Workgroup: Network Working Group
Internet-Draft:
draft-wang-bess-sbfd-discriminator-04
Published: 3 December 2022
Intended Status: Standards Track
Expires: 6 June 2023
Authors: H. Wang J. Dong G. Mirsky Y. Huang
Huawei Huawei Ericsson Huawei

Advertising S-BFD Discriminators in BGP

Abstract

Seamless Bidirectional Forwarding Detection (S-BFD) is a simplified BFD mechanism. It eliminates most negotiation aspects and provides advantages such as fast configuration injection. S-BFD is especially useful in multi-homing PE scenarios and reduces resource overheads on the dual-homing PEs. Although S-BFD is simpler than BFD, a large number of manual configurations are required when there are a large number of PEs.

This document provides the mechanism of distributing S-BFD discriminators with VPN service routes, which simplifies S-BFD deployment for VPN services.

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 RFC 2119 [RFC2119].

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 6 June 2023.

Table of Contents

1. Introduction
2. Terminology
3. Typical Scenarios
   3.1. EVPN Layer 3 Service Over SRv6 BE Use Case
   3.2. EVPN Layer 3 Service Over SPv6 Policy Use Case
4. Procedure
   4.1. BGP Encoding
   4.2. Router Procedure
      4.2.1. Egress Node Process
      4.2.2. Transit Node Process
      4.2.3. Ingress Node Process
5. Error handling
6. IANA Considerations
7. Security Considerations
8. Acknowledgements
9. References
   9.1. Normative References
   9.2. References
Authors' Addresses

1. Introduction

In deploying some network services, carriers usually deploy active and standby nodes to improve network reliability. In this way, a protection switchover can be performed quickly when there is a fault at the active node or a failure on the working path between PEs. To accelerate fault detection, BFD is usually used. Traditionally, BFD sessions need to be provisioned on both ends of the BFD session, which occupies the resources of both PEs.

[RFC7880] defines Seamless Bidirectional Forwarding Detection (S-BFD), a simplified mechanism for using BFD with a large proportion

of negotiation aspects eliminated, thus providing benefits such as quick provisioning, as well as improved control and flexibility for network nodes initiating path monitoring. This mechanism is especially useful in asymmetric scenarios, such as the 3PE scenarios. At the dual-homing PE nodes, BFD does not need to be used to detect the single-homed PE node. In this scenario, S-BFD greatly saves resources on the dual-homing nodes.

To deploy S-BFD, the initiator needs to know the reflector's ip address and discriminator . When a large number of PEs need to be deployed, even the deployment of S-BFD can become complicated. [RFC7883] and [RFC7884] introduced an IGP-based S-BFD discriminator advertisement mechanism to simplify S-BFD deployment. Since VPN services may be deployed across an area or AS boundaries, the IGP-based S-BFD mechanism does not apply in this case.

This document proposes a mechanism to distribute S-BFD discriminator information with BGP service routes. It allows advertising S-BFD discriminator across multiple domains and achieves on-demand end-to-end S-BFD session provisioning for specific BGP service routes.

2. Terminology

BFD : Bidirectional Forwarding Detection

S-BFD : Seamless Bidirectional Forwarding Detection

3PE : One PE connect to dual-homed PEs scenario

APE : Access PE, used to access users

SPE: Service PE, used to support service for users

UCE: User CE

SCE: Service CE

3. Typical Scenarios

In some inter-domain VPN service deployments, only one of a pair of interconnected PEs benefits from monitoring the status of the connection. In such a case, using S-BFD [RFC7880] is advantageous as it reduces the load on one of the PEs while providing the benefit of faster convergence. The following subsections provide some examples of SRv6-based VPN services to show the benefits of using S-BFD.

For SRv6 services, two types of paths are used to support the service. One is service over SRv6 BE, the other is service over SRv6 Policy. For the service over SRv6 BE case, the VPNSID is used to resolve the forwarding information. Thus an S-BFD session is needed

to detect the reachability of the VPNSID. The session is an IP-routed S-BFD, and the SRv6 locator of the remote VPN SID can be used as the destination identifier of the S-BFD session. For the service over SRv6 Policy, the <nexthop, color> in the service route is used to resolve to an SRv6 Policy. Then an S-BFD session is needed to detect the reachability of the SRv6 Policy.

3.1. EVPN Layer 3 Service Over SRv6 BE Use Case

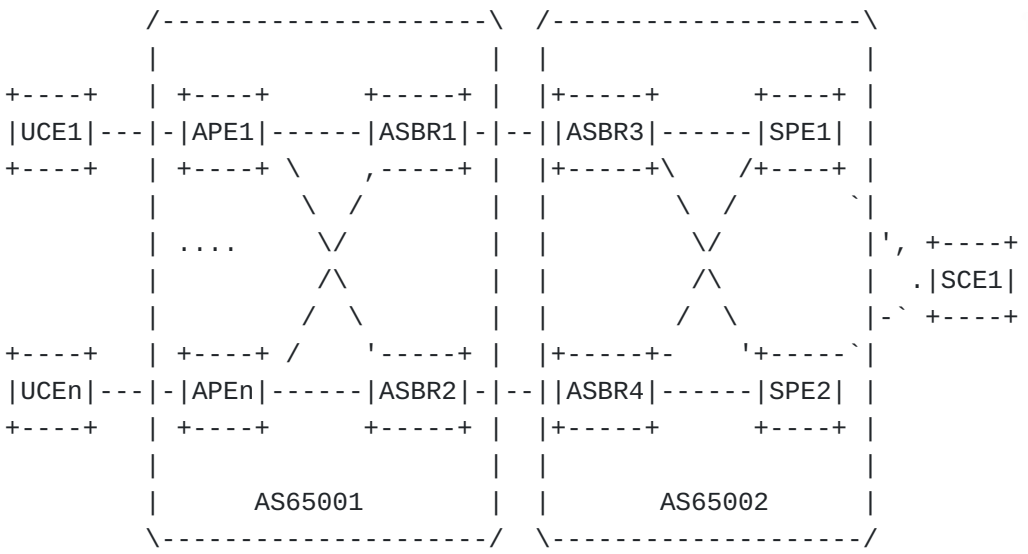

Figure 1: EVPN Layer 3 Service Over SRv6 BE

Figure 1 shows a SRv6 BE based seamless scenario. UCE is single-homed to APE, and SCE is dual-homed to SPE1 and SPE2. The service is across multiple ASes.

SCE1 accesses SPE1 and SE2 through Layer 3 and advertises its private network routes to them. SPE1 and SPE2 encapsulate the routes into Type 5 routes in the EVPN format and sends them to APE1. After receiving Type 5 routes advertised by SPE1 and SPE2, APE1 generates primary and backup entries for the routes to speed up service switchover. In this scenario, the SRv6 BE service mode is used. APE1 will resolve SPE1's VPN routes reachability through the VPN SID. To ensure that APE1 can properly route to PE1, PE1 needs to advertise its own locator route. The advertisement of the locator route is not in the scope of this document.

To speed up the fault detection, we may configure an S-BFD session on APE1 to detect SPE1 or SPE2's reachability. In traditional mode, a discriminator needs to be assigned by SPE1 and SPE2, and two S-BFD sessions need to be configured on APE1 to detect the VPN SID's reachability of SPE1 and SPE2. It needs to generate an S-BFD session with the destination set to the VPN SID. To reduce the number of S-BFD sessions, locator-based S-BFD sessions can be used instead of S-BFD sessions for VPNSIDs.

There are a large number of such APEs that exist on the network. Each APE is configured with several S-BFD sessions to detect PE1 and PE2, which increases the deployment complexity.

3.2. EVPN Layer 3 Service Over SPv6 Policy Use Case

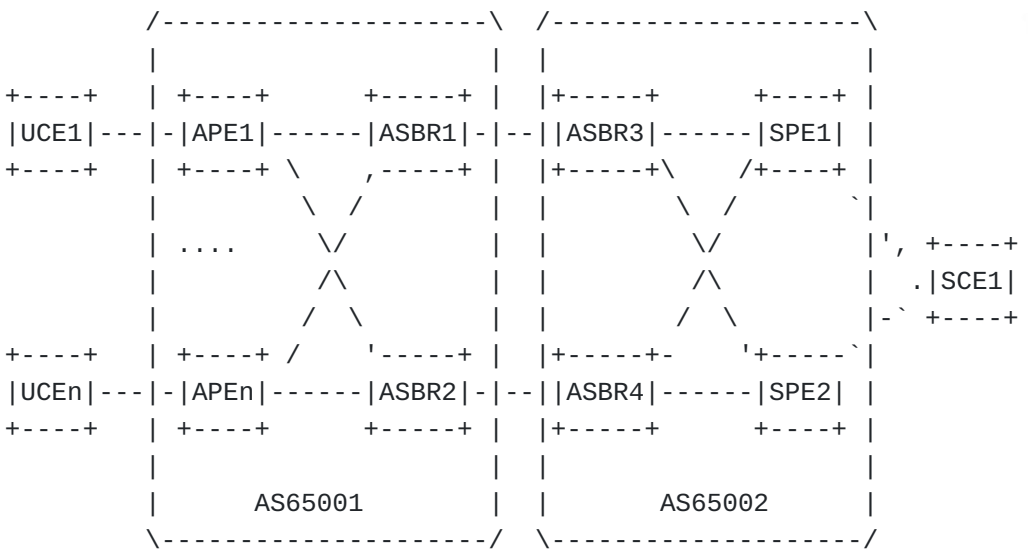

Figure 2: EVPN Layer 3 Service Over SRv6 Policy

Figure 2 shows an SRv6 Policy scenario. SCE1 is dual-homed to SPE1 and SPE2, and UCE1 is accessed to APE1. SPE1, SPE2, and APE1 are cross BGP ASes.

SCE1 accesses SPE1 and SPE2 through Layer 3 and advertises its private network routes to APE1. SPE1 and SPE2 encapsulate the routes into Type 5 routes in the EVPN format and sends them to APE1.

After receiving Type 5 routes advertised by SPE1 and SPE2, APE1 generates primary and backup entries for the routes, speeding up service switchover. APE1 parses the tunnel based on the <nexthop, color> of the service routes advertised by SPE1 and SPE2, and matches an SRv6 Policy. After receiving the traffic from UCE1 to SCE1, APE1 encapsulates and forwards the traffic based on the SRv6 Policy.

An S-BFD session needs to be established for these SRv6 Policy-based forwarding paths to swiftly detect the availability of the paths. When detecting a fault on the SRv6 Policy path of the primary service route, services can be swiftly switched to the backup path, providing more reliable protection for services.

There are a large number of such PEs that exist on the network. Each PE is configured with several S-BFD sessions to detect PE1 and PE2, which increases the deployment complexity.

Certainly, this scenario may also be implemented in other methods. For example, when provisioning an SRv6 policy, an S-BFD session can be provisioned. While in some cases, it would be more efficient if the S-BFD session could be provisioned based on the demand of the services.

4. Procedure

4.1. BGP Encoding

[RFC9026] specifies the "BFD Discriminators" (38) attribute, which is an optional transitive BGP attribute that conveys the Discriminators and other optional attributes used to establish BFD sessions.

In [RFC9026], the BFD Discriminators attribute is used to transmit P2MP BFD session creation information MVPN scenarios. For non-multicast services, such as L3VPN services, L2VPN services, EVPN services and native IP services, BFD discriminators are also required for creating an S-BFD session. This document reuses the BFD Discriminators attribute and defines new BFD modes for some of these services.

The format of BFD Discriminator attribute is shown as follows:

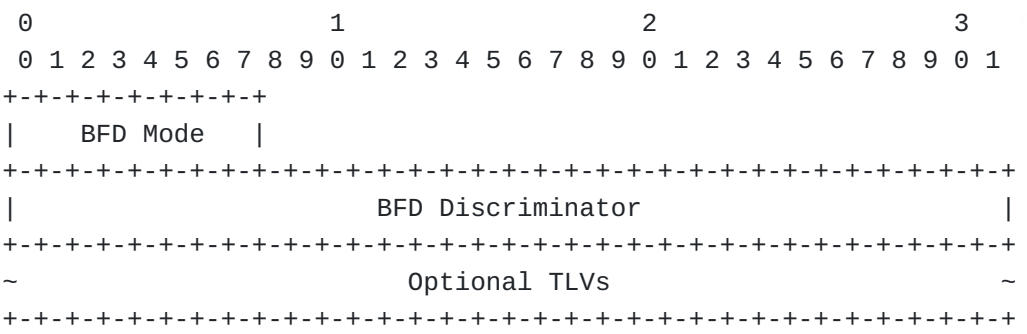

Figure 3: Format of the BFD Discriminator Attribute

o BFD Mode:

The BFD Mode field is 1 octet. [RFC9026] defines only the P2MP BFD session for MVPN. This document defines two new types of S-BFD session types for the scenarios in Section 3.

As described in the preceding scenarios, there are two types of S-BFD sessions for SRv6 services. For service over SRv6 BE, an IP-routed S-BFD session needs to be created to detect the reachability of the SRv6 locator. For service over SRv6 Policy, an S-BFD session

for SRv6 Policy needs to be created to detect the reachability of the SRv6 Policy. Thus two new BFD modes are introduced:

- *S-BFD for SRv6 Locator Session Mode, which is dedicated to detecting the locator. The type value is to be allocated as described in Section 6.

- *S-BFD for Common Session Mode, which is for general S-BFD session. The type value is to be allocated as described in Section 6. This mode is not only for SRv6, but also can be used for other scenarios.

o BFD Discriminators:

The field length is 4 octets. Used to specify the discriminator for S-BFD session.

o Optional TLVs:

Variable-length fields are optional. Indicates the additional information required for creating a S-BFD session. The format is as follows:

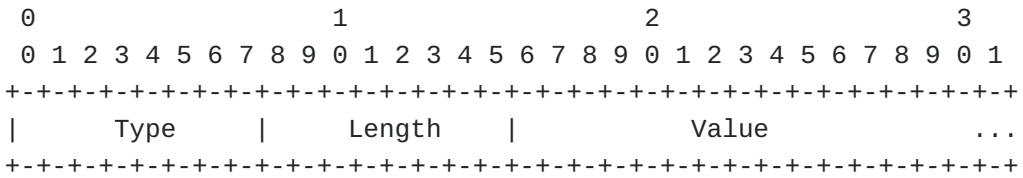

Figure 4: Format of the Optional TLV

If a transit node changes the next hop or reassigns a VPN SID when advertising a route, the transit node needs to use the locally allocated S-BFD discriminator for the S-BFD discriminator attribute. Suppose the transit node does not recognize the S-BFD Discriminator attribute in the received route and continues to advertise the route to the remote PE. In that case, the receiver may use incorrect information when creating an S-BFD session. Therefore, the advertised S-BFD Discriminator attribute also needs to carry the IP address of the originator of the discriminator for receiver side verification.

For the two BFD modes defined in this document, the "Source IP Address" TLV as defined in [RFC9026] MUST be carried in the BFD Discriminator attribute. If the mode is "S-BFD for SRv6 Locator Session", the SRv6 Locator address MUST be used for the service is carried in the TLV. If the mode is "S-BFD for Common Session", the next-hop address MUST be used for the service is carried in the TLV.

4.2. Router Procedure

In BGP address families, such as L3VPN or EVPN, routes can carry the S-BFD Discriminator attribute as required so that S-BFD sessions can be established based on the attribute. The following uses S-BFD for SRv6 Locator as an example. If mode is set to S-BFD for Common Session, the processing method is similar.

4.2.1. Egress Node Process

As shown in figure 1, the S-BFD discriminator is configured on PE1. After obtaining the information, BGP encapsulates the attribute into the EVPN route and sets the BFD Mode to S-BFD for Locator Session, when advertising the EVPN route. The Discriminator value is local discriminator value. The optional TLV carries the local PE's locator address used by the VPN.

4.2.2. Transit Node Process

Here is the end-to-end SRv6 BE scenario. The ASBR does not re-allocate the VPN SID. Thus, the ASBR does not require to modify the VPN SID, and not to alter the BFD discriminator attribute.

4.2.3. Ingress Node Process

After receiving the EVPN Type 5 routes from PE1 and PE2, PE3 imports the routes to the VRF of PE3 based on the route targets. Routes triggers establish the S-BFD sessions based on <S-BFD discriminator, locator ip>.

Then, routes with the same prefix from PE1 and PE2 form primary and backup paths. When the primary path or the egress node is in fault, S-BFD detects that fault and forms switch to backup path quickly.

To avoid the waste of redundant resources, assume that the ASBR re-assigns the SID in Option B and the ASBR does not recognize the attribute. In this case, the SID and locator carried in the route received by PE3 do not match the Source IP carried in the Optional TLV in the BFD attribute. Therefore, PE3 does not need to establish an S-BFD session to remote PE, which can avoid resource waste.

5. Error handling

Error handling complies with [RFC7606]. In this document, the BFD discriminator information is used only to establish an S-BFD session. Therefore, if the BFD discriminator information is invalid, the BFD attribute will be discard and not transmit to other devices.

For BFD discriminator attribute, the following case will be processed:

o The BFD Discriminator value in receiving BFD Discriminator attribute is 0, the attribute is invalid.

For the BFD mode type "S-BFD for SRv6 Locator Session", the following case will be processed:

o If the BFD discriminator attribute doesn't contain optional TLV with type set to 1, the attribute is invalid.

o If the optional TLV type is 1 but the length is not 16, the attribute is invalid.

o If the optional TLV type is 1 but the value is all 0, the attribute is invalid.

o If multiple Source IP Optional TLVs are carried, the first source IP address should be used as the destination to establish an S-BFD session. For EVPN type 2 MAC-IP routes may use the first and the second IP address because it may carry two SRv6 SIDs with different locators. Other source IP addresses should be ignored.

o If a non-Source IP Optional TLV is carried, the Optional TLV will be ignored.

For the BFD mode type "S-BFD for Common Session", the following case will be processed:

o If the BFD discriminator attribute doesn't contain optional TLV with type set to 1, the attribute is invalid.

o If the optional TLV type is 1 but the length is not 4 or 16, the attribute is invalid.

o It the optional TLV type is 1 but the value is all 0, the attribute is invalid.

o If multiple Source IP Optional TLVs are carried, only the first source IP address should be used as the destination to establish an S-BFD session. Other source IP addresses should be ignored.

o If a non-Source IP Optional TLV is carried, the Optional TLV will be ignored.

6. IANA Considerations

IANA is requested to assign two new code points from the "BFD Mode" subregistry in the "Border Gateway Protocol (BGP) Parameters" registry.

Value	Description
TBA1	S-BFD for SRv6 Locator Session
TBA2	S-BFD for Common Session

7. Security Considerations

The new S-BFD modes introduced in this document does not introduce any new security risks for BGP.

The BFD attribute is an optional attribute and is mainly used for network services within a single administrative domain. The operator SHOULD ensure this attribute does not propagate across the boundary of the administrative domain. For VPN services, the advertisement range of this attribute is the same as that of VPN routes. Therefore, this attribute is not advertised outside the management domain. For public IPv4 and IPv6 services, the border node of the administrative domain SHOULD be configured not to propagate the BFD attribute to other domains.

When creating an S-BFD session, the initiator verifies the S-BFD session based on routing information. This reduces the number of invalid S-BFD sessions and avoid attribute attack.

8. Acknowledgements

NA

9. References

9.1. 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>.

9.2. References

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <https://www.rfc-editor.org/info/rfc7606>.

[RFC7880] Pignataro, C., Ward, D., Akiya, N., Bhatia, M., and S. Pallagatti, "Seamless Bidirectional Forwarding Detection

(S-BFD)", RFC 7880, DOI 10.17487/RFC7880, July 2016, <https://www.rfc-editor.org/info/rfc7880>.

[RFC7883] Ginsberg, L., Akiya, N., and M. Chen, "Advertising Seamless Bidirectional Forwarding Detection (S-BFD) Discriminators in IS-IS", RFC 7883, DOI 10.17487/RFC7883, July 2016, <https://www.rfc-editor.org/info/rfc7883>.

[RFC7884] Pignataro, C., Bhatia, M., Aldrin, S., and T. Ranganath, "OSPF Extensions to Advertise Seamless Bidirectional Forwarding Detection (S-BFD) Target Discriminators", RFC 7884, DOI 10.17487/RFC7884, July 2016, <https://www.rfc-editor.org/info/rfc7884>.

[RFC9026] Morin, T., Ed., Kebler, R., Ed., and G. Mirsky, Ed., "Multicast VPN Fast Upstream Failover", RFC 9026, DOI 10.17487/RFC9026, April 2021, <https://www.rfc-editor.org/info/rfc9026>.

Authors' Addresses

Haibo Wang
Huawei
No. 156 Beiqing Road
Beijing
100095
P.R. China

Email: rainsword.wang@huawei.com

Jie Dong
Huawei
No. 156 Beiqing Road
Beijing
100095
P.R. China

Email: jie.dong@huawei.com

Greg Mirsky
Ericsson

Email: gregimirsky@gmail.com

Yang Huang
Huawei
No. 156 Beiqing Road
Beijing
100095
P.R. China

Email: yang.huang@huawei.com