Network Working Group
Internet-Draft
Intended status: Best Current Practice
Expires: June 17, 2016

R. Bush
Internet Initiative Japan

December 15, 2015

BGPsec Operational Considerations
draft-ietf-sidr-bgpsec-ops-07

Abstract

Deployment of the BGPsec architecture and protocols has many operational considerations. This document attempts to collect and present the most critical and universal. 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" 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 June 17, 2016.

Table of Contents

1. Introduction

BGPsec, [I-D.ietf-sidr-bgpsec-overview], 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/or modern CPUs, it is thought that origin validation based on the RPKI, [RFC6811], will occur over the next twp to three years and that BGPsec will start to deploy well after that.

BGPsec relies on widespread propagation of the Resource Public Key Infrastructure (RPKI) [RFC6480]. 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 an 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. This has special operational considerations.

Different prefixes may have different timing and replay protection considerations.

2. Suggested Reading

It is assumed that the reader understands BGP, [RFC4271], BGPsec, [I-D.ietf-sidr-bgpsec-overview], the RPKI, see [RFC6480], the RPKI Repository Structure, see [RFC6481], and ROAs, see [RFC6482].

3. RPKI Distribution and Maintenance

All non-ROA considerations in the section on RPKI Distribution and Maintenance of [RFC7115] apply.

4. AS/Router Certificates

As described in [I-D.ietf-sidr-rtr-keying] BGPsec-speaking routers are either capable of generating their own public/private key-pairs and having their certificates signed and published in the RPKI by the RPKI CA system, and/or are given public/private key-pairs by the operator.

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 other 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 choose to use a single certificate/key.

In anticipation of possible key compromise, a prudent operator will pre-provision each router's 'next' key in the RPKI so there is no propagation delay for provisioning the new key.

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, i.e. an RR client or the transitive closure of their customers' customers' customers' ....

A BGPsec capable router MAY use the data it receives to influence local policy within its network, see Section 7. In deployment this policy should fit into the AS's existing policy, preferences, etc.

This allows a network to incrementally deploy BGPsec enabled border routers.

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

The operator should be aware that BGPsec, 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.

On the other hand, an operator wanting to monitor router loading, shifts in traffic, etc. might deploy incrementally while watching those and similar effects.

As they are not formally verifiable, an eBGP listener SHOULD NOT strongly trust unsigned security 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 without being able to accept signed data. Thus a smallish edge router may hold only its own signing key(s), sign it's announcements, but not receive signed announcements and therefore not need to deal with the majority of the RPKI. Thus such routers CPU, RAM, and crypto needs are trivial and additional hardware should not be needed.

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

7. Routing Policy

Unlike origin validation based on the RPKI, BGPsec marks a received announcement as Valid or Invalid, there is no explicit NotFound state. In some sense, an unsigned BGP4 path is the equivalent of NotFound. How this is used in routing is up to the operator's local policy. See [RFC6811].

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 paths and giving very low preference, but still using, Invalid paths.

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 Valid announcement from neighbor V for prefix 10.0.0.0/16 and an Invalid announcement for 10.0.666.0/24 from neighbor I. If local policy on the router is not configured to discard the Invalid announcement from I, then longest match forwarding will send packets to neighbor I no matter the value of local preference.

A BGPsec speaker validates signed paths at the eBGP edge.

Local policy on the eBGP edge MAY convey the validation state of a BGP signed path through normal local policy mechanisms, e.g. setting a BGP community for internal use, 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.

Because of possible RPKI version skew, an AS Path which does not validate at router R0 might validate at R1. Therefore, signed paths that are Invalid and yet propagated (because they are chosen as best path) SHOULD have their signatures kept intact and MUST be signed if sent to external BGPsec speakers.

This implies that updates which a speaker judges to be Invalid MAY be propagated to iBGP peers. Therefore, unless local policy ensures otherwise, a signed path learned via iBGP MAY be Invalid. If needed, the validation state should be signaled by normal local 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 Invalid.

A BGPsec speaker receiving a path SHOULD perform origin validation per [RFC6811] and [RFC7115].

A route server is usually 'transparent', most importantly not inserting its own AS into the AS_Path, to not lengthen the AS hop count and thereby reduce the likelihood of best path selection. See 2.2.2 of [I-D.ietf-idr-ix-bgp-route-server]. A BGPsec-aware route server needs to validate the incoming BGPSEC_Path, and to forward updates which can be validated by clients which know the route

server's AS. The route server uses pCount of zero to not increase the effective AS hop count.

If it is known that a BGPsec neighbor is not a transparent route server, and the router provides a knob to disallow a received pCount (prepend count, zero for transparent route servers) of zero, that knob SHOULD be applied. Routers should default to this knob disallowing pCount 0.

To prevent exposure of the internals of BGP Confederations [RFC5065], a BGPsec speaker which is a Member-AS of a Confederation MUST NOT sign updates sent to another Member-AS of the same Confederation.

8. Notes

For protection from attacks replaying BGP data on the order of a day or longer old, re-keying routers with new keys (previously) provisioned in the RPKI is sufficient. For one approach, see [I-D.ietf-sidr-bgpsec-rollover]

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 or router than another cache or router. There is no 'fix' for this, it is the nature of distributed data with distributed caches.

Operators who manage certificates SHOULD have RPKI GhostBuster Records (see [RFC6493]), signed indirectly by End Entity certificates, for those certificates on which others' routing depends for certificate and/or ROA validation.

Operators should be aware of impending algorithm transitions, which will be rare and slow-paced, see see [RFC6916]. They should work with their vendors to ensure support for new algorithms.

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

If a router has reason to believe its clock is seriously incorrect, e.g. it has a time earlier than 2011, it SHOULD NOT attempt to validate incoming updates. It SHOULD defer validation until it believes it is within reasonable time tolerance.

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

9. Security Considerations

The major security considerations for the BGPsec protocol are described in [I-D.ietf-sidr-bgpsec-protocol].

10. IANA Considerations

This document has no IANA Considerations.

11. Acknowledgments

The author wishes to thank the BGPsec design group, Thomas King, and Arnold Nipper.

12. References

12.1. Normative References

[I-D.ietf-sidr-bgpsec-overview]
Lepinski, M. and S. Turner, "An Overview of BGPSEC", draft-ietf-sidr-bgpsec-overview-02 (work in progress), May 2012.

[I-D.ietf-sidr-bgpsec-protocol]
Lepinski, M., "BGPSEC Protocol Specification", draft-ietf-sidr-bgpsec-protocol-07 (work in progress), February 2013.

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

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

[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.

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

[RFC7115] Bush, R., "Origin Validation Operation Based on the Resource Public Key Infrastructure (RPKI)", BCP 185, RFC 7115, DOI 10.17487/RFC7115, January 2014, <http://www.rfc-editor.org/info/rfc7115>.

12.2. Informative References

[I-D.ietf-idr-ix-bgp-route-server]
Jasinska, E., Hilliard, N., Raszuk, R., and N. Bakker, "Internet Exchange Route Server", draft-ietf-idr-ix-bgp-route-server-02 (work in progress), February 2013.

[I-D.ietf-sidr-bgpsec-rollover]
Gagliano, R., Patel, K., and B. Weis, "BGPSEC router key rollover as an alternative to beaconing", draft-ietf-sidr-bgpsec-rollover-01 (work in progress), October 2012.

[I-D.ietf-sidr-rtr-keying]
Turner, S., Patel, K., and R. Bush, "Router Keying for BGPsec", draft-ietf-sidr-rtr-keying-01 (work in progress), February 2013.

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

[RFC5065] Traina, P., McPherson, D., and J. Scudder, "Autonomous System Confederations for BGP", RFC 5065, August 2007.

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

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

[RFC6916] Gagliano, R., Kent, S., and S. Turner, "Algorithm Agility Procedure for the Resource Public Key Infrastructure (RPKI)", BCP 182, RFC 6916, DOI 10.17487/RFC6916, April 2013, <http://www.rfc-editor.org/info/rfc6916>.

Author's Address

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

Email: randy@psg.com