SPRING Working Group W. Cheng
Internet Draft China Mobile
Intended status: Standards Track C. Lin
Expires: September 14, 2023 New H3C Technologies
March 13, 2023

Service Interworking between SRv6
draft-cheng-spring-service-interworking-srv6-00

Abstract

When operators provide services through SRv6, such as L3VPN and L2VPN, there may be cross-domain scenarios of multiple ASs, or multiple admin domain scenarios within the same AS. This document describes how to implement interworking of services for such scenarios.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

When operators begin to deploy SRv6, they cannot deploy a single SRv6 domain due to the original underlay network planning, or due to management considerations

Different ASs may belong to different SRv6 domains, or the same AS may be divided into multiple SRv6 domains. Between SRv6 domains, locator routes are not advertised to each other. When providing

services to customers, cooperation between multiple SRv6 domains is required to provide end-to-end services.

This document describes how to achieve interworking between SRv6 domains, in such scenarios when VPN services (L3VPN or L2VPN) are provided by the SRv6 service SID as per [I-D.ietf-bess-srv6-services].

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. Scenarios of Inter-domain interworking

When an operator provides VPN services, its transport network may contain multiple ASs. Due to the IPv6 feature of SRv6, BGP neighbors can be directly established between PEs and VPN routes can be advertised. Locator routes are advertised between ASs, or public network tunnels are established through SRv6 Policy to implement inter-AS forwarding based on SRv6 BE or SRv6 Policy.

Due to historical or practical reasons, operators may not be able to implement this SRv6 cross-domain solution. When VPN information is restricted within the AS, the cross-domain solution of SRv6 needs to be considered.

Referring to the Section 10 of [RFC4364], there are three ways to provide VPN service through BGP/MPLS, namely OptionA/B/C. When operators deploy VPN services through SRv6, there are also three cross-domain VPN ways.

Referring to the topology in the figure below, taking the service traffic as IPv4 as an example, the following section describe the three SRv6 cross-domain methods respectively.

```
              +--------+        +--------+        +--------+
              |  AS1   |        |  AS2   |        |  AS3   |
   +---+   +---+     +---+    +---+    +---+    +---+    +---+    +---+
   |CE1+---+PE1+=====+PE2+----+PE3+====+PE4+----+PE5+====+PE6+----+CE2|
   +---+   +---+     +---+    +---+    +---+    +---+    +---+    +---+
              |        |        |        |        |        |
              +--------+        +--------+        +--------+
                    ASBR1     ASBR2    ASBR3     ASBR4
```

Figure 1: reference topology for inter-domain

2.1. Option A (VRF-to-VRF)

In this way, the PE router as ASBR of one AS is directly connected to the PE router of another AS.

The two PE routers will be attached by multiple sub-interfaces, and associate each such sub-interface with a VRF. Each PE will treat the other as if it were a CE router.

iBGP neighbors are established between PEs in the same AS, and VPN routes are advertised. eBGP neighbors are established between ASBRs of the adjacent AS, and IPv4 unicast routes are advertised.

2.1.1. SRv6 BE

For SRv6 BE forwarding, the single-domain and cross-domain processing are the same, and only route advertisement and SRv6 forwarding are completed within each AS.

Take PE6 to advertise VPN routes to PE1 as an example, the route advertisement process is as follows:

o Each AS internally advertises the locator routes of each Endpoint through IGP

o @PE6 assigns VPNSID1 (End.DT4 segment) to it after learning the VPN route. Then advertise the VPN route and VPNSID1 to @ASBR4 via iBGP

o @ASBR4 learns the VPN route and VPNSID1 in the corresponding VPN instance routing table, and advertises it as an IPv4 unicast route to @ASBR3 through eBGP.

- @ASBR3 regards @ASBR4 as its own CE device, adds the routes learned through eBGP to the routing table of the corresponding VPN instance, and assigns VPNSID2 to it. Then advertise the VPN route and VPNSID2 to @ASBR2 via iBGP.

- @ASBR2 behaves like @ASBR4 and advertises VPN routes as IPv4 unicast routes to @ASBR1 via eBGP.

- @ASBR1 regards @ASBR2 as its own CE device, adds the routes learned through eBGP to the routing table of the corresponding VPN instance, and assigns VPNSID3 to it. Then advertise the VPN route and VPNSID3 to @PE1 via iBGP.

- @PE1 learns the VPN route and VPNSID3 in the corresponding VPN instance routing table.

```
             iBGP        eBGP       iBGP       eBGP       iBGP
           +------+    +-----+   +-------+   +------+   +------+
          /        \  /       \ /         \ /        \ /        \
         PE1-------PE2--------PE3---------PE4--------PE5--------PE6
          |       (ASBR1)   (ASBR2)     (ASBR3)     (ASBR4)     |
          |<--------|<---------|<----------|<---------|<--------|
          |  VPNv4  |   IPv4   |   VPNv4   |   IPv4   |  VPNv4  |
          |  route  |   route  |   route   |   route  |  route  |
```

Figure 2: process of route advertisement for SRv6 BE in option A

VPN traffic is forwarded through SRv6 within AS, and forwarded between ASBRs through IPv4 forwarding.

Taking the packet sent from CE1 to CE2 as an example, the packet forwarding process in SRv6 BE mode is as follows:

- @PE1 searches the routing table in the corresponding VPN after receiving the packet from CE1. Add IPv6 encapsulation to the original packet, the IPv6 destination address is VPNSID3, and forward the packet to @ASBR1.

- @ASBR1 removes the outer IPv6 encapsulation, searches the routing table bound to VPNSID3, and forwards the original packet to @ASBR2 according to the search result.

- @ASBR2 receives the packet, it adds IPv6 encapsulation to the packet, after receives the packet, the outer IPv6 destination address is VPNSID2, and forwards the packet to @ASBR3, similar to the processing process of PE1.

- @ASBR3 removes the outer IPv6 encapsulation after receiving the packet, searches the routing table bound to VPNSID2 for the route, and forwards the original packet to @ASBR4 according to the search result.

- @ASBR4 adds IPv6 encapsulation to the packet after receives the packet, the outer IPv6 destination address is VPNSID3, and forwards the packet to @PE6, similar to the processing process of PE1.

- @PE6 removes the outer IPv6 encapsulation after receiving the packet, searches the routing table bound to VPNSID3 for the route, and forwards the original packet to CE2 according to the search result.

```
            +---------+         +---------+         +--------+
            |   AS1   |         |   AS2   |         |  AS3   |
   CE1-----PE1-------PE2-------PE3-------PE4-------PE5------PE6----CE2
                    (ASBR1)   (ASBR2)   (ASBR3)   (ASBR4)

            +-------+           +-------+            +-------+
            | IPv6  |           | IPv6  |            | IPv6  |
            |VPNSID3|           |VPNSID2|            |VPNSID1|
   +-----+  +-------+  +-----+  +-------+  +-----+   +-------+  +-----+
   |/////|->|///////|->|/////|->|///////|->|/////| ->|///////|->|/////|
   +-----+  +-------+  +-----+  +-------+  +-----+   +-------+  +-----+
```

Figure 3: Process of forwarding for option A BE

2.1.2. SRv6 TE

For SRv6 TE of Option A, when packets are forwarded within each AS, SRH is encapsulated on the ingress PE and decapsulated on the egress PE. Neither the control plane routing nor the forwarding plane involves inter-AS interoperability.

2.1.3. Summary of Option A

Implementing SRv6 cross-domain forwarding through Option A has no special functional requirements for ASBR and PE nodes. This document only describes the main workflow of Option A.

2.2. Option B

For Option B, the interfaces between ASBRs of different ASs do not need to be bound to a VPN, and the VPN routes are republished between ASBRs through eBGP. Between the ingress and egress PEs, multi-segment tunnels from PE to ASBR, ASBR to ASBR, and ASBR to PE need to be established to guide traffic forwarding. There is a difference in processing for BE and TE of SRv6.

2.2.1. SRv6 BE

In the SRv6 BE mode, only one IPv6 encapsulation is added to the VPN traffic, and the VPN traffic is forwarded to the egress PE through the IPv6-encapsulated destination address (VPNSID).

For Option B, traffic can only be forwarded within the domain through the destination address. Therefore, when the ASBR republishes the VPN route, a new segment needs to be created locally, and the VPNSID of the VPN route needs to be advertised to the PE in the AS or the ASBR of other ASes. The new segment leads the traffic to be forwarded to the current ASBR. At the same time, the new segment needs to be associated with the original VPNSID, which is used for replacement during forwarding and directs the traffic to the next ASBR.

Take PE6 to advertise VPN routes to PE1 as an example, the route advertisement process is as follows

o @PE6 assigns VPNSID1 (End.DT4 segment) to it after learning the VPN route. Then advertise the VPN route and VPNSID1 to ASBR4 via iBGP.

o @ASBR4 learns the VPN route in the corresponding VPN instance routing table, and assigns a segment SID2. ASBR4 associates SID2 with VPNSID1. SID2 can be a segment of a new behavior, or a newly defined flavor for a segment of End. Its definition and specific behavior will be described in subsequent versions. @ASBR4 advertises VPN route and SID2 to ASBR3 via eBGP.

o @ASBR3 stores the VPN routes received from eBGP neighbors in the corresponding VPN instance routing table, and assigns a SID3 to associate with SID2. Continue to advertise VPN routes and SID3 to ASBR2 via iBGP.

o @ASBR2 behaves like @ASBR4, newly assigns SID4 to associate with SID3, and advertises VPN route and SID4 to @ASBR1 via eBGP.

- @ASBR1 behaves like @ASBR3, newly assigns SID5 to associate with SID4, and advertises VPN route and SID5 to @PE1 via iBGP.

- @PE1 learns the VPN route and VPNSID (SID5) in the corresponding VPN instance routing table.

```
               iBGP        eBGP       iBGP       eBGP       iBGP
            +------+    +-----+   +-------+   +------+   +------+
           /        \  /       \ /         \ /        \ /        \
         PE1-------PE2--------PE3---------PE4--------PE5--------PE6
                      (ASBR1)   (ASBR2)     (ASBR3)     (ASBR4)
           |<--------|<---------|<----------|<---------|<--------|
           |  VPNv4  |   VPNv4  |   VPNv4   |   VPNv4  |  VPNv4  |
           |  route  |   route  |   route   |   route  |  route  |
                   SID5       SID4        SID3       SID2   VPNSID1
```

Figure 4: process of route advertisement for SRv6 BE In option B

Taking the packet sent from CE1 to CE2 as an example, the packet forwarding process in SRv6 BE mode is as follows:

- @PE1 searches the routing table in the corresponding VPN after receiving the packet from CE1. Then add IPv6 encapsulation to the original packet, and the outer IPv6 destination address is SID5. The encapsulated packet is forwarded to ASBR1.

- @ASBR1 finds the SID4 associated with it through SID5 after receiving the packet, replaces the destination address of the packet with SID4, and forwards the packet to ASBR2.

- @ASBR2 finds the SID3 associated with it through SID4 after receiving the packet, replaces the destination address of the packet with SID3, and forwards the packet to ASBR3.

- @ASBR3 finds the SID2 associated with it through SID3 after receiving the packet, replaces the destination address of the packet with SID2, and forwards the packet to ASBR4.

- @ASBR4 finds the VPNSID1 associated with it through SID2 after receiving the packet, replaces the destination address of the packet with VPNSID1, and forwards the packet to PE6.

- @PE6 removes the outer IPv6 encapsulation after receiving the packet, searches for the route in the routing table bound to VPNSID1, and forwards the original packet to CE2 according to the search result.

```
            +---------+         +---------+         +--------+
            |   AS1   |         |   AS2   |         |  AS3   |
   CE1-----PE1-------PE2-------PE3-------PE4-------PE5------PE6----CE2
                    (ASBR1)   (ASBR2)   (ASBR3)   (ASBR4)

            +------+  +------+  +------+  +------+   +-------+
            | IPv6 |  | IPv6 |  | IPv6 |  | IPv6 |   | IPv6  |
            | SID5 |  | SID4 |  | SID3 |  | SID2 |   |VPNSID1|
   +-----+  +------+  +------+  +------+  +------+   +-------+  +-----+
   |/////|->|//////|->|//////|->|//////|->|//////| ->|///////|->|/////|
   +-----+  +------+  +------+  +------+  +------+   +-------+  +-----+
```

Figure 5: Process of forwarding for option B BE

2.2.2. SRv6 TE

For Option B, due to its deployment mode, there is usually no cross-domain controller, so an end-to-end SRv6 Policy cannot be created on the ingress PE. It is necessary to plan the path (segment list) independently according to the SLA requirements in each AS.

The PE needs to iterate the VPNSID to the segment list of the current AS.ASBR needs to be able to associate the segment lists on the left and right sides of itself

When forwarding VPN traffic, the paths passing through the AS need to be assembled to generate end-to-end paths between ingress and egress PEs.

Take PE6 to advertise VPN routes to PE1 as an example, the route advertisement process is as follows:

o @PE6 assigns VPNSID1 (End.DT4 segment) to it after learning the VPN route. Then advertise the VPN route and VPNSID1 to ASBR4 through iBGP, and the next hop address is the address of @PE6.

o @ASBR4 first learns the VPN route in the corresponding VPN instance routing table. ASBR4 then creates a segment list1 destined for PE6, uses VPNSID1 and PE6 addresses as the index of the segment list, and assigns a bindingSID (BSID1) to the segment list at the same time. If the corresponding segment list (with the same index) already exists, its bindingSID (BSID1) is used directly. ASBR4 advertises the VPN route, VPNSID1, and BSID1 to @ASBR3 through eBGP, and modifies the next hop to the address of ASBR4.

- @ASBR3 learns the VPN route in the routing table of the corresponding VPN instance, and then uses the addresses of BSID1 and ASBR4 as indexes to create a segment list2 destined for ASBR4. The list only contains the EPESIDs destined for ASBR4, and assigns BSID2 to segment list2. ASBR3 associates BSID2 with BSID1, then advertises the VPN route, VPNSID1 and BSID2 to ASBR2 through iBGP, and modifies the next hop to the address of ASBR3.

- @ASBR2 behaves like ASBR4, creates segment list3, assigns BSID3 to it, and associates it with BSID2. Then, the VPN route, VPNSID1 and BSID3 are advertised to ASBR1 through eBGP, and the next hop is changed to the address of ASBR2.

- @ASBR1 behaves similarly to ASBR3, creating segment list4 that only contains EPESID2 to ASBR2. BSID4 is allocated and associated with BSID3. Finally, the VPN route, VPNSID1 and BSID4 are advertised to PE1 through iBGP, and the next hop is changed to the address of ASBR1.

- @PE1 behaves like ASBR4, creates segment list5 to ASBR1, assigns it BSID5 and associates it with BSID4. Finally, PE1 stores BSID5 as the next hop of the newly learned VPN route in the VPN instance routing table.

BSID5/BSID4/BSID3/BSID2 are segments that need a new definition, temporarily named End.B6R for identification. Similar to End.B6, this type of segment is bound to a segment list, but is also associated with another segment.

When forwarding a message, if the destination address of the received message is a locally instantiated End.B6R segment, the SHR.SL field is not updated, but the End.B6R segment in the SRH is replaced with the associated segment. And continue to use the segment list bound by End.B6R to forward packets.

The specific definition and detailed description of End.B6R will be added in subsequent editions of this document.

```
             iBGP        eBGP       iBGP       eBGP       iBGP
           +------+    +-----+   +-------+   +------+   +------+
          /        \  /       \ /         \ /        \ /        \
         PE1-------PE2--------PE3---------PE4--------PE5--------PE6
                 (ASBR1)   (ASBR2)     (ASBR3)     (ASBR4)
          |<--------|<---------|<----------|<---------|<--------|
          |  VPNv4  |   VPNv4  |   VPNv4   |   VPNv4  |  VPNv4  |
          |  route  |   route  |   route   |   route  |  route  |
         BSID5     BSID4     BSID3       BSID2      BSID1   VPNSID1
```

Figure 6: process of route advertisement for SRv6 TE in option B

Taking the packet sent from CE1 to CE2 as an example, the packet forwarding process in SRv6 TE mode is as follows:

o After @PE1 receives the packet from CE1, it searches the routing table in the corresponding VPN. The next hop and service SID of the corresponding VPN route are BSID5 and VPNSID1, respectively. PE1 adds SRv6 encapsulation to the original packet. The segment list in the SRH is <BSID5, VPNSID1>, and the destination address of the outer IPv6 header is BSID5. Since BSID5 is the local segment of PE1, it continues to process the packet on PE1.

o @PE1 replaces BSID5 in SRH with BSID4 associated with BSID5, and modifies the destination address to BSID4. Use segment list5 associated with BSID5 to forward packets. Add IPv6 and SRH to the packet, and encapsulate segment list5 in the SRH. Forward the packet in AS1 to ASBR1

o Before the packet reaches ASBR1, the outer IPv6 and SRH may have been de-encapsulated by the penultimate hop, or the outer encapsulation may have been de-encapsulated by ASBR1. ASBR1 continues to process the packet whose outer encapsulation has been de-encapsulated, and the destination address of the packet is BSID4 at this time. ASBR1 replaces BSID4 in the SRH with BSID3 associated with BSID4, and modifies the IPv6 destination address to BSID3. ASBR1 continues to use segment list4 associated with BSID4 to forward packets. Since there is only one EPESID in segment list4 and it is a segment of End.x type, there is no need to add encapsulation, and the packet is forwarded to ASBR2 according to the EPESID.

- After @ASBR2 receives the packet, the destination address of the packet is now BSID3. ASBR2 replaces BSID3 in the SRH with BSID2 associated with BSID3, and modifies the IPv6 destination address to BSID2. ASBR2 continues to use segment list3 associated with BSID3 to forward packets, adds IPv6 and SRH to the packets, and encapsulates segment list3 in SRH. The packet is forwarded in AS2 to ASBR3.

- The behavior of @ASBR3 is similar to that of ASBR1. The destination address of the packet after removing the outer encapsulation is BSID3, the destination address of the continued packet is updated to BSID1, and the packet is forwarded to ASBR4 according to the EPESID.

- After @ASBR4 receives the packet, the destination address of the packet is BSID1, and BSID1 is a normal bindingSID. Therefore, ASBR4 performs the normal bindingSID forwarding behavior, updates SHR.SL, and updates the destination address of the packet to VPNSID1. ASBR4 forwards the packet according to the segment list1 associated with BSID1, adds IPv6 and SRH to the packet, and encapsulates segment list1 in the SRH. The packet is forwarded to PE6 in AS3.

- After receiving the packet, @PE6 removes the SRv6 encapsulation, searches for the route in the routing table bound to VPNSID1, and forwards the original packet to CE2 according to the search result.

```
            +---------+         +---------+         +--------+
            |   AS1   |         |   AS2   |         |  AS3   |
   CE1-----PE1-------PE2-------PE3-------PE4-------PE5------PE6----CE2
                    (ASBR1)   (ASBR2)   (ASBR3)   (ASBR4)

          +-------+             +-------+             +-------+
          | IPv6  |             | IPv6  |             | IPv6  |
          +-------+             +-------+             +-------+
          |  SRH  |             |  SRH  |             |  SRH  |
          |segment|             |segment|             |segment|
          |list5, |             |list3, |             |list1, |
          +-------+  +-------+  +-------+  +-------+  +-------+
          | IPv6  |  | IPv6  |  | IPv6  |  | IPv6  |  | IPv6  |
          +-------+  +-------+  +-------+  +-------+  +-------+
          |  SRH  |  |  SRH  |  |  SRH  |  |  SRH  |  |  SRH  |
          | SL= 1 |  | SL = 1|  | SL = 1|  | SL = 1|  | SL = 0|
          |VPNSID1|  |VPNSID1|  |VPNSID1|  |VPNSID1|  |VPNSID1|
          | BSD4  |  | BSID3 |  | BSID2 |  | BSID1 |  | BSID1 |
   +---+  +-------+  +-------+  +-------+  +-------+  +-------+  +---+
   |///|->|///////|->|///////|->|///////|->|///////|->|///////|->|///|
   +---+  +-------+  +-------+  +-------+  +-------+  +-------+  +---+
```

Figure 7: Process of forwarding for SRv6 TE in option B

2.2.3. Summary of Option B

For packets forwarded in SRv6 BE mode, since only IPv6 encapsulation is added to service traffic, the destination IPv6 address (VPNSID) is used to guide traffic to the egress PE. To ensure that the VPNSID is reachable, the ASBR needs to replace the original VPNSID with the reachable SID of the AS when republishing the VPN route. During the forwarding process, the ASBR at the AS boundary needs to replace the destination IPv6 address of the packet.

For packets forwarded in SRv6 TE mode, the forwarding logic is different from that of diverting VPN traffic to the corresponding SRv6 Policy based on color. In Option B mode, the processing logic of forwarding packets in SRv6 TE mode is similar to that of SRv6 BE, except that special processing is added to iterate BE forwarding to segment lists.

For ASBR and ingress PE, it behaves differently for BE and TE

o For SRv6 BE: ASBR needs to allocate a new SID, associate the original VPNSID, and replace the original VPNSID with the newly allocated SID when republishing VPN routes

- For SRv6 TE: ASBRs and ingress PEs need to create segment lists and assign BSIDs with special behaviors to them. And when the ASBR republishes the VPN route, it needs to advertise the BSID and the original VPNSID at the same time. Therefore, a special TLV needs to be added to carry the TSID, and the related extensions are described in subsequent versions of this document.

2.3. Option C

For Option C, through multi-hop EBGP, the egress PE directly advertises the VPN route and VPNSID to the ingress PE in other AS.

2.3.1. SRv6 BE

In the SRv6 BE scenario, for Option C, the ASBR needs to advertise the locator of the egress PE to the AS where the ingress PE is located, so that the ingress PE can learn the locator route of the egress PE, and the VPNSID is reachable to the ingress PE.

The Locator network segment can be planned for the entire network, and the ASBR can be configured to aggregate routes before advertising to reduce the number of other AS routes.

Take PE6 advertises VPN routes to PE1 as an example, the route advertisement process is as follows:

1. Advertising locator route

- @PE6 advertises its own locator route to @ASBR4 via IGP or iBGP

- After @ASBR4 learns the locator route of PE6, it advertises the locator route of PE6 to ASBR3 through eBGP, and specifies the next hop as ASBR4.

- After @ASBR3 receives the locator route, it advertises the locator route of PE6 to ASBR2 through IGP or iBGP, and specifies the next hop as ASBR3

- After @ASBR2 learns the locator route of PE6, it advertises the locator route of PE6 to ASBR1 through eBGP, and specifies the next hop as ASBR2

- After @ASBR1 receives the locator route, it advertises the locator route of PE6 to PE1 through IGP or iBGP, and specifies the next hop as ASBR1

- @PE1 learns the locator route to PE6 and iterates to the real next hop according to the route.

2. Advertising VPN route

- @PE6 assigns VPNSID1 (End.DT4 segment) to it after learning the VPN route. Then advertise the VPN route and VPNSID1 to PE1 through eBGP, and the next hop address is the IP address of PE6.

- @PE1 learns the VPN route and VPNSID1 in the corresponding VPN instance routing table, and iterates the real next hop through the learned locator route

```
                              Multi-hop EBGP
           +----------------------------------------------------+
          /             eBGP                     eBGP            \
         /           +---------+              +--------+          \
        /           /           \            /          \          \
      PE1---------PE2---------PE3---------PE4---------PE5---------PE6
       |        (ASBR1)     (ASBR2)     (ASBR3)     (ASBR4)        |
       |           |           |           |           |           |
       |    IGP    |           |    IGP    |           |    IGP    |
       |<-locator->|<-locator->|<-locator->|<-locator->|<-locator->|
       |  route    |   route   |   route   |   route   |  route    |
       |                                                           |
       |  <-------------------- VPNv4 route ------------------->   |
```

Figure 8: process of route advertisement for SRv6 BE in option C

Taking the packet sent from CE1 to CE2 as an example, the packet forwarding process in SRv6 BE mode is as follows:

- After receiving the packet from CE1, @PE1 searches the routing table in the corresponding VPN. PE1 adds an IPv6 header to the original packet, and the destination address is VPNSID1. According to the locator route of PE6, forward the packet to ASBR1

- After receiving the packets, @ASBR1, @ASBR2, @ASBR3, and @ASBR4 all forward the packets according to the locally learned locator route of PE6.

- After receiving the packet, @PE6 removes the outer IPv6 encapsulation, searches for the route in the routing table bound to VPNSID1, and forwards the original packet to CE2 according to the search result.

```
            +---------+         +---------+         +--------+
            |   AS1   |         |   AS2   |         |  AS3   |
   CE1-----PE1-------PE2-------PE3-------PE4-------PE5------PE6----CE2
                    (ASBR1)   (ASBR2)   (ASBR3)   (ASBR4)

          +-------+  +-------+  +-------+  +-------+  +-------+
          | IPv6  |  | IPv6  |  | IPv6  |  | IPv6  |  | IPv6  |
          |VPNSID1|  |VPNSID1|  |VPNSID1|  |VPNSID1|  |VPNSID1|
   +---+  +-------+  +-------+  +-------+  +-------+  +-------+  +----+
   |///|->|///////|->|///////|->|///////|->|///////|->|///////|->|////|
   +---+  +-------+  +-------+  +-------+  +-------+  +-------+  +----+
```

Figure 9: Process of forwarding for SRv6 BE in option C

2.3.2. SRv6 TE

For Option C mode, the AS is usually divided to control the scope of the IGP, and multiple ASs are in the same management domain. It is therefore possible to deploy cross-domain controllers, or hierarchical controllers consisting of intra-domain controllers and cross-domain controllers. The controller has the ability to directly deliver the end-to-end SRv6 Policy on the ingress PE, thereby implementing SRv6 TE forwarding in Option C mode.

If the scenario without a controller is considered, since VPN routes are advertised directly between PEs through BGP, the logical next hop of the VPN route learned by the ingress PE is the special address of the egress PE. In order to implement SRv6 TE forwarding, VPN routes need to be iterated to the segment list on the ingress PE, and a public network tunnel to the egress PE needs to be constructed through ASBR

Take PE6 to advertise VPN routes to PE1 as an example, the route advertisement process is as follows:

1.   Advertising VPN route

- o @PE6 assigns VPNSID1 (End.DT4 type segment) to it after learning the VPN route. Then, the VPN route and VPNSID1 are advertised to PE1 through multi-hop eBGP, and the next hop address of the route is specified as the address of PE6, NXHPE6

- o @PE1 learns the VPN route in the corresponding VPN instance routing table, and uses NXHPE6 to iterate the real next hop

2.   Advertising NXHPE6 route

- o @PE6 and @ASBR4 establish an iBGP neighbor relationship. PE6 advertises the route of NXHPE6 to ASBR4, carrying the prefix SID as PSID1, and the next hop is the address of PE6.

- o o @ASBR4 learns the routes of NXHPE6 in the public network routing table. At the same time, ASBR4 creates a segment list1 destined for PE6, uses the PSID1 and PE6 addresses as the index of the segment list, and assigns a bindingSID (BSID1) to the segment list. If the corresponding segment list (with the same index) already exists, its bindingSID (BSID1) is used directly. Associate BSID1 with PSID1, ASBR4 advertises the route of NXHPE6 and BSID1 to ASBR3 through eBGP, and modifies the next hop to the address of ASBR4.

- o @ASBR3 learns the NXHPE6 route in the corresponding public network routing table, and then uses the addresses of BSID1 and ASBR4 as indexes to create a segment list2 destined for ASBR4, the list only contains the EPESID destined for ASBR4, and assigns BSID2 to segment list2 . ASBR3 associates BSID2 with BSID1, then advertises the route of NXHPE6 and BSID2 to ASBR2 through iBGP, and modifies the next hop to the address of ASBR3.

- o @ASBR2 behaves like ASBR4, creates segment list3, assigns BSID3 to it, and associates it with BSID2. Then, the route of NXHPE6 and BSID3 are advertised to ASBR1 through eBGP, and the next hop is changed to the address of ASBR2.

- o @ASBR1 behaves like ASBR3, creating segment list4, which only contains EPESIDs to ASBR2. BSID4 is allocated and associated with BSID3. Finally, the route of NXHPE6 and BSID4 are advertised to PE1 through iBGP, and the next hop is changed to the address of ASBR1

- o @PE1 behaves like ASBR4, creates segment list5 to ASBR1, assigns BSID5 to segment list5, and associates BSID5 with BSID4. Finally, PE1 records BSID5 as the next hop of the newly learned VPN route in the corresponding VPN instance routing table.

3.   Iterate the real next hop for the VPN route

- o @PE1 uses the route of NXHPE6 to iterate the real next hop for the VPN route. The VPN route finally learned from PE6 has the service SID of VPNSID1 and the next hop of BSID5.

For the relevant definitions of BSID5/BSID4/BSID3/BSID2/BSID1, please refer to the description of End.B6R in Section 2.2.2.

```
                           Multi-hop EBGP
        +---------------------------------------------------+
       /                                                     \
      / iBGP        eBGP         iBGP        eBGP       iBGP  \
     /+-------+   +---------+   +------+   +---------+ +------+\
    //         \ /           \ /        \ /          \/        \\
   PE1---------PE2---------PE3---------PE4---------PE5---------PE6
    |        (ASBR1)     (ASBR2)     (ASBR3)     (ASBR4)        |
    |BSID5      |BSID4      |BSID3      |BSID2      |BSID1      |
    |<----------|<----------|<----------|<----------|<----------|
    |  NXHPE6   |  NXHPE6   |  NXHPE6   |   NXHPE6  |  NXHPE6   |
    |  route +  |  route +  |  route +  |  route +  |  route +  |
    |  BSID4    |   BSID3   |  BSID2    |   BSID1   | PrefixSID |
    |                                                           |
    |  <------------------------------------------------------  |
    |                     VPNv4 route                           |
                          NextHop = NXHPE6
```

Figure 10: process of route advertisement for SRv6 TE in option C

Taking the packet sent from CE1 to CE2 as an example, the packet forwarding process in SRv6 TE mode is as follows:

- After @PE1 receives the packet from CE1, it searches the routing table in the corresponding VPN. The next hop and service SID of the corresponding VPN route are BSID5 and VPNSID1, respectively. PE1 adds SRv6 encapsulation to the original packet. The segment list in the SRH is <BSID5, VPNSID1>, and the destination address of the outer IPv6 header is BSID5. Since BSID5 is the local segment of PE1, it continues to process the packet on PE1.

- @PE1 replaces BSID5 in SRH with BSID4 associated with BSID5, and modifies the destination address to BSID4. Use segment list5 associated with BSID5 to forward packets. Add IPv6 and SRH to the packet, and encapsulate segment list5 in the SRH. Forward the packet in AS1 to ASBR1

- Before the packet reaches ASBR1, the outer IPv6 and SRH may have been de-encapsulated by the penultimate hop, or the outer encapsulation may have been de-encapsulated by ASBR1. ASBR1 continues to process the packet whose outer encapsulation has been de-encapsulated, and the destination address of the packet is BSID4 at this time. ASBR1 replaces BSID4 in the SRH with BSID3 associated with BSID4, and modifies the IPv6 destination address to BSID3. ASBR1 continues to use segment list4 associated with BSID4 to forward packets. Since there is only one EPESID in segment list4 and it is a segment of End.x type, there is no need to add encapsulation, and the packet is forwarded to ASBR2 according to the EPESID.

- After @ASBR2 receives the packet, the destination address of the packet is now BSID3. ASBR2 replaces BSID3 in the SRH with BSID2 associated with BSID3, and modifies the IPv6 destination address to BSID2. ASBR2 continues to use segment list3 associated with BSID3 to forward packets, adds IPv6 and SRH to the packets, and encapsulates segment list3 in SRH. The packet is forwarded in AS2 to ASBR3.

- The behavior of @ASBR3 is similar to that of ASBR1. The destination address of the packet after removing the outer encapsulation is BSID3, the destination address of the continued packet is updated to BSID1, and the packet is forwarded to ASBR4 according to the EPESID.

- After @ASBR4 receives the packet, the destination address of the packet is BSID1, and BSID1 is a normal bindingSID. Therefore, ASBR4 performs the normal bindingSID forwarding behavior, updates SHR.SL, and updates the destination address of the packet to VPNSID1. ASBR4 forwards the packet according to the segment list1 associated with BSID1, adds IPv6 and SRH to the packet, and encapsulates segment list1 in the SRH. The packet is forwarded to PE6 in AS3.

- After receiving the packet, @PE6 removes the SRv6 encapsulation, searches for the route in the routing table bound to VPNSID1, and forwards the original packet to CE2 according to the search result.

```
            +---------+         +---------+         +--------+
            |   AS1   |         |   AS2   |         |  AS3   |
   CE1-----PE1-------PE2-------PE3-------PE4-------PE5------PE6----CE2
                    (ASBR1)   (ASBR2)   (ASBR3)   (ASBR4)

          +-------+                     +-------+                     +-------+
          | IPv6  |                     | IPv6  |                     | IPv6  |
          +-------+                     +-------+                     +-------+
          |  SRH  |                     |  SRH  |                     |  SRH  |
          |segment|                     |segment|                     |segment|
          |list5, |                     |list3, |                     |list1, |
          +-------+  +-------+  +-------+  +-------+  +-------+
          | IPv6  |  | IPv6  |  | IPv6  |  | IPv6  |  | IPv6  |
          +-------+  +-------+  +-------+  +-------+  +-------+
          |  SRH  |  |  SRH  |  |  SRH  |  |  SRH  |  |  SRH  |
          | SL= 1 |  | SL = 1|  | SL = 1|  | SL = 1|  | SL = 0|
          |VPNSID1|  |VPNSID1|  |VPNSID1|  |VPNSID1|  |VPNSID1|
          | BSD4  |  | BSID3 |  | BSID2 |  | BSID1 |  | PSID1 |
   +---+  +-------+  +-------+  +-------+  +-------+  +-------+  +---+
   |///|->|///////|->|///////|->|///////|->|///////|->|///////|->|///|
   +---+  +-------+  +-------+  +-------+  +-------+  +-------+  +---+
```

Figure 11: Process of forwarding for SRv6 TE in option C

2.3.3. Summary of Option C

For SRv6 BE, locator routes can be advertised across domains to simply implement BE forwarding.

For SRv6 TEs, end-to-end SRv6 Policy can also be easily deployed when there is a controller.

3. Scenario of Intra-domain interworking

A typical scenario for intra-domain interworking is HVPN (Hierarchical VPN). In order to reduce the pressure on PE nodes, HVPN distributes the functions of PE to multiple PE devices, and multiple PE devices assume different roles.

UPE: A device directly connected to a user is called an Under-layer PE or User-end PE, or UPE for short. UPE mainly completes the user access function.

SPE: The device that is connected to the UPE and located in the network is called the superstratum PE (Superstratum PE) or the Service Provider-end PE (Service Provider-end PE), or SPE for short. SPE mainly manages and advertises VPN routes.

```
                      +----------------------------+
                      |          AS100             |
                      |                            |
          +---+  +----+         +-----+        +-----+    +---+
          |CE1+--+UPE1+=========+SPE1+========+UPE2+----+CE2|
          +---+  +----+   SRv6  +-----+  SRv6  +-----+    +---+
                      |                            |
                      +----------------------------+
```

Figure 12: HVPN reference topology

UPE only establishes BGP neighbor relationship with SPE. When UPE and SPE are in the same AS, UPE and SPE establish iBGP neighbor relationship. In H-VPN mode, PE can advertise detailed routes to UPE. As the client of the reflector SPE, the UPE receives detailed routes reflected by the SPE.

If the SPE and UPE are separated by an MPLS network, take UPE1 to UPE2 as an example when advertising VPN routes, UPE1 first advertises the VPN route to the SPE, and carries the VPN label assigned to the VPN route. The SPE first assigns a VPN label to the VPN route, replacing the VPN label assigned by UPE1, and sends the connected VPN route to other UPEs. The SPE needs to associate the VPN label assigned by itself with the VPN label assigned by the UPE.

When VPN packets go from CE2 to CE1, UPE2 adds MPLS encapsulation to them. The inner VPN label is the VPN label assigned by SPE1, and the outer label is the public network label destined for SPE1. After the

packet reaches SPE1, SPE1 strips the outer public network tunnel label, replaces the VPN label with the VPN label assigned by UPE1, and sends the packet to UPE1. Finally, after receiving the packet, UPE1 strips the public network label and VPN label, and forwards the packet to CE1

If the provider upgrades MPLS to SRv6 on this basis, the SPE also needs to implement the interworking of the SRv6 domain within the domain.

The intra-domain SRv6 interworking represented by HVPN is similar to the cross-domain processing behavior of Option B.

3.1.1. SRv6 BE

Taking UPE2 to advertise VPN routes to UPE1 as an example, the route advertisement process is as follows:

- o @UPE2 assigns VPNSID1 (segment of End.DT4 type) to it after learning the VPN route. Then advertise the VPN route and VPNSID1 to SPE1 via iBGP.

- o @SPE1 learns VPN routes in the routing table of the corresponding VPN instance, and assigns a SID2 to associate it with VPNSID1. @SPE1 advertises VPN route and SID2 to @uPE1 via iBGP. SID2 has the same behavior as SIDs created by ASBR described in section2.2.1

- o @PE1 learns the VPN route and SID2 in the corresponding VPN instance routing table

```
                  iBGP              iBGP
             +-----------+    +------------+
            /             \  /              \
          UPE1-------------SPE1-------------UPE2
           |                |                 |
           |<-VPN4 route->  |<- VPNv4 route ->|
```

Figure 13: process of route advertisement for HVPN

Taking the packet sent from CE1 to CE2 as an example, the packet forwarding process in SRv6 BE mode is as follows:

- o After receiving the packet from CE1, @UPE1 searches the routing table in the corresponding VPN. Add IPv6 encapsulation to the original packet, and the outer IPv6 destination address is SID2

- o After receiving the packet, SPE1 finds the VPNSID1 associated with it according to the destination address SID2, replaces SID2 in the packet with VPNSID1, and forwards the packet to UPE2.

- o After receiving the packet, UPE2 removes the outer IPv6 encapsulation, searches for the route in the routing table bound to VPNSID1, and forwards the original packet to CE2 according to the search result

```
             CE1------UPE1--------SPE1---------UPE2---------CE2
                          +-------+    +-------+
                          | SRv6  |    | SRv6  |
                          | SID2  |    |VPNSID1|
              +------+    +-------+    +-------+    +-------+
              |//////| -> |///////| -> |///////| -> |///////|
              +------+    +-------+    +-------+    +-------+
```

Figure 14: Process of forwarding for HVPN

3.1.2. SRv6 TE

The processing process of SRv6 TE of HVPN is similar to that of Option B inter-domain. When SPE republishes routes, it needs to undertake functions similar to ASBR, which will not be described too much.

4. IANA Considerations

This document has no IANA actions.

5. Security Considerations

The security requirements and mechanisms described in [RFC8402] and [RFC8754] also apply to this document.

This document does not introduce any new security consideration.

6. References

6.1. 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>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[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>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

Contributors

xxx contributed to the content of this document.

Authors' Addresses

Weiqiang Cheng
China Mobile
China
Email: chengweiqiang@chinamobile.com

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