

BESS Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: 4 September 2025

K. Ananthamurthy  
A. Sajassi  
L. Krattiger  
Cisco  
3 March 2025

Per-flow based EVPN DF Election  
draft-kriswamy-bess-evpn-perflow-df-00

## Abstract

The Ethernet Virtual Private Network (EVPN) solution offers procedures for electing a Designated Forwarder (DF) for multihomed Ethernet Segments. In this context, the Provider Edge (PE) router is responsible for sending Broadcast, Unknown Unicast, and Multicast (BUM) traffic to a multihomed device or network. This applies in cases of an all-active multihomed Ethernet Segment (ES) as well as for BUM and unicast traffic in the case of single-active multihoming. While the Default Algorithm provides an efficient and automated way of selecting the DF across different Ethernet Tags in the Ethernet Segment, but lacks in providing per flow based DF within the given EVPN Instances (EVIs).

## Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

## 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 4 September 2025.

## Copyright Notice

Copyright (c) 2025 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

1. Introduction . . . . .	2
2. Requirements Language . . . . .	3
3. Designated Forwarder Election Extended Community Extensions . . . . .	3
4. Operation . . . . .	3
5. Security Considerations . . . . .	4
6. IANA Considerations . . . . .	4
7. Normative References . . . . .	4
Authors' Addresses . . . . .	5

## 1. Introduction

The Ethernet Virtual Private Network (EVPN) solution offers procedures for electing a Designated Forwarder (DF) for multihomed Ethernet Segments. In this context, the Provider Edge (PE) router is responsible for sending Broadcast, Unknown Unicast, and Multicast (BUM) traffic to a multihomed device or network. This applies in cases of an all-active multihomed Ethernet Segment (ES) as well as for BUM and unicast traffic in the case of single-active multihoming. While the Default Algorithm provides an efficient and automated way of selecting the DF across different Ethernet Tags in the Ethernet Segment, but lacks in providing per flow based DF within the given EVPN Instances (EVis).

This document proposes a Per-flow DF Election algorithm to provide optimal load sharing. This document updates RFC8584 by modifying the definition of the DF Election Extended Community with new DF Algorithm type.

## 2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

## 3. Designated Forwarder Election Extended Community Extensions

This solution reuses and extends the Designated Forwarder Election Extended Community defined in [RFC8584] that is advertised along with the Ethernet Segment route. This document defines a new DF Algorithm type Per-flow. The format of the DF Election Extended Community that is used in this document follows:

```

 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=0x06      | Sub-Type(0x06) | RSV | DF Alg |      Bitmap      ~
+-----+-----+-----+-----+-----+-----+-----+-----+
~      Bitmap    |      Reserved      |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Figure 1: DF Election Extended Community

DF Alg will carry the DF alg type.

New DF Alg type called per-flow df with value 6 (to be allocated by IANA) will be defined in the draft and same will be encoded in DF algorithm EC.

The procedures defined in [RFC8584] to handle Designated Forwarder Election Extended Community are applicable.

When a PE receives the ES routes from all the other PEs for the ES in question, it checks to see if all the advertisements have the Extended Community with the same DF Alg, then it will operate in per-flow DF algorithm.

## 4. Operation

When PE supports the Per-flow DF algorithm, the ES route is sent with the DF Election Extended Community, with DF-Alg set to Per-flow.

When PE supports the Per-flow DF algorithm, it should use the DF election as per the flow when it receives the ES route with the DF Election Extended Community and the DF-Alg set to Per-flow.

PE, which does not support Per-flow DF, will revert to the default modulo-based DF.

When a Provider Edge (PE) router operates in Per-Flow Designated Forwarder (DF) mode, it computes a hash based on the five-tuple of the packet header. The hashing algorithm can use a Cyclic Redundancy Check (CRC). The number of PEs involved in redundancy multihoming will be used in the calculation as follows:  $I$ , which is  $\text{HASH mod } N$ . This approach allows multiple flows within a given VLAN to be forwarded by different PEs acting as DF for a specific Ethernet Segment (ES).

## 5. Security Considerations

TBD

## 6. IANA Considerations

This document solicits:

- \* The allocation of two new values in the "DF Alg" registry created by [RFC8584] as follows:

Alg	Name	Reference
----	-----	-----
6	Per-flow Algorithm	This document

## 7. Normative References

- [I-D.ietf-bess-evpn-virtual-eth-segment]  
Sajassi, A., Brissette, P., Schell, R., Drake, J., and J. Rabadan, "EVPN Virtual Ethernet Segment", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-virtual-eth-segment-19, 9 December 2024, <<https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-virtual-eth-segment-19>>.
- [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>>.
- [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>>.

- [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>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8584] Rabadan, J., Ed., Mohanty, S., Ed., Sajassi, A., Drake, J., Nagaraj, K., and S. Sathappan, "Framework for Ethernet VPN Designated Forwarder Election Extensibility", RFC 8584, DOI 10.17487/RFC8584, April 2019, <<https://www.rfc-editor.org/info/rfc8584>>.

## Authors' Addresses

Krishnaswamy Ananthamurthy  
Cisco  
170 W. Tasman Drive  
San Jose, CA 95134  
United States of America  
Email: kriswamy@cisco.com

Ali Sajassi  
Cisco  
170 W. Tasman Drive  
San Jose, CA 95134  
United States of America  
Email: sajassi@cisco.com

Lukas Krattiger  
Cisco  
170 W. Tasman Drive  
San Jose, CA 95134  
United States of America  
Email: lkrattig@cisco.com