Routing Security and RPKI

Presenters:

Sandra Murphy (sandy@tislabs.com)
Parsons

Channeling:

Randy Bush (Randy@psg.com)

Rob Austein (sra@hactrn.net)

Dragon Research

Michael Elkins (melkins@tislabs.com)

Parsons

Randy/Rob slides

- Based on and some extracted from
- https://psg.com/140220.pdf
- https://nsrc.org/workshops/2014/sanog23-security/raw-attachment/wiki/Agenda/2-4-1.routing-protocols.pdf
- https://nsrc.org/workshops/2014/sanog23-security/raw-attachment/wiki/Agenda/2-4-1.RPKI-Lab.pdf

History of Routing Incidents

- **Apr 1997 – AS 7007 announced routes to all the Internet**
- Apr 1998 – AS 8584 mis-announced 100K routes
- Dec 1999 – AT&T's server network announced by another ISP – misdirecting their traffic (made the Wall Street Journal)
- May 2000 – Sprint addresses announced by another ISP
- Apr 2001 – Flag Telecom in London mis-announced 5K routes
- **Dec 24, 2004 – thousands of networks misdirected to Turkey**
- Feb 10, 2005: Estonian ISP announced a part of Merit address space
- **Sep 9, 2005 – AT&T, XO and Bell South (12/8, 64/8, 65/8) misdirected to Bolivia [the next day, Germany – prompting AT&T to deaggregate]**
- **Jan 22, 2006 – Many networks, including PANIX and Walrus Internet, misdirected to NY ISP (Con Edison)**
- Feb 26, 2006 - Sprint and Verio briefly passed along TTNET (Turkey again) announcements that it was the origin for 4/8, 8/8, and 12/8
- Jul 07, 2007 – Yahoo unreachable for an hour due to mis-origination to L3 from Hanaro Telecom
- **Feb 24, 2008 –Pakistan Telecom announces a part of YouTube's address blocks**
- Mar - Nov 2008 – various addresses within DoD address blocks announced by various ISPs (one in Russia, one in Argentina, others in Australia, Turkey, Indonesia, etc.) for periods up to 3 weeks
- Dec 2008 – Axtel in San Pedro, MX announces unallocated address block, and then sends a large amount of mail traffic (spam).
- Mar 2010 - For three weeks, the address of China's own internal version of the DNS root zone was advertised outside China. This made the altered China version of the root zone visible outside China (Asia, Chile, US, etc.)
- **April 2010 - China Telecom mis-originated about 15% of Internet address blocks**
- Jun 2010 – BGPmon reports bogon IPv6 announcements mis-originated by multiple ISPs to Cogent – no explanation
- Frequent full table leaks, e.g., Sep08 (Moscow), Nov08 (Brazil), Jan09(Russia), Jul 09 (Sweden), … say "when"
- Frequent route leaks: violation of routing policy of provider or peer
- Recent complaints of misbehavior in IRR registration causing routing misbehavior (e.g., RIPE Routing and Anti-Abuse wg discussion Nov 2014)

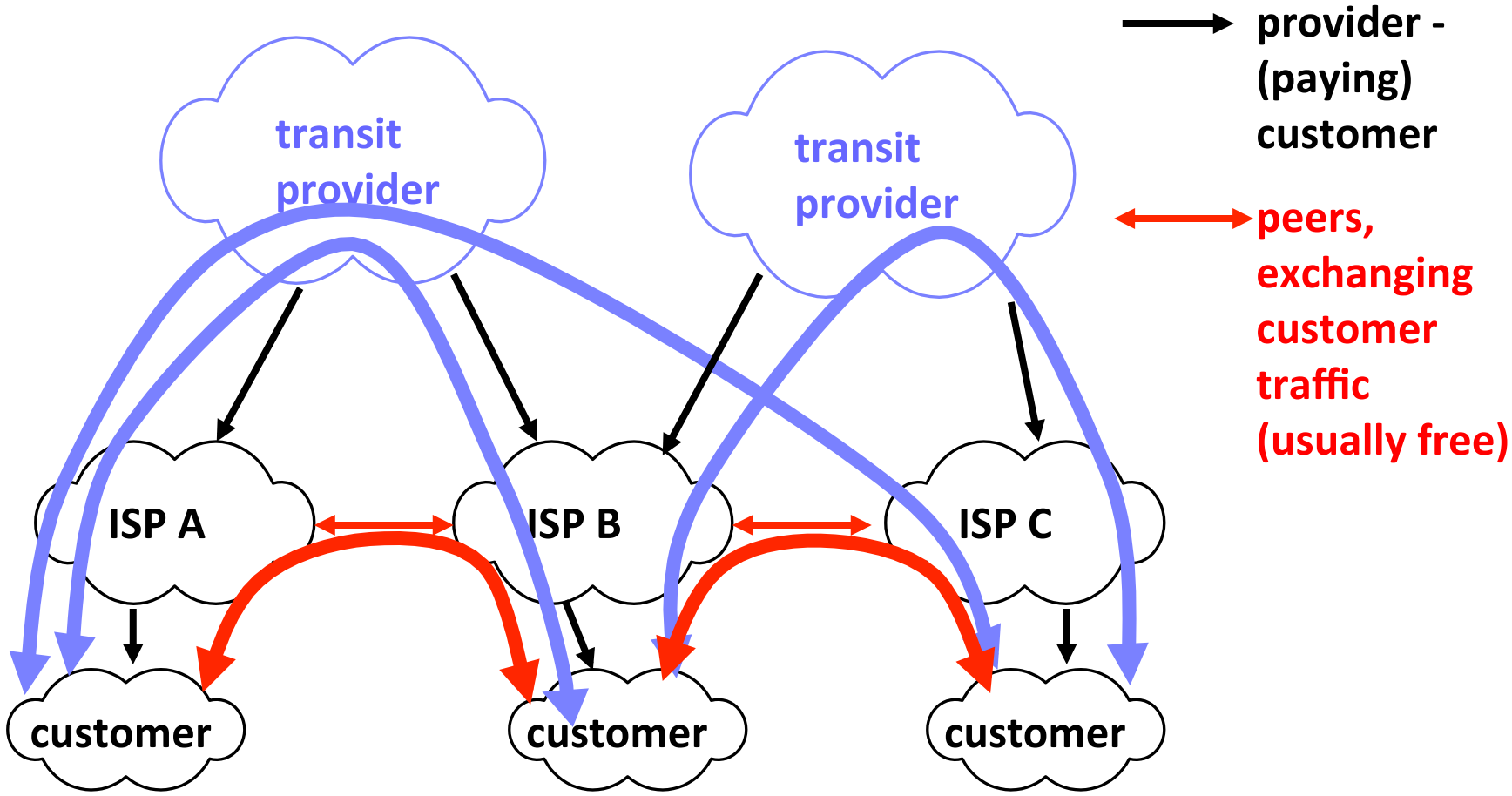
In the Last Two Years

- See Andree Toonk's presentation: https://www.nanog.org/sites/default/files/monday_general_bgp_toonk_63.18.pdf
  - Turkey and 8.8.8.8 (not BGP, example of control of routing)
  - Bitcoin hijack
  - Spammers
    - http://www.bgpmon.net/using-bgp-data-to-find-spammers/ for analysis (that and more)
    - Suggestion of spoofed IRR registration to make it work
  - Syria Telecom hijack of 1400 prefixes
  - Route Leak affecting Cloudflare
- Nov 2013 Renesys about targetted redircention - eg Iceland and Belarus
- April 2014: AS4761 Indostat misoriginates 400K prefixes (damage zone varies)
- Renesys about "attack in progress" – covered by route object, still originating same org's prefixes, prefix now originated by another AS.
- Victim reported on NANOG – announcement of unused space – could be a spammer – Andree Toonk analysis "AS Number 43239…Has started hijacking our IPv4 prefix … 103.20.212.0/22 <- This belongs to us."
- US NOAA-NCDC originated from China for 25 hours
- IRRs – some IRRs (RADB, Level3, Savvis, etc.) have "lots" of "proxy-registered" objects by very rough analysis
- European ISP says China ISP registered prefix belonging to another customer – origination succeeded – valid customer got blamed for spam.
- NANOG Oct 16 2014:"AS6983 is announcing a /24 out of space allocated to AS7922." – Earthlink and Comcast
- March 2015: Tier2 announces v6 /25 in Tier1's v6 /24
- March 2015: Enzu, route leak of more specifics, 7000 prefixes, 280 ASNs impacted
- 12 June 2015: AS4788 Telekom Malaysia leaked 170K prefixes, Level3 propagated, BGP sessions flapped, etc.
- 29 June 2015: NTT propagates route leak of HE prefixes, HE complains
- 30 June2015 : HE propagates hijack: 28,000 prefixes from 4,477 AANs impacted
- July 2015: prefix hijack by AS7514
- Nov 2015: AS9498 (BHARTI Airtel Ltd.) hijack, 16K prefixes, 3K ASNs impacted

So Maybe It's Not So Bad …

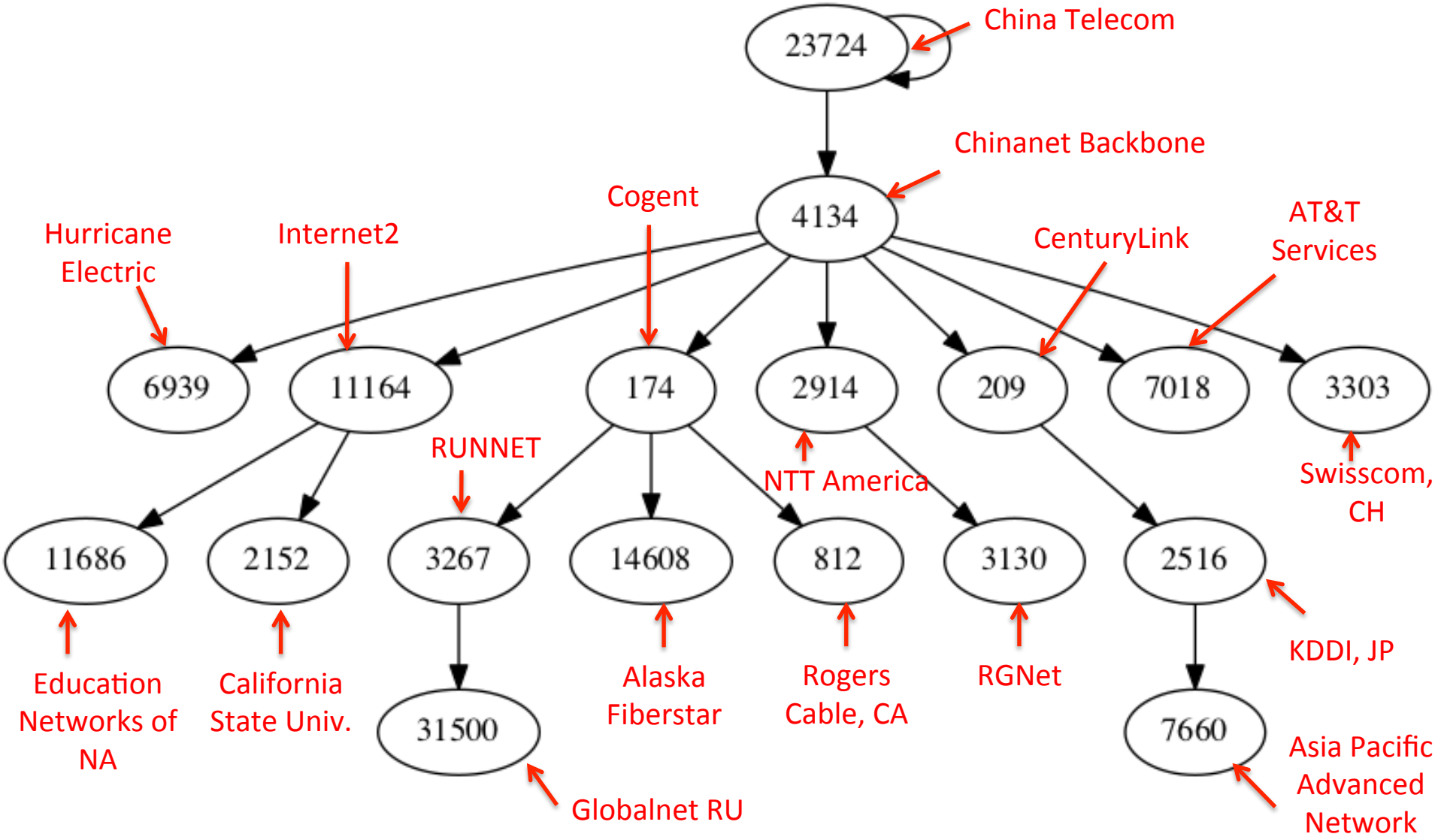
- Response is sometimes under an hour!
  - ***ONLY if someone notices***
  - ***Would you call that RELIABLE networking?***
  - ***Damage to applications and infrastructure***
- These are human mistakes, not attacks
  - ***Anything possible through human error is possible through human intent***
  - ***And some were deliberate***
- There are bigger outages due to hardware and software failures
  - ***But those aren't exploitable deterministically and remotely (mostly)***

AS relationships
(Why On Earth Does is Spread So Far?



Note: Traffic A <-> C does not go through B! (but path exists)

ASNs Propagated China Telecom's Routes



Common Wisdom
"Don't be That Guy (Gal)"

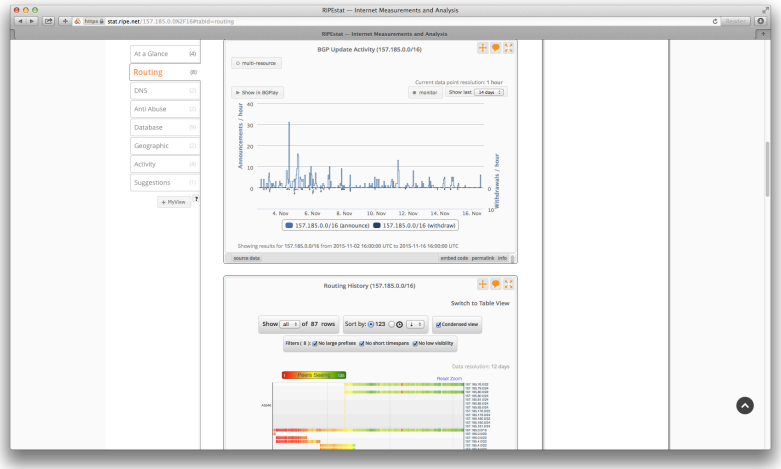
- Filter bogons and martian prefixes
- Inbound prefix filter on customers
  - Use IRR based prefix filters
    - Get your downstreams to create route objects before you turn them up.
  - Get your provisioning teams to validate the prefixes being provided by your downstreams.
  - Use both prefix- and AS_PATH-based filters for your downstreams.
  - fully automate ingress prefix management
- outbound prefix-filter on all transit & peering sessions
  - Outbound AS_Path filter for route leaks (check for transit and peer)
  - Use BGP community based route filtering in outbound policy.
- Max-prefix to catch massive problems
  - use maxprefixes with manual reenable on all ebgp sessions
- No exceptions.

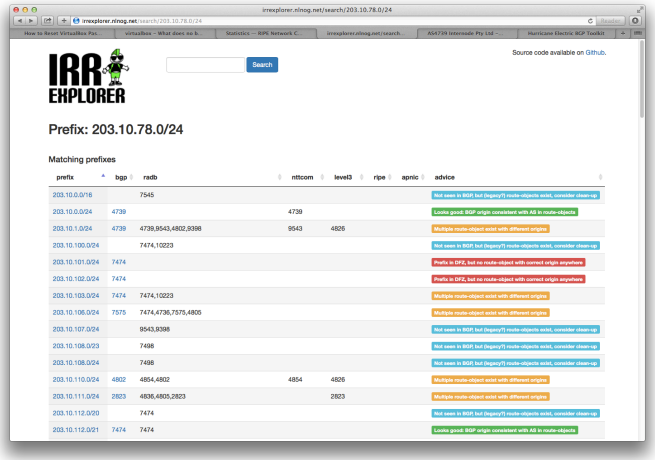
Current Practice: Internet Routing Registry based filtering

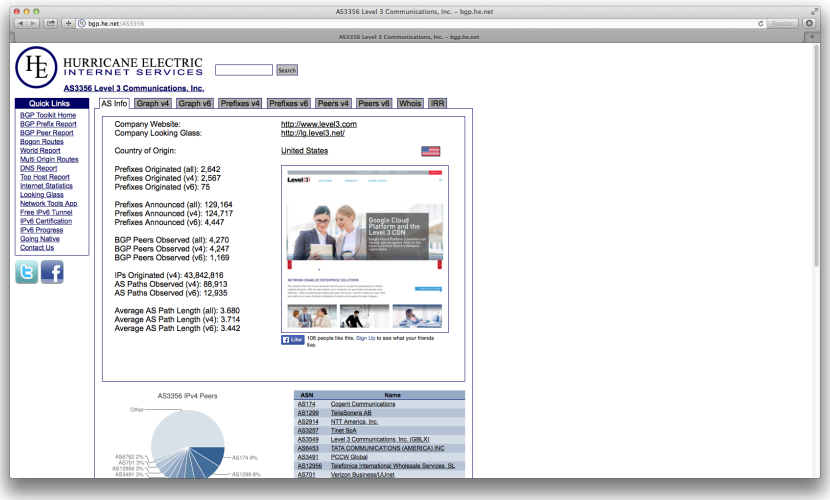
- IRRs are databases
  - Register an AS's routing policy
  - route objects – prefixes the AS asserts it may originate
- 30+ IRRs, some associated with RIRs, some not
- There is a trust model – RFC2725 (allocate only out of your allocation, can create route object only for your AS and your prefix)
- RIR based IRRs can tie allocation to registration of objects
  - Know whether registrant is authorized to speak for prefix/AS
  - CAN follow RFC2725 for resources in their regions, CAN NOT for outside region
- Non RIR based IRRs (RADB, Level3, Savvis,…) can not tell if registrant is authorized
  - Can NOT follow RFC2725
- Trust model doesn't scale – channel security
- Use doesn't scale. See Jared Mauch (260K lines of prefix list, 96% of config is prefix lists, 5 min commit times) Mar 14 IEPG
  - http://iepg.org/2014-03-02-ietf89/ietf89_iepg_jmauch.pdf
  - In Jun 2015, NTT reports config file has grown another 100K lines

Good Tools Abound

- http://bgp.he.net
- https://stat.ripe.net
- http://irrexplorer.nlnog.net
- http://www.routeviews.org
  - https://github.com/cmu-sei/bgpuma



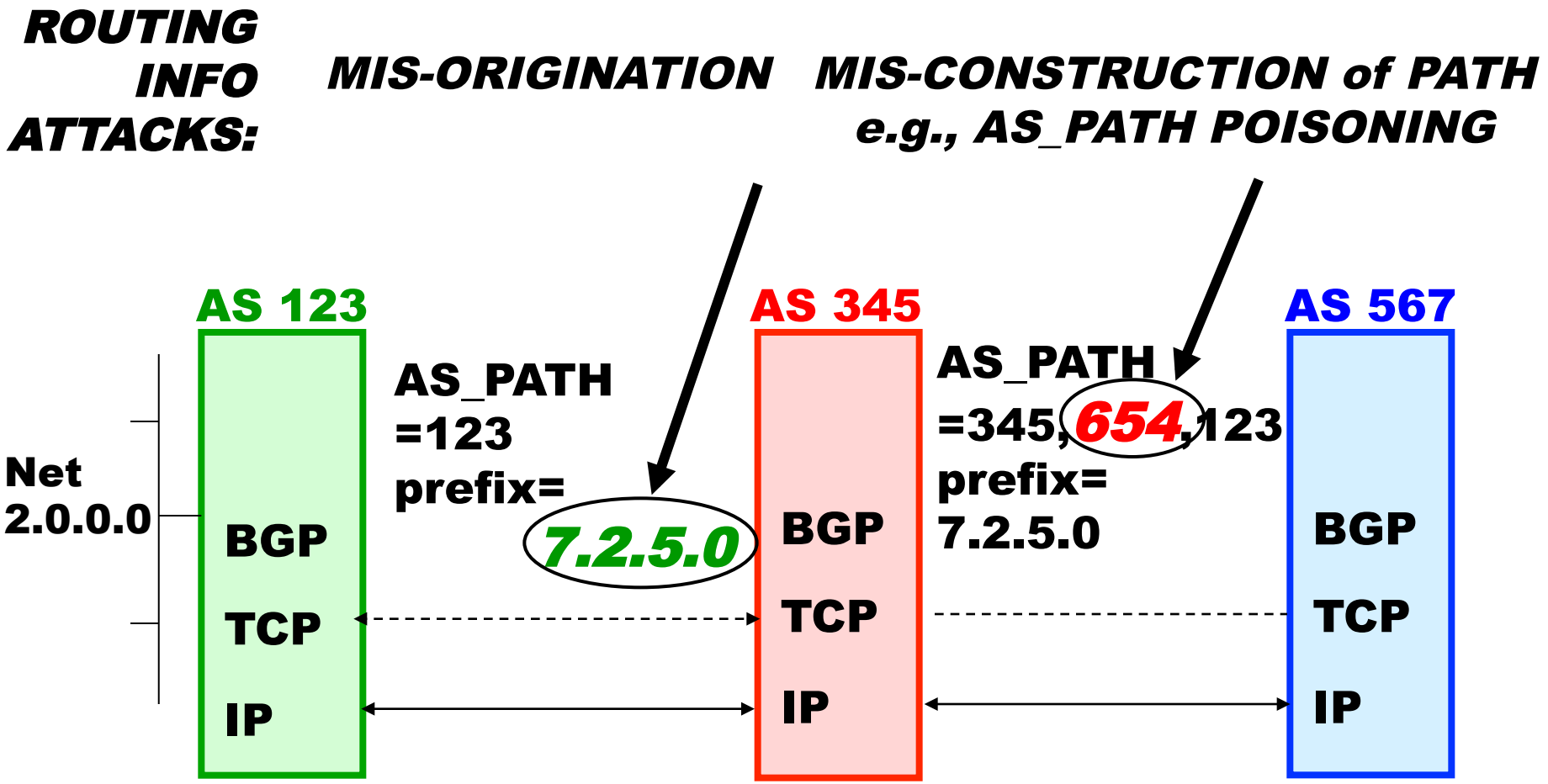




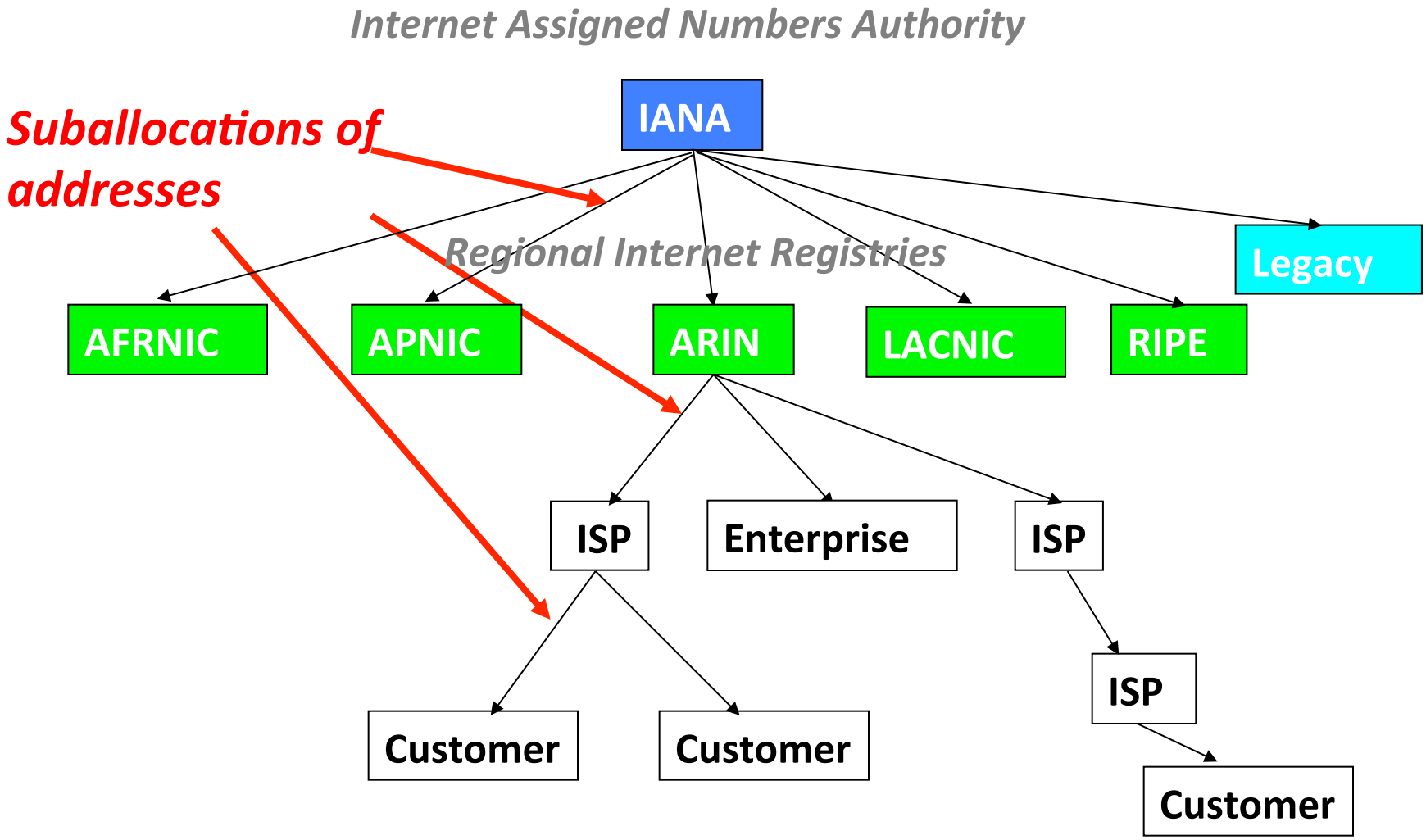
A Stronger Solution in Three Parts

- Prefix Holder: Who has the right to use a prefix?
  - Resource Public Key Infrastructure – RPKI
- Origin Validation: Who is authorized to originate a route to a prefix?
  - Based on the RPKI: only the prefix holder can say
  - Prevent mis-originations – common hijacks
- Path Validation: Who has the right to propagate a route?
  - Based on the RPKI: only the AS who propagates can say
  - Prevent path problems: bogus first hop, maybe route leaks

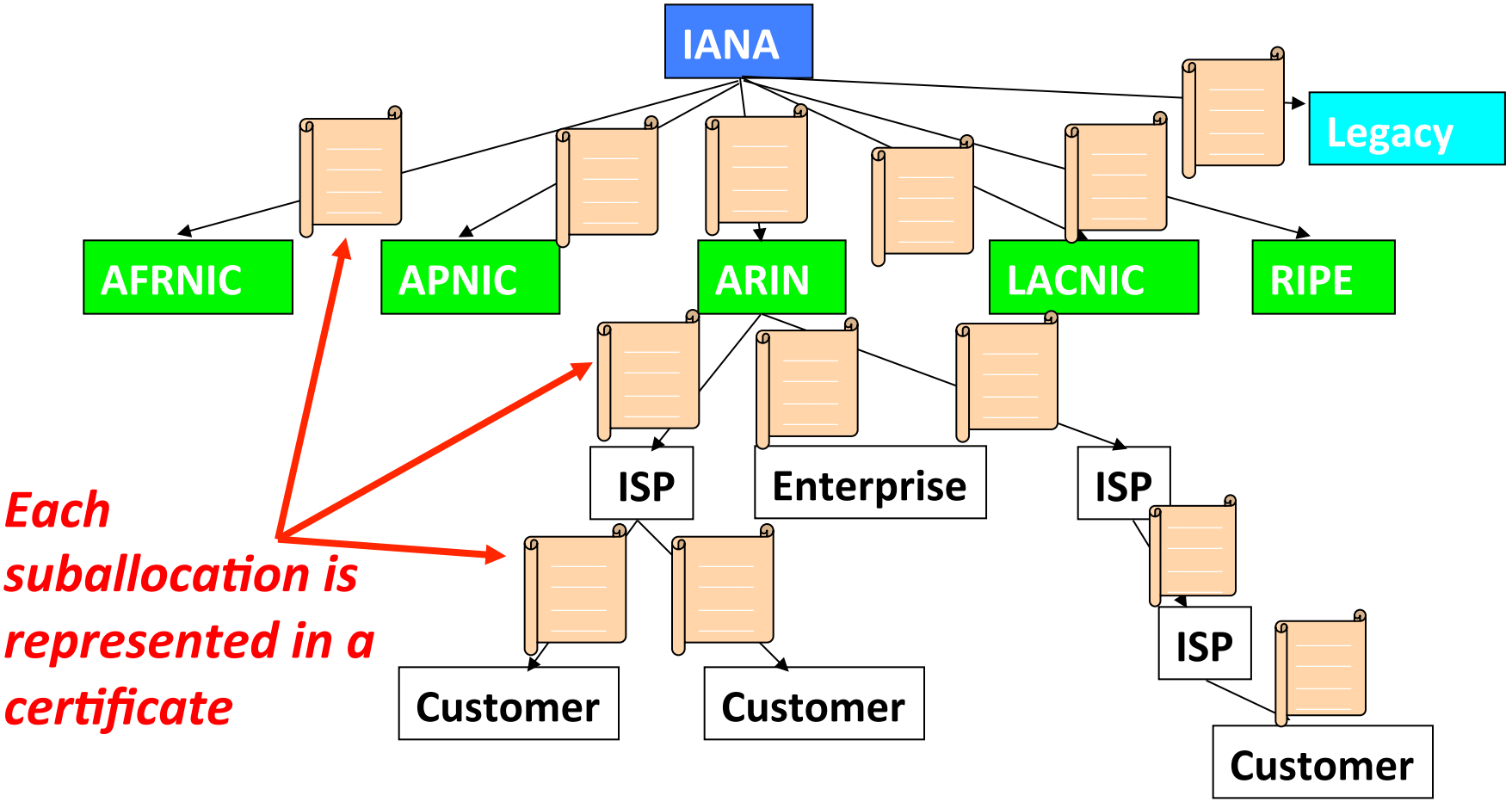
BGP Vulnerabilities



Just Who Does Hold an Address?

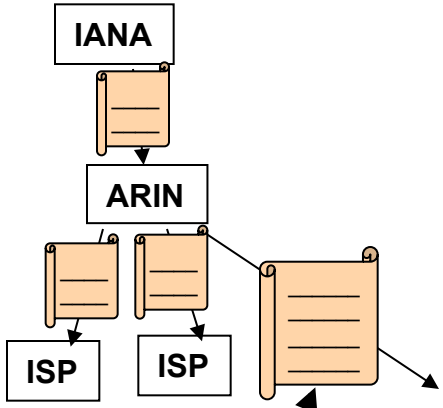


RPKI - Resource Certificates



Resource **certificate, not identity certificate**

Origin Validation: Certs & Route Origin Authorization



Sign a Route Origin Authorization (ROA) for your address space
Your certificate validates the signature

Enterprise

Certificate lists the addresses you hold and who gave them to you

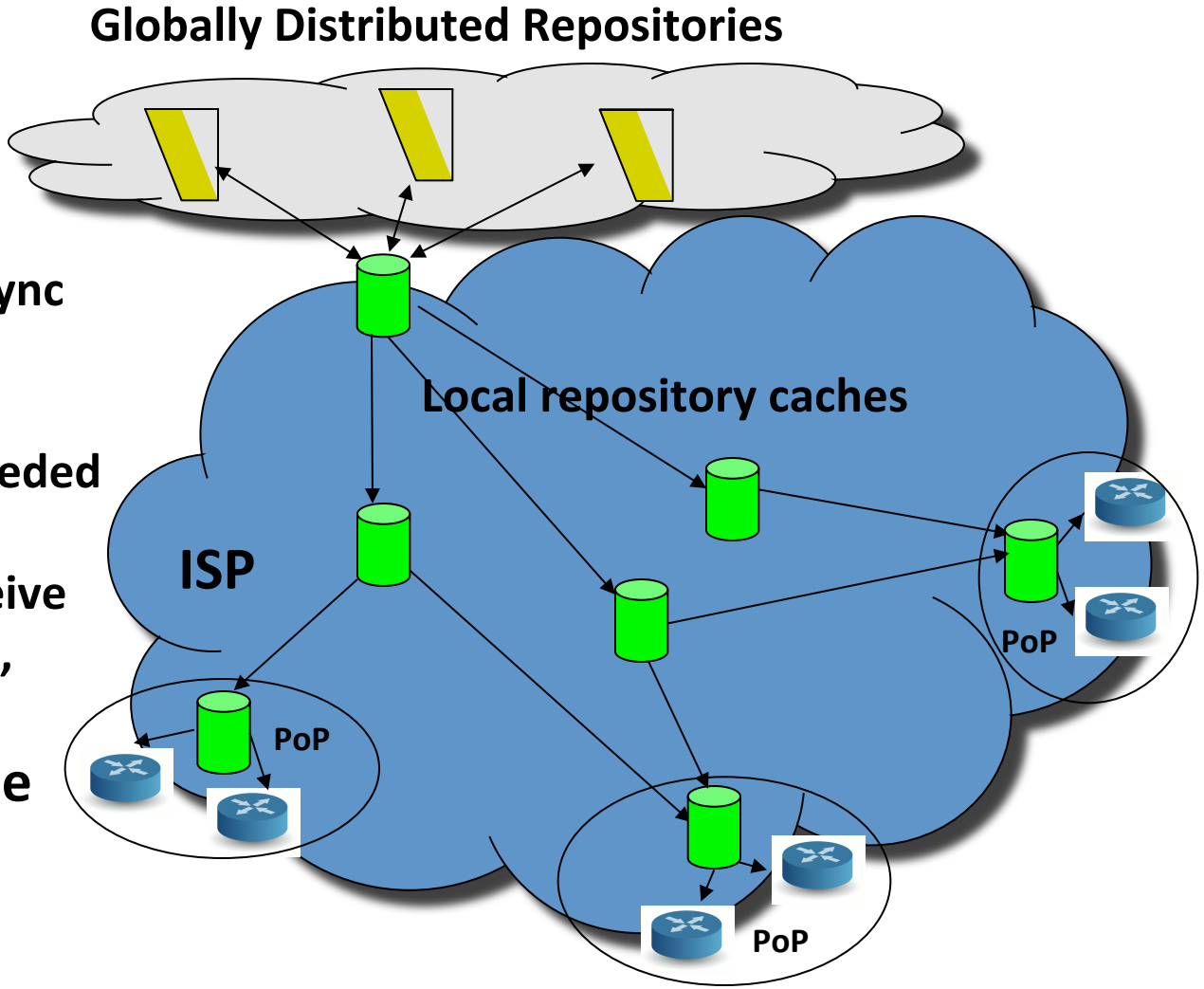
CA certificate
Key: EnterpriseKey
Signed by: ARIN
Addresses: 10.2/16
(10.2.0.0 – 10.2.255.255)

ROASignedObject
Signed by: EnterpriseKey
Addresses: someofyouraddresses
Valid Origin: some ASn

The ROA lists the valid origins for those addresses

RPKI Architecture in Single AS

- Local cache is kept in sync with global distributed repositories
- Local cache does all needed crypto
- Routers need only receive list of (authorized origin, address) pairs
- *N*O* crypto in the routers



Two Sides of This

Thinking "Wow, Lots of WORK!"? Don't Panic

- **Securing routes to your addresses**
  - Get certificates for your address space
  - Sign ROAs
  - Maintain a CA repository
  - Create certificates for your customers
    - If you give them addresses
- *Think of this as signing the back of your credit card*

- **Securing routes to others' addresses**
  - Retrieve ROAs from other CA repositories
  - Validate received routes against the RPKI data
- *Think of this as checking the back of a credit card tendered to you for a sale*

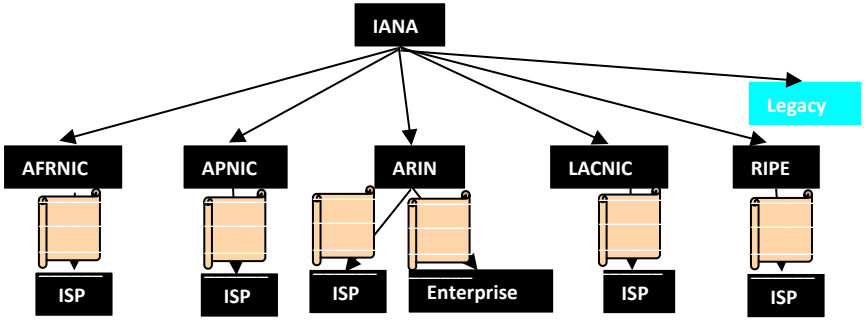
Hosted service

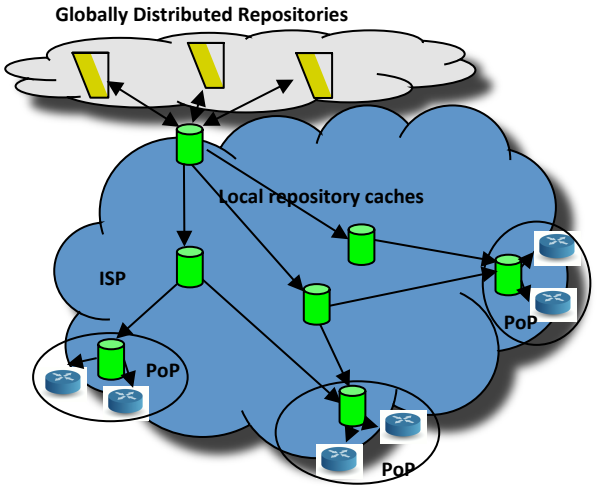
Outsourced service

Offline retrieval & crypto

Status on Multiple Fronts: Specs

- IETF SIDR RFCs
  - 24 documents published as RFCs

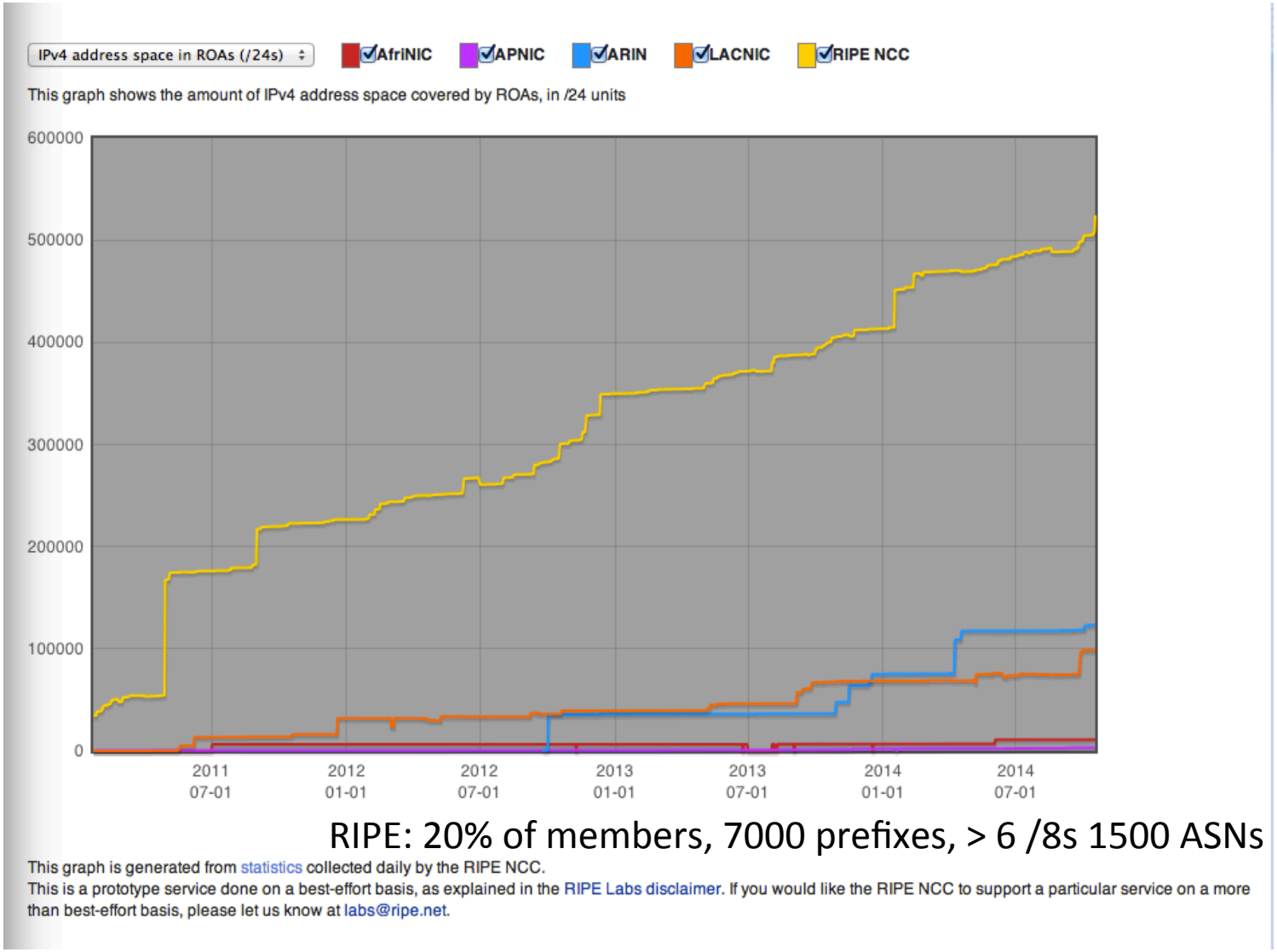




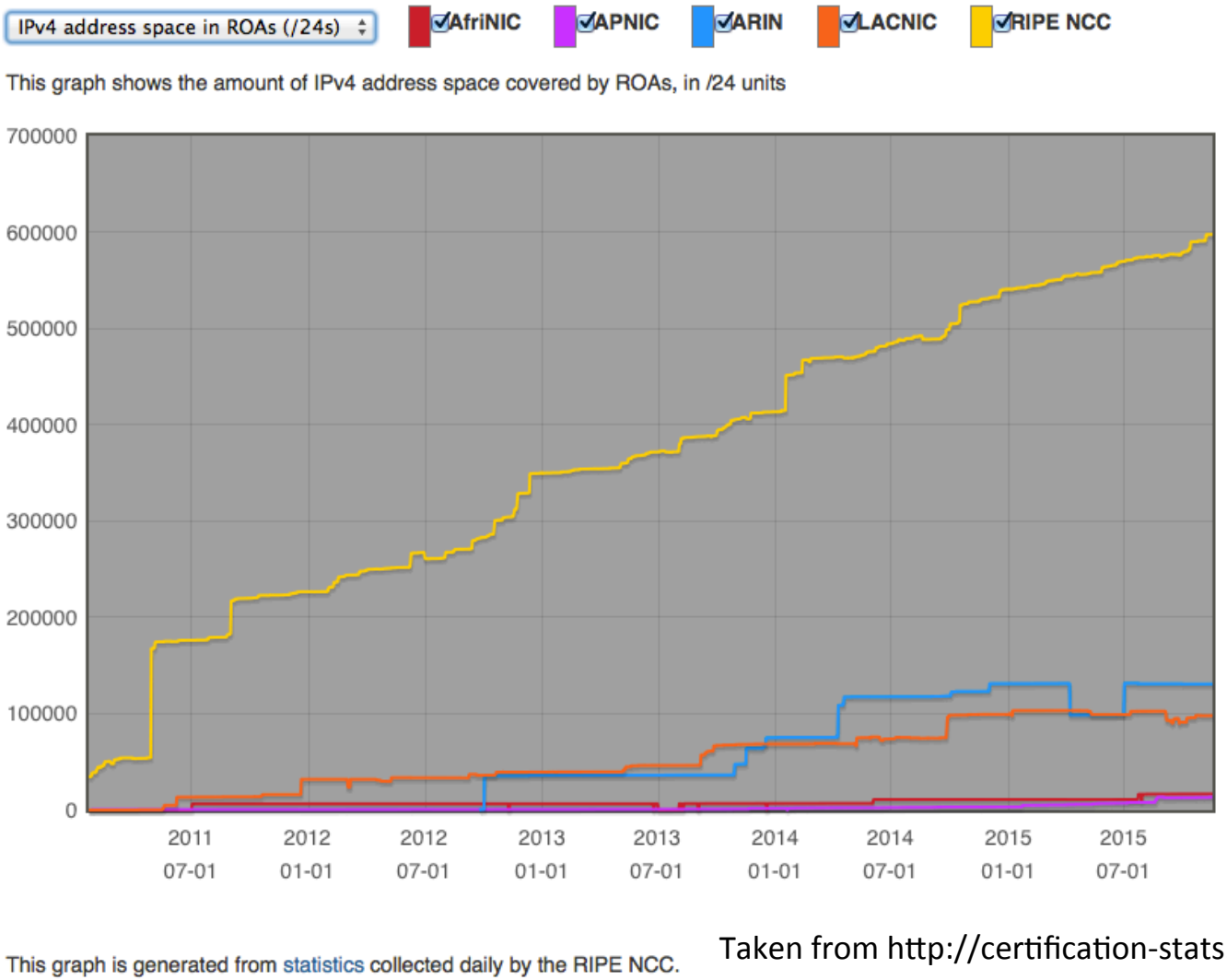
Certs, ROAs, certificate policy, repository structure, certificate management protocol (aka “up/down”), etc.

route validation, RPKI-to-router protocol, common operations, MIB, etc.

Status on Multiple Fronts: RPKI



Status on Multiple Fronts - RPKI



Taken from http://certification-stats.ripe.net/

RPKI stats and monitors

- http://www.labs.lacnic.net/rpkitools/looking_glass/
- http://www-x.antd.nist.gov/rpki-monitor/
- http://certification-stats.ripe.net/
- http://rpki.surfnet.nl/index.html
- http://www.hactrn.net/opaque/rcynic/

Status on Multiple Fronts: Origin Validation

- Cisco:
  - High-end & mid-range routers running IOS-XR
    - Minimum release XR 4.2.1
  - Access/Enterprise routers running IOS-XE
    - Minimum release XE 3.5
- Juniper
  - Juniper provides official support for RPKI since release 12.2.
- Alcatel-Lucent

Origin Validation Configuration

- See examples at RIPE
  https://www.ripe.net/manage-ips-and-asns/resource-management/certification/router-configuration
- JunOS
  - First: Set up communication with local RPKI cache
  - Second: Assign a local-preference based on the RPKI validity attribute

```
policy-options {
   policy-statement validation {
      term valid {
         from {
            protocol bgp;
            validation-database valid;
         }
         then {
            validation-state valid;
            community add origin-validation-state-valid;
            next policy;
         }
      }
 }
}
```

Origin Validation Configurations

- See examples at https://www.ripe.net/manage-ips-and-asns/resource-management/certification/router-configuration
- CISCO
  - First: Set up communication with local RPKI cache
  - Second: Assign a local-preference based on the RPKI validity attribute

```
!
route-map rpki-loc-pref permit 10
 match rpki invalid
 set local-preference 90
!
route-map rpki-loc-pref permit 20
 match rpki not-found
 set local-preference 100
!
route-map rpki-loc-pref permit 30
 match rpki valid
 set local-preference 110
```

More CISCO Config Options

Fairly Secure

```
route-map validity-0
  match rpki valid
  set local-preference 100
route-map validity-1
  match rpki not-found
  set local-preference 50
! invalid is dropped
```


Paranoid

```
route-map validity-0
  match rpki valid
  set local-preference 110
! everything else dropped
```

Junos Show Validation

```
195.24.160.0/19    *[BGP/170] 00:03:59, MED 2000, localpref 50, from 87.238.63.5
                          AS path: 3356 3549 4788 6939 39648 I, validation-state: unverified
                        > to 87.238.63.56 via ae0.0
                        [BGP/170] 00:05:24, MED 0, localpref 50, from 87.238.63.2
                          AS path: 3356 3549 4788 6939 39648 I, validation-state: unverified
                        > to 87.238.63.56 via ae0.0
                        [BGP ] 01:16:00, MED 25245, localpref 100
                          AS path: 3549 4788 6939 39648 I, validation-state: unverified
                        > to 64.210.69.85 via xe-1/1/0.0
```

Cisco Show Validation

Valid!

```
r0.sea#show bgp 192.158.248.0/24
BGP routing table entry for 192.158.248.0/24, version 3043542
Paths: (3 available, best #1, table default)
 6939 27318
    206.81.80.40 (metric 1) from 147.28.7.2 (147.28.7.2)
      Origin IGP, metric 319, localpref 100, valid, internal,
best
      Community: 3130:391
      path 0F6D8B74 RPKI State valid
 2914 4459 27318
    199.238.113.9 from 199.238.113.9 (129.250.0.19)
      Origin IGP, metric 43, localpref 100, valid, external
      Community: 2914:410 2914:1005 2914:3000 3130:380
      path 09AF35CC RPKI State valid
```

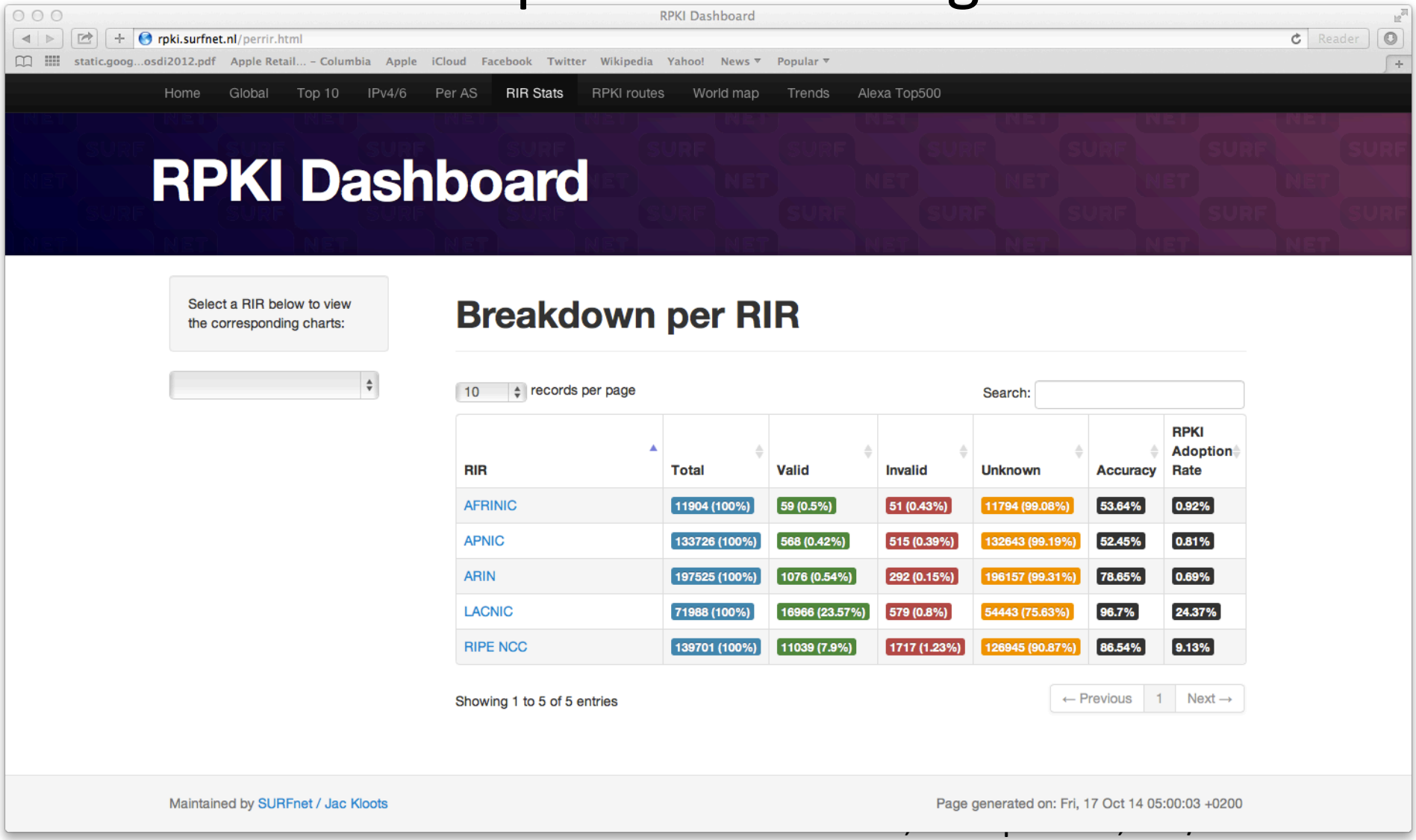
DRL RPKI Origin Validation 60

Invalid!

```
r0.sea#show bgp 198.180.150.0
BGP routing table entry for 198.180.150.0/24, version 2546236
Paths: (3 available, best #2, table default)
  Advertised to update-groups:
     2          5          6          8
  Refresh Epoch 1
  1239 3927
    144.232.9.61 (metric 11) from 147.28.7.2 (147.28.7.2)
      Origin IGP, metric 759, localpref 100, valid, internal
      Community: 3130:370
      path 1312CA90 RPKI State invalid
```

DRL RPKI Origin Validation 61

Status on Multiple Fronts: Origin Validation



Origin Validation Deployment

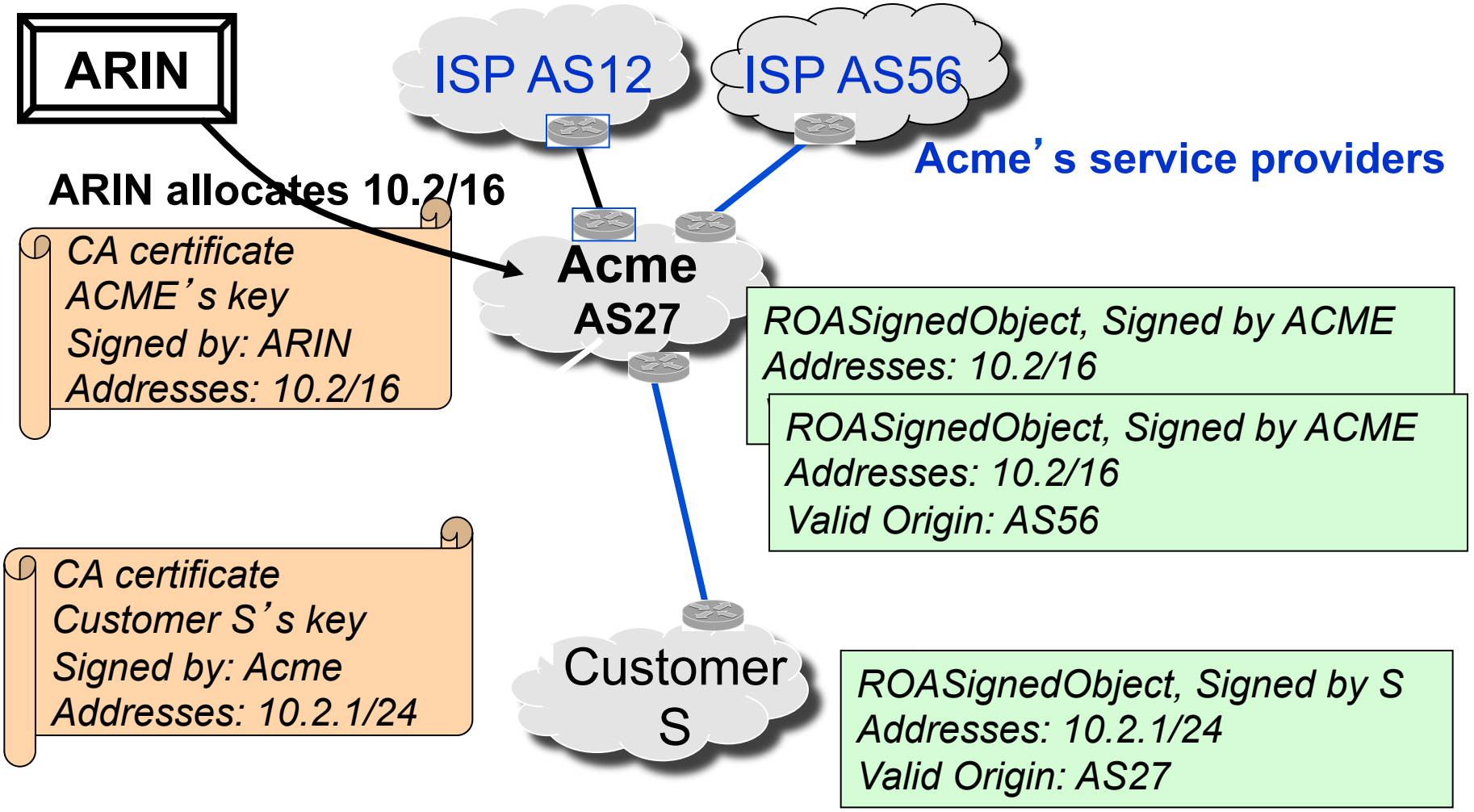
- IETF has used rpki.net for several IETFs in a row (sees few invalids)
- IXPs
  - Sep 2015: AMS-IX beginning to offer RPKI based filtering in their route servers
  - Oct 2014: French IXP announces they have begun to use RPKI for filtering
  - IXPs in RIPE have suggested RPKI as service for members
- Esnet doing RPKI based origin validation – pref valid
- Major European ISP testing in internal lab, requests for features
- Rpki.net virtual testbed and AltCA – a dozen or so active participants (Comcast, ATT, ESnet, LACNIC, European folk, Google)
- FCC CSRIC III WG 6 report 2013 "Cautious, staged deployment of RPKI Route Origin Validation"
- French ANSSI agency 2014 recommends use of RPKI and ROAs

Current Issues

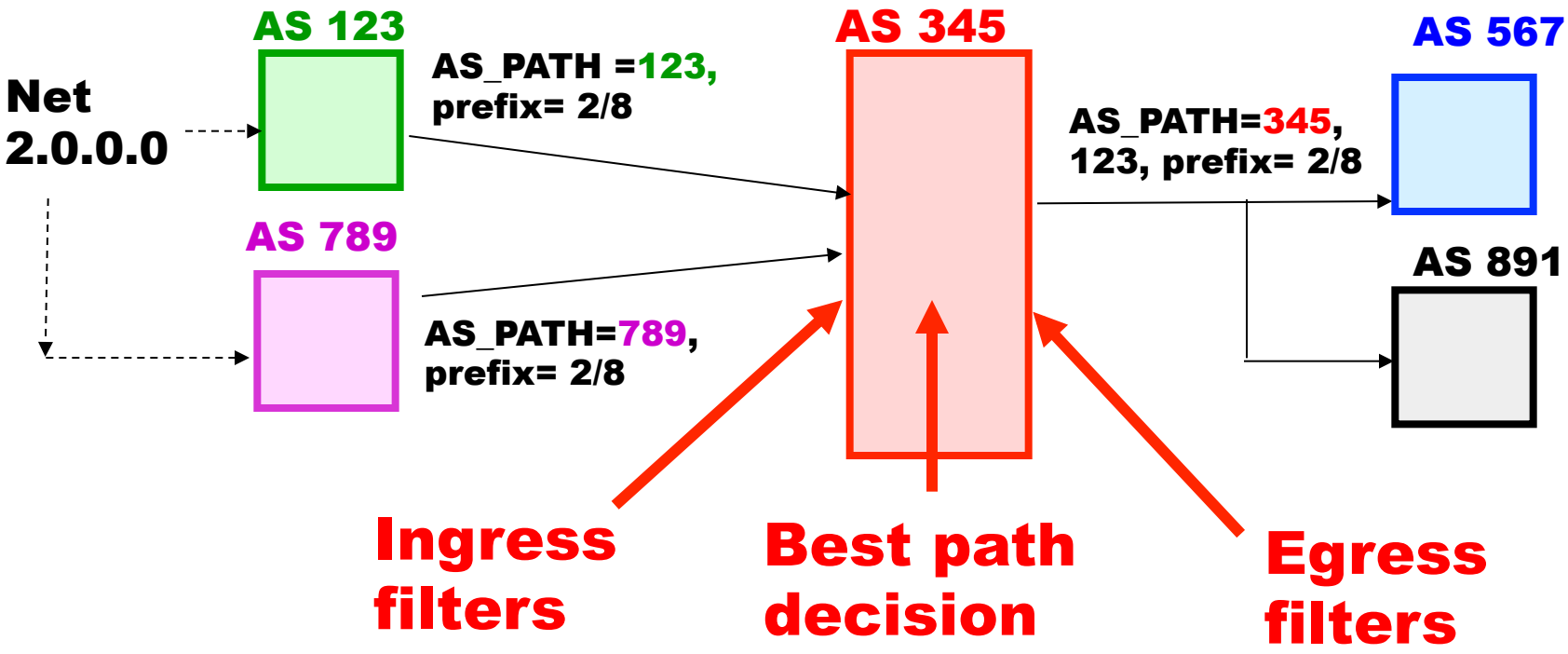
- Technical
  - Legacy space (44% of orgs in ARIN, 56% of addresses)
  - Rsync performance
  - Validation reconsidered
  - Legacy space
- Non-technical
  - Mis-use of hierarchical authority (errors, court orders)
  - Impact on routing from RIR actions, service level, etc.
  - The usual problems with new technology – effort and cost –
    - and usual problem with new security technology – hard for users to see immediate direct benefit –
    - and infrastructure technology – no one is in charge
- See Wes George talk at https://www.nanog.org/sites/default/files/wednesday_george_adventuresinrpki_62.9.pdf

Extra slides

The Way This Goes...



BGP Process



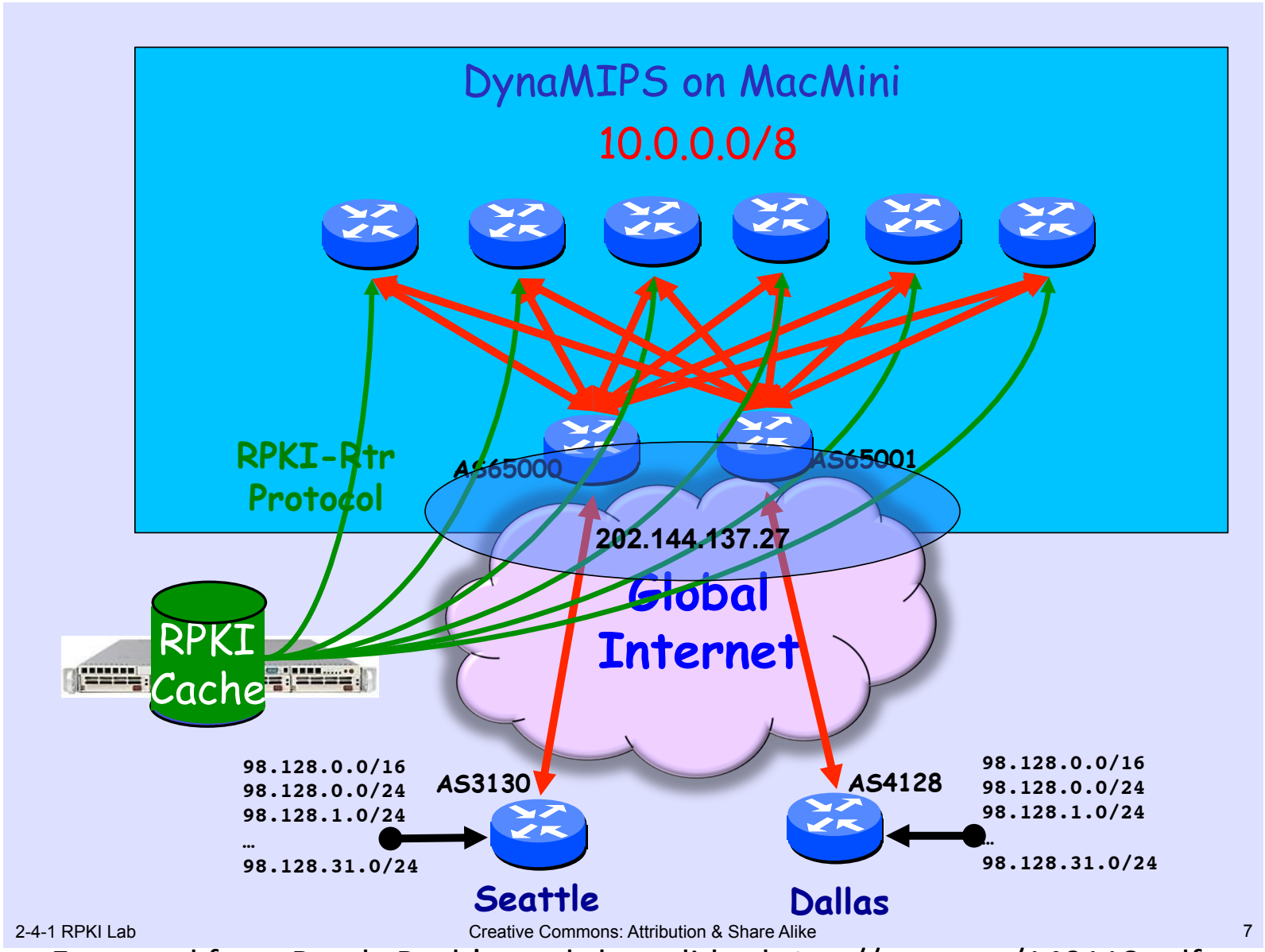
- BGP receives many routes to the same prefix
- Ingress filter decides what routes to consider
- Decision process picks just one best route
- Egress filter decides what neighbors receive an update

IRR Based Filters

- Registries could be used to check NLRI origination, AS_PATHs, etc.
- Level of protection from use of registry relies on registry containing complete and accurate information, including peering and policy
- Communication with registry would have to be protected
- IRRs are known to be inaccurate, incomplete, stale, and many have little to no security applied

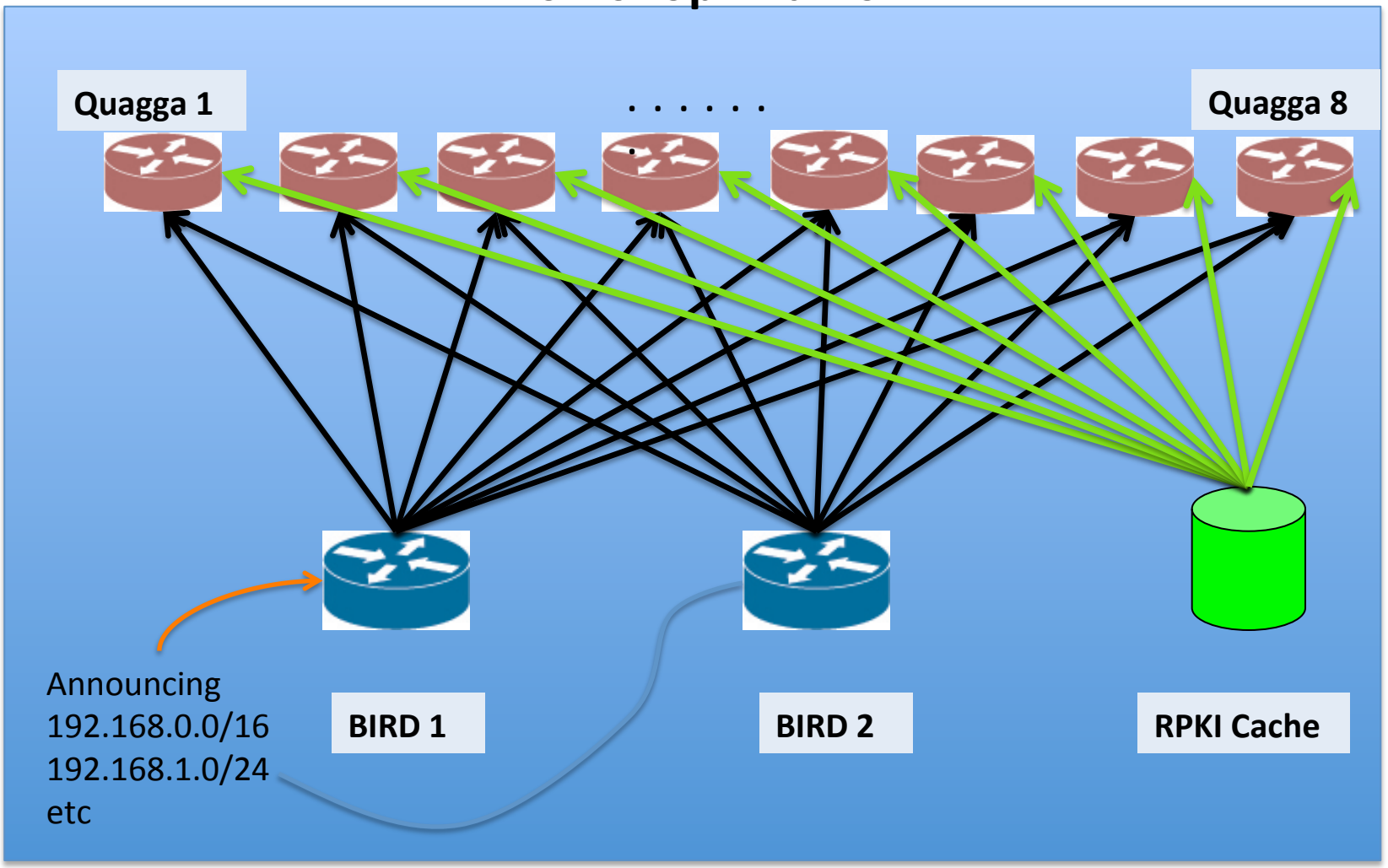
Workshop in a Box

Randy Bush's World Traveled Workshop Set-Up

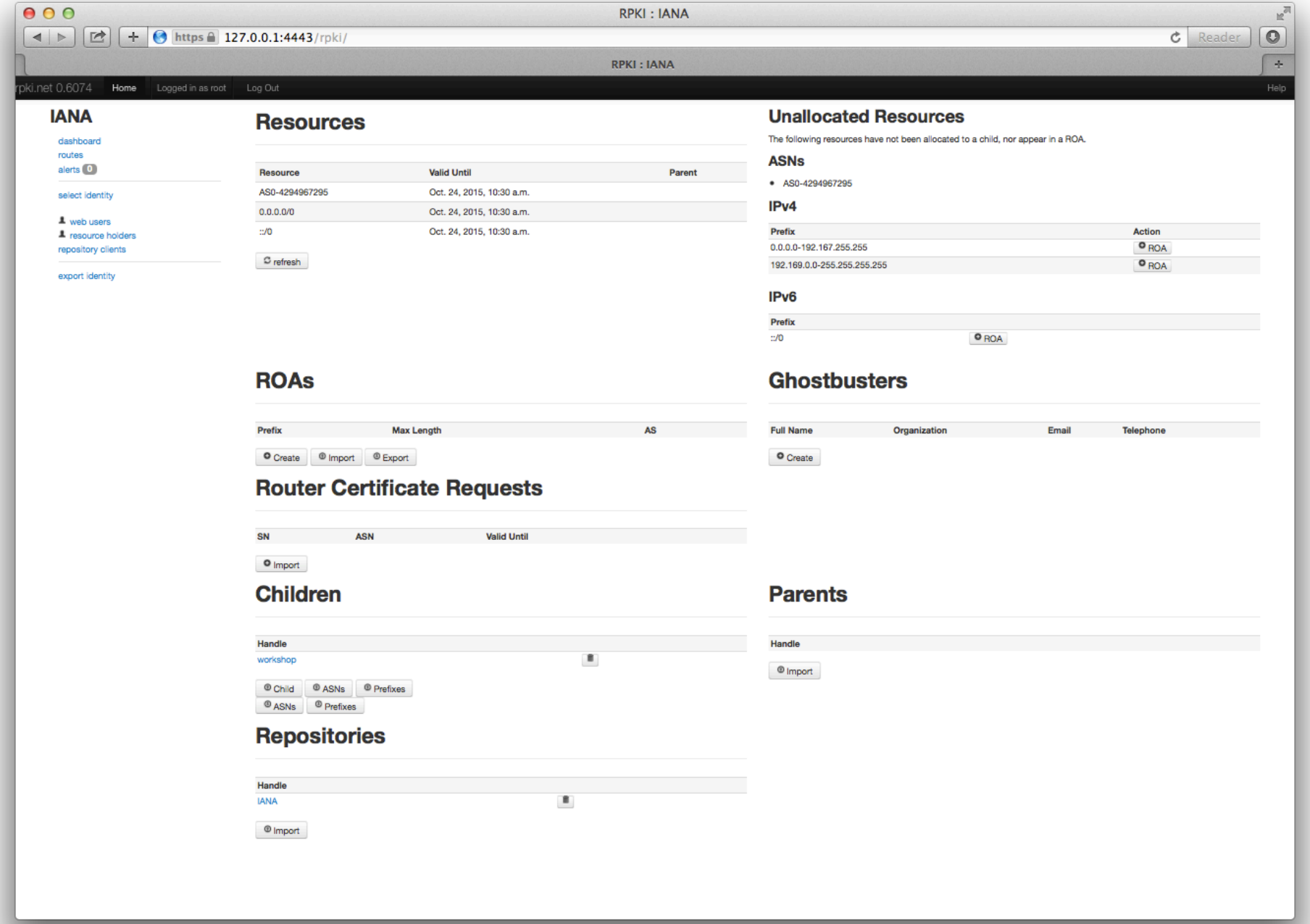


Extracted from Randy Bush's workshop slides https://psg.com/140118.pdf

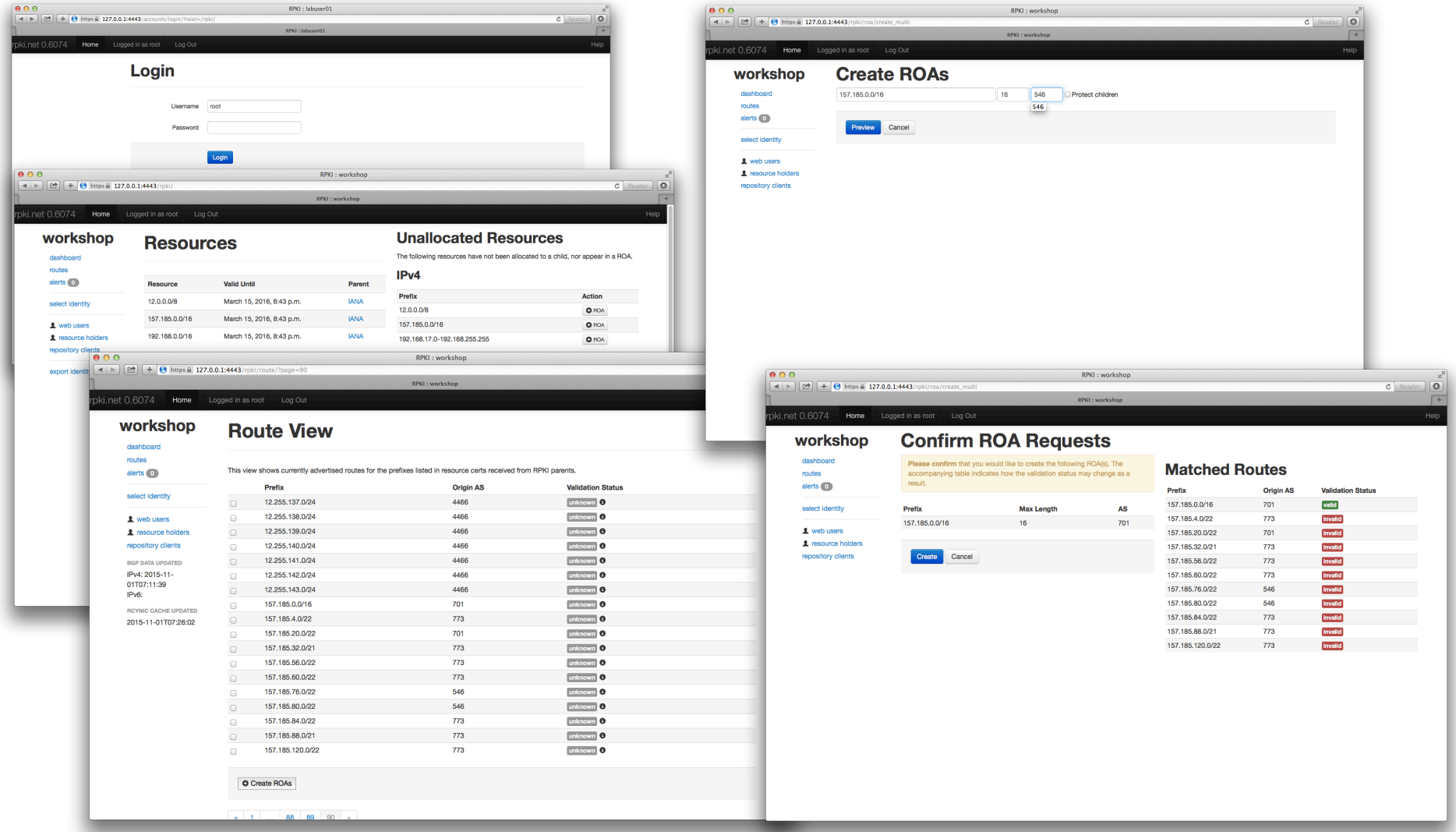
Workshop in a Box

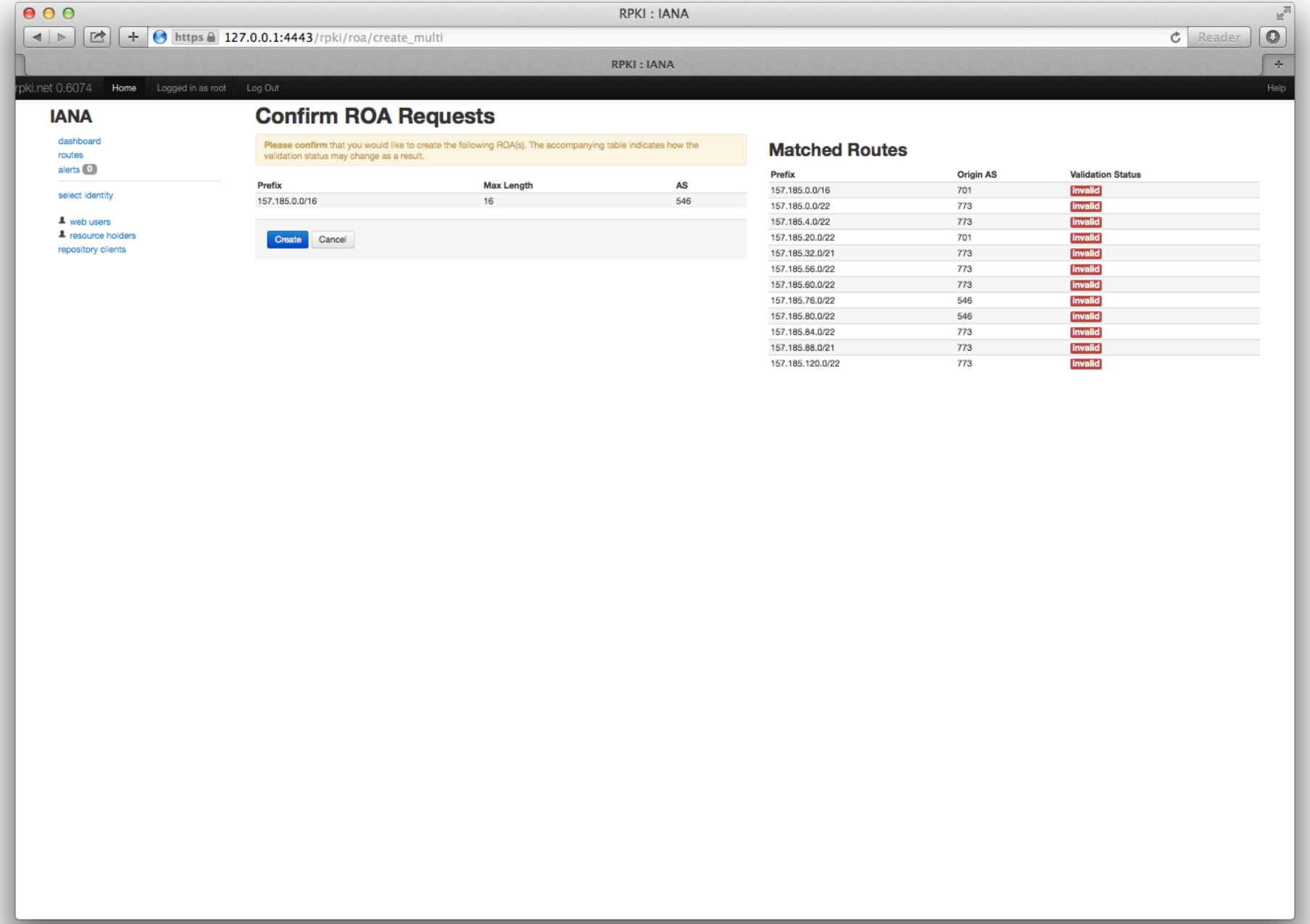


VM totally self-contained environment – no outside dependencies
Comes with local trust anchor so you can generate certs for your own prefixes
Use for experimentation, training, testing, whatever


RPKI : IANA
https 127.0.0.1:4443/rpki/
Reader
RPKI : IANA
rpki.net 0.6074
Home
Logged in as root
Log Out
Help
IANA
dashboard
routes
alerts 0
select identity
web users
resource holders
repository clients
export identity
Resources
Resource
Valid Until
Parent
AS0-4294967295
Oct. 24, 2015, 10:30 a.m.
0.0.0.0/0
Oct. 24, 2015, 10:30 a.m.
::/0
Oct. 24, 2015, 10:30 a.m.
refresh
Unallocated Resources
The following resources have not been allocated to a child, nor appear in a ROA.
ASNs
AS0-4294967295
IPv4
Prefix
Action
0.0.0.0-192.167.255.255
ROA
192.169.0.0-255.255.255.255
ROA
IPv6
Prefix
::/0
ROA
ROAs
Prefix
Max Length
AS
Create
Import
Export
Ghostbusters
Full Name
Organization
Email
Telephone
Create
Router Certificate Requests
SN
ASN
Valid Until
Import
Children
Handle
workshop
Child
ASNs
Prefixes
ASNs
Prefixes
Parents
Handle
Import
Repositories
Handle
IANA
Import

Workshop GUI




RPKI : IANA
https 127.0.0.1:4443/rpki/roa/create_multi
Reader
RPKI : IANA
rpki.net 0.6074
Home
Logged in as root
Log Out
Help
IANA
dashboard
routes
alerts 0
select identity
web users
resource holders
repository clients
Confirm ROA Requests
Please confirm that you would like to create the following ROA(s). The accompanying table indicates how the validation status may change as a result.
Prefix
Max Length
AS
157.185.0.0/16
16
546
Create
Cancel
Matched Routes
Prefix
Origin AS
Validation Status
157.185.0.0/16 701 invalid
157.185.0.0/22 773 invalid
157.185.4.0/22 773 invalid
157.185.20.0/22 701 invalid
157.185.32.0/21 773 invalid
157.185.56.0/22 773 invalid
157.185.60.0/22 773 invalid
157.185.76.0/22 546 invalid
157.185.80.0/22 546 invalid
157.185.84.0/22 773 invalid
157.185.88.0/21 773 invalid
157.185.120.0/22 773 invalid