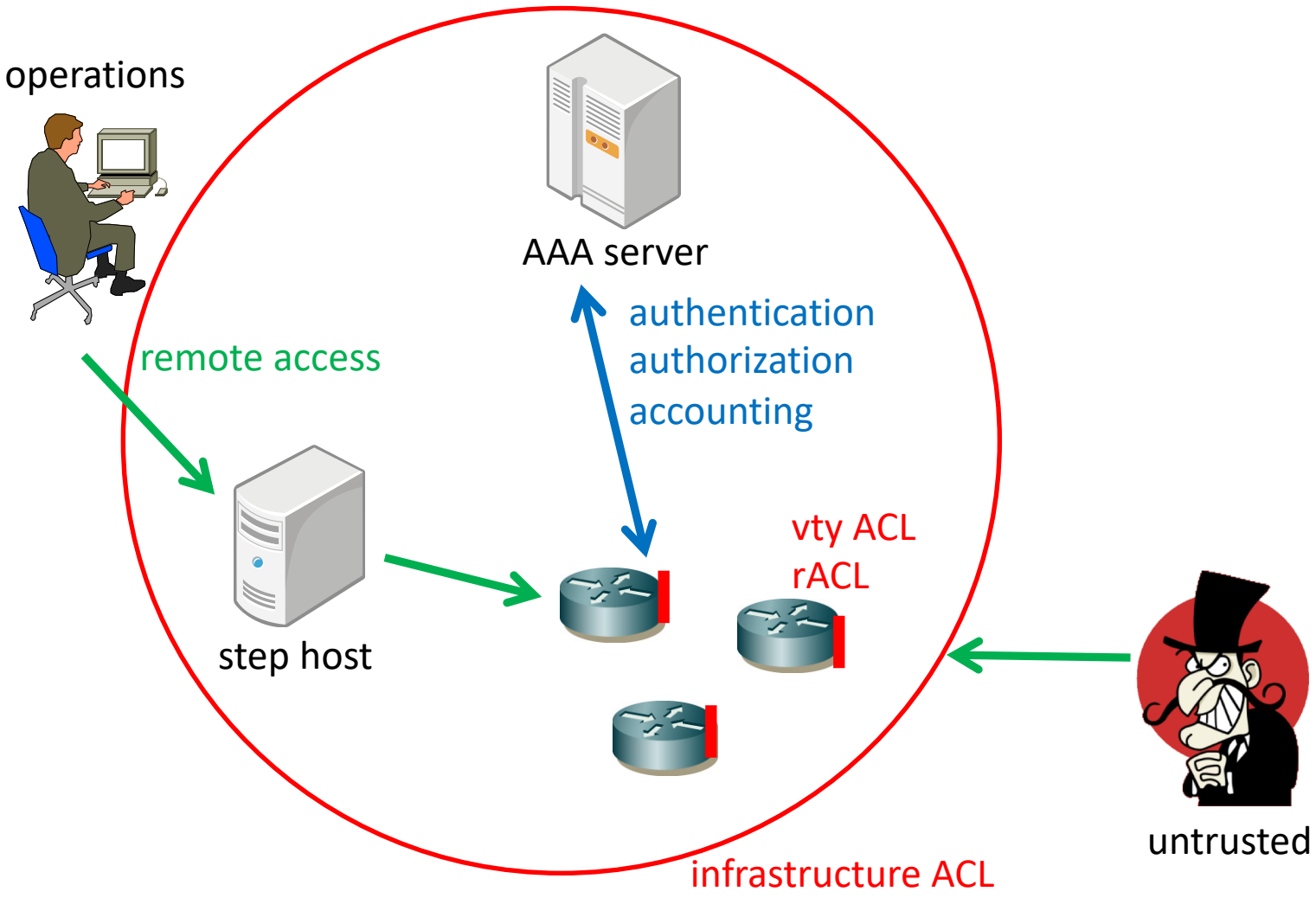
- enforce a policy on the network edge



multiple ACLs to protect Devices



protecting devices



config audit

- Maintain history of configuration changes
  - expect
  - RANCID
    - https://www.shrubbery.net/rancid/
- sanity check
  - Access controls
  - Routing policies
  - Filtering rules

Monitoring

- what's happened in the past
- syslog
  - to record messages from devices/softwares
- snmp
  - to monitor resources
- netflow
  - to monitor packet flows

syslog messages

- Nov 9 15:19:14.390 UTC: config[65775]: %MGBL-SYS-5-CONFIG_I : Configured from console by maz on vty0 (2001:db8:120:100:e1dd:97f3:fd98:a51f)

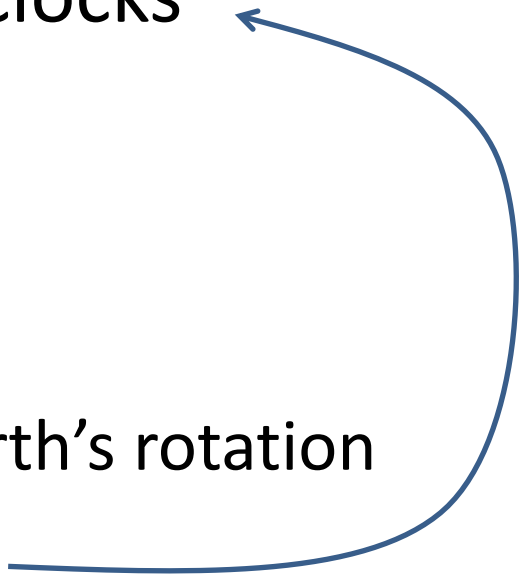
- Nov 12 13:53:38 maz sudo: maz : user NOT in sudoers ; TTY=pts/3 ; PWD=/home/maz ; USER=root ; COMMAND=/bin/bash

synced timestamp

- To make log messages useful
  - To compare incidents among devices
  - To compare time-related events
- Use NTP to sync clocks
  - Good enough for our purpose (around 1sec order)
  - If you need more accurate time clock, you might consider PTP instead

clock = oscillation + counter

- TAI = weighted average of atom clocks
  - TAI: International Atomic Time
- UTC = TAI + leap seconds
  - UTC: Coordinated Universal Time
  - leap seconds: to adjust clock to Earth's rotation
- atom clocks are adjusted to TAI
- localtime = UTC + timezone (+ summer time)

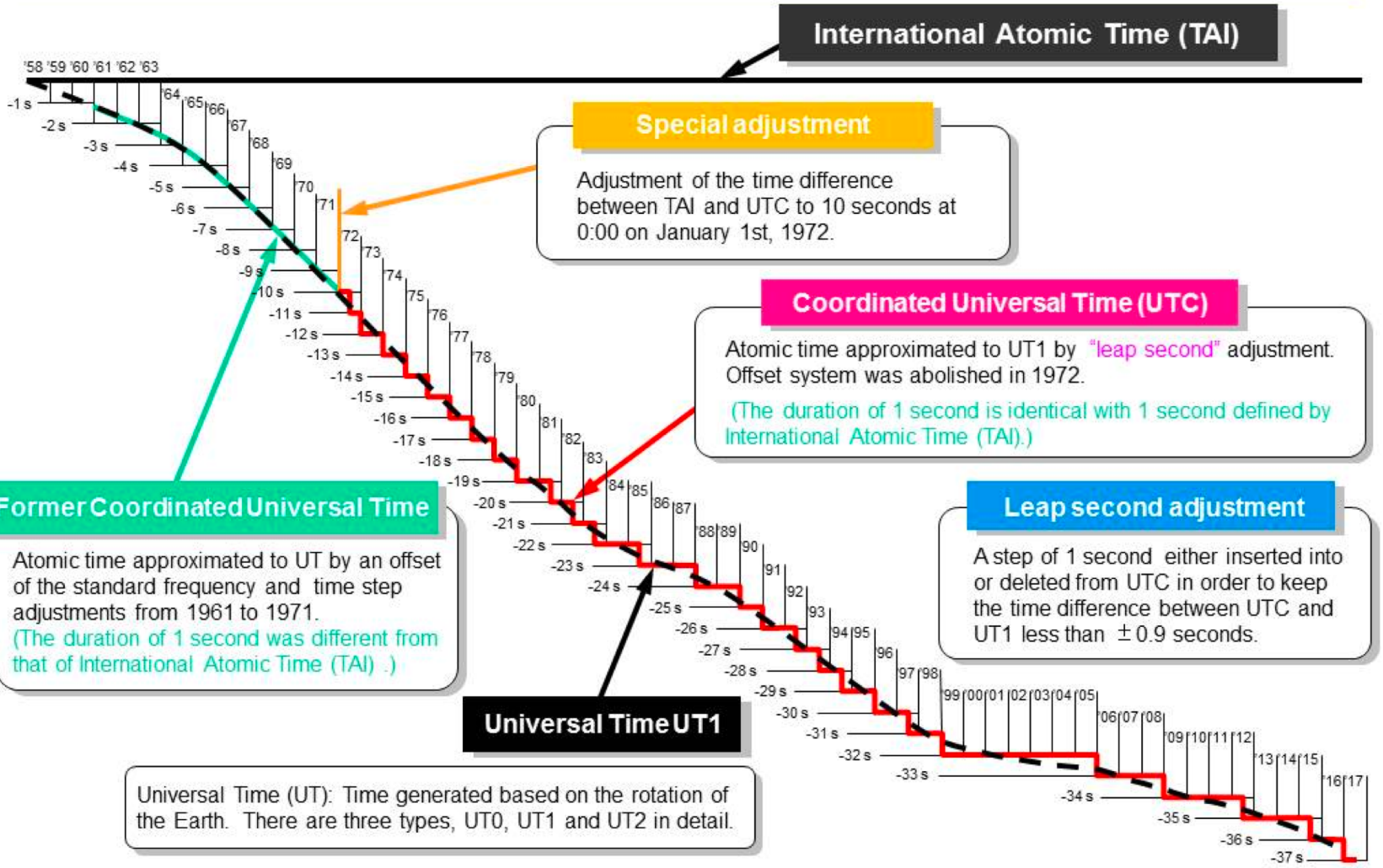


leap second

- There is no scheduled leap second in 2020
- Your applications should work as usual even the leap second introduced….
  - https://git.kernel.org/cgit/linux/kernel/git/torvalds/linux.git/commit/?id=6b43ae8a619d17c4935c3320d2ef9e92bdeed05d

Atomic Time and Leap Seconds

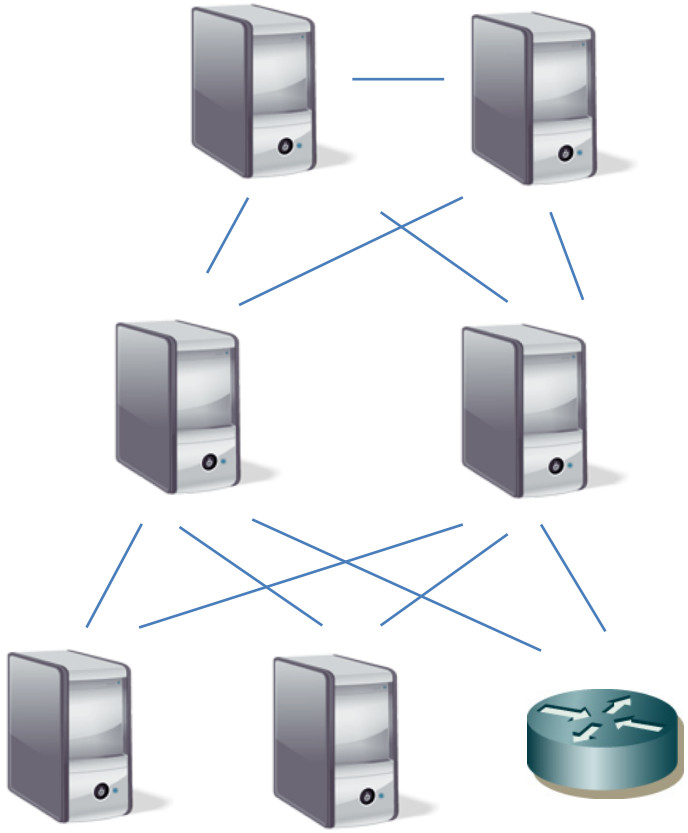
Japan Standard Time Group
National Institute of Information and Communications Technology



http://jjy.nict.go.jp/mission/page1-e.html

NTP strata

- Stratum 1
  - Primary time servers that are synchronized with their attached clock source (ATOM clock, GPS, radio)
- Stratum 2
  - Computers that are synchronized with Stratum 1 servers
- Stratum 3
  - Computers that are synchronized with Stratum 2 servers



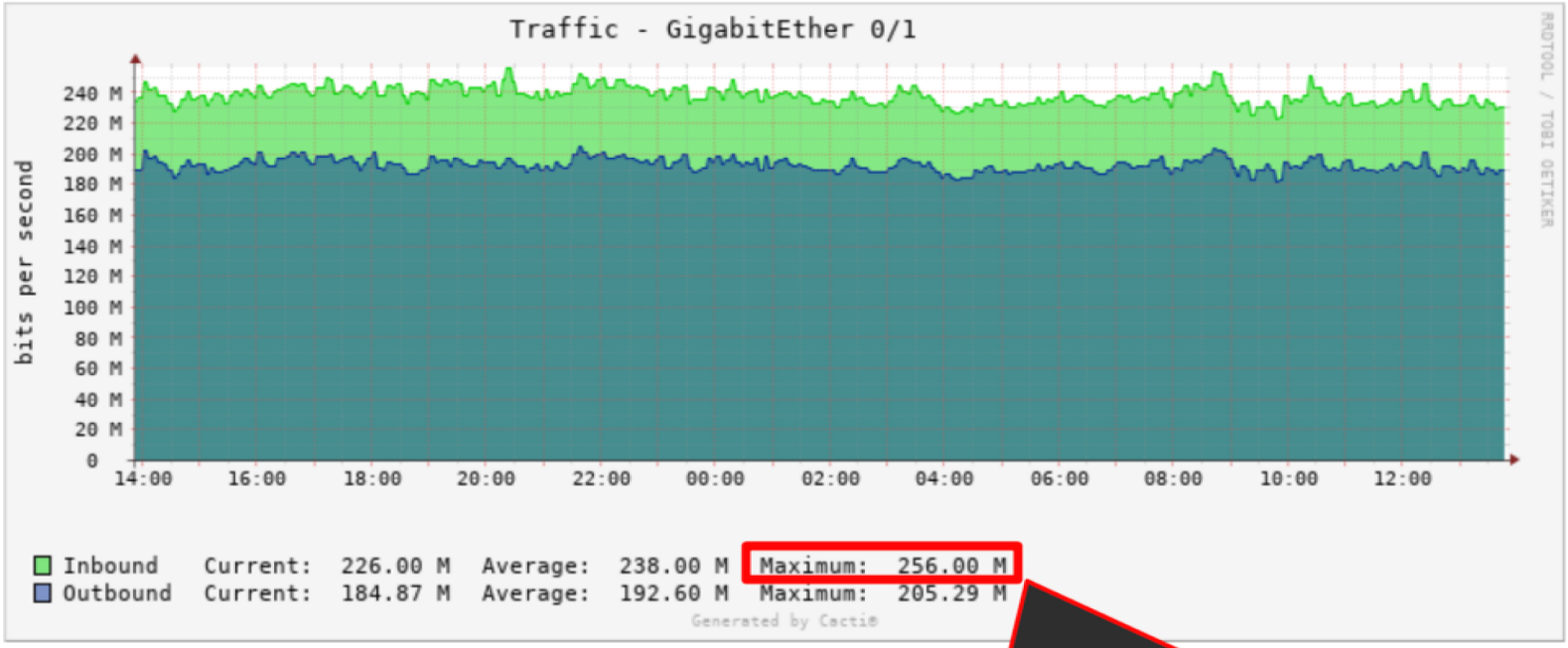
Choosing upstream clock source

- Choose appropriate clock source
  - ATOM clock, Radio, GPS
  - National NTP services
  - Public NTP services

- Reviewing your setting is important
  - Fukuoka Univ. started their public NTP service in 1993, and this gets too popular somehow
  - Fukuoka Univ. is trying to terminate the service now

NTP Traffic

- Only NTP



Maximum: 256 Mbps
340,000 Packet / s !!

人をつくり、時代を拓く。
福岡大學

https://www.lacnic.net/innovaportal/file/4016/1/fukuoka-university-public-ntp-service-deployment-use-case.pdf

Why is it so popular in the world?

- written in manual as setting example
  - Network devices such as L2, L3 switch
  - Multifunction device, etc.

Example

Configure the system time mode as NTP, the time zone is UTC-12:00, the primary NTP server is 133.100.9.2 and the secondary NTP server is 139.78.100.163, the fetching-rate is 11 hours:

TL-SG3424(config)# system-time ntp UTC-12:00 133.100.9.2 139.79.100.163 11

11

https://www.lacnic.net/innovaportal/file/4016/1/fukuoka-university-public-ntp-service-deployment-use-case.pdf

Why is it so popular? (2)

- It's embedded as default setting
- But this is just the tip of the iceberg. We believe that this is not the only reason for the increase in traffic.

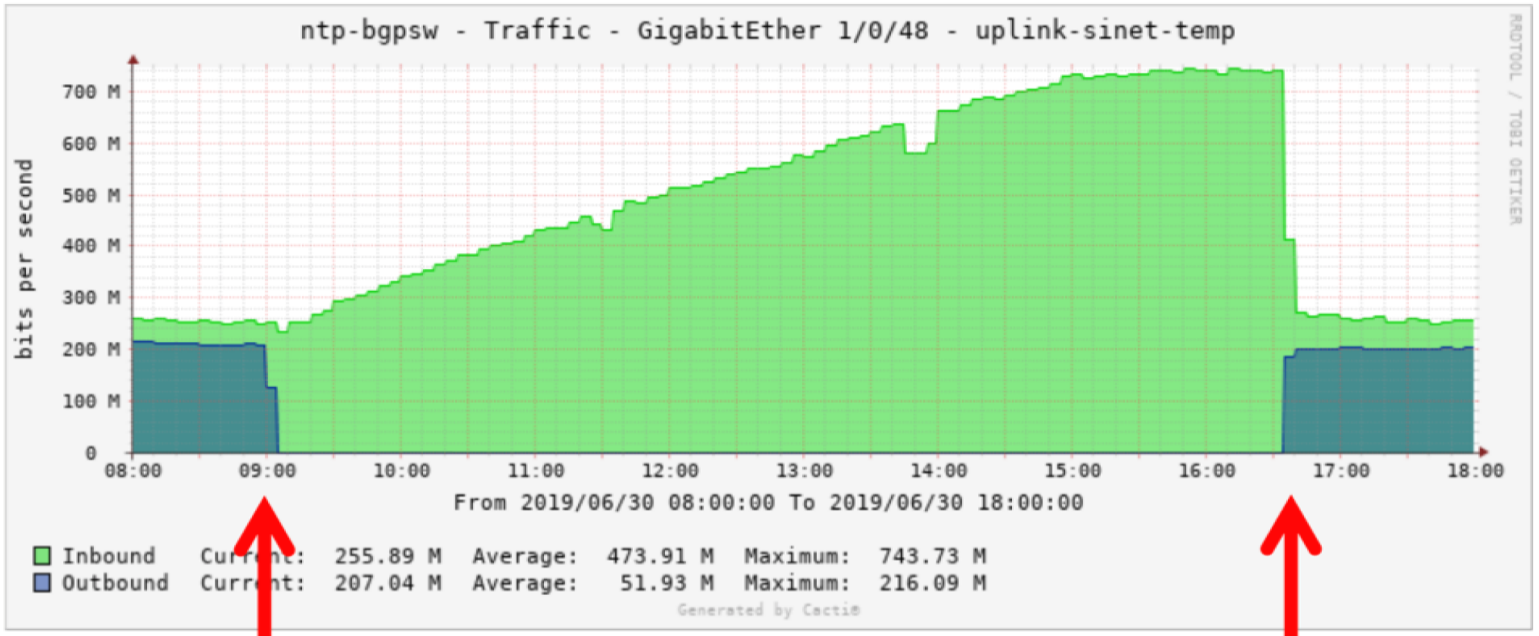
```
    93 77.444013   192.168.2.2      133.100.9.2      NTP      90 NTP Version 3, client
    94 77.658785   133.100.9.2      192.168.2.2      NTP      90 NTP Version 3, server
    95 88.761313   192.168.2.2      192.168.2.1      DNS      78 Standard query 0x04d2 A
    96 88.762061   192.168.2.1      192.168.2.2      DNS      94 Standard query respons
▶ Frame 93: 90 bytes on wire (720 bits), 90 bytes captured (720 bits) on interface 0
▶ Ethernet II, Src: Tp-LinkT_ae:ee:53 (30:b5:c2:ae:ee:53), Dst: MS-NLB-PhysServer-32_05:4b:2d:72:64
▶ Internet Protocol Version 4, Src: 192.168.2.2, Dst: 133.100.9.2
▶ User Datagram Protocol, Src Port: 42336 (42336), Dst Port: 123 (123)
▼ Network Time Protocol (NTP Version 3, client)
  ▶ Flags: 0x1b, Leap Indicator: no warning, Version number: NTP Version 3, Mode: client
    Peer Clock Stratum: unspecified or invalid (0)
    Peer Polling Interval: 4 (16 sec)
    Peer Clock Precision: 0.015625 sec
    Root Delay:    1.0000 sec
    Root Dispersion:    1.0000 sec
    Reference ID: NULL
    Reference Timestamp: Jan  1, 1970 00:00:00.000000000 UTC
    Origin Timestamp: Jan  1, 1970 00:00:00.000000000 UTC
    Receive Timestamp: Jan  1, 1970 00:00:00.000000000 UTC
    Transmit Timestamp: Jan  1, 2014 00:01:16.005072000 UTC
```

人をつくり、時代を拓く。
福岡大学

https://www.lacnic.net/innovaportal/file/4016/1/fukuoka-university-public-ntp-service-deployment-use-case.pdf

Results of Experiment

- Increased to 740Mbps, 950,000pps
- Increase in bandwidth approx. 500Mbps, 500,000pps



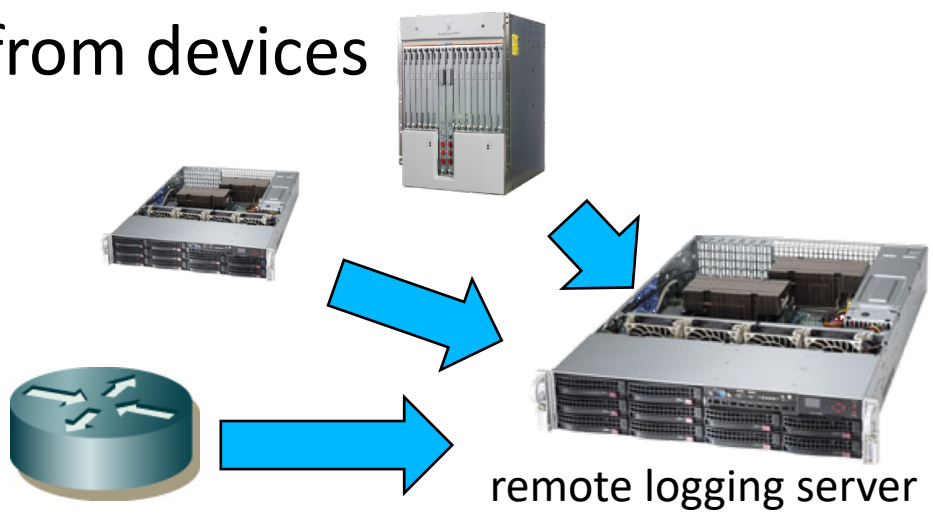
Service stopped

Service resumed

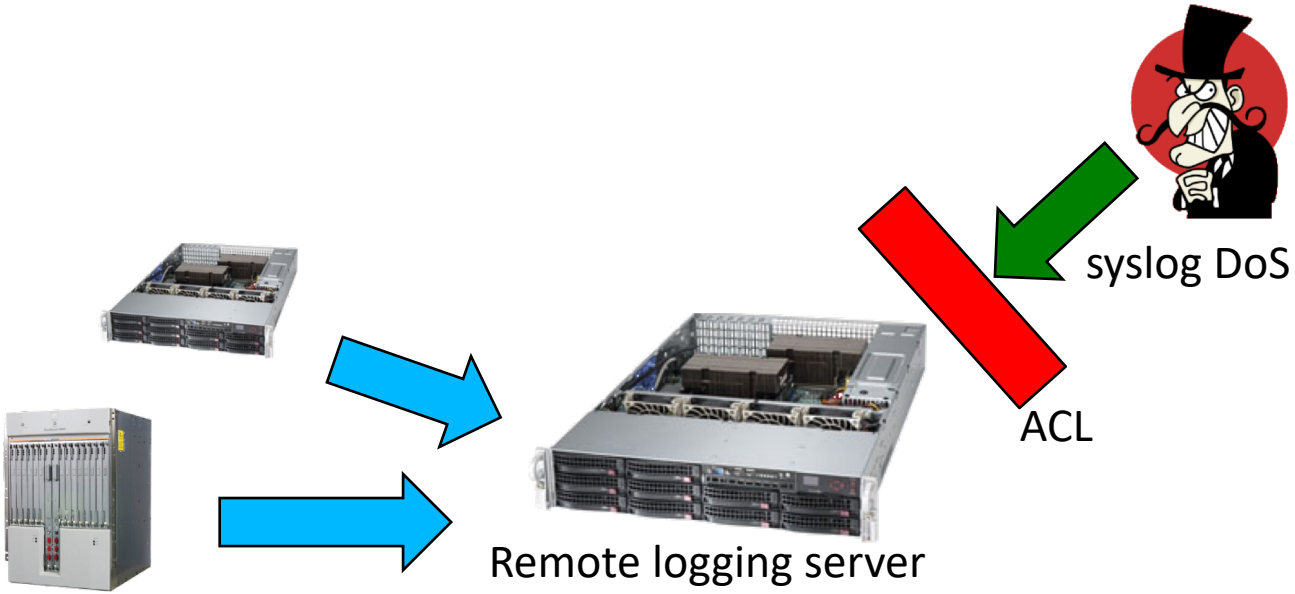
https://www.lacnic.net/innovaportal/file/4016/1/fukuoka-university-public-ntp-service-deployment-use-case.pdf

Remote logging

- Log messages could be modified/deleted
  - If the system is compromised
  - Limited memory buffered log messages
- Remote logging server
  - Receive log messages from devices
  - Syslog-ng
  - Enough storage there



Protecting syslog

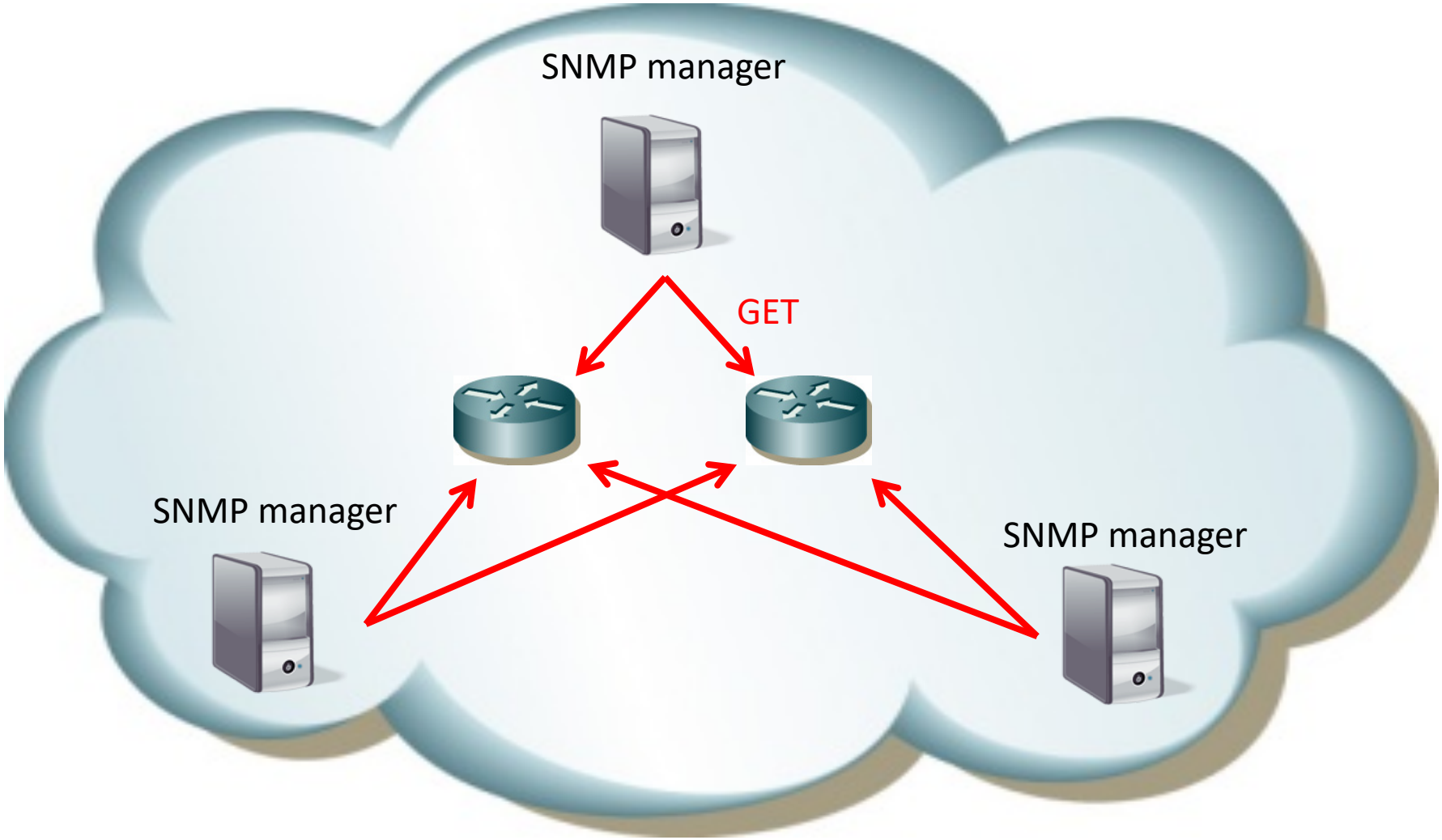


Ensure the correctness of log entries

snmp

- can read/write information and send a trap
  - use version 3, and set password
  - prevent ‘write’ function, or just disable it on agents
  - put ACL to prevent unauthorized access
- require a little disk space on snmp manager
  - useful to check long-term trend

snmp monitoring system



snmp MIB

- Management information base
  - MIB-2, IF-MIB, vender-specific MIB
  - you can get information if an agent supports the MIB you want
- you can specify the information by OIDs
  - ifHCinOctets = .1.3.6.1.2.1.31.1.1.1.6
  - ifHCOutOctets = .1.3.6.1.2.1.31.1.1.1.10

snmp counters

- Frequency of updating counters
  - depends on agents (0-30sec)
  - 5min is widely used as snmp polling time
- Counter overflow
  - 32bit counters(ifIn/OutOctets) could wrap in 5.7min at 100Mbps
  - consider 64bit counters(ifHCInOctets) for 1Gbps or more interfaces

Useful information via SNMP MIBs

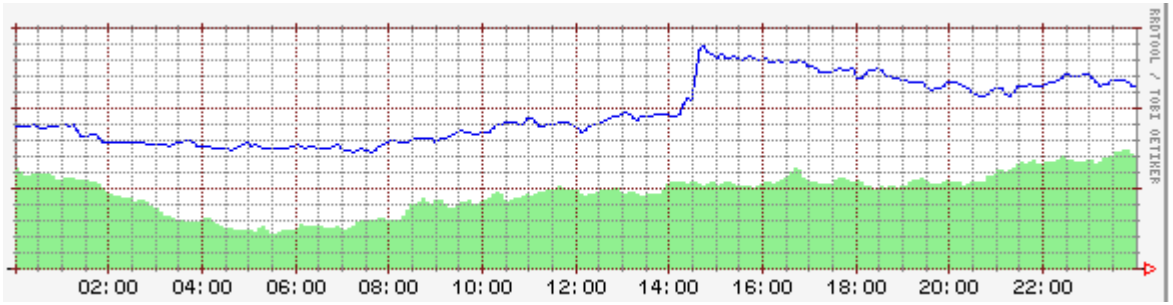
- interface
  - bytes, packets, errors
- system
  - cpu load
  - memory usage
  - temperature
  - icmp, udp
  - ntp

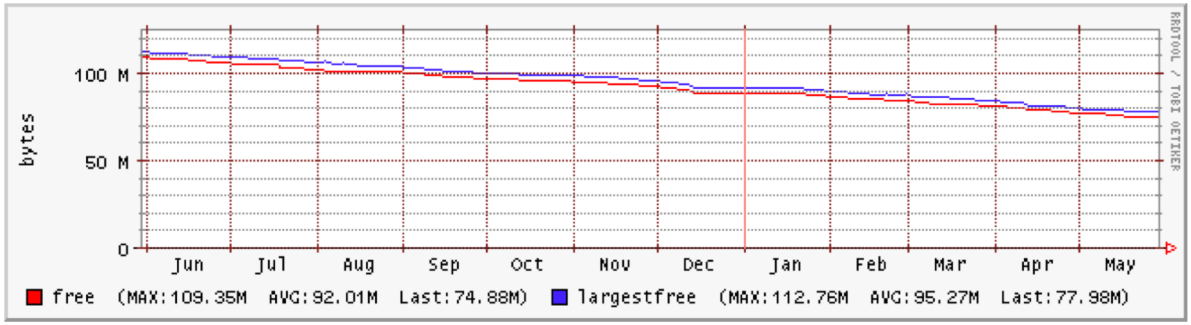
snmp use case

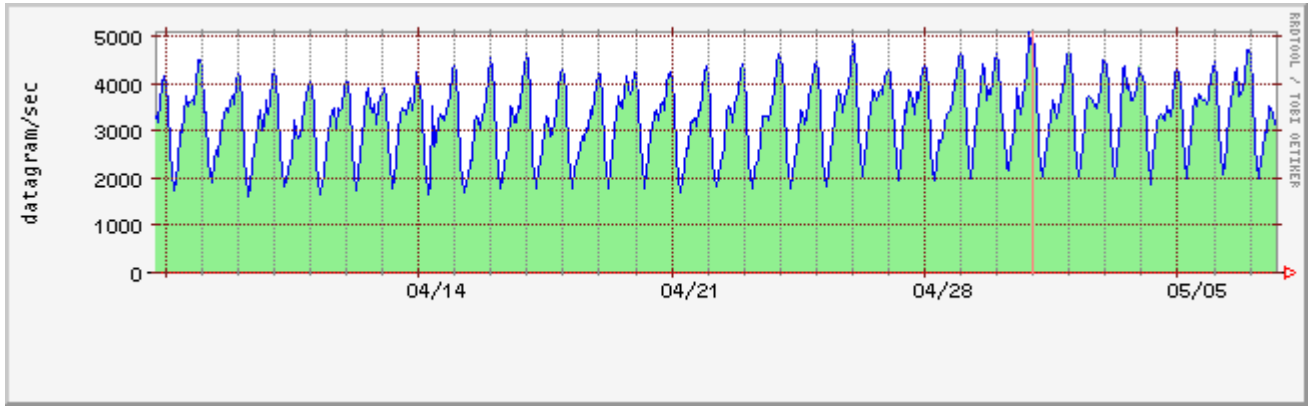
- Usage monitoring
  - bandwidth and traffic volume
- Visualize
  - Observium - https://www.observium.org/
    - LibraNMS - https://www.librenms.org/
  - Cacti - https://www.cacti.net/
  - RRDtool - https://oss.oetiker.ch/rrdtool/

Visualize

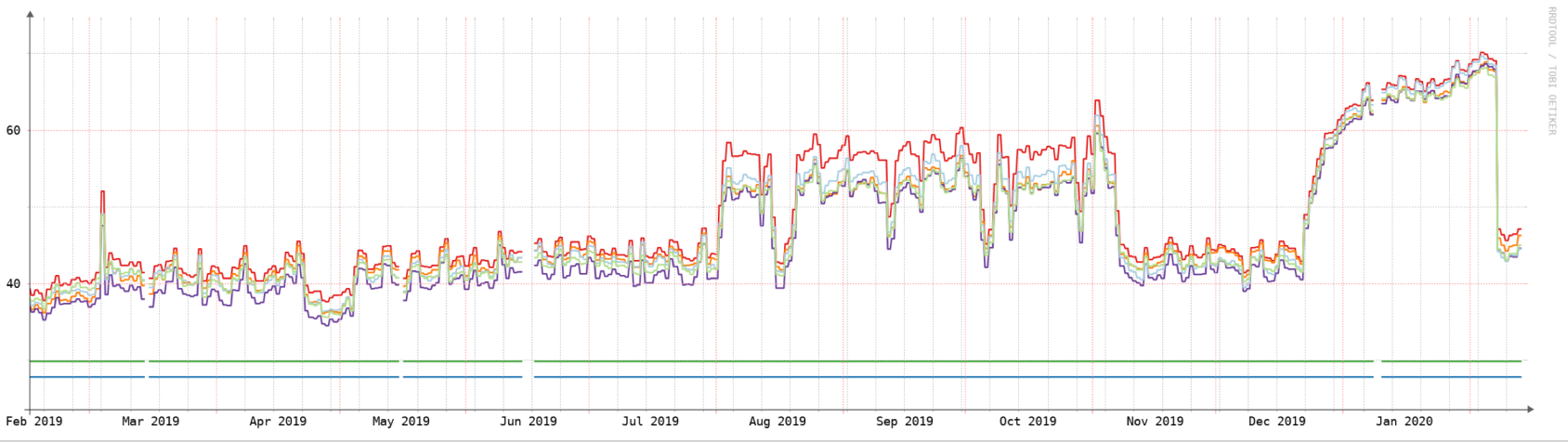
- RRDtool







Example: Temperature



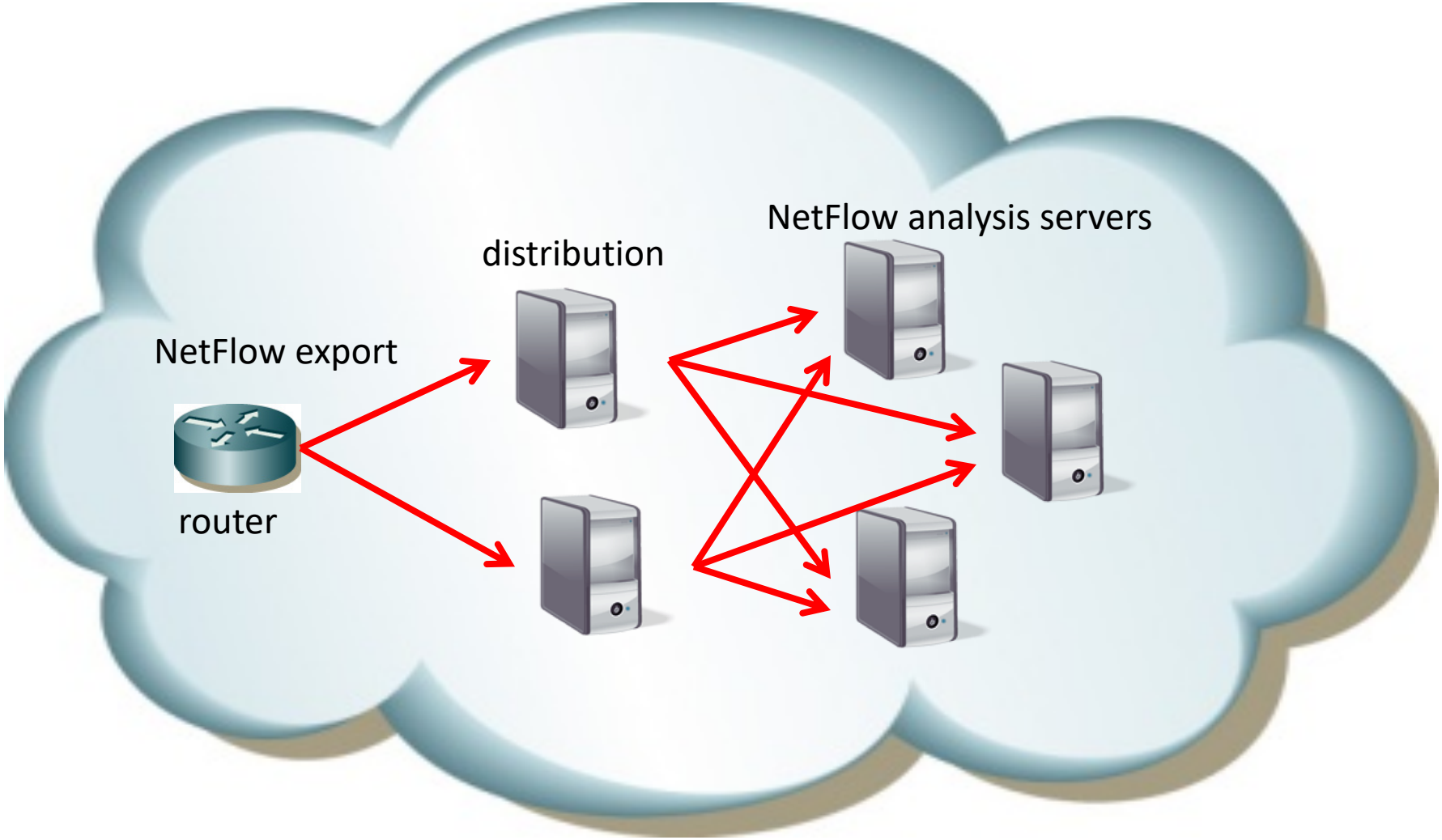
Netflow

- To monitor flow information
  - Packet header
  - Most routers support it
- Require more storage
  - even with sampling, still need to expect huge data
  - not for long term monitoring
- Useful for analysis and anomaly detection

Netflow and sampling

- Sampled netflow is widely used
  - just to know your trend
  - to reduce data
- Margin of error
  - sampled netflow and actual traffic
  - depends on routers
  - worst case: 20%
- IIJ uses magic numbers as sampling rate
  - 1/16382 or 1/8192 where possible

Netflow monitoring system



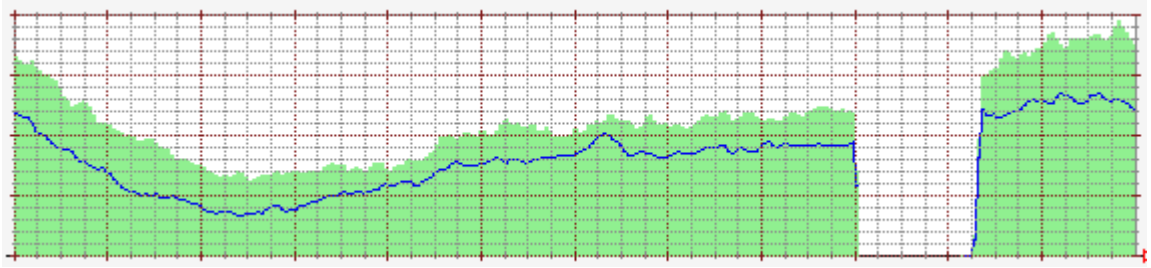
Netflow analysis

- Combination of parameters
  - AS, IP address, protocol, port number
  - too many patterns to pre-generate every graphs
- Visualizing
  - NfSen - http://nfsen.sourceforge.net/

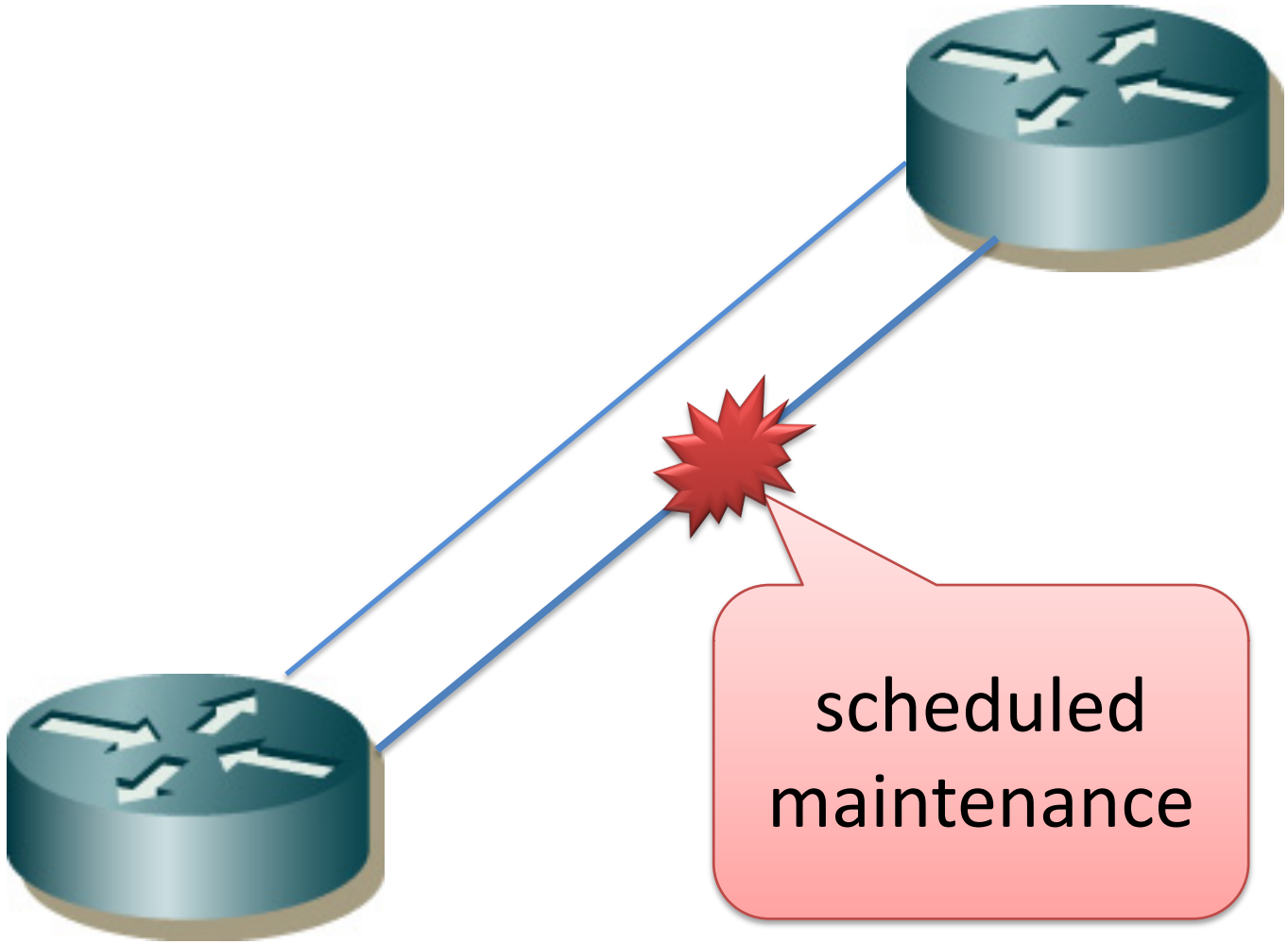
case 1: bps

- traffic was suddenly doubled on a link

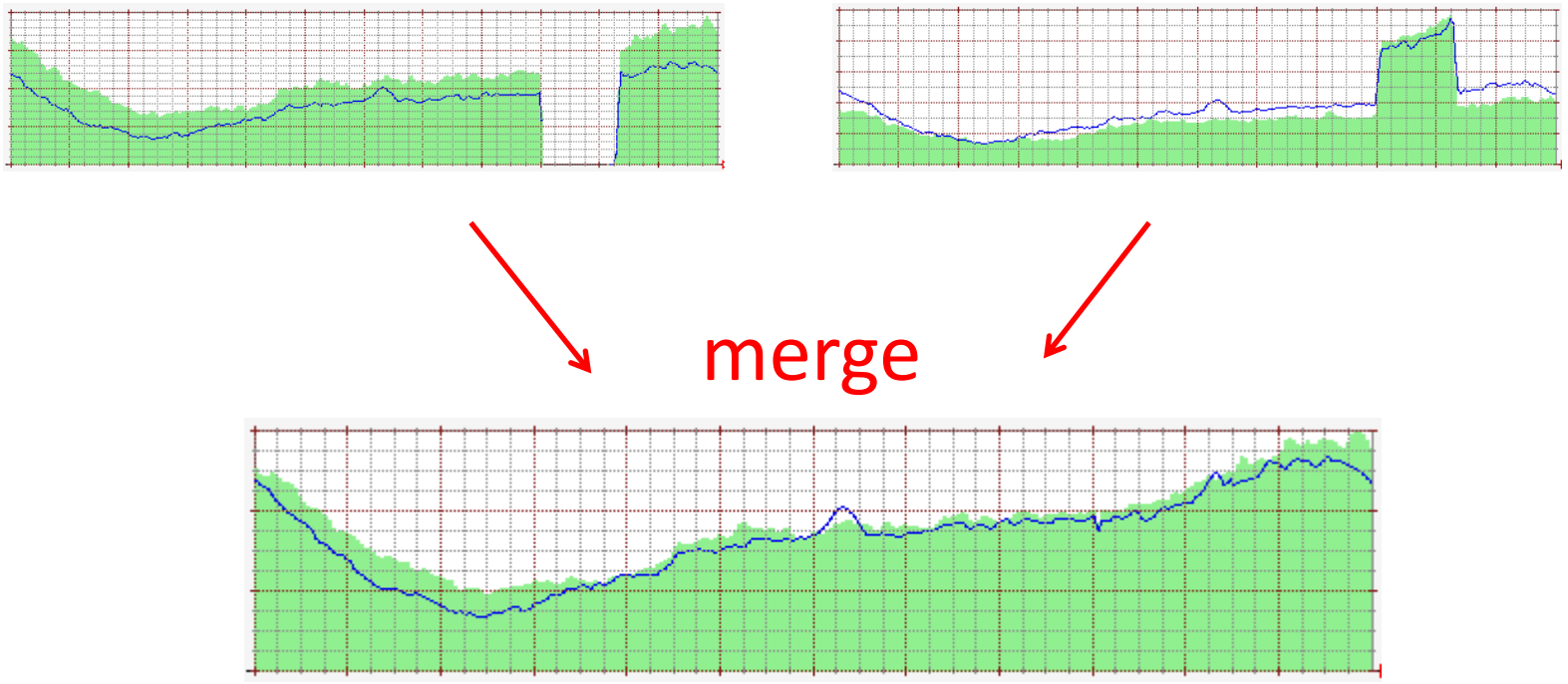
- also found a missing traffic



case 1: 2 links between routers

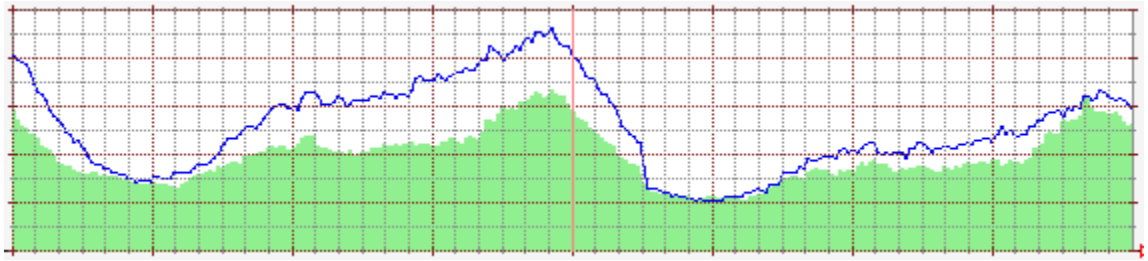


case 1: total traffic: bps

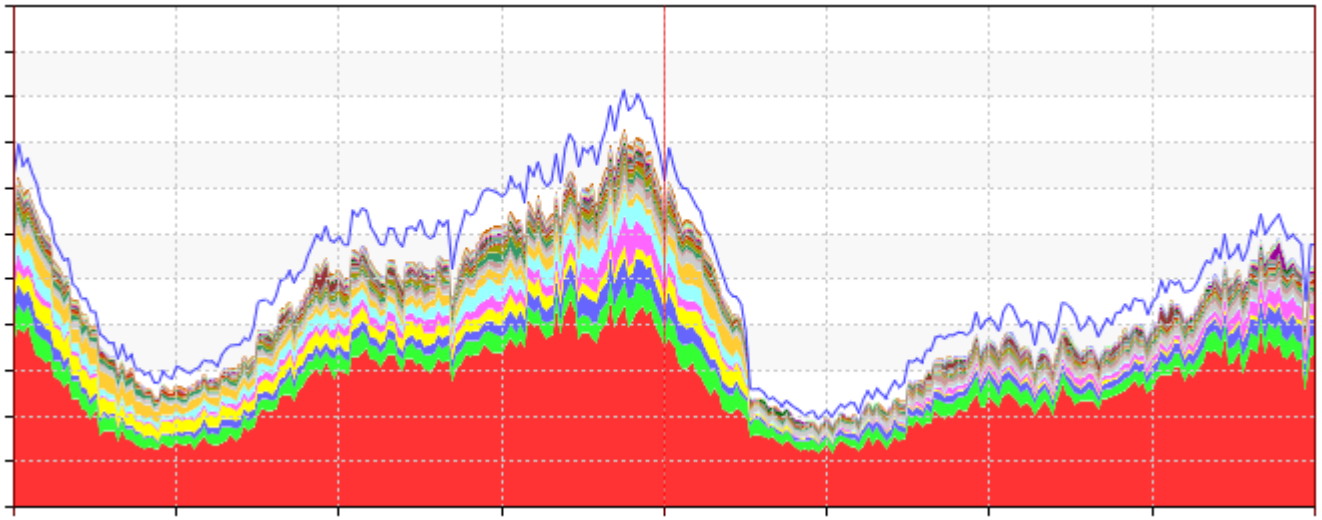


case 2: bps

- traffic decreased
- There is no routing change in the network

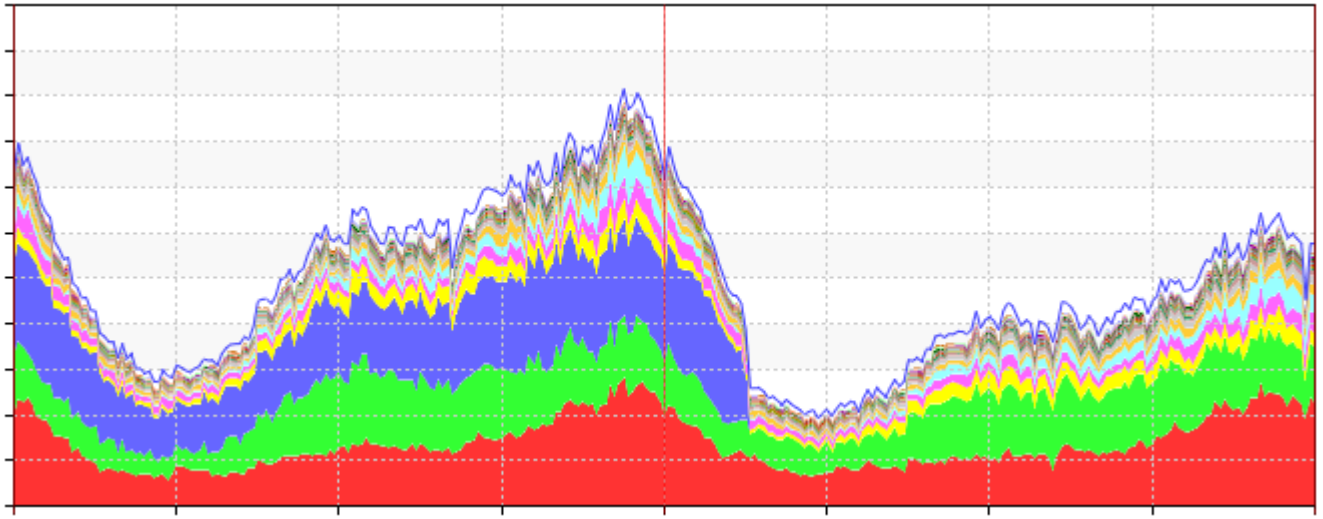


case 2: netflow graph(dst AS)



- the dst AS based graph shows
  - missing traffic to several ASes
  - traffic to the other ASes also a bit decreased

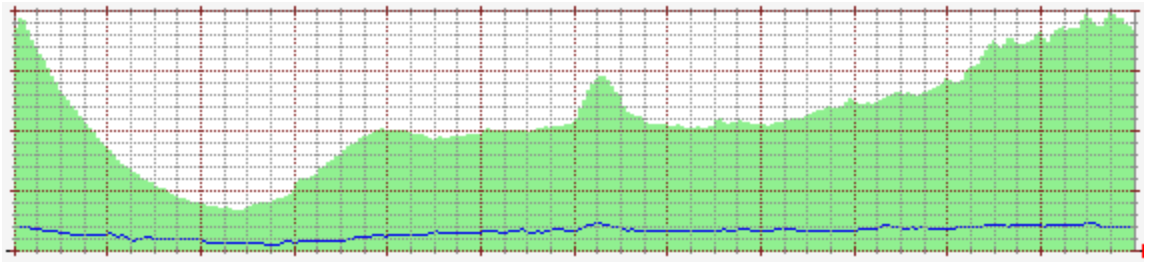
case 2: netflow graph(src AS)



- traffic from a particular AS(blue) was gone
- probably something was happened on the AS(blue)
  - trouble or route change

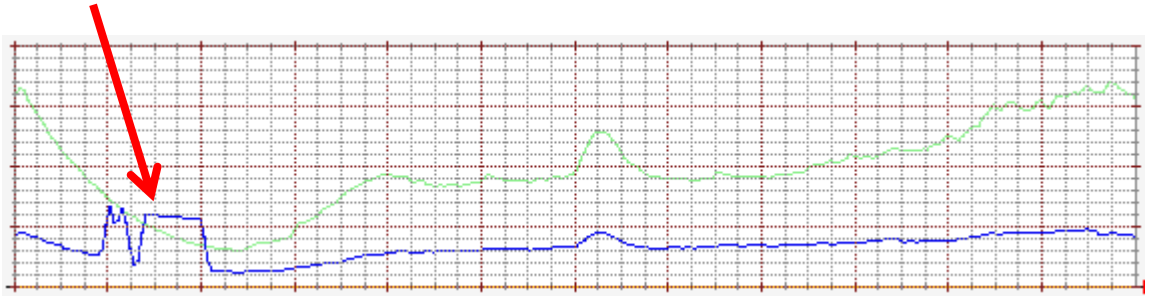
case 3: bps

- traffic looks stable



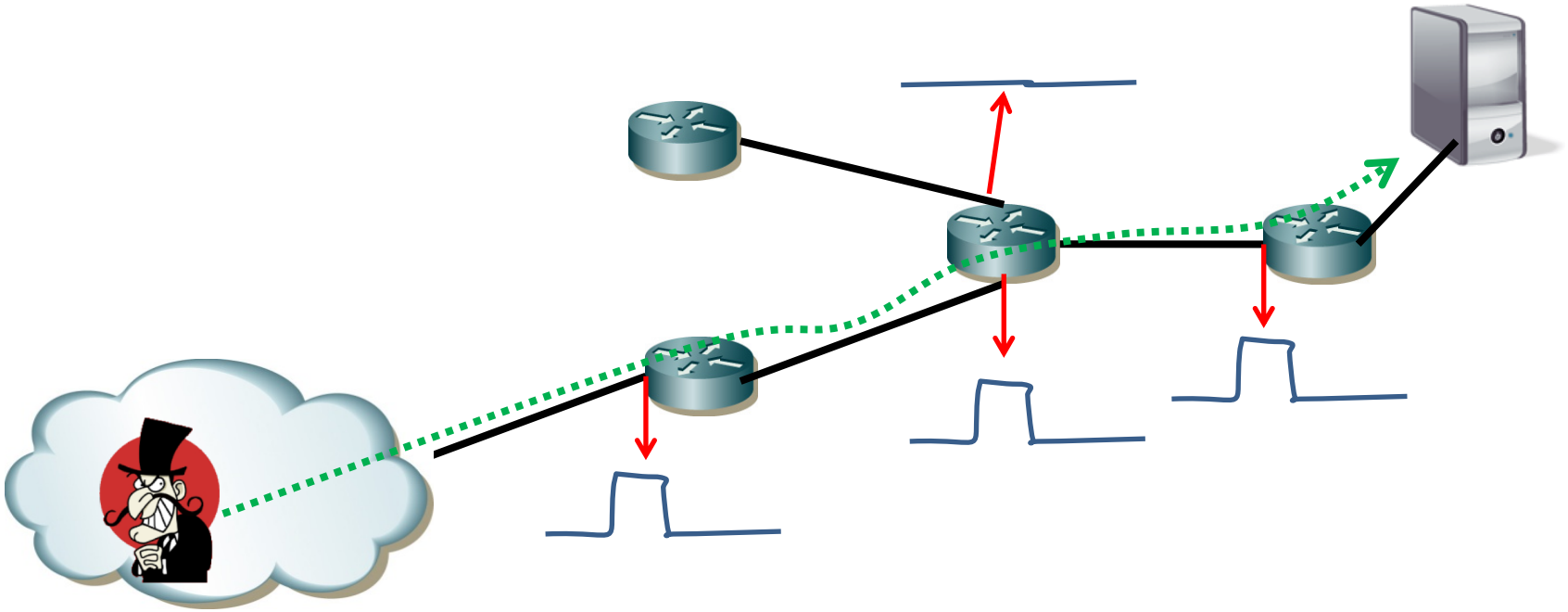
case 3: pps

- pps(packets/sec) graph shows something anomaly

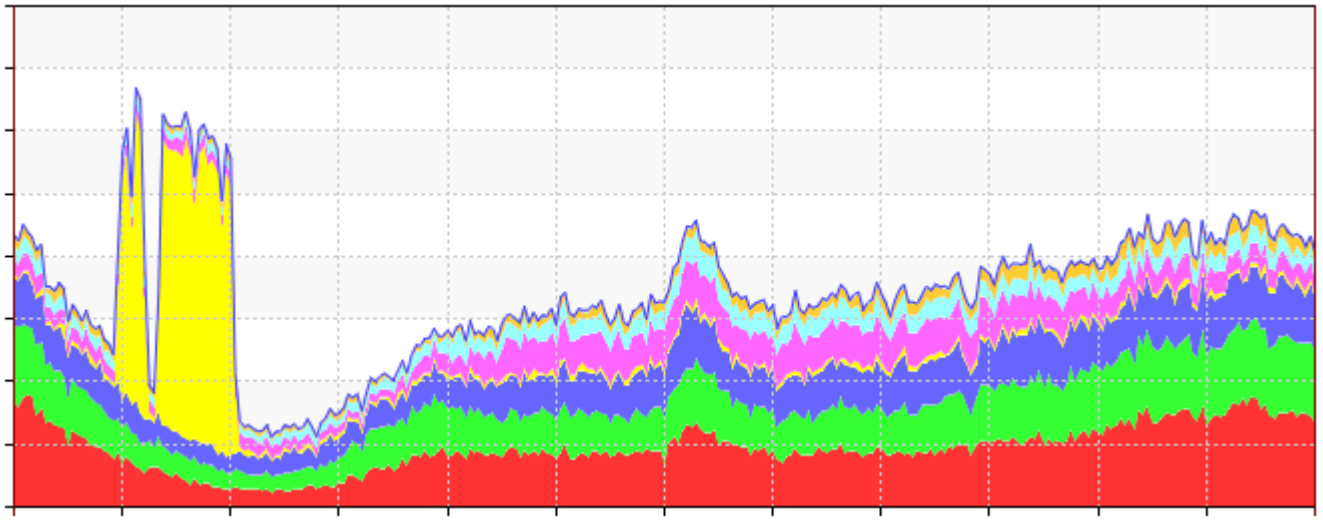


traceback by a shape

- if the traffic pattern is enough characteristic, you can traceback to the inbound interface

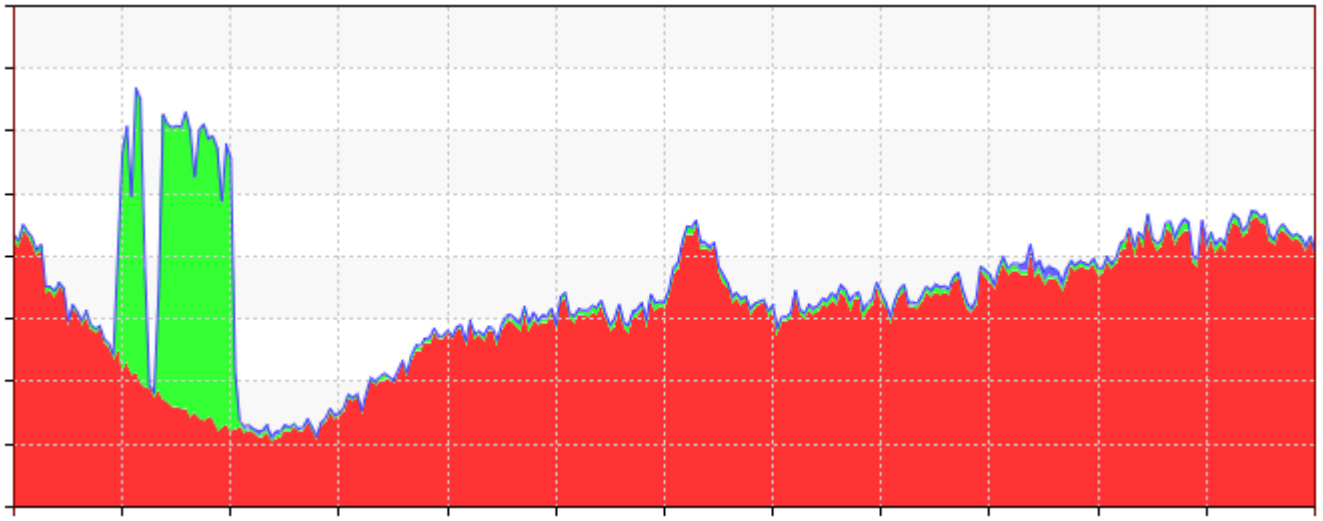


case 3: netflow graph(dst AS, pps)



- according to dst AS based graph, the anomaly traffic was directed to a particular AS(yellow)

case 3: netflow graph(protocol, pps)

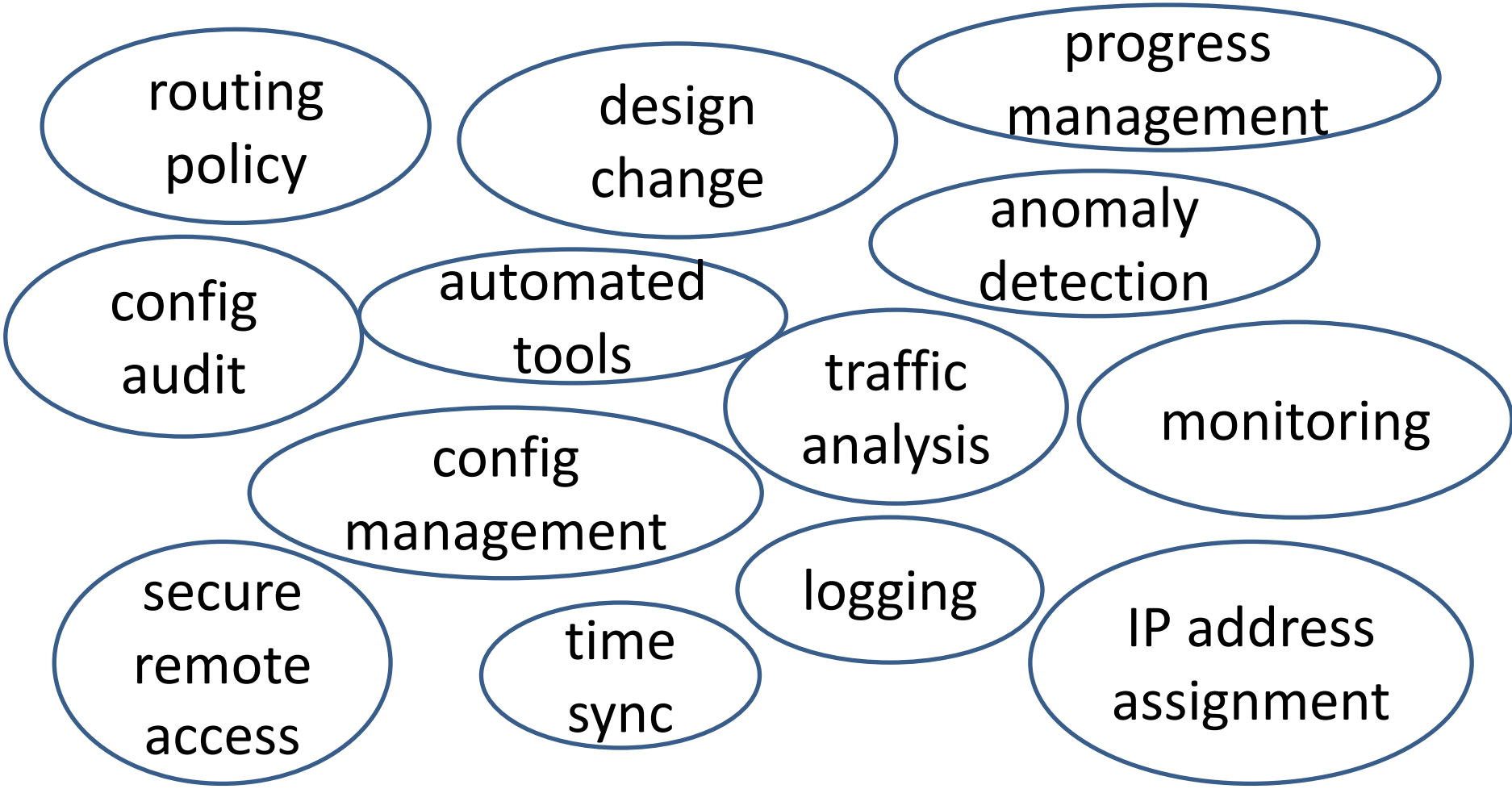


- the traffic profile was mostly UDP

monitoring and detection

- snmp is useful to check
  - trend
  - threshold
- netflow is useful to analysis
  - anomaly
  - change

Operational Design



Be aware of security advisories

DEALING WITH VULNERABILITIES

CVE

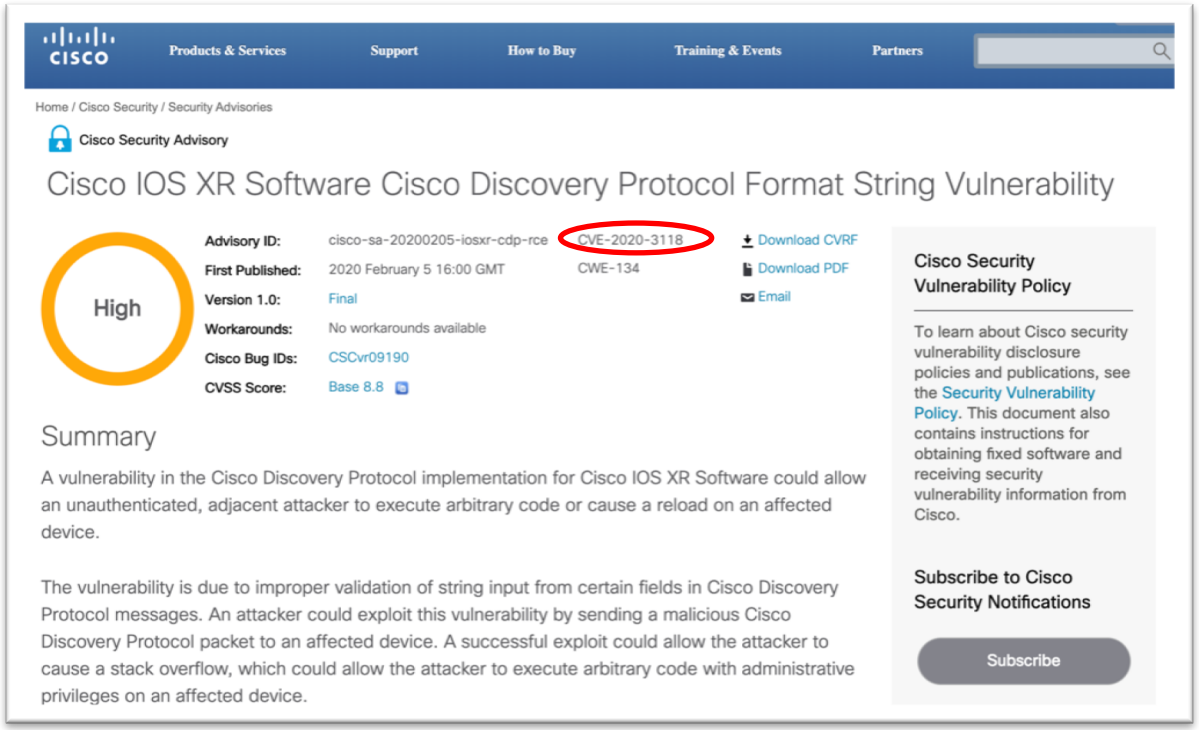
- Common Vulnerabilities and Exposures
- Dictionary of common names (ex. CVE identifiers) for publicly known security vulnerabilities
- https://cve.mitre.org/
  - https://twitter.com/CVEnew
- We can use a common name to specify a security vulnerability

example: CVE-2020-3119

- http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2020-3119
- target
  - CDP enabled IOS-XR devices
- impact
  - Could allow an unauthenticated, adjacent attacker to execute arbitrary code or cause a reload on an affected device

Cisco web site

- https://tools.cisco.com/security/center/content/CiscoSecurityAdvisory/cisco-sa-20200205-iosxr-cdp-rce



CVSS

- Common Vulnerability Scoring System
- https://www.first.org/cvss/
  - CVSSv3 is released in 2015
- An open framework for communicating the characteristics and impact of IT vulnerabilities

CVSS Scores

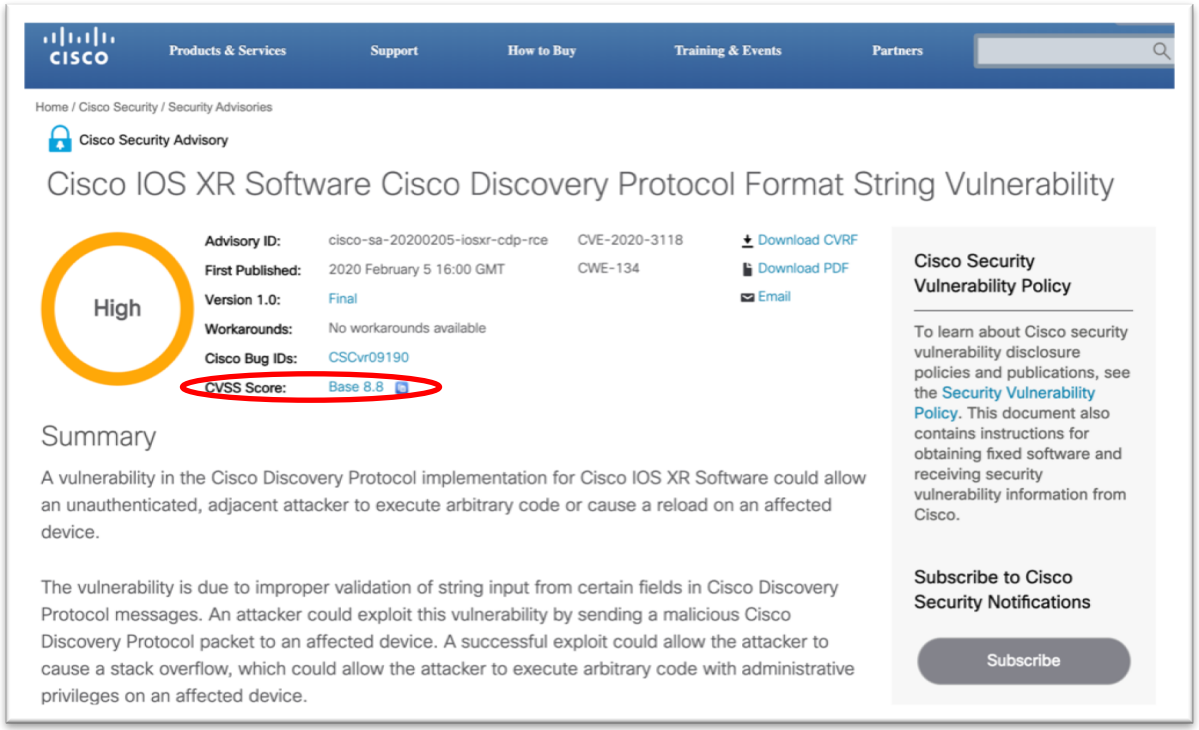
- Base Score
  - technical evaluation
- Temporal Score
  - environmental evaluation
  - proof of concept code/attack code
  - could be changed over the time

CVSS Scores

Security Level	Score
Critical	9 - 10
High	7 - 8.9
Medium	4 - 6.9
Low	0.1 - 3.9
Info	0

CVSS score of the

- https://tools.cisco.com/security/center/content/CiscoSecurityAdvisory/cisco-sa-20200205-iosxr-cdp-rce

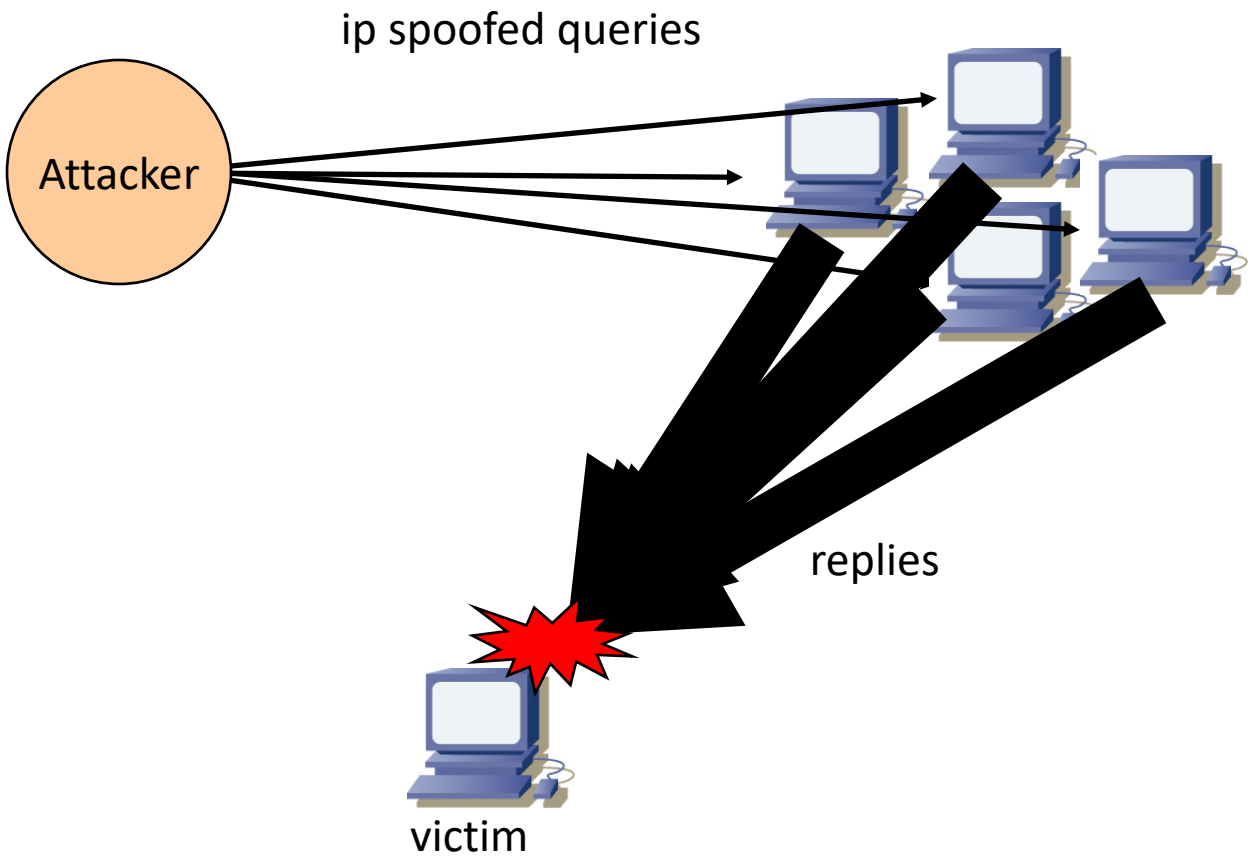


Security Advisories

- Cisco
  - https://tools.cisco.com/security/center/publicationListing.x
- Juniper
  - https://advisory.juniper.net

MITIGATING ATTACKS

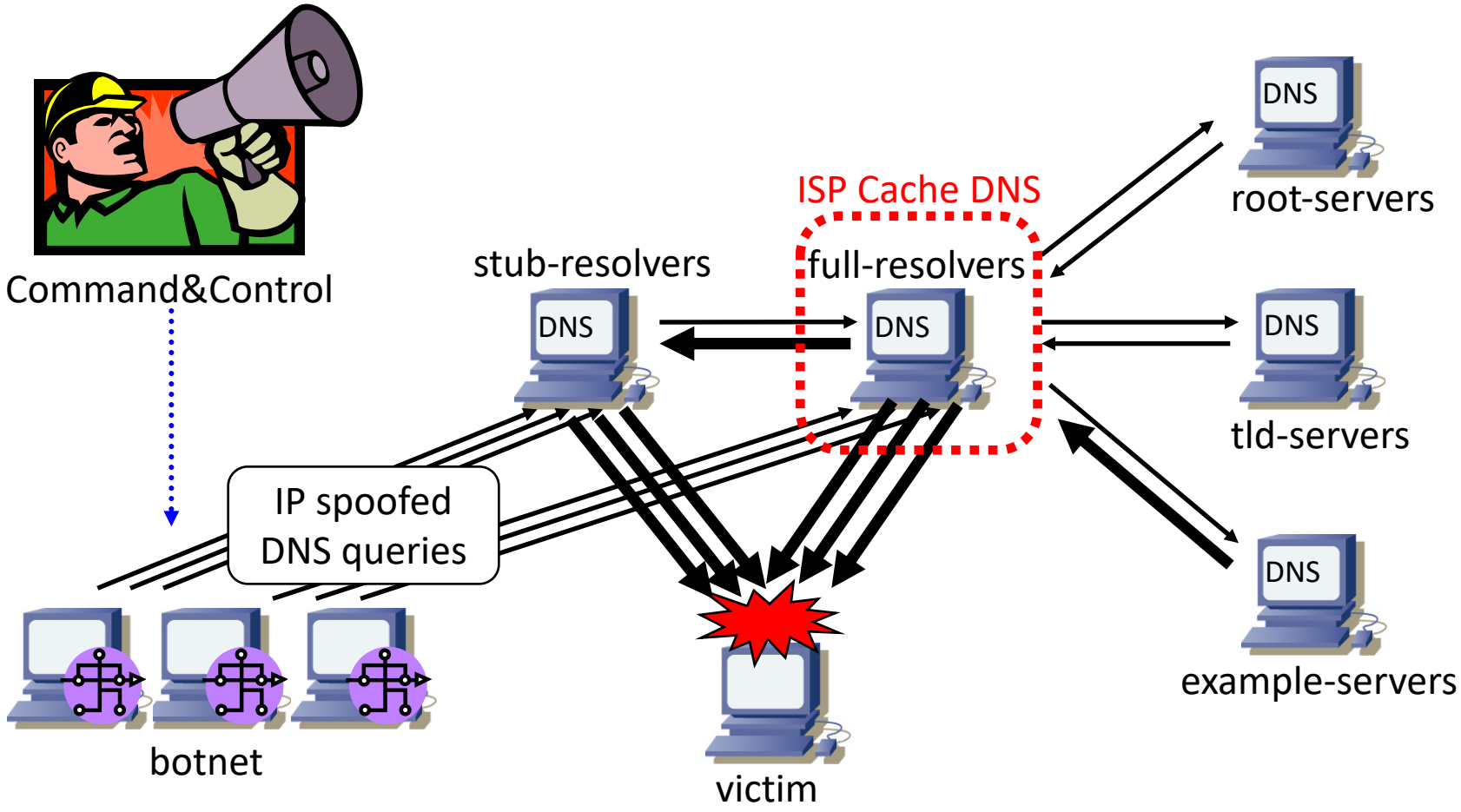
reflection attacks



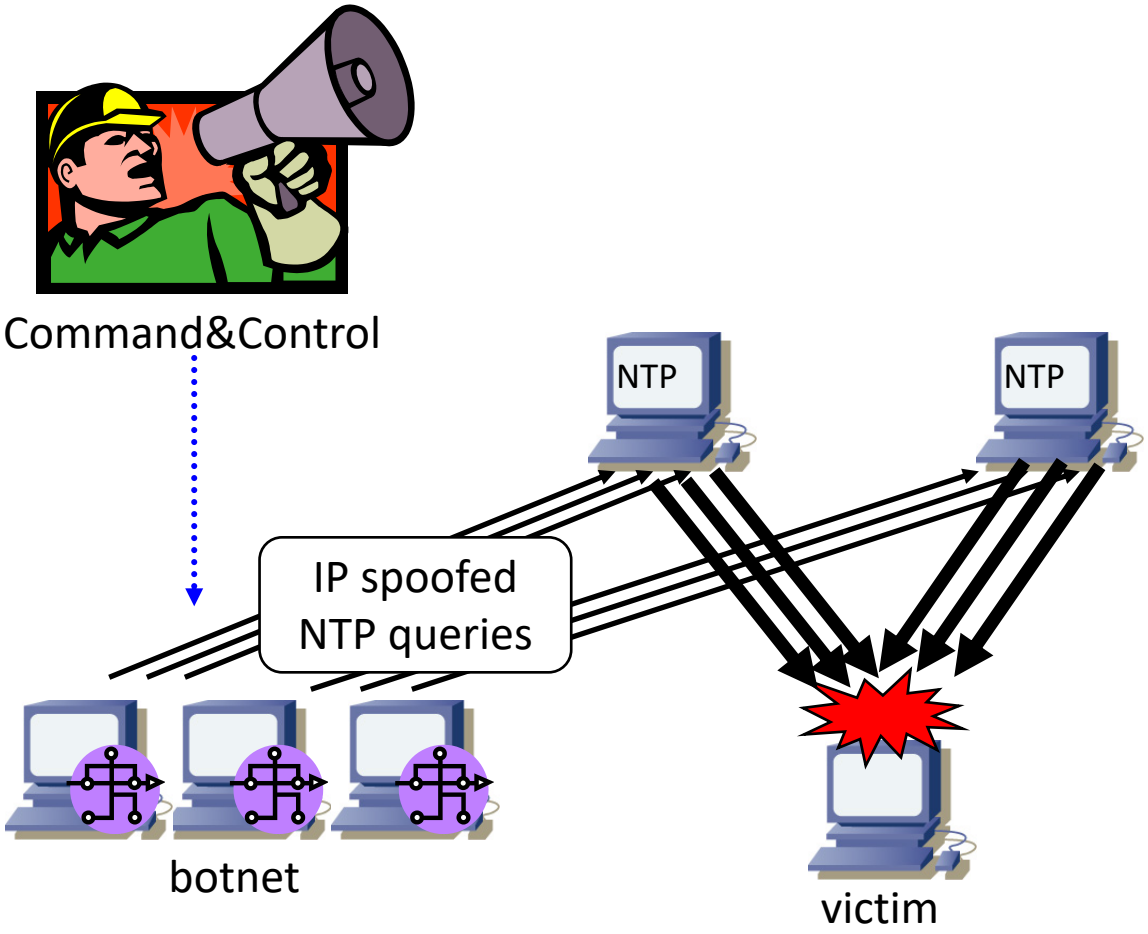
amplifiers

- smurf attack
  - directed broadcast
  - amplification ratio: ~100
- dns amplification attack
  - a huge size record
  - amplification ratio: ~60
- ntp amplification attack
  - monlist query
  - amplification ratio: ~200

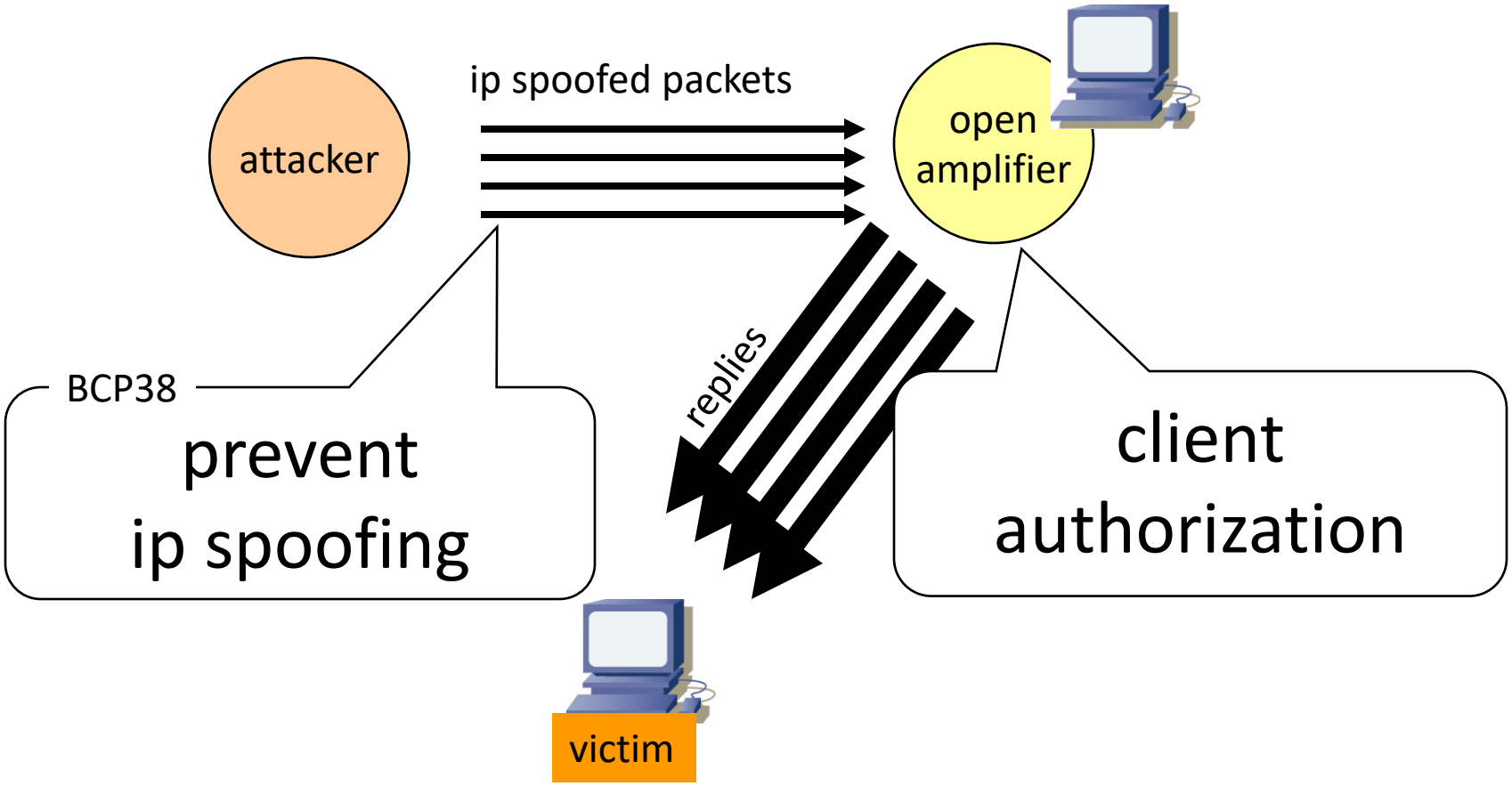
dns amp attack



ntp amp attack



solutions against ip reflection attacks



client authorization

- Incoming interface based
  - useful for home users and enterprises
  - allow from inside, deny from outside
- source IP address based
  - useful for service providers
  - allow from customer network

- you can simply disable the service if it's not necessary

BCP38

- A “Best Current Practice” document of the IETF. BCP38(RFC2827) is intended to limit the impact of DDoS attacks by:
  - Denying traffic with spoofed source address
  - Helping to ensure that traffic is traceable to its correct source network

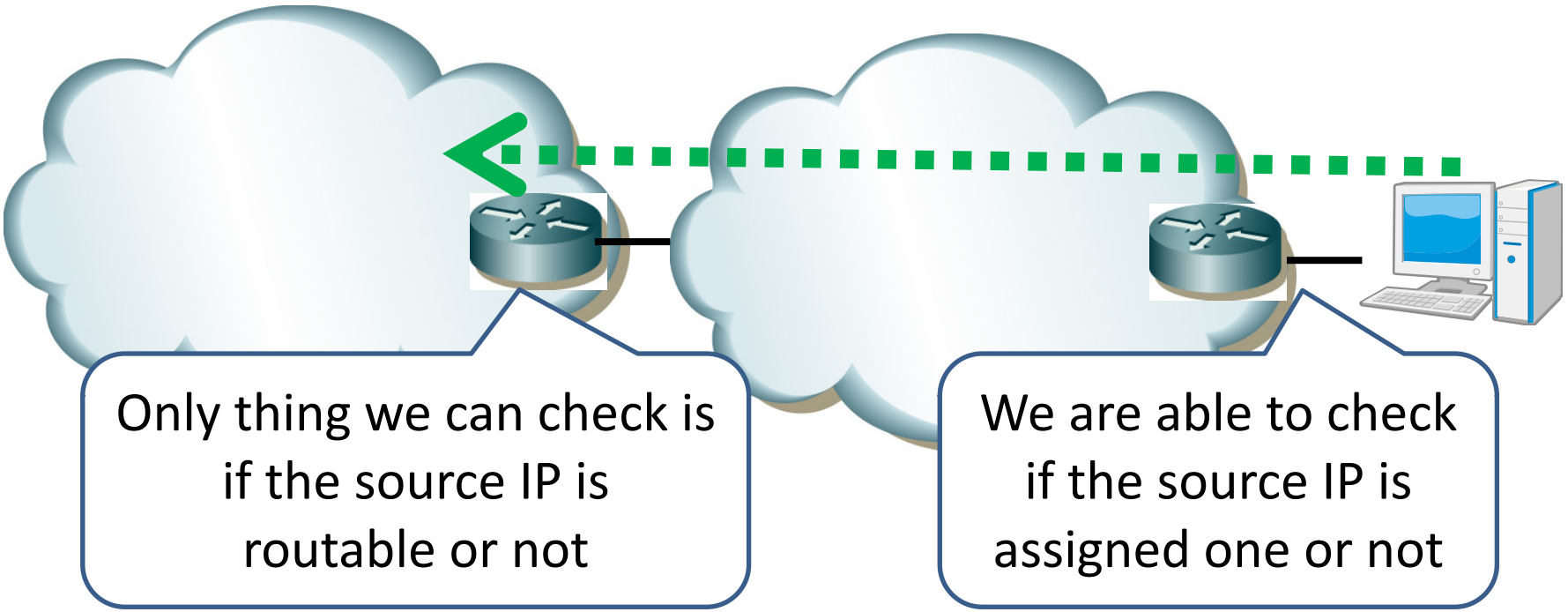
Addressing and Users

- ISP/network administrator assigns IP prefix(es) to their users
  - dynamic or static
  - DHCP, PPP, RA
- Users should use these assigned IP prefixes as their source IP address

BCP38 implementation

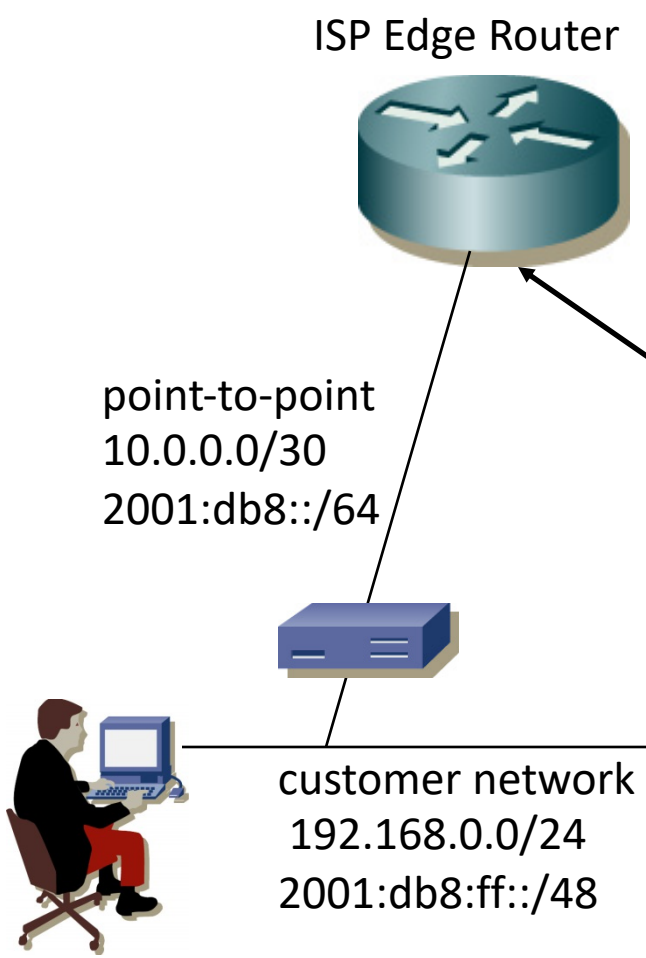
- ACL
  - packet filter
  - permit valid-source, then drop any
- uRPF check
  - checks incoming packets using ‘routing table’
  - look-up a return path for the source IP address
  - loose mode can’t stop most misuse
    - use strict mode

deployment point



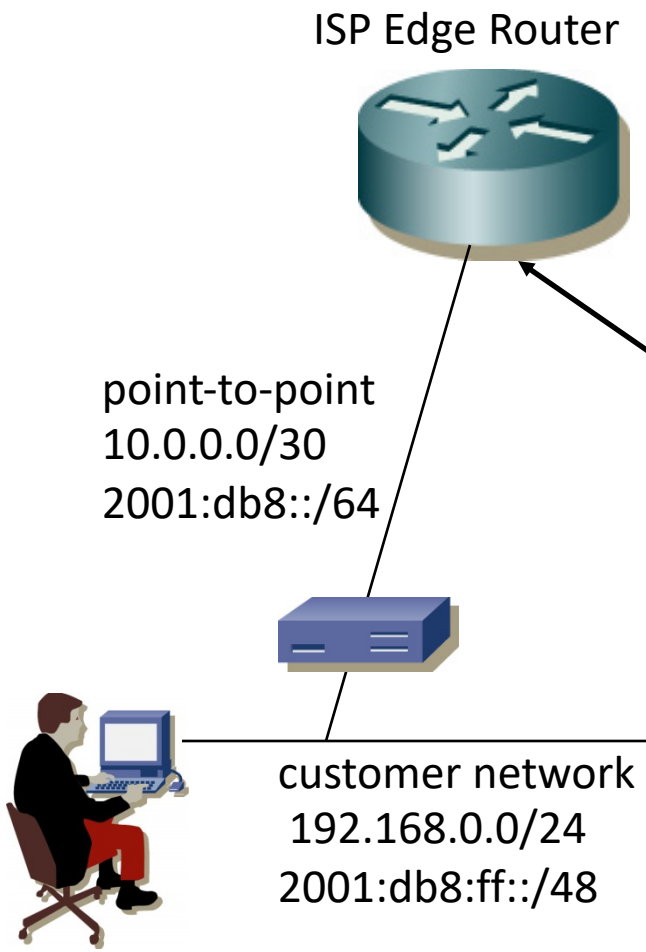
- ISP Edge (customer aggregation) router
  - close to packet source as possible

cisco ACL example



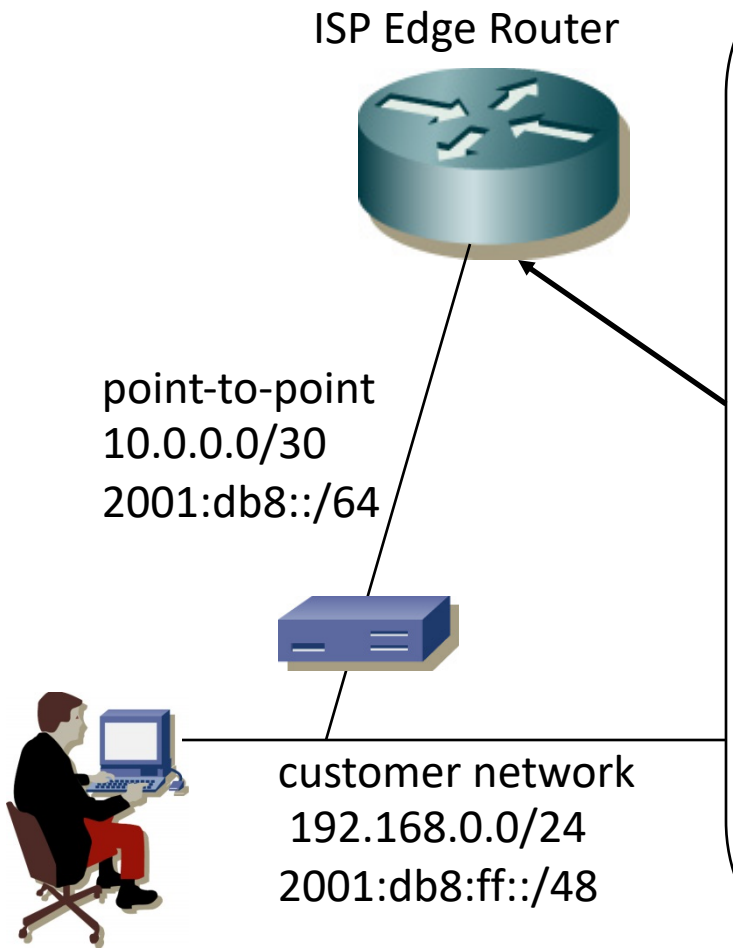

```
ip access-list extended fromCUSTMER4
 permit ip 192.168.0.0 0.0.255.255 any
 permit ip 10.0.0.0 0.0.0.3 any
 deny   ip  any any
!
IPv6 access-list fromCUSTMER6
 permit ipv6 2001:db8::/64 any
 permit ipv6 any 2001:db8::/64 any
 permit ipv6 2001:db8:ff::/48 any
 permit ipv6 fe80::/10 fe80::/10
 permit ipv6 fe80::/10 ff02::/16
 deny ipv6 any any
!
interface Gigabitethernet0/0
 ip access-group fromCUSTOMER4 in
 ipv6 traffic-filter fromCUSTOMER6 in
```

juniper IPv4 ACL example



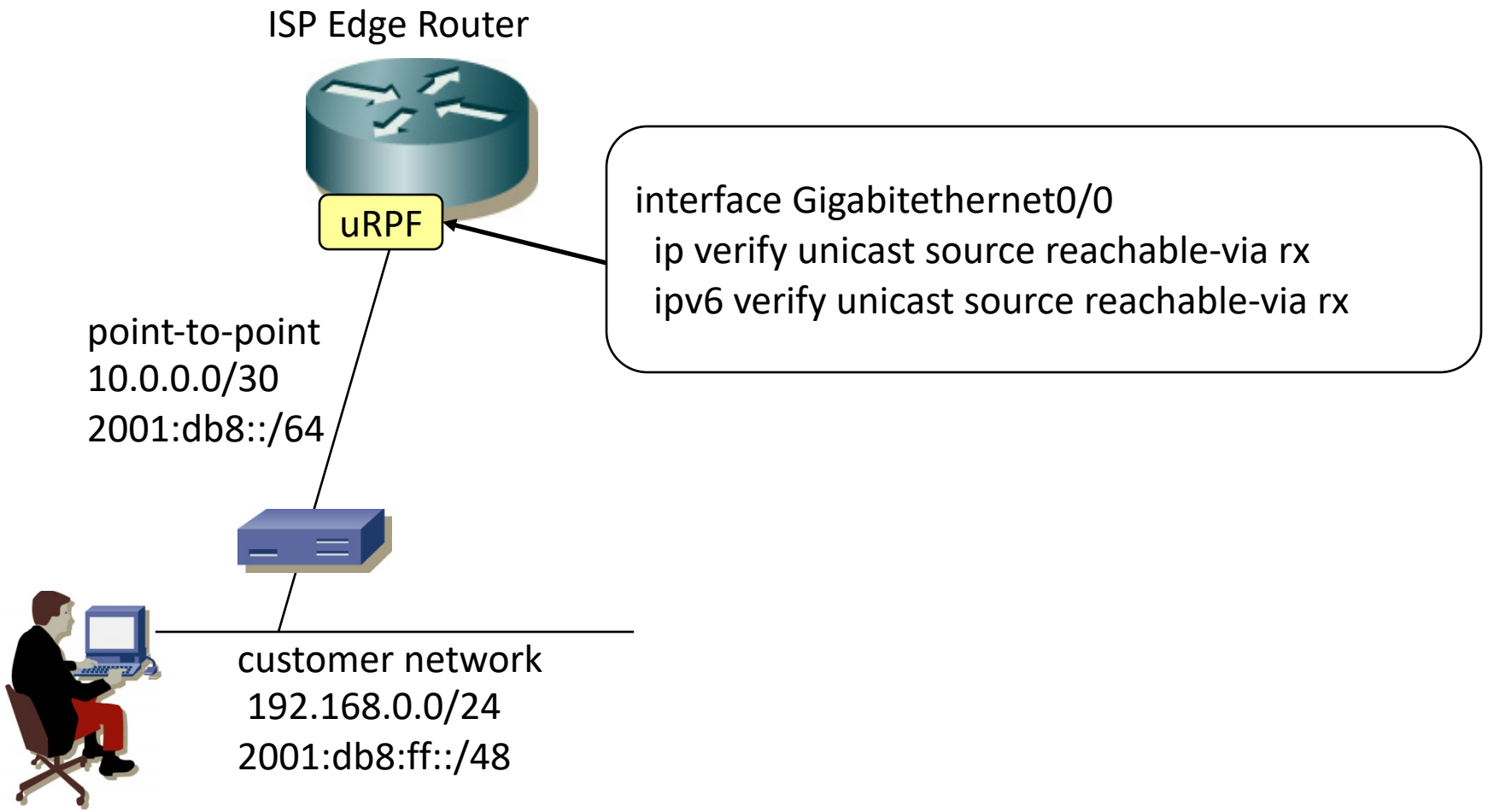

```
firewall family inet {
 filter fromCUSTOMER4 {
  term CUSTOMER4 { from
   source-address {
    192.168.0.0/16;
    10.0.0.0/30;
   }
   then accept;
  }
  term Default {
   then discard;
}}}
[edit interface ge-0/0/0 unit 0 family inet]
filter {
 input fromCUSTOMER;
}
```

juniper IPv6 ACL example

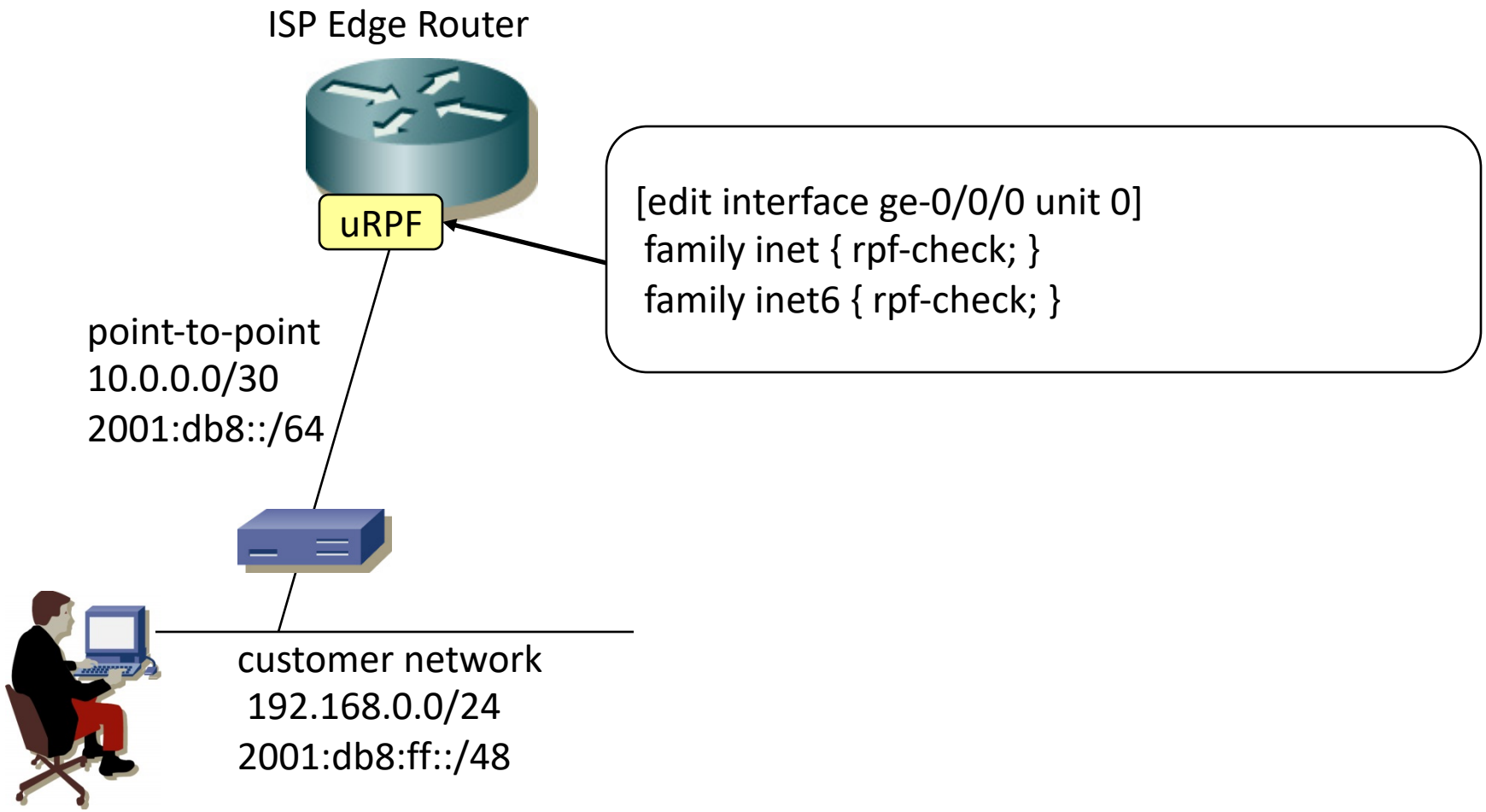



```
firewall family inet6 {
 filter fromCUSTOMER6 {
  term CUSTOMER6 { from
   source-address {
    2001:db8::/64;
    2001:db8:ff::/48;
   }
   then accept;
  }
  term LINKLOCAL { from
    source-address {
      fe80::/10;
    } destination-address {
      fe80::/10;
      ff02::/16;
    }
    then accept;
  }
  term Default {
   then discard;
}}}
[edit interface ge-0/0/0 unit 0 family inet6]
filter {
 input fromCUSTOMER6;
}
```

cisco uRPF example

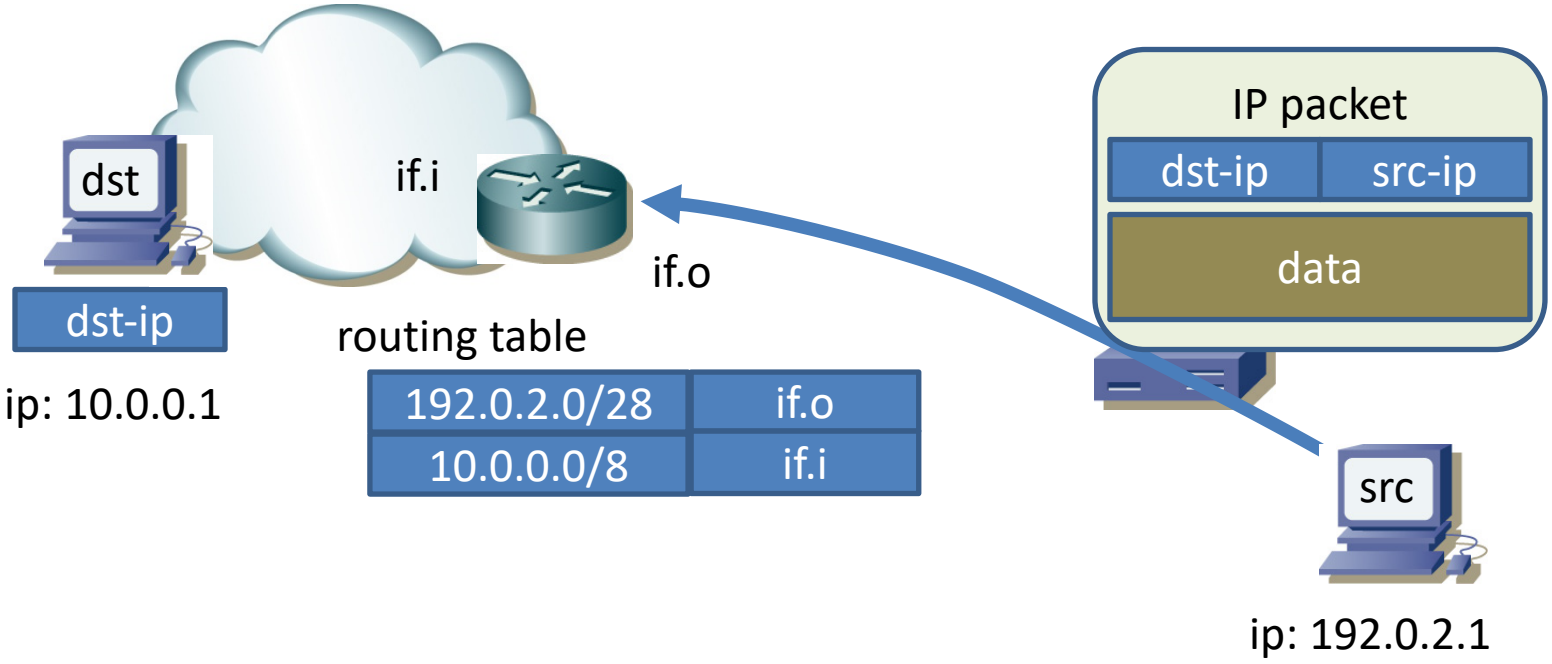


juniper uRPF example



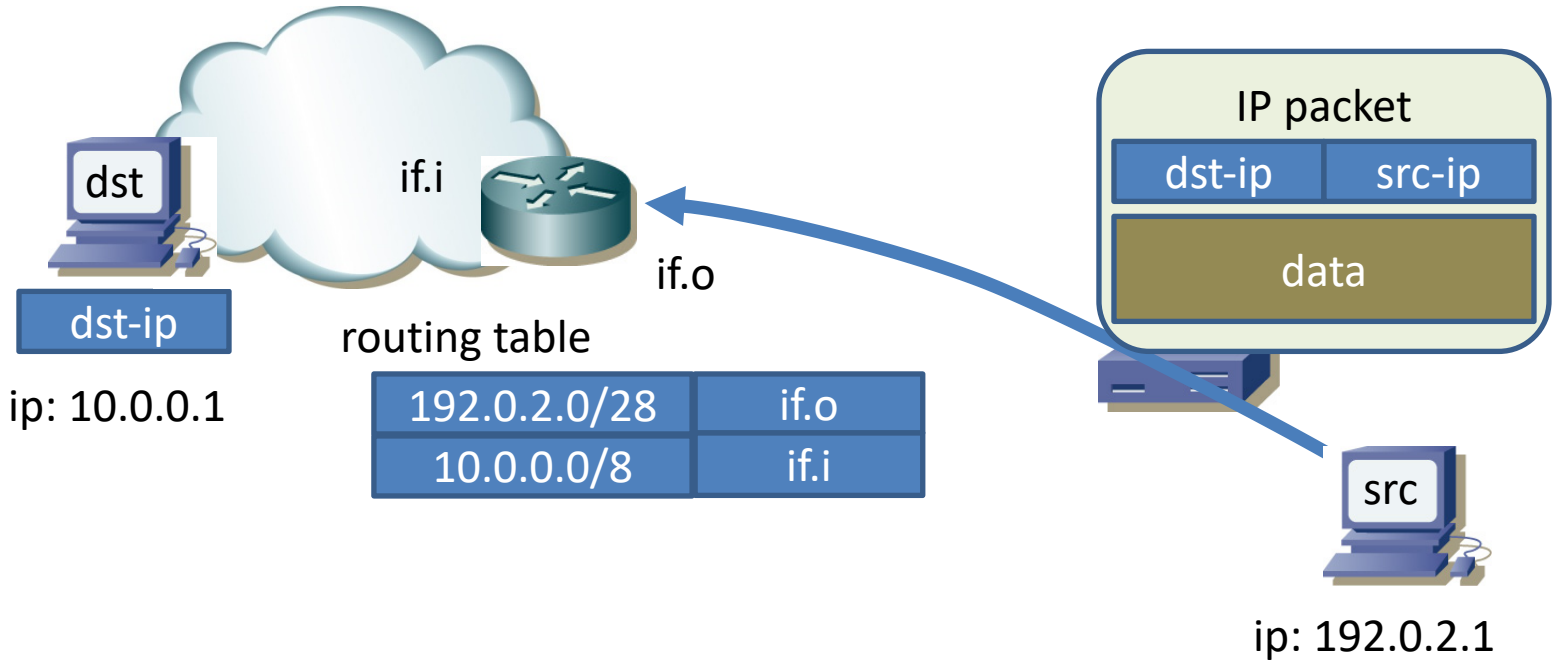
packet forwarding – dst-ip based

- routing_table(dst-ip) => outgoing interface
  - lookup by 10.0.0.1 => if.i
  - then router forwards the packet

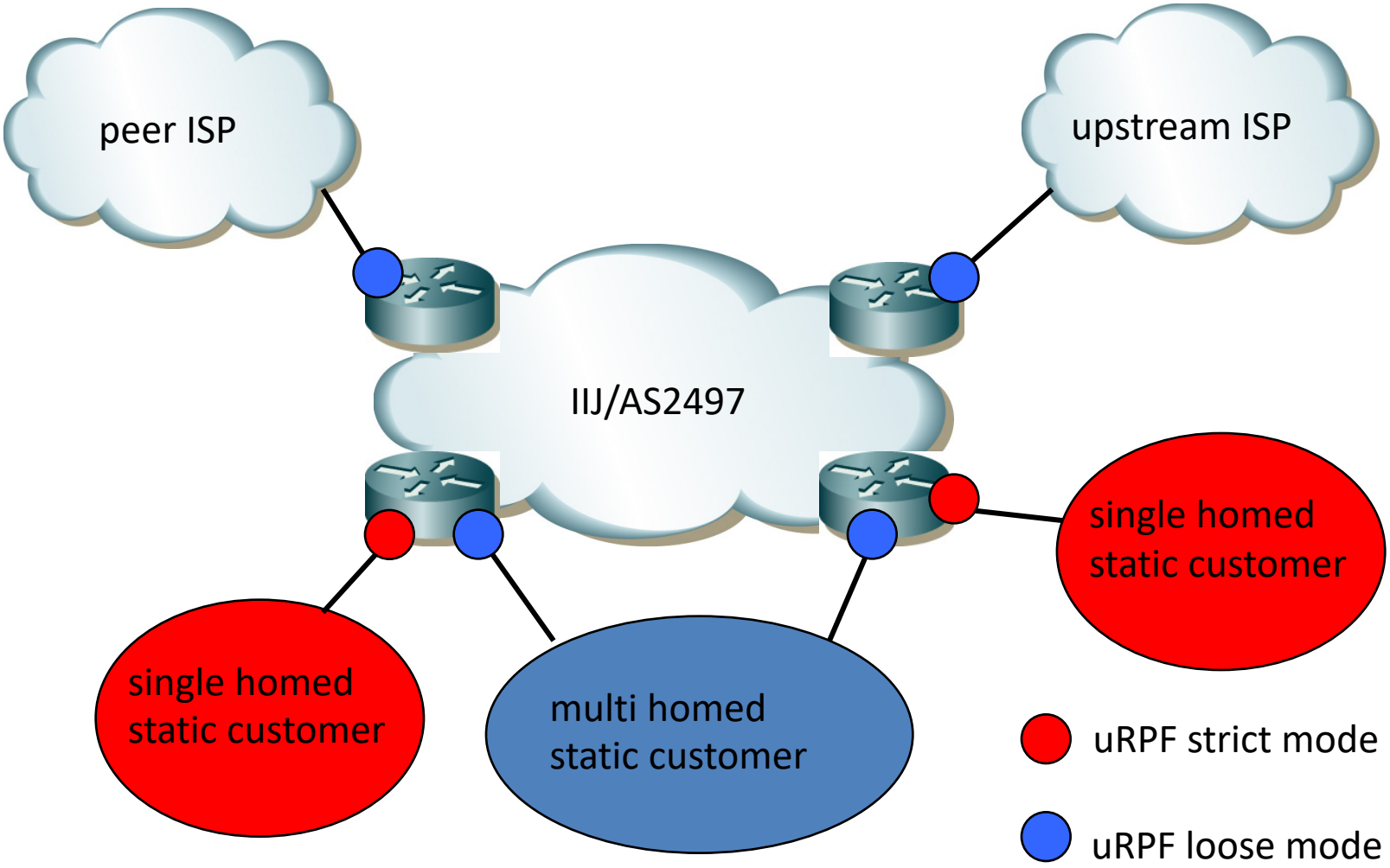


uRPF check – lookup by the src-ip

- routing_table(src-ip) => interface
  - lookup by 192.0.2.1 => if.o
  - The result MUST match the incoming interface



IIJ's policy

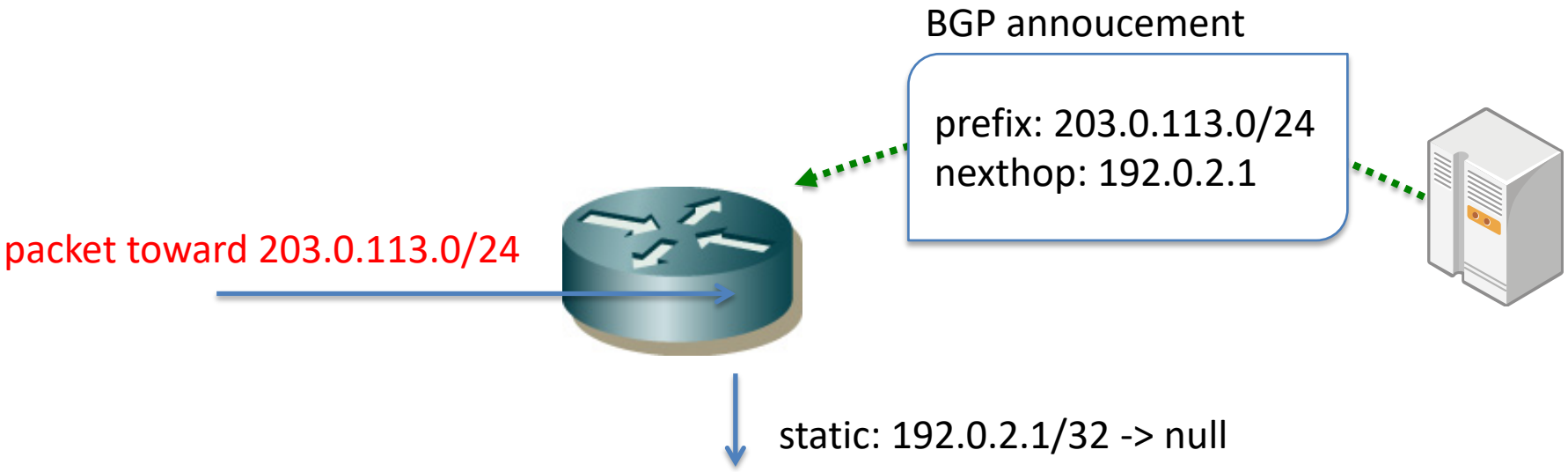

peer ISP
upstream ISP
IIJ/AS2497
single homed
static customer
single homed
static customer
multi homed
static customer
uRPF strict mode
uRPF loose mode

blackhole routing

- Routers are good at forwarding
  - not packet filtering
- Use the forwarding capability to discard packets
  - Null/discard routing

- Remote Triggered Black Hole
  1. Assign IP address for RTBH
  2. Configure static route to discard packets toward the IP address which was assigned at 1)
  3. Announce an blackholing prefix with its netxhop as 1) by BGP

RTBH diagram

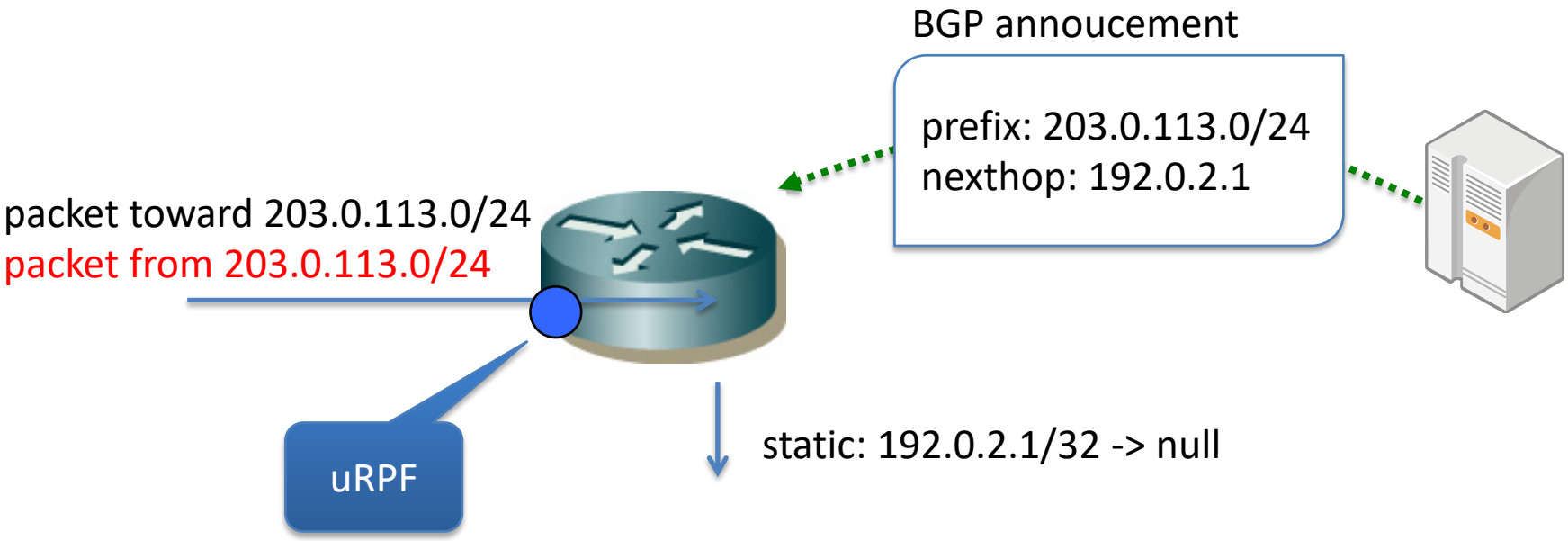


uRPF and blackhole routing

- You can drop a packet that has source ip matches those blackhole route
  - cisco and juniper
- Source IP address based filtering

RTBH w/ uRPF

- Remote Triggered Black Hole



packet filtering for transit traffic

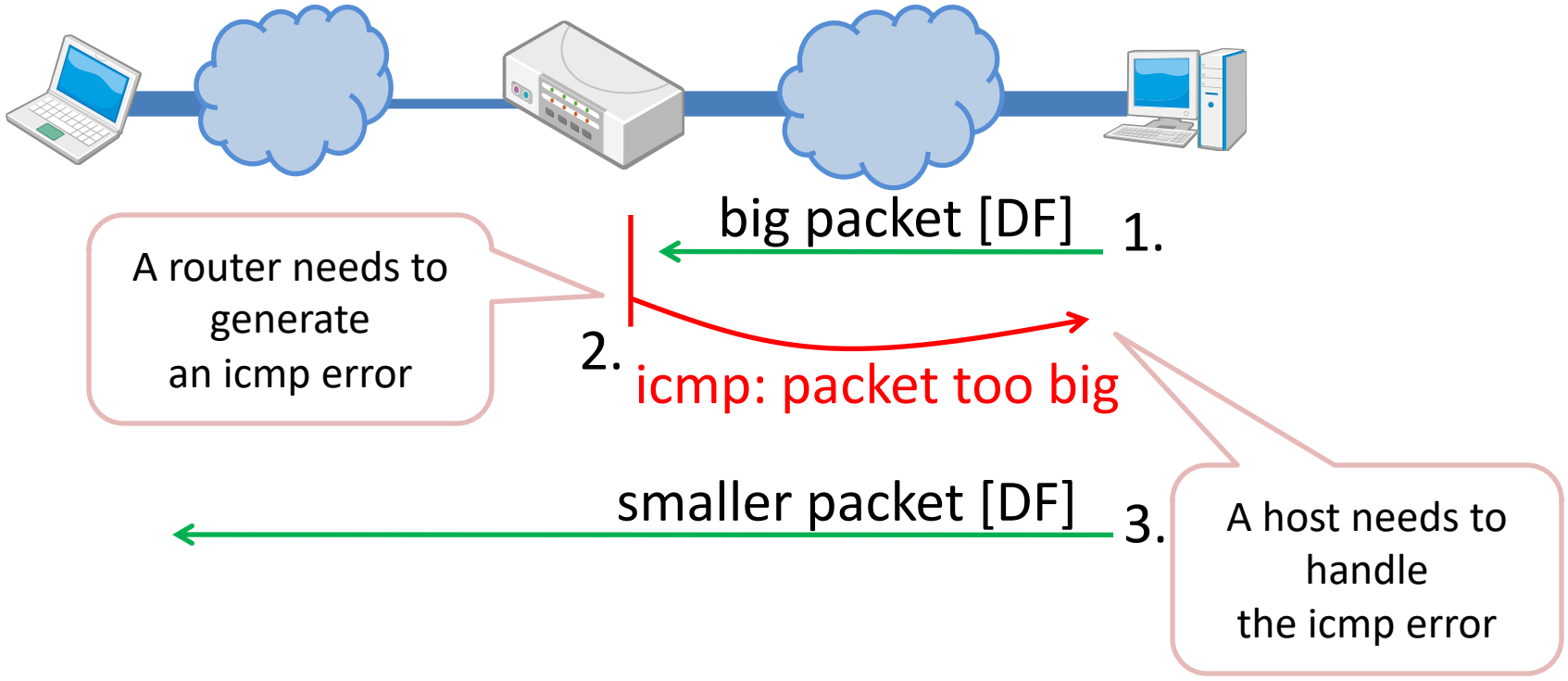
- IP is not that simple
  - IP fragments
  - path MTU discovery
- IPv6, DNSSEC and so on

Path MTU Discovery

- Path MTU discovery [RFC1191]
- Path MTU discovery for IPv6 [RFC1981]

- IPv4 minimum link MTU [RFC791] == 68
  - 576 is widely accepted though
- IPv6 minimum link MTU [RFC2460] == 1280

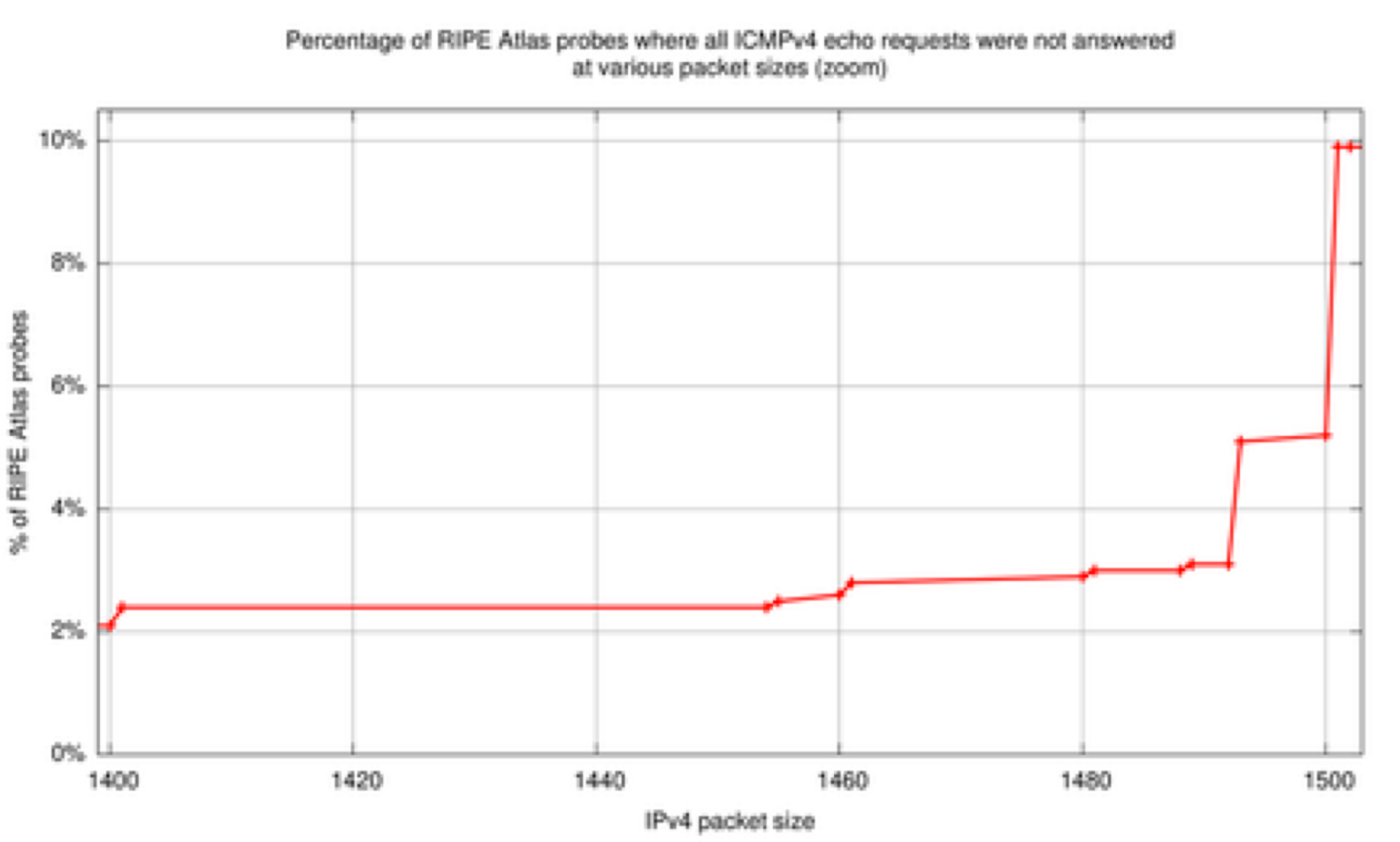
path MTU discovery scenario



icmp originating-limit

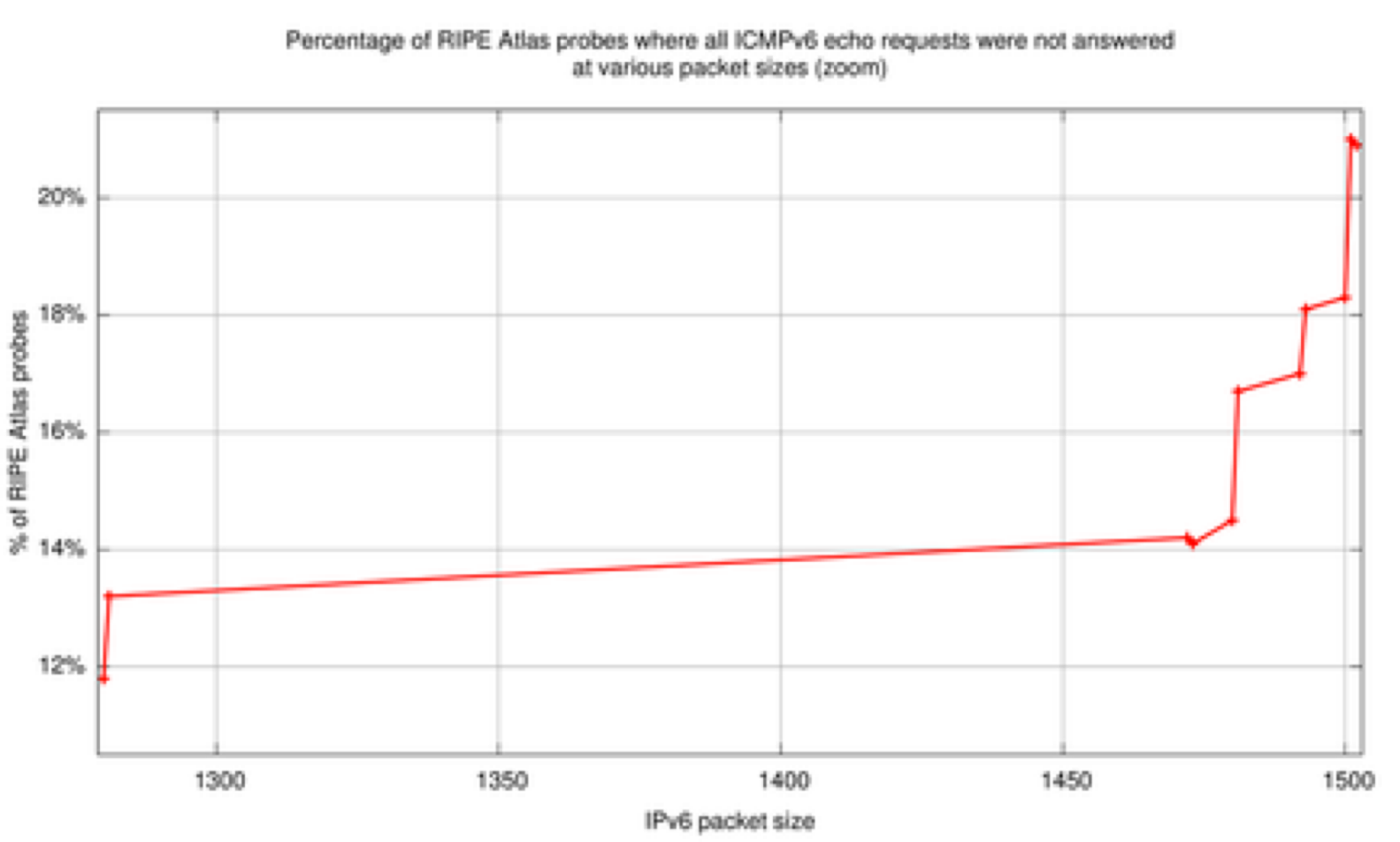
- cisco ios
  - ip icmp rate-limit unreachable 500
    - icmp errors are limited to one every 500msec
  - ipv6 icmp error-interval 100
    - icmp errors are limited to one every 100msec
- juniper junos
  - icmpv4-rate-limit {packet-rate 1000;};
    - up to 1000pps icmp packets to/from RE
  - icmpv6-rate-limit {packet-rate 1000;};
    - up to 1000pps icmp packets to/from RE

IPv4 pMTUd fails



https://labs.ripe.net/Members/emileaben/ripe-atlas-packet-size-matters

IPv6 as well

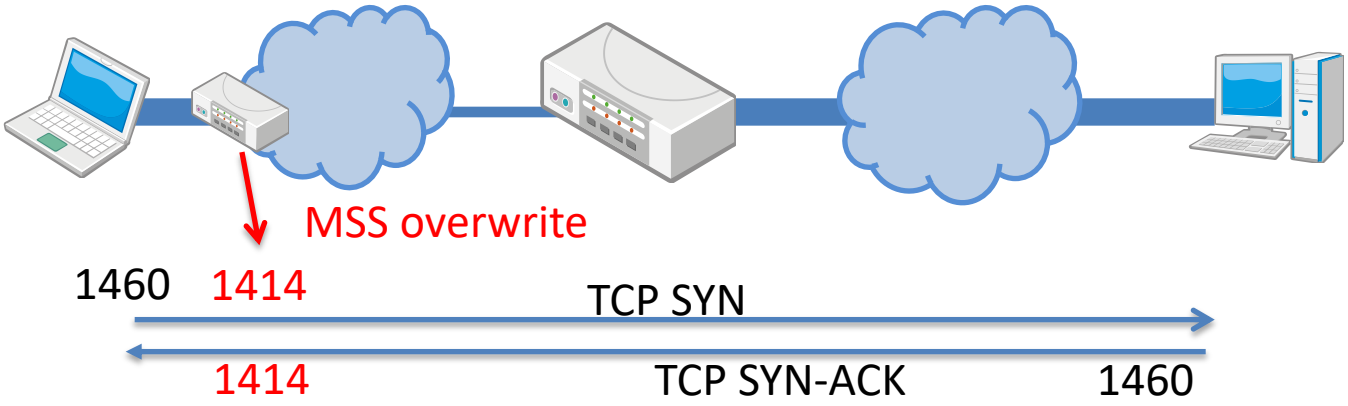


https://labs.ripe.net/Members/emileaben/ripe-atlas-packet-size-matters

learning from IPv4

- Almost of all broadband routers have a TCP MSS hack capability
- It chokes TCP MSS on a tunnel link
  - PPPoE, or whatever the link MTU is less than 1500
  - to avoid unnecessary fallbacks
- The TCP MSS hack works fine
  - No complaint from customers

TCP MSS hack



- both ends agree to use 1414 as MSS size

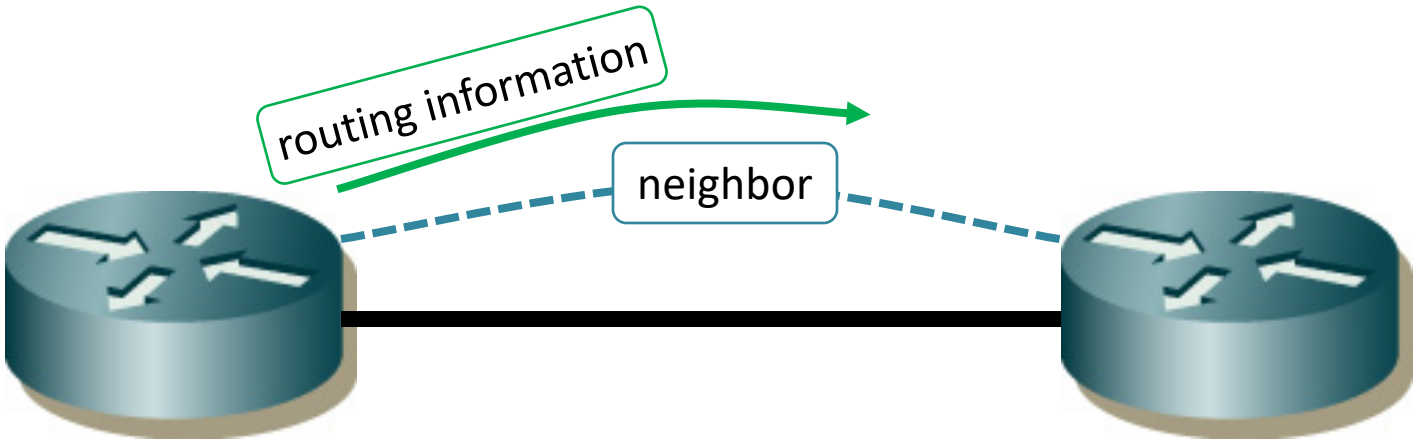
still we need pMTUd

- MSS hack work only for TCP
  - UDP, and any other protocols
- Do not filter ICMP error messages without consideration

Protect Routing

- To keep your network working
  - as you designed
  - as you configured
- Static Routing
  - mostly depends on design
- Dynamic Routing
  - possibility of remote attacks

Routing Protocol



- Routers exchange routing information over a neighboring relationship.

Threat Model for Routing

- Neighboring Relationship
  - Unexpected Neighboring
  - Shutdown by Someone else
  - Spoofed Neighbor
- Routing Information
  - Propagation of Wrong Information
  - Unintended Routing Policy
  - Hit a Hardware Limitation

OSPF Neighbors

- Establishing a relationship among trusted neighbors only
- Disabled by default
  - Especially on a link to other parties (IX,customer)
    - to avoid unexpected neighbors
    - if you have to enable on these links, use ‘passive’ feature
  - Enabled where it is needed like backbone
- Authentication
  - MD5 authentication (OSPFv2, RFC2328)

OSPF md5 configuration

cisco

```
interface <interface_name>
 ip ospf authentication message-digest
 ip ospf message-digest-key <keyid#> md5 <md5_key>
```

juniper

```
protocols ospf  {
 area <area#> {
  interface <interface_name> {
     authentication {
          md5  <keyid#> key “<md5_key>”;
     }
  }
}}
```

BGP4 Neighbors

- Protecting TCP sessions
  - md5 authentication
- Peering with other parties
  - possibility of injection
  - needs more attention about routing information

BGP md5 configuration

cisco

```
router  bgp <as#>
 neighbor <neighbor_ip> password <md5_key>
```

juniper

```
protocols bgp {
 neighbor <neighbor_ip>  {
    authentication-key “<md5_key>”;
 }
}
```

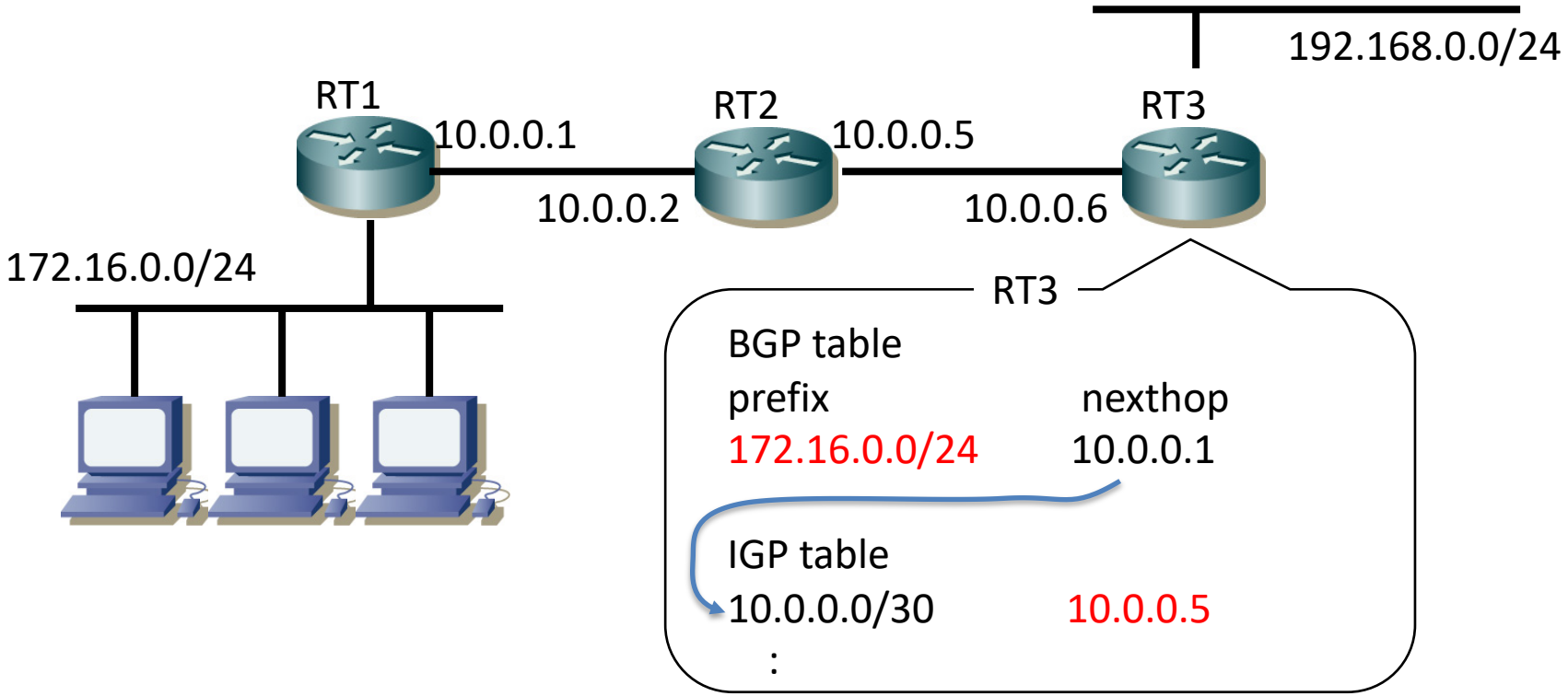
Protecting routing information

- OSPF
  - mostly relies on neighboring
  - IGP should be used for internal purpose
    - should not be used to share routing information with your customers
- BGP
  - routing information is more problematic

critical routing information inside AS

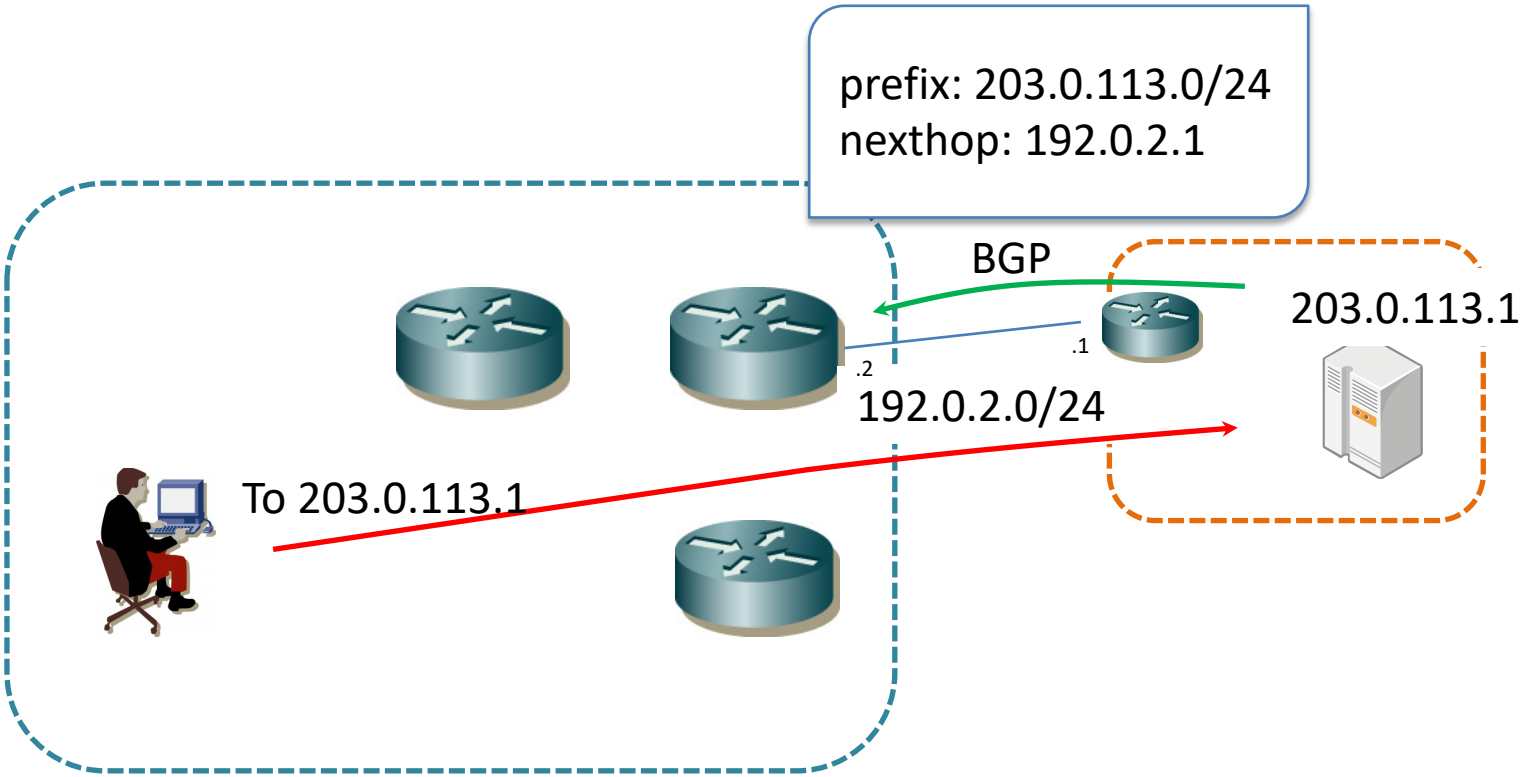
- iBGP neighbor
  - usually loopback interface
  - /32 announcement by IGP
    - the most preferred
- BGP nexthop
  - typical BGP nexthop
    - IX segment
    - peering link
    - customer link
  - route filtering on eBGP sessions
    - needs care about more-specifics

BGP and recursive lookup

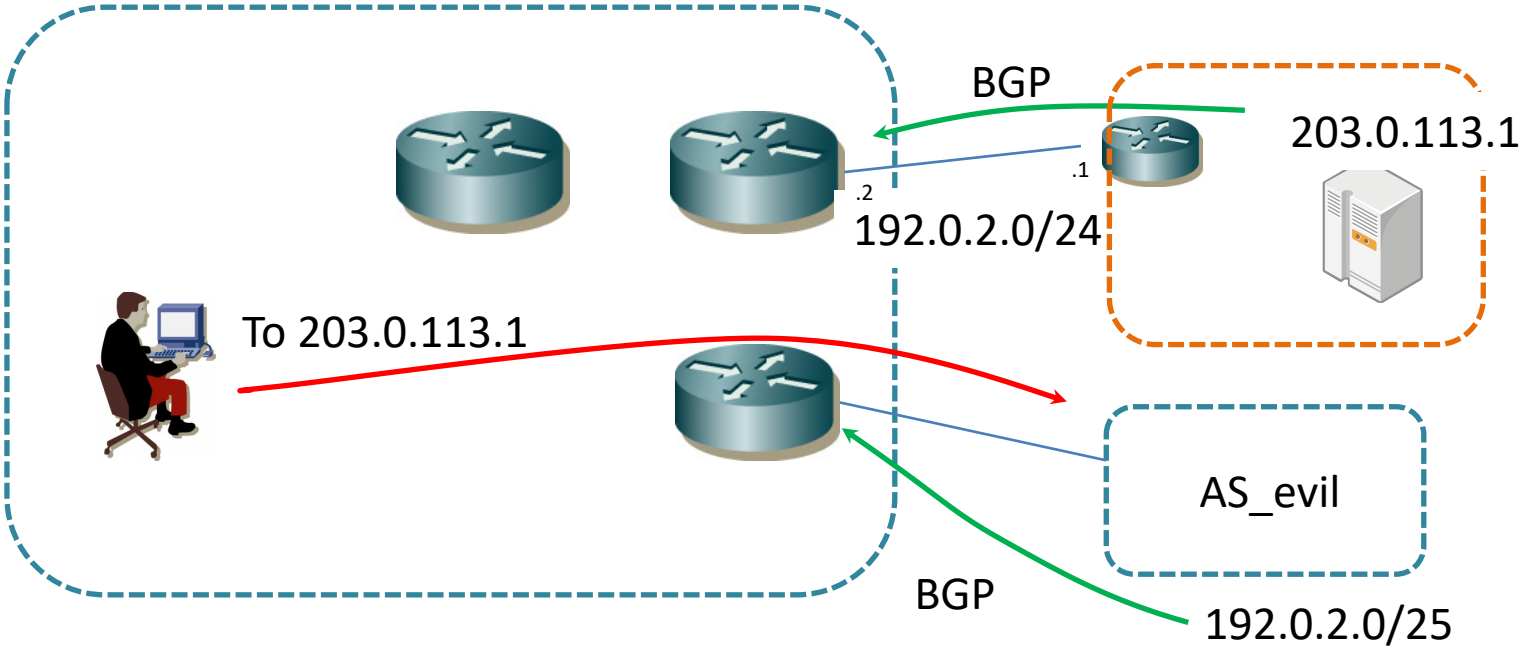


To determine the immediate nexthop for the destination, BGP lookup its routing table recursively.
To get 172.16.0.0/24, the immediate netxthop is 10.0.0.5(RT2)

Usual BGP


prefix: 203.0.113.0/24
nexthop: 192.0.2.1
BGP
203.0.113.1
.1
.2
192.0.2.0/24
To 203.0.113.1

In case of nexthop injection



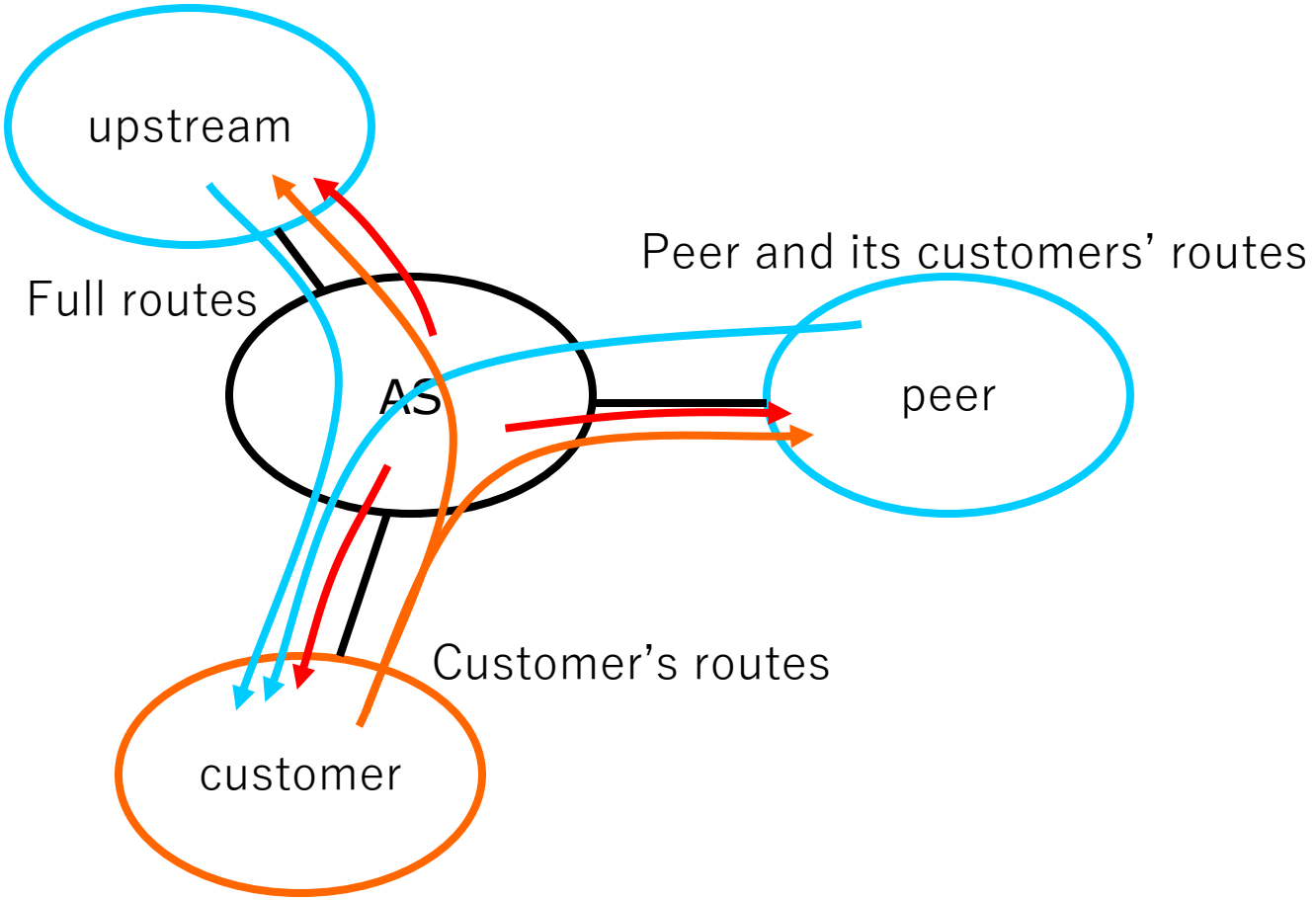
BGP announcement from you

- To peer/upstream AS
  - Your aggregated prefixes
    - More specifics should stay inside your AS
  - Customer ASes' announcements
- To customer AS
  - Full-routes
  - Default route if necessary

BGP announcements from peers

- From upstream AS
  - Full-routes
  - Default route if necessary
- From peer/customer AS
  - Prefixes of the neighbor and its customer AS

BGP announcements diagram



Prefix Filtering

- From peer/upstream AS
  - Deny prefixes those should not be accepted
    - Private, linklocal, testnet
    - Your critical prefixes
  - Then accept others
- From customer AS
  - Accept prefixes that are authorized to announce from the customer AS
    - Prefixes allocated to the customer
    - Some provider independent (PI) spaces
  - Then deny all

BGP Prefix hijack

- Announcing a prefix without permission from the resource holder of the prefix
- It happens many times
  - Mostly it seems unintentional – typo, mitake
  - There are some suspicious cases as well
    - Announcing more specifics only to free mail providers to send SPAM
      - https://ripe72.ripe.net/presentations/45-Invisible_Hijacking.pdf

Our case

- [janog:12845] IIJ to the white courtesy phone.
  - notifying strange BGP announcements
  - also stating the prefix was listed at the Spamhaus SBL

- Thanks for the heads-up!

the /16 IPv4 prefix

- was transferred to IIJ
  - on 21/Oct/2014
- IIJ kept it in stock for future use
  - IIJ didn't start to announce it at that time ☹
  - whois information at JPNIC was updated, but no IRR registration ☹
- An ISP in U.S. started to announce the IP block as 2 x /17s on 5/Jan/2015
  - No, IIJ didn't ask that

to stop the wrong announcements

- IIJ contacted the announcing ISP immediately
  - e-mail to their NOC followed by a phone call
  - and started BGP announcements by ourselves
- The first contact:
  - got ACK and the person on the call agreed to deal with the announcements, but nothing was happened in the next 48 hours
- The second contact:
  - convinced the (different) person on the call, and got a ticket # to track the progress of handling
  - the announcements were finally stopped ☺

lesson learned #1

- ask for a ticket #
  - especially in case the ISP has a ticket system to track their jobs
- keep whois DB up-to-date
  - To prove your correctness
  - I sent our whois information to the NOC by e-mail, and also asked the NOC person to query the prefix by himself

the progress

- 4/Feb/2015 - the post to JANOG
  - the first contact to the ISP
- 6/Feb/2015 - the second contact to the ISP
- 7/Feb/2015 - the routes were withdrawn
- 12/Feb/2015 - contacted Spamhaus to delist
- 13/Feb/2015 - the prefix was delisted from SBL

the cause of the announcements

- A customer of the ISP submitted a LoA (Letter of Authority) to use the prefix, and asked the ISP to originate the BGP announcements

- No, IIJ didn't submit such a document

An Example of Letter of Authority

<Company Name>
<Address>

<date>

To: <the Customer>

We authorize <the Customer> or <the ISP> to announce the following IP blocks -

<IP address blocks>

This authorization shall be valid until revoked by us in writing or by e-mail from <e-mail address>.

I may be contacted at <Tel#> or <e-mail address>

Sincerely,

<signature>

<signer's name in print>

<Company Name>

the actual LoA looks ... strange

- The company name was a family company of the previous resource holder
- Suspicious
  - the domain name used as a contact e-mail address was different from the actual one
    - The domain name was newly registered in 2014
  - the Tel# was wrong - missing a country code
    - As the previous holder registered it wrongly at the whois DB before

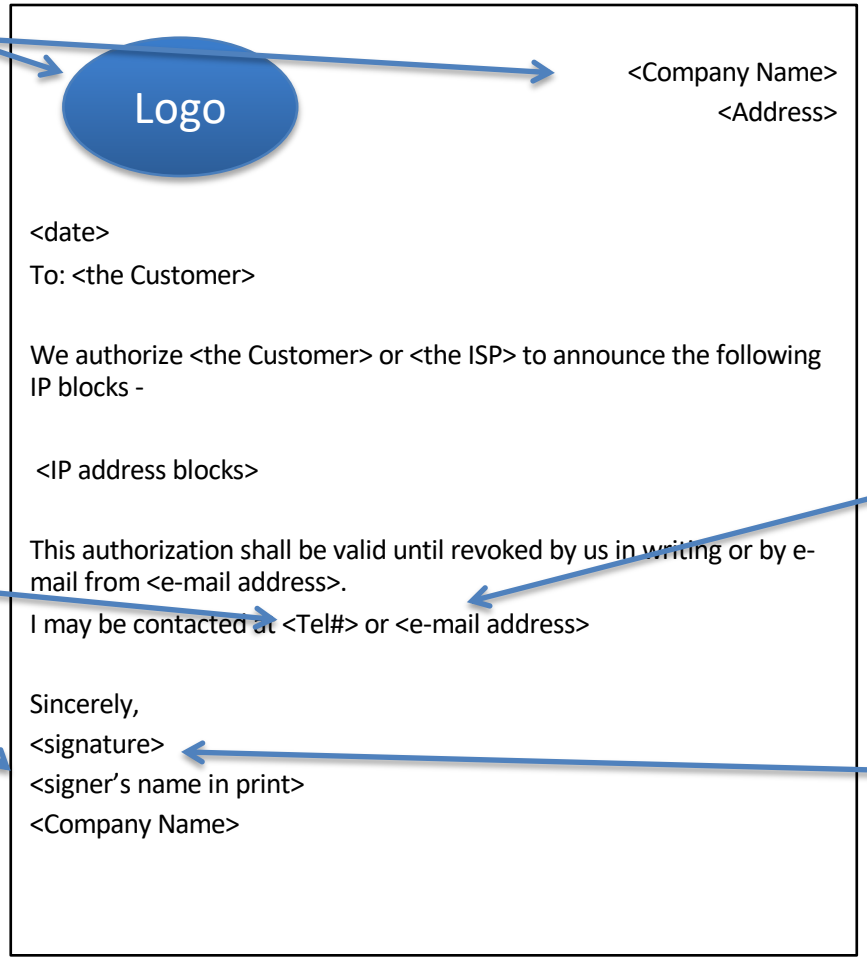
visited the previous resource holder

- met a person who was previously the contact person of their whois DB entry
  - and also his name was used as a signer in the LoA
- No, he didn't sign the document, and their company wasn't aware of the LoA and even the domain name which was used in the LoA

- A fake LoA!!

the fake LoA

Copied from a web site of a family company of the previous resource holder

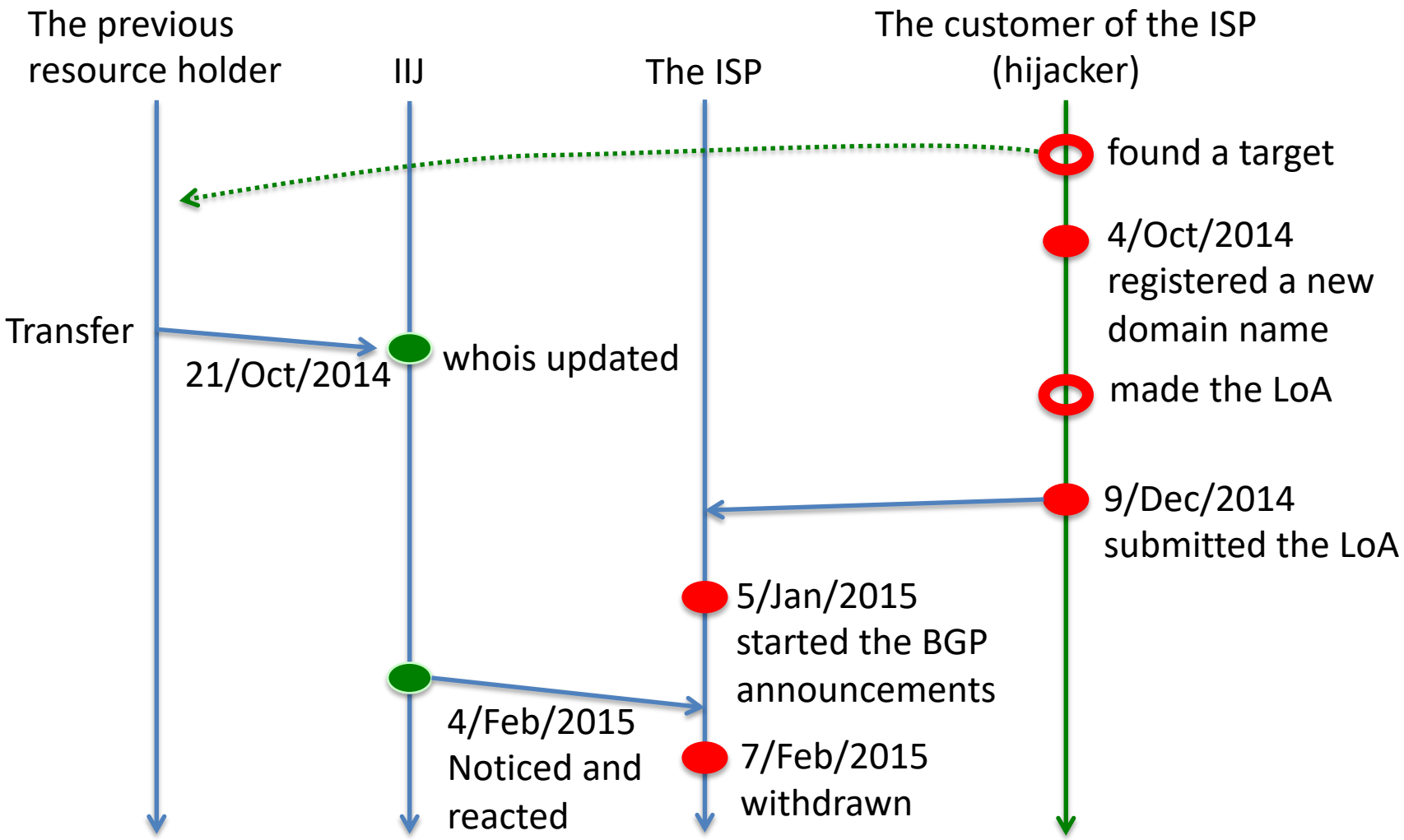


Registered a new domain name looks like related to the organization

Copied from previous whois DB entry

A fake signature

timeline


The previous resource holder
IIJ
The ISP
The customer of the ISP (hijacker)
found a target
4/Oct/2014 registered a new domain name
made the LoA
Transfer
21/Oct/2014
whois updated
9/Dec/2014 submitted the LoA
5/Jan/2015 started the BGP announcements
4/Feb/2015 Noticed and reacted
7/Feb/2015 withdrawn

the hijacker

- We don’t know how they used the network
  - no evidence so far
  - no spam compliant related to the prefix
- After stopping the announcement, they started to use ‘the next’ prefix by using the same technique - by submitting a fake LoA ☹
  - and it was noticed and stopped by the actual resource holder a few months later

looking back

- IIJ should announce all holding prefixes
  - We changed our policy to announce all of them
  - Before announcements, IIJ registers route objects to IRRs - JPIRR and RADB. By registering a route object at JPIRR, a route monitoring service named ‘keiro bugyo’ automatically starts to monitor malicious announcement related to the route object. ☺
- The ISP should carefully check IP blocks before announcements
  - As whois DB was already changed - indicating IIJ as a resource holder at that time

lesson learned #2

- announce all holding prefixes
  - register route objects to an IRR for reference
- RPKI is the next choice for us
  - we need to promote RPKI more, and train engineers to be aware of public-key cryptography
  - signing and verifying by using public-key cryptography is a key technology now days

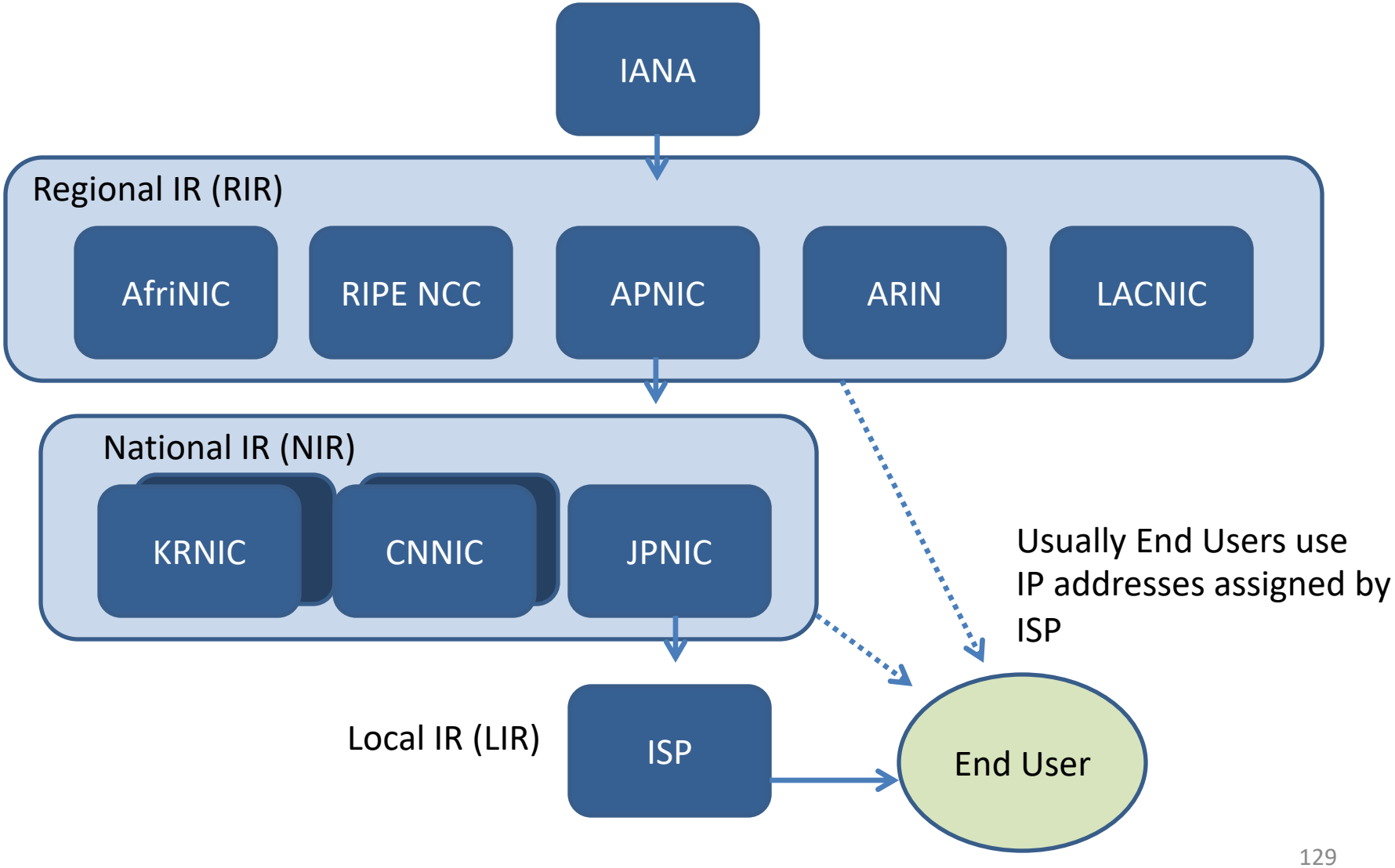
RPKI

- Public Key Infrastructure for Number Resources
  - such as IP addresses and AS numbers
  - a digital certificate can prove that you are the current resource holder of specific number resource
  - you can add digital signature to your documents like LoA or transfer agreement
- You can issue ROAs to indicate originating AS for prefixes

Internet Registry (IR)

- Maintains Internet Resources such as IP addresses and ASNs, and publish the registration information
  - allocations for Local Internet Registries
  - assignments for end-users
- APNIC is the Regional Internet Registry(RIR) in the Asia Pacific region
  - National Internet Registry(NIR) exists in several economies

management of IP addresses



whois at IANA

```
$ whois -h whois.iana.org '160.13.0.0'
% IANA WHOIS server
% for more information on IANA, visit http://www.iana.org
% This query returned 1 object

refer:        whois.arin.net

inetnum:      160.0.0.0 - 160.255.255.255
organisation: Administered by ARIN
status:       LEGACY

whois:        whois.arin.net

changed:      1993-05
source:       IANA
```

whois at ARIN

```
$ whois -h whois.arin.net '160.13.0.0'

#
# ARIN WHOIS data and services are subject to the Terms of Use
# available at: https://www.arin.net/whois_tou.html
#
# If you see inaccuracies in the results, please report at
# http://www.arin.net/public/whoisinaccuracy/index.xhtml
#


#
# Query terms are ambiguous.  The query is assumed to be:
#     "n 160.13.0.0"
#
# Use "?" to get help.
#

#
# The following results may also be obtained via:
#
http://whois.arin.net/rest/nets;q=160.13.0.0?showDetails=true&showARIN=false&showNonArinTopLeve
lNet=false&ext=netref2
#

NetRange:       160.11.0.0 - 160.30.255.255
CIDR:           160.24.0.0/14, 160.11.0.0/16, 160.30.0.0/16, 160.28.0.0/15, 160.12.0.0/14, 160.16.0.0/13
NetName:        APNIC-ERX-160-11-0-0
NetHandle:      NET-160-11-0-0-1
Parent:         NET160 (NET-160-0-0-0-0)
NetType:        Early Registrations, Transferred to APNIC
OriginAS:
Organization:   Asia Pacific Network Information Centre (APNIC)
RegDate:        2004-04-05
Updated:        2009-10-08
Comment:        This IP address range is not registered in the ARIN database.
Comment:        This range was transferred to the APNIC Whois Database as
Comment:        part of the ERX (Early Registration Transfer) project.
Comment:        For details, refer to the APNIC Whois Database via
Comment:        WHOIS.APNIC.NET or http://wq.apnic.net/apnic-bin/whois.pl
Comment:
Comment:        ** IMPORTANT NOTE: APNIC is the Regional Internet Registry
Comment:        for the Asia Pacific region.  APNIC does not operate networks
Comment:        using this IP address range and is not able to investigate
Comment:        spam or abuse reports relating to these addresses.  For more
Comment:        help, refer to http://www.apnic.net/apnic-info/whois_search2/abuse-and-spamming
Ref:            http://whois.arin.net/rest/net/NET-160-11-0-0-1

ResourceLink:  http://wq.apnic.net/whois-search/static/search.html
ResourceLink:  whois.apnic.net

OrgName:        Asia Pacific Network Information Centre
OrgId:          APNIC
Address:        PO Box 3646
City:           South Brisbane
StateProv:      QLD
PostalCode:     4101
Country:        AU
RegDate:
Updated:        2012-01-24
Ref:            http://whois.arin.net/rest/org/APNIC

ReferralServer:  whois://whois.apnic.net
ResourceLink:  http://wq.apnic.net/whois-search/static/search.html

OrgAbuseHandle: AWC12-ARIN
OrgAbuseName:   APNIC Whois Contact
OrgAbusePhone:  +61 7 3858 3188
OrgAbuseEmail:  search-apnic-not-arin@apnic.net
OrgAbuseRef:    http://whois.arin.net/rest/poc/AWC12-ARIN

OrgTechHandle: AWC12-ARIN
OrgTechName:   APNIC Whois Contact
OrgTechPhone:  +61 7 3858 3188
OrgTechEmail:  search-apnic-not-arin@apnic.net
OrgTechRef:    http://whois.arin.net/rest/poc/AWC12-ARIN


#
# ARIN WHOIS data and services are subject to the Terms of Use
# available at: https://www.arin.net/whois_tou.html
#
# If you see inaccuracies in the results, please report at
# http://www.arin.net/public/whoisinaccuracy/index.xhtml
#
```

whois at APNIC

```
$ whois -h whois.apnic.net '160.13.0.0'
% [whois.apnic.net]
% Whois data copyright terms    http://www.apnic.net/db/dbcopyright.html

% Information related to '160.13.0.0 - 160.13.255.255'

inetnum:        160.13.0.0 - 160.13.255.255
netname:        IIJ
descr:          Internet Initiative Japan Inc.
descr:          Iidabashi Grand Bloom,
descr:          2-10-2 Fujimi, Chiyoda-ku,
descr:          Tokyo, 102-0071 Japan
country:        JP
admin-c:        JNIC1-AP
tech-c:         JNIC1-AP
status:         ALLOCATED PORTABLE
remarks:        Email address for spam or abuse complaints : abuse-contact@iij.ad.jp
mnt-irt:        IRT-JPNIC-JP
mnt-by:         MAINT-JPNIC
mnt-lower:      MAINT-JPNIC
changed:        hm-changed@apnic.net 20050712
changed:        ip-apnic@nic.ad.jp 20141021
source:         APNIC

irt:            IRT-JPNIC-JP
address:        Urbannet-Kanda Bldg 4F, 3-6-2 Uchi-Kanda
address:        Chiyoda-ku, Tokyo 101-0047, Japan
e-mail:         hostmaster@nic.ad.jp
abuse-mailbox:  hostmaster@nic.ad.jp
admin-c:        JNIC1-AP
tech-c:         JNIC1-AP
auth:           # Filtered
mnt-by:         MAINT-JPNIC
changed:        abuse@apnic.net 20101108
changed:        hm-changed@apnic.net 20101111
changed:        ip-apnic@nic.ad.jp 20140702
source:         APNIC

role:           Japan Network Information Center
address:        Urbannet-Kanda Bldg 4F
address:        3-6-2 Uchi-Kanda
address:        Chiyoda-ku, Tokyo 101-0047,Japan
country:        JP
phone:          +81-3-5297-2311
fax-no:         +81-3-5297-2312
e-mail:         hostmaster@nic.ad.jp
admin-c:        JI13-AP
tech-c:         JE53-AP
nic-hdl:        JNIC1-AP
mnt-by:         MAINT-JPNIC
changed:        hm-changed@apnic.net 20041222
changed:        hm-changed@apnic.net 20050324
changed:        ip-apnic@nic.ad.jp 20051027
changed:        ip-apnic@nic.ad.jp 20120828
source:         APNIC

% Information related to '160.13.0.0 - 160.13.15.255'

inetnum:        160.13.0.0 - 160.13.15.255
netname:        IIJNET
descr:          IIJ Internet
country:        JP
admin-c:        JP00010080
tech-c:         JP00010080
remarks:        This information has been partially mirrored by APNIC from
remarks:        JPNIC. To obtain more specific information, please use the
remarks:        JPNIC WHOIS Gateway at
remarks:        http://www.nic.ad.jp/en/db/whois/en-gateway.html or
remarks:        whois.nic.ad.jp for WHOIS client. (The WHOIS client
remarks:        defaults to Japanese output, use the /e switch for English
remarks:        output)
changed:        apnic-ftp@nic.ad.jp 20150417
changed:        apnic-ftp@nic.ad.jp 20150424
source:         JPNIC

% This query was served by the APNIC Whois Service version 1.69.1-APNICv1r7-SNAPSHOT (WHOIS4)
```

whois at JPNIC

```
$ whois -h whois.nic.ad.jp '160.13.0.0 /e'
[ JPNIC database provides information regarding IP address and ASN. Its use   ]
[ is restricted to network administration purposes. For further information,  ]
[ use 'whois -h whois.nic.ad.jp help'. To only display English output,        ]
[ add '/e' at the end of command, e.g. 'whois -h whois.nic.ad.jp xxx/e'.      ]

Network Information:
a. [Network Number]           160.13.0.0/20
b. [Network Name]             IIJNET
g. [Organization]             IIJ Internet
m. [Administrative Contact]   JP00010080
n. [Technical Contact]        JP00010080
p. [Nameserver]               dns0.iij.ad.jp
p. [Nameserver]               dns1.iij.ad.jp
[Assigned Date]               2015/04/17
[Return Date]
[Last Update]                 2015/04/24 11:47:06(JST)

Less Specific Info.
----------
Internet Initiative Japan Inc.
                [Allocation]                  160.13.0.0/16

More Specific Info.
----------
No match!!
```

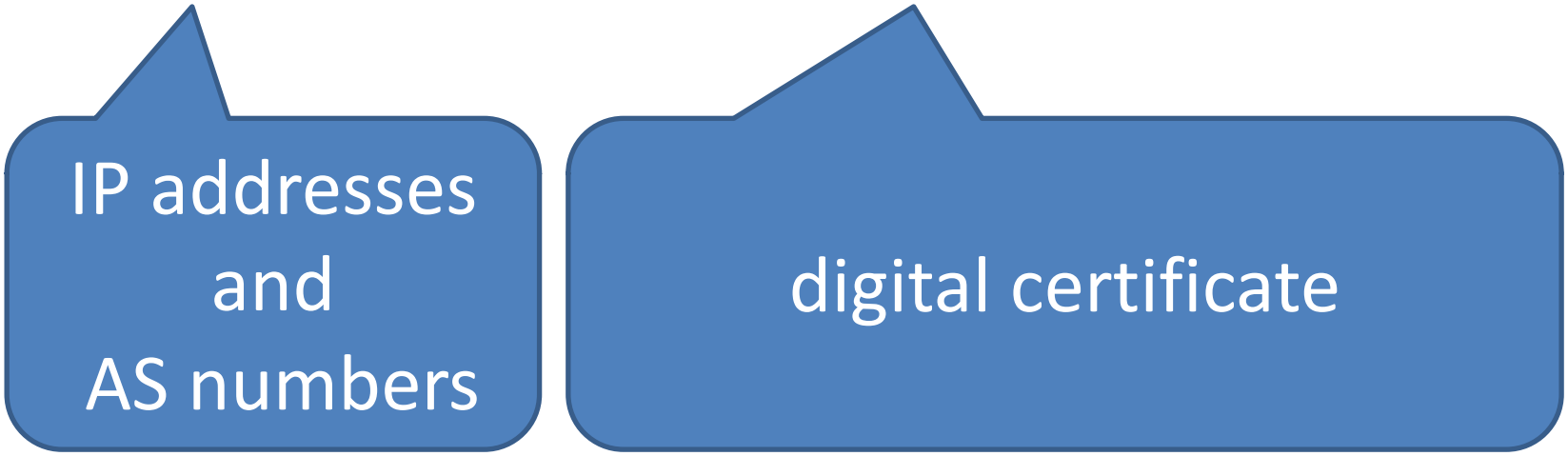
whois at JPNIC again

```
$ whois -h whois.nic.ad.jp '160.13.0.0/16 /e'
[ JPNIC database provides information regarding IP address and ASN. Its use   ]
[ is restricted to network administration purposes. For further information,  ]
[ use 'whois -h whois.nic.ad.jp help'. To only display English output,        ]
[ add '/e' at the end of command, e.g. 'whois -h whois.nic.ad.jp xxx/e'.      ]

Network Information:
[Network Number]              160.13.0.0/16
[Network Name]
[Organization]                Internet Initiative Japan Inc.
[Administrative Contact]      JP00010080
[Technical Contact]           JP00010080
[Abuse]                       abuse-contact@iij.ad.jp
[Allocated Date]              2014/10/21
[Last Update]                 2014/10/21 15:04:47(JST)

Less Specific Info.
----------
No match!!

More Specific Info.
----------
IIJ Internet
        IIJNET [Assignment]                   160.13.0.0/20
IIJ Internet
        IIJNET [Assignment]                   160.13.16.0/24
 [...]
```

allocations

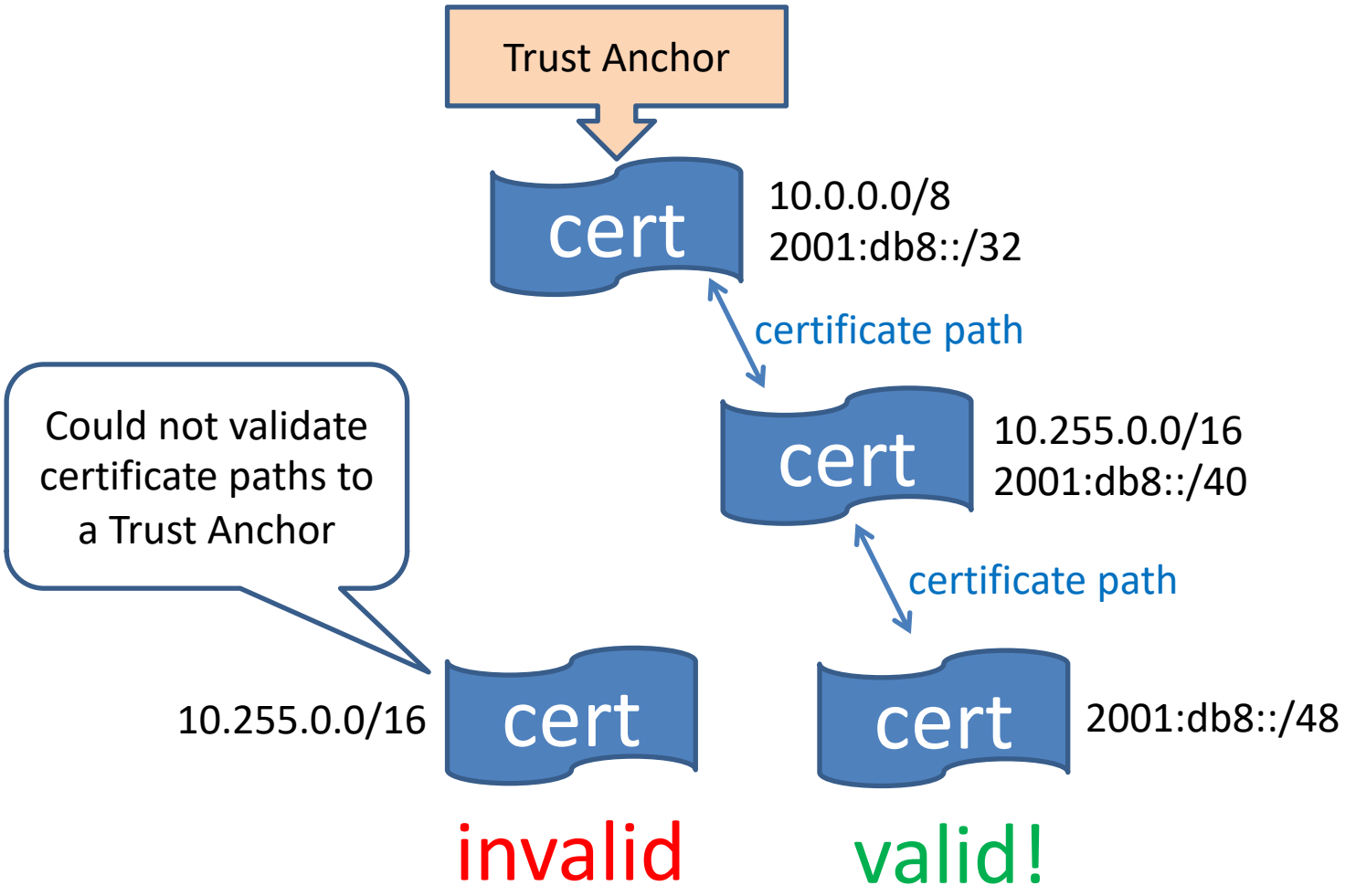
- It’s already complicated
  - and getting more complicated
- IR whois is not so human friendly nor machine friendly
  - You need to train engineers about every whois DB’s expressions, history of the Internet, the current resource policies. Yes, it’s important though...
  - And probably that’s why we have IRRs to register routing related information
- We need something better to prove our holding resources

Resource Public Key Infrastructure

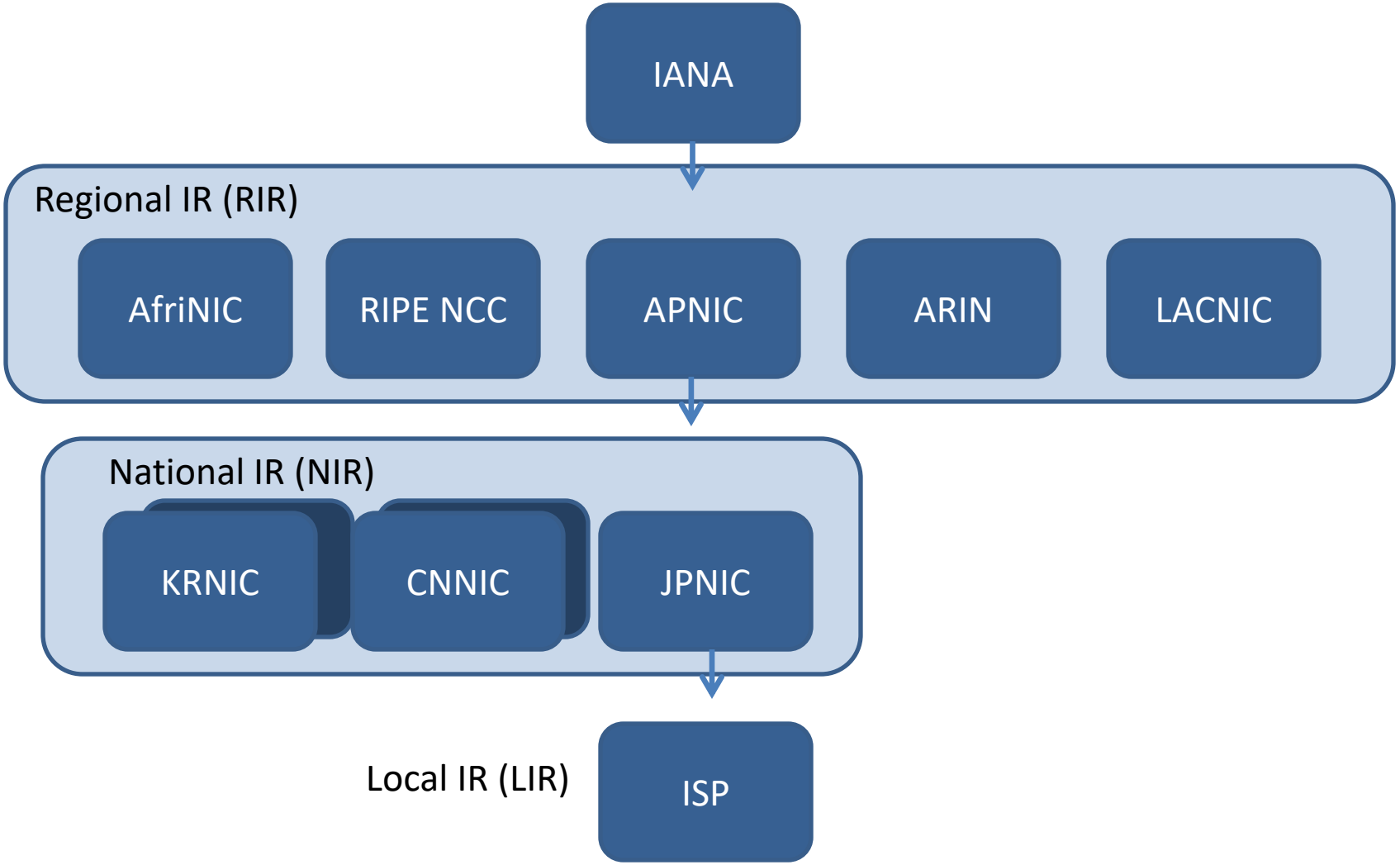


- so-called RPKI
- a PKI for Internet Resources
  - based on Public-key cryptography technology (X.509)
  - enables users to verify the authenticity of Internet Resources

RPKI structure



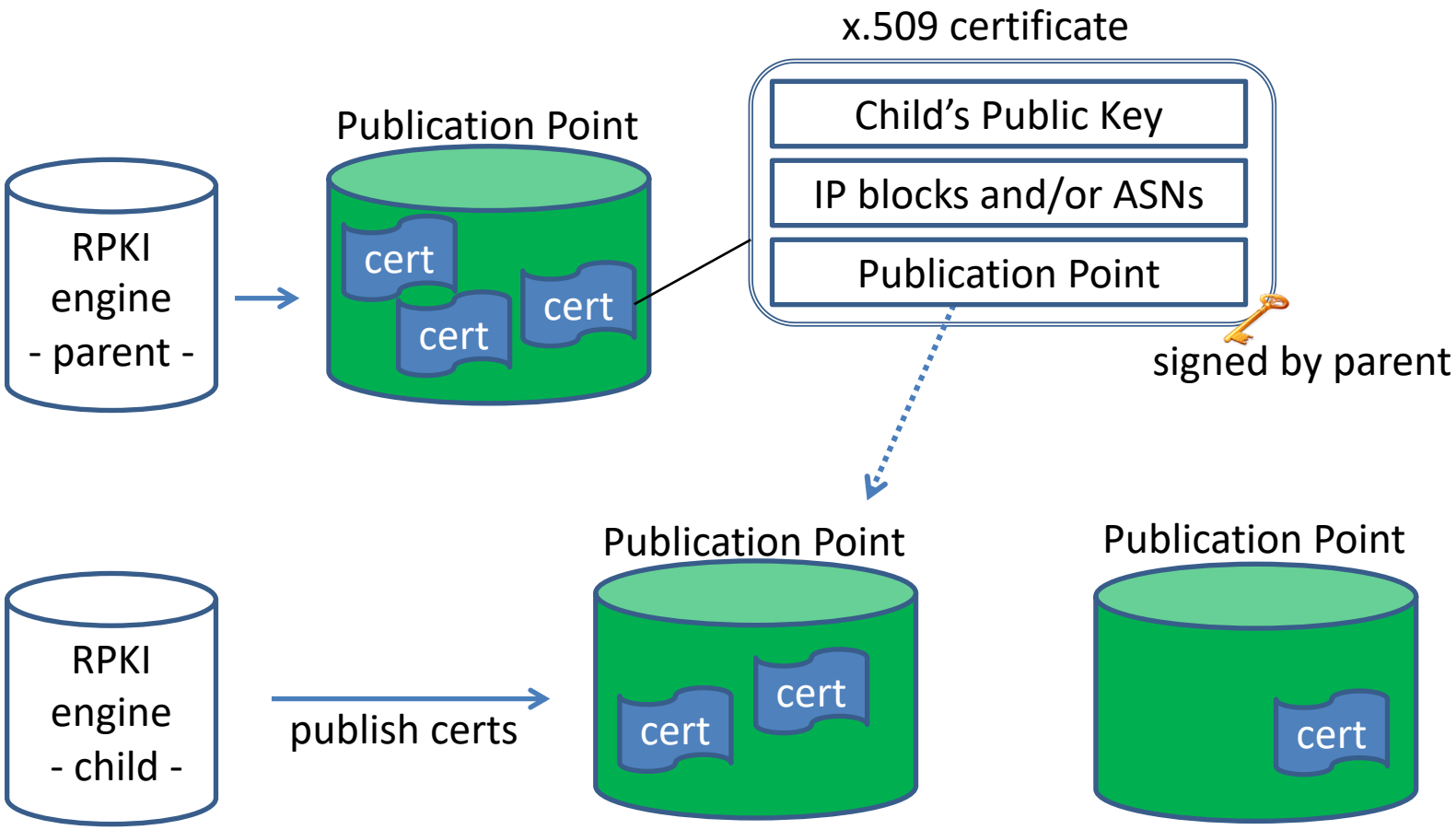
certificate and allocation hierarchy



Trust Anchor Locations (TALs)

- A rsync URL and Public Key information
  - RFC6490
- 5 RIRs support RPKI already
  - each RIR publishes TAL for their resources
  - https://www.ripe.net/lir-services/resource-management/certification/rir-trust-anchor-statistics

RPKI publication



certificate

```
$ openssl x509 -inform DER -text -in nUoKQJmirKA2dIS40zY34cs7tKc.cer
 :
Subject Information Access:
      CA Repository - URI:rsync://rpki.apnic.net/member_repository/XXX/XX/
 :
sbgp-autonomousSysNum: critical
            Autonomous System Numbers:
              2497-2528
              2554
 :
sbgp-ipAddrBlock: critical
      IPv4:
            1.0.16.0/20
            1.0.64.0/18
:
```

publication point

Route Origin Attestations (ROAs)

- a signed object contains an AS and IP prefixes
  - the AS is authorized to originate routes to the given IP prefixes
    - similar to IRR's route and route6 object
  - an IP address block holder can issue a ROA within that block
- 'maximum length' option
  - specifies the maximum length of an IP prefix that the AS is authorized to originate

ROA

```
$ print_roa FksMMjbAOUZnFeuDv2yZmcAXJeY.roa
 :
asID:          2497
 addressFamily: 2
    IPaddress: 2001:240::/32
```

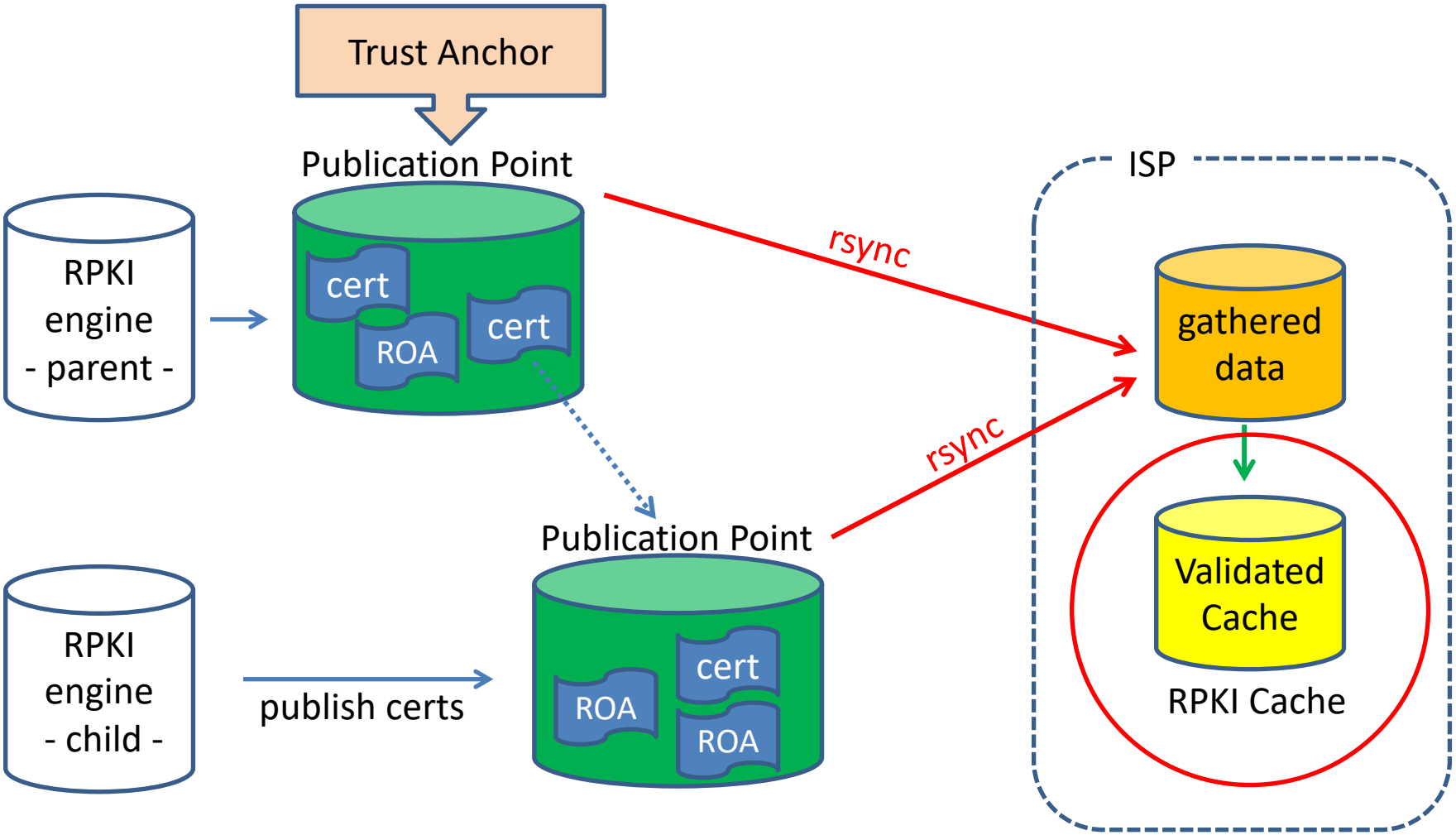
You can issue multiple ROAs to originate a prefix from different ASes

ROA example

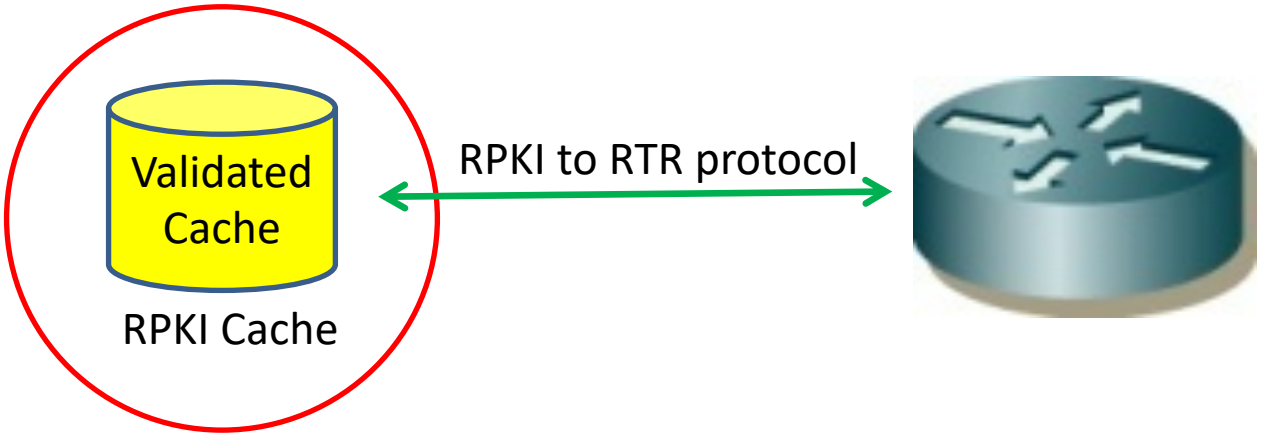
prefix	Max length	AS
2001:db8::/32	32	65551
2001:db8::/48	64	4200000000
192.0.2.0/24	24	65536
192.0.2.0/24	32	65537

- Note: Special care is needed if you allow your customer to announce an assigned prefix from their AS

RPKI cache


Trust Anchor
Publication Point
RPKI
engine
- parent -
cert
ROA
cert
rsync
ISP
gathered
data
rsync
Publication Point
Validated
Cache
RPKI Cache
RPKI
engine
- child -
publish certs
ROA
cert
ROA

Origin Validation

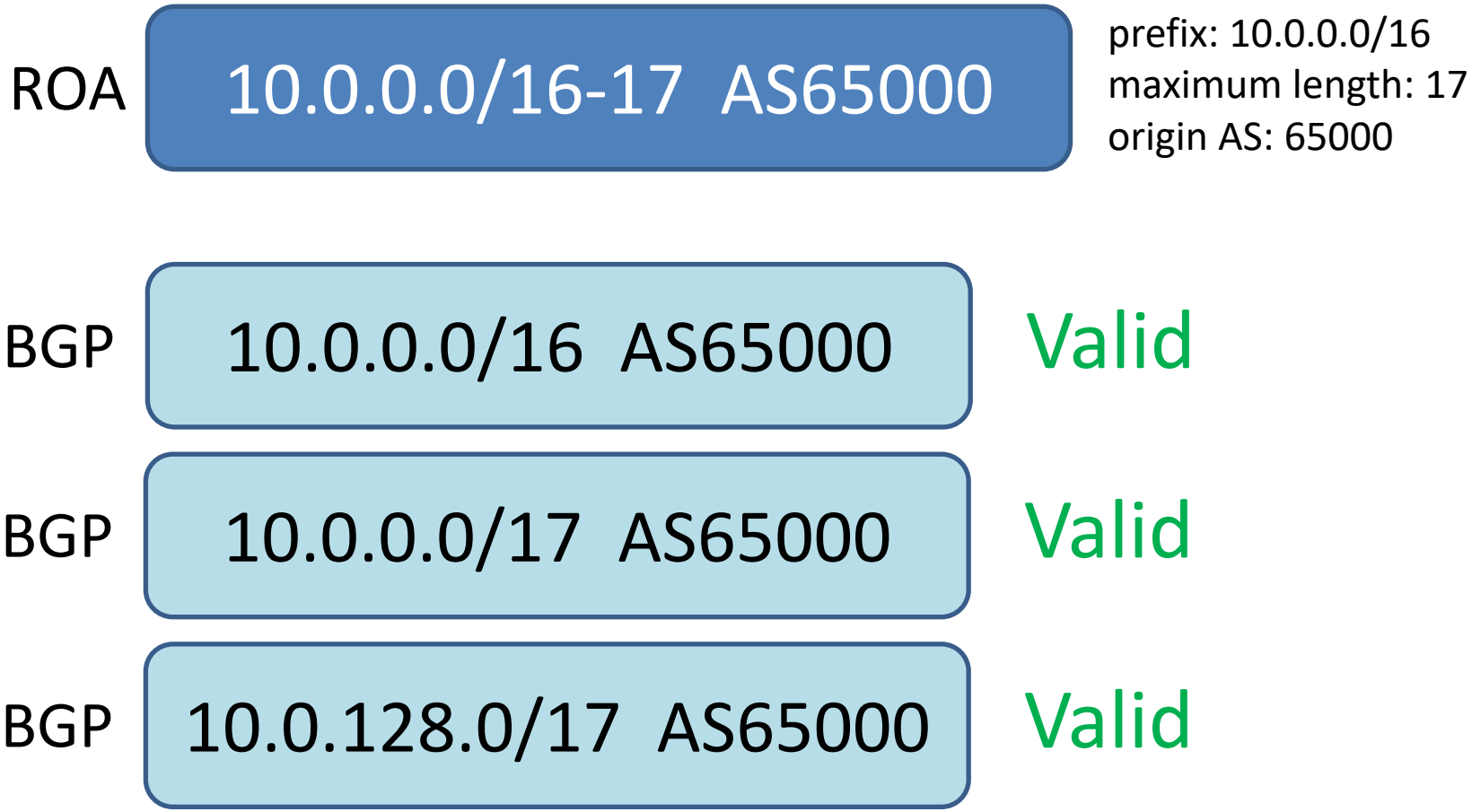


- Router gets ROA information from the RPKI Cache
  - RPKI verification is done by the RPKI Cache
- The BGP process will check each announcement with the ROA information and label the prefix

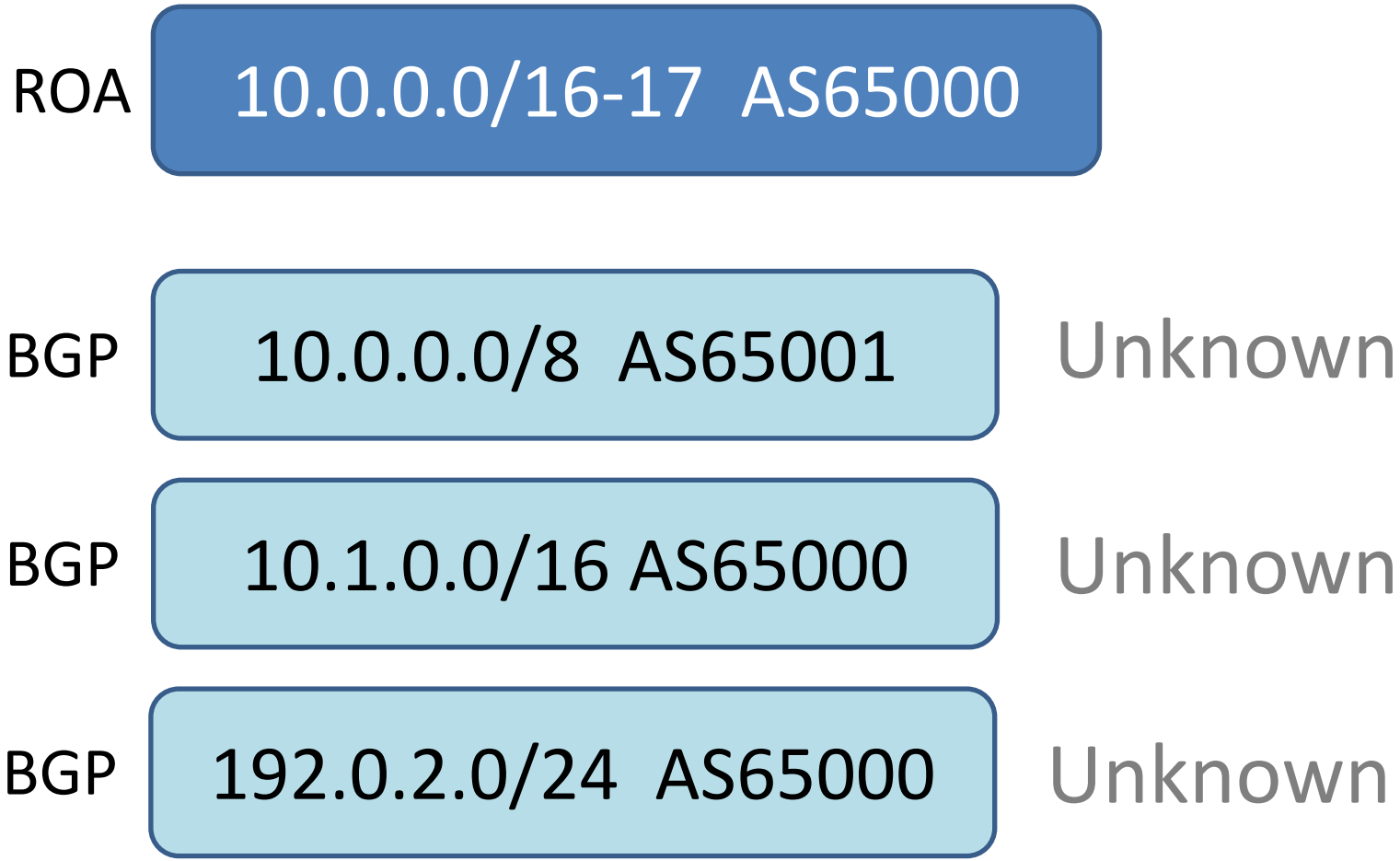
possible outcomes

- Valid
  - a ROA matching the prefix and ASN is found
- Unknown (Not found)
  - There is no covering ROA for the prefix
- Invalid
  - There are ROAs covering the prefix, but none of them matches the ASN or the prefix length

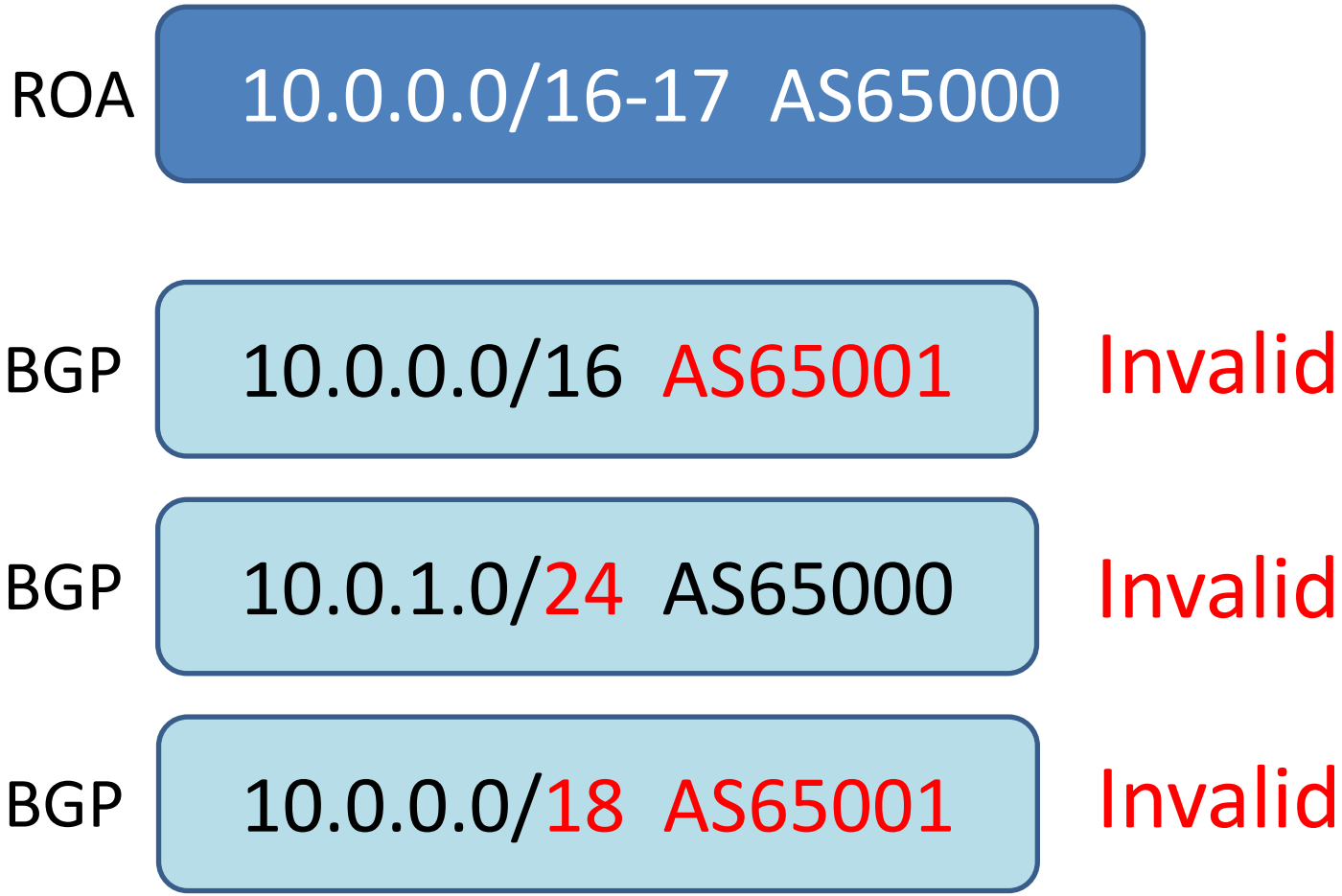
example - valid



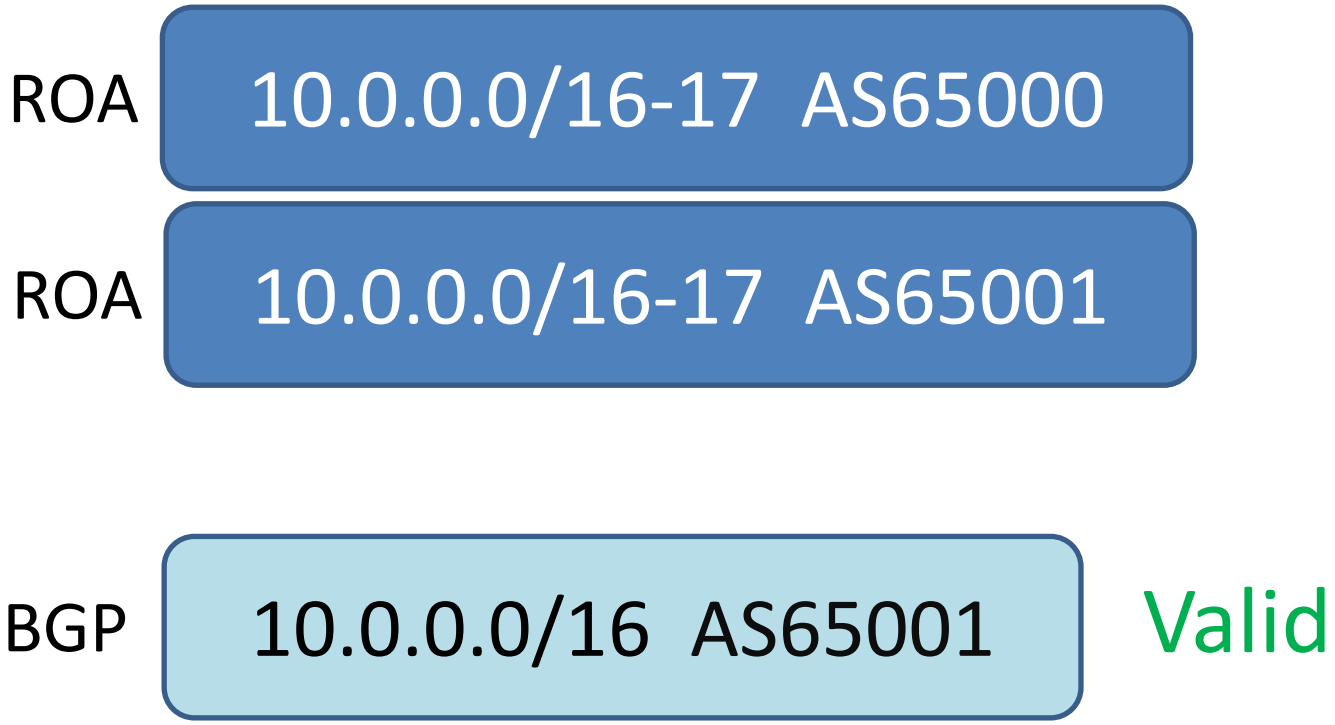
example - unknown



example - invalid



example - multiple origin ROA



local policy

- You can define your policy based on the outcomes
  - do nothing
  - just logging
  - label BGP communities
  - modify preference values
  - rejecting the announcement

RPKI running codes

- RPKI Validator
  - Routinator
    - https://nlnetlabs.nl/projects/rpki/routinator/
  - RIPE NCC Validator
    - https://www.ripe.net/manage-ips-and-asns/resource-management/certification/tools-and-resources
  - FORT
    - https://fortproject.net/validator
  - OctoRPKI/GoRTR
    - https://github.com/cloudflare/cfrpki
- Routers
  - Cisco, Juniper and Quagga