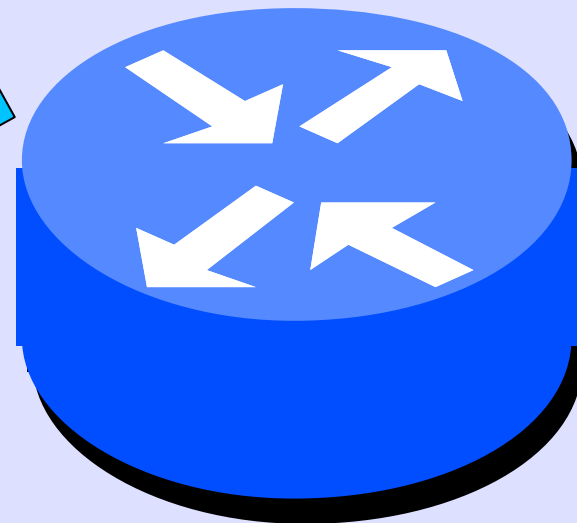# Protect
# Routing Protocols

# Remember This One?

- **Routing was Designed With no Concern for Security**

- **Attacks can be Close or Remote, e.g. YouTube Incident**

**Routing Protocol Attacks**
IS-IS
OSPF
BGP
LDP

Control Plane

- **IS-IS a bit Less Vulnerable as it is not Over IP, it is CLNP**

- **Use MD5 Auth for Authenticity**

- **Other Protections Very Active in IETF**

# What is Routing Security?

- **Defending routers against attacks that are similar to attacks on hosts**

- **But the unique threat is attackers using routing protocols**

  - **To divert traffic**

  - **To alter traffic**

- **We have some ability to lessen the danger, but not enough!**

# Protocol Attacks

- **The Router is Secured Against Attack**

- **The Routing Code is Good**

- **The Attacker is 'Gaming' the Protocol**

- **Sending Lies Over BGP is the Big Threat**

- **But IGP, OSPF or IS-IS may also be Attacked**

# History of Routing Security

- **Radia Perlman dissertation:** *Network Layer Protocols with Byzantine Robustness*, 1988
- **Bellovin:** *Security Problems in the TCP/IP Protocol Suite*, 1989
- **Work Begins in 1996**
- **Kent et alia two papers in 2000**
- **Endless Talk in the IETF**
- **2005 Serious Work Outside IETF**
- **2010-12 RFCs Published, Code by C & J**

# Why so Little Progress

- **The Problems are Technically Very Difficult**

- **Simple Routing is Already a Very Complex Operational Issue**

- **It is Not Traditional Communications Security**

- **Installed Base & Transition Problem**

# Normal Ops Security

- **Protect Router Itself, Like a Host**
- **TCP/MD5 Session Protection**
- **ACLs on Everything**
- **ssh, not telnet.  no http, …**
- **Route Filtering (based on IRR),**
- **…**

# A Quick Commercial

# Why I Prefer IS-IS

- **Simpler than OSPF**

- **Used in ISPs, not many End Sites**

- **Over CLNP, the Link Layer, not IP, so Harder to Attack**

- **IPv4 and IPv6 in One Protocol**

- **Older and Less Buggy**

- **Biggest ISPs Used it so Well Debugged**

# IS-IS

- **Is Over the Link Layer**
- **So Attacker Must be On The Link**
- **Of Course, a Compromised Neighbor Router is On The Link**
- **OSPF, BGP, and LDP may be Attacked from a Long Distance**

# But it Makes No Difference

- **Use Either IS-IS or OSPF**

- **But Protect Them**

- **Use MD5 Auth/Password to be Sure you have Connected to Right Peer**

- **Use MD5 Auth to Protect from on-the-wire Attack**

- **Maybe even ACL Filter Who May Exchange IGP with Each Router**
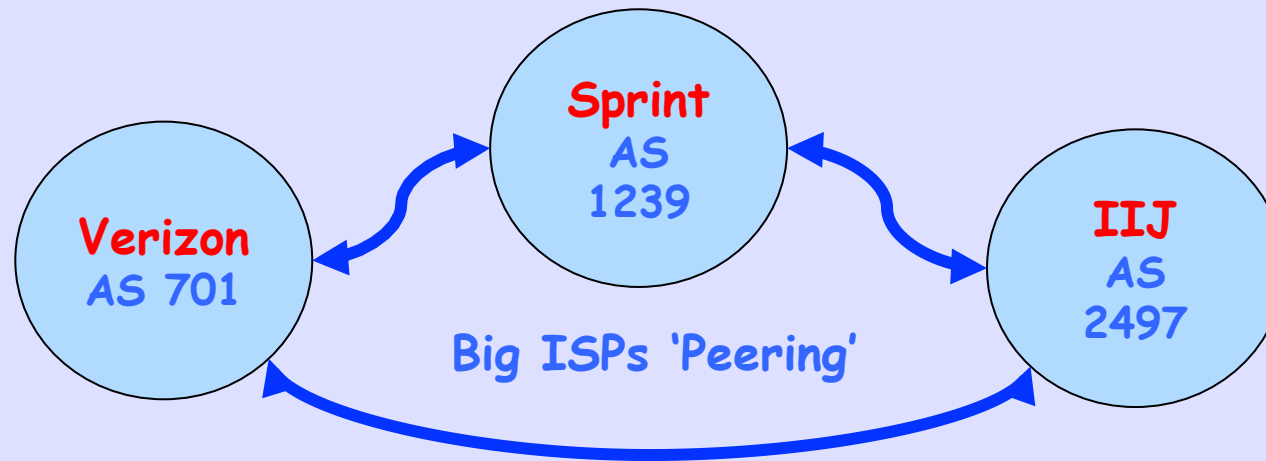
# BGP
# is the
# Big Vulnerability

# Basic Protection Same as IGP

- **Use MD5 Auth/Password to be Sure you have Connected to Right Peer**

- **Use MD5 Auth to Protect from on-the-wire Attack**

- **Maybe even ACL Filter Who May Exchange BGP with Each Router**

# What is an AS?
# An ISP or End Site



Sprint
AS
1239

Verizon
AS 701

IIJ
AS
2497

Big ISPs 'Peering'

# What is an AS?
## An ISP or End Site

Sprint
AS
1239

Verizon
AS 701

IIJ
AS
2497

Big ISPs 'Peering'

Amazon
AS
16509

Customers buy 'Transit'

GoJ
AS 234

Sakura
AS
9370

# An IP Prefix is Announced & Propagated

**Sprint**
AS 1239

**Verizon**
AS 701

**IIJ**
AS 2497

Big ISPs 'Peering'

**Amazon**
AS 16509

147.28.0.0/16
**GoJ**
AS 234

**Sakura**
AS 9370

Customers buy 'Transit'

# From Inside a Router

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

**Origin AS**

**16509   1239   2497   234**

**The AS-Path**

# 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
```
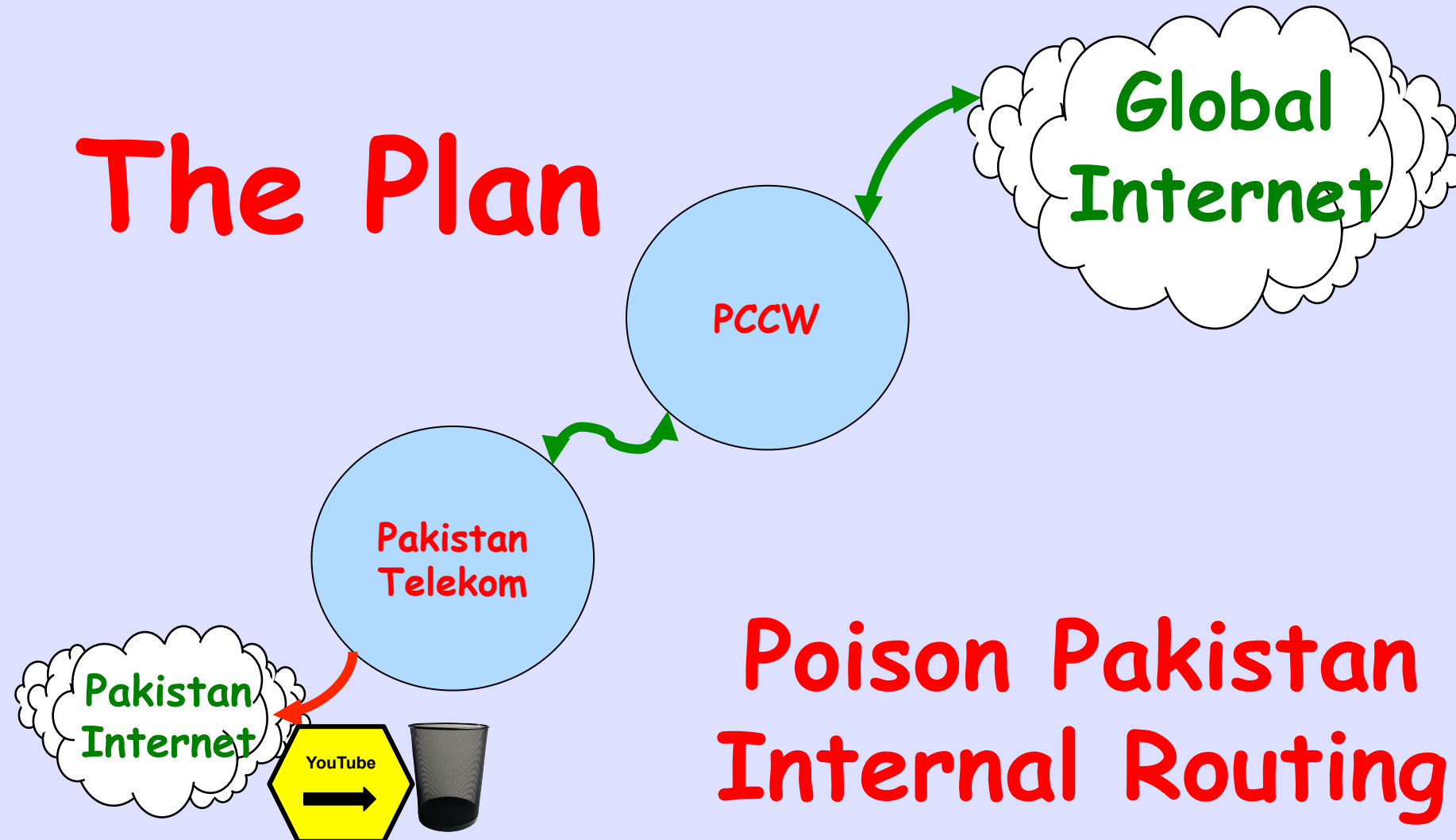
# The YouTube Incident

## The Plan



Global Internet

PCCW

Pakistan Telekom

Pakistan Internet

YouTube →

## Poison Pakistan Internal Routing

# The YouTube Incident

**What Happened**

**Global Internet**

PCCW

**YouTube**

**Oops!**

**Pakistan Telekom**

**Pakistan Internet**

**Poisoned the Global Internet**

# 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, BGPsec

# We Need to be Able to <u>Authoritatively</u> 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)

# Agenda

- **Some Technical Background**

- **Mis-Origination - YouTube Incident**

- **The RPKI – Needed Infrastructure**

- **RPKI-Based Origin Validation**

- **Use the GUI to make ROAs and look at the result on a router**

- **Build your own Relying Party Server**

- **Discussion**

# This is Not New

- 1986 – Bellovin & Perlman identify the vulnerability in DNS and Routing

- 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 – RPKI.NET(for ARIN) & APNIC start work on RPKI.  RIPE starts in 2008.

- 2009 – RPKI.NET Open Testbed and running code in test routers

# What is an AS?
# An ISP or End Site

**Sprint**
**AS**
**1239**

**Verizon**
**AS 701**

**IIJ**
**AS**
**2497**

**Big ISPs 'Peering'**

# What is an AS?
## An ISP or End Site



Sprint
AS 1239

Verizon
AS 701

IIJ
AS 2497

Big ISPs 'Peering'

Amazon
AS 16509

Customers buy 'Transit'

GoJ
AS 234

Sakura
AS 9370

# An IP Prefix is Announced & Propagated



Sprint
AS 1239

Verizon
AS 701

IIJ
AS 2497

Big ISPs 'Peering'

Amazon
AS 16509

147.28.0.0/16
GoJ
AS 234

Sakura
AS 9370

Customers buy 'Transit'

# From Inside a Router

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

**Origin AS**

**16509    1239    2497    234**

**The AS-Path**

# 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

**Global Internet**

**PCCW**

**Pakistan Telekom**

**Pakistan Internet**

YouTube

**Poison Pakistan Internal Routing**

# The YouTube Incident

**What Happened**

**Poisoned the Global Internet**

Global Internet

PCCW

YouTube

Oops!

Pakistan Telekom

Pakistan Internet

# 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 (deployed at all RIRs)

- **Origin Validation** – Using the RPKI to detect and prevent mis-originations of someone else's prefixes (in deployment)

- **AS-Path Validation** AKA **BGPsec** – Prevent Path Attacks on BGP (in IETF)

# 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*, 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, …

# X.509-Based IP Resource PKI

# RFCs Have Been Long Published

# Deployed by All RIRs

# ROAs Registered by > 1,000 Operators

# In Live Routers

# X.509 Certificate w/ 3779 Ext

**Signed by Parent's Private Key**

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

**Cert/ARIN** CA

98.128.0.0/16

Public Key

SIA → **Cert/ISC** CA

98.128.0.0/20

Public Key

SIA → **Cert/RGnet** CA

98.128.16.0/20

Public Key

SIA → **Cert/PSGnet** CA

98.128.32.0/19

Public Key

SIA → **Cert/Randy** CA

98.128.16.0/24

Public Key

SIA → **Cert/Rob** CA

98.128.17.0/24

Public Key

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

# Route Origin Authorization (ROA)

**Owning Cert** [CA]

98.128.0.0/16
147.28.0.0/16

Public Key

**EE Cert**

98.128.0.0/16

Public Key

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

**ROA**

98.128.0.0/16

AS 42

This is not a Cert
It is a signed blob

# How RPKI is Generated

**Up / Down to Parent**

**RPKI Certificate Engine**



**GUI**

**Publication Protocol**

**Resource PKI**

IP Resource Certs
ASN Resource Certs
Route Origin Attestations

**Up / Down to Child**

# Issuing Parties



Trust Anchor

IANA

GUI

Publication Protocol

Resource PKI

Please Issue My Cert

Up/Down

Cert Issuance

APNIC

GUI

Publication Protocol

Resource PKI

Please Issue My Cert

Up/Down

Cert Issuance

JPNIC

GUI

Publication Protocol

Resource PKI

Please Issue My Cert

Up/Down

Cert Issuance

# Issuing Parties



Trust Anchor

IANA

GUI

Publication Protocol

Resource PKI

Up Down

SIA Pointers

APNIC

GUI

Publication Protocol

Resource PKI

Up Down

SIA Pointers

JPNIC

GUI

Publication Protocol

Resource PKI

# Issuing Parties    Relying Parties

Trust Anchor

**IANA**

GUI

Publication Protocol

Resource PKI

Up Down

Publication Protocol

**APNIC**

GUI

Resource PKI

SIA Pointers

Up Down

Publication Protocol

**JPNIC**

GUI

Resource PKI

SIA Pointers

**RCynic Gatherer**

**Validated Cache**

**NOC Tools**

## Pseudo IRR

```
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
```

**BGP Decision Process**

# How Do ROAs Affect BGP Updates?

# ROAs Become Router ROAs

**Want to Run on Current Hardware**

**RPKI-Rtr Protocol**

**RPKI Cache**

**Crypto is Stripped**

**RPKI ROAs**

# 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                          |
 |                                                             |
 `-------------------------------------------------------------'
```
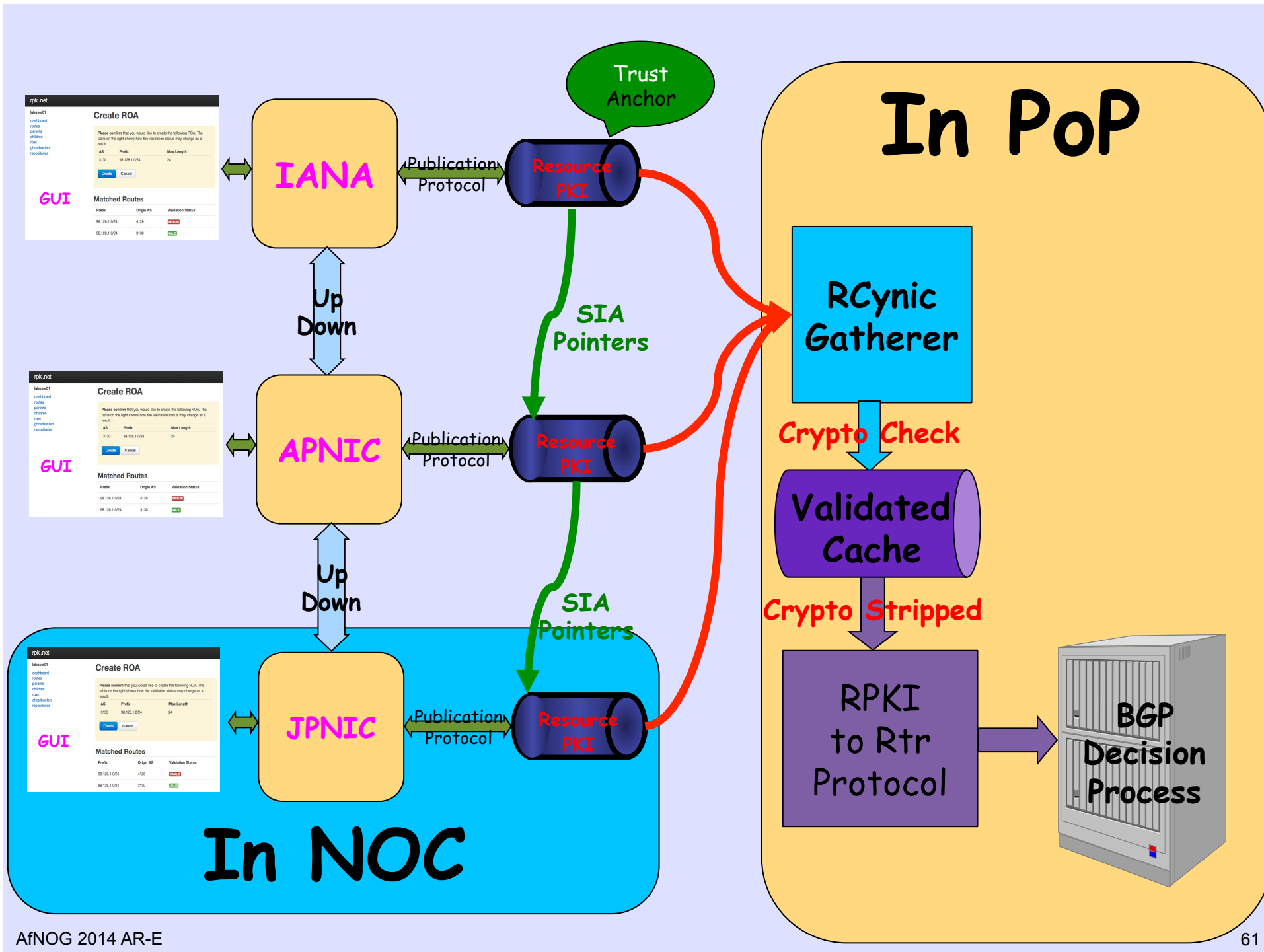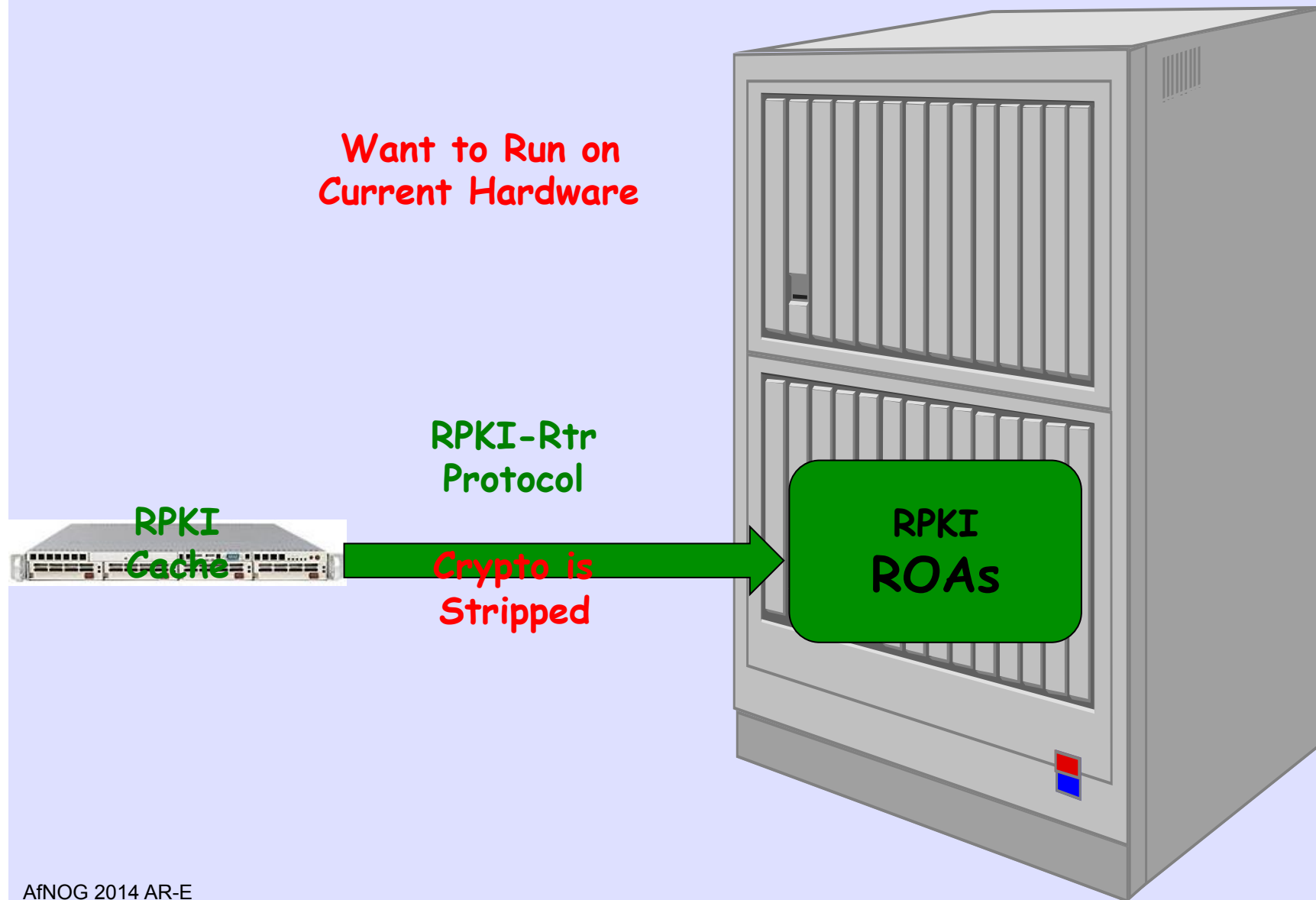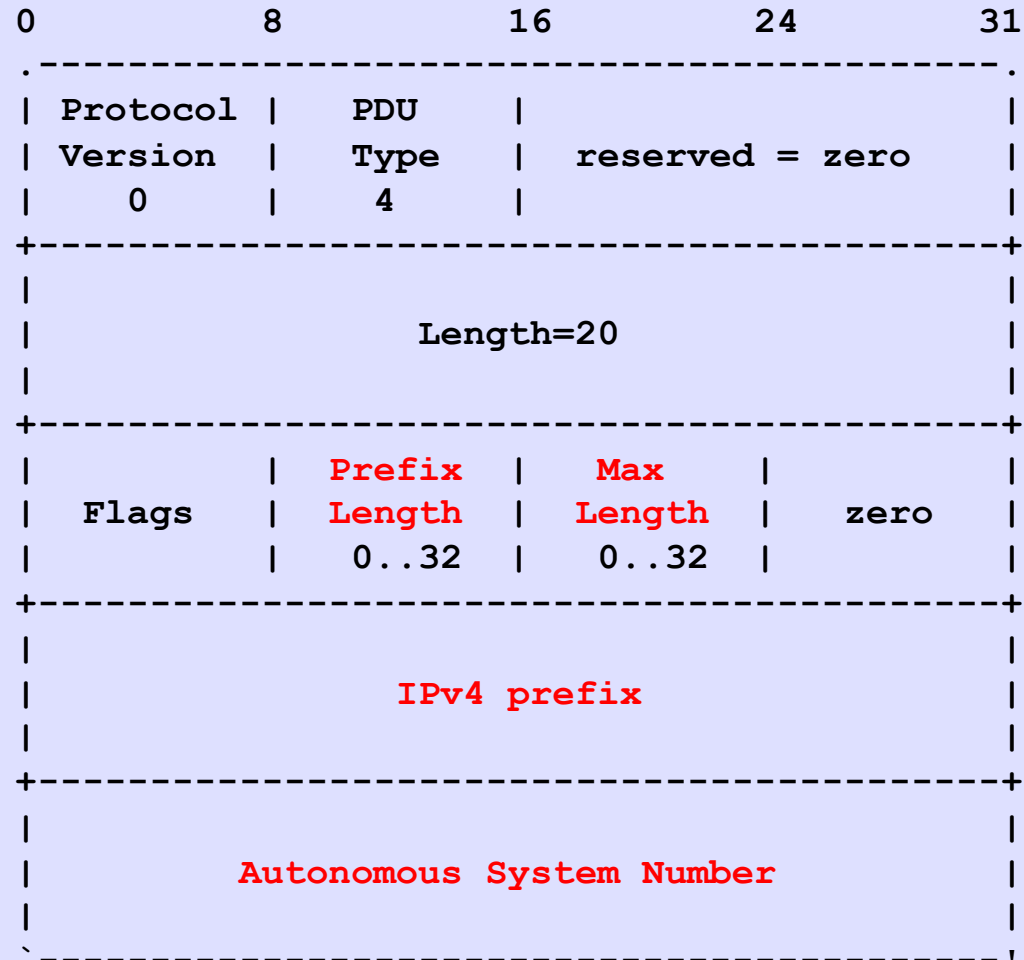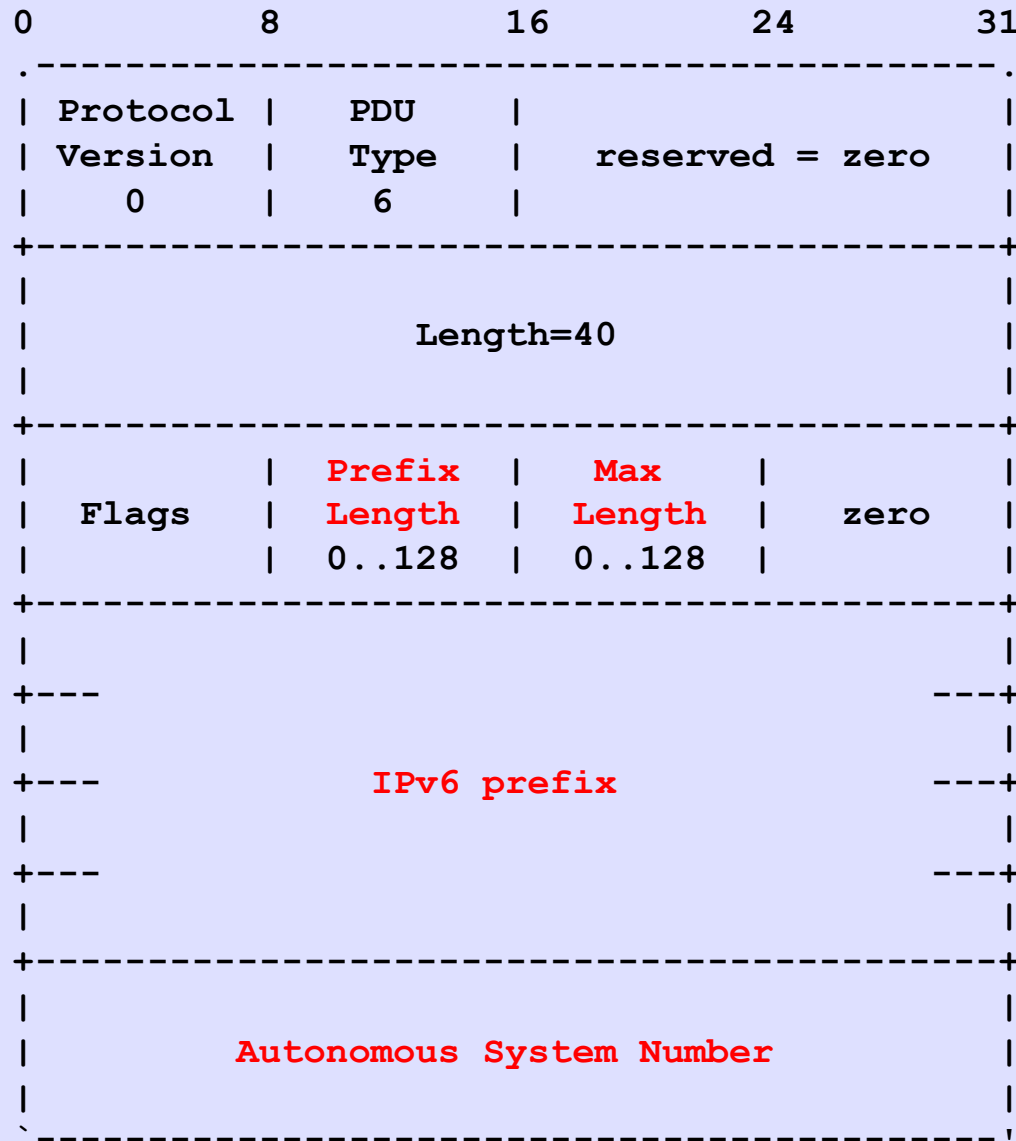
**96 More Bits
No Magic**

# Marking BGP Updates



BGP Data

BGP Peer

BGP Updates

Valid
Invalid
NotFound

Mark

RPKI-Rtr Protocol

RPKI Cache

RPKI ROAs

# Result of Check

- **Valid** – A matching/covering ROA was found with a matching AS number

- **Invalid** – A covering ROA was found, but the AS number did not match, and there was no other matching one

- **NotFound** – No matching or covering ROA was found, <u>same</u> <u>as</u> <u>today</u>

# Configure Router to Get ROAs

```
router bgp 651nn

  …

  bgp rpki server tcp 192.168.179.3 port 43779 refresh 60

  bgp rpki server tcp 147.28.0.84 port 93920 refresh 60

  …
```

# 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
```

# The Operator
# Tests the Mark
# and then
# Applies 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
```

# Security Geek

```
route-map validity-0
    match rpki invalid
    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
```

# 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
```
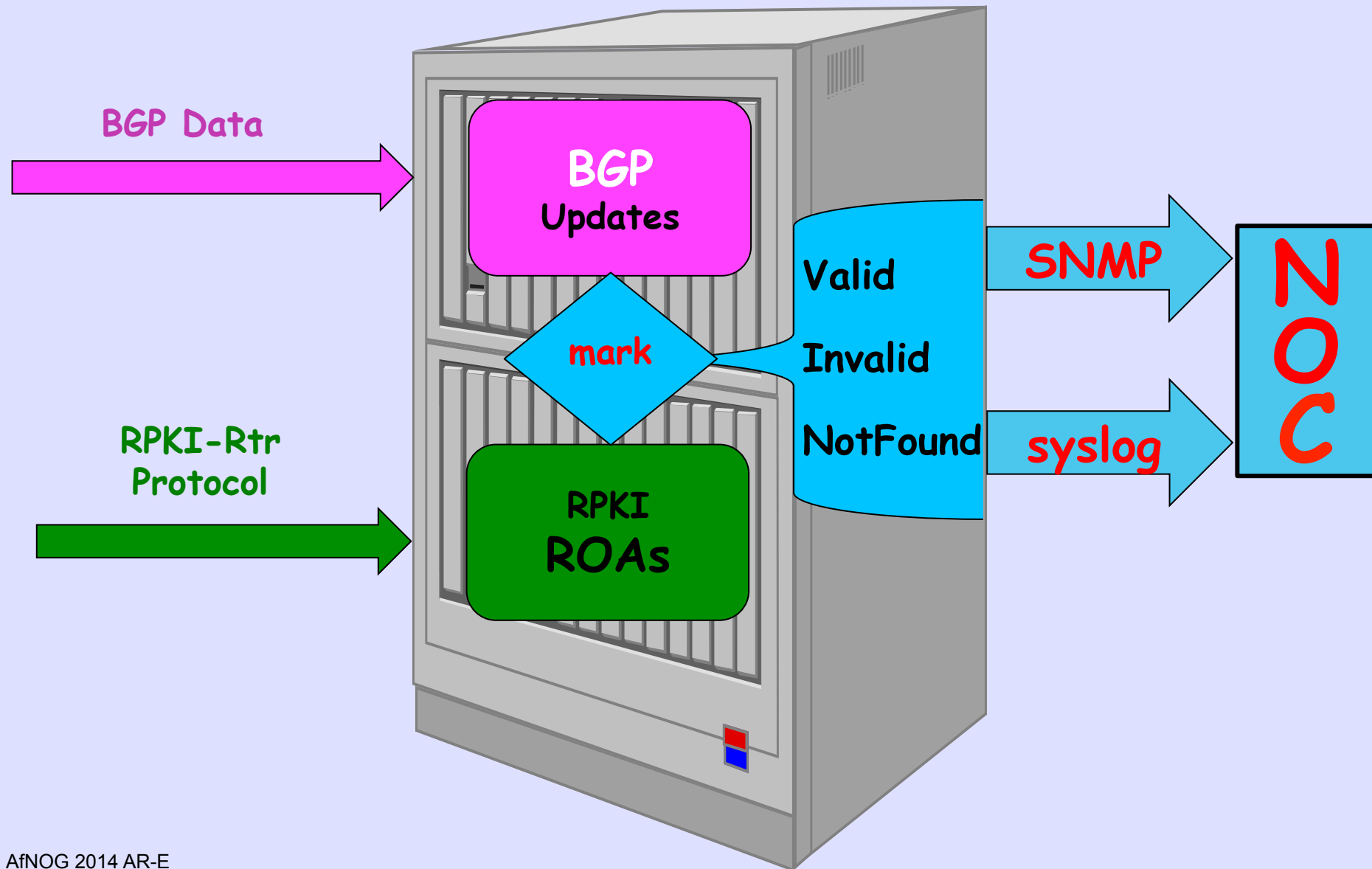
# And it is All Monitored



BGP Data

RPKI-Rtr Protocol

BGP Updates

mark

RPKI ROAs

Valid
Invalid
NotFound

SNMP

syslog

NOC

# 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.  In such circumstances, the announcement SHOULD have a lower preference than that given to Valid or NotFound."

-- draft-ietf-sidr-origin-ops

# RPKI at the Registries

- RIPE seriously deployed with a few thousand LIRs and thousands of ROAs

- APNIC is operational and moving forward, moving to RIPE's GUI

- ARIN is doing their best to make RPKI deployment very hard

- LACNIC is deployed and has 100s of LIRs

- AFRINIC is deployed with O(25) LIRs

# RIPE Progress

- Policy just passed to allow registration of legacy space without having to become a member or sign away all rights

- Policy just passed to allow registration of 40,000 PI allocations to end sites without having to become a full member

- And RIPE already had thousands of RPKI registrations

# LACNIC / Ecuador

- LACNIC working with the Ecuador Internet Exchange

- All ISPs and lmost all address space in Ecuador is certified and has ROAs

- All members of the exchange are using RPKI-Rtr protocol to get ROAs from a cache at the exchange

- Watching routers to see markings

# Per-RIR Statistics

ome    Global    Top 10    IPv4/6    Per AS    **RIR Stats**    RPKI routes    World map    Trends    Alexa Top500

Select a RIR below to view the corresponding charts:

## Breakdown per RIR

10 ▾ records per page                                    Search:

| RIR | Total | Valid | Invalid | Unknown | Accuracy | RPKI Adoption Rate |
|-----|-------|-------|---------|---------|----------|--------------------|
| AFRINIC | 11585 (100%) | 41 (0.35%) | 42 (0.36%) | 11502 (99.28%) | 49.4% | 0.72% |
| APNIC | 119646 (100%) | 167 (0.14%) | 240 (0.2%) | 119239 (99.66%) | 41.03% | 0.34% |
| ARIN | 187547 (100%) | 666 (0.36%) | 105 (0.06%) | 186776 (99.59%) | 86.38% | 0.41% |
| LACNIC | 63054 (100%) | 11361 (18.02%) | 1156 (1.83%) | 50537 (80.15%) | 90.76% | 19.85% |
| RIPE NCC | 131995 (100%) | 7709 (5.84%) | 1189 (0.9%) | 123097 (93.26%) | 86.64% | 6.74% |

Much Better Than IPv6

Half are Two LIRs

Embarrassing

# Router Origin Validation

- **Cisco IOS – solid in 15.2**

- **Cisco IOS/XR – shipped in 4.3.2**

- **Juniper – shipped in 12.2**

- **AlcaLu – in development**

# RPKI Implementations

- **RIPE/NCC – CA (partial closed) & RP (partial open)**

- **APNIC – CA only – Closed Source**

- **RTRlib/Berlin – RP only – Open Source**

- **BBN – RP Only – Open Source**

- **RPKI.NET – CA & RP – Open Source**

# RPKI.NET / Dragon Labs

- **Open Source BSD License**

- **CA – Hosted and Delegated Models, GUI**

- **RP – RPKI-RTR, NOC Tools, IRR Gen**

- **FreeBSD, Ubuntu, Debian, … Packaged (docs still catching up)**