Configuring Proportional Multipath for VNF

This chapter contains the following sections:

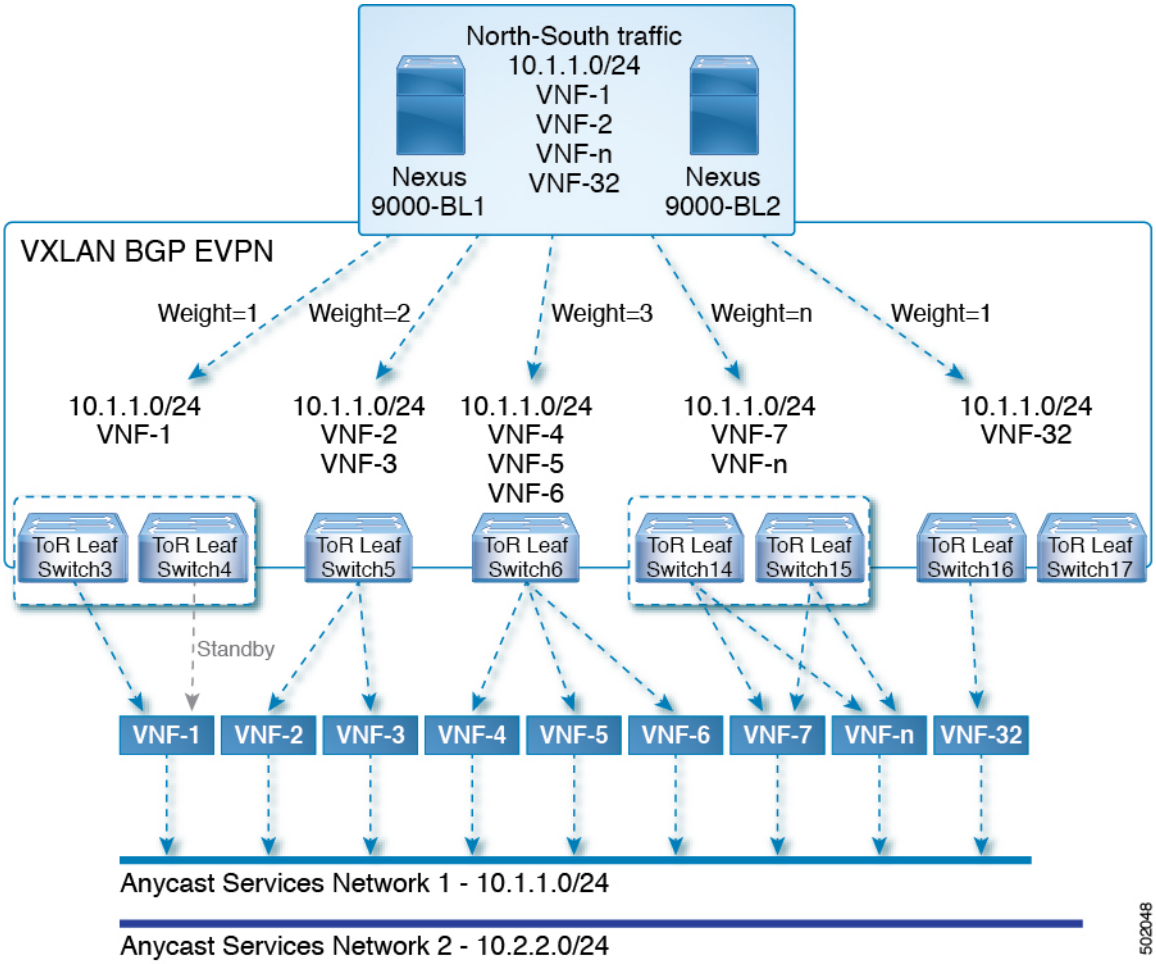
- About proportional multipath for VNF, on page 1
- Proportional Multipath for VNF with Multi-Site, on page 5
- Prerequisites for Proportional Multipath for VNF, on page 5
- Guidelines and limitations for proportional multipath for VNF, on page 6
- Configuring the Spine or Route Reflector, on page 7
- Configuring the Leaf or ToR, on page 8
- Configuring the Border Leaf or Border Gateway, on page 14
- Configuring the BGP Legacy Peer, on page 20
- Configuring a User-Defined Profile for Maintenance Mode, on page 21
- Configuring a User-Defined Profile for Normal Mode, on page 21
- Configuring a Default Route Map, on page 22
- Applying a Route Map to a Route Reflector, on page 23
- Verifying Proportional Multipath for VNF, on page 24
- Configuration Example for Proportional Multipath for VNF with Multi-Site, on page 27

About proportional multipath for VNF

In Network Function Virtualization Infrastructures (NFVi), anycast services networks are advertised from multiple Virtual Network Functions (VNFs). The Proportional Multipath for VNF feature enables advertising of all the available next hops to a given destination network. This feature enables the switch to consider all paths to a given route as equal cost multipath (ECMP) allowing the traffic to be forwarded using all the available links stretched across multiple ToRs.

Configuring Proportional Multipath for VNF

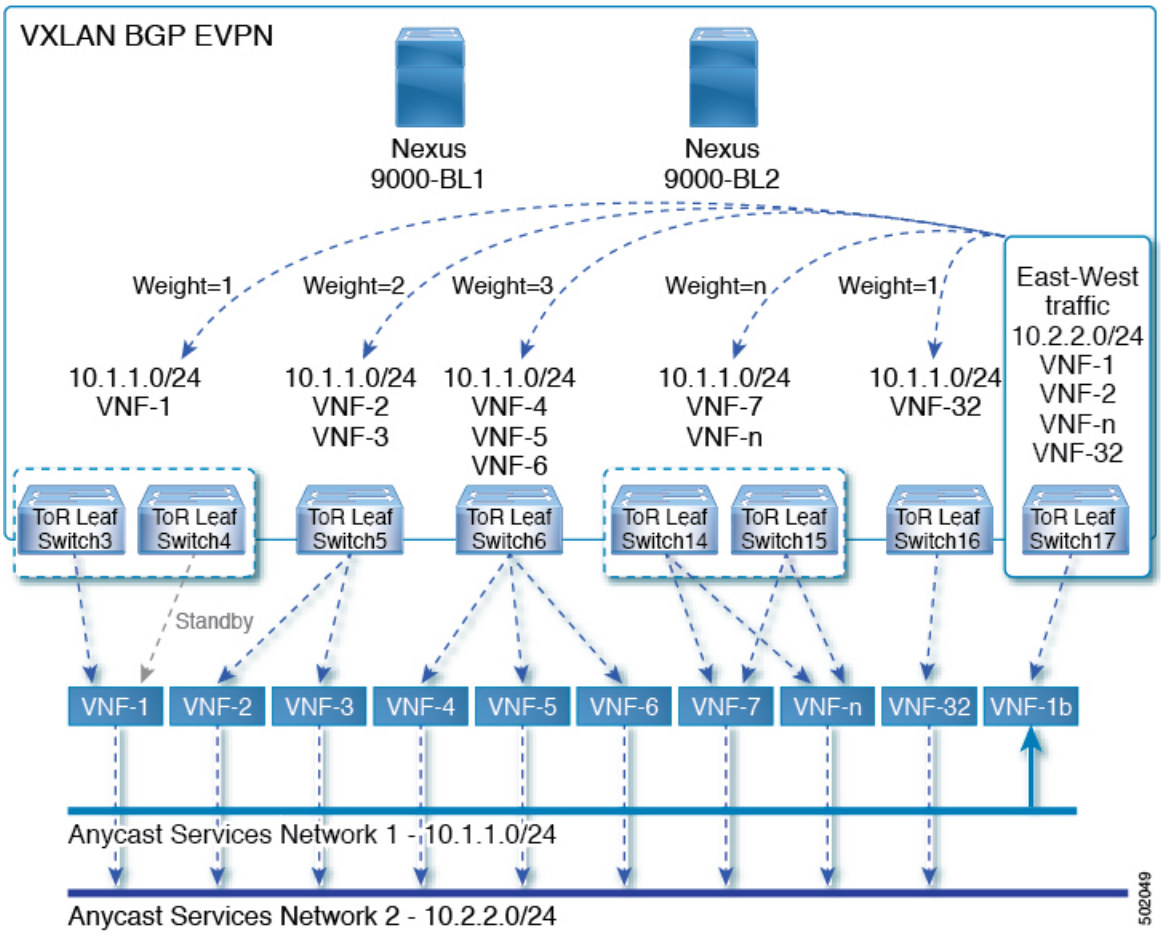
Figure 1: Sample Topology (North-South Traffic)



In the preceding diagram, North-South traffic that enters the VXLAN fabric at a border leaf is sent across all egress endpoints with the traffic forwarded proportional to the number of links from the egress top of rack (ToR) to the destination network.

Configuring Proportional Multipath for VNF

Figure 2: Sample Topology (East-West Traffic)



East-West traffic is forwarded between the VXLAN Tunnel Endpoints (VTEPs) proportional to the number of next hops advertised by each ToR switch to the destination network.

The switch uses BGP to advertise reachability within the fabric using the Layer 2 VPN (L2VPN)/Ethernet VPN (EVPN) address family. If all ToR switches and border leafs are within the same Autonomous System (AS), a full internal BGP (iBGP) mesh is configured by using route reflectors or by having each BGP router peer with every other router.

Each ToR and border leaf constitutes a VTEP in the VXLAN fabric. You can use a BGP route reflector to reduce the full mesh BGP sessions across the VTEPs to a single BGP session between a VTEP and the route reflector. Virtual Network Identifiers (VNIs) are globally unique within the overlay. Each Virtual Routing and Forwarding (VRF) instance is mapped to a unique VNI. The inner destination MAC address in the VXLAN header belongs to the receiving VTEP that does the routing of the VXLAN payload. This MAC address is distributed as a BGP attribute along with the EVPN routes.

Advertisement of customer networks

Customer networks are configured statically or learned locally by using an interior gateway protocol, (IGP) or external BGP (eBGP), over a Provider Edge(PE)-Customer Edge(CE) link. These networks are redistributed into BGP and advertised to the VXLAN fabric.

The networks advertised to the ToRs by the virtual machines (VMs) attached to them are advertised to the VXLAN fabric as EVPN Type-5 routes with the following:

- The route distinguisher (RD) will be the Layer 3 VNI's configured RD.
- The gateway IP field will be populated with the next hop.
- The next hop of the EVPN route will continue to be the VTEP IP.
- The export route targets of the routes will be derived from the configured export route targets of the associated Layer 3 VNI.

Multiple VRF routes may generate the same Type-5 Network Layer Reachability Information (NLRI) differentiated only by the gateway IP field. The routes are advertised with the L3VNI's RD, and the gateway IP isn't part of the Type-5 NLRI's key. The NLRI is exchanged between BGP routers using update messages. These routes are advertised to the EVPN AF by extending the BGP export mechanism to include ECMPs and using the addpath BGP feature in the EVPN AF.

Each Type-5 route within the EVPN AF that is created by using the Proportional Multipath for VNF feature may have multiple paths that are imported into the corresponding VRF based on the matching of the received route targets and by having ECMP enabled within the VRF and in the EVPN AF. Within the VRF, the route is a single prefix with multiple paths. Each path represents a Type-5 EVPN path or those learned locally within the VRF. The EVPN Type-5 routes that are enabled for the Proportional Multipath for VNF feature will have their next hop in the VRF derived from their gateway IP field. Use the **export-gateway-ip** command to enable BGP to advertise the gateway IP in the EVPN Type-5 routes.

Use the **maximum-paths mixed** command to enable BGP and the Unicast Routing Information Base (URIB) to consider the following paths as ECMP:

- iBGP paths
- eBGP paths
- Paths from other protocols (such as static) that are redistributed or injected into BGP

The paths can be either local to the device (static, iBGP, or eBGP) or remote (eBGP or iBGP learned over BGP-EVPN). This overrides the default route selection behavior in which local routes are preferred over remote routes. URIB downloads all next hops of the route, including locally learned and user-configured routes, to the Unicast FIB Distribution Module (uFDM)/Forwarding Information Base (FIB).

Beginning with Cisco NX-OS Release 9.3(5), you don't need to use mixed paths. You can choose to have only eBGP or iBGP filter the ECMP paths.

When you enter the **maximum-paths mixed** command beginning with Cisco NX-OS Release 9.3(5), BGP checks for the AS-path length by default. If you want to ignore the AS-path length (for example, on nodes that participate in packet forwarding such as BGWs and VTEPs), you must enter the **bestpath as-path ignore** command. When the **maximum-paths mixed** command is enabled for earlier releases, BGP ignores the AS-path length, and URIB ignores the administrative distance when choosing ECMPs. To ensure that no impact is observed, we recommend upgrading to Cisco NX-OS Release 9.3(5) prior to entering this command.

Legacy peer support

Use the **advertise-gw-ip** command to advertise EVPN Type-5 routes with the gateway IP set. ToRs then advertise the gateway IP in the Type-5 NLRI. However, legacy peers running on NX-OS version older than Cisco NX-OS Release 9.2(1) can't process the gateway IP which might lead to unexpected behavior. To prevent this scenario from occurring, use the **no advertise-gw-ip** command to disable the Proportional

Multipath for VNF feature for a legacy peer. BGP sets the gateway IP field of the Type-5 NLRI to zero even if the path being advertised has a valid gateway IP.

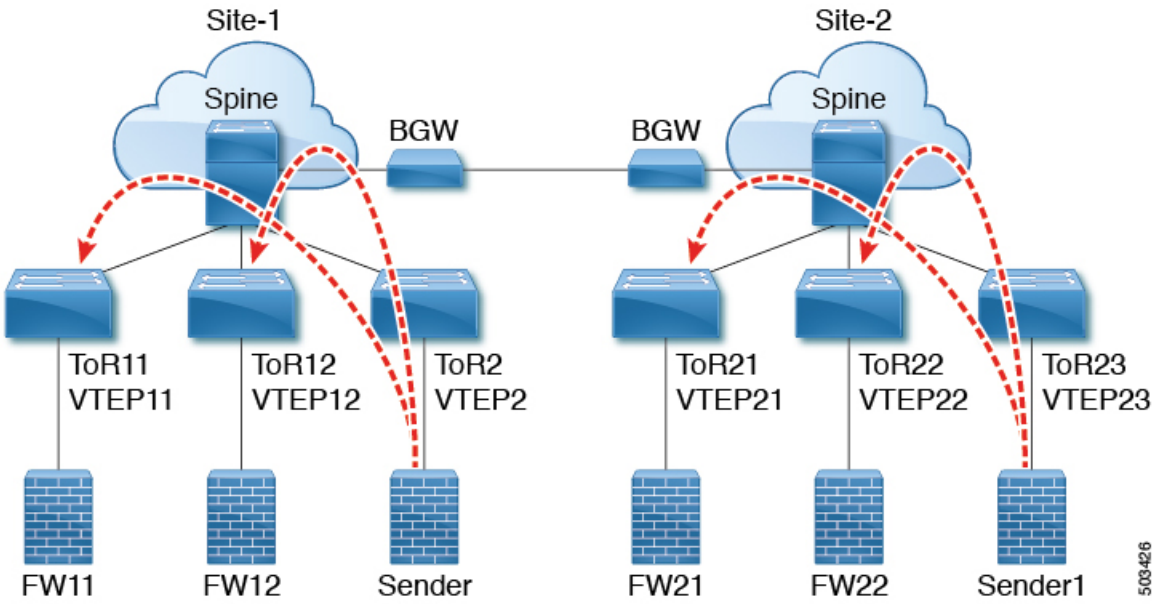
The **no advertise-gw-ip** command flaps the specified peer session as gracefully as possible. The remote peer triggers a graceful restart if the peer supports this capability. When the session is re-established, the local peer advertises EVPN Type-5 routes with the gateway IP set or with the gateway IP as zero depending on whether the **advertise-gw-ip** command has been used. By default, this knob is enabled and the gateway IP field is populated with the appropriate next hop value.

Proportional Multipath for VNF with Multi-Site

Cisco NX-OS Release 9.3(6) and later releases support Proportional Multipath for VNF with Multi-Site. This feature allows traffic to be sent across sites if a local VNF isn't available.

ToRs prefer to use local VNFs. However, if local VNFs aren't available, they can use VNFs in a different site. In the following topology, the ToRs in site 2 would use VNFs 21 and 22. However, if these VNFs aren't available, sender 1 in site 2 could send traffic to VNFs 11 and 12 in site 1.

Figure 3: Proportional Multipath for VNF with Multi-Site Topology



To use this feature, simply configure Proportional Multipath for VNF and enable Multi-Site. For a sample configuration, see Configuration Example for Proportional Multipath for VNF with Multi-Site, on page 27.

Prerequisites for Proportional Multipath for VNF

If desired, take the following actions before upgrading to Cisco NX-OS Release 9.3(5):

- Configure a route map for redistributed paths and use the **set ip next-hop redist-unchanged** command when using locally redistributed paths to export the gateway IP address. This command preserves the next hop for locally redistributed paths. For example:

```
route-map redist-rtmap permit 10
match ip prefix-list vm-pfx-list
set ip next-hop redist-unchanged
```

- Enter the **bestpath as-path ignore** command on nodes that participate in packet forwarding, such as BGWs and VTEPs. This command causes BGP to ignore the AS-path length.

Guidelines and limitations for proportional multipath for VNF

Proportional Multipath for VNF has the following guidelines and limitations:

- If the Proportional Multipath for VNF feature is enabled, maintenance mode isolation doesn't work because BGP installs all the paths in mixed multipath mode. Alternatively, a route-map is used to deny outbound BPG updates when a switch goes into maintenance mode by using user-defined profiles.
- This feature is supported for Cisco Nexus 9364C, 9300-EX, and 9300-FX/FX2/FX3 platform switches and Cisco Nexus 9500 platform switches with the N9K-C9508-FM-E2 fabric module and an -EX or -FX line card.
- Beginning with Cisco NX-OS Release 10.2(3)F, the Proportional Multipath for VNF feature is supported on Cisco Nexus 9300-GX/GX2B platform switches.
- Static and direct routes have to be redistributed into the BGP when the Proportional Multipath for VNF feature is enabled.
- If OSPF or EIGRP is being used as an IGP, routes can't be redistributed into BGP.
- If Proportional Multipath for VNF is enabled and routes aren't redistributed into BGP, asymmetric load balancing of traffic may occur as the local routes from URIB may not show up in BGP and on remote TORs as EVPN paths.
- Devices on which mixed-multipath is enabled must support the same load-balancing algorithm.
- If a VNF instance is multi-homed to multiple TORs, policies have to be configured or BGP routes have to be originated using a network command. As a result, each TOR connection to the VNF is displayed in the BGP routing table. Each TOR can now see the VNF's direct routes to the other TORs in which the VNF is multi-homed. Consequently, each TOR can advertise paths to the Gateway IPs through other TORs leading to a next hop resolution loop.

  Consider a scenario in which a VNF is multi-homed to two TORs, TOR1 and TOR2. Individual links to the TORs are addressed as 1.1.1.1 and 2.2.2.2. If the VNF advertises a service 192.168.1.0/24 through the TORs, the TORs advertise EVPN routes to 192.168.1.0/24 with Gateway IPs of 1.1.1.1 and 2.2.2.2 respectively.

  As a result, an issue occurs with the Recursive Next Hop (RNH) resolution on a remote TOR (for example, TOR3). The gateway IP is resolved to a /24 route pointing to another gateway IP. That second gateway IP is resolved by a route pointing to the first gateway IP. So, in our scenario, the gateway IP 1.1.1.1 is resolved by 1.1.1.0/24 which points to 2.2.2.2. And 2.2.2.2 is resolved by 2.2.2.0/24 which points to 1.1.1.1.

  This condition occurs as both TORs connected to the VNF are advertising the VNF’s connected routes. TOR1 is advertising 1.1.1.0/24 and 2.2.2.0/24. However, 1.1.1.0 is advertised without a gateway IP as it's a connected subnet on TOR1. Also, 2.2.2.0 is an OSPF route pointing to 1.1.1.1 which is the VNF’s address connected to TOR1.

Similarly, TOR2 advertises both subnets and 2.2.2.0/24 is sent without a gateway IP as it is directly connected to TOR2. 1.1.1.0 is learned via OSPF and is sent with a gateway IP of 2.2.2.2 which is the VNF's address connected to TOR2. 1.1.1.1/32 and 2.2.2.2/32 won't be advertised as they are Adjacency Manager (AM) routes on each TOR.

This issue doesn't have a resolution when Type-5 routes are involved. However, this scenario can be avoided if the TORs advertise the gateway IP's /32 address using a network command. And if the gateway IPs are being resolved by Type-2 EVPN MAC/IP routes, this scenario can be avoided as the gateway IP will be resolved by the /32 IP route.

- The following guidelines and limitations apply to Proportional Multipath for VNF with Multi-Site:
  - This feature is supported for Cisco Nexus 9364C, 9300-EX, and 9300-FX/FX2/FX3 platform switches and Cisco Nexus 9500 platform switches with the N9K-C9508-FM-E2 fabric module and an -EX or -FX line card.
  - VNF moves across sites aren't supported.
- Proportional multipath with max-path mixed configuration is not supported for VNFs attached to vPC leaf switches, and vPC is supported when the max-path mixed configuration is not used.

Configuring the Spine or Route Reflector

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *number*
3. **address-family l2vpn evpn**
4. **additional-paths send**
5. **additional-paths receive**
6. **additional-paths selection route-map passall**
7. **route-map passall permit** *seq-num*
8. **set path-selection all advertise**

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter global configuration mode.
Step 2	**router bgp** *number* **Example:** `switch(config)# router bgp 2`	Configure BGP.

	Command or Action	Purpose
Step 3	**address-family l2vpn evpn** **Example:** `switch(config-router)# address-family l2vpn evpn`	Configure address family Layer 2 VPN EVPN under **router bgp** context.
Step 4	**additional-paths send** **Example:** `switch(config-router-af)# additional-paths send`	The additional-paths configuration for sending..
Step 5	**additional-paths receive** **Example:** `switch(config-router-af)# additional-paths receive`	The additional-paths configuration for receiving.
Step 6	**additional-paths selection route-map passall** **Example:** `switch(config-router-af)# additional-paths selection route-map passall`	The additional-paths configuration applied the route map.
Step 7	**route-map passall permit** *seq-num* **Example:** `switch(config)# route-map passall permit 10`	Configure the route map.
Step 8	**set path-selection all advertise** **Example:** `switch(config-route-map)# set path-selection all advertise`	Sets the route-map related to the additional-paths feature.

Configuring the Leaf or ToR

This procedure describes how to configure the ToR.

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *number*
3. **address-family l2vpn evpn**
4. [**no**] **maximum-paths** [*eBGP max-paths* |**mixed** | **ibgp** |**local** | **eibgp**] *mpath-count*
5. **additional-paths send**
6. **additional-paths receive**
7. **additional-paths selection route-map passall**
8. **exit**
9. **vrf** *vrf-name*
10. **address-family ipv4 unicast**
11. **export-gateway-ip**

12. [**no**] **maximum-paths** [*eBGP max-paths* |**mixed** | **ibgp** |**local** | **eibgp**] *mpath-count*
13. **redistribute static route-map** *route-map-name*
14. **maximum-paths local** *number*
15. **exit**
16. **address-family ipv6 unicast**
17. **export-gateway-ip**
18. [**no**] **maximum-paths** [*eBGP max-paths* |**mixed** | **ibgp** |**local** | **eibgp**] *mpath-count*
19. **redistribute static route-map** *route-map-name*
20. **maximum-paths local** *number*
21. **exit**
22. **route-map passall permit** *seq-num*
23. **set path-selection all advertise**
24. **ip load-sharing address source-destination rotate** *rotate* **universal-id** *seed*

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter global configuration mode.
Step 2	**router bgp** *number* **Example:** `switch(config)# router bgp 2`	Configure BGP.
Step 3	**address-family l2vpn evpn** **Example:** `switch(config-router)# address-family l2vpn evpn`	Configure address family Layer 2 VPN EVPN under **router bgp** context.

Step 4	[**no**] **maximum-paths** [*eBGP max-paths* |**mixed**	**ibgp** |**local**	**eibgp**] *mpath-count*
			Example:
			`switch(config-router-af)# maximum-paths ?`
			`  <1-64>  Number of parallel paths`
			`  *Default value is 1`
			`  eibgp   Configure multipath for both EBGP and IBGP paths`
			`  ibgp    Configure multipath for IBGP paths`
			`  local   Configure multipath for local paths`
			`  mixed   Configure multipath for local and remote paths`
			`switch(config-router-af)# maximum-paths mixed 32`
			Example:
				The following options are available:
				• *eBGP max-path*–Enables the eBGP maximum paths. The range is from 1 to 64 parallel paths. The default value is 1.
				• **mixed**–Enables BGP and the Unicast Routing Information Base (URIB) to consider the following paths as Equal Cost Multi Path (ECMP):
				  • eBGP paths
				  • eiBGP paths
				  • iBGP paths

	Command or Action	Purpose
	`switch(config-router-af)# maximum-paths ibgp 32`	• Paths from other protocols (such as static) that are redistributed or injected into BGP • **ibgp**–Uses iBGP to filter the ECMP paths. • **local**–Enables the multipath for local paths. • If you enter the command without the **mixed** or **ibgp** option, eBGP is used to filter the ECMP paths. **Note** Use the **no** form of this command if you want to use a single path instead of maximum paths.
Step 5	**additional-paths send** **Example:** `switch(config-router-af)# additional-paths send`	The additional-paths configuration for sending.
Step 6	**additional-paths receive** **Example:** `switch(config-router-af)# additional-paths receive`	The additional-paths configuration for receiving.
Step 7	**additional-paths selection route-map passall** **Example:** `switch(config-router-af)# additional-paths selection route-map passall`	The additional-paths configuration applied the route map.
Step 8	**exit** **Example:** `switch(config-router-af)# exit`	Exits command mode.
Step 9	**vrf** *vrf-name* **Example:** `switch(config-router)# vrf evpn-tenant-1001`	Switch to the VRF configuration mode.
Step 10	**address-family ipv4 unicast** **Example:** `switch(config-router)# address-family ipv4 unicast`	Configure address family for IPv4.
Step 11	**export-gateway-ip** **Example:** `switch(config-router-vrf-af)# export-gateway-ip`	Enables BGP to advertise the gateway IP in the EVPN Type-5 routes. It exports the gateway IP for all prefixes in that VRF. **Note** If you want choose specific prefixes for which to export the gateway IP, use the following configuration instead of the **export-gateway-ip** command:

	Command or Action	Purpose
		route-map *name* permit *sequence* match ip address prefix-list *name* set evpn gateway-ip use-next-hop vrf context *vrf* address-family ipv4 unicast export map *name*
Step 12	[**no**] **maximum-paths** [*eBGP max-paths* |**mixed**	**ibgp** |**local**	**eibgp**] *mpath-count* **Example:** switch(config-router-vrf-af)# maximum-paths ? <1-64> Number of parallel paths *Default value is 1 eibgp Configure multipath for both EBGP and IBGP paths ibgp Configure multipath for IBGP paths local Configure multipath for local paths mixed Configure multipath for local and remote paths switch(config-router-vrf-af)# **maximum-paths mixed 32** **Example:** switch(config-router-vrf-af)# **maximum-paths ibgp 32**	The following options are available: • *eBGP max-path*–Enables the eBGP maximum paths. The range is from 1 to 64 parallel paths. The default value is 1. • **mixed**–Enables BGP and the Unicast Routing Information Base (URIB) to consider the following paths as Equal Cost Multi Path (ECMP): • eBGP paths • eiBGP paths • iBGP paths • Paths from other protocols (such as static) that are redistributed or injected into BGP • **ibgp**–Uses iBGP to filter the ECMP paths. • **local**–Enables the multipath for local paths. • If you enter the command without the **mixed** or **ibgp** option, eBGP is used to filter the ECMP paths. **Note** Use the **no** form of this command if you want to use a single path instead of maximum paths.
Step 13	**redistribute static route-map** *route-map-name* **Example:** switch(config-router-vrf-af)# **redistribute static route-map redist-rtmap**	Preserves the next-hop of the redistributed paths.
Step 14	**maximum-paths local** *number* **Example:** switch(config-router-vrf-af)# **maximum-paths local 32**	Specifies the number of local paths to be redistributed as the BGP best path for a route. The range is from 0 to 32. The default value is 1. **Note** This command isn't supported with the **maximum-paths mixed** *mpath-count* command. An error message appears if you try to configure them together. **Note**

	Command or Action	Purpose
		The **set ip next-hop redist-unchanged** command is required in order for the **maximum-paths local** command to work.
Step 15	**exit** **Example:** `switch(config-router-vrf-af)# exit`	Exits command mode.
Step 16	**address-family ipv6 unicast** **Example:** `switch(config-router-vrf)# address-family ipv6 unicast`	Configure address family for IPv6.

Step 17	**export-gateway-ip**
	Example:
	`switch(config-router-vrf-af)# export-gateway-ip`
		Enables BGP to advertise the gateway IP in the EVPN Type-5 routes. It exports the gateway IP for all prefixes in that VRF.
		Note
		If you want choose specific prefixes for which to export the gateway IP, use the following configuration instead of the **export-gateway-ip** command:
		`route-map name permit sequence`
		`  match ip address prefix-list name`
		`  set evpn gateway-ip use-next-hop`
		
		`vrf context vrf`
		`  address-family ipv4 unicast`
		`  export map name`

Step 18	[**no**] **maximum-paths** [*eBGP max-paths* |**mixed** | **ibgp** |**local** | **eibgp**] *mpath-count*
	Example:
	`switch(config-router-vrf-af)# maximum-paths ?`
	`  <1-64>  Number of parallel paths`
	`  *Default value is 1`
	`  eibgp  Configure multipath for both EBGP and IBGP paths`
	`  ibgp  Configure multipath for IBGP paths`
	`  local  Configure multipath for local paths`
	`  mixed  Configure multipath for local and remote paths`
	`switch(config-router-vrf-af)# maximum-paths mixed 32`
	Example:
	`switch(config-router-vrf-af)# maximum-paths ibgp 32`
		The following options are available:
		• *eBGP max-path*–Enables the eBGP maximum paths. The range is from 1 to 64 parallel paths. The default value is 1.
		• **mixed**–Enables BGP and the Unicast Routing Information Base (URIB) to consider the following paths as Equal Cost Multi Path (ECMP):
		  • eBGP paths
		  • eiBGP paths
		  • iBGP paths
		  • Paths from other protocols (such as static) that are redistributed or injected into BGP
		• **ibgp**–Uses iBGP to filter the ECMP paths.
		• **local**–Enables the multipath for local paths.

	Command or Action	Purpose
		• If you enter the command without the **mixed** or **ibgp** option, eBGP is used to filter the ECMP paths. **Note** Use the **no** form of this command if you want to use a single path instead of maximum paths.
Step 19	**redistribute static route-map** *route-map-name* **Example:** `switch(config-router-vrf-af)# redistribute static route-map redist-rtmap`	Preserves the next-hop of the redistributed paths.
Step 20	**maximum-paths local** *number* **Example:** `switch(config-router-vrf-af)# maximum-paths local 32`	Specifies the number of local paths to be redistributed as the BGP best path for a route. The range is from 0 to 32. The default value is 1. **Note** This command isn't supported with the **maximum-paths mixed** *mpath-count* command. An error message appears if you try to configure them together.
Step 21	**exit** **Example:** `switch(config-router-vrf-af)# exit`	Exits command mode.
Step 22	**route-map passall permit** *seq-num* **Example:** `switch(config)# route-map passall permit 10`	Configure the route map.
Step 23	**set path-selection all advertise** **Example:** `switch(config-route-map)# set path-selection all advertise`	Sets the route-map related to the additional-paths feature.
Step 24	**ip load-sharing address source-destination rotate** *rotate* **universal-id** *seed* **Example:** `ip load-sharing address source-destination rotate 32 universal-id 1`	Configures the unicast FIB load-sharing algorithm for data traffic. • The **universal-id** option sets the random seed for the hash algorithm and shifts the flow from one link to another. You do not need to configure the universal ID. Cisco NX-OS chooses the Universal ID if you do not configure it. The *seed* range is from 1 to 4294967295. • The **rotate** option causes the hash algorithm to rotate the link picking selection so that it does not continually choose the same link across all nodes in the network. It does so by influencing the bit pattern

	Command or Action	Purpose
		for the hash algorithm. This option shifts the flow from one link to another and load balances the already load-balanced (polarized) traffic from the first ECMP level across multiple links. If you specify a **rotate** value, the 64-bit stream is interpreted starting from that bit position in a cyclic rotation. The **rotate** range is from 1 to 63, and the default is 32. **Note** With multi-tier Layer 3 topology, polarization is possible. To avoid polarization, use a different rotate bit at each tier of the topology. **Note** To configure a rotation value for port channels, use the **port-channel load-balance src-dst ip-l4port rotate** *rotate* command. For more information on this command, see the *Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 9.x*.

Configuring the Border Leaf or Border Gateway

This procedure describes how to configure the border leaf.

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *number*
3. **address-family l2vpn evpn**
4. [**no**] **maximum-paths** [*eBGP max-paths* |**mixed** | **ibgp** |**local** | **eibgp**] *mpath-count*
5. **additional-paths send**
6. **additional-paths receive**
7. **additional-paths selection route-map passall**
8. **exit**
9. **vrf** *vrf-name*
10. **address-family ipv4 unicast**
11. **export-gateway-ip**
12. [**no**] **maximum-paths** [*eBGP max-paths* |**mixed** | **ibgp** |**local** | **eibgp**] *mpath-count*
13. **redistribute static route-map** *route-map-name*
14. **maximum-paths local** *number*
15. **address-family ipv6 unicast**
16. **export-gateway-ip**
17. [**no**] **maximum-paths** [*eBGP max-paths* |**mixed** | **ibgp** |**local** | **eibgp**] *mpath-count*

18. **redistribute static route-map** *route-map-name*
19. **maximum-paths local** *number*
20. **exit**
21. **route-map passall permit** *seq-num*
22. **set path-selection all advertise**
23. **ip load-sharing address source-destination rotate** *rotate* **universal-id** *seed*

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter global configuration mode.
Step 2	**router bgp** *number* **Example:** `switch(config)# router bgp 2`	Configure BGP.
Step 3	**address-family l2vpn evpn** **Example:** `switch(config-router)# address-family l2vpn evpn`	Configure address family Layer 2 VPN EVPN under **router bgp** context.

Step 4	[**no**] **maximum-paths** [*eBGP max-paths* |**mixed**	**ibgp** |**local**	**eibgp**] *mpath-count*
			Example:
			`switch(config-router-af)# maximum-paths ?`
			`  <1-64>  Number of parallel paths`
			`  *Default value is 1`
			`  eibgp  Configure multipath for both EBGP and IBGP paths`
			`  ibgp  Configure multipath for IBGP paths`
			`  local  Configure multipath for local paths`
			`  mixed  Configure multipath for local and remote paths`
			`switch(config-router-af)# maximum-paths mixed 32`
			Example:
			`switch(config-router-af)# maximum-paths ibgp 32`
				The following options are available:
				• *eBGP max-path*–Enables the eBGP maximum paths. The range is from 1 to 64 parallel paths. The default value is 1.
				• **mixed**–Enables BGP and the Unicast Routing Information Base (URIB) to consider the following paths as Equal Cost Multi Path (ECMP):
				  • eBGP paths
				  • eiBGP paths
				  • iBGP paths
				  • Paths from other protocols (such as static) that are redistributed or injected into BGP
				• **ibgp**–Uses iBGP to filter the ECMP paths.
				• **local**–Enables the multipath for local paths.
				• If you enter the command without the **mixed** or **ibgp** option, eBGP is used to filter the ECMP paths.

	Command or Action	Purpose
		Note Use the **no** form of this command if you want to use a single path instead of maximum paths.
Step 5	**additional-paths send** **Example:** `switch(config-router-af)# additional-paths send`	The additional-paths configuration for sending.
Step 6	**additional-paths receive** **Example:** `switch(config-router-af)# additional-paths receive`	The additional-paths configuration for receiving.
Step 7	**additional-paths selection route-map passall** **Example:** `switch(config-router-af)# additional-paths selection route-map passall`	The additional-paths configuration enables the additional-paths feature.
Step 8	**exit** **Example:** `switch(config-router-af)# exit`	Exits command mode.
Step 9	**vrf** *vrf-name* **Example:** `switch(config-router)# vrf evpn-tenant-1001`	Switch to the VRF configuration mode.
Step 10	**address-family ipv4 unicast** **Example:** `switch(config-router)# address-family ipv4 unicast`	Configure address family for IPv4.

Step 11	**export-gateway-ip**
	Example:
	`switch(config-router-vrf-af)# export-gateway-ip`
		Enables BGP to advertise the gateway IP in the EVPN Type-5 routes. It exports the gateway IP for all prefixes in that VRF.
		Note
		If you want choose specific prefixes for which to export the gateway IP, use the following configuration instead of the **export-gateway-ip** command:
		`route-map name permit sequence`
		`  match ip address prefix-list name`
		`  set evpn gateway-ip use-next-hop`
		
		`vrf context vrf`
		`  address-family ipv4 unicast`
		`  export map name`

	Command or Action	Purpose

Step 12	[**no**] **maximum-paths** [*eBGP max-paths* |**mixed**	**ibgp** |**local**	**eibgp**] *mpath-count*
			Example:
			`switch(config-router-af)# maximum-paths ?`
			`  <1-64>  Number of parallel paths`
			`  *Default value is 1`
			`  eibgp   Configure multipath for both EBGP and IBGP paths`
			`  ibgp    Configure multipath for IBGP paths`
			`  local   Configure multipath for local paths`
			`  mixed   Configure multipath for local and remote paths`
			`switch(config-router-vrf-af)# maximum-paths mixed 32`
			Example:
			`switch(config-router-vrf-af)# maximum-paths ibgp 32`
				The following options are available:
				• *eBGP max-path*–Enables the eBGP maximum paths. The range is from 1 to 64 parallel paths. The default value is 1.
				• **mixed**–Enables BGP and the Unicast Routing Information Base (URIB) to consider the following paths as Equal Cost Multi Path (ECMP):
				  • eBGP paths
				  • eiBGP paths
				  • iBGP paths
				  • Paths from other protocols (such as static) that are redistributed or injected into BGP
				• **ibgp**–Uses iBGP to filter the ECMP paths.
				• **local**–Enables the multipath for local paths.
				• If you enter the command without the **mixed** or **ibgp** option, eBGP is used to filter the ECMP paths.
				Note
				Use the **no** form of this command if you want to use a single path instead of maximum paths.

Step 13	**redistribute static route-map** *route-map-name* **Example:** `switch(config-router-vrf-af)# redistribute static route-map redist-rtmap`	Preserves the next-hop of the redistributed paths.
Step 14	**maximum-paths local** *number* **Example:** `switch(config-router-vrf-af)# maximum-paths local 32`	Specifies the number of local paths to be redistributed as the BGP best path for a route. The range is from 0 to 32. The default value is 1. **Note** This command isn't supported with the **maximum-paths mixed** *mpath-count* command. An error message appears if you try to configure them together.
Step 15	**address-family ipv6 unicast** **Example:** `switch(config-router-vrf)# address-family ipv6 unicast`	Configure address family for IPv6.

	Command or Action	Purpose

Step 16	**export-gateway-ip**
	Example:
	`switch(config-router-vrf-af)# export-gateway-ip`
		Enables BGP to advertise the gateway IP in the EVPN Type-5 routes. It exports the gateway IP for all prefixes in that VRF.
		Note
		If you want choose specific prefixes for which to export the gateway IP, use the following configuration instead of the **export-gateway-ip** command:
		`route-map name permit sequence`
		`  match ip address prefix-list name`
		`  set evpn gateway-ip use-next-hop`
		
		`vrf context vrf`
		`  address-family ipv4 unicast`
		`  export map name`

Step 17	[**no**] **maximum-paths** [*eBGP max-paths* |**mixed**	**ibgp** |**local**	**eibgp**] *mpath-count*
			Example:
			`switch(config-router-vrf-af)# maximum-paths ?`
			`  <1-64>  Number of parallel paths`
			`  *Default value is 1`
			`  eibgp  Configure multipath for both EBGP and IBGP paths`
			`  ibgp  Configure multipath for IBGP paths`
			`  local  Configure multipath for local paths`
			`  mixed  Configure multipath for local and remote paths`
			`switch(config-router-vrf-af)# maximum-paths mixed 32`
			Example:
			`switch(config-router-vrf-af)# maximum-paths ibgp 32`
				The following options are available:
				• *eBGP max-path*–Enables the eBGP maximum paths. The range is from 1 to 64 parallel paths. The default value is 1.
				• **mixed**–Enables BGP and the Unicast Routing Information Base (URIB) to consider the following paths as Equal Cost Multi Path (ECMP):
				  • eBGP paths
				  • eiBGP paths
				  • iBGP paths
				  • Paths from other protocols (such as static) that are redistributed or injected into BGP
				• **ibgp**–Uses iBGP to filter the ECMP paths.
				• **local**–Enables the multipath for local paths.
				• If you enter the command without the **mixed** or **ibgp** option, eBGP is used to filter the ECMP paths.
				Note
				Use the **no** form of this command if you want to use a single path instead of maximum paths.

Step 18	**redistribute static route-map** *route-map-name* **Example:** `switch(config-router-vrf-af)# redistribute static route-map redist-rtmap`	Preserves the next-hop of the redistributed paths.

	Command or Action	Purpose
Step 19	**maximum-paths local** *number* **Example:** `switch(config-router-vrf-af)# maximum-paths local 32`	Specifies the number of local paths to be redistributed as the BGP best path for a route. The range is from 0 to 32. The default value is 1. **Note** This command isn't supported with the **maximum-paths mixed** *mpath-count* command. An error message appears if you try to configure them together.
Step 20	**exit** **Example:** `switch(config-router-vrf-af)# exit`	Exits command mode.
Step 21	**route-map passall permit** *seq-num* **Example:** `switch(config)# route-map passall permit 10`	Configure the route map.
Step 22	**set path-selection all advertise** **Example:** `switch(config-route-map)# set path-selection all advertise`	Sets the route-map related to the additional-paths feature.
Step 23	**ip load-sharing address source-destination rotate** *rotate* **universal-id** *seed* **Example:** `ip load-sharing address source-destination rotate 32 universal-id 1`	Configures the unicast FIB load-sharing algorithm for data traffic. • The **universal-id** option sets the random seed for the hash algorithm and shifts the flow from one link to another. You do not need to configure the universal ID. Cisco NX-OS chooses the Universal ID if you do not configure it. The *seed* range is from 1 to 4294967295. • The **rotate** option causes the hash algorithm to rotate the link picking selection so that it does not continually choose the same link across all nodes in the network. It does so by influencing the bit pattern for the hash algorithm. This option shifts the flow from one link to another and load balances the already load-balanced (polarized) traffic from the first ECMP level across multiple links. If you specify a **rotate** value, the 64-bit stream is interpreted starting from that bit position in a cyclic rotation. The **rotate** range is from 1 to 63, and the default is 32. **Note**

	Command or Action	Purpose
		With multi-tier Layer 3 topology, polarization is possible. To avoid polarization, use a different rotate bit at each tier of the topology. **Note** To configure a rotation value for port channels, use the **port-channel load-balance src-dst ip-l4port rotate** *rotate* command. For more information on this command, see the *Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 9.x*.

Configuring the BGP Legacy Peer

If you are running a Cisco Nexus Release prior to 9.2(1), follow this procedure to disable sending the gateway IP address to that peer.

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *number*
3. **neighbor** *address* **remote-as** *number*
4. **address-family l2vpn evpn**
5. **no advertise-gw-ip**

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter global configuration mode.
Step 2	**router bgp** *number* **Example:** `switch(config)# router bgp 2000000`	Configure BGP.
Step 3	**neighbor** *address* **remote-as** *number* **Example:** `switch(config-router)# neighbor 8.8.8.8 remote-as 2000000`	Define neighbor.
Step 4	**address-family l2vpn evpn** **Example:**	Configure address family Layer 2 VPN EVPN.

	Command or Action	Purpose
	`switch(config-router-neighbor)# address-family l2vpn evpn`	
Step 5	**no advertise-gw-ip** **Example:** `switch(config-router-neighbor-af)# no advertise-gw-ip`	Disables the BGP EVPN Mixed-path and Proportional Layer-3 Multipath feature for a legacy peer.

Configuring a User-Defined Profile for Maintenance Mode

SUMMARY STEPS

1. **configure terminal**
2. **configure maintenance profile maintenance-mode**
3. **route-map** *name* **deny** *sequence*

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter global configuration mode.
Step 2	**configure maintenance profile maintenance-mode** **Example:** `switch(config)# configure maintenance profile maintenance-mode`	Configure maintenance mode profile.
Step 3	**route-map** *name* **deny** *sequence* **Example:** `switch(config-mm-profile)# route-map GIR deny 5`	Configure route map. The value of *sequence* is from 0 to 65535. Default is 10.

Configuring a User-Defined Profile for Normal Mode

SUMMARY STEPS

1. **configure terminal**
2. **configure maintenance profile normal-mode**
3. **route-map** *name* **permit** *sequence*

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter global configuration mode.
Step 2	**configure maintenance profile normal-mode** **Example:** `switch(config)# configure maintenance profile normal-mode`	Configure maintenance mode.
Step 3	**route-map** *name* **permit** *sequence* **Example:** `switch(config-mm-profile)# route-map GIR permit 5`	Configure route map. The value of *sequence* is from 0 to 65535. Default is 10.

Configuring a Default Route Map

SUMMARY STEPS

1. **configure terminal**
2. **route-map** *name* **permit** *sequence*

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter global configuration mode.
Step 2	**route-map** *name* **permit** *sequence* **Example:** `switch(config-mm-profile)# route-map GIR permit 5`	Configure route map. The value of *sequence* is from 0 to 65535. Default is 10.

Applying a Route Map to a Route Reflector

SUMMARY STEPS

1. **configure terminal**
2. **router bgp** *number*
3. **neighbor** *ip-address*
4. **address-family l2vpn evpn**
5. **route-map** *name* **out**

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enter global configuration mode.
Step 2	**router bgp** *number* **Example:** `switch(config)# router bgp 2`	Configure BGP.
Step 3	**neighbor** *ip-address* **Example:** `switch(config-router)# neighbor 10.1.1.1`	Configure the IP address of a BGP neighbor which is the route reflector. *ip-address* can be an IPv4 or IPv6 address or prefix.
Step 4	**address-family l2vpn evpn** **Example:** `switch(config-router-neighbor)# address-family l2vpn evpn`	Configure a Layer 2 VPN EVPN address family.
Step 5	**route-map** *name* **out** **Example:** `switch(config-router-neighbor-af)# route-map GIR out`	Apply the route map to the neighbor route reflector.

Verifying Proportional Multipath for VNF

Command	Purpose
show bgp ipv4 unicast	Displays Border Gateway Protocol (BGP) information for the IPv4 unicast address family.
show bgp l2vpn evpn	Displays BGP information for the Layer-2 Virtual Private Network (L2VPN) Ethernet Virtual Private Network (EVPN) address family.
show ip route	Displays routes from the unicast RIB.
show maintenance profile maintenance-mode	Displays the GIR user-defined profile for the maintenance mode.
show maintenance profile normal-mode	Displays the GIR user-defined profile for the normal mode.

The following example shows how to display BGP information for the L2VPN EVPN address family:

```
switch# show bgp l2vpn evpn 11.1.1.0
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 13.13.13.13:3       //     Remote route
BGP routing table entry for [5]:[0]:[0]:[24]:[11.1.1.0]/224, version 1341
Paths: (3 available, best #1)
Flags: (0x000002) on xmit-list, is not in l2rib/evpn, is not in HW
Multipath: eBGP

  Advertised path-id 1
  Path type: external, path is valid, is best path
             Imported to 2 destination(s)
  Gateway IP: 11.1.1.133
  AS-Path: 2000000 100000 , path sourced external to AS
    11.11.11.11 (metric 5) from 102.102.102.102 (102.102.102.102)
      Origin incomplete, MED not set, localpref 100, weight 0
      Received label 22001
      Received path-id 3
      Extcommunity: RT:23456:22001 Route-Import:11.11.11.11:2001 ENCAP:8
          Router MAC:003a.7d7d.1dbd

  Path type: external, path is valid, not best reason: Neighbor Address, multipath
             Imported to 2 destination(s)
  Gateway IP: 11.1.1.233
  AS-Path: 2000000 100 , path sourced external to AS
    33.33.33.33 (metric 5) from 102.102.102.102 (102.102.102.102)
      Origin incomplete, MED not set, localpref 100, weight 0
      Received label 22001
      Received path-id 2
      Extcommunity: RT:23456:22001 Route-Import:33.33.33.33:2001 ENCAP:8
          Router MAC:e00e.da4a.589d

  Path type: external, path is valid, not best reason: Neighbor Address, multipath
             Imported to 2 destination(s)
  Gateway IP: 11.1.1.100
  AS-Path: 2000000 500000 , path sourced external to AS
```

```
    22.22.22.22 (metric 5) from 102.102.102.102 (102.102.102.102)
      Origin incomplete, MED not set, localpref 100, weight 0
      Received label 22001
      Received path-id 1
      Extcommunity: RT:23456:22001 Route-Import:22.22.22.22:2001 ENCAP:8
          Router MAC:e00e.da4a.62a5

  Path-id 1 not advertised to any peer

Route Distinguisher: 4.4.4.4:3    (L3VNI 22001)       //   Local L3VNI
BGP routing table entry for [5]:[0]:[0]:[24]:[11.1.1.0]/224, version 3465
Paths: (3 available, best #1)
Flags: (0x000002) on xmit-list, is not in l2rib/evpn, is not in HW
Multipath: eBGP

  Advertised path-id 1
  Path type: external, path is valid, is best path
             Imported from 13.13.13.13:3:[5]:[0]:[0]:[24]:[11.1.1.0]/224
  Gateway IP: 11.1.1.100
  AS-Path: 2000000 500000 , path sourced external to AS
    22.22.22.22 (metric 5) from 102.102.102.102 (102.102.102.102)
      Origin incomplete, MED not set, localpref 100, weight 0
      Received label 22001
      Received path-id 1
      Extcommunity: RT:23456:22001 Route-Import:22.22.22.22:2001 ENCAP:8
          Router MAC:e00e.da4a.62a5

  Path type: external, path is valid, not best reason: newer EBGP path, multipat
h
             Imported from 13.13.13.13:3:[5]:[0]:[0]:[24]:[11.1.1.0]/224
  Gateway IP: 11.1.1.233
  AS-Path: 2000000 100 , path sourced external to AS
    33.33.33.33 (metric 5) from 102.102.102.102 (102.102.102.102)
      Origin incomplete, MED not set, localpref 100, weight 0
      Received label 22001
      Received path-id 2
      Extcommunity: RT:23456:22001 Route-Import:33.33.33.33:2001 ENCAP:8
          Router MAC:e00e.da4a.589d

  Path type: external, path is valid, not best reason: newer EBGP path, multipat
h
             Imported from 13.13.13.13:3:[5]:[0]:[0]:[24]:[11.1.1.0]/224
  Gateway IP: 11.1.1.133
  AS-Path: 2000000 100000 , path sourced external to AS
    11.11.11.11 (metric 5) from 102.102.102.102 (102.102.102.102)
      Origin incomplete, MED not set, localpref 100, weight 0
      Received label 22001
      Received path-id 3
      Extcommunity: RT:23456:22001 Route-Import:11.11.11.11:2001 ENCAP:8
          Router MAC:003a.7d7d.1dbd

  Path-id 1 not advertised to any peer
```

The following example shows how to display BGP information for the IPv4 unicast address family:

```
switch# show bgp ipv4 unicast 11.1.1.0 vrf cust_1
BGP routing table information for VRF cust_1, address family IPv4 Unicast
BGP routing table entry for 11.1.1.0/24, version 4
Paths: (3 available, best #1)
Flags: (0x80080012) on xmit-list, is in urib, is backup urib route, is in HW
  vpn: version 1093, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
```

```
  Path type: external, path is valid, is best path, in rib
             Imported from 13.13.13.13:3:[5]:[0]:[0]:[24]:[11.1.1.0]/224
  AS-Path: 2000000 500000 , path sourced external to AS
    11.1.1.100 (metric 5) from 102.102.102.102 (102.102.102.102)
      Origin incomplete, MED not set, localpref 100, weight 0
      Received label 22001
      Received path-id 1
      Extcommunity: RT:23456:22001 Route-Import:22.22.22.22:2001 ENCAP:8
          Router MAC:e00e.da4a.62a5

  Path type: external, path is valid, not best reason: Neighbor Address, multipath, in rib
             Imported from 13.13.13.13:3:[5]:[0]:[0]:[24]:[11.1.1.0]/224
  AS-Path: 2000000 100 , path sourced external to AS
    11.1.1.233 (metric 5) from 102.102.102.102 (102.102.102.102)
      Origin incomplete, MED not set, localpref 100, weight 0
      Received label 22001
      Received path-id 2
      Extcommunity: RT:23456:22001 Route-Import:33.33.33.33:2001 ENCAP:8
          Router MAC:e00e.da4a.589d

  Path type: external, path is valid, not best reason: Neighbor Address, multipath, in rib
             Imported from 13.13.13.13:3:[5]:[0]:[0]:[24]:[11.1.1.0]/224
  AS-Path: 2000000 100000 , path sourced external to AS
    11.1.1.133 (metric 5) from 102.102.102.102 (102.102.102.102)
      Origin incomplete, MED not set, localpref 100, weight 0
      Received label 22001
      Received path-id 3
      Extcommunity: RT:23456:22001 Route-Import:11.11.11.11:2001 ENCAP:8
          Router MAC:003a.7d7d.1dbd

  VRF advertise information:
  Path-id 1 not advertised to any peer

  VPN AF advertise information:
  Path-id 1 not advertised to any peer
```

The following example shows how to display routes from the unicast RIB after the Proportional Multipath for VNF feature has been configured:

```
switch# show ip route 1.1.1.0 vrf cust_1
IP Route Table for VRF "cust_1"
…
1.1.1.0/24, ubest/mbest: 22/0, all-best (0x300003d)
    *via 3.0.0.1, [1/0], 08:13:17, static
         recursive next hop: 3.0.0.1/32
    *via 3.0.0.2, [1/0], 08:13:17, static
         recursive next hop: 3.0.0.2/32
    *via 3.0.0.3, [1/0], 08:13:16, static
         recursive next hop: 3.0.0.3/32
    *via 3.0.0.4, [1/0], 08:13:16, static
         recursive next hop: 3.0.0.4/32
   *via 2.0.0.1, [200/0], 06:09:19, bgp-2, internal, tag 2 (evpn) segid: 3003802 tunnelid:
 0x300003e encap: VXLAN
          BGP-EVPN: VNI=3003802 (EVPN)
         client-specific data: 3b
         recursive next hop: 2.0.0.1/32
         extended route information: BGP origin AS 2 BGP peer AS 2
   *via 2.0.0.2, [200/0], 06:09:19, bgp-2, internal, tag 2 (evpn) segid: 3003802 tunnelid:
 0x300003e encap: VXLAN
          BGP-EVPN: VNI=3003802 (EVPN)
         client-specific data: 3b
         recursive next hop: 2.0.0.2/32
         extended route information: BGP origin AS 2 BGP peer AS 2
```

The following example shows how to display the GIR user-defined profile for the maintenance mode:

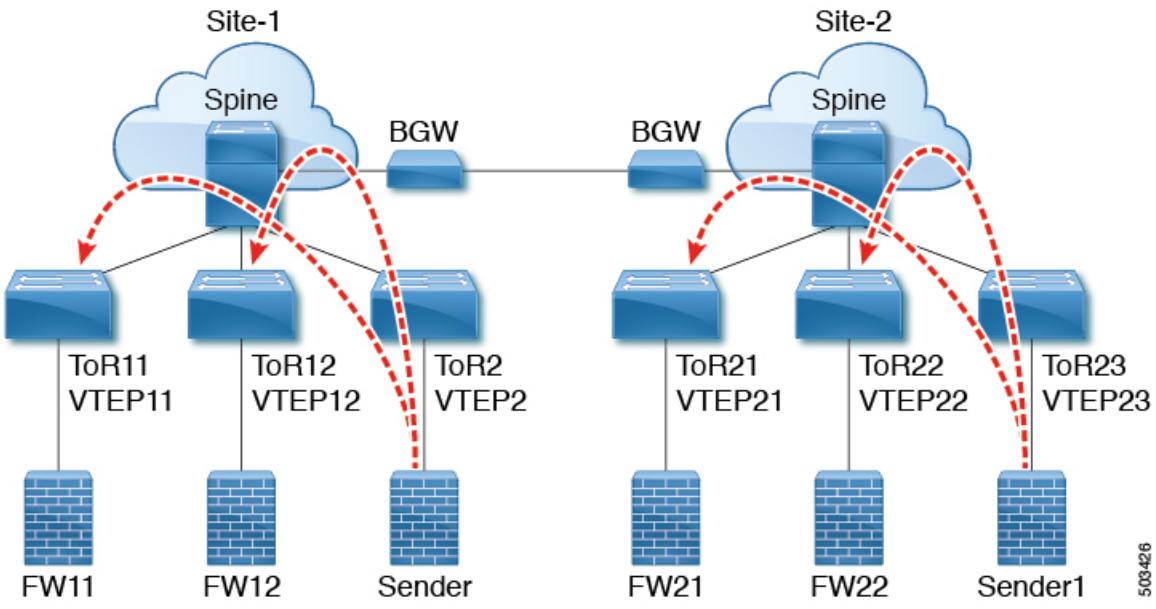
```
switch# show maintenance profile maintenance-mode
  [Maintenance Mode]
  ip pim isolate
  router bgp 2
   isolate
  router isis 1
   isolate
  route-map GIR deny 5
```

The following example shows how to display the GIR user-defined profile for the normal mode:

```
switch# show maintenance profile normal-mode
  [Normal Mode]
  no ip pim isolate
  router bgp 2
   no isolate
  router isis 1
   no isolate
  route-map GIR permit 5
```

Configuration Example for Proportional Multipath for VNF with Multi-Site

Figure 4: Proportional Multipath for VNF with Multi-Site Topology



The following configuration example allows traffic to be sent across sites if a local VNF isn't available.

```
feature telnet
feature nxapi
feature bash-shell
feature scp-server
nv overlay evpn
```

```
feature ospf
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature bfd
feature nv overlay

no password strength-check
username admin password 5 password role network-admin
ip domain-lookup
copp profile strict
evpn multisite border-gateway 1
  delay-restore time 30
snmp-server user admin network-admin auth md5 0x66a8185ad28d9df13d9214f6e19aad37 priv
0x66a8185ad28d9df13d9214f6e19aad37 localizedkey

fabric forwarding anycast-gateway-mac 0000.2222.3333
ip pim ssm range 232.0.0.0/8
vlan 1,14,24,100-110,120-150,1000-1010,1100-1110,2000-2010,2100-2110,3000-3010
vlan 100
  name l2-vni-vlan-0-for-vrf100
  vn-segment 2000100
vlan 101
  name l2-vni-vlan-0-for-vrf101
  vn-segment 2000101
vlan 1100
  name l2-vni-vlan-1-for-vrf100
  vn-segment 2001100
vlan 1101
  name l2-vni-vlan-1-for-vrf101
  vn-segment 2001101
vlan 2100
  name l3-vni-vlan-for-vrf100
  vn-segment 3000100
vlan 2101
  name l3-vni-vlan-for-vrf101
  vn-segment 3000101

route-map passall permit 10
  set path-selection all advertise
route-map permit-all permit 10
  set path-selection all advertise
route-map permit-all-v6 permit 10

vrf context vrf100
  vni 3000100
  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
vrf context vrf101
 vni 3000101
  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
```

```
interface Vlan14
  no shutdown
  vrf member vrf100
  ip address 192.14.0.1/24
  ipv6 address 192:14::1/64

interface Vlan24
  no shutdown
  vrf member vrf101
  ip address 192.24.0.1/24
  ipv6 address 192:24::1/64

interface Vlan100
  description "L3VRF.VLANNUM.0.222"
  no shutdown
  vrf member vrf100
  ip address 100.0.0.222/24
  ipv6 address 100::222/64
  fabric forwarding mode anycast-gateway

interface Vlan101
  description "L3VRF.VLANNUM.0.222"
  no shutdown
  vrf member vrf101
  ip address 101.0.0.222/24
  ipv6 address 101::222/64
  fabric forwarding mode anycast-gateway

interface Vlan1100
  description "L3VRF.VLANNUM.0.222"
  no shutdown
  vrf member vrf100
  ip address 100.1.0.222/16
  ipv6 address 100:1::222/64
  fabric forwarding mode anycast-gateway

interface Vlan1101
  description "L3VRF.VLANNUM.0.222"
  no shutdown
  vrf member vrf101
  ip address 101.1.0.222/16
  ipv6 address 101:1::222/64
  fabric forwarding mode anycast-gateway

interface Vlan2100
  no shutdown
  vrf member vrf100
  ip forward
  ipv6 address use-link-local-only

interface Vlan2101
  no shutdown
  vrf member vrf101
  ip forward
  ipv6 address use-link-local-only

interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback1
  multisite border-gateway interface loopback2
  member vni 2000100-2000110
    suppress-arp
    mcast-group 227.1.1.1
```

```
  member vni 2000120-2000150
    suppress-arp
    mcast-group 227.1.1.1
  member vni 2001100-2001110
    suppress-arp
    mcast-group 227.1.1.1
  member vni 3000100-3000110 associate-vrf
  member vni 3100100-3100110 associate-vrf

interface Ethernet1/22
  description "BGW11 to BGW2"
  medium p2p
  ip unnumbered loopback0
  ip ospf cost 40
  ip ospf network point-to-point
  ip router ospf 12 area 0.0.0.0
  no shutdown
  evpn multisite dci-tracking

interface Ethernet1/25
  description "BGW11 to Spine11"
  medium p2p
  ip unnumbered loopback0
  ip ospf cost 40
  ip ospf network point-to-point
  ip router ospf 1 area 0.0.0.0
  no shutdown
  evpn multisite fabric-tracking

interface Ethernet1/27
  description "BGW11 to Spine12"
  medium p2p
  ip unnumbered loopback0
  ip ospf cost 40
  ip ospf network point-to-point
  ip router ospf 1 area 0.0.0.0
  no shutdown
  evpn multisite fabric-tracking

interface Ethernet1/34
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 14,24
  no shutdown

interface loopback0
  ip address 1.1.11.0/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode

interface loopback1
  ip address 1.1.11.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode

interface loopback2
  ip address 11.11.11.11/32
  ip router ospf 12 area 0.0.0.0
  ip pim sparse-mode

router ospf 1
  redistribute direct route-map permit-all
router ospf 12
  redistribute direct route-map permit-all
```

```
ip load-sharing address source-destination rotate 32 universal-id 1

router bgp 1
  log-neighbor-changes
  address-family l2vpn evpn
    maximum-paths 8
    maximum-paths ibgp 8
    additional-paths send
    additional-paths receive
    additional-paths selection route-map passall
  neighbor 1.2.11.1
    remote-as 1
    description "SPINE-11"
    update-source loopback1
    address-family l2vpn evpn
      send-community extended
  neighbor 1.2.12.1
    remote-as 1
    description "SPINE-12"
    update-source loopback1
    address-family l2vpn evpn
      send-community extended
  neighbor 2.1.2.1
    remote-as 2
    description "BGW-2"
    update-source loopback1
    ebgp-multihop 3
    peer-type fabric-external
    address-family ipv4 unicast
    address-family l2vpn evpn
      send-community extended
      rewrite-evpn-rt-asn
  vrf vrf100
    address-family ipv4 unicast
      redistribute direct route-map permit-all
      maximum-paths 8
      maximum-paths ibgp 8
      export-gateway-ip
    address-family ipv6 unicast
      redistribute direct route-map permit-all
      maximum-paths 8
      maximum-paths ibgp 8
      export-gateway-ip
  vrf vrf101
    address-family ipv4 unicast
      redistribute direct route-map permit-all
      maximum-paths 8
      maximum-paths ibgp 8
      export-gateway-ip
    address-family ipv6 unicast
      redistribute direct route-map permit-all
      maximum-paths 8
      maximum-paths ibgp 8
      export-gateway-ip
evpn
  vni 2000100 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2000101 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2001100 l2
```

```
    rd auto
    route-target import auto
    route-target export auto
  vni 2001101 l2
    rd auto
    route-target import auto
    route-target export auto
```

The following example shows that the VTEP in site 1 prefers the local VNF (FW).

```
leaf1# show bgp l2vpn evpn 200.100.1.1
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 1.3.12.0:3
BGP routing table entry for [5]:[0]:[0]:[32]:[200.100.1.1]/224, version 77902
Paths: (4 available, best #2)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW
Multipath: eBGP iBGP Local

  Path type: internal, path is valid, not best reason: Neighbor Address, no labeled nexthop

  Gateway IP: 100.0.0.12
  AS-Path: 99 100 , path sourced external to AS
    1.3.12.1 (metric 81) from 1.2.12.1 (1.2.12.0)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 3000100
      Received path-id 2
      Extcommunity: RT:1:3000100 ENCAP:8 Router MAC:00be.7547.13bf
      Originator: 1.3.12.0 Cluster list: 1.2.12.0

  Advertised path-id 2
  Path type: local, path is valid, not best reason: Locally originated, multipath, no labeled
 nexthop
  Gateway IP: 100.0.0.11
  AS-Path: 99 100 , path sourced external to AS
    1.3.11.1 (metric 0) from 0.0.0.0 (1.3.11.0)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 3000100
      Received path-id 1
      Extcommunity: RT:1:3000100 ENCAP:8 Router MAC:d478.9bb3.c1a1
```

The following example shows how the local VNF is disabled so that the VNF from site 2 is used. The BGP adjacency is shut down between site 1's VTEP11 to FW11 and between VTEP12 to FW12.

```
leaf1(config-router)# vrf vrf100
leaf1(config-router-vrf)# neighbor 100::11
leaf1(config-router-vrf-neighbor)# shut
leaf1(config-router-vrf-neighbor)# neighbor 100::12
leaf1(config-router-vrf-neighbor)# shut
leaf1(config-router-vrf-neighbor)# neighbor 100:1::11
leaf1(config-router-vrf-neighbor)# shut
leaf1(config-router-vrf-neighbor)# neighbor 100:1::12
leaf1(config-router-vrf-neighbor)# shut
leaf1(config-router-vrf-neighbor)# neighbor 100.0.0.11
leaf1(config-router-vrf-neighbor)# shut
leaf1(config-router-vrf-neighbor)# neighbor 100.0.0.12
leaf1(config-router-vrf-neighbor)# shut
leaf1(config-router-vrf-neighbor)# neighbor 100.1.0.11
leaf1(config-router-vrf-neighbor)# shut
leaf1(config-router-vrf-neighbor)# neighbor 100.1.0.12
leaf1(config-router-vrf-neighbor)# shut
leaf1(config-router-vrf-neighbor)# end
```

The following example shows that the prefix now uses the VNF (FW) from site 2.

```
leaf1# show bgp l2vpn evpn 200.100.1.1
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 1:3000100
BGP routing table entry for [5]:[0]:[0]:[32]:[200.100.1.1]/224, version 97269
Paths: (3 available, best #3)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW
Multipath: eBGP iBGP Local

  Path type: internal, path is valid, not best reason: Neighbor Address, no labeled nexthop

  Gateway IP: 100.1.0.21
  AS-Path: 2 99 100 , path sourced external to AS
    11.11.11.11 (metric 20) from 1.2.12.1 (1.2.12.0)
      Origin IGP, MED 2000, localpref 100, weight 0
      Received label 3000100
      Received path-id 2
      Extcommunity: RT:1:3000100 SOO:03030100:00000000 ENCAP:8
          Router MAC:0200.0b0b.0b0b
      Originator: 1.1.12.0 Cluster list: 1.2.12.0
```

Configuring Proportional Multipath for VNF