BESS Workgroup
Internet Draft
Intended status: Standards Track

J. Rabadan, Ed.
S. Sathappan
K. Nagaraj
Nokia

M. Miyake
T. Matsuda
Softbank

Expires: July 26, 2019

January 22, 2019

PBB-EVPN ISID-based CMAC-Flush
draft-snr-bess-pbb-evpn-isid-cmacflush-05

Abstract

Provider Backbone Bridging (PBB) can be combined with Ethernet VPN (EVPN) to deploy ELAN services in very large MPLS networks (PBB-EVPN). Single-Active Multi-homing and per-ISID Load-Balancing can be provided to access devices and aggregation networks. In order to speed up the network convergence in case of failures on Single-Active Multi-Homed Ethernet Segments, PBB-EVPN defines a CMAC-Flush mechanism that works for different Ethernet Segment BMAC address allocation models. This document complements those CMAC-Flush procedures for cases in which no PBB-EVPN Ethernet Segments are defined (ESI 0) and an ISID-based CMAC-Flush granularity is desired.

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 http://www.ietf.org/ietf/1id-abstracts.txt

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

This Internet-Draft will expire on July 26, 2019.

Table of Contents

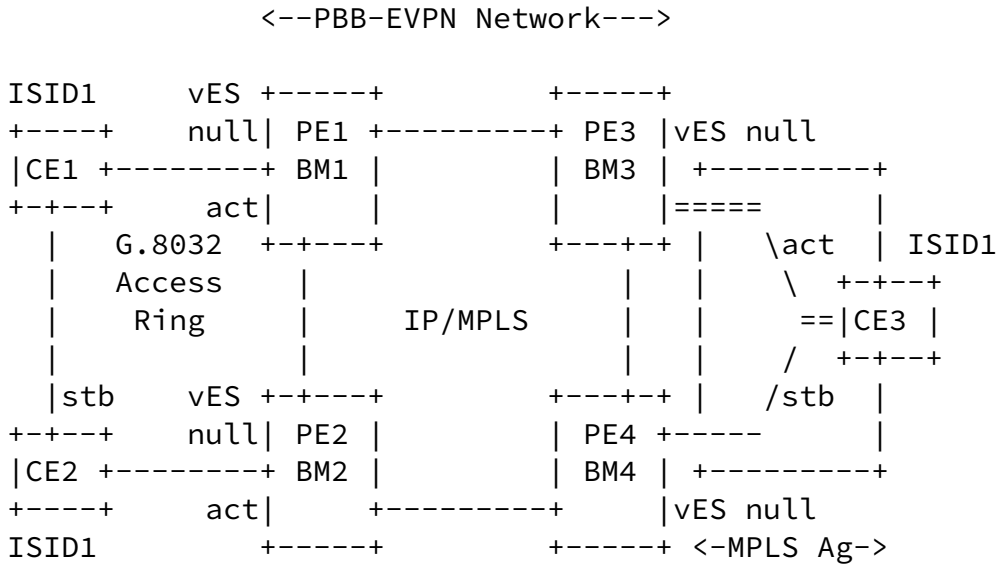
1. Problem Statement 2
2. Solution requirements 4
3. EVPN BGP Encoding for ISID-based CMAC-flush 5
4. Solution description 6
4.1 ISID-based CMAC-Flush activation procedures 7
4.2 CMAC-Flush generation 7
4.3 CMAC-Flush process upon receiving a CMAC-Flush notification 8
5. Conclusions 8
6. Conventions used in this document 9
7. Security Considerations 9
8. IANA Considerations 9
9. References 9
9.1 Normative References 9
9.2 Informative References 9
10. Acknowledgments 9
11. Contributors 9
17. Authors' Addresses 10

1. Problem Statement

[RFC7623] defines how Provider Backbone Bridging (PBB) can be

combined with Ethernet VPN (EVPN) to deploy ELAN services in very large MPLS networks. [RFC7623] also describes how Single-Active Multi-homing and per-ISID Load-Balancing can be provided to access devices and aggregation networks. When Access Ethernet/MPLS Networks exists, [vES] describes how virtual ES can be associated to a group of Ethernet Virtual Circuits (EVCs) or even Pseudowires (PWs). In order to speed up the network convergence in case of failures on Single-Active Multi-Homed Ethernet Segments, [RFC7623] defines a CMAC-Flush mechanism that works for different Ethernet Segment BMAC address allocation models.

In some cases, the administrative entities that manage the access devices or aggregation networks, don't demand Multi-Homing Ethernet Segments (ES) from the PBB-EVPN provider, but simply multiple single-homed ES. If that is the case, the PBB-EVPN network is no longer aware of the redundancy offered by the access administrative entity. Figure 1 shows an example where the PBB-EVPN network provides four different Attachment Circuits (ACs) for ISID1, with those ACs not being part of any ES or vES (therefore they are referred to as null vES).



Network

Figure 1 PBB-EVPN and non-ES based redundancy

In the example in Figure 1, CE1 and CE2 provide redundant connectivity for ISID1 through the use of G.8032 Ethernet Ring Protection Switching. CE3 provides redundant active-standby PW connectivity for ISID1. In the two cases the ACs are connected to null ES, hence the PEs will keep their ACs active and the CEs will be responsible for the per-ISID load balancing while avoiding loops.

For instance, CE2 will block its link to CE1 and CE3 will block its

forwarding path to PE4. In this situation, a failure in one of the redundant ACs will make the CEs to start using their redundant paths, however those failures will not trigger any CMAC-Flush procedures in the PEs. For example, if the active PW from CE3 fails, PE3 will not issue any CMAC-Flush message and therefore the remote PEs will continue pointing at PE3's BMAC to reach CE3's CMACs, until the CMACs age out in the ISID1 FDBs.

[RFC7623] provides a CMAC-Flush solution based on a shared BMAC update along with the MAC Mobility extended community where the sequence number is incremented. However, while that procedure could be used in the example of Figure 1, it would result in unnecessary flushing of unaffected ISIDs on the remote PEs, and subsequent flooding.

This document describes an extension of the [RFC7623] CMAC-Flush procedures, so that in the above failure example, PE3 can trigger a CMAC-Flush notification that makes PE1, PE2 and PE4 flush all the CMACs associated to PE3's BMAC and (only) ISID1. This new CMAC-Flush procedure explained in this document will be referred to as "PBB-EVPN ISID-based CMAC-Flush" and can be used in PBB-EVPN networks with null or non-null (virtual) Ethernet Segments.

2. Solution requirements

The following requirements must be met by the CMAC-Flush solution described in this document:

a) The solution MUST solve black-hole scenarios in case of failures on null ES ACs (Attachment Circuits not associated to ES, that is, ESI=0) when the access device/network is responsible for the redundancy.

b) This extension SHOULD work with Single-Active non-null ES and virtual ES, irrespective of the PE BMAC address assignment (dedicated per-ES BMAC or shared BMAC).

c) In case of failure on the egress PE, the solution MUST provide a CMAC-Flush notification at BMAC AND ISID granularity level.

d) The solution MUST provide a reliable CMAC-Flush notification in PBB-EVPN networks that use Route-Reflectors (RRs).

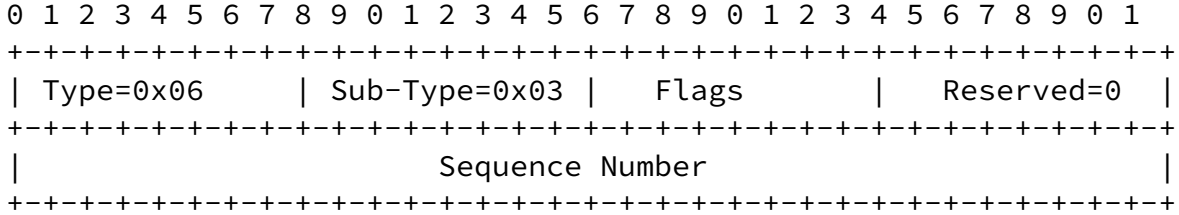
e) The solution MUST coexist in [RFC7623]-compliant networks where there are systems not supporting this extension.

f) The solution SHOULD be enabled/disabled by an administrative

option on a per-PE and per-ISID basis.

3. EVPN BGP Encoding for ISID-based CMAC-flush

The solution does not use any new BGP attributes but reuses the MAC Mobility extended community as an indication of CMAC-Flush (as in [RFC7623]) and encodes the ISID in the Ethernet Tag field of the MAC/IP route. As a reference, Figure 2 shows the MAC Mobility extended community and the MAC/IP route that are used in this document as a CMAC-Flush notification.



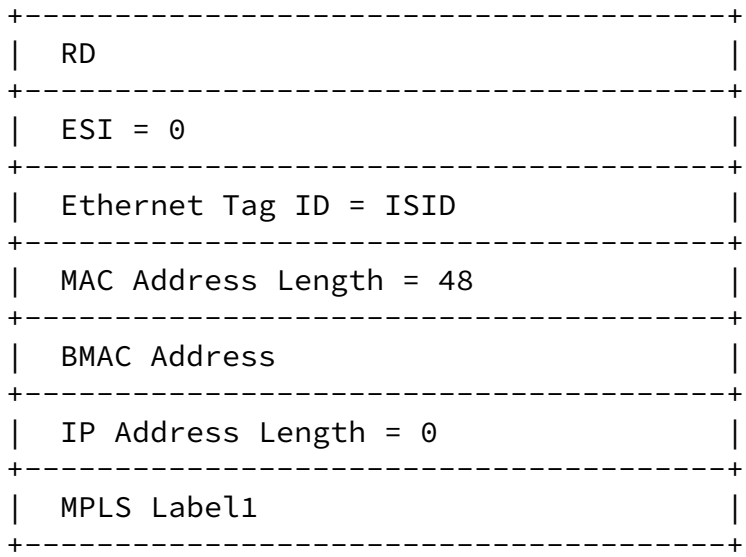


Figure 2 CMAC-Flush notification encoding: BMAC/ISID route

Where:

- The route's RD and RT are the ones corresponding to its EVI. Alternatively to the EVI's RT, the route MAY be tagged with an RT auto-derived from the Ethernet Tag (ISID) instead. [RFC7623] describes how the RT can be derived from the ISID.

- The Ethernet Tag encodes the ISID for which the PE that receives

the route must flush the CMACs upon reception of the route.

- The MAC address field encodes the BMAC Address for which the PE that receives the route must flush the CMACs upon reception of the route.

- The MAC Mobility extended community is used as in [RFC7623], where a delta in the sequence number between two updates for the same BMAC/ISID will be interpreted as a CMAC-flush notification for the corresponding BMAC and ISID.

All the other fields are set and used as defined in [RFC7623]. This document will refer to this route as the BMAC/ISID route, as opposed to the [RFC7623] BMAC/0 route (BMAC route sent with Ethernet Tag = 0).

Note that this BMAC/ISID route will be accepted and reflected by any RFC7432-compliant RR, since no new attributes or values are used. A PE receiving the route will process the received BMAC/ISID update only in case of supporting the procedures described in this document.

4. Solution description

Figure 1 will be used in the description of the solution. CE1, CE2 and CE3 are connected to ACs associated to ISID1, where no (Multi-Homed) Ethernet Segments have been enabled. All the ACs are operationally active and ready to forward frames.

Enabling or disabling ISID-based CMAC-Flush SHOULD be an administrative choice on the system that MAY be configured per ISID (I-Component). When enabled on a PE:

a) The PE will be able to generate BMAC/ISID routes as CMAC-Flush notifications for the remote PEs.

b) The PE will be able to process BMAC/ISID routes received from remote PEs.

When ISID-based CMAC-Flush is disabled, the PE will follow the [RFC7623] procedures for CMAC-flush.

These new CMAC-flush procedures are described in sections 4.1, 4.2 and 4.3 respectively:

o ISID-based CMAC-flush activation
o CMAC-flush notification generation upon AC failures

o CMAC-flush process upon receiving a CMAC-Flush notification

4.1 ISID-based CMAC-Flush activation procedures

The following behavior MUST be followed by the PBB-EVPN PEs (see Figure 1):

- As in [RFC7623], each PE has previously advertised a shared BMAC in a BMAC/0 route (BM1, BM2, BM3 and BM4 respectively). This is the BMAC that each PE will use as BMAC SA (Source Address) when encapsulating the frames received on any local single-homed AC. Each PE will import the received BMAC/0 routes from the remote PEs and will install the BMACs in its B-component MAC-VRF. For instance, PE1 will advertise BM1/0 and will install BM2, BM3 and BM4 in its MAC-VRF.

- Assuming ISID-based CMAC-Flush is activated for ISID 1, the PEs will advertise the shared BMAC with ISID 1 encoded in the Ethernet Tag. That is, PE1 will advertise BM1/1 and will receive BM2/1, BM3/1 and BM4/1. The receiving PEs MUST use these BMAC/ISID routes only for CMAC-Flush procedures and they MUST NOT be used to add/withdraw any BMAC entry in the MAC-VRFs. As per [RFC7623], only BMAC/0 routes can be used to add/withdraw BMACs in the MAC-VRFs.

- The above procedure MAY also be used for dedicated BMACs.

4.2 CMAC-Flush generation

If, for instance, there is a failure on PE1's AC, PE1 will generate an update including BM1/1 along with the MAC Mobility extended community where the Sequence Number has been incremented. The reception of the BM1/1 with a delta in the sequence number will trigger the CMAC-Flush procedures on the receiving PEs.

- An AC going operationally down MUST generate a BMAC/ISID with a higher Sequence Number. If the AC going down makes the entire local ISID go operationally down, the PE will withdraw the BMAC/ISID route for the ISID.

- An AC going operationally up SHOULD NOT generate any BMAC/ISID update, unless it activates its corresponding ISID, in which case the PE will advertise the BMAC/ISID route.

- An AC receiving a CMAC-Flush notification from the access network, e.g. by G.8032, MAY propagate it to the remote PEs by generating a BMAC/ISID update with higher Sequence Number.

4.3 CMAC-Flush process upon receiving a CMAC-Flush notification

A PE receiving a CMAC-Flush notification will follow these procedures:

- A received BMAC/ISID route (with non-zero ISID) MUST NOT add/remove any BMAC to/from the MAC-VRF.

- An update of a previously received BMAC/ISID route with a delta Sequence Number, MUST flush all the CMACs associated to that ISID and BMAC. CMACs associated to the same ISID but different BMAC MUST NOT be flushed.

- A received BMAC/ISID withdraw (with non-zero ISID) MUST flush all the CMACs associated to that BMAC and ISID.

Note that the CMAC-Flush procedures described in [RFC7623] for BMAC/0 routes are still valid and a PE receiving [RFC7623] CMAC-flush notification messages MUST observe the behavior specified in [RFC7623].

5. Conclusions

The ISID-based CMAC-Flush solution described in this document has the following benefits:

a) The solution solves black-hole scenarios in case of failures on null ES ACs, since the CMAC-flush procedures are independent of the Ethernet Segment definition.

b) This extension can also be used with Single-Active non-null ES and virtual ES, irrespective of the PE BMAC address assignment (dedicated per-ES BMAC or shared BMAC).

c) It provides a CMAC-Flush notification at BMAC AND ISID granularity level, therefore flushing a minimum number of CMACs and reducing the amount of flooding in the network.

d) It provides a reliable CMAC-Flush notification in PBB-EVPN networks that use RRs. RRs will propagate the CMAC-flush notifications for all the affected ISIDs and irrespective of the order in which the notifications make it to the RR.

e) The solution can coexist in a network with systems supporting or not supporting the CMAC-flush extensions.

6. Conventions used in this document

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 BCP14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

7. Security Considerations

This section will be added in future versions.

8. IANA Considerations

9. References

9.1 Normative References

[RFC7623] Sajassi, A., Ed., Salam, S., Bitar, N., Isaac, A., and W. Henderickx, "Provider Backbone Bridging Combined with Ethernet VPN (PBB-EVPN)", RFC 7623, DOI 10.17487/RFC7623, September 2015, <https://www.rfc-editor.org/info/rfc7623>.

[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>.

9.2 Informative References

[vES] Sajassi et al. "EVPN Virtual Ethernet Segment", draft-ietf-bess-evpn-virtual-eth-segment-04, work-in-progress, January, 2019.

10. Acknowledgments

The authors want to thank Vinod Prabhu, Sriram Venkateswaran, Laxmi Padakanti, Ranganathan Boovaraghavan for their review and contributions.

11. Contributors

17. Authors' Addresses

Jorge Rabadan
Nokia
777 E. Middlefield Road
Mountain View, CA 94043 USA
Email: jorge.rabadan@nokia.com

Senthil Sathappan
Nokia
701 E. Middlefield Road
Mountain View, CA 94043 USA
Email: senthil.sathappan@nokia.com

Kiran Nagaraj
Nokia
701 E. Middlefield Road
Mountain View, CA 94043 USA
Email: kiran.nagaraj@nokia.com

Masahiro Miyake
Softbank
Email: masahiro.miyake@g.softbank.co.jp

Taku Matsuda
Softbank
Email: taku.matsuda@g.softbank.co.jp