Bidirectional Forwarding Detection (BFD) in SONiC using GNS3

OCT 26, 2023

Table of Contents

Introduction

The document provides a comprehensive guide on Bidirectional Forwarding Detection BFD configuration in SONiC.

Intended Audience

This document is tailored for network administrators and network engineers interested in configuring **BFD** in SONiC. It is designed for individuals with a solid understanding of networking principles. Whether you are a network engineer, network operator, or vendor, this document aims to provide you with practical, step-by-step guidance, and best practices for deploying, configuring, and setting up BFD for SONiC devices using the GNS3 network simulation tool.

Bidirectional Forwarding Detection (BFD)

Bidirectional Forwarding Detection (BFD) is a network protocol designed to quickly identify and respond to the faults in the network path, including link failures in milliseconds or even microseconds, and provide this information to routing protocols, enabling them to react swiftly to reroute traffic and maintain network availability.

By sending rapid failure detection notices to the routing protocols BFD runs independently from any other (routing) protocols. Once it's up and running, routing protocols like OSPF, EIGRP, BGP, HSRP, MPLS LDP, etc. can be configured to use BFD for link failure detection instead of their mechanisms. When the link fails, BFD will inform the protocol in the local router to initiate the routing table recalculation process. BFD contributes to greatly reducing overall network convergence time.

Operating Modes of BFD

Bidirectional Forwarding Detection (BFD) operates in three modes: Asynchronous (Independent) Mode and Demand Mode. These modes define how BFD packets are transmitted and received between network devices. Here's an overview of each mode:

1. Asynchronous (Independent) Mode

In the Asynchronous Mode, BFD packets are sent periodically at a fixed interval, independent of the traffic flow on the network link. This mode is typically used for proactive fault detection and monitoring. Key characteristics of Asynchronous Mode include:

- **Periodic Packet Transmission**: BFD packets are sent at regular intervals, determined by the configured timers (e.g., "Required Minimum Echo Transmit Interval" and "Desired Minimum Echo Receive Interval").
- **Constant Monitoring**: BFD continuously monitors the link for faults. If the link or the remote endpoint becomes unresponsive, BFD can quickly detect the failure and trigger a reaction, such as routing protocol convergence or link failover.
- **Low Latency Detection**: Asynchronous BFD is designed for low-latency fault detection, making it suitable for real-time applications and services that require minimal network interruption.

2. Demand Mode

In Demand Mode, BFD packets are not sent periodically but are transmitted on-demand or in response to specific network events or requests. This mode is typically used for scenarios where resources need to be conserved or when it's not necessary to monitor the link continually. Key characteristics of Demand Mode include:

- **Packet Transmission on Request**: BFD packets are sent when explicitly requested by the network administrator or in response to network events. This conserves bandwidth and reduces the overhead associated with periodic packet transmission.
- **Resource-Efficient**: Demand Mode is suitable for scenarios where network resources need to be conserved, as it avoids the continuous transmission of BFD packets when not actively required.
- **Selective Monitoring**: Demand BFD allows for selective monitoring of specific links or paths as needed, rather than monitoring all links continuously

3. Echo Mode

In Echo Mode, BFD sessions are established with a remote peer, similar to other BFD modes. However, in Echo Mode, the primary purpose is to verify the bidirectional reachability of a network path rather than just monitoring the path for faults. Echo Mode is often used to ensure that not only can BFD packets be exchanged between peers, but also that these packets can be properly received and echoed back by the remote device. This helps verify the health and round-trip functionality of the entire network path. Key characteristics of Echo Mode include:

- The local BFD device sends an "echo" packet to the remote BFD device.
- The remote BFD device, upon receiving the echo packet, echoes it back to the local device.
- The local BFD device confirms the successful receipt of the echoed packet. If it is received as expected, bidirectional reachability is verified.

BFD Intervals

Detect-multiplier (2-255)

Configures the detection multiplier to determine packet loss. The remote transmission interval will be multiplied by this value to determine the connection loss detection timer. The default value is 3.

Receive-interval (10-60000)ms

Configures the minimum interval that this system can receive control packets. The default value is 300 milliseconds.

Transmit-interval (10-60000)ms

The minimum transmission interval (less jitter) that this system wants to use to send BFD control packets. Defaults to 300ms.

Echo receive-interval <disabled|(10-60000)ms>

Configures the minimum interval that this system can receive echo packets. Disabled means that this system doesn’t want to receive echo packets. The default value is 50 milliseconds.

Echo transmit-interval (10-60000)ms

The minimum transmission interval (less jitter) that this system wants to use to send BFD echo packets. Defaults to 50ms.

Why Do We Need BFD?

To minimize the impact of devices and link failures on services and enhance network reliability, a network device must be able to quickly detect faults in communication with adjacent devices. Measures can then be taken to promptly rectify the faults to ensure service continuity.

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 provide the link failure detection mechanism. In this case, applications use the Hello mechanism of the upper-layer protocol to detect faults, which usually take seconds. This long detection period causes severe packet loss when traffic is transmitted at gigabit rates. On a Layer 3 network, the Hello mechanism cannot detect faults for all routes, such as static routes. This means that a fault between interconnected systems is difficult to locate.

BFD addresses these issues and provides fast fault detection independent of media and routing protocols. BFD is useful because it can:

- Implement light-load fault detection, which takes only milliseconds, enhancing reliability.
- Quickly detect a broad range of faults, including interface, data link, and forwarding engine faults.
- Provide uniform detection for all media and protocol layers in real-time, without depending on hardware.

BFD for OSPF

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

BFD Session	Link Fault Detection	Convergence Time
Not bound	Tiomeout of the OSPF Hello keepalive timer.	Second
Bound	BFD session in Downstate	Milliseconds

BFD for OSPF associates a BFD session with OSPF. The BFD session quickly detects a link fault and notifies OSPF of the fault. OSPF then quickly responds to the network topology change. Figure 1 shows the OSPF convergence time.

R1 and R2 are configured to use BFD and will send control packets to each other. OSPF remains the same, it's sending its OSPF packets.

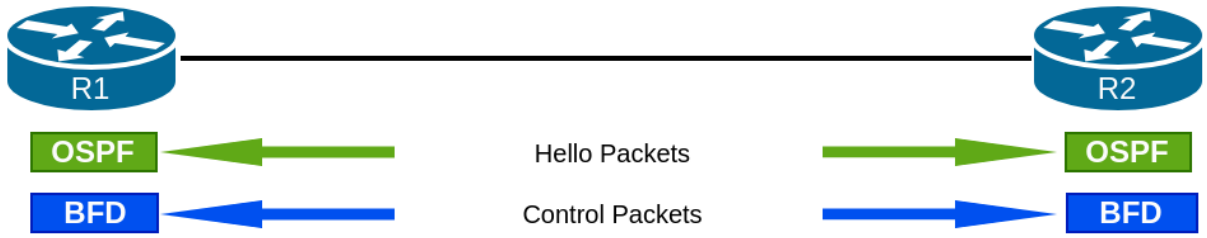

Figure1: BFD Mechanism

Once the link fails, this will happen:

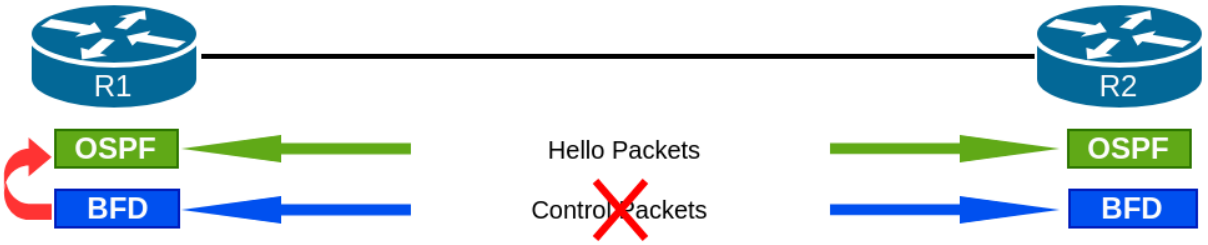

Figure2: BFD Mechanism (Link failure)

When BFD doesn't receive its control packets anymore it realizes we have a link failure and it will report this to OSPF. OSPF will then tear down the neighbor's adjacency.

Network Topology

The network topology shown in Figure 3, three SONiC switches, specifically designated as BFD-1, BFD-2, and BFD-3, are strategically used. OSPF (Open Shortest Path First) protocol for unicast routing is configured between switches, with all switches residing within OSPF Area 0. This setup ensures fast and efficient network convergence by configuring intervals (hello interval, dead interval, and transmission interval) more efficiently providing low overhead, and a short duration of detecting a failure in the forwarding path between two adjacent routers.

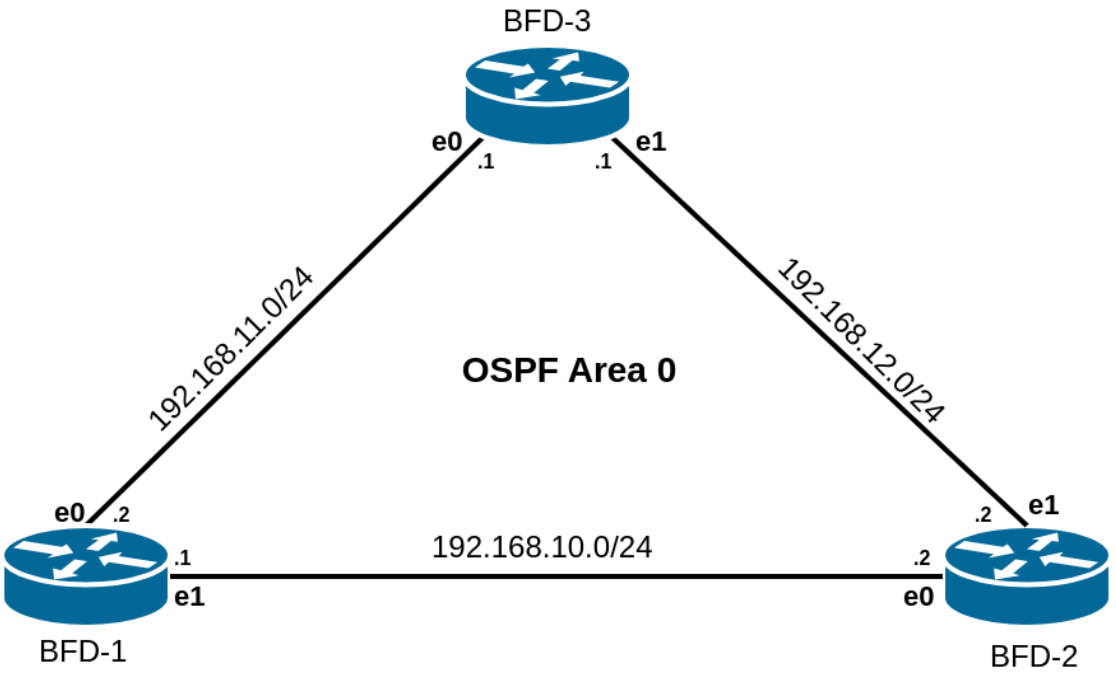

Figure 3: BFD Topology

BFD Configurations

Configurations of the BFD-1 peer are shown below. BFD-2 and BFD-3 can be configured in a similar way.

BFD-1

SONiC Native Configurations

Step 1. Configure IPs on interfaces.

```
admin@sonic:~$ sudo config interface ip rem Ethernet0 10.0.0.0/31
admin@sonic:~$ sudo config interface ip rem Ethernet4 10.0.0.2/31
admin@sonic:~$ sudo config interface ip add Ethernet0 192.168.11.2/24
admin@sonic:~$ sudo config interface ip add Ethernet4 192.168.10.1/24
```

FRR Configurations

Step 2. Configure OSPF.

```
sonic# conf terminal
sonic(config)# router ospf
sonic(config-router)# network 192.168.11.0/24 area 0
sonic(config-router)# network 192.168.10.0/24 area 0
sonic(config-router)# exit
sonic(config)# exit
sonic# write
```

Step 3. Configure BFD.

- Establish a BFD session between BFD Peers.
- Enable BFD protocol on BFD peer interfaces.

```
sonic# conf
sonic(config)# bfd
sonic(config-bfd)# peer 192.168.11.1
sonic(config-bfd-peer)# exit
sonic(config-bfd)# peer 192.168.10.2
sonic(config-bfd-peer)# exit
sonic(config-bfd)# exit
sonic(config)# exit
sonic(config)# interface Ethernet4
sonic(config-if)# ip ospf bfd
sonic(config-if)# exit
sonic(config)# interface Ethernet0
sonic(config-if)# ip ospf bfd
sonic(config-if)# exit
sonic(config)# exit
sonic# write
```

Results

- First, OSPF neighborship among the three switches was configured as shown in figure 3, confirming that they are communicating and sharing routing information effectively.

```
sonic# show ip ospf
                    OSPF Routing Process, Router ID: 192.168.11.2
                    Supports only single TOS (TOS0) routes
                    This implementation conforms to RFC2328
                    RFC 1583 Compatibility flag is disabled
                    OpaqueCapability flag is disabled
                    Initial SPF scheduling delay 0 millisec(s)
                    Minimum hold time between consecutive SPFs 50
                   millisec(s)
                    Maximum hold time between consecutive SPFs 5000
                   millisec(s)
                    Hold time multiplier is currently 1
                    SPF algorithm last executed 4m48s ago
                    Last SPF duration 0.032s
                    SPF timer is inactive
                    LSA minimum interval 5000 msecs
                    LSA minimum arrival 1000 msecs
                    Write Multiplier set to 20
                    Refresh timer 10 secs
```

- The direct link between BFD-1 and BFD-2 was disabled. However, the default behavior of OSPF at BFD-1 led to a delay of over 30 seconds before it recognized the link failure and rerouted traffic through an alternative path. This 30-second delay resulted in significant data loss within the network, highlighting the need for a more efficient failover mechanism.

- The BFD peer's information provides statistics of the BFD peer's status, uptime, and various timing intervals. These intervals include transmit and receive intervals, a detection multiplier, and the intervals associated with the remote peer. It essentially gives you detailed information on the status and communication parameters of BFD peers.

```
sonic# show bfd peer
                    BFD Peers:
                          peer 192.168.11.1 vrf default
                                ID: 193976014
                                Remote ID: 4246948269
                                Active mode
                                Status: up
                                Uptime: 14 second(s)
                                Diagnostics: ok
                                Remote diagnostics: ok
                                Peer Type: configured
                    Local timers:
                                Detect-multiplier: 3
                                Receive interval: 300ms
                                Transmission interval: 300ms
                                Echo receive interval: 50ms
                                Echo transmission interval: disabled
                    Remote timers:
                                Detect-multiplier: 3
                                Receive interval: 300ms
                                Transmission interval: 300ms
                                Echo receive interval: 50ms

                    peer 192.168.10.2 vrf default
                                ID: 2173032337
                                Remote ID: 178524670
                                Active mode
                                Status: up
                                Uptime: 14 second(s)
                                Diagnostics: ok
                                Remote diagnostics: ok
                                Peer Type: configured
                          Local timers:
                                Detect-multiplier: 3
                                Receive interval: 300ms
                                Transmission interval: 300ms
                                Echo receive interval: 50ms
                                Echo transmission interval: disabled
                          Remote timers:
                                Detect-multiplier: 3
                                Receive interval: 300ms
                                Transmission interval: 300ms
                                Echo receive interval: 50ms
```

- After enabling BFD on all three switches, the direct link between peer BFD-1 and peer BFD-2 was once again deactivated. This time, BFD-1 reacted swiftly, taking only a fraction of a second to redirect the traffic through an alternative path. This improvement demonstrates a substantial reduction in the time required to reroute network traffic in the event of a link failure.

- Ping results from BFD-1 to BFD-2.

```
sonic# ping 192.168.11.2
PING 192.168.11.2 (192.168.11.2) 56(84) bytes of data.
64 bytes from 192.168.11.2: icmp_seq=1 ttl=64 time=0.053 ms
64 bytes from 192.168.11.2: icmp_seq=2 ttl=64 time=0.114 ms
64 bytes from 192.168.11.2: icmp_seq=3 ttl=64 time=0.156 ms
64 bytes from 192.168.11.2: icmp_seq=4 ttl=64 time=0.071 ms
```

References

1. https://docs.frrouting.org/en/latest/bfd.html
2. https://www.cisco.com/c/en/us/td/docs/ios/12_0s/feature/guide/fs_bfd.html
3. https://www.packetcoders.io/what-is-bfd/
4. https://networklessons.com/cisco/ccie-routing-switching/bidirectional-forwarding-detection-bfd
5. https://info.support.huawei.com/info-finder/encyclopedia/en/BFD.html