Signaling Prefix Origin Validation Results from a Route Server to Peers
draft-ietf-sidrops-route-server-rpki-light-02

Abstract

This document defines the usage of the BGP Prefix Origin Validation
State Extended Community [RFC8097] to signal prefix origin validation
results from a route server to its peers. Upon reception of prefix
origin validation results peers can use this information in their
local routing decision process.

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 [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 October 12, 2017.
1. Introduction

RPKI-based prefix origin validation [RFC6480] can be a significant operational burden for BGP peers to implement and adopt. In order to boost acceptance and usage of prefix origin validation and ultimately increase the security of the Internet routing system, IXPs may provide RPKI-based prefix origin validation at the route server [RFC7947]. The result of this prefix origin validation is signaled to peers by using the BGP Prefix Origin Validation State Extended Community as introduced in [RFC8097].

Peers receiving the prefix origin validation result from the route server(s) can use this information in their local routing decision...
process for acceptance, rejection, preference, or other traffic engineering purposes of a particular route.

2. BGP Prefix Origin Validation State Utilized at Route-Servers

A route server that is aware of a BGP Prefix Origin Validation state for a certain route can handle this information in one of the following modes of operation:

Simple Tagging: The prefix origin validation state is tagged to the route as described in Section 3. This mode of operation is like the traditional way route servers work, however, the prefix origin validation state information is additionally available for peers.

Dropping and Tagging: Routes for which the prefix origin validation state is "invalid" (according to [RFC6811]) are dropped by the route server. Routes which show a prefix origin validation state of "not found" and "valid" (according to [RFC6811]) are tagged accordingly to Section 3. Security is higher rated than questionable reachability of a prefix by this mode of operation.

Prioritizing and Tagging: If the route server learned for a particular prefix more than one route it removes firstly the set of "invalid" routes and secondly the "not found" routes unless the set of routes is empty. Based on the set of routes left over the BGP best path section algorithm is executed. The selected route is marked accordingly to Section 3. The BGP best path selection algorithm is changed by this mode of operation in such a way that "valid" routes are preferred even if they are unfavorable by the traditional best path selection algorithm. This puts prefix origin validation on top of the best path selection.

A route server MUST support the Simple Tagging mode of operation. Other modes of operation are OPTIONAL. The mode of operation MAY be configured by the route server operator for a route server instance or for each BGP session with a peer separately.

These mode of operations might be used in combination with [RFC7911] in order to allow a peer to receive all routes and take the routing decision by itself.
3. Signaling Prefix Origin Validation Results from a Route Server to Peers

The BGP Prefix Origin Validation State Extended Community (as defined in [RFC8097]) is utilized for signaling prefix origin validation result from a route server to peers.

[RFC8097] proposes an encoding of the prefix origin validation result [RFC6811] as follows:

<table>
<thead>
<tr>
<th>Value</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Lookup result = &quot;valid&quot;</td>
</tr>
<tr>
<td>1</td>
<td>Lookup result = &quot;not found&quot;</td>
</tr>
<tr>
<td>2</td>
<td>Lookup result = &quot;invalid&quot;</td>
</tr>
</tbody>
</table>

Table 1

This encoding is re-used. Route servers providing RPKI-based prefix origin validation set the validation state according to the prefix origin validation result (see [RFC6811]).

4. Operational Recommendations

4.1. Local Routing Decision Process

A peer receiving prefix origin validation results from the route server MAY use the information in its own local routing decision process. The local routing decision process SHOULD apply to the rules as described in section 5 [RFC6811].

A peer receiving a prefix origin validation result from the route server MAY redistribute this information within its own AS.

4.2. Route Server Receiving the BGP Prefix Origin Validation State Extended Community

An IXP route server receiving routes from its peers containing the BGP Prefix Origin Validation State Extended Community MUST remove the extended community before the route is re-distributed to its peers. This is required regardless of whether the route server is executing prefix origin validation or not.

Failure to do so would allow opportunistic peers to advertise routes tagged with arbitrary prefix origin validation results via a route
4.3. Information about Validity of a BGP Prefix Origin Not Available at a Route-Server

In case information about the validity of a BGP prefix origin is not available at the route server (e.g., error in the ROA cache, CPU overload) the route server MUST NOT add the BGP Prefix Origin Validation State Extended Community to the route.

4.4. Error Handling at Peers

A route sent by a route server SHOULD only contain none or one BGP Prefix Origin Validation State Extended Community.

A peer receiving a route from a route server containing more than one BGP Prefix Origin Validation State Extended Community SHOULD only consider the largest value (as described in Table 1) in the validation result field and disregard the other values. Values larger than two in the validation result field MUST be disregarded.

5. IANA Considerations

None.

6. Security Considerations

All security considerations described in RFC RFC6811 [RFC6811] fully apply to this document.

Additionally, threat agents polluting ROA cache server(s) run by IXP operators could cause significant operational impact, since multiple route server clients could be affected. Peers should be vigilant as to the integrity and authenticity of the origin validation results, as they are provided by a third party, namely the IXP operator hosting both the route server as well as any ROA cache server(s).

Therefore, a route server could be misused to spread malicious prefix origin validation results. However, peers already trust the route server for the collection, filtering (e.g. IRR database filtering), and redistribution of BGP routing information to other peers. So, no change in the trust level is needed for this proposal.

To facilitate trust and help with peers establishing appropriate controls in mitigating the risks mentioned above, IXPs SHOULD provide out-of-band means for peers to ensure that the ROA validation process has not been compromised or corrupted.
While being under DDoS attacks, it is a common practice for peers connected to an IXP to make use of blackholing services (see [RFC7999]). Peers are using blackholing to drop traffic, typically by announcing a more specific prefix, which is under attack. A peer SHOULD make sure that this prefix is covered by an appropriate ROA.

7. References

7.1. Normative References


7.2. Informative References


Internet-DraSignaling Prefix Origin Validation Results from April 2017


Authors’ Addresses

Thomas King
DE-CIX Management GmbH
Lichtstrasse 43i
Cologne 50825
DE

Email: thomas.king@de-cix.net

Daniel Kopp
DE-CIX Management GmbH
Lichtstrasse 43i
Cologne 50825
DE

Email: daniel.kopp@de-cix.net

Aristidis Lambrianidis
Amsterdam Internet Exchange
Frederiksplein 42
Amsterdam 1017 XN
NL

Email: aristidis.lambrianidis@ams-ix.net

Arnaud Fenioux
France-IX
88 Avenue Des Ternes
Paris 75017
FR

Email: afenioux@franceix.net
Abstract

It is useful for RPKI-based Origin Validation to classify and mark prefixes for all ingress, redistribution, and egress policies. For egress policy, it is important that the classification uses the effective origin AS of the processed route, which may specifically be altered by the commonly available knobs such as removing private ASs, confederation handling, and other modifications of the origin AS.

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 [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 https://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 July 7, 2019.
1. Introduction

As the origin AS may be modified by outbound policy, policy semantics based on RPKI Origin Validation state MUST be able to be applied separately on distribution into BGP and on egress.

When applied to egress policy, the effective origin AS MUST be used to determine the Origin Validation state. The effective origin AS is that which will actually be the origin AS in the announcement. It might be affected by removal of private AS(s), confederation, AS migration, etc. If there are any AS_PATH modifications resulting in origin AS change, then these MUST be taken into account.

2. Suggested Reading

It is assumed that the reader understands BGP, [RFC4271], the RPKI, [RFC6480], Route Origin Authorizations (ROAs), [RFC6482], RPKI-based Prefix Validation, [RFC6811], and Origin Validation Clarifications, [RFC8481].

3. Egress Processing

BGP implementations supporting RPKI-based origin validation SHOULD provide the same policy configuration primitives for decisions based on validation state available for use in ingress, redistribution, and egress policies. When applied to egress policy, validation state MUST be determined using the effective origin AS of the route as it will (or would) be announced to the peer. The effective origin AS may differ from that of the route in the RIB due to commonly available knobs such as: removal of private ASs, AS path manipulation, confederation handling, etc.
Egress policy handling can provide more robust protection for outbound eBGP than relying solely on ingress (iBGP, eBGP, connected, static, etc.) redistribution being configured and working correctly — better support for the robustness principle.

4. Security Considerations

This document does not create security considerations beyond those of [RFC6811] and [RFC8481].

5. IANA Considerations

This document has no IANA Considerations.

6. References

6.1. Normative References


6.2. Informative References


Authors’ Addresses

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

Email: randy@psg.com

Ruediger Volk
Deutsche Telekom

Jakob Heitz
Cisco Systems, Inc.
Origin Validation Signaling
draft-ymbk-sidrops-ov-signal-02

Abstract

Within a trust boundary, e.g. an operator’s PoP, it may be useful to have only a few central devices do full Origin Validation using the Resource Public Key Infrastructure, and be able to signal to an internal sender that a received route fails Origin Validation. E.g. route reflectors could perform Origin Validation for a cluster and signal back to a sending client that it sent an invalid route. Routers capable of sending and receiving this signal can use the extended community described in [RFC8097]

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 [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 https://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 July 2, 2019.
1. Introduction

Within a routing trust boundary, e.g. an operator's Point of Presence (PoP), it may not be desirable or necessary for all routers to perform Origin Validation using the Resource Public Key Infrastructure (RPKI) per [RFC6811]. A good example is route reflectors (see [RFC4456]).

An RPKI-enabled device, an Evaluator, SHOULD signal receipt of an Invalid route back to the sender by announcing that route back to the sender marked with the BGP Prefix Origin Validation State Extended Community as defined in [RFC8097] with a last octet having the value 2, meaning "Invalid." In the rest of this document we take the liberty of calling it the "community."

We use the term "Sender" to refer to the router announcing routes to the device evaluating the Origin Validation of the announcements. Beware that the Sender receives signaling back from the Evaluator, which can be somewhat confusing.

We use the term "Evaluator" to describe the device receiving routing announcements from senders, applying RPKI-based Origin Validation, and possibly signaling route Invalidity back to the sender(s).

2. Suggested Reading

It is assumed that the reader understands BGP, [RFC4271], the RPKI, [RFC6480], RPKI-based Prefix Validation, [RFC6811], and the BGP Prefix Origin Validation State Extended Community as described in [RFC8097].
3. Trust Boundary

As a general rule, we discourage ‘outsourcing trust,’ i.e. letting others make security decisions for us. But there are operational environments with a somewhat wide trust boundary, a single operator’s PoP for example.

This is not outsourcing trust; this is remote decision making. It is not letting a third party make the decision; it is simply doing it on a different computer. It’s trust in a distributed system, where what is (sometimes) called the Policy Decision Point is not the same as the Policy Enforcement Point.

As described in [RFC7115], a PoP might have a single RPKI Cache, hence all trust is vested in it. So it is reasonable that routers in that PoP could share Origin Validation results instead of each doing full validation.

An [RFC4456] Route Reflector Cluster is an obvious candidate for this approach. The route reflector(s) would perform Origin Validation and signal an Invalid route back to the sending client.

[RFC8097] provides the obvious signaling mechanism, the BGP Prefix Origin Validation State Extended Community. The device performing OV SHOULD signal back to the sender by announcing the offending prefix marked with the extended community with the last octet having the value 2, indicating an Invalid route.

4. The OV Signaling Capability

Unfortunately, the router sending the Invalid announcement is not normally expecting to receive it back. Therefore, both parties MUST agree on this feature by using a BGP Capability [RFC5492].

To advertise the OV Signaling Capability to a peer, a BGP speaker uses BGP Capabilities Advertisement [RFC5492]. By advertising the OV Signaling Capability to a peer, a BGP speaker conveys that it is able to send, receive, and properly handle OV Signaling using the community.

A peer which does not advertise this capability MUST NOT send OV Signaling, and BGP OV Signaling MUST NOT be sent to it.

The OV Signaling Capability is a new BGP Capability defined with Capability code [TBD] and Capability length 0.
5. Recommended Action

This section assumes that the OV Signaling Capability has been negotiated by the sending and receiving routers.

An Evaluating device which performs Origin Validation on a route received from a capable sender and finds a prefix with a particular origin AS to be Invalid (in the [RFC6811] sense), MUST announce that prefix back to the sending router from which it was received with the Invalid origin AS and the addition of the community with the last octet being 2.

A sender receiving the returned prefix announcement so marked MUST treat it the way it would treat an Invalid origin that it itself detected. It should withdraw all routes it had announced to that prefix with the Invalid origin AS. This includes withdrawing any instances of additional paths with that origin AS advertised under [RFC7911].

For a sender to properly evaluate the community returned by the evaluator, the sender MUST recognize the community before loop detection. This is a change to the Phase 2 Route Selection process of [RFC4271] Section 9.1.2.

If a sender originally received the Invalid route from an evaluator within its trust boundary with which it has negotiated the OV Signaling Capability, it MAY also propagate that signal to the original sender.

6. Security Considerations

As with all communities which cause semantic change, this use of the community may be abused as an attack vector. Therefore the operator MUST configure their incoming external border to strip the community.

As the BGP sessions are already established using whatever channel security the operator chooses or not, this change specifies no additional channel or object security. Of course, the BGP transport should be protected for integrity and authentication. TCP-MD5 [RFC2385] is available on almost all platforms. If more modern methods are available, they should be used.

Outsourcing security is usually considered bad policy. Section Section 3 above discusses why that is not really the case here.

Otherwise, this document does not create security considerations beyond those of [RFC6811].
7. IANA Considerations

This document requests the IANA assign the "OV Signaling Capability" to the BGP Capabilities described in Section 2.1 in the "Capability Codes" registry's "IETF Review" range [RFC8126]. This document is the reference for the new capability.

8. Acknowledgments

Thanks to Steve Bellovin for a serious security review, and Rob Austein for a useful security snark.

9. References

9.1. Normative References


9.2. Informative References


Authors’ Addresses

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

Email: randy@psg.com

Keyur Patel
Arrcus
2077 Gateway Place, Suite #250
San Jose, CA 95119
United States of America

Email: keyur@arrcus.com