Introduction to MPLS

Santanu Dasgupta
santanu@cisco.com
Goals of this Session

- Understand history and business drivers for MPLS
- Learn about MPLS customer and market segments
- Understand the problems MPLS is addressing
- Understand benefits of deploying MPLS
- Understand the major MPLS technology components
- Learn the basics of MPLS technology
- Understand typical applications of MPLS
The Big Picture

End-to-end Services
- Layer-3 VPNs
- Layer-2 VPNs

MPLS Network Services
- MPLS QoS
- MPLS TE
- MPLS OAM/MIBs

Core MPLS
- MPLS Signaling and Forwarding

Network Infrastructure
Agenda

- Introduction
- MPLS Network Components
  - MPLS VPNs
    - MPLS Layer-3 VPNs
    - MPLS Layer-2 VPNs
- MPLS QoS
- MPLS Traffic Engineering
- MPLS Management
- Summary

Core MPLS
End-to-end MPLS Services
MPLS Network Services
Introduction

The business drivers for MPLS
Why Multi Protocol Label Switching?

- **SP/Carrier perspective**
  - Reduce costs (CAPEX); consolidate networks
  - Consolidated network for multiple Layer-2/3 services
  - Support increasingly stringent SLAs
  - Handle increasing scale/complexity of IP-based services

- **Enterprise/end-user perspective**
  - Campus/LAN
    - Need for network segmentation (users, applications, etc.)
  - WAN connectivity (connecting enterprise networks)
    - Need for easier configuration of site-to-site WAN connectivity
What Is MPLS Technology?

- It’s all about labels …
- Use the best of both worlds
  - Layer-2 (ATM/FR): efficient forwarding and traffic engineering
  - Layer-3 (IP): flexible and scalable
- MPLS forwarding plane
  - Use of labels for forwarding Layer-2/3 data traffic
  - Labeled packets are being switched instead of routed
    - Leverage layer-2 forwarding efficiency
- MPLS control/signaling plane
  - Use of existing IP control protocols extensions + new protocols to exchange label information
    - Leverage layer-3 control protocol flexibility and scalability
Evolution of MPLS

- Evolved from tag switching in 1996 to full IETF standard, covering over 130 RFCs
- Key application initially were Layer-3 VPNs, followed by Traffic Engineering (TE), and Layer-2 VPNs
# MPLS Applications

## Key Features

<table>
<thead>
<tr>
<th>Applications</th>
<th>Enterprise Data Center</th>
<th>Data center interconnects</th>
<th>EWAN Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Service Providers</td>
<td>Data center interconnects</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hosted Data centers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data center interconnect</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Segmentations for IT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mergers, Acquisitions, spinoffs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L2/L3VPN’s</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TE/FRR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>QoS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High Availability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Departmental segmentation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Service multiplexing</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Security</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mergers, Acquisitions, spinoffs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Disaster Recovery</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vmotion support</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Branch Interconnects</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Internet Access</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Branch Connectivity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Applications

- **Network Consolidation** – Merging multiple parallel networks into a shared infrastructure
- **Network segmentation** – By user groups or business function
- **Service and policy centralization** – Security policies and appliances at a central location
- **New applications readiness** – Converged multi-service network
- **Increased network security** – User groups segmentation with VPNs

© 2010 Cisco and/or its affiliates. All rights reserved.
Enterprise MPLS Customers

- Two types of enterprise customers for MPLS technology

- MPLS indirectly used as subscribed WAN service
  
  Enterprise subscribes to WAN connectivity data service offered by external Service Provider
  
  Data connectivity service implemented by Service Provider via MPLS VPN technology (e.g., layer-2 and layer-3 VPNs)
  
  VPN Service can be managed or unmanaged

- MPLS used as part of self managed network
  
  Enterprise deploys MPLS in it’s own network
  
  Enterprise manages it’s own MPLS-based network
MPLS Technology Framework

- **End-to-end Services**
  - Layer-3 VPNs
  - Layer-2 VPNs

- **MPLS Network Services**
  - MPLS QoS
  - MPLS TE
  - MPLS OAM/MIBs

- **Core MPLS**
  - MPLS Signaling and Forwarding

- **Network Infrastructure**
MPLS Technology Components

Basic building blocks of MPLS
MPLS Forwarding and Signaling

- MPLS label forwarding and signaling mechanisms

- Layer-3 VPNs
  - MPLS QoS

- Layer-2 VPNs
  - MPLS TE
  - MPLS OAM/MIBs

- Core MPLS
  - MPLS Signaling and Forwarding

- Network Infrastructure
Basic Building Blocks

- The big picture
  - MPLS-enabled network devices
  - Label Switched Paths (LSPs)

- The internals
  - MPLS labels
  - Processing of MPLS labels
  - Exchange of label mapping information
  - Forwarding of labeled packets

- Other related protocols and protocols to exchange label information
  - Between MPLS-enabled devices
MPLS Network Overview

- **P** (Provider) router = label switching router = core router (LSR)
  - Switches MPLS-labeled packets
- **PE** (Provider Edge) router = edge router (LSR)
  - Imposes and removes MPLS labels
- **CE** (Customer Edge) router
  - Connects customer network to MPLS network
MPLS Label and Label Encapsulation

MPLS Label

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
Label # – 20bits           EXP     S     TTL-8bits
```

COS/EXP = Class of Service: 3 Bits; S = Bottom of Stack; TTL = Time to Live

MPLS Label Encapsulation

PPP Header
(Packet over SONET/SDH)

LAN MAC Label Header

One or More Labels Appended to the Packet
(Between L2/L3 packet header and link layer header)
MPLS Label Operations

- **Label imposition (Push)**
  By ingress PE router; classify and label packets

- **Label swapping or switching**
  By P router; forward packets using labels; indicates service class & destination

- **Label disposition (PoP)**
  By egress PE router; remove label and forward original packet to destination CE
Forwarding Equivalence Class

- Mechanism to map ingress layer-2/3 packets onto a Label Switched Path (LSP) by ingress PE router
  - Part of label imposition (Push) operation
- Variety of FEC mappings possible
  - IP prefix/host address
  - Groups of addresses/sites (VPN x)
    - Used for L3VPNs
  - Layer 2 circuit ID (ATM, FR, PPP, HDLC, Ethernet)
    - Used for Pseudowires (L2VPNs)
  - A bridge/switch instance (VSI)
    - Used for VPLS (L2VPNs)
  - Tunnel interface
    - Used for MPLS traffic engineering (TE)
Label Distribution Protocol

- MPLS nodes need to exchange label information with each other
  - Ingress PE node (Push operation)
    Needs to know what label to use for a given FEC to send packet to neighbor
  - Core P node (Swap operation)
    Needs to know what label to use for swap operation for incoming labeled packets
  - Egress PE node (Pop operation)
    Needs to tell upstream neighbor what label to use for specific FEC type LDP used for exchange of label (mapping) information

- Label Distribution Protocol (LDP)
  Defined in RFC 3035 and RFC3036; updated by RFC5036
  LDP is a superset of the Cisco-specific Tag Distribution Protocol

- Note that, in addition LDP, also other protocols are being used for label information exchange
  Will be discussed later
Some More LDP Details

- Assigns, distributes, and installs (in forwarding) labels for prefixes advertised by unicast routing protocols: OSPF, IS-IS, EIGRP, etc.
- Also used for Pseudowire/PW (VC) signaling
  Used for L2VPN control plane signaling
- Uses UDP (port 646) for session discovery and TCP (port 646) for exchange of LDP messages
- LDP operations
  LDP Peer Discovery
  LDP Session Establishment
  MPLS Label Allocation, Distribution, and Updating MPLS forwarding
- Information repositories used by LDP
  LIB: Label Information Database (read/write)
  RIB: Routing Information Database/routing table (read-only)
LDP Operations

- **LDP startup**
  - Local labels assigned to RIB prefixes and stored in LIB
  - Peer discovery and session setup
  - Exchange of MPLS label bindings

- **Programming of MPLS forwarding**
  - Based on LIB info
  - CEF/MFI updates
MPLS Control and Forwarding Plane

- **MPLS control plane**
  - Used for distributing labels and building label-switched paths (LSPs)
  - Typically supported by LDP; also supported via RSVP and BGP
  - Labels define destination and service

- **MPLS forwarding plane**
  - Used for label imposition, swapping, and disposition
  - Independent of type of control plane
  - Labels separate forwarding from IP address-based routing
IP Packet Forwarding Example

<table>
<thead>
<tr>
<th>Address</th>
<th>Prefix</th>
<th>I/F</th>
<th>FIB</th>
</tr>
</thead>
<tbody>
<tr>
<td>128.89</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>171.69</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Packets Forwarded Based on IP Address (via RIB lookup)
**Step 1: IP Routing (IGP) Convergence**

<table>
<thead>
<tr>
<th>In Label</th>
<th>Address Prefix</th>
<th>Out l’face</th>
<th>Out Label</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>128.89</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>171.69</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

You Can Reach 128.89 and 171.69 Thru Me

Routing Updates (OSPF, EIGRP, …)

You Can Reach 128.89 Thru Me

You Can Reach 171.69 Thru Me

© 2010 Cisco and/or its affiliates. All rights reserved.
Step 2a: LDP Assigns Local Labels

<table>
<thead>
<tr>
<th>In Label</th>
<th>Address Prefix</th>
<th>Out I'face</th>
<th>Out Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>128.89</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>171.69</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Step 2b: LDP Assigns Remote Labels

**Label Distribution Protocol (LDP) (Downstream Allocation)**

Use Label 4 for 128.89 and Use Label 5 for 171.69

Use Label 7 for 171.69

Use Label 9 for 128.89
Step 3: Forwarding MPLS Packets

Label Switch Forwards Based on Label
Summary Steps For MPLS Forwarding

- Each node maintains IP routing information via IGP
  IP routing table (RIB) and IP forwarding table (FIB)
- LDP leverages IGP routing information
- LDP label mapping exchange (between MPLS nodes) takes place after IGP has converged
  LDP depends on IGP convergence
  Label binding information stored in LIB
- Once LDP has received remote label binding information MPLS forwarding is updated
  Label bindings are received from remote LDP peers
  MPLS forwarding via MFI
MPLS Network Protocols

- IGP: OSPF, EIGRP, IS-IS on core facing and core links
- RSVP and/or LDP on core and/or core facing links
- MP-iBGP on PE devices (for MPLS services)
Label Stacking

- More than one label can be used for MPLS packet encapsulation
  Creation of a label stack

- Recap: labels correspond to Forwarding Equivalence Class (FEC)
  Each label in stack used for different purposes

- Outer label always used for switching MPLS packets in network

- Remaining inner labels used to specific services/FECs, etc.

- Last label in stack marked with EOS bit

- Allows building services such as
  - MPLS VPNs; LDP + VPN label
  - Traffic engineering (FRR): LDP + TE label
  - VPNs over TE core: LDP + TE + VPN label
  - Any transport over MPLS: LDP + PW label

![Diagram of label stacking]

© 2010 Cisco and/or its affiliates. All rights reserved.
Summary

- MPLS uses labels to forward traffic
- More than one label can be used for traffic encapsulation; multiple labels make up a label stack
- Traffic is encapsulated with label(s) at ingress and at egress labels are removed in MPLS network
- MPLS network consists of PE router at ingress/egress and P routers in the core
- MPLS control plane used for signaling label mapping information to set up end-to-end Label Switched Paths
- MPLS forwarding plane used for label imposition (PUSH), swapping, and disposition (POP) operation
MPLS VPNs

Overviews
MPLS Technology Framework

- End-to-end data connectivity services across MPLS networks (from PE to PE)

**End-to-end Services**

- Layer-3 VPNs
- Layer-2 VPNs

**Key Components**

- MPLS QoS
- MPLS TE
- MPLS OAM/MIBs

- MPLS Signaling and Forwarding

**Network Infrastructure**
What Is a Virtual Private Network?

- VPN is a set of sites or groups which are allowed to communicate with each other in a secure way
  - Typically over a shared public or private network infrastructure
- VPN is defined by a set of administrative policies
  - Policies established by VPN customers themselves (DIY)
  - Policies implemented by VPN service provider (managed/unmanaged)
- Different inter-site connectivity schemes possible
  - Ranging from complete to partial mesh, hub-and-spoke
- Sites may be either within the same or in different organizations
  - VPN can be either intranet or extranet
- Site may be in more than one VPN
  - VPNs may overlap
- Not all sites have to be connected to the same service provider
  - VPN can span multiple providers
MPLS VPN Example

- **PE-CE link**
  Connect customer network to SP network; layer-2 or layer-3

- **VPN**
  Dedicated secure connectivity over shared infrastructure
MPLS VPN Benefits

- SP/Carrier perspective
  - Reduce costs (CAPEX)
    - Leverage same network for multiple services and customers
    - Migrate legacy networks onto single converged network
  - Reduce costs (OPEX)
    - Easier service enablement; only edge node configuration

- Enterprise/end-user perspective
  - Enables site/campus network segmentation
    - Allows for dedicated connectivity for users, applications, etc.
  - Enables easier setup of WAN connectivity
    - Easier configuration of site-to-site WAN connectivity (for L3VPN and VPLS); only one WAN connection needed
MPLS VPN Options

MPLS VPN Models

Layer-2 VPNs
- Point-to-Point Layer-2 VPNs
  - CPE connected to PE via p2p Layer-2 connection (FR, ATM)
  - CEs peer with each other (IP routing) via p2p layer-2 VPN connection
  - CE-CE routing; no SP involvement
- Multi-Point Layer-2 VPNs
  - CPE connected to PE via Ethernet connection (VLAN)
  - CEs peer with each other via fully/partial mesh Layer-2 VPN connection
  - CE-CE routing; no SP involvement

Layer-3 VPNs
- CPE connected to PE via IP-based connection (over any layer-2 type)
  - Static routing
  - PE-CE routing protocol; eBGP, OSPF, IS-IS
- CEs peer with PE router
- PE routers maintain customer-specific routing tables and exchange customer-specific routing information
- Layer-3 VPN provider’s PE routers are part of customer routing
MPLS Layer-3 VPNs

Technology Overview and Applications
MPLS L3 VPN Overview

- Customer router (CE) has an IP peering connection with PE/edge router in MPLS network
  - IP routing/forwarding across PE-CE link
- MPLS VPN network responsible for distributing routing information to remote VPN sites
  - MPLS VPN part of customer IP routing domain
- MPLS VPNs enable full-mesh, hub-and-spoke, and hybrid connectivity among connected CE sites
- MPLS VPN service enablement in MPLS networks only requires VPN configuration at edge/PE nodes
  - Connectivity in core automatically established via BGP signaling
MPLS L3 VPN Technology Components

- **PE-CE link**
  Can be any type of layer-2 connection (e.g., FR, Ethernet)
  CE configured to route IP traffic to/from adjacent PE router
  Variety of routing options; static routes, eBGP, OSPF, IS-IS

- **MPLS L3VPN Control Plane**
  Separation of customer routing via virtual VPN routing table
  In PE router: customer I/Fs connected to virtual routing table
  Between PE routers: customer routes exchanged via BGP

- **MPLS L3VPN Forwarding Plane**
  Separation of customer VPN traffic via additional VPN label
  VPN label used by receiving PE to identify VPN routing table
Virtual Routing and Forwarding Instance

- Virtual Routing and Forwarding Instance (VRF)
- Typically one VRF created for each customer VPN on PE router
- VRF associated with one or more customer interfaces
- VRF has its own instance of routing table (RIB) and forwarding table (CEF)
- VRF has its own instance for PE-CE configured routing protocols
VPN Route Distribution

- Full mesh of BGP sessions among all PE routers
  - Multi-Protocol BGP extensions (MP-iBGP)
  - Typically BGP Route Reflector (RR) used for improved scalability
VPN Control Plane Processing

**Make customer routes unique:**
- **Route Distinguisher (RD):** 8-byte field, VRF parameters; unique value assigned by a provider to each VPN to make different VPN routes unique
- **VPNv4 address:** RD+VPN IP prefix

**Selective distribute customer routes:**
- **Route Target (RT):** 8-byte field, VRF parameter, unique value to define the import/export rules for VPNv4 routes
- **MP-iBGP:** advertises VPNv4* prefixes + labels

**Processing Steps:**
1. CE1 redistribute IPv4 route to PE1 via eBGP.
2. PE1 allocates VPN label for prefix learnt from CE1 to create unique VPNv4 route.
3. PE1 redistributes VPNv4 route into MP-iBGP, it sets itself as a next hop and relays VPN site routes to PE2.
4. PE2 receives VPNv4 route and, via processing in local VRF (green), it redistributes original IPv4 route to CE2.
VPN Forwarding Plane Processing

Processing Steps:
1. CE2 forwards IPv4 packet to PE2.
2. PE2 imposes pre-allocated VPN label (learned via MP-IBGP) to IPv4 packet received from CE2.
3. PE2 imposes outer IGP label (learned via LDP) and forwards labeled packet to next-hop P-router P2.
4. P-routers P1 and P2 swap outer IGP label and forward label packet to PE1.
5. Router PE1 strips VPN label and forwards IPv4 packet to CE1.
Use Case 1: Traffic Separation

**Requirement:** Need to ensure data separation between Aerospace, Cosmetics and Financial Services, while leveraging a shared infrastructure

**Solution:** Create MPLS VPN for each group

![Diagram showing traffic separation using MPLS VPNs for Aerospace, Cosmetics, and Financial Services.]
Use Case 2: Simplify Hub Site Design

**Requirement:** To ease the scale and design of head-end site

**Solution:** Implement MPLS Layer 3 VPNs, which reduces the number of routing peers of the central site

**Without MPLS**
- Central site has high number of routing peers – creates a complicated headend design

**With MPLS**
- Central site has a single routing peer – enhancing head-end design
# Enterprise Network Segmentation

<table>
<thead>
<tr>
<th>VRF-lite + 802.1Q VLANs</th>
<th>Distribution</th>
<th>Core</th>
<th>End-to-end Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>VRF lite configured on distribution nodes</td>
<td>VRF lite configured on core nodes</td>
<td>VRF lite configured on core nodes</td>
<td>Device Separation: VRF</td>
</tr>
<tr>
<td>VLAN mapping onto VRFs</td>
<td>802.1Q VLAN ID mapping onto VRFs</td>
<td>802.1Q VLAN ID</td>
<td>Data Path Separation: 802.1Q VLAN ID</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VRF-lite + GRE tunnels</th>
<th>Distribution</th>
<th>Core</th>
<th>End-to-end Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>VRF lite configured on distribution nodes</td>
<td>Core nodes forward IP packets (GRE IP Packets)</td>
<td>Core nodes forward IP packets (GRE IP Packets)</td>
<td>End-to-end GRE tunnels between distribution nodes</td>
</tr>
<tr>
<td>VRFs associated with GRE tunnels</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer-3 MPLS VPNs</th>
<th>Distribution</th>
<th>Core</th>
<th>End-to-end Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution nodes configured as PE routers with VRF(s)</td>
<td>Core nodes forward MPLS packets (via LFIB)</td>
<td>Core nodes forward MPLS packets (via LFIB)</td>
<td>End-to-end label switched paths (LSPs) between distribution nodes (PE routers)</td>
</tr>
</tbody>
</table>
Option 1: VRF-lite + 802.1Q

- Layer-2 access
- No BGP or MPLS
- VRF-lite configured on core and distribution nodes
- MPLS labels substituted by 802.1q tags end-to-end
- Every link is a 802.1Q trunk
- Many-to-Many model
- Restricted scalability
- Typical for department inter-connectivity
Option 2: VRF-lite + GRE

- L2 access
- No BGP or MPLS
- VRF-lite only configured on distribution nodes
- VLANs associated with end-to-end GRE Tunnels
- Many-to-One model
- Restricted scalability
- Typical for user-specific VPN connectivity
Option 3: Layer-3 MPLS VPNs

- L2 access
- Distribution nodes configured as PE routers with VRFs
- MP-iBGP between distribution nodes
- MPLS packet forwarding by core nodes
- Many-to-Many model
- High scalability
MPLS Layer-3 VPN Summary

- Provide layer-3 connectivity among CE sites via IP peering (across PE-CE link)
- Implemented via VRFs on edge/PE nodes providing customer route and forwarding segmentation
- BGP used for control plane to exchange customer VPN (VPNv4) routes between PE routers
- MPLS VPNs enable full-mesh, hub-and-spoke, and hybrid IP connectivity among connected CE sites
- L3 VPNs for enterprise network segmentation can also be implemented via VRFs + GRE tunnels or VLANs
MPLS Layer-2 VPNs

Technology Overview and Applications
L2VPN Options

L2VPN Models

VPWS
Virtual Private Wire Service
Point to Point

VPWS Options:
- L2TPv3
- IP Core
- Ethernet
- Frame Relay
- ATM (AAL5 and Cell)
- PPP and HDLC

VPLS
Virtual Private LAN Service
Point to Multipoint

VPLS Options:
- AToM
- MPLS Core
- Ethernet
- Frame Relay
- ATM (AAL5 and Cell)
- PPP and HDLC

Any Transport over MPLS: AToM

MPLS Layer-2 VPNs

© 2010 Cisco and/or its affiliates. All rights reserved.
Layer-2 VPN Overview

- Enables transport of any Layer-2 traffic over MPLS network

  Includes label encapsulation and translation

  - Ethernet
  - ATM
  - HDLC
  - PPP
  - FR
  - Pseudo Wire
  - Many Subscriber Encapsulations Supportable
Any Transport over MPLS Architecture

- Based on IETF’s Pseudo-Wire (PW) Reference Model
- PW is a connection (tunnel) between 2 PE Devices, which connects 2 PW End-Services
  - PW connects 2 Attachment Circuits (ACs)
  - Bi-directional (for p2p connections)
  - Use of PW/VC label for encapsulation
AToM Technology Components

- PE-CE link
  - Referred to as Attachment Circuit (AC)
  - Can be any type of layer-2 connection (e.g., FR, Ethernet)

- AToM Control Plane
  - Targeted LDP (Label Distribution Protocol) Session
    - Virtual Connection (VC)-label negotiation, withdrawal, error notification

- AToM Forwarding Plane
  - 2 labels used for encapsulation + control word
  - Outer tunnel (LDP) label
    - To get from ingress to egress PE using MPLS LSP
  - Inner de-multiplexer (VC) label
    - To identify L2 circuit (packet) encapsulated within tunnel label
  - Control word
    - Replaces layer-2 header at ingress; used to rebuild layer-2 header at egress
AToM Control Plane Processing

Processing Steps (for both P1 and P2):
1. CE1 and CE2 are connected to PE routers via layer-2 connections.
2. Via CLI, a new virtual circuit cross-connect is configured, connecting customer interface to manually provided VC ID with target remote PE.
3. New targeted LDP session between PE routers established, in case one does not already exist.
4. PE binds VC label with customer layer-2 interface and sends label-mapping message to remote PE over LDP session.
5. Remote PE receives LDP label binding message and matches VC ID with local configured cross-connect.
AToM Forwarding Plane Processing

Processing Steps:
1. CE2 forwards layer-2 packet to PE2.
2. PE2 imposes VC (inner) label to layer-2 packet received from CE2 and optionally a control word as well (not shown).
3. PE2 imposes Tunnel outer label and forwards packet to P2.
4. P2 and P1 router forwards packet using outer (tunnel) label.
5. Router PE2 strips Tunnel label and, based on VC label, layer-2 packet is forwarded to customer interface to CE1, after VC label is removed.
   
   In case control word is used, new layer-2 header is generated first.
Use Case: L2 Network Interconnect

**Requirement:** Need to create connectivity between remote customer sites, currently interconnected via Frame Relay WAN connectivity. Only point-to-point connectivity required.

**Solution:** Interconnect AToM PW between sites, enabling transparent Frame Relay WAN connectivity.

VC1 – Connects DLCI 101 to DLCI 201

<table>
<thead>
<tr>
<th>Label Exchange for VC1 – Label 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE1</td>
</tr>
<tr>
<td>DLCI 101</td>
</tr>
<tr>
<td>CPE Router, FRAD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Directed LDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE2</td>
</tr>
<tr>
<td>DLCI 201</td>
</tr>
<tr>
<td>CPE Router, FRAD</td>
</tr>
</tbody>
</table>

Directed LDP – Label 50
Directed LDP – Label 90

VC1 – Connects DLCI 101 to DLCI 201
Virtual Private LAN Service Overview

- Architecture for Ethernet Multipoint Services (EMS) over MPLS
- Emulates IEEE Ethernet bridge; VPLS network acts like a virtual switch that emulates conventional L2 bridge
- Fully meshed or Hub-Spoke topologies supported
VPLS Technology Components

- PE-CE link
  Referred to as Attachment Circuit (AC)
  Ethernet VCs are either port mode or VLAN ID

- VPLS Control Plane
  Full mesh of targeted LDP sessions
  Virtual Connection (VC)-label negotiation, withdrawal, error notification

- VPLS Forwarding Plane
  Virtual Switching Instance: VSI or VFI (Virtual Forwarding Instance)
  VPN ID: Unique value for each VPLS instance
  PWs for interconnection of related VSI instances
VPLS Overview

Full Mesh of PWs Between VSIs

Directed LDP Session Between Participating PEs

Full Mesh of Targeted-LDP Sessions
Exchange VC Labels
Use Case: VPLS Network Interconnect

**Requirement:** Need to create full-mesh connectivity between separate metro networks.

**Solution:** Use VPLS to create transparent bridge layer-2 Ethernet connectivity between ethernet networks.

---

**VPLS VPN Name:** VPLS-CarrierA  
**VPN ID:** 1100  
**VCID:** 1234  
Each PE points to other peer PE’s loopback address
Layer-2 VPN Summary

- Enables transport of any Layer-2 traffic over MPLS network
- Two types of L2 VPNs; AToM for point-to-point and VPLS point-to-multipoint layer-2 connectivity
- Layer-2 VPN forwarding based on Pseudo Wires (PW), which use VC label for L2 packet encapsulation. LDP used for PW signaling
- AToM PWs suited for implementing transparent point-to-point connectivity between Layer-2 circuits
- VPLS suited for implementing transparent point-to-multipoint connectivity between Ethernet links/sites
MPLS QoS

Technology Overview and Applications
MPLS Technology Framework

- MPLS QoS support for traffic marking and classification to enable differentiated services

- Layer-3 VPNs
- Layer-2 VPNs
- MPLS QoS
- MPLS TE
- MPLS OAM/MIBs
- MPLS Signaling and Forwarding

Network Infrastructure
Why MPLS QoS?

- Typically different traffic types (packets) sent over MPLS networks
  
  E.g., Web HTTP, VoIP, FTP, etc.

- Not all application traffic types/flows are the same …
  
  Some require low latency to work correctly; e.g., VoIP

- MPLS QoS used for traffic prioritization to guarantee minimal traffic loss and delay for high priority traffic
  
  Involves packet classification and queuing

- MPLS leverages mostly existing IP QoS architecture
  
  Based on Differentiated Services (DiffServ) model; defines per-hop behavior based on IP Type of Service (ToS) field
MPLS QoS Operations

- MPLS EXP bits used for packet classification and prioritization instead of IP Type of Service (ToS) field
  DSCP values mapped into EXP bits at ingress PE router
- Most providers provide 3–5 service classes
- Different DSCP <-> EXP mapping schemes
  Uniform mode, pipe mode, and short pipe mode
MPLS Uniform Mode

- End-to-end behavior: original IP DSCP value not preserved
  - At ingress PE, IP DSCP value copied in EXP value
  - EXP value changed in the MPLS core
  - At egress PE, EXP value copied back into IP DSCP value
MPLS Pipe Mode

- End-to-end behavior: original IP DSCP is preserved
  - At ingress PE, EXP value set based on ingress classification
  - EXP changed in the MPLS core
  - At egress PE, EXP value not copied back into IP DSCP value
MPLS Short Pipe Mode

- End-to-end behavior: original IP DSCP is preserved
  - At ingress PE, EXP value set based on ingress classification
  - EXP changed in the MPLS core
  - At egress PE, original IP DSCP value used for QoS processing
MPLS QoS Summary

- MPLS QoS used for MPLS packet-specific marking and classification
  - Based on EXP bits

- Different schemes for mapping between IP (ToS/DSCP) and MPLS packet (EXP) classification
  - At ingress and egress PE router
  - MPLS pipe mode mostly used; preserves end-to-end IP QoS

- Enables traffic prioritization to guarantee minimal traffic loss and delay for high priority traffic
  - Useful when packet loss and delay guarantees must be provided for high priority traffic across MPLS network
MPLS Traffic Engineering

Technology Overview and Applications
MPLS Technology Framework

- Traffic engineering capabilities for bandwidth management and network failure protection

- Layer-3 VPNs
- Layer-2 VPNs
- MPLS QoS
- MPLS TE
- MPLS OAM/MIBs
- MPLS Signaling and Forwarding
- Network Infrastructure
Why Traffic Engineering?

- Congestion in the network due to changing traffic patterns
  - Election news, online trading, major sports events
- Better utilization of available bandwidth
  - Route on the non-shortest path
- Route around failed links/nodes
  - Fast rerouting around failures, transparently to users
  - Like SONET APS (Automatic Protection Switching)
- Build new services—virtual leased line services
  - VoIP toll-bypass applications, point-to-point bandwidth guarantees
- Capacity planning
  - TE improves aggregate availability of the network
The Problem with Shortest-Path

- Massive (44%) packet loss at router B→router E!
- Some links are DS3, some are OC-3
- Router A has 40M of traffic for router F, 40M of traffic for router G
- Changing to A→C→D→E won’t help

IP (Mostly) Uses Destination-Based Least-Cost Routing
Alternate Path Under Utilized

<table>
<thead>
<tr>
<th>Node</th>
<th>Next-Hop</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>B</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>10</td>
</tr>
<tr>
<td>D</td>
<td>C</td>
<td>20</td>
</tr>
<tr>
<td>E</td>
<td>B</td>
<td>20</td>
</tr>
<tr>
<td>F</td>
<td>B</td>
<td>30</td>
</tr>
<tr>
<td>G</td>
<td>B</td>
<td>30</td>
</tr>
</tbody>
</table>
How MPLS TE Solves the Problem

- Router A sees all links
- Router A computes paths on properties other than just shortest cost; creation of 2 tunnels
- No link oversubscribed!

<table>
<thead>
<tr>
<th>Node</th>
<th>Next-Hop</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>B</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>10</td>
</tr>
<tr>
<td>D</td>
<td>C</td>
<td>20</td>
</tr>
<tr>
<td>E</td>
<td>B</td>
<td>20</td>
</tr>
<tr>
<td>F</td>
<td>Tunnel 0</td>
<td>30</td>
</tr>
<tr>
<td>G</td>
<td>Tunnel 1</td>
<td>30</td>
</tr>
</tbody>
</table>
How MPLS TE Works

- Link information Distribution*
  - ISIS-TE
  - OSPF-TE
- Path Calculation (CSPF)*
- Path Setup (RSVP-TE)
- Forwarding Traffic down Tunnel
  - Auto-route
  - Static
  - PBR
  - CBTS / PBTS
  - Forwarding Adjacency
  - Tunnel select

* Optional
Link Information Distribution

- Additional link characteristics
  - Interface address
  - Neighbor address
  - Physical bandwidth
  - Maximum reservable bandwidth
  - Unreserved bandwidth (at eight priorities)
  - TE metric
  - Administrative group (attribute flags)

- IS-IS or OSPF flood link information

- TE nodes build a topology database

- Not required if using off-line path computation
Path Calculation

- TE nodes can perform constraint-based routing
- Constraints and topology database as input to path computation
- Shortest-path-first algorithm ignores links not meeting constraints
- Tunnel can be signaled once a path is found
- Not required if using offline path computation

Find shortest path to R8 with 8Mbps

Link with insufficient bandwidth
Link with sufficient bandwidth

IP/MPLS

TE nodes can perform constraint-based routing
Constraints and topology database as input to path computation
Shortest-path-first algorithm ignores links not meeting constraints
Tunnel can be signaled once a path is found
Not required if using offline path computation
TE LSP Signaling

- Tunnel signaled with TE extensions to RSVP
- Soft state maintained with downstream PATH messages
- Soft state maintained with upstream RESV messages
- New RSVP objects
  - LABEL_REQUEST (PATH)
  - LABEL (RESV)
  - EXPLICIT_ROUTE
  - RECORD_ROUTE (PATH/RESV)
  - SESSION_ATTRIBUTE (PATH)
- LFIB populated using RSVP labels allocated by RESV messages
MPLS TE FRR - Link Protection

- Primary tunnel: A → B → D → E
- Backup tunnel: B → C → D (preprovisioned)
- Recovery = ~ 50 ms

*Actual Time Varies—Well Below 50 ms in Lab Tests, Can Also Be Higher
Use Case 1: Tactical TE Deployment

**Requirement:** Need to Handle Scattered Congestion Points in the Network

**Solution:** Deploy MPLS TE on Only Those Nodes that Face Congestion

- MPLS Traffic Engineering Tunnel Relieves Congestion Points
- Bulk of Traffic Flow e.g. Internet Download
- Oversubscribed Shortest Links
Use Case 2: 1-Hop Tunnel Deployment

**Requirement:** Need Protection Only — Minimize Packet Loss of Bandwidth in the Core

**Solution:** Deploy MPLS Fast Reroute for Less than 50ms Failover Time with 1-Hop Primary TE Tunnels and Backup Tunnel for Each Service Provider
MPLS TE Summary

- MPLS TE can be used to implement traffic engineering to enable enhanced network availability, utilization, and performance.
- Enhanced network availability can be implemented via MPLS TE Fast Re-Route (FRR):
  - Link, node, and path protection
  - Automatically route around failed links/nodes; like SONET APS
- Better network bandwidth utilization can be implemented via creation of MPLS TE tunnels using explicit routes:
  - Route on the non-shortest path
- MPLS TE can be used for capacity planning by creation of bandwidth-specific tunnels with explicit paths through the network:
  - Bandwidth management across links and end-to-end paths
MPLS Management

Technology Overview and Applications
MPLS Technology Framework

- MPLS management using SNMP MPLS MIB and MPLS OAM capabilities

- Layer-3 VPNs
- Layer-2 VPNs
- MPLS QoS
- MPLS TE
- MPLS OAM/MIBs
- MPLS Signaling and Forwarding

Network Infrastructure
What’s Needed for MPLS management?

- What’s needed beyond the basic MPLS CLI?
  
  CLI used for basic configuration and trouble shooting (show commands)

Traditional management tools:

- MIBs to provide management information for SNMP management applications (e.g., HPOV)
  
  MIB counters, Trap notifications, etc.

New management tools:

- MPLS OAM -> for reactive trouble shooting
  
  Ping and trace capabilities of MPLS label switched paths

- Automated MPLS OAM -> for proactive trouble shooting
  
  Automated LSP ping/trace via Auto IP SLA
MPLS Operations Lifecycle

- Build and plan the network
  - Capacity planning and resource monitoring

- Monitor the network
  - Node/link failure detection
  - May impact multiple services

- Provision new services and maintain existing services
  - Edge/service node configuration

- Monitor service
  - End-to-end monitoring
  - Linked to customer SLAs
## MPLS MIBs and OAM

<table>
<thead>
<tr>
<th>MPLS MIBs</th>
<th>Management Feature</th>
<th>Key Functionality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MPLS-LDP-STD-MIB</td>
<td>LDP session status Trap notifications</td>
</tr>
<tr>
<td></td>
<td>MPLS-L3VPN-STD-MIB</td>
<td>VRF max-route Trap notifications</td>
</tr>
<tr>
<td></td>
<td>MPLS-TE-STD-MIB</td>
<td>TE Tunnel status Trap notifications</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MPLS OAM</th>
<th>Management Feature</th>
<th>Key Functionality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MPLS LSP Ping/Trace for LDP-based LSPs</td>
<td>Validate end-to-end connectivity of LDP-signaled LSPs</td>
</tr>
<tr>
<td></td>
<td>MPLS LSP Ping/Trace for TE tunnels</td>
<td>Validate end-to-end connectivity of TE tunnels</td>
</tr>
<tr>
<td></td>
<td>LSP Multipath (ECMP) Tree Trace</td>
<td>Discovery of all available equal cost LSP paths between PEs</td>
</tr>
</tbody>
</table>
LDP Event Monitoring Using LDP Traps

Interface Shutdown (E1/0 on PE1)

```
Time = t: Received SNMPv2c Trap from pe1:
sysUpTimeInstance = 8159606
snmpTrapOID.0 = mplsLdpSessionDown
mplsLdpSessionState.<index> = nonexistent(1)
```

```
Time = t+1: Received SNMPv2c Trap from pe1:
sysUpTimeInstance = 8159606
snmpTrapOID.0 = mplsLdpSessionDown
mplsLdpSessionState.<index> = nonexistent(1)
```

```
Time = t+2: Received SNMPv2c Trap from p01:
sysUpTimeInstance = 8160579
snmpTrapOID.0 = mplsLdpSessionDown
mplsLdpSessionState.<index> = nonexistent(1)
```

LDP Session Down (PE1 – P01)

```
Time = t: Received SNMPv2c Trap from pe1:
sysUpTimeInstance = 8159606
snmpTrapOID.0 = mplsLdpSessionDown
mplsLdpSessionState.<index> = nonexistent(1)
```

```
Time = t+1: Received SNMPv2c Trap from pe1:
sysUpTimeInstance = 8159606
snmpTrapOID.0 = mplsLdpSessionDown
mplsLdpSessionState.<index> = nonexistent(1)
```

```
Time = t: Received SNMPv2c Trap from p01:
sysUpTimeInstance = 8160579
snmpTrapOID.0 = mplsLdpSessionDown
mplsLdpSessionState.<index> = nonexistent(1)
```

IfIndex.5 = 5

Interface goes down

LDP session goes down
Validation of PE-PE MPLS Connectivity

- Connectivity of LSP path(s) between PE routers can be validated using LSP ping (ping mpls command via CLI)

pe1>ping mpls ipv4 10.1.2.249/32
Sending 5, 100-byte MPLS Echos to 10.1.2.249/32,
    timeout is 2 seconds, send interval is 0 msec:

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
'L' - labeled output interface, 'B' - unlabeled output interface,
'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
'M' - malformed request, 'm' - unsupported tlvs,
'N' - no label entry, 'P' - no rx intf label prot,
'R' - transit router, 'I' - unknown upstream index,
'X' - unknown return code, 'x' - return code 0

Type escape sequence to abort.

!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 284/294/300 ms
Automated MPLS OAM

- Automatic MPLS OAM probes between PE routers
  - Automatic discovery of PE targets via BGP next-hop discovery
  - Automatic discovery of all available LSP paths for PE targets via LSP multi-path trace
  - Scheduled LSP pings to verify LSP path connectivity
  - 3 consecutive LSP ping failures result in SNMP Trap notification
MPLS Management Summary

- MPLS management operations include MPLS node and service configuration, and monitoring
- In addition to CLI, SNMP MIBs and OAM capabilities are available for MPLS management
- MPLS MIBs provide LDP, VPN, and TE management information, which can be collected by SNMP tools
  - MIB counters, Trap notifications
- Advanced MPLS management capabilities can be implemented via MPLS OAM
  - LSP path discovery and connectivity validation
  - Proactive monitoring via automated MPLS OAM
Summary

Final Notes and Wrap Up
Summary and Key Takeaways

- It’s all about labels …
  - Label-based forwarding and IP protocol extensions for label exchange
  - Best of both worlds … L2-type forwarding and L3 control plane

- Key application of MPLS is to implement VPN services
  - Secure and scalable layer 2 and 3 VPN connectivity

- MPLS supports advanced traffic engineering capabilities
  - QoS, bandwidth control, and failure protection

- MPLS is a mature technology with widespread deployments
  - Both SP and enterprise networks

- Two types of MPLS users
  - Indirect (Subscriber): MPLS used as transport for subscribed service
  - Direct (DIY): MPLS implemented in (own) SP or enterprise network
### MPLS Applications

<table>
<thead>
<tr>
<th>Key Features</th>
<th>Service Providers</th>
<th>Enterprise Data Center</th>
<th>Data center interconnects</th>
<th>EWAN Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2/L3VPN’s</td>
<td></td>
<td></td>
<td></td>
<td>VPN’s / VRF’s</td>
</tr>
<tr>
<td>TE/FRR</td>
<td></td>
<td></td>
<td></td>
<td>VRF Aware Security</td>
</tr>
<tr>
<td>QoS</td>
<td></td>
<td></td>
<td></td>
<td>High Availability</td>
</tr>
<tr>
<td>High Availability</td>
<td></td>
<td></td>
<td></td>
<td>High Availability</td>
</tr>
<tr>
<td>Hosted Data centers</td>
<td></td>
<td></td>
<td></td>
<td>VPN’s / VRF’s</td>
</tr>
<tr>
<td>Data center interconnect</td>
<td></td>
<td></td>
<td></td>
<td>VRF Aware Security</td>
</tr>
<tr>
<td>Segmentation for IT Mergers, Acquisitions, spinoffs</td>
<td></td>
<td></td>
<td></td>
<td>Internet Access</td>
</tr>
<tr>
<td>Departmental segmentation</td>
<td></td>
<td></td>
<td></td>
<td>Branch Connectivity</td>
</tr>
<tr>
<td>Service multiplexing Security Mergers, Acquisitions, spinoffs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disaster Recovery Vmotion support</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branch Interconnects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Applications

- **Network Consolidation** – Merging Multiple parallel network into a shared infrastructure
- **Network segmentation** – By user groups or business function
- **Service and policy centralization** – Security policies and appliances at a central location
- **New applications readiness** – Converged multi-service network
- **Increased network security** – User groups segmentation with VPNs

© 2010 Cisco and/or its affiliates. All rights reserved.
Consider MPLS When …

- There’s a need for network segmentation
  - Segmented connectivity for specific locations, users, applications, etc.
  - Full-mesh and hub-and-spoke connectivity

- There’s a need for network realignment/migration
  - Consolidation of (multiple) legacy networks
  - Staged network consolidation after company merger/acquisition

- There’s a need for optimized network availability and performance
  - Node/link protection, pro-active connectivity validation
  - Bandwidth traffic engineering and QoS traffic prioritization
Q and A