Network Working Group   T. King
Internet-Draft   C. Dietzel
Intended status: Standards Track   D. Kopp
Expires: September 22, 2018   DE-CIX
                                                     A. Lambrianidis
                                                             AMS-IX
                                                          A. Fenioux
                                                           France-IX
                                                      March 21, 2018

        Signaling Prefix Origin Validation Results from an RPKI Origin
                Validating BGP Speaker to BGP Peers
                draft-ietf-sidrops-validating-bgp-speaker-02

Abstract

   This document defines a new BGP transitive extended community, as
   well as its usage, to signal prefix origin validation results from an
   RPKI Origin validating BGP speaker to other BGP 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", "NOT RECOMMENDED", "MAY", and
   "OPTIONAL" in this document are to be interpreted as described in BCP
   14 [RFC2119] [RFC8174] when, and only when, they appear in all
   capitals, as shown here.

Status of This Memo

Table of Contents

[1](1). Introduction

   RPKI-based prefix origin validation [[RFC6480](RFC6480)] can be a significant
   operational burden for BGP peers to implement and adopt. To
   facilitate acceptance and usage of prefix origin validation and
   ultimately increase the security of the Internet routing system,
   Autonomous Systems may provide RPKI-based prefix origin validation at
   certain vantage points. The result of this prefix origin validation
   is signaled to peers by using the EBGP Prefix Origin Validation State
   Extended Community as introduced in this document.

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

2. EBGP Prefix Origin Validation Extended Community

The origin validation state extended community is a transitive Four-
octet AS Specific Extended Community [RFC5668] with the following
encoding:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      0x02    |TBD1 (Sub-Type)|     Global Administrator     :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:  Global Administrator (cont.) |      Validation State       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
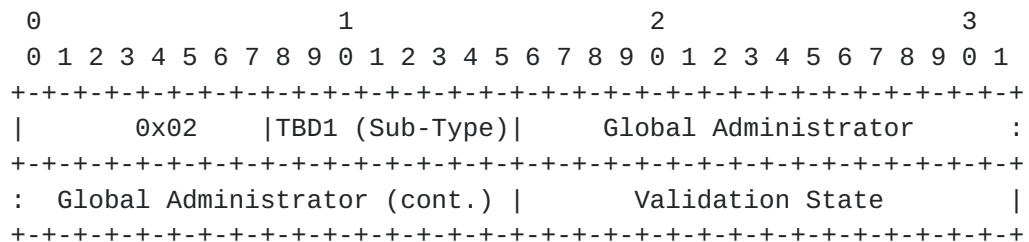
                            Figure 1

The value of the high-order octet of the extended Type field is 0x02,
which indicates it is transitive. The value of the low-order octet
(Sub-Type) of the extended Type field as assigned by IANA is TBD1.
The Reserved field MUST be set to 0 and ignored upon receipt of this
community. The Global Administrator field MUST be set to the AS
number of the validating BGP speaker conducting the prefix origin
validation. The last octets of the extended community is an unsigned
integer that gives the route's validation state as described in
Section 4.

If the validating BGP speaker is configured to support the extensions
defined in this document, it SHOULD attach the origin validation
state extended community to BGP UPDATE messages sent to EBGP peers by
mapping the computed validation state in the last octet of the
extended community. A receiving BGP speaker, in the absence of a
local validation state, SHOULD derive a validation state from the
last octet of the received extended community, if present.

An implementation SHOULD NOT send more than one instance of the
origin validation state extended community. However, if more than
one instance is received, an implementation MUST disregard all
instances other than the one with the numerically greatest validation
state value. If the value received is greater than the largest
specified value (2), the implementation MUST apply a strategy similar
to attribute discard [RFC7606] by discarding the erroneous community
and logging the error for further analysis.

3. **BGP Prefix Origin Validation State Utilized at Validating Peers**

   A validating BGP speaker that is aware of a BGP Prefix Origin
   Validation state (see Section 4) for a certain route can handle this
   information in one of the following modes of operation, attaching
   validation state to routes as discussed in Section 2:

   Simple Tagging: In this mode of operation, the BGP best path
      selection algorithm is executed. The prefix origin validation
      state is tagged accordingly.

   Dropping and Tagging: Routes for which the prefix origin validation
      state is "invalid" (according to [RFC6811]) are dropped by the
      validating BGP speaker. Based on the remaining set of routes,
      the BGP best path selection algorithm is executed. The prefix
      origin validation state of "not found" or "valid" (according to
      [RFC6811]) is tagged accordingly.

   Strict Dropping and Tagging: Routes for which the prefix origin
      validation state is "invalid" or "not found" (according to
      [RFC6811]) are dropped by the validating BGP speaker. Based on
      the remaining set of routes, the BGP best path selection
      algorithm is executed. The prefix origin validation state of
      "valid" is tagged to the advertised route.

   A validating BGP speaker MUST support the Simple Tagging operation
   mode. Other modes of operation are OPTIONAL. The mode of operation
   MAY be configured by the validating BGP speaker operator for all
   connected peers, or for each BGP session with a peer separately.

   Path hiding, as originally discussed in [RFC7947], may impact end-to-
   end connectivity for peers receiving prefixes via validating BGP
   speakers, if the best path selected contains a prefix with a prefix
   origin validation state which is subsequently dropped.

   However, these modes of operation might be used in combination with
   any options discussed in Section 2.3.2 of [RFC7947] in order to allow
   a peer to receive one or more routes and take the routing decision by
   itself, or with implementations who support sending the next best
   available path.

4. **Signaling Prefix Origin Validation Results from a Validating Peer to
      Peers**

   The EBGP Prefix Origin Validation State Community is utilized for
   signaling prefix origin validation result from a validating BGP
   speaker to other peers.

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

```
           +-------+-----------------------------+
           | Value | Meaning                     |
           +-------+-----------------------------+
           |   0   | Lookup result = "valid"     |
           |   1   | Lookup result = "not found" |
           |   2   | Lookup result = "invalid"   |
           +-------+-----------------------------+
```

                            Table 1

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

5. Operational Recommendations

5.1. 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.

In cases where multiple ASes are being administered by the same
authority, peers MAY also redistribute this information across EBGP
boundaries of the authority in question.

5.2. Validating Peers Receiving the EBGP Prefix Origin Validation State
      Extended Community

A validating BGP speaker receiving routes from peers containing the
EBGP 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 validating BGP
speaker 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 validating
peers, influencing maliciously the decision process of other, non-
validating BGP speakers.

5.3. **Information about Validity of a BGP Prefix Origin Not Available at a Validating Peer**

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

5.4. **Error Handling at Peers**

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

   A peer receiving a route from a validating BGP speaker containing
   more than one EBGP 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.

6. **IANA Considerations**

   IANA is asked to assign a transitive Four-octet AS Specific Extended
   Community Sub-Type as defined in Section 5.2.4 of [RFC7153].

7. **Security Considerations**

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

   Additionally, threat agents polluting ROA cache server(s) run by AS
   operators could cause significant operational impact, since multiple
   validating BGP speaker 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
   AS operator hosting both the validating BGP speaker as well as any
   ROA cache server(s).

   Therefore, a validating BGP speaker could be misused to spread
   malicious prefix origin validation results. However, in the case of
   IXPs, peers already trust the route server for the collection,
   filtering (e.g., IRR database filtering), and redistribution of BGP
   routing information to other peers.

   To facilitate trust and support with peers establishing appropriate
   controls in mitigating the risks mentioned above, AS operators 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 other Autonomous Systems and make use of blackholing
services. Peers are using blackholing to drop traffic, typically by
announcing a more specific prefix, which is under attack. A peer
SHOULD ensure that this prefix is covered by an appropriate ROA.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
           Requirement Levels", BCP 14, RFC 2119,
           DOI 10.17487/RFC2119, March 1997,
           <https://www.rfc-editor.org/info/rfc2119>.

[RFC5668] Rekhter, Y., Sangli, S., and D. Tappan, "4-Octet AS
           Specific BGP Extended Community", RFC 5668,
           DOI 10.17487/RFC5668, October 2009,
           <https://www.rfc-editor.org/info/rfc5668>.

[RFC6811] Mohapatra, P., Scudder, J., Ward, D., Bush, R., and R.
           Austein, "BGP Prefix Origin Validation", RFC 6811,
           DOI 10.17487/RFC6811, January 2013,
           <https://www.rfc-editor.org/info/rfc6811>.

[RFC7153] Rosen, E. and Y. Rekhter, "IANA Registries for BGP
           Extended Communities", RFC 7153, DOI 10.17487/RFC7153,
           March 2014, <https://www.rfc-editor.org/info/rfc7153>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K.
           Patel, "Revised Error Handling for BGP UPDATE Messages",
           RFC 7606, DOI 10.17487/RFC7606, August 2015,
           <https://www.rfc-editor.org/info/rfc7606>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
           2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
           May 2017, <https://www.rfc-editor.org/info/rfc8174>.

8.2. Informative References

[RFC6480] Lepinski, M. and S. Kent, "An Infrastructure to Support
           Secure Internet Routing", RFC 6480, DOI 10.17487/RFC6480,
           February 2012, <https://www.rfc-editor.org/info/rfc6480>.

[RFC7947] Jasinska, E., Hilliard, N., Raszuk, R., and N. Bakker,
           "Internet Exchange BGP Route Server", RFC 7947,
           DOI 10.17487/RFC7947, September 2016,
           <https://www.rfc-editor.org/info/rfc7947>.

Authors' Addresses

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

    Email: thomas.king@de-cix.net

    Christoph Dietzel
    DE-CIX Management GmbH
    Lichtstrasse 43i
    Cologne 50825
    DE

    Email: christoph.dietzel@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: ietf@afenioux.fr