BESS Workgroup
Internet-Draft
Intended status: Standards Track
Expires: May 22, 2022

J. Rabadan, Ed.
S. Sathappan
K. Nagaraj
Nokia
M. Miyake
T. Matsuda
Softbank

November 18, 2021

PBB-EVPN ISID-based CMAC-Flush
draft-ietf-bess-pbb-evpn-isid-cmacflush-04

Abstract

Provider Backbone Bridging (PBB) can be combined with Ethernet VPN (EVPN) to deploy Ethernet Local Area Network (ELAN) services in large Multi-Protocol Label Switching (MPLS) networks (PBB-EVPN). Single-Active Multi-homing and per-I-SID (per Service Instance Identifier) 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 flush mechanism for Customer MACs (CMAC-flush) that works for different Ethernet Segment Backbone MAC (BMAC) address allocation models. This document complements those CMAC-flush procedures for cases in which no PBB-EVPN Ethernet Segments are defined (the attachment circuit is associated to a zero Ethernet Segment Identifier) and a Service Instance Identifier based (I-SID-based) CMAC-flush granularity is required.

Table of Contents

1. Introduction

[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-I-SID Load-Balancing can be provided to access devices and aggregation networks. When Access Ethernet/MPLS Networks exists, [I-D.ietf-bess-evpn-virtual-eth-segment] describes how virtual Ethernet Segments 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 do not 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 I-SID1, with those ACs not being part of any ES or vES (therefore they are referred to as null vES).

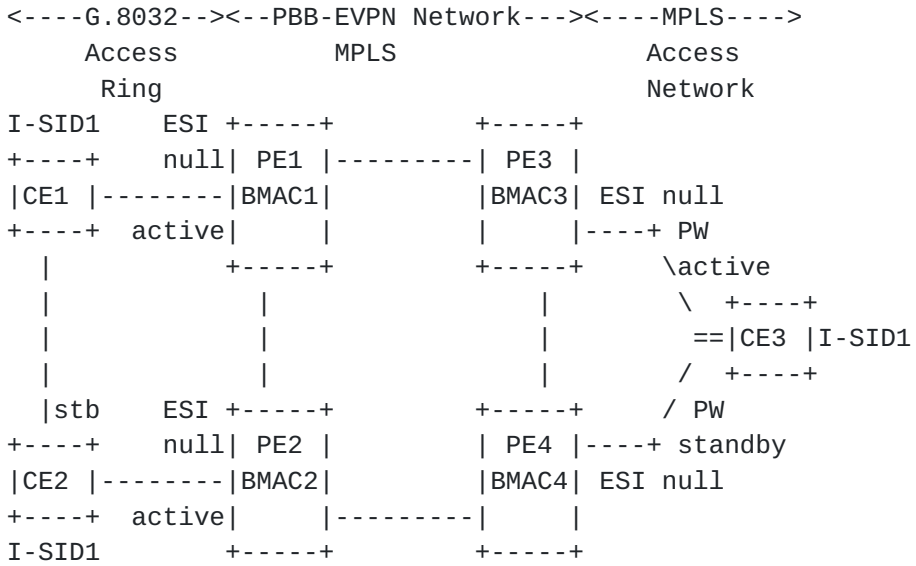

Figure 1: PBB-EVPN and non-ES based redundancy

In the example in Figure 1, CE1, CE2 and CE3 are attached to the same service, identified by I-SID1 in the PBB-EVPN PEs. CE1 and CE2 are connected to the PEs via G.8032 Ethernet Ring Protection Switching, and their ACs to PE1 and PE2 are represented by a port and VLAN identifier. CE3 is dual-homed to PE3 and PE4 through an active-standby PW, and its AC to the PEs is represented by a PW. Each of the four PEs uses a dedicated BMAC address as source MAC address (BMAC1, BMAC2, BMAC3 and BMAC4, respectively) when encapsulating customer frames in PBB packets and forwarding those PBB packets to the remote PEs as per [RFC7623]. There are no multi-homed Ethernet Segments defined in the PBB-EVPN network of the example, that is why the four ACs in Figure 1 show the text "ES null", which means the Ethernet Segment Identifier on those ACs is zero. Since there are no multi-homed ES defined, the PEs keep their ACs active as long as the physical connectivity is established and the CEs are responsible for

managing the redundancy, avoiding loops and providing per-I-SID load balancing to the PBB-EVPN network.

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 trigger the CEs to start using their redundant paths, however those failures will not trigger any CMAC-flush procedures in the PEs that implement [RFC7623], since the PEs are not using the PBB-EVPN multi-homing procedures. For example, if the active PW from CE3 (to PE3) 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 I-SID1 forwarding tables.

[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, the procedure is only used along with multi-homed Ethernet Segments. Even if that procedure could be used for null Ethernet Segments, as in the example of Figure 1, the [RFC7623] CMAC-flush procedure would result in unnecessary flushing of unaffected I-SIDs on the remote PEs, and subsequent flooding of unknown unicast traffic in the network.

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 BMAC3 and (only) I-SID1. This new CMAC-flush procedure explained in this document will be referred to as "PBB-EVPN I-SID-based CMAC-flush" and can be used in PBB-EVPN networks with null or non-null (virtual) Ethernet Segments.

1.1. Terminology and Conventions

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.

AC: Attachment Circuit.

B-Component: Backbone Component, as in [RFC7623].

BMAC: Backbone MAC address.

BMAC/0 route: an EVPN MAC/IP Advertisement route that uses a BMAC in the MAC address field and a zero Ethernet Tag ID.

BMAC/I-SID route: an EVPN MAC/IP Advertisement route that uses a BMAC in the MAC address field and an I-SID in the Ethernet Tag field, and it is used to notify remote PEs about the required CMAC-flush procedure for the CMACs associated with the advertised BMAC and I-SID.

CE: Customer Edge router.

CMAC: Customer MAC address.

ES and ESI: Ethernet Segment and Ethernet Segment Identifier.

EVI: EVPN Instance.

EVPN: Ethernet Virtual Private Networks, as in [RFC7432].

G.8032: Ethernet Ring Protection.

I-Component: Service Instance Component, as in [RFC7623].

I-SID: Service Instance Identifier.

MAC-VRF: A Virtual Routing and Forwarding table for MAC addresses.

PBB-EVPN: Provider-Backbone-Bridging and EVPN, as in [RFC7623].

PE: Provider Edge router.

RD: Route Distinguisher.

RT: Route Target.

Familiarity with the terminology in [RFC7623] is expected.

2. Solution requirements

The following requirements are followed by the CMAC-flush solution described in this document:

a. The solution solves 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 works with Single-Active non-null ES and virtual ES, irrespective of the PE BMAC address assignment (dedicated per-ES BMAC or shared BMAC, as in [RFC7623]).

c. In case of failure on the egress PE, the solution provides a CMAC-flush notification at BMAC and I-SID granularity level.

d. The solution provides a reliable CMAC-flush notification in PBB-EVPN networks that use Route-Reflectors (RRs), without causing "double flushing" or no flushing for certain I-SIDs due to the notification messages being aggregated at the RR.

e. The solution coexists in [RFC7623] networks where there are PEs that do not support this specification.

f. The solution SHOULD be enabled/disabled by an administrative option on a per-PE and per-I-SID 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 I-SID in the Ethernet Tag field of the EVPN MAC/IP advertisement route. As a reference, Figure 2 shows the MAC Mobility extended community and the EVPN MAC/IP advertisement route that are used specified in [RFC7432] and used in this document as a CMAC-flush notification message.

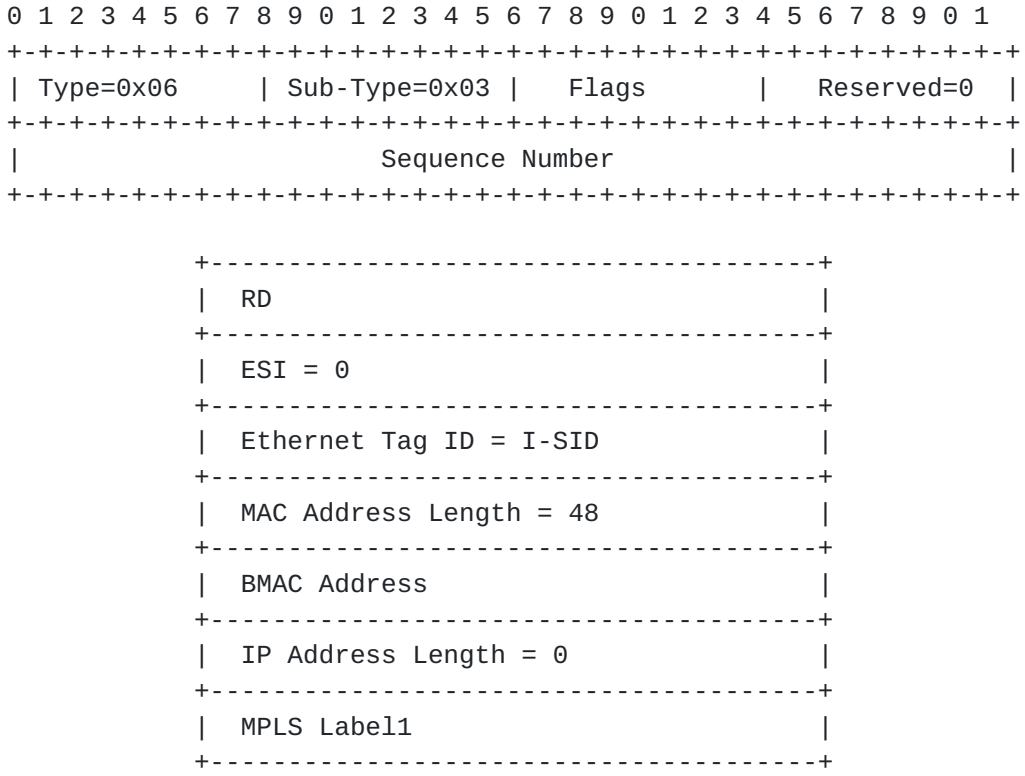

Figure 2: CMAC-Flush notification encoding: BMAC/ISID route

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 (I-SID) instead. [RFC7623] describes how the EVPN MAC/IP Advertisement routes can be advertised along with the EVI RT or an RT that is derived from the I-SID.

- The Ethernet Tag encodes the I-SID 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/I-SID will be interpreted as a CMAC-flush notification for the corresponding BMAC and I-SID.

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

Note that this BMAC/I-SID route will be accepted and reflected by any [RFC7432] RR, since no new attributes or values are used. A PE receiving the route will process the received BMAC/I-SID 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 I-SID1, where no (Multi-Homed) Ethernet Segments have been enabled, and the ACs and PWs are in active or standby state as per Figure 1.

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

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

b. The PE will be able to process BMAC/I-SID routes received from remote PEs.

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

This CMAC-flush specification is described in three sets of procedures:

- I-SID-based CMAC-flush activation
- CMAC-flush notification generation upon AC failures
- 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 following this specification. Figure 1 is used as a reference.

- As in [RFC7623], each PE advertises a shared BMAC in a BMAC/0 route (with BMAC1, BMAC2, BMAC3 and BMAC4 in the MAC address field, 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 BMAC1/0 and will install BMAC2, BMAC3 and BMAC4 in its MAC-VRF.

- Assuming I-SID-based CMAC-flush is activated for I-SID 1, the PEs will advertise the shared BMAC with I-SID 1 encoded in the Ethernet Tag. That is, PE1 will advertise BMAC1/1 and will receive BMAC2/1, BMAC3/1 and BMAC4/1. The receiving PEs MUST use these BMAC/I-SID routes only for CMAC-flush procedures and they MUST NOT be used them 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 (BMACs allocated per Ethernet Segment).

4.2. CMAC-Flush generation

If, for instance, there is a failure on PE1's AC, PE1 will generate an update including BMAC1/1 along with the MAC Mobility extended community where the Sequence Number has been incremented. The reception of the BMAC1/1 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/I-SID with a higher Sequence Number. If the AC going down makes the entire

local I-SID go operationally down, the PE will withdraw the BMAC/I-SID route for the I-SID.

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

- An AC receiving a G.8032 flush notification or a flush message in any other protocol from the access network MAY propagate it to the remote PEs by generating a BMAC/I-SID route 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/I-SID route (with non-zero I-SID) MUST NOT add/remove any BMAC to/from the MAC-VRF.

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

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

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 I-SID-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 I-SID granularity level, therefore flushing a minimum number of CMACs

and reducing the amount of unknown unicast 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 I-SIDs 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 this specification.

6. Security Considerations

Security considerations described in [RFC7623] apply to this document.

In addition, this document suggests additional procedures, that can be activated on a per I-SID basis, and generate additional EVPN MAC/IP Advertisement routes in the network. The format of these additional EVPN MAC/IP Advertisement routes is backwards compatible with [RFC7623] procedures and should not create any issues on receiving PEs not following this specification, however, the additional routes may consume extra memory and processing resources on the receiving PEs. Because of that, it is RECOMMENDED to activate this feature only when necessary (when multi-homed networks or devices are attached to the PBB-EVPN PEs), and not by default in any PBB-EVPN PE.

7. IANA Considerations

8. Acknowledgments

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

9. Contributors

10. References

10.1. 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>.

[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>.

[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>.

10.2. Informative References

[I-D.ietf-bess-evpn-virtual-eth-segment]
Sajassi, A., Brissette, P., Schell, R., Drake, J. E., and J. Rabadan, "EVPN Virtual Ethernet Segment", draft-ietf-bess-evpn-virtual-eth-segment-07 (work in progress), July 2021.

Authors' Addresses

Jorge Rabadan (editor)
Nokia
777 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

M. Miyake
Softbank

Email: masahiro.miyake@g.softbank.co.jp

T. Matsuda
Softbank

Email: taku.matsuda@g.softbank.co.jp