- Multiprotocol Label Switching -

Multiprotocol Label Switching

Multiprotocol Label Switching (MPLS) is a Layer-2 *switching* technology. MPLS-enabled routers apply numerical **labels** to packets, and can make forwarding decisions based on these labels. The MPLS architecture is detailed in RFC 3031.

MPLS reduces CPU-usage on routers, by allowing routers to make forwarding decisions *solely* on the attached label, as opposed to parsing the full routing table.

Labels can based on a variety of parameters:
* Destination IP network
* Source IP address
* QoS parameters
* VPN destination
* Outgoing interface
* Layer-2 circuit

MPLS is not restricted to IP, or any specific Layer-2 technology, and thus is essentially **protocol-independent**.

Labels are applied to and removed from packets on **edge Label Switch Routers (edge LSRs)**. Only edge routers perform a route-table lookup on packets. All **core routers (**identified simply as **LSRs**) in the MPLS network forward solely based on the label.

As a packet traverses the core MPLS network, core routers will swap the label on hop-by-hop basis.

MPLS is completely dependent on **Cisco Express Forwarding (CEF)** to determine the next hop.

(Reference: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/mpls_tsw.htm; http://www.cisco.com/en/US/tech/tk436/tk428/technologies_q_and_a_item09186a00800949e5.shtml)

Cisco Express Forwarding (CEF)

Multilayer switches contain both a switching and routing engine. A packet must first be "routed," allowing the switching engine to cache the IP traffic flow. After this cache is created, subsequent packets destined for that flow can be "switched" as opposed to "routed," reducing latency.

This concept is often referred to as **route once, switch many.** Cisco refers to this type of Multilayer switching as **NetFlow switching** or **route cache switching**.

As is their habit, Cisco replaced NetFlow multilayer switching with a more advanced method called **Cisco Express Forwarding (CEF)**. CEF is **enabled by default** on all Catalyst multi-layer switches (at least, those that support CEF). CEF cannot even be disabled on the Catalyst 3550, 4500 and 6500.

CEF contains two basic components:
- **Layer 3 Engine –** Builds the routing table and then "routes" data
- **Layer 3 Forwarding Engine –** "Switches" data based on the FIB.

The Layer 3 Engine builds its routing table using either static routes, or routes dynamically learned through a routing protocol (such as RIP or OSPF).

The routing table is then reorganized into a more efficient table called the **Forward Information Base (FIB)**. The most specific routes are placed at the top of the FIB. The Layer 3 Forwarding Engine utilizes the FIB to then "switch" data in hardware, as opposed to "routing" it through the Layer 3 Engine's routing table.

The FIB contains the following information:
- Destination networks
- Destination masks
- Next-hop addresses
- The MAC addresses of each next hop (called the **Adjacency Table**)

To view the CEF FIB table:

> **Switch#** *show ip cef*

The MPLS Label

Two forms of MPLS exist:
- **Frame Mode MPLS –** utilizes a 32-bit label that is injected between the Layer-2 and Layer-3 headers.
- **Cell Mode MPLS –** used with ATM, and utilizes the VPI/VCI fields ATM header as the label.

This guide will concentrate on Frame Mode MPLS. The 32-bit label has the following format:

Label	Experimental	Bottom-of-Stack	TTL
20 bits	*3 bits*	*1 bit*	*8 bits*

- **Label (20 bits) –**

- **Experimental (3 bits)** – This field is *officially* undefined, but is used by Cisco as an IP precedence value.

- **Bottom-of-Stack (1 bit)** – This field indicates the *last* label, as multiple labels are supported in the same packet. A value of **1** identifies the last label in the stack.

- **TTL (8 bits)** – This field indicates the number of router this label can 'live' through.

An Ethernet header is modified to indicate the presence of an MPLS label.
- 0x8847 – indicates a labeled unicast IP packet
- 0x8848 – indicates a labeled multicast IP packet

(Reference: http://www.iec.org/online/tutorials/mpls/topic03.html; http://www.rfc-editor.org/rfc/rfc3032.txt)

* * *

The MPLS Components

MPLS router designations include:

- **LSR (Label Switch Router)** – responsible for forwarding packets through the provider core based on the packet's *label*. Cisco refers to this as a **P** (provider) **router.**

- **Edge LSR (Label Edge Router)** – responsible for adding or removing labels from packets. Cisco refers to this as a **PE** (provider edge) **router.**

- **Non**-**Label Routers –** Cisco refers to this as **C** (customer) **routers**

LSRs perform the following functions:
- **Control Plane** - exchanges *routing* and *label* information
- **Data Plane** - forwards actual packets based on label information

The Control Plane, in charge of information exchange, builds and maintains the following tables:
- **Routing Table –** routing information is exchanged between LSRs using a routing protocol, such as IGRP, EIGRP, IS-IS, OSPF, or BGP.
- **Label Information Base (LIB) –** label information is exchanged between LSRs using a label protocol, such as **LDP (Label Distribution Protocol)** or **TDP (Tag Distribution Protocol)**

LDP is now default on Cisco devices, and uses TCP port 646. TDP is a Cisco-proprietary label protocol, and uses TCP port 711.

Label convergence will occur *after* routing convergence is completed. Label protocols require the underlying routing infrastructure in order to function.

The Data Plane, in charge of information forwarding, maintains the following tables:
- **The CEF Forwarding Information Base (FIB) –**for forwarding unlabeled packets. Contains destination IP networks, and the Layer-2 address of the next-hop router.
- **Label Forwarding Information Base (LFIB)** – for forwarding labeled packets.

The MPLS Process

There are four scenarios detailing how LSRs forward packets:

1. An *unlabeled* IP packet is received, and is **routed** unlabeled to the next hop.
2. An *unlabeled* IP packet is received, a label is inserted in the header, and is **switched** to the next hop.
3. A *labeled* IP packet is received, the label is swapped, and is **switched** to the next hop.
4. A *labeled* IP packet is received, the label is stripped off, and is routed to the next hop or destination.

Frame-mode MPLS performs as follows:
1. An edge LSR receives a packet.
2. The edge LSR performs a routing table lookup to determine the next hop (or exit interface).
3. If destined for the MPLS network, the edge router inserts the label between the Layer-2 and Layer-3 headers.
4. The edge LSR forwards the labeled packet to the core LSR.
5. Core LSRs will route solely based on the label, and will not perform a routing table lookup.

Configuring Basic MPLS

The first step in configuring MPLS is enabling CEF switching, which can be accomplished globally (for all interfaces), or on a per interface basis:

> **Router(config)#** *ip cef*
>
> **Router(config)#** *interface serial0/0*
> **Router(config-if)#** *ip route-cache cef*

To view CEF information:

> **Router(config)#** *show ip cef*

Next, MPLS must be enabled on the interface:

> **Router(config)#** *interface serial0/0*
> **Router(config-if)#** *mpls ip*

The desired *label protocol* can then be specified (either *ldp, tdp*, or *both*):

> **Router(config)#** *interface serial0/0*
> **Router(config-if)#** *mpls label protocol ldp*
>
> **Router(config)#** *interface serial0/0*
> **Router(config-if)#** *mpls label protocol tdp*
>
> **Router(config)#** *interface serial0/0*
> **Router(config-if)#** *mpls label protocol both*

If a label protocol is *not* specified, Cisco devices will default to *ldp* (as of the IOS version 12.4).

Because the MPLS label increases the size of the frame by 32 bits (4 bytes), the *mtu* for *mpls* packets should be adjusted to accommodate this:

> **Router(config)#** *interface serial0/0*
> **Router(config-if)#** *mpls mtu 1504*

MPLS VPNs require *two* labels, thus the *mtu* should be adjusted accordingly:

> **Router(config)#** *interface serial0/0*
> **Router(config-if)#** *mpls mtu 1508*

(Reference: http://www.cisco.com/en/US/products/ps6350/products_configuration_guide_chapter09186a008045543c.html)

* * *

MPLS VPNs

Cisco identifies two key categories of VPNs:

- **Overlay –** connections are set up and maintained by the service provider. However, the provider has no knowledge of, and does not participate in, the customer's routing infrastructure.

- **Peer-to-Peer** – the provider directly participates in routing the customer's infrastructure. Routes from multiple customers are not kept separate. This may require customers to readdress their networks.

MPLS VPNs provide the best of both words. Advantages of MPLS VPNs include:
- The provider directly participates in routing the customer infrastructure.
- Peer-to-peer peering is not required, leading to a scalable infrastructure.
- Customer networks do not need to be readdressed
- Routes from multiple customers are kept separate.

MPLS VPNs use **two labels**.

MPLS separates customer routes by assigning each a unique **Virtual Routing Instance (VRI)**, stored in a **Virtual Routing and Forwarding** (VRF) table.

If the addressing structure of multiple clients overlaps, each network is assigned a unique 64-bit **route distinguisher (RD).** The IP network and RD combination is called a **VPNv4 address**.

Route targets are BGP attributes that designate MPLS VPN membership of routes.