The RPKI and BGP Origin Validation

APRICOT / New Delhi
2012.02.27

Randy Bush <randy@psg.com>
Rob Austein <sra@isc.org>
Steve Bellovin <smb@cs.columbia.edu>

And a cast of thousands! Well, dozens :)

2012.02.27 APRICOT RtgSec
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
Prefix Has Origin AS

BGP routing table entry for 98.128.1.0/24

Paths: (32 available, best #21, table Default-IP-Routing-Table)

1221 4637 3561 2914 4128

Origin AS

The AS-Path
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)
Resource Public Key Infrastructure (RPKI)
X.509 RPKI Being Developed & Deployed by IANA, RIRs, and Operators
Private/Public Keys

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

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

Unencrypted message

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

Unencrypted message
Digital Signature

Diagram:
- Sender
  - Message
    - To be, or not to be, that is the question, whether tis nobler in the...
  - Message Digest Algorithm
    - Encryption Algorithm
      - Message Digest
        +"z["~\@!"
        ]":>>
  - Sender’s Private Key
    - Encrypted Message Digest
      *^=/e
      \%;<=+$=\$1

Receiver
- Message
  - To be, or not to be, that is the question, whether tis nobler in the...
  - Message Digest Algorithm
  - Sender’s Public Key
    - Encrypted Message Digest
      +"z["~\@!"
      ]":>>
  - Message
    - equal?
      yes
      - Message transmitted correctly
      no
      - Error! Message has been modified!
X.509 Certificate w/ 3779 Ext

Signed by Parent's Private Key

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
Multiple ROAs
Make Before Break

I Plan to Switch Providers
ROA Aggregation Using Max Length
RPKI-Based
Origin Validation
RPKI Certificate Engine

GUI

Up / Down to Parent

Up / Down to Child

Publication Protocol

Resource PKI
- IP Resource Certs
- ASN Resource Certs
- Route Origin Attestations

Create ROA

Matched Routes
- Prefix: 98.136.1.0/24
- Origin AS: 4128
- Validation Status: Pending

Matched Routes
- Prefix: 98.136.1.0/24
- Origin AS: 3100
- Validation Status: Pending
Warning What ROA Will Do

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.

<table>
<thead>
<tr>
<th>AS</th>
<th>Prefix</th>
<th>Max Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>3130</td>
<td>98.128.1.0/24</td>
<td>24</td>
</tr>
</tbody>
</table>

Create  Cancel

Matched Routes

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Origin AS</th>
<th>Validation Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>98.128.1.0/24</td>
<td>4128</td>
<td>INVALID</td>
</tr>
<tr>
<td>98.128.1.0/24</td>
<td>3130</td>
<td>VALID</td>
</tr>
</tbody>
</table>
Issuing Parties
Extremely Large ISP Deployment

Global RPKI

Caches Feed Caches

Asia Cache

NoAm Cache

Euro Cache

in-PoP Cache in-PoP Cache in-PoP Cache in-PoP Cache in-PoP Cache in-PoP Cache in-PoP Cache in-PoP Cache

Cust Facing Cust Facing Cust Facing Cust Facing Cust Facing Cust Facing Cust Facing Cust Facing

High Priority Lower Priority
How Do ROAs Affect BGP Updates?
In NOC

IANA

APNIC

IIJ

Trust Anchor

SIA Pointers

SIA Pointers

Resource PKI

Resource PKI

Resource PKI

Up

Down

Publication Protocol

Up

Down

Publication Protocol

Publication Protocol

RCynic Gatherer

Validated Cache

RPKI to Rtr Protocol

BGP Decision Process

In PoP

2012.02.27 APRICOT RtgSec
### IPv4 Prefix

<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>Version</td>
<td>Type</td>
<td>reserved = zero</td>
</tr>
<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>Flags</td>
<td>Length</td>
<td>Length</td>
</tr>
<tr>
<td>0..32</td>
<td>0..32</td>
<td>zero</td>
</tr>
</tbody>
</table>

IPv4 prefix

Autonomous System Number

---

2012.02.27 APRICOT RtgSec
IPv6 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></td>
<td></td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

Length=40

<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..128</td>
<td>0..128</td>
</tr>
</tbody>
</table>

---

IPv6 prefix

---

Autonomous System Number
BGP Updates are compared with ROAs loaded from the RPKI
Marking BGP Updates

BGP Peer

BGP Data

BGP Updates

mark

Valid
Invalid
NotFound

RPKI-Rtr Protocol

RPKI Cache

RPKI ROAs
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
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
...```
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
What are the BGP / ROA Matching Rules?
A Prefix is Covered by a ROA when the ROA prefix length is less than or equal to the Route prefix length.

- BGP: 98.128.0.0/16
- ROA: 98.128.0.0/12-16
  - Covers
- ROA: 98.128.0.0/16-24
  - Covers
- ROA: 98.128.0.0/20-24
  - No. It’s Longer
Prefix is matched by a ROA when the prefix is covered by that ROA, prefix length is less than or equal to the ROA max-len, and the Route Origin AS is equal to the ROA’s AS.

<table>
<thead>
<tr>
<th>BGP</th>
<th>98.128.0.0/16  AS 42</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROA</td>
<td>98.128.0.0/12-16 AS 42</td>
</tr>
<tr>
<td>ROA</td>
<td>98.128.0.0/16-24 AS 666</td>
</tr>
<tr>
<td>ROA</td>
<td>98.128.0.0/20-24 AS 42</td>
</tr>
</tbody>
</table>
Matching and Validity

ROA₀  98.128.0.0/16-24  AS 6

ROA₁  98.128.0.0/16-20  AS 42

BGP  98.128.0.0/12  AS 42  NotFound, shorter than ROAs
BGP  98.128.0.0/16  AS 42  Valid, Matches ROA₁
BGP  98.128.0.0/20  AS 42  Valid, Matches ROA₁
BGP  98.128.0.0/24  AS 42  Invalid, longer than ROAs
BGP  98.128.0.0/24  AS 6    Valid, Matches ROA₀
The Operator Tests and then Sets 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 100

route-map validity-1
match rpki invalid
set metric 150

route-map validity-2
set metric 50
Allocation in Reality

/16 Assignment from RIR
ROA Use

My Aggregate ROA

Customer ROAs

I Generate for 'Lazy' Customer

My Infrastructure

BGP Cust

Static (non BGP) Cust

Unused

2012.02.27 APRICOT RtgSec
Covering a Customer

I Issue a ROA for the Covering Prefix

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

These are not Identity Certs
So Who You Gonna Call?

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 You Gonna Call?
Ghostbusters!

![Diagram showing the relationship between various internet address blocks and public keys, with a vCard for a character named Human's Name. The vCard includes details such as name, organizational entity, address, and contact information.]
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?
Open Source (BSD Lisc)
Running Code
https://rpki.net/

Test Code in Routers
Talk to C & J
Vendor Code

- Cisco IOS and XR test code have Origin Validation now, shipping some code now
- Juniper has test code now, ship 2Q2012
- Work continues daily in test routers
- Compute load much less than ACLs from IRR data, 10µsec per update!
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
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

- Hash Signed by Router Key AS0.rtr-xx
- Hash Signed by Router Key AS1-rtr-yy

Signed Forward Reference