Understanding MPLS OAM capabilities to troubleshoot MPLS Networks

Mukhtiar A. Shaikh (mshaikh@cisco.com)

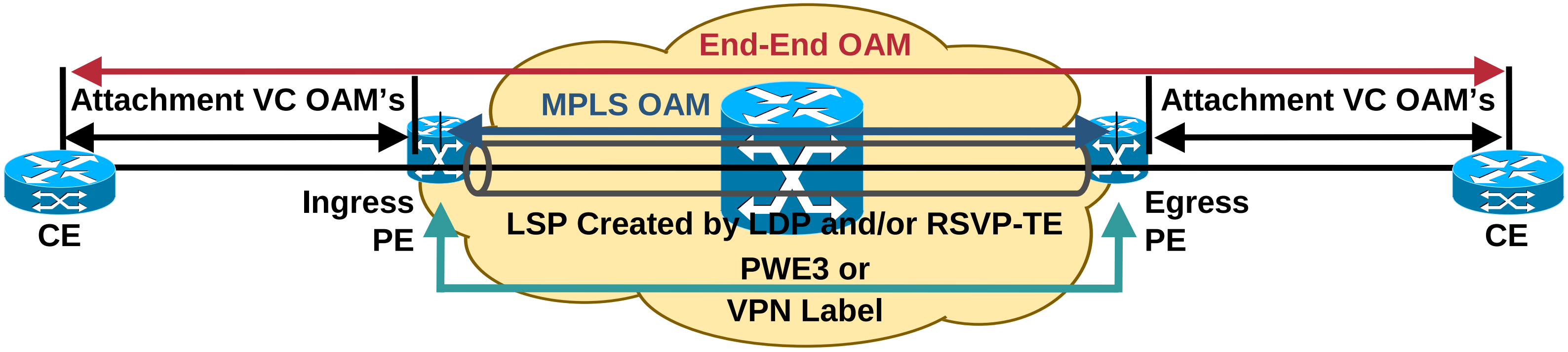
Moiz Moizuddin (mmoizudd@cisco.com)

Agenda

- MPLS Overview
- Existing Ping/Trace Capabilities
- LSP Ping/Trace
    - Theory of Operation
    - MPLS Echo Packet
    - Configuration and Troubleshooting Using LSP Ping/Trace
        - LSP Ping
        - LSP Trace
    - AToM VCCV
- Summary

MPLS OAM Overview

- **Converged network implies a wide range of applications and OAM needs**
- **IP Based Tools**
  - **A flexible set of tools**
  - **LSP Ping / Traceroute**



Agenda

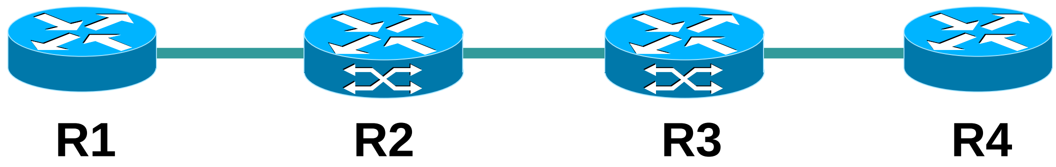
- MPLS Overview
- Existing Ping/Trace Capabilities
- LSP Ping/Trace
    - Theory of Operation
    - MPLS Echo Packet
    - Configuration and Troubleshooting Using LSP Ping/Trace
        - LSP Ping
        - LSP Trace
    - AToM VCCV
- Summary

IP Ping/Trace

- PING makes use of the Internet Control Message Protocol (ICMP) protocol
- Ping message of 2 types

  type=8: ICMP echo request messages
  type=0: ICMP echo reply message
- Traceroute makes use of the Internet Control Message Protocol (ICMP) protocol and TTL field on the IP header
- Traceroute is sent in a UDP packet encapsulated on an IP packet
- TTL-field of an IP datagram is decremented by each hop

Traceroute from R1 with Destination R4



IP Datagram with
Destination R4 and TTL=1

R2 Drops the Packet and Sends TTL
Expired ICMP Message Back to R1

IP Datagram with Destination R4 and TTL=2, R2
Decrements TTL by 1 and Forwards It to R3

R3 Drops the Packet and Sends TTL
Expired ICMP Message Back to R1

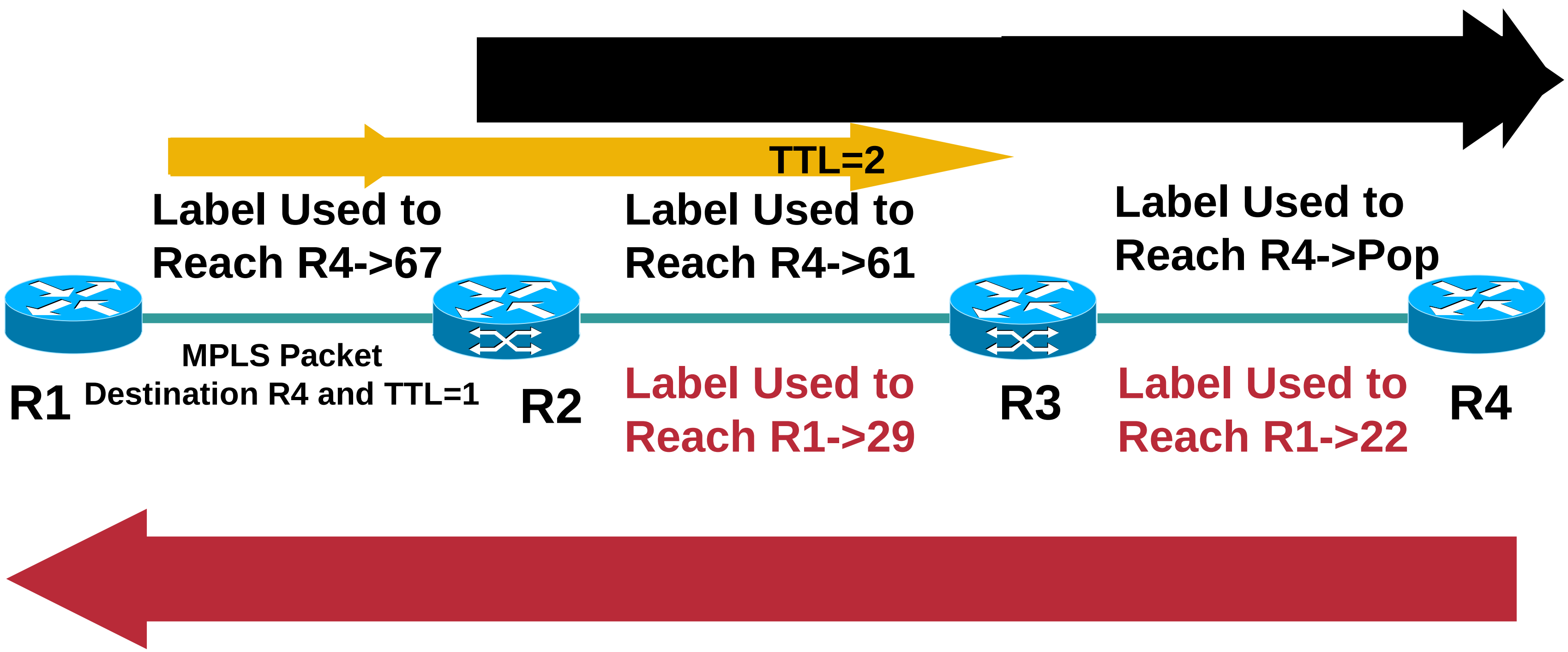
IP datagram with Destination R4 and TTL=3, Datagram Reaches R4

R4 Responds with the ICMP Message

R1 Now Has All the ICMP Error Messages with the Corresponding Source Addresses and Hence Has Got the Complete Route to the Destination

Traceroute from R1 to R4 in MPLS Environment

IP Packet’s TTL Field Is Copied onto the TTL Field of Label Header



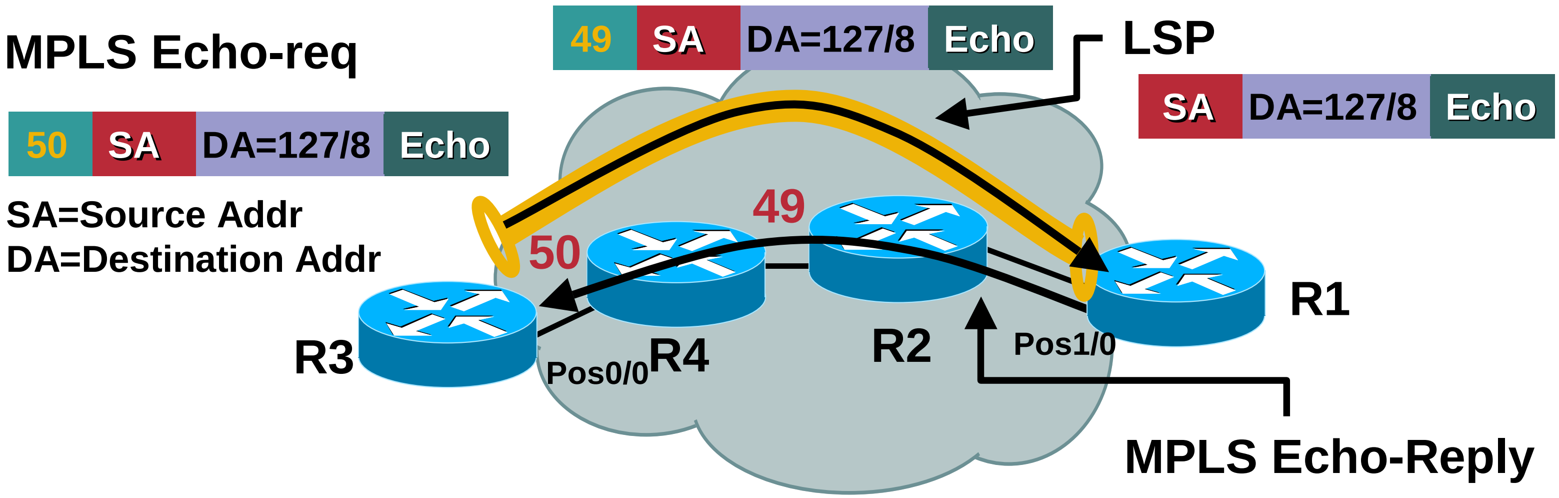
Agenda

- MPLS Overview
- Existing Ping/Trace Capabilities
- LSP Ping/Trace
    - Theory of Operation
    - MPLS Echo Packet
    - Configuration and Troubleshooting Using LSP Ping/Trace
        - LSP Ping
        - LSP Trace
    - AToM VCCV
- Summary

LSP Ping

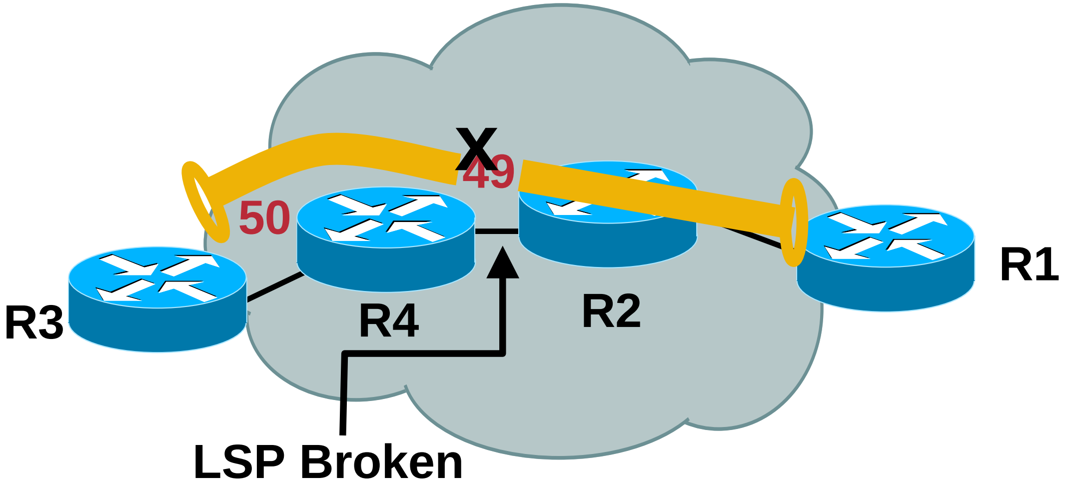
- **LSP Ping/Trace, like the traditional IP Ping, is based on echo request and echo reply**
- **LSP Ping/Trace doesn't use an ICMP packet**
- **Relies on IPv4(or IPv6) UDP packets with port 3503**
- **UDP packets received with port 3503 are either an MPLS echo or an MPLS echo-reply**

Theory of Operation



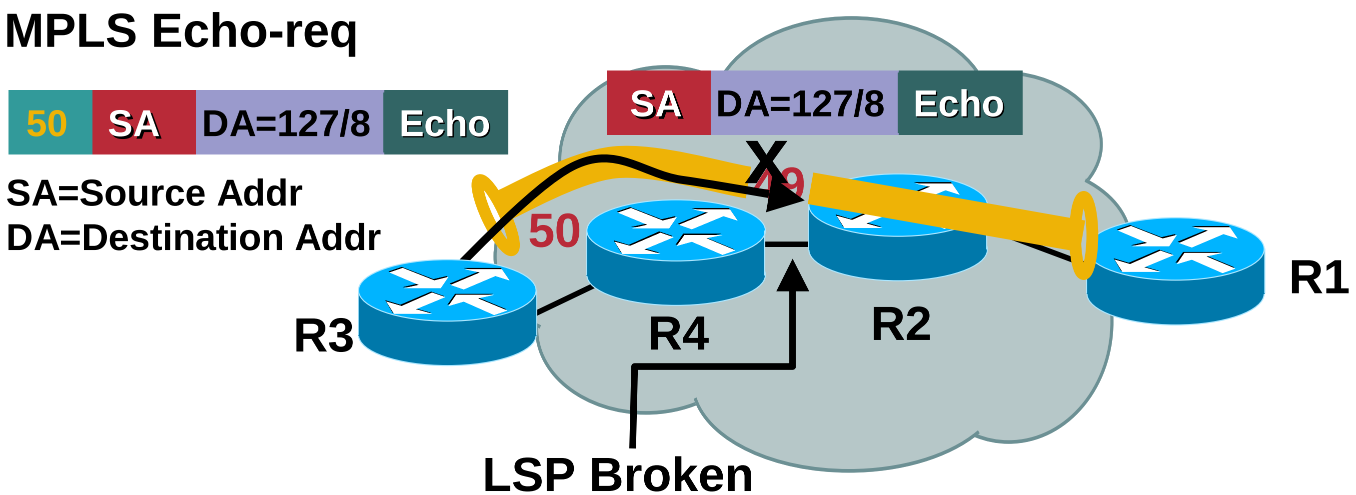
- We use the same label stack as used by the LSP and this makes the echo to be switched inband of LSP
- The IP header destination address field of the echo request is a 127/8 address
- An Echo reply, which may or not be labelled, has outgoing interface IP address as the source; destination IP address/port are copied from the echo-request's source address/port

Theory of Operation (Cont.)



- **Various reasons for LSP to break**
  - **Broken LDP adjacency**
  - **MPLS not enabled**
  - **Mismatch labels**
  - **Software/hardware corruption**
- **Regular IP ping will be successful**

Theory of Operation (Cont.)



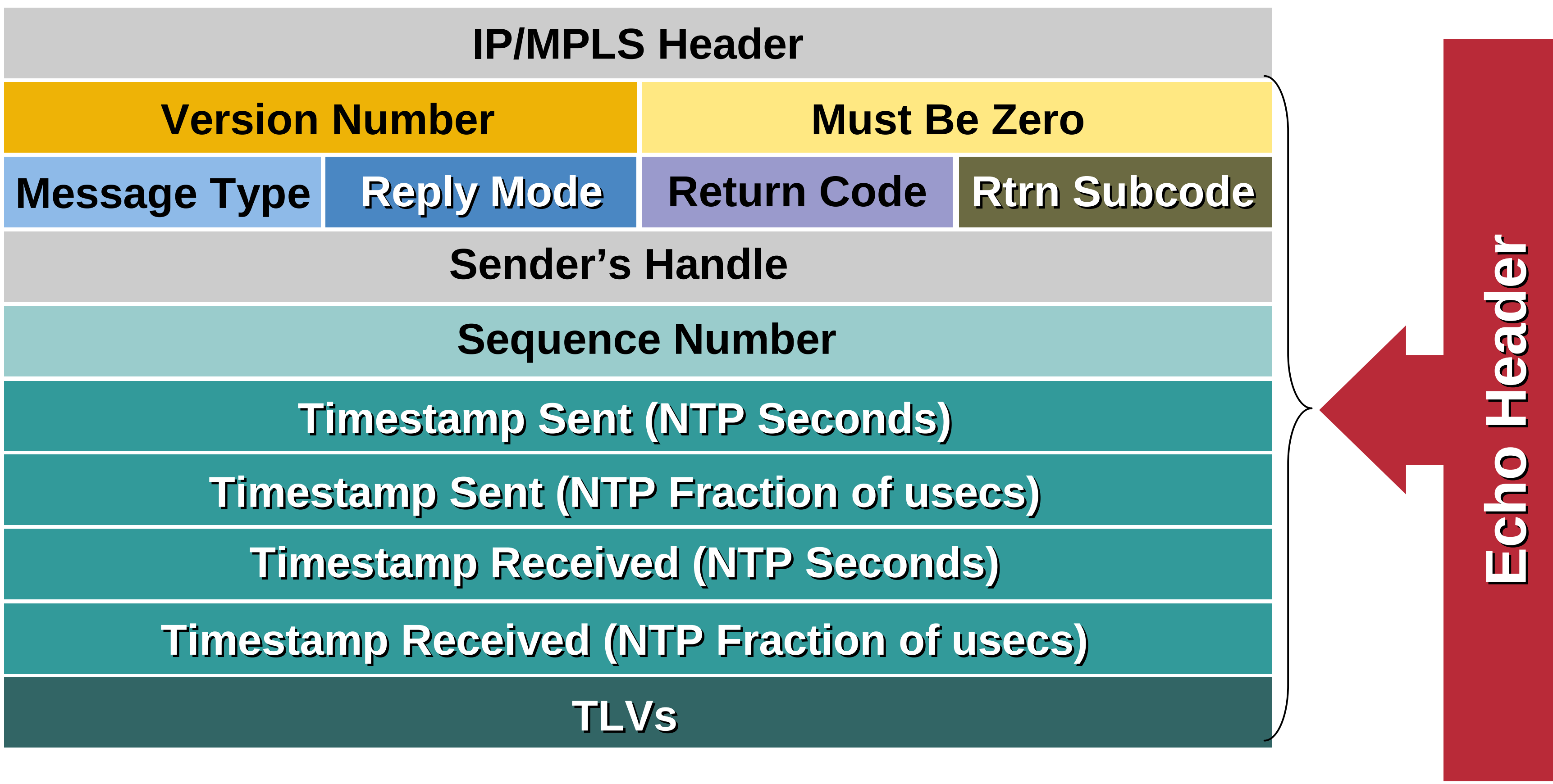
- **Presence of the 127/8 address in the IP header destination address field causes the packet to be consumed by any routers trying to forward the packet using the ip header**
- **In this case R2 would not forward the echo-req to R1 but rather consumes the packet and sends a reply to R3 accordingly**

Agenda

- MPLS Overview
- Existing Ping/Trace Capabilities
- LSP Ping/Trace
    - –Theory of Operation
    - –MPLS Echo Packet
    - –Configuration and Troubleshooting Using LSP Ping/Trace
        - •LSP Ping
        - •LSP Trace
    - –AToM VCCV
- Summary

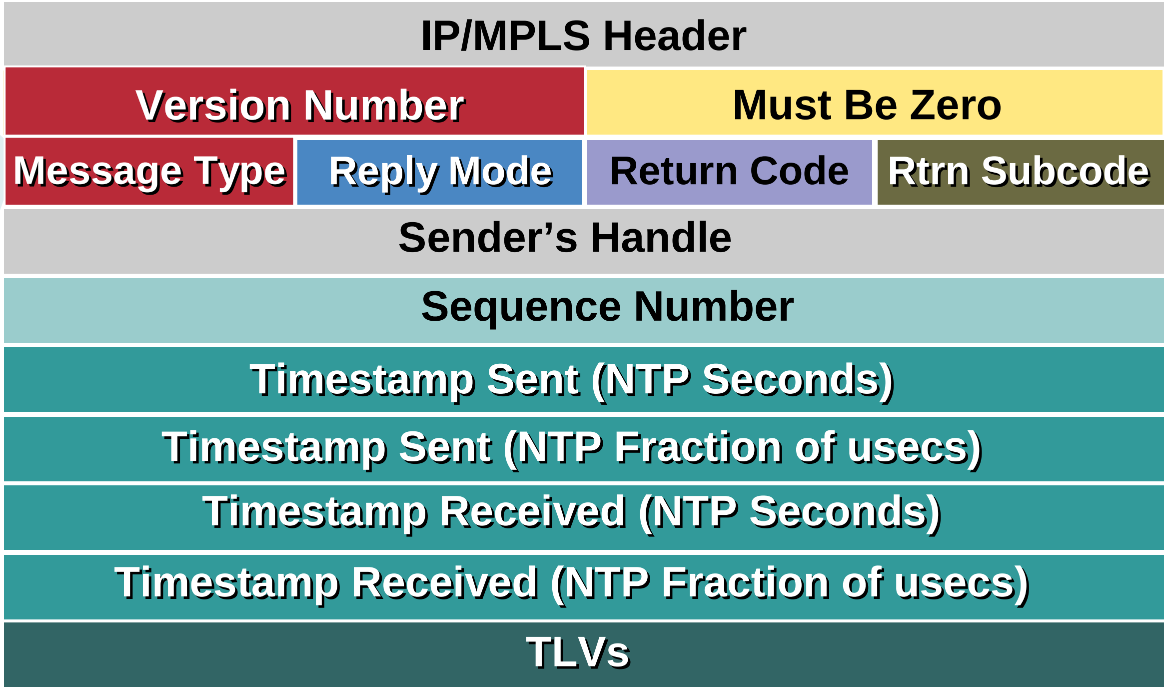
Packet Format of an MPLS LSP Echo

MPLS LSP Echo Request and Replies Are UDP Packets with Header and TLVs



Packet Format of an MPLS LSP Echo (Cont.)

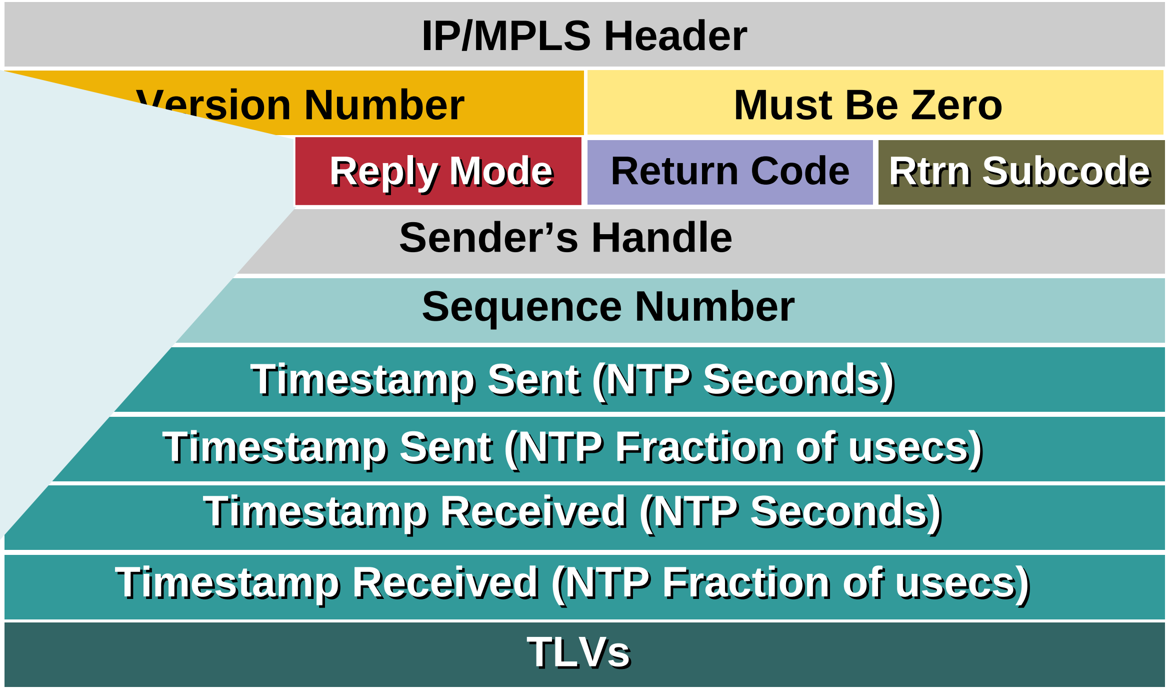
Value	Meaning
1	MPLS Echo Request
2	MPLS Echo Reply



- Version Number: It's Set to One
- Message Type: Message Type Field Tells Whether the Packet Is an MPLS Echo Request or MPLS Echo Reply

Packet Format of an MPLS LSP Echo (Cont.)

Value	Meaning
1	Do Not Reply
2	Reply via an IPv4 UDP Packet
3	Reply via an IPv4 UDP packet with Router Alert



Reply Mode: The Reply Mode Is Used to Control How the Target Router Replies to MPLS Echo Request

Return Code

IP/MPLS Header		
Version Number	Must Be Ze	
Message Type	Reply Mode	Return Code
Sender's Handle		
Sequence Number		
Timestamp Sent (NTP Seconds)		
Timestamp Sent (NTP Fraction of usecs)		
Timestamp Received (NTP Seconds)		
Timestamp Received (NTP Fraction of usecs)		
TLVs		

Value	Meaning
0	The Error Code Is Contained in the Error Code TLV
1	Malformed Echo Request Received
2	One Or More of the TLVs Was Not Understood
3	Replying Router Is an Egress for the FEC
4	Replying Router Has No Mapping for the FEC
5	Replying Router Is Not One of the "Downstream Routers"
6	Replying Router Is One of the "Downstream Routers", and Its Mapping for this FEC on the Received Interface Is the Given Label

- The router initiating the LSP ping/trace would set the return code to zero
- The replying router would set it accordingly based on the table shown

Target FEC Stack TLV

Value	Meaning
1	Target FEC Stack
2	Downstream Mapping
3	Pad
4	Error Code
5	Vendor Enterprise Code

Sub Type	Length	ValueField
1	5	LDP IPv4 Prefix
2	17	LDP IPv6 Prefix
3	20	RSVP IPv4 Session Query
4	56	RSVP IPv6 Session Query
5		Reserved
6	13	VPN IPv4 Prefix
7	25	VPN IPv6 prefix
9	10	L2 Circuit ID

Sub-TLVs

- **LDP IPv4 Prefix Sub-tlv**

0 7 8 15	16
0x0001	Length = 5
Ipv4 Prefix	31
Prefix Length	

- **RSVP IPv4 Prefix Sub-tlv**

0 15	16 31
0x0003	Length = 20
IPv4 Tunnel Endpoint Address	
Must Be Zero	Tunnel ID
Extended Tunnel ID	
IPv4 Tunnel Sender Address	
Must Be Zero	LSP ID

L2 Circuit Type (Sub-TLV)

- **L2 Circuit Type Sub-tlv**

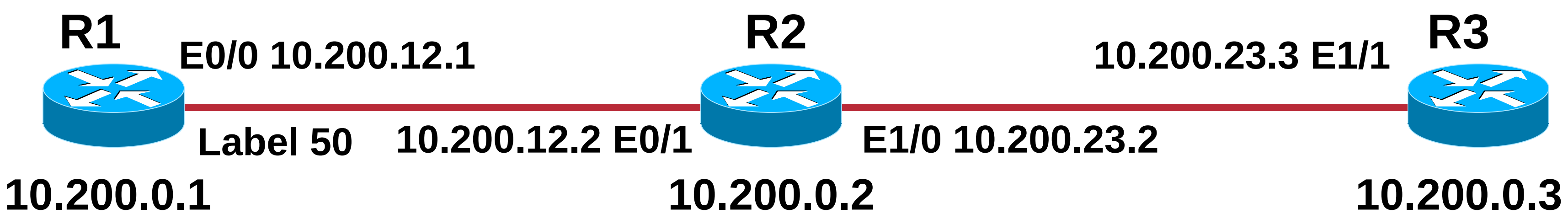
0x0009		Length = 16
Remote PE Address		
Source PE Address		
PWID Type	PWID Length=4	PWID

- **L3VPN (VPN IPv4 prefix) Sub-tlv**

0 7	8 15	16 31
0x0006		Length = 13
Route-Distinguisher		
(8-octets)		
IPv4 Prefix (4- Octets)		
Prefix Length		

Downstream Mapping TLV

Value	Meaning
1	Target FEC Stack
2	Downstream Mapping
3	Pad
4	Error Code
5	Vendor Enterprise Code



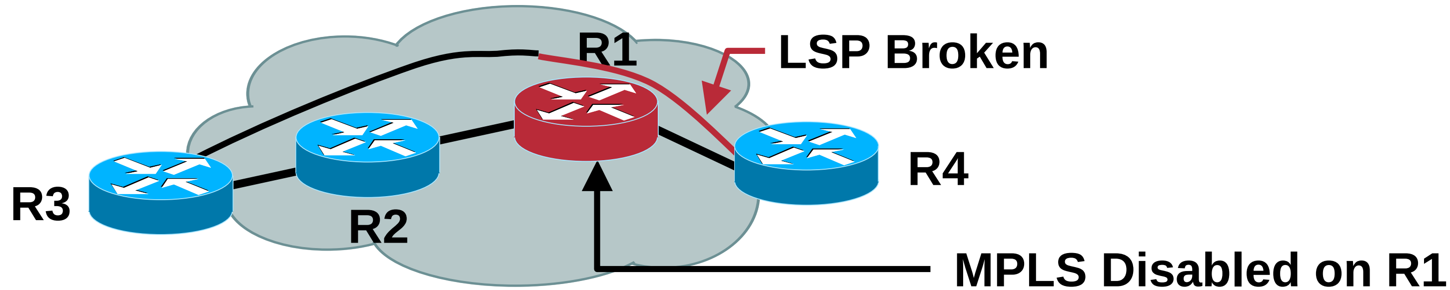
R1's Downstream Mapping for 10.200.0.3
Common_Header
MTU: Mtu of E0/0
Address Type 1
Downstream Intf Addr 10.200.12.1
Downstream Label 50

R2's Downstream Mapping for 10.200.0.3
Common_Header
MTU: Mtu of E1/0
Address Type 1
Downstream Intf Addr 10.200.23.2

Agenda

- MPLS Overview
- Existing Ping/Trace Capabilities
- LSP Ping/Trace
    - –Theory of Operation
    - –MPLS Echo Packet
    - –Configuration and Troubleshooting Using LSP Ping/Trace
        - •LSP Ping
        - •LSP Trace
    - –AToM VCCV
- Summary

Troubleshooting Using LSP Ping (IPv4)
MPLS Disabled at the P Router (R1)



- **If a Regular Ping Is Done from R3 to R4, It Would Be Successful But an LSP Ping Would Fail**
- **The Response Would Come from R1**

```
R3#ping 10.200.0.4
!!!!!
Success rate is 100 percent (5/5), round-trip
min/avg/max = 24/28/32 ms
```

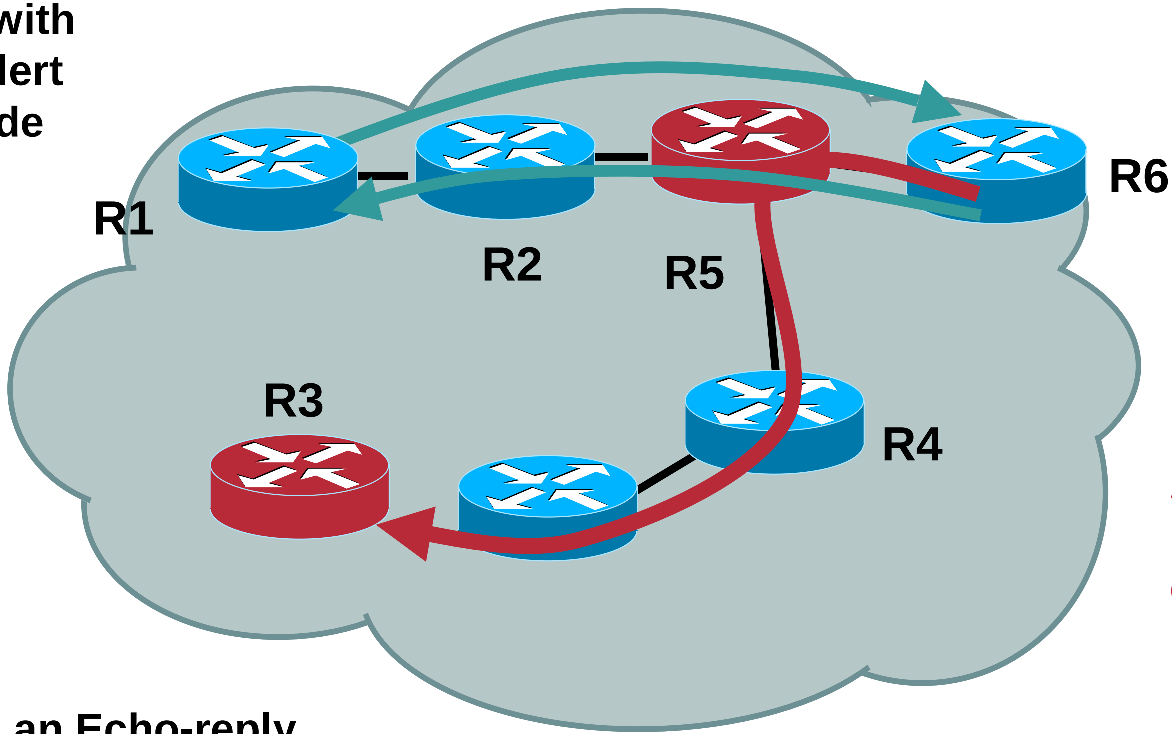
```
R3#ping mpls ip 10.200.0.4/32
Sending 5, 100-byte MPLS Echos to
10.200.0.4/32,
     timeout is 2 seconds, send interval is 0 msec:

Codes: '!' - success, 'Q' - request not transmitted,
      '.' - timeout, 'U' - unreachable,
      'R' - downstream router but not target

Type escape sequence to abort.
UUUUU
Success rate is 0 percent (0/5)
```

```
R3#ping mpls ipv4 10.200.0.4/32 verbose
Sending 5, 100-byte MPLS Echos to 10.200.0.4/32,
     timeout is 2 seconds, send interval is 0 msec:

Codes: '!' - success, 'Q' - request not transmitted,
      '.' - timeout, 'U' - unreachable,
      'R' - downstream router but not target

Type escape sequence to abort.
U   10.200.21.1, return code 4
U   10.200.21.1, return code 4
U   10.200.21.1, return code 4
U   10.200.21.1, return code 4
U   10.200.21.1, return code 4

Success rate is 0 percent (0/5)
```

Troubleshooting Using LSP Ping (IPv4) (Using Router Alert)

Send a ping from R1 with Router Alert reply mode option

Echo-req Is Sent from R1-R6



R5 Receives a Reply with RA

It Process Switch the Packet Correctly to R2

R6 Issues an Echo-reply

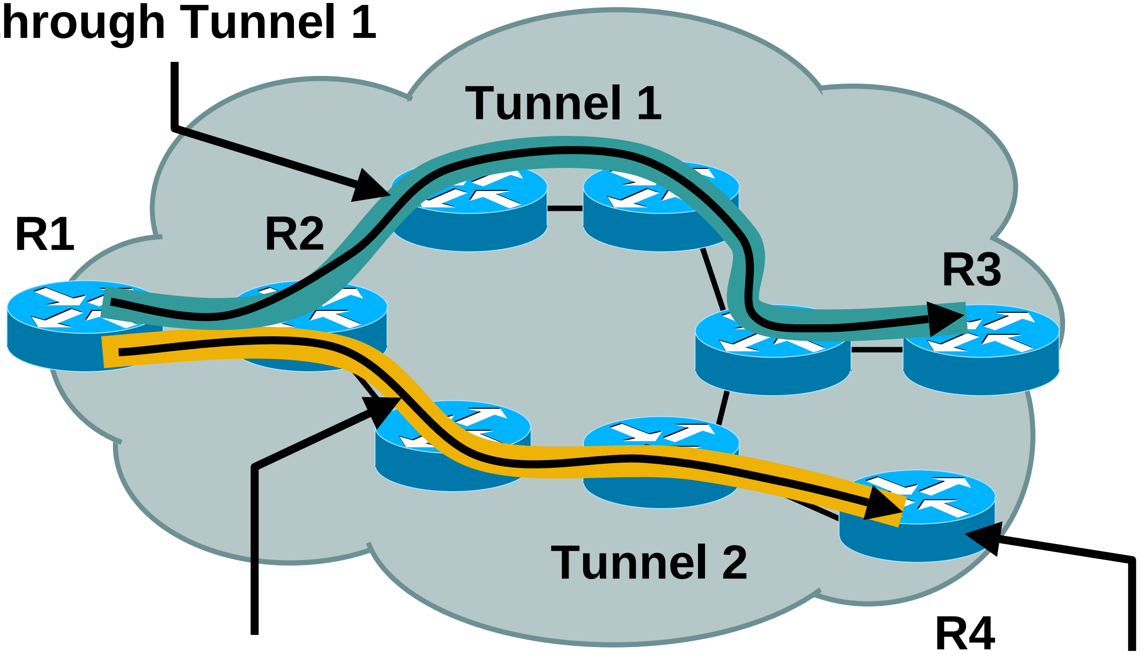
R5 Has a Wrong Label Binding and Forward the Packet to R4

R3 Would Drop the Packet

So LSP Ping Fails

Troubleshooting Using LSP Ping (RSVP IPv4)

LSP Ping Is Initiated from R1 through Tunnel 1



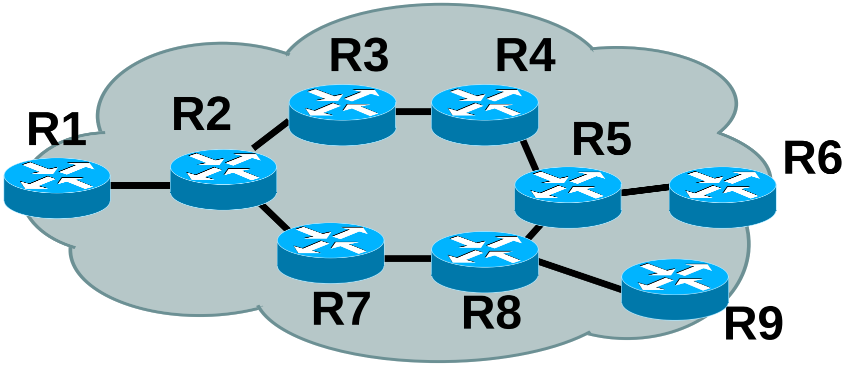
Due to an Error on R2 the LSP Ping Is Switched into Tun 2

R4 Would Recognize that **dest addr, LSP id and Tu id** Are Different and Would Reply with a Return Code 4

Agenda

- MPLS Overview
- Existing Ping/Trace Capabilities
- LSP Ping/Trace
  - Theory of Operation
  - MPLS Echo Packet
  - Configuration and Troubleshooting Using LSP Ping/Trace
    - LSP Ping
    - LSP Trace
  - AToM VCCV
- Summary

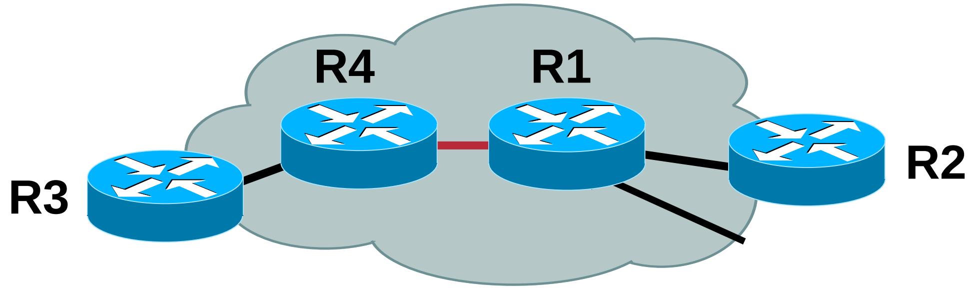
LSP Trace: Path/Tree Trace (Cont.)



Trace Can Be Divided into Two Types

- **Path trace** would give us information of only one path out of all the possible ECMP paths
- In the above example if I do a path trace from R1 to R6; I might only be reported about R1-R2-R3-R4-R5-R6
- **Tree trace** returns ALL of the possible paths between one source and destination
- So in the above case the LSP (tree) trace would give us information about both the paths R1-R2-R3-R4-R5-R6 and R1-R2-R7-R8-R5-R6

Troubleshooting Using LSP Trace (IPv4)



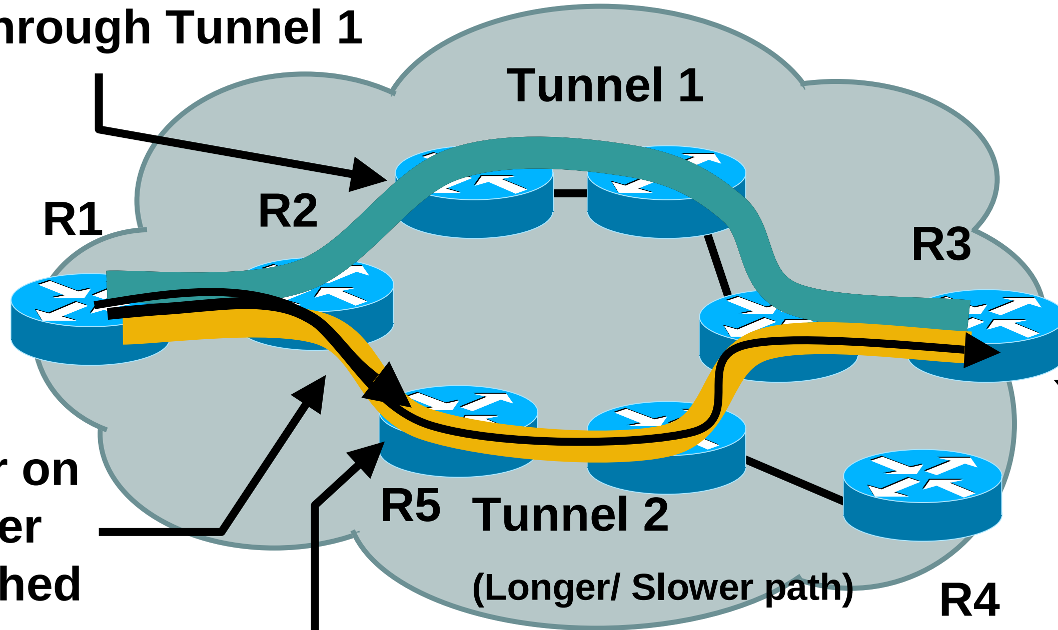
- There is an intermittent response for the data traffic using the LSP R3-R4-R1-R2
- Sweeping LSP ping tells us that packets over 1500 are failing

Output with regular trace..	But if an LSP trace is done, output looks as follows
R3#tracer 10.200.0.2 Type escape sequence to abort. Tracing the route to 10.200.0.2 1 10.200.34.4 [MPLS: Label 44 Exp 0] 0 msec 0 msec 0 msec 2 10.200.14.1 [MPLS: Label 22 Exp 0] 0 msec 0 msec 0 msec 3 10.200.12.2 0 msec * 0 msec R3#	R3#tracer mpls ip 10.200.0.2/32 Tracing MPLS Label Switched Path to 10.200.0.2/32, timeout is 2 seconds Codes: '!' - success, 'Q' - request not transmitted, '.' - timeout, 'U' - unreachable, 'R' - downstream router but not target Type escape sequence to abort. 0 10.200.34.3 MRU 4470 [Labels: 44 Exp: 0] R 1 10.200.14.4 MRU 1500 [Labels: 22 Exp: 0] 4 ms R 2 10.200.12.1 MRU 4474 [implicit-null] 15 ms ! 3 10.200.12.2 20 ms

Troubleshooting Using LSP Trace (RSVP IPv4)

Customer Complains That He's Seeing Latency; Customer Traffic Is Going Through Tunnel 1

R1#ping mpls traffic-eng tunnel tunnel1



Due to an Error on R2 the Customer Traffic Is Switched into Tunnel 2

When We Do LSP Trace R5 Would Not Be Able to Match the 5 Tuples and Would Reply with a Return Code of 4

LSP Ping from R1 Would Work as All the Five Values in the LSP Ping Would Be Correct

R1#trace mpls traffic-eng tunnel tunnel1

Loadbalancing

R3#sh mpls forwarding-table 10.200.0.1

Local tag	Outgoing tag or VC	Prefix or Tunnel Id	Bytes tag switched	Outgoing interface	Next Hop
27	20	10.200.0.1/32	0	PO0/0	point2point
	23	10.200.0.1/32	0	PO1/0	point2point

R3#

```
R3#trace mpls ip 10.200.0.1/32  destination
127.0.0.1
Tracing MPLS Label Switched Path to
10.200.0.1/32, timeout is 2 seconds

Codes: '!' - success, 'Q' - request not
transmitted,
       '.' - timeout, 'U' - unreachable,
       'R' - downstream router but not target

Type escape sequence to abort.
  0 10.200.123.3 MRU 4470 [Labels: 20 Exp: 0]
R 1 10.200.12.2 MRU 1504 [implicit-null] 12 ms
! 2 10.200.12.1 3 ms
```

```
R3#trace mpls ip 10.200.0.1/32  destination
127.0.0.3
Tracing MPLS Label Switched Path to
10.200.0.1/32, timeout is 2 seconds

Codes: '!' - success, 'Q' - request not
transmitted,
       '.' - timeout, 'U' - unreachable,
       'R' - downstream router but not target

Type escape sequence to abort.
  0 10.200.134.3 MRU 4470 [Labels: 23 Exp: 0]
R 1 10.200.14.4 MRU 1504 [implicit-null] 14 ms
! 2 10.200.14.1 5 ms
```

Agenda

- MPLS Overview
- Existing Ping/Trace Capabilities
- LSP Ping/Trace
  - –Theory of Operation
  - –MPLS Echo Packet
  - –Configuration and Troubleshooting Using LSP Ping/Trace
    - •LSP Ping
    - •LSP Trace
  - –AToM VCCV
- Summary

VCCV Switching Types

Two Types of Switching Modes

- Type 1 involves defining the upper nibble of the control word as a Protocol Id (PID) field

Control Word Use Is Signalled in LDP—Standard Form:

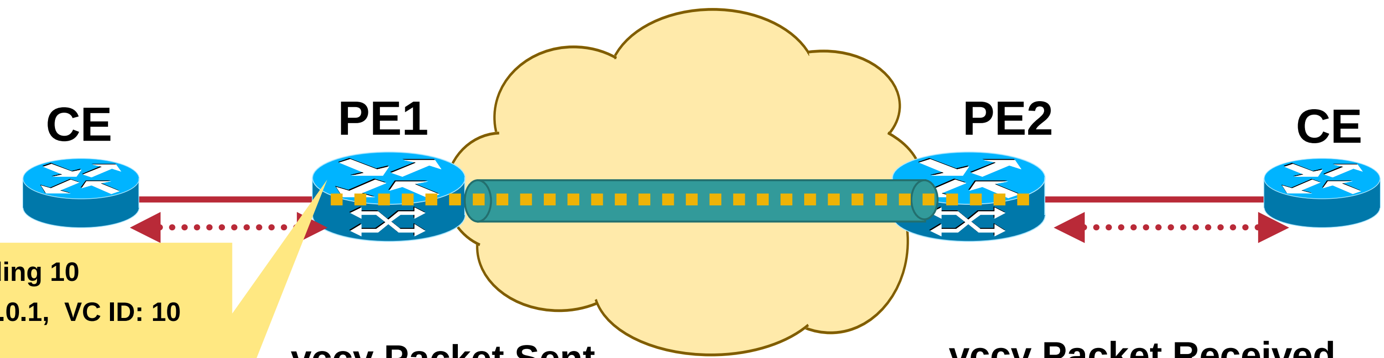
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

0 0 0 0	Flags	FRG	Length	Sequence Number

OAM Uses a different 1st Nibble

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 0 0 1	Reserved	PPP DLL Protocol Number=IPvx
IP OAM Packet: Ping/BFD/LSP Ping		




```
PE1#sh mpls l2transport binding 10
  Destination Address: 10.200.0.1,  VC ID: 10
    Local Label:  16
        Cbit: 0,    VC Type: Ethernet,    GroupID: 0
        MTU: 1500,   Interface Desc: n/a
        VCCV Capabilities: Type 1
    Remote Label: 69
        Cbit: 0,    VC Type: Ethernet,    GroupID: 0
        MTU: 1500,   Interface Desc: n/a
        VCCV Capabilities: Type 1
```

vccv Packet Sent from PE1 to PE2

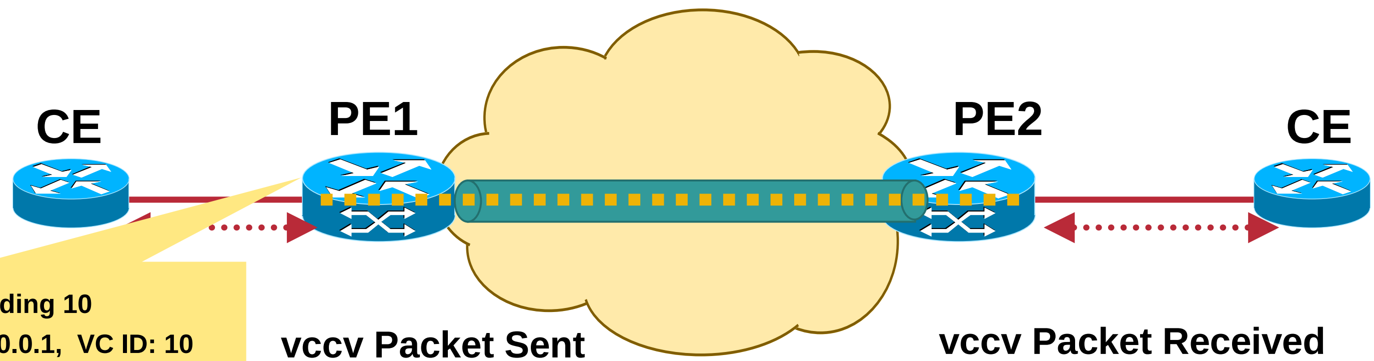
IGP Label TTL=255
vc Label+CW
IP Payload

vccv Packet Received from PE1 to PE2

vc Label+CW
IP Payload

VCCV Switching Types (Cont.)

- **Type 2 involves shimming a MPLS router alert label between the IGP label stack and VC label**




```
PE1#sh mpls l2transport binding 10
  Destination Address: 10.200.0.1,  VC ID: 10
    Local Label:  16
        Cbit: 0,    VC Type: Ethernet,    GroupID: 0
        MTU: 1500,   Interface Desc: n/a
        VCCV Capabilities: Type 2
    Remote Label: 69
        Cbit: 0,    VC Type: Ethernet,    GroupID: 0
        MTU: 1500,   Interface Desc: n/a
        VCCV Capabilities: Type 2
```

vccv Packet Sent from PE1 to PE2

IGP Label TTL=255
Rtr Alert Label 0x0001
vc Label+CW
IP Payload

vccv Packet Received from PE1 to PE2

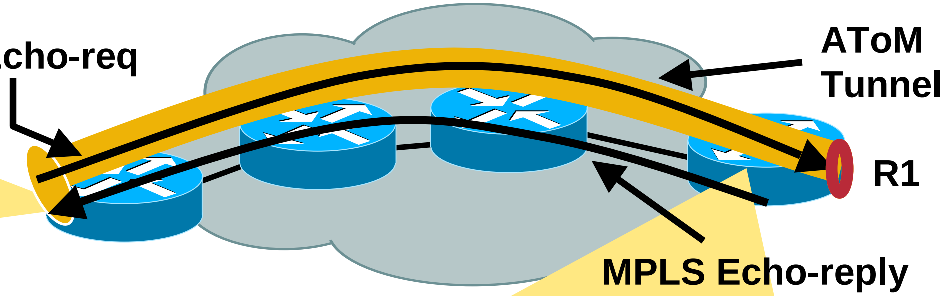
Rtr Alert Label 0x0001
vc Label+CW
IP Payload

Troubleshooting Using LSP Ping (L2 CKT)

Pinging from R3 to R1 through AToM Tunnel
R3#ping mpls pseudowire 10.200.0.1 10

R3#ping mpls pseudowire <IPv4 peer IP addr > <VC ID>?

destination	Destination address or address range
exp	EXP bits in mpls header
interval	Send interval between requests in Routerc
pad	Pad TLV pattern
repeat	Repeat count
reply	Reply mode
size	Packet size
source	Source specified as an IP address
sweep	Sweep range of sizes
timeout	Timeout in seconds
ttl	Time to live
verbose	verbose mode for ping output



R1#

*Jan 19 19:32:17.726: LSPV: AToM echo request rx packet handler

*Jan 19 19:32:17.726: LSPV: Echo packet received: src 10.200.0.3, dst 127.0.0.1, size 122

*Jan 19 19:32:17.734: LSPV: Echo Hdr decode: version 1, msg type 1, reply mode 2 , return_code 0, return_subcode 0, sender handle 850000D1, sequence number 1, ti mestamp sent 20:22:30 UTC Mon Jan 19 2004, timestamp rcvd 00:00:00 UTC Mon Jan 1 1900

*Jan 19 19:32:17.734: LSPV: tlvtype 1, tlvlength 20

*Jan 19 19:32:17.734: LSPV: AToM FEC decode: srcaddr 10.200.0.1, destaddr 10.200 .0.3, vcid 10, vctype 5

*Jan 19 19:32:17.734: LSPV: Target FEC stack length = 20, retcode = 3

*Jan 19 19:32:17.734: LSPV: tlvtype 3, tlvlength 8

*Jan 19 19:32:17.734: LSPV: Pad TLV decode: type 1, size 8

*Jan 19 19:32:17.734: LSPV: Echo Hdr encode: version 1, msg type 2, reply mode 2 , return_code 4, return_subcode 0, sender handle 850000D1, sequence number 1, ti mestamp sent 20:22:30 UTC Mon Jan 19 2004, timestamp rcvd 19:32:17 UTC Mon Jan 1 9 2004

- **Return code 4 sent due to some error condition either of the following has occurred**
  - Wrong VC ID
  - Wrong VC Type
  - Wrong Source Address

Agenda

- MPLS Overview
- Existing Ping/Trace Capabilities
- LSP Ping/Trace
  - –Theory of Operation
  - –MPLS Echo Packet
  - –Configuration and Troubleshooting Using LSP Ping/Trace
    - •LSP Ping
    - •LSP Trace
  - –AToM VCCV
- Summary

Summary

- **Traditional ping/trace not able to detect the problems in the MPLS networks.**
- **LSP ping/trace brings a new set of tools to troubleshoot MPLS forwarding plane problems**
- **VCCV adds new capability to help troubleshoot layer2 VPN issues**

THANK YOU