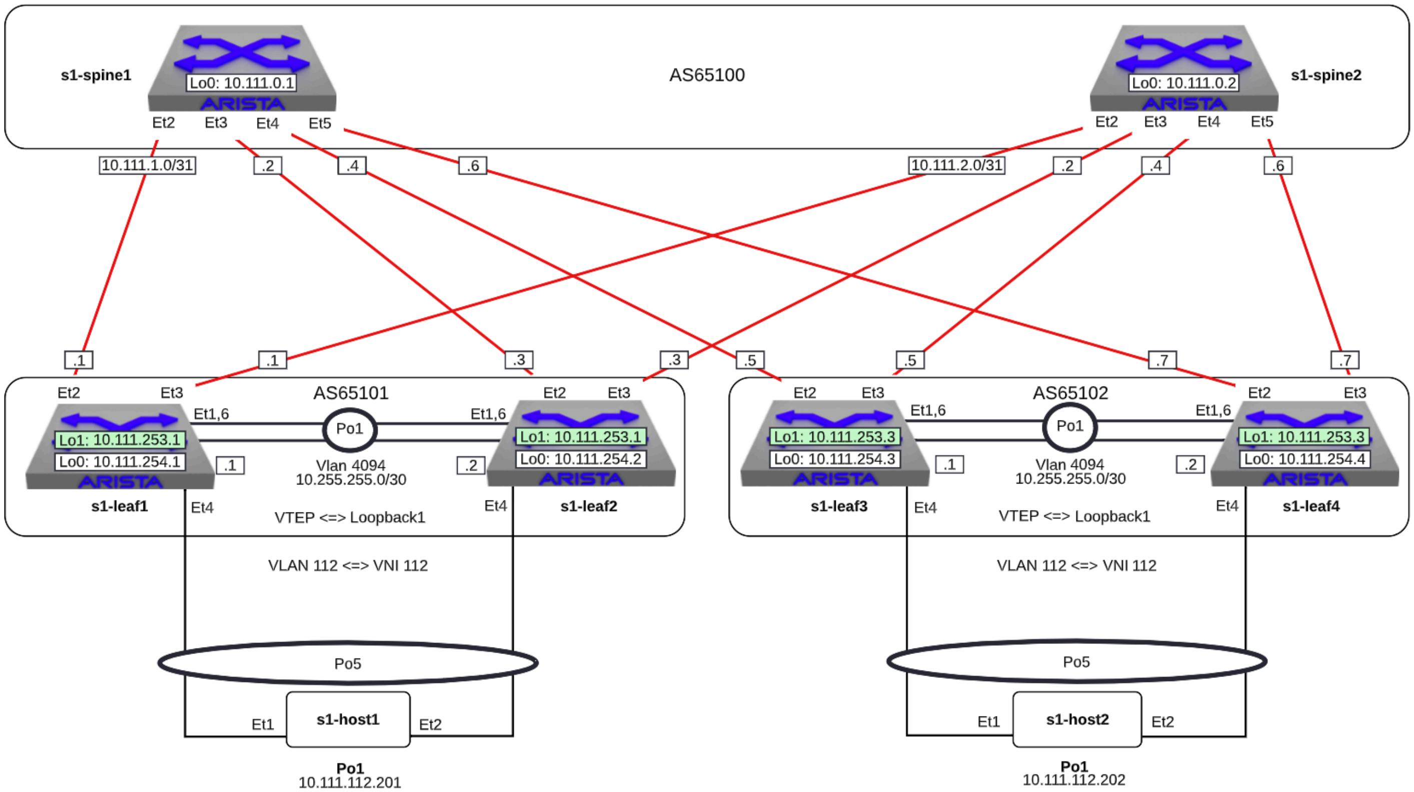
L2 EVPN Services



(_images/nested_l2evpn_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 `l2evpn` for the Layer 2 EVPN lab. The script will configure the datacenter with the exception of **s1-leaf4**.

   > **Note**
   >
   > Did you know the “l2evpn” 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(config-router-bgp)#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              6         5    0    0 00:00:03 Estab   2      2
10.111.0.2 4 65100              6         4    0    0 00:00:03 Estab   2      2
```

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.

```
interface Vxlan1
   vxlan source-interface Loopback1
```

6. Configure a Layer 2 EVPN service on **s1-leaf4**.

   a. Add the local Layer 2 VLAN with an ID of 112.

```
vlan 112
   name Host_Network_112
```

b. Map the local Layer 2 VLAN with a matching VNI.

> **Note**
>
> This is how the switch understands which local Layer 2 VLAN maps to which VNI in the overlay. The example shows matching them one to one, but any scheme or method is valid, such as adding 10000 to the VLAN ID.

```
interface Vxlan1
   vxlan vlan 112 vni 112
```

c. Add the mac-vrf EVPN configuration for VLAN 112.

> **Note**
>
> Here we configure a VLAN-based 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** : **VLAN ID** or as we do in this case, let EOS automatically allocate one.
>
> 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 mac-vrf (or VLAN). If they do not, they ignore the route.

```
router bgp 65102
   !
   vlan 112
      rd auto
      route-target both 112:112
      redistribute learned
```

d. Configure the host-facing MLAG port.

```
interface Port-Channel5
   description MLAG Downlink - s1-host2
   switchport access vlan 112
   mlag 5
!
interface Ethernet4
   description MLAG Downlink - s1-host2
   channel-group 5 mode active
```

7. With the Layer 2 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), VLAN to VNI mappings and discovered VTEPs.

```
s1-leaf4#show vxlan config-sanity detail
Category                            Result  Detail
---------------------------------- -------- ------------------------------------------
Local VTEP Configuration Check        OK
  Loopback IP Address                 OK
  VLAN-VNI Map                        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    VXLAN Routing not enabled
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

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: EVPN
Remote MAC learning via EVPN
VNI mapping to VLANs
Static VLAN to VNI mapping is
  [112, 112]
Note: All Dynamic VLANs used by VCS are internal VLANs.
      Use 'show vxlan vni' for details.
Static VRF to VNI mapping is not configured
Headend replication flood vtep list is:
112 10.111.253.1
MLAG Shared Router MAC is 0000.0000.0000
```

b. On **s1-leaf1** (and/or **s1-leaf2**) verify the IMET table to ensure **s1-leaf4** has been discovered in the overlay.

Note

The Inclusive Multicast Ethernet Tag, or **IMET**, route is how a VTEP advertises membership in a given Layer 2 service, or VXLAN segment. This is also known as the EVPN Type 3 Route. Other leaves receive this route, evaluate the **RT** to see if they have a matching configuration and, if so, import the advertising VTEP into their flood list for BUM traffic.

```
s1-leaf1#show bgp evpn route-type imet
BGP routing table information for VRF default
Router identifier 10.111.254.1, local AS number 65101
Route status codes: s - suppressed, * - valid, > - active, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup
                    % - 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
* >Ec   RD: 10.111.254.3:112 imet 10.111.253.3
                                 10.111.253.3          -       100     0       65100 651
*  ec   RD: 10.111.254.3:112 imet 10.111.253.3
                                 10.111.253.3          -       100     0       65100 651
* >Ec   RD: 10.111.254.4:112 imet 10.111.253.3
                                 10.111.253.3          -       100     0       65100 651
*  ec   RD: 10.111.254.4:112 imet 10.111.253.3
                                 10.111.253.3          -       100     0       65100 651
* >     RD: 10.111.254.1:112 imet 10.111.253.1
                                 -                     -       -       0       i
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: EVPN
  Remote MAC learning via EVPN
  VNI mapping to VLANs
  Static VLAN to VNI mapping is
    [112, 112]
  Note: All Dynamic VLANs used by VCS are internal VLANs.
        Use 'show vxlan vni' for details.
  Static VRF to VNI mapping is not configured
  Headend replication flood vtep list is:
  112 10.111.253.3
  MLAG Shared Router MAC is 0000.0000.0000
```

c. Log into **s1-host1** and ping **s2-host2** to populate the network's MAC tables.

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

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

> **Note**
>
> The MAC addresses in your lab may differ as they are randomly generated during the lab build. We see here that the MAC of **s1-host2** has been learned via the Vxlan1 interface on **s1-leaf1**.

```
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       0:00:41 ago
112    001c.73c0.c617    DYNAMIC     Vx1        1       0:00:41 ago
Total Mac Addresses for this criterion: 2
      Multicast Mac Address Table
------------------------------------------------------------------


Vlan    Mac Address       Type        Ports
----    -----------       ----        -----
Total Mac Addresses for this criterion: 0
```

e. On **s1-leaf1**, check the EVPN control-plane for the associated host MAC.

> **Note**
>
> We see the MAC of **s1-host2** multiple times in the control-plane due to our redundant MLAG and ECMP design. Both **s1-leaf3** and **s1-leaf4** are attached to **s1-host2** and therefore will generate this Type 2 EVPN route for its MAC. They each then send this route up to the redundant Spines (or EVPN Route Servers) which provides an ECMP path to the host.

```
s1-leaf1#show bgp evpn route-type mac-ip
BGP routing table information for VRF default
Router identifier 10.111.254.1, local AS number 65101
Route status codes: s - suppressed, * - valid, > - active, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup
                    % - 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: 10.111.254.1:112 mac-ip 001c.73c0.c616
                                 -                     -       -       0       i
* >      RD: 10.111.254.1:112 mac-ip 001c.73c0.c616 10.111.112.201
                                 -                     -       -       0       i
* >Ec    RD: 10.111.254.3:112 mac-ip 001c.73c0.c617
                                 10.111.253.3          -       100     0       65100 651
*  ec    RD: 10.111.254.3:112 mac-ip 001c.73c0.c617
                                 10.111.253.3          -       100     0       65100 651
* >Ec    RD: 10.111.254.4:112 mac-ip 001c.73c0.c617
                                 10.111.253.3          -       100     0       65100 651
*  ec    RD: 10.111.254.4:112 mac-ip 001c.73c0.c617
                                 10.111.253.3          -       100     0       65100 651
```

f. On **s1-leaf1**, check the VXLAN data-plane for MAC address.

> **Note**
>
> Though both **s1-leaf3** and **s1-leaf4** are advertising the MAC of **s1-host2** recall that they have a shared MLAG VTEP IP for VXLAN reachability. Therefore we only see one possible destination for this host MAC. The `show l2rib output mac <MAC of remote host>` command then allows us to see the VTEP info in the hardware. Finally we can verify the ECMP path to the remote MLAG VTEP via **s1-spine1** and **s1-spine2** with a simple `show ip route 10.111.253.3` command.

```
s1-leaf1#show vxlan address-table evpn
  Vxlan Mac Address Table
----------------------------------------------------------------------

VLAN  Mac Address     Type      Prt  VTEP             Moves   Last Move
----  -----------     ----      ---  ----             -----   ---------
112  001c.73c0.c617  EVPN      Vx1  10.111.253.3     1       0:00:57 ago
Total Remote Mac Addresses for this criterion: 1
s1-leaf1#show l2rib output mac 001c.73c0.c617
001c.73c0.c617, VLAN 112, seq 1, pref 16, evpnDynamicRemoteMac, source: BGP
   VTEP 10.111.253.3
s1-leaf1#show ip route 10.111.253.3

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

 B E      10.111.253.3/32 [200/0] via 10.111.1.0, Ethernet2
                                  via 10.111.2.0, Ethernet3
```

LAB COMPLETE!

Back to top