

BESS WG
Internet-Draft
Intended status: Standards Track
Expires: September 6, 2019

Z. Zhang
Y. Wang
G. Mirsky
ZTE Corporation
March 5, 2019

**Bidirectional Forwarding Detection (BFD) for EVPN Ethernet Segment
Failover Use Case
draft-zwm-bess-es-failover-00.txt**

Abstract

This document introduces a method for fast switchover of Designated Forwarder for Ethernet Segment failover by using Bidirectional Forwarding Detection protocol.

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 September 6, 2019.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

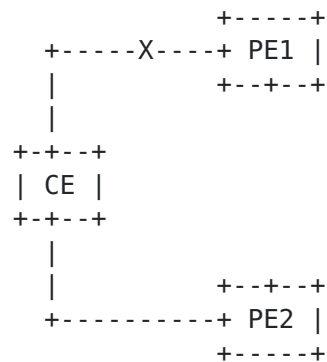
Table of Contents

1.	Introduction	2
2.	Proposal	3
3.	Specification	3
4.	Security Considerations	4
5.	IANA Considerations	4
6.	Normative References	4
	Authors' Addresses	4

[1.](#) Introduction

[RFC7432] introduces Ethernet Virtual Private Network (EVPN) technology. Designated Forwarder (DF) election procedures for multi-homing Ethernet Segments has been described in it. When PE (provider edge) receives BUM (Broadcast, Unknown Unicast and Multicast) flows, only DF forwards the BUM flows to CE (customer edge). Non-DFs do not forward the BUM flows in order to avoid duplication. If the link between DF and CE fails, another PE will forward the BUM flows after it is elected as DF.

[I-D.ietf-bess-evpn-df-election-framework] defines the DF election framework, including that Backup Designated Forwarder (BDF) can be elected as the next best for the role. But before the BDF is elected as DF, the BUM flows are discarded after the link between DF and CE fails.



For example, CE is multihomed to PE1 and PE2. PE1 is elected as DF. All BUM flows are forwarded by PE1 when the link between PE1 and CE is operational. When the link between PE1 and CE fails, the BUM flows are discarded until PE2 is elected as DF.

This document will use terminology defined in [[RFC7432](#)] and [[I-D.jain-bess-evpn-lsp-ping](#)].

2. Proposal

In order to avoid the BUM packet loss on BDF after the link between DF and CE fails, a data-plane detection function is needed for DF fast switchover. [RFC5884] provides mechanisms for using LSP Ping to bootstrap a BFD session. [I-D.jain-bess-evpn-lsp-ping] introduces four new Target FEC Stack sub-TLVs that are included in the LSP-Ping Echo Request packet. This document uses the mechanisms defined in [RFC5884] and the EVPN Ethernet Auto-Discovery (AD) sub-TLV defined in [I-D.jain-bess-evpn-lsp-ping] to provide DF fast switchover by data-plane failure detection.

An LSP-Ping Echo Request message which carries EVPN AD Sub-TLV associated with the DF-CE Ethernet Segment Identifier (ESI) is used to bootstrap the BFD session between BDF and DF. After the BFD session is built, when the ES fault occurs on DF-CE link, BDF detects the fault by the state change BFD control packet sent by DF, or BDF detects the fault when the detection timer expires. Then BDF becomes DF and will forward the BUM flows to CE.

3. Specification

[I-D.jain-bess-evpn-lsp-ping] [section 4.3](#) defines an Ethernet AD sub-TLV as a new Target FEC Stack sub-TLV. It is carried in the LSP-Ping Echo Request message. BDF generates an LSP-Ping Echo Request message which carries the associated ES AD sub-TLV. And BDF sends the message with a local discriminator assigned by BDF for this BFD session to DF. DF responds with the BFD control packet with 'Your discriminator' set to the discriminator value received in the Echo request message from the BDF. BDF can demultiplex the BFD session based on the received 'Your Discriminator' field.

After the BFD session is established, when the link between DF and CE fails, DF MUST send a BFD control packet with the value of State field set to AdminDown through the established BFD session to BDF. If DF is not operational, BDF also detects the failure when the BFD detection time expires. Then BDF becomes DF immediately and forwards the BUM flows to CE.

When the ES between 'old' DF and CE recovers, the BFD session MAY be reused or a new BFD session can be established for the ES failover monitor.

For the same example in last section, PE2 generates an LSP-Ping Echo Request message which carries the associated ES AD sub-TLV and sends the message with an assigned local discriminator to DF. PE1 responds with a BFD control packet with 'Your discriminator' set to the

received discriminator from PE2. PE2 can demultiplex the BFD session based on the received 'Your Discriminator' field.

When the link between PE1 and CE fails, PE1 sends a BFD control packet with the state set to AdminDown to PE2 through the BFD session. If the packet is lost, PE2 also can detect the fault by the session detection time expiration. PE2 becomes DF immediately, then the BUM packets can be forwarded to CE.

4. Security Considerations

This document does not introduce any new security considerations other than already discussed in [[RFC7432](#)] and [[RFC5884](#)].

5. IANA Considerations

There is no IANA consideration.

6. Normative References

- [I-D.ietf-bess-evpn-df-election-framework]
Rabadan, J., satyamoh@cisco.com, s., Sajassi, A., Drake, J., Nagaraj, K., and S. Sathappan, "Framework for EVPN Designated Forwarder Election Extensibility", [draft-ietf-bess-evpn-df-election-framework-09](#) (work in progress), January 2019.
- [I-D.jain-bess-evpn-lsp-ping]
Jain, P., Salam, S., Sajassi, A., Boutros, S., and G. Mirsky, "LSP-Ping Mechanisms for EVPN and PBB-EVPN", [draft-jain-bess-evpn-lsp-ping-08](#) (work in progress), December 2018.
- [RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", [RFC 5884](#), DOI 10.17487/RFC5884, June 2010, <<https://www.rfc-editor.org/info/rfc5884>>.
- [RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", [RFC 7432](#), DOI 10.17487/RFC7432, February 2015, <<https://www.rfc-editor.org/info/rfc7432>>.

Authors' Addresses

Zheng(Sandy) Zhang
ZTE Corporation
No. 50 Software Ave, Yuhuatai Distinct
Nanjing
China

Email: zzhang_ietf@hotmail.com

Yubao Wang
ZTE Corporation
No. 50 Software Ave, Yuhuatai Distinct
Nanjing
China

Email: wang.yubao2@zte.com.cn

Greg Mirsky
ZTE Corporation

Email: gregimirsky@gmail.com