

BGP Origin Validation

2013.02.23 / APOPS Singapore

Randy Bush <randy@psg.com>

Rob Austein <sra@isc.org>

Philip Smith <pfs@apnic.net>

Steve Bellovin <smb@cs.columbia.edu>

And a cast of thousands! Well, dozens :)

Agenda

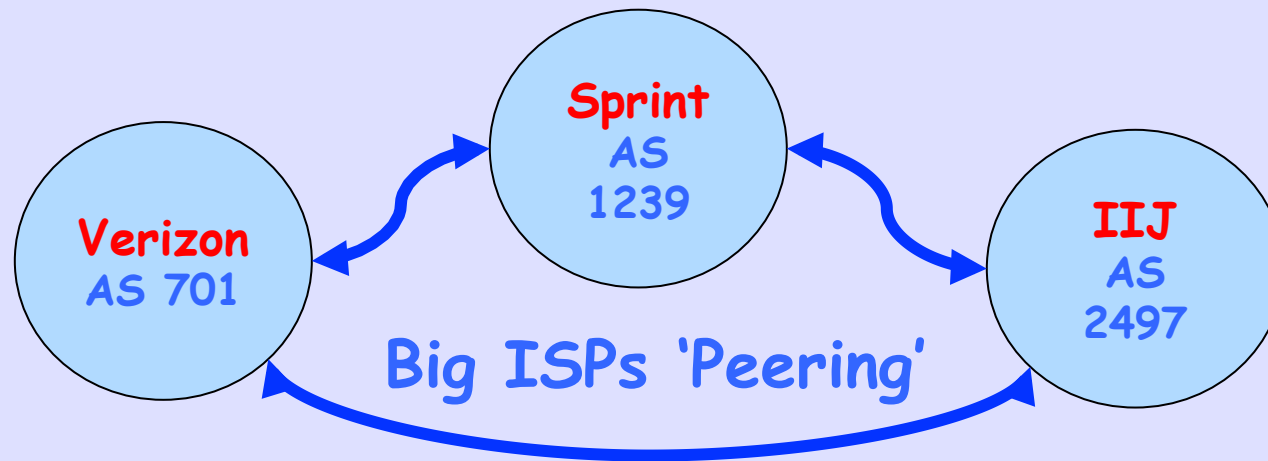
- This Preso
 - Some Technical Background
 - Mis-Origination - YouTube Incident
 - The RPKI - Needed Infrastructure
 - BGP Origin Validation
- RIPE/NCC-Hosted GUI Demo
- Common Errors
- RPKI Child Under RIPE/NCC
- Build Your Own RPKI Instance
- Playing at Scale with AutoNetKit

This is Not New

- 1986 - Bellovin & Perlman identify the vulnerability
- 1999 - National Academies study called it out
- 2000 - S-BGP - X.509 PKI to support Secure BGP - Kent, Lynn, et al.
- 2003 - NANOG S-BGP Workshop
- 2006 - ARIN & APNIC start work on RPKI. RIPE starts in 2008.
- 2009 - RPKI Open Testbed and running code in test routers

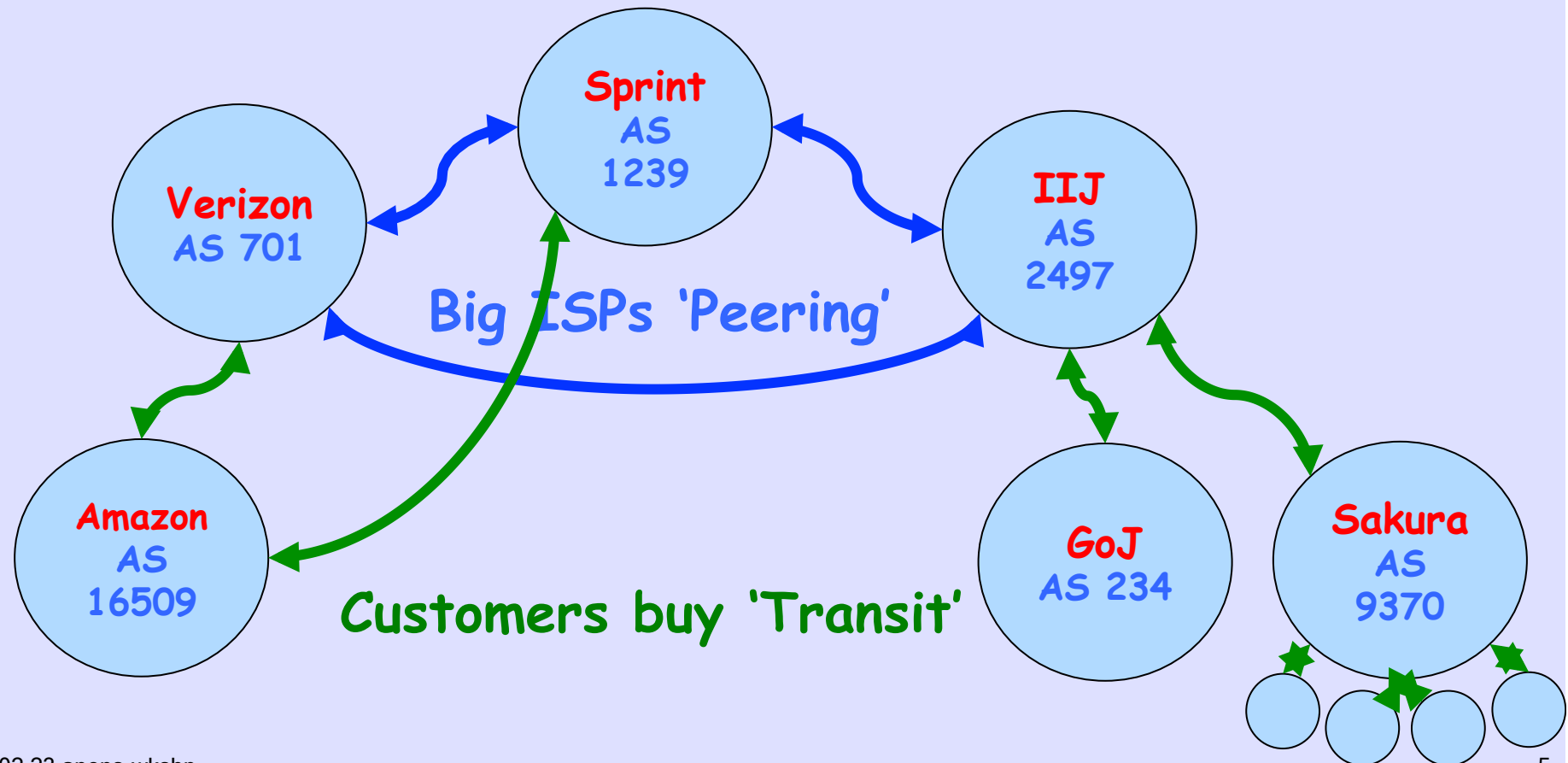
What is an AS?

An ISP or End Site

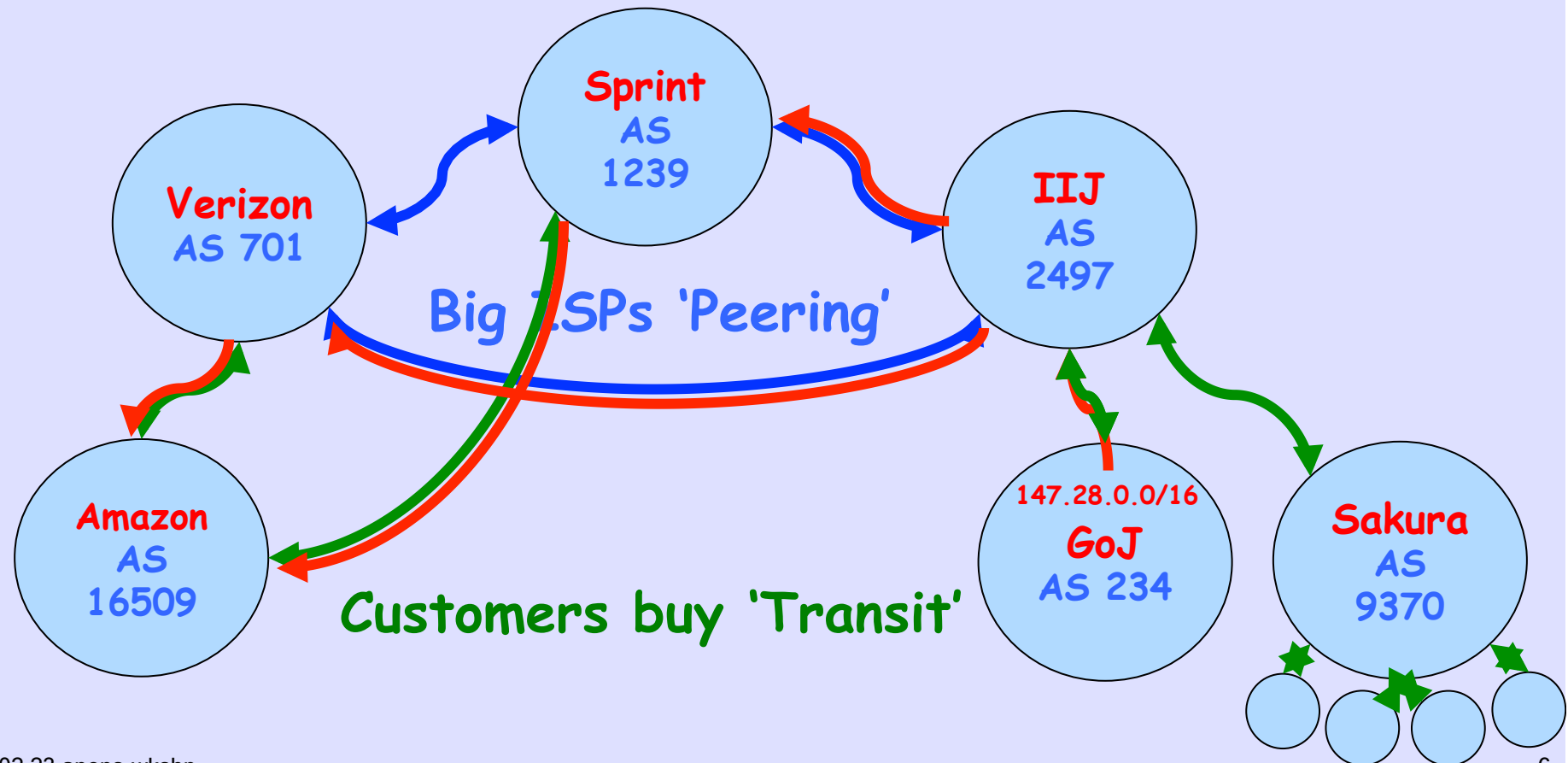


What is an AS?

An ISP or End Site



An IP Prefix is Announced & Propagated



From Inside a Router

BGP routing table entry for **147.28.0.0/16**



Of Course it's Uglier ☺

```
r1.iad#sh ip bgp 147.28.0.0/16
```

```
BGP routing table entry for 147.28.0.0/16, version 21440610
```

```
Paths: (2 available, best #1, table default)
```

```
Advertised to update-groups:
```

```
1
```

```
Refresh Epoch 1
```

```
16509 1239 2497 234
```

```
144.232.18.81 from 144.232.18.81 (144.228.241.254)
```

```
Origin IGP, metric 841, localpref 100, valid, external, best
```

```
Community: 3297:100 3927:380
```

```
path 67E8FFCC RPKI State valid
```

```
Refresh Epoch 1
```

```
16509 701 2497 234
```

```
129.250.10.157 (metric 11) from 198.180.150.253 (198.180.150.253)
```

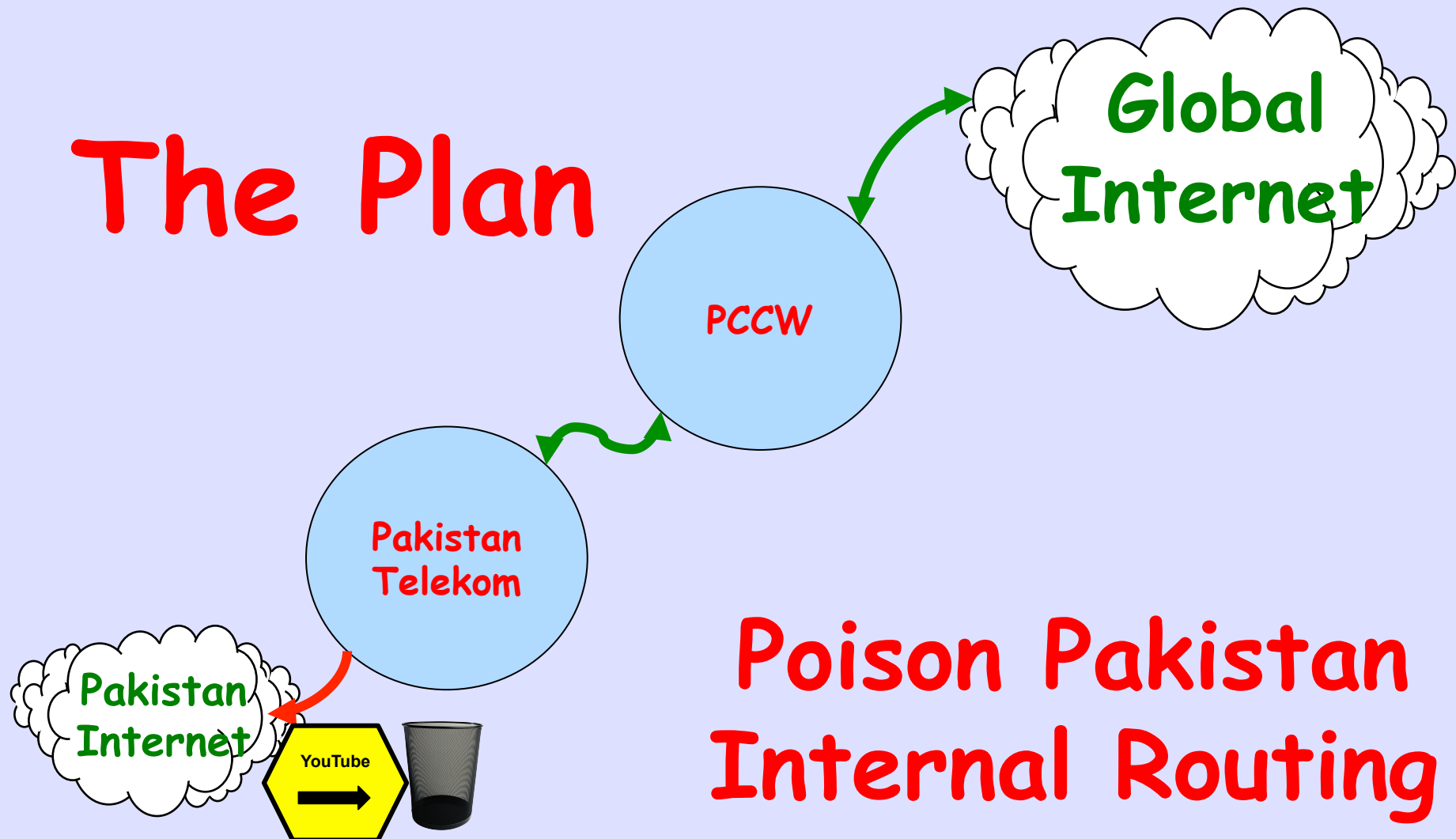
```
Origin IGP, metric 95, localpref 100, valid, internal
```

```
Community: 2914:410 2914:1007 2914:2000 2914:3000 3927:380
```

```
path 699A867C RPKI State valid
```


The YouTube Incident

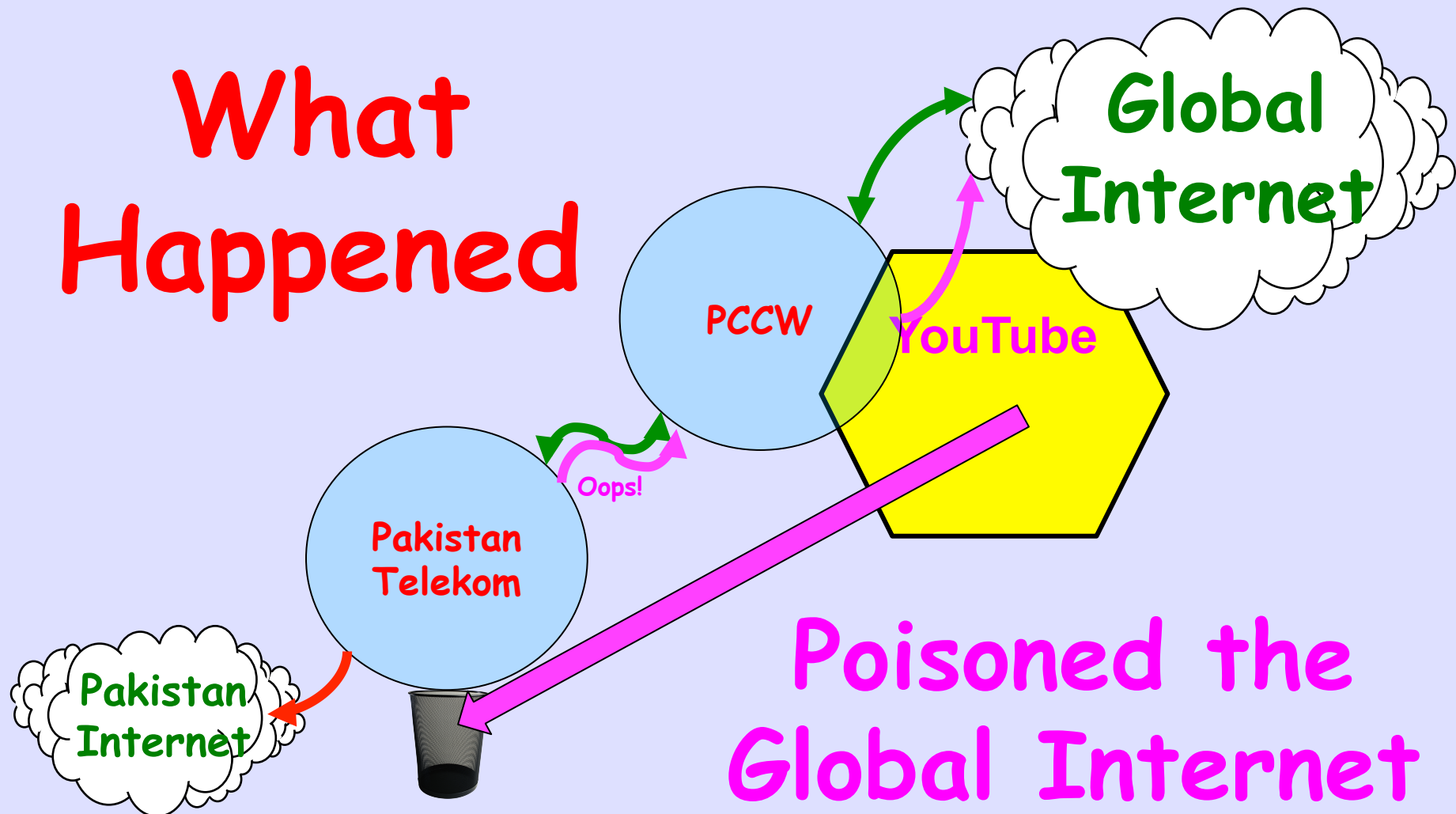
The Plan



Poison Pakistan
Internal Routing

The YouTube Incident

What Happened



We Call this
Mis-Origination

a Prefix is Originated
by an AS Which Does
Not Own It

I Do Not Call it
Hijacking

Because that Assumes
Negative Intent

And These Accidents
Happen Every Day

Usually to Small Folk
Sometimes to Large

So,

What's the Plan?

Three Pieces

- **RPKI** - Resource Public Key Infrastructure, the Certificate Infrastructure to Support the other Pieces (starting last year)
- **Origin Validation** - Using the RPKI to detect and prevent mis-originations of someone else's prefixes (early 2012)
- **AS-Path Validation AKA BGPsec** - Prevent Attacks on BGP (future work)

Why Origin Validation?

- Prevent YouTube accident & Far Worse
- Prevent 7007 accident, UU/Sprint 2 days!
- Prevents most accidental announcements
- Does not prevent malicious path attacks such as the Kapela/Pilosov DefCon attack
- That requires 'Path Validation' and locking the data plane to the control plane, the third step, a few years away

We Need to be Able to
Authoritatively Prove
Who Owns an IP Prefix
And What AS(s) May
Announce It

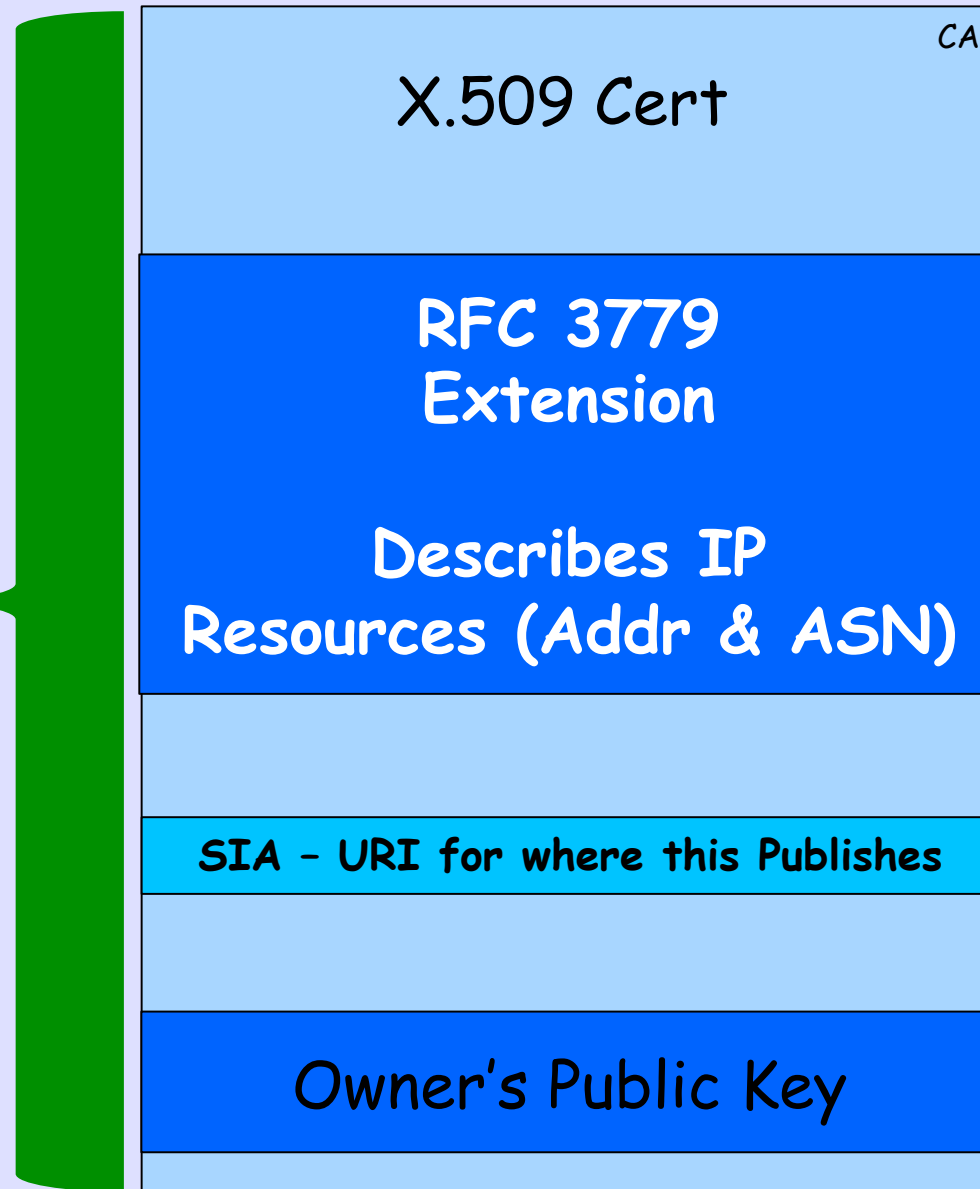
Prefix Ownership Follows the Allocation Hierarchy IANA, RIRs, ISPs, ...

Resource Public Key Infrastructure (RPKI)

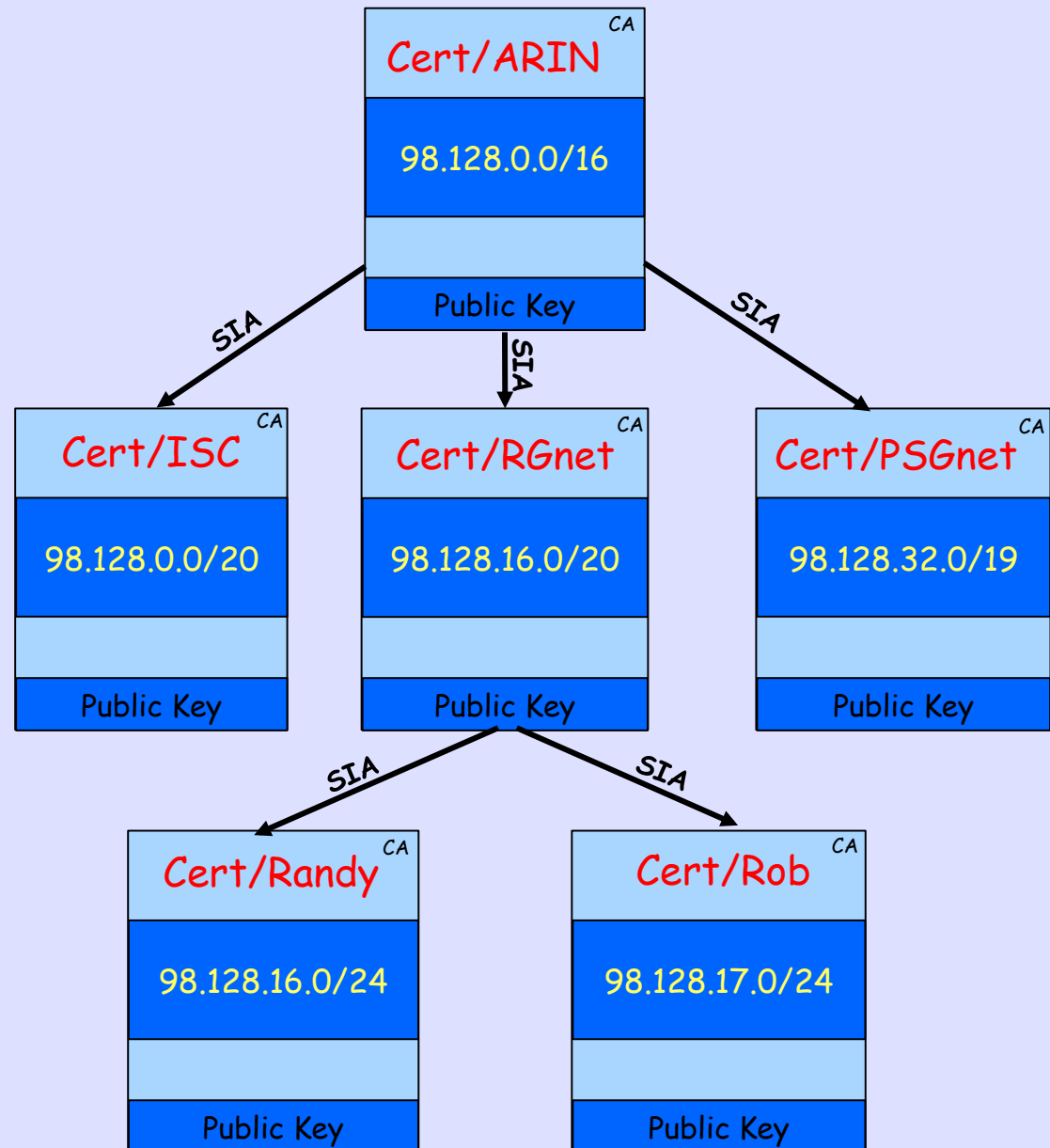
X.509 RPKI Being Developed & Deployed by IANA, RIRs, and Operators

X.509 Certificate w/ 3779 Ext

Signed
by
Parent's
Private
Key

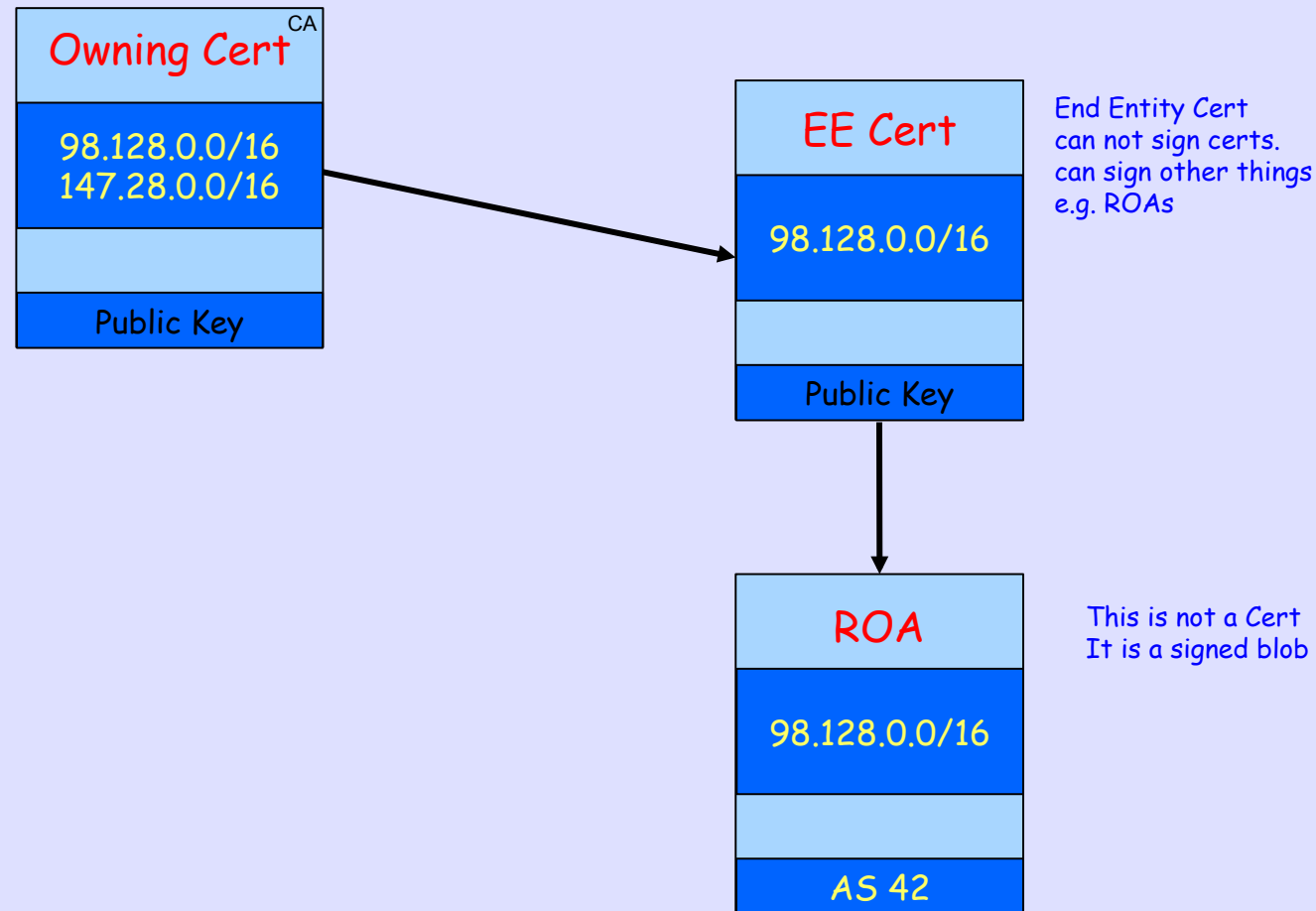


Certificate Hierarchy follows Allocation Hierarchy

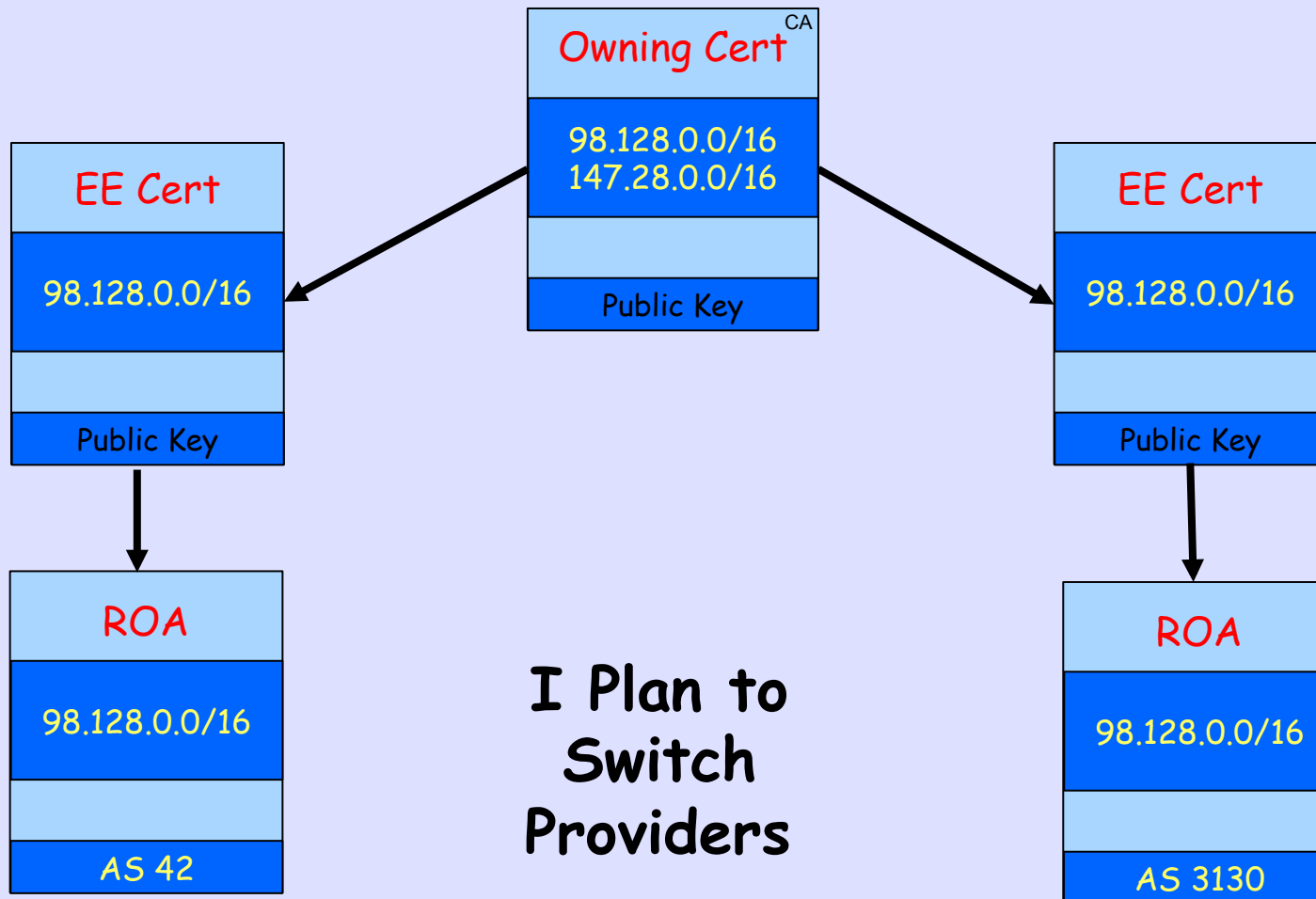


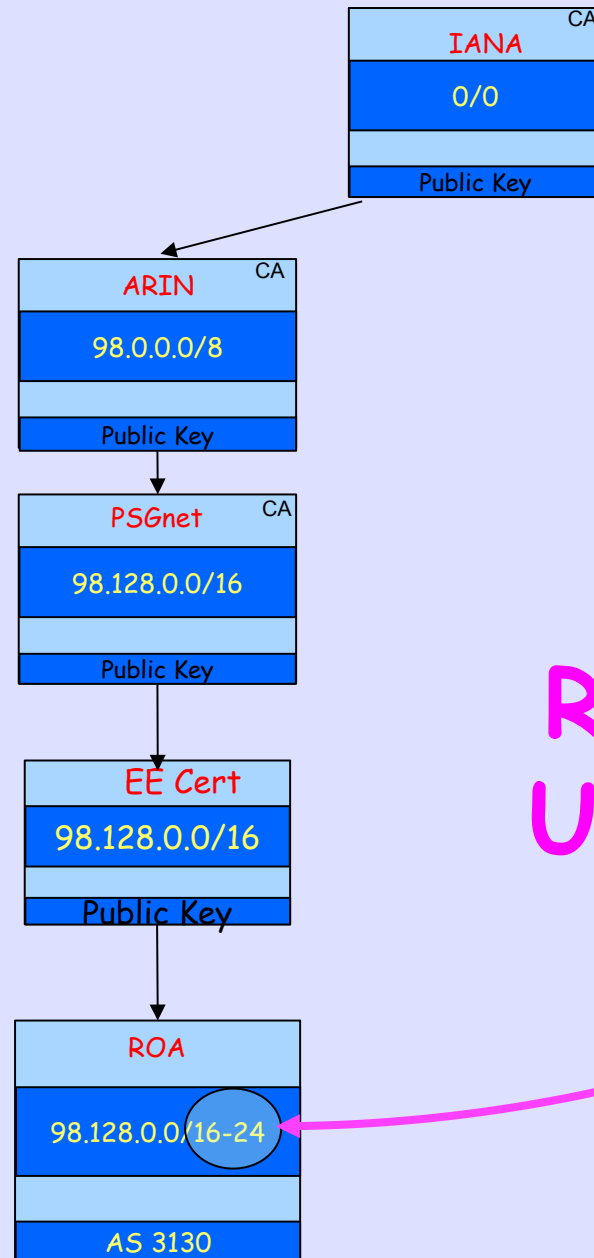
That's Who Owns It
but
Who May Route It?

Route Origin Authorization (ROA)



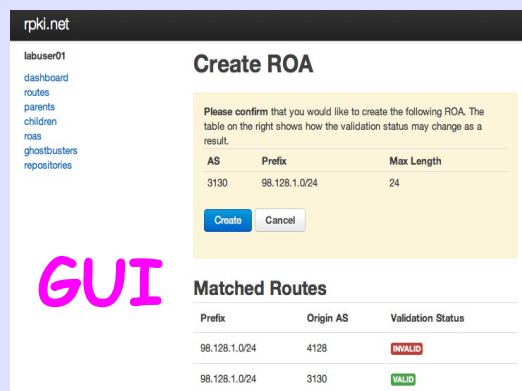
Multiple ROAs Make Before Break



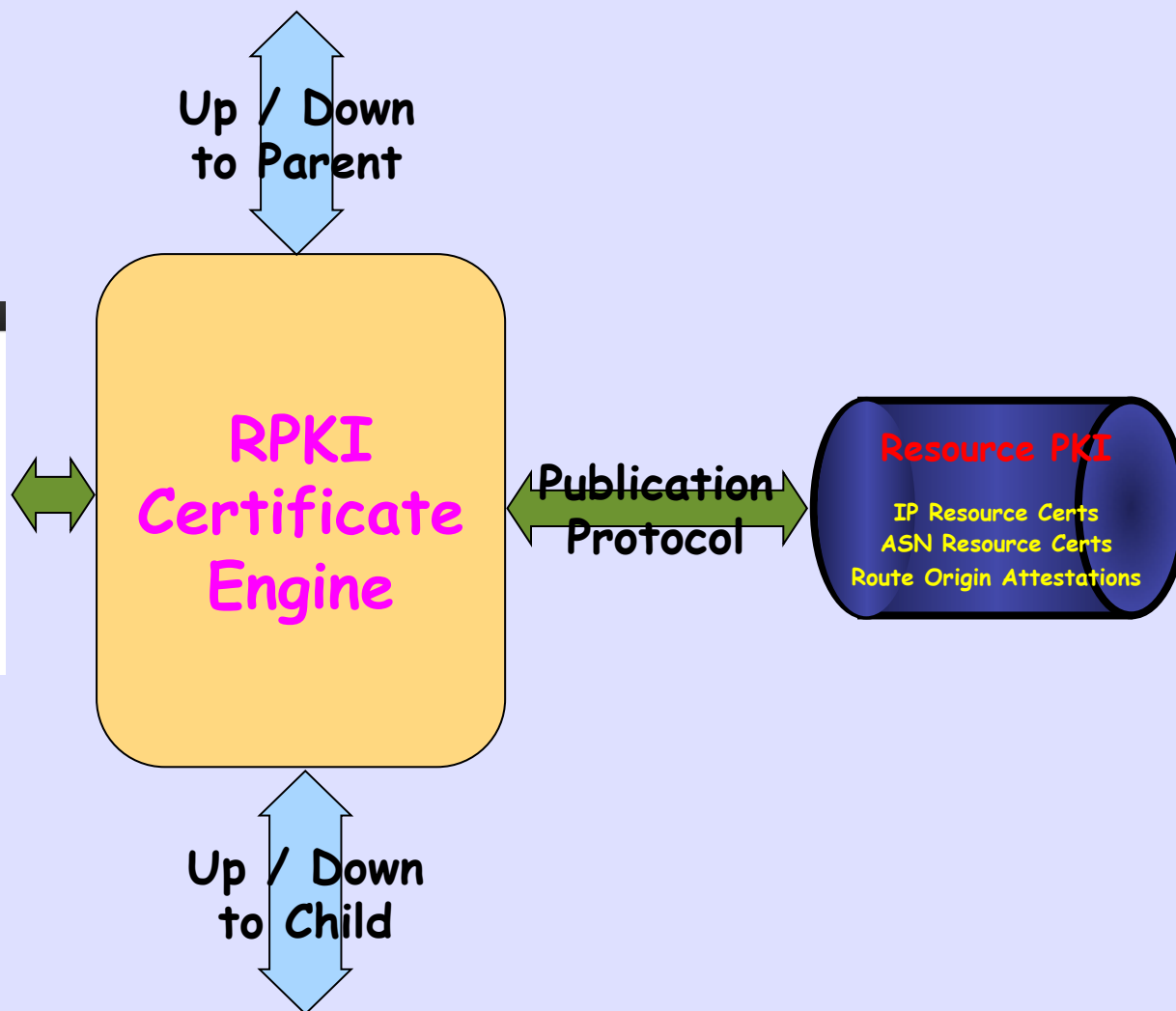


ROA Aggregation Using Max Length

RPKI-Based Origin Validation



GUI



Warning What ROA Will Do

rpki.net

labuser01

[dashboard](#)

[routes](#)

[parents](#)

[children](#)

[roas](#)

[ghostbusters](#)

[repositories](#)

Create ROA

Please confirm that you would like to create the following ROA. The table on the right shows how the validation status may change as a result.

AS	Prefix	Max Length
3130	98.128.1.0/24	24

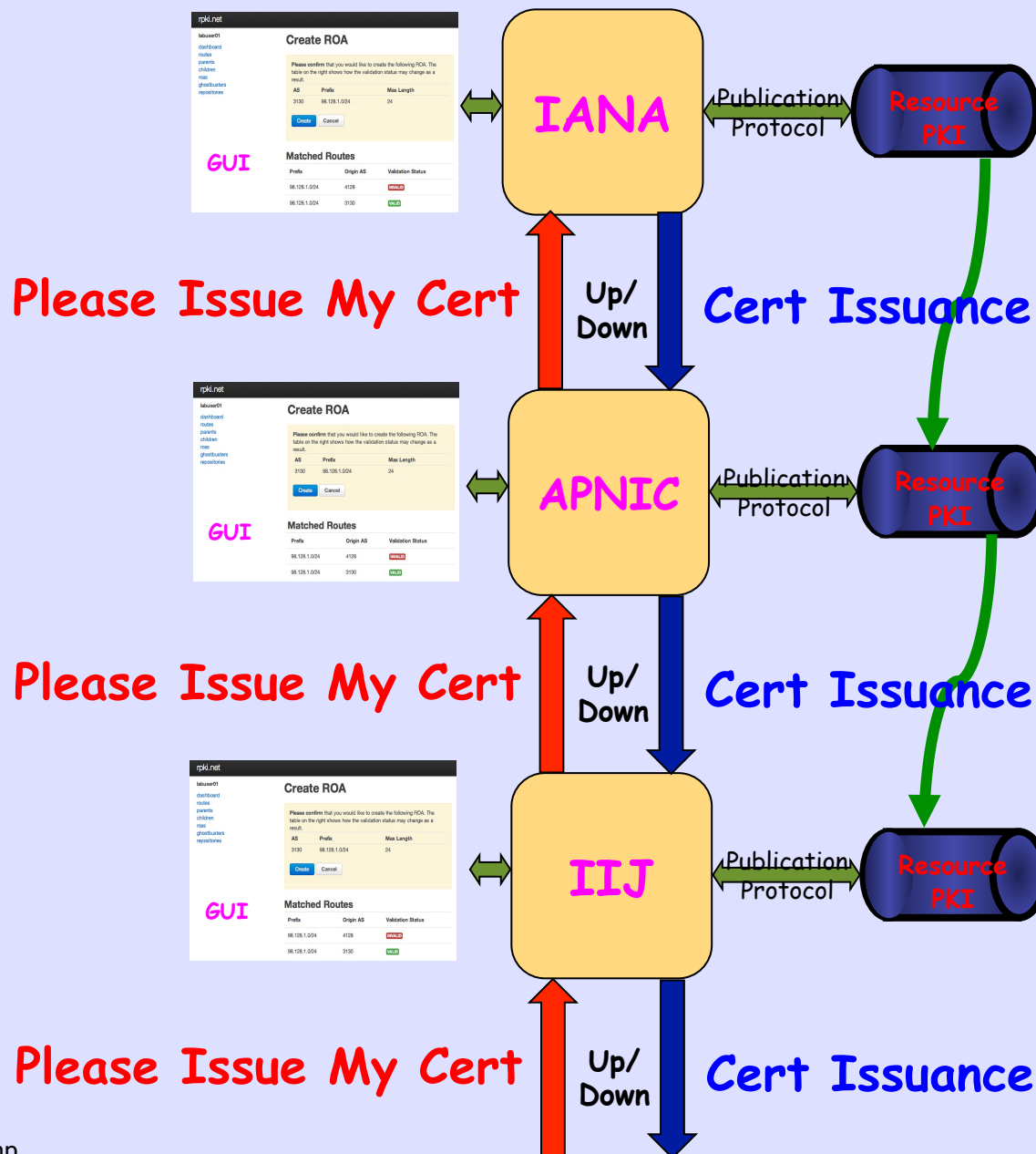
Create

Cancel

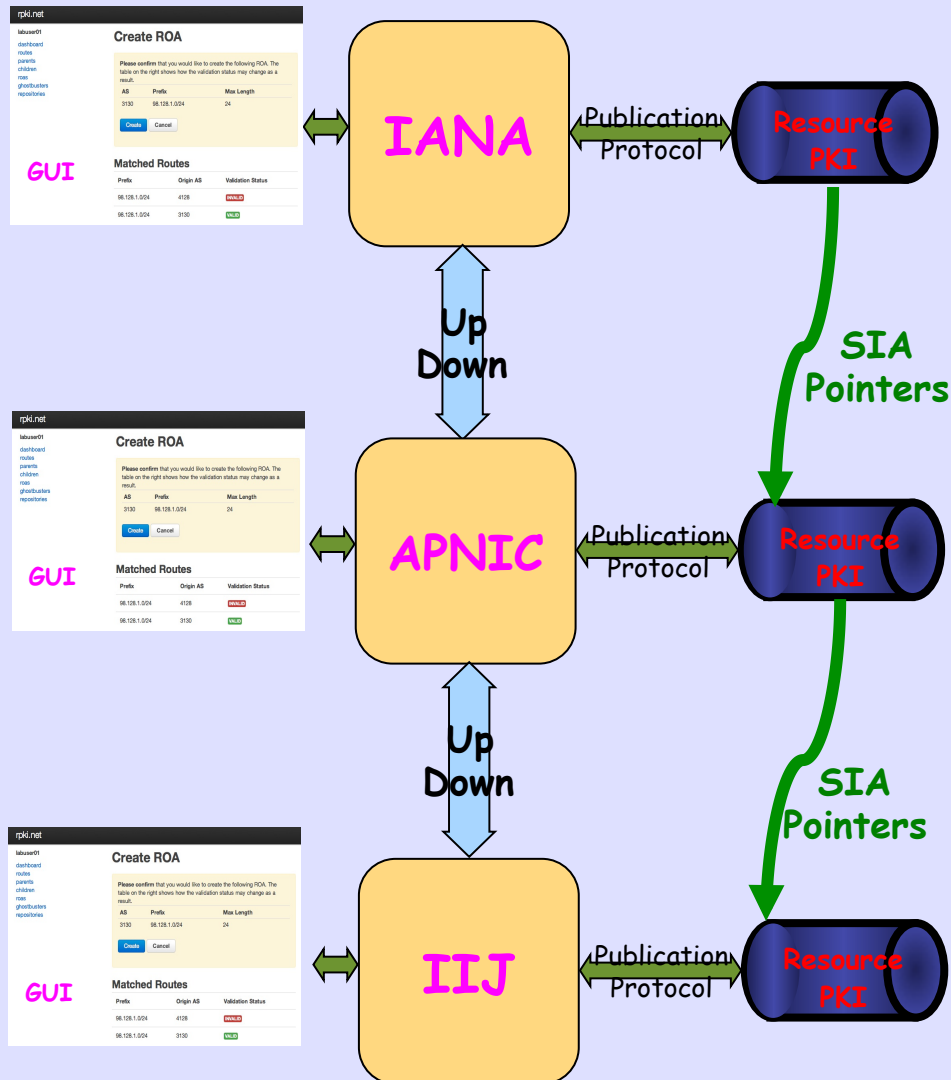
Matched Routes

Prefix	Origin AS	Validation Status
98.128.1.0/24	4128	INVALID
98.128.1.0/24	3130	VALID

Issuing Parties

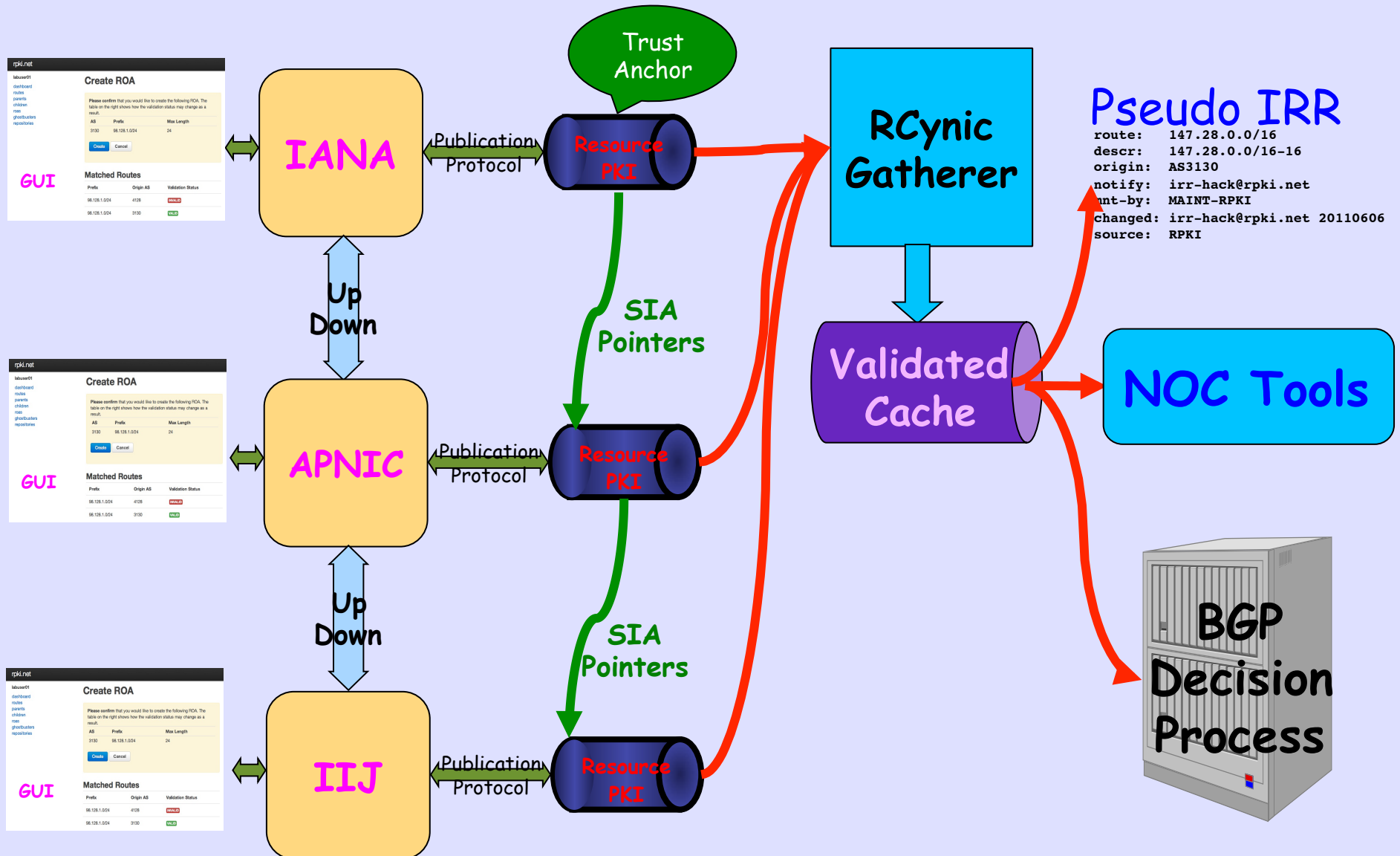


Issuing Parties

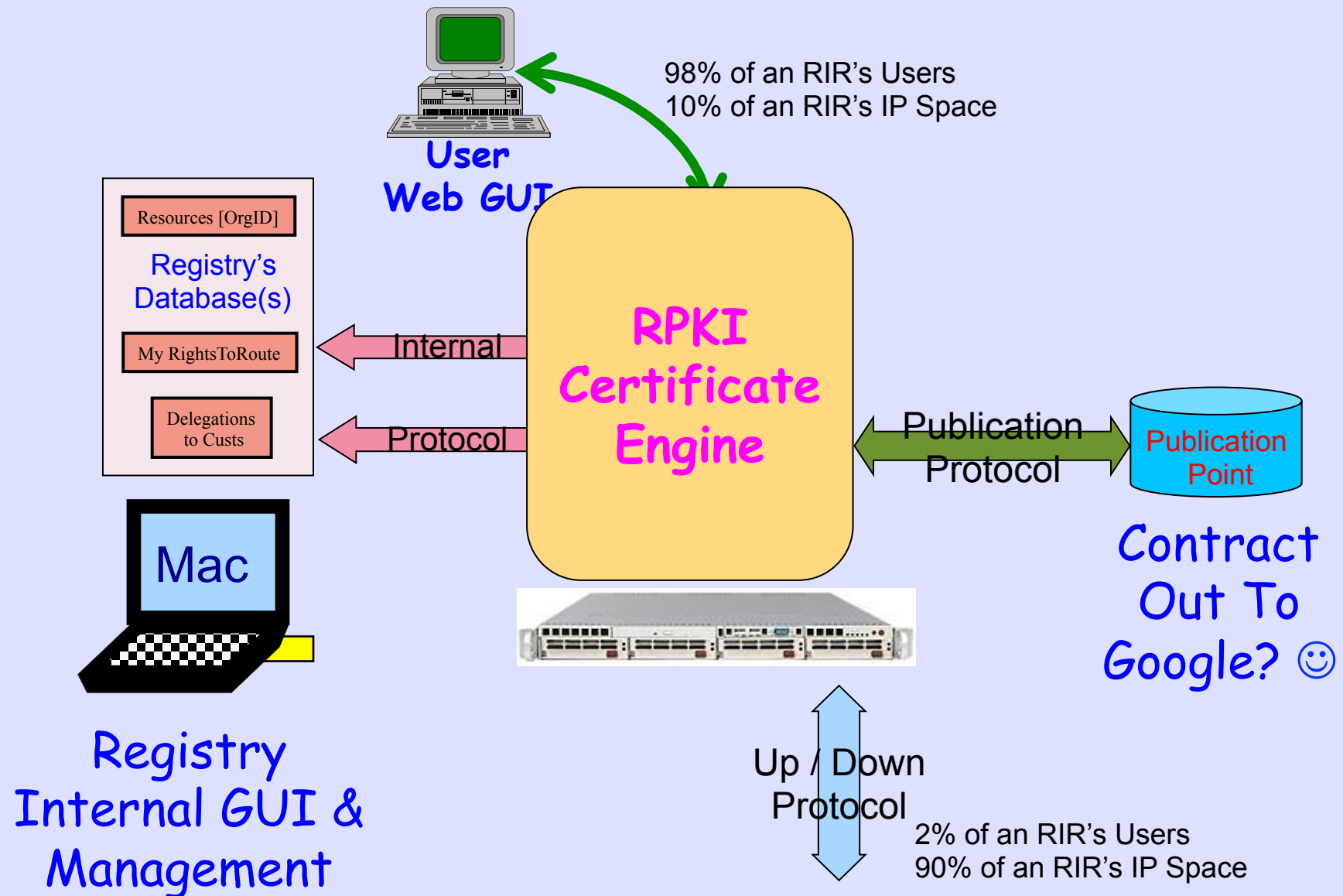


Issuing Parties

Relying Parties

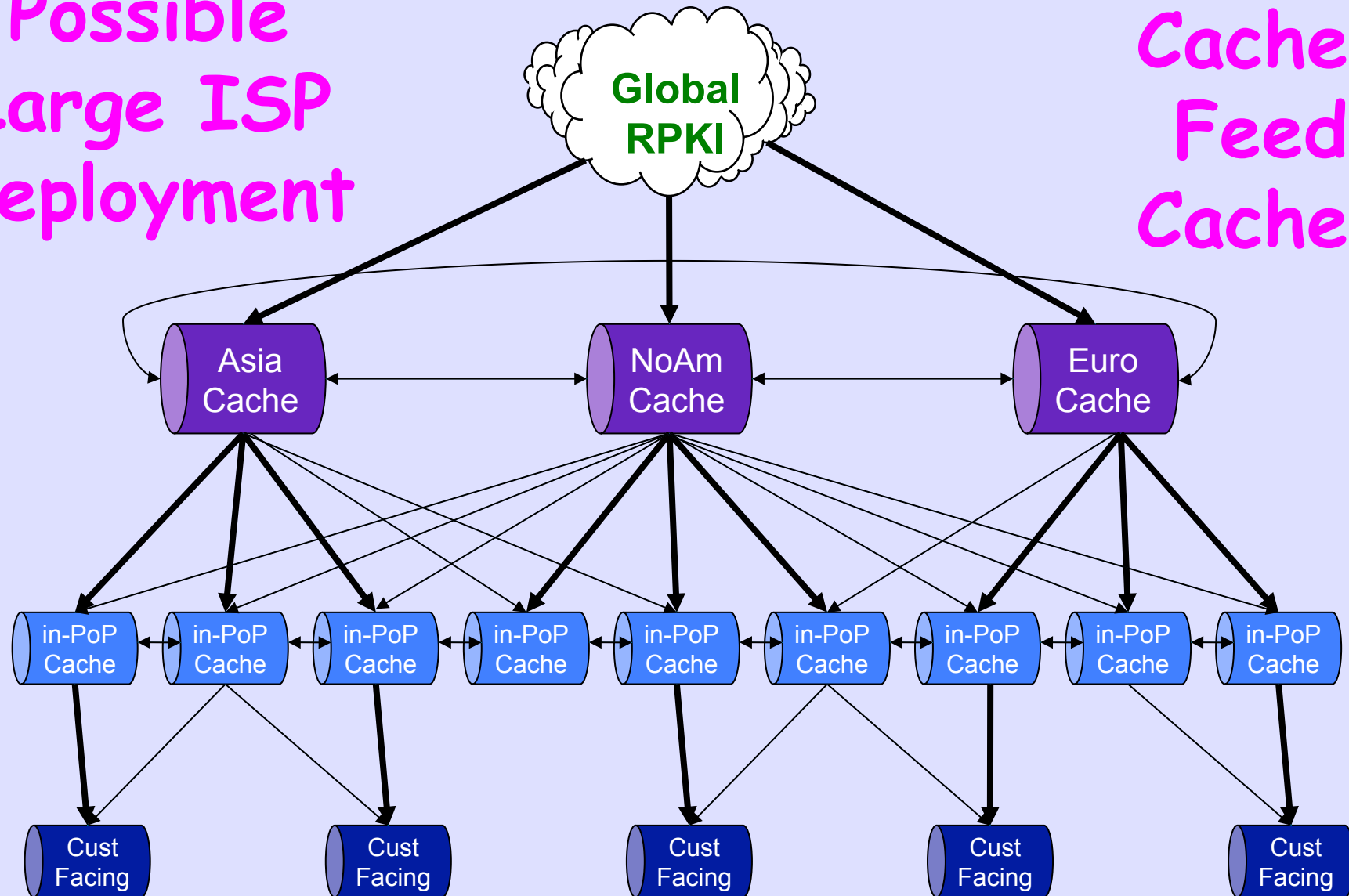


A Usage Scenario

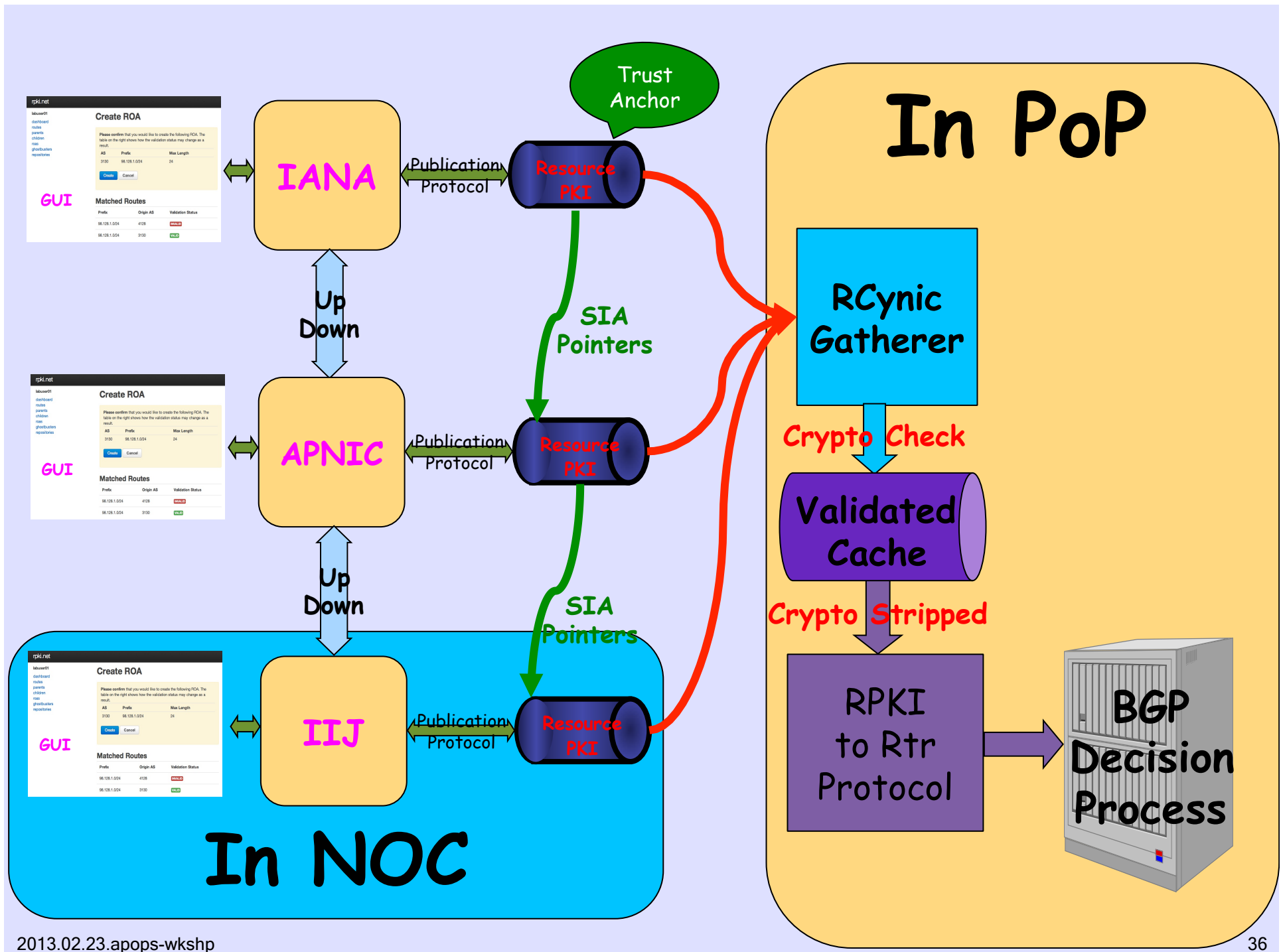


Possible
Large ISP
Deployment

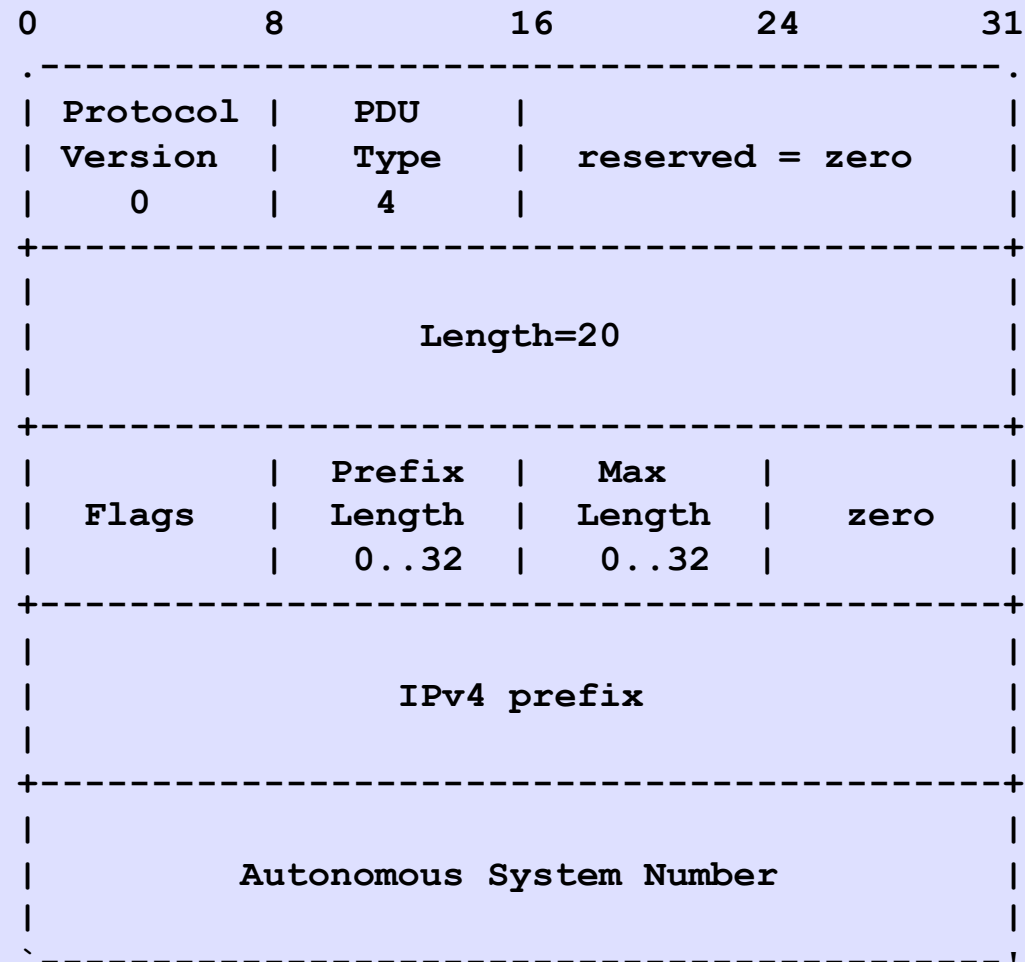
Caches
Feed
Caches



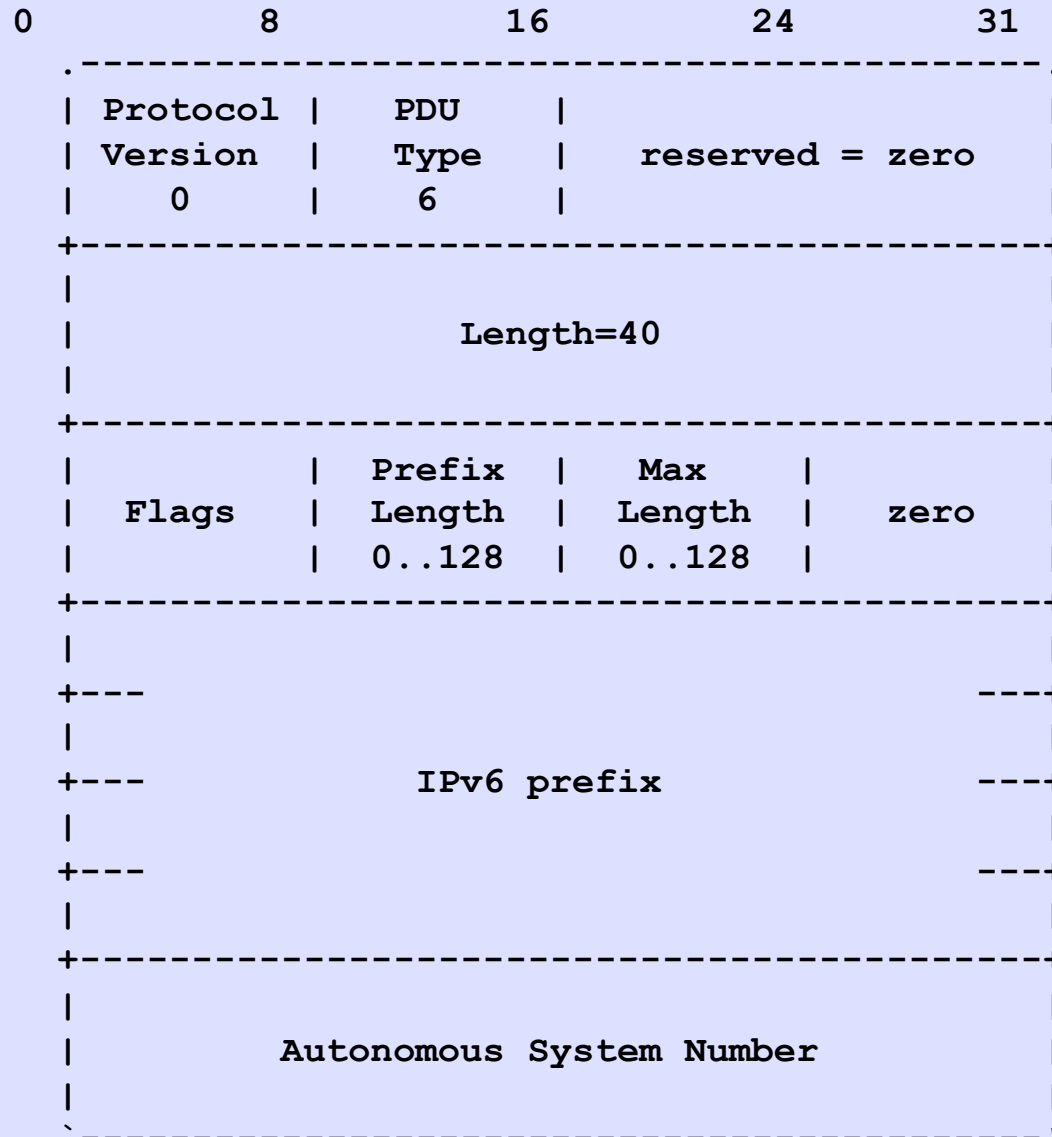
How Do ROAs Affect BGP Updates?



IPv4 Prefix

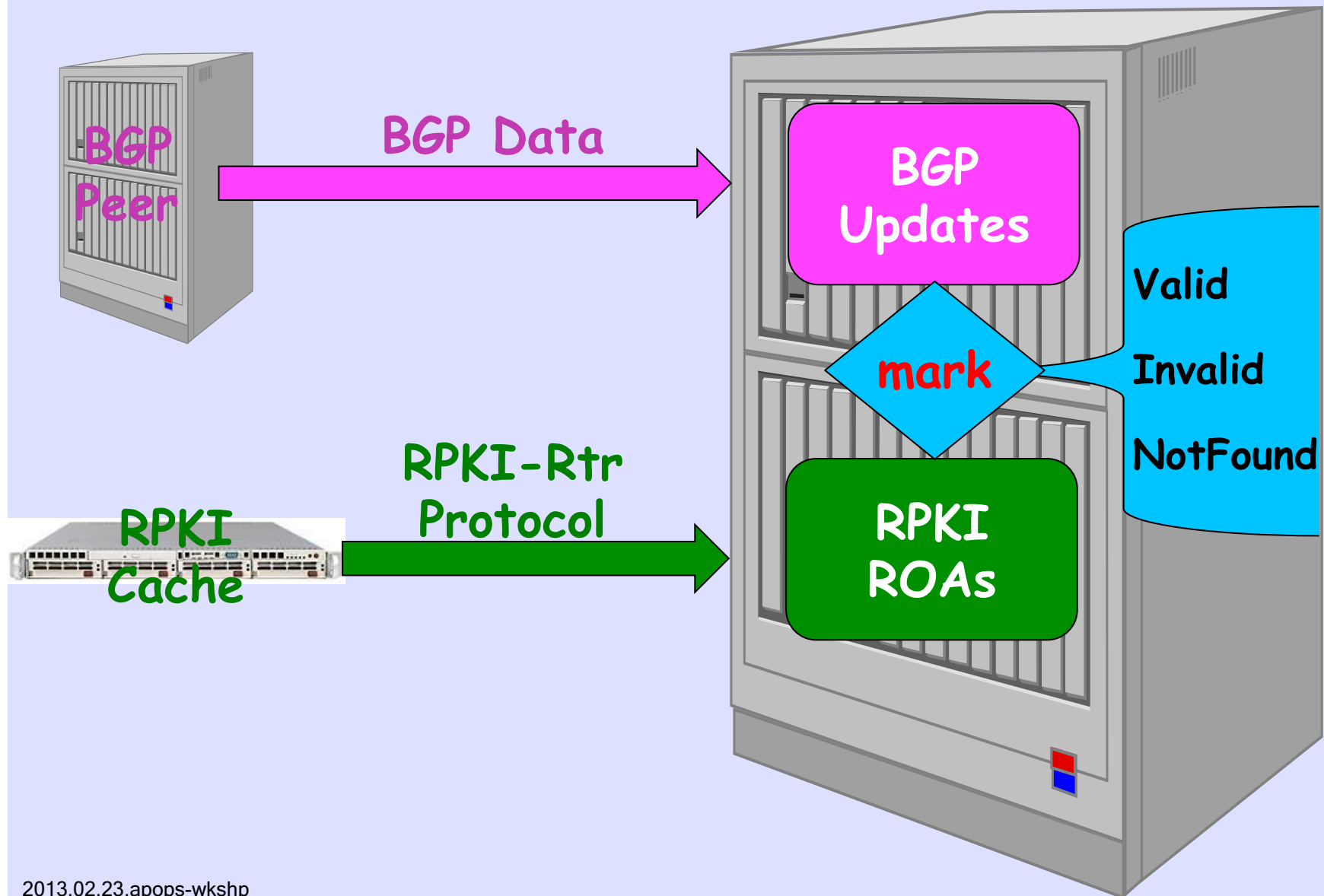


IPv6 Prefix



BGP Updates are
compared with
ROA Data loaded
from the RPKI

Marking BGP Updates



Configure Router to Get ROAs

```
router bgp 3130
```

```
...
```

```
bgp rpki server tcp 198.180.150.1 port 42420 refresh 3600
```

```
bgp rpki server tcp 147.28.0.35 port 93920 refresh 3600
```

```
...
```

Result of Check

- **Valid** - A matching/covering ROA was found with a matching AS number
- **Invalid** - A matching or covering ROA was found, but AS number did not match, and there was no valid one
- **Not Found** - No matching or covering ROA was found, same as today

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
```

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
```

NotFound

```
r0.sea#show bgp 64.9.224.0
```

```
BGP routing table entry for 64.9.224.0/20, version 35201
```

```
Paths: (3 available, best #2, table default)
```

```
  Advertised to update-groups:
```

```
      2          5          6
```

```
Refresh Epoch 1
```

```
1239 3356 36492
```

```
  144.232.9.61 (metric 11) from 147.28.7.2 (147.28.7.2)
```

```
    Origin IGP, metric 4, localpref 100, valid, internal
```

```
    Community: 3130:370
```

```
    path 11861AA4 RPKI State not found
```

What are the BGP / VRP¹ Matching Rules?

¹ Validated ROA Payload

A Prefix is Covered by a VRP when the VRP prefix length is less than or equal to the Route prefix length

BGP

98.128.0.0/16

VRP

98.128.0.0/12-16

Covers

VRP

98.128.0.0/16-24

Covers

VRP

98.128.0.0/20-24

No. It's Longer

Prefix is Matched by a VRP when the Prefix is Covered by that VRP , prefix length is less than or equal to the VRP max-len, and the Route Origin AS is equal to the VRP's AS

BGP

98.128.0.0/16 AS 42

VRP

98.128.0.0/12-16 AS 42

Matched

VRP

98.128.0.0/16-24 AS 666

No. AS Mismatch

VRP

98.128.0.0/20-24 AS 42

No. VRP Longer

Matching and Validity

VRP₀

98.128.0.0/16-24 AS 6

VRP₁

98.128.0.0/16-20 AS 42

BGP	98.128.0.0/12 AS 42	NotFound, shorter than VRPs
BGP	98.128.0.0/16 AS 42	Valid, Matches VRP ₁
BGP	98.128.0.0/20 AS 42	Valid, Matches VRP ₁
BGP	98.128.0.0/24 AS 42	Invalid, longer than VRP with AS 42
BGP	98.128.0.0/24 AS 6	Valid, Matches VRP ₀

The Operator Tests and then Sets Local Policy

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
```

Set a Community

```
route-map validity-0
```

```
  match rpki valid
```

```
    set community 3130:400
```

```
route-map validity-1
```

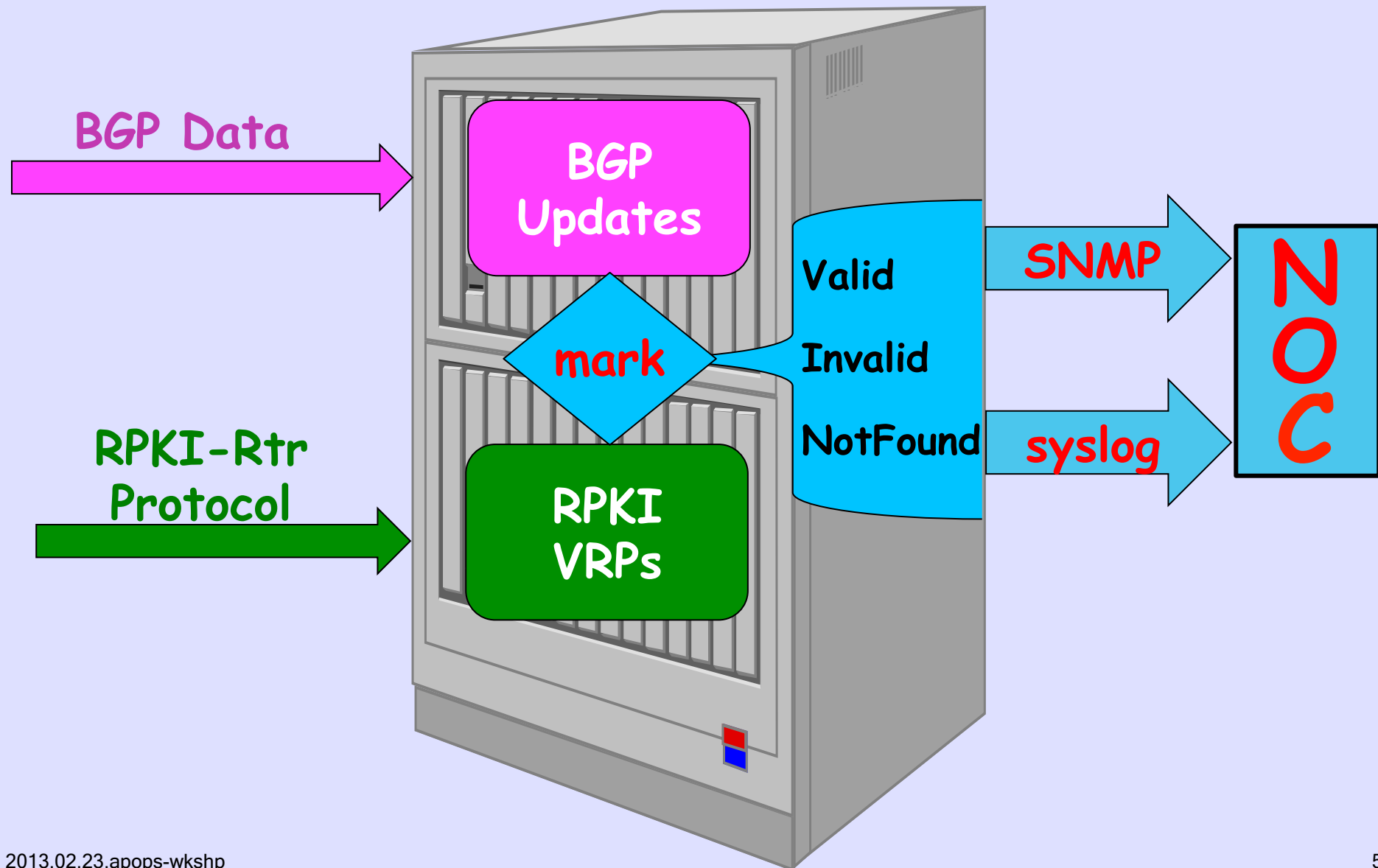
```
  match rpki invalid
```

```
    set community 3130:200
```

```
route-map validity-2
```

```
  set community 3130:300
```

And it is All Monitored



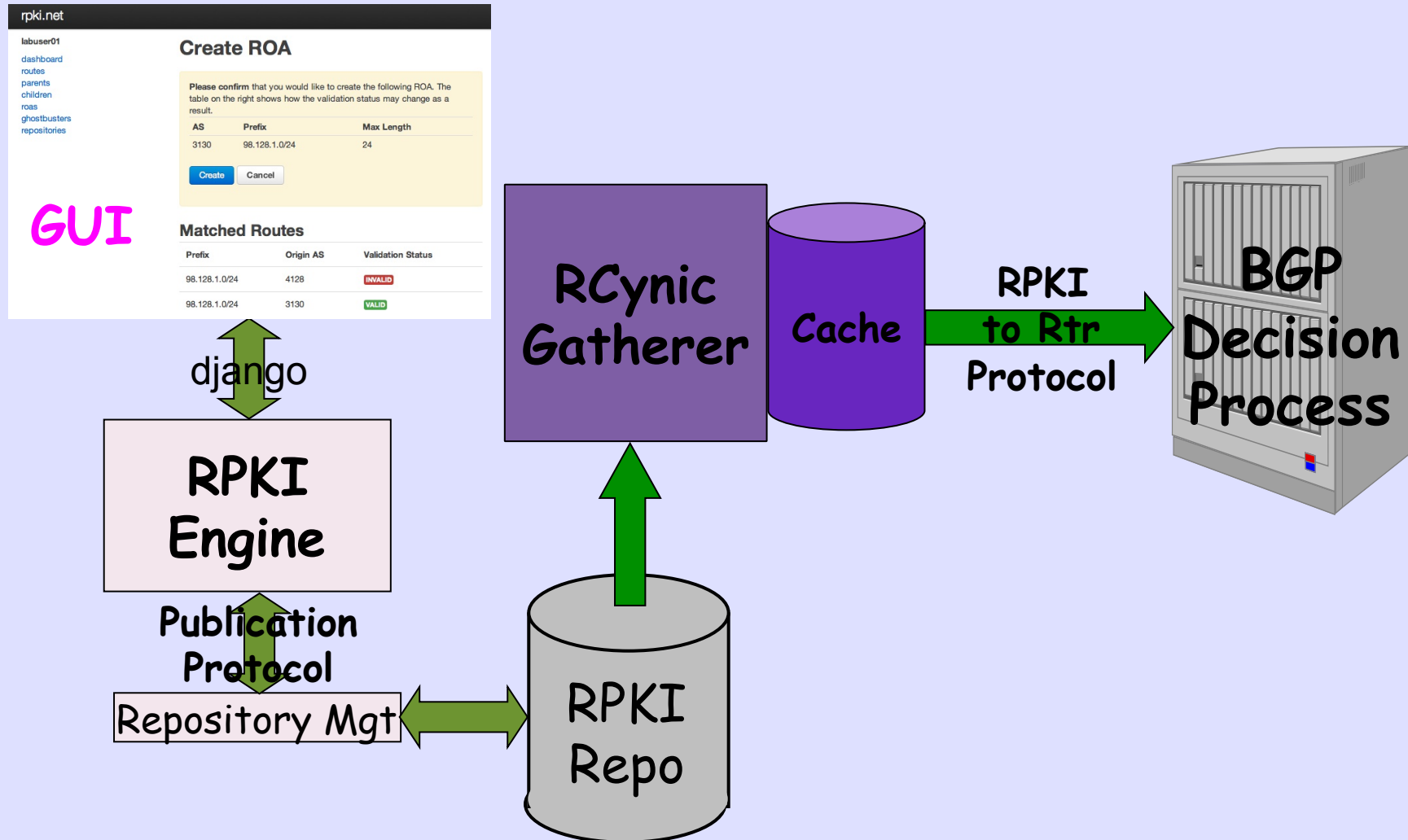
But in the End, You Control Your Policy

"Announcements with Invalid origins MAY be used, but SHOULD be less preferred than those with Valid or NotFound."

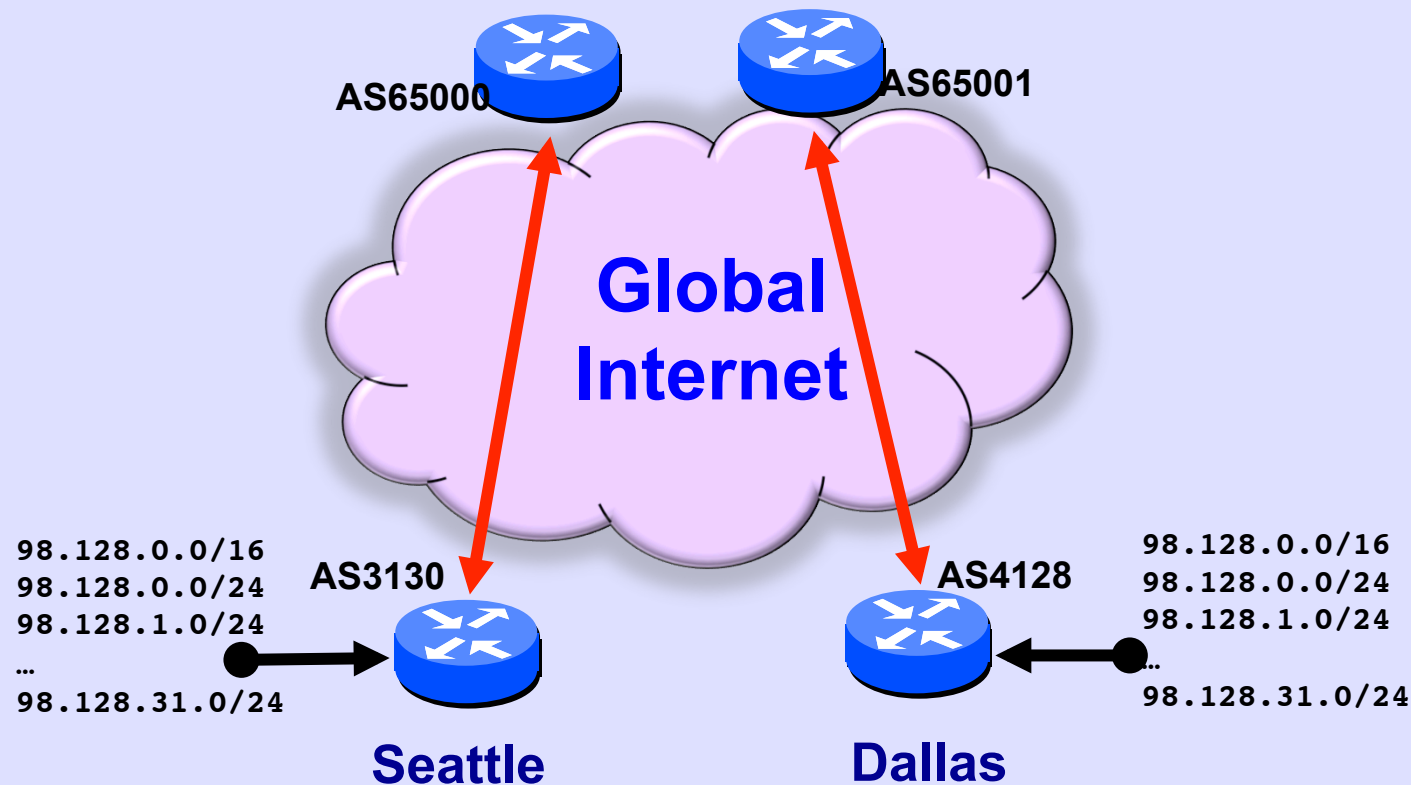
-- draft-ietf-sidr-origin-ops

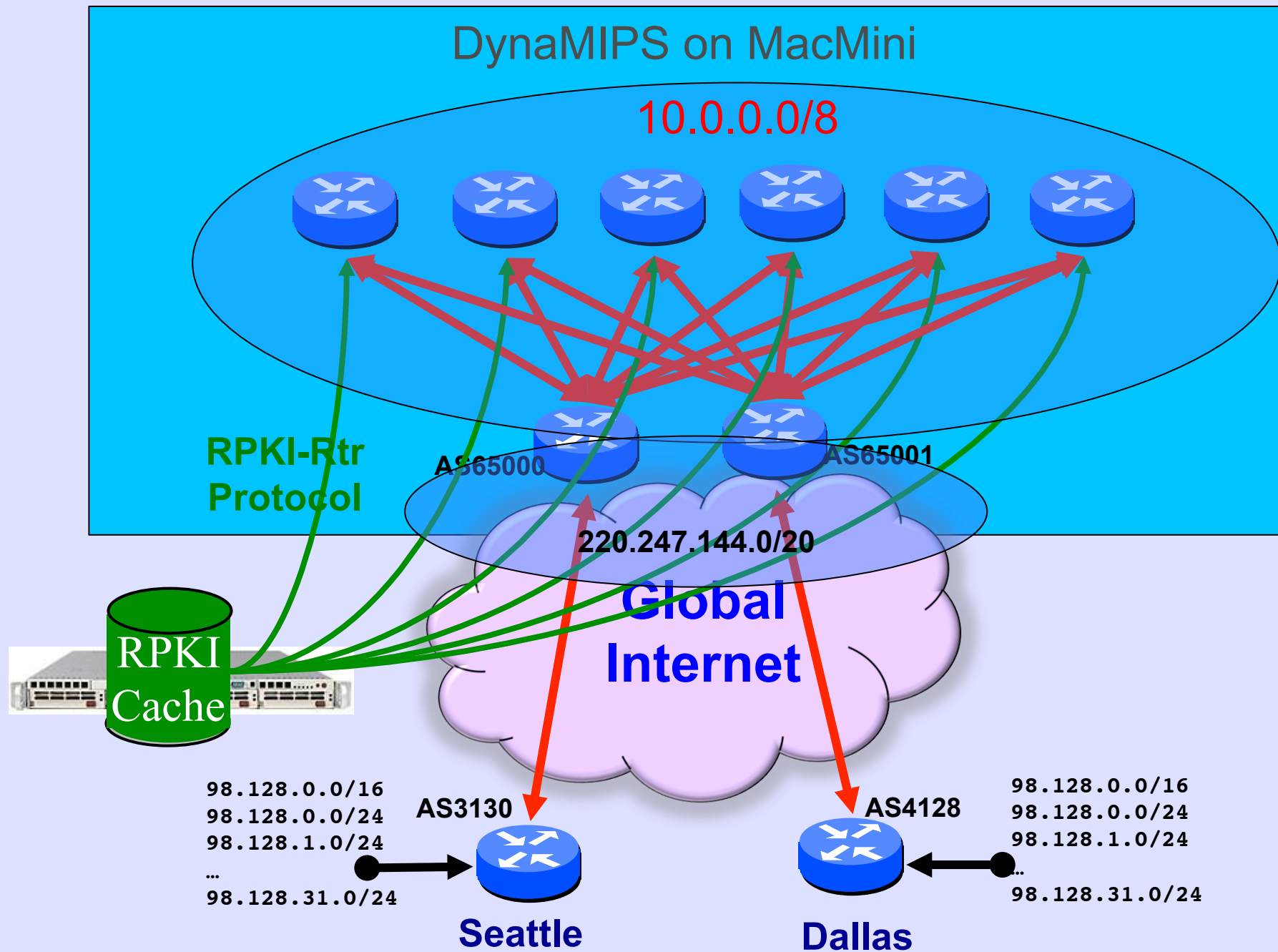
But if I do not reject Invalid, what is all this for?

Lab Overview



BGP Feeds into Lab





IP Address Allocation

98.128.0.0/16 ARIN Experimental Allocation

98.128.0.0/24 Instructors Play

98.128.1.0/24 labuser01

98.128.2.0/24 labuser02

...

98.128.30.0/24 labuser30

GUI Accounts

<https://demo.rpki.net/rpki>

<u>UserID</u>	<u>Password</u>
labuser01	fnord
labuser02	fnord
labuser03	fnord
...	
labuser30	fnord

https://demo.rpki.net/

The screenshot shows a web browser window with the address bar displaying `https://demo.rpki.net/accounts/login/?next=/rpki/`. The page has a dark header with the text "rpki.net". Below the header, there is a login form with two input fields: "Username" containing the text "randy" and "Password" containing seven dots. Both input fields have a red asterisk icon to their right. Below the input fields, there is a blue "Login" button. The background of the page is white, and the overall layout is clean and simple.


The Front Page

rpki.net

Logged in

labuser01

[dashboard](#)

 [routes](#)

[parent](#)

[children](#)

[roas](#)

[ghostbusters](#)

[repositories](#)

[export identity](#)

[select identity](#)

[refresh](#)

Dashboard

Resources

Resource	Valid Until	Parent
98.128.1.0/24	Feb. 16, 2013, 2:38 a.m.	rgnet

Unallocated Resources

The following resources have not been allocated to a child, nor appear in a ROA.

- 98.128.1.0/24

rpki.net

labuser01

[dashboard](#)

[routes](#)

[parents](#)

[children](#)

[roas](#)

[ghostbusters](#)

[repositories](#)

Roas

There are **no items** in this list.



Create

Issue a ROA

rpki.net

labuser01

- dashboard
- routes
- parents
- children
- roas
- ghostbusters
- repositories

Create ROA

AS

Prefix

Max Prefix Length

What Will Happen?

rpki.net

labuser01

[dashboard](#)

[routes](#)

[parents](#)

[children](#)

[roas](#)

[ghostbusters](#)

[repositories](#)

Create ROA

Please confirm that you would like to create the following ROA. The table on the right shows how the validation status may change as a result.

AS	Prefix	Max Length
3130	98.128.1.0/24	24



Create

Cancel

Matched Routes

Prefix	Origin AS	Validation Status
98.128.1.0/24	4128	INVALID
98.128.1.0/24	3130	VALID

ROA View

rpki.net

Logged in as labus

labuser01

dashboard

routes

parents

children

roas

ghostbusters

repositories

Roas

Prefix	Max Length	ASN	Action
98.128.1.0/24	24	3130	<div>Delete</div>

Create

Route View

rpki.netLogged in as labuser01 Log O

labuser01
[dashboard](#)
[routes](#)
[parents](#)
[children](#)
[roas](#)
[ghostbusters](#)
[repositories](#)

BGP data updated
IPv4: 2012-02-23T05:03:31
IPv6:

rcynic cache updated
2012-02-23T05:30:26

Route View

This view shows currently advertised routes for the prefixes listed in resource certs received from RPKI parents.

Prefix	Origin AS	Validation Status
98.128.1.0/24	4128	INVALID roas
98.128.1.0/24	3130	VALID roas

Router Accounts

telnet ws-routing-b 20xx (Users 1-18)

telnet ws-routing-a 20xx (Users 19-30)

e.g.

% telnet ws-routing-b 20xx

user: cisco

password: cisco

enable: cisco

Configure RPKI Server

```
router bgp <AS>
```

```
bgp rpki server tcp <ip address> port <port>  
refresh <time in sec>
```

```
router bgp 651xx
```

```
bgp rpki server tcp 198.180.150.1 port \  
42420 refresh 180
```

Check Server

```
r0.sea#show ip bgp rpki servers
```

```
BGP SOVC neighbor is 198.180.150.1/42420 connected to port 42420
```

```
Flags 0, Refresh time is 600, Serial number is 1304239609
```

```
InQ has 0 messages, OutQ has 0 messages, formatted msg 345
```

```
Session IO flags 3, Session flags 4008
```

```
Neighbor Statistics:
```

```
Nets Processed 624
```

```
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
```

```
Connection is ECN Disabled
```

```
Mininum incoming TTL 0, Outgoing TTL 255
```

```
Local host: 199.238.113.10, Local port: 57932
```

```
Foreign host: 198.180.150.1, Foreign port: 42420
```

```
Connection tableid (VRF): 0
```

Look at Table

```
router1#show ip bgp rpki table
```

```
76 BGP sovc network entries using 6688 bytes of memory
```

```
78 BGP sovc record entries using 1560 bytes of memory
```

Network	Maxlen	Origin-AS	Source	Neighbor
98.128.0.0/24	24	3130	0	198.180.150.1/4242
98.128.0.0/16	16	3130	0	198.180.150.1/4242
98.128.6.0/24	24	4128	0	198.180.150.1/4242
98.128.9.0/24	24	3130	0	198.180.150.1/4242
98.128.30.0/24	24	1234	0	198.180.150.1/4242
128.224.1.0/24	24	3130	0	198.180.150.1/4242
129.6.0.0/17	17	49	0	198.180.150.1/4242
129.6.112.0/24	24	10866	0	198.180.150.1/4242
129.6.128.0/17	17	49	0	198.180.150.1/4242
147.28.0.0/16	16	3130	0	198.180.150.1/4242

Look at BGP Table

```
r0.sea#sh ip bgp
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
* i	I198.180.150.0	144.232.9.61	100	0	1239	3927 i
*>	I	199.238.113.9		0	2914	3927 i
*	I	129.250.11.41		0	2914	3927 i
*>	V198.180.152.0	199.238.113.9		0	2914	4128 i
*	V	129.250.11.41		0	2914	4128 i
*>	N198.180.155.0	199.238.113.9		0	2914	22773 i
*	N	129.250.11.41		0	2914	22773 i
*>	N198.180.160.0	199.238.113.9		0	2914	23308 13408 5752 i
*	N	129.250.11.41		0	2914	23308 13408 5752 i

Look at a Prefix

```
R3#show ip bgp 98.128.0.0/24
```

```
BGP routing table entry for 98.128.0.0/24, version 360
```

```
Paths: (2 available, best #1, table default)
```

```
65000 3130
```

```
10.0.0.1 from 10.0.0.1 (193.0.24.64)
```

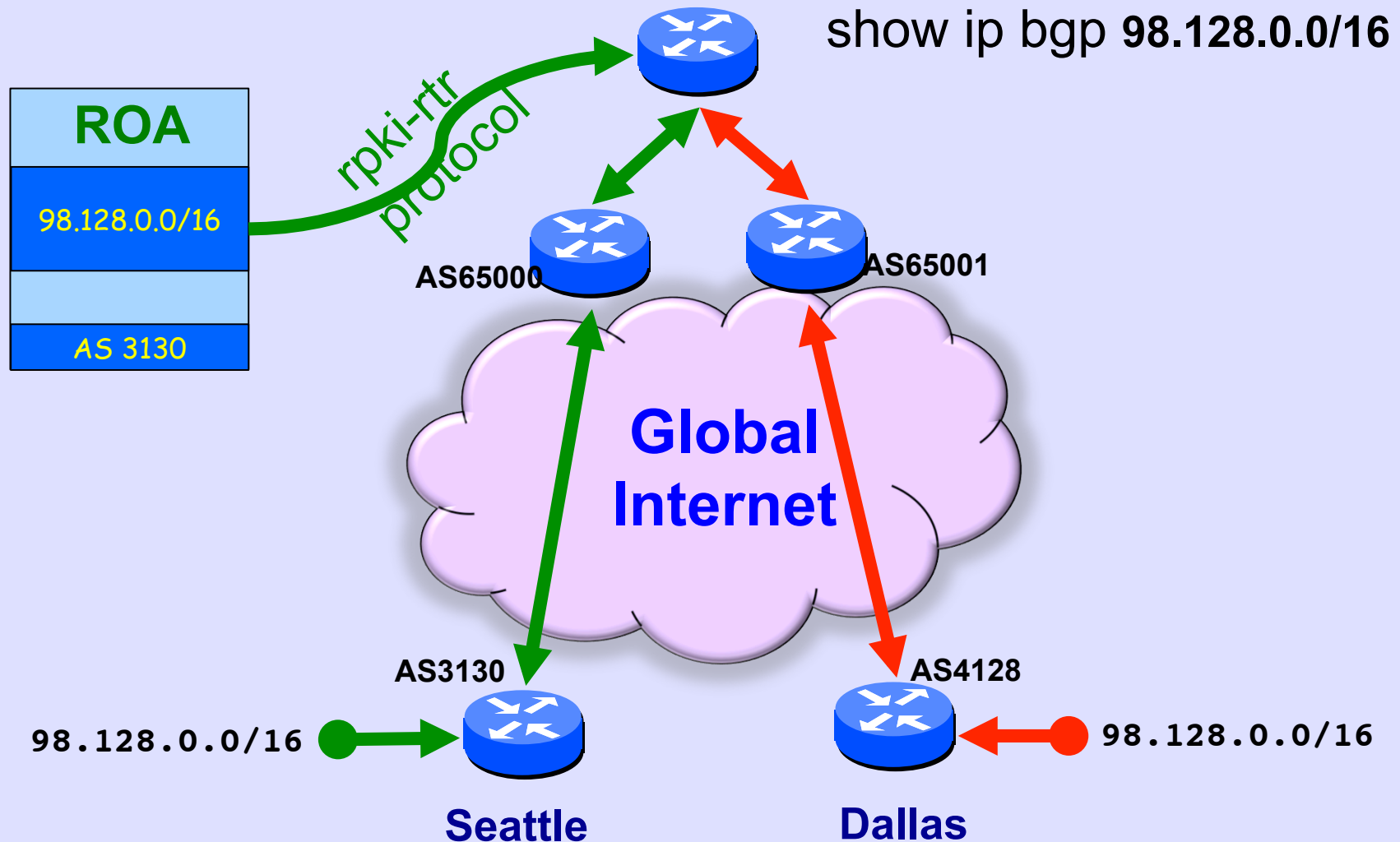
```
Origin IGP, localpref 100, valid, external, best  
path 680D859C RPKI State valid
```

```
65001 4128
```

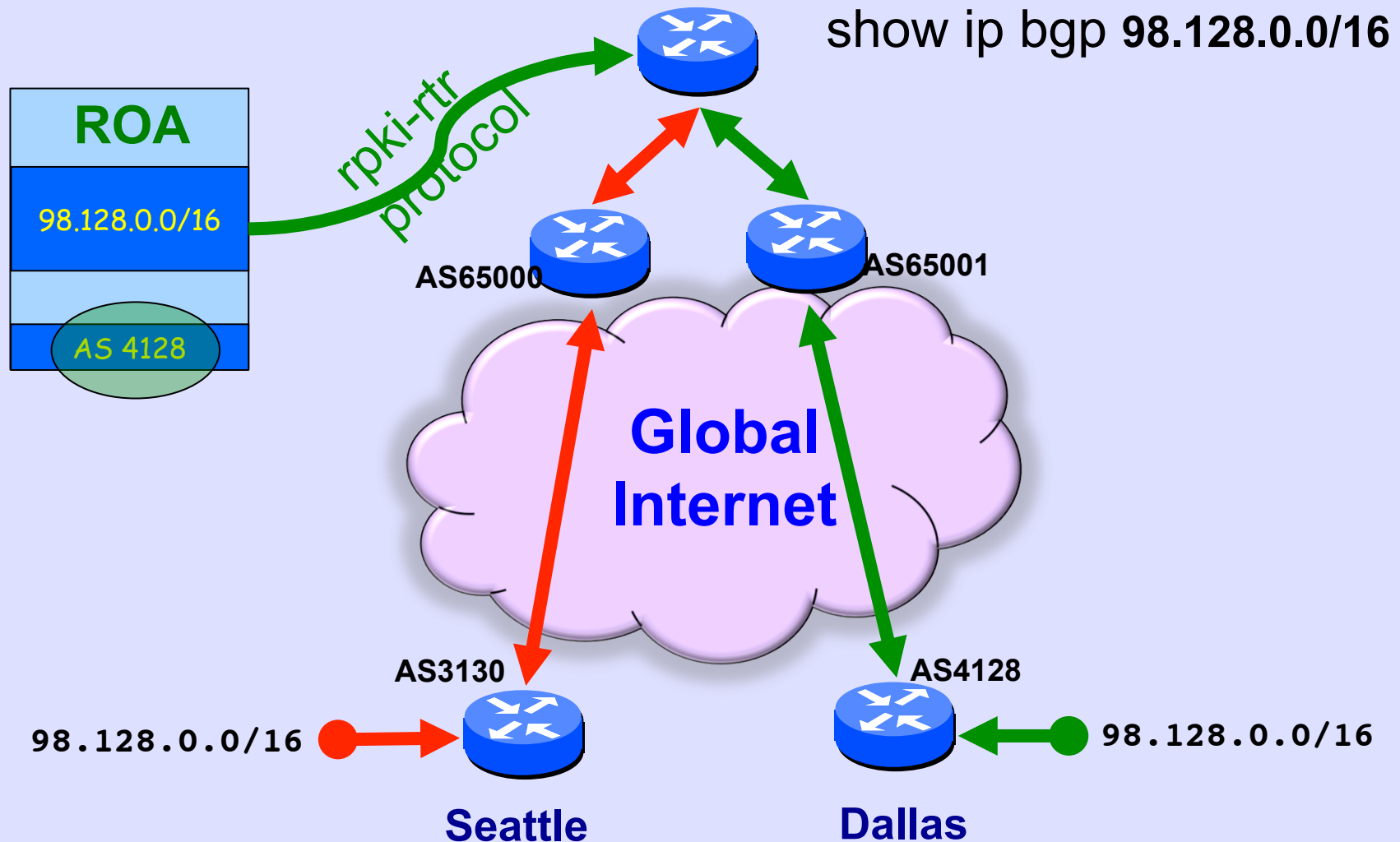
```
10.0.1.1 from 10.0.1.1 (193.0.24.65)
```

```
Origin IGP, localpref 100, valid, external  
path 680D914C RPKI State invalid
```

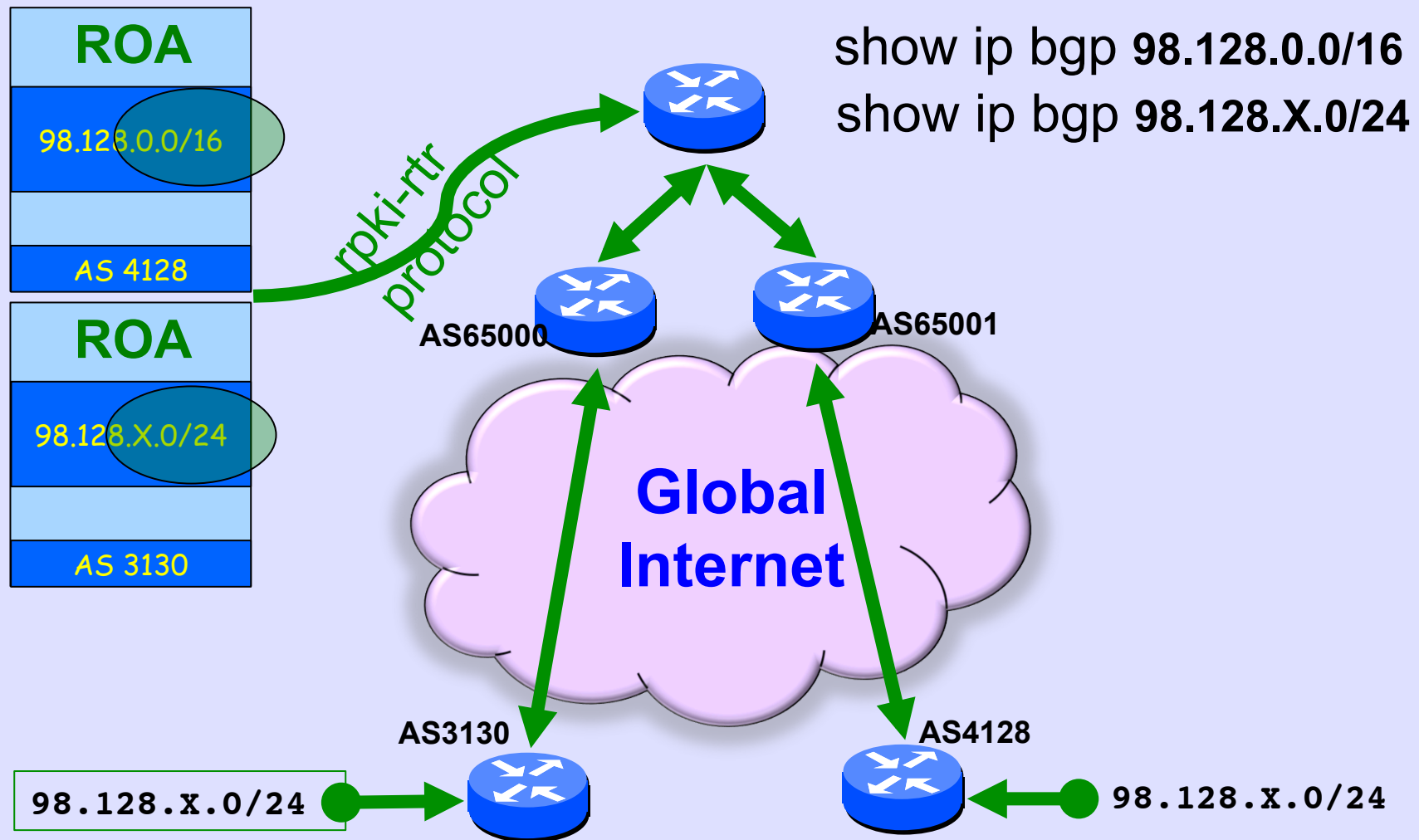
Fat-Finger Detected



ROA Controls Validity



Try Your Own /24



Mis-Origination Caught

```
R3#sh ip bgp 98.128.0.0/24
```

```
BGP routing table entry for 98.128.0.0/24, version 94
```

```
Paths: (2 available, best #2, table default)
```

```
Advertised to update-groups:
```

```
1
```

```
Refresh Epoch 1
```

```
65000 3130
```

```
10.0.0.1 from 10.0.0.1 (65.38.193.12)
```

```
Origin IGP, localpref 100, valid, external
```

```
path 6802D4DC RPKI State invalid
```

```
Refresh Epoch 1
```

```
65001 4128
```

```
10.0.1.1 from 10.0.1.1 (65.38.193.13)
```

```
Origin IGP, localpref 100, valid, external, best
```

```
path 6802D7C8 RPKI State valid
```