Configure Layer3 EVPN over SR MPLS [eBGP] in Nexus 9300

Contents

Introduction

Prerequisites

Requirements

Components Used

Background Information

MPLS L3VPN Recap

Overview of EVPN with L3VPN (MPLS SR)

Network Diagram

High-Level Configuration

Verify

Troubleshoot

Introduction

This document describes how to deploy L3 Ethernet VPN (EVPN) over Segment Routing (SR) Multiprotocol Label Switching on Nexus 9300 with external BGP.

Prerequisites

Requirements

Cisco recommends that you have knowledge of these topics:

- Border Gateway Protocol (BGP)
- L3VPN
- EVPN
- SR

Components Used

The information in this document is based on these software and hardware versions:

- SPINE Hardware 9336C-FX that runs Release 10.2(2)
- LEAF Hardware 93240YC-FX2 that runs Release 10.2(2)
- CLIENT 92160YC-X (Host-1), Catlyst-3850 (Host-2)

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.

Background Information

MPLS L3VPN Recap

A VPN is:

- An IP-based network that delivers private network services over a public infrastructure.
- A set of sites that are allowed to communicate with each other privately over the Internet or other public or private networks.

Conventional VPNs are created by the configuration of a full mesh of tunnels or permanent virtual circuits (PVCs) to all sites in a VPN. This type of VPN is not easy to maintain or expand, as the addition of a new site requires a change to each edge device in the VPN.

MPLS-based VPNs are created in Layer 3 and are based on the peer model. The peer model enables the service provider and the CE to exchange Layer 3 routing information. The service provider relays the data between the CE sites without CE involvement.

MPLS VPNs are easier to manage and expand than conventional VPNs. When a new site is added to an MPLS VPN, only the edge router of the service provider that provides services to the customer site needs to be updated.

These are the components of the MPLS VPN:

- Provider (P) router- Router in the core of the provider network. PE routers run MPLS switching
 and do not attach VPN labels to routed packets. VPN labels are used to direct data packets to
 the correct private network or CE edge router.
- Provider Edge (PE) router- Router that attaches the VPN label to incoming packets based on the interface or subinterface on which they are received, and also attaches the MPLS core labels. A PE router attaches directly to a router.
- Customer (C) router- Router in the Internet service provider (ISP) or enterprise network.
- Customer Edge (CE) router- Edge router on the network of the ISP that connects to the PE router on the network. A CE router must interface with a PE router.

Overview of EVPN with L3VPN (MPLS SR)

Data Center (DC) deployments have adopted Virtual Extensible LAN (VXLAN) EVPN or MPLS EVPN for its benefits such as EVPN control-plane learning, multitenancy, seamless mobility, redundancy, and easier POD additions. Similarly, the CORE is either a Label Distribution Protocol (LDP)-based MPLS L3VPN network or a transition from the traditional MPLS L3VPN LDP-based underlay to a more sophisticated solution like SR.

SR is adopted for its benefits such as:

- Unified interior gateway protocol (IGP) and MPLS control planes
- Simpler traffic engineering methods

- Easier configuration
- Software-Defined Networking (SDN) adoption

EVPN (RFC 7432) is BGP MPLS-based solution that has been used for next-generation Ethernet services in a virtualized data center network. It uses several blocks such as Route.

Distinguisher (RD), Route Target (RT), and Virtual Routing and Forwarding (VRF) from MPLS technologies that exist.

L3 EVPN over SR which was introduced in NXOS 7.0(3)I6(1) release uses the EVPN Type-5 route with MPLS encapsulation. It offers Multi-tenant, Scalability, and High Performance for evolved data center services.

Note: In DC, the data plane can be VXLAN or MPLS.

Traditional MPLS L3 VPN

MPLS L3 VPN over SR

Main build blocks: RD, RT, and VRF

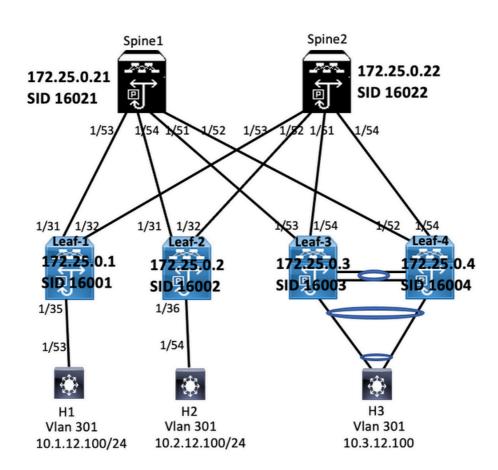
Main build blocks: RD, RT, and VRF

Underlay Layer for Transport: IGP, LDP, and RSVP-TE Underlay Layer for Transport: IGP/BGP-LU and S

Overlay Layer for Service: VPNv4 and VPNv6

Overlay Layer for Service: EVPN

Network Diagram



High-Level Configuration

1. Install Features

- 2. Configure IP address Underlay
- 3. Configure IGP/MP BGP
- 4. Configure VLAN and EVPN Overlay
- 5. Configure e-BGP between Hosts and LEAFs

install feature-set mpls feature-set mpls feature bgp feature mpls segment-routing feature mpls evpn feature interface-vlan nt-routing traffic-engineering vlan 1,301-310 global-block 16000 24000 global-block 16000 24000 connected-prefix-sid-map address-family ipv4 172.25.0.1/32 absolute 16001 ip prefix-list node-sid-loopback seq 10 permit 172.25.0.1/32 ip as-path access-list LOCALLY-ORIGINATE seq 1 permit "^5534" ip as-path access-list LOCALLY-ORIGINATE seq 2 permit "^5" crute-map NODE-SID-MED permit 10 match ip address prefix-list node-sid-loopback set metric 100 route-map NODE-SID-MED permit 20 route-map SET_NH permit 5

match community MATCH-65534:65534.
set ip next-hop unchanged
route-map SET_NH permit 10
match as-path LOCALLY-ORIGINATE
set ip next-hop 172.25.0.1

vrf context VPN-A of auto address-family ipv4 unicast route-target import 301:301 route-target import 301:301 evpn route-target export 301:301 evpn route-target export 301:301 evpn vrf context VPN-B

ort context virves or disute address-family ipv4 unicast route-target import 302:302 route-target import 302:302 evpn route-target export 302:302 evpn route-target export 302:302 evpn

interface Vlan301 ip access-group deny-to-core_ra in vrf member VPN-A no ip redirects ip address 10.1.12.1/24 ip arp timeout 720

interface Vlan302 ip access-group deny-to-core_ra in vrf member VPN-B no ip redirects ip address 10.1.13.1/24 ip directed-broadca ip arp timeout 720

interface Ethernet1/31 ed to spine1 - 1/53 - 192.168.1.10

mtu 9216 mtu 9216 logging event port link-status no ip redirects ip address 192.168.1.9/30 ip arp timeout 14400 mpls ip forwarding

interface Ethernet1/32 description connected to spine2 - 1/53 - 192.168.1.14 mtu 9216 logging event port link-status no ip redirects ip address 192.168.1.13/30 ip art himografi ip arp timeout 14400 mpls ip forwarding no shutdown

interface Ethernet1/35 switchport switchport mode trunk switchport trunk allowed vlan 301-310 no shutdown

interface loopback0 ip address 172.25.0.1/32

router bgp 65534 router-id 172.25.0.1 disable-policy-batching bestpath as-path multipath-relat bestpath med missing-as-worst log-neighbor-changes event-history detail size large nexthop suppress-default-res address-family ipv4 unicast network 172.25.0.1/32 maximum-paths 4 maximum-paths ibgp 4 allocate-label route-mag allocate-label route-map node-sid-address-family ipv4 labeled-unicast prefix-priority high address-family I2vpn evpn template peer EBGP-SPINE remote-as 64087 remote-as 6-4087 description EBGP-PEERING-to-AGG address-family ipv4 unicast allowas-in 1 send-community

no advertise local-labeled-route soft-reconfiguration inbound address-family jovd labeled-unicast allowas-in 1 send-community send-community send-community send-community extended route-map MODE-SID-MED out soft-reconfiguration inbound alway address-family 12 ypn evpn allowas-in 1 send-community send-community send-community extended filter-list LOCALIY-ORBINATE out route-map SET, MH out niter-list EUCALT-ONG inter-list EUCALT-ONG encapsulation mpls neighbor 192.168.1.10 inherit peer EBGP-SPINE neighbor 192.168.1.14 inherit peer EBGP-SPINE

map NODE-SID-MED out

vrf VPN-A wir vin-A. bestpath as-path multipath-relax allocate-index 2001 address-family ipv4 unicast network 10.1.12.0/24 advertise I2vpn ev maximum-paths 4 vrf VPN-B bestpath as-path multipath-allocate-index 2002 address-family ipv4 unicast network 10.1.13.0/24

install feature-set mpls

feature-set mpls feature-set mpis
feature bgp
feature mpis segment-routing
feature mpis evpn
feature interface-vlan
feature lacp
feature mpis oam
feature mpis oam ent-routing traffic-engineering vlan 1,301-310 global-block 16000 24000 global-block 1,6000 24000 connected-prefix-sid-map address-family lpv4 172.25.0,2/32 absolute 1,6002 ip perfix-list node-sid-loopback seq 10 permit 172.25.0,2/32 ip as-path access-list LOCALLY-ORIGINATE seq 2 permit "^65534" ip as-path access-list LOCALLY-ORIGINATE seq 2 permit "^6" route-map NODE-SID-MED permit 10 match ip address prefix-list node-sid-loopback set metric 100 route-map NODE-SID-MED permit 20 route-map SET_NH permit 5

match community MATCH-65534:65534
set ip next-hop unchanged
route-map SET_NH permit 10
match as-path LOCALLY-ORIGINATE
set ip next-hop 172.25.0.2
vrf context VPN-A address-family ipv4 unicast address-family ipv4 unicast route-target import 301:301 route-target import 301:301 evpn route-target export 301:301 route-target export 301:301 evpn vf context VPN-B rd auto address-family ipv4 unicast route-target import 302:302 route-target import 302:302 evpn route-target export 302:302 route-target export 302:302 evpn

interface Vlan301 ip access-group deny-to-core_ra in vrf member VPN-A no ip redirects ip address 10.2.12.1/24 ip directed-broadcast ip-dir-bcast ip arp timeout 720

interface Vlan302 no shutdown ip access-group deny-to-core_ra in vrf member VPN-B no ip redirects ip address 10.2.13.1/24 ip directed-broadcast ip-dir-bcast ip arp timeout 720

interface Ethernet1/3 switchport switchport mode trunk switchport trunk allowed vlan 301-310

logging event port link-status no ip redirects ip address 192.168.2.9/30

logging event port link-status no ip redirects ip address 192.168.2.13/30

ip arp timeout 14400 mpls ip forwarding interface Ethernet1/32

ip arp timeout 14400 mpls ip forwarding interface Ethernet1/36

switchport mode trunk

switchport trunk allowed vlan 301-310 ip address 172.25.0.2/32

mtu 9216

description connected to spine1 - 1/54 - 192.168.2.10 mtu 9216

description connected to Spine2 - 1/52 - 192,168,2,14

router bgp 65534 router-id 172.25.0.2 disable-policy-batching bestpath as-path multipath-relax bestpath as-path multipath-relabestpath med missing-as-worst log-neighbor-changes event-history detail size large nexthop suppress-default-resol address-family jpv4 unicast address-family ipv4 unica network 172.25.0.2/32 maximum-paths 4 maximum-paths ibgp 4 maximum-paths logp 4
allocate-label route-map node-sid-label
address-family joy4 labeled-unicast
perfix-priority high
address-family I2vpn evpn
template peer EBGP-SPINE
remote-as 64082F RRING-to-AGG
address-family losd unicast address-family ipv4 unicast allowas-in 1 send-community extended route-map NODE-SID-MED out

no advertise local-labeled-route soft-reconfiguration inbou address-family ipv4 labeled allowas-in 1 send-community send-community extended route-map NODE-SID-MED out

soft-reconfiguration inbound always address-family I2vpn evpn address-ramily (2vpn evpn allowas-in 1 send-community send-community extended filter-list LOCALLY-ORIGINATE out route-map SET_NH out encapsulation mpls neighbor 192.168.2.10 inherit peer EBGP-SPINE neighbor 192.168.2.14 inherit peer EBGP-SPINE

vrf VPN-A

rrt VPN-A
bestpath as-path multipath-relax
allocate-index 2001
address-family igw4 unicast
network 10.2.12.0/24
advertise I2vpn evpn
maximum-paths 4
vrf VPN-B
bestpath as-path multipath-relax bestpath as-path multipath-relax allocate-index 2002 address-family ipv4 unicast network 10.2.13.0/24

install feature-set mpls feature-set mpls feature bgp feature mpls segment-routing feature mpls evpn feature interface-vlan feature lacp feature mpls oam feature mpls segment-routing traffic-engineering vlan 1 segment-routing mpls global-block 16000 24000 connected-prefix-sid-map address-family ipv4 172.25.0.21/32 absolute 16021 ip prefix-list NH-RESTRICT seq 5 permit 0.0.0.0/0 ip prefix-list node-sid-loopback seq 5 permit 172.25.0.21/32 route-map NH-RESTRICT deny 10 match in address prefix-list NH-RESTRICT route-map NH-RESTRICT permit 20 route-map NH UNCHG permit 10 set ip next-hop unchanged

Spineterface Configuration

interface Ethernet1/53 description connected to Leaf1 - 1/31 - 192.168.1.9 mtu 9216 logging event port link-status no ip redirects ip address 192.168.1.10/30 ip arp timeout 14400 mpls ip forwarding

interface Ethernet1/54 description connected to Leaf2- 1/31 - 192.168.2.9 mtu 9216

logging event port link-status no ip redirects ip address 192.168.2.10/30 ip arp timeout 14400 mpls ip forwarding no shutdown

no shutdown

interface loopback0 ip address 172.25.0.21/32 no shutdown BGP/EVPN Configuration

router bgp 64087 router-id 172.25.0.21 bestpath as-path multipath-relax bestpath med missing-as-worst log-neighbor-changes nexthop suppress-default-reso address-family joy4 unicast network 172.25.0.21/32 maximum-paths 4 nexthop route-map NH-RESTRICT allocate-label route-map node-sid-label address-family ipv4 labeled-unicast prefix-priority high address-family I2vpn evpn retain route-target all template peer EBGP-ACCESS remote-as 65534 description EBGP-PEERING-to-ACCESS address-family ipv4 unicast disable-peer-as-check send-community send-community extended default-originate no advertise local-labeled-route

soft-reconfiguration inbound address-family ipv4 labeled-unicast disable-peer-as-check send-community send-community send-community extended soft-reconfiguration inbound address-family IZvpn evpn disable-peer-as-check send-community send-community send-community extended route-map NH_UNCHG out encapsulation mpls neighbor 192.168.1.9 inherit peer EBGP-ACCESS neighbor 192.168.2.9 inherit peer EBGP-ACCESS

Spine-2

description connected to Leaf2 - 1/31 - 192.168.2.13

interface Ethernet1/52

mtu 9216

hling Feature Interface Configurati

install feature-set mpls feature-set mpls feature bgp feature mpls segment-routing feature mpls evpn feature interface-vlan feature lacp feature mpls oam feature mpls segment-routing traffic-engineering segment-routing mpls global-block 16000 24000 connected-prefix-sid-map address-family ipv4 172.25.0.22/32 absolute 16021 ip prefix-list NH-RESTRICT seq 5 permit 0.0.0.0/0 ip prefix-list node-sid-loopback seq 5 permit 172.25.0.22/32 route-map NH-RESTRICT deny 10 match ip address prefix-list NH-RESTRICT route-map NH-RESTRICT permit 20 route-map NH UNCHG permit 10

set ip next-hop unchanged

logging event port link-status
no ip redirects
ip address 192.168.2.14/30
ip arp timeout 14400
mpls ip forwarding
no shutdown

interface Ethernet1/53
description connected to Leaf2- 1/32 - 192.168.1.13
mtu 92.16
logging event port link-status
no ip redirects
ip address 192.168.1.14/30
ip arp timeout 14400
mpls ip forwarding
no shutdown

interface loopback0 ip address 172.25.0.22/32

BGP/EVPN Configuration

router bgp 64087 router-id 172.25.0.22 bestpath as-path multipath-relax bestpath med missing-as-worst log-neighbor-changes nexthop suppress-default-resolution address-family ipv4 unicast network 172.25.0.22/32 maximum-paths 4 nexthop route-map NH-RESTRICT allocate-label route-map node-sid-label address-family ipv4 labeled-unicast prefix-priority high address-family I2vpn evpn retain route-target all template peer EBGP-ACCESS remote-as 65534 description EBGP-PEERING-to-ACCESS address-family ipv4 unicast disable-peer-as-check send-community send-community extended

soft-reconfiguration inbound address-family ipv4 labeled-unicast disable-peer-as-check send-community send-community extended soft-reconfiguration inbound address-family I2vpn evpn disable-peer-as-check send-community send-community extended route-map NH UNCHG out encapsulation mpls neighbor 192.168.1.13 inherit peer EBGP-ACCESS neighbor 192.168.2.13 inherit peer EBGP-ACCESS

default-originate no advertise local-labeled-route install feature-set mpls feature mpls interface Ethernet1/53

switchport

switchport mode trunk

switchport trunk allowed vlan 301-310

no shut

interface vlan 301 no shutdown no ip redirects ip address 10.1.12.100/24 ip directed-broadcast ip-dir-bcast

ip arp timeout 720

install feature-set mpls feature mpls interface Ethernet1/54 switchport

switchport mode trunk

switchport trunk allowed vlan 301-310

no shut

interface vlan 301 no shutdown no ip redirects

ip address 10.2.12.100/24 ip directed-broadcast ip-dir-bcast

ip arp timeout 720

Verify

Use this section in order to confirm that your configuration works properly.

H1(config)# show ip int brief

IP Interface Status for VRF "default"(1)
Interface IP Address Interface Status

Vlan301 10.1.12.100 protocol-up/link-up/admin-up

H1(config)# ping 10.2.12.100 PING 10.2.12.100 (10.2.12.100): 56 data bytes

64 bytes from 10.2.12.100: icmp_seq=0 ttl=251 time=0.994 ms 64 bytes from 10.2.12.100: icmp_seq=1 ttl=251 time=0.586 ms 64 bytes from 10.2.12.100: icmp_seq=2 ttl=251 time=0.677 ms 64 bytes from 10.2.12.100: icmp_seq=3 ttl=251 time=0.615 ms 64 bytes from 10.2.12.100: icmp_seq=4 ttl=251 time=0.597 ms

--- 10.2.12.100 ping statistics ---

5 packets transmitted, 5 packets received, 0.00% packet loss round-trip min/avg/max = 0.586/0.693/0.994 ms

H2(config)# show ip int brief

IP Interface Status for VRF "default"(1)
Interface IP Address Interface Status

Vlan301 10.2.12.100 protocol-up/link-up/admin-up

H2(config)# ping 10.1.12.100

PING 10.1.12.100 (10.1.12.100): 56 data bytes

64 bytes from 10.1.12.100: icmp_seq=0 ttl=251 time=1.043 ms 64 bytes from 10.1.12.100: icmp_seq=1 ttl=251 time=1.933 ms 64 bytes from 10.1.12.100: icmp_seq=2 ttl=251 time=0.865 ms 64 bytes from 10.1.12.100: icmp_seq=3 ttl=251 time=0.668 ms 64 bytes from 10.1.12.100: icmp_seq=4 ttl=251 time=0.713 ms

--- 10.1.12.100 ping statistics ---

5 packets transmitted, 5 packets received, 0.00% packet loss round-trip min/avg/max = 0.668/1.044/1.933 ms

Troubleshoot

This section provides information you can use to troubleshoot your configuration.

spine1(config-router-af)# show mpls switching

Legend:

(P)=Protected, (F)=FRR active, (*)=more labels in stack.

IPV4:

In-Label Out-Label FEC name Out-Interface Next-Hop VRF default

 16001
 Pop Label 172.25.0.1/32
 Eth1/53
 10.1.1.9

 16002
 Pop Label 172.25.0.2/32
 Eth1/54
 10.2.1.9

In-Label VRF 492287 default

Block Label-Range 1 16000 - 24000 spine1(config-router-af)# show bgp I2vpn evpn BGP routing table information for VRF default, address family L2VPN EVPN

BGP table version is 17, Local Router ID is 172.25.0.21

Network Next Hop Metric LocPrf Weight Path Route Distinguisher: 172.25.0.1:3

*>e[5]:[0]:[0]:[24]:[12.1.12.0]/224

172.25.0.1 4294967295 0 65534 i

Route Distinguisher: 172.25.0.1:4 *>e[5]:[0]:[0]:[24]:[12.1.13.0]/224

172.25.0.1 4294967295 0 65534 i

Route Distinguisher: 172.25.0.2:3
*>e[5]:[0]:[0]:[24]:[10.2.12.0]/224

172.25.0.2 4294967295 0 65534 i

Route Distinguisher: 172.25.0.2:4
*>e[5]:[0]:[0]:[24]:[10.2.13.0]/224

172.25.0.2 4294967295 0 65534 i

H18 show ip int brie

IP interface Status for VRF "default" (1)
Interface

IP Address Interface Status
Vlan301

10.1.12.100

protocol-up/link-up/admin-up

H18 show mac address-table
Legend: primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age -seconds since last seen. - primary entry using vPC Peet-Link,
(T) - True, (F) - Faile, C - ControlPlane MAC, - visan
VLAN MACA dones - Type age Secure NTP Ports

* 301 0000.0000.1111 dynamic 0 F F Po30

* 301 000e.3d/27.86Ed dynamic 0 F F Po30
G - Otoe.abd/27.86E static - F F sup-eth/LR)
G 301 00ea.bd/27.62ES static - F F sup-eth/LR)

Route Distinguisher: 172.25.0.3.37164

**e[2](0)(0)(48)[400:abd77.86ef](0)(0.0.0)/216

172.25.0.3 429467255

**e[2](0)(0)(48)[(00eabd7.86ef](32)[10.112.00)/272

172.25.0.3 4294967255

0.655341

**e[3](0)[2](2](127.25.0.3)/88

172.25.0.3 4294967295

0.655341

spine-1# show ip int bri

| Pinterface Status for VRE "default" (1) | Interface | P Address | Interface Status | Interface Status | Interface | P Address | Interface Status | Interface | I

BGP routing table information for VRF default, address family L2VPN

EVPN
BGP table version is 188, Local Router ID is 172.25.0.22
Status: s-suppressed, x-deleted, 5-stale, d-dampened, h-history, *-valid,

Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, I-local, a-aggregate, r-redist, I-

Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 -

Route Distinguisher: 172.25.0.2:37164
*>e[2] {0} {0} {148} {0.0ea.bd27.628} {10} {1.0.0.0} {216}
\$12.25.0.15 429495295 0.65534 {1}
*>e[2] {0} {0} {48} {0.0ea.bd27.628} {32} {1.1.12.100} {272}

Route Distinguisher: 172.25.0.337164
**e[2](0](0](48)[00ea bd27.86ef](0](0.0.0)[216
**e[2](0](0)[48][00ea bd27.86ef](0][10.0.0][216
**e[2](0](0](48)[00ea bd27.86ef][12][10.112.200][727
**e[3][0](2][17.25.0.3 4294672795
**e[3][0](2][17.25.0.3 4294672795
0655341

| P Interface Status for VRF "default"(1) | Interface | P Address | Interface Status | LoO | 172.25.0.22 | protect—plifts-up/admin-up | LoO | 172.25.0.22 | protect—plifts-up/admin-up | LoO | L