Deploy Layer3 EVPN over Segment Routing MPLS [Ospf / iBGP] in Nexus 3000

Contents

Introduction

This document describes how to deploy/configure Layer3 EVPN over Segment Routing MPLS on Nexus 3000 products.

Prerequisites

Requirements

Cisco recommends that you have knowledge of these topics:

- Border Gateway Protocol (BGP)
- L3VPN
- EVPN
- Segment Routing

Components Used

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

- SPINE Hardware - N9K-C92160YC-X running with 9.2(3)
- LEAF Hardware - N3K-C31108PC-V running with 9.3(3)

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, make sure that you understand the potential impact of any command.

Background Information

MPLS L3VPN Recap

A VPN is:

- An IP-based network delivering 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 configuring 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 adding a new site requires changing 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 customer to exchange Layer 3 routing information. The service provider relays the data between the customer sites without customer 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 customer edge router.

- 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 CE 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 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 transitioning from the traditional MPLS L3VPN LDP-based underlay to a more sophisticated solution like Segment Routing (SR).

Segment Routing is adopted for its benefits such as:

- Unified IGP and MPLS control planes
- Simpler traffic engineering methods

- Easier configuration
- 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 building blocks such as RD, RT, and VRF from existing MPLS technologies.

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.

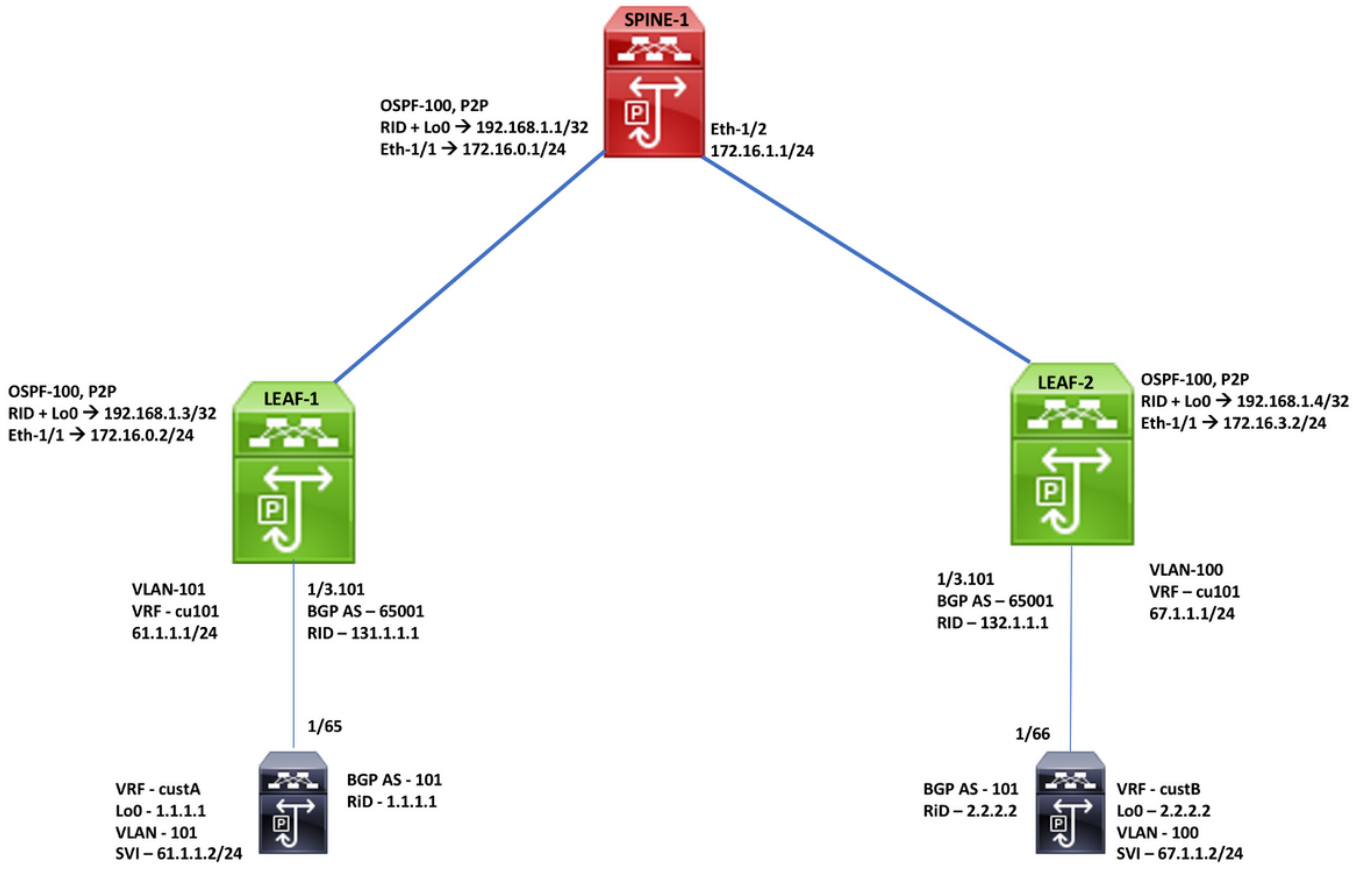
> **Note**: In DC, the data plane can be VXLAN or MPLS.

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 SR-
Overlay Layer for Service: VPNv4 and VPNv6	Overlay Layer for Service: EVPN

Limitations

L2-EVPN is not supported in **Nexus C31108PC-V**, N9K Cloud-Scale is suitable for any SR deployment due to Scale considerations.

Network Diagram



Configuration

High-Level Configuration

1. Install Features
2. Configure IP address -Underlay
3. Configure IGP -OSPF
4. Configure MP-BGP
5. Configure VLAN and EVPN Overlay
6. Configure e-BGP between Hosts & LEAFs

SPINE-1 Configuration

Enabling Features, Label-Range, Route-map, Label-Index

```
feature-set mpls
feature ospf
feature bgp
feature mpls segment-routing
feature mpls evpn
feature interface-vlan
feature mpls oam


mpls label range 5000 45000


segment-routing mpls
  global-block 16000 25000
  connected-prefix-sid-map
    address-family ipv4
      192.168.1.1/32 index 211


route-map label-index-spine1 permit 10
  set label-index 211
```

OSPF Configuration

```
interface Ethernet1/1
  ip address 172.16.0.1/24
  ip ospf network point-to-point
  ip router ospf 100 area 0.0.0.0
  mpls ip forwarding
  no shutdown


interface Ethernet1/2
  ip address 172.16.1.1/24
  ip ospf network point-to-point
  ip router ospf 100 area 0.0.0.0
  mpls ip forwarding
  no shutdown


interface loopback0
  ip address 192.168.1.1/32
  ip router ospf 100 area 0.0.0.0


router ospf 100
  segment-routing mpls
  router-id 192.168.1.1
```

BGP/EVPN Configuration

```
router bgp 65001
  router-id 192.168.1.1
  address-family ipv4 unicast
    network 192.168.1.1/32 route-map label-index-spine1
    allocate-label all
  address-family ipv4 labeled-unicast
  address-family l2vpn evpn
  template peer EVPN
    remote-as 65001
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
      encapsulation mpls
  template peer Labeled-unicast
    remote-as 65001
    address-family ipv4 labeled-unicast
      send-community extended
      route-reflector-client
      next-hop-self
      soft-reconfiguration inbound always
  neighbor 172.16.0.2
    inherit peer Labeled-unicast
  neighbor 172.16.1.2
    inherit peer Labeled-unicast
  neighbor 192.168.1.3
    inherit peer EVPN
  neighbor 192.168.1.4
    inherit peer EVPN
```

LEAF-1 Configuration

Enabling Features, Label-Range, Route-map, Label-Index

```
feature-set mpls
feature ospf
feature bgp
feature mpls segment-routing
feature mpls evpn
feature interface-vlan
feature lacp
feature mpls oam


mpls label range 5000 450000


segment-routing
  mpls
    global-block 16000 25000
    connected-prefix-sid-map
      address-family ipv4
        192.168.1.3/32 index 311


route-map label-index-leaf-1 permit 10
  set label-index 311
```

OSPF, VRF Configuration

```
interface Ethernet1/1
  no switchport
  ip address 172.16.0.2/24
  ip ospf network point-to-point
  ip router ospf 100 area 0.0.0.0
  mpls ip forwarding
  no shutdown

interface loopback0
  ip address 192.168.1.3/32
  ip router ospf 100 area 0.0.0.0

router ospf 100
  segment-routing mpls
  router-id 192.168.1.3


interface Ethernet1/3
  no switchport
  no shutdown

interface Ethernet1/3.101
  encapsulation dot1q 101
  vrf member cu101
  ip address 61.1.1.1/24
  ip ospf network point-to-point
  ip router ospf 200 area 0.0.0.0
  no shutdown

vrf context cu101
  rd auto
  address-family ipv4 unicast
    route-target import 1:101
    route-target import 1:101 evpn
```

BGP/EVPN Configuration

```
router bgp 65001
  router-id 192.168.1.3
  address-family ipv4 unicast
    network 192.168.1.3/32 route-map label-index-leaf-1
    allocate-label all
  address-family ipv4 labeled-unicast
  address-family l2vpn evpn
  template peer EVPN
    remote-as 65001
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      encapsulation mpls
  template peer Labeled-unicast
    remote-as 65001
    address-family ipv4 labeled-unicast
      send-community extended
      soft-reconfiguration inbound always
  template peer cu1
    address-family ipv4 unicast
      as-override
      send-community
      soft-reconfiguration inbound always
  neighbor 172.16.0.1
    inherit peer Labeled-unicast
  neighbor 192.168.1.1
    inherit peer EVPN

  vrf cu101
    router-id 131.1.1.1
    address-family ipv4 unicast
      advertise l2vpn evpn
    neighbor 61.1.1.2
      inherit peer cu1
      remote-as 101
```

LEAF-2 Configuration

Enabling Features, Label-Range, Route-map, Label-Index

```
feature-set mpls
feature ospf
feature bgp
feature mpls segment-routing
feature mpls evpn
feature interface-vlan
feature mpls oam


mpls label range 5000 450000


segment-routing
  mpls
    global-block 16000 25000
    connected-prefix-sid-map
      address-family ipv4
        192.168.1.4/32 index 321


route-map label-index-Leaf2 permit 10
  set label-index 321
```

OSPF, VRF Configuration

```
interface Ethernet1/1
  no switchport
  ip address 172.16.1.2/24
  ip ospf network point-to-point
  ip router ospf 100 area 0.0.0.0
  mpls ip forwarding
  no shutdown

interface loopback0
  ip address 192.168.1.4/32
  ip router ospf 100 area 0.0.0.0

router ospf 100
  segment-routing mpls
  router-id 192.168.1.4


interface Ethernet1/3
  no switchport
  no shutdown

interface Ethernet1/3.101
  encapsulation dot1q 100
  vrf member cu101
  ip address 67.1.1.1/24
  no shutdown

vrf context cu101
  rd auto
  address-family ipv4 unicast
    route-target import 1:101
    route-target import 1:101 evpn
```

BGP/EVPN Configuration

```
router bgp 65001
  router-id 192.168.1.4
  address-family ipv4 unicast
    network 192.168.1.4/32 route-map label-index-Leaf2
    allocate-label all
  address-family ipv4 labeled-unicast
  address-family l2vpn evpn
  template peer EVPN
    remote-as 65001
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      encapsulation mpls
  template peer Labeled-unicast
    remote-as 65001
    address-family ipv4 labeled-unicast
      send-community extended
      soft-reconfiguration inbound always
  template peer cu1
    address-family ipv4 unicast
      as-override
      send-community
      soft-reconfiguration inbound always

  neighbor 172.16.1.1
    inherit peer Labeled-unicast
  neighbor 192.168.1.1
    inherit peer EVPN


  vrf cu101
    router-id 132.1.1.1
    address-family ipv4 unicast
      advertise l2vpn evpn
    neighbor 67.1.1.2
      inherit peer cu1
      remote-as 101
```

END-Host Configuration

Enabling Features, , Route-map, VRF-A Configuration

```
feature bgp
feature interface-vlan

vlan 1,100-101

route-map twist permit 10
  set metric 10


vrf context custA
  rd 101:1
  address-family ipv4 unicast

interface loopback0
  vrf member custA
  ip address 1.1.1.1/32

interface Vlan101
  no shutdown
  vrf member custA
  ip address 61.1.1.2/24

interface Ethernet1/65
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 101
  no shutdown
```

BGP Configuration

```
router bgp 101
  vrf custA
    router-id 1.1.1.1
    address-family ipv4 unicast
      network 1.1.1.1/32
      redistribute direct route-map twist
    neighbor 61.1.1.1
      remote-as 65001
      address-family ipv4 unicast
        send-community
        send-community extended

  vrf custB
    router-id 2.2.2.2
    address-family ipv4 unicast
      network 2.2.2.2/32
      redistribute direct route-map twist
    neighbor 67.1.1.1
      remote-as 65001
      address-family ipv4 unicast
        send-community
        send-community extended
        soft-reconfiguration inbound
```

VRF-B Configuration

```
vrf context custB
  rd 101:2
  address-family ipv4 unicast

interface loopback1
  vrf member custB
  ip address 2.2.2.2/32

interface Vlan100
  no shutdown
  vrf member custB
  ip address 67.1.1.2/24


interface Ethernet1/66
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 100
  no shutdown
```

Verify

```
Leaf2(config)# show bgp l2vpn evpn
BGP routing table information for VRF default, address family L2VPN EVPN
BGP table version is 14, Local Router ID is 192.168.1.4
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 - best2

   Network            Next Hop            Metric     LocPrf     Weight Path
Route Distinguisher: 192.168.1.3:4
*>i[5]:[0]:[0]:[24]:[61.1.1.0]/224
                      192.168.1.3              10        100          0 101 ?
*>i[5]:[0]:[0]:[32]:[1.1.1.1]/224
                      192.168.1.3                        100          0 101 i

Route Distinguisher: 192.168.1.4:3
*>i[5]:[0]:[0]:[24]:[61.1.1.0]/224
                      192.168.1.3              10        100          0 101 ?
*>l[5]:[0]:[0]:[24]:[67.1.1.0]/224
                      0.0.0.0                  10                     0 101 ?
*>i[5]:[0]:[0]:[32]:[1.1.1.1]/224
                      192.168.1.3                        100          0 101 i
*>l[5]:[0]:[0]:[32]:[2.2.2.2]/224
                      0.0.0.0                                         0 101 i
```

```
Leaf2(config)# show bgp ipv4 labeled-unicast
BGP routing table information for VRF default, address family IPv4 Label Unicast
BGP table version is 8, Local Router ID is 192.168.1.4
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 - best2

   Network            Next Hop            Metric     LocPrf     Weight Path
*>i192.168.1.1/32     172.16.1.1                       100          0 i
*>i192.168.1.3/32     172.16.0.2                       100          0 i
*>l192.168.1.4/32     0.0.0.0                          100      32768 i
```

```
Leaf2(config)# show ip int brief vrf all
IP Interface Status for VRF "default"(1)
Interface            IP Address      Interface Status
Lo0                  192.168.1.4     protocol-up/link-up/admin-up
Eth1/1               172.16.1.2      protocol-up/link-up/admin-up
Eth1/2               172.16.5.2      protocol-up/link-up/admin-up

IP Interface Status for VRF "management"(2)
Interface            IP Address      Interface Status
mgmt0                10.82.139.100   protocol-up/link-up/admin-up

IP Interface Status for VRF "cu101"(3)
Interface            IP Address      Interface Status
Eth1/3.101           67.1.1.1        protocol-up/link-up/admin-up
```

```
Leaf2(config)# show forwarding 1.1.1.1/32 vrf cu101
slot  1
=======
IPv4 routes for table cu101/base
------------------+-------------------------------------------+--------------------+---------------+-----------------
Prefix            | Next-hop                                  | Interface          | Labels        | Partial Install
------------------+-------------------------------------------+--------------------+---------------+-----------------
*1.1.1.1/32       172.16.1.1                                  Ethernet1/1          PUSH 16311 492288
```

```
Leaf2(config)# show forwarding 192.168.1.3/32
slot  1
=======
IPv4 routes for table default/base
------------------+-------------------------------------------+--------------------+---------------+-----------------
Prefix            | Next-hop                                  | Interface          | Labels        | Partial Install
------------------+-------------------------------------------+--------------------+---------------+-----------------
192.168.1.3/32    172.16.1.1                                  Ethernet1/1          PUSH 16311
```

```
Leaf2(config)# show ip route vrf 101
No IP Route Table for VRF "101"
Leaf2(config)# show ip route vrf cu101
IP Route Table for VRF "cu101"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

1.1.1.1/32, ubest/mbest: 1/0
    *via 192.168.1.3%default, [200/0], 00:15:39, bgp-65001, internal, tag 101 (mpls-vpn)
2.2.2.2/32, ubest/mbest: 1/0
    *via 67.1.1.2, [20/0], 00:36:44, bgp-65001, external, tag 101
61.1.1.0/24, ubest/mbest: 1/0
    *via 192.168.1.3%default, [200/10], 00:15:39, bgp-65001, internal, tag 101 (mpls-vpn)
67.1.1.0/24, ubest/mbest: 1/0, attached
    *via 67.1.1.1, Eth1/3.101, [0/0], 00:39:32, direct
67.1.1.1/32, ubest/mbest: 1/0, attached
    *via 67.1.1.1, Eth1/3.101, [0/0], 00:39:32, local
```

```
host1# show ip route vrf custA
IP Route Table for VRF "custA"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

1.1.1.1/32, ubest/mbest: 2/0, attached
    *via 1.1.1.1, Lo0, [0/0], 00:40:10, local
    *via 1.1.1.1, Lo0, [0/0], 00:40:10, direct
2.2.2.2/32, ubest/mbest: 1/0
    *via 61.1.1.1, [20/0], 00:37:21, bgp-101, external, tag 65001
61.1.1.0/24, ubest/mbest: 1/0, attached
    *via 61.1.1.2, Vlan101, [0/0], 00:37:38, direct
61.1.1.2/32, ubest/mbest: 1/0, attached
    *via 61.1.1.2, Vlan101, [0/0], 00:37:38, local
67.1.1.0/24, ubest/mbest: 1/0
    *via 61.1.1.1, [20/0], 00:37:21, bgp-101, external, tag 65001
RTP_host1#
```

```
host1# ping 2.2.2.2 vrf custA
PING 2.2.2.2 (2.2.2.2): 56 data bytes
64 bytes from 2.2.2.2: icmp_seq=0 ttl=251 time=0.737 ms
64 bytes from 2.2.2.2: icmp_seq=1 ttl=251 time=0.579 ms
64 bytes from 2.2.2.2: icmp_seq=2 ttl=251 time=0.513 ms
64 bytes from 2.2.2.2: icmp_seq=3 ttl=251 time=0.472 ms
64 bytes from 2.2.2.2: icmp_seq=4 ttl=251 time=0.466 ms

--- 2.2.2.2 ping statistics ---
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 0.466/0.553/0.737 ms
RTP_host1#
```

```
host2# show ip route vrf custB
IP Route Table for VRF "custB"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

1.1.1.1/32, ubest/mbest: 1/0
    *via 67.1.1.1, [20/0], 00:37:25, bgp-101, external, tag 65001
2.2.2.2/32, ubest/mbest: 2/0, attached
    *via 2.2.2.2, Lo1, [0/0], 00:40:14, local
    *via 2.2.2.2, Lo1, [0/0], 00:40:14, direct
61.1.1.0/24, ubest/mbest: 1/0
    *via 67.1.1.1, [20/0], 00:37:25, bgp-101, external, tag 65001
67.1.1.0/24, ubest/mbest: 1/0, attached
    *via 67.1.1.2, Vlan100, [0/0], 00:38:08, direct
67.1.1.2/32, ubest/mbest: 1/0, attached
    *via 67.1.1.2, Vlan100, [0/0], 00:38:08, local
host2#
```

```
host2# ping 1.1.1.1 vrf custB
PING 1.1.1.1 (1.1.1.1): 56 data bytes
64 bytes from 1.1.1.1: icmp_seq=0 ttl=251 time=0.786 ms
64 bytes from 1.1.1.1: icmp_seq=1 ttl=251 time=0.526 ms
64 bytes from 1.1.1.1: icmp_seq=2 ttl=251 time=0.604 ms
64 bytes from 1.1.1.1: icmp_seq=3 ttl=251 time=0.568 ms
64 bytes from 1.1.1.1: icmp_seq=4 ttl=251 time=0.522 ms

--- 1.1.1.1 ping statistics ---
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 0.522/0.601/0.786 ms
RTP_host1#
```

Related Information

- Multiprotocol BGP MPLS VPN
- Segment Routing on Cisco Nexus 9500, 9300, 9200, 3200, and 3100 Platform Switches White paper
- Configuring Layer 3 EVPN and Layer 3 VPN over Segment Routing MPLS