

VXLAN MP-BGP EVPN High-Level Config Steps

Tuesday, February 21, 2023 18:20

SECTION 0 - FEATURES CONFIG:

For Spine & Leaf:

- Step 1: Enable the OSPF feature
- Step 2: Enable the PIM feature
- Step 3: Enable the BGP feature
- Step 4: Enable the NV Overlay feature
- Step 5: Enable NV Overlay EVPN

For Leaf only:

- Step 6: Enable the VN Segment VLAN-based feature
- Step 7: Enable the Fabric Forwarding feature
- Step 8: Enable the Interface VLAN feature

For vPC (optional, only if needed) on Leaf only:

- Step 9: Enable the vPC feature.
- Step 10 (optional, only if needed): Enable the LACP feature.

*For 5600 leaf switches, the switch must be in store-and-forward mode, and you must also install and enable feature-set fabric and feature-set virtualization.

SECTION 1 - UNDERLAY CONFIG :

For Spine & Leaf:

- Step 1: Configure IGP (OSPF, IS-IS, EIGRP)
- Step 2: Configure Multicast (PIM/Bidir PIM)

For Spine only:

- Step 3: Configure Multicast Redundancy (Anycast RP or Phantom RP)

For Spine & Leaf:

- Step 4: Configure BGP (iBGP or eBGP; use L2VPN EVPN Address Family and send both community attributes; make spines route-reflectors)

*Don't forget to change the MTU size on fabric links!

SECTION 2 - OVERLAY CONFIG (L2):

For Leaf only:

- Step 1: Configure VLANs and map them to VNIs.
- Step 2: Configure downstream ports as access or trunk ports.
- Step 3: Configure NVE interface to use BGP for host reachability, and then configure its source interface.

Also, configure the required VNI members and necessary options such as multicast group and ARP suppression.

Step 4: Define the L2 VNIs under the EVPN configuration mode and auto-assign route distinguishers and route-targets.

SECTION 3 - OVERLAY CONFIG (L3):

For Leaf only:

Step 1: In all VTEPs, configure the required VRF(s) used for overlay traffic and define its L3 VNI; auto-assign route distinguishers and under the IPv4 address family, auto-assign route-targets.

Step 2: In all VTEPs, configure the VRF-shared VLAN used for L3 forwarding and associate it with the VRF's L3 VNI.

Step 3: In all VTEPs, configure the distributed anycast gateway MAC (same address on all VTEPs).

Step 4: In all VTEPs, configure the required SVIs, including the VRF-shared interface VLAN (defined in Step 2). Enable IP forwarding on the VRF-shared interface VLAN, and configure other SVIs in anycast gateway mode where needed. Optionally, add a route tag to the IP address configured on the anycast gateway SVIs.

Step 5: In all VTEPs, under the NVE interface, associate the L3 VNI defined in the VRF context configuration.

Step 6: In all VTEPs, redistribute the SVI subnets into the BGP process under the VRF address family (redistribute direct route-map under IPv4 address family).

*For 5600 leaf switches, the "advertise l2vpn evpn" command must also be used in Step 6.

SECTION 4 - vPC (Optional, only if needed):

For Leaf only:

Step 1: Configure the Anycast VTEP IP in both switches (secondary IP under NVE source interface).

Step 2: Configure the NVE peer-link VLAN in both vPC switches.

Step 3 (Optional, but recommended): Configure backup link across the NVE peer-link VLAN in both vPC switches (for underlay network).

Step 4 (Optional, only if needed): Configure routing across the peer link in both vPC switches (per overlay VRF, used for external routing).

SECTION 5 - MULTI-SITE (Optional, only if needed):

For BGW Leaf only:

Step 1: Define a loopback interface as the EVPN multi-site source interface (BGW VTEP), then advertise its IP internally in the site and to the neighboring BGW. Also, advertise the VTEP IP (lo0) to the neighboring BGW. Use the pre-existing underlay routing protocol for the advertisements.

Step 2: Define eBGP peering between the BGWs using site-specific autonomous systems. Increase the eBGP multi-hop count as needed, also be sure that the peer-type is fabric-external and under the L2VPN EVPN address family, rewrite the EVPN RT ASN.

Step 3: Define a leaf node in each site as an EVPN multi-site BGW using the appropriate site ID. Set the delay-restart time as needed.

Step 4: Specify EVPN multi-site interface tracking for the site-internal fabric links and site-external DCI links on the BGWs.

Step 5: Under the NVE1 interface, define the previously configured loopback interface (defined in Step 1) as the EVPN multi-site border-gateway interface, and for the extended VNIs make sure to suppress ARP and use multi-site ingress replication.

*Don't forget to change the MTU size on DCI links!