BFD Workgroup
Internet-Draft
Intended status: Standards Track
Expires: July 27, 2020

S. Boutros, Ed.
Ciena
A. Dubey
VMware
R. Rahman
Cisco

January 24, 2020

Service Redundancy using BFD
draft-adubey-bfd-service-redundancy-03

Abstract

In a data center, when multiple routing/service nodes are providing single active redundancy for a set of L2, L3 and/or L4-L7 services. Both non-revertive and revertive fail over modes are required for the services. This draft describes a method to achieve the non-revertive and revertive fail over modes for services using Bidirectional Forwarding Detection (BFD).

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 July 27, 2020.

Copyright Notice

Table of Contents

1. Introduction

This document describes how can a group of service/routing nodes in a data center providing single active redundancy for multiple L2/L3 and/or L4/L7 services, can use BFD protocol to support non-revertive as well as revertive fail over mode.

Typically, BFD is used between the group of service nodes to verify the connectivity as well as the aliveness of the service nodes. The assignment of which node in the group is the primary designated forwarder for a given service can be determined using a centralized or distributed control plane.

The use of BFD will be to communicate the set of services that are being currently active on a given service node to the other service nodes. On a given node failure, for a given service the backup node will take over. If the service was configured to have a non-revertive fail over mode, then the backup node should continue to perform the service forwarding even after the primary node recovers and comes back up. In order to do that, the backup node MUST inform the primary node that it is currently active for the service. This is achieved through the extension we are proposing to the BFD protocol as will be described in the following sections.

It is to be noted that for revertive fail over mode of operation, the primary node should be able to take over the active role from the backup node when the primary node goes back to an operational state.

This can be as well communicated using the BFD session establishment between the primary node and the backup node.

1.1. Terminology

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 [RFC2119].

2. Solution Overview

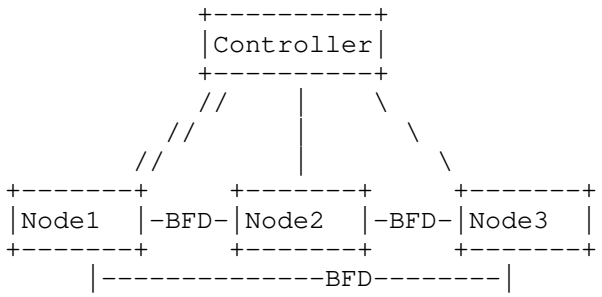

Figure 1: Solution Overview

Figure 1 shows 3 routing nodes using BFD to implement the single active redundancy for revertive and non-revertive services. More than 3 routing nodes can be used.

Multiple L2/L3 and/or L4/L7 services are offered in a data center by a set of routing/service nodes providing single active redundancy. The provisioning of the services can be done using a centralized control plane implemented in a controller or using a distributed dynamic control plane.

2.1. Node failover

An implementation MAY choose to support only node failover and not a per service failover. A node can be primary or backup for a given service. On a primary node failure, all non-revertive and revertive services will become active on the backup node.

In figure 1, lets assume that Node1 is the primary node for a set A of non-revertive services with node2 as backup, and another set B of non-revertive services with Node3 as backup. As well, Node1 is primary for a set C of revertive services with Node2 as backup and, another set D of revertive services with Node3 as backup.

If Node1 fails, Node2 and Node3 will set a new diag code in the BFD control packet. This diag code will inform Node1 that both Node2 and Node3 didn't fail, and Node1 MUST NOT activate the non-revertive set of services A and B respectively, when it comes back up. The BFD control packet with the new diag code will be sent after the BFD session came up for at least twice the detection multiplier count.

Therefore, Node1 upon receiving the BFD control packet with the new diag code, MUST NOT attempt to activate the non-revertive services, but remain in standby state for the non-revertive services until the Node2 or Node3 that took over fails.

Revertive services are assumed to revert back to the primary node Node1, after the node recovers. Once the BFD session comes up between the primary and backup nodes, the backup node should stop forwarding for any revertive services. A node MUST start forwarding all revertive services for which it is configured as a primary once the BFD session comes up with the corresponding backup nodes. A node MUST stop forwarding for revertive services for which it is a backup once the BFD session comes up with the corresponding primary.

2.2. Per service failover for non-revertive services

An implementation MAY choose to support per service failover for non-revertive services. For example, in figure1, some non-revertive services could be active on Node1 while some non-revertive services could be active on Node2 or Node3 for better load balancing of services traffic. In this mode, every L2/L3 and/or L4/L7 non-revertive service will be identified by a unique ID known across the routing/service nodes providing the services.

A bitmap will be used to represent the non-revertive services, where each non-revertive service is represented by one bit in the bitmap. All the service nodes MUST have the same mapping of the bit position to the non-revertive service unique ID. The bitmap position and the unique service ID could be maintained by a network controller.

A node that is assigned as backup for a given non-revertive service node will take over as active in either of the following cases: 1) The node assigned as primary for this service failed. 2) This specific service failed on the primary node for this service.

In case 1, the BFD session will go down since it is a node failure. In case 2, BFD session between the nodes will remain up. In either scenarios, the node assigned as secondary will become active for the non-revertive service. In case 1, the secondary node will set the new diag code in the BFD control packets once the BFD session is established. The new diag code will be set in the BFD control

packets for at least twice the detection multiplier count. In case 2, this diag code will be set in the next BFD control packets sent after the node takes over as Active for a given non-revertive service. If there is at least one non-revertive service for which this node is not active AND at least 1 non-revertive service for which it is active, the node will also send the bitmap in the BFD control packets payload. The bits identifying the active non-revertive services will be set in this bitmap. The new diag code and the optional bitmap payload will be sent in the BFD control packets for at least twice the detection multiplier count.

Therefore, if a node receives a BFD control packet with the new diag code set but no payload in the BFD control packet, this means that it MUST NOT activate all non-revertive services for which this node is primary. Whereas, if a payload is present in the BFD control packet that has the new diag code set, the receiving node MUST NOT activate the non-revertive services indicated by the set bits in the bitmap.

Per service failover is not applicable to revertive services. They will behave the same way as described in section 2.1

3. Security Considerations

This document does not introduce any additional security constraints.

4. IANA Considerations

IANA is requested to assign a new diag code from the "BFD Diagnostic Codes"

Value	BFD Diagnostic Code Name
0xNN	Out-lived and optional BitMap BFD control packet payload for non-revertive services.

5. Acknowledgments

6. References

6.1. Normative References

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <https://www.rfc-editor.org/info/rfc5880>.

6.2. Informative 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>.

Authors' Addresses

Sami Boutros (editor)
Ciena
USA

Email: sboutros@ciena.com

Ankur Dubey
VMware
USA

Email: adubey@vmware.com

Reshad Rahman
Cisco
USA

Email: rrahman@cisco.com