The RPKI & Origin Validation

AfNOG / Dar es Salaam

2011.06.03

Randy Bush <randy@psg.com>

Rob Austein <sra@isc.org>

Steve Bellovin <smb@cs.columbia.edu>

Michael Elkins <me@sigpipe.net>

And a cast of thousands! Well, dozens :)

Routing is Very Fragile

- How long can we survive on *The Web as Random Acts of Kindness*, TED Talk by Jonathan Zittrain?

- 99% of mis-announcements are accidental originations of someone else's prefix -- Google, UU, IIJ, …

Why Origin Validation?

- Prevent YouTube accident
- 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 next steps, last talk today

The Goal

- Keep the Internet working!!!
- Seriously reduce routing damage from mis-configuration, mis-origination

Non-Goals

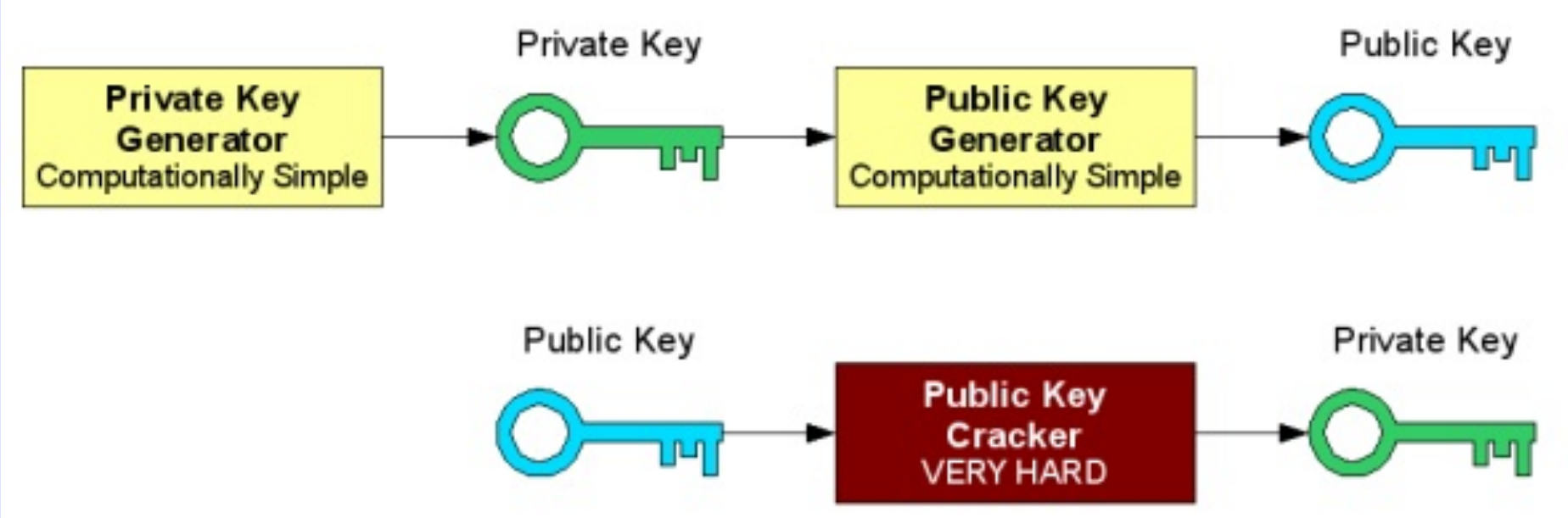
- Prevent Malicious Attacks
- Keep RIRs in business by selling X.509 Certificates

Resource Public Key Infrastructure (RPKI)

Public-Key Concept

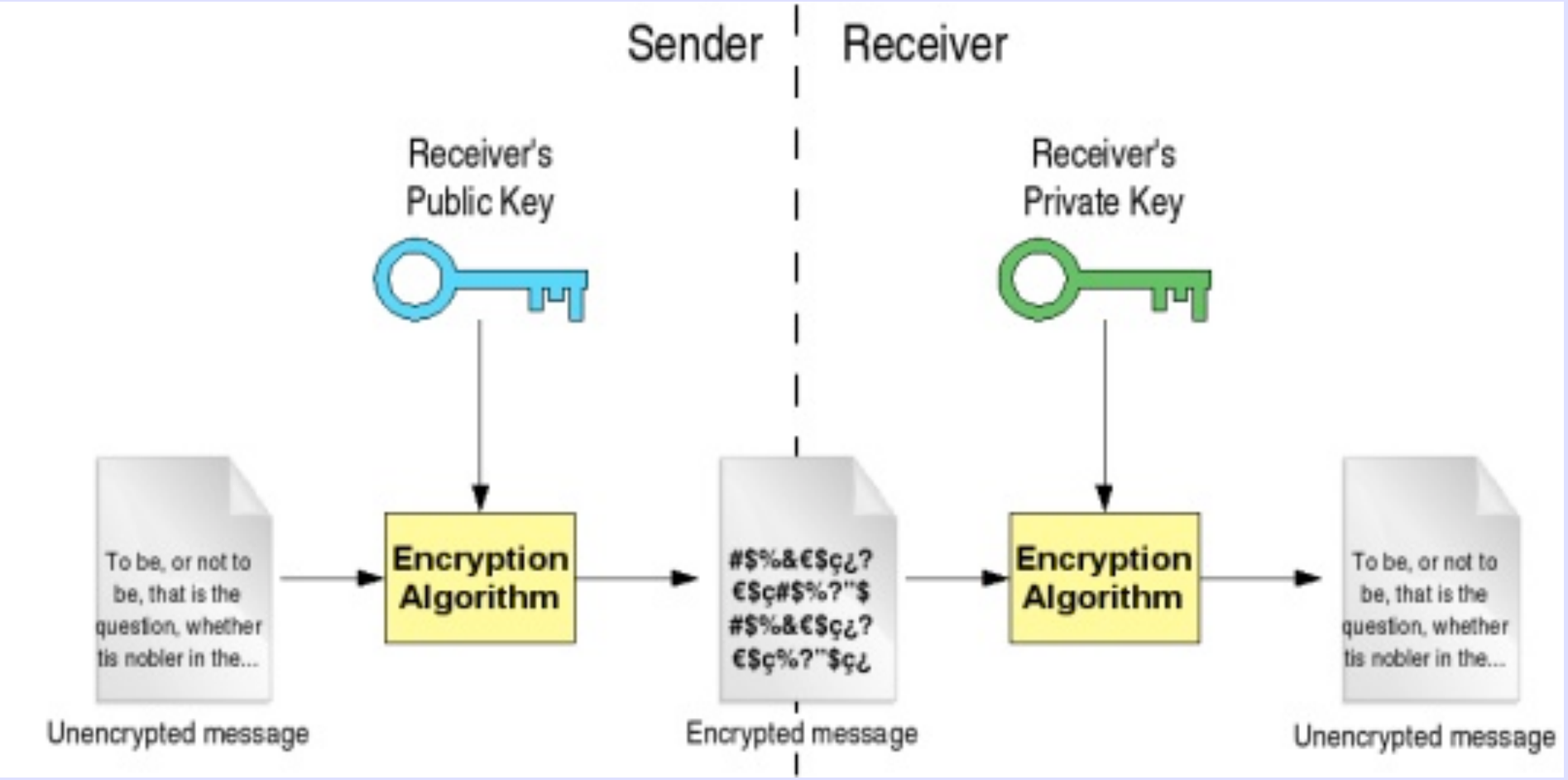
- **Private key**: This key must be known *only* by its owner.
- **Public key**: This key is known to everyone (it is *public*)
- **Relation between both keys**: What one key encrypts, the other one decrypts, and vice versa. That means that if you encrypt something with my public key (which you would know, because it's public :-), I would need my private key to decrypt the message.

Key Generation

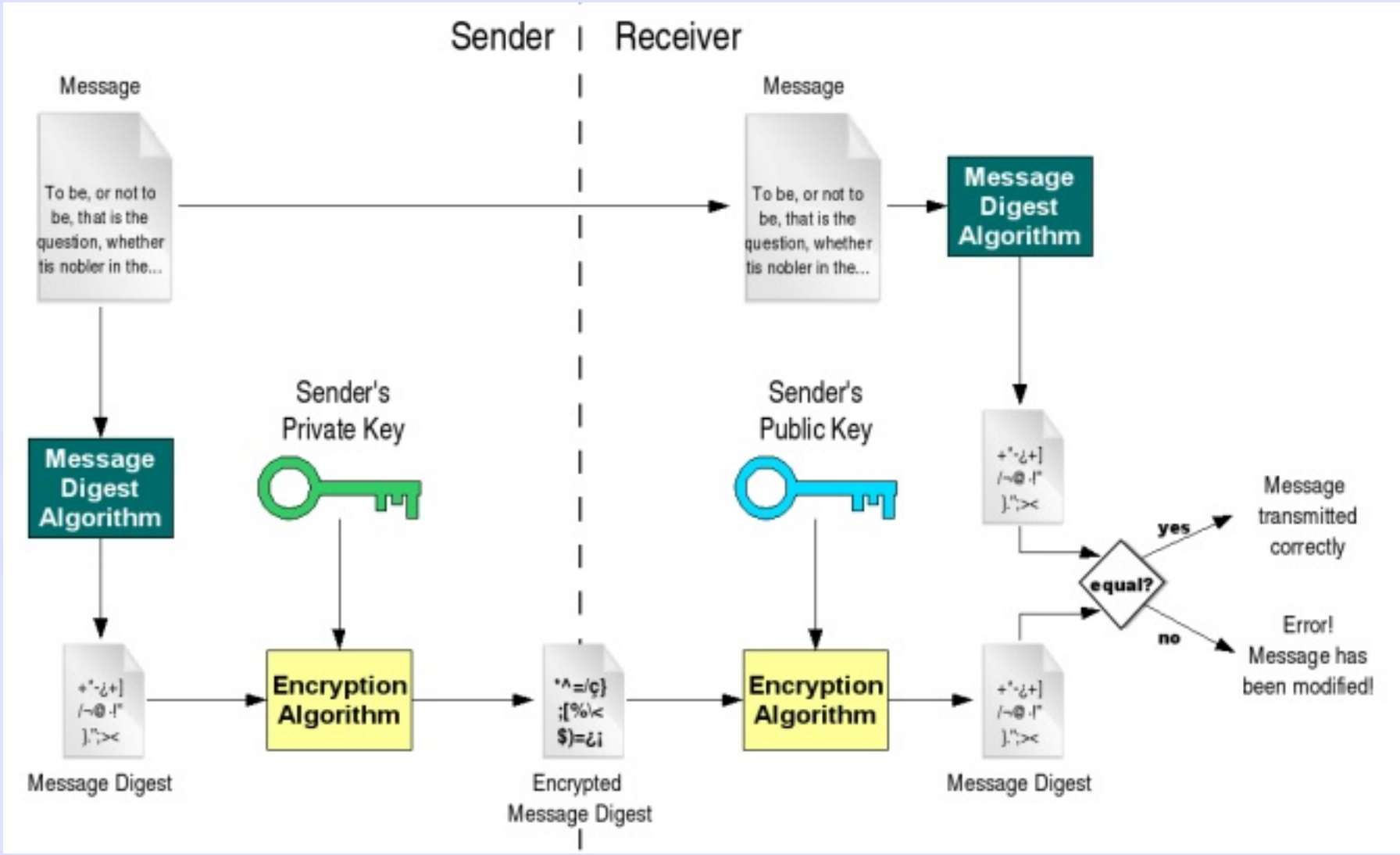


Stolen from - http://gdp.globus.org/gt4-tutorial/multiplehtml/ch09s03.html

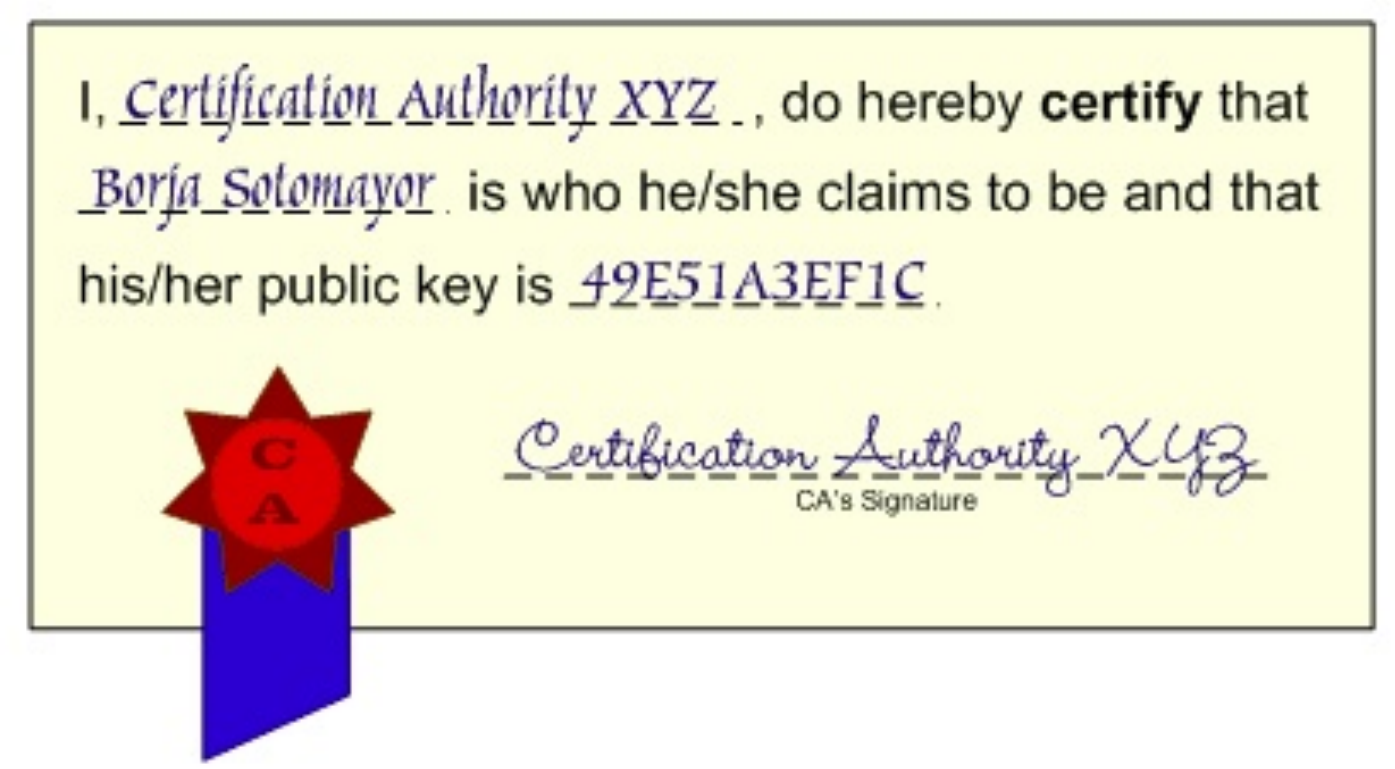
En/DeCryption



Digital Signature

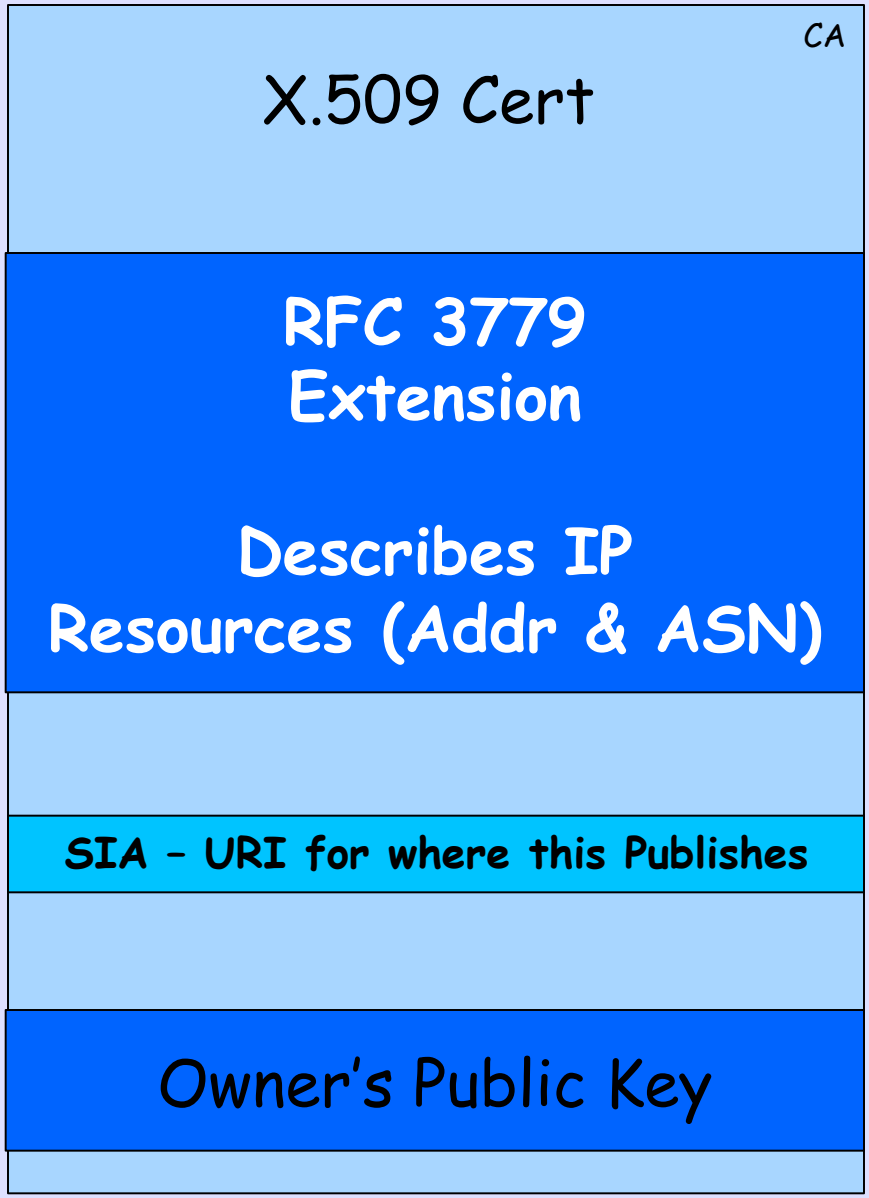


Certificate

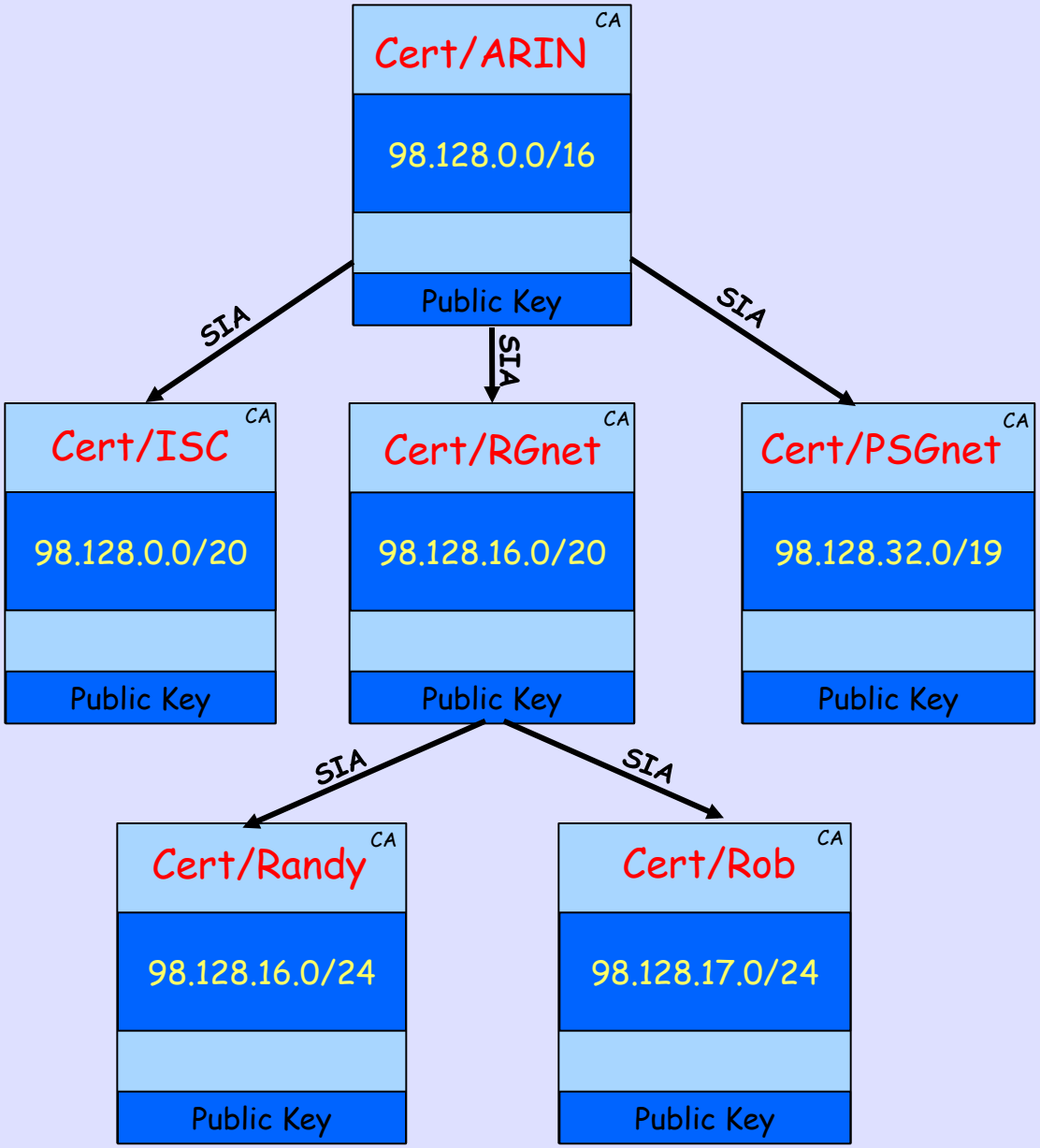


X.509 RPKI Being Developed & Deployed by IANA, RIRs, and Operators

X.509 Certificate w/ 3779 Ext

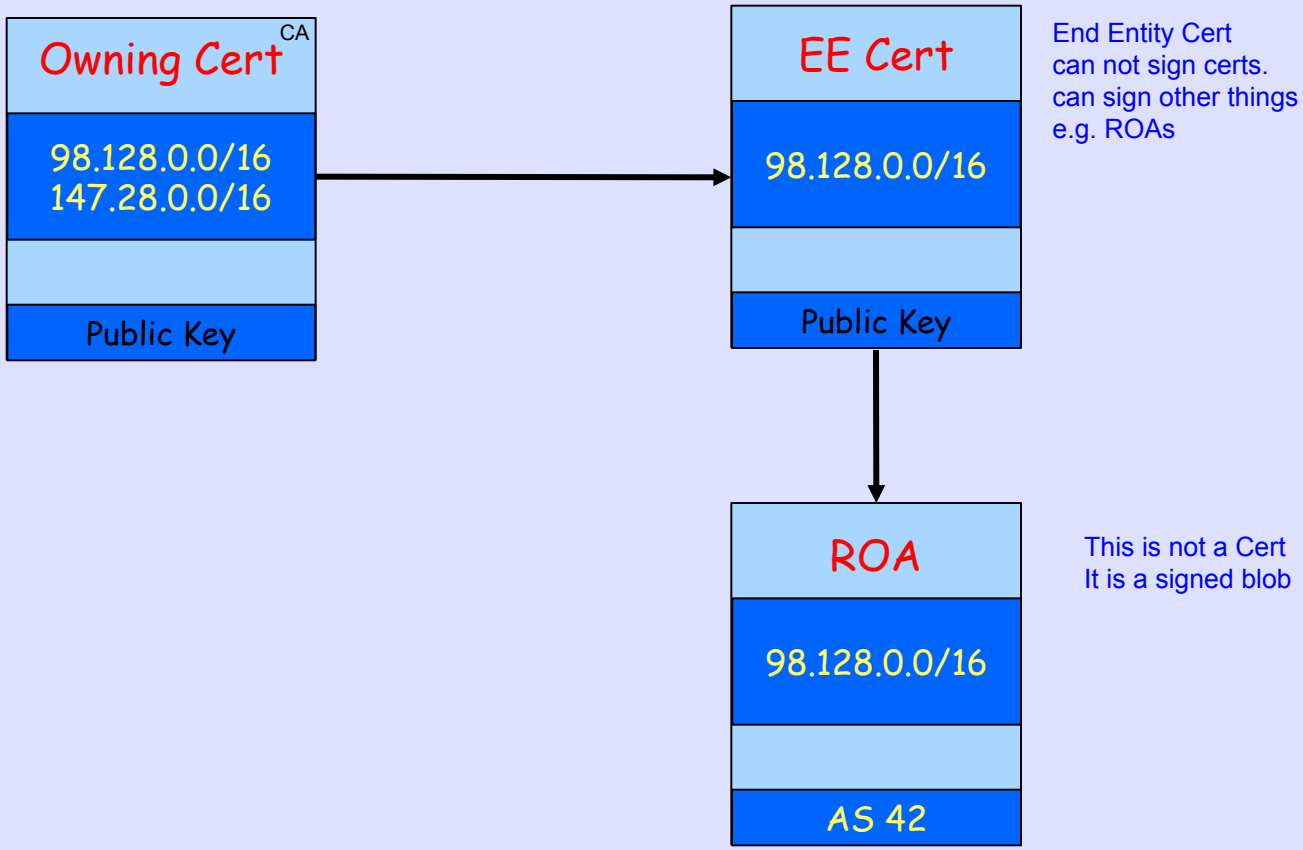


Certificate Hierarchy follows Allocation Hierarchy



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

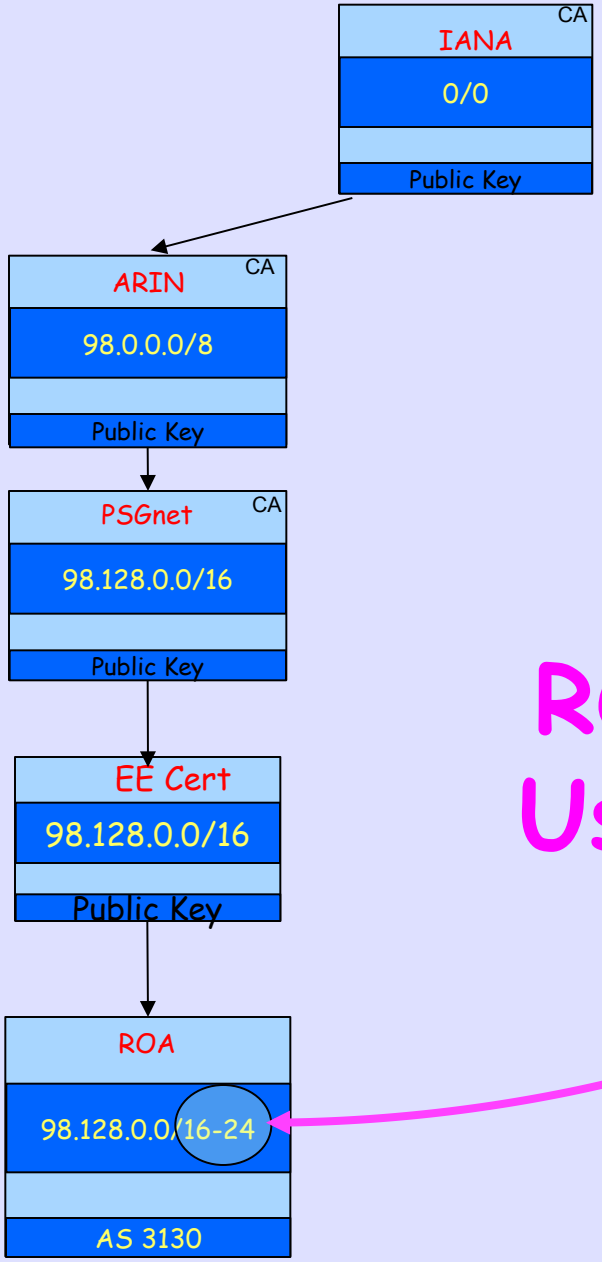
Route Origin Authorization (ROA)



PSGnet /16 Experimental Allocation from ARIN

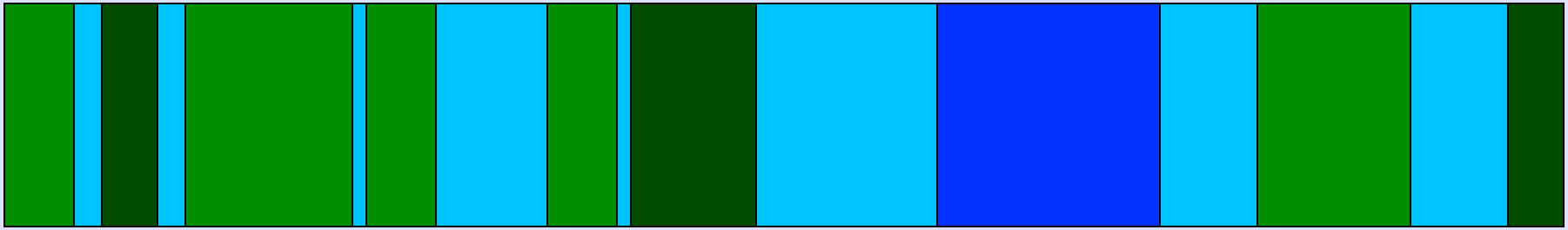
Announces 256 /24s

Too Many EE Certs and ROAs, Yucchhy!



ROA Aggregation Using Max Length

Allocation in Reality



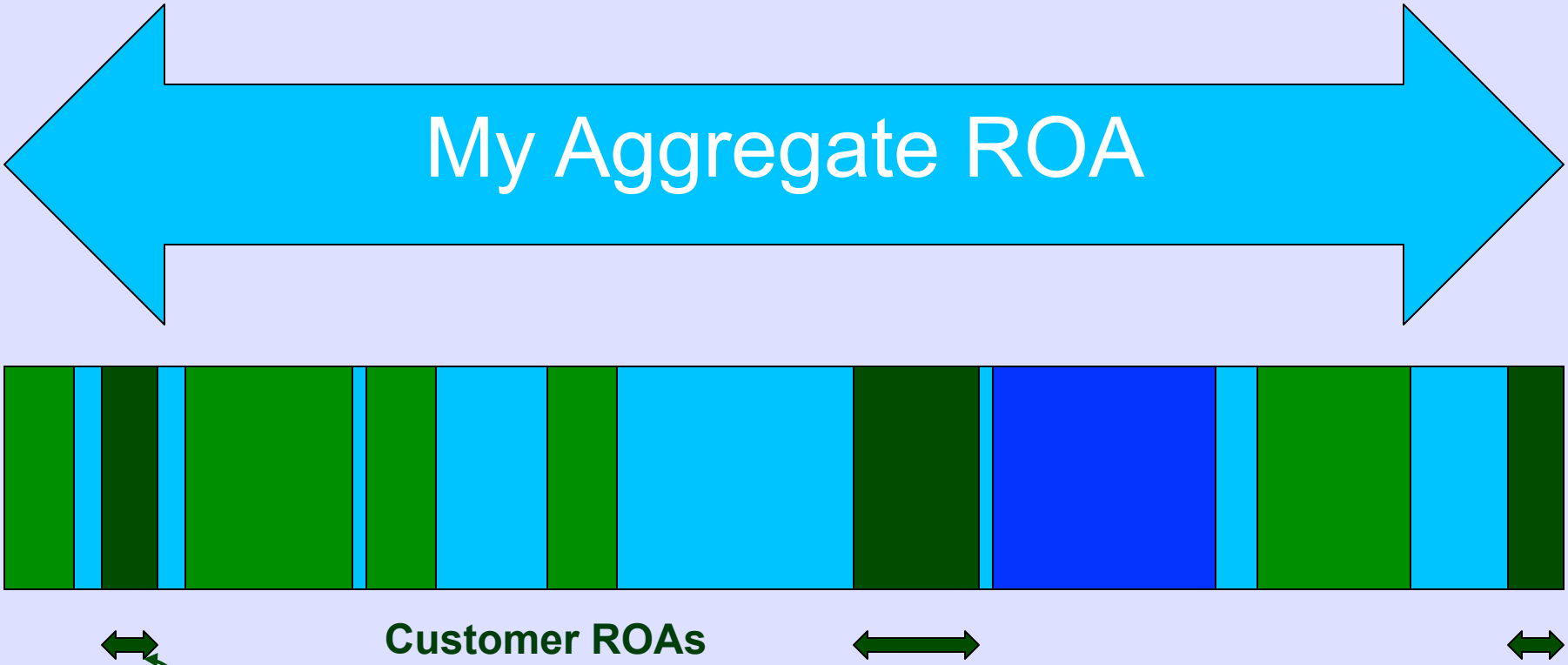
My Infrastructure

BGP Cust

Static (non BGP) Cust

Unused

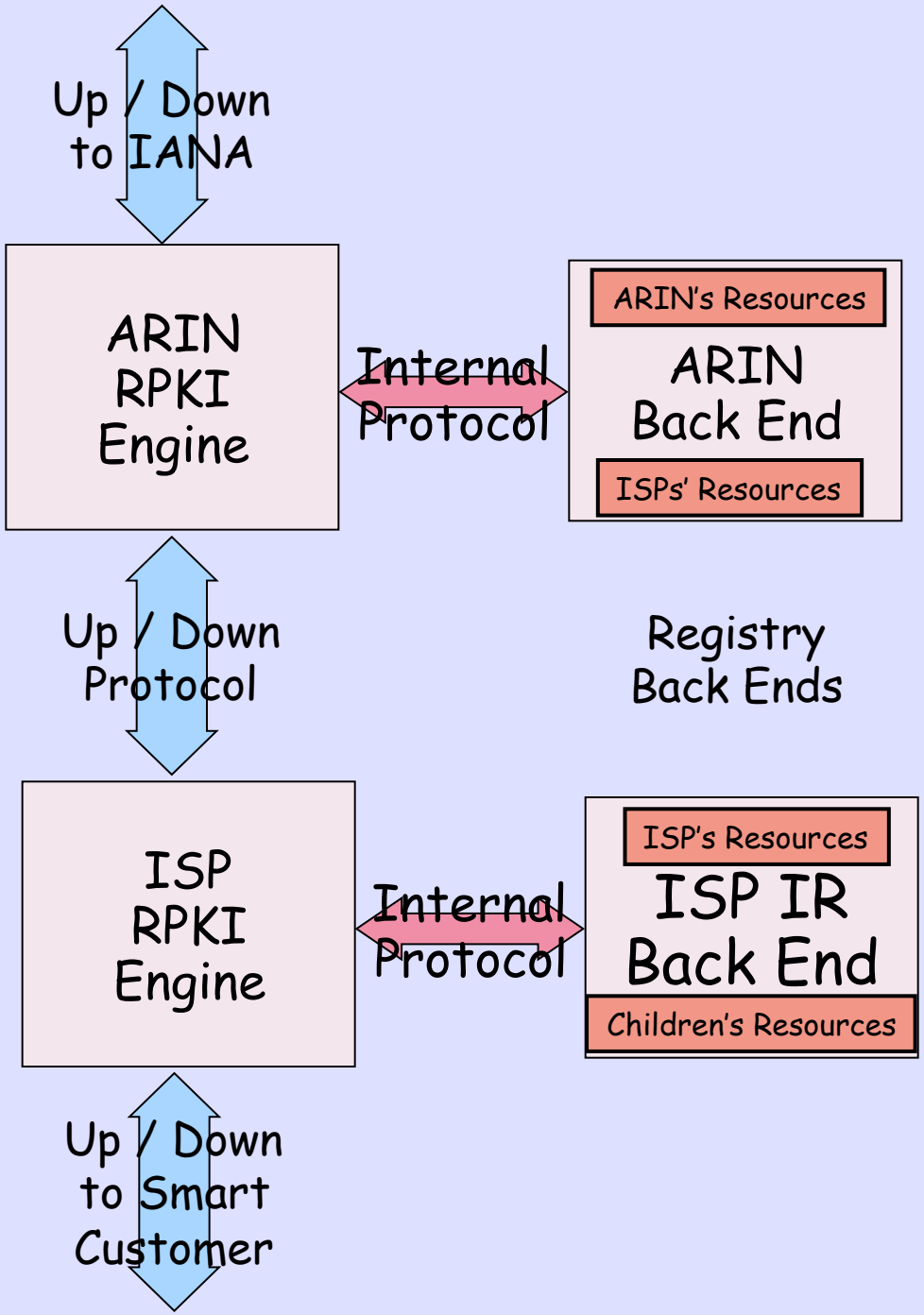
ROA Use

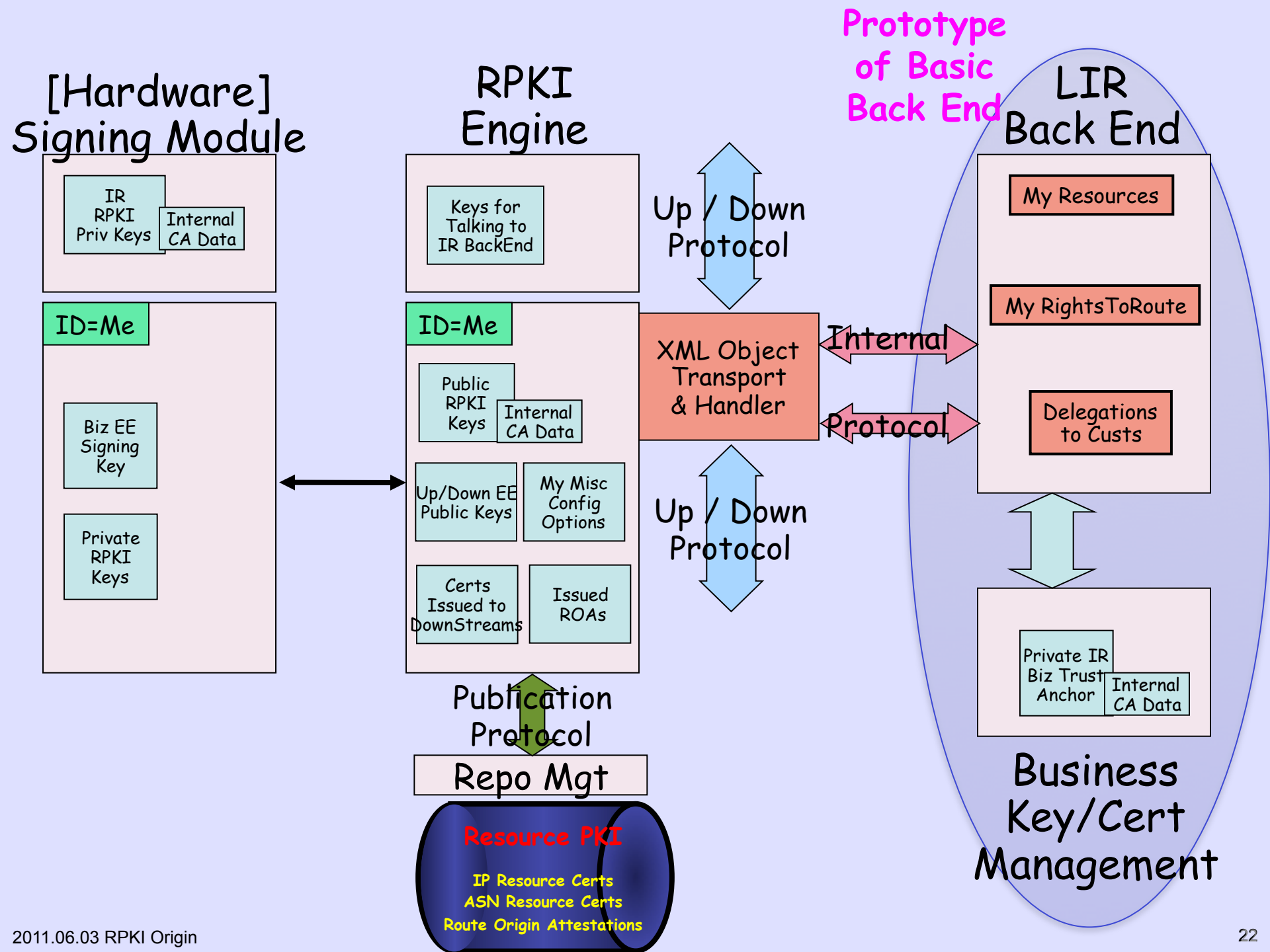

My Aggregate ROA
Customer ROAs
I Generate for 'Lazy' Customer
My Infrastructure
BGP Cust
Static (non BGP) Cust
Unused

Running Code

And the Three RPKI Protocols

Parent and Child

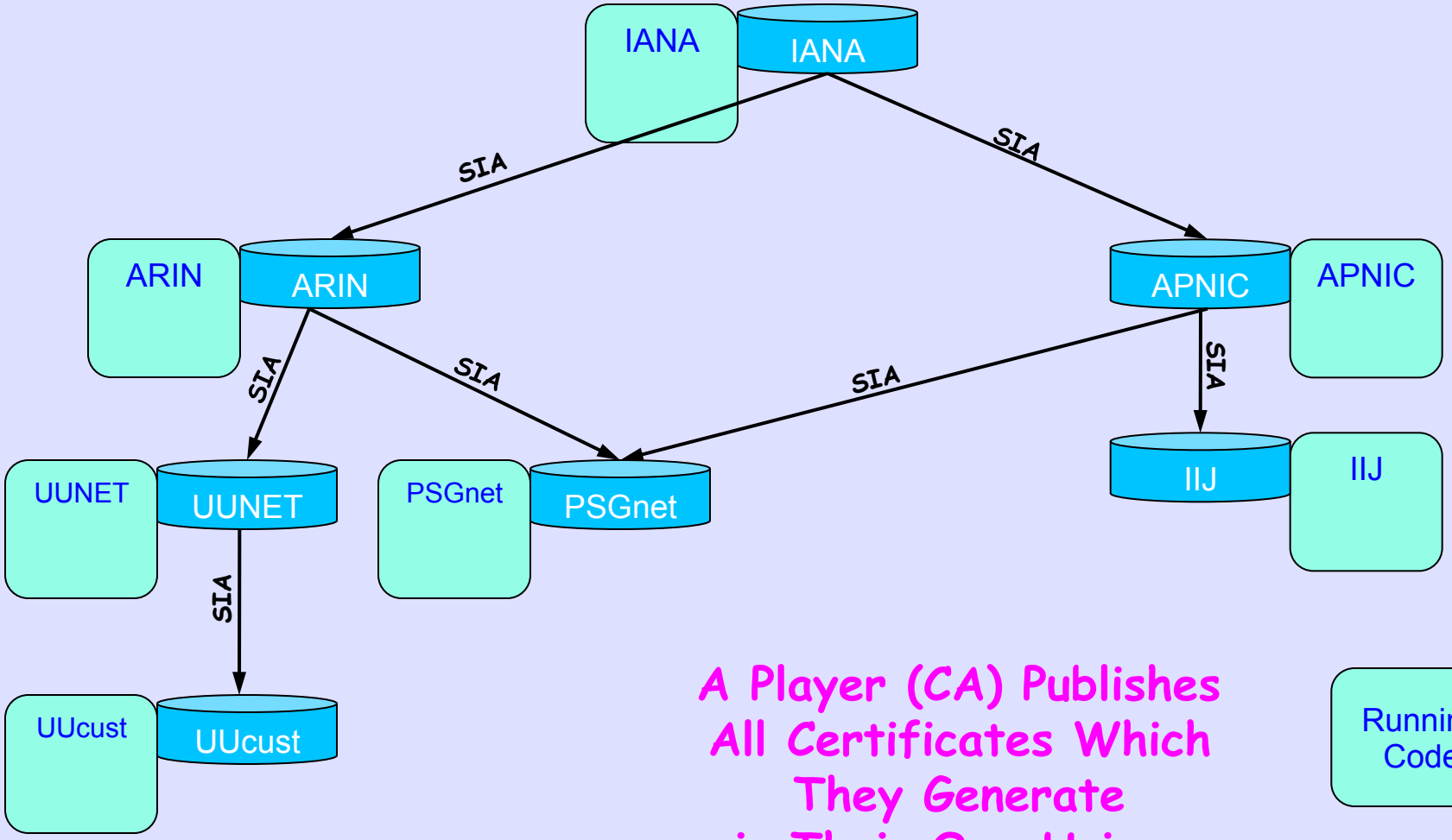



Prototype of Basic Back End
[Hardware] Signing Module
IR RPKI Priv Keys
Internal CA Data
ID=Me
Biz EE Signing Key
Private RPKI Keys
RPKI Engine
Keys for Talking to IR BackEnd
ID=Me
Public RPKI Keys
Internal CA Data
Up/Down EE Public Keys
My Misc Config Options
Certs Issued to DownStreams
Issued ROAs
Up / Down Protocol
XML Object Transport & Handler
Internal
Protocol
Up / Down Protocol
LIR Back End
My Resources
My RightsToRoute
Delegations to Custs
Private IR Biz Trust Anchor
Internal CA Data
Business Key/Cert Management
Publication Protocol
Repo Mgt
Resource PKI
IP Resource Certs
ASN Resource Certs
Route Origin Attestations

Big, Centralized, & Scary

We Don't Do This

Distributed RPKI DataBase



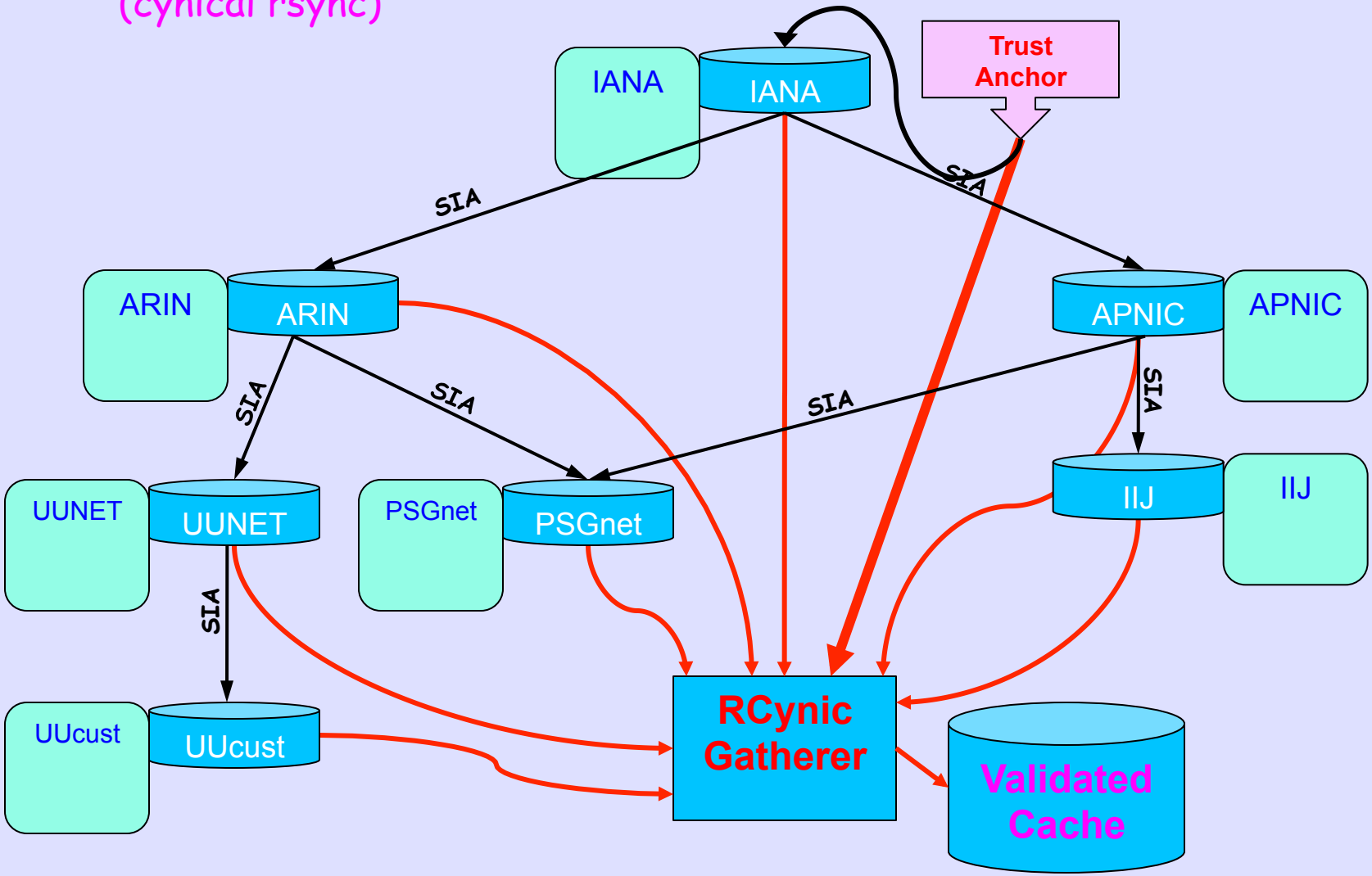
A Player (CA) Publishes
All Certificates Which
They Generate
in Their Own Unique
Publication Point

Running Code

Repository

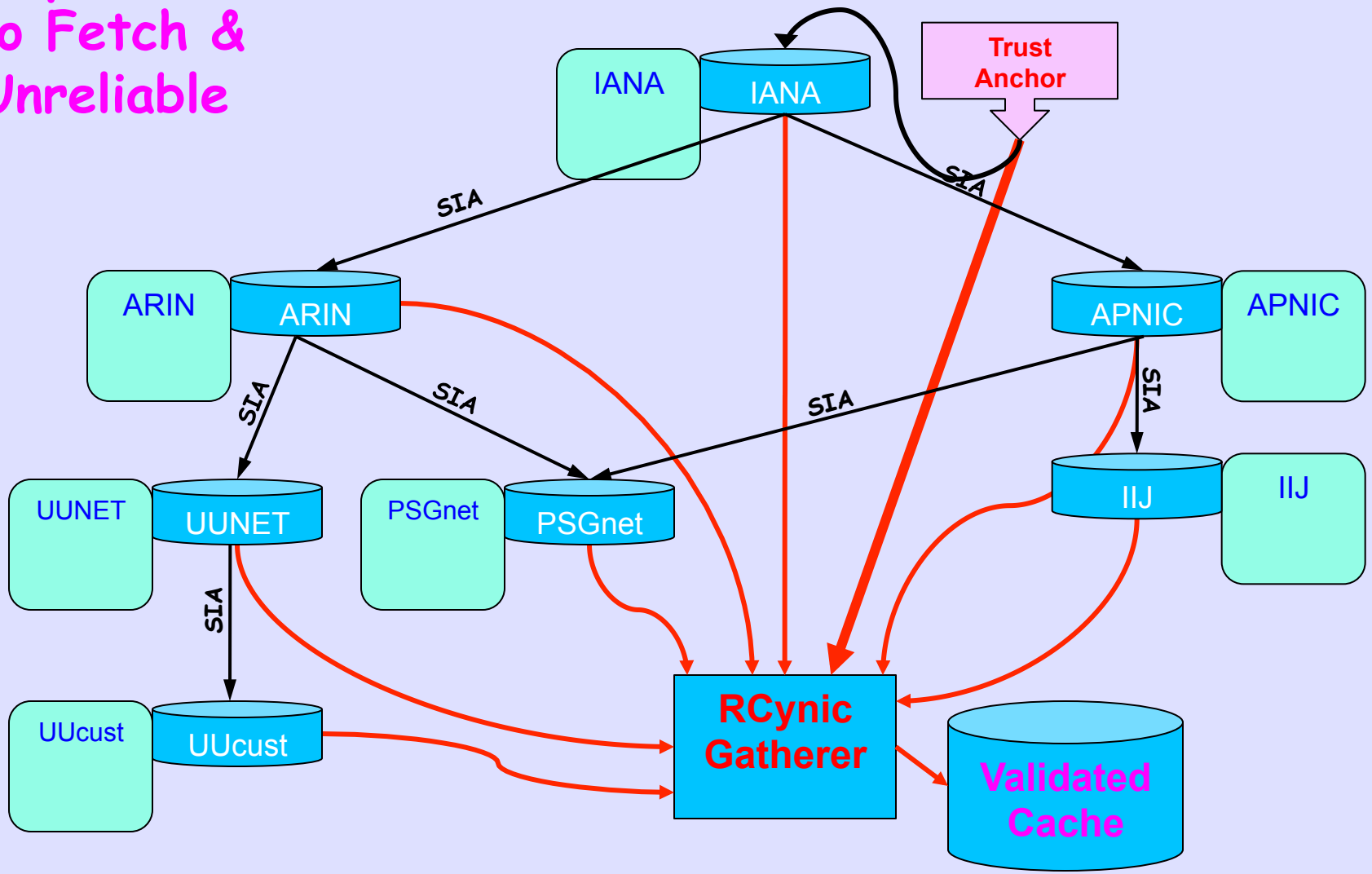
RCynic Cache Gatherer

(cynical rsync)

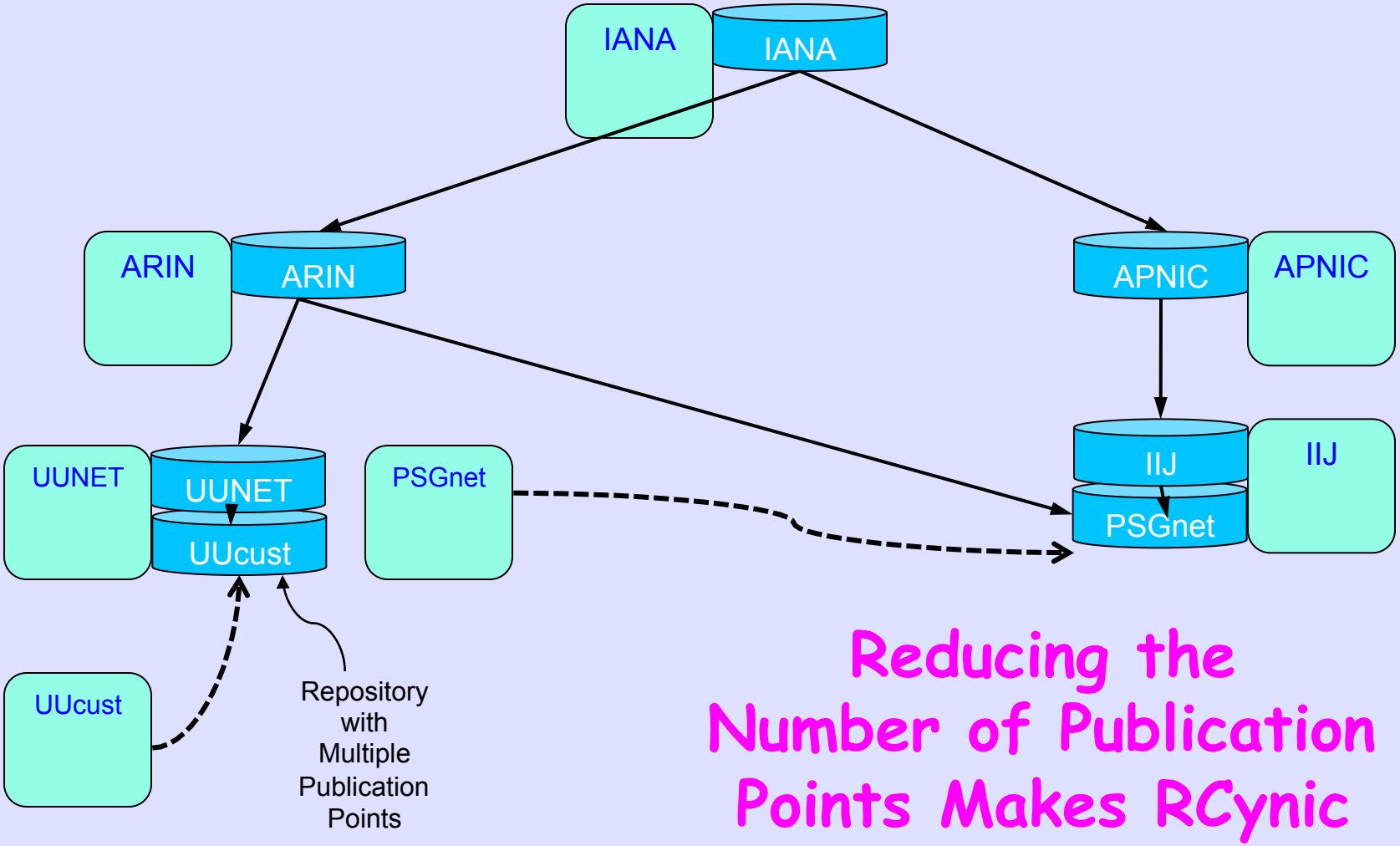


Reliability Issue

Expensive
To Fetch &
Unreliable

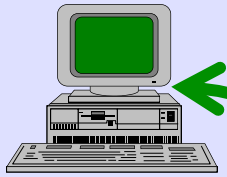


Reliability Via Hosted Publication



Reducing the Number of Publication Points Makes RCynic More Efficient

A Usage Scenario



User Web GUI

98% of an RIR's Users
10% of an RIR's IP Space

RPKI Engine

Resources [OrgID]

IR's Database(s)

My RightsToRoute

Delegations to Custs

Internal

Protocol

Keys for Talking to IR BackEnd

ID=Me

Public RPKI Keys

Internal CA Data

Up/Down EE Public Keys

My Misc Config Options

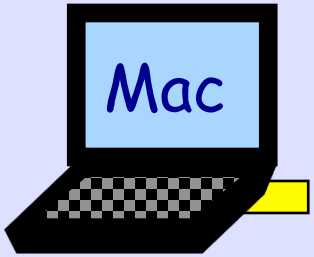
Certs Issued to DownStreams

Issued ROAs

Publication Protocol

Publication Point

Contract Out To Google



Front End GUI & Management

Up / Down Protocol

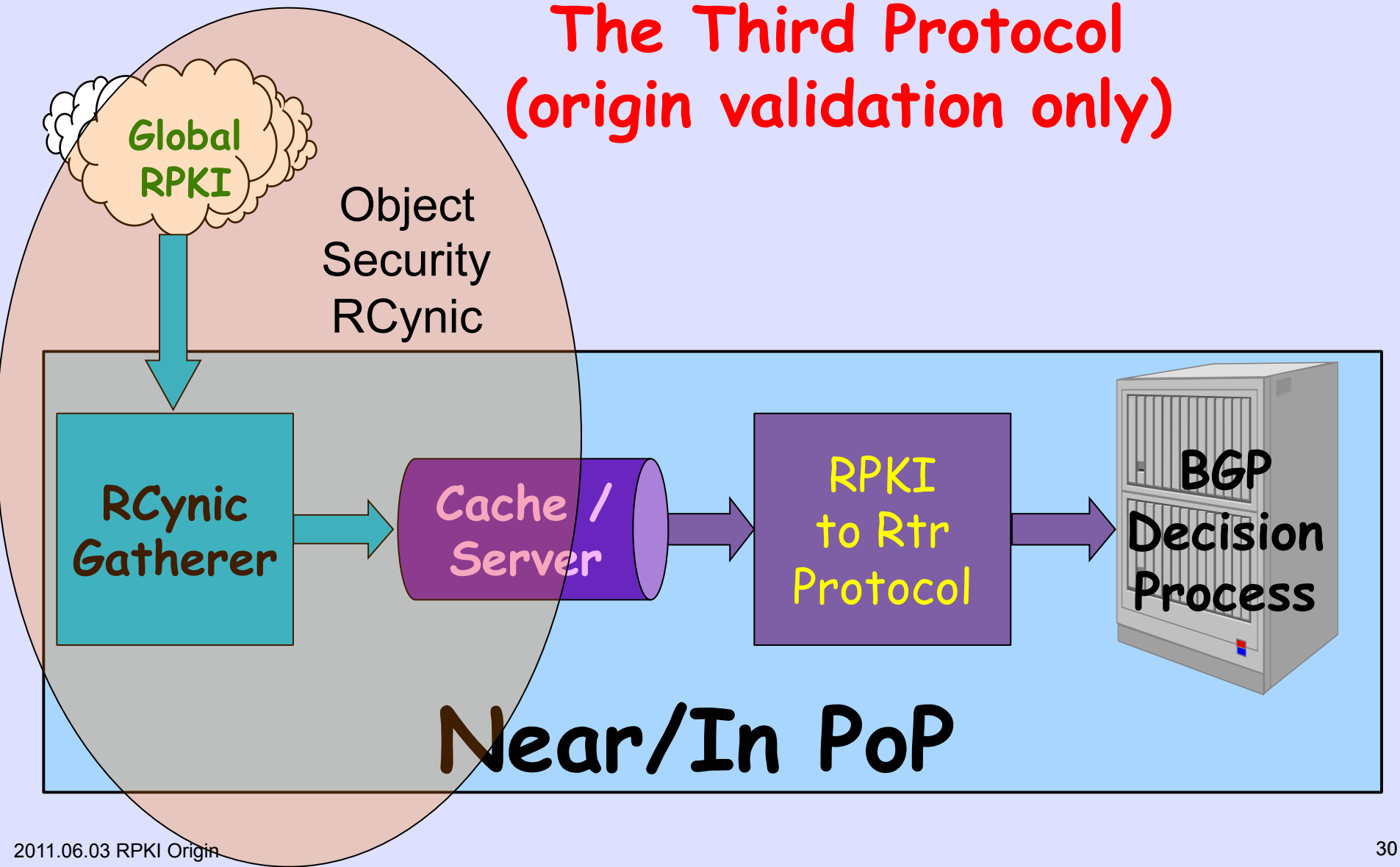
2% of an RIR's Users
90% of an RIR's IP Space

Origin Validation

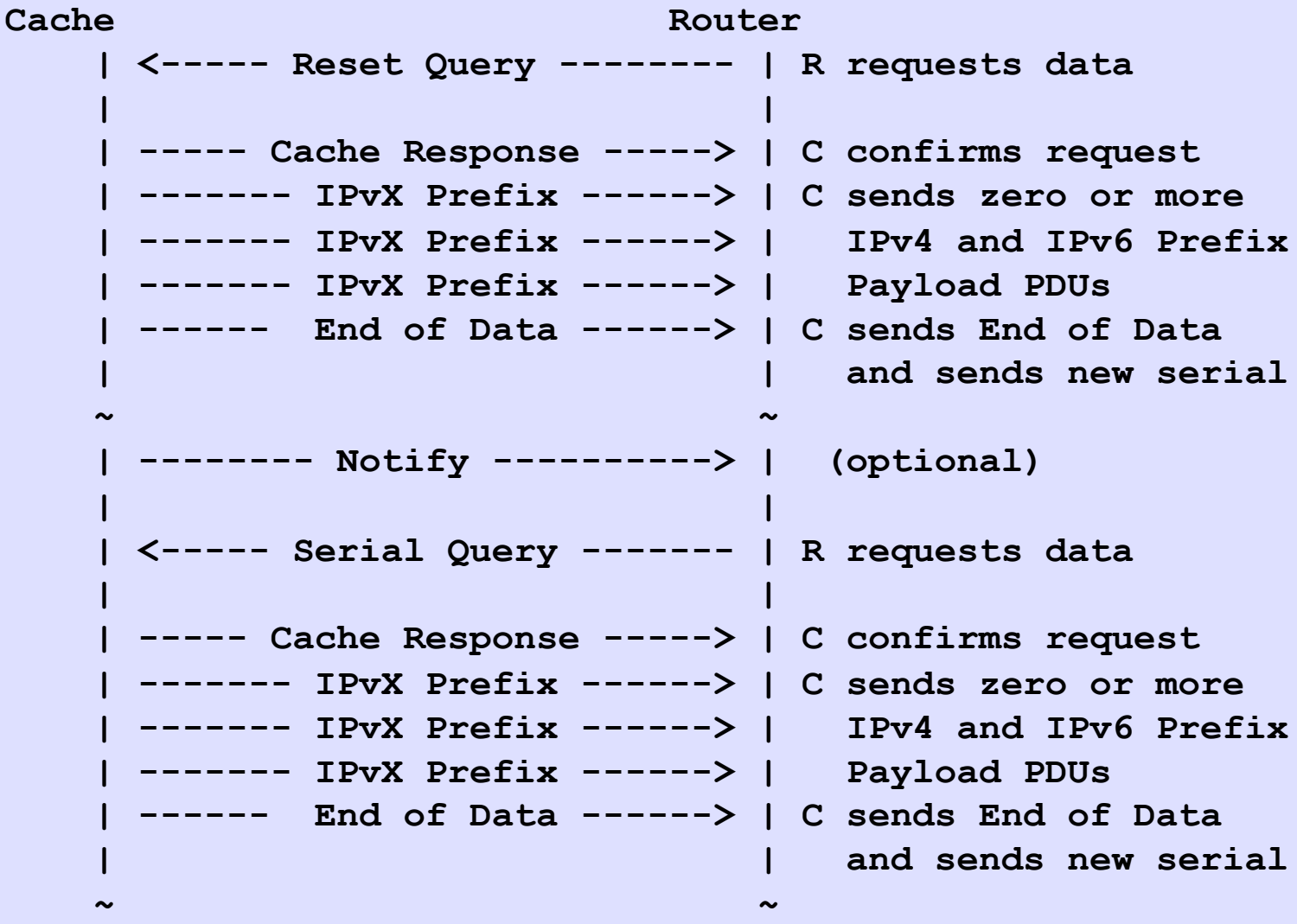
- Cisco IOS and IOS-XR test code have Origin Validation now
- Juniper has test code now
- Work continues daily in test routers
- Compute load much less than ACLs from IRR data, **10µsec per update!**

RPKI -> Router

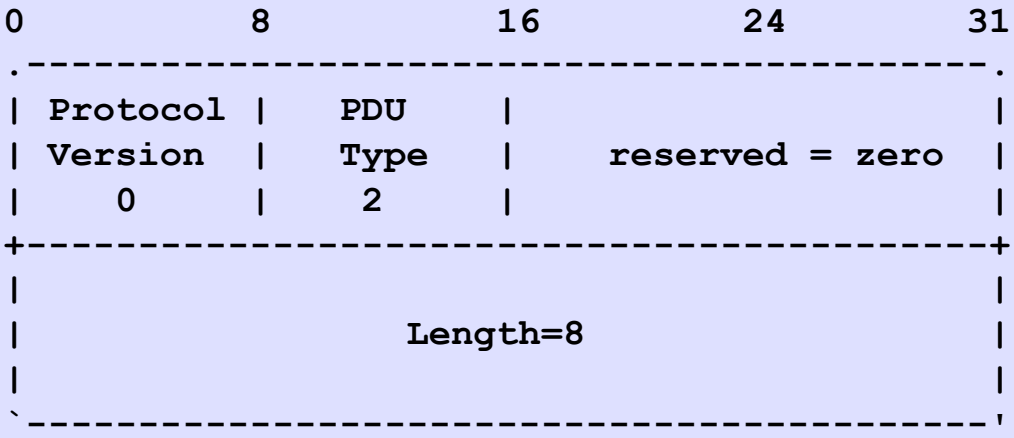
The Third Protocol (origin validation only)



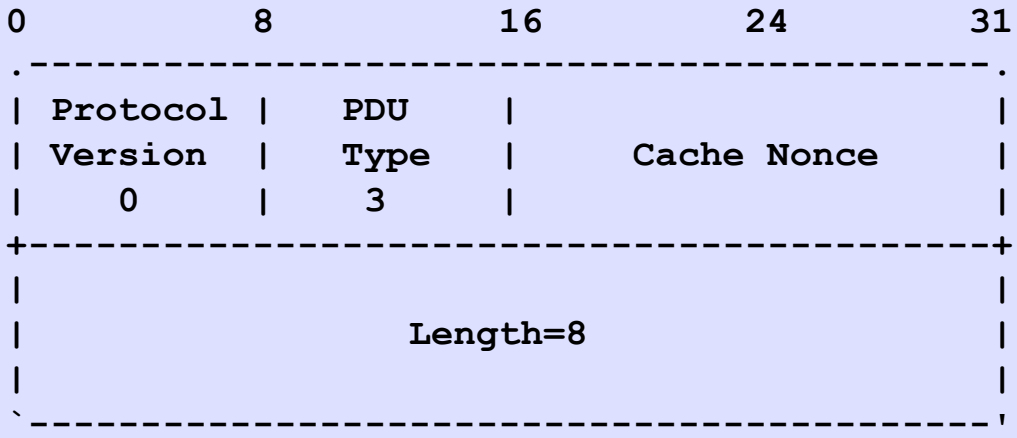
Typical Exchange



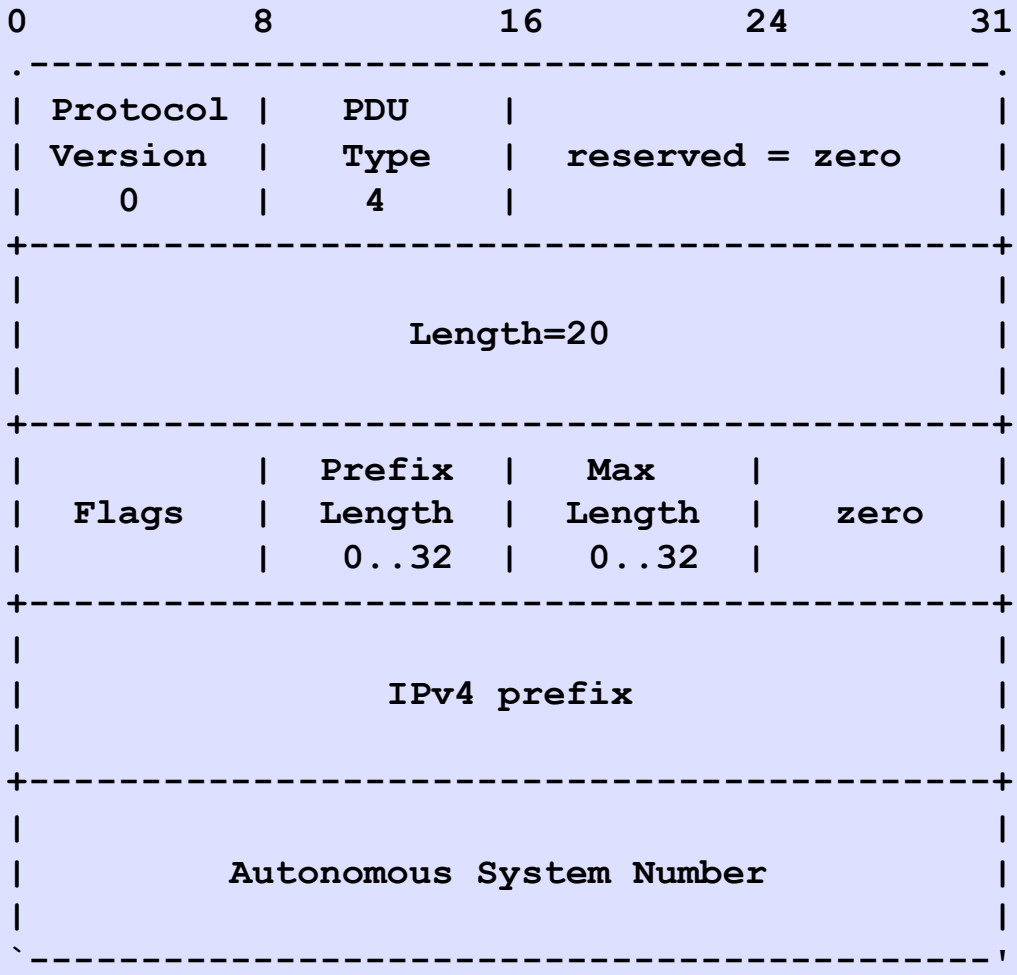
Reset Query


0
8
16
24
31
Protocol Version
0
PDU Type
2
reserved = zero
Length=8

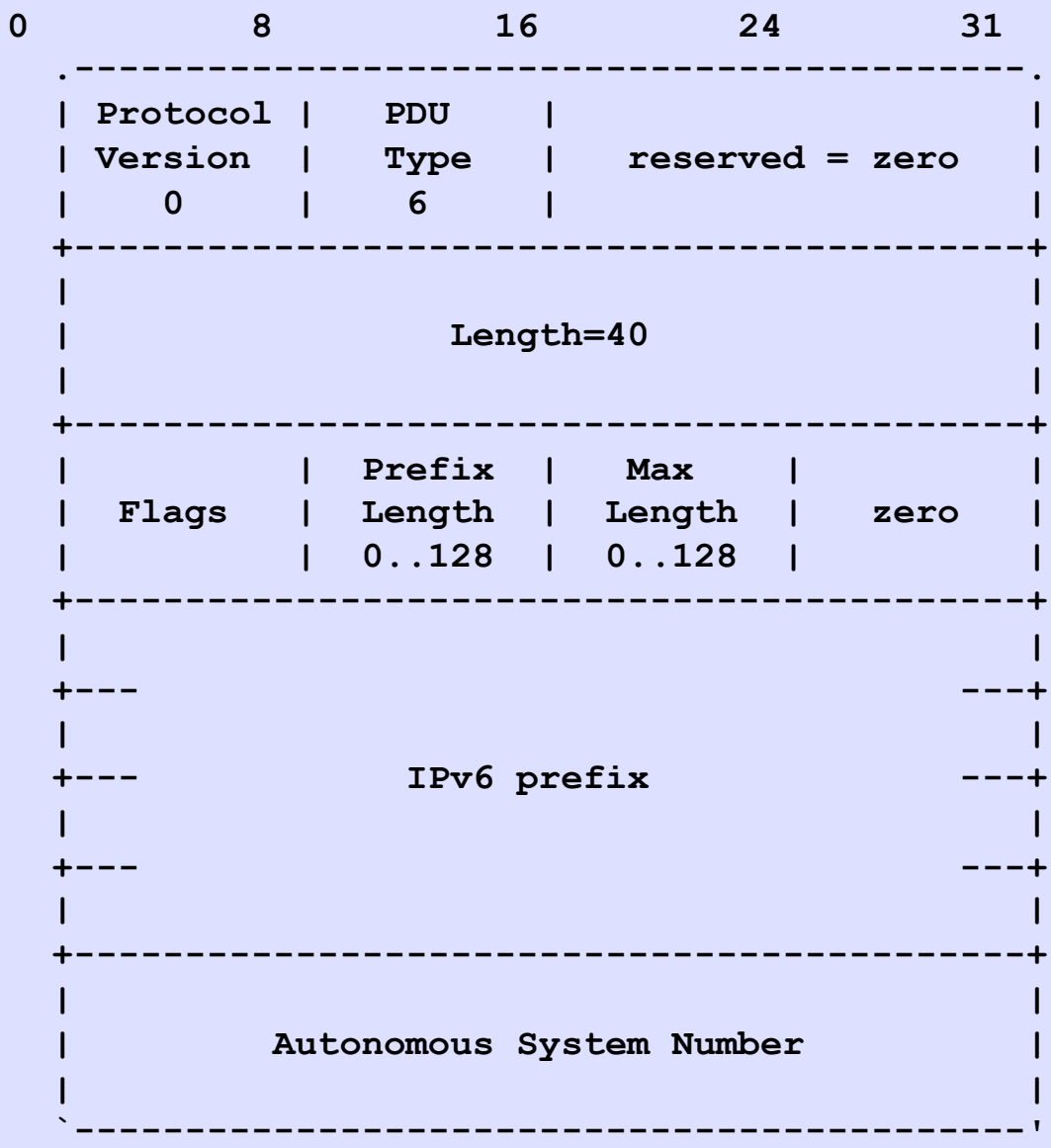
Cache Response



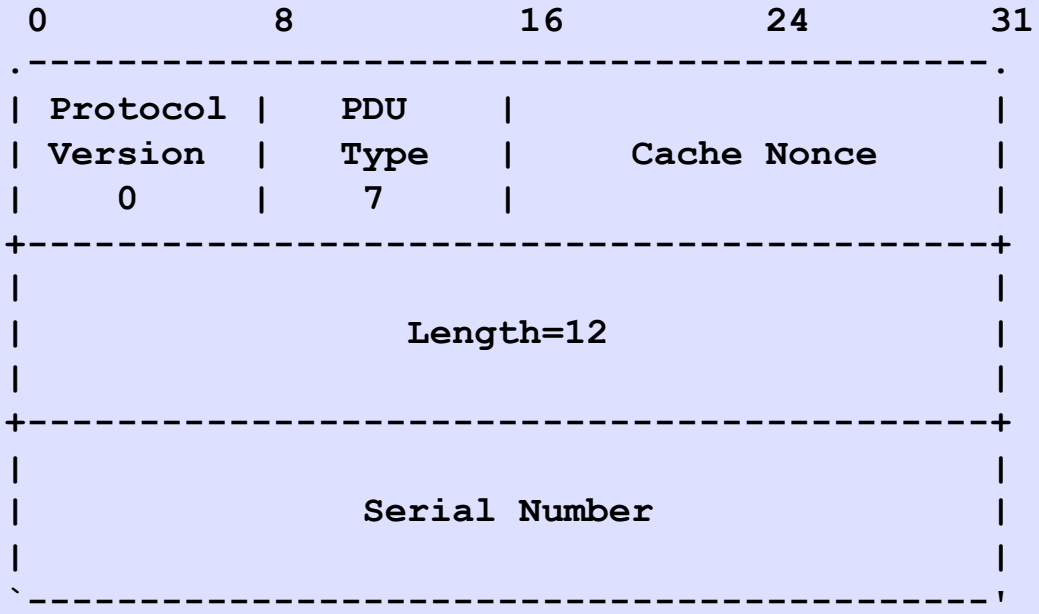
IPv4 Prefix


0
8
16
24
31
Protocol Version 0
PDU Type 4
reserved = zero
Length=20
Flags
Prefix Length 0..32
Max Length 0..32
zero
IPv4 prefix
Autonomous System Number

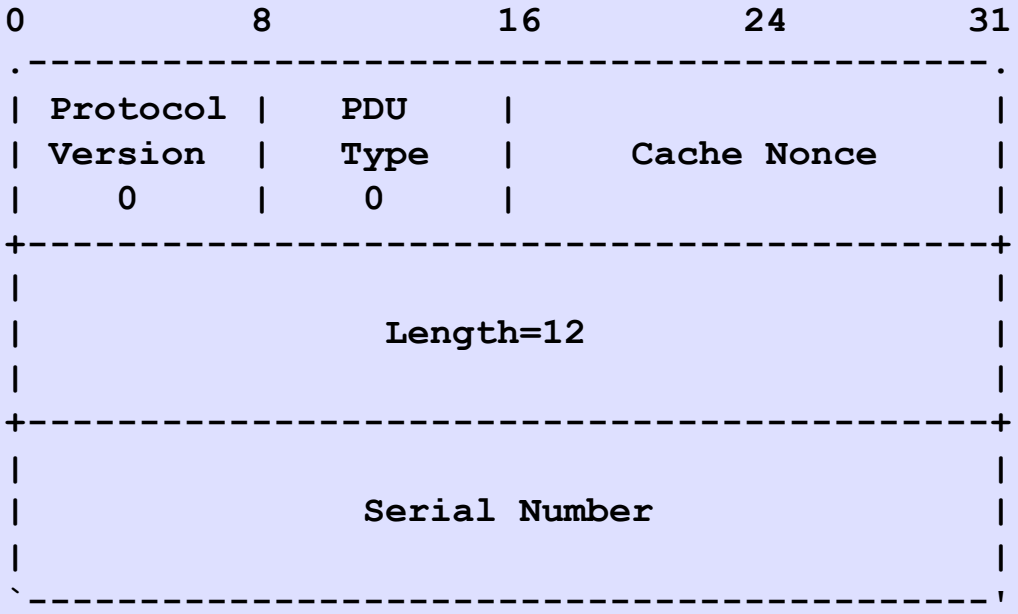
IPv6 Prefix


0
8
16
24
31
Protocol Version
0
PDU Type
6
reserved = zero
Length=40
Flags
Prefix Length
0..128
Max Length
0..128
zero
IPv6 prefix
Autonomous System Number

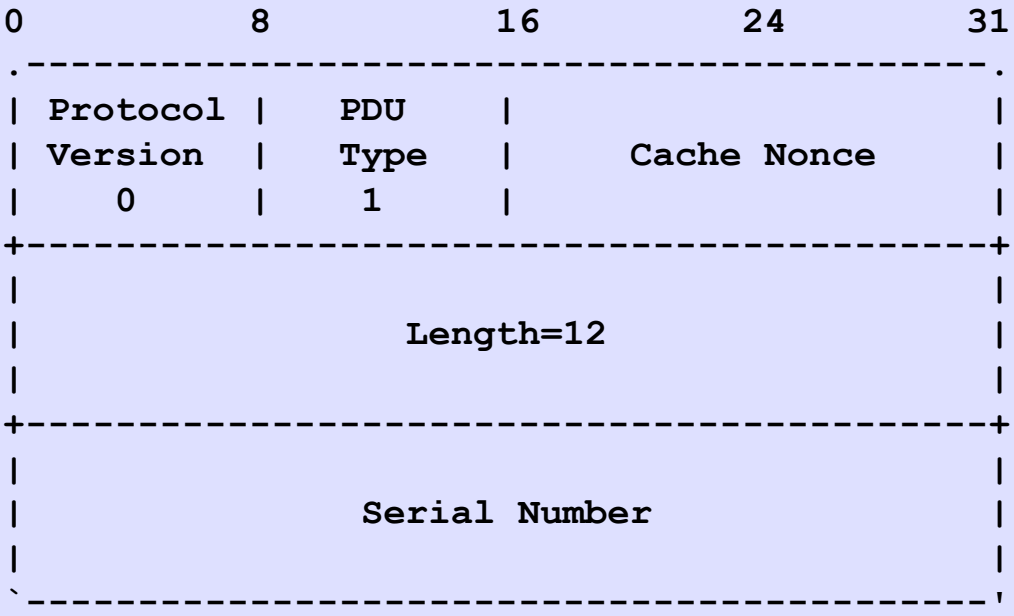
End of Data


0
8
16
24
31
Protocol Version
0
PDU Type
7
Cache Nonce
Length=12
Serial Number

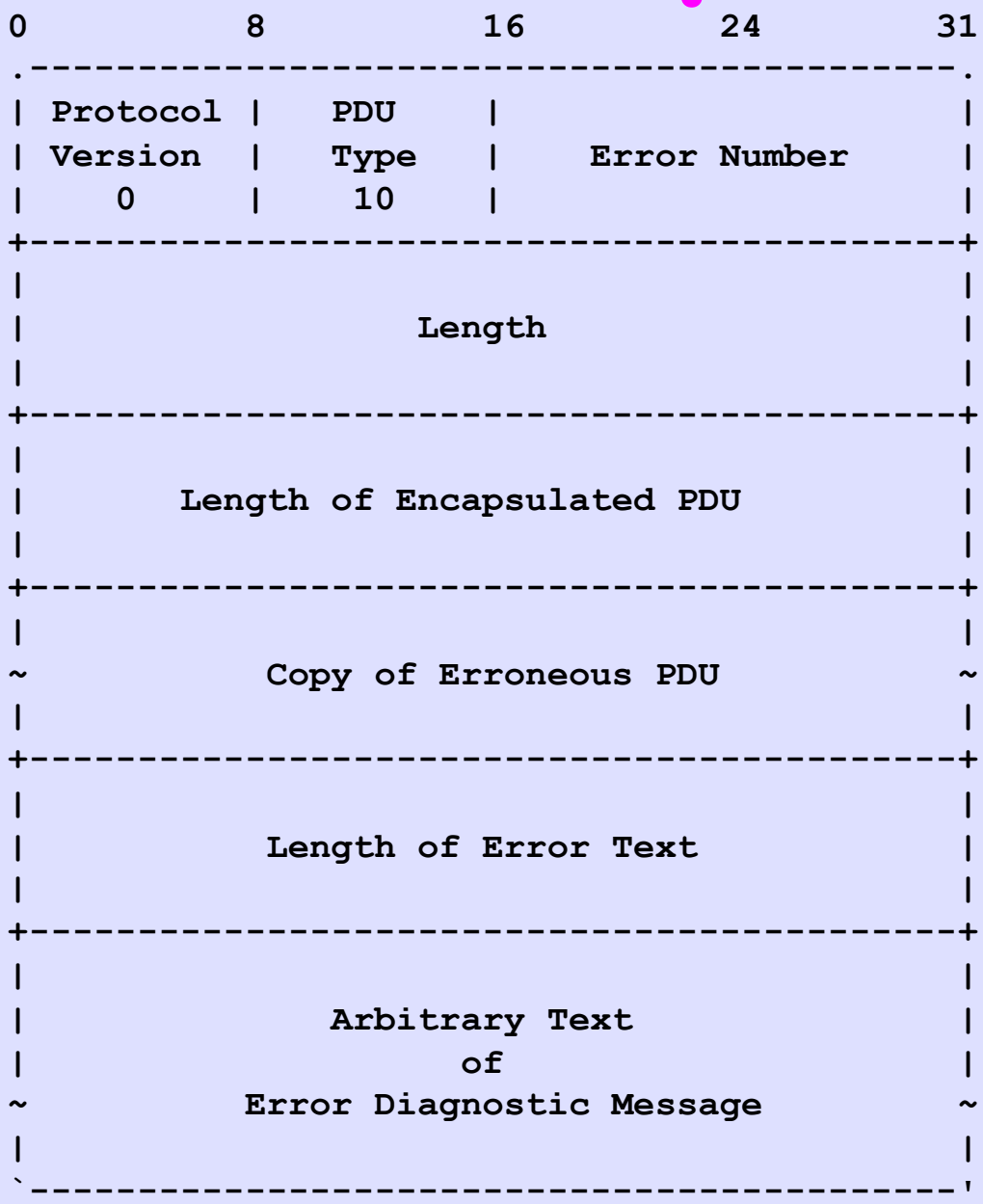
Notify (Think DNS)



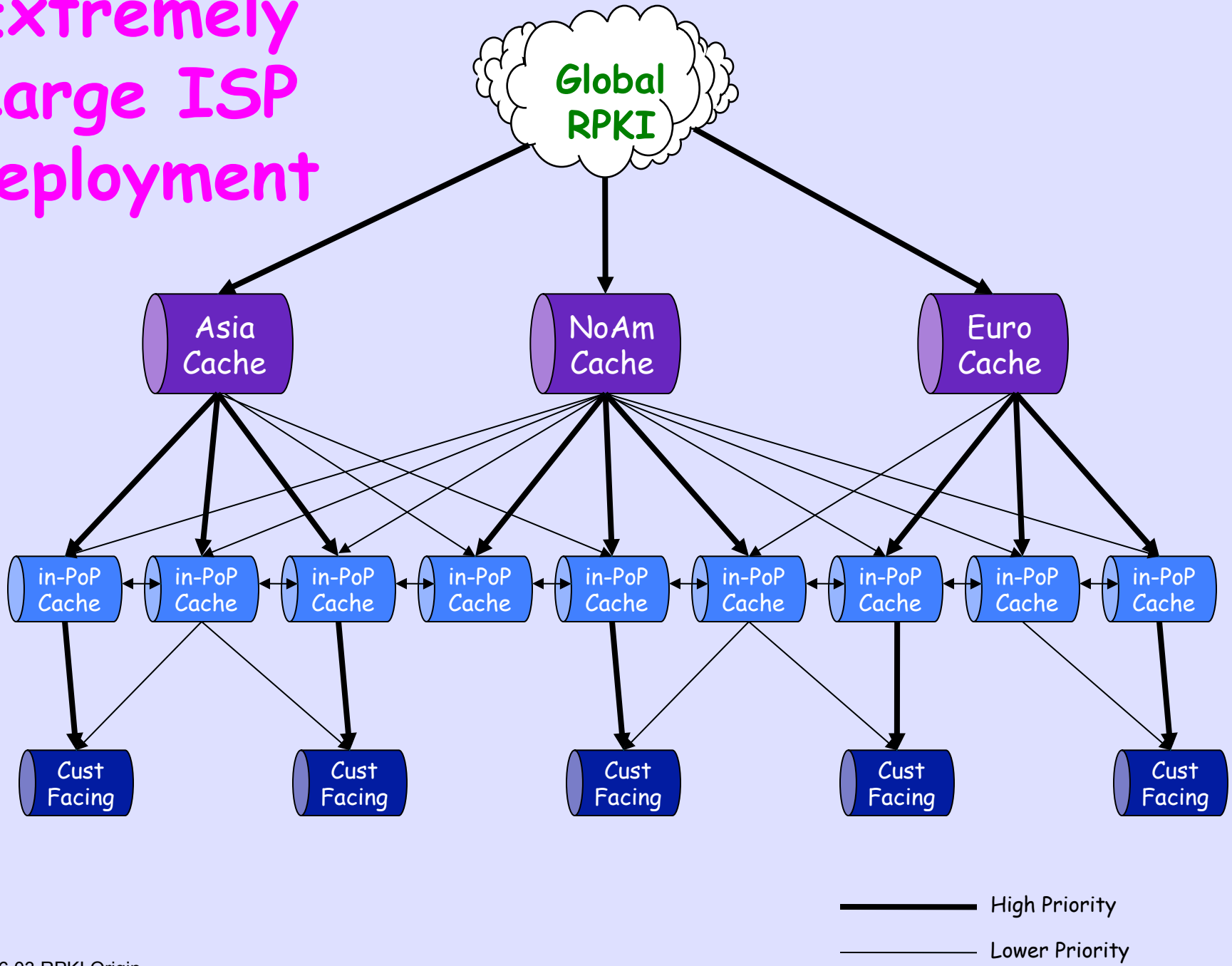
Serial Query



Error Response



Extremely Large ISP Deployment



Configure

```
router bgp 3130
 …
 bgp rpki server tcp 198.180.150.1 port 42420 refresh 120
 bgp bestpath prefix-validate allow-invalid
```

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

Policy Override Knobs

- Disable Validity Check Completely
- Disable Validity Check for a Peer
- Disable Validity Check for Prefixes

When check is disabled, the result is "Not Found," i.e. as if there was no ROA

Look at Table

```
r0.sea#show ip bgp rpki table
76 BGP sovc network entries using 6688 bytes of memory
422 BGP sovc record entries using 8440 bytes of memory

Network          Maxlen  Origin-AS  Source  Neighbor
67.21.36.0/24    24      3970       0       198.180.150.1/424
98.128.0.0/24    24      4128       0       198.180.150.1/424
98.128.0.0/16    16      3130       0       198.180.150.1/424
98.128.6.0/24    24      4128       0       198.180.150.1/424
98.128.9.0/24    24      3130       0       198.180.150.1/424
98.128.30.0/24   24      1234       0       198.180.150.1/424
129.6.128.0/17   17      49         0       198.180.150.1/424
147.28.0.0/16    16      3130       0       198.180.150.1/424
147.28.224.0/19  19      4128       0       198.180.150.1/424
```

Defaults

- Origin Validation is Enabled if you have configured a cache server peering
- Default Poll Interval is 30 Minutes
- No Effect on Policy unless you have configured it

Good Dog!

```
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
```

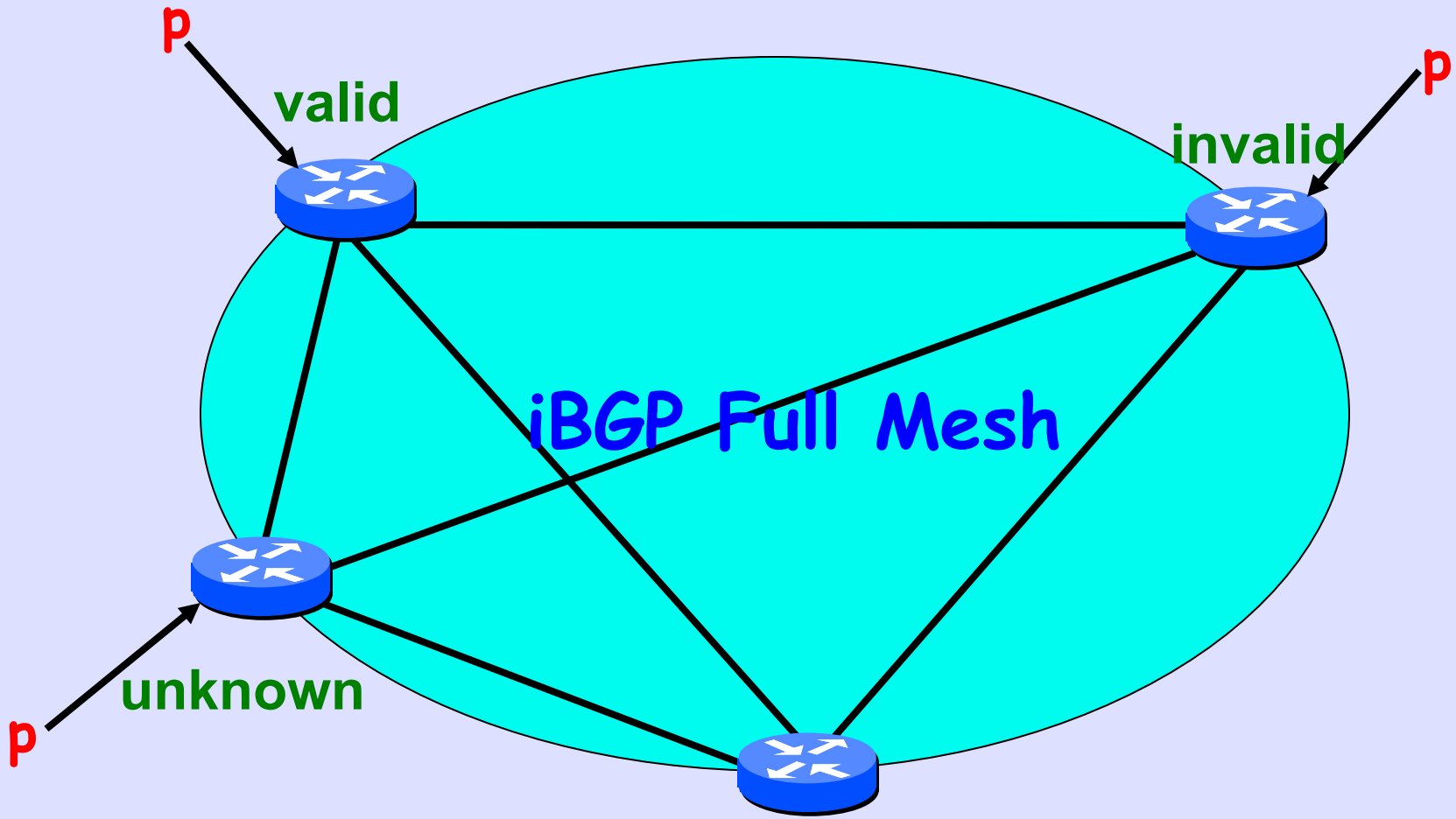
Bad Dog!

```
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
```

Strange Dog!

```
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
```

iBGP Hides Validity State



which do i choose?
why do i choose it?

The Solution is to Allow Operator to Test and then Set 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
```

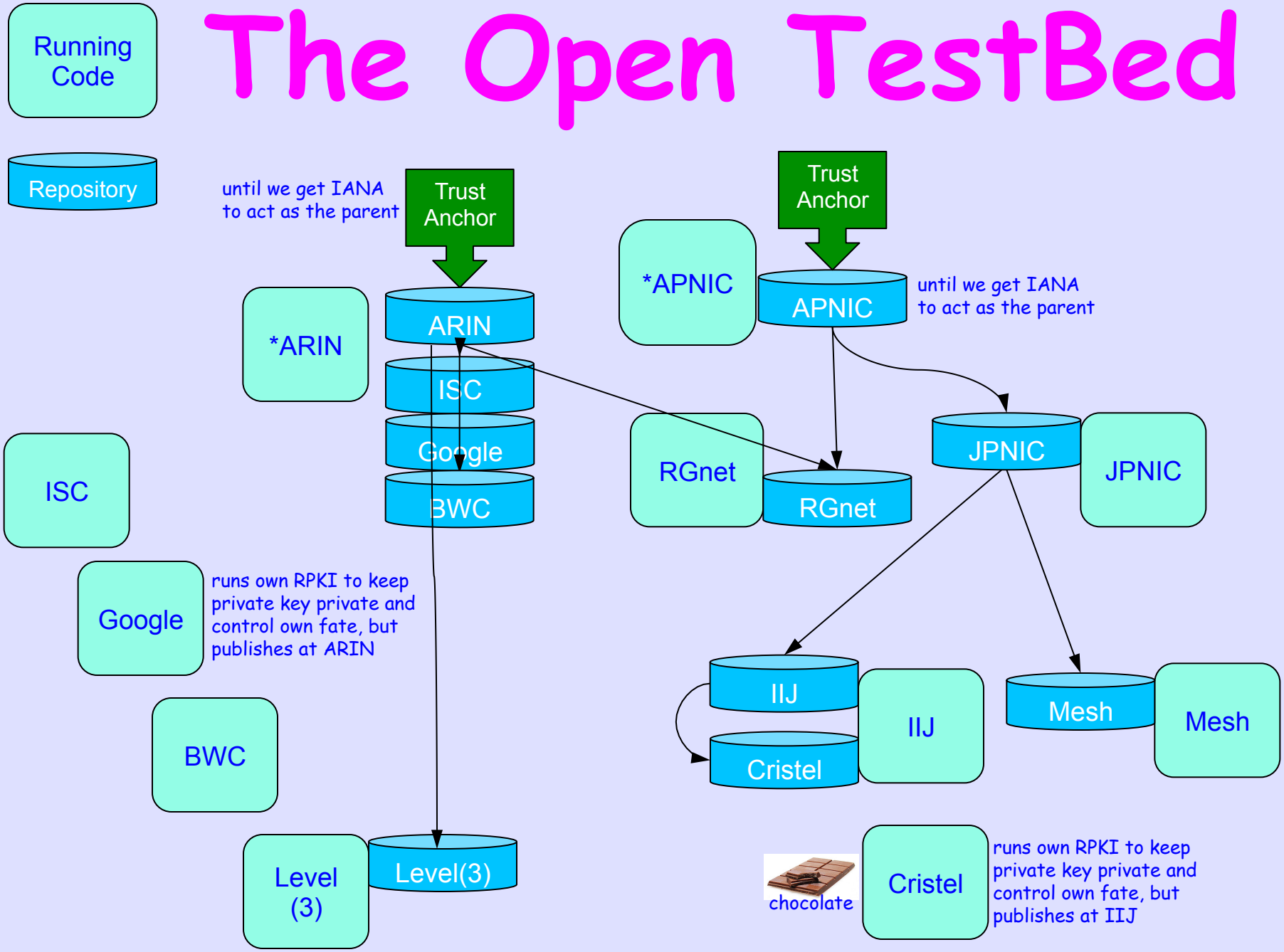
Paranoid

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

After AS-Path

```
route-map validity-0
 match rpki not-found
   set metric 50
route-map validity-1
   match rpki invalid
   set metric 25
route-map validity-2
   set metric 100
```

The Open TestBed



* APNIC and ARIN are simulations constructed from public data

The Big Speedbump

But Who Do We Trust?

Two digital certificates have been mistakenly issued in Microsoft's name that could be used by virus writers to fool people into running harmful programs, the software giant warned Thursday.

According to Microsoft, someone posing as a Microsoft employee tricked VeriSign, which hands out so-called digital signatures, into issuing the two certificates in the software giant's name on Jan. 30 and Jan. 31.

FAQ: Microsoft's security breach and how it affects you
story

Such certificates are critical for businesses and consumers who download patches, updates and other pieces of software from the Internet, because they verify that the software is being supplied from a particular company, such as Microsoft.

http://news.cnet.com/2100-1001-254586.html

Open Source (BSD Lisc) Running Code

https://rpki.net/

Test Code in Routers

Talk to C & J

Work Supported By

- **US Government**

  THIS PROJECT IS SPONSORED BY THE DEPARTMENT OF HOMELAND SECURITY UNDER AN INTERAGENCY AGREEMENT WITH THE AIR FORCE RESEARCH LABORATORY (AFRL). **[0]**

  [0] – they take your scissors away and we turn them into plowshares

- **ARIN**
- **Internet Initiative Japan**
- **Cisco, Juniper, Google, NTT, Equinix**