The RPKI, Origin Validation, & BGPsec

6th Slo IPv6 Summit / Ljubljana

2011.11.08

Randy Bush <randy@psg.com>

Rob Austein <sra@isc.org>

Steve Bellovin <smb@cs.columbia.edu>

And a cast of thousands! Well, dozens :)

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)

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, ...

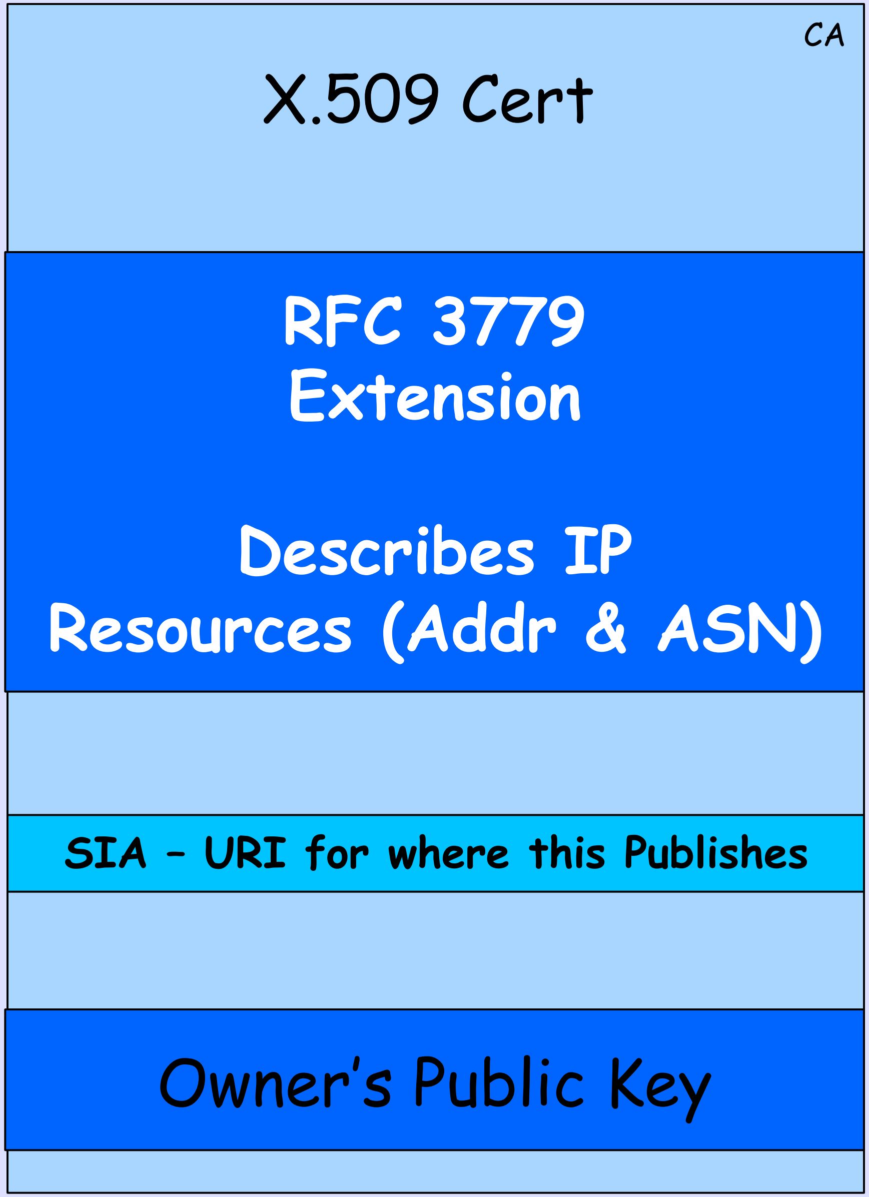
Why Origin Validation?

- 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 third step, a few years away

Resource
Public
Key
Infrastructure
(RPKI)

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

X.509 Certificate w/ 3779 Ext



CA

X.509 Cert

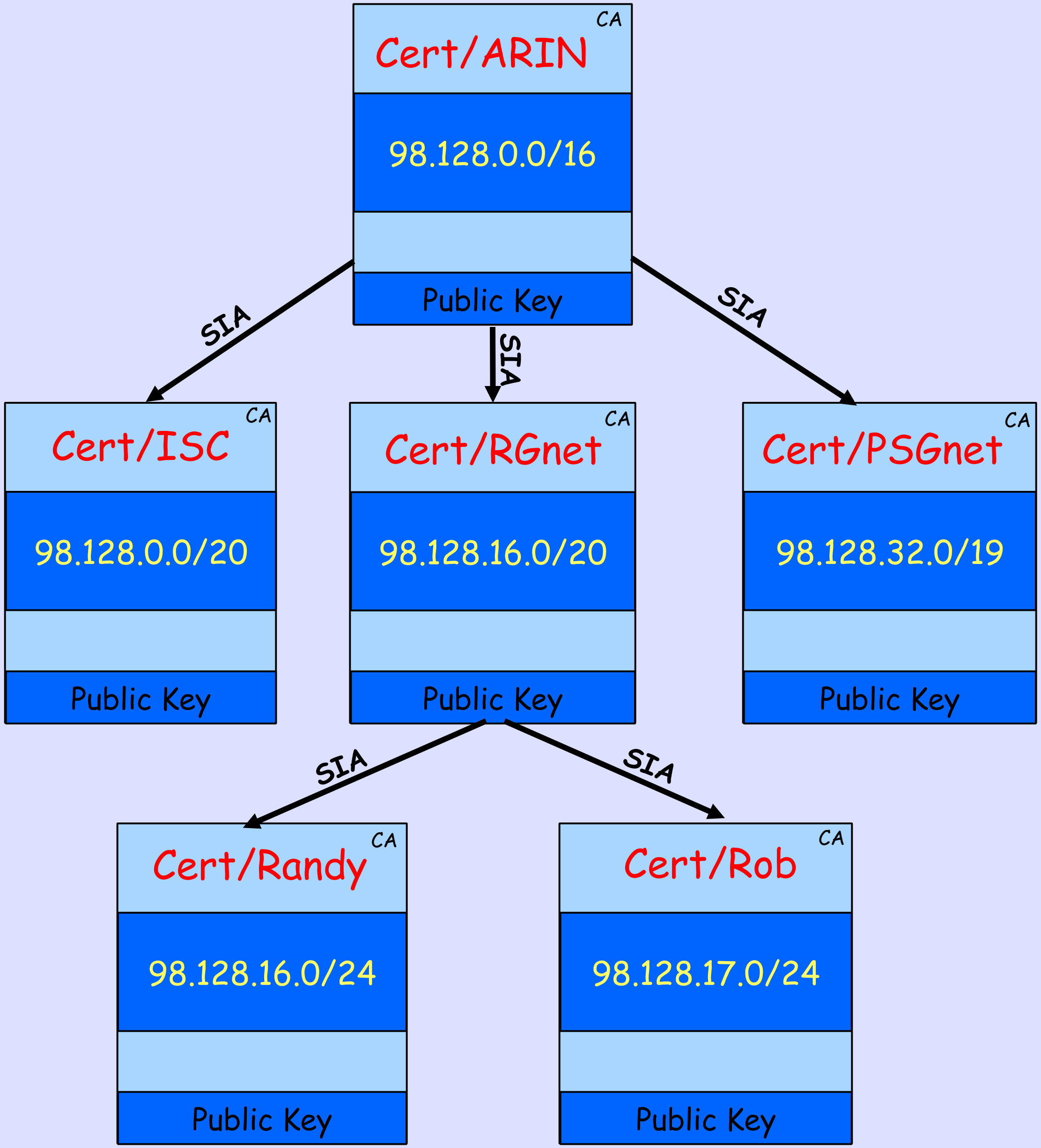
RFC 3779
Extension

Describes IP
Resources (Addr & ASN)

SIA – URI for where this Publishes

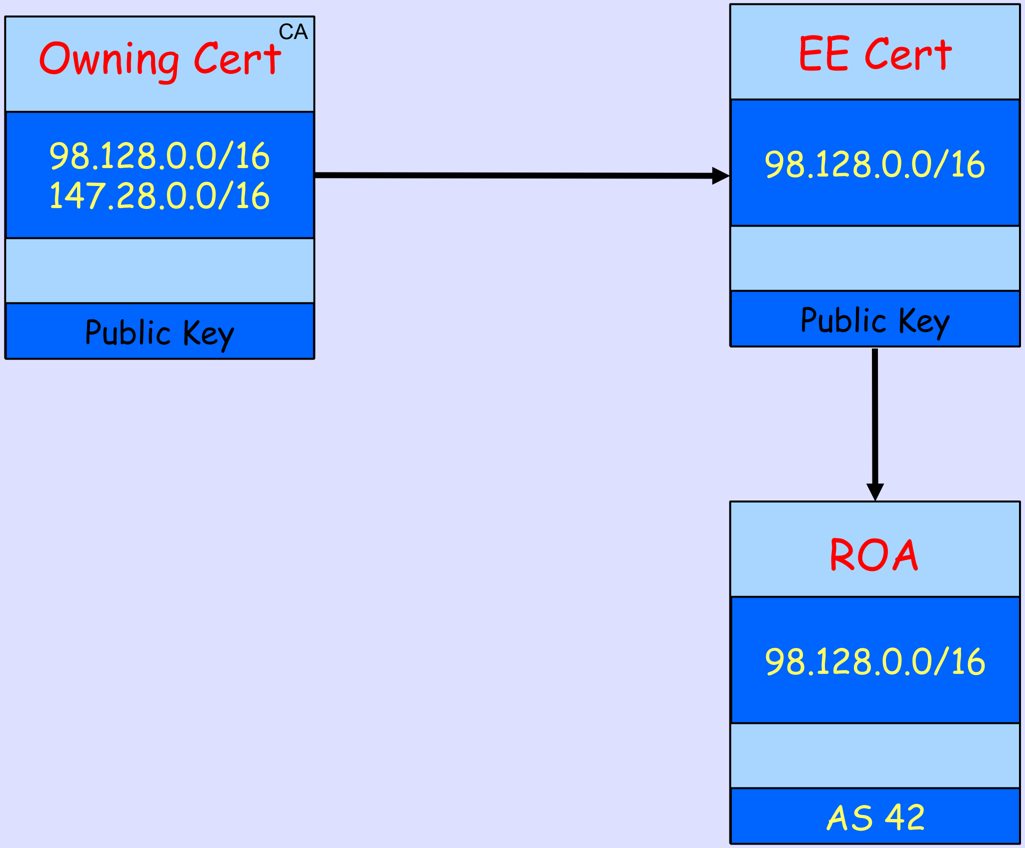
Owner's Public Key

Certificate Hierarchy follows Allocation Hierarchy



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

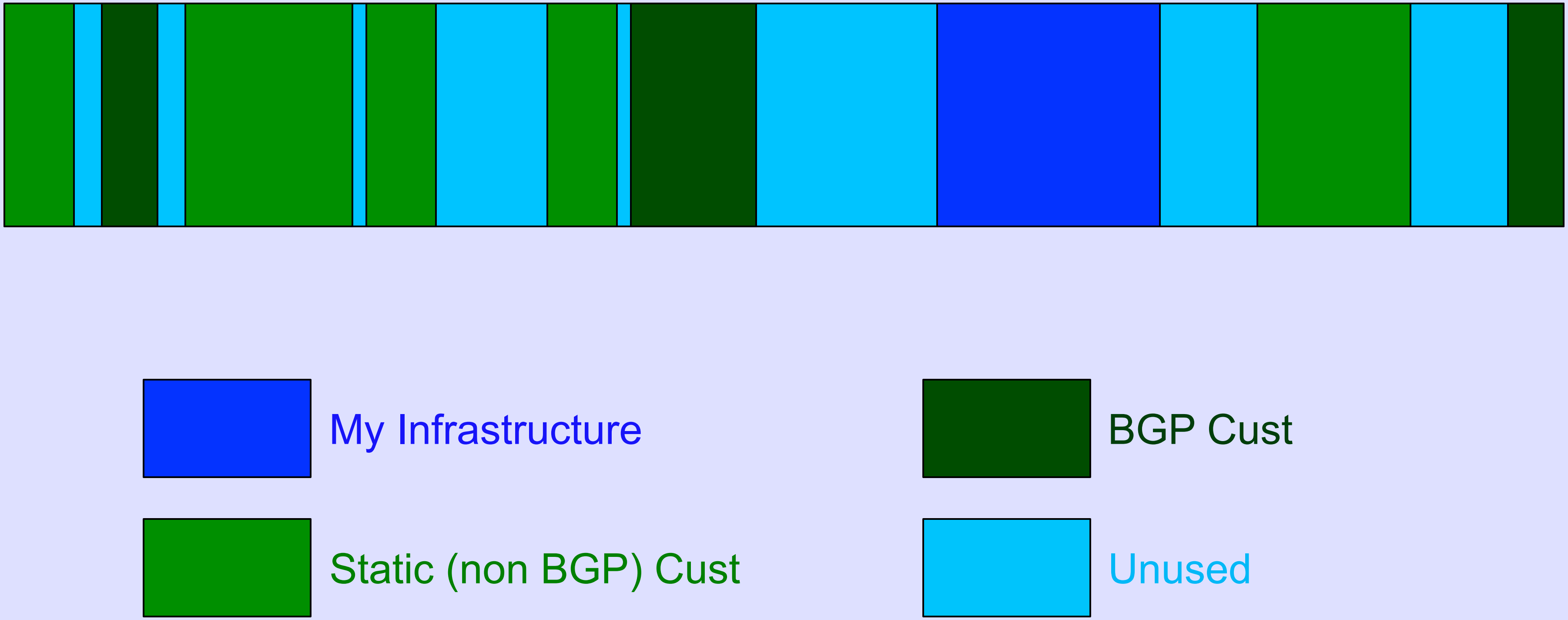
Route Origin Authorization (ROA)



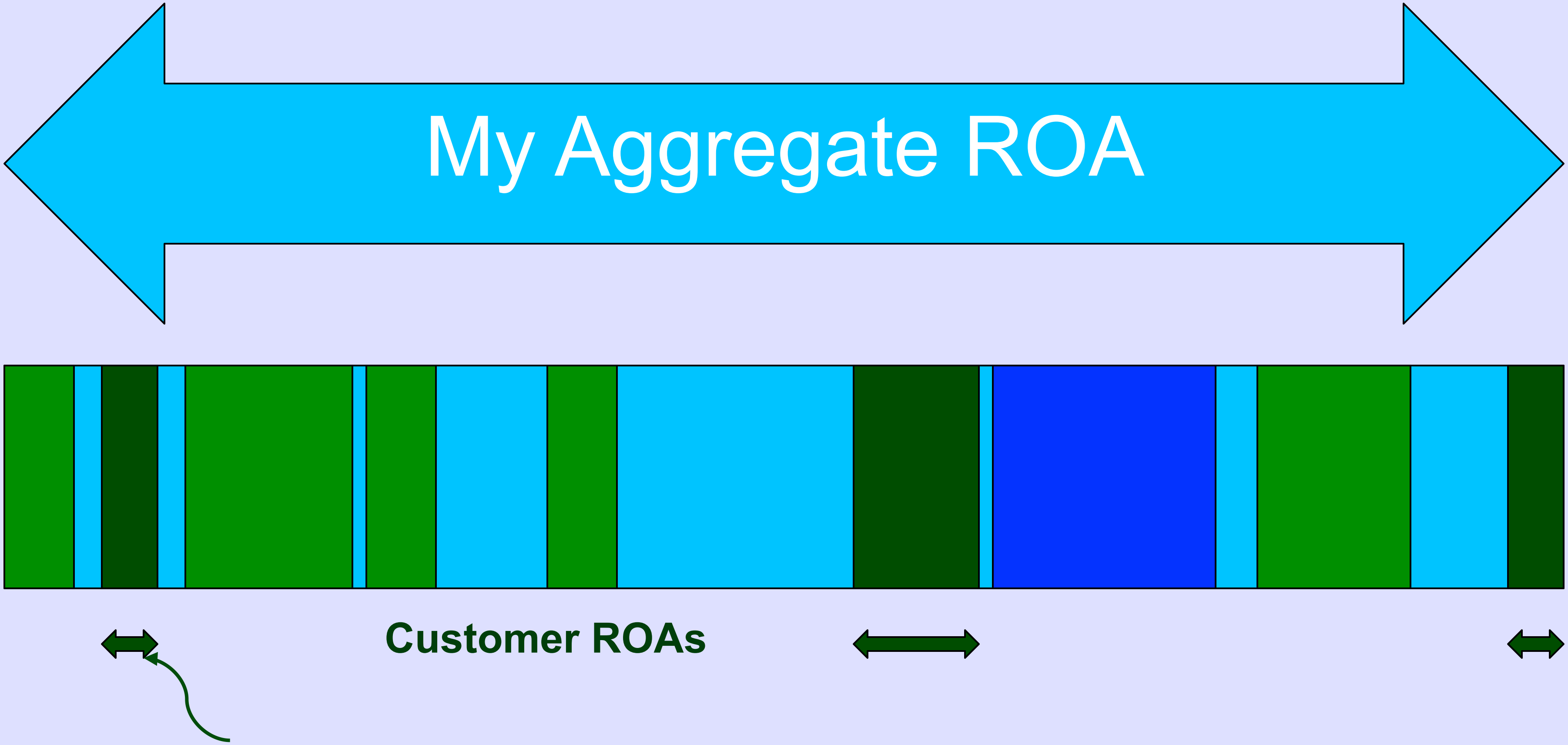
End Entity Cert
can not sign certs.
can sign other things
e.g. ROAs

This is not a Cert
It is a signed blob

Allocation in Reality



ROA Use



I Generate for
'Lazy' Customer

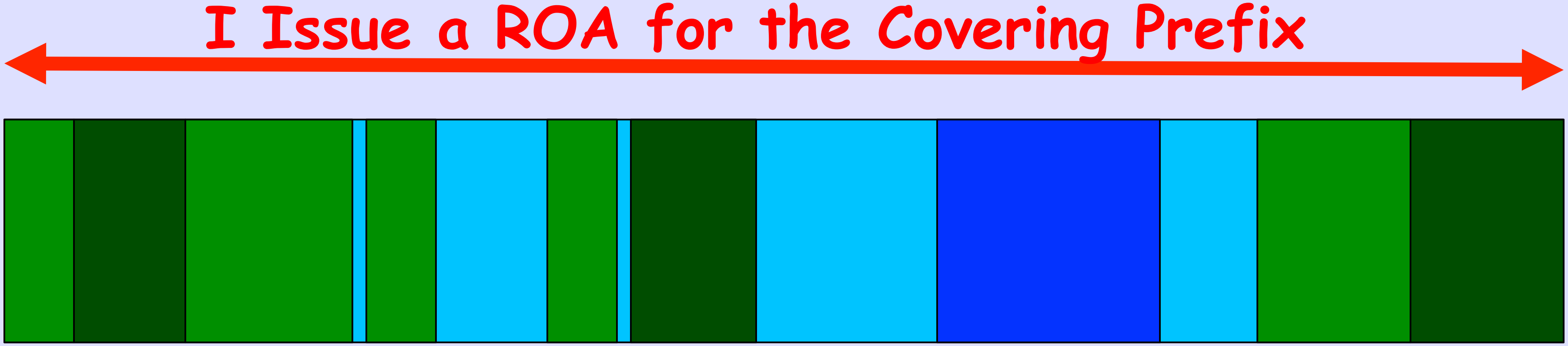
🟦 My Infrastructure

🟩 BGP Cust

🟩 Static (non BGP) Cust

🟦 Unused

Covering a Customer

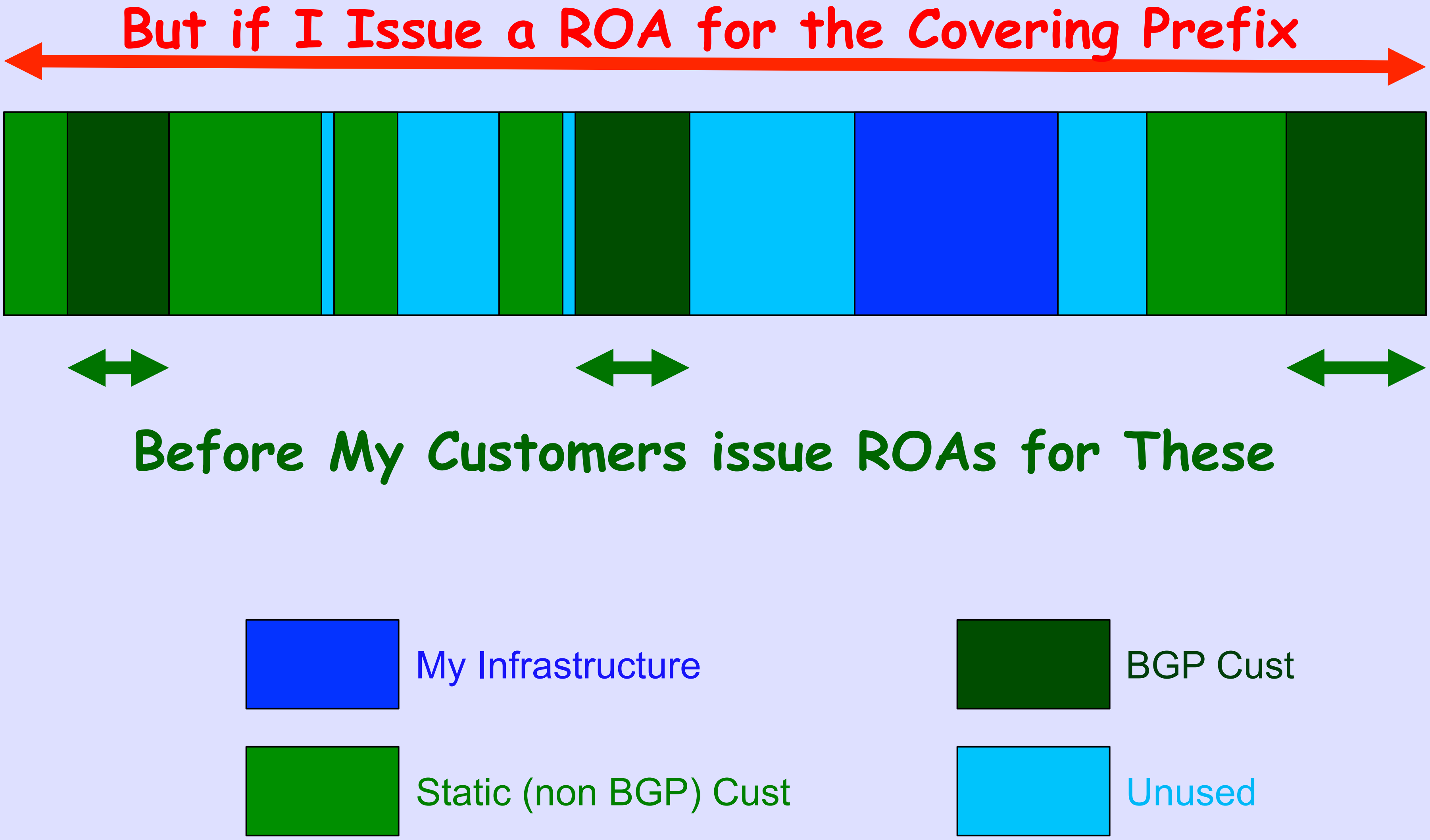


I Issue a ROA for the Covering Prefix

I need to do this to protect
Static Customers and my Infrastructure

🟦	My Infrastructure	⬛	BGP Cust
🟩	Static (non BGP) Cust	🟦	Unused

Covering a Customer

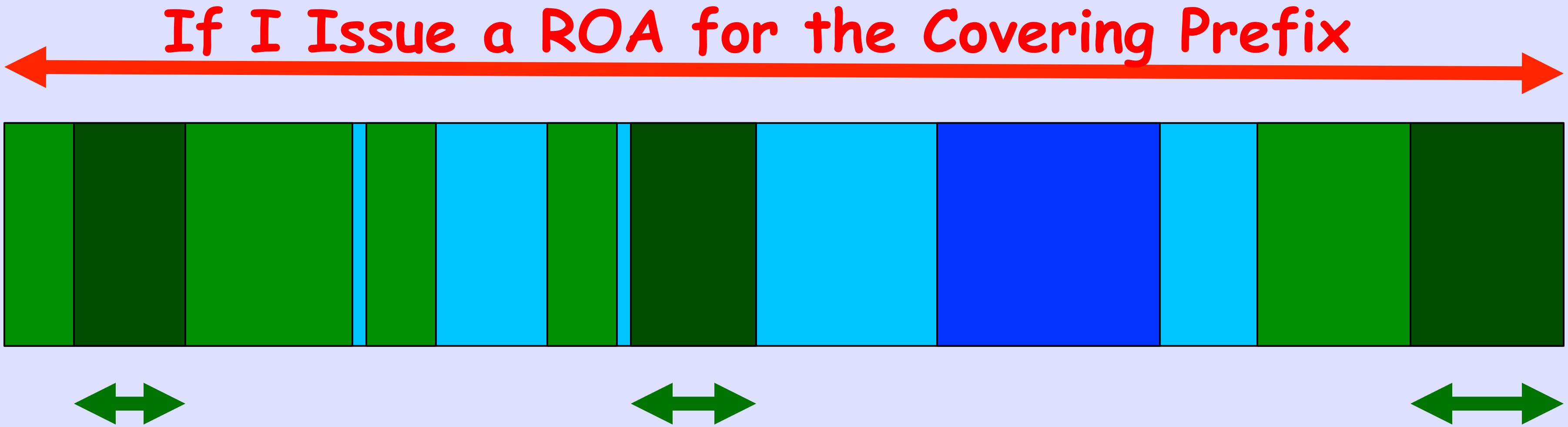


But if I Issue a ROA for the Covering Prefix

Before My Customers issue ROAs for These

🟦	My Infrastructure	■	BGP Cust
🟩	Static (non BGP) Cust	🟦	Unused

Covering a Customer



If I Issue a ROA for the Covering Prefix

Before My Customers issue ROAs for These

Their Routing Becomes Invalid!

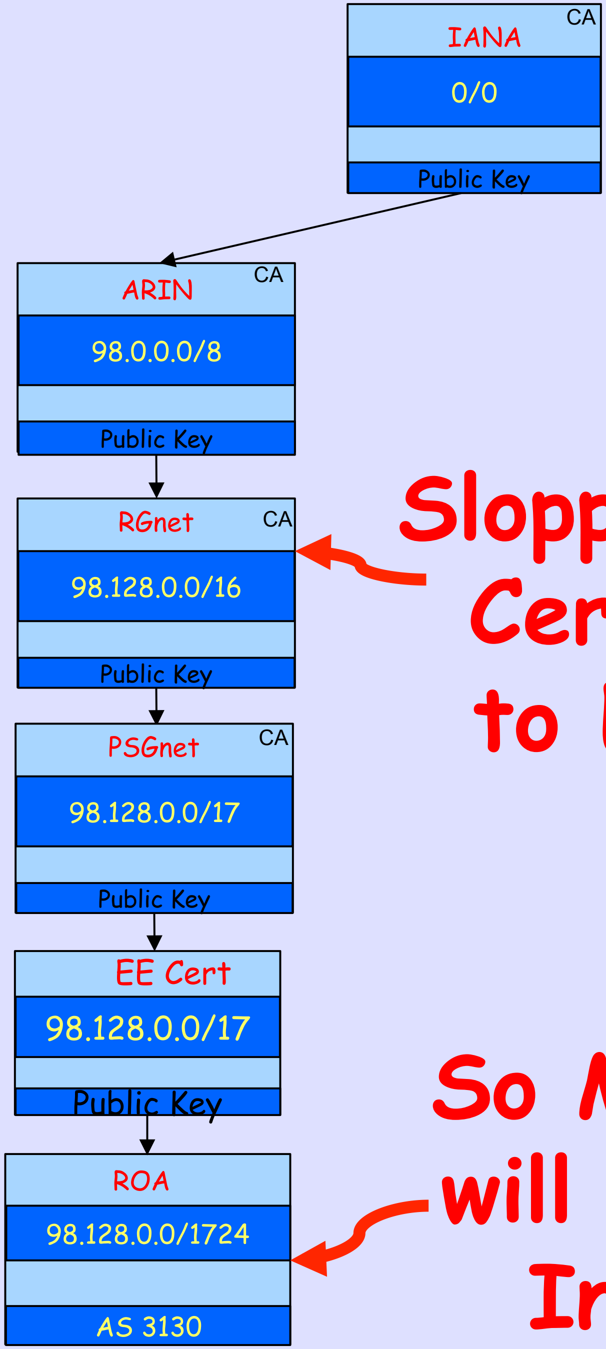
🟦 My Infrastructure

🟩 BGP Cust

🟩 Static (non BGP) Cust

🟦 Unused

Up-Chain Expiration



Sloppy Admin Cert Soon to Expire!

These are not Identity Certs

So Who You Gonna Call?

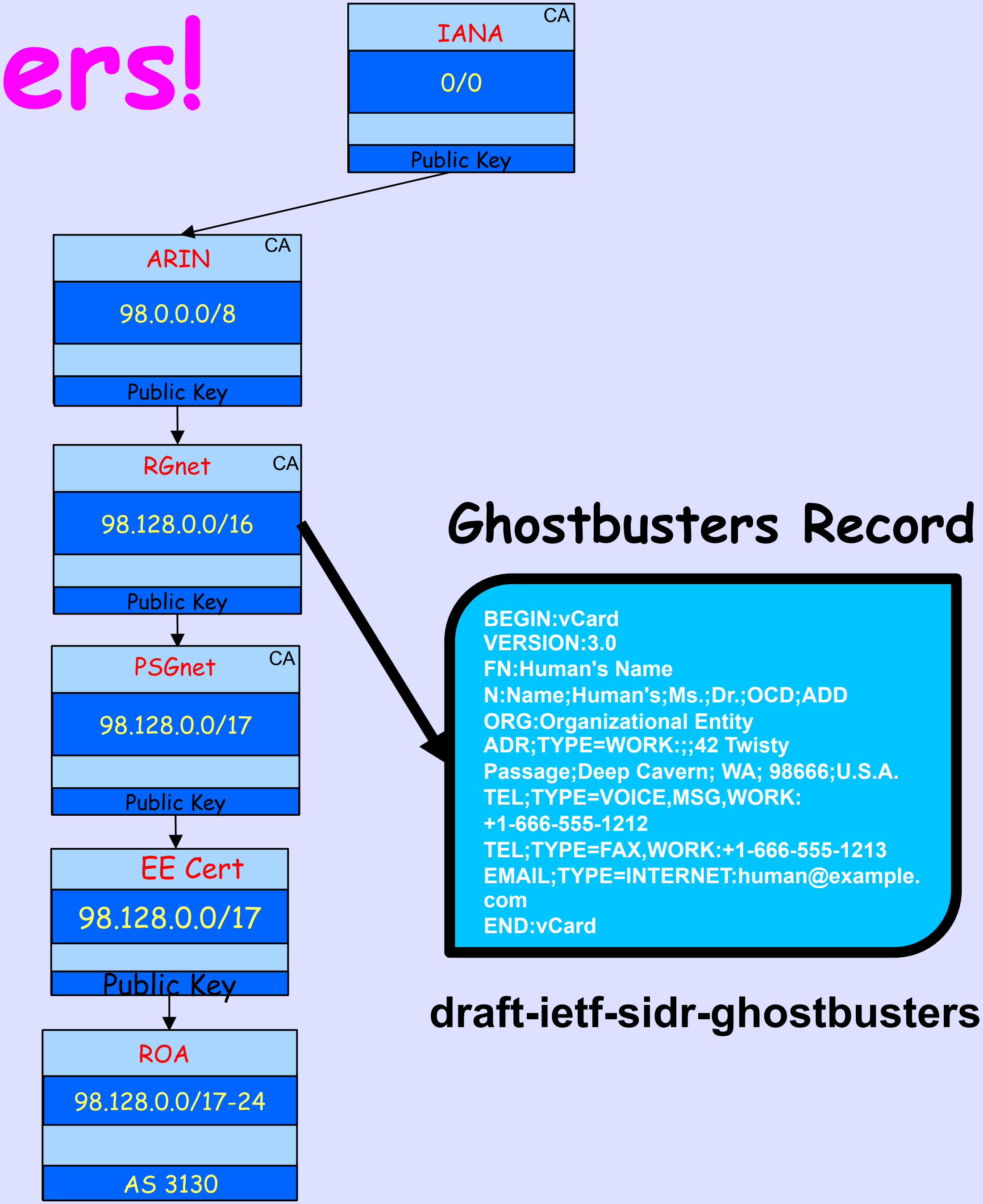
So My ROA will become Invalid!

ROA Invalid but I Can Route

- The ROA will become Invalid
- My announcement will just become NotFound, not Invalid
- Unless my upstream has a ROA for the covering prefix, which is likely

So Who You Gonna Call?

Ghostbusters!



Ghostbusters Record

```
BEGIN:vCard
VERSION:3.0
FN:Human's Name
N:Name;Human's;Ms.;Dr.;OCD;ADD
ORG:Organizational Entity
ADR;TYPE=WORK:;;42 Twisty
Passage;Deep Cavern; WA; 98666;U.S.A.
TEL;TYPE=VOICE,MSG,WORK:
+1-666-555-1212
TEL;TYPE=FAX,WORK:+1-666-555-1213
EMAIL;TYPE=INTERNET:human@example.
com
END:vCard
```

draft-ietf-sidr-ghostbusters

But in the End, You Control Your Policy

"Announcements with Invalid origins SHOULD NOT be used, but MAY be used to meet special operational needs."

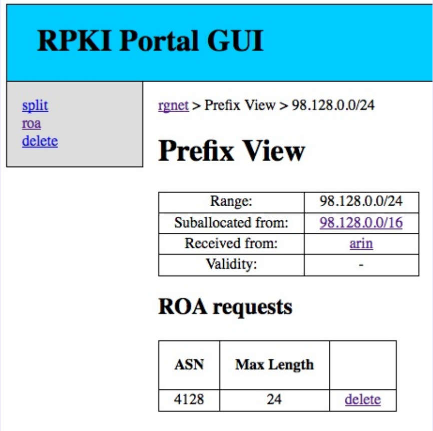
-- draft-ietf-sidr-origin-ops

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

RPKI-Based

Origin Validation

And the Three RPKI Protocols



RPKI Portal GUI

split
roa
delete

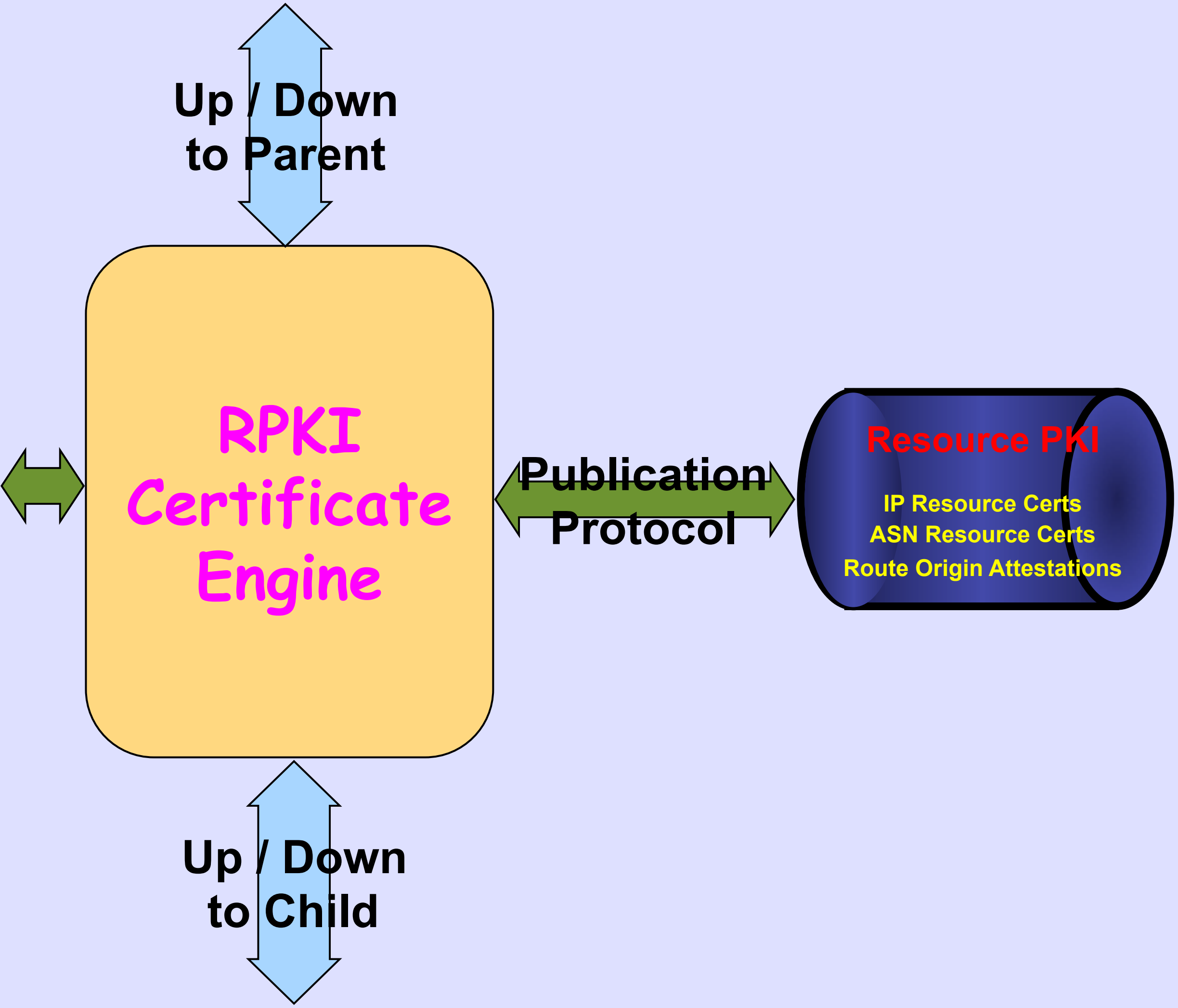
rgnet > Prefix View > 98.128.0.0/24

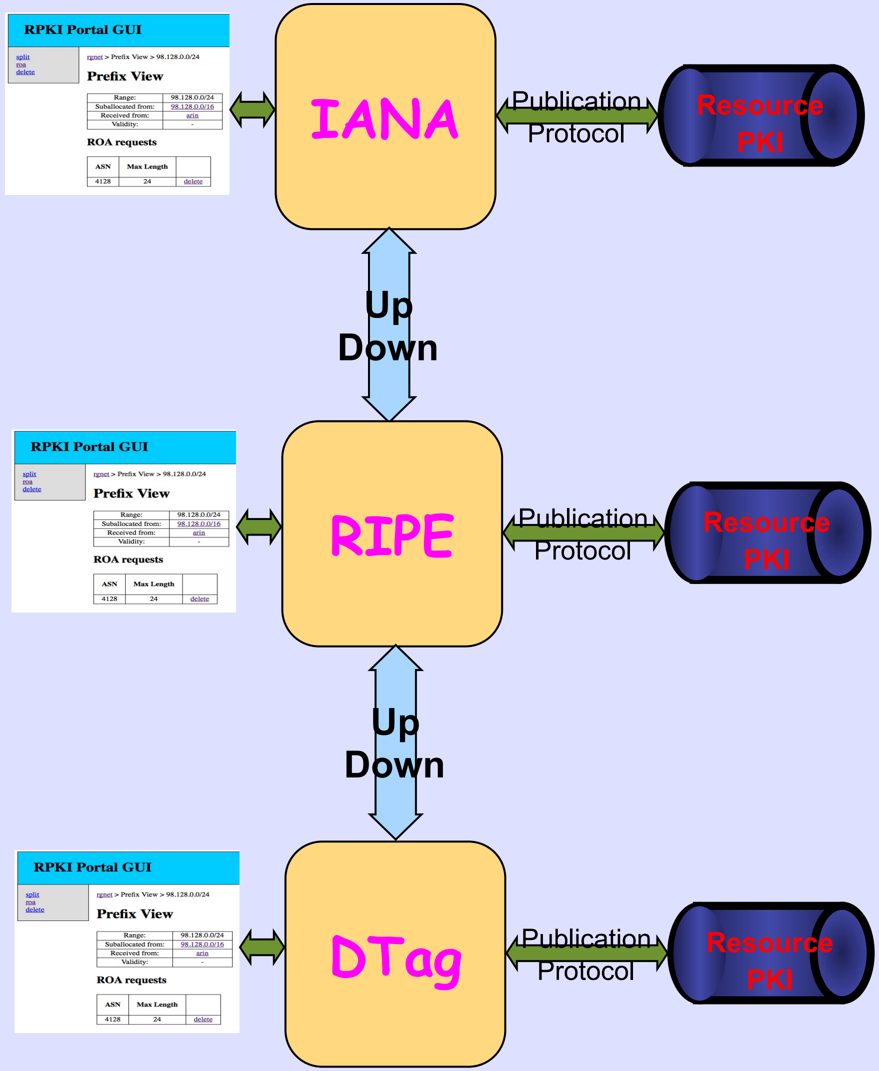
Prefix View

Range:	98.128.0.0/24
Suballocated from:	98.128.0.0/16
Received from:	arin
Validity:	-

ROA requests

ASN	Max Length	
4128	24	delete





RPKI Portal GUI

split
roa
delete

rgnet > Prefix View > 98.128.0.0/24

Prefix View

Range:	98.128.0.0/24
Suballocated from:	98.128.0.0/16
Received from:	arin
Validity:	-

ROA requests

ASN	Max Length	
4128	24	delete

RPKI Portal GUI

split
roa
delete

rgnet > Prefix View > 98.128.0.0/24

Prefix View

Range:	98.128.0.0/24
Suballocated from:	98.128.0.0/16
Received from:	arin
Validity:	-

ROA requests

ASN	Max Length	
4128	24	delete

RPKI Portal GUI

split
roa
delete

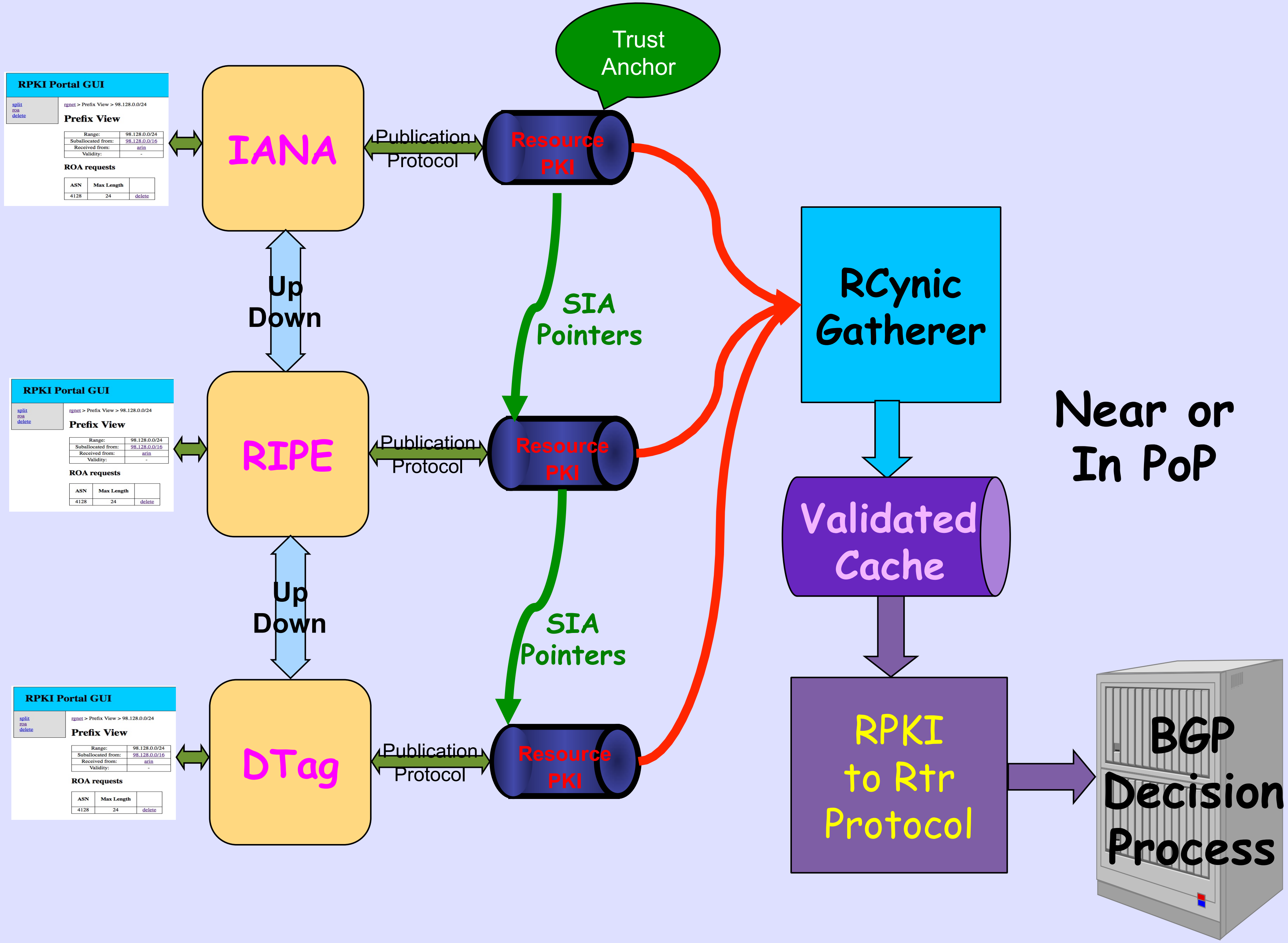
rgnet > Prefix View > 98.128.0.0/24

Prefix View

Range:	98.128.0.0/24
Suballocated from:	98.128.0.0/16
Received from:	arin
Validity:	-

ROA requests

ASN	Max Length	
4128	24	delete



Near or In PoP

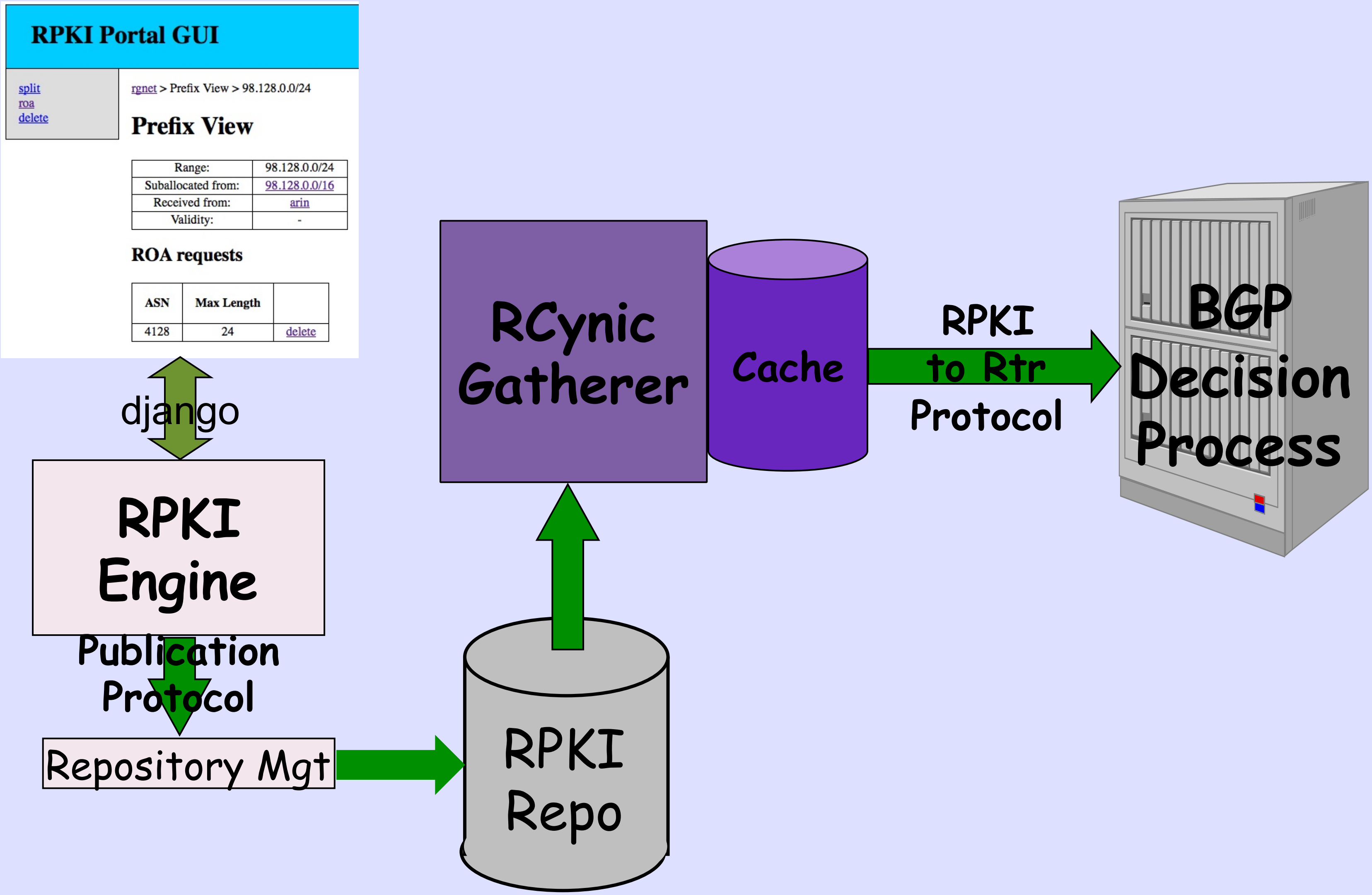
RPSL Your WorkFLow?

```
route:   147.28.0.0/16
descr:   147.28.0.0/16-16
origin:  AS3130
notify:  irr-hack@rpki.net
mnt-by:  MAINT-RPKI
changed: irr-hack@rpki.net 20110606
source:  RPKI
```

CSV Your WorkFlow?

67.21.36.0/24	3970
192.169.0.0/23	3970
207.34.0.0/24	3970
216.21.0.0/24	3970
216.21.14.0/24	3970
216.21.16.0/24	3970
216.151.34.0/24	3970
147.28.0.0/16	3130
192.83.230.0/24	3130

RPKI-Rtr Protocol



RPKI Portal GUI

split
roa
delete

rgnet > Prefix View > 98.128.0.0/24

Prefix View

Range:	98.128.0.0/24
Suballocated from:	98.128.0.0/16
Received from:	arin
Validity:	-

ROA requests

ASN	Max Length	
4128	24	delete

Typical Exchange

```
Cache                         Router
    | <----- Reset Query --------- | R requests data
    |                              |
    | ----- Cache Response ----->  | C confirms request
    | ------- IPvX Prefix ------>  | C sends zero or more
    | ------- IPvX Prefix ------>  |   IPv4 and IPv6 Prefix
    | ------- IPvX Prefix ------>  |   Payload PDUs
    | ------  End of Data ------>  | C sends End of Data
    |                              |   and sends new serial
    ~                              ~
    | -------- Notify ---------->  |  (optional)
    |                              |
    | <----- Serial Query -------- | R requests data
    |                              |
    | ----- Cache Response ----->  | C confirms request
    | ------- IPvX Prefix ------>  | C sends zero or more
    | ------- IPvX Prefix ------>  |   IPv4 and IPv6 Prefix
    | ------- IPvX Prefix ------>  |   Payload PDUs
    | ------  End of Data ------>  | C sends End of Data
    |                              |   and sends new serial
    ~                              ~
```

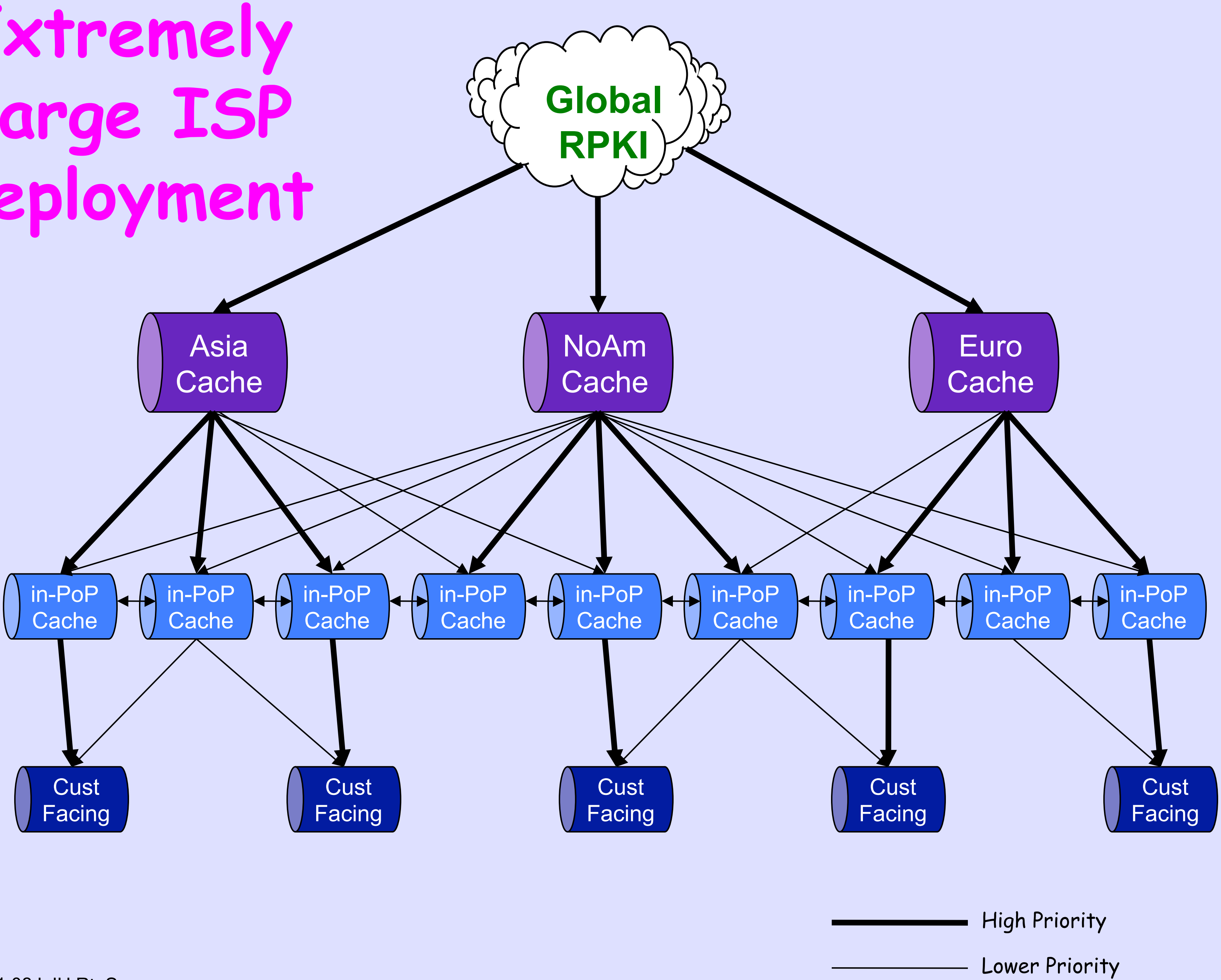
IPv4 Prefix

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

IPv6 Prefix

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

Extremely Large ISP Deployment



Origin Validation

- Cisco IOS and IOS-XR test code have Origin Validation now, ship 1Q2012
- Juniper has test code now, ship 1Q2012
- Work continues daily in test routers
- Compute load much less than ACLs from IRR data, **10µsec per update!**

Configure

```
router bgp 3130
 …
 bgp rpki server tcp 198.180.150.1 port 42420 refresh 3600
 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

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
```

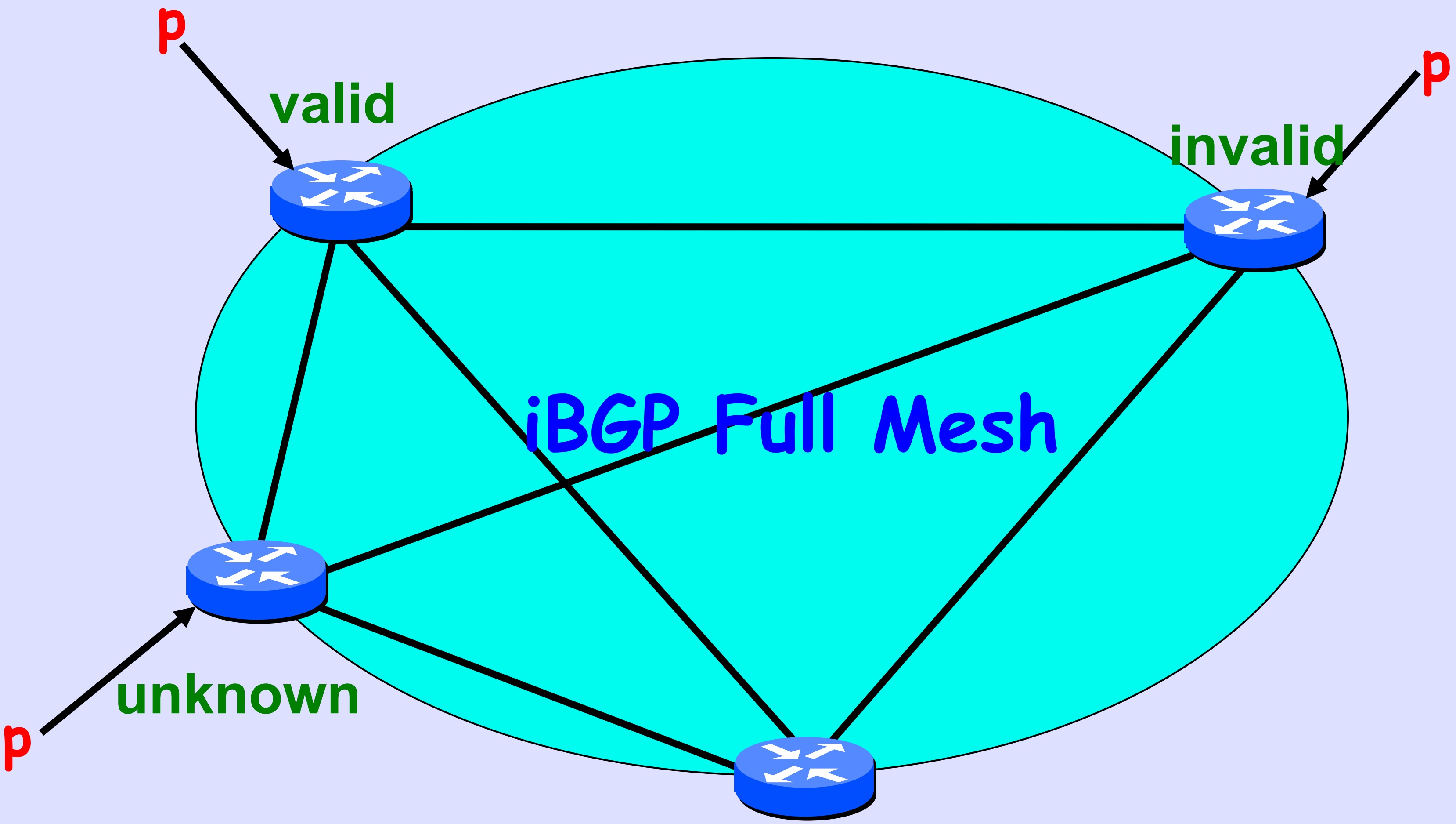
Bad Dog!

```
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 100
route-map validity-1
  match rpki invalid
  set metric 150
route-map validity-2
  set metric 50
```

Open Source (BSD Lisc) Running Code

https://rpki.net/

Test Code in Routers

Talk to C & J

BGPsec AS-Path Validation

Future Work

Origin Validation is Weak

- RPKI-Based Origin Validation only stops accidental misconfiguration, which is very useful. But ...
- A malicious router may announce as any AS, i.e. forge the ROAed origin AS.
- This would pass ROA Validation as in draft-ietf-sidr-pfx-validate.

Full Path Validation

- Rigorous per-prefix AS path validation is the goal
- Protect against origin forgery and AS-Path monkey in the middle attacks
- Not merely showing that a received AS path is not impossible

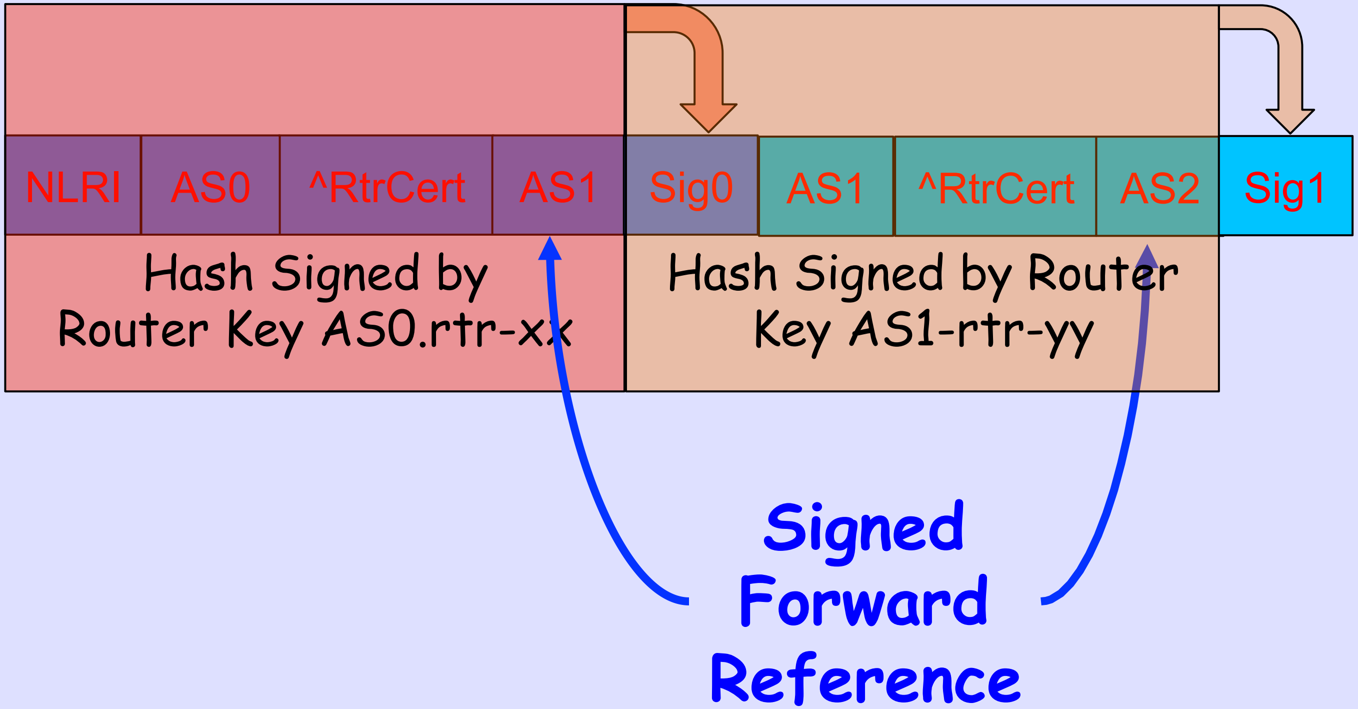
Protocol Not Policy

- We can not know intent, **should** Mary have announced the prefix to Bob
- But Joe can formally validate that Mary **did** announce the prefix to Bob
- Policy on the global Internet changes every 36ms, new peers, new customers, new circuits, etc.
- We already have a protocol to distribute policy or its effects, it is called BGP
- BGPsec validates that the protocol has not been violated, and is not about intent or business policy

Forward Path Signing

AS hop N signing (among other things) that it is sending the announcement to AS hop N+1 by AS number, is believed to be fundamental to protecting against monkey in the middle attacks

Forward Path Signing


NLRI
AS0
^RtrCert
AS1
Sig0
AS1
^RtrCert
AS2
Sig1
Hash Signed by Router Key AS0.rtr-xx
Hash Signed by Router Key AS1-rtr-yy
Signed Forward Reference

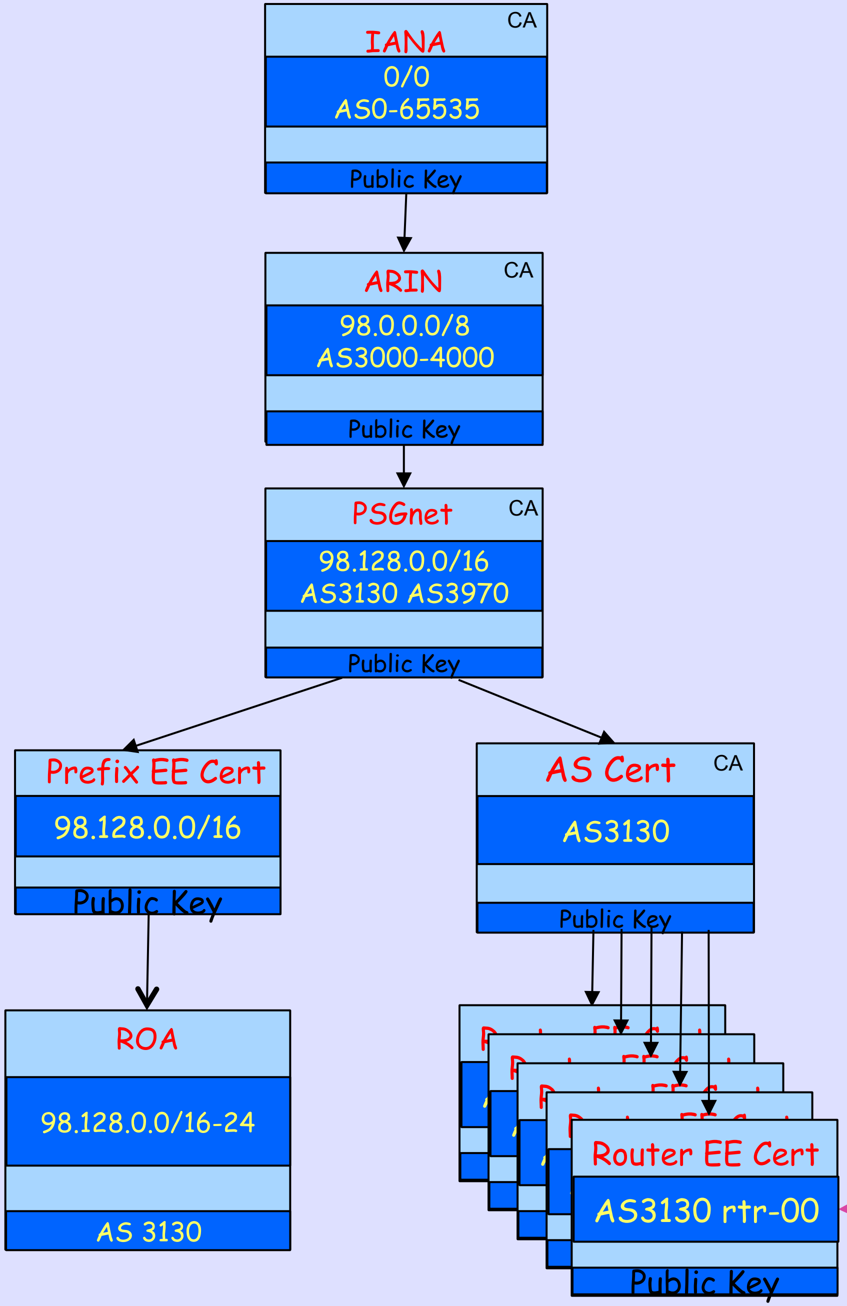
Capability Negotiation

- It is assumed that consenting routers will use BGP capability exchange to agree to run BGPsec between them
- The capability will, among other things remove the 4096 PDU limit for updates
- If BGPsec capability is not agreed, then only traditional BGP data are sent

Per-Router Keys

- Needed to deal with compromise of one router exposing an AS's private key
- Implies a more complex certificate and key distribution mechanism
- A router could generate key pair and send certificate request to RPKI for signing
- Certificate, or reference to it, must be in each signed path element
- If you want one per-AS key, share a router key

Cert / Key Structure for an ISP

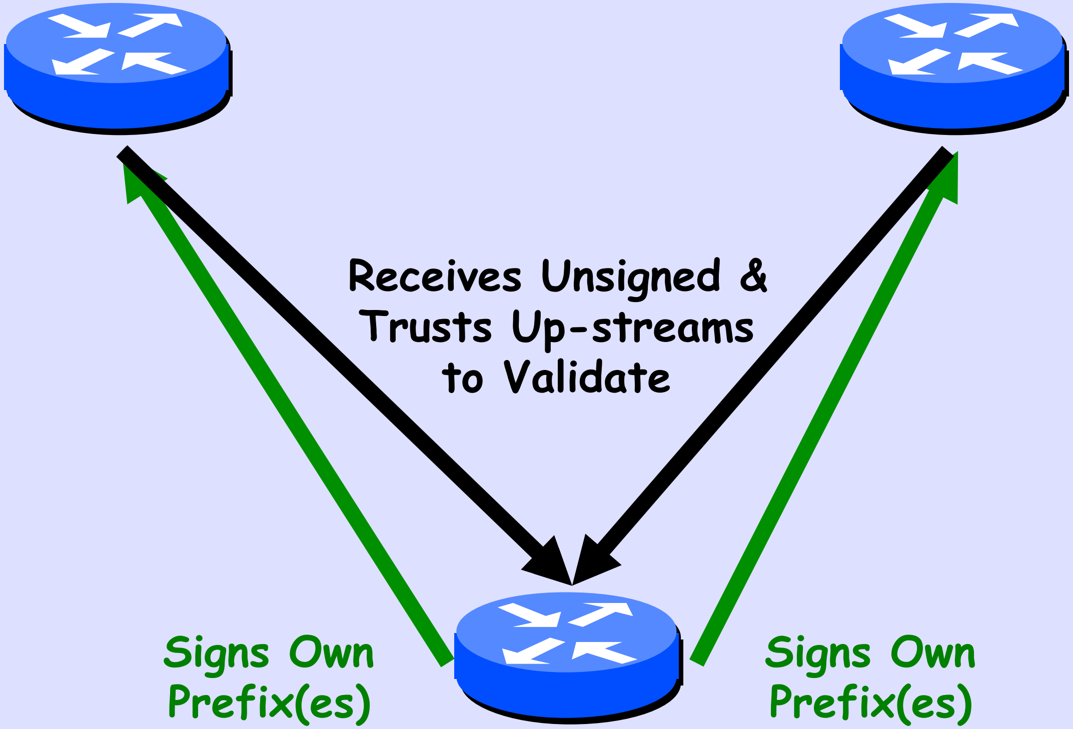


Encodes ASN and Router ID

Only at Provider Edges

- This design protects only inter-domain routing, not IGPs, not even iBGP
- BGPsec will be used inter-provider, only at the providers' edges
- Of course, the provider's iBGP will have to carry the BGPsec information
- Providers and inter-provider peerings might be heterogeneous

Simplex End Site



Receives Unsigned & Trusts Up-streams to Validate

Signs Own Prefix(es)

Signs Own Prefix(es)

Only Needs to Have Own Private Key, No Other Crypto or RPKI Data

No Hardware Upgrade!!

Incremental Deployment

Meant to be incrementally deployable in today's Internet, and does not require global deployment, a flag day, etc.

Incremental Deployment will form Islands

No Increase of Operator Data Exposure

Operators wish to minimize any increase in visibility of information about peering and customer relationships etc.

No IRR-style publication of customer or peering relationships is needed

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 & ISC

- Cisco, Juniper, Google, NTT, Equinix