BESS Working Group
Internet-Draft
Intended status: Standards Track
Expires: 21 January 2026

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

20 July 2025

Layer-3 Accessible EVPN Services
draft-wang-bess-l3-accessible-evpn-10

Abstract

This draft describes layer-3 accessible EVPN service interfaces, which aim is to connect the layer 2 customers to one EVPN backbone, via the layer 3 network, and keep the traffic isolation among different layer 2 customers. It proposes to extend the VxLAN packet format to transfer the customer’s Virtual Network Identifier(VNI) information, and also the related control plane extension.

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/.

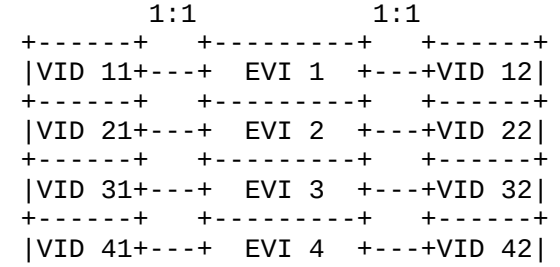
Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 21 January 2026.

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.



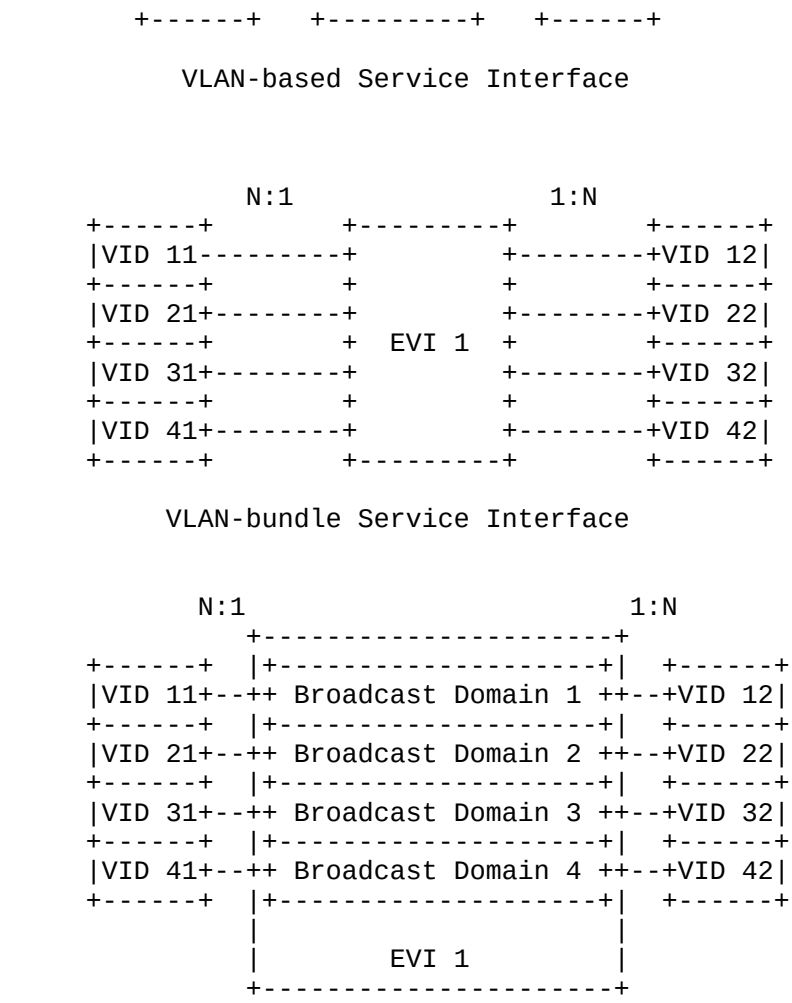


Figure 1: EVPN Service Interfaces Overview

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.

But, 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. Figure 2 shows the typical topology within the operator's network.

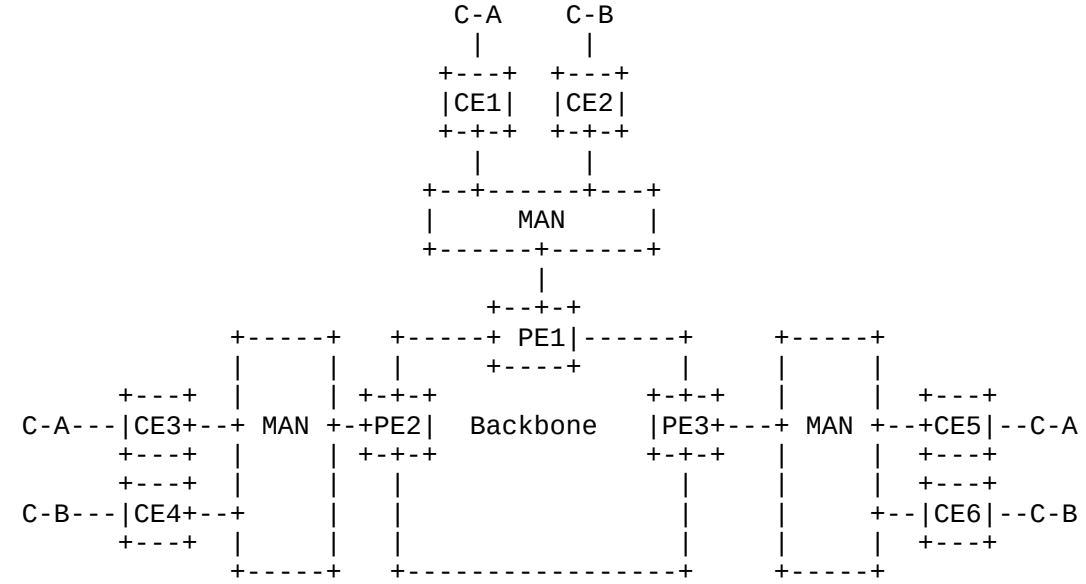


Figure 2: LSI-aware bundle service interface scenario

Assuming that the customer is a cross-regional enterprise, CEs represents the devices that connect its branches into the nearby Metro Area Network(MAN), which is one layer 3 network, and is connected each other via the service provider's backbone network.

The customer wants to connect its branch sites together via the service provider's backbone network. The service provider deploy one EVPN instance for this customer within its backbone network, but can't connect each of these branches via the traditional layer 2 access EVPN interfaces [RFC7432], because the MAN is one layer 3 network, and there is no user's VLAN information that can be used to isolate the traffic of its different divisions.

This scenario may involve the following situations:

* point-to-point communication (e.g. the communication between C-A connected to CE3 and C-A connected to CE5.)

* point-to-multipoint communication (e.g. C-A connected to CE-3 needs to communicate with both C-A connected to CE5 and C-A connected to CE-1 simultaneously.)

* multipoint-to-multipoint communication (e.g. mutual communication among the three C-As in Figure 2.)

In this draft, we describe three layer-3 accessible interfaces for EVPN, the above problem can be solved by using these L3 accessible interfaces.

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

* EVPN: BGP/MPLS Ethernet VPN, defined in [RFC7432]

* IPSec: Internet Protocol Security, defined in [RFC4301]

* Layer-3 accessible interface for EVPN: The interface, which is tunnelled over one layer 3 network, can be used to access the EVPN service, and keep the traffic within the EVPN forwarding plane isolated among different customer domains.

* PE: Provider Edge

* SPI: Security Parameters Index, defined in [RFC4301]

* VNI: VXLAN Network Identifier (or VXLAN Segment ID), defined in [RFC7348]

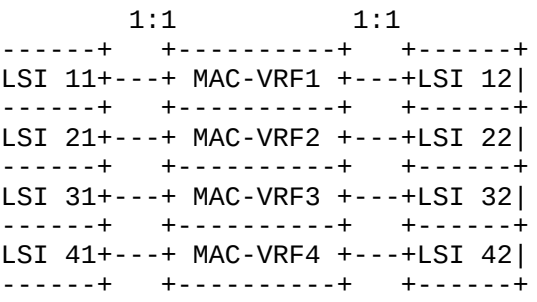
* VxLAN: Virtual eXtensible Local Area Network, defined in [RFC7348]

4. Service Interfaces in layer-3 accessible EVPN

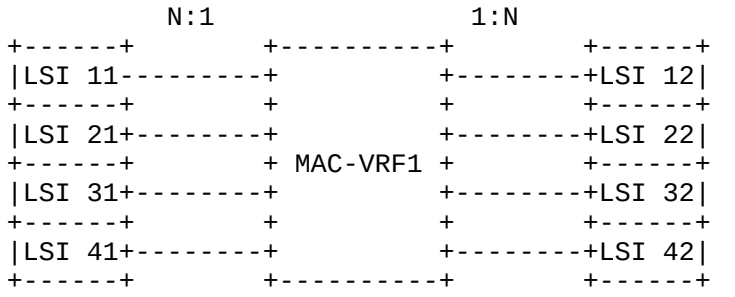
In most of provider network, CE-PE need to cross a Layer-3 network. With this scenario, service interfaces defined in [RFC7432] should be extended to adapt to the layer-3 network. To achieve the traffic isolation, tunnel encapsulation technologies can be used.

We define Logical Session Identifier(LSI) to distinguish the customer’s packets from different tunnels, which is VNI when the tunnel is VxLAN. The length of LSI is 16 bits.

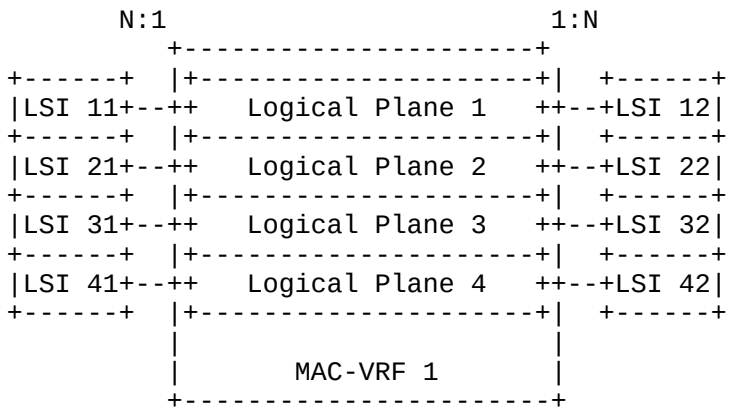
The concepts of layer-3 accessible interfaces for EVPN are shown in Figure 3, refer to [RFC7432]



LSI-based Service Interface



LSI-bundle Service Interface



LSI-Aware Bundle Service Interface

Figure 3: Layer-3 accessible EVPN Service Interfaces Overview

For LSI-based service interface, there is a one to one mapping between LSI and MAC-VRF. Each MAC-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 MAC-VRF. Each MAC-VRF has a single logical plane, but the MAC 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 MAC-VRF. Each MAC-VRF has multiple logical planes while the MAC 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

For LSI-Aware Bundle service interface, the PE should maintain one MAC-VRF that be sub-divided into different logical plane. Similar with the VLAN-Aware Bundle service, it needs the forwarding plane of the customer’s packet to carry the customer’s LSI information, and also the control plane extension to transfer the required the LSI information of the communication peer.

6. Protocol Extensions

6.1. Forwarding Plane

6.1.1. Extensions to VxLAN

When the forwarding plane uses VxLAN tunnel technologies, we should extend the VxLAN header[RFC7348] to carry the LSI information, the extentions to the VxLAN header is shown in Figure 4:

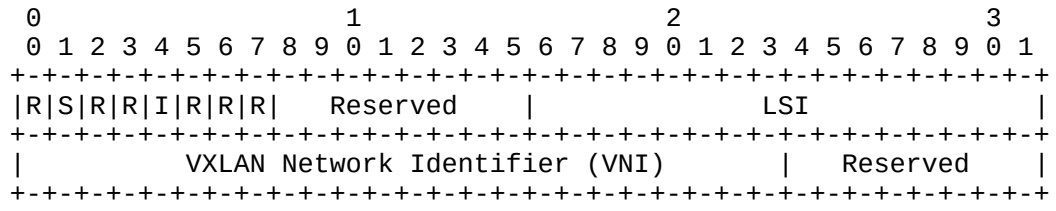


Figure 4: The extentions to VxLAN header

We define one flag “S” from the reserved bits of the current VxLAN header, to indicate the last 16 bits of first 4-bytes indicates the value of “LSI”

6.2. Control Plane

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

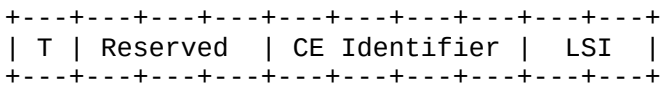


Figure 5: The format of new ESI type

Where:

* T (1 octet): specify 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 bits MUST set to 0.

7. Modification of MAC address storage mode on PE

LSI-aware bundle service interface also changes the storage mode of MAC address on PE, as shown in Figure 6.

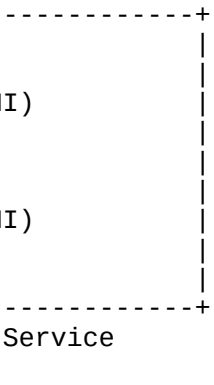


Figure 6: Modification of MAC/IP address storage mode on PE

For end-to-end layer-2 data transmission, the storage mode of MAC address in MAC-VRF is similar to VLAN-aware bundle service, the only change is that different bridge domains are distinguished by LSI.

8. Gap analysis

This section specifies the differences among L3 accessible EVPN with EVPN([RFC7432]), EVPN-VPWS ([RFC8214]) and EVPN-Etree.

8.1. Differences with EVPN

[RFC7432] defines three types of Layer-2 access EVPN interfaces (VLAN-Based, VLAN-Bundle, VLAN-Aware Bundle), which are applicable to the scenarios where customers directly access the provider’s PE through Layer-2 links (such as an enterprise network connecting to the PE via a dedicated line).

The L3 accessible EVPN is designed for the scenarios where CE-PE needs to cross a Layer-3 network (such as a customer branch accessing the backbone network PE through an IP Metropolitan Area Network / MAN), it proposes three types of Layer-3 accessible interfaces (LSI-Based, LSI-Bundle, LSI-Aware Bundle). These interfaces support accessing EVPN through Layer-3 tunnels (such as VxLAN) and solve the problem of traffic isolation across Layer-3 networks.

8.2. Differences with EVPN-VPWS

[RFC8214] defines the EVPN Virtual Private Wide Area Network Service (VPWS), which provides point-to-point (P2P) Layer-2 dedicated line connections. The core is to advertise Ethernet A-D routes through BGP to establish dedicated line forwarding tunnels, without dealing with the traffic isolation problem of multiple customers on the same PE.

L3 accessible EVPN focuses on the traffic isolation when multiple customers access the same PE through a Layer 3 network. It achieves logical isolation of Layer-2 traffic of different customers through LSI/MAC-VRF mapping (similar to the multi-tenant scenario). In contrast, VPWS focuses more on the establishment of dedicated line connections and does not handle multi-tenant isolation.

8.3. Differences with EVPN-ETree

[RFC8317] defines the EVPN-ETree service, which supports point-to-multipoint tree-shaped topologies (such as video conferencing, content distribution). It optimizes traffic forwarding through E-Tree routes to reduce duplicate traffic.

L3 accessible EVPN aims at the isolation of traffic for multipoint-to-multipoint connections in unicast scenarios (such as different branches of an enterprise accessing the same EVPN instance through a Layer-3 network). It divides logical planes through the mapping of

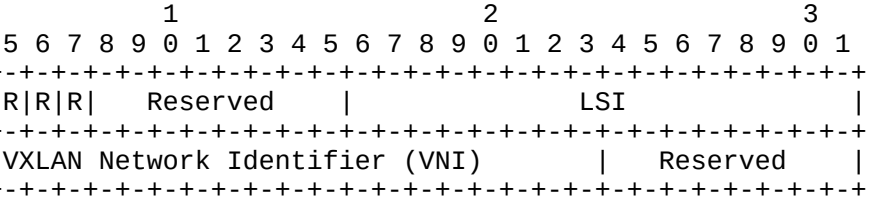
LSI and MAC-VRF to ensure that the traffic of different customers does not communicate with each other, which has nothing to do with multicast.

9. Security Considerations

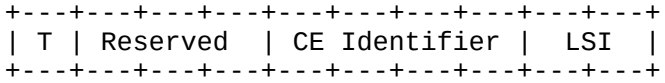
TBD

10. IANA Considerations

This document creates a 1-bit registry called "S bit". New registrations will be made through the "RFC Required" procedure defined in [RFC8126]. Initial registrations are as follows: The second bit on the left side of the VXLAN header is defined as the "S bit," and the reserved field occupying bits 17 to 32 is defined as "LSI" field. When S bit is set to 1, the "LSI" field carries the value of LSI; otherwise, the value of "LSI" field should not be see as the value of LSI.



This draft also define a new ESI type:



11. Acknowledgements

Thanks Jeffrey Zhang for its review and discussion to improve this document.

12. Normative References

[I-D.ietf-bess-evpn-prefix-advertisement]
Rabadan, J., Henderickx, W., Drake, J., Lin, W., and A. Sajassi, "IP Prefix Advertisement in Ethernet VPN (EVPN)", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-prefix-advertisement-11, 18 May 2018, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-prefix-advertisement-11>.

[I-D.ietf-bess-mvpn-evpn-aggregation-label]
Zhang, Z. J., Rosen, E. C., Lin, W., Li, Z., and I. Wijnands, "MVPN/EVPN Tunnel Aggregation with Common Labels", Work in Progress, Internet-Draft, draft-ietf-bess-mvpn-evpn-aggregation-label-14, 4 October 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-mvpn-evpn-aggregation-label-14>.

[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>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[RFC8214] Boutros, S., Sajassi, A., Salam, S., Drake, J., and J. Rabadan, "Virtual Private Wire Service Support in Ethernet VPN", RFC 8214, DOI 10.17487/RFC8214, August 2017, <https://www.rfc-editor.org/info/rfc8214>.

[RFC8317] Sajassi, A., Ed., Salam, S., Drake, J., Uttaro, J., Boutros, S., and J. Rabadan, "Ethernet-Tree (E-Tree) Support in Ethernet VPN (EVPN) and Provider Backbone Bridging EVPN (PBB-EVPN)", RFC 8317, DOI 10.17487/RFC8317, January 2018, <https://www.rfc-editor.org/info/rfc8317>.

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