Network Working Group
Internet-Draft
Intended status: BCP
Expires: September 16, 2011

R. Bush
Internet Initiative Japan

March 15, 2011

BGPsec Operational Considerations
draft-ymbk-bgpsec-ops-01

Abstract

Deployment of the BGPsec architecture and protocols has many operational considerations. This document attempts to collect and present them. It is expected to evolve as BGPsec is formalized and initially deployed.

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].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79. This document may not be modified, and derivative works of it may not be created, and it may not be published except as an Internet-Draft.

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 September 16, 2011.

Table of Contents

1. Introduction 3
2. Suggested Reading 3
3. RPKI Distribution and Maintenance 3
4. AS/Router Certificates 4
5. Within a Network 4
6. Considerations for Edge Sites 5
7. Beaconing Considerations 5
8. Routing Policy 6
9. Notes 7
10. Security Considerations 7
11. IANA Considerations 7
12. Acknowledgments 8
13. References 8
13.1. Normative References 8
13.2. Informative References 8
Author's Address 9

1. Introduction

BGPsec is a new protocol with many operational considerations. It is expected to be deployed incrementally over a number of years. As core BGPsec-capable routers may require large memory and crypto assist, it is thought that origin validation based on the RPKI will occur over the next two to five years and that BGPsec will start to deploy late in that window.

BGPsec relies on widespread propagation of the Resource Public Key Infrastructure (RPKI) [I-D.ietf-sidr-arch]. How the RPKI is distributed and maintained globally and within an operator's infrastructure may be different for BGPsec than for origin validation.

BGPsec need be spoken only by a AS's eBGP speaking, AKA border, routers, and is designed so that it can be used to protect announcements which are originated by small edge routers, and this has special operational considerations.

Different prefixes have different timing and replay protection considerations.

2. Suggested Reading

It is assumed that the reader understands BGP, [RFC4271], BGPsec, [I-D.lepinski-bgpsec-overview], the RPKI, see [I-D.ietf-sidr-arch], the RPKI Repository Structure, see [I-D.ietf-sidr-repos-struct], and ROAs, see [I-D.ietf-sidr-roa-format].

3. RPKI Distribution and Maintenance

The RPKI is a distributed database containing certificates, CRLs, manifests, ROAs, and Ghostbuster 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 and a validation tool such as 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.

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. A router can peer with one or more nearby caches.

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 SHOULD also trust the RPKI data those upstreams cache, and SHOULD peer with those caches. Note that this places an obligation on those upstreams to maintain fresh and reliable caches.

A transit provider or a network with peers SHOULD validate NLRI in announcements made by upstreams, downstreams, and peers. They still SHOULD trust the caches provided by their upstreams.

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].

4. AS/Router Certificates

A site/operator MAY use a single certificate/key in all their routers, one certificate/key per router, or any granularity in between.

A large operator, concerned that a compromise of one router's key would make many routers vulnerable, MAY accept a more complex certificate/key distribution burden to reduce this exposure.

On the other extreme, an edge site with one or two routers MAY use a single certificate/key.

Routers MAY be capable of generating their own keys and having their certificates signed and published in the RPKI by their NOC. This would mean that a router's private key need never leave the router.

5. Within a Network

BGPsec is spoken by edge routers in a network, those which border other networks/ASs.

In a fully BGPsec enabled AS, Route Reflectors MUST have BGPsec

enabled if and only if there are eBGP speakers in their client cone.

A BGPsec capable router MAY use the data it receives to influence local policy within its network, see Section 8. In deployment this policy should fit into the AS's existing policy, preferences, etc. This allows a network to incrementally deploy BGPsec capable border routers.

eBGP speakers which face more critical peers or up/downstreams would be candidates for the earliest deployment. Both securing one's own announcements and validating received announcements should be considered in partial deployment.

An eBGP listener MUST NOT trust non-BGPsec markings such as communities received across a trust boundary.

6. Considerations for Edge Sites

An edge site which does not provide transit and trusts its upstream(s) SHOULD only originate a signed prefix announcement and need not validate received announcements.

BGPsec protocol capability negotiation provides for a speaker signing the data it sends but being unable to accept signed data. Thus a smallish edge router may hold only its own signing key(s) and sign it's announcement but not receive signed announcements and therefore not need to deal with the majority of the RPKI.

As the vast majority (84%) of ASs are stubs, and they announce the majority of prefixes, this allows for simpler and cheaper early incremental deployment. It may also mean that edge sites concerned with routing security will be attracted to upstreams which support BGPsec.

7. Beaconing Considerations

The BGPsec protocol attempts to reduce exposure to replay attacks by allowing the route originator to sign an announcement with a validity period and re-announce well within that period.

This re-announcement is termed 'beaconing'. All timing values are, of course, jittered.

It is only the originator of an NLRI which signs the announcement with a validity period.

To reduce vulnerability to a lost beacon announcement, a router SHOULD beacon at a rate somewhat greater than half the signature validity period it uses.

As beaconing places a load on the entire global routing system, careful thought MUST be given to any need to beacon frequently. This would be based on a conservative estimation of the vulnerability to a replay attack.

Beacon timing and signature validity periods SHOULD be as follows:

The Exemplary Citizen: Prefix originators who are not overly concerned about replay attacks might announce with a signature validity of multiple weeks and beacon one third of the validity period.

Normal Prefix: Most prefixes SHOULD announce with a signature validity of a week and beacon every three days.

Critical Prefix: Of course, we all think what we do is critical. But prefixes of top level DNS servers, and RPKI publication points are actually critical to large swaths of the Internet and are therefore tempting targets for replay attacks. It is suggested that the beaconing of these prefixes SHOULD be two to four hours, with a signature validity of six to twelve hours.

Note that this may incur route flap damping (RFD) with current default but deprecated RFD parameters, see [I-D.ymbk-rfd-usable].

8. Routing Policy

Unlike origin validation based on the RPKI, BGPsec marks a received announcement as Valid or Invalid, there is no NotFound state. How this is used in routing is up to the operator's local policy. See [I-D.pmohapat-sidr-pfx-validate].

As BGPsec will be rolled out over years and does not allow for intermediate non-signing edge routers, coverage will be spotty for a long time. Hence a normal operator's policy SHOULD NOT be overly strict, perhaps preferring valid announcements and giving very low preference, but still using, invalid announcements.

Local policy on the eBGP edge MAY convey the validation status of a BGP signed path through various pre-existing mechanisms, e.g. setting a BGP community, or modifying a metric value such as local-preference or MED. Some MAY choose to use the large Local-Pref hammer. Others MAY choose to let AS-Path rule and set their internal metric, which

comes after AS-Path in the BGP decision process.

A BGPsec speaker validates signed paths at the eBGP edge.

Because of possible RPKI version skew, an AS Path which does not validate at router R0 might validate at R1. Therefore, signed paths that can not be validated SHOULD have their signatures kept intact and should be signed when sent to external BGPsec speakers.

This implies that AS Paths with non-validated signatures MAY be propagated to iBGP peers. Therefore, unless local policy ensures otherwise, a signed path learned via iBGP MAY NOT have been validated. If needed, the validation state SHOULD be signaled by normal policy mechanisms such as communities or metrics.

On the other hand, local policy on the eBGP edge might preclude iBGP or eBGP announcement of signed AS Paths which are not validated.

If a BGPsec speaker receives an unsigned path, it SHOULD perform origin validation per [I-D.pmohapat-sidr-pfx-validate].

9. 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.

Operators which manage certificates SHOULD have RPKI Ghostbuster Records (see [I-D.ietf-sidr-ghostbusters]), signed indirectly by End Entity certificates, for those certificates on which others' routing depends for certificate and/or ROA validation.

10. Security Considerations

BGPsec is all about security, routing security. The major security considerations for the protocol are described in [BGPsec].

11. IANA Considerations

This document has no IANA Considerations.

12. Acknowledgments

The author wishes to thank the entire BGPsec foundation team.

13. References

13.1. Normative References

[I-D.ietf-sidr-ghostbusters]
Bush, R., "The RPKI Ghostbusters Record", draft-ietf-sidr-ghostbusters-02 (work in progress), March 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-10 (work in progress), February 2011.

[I-D.lepinski-bgpsec-overview]
Lepinski, M. and S. Turner, "An Overview of BGPSEC", draft-lepinski-bgpsec-overview-00 (work in progress), March 2011.

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

13.2. Informative References

[I-D.ietf-sidr-arch]
Lepinski, M. and S. Kent, "An Infrastructure to Support Secure Internet Routing", draft-ietf-sidr-arch-12 (work in progress), February 2011.

[I-D.ietf-sidr-ltamgmt]
Kent, S. and M. Reynolds, "Local Trust Anchor Management for the Resource Public Key Infrastructure", draft-ietf-sidr-ltamgmt-00 (work in progress), November 2010.

[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-07 (work in progress), February 2011.

[I-D.pmohapat-sidr-pfx-validate]

Mohapatra, P., Scudder, J., Ward, D., Bush, R., and R. Austein, "BGP Prefix Origin Validation", draft-pmohapat-sidr-pfx-validate-07 (work in progress), April 2010.

[I-D.ymbk-rfd-usable] Pelsser, C., Bush, R., Patel, K., Mohapatra, P., and O. Maennel, "Making Route Flap Damping Usable", draft-ymbk-rfd-usable-00 (work in progress), March 2011.

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

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