Optical Networking: Recent Developments, Issues, and Trends

Raj Jain
Co-founder and Sr. VP
Nayna Networks, Inc.
481 Sycamore Dr, Milpitas, CA 95035
Email: jain@acm.org
www.nayna.com and http://www.cis.ohio-state.edu/~jain/

Overview

1. Networking trends
2. Optical Transmission and Switching
3. Carrier Networking Technologies: SONET, SDH, OTN
4. Gigabit and 10 G Ethernet, Next Gen SONET
5. Passive Optical Networks
6. IP over DWDM: MPλS, GMPLS, UNI

1. Networking Trends

- Life Cycles of Technologies
- Traffic vs Capacity Growth
- Technology Failures vs Successes
- Trend: LAN - WAN Convergence
- Trend: Everything over IP

2. Optical Transmission and Switching

Physical Layer

- Recent DWDM Records
- Amplifiers
- Transmission Products
- OEO vs OOO Switches
- Higher Speed: 40 Gbps
- More Wavelengths per fiber
- Ultra-Long Haul Transmission
- Free space Optical Communications

3. SONET, SDH, OTN

- SONET: Components, Frame Format
- Multiplexing and Concatenation
- Protection: Rings
- Synchronous Digital Hierarchy (SDH)
- Optical Transport Network (OTN)

4. GbE, 10GbE, Next Gen SONET

- Distance-B/W Principle
- GbE: Key features, PMD types
- 10 GbE: Key Features, PMD Types
- Resilient Packet Rings
- Beyond 10 GbE
- Next Gen SONET:
  - Virtual Concatenation
  - Generic Framing Protocol (GFP)
  - Link Capacity Adjustment Scheme (LCAS)

5. Passive Optical Networks

- Why PONs?
- PON Operation
- PON Design Issues
- APON, BPON, GPON, EPON
- PON Applications

6. IP over DWDM

- Why IP over DWDM?
- How to IP over DWDM?
  - What changes are required in IP?
  - MPλS and GMPLS
  - UNI, LDP, RSVP, LMP
- Upcoming Optical Technologies

Schedule (Tentative)

9:00 - 9:30	Course Introduction
9:30 -10:15	Networking Trends
10:15 -10:30	*Coffee Break*
10:30 -11:15	Transmission and Switching
11:15 - 12:00	SONET, SDH, OTN I
12:00 - 1:00	*Lunch Break*
1:00 - 1:15	SONET, SDH, OTN II
1:15 - 2:00	GbE, 10 GbE, Next Gen SONET I
2:00 - 2:15	*Break*
2:15 - 2:45	GbE, 10 GbE, Next Gen SONET II
2:45 - 3:30	Passive Optical Networks
3:30 - 3:45	*Break*
3:45 - 5:00	IP over DWDM

References

- You can get to all on-line references via: http://www.cis.ohio-state.edu/~jain/refs/opt_refs.htm

Pre-Test

Check if you know the difference between:

- SONET and Ethernet Frame Format
- BLSR and UPSR
- SDH and OTN
- GFP and LCAS
- Token Ring and Resilient Packet Ring
- Ring wrapping vs Steering
- MPλS and GMPLS
- 10GBASE-LR and 10GBASE-EW
- 1000BASE-BX10 and 1000BASE-PX10
- GPON and EPON

Number of items checked ______

- If you checked more than 5 items,
you may not gain much from this tutorial.
- If you checked only a few or none, don't worry. This course will cover all this and much more.

Disclaimer

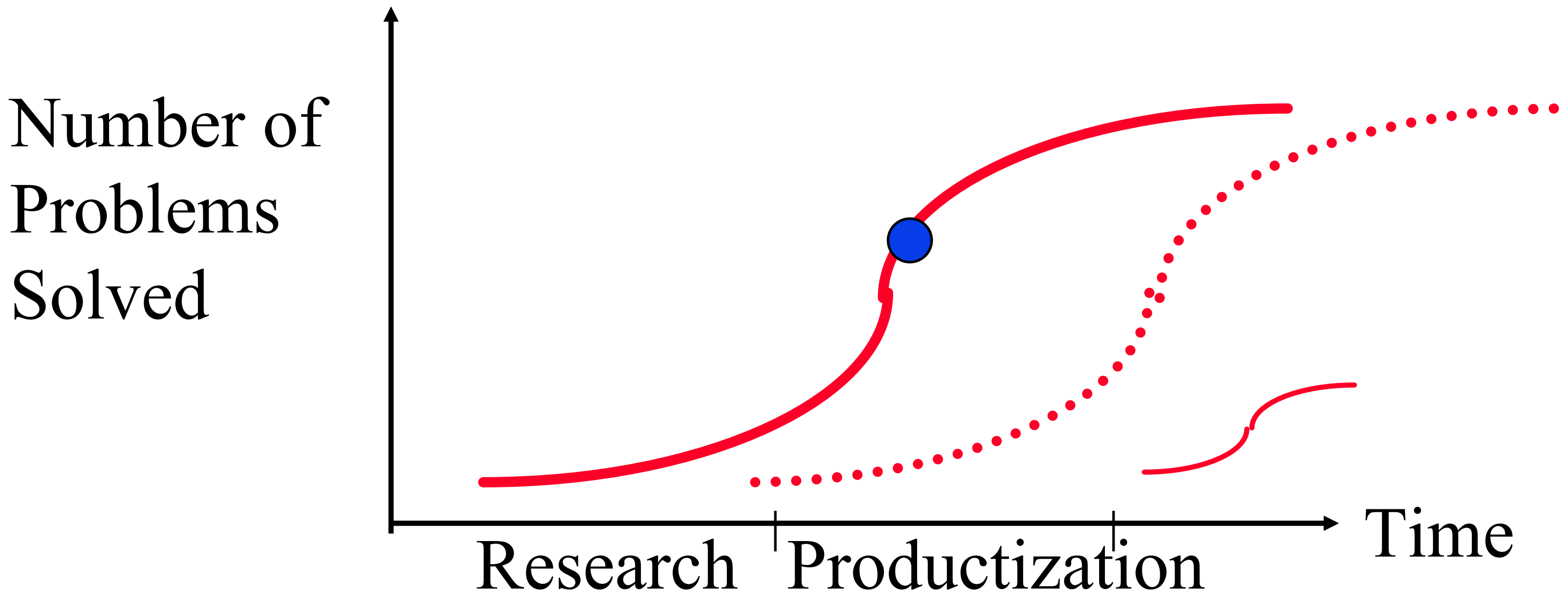
- The technologies are currently evolving.
  $\Rightarrow$ Many statements are subject to change.
- Features not in a technology may be implemented later in that technology.
- Problems claimed to be in a technology may later not be a problem.

Networking Trends and Their Impact

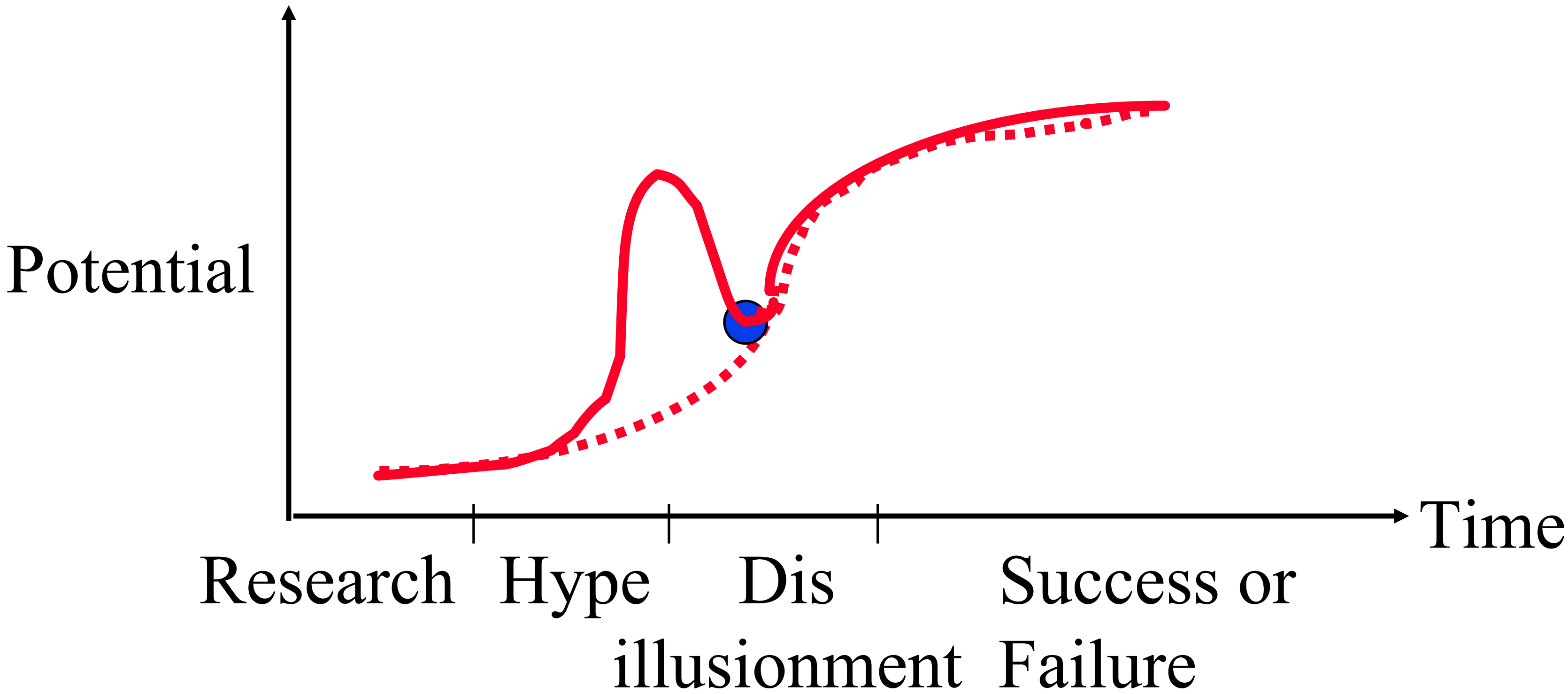
Raj Jain
Co-founder and Sr. VP
Nayna Networks, Inc.
481 Sycamore Dr, Milpitas, CA 95035
Email: jain@acm.org
www.nayna.com and http://www.cis.ohio-state.edu/~jain/

- Life Cycles of Technologies
- Traffic vs Capacity Growth
- Technology Failures vs Successes
- Trend: LAN - WAN Convergence
- Trend: Everything over IP

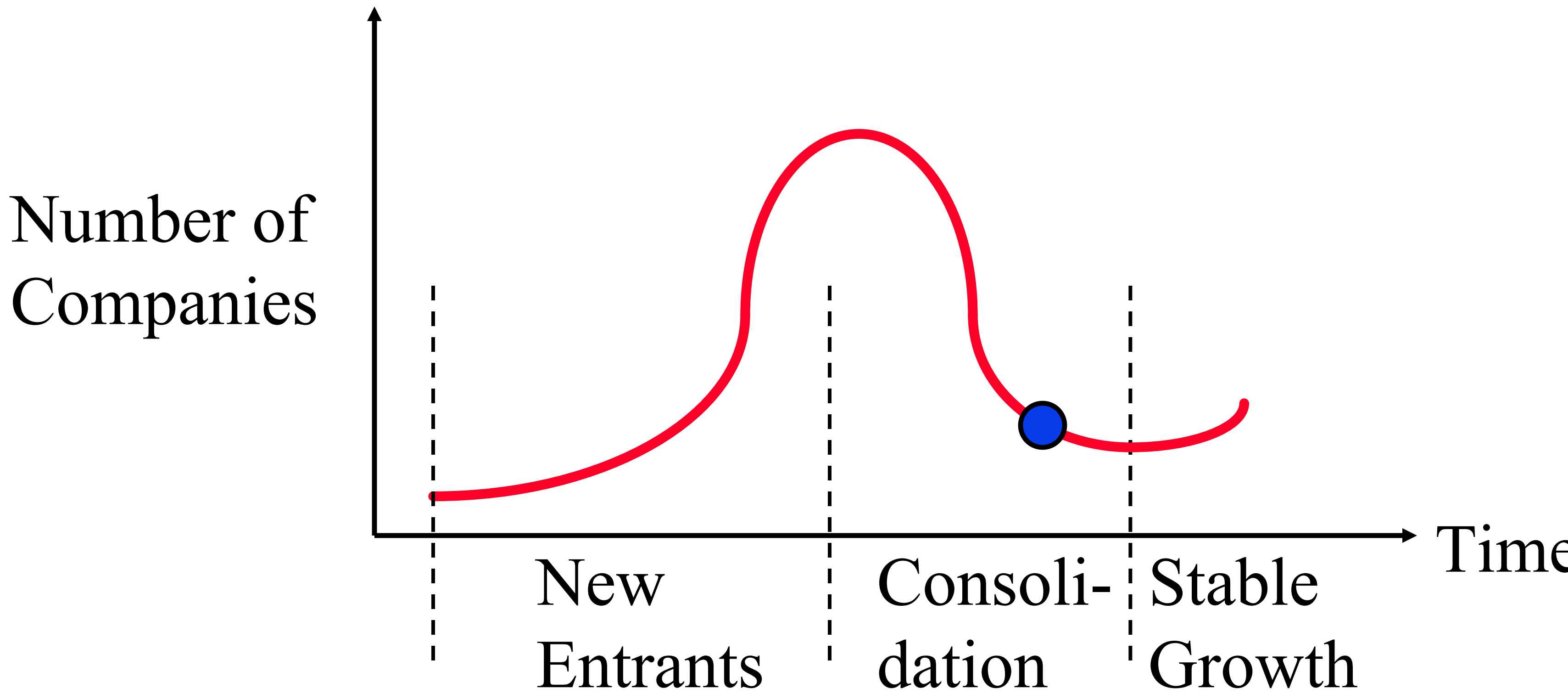
Life Cycles of Technologies



Hype Cycles of Technologies



Industry Growth



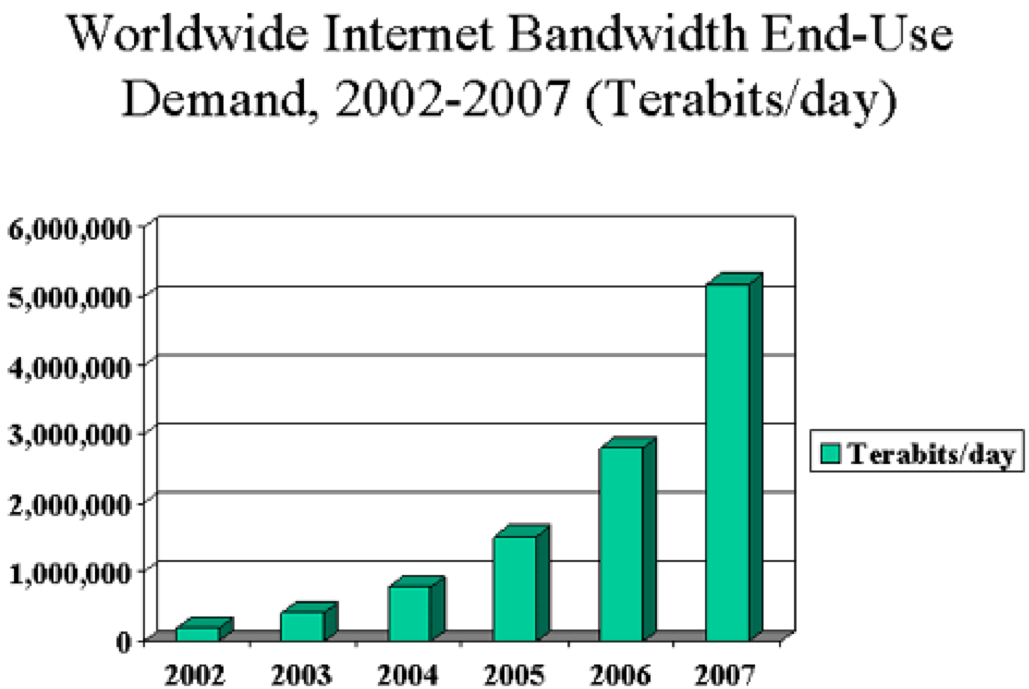
Trend: Back to ILECs

1. CLECs to ILECs
   ILEC: Slow, steady, predictable.
   CLEC: Aggressive, Need to build up fast
   New networks with newest technology
   No legacy issues
2. Back to Voice
   CLECs wanted to *start* with data
   ILECs want to *migrate* to data
   ⇒ Equipment that support voice circuits but allow packet based (hybrids) are more important than those that allow only packet based

Traffic vs Capacity Growth

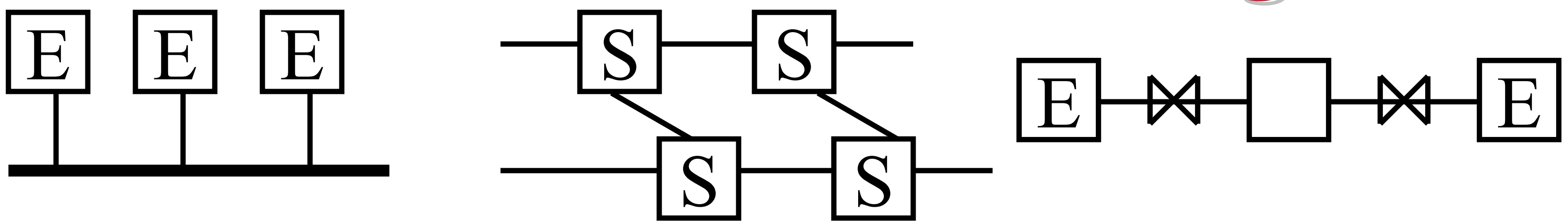
Expensive Bandwidth	Cheap Bandwidth
❑ Sharing	❑ No sharing
❑ Multicast	❑ Unicast
❑ Virtual Private Networks	❑ Private Networks
❑ Need QoS	❑ QoS less of an issue
❑ Likely in WANs, L3	❑ Possible in LANs, L1/L2

Traffic Growth



- Internet traffic to nearly double every year for the next five years
- IDC Report, "Broadband Drives Internet Growth," February 27, 2003.

Trend: LAN - WAN Convergence



- Past: Shared media in LANs. Point to point in WANs.
- Future: No media sharing by multiple stations
  - Point-to-point links in LAN and WAN
  - No distance limitations due to MAC. Only Phy.
  - Datalink protocols limited to frame formats
- 10 GbE over 40 km without repeaters
- Ethernet End-to-end.
- Ethernet carrier access service:$1000/mo 100Mbps

Networking: Failures vs Successes

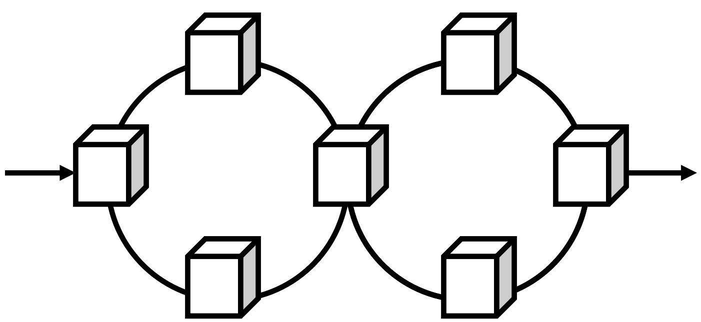
- 1980: Broadband (vs baseband)
- 1984: ISDN (vs Modems)
- 1986: MAP/TOP (vs Ethernet)
- 1988: OSI (vs TCP/IP)
- 1991: DQDB
- 1994: CMIP (vs SNMP)
- 1995: FDDI (vs Ethernet)
- 1996: 100BASE-VG or AnyLan (vs Ethernet)
- 1997: ATM to Desktop (vs Ethernet)
- 1998: Integrated Services (vs MPLS)
- 1999: Token Rings (vs Ethernet)

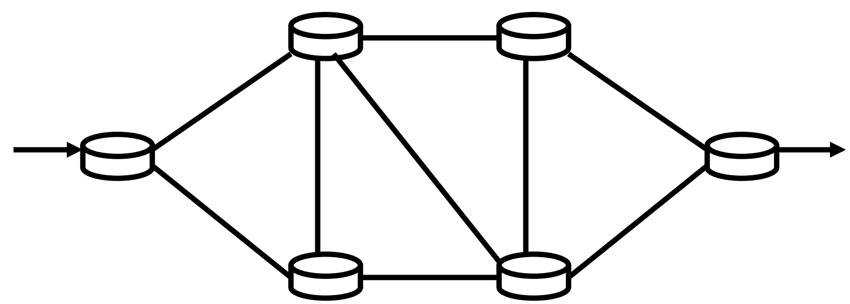
Requirements for Success

- Low Cost: Low startup cost ⇒ Evolution
- High Performance
- Killer Applications
- Timely completion
- Manageability
- Interoperability
- Coexistence with legacy LANs
  Existing infrastructure is more important than new technology

Trend: Everything over IP

	Telecom Networks	Data Networks
Topology Discovery	Manual	Automatic
Path Determination	Manual	Automatic
Circuit Provisioning	Manual	No Circuits
Transport & Control Planes	Separate	Mixed
User and Provider Trust	No	Yes
Protection	Static using Rings	No Protection



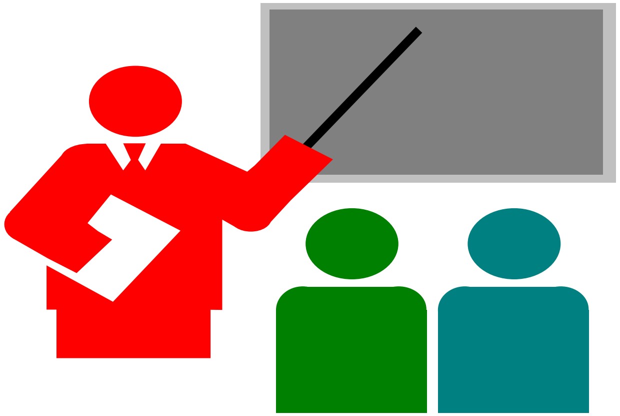


- IP needs circuits, signaling, protection, data and control plane separation

Other Networking Trends

- Hottest Technologies: Storage, IP, Ethernet, Wireless, Optical
- Hottest Applications: Peer-to-peer (no money to be made by carriers), Storage, VOIP
- LAN:WAN Traffic Ratio : From 80/20 towards 20/80 (IP addressing and distance independent billing)
- Enterprise Market > Access > Metro > Core
- Financial Markets: No CLECs
- Glut of Fiber in long haul but shortage in Metro/Access
- Emergence of Metro Ethernet
- Emphasis on Security, Mobility

Summary



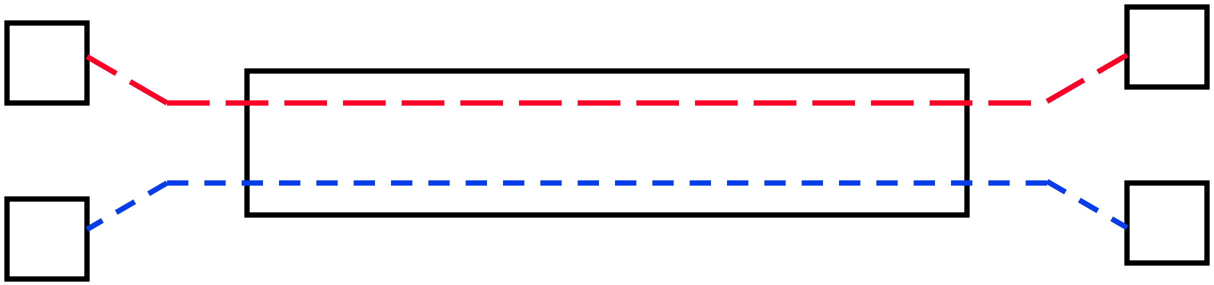
- Bottom of the hype cycle ⇒ Better days are ahead
- ILEC vs CLECs ⇒ Evolution vs Revolution
- Ethernet and IP in telecom ⇒ Need reliability, protection, and circuits
- More activity in access and metro than in core

Optical Tranmission and Switching

Raj Jain
Co-founder and Sr. VP
Nayna Networks, Inc.
481 Sycamore Dr, Milpitas, CA 95035
Email: jain@acm.org
www.nayna.com and http://www.cis.ohio-state.edu/~jain/

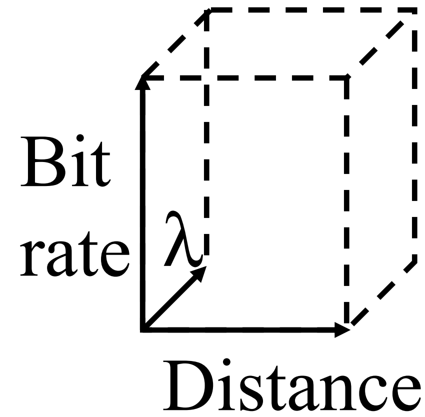
- Recent DWDM Records
- Transmission Products
- OEO vs OOO Switches
- Higher Speed: 40 Gbps
- More Wavelengths per fiber
- Ultra-Long Haul Transmission
- Free space optical communication

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 ≈ 0.1 - 2 nm separation
- Coarse = 2 to 25 nm = 4 to 12 λ's
- Wide = Different Wavebands

Recent DWDM Records

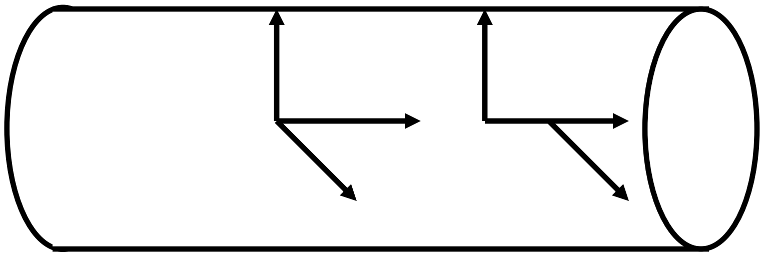


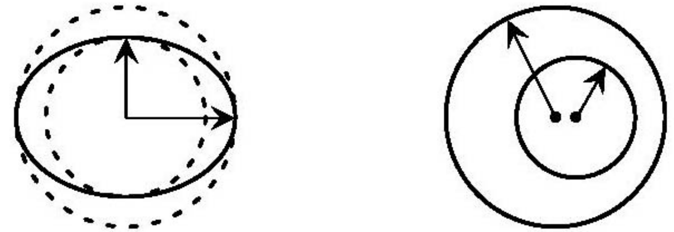
- 32λ× 5 Gbps to 9300 km (1998)
- 16λ× 10 Gbps to 6000 km (NTT'96)
- 160λ× 20 Gbps (NEC'00)
- 128λ× 40 Gbps to 300 km (Alcatel'00)
- 64λ× 40 Gbps to 4000 km (Lucent'02)
- 19λ× 160 Gbps (NTT'99)
- 7λ× 200 Gbps (NTT'97)
- 1λ×1200 Gbps to 70 km using TDM (NTT'00)
- 1022 Wavelengths on one fiber (Lucent'99)

Potential: 58 THz = 50 Tbps on 10,000 λ's

Ref: IEEE J. on Selected Topics in Quantum Electronics, 11/2000.

Polarization Mode Dispersion

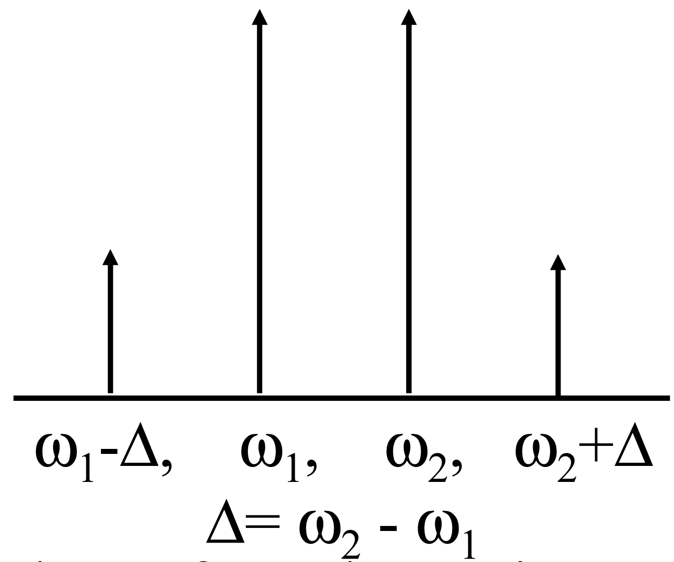




(a) Circularity (b) Concentricity

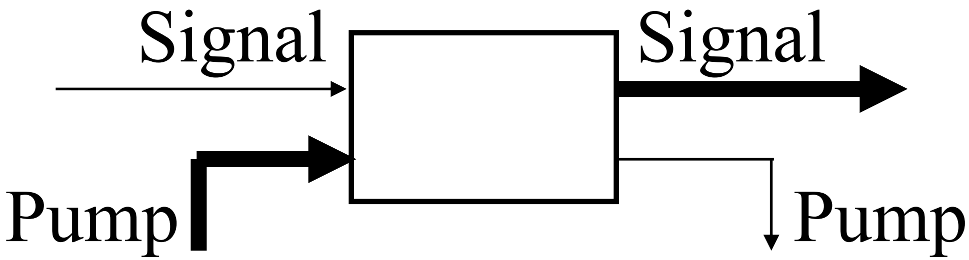
- Two polarization modes may travel at different speeds
- Non-circular core may increase PMD
- High winds may induce time-varying PMD on above-ground cables
- Polarization Mode Dispersion (PMD) limits distances to square of the bit rate
  $\Rightarrow$ 6400 km at 2.5 Gbps, 400 km at 10 Gbps, 25 km at 40 Gbps

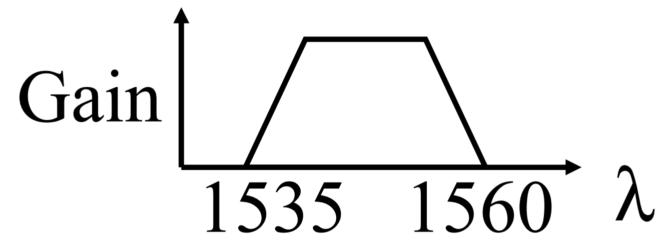
Four-Wave Mixing



- If two signals travel in the same phase for a long time, new signals are generated.

Amplifiers





- Erbium-Doped Fiber Amplifiers (EDFAs)
- Up to 30 dB amplification
- Flat response in 1535-1560 nm
  Fiber loss is minimum in this region
  Can be expanded to 40 nm width
- Raman Amplifiers: less noise, more expensive, and less gain than EDFA
  - Less noise ⇒ Critical for ultra-high bit rate systems

Transmission Products

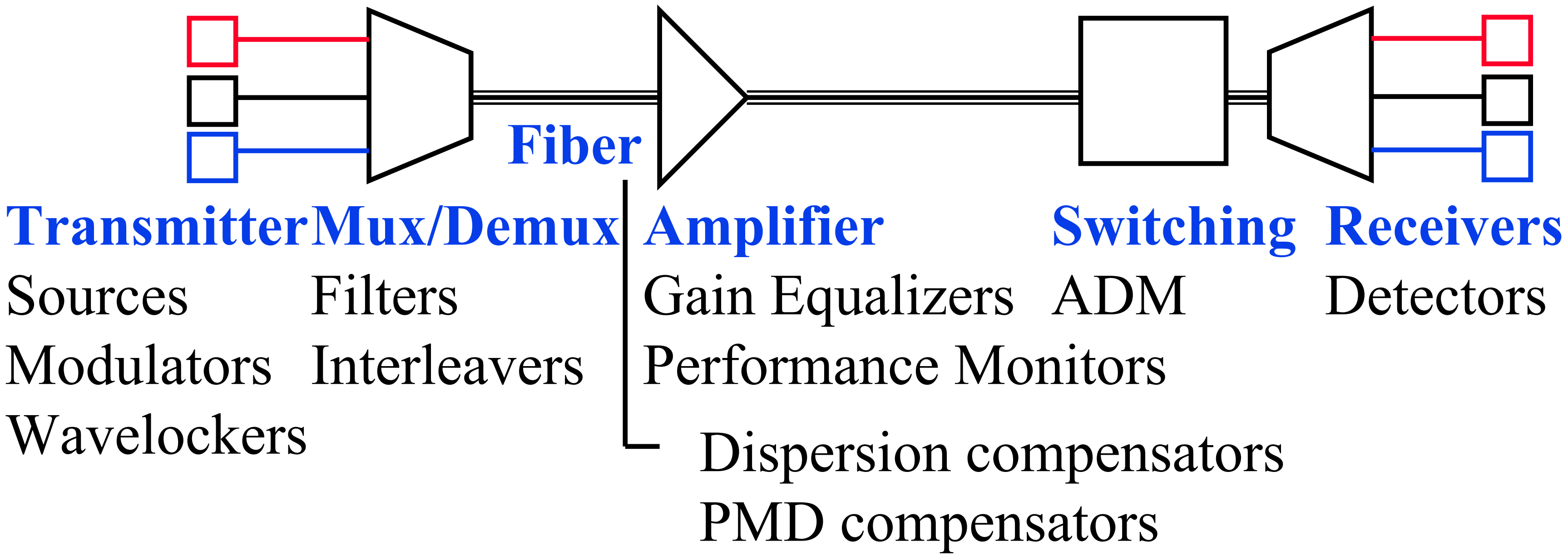
Product	λ's	Gb/s	km	Avail-ability
Siemens/Optisphere TransXpress	80	40	250	2001
	160	10	250	2001
Alcatel 1640 OADM	160	2.5	2300	2001
	80	10	330	2001
Corvis Optical Network Gateway	160	2.5	3200	2000
	40	10	3200	2000
Ciena Multiwave CoreStream	160	10	1600	2001
Nortel Optera LH4000	56	10	4000	2000
Optera LH 5000	104	40	1200	2002
Sycamore SN10000	160	10	800	2001
	40	10	4000	2001
Cisco ONS 15800	160	10	2000	2002

- Ref: "Ultra everything," Telephony, October 16, 2000

OEO vs OOO Switches

- OEO:
  - Requires knowing data rate and format, e.g., 10 Gbps SONET
  - Can multiplex lower rate signals
  - Cost/space/power increases linearly with data rate
- OOO:
  - Data rate and format independent
    $\Rightarrow$ Data rate easily upgraded
  - Sub-wavelength mux/demux difficult
  - Cost/space/power relatively independent of rate
  - Can switch multiple ckts per port (waveband)
  - Issues: Wavelength conversion, monitoring

40 Gbps

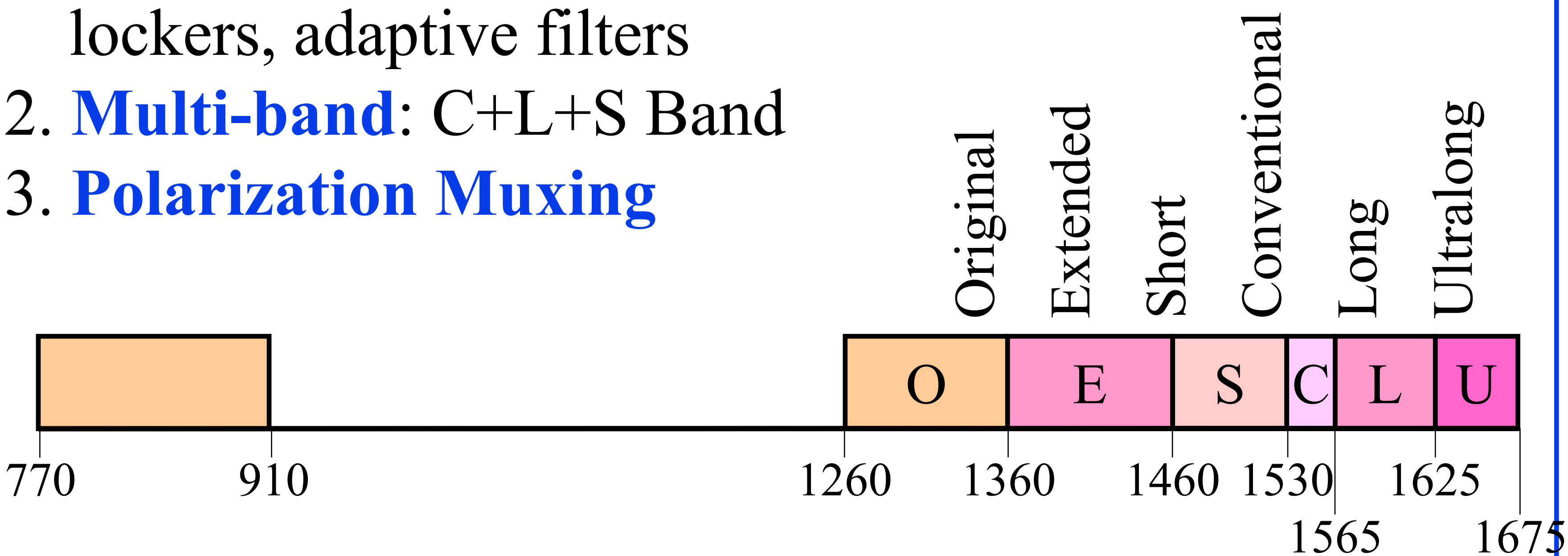


- Need all new optical and electronic components
- Non-linearity's reduce distance by square of rate.
- Deployment may be 2-3 years away
- Development is underway. To avoid 10 Gbps mistake.
- Cost goal: 2.5×10 Gbps

More Wavelengths

- C-Band (1535-1560nm), 1.6 nm (200 GHz) ⇒ 16 λ's
- Three ways to increase # of wavelengths:

1. **Narrower Spacing**: 100, 50, 25, 12.5 GHz
   Spacing limited by data rate. Cross-talk (FWM)
   Tight frequency management: Wavelength monitors, lockers, adaptive filters
2. **Multi-band**: C+L+S Band
3. **Polarization Muxing**



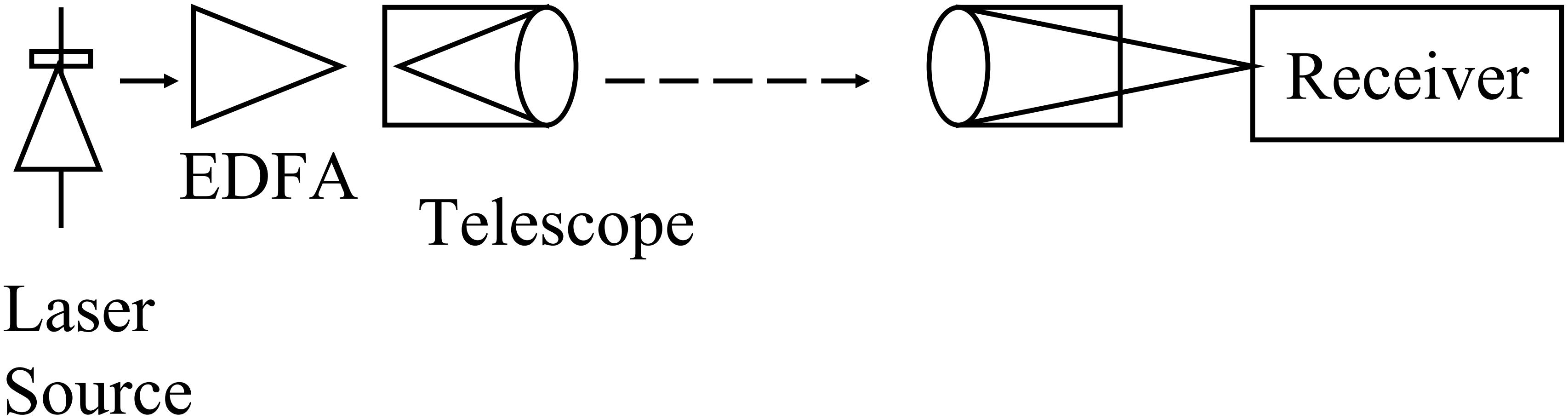
More Wavelengths (Cont)

- More wavelengths ⇒ More Power
  ⇒ Fibers with large effective area
  ⇒ Tighter control of non-linearity's
  ⇒ Adaptive tracking and reduction of polarization mode dispersion (PMD)

Ultra-Long Haul Transmission

1. Strong out-of-band Forward Error Correction (FEC)
   Changes regeneration interval from 80 km to 300km
   Increases bit rate from 40 to 43 Gbps
2. Dispersion Management: Adaptive compensation
3. More Power: Non-linearity's $\Rightarrow$ RZ coding
   Fiber with large effective area
   Adaptive PMD compensation
4. Distributed Raman Amplification:
   Less Noise than EDFA
5. Noise resistant coding: 3 Hz/bit by Optimight

Free Space Optical Comm



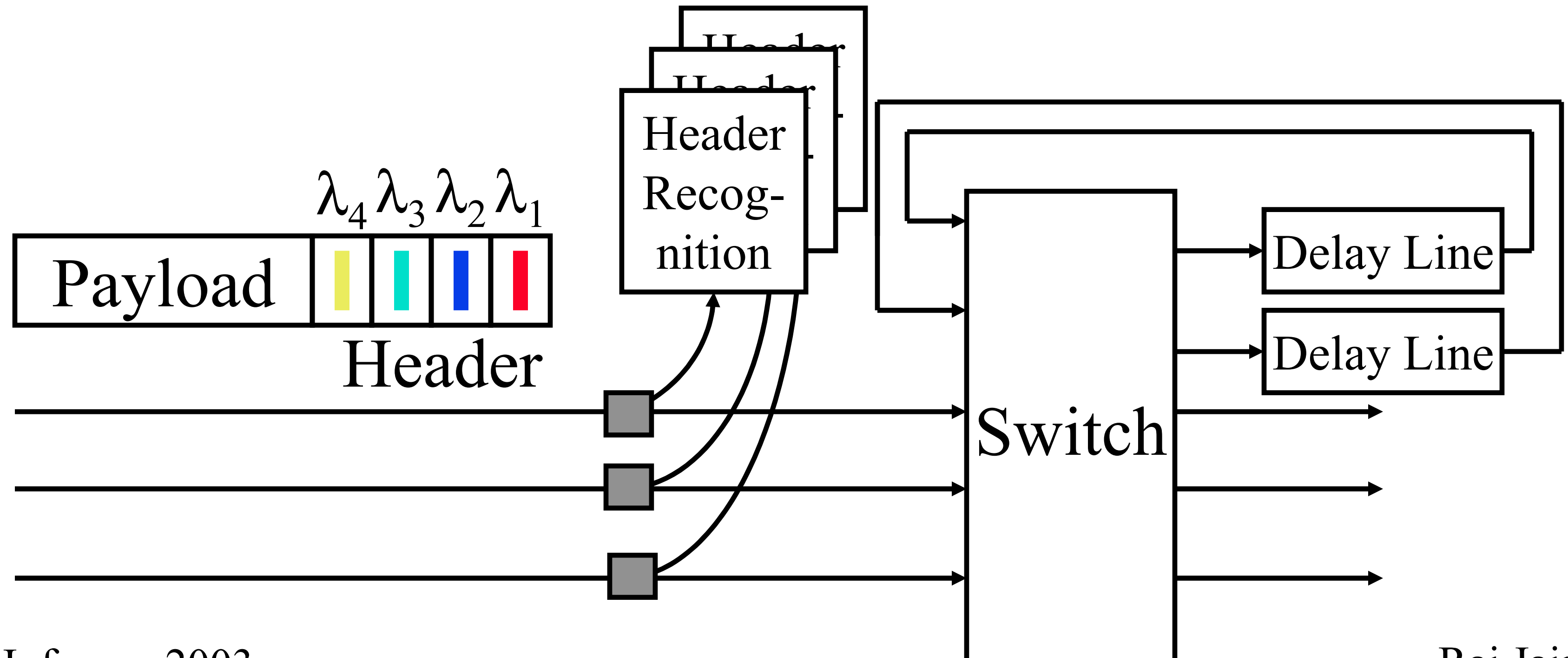
- Uses WDM in open air
- Sample Product:
  Lucent WaveStar OpticAir: 4×2.5Gbps to 5 km
  Available March'00.
- EDFA = Erbium Doped Fiber Amplifier

Free Space Optical Comm

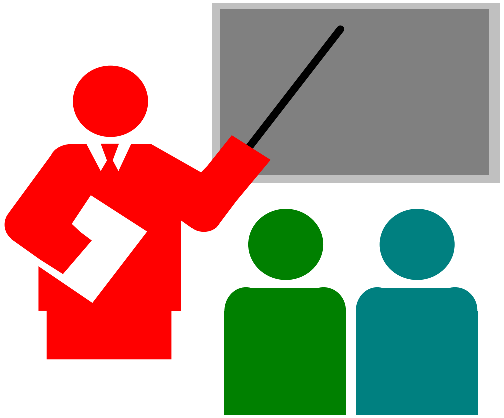
- No FCC Licensing required
- Immunity from interference
- Easy installation
  ⇒ Unlimited bandwidth, Easy Upgrade
- Transportable upon service termination or move
- Affected by weather (fog, rain)
  ⇒ Need lower speed Microwave backup
- Example Products: Optical Crossing Optibridge 2500 2.5Gbps to 2km, Texas Instruments TALP1135 Chipset for 10/100 Mbps up to 50m

Optical Packet Switching

- Header Recognition: Lower bit rate or different λ
- Switching
- Buffering: Delay lines, Dispersive fiber



Summary



- DWDM systems use 1550 nm band due to EDFA
- Raman Amplifiers for long distance applications
- O/O/O switches are bit rate and data format independent
- Tighter control of PMD is required for higher speed or longer distance transmission

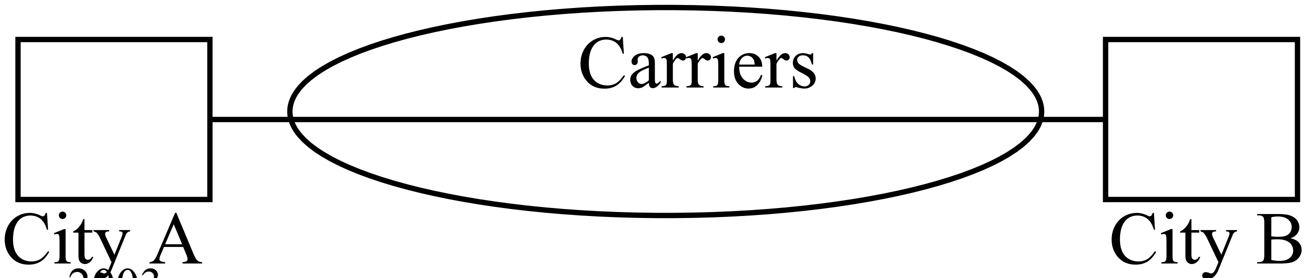
Carrier Networking Technologies

Raj Jain
Co-founder and Sr. VP
Nayna Networks, Inc.
481 Sycamore Dr, Milpitas, CA 95035
Email: jain@acm.org
www.nayna.com and http://www.cis.ohio-state.edu/~jain/

- SONET: Components, Frame Format
- Multiplexing and Concatenation
- Protection: Rings
- SDH, OTN

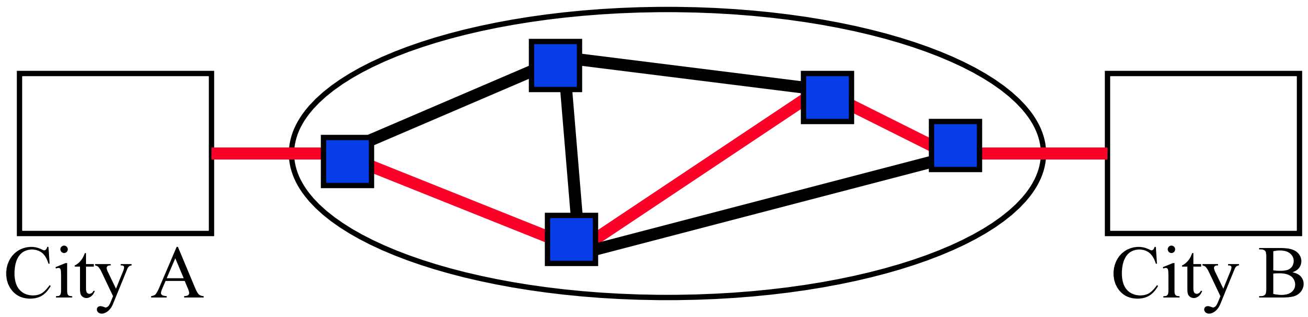
SONET

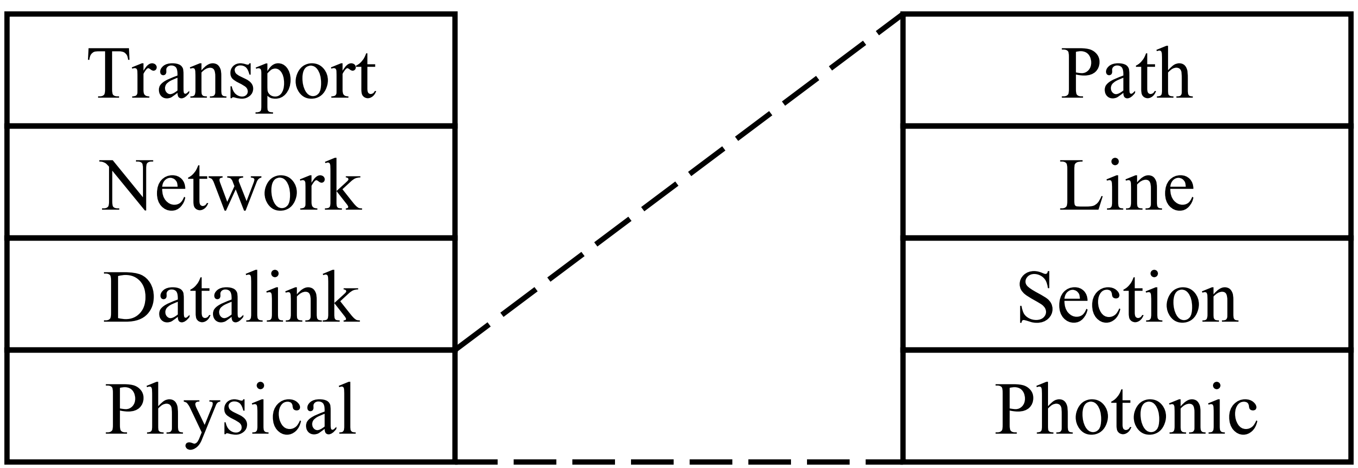
- Synchronous optical network
- Standard for digital optical transmission (bit pipe)
- Developed originally by Bellcore to allow mid-span meet between carriers: MCI and AT&T. Standardized by ANSI and then by ITU
  $\Rightarrow$ Synchronous Digital Hierarchy (SDH)
- You can lease a SONET connection from carriers



SONET Protocols

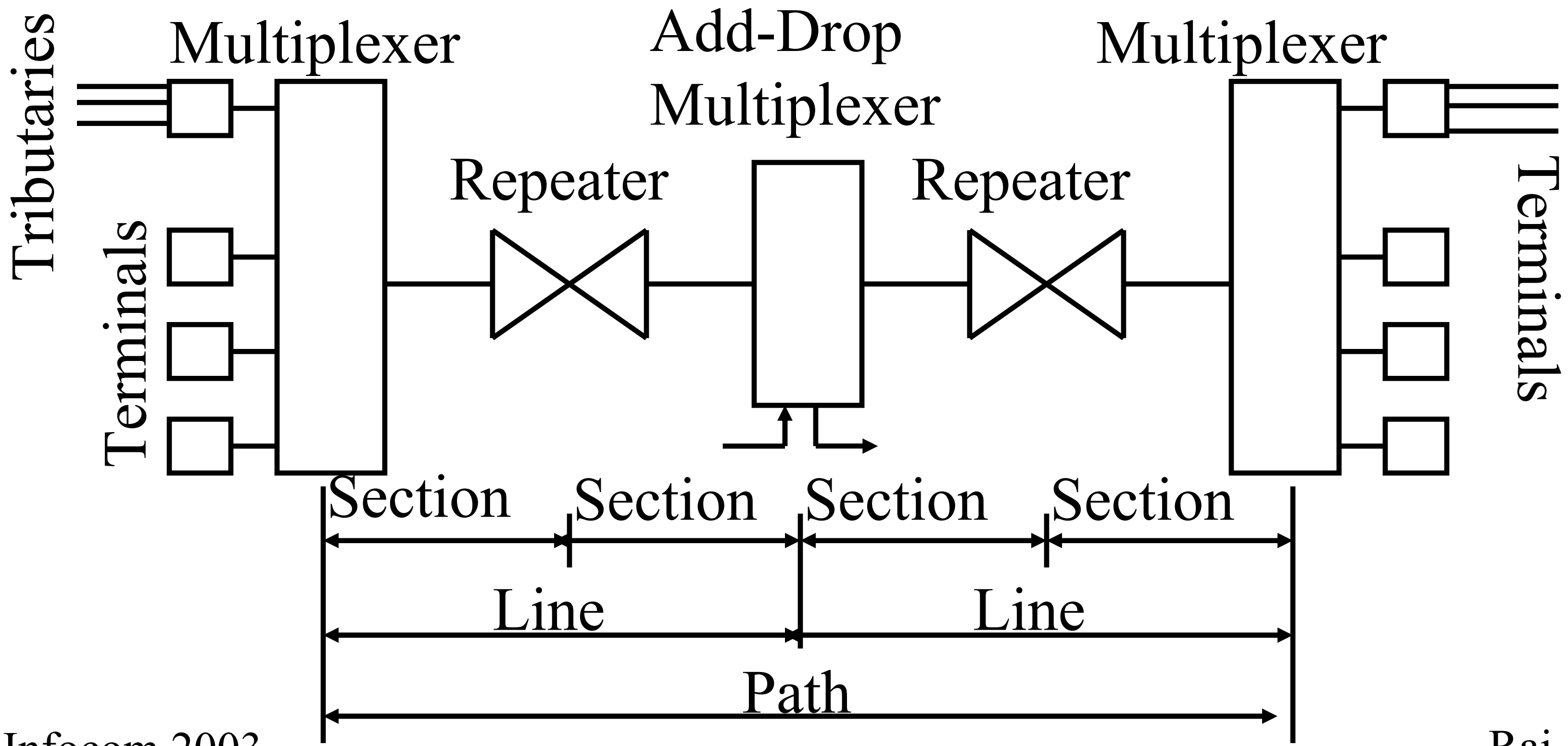
- Synchronous Optical **Network**





Physical Components

- Section = Single run of fiber
- Line = Between multiplexers
- Path = End-to-end



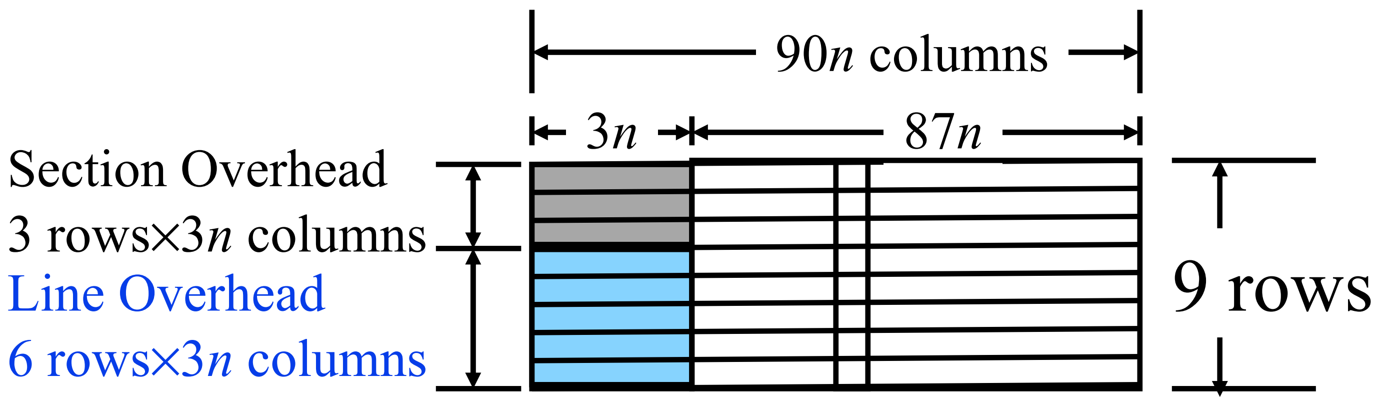
Signal Hierarchy

Synchronous Transport Signal Level n = STS-n = $n \times 51.84$ Mbps
STM=Synchronous Transport Module, OC=Optical Carrier level

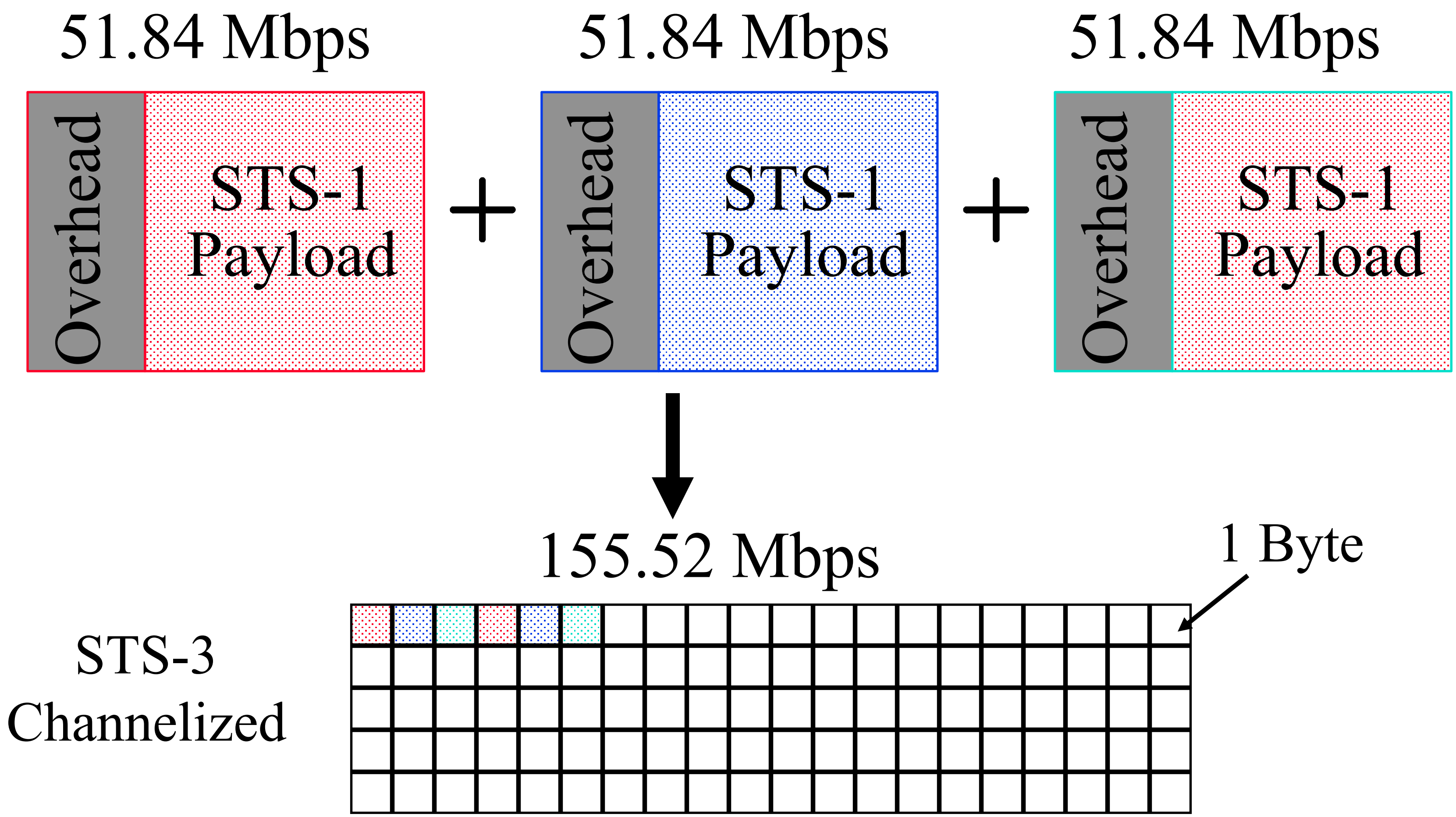
ANSI Designation	Optical Signal	CCITT Designation	Data Rate (Mbps)	Payload Rate (Mbps)
STS-1	OC-1		51.84	50.112
STS-3	OC-3	STM-1	155.52	150.336
STS-9	OC-9	STM-3	466.56	451.008
STS-12	OC-12	STM-4	622.08	601.344
STS-18	OC-18	STM-6	933.12	902.016
STS-24	OC-24	STM-8	1244.16	1202.688
STS-36	OC-36	STM-12	1866.24	1804.032
STS-48	OC-48	STM-16	2488.32	2405.376
STS-96	OC-96	STM-32	4976.64	4810.176
STS-192	OC-192	STM-64	9953.28	9620.928

SONET Frame Format

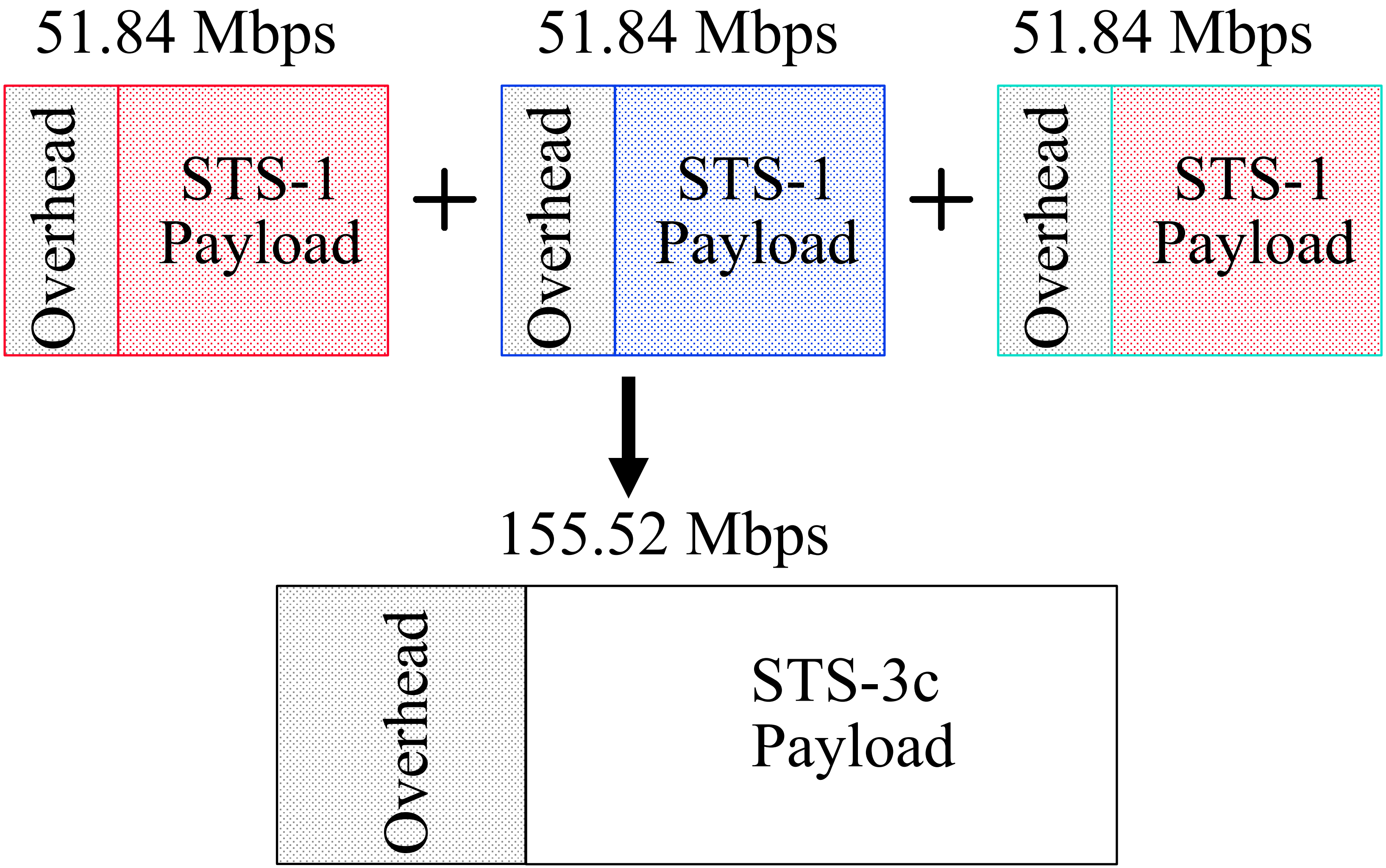
- OC-1 = 51.84 Mbps (payload and overhead)
- OC-n = $n \times$ 51.84 Mbps
  e.g., OC-3 = 3 × 51.84 = 155.54 Mbps
- All SONET frames are 125 μs long.
  E.g., OC-3 frames are 2430 (125 × 155.54) bytes
- Represented as 2D arrays of bytes.
  9 rows × 90n columns. Transmitted row-wise



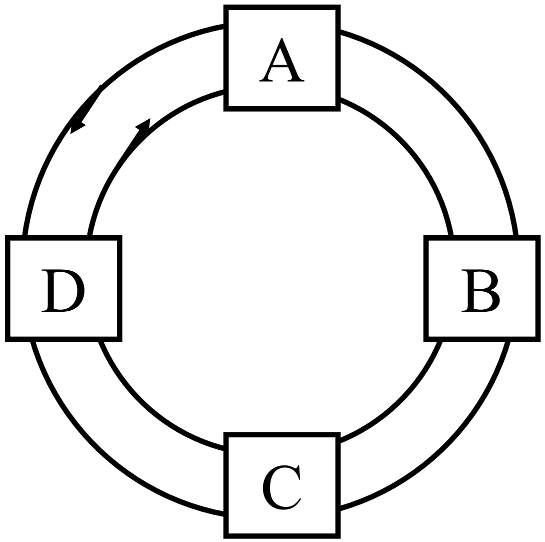
Multiplexing

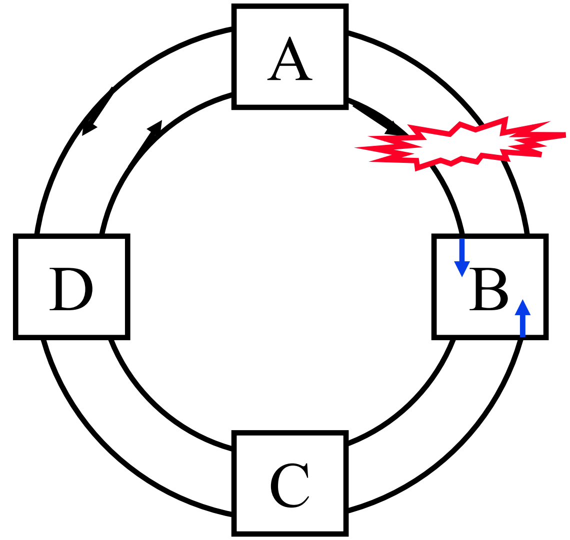


Concatenation



Unidirectional Path Switched Ring



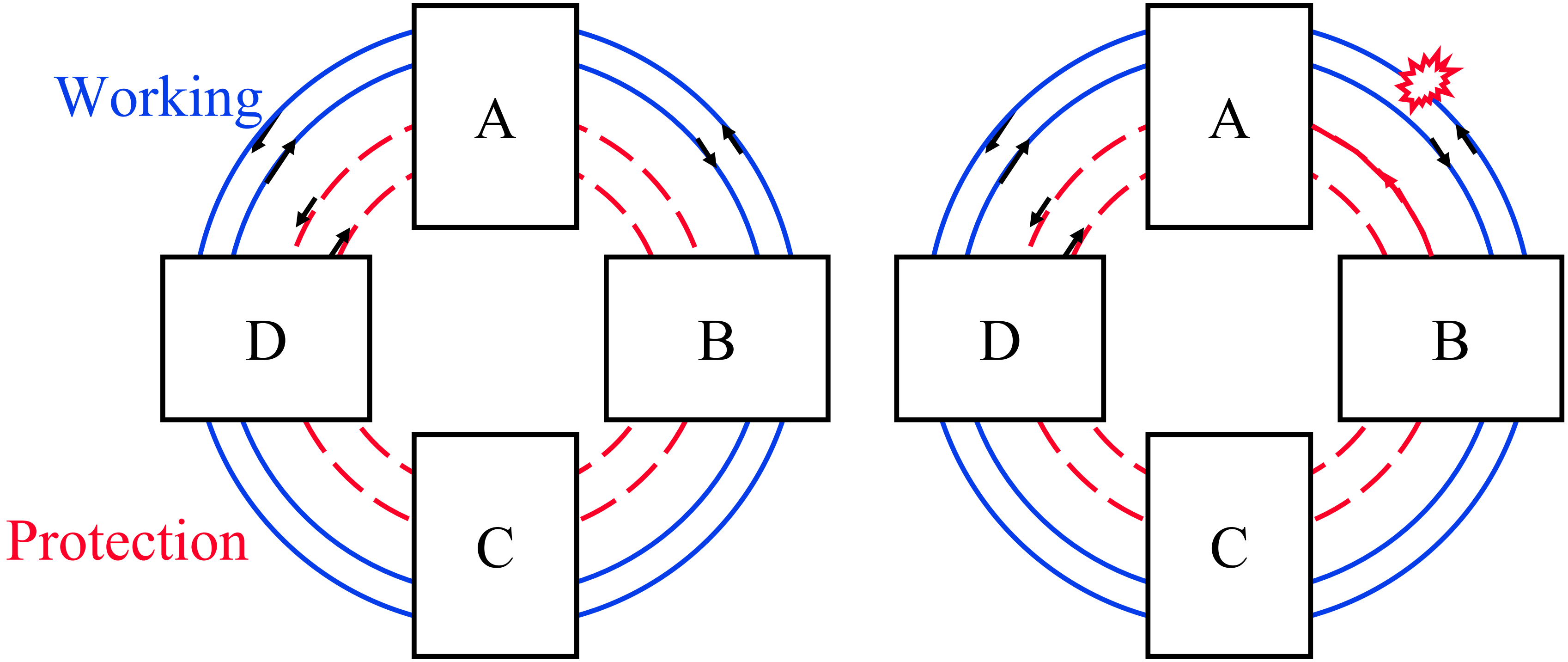


- Two counter rotating fibers: working+protection
- 1+1 ⇒ Signal is sent on both fibers, receiver takes the stronger signal
- Unidirectional: Working ring is in one direction

UPSR (Cont)

- Path-Switched: the path changes on a link failure. SONET Path overhead is used.
- Receiver controls the switching. No transmitter involvement $\Rightarrow$ Fast
- No APS signaling channel required

Bi-directional Line-Switched Ring

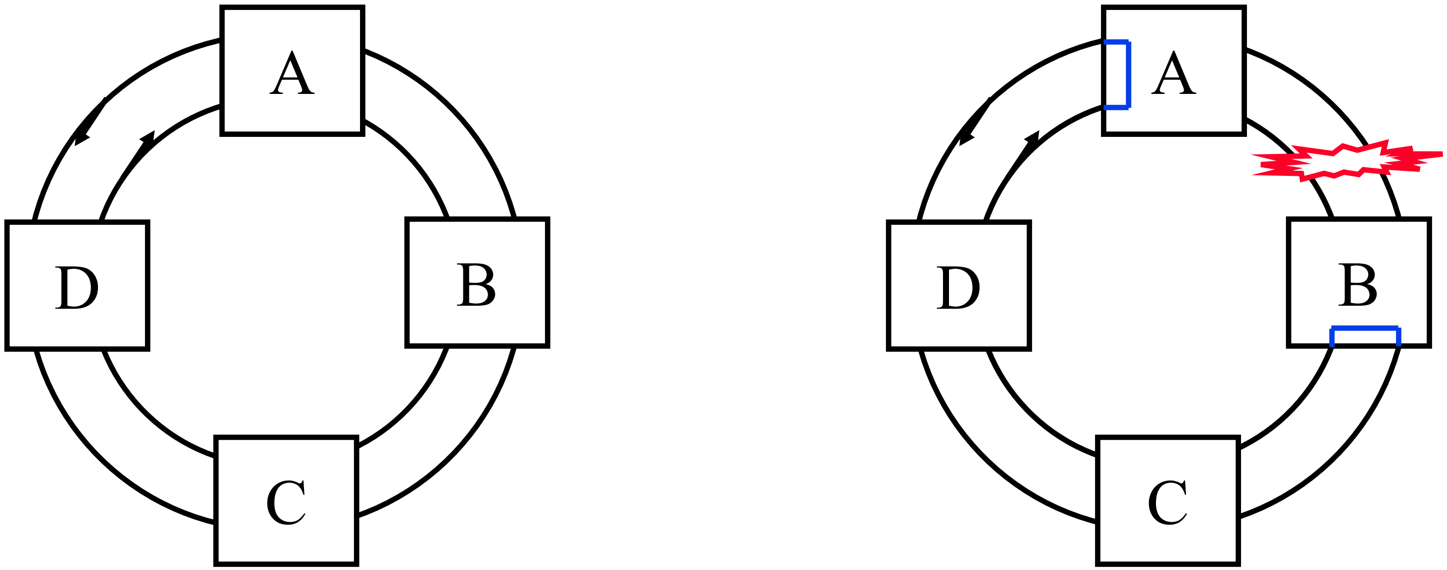


- Two working rings in counter-rotating directions
- Two protection rings in counter-rotating directions
- Bi-directional: Working signals between two nodes on shortest path in both directions

4-Fiber BLSR (Cont)

- Line Switched: If only one fiber is cut, traffic is switched from working to protection fiber in the same direction
- SONET line overhead is used for APS signaling
- Ring Switched: If both fibers are cut, traffic is switched to protection ring
- 1:1 Protection: APS signaling channel is required
- Signaling ⇒ Restoration time more than UPSR
- Preferred by long-haul carriers.

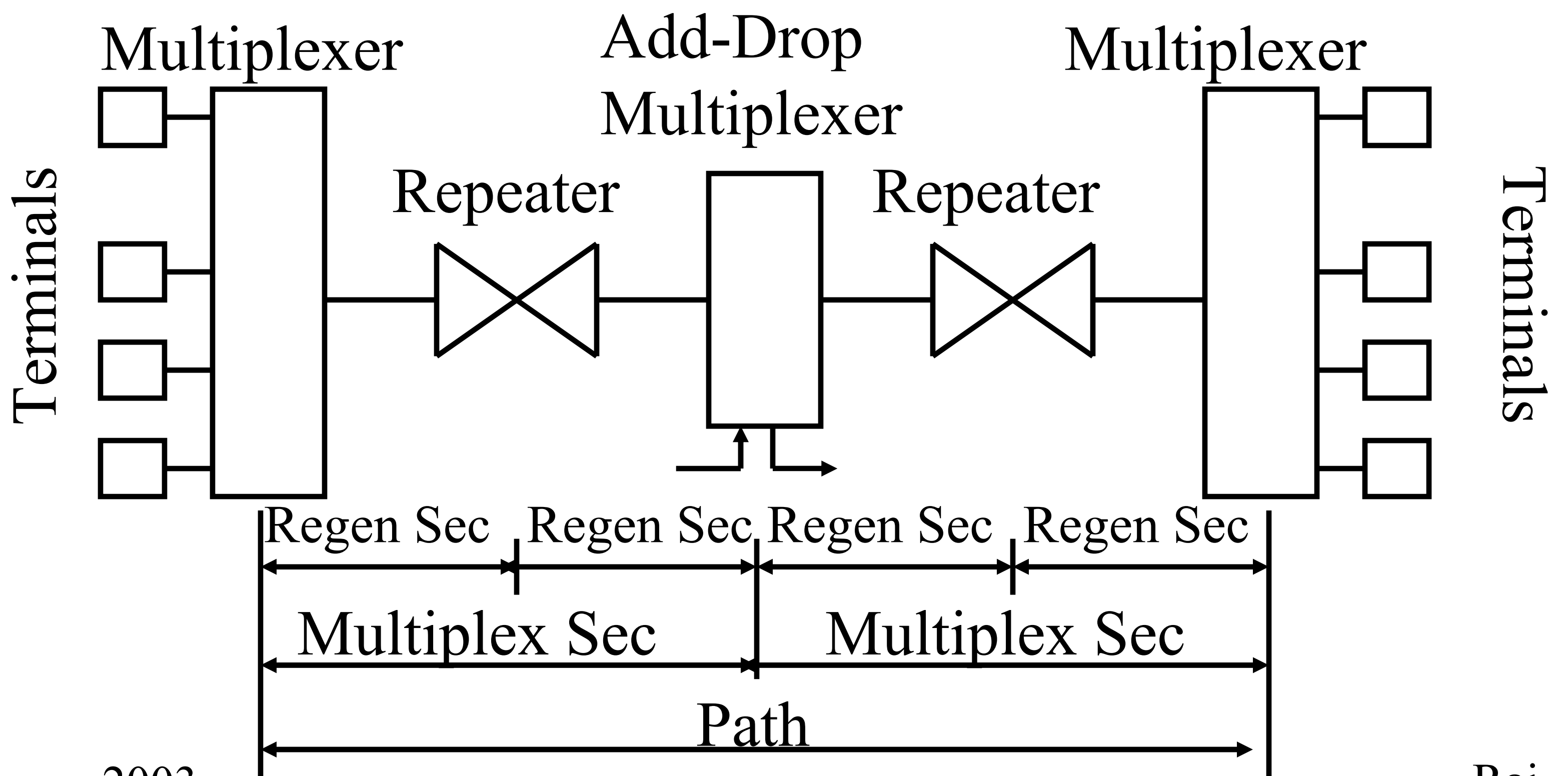
Bi-directional Line-Switched Ring (2-Fiber)



- Two counter rotating rings: both 1/2 working and 1/2 protection using TSI
- Allows only ring switching if one fiber is cut
- Ring wraps if both fibers are but

Synchronous Digital Hierarchy

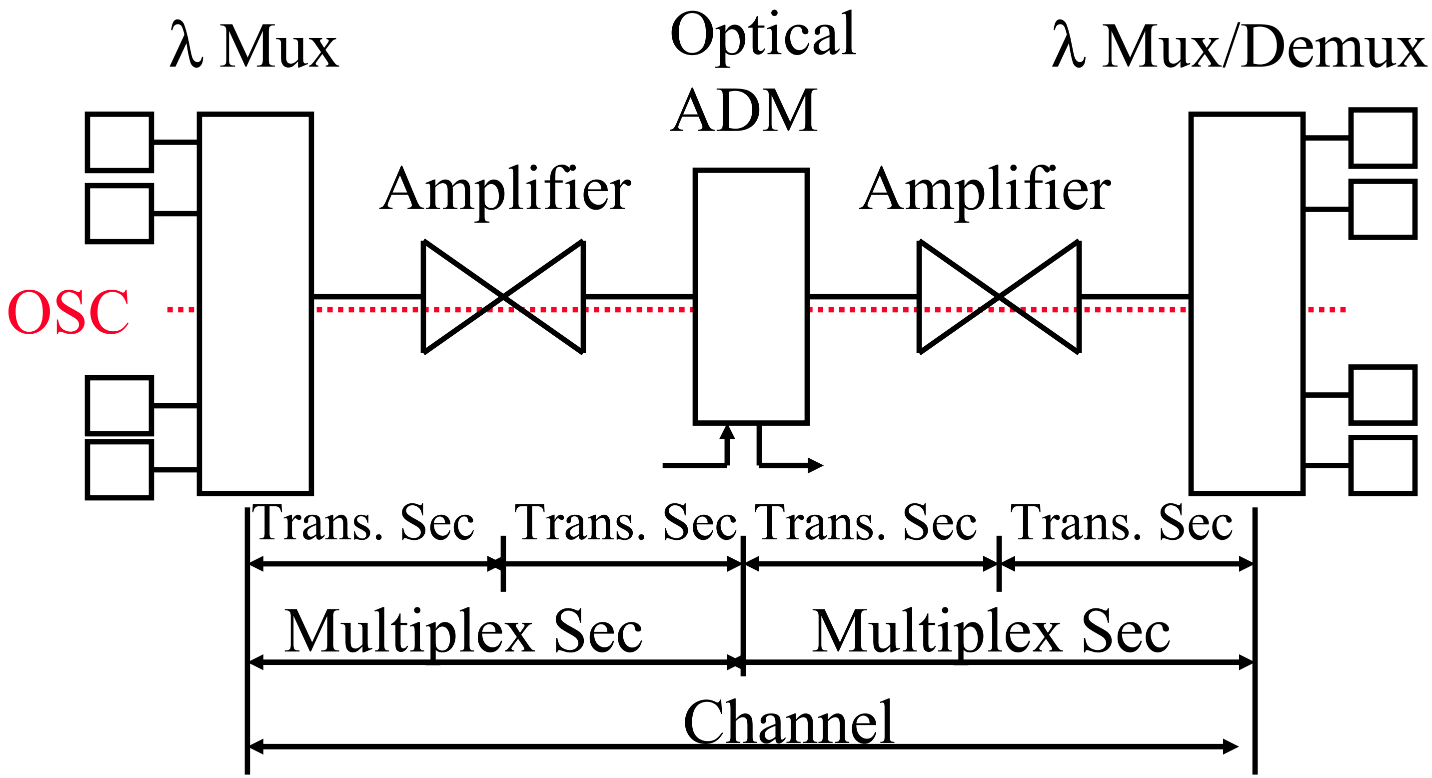
- Regenerator Section = Single run of fiber
- Multiplex Section = Between multiplexers
- Path = End-to-end



SONET vs SDH

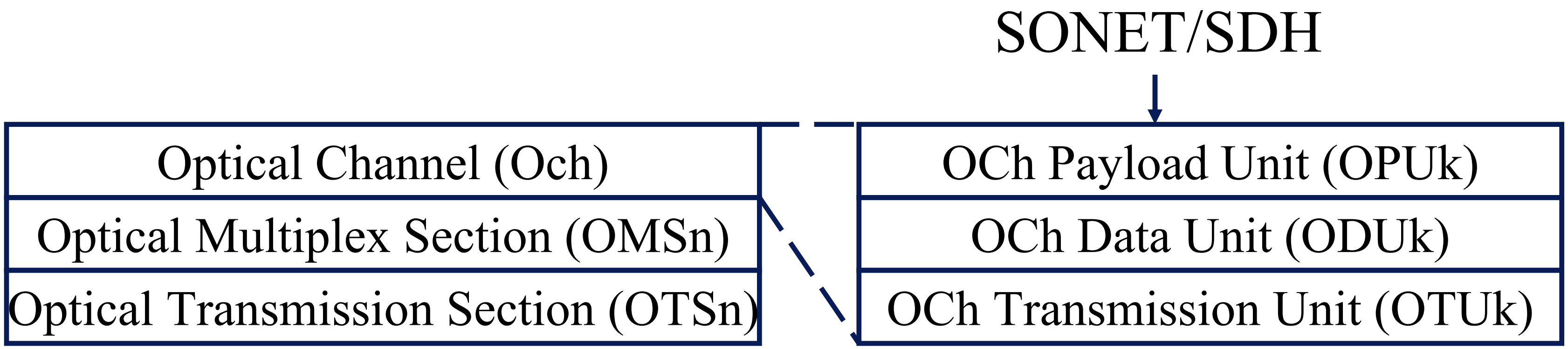
SONET	SDH
Section	Regeneration Section
Line	Multiplex Section
Path	Path
Byte	Octet
Tributary	Container
	Virtual Container
Virtual Tributary	Tributary Unit
Virtual Tributary Group	Tributary Unit Group
	Administrative Unit
UPSR	SNCP
BLSR	MS-SPRing

Optical Transport Network (OTN)

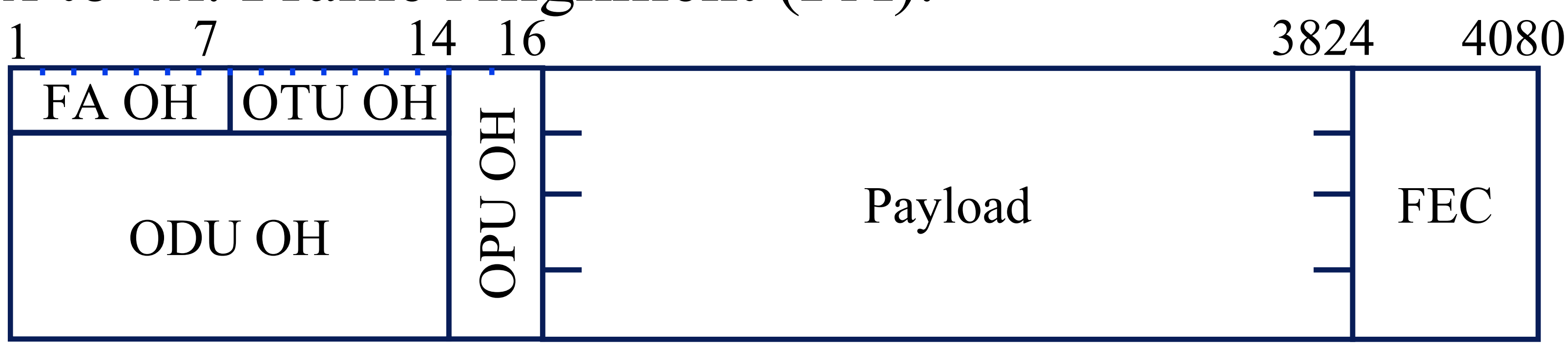


- G.709 Digital Wrapper designed for WDM networks
- OTN*n.k* = *n* wavelengths at *k*th rate, 2.5, 10, 40 Gbps plus one Optical Supervisory Channel (OSC)
- OTNnr.k = Reduced OTNn.k ⇒ Without OSC

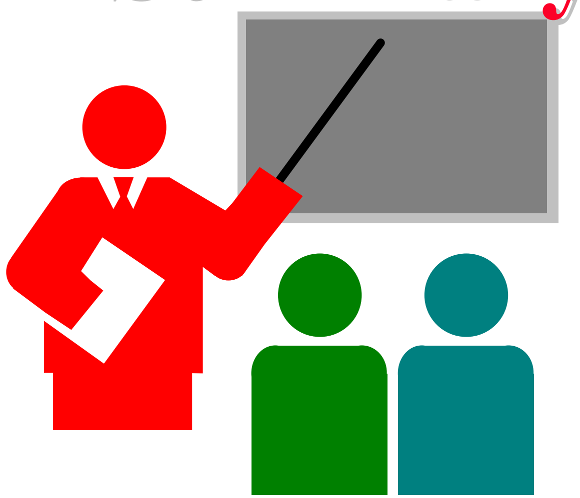
OTN Layers and Frame Format



- **OTU1 Frame Format:** 4×4080 Octets/125 ms
  Forward Error Correction (FEC) increases distance by 2x to 4x. Frame Alignment (FA).



Summary



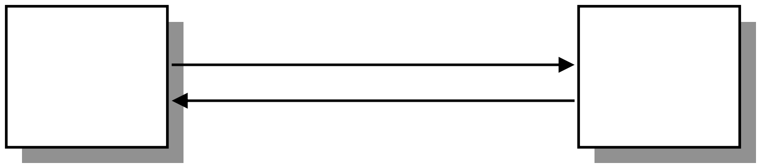
- STS-n = OC-n = n X 51. Mbps line rate
  STM-n = STS-3n is used in Europe
- SONET/SDH have ring based protection
- OTN uses FEC digital wrapper and allows WDM

Next Generation Data Networking

Raj Jain
Co-founder and Sr. VP
Nayna Networks, Inc.
481 Sycamore Dr, Milpitas, CA 95035
Email: jain@acm.org
www.nayna.com and http://www.cis.ohio-state.edu/~jain/

- Gigabit Ethernet
- 10 G Ethernet
- Resilient Packet Rings
- Next Generation SONET: VCAT, GFP, LCAS

Full-Duplex Ethernet



- Uses point-to-point links between TWO nodes
- Full-duplex bi-directional transmission
- Transmit any time
- Many vendors are shipping switch/bridge/NICs with full duplex
- No collisions $\Rightarrow$ 50+ Km on fiber.
- Between servers and switches or between switches

1 GbE: Key Design Decisions

- P802.3z ⇒ Update to 802.3
  Compatible with 802.3 frame format, services, management
- 1000 Mb vs. 800 Mb Vs 622 Mbps
  Single data rate
- LAN distances only
- No Full-duplex only ⇒ Shared Mode
  Both hub and switch based networks
- Same min and max frame size as 10/100 Mbps
  ⇒ Changes to CSMA/CD protocol
  Transmit longer if short packets

1000Base-X

- 1000Base-LX: 1300-nm laser transceivers
  - 2 to 550 m on 62.5-μm or 50-μm multimode, 2 to 5000 m on 10-μm single-mode
- 1000Base-SX: 850-nm laser transceivers
  - 2 to 275 m on 62.5-μm, 2 to 550 m on 50-μm. Both multimode.
- 1000Base-CX: Short-haul copper jumpers
  - 25 m 2-pair shielded twinax cable in a single room or rack.

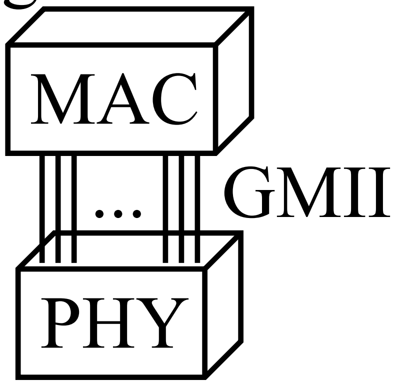
    Uses **8b/10b** coding ⇒ 1.25 GBaud/s line rate
- *1000Base-ZX: Long haul lasers to 70 km (not Std)*

1000Base-T

- 100 m on 4-pair Cat-5 UTP
  ⇒ Network diameter of 200 m
- Applications: Server farms, High-performance workgroup, Network computers
- Supports CSMA/CD (Half-duplex):
  Carrier Extension, Frame Bursting
- 250 Mbps/pair full-duplex DSP based PHY
  ⇒ Requires new 5-level (PAM-5) signaling
  with 4-D 8-state Trellis code FEC
- FEC coded symbols.
  Octet data to 4 quinary (5-level) symbols and back,
  e.g., 001001010 = {0, -2, 0, -1}

1000BASE-T (Cont)

- Inside PHY, before coding, the data is scrambled using $x^{33}+x^{20}+1$ in one direction and $x^{33}+x^{13}+1$ self-synchronizing scrambler in the other direction
- Automatically detects and corrects pair-swapping, incorrect polarity, differential delay variations across pairs
- Autonegotiation ⇒ Compatibility with 100Base-T
- Complies with Gigabit Media Independent Interface
- 802.3ab-1999



10 GbE: Key Design Decisions

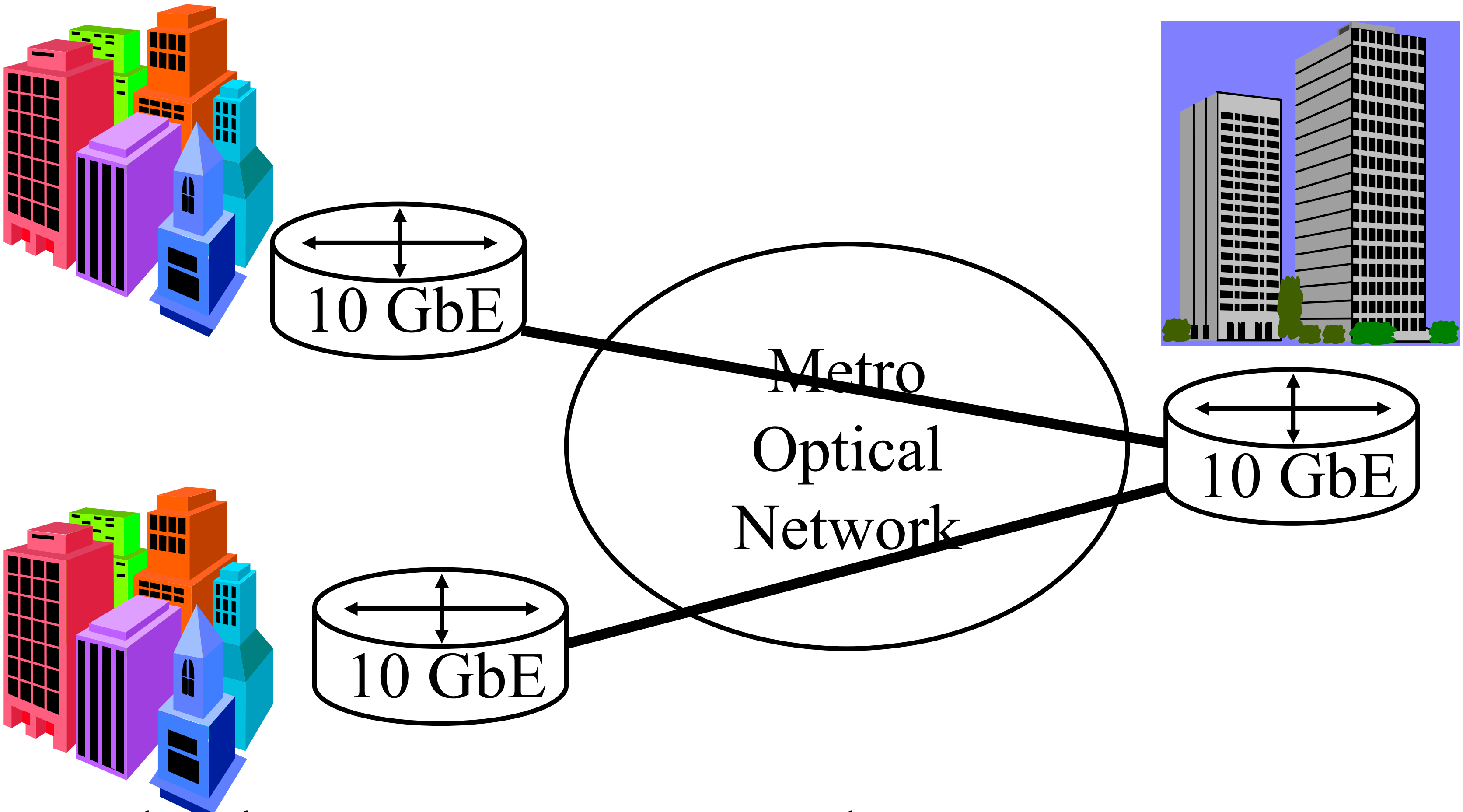
- P802.3ae ⇒ Update to 802.3
  Compatible with 802.3 frame format, services, management
- 10 Gbps vs. 9.5 Gbps. **Both** rates.
- LAN and MAN distances
- Full-duplex only ⇒ No Shared Mode
  Only switch based networks. No Hubs.
- Same min and max frame size as 10/100/1000 Mbps
  Point-to-point ⇒ No CSMA/CD protocol
- 10.000 Gbps at MAC interface
  ⇒ Flow Control between MAC and PHY

10 GbE PMD Types

PMD	Description	MMF	SMF
10GBASE-R:			
10GBASE-SR	850nm Serial LAN	300 m	N/A
10GBASE-LR	1310nm Serial LAN	N/A	10 km
10GBASE-ER	1550nm Serial LAN	N/A	40 km
10GBASE-X:			
10GBASE-LX4	1310nm WWDM LAN	300 m	10 km
10GBASE-W:			
10GBASE-SW	850nm Serial WAN	300 m	N/A
10GBASE-LW	1310nm Serial WAN	N/A	10 km
10GBASE-EW	1550nm Serial WAN	N/A	40 km
10GBASE-LW4	1310nm WWDM WAN	300 m	10 km

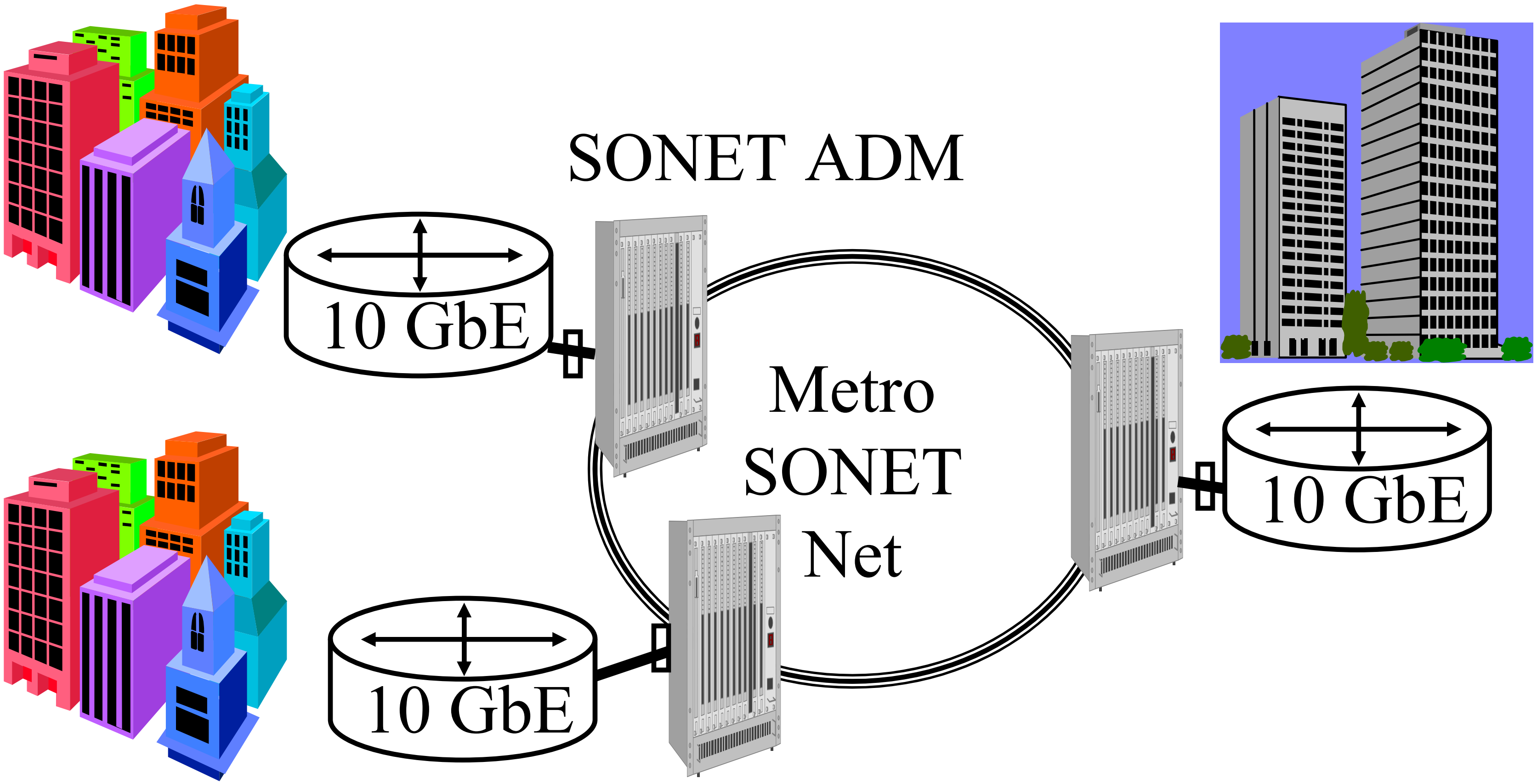
- S = Short Wave, L=Long Wave, E=Extra Long Wave
- R = Regular reach (64b/66b), W=WAN (64b/66b + SONET Encapsulation), X = 8b/10b
- 4 = 4 λ's

10 GbE over Dark Fiber



- Need only LAN PMD up to 40 km. No SONET overhead. No protection.

10 GbE over SONET/SDH



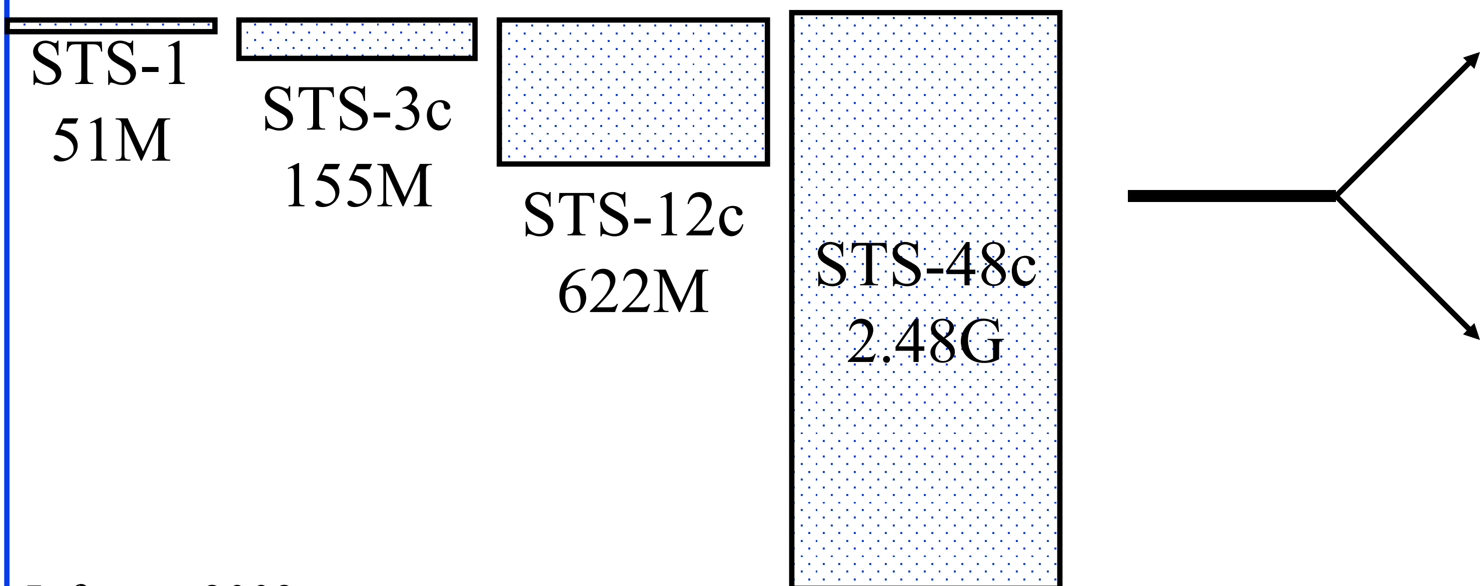
- Using WAN PMD.
  Legacy SONET. Protection via rings.
  ELTE = Ethernet Line Terminating Equipment

Future Possibilities

- 40 Gbps
- 100 Gbps:
  - 16λ×6.25 Gbps
  - 8λ × 12.5 Gbps
  - 4λ × 12.5 using PAM-5
- 160 Gbps
- 1 Tbps:
  - 12 fibers with 16λ × 6.25 Gbps
  - 12 fibers with 8λ × 12.5 Gbps
- 70% of 802.3ae members voted to potentially start 40G in 2004

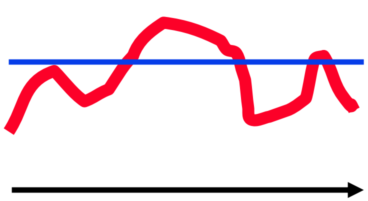
Data over SONET: Problems

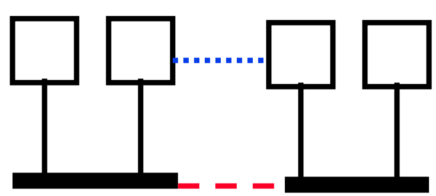
1. Rates highly discrete: In units of STS-3c's. Can't do STS-2c.
2. Entire payload on one path. No splitting, no multipath.
3. Size mismatch: 10 Mbps over 51.84, 100 Mbps over 155 Mbps, 1 Gbps over 1.24 Gbps

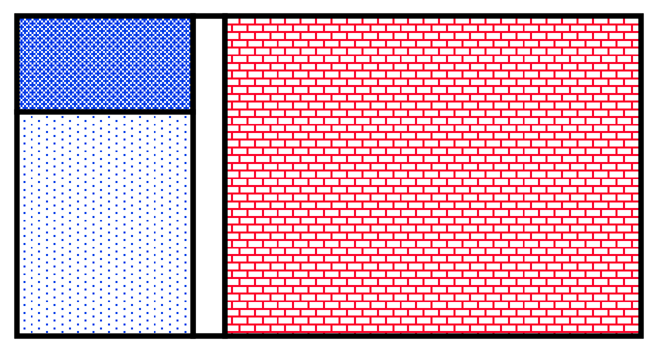


SONET Problems (Cont)

4. Data is bursty (Dynamic). SONET is fixed (static).
5. Inefficient Transparent Connections:
   1 GE = 1.25 Gbps at PHY layer ⇒ Needs OC-48c
6. Only one type of payload per stream: TDM, ATM, FDDI, Packets, Ethernet, Fiber Channel







Feature	SONET	Ethernet
Payload Rates	51M, 155M, 622M, 2.4G, 9.5G	10M, 100M, 1G, 10G
Payload Rate Granularity	Fixed	√Any
Bursty Payload	No	√Yes
Payload Count	One	√Multiple
Protection	√Ring	Mesh
OAM&P	√Yes	No
Synchronous Traffic	√Yes	No
Restoration	√50 ms	Minutes
Cost	High	√Low
Used in	Telecom	Enterprise

Feature	SONET	Ethernet	Remedy
Payload Rates	51M, 155M, 622M, 2.4G, 9.5G	10M, 100M, 1G, 10G	10GE at 9.5G
Payload Rate Granularity	Fixed	√Any	Virtual Concatenation
Bursty Payload	No	√Yes	Link Capacity Adjustment Scheme
Payload Count	One	√Multiple	Packet GFP
Protection	√Ring	Mesh	Resilient Packet Ring (RPR)
OAM&P	√Yes	No	In RPR
Synchronous Traffic	√Yes	No	MPLS + RPR
Restoration	√50 ms	Minutes	Rapid Spanning Tree
Cost	High	√Low	Converging
Used in	Telecom	Enterprise	

Rapid Spanning Tree Protocol (RSTP)

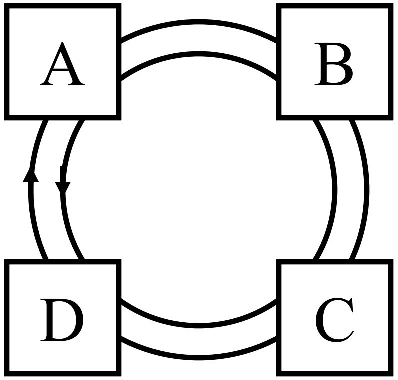
1. Builds upon the known topology rather than starting fresh.
2. Topology change is sent along designated ports (to sub-tree). Not all ports.
3. If the root port becomes disabled, alternate port becomes root port.
4. The learned address database (stations towards the root) is not flushed but transferred.
5. Edge ports and point-to-point LANs are treated efficiently. Old STP assumed all LANs are shared and have multiple bridges.

Ref: IEEE 802.1w-2001, October 25, 2001

Networking and Religion

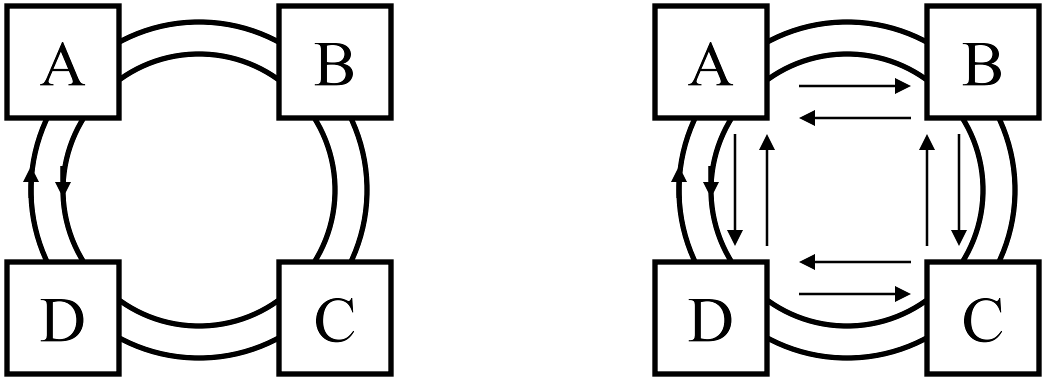
Both are based on a set of beliefs

RPR: Key Features



- Dual Ring topology
- Supports broadcast and multicast
- Packet based ⇒ Continuous bandwidth granularity
- Max 256 nodes per ring
- MAN distances: Several hundred kilometers.
- Gbps speeds: Up to 10 Gbps

RPR Features (Cont)



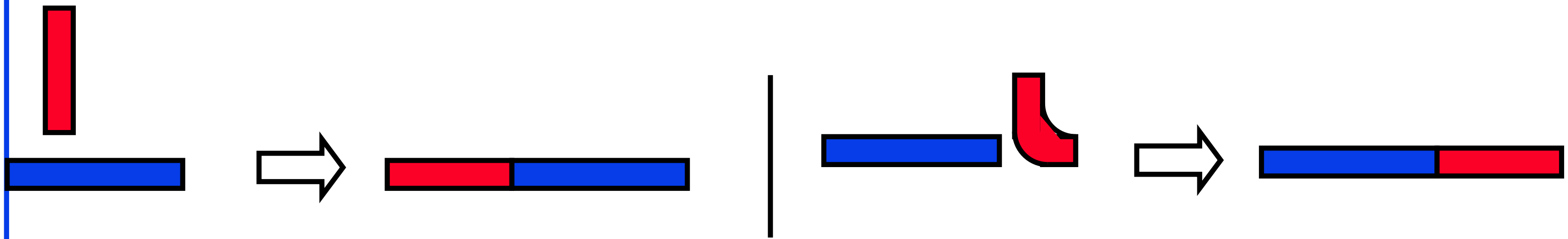
- Both rings are used (unlike SONET)
- Normal transmission on the shortest path
- Destination stripping ⇒ Spatial reuse
  Multicast packets are source stripped

RPR Service Classes

CoS		QoS				
Class	**Rate**	**Bandwidth**	**Delay Jitter**	**Subtype**	**Type**	**Rate Limiting**
A	A0	Guaranteed	Low	Committed	Allocated	Allocation Shapers
	A1	Guaranteed	Low	Reclaimable	Allocated	Allocation Shapers
B	CIR	Guaranteed	Bounded	Reclaimable	Allocated	Allocation Shapers
	EIR	Residual	Unbounded	Opportunistic	Allocated	Fairness Shapers
C	-	Residual	Unbounded	Opportunistic		Fairness Shapers

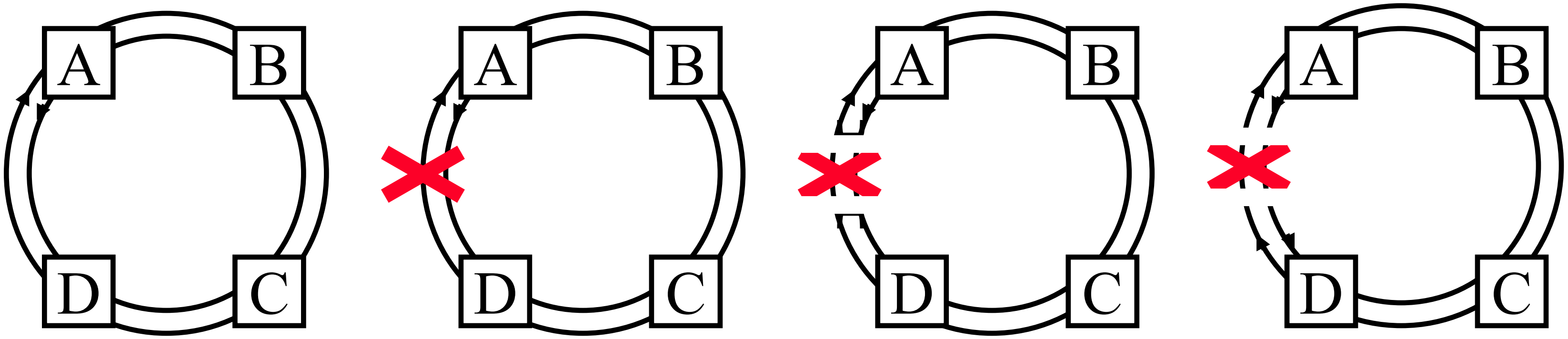
- Committed Information Rate (CIR), Excess Information Rate (EIR)
- Reclaimed: Unused bandwidth can be used by opportunistic traffic
- Committed: Can not be reclaimed around the ring.

RPR (Cont)



- Buffer Insertion Ring: Absolute but non-preemptive priority to pass-through traffic
- Cut-through of transit packets optional.
- Bandwidth management: Unused bandwidth is advertised so that others can use it
- Fairness Algorithm for fair and efficient bandwidth use
- Physical Layer Independent: GbE/10GE or SONET with GFP or PoS

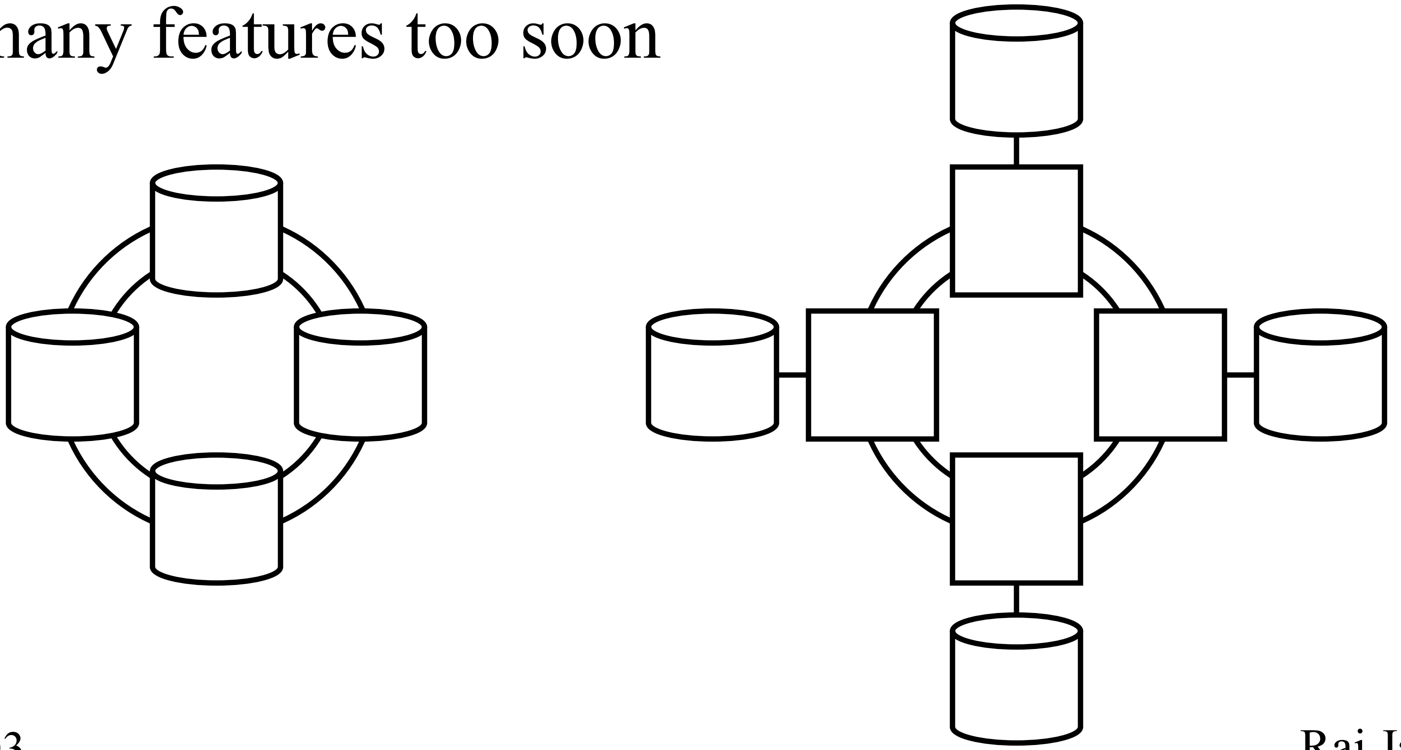
RPR Protection Mechanisms



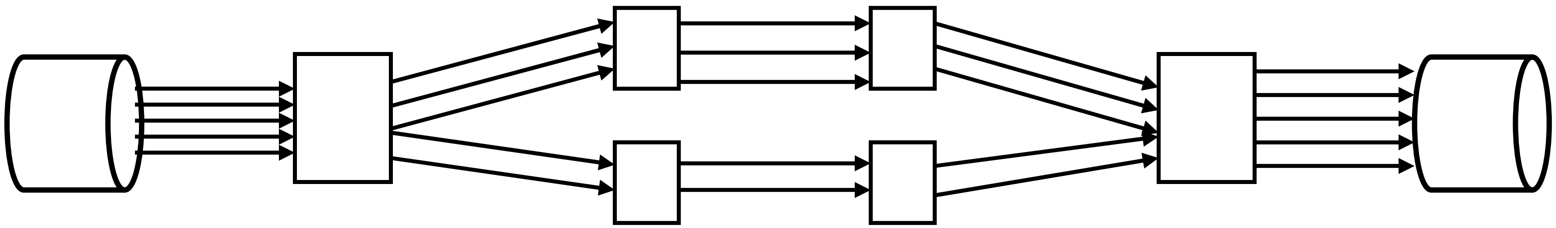
1. **Wrapping**: Stations adjacent to failure wrap. After re-org, packets sent on shortest path. Multicast packets are sent on **one** ring with TTL=Total number of stations.
2. **Source Steering**: Failure detecting station sends a Protection Request message to every station. Sources select appropriate ringlet to reach their destination. Multicast packets are sent on **both** rings with TTL=Total number of stations

RPR Issues

- Ring vs Mesh (Atrica)
- Router Feature vs Dedicated RPR Node (Cisco, Redback, Riverstone vs Luminous)
- Too many features too soon

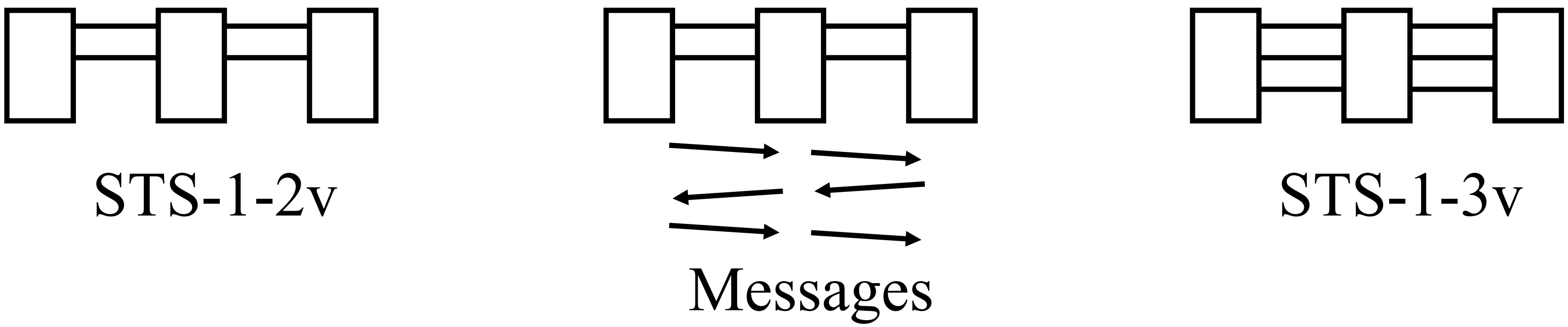


SONET Virtual Concatenation



- VCAT: Bandwidth in increments of VT1.5 or STS-1
- For example: 10 Mbps Ethernet in 7 T1's = VT1.5-7v
  100 Mbps Ethernet in 2 OC-1 = STS-1-2v,
  1GE in 7 STS-3c = STS-3c-7v
- The concatenated channels can travel different paths
  ⇒ Need buffering at the ends to equalize delay
- All channels are administered together.
  Common processing only at end-points.

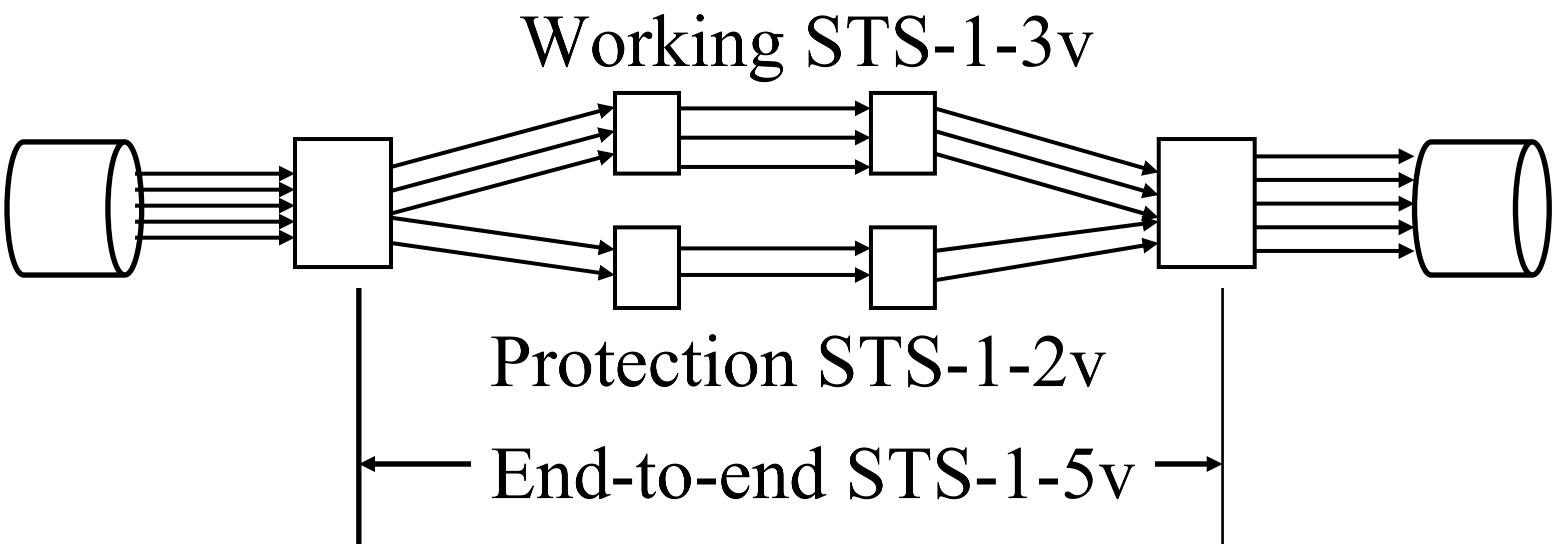
SONET LCAS



- Link Capacity Adjustment Scheme for Virtual Concatenation
- Allows hitless addition or deletion of channels from virtually concatenated SONET/SDH connections
- Control messages are exchanged between end-points to accomplish the change

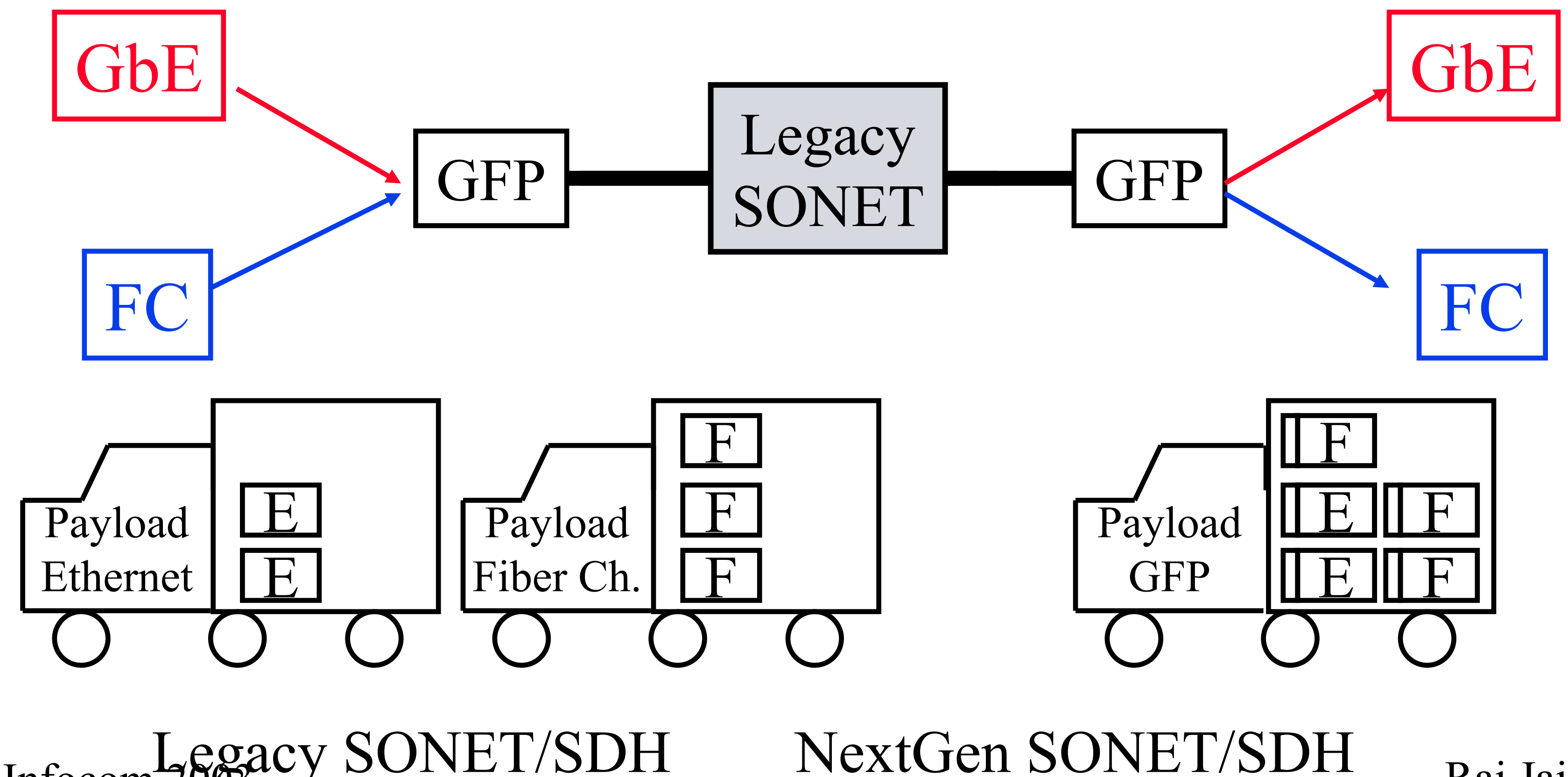
LCAS (Cont)

- Provides enhanced reliability. If some channels fail, the remaining channels can be recombined to produce a lower speed stream



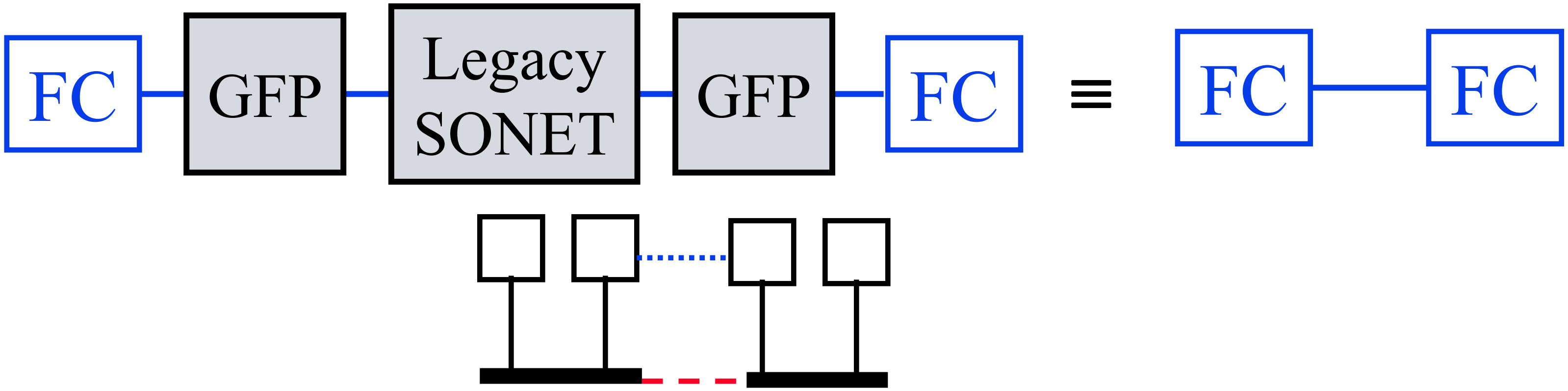
Generic Framing Procedure (GFP)

- Allows multiple payload types to be aggregated in one SONET path and delivered separately at destination



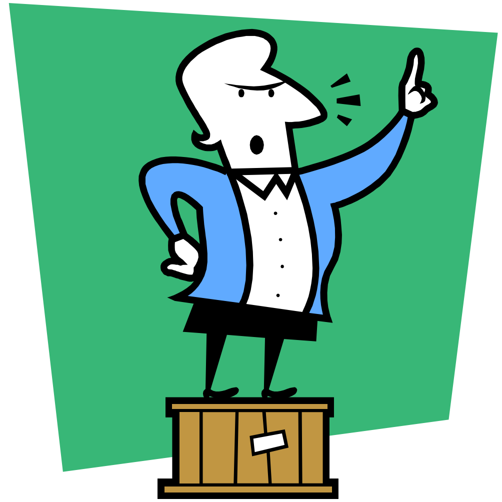
Transparent GFP

- Allows LAN/SAN PHY extension over SONET links Control codes carried as if it were a dark fiber.



- Problem: 8b/10b results in 1.25 Gb stream for 1 GbE
- Solution: Compress 80 PHY bits to 65 bits
  $\Rightarrow$ 1.02 Gbps SONET payload per GbE

Summary



- Gigabit Ethernet runs at 1000 Mbps
- 10 GbE for full duplex LAN and WAN links
- 1000 Mbps and 9,584.640 Mbps
- RPR will make it more suitable for Metro

Summary (Cont)

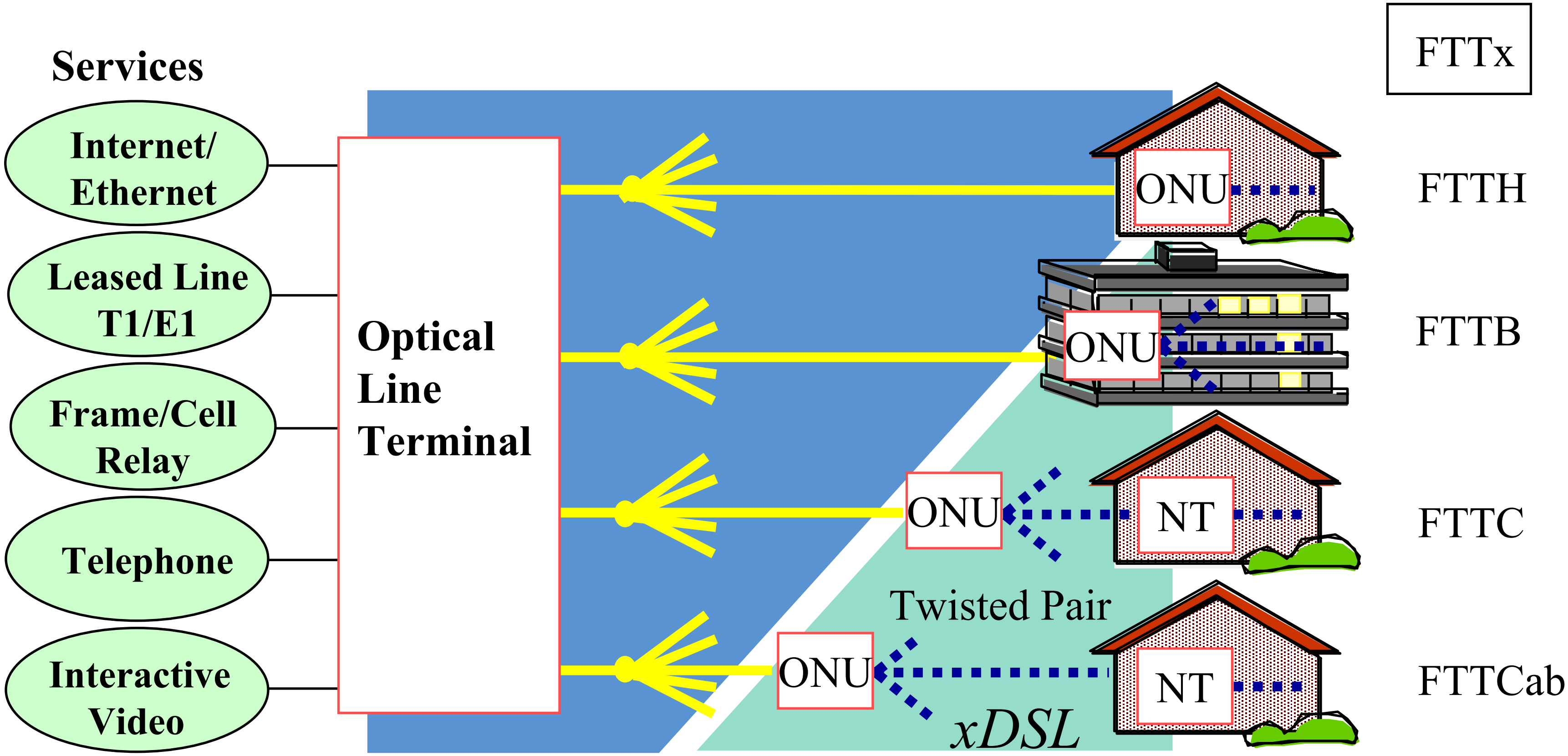
- Virtual concatenation allows a carrier to use any arbitrary number of STS-1's or T1's for a given connection. These STS-1's can take different paths.
- LCAS allows the number of STS-1's to be dynamically changed
- Frame-based GFP allows multiple packet types to share a connection
- Transparent GFP allows 8b/10 coded LANs/SANs to use PHY layer connectivity at lower bandwidth.

Passive Optical Access Networks

Raj Jain
Co-founder and Sr. VP
Nayna Networks, Inc.
481 Sycamore Dr, Milpitas, CA 95035
Email: jain@acm.org
www.nayna.com and http://www.cis.ohio-state.edu/~jain/

- Why PONs?
- PON Operation
- PON Design Issues
- APON, BPON, GPON, EPON
- PON Applications

Access: Fiber To The X(FTTx)



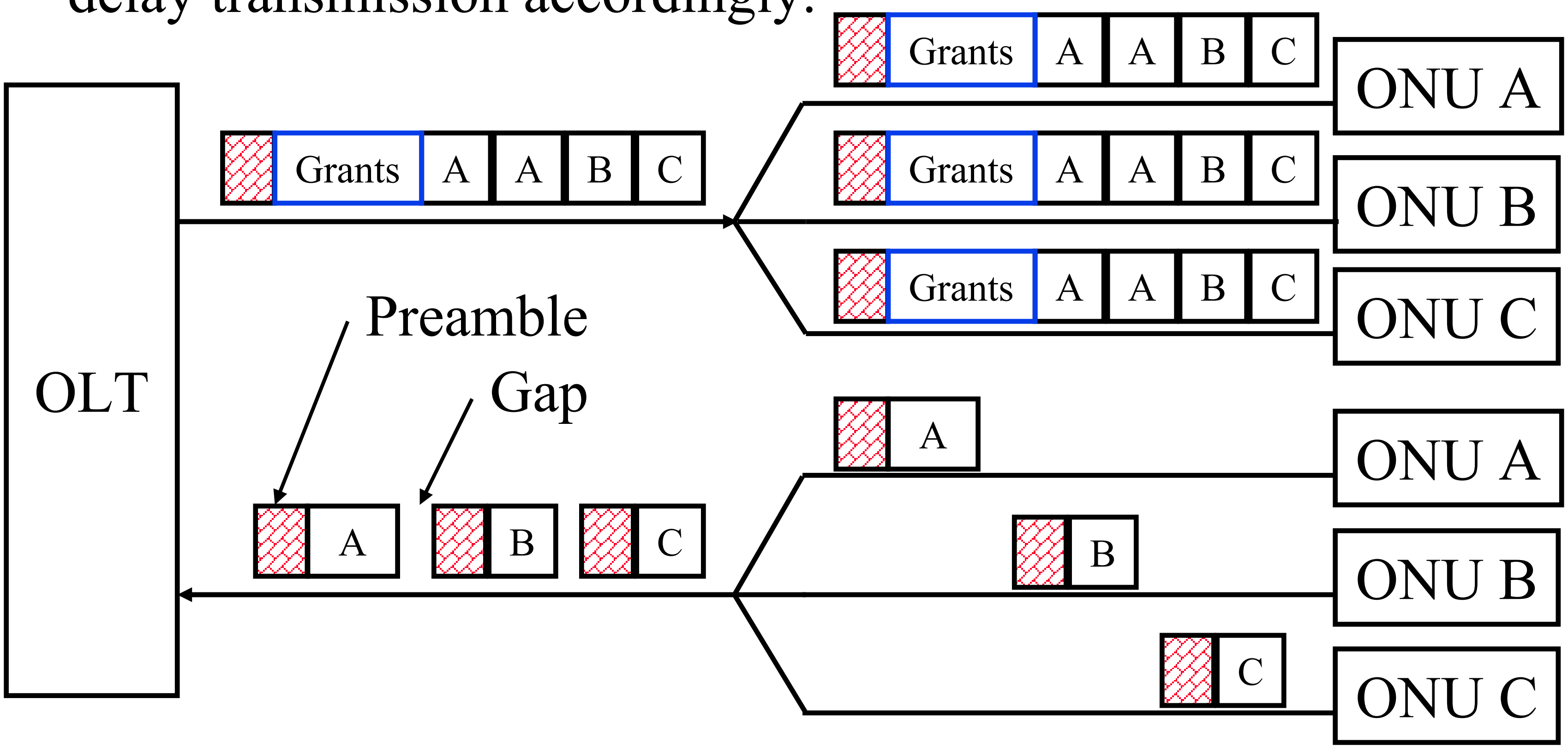
FTTH :Fiber To The Home
FTTB :Fiber To The Building
FTTC:Fiber To The Curb
FTTCab :Fiber To The Cabinet

Why PONs?

1. Passive ⇒ No active electronics or regenerators in distribution network ⇒ Very reliable. Easy to maintain. Reduced truck rolls. Shorter installation times. Reduced power expences. ⇒ Lower OpEx.
2. Single fiber for bi-directional communication ⇒ Reduced cabling and plant cost ⇒ Lower CapEx
3. A single fiber is shared among 16 to 64 customers ⇒ Relieves fiber congestion
4. Single CO equipment is shared among 16 to 64 customers 2N fibers + 2N transceivers vs 1 fiber + (N+1) transceivers ⇒ Significantly lower CapEx.
5. Scalable ⇒ New customers can be added. Exisiting Customer bandwidth can be changed
6. Multi-service: Voice, T1/E1, SONET/SDH, ATM, Video, Ethernet. Most pt-pt networks are single service.

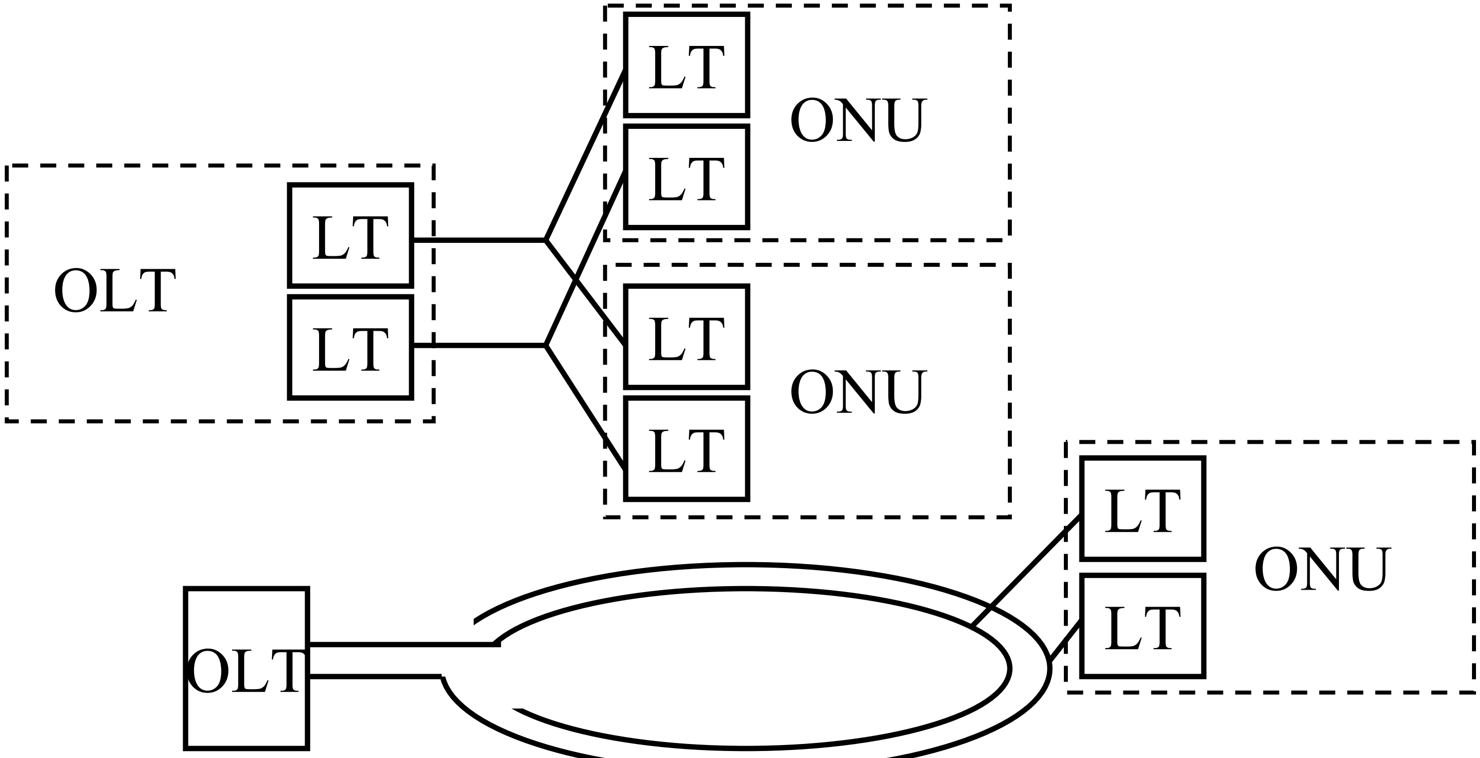
PON Operation

- All downstream traffic is broadcast
- Each ONU transmits in its allocated time slot ⇒ no collisions
- OLT accurately measures delay to each ONU and asks each to delay transmission accordingly.



Protection and Redundancy

- Redundant fibers, ONU ports, OLT ports allow operation to continue after a single failure
- Ring, tree, or bus topology



PON Design Issues

- Variable distance compensation: Ranging
- Line coding: NRZ with scrambling, 8b/10b
- Burst synchronization: Fast laser turn on/off.
- Bit/Byte alignment: Fast CDR.
- Frame alignment: Framing delimiter bits
- Framing structure
- Slot allocations: Per ONU, Per Port, Per VP
- Bandwidth allocation: Static or Dynamic
- Information integrity: CRC
- Privacy/security: Churning
- Operations Communications: OAM

PON: Types

- Full Service Access Network (FSAN) designed an ATM based PON (APON)
- APON adopted by ITU Broadband PON (BPON)
- IEEE is developing Ethernet PON (EPON)
- ITU is developing a gigabit rate PON (GPON)
- VPON = PON with Video
- HPON = Hybrid PON (Analog Video + Digital PON)

BPON

- Broadband Passive Optical Networks (BPON)
- 155 and 622 M down and 155 or 622 M up (total 3 combination. 155 Down and 622 Up not allowed)
- Split Ratio: 32
- 10 or 20 km Differential.
- Services: Ethernet, POTS, ISDN PRI, ISDN BRI, T1, E1 DS3, E3, ATM (25M, STM-1), Digital Video
- Maximum mean signal transfer delay 1.5ms

GPON

- Gigabit-Capable Passive Optical Networks (GPON)
- 1.244 and 2.488 G down and 0.155, 0.622, 1.244 and 2.488 G up (total 7 combination. 1.244 Down and 2.488 Up not allowed)
- Not compatible with BPON even at 1.244 Down/0.155 Up or 0.622 Up
- Split Ratio: 64, 128 in future
- 60 km Logical, 10 or 20 km Differential.
  10km can use FP lasers.
- Services: Ethernet, POTS, ISDN PRI, ISDN BRI, T1, E1 DS3, E3, ATM (25M, STM-1,4,16), Digital Video
- Maximum mean signal transfer delay 1.5ms

EPON

- 1 Gbps only. 8b/10 => 1.25 Gbps on the fiber.
- Covers 10 and 20 km. Min Split ratio = 1:16.
- New PX PHYs have been added:
  1000BASE-PX10: 1310 Upstream/1490 downstream, 10 km
  1000BASE-PX20: 1310 Upstream/1490 downstream, 20 km
- LT and NT port types. LT = Master. LT configures NT.
- PHYs for downstream/upstream are different
- Preamble contains logical link ID (ONU ID).
- Inter-frame gap has been extended to allow FEC
- OAM has been added. OLT can set ONT configuration.
- Unidirectional OAM ability allows transmission regardless of the link status.

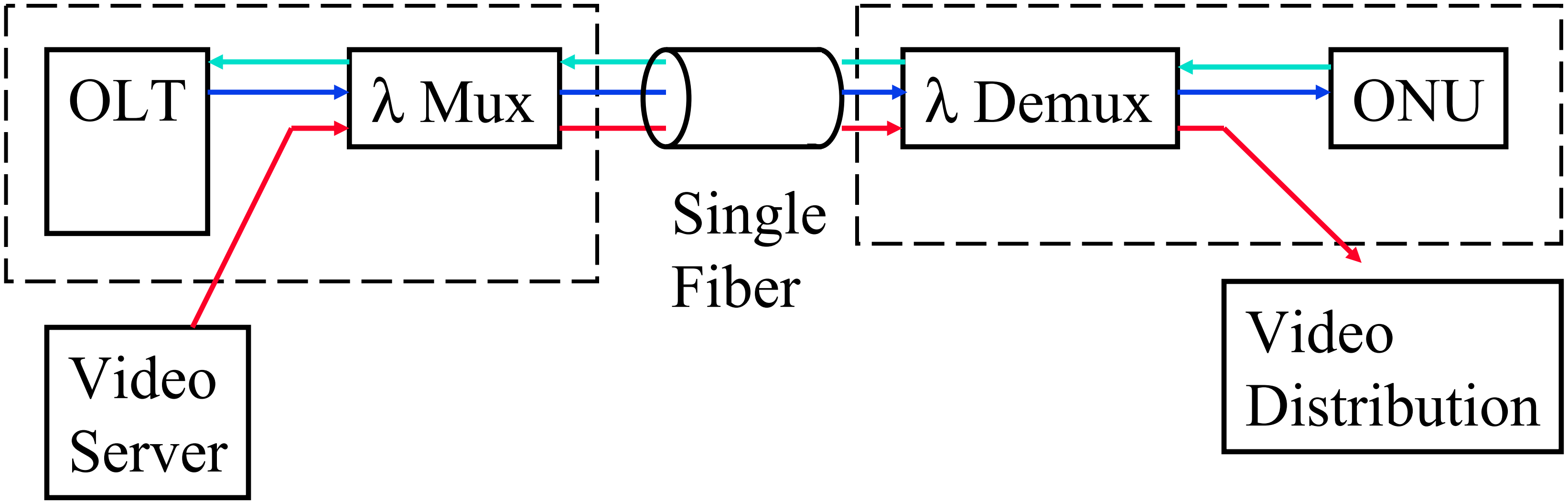
PON Applications

- Fiber to the home or Businesses
- Video Distribution
- PONs with xDSL
- PON for Cellular Backhaul Applications

Useful if customers are clustered $\Rightarrow$ PONs are succeeding in Asia (Korea, China) because of high-rise living/business

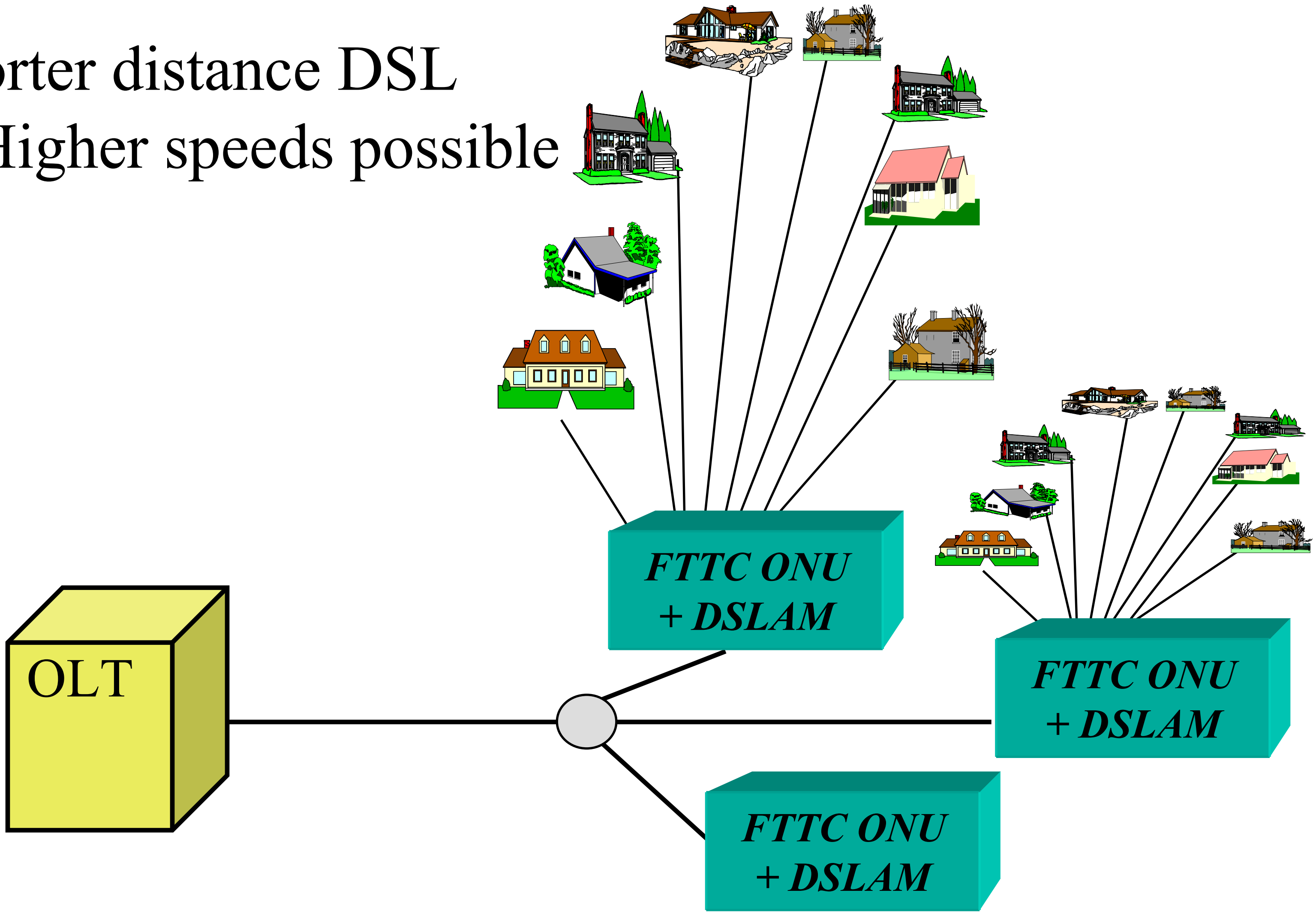
Video over PONs

- Analog and/or digital video can be sent over 3rd wavelength
- ITU has standardized 1480-1500nm for downstream data and 1540-1560 for "enhanced services"
- Video over IP feasible with Gigabit PONs



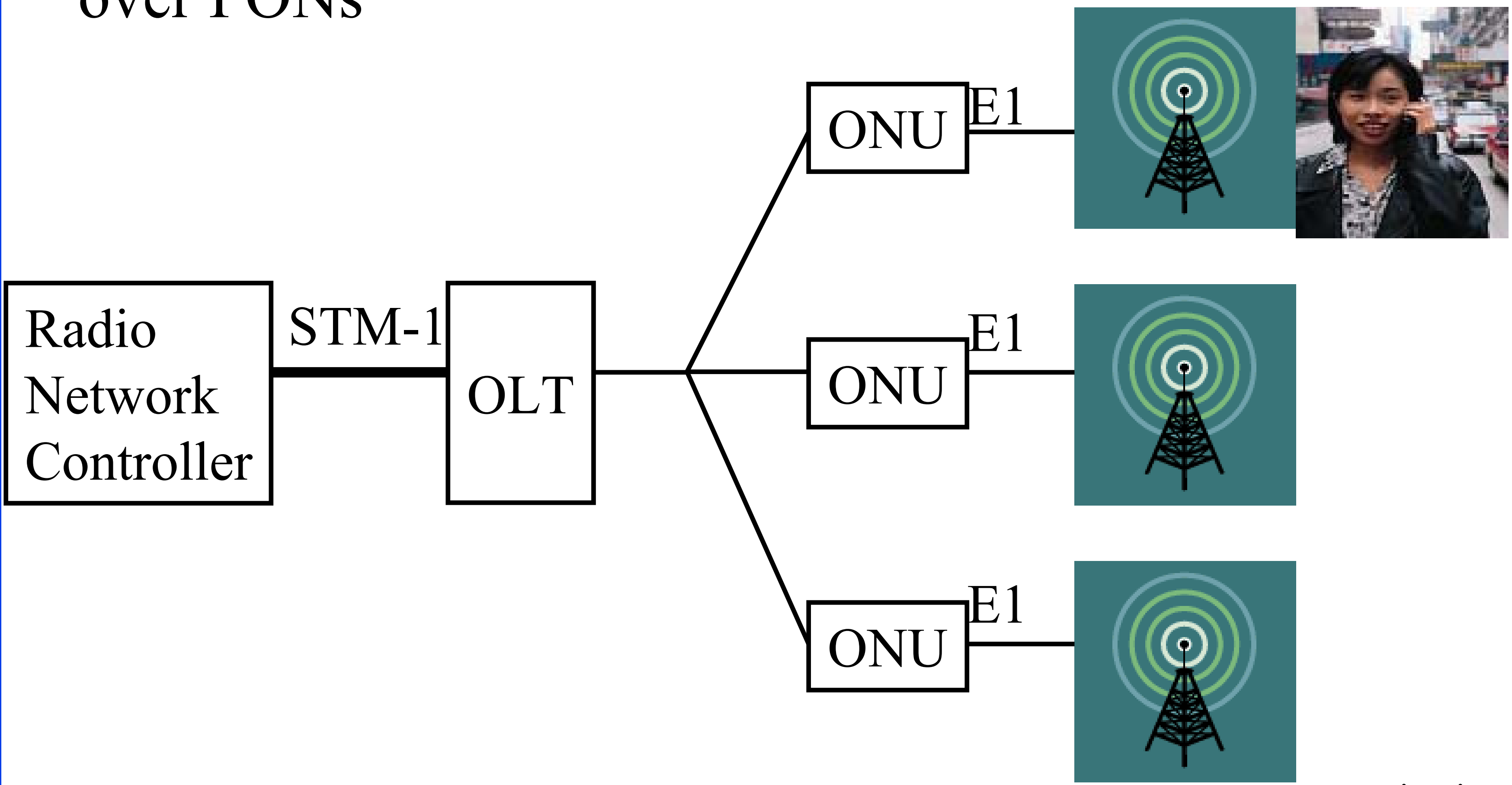
PONs With xDSL

- Shorter distance DSL
  ⇒Higher speeds possible

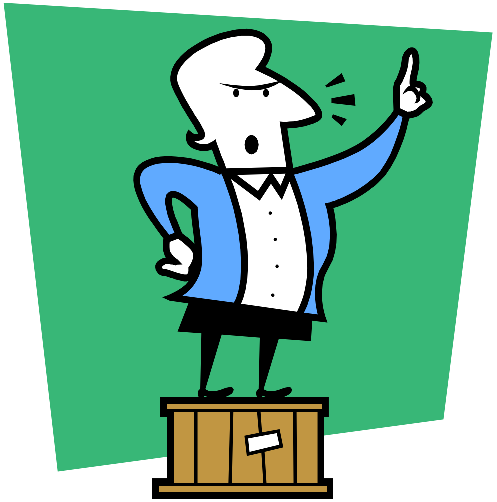


PONs for Cellular Backhaul

- Cellular antenna sites use T1/E1 which can be carried over PONs



Summary



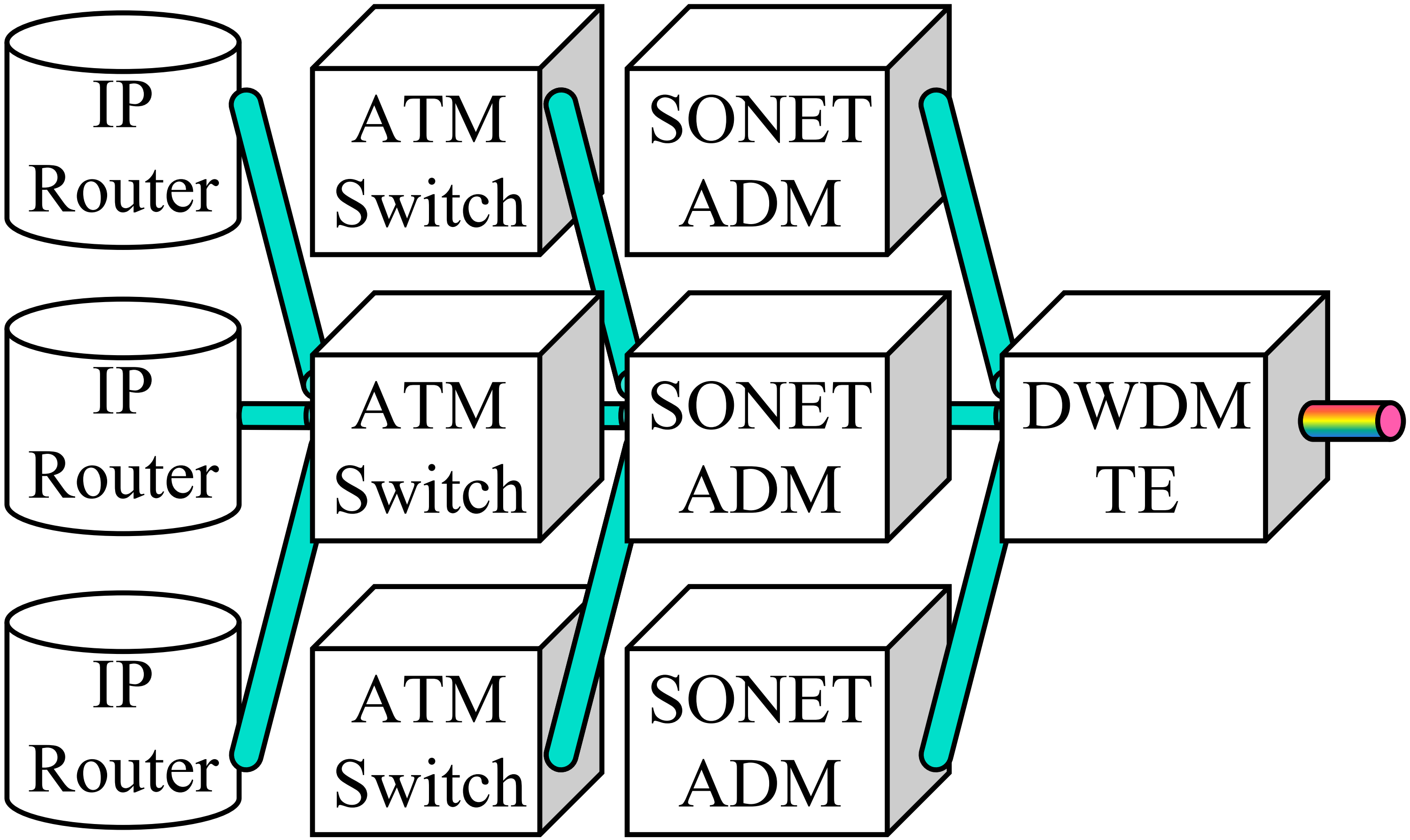
- Passive distribution network ⇒ Reliable ⇒ Lower OpEx
- Shared central office equipment ⇒ Lower CapEx
- Easy to change customer data rate or add customers
- Three Standards: BPON, GPON, and EPON
- Useful for data, video, cellular back-haul applications.

IP Over DWDM

Raj Jain
Co-founder and Sr. VP
Nayna Networks, Inc.
481 Sycamore Dr, Milpitas, CA 95035
Email: jain@acm.org
www.nayna.com and http://www.cis.ohio-state.edu/~jain/

- IP over DWDM
- UNI
- ASTN/ASON
- MPLS, MPλS, GMPLS
- Upcoming optical technologies

IP over DWDM (Past)

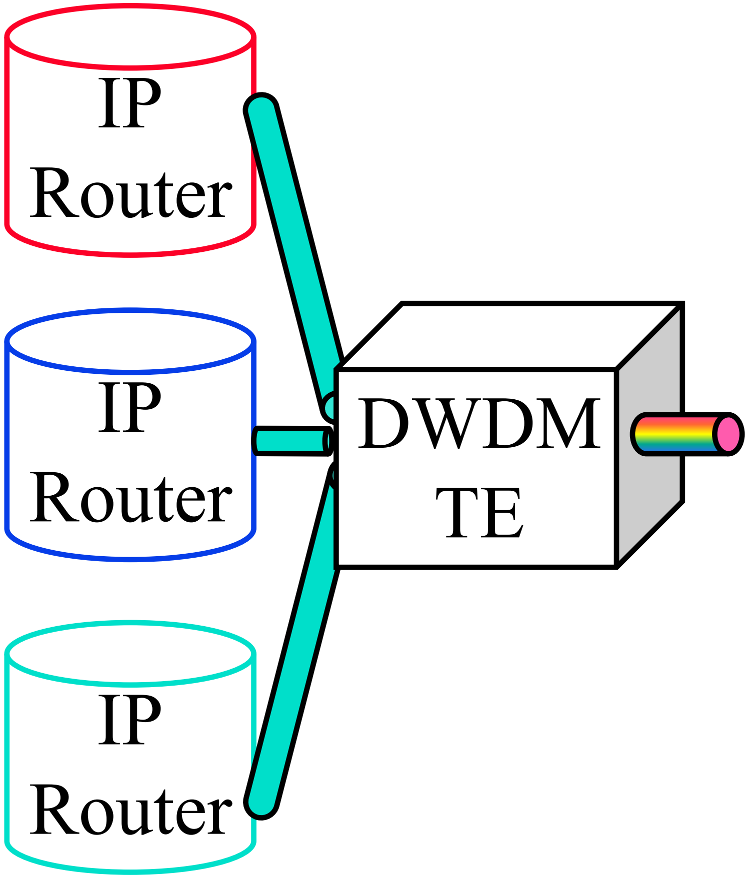


IP over DWDM: Protocol Layers

1993	1996	1999	2001	2005
IP	IP	IP/MPλS	IP/GMPLS	IP/GMPLS
ATM	PPP	PPP	Ethernet	Ethernet
SONET	SONET	SONET Framing	SONET Framing	
DWDM	DWDM	DWDM	DWDM	DWDM
Fiber	Fiber	Fiber	Fiber	Fiber

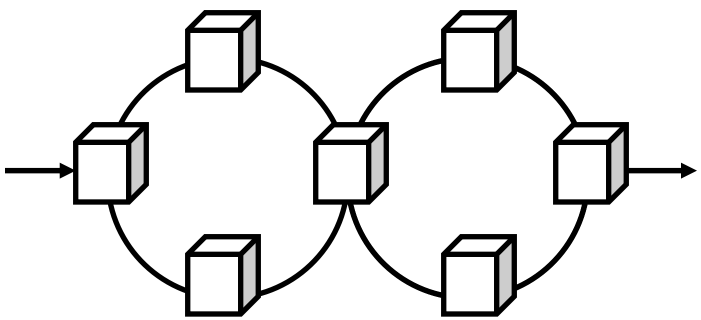
- IP is good for routing, traffic aggregation, resiliency
- ATM for multi-service integration, QoS/signaling
- SONET for traffic grooming, monitoring, protection
- DWDM for capacity
- Problem: Restoration in multiple layers, Sonet Manual
  ⇒ Intersection of features and union of problems

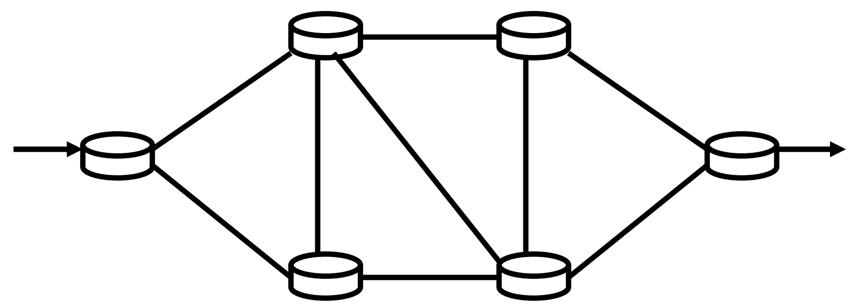
IP over DWDM (Future)



Telecom vs Data Networks

	Telecom Networks	Data Networks
Topology Discovery	Manual	Automatic
Path Determination	Manual	Automatic
Circuit Provisioning	Manual	No Circuits
Transport & Control Planes	Separate	Mixed
User and Provider Trust	No	Yes
Protection	Static using Rings	No Protection



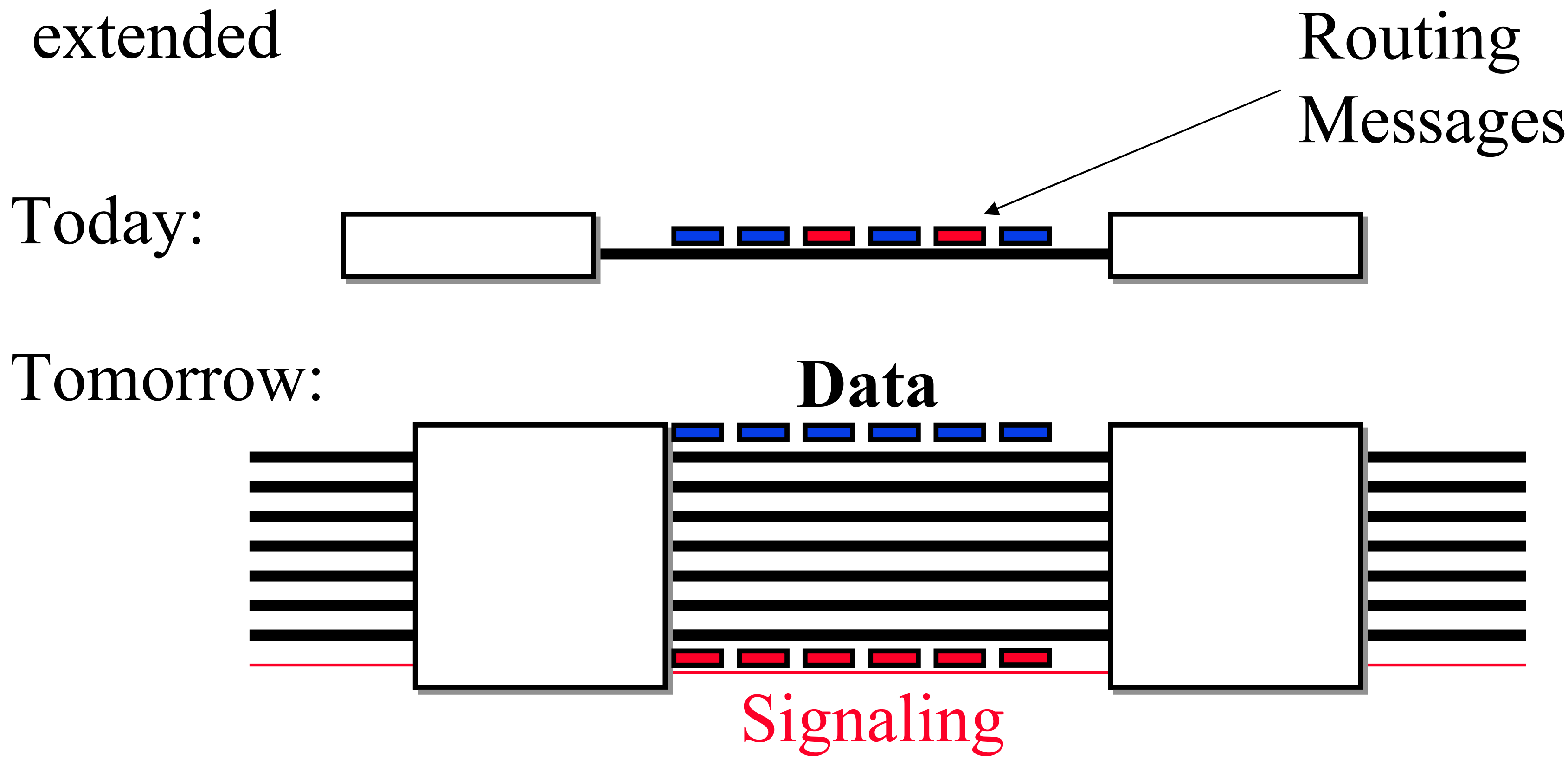


IP over DWDM Issues

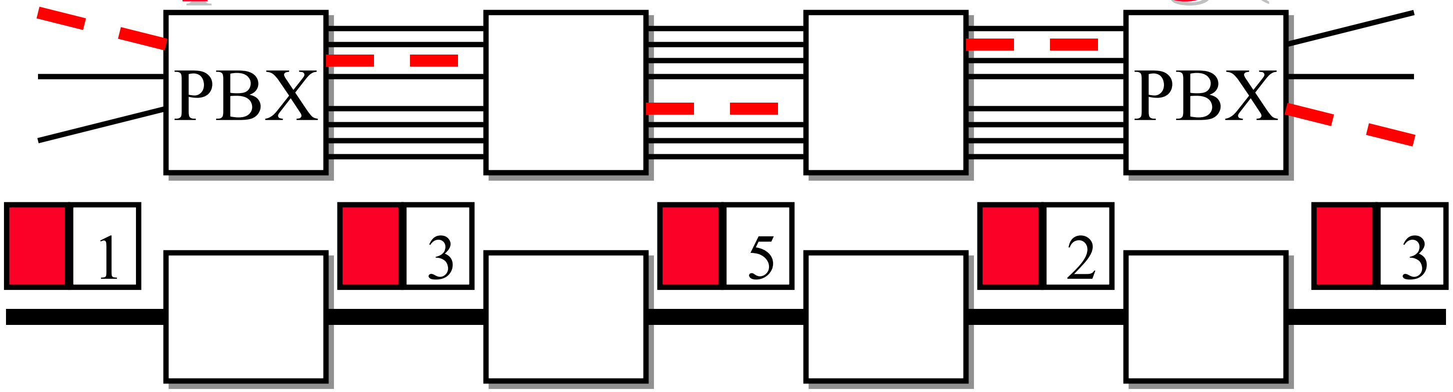
1. Data and Control plane separation
2. Circuits
3. Signaling
4. Addressing
5. Protection and Restoration

Issue: Control and Data Plane Separation

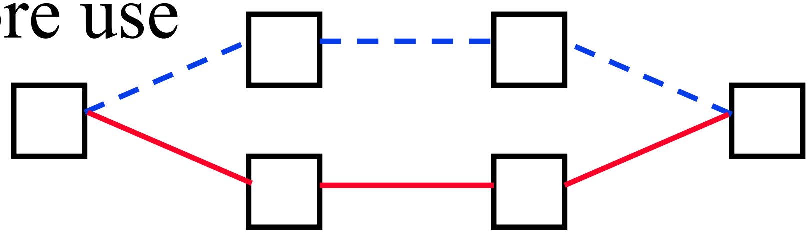
- Separate control and data channels
- IP routing protocols (OSPF and IS-IS) are being extended



Multiprotocol Label Switching (MPLS)

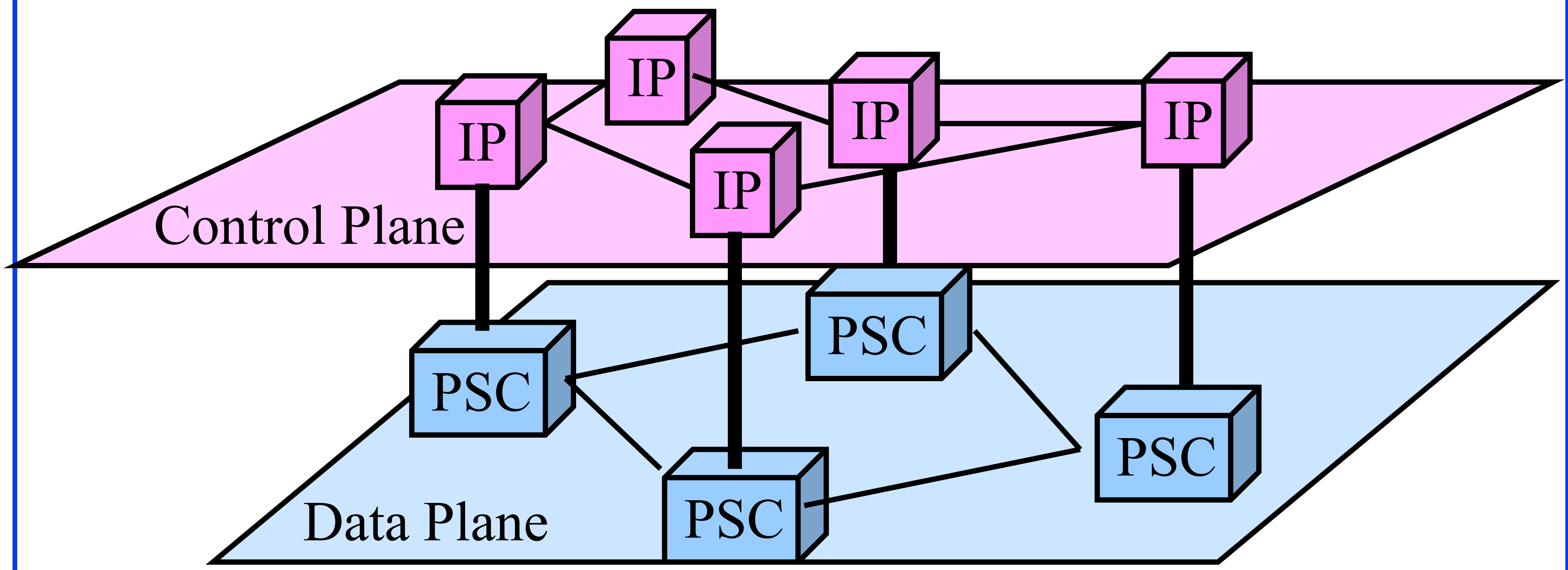


- Allows virtual circuits in IP Networks (May 1996)
- Each packet has a virtual circuit number called 'label'
- Label determines the packet's queuing and forwarding
- Circuits are called Label Switched Paths (LSPs)
- LSP's have to be set up before use
- Allows traffic engineering



IP-Based Control Plane

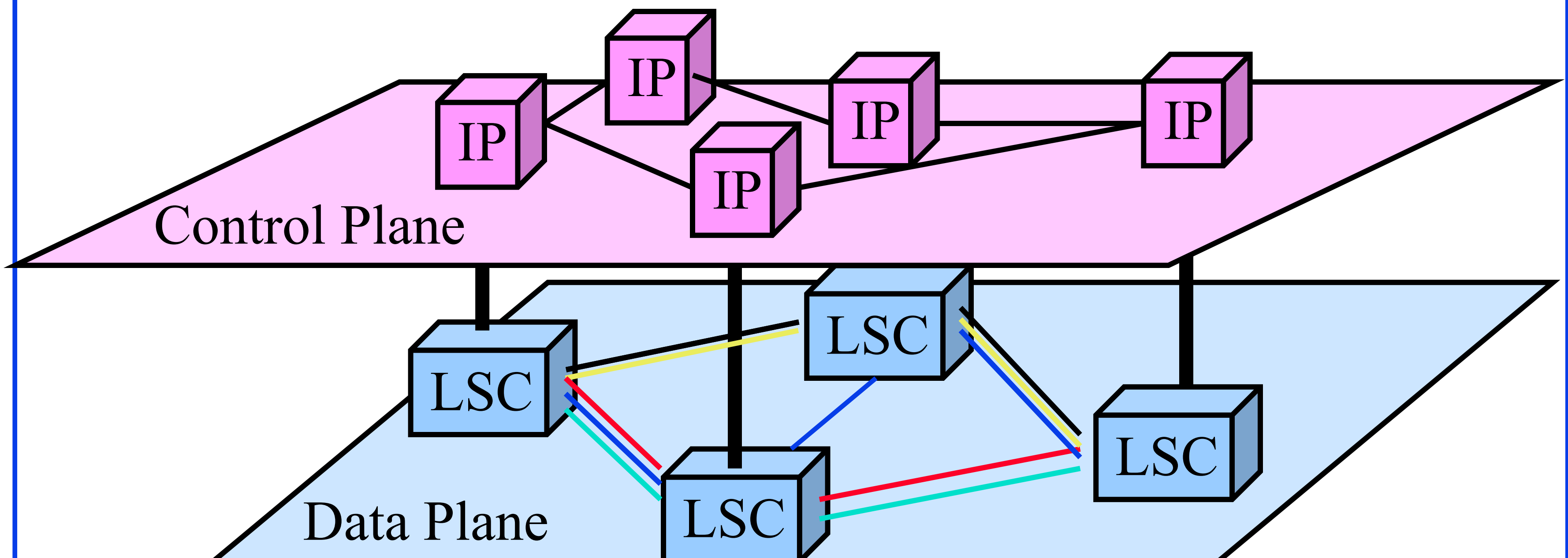
- Control is by IP packets (electronic).
  Data can be any kind of packets (IPX, ATM cells).
  ⇒ MPLS



PSC = Packet Switch Capable Nodes

MPλS

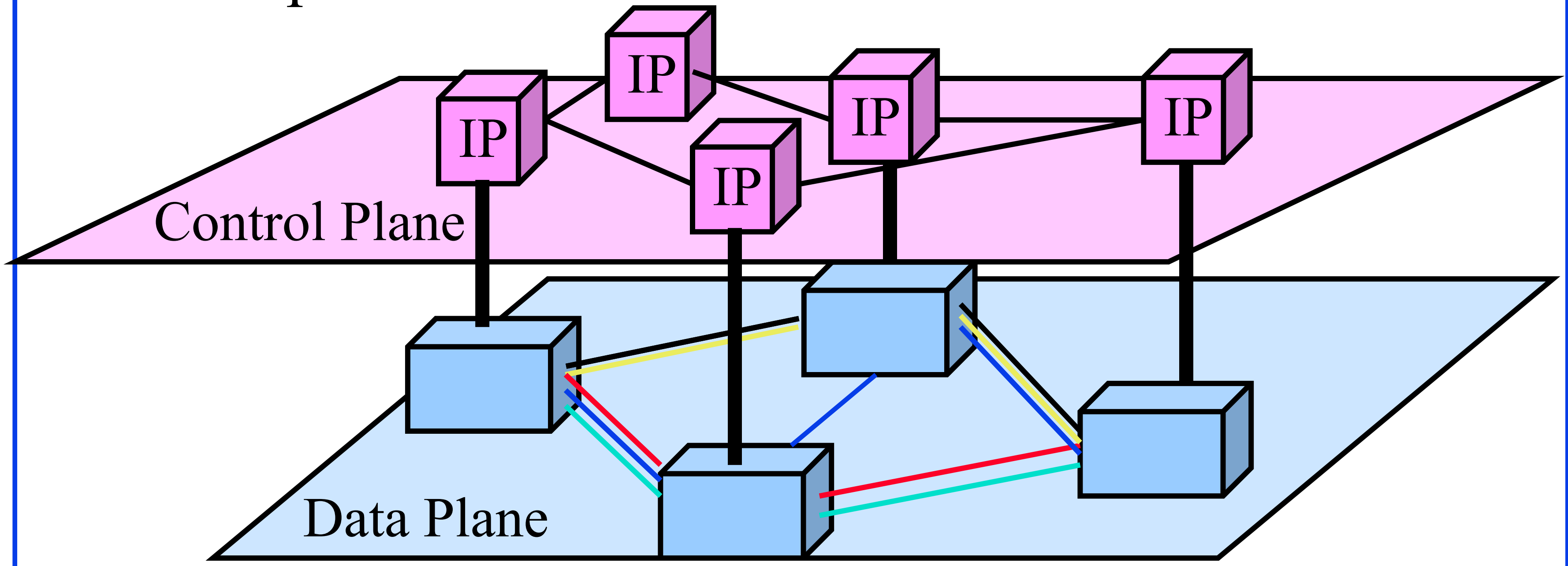
- Control is by IP packets (electronic).
  Data plane consists of wavelength circuits
  ⇒ Multiprotocol Lambda Switching (October 1999)



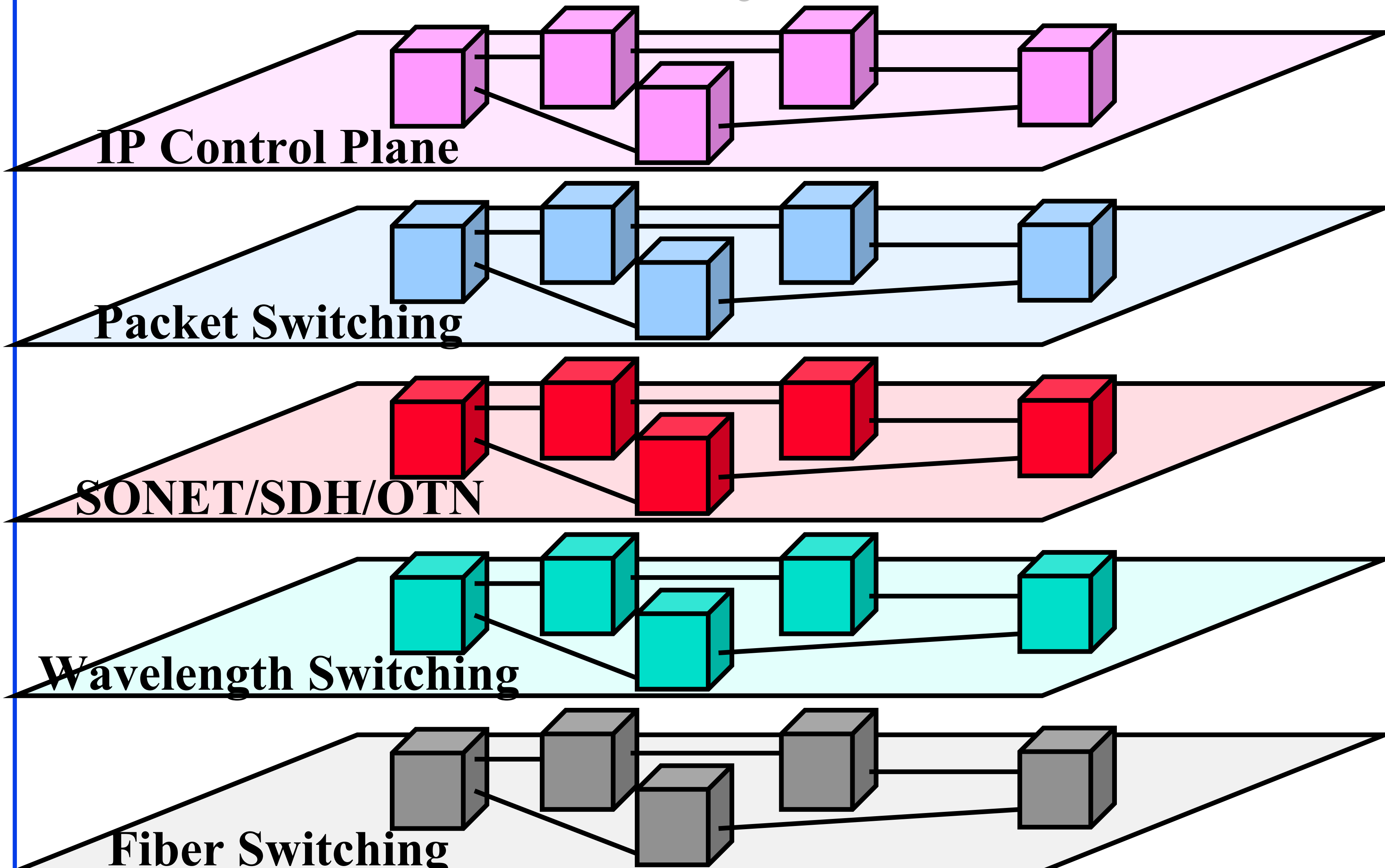
LSC = Lambda Switch Capable Nodes
= Optical Cross Connects = OXC

GMPLS

- Data Plane = Wavelengths, Fibers, SONET Frames, Packets (October 2000)
- Two separate routes: Data route and control route

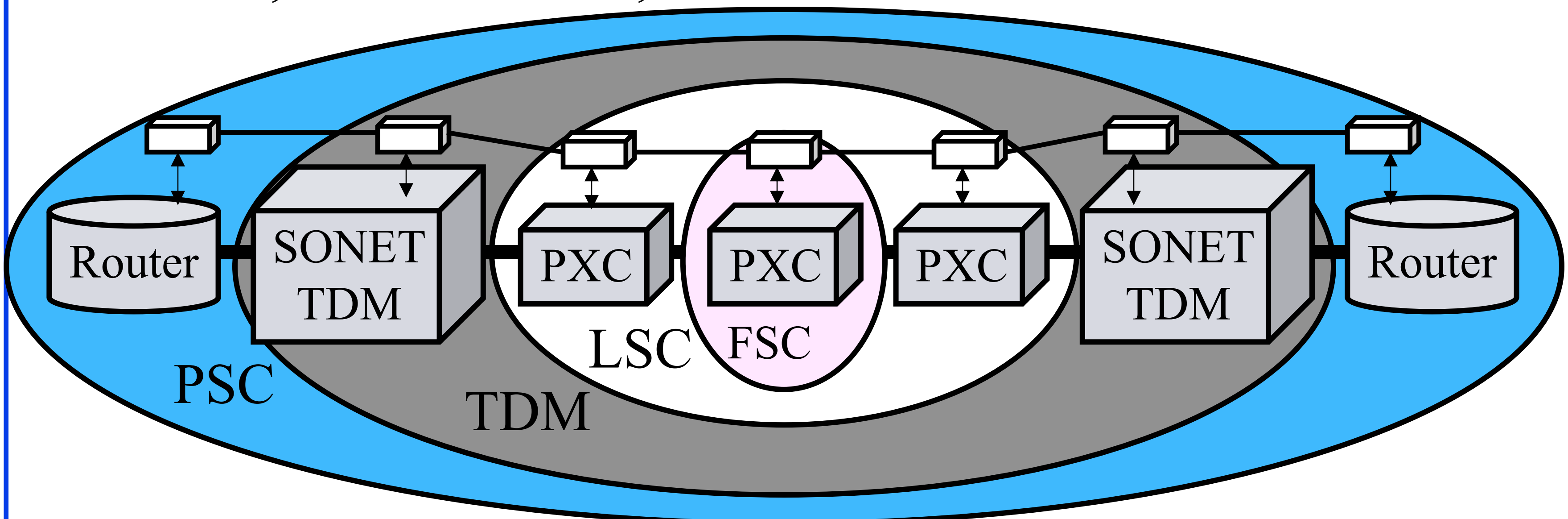


GMPLS: Layered View



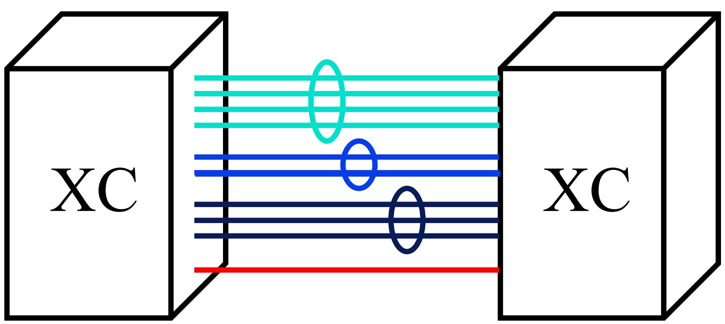
GMPLS: Hierarchical View

- Packets over SONET over Wavelengths over Fibers
- Packet switching regions, TDM regions, Wavelength switching regions, fiber switching regions
- Allows data plane connections between SONET ADMs, PXCs. FSCs, in addition to routers



MPLS vs GMPLS

Issue	MPLS	GMPLS
Data & Control Plane	Same channel	Separate
Types of Nodes and labels	Packet Switching	PSC, TDM, LSC, FSC, …
Bandwidth	Continuous	Discrete: OC-n, λ's, ..
# of Parallel Links	Small	100-1000's
Port IP Address	One per port	Unnumbered
Fault Detection	In-band	Out-of-band or In-Band



Draft Martini

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

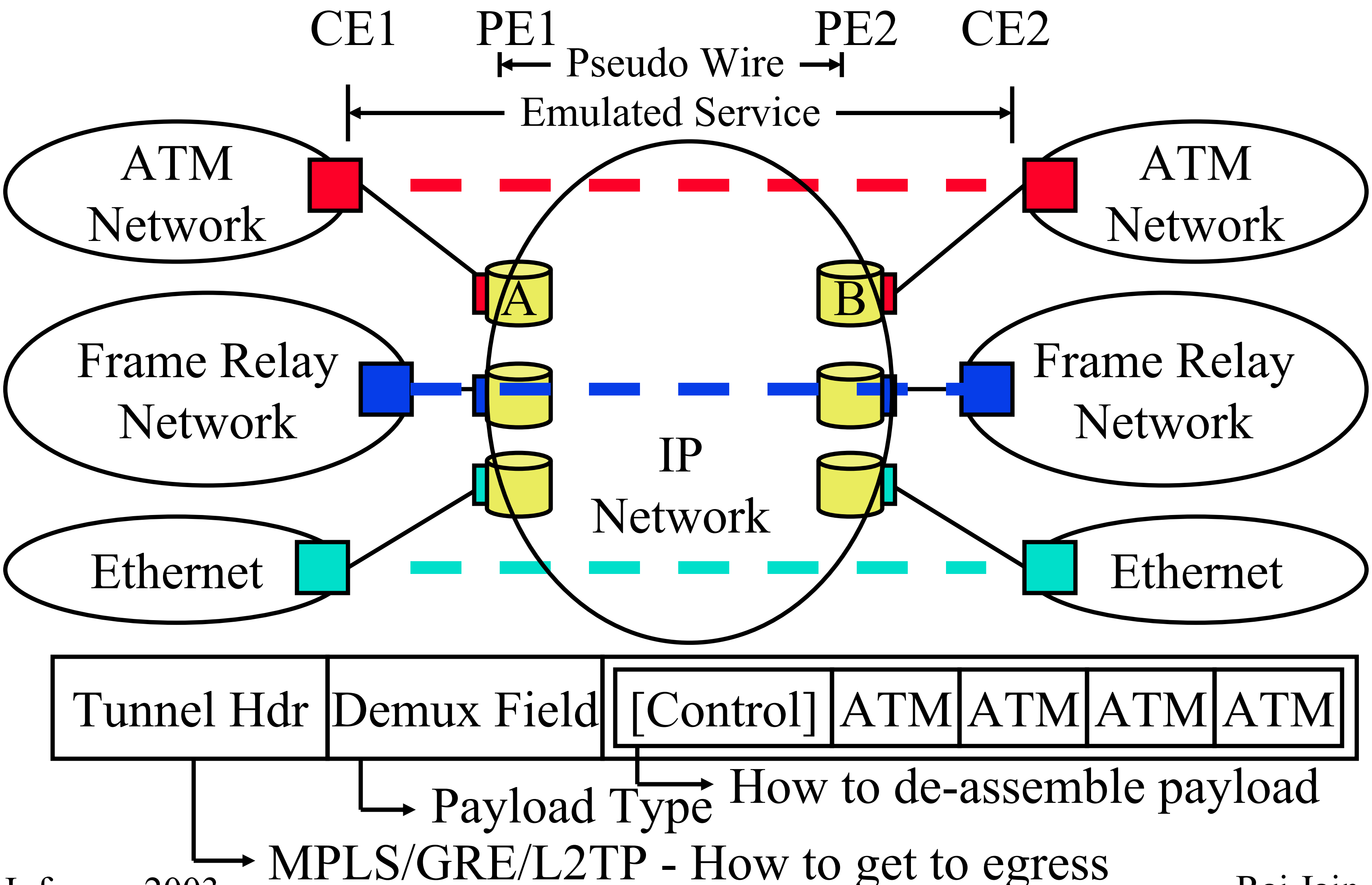
IP		
Ethernet	ATM	PPP

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

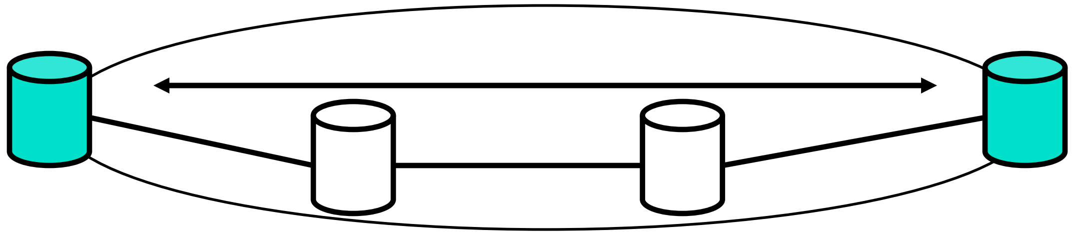
Ethernet	ATM	PPP
IP		

- Ref: draft-ietf-pwe3-*.txt

L2 Circuits over IP



PW Label



- PW Label bindings distributed using LDP downstream unsolicited mode between ingress and egress LSRs
- Circuit specific parameters such as MTU, options are exchanged at the time VC Label exchange
- PW Label: S=1 $\Rightarrow$ Bottom of stack, TTL=2
- PW Type:

1 Frame Relay DLCI	6 HDLC
2 ATM AAL5 VCC Transport	7 PPP
3 ATM Transparent Cell Transport	8 Circuit Emulation
4 Ethernet VLAN	9 ATM VCC Cell Transport
5 Ethernet	10 ATM VPC Cell Transport

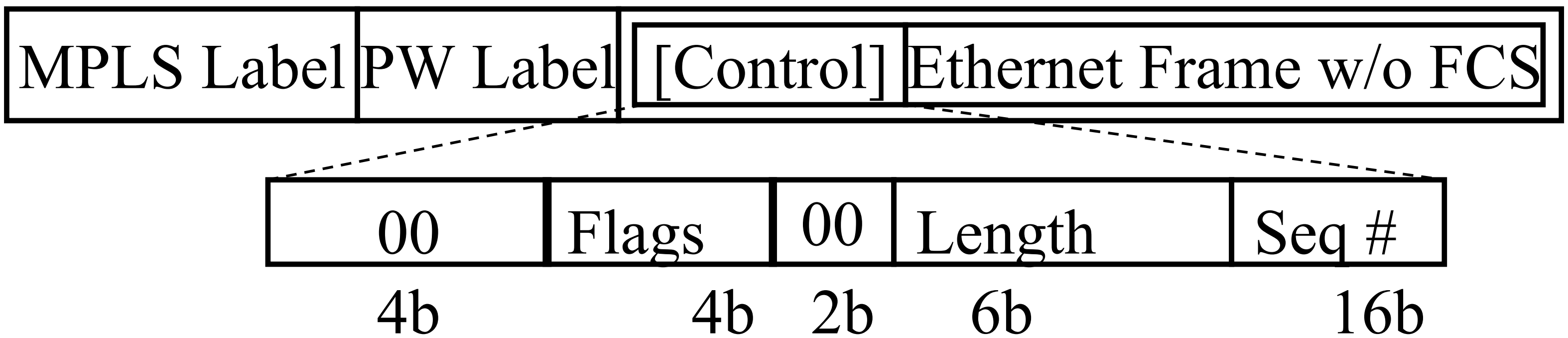
ATM over MPLS

MPLS Label	PW Label	[Control]	ATM	ATM	ATM	ATM

MPLS Label	PW Label	Control	ATM AAL5 SDU

- Multiple VCCs in the PW: HEC is stripped. 52B/Cell.
- Only one VCC in the PW: VPI/VCI are stripped. 48 bytes of payload + 1 byte containing PTI, CLP = 49B/Cell
- Only one VPC in the PW: VPI is stripped. 48 bytes of payload + 3 bytes of VCI, PTI, CLP = 51B/Cell
- Control word is optional in above cases
- AAL5 SDU Mode: control word indicates length and other info
- Ref: draft-ietf-pwe3-atm-encap-01.txt, Feb 2003

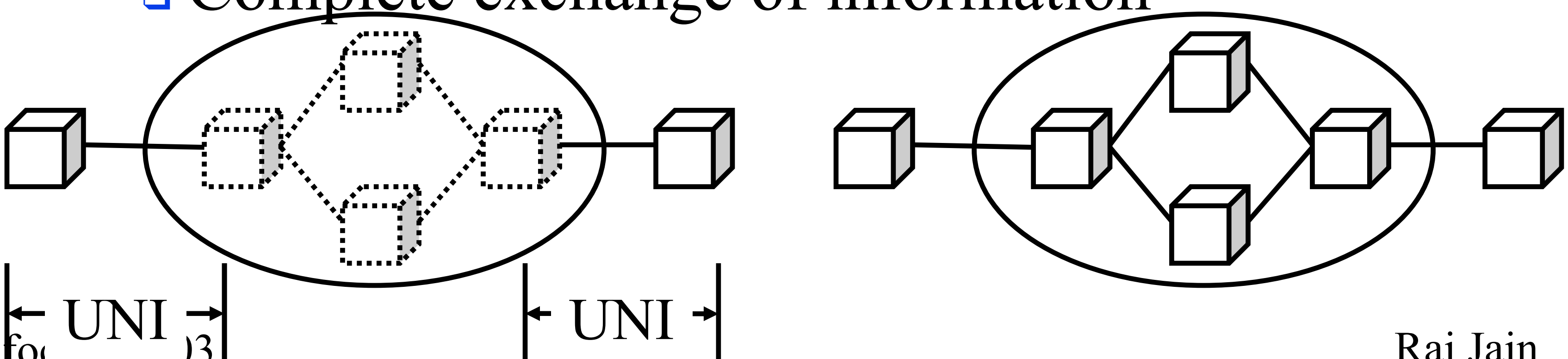
Ethernet over MPLS



- Control word is optional
- Flags are not used
- May put 802.1p priority in exp field of MPLS label
- Frame ordering is optional. If enabled, all out-of-order frames are discarded at exit.
- Pause frames are obeyed locally. Not transported.
- Ref: draft-ietf-pwe3-ethernet-encap-02.txt

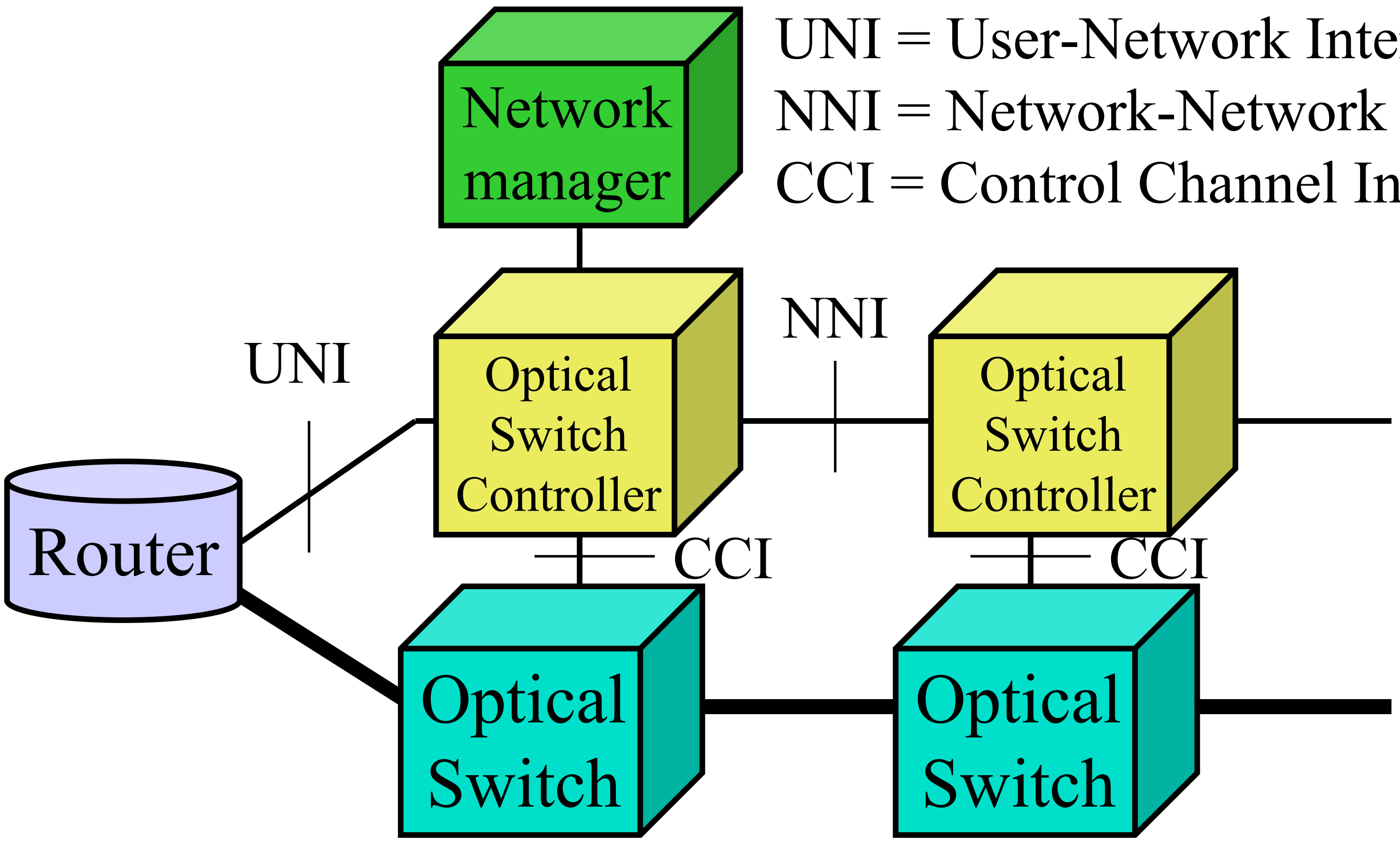
Issue: UNI vs Peer-to-Peer Signaling

- Two Business Models:
  - Carrier: Overlay or cloud
    - Network is a black-box
    - User-to-network interface (UNI) to create/destroy light paths (in OIF)
  - Enterprise: Peer-to-Peer
    - Complete exchange of information



ASTN/ASON

- Automatically Switched Transport Networks

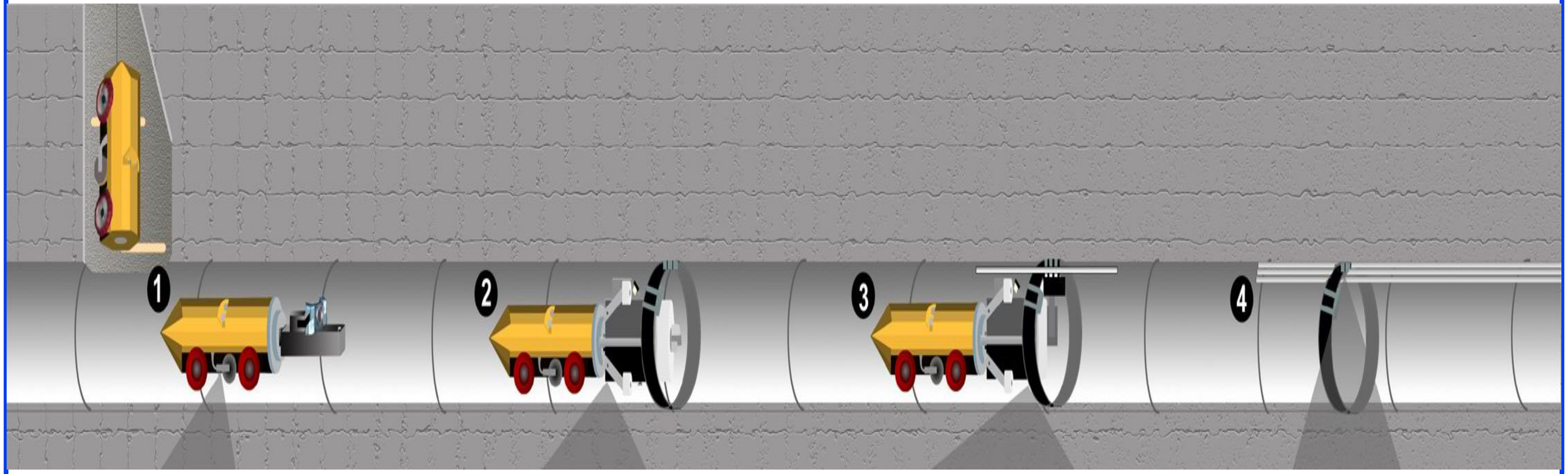


UNI = User-Network Interface
NNI = Network-Network Interface
CCI = Control Channel Interface

Fiber Access Thru Sewer Tubes (FAST)

- Right of ways is difficult in dense urban areas
- Sewer Network: Completely connected system of pipes connecting every home and office
- Municipal Governments find it easier and more profitable to let you use sewer than dig street
- Installed in Zurich, Omaha, Albuquerque, Indianapolis, Vienna, Ft Worth, Scottsdale, ...
- Corrosion resistant inner ducts containing up to 216 fibers are mounted within sewer pipe using a robot called Sewer Access Module (SAM)
- Ref: http://www.citynettelecom.com, NFOEC 2001, pp. 331

FAST Installation



1. Robots map the pipe
2. Install rings
3. Install ducts
4. Thread fibers

Fast Restoration: Broken sewer pipes replaced with minimal disruption

Summary

1. High speed routers
   ⇒ IP directly over DWDM
2. Separation of control and data plane
   ⇒ IP-Based control plane
3. Transport Plane = Packets ⇒ MPLS
   Transport Plane = Wavelengths
   ⇒ MPλS
   Transport Plane = λ, SONET, Packets
   ⇒ GMPLS
4. UNI allows users to setup paths on demand

Standards Organizations

- IETF: www.ietf.org
  - Multiprotocol Label Switching (MPLS)
  - IP over Optical (IPO)
  - Traffic Engineering (TE)
  - Common Control and Management Plane (CCAMP)
- Optical Internetworking Forum (OIF): www.oiforum.com
- ANSI T1X1.5: http://www.t1.org/t1x1/_x15-hm.htm
- ITU, www.itu.ch, Study Group 15 Question 14 and Question 12
- Optical Domain Service Interface (ODSI) - Completed December 2000

References

- Detailed 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 Optical: A summary of issues, (internet draft) http://www.cis.ohio-state.edu/~jain/ietf/issues.html
- Lightreading, http://www.lightreading.com