BESS Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 13, 2021

W. Wang
A. Wang
China Telecom
H. Wang
Huawei Technologies

March 12, 2021

Layer-3 Accessible EVPN Services
draft-wang-bess-l3-accessible-evpn-04

Abstract

This draft describes layer-3 accessible EVPN service interfaces according to [RFC7432], and proposes a new solution which can simplify the deployment of layer-3 accessible EVPN service. This solution allows each PE in EVPN network to maintain only one IP-VRF.

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). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

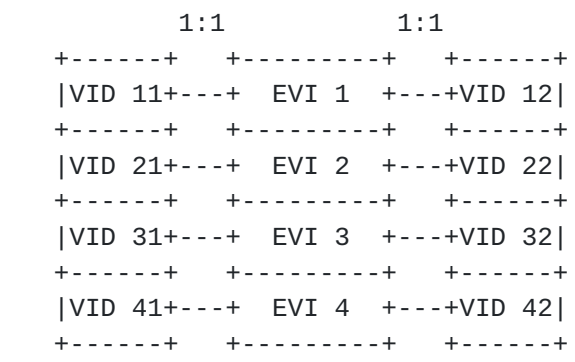
Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 13, 2021.

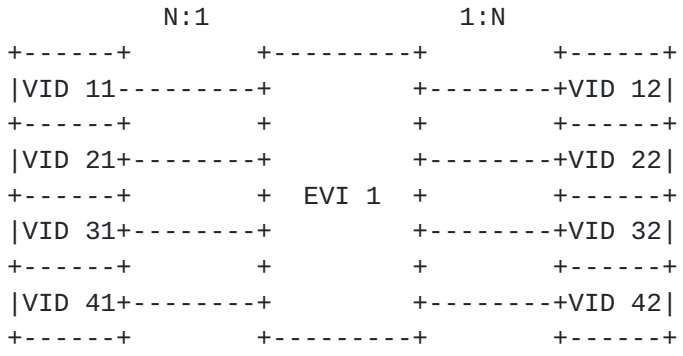
Table of Contents

1. Introduction

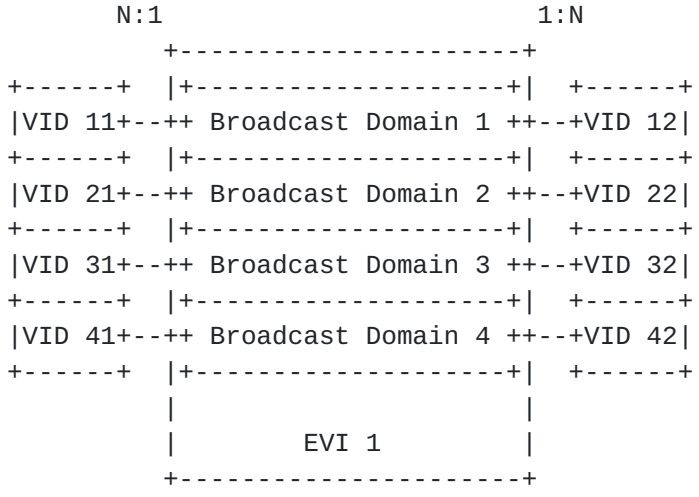
[RFC7432]defines three service interfaces for layer-2 accessible EVPN: VLAN-Based Service Interface, VLAN-Bundle Service Interface and VLAN-Aware Bundle Service Interface. These three types of service interfaces can realize the isolation of layer-2 traffic of customers in different ways, as shown in Figure 1.



VLAN-based Service Interface



VLAN-bundle Service Interface



VLAN-Aware Bundle Service Interface

Figure 1: EVPN Service Interfaces Overview

For VLAN-based service interface, there is a one to one mapping between VID and EVI. Each EVI has a single broadcast domain so that traffic from different customers can be isolated.

For VLAN-bundle service interface, there is a N to one mapping between VID and EVI. Each EVI has a single broadcast domain, but the MAC address MUST be unique that can be used for customer traffic isolation.

For VLAN-aware bundle service interface, there is a N to one mapping between VID and EVI. Each EVI has multiple broadcast domains while the MAC address can overlap. One broadcast domain corresponds to one VID, which can be used to customer traffic isolation.

In the scenarios corresponding to these service interfaces, CE-PE should be placed in the same Layer-2 network. In most of provider network, CE-PE need to cross a Layer-3 network, then the above service interfaces should be extended to adapt to the layer-3 network.

In this draft, we describe three layer-3 accessible interfaces for EVPN, summarize the existing layer-3 accessible EVPN solutions, and propose a new solution which can simplify the depolyment of layer-3 accessible EVPN service.

2. Conventions used in this document

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. Terminology

The following terms are defined in this draft:

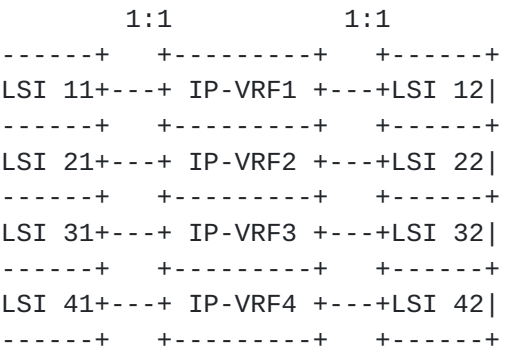
- CE: Client Edge
- PE: Provider Edge
- EVPN: BGP/MPLS Ethernet VPN, defined in [RFC7432]
- VxLAN: Virtual eXtensible Local Area Network, defined in [RFC7348]
- IPSec: Internet Protocol Security, defined in [RFC4301]

4. Service Interfaces in layer-3 accessible EVPN

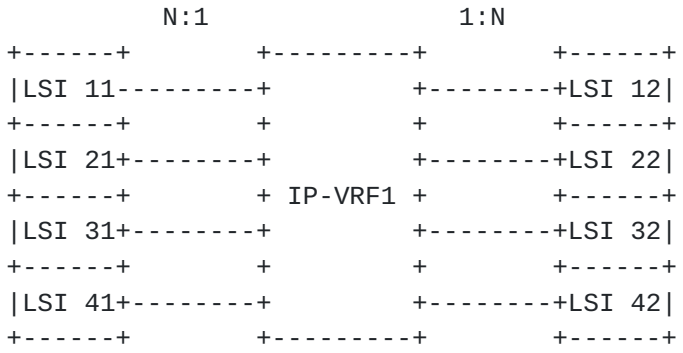
In most of provider network, CE-PE need to cross a Layer-3 network. With this scenario, service interfaces defined in [RFC7432] should be extended to adapt to the layer-3 network. To achieve the traffic isolation, tunnel encapsulation technologies can be used.

We define Logical Session Identifier(LSI) to distinguish the packets from different tunnels, which is related to VNI/SPI. The length of LSI is 16 bits.

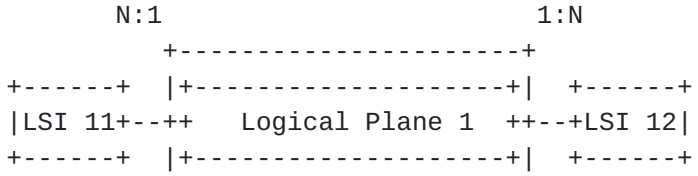
The layer-3 accessible interfaces for EVPN are shown in Figure 2, refer to [RFC7432]



LSI-based Service Interface



LSI-bundle Service Interface



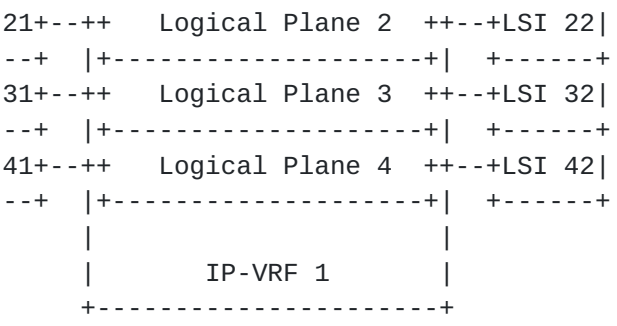


Figure 2: Layer-3 accessible EVPN Service Interfaces Overview

For LSI-based service interface, there is a one to one mapping between LSI and IP-VRF. Each IP-VRF has a single logical plane so that traffic from different customers can be isolated.

For LSI-bundle service interface, there is a N to one mapping between LSI and IP-VRF. Each IP-VRF has a single logical plane, but the IP address MUST be unique that can be used for customer traffic isolation.

For LSI-aware bundle service interface, there is a N to one mapping between LSI and IP-VRF. Each IP-VRF has multiple logical planes while the IP address can overlap. One logical plane corresponds to one LSI, which can be used to customer traffic isolation.

5. Solutions of LSI-aware bundle service interface

Let's assume a scenario as shown in Figure 3. PE1, PE2 and PE3 are EVPN peers, the customer data transmission between PEs relies on VxLAN. CE1, CE2 and CE3 are connected to the sites of customer for its department A and B.

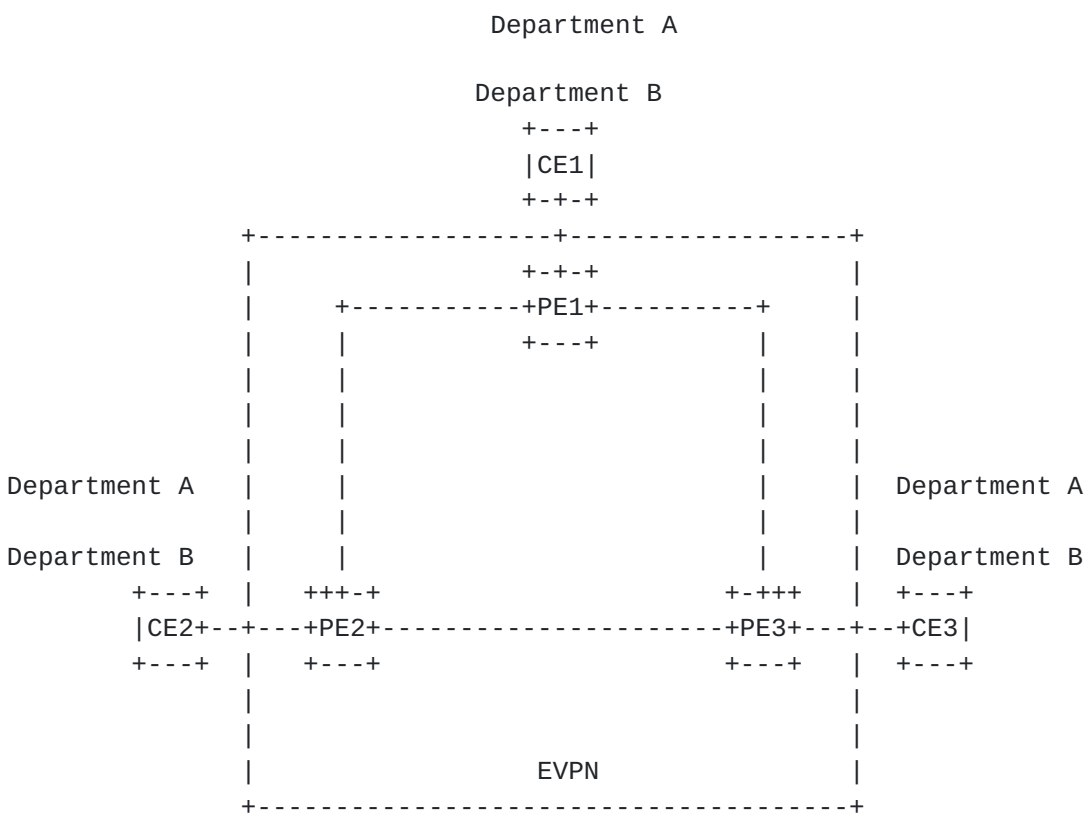


Figure 3: LSI-aware bundle service interface scenario

If each VNI has its own IP-VRF, each PE and CE maintain an IP-VRF for each deployment. In this situation, customer traffic can be isolated by different VNIs, and there is no need for extending control plane/ forwarding plane protocols.

For deployment, we expect a simpler way, such as assign an IP-VRF to each customer, not to each department. That is to say, all VNIs share one IP-VRF on PEs. In this situation, each CE still maintain an IP-VRF for each deployment, but each PE maintains only one VRF for all deployments. In this situation, customer traffic cannot be isolated by VNIs. We propose a solution for this scenario:

- Using LSI information to identify different customer routes / traffic. As described above, LSI can be generated by VNI/SPI, and there is a one to one mapping between LSI and VNI/SPI. PEs should maintain the mapping table of LSI and VNI/SPI, so that they can distinguish different customer routes / traffic. LSI information can be transmitted by using Ethernet Tag ID or a newly defined ESI type.

- TBD (more solutions are welcome).

6. Protocol Extensions

6.1. Forwarding Plane

6.1.1. Extensions to VxLAN

When the forwarding plane uses VxLAN tunnel technologes, we should extend the VxLAN GPE header to carry the LSI information, the extentions to the VxLAN GPE header is shown in Figure 4:

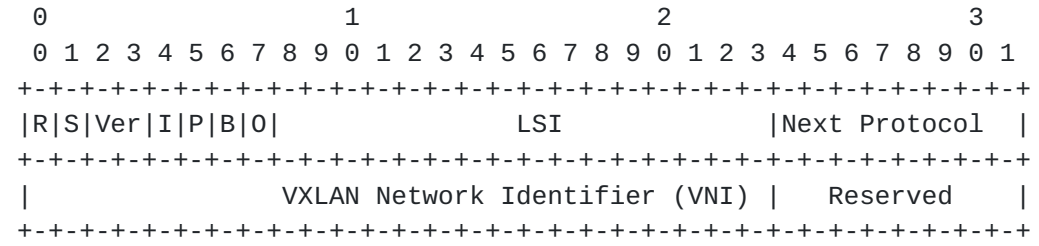


Figure 4: The extentions to VxLAN GPE header

We define a S bit. If S is set to 1, it means the field after O bit contains LSI information.

6.2. Control Plane

We proposed two methods to identify the routes that related to different LSI information:

- Reusing the Ethernet Tag ID. This method requires the update of [I-D.ietf-bess-evpn-prefix-advertisement] (Etherenet Tag ID is set to 0 for route type 5), and may arises some confuse with the original defination of Ethernet Tag ID.

- Using the newly defined ESI type as shown in Figure 5. This method can preserve the original purpose of ESI defination (multi-homing).

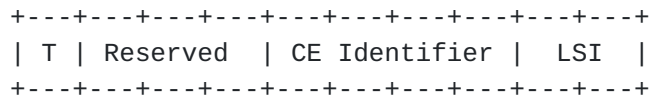


Figure 5: The format of new ESI type

Where:

- T (1 octet): specifys the ESI Type. The recommended value is 0x06.

- CE Identifier (3 octets): the route ID/IPv4 address of CE.

- LSI (2 octets): the LSI information.

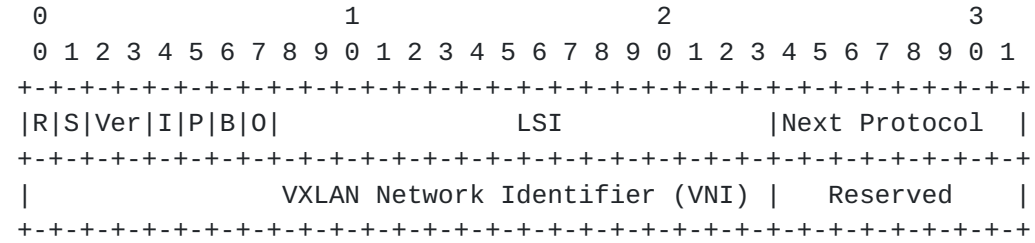
Since the length of LSI is 16 bits, while the length of Ethernet Tag ID and ESI are 80 bits and 32 bits, respectively. We can only use the lower 16 bits of Ethernet Tag ID / ESI field to carry LSI information, the other locations MUST set to 0.

7. Security Considerations

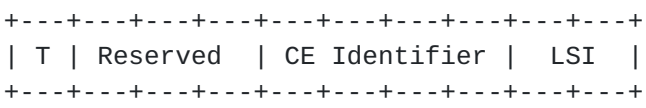
TBD

8. IANA Considerations

This draft extends the VxLAN GPE header, S bit of Flag and LSI field are added:



This draft also define a new ESI type:



9. Normative References

[I-D.ietf-bess-evpn-prefix-advertisement]
Rabadan, J., Henderickx, W., Drake, J., Lin, W., and A. Sajassi, "IP Prefix Advertisement in EVPN", draft-ietf-bess-evpn-prefix-advertisement-11 (work in progress), May 2018.

[I-D.ietf-bess-mvpn-evpn-aggregation-label]
Zhang, Z., Rosen, E., Lin, W., Li, Z., and I. Wijnands, "MVPN/EVPN Tunnel Aggregation with Common Labels", draft-ietf-bess-mvpn-evpn-aggregation-label-05 (work in progress), January 2021.

[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>.

[RFC2890] Dommety, G., "Key and Sequence Number Extensions to GRE", RFC 2890, DOI 10.17487/RFC2890, September 2000, <https://www.rfc-editor.org/info/rfc2890>.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI 10.17487/RFC4301, December 2005, <https://www.rfc-editor.org/info/rfc4301>.

[RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <https://www.rfc-editor.org/info/rfc7348>.

[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>.

Authors' Addresses

Wei Wang
China Telecom
Beiqijia Town, Changping District
Beijing, Beijing 102209
China

Email: weiwang94@foxmail.com

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing, Beijing 102209
China

Email: wangaj3@chinatelecom.cn

Haibo Wang
Huawei Technologies
Huawei Building, No.156 Beiqing Rd.
Beijing, Beijing 100095
China

Email: rainsword.wang@huawei.com