Multi Protocol Label Switching (MPLS)

Anjali Agarwal

Why do we need MPLS

- to address the problems faced by present-day networks—speed, scalability, quality-of-service (QoS) management, and traffic engineering
- Applying QoS on a flow-by-flow basis is not practical due to the huge numbers of IP traffic flows in carrier-sized networks.
- most of the routing protocols deployed today are based on algorithms designed to obtain the shortest path in the network for packet traversal and do not take into account additional metrics (such as delay, jitter, and traffic congestion), which can further diminish network performance.
- Layer-2 switching devices addressed the switching bottlenecks within the subnets of a local-area network (LAN) environment.
- Layer-3 switching devices helped alleviate the bottleneck in Layer-3 routing by moving the route lookup for Layer-3 forwarding to high-speed switching hardware.

What is MPLS

- MPLS is an Internet Engineering Task Force (IETF)–specified framework that provides for the efficient designation, routing, forwarding, and switching of traffic flows through the network.

- MPLS is a key development in Internet technologies that will assist in adding a number of essential capabilities to today's best effort IP networks, including

  - *Traffic Engineering*

  - *Providing traffic with different qualitative Classes of Service (CoS)*
  - *Providing traffic with different quantitative Quality of Service (QoS)*
  - *Providing IP based Virtual Private Networks (VPN's)*
- MPLS assists in addressing the ever-present scaling issues faced by the Internet as it continues to grow, and to address issues related to routing (based on QoS and service quality metrics)

MPLS functions

- specifies mechanisms to manage traffic flows of various granularities, such as flows between different hardware, machines, or even flows between different applications

- remains independent of the Layer-2 and Layer-3 protocols

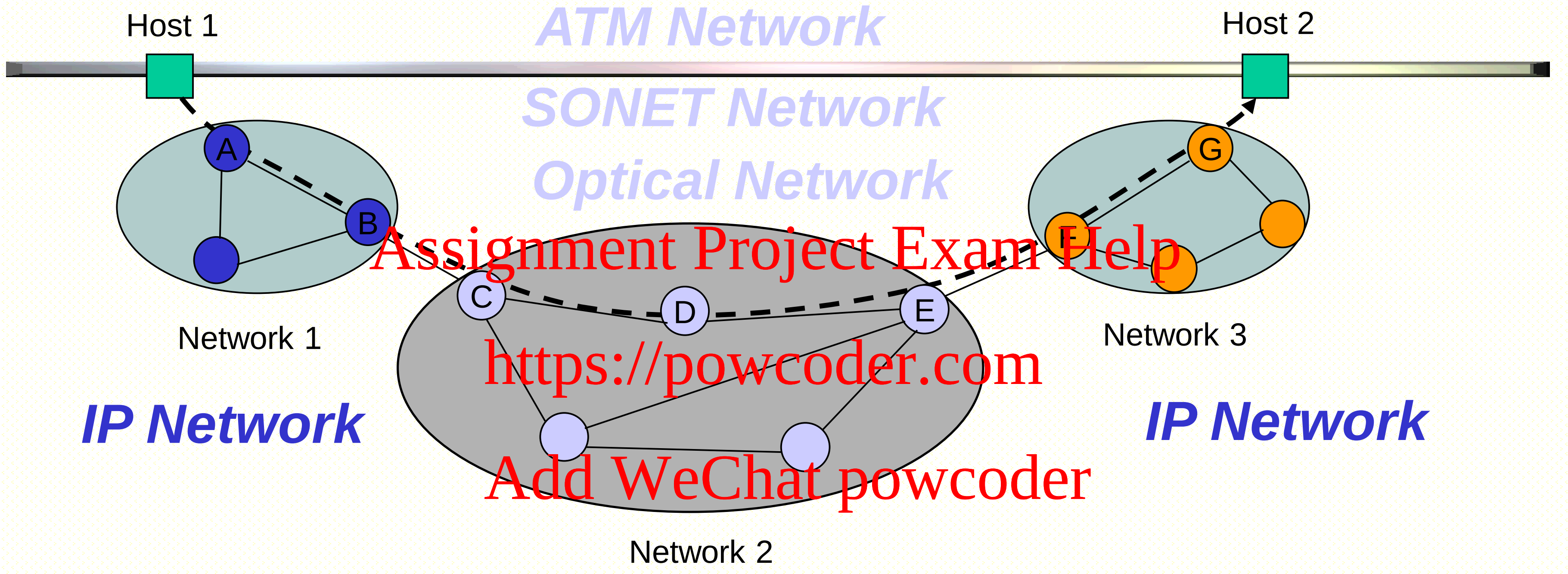
- provides a means to map IP addresses to simple, fixed-length labels used by different packet-forwarding and packet-switching technologies

- interfaces to existing routing protocols such as resource reservation protocol (RSVP) and open shortest path first (OSPF)
- supports the IP, ATM, and frame-relay Layer-2 protocols (glueing connectionless IP to connection-oriented networks)

What problems does it solve

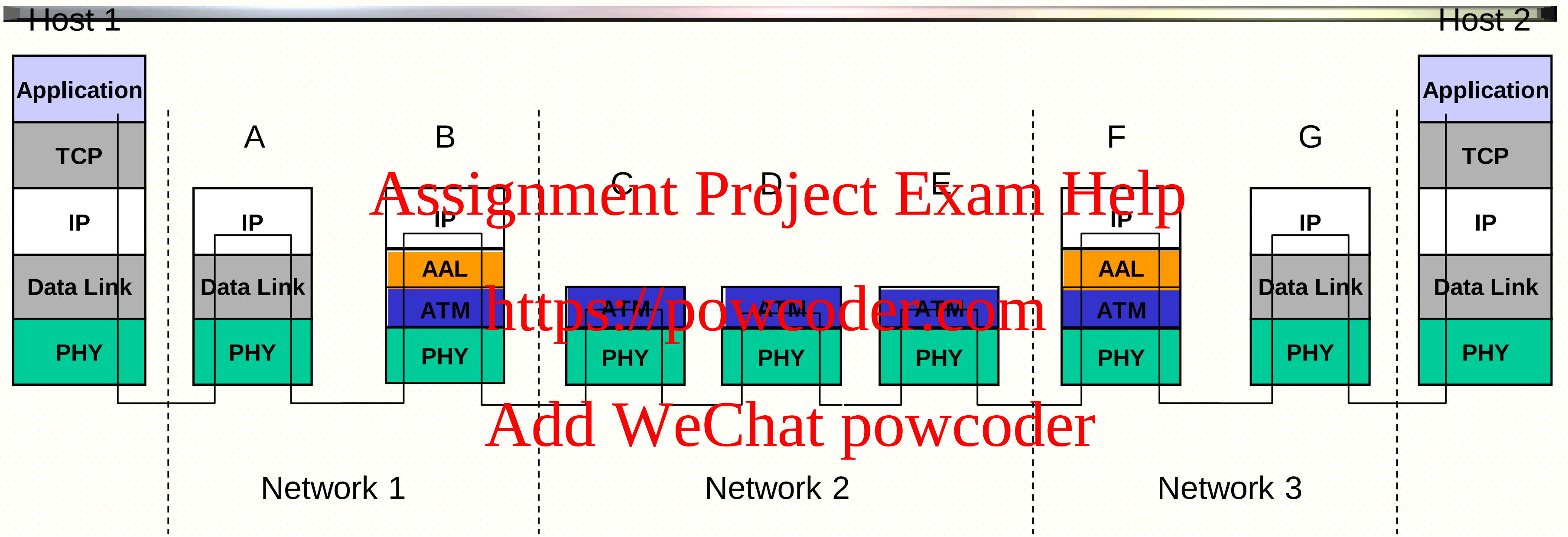
- The goal is to bring the speed of Layer 2 switching to Layer 3
- Routers make forwarding decisions based on the contents of a simple label, rather than by performing a complex route lookup based on destination IP address
- Elimination of multiple layers – typically an overlay model is employed where ATM is used at layer 2 to provide high-speed connectivity,and IP is used at layer 3 to provide the intelligence to forward IP datagrams.
  - » *complex mapping between two distinct architectures (connectionless vs. connection-oriented) that require the definition and maintenance of separate topologies, address spaces, routing protocols, signaling protocols, and resource allocation schemes*
- combining Layer 2 switching and Layer 3 routing into a fully integrated solution and eliminating inherent "cell-tax" in carrying IP traffic over ATM

Network Interconnection



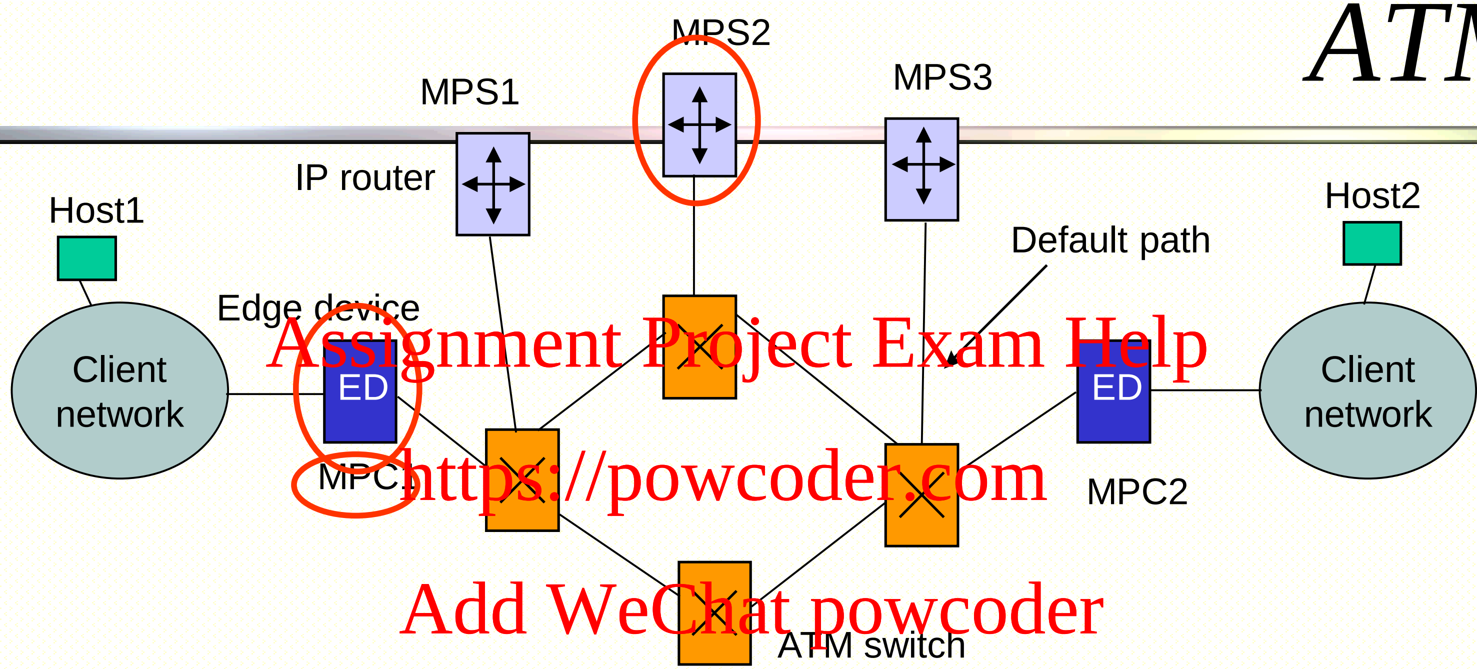
- *Server network* (Network 2) provides transport service to *Client networks* (Network 1 & Network 3)
- Control Plane Issues:
  - » *Server network & client networks may use different technologies*
  - » *What signaling is used and how are paths determined?*

End-to-End Protocol Stacks



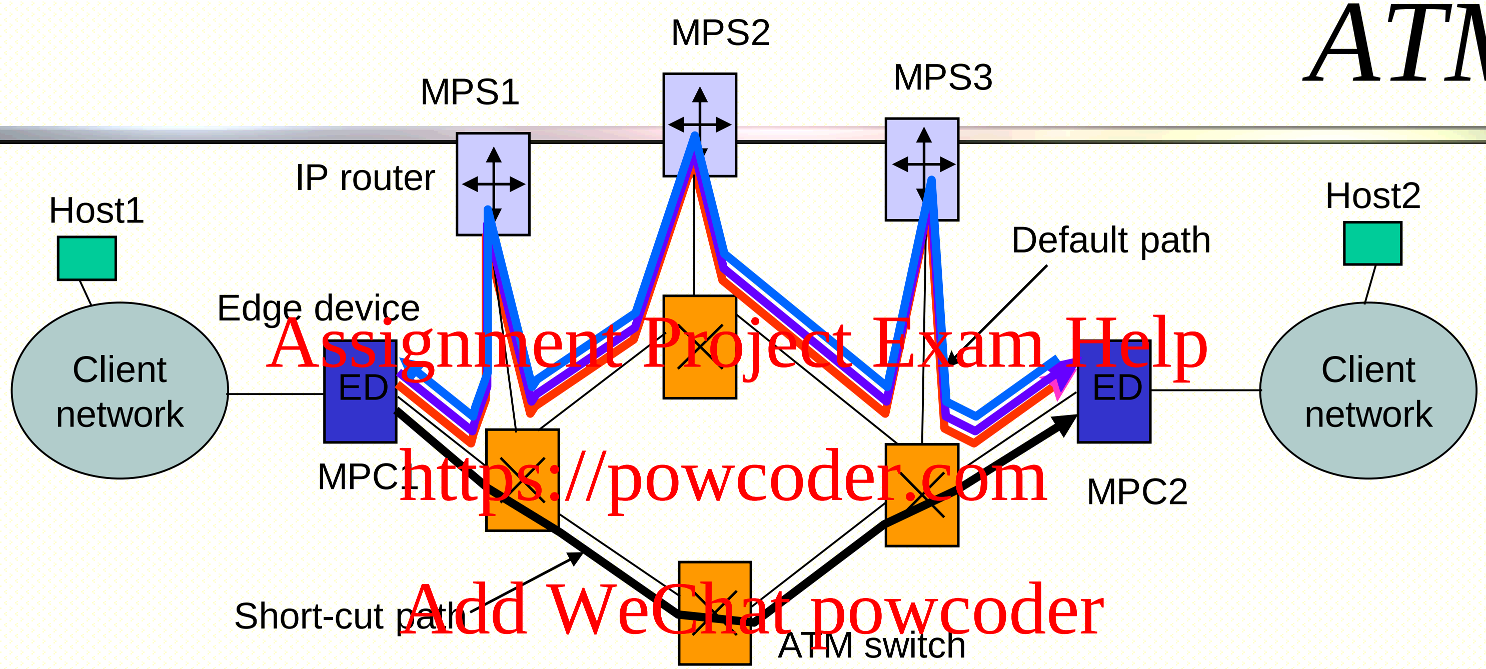
- Example: IP over ATM
- Hosts run TCP/IP
- Client networks are IP networks
- Server network is ATM

Overlay Example: IP over ATM



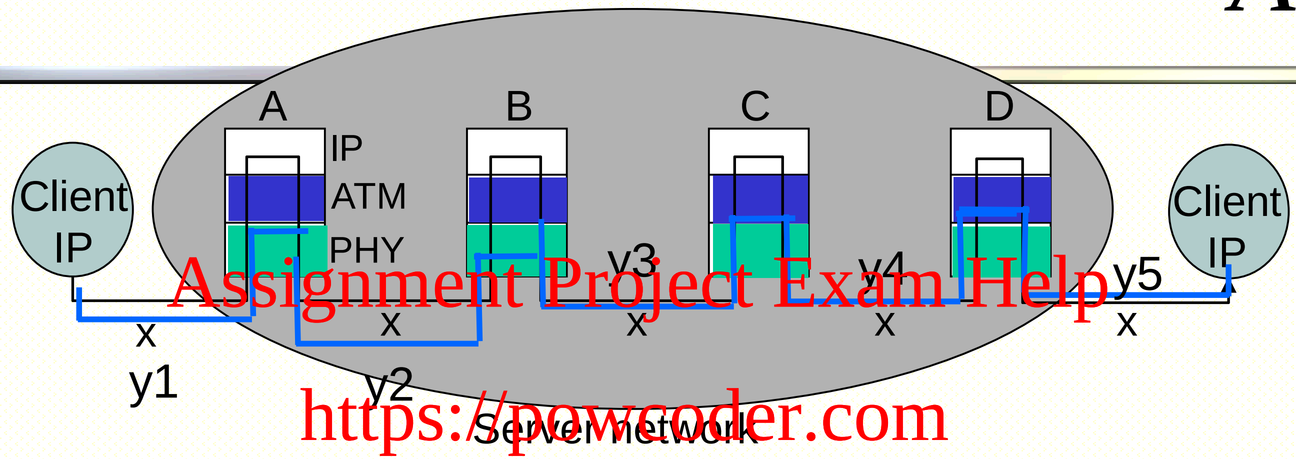
- Multiprotocol over ATM (MPOA) uses overlay approach
- *Edge Device (ED)* interposed between IP net & ATM net
- ED contains *MPOA client (MPC)* to set up & release VCs
- ATM has *MPOA servers (MPS)* for IP-ATM address resolution & IP packet forwarding

Overlay Example: IP over ATM



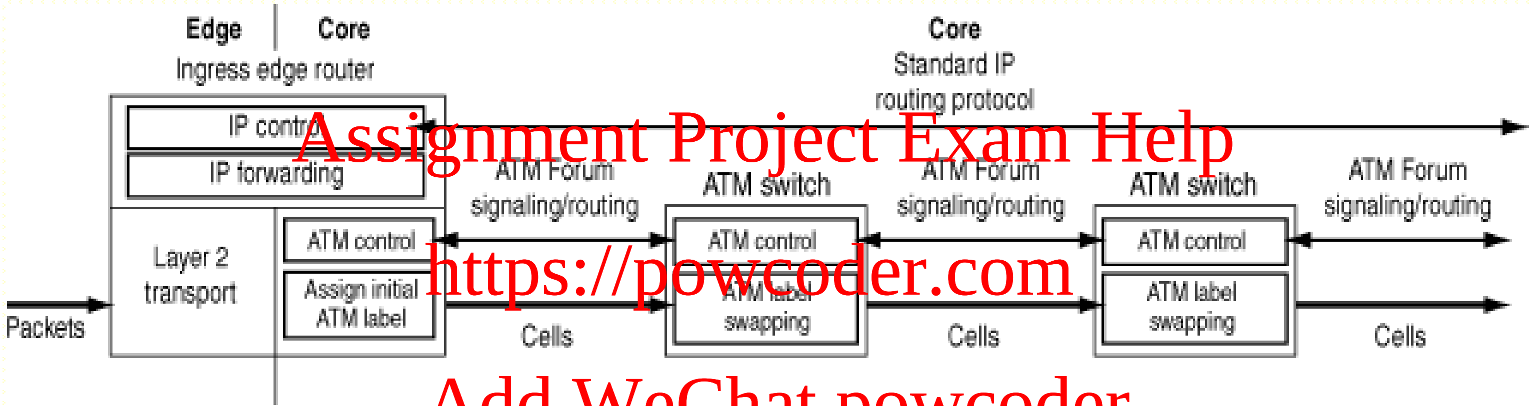
- First packets from Host 1 to Host 2 are routed using MPSs
- Ingress ED monitors packet flows
- When "long-lived" flow detected, MPD decides to set up VC
  - » *Sends ARP request to perform ATM address discovery of the corresponding egress MPC, which is routed along routed path*
  - » *Reply informs ingress ED of egress ED's ATM address*
  - » *VC set up & subsequent packet use ATM shortcut*

Peer-to-Peer Example: IP + ATM



- Nodes combine ATM switching & IP routing
- Initially packets are routed, hop by hop
  - » *Packets flow along default VCs "x"*
- When long-lived flow detected, node sets up shortcut
  - » *Client establishes VC shortcut y1*
  - » *Node A establishes VC shortcut y2*
  - » *And so on*

IP over ATM model



- The role of IP routing is limited to the edges of the network
- Layer 3 functionality is at the edges of the network and maximized network throughput is by relying on high-speed, label-swapping ATM switches and PVCs in the core
- overly complex approach that requires two separate sets of equipment

Multilayer Switching Alternatives to IP-over-ATM Model

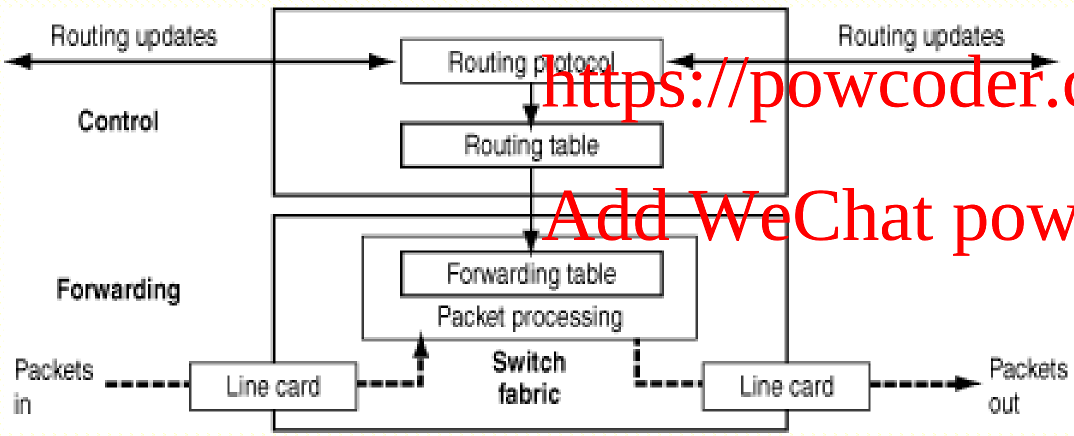
- By late 1996, number of vendors promoted proprietary multilayer switching solutions that integrated ATM switching and IP routing, including:
  - *IP Switching designed by Ipsilon/Nokia*
  - *Tag Switching developed by Cisco Systems*
  - *Aggregate Route-Based IP Switching (ARIS) designed by IBM*
  - *IP Navigator delivered by Cascade/Ascend/Lucent*
  - *Cell Switching Router (CSR) developed by Toshiba*
- These are not interoperable although they have a number of characteristics in common

Fundamental Building Blocks

Common to all multilayer switching solutions and MPLS:

» Separation of the control and forwarding components.

» Label-swapping forwarding algorithm.



♦ **The control component uses standard routing protocols (OSPF, IS-IS, and BGP-4) to exchange information with other routers to build and maintain a forwarding table**

♦ **When packets arrive, the forwarding component (based on a label-swapping forwarding algorithm), searches the forwarding table maintained by the control component to make a routing decision for each packet**

MPLS and Its Components

Label Switched Path (LSP):

- » are a sequence of labels at each and every node along the path from the source to the destination.

- » are established either prior to data transmission (control-driven) or upon detection of a certain flow of data (data-driven).

- »LSPs are simplex in nature (traffic flows in one direction from the head-end toward the tail-end), duplex traffic requires two LSPs, one LSP to carry traffic in each direction

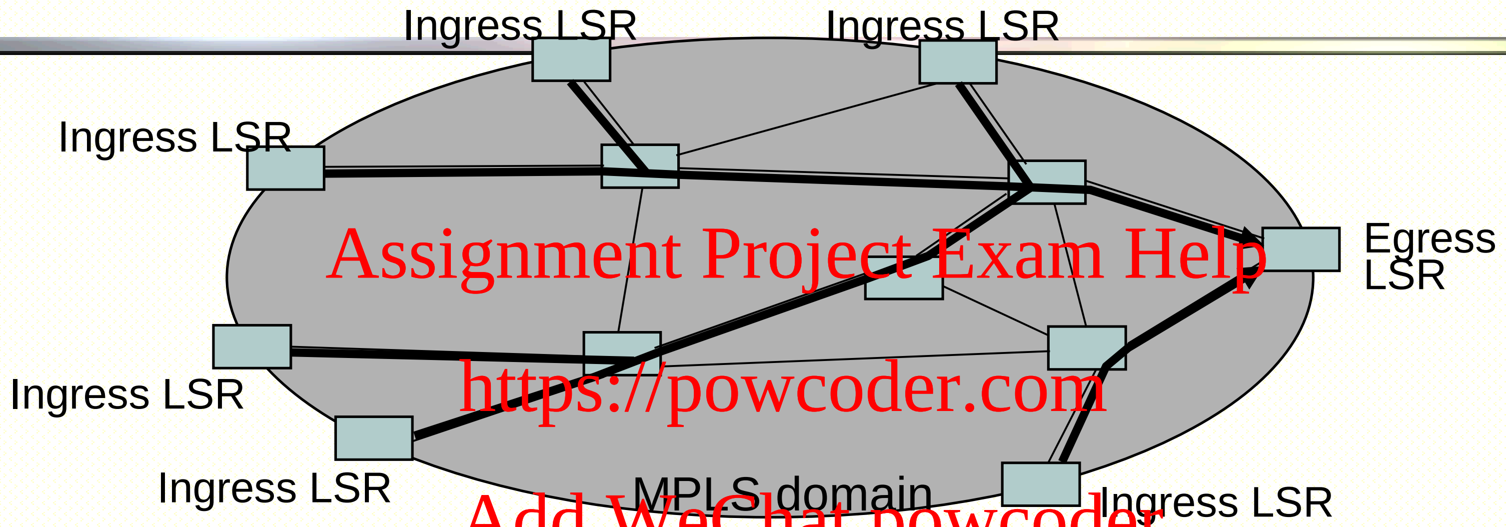
MPLS and Its Components

Label switching routers (LSRs) and Label edge routers (LERs):

- » LER operates at the edge of the access network and MPLS network and supports multiple ports connected to dissimilar networks (such as frame relay, ATM, and Ethernet)

- » LER plays a very important role in the assignment and removal of labels
- » LSR is in the core of an MPLS network and participates in the establishment of LSPs
- » LSR performs high-speed switching of the data traffic based on the established paths.

Labels and Paths



- Label-switched paths (LSPs) are *unidirectional*
- LSPs can be:
  - » *point-to-point*
  - » tree rooted in egress node *corresponds to shortest paths leading to a destination egress router*

MPLS and Its Components

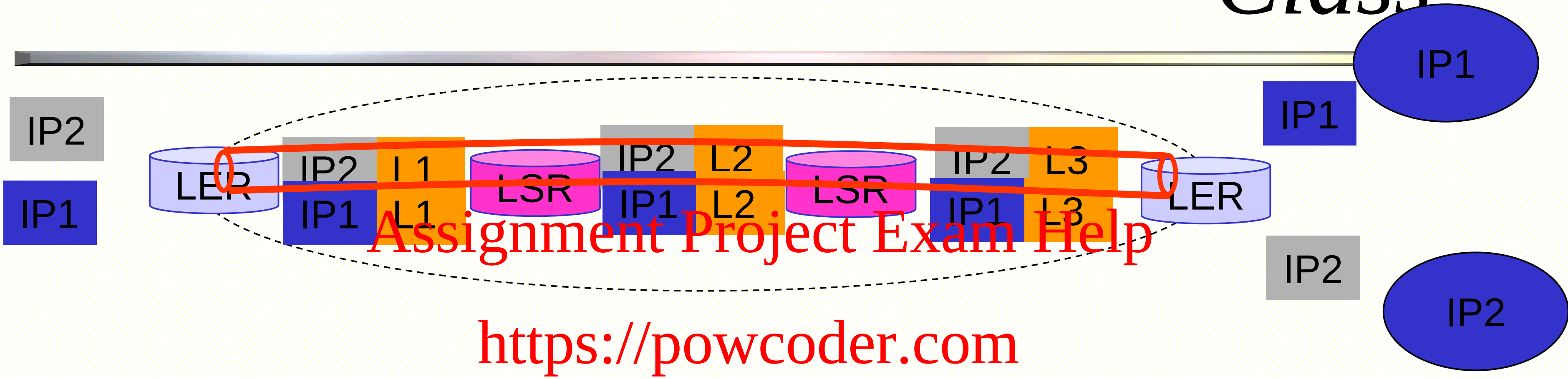
Forward equivalence class (FEC):

» group of packets that share the same requirements for their transport and are provided the same treatment en route to the destination

» a particular packet is assigned to a particular FEC just once, as the packet enters the network

» Each LSR builds a table, called a label information base (LIB), to specify how a packet must be forwarded, and is comprised of FEC–to-label bindings.

Forwarding Equivalence Class



- *FEC:* set of packets that are forwarded in the same manner
  - » *Over the same path, with the same forwarding treatment*
  - » *Packets in an FEC have same next-hop router*
  - » *Packets in same FEC may have different network layer header*
  - » *Each FEC requires a* single entry *in the forwarding table*
  - » *Coarse Granularity FEC: packets for all networks whose destination address matches a given address prefix*
  - » *Fine Granularity FEC: packets that belong to a particular application running between a pair of computers*

MPLS and Its Components

Labels and Label Bindings:

» A label identifies the path a packet should traverse

» A label is encapsulated in a Layer 2 technology supporting a label field such as the ATM VPI/VCI or the Frame Relay DLCI fields; or if the Layer 2 technology does not support a label field, the MPLS label is carried in a standardized MPLS header that is inserted between the Layer 2 and IP headers

» the packet journey through the backbone is based on label switching

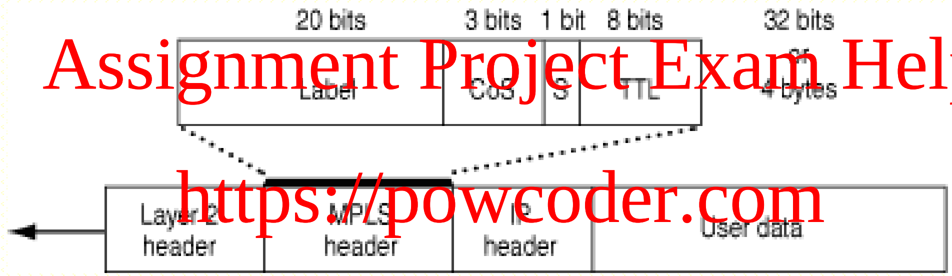
» label values are of local significance only; they pertain only to hops between LSRs

» Labels are bound to an FEC and their assignment decisions are based on forwarding criteria such as the following:

- destination unicast routing
- traffic engineering
- multicast
- QoS
- virtual private network (VPN)

MPLS and Its Components

MPLS Generic Label Format



- The label field (20-bits) carries the actual value of the MPLS label.
- The CoS field (3-bits) can affect the queuing and discard algorithms applied to the packet as it is transmitted through the network.
- The Stack (S) field (1-bit) supports a hierarchical label stack.
- The TTL (time-to-live) field (8-bits) provides conventional IP TTL functionality.

MPLS and Its Components

Label Creation

» topology-based method—uses normal processing of routing protocols

– *(such as OSPF and BGP)*

» request-based method—uses processing of request-based control traffic

– *(such as RSVP)*

» traffic-based method—uses the reception of a packet to trigger the assignment and distribution of a label

Label Distribution

» LDP—maps unicast IP destinations into labels, for explicit signaling and management of the label space

» RSVP, CR–LDP—used for traffic engineering and resource reservation, to support explicit routing based on QoS and CoS requirements

» protocol-independent multicast (PIM)—used for multicast states label mapping

» BGP—external labels (VPN)

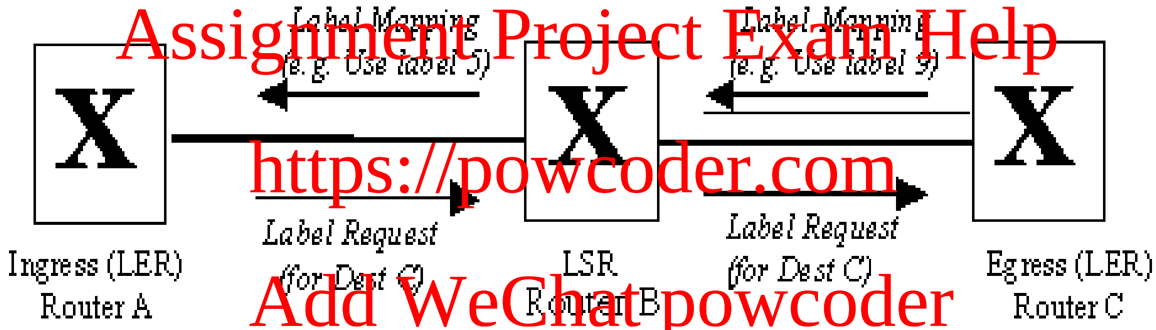
MPLS and Its Components

Setting up of LSPs

- *hop-by-hop routing —*
  - » Each LSR independently selects the next hop for a given FEC
  - » LSR uses any available routing protocols, such as OSPF, ATM's (PNNI)
- *explicit routing — similar to source routing*
  - » ingress LSR specifies the list of nodes through which the LSP traverses
  - » resources may be reserved along the path to ensure QoS

MPLS and Its Components

Signaling Mechanisms



- *an LSR requests a label from its downstream neighbor so that it can bind to a specific FEC*
- *In response to a label request, a downstream LSR will send a label to the upstream initiator using the label mapping mechanism*

MPLS Operation

Label creation and label distribution

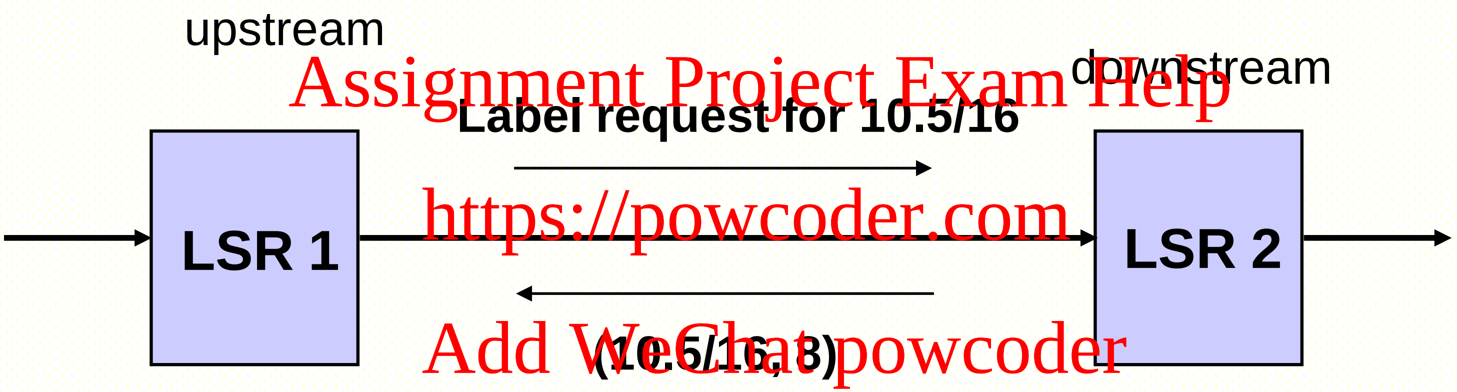
- Before any traffic begins the routers make the decision to bind a label to a specific FEC and build their tables.

- In LDP, downstream routers initiate the distribution of labels and the label/FEC binding.

- In addition, traffic-related characteristics and MPLS capabilities are negotiated using LDP.
- A reliable and ordered transport protocol should be used for the signaling protocol. LDP uses TCP.

Label Distribution

- Label Distribution Protocols distribute label bindings between LSRs



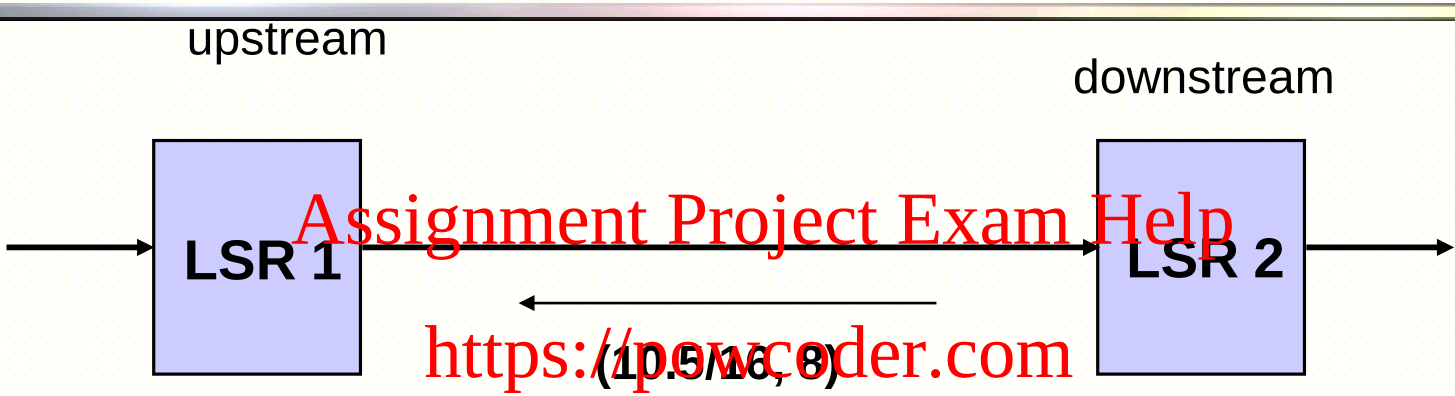
Downstream-on-Demand Mode

» *LSR1 becomes aware LSR2 is next-hop in an FEC*

» *LSR1 requests a label from LSR2 for given FEC*

» *LSR2 checks that it has next-hop for FEC, responds with label*

Label Distribution

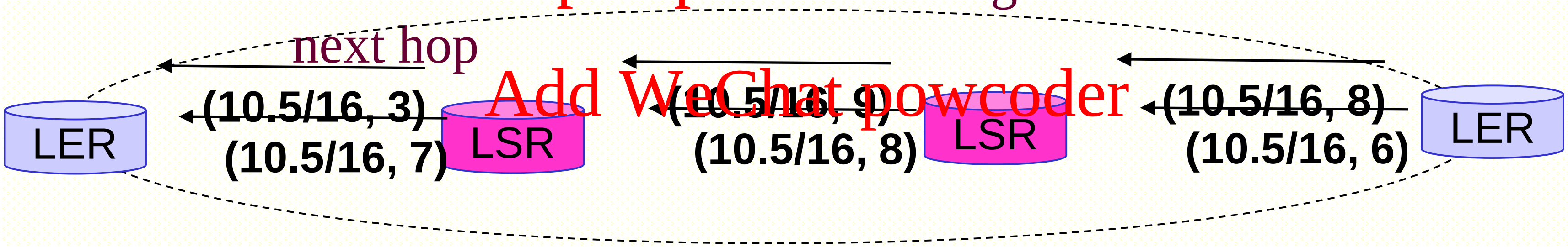


Downstream Unsolicited Mode

- *LSR2 becomes aware of a next hop for an FEC*
- *LSR2 creates a label for the FEC and forwards it to LSR1*
- *LSR2 can use this label if it finds that LSR2 is next-hop for that FEC*

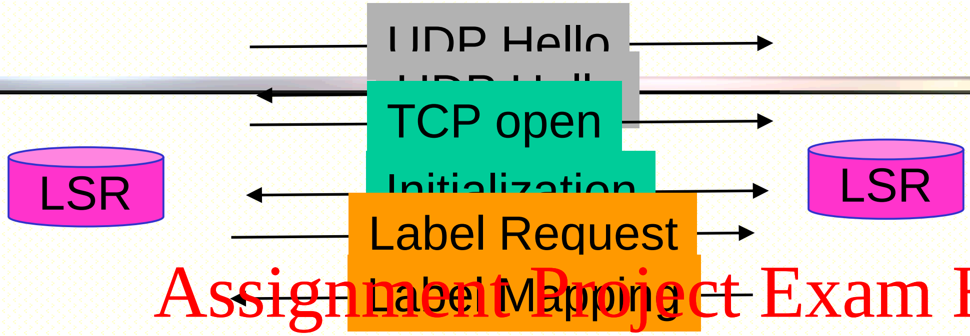
Independent vs. Order Label Distribution Control

- *Ordered Label Distribution Control*: LSR can distribute label if
  - – It is an egress LSR
  - – It has received FEC-label binding for that FEC from its next hop



- *Independent Label Distribution Control*: LSR independently binds FEC to label and distributes to its peers

Label Distribution Protocol



- *Label Distribution Protocol (LDP)*, RFC 3036
  - » *Topology-driven assignment (routes specified by routing protocol)*
  - » *Hello messages over UDP*
  - » *TCP connection & negotiation (session parameters & label distribution option, label ranges, valid timers)*
  - » *Message exchange (label request/mapping/withdraw)*

MPLS Operation (cont)

Table creation

- On receipt of label bindings each LSR creates entries in the label information base (LIB).

- The contents of the table will specify the mapping between a label and an FEC.

  - mapping between the input port and input label table to the output port and output label table.

  - The entries are updated whenever renegotiation of the label bindings occurs.

MPLS Operation (cont)

Label switched path creation

- the LSPs are created in the reverse direction to the creation of entries in the LIBs.

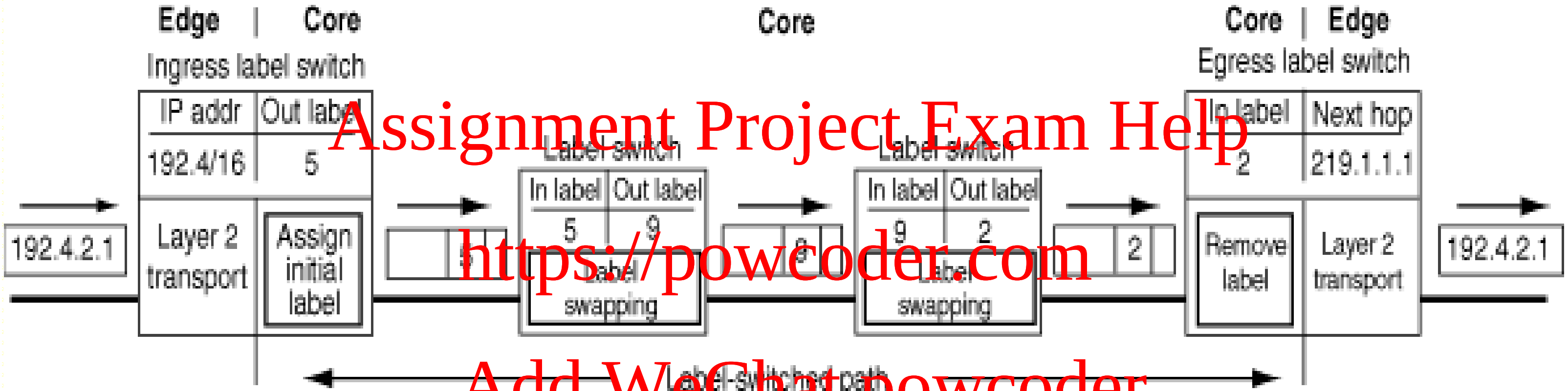
Label insertion/table-lookup

- The first router uses the LIB table to find the next hop and request a label for the specific FEC.

- Subsequent routers just use the label to find the next hop.
- Once the packet reaches the egress LSR, the label is removed and the packet is supplied to the destination.

MPLS Operation (cont)

Packet Traversing a Label Switched Path



the ingress label switch receives an unlabeled packet with a destination address of 192.4.2.1

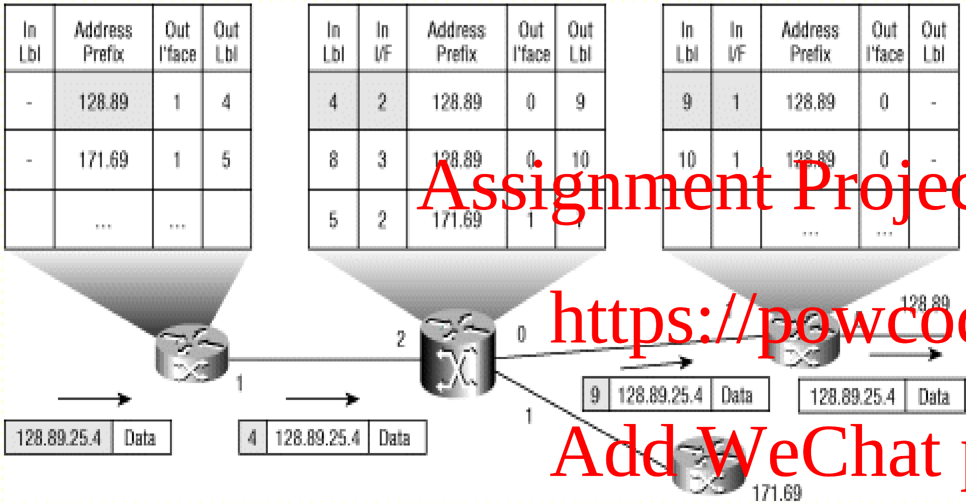
The label switch performs a longest-match routing table lookup and maps the packet to an FEC--192.4/16

The ingress label switch then assigns a label(with a value of 5) to the packet and forwards it to the next hop in the label-switched path (LSP)

Label switches ignore the packet's network layer header and simply forward the packet using the label-swapping algorithm

MPLS Packet Forwarding

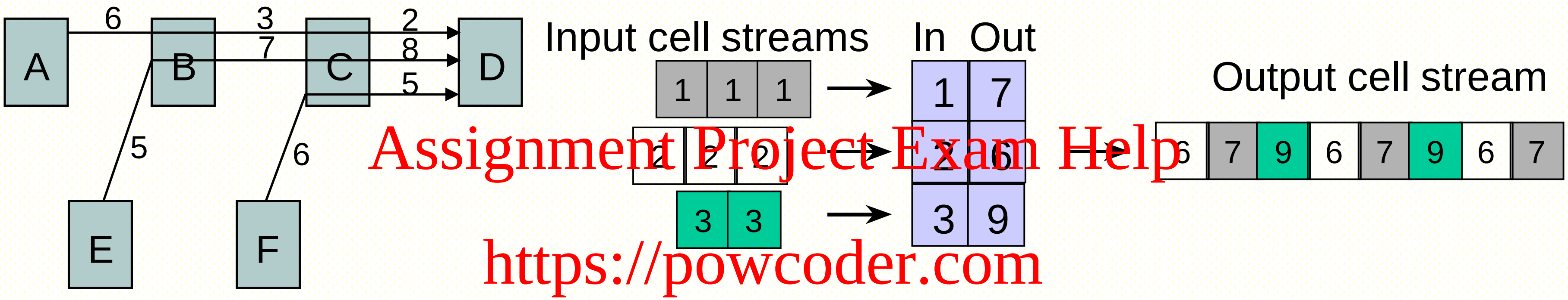
(another example)



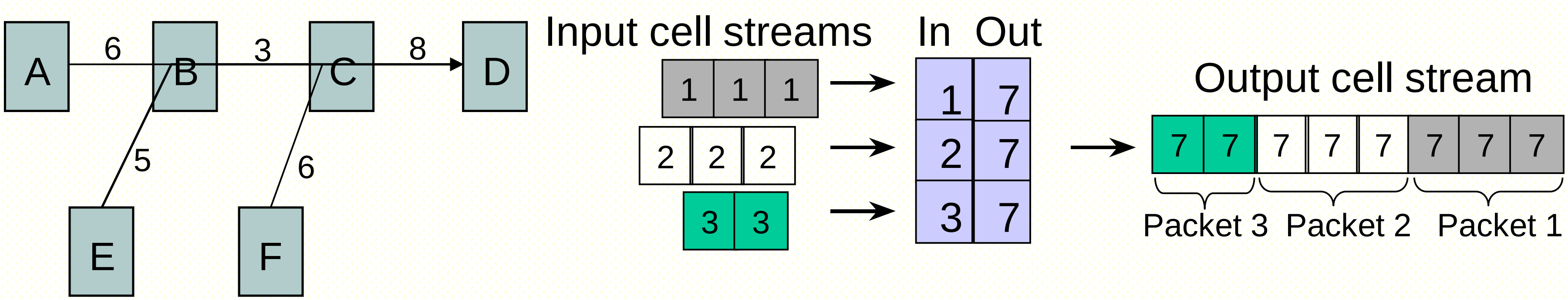
- an ingress packet arrives at the Edge LER, which reads the packet for the destination prefix, 128.89
- Edge LER looks up the destination address in the switching table and inserts the corresponding label 4, then forwards it out interface 1
- The LSR in the core reads the label and looks up its match in its switching table, then swaps incoming label with the outgoing label (label 4 with label 9), and forwards it out interface 0.
- The egress router reads and looks up label 9 in its table, which says to strip the label and forward the packet out interface 0.

VC Merging Conserves Labels

Non-VC merging



VC merging

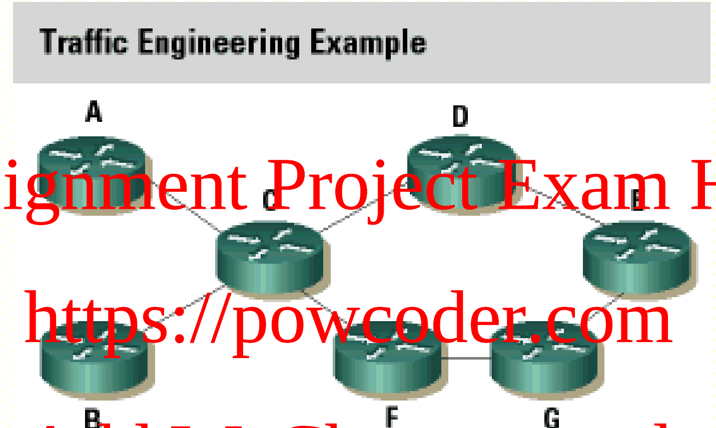


AAL5 End-of-Packet bit can be used to reassemble packets

MPLS APPLICATIONS
1. Traffic Engineering

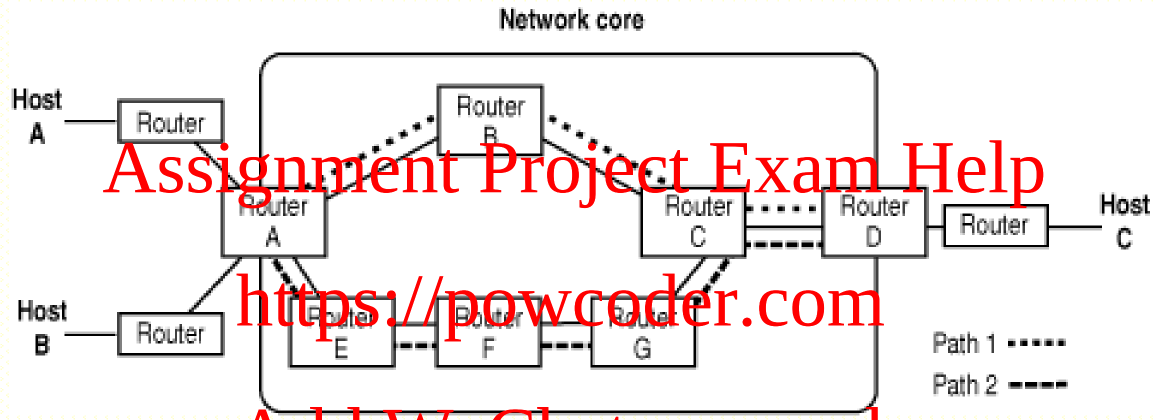
- refers to the ability to control where traffic flows in a network, with the goal of reducing congestion and getting the most use out of the available facilities.
- a way of managing traffic and link utilization in a routed network.

Traffic Engineering



- traffic engineering control module can establish a label-switched path from A to C to D to E and another from B to C to F to G to E.
- By defining policies that select certain packets to follow these paths, traffic flow across the network can be managed.
- the amount of load expected to flow between various points in the network (a traffic matrix) may be specified, and the routing system calculates the best paths to carry that load and establish explicit paths as a result.

Traffic Engineering

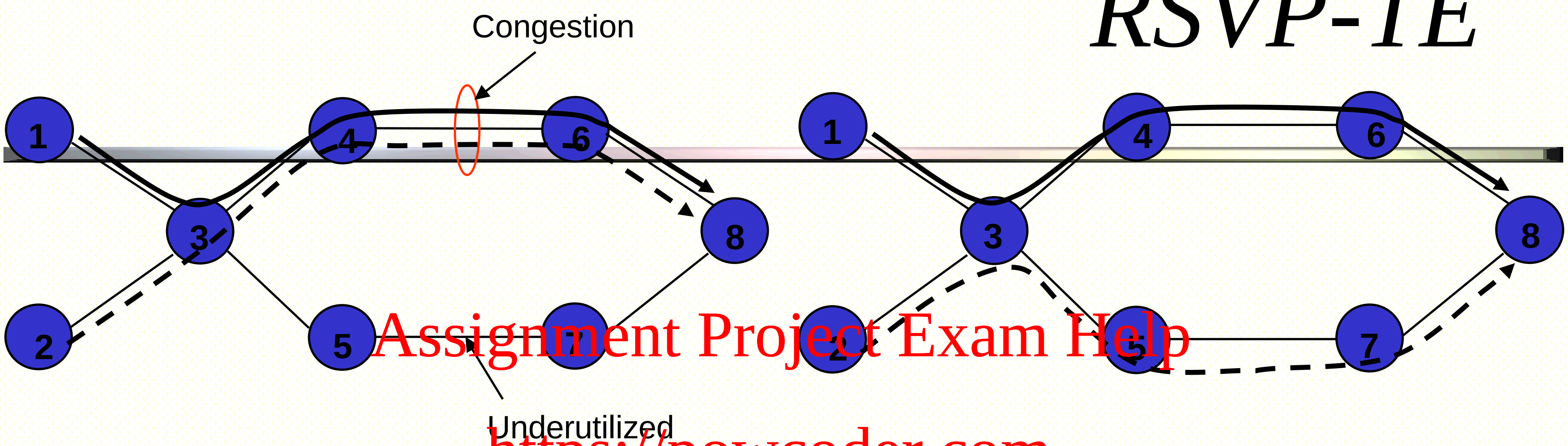


- Using conventional IP routing, traffic engineering cannot be implemented because all forwarding at Router A is based on the packet's destination address
- If core routers function as LSRs and LSP 1 and LSP 2 are configured as path 1 and path 2, MPLS provides ISPs an unprecedented level of control over traffic

Route Pinning

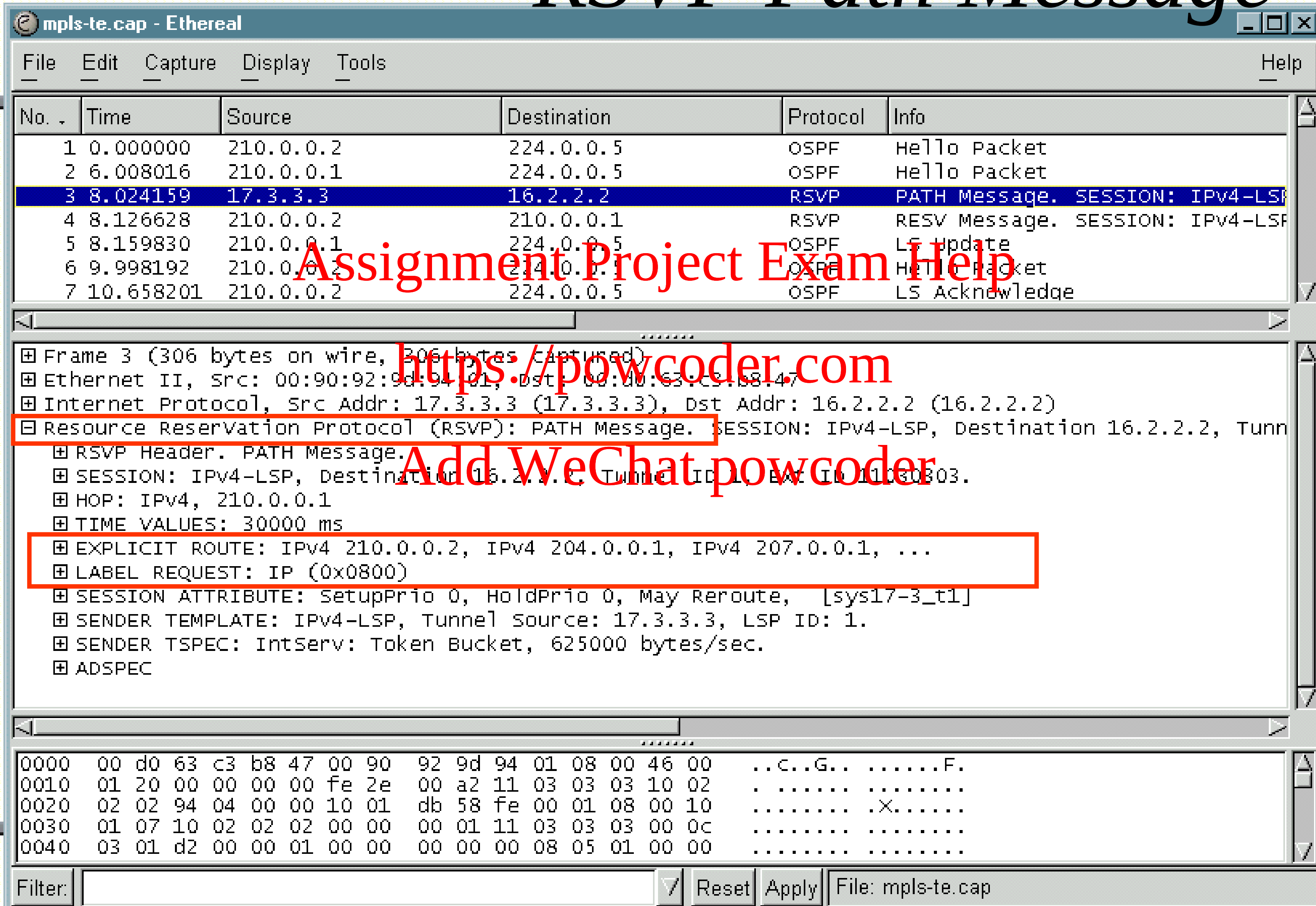
- Need for a specific and stable path through the network – route that has been *pinned*
  - » some applications are highly sensitive to changes in latency, an improvement in path may result in increasing/decreasing the latency
  - » LSP path does not change from the time it was established until it is disconnected

RSVP-TE

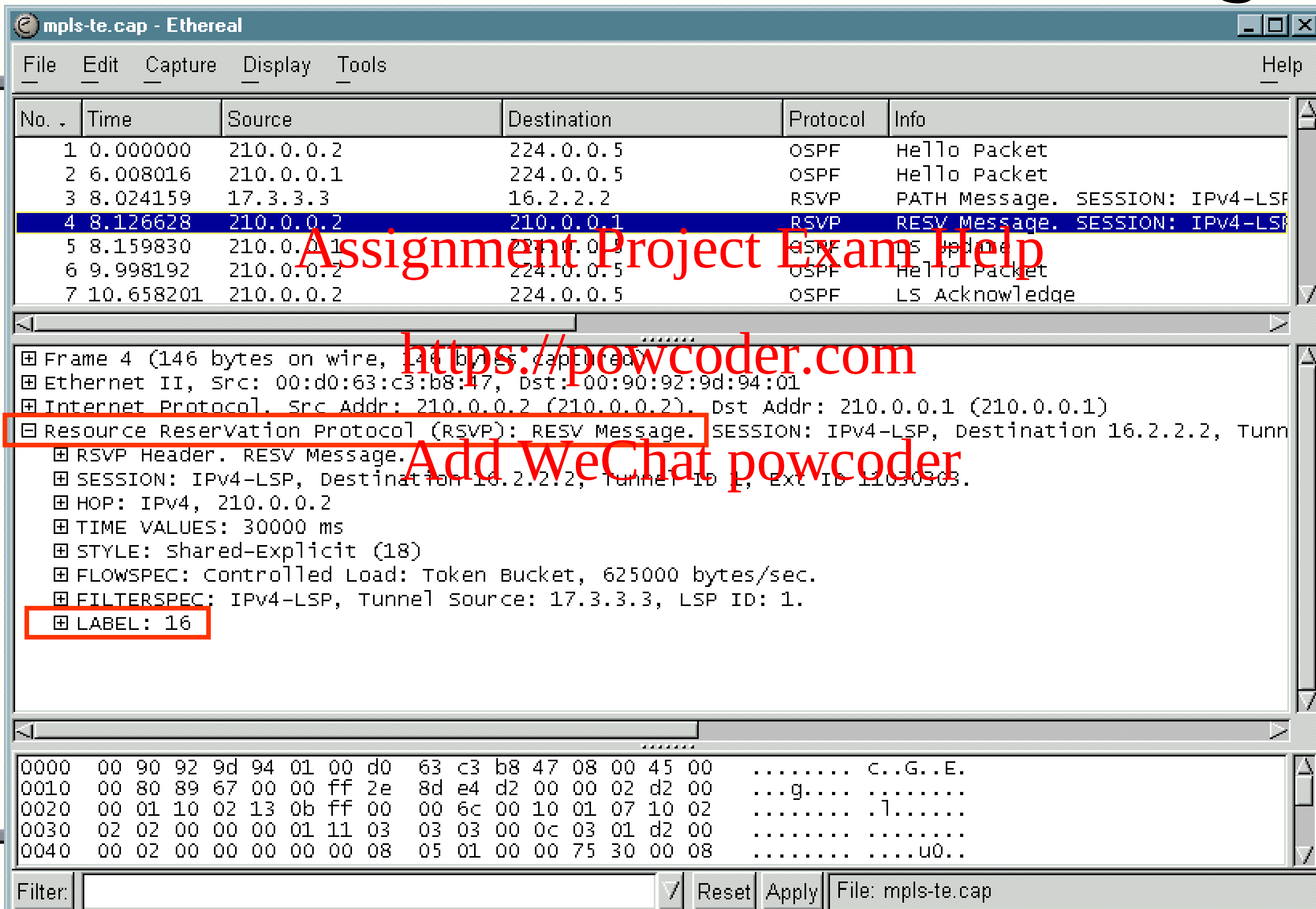


- Extensions to RSVP for *traffic-engineered LSPs*
  - » *Request-driven label distribution to create explicit route LSPs*
  - » *Single node (usually ingress) determines route*
  - » *Enables traffic engineering*
- RSVP Path message includes
  - » *label request object to request label binding*
  - » *Explicit route object (ERO)*
- RSVP Resv message includes label object

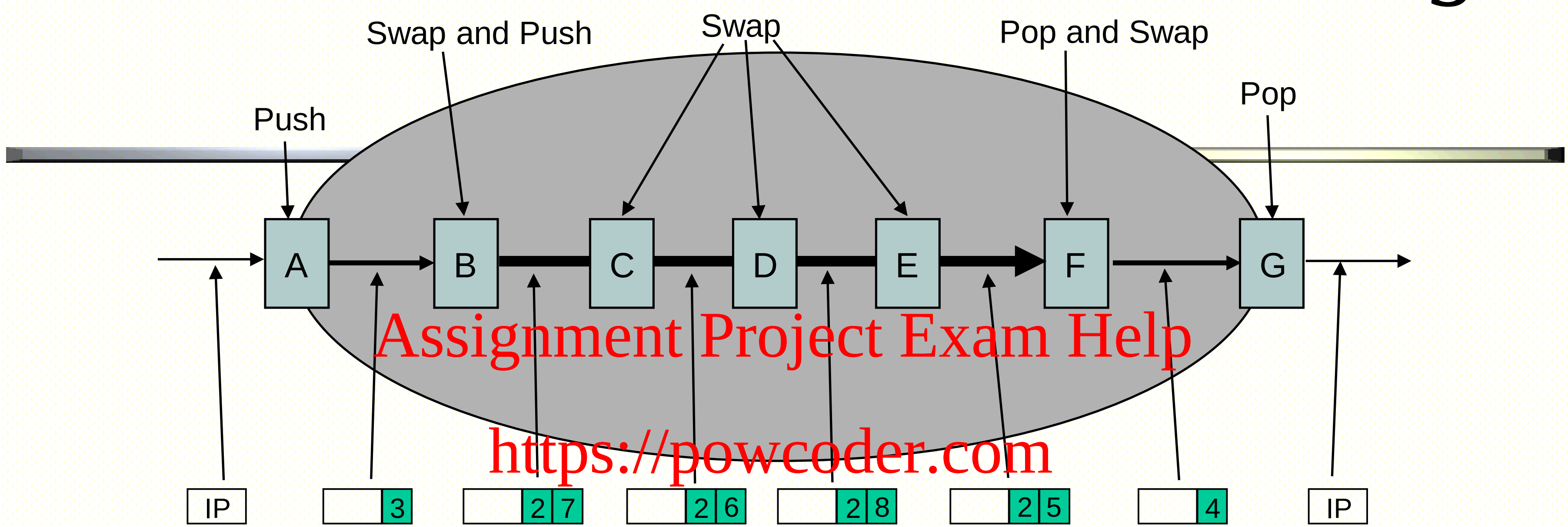
RSVP Path Message



RSVP Resv Message

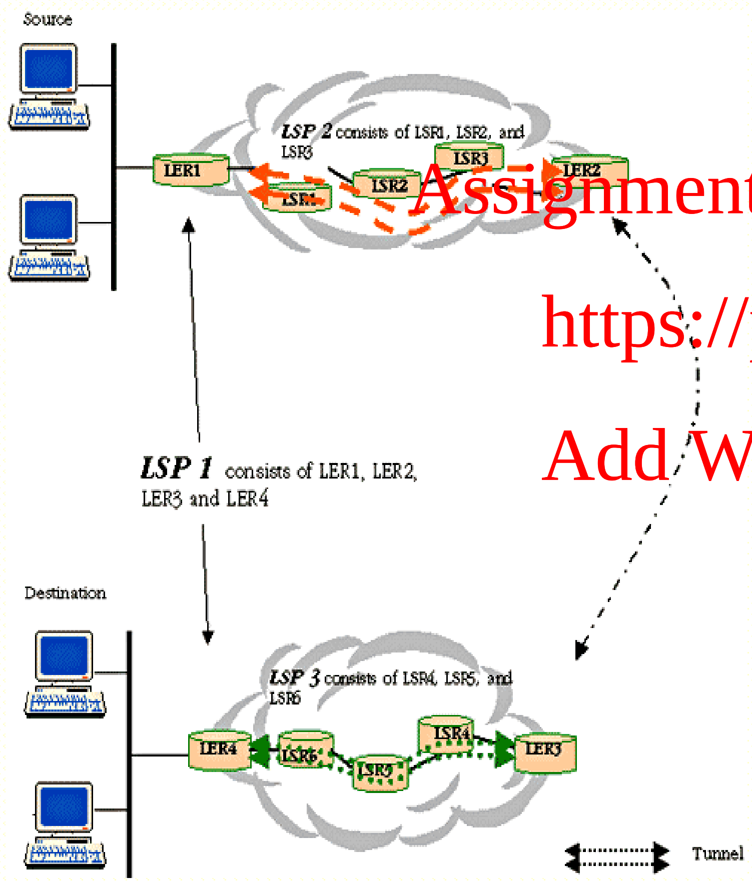


Label Stacking



- MPLS allows multiple labels to be stacked
  - » *Ingress LSR performs* label push *(S=1 in label)*
  - » *Egress LSR performs* label pop
  - » *Intermediate LSRs can perform additional pushes & pops (S=0 in label) to create tunnels*
  - » *Above figure has tunnel between A & G; tunnel between B&F*
  - » *All flows in a tunnel share the same outer MPLS label*

Tunneling in MPLS



- MPLS can control the entire path of a packet without explicitly specifying the intermediate routers by creating tunnels through the intermediary routers that can span multiple segments
- LERs (LER1, LER2, LER3, and LER4) create an LSP between them (LSP 1)
- separate LSP (LSP 2) is created between the two LERs (LER1 and LER2) that spans LSR1, LSR2, and LSR3
- the concept of a label stack is used
  - packet that travels through LSP 1, LSP 2, and LSP 3, carries two complete labels at a time
  - label for LSP 1 and LSP 2, and label for LSP 1 and LSP 3

MPLS APPLICATIONS
2. Virtual Private Networks

- *VPN simulates the operation of a private wide area network (WAN) over the public Internet*
- *an ISP must solve the problems of data privacy and support the use of non-unique, private IP addresses within a VPN*

possible because MPLS makes forwarding decisions based on the value of the label, not the destination address in the packet header.

Fundamental building blocks for VPNs:

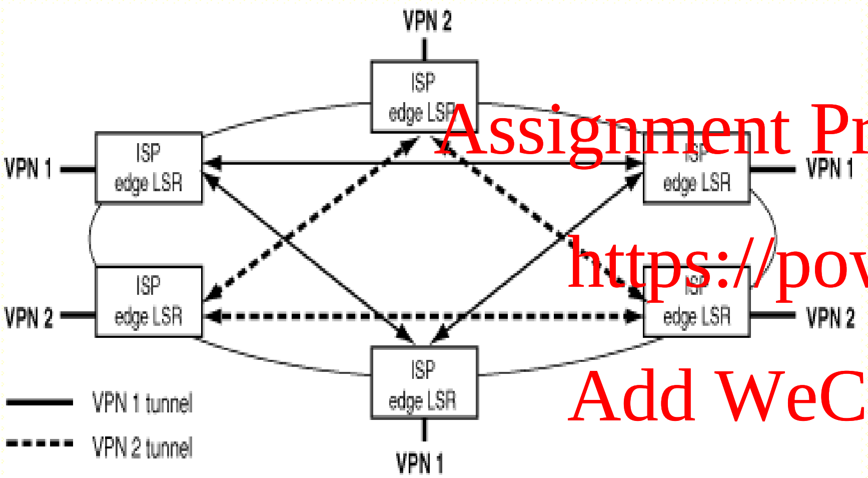
- Firewalls to protect each customer site and provide a secure interface to the Internet

- Authentication to verify that each customer site exchanges data with only validated remote sites

- Encryption to protect data from examination or manipulation as it is transported across the Internet

- Tunneling encapsulation to provide a multiprotocol transport service and enable the use of the private IP address space within a VPN

VPN Deployment

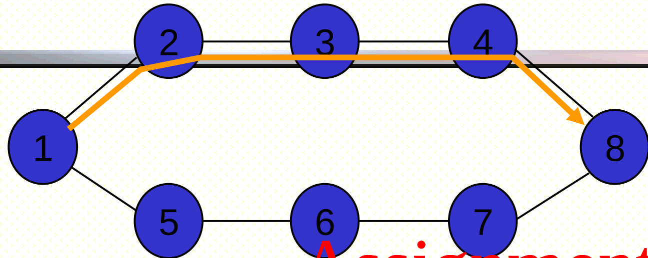


- *ISP can deploy a VPN by provisioning a set of LSPs to provide connectivity among the different sites in the VPN*
- *Each VPN site advertises to the ISP a set of prefixes that are reachable within the local site*
- *VPN Identifiers allow a single routing system to support multiple VPNs whose internal address spaces overlap with each other; for example 23:10.1.1.0 and 109:10.1.1.0*
- *each ingress LSR places traffic into LSPs based on a combination of a packet's destination address and VPN membership information.*

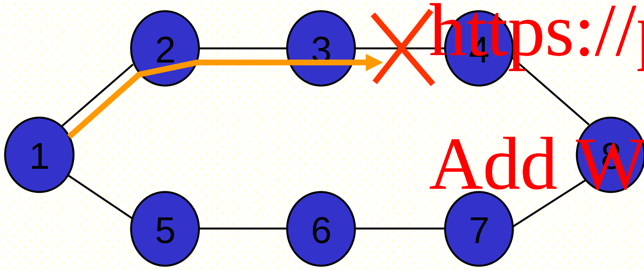
MPLS Survivability

- IP routing recovers from faults in seconds to minutes
- SONET recovers in 50 ms
- MPLS targets in-between path recovery times
- Basic approaches:
  - *Restoration: slower, but less bandwidth overhead*
  - *Protection: faster, but more protection bandwidth*
- Repair methods:
  - *Global repair: node that performs recovery (usually ingress node) may be far from fault, depends on failure notification message*
  - *Local repair: local node performs recovery (usually upstream from fault); does not require failure notification*

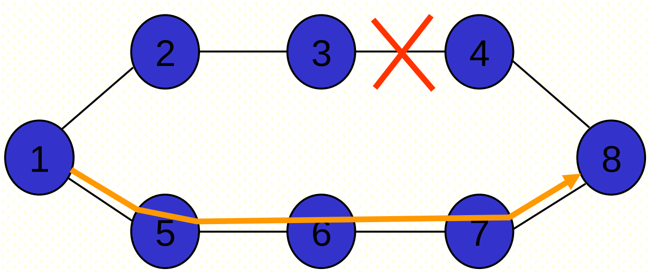
MPLS Restoration



Normal operation



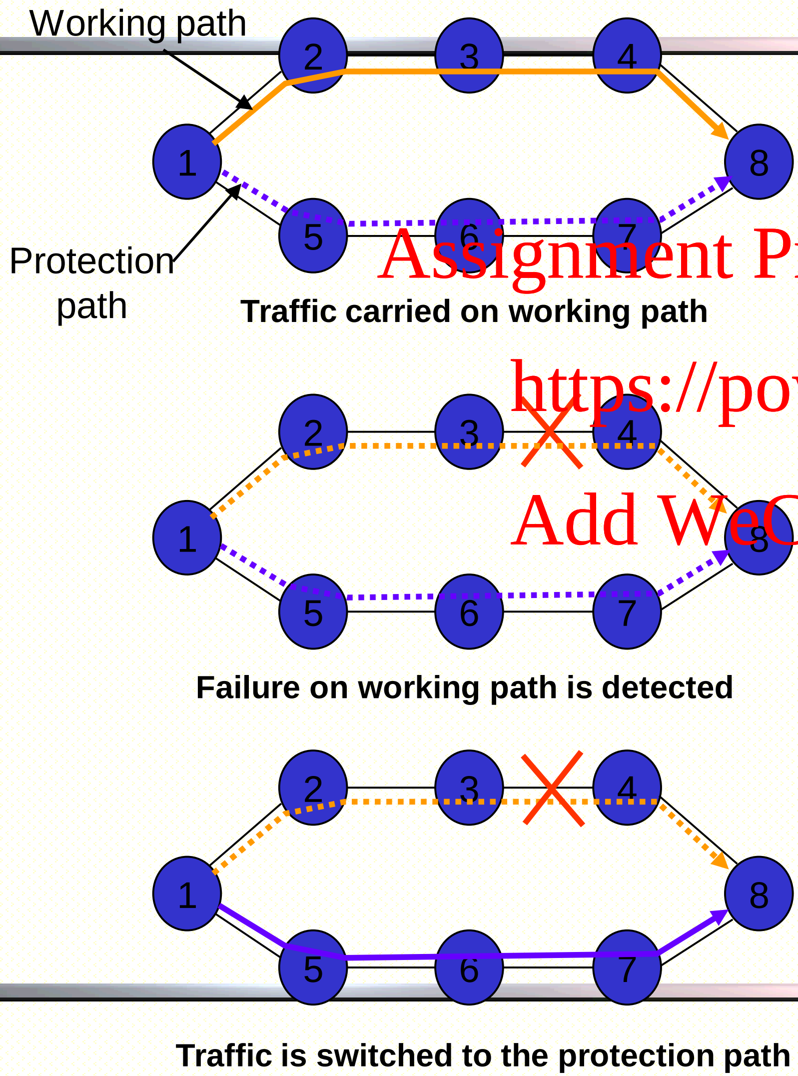
Failure occurs and is detected



Alternate path is established, and traffic is re-routed

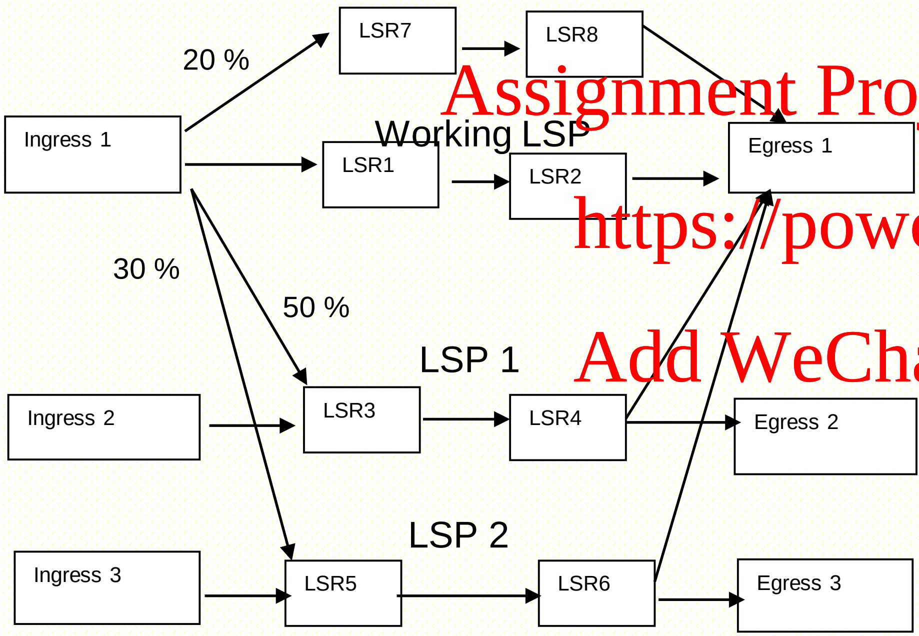
- No protection bandwidth allocated prior to fault
- New paths are established after a failure occurs
- Traffic is rerouted onto the new paths

MPLS Protection



- Protection paths are setup as backups for working paths
  - » *1+1: working path has dedicated protection path*
  - » *1:1: working path shares protection path*
- Protection paths selected so that they are disjoint from working path
- Faster recovery than restoration

MPLS Split-Path Protection



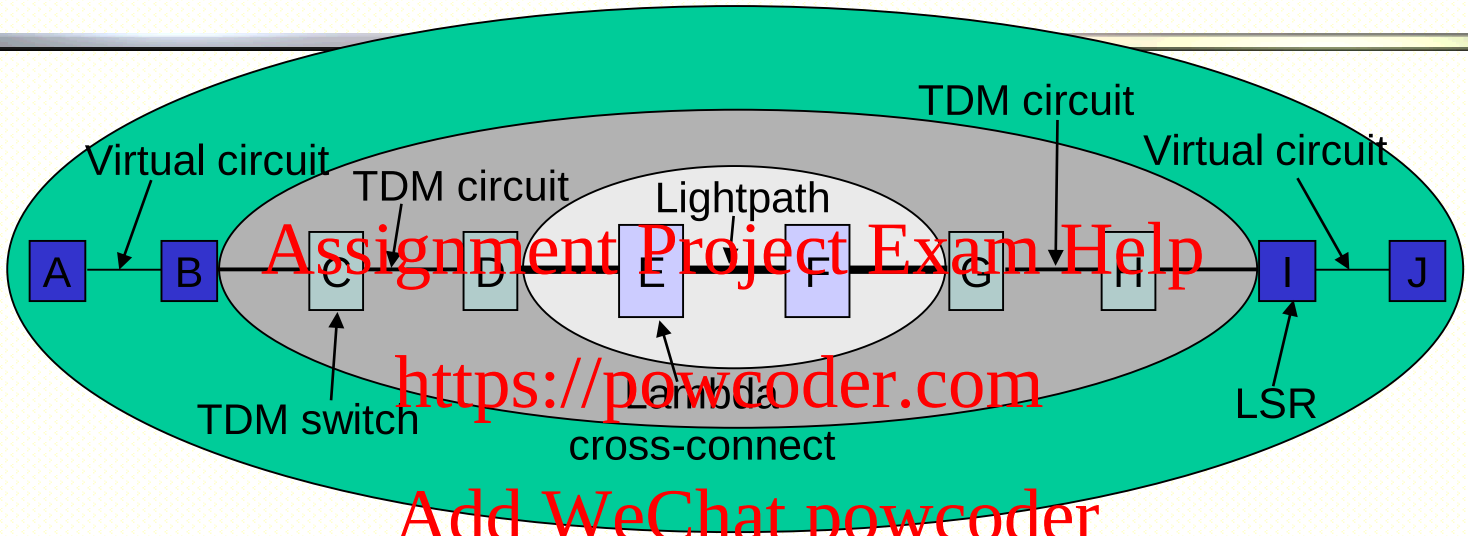
- multiple alternate paths are allowed to carry the traffic of a failed working path

LSP	Bandwidth Guaranteed to working path of other applications	Bandwidth Utilized in Backup path
1	50 %	50%
2	70 %	30 %
3	80 %	20 %

Generalized MPLS

- MPLS:
  - » *Connection-oriented*
  - » *Leverages IP routing protocols, with TE extensions, to provide means for selecting good paths*
  - » *Provides signaling for establishing paths*
- With appropriate extensions, Generalized MPLS can provide the control plane for other networks:
  - » *SONET networks that provide TDM connections*
  - » *WDM networks that provide end-to-end optical wavelength connection*
  - » *Optical networks that provide end-to-end optical fiber path*

Hierarchical LSPs



- GMPLS allows node with multiple switching technologies to be controlled by one control component
- Notion of "label" generalized:
  - » *TDM slot, WDM wavelength, optical fiber port*
- LSP Hierarchy extended to generalized labels"
  - » *MPLS LSP over SONET circuit over wavelength path over fiber*