BESS
Internet Draft
Intended status: Standards Track
Expires: August 24, 2023
Yisong Liu
China Mobile
C. Lin
M. Chen
New H3C Technologies
February 24, 2023

SRv6 Service SID Flag Extension for Multi-homed SRv6 BGP Services
draft-liu-bess-multihome-srv6-service-sid-flag-00

Abstract

In some multi-homed SRv6 L3VPN and EVPN scenarios, there are requirements for the egress PE to advertise multiple SRv6 Service SIDs for the same service, such as anycast Service SID and bypass Service SID. This document defines anycast flag and bypass flag for SRv6 Service SIDs carried in BGP messages.

Table of Contents

1. Introduction

[RFC9252] defines procedures and messages for SRv6-based BGP services, including Layer 3 Virtual Private Network (L3VPN), Ethernet VPN (EVPN), and Internet services. In some multi-homed scenarios, there are requirements for the egress PE to advertise multiple SRv6 Service SIDs for the same service, such as anycast Service SID and bypass Service SID. And those SIDs need to be identified in the BGP messages.

This document defines anycast flag and bypass flag for SRv6 Service SIDs carried in BGP messages.

1.1. 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. Use Case

2.1. Anycast SRv6 Service SID

In the multi-homed SRv6 L3VPN and EVPN scenarios, anycast Service SID may be used to advertise the same service at different egress PEs, which can improve service reliability and load balancing.

```
             +-----+             +-----+
             | CE1 |             | CE2 |
             +-----+             +-----+
                |                   |
             +-----+             +-----+
  ---------- | PE1 |             | PE2 |
      ^      +-----+             +-----+
      |             *           *
      |              *         *
    SRv6              +-------+
  L3VPN/EVPN          |BGP-RR |
      |               +-------+
      |              *         *
      |             *           *
      v      +-----+             +-----+
  ---------- | PE3 |             | PE4 |
             +-----+             +-----+
      1. Anycast    \           /  1. Anycast
       Service SID   \         /    Service SID
      2. Unicast      \       /    2. Unicast
       Service SID-1   +-----+      Service SID-2
                       | CE3 |
                       +-----+
```

Figure 1

As shown in Figure 1, PE3 and PE4 use the same anycast SRv6 Service SID for the VPN service of CE3. The ingress PE1 encapsulates the payload in an outer IPv6 header where the destination address is that anycast SRv6 Service SID. The packets from CE1 can reach CE3 through either PE3 or PE4. Assume that the path from PE1 to PE3 and the path from PE1 to PE4 have the same cost. The traffic flows will be load balanced between PE3 and PE4.

PE3 and PE4 also have unicast SRv6 Service SIDs, which are SID-1 and SID-2, for the VPN service of CE3. The ingress PE2 uses SID-1 as the primary SRv6 Service SID, and SID-2 as backup. The packets from CE2 will be forwarded to CE3 through PE3. If any failure occurs on the path to PE3, service will be switched to PE4.

Since ingress PE1 and PE2 have different strategies for the control of VPN traffics, egress PE3 and PE4 each need to advertise two SRv6 Service SIDs, an anycast SID for ingress PE1 and a unicast SID for ingress PE2. Local export policy may be used by egress PE3 and PE4 to control which SID is advertised to ingress PE1 and which is advertised to ingress PE2. However, if BGP Route Reflector is deployed, both the anycast Service SID and the unicast Service SID will be advertised to RR and reflected to ingress PEs, and the receiver has to choose which Service SID to use. In this case, it is required to identify which Service SID is anycast and which Service SID is unicast, when both two SIDs are advertised in BGP messages.

2.2. Bypass SRv6 Service SID

In the multi-homed SRv6 L3VPN and EVPN scenarios, two egress PEs may establish a bypass path between them and use it as the protection of PE-CE link failure.

As shown in Figure 2, PE2 and PE3 each has a normal SRv6 Service SID and bypass SRv6 Service SID for the L3VPN service of CE2.

The ingress PE1 encapsulates the payload in an outer IPv6 header where the destination address is the normal SRv6 Service SID. The packets from CE1 can reach CE2 through either PE2 or PE3.

Assume that PE2 is the primary egress PE, and PE3 is the backup one. If the link between PE2 and CE2 fails, the packets are still forwarded to PE2 before PE1 recalculates BGP routes. So, PE2 should forward the packets through the bypass path to PE3. Along the bypass path, the packets are steered by the bypass SRv6 Service SID of PE3.

The routes for the SRv6 Service SIDs are as following. Note that the bypass Service SID has no local backup protection, in order to avoid routing loops between PE2 and PE3 when their CE side links fail at the same time.

```
Routes on PE2:
  SID-21
    Primary Next-hop: CE2
    Backup Next-hop: Service SRv6 SID-32
  SID-22 (Bypass)
    Primary Next-hop: CE2

Routes on PE3:
  SID-31
    Primary Next-hop: CE2
    Backup Next-hop: Service SRv6 SID-22
  SID-32 (Bypass)
    Primary Next-hop: CE2
```

So, the egress PE needs to advertise two SRv6 Service SIDs, a normal SID for the ingress PE and a bypass SID for the other egress PE. Local export policy may be used to control which SID is advertised to ingress PE and which is advertised to the other egress PE. However, if BGP Route Reflector is deployed, both the normal Service SID and the bypass Service SID will be advertised to RR and reflected to other PEs, and the receiver needs to choose which Service SID to use. In this case, it is required to identify which Service SID is for bypass purpose, when both two SIDs are advertised in BGP messages.

```
                          +-----+
                          | CE1 |
                          +-----+
                             |
                          +-----+
      ------------------- | PE1 |**************
         ^                +-----+             *
         |               /       \            *
         |              /         \        +------+
     SRv6 L3VPN/EVPN   / **********\********|BGP-RR|
         |            / *           \       +------+
         |           / *             \         *
         v      +-----+   Bypass    +-----+    *
      --------- | PE2 |-------------| PE3 |*****
                +-----+    Path     +-----+
       1. Normal       \           /  1. Normal
        Service SID-21  \         /    Service SID-31
       2. Bypass         \       /    2. Bypass
        Service SID-22    +-----+      Service SID-32
                          | CE2 |
                          +-----+
```

Figure 2

3. Extensions for BGP

[RFC9252] defines the SRv6 SID Information Sub-TLV to carry SRv6 Service SID in BGP messages. Its encoding is as following:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| SRv6 Service  |    SRv6 Service               |               |
| Sub-TLV       |    Sub-TLV                    |               |
| Type=1        |    Length                     |  RESERVED1    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  SRv6 SID Value (16 octets)                                  //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Svc SID Flags |   SRv6 Endpoint Behavior      |   RESERVED2   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  SRv6 Service Data Sub-Sub-TLVs                              //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

This document defines two new flags in the SRv6 Service SID Flags field:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|A|B|           |
+-+-+-+-+-+-+-+-+
```

o A-flag: Anycast flag. When set, the associated SID is anycast.

o B-flag: Bypass flag. When set, the associated SID is for bypass usage, without local backup protection.

The new-defined flags can be used for the SRv6 Service SIDs of L3 and L2 services, such as End.DX4, End.DT4, End.DX6, End.DT6, End.DT46. End.DX2, End.DX2V, End.DT2U, End.DT2M, etc.

4. Backward Compatibility

According to [RFC9252],

o Any unknown flags in the SRv6 Service SID Flags field MUST be ignored by the receiver.

o When multiple SRv6 SID Information Sub-TLVs are present, the ingress PE SHOULD use the SRv6 SID from the first instance of the Sub-TLV.

If there are PE routers not supporting the flags defined in this document, the egress PE may expect those routers to use the first SID and ignore the new-defined flags.

5. Security Considerations

TBD.

6. IANA Considerations

This document defines the following bits in the SRv6 Service SID Flags field of SRv6 SID Information Sub-TLV:

TLV Code Point	Value
TBD	A-flag
TBD	B-flag

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017

[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

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

Changwang Lin
New H3C Technologies
China
Email: linchangwang.04414@h3c.com

Mengxiao Chen
New H3C Technologies
China
Email: chen.mengxiao@h3c.com