BGP EVPN

Foreword

- Standard BGP-4 supports only IPv4 unicast addresses. To support more network layer protocols, Multiprotocol Extensions for BGP-4 (MP-BGP) (RFC 4760) was proposed as an extension to BGP-4 to allow different types of address families to be distributed in BGP at the same time. The address families include IPv4 multicast, IPv6, L3VPN, and Ethernet Virtual Private Network (EVPN) address families.
- With the development and commercial use of software-defined networking (SDN), EVPN plays an important role in various solutions, covering all scenarios, including campus networks, data centers, IP WAN transport networks, and software-defined networking in a wide area network (SD-WAN).
- This course describes the concept of MP-BGP, development history of EVPN, common EVPN route types, and EVPN usage scenarios.

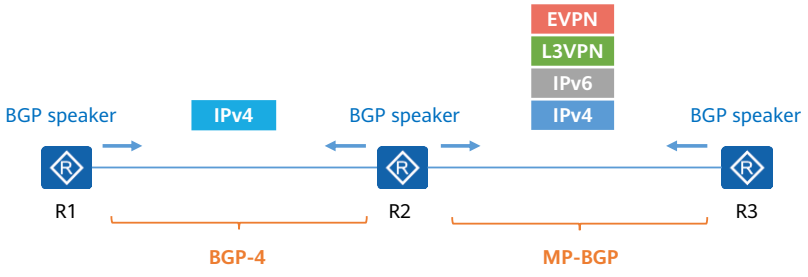
Objectives

- Upon completion of this course, you will be able to:
  - Understand basic MP-BGP concepts.
  - Understand the origin of EVPN.
  - Understand common EVPN route types.
  - Understand typical EVPN usage scenarios.

Contents

MP-BGP

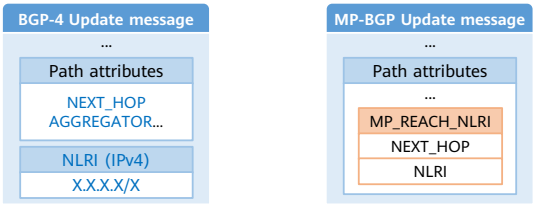
As defined in RFC 4760, MP-BGP is used to extend BGP-4 to allow BGP to carry multiple network layer protocols, such as IPv6, L3VPN, and EVPN. This extension has good backward compatibility. That is, an MP-BGP-capable router can interact with a router that supports only BGP-4.



- https://datatracker.ietf.org/doc/rfc4760/

BGP-4 Extensions

- BGP-4 has three IPv4-specific pieces of information: NEXT_HOP, AGGREGATOR, and IPv4 network layer reachable information (NLRI). To support multiple network layer protocols, BGP-4 has to provide the following abilities:
  - Ability of associating network layer protocols with next-hop information
  - Ability of associating network layer protocols with NLRIs
- The two abilities are collectively referred to as the address family (AF) defined by the Internet digital distribution agency (IANA).
- To implement forward compatibility, MP-BGP adds two new attributes: MP_REACH_NLRI and MP_UNREACH_NLRI, which are used to indicate reachable and unreachable destinations, respectively. The two attributes are optional non-transitive.



- According to BGP-4, NEXT_HOP and AGGREGATOR fields are contained in Path attributes of IPv4, and the IPv4 NLRI carries IPv4 routing entries.
- The Path attributes field is added in MP-BGP. MP_REACH_NLRI is a new field of path attributes. The NEXT_HOP and NLRI fields of the corresponding network layer protocol and the NLRI belong to MP_REACH_NLRI.

MP_REACH_NLRI

- MP_REACH_NLRI is carried in a BGP Update message and provides the following functions:
  - Advertises reachable routes to BGP peers.
  - Advertises the next-hop address of a reachable route to a BGP peer.
- It contains the following fields:

MP_REACH_NLRI Format	Field Description
Address Family Identifier (2 octets)	Network layer protocol. Value 2 indicates IPv6.
Subsequent Address Family Identifier (1 octet)	This field is used together with address family identifier (AFI). Value 1 indicates unicast, and value 2 indicates IPv6 unicast.
Length of Next Hop Network Address (1 octet)	Length of a next-hop IP address.
Network Address of Next Hop (variable)	Next-hop address. The format is determined by the AFI and SAFI.
Reserved (1 octet)	All 0s.
Network Layer Reachability Information (variable)	The length of this field is variable. This field can contain reachable routes.

- In the SAFI field, value 1 indicates unicast, and value 2 indicates multicast. The value is allocated by the IANA. The allocation rules are defined in RFC 2434 (titled "Guidelines for Writing an IANA Considerations Section in RFCs").
- In this section, the AFI of EVPN is 25 (L2VPN) and the SAFI is 70 (EVPN).

MP_UNREACH_NLRI

- MP_UNREACH_NLRI is carried in BGP Update messages to withdraw unreachable routes.
- It contains the following fields:

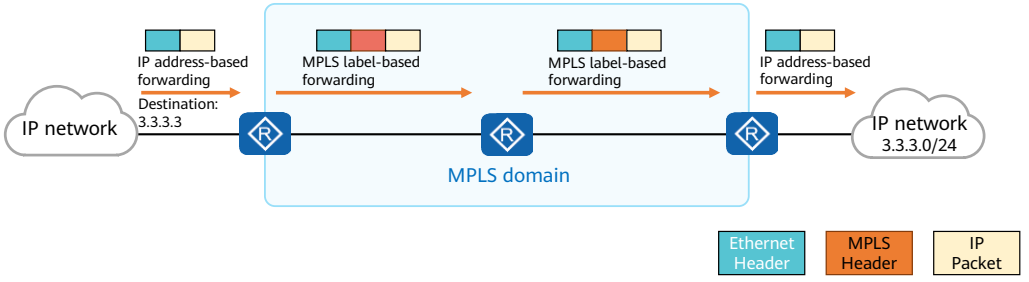
MP_UNREACH_NLRI Format	Field Description
Address Family Identifier (2 octets)	Network layer protocol. Value 2 indicates IPv6.
Subsequent Address Family Identifier (1 octet)	This field is used together with the AFI. Value 1 indicates unicast, and value 2 indicates IPv6 unicast.
Withdrawn Routes (variable)	The length of this field is variable. This field lists the routes that need to be withdrawn. The format of this field is determined by the AF and SAFI.

- The AFI of EVPN is 25 (L2VPN) and the subsequent address family identifier (SAFI) is 70 (EVPN).

Contents

MPLS Overview

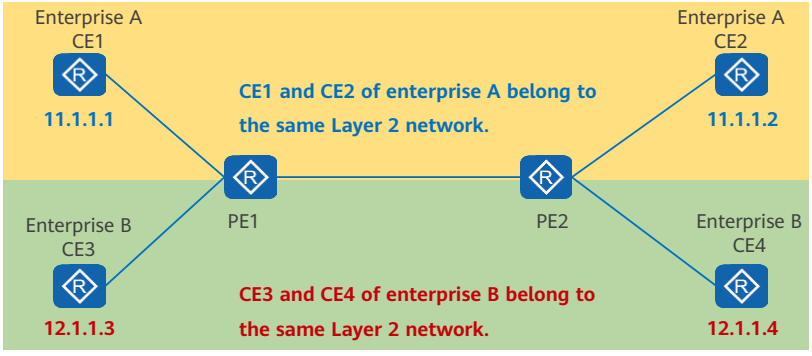
- Multiprotocol Label Switching (MPLS) is located between the data link layer and the network layer in the TCP/IP protocol stack. An MPLS header is added between the two layers. Packets are forwarded based on the MPLS header. The MPLS header is also called the MPLS label.
- MPLS replaces IP forwarding with label switching to implement label-based rapid forwarding.



- MPLS originates from IPv4 and its core technologies can be extended to multiple network protocols, including IPv6, Internet Packet Exchange (IPX), Appletalk, DECnet and Connectionless Network Protocol (CLNP). "Multiprotocol" in MPLS indicates that multiple network protocols are supported.
- MPLS replaces IP forwarding with label switching. A label is a short and fixed-length connection identifier that has only local significance. It is similar to the virtual path identifier (VPI)/virtual channel identifier (VCI) of Asynchronous Transfer Mode (ATM) and the data link connection identifier (DLCI) of Frame Relay.
- MPLS domain: An MPLS domain consists of a series of consecutive network devices that run MPLS.

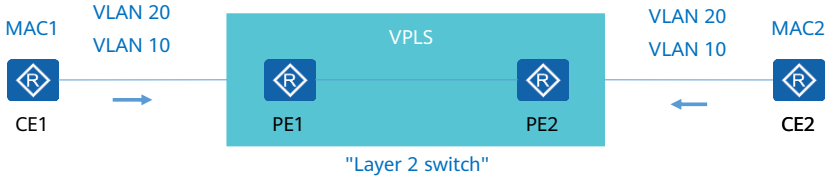
VPLS Overview

Virtual private LAN service (VPLS) is an Ethernet-based L2VPN technology. VPLS provides services similar to LAN services on an MPLS network and allows users to access the network from different locations.



Traditional L2VPN

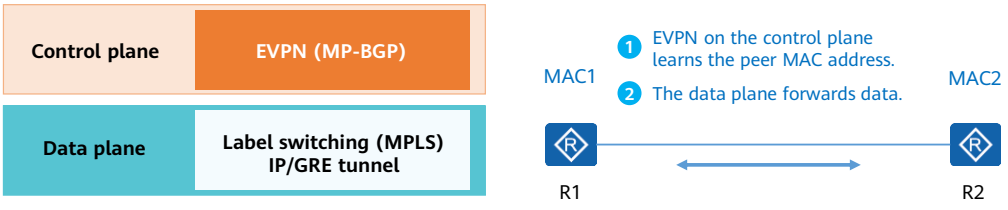
- Traditional L2VPN services, such as VPLS, provide Layer 2 connections between remote sites. An L2VPN network is built and functions like a Layer 2 switch to transparently transmit Ethernet packets. In this example, PE1 and PE2 form a VPLS network to transparently transmit VLAN traffic between CE1 and CE2.
- In a traditional L2VPN, remote MAC addresses are learned through ARP broadcast flooding, and therefore, PEs need to carry broadcast traffic. Broadcast consumes a large amount of interface bandwidth, which is a typical issue of traditional L2VPN.



- VPLS does not support all-active access or load balancing and implements slow fault convergence. For details, see materials of the HCIE-HCIE-Datacom Ethernet VPN and RFC 7209 titled "Requirements for Ethernet VPN (EVPN)."

Emergence of EVPN

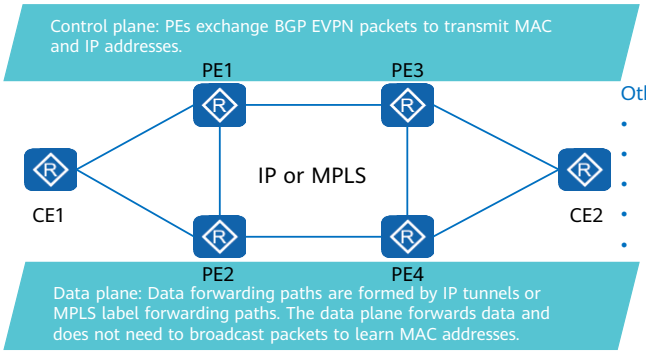
- With new technologies and scenarios emerging, VPLS cannot meet the requirements of L2VPN. The industry has reviewed the requirements for Ethernet VPN (RFC 7209) and proposed a new solution, that is, EVPN.
- EVPN was first defined in RFC 7432. EVPN introduces the control plane to better control MAC address learning.
- EVPN uses MP-BGP on the control plane and supports MPLS label switched paths (LSPs) or IP/Generic Routing Encapsulation (GRE) tunneling on the data plane.



- https://datatracker.ietf.org/doc/rfc7209/
- https://datatracker.ietf.org/doc/rfc7432/

Advantages of EVPN

- EVPN introduces the control plane to learn MAC and IP addresses to guide data forwarding, implementing forwarding-control separation.
- EVPN resolves typical problems in traditional L2VPNs and offers more benefits, such as active-active, rapid convergence, and simplified O&M.



Other advantages of EVPN:

- CEs can access PEs in all-active mode.
- Automatic discovery of PE members
- Loop prevention
- Broadcast traffic optimization
- Support for Equal-Cost Multi-Path (ECMP)

Contents

EVPN NLRI

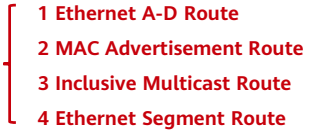
- EVPN defines a new type of BGP NLRI, known as EVPN NLRI, to carry all EVPN routes.
- EVPN NLRI is a new extension to MP-BGP. It is included in MP_REACH_NLRI. For the EVPN NLRI, the AFI is 25 and the SAFI is 70.

MP_REACH_NLRI format	MP_REACH_NLRI of BGP EVPN
Address Family Identifier (2 octets)	**AFI: 25**
Subsequent Address Family Identifier (1 octet)	**SAFI: 70**
Length of Next Hop Network Address (1 octet)	Length of a next-hop IP address.
Network Address of Next Hop (variable)	Next-hop IP address in an EVPN route.
Reserved (1 octet)	All 0s.
Network Layer Reachability Information (variable)	**EVPN NLRI**

EVPN Route

- The EVPN NLRI format uses the Type-Length-Value (TLV) structure, making packets highly flexible and scalable.
  - Route Type field: defines different EVPN routes. RFC 7432 defines four types of routes.
  - Length field: defines the length of a field.
  - Route Type Specific field: contains fields of a particular route type.

EVPN NLRI format



Four common EVPN routes defined in the Route Type field

- The NLRI field in the MP_REACH_NLRI/MP_UNREACH_NLRI attribute contains the EVPN NLRI (encoded as specified above).
- The EVPN NLRI is carried in BGP [RFC4271] using BGP Multiprotocol Extensions [RFC4760] with an Address Family Identifier (AFI) of 25 (L2VPN) and a Subsequent Address Family Identifier (SAFI) of 70 (EVPN). The NLRI field in the MP_REACH_NLRI/MP_UNREACH_NLRI attribute contains the EVPN NLRI (encoded as specified above).
- In order for two BGP speakers to exchange labeled EVPN NLRI, they must use BGP Capabilities Advertisements to ensure that they both are capable of properly processing such NLRI. This is done as specified in [RFC4760], by using capability code 1 (multiprotocol BGP) with an AFI of 25 (L2VPN) and a SAFI of 70 (EVPN).

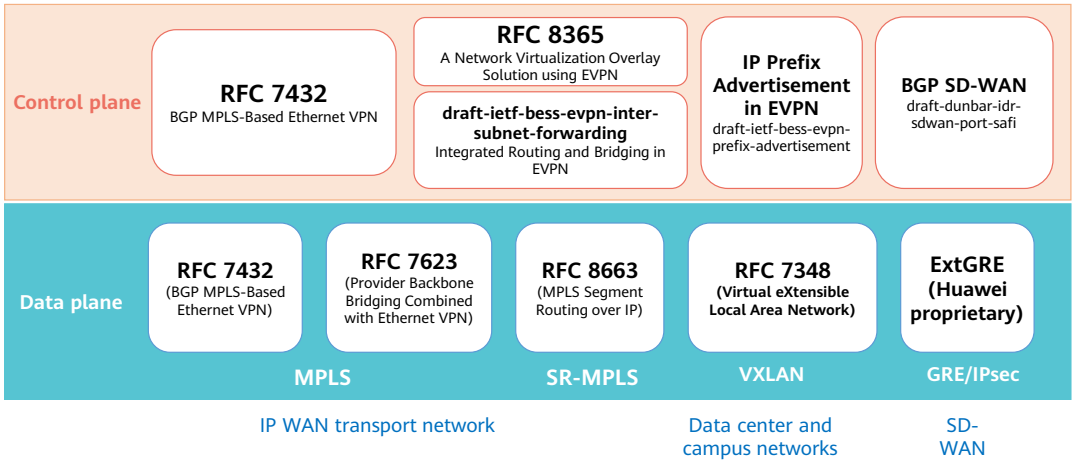
More Types of EVPN Routes and Their Functions

EVPN is not limited to L2VPN applications. With the increase of EVPN route types, more applications, such as L3VPN, are supported.

Type of Route	Function	RFC
(Type 1) Ethernet A-D Route	• Aliasing • MAC address batch withdraw • All-active flag • ESI label advertisement	RFC 7432
(Type 2) MAC/IP Advertisement Route	• MAC address learning notification • MAC/IP binding • MAC mobility	
(Type 3) Inclusive Multicast Route	Automatic discovery of multicast tunnel endpoints and multicast types	
(Type 4) Ethernet Segment Route	Automatic discovery of ES members DF election	
(Type 5) IP Prefix Route	IP prefix advertisement (support for L3VPN)	draft-ietf-bess-evpn-prefix-advertisement

- The Type 5 route (IP prefix route) related standard is in the draft phase, in draft-ietf-bess-evpn-prefix-advertisement.

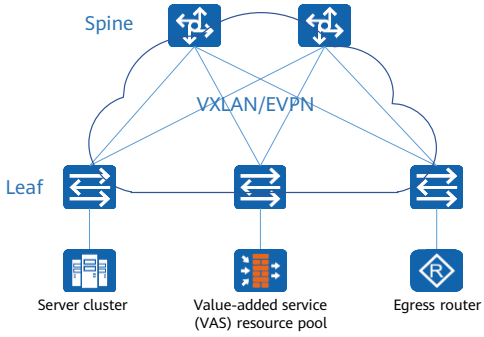
EVPN Protocol Standards



Contents

EVPN on a Data Center Network

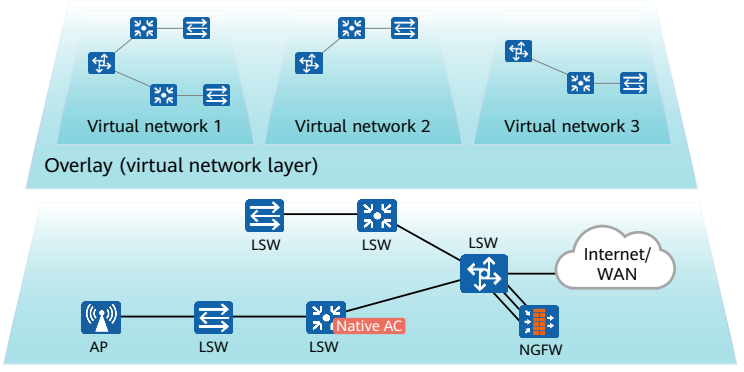
- The network virtualization overlay (NVO) solution (RFC 8365) of EVPN is used in cloud data centers.
- It is recommended that the data plane use Virtual Extensible LAN (VXLAN) encapsulation and the control plane use EVPN to construct a flexible data center overlay network.



- All services in the data center are carried by the VXLAN overlay network.
- The underlay network consisting of spine and leaf nodes performs high-speed forwarding.

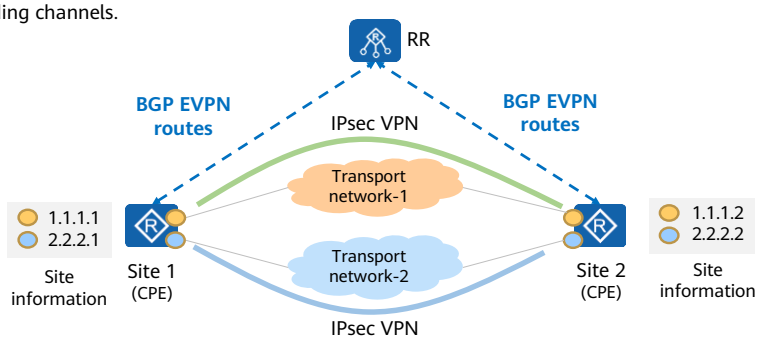
EVPN Application on a Campus Network

- The campus network virtualization solution is similar to that in the cloud data center. The EVPN NVO solution (RFC 8365) is used.
- VXLAN encapsulation and EVPN are used on different underlying networks to build a flexible overlay network.



EVPN Application in SD-WAN

- SD-WAN is a next-generation enterprise branch interconnection solution that supports intelligent dynamic traffic steering, Zero Touch Provisioning (ZTP), and visualization.
- In the SD-WAN solution, EVPN is deployed between route reflectors (RRs) and customer-premises equipment (CPE) devices to advertise SD-WAN overlay VPN routes on the control plane. IPsec VPN is used on the data plane to build secure forwarding channels.



- Overlay VPN routes include site VPN route prefixes, next-hop route information, and IPsec key pairs required for data encryption of data channels between CPEs.

Quiz

1. (Essay) Please describe the principles and common route types of EVPN.
2. (Essay) Please describe usage scenarios of EVPN.

1. EVPN is an extension to MP-BGP. EVPN provides five major types of routes and is used as the control plane of Layer 2 or Layer 3 tunnels.
2. EVPN can be widely used in all enterprise scenarios, such as SD-WAN, campus networks, data centers, and WANs. In data centers and campus networks, EVPN and VXLAN are used together to construct a service overlay network. In SD-WAN scenarios, EVPN and IPsec are used together to build enterprise branch interconnection networks. On a WAN, EVPN can be used with various underlying tunneling and label technologies, such as MPLS, Segment Routing (SR), VPLS, and virtual private wire service (VPWS).

Summary

- MP-BGP's extension to BGP-4 allows different types of address families, such as IPv4 multicast, IPv6, L3VPN, and EVPN, to be distributed in BGP.
- This course describes EVPN that is used to solve the Ethernet L2VPN problems. With the increase of usage scenarios and protocol extensions, EVPN can be used in various scenarios, including WANs, data centers, campus networks, and SD-WANs.