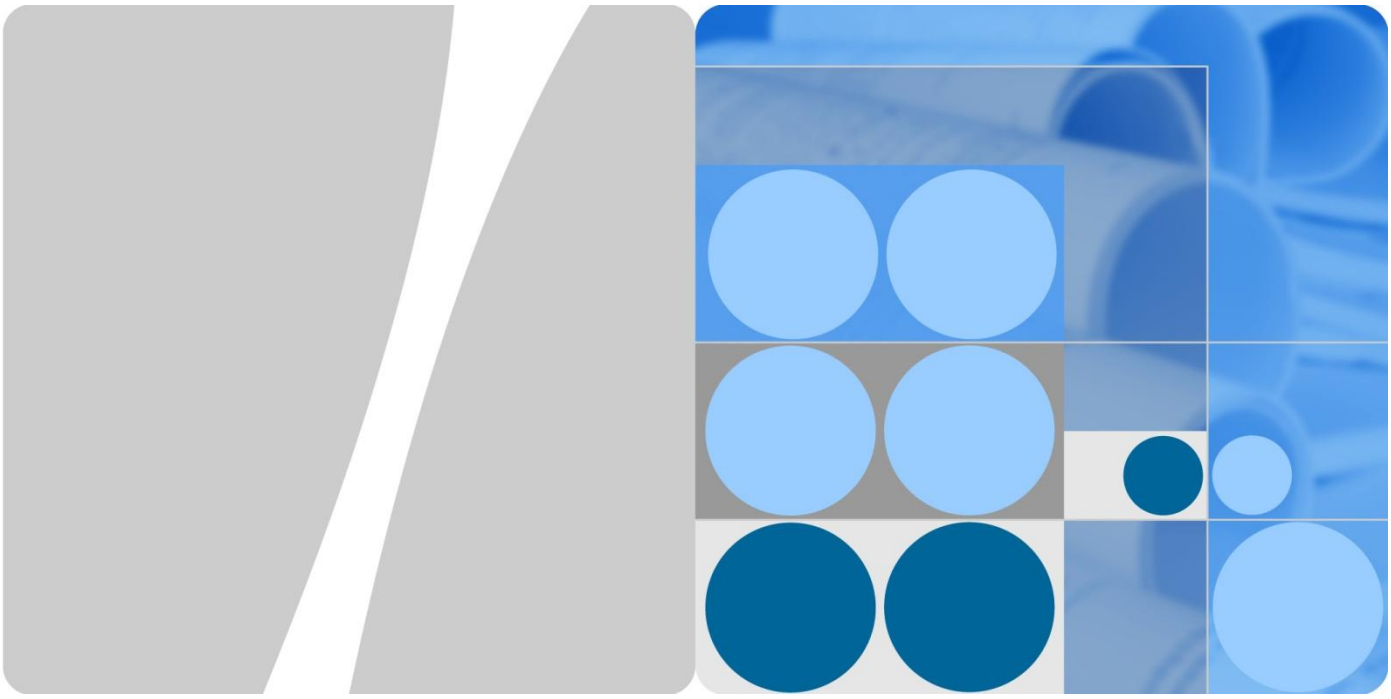


BFD Technology White Paper

Issue 01

Date 2012-08-07

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions

and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: http://www.huawei.com

Email: support@huawei.com

Contents

1 BFD

1.1 Introduction

Definition

Bidirectional Forwarding Detection (BFD) is a unified detection mechanism used to rapidly detect link faults and monitor IP connectivity.

Purpose

A network device must detect a communications fault between adjacent devices quickly so that the upper layer protocol can rectify the fault and prevent a service interruption. In practice, hardware detection is used to detect link faults. For example, Synchronous Digital Hierarchy (SDH) alarms are used to report link faults. However, not all media can provide the hardware detection mechanism. Applications use the Hello mechanism of the upper-layer routing protocol to detect faults. The detection duration is more than 1 second, which is too long for some applications. If no routing protocol is deployed on a small-scale Layer 3 network, the Hello mechanism cannot be used.

BFD provides fast fault detection independent of media and routing protocols. It has the following advantages:

- Rapidly detects link faults between neighboring network devices. The detected faults may occur on interfaces, data links, or forwarding engines.
- Provides uniform detection for all media and protocol layers in real time.

Benefits

BFD rapidly detects link faults and monitors IP connectivity, helping you improve network performance. Adjacent systems can quickly detect communication faults so that a standby channel can be created immediately to restore communication and ensure network reliability.

1.2 Reference Standards and Protocols

Table 1-1 Reference standards and protocols

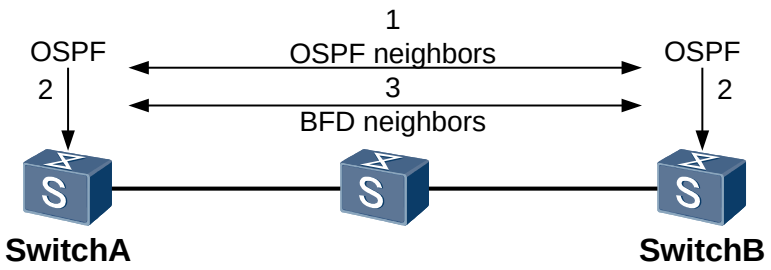
Document	Description	Remarks
RFC 5880	Bidirectional Forwarding Detection (BFD)	-
RFC 5881	Bidirectional Forwarding Detection (BFD) for IPv4 and IPv6 (Single Hop)	-
RFC 5882	Generic Application of Bidirectional Forwarding Detection (BFD)	-
RFC 5883	Bidirectional Forwarding Detection (BFD) for Multihop Paths	-
RFC 5884	Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)	-

1.3 Principles

1.3.1 BFD Implementation

Two network devices establish a BFD session to detect the forwarding path between them and serve upper-layer applications. BFD does not provide neighbor discovery. Instead, BFD obtains neighbor information from the upper-layer application BFD serves to establish a BFD session. After the BFD session is set up, the local device periodically sends BFD packets. If the local device does not receive a response from the peer system within the detection time, it considers the forwarding path faulty. BFD then notifies the upper-layer application for processing. The following uses association between OSPF and BFD as an example to describe the BFD session setup process.

Figure 1-1 BFD session setup

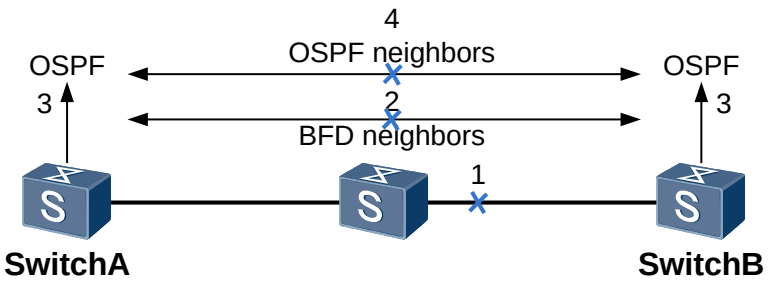


As shown in Figure 1-1, OSPF and BFD are configured on SwitchA and SwitchB. The BFD session setup process is as follows:

- OSPF uses the Hello mechanism to discover neighbors and establishes a neighbor relationship.

- OSPF notifies BFD of neighbor information including source and destination addresses.
- BFD sets up a BFD session based on received neighbor information. After the BFD session is set up, BFD starts to detect link faults and rapidly responds to link faults.

Figure 1-2 BFD detecting link faults



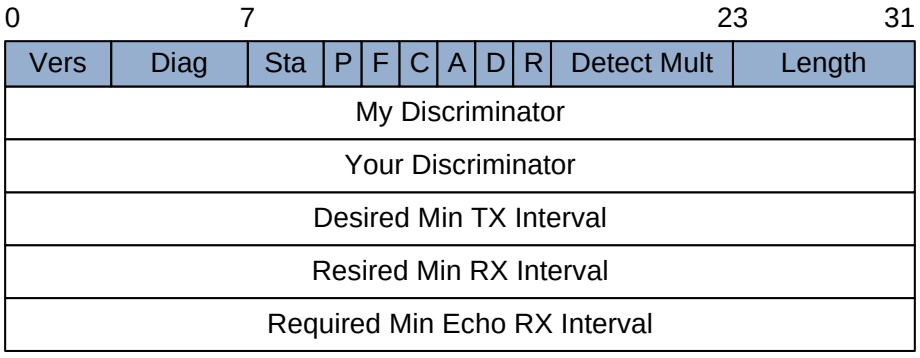
As shown in Figure 1-2:

- The detected link is faulty.
- BFD rapidly detects the link fault and the BFD session becomes Down.
- BFD notifies the local OSPF process that the neighbor is unreachable.
- The local OSPF process ends the OSPF neighbor relationship.

1.3.2 BFD Packet

BFD packets include BFD control packets and BFD echo packets.

Figure 1-3 BFD packet format



- Vers: indicates the BFD version number. The current version number is 1.
- Diag: indicates that the cause of the last session status change on the local BFD system. The following table describes the values and meanings of the Diag field.

Table 1-2 Meanings of the Diag field

Value	Meaning
0	No Diagnostic

Value	Meaning
1	Control Detection Time Expired
2	Echo Function Failed
3	Neighbor Signaled Session Down
4	Forwarding Plane Reset
5	Path Down
6	Concatenated Path Down
7	Administratively Down
8	Reverse Concatenated Path Down

- Sta: indicates the local BFD status. The following table describes the values and meanings of the Sta field.

Table 1-3 Meanings of the Sta field

Value	Meaning
0	AdminDown
1	Down
2	Init
3	Up

- P: If the P bit is set, the transmitting system requests verification of connectivity or of a parameter change. If it is cleared, the transmitting system does not request verification.
- F: If the F bit is set, the transmitting system responds to a received BFD packet that has the P bit set to 1. If it is cleared, the transmitting system does not respond to a BFD packet that has the P bit set to 1.
- **C**: indicates the forwarding plane/control plane independent flag. Once the C bit is set, the changes of the control plane do not affect BFD detection. For example, during the restart or GR of OSPF applied in the control plane, BFD can continue to detect link status.
- A: indicates authentication. If the A bit is set, a BFD session needs to be authenticated.
- D: indicates the demand bit. If this bit is set, the transmitting system wants to monitor links in demand mode.
- R: indicates the reserved bits. The bit is set to 0 during transmission and ignored during reception.
- Detect Mult: indicates the detection multiplier. In asynchronous mode, the detection time of the transmitting system is the negotiated transmit interval multiplied by this value.
- Length: indicates the length of the BFD packet, in bytes.

- My Discriminator: indicates the local discriminator of the BFD session. It is a unique, nonzero value generated by the transmitting system, used to demultiplex multiple BFD sessions between the same pair of systems.
- Your Discriminator: indicates the remote discriminator of the BFD session. It is the discriminator received from the corresponding remote system. This field reflects back the received value of My Discriminator, or is zero if that value is unknown.
- Desired Min Tx Interval: indicates the minimum interval for sending BFD packets on the local end, in milliseconds.
- Required Min RX Interval: indicates the minimum interval for receiving BFD packets on the local end, in milliseconds.
- Required Min Echo RX Interval: indicates the minimum interval for receiving Echo packets on the local end, in milliseconds. If the local end does not support the Echo function, the field is set to 0.
- Auth Type: indicates the authentication type. It takes effect when the A bit is set to 1. It indicates the authentication type used by BFD control packets. Currently, BFD supports the following authentication types.

Table 1-4 Authentication type

Authentication Type Field	Authentication Type
0	Reserved
1	Simple Password
2	Keyed MD5
3	Meticulous Keyed MD5
4	Keyed SHA1
5	Meticulous Keyed SHA1
6-255	Reserved for future use

- Auth Length: indicates the authentication field length, including authentication type and authentication length field, in bytes.
- Authentication Data: indicates the authentication data area.

The authentication field is optional in packets. The authentication data area changes according to the authentication type and is not mentioned here. As defined by the protocol, BFD control packets are encapsulated in UDP packets, using destination port 3784.

BFD echo packets provide a fault detection mechanism without using BFD control packets. One end sends a BFD echo packet to the peer end, which then returns the received BFD echo packet back without processing it. Therefore, the BFD protocol does not define the format of BFD echo packets. The only requirement is that the transmitting end can distinguish between sessions based on packet contents.

BFD echo packets are encapsulated in UDP packets, using destination port 3785. The IP address of the transmitting interface is the destination IP address. The source IP address is configured and it should not cause ICMP redirection.

1.3.3 BFD Detection Mechanism

Two systems set up a BFD session and periodically send BFD control packets along the path between them. If one system does not receive BFD control packets within a specified period, the system considers that a fault has occurred on the path.

BFD control packets are encapsulated in UDP packets. In the initial phase of a BFD session, the transmitting and receiving systems negotiate with each other over parameters carried in BFD control packets, such as discriminators, expected minimum intervals for sending and receiving BFD control packets, and local BFD session status. When negotiations are successful, both systems send BFD control packets at the negotiated intervals on the path between them.

To meet the requirement for fast detection, the BFD draft defines that the intervals for sending and receiving BFD control packets are expressed in microseconds. Limited by the current processing capability, BFD-enabled devices of most manufacturers process BFD control packets at millisecond level. The speed is transformed to microsecond during internal processing.

BFD provides the following detection modes:

- Asynchronous mode: In asynchronous mode, two systems periodically send BFD control packets to each other. If one system receives no packets consecutively, the system considers the BFD session Down.
- Demand mode: If multiple BFD sessions exist in a system, periodically sending costs of BFD control packets affects system running. To solve this problem, use the demand mode. In demand mode, after BFD sessions are set up, the system does not periodically send BFD control packets. The system detects connectivity using other mechanisms such as the Hello mechanism of a routing protocol and hardware detection to reduce the costs of BFD sessions.

An auxiliary function of two modes is the echo function. When the echo function is activated, a BFD control packet is sent as follows: The local system sends a BFD control packet, and the remote system loops the packet back through its forwarding channel. If consecutive echo packets are not received, the BFD session is declared Down. The echo function can work with the asynchronous mode or demand mode.

Currently, Huawei S series switches only support the passive echo function.

1.3.4 BFD Session Establishment Modes

BFD sessions can be set up statically and dynamically. Static and dynamic BFD sessions differ in that local and remote discriminators are configured in different modes. BFD uses local and remote discriminators in control packets to differentiate BFD sessions.

- Statically establishing a BFD session

  BFD session parameters, including the local and remote discriminators, are specified using commands. Then a request for BFD session establishment is distributed manually.

  The following example shows the configuration of single-hop BFD on a Layer 2 link.

Step 1 Enable BFD globally.

```
<Quidway> system-view
[Quidway] bfd
[Quidway-bfd] quit
```

Step 2 Configure a single-hop BFD session.

```
[Quidway] bfd atob bind peer-ip default-ip interface gigabitethernet 1/0/1
[Quidway-bfd-session-atob] discriminator local 1
[Quidway-bfd-session-atob] discriminator remote 2
[Quidway-bfd-session-atob] commit
[Quidway-bfd-session-atob] quit
```

Step 3 Configure a single-hop BFD session on the peer switch.

```
[Quidway] bfd btoa bind peer-ip default-ip interface gigabitethernet 1/0/1
[Quidway-bfd-session-btoa] discriminator local 2
[Quidway-bfd-session-btoa] discriminator remote 1
[Quidway-bfd-session-btoa] commit
[Quidway-bfd-session-btoa] quit
```

Step 4 After the preceding configurations are complete, run the **display bfd session** command on the switches. You can view that a single-hop BFD session is set up and the status is Up.

```
<Quidway> display bfd session all verbose
--------------------------------------------------------------------------
Session MIndex : 4097     (One Hop) State : Up       Name : atob
--------------------------------------------------------------------------
 Local Discriminator    : 1            Remote Discriminator   : 2
 Session Detect Mode    : Asynchronous Mode Without Echo Function
 BFD Bind Type          : Interface(GigabitEthernet1/0/1)
 Bind Session Type      : Static
 Bind Peer IP Address   : 224.0.0.184
 NextHop Ip Address     : 224.0.0.184
 Bind Interface         : GigabitEthernet1/0/1
 FSM Board Id           : 0              TOS-EXP              : 7
 Min Tx Interval (ms)   : 1000           Min Rx Interval (ms)   : 1000
 Actual Tx Interval (ms): 13000          Actual Rx Interval (ms): 13000
 Local Detect Multi     : 3              Detect Interval (ms)   : 3000
 Echo Passive           : Disable        Acl Number           : -
 Destination Port       : 3784           TTL                  : 255
 Proc Interface Status  : Disable        Process PST          : Disable
 WTR Interval (ms)      : -
 Active Multi           : 3
 Last Local Diagnostic  : No Diagnostic
 Bind Application       : No Application Bind
 Session TX TmrID       : -              Session Detect TmrID   : -
 Session Init TmrID     : -              Session WTR TmrID      : -
 Session Echo Tx TmrID  : -
 PDT Index              : FSM-0 | RCV-0 | IF-0 | TOKEN-0
 Session Description    : -
--------------------------------------------------------------------------

    Total UP/DOWN Session Number : 1/0
```

- Dynamically establishing a BFD session

  When a BFD session is set up dynamically, the system processes the local and remote discriminators as follows:

  - Dynamically allocated local discriminator

    When an application triggers dynamic setup of a BFD session, the system allocates a value as the local discriminator of the BFD session. Then the local system sends a BFD control packet with Remote Discriminator as 0 to the remote system to negotiate on the BFD session.

- Self-learned remote discriminator

  When one end of a BFD session receives a BFD control packet with Remote Discriminator as 0, this end checks the BFD control packet. If the packet matches the local BFD session, this end learns the value of Local Discriminator in the received BFD control packet to obtain the remote discriminator.

The following example shows the configuration of BFD for OSPF.

Step 5 Configure basic OSPF functions. The configuration details are not mentioned here.

Step 6 Enable BFD globally.

```
<Quidway> system-view
[Quidway] bfd
[Quidway-bfd] quit
```

Step 7 Configure BFD for OSPF.

```
[SwitchA] ospf
[SwitchA-ospf-1] bfd all-interfaces enable
[SwitchA-ospf-1] quit
```

Step 8 After the preceding configurations are complete, run the **display ospf bfd session all** command on SwitchA and SwitchB and you can view that the BFD session status is Up.

```
[Quidway] display ospf bfd session all
        OSPF Process 1 with Router ID 1.1.1.1
 Area 0.0.0.0 interface  3.3.3.1(Vlanif20)'s BFD Sessions

 NeighborId:2.2.2.2         AreaId:0.0.0.0         Interface:Vlanif20
 BFDState:up               rx    :1000           tx      :1000
 Multiplier:3              BFD Local Dis:8195     LocalIpAdd:3.3.3.1
 RemoteIpAdd:3.3.3.2        Diagnostic Info:No diagnostic information

 Area 0.0.0.0 interface 1.1.1.1(Vlanif10)'s BFD Sessions

 NeighborId:3.3.3.3         AreaId:0.0.0.0         Interface:Vlanif10
 BFDState:up               rx    :1000           tx      :1000
 Multiplier:3              BFD Local Dis:8194     LocalIpAdd1:1.1.1.1
 RemoteIpAdd:1.1.1.2        Diagnostic Info:No diagnostic information
```

----End

1.3.5 Link Types Detected by BFD

Generally, the link types detected by BFD are as follows:

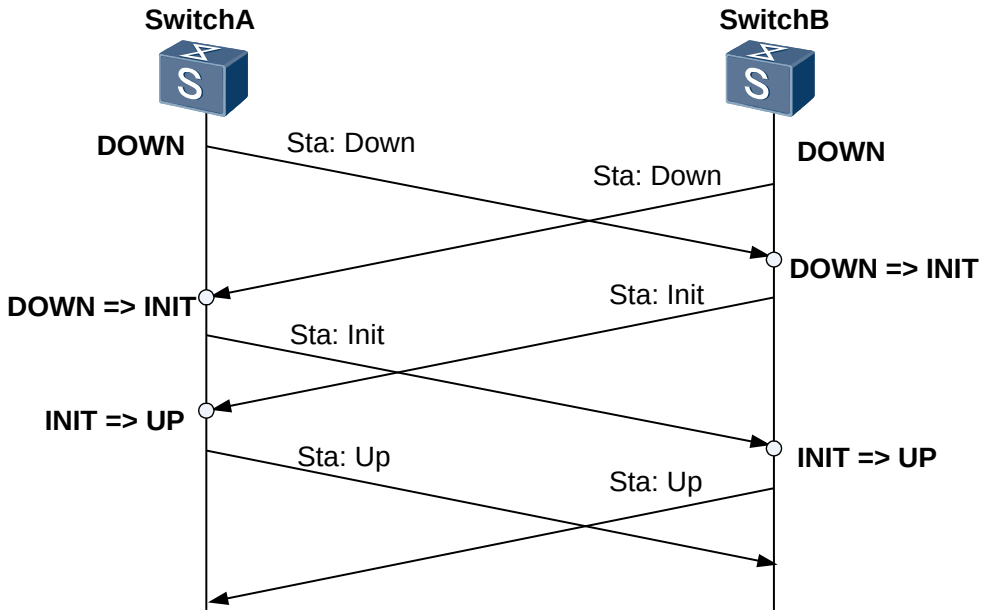
- IP link
- Eth-Trunk: includes the links of member interfaces and the aggregation interface in the same port group. BFD sessions used to detect the trunk member interfaces and the trunk interface are independent from each other and can detect links at the same time.
- VLANIF: includes the links of VLANIF interfaces and VLANIF member interfaces. BFD sessions used to detect the VLANIF interface and VLANIF member interfaces are independent from each other and can detect links at the same time.
- MPLS LSP: To detect connectivity of MPLS LSPs, two modes of BFD sessions can be used.
  Static BFD sessions: The negotiation of a BFD session is performed through the local

discriminator and remote discriminator of a BFD session that are configured manually. Dynamic BFD session: The negotiation of a BFD session is performed through the BFD discriminator TLV in an LSP ping packet.

1.3.6 BFD Session Management

The BFD session has the following status: Down, Init, Up, and Down. The State field of a BFD control packet shows the session status. The system changes the session status based on the local session status and the received session status of the peer. The BFD state machine implements a three-way handshake for BFD session setup or deletion to ensure that the two systems detect the status change. The following uses BFD session setup as an example to describe the state machine transition process.

Figure 1-4 BFD session setup



1. SwitchA and SwitchB start BFD state machines respectively. The initial state of BFD state machine is Down. SwitchA and SwitchB send BFD control packets with the State field as Down. If BFD sessions are configured statically, the values of Remote Discriminator in BFD packets are specified. If BFD sessions are configured dynamically, the value of Remote Discriminator is set to 0.
2. After receiving the BFD packet with the State field as Down, SwitchB switches the session status to Init and sends a BFD packet with State field as Init.
3. After the local BFD session status of SwitchB changes to Init, SwitchB no longer processes the received BFD packets with the State field as Down.
4. The BFD session status change on SwitchA is similar to that on SwitchB.
5. After receiving the BFD packet with the State field as Init, SwitchB changes the local BFD session status to Up.
6. The BFD session status change on SwitchA is similar to that on SwitchB.

1.4 Applications

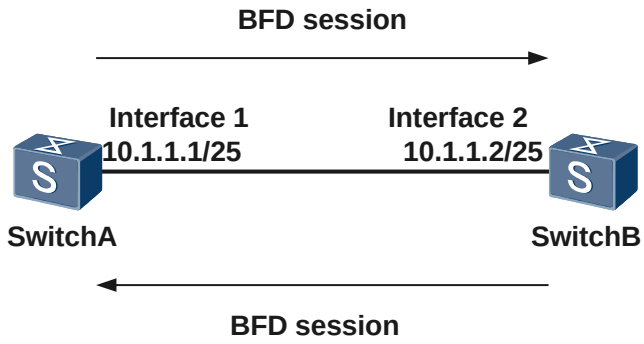
1.4.1 BFD for IP Links

You can create a single-hop or multi-hop BFD session on an IP link to rapidly detect faults:

- Single-hop BFD detects IP connectivity of the forwarding link between two directly connected systems.
- Multi-hop BFD detects IP connectivity of paths between two indirectly connected systems. These paths may span multiple hops or overlap.

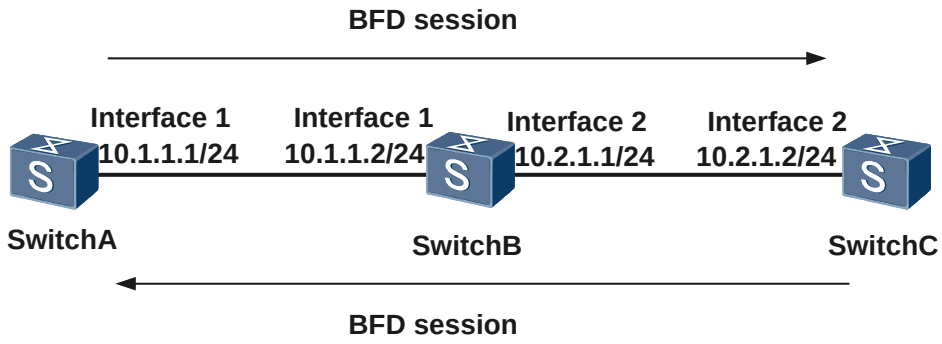
As shown in Figure 1-5, a BFD session detects a single-hop path between two devices and the BFD session is bound to the outbound interface.

Figure 1-5 Single-hop BFD for IP links



As shown in Figure 1-6, a BFD session detects a multi-hop path between SwitchA and SwitchC, and the BFD session is bound to the peer IP address but not the outbound interface.

Figure 1-6 Multi-hop BFD for IP links



1.4.2 BFD Echo Function

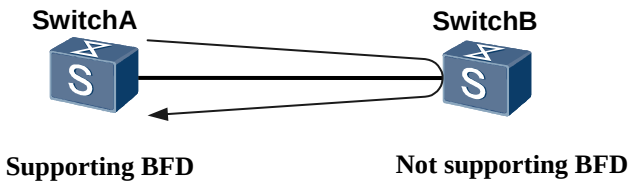
The BFD echo function detects connectivity of the forwarding link by looping back packets.

Among two directly connected devices, one device supports BFD, but the other device does not support BFD and supports only forwarding at the network layer. To rapidly detect forwarding failures between the two devices, the BFD echo function is configured on the

BFD-supporting device. The BFD-supporting device sends an Echo Request packet to the remote device. The remote device sends the Echo Request packet back along the same path to detect the connectivity of the forwarding link.

As shown in Figure 1-7, SwitchA supports BFD, whereas SwitchB does not support BFD. The BFD echo function is configured on SwitchA to detect connectivity of the single-hop path between SwitchA and SwitchB. After SwitchB receives a BFD echo packet from SwitchA, SwitchB loops back the packet at the network layer. This can rapidly detect connectivity of the direct link between SwitchA and SwitchB.

Figure 1-7 BFD echo function

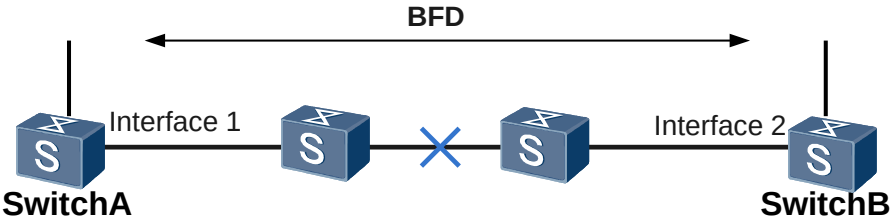


1.4.3 Association Between the BFD Session Status and the Interface Status

BFD for process interface status (PIS) associates the BFD session with the interface status. This improves sensitivity of interfaces to detect link faults and minimizes the impact of faults on indirectly connected links. When detecting a link fault, a BFD session immediately sends a Down message to the corresponding interface. The interface enters the BFD Down state. In BFD Down state, the interface can process only BFD packets. Therefore, the interface can rapidly detect link faults.

As shown in Figure 1-8, a transit device exists on a faulty link, it takes a long time for devices on two ends of the link to detect faults although they are directly connected at Layer 3. The reason is that the two devices are connected by multiple physical links. As a result, service interruption time is long. A BFD session is configured on SwitchA and SwitchB and the BFD session status is associated with the interface status. When detecting a link fault, a BFD session immediately sends a Down message to the corresponding interface. The interface enters the BFD Down state.

Figure 1-8 Association between the BFD session status and the interface status



1.4.4 BFD for Static Routes

Unlike dynamic routing protocols, static routes do not have a dedicated detection mechanism. After a fault occurs, static routes cannot detect the fault, and the network administrator must

delete the corresponding static route. BFD for static routes enables a BFD session to detect the status of the link of the static route on the public network.

Each static route can be bound to a BFD session. When a BFD session bound to a static route detects a fault (for example, the link changes from Up to Down) on a link, BFD reports the fault to the routing management module (RM). Then, the RM configures the route as inactive, indicating that the route is unavailable and deleted from the IP routing table. When the BFD session bound to the static route is successfully set up or the link of the static route recovers (that is, the link changes from Down to Up), BFD reports the event to the RM and the RM configures the static route as active, indicating that the route is available and added to the IP routing table.

1.4.5 BFD for OSPF

A link failure or topology change may lead to route recalculation; therefore, convergence of routing protocols must be shortened as much as possible to improve network performance. A feasible solution is to rapidly detect link faults and immediately notify routing protocols of the faults.

BFD for OSPF associates a BFD session with OSPF. The BFD session rapidly detects a link fault and notifies OSPF of the fault. By doing this, OSPF quickly responds to the network topology change. Table 1-5 lists OSPF convergence speed when a BFD session is bound or not.

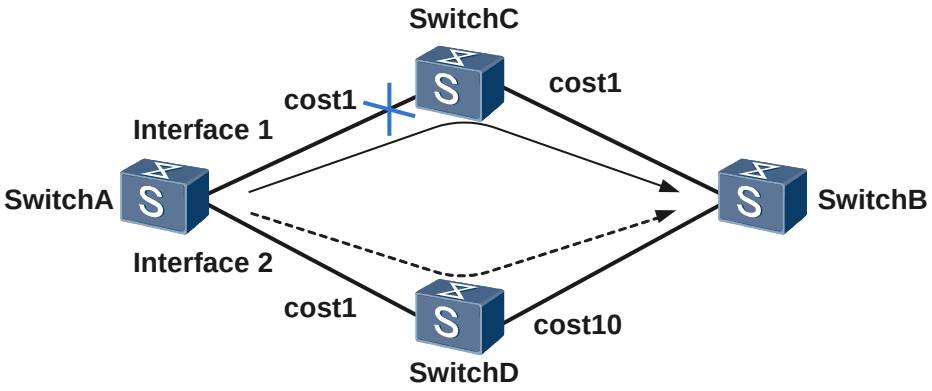
Table 1-5 OSPF convergence speed

Whether a BFD Session Is Bound	Link Fault Detection Mechanism	Convergence Speed
No	Timeout of the OSPF Hello keepalive timer	At the second level
Yes	BFD session in Down state	At the millisecond level

As shown in Figure 1-9, SwitchA establishes OSPF neighbor relationships with SwitchC and SwitchD. The outbound interface in the route from SwitchA to SwitchB is Interface 1. Packets from SwithA traverse SwitchC, and then reach SwitchB. When the OSPF neighbor is in Full state, the system instructs BFD to create a BFD session.

When a fault occurs on the link between Switch and SwitchC, the BFD session detects the fault and notifies SwitchA. SwitchA processes the neighbor Down event and recalculates the route. Then, the new outbound interface in the route is Interface 2. Packets from SwithA traverse SwitchD, and then reach SwitchB.

Figure 1-9 BFD for OSPF



1.4.6 BFD for IS-IS

Generally, the interval at which Intermediate System to Intermediate System (IS-IS) sends Hello packets is 10s. The holdtime of neighbors is three times the interval at which Hello packets are sent. If the Switch does not receive a Hello packet from its neighbor within the holddown time, the Switch deletes the neighbor relationship. That is, the Switch detects neighbor faults in seconds. The second-level detection leads to the loss of a large number of packets on a high-speed network.

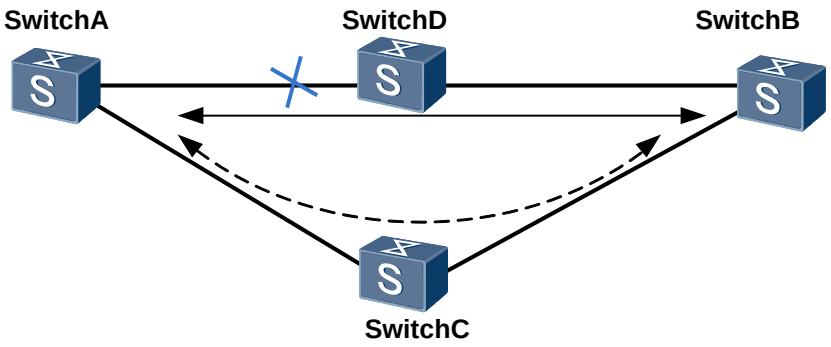
In BFD for IS-IS, BFD session setup is dynamically triggered by IS-IS but not configured manually. When detecting a fault, the BFD session notifies IS-IS through the RM. Then, IS-IS processes the neighbor Down event and rapidly updates the link state PDU (LSP) and performs the partial route calculation (PRC). This speeds up IS-IS route convergence. BFD is not used to replace the Hello mechanism of IS-IS. Instead, BFD works with IS-IS to rapidly detect link faults and to immediately notify IS-IS of route recalculation, which guides packet forwarding. Table 1-6 lists IS-IS convergence speed when a BFD session is bound or not.

Table 1-6 IS-IS convergence speed

Whether a BFD Session Is Bound	Link Fault Detection Mechanism	Convergence Speed
No	Hello mechanism	At the second level
Yes	BFD session in Down state	At the millisecond level

As shown in Figure 1-10, IS-IS is enabled on devices and association between BFD and IS-IS is enabled on SwitchA and SwitchB. When the link between SwitchA and SwitchB fails, BFD can rapidly detect the fault and report the fault to IS-IS. IS-IS then disconnects the neighbors of this interface, which triggers topology calculation. IS-IS updates LSPs so that the neighbors, for example, Switch B's neighbor SwitchC, can receive the updated LSPs from SwitchB. IS-IS fast convergence is implemented.

Figure 1-10 BFD for IS-IS



1.4.7 BFD for BGP

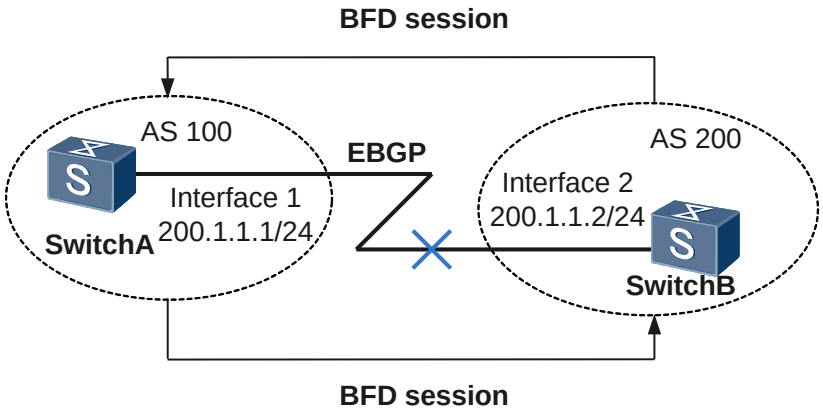
BGP enables the Switch to periodically send Keepalive packet to its peers for fault detection. Detecting a fault takes more than 1s. When traffic is transmitted at gigabit rates, long-time fault detection will cause packet loss. Association between BFD and BGP enables BFD to rapidly detect faults on links between BGP peers and reports faults to BGP, which implements fast BGP route convergence. Table 1-7 lists BGP convergence speed when a BFD session is bound or not.

Table 1-7 BGP convergence speed

Whether a BFD Session Is Bound	Link Fault Detection Mechanism	Convergence Speed
No	Keepalive packet mechanism	At the second level
Yes	BFD session in Down state	At the millisecond level

As shown in Figure 1-11, SwitchA belongs to AS 100, and SwitchB belongs to AS 200. SwitchA and SwitchB are directly connected and establish an EGBP connection. BFD detects the status of the EGBP connection between SwitchA and SwitchB. When the link between SwitchA and SwitchB becomes faulty, BFD can rapidly detect the fault and notify BGP of the fault.

Figure 1-11 BFD for BGP



1.4.8 BFD for MPLS LSPs

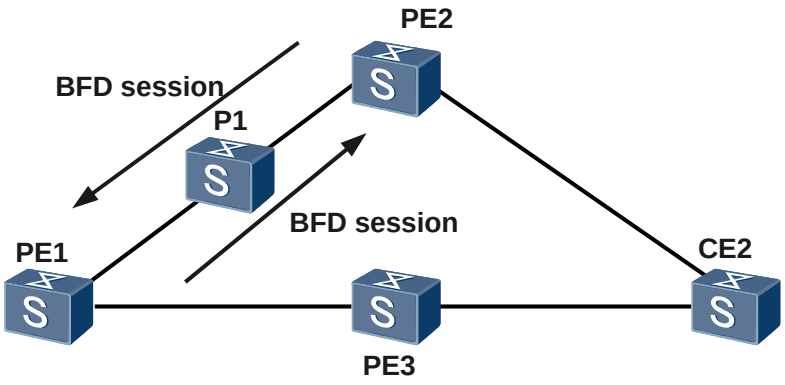
A BFD session that is established on an LSP can fast detect a fault on the LSP to provide end-to-end protection for the LSP. When a BFD session is associated with a unidirectional LSP, the reverse link can be an IP link, an LSP, or a TE tunnel. Both static and dynamic BFD sessions can detect connectivity of MPLS LSPs. Dynamic BFD sessions support only dynamic LSPs, and static BFD sessions support static and dynamic LSPs.

To detect connectivity of an LSP, the ingress and egress nodes periodically send BFD packets to each other. If the ingress or egress node does not receive BFD packets from the other end within the detection period, BFD considers the LSP as Down and sends an LSP Down message to the LSP management module (LSPM).

As shown in Figure 1-12, only traffic from PE1 to CE2 is involved in the application. When a fault occurs on the link between PE1 and P1, PE1 can detect the fault through its interface connected to P1. When a fault occurs on the link between P1 and PE2, PE1 cannot detect the fault through its interface connected to P1. BFD for dynamic LSPs needs to be configured to rapidly detect faults.

A dynamic LSP destined for PE2 is set up on PE1. BFD for dynamic LSPs is enabled and a BFD session is set up. In addition, policies of Virtual Private Network fast reroute (VPN FRR) are configured on PE1, and the protection path between PE1 and PE3 is specified. When a fault occurs on the link between PE1 and P1 or between P1 and PE2, PE1 fast detects the LSP fault and triggers VPN FRR switching. The traffic is then switched to the path PE1-PE3-CE2.

Figure 1-12 BFD for MPLS LSPs



1.4.9 BFD for MPLS TE

BFD for TE is an end-to-end fast detection mechanism in MPLS TE, and rapidly detects faults along the link through which an MPLS TE tunnel passes. Traditional detection mechanisms, including RSVP Hello mechanism or RSVP summary refresh (Srefresh) mechanism, detect faults at slow speeds. BFD uses the fast packet transmission mode to quickly detect faults on MPLS TE tunnels. When an MPLS TE tunnel fails, BFD triggers fast switchover to protect services.

BFD detects faults on the following types of TE tunnels:

- Static BFD for CR-LSPs

  Static BFD for CR-LSPs uses a manually configured BFD session to rapidly detect faults on CR-LSPs.
- Static BFD for TE tunnels

  Static BFD for TE tunnels uses a manually configured BFD session to monitor the whole TE tunnel and trigger traffic switchover of applications such as VPN Fast Reroute (FRR).
- Dynamic BFD for CR-LSPs

  Dynamic BFD for CR-LSPs has the same function as static BFD for CR-LSPs. However, dynamic BFD for CR-LSPs uses a dynamic BFD session.

BFD for TE tunnels and BFD for CR-LSPs report faults to different objects. In BFD for TE, BFD notifies applications such as VPN of faults and triggers traffic switchover between different tunnel interfaces. In BFD for CR-LSPs, BFD notifies TE tunnels of faults and triggers traffic switchover between different CR-LSPs in the same TE tunnel.

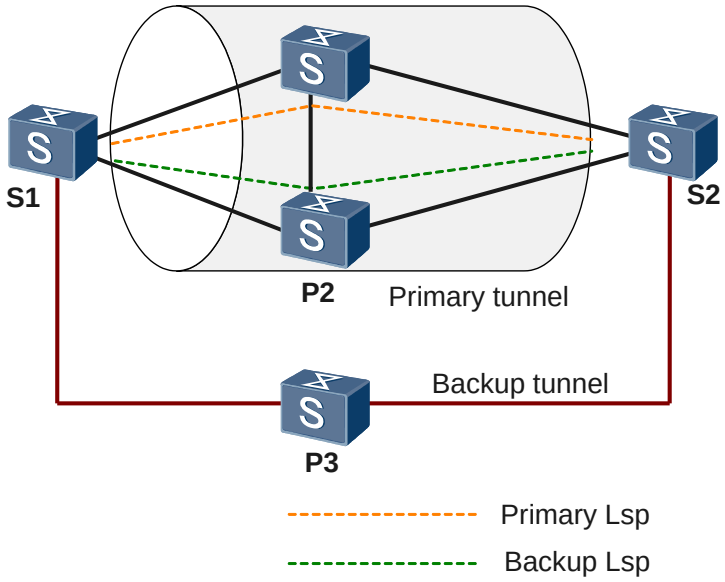
BFD is bound to an LSP and a BFD session is set up between the ingress and the egress. A BFD packet is sent by the ingress to the egress through an LSP. Then the egress responds to the BFD packet. In this manner, a BFD session at the ingress can rapidly detect the status of the path through which the LPS passes. After BFD detects a link fault, it notifies the LSP management module. Then, traffic is switched to the backup LSP.

- BFD for CR-LSPs

  As shown in Figure 1-13, a primary LSP and a backup LSP are set up between S1 and S2. On S1, a BFD session is set up from S1 to S2 to detect faults on the primary LSP of the TE tunnel. When a fault occurs on the primary LSP, the BFD session rapidly notifies S1. After learning the fault, S1 fast switches traffic to the backup LSP to ensure nonstop traffic transmission.
- BFD for TE tunnels

  As shown in Figure 1-13, the primary LSP is established along the path S1->P2->S2, and the backup LSP is established along the path S1->P3->S2. A BFD session is set up along the path S1->P2->S2 to monitor the primary LSP. When a fault occurs on the primary LSP, the BFD session rapidly notifies S1. After learning the fault, S1 fast switches traffic to the backup LSP to ensure nonstop traffic transmission.

Figure 1-13 BFD for MPLS TE



1.4.10 BFD for VRRP

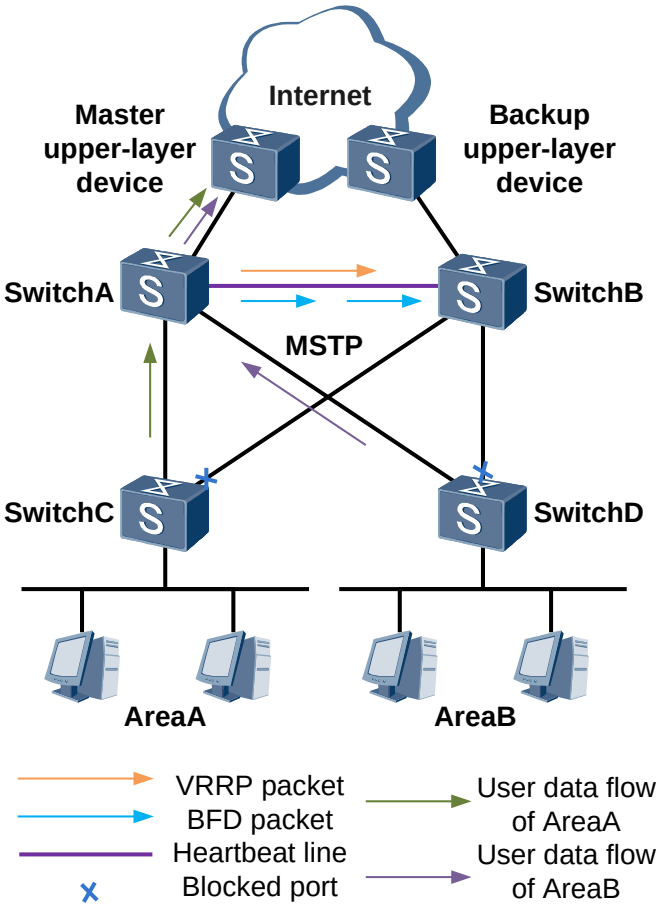
When the VRRP master fails, the VRRP backup with the highest priority should take over traffic within a short time to shorten service interruption.

When the VRRP master fails, VRRP determines whether to perform preemption based on the timeout interval. The switching takes more than 1s. BFD can be used to rapidly detect the master status and shorten traffic interruption. BFD detects real IP addresses of the master and backup devices during communication. If communication is abnormal, the backup device considers that the master device is Down and becomes the master device. a VRRP backup group implements a master/backup VRRP switchover rapidly by tracking the BFD session status. The switchover time is within 50 milliseconds.

As shown in Figure 1-14, SwitchA and SwitchB establish a VRRP group. SwitchA functions as the master and SwitchB functions as the backup. User traffic is transmitted through SwitchA. A BFD session is established between SwitchA and SwitchB. The VRRP group tracks the BFD session status. When the BFD session status changes, the priority of the VRRP group is changed and then a master/backup VRRP switchover is triggered.

When a BFD session detects a fault, BFD reports a Down event to VRRP. Then the priority of SwitchB increases above the priority of SwitchA. SwitchB becomes the master switch immediately and subsequent user traffic is forwarded through SwitchB. In this manner, fast master/backup VRRP switchover is performed.

Figure 1-14 BFD for VRRP



1.4.11 BFD for PIM

If a DR on the shared network segment becomes faulty, PIM neighbor relationships time out, and a new DR election is triggered among PIM neighbors. Consequently, multicast data transmission is interrupted. The interruption period, usually in seconds, is at least as long as the timeout interval of the neighbor relationship.

After detecting a fault on the peer, BFD immediately instructs the PIM module to trigger a new DR election without waiting for timeout of the neighbor relationship.

BFD for PIM can rapidly detect faults on the Assert winner and is also applicable to Assert election on a shared network segment. Table 1-8 lists PIM convergence speed when a BFD session is bound or not.

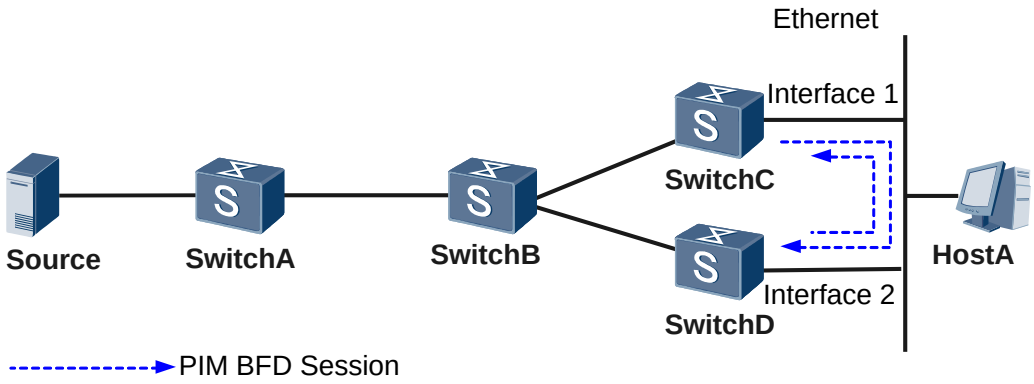
Table 1-8 PIM convergence speed

Whether a BFD Session Is Bound	Link Fault Detection Mechanism	Convergence Speed
No	Neighbor relationship timeout	At the second level
Yes	BFD session in Down state	At the millisecond level

As shown in Figure 1-15, on the shared network segment connected to user hosts, downstream interface Interface1 on SwitchC and downstream interface Interface2 on SwitchD establish a PIM BFD session and send BFD control packets to detect the link status.

SwitchC functions as the DR and its downstream interface Interface1 is responsible for forwarding multicast data. If Interface1 becomes faulty, BFD fast notifies the RM of the session status, and the RM notifies the PIM module. The PIM module then triggers a new DR election. SwitchD quickly begins functioning as the new DR and its downstream interface Interface2 forwards multicast data to the receivers.

Figure 1-15 BFD for PIM



1.5 Troubleshooting Cases

1.5.1 BFD Session Cannot Become Up

Common Causes

- The link carrying the BFD session is faulty. As a result, BFD packets cannot be exchanged.
- The BFD session flaps frequently.

Procedure

Step 1 Run the **display current-configuration configuration bfd-session** command to check whether the local and remote discriminators at both ends match.

- If the local and remote discriminators at both ends match, go to Step 2.
- If the local and remote discriminators at both ends do not match, run the **discriminator** command to correctly configure local and remote discriminators, and then run the **display bfd session all** command to check whether the BFD session is Up.
  - If the value of the State field is Up, the BFD session has been established.
  - If the value of the State field is not Up, go to Step 2.

Step 2 Run the **display current-configuration configuration bfd-session** command to check whether the BFD detection time is longer than the delay before the BFD session becomes Up.

Detection time = Received Detect Multi of the remote system x Max (Local DMTI/Received RMRI) Detect Multi is the local detection multiplier, which is set by using the **detect-multiplier** command. The Required Min Rx Interval (RMRI) is the minimum interval for receiving BFD packets, which is set by using the **min-rx-interval** command. The Desired Min Tx Interval (DMTI) is the minimum interval for sending BFD packets, which is set by using the **min-tx-interval** command.

The link delay can be obtained using the ping or tracert mechanism.

If the BFD detection time is shorter than the delay before the BFD session becomes Up, run the **detect-multiplier**, **min-rx-interval**, and **min-tx-interval** commands to increase the BFD detection time to be longer than the delay.

----End

1.5.2 BFD Detection Result Affects Forwarding on an Interface

Common Cases

The BFD session is associated with the interface status.

Procedure

Step 1 Run the **display interface** *interface-type interface-number* command to check the physical status of the interface bound to the BFD session.

- If the value of **Line protocol current state** is **UP (BFD status down)**, the interface status is affected by the BFD session status. When the BFD session detects a link fault, the interface enters the **BFD status down** state. Go to Step 2.
- If the value of **Line protocol current state** is **UP** but the interface cannot forward packets, the forwarding module is working properly.

Step 2 Run the **display bfd session all** command to check the BFD session status.

If the BFD session status is **Down**, go to Step 3.

Step 3 Run the **display current-configuration configuration bfd-session** command to check the BFD session configuration and check whether the **process-interface-status** command is used.

If the process-interface-status command is used, the interface enters the UP (BFD status down) state when the BFD session detects a link fault and enters the Down state. As a result, the interface cannot forward packets.

----End

1.6 FAQs

1.6.1 Association Between a BFD Session and the Interface Status Is Configured on the Devices Equipped with the FSU, and the WTR Time Is Set. Why Does BFD Flapping Occur Sometimes?

In V100R002, when the BFD session becomes Up, the blocked interface is unblocked. When the WTR time is set and a BFD session changes from Down to Up, BFD sends a notification to the upper-layer application only after the WTR time expires. The blocked interface cannot send packets. Consequently, the BFD session becomes Down again. BFD flapping occurs. V100R003 and later versions solve this problem.

1.6.2 What Is the BFD Detection Time?

In versions earlier than V100R006, the switch with the FPIC installed supports the BFD detection time of 10 ms, 20 ms, 30 ms, 50 ms, 100 ms, and 1s. The switch without the FPIC installed supports only the BFD detection time of 1s. In versions later than V100R006, the switch without the FPIC installed supports the BFD detection time of at least 100 ms because the switch provides super tasks.

1.7 Reference Standards and Protocols

Table 1-9 Acronym

Acronym	Full Name
ISIS	Intermediate System-Intermediate System
BFD	Bidirectional Forwarding Detection
PE	Provider Edge Router
CE	Customer Edge Router
OSPF	Open Shortest Path First
TE	Traffic Engineer
VRRP	Virtual Router Redundancy Protocol
MPLS	Multi Protocol Label Switching
LSP	Label switched path
PIM	Protocol Independent Multicast
BGP	Border Gateway Protocol