NETWORKERS 2004

# TROUBLESHOOTING MPLS NETWORKS

**SESSION RST-3605**

**PRESENTER: YUSUF HASSAN (YHASAN@CISCO.COM)**

**CONTENT DEVELOPER: RAJIV ASATI (RAJIVA@CISCO.COM)**





## Agenda

Cisco.com

- **Prerequisites**
- **MPLS Troubleshooting**
- **Conclusion**

## Prerequisites

Cisco.com

- IGP Routing protocols
- MPLS in general
- Other Sessions

| | |
|---|---|
| Troubleshooting MPLS VPN | RST-3606 |
| Deploying MPLS VPN | RST-2602 |
| Intro to MPLS | RST-1601 |
| Troubleshooting BGP | RST-3303 |





## Before We Begin…

Cisco.com

- This presentation will NOT cover MPLS VPN, TE, AToM, etc.
- What can you expect to get from this presentation?
  - Overview of MPLS/LDP/
  - Learn how to use show commands and debugs to troubleshoot MPLS problems
  - Go through various real-life troubleshooting examples

## Agenda

Cisco.com

- Prerequisites
- MPLS Troubleshooting
- Conclusion





## MPLS Troubleshooting

Cisco.com

### MPLS Troubleshooting from Two Different Perspectives

- Control plane
  - Involves LDP, LIB, etc.
- Forwarding Plane
  - Involves FIB, LFIB, etc.

## MPLS Control Plane

Cisco.com

- LDP is one of the primary ways, but not the only one, to enable MPLS on an interface; other ways are
  - TDP
  - BGP+Label
  - RSVP
- Enabling MPLS means—the ability to send/receive MPLS packets on an interface





## MPLS Control Plane

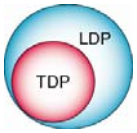
Cisco.com

### This Section Is All About LDP (and Its Related Components)

- LDP vs. TDP
- LDP (Discovery, Session Setup, Label Xchange)
- RIB/FIB/LIB/LFIB Relationship
- Troubleshooting Tips
- Troubleshooting Case Studies

## MPLS Control Plane: LDP vs. TDP

Cisco.com




- LDP is quite similar to TDP
- LDP is standardized by IETF
- LDP has more features such as abort, MD5 authentication, notification, backoff logic, etc.
- TDP is the default on Cisco routers
- LDP is the default with this global config—

  *"mpls ldp protocol ldp"*





## MPLS Control Plane

Cisco.com

- Control Plane
  - LDP vs. TDP
  - LDP (Discovery, Session Setup, Label Xchange)
  - RIB/FIB/LIB/LFIB Relationship
  - Troubleshooting Tips
  - Troubleshooting Case Studies
- Forwarding Plane

## MPLS Control Plane: LDP

Cisco.com

- LDP/TDP operates in three steps—
  - Neighbor Discovery
  - Session establishment
  - Label Distribution/exchange
- Once labels are exchanged, LIB is built
- LIB and FIB together helps to build LFIB

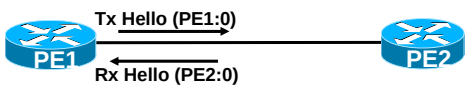




## MPLS Control Plane: TDP (i)

Cisco.com

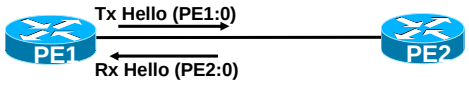
- TDP Neighbors are discovered via TDP Hellos (like most of the routing protocols)
- TDP Hellos are sent to 255.255.255.255
- TDP hellos are sent to UDP port = 711
- TDP hellos are sent only after “mpls ip” is configured on an interface



## MPLS Control Plane: LDP (i)

Cisco.com

- LDP Neighbors are discovered via LDP Hellos (like most of the routing protocols)
- LDP Hellos are sent to 224.0.0.2
- LDP hellos are sent to UDP port = 646
- LDP hellos are sent only after both “mpls ip” and “mpls label protocol ldp” are configured on an interface **




** If LDP is the global default, then interface-level LDP is not needed.





## MPLS Control Plane: LDP (i)

Cisco.com

- LDP_ID should be hardcoded via—

  “mpls ldp router-ID <*interface*>”

- The above won’t do any good unless

  <*interface*> is UP when LDP gets started

  Existing LDP_ID (usually an interface) is shut/unshut

- Following avoids both shortcomings—

  “mpls ldp router-ID <*interface*> force”

## MPLS Control Plane: LDP (i)

Cisco.com

- Use the same Loopback0 as the router-ID for LDP, IGP, BGP, etc.
- Assign an IP address to the Loopback0 from the separate IP address subnet (or space)
- Avoid the IGP summarization of prefixes that correspond to the router-ids





## MPLS Control Plane: LDP (i)

Cisco.com

- "sh mpls ldp discovery [detail]"

  Must show xmit/recv on LDP enabled interface

```
PE1#sh mpls ldp discovery
 Local LDP Identifier:
    10.13.1.61:0
    Discovery Sources:
    Interfaces:
        Ethernet0/0 (ldp): xmit/recv
            LDP Id: 10.13.1.101:0
        Ethernet1/0 (ldp): xmit/recv
            LDP Id: 10.13.1.101:0
PE1#
```

Local LDP_ID

Eth0/0 is configured with LDP

Xmited and Recvd Hellos on that interface

Discovered Neighbors' LDP_ID

- "debug mpls ldp transport connections"

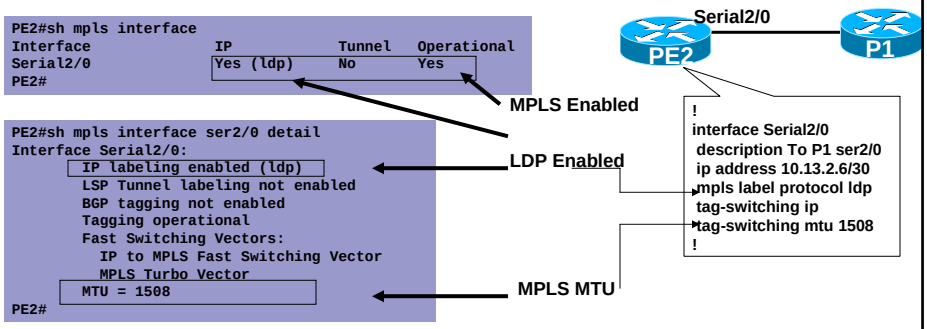
  Should give information regarding whether the HELLOS are advertised/received

# MPLS Control Plane: LDP (i)

Cisco.com

- "sh mpls interface [detail]"

  Lists whether MPLS is enabled and the application that enabled MPLS on the interface



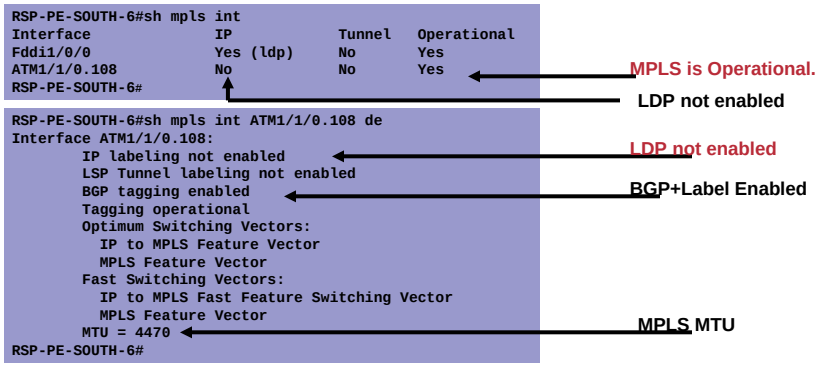





# MPLS Control Plane: LDP (i)

Cisco.com

- This slide is to show that BGPipv4+label (or MP-eBGP) is another application that can enable MPLS; WHAT'S DIFFERENT HERE—



## MPLS Control Plane: LDP (ii)

Cisco.com

- After discovering each other, they want to get cozy and establish the session.

  (Even routers have the dating concept) ☺

- LDP INITIALIZATION, KEEPALIVE and ADDRESS messages are exchanged to establish LDP session
- LSR_ID (Transport address) MUST be IP reachable

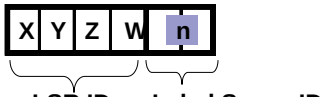





## MPLS Control Plane: LDP (ii)

Cisco.com

**LDP_ID =>**




**LSR_ID**

The LSR_ID is a four byte number that identifies a specific LSR. It is derived from an interface on the LSR. By default, it is the highest IP address, or highest IP address of a loopback– if it's available.

**Label_Space_Id**

A two byte number that identifies a specific label space on the LSR. 0x00 is reserved for the platform label space (i.e. frame-mode MPLS). Non-zero refers to the interface label space (i.e. cell-mode MPLS).

## MPLS Control Plane: LDP (ii)

Cisco.com

- LDP session is a TCP session (port = 646)
- Multiple links between two routers still mean single LDP session

```
PE1#sh mpls ldp neighbor
    Peer LDP Ident: 10.13.1.101:0; Local LDP Ident 10.13.1.61:0
        TCP connection: 10.13.1.101.11031 - 10.13.1.61.646
        State: Oper; Msgs sent/rcvd: 58/60; Downstream
        Up time: 00:39:27
        LDP discovery sources:
          Ethernet0/0, Src IP addr: 10.13.1.5
          Ethernet1/0, Src IP addr: 10.13.1.9
        Addresses bound to peer LDP Ident:
          10.13.1.9       10.13.1.5       10.13.2.5     10.13.1.101
PE1#
```

LDP_ID

Unsolicited Label Distribution*

Interfaces on Which Peer Is Discovered

Peer's Connected Int

```
PE1#sh tcp brief| i 646
43ABB020  10.13.1.101.11031       10.13.1.61.646         ESTAB
PE1#
```





## MPLS Control Plane: LDP (ii)

Cisco.com

### Relevant LDP Session Commands/Debugs:

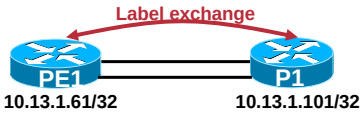
- "sh mpls ldp neighbor [neighbor]"
  - Shows LDP neighbor and relevant info
- "sh mpls ldp neighbor [interface]"
  - LDP neighbors discovered over this interface
- "Debug mpls ldp session io|state"
  - Useful when the session doesn't come up
- "Debug mpls ldp messages sent|receive"
  - Shows all the LDP messages sent or received

# MPLS Control Plane: LDP (iii)

Cisco.com

- Now, the LDP session is established, LDP neighbors start exchanging label bindings via LABEL MAPPING message (after the Keepalive gets exchanged)
- Label binding => prefix + Label
- Label bindings are stored in the LIB

  LIB => Label Information Base








# MPLS Control Plane: LDP (iii)

Cisco.com

- LIB entry can be verified with the following

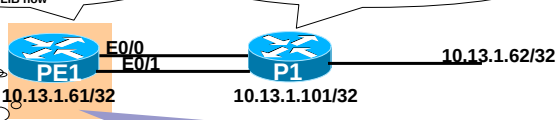
```
PE1#sh mpls ip bindings 10.13.1.62 32
 10.13.1.62/32
        in label:     20
        out label:    2001       lsr: 10.13.1.101:0
PE1#
```

Local binding
Remote binding

Ok. I hear you 10.13.1.101:0. I have the binding from you in my LIB now

This is 10.13.1.101:0. Use label 2001 to reach 10.13.1.62/32

But whether I use your binding or not will be dictated by RIB entry




Oh ok. Per RIB, 10.13.1.101 is the next-hop for 10.13.1.62/32. I have to use label 2001 in LFIB

```
PE1#sh mpls forwarding 10.13.1.62
Local  Outgoing    Prefix            Bytes tag  Outgoing    Next Hop
tag    tag or VC   or Tunnel Id      switched   interface
20     2001        10.13.1.62/32     0          Et0/0       10.13.1.5
       2001        10.13.1.62/32     0          Et1/0       10.13.1.9
PE1#
```

## MPLS Control Plane: LDP (iii)

Cisco.com

- "**sh mpls ip binding detail**"

  **Lists all prefixes with labels and LDP neighbors**

- "**sh mpls ip binding *<prefix>* <mask> det**"

  **Lists ACLs (if any), *prefix* bindings, and LDP neighbors**
  **Notice "Advertised to:" field**

- "**sh mpls ip binding advertisement-acls**"

  **Lists LDP filter, if there is any, on the first line. Prefixes followed by "Advert acl(s):" are advertised via LDP, others are not**





## MPLS Control Plane

Cisco.com

- **Control Plane**

  **LDP vs. TDP**

  **LDP (Discovery, Session Setup, Label Xchange)**

  **RIB/FIB/LIB/LFIB Relationship**

  **Troubleshooting Tips**

  **Troubleshooting Case Studies**

- **Forwarding Plane**

## RIB/FIB/LIB/LFIB

Cisco.com

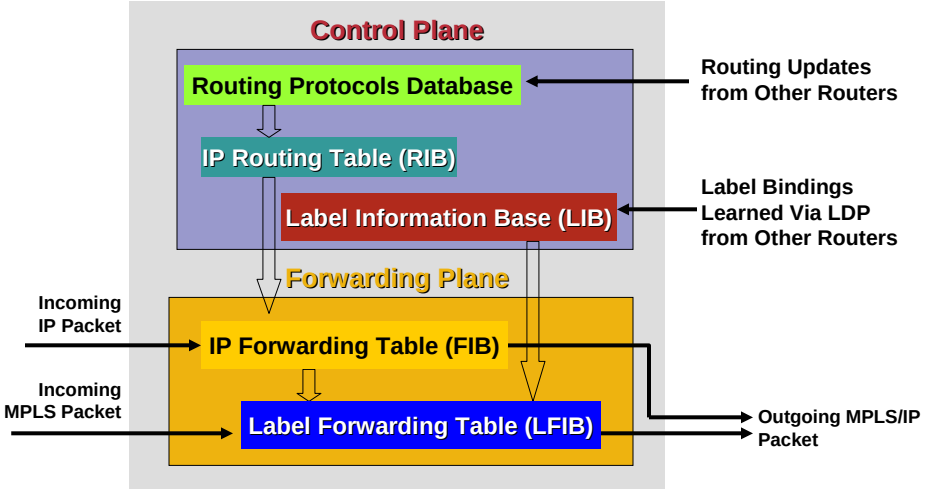
- **RIB** is the Routing Information Base that is analogous to the ip routing table
- **FIB** aka CEF is Forwarding information base that is derived from the ip routing table
- **LIB** is Label Information Base that contains all the label bindings learned via LDP
- **LFIB** is Label Forwarding Information Base that is derived from FIB entries and corresponding LIB entries
- Let's go through the pictorial view—





## MPLS Control Plane: RIB/FIB/LIB/LFIB

Cisco.com




**Population of RIB/FIB/LIB/LFIB in a LSR**

## MPLS Control Plane: Debugs

Cisco.com

### Be Careful on the Production Routers

- **"debug mpls ldp advertisements"**
  - **Useful to see label bindings that are advertised**
- **"debug mpls ldp binding"**
  - **Useful to see label bindings that are received**
- **"debug mpls ldp message sent|received"**
  - **Useful for the protocol understanding purposes**





## MPLS Control Plane

Cisco.com

- **Control Plane**
  - **LDP vs. TDP**
  - **LDP (Discovery, Session Setup, Label Xchange)**
  - **RIB/FIB/LIB/LFIB relationship**
  - **Troubleshooting Tips**
  - **Troubleshooting Case Studies**
- **Forwarding Plane**

## MPLS Control Plane: Troubleshooting Tips

Cisco.com

1. Check for same label protocol to be configured on both sides of the interface

   "Sh mpls ldp discovery | inc ldp|tdp"

2. Check whether correct local LSR_ID is used on both LSRs (sh mpls ldp disc)

   "sh mpls ldp discovery"—2nd l

```
PE1#sh mpls ldp disc | i ldp|tdp
        Ethernet0/0 (ldp): xmit/recv
PE1#
```

3. Don't assume that the neighbor discovery means everything is good

```
PE1#sh mpls ldp disco
 Local LDP Identifier:
    10.13.1.61:0
```

RST-3605
9808_06_2004_X2





## MPLS Control Plane: Troubleshooting Tips

Cisco.com

4. Check IP reachability to remote LSR_ID on both LSRs

   "ping <lsr_id>"

```
PE1#ping 10.13.1.101 source 10.13.1.61

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.13.1.101, timeout is 2
seconds:
Packet sent with a source address of 10.13.1.61
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max =
32/49/72 ms
PE1#
```

5. Check for ACL (

6. Untagged outgo
   loopbacks is almost always alarming

7. Check the label binding for a prefix on both LSRs

   "sh mpls ldp bind <prefix> <mask>"

```
PE1#sh mpls ldp bind 10.13.1.62 32
  tib entry: 10.13.1.62/32, rev 16
        local binding:  tag: 17
        remote binding: tsr: 10.13.1.101:0, tag: 2001
PE1#
```

RST-3605
9808_06_2004_X2

## MPLS Control Plane: Troubleshooting Tips

Cisco.com

8. **Make sure the LDP filtering (if configured) is correctly setup via ACL**

   **"sh mpls ip bind advertisement-acl | inc Prefix"**

9. **Good practice is to configure the Loopback0 as the router-ID for LDP**

   **"mpls ldp router-id loopback0 force"**

Configur ed filter

```
VXR-PE-WEST-5#sh mpls ip bind advertisement-acls  | i Prefix
    Prefix acl = raj-ti-avoid-mgmt; Peer acl = raj-ti-p
    Advert acl(s): Prefix acl raj-ti-avoid-mgmt; Peer acl raj-ti-p
VXR-PE-WEST-5#
```

Only one "Advert acl" appearance. Hence, only one prefix is getting advertised via LDP (rest are filtered).





## MPLS Control Plane

Cisco.com

- **Control Plane**
  - **LDP vs. TDP**
  - **LDP (Discovery, Session Setup, Label Xchange)**
  - **RIB/FIB/LIB/LFIB Relationship**
  - **Troubleshooting Tips**
  - **Troubleshooting Case Studies**
- **Forwarding Plane**

# MPLS Control Plane: LDP Troubles

Cisco.com

- Let's do some REAL Troubleshooting now ☺

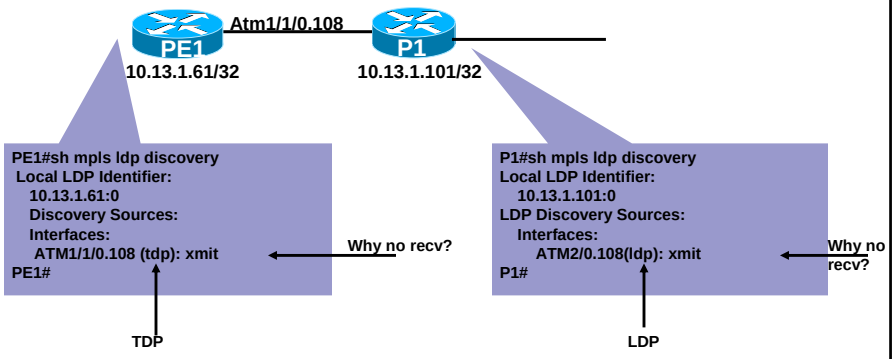




# MPLS Control Plane: LDP Problems

Cisco.com

## Prob #1—Session Establishment (Protocol Mismatch)



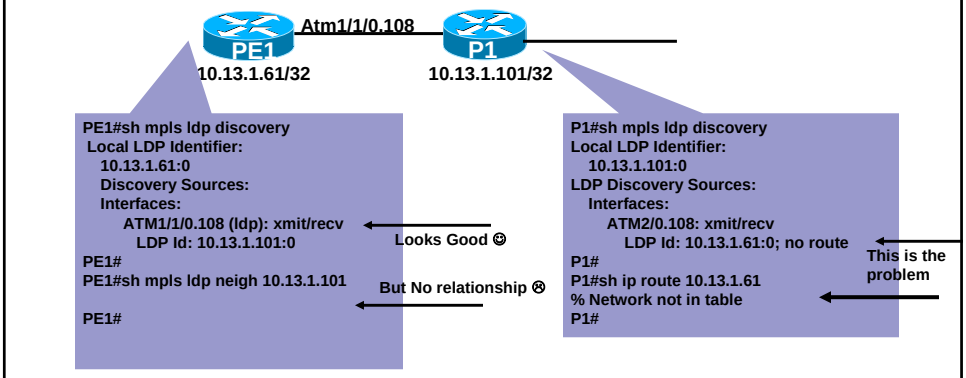

## TIP—Check for the Protocol Mismatch and Fix It

```
PE1(config)#int atm1/1/0.108
PE1(config-if)#mpls label protocol ldp
```

# MPLS Control Plane: LDP Troubles

Cisco.com

## Prob #2—Session Establishment (No Route to Peer)




**TIP—Check for IP reachability to LDP_ID; Fix It by Letting PE1 Advertise 10.13.1.61/32 via IGP to P1**

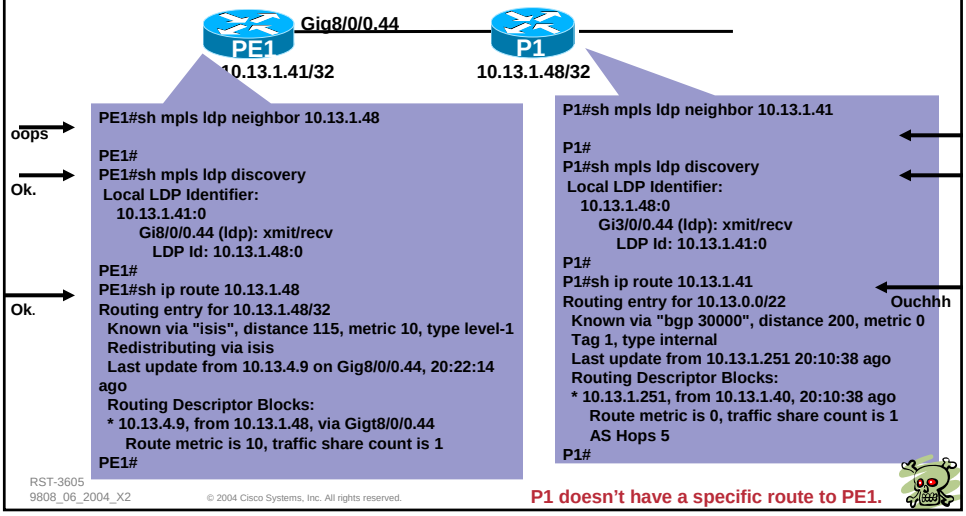




# MPLS Control Plane: LDP Troubles

Cisco.com

## Prob #3—Session Establishment (No Specific Route)







P1 doesn't have a specific route to PE1.

# MPLS Control Plane: LDP Troubles

Cisco.com

## Prob #3—Session Establishment (Cont.)


```
PE1#ping 10.13.1.48
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.13.1.48, timeout is 2
seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max =
1/1/4 ms
PE1#
```

```
P1#ping 10.13.1.41
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.13.1.41,
timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
P1#
```

Eeeekks !! It is an IP problem.

**TIP—Check for IP connectivity first. Unless Layer3 is up, Layer4 (TCP session for LDP) won't come up.**





# MPLS Control Plane: LDP Troubles

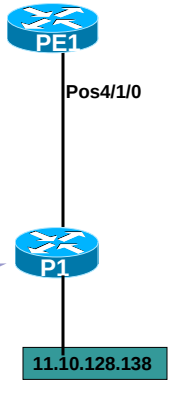
Cisco.com

## Prob #4—"Untagged" Problem

```
PE1#sh tag for 11.10.128.138
Local  Outgoing    Prefix          Bytes tag  Outgoing   Next Hop
tag    tag or VC   or Tunnel Id     switched  interface
16     Untagged    11.10.128.138/32  0        PO4/1/0    point2point
PE1#
```

```
PE1#sh mpls ldp  bind 11.10.128.138 32
 tib entry: 11.10.128.138/32, rev 14
      local binding:  tag: 16
PE1#
```

```
P1#sh mpls ldp bind 11.10.128.138 32
 tib entry: 11.10.128.138/32, rev 4849(no route)
      local binding:  tag: 630
      remote binding: tsr: 10.13.1.54:0,   tag: 16
      remote binding: tsr: 11.10.65.12:0, tag: 48
P1#
P1#sh ip route 11.10.128.138
Routing entry for 11.10.0.0/16
 Known via "isis", distance 115, metric 44, type level-2
 Redistributing via isis
 Last update from 11.10.65.13 on POS0/0, 1d00h ago
 Routing Descriptor Blocks:
 * 11.10.65.13, from 11.10.128.31, via POS0/0
     Route metric is 44, traffic share count is 1
P1#
```




Untagged ?

No remote binding. Huh

No route

But there is a RIB entry. Let's check FIB entry -

# MPLS Control Plane: LDP Troubles

Cisco.com

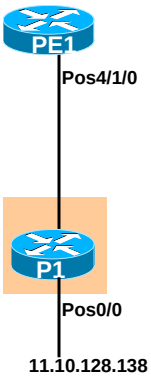
## Prob #4—"No Route" Problem (Cont.)

```
P1#sh ip cef 11.10.128.138
11.10.0.0/16, version 142, cached adjacency to POS0/0
0 packets, 0 bytes
  tag information set
    local tag: 307
    fast tag rewrite with PO0/0, point2point, tags imposed {48}
  via 11.10.65.13, POS0/0, 0 dependencies
    next hop 11.10.65.13, POS0/0
    unresolved        <<<<<<<<<<<<<<<<<<<<<<<<<<<
    valid cached adjacency
    tag rewrite with PO0/0, point2point, tags imposed {48}
P1#
```

FIB's local label is different from that of LIB

Unresolved ?

```
P1#clear ip route 11.10.128.138
P1#sh mpls ldp bind 11.10.128.138 32
  tib entry: 11.10.128.138/32, rev 4849
        local binding:  tag: 307
        remote binding: tsr: 10.13.1.54:0, tag: 16
        remote binding: tsr: 11.10.65.20:0,tag:48
P1#
```




**TIP—If Local Label for a Prefix Is Not Same in FIB and LIB, Then Issue "clear ip route <prefix>" to fix**





# MPLS Control Plane: LDP Troubles

Cisco.com

## Prob #5—LFIB Entry Disappears

- No LFIB entry
- This might occur if the RIB owner for an IPv4 routes changes from IGP to BGP
- LDP doesn't allocate labels for the BGP learned IPv4 routes
- Notice the absence of local binding in LIB for that route

# MPLS Control Plane: LDP Troubles

Cisco.com

```
7206-PE-SOUTH-1#sh mpls ldp bind 4.4.0.0 24
 tib entry: 4.4.0.0/24, rev 152
       remote binding: tsr: 10.13.1.69:0, tag: 213
       remote binding: tsr: 10.13.1.68:0, tag: 212
7206-PE-SOUTH-1#
```

**No Local Binding**

```
7206-PE-SOUTH-1#sh ip route 4.4.0.0
Routing entry for 4.4.0.0/24
  Known via "bgp 30000", distance 200, metric 0
  Tag 1, type internal
  Redistributing via isis, ospf 1
  Last update from 10.13.1.251 5d17h ago
  Routing Descriptor Blocks:
  * 10.13.1.251, from 10.13.1.40, 5d17h ago
      Route metric is 0, traffic share count is 1
      AS Hops 5
      Route tag 1

7206-PE-SOUTH-1#
```

**Because it is a BGP learned prefix**

**LDP doesn't allocate labels for the BGP learned IPv4 routes.**





# MPLS Troubleshooting

Cisco.com

- **Control plane**
  - **Involves LDP, LIB, etc.**
- **Forwarding Plane**
  - **Involves FIB, LFIB, etc.**

## MPLS Forwarding Plane

Cisco.com

- With MPLS, the idea is to de-couple the forwarding from the IP
- The forwarding decision is based on the MPLS header, not the IP header
- The above is true once the packet is inside the MPLS network
- Forwarding is still based on the IP header at the edge where the packet first enters the MPLS network





## MPLS Forwarding Plane

Cisco.com

- Control Plane
- Forwarding Plane—
  - CEF's Role
  - New Ethertype, What Is a Label, Types of Labels
  - Forwarding Explained
  - Loadsharing
  - Fragmentation and MTU
  - Troubleshooting Tips
  - Troubleshooting Case Studies
  - LSP Ping/Traceroute

## MPLS Forwarding Plane: CEF's Role

Cisco.com

- CEF must be configured on all the routers in a MPLS network
- CEF takes care of the crucial "recursion" and "resolution" operations
- MPLS relies on CEF
- CEF is must for the MPLS





## MPLS Forwarding Plane

Cisco.com

- Control Plane
- Forwarding Plane
  - CEF's Role
  - New Ethertype, What Is a Label, Types of Labels
  - Forwarding Explained
  - Loadsharing
  - Fragmentation and MTU
  - Troubleshooting Tips
  - Troubleshooting Case Studies
  - LSP Ping/Traceroute

## MPLS Forwarding Plane: Ethertype

Cisco.com

- Ethertype 0x0800 refers to IP
- Ethertype 0x8847 refers to MPLS
- Based on the Ethertype, the packet is handed over to the appropriate processing engine in the router

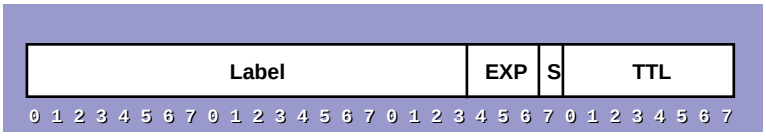




## MPLS Fwd Plane: What Is a Label

Cisco.com

- A LABEL is 4 bytes identifier, which is carried by the packet and used to identify a prefix

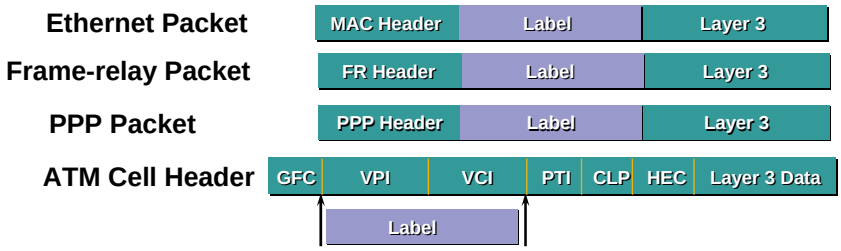



| | | |
|---|---|---|
| Label | = Actual Label, | = 20 bits |
| EXP/QoS | = Experimental bits, | = 3 bits |
| S | = End of Stack, | = 1 bit |
| TTL | = Time to Live, | = 8 bits |

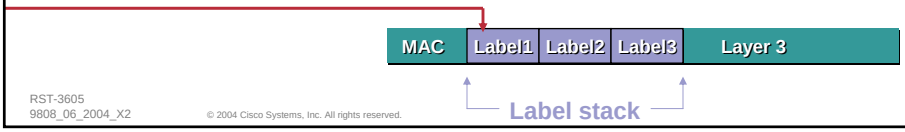
## MPLS Fwd Plane: Where Is Label?

Cisco.com

- Typically resides between L2 and L3 header




- Routers always makes forwarding-decision based on the topmost label i.e. label1 below-








## MPLS Forwarding Plane: Outgoing Labels

Cisco.com

```
RSP-PE-SOUTH-5#sh mpls forwarding 10.13.1.11
Local  Outgoing    Prefix           Bytes tag  Outgoing    Next Hop
tag    tag or VC   or Tunnel Id     switched   interface
59     46          10.13.1.11/32    0          Se10/0/0    point2point
RSP-PE-SOUTH-5#
```

- Outgoing label also conveys what treatment the packet is going to get; it could also be—
  - Pop —Pops the topmost label
  - Untagged —Untag the incoming MPLS packet
  - Aggregate —Untag and then do a FIB lookup
  - 0 —Nullify the top label (first 20bits)
- Label values 0-15 are reserved

# MPLS Forwarding Plane: Outgoing Labels

Cisco.com

```
PE1#sh mpls forwarding-table
Local  Outgoing    Prefix           Bytes tag  Outgoing    Next Hop
tag    tag or VC   or Tunnel Id     switched   interface
16     2002        10.13.1.22/32    0          Et0/0       10.13.1.5
       2002        10.13.1.22/32    0          Et1/0       10.13.1.9
17     2001        10.13.1.62/32    0          Et0/0       10.13.1.5
       2001        10.13.1.62/32    0          Et1/0       10.13.1.9
18     Pop tag     10.13.1.101/32   0          Et1/0       10.13.1.9
       Pop tag     10.13.1.101/32   0          Et0/0       10.13.1.5
19     Pop tag     10.13.2.4/30     0          Et1/0       10.13.1.9
       Pop tag     10.13.2.4/30     0          Et0/0       10.13.1.5
20     Untagged    5.5.5.5/32[V]    0          Se2/0       point2point
21     Pop tag     10.13.21.4/30    0          Et1/0       10.13.1.9
       Pop tag     10.13.21.4/30    0          Et0/0       10.13.1.5
22     Pop tag     10.13.22.4/30    0          Et1/0       10.13.1.9
       Pop tag     10.13.22.4/30    0          Et0/0       10.13.1.5
23     Aggregate   0.0.0.0/0[V]     0
24     Aggregate   200.1.61.4/30[V] 0
26     Untagged    30.30.30.1/32[V] 0          Se2/0       point2point
PE1#
```

V means it is a VPN prefix





# MPLS Forwarding Plane: Outgoing Labels

Cisco.com

- **Untagged**
  - Convert the incoming MPLS packet to an IP packet and forward it
- **Pop**
  - Pop the top label from the label stack present in an incoming MPLS packet and forward it as an MPLS packet.; if there was only one label in the stack, then forward it as an IP packet; SAME as imp-null label
- **Aggregate**
  - Convert the incoming MPLS packet to an IP packet and then do a FIB lookup for it to find out the outgoing interface
- **0 (zero)**
  - Same as exp-null label; simplify fills 0 in the first 20 bits of label; helps to preserve the EXP value of the top label

## MPLS Forwarding Plane

Cisco.com

- Control Plane
- Forwarding Plane
  - CEF's Role
  - New Ethertype, What Is a Label, Types of Labels
  - Forwarding Explained
  - Loadsharing
  - Fragmentation and MTU
  - Troubleshooting Tips
  - Troubleshooting Case Studies
  - LSP Ping/Traceroute





## MPLS Forwarding Plane

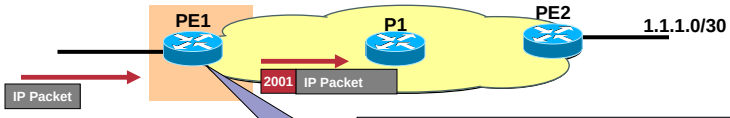
Cisco.com

- Three cases in the MPLS forwarding—
  1. Label Imposition—IP to MPLS conversion
  2. Label swapping—MPLS to MPLS
  3. Label disposition—MPLS to IP conversion
- So, depending upon the case, we need to check—
  1. FIB—For IP packets that get forwarded as MPLS
  2. LFIB—For MPLS packets that get fwded as MPLS
  3. LFIB—For MPLS packets that get fwded as IP

# MPLS Forwarding Plane

Cisco.com

## Case 1: IP Packets Get Forwarded as MPLS




```
PE1#sh ip cef 1.1.1.0
1.1.1.0/30, version 25, epoch 0, cached adjacency 10.13.1.5
0 packets, 0 bytes
  tag information set
    local tag: 20
    fast tag rewrite with Et0/0, 10.13.1.5, tags imposed: {2001}
  via 10.13.1.5, Ethernet0/0, 0 dependencies
    next hop 10.13.1.5, Ethernet0/0
    valid cached adjacency
    tag rewrite with Et0/0, 10.13.1.5, tags imposed: {2001}
PE1#
```

- PE1 does a FIB lookup for the incoming IP packet
- It imposes the label (if there is one)
- For troubleshooting, look at the FIB (not LFIB)

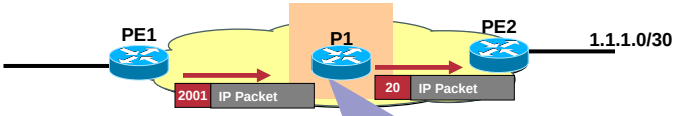




# MPLS Forwarding Plane

Cisco.com

## Case 2: MPLS Packets Get Forwarded as MPLS




```
P1#sh mpls for 1.1.1.0
Local  Outgoing    Prefix           Bytes tag  Outgoing    Next Hop
tag    tag or VC   or Tunnel Id     switched   interface
2001   20          1.1.1.1.0/30     0          Se2/0       point2point
P1#
```

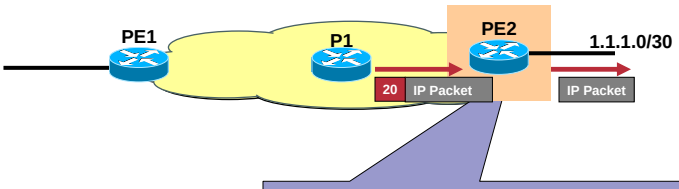
```
P1#sh mpls for 10.13.1.62
Local  Outgoing    Prefix           Bytes tag  Outgoing    Next Hop
tag    tag or VC   or Tunnel Id     switched   interface
2001   Pop tag     10.13.1.62/32    0          Se2/0       point2point
P1#
```

- P1 does the LFIB lookup for incoming MPLS packets
- P1 could swap (or dispose) the label
- For troubleshooting, look at the LFIB (not FIB)

# MPLS Forwarding Plane

Cisco.com

## Case 3: MPLS Packets Get Forwarded as IP




```
PE2#sh mpls for 1.1.1.0
Local  Outgoing    Prefix          Bytes tag  Outgoing    Next Hop
tag    tag or VC   or Tunnel Id    switched   interface
20     Untagged    1.1.1.1.0/30    0          Se2/0       point2point
PE2#
```

- Typically happen at the edge
- Could also happen at the PHP router
- For troubleshooting, look at the LFIB (not FIB)





# MPLS Forwarding Plane

Cisco.com

- Control Plane
- Forwarding Plane
  - CEF's Role
  - New Ethertype, What Is a Label, Types of Labels
  - Forwarding Explained
  - Loadsharing in MPLS
  - Fragmentation and MTU
  - Troubleshooting Tips
  - Troubleshooting Case Studies
  - LSP Ping/Traceroute

## MPLS Forwarding Plane: Loadsharing

Cisco.com

- Loadsharing (due to multiple paths to a prefix) in MPLS is no different from that of IP
- Hashing-algorithm is still the typical 'FIB based' i.e. per-dest loadsharing by default **
- So the below "show command" is still relevant
  - "Sh ip cef exact-route <source> <dest>" etc
- But the dest must be known in the FIB table, otherwise the command won't work
  - Won't work on P routers for the VPN prefixes





## MPLS Forwarding Plane

Cisco.com

- Control Plane
- Forwarding Plane
  - CEF's Role
  - New Ethertype, What Is a Label, Types of Labels
  - Forwarding Explained
  - Loadsharing
  - Fragmentation and MTU
  - Troubleshooting Tips
  - Troubleshooting Case Studies
  - LSP Ping/Traceroute

## MPLS Fwd Plane: Fragmentation

Cisco.com

- After the Layer 2 header is added to the IP packet, the resulting packet size shouldn't exceed the max packet size (IP MTU size) applicable; otherwise, packet will be fragmented
- MTU size needs to be tuned to avoid fragmentation in MPLS network
- MTU could be increased only for MPLS packets => MPLS MTU





## MPLS Fwd Plane: Fragmentation MTU Setting in MPLS

Cisco.com

- Two things to remember—
  1. DF bit of an incoming packet
  2. MTU size of an outgoing interface
- Label imposition(s) increases the packet size by 4 bytes/label, hence the outgoing packet size may exceed 'interface MTU' size, hence the need to tune MTU
- Q is: which MTU to tune in MPLS network? ☺

## Fragmentation
## MTU Setting in MPLS

Cisco.com

- Most of the interfaces (depending upon the hardware) support transmitting packets bigger than the “interface MTU” size
- “mpls mtu <bytes>” can be applied to an interface to change the MPLS MTU size on the interface
- MPLS MTU size is checked by the router

  While converting an IP packet into a labeled packet or transmitting a labeled packet





## Fragmentation
## MTU Setting in MPLS

Cisco.com

### Remember That:

- ‘mpls mtu <bytes>” command has no effect on “interface or IP” MTU size
- By default, MPLS MTU = interface MTU
- MPLS MTU setting doesn’t affect MTU handling for IP-to-IP packet switching

## Fragmentation
## MTU Setting in MPLS

Cisco.com

- If the label imposition makes the packet bigger than the MPLS MTU size of an outgoing interface, then
  - If the DF bit set, then discard the packet and send ICMP reply back (with code=4)
  - If the DF bit is not set, then fragment the IP packet (say, into 2 packets), and then impose the same label(s) on both the packets, and then transmit MPLS packets
- Fragmentation should be done at the edge itself





## MTU Setting in MPLS
## Configuring the MPLS MTU

Cisco.com

```
RSP-PE-WEST-4(config)#int fa1/1/0
RSP-PE-WEST-4(config-if)#mpls mtu 1508
RSP-PE-WEST-4(config-if)#^Z
RSP-PE-WEST-4#
```

## MTU Setting in MPLS
## Before setting the MPLS MTU

Cisco.com

- **Interface MTU is 1500 bytes (no change):**

```
RSP-PE-WEST-4#sh int fa1/1/0
  FastEthernet1/1/0 is up, line protocol is up
  Hardware is cyBus FastEthernet Interface, address is 0004.4e75.4828
  MTU 1500 bytes, BW 100000 Kbit, DLY 100 usec,
     ……..
RSP-PE-WEST-4#
```

- **MPLS MTU is 1508 bytes (changed):**

```
RSP-PE-WEST-4#sh mpls interface fa1/1/0 de
Interface FastEthernet1/1/0:
        IP tagging enabled
        TSP Tunnel tagging not enabled
       Tagging operational
        ……..
        MTU = 1508
RSP-PE-WEST-4#
```





## MPLS Fwd Plane: Show Commands

Cisco.com

- "**sh mpls forwarding**"

  Shows all LFIB entries (vpn, non-vpn, TE, etc.)

- "**sh mpls forwarding <prefix>**"

  LFIB lookup based on a prefix

- "**sh mpls forwarding label <label>**"

  LFIB lookup based on an incoming **label**

- "**sh mpls forwarding <prefix> detail**"

  Shows detailed info such as L2 encap, etc.

# MPLS Fwd Plane: Show Command

Cisco.com

```
RSP-PE-WEST-4#sh mpls for 10.13.1.11 detail
Local  Outgoing    Prefix            Bytes tag  Outgoing    Next Hop
tag    tag or VC   or Tunnel Id      switched   interface
45     51          10.13.1.11/32     0          Fa1/1/1     10.13.7.33
       MAC/Encaps=14/18, MRU=1500, Tag Stack{51}
       0003FD1C828100044E754829 00033000
       No output feature configured
    Per-packet load-sharing
RSP-PE-WEST-4#
```

Detail Is Optional

Only One Outgoing Label in the Label Stack

MRU – Max Receivable Unit. The received packet will be transmitted unfragmented on Fa1/1/1, if its size is not more than 1500B.

MAC header =0003FD1C828100044E754829

MPLS Ethtype= 0x8847

Label = 0x00033000

14/18 means that the L2 header is of 14 bytes, but L2+label header is 18 bytes (one label is 4 bytes)





# MPLS Fwd Plane: Debugs

Cisco.com

## Be Careful on the Production Routers

- **"Debug mpls lfib cef"**

  **Useful for seeing FIB and LFIB interaction when a label is missing for a prefix**

- **"debug mpls lfib struct"**

  **Shows changes in the LFIB structures when label is allocated/deallocated**

## MPLS Forwarding Plane

Cisco.com

- Control Plane
- Forwarding Plane
  - CEF's Role
  - New Ethertype, What Is a Label, Types of Labels
  - Forwarding Explained
  - Loadsharing
  - Fragmentation and MTU
  - Troubleshooting Tips
  - Troubleshooting Case Studies
  - LSP Ping/Traceroute





## MPLS Fwd Plane: Troubleshooting Tips

Cisco.com

1. If PXF-based platform, then check the PXF1
2. On distributed platforms, check the FIB/LFIB entries on the LC

3. On distributed platforms that have HW-based forwarding, check the FIB/LFIB on specific HW i.e. PSA (E2), Alpha(E3) on GSR, etc.

   Sh ip psa-cef, sh tag psa-tag, sh ip alpha-cef etc

[1] Not all PXF based platform support MPLS; they punt the MPLS packets to the CEF path.

## MPLS Fwd Plane: Troubleshooting Tips

Cisco.com

4. **Label imposition** is always done using **FIB**
5. **Label swapping and disposition** is always done using **LFIB**
6. Increase the MPLS MTU to accommodate the largest packet payload size

7. Make sure that baby giant/jumbo is enabled on the Ethernet switches





## MPLS Fwd Plane: Troubleshooting Tips

Cisco.com

8. Check that MPLS enabled interface has “TAG” adjacency via

   “sh adjacency <interface>”

9. Check that the LFIB’s outgoing label is same as the incoming label in neighbor’s LFIB
10. Check the LSP via traceroute that shows labels used by each router in the path **

    “traceroute <prefix>”

## MPLS Forwarding Plane: TAG adj

Cisco.com

11. Make sure that the interface has the “tag” adjacency along with “IP” adj, otherwise MPLS packets will not get switched on that interface

```
PE1#sh adjacency e0/0 de
Protocol Interface                 Address
TAG      Ethernet0/0               10.13.1.5(6)
                                   0 packets, 0 bytes
                                   AABBCC006500AABBCC0001008847
                                   mpls adj   never
                                   Epoch: 0
IP       Ethernet0/0               10.13.1.5(35)
                                   0 packets, 0 bytes
                                   AABBCC006500AABBCC0001000800
                                   ARP        03:46:13
                                   Epoch: 0
PE1#
```

L2 header for MPLS

L2 header for IP





## MPLS Forwarding Plane

Cisco.com

- Control Plane
- Forwarding Plane
  - CEF’s Role
  - New Ethertype, What Is a Label, Types of Labels
  - Forwarding Explained
  - Loadsharing
  - Fragmentation and MTU
  - Troubleshooting Tips
  - Troubleshooting Case Studies

## MPLS Fwd Plane: Troubles

Cisco.com

- Let's do some real trouble(shooting) ☺





## MPLS Fwd Plane: Troubles and Shooting

Cisco.com

### Prob #1—No Entries in LFIB

```
P1#sh mpls forwarding-table 10.13.1.61
Tag switching is not operational.
CEF or tag switching has not been enabled.
Local  Outgoing    Prefix            Bytes tag  Outgoing   Next Hop
tag    tag or VC   or Tunnel Id      switched   interface
P1#
```

```
P1#sh mpls ip binding
 10.13.1.61/32
        out label:    imp-null  lsr: 10.13.1.61:0
        out label:    21        lsr: 10.13.1.62:0
  10.13.1.62/32
        out label:    imp-null  lsr: 10.13.1.62:0
        out label:    17        lsr: 10.13.1.61:0
  10.13.1.101/32
        out label:    19        lsr: 10.13.1.62:0
        out label:    18        lsr: 10.13.1.61:0
  10.13.2.4/30
        out label:    imp-null  lsr: 10.13.1.62:0
        out label:    19        lsr: 10.13.1.61:0
P1#
```

```
P1#sh ip cef
%CEF not running
Prefix              Next Hop            Interface
P1#
```

### TIP—Enable CEF; It Is Must for MPLS

# MPLS Fwd Plane: Troubles and Shooting

Cisco.com

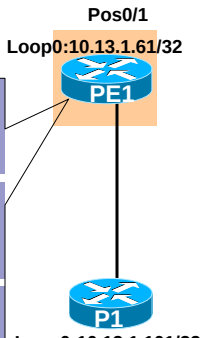
## Prob #2—"Untagged" Problem

- LDP session is UP;LIB has correct binding; but LFIB has "Untagged" ☹

```
PE1#sh mpls for 10.13.1.101
Local  Outgoing    Prefix           Bytes tag  Outgoing   Next Hop
tag    tag or VC   or Tunnel Id     switched   interface
20     Untagged    10.13.1.101/32   0          PO0/1      point2point
PE1#
```

```
PE1#sh mpls ip bind 10.13.1.101 32
  10.13.1.101/32
        in label:     20
        out label:    imp-null  lsr: 10.13.1.101:0
PE1#
```

```
PE1#sh adjacency pos0/1
Protocol Interface                 Address
TAG      POS0/1                    point2point(7) (incomplete)  <<=====Oops
IP       POS0/1                    point2point(39)
PE1#
```




**TAG ADJ for pos0/1 Is Incomplete; No Good.**





# MPLS Fwd Plane: Troubles and Shooting

Cisco.com

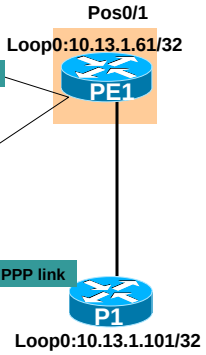
- Adj is incomplete; check the interface

```
PE1#sh mpls for 10.13.1.101 detail
Local  Outgoing    Prefix           Bytes tag  Outgoing   Next Hop
tag    tag or VC   or Tunnel Id     switched   interface
12318  Untagged    10.13.1.101/32   0          PO0/1      point2point
       MAC/Encaps=0/0, MRU=4474, Tag Sta
       No output feature configured
    Per-packet load-sharing, slots: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
PE1#
```

<<======= Another hint- why MAC/Encap is 0/0?

```
PE1#sh int pos0/1
POS0/1 is up, line protocol is up
  Hardware is Packet over SONET
  Description: OC48 to Redback
  Internet address is 10.1.17.1/24
  MTU 4470 bytes, BW 2488000 Kbit, DLY 100 usec, rely 255/255, load 1/255
  Encapsulation PPP, crc 32, loopback not set
  Keepalive not set
  Scramble disabled
  LCP Open
  Listen: TAGCP, CDPCP
  Open: IPCP
  Last input 00:00:01, output 00:00:03, output hang never
....
PE1#
```

<<========== TAGCP should also be in the Open state on PPP link



# MPLS Fwd Plane: Troubles and Shooting

Cisco.com

```
PE1#deb mpls adj
PE1#deb mpls lfib enc
PE1#
01:43:19: LFIB: finish res:inc tag=12318,outg=Imp_null,next_hop=0.0.0.0,POS0/1
01:43:19: LFIB: get ip adj: addr=0.0.0.0,is_p2p=1,fibidb=POS0/1,linktype=7
01:43:19: LFIB: get tag adj: addr=0.0.0.0,is_p2p=1,fibidb=POS0/1,linktype=90 INCOMPLETE
01:43:19: TAG ADJ: check 0.0.0.0, POS0/1 (537CF240/537CEE80)
01:43:19: LFIB: get ip adj: addr=0.0.0.0,is_p2p=1,fibidb=POS0/1,linktype=7
01:43:19: LFIB: get tag adj: addr=0.0.0.0,is_p2p=1,fibidb=POS0/1,linktype=90
01:43:19: LFIB: encaps:zero encaps,enc=0,mac=0,tag_adj POS0/1,itag=12318
```

**TIP—If the Interface Doesn't Have "TAG" adj, Then the Label Will Not Get Installed in LFIB; Fix PPP in This Case.**





# MPLS Fwd Plane: Troubles and Shooting

Cisco.com

## Prob #2— "Recursive Rewrite" Problem

- **If you ever see "Recursive rewrite via…" in the "sh ip cef .." output, then it might indicate a problem.**

```
2611-CE-30#sh ip cef 10.13.1.74
10.13.1.74/32, version 43, epoch 0, cached adjacency 5.5.5.14
0 packets, 0 bytes
  tag information set
    local tag: BGP route head
    fast tag rewrite with
      Recursive rewrite via 217.60.217.2/32, tags imposed {23}
  via 217.60.217.2, 0 dependencies, recursive
    next hop 5.5.5.14, Ethernet0/0.2 via 217.60.217.2/32
    valid cached adjacency
    tag rewrite with
      Recursive rewrite via 217.60.217.2/32, tags imposed {23}
2611-CE-30#
```

Problem with the 217.60.217.2. Check its label binding in FIB/LIB.

# MPLS Fwd Plane: Troubles and Shooting (Cont.)

Cisco.com

- **"Recursive rewrite" usually means that**
  - **(a) Either the label to the next-hop is not available**
  - **(b) Or there is an internal problem with the CEF recursion resolution process**
- **(a) usually turns out to be a LDP problem, and should be fixed by investigating into LDP**
- **(b) could be fixed by "clear ip route <prefix>" or "clear ip bgp *"**





# MPLS Fwd Plane: Troubles and Shooting (Cont.)

Cisco.com

- **In order to troubleshoot (a), check the label availability for the next-hop (in LIB). If it is missing, then fix LDP**

```
2611-CE-30#sh mpls for 217.60.217.2
Local  Outgoing    Prefix          Bytes tag  Outgoing    Next Hop
tag    tag or VC   or Tunnel Id    switched   interface
17     Untagged    217.60.217.2/32  0          Et0/0.2     5.5.5.14
2611-CE-30#
```

Untagged outgoing label

```
2611-CE-30#sh mpls ldp bind 217.60.217.2 32
  tib entry: 217.60.217.2/32, rev 14
        local binding:  tag: 17
2611-CE-30#
```

No remote label binding in LIB

```
2611-CE-30#sh mpls ldp dis
 Local LDP Identifier:
    217.60.217.3:0
    Discovery Sources:
    Interfaces:
        Ethernet0/0.2 (ldp): xmit
2611-CE-30#co
```

Because there is no LDP neighbor.

## MPLS Fwd Plane: Troubles and Shooting (Cont.)

Cisco.com

- LDP session needs to be established first
- It is an LDP (control plane) problem
- Troubleshoot for the LDP (as shown in the control plane section)





## Agenda

Cisco.com

- Prerequisites
- MPLS Troubleshooting
  - Control Plane
  - Forwarding Plane
- Conclusion

## Conclusion

Cisco.com

- MPLS seems cryptic, but it is not. ☺
- Whether to look at FIB or LFIB?
- Whether it is a BGP or MPLS problem?
- There is a tremendous work going on to ease the troubleshooting
  - LSP ping, MPLS MIBs, etc.





## Complete Your Online Session Evaluation!

Cisco.com

**WHAT:** Complete an online session evaluation and your name will be entered into a daily drawing

**WHY:** Win fabulous prizes! Give us your feedback!

**WHERE:** Go to the Internet stations located throughout the Convention Center

**HOW:** Winners will be posted on the onsite Networkers Website; four winners per day