Network Working Group
Internet-Draft
Intended status: Best Current Practice
Expires: May 25, 2014

R. Bush
Internet Initiative Japan

November 21, 2013

RPKI-Based Origin Validation Operation
draft-ietf-sidr-origin-ops-23

Abstract

Deployment of RPKI-based BGP origin validation has many operational considerations. This document attempts to collect and present those which are most critical. It is expected to evolve as RPKI-based origin validation continues to be deployed and the dynamics are better understood.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in RFC 2119 [RFC2119] only when they appear in all upper case. They may also appear in lower or mixed case as English words, without normative meaning.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 25, 2014.

Copyright Notice

Table of Contents

1. Introduction

RPKI-based origin validation relies on widespread deployment of the Resource Public Key Infrastructure (RPKI) [RFC6480]. How the RPKI is distributed and maintained globally is a serious concern from many aspects.

While the global RPKI is in the early stages of deployment, there is no single root trust anchor, initial testing is being done by the RIRs, and there are technical testbeds. It is thought that origin validation based on the RPKI will continue to be deployed incrementally over the next few years. It is assumed that eventually there must be a single root trust anchor for the public address space, see [iab].

Origin validation needs to be done only by an AS's border routers and is designed so that it can be used to protect announcements which are originated by any network participating in Internet BGP routing: large providers, upstreams and down-streams, and by small stub/ enterprise/edge routers.

Origin validation has been designed to be deployed on current routers without significant hardware upgrade. It should be used in border routers by operators from large backbones to small stub/entetprise/ edge networks.

RPKI-based origin validation has been designed so that, with prudent local routing policies, there is little risk that what is seen as today's normal Internet routing is threatened by imprudent deployment of the global RPKI, see Section 5.

2. Suggested Reading

It is assumed that the reader understands BGP, [RFC4271], the RPKI, see [RFC6480], the RPKI Repository Structure, see [RFC6481], Route Origin Authorizations (ROAs), see [RFC6482], the RPKI to Router Protocol, see [RFC6810], RPKI-based Prefix Validation, see [RFC6811], and Ghostbusters Records, see [RFC6493].

3. RPKI Distribution and Maintenance

The RPKI is a distributed database containing certificates, Certificate Revocation Lists (CRLs), manifests, ROAs, and Ghostbusters Records as described in [RFC6481]. Policies and considerations for RPKI object generation and maintenance are discussed elsewhere.

The RPKI repository design [RFC6481] anticipated a hierarchic organization of repositories, as this seriously improves the performance of relying parties gathering data over a non-hierarchic organization. Publishing parties MUST implement hierarchic directory structures.

A local relying party valid cache containing all RPKI data may be gathered from the global distributed database using the rsync protocol, [RFC5781], and a validation tool such as rcynic [rcynic].

A validated cache contains all RPKI objects that the RP has verified to be valid according to the rules for validation RPKI certificates and signed objects, see [RFC6487] and [RFC6488]. Entities that trust the cache can use these RPKI objects without further validation.

Validated caches may also be created and maintained from other validated caches. Network operators SHOULD take maximum advantage of this feature to minimize load on the global distributed RPKI database. Of course, the recipient relying parties should re-validate the data.

As Trust Anchor Locators (TALs), see [RFC6490], are critical to the RPKI trust model, operators should be very careful in their initial selection and vigilant in their maintenance.

Timing of inter-cache synchronization, and synchronization between caches and the global RPKI, is outside the scope of this document, and depends on things such as how often routers feed from the caches, how often the operator feels the global RPKI changes significantly, etc.

As inter-cache synchronization within an operator's network does not impact global RPKI resources, an operator may choose to synchronize quite frequently.

To relieve routers of the load of performing certificate validation, cryptographic operations, etc., the RPKI-Router protocol, [RFC6810], does not provide object-based security to the router. I.e. the router can not validate the data cryptographically from a well-known trust anchor. The router trusts the cache to provide correct data and relies on transport based security for the data received from the cache. Therefore the authenticity and integrity of the data from the cache should be well protected, see Section 7 of [RFC6810].

As RPKI-based origin validation relies on the availability of RPKI data, operators SHOULD locate RPKI caches close to routers that require these data and services in order to minimize the impact of likely failures in local routing, intermediate devices, long circuits, etc. One should also consider trust boundaries, routing bootstrap reachability, etc.

For example, a router should bootstrap from a chache which is reachable with minimal reliance on other infrastructure such as DNS or routing protocols. If a router needs its BGP and/or IGP to converge for the router to reach a cache, once a cache is reachable, the router will then have to reevaluate prefixes already learned via BGP. Such configurations should be avoided if reasonably possible.

If insecure transports are used between an operator's cache and their router(s), the Transport Security recommendations in [RFC6810] SHOULD be followed. In particular, operators MUST NOT use insecure transports between their routers and RPKI caches located in other Autonomous Systems.

For redundancy, a router should peer with more than one cache at the same time. Peering with two or more, at least one local and others remote, is recommended.

If an operator trusts upstreams to carry their traffic, they may also trust the RPKI data those upstreams cache, and SHOULD peer with caches made available to them by those upstreams. Note that this places an obligation on those upstreams to maintain fresh and reliable caches, and to make them available to their customers. And, as usual, the recipient SHOULD re-validate the data.

A transit provider or a network with peers SHOULD validate origins in announcements made by upstreams, down-streams, and peers. They still should trust the caches provided by their upstreams.

Before issuing a ROA for a super-block, an operator MUST ensure that all sub-allocations from that block which are announced by other ASs, e.g. customers, have correct ROAs in the RPKI. Otherwise, issuing a ROA for the super-block will cause the announcements of sub-allocations with no ROAs to be viewed as Invalid, see [RFC6811]. While waiting for all sub-allocatees to register ROAs, the owner of the super-block may use live BGP data to populate ROAs as a proxy, and then safely issue a ROA for the super-block.

Use of RPKI-based origin validation removes any need to originate more specifics into BGP to protect against mis-origination of a less specific prefix. Having a ROA for the covering prefix will protect it.

To aid translation of ROAs into efficient search algorithms in routers, ROAs should be as precise as possible, i.e. match prefixes as announced in BGP. E.g. software and operators SHOULD avoid use of excessive max length values in ROAs unless operationally necessary.

One advantage of minimal ROA length is that the forged origin attack does not work for sub-prefixes that are not covered by overly long max length. E.g. if, instead of 10.0.0.0/16-24, one issues 10.0.0.0/16 and 10.0.42.0/24, a forged origin attack can not succeed against 10.0.666.0/24. They must attack the whole /16, which is more likely to be noticed because of its size.

Therefore, ROA generation software MUST use the prefix length as the max length if the user does not specify a max length.

RFC EDITOR PLEASE REMOVE THIS PARAGRAPH: The above example does not use a standard documentation prefix as it needs a /16 so that a /24 can hole punch. As anything longer than a /24 is not globally routed, a /24 with a /25 (or whatever) hole would not be realistic and the ops reader would spend their energy on that anomaly instead of the example.

Operators should be conservative in use of max length in ROAs. E.g., if a prefix will have only a few sub-prefixes announced, multiple ROAs for the specific announcements should be used as opposed to one ROA with a long max length.

Operators owning prefix P should issue ROAs for all ASs which may announce P. If a prefix is legitimately announced by more than one AS, ROAs for all of the ASs SHOULD be issued so that all are considered Valid.

In an environment where private address space is announced in eBGP the operator may have private RPKI objects which cover these private spaces. This will require a trust anchor created and owned by that environment, see [I-D.ietf-sidr-ltamgmt].

Operators issuing ROAs may have customers which announce their own prefixes and ASs into global eBGP but who do not wish to go though the work to manage the relevant certificates and ROAs. Operators SHOULD offer to provision the RPKI data for these customers just as they provision many other things for them.

While an operator using RPKI data MAY choose any polling frequency they wish for ensuring they have a fresh RPKI cache. However, if they use RPKI data as an input to operational routing decisions, they SHOULD ensure local caches inside their AS are synchronized with each other at least every four to six hours.

Operators should use tools which warn them of any impending ROA or certificate expiry which could affect the validity of their own data. Ghostbuster Records, see [RFC6493], can be used to facilitate contact with upstream CAs to effect repair.

4. Within a Network

Origin validation need only be done by edge routers in a network, those which border other networks/ASs.

A validating router will use the result of origin validation to influence local policy within its network, see Section 5. In deployment this policy should fit into the AS's existing policy, preferences, etc. This allows a network to incrementally deploy validation-capable border routers.

The operator should be aware that RPKI-based origin validation, as any other policy change, can cause traffic shifts in their network. And, as with normal policy shift practice, a prudent operator has tools and methods to predict, measure, modify, etc.

5. Routing Policy

Origin validation based on the RPKI marks a received announcement as having an origin which is Valid, NotFound, or Invalid, see [RFC6811]. How this is used in routing should be specified by the operator's local policy.

Local policy using relative preference is suggested to manage the uncertainty associated with a system in early deployment, applying local policy to eliminate the threat of unreachability of prefixes due to ill-advised certification policies and/or incorrect certification data. E.g. until the community feels comfortable relying on RPKI data, routing on Invalid origin validity, though at a low preference, MAY occur.

Operators should be aware that accepting Invalid announcements, no matter how de-preffed, will often be the equivalent of treating them as fully Valid. Consider having a ROA for AS 42 for prefix 10.0.0.0/16-24. A BGP announcement for 10.0.666.0/24 from AS 666 would be Invalid. But if policy is not configured to discard it, then longest match forwarding will send packets toward AS 666 no matter the value of local preference.

As origin validation will be rolled out incrementally, coverage will be incomplete for a long time. Therefore, routing on NotFound validity state SHOULD be done for a long time. As the transition moves forward, the number of BGP announcements with validation state NotFound should decrease. Hence an operator's policy should not be overly strict, and should prefer Valid announcements, attaching a lower preference to, but still using, NotFound announcements, and dropping or giving a very low preference to Invalid announcements. Merely de-preffing Invalids is ill-advised, see previous paragraph.

Some providers may choose to set Local-Preference based on the RPKI validation result. Other providers may not want the RPKI validation result to be more important than AS-path length -- these providers would need to map RPKI validation result to some BGP attribute that is evaluated in BGP's path selection process after AS-path is evaluated. Routers implementing RPKI-based origin validation MUST provide such options to operators.

Local-Preference may be used to carry both the validity state of a prefix along with its traffic engineering (TE) characteristic(s). It is likely that an operator already using Local-Preference will have to change policy so they can encode these two separate characteristics in the same BGP attribute without negative impact or opening privilege escalation attacks. E.g. do not encode validation state in higher bits than used for TE.

When using a metric which is also influenced by other local policy, an operator should be careful not to create privilege upgrade vulnerabilities. E.g. if Local Pref is set depending on validity state, be careful that peer community signaling SHOULD NOT upgrade an Invalid announcement to Valid or better.

Announcements with Valid origins should be preferred over those with NotFound or Invalid origins, if Invalid origins are accepted at all.

Announcements with NotFound origins should be preferred over those with Invalid origins.

Announcements with Invalid origins SHOULD NOT be used, but may be used to meet special operational needs. In such circumstances, the announcement should have a lower preference than that given to Valid or NotFound.

When first deploying origin validation, it may be prudent to not drop announcements with Invalid orgins until inspection of logs, SNMP, or other data indicate that the correct result would be obtained.

Validity state signaling SHOULD NOT be accepted from a neighbor AS. The validity state of a received announcement has only local scope due to issues such as scope of trust, RPKI synchrony, and [I-D.ietf-sidr-ltamgmt].

6. Notes and Recommendations

Like the DNS, the global RPKI presents only a loosely consistent view, depending on timing, updating, fetching, etc. Thus, one cache or router may have different data about a particular prefix than another cache or router. There is no 'fix' for this, it is the nature of distributed data with distributed caches.

Operators should beware that RPKI caches are loosely synchronized, even within a single AS. Thus, changes to the validity state of prefixes could be different within an operator's network. In addition, there is no guaranteed interval from when an RPKI cache is updated to when that new information may be pushed or pulled into a set of routers via this protocol. This may result in sudden shifts of traffic in the operator's network, until all of the routers in the AS have reached equilibrium with the validity state of prefixes reflected in all of the RPKI caches.

It is hoped that testing and deployment will produce advice on relying party cache loading and timing.

There is some uncertainty about the origin AS of aggregates and what, if any, ROA can be used. The long range solution to this is the deprecation of AS-SETs, see [RFC6472].

As reliable access to the global RPKI and an operator's caches (and possibly other hosts, e.g. DNS root servers) is important, an operator should take advantage of relying party tools which report changes in BGP or RPKI data which would negatively affect validation of such prefixes.

Operators should be aware that there is a trade-off in placement of an RPKI repository in address space for which the repository's content is authoritative. On one hand, an operator will wish to maximize control over the repository. On the other hand, if there are reachability problems to the address space, changes in the repository to correct them may not be easily accessed by others.

Operators who manage certificates should associate RPKI Ghostbusters Records (see [RFC6493]) with each publication point they control. These are publication points holding the CRL, ROAs, and other signed objects issued by the operator, and made available to other ASs in support of routing on the public Internet.

Routers which perform RPKI-based origin validation must support Four-octet AS Numbers (see [RFC6793]), as, among other things, it is not reasonable to generate ROAs for AS 23456.

Software which produces filter lists or other control forms for routers where the target router does not support Four-octet AS Numbers (see [RFC6793]) must be prepared to accept Four-octet AS Numbers and generate the appropriate two-octet output.

As a router must evaluate certificates and ROAs which are time dependent, routers' clocks MUST be correct to a tolerance of approximately an hour.

Servers should provide time service, such as [RFC5905], to client routers.

7. Security Considerations

As the BGP origin AS of an update is not signed, origin validation is open to malicious spoofing. Therefore, RPKI-based origin validation is expected to deal only with inadvertent mis-advertisement.

Origin validation does not address the problem of AS-Path validation. Therefore paths are open to manipulation, either malicious or accidental.

As BGP does not ensure that traffic will flow via the paths it advertises, the data plane may not follow the control plane.

Be aware of the class of privilege escalation issues discussed in Section 5 above.

8. IANA Considerations

This document has no IANA Considerations.

9. Acknowledgments

The author wishes to thank Shane Amante, Rob Austein, Steve Bellovin, Jay Borkenhagen, Wes George, Seiichi Kawamura, Steve Kent, Pradosh Mohapatra, Chris Morrow, Sandy Murphy, Eric Osterweil, Keyur Patel, Heather and Jason Schiller, John Scudder, Kotikalapudi Sriram, Maureen Stillman, and Dave Ward.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC6481] Huston, G., Loomans, R., and G. Michaelson, "A Profile for Resource Certificate Repository Structure", RFC 6481, February 2012.

[RFC6482] Lepinski, M., Kent, S., and D. Kong, "A Profile for Route Origin Authorizations (ROAs)", RFC 6482, February 2012.

[RFC6490] Huston, G., Weiler, S., Michaelson, G., and S. Kent, "Resource Public Key Infrastructure (RPKI) Trust Anchor Locator", RFC 6490, February 2012.

[RFC6493] Bush, R., "The Resource Public Key Infrastructure (RPKI) Ghostbusters Record", RFC 6493, February 2012.

[RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet Autonomous System (AS) Number Space", RFC 6793, December 2012.

[RFC6810] Bush, R. and R. Austein, "The Resource Public Key Infrastructure (RPKI) to Router Protocol", RFC 6810, January 2013.

[RFC6811] Mohapatra, P., Scudder, J., Ward, D., Bush, R., and R. Austein, "BGP Prefix Origin Validation", RFC 6811, January 2013.

10.2. Informative References

[I-D.ietf-sidr-ltamgmt]
Reynolds, M., Kent, S., and M. Lepinski, "Local Trust Anchor Management for the Resource Public Key Infrastructure", draft-ietf-sidr-ltamgmt-08 (work in progress), April 2013.

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[RFC5781] Weiler, S., Ward, D., and R. Housley, "The rsync URI Scheme", RFC 5781, February 2010.

[RFC5905] Mills, D., Martin, J., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, June 2010.

[RFC6472] Kumari, W. and K. Sriram, "Recommendation for Not Using AS_SET and AS_CONFED_SET in BGP", BCP 172, RFC 6472, December 2011.

[RFC6480] Lepinski, M. and S. Kent, "An Infrastructure to Support Secure Internet Routing", RFC 6480, February 2012.

[RFC6487] Huston, G., Michaelson, G., and R. Loomans, "A Profile for X.509 PKIX Resource Certificates", RFC 6487, February 2012.

[RFC6488] Lepinski, M., Chi, A., and S. Kent, "Signed Object Template for the Resource Public Key Infrastructure (RPKI)", RFC 6488, February 2012.

[iab] , "IAB statement on the RPKI", , <http://www.iab.org/documents/correspondence-reports-documents/docs2010/iab-statement-on-the-rpki/>.

[rcynic] , "rcynic read-me", , <http://rpki.net/rcynic>.

Author's Address

Randy Bush
Internet Initiative Japan
5147 Crystal Springs
Bainbridge Island, Washington 98110
US

Email: randy@psg.com