# *Performance Measurements of MPLS Traffic Engineering and QoS*

By

Tamrat Bayle

Reiji Aibara

Kouji Nishimura

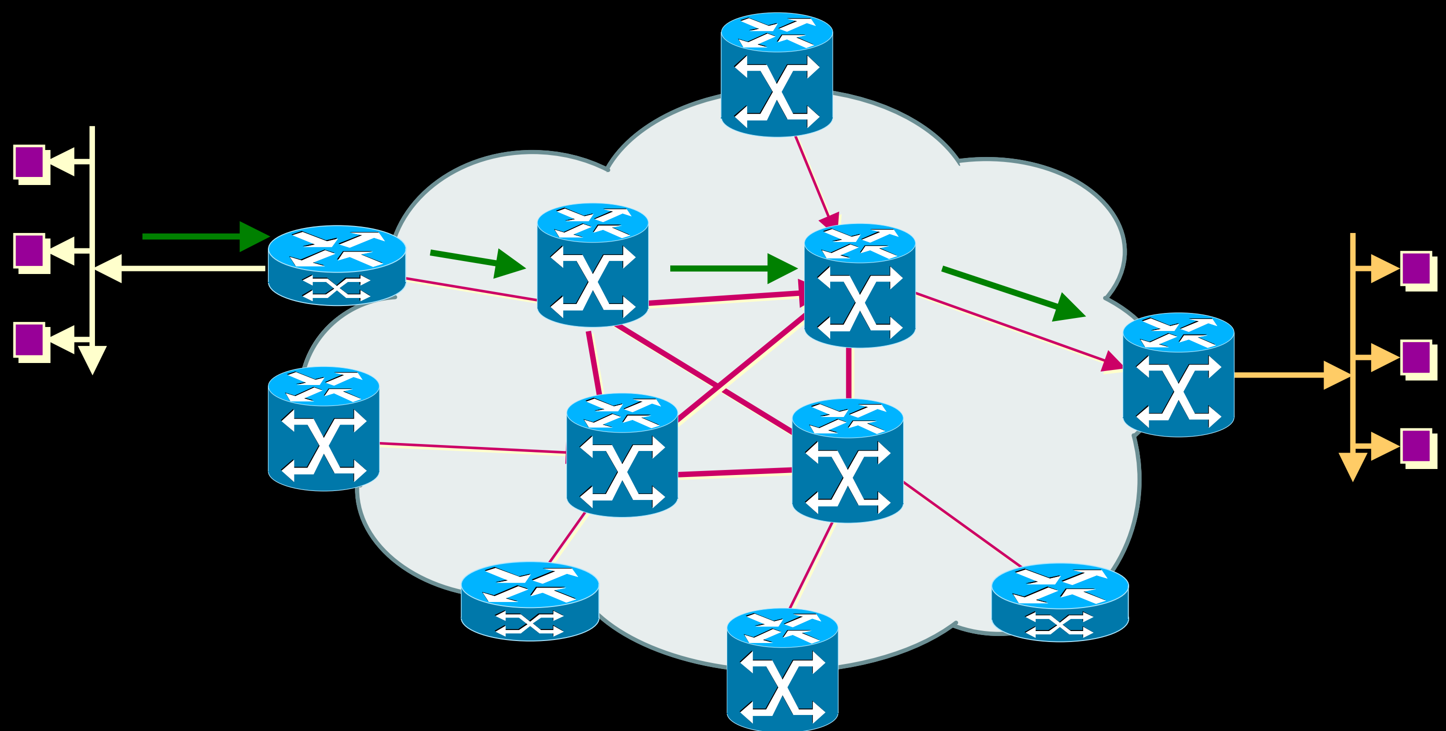
# *Multiprotocol Label Switching*

- Traditional IP Routing
- Disadvantages
- Need for MPLS
- MPLS basics and terminologies
- Experiments

# *Traditional IP Routing*

- Choosing the next hop
  - Open Shortest Path First (OSPF) to populate the routing table
  - Route look up based on the IP address
  - Find the next router to which the packet has to be sent
  - Replace the layer 2 address
- Each router performs these steps

# *Traditional IP Routing (contd)*

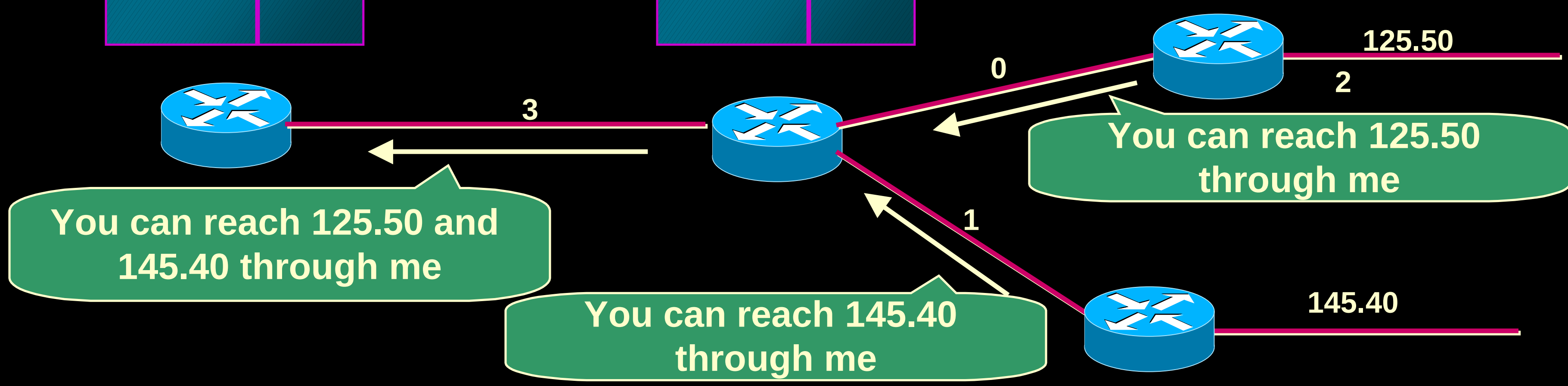


# *Distributing Routing Information*

| Address Prefix | Path |
|---|---|
| 125.50 | 3 |
| 145.40 | 3 |

| Address Prefix | Path |
|---|---|
| 125.50 | 0 |
| 145.40 | 1 |

| Address Prefix | Path |
|---|---|
| 125.50 | 2 |

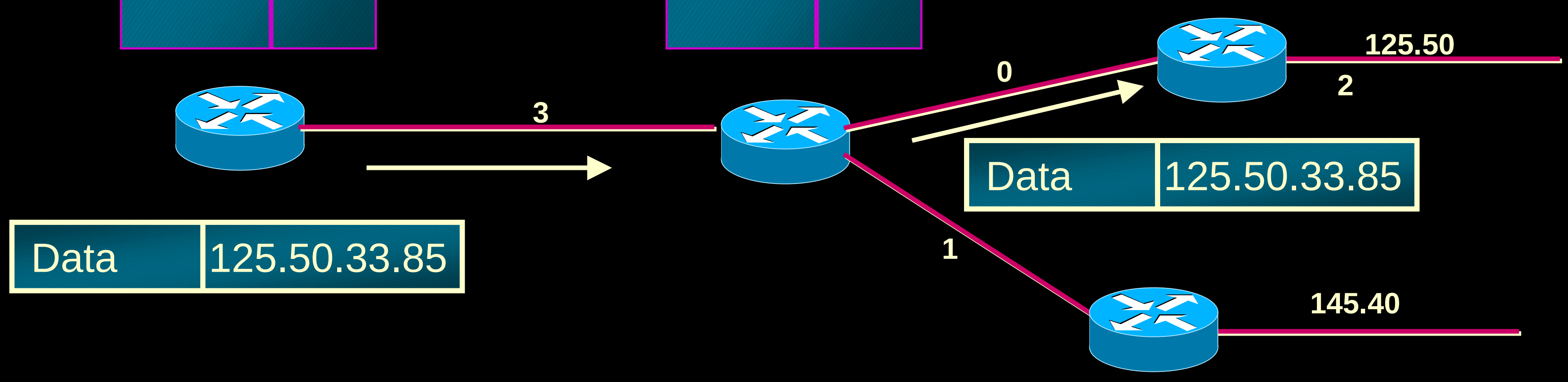


# *Distributing Routing Information(contd)*

| Address Prefix | Path |
|---|---|
| 125.50 | 3 |
| 145.40 | 3 |

| Address Prefix | Path |
|---|---|
| 125.50 | 0 |
| 145.40 | 1 |

| Address Prefix | Path |
|---|---|
| 125.50 | 2 |



# *Disadvantages*

- Header analysis performed at each hop
- Increased demand on routers
- Utilizes the best available path
- Some congested links and some underutilized links!
  - Degradation of throughput
  - Long delays
  - More losses
- No QoS
  - No service differentiation
  - Not possible with connectionless protocols

# *Need for MPLS*

- Rapid growth of Internet
- New *latency dependent* applications
- Quality of Service (QoS)
  - Less time at the routers
- Traffic Engineering
  - Flexibility in routing packets
- Connection-oriented forwarding techniques with connectionless IP
  - Utilizes the IP header information to maintain interoperability with IP based networks
  - Decides on the path of a packet before sending it

# *What is MPLS?*

- Multi Protocol – supports protocols even other than IP
  - Supports IPv4, IPv6, IPX, AppleTalk at the network layer
  - Supports Ethernet, Token Ring, FDDI, ATM, Frame Relay, PPP at the link layer
- Label – short fixed length identifier to determine a route
  - Labels are added to the top of the IP packet
  - Labels are assigned when the packet enters the MPLS domain
- Switching – forwarding a packet
  - Packets are forwarded based on the label value
  - NOT on the basis of IP header information

# *MPLS Background*

- Integration of layer 2 and layer 3
  - Simplified connection-oriented forwarding of layer 2
  - Flexibility and scalability of layer 3 routing
- MPLS does not replace IP; it supplements IP
- Traffic can be marked, classified and explicitly routed
- QoS can be achieved through MPLS

# *IP/MPLS comparison*

- Routing decisions
  - IP routing – based on destination IP address
  - Label switching – based on labels
- Entire IP header analysis
  - IP routing – performed at each hop of the packets path in the network
  - Label switching – performed only at the ingress router
- Support for unicast and multicast data
  - IP routing – requires special multicast routing and forwarding algorithms
  - Label switching – requires only one forwarding algorithm

# *Key Acronyms*

- MPLS – MultiProtocol Label Switching
- FEC – Forward Equivalence Class
- LER – Label Edge Router
- LSR – Label Switching Router
- LIB – Label Information Base
- LSP – Label Switched Path
- LDP – Label Distribution Protocol

# *Forwarding Equivalence Class (FEC)*

- A group of packets that require the same forwarding treatment across the same path
- Packets are grouped based on any of the following
  - Address prefix
  - Host address
  - Quality of Service (QoS)
- FEC is encoded as the label

# *FEC example*

Assume packets have the destination address as

- 124.48.45.20
- 143.67.25.77
- 143.67.84.22
- 124.48.66.90

| <u>FEC –1</u> <u>label x</u> | <u>FEC – 2</u> <u>label y</u> |
|---|---|
| 143.67.25.77 | 124.48.45.20 |
| 143.67.84.22 | 124.48.66.90 |

# *FEC example (contd)*

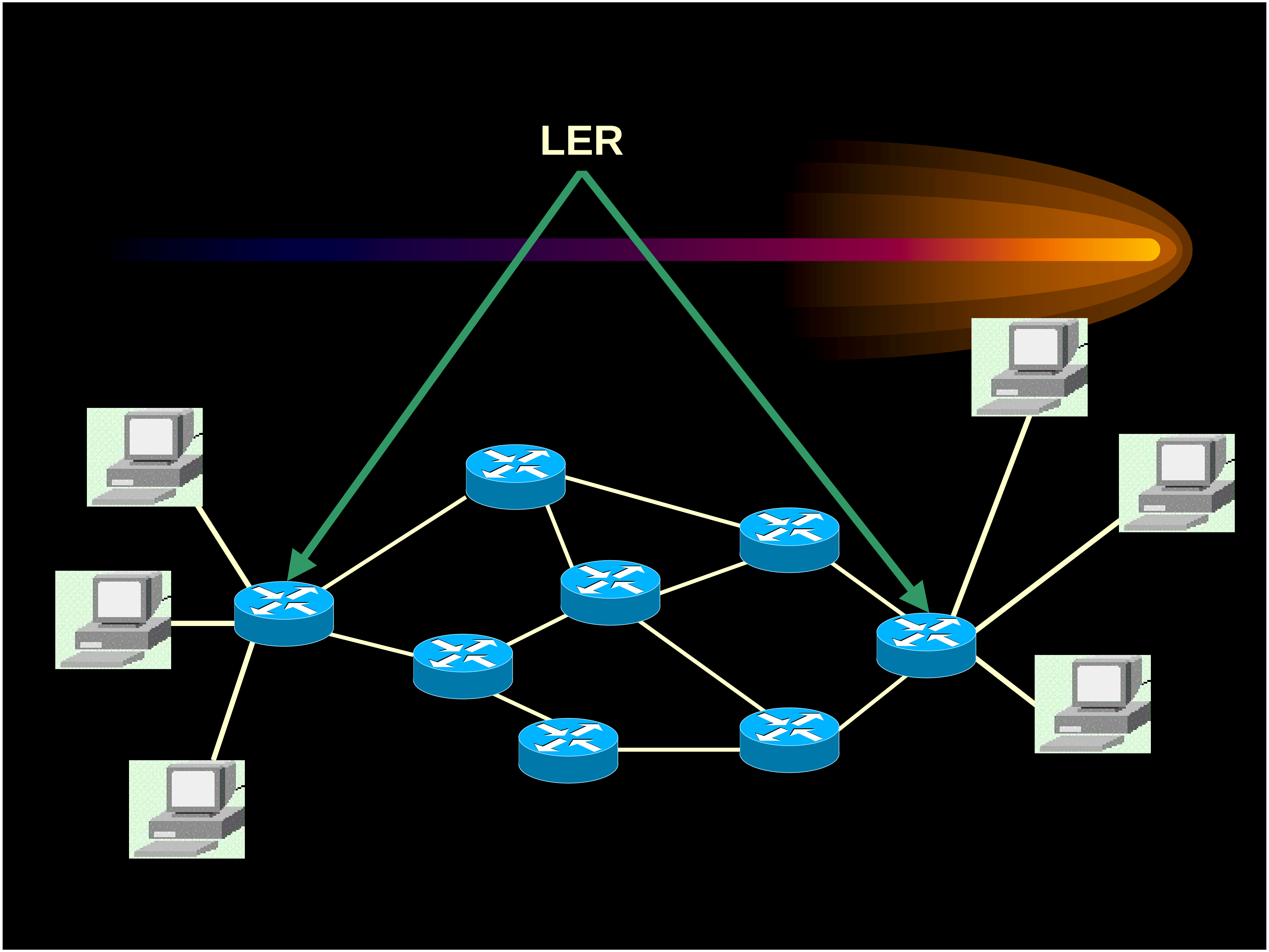
- Assume packets have the destination address and QoS requirements as

| | |
|---|---|
| • 124.48.45.20 | qos = 1 |
| • 143.67.25.77 | qos = 1 |
| • 143.67.84.22 | qos = 3 |
| • 124.48.66.90 | qos = 4 |
| • 143.67.12.01 | qos = 3 |

| FEC –1 label a | FEC – 2 label b | FEC – 3 label c | FEC – 4 label d |
|---|---|---|---|
| 143.67.25.77 | 124.48.45.20 | 143.67.84.22 | 124.48.66.90 |
| | | 143.67.12.01 | |

# *Label Edge Router (LER)*

- Can be an ATM switch or a router
- Ingress LER performs the following:
  - Receives the packet
  - Adds label
  - Forwards the packet into the MPLS domain
- Egress LER removes the label and delivers the packet

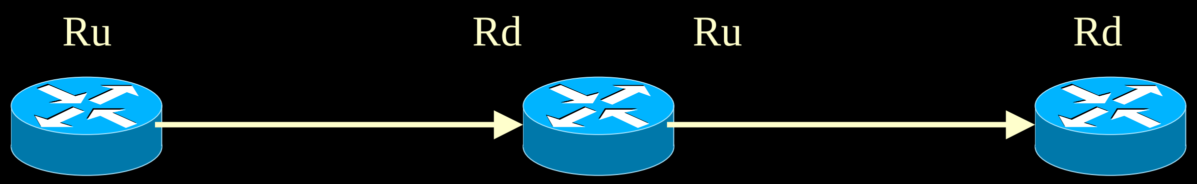

LER

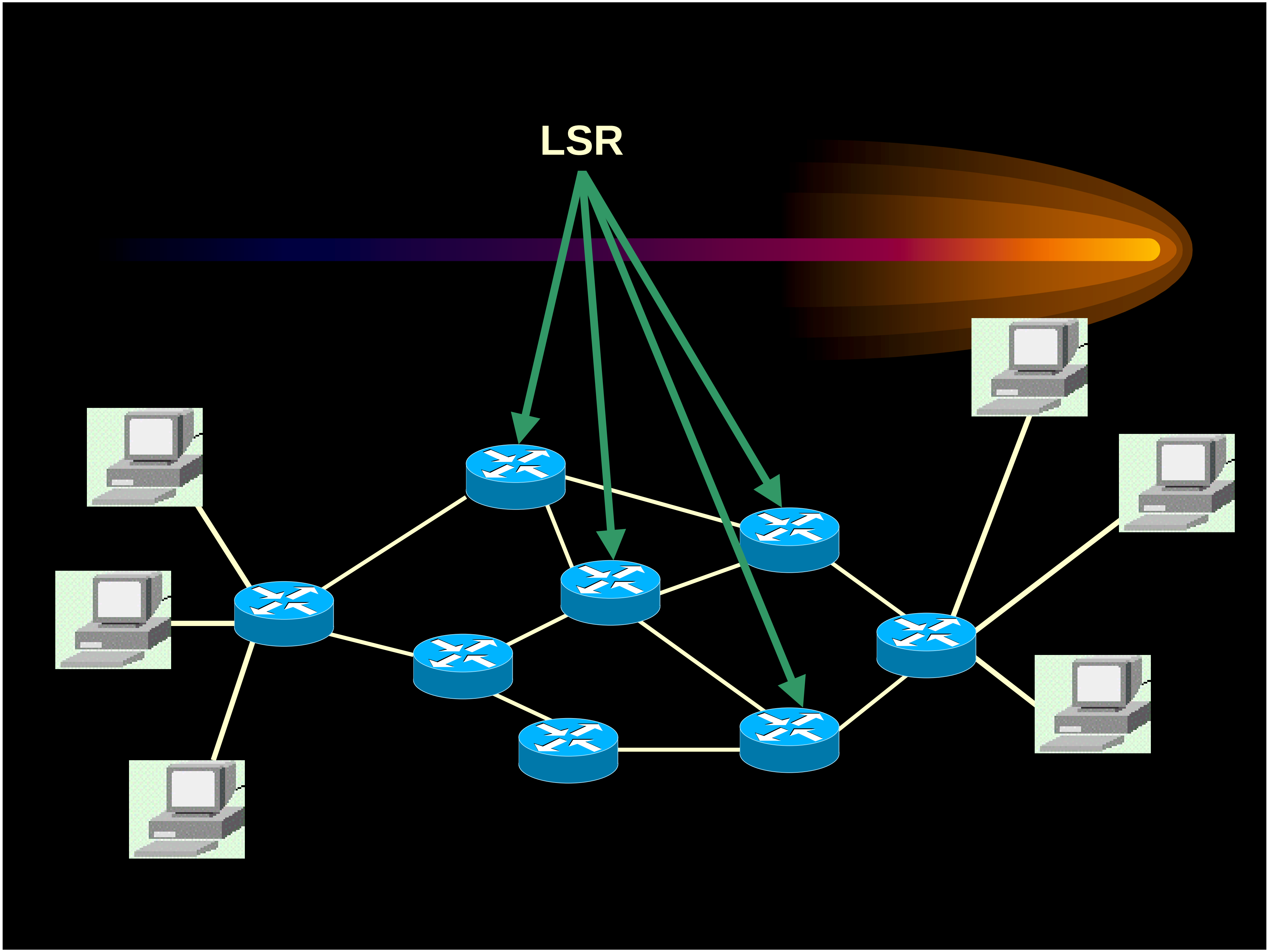
# *Label Switching Router (LSR)*

- A router/switch that supports MPLS
- Can be a router
- Can be an ATM switch + label switch controller
- Label swapping
  - Each LSR examines the label on top of the stack
  - Uses the Label Information Base (LIB) to decide the outgoing path and the outgoing label
  - Removes the old label and attaches the new label
  - Forwards the packet on the predetermined path

# *Label Switching Router (contd)*

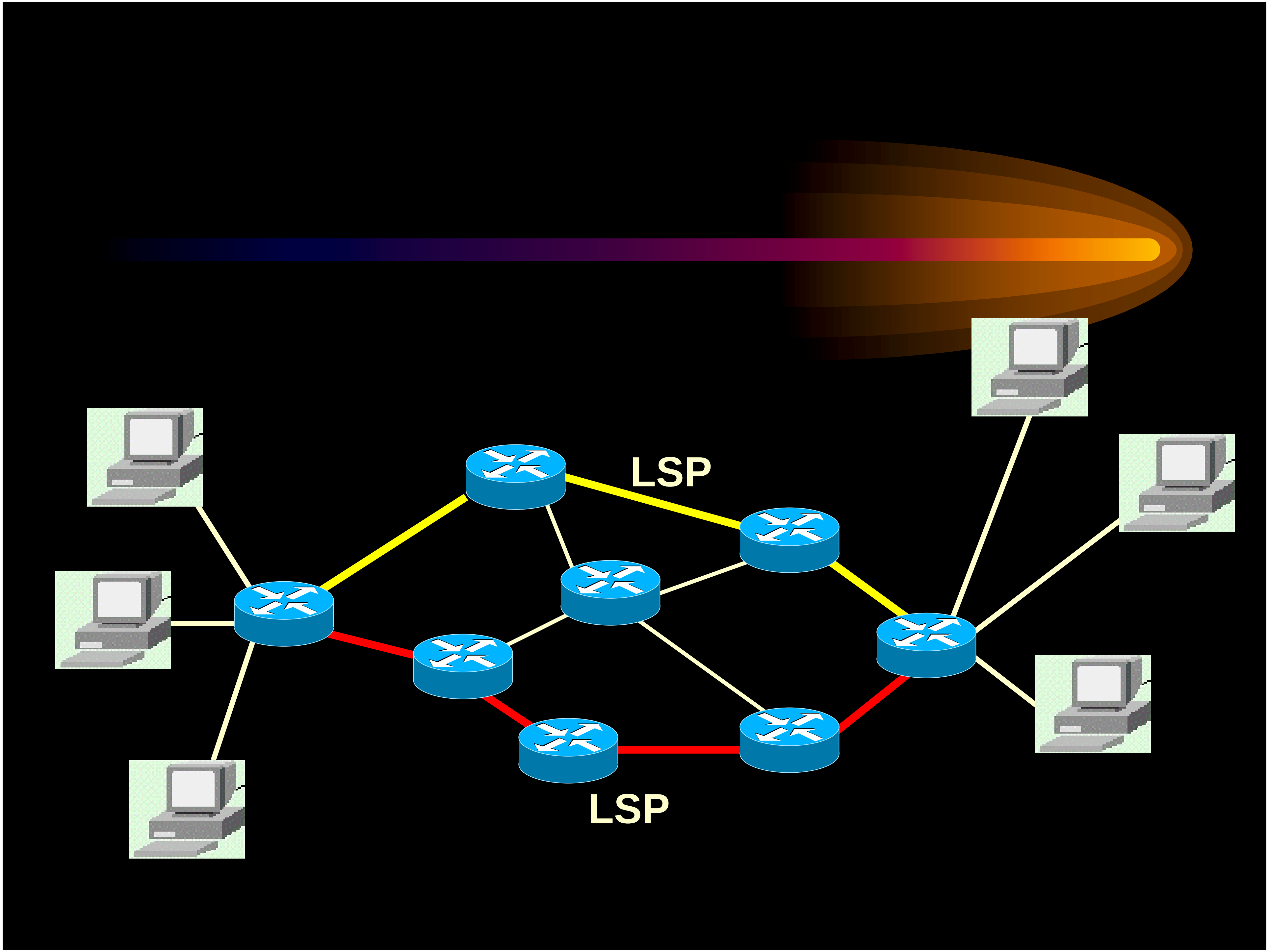
- Upstream Router (Ru) – router that sends packets
- Downstream Router(Rd) – router that receives packets
  - Need not be an end router
  - Rd for one link can be the Ru for the other




LSR

# *Label Switched Path(LSP)*

- LSP defines the path through LSRs from ingress to egress router
- FEC is determined at the LER-ingress
- LSPs are unidirectional
- LSP might deviate from the IGP shortest path


LSP
LSP

# *Label*

- A short, fixed length identifier (32 bits)
- Sent with each packet
- Local between two routers
- Can have different labels if entering from different routers
- One label for one FEC
- Decided by the downstream router
  - LSR binds a label to an FEC
  - It then informs the upstream LSR of the binding

# *Label (contd)*

- ATM
  - VCI/VPI field of ATM header
- Frame Relay
  - DLCI field of FR header
- PPP/LAN
  - ‘shim’ header inserted between layer 2 and layer 3

# *Label (contd)*

PPP Header

| | | |
|---|---|---|
| **Layer 3 Header** | **Label** | **PPP Header** |

LAN MAC Header

| | | |
|---|---|---|
| **Layer 3 Header** | **Label** | **MAC Header** |

ATM Cell Header

| | | | | | | |
|---|---|---|---|---|---|---|
| **DATA** | **HEC** | **CLP** | **PTI** | **VCI** | **VPI** | **GFC** |

**Label**

# *Shim Header*

| TTL | S | EXP | Label |
|---|---|---|---|
| 31 | 23 | 22 | 19 … 0 |

Label = 20 bits

EXP = Experimental bits, 3 bits

S = Bottom of stack, 1 bit

TTL = Time To Live, 8 bits

# *Shim Header (contd)*

- EXP field
  - Also known as Class of Service (CoS) bits
  - Used for experimentation to indicate packet's treatment
  - Queuing as well as scheduling
  - Different packets can receive different treatment depending on the CoS value
- S bit
  - Supports hierarchical label stack
  - 1 – if the label is the bottom most label in the label stack
  - 0 – for all other labels

# *Time To Live (TTL)*

- TTL value decremented by 1 when it passes through an LSR
- If TTL value = 0 before the destination, discard the packet
- Avoids loops may exist because of some misconfigurations
- Multicast scoping – limit the scope of a packet
- Supporting the *traceroute* command

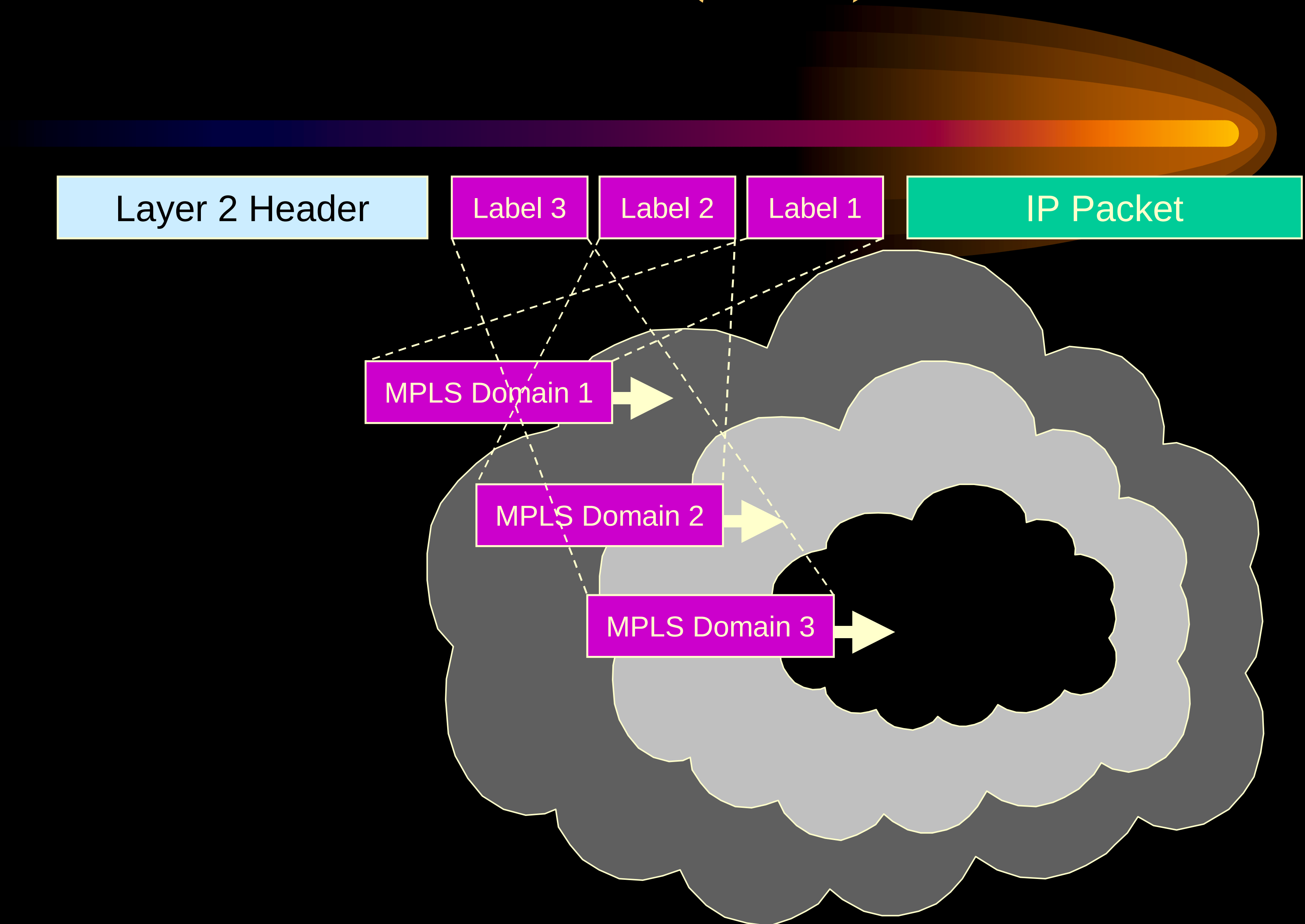
# *TTL (contd)*

- Shim header
  - Has an explicit TTL field
  - Initially loaded from the IP header TTL field
  - At the egress LER, value of TTL is copied into the TTL field of the IP header
- Data link layer header (e.g VPI/VCI)
  - No explicit TTL field
  - Ingress LER estimates the LSP length
  - Decrements the TTL count by the LSP length
  - If initial count of TTL less than the LSP length, discard the packet

# *Label stack*

- MPLS supports hierarchy
- A packet can carry a number of labels
- Each LSR processes the topmost label
  - Irrespective of the level of hierarchy
- If traffic crosses several networks, it can be tunneled across them
- Use stacked labels
- Advantage – reduces the LIB table of each router drastically

# *Label stack (contd)*



# *Labels – scope and uniqueness*

- Labels are local between two LSRs
- Rd might give label L1 for FEC F and distribute it to Ru1
- At the same time, it might give a label L2 to FEC F and distribute it to Ru2
- L1 might not necessarily be equal to L2
- Can there be a same label for different FECs?
  - Generally, NO
  - BUT no such specification
  - LSR must have different label spaces to accommodate both
  - SHIM header specifies that different label spaces used for unicast packets and multicast packets

# *Invalid labels*

- What should be done if an LSR receives an invalid label?
- Should it be forwarded as an unlabeled IP packet?
- Should it be discarded?

- MUST be discarded!
- Forwarding it can cause a loop
- Same treatment if there is no valid outgoing label

# *Route selection*

- Refers to the method of selecting an LSP for a particular FEC
- Done by LDP
  - Set of procedures and messages
  - Messages exchanged between LSRs to establish an LSP
  - LSRs associate an FEC with each LSP created
- Two types of LDP
  - Hop by hop routing
  - Explicit routing

# *Route selection (contd)*

- Hop by Hop
  - Allows each LSR to individually choose the next hop
  - This is the usual mode today in existing IP networks
  - No overhead processing as compared to IP
- Explicit routing
  - A single router, generally the ingress LER,specifies several or all of the LSRs in the LSP
  - Provides functionality for traffic engineering and QoS
    - Several: loosely explicitly routed
    - All: strictly explicitly routed
  - E.g. CR-LDP, TE-RSVP

# *Label Information Base (LIB)*

- Table maintained by the LSRs
- Contents of the table
  - Incoming label
  - Outgoing label
  - Outgoing path
  - Address prefix

# Label Information Base (LIB)

| Incoming label | Address Prefix | Outgoing Path | Outgoing label |
|---|---|---|---|
| | | | |

# *MPLS forwarding*

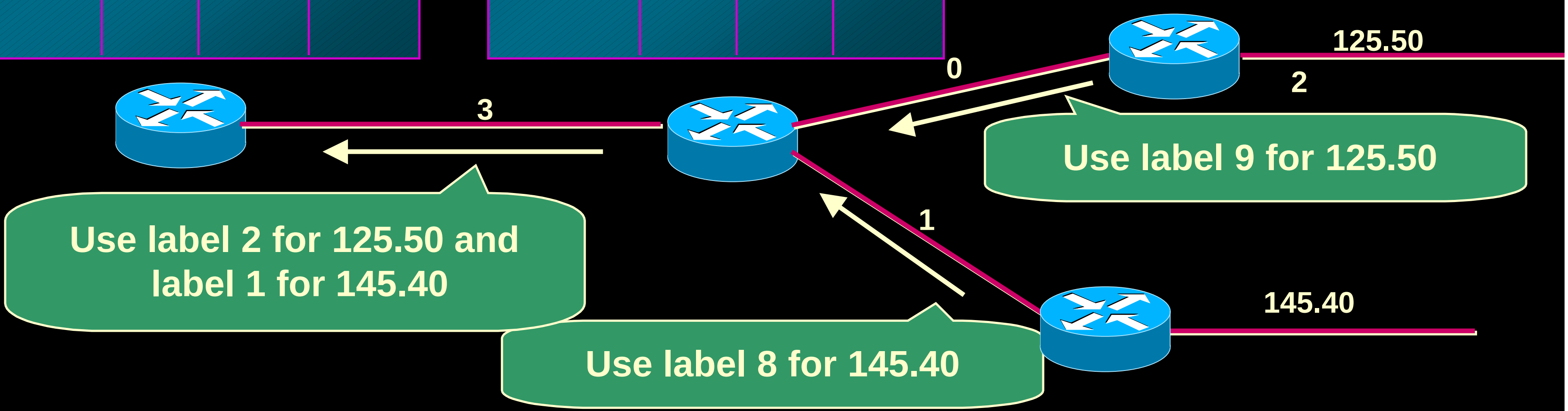
- Existing routing protocols establish routes
- LDP establishes label to route mappings
- LDP creates LIB entries for each LSR
- Ingress LER receives packet,adds a label
- LSRs forward labeled packets using label swapping
- Egress LER removes the label and delivers the packet

# *MPLS forwarding (contd)*

| Address Prefix | Out Path | In Label | Out Label |
|---|---|---|---|
| 125.50 | 3 | | 2 |
| 145.40 | 3 | | 1 |
| | | | |

| Address Prefix | Out Path | In Label | Out Label |
|---|---|---|---|
| 125.50 | 0 | 2 | 9 |
| 145.40 | 1 | 1 | 8 |
| | | | |

| Address Prefix | Out Path | In Label | Out Label |
|---|---|---|---|
| 125.50 | 2 | 9 | |
| | | | |

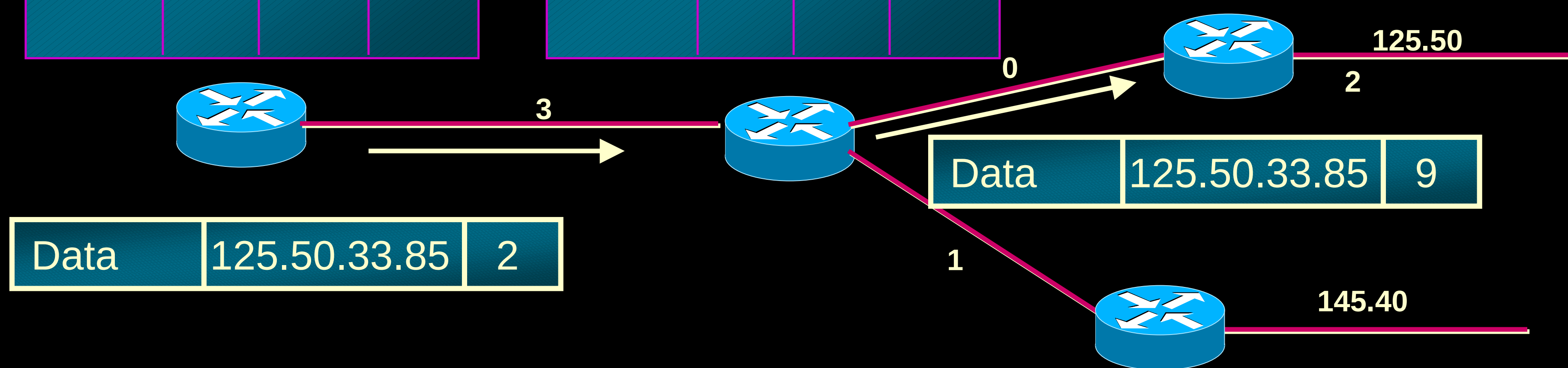


# *MPLS forwarding (contd)*

| Address Prefix | Out Path | In Label | Out Label |
|---|---|---|---|
| 125.50 | 3 | | 2 |
| 145.40 | 3 | | 1 |
| | | | |

| Address Prefix | Out Path | In Label | Out Label |
|---|---|---|---|
| 125.50 | 0 | 2 | 9 |
| 145.40 | 1 | 1 | 8 |
| | | | |

| Address Prefix | Out Path | In Label | Out Label |
|---|---|---|---|
| 125.50 | 2 | 9 | |
| | | | |



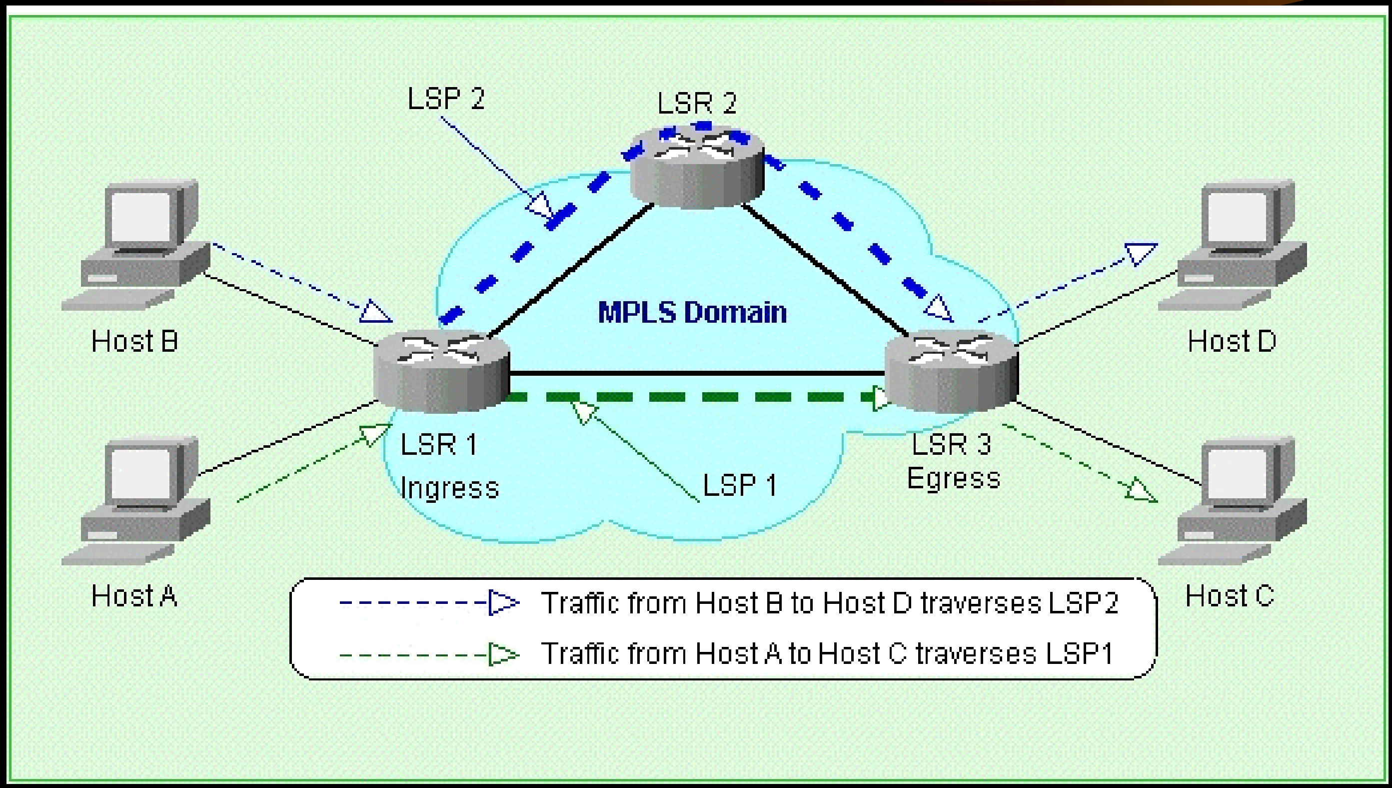
# *Multiprotocol Label Switching*

- Traditional IP Routing
- Disadvantages
- Need for MPLS
- MPLS basics and terminologies
- <u>Experiments</u>

# *Measurements of MPLS Traffic Engineering and QoS*

- Series of tests were run to evaluate the performance of TCP and UDP flows.
  - Tests include the effects of using different MPLS features on the performance of traffic flows.
- Goals:
  - Evaluating how well MPLS traffic engineering and QoS can improve the performance of today's Internet.
  - Identify opportunities for improvement and development of new mechanisms to ensure provision of traffic engineering as well as QoS/CoS features in future networks.

# *Experimental Network Configuration*



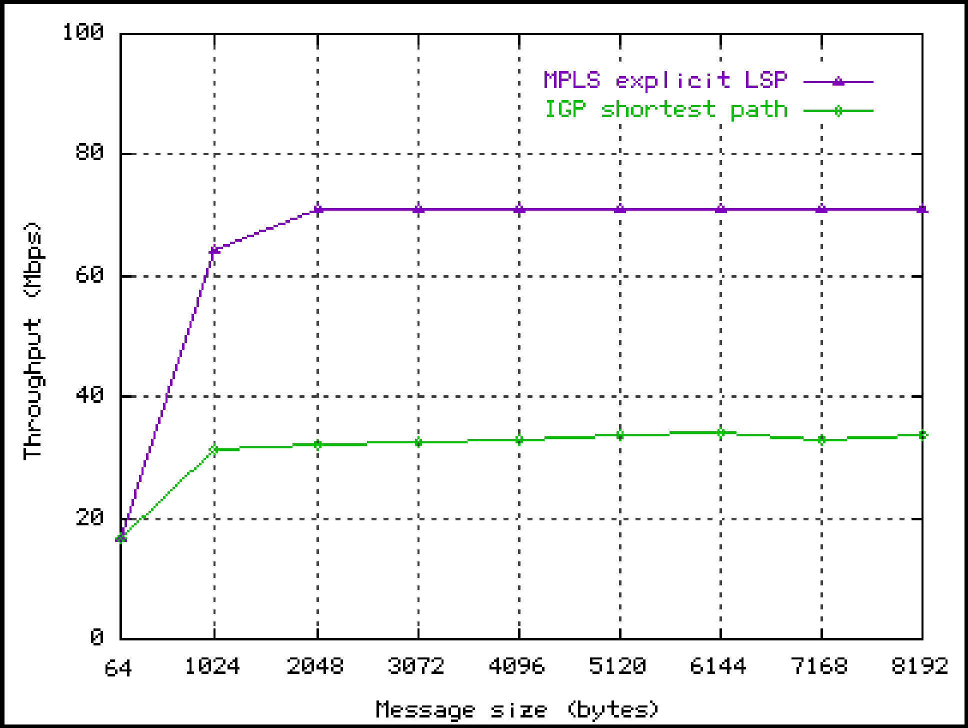
# *Network Description*

- Host Computers:
  - Intel Pentium II, 300MHz processors, 128 MB RAM.
  - Equipped with Fast Ethernet NICs and running FreeBSD 4.1.
  - Connected to the MPLS domain using 100Base-T connections via Gigabit Ethernet switches.
- Label Switched Routers:
  - Juniper Networks M40 routers running JUNOS Internet Software supporting Juniper Network's MPLS implementation.
  - Routers connected using OC-12 ATM links.
  - Distance between LSR1 and LSR3, LSR2 and LSR3 is about 40Km while LSR1 and LSR2 are 5Km apart.

# *Experiment Using MPLS Explicit LSPs*

- Minimize the effects of network congestion by using MPLS traffic engineering capability.
  - This is done by applying explicit routing.
- Scenario 1:
  - Two explicit LSPs are established between LSR1 and LSR3, both following the IGP shortest path.
- Scenario 2:
  - Two explicit LSPs set up again. However, traffic from host A to host C is made to traverse LSP2 while traffic from host B to host D flows across LSP1.

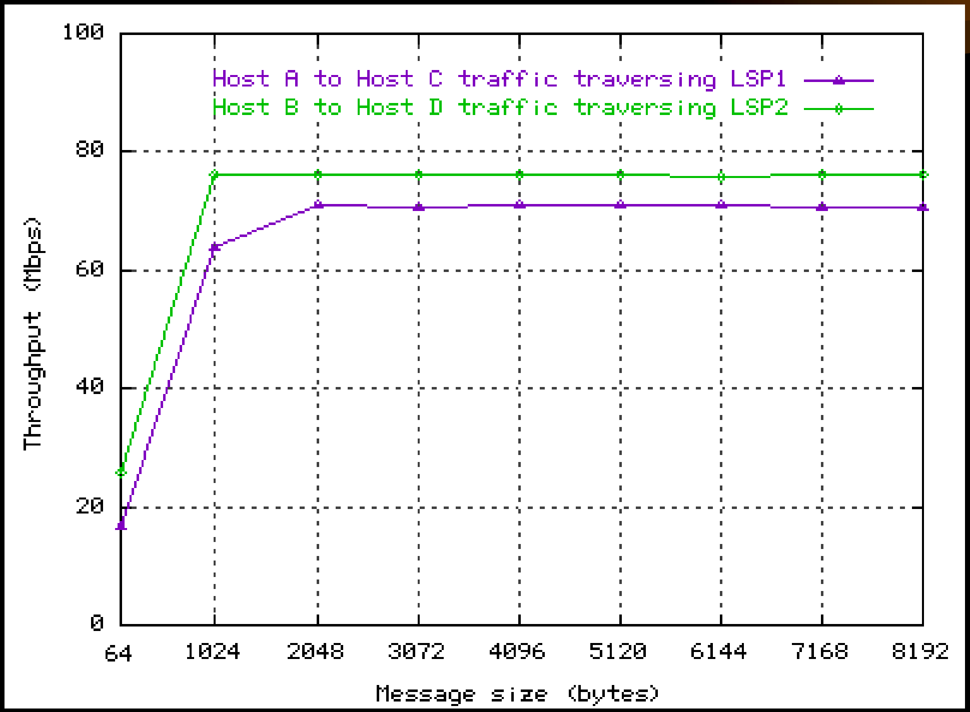
# *Results*




- Traffic from host A to host C is diverted to flow on the MPLS explicit path.
- Significant improvement of throughput over the IGP shortest path is observed.

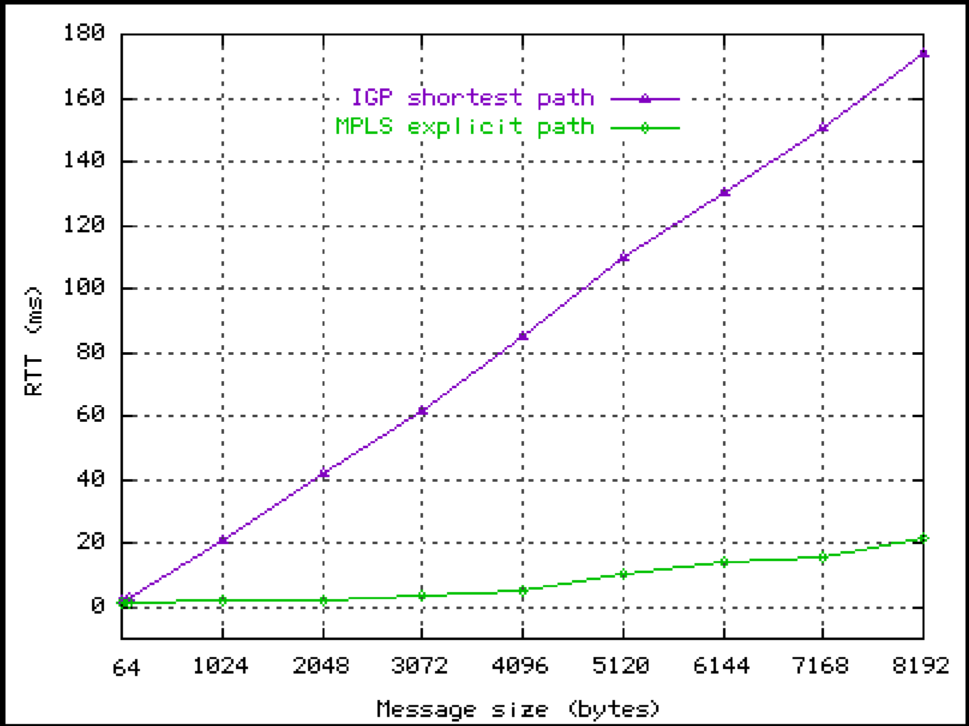
Throughput of TCP flow from Host A to Host C

# *Results (contd)*




Throughput of both flows

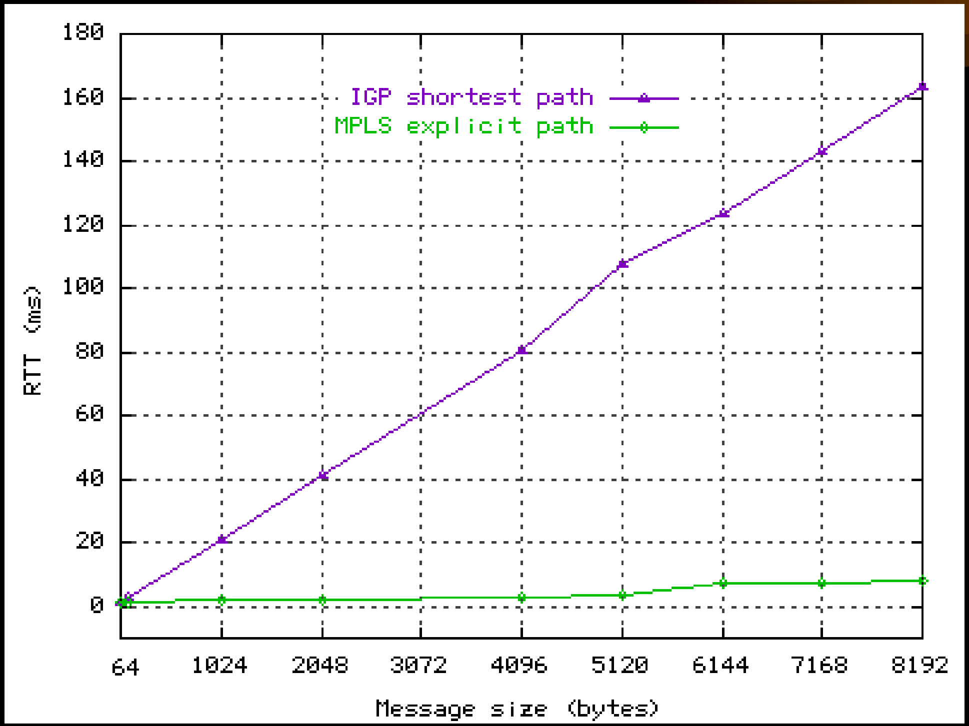
# Results(contd)




TCP average RTT

- Average RTT is measured using Netperf request/response method.
- RTT dramatically increases for congested IGP path, while it is minimal for packets traversing the MPLS explicit LSPs.

# *Results(contd)*




UDP average RTT

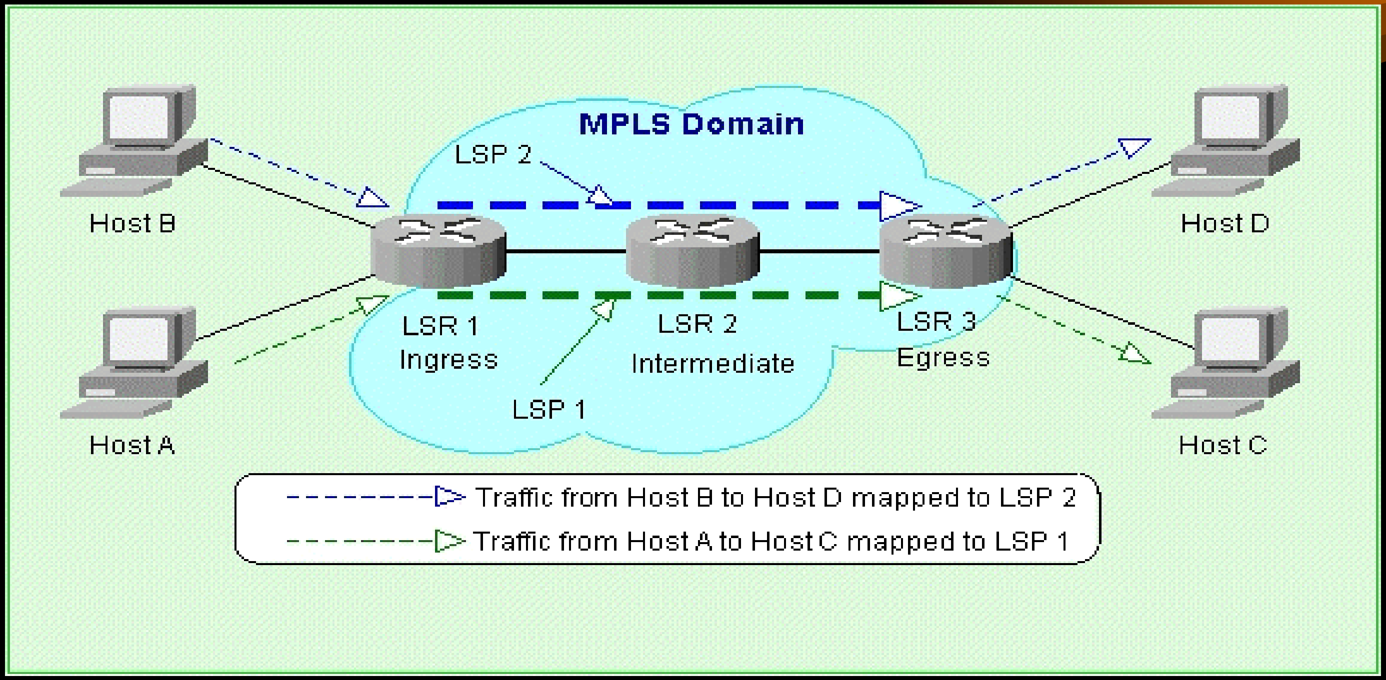
# *Experiment Using MPLS CoS/QoS*

- Study how MPLS can be used to provide guaranteed bandwidth and different levels of service for flows.
  - This is done by characterizing each LSP with a certain reserved bandwidth across the MPLS network.
  - Each LSP is also characterized with different CoS values.
- Network configuration is set up in such a way as to apply MPLS service differentiation along the same path.
- Reservation of bandwidth is done using the Committed Data Rate (CDR) QoS parameter in CR-LDP.

# *Assigning CoS Values*

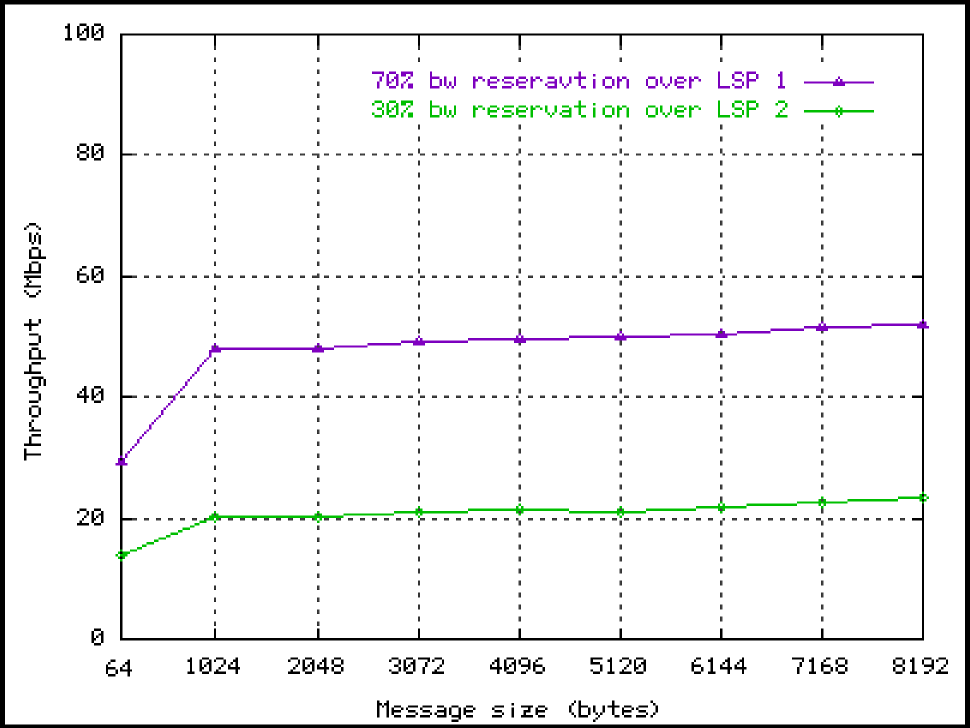
- EXP header is used. So, we have 8 different classes (0-7) to assign. A class indicates:
  - Output transmission queue to use, percent of the queue buffer to use, percent of link bandwidth to serve, packet loss priority to apply in presence of congestion.
  - Traffic with higher priority class receives better treatment than a lower priority class.
- Ingress router LSR1 is configured so that it can classify and map flows into LSP1 and LSP2 based on their destination address.
- The two LSPs are also configured with different CoS values.

# *Network Configuration For CoS Test*



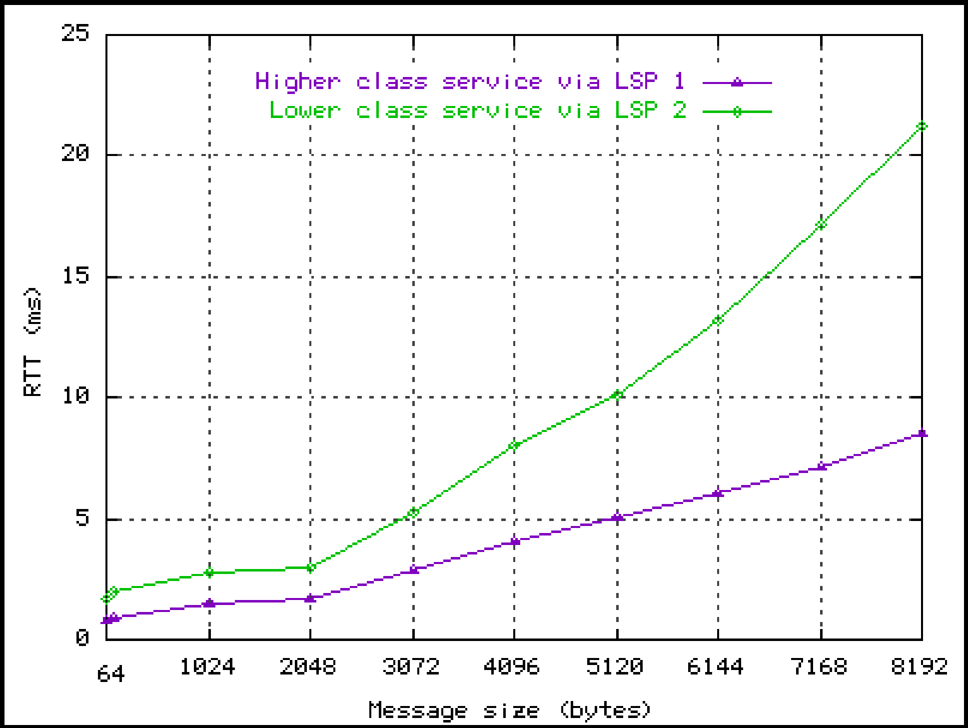

- 70 % bandwidth reserved for LSP1
- 30 % bandwidth reserved for LSP2

# *Bandwidth Reservation Over LSPs*




- This demonstrates how we can reserve resources in advance, as well as ensure guaranteed bandwidth.

# *Results*




- Traffic from LSP1 is offered a higher service level and delivered with lower latency.
- Service differentiation using MPLS CoS values has a significant impact on the performance of applications.

## *Conclusion*

- Providing QoS and traffic engineering capabilities in the Internet is very essential.
- For this purpose, the current Internet must be enhanced with new technologies such as MPLS.
- MPLS will play a key role in future service providers and carriers IP backbone networks.
- The use of MPLS in IP backbone networks will facilitate the development of new services such as real-time applications in the Internet.