Multiprotocol Label Switching (MPLS)

Engenharia de Computadores e Telemática
4º ano, 1º semestre, 2020/2021

Outcomes

- Understand the operation of MPLS, as a tool for improving transport efficiency
- Detail the LSP mechanisms
- Discuss QoS mechanisms applied to MPLS networks, and the difference between the basic concept
- Look into the complexity of the core network, and the multiple technologies that have been used in different contexts (and historically have evolved)

2

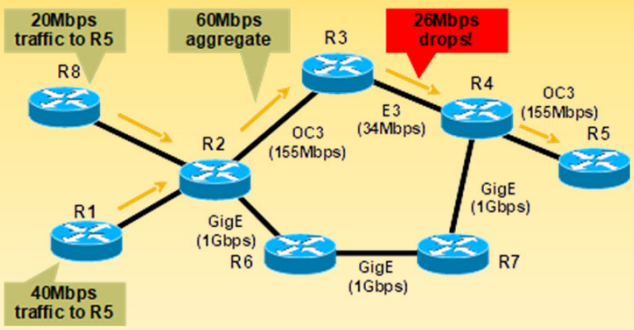
Traffic Engineering (TE)

- Network Engineering
  - Build your network to carry your predicted traffic!
  - Traffic patterns are impossible to predict!
  - Routing is based on the destination and does not allow to take the maximum possible advantage of the network resources.
  - IP source routing (using options field of IP header) is not usable in practice due to security reasons.
- Traffic Engineering
  - Manipulate your traffic path to fit your network!
    - Can be done with routing protocol costs (difficult deployment), or MPLS.
    - With RIP or OSPF or ANY OTHER IGP it is not possible to condition multiple traffic flows.
  - Increase efficiency of bandwidth resources.
    - Prevent over-utilized (congested) links whilst other links are under-utilized.
  - Ensure the most desirable/appropriate path for some/all traffic.
    - Override the shortest path selected by the routing protocols.

3

Example – avoiding congestion

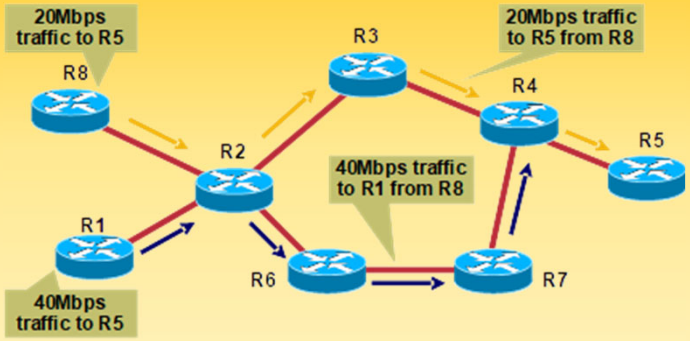
- On IP networks, *IntServ* and *DiffServ* are "routing independent architectures", retain the issues from routing
- IP network routing is based on the destination and does not allow to take the maximum possible advantage of the network resources
  - Shortest path will lead to congestion, even with available resources in the core
  - With **RIP or OSPF or ANY OTHER IGP** it is not possible to condition both flows.



6

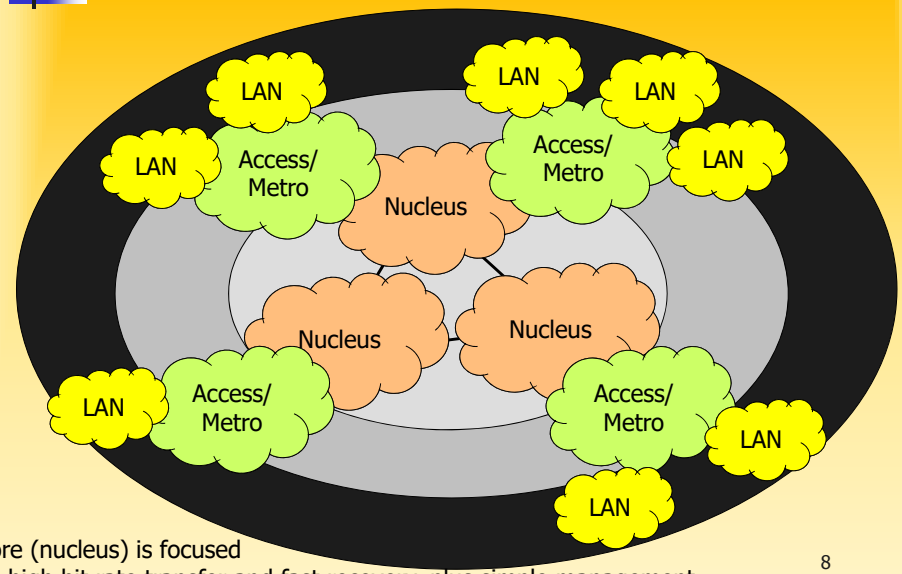
Using TE to solve… (source based routing)

- Tunnels (virtual entities) explore all capacity
  - Packets will transport, from their source, a list of routers' addresses that define their path to the destination (*Options* field of the IP datagram header)



7

Network architecture of a public operator



Core (nucleus) is focused
on high bit rate transfer and fast recovery, plus simple management

8

Leased lines are used to connect IP routers.

Network that offers leased line service is called "transport network" by telecom industry

Leased line

IP

Circuit technologies: time/frequency division multiplexing

PDH: T1, T3 switch: Digital Cross Connect (DCS)	SONET/SDH: OC3-OC768 Switch: SONET/SDH crossconnects	DWDM: OTU1-OTU3 Switch: optical WDM crossconnects

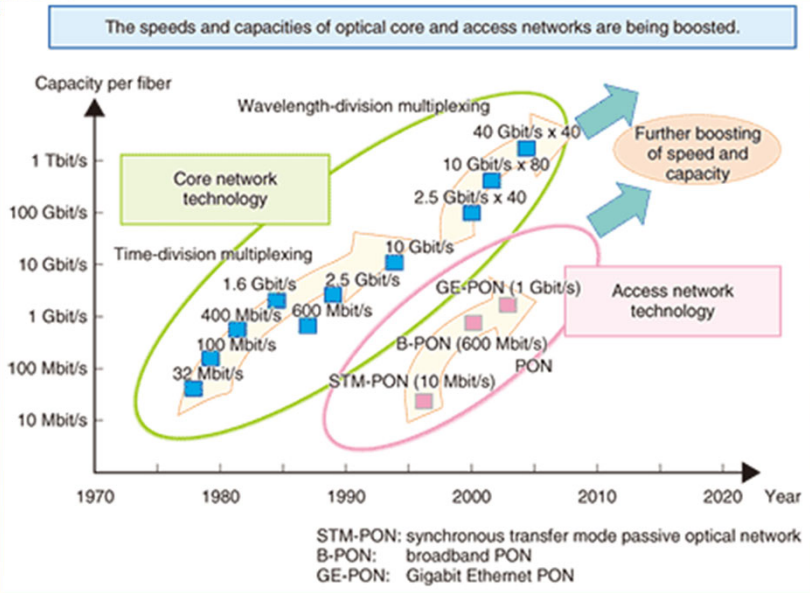
Packet technologies: virtual circuit switches

ATM	MPLS	Carrier-grade Ethernet

All the above technologies refer to data-plane technologies (control plane may be different)

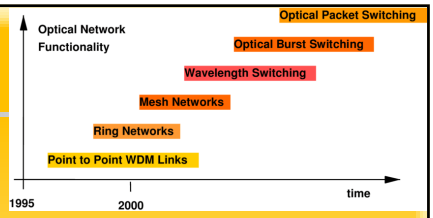
9

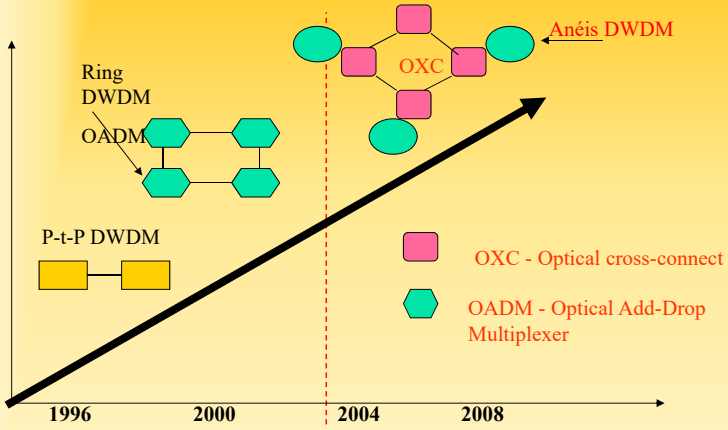
Optical Systems



10

Optical System Evolution



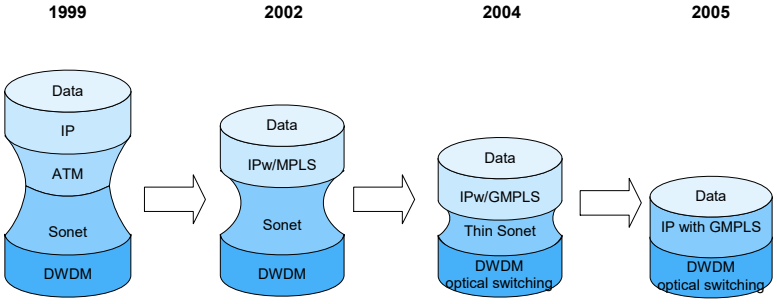


Improved optical infrastructure, increasingly optical switching.

11

Core network evolution

Increased simplification of the optical infrastructure



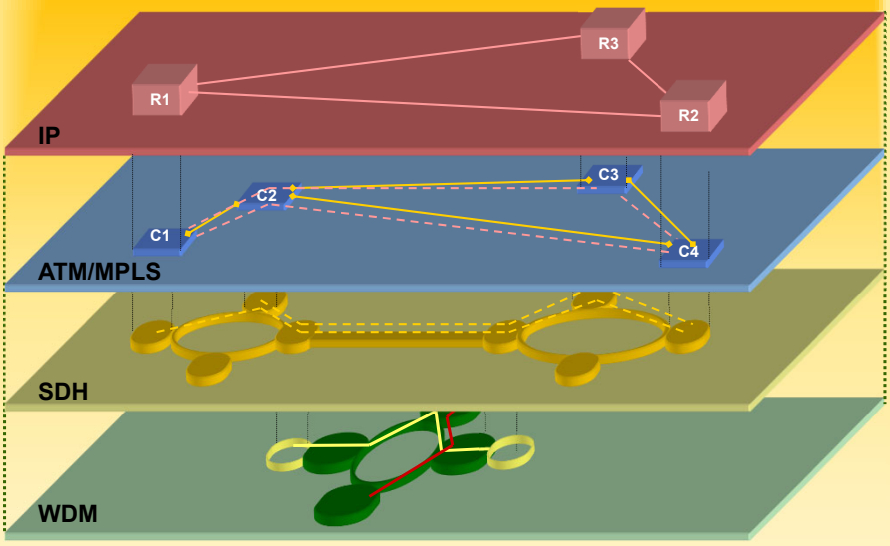
"GMPLS: An Overview of Routing and Management Enhancements, Ayan Banerjee et al"

12

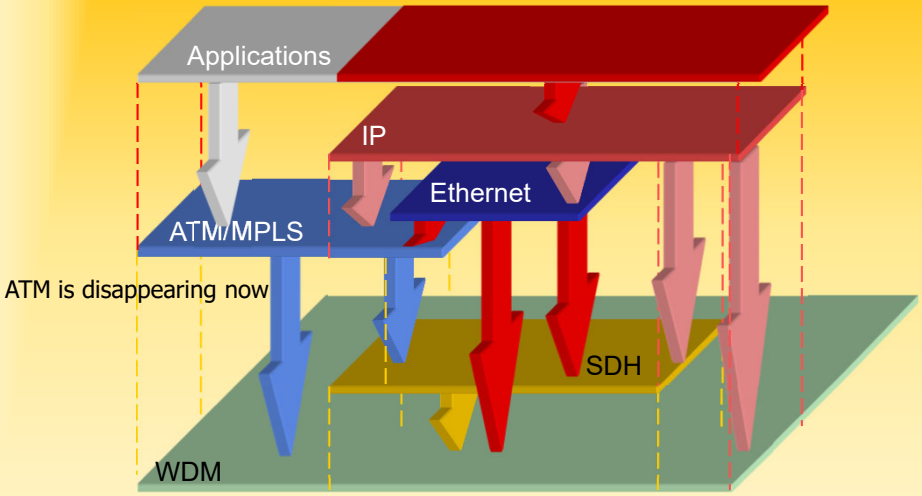
Why Migrating Core networks towards packet-based networks

- **Substitution of SDH/SONET circuit-switching transport network:**
  - Removing:
    - Inflexible bandwidth allocation of circuit switching
  - Keeping:
    - Connection Oriented
    - Deterministic performance and QoS
    - High resiliency and availability based on:
      - In-band OA&M
      - Protection with 50 ms recovery
    - Security
    - Simple, centralized operational model

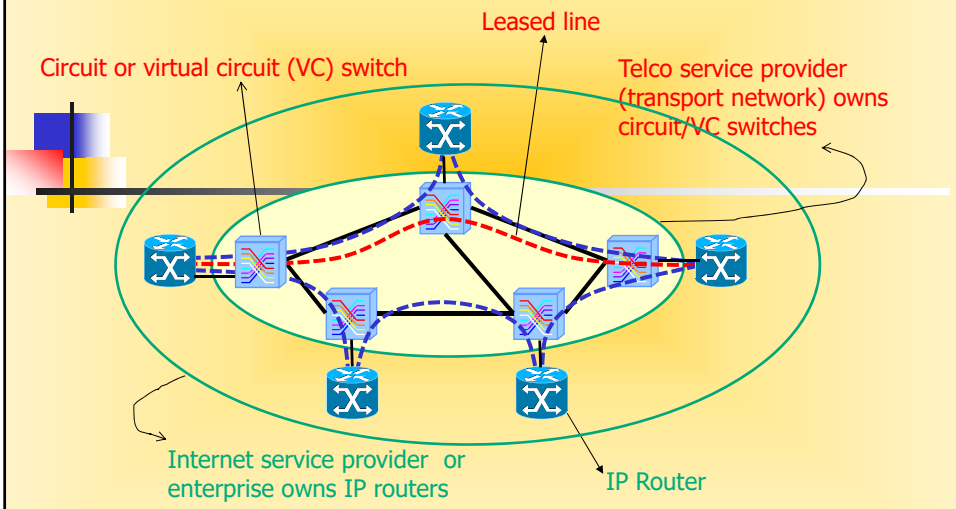
The physical reality: Overlay Networks



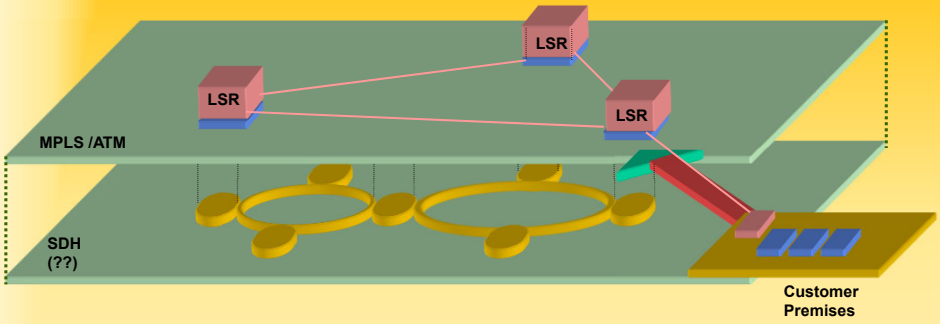
IP Transport Network as a Overlay



IP and leased line service deployment

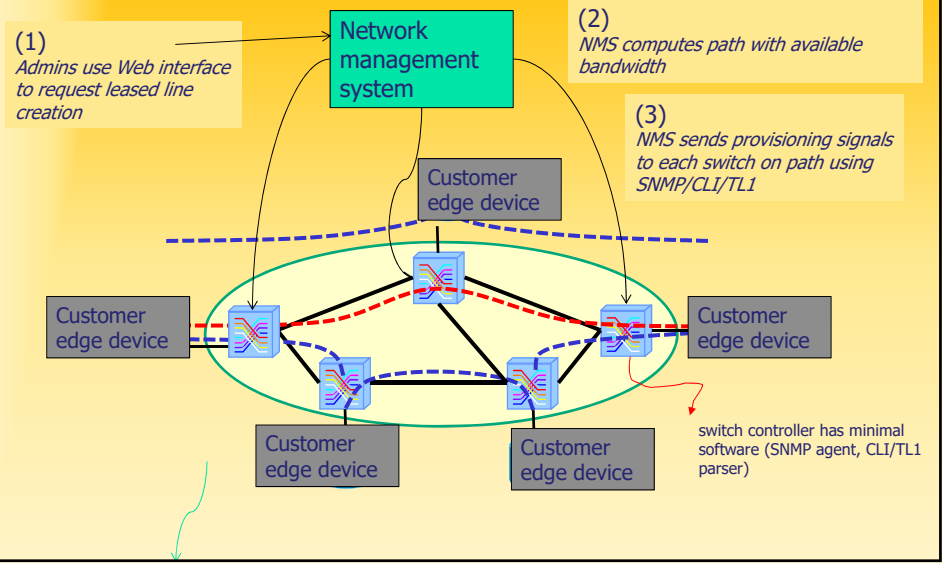


Switching Capacity not an issue any more, but some technology (ATM,MPLS) still needed



- Quality of Service
- Evolved VPN
- Traffic Engineering, Protection
- Multicast

Management plane (in transport network)



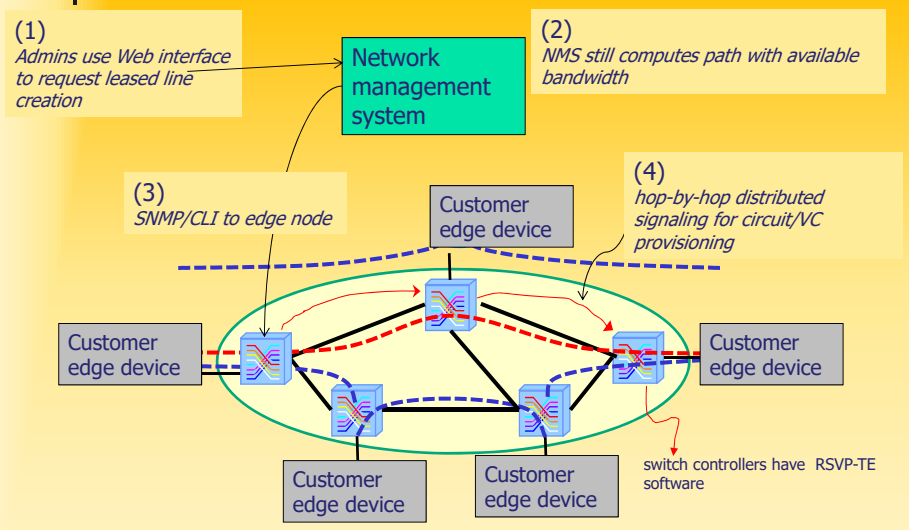
Migration towards packet-based networks

- **IP over MPLS/ATM**

  improves/guarantees packet-switching performance and survivability:

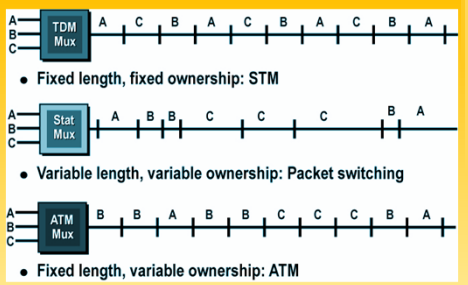
  - Removing:
    - IP Connectionless;
    - IP best-effort delivery with limited Quality of Service (QoS);
  - Keeping:
    - IP Statistical Multiplexing;
    - IP Dynamic Control Plane;
  - Offering:
    - Connection Oriented, Client agnostic;
    - Efficient bandwidth management with QoS;
    - Provisioning using centralized management;
    - Resiliency based on control-plane OA&M mechanisms and MPLS fast reroute;
    - High Scalability;

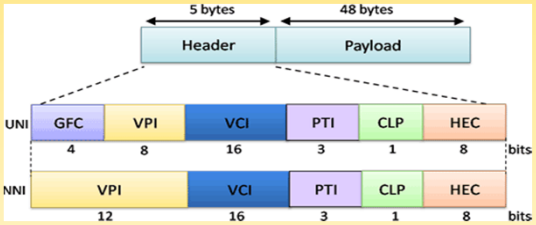
Management plane + control plane



Asynchronous Transfer Mode (ATM)

- ATM is a blend of Synchronous Transfer Mode (STM) and packet switching.
  - It has variable assignment, based on the arrival rate and delay sensitivity of the traffic.
  - After the assignment occurs, uses fixed-length time slots called cells.
    - Delay-sensitive traffic has immediate assignment
    - Data traffic can be temporarily buffered before being transmitted.
- Is a form of cell switching using small fixed-sized data units called cells.
  - 53 bytes: 5 header and 48 data
  - Developed to overcome switch eletronics limitations at the time





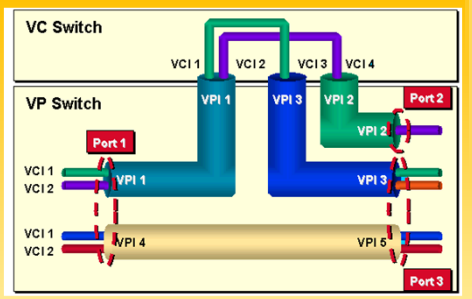
UNI (User-Network Interface).
NNI (Network-Network Interface)

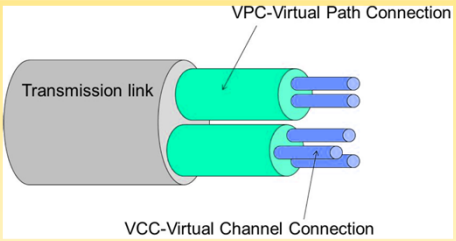
21

ATM Connections and Switching

- ATM is **connection-oriented.**
  - A connection (an ATM channel) must be established before any cells are sent.
  - Two levels of ATM connections:
    - **Virtual path connections.**
    - **Virtual channel connections**.
    - Indicated by two fields in the cell header:
      - Virtual Path Identifier: VPI.
      - Virtual Channel Identifier: VCI.
- Switching (very simple to do here) is based on VPI/VCI

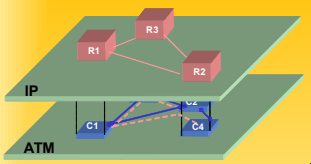
Port in	VPI/VCI	Port out	VPI/VCI
1	1/1	2	2/4
1	1/2	2	3/3
1	4/1	3	5/1
3	4/2	3	5/2





22

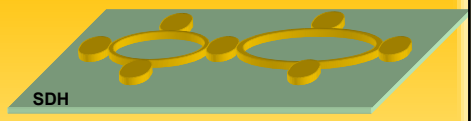
Main trends



LSR

IP and ATM integration
Label Swapping Paradigm

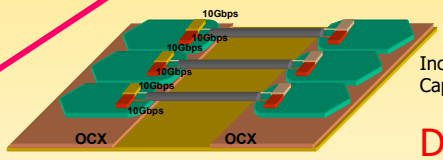
MPLS



Rapid and Predictable Restoration
Standard Time Division Multiplexing

SONET/SDH
Dynamic Allocation and Control?

G-MPLS



Increasing
Capacity Requirements

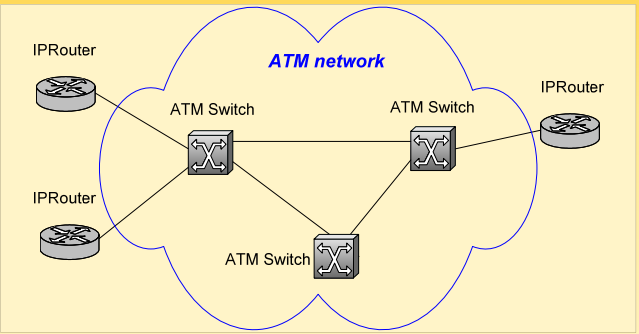
DWDM
Dynamic Allocation and Control?

MPLS

27

IP networks over ATM

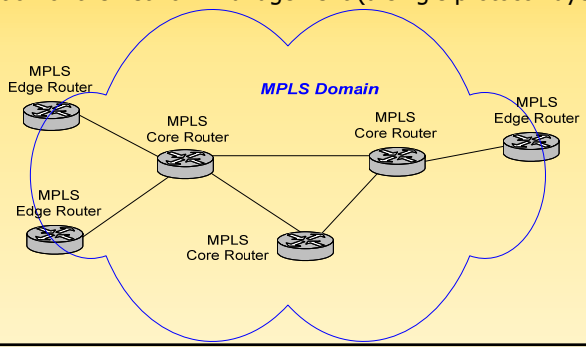
- IP routers are interconnected by an ATM network
- Connections between IP routers are implemented through virtual circuits (VCCs) or virtual paths (VPCs) on the ATM network
- It is necessary to manage two protocol layers
  (ATM is not available anymore on new networks)



28

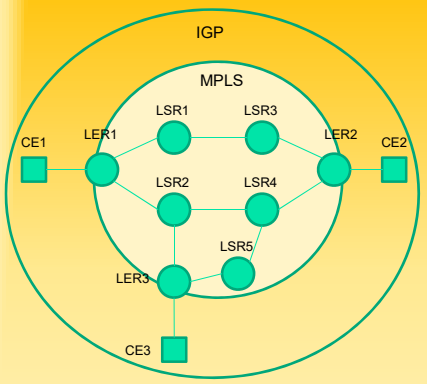
MPLS networks

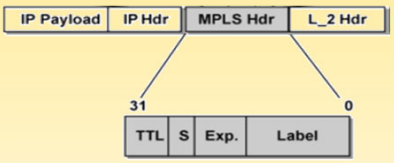
- Packets are labeled at the source with the label of the first hop
- Routers route packets based on their labels, just like ATM does with the VCI and VPI fields
- Advantages
  - Simplification of the packet routing process on routers
  - Traffic engineering capability equivalent to ATM
  - Simplification of the network management (a single protocol layer)



29

IP/MPLS





- **FEC – Forwarding Equivalence Class**
  - Identifies group of MPLS packets that are handled in the same way;
- **LABEL**
  - Local identifier based on the concept of FEC;
- **LSR – Label Switching Router**
  - LSRs in the same domain form an MPLS domain, using label forwarding;
- **LER – Label Edge Router**
  - LERs interacts with devices outside the domain, using IP technologies;

IP/MPLS

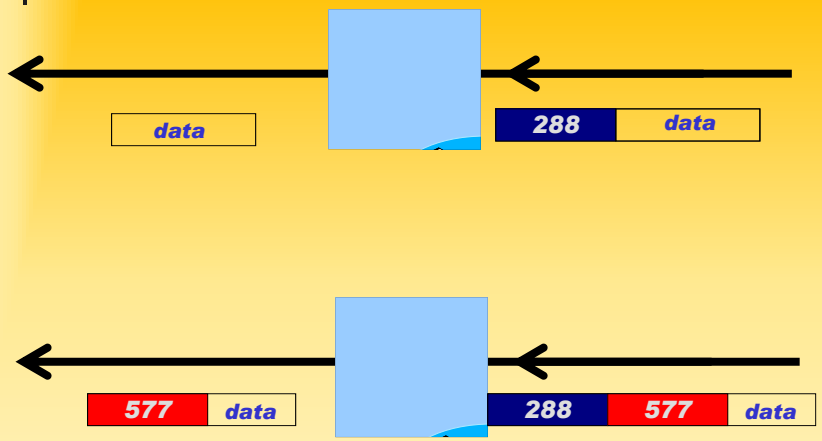
- **LSP – Label Switched Path**
  - Path along which FEC travels through an MPLS network;
- **LDP – Label Distribution Protocol**
  - The protocol used by MPLS for control that classifies FECs, distributes labels, and establish and maintains LSPs;
- **LFIB – Label Forwarding Information Base**
  - Forwarding table created by routing protocols and labels;
- **LSP tunneling**
  - Explicit route LSP between two LSRs that are not directly connected (indicated for TE);
- **Multi-level label stack**
  - Label stack carried by MPLS packets organized as LIFO (last-in first-out) stack;

Forwarding via Label Swapping

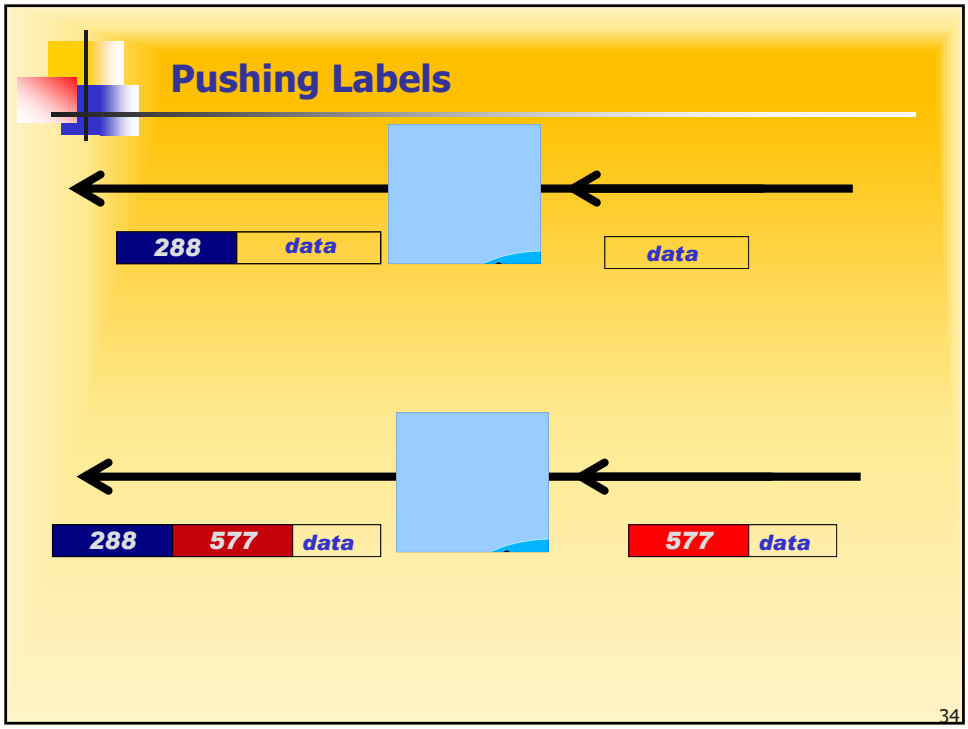
Labels are short, fixed-length values.

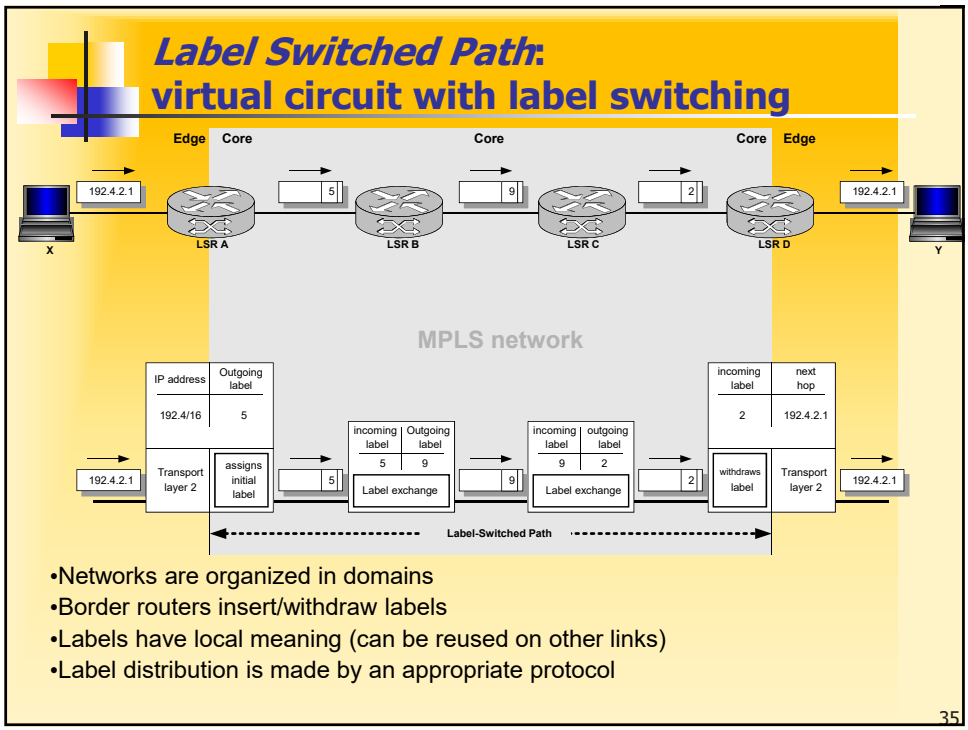
32

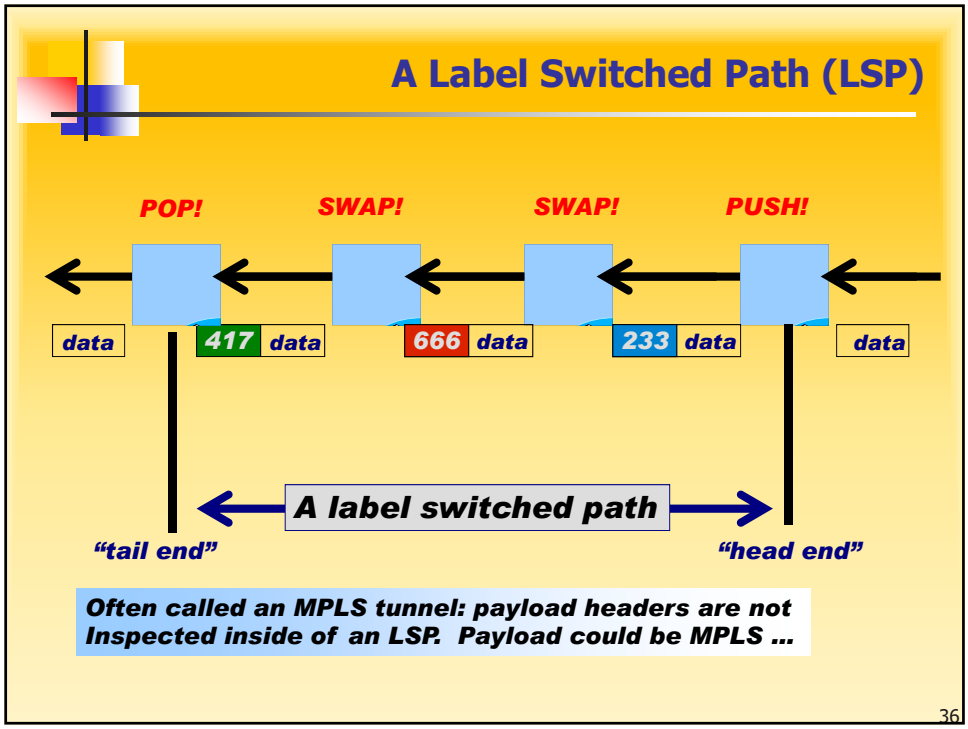
Popping Labels

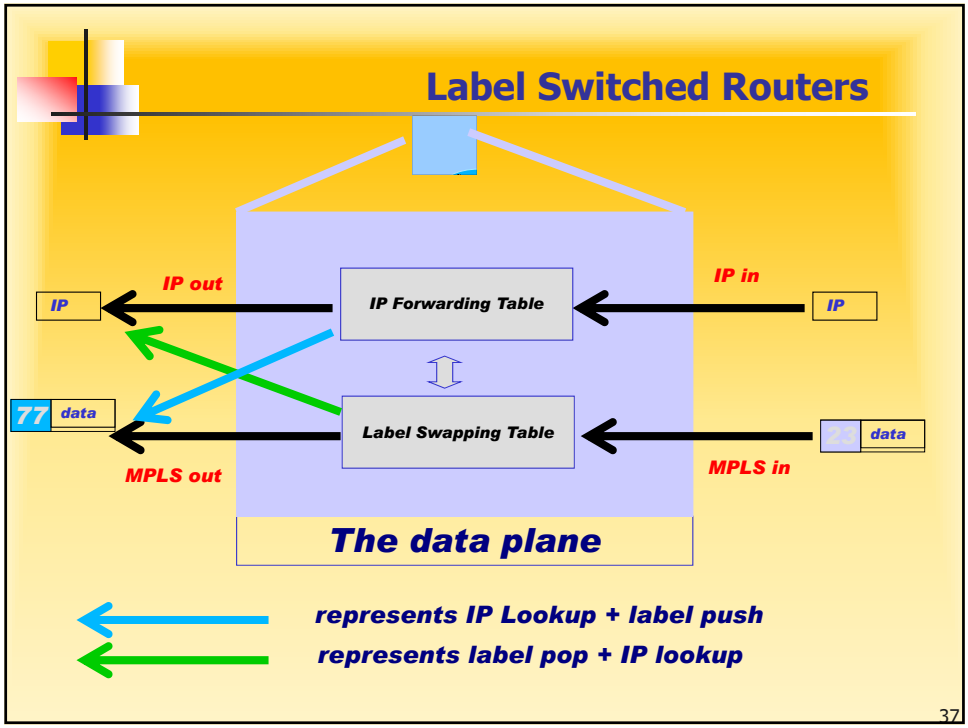


33

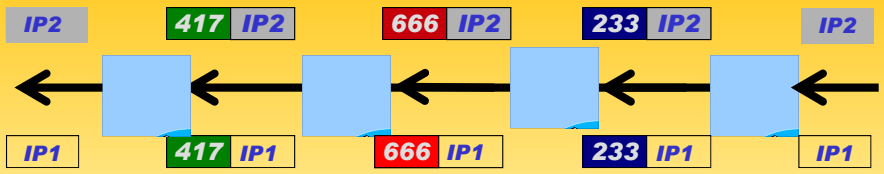

Pushing Labels
288
data
data
288
577
data
577
data
34


Label Switched Path:
virtual circuit with label switching
Edge
Core
Core
Core
Edge
192.4.2.1
5
9
2
192.4.2.1
X
LSR A
LSR B
LSR C
LSR D
Y
MPLS network
IP address
Outgoing label
192.4/16
5
incoming label
Outgoing label
5
9
incoming label
outgoing label
9
2
incoming label
next hop
2
192.4.2.1
192.4.2.1
Transport layer 2
assigns initial label
5
Label exchange
9
Label exchange
2
withdraws label
Transport layer 2
192.4.2.1
Label-Switched Path
•Networks are organized in domains
•Border routers insert/withdraw labels
•Labels have local meaning (can be reused on other links)
•Label distribution is made by an appropriate protocol
35


A Label Switched Path (LSP)
POP!
SWAP!
SWAP!
PUSH!
data
417 data
666 data
233 data
data
A label switched path
"tail end"
"head end"
Often called an MPLS tunnel: payload headers are not Inspected inside of an LSP. Payload could be MPLS ...
36


Label Switched Routers
IP out
IP
IP Forwarding Table
IP in
IP
77 data
Label Swapping Table
23 data
MPLS out
MPLS in
The data plane
represents IP Lookup + label push
represents label pop + IP lookup
37

Forwarding Equivalence Class (FEC)

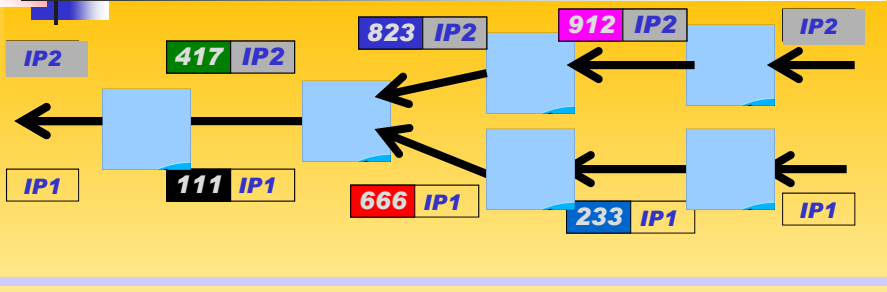


Packets IP1 and IP2 are forwarded in the same way --- they are in the same FEC.

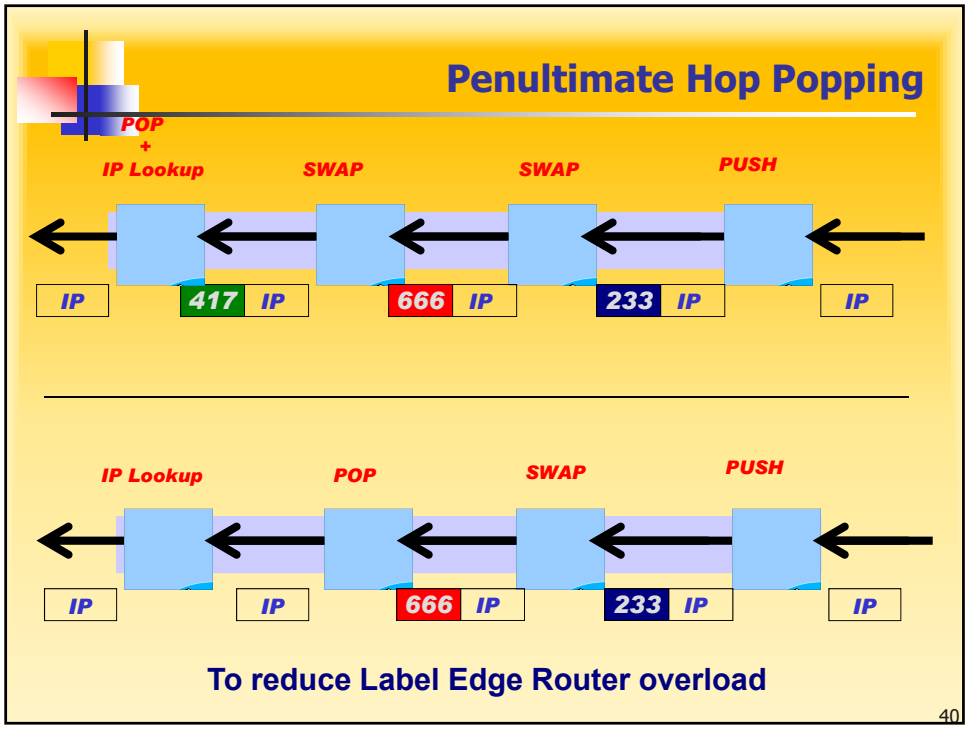
Network layer headers are not inspected inside an MPLS LSP. This means that inside of the tunnel the LSRs do not need full IP forwarding table.

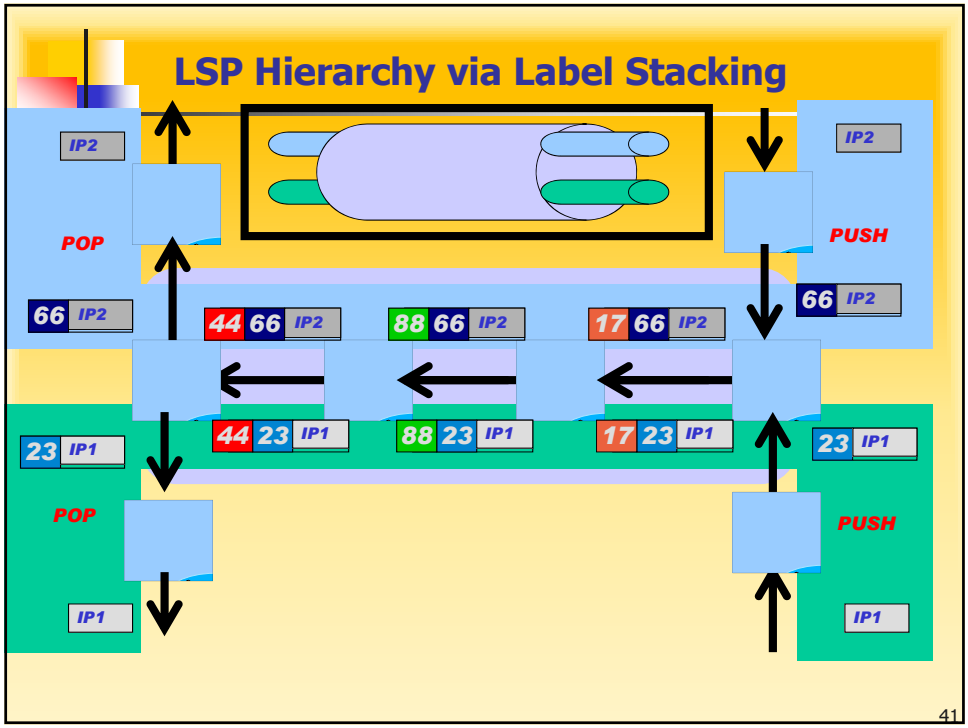
38

LSP Merge



39


Penultimate Hop Popping
POP
+
IP Lookup
SWAP
SWAP
PUSH
IP
417 IP
666 IP
233 IP
IP
IP Lookup
POP
SWAP
PUSH
IP
IP
666 IP
233 IP
IP
To reduce Label Edge Router overload
40


LSP Hierarchy via Label Stacking
IP2
POP
66 IP2
44 66 IP2
88 66 IP2
17 66 IP2
IP2
PUSH
66 IP2
23 IP1
44 23 IP1
88 23 IP1
17 23 IP1
23 IP1
POP
IP1
PUSH
IP1
41

IP/MPLS Network Establishment

- **Discovery link and topology**
  - IP routing table is built
    - LSRs and LERs use routing protocols to discover network topology eg. OSPF, ISIS, (BGP);
    - CEs advertise their addresses using routing protocols into MPLS cloud;
  - Forwarding Information Base (FIB) is built, initially without label information
- **Label Assignment**
  - FECs creation
    - LSRs classify with the same FEC all packets handled on the same way;
  - Allocate Labels
    - Every LSR allocates locally labels for every destination in the IP routing table (LIB and LFIB setup);

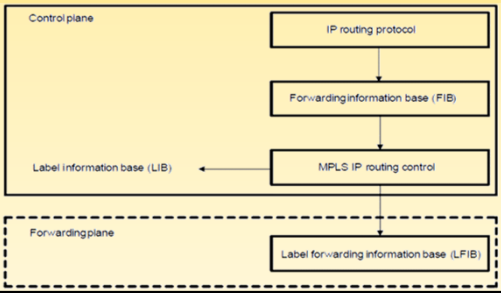
Label Distribution Protocols

- Unconstrained routing
  - Label Distribution Protocol (LDP).
  - Path is chosen based on IGP shortest path.
- Constrained routing
  - Constrained by explicit path definition and/or performance requirements (e.g., available bandwidth).
  - Resource Reservation Protocol with Traffic Engineering (RSVP-TE).
    - Evolution of RSVP to support traffic engineering and label distribution.
  - Constrained based Routing LDP (CR-LDP).
    - Evolution of LDP to support constrained routing.
    - Deprecated!
- MPLS VPN scope
  - MP-BGP using address family VPN IPv4 and family specific MP_REACH_NLRI attribute.

43

IP/MPLS Network Establishment

- **Label distribution operation and LSP Establishment**
  - Discovery
    - Basic Discovery – LSRs send LDP link Hellos UDP (multicast) for directly connected peers.
    - Extended Discovery – LSRs send LDP targeted Hellos UDP for a specific (remote) IP peer.
  - Session Establish and Maintenance
    - TCP session is established and it is maintained through periodically Keep-Alive messages
  - LFIBs are established accordingly with routing and Label tables.



Label Distribution Protocol (LDP)

- Dynamic distribution of label binding information.
- LSR discovery.
- Reliable transport with TCP.
- Incremental maintenance of label swapping tables (only deltas are exchanged).
- Designed to be extensible with Type-Length-Value (TLV) coding of messages.
- Modes of behavior that are negotiated during session initialization
  - Label distribution control (ordered or independent).
  - Label retention (liberal or conservative).
  - Label advertisement (unsolicited or on-demand).

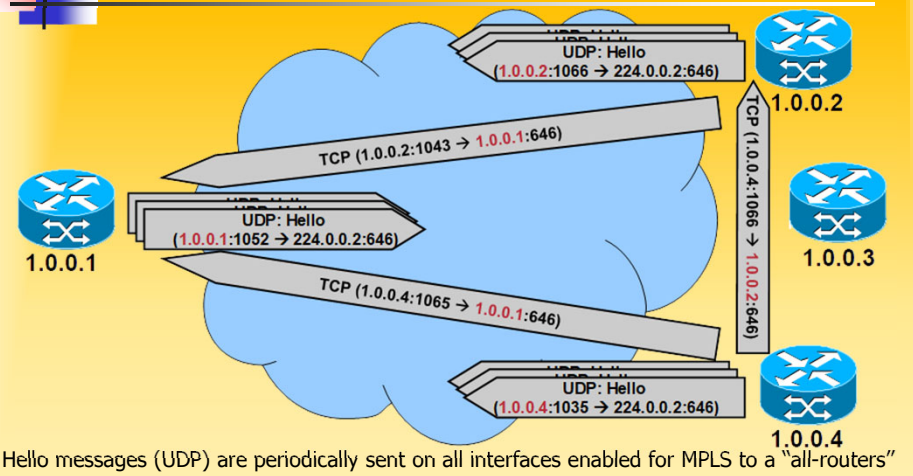
45

LDP Messages

- Discovery messages
  - Announce and maintain the presence of an LSR in a network.
  - **Hello Messages** (UDP) sent to "all-routers" multicast address.
  - Once neighbor is discovered, a LDP session is established over TCP.
- Session messages
  - Establish (**Initialization Message**) and maintain (**KeepAlive Message**) sessions between LDP peers.
- Advertisement messages
  - When a new LDP session is initialized and before sending label information an LSR advertises its interface addresses with one or more **Address Messages**.
  - An LSR withdraw previously advertised interface addresses with **Address Withdraw Messages**.
  - Create, change, and delete label mappings for FECs.
    - **Label Mapping, Label Request, Label Abort Request, Label Withdraw,** and **Label Release Messages.**
- Notification messages
  - Provide advisory information and to signal error information.

46

LDP Neighbour Discovery



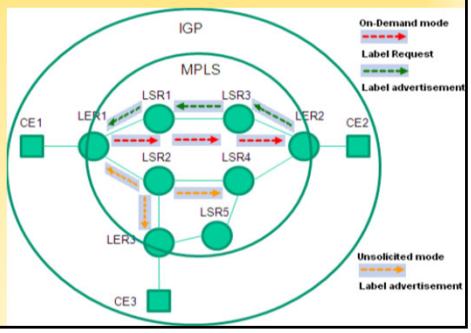
- Hello messages (UDP) are periodically sent on all interfaces enabled for MPLS to a "all-routers" multicast address (224.0.0.2).
- If there is another router on that interface it will respond by trying to establish a LDP/TCP session with the source of the hello messages.
- Both TCP and UDP messages use well-known LDP port number 646
- LDP Session is started by the router with higher IP address.

47

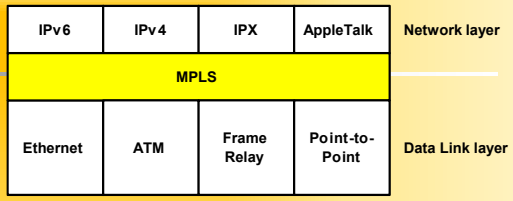
IP/MPLS Network Establishment

- LSP Establishment and Maintenance
  - Downstream On-demand mode– Upstream LSR sends a label request message (with FEC description) to its Downstream:
    - Ordered mode – a LSR only sends label (response) to its Upstream when it receives the label from its Downstream;
    - Independent mode – a LSR sends the label (response) when receives any request label;
  - Downstream Unsolicited mode – a Downstream LSR advertises label binding information to its Upstream LSR unsolicited after session to be established, without request ;

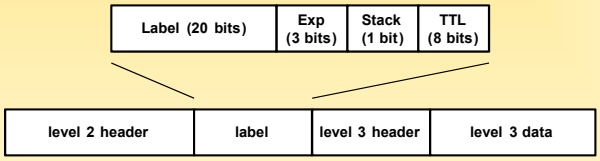
- Label retention mode
  - Conservative mode
    - LSRs keeps only the labels from next hops
    - Indicated for limited label space
  - Liberal mode
    - LSRs keeps any labels, even if those are not from next hops
    - Indicated for quick adaptation of route changes



Labels

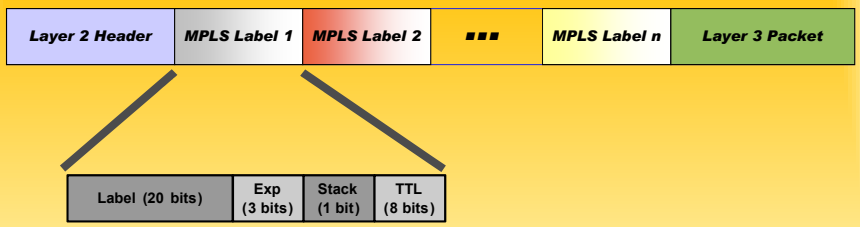


- On some Data Link (level 2) technologies, label is given by the appropriate fields of their header
  - ATM technology : VPI (Virtual Path ID) and VCI (Virtual Channel ID) fields
  - Frame Relay technology: DLCI (Data Link Connection Identifier) field
- On other Data Link technologies (Point-to-Point, Ethernet), the label is inserted between layer 2 and layer 3 headers



49

Generic MPLS Encapsulation



RFC 3032. MPLS Label Stack Encoding

- *Label: Label Value, 20 bits*
- *Exp: Experimental, 3 bits*
- *S: Bottom of Stack, 1 bit*
- *TTL: Time to Live, 8 bits*

50