L2VPN Working Group
Internet Draft
Intended status: Standards Track
Expires: January 2014

Dave Allan, Jeff Tantsura
Ericsson
Don Fedyk
Alcatel-Lucent
Ali Sajassi
Cisco

July 2013

802.1aq Support over EVPN
draft-allan-l2vpn-spbm-evpn-04

Abstract

This document describes how Ethernet Shortest Path Bridging MAC mode (802.1aq) can be combined with EVPN in a way that interworks with PBB-PEs as described in the PBB-EVPN solution in a way that permits operational isolation of each Ethernet network subtending an EVPN core while supporting full interworking between the different variations of Ethernet operation.

Status of this Memo

This Internet-Draft is submitted to IETF 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 January 2014.

Copyright and License Notice

Table of Contents

1. Introduction

This document describes how Ethernet Shortest Path Bridging MAC mode (802.1aq) along with PBB-PEs and PBBNs (802.1ah) can be supported by EVPN such that each island is operationally isolated while providing full L2 connectivity between them. Each island can use its own control plane instance and multi-pathing design, be it multiple ECT sets, or multiple spanning trees.

The intention is to permit both past, current and emerging future versions of Ethernet to be seamlessly integrated to permit large scale, geographically diverse numbers of Ethernet end systems to be fully supported with EVPN as the unifying agent.

1.1. Authors

David Allan, Jeff Tantsura, Don Fedyk, Ali Sajassi

1.2. Requirements Language

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 [1].

2. Conventions used in this document

2.1. Terminology

BCB: Backbone Core Bridge
BEB: Backbone Edge Bridge
BU: Broadcast/Unknown
B-MAC: Backbone MAC Address
B-VID: Backbone VLAN ID
CE: Customer Edge
C-MAC: Customer/Client MAC Address
DF: Designated Forwarder
ESI: Ethernet Segment Identifier
EVPN: Ethernet VPN
ISIS-SPB: IS-IS as extended for SPB
I-SID: I-Component Service ID
MP2MP: Multipoint to Multipoint

MVPN: Multicast VPN
NLRI: Network Layer Reachability Information
PBBN: Provider Backbone Bridged Network
PBB-PE: Co located BEB and PE
PE: provider edge
P2MP: Point to Multipoint
P2P: Point to Point
RD: Route Distinguisher
SPB: Shortest path bridging
SPBM: Shortest path bridging MAC mode

3. Changes since previous version

1) Removal of reference to 802.1Qbp. This will be addressed in separate document.

2) Determining ESI value exclusively requires configuration. This was an open item in previous drafts.

4. Solution Overview

The EVPN solution for 802.1aq SPBM incorporates control plane interworking in the PE to map ISIS-SPB [2] information elements into the EVPN NLRI information and vice versa. This requires each PE to act both as an EVPN BGP speaker and as an ISIS-SPB edge node. Associated with this are procedures for configuring the forwarding operations of the PE such that an arbitrary number of EVPN subtending SPBM islands may be interconnected without any topological or multipathing dependencies. This model also permits PBB-PEs as defined in draft-l2vpn-pbb-evpn-02[8] to seamlessly communicate with the SPB islands.

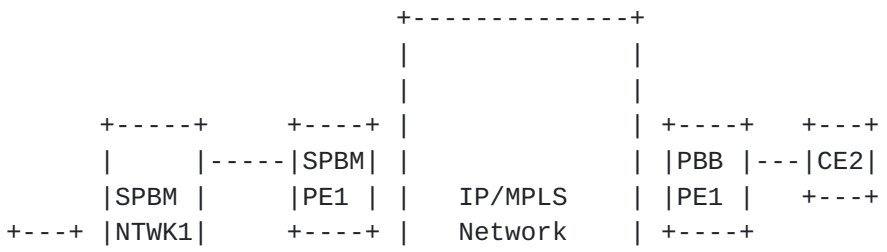

Figure 1: PBB and SPBM EVPN Network

Each EVPN is identified by a route target. The route target identifies the set of SPBM islands and BEB-PEs that are allowed to communicate. Each SPBM island is administered to have an associated Ethernet Segment ID (ESI) associated with it. This manifests itself as a set of Ethernet segments, where each Ethernet segment ID is unique within the route target.

BGP acts as a common repository of the I-SID attachment points for the set of subtending PEs/SPBM islands. This is in the form of B-MAC address/I-SID/Tx-Rx-attribute tuples. BGP leaks I-SID information into each SPBM island on the basis of locally registered interest. If an SPBM island has no BEBs registering interest in an I-SID, information about that I-SID from other SPBM islands, PBB-PEs or PBBNs will not be leaked into the local ISIS-SPB routing system.

For each B-VID in an SPBM island, a single SPBM-PE is elected the designated forwarder for the B-VID. An SPBM-PE may be a DF for more than one B-VID. This is described further in section 4.2. The SPBM-PE originates IS-IS advertisements as if it were an I-BEB or IB-BEB that proxy for the other SPBM islands and PBB PEs in the EVPN defined by the route target, but the PE typically will not actually host any I-components.

An SPBM-PE that is a DF for a B-VID strips the B-VID tag information from frames relayed towards the EVPN. The DF also inserts the appropriate B-VID tag information into frames relayed towards the SPBM island on the basis of the local I-SID/B-VID bindings advertised in ISIS-SPB.

5. Elements of Procedure

5.1. PE Configuration

At SPBM island commissioning a PE is configured with:

1) The route target for the service instance. Where a route target is defined as identifying the set of SPBM islands, PBBNs and PBB-PEs to be interconnected by the EVPN.

2) The unique ESI for the SPBM island.

And the following is configured as part of commissioning an ISIS-SPB node:

1) A Shortest Path Source ID (SPSourceID) used for algorithmic construction of multicast DA addresses. Note this is required for SPBM BEBs independent of the EVPN operation.

2) The set of VLANs (identified by B-VIDs) used in the SPBM island and multi-pathing algorithm IDs to use. The B-VID may be different in different domains and may be removed as carried over the IP/MPLS network.

A type-1 Route Distinguisher (RD) for the node can be auto-derived. This will be described in a future version of the document.

5.2. DF Election

PEs self appoint in the role of DF for a B-VID for a given SPBM island. The procedure used is as per section 9.5 of draft-ietf-l2vpn-evpn-03[4] "DF election".

5.3. Control plane interworking ISIS-SPB to EVPN

When a PE receives an SPBM service identifier and unicast address sub-TLV as part of an ISIS-SPB MT capability TLV it checks if it is the DF for the B-VID in the sub-TLV.

If it is the DF, and there is new or changed information then a MAC advertisement route NLRI is created for each new I-SID in the sub-TLV.

- the Route Distinguisher (RD) is set to that of the PE.

- the ESI is that of the SPBM island.

- the Ethernet tag ID contains the I-SID (including the Tx/Rx attributes). The encoding of I-SID information is as per figure 2.

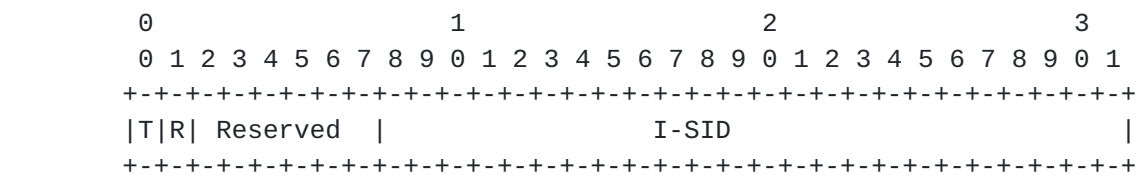

Figure 2: I-SID encoding in the Ethernet tag-ID field

- the MAC address from the sub-TLV

- an MPLS label

Similarly in the scenario where a PE became elected DF for a B-VID in an operating network, the IS-IS database would be processed in order to construct the NLRI information associated with the new role of the PE.

If the BGP database has NLRI information for the I-SID, and this is the first instance of registration of interest in the I-SID from the SPB island, the NLRI information with that tag is processed to construct an updated set of SPBM service identifier and unicast address sub-TLVs to be advertised by the PE.

The ISIS-SPB information is also used to keep a local table indexed by I-SID current to indicate the associated B-VID for processing of frames received from EVPN. When an I-SID is associated with more than one B-VID, only one entry is allowed in the table. Rules for this will be in a future version of the document.

5.4. Control plane interworking EVPN to ISIS-SPB

When a PE receives a BGP NLRI that has new information, it checks if the I-SID in the Ethernet Tag ID locally maps to the B-VID that are an elected DF. Note that if no BEBs in the SPB island have advertised any interest in the I-SID, it will not be associated with any B-VID locally, and therefore not of interest. If the I-SID is of local interest to the SPBM island and the PE is the DF for the B-VID that that I-SID is locally mapped to, a SPBM service identifier and unicast address sub-TLV is constructed/updated for advertisement into ISIS-SPB.

The NLRI information advertised into ISIS-SPB is also used to locally populate a forwarding table indexed by B-MAC+I-SID that points to the

label stack associated with the SPBM frame. The bottom label in the stack being that offered in the NLRI.

5.5. Data plane Interworking 802.1aq SPBM island or PBB-PE to EVPN

When an PE receives a frame from the SPBM island in a B-VID for which it is a DF, it looks up the B-MAC/I-SID information to determine the label stack to be added to the frame for forwarding in the EVPN. The PE strips the B-VID information from the frame, adds the label information to the frame and forwards the resulting MPLS packet.

5.6. Data plane Interworking EVPN to 802.1aq SPBM island

When a PE receives a packet from the EVPN it may infer the B-VID to overwrite in the SPBM frame from the I-SID or by other means (such as via the bottom label in the MPLS stack).

If the frame has a local multicast DA, it overwrites the SPsourceID in the frame with the local SPsourceID.

5.7. Data plane interworking EVPN to 802.1ah PBB-PE

A PBB-PE actually has no subtending PBBN nor concept of B-VID so no frame processing is required.

A PBB-PE is required to accept SPBM encoded multicast DAs as if they were 802.1ah encoded multicast DAs. The only information of interest being that it is a multicast frame, and the I-SID encoded in the lower 24 bits.

5.8. Multicast Support

Refer to "mLDP extensions for integrating EVPN and multicast"[5].

6. Other Aspects

6.1. Flow Ordering

When per I-SID multicast is implemented via PE replication, a stable network will preserve frame ordering between known unicast and BU traffic (e.g. race conditions will not exist). This cannot be guaranteed when multicast is used in the EVPN.

6.2. Transit

Any PE that does not need to participate in the tandem calculations may use the IS-IS overload bit to exclude SPBM tandem paths and behave as pure interworking platform (I-BEB).

7. Acknowledgements

The authors would like to thank Peter Ashwood-Smith, Martin Julien and Janos Farkas for their detailed review of this draft.

8. Security Considerations

For a future version of this document.

9. IANA Considerations

For a future version of this document.

10. References

10.1. Normative References

[1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[2] Fedyk et.al. "IS-IS Extensions Supporting IEEE 802.1aq Shortest Path Bridging", IETF RFC 6329, April 2012

[3] Rosen et.al., "BGP/MPLS IP Virtual Private Networks (VPNs)", IETF RFC 4364, February 2006

[4] Aggarwal et.al. "BGP MPLS Based Ethernet VPN", IETF work in progress, draft-ietf-l2vpn-evpn-02, October 2012

[5] Allan et.al. "mLDP extensions for integrating EVPN and multicast", IETF work in progress draft-allan-l2vpn-mldp-evpn-01, May 2013

10.2. Informative References

[6] 802.1aq(2012) IEEE Standard for Local and Metropolitan Area Networks: Bridges and Virtual Bridged Local Area Networks - Amendment 9: Shortest Path Bridging

[7] Sajassi et.al. "PBB E-VPN", IETF work in progress, draft-ietf-l2vpn-pbb-evpn-04, February 2013

[8] 802.1Q (2011) IEEE Standard for Local and metropolitan area networks--Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks

11. Authors' Addresses

Dave Allan (editor)
Ericsson
300 Holger Way
San Jose, CA 95134
USA
Email: david.i.allan@ericsson.com

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, CA 95134
Email: jeff.tantsura@ericsson.com

Don Fedyk
Alcatel-Lucent
Groton, MA 01450
USA
EMail: Donald.Fedyk@alcatel-lucent.com

Ali Sajassi
Cisco
170 West Tasman Drive
San Jose, CA 95134, US
Email: sajassi@cisco.com