Configuring Policy-Based Routing

This chapter describes how to configure policy based routing on the Cisco NX-OS device.

This chapter includes the following sections:

- About Policy-Based Routing, on page 1
- Prerequisites for Policy-Based Routing, on page 3
- Guidelines and Limitations for Policy-Based Routing, on page 3
- Default Settings, on page 4
- Configuring Policy-Based Routing, on page 4
- Verifying the Policy-Based Routing Configuration, on page 7
- Configuration Examples for Policy Based-Routing, on page 7

About Policy-Based Routing

With policy-based routing, you can configure a defined policy for IPv4 and IPv6 traffic flows that lessens the reliance on routes derived from routing protocols. All packets received on an interface with policy-based routing enabled are passed through enhanced packet filters or route maps. The route maps dictate the policy that determines where to forward packets.

Policy-based routing includes the following features:

- Source-based routing—Routes traffic that originates from different sets of users through different connections across the policy routers.
- Quality of Service (QoS)—Differentiates traffic by setting the precedence or type of service (ToS) values in the IP packet headers at the periphery of the network and leveraging queuing mechanisms to prioritize traffic in the core or backbone of the network (see the Cisco Nexus 3600 NX-OS Quality of Service Configuration Guide).
- Load sharing—Distributes traffic among multiple paths based on the traffic characteristics.

Policy Route Maps

Each entry in a route map contains a combination of match and set statements. The match statements define the criteria for whether appropriate packets meet the particular policy (that is, the conditions to be met). The set clauses explain how the packets should be routed once they have met the match criteria.

You can mark the route-map statements as permit or deny. You can interpret the statements as follows:

- If the statement is marked as permit and the packets meet the match criteria, the set clause is applied. One of these actions involves choosing the next hop.
- If a statement is marked as deny, the packets that meet the match criteria are sent back through the normal forwarding channels, and destination-based routing is performed.
- If the statement is marked as permit and the packets do not match any route-map statements, the packets are sent back through the normal forwarding channels, and destination-based routing is performed.

Note Policy routing is specified on the interface that receives the packets, not on the interface from which the packets are sent.

Set Criteria for Policy-Based Routing

The Cisco Nexus 3600 platform switches support the following set commands for route maps used in policy-based routing:

- set {ip | ipv6} next-hop address1 [address2...] [load-share]
- set interface null0

These set commands are mutually exclusive within the route-map sequence.

In the first command, the IP address specifies the adjacent next-hop router in the path toward the destination to which the packets should be forwarded. The first IP address associated with a currently up connected interface is used to route the packets.

Note You can optionally configure this command for next-hop addresses to load balance traffic for up to 32 IP addresses. In this case, Cisco NX-OS sends all traffic for each IP flow to a particular IP next-hop address.

If the packets do not meet any of the defined match criteria, those packets are routed through the normal destination-based routing process.

Route-Map Processing Logic

When a packet is received on an interface that is configured with a route map, the forwarding logic processes each route-map statement according to the sequence number.

If the route-map statement encountered is a route-map...permit statement, the packet is matched against the criteria in the match command. This command may refer to an ACL that has one or more access control entries (ACEs). If the packet matches the permit ACEs in the ACL, the policy-based routing logic executes the action specified by the set command on the packet.

If the route-map statement encountered is a route-map... deny statement, the packet is matched against the criteria in the match command. This command may refer to an ACL that has one or more ACEs. If the packet matches the permit ACEs in the ACL, policy-based routing processing terminates, and the packet is routed using the default IP routing table.

Note The set command has no effect inside a route-map... deny statement.

- If the route-map configuration does not contain a match statement, the policy-based routing logic executes the action specified by the set command on the packet. All packets are routed using policy-based routing.
- If the route-map configuration references a match statement but the match statement references a non-existing ACL or an existing ACL without any access control entries (ACEs), the packet is routed using the default routing table.
- If the next-hop specified in the set {ip | ipv6} next-hop command is down, is not reachable, or is removed, the packet is routed using the default routing table.

Beginning Cisco NX-OS Release 9.2(3), you can balance policy-based routing traffic on the Cisco Nexus 36180YC-R switch, if the next hop is recursive over ECMP paths using the next-hop ip-address load-share command. For all the next hop routing requests, the Routing Profile Manager (RPM) resolves them using unicast Routing Information Base (uRIB) and program all ECMP paths, which helps to uniformly load balance all the ECMP paths. PBR over ECMP is supported only on IPv4.

Policy-Based Routing Filtering Options

You can identify traffic by using additional options. The following list includes most but not all additional filtering options.

Policy-based routing ACLs support the following additional filtering options:

- Layer 3 source and/or destination address
- TCP and UDP ports

Prerequisites for Policy-Based Routing

You can identify traffic by using additional options. The following list includes most but not all additional filtering options.

Policy-based routing ACLs support the following additional filtering options:

- Layer 3 source and/or destination address
- TCP and UDP ports

Guidelines and Limitations for Policy-Based Routing

Policy-based routing has the following configuration guidelines and limitations:

- Beginning with Cisco NX-OS Release 7.0(3)F3(3), Cisco Nexus 3600 platform switches support IPv4 and IPv6 policy-based routing. For these switches, PBR policy has a higher priority over attached and local routes. Explicit allowed listing might be required if protocol neighbors are directly attached.
- A policy-based routing route map can have only one match statement per route-map statement.

- A match command cannot refer to more than one ACL in a route map used for policy-based routing.
- The same route map can be shared among different interfaces for policy-based routing as long as the interfaces belong to the same virtual routing and forwarding (VRF) instance.
- Using a prefix list as a match criteria is not supported. Do not use a prefix list in a policy-based routing route map.
- Policy-based routing supports only unicast traffic. Multicast traffic is not supported.
- Policy-based routing is supported with Layer 3 port-channel subinterfaces.
- An ACL used in a policy-based routing route map cannot include deny access control entries (ACEs).
- Policy-based routing is supported only in the default system routing mode.
- Policy-based routing traffic cannot be balanced if the next hop is recursive over ECMP paths. Instead, use the **set {ip | ipv6} next-hop** *ip-address* **load-share** command to specify the adjacent next hops.
- Policy-based routing is not supported with VXLAN.
- Policy-based routing policy statistics are not supported.

Default Settings

Table below lists the default settings for policy-based routing parameters.

Table 1: Default Policy-based Routing Parameters

Parameters	Default
Policy-based routing	Disabled

Configuring Policy-Based Routing

Enabling the Policy-Based Routing Feature

You must enable the policy-based routing feature before you can configure a route policy.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **feature pbr**	Enables the policy-based routing feature.

	Command or Action	Purpose
	Example: `switch(config)# feature pbr`	Use the no form of this command to disable the policy-based routing feature. **Note** The no feature pbr command removes the policies applied under the interfaces. It does not remove the ACL or route-map configuration nor does it create a system checkpoint.
Step 3	(Optional) **show feature** **Example:** `switch(config)# show feature`	Displays enabled and disabled features.
Step 4	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Saves this configuration change.

Enabling the Policy-Based Routing over ECMP

PBR over ECMP is not enabled by default. You must enable the policy-based routing feature before you can configure a route policy.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **feature pbr** **Example:** `switch(config)# feature pbr`	Enables the policy-based routing feature. Use the **no** form of this command to disable the policy-based routing feature. **Note** The **no feature pbr** command removes the policies applied under the interfaces. It does not remove the ACL or route-map configuration nor does it create a system checkpoint.
Step 3	(Optional) **show feature** **Example:**	Displays enabled and disabled features.

	Command or Action	Purpose
	`switch(config)# show feature`	

Step 4	**[no] hardware profile pbr ecmp paths max-paths**
	Example:
	`switch(config)# hardware profile pbr ecmp paths max-paths 12`
	`Warning!!: The pbr ecmp path limits have been changed.`
	`Please reload the switch now for the change to take effect.`
	`switch(config)#`
	`switch(config)# no hardware profile pbr ecmp paths max-paths 12`
	`Warning!!: The pbr ecmp path limits have been changed.`
	`Please reload the switch now for the change to take effect.`
	`switch(config)#`
		Configure the number of ECMP paths for IP next hop. However, the traffic may not go through all the paths unless you explicitly configure the load share in the set IP next hop. Whenever you remove or modify the PBR ECMP paths, the changes will take effect only after next reload. The range is from 1 through 64.

Step 5	**show system internal rpm state**	Displays the currently configured and operational values of PBR ECMP paths.

Configuring a Route Policy

You can use route maps in policy-based routing to assign routing policies to the inbound interface.

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 1/2`
	`switch(config-if)#`
		Enters interface configuration mode.

Step 3	{ **ip** | **ipv6** } **policy route-map** *map-name* **Example:** `switch(config-if)# ip policy route-map Testmap`	Assigns a route map for IPv4 or IPv6 policy-based routing to the interface.

Step 4	**route-map map-name [permit | deny] [seq]**
	Example:
	`switch(config-if)# route-map Testmap`
	`switch(config-route-map)#`
		Creates a route map or enters route-map configuration mode for an existing route map. Use seq to order the entries in a route map.

	Command or Action	Purpose
Step 5	**match {ip | ipv6} address access-list-name name** [*name...*] **Example:** `switch(config-route-map)# match ip address access-list-name ACL1`	Matches an IPv4 or IPv6 address against one or more IP or IPv6 access control lists (ACLs). This command is used for policy-based routing and is ignored by route filtering or redistribution.
Step 6	(Optional) **set ip next-hop address1 [address2...]** [*load-share*] **Example:** `switch(config-route-map)# set ip next-hop 192.0.2.1`	Sets the IPv4 next-hop address for policy-based routing. This command uses the first valid next-hop address if multiple addresses are configured. Use the optional load-share keyword to load balance traffic across a maximum of 32 next-hop addresses.
Step 7	**set ipv6 next-hop address1 [address2...]** [*load-share*] **Example:** `switch(config-route-map)# set ipv6 next-hop 2001:0DB8::1`	Sets the IPv6 next-hop address for policy-based routing. This command uses the first valid next-hop address if multiple addresses are configured. Use the optional load-share keyword to load balance traffic across a maximum of 32 next-hop addresses.
Step 8	(Optional) **set interface null0** **Example:** `switch(config-route-map)# set interface null0`	Sets the interface used for routing. Use the null0 interface to drop packets.
Step 9	(Optional) **copy running-config startup-config** **Example:** `switch(config-route-map)# copy running-config startup-config`	Saves this configuration change.

Verifying the Policy-Based Routing Configuration

To display policy-based routing configuration information, perform one of the following tasks:

Command	Purpose
show [ip | ipv6] policy [*name*]	Displays information about an IPv4 or IPv6 policy.

Configuration Examples for Policy Based-Routing

This example shows how to configure a simple route policy on an interface:

```
ip access-list pbr-sample
permit tcp host 10.1.1.1 host 192.168.2.1 eq 80
!
route-map pbr-sample
match ip address pbr-sampl
set ip next-hop 192.168.1.1
!
route-map pbr-sample
interface ethernet 1/2
ip policy route-map pbr-sample
```

The following output verifies this configuration:

```
switch# show route-map pbr-sample
route-map pbr-sample, permit, sequence 10
Match clauses:

ip address (access-lists): pbr-sample
Set clauses:

ip next-hop 192.168.1.1
switch# show ip policy
Interface Route-map Status VRF-Name
Ethernet1/2 pbr-sample Active --
```

This example shows load sharing between ECMP and non ECMP paths:

```
switch# show run rpm
!Command: show running-config rpm
!Running configuration last done at: Sun Dec 23 16:02:32 2018
!Time: Sun Dec 23 16:06:13 2018

version 9.2(3) Bios:version 08.35
feature pbr

route-map policy1 pbr-statistics
route-map policy1 permit 10
  match ip address acl2
  set ip next-hop 131.1.1.2 load-share
route-map policy2 pbr-statistics
route-map policy2 permit 10
  match ip address acl2
  set ip next-hop verify-availability 131.1.1.2 track 1
  set ip next-hop verify-availability 30.1.1.2 track 2 load-share


interface Ethernet1/31
  ip policy route-map policy2
```

This example displays information about next hop routing request:

```
switch# show system internal rpm pbr ip nexthop
PBR IPv4 nexthop table for vrf default

30.1.1.2 Usable
  via 28.1.1.2 Ethernet1/18 a46c.2ae3.02a7

131.1.1.2 Usable
  via 111.1.1.2 Vlan81 8478.ac58.afc1
Usable
  via 112.1.1.2 Vlan82 8478.ac58.afc1
Usable
  via 113.1.1.2 Vlan83 8478.ac58.afc1
Usable
```

```
  via 114.1.1.2 Vlan84 8478.ac58.afc1
Usable
  via 115.1.1.2 Vlan85 8478.ac58.afc1
Usable
  via 116.1.1.2 Vlan86 8478.ac58.afc1
Usable
  via 117.1.1.2 Vlan87 8478.ac58.afc1
Usable
  via 118.1.1.2 Vlan88 8478.ac58.afc1
```

This example display routes from the unicast RIB:

```
switch# show ip route 130.1.1.2
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>

130.1.1.0/24, ubest/mbest: 8/0
    *via 111.1.1.2, Vlan81, [110/120], 00:07:57, ospf-1, inter
    *via 112.1.1.2, Vlan82, [110/120], 00:07:57, ospf-1, inter
    *via 113.1.1.2, Vlan83, [110/120], 00:07:57, ospf-1, inter
    *via 114.1.1.2, Vlan84, [110/120], 00:07:57, ospf-1, inter
    *via 115.1.1.2, Vlan85, [110/120], 00:07:57, ospf-1, inter
    *via 116.1.1.2, Vlan86, [110/120], 00:07:57, ospf-1, inter
    *via 117.1.1.2, Vlan87, [110/120], 00:07:57, ospf-1, inter
    *via 118.1.1.2, Vlan88, [110/120], 00:07:57, ospf-1, inter


switch# show ip route 30.1.1.2
IIP 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>

30.1.1.0/24, ubest/mbest: 1/0
    *via 28.1.1.2, [1/0], 00:38:36, static
```