Network Working Group
Internet-Draft
Intended status: BCP
Expires: May 17, 2012

R. Bush
Internet Initiative Japan
November 14, 2011

RPKI-Based Origin Validation Operation
draft-ietf-sidr-origin-ops-13

Abstract

Deployment of RPKI-based BGP origin validation has many operational considerations. This document attempts to collect and present them. It is expected to evolve as RPKI-based origin validation is 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" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Table of Contents

1. Introduction

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

The global RPKI is in very initial stages of deployment, there is no single root trust anchor, initial testing is being done by the IANA and the RIRs, and there is a technical testbed. It is thought that origin validation based on the RPKI will be deployed incrementally over the next year to five years.

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 [I-D.ietf-sidr-arch], the RPKI Repository Structure, see [I-D.ietf-sidr-repos-struct], ROAs, see [I-D.ietf-sidr-roa-format], the RPKI to Router Protocol, see [I-D.ietf-sidr-rpki-rtr], RPKI-based Prefix Validation, see [I-D.ietf-sidr-pfx-validate], and Ghostbusters Records, see [I-D.ietf-sidr-ghostbusters].

3. RPKI Distribution and Maintenance

The RPKI is a distributed database containing certificates, CRLs, manifests, ROAs, and Ghostbusters Records as described in [I-D.ietf-sidr-repos-struct]. Policies and considerations for RPKI object generation and maintenance are discussed elsewhere.

A local 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].

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 SHOULD re-validate the data.

Timing of inter-cache synchronization is outside the scope of this document, but 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 does not impact global RPKI resources, an operator MAY choose to synchronize quite frequently.

As RPKI-based origin validation relies on the availability of RPKI data, operators SHOULD locate caches close to routers that require these data and services. 'Close' is, of course, complex. One should consider trust boundaries, routing bootstrap reachability, latency, etc.

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 any 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 [I-D.ietf-sidr-pfx-validate].

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 should 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 10.0.66.0/24. 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.

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.

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.

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 owning prefix P should issue ROAs for all ASs which may announce P.

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 cache freshness at least every four to six hours.

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.

eBGP speakers which face more critical peers or up/down-streams are candidates for the earliest deployment. Validating more critical received announcements should be considered in partial deployment.

5. Routing Policy

Origin validation based on the RPKI marks a received announcement as having an origin which is Valid, NotFound, or Invalid. See [I-D.ietf-sidr-pfx-validate]. 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 unroutability 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.

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, preferring Valid announcements, attaching a lower preference to, but still using, NotFound announcements, and dropping or giving very low preference to Invalid announcements.

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 it's traffic engineering 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 negatively impact or opening privilege escalation attacks.

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 MAY NOT upgrade an Invalid announcement to Valid or better.

Announcements with Valid origins SHOULD be preferred over those with NotFound or Invalid origins, if the latter 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.

Validity state signialing 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

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.

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 [I-D.wkumari-deprecate-as-sets].

Operators who manage certificates SHOULD associate RPKI Ghostbusters Records (see [I-D.ietf-sidr-ghostbusters]) 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.

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 designed 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, Steve Kent, Pradosh Mohapatra, Chris Morrow, Sandy Murphy, Keyur Patel, Heather and Jason Schiller, John Scudder, Kotikalapudi Sriram, Maureen Stillman, and Dave Ward.

10. References

10.1. Normative References

[I-D.ietf-sidr-arch]
Lepinski, M. and S. Kent, "An Infrastructure to Support Secure Internet Routing", draft-ietf-sidr-arch-13 (work in progress), May 2011.

[I-D.ietf-sidr-ghostbusters]
Bush, R., "The RPKI Ghostbusters Record", draft-ietf-sidr-ghostbusters-15 (work in progress), October 2011.

[I-D.ietf-sidr-ltamgmt]
Reynolds, M. and S. Kent, "Local Trust Anchor Management for the Resource Public Key Infrastructure", draft-ietf-sidr-ltamgmt-02 (work in progress), June 2011.

[I-D.ietf-sidr-pfx-validate]
Mohapatra, P., Scudder, J., Ward, D., Bush, R., and R. Austein, "BGP Prefix Origin Validation", draft-ietf-sidr-pfx-validate-03 (work in progress), October 2011.

[I-D.ietf-sidr-repos-struct]
Huston, G., Loomans, R., and G. Michaelson, "A Profile for Resource Certificate Repository Structure", draft-ietf-sidr-repos-struct-09 (work in progress), July 2011.

[I-D.ietf-sidr-roa-format]
Lepinski, M., Kent, S., and D. Kong, "A Profile for Route Origin Authorizations (ROAs)", draft-ietf-sidr-roa-format-12 (work in progress), May 2011.

[I-D.ietf-sidr-rpki-rtr]
Bush, R. and R. Austein, "The RPKI/Router Protocol", draft-ietf-sidr-rpki-rtr-19 (work in progress), October 2011.

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

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

10.2. Informative References

[I-D.wkumari-deprecate-as-sets]
Kumari, W., "Deprecation of BGP AS_SET, AS_CONFED_SET.", draft-wkumari-deprecate-as-sets-01 (work in progress), September 2010.

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

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

Author's Address

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

Phone: +1 206 780 0431 x1
Email: randy@psg.com