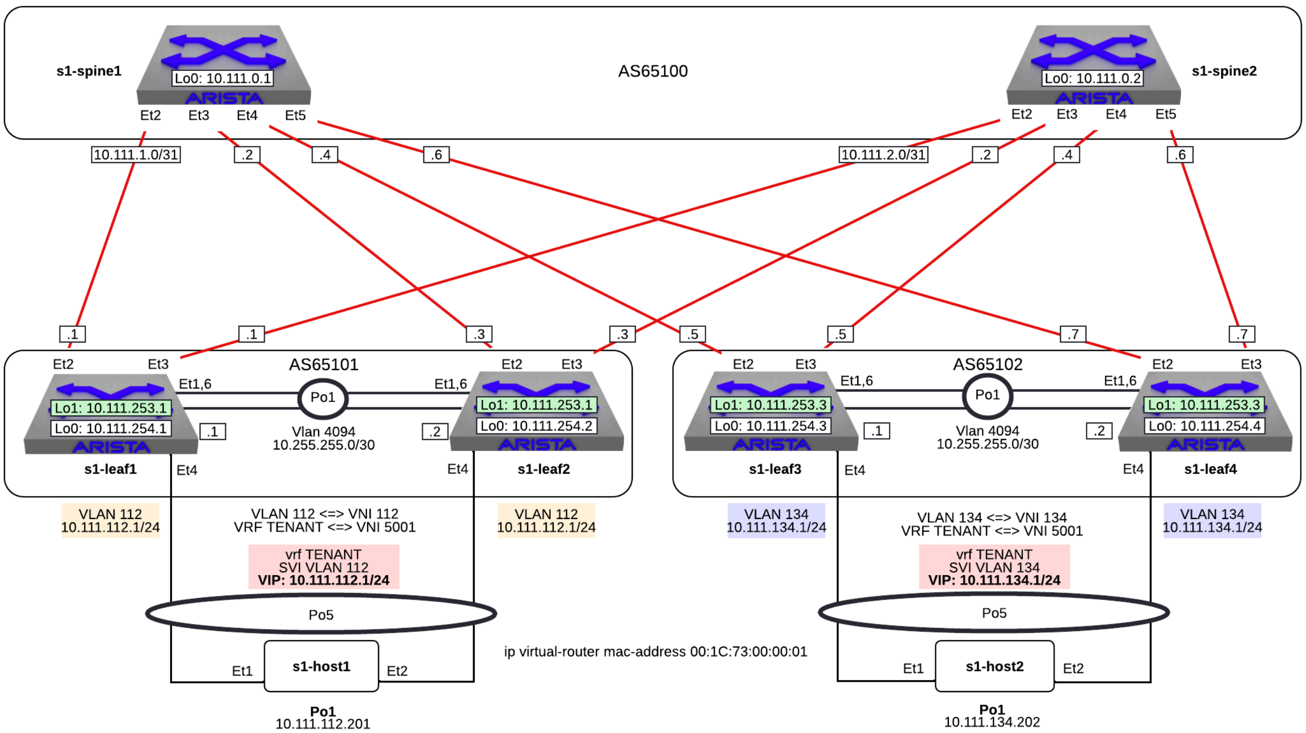
L3 EVPN Services



(_images/nested_l3evpn_topo_dual_dc.png)

> **Note**
>
> This lab exercise is focused on the VXLAN EVPN configuration. IP addresses, MLAG and BGP Underlay are already configured.

1. Log into the **LabAccess** jumpserver:

   a. Type `97` to access additional lab, then `evpn-labs` at the prompt to access the EVPN VXLAN content. Then type `l3evpn` for the Layer 3 EVPN lab. The script will configure the datacenter with the exception of **s1-leaf4**.

Note

Did you know the “l3evpn” script is composed of Python code that uses the CloudVision Portal REST API to automate the provisioning of CVP Configlets. The configlets that are configured via the REST API are `L2EVPN_s1-spine1`, `L2EVPN_s1-spine2`, `L2EVPN_s1-leaf1`, `L2EVPN_s1-leaf2`, `L2EVPN_s1-leaf3`, `L2EVPN_s1-leaf4`.

2. On **s1-leaf4**, check if Multi-Agent Routing Protocols are enabled.

```
s1-leaf4#show run section service
service routing protocols model multi-agent
s1-leaf4#show ip route summary

Operating routing protocol model: multi-agent
Configured routing protocol model: multi-agent

VRF: default
   Route Source                                Number Of Routes
------------------------------------- -------------------------
   connected                                                  4
   static (persistent)                                        0
   static (non-persistent)                                    0
   VXLAN Control Service                                      0
   static nexthop-group                                       0
   ospf                                                       0
     Intra-area: 0 Inter-area: 0 External-1: 0 External-2: 0
     NSSA External-1: 0 NSSA External-2: 0
   ospfv3                                                     0
   bgp                                                        9
     External: 7 Internal: 2
   isis                                                       0
     Level-1: 0 Level-2: 0
   rip                                                        0
   internal                                                  11
   attached                                                   3
   aggregate                                                  0
   dynamic policy                                             0
   gribi                                                      0

   Total Routes                                              27

Number of routes per mask-length:
   /8: 2         /24: 3        /30: 1        /31: 2        /32: 19
```

Note

By default, EOS is using the GateD routing process. Activating (ArBGP) is requiring a reboot. This has been done prior to the lab buildout so no reboot is required here.

3. On **s1-leaf4**, check the following operational states before configuring EVPN constructs:

   a. Verify EOS MLAG operational details.

   Note

   The MLAG state between **s1-leaf4** and its peer **s1-leaf3** will be inconsistent. This is expected as **s1-leaf3** is fully configured and **s1-leaf4** is not as of yet.

```
s1-leaf4#show mlag
MLAG Configuration:
domain-id                          :                MLAG
local-interface                    :            Vlan4094
peer-address                       :        10.255.255.1
peer-link                          :        Port-Channel1
peer-config                        :        inconsistent

MLAG Status:
state                              :              Active
negotiation status                 :           Connected
peer-link status                   :                  Up
local-int status                   :                  Up
system-id                          :   02:1c:73:c0:c6:14
dual-primary detection             :            Disabled
dual-primary interface errdisabled :               False

MLAG Ports:
Disabled                           :                   0
Configured                         :                   0
Inactive                           :                   0
Active-partial                     :                   0
Active-full                        :                   0
```

   b. Verify BGP operational details for Underlay:

   Note

   You should see 3 underlay sessions; one to each spine and one to the MLAG peer for redundancy.

```
s1-leaf4#show ip bgp summary
BGP summary information for VRF default
Router identifier 10.111.254.4, local AS number 65102
Neighbor Status Codes: m - Under maintenance
Neighbor      V AS          MsgRcvd   MsgSent  InQ OutQ  Up/Down State   PfxRcd PfxAcc
10.111.1.6    4 65100             9        12    0    0 00:00:07 Estab   6      6
10.111.2.6    4 65100             9        12    0    0 00:00:07 Estab   5      5
10.255.255.1 4 65102             8        10    0    0 00:00:07 Estab   10     10
```

c. Check the IP routing table:

Note

Notice that **s1-leaf4** has 2 ECMP paths for reaching **s1-leaf1** or **s1-leaf2** loopacks.

```
s1-leaf4#show ip route

VRF: default
Codes: C - connected, S - static, K - kernel,
      O - OSPF, IA - OSPF inter area, E1 - OSPF external type 1,
      E2 - OSPF external type 2, N1 - OSPF NSSA external type 1,
      N2 - OSPF NSSA external type2, B - Other BGP Routes,
      B I - iBGP, B E - eBGP, R - RIP, I L1 - IS-IS level 1,
      I L2 - IS-IS level 2, O3 - OSPFv3, A B - BGP Aggregate,
      A O - OSPF Summary, NG - Nexthop Group Static Route,
      V - VXLAN Control Service, M - Martian,
      DH - DHCP client installed default route,
      DP - Dynamic Policy Route, L - VRF Leaked,
      G  - gRIBI, RC - Route Cache Route

Gateway of last resort is not set

B E      10.111.0.1/32 [200/0] via 10.111.1.6, Ethernet2
B E      10.111.0.2/32 [200/0] via 10.111.2.6, Ethernet3
C        10.111.1.6/31 is directly connected, Ethernet2
B E      10.111.1.0/24 [200/0] via 10.111.1.6, Ethernet2
C        10.111.2.6/31 is directly connected, Ethernet3
B E      10.111.2.0/24 [200/0] via 10.111.2.6, Ethernet3
B I      10.111.112.0/24 [200/0] via 10.255.255.1, Vlan4094
B E      10.111.253.1/32 [200/0] via 10.111.1.6, Ethernet2
                                 via 10.111.2.6, Ethernet3
B I      10.111.253.3/32 [200/0] via 10.255.255.1, Vlan4094
B E      10.111.254.1/32 [200/0] via 10.111.1.6, Ethernet2
                                 via 10.111.2.6, Ethernet3
B E      10.111.254.2/32 [200/0] via 10.111.1.6, Ethernet2
                                 via 10.111.2.6, Ethernet3
B I      10.111.254.3/32 [200/0] via 10.255.255.1, Vlan4094
C        10.111.254.4/32 is directly connected, Loopback0
C        10.255.255.0/30 is directly connected, Vlan4094
C        192.168.0.0/24 is directly connected, Management0
```

4. On **s1-leaf4**, configure the BGP EVPN control-plane.

   a. Configure the EVPN control plane.

Note

In this lab, the Spines serve as EVPN Route Servers. They receive the EVPN Routes from each leaf and, due to our eBGP setup, will naturally pass them along the other leaves.

Also note that BGP standard and extended communities are explicitly enabled on the peering. EVPN makes use of extended BGP communities for route signaling and standard communities allow for various other functions such as BGP maintenance mode.

Lastly, note in this setup we use eBGP-multihop peerings with the Loopback0 interfaces of each switch. This follows Arista best-practice designs for separation of Underlay (peerings done using physical Ethernet interfaces) and Overlay (peerings done using Loopbacks) when leveraging eBGP. Other options exist and can be discussed with your Arista SE.

```
router bgp 65102
    neighbor SPINE-EVPN peer group
    neighbor SPINE-EVPN remote-as 65100
    neighbor SPINE-EVPN update-source Loopback0
    neighbor SPINE-EVPN ebgp-multihop 3
    neighbor SPINE-EVPN send-community standard extended
    neighbor 10.111.0.1 peer group SPINE-EVPN
    neighbor 10.111.0.2 peer group SPINE-EVPN
    !
    address-family evpn
       neighbor SPINE-EVPN activate
```

b. Verify the EVPN Control-Plane is established to both Spine peers.

```
s1-leaf4#show bgp evpn summary
BGP summary information for VRF default
Router identifier 10.111.254.4, local AS number 65102
Neighbor Status Codes: m - Under maintenance
  Neighbor    V AS           MsgRcvd   MsgSent  InQ OutQ  Up/Down State   PfxRcd PfxAcc
  10.111.0.1 4 65100              8         6    0    0 00:00:14 Estab   4      4
  10.111.0.2 4 65100              8         4    0    0 00:00:14 Estab   4      4
```

5. On **s1-leaf4**, configure the VXLAN data-plane for transport.

   a. Configure Loopback1 with the shared IP of **s1-leaf3**.

   Note

   This is referred to as an MLAG VTEP. The MLAG peer leafs provide redundancy by sharing the Loopback1 IP and jointly advertising reachability for it. Route redistribution has already been configured for the underlay.

```
interface Loopback1
   description VTEP
   ip address 10.111.253.3/32
```

   b. Configure the Vxlan1 interface with the Loopback1 as the source.

   Note

   This is the logical interface that will provide VXLAN header encap and decap functions. In this lab, since we are leveraging VXLAN routing, we can able the use of a virtual-router MAC address. This instructs the device to use the shared MLAG System ID as the router MAC when performing VXLAN routing operations and ensures that whichever switch in the MLAG receives the VXLAN Routed packet can provide forwarding of that traffic without shunting it over the MLAG peer-link.

```
interface Vxlan1
   vxlan source-interface Loopback1
   vxlan virtual-router encapsulation mac-address mlag-system-id
```

6. Configure a Layer 3 EVPN service on **s1-leaf4**.

   a. Add the local Layer 2 VLAN with an ID of 134 that the host will attach to.

   ```
   vlan 134
      name Host_Network_134
   ```

   b. Create the VRF, or logical routing instance, for the Tenant Layer 3 Network.

   Note

   In EOS, by default, VRFs are created with inter-subnet routing disabled. Always be sure to enable IP routing in user-defined VRFs.

   ```
   vrf instance TENANT
   !
   ip routing vrf TENANT
   ```

   c. Create the SVI for default gateway function for the host network as an Anycast Gateway.

   Note

   With VXLAN, we can leverage a shared IP using Anycast Gateway. This allows a single IP to be shared without any other dedicated IPs per switch.

   ```
   ip virtual-router mac-address 00:1C:73:00:00:01
   !
   interface Vlan134
      description Host Network 134
      vrf TENANT
      ip address virtual 10.111.134.1/24
   ```

   d. Map the local Layer 3 VRF with a matching VNI.

   Note

   For the Layer 3 Service, the VRF requires what is referred to as the Layer 3 VNI, which is used for VXLAN Routing in a Symmetric IRB deployment between VTEPs. Any unique ID number will serve here.

   ```
   interface Vxlan1
      vxlan vrf TENANT vni 5001
   ```

   e. Add the IP VRF EVPN configuration for the TENANT VRF.

   Note

Here we configure a Layer 3 VRF service with EVPN. It has two components. The first is a route-distinguisher, or **RD** to identify the router (or leaf switch) that is originating the EVPN routes. This can be manually defined in the format of **Number** : **Number**, such as **Loopback0** : **VRF ID** or as we do in this case, let EOS automatically allocate one. The Auto RD function is enabled globally for all VRFs under the BGP process.

Second is the route-target, or **RT**. The **RT** is used by the leaf switches in the network to determine if they should import the advertised route into their local table(s). If they receive an EVPN route, they check the **RT** value and see if they have a matching **RT** configured in BGP. If they do, they import the route into the associated VRF. If they do not, they ignore the route.

```
router bgp 65102
   rd auto
   !
   vrf TENANT
      route-target import evpn 5001:5001
      route-target export evpn 5001:5001
      redistribute connected
```

f. Configure the host-facing MLAG port.

```
interface Port-Channel5
   description MLAG Downlink - s1-host2
   switchport access vlan 134
   mlag 5
!
interface Ethernet4
   description MLAG Downlink - s1-host2
   channel-group 5 mode active
```

7. With the Layer 3 EVPN Service configured, verify the operational state.

   a. Check the VXLAN data-plane configuration.

Note

Here we can see some useful commands for VXLAN verification. `show vxlan config-sanity detail` verifies a number of standard things locally and with the MLAG peer to ensure all basic criteria are met. `show interfaces Vxlan1` provides a consolidated series of outputs of operational VXLAN data such as control-plane mode (EVPN in this case), VRF to VNI mappings and MLAG Router MAC.

```
s1-leaf4#show vxlan config-sanity detail
Category                            Result  Detail
---------------------------------- -------- ------------------------------------------
Local VTEP Configuration Check        OK
  Loopback IP Address                 OK
  VLAN-VNI Map                        OK
  Flood List                          OK
  Routing                             OK
  VNI VRF ACL                         OK
  Decap VRF-VNI Map                   OK
  VRF-VNI Dynamic VLAN                OK
Remote VTEP Configuration Check       OK
  Remote VTEP                         OK
Platform Dependent Check              OK
  VXLAN Bridging                      OK
  VXLAN Routing                       OK
CVX Configuration Check               OK
  CVX Server                          OK    Not in controller client mode
MLAG Configuration Check              OK    Run 'show mlag config-sanity' to verify MLA
  Peer VTEP IP                        OK
  MLAG VTEP IP                        OK
  Peer VLAN-VNI                       OK
  Virtual VTEP IP                     OK
  MLAG Inactive State                 OK

s1-leaf4#show interfaces Vxlan1
Vxlan1 is up, line protocol is up (connected)
  Hardware is Vxlan
  Source interface is Loopback1 and is active with 10.111.253.3
  Replication/Flood Mode is headend with Flood List Source: CLI
  Remote MAC learning is disabled
  VNI mapping to VLANs
  Static VLAN to VNI mapping is
  Dynamic VLAN to VNI mapping for 'evpn' is
    [4092, 5001]
  Note: All Dynamic VLANs used by VCS are internal VLANs.
        Use 'show vxlan vni' for details.
  Static VRF to VNI mapping is
   [TENANT, 5001]
  MLAG Shared Router MAC is 021c.73c0.c614
```

b. On **s1-leaf1** (and/or **s1-leaf2**) verify the BGP and Route table to ensure the Tenant network on **s1-leaf4** has been learned in the overlay.

Note

The output below shows learned **IP Prefix** routes from EVPN. These are referred to as EVPN Type 5 routes. Other leaves receive this route, evaluate the **RT** to see if they have a matching configuration and, if so, import the contained prefix into their VRF Route Table. Note that IPv4 and IPv6 are supported.

Note on the route table for the TENANT VRF, we see a single route entry for the remote tenant subnet. This route is directed to the shared MLAG VTEP IP and Router MAC. It will be ECMPed via the Spines providing a dual path for load-balancing and redundancy.

```
s1-leaf1#show bgp evpn route-type ip-prefix ipv4
BGP routing table information for VRF default
Router identifier 10.111.254.1, local AS number 65101
Route status codes: * - valid, > - active, S - Stale, E - ECMP head, e - ECMP
                    c - Contributing to ECMP, % - Pending BGP convergence
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Loca

          Network                Next Hop              Metric  LocPref Weight  Path
 * >      RD: 10.111.254.1:1 ip-prefix 10.111.112.0/24
                                 -                     -       -       0       i
 * >Ec    RD: 10.111.254.3:1 ip-prefix 10.111.134.0/24
                                 10.111.253.3          -       100     0       65100 65
 *  ec    RD: 10.111.254.3:1 ip-prefix 10.111.134.0/24
                                 10.111.253.3          -       100     0       65100 65
 * >Ec    RD: 10.111.254.4:1 ip-prefix 10.111.134.0/24
                                 10.111.253.3          -       100     0       65100 65
 *  ec    RD: 10.111.254.4:1 ip-prefix 10.111.134.0/24
                                 10.111.253.3          -       100     0       65100 65

s1-leaf1#show ip route vrf TENANT

VRF: TENANT
Codes: C - connected, S - static, K - kernel,
       O - OSPF, IA - OSPF inter area, E1 - OSPF external type 1,
       E2 - OSPF external type 2, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type2, B - Other BGP Routes,
       B I - iBGP, B E - eBGP, R - RIP, I L1 - IS-IS level 1,
       I L2 - IS-IS level 2, O3 - OSPFv3, A B - BGP Aggregate,
       A O - OSPF Summary, NG - Nexthop Group Static Route,
       V - VXLAN Control Service, M - Martian,
       DH - DHCP client installed default route,
       DP - Dynamic Policy Route, L - VRF Leaked,
       G  - gRIBI, RC - Route Cache Route

Gateway of last resort is not set

 C        10.111.112.0/24 is directly connected, Vlan112
 B E      10.111.134.0/24 [200/0] via VTEP 10.111.253.3 VNI 5001 router-mac 02:1c:73:c0
```

c. Log into **s1-host1** and ping **s2-host2** to verify connectivity.

```
s1-host1#ping 10.111.134.202
PING 10.111.112.202 (10.111.134.202) 72(100) bytes of data.
80 bytes from 10.111.134.202: icmp_seq=1 ttl=64 time=16.8 ms
80 bytes from 10.111.134.202: icmp_seq=2 ttl=64 time=14.7 ms
80 bytes from 10.111.134.202: icmp_seq=3 ttl=64 time=16.8 ms
80 bytes from 10.111.134.202: icmp_seq=4 ttl=64 time=16.7 ms
80 bytes from 10.111.134.202: icmp_seq=5 ttl=64 time=15.2 ms
--- 10.111.134.202 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 61ms
```

d. On **s1-leaf1**, check the local MAC address-table and ARP Table.

> **Note**
>
> The MAC addresses in your lab may differ as they are randomly generated during the lab build. We see here that the ARP and MAC of **s1-host1** has been learned locally **s1-leaf1**. We also see the remote MAC for the shared MLAG System ID of **s1-leaf3** and **s1-leaf4** associated with VLAN 4092 and the Vxlan1 interface. This is how the local VTEP knows where to send routed traffic when destined to the remote MLAG pair. We can see this VLAN is dynamically created in the VLAN database and is mapped to our Layer 3 VNI (5001) in our VXLAN interface output. **Be aware that since this VLAN is dynamic, the ID used in your lab may be different.**
>
> Since we are using VXLAN ONLY for Layer 3 VRF services and not extending any local VLANs, **s1-host2**'s MAC and ARP are not learned. It is reached via the IP Prefix route only.

```
s1-leaf1#show ip arp vrf TENANT
Address         Age (sec)  Hardware Addr   Interface
10.111.112.201    0:08:01  001c.73c0.c616  Vlan112, not learned
s1-leaf1#show mac address-table dynamic
          Mac Address Table
------------------------------------------------------------------

Vlan    Mac Address       Type        Ports      Moves   Last Move
----    -----------       ----        -----      -----   ---------
 112    001c.73c0.c616    DYNAMIC     Po5        1       0:00:05 ago
4092    021c.73c0.c614    DYNAMIC     Vx1        1       3:25:13 ago
Total Mac Addresses for this criterion: 1

          Multicast Mac Address Table
------------------------------------------------------------------

Vlan    Mac Address       Type        Ports
----    -----------       ----        -----
Total Mac Addresses for this criterion: 0
s1-leaf1#show vlan 4092
VLAN  Name                             Status    Ports
----- -------------------------------- --------- -------------------------------
4092* VLAN4092                         active    Cpu, Po1, Vx1

* indicates a Dynamic VLAN
s1-leaf1#show interfaces Vxlan1
Vxlan1 is up, line protocol is up (connected)
  Hardware is Vxlan
  Source interface is Loopback1 and is active with 10.111.253.1
  Replication/Flood Mode is headend with Flood List Source: CLI
  Remote MAC learning is disabled
  VNI mapping to VLANs
  Static VLAN to VNI mapping is
  Dynamic VLAN to VNI mapping for 'evpn' is
    [4092, 5001]
  Note: All Dynamic VLANs used by VCS are internal VLANs.
        Use 'show vxlan vni' for details.
  Static VRF to VNI mapping is
   [TENANT, 5001]
  MLAG Shared Router MAC is 021c.73c0.c612
```

LAB COMPLETE!

Back to top