



Configuring MPLS Layer 3 VPN Load Balancing

This chapter describes how to configure load balancing for Multiprotocol Label Switching (MPLS) Layer 3 virtual private networks (VPNs) on Cisco NX-OS devices.

This chapter includes the following sections:

- [Finding Feature Information, page 22-398](#)
- [Information About MPLS Layer 3 VPN Load Balancing, page 22-398](#)
- [Licensing Requirements for MPLS Layer 3 VPN Load Balancing, page 22-404](#)
- [Prerequisites for MPLS Layer 3 VPN Load Balancing, page 22-404](#)
- [Guidelines and Limitations for MPLS Layer 3 VPN Load Balancing, page 22-404](#)
- [Default Settings for MPLS Layer 3 VPN Load Balancing, page 22-405](#)
- [Configuring MPLS Layer 3 VPN Load Balancing, page 22-405](#)
- [Verifying the MPLS Layer 3 VPN Load-Balancing Configuration, page 22-410](#)
- [Configuration Examples for MPLS Layer 3 VPN Load Balancing, page 22-410](#)
- [Additional References for MPLS Layer 3 VPN Load Balancing, page 22-411](#)
- [Feature History for MPLS Layer 3 VPN Load Balancing, page 22-412](#)

Finding Feature Information

Your software release might not support all the features documented in this module. For the latest caveats and feature information, see the Bug Search Tool at <https://tools.cisco.com/bugsearch/> and the release notes for your software release. To find information about the features documented in this module, and to see a list of the releases in which each feature is supported, see the “New and Changed Information” chapter or the Feature History table below.

Information About MPLS Layer 3 VPN Load Balancing

Load balancing distributes traffic so that no individual router is overburdened. In an MPLS Layer 3 network, you can achieve load balancing by using the Border Gateway Protocol (BGP). When multiple iBGP paths are installed in a routing table, a route reflector advertises only one path (next hop). If a router is behind a route reflector, all routes that are connected to multihomed sites are not advertised

unless a different route distinguisher is configured for each virtual routing and forwarding instance (VRF). (A route reflector passes learned routes to neighbors so that all iBGP peers do not need to be fully meshed.)

This section includes the following topics:

- [iBGP Load Balancing, page 22-399](#)
- [eBGP Load Balancing, page 22-399](#)
- [Layer 3 VPN Load Balancing, page 22-399](#)
- [BGP VPNv4 Multipath, page 22-401](#)

iBGP Load Balancing

When a BGP-speaking router configured with no local policy receives multiple network layer reachability information (NLRI) from the internal BGP (iBGP) for the same destination, the router chooses one iBGP path as the best path and installs the best path in its IP routing table. iBGP load balancing enables the BGP-speaking router to select multiple iBGP paths as the best paths to a destination and to install multiple best paths in its IP routing table.

eBGP Load Balancing

When a router learns two identical eBGP paths for a prefix from a neighboring autonomous system, it chooses the path with the lower route ID as the best path. The router installs this best path in the IP routing table. You can enable eBGP load balancing to install multiple paths in the IP routing table when the eBGP paths are learned from a neighboring autonomous system instead of picking one best path.

During packet switching, depending on the switching mode, the router performs either per-packet or per-destination load balancing among the multiple paths.

Layer 3 VPN Load Balancing

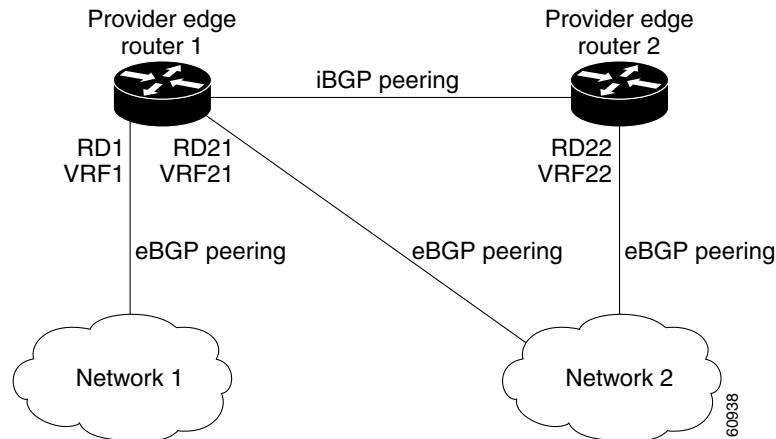
Layer 3 VPN load balancing for both eBGP and iBGP allows you to configure multihomed autonomous systems and provider edge (PE) routers to distribute traffic across both external BGP (eBGP) and iBGP multipaths.

Layer 3 VPN load balancing supports IPv4 and IPv6 for the PE routers and VPNs.

BGP installs up to the maximum number of multipaths allowed. BGP uses the best path algorithm to select one path as the best path, inserts the best path into the routing information base (RIB) and advertises the best path to BGP peers. The router can insert other paths into the RIB but selects only one path as the best path.

Layer 3 VPNs load balance on a per-packet or per-source or destination pair basis. To enable load balancing, configure the router with Layer 3 VPNs that contain VPN routing and forwarding instances (VRFs) that import both eBGP and iBGP paths. You can configure the number of paths separately for each VRF.

[Figure 22-1](#) shows an MPLS provider network that uses BGP. In the figure, two remote networks are connected to PE1 and PE2, which are both configured for VPN unicast iBGP peering. Network 2 is a multihomed network that is connected to PE1 and PE2. Network 2 also has extranet VPN services configured with Network 1. Both Network 1 and Network 2 are configured for eBGP peering with the PE routers.

Figure 22-1 Provider MPLS Network Using BGP

You can configure PE1 so that it can select both iBGP and eBGP paths as multipaths and import these paths into the VPN routing and forwarding instance (VRF) of Network 1 to perform load balancing.

Traffic is distributed as follows:

- IP traffic that is sent from Network 2 to PE1 and PE2 is sent across the eBGP paths as IP traffic.
- IP traffic that is sent from PE1 to PE2 is sent across the iBGP path as MPLS traffic.
- Traffic that is sent across an eBGP path is sent as IP traffic.

Any prefix that is advertised from Network 2 will be received by PE1 through route distinguisher (RD) 21 and RD22.

- The advertisement through RD21 is carried in IP packets.
- The advertisement through RD22 is carried in MPLS packets.

The router can select both paths as multipaths for VRF1 and insert these paths into the VRF1 RIB.

Layer 3 VPN Load Balancing with Route Reflectors

Route reflectors reduce the number of sessions on PE routers and increase the scalability of Layer 3 VPN networks. Route reflectors hold on to all received VPN routes to peer with PE routers. Different PEs can require different route target-tagged VPNv4 and VPNv6 routes. The route reflector may also need to send a refresh for a specific route target to a PE when the VRF configuration has changed. Storing all routes increases the scalability requirements on a route reflector. You can configure a route reflector to only hold routes that have a defined set of route target communities.

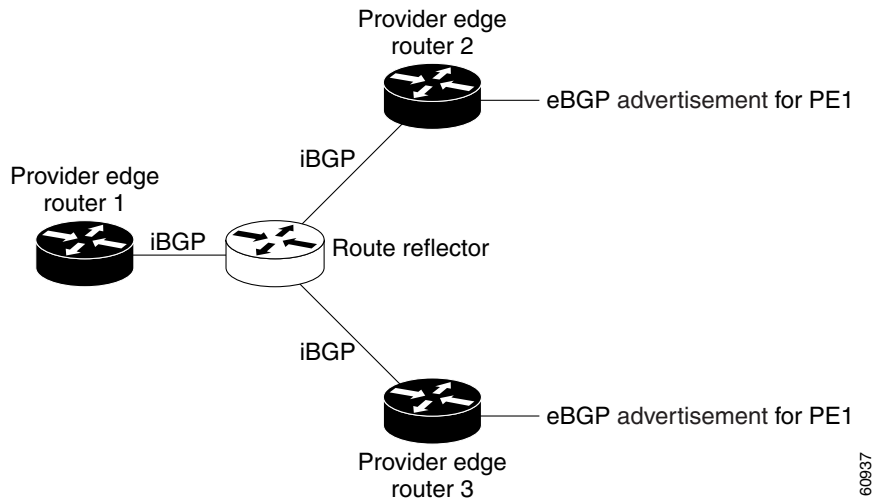
You can configure route reflectors to service a different set of VPNs and configure a PE to peer with all route reflectors that service the VRFs configured on the PE. When you configure a new VRF with a route target that the PE does not already hold routes for, the PE issues route refreshes to the route reflectors and retrieves the relevant VPN routes.



Note

The route reflectors do not need to be in the forwarding path, but you must configure unique route distinguisher (RDs) for VPN sites that are multihomed.

Figure 22-2 shows a topology that contains three PE routers and a route reflector, all configured for iBGP peering. PE2 and PE3 each advertise an equal preference eBGP path to PE1. By default, the route reflector chooses only one path and advertises PE1.

Figure 22-2 Topology with a Route Reflector

For all equal preference paths to PE1 to be advertised through the route reflector, you must configure each VRF with a different RD. The prefixes received by the route reflector are recognized differently and advertised to PE1.

Layer 2 Load Balancing Coexistence

The load balance method that is required in the Layer 2 VPN is different from the method that is used for Layer 3 VPN. Layer 3 VPN and Layer 2 VPN forwarding is performed independently using two different types of adjacencies. The forwarding is not impacted by using a different method of load balancing for the Layer 2 VPN.



Note

Load balancing is not supported at the ingress PE for Layer 2 VPNs.

BGP VPNv4 Multipath

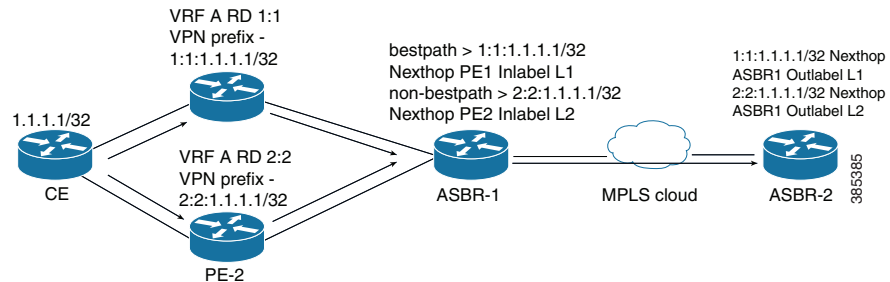
Starting from Cisco NX-OS Release 8.0(1), use the BGP VPNv4 Multipath feature to achieve Equal Cost Multi-Path (ECMP) for traffic flowing from an Autonomous System Border Router (ASBR) towards the Provider Edge (PE) device in an Multi-Protocol Label Switching (MPLS) cloud network by using a lower number of prefixes and MPLS labels. This feature configures the maximum number of multipaths for both eBGP and iBGP paths. This feature can be configured on PE devices and Route Reflectors in an MPLS topology.

Consider a scenario in which a dual homed Customer Edge (CE) device is connected to 2 PE devices and you have to utilize both the PE devices for traffic flow from ASBR-2 to the CE device.

Currently, as shown in Figure 22-3, Virtual Routing and Forwarding (VRF) on each PE is configured using separate Route Distinguishers (RD). The CE device generates a BGP IPv4 prefix. The PE devices are configured with 2 separate RDs and generate two different VPN-IPv4 prefixes for the BGP IPv4 prefix sent by the CE device. ASBR-1 receives both the VPN-IPv4 prefixes and adds them to the routing

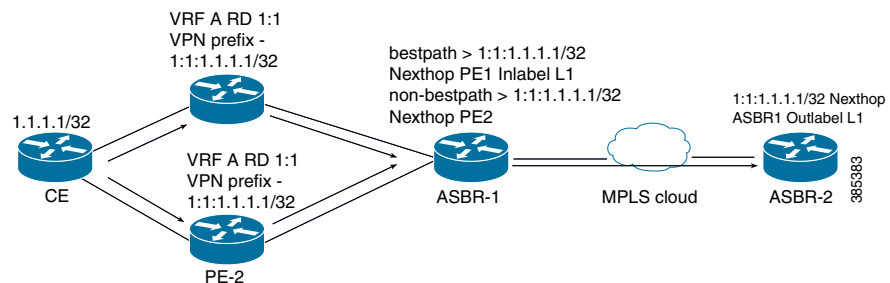
table. ASBR-1 allocates Inter-AS option-B labels, Inlabel L1 and Inlabel L2, to both the VPN routes and then advertises both VPN routes to ASBR-2. To use both PE devices to maintain traffic flow, ASBR-1 has to utilize two Inter-AS option-B labels and two prefixes which limits the scale that can be supported.

Figure 22-3 Virtual Routing and Forwarding (VRF) on each PE configured using separate Route Distinguishers



Using the BGP VPN Multipath feature, as shown in Figure 22-4, you can enable the VRF on both PE devices to use the same RD. In such a scenario, ASBR-1 receives the same prefix from both the PE devices. ASBR-1 allocates only one Inter-AS option-B label, Inlabel L1, to the received prefix and advertises the VPN route to ASBR-2. In this case, the scale is enhanced as traffic flow using both PE devices is established with only one prefix and label on ASBR-1.

Figure 22-4 Enabling the VRF on both PE devices to use the same RD



BGP Cost Community

The BGP cost community is a nontransitive extended community attribute that is passed to iBGP and confederation peers but not to eBGP peers. (A confederation is a group of iBGP peers that use the same autonomous system number to communicate to external networks.) The BGP cost community attributes includes a cost community ID and a cost value. You can customize the BGP best path selection process for a local autonomous system or confederation by configuring the BGP cost community attribute. You configure the cost community attribute in a route map with a community ID and cost value. BGP prefers the path with the lowest community ID, or for identical community IDs, BGP prefers the path with the lowest cost value in the BGP cost community attribute.

BGP uses the best path selection process to determine which path is the best where multiple paths to the same destination are available. You can assign a preference to a specific path when multiple equal cost paths are available.

Since the administrative distance of iBGP is worse than the distance of most Interior Gateway Protocols (IGPs), the unicast Routing Information Base (RIB) may apply the same BGP cost community compare algorithm before using the normal distance or metric comparisons of the protocol or route. VPN routes that are learned through iBGP can be preferred over locally learned IGP routes.

The cost extended community attribute is propagated to iBGP peers when an extended community exchange is enabled.

How the BGP Cost Community Influences the Best Path Selection Process

The cost community attribute influences the BGP best path selection process at the point of insertion (POI). The POI follows the IGP metric comparison. When BGP receives multiple paths to the same destination, it uses the best path selection process to determine which path is the best path. BGP automatically makes the decision and installs the best path into the routing table. The POI allows you to assign a preference to a specific path when multiple equal cost paths are available. If the POI is not valid for local best path selection, the cost community attribute is silently ignored.

You can configure multiple paths with the cost community attribute for the same POI. The path with the lowest cost community ID is considered first. All of the cost community paths for a specific POI are considered, starting with the one with the lowest cost community ID. Paths that do not contain the cost community (for the POI and community ID being evaluated) are assigned with the default community cost value.

Applying the cost community attribute at the POI allows you to assign a value to a path originated or learned by a peer in any part of the local autonomous system or confederation. The router can use the cost community as a tie breaker during the best path selection process. You can configure multiple instances of the cost community for separate equal cost paths within the same autonomous system or confederation. For example, you can apply a lower cost community value to a specific exit path in a network with multiple equal cost exits points, and the BGP best path selection process prefers that specific exit path.

Cost Community and EIGRP PE-CE with Back-Door Links

BGP prefers back-door links in an Enhanced Interior Gateway Protocol (EIGRP) Layer 3 VPN topology if the back-door link is learned first. A back-door link, or a route, is a connection that is configured outside of the Layer 3 VPN between a remote and main site.

The pre-best path point of insertion (POI) in the BGP cost community supports mixed EIGRP Layer 3 VPN network topologies that contain VPN and back-door links. This POI is applied automatically to EIGRP routes that are redistributed into BGP. The pre-best path POI carries the EIGRP route type and metric. This POI influences the best-path calculation process by influencing BGP to consider this POI before any other comparison step.

**Note**

You can configure BGP path selection to ignore the cost community but this is not recommended for configuration with a back-door link.

Licensing Requirements for MPLS Layer 3 VPN Load Balancing

Product	License Requirement
Cisco NX-OS	MPLS Layer 3 VPN load balancing requires an MPLS license. For a complete explanation of the Cisco NX-OS licensing scheme and how to obtain and apply licenses, see the <i>Cisco NX-OS Licensing Guide</i> .

Prerequisites for MPLS Layer 3 VPN Load Balancing

MPLS Layer 3 VPN load balancing has the following prerequisites:

- You must enable the MPLS and L3VPN features.
- You must install the correct license for MPLS.

Guidelines and Limitations for MPLS Layer 3 VPN Load Balancing

MPLS Layer 3 VPN load balancing has the following configuration guidelines and limitations:

- If you place a router behind a route reflector and it is connected to multihomed sites, the router will not be advertised unless separate VRFs with different RDs are configured for each VRF.
- Each IP routing table entry for a BGP prefix that has multiple iBGP paths uses additional memory. We recommend that you do not use this feature on a router with a low amount of available memory or when it is carrying a full Internet routing table.
- You should not ignore the BGP cost community when a back-door link is present and EIGRP is the PE-CE routing protocol.
- Starting with Cisco NX-OS Release 8.0(1), a maximum of 500K VPN prefixes is supported on M3-Series I/O modules. For M2-Series I/O modules, a maximum of 125K VPN prefixes is supported.
- Starting with Cisco NX-OS Release 8.0(1), 4K VRFs are supported.

Default Settings for MPLS Layer 3 VPN Load Balancing

Table 22-1 lists the default settings for MPLS Layer 3 VPN load balancing parameters.

Table 22-1 *Default MPLS Layer 3 VPN Load Balancing Parameters*

Parameters	Default
Layer 3 VPN feature	Disabled
BGP cost community ID	128
BGP cost community cost	2147483647
maximum multipaths	1
BGP VPNv4 Multipath	Disabled

Configuring MPLS Layer 3 VPN Load Balancing

This section includes the following topics:

- [Configuring BGP Load Balancing for eBGP and iBGP, page 22-405](#)
- [Configuring BGPv4 Multipath, page 22-407](#)

Configuring BGP Load Balancing for eBGP and iBGP

You can configure a Layer 3 VPN load balancing for an eBGP or iBGP network.

Prerequisites

Ensure that you are in the correct VDC (or use the **switchto vdc** command).

SUMMARY STEPS

1. **configure terminal**
2. **feature-set mpls**
3. **feature mpls l3vpn**
4. **feature bgp**
5. **router bgp** *as-number*
6. (Optional) **bestpath cost-community ignore**
7. **address-family {ipv4 | ipv6} unicast**
8. **maximum-paths [ibgp] number-of-paths**
9. (Optional) **show running-config bgp**
10. (Optional) **copy running-config startup-config**

DETAILED STEPS

	Command	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	feature-set mpls Example: switch(config)# feature-set mpls	Enables the MPLS feature-set.
Step 3	feature mpls l3vpn Example: switch(config)# feature mpls l3vpn	Enables the Layer 3 VPN feature.
Step 4	feature bgp Example: switch(config)# feature bgp	Enables the BGP feature.
Step 5	router bgp <i>as-number</i> Example: switch(config)# router bgp 1.1 switch(config-router)#	Configures a BGP routing process and enters router configuration mode. The <i>as-number</i> argument indicates the number of an autonomous system that identifies the router to other BGP routers and tags the routing information passed along. The AS number can be a 16-bit integer or a 32-bit integer in the form of a higher 16-bit decimal number and a lower 16-bit decimal number in xx.xx format.
Step 6	bestpath cost-community ignore Example: switch(config-router)# bestpath cost-community ignore#	(Optional) Ignores the cost community for BGP bestpath calculations.
Step 7	address-family {ipv4 ipv6} unicast Example: switch(config-router)# address-family ipv4 unicast switch(config-router-af)#	Enters address family configuration mode for configuring IP routing sessions.
Step 8	maximum-paths [<i>ibgp</i>] <i>number-of-paths</i> Example: switch(config-router-af)# maximum-paths 4	Configures the maximum number of multipaths allowed. Use the ibgp keyword to configure iBGP load balancing. The range is from 1 to 16.
Step 9	show running-config bgp Example: switch(config-router-af)# show running-config bgp	(Optional) Displays the running configuration for BGP.
Step 10	copy running-config startup-config Example: switch(config-router-af)# copy running-config startup-config	(Optional) Copies the running configuration to the startup configuration.

Configuring BGPv4 Multipath

SUMMARY STEPS

1. **configure terminal**
2. **feature bgp**
3. **router bgp** *as-number*
4. **address-family vpnv4 unicast**
5. **maximum-paths eibgp** *parallel-paths*

DETAILED STEPS

	Command	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	feature bgp Example: switch(config)# feature bgp	Enables the BGP feature.
Step 3	router bgp <i>as-number</i> Example: switch(config)# router bgp 2 switch(config-router)#	Assigns an autonomous system (AS) number to a router and enter the router BGP configuration mode.
Step 4	address-family vpnv4 unicast Example: switch(config-router)# address-family vpnv4 unicast switch(config-router-af)#	Enters address family configuration mode for configuring routing sessions, such as BGP, that use standard VPNv4 address prefixes:
Step 5	maximum-paths eibgp <i>parallel-paths</i> Example: switch(config-router-af)# maximum-paths eibgp 3	Specifies the maximum number of BGP VPNv4 multipaths for both eBGP and iBGP paths. The range is from 1 to 32.

Configuring BGP Cost Community

You can configure the BGP cost community for routes in a Layer 3 VPN.

Prerequisites

Ensure that you are in the correct VDC (or use the **switchto vdc** command).

SUMMARY STEPS

1. **configure terminal**

2. **feature-set mpls**
3. **feature mpls l3vpn**
4. **feature bgp**
5. **route-map *map-name* {permit | deny} [*sequence-number*]**
6. **set extcommunity cost [igp | pre-bestpath] *community-id cost-value***
7. **router bgp *as-number***
8. (Optional) **router-id *ip-address***
9. **neighbor *ip-address* remote-as *as-number***
10. **address-family {vpn4 | vpn6} unicast**
11. **send-community extended**
12. **route-map *map-name* {in | out}**
13. (Optional) **show running-config bgp**
14. (Optional) **copy running-config startup-config**

DETAILED STEPS

	Command	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	feature-set mpls Example: switch(config)# feature-set mpls	Enables the MPLS feature-set.
Step 3	feature mpls l3vpn Example: switch(config)# feature mpls l3vpn	Enables the Layer 3 VPN feature.
Step 4	feature bgp Example: switch(config)# feature bgp	Enables the BGP feature.
Step 5	route-map <i>map-name</i> {permit deny} [<i>sequence-number</i>] Example: switch(config)# route-map costMap permit 10	Enters route map configuration mode to create or configure a route map.

	Command	Purpose
Step 6	set extcommunity cost [<i>igp</i> <i>pre-bestpath</i>] <i>community-id cost-value</i> Example: switch(config-route-map)# set extcommunity cost 1 100	Creates a set clause to apply the cost community attribute. Multiple cost community set clauses can be configured in each route map block or sequence. Each cost community set clause must have a <i>community-id</i> (0 to 255). The cost community set clause with the lowest <i>cost-value</i> is preferred by the best path selection process when all other attributes are equal. The <i>cost-value</i> range is from 0 to 4294967295. Paths that are not configured with the cost community attribute are assigned the default <i>cost-value</i> of 2147483647.
Step 7	router bgp <i>as-number</i> Example: switch(config)# router bgp 1.1 switch(config-router)#	Configures a BGP routing process and enters router configuration mode. The <i>as-number</i> argument indicates the number of an autonomous system that identifies the router to other BGP routers and tags the routing information passed along. The AS number can be a 16-bit integer or a 32-bit integer in the form of a higher 16-bit decimal number and a lower 16-bit decimal number in xx.xx format.
Step 8	router-id <i>ip-address</i> Example: switch(config-router)# router-id 192.0.2.255	(Optional) Configures the BGP router ID. This IP address identifies this BGP speaker. This command triggers an automatic notification and session reset for the BGP neighbor sessions.
Step 9	neighbor <i>ip-address</i> remote-as <i>as-number</i> Example: switch(config-router)# neighbor 192.0.2.1 remote-as 100 switch(config-router-neighbor)#	Configures a neighbor and enters address family neighbor configuration mode. The <i>ip-address</i> is an IPv4 or IPv6 address. The <i>as-number</i> can be a 16-bit integer or a 32-bit integer in the form of a higher 16-bit decimal number and a lower 16-bit decimal number in xx.xx format.
Step 10	address-family { <i>vpn4</i> <i>vpn6</i> } unicast Example: switch(config-router-neighbor)# address-family vpn4 unicast switch(config-router-neighbor-af)#	Enters address family configuration mode for configuring IP VPN sessions.
Step 11	send-community extended Example: switch(config-router-neighbor-af)# send-community extended	Specifies that a communities attribute should be sent to a BGP neighbor.
Step 12	route-map <i>map-name</i> { <i>in</i> <i>out</i> } Example: switch(config-router-neighbor-af)# route-map costMap in	Configures a route map for this address family.

	Command	Purpose
Step 13	show running-config bgp Example: switch(config-router-neighbor-af)# show running-config bgp	(Optional) Displays the running configuration for BGP.
Step 14	copy running-config startup-config Example: switch(config-router-neighbor-af)# copy running-config startup-config	(Optional) Copies the running configuration to the startup configuration.

Verifying the MPLS Layer 3 VPN Load-Balancing Configuration

To display an MPLS Layer 3 VPN load balancing configuration, perform the following task:

Command	Purpose
show running-config bgp	Displays the running configuration for BGP.
show bgp vpnv4 unicast	Display the VPNv4 routes from the BGP table.

For detailed information about the fields in the output from these commands, see the *Cisco NX-OS MPLS Command Reference*.

Configuration Examples for MPLS Layer 3 VPN Load Balancing

This section includes the following topics:

- [Example: MPLS Layer 3 VPN Load Balancing, page 22-410](#)
- [Example: BGP VPNv4 Multipath, page 22-410](#)
- [Example: MPLS Layer 3 VPN Cost Community, page 22-411](#)

Example: MPLS Layer 3 VPN Load Balancing

The following example shows how to configure iBGP load balancing:

```
configure terminal
feature-set mpls
feature mpls l3vpn
feature bgp
router bgp 1.1
  bestpath cost-community ignore
  address-family ipv6 unicast
    maximum-paths ibgp 4
```

Example: BGP VPNv4 Multipath

The following example shows how to configure a maximum of 3 BGP VPNv4 multipaths:

```
configure terminal
router bgp 100
address-family vpnv4 unicast
```

```
maximum-paths eibgp 3
```

Example: MPLS Layer 3 VPN Cost Community

The following example shows how to configure the BGP cost community:

```
configure terminal
feature-set mpls
feature mpls l3vpn
feature bgp
route-map CostMap permit
set extcommunity cost 1 100
router bgp 1.1
router-id 192.0.2.255
neighbor 192.0.2.1 remote-as 1.1
address-family vpnv4 unicast
send-community extended
route-map CostMap in
```

Additional References for MPLS Layer 3 VPN Load Balancing

For additional information related to implementing an Layer 3 VPN load balancing, see the following sections:

- [Related Documents, page 22-412](#)
- [MIBs, page 22-412](#)

Related Documents

Related Topic	Document Title
CLI commands	<i>Cisco Nexus 7000 Series NX-OS MPLS Command Reference</i>

MIBs

MIBs	MIBs Link
MPLS-Layer 3 VPN-STD-MIB	To locate and download Cisco MIBs, go to the following URL: http://www.cisco.com/public/sw-center/netmgmt/cmtk/mibs.shtml

Feature History for MPLS Layer 3 VPN Load Balancing

Table 22-2 lists the release history for this feature.

Table 22-2 Feature History for MPLS Layer 3 VPN Load Balancing

Feature Name	Releases	Feature Information
BGP VPNv4 Multipath	8.0(1)	This feature was introduced.
Layer 2 and Layer 3 load balancing co-existence	6.2(2)	This feature was introduced.
Layer 3 VPN load balancing	5.2(1)	This feature was introduced.
BGP cost community	5.2(1)	This feature was introduced.

