Resource Public Key Infrastructure (RPKI)

Nurul Islam Roman, APNIC
Target Audience

- Knowledge of Internet Routing (specially BGP)
- Fair idea on Routing Policy
- No need to know Cryptography
- Basic knowledge of PKI (Public Key Infrastructure)
Agenda

• BGP 101
• Routing Policy
• RPKI
  – Configuration & Hands on router config demo
BGP (AS)

Send a packet to 2001:DB8::1

I have 2001:DB8::/32

65551

65549
AS Path

2001:DB8::/32 65551 65550 65549 i

Send a packet to 2001:DB8::1

I have 2001:DB8::/32
AS Path

2001:DB8::/32  65551  65550  65549  i

2001:DB8::/48  65551  65550  65536  i

Send a packet to 2001:DB8::1

I have 2001:DB8::/48

I have 2001:DB8::/32

I have 2001:DB8::/48
BGP Best Path Calculation

- Drop if own AS in AS-Path
- Prefer path with highest Weight
- Highest Local Preference
- Shortest AS-Path
- Lowest MED
- Path with shortest next hop metric (minimum IGP cost)
- Oldest received path
- Path from lowest neighbour address
Constructing the Forwarding Table

- **BGP in process**
  - Accepted
  - Discarded

- **BGP out process**
  - Best path

- **BGP table**

- **Routing table**

- **Forwarding table**

- **BGP peer**

- **In**
  - Everything
Control Plane and Forwarding Plane
Routing Incidents Types

- Incidents
  - Misconfiguration
  - Malicious
  - Targeted Traffic Misdirection

- For theory of positivity let's call all these as Mis-Origination

- Traffic Hijacking or Prefix Hijacking assumes Negative intent
Historical Incident

- April 1997: The "AS 7007 incident" UU/Sprint for 2 days
- February 24, 2008: Pakistan's attempt to block YouTube access within their country takes down YouTube entirely.[6]
- November 11, 2008: The Brazilian ISP CTBC - Companhia de Telecomunicações do Brasil Central leaked their internal table into the global BGP table.
- April 8, 2010: China Telecom originated 37,000 prefixes not belonging to them in 15 minutes, causing massive outage of services globally.

Securing Internet Routing

To Secure Internet Routing; we need to check:

A network should only originate his own prefix
1. How do we verify?
2. How do we avoid false advertisement?

A transit network should filter customer prefix
1. Check customer prefix and ASN delegation
2. Transitive trust
Routing Policy

• Public description of the relationship between external BGP peers
• Can also describe internal BGP peer relationship
• Usually registered at an IRR (Internet Routing Registry) such as RADB or APNIC
Routing Policy

• Who are my BGP peers

• What routes are
  – Originated by a peer
  – Imported from each peer
  – Exported to each peer
  – Preferred when multiple routes exist

• What to do if no route exists
Prefix Advertise to Internet

• Ingress prefix from downstream:
  – Option 1: Customer single home and non portable prefix
    • Customer is not APNIC member prefix received from upstream ISP
  – Option 2: Customer single home and portable prefix
    • Customer is APNIC member receive allocation as service provider but no AS number yet
  – Option 3: Customer multihome and non portable prefix
    • Customer is not APNIC member both prefix and ASN received from upstream ISP
  – Option 4: Customer multihome and portable prefix
    • Customer is APNIC member both prefix and ASN received from APNIC
Prefix Filtering BCP [Single home]

- Option 1: Customer **single home and non portable prefix**

ISP Prefix
3fff:ffff::/32

Customer Prefix
3fff:ffff:dcdc::/48

AS17821
Static 3fff:ffff:dcdc::/48 to customer WAN Interface
No LoA Check of Cust prefix upstream

NO BGP
Static Default to ISP WAN Interface downstream
Prefix Filtering BCP [Single home]

• Option 2: Customer single home and portable prefix

ISP Prefix
3ffe:ffff::/32

Customer Prefix
2001:0DB8::/32

AS17821
Static 2001:0DB8::/32 to customer WAN Interface
BGP network 2001:0DB8::/32 AS17821
Check LoA of Cust prefix

NO BGP
Static Default to ISP WAN Interface
Static 2001:0DB8::/32 null0

upstream

downstream

Internet
Prefix Filtering [Multihome]

- Option 3: Customer multihome and non portable prefix

Internet

AS131107
- Check LoA of Cust prefix
- Manual process e-mail to tech-c
- Automated process route object or RPKI
- Nearly same filter requirement as other ISP

ISP Prefix
3fff:ffff::/32
 upstream
 can change

AS64500
- eBGP peering with both ISP WAN interface
- BGP network 3fff:ffff::/48 AS64500

AS17821
- eBGP peering with customer WAN interface
- No LoA Check of Cust prefix

Customer Prefix
3fff:ffff:dcdc::/48
 upstream
 can not change

BGP network 3fff:ffff:dcdc::/48 AS64500 i or aggregate address from gateway router

APNIC
Prefix Filtering [Multihome]

• Option 4: Customer **multihome** and **portable prefix**

- **ISP Prefix**: 3fff:ffff::/32
- **Customer Prefix**: 2001:0DB8::/32

**AS131107**
- Check LoA of Cust prefix
- Manual process e-mail to tech-c
- Automated process route object or RPKI
- Nearly same filter requirement as other ISP

**AS17821**
- Check LoA of Cust prefix
- Manual process e-mail to tech-c
- Automated process route object or RPKI

**AS64500**
- eBGP peering with both ISP WAN Interface
- BGP network 2001:0DB8::/32 AS64500 i or aggregate address from gateway router

**upstream can change**
Secure Internet Routing

Routing Policy System (RPS) Working Group’s model

Secure Inter-Domain Routing (SIDR) Working Group’s model
Purpose of RPKI

• RPKI replaces IRR or lives side by side?
  – Side by side: different advantages
  – Security, almost real time, simple interface: RPKI

• Purpose of RPKI
  – Is that ASN authorized to originate that address range?
RPKI Origin Validation

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Origin1</th>
<th>Origin2</th>
<th>Origin3</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001:DB8::/32</td>
<td>65551</td>
<td>65550</td>
<td>65549</td>
<td>VALID</td>
</tr>
<tr>
<td>2001:DB8::/48</td>
<td>65551</td>
<td>65550</td>
<td>65536</td>
<td>INVALID</td>
</tr>
</tbody>
</table>

Send a packet to 2001:DB8::1

I have 2001:DB8::/32

I have 2001:DB8::/48

65551

65550

65549

65536
RPKI Deployment

Phase 1
Origin Validation

Phase 2
Path Validation

Send a packet to 2001:DB8::1

I have 2001:DB8::/32

65550

65549

65551

APNIC
Internet Registry (IR) / RIR

- Maintains Internet Resources such as IP addresses and ASNs, and publish the registration information
  - Allocations for Local Internet Registries
  - Assignments for end-users

- APNIC is the Regional Internet Registry (RIR) in the Asia Pacific region
  - National Internet Registry (NIR) exists in several economies
The Eco-System

Internet Assigned Numbers Authority

Regional IR (RIR)

National IR (NIR)

Internet Service Provider

End User
Goals of RPKI

• Able to authoritatively prove who owns an IP Prefix and what AS(s) may Announce It
  – Reducing routing leaks
  – Attaching digital certificates to network resources (AS Number & IP Address)

• Prefix Ownership Follows the Allocation Hierarchy IANA, RIRs, ISPs, …
Advantage of RPKI

• Useable toolset
  – No installation required
  – Easy to configure manual overrides

• Tight integration with routers
  – Supported routers have awareness of RPKI validity states

• Stepping stone for AS-Path Validation
  – Prevent Attacks on BGP
RPKI Implementation

• Two RPKI implementation type
  – **Delegated**: Each participating node becomes a CA and runs their own RPKI repository, delegated by the parent CA.
  – **Hosted**: The RIR runs the CA functionality for interested participants.
Two Components

- Certificate Authority (CA)
  - Internet Registries (RIR, NIR, Large LIR)
  - Issue certificates for customers
  - Allow customers to use the CA’s GUI to issue ROAs for their prefixes

- Relying Party (RP)
  - Software which gathers data from CAs
Issuing Party

- Internet Registries (RIR, NIR, Large LIRs)
- Acts as a Certificate Authority and issues certificates for customers
- Provides a web interface to issue ROAs for customer prefixes
- Publishes the ROA records
Relying Party (RP)

Software which gathers data from CAs
Also called RP cache or validator
RPKI Building Blocks

1. Trust Anchors (RIR’s)
2. Route Origination Authorizations (ROA)
3. Validators
1. PKI & Trust Anchors
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.
- **Same alike http with SSL aka https**
X.509 Certificates 3779 EXT

Certificates are X.509 certificates that conform to the PKIX profile [PKIX]. They also contain an extension field that lists a collection of IP resources (IPv4 addresses, IPv6 addresses and AS Numbers) [RFC3779]
Trust Anchor

Resource Allocation Hierarchy

AfriNIC
RIPE NCC
APNIC
IANA
Trust Anchor Certificate

Issued Certificates match allocation actions

NIR
ISP
ISP
ISP
ISP
ISP
ISP

Source: http://isoc.org/wp/ietfjournal/?p=2438
RPKI Chain of Trust

• The RIRs hold a self-signed root certificate for all the resources that they have in the registry
  – They are the trust anchor for the system

• That root certificate is used to sign a certificate that lists your resources

• You can issue child certificates for those resources to your customers
  – When making assignments or sub allocations
2. ROA
Route Origin Authorizations
Route Origination Authorizations (ROA)

- A ROA is a digitally signed object that provides a means of verifying that an IP address block holder has authorized an Autonomous System (AS) to originate routes to one or more prefixes within the address block.
- With a ROA, the resource holder is attesting that the origin AS number is authorized to announce the prefix(es). The attestation can be verified cryptographically using RPKI.
Route Origination Authorizations (ROA)

• Next to the prefix and the ASN which is allowed to announce it, the ROA contains:
  – A minimum prefix length
  – A maximum prefix length
  – An expiry date
  – Origin ASN

• Multiple ROAs can exist for the same prefix

• ROAs can overlap
3. Validators
Origin Validation

• Router gets ROA information from the RPKI Cache
  – RPKI verification is done by the RPKI Cache
• The BGP process will check each announcement with the ROA information and label the prefix
Result of Check

- **Valid** – Indicates that the prefix and AS pair are found in the database.
- **Invalid** – Indicates that the prefix is found, but either the corresponding AS received from the EBGP peer is not the AS that appears in the database, or the prefix length in the BGP update message is longer than the maximum length permitted in the database.
- **Not Found / Unknown** – Indicates that the prefix is not among the prefixes or prefix ranges in the database.

Valid > Unknown > Invalid
## ROA Example

<table>
<thead>
<tr>
<th>ROA</th>
<th>Origin AS</th>
<th>Prefix</th>
<th>Max Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>VALID</td>
<td>AS65420</td>
<td>10.0.0.0/16</td>
<td></td>
</tr>
<tr>
<td>VALID</td>
<td>AS65420</td>
<td>10.0.128.0/17</td>
<td></td>
</tr>
<tr>
<td>INVALID</td>
<td>AS65421</td>
<td>10.0.0.0/16</td>
<td></td>
</tr>
<tr>
<td>INVALID</td>
<td>AS65420</td>
<td>10.0.10.0/24</td>
<td></td>
</tr>
<tr>
<td>UNKNOWN</td>
<td>AS65430</td>
<td>10.0.0.0/8</td>
<td></td>
</tr>
</tbody>
</table>
Local Policy

- You can define your policy based on the outcomes
  - Do nothing
  - Just logging
  - Label BGP communities
  - Modify preference values
  - Rejecting the announcement
In summary

• As an announcer/LIR
  – You choose if you want certification
  – You choose if you want to create ROAs
  – You choose AS, max length

• As a Relying Party
  – You can choose if you use the validator
  – You can override the lists of valid ROAs in the cache, adding or removing valid ROAs locally
  – You can choose to make any routing decisions based on the results of the BGP Verification (valid/invalid/unknown)
RPKI Caveats

• When RTR session goes down, the RPKI status will be not found for all the bgp route after a while
  – Invalid => not found
  – we need several RTR sessions or care your filtering policy

• In case of the router reload, which one is faster, receiving ROAs or receiving BGP routes?
  – If receiving BGP is match faster than ROA, the router propagate the invalid route to others
  – We need to put our Cache validator within our IGP scope
RPKI Further Reading

- RFC 5280: X.509 PKI Certificates
- RFC 3779: Extensions for IP Addresses and ASNs
- RFC 6481-6493: Resource Public Key Infrastructure
RPKI Configuration
RPKI Configuration

• Resources:
  – AS: 45192 [APNIC-TRAINING-DC-AS-AP]
  – IPv4: 202.125.97.0/24
  – IPv4: 203.176.189.0/24
  – IPv6: 2001:DF2:ee01::/48

• Process
  – Create ROA
  – Setup cache validation server
  – Validate the ROA
Phase I - Publishing ROA

• Login to your MyAPNIC portal
• Required valid certificate
• Go to Resources > Certification Tab
Phase I - Publishing ROA

1. Click on Certification.

2. Select the option:
   - I want to operate in the MyAPNIC RPKI portal.
   - I want to host my own certification authority and run an RPKI engine myself.

3. Click Next.

4. Accept the Terms and Conditions of APNIC Certification Authority.

5. Create my Certification Authority.
Phase I - Publishing ROA

- Show available prefix for which you can create ROA

<table>
<thead>
<tr>
<th>Origin AS</th>
<th>Prefix</th>
</tr>
</thead>
<tbody>
<tr>
<td>45192</td>
<td>2001:df0:a::/48</td>
</tr>
<tr>
<td>45192</td>
<td>203.176.189.0/24</td>
</tr>
</tbody>
</table>
Phase I - Publishing ROA

• Create ROA for smaller block.

ROA Configuration

1. Write your ASN
2. Your IP Block
3. Subnet
4. Click Add

- Origin ASN: 45192
- Prefix: 2001:df0:a::/48
- Max Length: 48

APNIC
Phase I - Check your ROA

fakrul@www:~$ whois -h whois.bgpmon.net 2001:df0:a::/48

<table>
<thead>
<tr>
<th>Prefix:</th>
<th>2001:df0:a::/48</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prefix description:</td>
<td>APNIC Training data centre</td>
</tr>
<tr>
<td>Country code:</td>
<td>AU</td>
</tr>
<tr>
<td><strong>Origin AS:</strong></td>
<td><strong>45192</strong></td>
</tr>
<tr>
<td>Origin AS Name:</td>
<td>Two-byte AS number for APNIC Training Data Centre</td>
</tr>
<tr>
<td><strong>RPKI status:</strong></td>
<td><strong>ROA validation successful</strong></td>
</tr>
<tr>
<td>First seen:</td>
<td>2013-12-11</td>
</tr>
<tr>
<td>Last seen:</td>
<td>2016-01-03</td>
</tr>
<tr>
<td>Seen by #peers:</td>
<td>170</td>
</tr>
</tbody>
</table>
Phase I - Check your ROA

fakrul@www:~$ whois -h whois.bgpmon.net " --roa 45192 2001:df0:a::/48"

0 - Valid
------------------------
ROA Details
------------------------

**Origin ASN:** AS45192

Not valid Before: 2016-01-02 02:30:14
Not valid After: 2020-07-27 00:00:00  Expires in 4y204d23h46m30.4000000059605s

Trust Anchor: rpki.apnic.net

**Prefixes:** 2001:df0:a::/48 (max length /48)
203.176.189.0/24 (max length /24)
Phase II - RPKI Validator

- Download RPKI Validator

Tools and Resources

Here you can find an overview of all information and tools for the Resource Certification (RPKI) service.

RIPE NCC RPKI Validator 2.21 (Updated 3 November 2015)

This application allows operators to download and validate the global RPKI data set for use in their BGP decision making process and router configuration.

System requirements: a UNIX-like OS, Java 7, rsync and 2GB free memory. To install, simply unpack the archive and run "rpki-validator.sh" from the base folder.

For more information, view the release notes. You can also contribute to the project on GitHub.
Phase II - RPKI Validator

# tar -zxvf rpki-validator-app-2.21-dist.tar.gz
# cd rpki-validator-app-2.21
# ./rpki-validator.sh start
Phase II - RPKI Validator

http://ip_address:8080
# Phase III - Router Configuration

## 1. Establish session with RPKI Validator

### Junos

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>set routing-options validation group RPKI session 202.4.96.221 refresh-time 120</td>
<td>Establish RPKI session with RP Validator</td>
</tr>
<tr>
<td>set routing-options validation group RPKI session 202.4.96.221 hold-time 180</td>
<td>Establish RPKI session with RP Validator</td>
</tr>
<tr>
<td>set routing-options validation group RPKI session 202.4.96.221 port 8282</td>
<td>Establish RPKI session with RP Validator</td>
</tr>
<tr>
<td>set routing-options validation group RPKI session 202.4.96.221 local-address 103.21.75.1</td>
<td>Establish RPKI session with RP Validator</td>
</tr>
</tbody>
</table>

### IOS

```bash
router bgp 64500
    bgp log-neighbor-changes
    bgp rpki server tcp 202.4.96.221 port 8282 refresh 120
```
## Phase III - Router Configuration

### 2. Configure policy to tag ROA

**Junos**

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term valid from protocol bgp</code></td>
<td>Configure policy with valid terms based on validation database.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term valid from validation-database valid</code></td>
<td>Specify local preference for valid validation.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term valid then local-preference 110</code></td>
<td>Set local preference for valid routes.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term valid then validation-state valid</code></td>
<td>Accept routes with valid validation state.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term invalid from protocol bgp</code></td>
<td>Configure policy with invalid terms based on validation database.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term invalid from validation-database invalid</code></td>
<td>Specify local preference for invalid validation.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term invalid then local-preference 90</code></td>
<td>Set local preference for invalid routes.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term invalid then validation-state invalid</code></td>
<td>Accept routes with invalid validation state.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term unknown from protocol bgp</code></td>
<td>Configure policy with unknown terms based on validation database.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term unknown from validation-database unknown</code></td>
<td>Specify local preference for unknown validation.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term unknown then local-preference 100</code></td>
<td>Set local preference for unknown routes.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term unknown then validation-state unknown</code></td>
<td>Accept routes with unknown validation state.</td>
</tr>
<tr>
<td><code>set policy-options policy-statement ROUTE-VALIDATION term unknown then accept</code></td>
<td>Accept routes with unknown terms.</td>
</tr>
</tbody>
</table>
Phase III - Router Configuration

2. Configure policy to tag ROA

<table>
<thead>
<tr>
<th>IOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>!</td>
</tr>
<tr>
<td>route-map ROUTE-VALIDATION permit 10</td>
</tr>
<tr>
<td>match rpki invalid</td>
</tr>
<tr>
<td>set local-preference 90</td>
</tr>
<tr>
<td>!</td>
</tr>
<tr>
<td>route-map ROUTE-VALIDATION permit 20</td>
</tr>
<tr>
<td>match rpki not-found</td>
</tr>
<tr>
<td>set local-preference 100</td>
</tr>
<tr>
<td>!</td>
</tr>
<tr>
<td>route-map ROUTE-VALIDATION permit 30</td>
</tr>
<tr>
<td>match rpki valid</td>
</tr>
<tr>
<td>set local-preference 110</td>
</tr>
</tbody>
</table>
Phase III - Router Configuration

3. Push policy to the BGP neighbour

<table>
<thead>
<tr>
<th>Junos</th>
</tr>
</thead>
<tbody>
<tr>
<td>set protocols bgp import ROUTE-VALIDATION</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>router bgp 64500</td>
</tr>
<tr>
<td>bgp log-neighbor-changes</td>
</tr>
<tr>
<td>!other neighbour related configuration</td>
</tr>
<tr>
<td>neighbor 10.1.1.2 route-map ROUTE-VALIDATION in</td>
</tr>
</tbody>
</table>
Check your prefix

Junos

show route protocol bgp 203.176.189.0

inet.0: 575802 destinations, 575803 routes (575802 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both

203.176.189.0/24  *[BGP/170] 1w3d 23:50:49, localpref 100, from 79.141.168.1
                         AS path: 33926 2828 7545 24130 4608 45192 17821 I, validation-state: unknown
                         > to 193.34.50.1 via em0.0
Check your prefix

**IOS**

```bash
rpki-rtr>show ip bgp 203.176.189.0/24
BGP routing table entry for 203.176.189.0/24, version 70470025
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 1
  3333 1273 4637 1221 4608 45192
    193.0.19.254 from 193.0.3.5 (193.0.0.56)
      Origin IGP, **localpref 110**, valid, external
      Community: 83449328 83450313
      path 287058B8 RPKI State valid
```
# Commands

<table>
<thead>
<tr>
<th>Command (Junos)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show validation session detail</code></td>
<td>Check session status of cache validator server</td>
</tr>
<tr>
<td><code>show validation statistics</code></td>
<td>Statistics on valid/invalid prefixes</td>
</tr>
<tr>
<td><code>show validation database</code></td>
<td>Full validation database</td>
</tr>
<tr>
<td><code>show route protocol bgp</code> validation-state</td>
<td>Find valid/invalid/unknown routes</td>
</tr>
<tr>
<td><code>valid/invalid/unknown</code></td>
<td></td>
</tr>
</tbody>
</table>
!Caution!
Testbed

• Cisco (hosted by the RIPE NCC)
  – Public Cisco router: rpki-rtr.ripe.net
  – Telnet username: ripe / No password

• Juniper (hosted by Kaia Global Networks)
  – Public Juniper routers: 193.34.50.25, 193.34.50.26
  – Telnet username: rpki / Password: testbed
Configuration - Reference Link

• Cisco

• Juniper
www.apnic.net/roa