EVPN Multicast Synchronization Enhancement
draft-liu-bess-evpn-mcast-synch-enhance-01

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 June 10, 2019.

Copyright Notice

Abstract

Ethernet Virtual Private Network (EVPN, RFC7432) solutions is becoming prevalent in Data Centers, Data Center Interconnect (DCI) and Service Provider VPN applications. In the scenario of multi-homing from a CE to multiple PEs with links used in a single-active or all-active redundancy mode, IGMP & MLD Join/Leave and PIM Join/RPT-Prune synchronization procedures have been provided. This document specifies an enhancement for IGMP/PIM signaling synchronization when using different Ethernet tag encapsulation mode.

Table of Contents

1. Introduction

Ethernet Virtual Private Network (EVPN [RFC7432]) solutions are becoming prevalent in Data Centers, Data Center Interconnect (DCI) and Service Provider VPN applications. In the scenario of multi-homing from a CE to multiple PEs with links used in a single-active or all-active redundancy mode, [EVPN-IGMP-MLD-PROXY] for IGMP&MLD Join/Leave and [EVPN-PIM-PROXY] for PIM Join/RPT-Prune synchronization procedures have been provided.

However, there may be multiple ways for a CE to connect to PE, including Ethernet tag termination and Q-in-Q termination. When IGMP/PIM synch routes are transmitted from one PE to another PE both multi-homing to the same CE, the appropriate VLAN tags should be included in the routes. This document specifies an improvement in IGMP/PIM synchronization process.

1.1. 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.

1.2. Terminology

CE - Customer Edge equipment

EVPN: Ethernet Virtual Private Network

Ethernet Segment (ES): When a customer site (device or network) is
connected to one or more PEs via a set of Ethernet links, then that
set of links is referred to as an 'Ethernet segment'.

Ethernet Segment Identifier (ESI): A unique non-zero identifier that
identifies an Ethernet segment is called an 'Ethernet Segment
Identifier'.

IGMP: Internet Group Management Protocol

MLD: Multicast Listener Discovery

PE: Provider Edge device

PIM: Protocol Independent Multicast

2. Enhancement of IGMP/PIM Synchronization

Consider the situation where one CE is dual-homed to two PEs. If the
CE is connected through one VLAN and only one VLAN under the EVPN
instance at the PE, the IGMP/PIM synch message can be directly sent
from one PE to another PE with the ESI, which can represent the
appropriate interface and can be encapsulated with the appropriate
VLAN ID.

However, there may be Ethernet tag termination and Q-in-Q
termination on the PE interface connected to CE. When IGMP/PIM
synchronization procedures are implemented between the two dual-
homed PEs, the receiver PE cannot obtain the appropriate VLANs
because of the information lacking in the IGMP/PIM synch route
message. For example, PE1 and PE2 connect to the CE dual-homed for
the same ES, and VLAN tag 1 to 10 termination is configured on the
interface for CE connecting to PE1 and PE2. After receiving a

IGMP/PIM join with VLAN tag 2 from the interface, PE1 sends a
IGMP/PIM join synch route to PE2 with the ESI. But PE2 cannot get
the VLAN tag information from the route, so that PE doesn't create
the IGMP Join state accurately. It can only know the VLAN tag 1-10
termination on the corresponding interface according to the [ES,BD].
Especially Q-in-Q termination is configured on the interface for CE
connecting to PE1 and PE2, outer and inner VLAN tag information will
be both needed when IGMP/PIM synch routes are advertised between the
dual-homed PE1 and PE2.

This document extends the IGMP/PIM synch routes to solve the problem
mentioned above. VLAN tag information from the original IGMP/PIM
message is added in the synch routes NLRI as a part of the route
key.

3. **BGP Encoding Improvement**

This document extends the following routes defined in [EVPN-IGMP-
MLD-PROXY] and [EVPN-PIM-PROXY]

+ Type 6 - IGMP/PIM Join Synch Route

+ Type 7 - IGMP Leave Synch Route

+ Type TBD - PIM RPT-Prune Synch Route

  3.1. IGMP/PIM Join Synch Route

This document extends the IGMP/PIM Join Synch Route defined in
[EVPN-PIM-PROXY] with new fields and Flags as shown in Figure 1.

```
          +----------------------------------------------+
          |  RD (8 octets)                               |
          +----------------------------------------------+
          |  Ethernet Segment Identifier (10 octets)     |
          +----------------------------------------------+
          |  Ethernet Tag ID  (4 octets)                 |
          +----------------------------------------------+
          |  Multicast Source Length (1 octet)           |
          +----------------------------------------------+
          |  Multicast Source Address (variable)         |
          +----------------------------------------------+
          |  Multicast Group Length (1 octet)            |
          +----------------------------------------------+
          |  Multicast Group Address (Variable)          |
          +----------------------------------------------+
          |  Originator Router Length (1 octet)          |
```

```
           +----------------------------------------------+
           |  Originator Router Address (variable)        |
           +----------------------------------------------+
           |  Flags (1 octet)                             |
           +----------------------------------------------+
           |  Upstream Router Length (1B)(optional)       |
           +----------------------------------------------+
           |  Upstream Router Addr (variable)(opt)        |
           +----------------------------------------------+
           |  Outer VLAN (2 octets)                       |
           +----------------------------------------------+
           |  Inner VLAN (2 octets)                       |
           +----------------------------------------------+


           Flags:

           0  1  2  3  4  5  6  7
           +--+--+--+--+--+--+--+--+
           |  |  | V| P|IE|v3|v2|v1|
           +--+--+--+--+--+--+--+--+
```

        Figure 1 IGMP/PIM Join Synch Route

   This route will be added the following extra fields:

   Outer VLAN Can be used for single layer VLAN tag and the outer layer
   of double layer VLAN tags, and can be 0. When 0, it means there is
   no VLAN tag information needed to be advertised.

   Inner VLAN Can be used for the inner layer of double layer VLAN
   tags, and can be 0. When 0, it means there is no inner VLAN.

   Flags: This field encodes Flags that are already relevant to IGMP
   and PIM. The following new Flag is defined:

   - Flag V: Indicates the Join Synch route carries VLAN tag
     information by the encapsulation of the received protocol packet.
     When V=1, the Outer VLAN and Inner VLAN fields are present in the
     route. Otherwise the two fields will not be present.

   Compared to [EVPN-PIM-PROXY] there are two additional fields
   considered part of the route key for BGP processing when Flag V is
   set 1.

    3.2. IGMP Leave Synch Route

   This document extends the IGMP Leave Synch Route defined in [EVPN-

IGMP-MLD-PROXY] as shown in Figure 2

```
               +----------------------------------------------+
               |  RD (8 octets)                               |
               +----------------------------------------------+
               |  Ethernet Segment Identifier (10 octets)     |
               +----------------------------------------------+
               |  Ethernet Tag ID  (4 octets)                 |
               +----------------------------------------------+
               |  Multicast Source Length (1 octet)           |
               +----------------------------------------------+
               |  Multicast Source Address (variable)         |
               +----------------------------------------------+
               |  Multicast Group Length (1 octet)            |
               +----------------------------------------------+
               |  Multicast Group Address (Variable)          |
               +----------------------------------------------+
               |  Originator Router Length (1 octet)          |
               +----------------------------------------------+
               |  Originator Router Address (variable)        |
               +----------------------------------------------+
               |  Leave Group Synchronization  # (4 octets)   |
               +----------------------------------------------+
               |  Maximum Response Time (1 octet)             |
               +----------------------------------------------+
               |  Flags (1 octet)                             |
               +----------------------------------------------+
               |  Outer VLAN (2 octets)                       |
               +----------------------------------------------+
               |  Inner VLAN (2 octets)                       |
               +----------------------------------------------+


               Flags:

               0  1  2  3  4  5  6  7
               +--+--+--+--+--+--+--+--+
               |  |  |  | V|IE|v3|v2|v1|
               +--+--+--+--+--+--+--+--+
```


Figure 2 IGMP Leave Synch Route

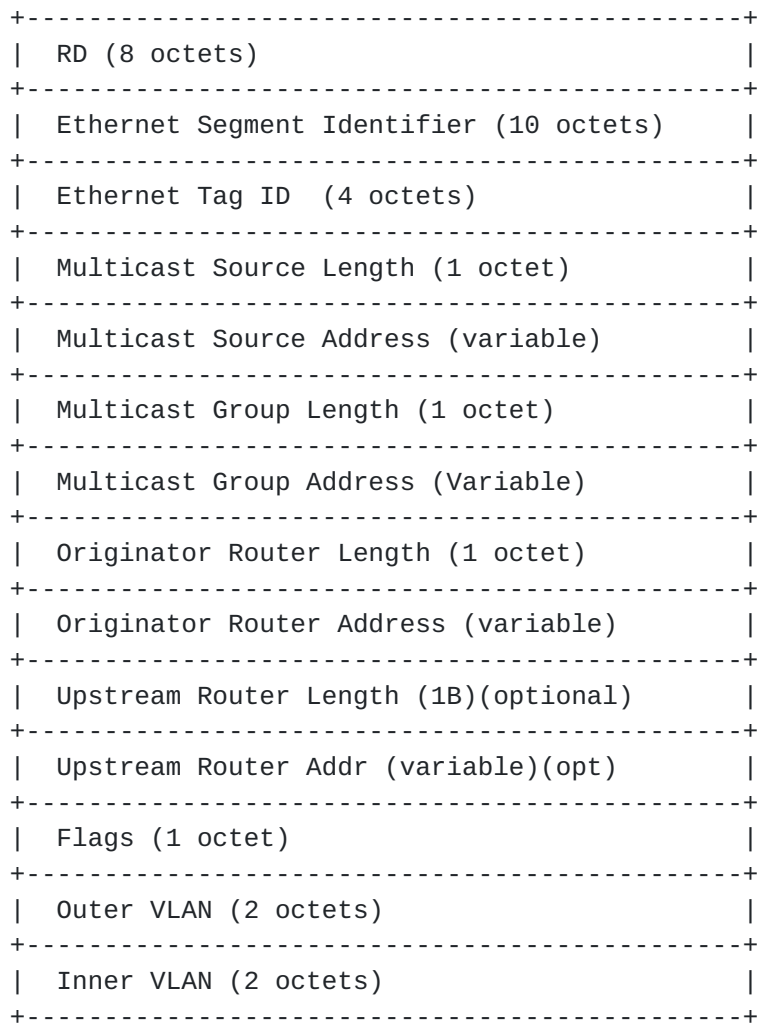This route will have the same extra fields as described in section 3.1 added.

Compared to [EVPN-IGMP-MLD-PROXY] there are two additional fields considered part of the route key for BGP processing when Flag V is set 1.

3.3. PIM RPT-Prune Synch Route

This document extends the PIM RPT-Prune Synch Route defined in
[EVPN-PIM-PROXY] as shown in Figure 3

```
        +----------------------------------------------+
        |  RD (8 octets)                               |
        +----------------------------------------------+
        |  Ethernet Segment Identifier (10 octets)     |
        +----------------------------------------------+
        |  Ethernet Tag ID  (4 octets)                 |
        +----------------------------------------------+
        |  Multicast Source Length (1 octet)           |
        +----------------------------------------------+
        |  Multicast Source Address (variable)         |
        +----------------------------------------------+
        |  Multicast Group Length (1 octet)            |
        +----------------------------------------------+
        |  Multicast Group Address (Variable)          |
        +----------------------------------------------+
        |  Originator Router Length (1 octet)          |
        +----------------------------------------------+
        |  Originator Router Address (variable)        |
        +----------------------------------------------+
        |  Upstream Router Length (1B)(optional)       |
        +----------------------------------------------+
        |  Upstream Router Addr (variable)(opt)        |
        +----------------------------------------------+
        |  Flags (1 octet)                             |
        +----------------------------------------------+
        |  Outer VLAN (2 octets)                       |
        +----------------------------------------------+
        |  Inner VLAN (2 octets)                       |
        +----------------------------------------------+


        Flags:

        0  1  2  3  4  5  6  7
        +--+--+--+--+--+--+--+--+
        |  |  |  |  |  |  |  | V|
        +--+--+--+--+--+--+--+--+
```
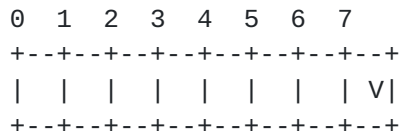
Figure 3 PIM RPT-Prune Synch Route

This route will have the same Outer VLAN and Inner VLAN fields as
described in section 3.1 added and an extra Flags field as IGMP/PIM
Join Synch Route and IGMP Leave Synch Route to indicate the existing

VLAN information.

Compared to [EVPN-PIM-PROXY] there are two additional fields
considered part of the route key for BGP processing when Flag V is
set 1.

4. Security Considerations

TBD

For general EVPN Security Considerations, see [RFC7432].

5. IANA Considerations

TBD

6. References

6.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
          Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC7432] A. Sajassi, Ed., R. Aggarwal, N. Bitar, A. Isaac, J.
          Uttaro, J. Drake, and W. Henderickx, "BGP MPLS-Based
          Ethernet VPN", RFC 7432, February 2015

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
          2119 Key Words", BCP 14, RFC 8174, May 2017

[EVPN-IGMP-MLD-PROXY] Ali Sajassi, Samir Thoria, Keyur Patel, Derek
          Yeung, John Drake and Wen Lin, "IGMP and MLD Proxy for
          EVPN", June 2018, work-in-progress, draft-ietf-bess-evpn-
          igmp-mld-proxy-02.

[EVPN-PIM-PROXY] J. Rabadan, Ed., J. Kotalwar, S. Sathappan, Z.
          Zhang and A. Sajassi, "PIM Proxy in EVPN Networks",
          October 2017, expired, draft-skr-evpn-bess-pim-proxy-01.

6.2. Informative References

TBD

7. Acknowledgments

The authors would like to thank the following for their valuable
contributions of this document:

TBD

Authors' Addresses

Yisong Liu
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: liuyisong@huawei.com

Donald E. Eastlake, 3rd
Huawei Technologies
155 Beaver Street
Milford, MA 01757 USA

Phone: +1-508-333-2270
Email: d3e3e3@gmail.com