The Multi-Protocol Label Switching Architecture

Harry Perros

Computer Science Department

NC State University

Raleigh, NC, USA

Email: hp@csc.ncsu.edu

URL: http://www.csc.ncsu.edu/faculty/perros//index.html

TOPICS

- The MPLS architecture
  - Label allocation schemes
  - Explicit routing
  - Label stack
  - Schemes for setting up an LSP
- Label distribution protocols
  - LDP and CR-LDP
  - RSVP-TE

The IP router

- In order to understand the basic concepts behind MPLS, we need to take a look at the structure of an IP router.
- An IP router has a
  - *control* component, and a
  - *forwarding* component.

The control component

- It consists of routing protocols, such as OSPF, BGP, and PIM, which are used to construct routes and exchange routing information among routers.
- This information is used by the routers to construct a forwarding table (routing table), known as the *forwarding information base* (FIB).

The forwarding component

- It consists of procedures for forwarding an IP packet.
- The IP router uses the destination IP address to find an entry in the FIB, using the longest match algorithm. From this, it obtains an interface number, which is the output port connecting the IP router to the next-hop router, to which the IP packet should be sent.

Forwarding equivalent class (FEC)

- A FEC (pronounced *fec)* is the set of all forwarding addresses that have the same prefix.
- Thus, addresses in a router can be grouped into a number of disjoint FECs.
- IP packets belonging to the same FEC have the same output interface.
- In MPLS, an FEC is associated with a *label.*

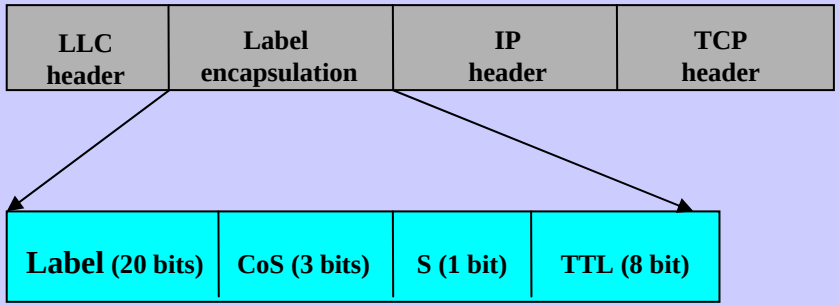
Labels

- A label is a short, fixed-length identifier that has local significance (i.e. it is valid on a single hop interconnecting two routers).
- A label in a packet represents the FEC to which the packet has been assigned.
- The label assigned to a packet is not an encoding of its destination address.

Where is the label carried?

- No space in IPv4 packet for a label.
- If the IP network is running on top of an ATM network/ Frame Relay, the label is carried in the VPI/VCI field/DLCI field.
- For Ethernet, and point-to-point connections running a link layer protocol, such as PPP, the label is encapsulated and inserted between the LLC header and the IP header (*shim header*).

Label encapsulation (Shim header)



CoS (class of service)
S - stack of labels
TTL (time to live)

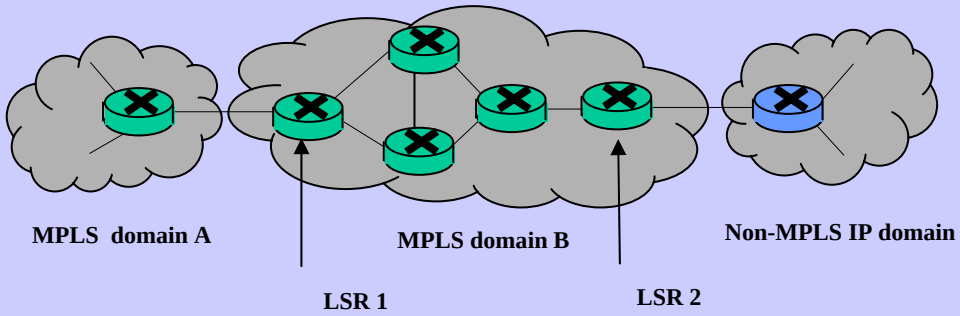
Label stack

Label (20 bits)	CoS (3 bits)	S =0	TTL (8 bits)
Label (20 bits)	CoS (3 bits)	S =0	TTL (8 bits)
⋮			
Label (20 bits)	CoS (3 bits)	S = 1	TTL (8 bits)

Label Switching Routers (LSR)

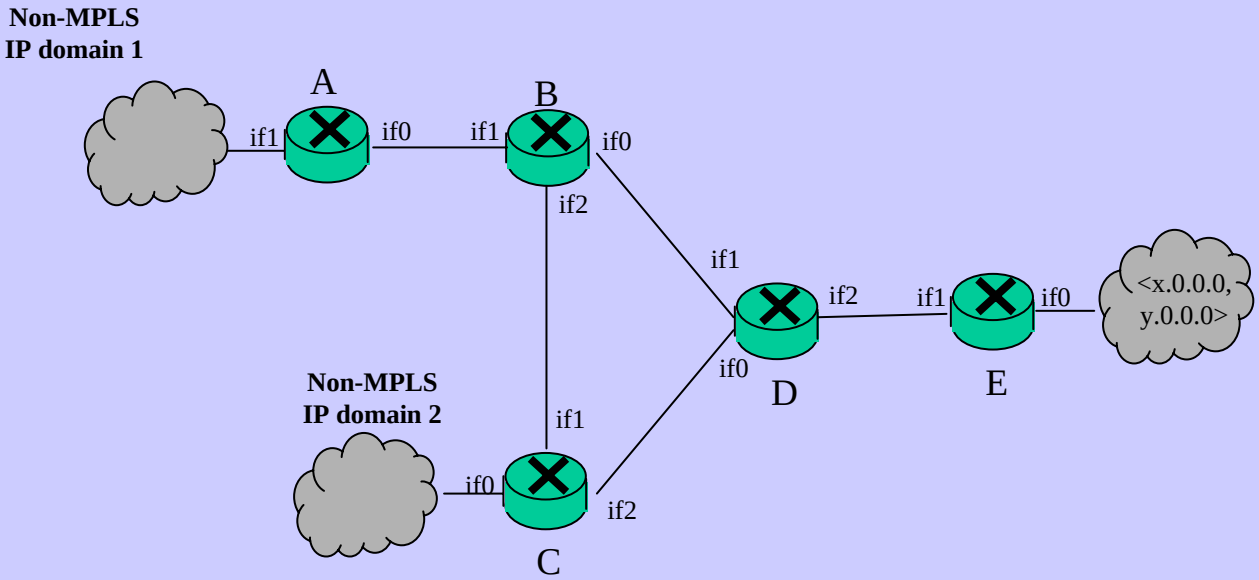
- An *LSR* is an IP router that runs MPLS. It is aware of MPLS control protocols and it operates one or more layer 3 routing protocols
  - It binds labels to FECs,
  - forwards packets based on their labels, and
  - it carries out the customary IP forwarding decision based on prefixes.

An MPLS domain

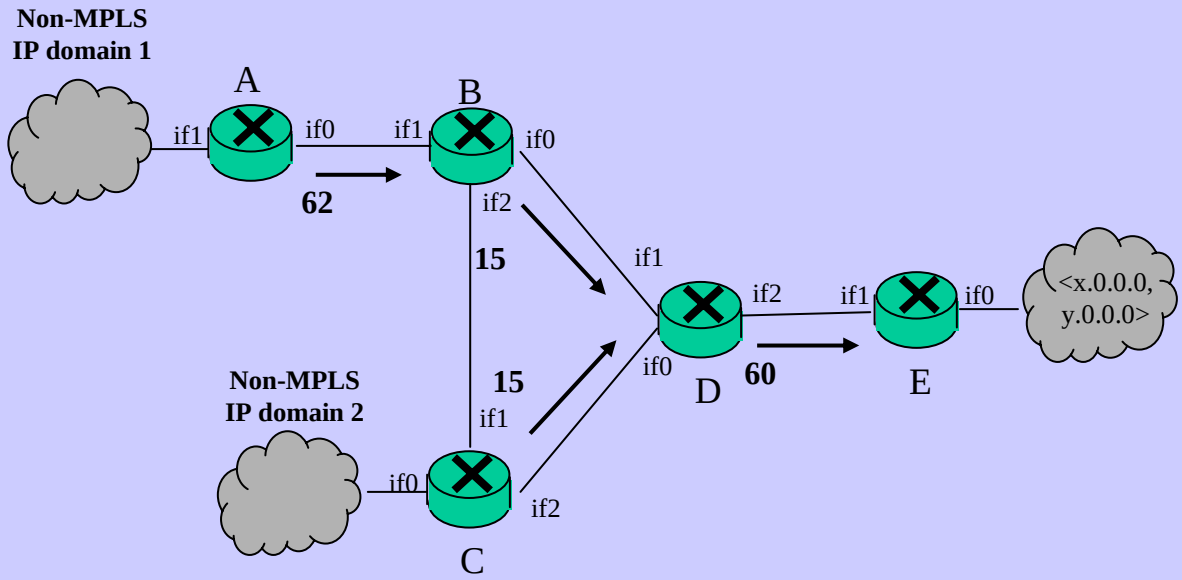


- An MPLS domain is a contiguous set of MPLS nodes which are in the same routing or administrative domain.
- Within an MPLS domain, IP packets are switched using their labels.
- An MPLS domain maybe connected to other MPLS or non-MPLS domains.

An example of label switching



The labels allocated by the LSRs are as follows:

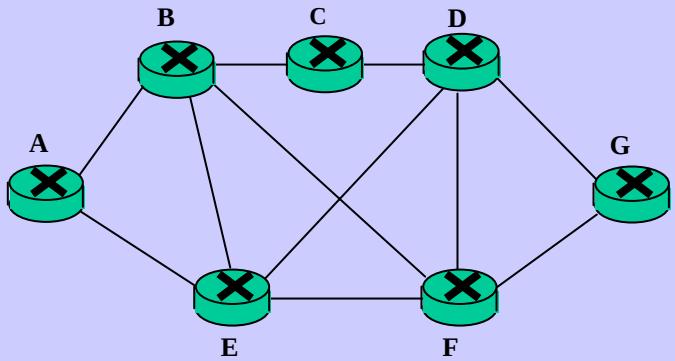


- Now, once the labels have been distributed and the entries have been updated, the forwarding of a packet belonging to this particular FEC is done using solely the labels in the LFIBs.

- Let us assume that A receives a packet with a label 100. A uses this label in its LFIB to locate the new outgoing label and interface. The old label is swapped with the new one and the packet is forwarded to interface 1.

Point-to-point route selection of a label switched path (LSP)

- In general, there are two methods for selecting an LSP for a particular FEC
  - *hop-by-hop routing*
  - *explicit routing.*

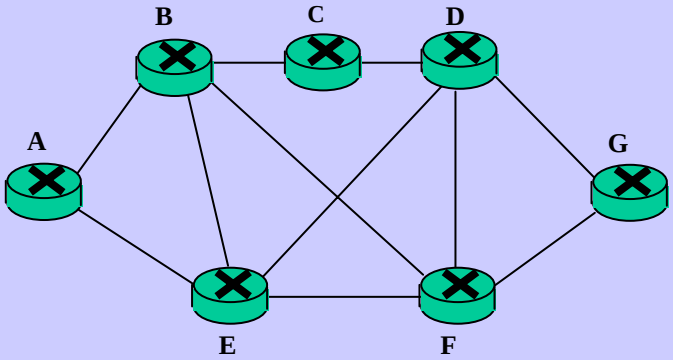


Hop-by-hop routing

- Each node chooses independently the next-hop for a FEC, as in the existing IP networks.
- The information for the next hop is typically provided by a routing protocol such as OSPF, BGP, etc

Explicit routing

- An explicit route is a pre-defined path through the network, and it is known as *constrained-based routed LSP (CR-LSP).*
- The route may be different to those advertised by the routing protocols.
- An LSR determines its next hop for the FEC based on the explicit route.



- In this MPLS network, router A wants to set-up a CR-LDP (i.e., a point-to-point unicast connection) to router G.
- This path can be calculated so that it
  - *minimizes the number of hops (as in IP), or*
  - *minimizes the total end-to-end delay, or*
  - *maximizes throughput, or*
  - *path is pre-calculated to achieve load-balancing, etc.*

An LSP therefore, can be used for a variety of reasons, such as:

- *Evenly distribute traffic among links by moving some of the traffic from highly utilized links to less utilized links (load balancing),*
- *create tunnels for MPLS-based VPNs,*
- *introduce routes based on a QoS criterion, such as minimum number of hops, minimmum total end-to-end delay, and maximum throughput.*

Label Distribution Protocols

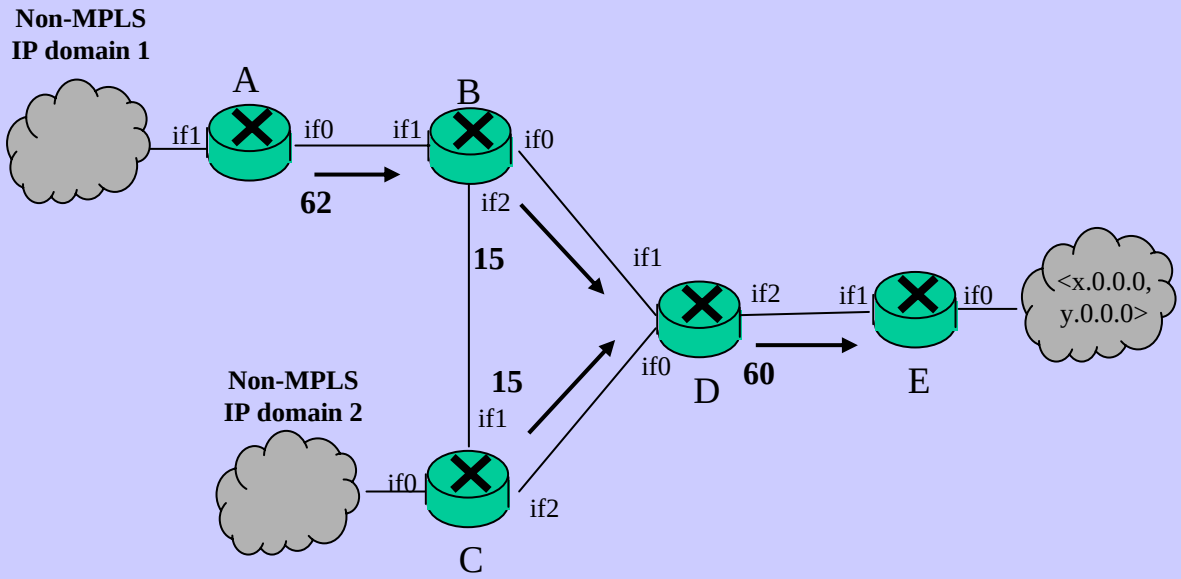
Popular protocols:

- Label Distribution Protocol (LDP)
- Constrained Routing-LDP (CR-LDP)
- RSVP Traffic Engineering (RSVP-TE)

- MPLS requires a set of procedures for the reliable distribution of label bindings between LSRs.
- MPLS does not require a single label distribution protocol.
- LDP/CR-LDP and RSVP-TE are the most popular label distribution protocols

LDP

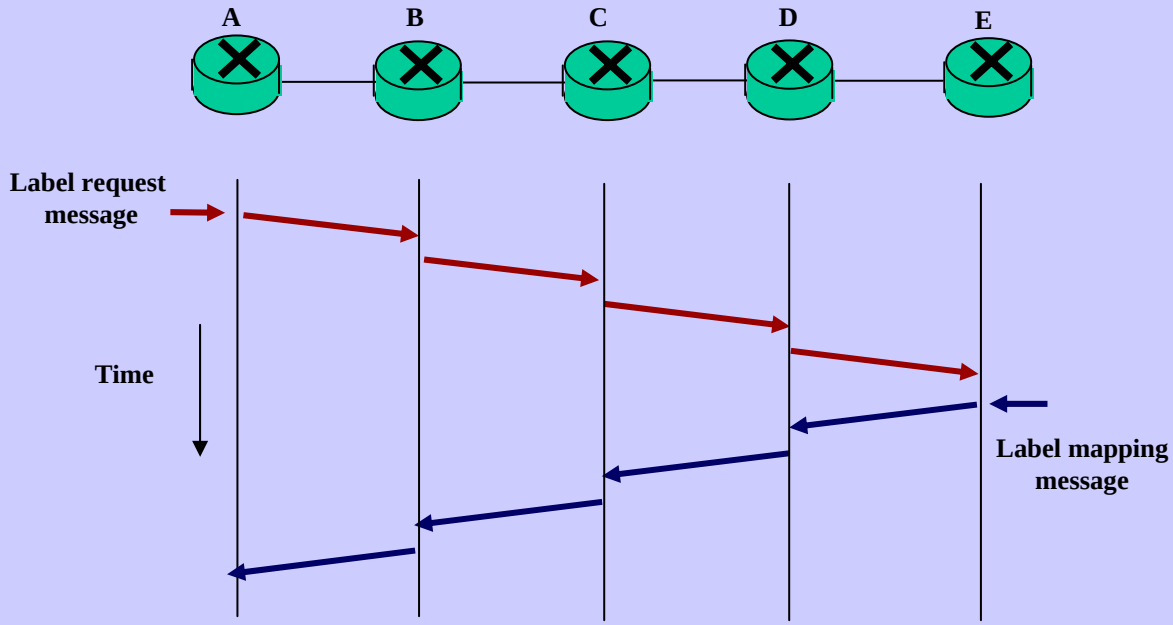
- LDP is a *signalling protocol* used to setup LSPs as in the previous example:



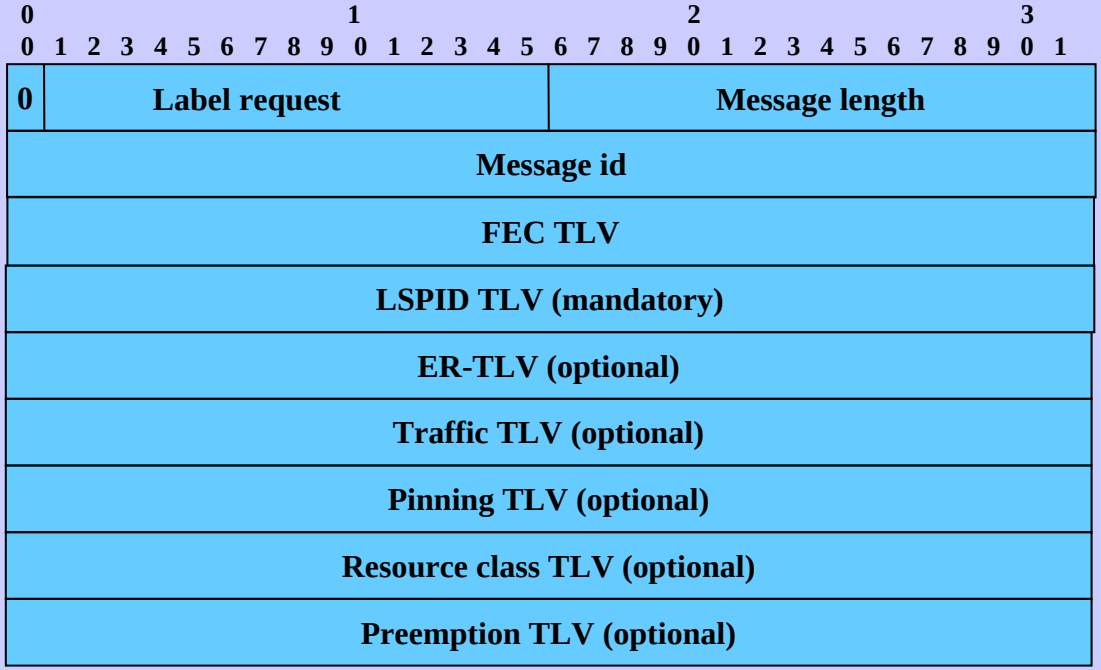
CR-LDP

- CR-LDP is a *signalling protocol* based on LDP.
- It is used to set-up a *unidirectional* point-to-point LSP, referred to as *constrained-routed label-switched path* (CR-LSP).
- A bidirectional LSP is setup by creating two separate LSPs, one in each direction.

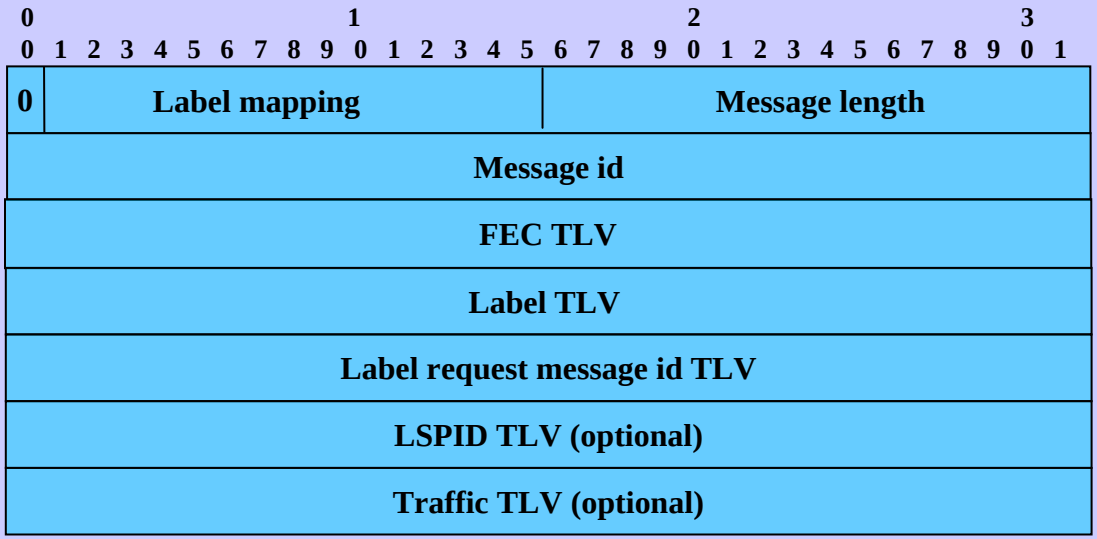
An example of how an CR-LSP is setup



The label request message



The label mapping message



RSVP

- RSVP was designed to support the *integrated services (intserv)* architecture.
- The intserv architecture was developed by IETF in the mid 1990s with a view to introducing QoS in the IP network.
- Intserv was never widely accepted. It has been superceded by the *DiffServ* architecture, which has been successfully deployed in the IP network.

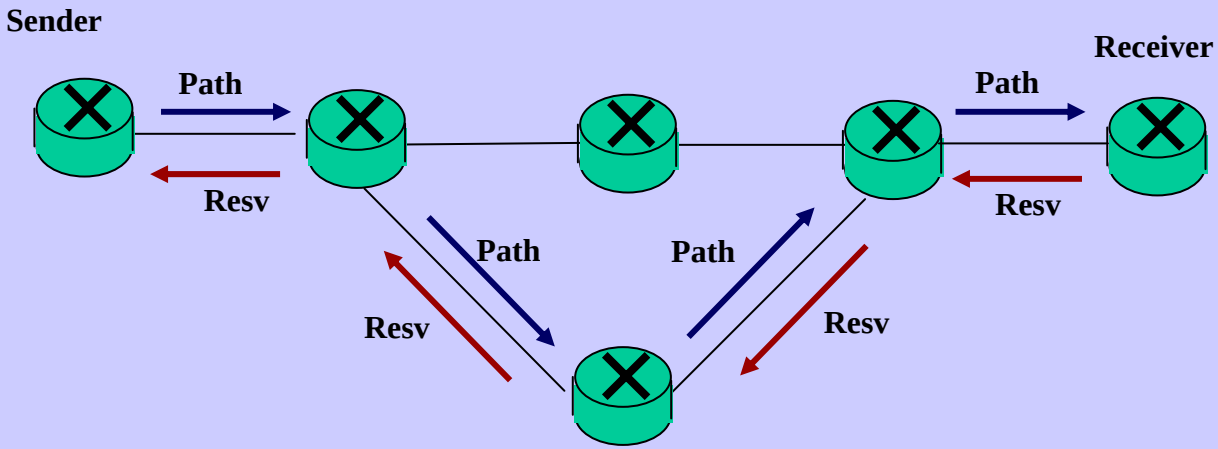
- RSVP is a signaling protocol used for the reliable establishment and maintenance of resource reservations for both unicast and many-to-many multicast applications
- RSVP can be used to carry other types of control information since it is not aware of the content of the RSVP protocol fields.
- In view of this, it was proposed to be used in MPLS.

Path and Resv messages

RSVP makes use of two messages:

- *Path*: This message originates from the sender and travels to the receiver.
- *Resv*: Upon receipt of the Path message, the receiver responds with a *Resv* message, which travels on the reverse path of the Path message, and reserves bandwidth on each router along the path.

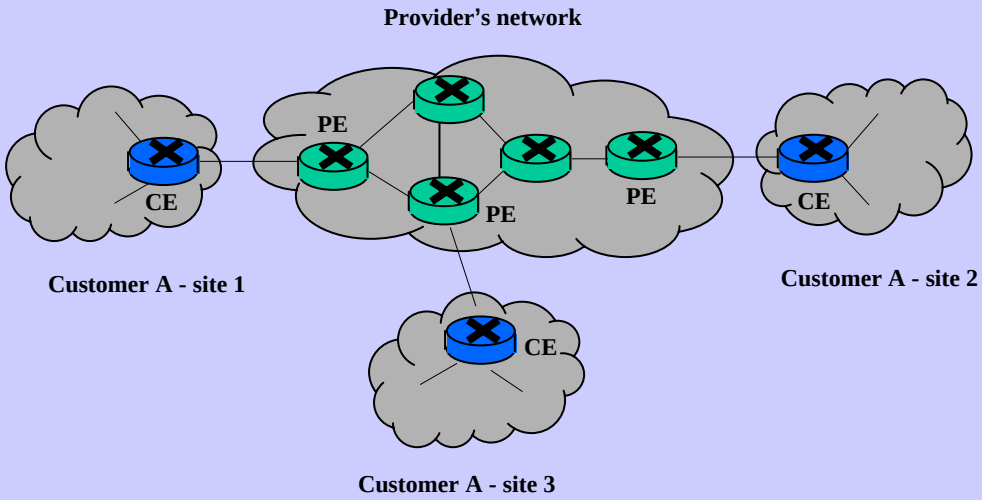
An example



RSVP - TE

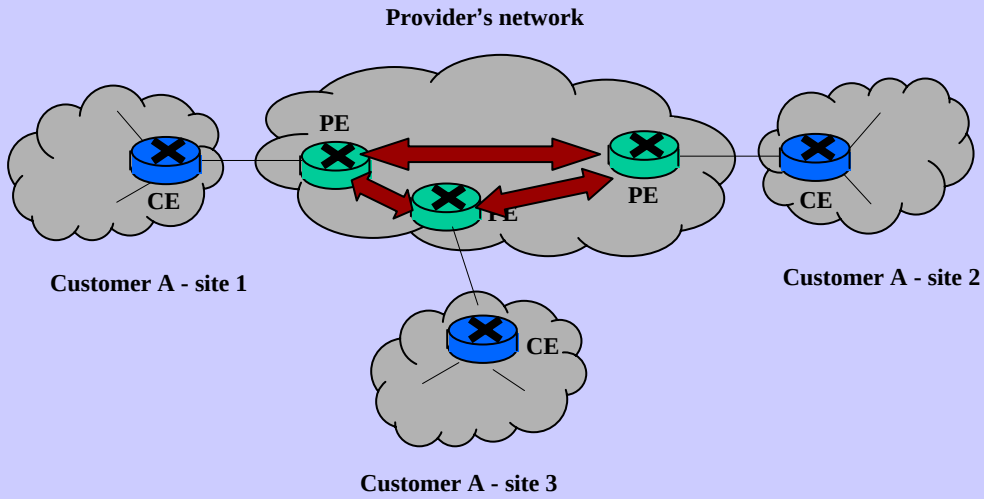
- RSVP-TE uses downstream-on-demand label allocation with ordered control to setup an LSP.
- This is implemented using the Path and Resv messages of RSVP which have been augmented with new objects.

Virtual Private Networks



Physical topology

Logical topology

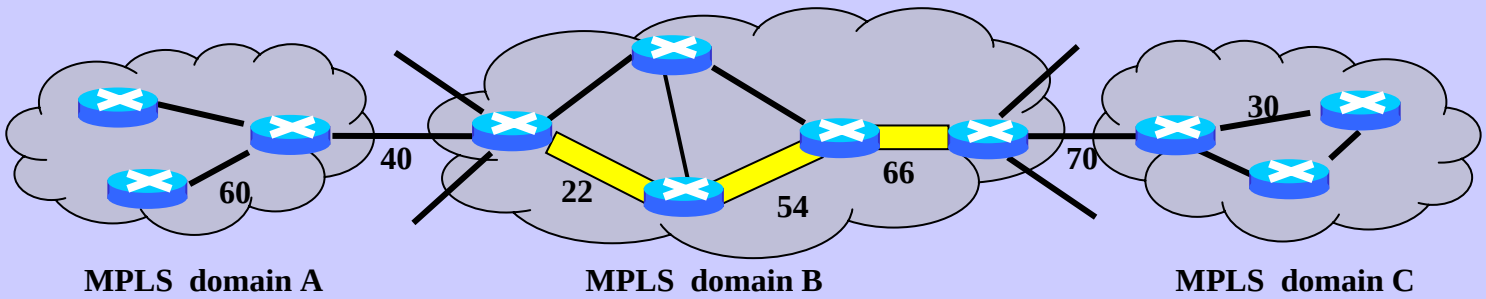


The label stack

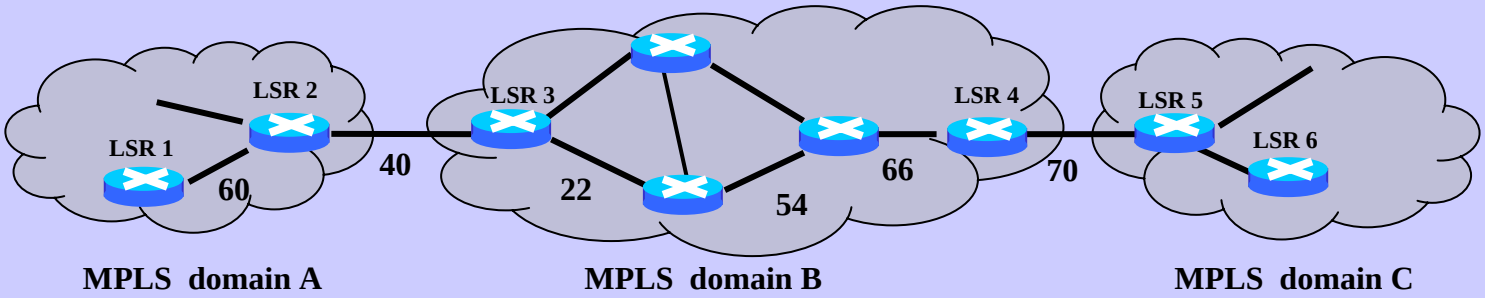
- MPLS allows a packet to carry a set of labels organized as a *stack*.
- When the packet is forwarded within a domain, it contains two labels. The label at the top of the stack is used for label switching within the interior LSRs. The label in the next level is used by the egress LSP LSR to forward the packet to the next ingress LSP LSR.

An example

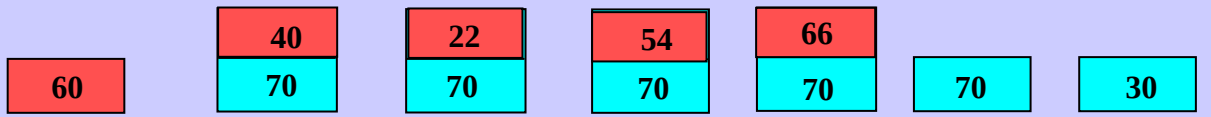
The labels in MPLS B domain form a tunnel. At the end of the tunnel, the LSR may not know where to forward the packet. This can be easily resolved using a label stack.



The label stack



Label stack carried in the shim header of an IP packet



Level 1 — Level 2 — Level 1

Operation in the NHLFE

- **LSR2:** Replace label, Push new label
- **LSR 3 - LSR 4:** Replace label
- **LSR 4:** Pop label stack
- **LSR 5:** Replace label