Stream: Internet Engineering Task Force (IETF)
RFC: 9117
Updates: 8955
Category: Standards Track
Published: August 2021
ISSN: 2070-1721
Authors: J. Uttaro, *AT&T*; J. Alcaide, *Cisco*; C. Filsfils, *Cisco*; D. Smith, *Cisco*; P. Mohapatra, *Sproute Networks*

RFC 9117
Revised Validation Procedure for BGP Flow Specifications

Abstract

This document describes a modification to the validation procedure defined for the dissemination of BGP Flow Specifications. The dissemination of BGP Flow Specifications as specified in RFC 8955 requires that the originator of the Flow Specification match the originator of the best-match unicast route for the destination prefix embedded in the Flow Specification. For an Internal Border Gateway Protocol (iBGP) received route, the originator is typically a border router within the same autonomous system (AS). 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 from any place within the autonomous system itself, for example, from a centralized BGP route controller. However, the validation procedure described in RFC 8955 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 the AS_PATH validation rules so Flow Specifications received from an External Border Gateway Protocol (eBGP) peer can be validated when such a peer is a BGP route server.

This document updates the validation procedure in RFC 8955.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9117.

Copyright Notice

Table of Contents

1. Introduction

[RFC8955] defines BGP Network Layer Reachability Information (NLRI) [RFC4760] 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 systems avoid carrying the traffic to the downstream autonomous systems only to be discarded. [RFC8955] also enables more granular traffic filtering based upon upper-layer protocol information (e.g., protocol or port numbers) as opposed to coarse IP destination prefix-based filtering. Flow Specification NLRIs received from a BGP peer is 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 match the originator of the best-match unicast route for the destination prefix embedded in the Flow Specification. The aim is to make 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 Local Domain as the speaker performing the validation (for example, by a centralized BGP route controller), and the best-match unicast route is originated in another Local Domain. In order for the validation to succeed for a Flow Specification received from an iBGP peer, it would be necessary to disseminate such Flow Specification NLRI directly from the specific border router (within the Local Domain) that is advertising the corresponding best-match unicast route to the Local Domain. Those border routers would be acting as de facto route controllers. This approach would be, however, operationally cumbersome in a Local Domain with numerous border routers having complex BGP policies.

Figure 1 illustrates this principle. R1 (the upstream router) and RR (a route reflector) 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 (from their point of view, the originator of both the Flow Specification and the best-match unicast route will be ASBR1). Sometimes the Flow Specification needs to be originated within AS1. ASBR1 could originate it, and the 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 Flow Specification rules on each ASBR in AS1 based on which ASBR each dest-route is learned from. To make the situation more tenable, the objective is to advertise all the Flow Specifications from the same route controller.

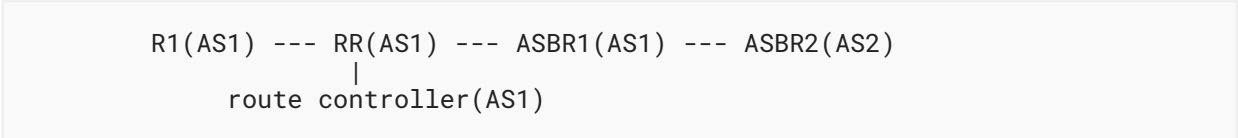

Figure 1

This document describes a modification to the validation procedure described in [RFC8955], by allowing Flow Specification NLRIs to be originated from a centralized BGP route controller located within the Local Domain and not necessarily 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 a Local Domain that has

numerous 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 the Local Domain.

Throughout this document, some references are made to AS_CONFED_SEQUENCE segments; see Sections 4.1 and 5. If AS_CONFED_SET segments are also present in the AS_PATH, the same considerations apply to them. Note, however, that the use of AS_CONFED_SET segments is not recommended [RFC6472]. Refer to [CONFED-SET] as well.

2. Definitions of Terms Used in This Memo

Local Domain: the local AS or the local confederation of ASes [RFC5065].

eBGP: BGP peering to a router not within the Local Domain.

iBGP: Both classic iBGP and any form of eBGP peering with a router within the same confederation (i.e., iBGP peering is a peering that is not eBGP as defined above).

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.

3. Motivation

Step (b) of the validation procedure in Section 6 of [RFC8955] 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 the 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

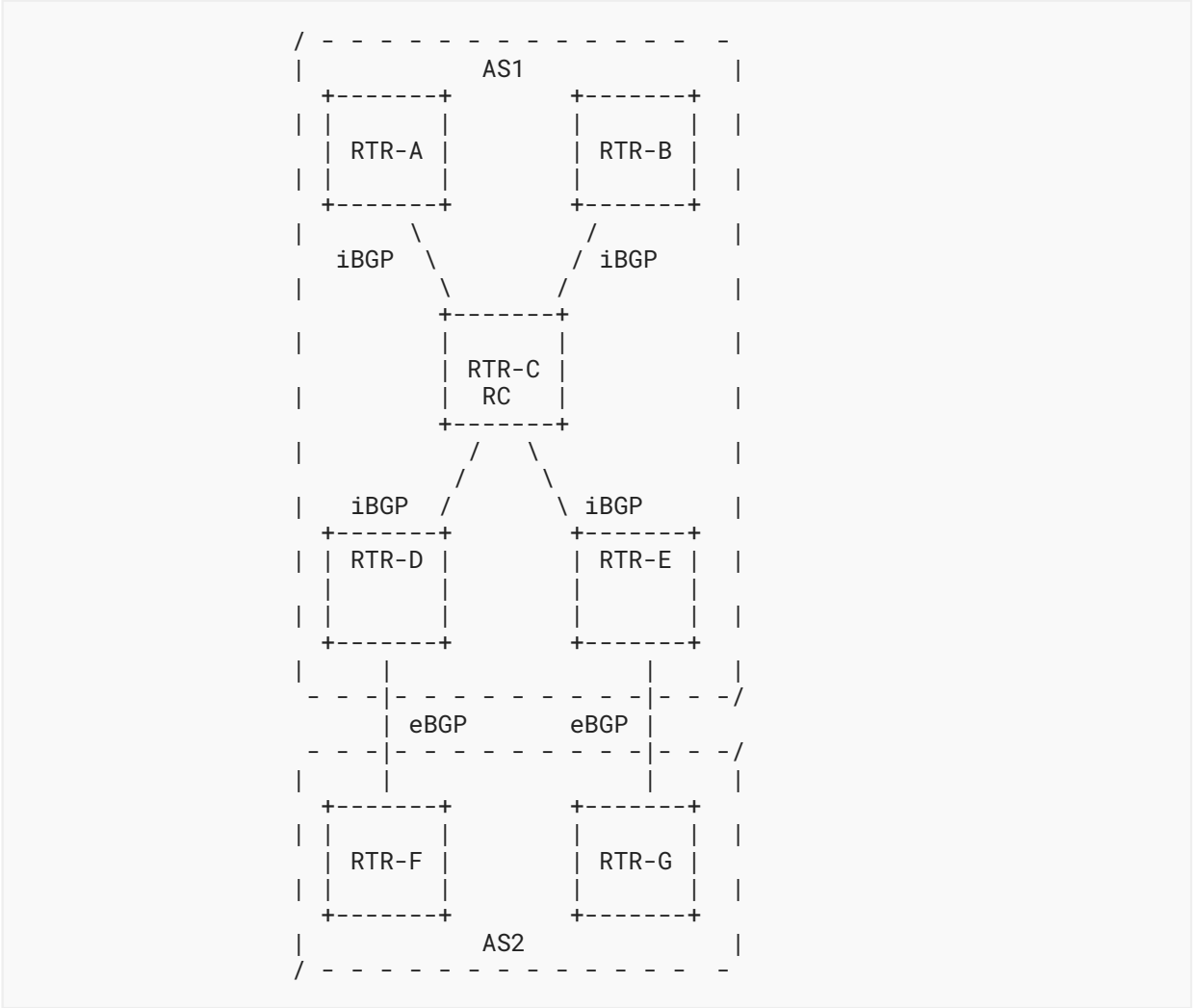

Figure 2

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 pipe-cleaning location, etc.).

In addition, an operator may extend the requirements above for a group of ASes via policy. This is described in Section 4.1 (b.2.3) of the validation procedure.

A central BGP route controller that originates 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 Flow Specification NLRI,

the rest of the speakers within the AS will see the BGP route 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 validation procedure described in [RFC8955] 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 Section 6 of [RFC8955] 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 an AS_CONFED_SEQUENCE segment [RFC5065].
>
>   1. This condition **SHOULD** be enabled by default.
>   2. This condition **MAY** be disabled by explicit configuration on a BGP speaker.
>   3. As an extension to this rule, a given non-empty AS_PATH (besides AS_CONFED_SEQUENCE segments) **MAY** be permitted by policy.

Explanation:

Receiving either an empty AS_PATH or one with only an AS_CONFED_SEQUENCE segment 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 (see step b.2.2 in Section 4.1) could be a good practice if the operator knew with certainty that a Flow Specification would not be originated inside the Local Domain. An additional case would be if it was known for a fact that only the right egress border routers (i.e., those that were also egress border routers for the best routes) were originating Flow Specification NLRI.

Also, policy may be useful to permit a specific set of non-empty AS_PATHs (see step b.2.3 in Section 4.1). For example, it could validate a Flow Specification whose AS_PATH contained only an AS_SEQUENCE segment with ASes that were all known to belong to the same administrative domain.

4.2. Revision of AS_PATH Validation

Section 6 of [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. While this rule is optional in the BGP specification, it becomes necessary to enforce it here for security reasons.

This rule prevents the exchange of BGP Flow Specification NLRIs at Internet exchanges with BGP route servers, which by design don't insert their own AS number into the AS_PATH (Section 2.2.2.1 of [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 an 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 left-most AS in the AS-PATH for eBGP-learned Flow Specification NLRIs is roughly equivalent to checking the neighboring AS. If the peer is a route server, security is necessarily weakened for the Flow Specification NLRI, as it is for any unicast route advertised from a route server. An example is discussed in the Security Considerations section.

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 Section 6.3 of [RFC4271]. That is, a BGP speaker can enforce the first AS in the AS_PATH to be the same as the neighbor AS for a route belonging to any Address Family (including Flow Specification Address Family). If the BGP speaker peer is not a route server, when enforcing this optional rule, the security characteristics are exactly equivalent to those specified in [RFC8955].

Alternatively, enforcing this optional rule for unicast routes (even if not enforced on Flow Specification NLRIs) achieves exactly the same security characteristics. The reason is that, after all validations, the neighboring AS will be the same as the left-most AS in the AS-PATH for the unicast route, and the left-most AS in the AS_PATH for the unicast route will be the same as the left-most AS in the AS_PATH for the Flow Specification NLRI. Therefore, the neighboring AS will be the same as the left-most AS in the AS_PATH for the Flow Specification NLRI (as the original AS_PATH validation rule in [RFC8955] states).

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 BGP security threat, already present in [RFC8955], but out of the scope of this document.

Using the new rule to validate a Flow Specification route received from a peer belonging to the same Local Domain is out of the scope of this document. Note that although it's possible, its utility is dubious. Although it is conceivable that a router in the same Local Domain could send a rogue update, only eBGP risk is considered within this document (in the same spirit as the aforementioned 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 unicast routes and Flow Specification routes. By congruent topology, it is understood that the two routes (i.e., the Flow Specification route and its best-match unicast route) are learned from the same peer across the AS. That would likely not be true, for instance, if some peers only negotiated one Address Family or if each Address Family peering had a different set of policies. Failing to have a congruent topology would result in step (b.1) of the validation procedure to fail.

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 following scenarios of a non-congruent topology without the second condition (b.2) being added to the validation procedure:

1. Consider a topology with two BGP speakers with two iBGP 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 confederation of ASes with local AS X and local AS Y (both belonging to the same Local Domain), and a given BGP speaker X1 inside local AS X. The ORIGINATOR_ID attribute is not advertised when propagating routes across local ASes. Let's assume the Flow Specification route is received from peer Y1 and the best-match unicast route is

received from peer Y2. Both peers belong to local AS Y. The Flow Specification validation procedure will also fail because of the first condition (b.1).

Consider now that the second condition (b.2) is added to the validation procedure. In the scenarios above, if Flow Specifications are originated in the same Local Domain, the AS_PATH will be empty or contain only an AS_CONFED_SEQUENCE segment. Condition (b.2) will evaluate 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 Specifications originated in a different Local Domain sill need a congruent topology. The reason is that in a non-congruent topology, the second condition (b.2) evaluates to false and only the first condition (b.1) is evaluated.

6. IANA Considerations

This document has no IANA actions.

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 described in [RFC8955] and, thus, security characteristics remain essentially equivalent to the existing security properties of BGP unicast routing, except as detailed below.

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

This document makes the original AS_PATH validation rule (Section 6.3 of [RFC4271]) again **OPTIONAL** (Section 4.2) for Flow Specification Address Family (the rule is no longer mandatory as had been specified by [RFC8955]). If that original rule is not enforced for Flow Specification, it may introduce some new security risks. A speaker in AS X peering with a route server 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. When the route server advertises the Flow Specification to a speaker in AS Z, it will be validated by that speaker. This risk is impossible to prevent if the Flow Specification route is received from a route server peer. If configuration (or other means beyond the scope of this document) indicates that the peer is not a route server, that optional rule **SHOULD** be enforced for unicast and/or for Flow Specification routes (as discussed in the Revision of AS_PATH Validation section, just enforcing it in one of those Address Families is enough). If the indication is that the peer is not a route server or there is no conclusive indication, that optional rule **SHOULD NOT** be enforced.

A route server itself may be in a good position to enforce the AS_PATH validation rule described in the previous paragraph. If it is known that a route server is not peering with any other route server, it can enforce the AS_PATH validation rule across all its peers.

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 destination prefix embedded in that Flow Specification. Note that this trustworthiness consideration is not absolute and the new possibility that an iBGP speaker could send a rogue Flow Specification is introduced.

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

It's worth mentioning that allowing (or making operationally feasible) Flow Specifications to originate within the Local Domain makes the network overall more secure. Flow Specifications can be originated more readily during attacks and improve the stability and security of the network.

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>.

[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>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[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>.

[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>.

8.2. Informative References

[CONFED-SET] Kumari, W., Sriram, K., Hannachi, L., and J. Haas, "Deprecation of AS_SET and AS_CONFED_SET in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-deprecate-as-set-confed-set-05, 12 March 2021, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-deprecate-as-set-confed-set-05>.

[RFC6472] Kumari, W. and K. Sriram, "Recommendation for Not Using AS_SET and AS_CONFED_SET in BGP", BCP 172, RFC 6472, DOI 10.17487/RFC6472, December 2011, <https://www.rfc-editor.org/info/rfc6472>.

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, Shyam Sethuram, Susan Hares, Alvaro Retana, and John Scudder for their review and comments.

Authors' Addresses

James Uttaro
AT&T
200 S. Laurel Ave
Middletown, NJ 07748
United States of America
Email: ju1738@att.com

Juan Alcaide
Cisco
Research Triangle Park
7100 Kit Creek Road
Morrisville, NC 27709
United States of America
Email: jalcaide@cisco.com

Clarence Filsfils
Cisco
Email: cf@cisco.com

David Smith
Cisco
111 Wood Ave South
Iselin, NJ 08830
United States of America
Email: djsmith@cisco.com

Pradosh Mohapatra
Sproute Networks
Email: mpradosh@yahoo.com