BGP PIC (Prefix Independent Convergence) Edge for IP and MPLS-VPN

The BGP PIC (Prefix Independent Convergence) Edge for IP and MPLS-VPN feature improves BGP convergence after a network failure. This convergence is applicable to both core and edge failures and can be used in both IP and MPLS networks. The BGP PIC Edge for IP and MPLS-VPN feature creates and stores a backup or alternate path in the routing information base (RIB), forwarding information base (FIB), and Cisco Express Forwarding. When a failure is detected, the backup or alternate path immediately takes over, thus enabling fast failover.

Note In this document, the BGP PIC Edge for IP and MPLS-VPN feature is called by the short name BGP PIC.

Prerequisites for BGP PIC

- Ensure that the Border Gateway Protocol (BGP) and the IP or Multiprotocol Label Switching (MPLS) network is up and running with the customer site connected to the provider site by more than one path (multihomed).
- Ensure that the backup/alternate path has a unique next hop that is not the same as the next hop of the best path.
- Enable the Bidirectional Forwarding Detection (BFD) protocol to quickly detect link failures of directly connected neighbors.

Restrictions for BGP PIC

- The uLoop during cutover may delay the label switched paths propagation and eventually impact BGP PIC edge Convergence. To avoid uLoop, you can configure MPLS-TE tunnels from local PE to the RRs.
- If multiple features are configured with their respective scales, then it affects the BGP PIC edge Convergence.
- With BGP Multipath, the BGP Prefix-Independent Convergence (PIC) feature is already supported.
- In MPLS VPNs, the BGP PIC feature is not supported with MPLS VPN Inter-Autonomous Systems Option B.
- The BGP PIC feature supports prefixes only for IPv4, IPv6, VPNv4, and VPNv6 address families.
- The BGP PIC feature cannot be configured with Multicast or L2VPN Virtual Routing and Forwarding (VRF) address families.
- If the route reflector is only in the control plane, then you do not need BGP PIC, because BGP PIC addresses data plane convergence.
- When two PE routers become each other's backup/alternate path to a CE router, traffic might loop if the CE router fails. Neither router will reach the CE router, and traffic will continue to be forwarded between the PE routers until the time-to-live (TTL) timer expires.
- The BGP PIC feature solves the traffic forwarding only for a single network failure at both the edge and the core.
- The BGP PIC feature does not work with the BGP Best External feature. If you try to configure the BGP PIC feature after configuring the BGP Best External feature, you receive an error.
- BGP PIC over bridge domain interface (BDI) core interfaces can have only one Ethernet Flow Point (EFP) associated with each of the BDI interfaces.
- The maximum number of bridge domain interfaces (BDI) that can act as protected or protecting interfaces via BGP PIC is 24.
- BGP PIC Core is enabled only if a minimum of 15 BGP prefixes are received from the peer.
- BGP PIC edge works only when BGP PIC core is enabled.
- Each of the 2 BGP peers must send the same set of 15 BGP prefixes to the DUT.
- BGP PIC edge provides sub-second convergence for Global prefixes, EoMPLS over BGP LU, VPLS, and 6PE.
- BGP PIC edge provides sub-second convergence for VPNv4/VPNv6 prefixes with a maximum of 4000 (including global). Sub-second convergence is not supported for BGP PIC Edge for prefixes beyond 4000.
- Starting with Cisco IOS XE Everest 16.6.1 release, **bgp mpls-local-label** command must be enabled for BGP PIC edge with EoMPLS over BGP labeled unicast (RFC3107) configuration.
- Starting with Cisco IOS XE Release 16.9.1, BGP PIC Edge is supported over TE-FRR with Next Hops.
- BGP PIC edge is supported only on dual-rate ports. It is not supported on the AMS ports.

- Targeted LDP (MPLS IP over tunnels) is not supported.
- VPLS over MPLS TE and MPLS TE FRR is not supported.
- Tunnel configuration under two simultaneous IGPs is not supported.
- Tunnel statistics is not supported, if one of the labels (Primary/Backup) is an implicit-null. For the tunnel statistics to work as expected:
  - Always configure tunnel which is at least one hop away and not the immediate next-hop.
  - In case of TE-FRR, the backup tunnel should end at least one hop before the primary tunnel's destination and not on the same hop as that of the primary tunnel's destination. However, for ping packets, statistics will increment.
- Sub second convergence is not guaranteed when more than 100 primary and backup tunnels are terminating on the same node.
- 500 MPLS TE Headend Tunnels are supported.
- Primary tunnels configured for link or node protection cannot go over port channel interfaces.
- P2MP TE Tunnels are not supported.
- Multicast over PtoP Tunnel (MPLS TE) is not supported.
- Inter AS TE tunnels are not supported.
- MPLS TE path protection is not supported.
- MPLS TE auto route destination not supported for inter-area tunnels.
- Use the following commands to avoid high convergence issues. Actual values for *delay installation* and *delay cleanup* can be configured.
  - **mpls traffic-eng reoptimize timers delay installation** *delay installation*
  - **mpls traffic-eng reoptimize timers delay cleanup** *delay cleanup*
- MPLS TE tunnel statisticss are supported only for tunnels with real label.
- MPLS TE explicit-null is not supported.

About BGP PIC

In the following sections, we describe the BGP PIC feature in details, how to detect a failure, a scenario and how to configure it.

Benefits

- An extra path for failover allows faster restoration of connectivity when a primary path is invalid or withdrawn.
- Reduction of traffic loss.

- Constant convergence time so that the switching time is the same for all prefixes.

BGP Convergence

Under normal circumstances, BGP can take several seconds to a few minutes to converge after a change in the network. At a high level, BGP goes through the steps of the following process:

1. BGP learns of failures through either Interior Gateway Protocol (IGP) or BFD events or interface events.
2. BGP withdraws the routes from the routing information base (RIB), and the RIB withdraws the routes from the forwarding information base (FIB) and distributed FIB (dFIB). This process clears the data path for the affected prefixes.
3. BGP sends withdrawn messages to its neighbors.
4. BGP calculates the next best path to the affected prefixes.
5. BGP inserts the next best path for affected prefixes into the RIB, and the RIB installs them in the FIB and dFIB.

This process may take from few seconds to a few minutes to complete. It depends on, the latency of the network, the convergence time across the network, and the local load on the devices. The data plane converges only after the control plane converges.

Improve Convergence

The BGP PIC functionality is achieved by an extra functionality in the BGP, RIB, Cisco Express Forwarding, and MPLS.

- BGP Functionality

BGP PIC affects prefixes under IPv4 and VPNv4 address families. For those prefixes, BGP calculates an extra second best path, along with the primary best path. (The second best path is called the backup or alternate path.) BGP installs the best and backup or alternate paths for the affected prefixes into the BGP RIB. The backup or alternate path provides a fast reroute mechanism to counter a singular network failure. BGP also includes the alternate or backup path in its application programming interface (API) to the IP RIB.

- RIB Functionality

For BGP PIC, RIB installs an alternate path per route if one is available. If the RIB selects a BGP route containing a backup or alternate path, it installs the backup or alternate path with the best path. The RIB also includes the alternate path in its API with the FIB.

- Cisco Express Forwarding Functionality

With BGP PIC, Cisco Express Forwarding stores an alternate path per prefix. When the primary path goes down, Cisco Express Forwarding searches for the backup or alternate path in a prefix-independent manner. Cisco Express Forwarding also listens to BFD events to rapidly detect local failures.

- MPLS Functionality

MPLS Forwarding is similar to Cisco Express Forwarding in that it stores alternate paths and switches to an alternate path if the primary path goes down.

When the BGP PIC feature is enabled, BGP calculates a backup or alternate path per prefix and installs it into BGP RIB, IP RIB, and FIB. This improves convergence after a network failure. There are two types of network failures that the BGP PIC feature detects:

- Core node or link failure (internal Border Gateway Protocol [iBGP] node failure): If a PE node or link fails, then the failure is detected through IGP convergence. IGP conveys the failure through the RIB to the FIB.
- Local link or immediate neighbor node failure (external Border Gateway Protocol [eBGP] node or link failure): To detect a local link failure or eBGP single-hop peer node failure in less than a second, you must enable BFD. Cisco Express Forwarding looks for BFD events to detect a failure of an eBGP single-hop peer.

Convergence in the Data Plane

Upon detecting a failure, Cisco Express Forwarding detects the alternate next hop for all prefixes that are affected by the failure. The data plane convergence is achieved in subseconds depending on whether the BGP PIC implementation exists in the software or hardware.

Convergence in the Control Plane

Upon detecting a failure, BGP learns about the failure through IGP convergence or BFD events and sends withdrawn messages for the prefixes, recalculating the best and backup or alternate paths, and advertising the next best path across the network.

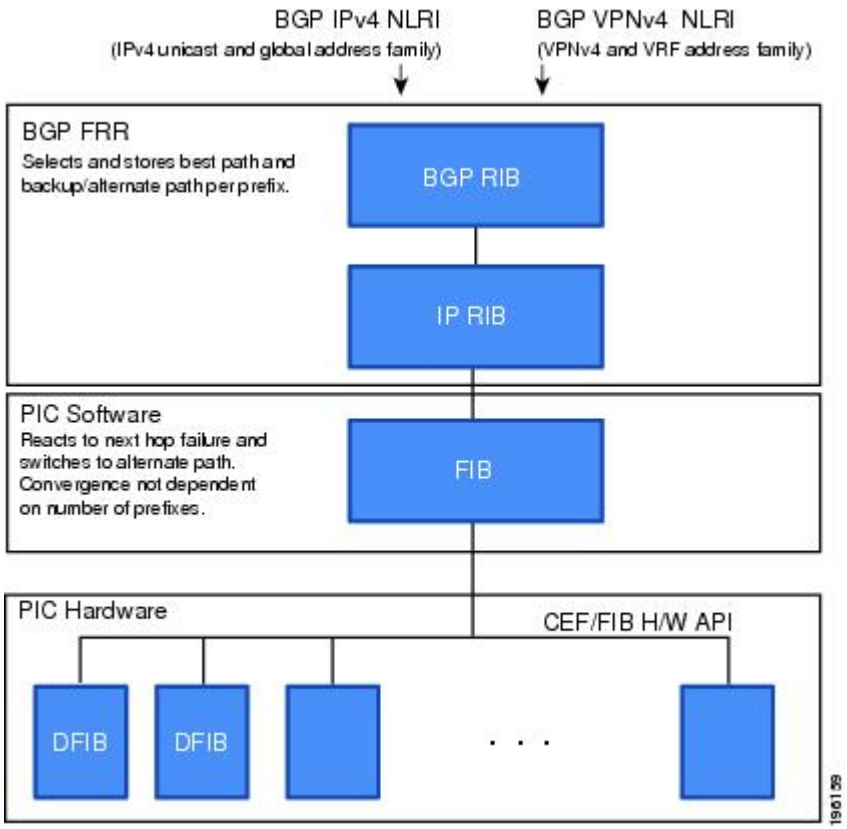
BGP Fast Reroute

BGP Fast Reroute (FRR) provides a best path and a backup or alternate path in BGP, RIB, and Cisco Express Forwarding. BGP FRR provides a fast reroute mechanism into the RIB and Cisco Express Forwarding (CEF) on the backup BGP next hop to reach a destination when the current best path is not available.

BGP FRR precomputes a second best path in BGP and gives it to the RIB and Cisco Express Forwarding as a backup or alternate path, and CEF programs it into line cards.

The BGP PIC feature provides the ability for CEF to quickly switch the traffic to the other egress ports if the current next hop or the link to this next hop goes down.

Figure 1: BGP PIC Edge and BGP FRR



Detect a Failure

IGP detects a failure in the iBGP (remote) peer; it may take a few seconds to detect the failure. Convergence can occur in subseconds or seconds, depending on whether PIC is enabled on the line cards.

If the failure is among the directly connected neighbors (eBGP), and if you use BFD to detect when a neighbor has gone down. Depending on whether PIC is enabled on the line cards, the detection may happen within subseconds and the convergence can occur in subseconds or few seconds.

How BGP PIC Achieves Subsecond Convergence

MPLS VPN–BGP Local Convergence

The BGP PIC is an enhancement to the MPLS VPN–BGP Local Convergence feature. It provides a failover mechanism that recalculates the best path after a link failure. It then installs the new path in forwarding. To minimize traffic loss, the feature maintains the local label for 5 minutes to ensure that the traffic uses the backup or alternate path.

The BGP PIC improves the LoC time to under a second by calculating a backup or alternate path in advance. When a link failure occurs, the traffic is sent to the backup or alternate path.

When you configure BGP PIC, it overrides the functionality of the MPLS VPN--BGP Local Convergence feature. Do not remove the **protection local-prefixes** command from the configuration.

Overview of IPv6 VPN Provider Edge (6PE/VPE)

IPv6 VPN Provider Edge (6PE/VPE) uses the existing MPLS IPv4 core infrastructure for IPv6 transport. 6PE/VPE enables IPv6 sites to communicate with each other over an MPLS IPv4 core network using MPLS label switched paths (LSPs). This feature relies on multiprotocol Border Gateway Protocol (BGP) extensions in the IPv4network configuration on the provider edge (PE) router to exchange IPv6 reachability information, in addition to an MPLS label) for each IPv6 address prefix. Edge routers are configured as dual-stack, running both IPv4 and IPv6, and use the IPv4 mapped IPv6 address for IPv6 prefix reachability exchange.

A hierarchy of labels are imposed on the 6PE ingress routers to keep the IPv6 traffic transparent to all the core routers. The top label provides connectivity inside the IPv4 MPLS core network and the label is distributed by Label Distribution Protocol (LDP). The bottom label, assigned to the IPv6 prefix of the destination, is distributed by multiprotocol BGP and used at each 6PE egress router for IPv6 forwarding.

All 6PE and core routers within the MPLS domain share a common IPv4 Interior Gateway Protocol (IGP) such as Open Shortest Path First (OSPF) or Integrated Intermediate System-to-Intermediate System (IS-IS).

IPv6 VPN uses the coexistence between IPv6 and IPv4 by leveraging an existent MPLS IPv4 core network. This approach is called 6VPE. The routing component of the VPN operation is divided into core routing and edge routing.

- Core routing, which involves PE routers and P routers, typically is performed by an IPv4 Interior Gateway Protocol (IGP) such as Open Shortest Path First (OSPF) or Intermediate System-to-Intermediate System (IS-IS).
- Edge routing takes place in two directions—routing between PE pairs and routing between a PE and a CE.

  Routing between PE pairs is achieved using multiprotocol internal BGP (iBGP) using the IPv6 VPN address family. This method distributes routes learned from CEs through PE-CE routing, using appropriate route export policies at the ingress PE router and appropriate route import policies at the egress PE router. Routing between the CE and its PE is achieved using a routing protocol that is VPN routing and forwarding (VRF) aware.

Static routes and external BGP (eBGP) are VRF instance aware.

BGP PIC Scenario

You can configure the BGP PIC functionality to achieve fast convergence.

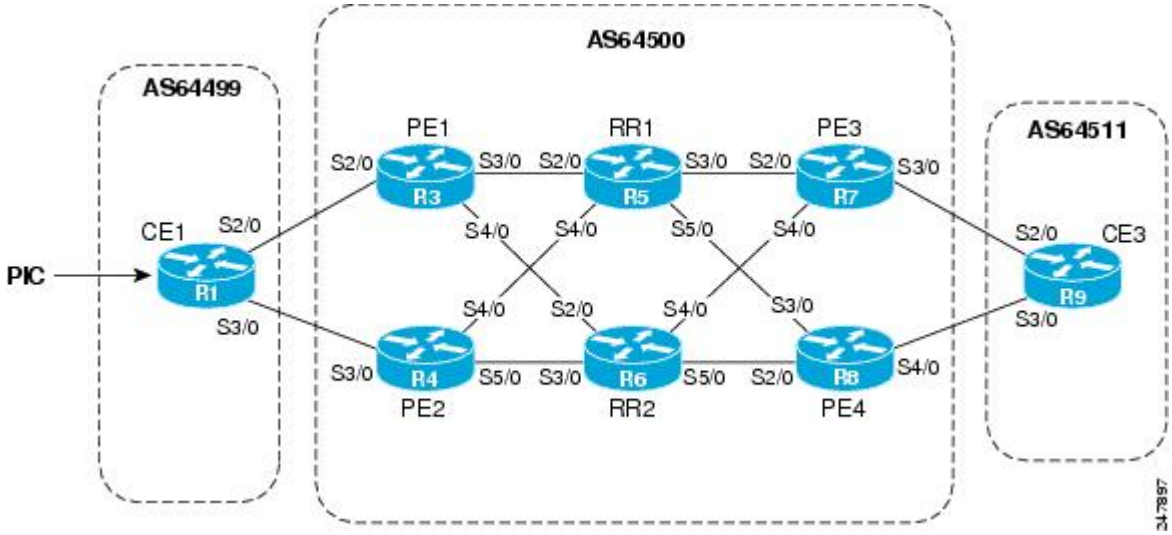
IP PE-CE Link and Node Protection

The network includes the following components:

- eBGP sessions exist between the PE and CE routers.
- Traffic from CE1 uses PE1 to reach network 192.168.9.0/24 through router CE3.
- CE1 has two paths:
  - PE1 as the primary path.
  - PE2 as the backup or alternate path.

CE1 is configured with the BGP PIC feature. BGP computes PE1 as the best path and PE2 as the backup or alternate path. It installs both routes into the RIB and Cisco Express Forwarding plane. When the CE1-PE1 link goes down, Cisco Express Forwarding detects the link failure and points the forwarding object to the backup or alternate path. Traffic is quickly rerouted due to local fast convergence in Cisco Express Forwarding.

Figure 2: Using BGP PIC to Protect the PE-CE Link



IP PE-CE Link and Node Protection on the CE Side (Dual CEs and Dual PE Primary and Backup Nodes)

The figure below shows a network that uses the BGP PIC feature on CE1. The network includes the following components:

- eBGP sessions exist between the PE and CE routers.
- Traffic from CE1 uses PE1 to reach network 192.168.9.0/24 through router CE3.
- CE1 has two paths:
  - PE1 as the primary path.
  - PE2 as the backup/alternate path.
- An iBGP session exists between the CE1 and CE2 routers.

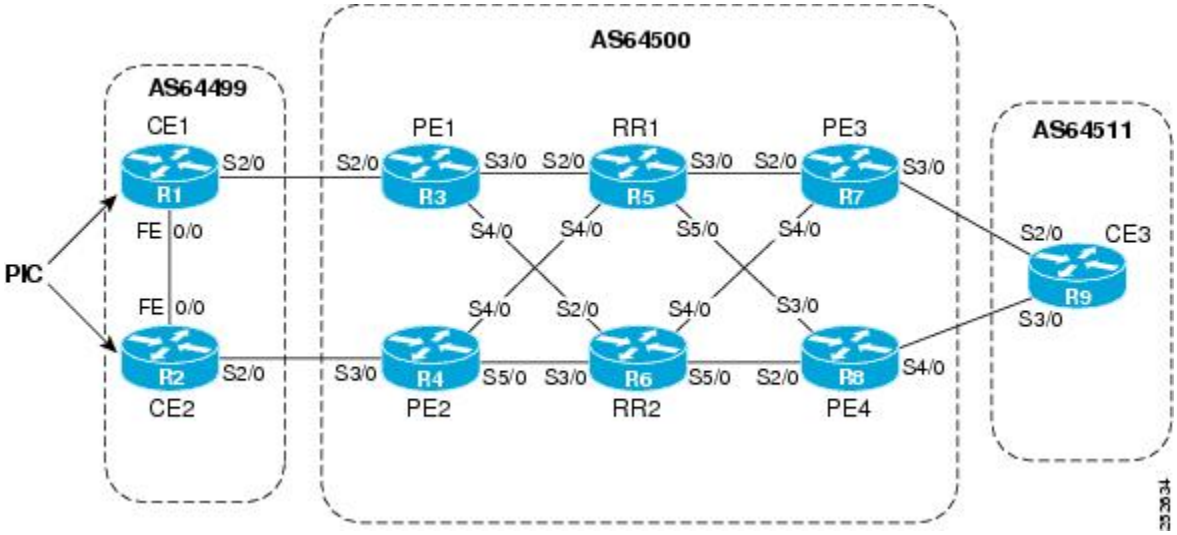
In this example, CE1 and CE2 are configured with the BGP PIC feature. BGP computes PE1 as the best path and PE2 as the backup/alternate path and installs both the routes into the RIB and Cisco Express Forwarding plane.

There should not be any policies set on CE1 and CE2 for the eBGP peers PE1 and PE2. Both CE routers must point to the eBGP route as next hop. On CE1, the next hop to reach CE3 is through PE1, so PE1 is the best path to reach CE3. On CE2, the best path to reach CE3 is PE2. CE2 advertises itself as the next hop to CE1, and CE1 does the same to CE2. As a result, CE1 has two paths for the specific prefix and it usually selects the directly connected eBGP path over the iBGP path according to the best path selection rules. Similarly, CE2 has two paths--an eBGP path through PE2 and an iBGP path through CE1-PE1.

When the CE1-PE1 link goes down, Cisco Express Forwarding detects the link failure and points the forwarding object to the backup/alternate node CE2. Traffic is quickly rerouted due to local fast convergence in Cisco Express Forwarding.

If the CE1-PE1 link or PE1 goes down and BGP PIC is enabled on CE1, BGP recomputes the best path, removing the next hop PE1 from RIB and reinstalling CE2 as the next hop into the RIB and Cisco Express Forwarding. CE1 automatically gets a backup/alternate repair path into Cisco Express Forwarding and the traffic loss during forwarding is now in subseconds, thereby achieving fast convergence.

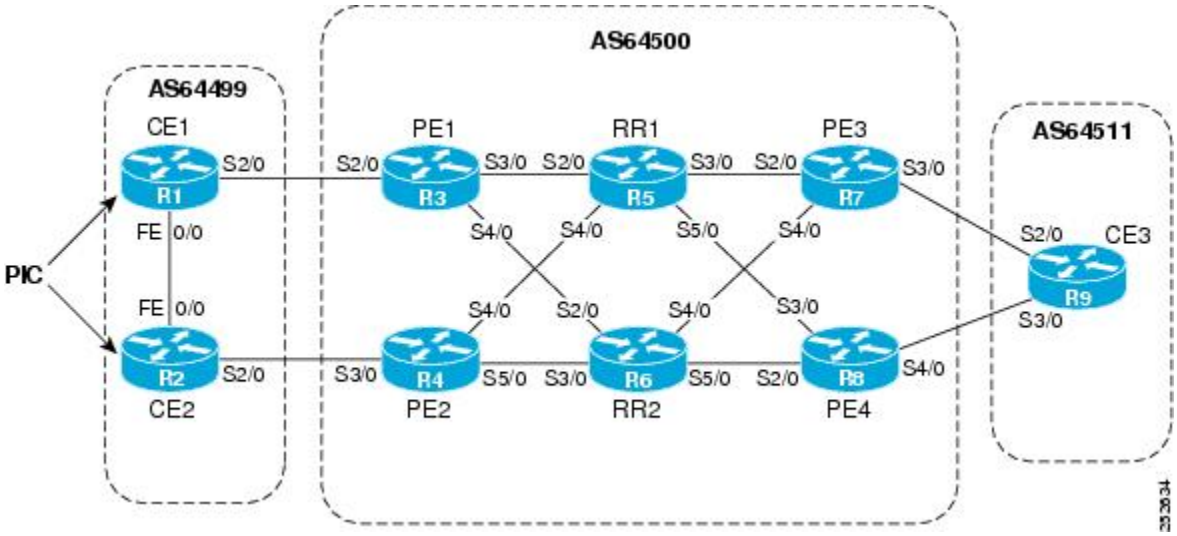
Figure 3: Using BGP PIC in a Dual CE, Dual PE Network



IP MPLS PE-CE Link Protection for the Primary or Backup-Alternate Path

The figure below shows a network that uses the BGP PIC feature on CE1 and CE2. The network includes the following components:

Figure 4: Using BGP PIC in a Dual CE, Dual PE Network



- eBGP sessions exist between the PE and CE routers.
- The PE routers are VPNv4 iBGP peers with reflect routers in the MPLS network.
- Traffic from CE1 uses PE1 to reach the network 192.168.9.0/24 through router CE3.
- CE3 is dual-homed with PE3 and PE4.
- PE1 has two paths to reach CE3 from the reflect routers:
  - PE3 is the primary path with the next hop as a PE3 address.
  - PE4 is the backup/alternate path with the next hop as a PE4 address.

In this example, all the PE routers can be configured with the BGP PIC feature under IPv4 or VPNv4 address families.

For BGP PIC to work in BGP for PE-CE link protection, set the policies on PE3 and PE4 for prefixes received from CE3 so that one of the PE routers acts as the primary and the other as the backup/alternate. Usually, this is done using local preference and giving better local preference to PE3. In the MPLS cloud, traffic internally flows through PE3 to reach CE3. Thus, PE1 has PE3 as the best path and PE4 as the second path.

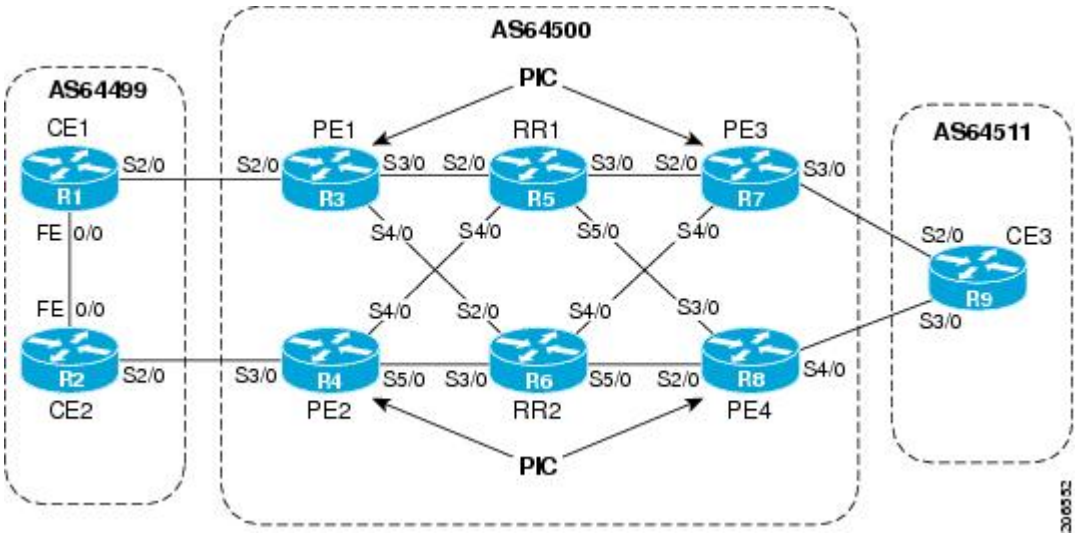
When the PE3-CE3 link goes down, Cisco Express Forwarding detects the link failure, and PE3 recomputes the best path, selects PE4 as the best path, and sends a withdraw message for the PE3 prefix to the reflect routers. Some of the traffic goes through PE3-PE4 until BGP installs PE4 as the best path route into the RIB and Cisco Express Forwarding. PE1 receives the withdraw, recomputes the best path, selects PE4 as the best path, and installs the routes into the RIB and Cisco Express Forwarding plane.

Thus, with BGP PIC enabled on PE3 and PE4, Cisco Express Forwarding detects the link failure and does in-place modification of the forwarding object to the backup/alternate node PE4 that already exists in Cisco Express Forwarding. PE4 knows that the backup/alternate path is locally generated and routes the traffic to the egress port connected to CE3. This way, traffic loss is minimized and fast convergence is achieved.

IP MPLS PE-CE Node Protection for Primary or Backup-Alternate Path

The figure below shows a network that uses the BGP PIC feature on all the PE routers in an MPLS network.

Figure 5: Enabling BGP PIC on All PEs Routers in the MPLS Network



The network includes the following components:

- eBGP sessions exist between the PE and CE routers.
- The PE routers are VPNv4 iBGP peers with reflect routers in the MPLS network.
- Traffic from CE1 uses PE1 to reach the network 192.168.9.0/24 through router CE3.
- CE3 is dual-homed with PE3 and PE4.
- PE1 has two paths to reach CE3 from the reflect routers:
  - PE3 is the primary path with the next hop as a PE3 address.
  - PE4 is the backup/alternate path with the next hop as a PE4 address.

In this example, all the PE routers are configured with the BGP PIC feature under IPv4 and VPNv4 address families.

For BGP PIC to work in BGP for the PE-CE node protection, set the policies on PE3 and PE4 for the prefixes received from CE3 such that one of the PE routers acts as primary and the other as backup/alternate. Usually, this is done using local preference and giving better local preference to PE3. In the MPLS cloud, traffic internally flows through PE3 to reach CE3. So, PE1 has PE3 as the best path and PE4 as the second path.

When PE3 goes down, PE1 knows about the removal of the host prefix by IGPs in subseconds, recomputes the best path, selects PE4 as the best path, and installs the routes into the RIB and Cisco Express Forwarding plane. Normal BGP convergence will happen while BGP PIC is redirecting the traffic through PE4, and packets are not lost.

Thus, with BGP PIC enabled on PE3, Cisco Express Forwarding detects the node failure on PE3 and points the forwarding object to the backup/alternate node PE4. PE4 knows that the backup/alternate path is locally generated and routes the traffic to the egress port using the backup/alternate path. This way, traffic loss is minimized.

No Local Policies Set on the PE Routers

PE1 and PE2 point to the eBGP CE paths as the next hop with no local policy. Each of the PE routers receives the other's path, and BGP calculates the backup/alternate path and installs it into Cisco Express Forwarding, along with its own eBGP path towards CE as the best path. The limitation of the MPLS PE-CE link and node protection solutions is that you cannot change BGP policies. They should work without the need for a best-external path.

Local Policies Set on the PE Routers

Whenever there is a local policy on the PE routers to select one of the PE routers as the primary path to reach the egress CE, the **bgp advertise-best-external** command is needed on the backup/alternate node PE3 to propagate the external CE routes with a backup/alternate label into the route reflectors and the far-end PE routers.

Enable BGP PIC

Because many service provider networks contain many VRFs, the BGP PIC allows you to configure, at a time, the BGP PIC feature for all VRFs.

- VPNv4 address family configuration mode protects all VRFs.
- VRF-IPv4 address family configuration mode protects only IPv4 VRFs.

- Router configuration mode protects prefixes in the global routing table.

Cisco Express Forwarding Recursion

Recursion is the ability to find the next longest matching path when the primary path goes down.

If BGP PIC is not installed, and if the next hop to a prefix fails, Cisco Express Forwarding finds the next path to reach the prefix by recursing through the FIB to find the next longest matching path to the prefix. This recurstion mechanism is useful when the next hop is multiple hops away and there is more than one way of reaching the next hop.

However, with the BGP PIC feature, you may want to disable Cisco Express Forwarding recursion for the following reasons:

- Recursion slows down convergence when Cisco Express Forwarding searches all the FIB entries.
- BGP PIC Edge already precomputes an alternate path. It therefore eliminates the need for Cisco Express Forwarding recursion.

When the BGP PIC functionality is enabled, Cisco Express Forwarding recursion is disabled by default for two conditions:

- For next hops learned with a /32 network mask (host routes)
- For next hops that are directly connected.

For all other cases, Cisco Express Forwarding recursion is enabled.

You can issue the **bgp recursion host** command to disable or enable Cisco Express Forwarding recursion for BGP host routes. This provision is part of the BGP PIC functionality.

Note When the BGP PIC feature is enabled, by default, **bgp recursion host** is configured for VPNv4 and VPNv6 address families and disabled for IPv4 and IPv6 address families.

To disable or enable Cisco Express Forwarding recursion for BGP directly connected next hops, run the **disable-connected-check** command.

How to Configure BGP PIC

Configuring BGP PIC

Because many service provider networks contain many VRFs, the BGP PIC feature allows you to configure the BGP PIC feature for all VRFs at once.

- VPNv4 address family configuration mode protects all the VRFs.
- VRF-IPv4 address family configuration mode protects only IPv4 VRFs.
- Router configuration mode protects prefixes in the global routing table.

For a full configuration example that includes configuring multiprotocol VRFs and shows output to verify that the feature is enabled, see the Example: Configuring BGP PIC.

Before you begin

- If you are implementing the BGP PIC feature in an MPLS VPN, ensure that the network is working properly before configuring the BGP PIC feature. See the *MPLS: Layer 3 VPNs Configuration Guide* for more information.
- If you are implementing the BGP PIC feature in an MPLS VPN, configure multiprotocol VRFs, which allow you to share route-target policies (import and export) between IPv4 and IPv6 or to configure separate route-target policies for IPv4 and IPv6 VPNs. For information about configuring multiprotocol VRFs, see MPLS VPN--VRF CLI for IPv4 and IPv6 VPNs .
- Ensure that the CE router is connected to the network by at least two paths.

Procedure

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device> enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	**router bgp** *autonomous-system-number* **Example:** `Device(config)# router bgp 40000`	Enters router configuration mode for the specified routing process.

Step 4	Do one of the following:
	• **address-family ipv4** [**unicast** | **vrf** *vrf-name*]
	• or
	• **address-family vpnv4** [**unicast**]
	Example:
	`Device(config-router)# address-family ipv4 unicast`
	Example:
	`Device(config-router)# address-family vpnv4`
		Specifies the IPv4 or VPNv4 address family and enters address family configuration mode.
		• The **unicast** keyword specifies the IPv4 or VPNv4 unicast address family.
		• The **vrf** keyword and *vrf-name* argument specify the name of the virtual routing and forwarding (VRF) instance to associate with subsequent IPv4 address family configuration mode commands.

	Command or Action	Purpose
Step 5	**bgp additional-paths install** **Example:** `Device(config-router-af)# bgp additional-paths install`	Calculates a backup/alternate path and installs it into the RIB and Cisco Express Forwarding.
Step 6	**neighbor** *ip-address* **remote-as** *autonomous-system-number* **Example:** `Device(config-router-af)# neighbor 192.168.1.1 remote-as 45000`	Adds the IP address of the neighbor in the specified autonomous system to the IPv4 multiprotocol BGP neighbor table of the local router. • By default, neighbors that are defined using the **neighbor remote-as** command in router configuration mode exchange only IPv4 unicast address prefixes. To exchange other address prefix types, neighbors must also be activated using the **neighbor activate** command in address family configuration mode for the other prefix types.
Step 7	**neighbor** *ip-address* **activate** **Example:** `Device(config-router-af)# neighbor 192.168.1.1 activate`	Enables the neighbor to exchange prefixes for the IPv4 unicast address family with the local router.
Step 8	**bgp recursion host** **Example:** `Device(config-router-af)# bgp recursion host`	(Optional) Enables the recursive-via-host flag for IPv4, VPNv4, and VRF address families. • When the BGP PIC feature is enabled, Cisco Express Forwarding recursion is disabled. Under most circumstances, you do not want to enable recursion when BGP PIC is enabled.
Step 9	**neighbor** *ip-address* **fall-over** [**bfd** |**route-map** *map-name*] **Example:** `Device(config-router-af)# neighbor 192.168.1.1 fall-over bfd`	Enables BFD protocol support to detect when a neighbor has gone away, which can occur within a subsecond.
Step 10	**end** **Example:** `Device(config-router-af)# end`	Exits address family configuration mode and returns to privileged EXEC mode.

Disabling BGP PIC Core

BGP PIC core feature is enabled by default. Use the following configuration to disable the BGP PIC core feature.

Note Use the **cef table output-chain build favor convergence-speed** command in global configuration mode to re-enable the BGP PIC core feature.

Procedure

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device> enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	**cef table output-chain build favor memory-utilization** **Example:** `Device(config)# cef table output-chain build favor memory-utilization`	Configures memory characteristics for Cisco Express Forwarding table output chain building for the forwarding of packets through the network.
Step 4	**end** **Example:** `Device(config)# end`	Exits global configuration mode and returns to privileged EXEC mode.

Configuration Examples for BGP PIC

Example: Configuring BGP PIC

The following example shows how to configure the BGP PIC feature in VPNv4 address family configuration mode, which enables the feature on all VRFs. In the following example, there are two VRFs defined: blue and green. All the VRFs, including those in VRFs blue and green, are protected by backup/alternate paths.

```
vrf definition test1
 rd 400:1
 route-target export 100:1
 route-target export 200:1
```

```
 route-target export 300:1
 route-target export 400:1
 route-target import 100:1
 route-target import 200:1
 route-target import 300:1
 route-target import 400:1
 address-family ipv4
 exit-address-family
 exit
!
 interface GigabitEthernet 0/0
 vrf forwarding test1
 ip address 10.0.0.1 255.0.0.0
 exit
router bgp 3
 no synchronization
 bgp log-neighbor-changes
 redistribute static
 redistribute connected
 neighbor 10.6.6.6 remote-as 3
 neighbor 10.6.6.6 update-source Loopback0
 neighbor 10.7.7.7 remote-as 3
 neighbor 10.7.7.7 update-source Loopback0
 no auto-summary
 !
 address-family vpnv4
  bgp additional-paths install
  neighbor 10.6.6.6 activate
  neighbor 10.6.6.6 send-community both
  neighbor 10.7.7.7 activate
  neighbor 10.7.7.7 send-community both
 exit-address-family
 !
 address-family ipv4 vrf blue
  import path selection all
  import path limit 10
  no synchronization
  neighbor 10.11.11.11 remote-as 1
  neighbor 10.11.11.11 activate
 exit-address-family
 !
 address-family ipv4 vrf green
  import path selection all
  import path limit 10
  no synchronization
  neighbor 10.13.13.13 remote-as 1
  neighbor 10.13.13.13 activate
 exit-address-family
```

The following **show vrf detail** command output shows that the BGP PIC feature is enabled:

```
Router# show vrf detail
VRF test1 (VRF Id = 1); default RD 400:1; default VPNID <not set>
  Interfaces:
    Se4/0
Address family ipv4 (Table ID = 1 (0x1)):
  Export VPN route-target communities
    RT:100:1                RT:200:1                RT:300:1
    RT:400:1
  Import VPN route-target communities
    RT:100:1                RT:200:1                RT:300:1
    RT:400:1
  No import route-map
  No export route-map
```

```
  VRF label distribution protocol: not configured
  VRF label allocation mode: per-prefix
  Prefix protection with additional path enabled
Address family ipv6 not active.
```

Example: Configuring IPv6 BGP PIC Edge

On Primary PE

```
router bgp 100
address-family ipv6 vrf V1
 bgp  additional-paths install
 neighbor 2.2.2.2 activate
  neighbor 2.2.2.2 send-community both
  neighbor 2.2.2.2 next-hop-self
  neighbor 2.2.2.2 send-label
  exit-address-family
```

On Backup PE

```
router bgp 100
address-family ipv6 vrf V1
bgp advertise-best-external
neighbor 10.0.0.1 activate
neighbor 10.0.0.1 send-community both
neighbor 10.0.0.1 next-hop-self
neighbor 10.0.0.1 send-label
```

Example: Disabling BGP PIC Core

The following example shows how to disable the BGP PIC core in global configuration mode.

```
Device> enable
Device# configure terminal
Device(config)# cef table output-chain build favor memory-utilization
Device(config)# end
```

Verification Examples for BGP PIC

Example: Displaying Backup Alternate Paths for BGP PIC

The command output in the following example shows that the VRFs in VRF blue have backup/alternate paths:

```
Device# show ip bgp vpnv4 vrf blue 10.0.0.1

BGP routing table entry for 10:12:10.0.0.1/24, version 88
Paths: (4 available, best #1, table blue)
  Additional-path
  Advertised to update-groups:
     6
  1, imported path from 12:23:10.0.0.1/24
    10.3.3.3 (metric 21) from 10.6.6.6 (10.6.6.6)
      Origin incomplete, metric 0, localpref 200, valid, internal, best
```

```
      Extended Community: RT:12:23
      Originator: 10.3.3.3, Cluster list: 10.0.0.1 , recursive-via-host
      mpls labels in/out nolabel/37
  1, imported path from 12:23:10.0.0.1/24
    10.13.13.13 (via green) from 10.13.13.13 (10.0.0.2)
      Origin incomplete, metric 0, localpref 100, valid, external
      Extended Community: RT:12:23 , recursive-via-connected
  1, imported path from 12:23:10.0.0.1/24
    10.3.3.3 (metric 21) from 10.7.7.7 (10.7.7.7)
      Origin incomplete, metric 0, localpref 200, valid, internal
      Extended Community: RT:12:23
      Originator: 10.3.3.3, Cluster list: 10.0.0.1 , recursive-via-host
      mpls labels in/out nolabel/37
  1
    10.11.11.11 from 10.11.11.11 (1.0.0.1)
      Origin incomplete, metric 0, localpref 100, valid, external, backup/repair
      Extended Community: RT:11:12 , recursive-via-connected
```

The command output in the following example shows that the VRFs in VRF green have backup/alternate paths:

```
Device# show ip bgp vpnv4 vrf green 10.0.0.1

BGP routing table entry for 12:23:10.0.0.1/24, version 87
Paths: (4 available, best #4, table green)
  Additional-path
  Advertised to update-groups:
     5
  1, imported path from 11:12:10.0.0.1/24
    10.11.11.11 (via blue) from 10.11.11.11 (1.0.0.1)
      Origin incomplete, metric 0, localpref 100, valid, external
      Extended Community: RT:11:12 , recursive-via-connected
  1
    10.3.3.3 (metric 21) from 10.7.7.7 (10.7.7.7)
      Origin incomplete, metric 0, localpref 200, valid, internal
      Extended Community: RT:12:23
      Originator: 10.3.3.3, Cluster list: 10.0.0.1 , recursive-via-host
      mpls labels in/out nolabel/37
  1
    10.13.13.13 from 10.13.13.13 (10.0.0.2)
      Origin incomplete, metric 0, localpref 100, valid, external, backup/repair
      Extended Community: RT:12:23 , recursive-via-connected
  1
    10.3.3.3 (metric 21) from 10.6.6.6 (10.6.6.6)
      Origin incomplete, metric 0, localpref 200, valid, internal, best
      Extended Community: RT:12:23
      Originator: 10.3.3.3, Cluster list: 10.0.0.1 , recursive-via-host
      mpls labels in/out nolabel/37
```

The command output in the following example shows the BGP routing table entries for the backup and alternate paths:

```
Device# show ip bgp 10.0.0.1 255.255.0.0

BGP routing table entry for 10.0.0.1/16, version 123
Paths: (4 available, best #3, table default)
  Additional-path
  Advertised to update-groups:
     2          3
  Local
    10.0.101.4 from 10.0.101.4 (10.3.3.3)
      Origin IGP, localpref 100, weight 500, valid, internal
  Local
```

```
    10.0.101.3 from 10.0.101.3 (10.4.4.4)
      Origin IGP, localpref 100, weight 200, valid, internal
  Local
    10.0.101.2 from 10.0.101.2 (10.1.1.1)
      Origin IGP, localpref 100, weight 900, valid, internal, best
  Local
    10.0.101.1 from 10.0.101.1 (10.5.5.5)
      Origin IGP, localpref 100, weight 700, valid, internal, backup/repair
```

The command output in the following example shows the routing information base entries for the backup and alternate paths:

```
Device# show ip route repair-paths 10.0.0.1 255.255.0.0

Routing entry for 10.0.0.1/16
  Known via "bgp 10", distance 200, metric 0, type internal
  Last update from 10.0.101.2 00:00:56 ago
  Routing Descriptor Blocks:
  * 10.0.101.2, from 10.0.101.2, 00:00:56 ago
      Route metric is 0, traffic share count is 1
      AS Hops 0
      MPLS label: none
    [RPR]10.0.101.1, from 10.0.101.1, 00:00:56 ago
      Route metric is 0, traffic share count is 1
      AS Hops 0
      MPLS label: none
```

The command output in the following example shows the Cisco Express Forwarding/forwarding information base entries for the backup and alternate paths:

```
Device# show ip cef 10.0.0.1 255.255.0.0 detail

10.0.0.1/16, epoch 0, flags rib only nolabel, rib defined all labels
  recursive via 10.0.101.2
    attached to  GigabitEthernet0/2
  recursive via 10.0.101.1, repair
    attached to  GigabitEthernet0/2
```

Example: Verifying BGP PIC Edge

show ip bgp all summary

```
Router# show ip bgp all summary

For address family: IPv4 Unicast
BGP router identifier 65.1.160.1, local AS number 100
BGP table version is 1, main routing table version 1
 Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
3.3.3.3         4          100     154      50        1    0    0 00:33:33        0
4.4.4.4         4          100   16158   10579        1    0    0 6d14h           0

For address family: IPv6 Unicast
BGP router identifier 65.1.160.1, local AS number 100
BGP table version is 363156, main routing table version 363156

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
3.3.3.3         4          100     154      50   363156    0    0 00:33:33        0
4.4.4.4         4          100   16158   10579   363156    0    0 6d14h           0
For address family: VPNv4 Unicast
BGP router identifier 65.1.160.1, local AS number 100
```

```
BGP table version is 120, main routing table version 120
1 network entries using 156 bytes of memory
1 path entries using 80 bytes of memory
1/1 BGP path/bestpath attribute entries using 168 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
1 BGP extended community entries using 24 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 452 total bytes of memory
BGP activity 19514/15612 prefixes, 121023/113221 paths, scan interval 60 secs
 Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
3.3.3.3         4          100     154      50      120    0    0 00:33:34        1
4.4.4.4         4          100   16158   10579      120    0    0 6d14h           0
For address family: VPNv6 Unicast
BGP router identifier 65.1.160.1, local AS number 100
BGP table version is 291083, main routing table version 291083
3901 network entries using 702180 bytes of memory
7801 path entries using 842508 bytes of memory
2/2 BGP path/bestpath attribute entries using 336 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
1 BGP extended community entries using 24 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 1545072 total bytes of memory
BGP activity 19514/15612 prefixes, 121023/113221 paths, scan interval 60 secs

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down State/PfxRcd
3.3.3.3         4          100     154      50   291083    0    0 00:33:35     3900
4.4.4.4         4          100   16159   10579   291083    0    0 6d14h        3900
```

show ipv6 route vrf

```
Router# show ipv6 route vrf vpn1 800::0/64

Routing entry for 800::/64
  Known via "bgp 100", distance 200, metric 0, type internal
  Route count is 1/1, share count 0
  Routing paths:
    3.3.3.3%default indirectly connected [Host Res]
      MPLS label: 166
      Last updated 00:43:36 ago
    4.4.4.4%default indirectly connected [Repair] [Host Res]
      MPLS label: 7943
      Last updated 01:08:22 ago
```

show ipv6 cef

```
Router# show ipv6 cef 900::0 in

900::/64, epoch 2, flags [rlbls], RIB[B], refcnt 5, per-destination sharing
  sources: RIB
  feature space:
    IPRM: 0x00018000
    Broker: linked, distributed at 4th priority
    LFD: 900::/64 0 local labels
        contains path extension list
  ifnums: (none)
  path list 3C5993C8, 3801 locks, per-destination, flags 0x34D [shble, hvsh, rif, hwcn,
bldmp, bgp]
    path 3C594928, share 1/1, type recursive, for IPv6, flags [must-be-lbld]
      MPLS short path extensions: MOI flags = 0x0 label 2316
      recursive via 3.3.3.3[IPv4:Default] label 2316, fib 3D36FC7C, 1 terminal fib,
```

```
v4:Default:3.3.3.3/32
      path list 3C599968, 127 locks, per-destination, flags 0x4D [shble, hvsh, rif, hwcn]
          path 3C594ED8, share 1/1, type attached nexthop, for IPv4
            MPLS short path extensions: MOI flags = 0x0 label 84
           nexthop 21.1.1.2  GigabitEthernet0/1 label 84, IP adj out of  GigabitEthernet0/1
 , addr 21.1.1.2 3CCB3320
 path 3C594AC8, share 1/1, type recursive, for IPv6, flags [must-be-lbld, rpr]
      MPLS short path extensions: MOI flags = 0x0 label 57
      recursive via 4.4.4.4[IPv4:Default] label 57, repair, fib 3D4BCD74, 1 terminal fib,
v4:Default:4.4.4.4/32
      path list 3C599668, 129 locks, per-destination, flags 0x4D [shble, hvsh, rif, hwcn]
          path 3C594B98, share 1/1, type attached nexthop, for IPv4
            MPLS short path extensions: MOI flags = 0x0 label 19
           nexthop 21.1.6.1  GigabitEthernet0/0 label 19, IP adj out of GigabitEthernet0/0/0,
 addr 21.1.6.1 3D4D8BE0
  output chain:
loadinfo 3DCFE310, per-session, 2 choices, flags 0005, 4 locks
      flags [Per-session, for-rx-IPv6]
      translation map 3CF06908 owned by path list 3C5993C8, 1902 locks
        2 choices, 16 buckets, flags 0x1
        Path index      [  0  1  0  1  0  1  0  1  0  1  0  1  0  1  0  1 ]
        Repair path     [  -  X  -  X  -  X  -  X  -  X  -  X  -  X  -  X ]

        Path available  [  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X ]
        Translation map [  0  0  2  2  4  4  6  6  8  8 10 10 12 12 14 14 ]
      16 hash buckets
        < 0 > label 2316
              loadinfo 3D572F10, per-session, 1 choice, flags 0111, 3904 locks
                flags [Per-session, for-mpls-not-at-eos, indirection]
                1 hash bucket
                  < 0 > label 84
                        TAG adj out of  GigabitEthernet0/1, addr 21.1.1.2 3CCB3180
                Subblocks:
                  None
        < 1 > label 2316
              loadinfo 3D572F10, per-session, 1 choice, flags 0111, 3904 locks
                flags [Per-session, for-mpls-not-at-eos, indirection]
                1 hash bucket
                  < 0 > label 84
                        TAG adj out of  GigabitEthernet0/1, addr 21.1.1.2 3CCB3180
                Subblocks:
                  None
```

show vrf detail

```
Router# show vrf detail

VRF Mgmt-intf (VRF Id = 1); default RD <not set>; default VPNID <not set>
  New CLI format, supports multiple address-families
  Flags: 0x1808
  Interfaces:
    Gi0
Address family ipv4 unicast (Table ID = 0x1):
  Flags: 0x0
  No Export VPN route-target communities
  No Import VPN route-target communities
  No import route-map
  No global export route-map
  No export route-map
  VRF label distribution protocol: not configured
  VRF label allocation mode: per-prefix
Address family ipv6 unicast (Table ID = 0x1E000001):
  Flags: 0x0
```

```
  No Export VPN route-target communities
  No Import VPN route-target communities
  No import route-map
  No global export route-map
  No export route-map
  VRF label distribution protocol: not configured
VRF label allocation mode: per-prefix
Address family ipv4 multicast not active

VRF vpn1 (VRF Id = 2); default RD 100:1; default VPNID <not set>
  New CLI format, supports multiple address-families
  Flags: 0x180C
  Interfaces:
    BD13
Address family ipv4 unicast (Table ID = 0x2):
  Flags: 0x8000
  Export VPN route-target communities
    RT:100:1                RT:100:10
  Import VPN route-target communities
    RT:100:1                RT:100:10
  No import route-map
  No global export route-map
  No export route-map
  VRF label distribution protocol: not configured
  VRF label allocation mode: per-prefix
  Prefix protection with additional path enabled
Address family ipv6 unicast (Table ID = 0x1E000002):
  Flags: 0x8000
  Export VPN route-target communities
    RT:100:1                RT:100:10
  Import VPN route-target communities
    RT:100:1                RT:100:10
```

show platform hardware pp active feature cef database

```
Router# show platform hardware pp active feature cef database ipv6 800::0/64 0x1E000002

=== CEF Prefix ===
800::/64 -- next hop: UEA Load Balance (PI:0x10844c30, PD:0x1543d948)
               Route Flags: (0)
               Handles (PI:0x105f4d30) (PD:0x16042c80)
   HW Info:
    TCAM handle: 0x00000270    TCAM index: 0x000067af
    FID index  : 0x00008abb    EAID      : 0x00005608
    MET        : 0x0002a8aa    FID Count : 0x00000000
 === Load Balance OCE ===
   PI:0x10844c30, PD:0x1543d948
   FID Count: 0x00000001
      Load Balance HW Info:
     Hardware Index: 0
       FID Index: 0x00008abb    RW Index  : 0x00000000
       MET      : 0x0002a8aa    EAID      : 0x00005608
       EL3 Index : 0x00001529   EL2 Index : 0x00000000
       Hardware Rewrite Str: 00:00:00:00:00:00:00:00:00:00:00:00
 === Label OCE ===
  Label flags: 4
  Num Labels: 1

Out Labels: 166
  Out Backup Labels: 1048577
  Next OCE Type: Loadbalance OCE; Next OCE handle: 0x121452f8
 === Load Balance OCE ===
   PI:0x1067bf38, PD:0x121452f8
```

```
   FID Count: 0x00000001
      Load Balance HW Info:
     Hardware Index: 0
       FID Index: 0x0000609c    RW Index  : 0x00000000
       MET      : 0x0002a8a8    EAID      : 0x0000ee72
       EL3 Index : 0x00001527   EL2 Index : 0x00000000
       Hardware Rewrite Str: 00:00:00:00:00:00:00:00:00:00:00:00

=== Label OCE ===
**This is selected Label OCE**
  Label flags: 4
  Num Labels: 1
  Num Bk Labels: 1
  Out Labels: 18
  Out Backup Labels: 1048577
  Next OCE Type: Adjacency; Next OCE handle: 0x119f3800

=== Adjacency OCE ===
  Adj State: COMPLETE(0)   Address: 21.1.1.2
  Interface:  GigabitEthernet0/1   Protocol: TAG
  mtu:1500, flags:0x0, fixups:0x0, encap_len:14
  Handles (adj_id:0x000000f9) (PI:0x105f18d8) (PD:0x119f3800)
  Rewrite Str: d0:c2:82:cc:a5:cc:d0:c2:82:16:c4:8c:88:47

  HW Info:
    FID index: 0x0000605a    EL3 index: 0x0000100c   EL2 index: 0x00000000
    El2RW    : 0x00000108    MET index: 0x00032029    EAID     : 0x00001012
    HW ADJ FLAGS: 0x40
    Hardware MAC Rewrite Str: 00:00:00:00:00:00:00:00:00:00:00:00

=== Label OCE ===
  Label flags: 4
  Num Labels: 1
  Num Bk Labels: 1
  Out Labels: 7943
  Out Backup Labels: 1048577
  Next OCE Type: Loadbalance OCE; Next OCE handle: 0x12124f18

=== Load Balance OCE ===
   PI:0x10833fe8, PD:0x12124f18
   FID Count: 0x00000001
      Load Balance HW Info:
     Hardware Index: 0
       FID Index: 0x0000615c    RW Index  : 0x00000000
       MET      : 0x0002a894    EAID      : 0x0000aac7
       EL3 Index : 0x0000152b   EL2 Index : 0x00000000
       Hardware Rewrite Str: 00:00:00:00:00:00:00:00:00:00:00:00

=== Label OCE ===
**This is selected Label OCE**
  Label flags: 4
  Num Labels: 1
  Num Bk Labels: 1
  Out Labels: 19
  Out Backup Labels: 1048577
  Next OCE Type: Adjacency; Next OCE handle: 0x1211a648

=== Adjacency OCE ===
  Adj State: COMPLETE(0)   Address: 21.1.6.1
  Interface:  GigabitEthernet0/0  Protocol: TAG
  mtu:1500, flags:0x0, fixups:0x0, encap_len:14
  Handles (adj_id:0x0000110b) (PI:0x106a94f0) (PD:0x1211a648)
  Rewrite Str: d0:c2:82:17:71:00:d0:c2:82:16:c4:80:88:47
```

```
HW Info:
  FID index: 0x00006fdf    EL3 index: 0x00001001    EL2 index: 0x00000000
  El2RW    : 0x0000010c    MET index: 0x0003202d    EAID     : 0x00001009
  HW ADJ FLAGS: 0x40
  Hardware MAC Rewrite Str: 00:00:00:00:00:00:00:00:00:00:00:00
```

Additional References

Related Documents

Related Topic	Document Title
Cisco IOS commands	Cisco IOS Master Command List, All Releases
BGP commands	Cisco IOS IP Routing: BGP Command Reference
Basic MPLS VPNs	Configuring MPLS Layer 3 VPNs
A failover feature that creates a new path after a link or node failure	MPLS VPN--BGP Local Convergence
Configuring multiprotocol VRFs	MPLS VPN--VRF CLI for IPv4 and IPv6 VPNs

MIBs

MIB	MIBs Link
None	To locate and download MIBs for selected platforms, Cisco software releases, and feature sets, use Cisco MIB Locator found at the following URL: http://www.cisco.com/go/mibs

RFCs

RFC	Title
RFC 1771	*A Border Gateway Protocol 4 (BGP-4)*
RFC 2547	*BGP/MPLS VPNs*

Technical Assistance

Description	Link
The Cisco Support and Documentation website provides online resources to download documentation, software, and tools. Use these resources to install and configure the software and to troubleshoot and resolve technical issues with Cisco products and technologies. Access to most tools on the Cisco Support and Documentation website requires a Cisco.com user ID and password.	http://www.cisco.com/cisco/web/support/index.html