



External Connectivity—VRF Lite

- [External Connectivity](#), on page 1

External Connectivity

Layer-3 DCI for VXLAN BGP EVPN fabrics—VRF lite

VXLAN BGP EVPN data center fabrics can be connected across Layer-3 boundaries using MPLS L3VPN, VRF IP Routing (VRF lite), or LISP as the mechanism of transport outside the VXLAN fabric. The VRF lite scenario is explained in this chapter.

VM mobility across datacenters

VM mobility across VXLAN BGP EVPN datacenter fabrics works the same way as it does within the datacenter fabric. When VM mobility takes place, the VM generates RARP and GARP messages. You should enable a Layer-2 DCI such as OTV, Classical Ethernet or VPLS to transport broadcast RARP and GARP packets generated due to the VM movement.



Note

Additional configuration is not required to support VM movement across fabrics.

VM mobility across fabrics cannot take place in these scenarios:

1. VM movement takes place when MAC chaining (multiple IP addresses mapped to the same MAC address) is in effect.
2. When an end host sends a non broadcast packet such as ARP on VM move.

VXLAN BGP EVPN - VRF lite scenario - In brief

Some pointers are given below:

- The VXLAN BGP EVPN fabric pods are depicted at the left and the right of the below image (*VRF Lite DCI hand-off topology*). Routes within a fabric pod are shared between all VTEPs within the pod, including with the Cisco Nexus 7000 Series border leaf switches.
- The border leaf switches and the WAN ASBR/PE routers are configured to pass on routes between each other through eBGP sessions, using VRF Lite.

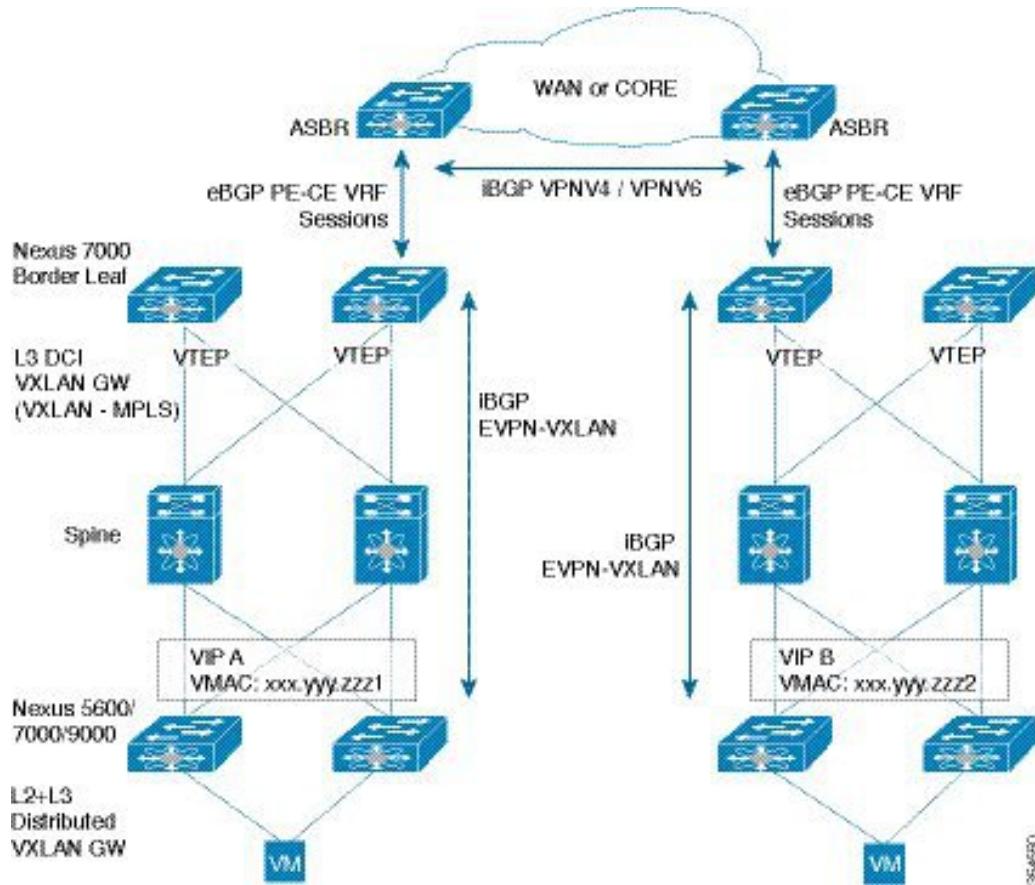
Layer-3 DCI for VXLAN BGP EVPN fabrics—VRF lite

For example, tenant VRF routes within VXLAN BGP EVPN fabric (left) are sent to the WAN ASBR routers, and (necessary) reachability routes within the WAN are sent to the border leaf switches on the VXLAN fabric (left).

- The WAN and the VXLAN fabric (right) can also be connected in a similar way.

As a result, the data center pods depicted in the left and right of the image are seamlessly connected through the WAN using VRF lite.

Figure 1: VRF Lite DCI hand-off topology



VXLAN BGP EVPN – VRF lite scenario – In more detail

Route distribution between the VXLAN pod (left) and the WAN are explained in the order given below:

1. Route distribution within the VXLAN pod, and subsequent export of VXLAN pod routes from the Border Leaf switch to the WAN.
2. Importing of VXLAN BGP EVPN fabric routes on the WAN edge device.
3. Importing of WAN routes into the VXLAN BGP EVPN border leaf switch.

Step 1 - Route distribution within the VXLAN pod, and subsequent export of VXLAN pod routes to the WAN

The routes within the VXLAN BGP EVPN pod can be exported to the WAN by the following process, thereby extending Layer-3 reachability from the WAN to the ToRs in the VXLAN BGP EVPN fabric.

- The BGP EVPN control plane in the VXLAN BGP EVPN fabric ensures distribution of routes between VTEPs (and to the border leaf switches) within the fabric. ToRs will forward the attached host IP and MAC addresses (/32 Host IP [or /128, for IPv6 addresses] + MAC routes) using the EVPN Route Type 5 option and the border leaf switch will import the /32 (or /128) routes into local VRF tables.



Note For VRF Lite extension, configure only those VRF instances on the border leaf switch that need to be extended outside the fabric. You need to configure a Layer-3 sub interface towards the WAN ASBR/PE and establish an eBGP session over that. Based on the configurations on the border leaf switch, the data is forwarded towards the WAN.

- If configured to do so, the border leaf switch advertises a 0/0 default route (IPv4/IPv6) per VRF to the other leaf switches/ToRs. When ToR switch nodes receive the same route from multiple border leaf switches, it results in ECMP at the ingress ToR/leaf switch nodes.

Step 2 - Importing of VXLAN BGP EVPN fabric routes on the WAN ASBR/PE device

- To receive VXLAN fabric routes from the border leaf switch, the WAN ASBR/PE routers should also have Layer-3 sub interfaces configured. The 802.1Q ID on each sub interface needs to be the same for a tenant on the WAN ASBR/PE router and the fabric border leaf switch.
- Similar configurations need to be implemented on the WAN ASBR/PE router (right) designated to communicate with the border leaf switch(es) on the VXLAN fabric (right).
- The configurations on the WAN ASBR/PE routers and the respective border leaf switches ensure that fabric routes are sent to the WAN and WAN routes to the fabric, for the configured VRFs.

Step 3 -Importing of WAN ASBR/PE routes into the VXLAN BGP EVPN border leaf switch

Routes arriving at the border leaf switch need to be re-advertised to the ToRs in the VXLAN BGP EVPN fabric, with the border leaf VTEP as the next hop for the ToR switches. As a result, Layer-3 reachability is extended from the ToR switches to the WAN. The process is given below.

- The WAN routes arrive at the relevant sub interfaces on the border leaf switch.
- The border leaf switches can be configured to re-originate these imported VRF prefixes towards the EVPN control plane (on the fabric side) or can be configured to originate a 0/0 default route in each VRF. If a default route is configured, the ToRs/leaf switches will import this default route, resulting in a VRF default route towards the border leaf switches in all relevant tenant VRFs on the ToRs.
- If configured to distribute default routes, necessary configuration knobs need to be added in BGP under **VRF**, and under the **neighbor evpn** address family to originate a default route towards EVPN neighbors and drop all other routes.

Data Flow

Data flow for Layer-3 DCI using VRF Lite is given below:

Route distribution within the VXLAN pod, and export of VXLAN pod routes to the WAN

Let us say a host in the VXLAN fabric (left) initiates communication with a host in the other VXLAN pod (right). A high level data plane flow is depicted below:

- The VXLAN packet reaches the border leaf switch. If a default route is configured, (the leaf switch uses the default IP route pointing to the BL switch as the next hop).



Note For the leaf switch, any destination outside the fabric will have the VTEP of the border leaf as the next hop.

- A VXLAN VNI lookup happens. This lookup points to the appropriate bridge domain and VRF interface.
- The IP lookup in the VRF IP table points to the WAN IP adjacency on the VRF interface.
- The IP packet is routed to the WAN PE or ASBR.

Importing of WAN ASBR/PE routes into the VXLAN BGP EVPN border leaf switch

Traffic from the WAN ASBR/PE device arrives at the border leaf switch on corresponding sub interfaces. A high level flow is depicted below:

- The IP packet arrives with the 802.1Q tag associated with the sub interface. The sub interface points to the correct VRF table.
- The (VRF, IP) lookup results in a /32 (or /128) route that points to the VXLAN tunnel end point adjacency for the ToR VTEP, on the fabric facing per VRF BDI interface.
- The packet is (VXLAN) encapsulated with the router MAC (RMAC) address of the remote ToR VTEP and sent on the VRF BDI interface, and the FIB lookup drives the VXLAN encapsulation towards the designated remote ToR VTEP.

The VXLAN fabrics are stitched together

- Once the VXLAN BGP EVPN fabric routes of the left and right datacenters are exchanged between each other, connectivity is complete.

Border leaf switches in a vPC scenario

In a vPC scenario, the two border leaf switches are configured for a Layer-2 handoff scenario, and a common, virtual VTEP IP address is used for communication. To use the same VTEP for Layer-2 and Layer-3 handoff scenarios, a common VTEP IP address and router MAC address should be configured on the border leaf switches. The Layer-3 prefixes or default route from the border leaf switch will be advertised with this virtual VTEP as the next-hop, and the ToR VTEPs will install the default route or prefix route with a single BGP path to the border leaf (virtual) VTEP.

Border leaf to WAN ASBR/PE link failure scenario

If there is a link failure between the designated border leaf switch and the WAN ASBR/PE device, the switch will withdraw the BGP routes that are being advertised towards the fabric and traffic re-convergence happens through the redundant border leaf switch.



Attention

Remove the default route manual configuration on the failed switch interface. Else, BGP re-convergence will fail since a default route will originate from both the border leaf switches.

Configuration for the VXLAN BGP EVPN—VRF lite scenario

Typically, most of the VXLAN BGP EVPN configuration is done during initial configuration of the fabric. Refer to the chapters *Forwarding concepts* and *IP Fabric Underlay* for more details.

On border leaf switch 1, configure a bridge domain and associate a Layer-3 network VNI

(config)#

```
system bridge-domain 2500-3500
system fabric bridge-domain 2500-2999
vni 50000
bridge-domain 2500
    member vni 50000
```

Create a VRF and associate the Layer-3 VNI to the VRF. Then, enable importing and exporting of VXLAN fabric routes on the border leaf switch

(config)#

```
vrf context vni-50000:p1
  vni 50000
  rd auto
    address-family ipv4 unicast
      route-target both auto
      route-target both auto evpn
    address-family ipv6 unicast
      route-target both auto
      route-target both auto evpn
```

Configuration towards the WAN

Create a subinterface and associate an 802.1Q tag to it

(config)#

```
interface port-channel 1.100
```



Note

You can implement configurations on interfaces, that is, associate an interface to an 802.1Q tag.

(config-subif)#

```
encapsulation dot1q 100
vrf member vni-50000:p1
ip address 192.0.2.1/24
ipv6 address 2001:DB8::1
```

The 802.1Q tag (100) acts as the distinguisher that tells the border leaf switch which VRF the packet belongs to.

Configure BGP and associate the tenant VRF under BGP

Configuration for the VXLAN BGP EVPN—VRF lite scenario

```
(config)#
router bgp 65536
vrf vni-50000:p1
  address-family ipv4 unicast
    advertise l2vpn evpn
    maximum-paths ibgp 2
    maximum-paths 2
  address-family ipv6 unicast
    advertise l2vpn evpn
    maximum-paths ibgp 2
    maximum-paths 2
  neighbor 192.0.2.3 remote-as 65551
    address-family ipv4 unicast
      send-community both
    address-family ipv6 unicast
      send-community both
```

65551 is the BGP autonomous system (AS) ID of the WAN.

Add the Layer-3 VRF VNI to the overlay

```
(config)#
interface nve 1
  source-interface loopback 0
  member vni 50000 associate-vrf
    host-reachability protocol bgp
```

Configure a BDI and associate the Layer-3 VRF to it

```
(config)#
interface BDI 2500
  ip forwarding enable
  vrf member vni-50000:p1
```

Establish a BGP session within the VXLAN BGP EVPN fabric to enable L2VPN routes' distribution

```
(config)#
router bgp 65536
  neighbor 10.2.2.1 remote-as 65536
    address-family l2vpn evpn
```



Attention

The above configurations contain VRF Lite configuration without default route creation. To create default routes and distribute them to the leaf/ToR switches, add the following configurations

Enable default route distribution

Configure IPv4 and IPv6 default routes for the corresponding VRF

```
(config)#
  address-family ipv4
    redistribute connected
  address-family ipv6
    redistribute connected
```

```
vrf context vni-50000:p1
  vni 50000
  rd auto
    ip route 0.0.0.0/0 192.0.2.3
    ipv6 route ::/0 2001:DB8:1::1
```



Note The **ip route** and **ipv6 route** commands are specific to default route configuration.

Create a prefix list to enable IP prefix filtering

(config)#

```
ip prefix-list default-route seq 5 permit 0.0.0.0/0 le 1
ipv6 prefix-list default-route-v6 seq 5 permit ::/0
route-map DENY-DEFAULT-ROUTE deny 10
  match ip address prefix-list default-route
```

Create route maps to restrict default routes generated in the fabric from being distributed to external neighbors

(config)#

```
route-map DENY-DEFAULT-ROUTE permit 1000
route-map DENY-DEFAULT-ROUTE-V6 deny 100
  match ipv6 address prefix-list default-route-v6
route-map DENY-DEFAULT-ROUTE-V6 permit 1000
route-map FABRIC-RMAP-REDIST-STATIC permit 10
  match ip address prefix-list default-route
route-map FABRIC-RMAP-REDIST-STATIC-V6 permit 100
  match ipv6 address prefix-list default-route-v6
```

The **DENY-DEFAULT-ROUTE** route map is created to restrict default routes from being sent to external networks

Enable BGP configuration and apply appropriate route maps

(config)#

```
router bgp 65536
  vrf vni-50000:p1
    address-family ipv4 unicast
      redistribute static route-map FABRIC-RMAP-REDIST-STATIC
      default-information originate
    address-family ipv6 unicast
      redistribute static route-map FABRIC-RMAP-REDIST-STATIC-V6
      default-information originate
    neighbor 192.0.2.3 remote-as 65551
    address-family ipv4 unicast
      send-community both
      route-map DENY-DEFAULT-ROUTE out
    address-family ipv6 unicast
      send-community both
      route-map DENY-DEFAULT-ROUTE-V6 out
```



Note The **redistribute static route-map**, **default-information originate**, and **route-map** commands are specific to default route configuration.

Some pointers for the above configurations are given below:

- The configured next hop IPv4 static route 192.0.2.3 (and next hop IPv6 route 2001:DB8:1::1) of the 0/0 default route is the IP address of the WAN ASBR/PE local interface facing the border leaf switch. By pointing the static route towards the interface facing the WAN routers, it is ensured that the default route will be withdrawn if the interface(s) facing the WAN routers are down.
- The default IP route configuration needs to be enabled only on the border leaf switch nodes, and not on any other leaf switch.
- The **redistribute static route-map** and **default-information originate** commands enable distribution of IPv4 and IPv6 default routes to ToR/leaf switches.
- *65551* is the BGP autonomous system (AS) ID of the WAN.