

BESS Workgroup
Internet Draft
Intended status: Standards Track

J. Chen

H3C

Expires: May 20, 2018

Nov 27, 2017

EVPN DF Election Algorithm Selection
draft-chen-bess-evpn-df-algorithm-selection-00

Abstract

This document describes a new EVPN Designated Forwarder Election (DF) method which can be used to select a proper DF Election algorithm.

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <https://www.ietf.org/1id-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This Internet-Draft will expire on May 20, 2018.

Copyright Notice

Copyright (c) 2017 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

(http://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. Specification of Requirements 3
3. Solution requirements 3
4. EVPN BGP Attributes for DF Election 3
5. Solution description 4
5.1 Use of the DF Type Priority 4
6. IANA Considerations 5
7. Security Considerations 5
8. Acknowledgments 5
9. References 5
9.1 Normative References 5
9.2 Informative References 6
Authors' Addresses 6

1. Introduction

IN RFC7432, the Designated Forwarder (DF) in (PBB-)EVPN networks is the PE responsible for ccertain actions:

o Sending multicast and broadcast traffic, on a given Ethernet tag on a particular Ethernet segment, to the CE.

o Flooding unknown unicast traffic (i.e., traffic for which a PE does not know the destination MAC address), on a given Ethernet tag on a particular Ethernet segment to the CE, if the environment requires flooding of unknown unicast traffic.

The DF is selected out of a candidate list of PEs at the granularity of <ES, VLAN> or <ES, VLAN bundle> as the default procedure. This is refereed to as "service carving".

Besides 'service-carving', there are also many other efficient and automated way of selecting the DF.

This document proposes an extension to the current [RFC7432](#) DF election procedures to select a proper way for DF election.

2. Specification of Requirements

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 [\[RFC2119\]](#).

3. Solution requirements

This document proposes an extension for 'service-carving' with the following requirements:

- a) An administrative priority option provided for controlling the order of candidate DF Election Algorithms MAY be used.

4. EVPN BGP Attributes for DF Election

This solution reuses and extends the DF Election Extended Community defined in [\[EVPN-HRW-DF\]](#) that is advertised along with the ES route:

```

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(TBD) | DF Type      | Reserved = 0 |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| DF Type Prio  | Reserved = 0 | Reserved = 0           |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

o The definition of DF Type is outside the scope of this document and can have the following values :

- Type 0 - Default, mod based DF election as per [RFC7432](#).
- Type 1 - HRW algorithm as per [\[EVPN-HRW-DF\]](#)
- Type 2 and so on - To be determined

o DF Type Prio or DF Type Priority defines a 1-octect value indicates the priority that PE uses the DF Type. The allowed values are within the range 0-255 inclusive, and default value MUST be 0. This value stands for lowest priority. If PEs advertise different priority for one ES, the minimum will be used.

5. Solution description

Figure 1 gives an example that will be used in the description of the solution.

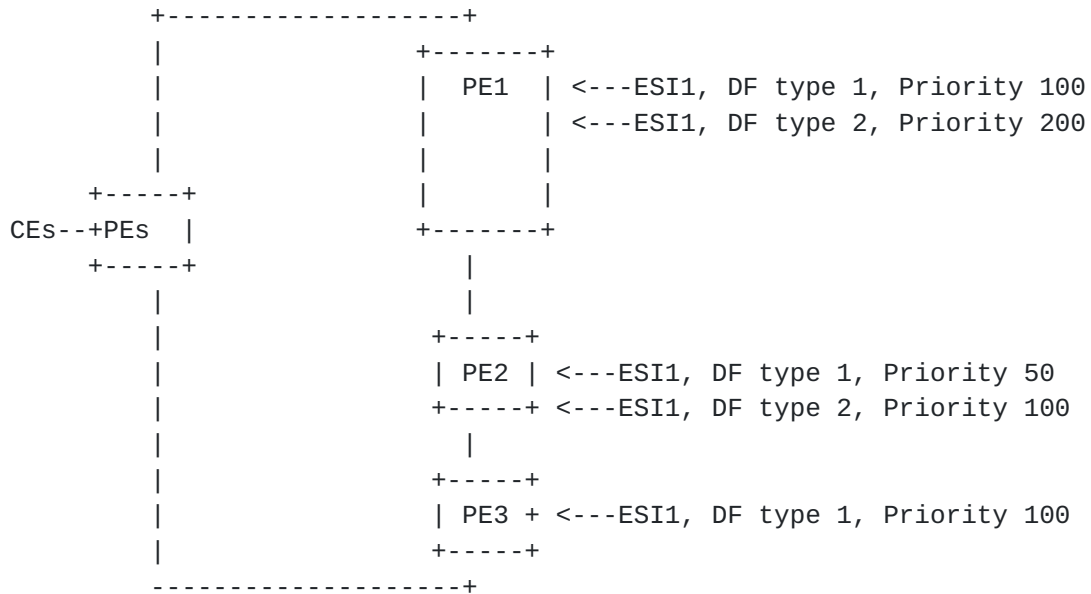


Figure 1 DF Election Method

Figure 1 shows three PEs that are connecting to the same ES in the EVPN network.

If all the PEs use the DF election method described in this document, DF type 1 will become the DF election algorithm. The following sections provide some examples of the new defined procedures and how they are applied in the use-case in Figure 1.

5.1 Use of the DF Type Priority

Assuming the operator wants to control globally - in a flexible way - what DF election algorithm to be used for a given ES . The following procedure may be used:

- a) PEs (or ES1 in PEs) are now configurable with two optional parameters that are signaled in the DF Election extended community. These parameters are the DF algorithm type and Priority. We will represent these parameters as [(DF Type, DF Type Priority), (DF Type, DF Type Priority) ...]. Let's assume PE1 is configured as [(1, 100), (2, 200)], [(1, 50), (2, 100)] in PE2 and [(1, 100)] in PE3.

- b) The PEs will advertise an ES route for each ES (here ESI 1), including the 2 parameters in one or more DF Election Extended Community. Each DF type needs one DF Election Extended Community except the DF type 0 with the priority 0.
- c) According to [RFC7432](#), each PE will wait for the DF timer to expire before running the DF election algorithm. After the timer expires, each PE runs the DF election method as follows:
 - o Each PE builds a list of candidate DF election algorithms that are supported by all PEs, ordered based on the Priority. E.g. PE1 will build a list of candidate DF election algorithms for ESI1 by the Priority, from high to low: (DF type 1, 50), (DF type 0, 0). Hence the DF type 1 will be used as DF election algorithm.
 - o Each PE uses the selected DF election algorithm to elect a proper DF for ESI1.
- d) In case of equal Priority for two or more DF types, the tie-breakers will be the highest DF type number in that order. For instance:
 - o If the DF list is [(2, 50), (1, 50), (0, 0)], DF type 2 will be used.
- e) The DF type and Priority are administrative options that may be configured on a per-ES/per-device basis from the management plane or fixed by device capability. The definition of the actual local policies is out of scope of this document.

6. IANA Considerations

This document solicits the addition of DFT Priority field in the registry created by [[EVPN-HRW-DF](#)].

7. Security Considerations

This section will be added in future versions.

8. Acknowledgments

TBD.

9. References

9.1 Normative References

[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, <<http://www.rfc-editor.org/info/rfc7432>>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.

9.2 Informative References

[EVPN-HRW-DF] Mohanty S. et al. "A new Designated Forwarder Election for the EVPN", [draft-mohanty-bess-evpn-df-election-02](#), work-in-progress, October 19, 2015.

[EVPN-PREF-DF] Rabadan J., Ed. et al. "Preference-based EVPN DF Election", [draft-ietf-bess-evpn-pref-df-00](#), work-in-progress, June 21, 2017.

Authors' Addresses

Jian Chen
H3C
Email: jian_chen@h3c.com

