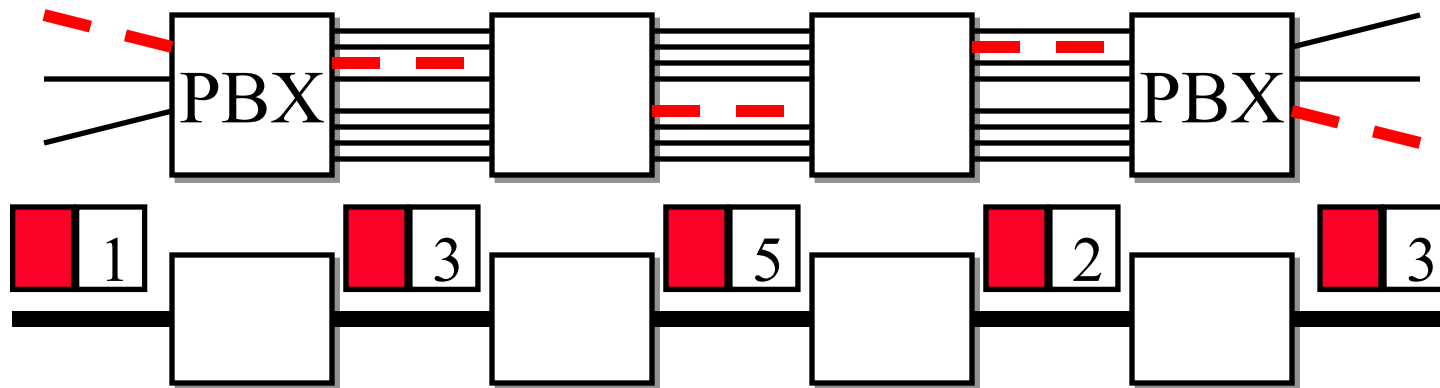


Carrier IP Networks: MPLS



Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides and audio/video recordings of this class lecture are at:
<http://www.cse.wustl.edu/~jain/cse570-15/>



1. Multiprotocol Label Switching (MPLS)
2. GMPLS, T-MPLS, MPLS-TP
3. Pseudo Wire: L2 Circuits over IP
4. Differentiated Services

Options to Connect Two Data Centers?



Danforth Campus



Medical Campus

1. Dedicated Optical fiber (leased from the phone company)
2. Ethernet over Optical Transport Network (all-Optical Switches)
3. Ethernet over Wavelength Division Multiplexing (DWDM)
4. Ethernet over Synchronous Digital Hierarchy (SDH)
5. Ethernet over Plesiochronous Hierarchy (PDH)
6. Ethernet over Pseudo-wire over MPLS
7. Ethernet over Micro-wave
8. Single Pair High-Speed Digital Subscriber Line (SHDSL)
9. Ethernet with enhancements

Plesiochronous Digital Hierarchy (PDH)

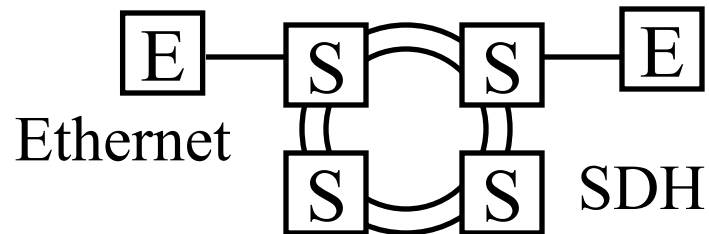
- ❑ Plesios + Synchronous = Near synchronous
- ❑ Phone Line = 64 kbps = 1 User channel
- ❑ North America
 - T1 = 1.544 Mbps = 24 User channels
 - T2 = 6.312 Mbps = 96 Channels
 - T3 = 44.736 Mbps = 480 Channels
- ❑ Europe:
 - E1 = 2.048 Mbps = 32 Channels
 - E2 = 8.448 Mbps = 128 Channels
 - E3 = 139.264 Mbps = 2048 Channels



T1

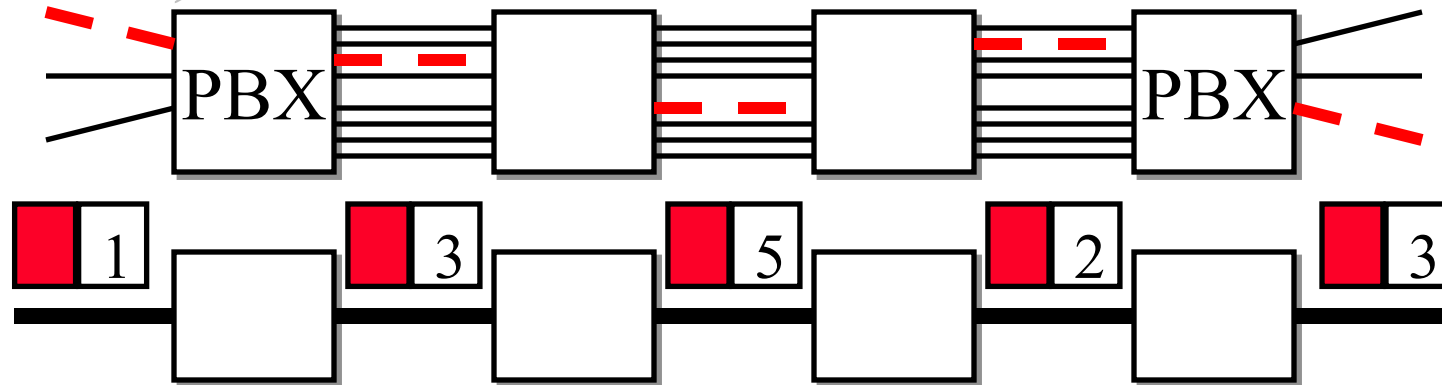


SONET/SDH

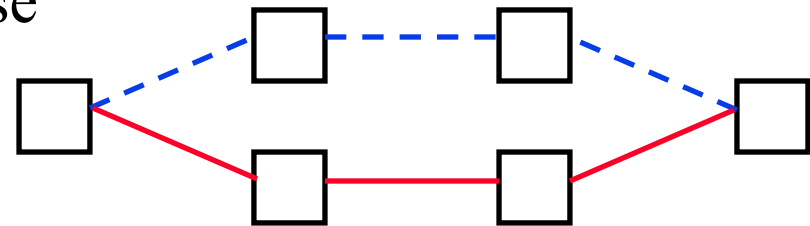


- ❑ SONET=Synchronous optical network
- ❑ Standard for digital optical transmission
- ❑ Standardized by ANSI and then by ITU
⇒ Synchronous Digital Hierarchy (SDH)
- ❑ Protection: Allows redundant Line or paths
- ❑ Fast Restoration: 50ms using rings
- ❑ Sophisticated management
- ❑ Ideal for Voice: No queues. Guaranteed delay
- ❑ Fixed Payload Rates: OC1=51.84 Mbps, OC3=155M, OC12=622M, OC48=2.4G, OC192=9.5G
Rates do not match data rates of 10M, 100M, 1G, 10G
- ❑ Static rates not suitable for bursty traffic
- ❑ One Payload per Stream ⇒ High Cost

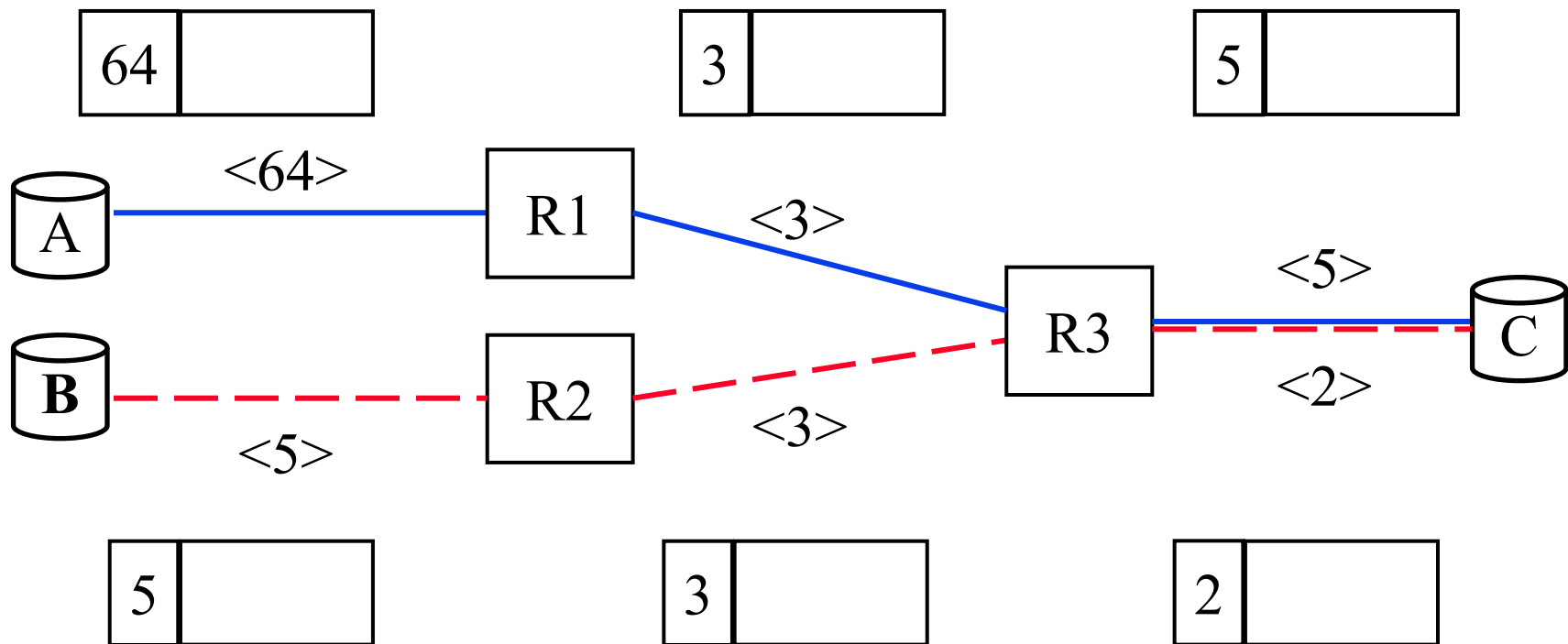
Multiprotocol Label Switching (MPLS)



- ❑ Allows virtual circuits in IP Networks (May 1996)
- ❑ Each packet has a virtual circuit number called 'label'
- ❑ Label determines the packet's queuing and forwarding
- ❑ Circuits are called Label Switched Paths (LSPs)
- ❑ LSP's have to be set up before use
- ❑ Allows traffic engineering

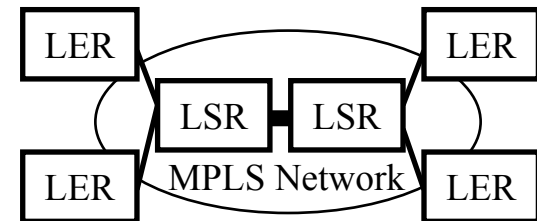


Label Switching Example



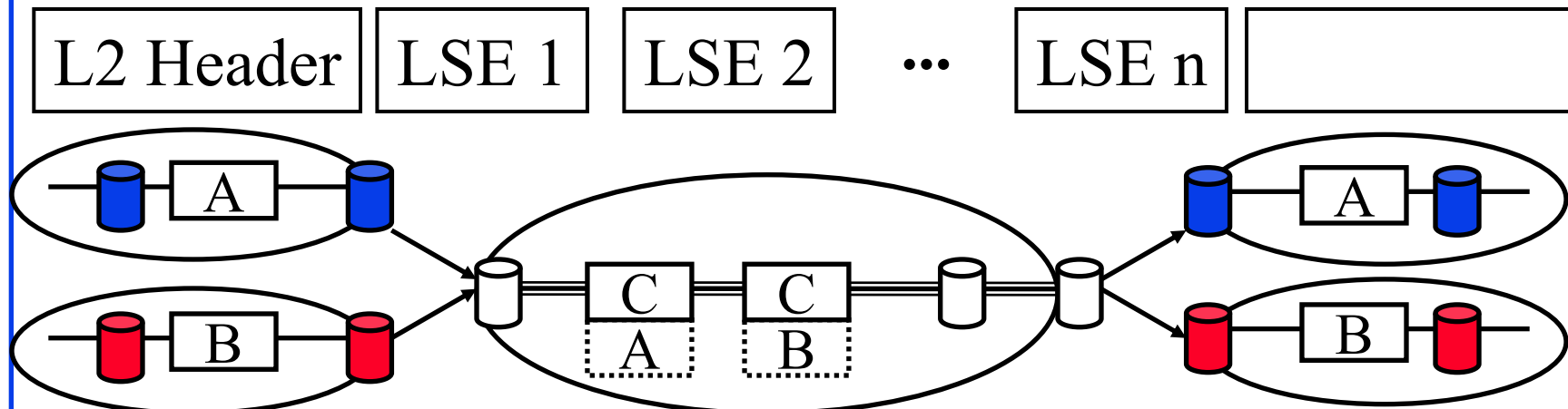
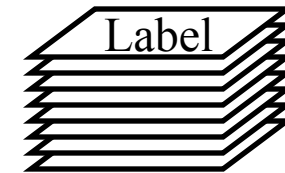
MPLS Concepts

- ❑ **Forwarding Equivalence Class (FEC):** All packets with the same top label
- ❑ **Label Switched Path (LSP):** End-to-end path from label push to label pop
- ❑ **Label Edge Router (LER):** Routers that push labels at the beginning of LSP and pop at the end
- ❑ **Label Switch Router (LSR):** Core routers that forward using the label
- ❑ **Label Forwarding Information Base (LFIB):** Forwarding table created using routing protocols, e.g., OSPF, BGP
- ❑ **Label Distribution Protocol (LDP):** Protocol to discover other MPLS routers and set up LSPs.
- ❑ **Resource ReSerVation Protocol with Traffic Engineering (RSVP-TE):** Alternative to LDP. BGP is also an alternative.



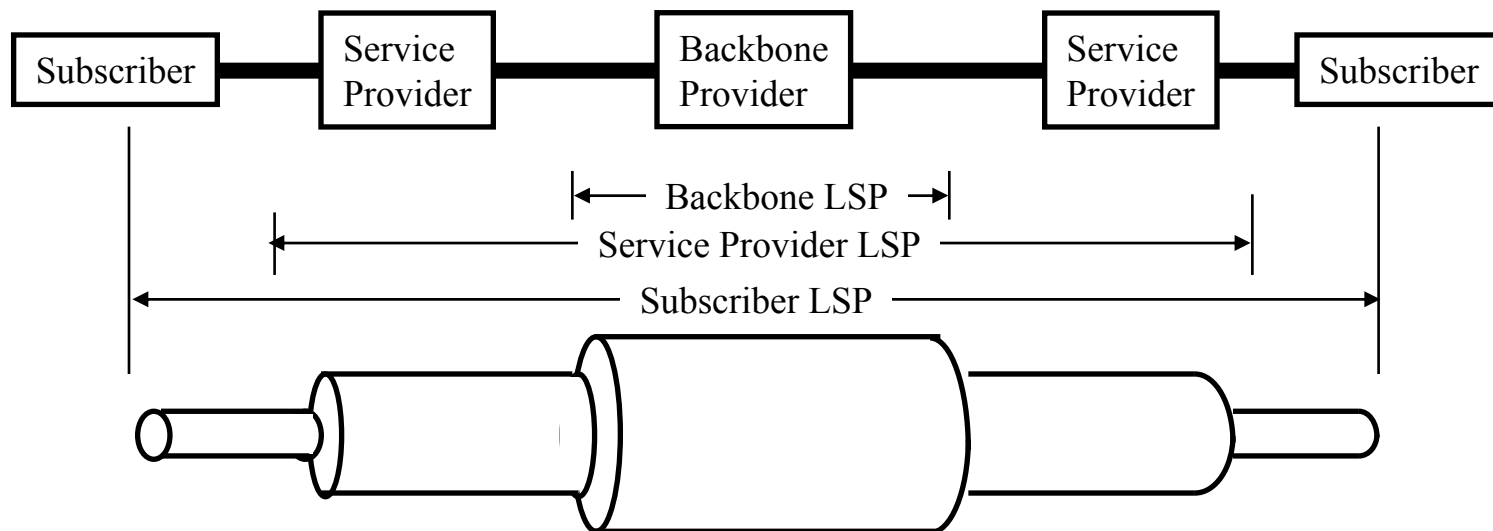
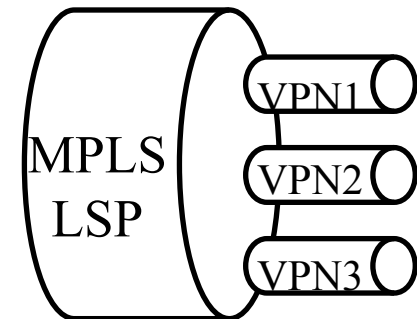
Label Stacks

- ❑ Labels are pushed/popped as they enter/leave MPLS domain
- ❑ Routers in the interior will use Interior Gateway Protocol (IGP) labels. Border gateway protocol (BGP) labels outside.
- ❑ Bottom label may indicate protocol (0=IPv4, 2=IPv6)



MPLS Label Stacking

- Label stacking allows:
 - Multiple levels of carriers.
 - Multiple VPNs in a single LSP
 - Multiple types of traffic in a single LSP



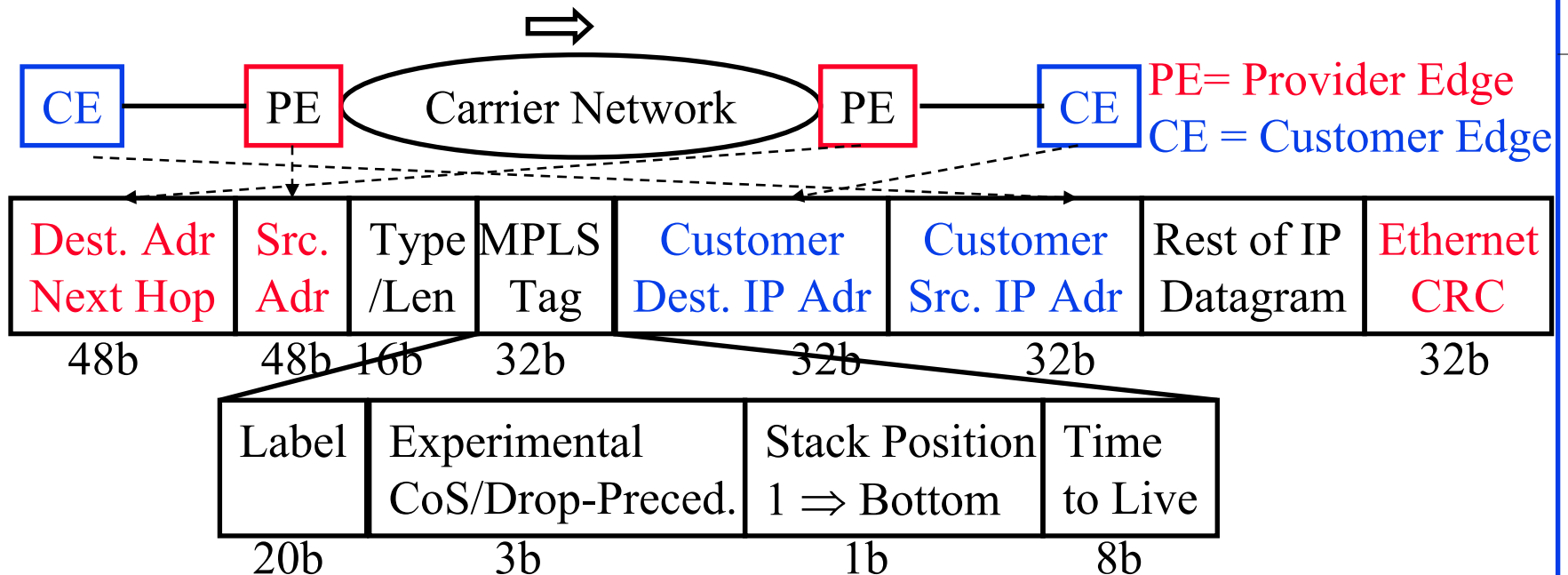
MPLS Traffic Engineering

- ❑ MPLS paths can be provisioned to follow a specific path (no need to use shortest path)
- ❑ Resources on the path can be reserved
- ❑ Multiple parallel LSPs can be established between the same pair of nodes
- ❑ Fault recovery via shifting traffic to standby LSPs

Label Assignment

- ❑ Unsolicited: Topology driven \Rightarrow Routing protocols exchange labels with routing information.
Many existing routing protocols are being extended: BGP, OSPF
- ❑ On-Demand: Label assigned when requested, e.g., when a packet arrives \Rightarrow latency
- ❑ Common MPLS Control Protocols:
 - Label Distribution Protocol called LDP
 - RSVP has been extended to allow label request and response (RSVP-TE)
 - Border Gateway Protocol (BGP): For signaling and discovery

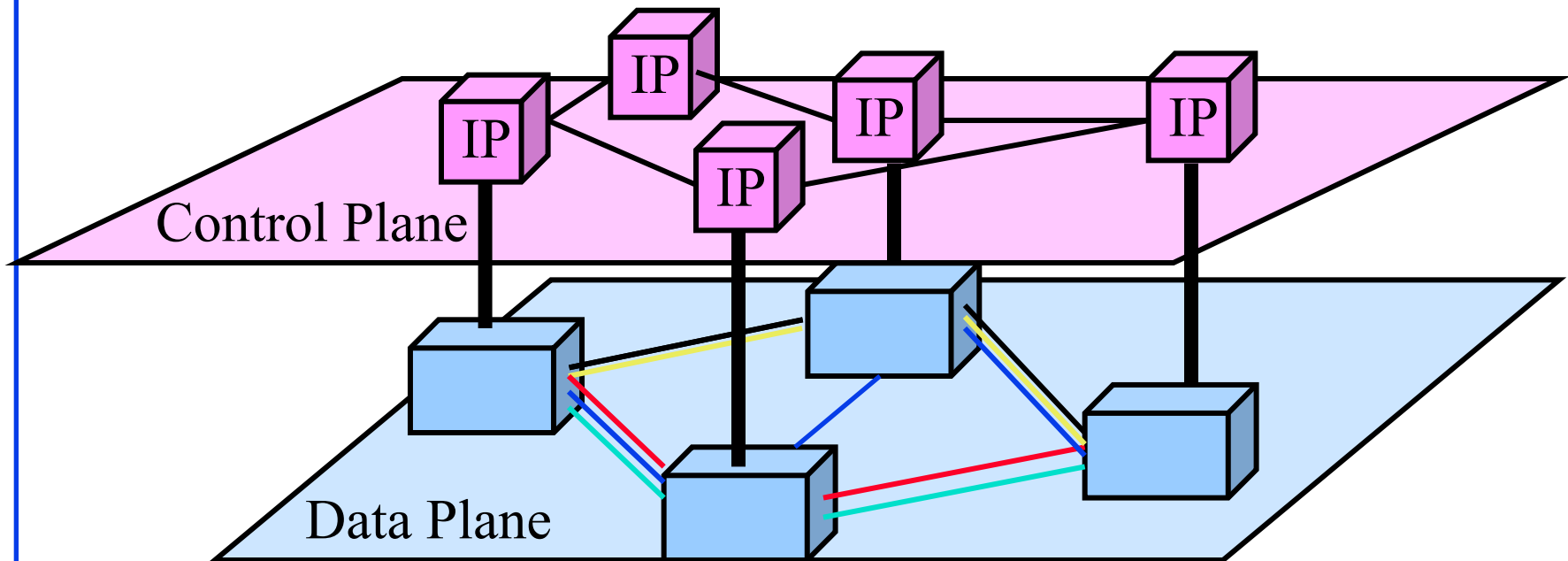
IP over MPLS over Ethernet



- ❑ Allows 2^{20} Label switched paths (LSP)
- ❑ Each path can have reserved capacity ⇒ Guaranteed **QoS**
- ❑ Explicit paths can be designed for specific traffic going to the same destination ⇒ **Traffic Engineering**
- ❑ Alternate paths can be set up for use if anything on the primary path fails ⇒ **Fast Reroute** ⇒ MPLS became a very popular

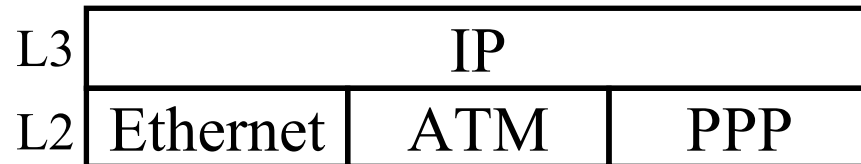
GMPLS

- ❑ Data Plane = Wavelengths, Fibers, SONET Frames, Packets (October 2000)
- ❑ Two separate routes: Data route and control route
- ❑ Allows data plane connections between SONET ADMs, PXCs, FSCs, in addition to routers

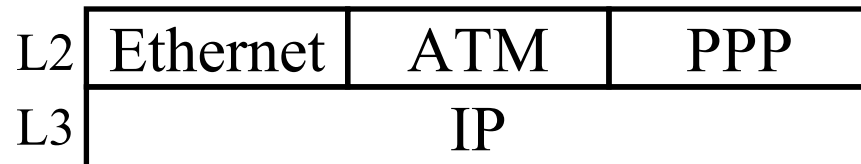


Martini Draft

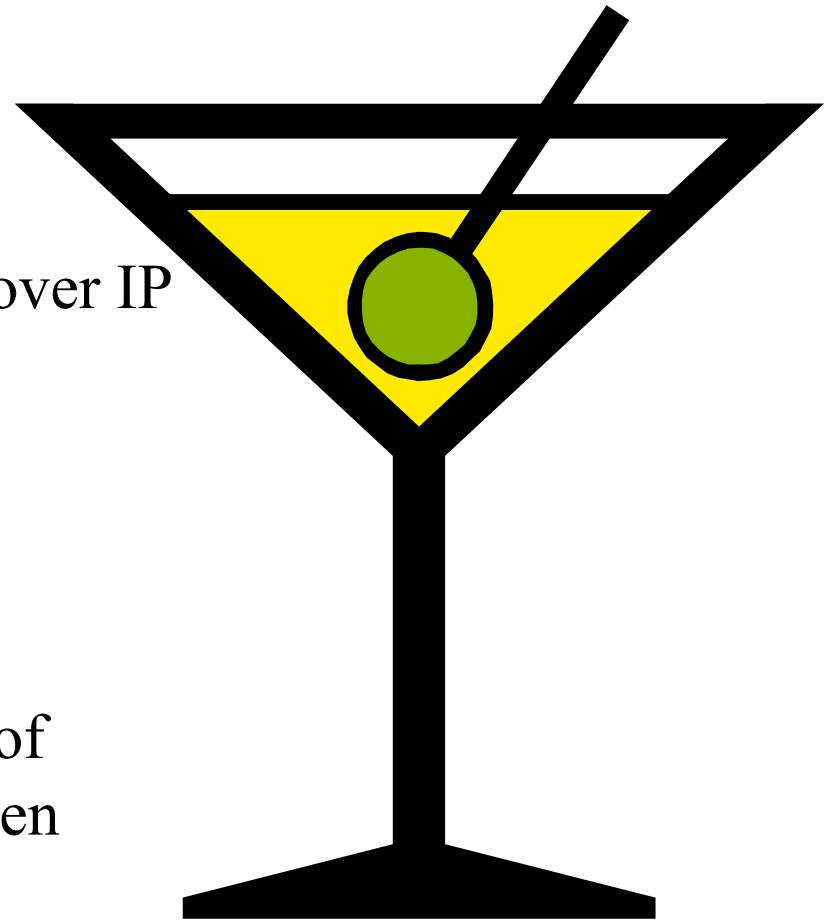
- 1995-1999: IP over ATM, IP over Ethernet



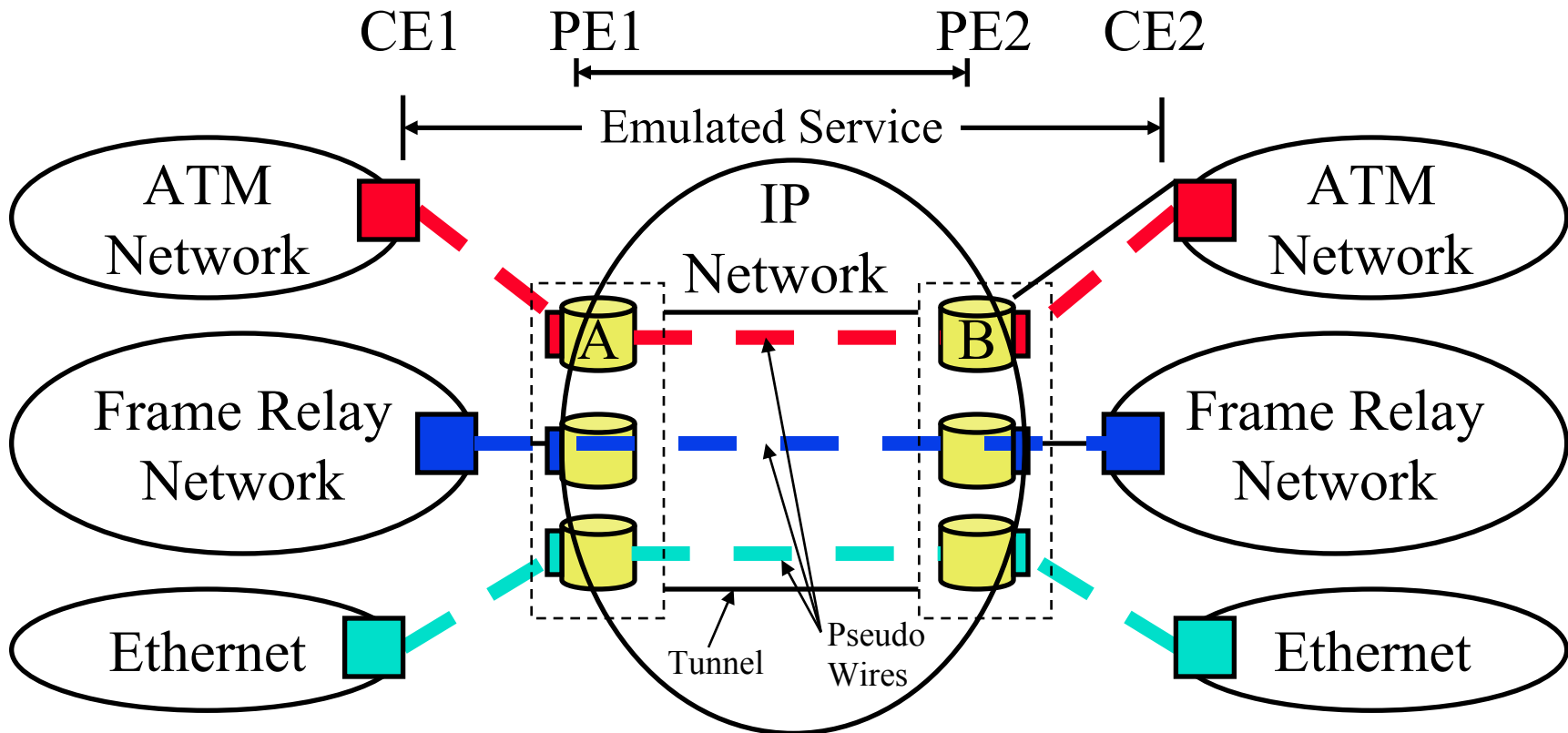
- 2000+: ATM over IP, Ethernet over IP
SONET over IP



- Idea proposed by *Luca Martini* of Level 3 Communications and then Cisco

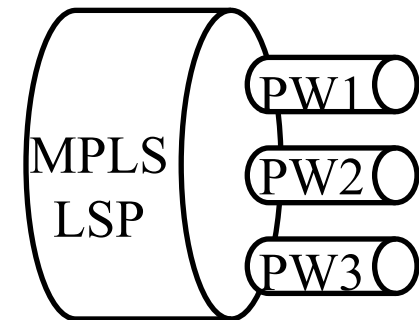
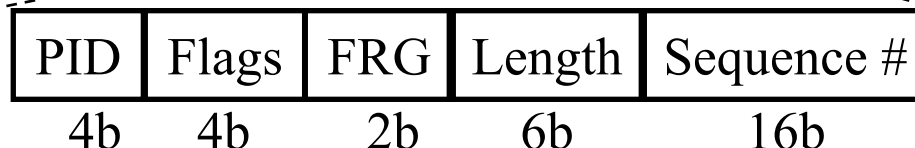


Pseudo Wire: L2 Circuits over IP



→ Payload Type → How to de-assemble payload
 → MPLS/GRE/L2TP - How to get to egress

Ethernet over PWE3 over MPLS



- ❑ Pseudo-Wire Emulation Edge-to-Edge (PWE3)
- ❑ Multiple pseudo-wires per MPLS LSP
 - ⇒ Core routers use outer “*path*” label and not inner “*VC*” label
- ❑ PW (VC) label format is same as MPLS label with End-of-Stack=1 and TTL=1. PW label is inserted/removed at the edge.
- ❑ Payload ID (PID): 5=Untagged Ethernet, 4=VLAN tagged, ...
- ❑ 4⇒VLAN tag put by carrier and customers may or may not be relevant for forwarding. Determined administratively by PE.
- ❑ Flags: Payload specific. FRG: Used for fragmentation
- ❑ Pause frames are obeyed locally. Not transported.

T-MPLS

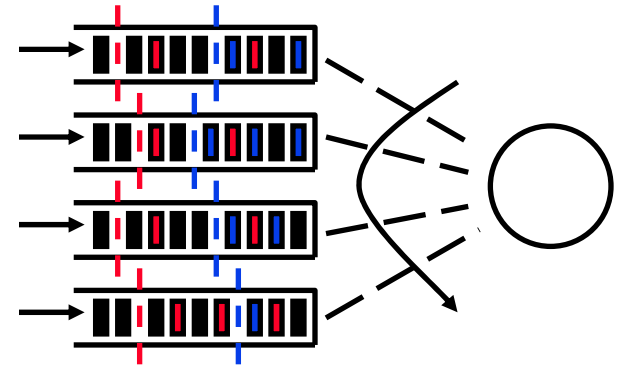
- ❑ A new profile for MPLS designed by ITU for carriers.
- ❑ No connectionless mode. No IP forwarding.
- ❑ Minimum IP-based control plane \Rightarrow Reduce cost
- ❑ All LSPs are bidirectional
- ❑ No penultimate hop option (PHP): PHP \Rightarrow Last LSR pops the stack before giving it to LER
- ❑ No equal cost multiple path (ECMP)
- ❑ Primary LSP and Backup LSP. Switching within 50 ms.
- ❑ Protection can be linear or ring

MPLS-TP

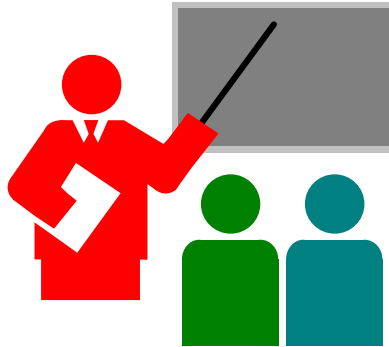
- ❑ Joint IETF and ITU effort to harmonize T-MPLS and MPLS-TE.
- ❑ Network provisioning via centralized network management system or distributed.
- ❑ Generalized Multiprotocol Label Switching (GMPLS), which is used for other transports, can be used for MPLS also.
- ❑ Comprehensive OAM for
 - fast detection, localization, troubleshooting, and
 - end-to-end SLA verification
 - Linear and ring protection with sub-50 ms recovery
 - Separation of control and data plane
 - Fully automated operation using NMS without control plane
⇒ No Label distribution protocol (LDP) or Resource Reservation Protocol with Traffic Engineering (RSVP-TE)

Differentiated Services

- ❑ A way for IP routers to provide QoS
- ❑ Expedited Forwarding (EF): Also known as Premium Service
 - Virtual leased line \Rightarrow Guaranteed minimum service rate
 - Policed: Arrival rate $<$ Minimum Service Rate
 - Not affected by other forwarding classes
- ❑ Assured Forwarding (AF):
 - Four Classes: No particular ordering
 - Three drop preference per class: Low, Medium, High
- ❑ Best Effort Service
- ❑ Differentiated Service Code Point (6 bits) encode the service, E.g., 101110 = EF



Summary



1. SONET, SDH, and PDH networks were designed for voice traffic
2. MPLS is used carriers to provide reliability and throughput guarantees similar to their previous networks
3. GMPLS extends MPLS to optical wavelengths
4. MPLS-TP is designed with OAM required for carriers
5. Differentiated services provide relative QoS guarantees using DSCP byte in the IP header

Reading List

- ❑ Bruce S. Davie, Adrian Farrel, "MPLS: Next Steps," Morgan Kaufmann, June 2008, ISBN: 978-0-12-374400-5, 432pp.
(Safari Book)

Wikipedia Links

- ❑ http://en.wikipedia.org/wiki/Differentiated_services
- ❑ http://en.wikipedia.org/wiki/Label-switched_path
- ❑ http://en.wikipedia.org/wiki/Label_Distribution_Protocol
- ❑ http://en.wikipedia.org/wiki/Link_protection
- ❑ <http://en.wikipedia.org/wiki/MPLS-TP>
- ❑ http://en.wikipedia.org/wiki/MPLS_local_protection
- ❑ http://en.wikipedia.org/wiki/MPLS_VPN
- ❑ http://en.wikipedia.org/wiki/Multiprotocol_Label_Switching
- ❑ http://en.wikipedia.org/wiki/Operations,_administration_and_management
- ❑ http://en.wikipedia.org/wiki/Optical_Carrier_transmission_rate

S

- ❑ http://en.wikipedia.org/wiki/Optical_Transport_Network

Wikipedia Links (Cont)

- ❑ <http://en.wikipedia.org/wiki/Pseudo-wire>
- ❑ http://en.wikipedia.org/wiki/Resilient_Packet_Ring
- ❑ http://en.wikipedia.org/wiki/Synchronous_optical_networking
- ❑ http://en.wikipedia.org/wiki/Traffic_policing
- ❑ http://en.wikipedia.org/wiki/Traffic_shaping
- ❑ http://en.wikipedia.org/wiki/Virtual_Private_LAN_Service
- ❑ http://en.wikipedia.org/wiki/Wavelength-division_multiplexing

Acronyms

- ❑ ADM Add-Drop Multiplexer
- ❑ AF Assured Forwarding
- ❑ ANSI American National Standards Institute
- ❑ ATM Asynchronous Transfer Mode
- ❑ BGP Border Gateway Protocol
- ❑ CoS Class of Service
- ❑ CRC Cyclic Redundancy Check
- ❑ DSCP Differentiated Services Code Points
- ❑ DWDM Dense Wavelength Division Multiplexing
- ❑ ECMP Equal-cost Multipathing
- ❑ EF Expedited Forwarding
- ❑ FCS Frame Check Sequence
- ❑ FEC Frame Equivalence Class
- ❑ FRG Fragment Bit
- ❑ FSC Fiber Switch Capable
- ❑ GMPLS Generalized Multi-Protocol Label Switching

Acronyms (Cont)

- ❑ GRE Generic Routing Encapsulation
- ❑ ID Identifier
- ❑ IETF Internet Engineering Task Force
- ❑ IGP Interior Gateway Protocol
- ❑ IP Internet Protocols
- ❑ ITU International Telecommunications Union
- ❑ LDP Label Distribution Protocol
- ❑ LER Label Edge Router
- ❑ LFIB Label Forwarding Information Base
- ❑ LSE Label Stack Entry
- ❑ LSP Label Switched Paths
- ❑ LSR Label Switching Router
- ❑ MPLS Multi-Protocol Label Switching
- ❑ NMS Network Management System
- ❑ OAM Operation, Administration and Maintenance
- ❑ OC Optical Carrier

Acronyms (Cont)

- ❑ OSPF Open Shortest Path First
- ❑ PBX Private Branch Exchange
- ❑ PDH Plesiochronous Digital Hierarchy
- ❑ PE Provider Edge
- ❑ PHP Penultimate Hop Option
- ❑ PW Pseudo-Wire
- ❑ PWE3 Pseudo-Wire Emulation Edge-to-Edge
- ❑ PXC Photonic Cross-Connect
- ❑ QoS Quality of Service
- ❑ RSVP Resource Reservation Protocol
- ❑ SDH Synchronous Digital Hierarchy
- ❑ SHDSL Single Pair High-Speed Digital Subscriber Line
- ❑ SLA Service Level Agreement
- ❑ SONET Synchronous optical network
- ❑ TE Traffic Engineering
- ❑ TP Transport Profile

Acronyms (Cont)

- ❑ TTL Time to Live
- ❑ VC Virtual Circuit
- ❑ VLAN Virtual Local Area Network
- ❑ VPN Virtual Private Network