This chapter covers the following topics:

- Multilayer Switching Architecture on Catalyst Switches
- MLS Memory Table Architecture
- CEF-Based MLS Configuration, Verification, and Troubleshooting

Understanding and Configuring Multilayer Switching

The purpose of this chapter is to provide you with details, architecture, and methods of multilayer switching on Catalyst switches. An understanding of multilayer switching is necessary for network designers, administrators, and operators for deployment and troubleshooting purposes.

The term *multilayer switching* refers to the ability of a Catalyst switch to support switching and routing of packets in hardware, with optional support for Layers 4 through 7 switching in hardware as well. As mentioned in Chapter 1, "Introduction to Building Cisco Multilayer Switched Networks," switching and routing in hardware (hardware switching) yields high throughput at or near line rate even with all ports sending traffic simultaneously.

With multilayer switches, the prime area of focus is often raw performance. Multilayer switches tend to have packet-switching throughputs in the millions of packets per second (pps), while traditional general-purpose routers have evolved from the 100,000 pps range to just over a million pps. Cisco Catalyst switches achieve this high rate of performance by using hardware switching.

In order for Catalyst switches to perform hardware switching, a route processor (Layer 3 engine) must download software-based routing, switching, access lists, QoS, and other information to the hardware for packet processing. To accomplish multilayer switching (packet processing in hardware), Cisco Catalyst switches use either the traditional multilayer switching (traditional MLS) or the Cisco Express Forwarding (CEF)-based MLS architecture. Traditional MLS is a legacy feature, whereas all leading-edge Catalyst switches support CEF-based multilayer switching (CEF-based MLS). Table 9-1 illustrates Catalyst switch support of traditional MLS and CEF-based MLS.

In terms of CCNP BCMSN exam preparation, you should focus on the following:

- The path of a packet with CEF-based MLS
- Understanding the differences between centralized and distributed switching
- CEF troubleshooting commands

The section dealing with traditional MLS is outside the scope of the current CCNP BCMSN exam.

Table 9-1   *Layer 3 Catalyst Switch Support of Traditional MLS and CEF-Based MLS*

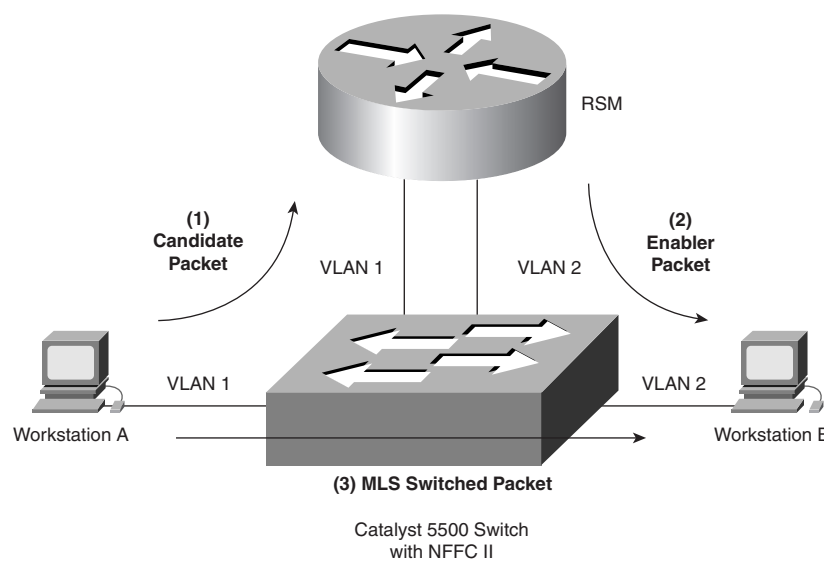| Catalyst Switch Family	MLS Mode of Operation
Catalyst 3550, 3560, and 3750	CEF-based MLS
Catalyst 4000 or 4500 running Cisco IOS and using Supervisor Engine II+, III, IV, or V	CEF-based MLS
Catalyst 5000 or 5500 with RSM or RSFC and NFFC or NFFC II	Traditional MLS
Catalyst 6500 with Supervisor Engine I with MSFC	Traditional MLS
Catalyst 6500 with Supervisor Engine II, Supervisor 32, or Supervisor 720 with MSFC	CEF-based MLS

Understanding Traditional MLS

MLS enables specialized application-specific integrated circuits (ASIC) to perform Layer 2 rewrite operations of routed packets. Layer 2 rewrites include rewriting the source and destination MAC addresses and writing a recalculated cyclic redundancy check (CRC). Because the source and destination MAC addresses change during Layer 3 rewrites, the switch must recalculate the CRC for these new MAC addresses.

For Catalyst switches that support traditional MLS, the switch learns Layer 2 rewrite information from an MLS router via an MLS protocol. Another name for traditional MLS is NetFlow-based switching. With traditional MLS, the Layer 3 engine (route processor) and switching ASICs work together to build Layer 3 entries on the switch. Each of these Layer 3 entries is populated in one of the following ways:

- Source IP address exclusively (S)
- Source and destination IP address (S/D)
- Full Flow Information with Layer 4 protocol information (FFI)

With traditional MLS, the switch forwards the first packet in any flow to the Layer 3 engine for processing using software switching. After the routing of the first packet in the flow, the Layer 3 engine programs the hardware-switching components for routing for subsequent packets. Figure 9-1 illustrates these fundamentals of traditional MLS.

In Figure 9-1, when workstation A sends a packet to workstation B, workstation A sends the packet to its default gateway. In Figure 9-1, the default gateway is the RSM. The switch (MLS-SE) recognizes this packet as an MLS candidate packet because the destination MAC address matches the MAC address of the MLS router (MLS-RP). As a result, the switch creates a candidate entry for this flow. Next, the router accepts the packets from workstation A, rewrites the Layer 2 destination MAC address and CRC, and forwards the packet to workstation B. The switch refers to the routed packet from the RSM as the enabler packet. The switch, upon seeing both the candidate and enabler packets, creates an MLS entry in hardware such that the switch rewrites and forwards all future packets matching this flow. The MLS switched packet arrow in Figure 9-1 indicates this flow. This behavior is very different from CEF-based MLS.

Figure 9-1   *Fundamentals of Traditional MLS*

Catalyst 5500 Switch
with NFFC II

On the Catalyst 5000 family of switches, MLS requires specialized supervisors and line modules and a router. In addition, MLS on Catalyst 5000 requires manual configuration. On the Catalyst 6500 family of switches with a Supervisor I Engine with an MSFC, all the line modules support MLS. In addition, this hardware combination uses MLS by default and configuration of MLS is not required. For more details on and examples of the MLS architecture, consult the following technical documents at Cisco.com.

"Configuring and Troubleshooting IP MLS on Catalyst 6000 Switches with an MSFC," Document ID: 10566

"Troubleshooting IP Multilayer Switching," Document ID: 10554

Traditional MLS is a legacy feature; all leading-edge and future Catalyst switches support CEF-based MLS. The next section and the remainder of this chapter discuss CEF-based MLS and its architecture.

Understanding CEF-Based MLS

The leading-edge Catalyst switches, as listed in Table 9-1, use the CEF-based MLS forwarding model to download the control-plane information such as the access lists to the data-plane on the supervisor, port, or line card for hardware switching of packets. In the context of this chapter, the *control-plane* represents the Layer 3 engine (route processor), and the *data-plane* represents the hardware components such as ASICs used by the switch for hardware switching.

CEF is a topology-based forwarding model in which all routing information is pre-populated into a forwarding information base (FIB) and Layer 2 rewrite information is dynamically updated in an adjacency table. As a result of the pre-population of routing
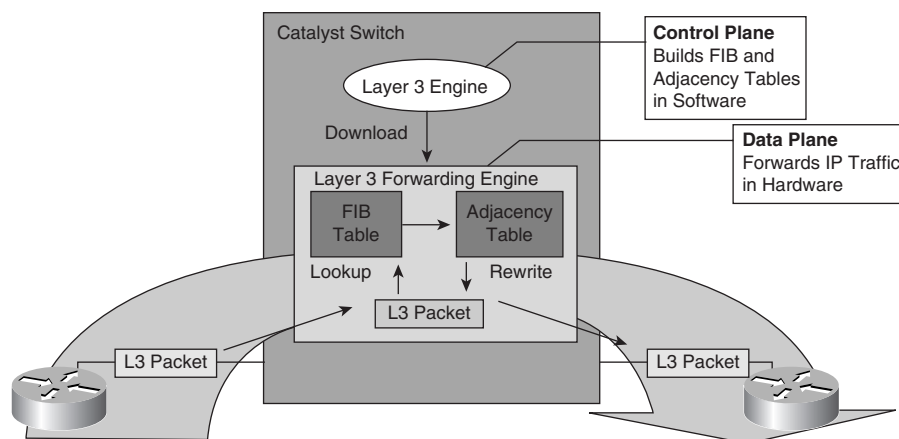
information, Catalyst switches can quickly look up routing information such as IP adjacencies and next-hop IP and MAC addresses.

The two main components of CEF are as follows:

- **Forwarding information base (FIB)**— CEF uses an FIB to make IP destination prefix-based switching decisions. The FIB is conceptually similar to a routing table or information base. It maintains a mirror image of the forwarding information contained in the IP routing table. When routing or topology changes occur in the network, the IP routing table is updated, and those changes are reflected in the FIB. The FIB maintains next-hop address information based on the information in the IP routing table. In the context of CEF-based MLS, both the Layer 3 engine and the hardware-switching components maintain an FIB.

- **Adjacency tables**—Network nodes in the network are said to be adjacent if they can reach each other with a single hop across a link layer. In addition to the FIB, CEF uses adjacency tables to store Layer 2 addressing information. The adjacency table maintains Layer 2 addresses for all FIB entries. As with the FIB, in the context of CEF-based MLS, both the Layer 3 engine and the hardware-switching components maintain an adjacency table.

Figure 9-2 depicts CEF-based MLS logically, applied to the Catalyst 6500 family of switches.

Figure 9-2 *Logical Depiction of CEF-Based MLS on the Catalyst 6500 Family of Switches*

CEF-based MLS separates the control-plane hardware from the data-plane hardware switching. Nevertheless, the control-plane is responsible for building the FIB table and adjacency tables in software and then downloading this information to the data-plane.

On the Catalyst 6500 family of switches, the control-plane hardware is easily distinguishable from the data-plane hardware, especially in the case of hybrid software. The Catalyst 6500 MSFC daughter card is responsible for the control-plane operations, while the Supervisor and PFC module are responsible for the data-plane operations. The Catalyst 3550, 3560, 3750, and 4500 families of switches do not have distinguishable control-plane and data-plane modules or daughter cards.

Catalyst switches do not support routing of all types of frames in hardware. For example, the following list details common frame types that are not supported by hardware switching:

- Packets with IP header options
- Packets sourced from or destined to tunnel interfaces
- Packets using Ethernet encapsulation types other than ARPA
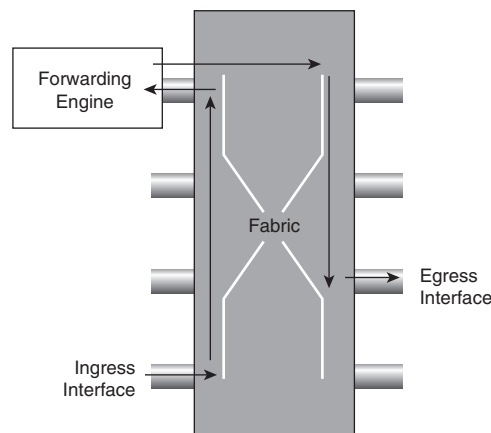- Packets that require fragmentation

Upcoming models of Catalyst switches may support these frame types in hardware. In addition, each Catalyst switch family has its own distinct list of frames not supported by hardware switching.

Centralized and Distributed Switching

CEF-based Catalyst switches support one of two methods of hardware switching at Layer 3:

- **Centralized switching**—Centralized switching carries out forwarding decisions on a specialized ASIC that is *central* to all interfaces of a Layer 3 switch. With centralized switching, routing, ACL, QoS, and forwarding decisions are made on the Supervisor Engine in a modular chassis or by Layer 3 engines in fixed port density Layer 3 switches. As a result, all frames to be routed or switched must pass through the centralized engine via a fabric or bus. Furthermore, with centralized switching, the hardware-switching performance of the Catalyst switch is based on the central forwarding engine and the fabric or bus architecture. Figure 9-3 illustrates centralized switching, logically. Note in Figure 9-3 how frames must pass through the centralized switching engine.
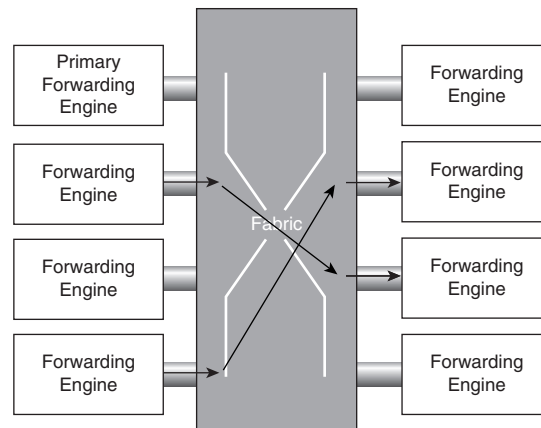
Figure 9-3   *Logical Representation of Centralized Switching*

Examples of Catalyst switches that are engineered for centralized switching are the Catalyst 4500 family of switches and the Catalyst 6500 family of switches without the use of Distributed Forwarding Cards (DFC).

- **Distributed switching**—With distributed switching, interfaces or line modules on Layer 3 switches handle forwarding decisions independently. With distributed switching, a centralized switching engine synchronizes Layer 3 forwarding, routing, and rewrite tables to local tables on distributed switching–capable modules. As a result, individual line cards or ports make forwarding decisions without the aid of the centralized switching engine; frames pass between ports directly across the fabric. In other words, switches using distributed switching place additional copies of the CEF FIB and adjacency table on line modules or interfaces for routing and switching of frames. System performance with distributed switching is equal to the aggregate of all forwarding engines. Distributed forwarding enables Catalyst switches to achieve rates of over 100 million pps. The Catalyst 6500 supports distributed switching through the use of the Switch Fabric module or with a Supervisor 720 that has an integrated fabric and DFC line modules. The Catalyst 6500 maintains use of a centralized distributing switching engine even when using distributed switching–capable line modules for backward-compatibility. Figure 9-4 illustrates distributed switching logically.

Figure 9-4 *Logical Representation of Distributed Switching*

| NOTE | Traditional MLS supports centralized switching only. |

Address Resolution Protocol Throttling

An important feature of CEF-based Catalyst switches is Address Resolution Protocol (ARP) throttling; this feature requires explanation before CEF-based MLS is explored further. Make note that the ARP table builds the CEF adjacency table. This concept is explored in more detail throughput this chapter; however, it is important to consider when reading this section.
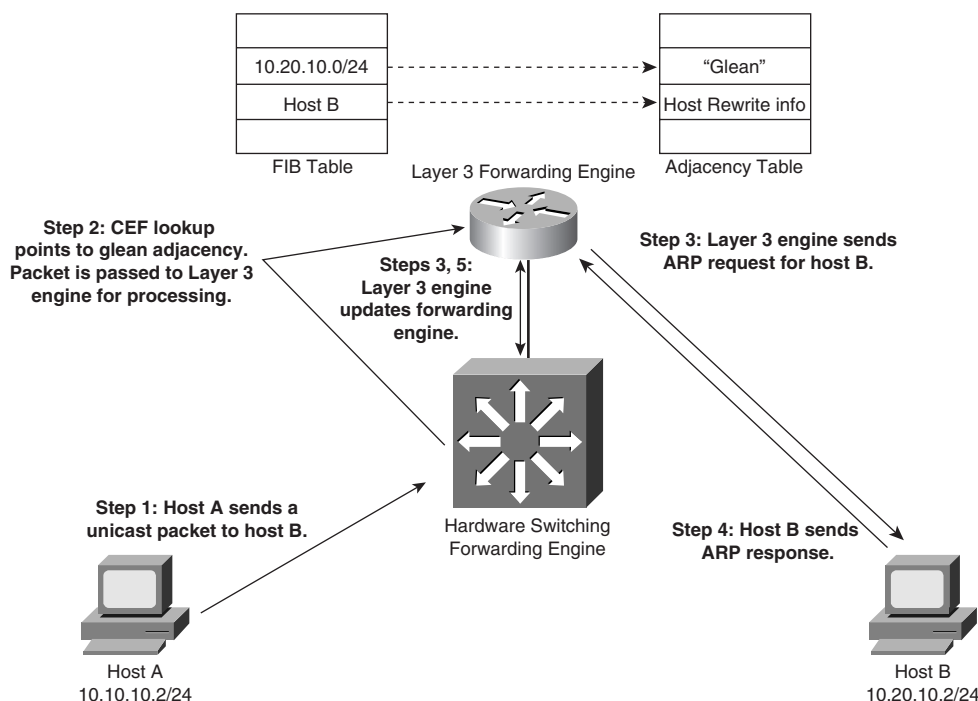
When a router is directly connected to a segment shared by multiple hosts such as Ethernet interfaces, the router maintains an additional prefix for the subnet. This subnet prefix points

to a glean adjacency. When a router receives packets that need to be forwarded to a specific host, the adjacency database is gleaned for the specific prefix. If the prefix does not exist, the subnet prefix is consulted, and the glean adjacency indicates that any addresses within this range should be forwarded to the Layer 3 engine ARP processing.

One example where glean adjacencies are used is where a Catalyst switch receives a packet for which no rewrite information exists. In order to obtain rewrite information, the Layer 3 engine sends ARP requests to obtain the rewrite information. Catalyst switches using CEF-based MLS forward only the first several packets to the Layer 3 engine for new destinations without any rewrite information. The switch installs a throttling adjacency such that the switch drops subsequent packets to the specific destination address in hardware until an ARP response is received. The switch removes the throttling adjacency when an ARP reply is received from the Layer 3 engine (and a complete rewrite adjacency is installed for the host). The switch removes the throttling adjacency if no ARP reply is seen within 2 seconds (to allow more packets through to the Layer 3 engine to reinitiate ARP). This relieves the Layer 3 engine from excessive ARP processing (or ARP-based Denial-of-Service [DoS] attacks).

Figure 9-5 shows an example of ARP throttling; an explanation of its stepwise behavior follows. Figure 9-5 depicts the Layer 3 forwarding engine and hardware switching forwarding engine as two separate hardware components for illustrative purposes.

Figure 9-5 *ARP Throttling Example*

ARP throttling consists of the following steps:

Step 1 Host A sends a packet to host B.

Step 2 The switch forwards the packet to the Layer 3 engine based on the "glean" entry in the FIB.

Step 3 The Layer 3 engine sends an ARP request for host B and installs the drop adjacency for host B in hardware.

Step 4 Host B responds to the ARP request.

Step 5 The Layer 3 engine installs adjacency for host B and removes the drop adjacency.

Figure 9-9, later in this chapter, builds on Figure 9-5 to show CEF-based MLS in a larger context.

Switching Table Architectures

Multilayer switches build routing (CEF FIB and adjacency), bridging, QoS, and ACL tables for centralized or distributed switching in hardware using high-speed memory tables. Switches perform lookups in these tables for result information, such as to determine whether a packet with a specific destination IP address is supposed to be dropped according to an ACL. These tables support high-performance lookups and search algorithms such that multilayer switches maintain line-rate performance.

Multilayer switches deploy these memory tables using specialized memory architectures, referred to as content addressable memory (CAM) and ternary content addressable memory (TCAM). CAM tables provide only two results: 0 (true) or 1 (false). CAM is most useful for building tables that search on exact matches such as MAC address tables. TCAM provides three results: 0, 1, and "don't care." TCAM is most useful for building tables for searching on longest matches such as IP routing tables organized by IP prefixes.
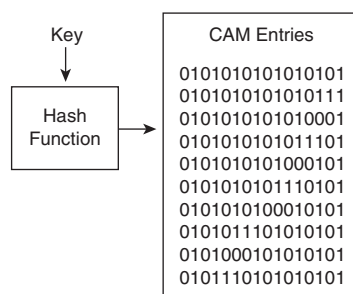
In addition, Catalyst switch architecture supports the ability to perform multiple lookups into multiple distinct CAM and TCAM regions in parallel. As a result of this ability to perform multiple lookups simultaneously, Catalyst switches do not suffer any performance degradation by enabling additional hardware-switching features such as QoS and IP ACL processing.

CAM

Catalyst switches use CAM tables to house, for example, Layer 2 switching tables. Switches match results in CAM tables in binary (0 or 1 operations). With CAM tables, switches must find exact matches or the switches use a default behavior. For example, in the case of Layer 2 switching tables, the switch must find an exact match to a destination MAC address or the switch floods the packet out all ports in the VLAN.

The information a switch uses to perform a lookup in a CAM table is called a key. For example, a Layer 2 lookup would use a destination MAC address and a VLAN ID as a key. Figure 9-6 depicts the use of keys in determining results using CAM.

Figure 9-6   *Content Addressable Memory*

The following steps detail the process of determining a result based on a key:

Step 1   The switch feeds the key into a hashing algorithm for searching the CAM for a matching key.

Step 2   The hashing algorithm returns a pointer into the CAM table for the matched entry.

Step 3   The switch accesses the pointer and finds the result without searching the entire table sequentially.

In the case of Layer 2 tables, CAM tables contain information such as destination VLAN, destination MAC address, and destination ports.

Switches do not always search for an exact match in memory tables. For example, a switch that is looking up an IP route destination subnet with a 24-bit mask is only concerned with the first 24 bits of an IP address (i.e., the prefix). In this case, the switch is not looking for an exact match in CAM but rather a match on the first 24 bits of the IP address.

TCAM

TCAM is a specialized CAM designed for rapid table lookups. For example, the Catalyst 2950, 3550, 3560, 3750, 4500, and 6500 families of switches use TCAM to handle ACL lookups at line rate. As a result of using TCAM, applying ACLs does not affect the performance of the switch.

TCAM populates a limited number of entries, which varies per platform, with pattern values and mask values, each with an associated result. These are known as value, mask, and result (VMR) entries, respectively. The term VMR simply refers to the format of entries in TCAM.

The "value" in VMR refers to the pattern that is to be matched; examples include IP addresses, protocol ports, DSCP values, and so on. The "mask" refers to the mask bits associated with the pattern and determines the prefix. The "result" refers to the result or

action that occurs in the case where a lookup returns a hit for the pattern and mask. This result might be a "permit" or "deny" in the case of a TCAM for ACLs. Another example of a result is a pointer to an entry in the hardware adjacency table that contains the next-hop MAC rewrite information in the case of a TCAM used for IP routing.

The TCAM structure is broken into a series of patterns and masks. Masks are shared among a specific number of patterns by using wildcard-specific content fields. To perform a lookup in a TCAM table, the switch checks all entries in parallel. The performance is independent of the number of entries. This allows a switch to use the longest match lookup when needed, and it provides fixed latency to unused fields.

Figure 9-7 illustrates a sample logical depiction of TCAM used for access lists in hardware. This figure represents the access list shown in Example 9-1.

Figure 9-7   *Sample TCAM Illustration*

Mask	Patterns
Mask 1 Match: All 32 Bits of Source IP Address	Src IP = 10.1.1.1
	Empty 2
	Empty 3
	Empty 4
	Empty 5
Don't Care: All Remaining Bits	Empty 6
	Empty 7
	Empty 8
Mask 2 Match: Most Significant 24 Bits of Source IP Addr	Src IP = 10.1.1.0
	Empty 2
	Empty 3
	Empty 4
	Empty 5
Don't Care: All Remaining Bits	Empty 6
	Empty 7
	Empty 8

Example 9-1   *Sample ACL for Figure 9-7*

```
access-list 101 permit ip host 10.1.1.1 any

access-list 101 deny ip 10.1.1.0 0.0.0.255 any log
```

Moreover, TCAM defines three different match options that correlate to specific match regions. These match regions are as follows:

- **Exact-match region**—Consists of Layer 3 entries for regions such as IP adjacencies. IP adjacencies are the next-hop information (MAC address) for an IP address. Other examples of exact-match regions are Layer 2 switching tables and UDP flooding tables.

- **Longest-match region**—Consists of multiple "buckets" or groups of Layer 3 address entries organized in decreasing order by mask length. All entries within a bucket share the same mask value and key size. The buckets change their size dynamically by borrowing address entries from neighboring buckets. Although the size of the whole protocol region is fixed, several platforms support configuration of the region size. For most platforms, the reconfigured size of the protocol region is effective only after the next system reboot.

- **First-match region**—Consists of regions that stop lookups after the first match of the entry. An example of when a first-match region is used is for ACL entries.

Table 9-2 illustrates the common protocol regions, lookup type, and key size found on Catalyst switches. The size of the regions and the ability to configure the region varies on each Catalyst switch family.

Table 9-2   *Common TCAM Protocol Regions*

Region Name	Cisco IOS Region Name	Lookup Type	Key Size	Sample Result
IP adjacency	ip-adjacency	Exact-match	32 bits	MAC address rewrite info
IP prefix	ip-prefix	Longest-match	32 bits	Next-hop routing information
IP multicast	ip-mcast	Longest-match	64 bits	Next-hop routing information
Layer 2 switching	l2-switching	Exact-match	64 bits	Destination interface and VLAN
UDP flooding	udp-flooding	Exact-match	64 bits	Next-hop routing or MAC address rewrite info
Access lists	access-list	First-match	128 bits	Permit, deny, or wildcard

CEF-Based MLS Operation and Use of TCAM

The following list details the characteristics of CEF-based MLS operation and its use of the TCAM:

- Longest-match lookups in the FIB table are done for the Layer 3 destination address prefixes.

- CEF uses the IP routing table on the Layer 3 forwarding engine to build the FIB. Arrangement of the FIB is for maximum lookup throughput.

- CEF builds the adjacency table from the ARP table. The adjacency table contains Layer 2 rewrite (MAC) information for the next hop.

- FIB entries in the TCAM table are populated from the most specific to the least specific entry.

- Adjacency (rewrite) information and statistics are maintained by specialized components.

- CEF maintains one-to-one CEF-to-adjacency mappings for accurate statistics tracking.

- When the FIB table in TCAM is full, a wildcard entry redirects unmatched entries to the software switching Layer 3 forwarding engine.

- When the adjacency table in TCAM is full, an entry in the FIB table points to the Layer 3 forwarding engine to redirect the adjacency lookup.

- FIB and adjacency tables are dynamically updated when an ARP entry for a destination next hop changes, ages out, or is removed; the routing table changes; or next-hop rewrite information changes.

Sample CEF-Based MLS Operation

Figure 9-8 provides an example of CEF-based MLS operation.

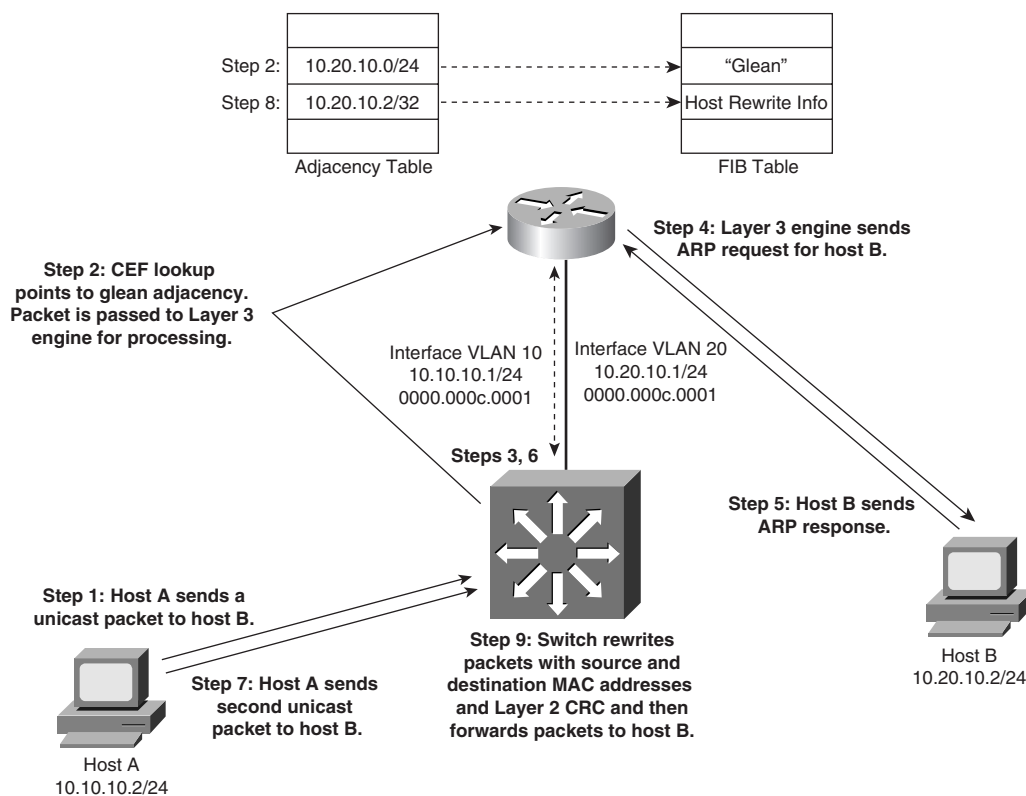**Figure 9-8** *Sample of CEF-Based MLS Operation*

1. Layer 3 packets initiate TCAM lookup.
2. The longest match returns adjacency with rewrite information.
3. The packet is rewritten per adjacency information and forwarded.

Before a multilayer switch can route frames in hardware, the switch sets up the necessary routing information in hardware. After the switch has set up the necessary routing information in the hardware, frame routing in hardware can start. The following steps illustrate an example of frame routing via hardware switching based on Figure 9-8. Figure 9-9 illustrates the steps. These steps assume the switch does not initially have rewrite information for the destination:

Step 1 Host A sends a packet to host B. The switch recognizes the frame as a Layer 3 packet because the destination MAC (0000.000c.0001) matches the Layer 3 engine MAC.

Step 2 The switch performs a CEF lookup based on the destination IP address (10.20.10.2) in hardware. The packet hits the CEF FIB entry for the connected (VLAN 20) network and is redirected to the Layer 3 engine using a "glean" adjacency. The hardware-switching CEF table cannot forward this packet because it does not have any rewrite information.

Figure 9-9   *Stepwise Example of CEF-Based MLS Operation*

Step 3   The Layer 3 engine installs an ARP throttling adjacency in the switch for host B's IP address, because an ARP entry does not exist for host B since host A and B have not previously communicated.

Step 4   The Layer 3 engine sends an ARP request for host B on VLAN 20.

Step 5   Host B sends an ARP response to the Layer 3 engine.

Step 6   The Layer 3 engine installs the resolved adjacency in its local adjacency table and the hardware-switching components install the adjacency as well (removing the ARP throttling adjacency).

Step 7   The switch receives another packet for host B (10.20.10.2).

Step 8   The switch performs a Layer 3 lookup and finds a CEF entry for host B. The entry points to the adjacency with rewrite information for host B.

Step 9   The switch rewrites packets per the adjacency information (source and destination MAC address) and forwards the packet to host B on VLAN 20.

NOTE   The ARP throttling adjacency drops the first packet in any communication between hosts A
and B due to ARP processing when no ARP entries exist for hosts A and B.

CEF-Based MLS Load Sharing

CEF-based MLS does support load sharing (equal-cost or nonequal-cost). However, CEF-based MLS does not support all the load-sharing features found in software-based CEF. With the current version of software on a Catalyst 6500 switch, a single FIB entry may have up to six adjacencies for load sharing per destination.

To achieve evenly distributed load balancing across multiple interfaces, CEF-based MLS selects a particular adjacency based on the hash (mathematical equivalent) of the following packet characteristics:

- Source IP address
- Destination IP address
- Source and destination IP Layer 4 ports

The load-sharing method and hashing algorithms vary slightly per Catalyst switch family. Consult the product documentation for specifics about load-sharing support on each Catalyst switch.

CEF-Based MLS Configuration, Verification, and Troubleshooting

In terms of designing multilayer switched networks, different modes of multilayer switching meet the requirements for different performance and scalability aspects. In terms of troubleshooting, understanding the basic operation of multilayer switches is paramount; multilayer switching requires a hierarchical approach to troubleshooting because several switching components determine the packet flow, not a single processing engine. CEF-based MLS configuration, verification, and troubleshooting vary slightly per Catalyst switch family but do have some commonalities.

CEF-Based MLS Configuration

Cisco Catalyst switches that use the CEF-based MLS architecture use CEF by default. Consequently, Catalyst switches perform hardware Layer 3 switching by default as well. For Catalyst switches that support CEF-based MLS, CEF-based MLS and per-destination load balancing with CEF-based MLS are enabled by default. As a result, no configuration is required for CEF-based MLS.

You should not disable CEF on Catalyst switches for any reason except under the supervision of a Cisco TAC engineer for the specific purpose of troubleshooting. Disabling

CEF on Cisco Catalyst switches yields low switching performance and may result in undesirable behavior.

CEF-Based MLS Verification and Troubleshooting

To display information about the Catalyst 6500 family of switches running Cisco IOS, use the **show interface** command as shown in Example 9-2.

Example 9-2 *Displaying Layer 3 Switching Statistics on the Cisco IOS–Based Catalyst 6500 Family of Switches*

```
Router#show interface port-channel 9

Port-channel9 is up, line protocol is up (connected)

  Hardware is EtherChannel, address is 00d0.039b.e80a (bia 00d0.039b.e800)

  Description: POINT-TO-POINT TO CORE-4

! Output omitted for brevity

  Output queue: 0/40 (size/max)

  5 minute input rate 0 bits/sec, 0 packets/sec

  5 minute output rate 0 bits/sec, 0 packets/sec

  L2 Switched: ucast: 205744 pkt, 34282823 bytes - mcast: 216245 pkt, 66357101 bytes

  L3 in Switched: ucast: 367825 pkt, 361204150 bytes - mcast: 0 pkt, 0 bytes mcast

  L3 out Switched: ucast: 248325 pkt, 243855150 bytes

      682964 packets input, 431530341 bytes, 0 no buffer

      Received 311465 broadcasts (50899 IP multicast)

      0 runts, 0 giants, 0 throttles

      0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored

      0 watchdog, 0 multicast, 0 pause input

      0 input packets with dribble condition detected

      554167 packets output, 309721969 bytes, 0 underruns

      0 output errors, 0 collisions, 8 interface resets

      0 babbles, 0 late collision, 0 deferred

      0 lost carrier, 0 no carrier, 0 PAUSE output

      0 output buffer failures, 0 output buffers swapped out
```

In Example 9-2, the lines beginning L2 Switched, L3 in Switched, and L3 out Switched indicate the Layer 3 hardware-switching statistics. Catalyst switches do not instantaneously update hardware-switching statistics for CLI **show** commands. For example, on the Catalyst 4500 family of switches running Cisco IOS, interface statistics may take up to 30 seconds to be updated in **show** commands. Furthermore, each Catalyst family of switches has its own troubleshooting methodology and commands for viewing hardware-switching statistics. Consult Cisco.com for more details.

Viewing the Layer 3 Engine CEF Table

The Layer 3 Engine CEF table determines the hardware-switching CEF table. As such, when troubleshooting any CEF issues, the first step is to view the software CEF table. To view the CEF FIB table as per the Layer 3 switching engine in software, use the following commands:

```
show ip cef
show ip cef detail
```

Both commands display the FIB table per the Layer 3 switching engine. These commands do not display the FIB table according to the hardware-switching components such as TCAM, but instead display a "shadow" copy in software of the hardware-switching table. The "receive" entry in the Next Hop column indicates that the destinations for that specific prefix are passed to a normal Cisco IOS routing engine for processing. Examples 9-3 and 9-4 illustrate the use of these commands, respectively, on a Catalyst 3550 switch.

Example 9-3 *Sample Use of the* **show ip cef** *Command*

```
Router#show ip cef

Prefix              Next Hop           Interface

0.0.0.0/32          receive

10.5.1.1/32         10.20.248.2        Port-channel1

10.5.1.140/30       10.20.248.202      GigabitEthernet3/16

                    10.20.248.2        Port-channel1

224.0.0.0/4         0.0.0.0

224.0.0.0/24        receive

255.255.255.255/32  receive
```

Example 9-4   *Sample Use of the* **show ip cef detail** *Command*

```
Switch#show ip cef detail

IP Distributed CEF with switching (Table Version 8548), flags=0x0

  4156 routes, 0 reresolve, 0 unresolved (0 old, 0 new), peak 1000

  4156 leaves, 56 nodes, 680744 bytes, 6513 inserts, 2357 invalidations

  1023 load sharing elements, 343728 bytes, 1023 references

  universal per-destination load sharing algorithm, id 0C2C7438

  3(0) CEF resets, 3099 revisions of existing leaves

  Resolution Timer: Exponential (currently 1s, peak 2s)

  2048 in-place/0 aborted modifications

  refcounts:  11471 leaf, 6319 node


  Table epoch: 0 (4156 entries at this epoch)


Adjacency Table has 8 adjacencies

0.0.0.0/32, version 0, epoch 0, receive

10.5.1.1/32, version 52, epoch 0, cached adjacency 10.20.248.2

0 packets, 0 bytes

  via 10.20.248.2, Port-channel1, 0 dependencies

    next hop 10.20.248.2, Port-channel1

    valid cached adjacency

! Output omitted for brevity

  via 10.20.248.2, Port-channel1, 0 dependencies

    traffic share 1

    next hop 10.20.248.2, Port-channel1

    valid adjacency

  0 packets, 0 bytes switched through the prefix
```

continues

Example 9-4   *Sample Use of the* **show ip cef detail** *Command (Continued)*

```
224.0.0.0/4, version 30, epoch 0

0 packets, 0 bytes, Precedence routine (0)

  via 0.0.0.0, 0 dependencies

    next hop 0.0.0.0

    valid punt adjacency

224.0.0.0/24, version 2, epoch 0, receive

255.255.255.255/32, version 1, epoch 0, receive
```

Viewing the Layer 3 Engine Adjacency Table

CEF populates the adjacency tables when MAC addresses are learned via the ARP process. As a result, the adjacency table includes the MAC address rewrite information and destination interface for the adjacent node. All IP routing entries in the CEF table correspond to a next-hop address (adjacency).

When a switch hardware-switches an ingress packet with the destination MAC address as itself, it looks up the destination IP address in the CEF table. The first match in the CEF table points to an adjacency entry that contains the MAC rewrite information and destination interface. The switch rewrites the packet accordingly and sends it out the destination interface.

A single CEF entry may point to multiple adjacency entries when multiple paths to a destination exist. In addition, when a router is connected directly to several hosts, the FIB table on the router maintains a prefix for the subnet rather than for the individual host prefixes. The subnet prefix points to a glean adjacency. When packets need to be forwarded to a specific host, the adjacency database is gleaned for the specific prefix.

Certain IP prefixes or addresses require exception processing. Switches require exception processing of packets in cases where hardware switching does not support routing of the frame or in cases where the Layer 3 engine requires processing of the packet. Examples of packets that require exception processing include interfaces configured for NAT and received packets with IP options. Recall from Chapters 1 and 2 that support for hardware switching of frames varies on a per-switch basis.

The following types of adjacencies exist for special processing:

- **Punt adjacency**—Used for packets that require special handling by the Layer 3 engine or for features that are not yet supported by hardware switching.

- **Drop or discard adjacency**—Used to drop ingress packets.

- **Null adjacency**—Used to drop packets destined for a Null0 interface. The use of a Null0 interface is for access filtering of specific source IP packets.

To display adjacency table information, use the **show adjacency** command. The optional **detail** keyword displays detailed adjacency information, including Layer 2 information. Adjacency statistics are updated approximately every 60 seconds. Examples 9-5 and 9-6 illustrate sample output from the **show adjacency** and **show adjacency detail** commands, respectively.

Example 9-5 *Sample Use of the* **show adjacency** *Command*

```
Router#show adjacency

Protocol Interface          Address

IP       GigabitEthernet3/16    10.20.248.202(11120)

IP       FastEthernet9/48       10.192.15.254(7)

IP       FastEthernet9/1        100.1.1.1(2005)

IP       EOBC0/0                127.0.0.12(3)

IP       FastEthernet9/48       10.192.15.100(5)

IP       Port-channel1          10.20.248.2(11162)
```

Example 9-6 *Sample Use of the* **show adjacency detail** *Command*

```
Router#show adjacency detail

Protocol Interface          Address

IP       GigabitEthernet3/16    10.20.248.202(11120)

                               0 packets, 0 bytes

                               00D00178780A

                               00D00624440A0800

                               ARP        00:05:40

                               Epoch: 0

IP       FastEthernet9/48       10.192.15.254(7)

                               0 packets, 0 bytes

                               00D0BCF107C8

                               00D00624440A0800

                               ARP        00:39:20

                               Epoch: 0

                                                              continues
```

Example 9-6 *Sample Use of the* **show adjacency detail** *Command (Continued)*

```
IP      FastEthernet9/1          100.1.1.1(2005)

                                 0 packets, 0 bytes

                                 0007ECC57000

                                 00D00624440A0800

                                 ARP        03:19:20

                                 Epoch: 0
IP      EOBC0/0                  127.0.0.12(3)

                                 0 packets, 0 bytes

                                 000021000000

                                 0000210000000800
```

Debugging CEF on the Layer 3 Engine

To debug CEF on the Layer 3 engine, use the **debug ip cef** EXEC command. These debug commands display information about the CEF from the perspective of the Layer 3 engine. The interprocess communication (IPC) debugs provide information on CEF data to and from centralized and distributed switching engines. The syntax for these **debug ip cef** commands is as follows:

```
debug ip cef {drops [access-list] | receive [access-list] | events [access-list] |
  prefix-ipc [access-list] | table [access-list]} debug ip cef {ipc | interface-ipc}
```

The optional constraints for the **debug ip cef** command are as follows. Use the options to debug specific CEF operations on the Layer 3 engine:

- **drops**—Records dropped packets.
- **receive**—Records packets that are not switched using information from the FIB table, but are received and sent to the next switching layer.
- **events**—Records general CEF events.
- **prefix-ipc**—Records updates related to IP prefix information, including the following:
  - Debugging of IP routing updates in a line card
  - Reloading of a line card with a new table
  - Adding a route update from the route processor to a DFC-enabled line card where the CEF table is already full
  - Control messages related to FIB table prefixes

- **table**—Produces a table showing events related to the FIB table. Possible types of events include the following:

  — Routing updates that populate the FIB table

  — Flushing of the FIB table

  — Adding or removing of entries to or from the FIB table

  — Table reloading process

- **ipc**—Records information related to interprocess communications in CEF. Possible types of events include the following:

  — Transmission status of IPC messages

  — Status of buffer space for IPC messages

  — IPC messages received out of sequence

  — Status of resequenced messages

  — Throttle requests sent from a line card to the route processor

- **interface-ipc**—Records IPC updates related to interfaces. Possible reporting includes an interface coming up or going down, and so on.

Specify an access list with the **debug** option to control collections of debugging information to specific IP address(es):

```
Switch#configure terminal
Switch(config)#access-list 99 permit host 192.168.1.100
Switch(config)#end
Switch#debug ip cef drops 99
```

CEF troubleshooting using the **debug ip cef** command should be coordinated with a Cisco TAC engineer because the **debug** output is not easily decoded.

CEF-Based MLS Troubleshooting Methodology

Previous subsections discussed several CEF troubleshooting techniques on the Layer 3 engine. Recall that the Layer 3 engine does not actually contain the hardware FIB and adjacency table. Rather, these tables are located in specialized hardware components in Supervisor Engines and line cards. The following highlights a stepwise approach to troubleshooting a unicast route on a CEF-based Catalyst switch. The troubleshooting steps are not inclusive but do review the hierarchical approach to troubleshooting CEF-based MLS:

Step 1 Verify that the IP routing information on the Layer 3 engine is correct. Use the **show ip route** or **show ip route** *destination-network* command to verify that the destination network routing entry exists and is associated with a valid next-hop address. If the route does not exist or the next-hop address is incorrect, troubleshooting of routing protocol, next-hop interfaces, or route configuration is required.

Step 2 Verify that the next-hop address has a valid next-hop MAC address by using the **show ip arp** *IP-address* command. If the entry is incomplete, troubleshooting of the ARP process is required.

Step 3 Verify that the IP route entry in the FIB on the Layer 3 engine contains the same next-hop address as in Step 1 by using the **show ip cef** *destination-network* command.

Step 4 Verify that the CEF adjacency table contains the same rewrite information as the ARP table from Step 2 by using the **show adjacency detail | begin** *next_hop_IP_address* command.

Step 5 When all other troubleshooting steps have been exhausted and the CEF-based MLS switch is still experiencing unicast routing issues, verify the population of the FIB and adjacency table in TCAM under the supervision of a TAC engineer.

This chapter contains an example of CEF-based MLS troubleshooting in the Configuration Exercise at the end of this chapter. For more information on troubleshooting CEF on Catalyst switches, refer to the following documents on Cisco.com:

- "How-To Troubleshoot Unicast IP Routing CEF on Catalyst 6000s with a Supervisor 2 in Hybrid Mode," Document ID: 20626

- "Troubleshooting Cisco Express Forwarding Routing Loops," Document ID: 26083

- "Troubleshooting Incomplete Adjacencies with CEF," Document ID: 17812

- "Troubleshooting Prefix Inconsistencies with Cisco Express Forwarding," Document ID: 14540

Study Tips

The following bullets review important BCMSN exam preparation points of this chapter. The bullets only briefly highlight the main points of this chapter related to the BCMSN exam. Consult the text of this chapter for additional information regarding these topics. Table 9-3 lists important commands to review for the BCMSN exam:

- In describing CEF-based MLS in a few sentences, CEF-based MLS is the Layer 3 forwarding mechanism deployed currently on all shipping Catalyst switches. CEF-based MLS consists of two logical devices: a control-plane and a data-plane. In CEF, the control-plane is generally referred to as the software-switching engine or Layer 3 forwarding engine. In addition, the data-plane carries several common names, including the hardware-switching component, "in hardware", hardware forwarding engine, and ASICs.

- The control-plane is responsible for building the IP routing table and ARP table. The IP routing table and ARP table ultimately build the IP CEF and CEF adjacency table. The CEF table contains prefixes of subnets, hosts, and so on, which are used to

quickly search for an adjacency index. The index into the adjacency table provides the next-hop address, interface, and Layer 3 rewrite information.

- The control-plane downloads its CEF and adjacency table to the data-plane (hardware forwarding engine). The hardware forwarding engine consists of hardware components including TCAM that enable high-speed switching and routing of frames. All changes, updates, additions, and so on to the CEF and adjacency table are controlled by the Layer 3 forwarding engine.

- With the Catalyst 6500 family of switches, the control-plane and data-plane are easily distinguishable (MSFC and PFC, respectively). However, the other Catalyst switches do not have separate components from an end-user standpoint, yet these switches still have a distinct control-plane and data-plane.

- Layer 3 switching (or multilayer switching) is the term used to describe the ability of a Cisco Catalyst switch to perform Layer 3 routing in hardware (hardware-switching) using high-speed hardware components.

- Layer 3 switching is beneficial in the campus network as it provides versatility, high-availability, large port density, and high-performance.

- Current shipping Layer 3 switches use CEF to populate IP routing tables and Layer 2 rewrite information in hardware.

- The terms "CEF-based MLS" or just plain "CEF" are used to describe the routing architecture of Catalyst switches.

- No specific configuration is needed for CEF-based MLS because it is enabled by default.

- ARP throttling is a feature that prevents ARP-based DoS attacks on Catalyst switches. The feature works by installing a throttling adjacency in the hardware CEF table such that hardware-switching drops subsequent identical ARP requests for a specific time period until the respective ARP response is received.

- With CEF, when a router is connected directly to several hosts, the FIB table on the router maintains a prefix for the subnet rather than for the individual host prefixes. The subnet prefix points to a glean adjacency. When packets need to be forwarded to a specific host, the adjacency database is gleaned for the specific prefix.

- With CEF, punt adjacencies in TCAM on Catalyst switch hardware are used to send packets to the Layer 3 engine (for example, MSFC) that require special handling or that are used for special situations where the hardware does not yet support the packet flow in question (for example, NAT).

- Catalyst switches support two methods for CEF-based MLS: centralized switching and distributing switching.

- Centralized switching describes hardware-switching using a centralized hardware-switching engine found on the Supervisor Engine of a Cisco Catalyst switch. With centralized switching, data to be routed or switched must pass through the switching fabric of the Catalyst switch to the centralized hardware-switching engine for

processing. After the forwarding decision has been made and the packet rewritten with new Layer 2 information, the centralized switching engine forwards the packet to the correct destination port and VLAN.

- The Catalyst 4500 family of switches is based on centralized switching.

- Distributed switching describes the hardware-switching method used by Catalyst switches that utilize hardware-switching engines on individual line cards and port groups. With distributed switching, individual line cards determine the routing destination and rewrite information for a packet. Therefore, with distributed switching line cards, a Catalyst switch is able to route a packet between source and destination without sending the packet to the Supervisor Engine for processing. In addition, the packet might never leave the line card if the source and destination of the line card are on the same switch.

- Review the "Sample CEF-Based MLS Operation" section and Figure 9-8, which illustrates the path of a packet during CEF-based MLS operation.

- Review the "CEF-Based MLS Troubleshooting Methodology" section of this chapter for details on how to approach CEF-based MLS troubleshooting.

Table 9-3 *Commands to Review*

Command	Description
debug ip cef	Enables debugging with IP CEF. Recommended practice is to use this **debug** option with optional parameters including ACLs to limit output. In addition, use this **debug** command in production only in coordination with a Cisco TAC engineer.
show adjacency detail	Displays the IP CEF Layer 2 rewrite information and statistical information from the perspective of the Layer 3 engine.
show arp	Displays the ARP table contents. This command is the first step in troubleshooting adjacency issues.
show ip cef	Displays the IP CEF information from the Layer 3 engine perspective. Displays the prefix, the next hop, and next-hop interface (outgoing interface).
show ip route	Displays the IP routing table. This command is the first step in troubleshooting IP routing or IP CEF issues.

Summary

Outside the scope of the BCMSN course, the following notes along with the study tips in the previous section summarize the chapter:

- CEF-based MLS is a forwarding model implemented on the latest generation of Cisco multilayer switches. CEF-based MLS is topology based; the control-plane and data-plane are separate. The control-plane downloads the routing table information to the

data-plane for hardware switching. CEF-based MLS scales to large networks and is not limited on the number of traffic flows.

- CEF-based MLS uses either centralized switching or distributed switching. Distributed switching provides higher performance over centralized switching.

- Switches use TCAM and other hardware-switching components not only for IP CEF but also for applying QoS and access lists to packets routed and switched using hardware switching.

- Cisco multilayer switches that support CEF-based MLS use CEF-based MLS by default.

- CEF-based MLS supports statistics via show commands with specific limitation.

Configuration Exercise: Troubleshooting CEF-Based MLS

Complete this configuration exercise to familiarize yourself with the basic troubleshooting steps of CEF-based MLS.

Required Resources

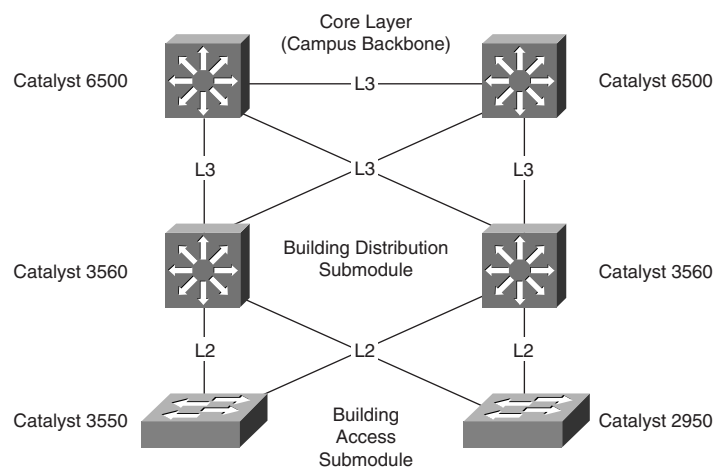The resources and equipment required to complete this exercise are as follows:

- A Cisco IOS–based, Layer 3 CEF-based MLS Catalyst switch such as a Catalyst 3550, 3560, 3750, 4500, or 6500 with Supervisor II or 720 with an MSFC in a Layer 3 Building Distribution submodule or Layer 3 routing topology

- A terminal server or workstation connected directly to the console port of the Catalyst switch

Exercise Objective

The purpose of this exercise is to demonstrate the first steps in troubleshooting CEF-based MLS on a Cisco IOS–based Catalyst switch. Troubleshooting CEF-based MLS is generally a last step in troubleshooting connectivity problems. Generally, connectivity problems are the result of a physical layer issue or configuration issue. Troubleshoot CEF-based MLS after all other troubleshooting steps have been exhausted.

Network Diagram

Figure 9-10 shows the network layout for this configuration exercise. The configuration exercise uses a Catalyst 3560 switch from the Building Distribution submodule to illustrate commands. This switch learns IP routing entries from the Campus Backbone switches.

Figure 9-10 *Network Diagram for Configuration Exercise*

Command List

In this configuration exercise, you use the commands listed in Table 9-4. These commands are in alphabetical order so that you can easily locate the information you need. Refer to this list if you need configuration command assistance during the exercise. The table includes only the specific parameters used in the example and not all the available options for the command.

Table 9-4 *Command List for Configuration Exercise*

Command	Description
ping *ip_address*	Sends ICMP echo requests to a specific IP address
show adjacency detail	Displays the CEF adjacency table and rewrite information
show ip arp *ip-address*	Displays the ARP table for a specific IP address
show ip cef *destination-network* **detail**	Displays the IP CEF FIB table for a specific destination network
show ip route *destination-subnet*	Displays the Cisco IOS IP routing table

Task 1: Establish Console (Out-of-Band) Connectivity or Telnet or SSH Connectivity to the Switch

This task is self-explanatory.

Task 2: Determine the IP Routing Entry or Subnet That Is Experiencing Connectivity Problems or a CEF-Based MLS Issue

For this configuration exercise, use the destination subnet, 10.20.30.0/24, as an example.

Task 3: Verify the Cisco IOS IP Routing Table and ARP Entries for the Route

Step 1   Verify that a route exists to the destination network in the IP routing table.

```
Switch#show ip route 10.20.30.0 255.255.255.0
Routing entry for 10.20.30.0/24
  Known via "ospf 3738", distance 110, metric 2, type intra area
  Last update from 10.20.248.6 on Port-channel2, 05:59:50 ago
  Routing Descriptor Blocks:
  * 10.20.248.34, from 10.20.255.14, 05:59:50 ago, via Port-channel4
      Route metric is 2, traffic share count is 1
    10.20.248.6, from 10.20.255.13, 05:59:50 ago, via Port-channel2
      Route metric is 2, traffic share count is 1
```

Step 2   Verify the ARP entries for the next-hop addresses for the destination subnet obtained in Step 1.

```
Switch#show ip arp 10.20.248.34
Protocol  Address          Age (min)  Hardware Addr   Type    Interface
Internet  10.20.248.34         1   00d0.03e5.840a  ARPA   Port-channel4

Switch#show ip arp 10.20.248.6
Protocol  Address          Age (min)  Hardware Addr   Type    Interface
Internet  10.20.248.6          1   00d0.03eb.b40a  ARPA   Port-channel2
```

Step 3   Using the ping command, send ICMP echoes to the next-hop address to verify connectivity.

```
Switch#ping 10.20.248.34

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.20.248.34, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms

Switch#ping 10.20.248.6

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.20.248.6, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
```

Task 4: Verify the IP CEF FIB and Adjacency Entries for the Route

Step 1   Verify the IP CEF FIB entry for the destination subnet in question.

```
Switch#show ip cef 10.20.30.0 detail
10.20.30.0/24, version 122, epoch 0, per-destination sharing
```

```
0 packets, 0 bytes
  via 10.20.248.34, Port-channel4, 0 dependencies
    traffic share 1
    next hop 10.20.248.34, Port-channel4
    valid adjacency
  via 10.20.248.6, Port-channel2, 0 dependencies
    traffic share 1
    next hop 10.20.248.6, Port-channel2
    valid adjacency
  0 packets, 0 bytes switched through the prefix
```

Step 2   Verify the CEF adjacency entry for the next-hop address obtained from Step 1 and compare the results to Step 2 in Task 3.

```
Switch#show adjacency detail | begin 10.20.248.34
IP       Port-channel4            10.20.248.34(170)
                                  0 packets, 0 bytes
                                  00D003E5840A
                                  00D00624440A0800
                                  ARP        00:05:42
                                  Epoch: 0


Switch#show adjacency detail | begin 10.20.248.6
IP       Port-channel2            10.20.248.6(166)
                                  0 packets, 0 bytes
                                  00D003EBB40A
                                  00D00624440A0800
                                  ARP        00:04:53
                                  Epoch: 0
```

Task 5: Debug the CEF FIB and Adjacency Table's Downloads to the Centralized Switching and Distributed Switching Engines, and Verify the TCAM Contents for FIB and Adjacency Tables

This task should be performed only under the supervision of a Cisco TAC engineer.

Review Questions

Note that more than one correct answer might be valid for multiple-choice questions.

1   True or False: CEF-based MLS Catalyst switches use the Layer 3 engine to route the first packet and build flows based in hardware for subsequent packets.

2   True or False: The performance of a Catalyst switch using distributed switching is the aggregate of all forwarding engines.

3 In CEF-based multilayer switching, which type of adjacency entry is used for features that require special handling or for features that are not yet supported in conjunction with CEF hardware-switching paths?

   a Glean adjacency

   b Punt adjacency

   c Drop adjacency

   d Default adjacency

4 Which statement describes a TCAM mask associated with the entry **access-list 101 deny ip 10.1.0.0 0.0.255.255?**

   a All 32 bits of source address

   b Most significant 8 bits of source address

   c Most significant 16 bits of source address

   d Most significant 24 bits of source address

5 When troubleshooting CEF-based MLS for issues such as being unable to reach a certain IP destination, what two tables need to be verified as a first step in troubleshooting?

   a The IP routing table and the IP route-cache table

   b The IP routing table and the CEF adjacency table

   c The IP CEF FIB table and the adjacency table in TCAM

   d The IP routing table and the ARP table

6 Which of the following statements is true with regard to CEF-based MLS?

   a CEF-based MLS Catalyst switches use the IP CEF FIB and adjacency tables to build FIB and adjacency tables in TCAM for hardware switching.

   b CEF-based MLS Catalyst switches use the IP routing and IP ARP tables to build FIB and adjacency tables in TCAM for software switching.

   c CEF-based MLS Catalyst switches use the IP CEF FIB and adjacency tables to build MLS flows on-demand in TCAM for hardware switching.

7 Describe the steps a multilayer switched packet takes when using CEF-based MLS.