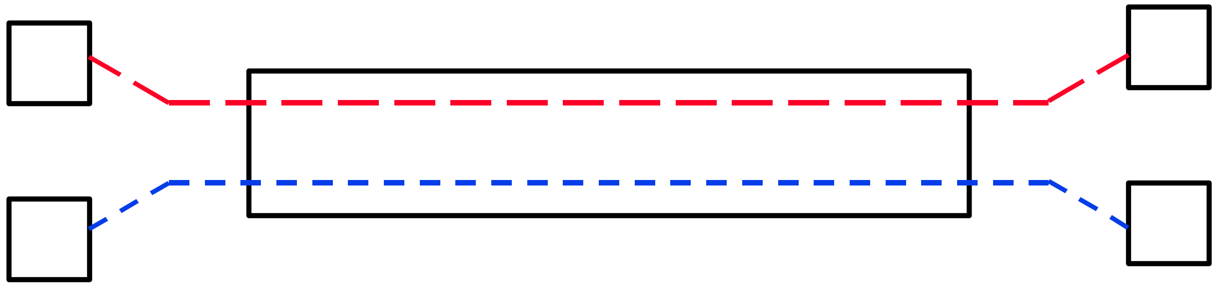
IP over DWDM Networks

Raj Jain
The Ohio State University
Columbus, OH 43210
Jain@CIS.Ohio-State.Edu

These slides are available at

http://www.cis.ohio-state.edu/~jain/talks/opt_gte.htm

Sparse and Dense WDM



- 10Mbps Ethernet (10Base-F) uses 850 nm
- 100 Mbps Ethernet (100Base-FX) + FDDI use 1310 nm
- Some telecommunication lines use 1550 nm
- WDM: 850nm + 1310nm or 1310nm + 1550nm
- Dense ⇒ Closely spaced ≈ 1nm separation

Recent WDM Records

- 1×40 G up to 65 km (Alcatel'98). PMD Limited.
- 32× 5 G to 9300 km (1998)
- 64× 5 G to 7200 km (Lucent'97)
- 100×10 G to 400 km (Lucent'97)
- 16×10 G to 6000 km (1998)
- 132×20 G to 120 km (NEC'96)
- 70×20 G to 600 km (NTT'97)
- 1022 Wavelengths on one fiber (Lucent 99)
- Ref: OFC'9x

WDM Applications

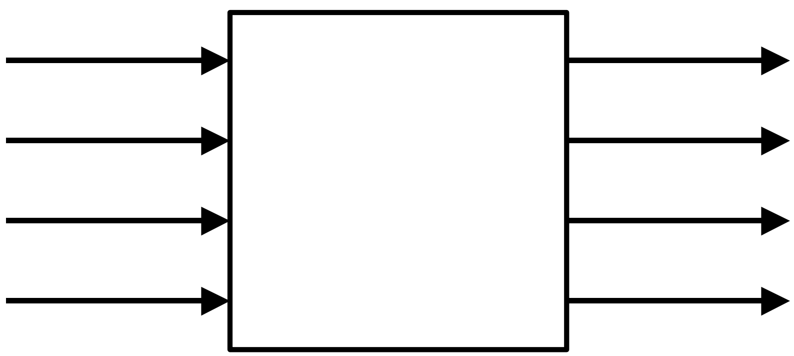
- WANs: Fiber links ⇒ WDM ⇒ DWDM Links
- Undersea Links: Amplifiers ⇒ High maintenance cost ⇒ Can't put too many fibers
- DWDM highly successful in long-haul market.
- Not yet cost-competitive in metro market.
  Bandwidth demand is low and more dynamic.
  Many new lower cost products for metro market.

Sample Products

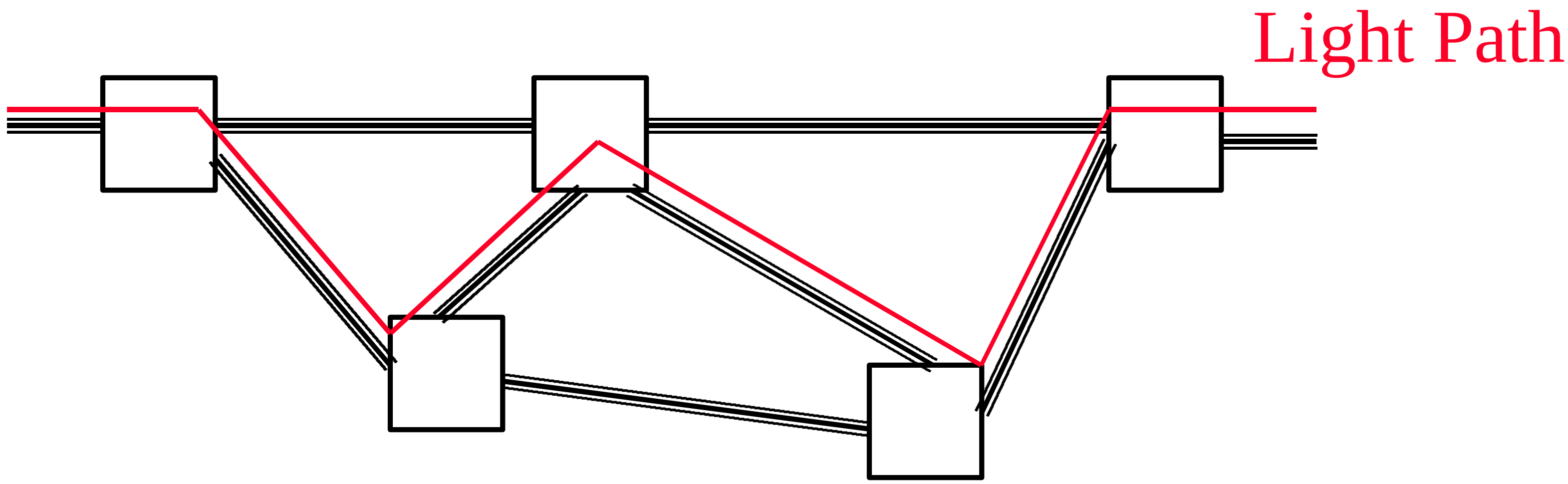
- **Nortel/Cambrian**: Optera Metro: 32 × 2.5G Optera LH: 2560×622Mbps, 1280×1.25Gbps (Gb Ethernet), 640×2.5Gbps, 160×10Gbps
- **Pirelli Optical Systems**: 128×10G TeraMuX WaveMux H-DWDM with Soliton
  OMDS 32λ WDM System
- **Monterey Networks**: Wavelength Router™
  256×256 OC-48 scalable to 160 Tbps
  Non-blocking any to any.
  Fully hot swappable w/o fiber swap
  1+1 or 1:N APS. Straight IP over DWDM.

Optical (Wavelength) Cross Connect

- Slow switching nodes.
- Configuration changed by management.
- May allow any wavelength on any fiber to go to any fiber.
- Programmable.
- Control channel could be electronic or optical.

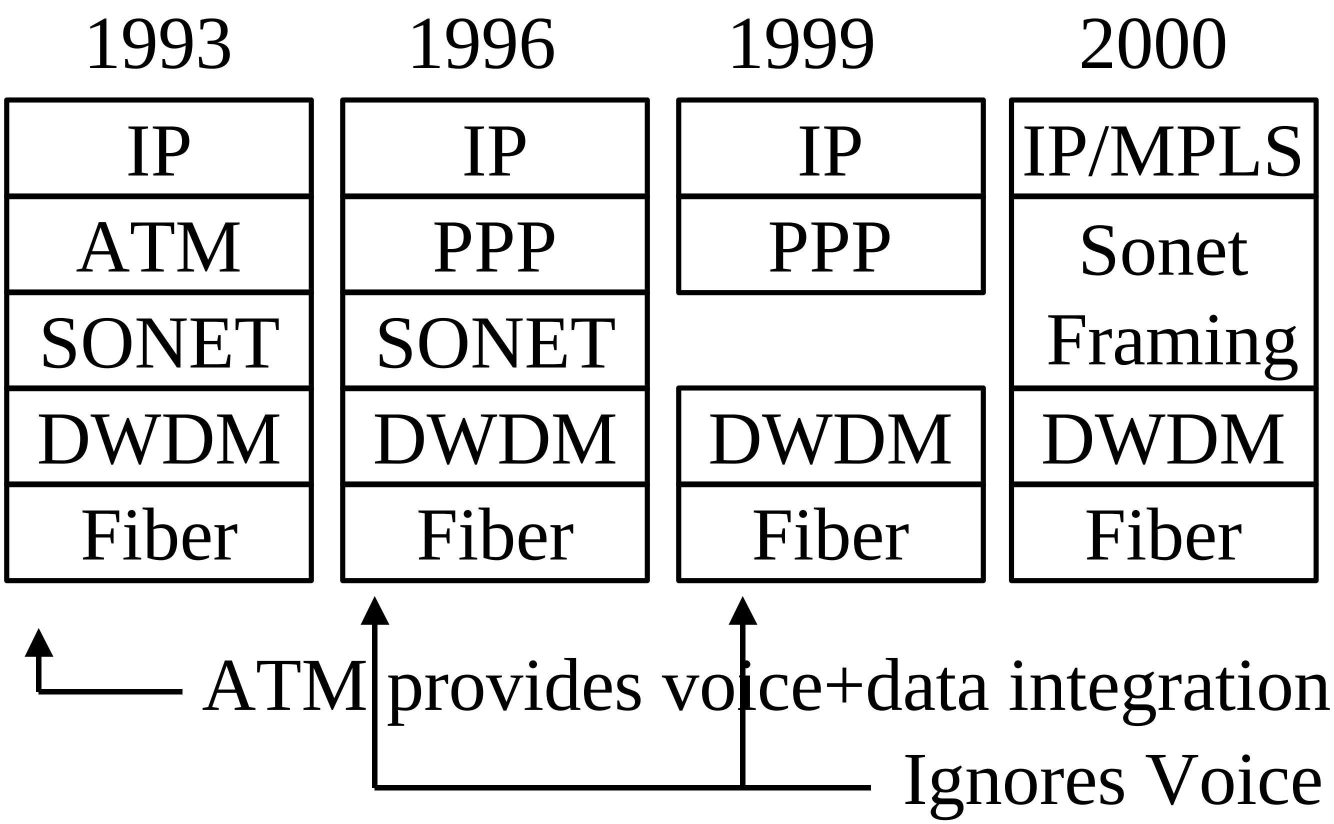


Wavelength Routed Networks



- Light path through a DWDM network
- Routing ⇒ Wavelength assignment problem
- Two wavelengths from different fibers should not be mixed ⇒ Need wavelength conversion

Stack Debate



PPP = Point to point protocol in HDLC-like framing

Multi-Layer Stack: Why?

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

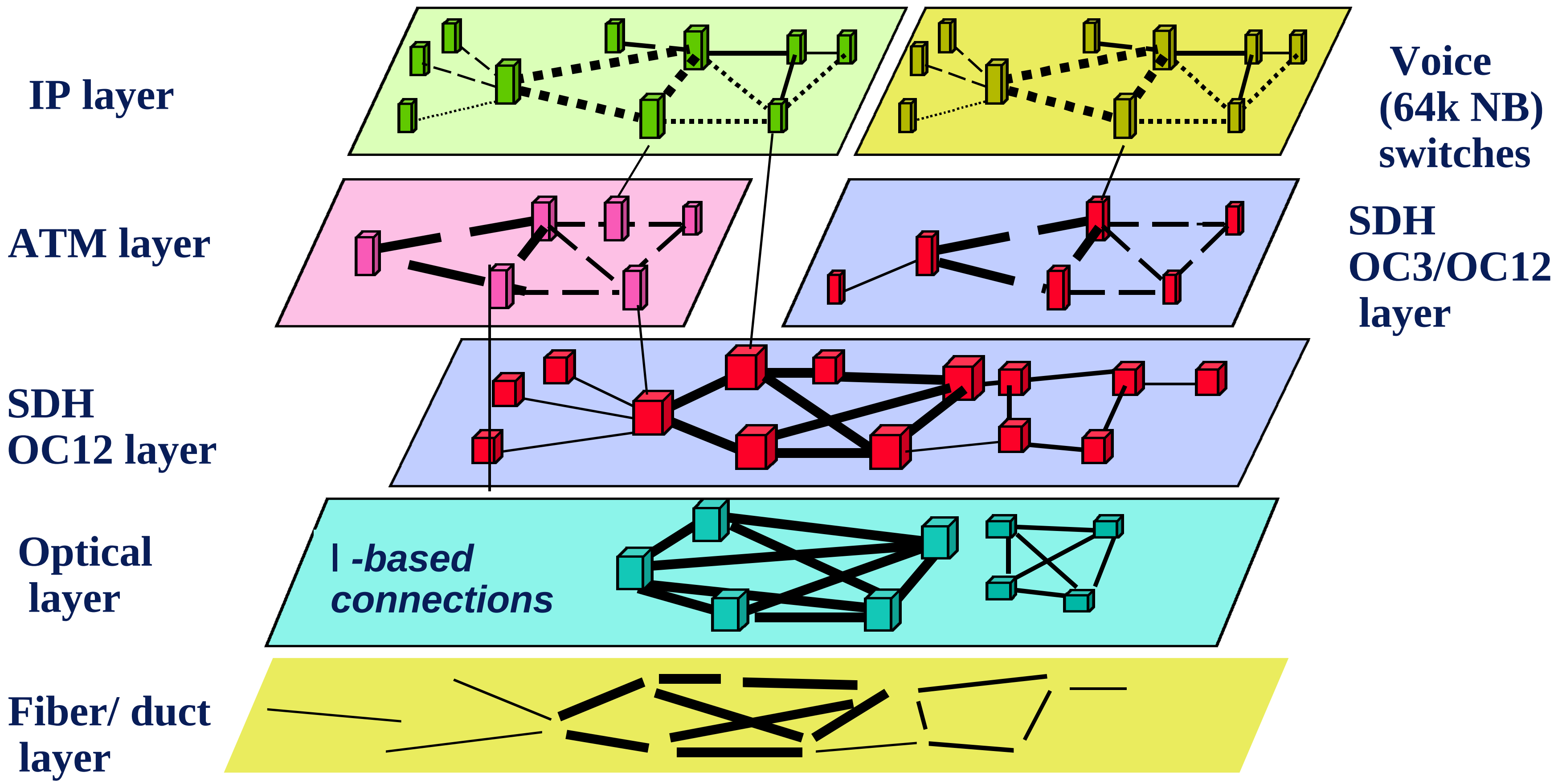
Multilayer Stack Problems (Cont)

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

IP Directly over DWDM: Why?

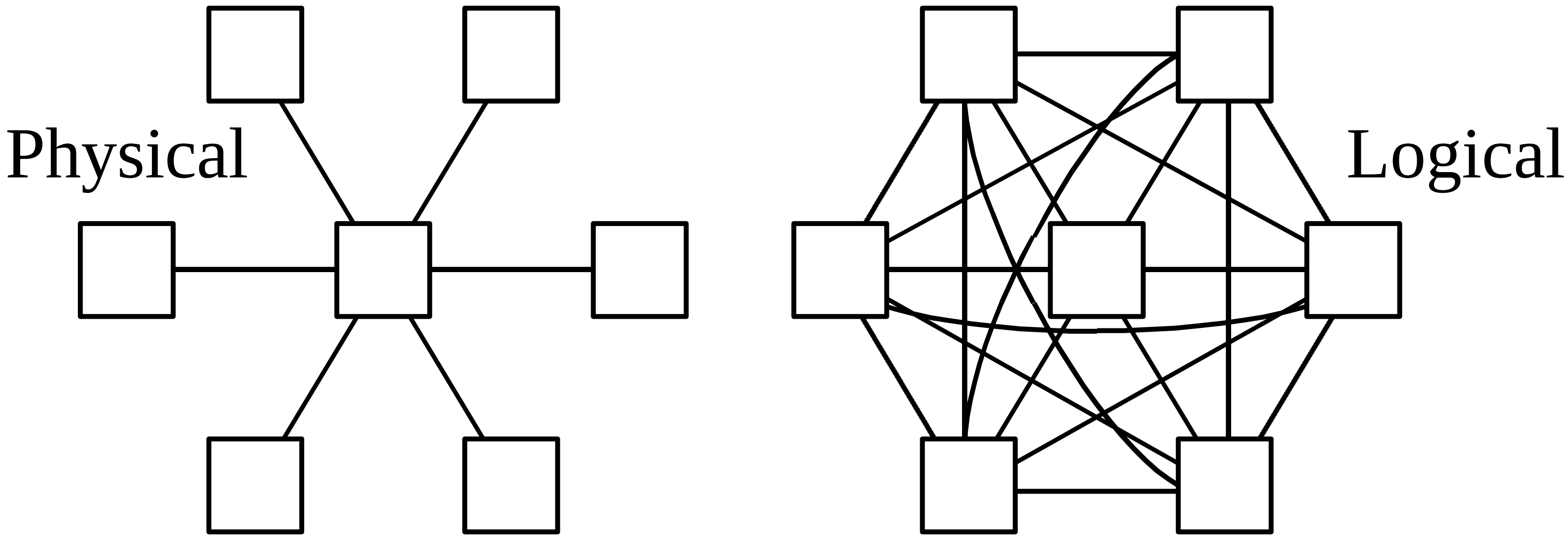
- IP $\Rightarrow$ revenue
  DWDM $\Rightarrow$ Cheap bandwidth
  IP and DWDM $\Rightarrow$ Winning combination
  Avoid the cost of SONET/ATM equipment
- IP routers at OC-192 (10 Gb/s)
  $\Rightarrow$ Don't need SONET multiplexing
- Coordinated restoration at optical/IP level
- Coordinated path determination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

Virtual Topology Issue



Ref: Dixit

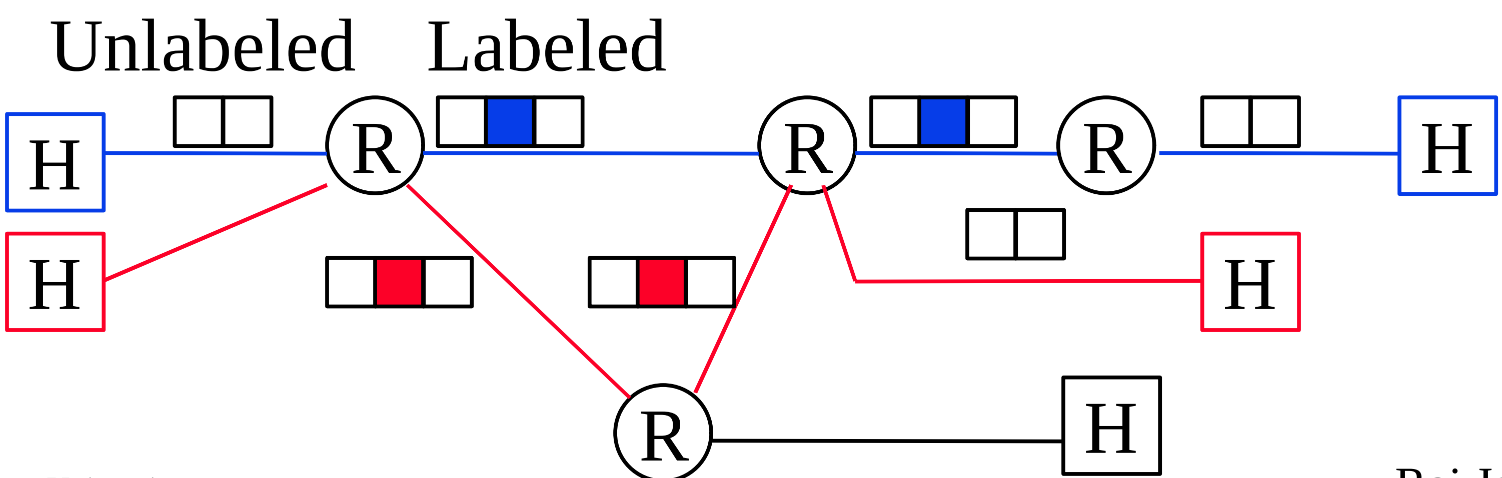
IP over ATM: Lessons



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

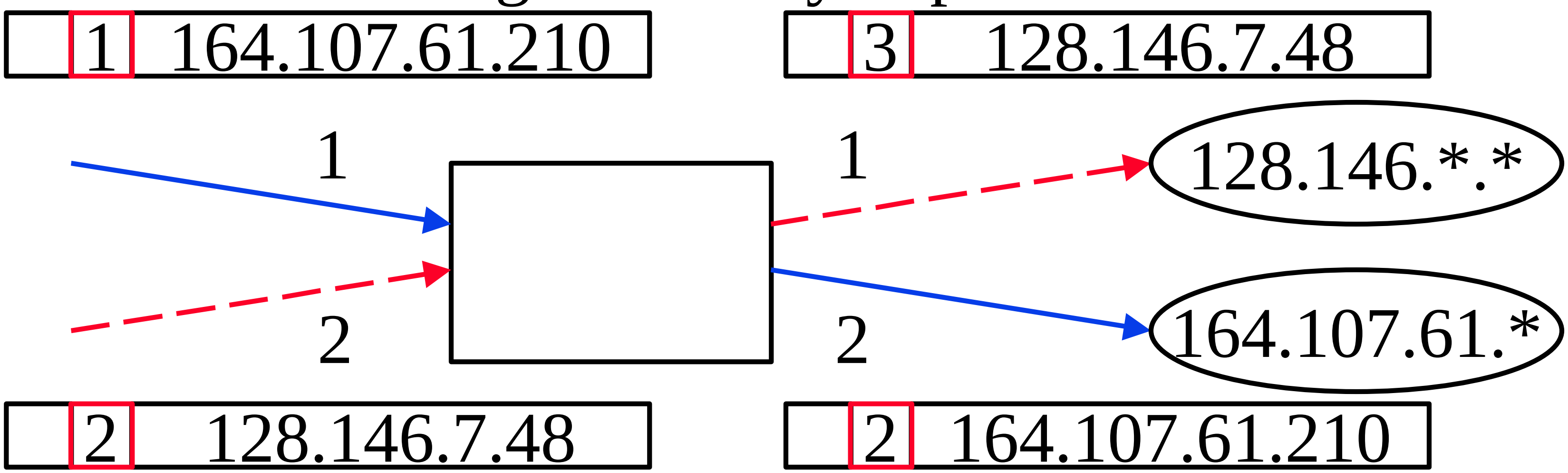
Label Switching

- 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 ⇒ 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

IP over MPLS over DWDM

- MPLS = Multi-Protocol Lambda Switching
- DWDM network ≈ 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 ⇒ No label stacks
- No TDM yet ⇒ No cells or packets
- No queueing ⇒ No scheduling, No Priority, No burstiness, No policing
- Need Shaping/grooming at entry
- Faster restoration via redundancy (rings/mesh)
- Vendor specific management
  ⇒ 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, ...

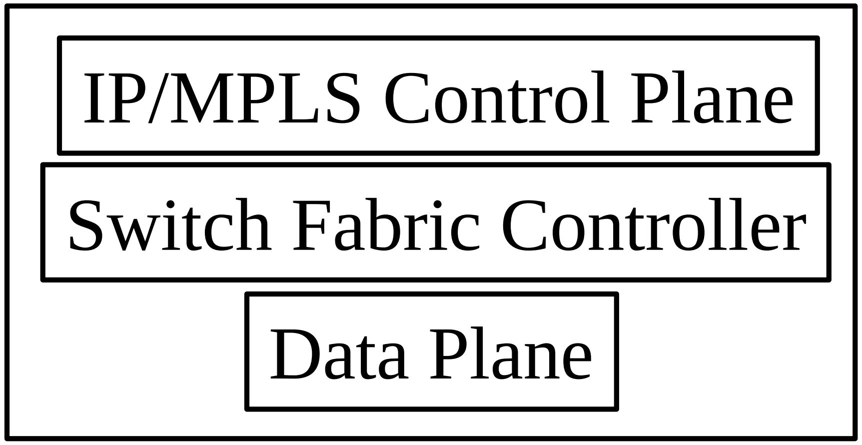
MPLS Control Plane: Tomorrow

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

MPLS Control Tomorrow (Cont)

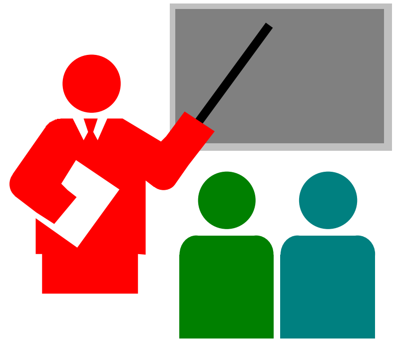
- Can add resilience (survivability) preemption, resource class affinity attributes to trails
- Each OXC will be an IP addressable device
- Control plane can be out-of-band IP channel, dedicated supervisory channel
- Need to build on concept of "Abstract Node" in IP routing $\Rightarrow$ Failures are 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)
  $\Rightarrow$ 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



- DWDM allows 32- to 128- channels per fiber
- High IP Routing speeds and volumes
  $\Rightarrow$ Need a full wavelength
  $\Rightarrow$ Many ATM/SONET functions not needed
- Need MPLS to provide QoS, Isolation
- Protection/Restoration/Routing should be coordinated between IP/MPLS and DWDM
- Need hierarchy/aggregation concepts for DWDM

References:

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

Standards Organization

- ITU:
  - **G.681** Functional characteristics of interoffice and long-haul line systems using optical amplifiers including optical multiplexing
  - **G.692** Optical Interfaces for multichannel systems with optical amplifiers (Oct 98): 50 and 100 GHz spacing centered at 193.1 THz (1553.5 nm)
  - **G.872** Architecture for Optical Transport Networks, 1999
  - Several others in preparation

Standards (Cont)

- ANSI T1X1.5: http://www.t1.org/t1x1/_x1-grid.htm
- IETF: MPLS over DWDM
- Optical Interoperability Forum (OIF): www.oiforum.com
  - Started April 1998 by CISCO, Ciena, ...Now over 128 members
  - Working groups on Architecture, Physical and Link Layer, OAM&P
  - Signaling protocols for rapid provisioning and restoration

Acronyms

- ADM Add-Drop Multiplexer
- PANDA Polarization maintaining AND Absorption reducing
- ANSI American National Standards Institute
- APS Automatic Protection Switching
- ATM Asynchronous Transfer Mode
- CDMA Code Division Multiple Access
- DARPA Defense Advanced Research Project Agency
- DCF Dispersion Compensating Fiber
- DPT Dynamic Packet Transport
- DSF Dispersion Shifted Fiber

- DFF Dispersion Flattened Fiber
- DSL Digital Subscriber Lines
- DWDM Digital Wavelength Division Multiplexing
- EDFAs Erbium-Doped Fiber Amplifiers
- FCC Federal Communications Commission
- FWM Four-Wave Mixing
- GHz Giga Hertz
- IEEE Institution of Electrical and Electronic Engineers
- IETF Internet Engineering Taskforce
- IPS Intelligent Protection Switching

ITU	International Telecommunications Union
KEOPS	Keys to Optical Packet Switching
LAN	Local Area Network
LED	Light Emitting Diode
MMF	Multimode Fiber
NRZ	Non-return to zero
NTONC	National Transparent Optical Network Consortium
OAM	Operation Administration and Maintenance
OC	Optical Carrier
OCh	Optical Channel Layer

- OFC Optical Fiber Conference
- OIF Optical Interoperability Forum
- OMS Optical Multiplex Section
- OPP Optical Packet Path
- SPP Secondary Packet Paths
- OSC Optical Supervisory Channel
- OSN Optical Service Networks
- OSPF Open Shortest Path First
- OTDM Optical Time Domain Multiplexing
- OTS Optical Tranmission Section
- OXC Opical cross connect
- PMD Polarization Mode Dispersion

PMF	Polarization Maintening Fiber
PMMA	PolyMethylMethyelAcrylate
RI	Refrective Index
RIP	Routing Information Protocol
SNMP	Simple Network Management Protocol
SNR	Signal to Noise Ratio
SONET	Synchronous Optical Network
SRP	Spatial Reuse Protoco
TDM	Time Division Multiplexing
WAN	Wide Area Network
WC	Wavelength converter

- WDM Wavelength Division Multiplexing
- WGR Wavelength Grafting Router
- WIXC Wavelength Interchanging Crossconnect
- WSXC Wavelength Selective Crossconnect
- ZBLAN Zirconium, barium, lanthanum, aluminium, and sodium