Network Working Group
Internet-Draft
Updates: 8955 (if approved)
Intended status: Standards Track
Expires: October 14, 2021

J. Uttaro
AT&T
J. Alcaide
C. Filsfils
D. Smith
Cisco
P. Mohapatra
Sproute Networks

April 12, 2021

Revised Validation Procedure for BGP Flow Specifications
draft-ietf-idr-bgp-flowspec-oid-13

Abstract

This document describes a modification to the validation procedure defined for the dissemination of BGP Flow Specifications. The dissemination of BGP Flow Specifications 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. For an iBGP received route, the originator is typically a border router within the same autonomous system. The objective is to allow only BGP speakers within the data forwarding path to originate BGP Flow Specifications. Sometimes it is desirable to originate the BGP Flow Specification any place within the autonomous system itself, for example, from a centralized BGP route controller. However, the validation procedure will fail in this scenario. The modification proposed herein relaxes the validation rule to enable Flow Specifications to be originated within the same autonomous system as the BGP speaker performing the validation. Additionally, this document revises AS_PATH validation rules so Flow Specifications received from an eBGP peer can be validated when such peer is a BGP route server.

This document updates the validation procedure in RFC8955.

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/.

Table of Contents

1. 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.

2. Introduction

[RFC8955] defines a BGP NLRI [RFC4271] that can be used to distribute traffic Flow Specifications amongst BGP speakers in support of traffic filtering. The primary intention of [RFC8955] 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. [RFC8955] 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 [RFC8955] 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. The aim is making sure that only speakers on the forwarding path can originate the Flow Specification. Let's consider the particular case where the Flow Specification is originated in any location within the same autonomous system than the speaker performing the validation (for example by a centralized BGP route controller), and the best-match unicast route is originated in another autonomous system. In order for validation to succeed for a Flow Specification received from an iBGP peer, it could be possible to disseminate such Flow Specification NLRIs directly from the specific border router (within the local autonomous system) that is advertising the corresponding best-match unicast route to the local autonomous system. This approach would be, however, operationally cumbersome in an autonomous system with a large number of border routers having complex BGP policies.

Figure 1 illustrates this principle. R1 (the upstream router) and RR need to validate the Flow Specification whose embedded destination prefix has a best-match unicast route (dest-route) originated by ASBR2. ASBR2 could originate the Flow Specification, and it would be validated when received by RR and R1. Sometimes the Flow Specification needs to be originated on AS1. ASBR1 could originate it, and Flow Specification would still be validated. In both cases, the Flow Specification is originated by a router in the same forwarding path as the dest-route. For the case where AS1 has thousands of ASBRs, it becomes impractical to originate different rules on each ASBR in AS1 based on which ASBR each dest- route is learned from. The objective is to advertise all the Flow Specifications from the same route-controller.

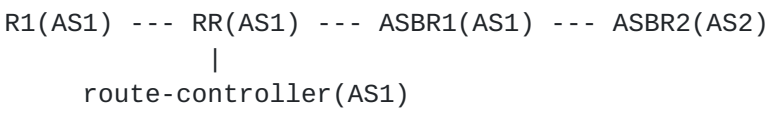

Figure 1

This document describes a modification to the [RFC8955] 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 (b) of the validation procedure in [RFC8955], 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 AS1 in Figure 2) matches the eBGP neighbor that advertised the longest match destination prefix (see RTR-F and RTR-G respectively in Figure 2).

Similarly, at the upstream routers of an AS (see RTR-A and RTR-B of AS1 in Figure 2), 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 2). This is true even when upstream routers select paths from different egress border routers as best route based upon IGP distance. For example, in Figure 2:

RTR-A chooses RTR-D as the best route

RTR-B chooses RTR-E as the best route

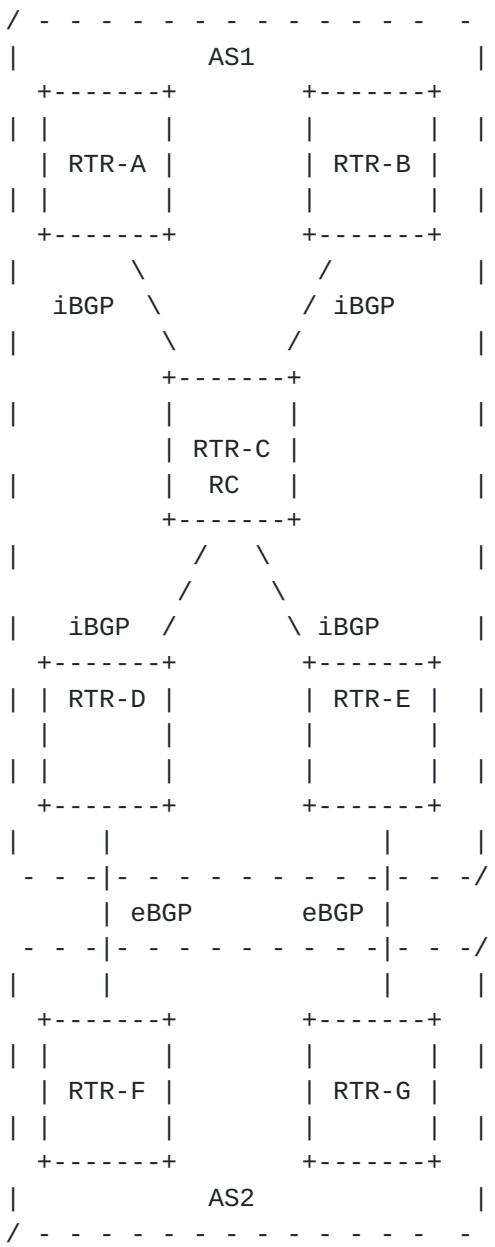

Figure 2

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

mitigation actions (drop the traffic, forward to a clean-pipe center, etc.).

In addition, an operator MAY extend the requirements above for a group of ASes via policy. This is described in section (b.2.3) of the validation procedure.

A central BGP route controller that originates a Flow Specification NLRI should be able to avoid the complexity of having to determine the egress border router whose path was chosen as the best for each of its neighbors. When a central BGP route controller originates a Flow Specification NLRI, the rest of the speakers within the AS will see the BGP controller as the originator of the Flow Specification in terms of the validation procedure rules. Thus, it is necessary to modify step (b) of the [RFC8955] 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 (b) of the validation procedure specified in [RFC8955], section 6 is redefined as follows:

b) 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 (this is the unicast route with the longest possible prefix length covering the destination prefix embedded in the Flow Specification).

2. The AS_PATH attribute of the Flow Specification is empty or contains only AS_CONFED_SEQUENCE and/or AS_CONFED_SET segments [RFC5065].

    1. This condition SHOULD be enabled by default (it may be disabled by explicit configuration as described on the next list item (b.2.1)).. an empty AS_PATH.

    2. This condition MAY be disabled by explicit configuration on a BGP speaker. A possible case would be if we know for a fact that only the right egress border routers (i.e. those that are also egress border routers for the best routes) are originating a Flow Specification NLRI.

3. As an extension to this rule, a given non-empty AS_PATHs (or AS_PATHS containing only AS_CONFED_SEQUENCE and/or AS_CONFED_SET segments), MAY be validated by policy. A possible case would be if the AS_SEQUENCE and AS_SET contained only ASes that are known to belong to our own administrative domain.

Explanation:

In this context, a local domain includes the local AS or the local confederation [RFC5065]. Receiving either an empty AS_PATH or one with only AS_CONFED_SEQUENCE and/or AS_CONFED_SET segments indicates that the Flow Specification was originated inside the local domain.

With the above modification to the [RFC8955] 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 (b.2.2) may be a good practice when the operator knows with certainty that a Flow Specification will not be originated inside the local domain.

Also, policy may be useful to validate a specific set of non-empty AS_PATHs (b.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

[RFC8955] states:

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 [RFC7947]. Therefore, this document also redefines the [RFC8955] AS_PATH validation procedure referenced above as follows:

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.

Note, however, that not checking the full AS_PATH allows any rogue or misconfigured AS the ability to originate undesired Flow Specifications. This is a security BGP threat, but out of the scope of this document.

Redefinition of this AS_PATH validation rule for a Flow Specification does not mean that the original rule in [RFC8955] cannot be enforced as well. Its enforcement remains optional per [RFC4271] section 6.3. That is, a BGP speaker 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 route received from an External Border Gateway Protocol (eBGP) peer belonging to the same local domain (in the case of a confederation) is out of the scope of this document. Note that although it's possible, its utility is dubious. Although it is conceivable that an router in the same local domain (both iBGP and eBGP within the same local domain) could send a rogue update, only eBGP (outside the local domain) risk is considered within this document (in the same spirit of the mentioned beforehand AS_PATH validation in [RFC4271]).

5. Topology Considerations

[RFC8955] 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 the BGP speaker received

the update. If the latter applies, a network should be designed so it has a congruent topology amongst ipv4 unicast routes and Flow Specification routes. By congruent topology, it is understood that for the two equivalent routes (i.e. the Flow Specification route and its best-match unicast route) are learned from the same peer accross the AS. That would likely not be true, for instance, if some peers only negotiated one type of address-family or if each address-family had a different set of policies.

With the additional second condition (b.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 validation procedure preceding this document. The second condition (b.2) does not exist. The two following scenarios have a non-congruent topology:

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 (b.1).

2. Consider a topology with a BGP speaker within a confederation of ASes, inside local AS X. The 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 to 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 (b.1).

In the scenarios above, if Flow Specifications are originated in the same local domain, the AS_PATH will be empty or contain just AS_CONFED_SEQUENCE and/or AS_CONFED_SET segments. Condition (b.2) evaluates to true. Therefore, using the second condition (b.2), as defined by this document, guarantees that the overall validation procedure will pass. Thus, non-congruent topologies are supported if the Flow Specification is originated in the same local domain.

Flow Specification originated in a different local domain needs a congruent topology. The reason is that the second condition (b.2)

evaluates to false and only the first condition (b.1) is evaluated.

6. IANA Considerations

This memo includes no request to IANA.

7. Security Considerations

This document updates the route feasibility validation procedures for Flow Specifications learned from iBGP peers and through route servers. This change is in line with the procedures in [rfc8955] and thus maintain security characteristics equivalent to the existing security properties of BGP unicast routing.

The security considerations discussed in [RFC8955] apply to this specification as well.

The original AS_PATH validation rule ([RFC4271] section 6.3) becomes hereby optional (section 4.2). If that original rule is actually not enforced it may introduce some security risks. A peer (or a client of a route server peer) in AS X could advertise a rogue Flow Specification route whose first AS in AS_PATH was Y (assume Y is the first AS in the AS_PATH of the best-match unicast route). This risk is impossible to prevent if the Flow Specification route is received from a route server peer. If that peer is known for a fact not to be a route server, that optional rule SHOULD be enforced for Flow Specification routes.

BGP updates learned from iBGP peers are considered trusted, so the Traffic Flow Specifications contained in BGP updates are also considered trusted. Therefore 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. Note that this trustworthy consideration is not absolute and the new possibility than an iBGP speaker could send a rouge Flow Specification is introduced.

The changes in Section 4.1 don't affect the validation procedures for eBGP-learned routes.

8. 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.

9. 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>.

[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>.

[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>.

[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>.

[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>.

[RFC8955] Loibl, C., Hares, S., Raszuk, R., McPherson, D., and M. Bacher, "Dissemination of Flow Specification Rules", RFC 8955, DOI 10.17487/RFC8955, December 2020, <https://www.rfc-editor.org/info/rfc8955>.

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