Revised Validation Procedure for BGP Flow Specifications
draft-ietf-idr-bgp-flowspec-oid-09

Abstract

   This document describes a modification to the validation procedure
   defined in RFC 5575bis for the dissemination of BGP Flow
   Specifications. RFC 5575bis requires that the originator of the Flow
   Specification matches the originator of the best-match unicast route
   for the destination prefix embedded in the Flow Specification. This
   allows only BGP speakers within the data forwarding path (such as
   autonomous system border routers) to originate BGP Flow
   Specifications. Though it is possible to disseminate such Flow
   Specifications directly from border routers, it may be operationally
   cumbersome in an autonomous system with a large number of border
   routers having complex BGP policies. The modification proposed
   herein enables Flow Specifications to be originated from a
   centralized BGP route controller.

Status of This Memo

Table of Contents

1. 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].

2. Introduction

   [RFC5575bis] defined a new BGP [RFC4271] capability that can be used
   to distribute traffic Flow Specifications amongst BGP speakers in
   support of traffic filtering. The primary intention of [RFC5575bis]
   is to enable downstream autonomous systems to signal traffic
   filtering policies to upstream autonomous systems. In this way,
   traffic is filtered closer to the source and the upstream autonomous
   system(s) avoid carrying the traffic to the downstream autonomous
   system only to be discarded. [RFC5575bis] also enables more granular
   traffic filtering based upon upper layer protocol information (e.g.,

protocol port numbers) as opposed to coarse IP destination prefix-
based filtering. Flow specification NLRIs received from a BGP peer
are subject to validity checks before being considered feasible and
subsequently installed within the respective Adj-RIB-In.

The validation procedure defined within [RFC5575bis] requires that
the originator of the Flow Specification NLRI matches the originator
of the best-match unicast route for the destination prefix embedded
in the Flow Specification. This allows only BGP speakers within the
data forwarding path (such as autonomous system border routers) to
originate BGP Flow Specification NLRIs. Though it is possible to
disseminate such Flow Specification NLRIs directly from border
routers, it may be operationally cumbersome in an autonomous system
with a large number of border routers having complex BGP policies.

This document describes a modification to the [RFC5575bis] validation
procedure allowing Flow Specification NLRIs to be originated from a
centralized BGP route controller within the local autonomous system
that is not in the data forwarding path. While the proposed
modification cannot be used for inter-domain coordination of traffic
filtering, it greatly simplifies distribution of intra-domain traffic
filtering policies within an autonomous system which has a large
number of border routers having complex BGP policies. By relaxing
the validation procedure for iBGP, the proposed modification allows
Flow Specifications to be distributed in a standard and scalable
manner throughout an autonomous system.

3. Motivation

Step (a) of the validation procedure in [RFC5575bis], section 6 is
defined with the underlying assumption that the Flow Specification
NLRI traverses the same path, in the inter-domain and intra-domain
route distribution graph, as that of the longest-match unicast route
for the destination prefix embedded in the Flow Specification.

In the case of inter-domain traffic filtering, the Flow Specification
originator at the egress border routers of an AS (e.g. RTR-D and
RTR-E of ASN1 in figure 1) matches the eBGP neighbor that advertised
the longest match destination prefix (see RTR-F and RTR-G
respectively in figure 1). Similarly, at the ingress border routers
of ASN (see RTR-A and RTR-B of ASN1 in figure 1), the Flow
Specification originator matches the egress iBGP border routers that
had advertised the unicast route for the best-match destination
prefix (see RTR-D and RTR-E respectively in figure 1). This is true
even when ingress border routers select paths from different egress
border routers as best path based upon IGP distance. For example, in
figure 1:

RTR-A chooses RTR-D's path as best

RTR-B chooses RTR-E as the best path

```
               / - - - - - - - - - - - - - -
               |           ASN1             |
                 +-------+         +-------+
               | |       |         |       | |
                 | RTR-A |         | RTR-B |
               | |       |         |       | |
                 +-------+         +-------+
               |       \           /        |
                 iBGP   \         / iBGP
               |         \       /          |
                        +-------+
               |        |       |           |
                        | RTR-C |
               |        | RC    |           |
                        +-------+
               |          /   \             |
                         /     \
               |   iBGP /       \ iBGP      |
                 +-------+         +-------+
               | | RTR-D |         | RTR-E | |
                 |       |         |       |
               | |       |         |       | |
                 +-------+         +-------+
               |     |                 |   |
                - - -|- - - - - - - - -|- - -/
                     | eBGP       eBGP |
                - - -|- - - - - - - - -|- - -/
               |     |                 |   |
                 +-------+         +-------+
               | |       |         |       | |
                 | RTR-F |         | RTR-G |
               | |       |         |       | |
                 +-------+         +-------+
               |           ASN2             |
               / - - - - - - - - - - - - - -
```

                           Figure 1

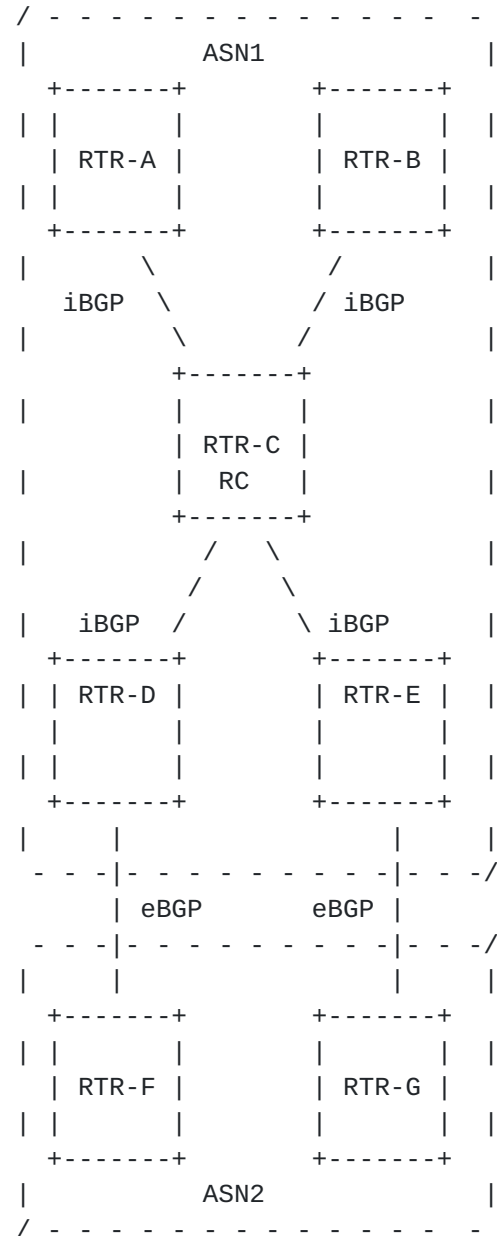   It is highly desirable that the mechanisms exist to protect each ASN
   independently from network security attacks using the BGP Flow
   Specification NLRI for intra-domain purposes only. Network operators
   often deploy a dedicated Security Operations Center (SOC) within
   their ASN to monitor and detect such security attacks. To mitigate
   attacks within a domain (AS or group of ASes), operators require the

ability to originate intra-domain Flow Specification NLRIs from a
central BGP route controller that is not within the data forwarding
plane. In this way, operators can direct border routers within their
ASN with specific attack mitigation actions (drop the traffic,
forward to a clean-pipe center, etc.).

To originate a Flow Specification NLRI, a central BGP route
controller must set itself as the originator in the Flow
Specification NLRI. This is necessary given the route controller is
originating the Flow Specification rather than reflecting it, and to
avoid the complexity of having to determine the egress border router
whose path was chosen as the best in each of the ingress border
routers. Thus, it is necessary to modify step (a) of the
[RFC5575bis] validation procedure such that an iBGP peer that is not
within the data forwarding plane may originate Flow Specification
NLRIs.

4. Revised Validation Procedure

4.1. Revision of Route Feasibility

Step (a) of the validation procedure specified in [RFC5575bis],
section 6 is redefined as follows:

a. One of the following conditions MUST hold true.

    1. The originator of the Flow Specification matches the
        originator of the best-match unicast route for the
        destination prefix embedded in the Flow Specification.

    2. The AS_PATH attribute of the Flow Specification does not
        contain AS_SET and/or AS_SEQUENCE segments.

        1. This condition SHOULD be enabled by default. This
            default behavior should validate an empty AS_PATH.

        2. This condition MAY be disabled by configuration on a BGP
            speaker.

        3. As an exception to this rule, a given AS_PATH with AS_SET
            and/or AS_SEQUENCE segments MAY be validated by policy.

Explanation:

    In this context, an empty AS_PATH means that it does not have
    AS_SET and/or AS_SEQUENCE segments, and local domain means the
    local AS [RFC4271] or the local confederation of ASes (in the case
    that the local AS belongs to a confederation of ASes [RFC5065]).

Thus, receiving a Flow Specification with an empty AS_PATH
indicates that the Flow Specification was originated inside the
local domain.

With the above modification to the [RFC5575bis] validation
procedure, a BGP peer within the local domain that is not within
the data forwarding path can originate a Flow Specification.

Disabling the new condition above (a.2.2) may be a good practice
when the operator knows with certainty that there is not a Flow
Specification originated inside the local domain.

Also, policy may be useful to validate a specific set of non-empty
AS_PATHs (a.2.3). For example, it could validate a Flow
Specification whose AS_PATH contains only an AS_SEQUENCE with ASes
that are all known to belong to the same administrative domain.

4.2. Revision of AS_PATH Validation

[RFC5575bis] states:

o BGP implementations MUST also enforce that the AS_PATH attribute
   of a route received via the External Border Gateway Protocol
   (eBGP) contains the neighboring AS in the left-most position of
   the AS_PATH attribute.

This rule prevents the exchange of BGP Flow Specification NLRIs at
Internet exchanges with BGP route servers. Therefore, this document
also redefines the [RFC5575bis] AS_PATH validation procedure
referenced above as follows:

o BGP Flow Specification implementations MUST enforce that the AS in
   the left-most position of the AS_PATH attribute of a Flow
   Specification route received via the External Border Gateway
   Protocol (eBGP) matches the AS in the left-most position of the
   AS_PATH attribute of the best-match unicast route for the
   destination prefix embedded in the Flow Specification NLRI.

Explanation:

For clarity, the AS in the left-most position of the AS_PATH means
the AS that was last added to the AS_SEQUENCE.

This proposed modification enables the exchange of BGP Flow
Specification NLRIs at Internet exchanges with BGP route servers
while at the same time, for security reasons, prevents an eBGP
peer from advertising an inter-domain Flow Specification for a

destination prefix that it does not provide reachability
information for.

Comparing only the last ASes added is sufficient for eBGP learned
Flow Specification NLRIs. Requiring a full AS_PATH match would
limit origination of inter-domain Flow Specifications to the
origin AS of the best-match unicast route for the destination
prefix embedded in the Flow Specification only. As such, a full
AS_PATH validity check may prevent transit ASes from originating
inter-domain Flow Specifications, which is not desirable.

Redefinition of this AS_PATH validation rule for a Flow
Specification does not mean that the original rule in [RFC5575bis]
cannot be enforced as well. Its enforcement remains optional per
[RFC4271]. That is, we can enforce the first AS in the AS_PATH to
be the same as the neighbor AS for any address-family route
(including a Flow Specification).

Using the new rule to validate a Flow Specification received from
an Internal Border Gateway Protocol (iBGP) peer is out of the
scope of this document. Note that in most scenarios such
validation would be redundant.

Using the new rule to validate a Flow Specification route received
from an External Border Gateway Protocol (eBGP) peer belonging to
the same local domain (in the case that the local AS belongs to a
confederation of ASes) is out of the scope of this document. Note
that although it's possible, its utility is dubious.

5. Other RFC5575bis Considerations

This section clarifies some of the terminology and rules referenced
in [RFC5575bis]. Namely:

o [RFC5575bis] references "the best-match unicast route for the
   destination prefix embedded in the Flow Specification". For
   clarity, this route is defined hereby as the best path of the
   unicast network with the longer prefix-length that covers the
   destination prefix embedded in the Flow Specification.

Explanation:

We consider only the best-match network and we do not consider
unicast non-best paths (even if it is received from the same peer
as the Flow Specification).

   o If a BGP Flow Specification does not embed a destination prefix
      component, then the default behavior SHOULD be not to perform the
      validation procedures.

   Explanation:

      The reason for this is because the best-match unicast route is not
      defined in this case.

   o In the context of this document and [RFC5575bis], AS_PATH
      attribute is defined as the reconstructed AS path information (by
      combining AS_PATH and AS4_PATH attributes, if the BGP speaker is a
      NEW speaker and receives the route from an OLD speaker), according
      to section 4.2.3 of [RFC6793].

   o Support for two-octet AS only implementations is out of the scope
      of this document (i.e. it's assumed that the BGP speaker supports
      [RFC6793]).

6. Topology Considerations

   [RFC5575bis] indicates that the originator may refer to the
   originator path attribute (ORIGINATOR_ID) or (if the attribute is not
   present) the transport address of the peer from which we received the
   update. If the latter applies, a network should be designed so it
   has a congruent topology.

   With the additional second condition (a.2) in the validation
   procedure, non-congruent topologies are supported within the local
   domain if the Flow Specification is originated within the local
   domain.

   Explanation:

      Consider the following scenarios without the second condition
      (a.2) being added to the validation procedure:

      1. Consider a topology with two BGP speakers with two peering
          sessions between them, one for unicast and one for Flow
          Specification. This is a non-congruent topology. Let's
          assume that the ORIGINATOR_ID attribute was not received (e.g.
          a route reflector receiving routes from its clients). In this
          case, the Flow Specification validation procedure will fail
          because of the first condition (a.1).

      2. Consider a topology with a BGP speaker within a confederation
          of ASes, inside local AS X. ORIGINATOR_ID attribute is not
          advertised within the local domain. Let's assume the Flow

Specification route is received from peer A and the best-match unicast route is received from peer B. Both peers belong in local AS Y. Both AS X and AS Y belong to the same local domain. The Flow Specification validation procedure will also fail because of the first condition (a.1).

In the examples above, if Flow Specifications are originated in the same local domain, AS_PATH will not contain AS_SET and/or AS_SEQUENCE segments. When the second condition (a.2) in the validation procedure is used, the validation procedure will pass. Thus, non-congruent topologies are supported if the Flow Specification is originated in the same local domain.

Even when the second condition (a.2) is used in the validation procedure, a Flow Specification originated in a different local domain needs a congruent topology. AS_SEQUENCE is not empty and the first condition (a.1) in the validation procedure needs to be evaluated. Because transport addresses for Flow Specification and unicast routes are different, the validation procedure will fail. This is true both across domains and within domains. Consider both cases:

*   Consider the first example. If the Flow Specification route is originated in another AS, the validation procedure will fail because the topology is non-congruent within the domain.

*   Consider the second example and modify it so AS X and AS Y belong to different local domains (no confederation of ASes exists). The validation procedure will fail because the topology is non-congruent across domains.

7. IANA Considerations

This memo includes no request to IANA.

8. Security Considerations

No new security issues are introduced by relaxing the validation procedure for iBGP learned Flow Specifications. With this proposal, the security characteristics of BGP Flow Specifications remain equivalent to the existing security properties of BGP unicast routing. Traffic Flow Specifications learned from iBGP peers are trusted, hence, it is not required to validate that the originator of an intra-domain traffic Flow Specification matches the originator of the best-match unicast route for the flow destination prefix. Conversely, this proposal continues to enforce the validation procedure for eBGP learned traffic Flow Specifications. In this way, the security properties of [RFC5575bis] are maintained such that an

eBGP peer cannot cause a denial-of-service attack by advertising an
inter-domain Flow Specification for a destination prefix that it does
not provide reachability information for.

9. Acknowledgements

The authors would like to thank Han Nguyen for his direction on this
work as well as Waqas Alam, Keyur Patel, Robert Raszuk, Eric Rosen
and Shyam Sethuram for their review comments.

10. Normative References

[I-D.ietf-idr-rfc5575bis]
           Hares, S., Loibl, C., Raszuk, R., McPherson, D., and M.
           Bacher, "Dissemination of Flow Specification Rules",
           draft-ietf-idr-rfc5575bis-14 (work in progress), April
           2019.

[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>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A
           Border Gateway Protocol 4 (BGP-4)", RFC 4271,
           DOI 10.17487/RFC4271, January 2006,
           <https://www.rfc-editor.org/info/rfc4271>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route
           Reflection: An Alternative to Full Mesh Internal BGP
           (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006,
           <https://www.rfc-editor.org/info/rfc4456>.

[RFC5065] Traina, P., McPherson, D., and J. Scudder, "Autonomous
           System Confederations for BGP", RFC 5065,
           DOI 10.17487/RFC5065, August 2007,
           <https://www.rfc-editor.org/info/rfc5065>.

[RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet
           Autonomous System (AS) Number Space", RFC 6793,
           DOI 10.17487/RFC6793, December 2012,
           <https://www.rfc-editor.org/info/rfc6793>.

Authors' Addresses

James Uttaro
AT&T
200 S. Laurel Ave
Middletown, NJ 07748
USA

Email: ju1738@att.com

Juan Alcaide
Cisco
7100 Kit Creek Road
Research Triangle Park, NC 27709
USA

Email: jalcaide@cisco.com

Clarence Filsfils
Cisco

Email: cf@cisco.com

David Smith
Cisco
111 Wood Ave South
Iselin, NJ 08830
USA

Email: djsmith@cisco.com

Pradosh Mohapatra
Sproute Networks

Email: mpradosh@yahoo.com