Understand MPLS L2VPN Pseudowire

Contents

Introduction

Background Information

Overview of L2VPN

Why is L2VPN Needed

MPLSL2VPN Models

Technology Options

1. VPWS Services

2. VPLS Services

3. EVPN

4. PBB-EVPN

VPWS - Pseudo Wire Reference Model

Layer 2 VPN Enabler: The Pseudowire

AToM Architecture

L2Transport over MPLS

VPWS Traffic Encapsulation

Signalling the Pseudowire

Control Word

Forwarding Plane Processing

Operation

Signalling the Status of PW

Basic AToM configuration

Pseudowire Packet Analysis

Topology

L2VPN Interworking

Interworking Possibilities

Related Information

Introduction

This document describes the Multiprotocol Label Switching (MPLS) based L2 Virtual Private Network (L2VPN) pseudowires.

Background Information

The signalling of the pseudowire and packet analysis in Cisco IOS®, Cisco IOS® XE in order to illustrate the behavior is covered.

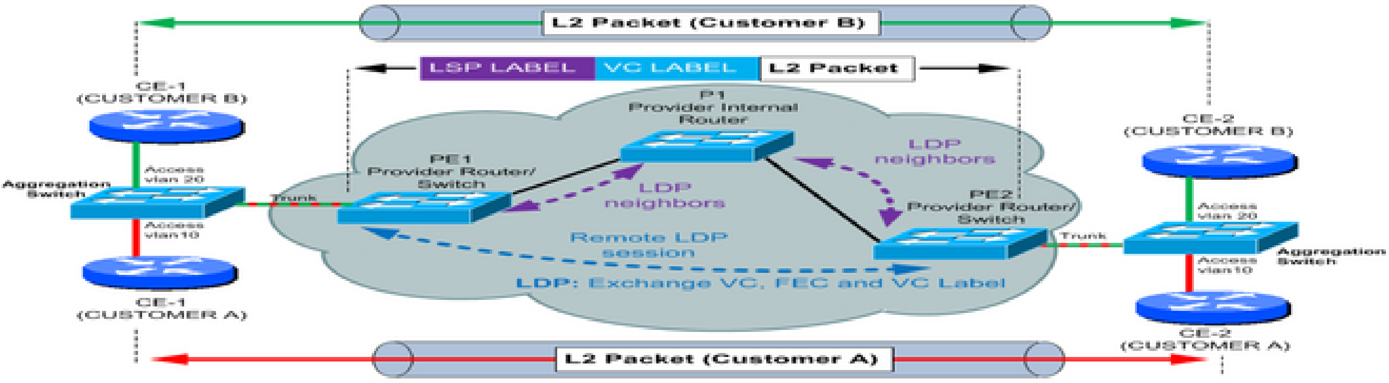
Overview of L2VPN

Layer 2 (L2) transport over MPLS and IP already exists for like-to-like attachment circuits, such as

Ethernet-to-Ethernet, PPP-to-PPP, High-Level Data Link Control (HDLC), and so on

L2VPNs employ L2 services over MPLS in order to build a topology of point-to-point connections that connect end you sites in a VPN. These L2VPNs provide an alternative to private networks that have been provisioned by means of dedicated leased lines or by means of L2 virtual circuits that employ ATM or Frame Relay. The service provisioned with these L2VPNs is known as Virtual Private Wire Service (VPWS).

- L2VPNs are built with Pseudowire (PW) technology.
- PWs provide a common intermediate format to transport multiple types of network services over a Packet Switched Network (PSN) – a network that forwards packets – IPv4, IPv6, MPLS, Ethernet.
- PW technology provides Like-to-Like transport and also Interworking (IW).
- Frames that are received at the PE router on the AC are encapsulated and sent across the PSW to the remote PE router.
- The egress PE router receives the packet from the PSW and removes their encapsulation.
- The egress PE extracts and forwards the frame to the AC.

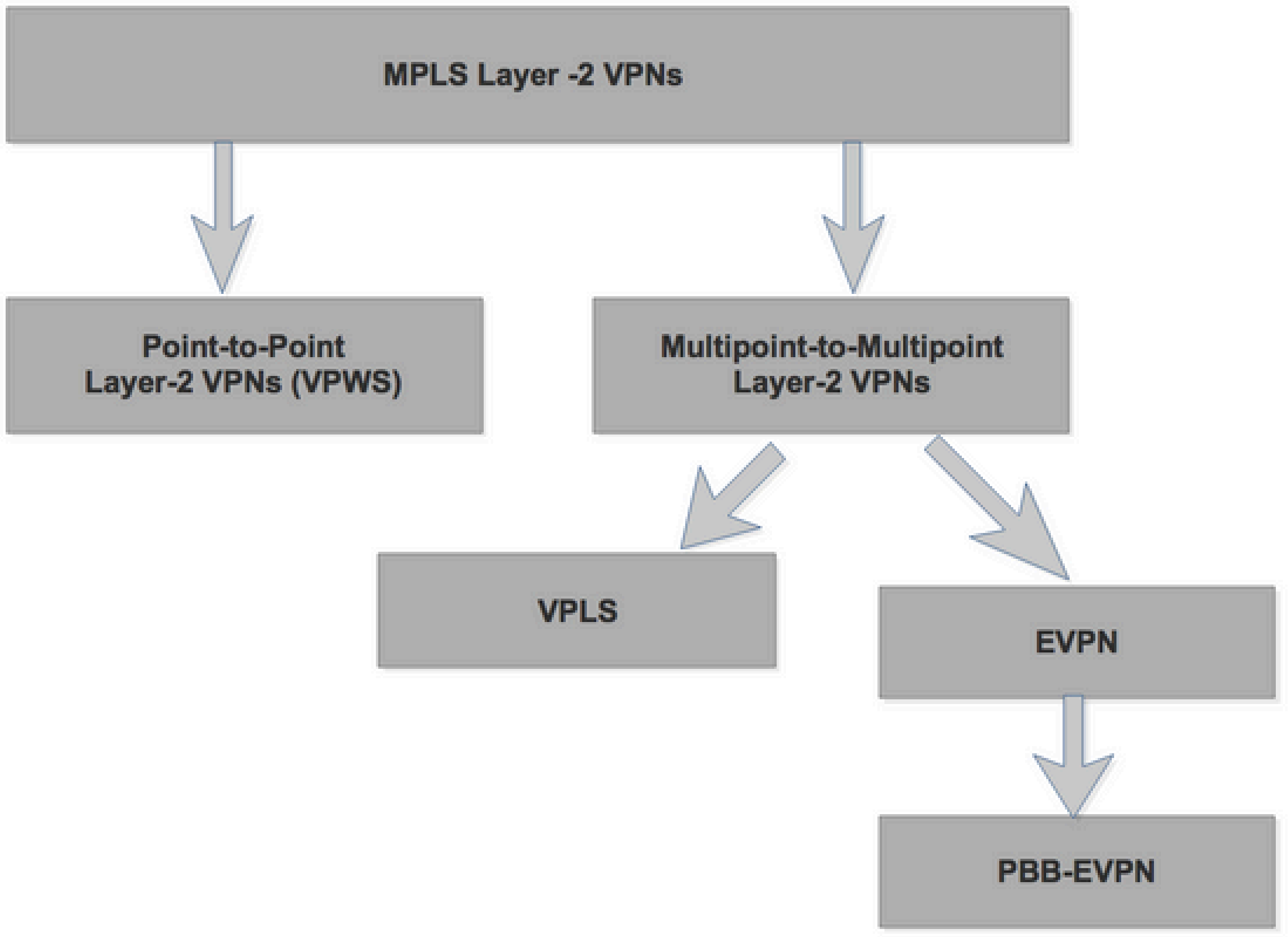


Why is L2VPN Needed

- Allows SP to have a single infrastructure for both IP and legacy services.
- Migrate legacy ATM and Frame Relay services to MPLS/IP core without interruption to existing services.
- Provisioning new L2VPN services are incremental (not from scratch) in existing MPLS/IP core.
- Capital and Operational savings of converged IP/MPLS network.
- SP provides new point-2-point or point-2-multi-point services You can have their own routing, QoS policies, security mechanisms, and so on.

MPLS L2 VPN Models

Technology Options


MPLS Layer -2 VPNs
Point-to-Point Layer-2 VPNs (VPWS)
Multipoint-to-Multipoint Layer-2 VPNs
VPLS
EVPN
PBB-EVPN

1. VPWS Services

• Point-to-point • Referred to as Pseudowires (PWs)

2. VPLS Services

• Multipoint

3. EVPN

• xEVPN family introduces next generation solutions for Ethernet services

a. BGP control-plane for Ethernet Segment and MAC distribution and learning over MPLS core

b. Same principles and operational experience of IP VPNs

• No use of Pseudowires

a. Uses MP2P tunnels for unicast

b. Multi-destination frame delivery via ingress replication (via MP2P tunnels) or LSM

• Multi-vendor solutions under IETF standardization

4. PBB-EVPN

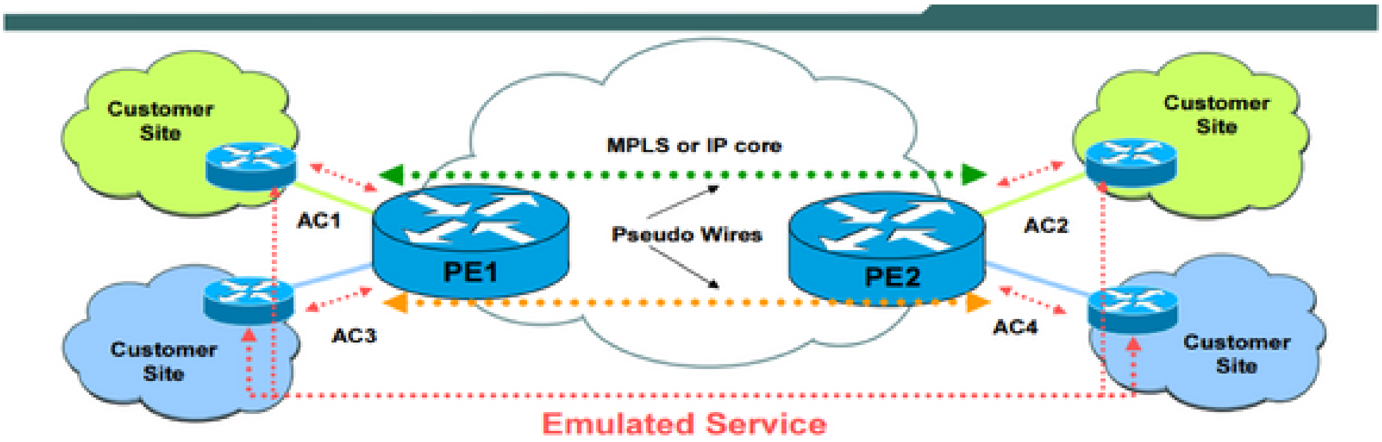
• Combines scale tools from PBB (aka MAC-in-MAC) with BGP-based MAC learning from EVPN

EVPN and Provider Backbone Bridging EVPN (PBB-EVPN) are next-generation L2VPN solutions based on BGP control plane for MAC distribution/learning over the core, designed to address these requirements:

- Per-Flow Redundancy and Load Balancing
- Simplified Provisioning and Operation
- Optimal Forwarding
- Fast Convergence
- MAC Address Scalability

VPWS - Pseudo Wire Reference Model

1. PW is a connection between two PE devices which connects two ACs, that carry L2 frames.
2. Any Transport Over MPLS (AToM) is Cisco's implementation of VPWS for IP/MPLS networks.
3. Attachment Circuit (AC) is the physical or virtual circuit attaching a CE to a PE, can be ATM, Frame Relay, HDLC, PPP and so on.
4. You Edge (CE) equipment perceives a PW as an unshared link or circuit.



Layer 2 VPN Enabler: The Pseudowire

L2VPNs are built with Pseudowire (PW) technology.

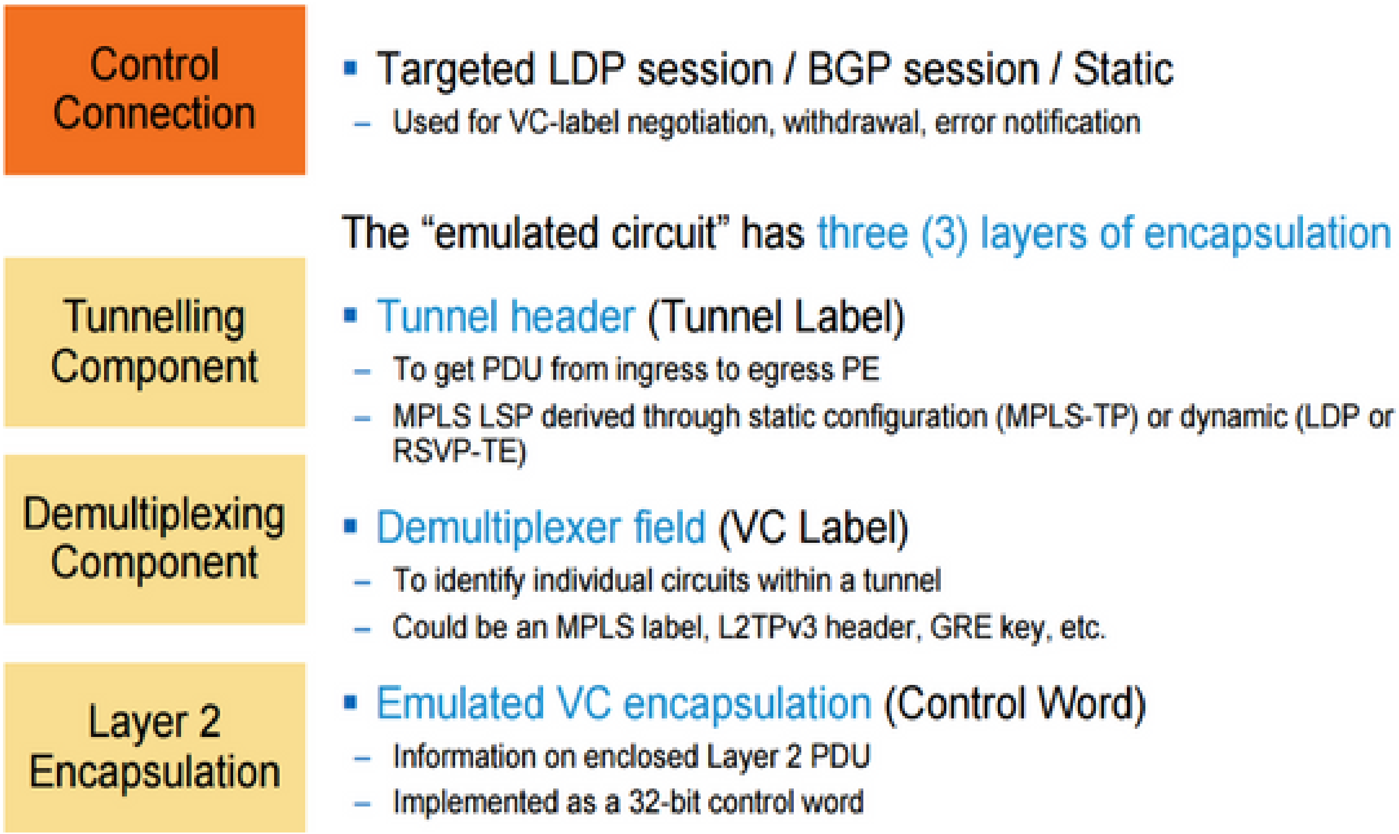
- PWs provide a common intermediate format to transport multiple types of network services over a Packet Switched Network (PSN) – a network that forwards packets – IPv4, IPv6, MPLS, Ethernet.
- PW technology provides Like-to-Like transport and also Interworking (IW).
- Frames that are received at the PE router on the AC are encapsulated and sent across the PSW to the remote PE router.
- The egress PE router receives the packet from the Pseudowire and removed their encapsulation.
- The egress PE extracts and forwards the frame to the AC.

AToM Architecture

- In AToM network, all the routers in the SP run MPLS and the PE router have an AC towards the CE router.
- In the case of AToM, the PSN tunnel is nothing other than a label switched path LSP between the two PE routers.

- As such the label that is associated with that LSP is called tunnel label in context to the AToM.
- First, the LDP signals hop by hop between the PE.
- Second, the LSP can be an MPLS TE tunnel that the RSVP signals with the extensions needed for TE.
- With this tunnel label, you can identify to which PSN tunnel the carried you frame belongs.
- This tunnel label also gets the frames from the local or ingress PE to the remote or egress PE across the MPLS backbone.
- To multiplex several Pseudowire onto one PSN tunnel the PE router uses another label to identify the Pseudowire.
- This label is called the VC or PW label because it identifies the VC or PW that the frame is multiplexed into.

L2 Transport over MPLS



Control Connection

- Targeted LDP session / BGP session / Static
  - Used for VC-label negotiation, withdrawal, error notification

The “emulated circuit” has three (3) layers of encapsulation

Tunnelling Component

- Tunnel header (Tunnel Label)
  - To get PDU from ingress to egress PE
  - MPLS LSP derived through static configuration (MPLS-TP) or dynamic (LDP or RSVP-TE)

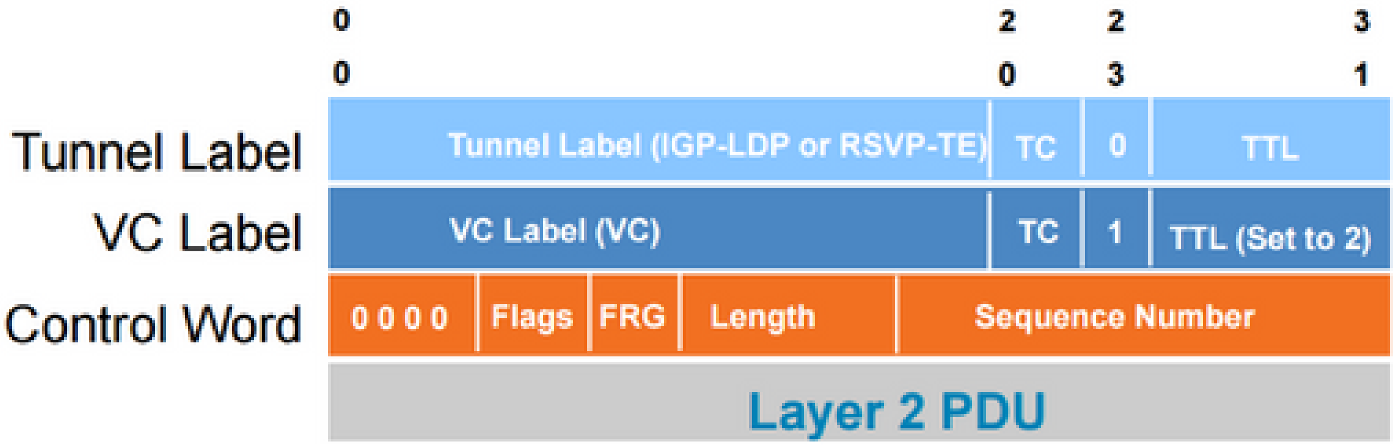
Demultiplexing Component

- Demultiplexer field (VC Label)
  - To identify individual circuits within a tunnel
  - Could be an MPLS label, L2TPv3 header, GRE key, etc.

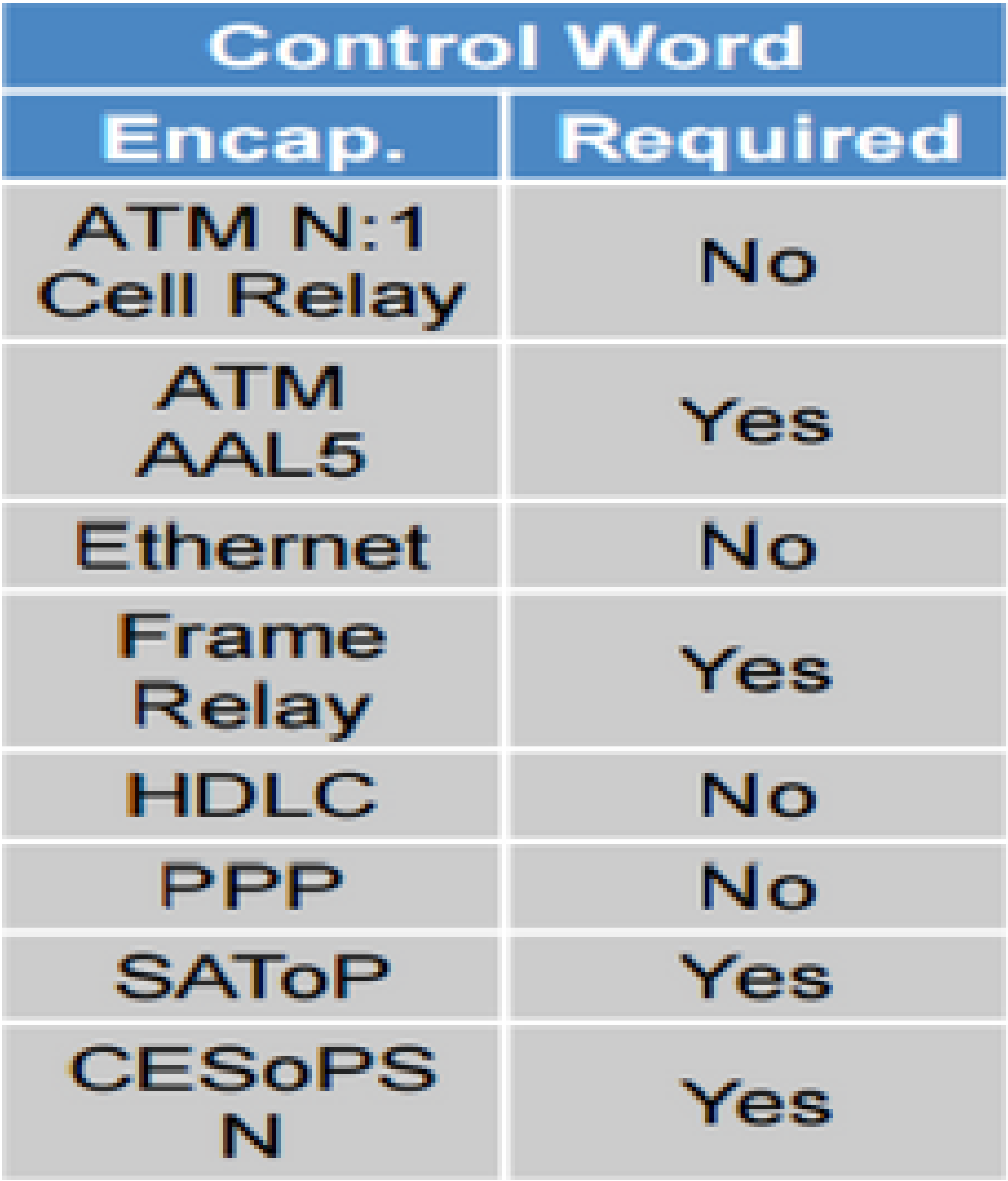
Layer 2 Encapsulation

- Emulated VC encapsulation (Control Word)
  - Information on enclosed Layer 2 PDU
  - Implemented as a 32-bit control word

VPWS Traffic Encapsulation



1. Three-level encapsulation used.
2. Packets switched between PEs using Tunnel label.
3. VC label identifies PW.
4. VC label signalled between PEs.
5. Optional Control Word (CW) carries Layer 2 control bits and enables sequencing.



Control Word	
Encap.	Required
ATM N:1 Cell Relay	No
ATM AAL5	Yes
Ethernet	No
Frame Relay	Yes
HDLC	No
PPP	No
SAToP	Yes
CESoPSN	Yes

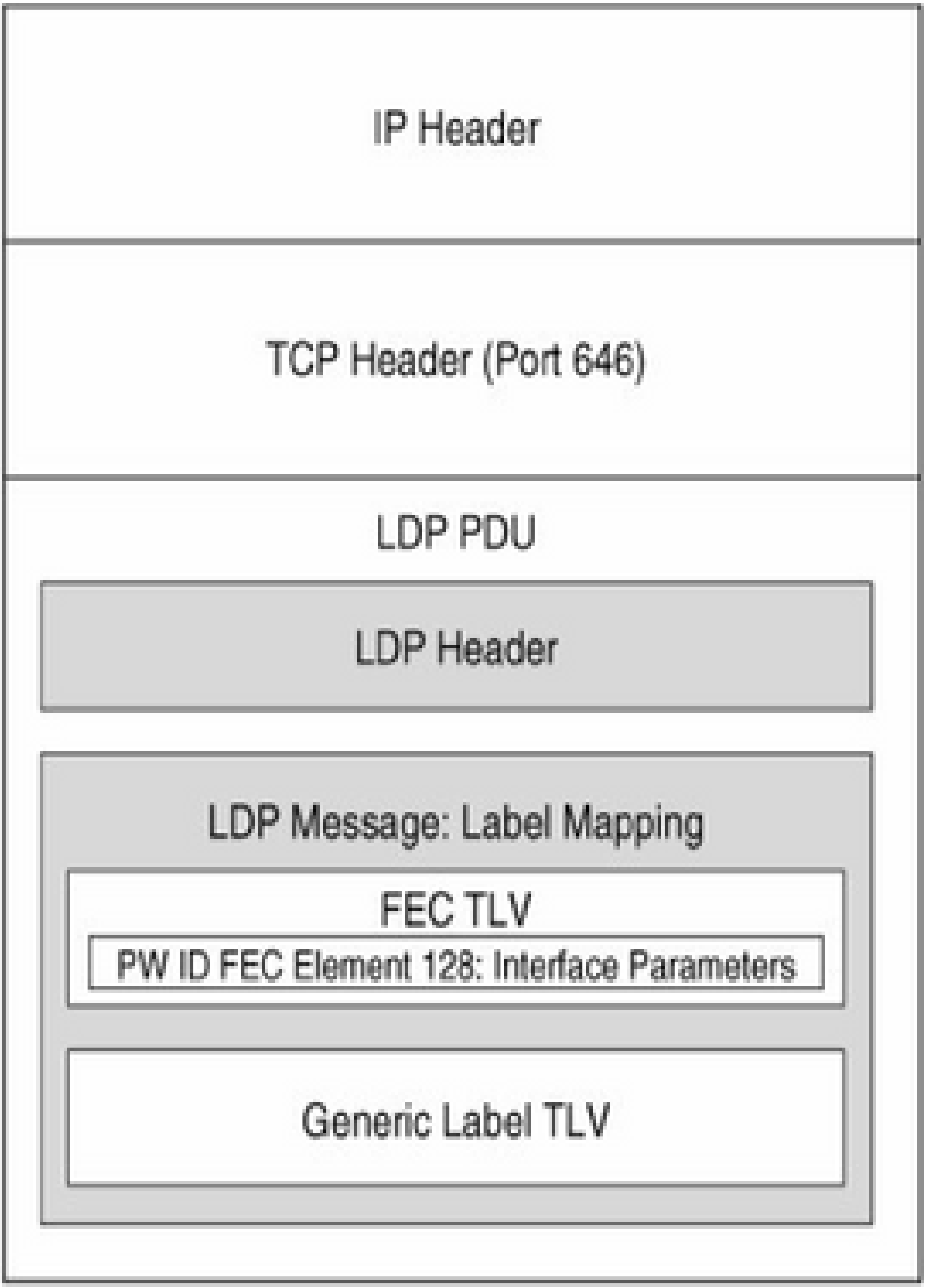
Signalling the Pseudowire

- A TLDP session between the PE router signals the Pseudowire.
- A T-LDP session between the PE routers is to advertise the VC label that is associated with the PSW.
- This label is advertised in a label mapping message that uses the downstream unsolicited label advertisement mode.
- VC label advertised by the egress PE to ingress PE for the AC over the TLDP session. # VC Label by TLDP
- Tunnel label advertised for the egress PE router to the ingress PE by LDP. # Tunnel Label by LDP

Notice that egress PE advertises label 3, which indicated that PHP is used.

The label mapping message that is advertised on the TLDP session contains some TLV :

LDP Label Mapping message:



IP Header

TCP Header (Port 646)

LDP PDU

LDP Header

LDP Message: Label Mapping

FEC TLV

PW ID FEC Element 128: Interface Parameters

Generic Label TLV

- If the other PE router does not support the PW status TLV method, both PE routers revert back to label withdraw method.
- After the pseudowire is singled, the PW status TLV is carried in an LDP notification message. The PW status TLV contains the 32-bit status code field.

Basic AToM configuration

Step 1. Select the encapsulation type.

Step 2. Enable specifying the connect command on the CE facing interface.

xocnnect peer-router-id vcid encapsulation mpls

Peer-router-id: LDP router id for the remote PE router.

VCID: identifier that you assigned to the PW.

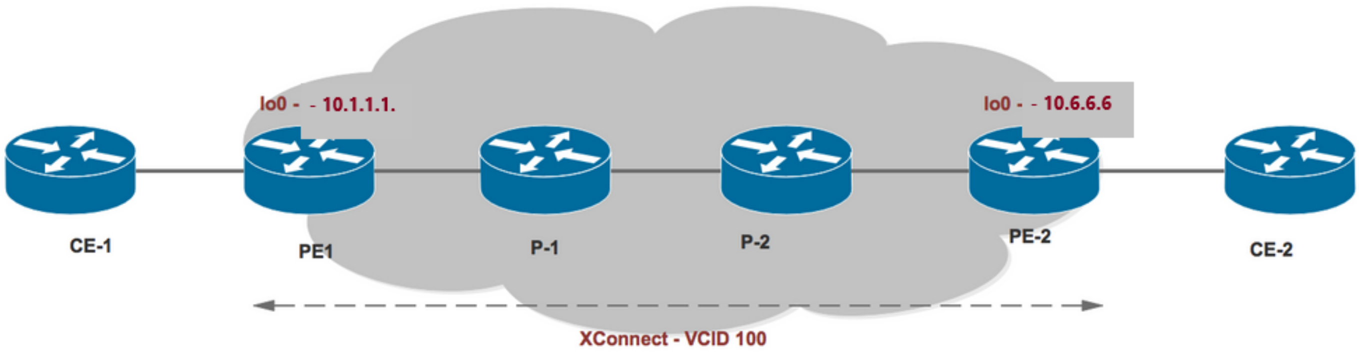
Step 3. As soon as xconnect in both the PE routers is configured, the targeted LDP session is established between the PE router.

Pseudowire Packet Analysis

Initiate a Pseudowire ping from Ingress PE to Egress PE.

MPLS Echo Request and Reply packets sent over point-to-point Pseudowire.

Topology



Ping from PE1 to PE2:

```
R1#ping mpls pseudowire 10.6.6.6 100

Sending 5, 100-byte MPLS Echos to 10.6.6.6,

     timeout is 2 seconds, send interval is 0 msec:

Type escape sequence to abort.

!!!!!
```

```
Success rate is 100 percent (5/5), round-trip min/avg/max = 48/61/80 ms
```

Observations made:

1. ECHO Request:

Carries 2 Labels - VPN and Transport

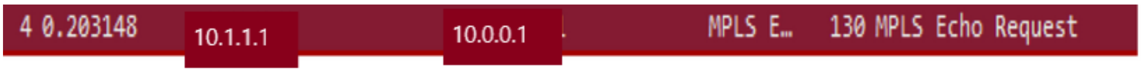
Sent as Labeled Packet that carry PW LABEL. This can be label switched (with Transport Label).

LABELS : 2
SRC IP : LOOPBACK IP (USED IN TARGETED LDP NEIGHBORSHIP)
DST IP : 127.0.0.1
L4 TYPE : UDP
SRC PORT : 3503
DST PORT : 3505
TOS BYTE : OFF
MPLS EXP : OFF
DF BIT : ON

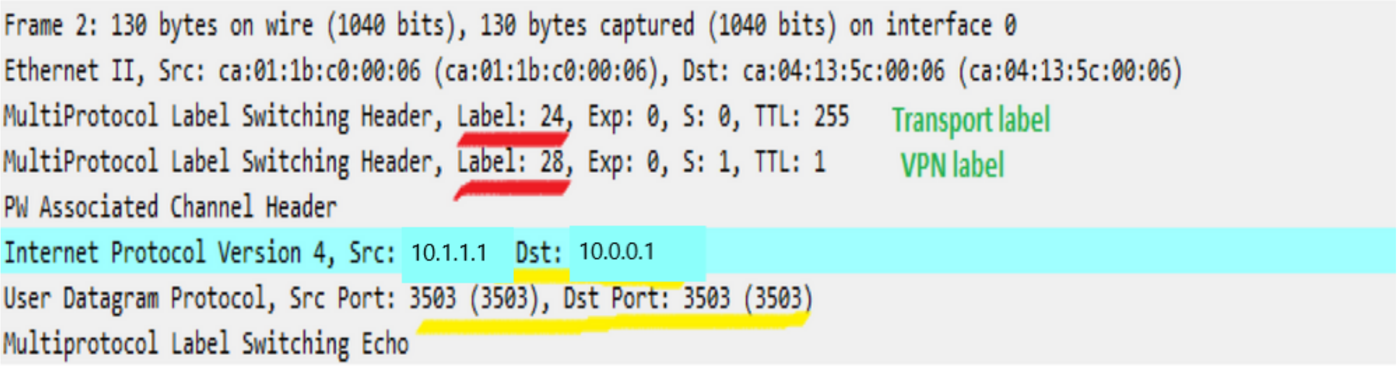
IPv4 OPTIONS Field is in USE: ROUTER ALERT OPTIONS FIELD (Punt to CPU)

UDP PAYLOAD can be MPLS LABEL SWITCHING ECHO REQUEST

Overview:

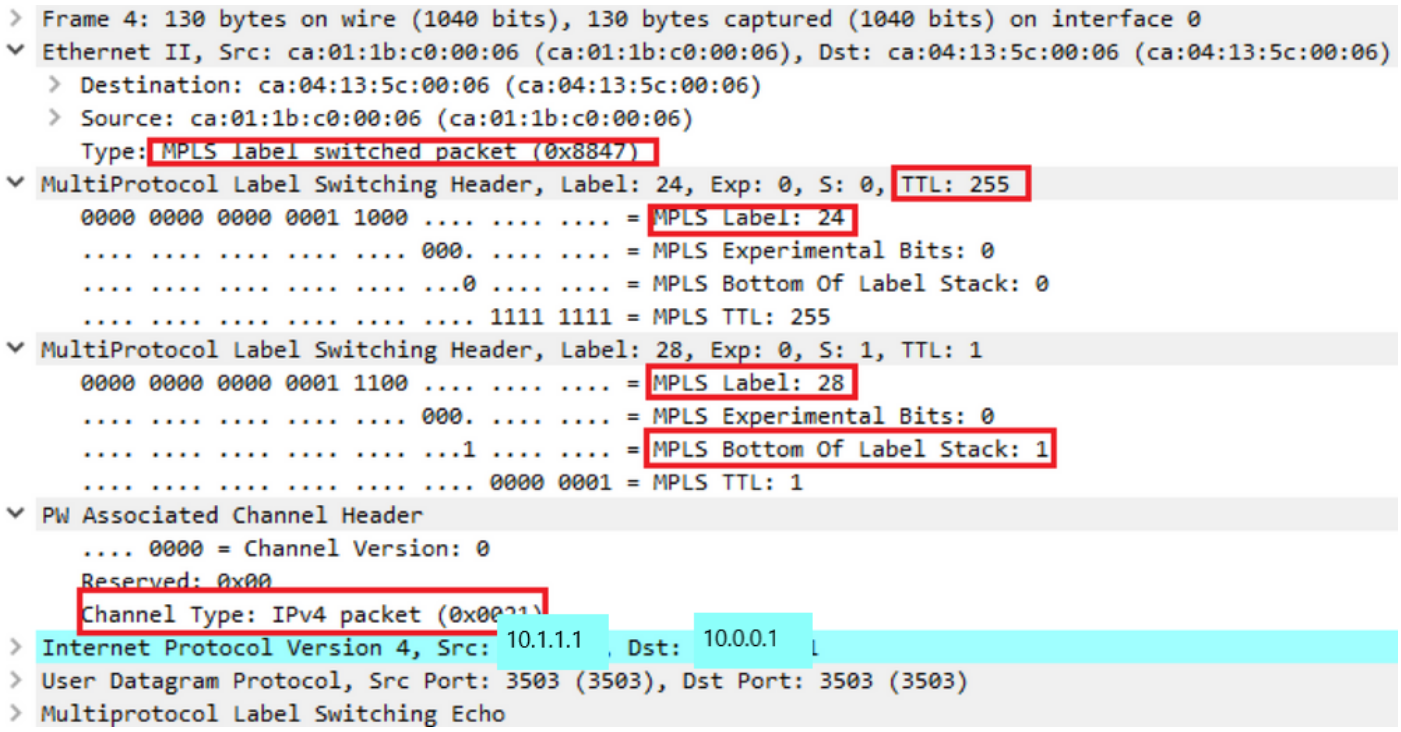


```
4 0.203148   10.1.1.1   10.0.0.1   MPLS E…   130 MPLS Echo Request
```



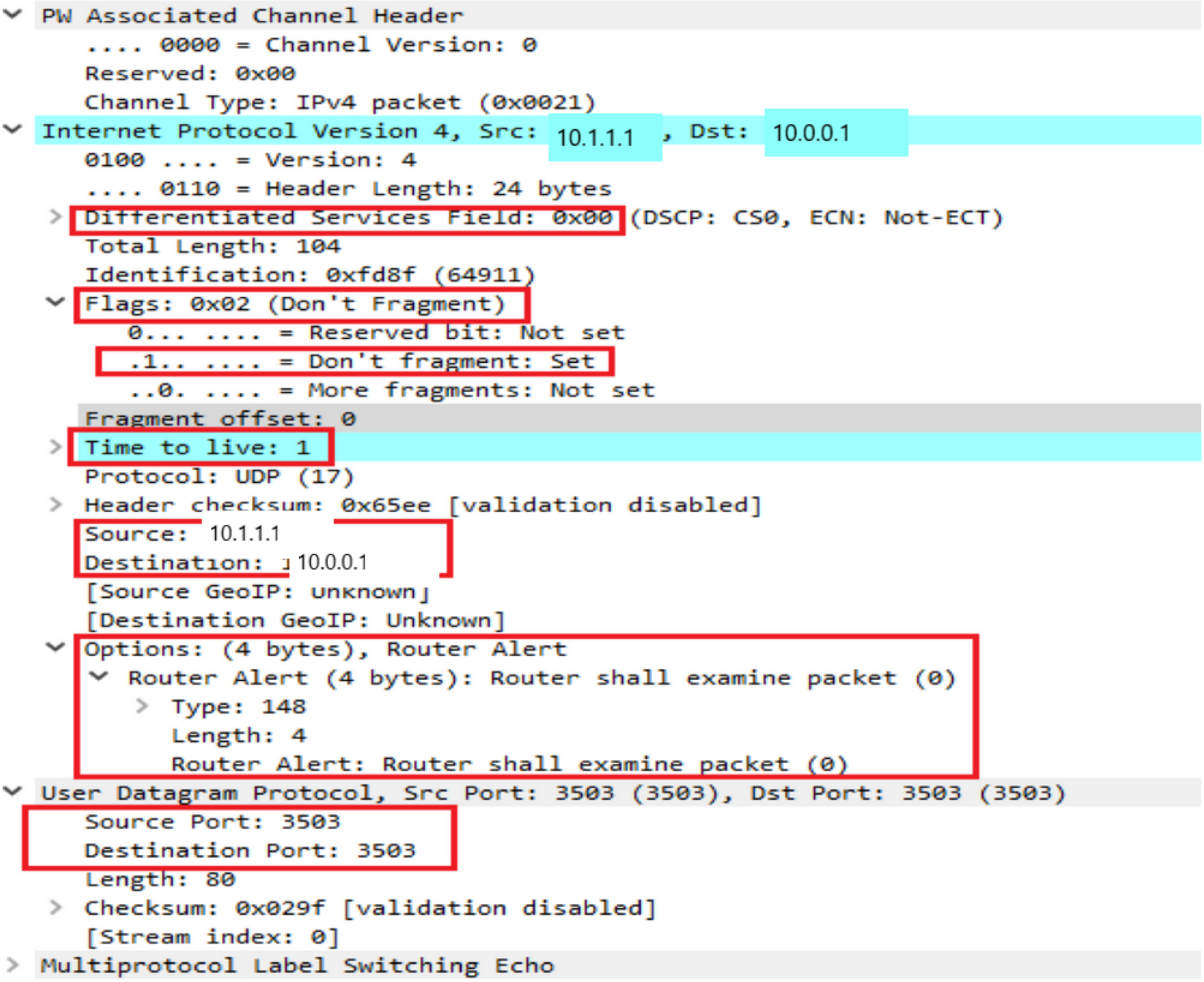
```
Frame 2: 130 bytes on wire (1040 bits), 130 bytes captured (1040 bits) on interface 0
Ethernet II, Src: ca:01:1b:c0:00:06 (ca:01:1b:c0:00:06), Dst: ca:04:13:5c:00:06 (ca:04:13:5c:00:06)
MultiProtocol Label Switching Header, Label: 24, Exp: 0, S: 0, TTL: 255     Transport label
MultiProtocol Label Switching Header, Label: 28, Exp: 0, S: 1, TTL: 1       VPN label
PW Associated Channel Header
Internet Protocol Version 4, Src: 10.1.1.1  Dst: 10.0.0.1
User Datagram Protocol, Src Port: 3503 (3503), Dst Port: 3503 (3503)
Multiprotocol Label Switching Echo
```

Layer 2/Labels:



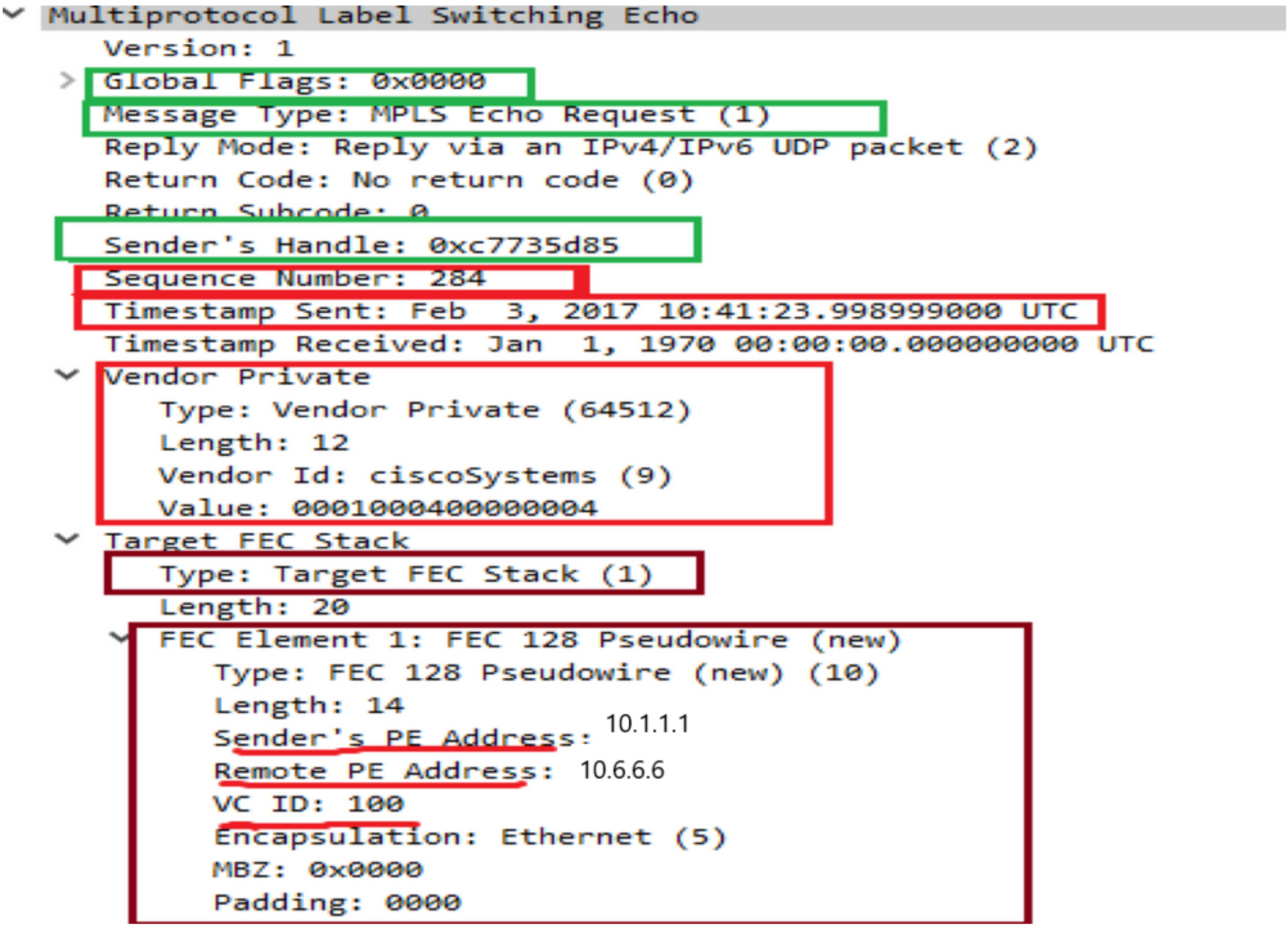
```
> Frame 4: 130 bytes on wire (1040 bits), 130 bytes captured (1040 bits) on interface 0
v Ethernet II, Src: ca:01:1b:c0:00:06 (ca:01:1b:c0:00:06), Dst: ca:04:13:5c:00:06 (ca:04:13:5c:00:06)
  > Destination: ca:04:13:5c:00:06 (ca:04:13:5c:00:06)
  > Source: ca:01:1b:c0:00:06 (ca:01:1b:c0:00:06)
    Type: MPLS label switched packet (0x8847)
v MultiProtocol Label Switching Header, Label: 24, Exp: 0, S: 0, TTL: 255
    0000 0000 0000 0001 1000 .... .... .... = MPLS Label: 24
    .... .... .... .... .... 000. .... .... = MPLS Experimental Bits: 0
    .... .... .... .... .... ...0 .... .... = MPLS Bottom Of Label Stack: 0
    .... .... .... .... .... .... 1111 1111 = MPLS TTL: 255
v MultiProtocol Label Switching Header, Label: 28, Exp: 0, S: 1, TTL: 1
    0000 0000 0000 0001 1100 .... .... .... = MPLS Label: 28
    .... .... .... .... .... 000. .... .... = MPLS Experimental Bits: 0
    .... .... .... .... .... ...1 .... .... = MPLS Bottom Of Label Stack: 1
    .... .... .... .... .... .... 0000 0001 = MPLS TTL: 1
v PW Associated Channel Header
    .... 0000 = Channel Version: 0
    Reserved: 0x00
    Channel Type: IPv4 packet (0x00
> Internet Protocol Version 4, Src: 10.1.1.1 , Dst: 10.0.0.1
> User Datagram Protocol, Src Port: 3503 (3503), Dst Port: 3503 (3503)
> Multiprotocol Label Switching Echo
```

L3/L4:



```
v PW Associated Channel Header
    .... 0000 = Channel Version: 0
    Reserved: 0x00
    Channel Type: IPv4 packet (0x0021)
v Internet Protocol Version 4, Src: 10.1.1.1 , Dst: 10.0.0.1
    0100 .... = Version: 4
    .... 0110 = Header Length: 24 bytes
  > Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    Total Length: 104
    Identification: 0xfd8f (64911)
  v Flags: 0x02 (Don't Fragment)
      0... .... = Reserved bit: Not set
      .1.. .... = Don't fragment: Set
      ..0. .... = More fragments: Not set
    Fragment offset: 0
  > Time to live: 1
    Protocol: UDP (17)
  > Header checksum: 0x65ee [validation disabled]
    Source: 10.1.1.1
    Destination: 10.0.0.1
    [Source GeoIP: Unknown]
    [Destination GeoIP: Unknown]
  v Options: (4 bytes), Router Alert
    v Router Alert (4 bytes): Router shall examine packet (0)
      > Type: 148
        Length: 4
        Router Alert: Router shall examine packet (0)
v User Datagram Protocol, Src Port: 3503 (3503), Dst Port: 3503 (3503)
    Source Port: 3503
    Destination Port: 3503
    Length: 80
  > Checksum: 0x029f [validation disabled]
    [Stream index: 0]
> Multiprotocol Label Switching Echo
```

The actual MPLS payload:



```
Multiprotocol Label Switching Echo
    Version: 1
  > Global Flags: 0x0000
    Message Type: MPLS Echo Request (1)
    Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
    Return Code: No return code (0)
    Return Subcode: 0
    Sender's Handle: 0xc7735d85
    Sequence Number: 284
    Timestamp Sent: Feb  3, 2017 10:41:23.998999000 UTC
    Timestamp Received: Jan  1, 1970 00:00:00.000000000 UTC
  Vendor Private
        Type: Vendor Private (64512)
        Length: 12
        Vendor Id: ciscoSystems (9)
        Value: 0001000400000004
  Target FEC Stack
        Type: Target FEC Stack (1)
        Length: 20
      FEC Element 1: FEC 128 Pseudowire (new)
            Type: FEC 128 Pseudowire (new) (10)
            Length: 14
            Sender's PE Address: 10.1.1.1
            Remote PE Address: 10.6.6.6
            VC ID: 100
            Encapsulation: Ethernet (5)
            MBZ: 0x0000
            Padding: 0000
```

2. Echo Reply:

Can carry 1 Label – Transport.

Sent as UNICAST PACKET. This can be label switched (with Transport Label) because of LDP in a core.

LABELS:1
SRC IP: EXIT INTERFACE IP ADDRESS (10.1.6.2 in our case)
DST IP: SOURCE IP SEEN IN ECHO REQUEST - LOOPBACK OF SOURCE ROUTER
L4 TYPE: UDP
SRC PORT:3503
DST PORT:3505
TOS BYTE: OFF
MPLS EXP: OFF
DF BIT: ON

UDP PAYLOAD can be MPLS LABEL SWITCHING ECHO REPLY

MPLS EXP is ON and SET to 6

DF BIT is ON

VC details for reference:

<#root>

```
R1#sh mpls l2transport vc detail

Local interface: Fa2/0 up, line protocol up, Ethernet up

  Destination address: 10.6.6.6

,

VC ID: 100, VC status: up

    Output interface: Fa0/1, imposed label stack {24 28}

    Preferred path: not configured

    Default path: active

    Next hop: 10.1.1.2

  Create time: 2d17h, last status change time: 2d17h

    Last label FSM state change time: 2d17h

  Signaling protocol: LDP, peer 10.6.6.6:0 up

    Targeted Hello: 10.1.1.1(LDP Id) -> 10.6.6.6, LDP is UP

    Status TLV support (local/remote)   : enabled/supported

      LDP route watch                   : enabled

      Label/status state machine        : established, LruRru

      Last local dataplane   status rcvd: No fault

      Last BFD dataplane     status rcvd: Not sent

      Last BFD peer monitor  status rcvd: No fault

      Last local AC  circuit status rcvd: No fault

      Last local AC  circuit status sent: No fault

      Last local PW i/f circ status rcvd: No fault

      Last local LDP TLV     status sent: No fault

      Last remote LDP TLV    status rcvd: No fault

      Last remote LDP ADJ    status rcvd: No fault

    MPLS VC labels: local 28, remote 28

    Group ID: local 0, remote 0

    MTU: local 1500, remote 1500

    Remote interface description:

  Sequencing: receive enabled, send enabled
  Sequencing resync disabled

  Control Word: On (configured: autosense)

  Dataplane:

    SSM segment/switch IDs: 4097/4096 (used), PWID: 1

  VC statistics:

    transit packet totals: receive 1027360, send 1027358

    transit byte totals:   receive 121032028, send 147740215

    transit packet drops:  receive 0, seq error 0, send 0
```

L2VPN Interworking

L2VPN Interworking builds on this functionality by allowing disparate attachment circuits to be connected. An interworking function facilitates the translation between different Layer 2 encapsulations. In earlier releases, the Cisco series router supported only bridged interworking, which is also known as Ethernet interworking.

Up to this point in this, the AC on both the sides has been the same encapsulation type, which is also referred to as like-to-like functionality.

L2VPN interworking is AToM feature allows different encapsulation type at both sides of the AToM network

- It is required to interconnect two heterogeneous attachment circuits (ACs).
- The two main L2VPN interworking (IW) functions supported in Cisco IOS Software are:

1. IP/Routed:MAC header is removed (and replaced with MPLS labels) at one end of the MPLS cloud and a new MAC header is constructed at the other PE. The IP header is retained as it is.

2. Ethernet/Bridged: MAC header is not removed at all. The MPLS labels are imposed on top of the MAC header and the MAC header is delivered as is to the other end of the MPLS cloud.

Interworking Possibilities

a. FR to Ethernet

b. FR to PPP

c. FR to ATM

d. Ethernet to VLAN

e. Ethernet to PPP

Related Information

- RFC Editor 4664
- RFC Editor 4667

- Technical Support & Documentation - Cisco Systems