          Revised Validation Procedure for BGP Flow Specifications
                  draft-ietf-idr-bgp-flowspec-oid-08

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.

Copyright Notice

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. 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, for example, the flow
specification originator at the egress border routers of ASN1 (RTR-D
and RTR-E in figure 1) matches the EBGP neighbor that advertised the
longest match destination prefix (RTR-F and RTR-G respectively).
Similarly, at the ingress border routers of ASN1 (RTR-A and RTR-B 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 (RTR-D and RTR-E respectively). This is
true even when ingress border routers select paths from different
egress border routers as best path based upon IGP distance (as an
example, 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             |
                  / - - - - - - - - - - - - -  -
```
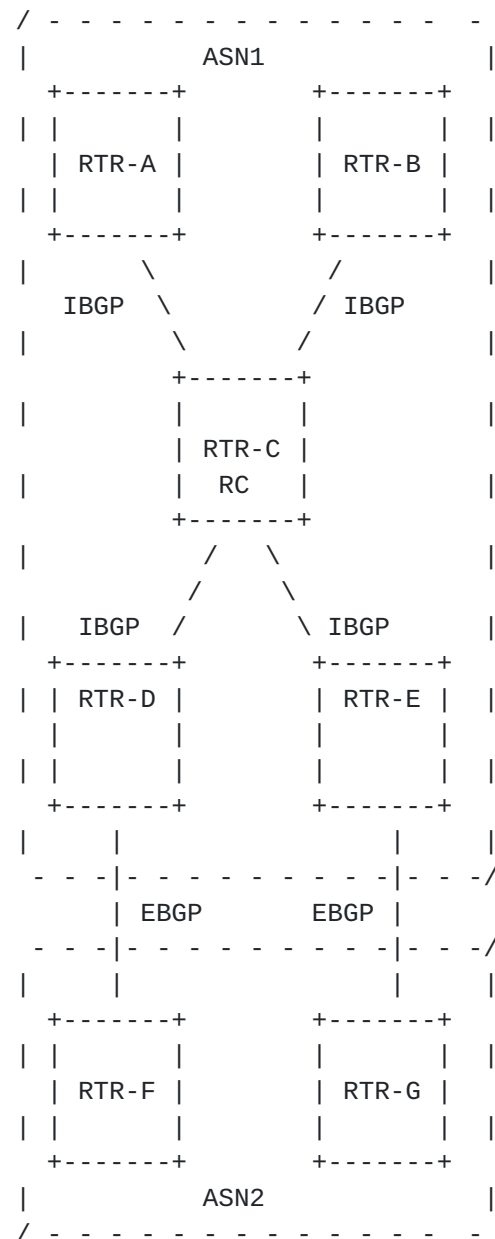
                         Figure 1

It is highly desirable that each ASN is able to protect itself
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 in a scalable intra-domain manner, 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 flowspec NLRI. This is necessary given the
route controller is originating the flow specification not 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. It thus becomes 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.

3. Introduction

[RFC5575bis] defined a new BGP 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 [RFC4271]
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 in an autonomous system with 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.

4. Revised Validation Procedure

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

a. One of the following conditions MUST hold true.

    * The originator of the flow specification matches the
       originator of the best-match unicast route for the destination
       prefix embedded in the flow specification.

    * The AS_PATH attribute of the flow specification does not
       contain AS_SET and/or AS_SEQUENCE segments.

An AS_PATH without AS_SET and/or AS_SEQUENCE segments indicates that
the flow specification was originated inside the local AS [RFC4271]
or inside the local confederation (in the case that the local AS
belongs to a confederation of ASes) [RFC5065]. With this
modification to the [RFC5575bis] validation procedure, it is now
possible for an IBGP peer that is not within the data forwarding path
to originate flow specification NLRIs. This applies whether the AS
belongs or not to a confederation of ASes. Checking the (newly
introduced) second condition above MAY be disabled by configuration
on a BGP speaker. However, it SHOULD be enabled by default.
Disabling the condition may be a good practice when the administrator
knows with certainty that there are not flow specification NLRI
originated inside the local AS (or local confederation). The default
behavior is thus to validate an empty AS_PATH. In this context, an
empty AS_PATH means that it does not have AS_SET and/or AS_SEQUENCE
segments. Optionally, an implementation MAY also validate a specific
non-empty AS_PATH. For instance, it could validate a flowspec NLRI
whose AS_PATH contains only an AS_SEQUENCE of ASes known (via
configuration) to belong to the same administrative domain.

Further, [RFC5575bis] states that "BGP (flow specification)
implementations MUST also enforce that 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 is not valid for all topologies. For example, it 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:

BGP flow specification implementations MUST enforce that the last AS
added within the AS_PATH attribute of a EBGP learned flow
specification NLRI MUST match the last AS added within the AS_PATH
attribute of the best-match unicast route for the destination prefix
embedded in the flow specification. 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. Note, 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 (or first) 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.

This document also clarifies proper handling when the BGP flow
specification does not embed a destination prefix component. The
default behavior SHOULD be not to perform any validation procedure.
Further, support for two-octet AS number space is out of the scope of
this document.

In this context, 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].

[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
that covers destination prefix embedded in the flow specification
with the longer prefix-length. In other words, 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 than the flowspec route).

Note that, per [RFC5575bis], originator may refer to the BGP
ORIGINATOR_ID attribute or the transport address of the peer from
which we received the update. If the later, a network must be
designed so it has a congruent topology. Otherwise, using two
peering sessions between the same pair of BGP speakers, one for
unicast and one for flowspec, will cause the flowspec validation
procedure to fail. Consider, for example, the case where a BGP route
reflector receives the NLRIs from a route reflector client, thus not
receiving the ORIGINATOR_ID attribute. If the speaker belongs to a
confederation [RFC5065] and we are receiving a flowspec route from
different peers than its best match unicast route, the flowspec

validation procedure will fail as well. Consider also a
misconfiguration where flowspec address-family is not configured for
a particular peering between different member-AS (but it is
configured for unicast). Even if we receive the flowspec route via a
redundant peer, we may receive the unicast route and the flowspec
from different peers, and thus flowspec validation will fail. Thus,
with the (newly introduced) second condition above applied,
incongruent topologies are supported.

Note that if the flowspec NLRI is learned from another AS (and thus
the AS_PATH is not empty), the original validation procedures defined
in [RFC5575bis] still apply and incongruent topologies may cause
validation rules to fail.

5. IANA Considerations

This memo includes no request to IANA.

6. 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.

7. 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.

8. 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