BGP Origin Validation

ISP Workshops

These materials are licensed under the Creative Commons Attribution-NonCommercial 4.0 International license (http://creativecommons.org/licenses/by-nc/4.0/)

Last updated 29th July 2019

Acknowledgements

- This material was built from contributions by Randy Bush, Mark Tinka and others

- Use of these materials is encouraged as long as the source is fully acknowledged and this notice remains in place

- Bug fixes and improvements are welcomed
  - Please email *workshop (at) bgp4all.com*

Philip Smith

Validating BGP Route Announcements

- How do we know that an AS is permitted to originate the prefix it is originating?
- Implicit trust?
- Because the Internet Routing Registry says so?
  - The Internet Routing Registry (IRR) only documents routing policy
  - And has a large amount of outdated/invalid information
- Is there something else?
  - Yes: Route Origin Authorisation

RPKI

- RPKI – Resource Public Key Infrastructure, the Certificate Infrastructure to Support the other Pieces
  - 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, etc)
  - 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)

BGP – Why Origin Validation?

- Prevent YouTube accident & Far Worse
- Prevents most accidental announcements
- Does not prevent malicious path attacks
- That requires ‘Path Validation’ and locking the data plane to the control plane, the third step, BGPsec

What is RPKI?

- Resource Public Key Infrastructure (RPKI)
  - A security framework for verifying the association between resource holder and their Internet resources
  - Created to address the issues discussed in RFC 4593 "Generic Threats to Routing Protocols" (Oct 2006)
- Helps to secure Internet routing by validating routes
  - Proof that prefix announcements are coming from the legitimate holder of the resource
  - RFC 6480 – An Infrastructure to Support Secure Internet Routing (Feb 2012)

Benefits of RPKI - Routing

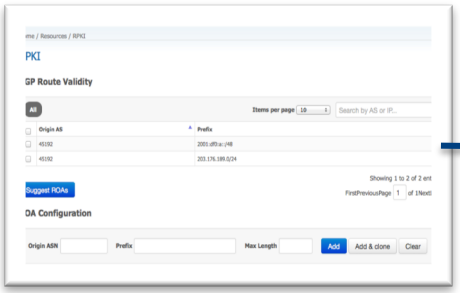
- Prevents route hijacking
  - A prefix originated by an AS without authorization
  - Reason: malicious intent
- Prevents mis-origination
  - A prefix that is mistakenly originated by an AS which does not own it
  - Also route leakage
  - Reason: configuration mistake / fat finger

BGP Security (BGPsec)

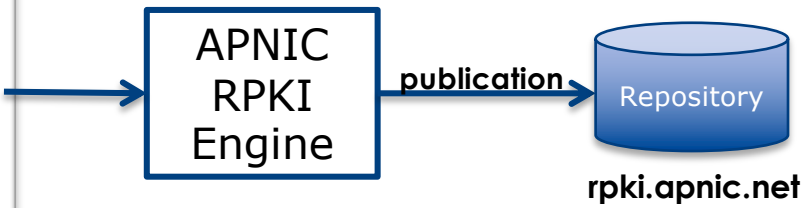
- Extension to BGP that provides improved security for BGP routing
- Being worked on by the SIDR Working Group at IETF
- Implemented via a new optional non-transitive BGP attribute that contains a digital signature
- Two components:
  - BGP Prefix Origin Validation (using RPKI)
  - BGP Path Validation

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



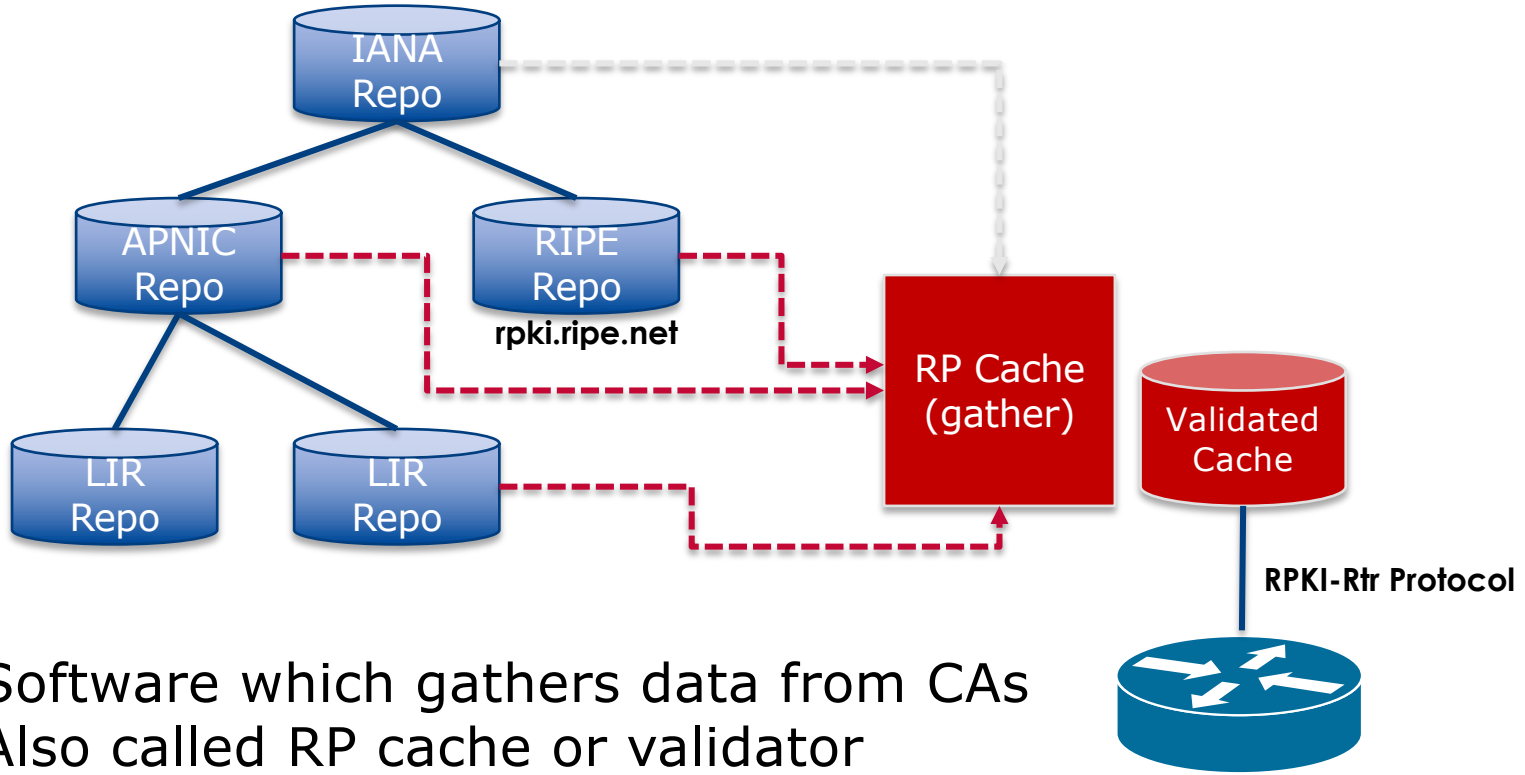
MyAPNIC GUI



rpki.apnic.net

Courtesy of APNIC: https://apnic.net

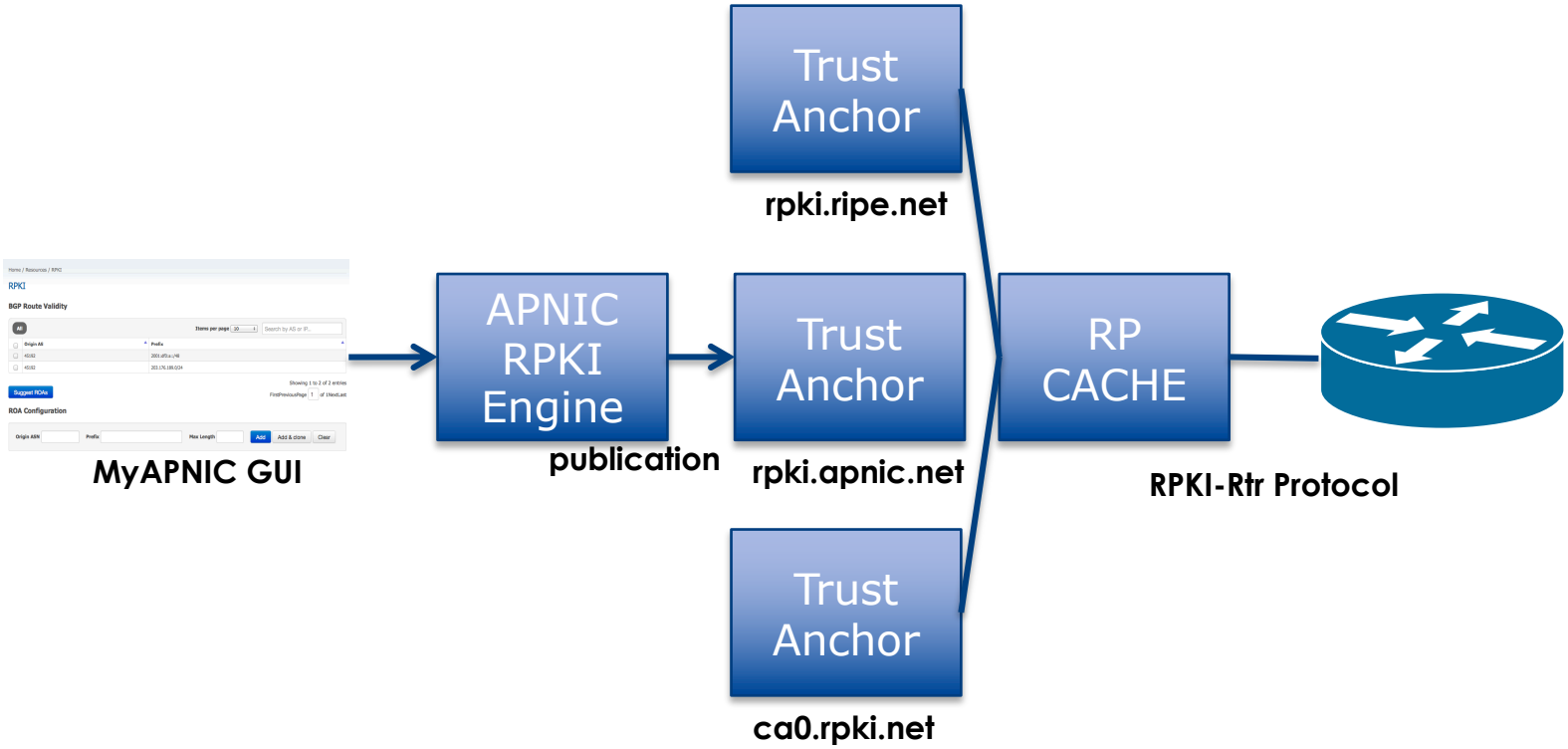
Relying Party (RP)



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

Courtesy of APNIC: https://apnic.net

RPKI Components



Courtesy of APNIC: https://apnic.net

Route Origin Authorisation (ROA)

- A digital object that contains a list of address prefixes and one AS number
- It is an authority created by a prefix holder to authorise an AS Number to originate one or more specific route advertisements
- Publish a ROA using your RIR member portal
  - Consult your RIR for how to use their member portal to publish your ROAs

Route Origin Validation

- Router must support RPKI
- Checks an RP cache / validator
- Validation returns 3 states:

State	Description
Valid	When authorisation is found for prefix X coming from ASN Y
Invalid	When authorisation is found for prefix X but **not** from ASN Y
Unknown	When no authorisation data is found for prefix X

Route Origin Validation

- Vendor support:
  - Cisco IOS – available in release 15.2
  - Cisco IOS/XR – available in release 4.3.2
  - Juniper – available in release 12.2
  - Nokia – available in release R12.0R4
  - Huawei – available in release V800R009C10
  - Brocade – available in release TBA
  - FRR – available in release 4.0

RPKI Validator Caches

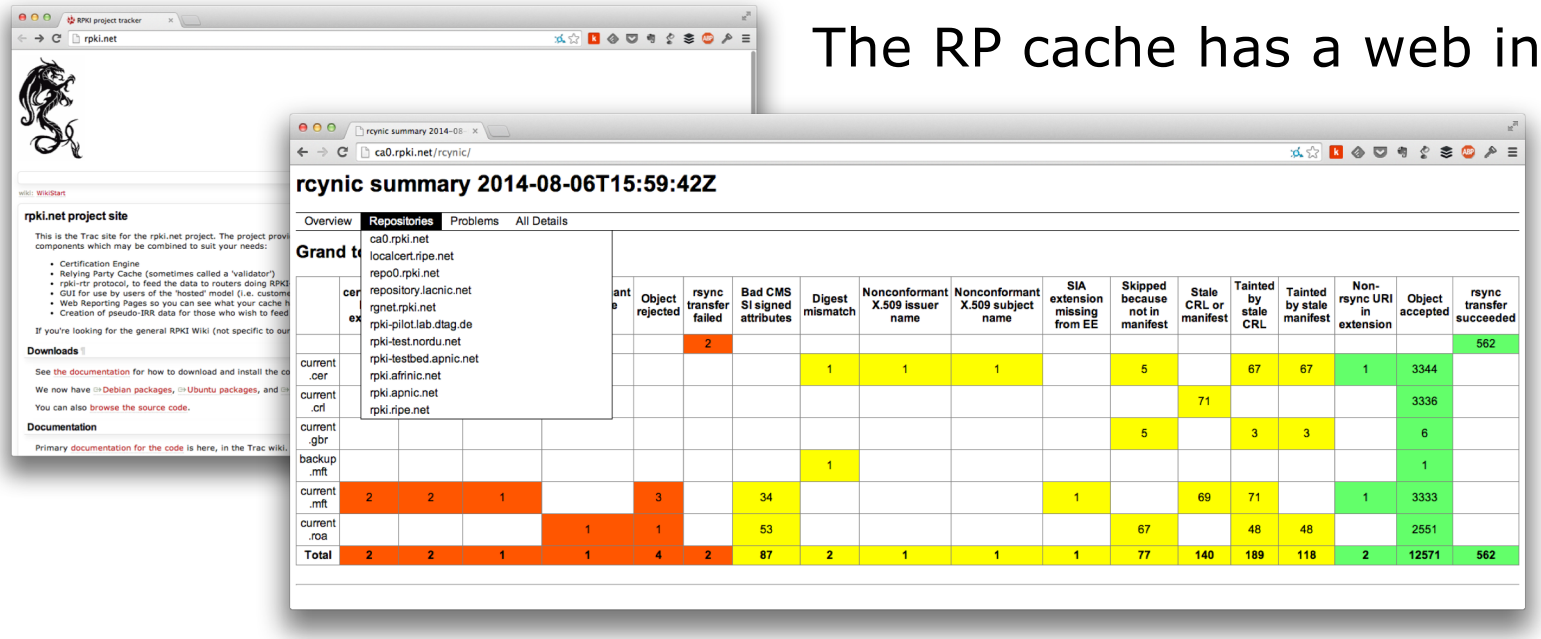
- NLnet Labs Routinator
  - https://www.nlnetlabs.nl/projects/rpki/routinator/
  - https://github.com/NLnetLabs/routinator
- Dragon Research validator
  - https://rpki.net
  - https://github.com/dragonresearch/rpki.net/
- RIPE NCC validator
  - https://github.com/RIPE-NCC/rpki-validator-3/wiki

Build an RP Cache – NLnet Labs

- Consult instructions at:
  - https://github.com/NLnetLabs/routinator
  - <screen shots needed>

Build an RP Cache – Dragon Research

- Download and install from http://rpki.net
  - Instructions here:
    - https://trac.rpki.net/wiki/doc/RPKI/Installation/UbuntuPackages



The RP cache has a web interface

	Object rejected	rsync transfer failed	Bad CMS SI signed attributes	Digest mismatch	Nonconformant X.509 issuer name	Nonconformant X.509 subject name	SIA extension missing from EE	Skipped because not in manifest	Stale CRL or manifest	Tainted by stale CRL	Tainted by stale manifest	Non-rsync URI in extension	Object accepted	rsync transfer succeeded
		2												562
current .cer				1	1	1		5		67	67	1	3344	
current .crl									71				3336	
current .gbr								5		3	3		6	
backup .mft				1									1	
current .mft	3		34				1		69	71		1	3333	
current .roa	1	1	53					67		48	48		2551	
Total	4	2	87	2	1	1	1	77	140	189	118	2	12571	562

RP Cache Deployment

- Network Operator design advice:
  - Deploy at least two Validator Caches
  - Geographically diverse
  - Consider two different validator cache implementations
    - Gives software independence
  - Implement on a Linux container so that the container can be moved between different server clusters as required
  - Make validator listen on both IPv4 and IPv6
  - Securing the validator: Only permit routers running EBGP to have access to the validators

Configure Router to Use Cache: Cisco IOS

- Point router to the local RPKI cache
  - Server listens on port 43779
  - Example:

```
router bgp 64512
 bgp rpki server tcp 10.0.0.3 port 43779 refresh 60
```

  - Once the router's RPKI table is populated, router indicates validation state in the BGP table

Some Cisco IOS commands

- `show ip bgp rpki servers`
  - Provide connection status to the RPKI server
- `show ip bgp rpki table`
  - Shows the VRPs (validated ROA payloads)
- `show ip bgp`
  - Shows the BGP table with status indication next to the prefix

Configure Router to Use Cache: JunOS

1. Connect to validation cache:

```
routing-options {
  validation {
    group ISP {
      session 10.0.0.3;
      port 43779;
      refresh-time 600;
      hold-time 1800;
    }
  }
}
```

- (using same parameters as for the Cisco IOS example)

Configure Router to Use Cache: JunOS

2. Configure validation policies:

```
policy-options {
  policy-statement RPKI-validation {
    term VALID {
      from {
        protocol bgp;
        validation-database valid;
      }
      then {
        validation-state valid;
        next policy;
      }
    }
    term INVALID {
      from {
        protocol bgp;
        validation-database invalid;
      }
      then {
        validation-state invalid;
        next policy;
      }
    }
```

```
(continued)...

    term UNKNOWN {
      from {
        protocol bgp;
        validation-database unknown;
      }
      then {
        validation-state unknown;
        next policy;
      }
    }
  }
}
```

Configure Router to Use Cache: JunOS

3. Apply policy to eBGP session:

```
protocols {
  bgp {
    group EBGP {
      type external;
      local-address 10.0.1.1;
      neighbor 10.1.15.1 {
        description "ISP Upstream";
        import [ RPKI-validation Upstream-in ];
        export LocalAS-out;
        peer-as 64511;
      }
    }
  }
}
```

- Note that policy options *Upstream-in* and *LocalAS-out* are the typical inbound and outbound filters needed for an eBGP session

Check Server

```
lg-01-jnb.za>sh ip bgp rpki servers
BGP SOVC neighbor is 105.16.112.2/43779 connected to port 43779
Flags 64, Refresh time is 300, Serial number is 1463607299
InQ has 0 messages, OutQ has 0 messages, formatted msg 493
Session IO flags 3, Session flags 4008
 Neighbor Statistics:
  Prefixes 25880
  Connection attempts: 44691
  Connection failures: 351
  Errors sent: 35
  Errors received: 0

Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Connection is ECN Disabled
Mininum incoming TTL 0, Outgoing TTL 255
Local host: 105.22.32.2, Local port: 27575
Foreign host: 105.16.112.2, Foreign port: 43779
Connection tableid (VRF): 0
```

Courtesy of SEACOM: http://as37100.net

RPKI Table (IPv4) – November 2018

```
51083 BGP sovc network entries using 4495304 bytes of memory
54231 BGP sovc record entries using 1084620 bytes of memory

Network              Maxlen  Origin-AS  Source  Neighbor
1.0.0.0/24           24      13335      0       105.16.160.2/43779
1.1.1.0/24           24      13335      0       105.16.160.2/43779
1.9.0.0/16           24      4788       0       105.16.160.2/43779
1.9.12.0/24          24      65037      0       105.16.160.2/43779
1.9.21.0/24          24      24514      0       105.16.160.2/43779
1.9.23.0/24          24      65120      0       105.16.160.2/43779
1.9.31.0/24          24      65077      0       105.16.160.2/43779
1.9.65.0/24          24      24514      0       105.16.160.2/43779
1.36.0.0/19          19      4760       0       105.16.160.2/43779
1.36.0.0/16          16      4760       0       105.16.160.2/43779
1.36.32.0/19         19      4760       0       105.16.160.2/43779
1.36.64.0/19         19      4760       0       105.16.160.2/43779
1.36.96.0/19         19      4760       0       105.16.160.2/43779
1.36.128.0/19        19      4760       0       105.16.160.2/43779
...
```

Courtesy of SEACOM: http://as37100.net

RPKI Table (IPv6) – November 2018

```
8639 BGP sovc network entries using 967568 bytes of memory
9583 BGP sovc record entries using 191660 bytes of memory

Network                Maxlen  Origin-AS  Source  Neighbor
2001:200::/32          32      2500       0       2C0F:FEB0:B:1::2/43779
2001:200:136::/48      48      9367       0       2C0F:FEB0:B:1::2/43779
2001:200:900::/40      40      7660       0       2C0F:FEB0:B:1::2/43779
2001:200:8000::/35     35      4690       0       2C0F:FEB0:B:1::2/43779
2001:200:C000::/35     35      23634      0       2C0F:FEB0:B:1::2/43779
2001:200:E000::/35     35      7660       0       2C0F:FEB0:B:1::2/43779
2001:370::/32          32      9607       0       2C0F:FEB0:B:1::2/43779
2001:3A0::/32          128     7521       0       2C0F:FEB0:B:1::2/43779
2001:500:4::/48        48      10745      0       2C0F:FEB0:B:1::2/43779
2001:500:13::/48       48      393225     0       2C0F:FEB0:B:1::2/43779
2001:500:14::/48       48      42         0       2C0F:FEB0:B:1::2/43779
2001:500:15::/48       48      715        0       2C0F:FEB0:B:1::2/43779
2001:500:15::/48       48      42         0       2C0F:FEB0:B:1::2/43779
2001:500:30::/48       48      10745      0       2C0F:FEB0:B:1::2/43779
...
```

Courtesy of SEACOM: http://as37100.net

BGP Table (IPv4)

```
RPKI validation codes: V valid, I invalid, N Not found

Network          Metric LocPrf Path
N*>  1.0.4.0/24       0         37100 6939 4637 1221 38803 56203 i
N*>  1.0.5.0/24       0         37100 6939 4637 1221 38803 56203 i
...
V*>  1.9.0.0/16       0         37100 4788 i
N*>  1.10.8.0/24      0         37100 10026 18046 17408 58730 i
N*>  1.10.64.0/24     0         37100 6453 3491 133741 i
...
V*>  1.37.0.0/16      0         37100 4766 4775 i
N*>  1.38.0.0/23      0         37100 6453 1273 55410 38266 i
N*>  1.38.0.0/17      0         37100 6453 1273 55410 38266 {38266} i
...
I*   5.8.240.0/23     0         37100 44217 3178 i
I*   5.8.241.0/24     0         37100 44217 3178 i
I*   5.8.242.0/23     0         37100 44217 3178 i
I*   5.8.244.0/23     0         37100 44217 3178 i
...
```

Courtesy of SEACOM: http://as37100.net

BGP Table (IPv6)

```
RPKI validation codes: V valid, I invalid, N Not found

Network                Metric LocPrf Path
N*>  2001::/32              0        37100 6939 i
N*   2001:4:112::/48        0        37100 112 i
...
V*>  2001:240::/32          0         37100 2497 i
N*>  2001:250::/48          0         37100 6939 23911 45
N*>  2001:250::/32          0         37100 6939 23911 23910 i
...
V*>  2001:348::/32          0         37100 2497 7679 i
N*>  2001:350::/32          0         37100 2497 7671 i
N*>  2001:358::/32          0         37100 2497 4680 i
...
I*   2001:1218:101::/48     0         37100 6453 8151 278 i
I*   2001:1218:104::/48     0         37100 6453 8151 278 i
N*   2001:1221::/48         0         37100 2914 8151 28496 i
N*>  2001:1228::/32         0         37100 174 18592 i
...
```

Courtesy of SEACOM: http://as37100.net

RPKI BGP State: Valid

```
BGP routing table entry for 2001:240::/32, version 109576927
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 1
  37100 2497
    2C0F:FEB0:11:2::1 (FE80::2A8A:1C00:1560:5BC0) from
                                   2C0F:FEB0:11:2::1 (105.16.0.131)
      Origin IGP, metric 0, localpref 100, valid, external, best
      Community: 37100:2 37100:22000 37100:22004 37100:22060
      path 0828B828 RPKI State valid
      rx pathid: 0, tx pathid: 0x0
```

Courtesy of SEACOM: http://as37100.net

RPKI BGP State: Invalid

```
BGP routing table entry for 2001:1218:101::/48, version 149538323
Paths: (2 available, no best path)
  Not advertised to any peer
  Refresh Epoch 1
  37100 6453 8151 278
    2C0F:FEB0:B:3::1 (FE80::86B5:9C00:15F5:7C00) from
                                    2C0F:FEB0:B:3::1 (105.16.0.162)
      Origin IGP, metric 0, localpref 100, valid, external
      Community: 37100:1 37100:12
      path 0DA7D4FC RPKI State invalid
      rx pathid: 0, tx pathid: 0
```

Courtesy of SEACOM: http://as37100.net

RPKI BGP State: Not Found

```
BGP routing table entry for 2001:200::/32, version 124240929
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 1
  37100 2914 2500
    2C0F:FEB0:11:2::1 (FE80::2A8A:1C00:1560:5BC0) from
                                  2C0F:FEB0:11:2::1 (105.16.0.131)
      Origin IGP, metric 0, localpref 100, valid, external, best
      Community: 37100:1 37100:13
      path 19D90E68 RPKI State not found
      rx pathid: 0, tx pathid: 0x0
```

Courtesy of SEACOM: http://as37100.net

Using RPKI

- Network operators can make decisions based on RPKI state:
  - Invalid – discard the prefix – several do this now!
  - Not found – let it through (maybe low local preference)
  - Valid – let it through (high local preference)
- Some operators even considering making "not found" a discard event
  - But then Internet IPv4 BGP table would shrink to about 55000 prefixes and the IPv6 BGP table would shrink to about 9600 prefixes!

RPKI Summary

- All AS operators must consider deploying:
  - **Signing ROAs**
  - **Dropping Invalids** (ROV)
- An important step to securing the routing system
- Doesn't secure the path, but that's the next hurdle to cross
- With origin validation, the opportunities for malicious or accidental mis-origination disappear
- FAQ:
  - https://nlnetlabs.nl/projects/rpki/faq/

Routing Security

- Implement the recommendations in https://www.manrs.org/manrs

1. Prevent propagation of incorrect routing information
   - Filter BGP peers, in & out!
2. Prevent traffic with spoofed source addresses
   - BCP38 – Unicast Reverse Path Forwarding
3. Facilitate communication between network operators
   - NOC to NOC Communication
4. Facilitate validation of routing information
   - Route Origin Authorisation using RPKI

Summary

- Deploy RPKI
  - It is in the Internet's best interest
- With wide deployment of RPKI it becomes possible to only allow validated prefix announcements into the Internet Routing System
  - Prevents mis-originations
  - Prevents prefix hijack
  - Makes the Internet infrastructure more reliable and more stable

BGP Origin Validation

ISP Workshops