

IP Over

DWDM

Raj Jain

The Ohio State University

Columbus, OH 43210

Jain@cse.ohio-State.Edu

These Slides are available on-line at:

http://www.cse.ohio-state.edu/~jain/talks/h_aipwd.htm



- ❑ Stack Debate: To SONET or Not to SONET?
- ❑ Why we have Multi-Layer Stack?
- ❑ What are the Problems with Multi-layer Stack?
- ❑ IP over DWDM Node Architecture and Issues
- ❑ Virtual Topology Issues
- ❑ Multiprotocol Lambda Switching
- ❑ IP/MPLS over DWDM

Stack Debate

1993	1996	1999	2000
IP	IP	IP	IP/MPLS
ATM	PPP	PPP	Sonet Framing
SONET	SONET		DWDM
DWDM	DWDM	DWDM	DWDM
Fiber	Fiber	Fiber	Fiber

SDL?

ATM provides voice+data integration
 Ignores Voice

PPP = Point to point protocol in HDLC-like framing

SDL = Simple Data Link

Simple Data Link

- ❑ Framing: How to tell where the frame begins and ends
- ❑ Two methods:
 - HDLC: 01111110 Flag
 - ❑ Need byte stuffing
 - ❑ Arbitrary increase in data rate
 - ❑ Need byte-level processing \Rightarrow slow
 - ATM: Header error check. Hunt and resync.
- ❑ SDL: Use HEC plus length (since variable size payload)

SONET Functions

- ❑ Clock Synchronization
- ❑ Rate Multiplexing/Traffic Grooming
- ❑ Rate Division/Inverse multiplexing
- ❑ Fault Tolerance
- ❑ Signal trace
- ❑ Error Monitoring
- ❑ Fault Isolation \Rightarrow Dual Ring
- ❑ Localized Decision \Rightarrow Fast Restoration

Multi-Layer Stack: Why?

- ❑ Speed: $\lambda > \text{SONET} > \text{ATM} > \text{IP}$
ATM < OC-12, IP < OC-3
Low speed devices \Rightarrow Not enough to fill a λ
SONET (1λ) limited to 10 Gbps
- ❑ Distance: End-system, Enterprise backbone, Carrier Access, Carrier Backbone, Core
- ❑ Some unique function in each layer
 - ATM = Access/Integration/Signaling/QoS/TM
 - SONET = Mux/Transport

Multi-layer Stack: Problems

- ❑ Increasing Bandwidth
 - ⇒ Core technologies move towards the edges
- ❑ Gigabit Routers ⇒ No need for grooming
 - One router port should be able to use all resources.
- ❑ Functional overlap:
 - Multiplexing:
 - DWDM $\lambda = \Sigma STM = \Sigma VC = \Sigma \text{Flows} = \Sigma \text{packets}$
 - Routing: DWDM, SONET, ATM, IP
 - QoS/Integration: ATM, IP
- ❑ Static division of bandwidth in SONET good for continuous traffic not for bursty traffic.

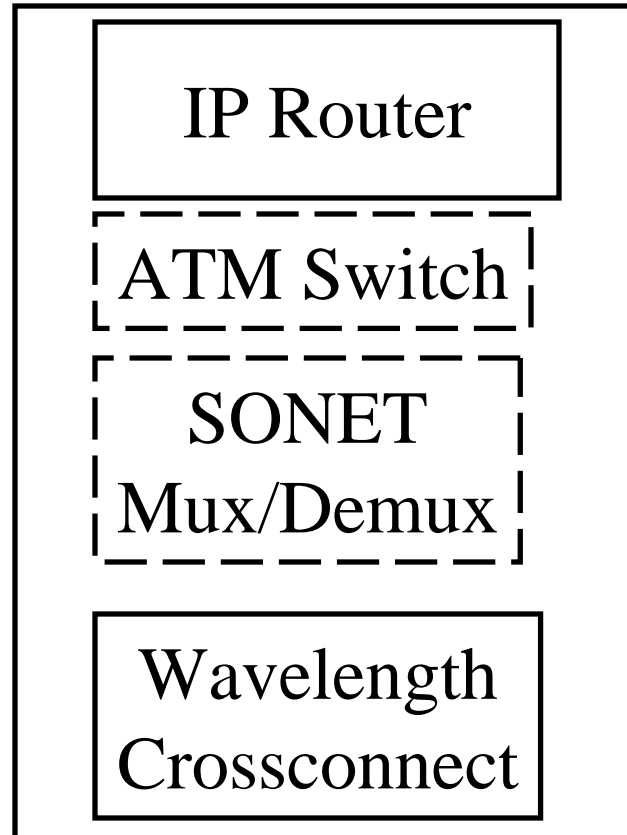
Multilayer Stack Problems (Cont)

- ❑ Failure affects multiple layers:
1 Fiber \Rightarrow 64 λ \Rightarrow 160Gbps = 1000 OC-3 \Rightarrow 10^5 VCs
 \Rightarrow 10^8 Flows
- ❑ Restoration at multiple layers:
DWDM \Rightarrow SONET \Rightarrow ATM \Rightarrow IP
- ❑ SONET \Rightarrow 50% lost = Inefficient Protection
- ❑ SONET \Rightarrow Manual (jumpers) \Rightarrow Slow provisioning
Need Bandwidth on all rings \Rightarrow months/connection
Bandwidth reserved during setup
- ❑ Any layer can bottleneck
 \Rightarrow Intersection of Features + Union of Problems

IP Directly over DWDM: Why?

- ❑ IP \Rightarrow revenue
DWDM \Rightarrow Cheap bandwidth
IP and DWDM \Rightarrow Winning combination
Avoid the cost of SONET/ATM equipment
- ❑ IP routers at OC-192 (10 Gb/s)
 \Rightarrow Don't need SONET multiplexing
- ❑ Coordinated restoration at optical/IP level
- ❑ Coordinated path determination at optical/IP level
- ❑ SONET Framing can remain for error monitoring
Two parts of a layer: Framing + Protocols

IP over DWDM Node

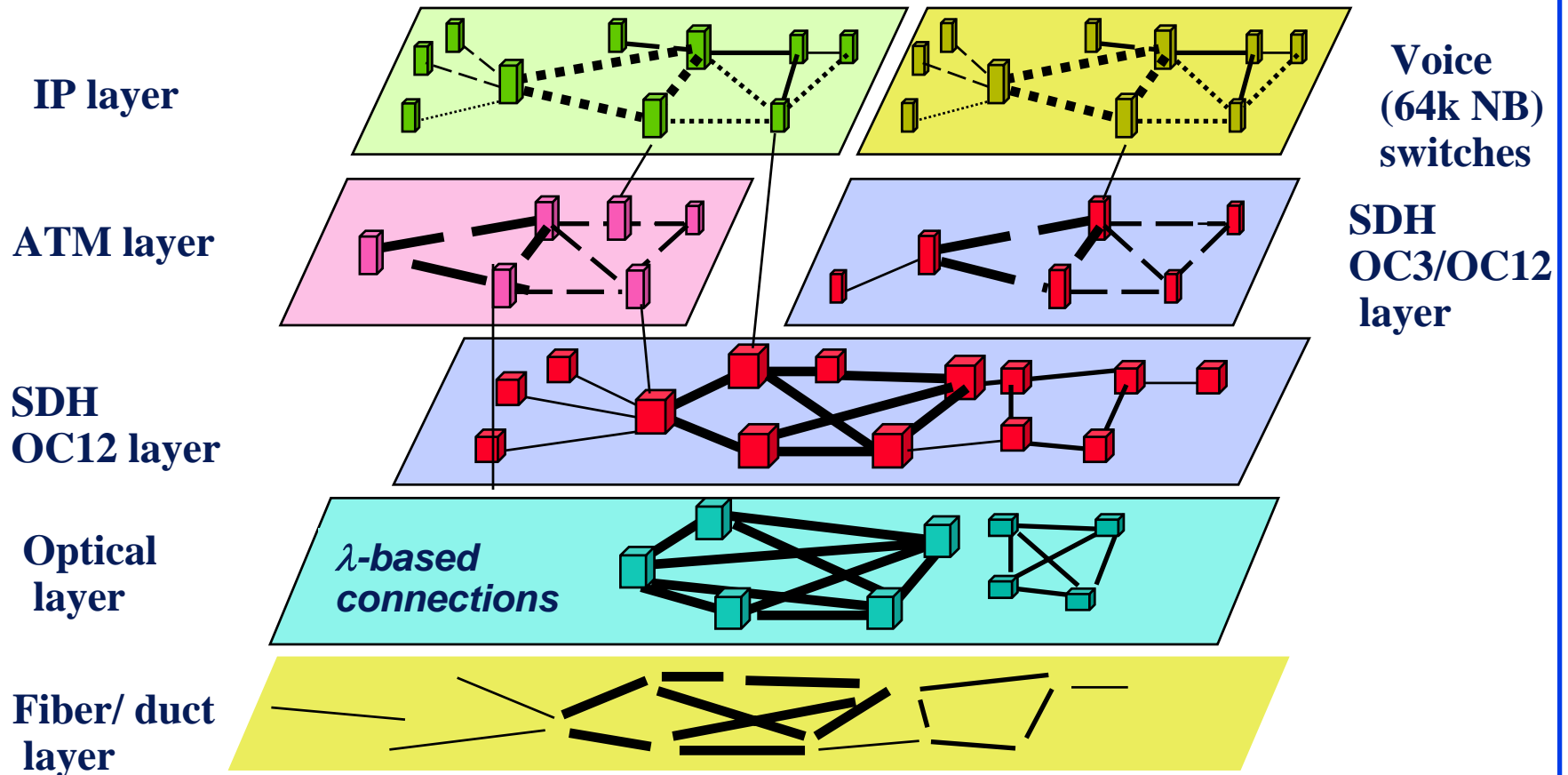


- ❑ Each optical node will be an IP addressable device
- ❑ Will implement OSPF/RIP/BGP, Protection, Wavelength Switching, QoS

IP over DWDM: Issues

- ❑ Routing Wavelength Assignment Algorithms
- ❑ Cheaper High-Speed Routers
- ❑ Topology design Algorithms
- ❑ Wavelength conversion devices
- ❑ Packet Switching Architecture
- ❑ Protection schemes
- ❑ Inverse multiplexing for higher speed pipes
- ❑ QoS
- ❑ Multicast

Virtual Topology Issue

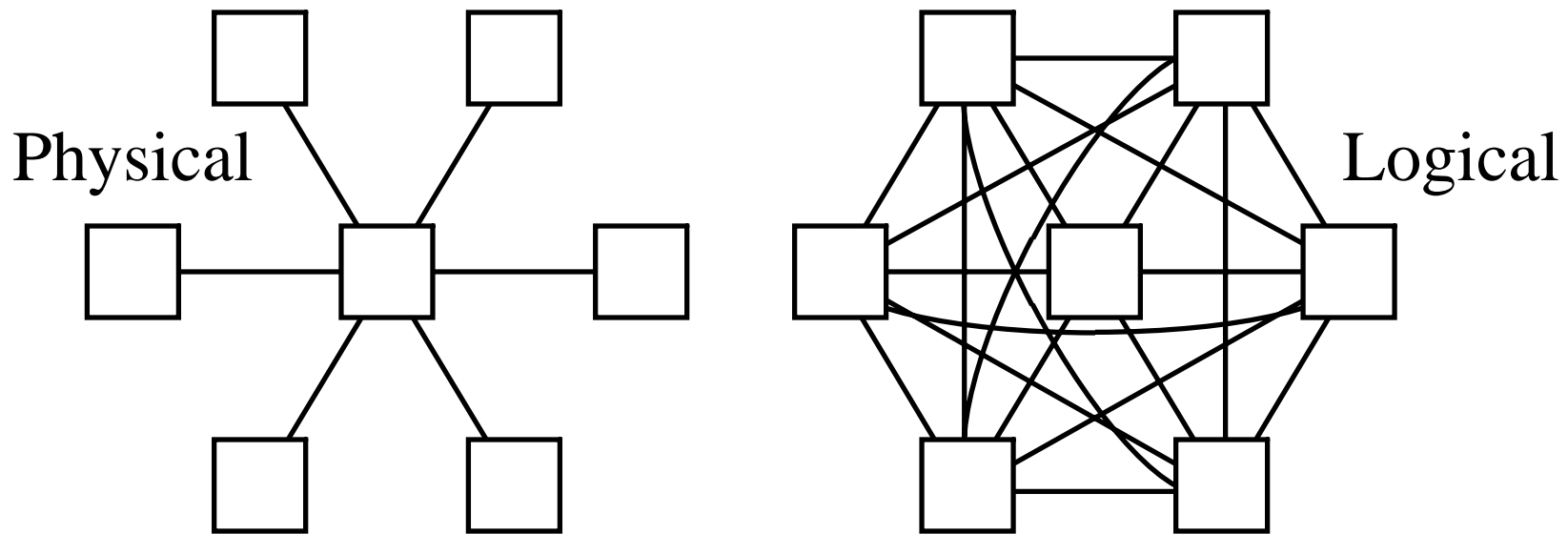


Ref: Dixit

The Ohio State University

Raj Jain

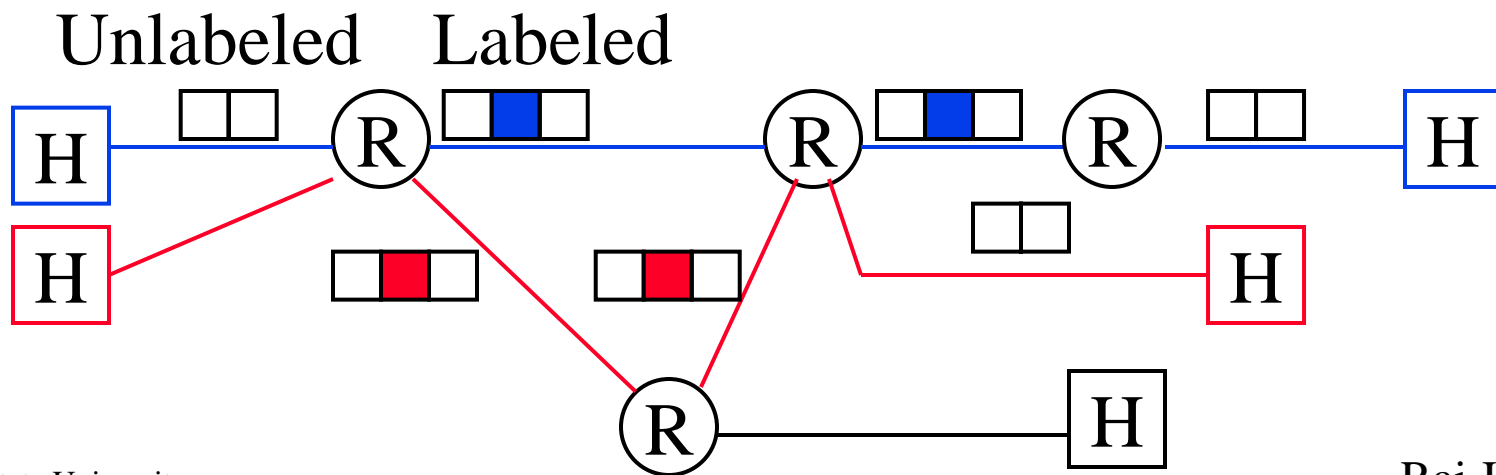
IP over ATM: Lessons



- ❑ Duplication between PNNI and OSPF
- ❑ Virtual topology $\Rightarrow n^2$ scaling problem
- ❑ Solutions:
 - IP Switching \Rightarrow Make every switch a router
 - MPLS \Rightarrow Make every switch an LSR

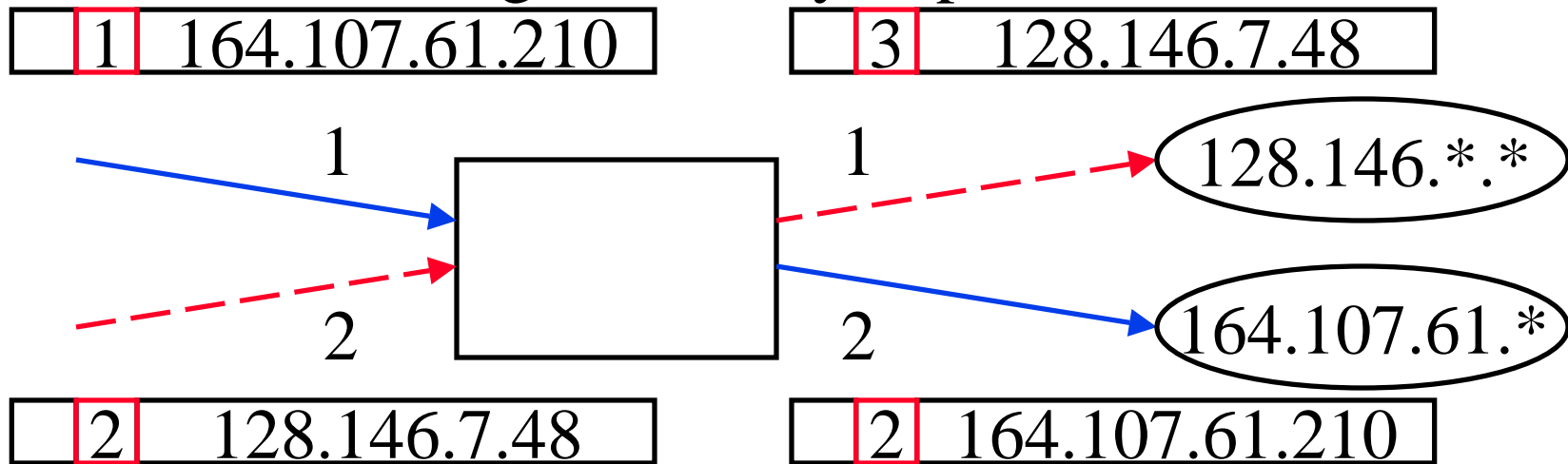
Label Switching

- ❑ Label = Circuit number = VC Id
- ❑ Ingress router/host puts a label. Exit router strips it off.
- ❑ Switches switch packets based on labels. Do not need to look inside \Rightarrow Fast.



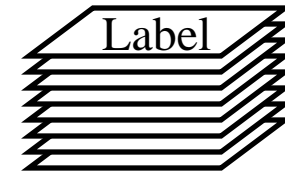
Label Switching (Cont)

- Labels have local significance
- Labels are changed at every hop

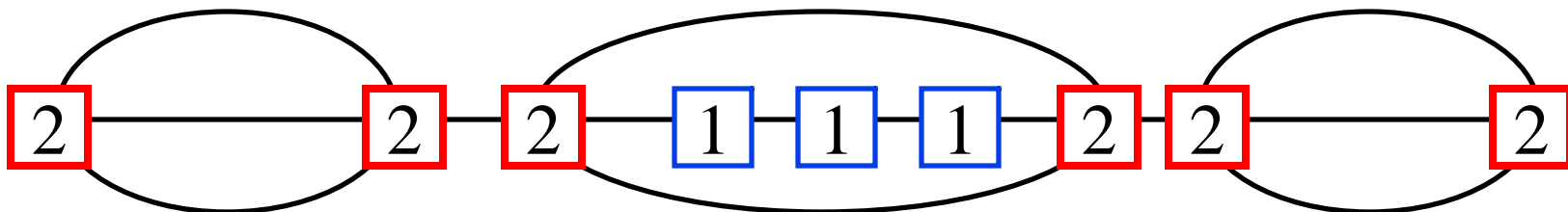
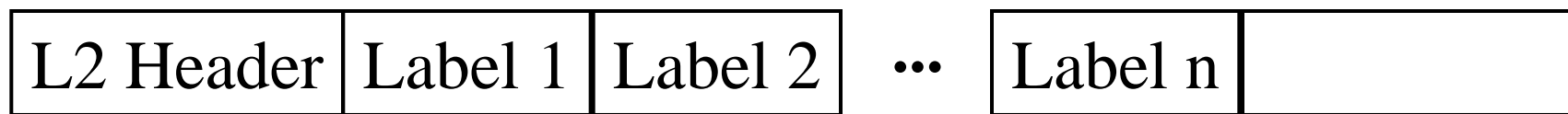


Input Port	Input Label	Adr Prefix	Output Port	Output Label
1	1	164.107.61.*	2	2
2	2	128.146.*.*	1	3

Label Stacks

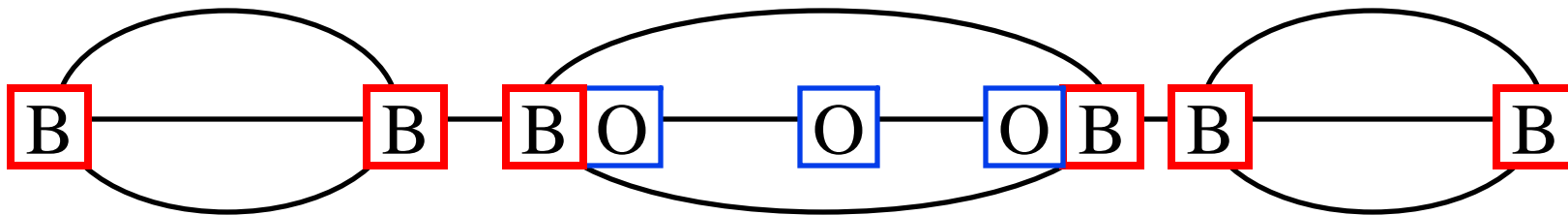


- ❑ A MPLS packet may have multiple labels
- ❑ Labels are pushed/popped as they enter/leave MPLS domain
- ❑ Stack allows hierarchy of MPLS domains
- ❑ Bottom label may indicate protocol (0=IPv4, 2=IPv6)

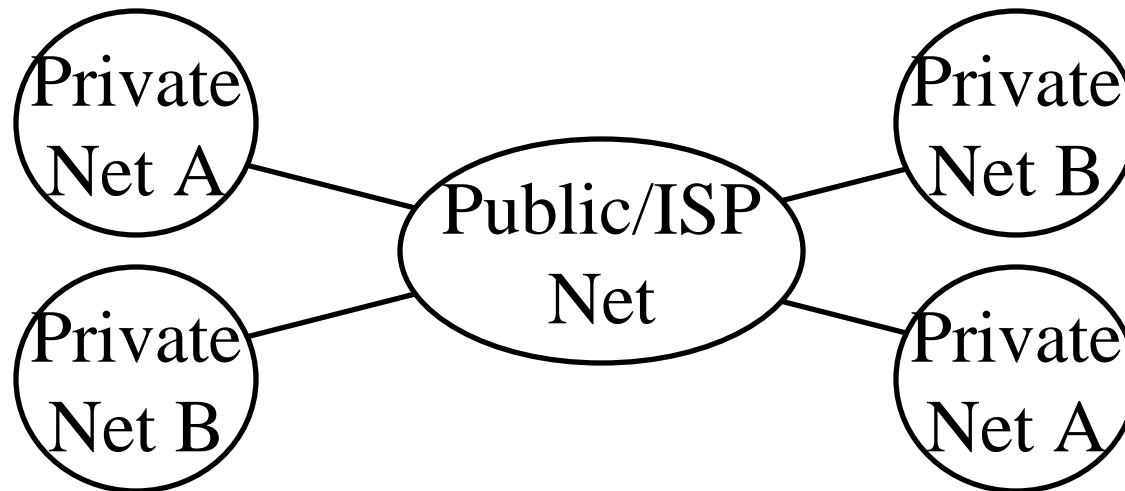


Label Stack Examples

1. BGP/OSPF Routing Hierarchy



2. VPN: Top label used in public network.
Net A and B can use the same private addresses.



Advantages of MPLS

- ❑ MPLS takes the best of both IP and ATM networks
- ❑ Works on both ATM and non-ATM networks
- ❑ Common routing and label distribution on all media
⇒ Easier management
- ❑ No routing over large cloud issue

IP over MPLS over DWDM

- ❑ MPLS = Multi-Protocol Lambda Switching
- ❑ DWDM network \approx ATM network with Limitations
- ❑ Optical Channel Trail = VC = LSPs = Traffic Trunk
- ❑ Fiber = Link
- ❑ Limited # of channels
- ❑ Global significance, if no λ conversion
- ❑ Local significance with λ conversion (still complex)
- ❑ Granularity = $\lambda \Rightarrow$ Fixed datarate
- ❑ No aggregation yet \Rightarrow No label merging

MPLS over DWDM (Cont)

- ❑ No hierarchy yet \Rightarrow No label stacks
- ❑ No TDM yet \Rightarrow No cells or packets
- ❑ No queueing \Rightarrow No scheduling, No Priority, No burstiness, No policing
- ❑ Need Shaping/grooming at entry
- ❑ Faster restoration via redundancy (rings/mesh)
- ❑ Vendor specific management
 \Rightarrow Interoperability issues

MPLS Control Plane: Today

- ❑ Resource Discovery: IGP (OSPF/PNNI)
- ❑ Path Computation: IGP (OSPF/PNNI)
- ❑ Connection Management: Label Distribution via IGP(OSPF), LDP, RSVP
- ❑ Survivability: Rerouting,...
- ❑ Constraint-based routing based on data rate, overbooking, delay, ...

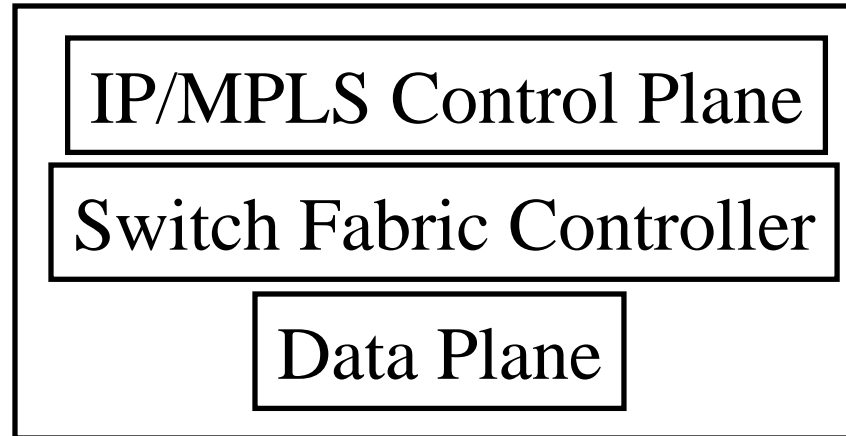
MPLS Control Plane: Tomorrow

- ❑ Next Hop Forwarding Label Entry (NHFLE)
 - = Preprogrammed λ switching
 - = Wavelength Forwarding Information Base matrix
 - \Rightarrow \langle Input port, λ \rangle to \langle output port, λ \rangle mapping
- ❑ Constraints: Data rate, Attenuation, Dispersion, Length, delay
- ❑ Topologies: Linear and rings to partial Mesh
- ❑ λ control plane via network management
 - \Rightarrow Permanent \Rightarrow Static routing
 - \Rightarrow Too slow for restoration

MPLS Control Tomorrow (Cont)

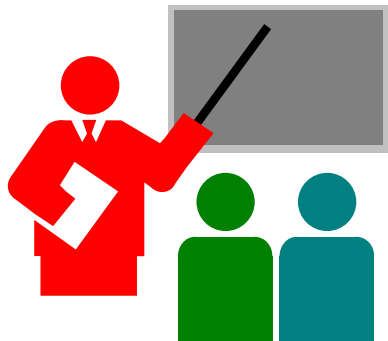
- ❑ Can add resilience (survivability) preemption, resource class affinity attributes to trails
- ❑ Each OXC will be an IP addressable device
- ❑ Control plane can be out-of-band IP channel, dedicated supervisory channel
- ❑ Need to build on concept of "Abstract Node" in IP routing \Rightarrow Failures are handled locally
- ❑ λ availability will be advertised by optical node/WRouter

Optical Node Architecture



- ❑ Pre-configured control wavelength upon initialization
- ❑ Need to develop hierarchical/aggregation concepts (label stacks or VPs)
 - ⇒ λ -Group (Optical channel, optical path, Light path)
- ❑ Add light path constraints to MPLS label distribution or explicit path requests
- ❑ Ref: draft-awduche-mpls-te-optical-00.txt

Summary



- ❑ High IP Routing speeds and volumes
 - ⇒ Need a full wavelength
 - ⇒ Many ATM/SONET functions not needed
- ❑ Need MPLS to provide QoS, Isolation
- ❑ Protection/Restoration/Routing should be coordinated between IP/MPLS and DWDM
- ❑ Need to develop hierarchy/aggregation concepts for DWDM

References:

- ❑ See references in http://www.cse.ohio-state.edu/~jain/refs/opt_refs.htm
- ❑ Recommended books on optical networking, http://www.cse.ohio-state.edu/~jain/refs/opt_book.htm
- ❑ Optical networking and DWDM, <http://www.cse.ohio-state.edu/~jain/cis788-99/dwdm/index.html>
- ❑ IP over DWDM, http://www.cse.ohio-state.edu/~jain/cis788-99/ip_dwdm/index.html
- ❑ Newsgroup: sci.optics.fiber

Acronyms

- ❑ ATM Asynchronous Transfer Mode
- ❑ BGP Border Gateway Protocol
- ❑ DWDM Digital Wavelength Division Multiplexing
- ❑ GHz Giga Hertz
- ❑ IGP Interior Gateway Protocol
- ❑ IP Internet Protocol
- ❑ IPv4 IP Version 4
- ❑ IPv6 IP Version 6
- ❑ MIP Millions of Instructions per second
- ❑ MPLS Multiprotocol Label Switching
- ❑ NHFLE Next Hop Forwarding Label Entry

Acronyms (Cont)

- ❑ OC Optical Carrier
- ❑ OSPF Open Shortest Path First
- ❑ OXC Optical cross connect
- ❑ PC Personal Computers
- ❑ PNNI Private Network to Node Interface
- ❑ PPP Point-to-point protocol
- ❑ SONET Synchronous Optical Network
- ❑ TDM Time Division Multiplexing
- ❑ VC Virtual Circuit
- ❑ VPs Virtual Paths
- ❑ WRouter Wavelength Router