BESS Working Group
Internet-Draft
Intended status: Experimental
Expires: 12 June 2026

Y. Liu
Z. Zheng
ZTE
E. Metz
KPN
Y. Liu
China Mobile

9 December 2025

SRv6-based BGP Service Capability
draft-lz-bess-srv6-service-capability-07

Abstract

RFC9252 specifies that implementations MUST provide a mechanism to control advertisement of SRv6-based BGP service routes on a per neighbor and per service basis. This document provides analysis on the problems that may be encountered if the SRv6-based service routes are received by the MPLS-only PEs. Some currently used SRv6-based service routes advertisement controlling methods by configuration or network planning are also described. And this document proposes an automatic advertisement controlling method for SRv6-based service routes by defining a new Capability Code for SRv6-based BGP service capability.

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/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 12 June 2026.

Copyright Notice

Table of Contents

1. Introduction

[RFC9252] defines procedures and messages for SRv6-based services. When an egress PE is enabled for BGP Services over SRv6 data plane, it signals one or more SRv6 Service SIDs enclosed in SRv6 Service TLV(s) within the BGP Prefix-SID Attribute[RFC8669] attached to MP-BGP NLRIs. In other words, instead of defining new AFI/SAFIs for SRv6-based service routes, existing AFI/SAFIs of MPLS-based service routes are re-used for SRv6-based service routes.

As specified in [RFC9252], there're two options to encode SRv6 service SIDs in the route advertisement:

* The first option is to encode the whole SRv6 Service SID in the SRv6 Service TLV and set the MPLS Label field(s) of the corresponding NLRI to Implicit NULL.

* The second option, which is referred to as the Transposition Scheme, is to put the function and/or the argument part of the SRv6 SID in the MPLS Label field of the NLRI and to encode the locator part of the SID in the SRv6 Service TLV.

However, [RFC8669] specifies that unknown TLVs in the BGP Prefix attribute MUST be ignored and propagated unmodified. If SRv6-based service routes are received by PEs that are only capable of MPLS-based services, the PEs may discard SRv6 Services TLV in the BGP Prefix attribute and process these routes wrongly, which may leads to service failure and/or abnormal extra traffic flows in the network.

To avoid these problems, [RFC9252] specifies that implementations MUST provide a mechanism to control advertisement of SRv6-based BGP service routes on a per neighbor and per service basis.

This document provides analysis on the problems that may be encountered in the MPLS and SRv6 co-existence scenario if the SRv6-based service routes are received by the MPLS-only PEs. Some currently used SRv6-based service routes advertisement controlling methods by configuration or network planning are also described. And this document proposes an automatic advertisement controlling method for SRv6-based service routes by defining a new Capability Code [RFC5492] for SRv6-based BGP service capability.

2. 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 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. the Co-existence Scenario

In the progress of network upgrading, some of the legacy devices that only support MPLS/SR-MPLS will coexist with the new devices capable of SRv6 for a long time.

3.1. Possible Problems for MPLS-only PEs receiving SRv6 Service Routes

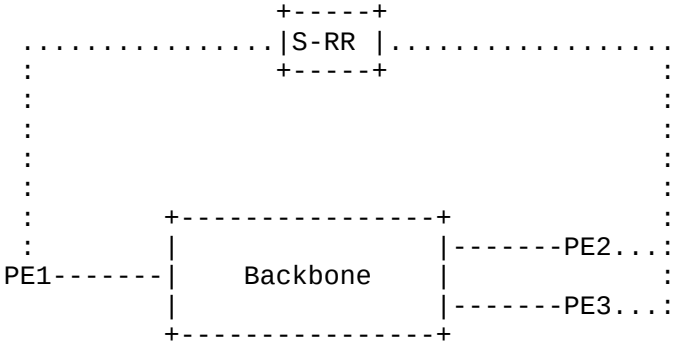

Figure 1: the Co-existence Scenario

As shown in Figure 1, PE1 is a legacy device that only supports MPLS-based services, PE2 supports both MPLS-based and SRv6-based services, and PE3 is an SRv6-only device. S-RR is a service route reflector that supports both MPLS and SRv6. On PE2, a SRv6 service SID sid-1 and a MPLS VPN route label label-1 are assigned for overlay service 1. On PE3, only SRv6-based service is enabled and configured for overlap service 2.

On PE2,the SRv6 service SID and a MPLS VPN route label for the service 1 are advertised in separate UPDATE messages. ADD-PATH[RFC7911] is used to avoid path hiding. S-RR reflects both SRv6-VPN route and MPLS-VPN route to PE1. Since PE1 only supports MPLS, it may discard the SRv6 Service TLV(s) in the BGP Prefix attribute and treat the SRv6-based route as a MPLS-based route for service 1, then there’re two MPLS-based routes for the same service 1 on PE1.

Depending on whether the Transposition Scheme is used, the following two scenarios are described separately.

* Scenario 1, the Transposition Scheme is used, the function and/or argument part of sid-1 is encoded in the MPLS Label field of the NLRI of the SRv6-based service route. Then PE1 may choose the route which is originally the SRv6-based route and use the label field in the NLRI of this route as MPLS VPN label for packet encapsulation.

  Unless the allocation of SRv6 SIDs and MPLS labels on PE2 is aligned to ensure compatibility, the interpretation of the function and/or argument of the SRv6 SID (sid-1 in the example) will lead to incorrect forwarding of the traffic. In the example above, on PE2 the packets may 1) be sent to the wrong service instance, in case the sid-1 function and/or argument value corresponds to an existing MPLS label, or 2) be dropped, in case the value of sid-1 does not correspond to an allocated MPLS label.

* Scenario 2, the entire sid-1 is encoded in the SRv6 Services TLV and the MPLS Label field of the corresponding NLRI is set to Implicit NULL. The SRv6 Services TLV in the UPDATE messages is discarded by PE1, and from PE1’s aspect, it has received a MPLS service route with an Implicit NULL label.

It should be noticed that how to deal with the MPLS-based route with an Implicit NULL label is not standardized, different vendors may have different processing procedures which are unpredictable, e.g, set the route to invalid, send the packet to service 1 without the service route label or something else.

On PE3, only SRv6 service SID sid-2 is configured for service 2. If the service routes from PE3 are received by PE1, the problems are similar.

* If the Transposition Scheme is used, PE1 may discard the SRv6 Service TLV(s) in the BGP Prefix attribute and treat the function and/or argument part of SRv6 service SID as a MPLS VPN route label. PE1 may 1) not send packets to PE2 since there's no LSP between PE1 and PE3 2) send packets encapsulated in IPv6 to PE3 if there's route to PE3.

* If the Transposition Scheme is not used and the label field in the NLRI is Implicit NULL, how PE1 deals with this route is unpredictable.

3.2. Some Current Methods for SRv6 Route Advertisement Controlling

[RFC9252] specifies that implementations MUST provide a mechanism to control advertisement of SRv6-based BGP service routes on a per neighbor and per service basis.

This can be done by configuration. First the network operator must obtain whether the PEs in the network are capable of SRv6-based services. Then the operator should config on PEs or route reflectors based on each PE's capability, the configuration is per neighbor.

* If there's a service route reflector, configurations on S-RR should ensure that the SRv6 service routes would not be reflected to MPLS-only legacy devices.

* If there's no route reflector in the network, which neighbors can the SRv6 service routes be advertised to should be specified when configuring SRv6 services on the PEs.

The above method may be feasible in small-scale networks, but are not applicable to large-scale networks. The main reasons are:

* The per neighbor configuration needs to be changed with the device capability. When a PE is upgraded to support SRv6-based services or rolled back to an older version that only supports MPLS, the configuration on its neighbors or the RR should be changed to add this PE to or exclude it from the advertisement of SRv6-based BGP

service routes. Although this may be done automatically by the network management system, it is still not a easy job in a large-scale network and is not flexible enough.

* The additional steps of device capability acquisition and capability based configuration increase the fault probability and troubleshooting difficulty. If the service from PE1 to PE2 fails, the operator needs to confirm the capability for SRv6-based service of the two devices, and then check the configuration on PE3 or RR to make sure that the SRv6-based service route is not advertised to PE1.

* There is no standard solution for the network operator to obtain the PE's capability for SRv6-based services. If there are devices from multiple vendors in the network, there may be interconnection problems.

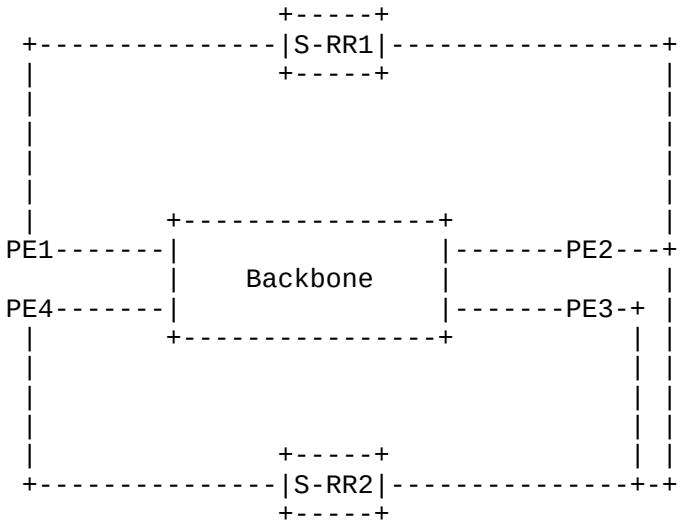

Figure 2: the Co-existence Example Topology 2

Some may implement service-RRs separately for MPLS and SRv6 when building the network. As shown in figure 2, S-RR1 is for MPLS service routes only and S-RR2 is for SRv6 service routes only. For MPLS-only PEs like PE1, they would only connect to S-RR1 and the situation is similar for SRv6-only PEs(e.g,PE3 and PE4). In this case, the configuration work is less than the scenario above, but,

* the configuration is at least required on dual-capability devices like PE2 to control the SRv6-based routes being advertised to the correct RR(i.e, S-RR2),

* the configuration on the PEs also needs to be changed to connect to the right S-RR if the PEs' capabilities of SRv6-based service routes change due to device upgrade.

4. SRv6-based BGP Service Capability

The basic idea is, if the BGP speaker can obtain the capability for SRv6-based services of its peers, the advertisement of SRv6-based BGP service routes can be automatically controlled.

[RFC5492] defines the "Capabilities Optional Parameter". A BGP speaker can include a Capabilities Optional Parameter in a BGP OPEN message. This allows BGP speakers to communicate capabilities. The Capabilities Optional Parameter is a triple that includes a one-octet Capability Code, a one-octet Capability length, and a variable-length Capability Value.

This document defines a Capability Code for SRv6-based BGP service capability. If a BGP speaker has not sent the SRv6-based BGP service capability in its BGP OPEN message on a particular BGP session, or if it has not received the SRv6-based BGP service capability in the BGP OPEN message from its peer on that BGP session, that BGP speaker MUST NOT send on that session any UPDATE message that includes the SRv6 service TLVs. Like any other BGP capabilities, if the capability for SRv6-based services is enabled or removed, an established session needs to be reset to resend the OPEN message.

In this way, the advertisement of SRv6-based BGP service routes is controlled without per neighbor or per-service configuration, which makes it easier to implement and manage in the network. In the co-existence scenario, the SRv6-based service routes would only be exchange between devices that support it based on this capability. There would be no UPDATE message that includes the SRv6 service TLV received by legacy devices.

PEs attached to the network, as BGP speakers, SHOULD indicate their ability to advertise and receive SRv6 based service routes through the SRv6 based BGP service capability. If service route reflectors are used in the network deploying SRv6-based services, they MUST support the SRv6-based BGP service capability if there're PEs in the network supporting this capability.

5. Operational Considerations

Considering that there’re already many implementations of [RFC9252] based on the method of per-neighbor configuration, only upgrading part of the SRv6-PEs/RRs in the network with the new BGP capability defined in this document would lead to interoperability problems since the SRv6-PEs/RRs that are not upgarded would not receive the SRv6-based service routes from the upgarded devices. A recommended method is to make this new BGP capability configurable on the devices, that is, to enable the function of this new BGP capability on the devices only when the network operator confirms that all the SRv6-PEs and the corresponding service RRs have been upgraded to support the new BGP capability.

6. IANA Considerations

This document defines a new Capability Codes option, named "SRv6 Service Capability" with an assigned value <TBD1> to indicate that a BGP speaker supports SRv6-based services. The length of this capability is 1.

7. Security Considerations

This extension to BGP does not change the underlying security issues inherent in [RFC5492] and [RFC9252].

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>.

[RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February 2009, <https://www.rfc-editor.org/info/rfc5492>.

[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

[RFC7911] Walton, D., Retana, A., Chen, E., and J. Scudder, "Advertisement of Multiple Paths in BGP", RFC 7911, DOI 10.17487/RFC7911, July 2016, <https://www.rfc-editor.org/info/rfc7911>.

[RFC8669] Previdi, S., Filsfils, C., Lindem, A., Ed., Sreekantiah, A., and H. Gredler, "Segment Routing Prefix Segment Identifier Extensions for BGP", RFC 8669, DOI 10.17487/RFC8669, December 2019, <https://www.rfc-editor.org/info/rfc8669>.

[RFC9252] Dawra, G., Ed., Talaulikar, K., Ed., Raszuk, R., Decraene, B., Zhuang, S., and J. Rabadan, "BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)", RFC 9252, DOI 10.17487/RFC9252, July 2022, <https://www.rfc-editor.org/info/rfc9252>.

Authors' Addresses

Yao Liu
ZTE
Nanjing
China
Email: liu.yao71@zte.com.cn

Zhang Zheng
ZTE
Nanjing
China
Email: zhang.zheng@zte.com.cn

Eduard Metz
KPN
Netherlands
Email: etmetz@gmail.com

Yisong Liu
China Mobile
China
Email: liuyisong@chinamobile.com