

# 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/>)

# 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
- 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 robust security framework for verifying the association between resource holder and their Internet resources
- Created to address the issues in RFC 4593 “Generic Threats to Routing Protocols”
- Helps to secure Internet routing by validating routes
  - Proof that prefix announcements are coming from the legitimate holder of the resource

# 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)

---

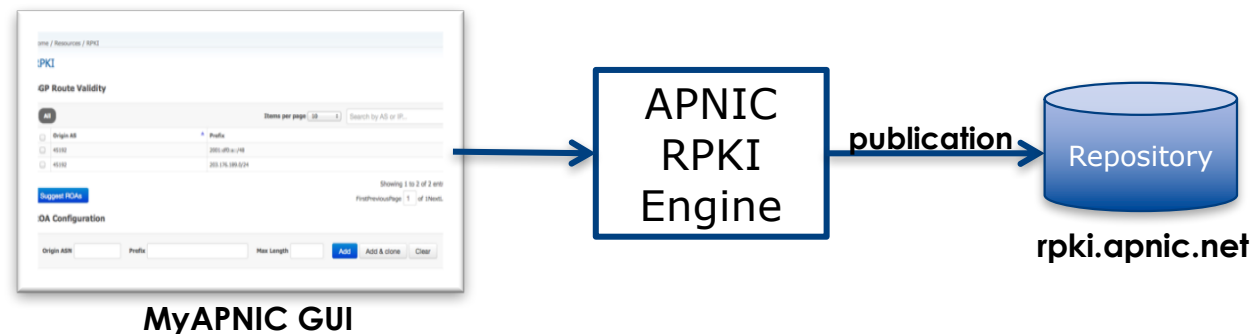
- ❑ Extension to BGP that provides improved security for BGP routing
- ❑ Being worked on by the SIDR Working Group at IETF
- ❑ Implemented via a new optional non-transitive BGP attribute that contains a digital signature
- ❑ Two components:
  - BGP Prefix Origin Validation (using RPKI)
  - BGP Path Validation



# Issuing Party

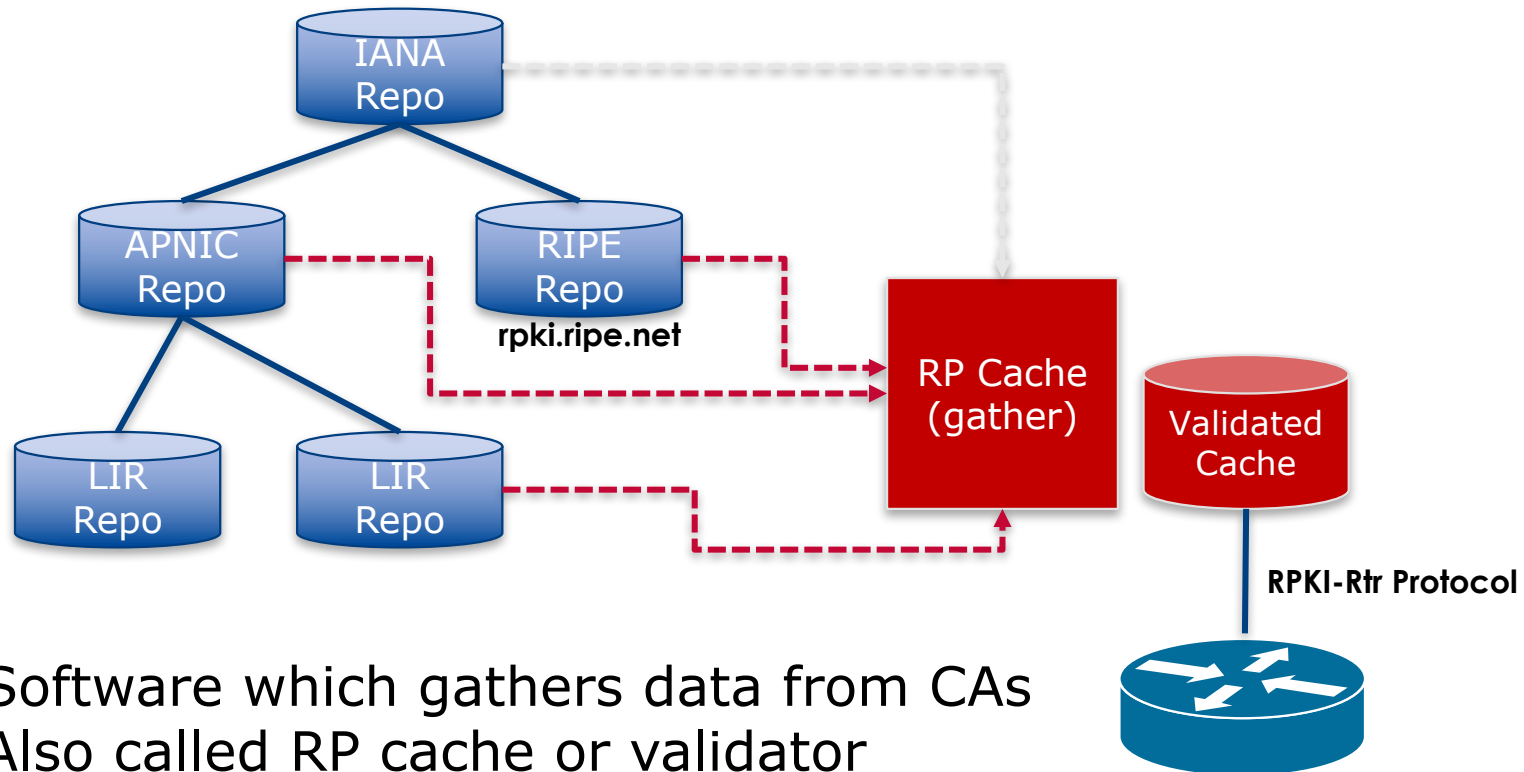
---

- ❑ 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



Courtesy of APNIC: <https://apnic.net>

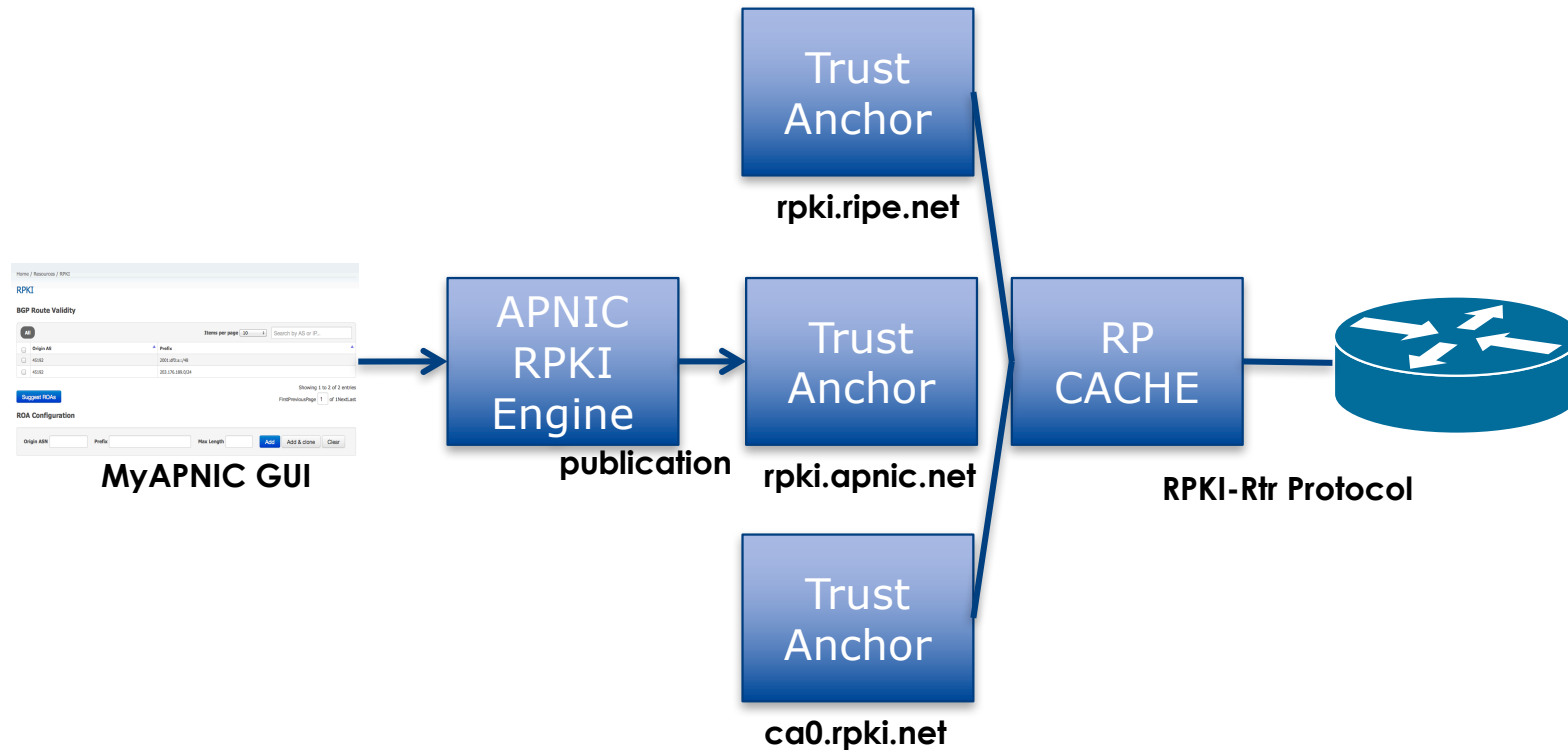
# Relying Party (RP)



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 Authorization (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 authorize an AS Number to originate one or more specific route advertisements
- ❑ Publish a ROA using MyAPNIC

# Router Origin Validation

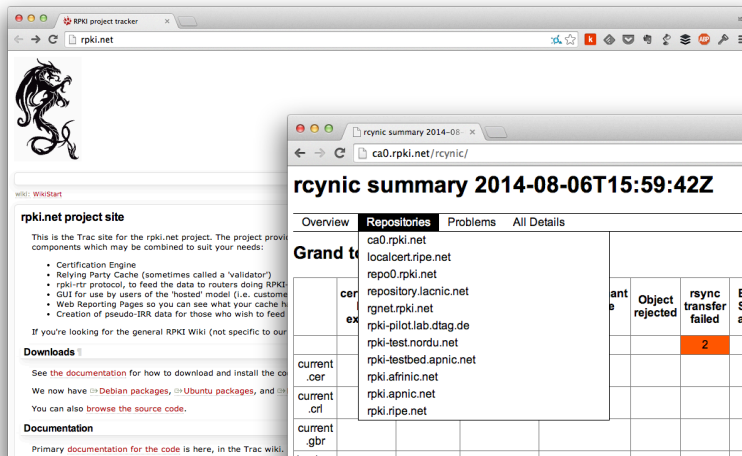
---

- ❑ Router must support RPKI
- ❑ Checks an RP cache / validator
- ❑ Validation returns 3 states:
  - Valid = when authorization is found for prefix X
  - Invalid = when authorization is found for prefix X but not from ASN Y
  - Unknown = when no authorization data is found
- ❑ 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 – newly available – release TBA

# Build an RP Cache

- 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



ca0.rpki.net/rcync/

rcync summary 2014-08-06T15:59:42Z

		Overview	Repositories	Problems	All Details													
Grand total	cert	ex	ant	Object rejected	rsync transfer failed	Bad CMS 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	
			ca0.rpki.net															
			localcert.ripe.net															
			repo0.rpki.net															
			repository.lacnic.net															
			rgnet.rpki.net															
			rpki-pilot.lab.dtag.de															
			rpki-testbed.nordu.net		2												562	
			rpki-testbed.apnic.net				1	1	1		5		67	67	1	3344		
			rpki.apnic.net									71				3336		
			rpki.ripe.net															
			current.gbr								5		3	3		6		
			backup.mnt				1									1		
			current.mnt	2	2	1				1		69	71		1	3333		
			current.roa			1					67		48	48		2551		
			<b>Total</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>77</b>	<b>140</b>	<b>189</b>	<b>118</b>	<b>2</b>	<b>12571</b>	<b>562</b>	

# Configure Router to Use Cache

---

- Point router to the local RPKI cache
  - Server listens on port 43779
  - Cisco IOS example:

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

# Some 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



# 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
Minimum 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)

---

21808 BGP sovc network entries using 1919104 bytes of memory  
22632 BGP sovc record entries using 452640 bytes of memory

Network	Maxlen	Origin-AS	Source	Neighbor
1.9.0.0/16	24	4788	0	105.16.112.2/43779
1.9.12.0/24	24	4788	0	105.16.112.2/43779
1.9.12.0/24	24	65037	0	105.16.112.2/43779
1.9.21.0/24	24	4788	0	105.16.112.2/43779
1.9.21.0/24	24	24514	0	105.16.112.2/43779
1.9.23.0/24	24	65120	0	105.16.112.2/43779
1.9.31.0/24	24	65077	0	105.16.112.2/43779
1.9.52.0/24	24	4788	0	105.16.112.2/43779
1.9.53.0/24	24	4788	0	105.16.112.2/43779
1.9.54.0/24	24	4788	0	105.16.112.2/43779
1.9.55.0/24	24	4788	0	105.16.112.2/43779
1.9.65.0/24	24	4788	0	105.16.112.2/43779
1.9.65.0/24	24	24514	0	105.16.112.2/43779
1.9.112.0/24	24	4788	0	105.16.112.2/43779
...				

Courtesy of SEACOM: <http://as37100.net>

# RPKI Table (IPv6)

---

3115 BGP sovc network entries using 348880 bytes of memory  
3249 BGP sovc record entries using 64980 bytes of memory

Network	Maxlen	Origin-AS	Source	Neighbor
2001:240::/32	32	2497	0	2C0F:FEB0:B:1::2/43779
2001:348::/32	64	7679	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
2001:500:30::/48	48	10745	0	2C0F:FEB0:B:1::2/43779
2001:500:31::/48	48	393220	0	2C0F:FEB0:B:1::2/43779
2001:500:F0::/48	48	53535	0	2C0F:FEB0:B:1::2/43779
2001:504:32::/48	48	21654	0	2C0F:FEB0:B:1::2/43779
2001:608::/32	32	5539	0	2C0F:FEB0:B:1::2/43779
2001:610::/32	48	1103	0	2C0F:FEB0:B:1::2/43779
2001:610:240::/42	42	3333	0	2C0F:FEB0:B:1::2/43779
2001:620::/32	32	559	0	2C0F:FEB0:B:1::2/43779
2001:620::/29	29	559	0	2C0F:FEB0:B:1::2/43779
2001:630::/32	48	786	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
  - 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 20k prefixes and the IPv6 BGP table would shrink to about 3k prefixes!

# RPKI Summary

---

- All AS operators should consider deploying
- An important step to securing the routing system
  - Origin validation
- 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

# Routing Security

---

- Implement the recommendations in <https://www.routingmanifesto.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