The RPKI & Origin Validation

NANOG / Denver
2011.06.12
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 :)

2011.06.12 RPKI Origin
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

Sender

Receiver

Receiver's Public Key

Receiver's Private Key

Encryption Algorithm

Encryption Algorithm

Unencrypted message

Encrypted message

To be, or not to be, that is the question, whether tis nobler in the...
Digital Signature

**Diagram Description:**
- **Sender:**
  - Message
  - Message Digest Algorithm
  - Message Digest
  - Encryption Algorithm
  - Encrypted Message Digest

- **Receiver:**
  - Message
  - Message Digest Algorithm
  - Encrypted Message Digest
  - Encryption Algorithm
  - Message Digest
  - Comparison: equal?
    - yes: Message transmitted correctly
    - no: Error! Message has been modified!
Certificate

I, Certification Authority XYZ, do hereby certify that Borja Sotomayor is who he/she claims to be and that his/her public key is 49E51A3EF1C.

CA’s Signature
X.509 RPKI Being Developed & Deployed by IANA, RIRs, and Operators
X.509 Certificate w/ 3779 Ext

- 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)

- **Owning Cert**
  - 98.128.0.0/16
  - 147.28.0.0/16
  - Public Key

- **EE Cert**
  - 98.128.0.0/16
  - Public Key

- **ROA**
  - 98.128.0.0/16
  - AS 42

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

This is not a Cert It is a signed blob
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

2011.06.12 RPKI Origin
Running Code
And the Three RPKI Protocols
Parent and Child

ARIN RPKI Engine

Up / Down to IANA

Internal Protocol

ARIN Back End

ARIN's Resources

ISPs' Resources

Registry Back Ends

ISP RPKI Engine

Up / Down Protocol

Internal Protocol

ISP IR Back End

ISP's Resources

Children's Resources

Up / Down to Smart Customer

2011.06.12 RPKI Origin
Big, Centralized, & Scary
We Don’t Do This

RPKI Database
IP Resource Certs
ASN Resource Certs
Route Origin Attestations
A Player (CA) Publishes All Certificates Which They Generate in Their Own Unique Publication Point
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

Resources [OrgID]
IR's Database(s)
My RightsToRoute
Delegations to Custs

Mac

Front End GUI & Management

IR's Database(s)

User GUI

Up / Down Protocol

Publication Protocol

Contract Out To Google

RPKI Engine

28% of an RIR's Users
10% of an RIR's IP Space

98% of an RIR's Users
90% of an RIR's IP Space

Publication Point

RPKI Engine

ID=Me
Public RPKI Keys
Up/Down EE Public Keys
Certs Issued to DownStreams

Internal CA Data
My Misc Config Options
Issued ROAs

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

90% of an RIR's IP Space

92% of an RIR's Users
2% of an RIR's Users

GUI & Management

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

28% of an RIR's Users
90% of an RIR's IP Space

Web GUI
Origin Validation

• Cisco IOS and IOS-XR test code have Origin Validation now

• Juniper has early test code now

• Work continues daily in test routers

• Compute load much less than ACLs from IRR data, 10µsec per update!
The Third Protocol (origin validation only)
Typical Exchange

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

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

2011.06.12 RPKI Origin
## Reset Query

<table>
<thead>
<tr>
<th>Protocol</th>
<th>PDU</th>
<th>Version</th>
<th>Type</th>
<th>reserved = zero</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Length=8

---

2011.06.12 RPKI Origin
Cache Response

0    8    16   24   31

| Protocol | PDU   |                     |
| Version  | Type  | Cache Nonce        |
| 0        | 3     |                    |

Length=8
# IPv4 Prefix

<table>
<thead>
<tr>
<th>Protocol</th>
<th>PDU</th>
<th>Version</th>
<th>Type</th>
<th>reserved = zero</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

Length = 20

<table>
<thead>
<tr>
<th>Flags</th>
<th>Prefix</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Length</td>
<td>Length</td>
</tr>
<tr>
<td></td>
<td>0..32</td>
<td>0..32</td>
</tr>
</tbody>
</table>

IPv4 prefix

Autonomous System Number
End of Data

<table>
<thead>
<tr>
<th>Protocol</th>
<th>PDU</th>
<th>Version</th>
<th>Type</th>
<th>Cache Nonce</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Length=12

Serial Number
# Notify (Think DNS)

<table>
<thead>
<tr>
<th>0</th>
<th>8</th>
<th>16</th>
<th>24</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol</td>
<td>PDU</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version</td>
<td>Type</td>
<td>Cache Nonce</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Length=12

Serial Number
### Serial Query

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
<td>24</td>
<td>31</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protocol</th>
<th>PDU Type</th>
<th>Cache Nonce</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

**Length=12**

**Serial Number**
Error Response

<table>
<thead>
<tr>
<th>Protocol</th>
<th>PDU</th>
<th>Version</th>
<th>Type</th>
<th>Error Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Length

Length of Encapsulated PDU

Copy of Erroneous PDU

Length of Error Text

Arbitrary Text of Error Diagnostic Message
Extremely Large ISP Deployment

Global RPKI

Asia Cache

NoAm Cache

Euro Cache

in-PoP Cache

Cust Facing

in-PoP Cache

Cust Facing

in-PoP Cache

Cust Facing

in-PoP Cache

Cust Facing

in-PoP Cache

Cust Facing

in-PoP Cache

Cust Facing

in-PoP Cache

Cust Facing

Extremely Large ISP Deployment

High Priority

Lower Priority
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
r0.sea#\texttt{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

<table>
<thead>
<tr>
<th>Network</th>
<th>Maxlen</th>
<th>Origin-AS</th>
<th>Source</th>
<th>Neighbor</th>
</tr>
</thead>
<tbody>
<tr>
<td>67.21.36.0/24</td>
<td>24</td>
<td>3970</td>
<td>0</td>
<td>198.180.150.1/424</td>
</tr>
<tr>
<td>98.128.0.0/24</td>
<td>24</td>
<td>4128</td>
<td>0</td>
<td>198.180.150.1/424</td>
</tr>
<tr>
<td>98.128.0.0/16</td>
<td>16</td>
<td>3130</td>
<td>0</td>
<td>198.180.150.1/424</td>
</tr>
<tr>
<td>98.128.6.0/24</td>
<td>24</td>
<td>4128</td>
<td>0</td>
<td>198.180.150.1/424</td>
</tr>
<tr>
<td>98.128.9.0/24</td>
<td>24</td>
<td>3130</td>
<td>0</td>
<td>198.180.150.1/424</td>
</tr>
<tr>
<td>98.128.30.0/24</td>
<td>24</td>
<td>1234</td>
<td>0</td>
<td>198.180.150.1/424</td>
</tr>
<tr>
<td>129.6.128.0/17</td>
<td>17</td>
<td>49</td>
<td>0</td>
<td>198.180.150.1/424</td>
</tr>
<tr>
<td>147.28.0.0/16</td>
<td>16</td>
<td>3130</td>
<td>0</td>
<td>198.180.150.1/424</td>
</tr>
<tr>
<td>147.28.224.0/19</td>
<td>19</td>
<td>4128</td>
<td>0</td>
<td>198.180.150.1/424</td>
</tr>
</tbody>
</table>
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
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
      path 09AF35CC RPKI State valid
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
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
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
*APNIC and ARIN are simulations constructed from public data

**The Open TestBed**

- **Trust Anchor**
  - **ARIN**
    - **ISC**
    - **Google**
    - **BWC**
  - **APNIC**
    - **RGnet**
    - **JPNIC**
  - **IIJ**
  - **Cristel**
  - **Mesh**

- **Level (3)**
- **Running Code**
- **Repository**
- **until we get IANA to act as the parent**

- **APNIC** runs own RPKI to keep private key private and control own fate, but publishes at IIJ until we get IANA to act as the parent.
- **APNIC** and **ARIN** are simulations constructed from public data.
The Big Speedbump

“We will never control Internet routing”
-- The RIRs
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.

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.

Up-Chain Expiration

These are not Identity Certs

Sloppy Admin, Cert Soon to Expire!

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 Do You Call?
Ghostbusters!

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

“All 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?
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 & ISC

• Cisco, Juniper, Google, NTT, Equinix