Workgroup: Network Working Group
Internet-Draft:
draft-wang-bess-sbfd-discriminator-00
Published: 19 October 2021
Intended Status: Standards Track
Expires: 22 April 2022
Authors: H. Wang Y. Huang J. Dong
Huawei Huawei Huawei

Advertising S-BFD Discriminators in BGP

Abstract

This document defines the method of transmitting S-BFD Discriminators through BGP attributes. This method helps services create S-BFD sessions more easily.

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

Copyright Notice

Table of Contents

1. Introduction

[RFC7880] defines Seamless Bidirectional Forwarding Detection (S-BFD) mechanism. S-BFD is 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. Currently, S-BFD can be used in service deployment to simplify the deployment.

2. Motivations

An important thing for S-BFD is to check the reachability of services, so that service interruption can be quickly detected when there is a failure on the service path and services can be switched to a backup path quickly.

[RFC7880] defines Seamless Bidirectional Forwarding Detection (S-BFD) mechanism. Generally, the administrator needs to manually deploy S-BFD discriminators on the device to create S-BFD sessions.

For the deployment of S-BFD in IPv4 network, the reflector can use the LSR-ID address as the discriminator. This reduces the number of discriminators deployed on the transmit end. This mode cannot be used for IPv6 because the discriminator has only four bytes.

[RFC7883] [RFC7884] defines IS-IS and OSPF to flood BFD discriminators. However, this mode is based on nodes and cannot traverse an IGP area. In addition, without the knowledge of services to be detected, a large number of unnecessary S-BFD sessions may be created.

It is suggested to use BGP to distribute BFD discriminator information. BGP can transmit routes across domains, and service routes can driven the establishment of end-to-end S-BFD sessions.

3. Scenarios

3.1. EVPN Layer 3 Service Over SRv6 BE Use Case

```
         +-----------------------+ +-------------------+
         |                       | |                   |
         | +-----+      +-----+  | | +-----+           |
         | , PE1 |------|ASBR1|------|ASBR3\           |
         |/+-----+      +-----+  | | +-----+\          |
         /                       | |         \         |
        /|                       | |          \        |
 +-----/ |                       | |           '-----+ | +-----+
 | CE1 | |                       | |           | PE3 |---| CE2 |
 +------ |                       | |           /-----+ | +-----+
        `,                       | |          /        |
         |.+-----+      +-----+  | | +-----+ /         |
         | ' PE2 |------|ASBR2|------|ASBR4|-          |
         | +-----+      +-----+  | | +-----+           |
         |                       | |                   |
         |        AS65001        | |      AS65002      |
         +-----------------------+ +-------------------+
```

Figure 1: EVPN Layer 3 Service Over SRv6 BE

Figure 1 shows a SRv6 BE-based seamless scenario, PE1 and PE2 are dual-homed to CE1, and PE3 is dual-homed to CE2. PE1, PE2, and PE3 cross BGP ASes.

CE1 accesses PE1 and PE2 through Layer 3 and advertises its private network routes to PE1. PE1 encapsulates the routes into Type 5 routes in the EVPN format and sends them to PE3. After receiving Type 5 routes advertised by PE1 and PE2, PE3 generates primary and backup entries for the routes to speed up service switchover.

To speed up fault detection, we may configure an S-BFD session on PE3 to detect PE1 and PE2. In traditional mode, a discriminator needs to be assigned to PE1 and PE2, and two S-BFD sessions needs to be configured on PE3 to detect the VPN SID's reachability of PE1 and

PE2. In this scenario, the ingress PE forward services based on the reachability of the VPN SID. To reduce the number of S-BFD sessions, we may detect SRv6 locator routes.

There are large number of such PEs exist on the network. Each PE 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

```
           +-----+      +-----+
           , PE1 |------| P1  |
          /+-----+      +-----+\
         /                      \
        /                        \
 +-----/         AS65001          '-----+   +-----+
 | CE1 |                          | PE3 |---| CE2 |
 +------                          /-----+   +-----+
        `,                       /
          .+-----+      +-----+ /
           ' PE2 |------| P2P |-
           +-----+      +-----+
```
Figure 2: EVPN Layer 3 Service Over SRv6 Policy

Figure 2 shows a SRv6 Policy scenario, CE1 is dual-homed to PE1 and PE2, and PE3 is dual-homed to PE1 and PE2.

CE1 accesses PE1 and PE2 through Layer 3 and advertises its private network routes to PE1. PE1 encapsulates the routes into Type 5 routes in the EVPN format and sends them to PE3.

After receiving Type 5 routes advertised by PE1 and PE2, PE3 generates primary and backup entries for the routes to speed up service switchover.

Configure S-BFD sessions on PE3 to detect PE1 and PE2 can speed up the fault detection. In traditional mode, a discriminator needs to be assigned to PE1 and PE2, and S-BFD sessions is configured on PE3 to detect the SRv6 Policy's endpoint of PE1 and PE2.

There are large number of such PEs exist on the network, each PE must be configured with S-BFD sessions to detect PE1 and PE2, which increases the deployment complexity.

4. Procedure

4.1. BGP Encoding

[RFC9026] defines 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.

The attribute defined at [RFC9026] is used to transmit P2MP BFD session creation information through the BFD Discriminator attribute in MVPN scenarios. For non-multicast services, such as L3VPN services, L2VPN services, and native IP services, BFD discriminators are also required to create an S-BFD session.

The format of the BFD Discriminator attribute is as follows:

```
 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
+-+-+-+-+-+-+-+-+
|    BFD Mode   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       BFD Discriminator                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                         Optional TLVs                         ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: Format of the BFD Discriminator Attribute

o BFD Mode:

The BFD Mode field is 1 octet long. [RFC9026] defines only the P2MP BFD session for MVPN. This document defines two new types of SBFD session types based on the preceding scenarios.

SBFD for SRv6 Locator Session Mode, which dedicated to detecting the locator. The temporary type is 176, and is to be allocated by IANA.

SBFD for Common Session Mode, which is for general SBFD session. The temporary type is 177, and is to be allocated by IANA. 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 describe 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:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type     |     Length    |           Value             ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Format of the Optional TLV

In this document, S-BFD for SRv6 Locator Session and S-BFD for Common Session must carry an IP addresses except discriminators, which reuse the Source IP Address TLV defined in [RFC9026].

If the mode is set to SBFD for SRv6 Locator Session, the SRv6 Locator address used for the service is carried.

If the mode is set to SBFD for Common Session, the next-hop address used for the service is carried.

For details about the error handling, see section "Error Handling".

4.2. Process

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

4.2.1. Egress 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 SBFD 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 Process

Here is the seamless scenario, the ASBR does not re-allocate the VPNSID. Therefore, the ASBR does not need to modify the VPNSID, and not to change the BFD discriminator attribute.

4.2.3. Ingress 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 <discriminator, locator ip> information to detect SRv6 BE connectivity.

In addition, 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 attirbute 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 BFD mode type is S-BFD for SRv6 Locator Session, the following case will be processed:

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

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

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

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

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

o 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

This document defines two new BFD modes in the BFD Discriminator attribute. The following values are recommended to be assigned by IANA:

Value	Description
176	S-BFD for SRv6 Locator Session
177	S-BFD for Common Session

7. Security Considerations

The new S-BFD Discriminators sub-TLV does not introduce any new security risks for BGP.

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

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

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

Email: yang.huang@huawei.com

Jie Dong
Huawei
No. 156 Beiqing Road
Beijing

100095
P.R. China

Email: jie.dong@huawei.com