Configuring Cisco Express Forwarding

This chapter describes Cisco Express Forwarding (CEF) on the switch. It also provides guidelines, procedures, and examples to configure this feature.

This chapter includes the following major sections:

Note For complete syntax and usage information for the switch commands used in this chapter, see the Cisco IOS Command Reference Guides for the Catalyst 4500 Series Switch.

If a command is not in the *Cisco Catalyst 4500 Series Switch Command Reference* , you can locate it in the Cisco IOS Master Command List, All Releases.

About CEF

This section contains information on the two primary components that comprise the CEF operation:

CEF Features

CEF is advanced Layer 3 IP switching technology that optimizes performance and scalability for large networks with dynamic traffic patterns or networks with intensive web-based applications and interactive sessions.

CEF provides the following features:

- Improves performance over the caching schemes of multilayer switches, which often flush the entire cache when information changes in the routing tables.

- Provides load balancing that distributes packets across multiple links based on Layer 3 routing information. If a network device discovers multiple paths to a destination, the routing table is updated with multiple entries for that destination. Traffic to that destination is then distributed among the various paths.

CEF stores information in several data structures rather than the route cache of multilayer switches. The data structures optimize lookup for efficient packet forwarding.

Forwarding Information Base

The Forwarding Information Base (FIB) is a table that contains a copy of the forwarding information in the IP routing table. When routing or topology changes occur in the network, the route processor updates the IP routing table and CEF updates the FIB. Because there is a one-to-one correlation between FIB entries and routing table entries, the FIB contains all known routes and eliminates the need for route cache maintenance that is associated with switching paths, such as fast switching and optimum switching. CEF uses the FIB to make IP destination-based switching decisions and maintain next-hop address information based on the information in the IP routing table.

On the Catalyst 4500 series switches, CEF loads the FIB in to the Integrated Switching Engine hardware to increase the performance of forwarding. The Integrated Switching Engine has a finite number of forwarding slots for storing routing information. If this limit is exceeded, CEF is automatically disabled and all packets are forwarded in software. In this situation, you should reduce the number of routes on the switch and then reenable hardware switching with the **ip cef** command.

Adjacency Tables

In addition to the FIB, CEF uses adjacency tables to prepend Layer 2 addressing information. Nodes in the network are termed *adjacent* if they are within a single hop from each other. The adjacency table maintains Layer 2 next-hop addresses for all FIB entries.

Adjacency Discovery

The adjacency table is populated as new adjacent nodes are discovered. Each time an adjacency entry is created (such as using the Address Resolution Protocol (ARP), a link-layer header for that adjacent node is stored in the adjacency table. Once a route is determined, the link-layer header points to a next hop and corresponding adjacency entry. The link-layer header is subsequently used for encapsulation during CEF switching of packets.

Adjacency Resolution

A route might have several paths to a destination prefix, such as when a router is configured for simultaneous load balancing and redundancy. For each resolved path, a pointer is added for the adjacency corresponding to the next-hop interface for that path. This method is used for load balancing across several paths.

Adjacency Types That Require Special Handling

In addition to adjacencies for next-hop interfaces (host-route adjacencies), other types of adjacencies are used to expedite switching when certain exception conditions exist. When the prefix is defined, prefixes requiring exception processing are cached with one of the special adjacencies listed in Table 35-1.

Table 35-1 Adjacency Types for Exception Processing

Adjacency Type	Processing Method
Null adjacency	Packets destined for a Null0 interface are dropped. A Null0 interface can be used as an effective form of access filtering.
Glean adjacency	When a router is connected directly to several hosts, the FIB table on the router maintains a prefix for the subnet rather than for each individual host. The subnet prefix points to a glean adjacency. When packets must be forwarded to a specific host, the adjacency database is gleaned for the specific prefix.
Punt adjacency	Features that require special handling or features that are not yet supported by CEF switching are sent (punted) to the next higher switching level.
Discard adjacency	Packets are discarded.
Drop adjacency	Packets are dropped.

Unresolved Adjacency

When a link-layer header is prepended to packets, FIB requires the prepend to point to an adjacency corresponding to the next hop. If an adjacency was created by FIB and was not discovered through a mechanism such as ARP, the Layer 2 addressing information is not known and the adjacency is considered incomplete. When the Layer 2 information is known, the packet is forwarded to the route processor, and the adjacency is determined through ARP.

Catalyst 4500 Series Switch Implementation of CEF

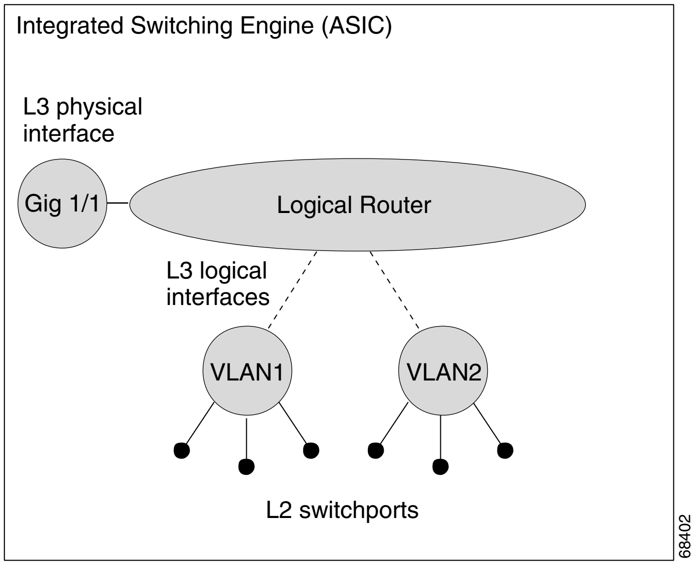
Catalyst 4500 series switch support an ASIC-based Integrated Switching Engine that provides these features:

- Ethernet bridging at Layer 2
- IP routing at Layer 3

Because the ASIC is specifically designed to forward packets, the Integrated Switching Engine hardware can run this process much faster than CPU subsystem software.

Figure 35-1 shows a high-level view of the ASIC-based Layer 2 and Layer 3 switching process on the Integrated Switching Engine.

Figure 35-1 Logical L2/L3 Switch Components



The Integrated Switching Engine performs inter-VLAN routing on logical Layer 3 interfaces with the ASIC hardware. The ASIC hardware also supports a physical Layer 3 interface that can be configured to connect with a host, a switch, or a router.

This section contains the following subsections:

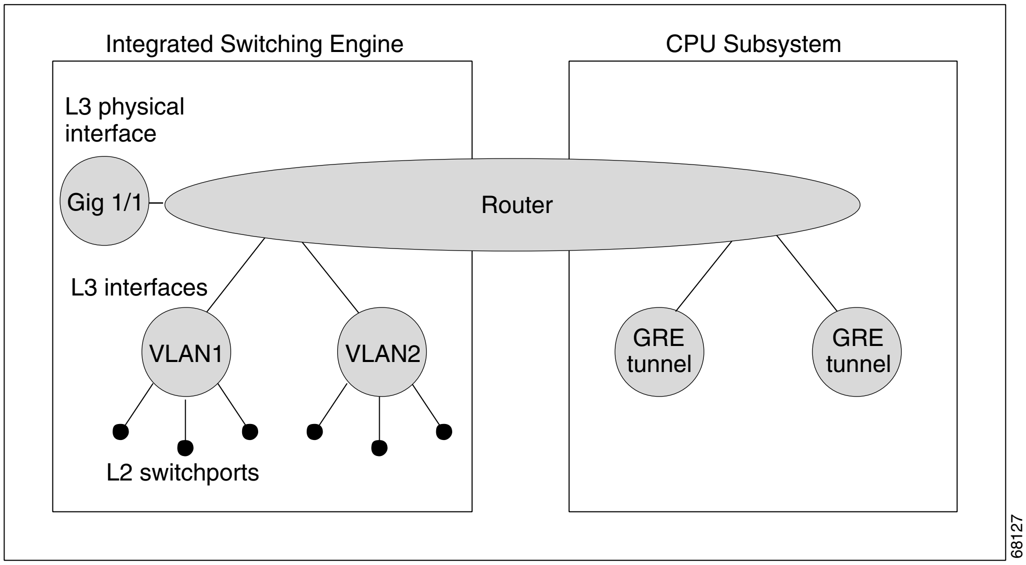
- Hardware and Software Switching, page 35-4
- Load Balancing, page 35-6
- Software Interfaces, page 35-6

Hardware and Software Switching

For the majority of packets, the Integrated Switching Engine performs the packet forwarding function in hardware. These packets are hardware-switched at very high rates. Exception packets are forwarded by the CPU subsystem software. Statistic reports should show that the Integrated Switching Engine is forwarding the vast majority of packets in hardware. Software forwarding is significantly slower than hardware forwarding, but packets forwarded by the CPU subsystem do not reduce hardware forwarding speed.

Figure 35-2 shows a logical view of the Integrated Switching Engine and the CPU subsystem switching components.

Figure 35-2 Hardware and Software Switching Components



The Integrated Switching Engine performs inter-VLAN routing in hardware. The CPU subsystem software supports Layer 3 interfaces to VLANs that use Subnetwork Access Protocol (SNAP) encapsulation. The CPU subsystem software also supports generic routing encapsulation (GRE) tunnels.

Hardware Switching

- Hardware switching is the normal operation for switches with supervisor engines.
- Beginning in Cisco IOS XE Release 3.7.1E, GRE tunnels are supported on the hardware on Catalyst 4500 series switches. When GRE is configured without tunnel options, packets are hardware-switched. Otherwise, packets are switched in the software.

Software Switching

Software switching occurs when traffic cannot be processed in hardware. The following types of exception packets are processed in software at a much slower rate:

- Packets that use IP header options

Note Packets that use TCP header options are switched in hardware because they do not affect the forwarding decision.

- Packets that have an expiring IP time-to-live (TTL) counter
- Packets that are forwarded to a tunnel interface.

Note When GRE tunnels are configured without tunnel options, packets are hardware-switched.

- Packets that arrive with non-supported encapsulation types
- Packets that are routed to an interface with non-supported encapsulation types

- Packets that exceed the MTU of an output interface and must be fragmented
- Packets that require an IGMP redirect for routing
- 802.3 Ethernet packets

Load Balancing

The Catalyst 4500 series switch supports load balancing for routing packets in the Integrated Switching Engine hardware. Load balancing is always enabled. It works when multiple routes for the same network with different next-hop addresses are configured. These routes can be configured either statically or through a routing protocol such as OSPF or EIGRP.

The hardware makes a forwarding decision by using a hardware load sharing hash function to compute a value, based on the source and destination IP addresses and the source and destination TCP port numbers (if available). This load sharing hash value is then used to select which route to use to forward the packet. All hardware switching within a particular flow (such as a TCP connection) is routed to the same next hop, which reduces the chance that packet reordering occurs. Up to eight different routes for a particular network are supported.

Software Interfaces

Cisco IOS for the Catalyst 4500 series switch supports GRE and IP tunnel interfaces that are not part of the hardware forwarding engine. All packets that flow to or from these interfaces must be processed in software and have a significantly lower forwarding rate than that of hardware-switched interfaces. Also, Layer 2 features are not supported on these interfaces.

CEF Configuration Restrictions

The CEF Integrated Switching Engine supports only ARPA and ISL/802.1q encapsulation types for Layer 3 switching in hardware. The CPU subsystem supports a number of encapsulations such as SNAP for Layer 2 switching that you can use for Layer 3 switching in software.

Configuring CEF

These sections describe how to configure CEF:

- Enabling CEF, page 35-6
- Configuring Load Balancing for CEF, page 35-7

Enabling CEF

By default, CEF is enabled globally on the Catalyst 4500 series switch. No configuration is required.

To reenable CEF, perform this task:

Command	Purpose
`Switch(config)#` **`ip cef distributed`**	Enables standard CEF operation.

Configuring Load Balancing for CEF

CEF load balancing is based on a combination of source and destination packet information; it allows you to optimize resources by distributing traffic over multiple paths for transferring data to a destination. You can configure load balancing on a per-destination basis. Load-balancing decisions are made on the outbound interface. You can configure per-destination load balancing for CEF on outbound interfaces.

The following topics are discussed:

- Configuring Per-Destination Load Balancing, page 35-7
- Configuring Load Sharing Hash Function, page 35-7
- Viewing CEF Information, page 35-8

Configuring Per-Destination Load Balancing

Per-destination load balancing is enabled by default when you enable CEF. To use per-destination load balancing, you do not perform any additional tasks once you enable CEF.

Per-destination load balancing allows the router to use multiple paths to achieve load sharing. Packets for a given source-destination host pair are guaranteed to take the same path, even if multiple paths are available. Traffic destined for different pairs tend to take different paths. Per-destination load balancing is enabled by default when you enable CEF; it is the load balancing method of choice in most situations.

Because per-destination load balancing depends on the statistical distribution of traffic, load sharing becomes more effective as the number of source-destination pairs increases.

Use per-destination load balancing to ensure that packets for a given host pair arrive in order. All packets for a certain host pair are routed over the same link or links.

Configuring Load Sharing Hash Function

When multiple unicast routes exist to a particular destination IP prefix, the hardware sends packets matching that prefix across all possible routes, which shares the load across all next hop routers. By default, the route used is chosen by computing a hash of the source and destination IP addresses and using the resulting value to select the route. This preserves packet ordering for packets within a flow by ensuring that all packets within a single IP source/destination flow are sent on the same route, but it provides a near-random distribution of flows to routes.

You can change the load-sharing hash function. So, in addition to the source and destination IP addresses, the source TCP/UDP port, the destination TCP/UDP port, or both can also be included in the hash.

To the configure load sharing hash function to use the source and/or destination ports, perform this task:

Command	Purpose
`Switch (config)# [no] ip cef load-sharing algorithm include-ports source destination]`	Enables load sharing hash function to use source and destination ports. Use the **no** keyword to set the switch to use the default Cisco IOS load-sharing algorithm.

Note The **include-ports** option does not apply to software-switched traffic on the Catalyst 4500 series switches.

Viewing CEF Information

You can view the collected CEF information. To view CEF information, perform this task:

Command	Purpose
`Switch# show ip cef`	Displays the collected CEF information.

Monitoring and Maintaining CEF

To display information about IP traffic, perform this task:

Command	Purpose
`Switch# show interface type slot/interface | begin L3`	Displays a summary of IP unicast traffic.

This example shows how to display information about IP unicast traffic on interface Fast Ethernet 3/3:

```
Switch# show interface fastethernet 3/3 | begin L3
  L3 in Switched: ucast: 0 pkt, 0 bytes - mcast: 12 pkt, 778 bytes mcast
  L3 out Switched: ucast: 0 pkt, 0 bytes - mcast: 0 pkt, 0 bytes
     4046399 packets input, 349370039 bytes, 0 no buffer
     Received 3795255 broadcasts, 2 runts, 0 giants, 0 throttles
<...output truncated...>
Switch#
```

Note The IP unicast packet count is updated approximately every five seconds.

Displaying IP Statistics

IP unicast statistics are gathered on a per-interface basis. To display IP statistics, perform this task:

Command	Purpose
`Switch#` **`show interface`** *`type number`* **`counters detail`**	Displays IP statistics.

This example shows how to display IP unicast statistics for fastethernet 3/1:

```
Switch# show interface fastethernet 3/1 counters detail

Port              InBytes       InUcastPkts       InMcastPkts       InBcastPkts
Fa3/1          7263539133           5998222           6412307               156

Port             OutBytes      OutUcastPkts      OutMcastPkts      OutBcastPkts
Fa3/1          7560137031           5079852          12140475                38

Port            InPkts 64        OutPkts 64    InPkts 65-127    OutPkts 65-127
Fa3/1               11274            168536           7650482          12395769

Port        InPkts 128-255   OutPkts 128-255   InPkts 256-511   OutPkts 256-511
Fa3/1               31191             55269             26923             65017

Port       InPkts 512-1023  OutPkts 512-1023
Fa3/1              133807            151582

Port      InPkts 1024-1518 OutPkts 1024-1518 InPkts 1519-1548 OutPkts 1519-1548
Fa3/1                  N/A               N/A              N/A               N/A

Port      InPkts 1024-1522 OutPkts 1024-1522 InPkts 1523-1548 OutPkts 1523-1548
Fa3/1              4557008           4384192                0                 0

Port      Tx-Bytes-Queue-1  Tx-Bytes-Queue-2 Tx-Bytes-Queue-3  Tx-Bytes-Queue-4
Fa3/1                   64                 0            91007        7666686162

Port      Tx-Drops-Queue-1  Tx-Drops-Queue-2 Tx-Drops-Queue-3  Tx-Drops-Queue-4
Fa3/1                    0                 0                0                 0

Port        Rx-No-Pkt-Buff     RxPauseFrames    TxPauseFrames   PauseFramesDrop
Fa3/1                    0                 0                0               N/A

Port      UnsupOpcodePause
Fa3/1                    0
Switch#
```

To display CEF (software switched) and hardware IP unicast adjacency table information, perform this task:

Command	Purpose
Switch# **show adjacency** [*interface*] [**detail** | **internal** | **summary**]	Displays detailed adjacency information, including Layer 2 information, when the optional **detail** keyword is used.

This example shows how to display adjacency statistics:

```
Switch# show adjacency gigabitethernet 3/5 detail
Protocol Interface                 Address
IP       GigabitEthernet9/5        172.20.53.206(11)
                                   504 packets, 6110 bytes
                                   00605C865B82
                                   000164F83FA50800
                                   ARP        03:49:31
```

Note Adjacency statistics are updated approximately every 10 seconds.