Getting Started with EVPN MPLS

This chapter introduces EVPN MPLS, covering key concepts, operational modes, and essential route types. Users can learn how EVPN MPLS enables scalable, simplified Layer 2 and Layer 3 VPN services, supports redundancy and load balancing, and understand important configuration and timer details for efficient network deployment.

EVPN MPLS

An Ethernet VPN (EVPN) is a networking technology that

- offers scalable and simplified Layer 2 and Layer 3 VPN services
- enables control-plane-based MAC learning using the MP-BGP protocol, and
- supports flexible data-plane encapsulations while maintaining a unified control plane.

Multiprotocol Border Gateway Protocol (MP-BGP) is a routing protocol designed to exchange routing and reachability information across multiple network layer protocols, enabling versatile and scalable network communication.

Ethernet architecture distinctly separates the control plane from the data plane, which significantly improves network scalability and simplifies operational management. EVPN supports efficient traffic load balancing and facilitates various service models, including E-LAN, E-LINE, and E-TREE, to meet diverse networking requirements.

Need for EVPN

Today, our networks have different protocols serving different purposes, which makes daily operations more complex than they need to be. This hinders the ability to deliver end-to-end services with speed and agility. As you deploy multiple geographically disparate data centers, they're looking for scalable and simplified network solutions to extend virtualization and cluster domains between multiple data centers.

The evolution of EVPN started due to the need for a scalable solution to bridge various layer 2 domains and overcome the limitations faced by VPLS, such as scalability, multihoming, and per-flow load balancing.

EVPN addresses the shortcomings of VPLS, including scalability, multihoming, and per-flow load balancing. EVPN provides secure and private connectivity for multiple sites within an organization spread across different geographical locations. EVPN uses MAC addresses as routable addresses and distributes them to all participating PEs through the MP-BGP EVPN control plane.

Note Starting with the Cisco IOS XR Release 24.3.1, line cards and routers equipped with the Q100, Q200, and P100-based Silicon One ASICs support EVPN MPLS.

To know more about EVPN, see https://e-vpn.io.

Benefits of EVPN

These are the key benefits of EVPN:

- Per flow-based load balancing—enables more granular load balancing of traffic across available paths.
- Scalability—offers a more scalable solution compared to previous L2VPN technologies.
- Reduced operational complexity—simplifies network operations by providing a unified approach for L2 and L3 VPN services.
- Improved network efficiency by eliminating flooding and learning—reduces or eliminates the need for flooding and learning in the data plane.
- Provides fast reroute, resiliency, fast reconvergence during link failure—offers mechanisms for fast reroute and reconvergence in case of link failures.
- Integrates L2 and L3 VPN services—integrates both Layer 2 and Layer 3 VPN services.

Key concepts of EVPN

Here are the key concepts related to EVPN fundamentals, including Ethernet Segments, EVPN instances, and associated mechanisms for traffic management and redundancy:

- Ethernet segment (ES)—is a set of Ethernet links that connects a multihomed device or network to two or more PEs, and uses Ethernet segment route (route type 4) for designated forwarder (DF) election for BUM traffic.
- Ethernet segment identifiers (ESI)—are unique non-zero identifiers that are assigned to Ethernet segments, and represent each Ethernet segment uniquely across the network.
- EVPN instances (EVI)—is a virtual network identifier (VNI) on a PE router that acts like a VPN Routing and Forwarding (VRF), uses Route Targets (RTs) for import and export policies, maps traffic from ports or VLANs to a Bridge Domain (BD), and forwards it to the MPLS core.
- Ethernet auto discovery routes per ES (EAD-ES)—is also referred to as route type 1 that is used to converge traffic faster during access failure scenarios, and has an Ethernet Tag of 0xFFFFFFFF.
- Ethernet auto discovery routes per EVI (EAD-EVI)— is also referred to as route type 1 that is used for aliasing and load balancing when traffic only hashes to one of the switches, and cannot have an Ethernet tag value of 0xFFFFFFFF to differentiate it from the EAD-ES route.

- Aliasing—is a process that is used for load balancing traffic to all connected switches for a given Ethernet segment using the route type 1 EAD-EVI route, and is performed irrespective of the switch where the hosts are actually learned.
- Mass withdrawal—is a process that is used for fast convergence during access failure scenarios using the route type 1 EAD-ES route.
- DF election—is a process that is used to prevent forwarding loops, and allows only a single router to decapsulate and forward traffic for a given Ethernet Segment.
- VPN membership—is the process where PEs exchange EVPN routes at startup to discover remote PE members of an EVI, which enables building flood lists for multicast ingress replication and exchanging BUM and unicast labels for MAC address learning.
- Ethernet segment reachability—is a process where, in multihoming scenarios, the PE auto-discovers remote PE and the corresponding redundancy mode (all-active or single-active), which allows PEs to withdraw the routes used at this stage in case of segment failures in order to trigger fast convergence by signaling a MAC mass withdrawal on remote PEs.
- Redundancy group membership—is a process where PEs connected to the same Ethernet segment (multihoming) automatically discover each other and elect a Designated Forwarder (DF) that is responsible for forwarding Broadcast, Unknown unicast, and Multicast (BUM) traffic for a given EVI.

EVPN route types

EVPN defines several route types that serve different purposes in delivering Layer 2 and Layer 3 services over an IP-MPLS network. Each route type carries specific information.

Route Type	Description	Usage
Route Type 1: Ethernet Auto-Discovery (AD) Route	The Ethernet Auto-Discovery (AD) routes are advertised on per EVI and per ESI basis. These routes are sent per ES. They carry the list of EVIs that belong to the ES. The ESI field is set to zero when a CE is single-homed. This route type is used for mass withdrawal of MAC addresses and aliasing for load balancing.	Few routes are sent per ES, carries the list of EVIs that belong to ES.
Route Type 2: MAC-IP Advertisement Route	These routes are per-VLAN routes, so only PEs that are part of a VNI require these routes. The host's IP and MAC addresses are advertised to the peers within NRLI. The control plane learning of MAC addresses reduces unknown unicast flooding.	Advertises MAC address reachability and IP-MAC binding.
Route Type 3: Inclusive Multicast Ethernet Tag Route	This route establishes the connection for broadcast, unknown unicast, and multicast (BUM) traffic from a source PE to a remote PE. This route is advertised on per VLAN and per ESI basis.	Discovers multicast tunnel end point.

Route Type	Description	Usage
Route Type 4: Ethernet Segment Route	Ethernet segment routes enable to connect a CE device to two or PE devices. ES route enables the discovery of connected PE devices that are connected to the same Ethernet Segment.	Discovers redundancy group and DF election.
Route Type 5: IP Prefix Route	The IP prefixes are advertised independently of the MAC-advertised routes. With EVPN IRB, host route /32 is advertised using RT-2 and subnet /24 is advertised using RT-5.	Advertises IP prefixes.

Note With EVPN IRB, host route /32 are advertised using RT-2 and subnet /24 are advertised using RT-5.

EVPN modes

EVPN modes address modern network needs like scalability, multi-tenancy, redundancy, and workload mobility, providing a unified solution for Layer 2 and Layer 3 services. These modes provide an efficient solution for Layer 2 and Layer 3 services, ensuring operational simplicity, high performance, and flexibility for data centers, service providers, and cloud environments.

These EVPN modes are supported:

- Single-homing—enables you to connect a customer edge (CE) device to one provider edge (PE) device.
- Multi-homing—enables you to connect a customer edge (CE) device to more than one provider edge (PE) device. Multi-homing ensures redundant connectivity, so that there is no traffic disruption when there is a network failure. These multi-homing modes are supported:
  - All-active
  - Single-active
  - Single-flow active
  - Port-active

How EVPN works

Summary

The key components involved in the process are:

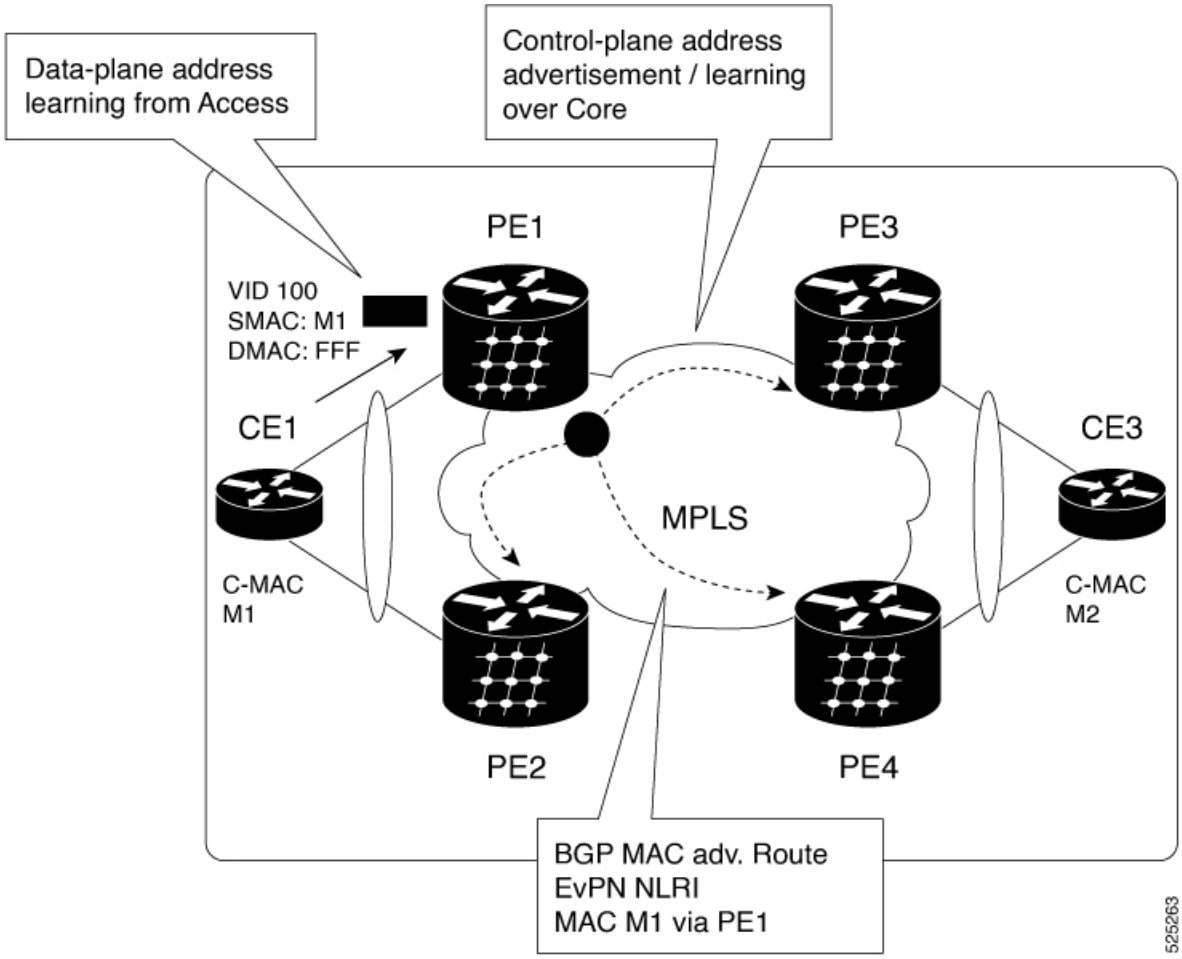
- VPN membership—PEs exchange EVPN routes to discover remote PE members of an EVI, build flood lists for multicast ingress replication, and exchange BUM and unicast labels during MAC learning.

- Ethernet segment reachability—PEs perform auto-discovery of remote PEs and determine redundancy modes in multihoming scenarios. During segment failures, routes are withdrawn to trigger fast convergence via MAC mass withdrawal.
- Redundancy group membership—PEs discover and elect a Designated Forwarder (DF) for forwarding BUM traffic in multihoming scenarios.

Note Starting with the Cisco IOS XR Release 24.2.1, the unknown unicast suppression (**unknown-unicast-suppression**) operation under a given EVI is not supported on the Cisco 8000 series routers.

Workflow

Figure 1: How EVPN works



The EVPN operation involves these stages:

1. Startup—PEs exchange EVPN routes for VPN membership, segment reachability, and redundancy group discovery.

2. MAC Advertisement—unknown unicast or BUM MAC addresses are advertised as route type 2.
3. Multihoming—route types 1, 3, and 4 are exchanged to discover redundancy modes (single-active or all-active) and elect a DF.

EVPN can be set up to operate in either single-homing or multihoming modes, providing various levels of redundancy and connectivity in network configurations.

Single-homing mode stages:

The EVPN single-homing mode involves these stages:

1. Route type 3 advertisement—PEs advertise route type 3 upon enabling EVPN.
2. Discovery—PEs discover all other member PEs within the EVPN instance.
3. BUM MAC handling—unknown unicast or BUM MACs are advertised as route type 2.
4. MAC route distribution—route type 2 advertises MAC routes to other PEs.

Multihoming mode stages:

The EVPN multihoming mode involves these stages:

1. Route types 1, 3, and 4 advertisement—PEs advertise these routes to discover redundancy modes and elect a DF
   a. Route type 1 usage—automatically discovers PEs hosting the same CE and enables fast rerouting during link failures.
   b. Route type 4 usage—facilitates DF election for BUM traffic.
2. Customer traffic handling—distributes MAC reachability information through route type 2.
3. Traffic delivery—remote PEs use MPLS labels to deliver traffic to the specified MAC address.

This process ensures efficient and reliable EVPN operation in both single-homing and multihoming scenarios, enabling robust connectivity, redundancy, and seamless traffic handling for customer networks.

Behavior change due to ESI label assignment

To adhere to RFC 7432 recommendations, the encoding or decoding of MPLS label is modified for extended community. Earlier, the lower 20 bits of extended community were used to encode the split-horizon group (SHG) label. The SHG label encoding now uses a higher 20 bits of extended community.

According to this change, routers running old and new software release versions in the same ethernet segment decodeextended communities differently. This change causes inconsistent SHG labels on peering EVPN PE routers. Almost always, the router drops BUM packets with incorrect SHG labels. However, remote PE may accept such packets and forward them to CE in certain conditions, potentially causing a loop. One such instance is when the label incorrectly reads NULL.

To overcome this problem, Cisco recommends you to:

- Minimize the time both PEs are running different software release versions.
- Before upgrading to a new release, isolate the upgraded node and shut down the corresponding AC bundle.

- After upgrading both the PEs to the same release, you can bring both into service.

Similar recommendations apply to peering PEs with different vendors with SHG label assignment that does not adhere to RFC 7432.

EVPN timers

EVPN timers ensure efficient and reliable exchange of MAC and IP information by controlling route propagation and convergence. Their significance lies in optimizing network stability, reducing downtime, and improving responsiveness during topology changes.

This table shows various EVPN timers:

Table 1: EVPN timers

Timer	Range	Default Value	Trigger	Applicability	Action	Sequence
startup-cost-in	30-86400	disabled	node recovered*	Single-Homed, All-Active, Single-Active	Postpone EVPN startup procedure and Hold AC link(s) down to prevent CE to PE forwarding. Startup-cost-in timer allows PE to set core protocols first.	1
recovery	20-3600s	30s	node recovered, interface recovered **	Single-Homed***, Single-Active	Postpone EVPN Startup procedure. Recovery timer allows PE to set access protocols (STP) before reachability towards EVPN core is advertised.	2

Timer	Range	Default Value	Trigger	Applicability	Action	Sequence
peering	0-3600s	3s	node recovered, interface recovered	All-Active, Single-Active	Starts after sending EVPN RT4 to postpone rest of EVPN startup procedure. Peering timer allows remote PE (multihoming AC with same ESI) to process RT4 before DF election will happen.	3

Note

- The timers are available in EVPN global configuration mode and in EVPN interface sub-configuration mode.
- Startup-cost-in is available in EVPN global configuration mode only.
- Timers are triggered in sequence (if applicable).
- Cost-out in EVPN global configuration mode brings down AC link(s) to prepare node for reload or software upgrade.

* indicates all required software components are loaded.

** indicates link status is up.

*** you can change the recovery timer on Single-Homed AC if you do not expect any STP protocol convergence on connected CE.