



Configuring Segment Routing

This chapter contains information on how to configure segment routing.

- [About Segment Routing, on page 1](#)
- [Guidelines and Limitations for Segment Routing, on page 3](#)
- [Overview of BGP Egress Peer Engineering With Segment Routing, on page 5](#)
- [Configuring Segment Routing, on page 7](#)
- [Configuring Layer 3 EVPN and Layer3 VPN over Segment Routing MPLS, on page 18](#)
- [Configuring Segment Routing with IS-IS Protocol, on page 26](#)
- [Configuring Segment Routing with OSPFv2 , on page 28](#)
- [About Segment Routing for Traffic Engineering, on page 30](#)
- [Configuring SR-TE, on page 31](#)
- [Configuration Example for an SR-TE ODN - Use Case, on page 32](#)
- [Verifying SR-TE for Layer 3 EVPN, on page 35](#)
- [Verifying the Segment Routing Configuration, on page 36](#)
- [Configuration Examples for Segment Routing, on page 36](#)
- [Additional References, on page 41](#)

About Segment Routing

Segment routing is a technique by which the path followed by a packet is encoded in the packet itself, similar to source routing. A node steers a packet through a controlled set of instructions, called segments, by prepending the packet with a segment routing header. Each segment is identified by a segment ID (SID) consisting of a flat unsigned 32-bit integer.

Border Gateway Protocol (BGP) segments, a subclass of segments, identify a BGP forwarding instruction. There are two groups of BGP segments: prefix segments and adjacency segments. Prefix segments steer packets along the shortest path to the destination, using all available equal-cost multi-path (ECMP) paths.

Adjacency segments steer packets onto a specific link to a neighbor.

The segment routing architecture is applied directly to the MPLS data plane.

BGP Prefix SID

In order to support segment routing, BGP requires the ability to advertise a segment identifier (SID) for a BGP prefix. A BGP prefix SID is always global within the segment routing BGP domain and identifies an

instruction to forward the packet over the ECMP-aware best path computed by BGP to the related prefix. The BGP prefix SID identifies the BGP prefix segment.

Segment Routing Global Block

The segment routing global block (SRGB) is the range of local labels reserved for MPLS segment routing. The default label range is from 16000 to 23999.

SRGB is the local property of a segment routing node. Each node can be configured with a different SRGB value, and hence the absolute SID value associated to a BGP prefix segment can change from node to node.

The SRGB must be a proper subset of the dynamic label range and must not overlap the optional MPLS static label range. If dynamic labels in the configured or defaulted SRGB range already have been allocated, the configuration is accepted, and the existing dynamic labels that fall in the SRGB range will remain allocated to the original client. If the BGP router attempts to allocate one of these labels, the SRGB mapping fails, and the BGP router reverts to dynamic label allocation. A change to the SRGB range results in the clients deallocating their labels independent of whether the new range can be allocated.

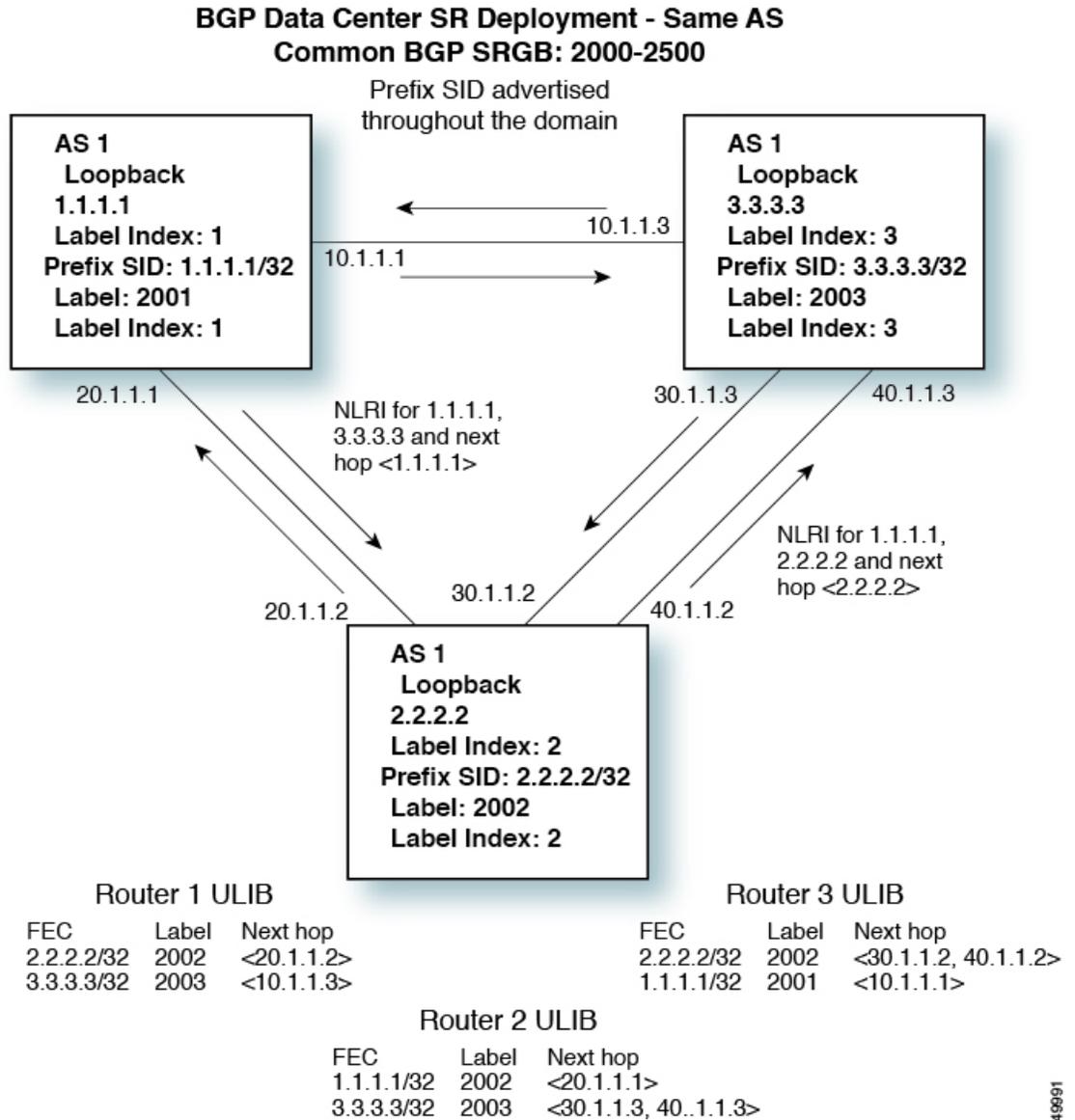
High Availability for Segment Routing

In-service software upgrades (ISSUs) are minimally supported with BGP graceful restart. All states (including the segment routing state) must be relearned from the BGP router's peers. During the graceful restart period, the previously learned route and label state are retained.

BGP Prefix SID Deployment Example

In the simple example below, all three routers are running iBGP and advertising Network Layer Reachability Information (NRLI) to one another. The routers are also advertising their loopback interface as the next hop, which provides the ECMP between routers 2.2.2.2 and 3.3.3.3.

Figure 1: BGP Prefix SID Simple Example



349691

Guidelines and Limitations for Segment Routing

Segment routing has the following guidelines and limitations:

- For notes on platform support see: [Platform Support for Label Switching Features](#).
- MPLS segment routing is not supported for FEX modules.
- When issuing the **feature mpls segment-routing** command to enable MPLS segment routing on a Cisco Nexus 9504 or 9508 switch with a -R series line card, you might find that BFD sessions may go down and come back up. BGP peerings, if configured with BFD, will also go down and come back up. When

a BGP session goes down, it will withdraw routes from the hardware. This results in packet loss until the BGP session is re-established and routes are re-installed. However, once the BFD comes up, no additional flaps should occur.

- Segment Routing Application (SR-APP) module is used to configure the segment routing functionality. Segment Routing Application (SR-APP) is a separate internal process that handles all the CLIs related to segment routing. It is responsible for reserving the SRGB range and for notifying the clients about it. It is also responsible for maintaining the prefix to SID mappings. For more information, see [Configuring Segment Routing Using Segment Routing Application Module, on page 7](#).
- BGP allocates a SRGB label for iBGP route-reflector clients only when next-hop-self is in effect (for example, the prefix is advertised with the next hop being one of the local IP/IPv6 addresses on RR). When you have configured next-hop-self on a RR, the next hop is changed for the routes that are being affected (subject to route-map filtering).
- Static MPLS, MPLS segment routing, and MPLS stripping cannot be enabled at the same time.
- Because static MPLS, MPLS segment routing, and MPLS stripping are mutually exclusive, the only segment routing underlay for multi-hop BGP is single-hop BGP. iBGP multi-hop topologies with eBGP running as an overlay are not supported.
- MPLS pop followed by a forward to a specific interface is not supported. The penultimate hop pop (PHP) is avoided by installing the Explicit NULL label as the out-label in the label FIB (LFIB) even when the control plane installs an IPv4 Implicit NULL label.
- BGP labeled unicast and BGP segment routing are not supported for IPv6 prefixes.
- BGP labeled unicast and BGP segment routing are not supported over tunnel interfaces (including GRE and VXLAN) or with vPC access interfaces.
- MTU path discovery (RFC 2923) is not supported over MPLS label switched paths (LSPs) or segment routed paths.
- For the Cisco Nexus 9500 Series switches, MPLS LSPs and segment routed paths are not supported on subinterfaces (either port channels or normal Layer 3 ports).
- For the Cisco Nexus 9500 platform switches, segment routing is supported only in the non-hierarchical routing mode.
- The BGP configuration commands **neighbor-down fib-accelerate** and **suppress-fib-pending** are not supported for MPLS prefixes.
- The uniform model as defined in RFC 2973 and RFC 3270 is not supported. Consequently, the IP DSCP bits are not copied into the imposed MPLS header.
- Reconfiguration of the segment routing global block (SRGB) results in an automatic restart of the BGP process to update the existing URIB and ULIB entries. Traffic loss will occur for a few seconds, so you should not reconfigure the SRGB in production.
- If the segment routing global block (SRGB) is set to a range but the route-map label-index delta value is outside of the configured range, the allocated label is dynamically generated. For example, if the SRGB is set to a range of 16000-23999 but a route-map label-index is set to 9000, the label is dynamically allocated.
- For network scalability, Cisco recommends using a hierarchical routing design with multi-hop BGP for advertising the attached prefixes from a top-of-rack (TOR) or border leaf switch.

- BGP sessions are not supported over MPLS LSPs or segment routed paths.
- The Layer 3 forwarding consistency checker is not supported for MPLS routes.
- Beginning with Cisco NX-OS Release 9.2(1), the following is applicable:
 1. You can configure segment routing traffic engineering with on-demand nexthop on Cisco Nexus 9000 Series switches
 2. You can configure OSPFv2 as an IGP control plane for Segment Routing on Cisco Nexus 9000 Series switches.
 3. Layer3 VPN and Layer3 EVPN Stitching for Segment Routing is supported on Cisco Nexus 9000 Series switches
 4. Layer3 VPN and Layer3 EVPN Stitching for Segment Routing is not supported on Cisco Nexus 9364C, Cisco Nexus 9200, Cisco Nexus9300-EX, and Cisco Nexus 9500 with 9700-EX line cards.
 5. The OSPF segment routing command and segment-routing traffic engineering with on-demand nexthop is not supported on Cisco Nexus 9364C (N9K-C9364C) switches.

Overview of BGP Egress Peer Engineering With Segment Routing

Cisco Nexus 9000 Series switches are often deployed in massive scale data centers (MSDCs). In such environments, there is a requirement to support BGP Egress Peer Engineering (EPE) with Segment Routing (SR).

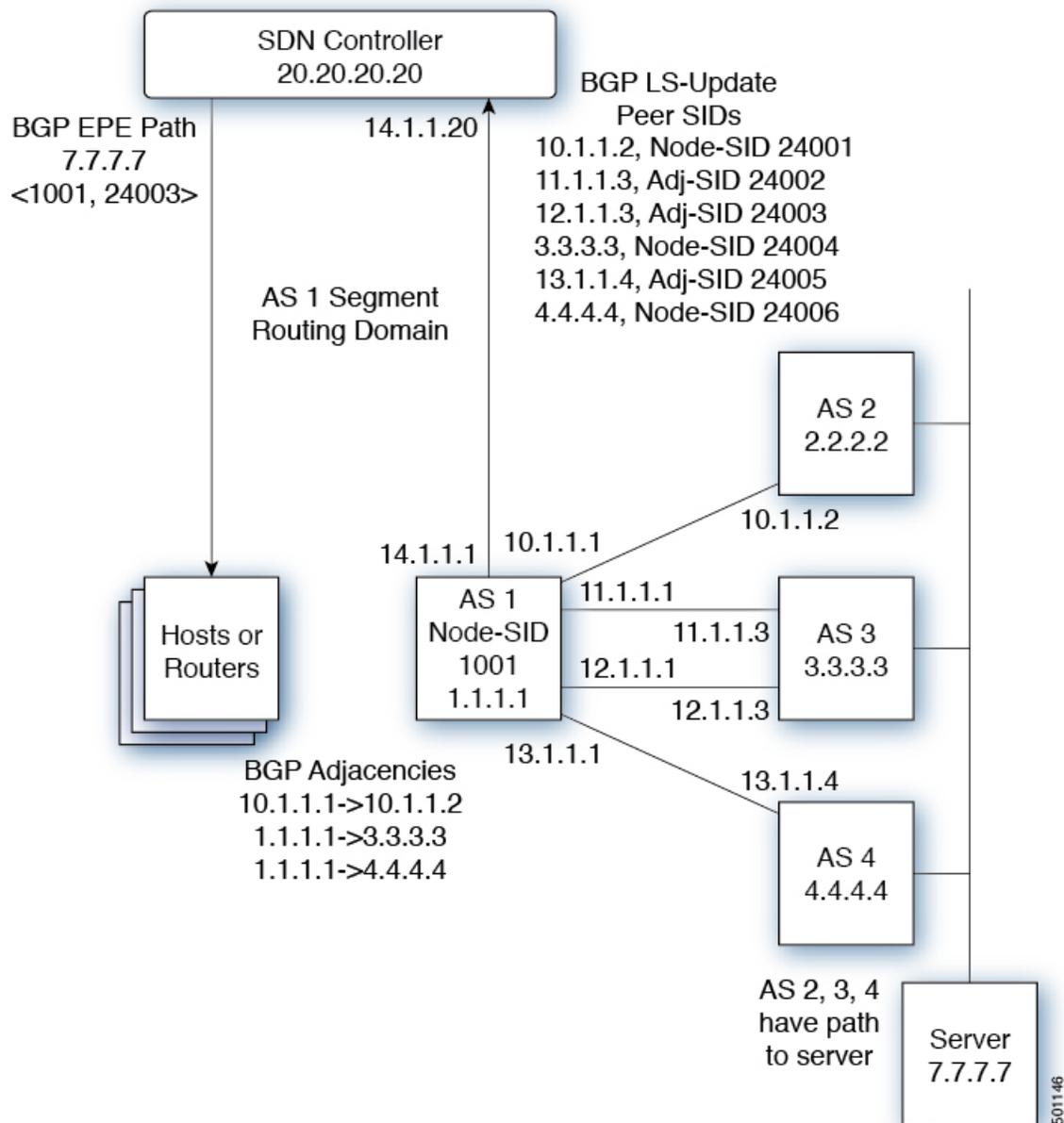
Segment Routing (SR) leverages source routing. A node steers a packet through a controlled set of instructions, known as segments, by prepending the packet with an SR header. A segment can represent any topological or service-based instruction. SR allows steering a flow through any topological path or any service chain while maintaining per-flow state only at the ingress node of the SR domain. For this feature, the Segment Routing architecture is applied directly to the MPLS data plane.

In order to support Segment Routing, BGP requires the ability to advertise a Segment Identifier (SID) for a BGP prefix. A BGP prefix is always global within the SR or BGP domain and it identifies an instruction to forward the packet over the ECMP-aware best-path that is computed by BGP to the related prefix. The BGP prefix is the identifier of the BGP prefix segment.

The SR-based Egress Peer Engineering (EPE) solution allows a centralized (SDN) controller to program any egress peer policy at ingress border routers or at hosts within the domain.

In the following example, all three routers run iBGP and they advertise NRL1 to one another. The routers also advertise their loopback as the next-hop and it is recursively resolved. This provides an ECMP between the routers as displayed in the illustration.

Figure 2: Example of Egress Peer Engineering



The SDN controller receives the Segment IDs from the egress router 1.1.1.1 for each of its peers and adjacencies. It can then intelligently advertise the exit points to the other routers and the hosts within the controller's routing domain. As displayed in the illustration, the BGP Network Layer Reachability Information (NLRI) contains both the Node-SID to Router 1.1.1.1 and the Peer-Adjacency-SID 24003 indicating that the traffic to 7.7.7.7 should egress over the link 12.1.1.1->12.1.1.3.

Guidelines and Limitations for BGP Egress Peer Engineering

BGP Egress Peer Engineering has the following guidelines and limitations:

- BGP Egress Peer Engineering is only supported for IPv4 BGP peers. IPv6 BGP peers are not supported.
- BGP Egress Peer Engineering is only supported in the default VPN Routing and Forwarding (VRF) instance.
- Any number of Egress Peer Engineering (EPE) peers may be added to an EPE peer set. However, the installed resilient per-CE FEC is limited to 32 peers.
- A given BGP neighbor can only be a member of a single peer-set. Peer-sets are configured. Multiple peer-sets are not supported. An optional **peer-set** name may be specified to add neighbor to a peer-set. The corresponding RPC FEC load-balances the traffic across all the peers in the peer-set. The peer-set name is a string that is a maximum length of 63 characters (64 NULL terminated). This length is consistent with the NX-OS policy name lengths. A peer can only be a member of a single peer-set.
- Adjacencies for a given peer are not separately assignable to different peer-sets.

Configuring Segment Routing

Configuring Segment Routing Using Segment Routing Application Module

Segment Routing Application (SR-APP) module is used to configure the segment routing functionality. Segment Routing Application (SR-APP) is a separate internal process that handles all the CLIs related to segment routing. It is responsible for reserving the SRGB range and for notifying the clients about it. It is also responsible for maintaining the prefix to SID mappings. The SR-APP support is also available for the BGP and IS-IS protocols.

Complete the following steps to configure segment routing:

Before you begin

Confirm that the following conditions are met before configuring Segment Routing using the Segment Routing Application (SR-APP) module.

- The **feature-set mpls** and **feature mpls segment-routing** commands should be present for configuring the **segment-routing mpls** command.
- The **feature mpls segment-routing** command starts the SR-APP process.
- If the global block is configured, the specified range is used. Otherwise, the default 16000 – 23999 range is used.
- With the introduction of SR-APP, all configuration is done under **segment-routing mpls** and the prefix SID configuration is handled by SR-APP.
- BGP now uses both **set label-index <value>** configuration and the new **connected-prefix-sid-map** CLI. In case of a conflict, the configuration in SR-APP is preferred.

Procedure

	Command or Action	Purpose
Step 1	configure terminal	Enters global configuration mode.
Step 2	segment-routing mpls	Activates the Segment Routing functionality
Step 3	global-block <min> <max> Example: global-block 201000 280000	Reserves the non-default SRGB range.
Step 4	connected-prefix-sid-map	Provides the SID label for the interface IP covered by the prefix-SID map.
Step 5	address-family ipv4	Enters global address family configuration mode for the IPv4 address family.
Step 6	<prefix>/<masklen> [index absolute] <label> Example: 2.1.1.5/32 absolute 201101 2.10.1.5/32 index 10001	The optional keywords index or absolute indicate whether the label value entered should be interpreted as an index into the SRGB or as an absolute value.

Example

See the following configuration examples of the show commands:

The SRGB allocation needs to be confirmed by an internal process that requires the clients to confirm their cleanup. The amount of time SR-APP waits for the clients to clean their labels, is determined by the cleanup interval. The default value for the cleanup interval is 60 seconds. It can be modified using the **timers srgb cleanup** <interval> CLI command.

Retry interval is amount of time for which SR-APP retries the allocation of the SRGB from the internal process if it fails. The default value for the retry interval is 180 and it can be modified using the **timers srgb retry** <interval> CLI command. The SR-APP module retries the SRGB allocation 10 times within the configured retry timer value, at equal intervals. See the **show segment-routing** CLI output as displayed in the following example:

```
switch# show segment-routing
Segment-Routing Global info

Service Name: segment-routing

State: Enabled

Process Id: 29123

Configured SRGB: 17000 - 24999

SRGB Allocation status: Alloc-Successful

Current SRGB: 17000 - 24999

Cleanup Interval: 60
```

```
Retry Interval: 180
```

The following CLI displays the clients that are registered with SR-APP. It lists the VRFs, for which the clients have registered interest.

```
switch# show segment-routing clients
      Segment-Routing Client Info

Client: isis-1
  PIB index: 1   UUID: 0x41000118   PID: 29463   MTS SAP: 412
  TIBs registered:
    VRF: default Table: base

Client: bgp-1
  PIB index: 2   UUID: 0x11b   PID: 18546   MTS SAP: 62252
  TIBs registered:
    VRF: default Table: base

Total Clients: 2
```

In the **show segment-routing ipv4 connected-prefix-sid-map** CLI command example, SRGB indicates whether the prefix SID is within the configured SRGB. The **Indx** field indicates that the configured label is an index into the global block. The **Abs** field indicates that the configured label is an absolute value.

If the SRGB field displays N, it means that the configured prefix SID is not within the SRGB range and it is not provided to the SR-APP clients. Only the prefix SIDs that fall into the SRGB range are given to the SR-APP clients.

```
switch# show segment-routing ipv4 connected-prefix-sid-map
      Segment-Routing Prefix-SID Mappings
Prefix-SID mappings for VRF default Table base
Prefix      SID   Type Range SRGB
13.11.2.0/24  713  Indx 1   Y
30.7.7.7/32   730  Indx 1   Y
59.3.24.0/30  759  Indx 1   Y
150.101.1.0/24 801  Indx 1   Y
150.101.1.1/32 802  Indx 1   Y
150.101.2.0/24 803  Indx 1   Y
1.1.1.1/32    16013 Abs 1   Y
```

The following CLI displays the **show running-config segment-routing** output.

```
switch# show running-config segment-routing

!Command: show running-config segment-routing
!Time: Thu Jan 25 10:13:53 2018

version 7.0(3)I7(3)
segment-routing mpls
  global-block 22000 35000
  connected-prefix-sid-map
  address-family ipv4
    42.11.11.0/24 index 251
    42.11.12.0/24 index 252
    42.11.13.0/24 index 253
    42.11.14.0/24 index 254
    42.11.15.0/24 index 255
    42.11.16.0/24 index 256
```

```

42.11.17.0/24 index 257
42.11.18.0/24 index 258
42.11.19.0/24 index 259
42.11.20.0/24 index 260
132.10.54.0/24 absolute 22101
2.2.2.9/32 index 202
2.2.2.10/32 index 203
2.2.2.11/32 index 204

```

Enabling MPLS Segment Routing

You can enable MPLS segment routing as long as mutually-exclusive MPLS features such as static MPLS are not enabled.

Before you begin

You must install and enable the MPLS feature set using the **install feature-set mpls** and **feature-set mpls** commands.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	[no] feature mpls segment-routing Example: switch(config)# feature mpls segment-routing	Enables the MPLS segment routing feature. The no form of this command disables the MPLS segment routing feature.
Step 3	(Optional) show running-config inc 'feature mpls segment-routing' Example: switch(config)# show running-config inc 'feature mpls segment-routing'	Displays the status of the MPLS segment routing feature.
Step 4	(Optional) copy running-config startup-config Example: switch(config)# copy running-config startup-config	Copies the running configuration to the startup configuration.

Enabling MPLS on an Interface

You can enable MPLS on an interface for use with segment routing.

Before you begin

You must install and enable the MPLS feature set using the **install feature-set mpls** and **feature-set mpls** commands.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	interface type slot/port Example: switch(config)# interface ethernet 2/2 switch(config-if)#	Enters the interface configuration mode for the specified interface.
Step 3	[no] mpls ip forwarding Example: switch(config-if)# mpls ip forwarding	Enables MPLS on the specified interface. The no form of this command disables MPLS on the specified interface.
Step 4	(Optional) copy running-config startup-config Example: switch(config-if)# copy running-config startup-config	Copies the running configuration to the startup configuration.

Configuring the Segment Routing Global Block

You can configure the beginning and ending MPLS labels in the segment routing global block (SRGB).

Before you begin

You must install and enable the MPLS feature set using the **install feature-set mpls** and **feature-set mpls** commands.

You must enable the MPLS segment routing feature. See [Enabling MPLS Segment Routing, on page 10](#).

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	[no] segment-routing mpls Example:	Enters the segment routing configuration mode and enables the default SRGB of 16000 to

	Command or Action	Purpose
	<pre>switch(config)# segment-routing mpls switch(config-segment-routing-mpls)#</pre>	<p>23999. The no form of this command unallocates that block of labels.</p> <p>If the configured dynamic range cannot hold the default SRGB, an error message appears, and the default SRGB will not be allocated. If desired, you can configure a different SRGB in the next step.</p>
Step 3	<p>[no] global-block <i>beginning-label ending-label</i></p> <p>Example:</p> <pre>switch(config-segment-routing-mpls)# global-block 16000 471804</pre>	<p>Specifies the MPLS label range for the SRGB. Use this command if you want to change the default SRGB label range that is configured with the segment-routing mpls command.</p> <p>The permissive values for the beginning MPLS label and the ending MPLS label are from 16000 to 471804. The mpls label range command permits 16 as the minimum label, but the SRGB can start only from 16000.</p> <p>Note The minimum value for the global-block command starts from 16000. If you upgrading from previous releases, you should modify the SRGB so that it falls within the supported range before triggering an upgrade.</p>
Step 4	<p>(Optional) show mpls label range</p> <p>Example:</p> <pre>switch(config-segment-routing-mpls)# show mpls label range</pre>	<p>Displays the SRGB, only if the SRGB allocation is successful.</p>
Step 5	<p>show segment-routing</p>	<p>Displays the configured SRGB.</p>
Step 6	<p>(Optional) copy running-config startup-config</p> <p>Example:</p> <pre>switch(config-segment-routing-mpls)# copy running-config startup-config</pre>	<p>Copies the running configuration to the startup configuration.</p>

Configuring the Label Index

You can set the label index for routes that match the **network** command. Doing so causes the BGP prefix SID to be advertised for local prefixes that are configured with a route map that includes the **set label-index** command, provided the route map is specified in the **network** command that specifies the local prefix. (For more information on the **network** command, see the "Configuring Basic BGP" chapter in the [Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide](#).)



Note Segment Routing Application (SR-APP) module is used to configure the segment routing functionality. BGP now uses both **set label-index <value>** configuration under route-map and the new **connected-prefix-sid-map** CLI for prefix SID configuration. In case of a conflict, the configuration in SR-APP is preferred.



Note Route-map label indexes are ignored when the route map is specified in a context other than the **network** command. Also, labels are allocated for prefixes with a route-map label index independent of whether the prefix has been configured by the **allocate-label route-map route-map-name** command.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	route-map map-name Example: switch(config)# route-map SRmap switch(config-route-map)#	Creates a route map or enters route-map configuration mode for an existing route map.
Step 3	[no] set label-index index Example: switch(config-route-map)# set label-index 10	Sets the label index for routes that match the network command. The range is from 0 to 471788. By default, a label index is not added to the route.
Step 4	exit Example: switch(config-route-map)# exit switch(config)#	Exits route-map configuration mode.
Step 5	router bgp autonomous-system-number Example: switch(config)# router bgp 64496 switch(config-router)#	Enables BGP and assigns the AS number to the local BGP speaker. 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	Required: address-family ipv4 unicast Example: switch(config-router)# address-family ipv4 unicast switch(config-router-af)#	Enters global address family configuration mode for the IPv4 address family.

	Command or Action	Purpose
Step 7	network <i>ip-prefix</i> [route-map <i>map-name</i>] Example: switch(config-router-af)# network 10.10.10.10/32 route-map SRmap	Specifies a network as local to this autonomous system and adds it to the BGP routing table.
Step 8	(Optional) show route-map [<i>map-name</i>] Example: switch(config-router-af)# show route-map	Displays information about route maps, including the label index.
Step 9	(Optional) copy running-config startup-config Example: switch(config-router-af)# copy running-config startup-config	Copies the running configuration to the startup configuration.

Configuring Neighbor Egress Peer Engineering Using BGP

With the introduction of RFC 7752 and draft-ietf-idr-bgppls-segment-routing-epe, you can configure Egress Engineering. The feature is valid only for external BGP neighbors and it is not configured by default. Egress Engineering uses RFC 7752 encoding.

Before you begin

- You must enable BGP.
- After an upgrade from Release 7.0(3)I3(1) or Release 7.0(3)I4(1), configure the TCAM region before configuring Egress Peer Engineering (EPE) on Cisco Nexus 9000 Series switches using the following commands:
 1. switch# **hardware access-list tcam region vpc-convergence 0**
 2. switch# **hardware access-list tcam region racl 0**
 3. switch# **hardware access-list tcam region mpls 256 double-wide**
- Save the configuration and reload the switch.

For more information, see the Using Templates to Configure ACL TCAM Region Sizes and Configuring ACL TCAM Region Sizes sections in the *Cisco Nexus 9000 Series NX-OS Security Configuration Guide*.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	router bgp < <i>bgp autonomous number</i> >	Specifies the autonomous router BGP number.

	Command or Action	Purpose
Step 3	neighbor <IP address>	Configures the IP address for the neighbor.
Step 4	<p>[no]default egress-engineering [peer-set peer-set-name]</p> <p>Example:</p> <pre>switch(config)# router bgp 1 switch(config-router)# neighbor 4.4.4.4 switch(config-router)# egress-engineering peer-set NewPeer</pre>	<p>Specifies whether a Peer-Node-SID is allocated for the neighbor and it is advertised in an instance of a BGP Link-State (BGP-LS) address family Link NLRI. If the neighbor is a multi-hop neighbor, a BGP-LS Link NLRI instance is also advertised for each Equal-Cost-MultiPath (ECMP) path to the neighbor and it includes a unique Peer-Adj-SID.</p> <p>Optionally, you can add the neighbor to a peer-set. The Peer-Set-SID is also advertised in the BGP-LS Link NLRI in the same instance as the Peer-Node-SID. BGP Link-State NLRI is advertised to all neighbors with the link-state address family configured.</p> <p>See RFC 7752 and draft-ietf-idr-bgppls-segment-routing-epe-05 for more information on EPE.</p>

Configuration Example for Egress Peer Engineering

See the Egress Peer Engineering sample configuration for the BGP speaker 1.1.1.1. Note that the neighbor 20.20.20.20 is the SDN controller.

```
hostname epe-as-1
install feature-set mpls
feature-set mpls

feature telnet
feature bash-shell
feature scp-server
feature bgp
feature mpls segment-routing

segment-routing mpls
vlan 1

vrf context management
ip route 0.0.0.0/0 10.30.97.1
ip route 0.0.0.0/0 10.30.108.1

interface Ethernet1/1
no switchport
ip address 10.1.1.1/24
no shutdown

interface Ethernet1/2
no switchport
ip address 11.1.1.1/24
no shutdown

interface Ethernet1/3
no switchport
```

```

ip address 12.1.1.1/24
no shutdown

interface Ethernet1/4
no switchport
ip address 13.1.1.1/24
no shutdown

interface Ethernet1/5
no switchport
ip address 14.1.1.1/24
no shutdown

interface mgmt0
ip address dhcp
vrf member management

interface loopback1
ip address 1.1.1.1/32
line console

line vty
ip route 2.2.2.2/32 10.1.1.2
ip route 3.3.3.3/32 11.1.1.3
ip route 3.3.3.3/32 12.1.1.3
ip route 4.4.4.4/32 13.1.1.4
ip route 20.20.20.20/32 14.1.1.20

router bgp 1
address-family ipv4 unicast
address-family link-state
neighbor 10.1.1.2
remote-as 2
address-family ipv4
egress-engineering
neighbor 3.3.3.3
remote-as 3
address-family ipv4
update-source loopback1
ebgp-multihop 2
egress-engineering
neighbor 4.4.4.4
remote-as 4
address-family ipv4
update-source loopback1
ebgp-multihop 2
egress-engineering
neighbor 20.20.20.20
remote-as 1
address-family link-state
update-source loopback1
ebgp-multihop 2
neighbor 124.11.50.5
bfd
remote-as 6
update-source port-channel50.11
egress-engineering peer-set pset2 <<<<<<<
address-family ipv4 unicast
neighbor 124.11.101.2
bfd
remote-as 6
update-source Vlan2401
egress-engineering

```

```
address-family ipv4 unicast
```

This example shows sample output for the **show bgp internal epe** command.

```
switch# show bgp internal epe
BGP Egress Peer Engineering (EPE) Information:
Link-State Server: Inactive
Link-State Client: Active
Configured EPE Peers: 26
Active EPE Peers: 3
EPE SID State:
RPC SID Peer or Set Assigned
ID Type Set Name ID Label Adj-Info, iod
1 Node 124.1.50.5 1 1600
2 Set pset1 2 1601
3 Node 6.6.6.6 3 1602
4 Node 124.11.50.5 4 1603
5 Set pset2 5 1604
6 Adj 6.6.6.6 6 1605 124.11.50.4->124.11.50.5/0x1600b031, 80
7 Adj 6.6.6.6 7 1606 124.1.50.4->124.1.50.5/0x16000031, 78
EPE Peer-Sets:
IPv4 Peer-Set: pset1, RPC-Set 2, Count 7, SID 1601
Peers: 124.11.116.2 124.11.111.2 124.11.106.2 124.11.101.2
124.11.49.5 124.1.50.5 124.1.49.5
IPv4 Peer-Set: pset2, RPC-Set 5, Count 5, SID 1604
Peers: 124.11.117.2 124.11.112.2 124.11.107.2 124.11.102.2
124.11.50.5
IPv4 Peer-Set: pset3, RPC-Set 0, Count 4, SID unspecified
Peers: 124.11.118.2 124.11.113.2 124.11.108.2 124.11.103.2
IPv4 Peer-Set: pset4, RPC-Set 0, Count 4, SID unspecified
Peers: 124.11.119.2 124.11.114.2 124.11.109.2 124.11.104.2
IPv4 Peer-Set: pset5, RPC-Set 0, Count 4, SID unspecified
Peers: 124.11.120.2 124.11.115.2 124.11.110.2 124.11.105.2
switch#
```

Configuring the BGP Link State Address Family

You can configure the BGP link state address family for a neighbor session with a controller to advertise the corresponding SIDs. You can configure this feature in global configuration mode and neighbor address family configuration mode.

Before you begin

You must enable BGP.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	router bgp <bgp autonomous number>	Specifies the autonomous router BGP number.

	Command or Action	Purpose
Step 3	<p>[no] address-family link-state</p> <p>Example:</p> <pre>switch(config)# router bgp 64497 switch (config-router af)# address-family link-state</pre>	<p>Enters address-family interface configuration mode.</p> <p>Note This command can also be configured in neighbor address-family configuration mode.</p>
Step 4	<p>neighbor <IP address></p>	<p>Configures the IP address for the neighbor.</p>
Step 5	<p>[no] address-family link-state</p> <p>Example:</p> <pre>switch(config)#router bgp 1 switch(config-router)#address-family link-state switch(config-router)#neighbor 20.20.20.20 switch(config-router)#address-family link-state</pre>	<p>Enters address-family interface configuration mode.</p> <p>Note This command can also be configured in neighbor address-family configuration mode.</p>

Configuring Layer 3 EVPN and Layer3 VPN over Segment Routing MPLS

This section describes tasks to configure the Layer 3 EVPN and stitching of L3 EVPN and L3VPN router. Perform the following tasks to complete the configuration:

- [Configuring the Features to Enable L3EVPN and L3VPN, on page 18](#)
- [Configuring VRF and Route Targets for Import and Export Rules, on page 19](#)
- [Configuring BGP EVPN and Label Allocation Mode, on page 20](#)
- [Configuring BGP L3 EVPN and L3 VPN Stitching, on page 22](#)
- [Configuring BGP L3 VPN over Segment Routing, on page 25](#)

Before you begin

Install the VPN Fabric license.

Make sure that the **feature interface-vlan** command is enabled.

Configuring the Features to Enable L3EVPN and L3VPN

Before you begin

Install the VPN Fabric license.

Make sure that the **feature interface-vlan** command is enabled.

Procedure

	Command or Action	Purpose
Step 1	feature bgp	Enables BGP feature and configurations.
Step 2	install feature-set mpls	Enables MPLS configuration commands.
Step 3	feature-set mpls	Enables MPLS configuration commands.
Step 4	feature mpls segment-routing	Enables segment routing configuration commands.
Step 5	feature mpls evpn	Enables EVPN over MPLS configuration commands. This command is mutually exclusive with the feature-nv CLI command.
Step 6	feature mpls l3vpn	Enables EVPN over MPLS configuration commands. This command is mutually exclusive with the feature-nv CLI command.

Configuring VRF and Route Targets for Import and Export Rules

Procedure

	Command or Action	Purpose
Step 1	configure terminal	Enters global configuration mode.
Step 2	vrf <i>vrf-name</i>	Defines a VPN routing and forwarding (VRF) instance and enters the VRF configuration mode.
Step 3	rd auto	Automatically assigns a unique route distinguisher (RD) to VRF.
Step 4	address-family { ipv4 ipv6 } unicast	Specifies either the IPv4 or IPv6 address family for the VRF instance and enters address family configuration submode.
Step 5	route-target import <i>route-target-id</i>	Configures importing of routes to the VRF from the L3VPN BGP NLRI that have the matching route-target value.
Step 6	route-target export <i>route-target-id</i>	Configures exporting of routes from the VRF to the L3VPN BGP NLRI and assigns the specified route-target identifiers to the L3VPN BGP NLRI.
Step 7	route-target import <i>route-target-id evpn</i>	Configures importing of routes from the L3 EVPN BGP NLRI that have the matching route-target value.

	Command or Action	Purpose
Step 8	route-target export <i>route-target-id</i> evpn	Configures exporting of routes from the VRF to the L3 EVPN BGP NLRIs and assigns the specified route-target identifiers to the BGP EVPN NLRIs.

Configuring BGP EVPN and Label Allocation Mode

You can use MPLS tunnel encapsulation using the **encapsulation mpls** command. You can configure the label allocation mode for the EVPN address family. The default tunnel encapsulation in EVPN for IP Route type in NX-OS is VXLAN.

Advertisement of (IP or Label) bindings from a Cisco Nexus 9000 Series switch via BGP EVPN enables a remote switch to send the routed traffic to that IP using the label for that IP to the switch that advertised the IP over MPLS.

The IP prefix route (Type-5) is:

- Type-5 route with MPLS encapsulation

```
RT-5 Route - IP Prefix

RD: L3 RD
IP Length: prefix length
IP address: IP (4 bytes)
Label1: BGP MPLS Label
Route Target
RT for IP-VRF
```

The default label allocation mode is per-VRF for Layer 3 EVPN over MPLS.

Complete the following steps to configure BGP EVPN and label allocation mode:

Before you begin

You must install and enable the MPLS feature set using the **install feature-set mpls** and **feature-set mpls** commands.

You must enable the MPLS segment routing feature.

Procedure

	Command or Action	Purpose
Step 1	configure terminal	Enters global configuration mode.
Step 2	[no] router bgp <i>autonomous-system-number</i> Example: switch(config)# router bgp 64496 switch(config-router)#	Enables BGP and assigns the AS number to the local BGP speaker. 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.

	Command or Action	Purpose
		Use the no option with this command to remove the BGP process and the associated configuration.
Step 3	<p>Required: address-family l2vpn evpn</p> <p>Example:</p> <pre>switch(config-router)# address-family l2vpn evpn switch(config-router-af)#</pre>	Enters global address family configuration mode for the Layer 2 VPN EVPN.
Step 4	<p>Required: exit</p> <p>Example:</p> <pre>switch(config-router-af)# exit switch(config-router)#</pre>	Exits global address family configuration mode.
Step 5	<p>neighbor ipv4-address remote-as autonomous-system-number</p> <p>Example:</p> <pre>switch(config-router)# neighbor 10.1.1.1 remote-as 64497 switch(config-router-neighbor)#</pre>	Configures the IPv4 address and AS number for a remote BGP peer.
Step 6	<p>address-family l2vpn evpn</p> <p>Example:</p> <pre>switch(config-router-neighbor)# address-family l2vpn evpn switch(config-router-neighbor-af)#</pre>	Advertises the labeled Layer 2 VPN EVPN.
Step 7	<p>encapsulation mpls</p> <p>Example:</p> <pre>router bgp 100 address-family l2vpn evpn neighbor NVE2 remote-as 100 address-family l2vpn evpn send-community extended encapsulation mpls vrf foo address-family ipv4 unicast advertise l2vpn evpn</pre> <p>BGP segment routing configuration:</p> <pre>router bgp 100 address-family ipv4 unicast network 200.0.0.1/32 route-map label_index_pol_100 network 192.168.5.1/32 route-map label_index_pol_101 network 101.0.0.0/24 route-map label_index_pol_103 allocate-label all</pre>	<p>Enables BGP EVPN address family and sends EVPN type-5 route update to the neighbors.</p> <p>Note The default tunnel encapsulation in EVPN for the IP route type in NX-OS is VXLAN. To override that, a new CLI is introduced to indicate MPLS tunnel encapsulation.</p>

	Command or Action	Purpose
	<pre>neighbor 192.168.5.6 remote-as 20 address-family ipv4 labeled-unicast send-community extended</pre>	
Step 8	vrf <customer_name>	Configures the VRF.
Step 9	address-family ipv4 unicast	Enters global address family configuration mode for the IPv4 address family.
Step 10	advertise l2vpn evpn	Advertises Layer 2 VPN EVPN.
Step 11	redistribute direct route-map DIRECT_TO_BGP	Redistributes the directly connected routes into BGP-EVPN.
Step 12	label-allocation-mode per-vrf	<p>Sets the label allocation mode to per-VRF. If you want to configure the per-prefix label mode, use the no label-allocation-mode per-vrf CLI command.</p> <p>For the EVPN address family, the default label allocation is per-vrf, compared to per-prefix mode for the other address-families where the label allocation CLI is supported. No form of CLI is displayed in the running configuration.</p>

Example

See the following example for configuring per-prefix label allocation:

```
router bgp 65000
[address-family l2vpn evpn]
neighbor 10.1.1.1
remote-as 100
address-family l2vpn evpn
send-community extended
neighbor 20.1.1.1
remote-as 65000
address-family l2vpn evpn
encapsulation mpls
send-community extended
vrf customer1
address-family ipv4 unicast
advertise l2vpn evpn
redistribute direct route-map DIRECT_TO_BGP
no label-allocation-mode per-vrf
```

Configuring BGP L3 EVPN and L3 VPN Stitching

In order to configure the stitching on the same router, configure the L3VPN neighbor relationship and router advertisement.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	[no] router bgp <i>autonomous-system-number</i> Example: switch# configure terminal switch(config)# router bgp 64496 switch(config-router)#	Enables BGP and assigns the AS number to the local BGP speaker. 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. Use the no option with this command to remove the BGP process and the associated configuration.
Step 3	address-family {<i>vpn4</i> <i>vpn6</i>} unicast Example: switch(config-router)# address-family vpn4 unicast switch(config-router-af)# address-family vpn6 unicast switch(config-router-af)#	Enters global address family configuration mode for the Layer 3 VPNv4 or VPNv6.
Step 4	exit Example: switch(config-router-af)# exit switch(config-router)#	Exits global address family configuration mode.
Step 5	neighbor <i>ipv4-address</i> remote-as <i>autonomous-system-number</i> Example: switch(config-router)# neighbor 20.1.1.1 remote-as 64498	Configures the IPv4 address and AS number for a remote BGP L3VPN peer.
Step 6	address-family {<i>vpn4</i> <i>vpn6</i>} unicast Example: switch(config-router)# address-family vpn4 unicast switch(config-router-af)# address-family vpn6 unicast switch(config-router-af)#	Configure the neighbor address-family for VPNv4 or VPNv6.
Step 7	send-community extended	Enables BGP VPN address family
Step 8	import l2vpn evpn reoriginate	Configures import of routing information from the L3VPN BGP NLRI that has route target identifier matching the normal route target identifier and exports this routing information after re-origination that assigns it with stitching

	Command or Action	Purpose
		route target identifier, to the BGP EVPN neighbor.
Step 9	neighbor ipv4-address remote-as autonomous-system-number Example: <pre>switch(config-router)# neighbor 10.1.1.1 remote-as 64497 switch(config-router-neighbor)#</pre>	Configures the IPv4 address and AS number for a remote L3EVPN BGP peer.
Step 10	address-family {l2vpn evpn} Example: <pre>switch(config-router-neighbor)# address-family l2vpn evpn switch(config-router-neighbor-af)#</pre>	Configure the neighbor address-family for L3EVPN.
Step 11	import vpn unicast reoriginate	Enables import of routing information from BGP EVPN NLRIs that has route target identifier matching the stitching route target identifier and exports this routing information after re-origination to the L3VPN BGP neighbor.
Step 12	vrf <customer_name>	Configures the VRF.
Step 13	address-family ipv4 unicast	Enters global address family configuration mode for the IPv4 address family.
Step 14	advertise l2vpn evpn	Advertises Layer 2 VPN EVPN.

Example

```
vrf context Customer1
  rd auto
  address-family ipv4 unicast
    route-target import 100:100
    route-target export 100:100
    route-target import 100:100 evpn
    route-target export 100:100 evpn

segment-routing mpls
  global-block 11000 20000
  connected-prefix-sid
    address-family ipv4 unicast
      200.0.0.1 index 101
!
int lol
  ip address 200.0.0.1/32
!
interface e1/13
  description "MPLS interface towards Core"
  ip address 192.168.5.1/24
  mpls ip forwarding
  no shut
```

```

router bgp 100
address-family ipv4 unicast
allocate-label all
address-family ipv6 unicast
address-family l2vpn evpn
address-family vpvv4 unicast
address-family vpvv6 unicast
neighbor 10.0.0.1 remote-as 200
    update-source loopback1
    address-family vpvv4 unicast
        send-community extended
    import l2vpn evpn reoriginate
address-family vpvv6 unicast
    import l2vpn evpn reoriginate
    send-community extended
neighbor 20.0.0.1 remote-as 300
    address-family l2vpn evpn
    send-community extended
    import vpn unicast reoriginate
    encapsulation mpls
neighbor 192.168.5.6 remote-as 300
    address-family ipv4 labeled-unicast
vrf Customer1
    address-family ipv4 unicast
    advertise l2vpn evpn
    address-family ipv6 unicast
    advertise l2vpn evpn

```

Configuring BGP L3 VPN over Segment Routing

Before you begin

You must install and enable the MPLS feature set using the **install feature-set mpls** and **feature-set mpls** commands.

You must enable the MPLS segment routing feature.

You must enable the MPLS L3 VPN feature using the **feature mpls l3vpn** command.

Procedure

	Command or Action	Purpose
Step 1	configure terminal Example: switch# configure terminal switch(config)#	Enters global configuration mode.
Step 2	[no] router bgp <i>autonomous-system-number</i> Example: switch(config)# router bgp 64496 switch(config-router)#	Enables BGP and assigns the AS number to the local BGP speaker. 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.

	Command or Action	Purpose
		Use the no option with this command to remove the BGP process and the associated configuration.
Step 3	address-family {vpn4 vpn6} unicast Example: switch(config-router)# address-family vpn4 unicast switch(config-router-af)# address-family vpn6 unicast switch(config-router-af)#	Enters global address family configuration mode for the Layer 3 VPNv4 or VPNv6.
Step 4	[no] allocate-label option-b	Disables the inter-AS option-b
Step 5	Required: exit Example: switch(config-router-af)# exit switch(config-router)#	Exits global address family configuration mode.
Step 6	neighbor ipv4-address remote-as autonomous-system-number Example: switch(config-router)# neighbor 20.1.1.1 remote-as 64498 switch(config-router-neighbor)#	Configures the IPv4 address and AS number for a remote BGP L3VPN peer.
Step 7	address-family {vpn4 vpn6 } unicast Example: switch(config-router-neighbor)# address-family vpn4 unicast switch(config-router-neighbor-af)#	Configure the neighbor address-family for VPNv4 or VPNv6.
Step 8	send-community extended	Enables BGP VPN address family.
Step 9	vrf <customer_name>	Configures the VRF.
Step 10	allocate-index x	Configure the allocate-index.
Step 11	address-family ipv4 unicast	Enters global address family configuration mode for the IPv4 address family.
Step 12	redistribute direct route-map DIRECT_TO_BGP	Redistributes the directly connected routes into BGP-L3VPN.

Configuring Segment Routing with IS-IS Protocol

You can configure segment routing with IS-IS protocol.

Before you begin

IS-IS segment routing is fully enabled when the following conditions are met:

- The **mpls segment-routing** feature is enabled.
- The IS-IS feature is enabled.
- Segment routing is enabled for at least one address family under IS-IS.

Procedure

	Command or Action	Purpose
Step 1	configure terminal	Enters global configuration mode.
Step 2	router isis <i>instance-tag</i>	Creates a new IS-IS instance with the configured instance tag.
Step 3	net <i>network-entity-title</i>	Configures the NET for this IS-IS instance.
Step 4	(Optional) is-type {<i>level-1</i> <i>level-2</i> <i>level-1-2</i>}	Configures the area level for this IS-IS instance. The default is level-1-2.
Step 5	log-adjacency-changes	Sends a system message whenever an IS-IS neighbor changes the state.
Step 6	address-family <i>ipv4</i> unicast	Enters address family configuration mode.
Step 7	segment-routing mpls	Configures segment routing with IS-IS protocol. Note <ul style="list-style-type: none"> • The IS-IS command is supported only on the IPv4 address family. It is not supported on the IPv6 address family. • Redistribution is not supported from any other protocol to ISIS for the SR prefixes. You need to enable ip router isis command on all the prefix SID interfaces.
Step 8	(Optional) show running-config segment-routing	Displays the status of the segment routing.

See the following configuration example for configuring segment routing with IS-IS protocol.

Example

```
switch# config t
router isis SR-ISIS-1
  bfd
```

```

net 31.0000.0000.0000.000e.00
is-type level-1-2
log-adjacency-changes
address-family ipv4 unicast
    segment-routing mpls          >>> # New command added for ISIS.
    address-family ipv6 unicast
    bfd

switch# show running-config segment-routing

!Command: show running-config segment-routing
!Time: Fri Dec 22 12:51:59 2017

version 7.0(3)I7(3)
segment-routing mpls
    global-block 201000 280000
    connected-prefix-sid-map
    address-family ipv4
        2.1.1.5/32 absolute 201101
        2.10.1.5/32 index 10001

switch# show running-config isis

!Command: show running-config isis
!Time: Thu Jan 25 10:18:19 2018

version 7.0(3)I7(3)
feature isis

router isis 10
    bfd
    net 56.0000.0000.0003.00
    is-type level-1-2
    maximum-paths 64
    log-adjacency-changes
    address-family ipv4 unicast
        segment-routing mpls

interface Vlan12
    ip router isis 10

interface Vlan13
    ip router isis 10

```

Configuring Segment Routing with OSPFv2

Beginning with Cisco NX-OS Release 9.2(1), you can configure segment routing with OSPFv2 protocol.

Before you begin

OSPFv2 segment routing is fully enabled when the following conditions are met:

- The mpls segment-routing feature is enabled.
- The OSPFv2 feature is enabled.
- Segment routing is enabled under OSPF.



Note Beginning with Cisco NX-OS Release 9.2(1), SR OSPF will advertise prefix SID for addresses associated with the loopback interfaces only.

Procedure

	Command or Action	Purpose
Step 1	<code>configure terminal</code>	Enters global configuration mode.
Step 2	<code>[no]router ospf</code>	Enables the OSPF mode.
Step 3	<code>segment-routing mpls</code>	Configures the Segment Routing functionality

See the following configuration example for configuring segment routing with OSPFv2.

Example

```
switch# show running-config ospf
!Command: show running-config ospf
!Running configuration last done at: Sun Jul 15 15:09:07 2018
!Time: Sun Jul 15 15:09:09 2018

version 9.2(1) Bios:version 07.60
feature ospf

router ospf SR_OSPF
  segment-routing mpls
  router-id 2.2.2.1

interface loopback1
  ip router ospf SR_OSPF area 0.0.0.0

switch# show running-config interface loopback 1
!Command: show running-config interface loopback1
!Running configuration last done at: Sun Jul 15 15:11:16 2018
!Time: Sun Jul 15 15:13:05 2018

version 9.2(1) Bios:version 07.60

interface loopback1
  ip address 2.2.2.1/32
  ip router ospf SR_OSPF area 0.0.0.0

switch# show running-config segment-routing
!Command: show running-config segment-routing
!Running configuration last done at: Sun Jul 15 15:11:16 2018
!Time: Sun Jul 15 15:11:54 2018

version 9.2(1) Bios:version 07.60
segment-routing mpls
  global-block 201000 400000
  connected-prefix-sid-map
  address-family ipv4
    2.2.2.1/32 absolute 201101
```

About Segment Routing for Traffic Engineering

Segment routing for traffic engineering (SR-TE) takes place through a tunnel between a source and destination pair. Segment routing for traffic engineering uses the concept of source routing, where the source calculates the path and encodes it in the packet header as a segment. A Traffic Engineered (TE) tunnel is a container of TE LSPs instantiated between the tunnel ingress and the tunnel destination. A TE tunnel can instantiate one or more SR-TE LSPs that are associated with the same tunnel.

With segment routing for traffic engineering (SR-TE), the network no longer needs to maintain a per-application and per-flow state. Instead, it simply obeys the forwarding instructions provided in the packet.

SR-TE utilizes network bandwidth more effectively than traditional MPLS-TE networks by using ECMP at every segment level. It uses a single intelligent source and relieves remaining routers from the task of calculating the required path through the network.

About SR-TE Policies

Segment routing for traffic engineering (SR-TE) uses a “policy” to steer traffic through the network. An SR-TE policy is a container that includes sets of segments or labels. This list of segments can be provisioned by an operator, a stateful PCE, or the SR-TE infra can dynamically calculate the path by applying Constrained Shortest Path First (CSPF) algorithm on its local IGP database. The headend imposes the corresponding MPLS label stack on traffic flow to be carried over the SR-TE policy. Each transit node along the SR-TE policy path uses the incoming top label to select the next-hop, pop, or swap the label, and forward the packet to the next node with the remainder of the label stack, until the packet reaches the ultimate destination.

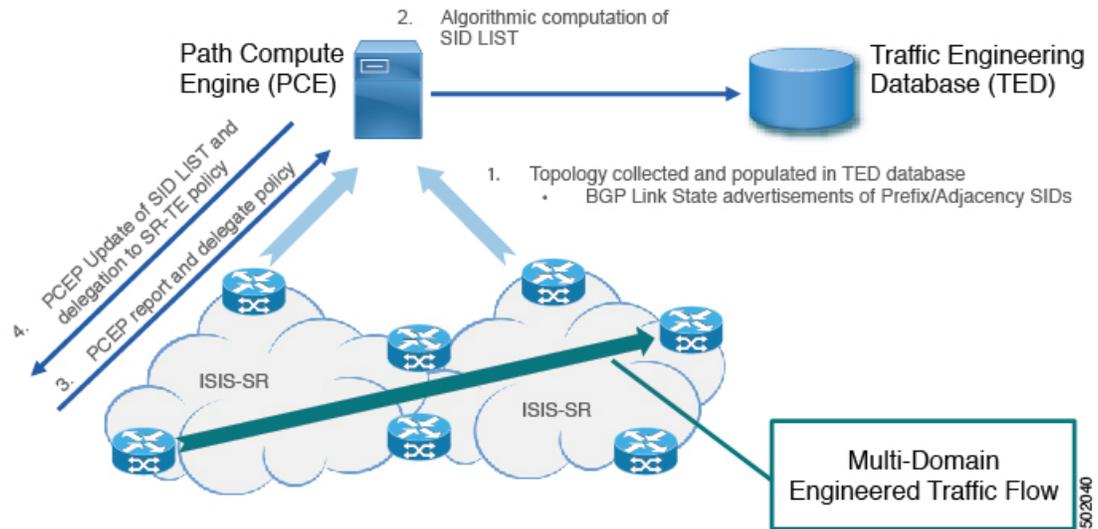
An SR-TE policy is uniquely identified by a tuple (color, endpoint). Color is represented as a 32-bit number and an endpoint is either an IPv4 and IPv6 address. Every SR-TE policy has a color value. Every policy between the same node pairs requires a unique color value. Multiple SR-TE policies can be created between the same two endpoints by choosing different colors for the policies.

Cisco NX-OS Release 9.2(1) supports the local dynamic SR-TE policy. When you configure local dynamic SR-TE, the headend locally calculates the path to the destination address. Dynamic path calculation results in a list of interface IP addresses that traffic engineering (TE) maps to adj-SID labels. Routes are learned by way of forwarding adjacencies over the TE tunnel.

Segment Routing On Demand Next Hop

On-Demand Next hop (ODN) leverages upon BGP Dynamic SR-TE capabilities and adds the path computation (PCE) ability to find and download the end to end path based on the requirements. ODN triggers an SR-TE auto-tunnel based on the defined BGP policy. As shown in the following figure, an end-to-end path between ToR1 and AC1 can be established from both ends based on IGP Metric. The work-flow for ODN is summarized as follows:

Figure 3: ODN Operation



Guidelines and Limitations for SR-TE On-Demand Next Hop

SR-TE ODN has the following guidelines and limitations:

- For notes on platform support see: [Platform Support for Label Switching Features](#).
- ODN for IPv6 is not supported.
- SR-TE ODN is supported only with ISIS Underlay.

Configuring SR-TE

Beginning with Cisco NX-OS Release 9.2(1), you can configure segment routing for traffic engineering.

Before you begin

You must ensure that the mpls segment routing feature is enabled.

Procedure

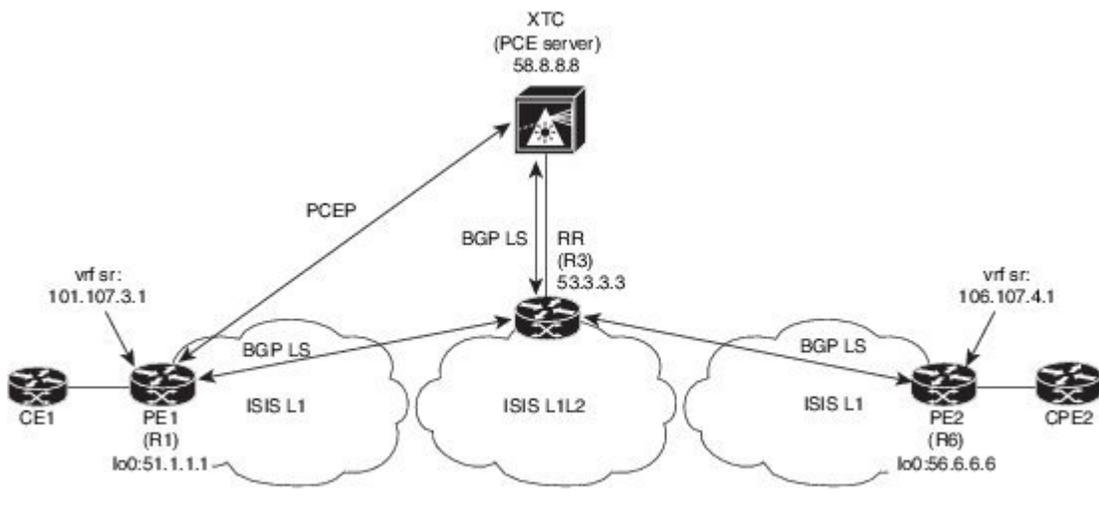
	Command or Action	Purpose
Step 1	<code>configure terminal</code>	Enters global configuration mode.
Step 2	<code>feature mpls segment-routing traffic-engineering</code>	Enables mpls SR-TE.
Step 3	<code>segment-routing</code>	Enters the segment-routing mode
Step 4	<code>traffic-engineering</code>	Enters the traffic engineering mode.

	Command or Action	Purpose
Step 5	encapsulation mpls source ipv4 <i>tunnel_ip_address</i>	Configures the source address for the SR TE Tunnel.
Step 6	pcc	Enters the PCC mode.
Step 7	source-address ipv4 <i>pcc_source_address</i>	Configure source address for the PCC
Step 8	pce-address ipv4 <i>pce_source_address</i> precedence num	Configure IP address of the PCE. The lowest numbered PCE will take precedence, and the other(s) be used as a backup.
Step 9	on-demand color <i>color_num</i>	Enters the on-demand mode to configure the color.
Step 10	metric-type <i>igp</i>	Configures the metric type.

Configuration Example for an SR-TE ODN - Use Case

Perform the following steps to configure ODN for SR-TE. The following figure is used as a reference to explain the configuration steps.

Figure 4: Reference Topology



1. Configure all links with IS-IS point-to-point session from PE1 to PE2. Also, configure the domains as per the above topology.
2. Enable “distribute link-state” for IS-IS session on R1, R3, and R6.

```
router isis 1
 net 31.0000.0000.0000.712a.00
 log-adjacency-changes
 distribute link-state
 address-family ipv4 unicast
```

```

bfd
segment-routing mpls
maximum-paths 32
advertise interface loopback0

```

3. Configure the router R1 (headend) and R6 (tailend) with a VRF interface.

VRF configuration on R1:

```

interface Ethernet1/49.101
encapsulation dot1q 201
vrf member sr
ip address 101.10.1.1/24
no shutdown

vrf context sr
rd auto
address-family ipv4 unicast
route-target import 101:101
route-target import 101:101 evpn
route-target export 101:101
route-target export 101:101 evpn
router bgp 6500
vrf sr
bestpath as-path multipath-relax
address-family ipv4 unicast
advertise l2vpn evpn

```

4. Tags VRF prefix with BGP community on R6 (tailend).

```

route-map color1001 permit 10
set extcommunity color 1001

```

5. Enable BGP on R6 (tailend) and R1 (headend) to advertise and receive VRF SR prefix and match on community set on R6 (tailend).

R6 < EVPN > R3 < EVPN > R1

BGP Configuration R6:

```

router bgp 6500
address-family ipv4 unicast
allocate-label all
neighbor 53.3.3.3
remote-as 6500
log-neighbor-changes
update-source loopback0
address-family l2vpn evpn
send-community extended
route-map Color1001 out
encapsulation mpls

```

BGP Configuration R1:

```

router bgp 6500
address-family ipv4 unicast
allocate-label all
neighbor 53.3.3.3
remote-as 6500
log-neighbor-changes
update-source loopback0
address-family l2vpn evpn
send-community extended
encapsulation mpls

```

6. Enable BGP configuration on R3 and BGP LS with XTC on R1, R3.abd

BGP Configuration R3:

```

router bgp 6500
  router-id 2.20.1.2
  address-family ipv4 unicast
  allocate-label all
  address-family l2vpn evpn
  retain route-target all
  neighbor 56.6.6.6
    remote-as 6500
    log-neighbor-changes
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
      route-map NH_UNCHANGED out
      encapsulation mpls
  neighbor 51.1.1.1
    remote-as 6500
    log-neighbor-changes
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
      route-map NH_UNCHANGED out
      encapsulation mpls
  neighbor 58.8.8.8
    remote-as 6500
    log-neighbor-changes
    update-source loopback0
    address-family link-state

route-map NH_UNCHANGED permit 10
  set ip next-hop unchanged

```

BGP Configuration R1:

```

router bgp 6500
neighbor 58.8.8.8
  remote-as 6500
  log-neighbor-changes
  update-source loopback0
  address-family link-state

```

BGP Configuration R6:

```

outer bgp 6500
  neighbor 58.8.8.8
  remote-as 6500
  log-neighbor-changes
  update-source loopback0
  address-family link-state

```

7. Enable PCE and SR-TE tunnel configurations on R1.

```

segment-routing
  traffic-engineering
  pcc
    source-address ipv4 51.1.1.1
    pce-address ipv4 58.8.8.8
    on-demand color 1001
    metric-type igp

```

Verifying SR-TE for Layer 3 EVPN

The ODN verifications are based on L3VPN VRF prefixes.

1. Verify that the PCEP session between R1 (headend and PCE server) is established.

```
R1# show srte pce ipv4 peer

PCC's peer database:
-----
Remote PCEP conn IPv4 addr: 58.8.8.8
Local PCEP conn IPv4 addr: 51.1.1.1
Precedence: 0
State: up
```

2. Verify BGP LS and BGP EVPN session on R1, R3, and R6 using the following commands:

- Show bgp l2vpn evpn summary
- Show bgp link-state summary

3. Verify that the R1 (headend) has no visibility to the R6 loopback address.

```
R1# show ip route 56.6.6.6
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
 '**' denotes best mcast next-hop
 '[x/y]' denotes [preference/metric]
 '%<string>' in via output denotes VRF <string>

56.6.6.6/32, ubest/mbest: 1/0
   *via Null0, [1/0], 1d02h, static
```

4. Verify that the VRF prefix is injected via MP-BGP in a R1 VRF SR routing table.

```
R1# show ip route vrf sr
106.107.4.1/32, ubest/mbest: 1/0
   *via binding label 100534%default, [20/0], 1d01h, bgp-6503, external, tag 6500
(mpls-vpn)
```

5. Verify the SR-TE Tunnel.

```
R1# show srte policy
Policy name: 51.1.1.1|1001
Source: 51.1.1.1
End-point: 56.6.6.6
Created by: bgp
State: UP
Color: 1001
Insert: FALSE
Re-opt timer: 0
Binding-sid Label: 100534
Policy-Id: 2
Flags:
Path type = MPLS          Path options count: 1
Path-option Preference:100 ECMP path count: 1
  1.      PCE              Weighted: No
      Delegated PCE: 58.8.8.8
          Index: 1          Label: 101104
          Index: 2          Label: 201102
          Index: 3          Label: 201103
```

Verifying the Segment Routing Configuration

To display the segment routing configuration, perform one of the following tasks:

Command	Purpose
<code>show bgp ipv4 labeled-unicast prefix</code>	Displays the advertised label index and the selected local label for the specified IPv4 prefix.
<code>show bgp paths</code>	Displays the BGP path information, including the advertised label index.
<code>show mpls label range</code>	Displays the configured SRGB range of labels.
<code>show route-map [map-name]</code>	Displays information about a route map, including the label index.
<code>show running-config inc 'feature mpls segment-routing'</code>	Displays the status of the MPLS segment routing feature.
<code>show running-config segment-routing</code>	Displays the status of the segment routing feature.

This example shows how the `show bgp ipv4 labeled-unicast` command can be used with a prefix specification to display the advertised label index and the selected local label:

```
switch# show bgp ipv4 labeled-unicast 19.19.19.19/32
BGP routing table information for VRF default, address family IPv4 Label Unicast
BGP routing table entry for 19.19.19.19/32, version 2
Paths: (1 available, best #1)
Flags: (0x20c0012) on xmit-list, is in urib, is backup urib route, has label
label af: version 2, (0x100002) on xmit-list
local label: 16010

Advertised path-id 1, Label AF advertised path-id 1
Path type: external, path is valid, is best path
AS-Path: 19 , path sourced external to AS
60.1.1.19 (metric 0) from 60.1.1.19 (100.100.100.100)
Origin IGP, MED not set, localpref 100, weight 0
Received label 3
Prefix-SID Attribute: Length: 10
Label Index TLV: Length 7, Flags 0x0 Label Index 10

Path-id 1 not advertised to any peer

Label AF advertisement
Path-id 1 not advertised to any peer
```

Configuration Examples for Segment Routing

The examples in this section show a common BGP prefix SID configuration between two routers.

This example shows how to advertise a BGP speaker configuration of 10.10.10.10/32 and 20.20.20.20/32 with a label index of 10 and 20, respectively. It uses the default segment routing global block (SRGB) range of 16000 to 23999.

```
hostname s1
install feature-set mpls
feature-set mpls

feature telnet
feature bash-shell
feature scp-server
feature bgp
feature mpls segment-routing

segment-routing
  mpls
  vlan 1
segment-routing
  mpls
  connected-prefix-sid-map
  address-family ipv4
  2.1.1.1/32 absolute 100100

route-map label-index-10 permit 10
  set label-index 10
route-map label-index-20 permit 10
  set label-index 20

vrf context management
  ip route 0.0.0.0/0 10.30.108.1

interface Ethernet1/1
  no switchport
  ip address 10.1.1.1/24
  no shutdown

interface mgmt0
  ip address dhcp
  vrf member management

interface loopback1
  ip address 10.10.10.10/32

interface loopback2
  ip address 20.20.20.20/32

line console
line vty

router bgp 1
  address-family ipv4 unicast
  network 10.10.10.10/32 route-map label-index-10
  network 20.20.20.20/32 route-map label-index-20
  allocate-label all
  neighbor 10.1.1.2 remote-as 2
  address-family ipv4 labeled-unicast
```

This example shows how to receive the configuration from a BGP speaker.

```
hostname s2
install feature-set mpls
feature-set mpls

feature telnet
feature bash-shell
feature scp-server
feature bgp
feature mpls segment-routing
```

```

segment-routing mpls
vlan 1

vrf context management
 ip route 0.0.0.0/0 10.30.97.1
 ip route 0.0.0.0/0 10.30.108.1

interface Ethernet1/1
 no switchport
 ip address 10.1.1.2/24
 ipv6 address 10:1:1::2/64
 no shutdown

interface mgmt0
 ip address dhcp
 vrf member management

interface loopback1
 ip address 2.2.2.2/32
 line console

line vty

router bgp 2
 address-family ipv4 unicast
  allocate-label all
 neighbor 10.1.1.1 remote-as 1
 address-family ipv4 labeled-unicast

```

This example shows how to display the configuration from a BGP speaker. The **show** command in this example displays the prefix 10.10.10.10 with label index 10 mapping to label 16010 in the SRGB range of 16000 to 23999.

```
switch# show bgp ipv4 labeled-unicast 10.10.10.10/32
```

```

BGP routing table information for VRF default, address family IPv4 Label Unicast
BGP routing table entry for 10.10.10.10/32, version 7
Paths: (1 available, best #1)
Flags: (0x20c001a) on xmit-list, is in urib, is best urib route, is in HW, , has label
 label af: version 8, (0x100002) on xmit-list
 local label: 16010

Advertised path-id 1, Label AF advertised path-id 1
Path type: external, path is valid, is best path, no labeled nexthop, in rib
AS-Path: 1 , path sourced external to AS
 10.1.1.1 (metric 0) from 10.1.1.1 (10.10.10.10)
  Origin IGP, MED not set, localpref 100, weight 0
  Received label 0
  Prefix-SID Attribute: Length: 10
    Label Index TLV: Length 7, Flags 0x0 Label Index 10

Path-id 1 not advertised to any peer
Label AF advertisement
Path-id 1 not advertised to any peer

```

This example shows how to configure egress peer engineering on a BGP speaker.

```

hostname epe-as-1
install feature-set mpls
feature-set mpls

feature telnet

```

```

feature bash-shell
feature scp-server
feature bgp
feature mpls segment-routing

segment-routing mpls
vlan 1

vrf context management
 ip route 0.0.0.0/0 10.30.97.1
 ip route 0.0.0.0/0 10.30.108.1

interface Ethernet1/1
 no switchport
 ip address 10.1.1.1/24
 no shutdown

interface Ethernet1/2
 no switchport
 ip address 11.1.1.1/24
 no shutdown

interface Ethernet1/3
 no switchport
 ip address 12.1.1.1/24
 no shutdown

interface Ethernet1/4
 no switchport
 ip address 13.1.1.1/24
 no shutdown

interface Ethernet1/5
 no switchport
 ip address 14.1.1.1/24
 no shutdown

```

The following is an example of show ip route vrf 2 command.

```

show ip route vrf 2
IP Route Table for VRF "2"
'*' denotes best ucast next-hop
 '**' denotes best mcast next-hop
 '[x/y]' denotes [preference/metric]
 '%<string>' in via output denotes VRF <string>

41.11.2.0/24, ubest/mbest: 1/0
   *via 1.1.1.9%default, [20/0], 13:26:48, bgp-2, external, tag 11 (mpls-vpn)
42.11.2.0/24, ubest/mbest: 1/0, attached
   *via 42.11.2.1, Vlan2, [0/0], 13:40:52, direct
42.11.2.1/32, ubest/mbest: 1/0, attached
   *via 42.11.2.1, Vlan2, [0/0], 13:40:52, local

```

The following is an example of show forwarding route vrf 2 command.

```

slot 1
=====

IPv4 routes for table 2/base

```

Prefix	Next-hop Partial Install	Interface	Labels
0.0.0.0/32	Drop	Null0	
127.0.0.0/8	Drop	Null0	
255.255.255.255/32	Receive	sup-eth1	
*41.11.2.0/24	27.1.31.4	Ethernet1/3	PUSH
30002 492529	27.1.32.4	Ethernet1/21	PUSH
30002 492529	27.1.33.4	port-channel23	PUSH
30002 492529	27.11.31.4	Ethernet1/3.11	PUSH
30002 492529	27.11.33.4	port-channel23.11	PUSH
30002 492529	37.1.53.4	Ethernet1/53/1	PUSH
29002 492529	37.1.54.4	Ethernet1/54/1	PUSH
29002 492529	37.2.53.4	Ethernet1/53/2	PUSH
29002 492529	37.2.54.4	Ethernet1/54/2	PUSH
29002 492529	80.211.11.1	Vlan801	PUSH
30002 492529			

The following is an example of **show bgp l2vpn evpn summary** command.

```
show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 2.2.2.3, local AS number 2
BGP table version is 17370542, L2VPN EVPN config peers 4, capable peers 1
1428 network entries and 1428 paths using 268464 bytes of memory
BGP attribute entries [476/76160], BGP AS path entries [1/6]
BGP community entries [0/0], BGP clusterlist entries [0/0]
476 received paths for inbound soft reconfiguration
476 identical, 0 modified, 0 filtered received paths using 0 bytes

Neighbor      V   AS MsgRcvd MsgSent  TblVer  InQ  OutQ  Up/Down  State/PfxRcd
1.1.1.1       4   11      0       0        0    0    23:01:53  Shut (Admin)
1.1.1.9       4   11    4637    1836  17370542  0    0    23:01:40  476
1.1.1.10      4   11      0       0        0    0    23:01:53  Shut (Admin)
1.1.1.11      4   11      0       0        0    0    23:01:52  Shut (Admin)
```

The following is an example of **show bgp l2vpn evpn** command.

```
show bgp l2vpn evpn 41.11.2.0
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 14.1.4.1:115
BGP routing table entry for [5]:[0]:[0]:[24]:[41.11.2.0]:[0.0.0.0]/224, version 17369591
Paths: (1 available, best #1)
Flags: (0x000002) on xmit-list, is not in l2rib/evpn, is not in HW

  Advertised path-id 1
```

```

Path type: external, path is valid, received and used, is best path
          Imported to 2 destination(s)
AS-Path: 11 , path sourced external to AS
  1.1.1.9 (metric 0) from 1.1.1.9 (14.1.4.1)
    Origin incomplete, MED 0, localpref 100, weight 0
    Received label 492529
    Extcommunity: RT:2:20

```

Path-id 1 not advertised to any peer

```

Route Distinguisher: 2.2.2.3:113
BGP routing table entry for [5]:[0]:[0]:[24]:[41.11.2.0]:[0.0.0.0]/224, version 17369595
Paths: (1 available, best #1)
Flags: (0x000002) on xmit-list, is not in l2rib/evpn, is not in HW

```

```

Advertised path-id 1
Path type: external, path is valid, is best path
          Imported from 14.1.4.1:115:[5]:[0]:[0]:[24]:[41.11.2.0]:[0.0.0.0]/224
AS-Path: 11 , path sourced external to AS
  1.1.1.9 (metric 0) from 1.1.1.9 (14.1.4.1)

```

Additional References

Related Documents

Related Topic	Document Title
BGP	<i>Cisco Nexus 9000 Series Unicast Routing Configuration Guide</i>

