Hot Topics in Networking

GMPLS
IP Switching
10 Gigabit Ethernet
Voice over IP
DWDM
VPNs
RSVP
Differentiated Services

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

What This Course is Not About?

- Hot (Adult) news groups on Internet
- Not for people who have no networking background This course does assume some familiarity with networking.
- It covers networking architectures rather than applications.

1. Networking Trends
2. QoS over Data Networks
3. Label Switching
4. Gigabit and 10 Gb Ethernet
5. IP Over DWDM
6. Voice over IP
7. Virtual Private Networks

1. Networking Trends

- Networking Bottleneck
- Traffic and Capacity growth
- Life Cycles of Technologies
- Laws of Technology Growth
- Networking and Telecom Convergence
- ATM: Why some technologies fail/succeed?
- Everything over IP
- LAN WAN Convergence: Ethernet Everywhere

2. QoS over Data Networks

- ATM QoS and Issues
- Integrated Services and RSVP
- Differentiated Services:
  Expedited and Assured Forwarding
- Subnet Bandwidth Manager (SBM)
- COPS Protocol for Policy
- IEEE 802.1D Model
- Comparison of QoS Approaches

3. Label Switching

- Routing vs Switching
- Tag Switching
- Multi-Protocol Label Switching
- Label Stacks
- Label Distribution Protocols: LDP, CR-LDP, RSVP-TE
- Traffic Engineering using MPLS

4. Gigabit and 10 Gb Ethernet

- Distance-B/W Principle
- Gigabit MAC issues: Carrier Extension, Frame Bursting
- 10 GbE: Key Features, PMD Types
- 1G/10G Ethernet Switch Features
- Flow Control, Link Aggregation, Jumbo Frames
- Resilient Packet Rings
- Beyond 10 GbE

5. IP over DWDM

- Recent DWDM Records and Product Announcements
- 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

6. Voice over IP

- Voice over IP: Why?
- Sample Products and Services
- 13 Technical Issues, 4 Other Issues
- H.323 Standard and Session Initiation Protocol (SIP)
- Media Gateway Control Protocol (MGCP) and Megaco
- Stream Control Transmission Protocol (SCTP)

7. Virtual Private Networks

- Types of VPNs
- When and why VPN?
- VPN Design Issues
- Security Issues
- VPN Examples: PPTP, L2TP, IPSec
- Authentication Servers: RADIUS and DIAMETER
- VPNs using MPLS, Virtual Routers

Day 1 (Tentative)

8:30- 9:15	Course Introduction
9:15-10:15	Trends I
10:15-10:30	*Coffee Break*
10:30-11:45	Trends II
11:45-12:00	QoS over data networks I
12:00 - 1:00	*Lunch Break*
1:00 - 2:15	QoS over data networks II
2:15 - 2:30	Tag Switching + MPLS I
2:30 - 2:45	*Coffee Break*
2:45 - 3:00	MPLS II
3:00 - 4:30	Gigabit and 10 Gb Ethernet

Day 2

8:30-10:00	IP over DWDM I
10:15-10:30	*Coffee Break*
10:30-11:15	IP Over DWDM II
11:15-12:00	Voice over IP I
12:00 - 1:00	*Lunch Break*
1:00 - 2:15	Voice over IP II
2:15 - 2:30	Virtual Private Networks I
2:30 - 2:45	*Coffee Break*
2:45 - 4:00	Virtual Private Networks II
4:00 - 4:30	Final Review

Announcements

- Materials: Copies of all information slides, A totebag
- Tutorial Changes: Until 1:00 PM today
  Go to the N+I Customer service/registration desk
- Evaluation:
  Forms will be distributed tomorrow during lunch
  Return to door monitor. Please do not fold.
- Certificates: Email requests to NIcert@key3media.com or fax to 650-578-6864
- Lunch will be provided
  Do not leave materials or personal belongings

References

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

Pre-Test

Check if you know the difference between:

- SONET and Ethernet Frame Format
- Moore's Law and Metcalf's Law
- Guaranteed quality and controlled load services
- Integrated vs Differentiated Services
- Expedited forwarding vs Assured Forwarding
- Tag Switching and Label Switching
- Min packet sizes on 10Base-T and 1000Base-T
- Token Ring and Resilient Packet Ring

Raj Jain

Pre-Test (Cont)

- ❑ Optical and All-Optical
- ❑ MPLS and GMPLS
- ❑ Soliton and Photon
- ❑ H.323 and Session Initiation Protocol
- ❑ Gatekeeper and Gateway
- ❑ Firewall and proxy server
- ❑ Digital signature and Digital Certificate
- ❑ Private Key and Public Key encryption

Number of items checked _______

- If you checked more than 8 items,
  you may not gain much from this course.
- 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 CTO
Nayna Networks, Inc.
481 Sycamore Dr, Milpitas, CA 95035
Email: raj@nayna.com
www.nayna.com
and http://www.cis.ohio-state.edu/~jain/

Future

White House Astrologer

Joan Quigly

All I want you to tell me is what will be the networking technology in the year 2002.

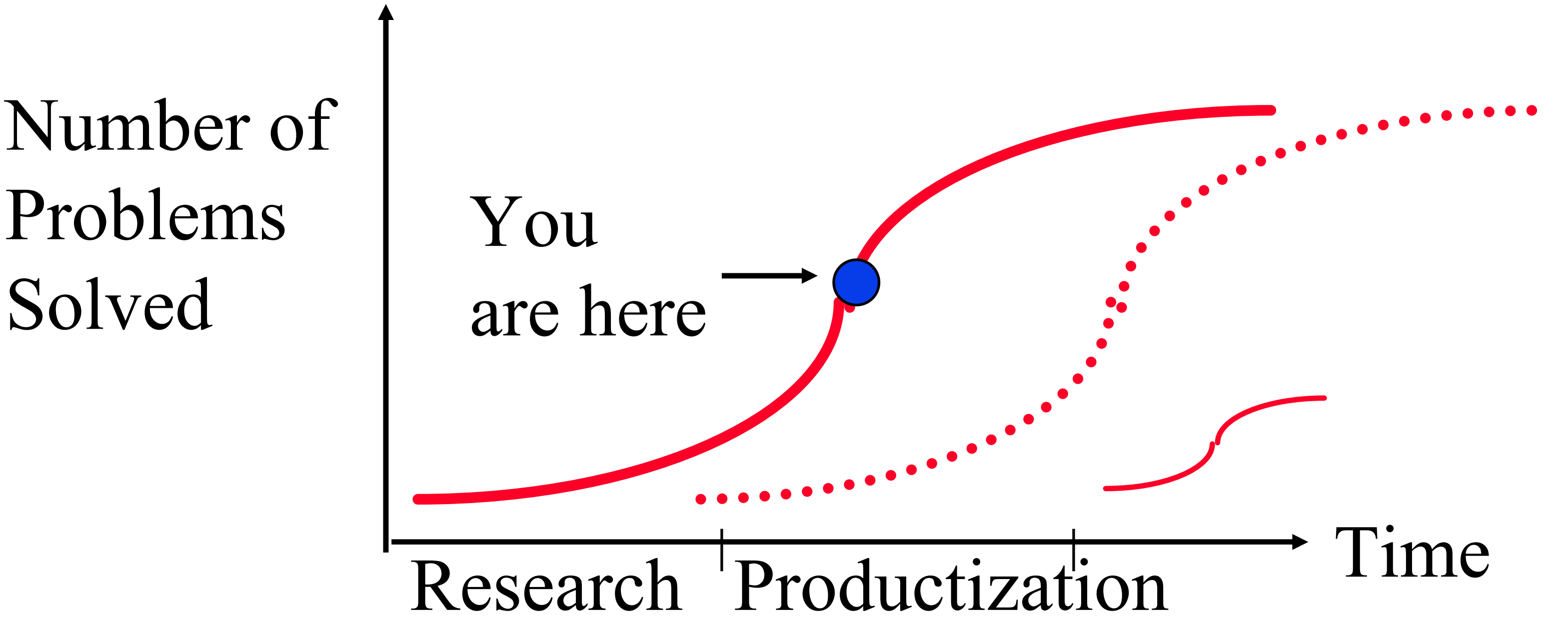
Raj Jain

Predictions

"And of all variations of multimedia, the one that will drive ATM is personal computers videoconferencing - interactive, two-way, real-time, integrated digital voice, video and data. Ethernet will remain as a legacy LAN." - Robert Metcalfe, Inventor of Ethernet

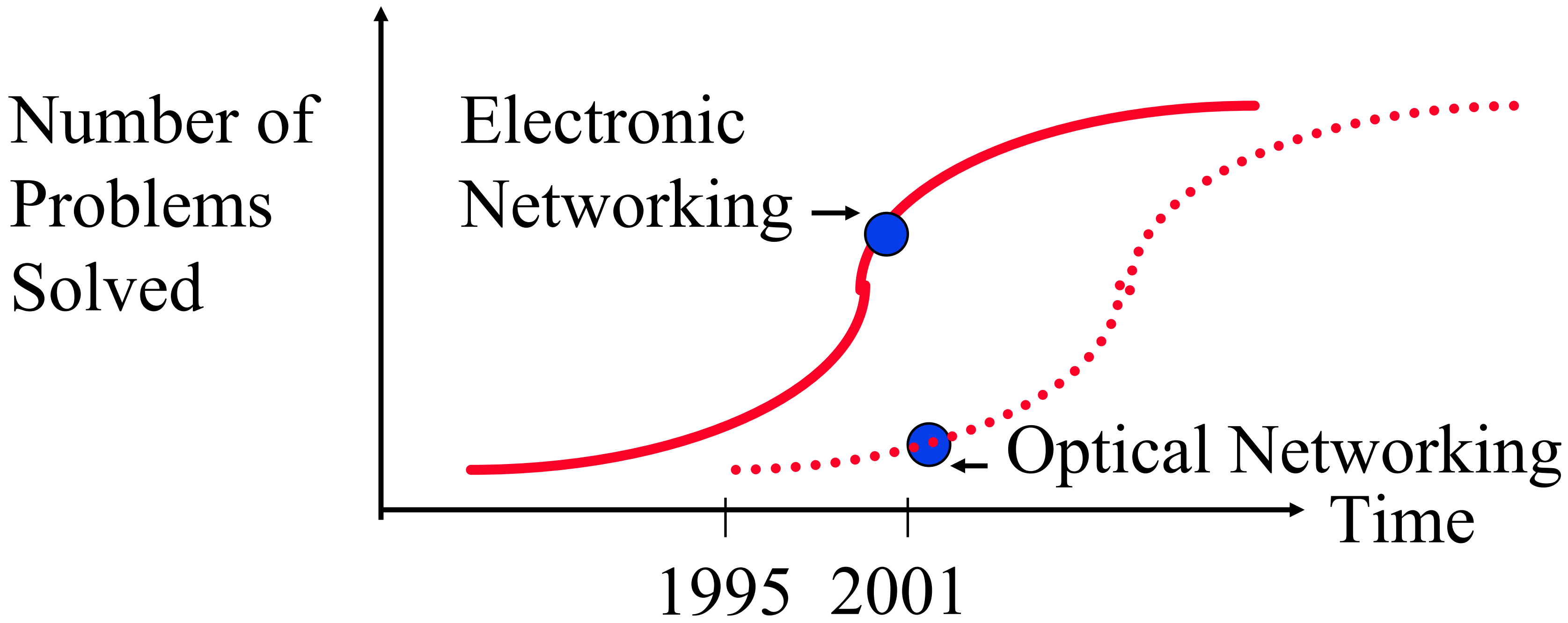
- Networking Bottleneck
- Traffic and Capacity growth
- Life Cycles of Technologies
- Laws of Technology Growth
- Networking and Telecom Convergence
- ATM: Why some technologies fail/succeed?
- Everything over IP
- LAN WAN Convergence: Ethernet Everywhere

Life Cycles of Technologies



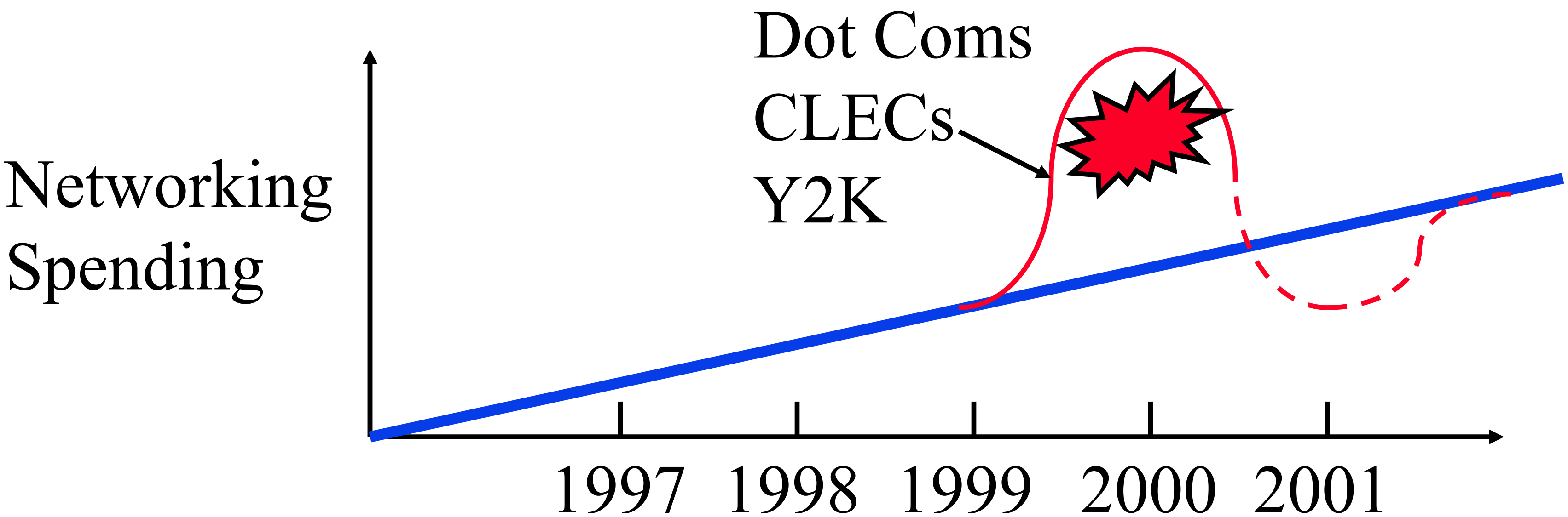
Raj Jain

Life Cycles of Technologies



Raj Jain

The Bubble



- Sidgmore: Internet Traffic doubling every 40 days, 30 days, …⇒ Over-projection data networking equipment
- Nearly 1/3 of all tech IPOs over the last 21 years happened in 1999 and 2000. Source:Morgan Stanely/Chi at Opticomm

Raj Jain

Trends: Networking Bottleneck

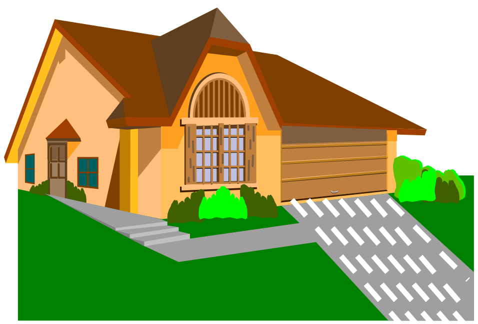
- Communication is more critical than computing
  - Greeting cards contain more computing power than all computers before 1950.
  - Genesis's game has more processing than 1976 Cray supercomputer.
- Networking speed is the key to productivity
- E-Commerce $\Rightarrow$ 20-30% of revenue spent on networking

Impact on R&D

- Too much growth in one year
  ⇒ Can't plan too much into long term
- Long term = 1_2 year or 10_2 years at most
- Products have life span of 1 year, 1 month, …
- Short product development cycles.
  Chrysler reduced new car design time
  from 6 years to 2.
- Information doubles every 18 months.
  Your knowledge becomes obsolete in 18 months.

Garden Path to I-Way

- Plain Old Telephone System (POTS) = 64 kbps = 3 ft garden path
- ISDN = 128 kbps = 6 ft sidewalk
- T1 Links to Businesses = 1.544 Mbps = 72 ft = 4 Lane roadway
- Cable Modem Service to Homes: = 10 Mbps = 470 ft = 26 Lane Driveway



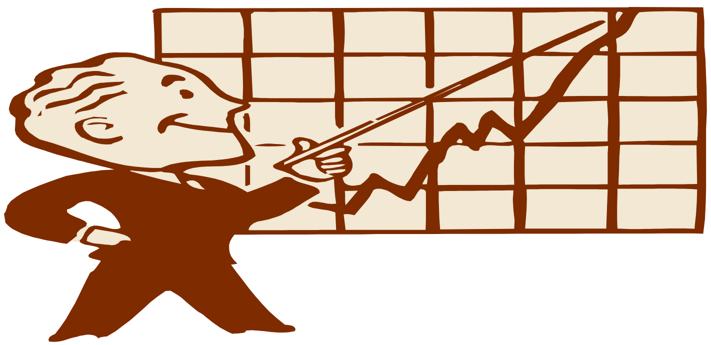
- OC3 = 155 Mbps = 1 Mile wide superhighway
- OC48 = 2.4 Gbps = 16 Mile wide superhighway
- OC768 = 38.4 Gbps = 256 Mile wide superhighway

Raj Jain

Trend: Faster Media

- One Gbps over 4-pair UTP-5 up to 100 m
  Was 1 Mbps (1Base-5) in 1984.
  10/100 Mbps over residential phone lines (actelis.com)
- Dense Wavelength Division Multiplexing (DWDM) allows 64 wavelengths in a single fiber
  $64 \times$OC-192 = 0.6 Tbps
  OC-768 = 40 Gbps demonstrated in 1998.
  Was 100 Mbps (FDDI) in 1993.
- 11 Mbps in-building wireless networks
  Was 1 Mbps (IEEE 802.11) in 1998.

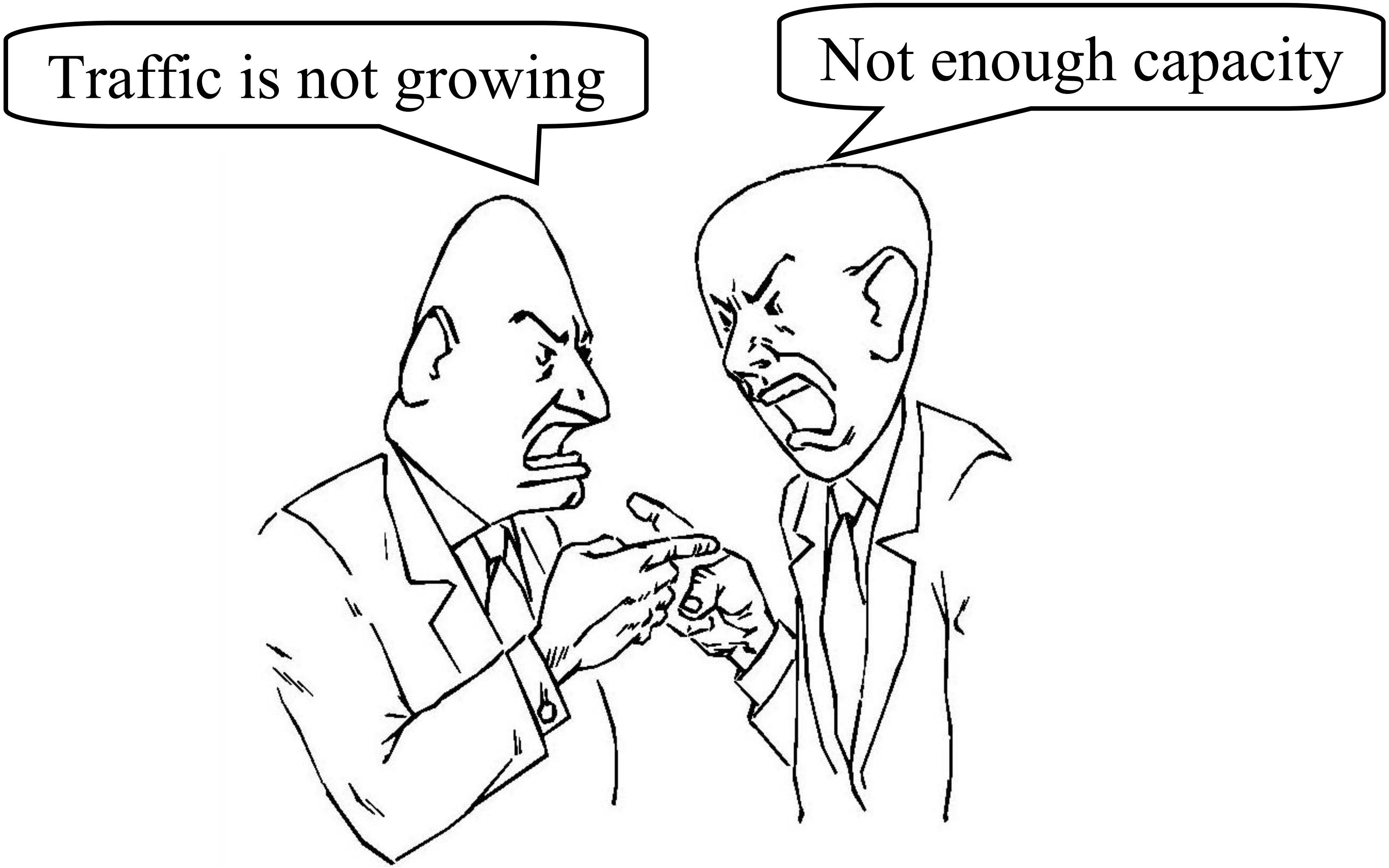
Trend: More Traffic



- Number of Internet hosts was growing super-exponentially.
- Traffic per host was increasing:
  - Cable modems allow 1 to 10 Mbps from home
  - 6-27 Mbps over phone lines using ADSL/VDSL
- Bandwidth requirements were doubling every ~~4 months~~ 100 days
- Actual demand has exceeded all past projections.

Raj Jain

Glut-sters vs Gap-sters



Raj Jain

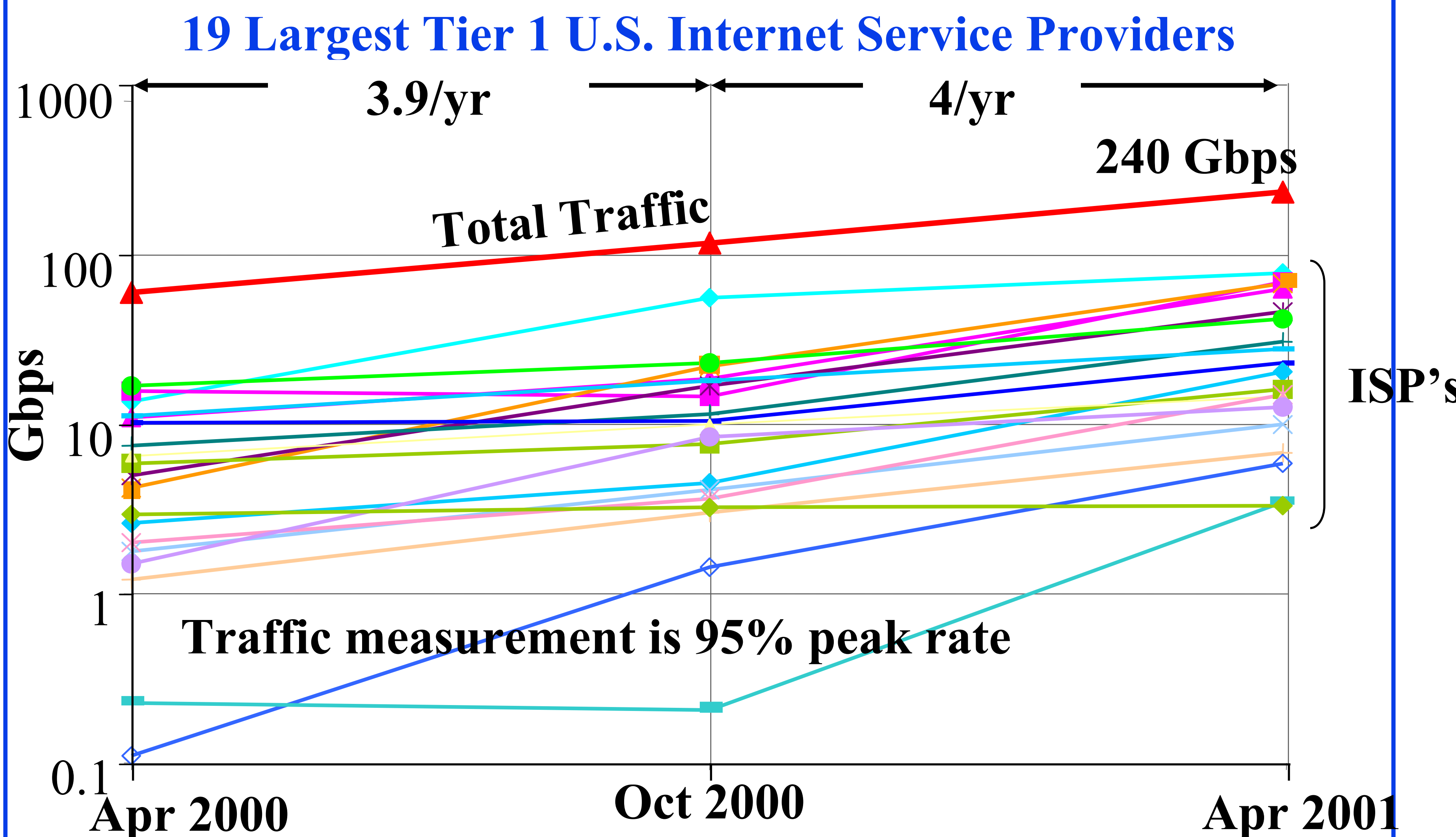
Is Internet Growing?

- IP Traffic Growth will slow down from 200-300% per year to 60% by 2005
  - McKinsey & Co and JP Morgan, May 16, 2001
- 98% of fiber is unlit - WSJ, New York Times, Forbes (Fiber is a small fraction of cost. Laying is expensive.)
- Nortel blamed sales decline on falling IP traffic
- Carriers are using only *avg* 2.7% of their total *lit* fiber capacity - Michael Ching, Marril Lynch & Co. in Wall Street Journal

Internet Growth (Cont)

- Demand on 14 of 22 most used routes exceeds 70%
  -Telechoice, July 19, 2001
- Traffic grew by a factor of 4 between April 2000-April 2001
  -Larry Roberts, August 15, 2001

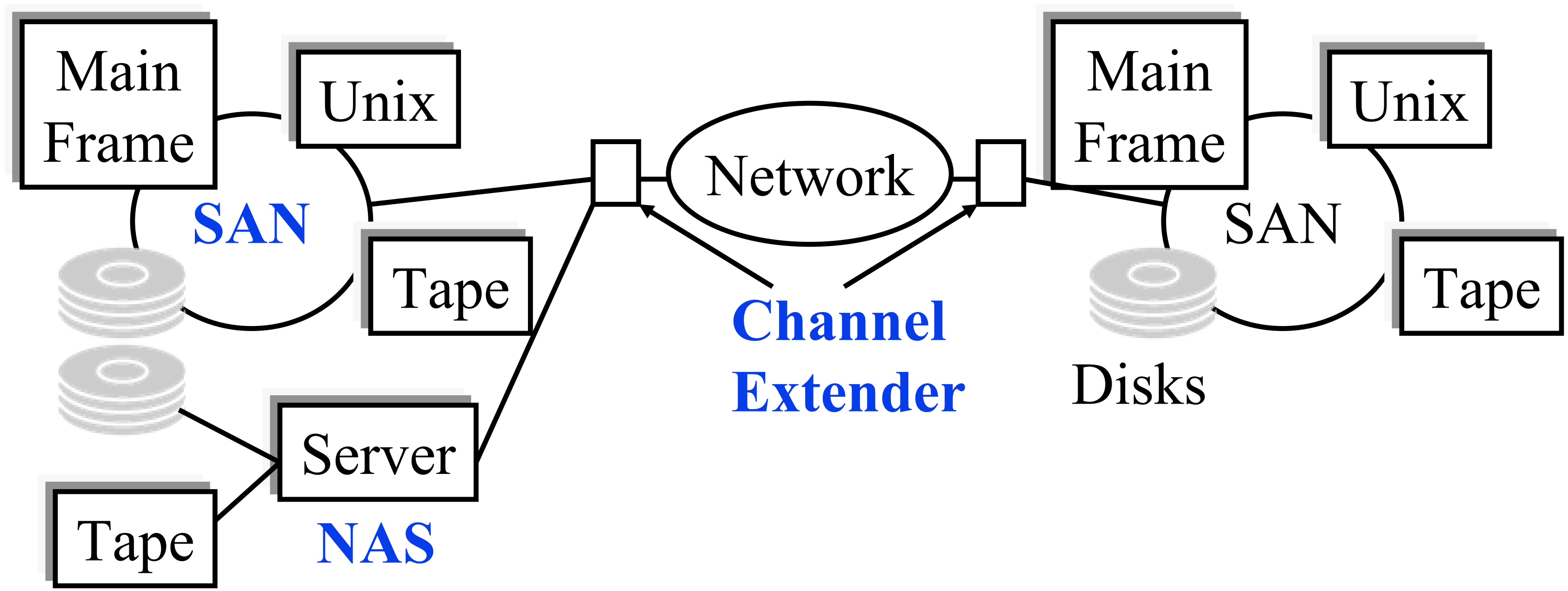
Robert's Traffic Measurements



Source: L. Roberts at Opticomm 2001

Raj Jain

Storage: New Traffic Demands



- Fiber Channel SAN limited to 10 km
- SAN extender switches allow connectivity over metro and long-haul optical networks ⇒ Outsourced storage
- Multiservice switches allow IP, ATM, Sonet, ESCON, ...

Trend: Traffic > Capacity

Expensive Bandwidth	Cheap Bandwidth
Sharing	No sharing
Multicast	Unicast
Virtual Private Networks	Private Networks
Need QoS	QoS less of an issue
Likely in WANs	Possible in LANs

Raj Jain

Laws of Technology Growth

1. Gordon Moore's Law:

$$P \approx P(0)[1+B]^t$$

B=0.5 $\Rightarrow$ Processor performance doubles every 18 mo

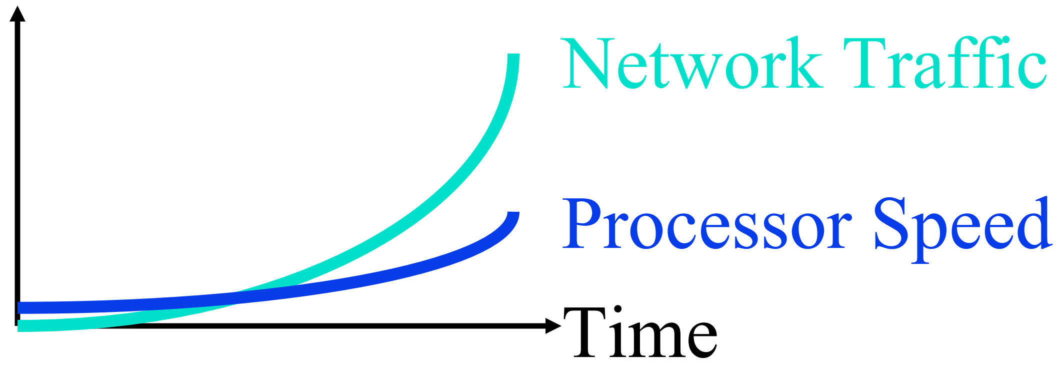
The next (double speed) processor has to be in 18 mo.

2. Bob Metcalf's Law:

$$N \approx N(0)n^2$$

The value (traffic, information) on a network is proportional to the square of number of users or computers. Law of increasing returns. Applies to other forms of communications - magazines, radio, … also.

Technology Growth (Cont)



3. Sidgmore's Law:

$$T \approx T(0)2^{4t}$$

Traffic on the net (UUNET) doubles every 3 months

$\Rightarrow$ Faster than Moore's Law

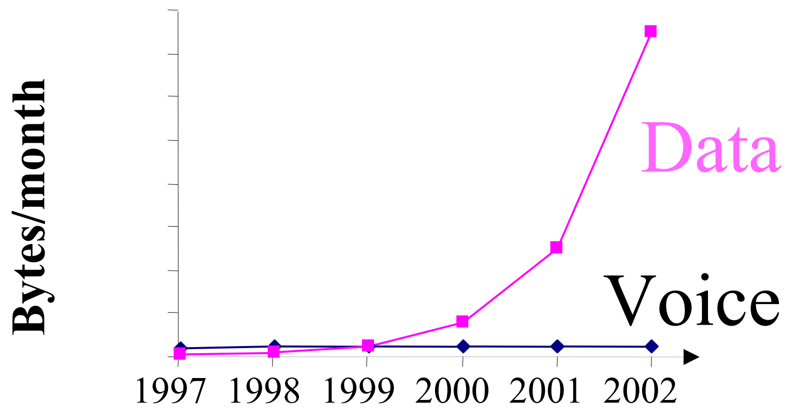
4. Cost of Bandwidth:

$$C \approx C(0)[1/2]4t$$

Cost is dropping exponentially.

The rates differ for enterprise and carrier spaces

Trend: Data > Voice

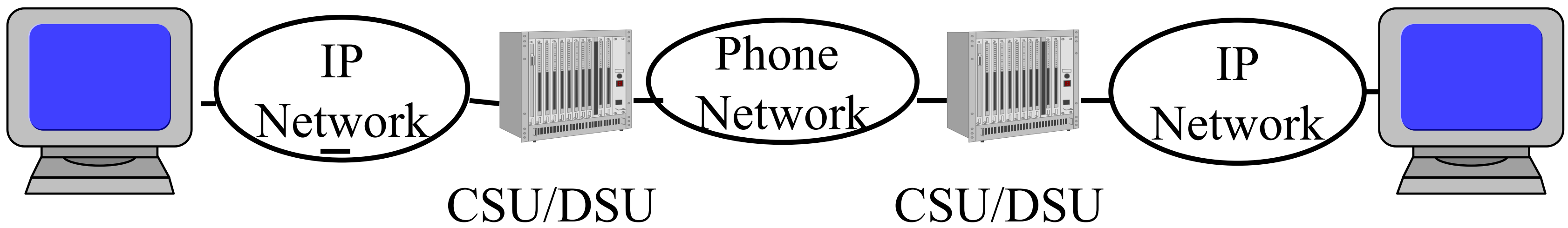


- Voice traffic is growing linearly
  Data traffic is growing exponentially
- In 1998-99, data traffic on carrier networks exceeded the voice traffic.
- Everyone is trying to get into the data business:
  - Phone Networks ⇒ High-speed frame relay
  - Video Networks ⇒ Cable Modems

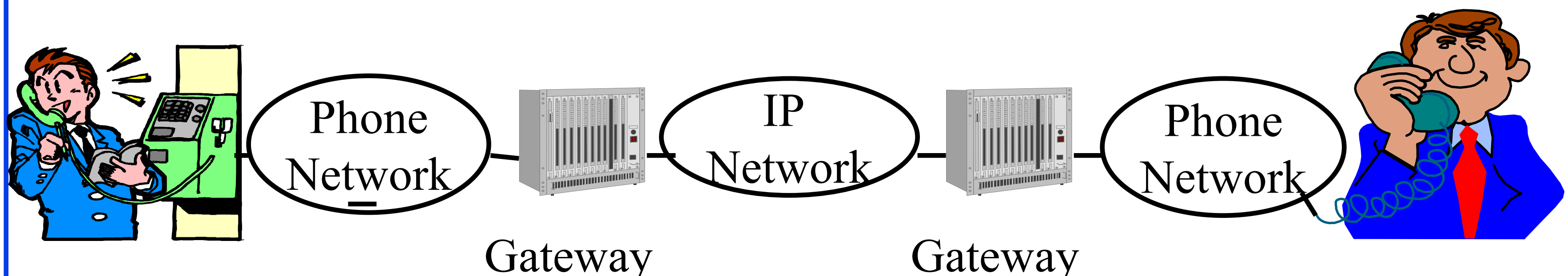
Raj Jain

Data > Voice (Cont)

- Past: Data over Voice



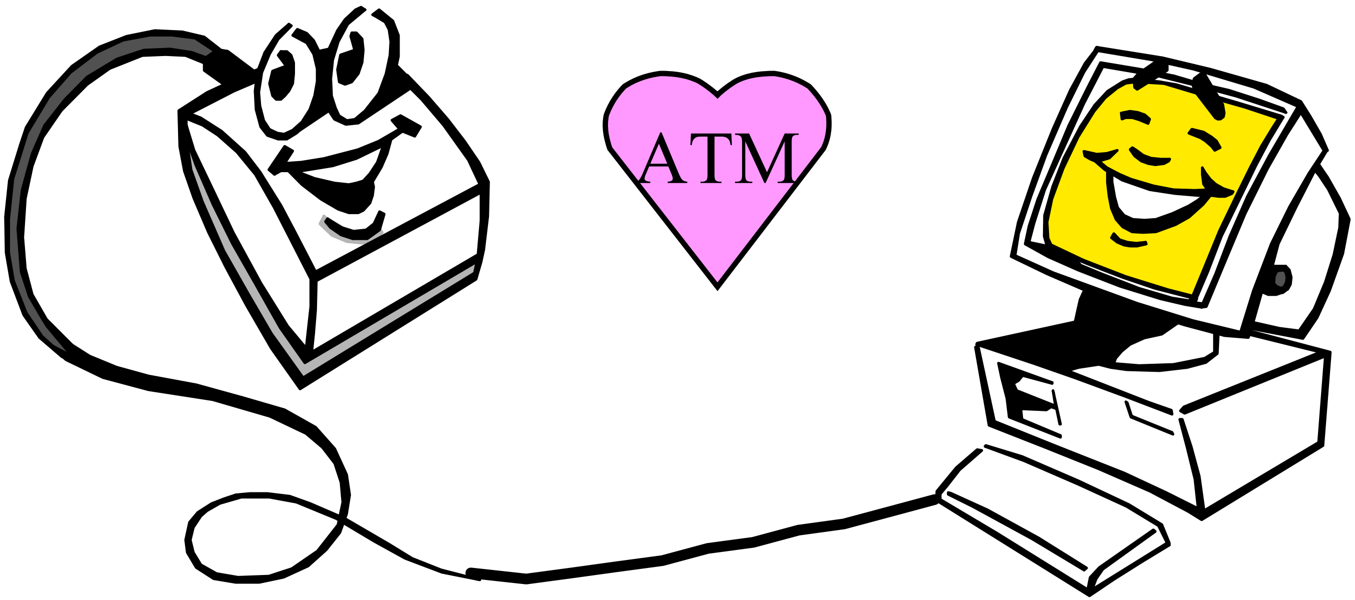
- Future: Voice over Data



- Convergence: Data+Voice+Video
  AT&T + TCI, Lucent+Ascend, Nortel+Baynetworks
- Voice over DSL: 16 lines + Data over 1 UTP by CLECs
- Voice is symmetric and local, Data is assymetric/global

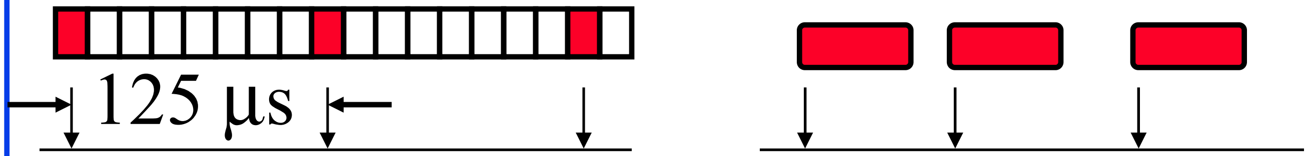
ATM

- Asynchronous Transfer Mode
- ATM Net = Data Net + Phone Net
- Combination of Internet method of communication (packet switching) and phone companies' method (circuit switching)



Raj Jain

ATM vs Phone Networks



- Current phone networks are synchronous (periodic). ATM = Asynchronous Transfer Mode
- Phone networks use circuit switching. ATM networks use "Packet" Switching
- In phone networks, all rates are multiple of 8 kbps. With ATM service, you can get any rate. You can vary your rate with time.
- With current phone networks, all high speed circuits are manually setup. ATM allows dialing any speed.

ATM vs Data Networks

- Traffic Management: Loss based in IP.
  ATM has 1996 traffic management technology.
  Required for high-speed and variable demands.
- Qos based Routing: Private Network to Node Interface (PNNI)
- Signaling: Internet Protocol (IP) is connectionless.
  You cannot reserve bandwidth in advance.
  ATM is connection-oriented.
  You declare your needs before using the network.
- Switching: In IP, each packet is addressed and processed individually.
- Cells: Fixed size or small size is not important

Trend: ATM in Carrier Backbone

- Most carriers including AT&T, MCI, Sprint, UUNET, have ATM backbone
- Over 80% of the internet traffic goes over ATM
- ATM provides:
  - Traffic management
  - Voice + Data Integration: CBR, VBR, ABR, UBR
  - Signaling
  - Quality of service routing: PNNI
- ATM can't reach desktop: Designed by carriers. Complexity in the end systems. Design favors voice.

Raj Jain

Networking: Failures vs Successes

- 1980: Broadband (vs baseband)
- 1981: PBX (vs Ethernet)
- 1984: ISDN (vs Modems)
- 1986: MAP/TOP (vs Ethernet)
- 1988: OSI (vs TCP/IP)
- 1991: DQDB
- 1992: XTP (vs TCP)
- 1994: CMIP (vs SNMP)

Requirements for Success

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

Raj Jain

Trend: Everything over IP

- Data over IP ⇒ IP needs Traffic engineering
- Voice over IP ⇒ Quality of Service, Signaling, virtual circuits (MPLS)
- Internet Engineering Task Force (IETF) is the center of action.
  Attendance at ITU is down.

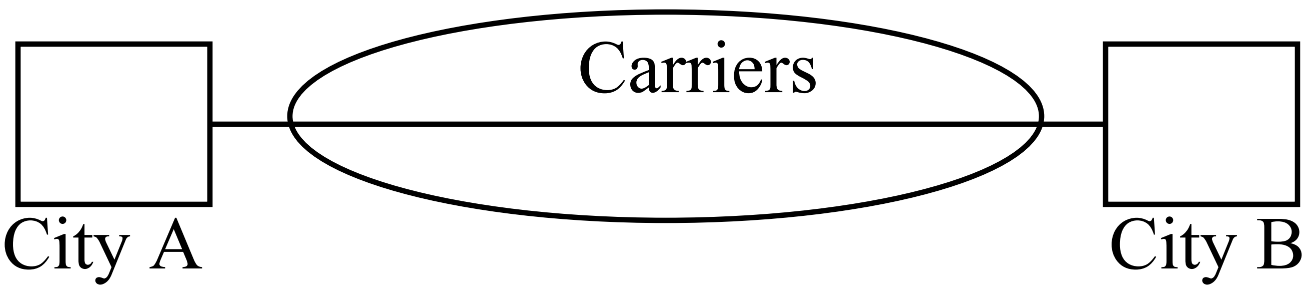
Trend: Traffic Engineering

- User's Performance Optimization
  ⇒ Maximum throughput, Min delay, min loss, min delay variation
- Efficient resource allocation for the provider
  ⇒ Efficient Utilization of all links
  ⇒ Load Balancing on parallel paths
  ⇒ Minimize buffer utilization
- Current routing protocols (e.g., RIP and OSPF) find the shortest path (may be over-utilized).
- QoS Guarantee: Selecting paths that can meet QoS
- Enforce Service Level agreements
- Enforce policies: Constraint based routing $\supseteq$ QoSR

Raj Jain

SONET

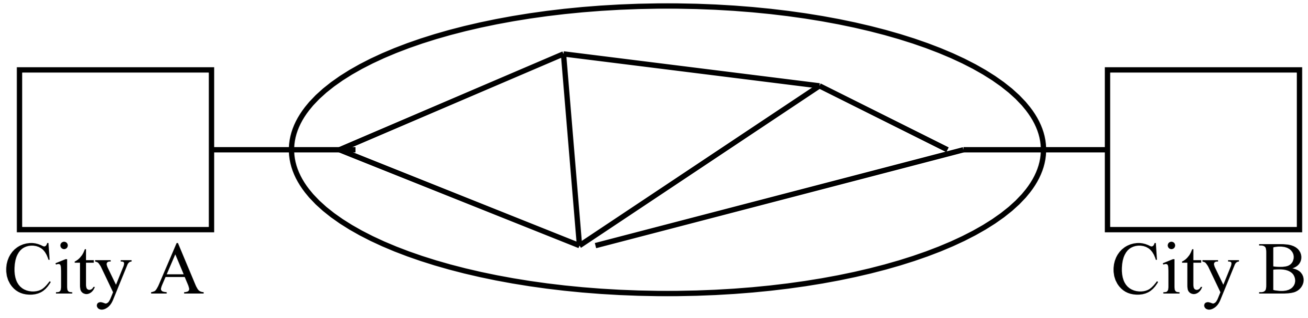
- Synchronous optical network
- Standard for digital optical transmission (bit pipe)
- Developed originally by Bellcore.
  Standardized by ANSI T1X1
  Standardized by CCITT
  $\Rightarrow$ Synchronous Digital Hierarchy (SDH)
- You can lease a SONET connection from carriers

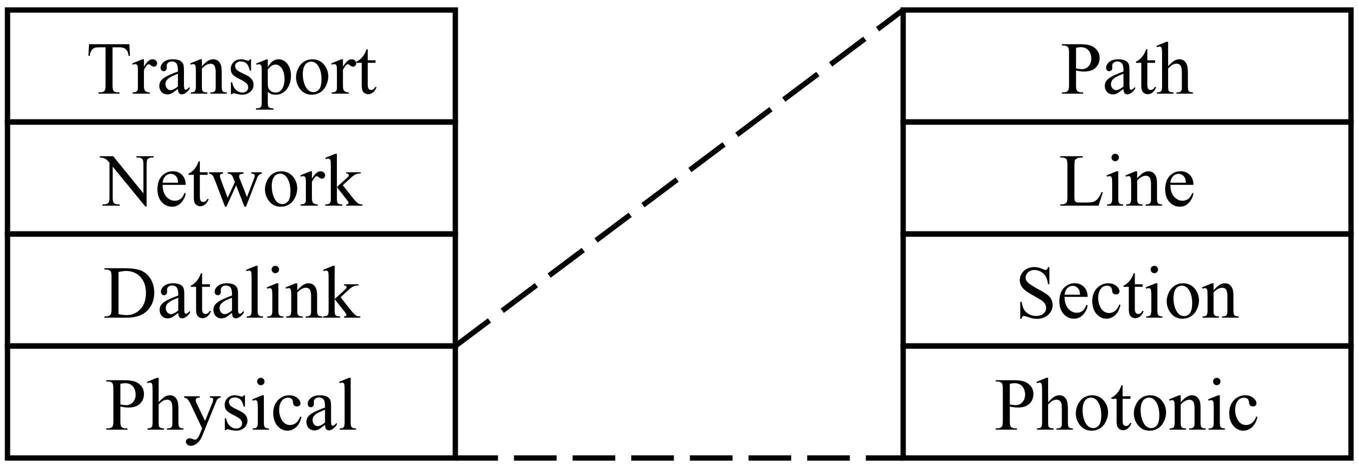


Raj Jain

SONET Protocols

- Synchronous Optical **Network**

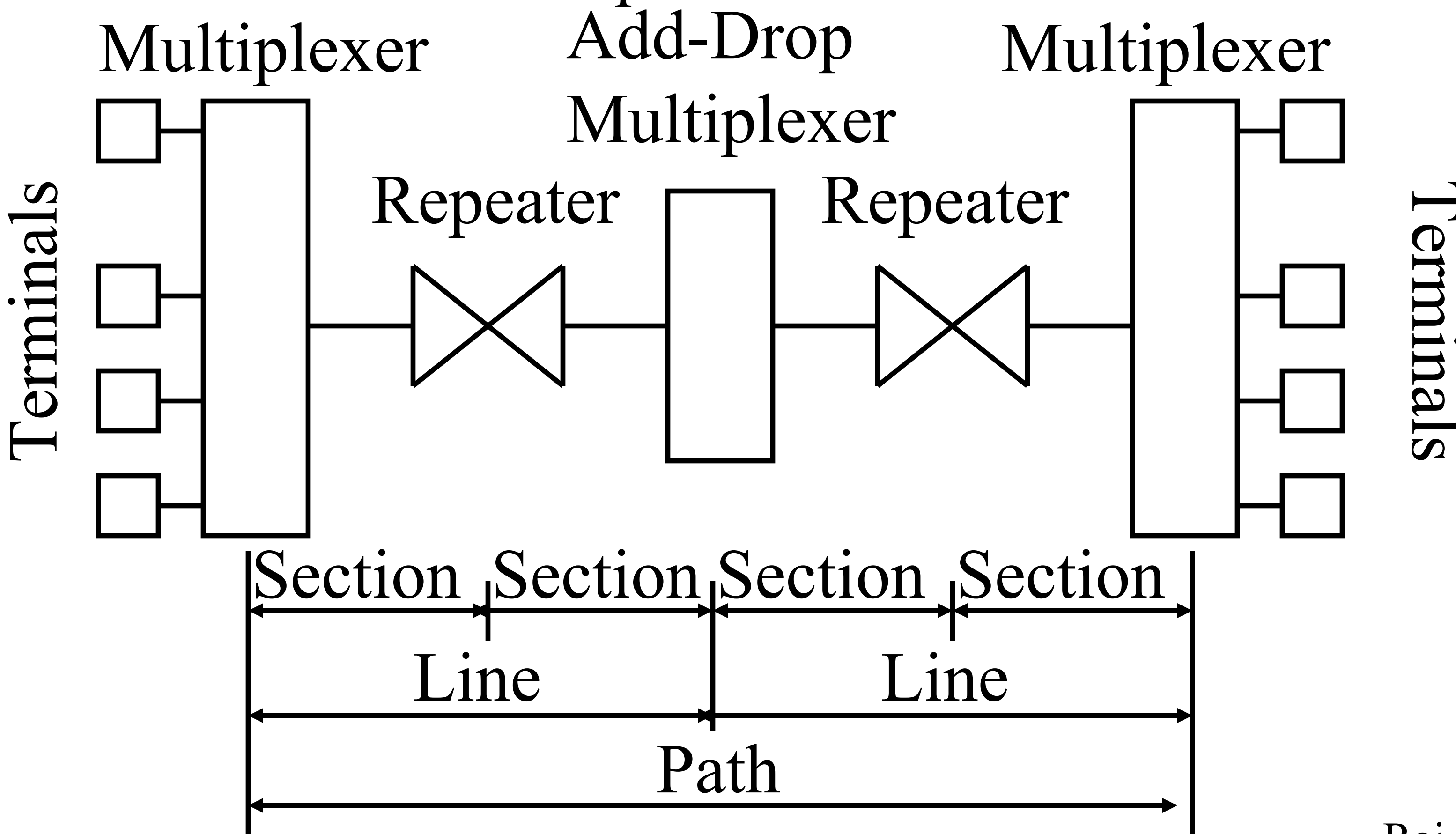




Raj Jain

Physical Components

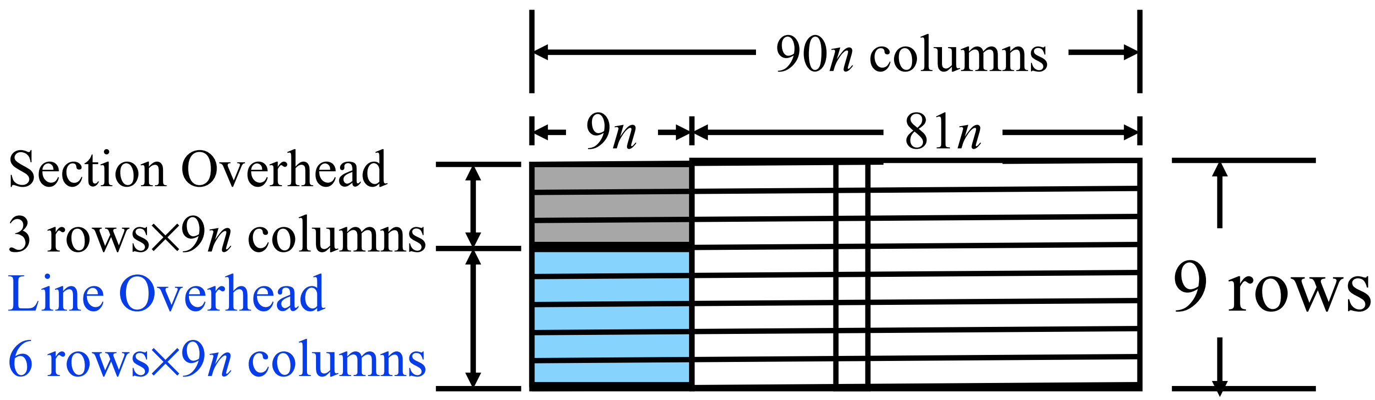
- Section = Single run of fiber
- Line = Between multiplexers



Raj Jain

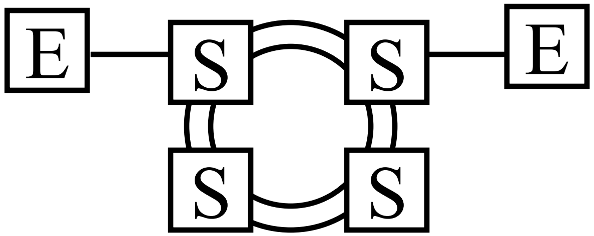
SONET Frame Format

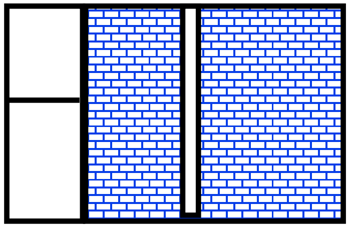
- OC-1 = 51.84 Mbps (payload and overhead)
- OC-n = $n \times$ 51.84 Mbps
  e.g., OC-3 = 3 × 51.84 = 155.54 Mbps
- All SONET frames are 125 μs long.
  E.g., OC-3 frames are 2430 (125 × 155.54) bytes
- Represented as 2D arrays of bytes.
  9 rows × 90n columns. Transmitted row-wise



Raj Jain

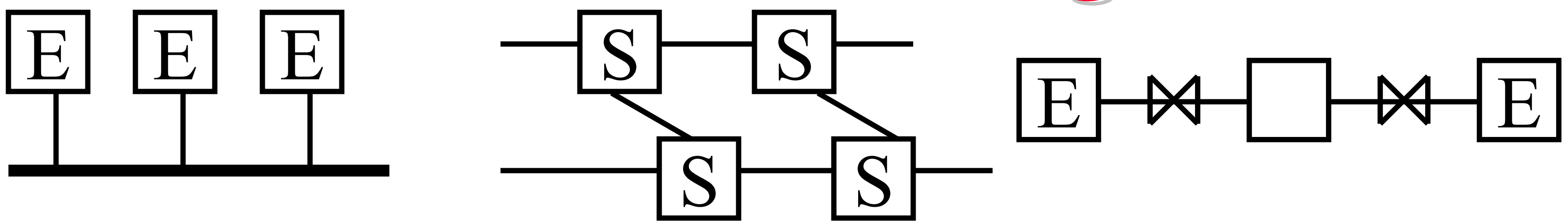
SONET Functions





- Grooming: T1's to OC-3, 4 OC-3 to OC-12, etc
- Protection: Allows redundant Line or path,
  - Signal on both paths: 1+1
  - Signal on working only: 1:1
- Fast Restoration: 50ms using rings
- Synchronous operation: No queues. Guaranteed delay
- Payload mapping for Ethernet, ATM, etc.

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

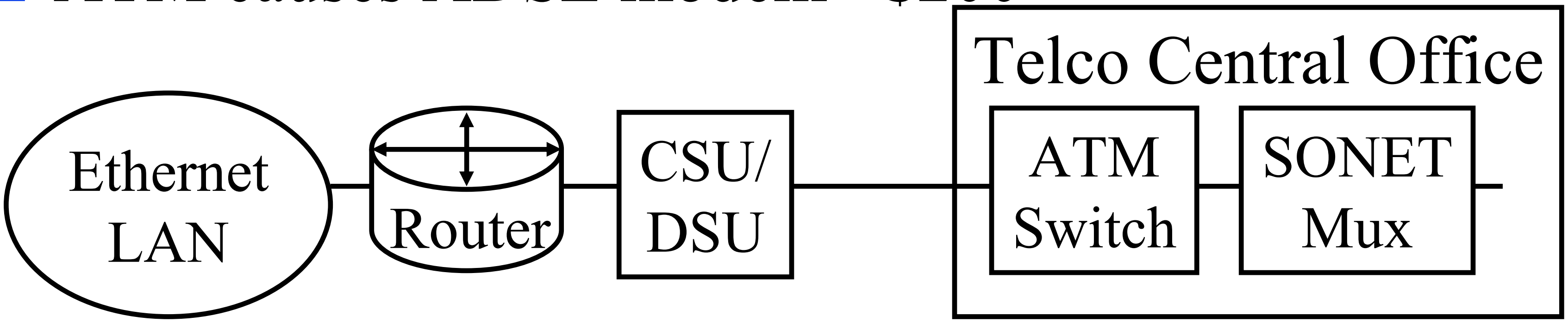
Raj Jain

Trend: Ethernet Everywhere

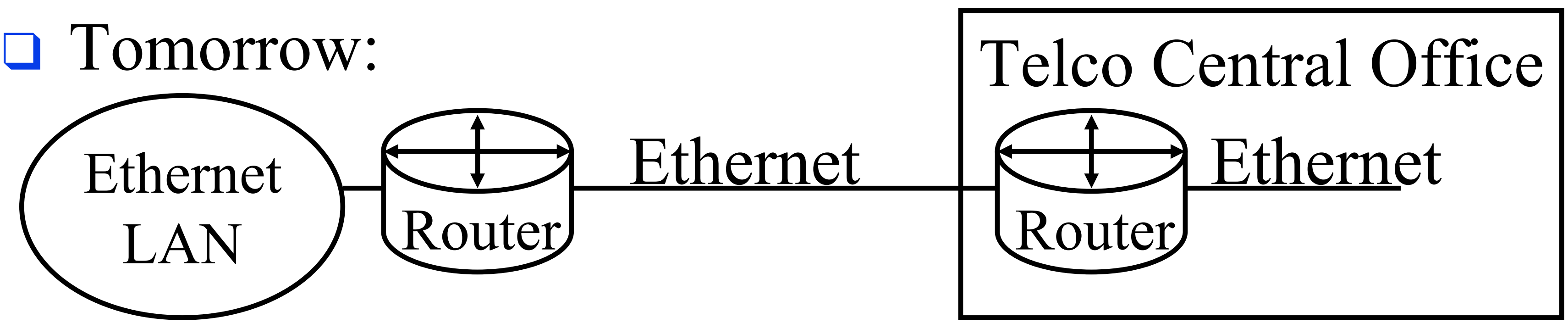
- Ethernet vs SONET in Metro:
  Survivability, Restoration
  Ring Topology
- Ethernet vs DSL in Access:
  Longer distances
- Ethernet vs ATM in Enterprise:
  Class of service
- Ethernet vs phone network in homes:
  Power over Ethernet

Ethernet in the First Mile

- IEEE 802.3 Study Group started November 2000
- Originally called Ethernet in the Last Mile
- ATM causes ADSL modem >$200



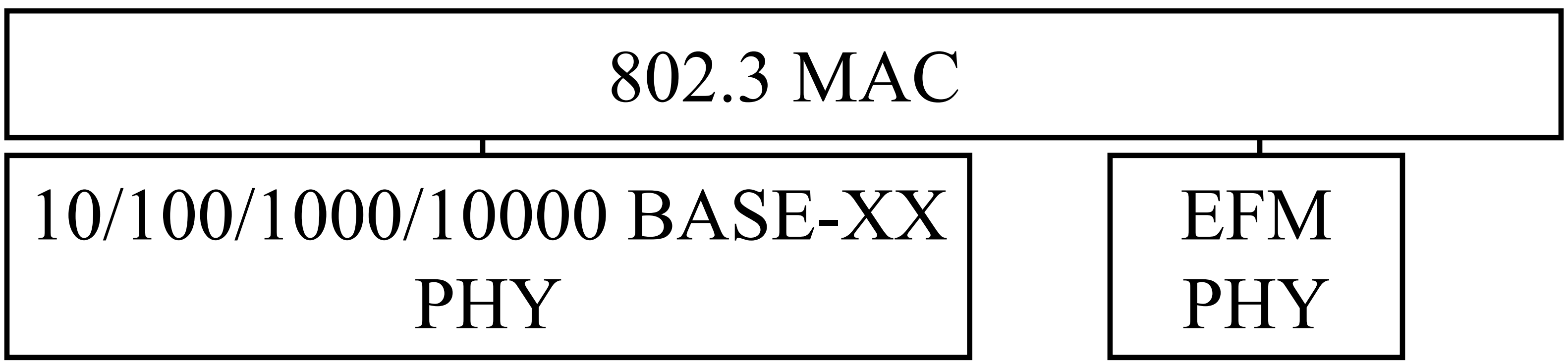
- Tomorrow:



Raj Jain

EFM

- Current Technologies: ISDN, xDSL, Cable Modem, Satellite, Wireless
- EFM Goals: (To be determined)
  - Media: Single pair UTP-3, 4-pair UTP-5, Fiber, Air
  - Speed: 125 kbps to 1 Gbps
  - Distance: 1500 ft, 18000 ft, 1 km - 40 km
- Compatibility:



- Ref: http://www.ieee802.org/3/efm/public/index.htm

Power over MDI

- IEEE 802.3af working group
- MDI = Media Dependent Interface
- Applications: Web Cams, PDAs, Intercoms, Ethernet Telephones, Wireless LAN Access points, Fire Alarms, Remote Monitoring, Remote entry
- Power over TP to a single Ethernet device: 10BASE-T, 100BASE-TX, 1000BASE-T (TBD)
- Interoperate with legacy RJ-45 Ethernet devices
- Allows:
  - Switch to Switch connections (both supplying power)

EFM (Cont)

- June 2001: Metro Ethernet Forum formed by 37 companies including vendors and carriers

Goals:

- Maximize deployment of Ethernet-based Metro nets
- Drive the standards and implementations
- Facilitate Interoperability
- Increase awareness

www.MetroEthernetForum.org

Power over MDI (Cont)

- Allows:
  - Cross-over cables
  - Shorted conductors, loopback plugs
- Approx 40V, 350mA at source
- One standard for worldwide use
- PAR approved: 30 January 2000
- Standard Expected: November 2002
- Email: subscribe stds-802-3-pwrviamdi <email> to majordomo@mail.ieee.org
- Ref: http://grouper.ieee.org/groups/802/3/power_study/public/nov99/802.3af_PAR.pdf

Summary

1. Traffic > Capacity ⇒ Traffic Engineering
2. Voice < Data ⇒ Voice over IP
3. Everything over IP ⇒ Quality of service in IP
4. Convergence: Voice+Video+Data Industries and Voice+Video+Data networks
5. Everything over Ethernet ⇒ LAN-WAN convergence

Raj Jain

Key References

- For a detailed list of references, see http://www.cis.ohio-state.edu/~jain/refs/ref_trnd.htm Also reproduced in the back of this tutorial handout.
- George Guilder, "Telecosm: How Infinite Bandwidth will Revolutionize Our World," Free Press, 2000, 352 pages.
- T. Lewis, "The Next $10{,}000_2$ years," IEEE Computer, April/May 1996
- "The Next 50 years," Special issue of Communications of the ACM, Feb 1997.

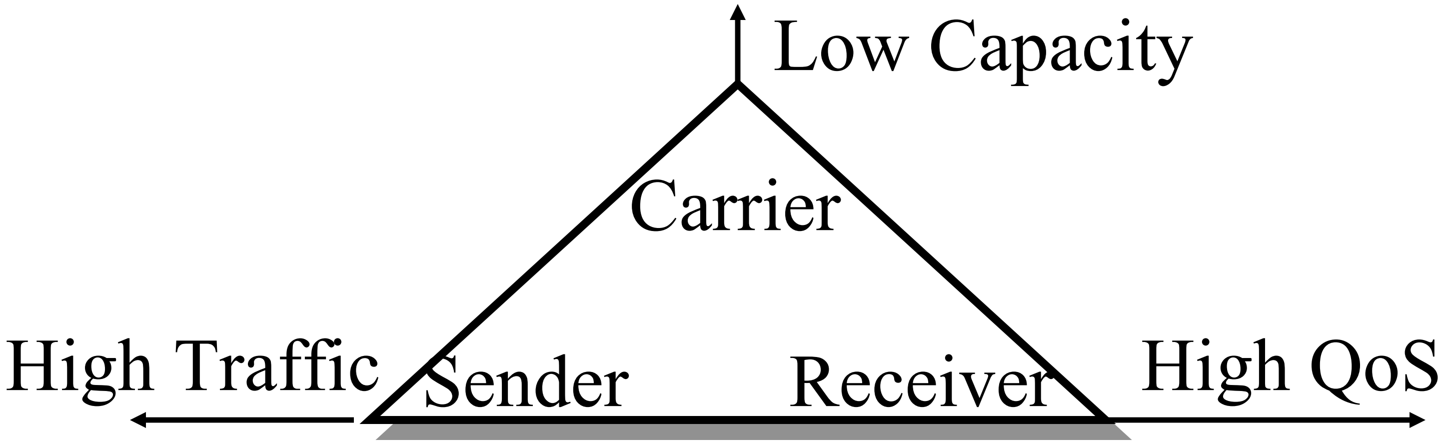
Quality of Service In Data Networks

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

- ATM QoS and Issues
- Integrated Services and RSVP
- Differentiated Services:
  Expedited and Assured Forwarding
- Subnet Bandwidth Manager (SBM)
- COPS Protocol for Policy
- IEEE 802.1D Model
- Comparison of QoS Approaches

Raj Jain

QoS Triangle



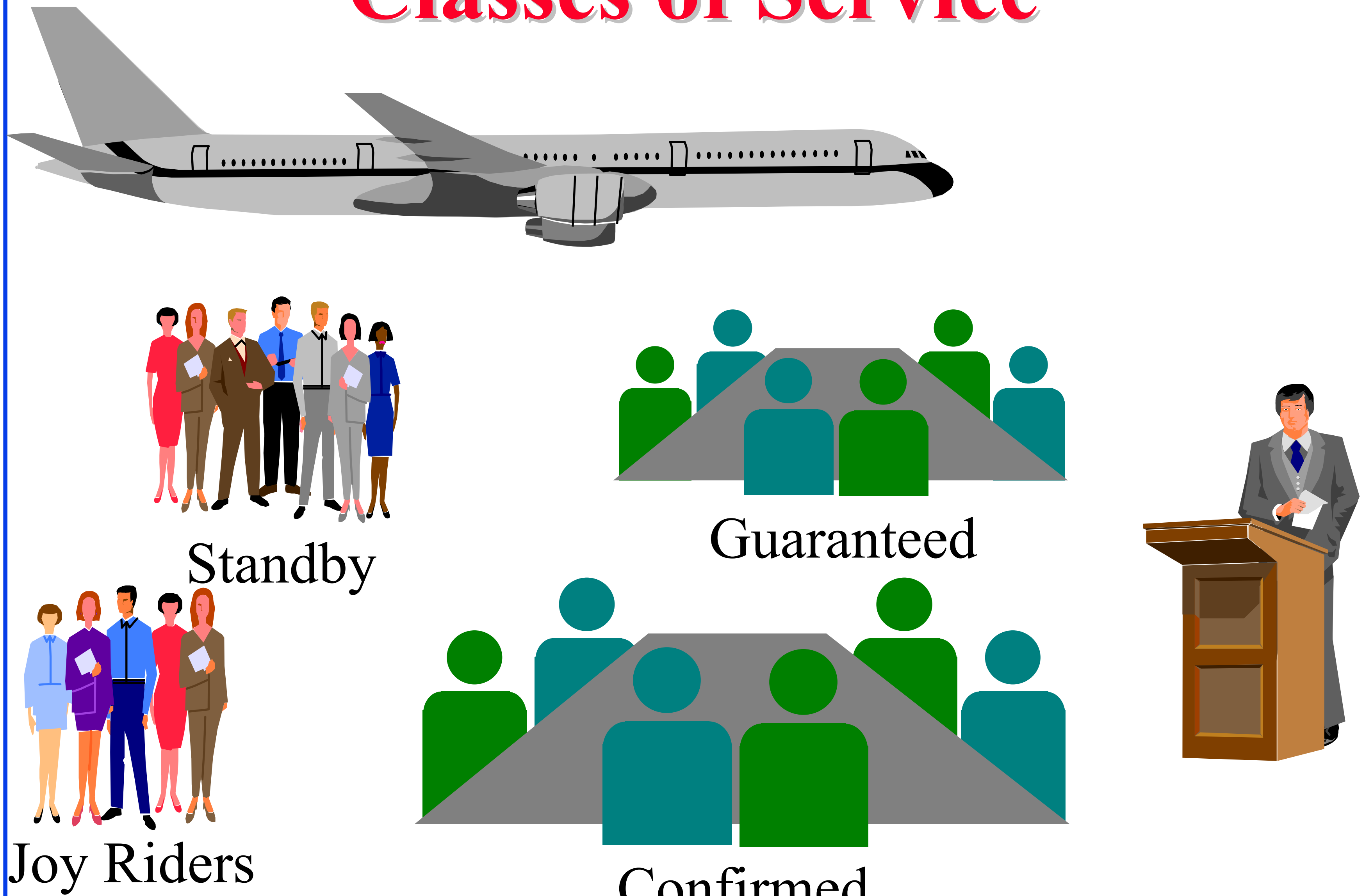
- Senders want to send traffic any time with high load, high burstiness
- Receivers expect low delay and high throughput
- Since links are expensive, providers want to minimize the infrastructure
- If one of the three gives in $\Rightarrow$ no problem

Raj Jain

What is QoS?

- Predictable Quality: Throughput, Delay, Loss, Delay jitter, Error rate
- Opposite of best effort or random quality
- Different treatment to different types of traffic
- Mechanisms:
  - Capacity Planning
  - Classification, Queueing, Scheduling, buffer management
  - QoS based path determination, Route pinning
  - Shaping, policing, admission control
  - Signaling

Classes of Service

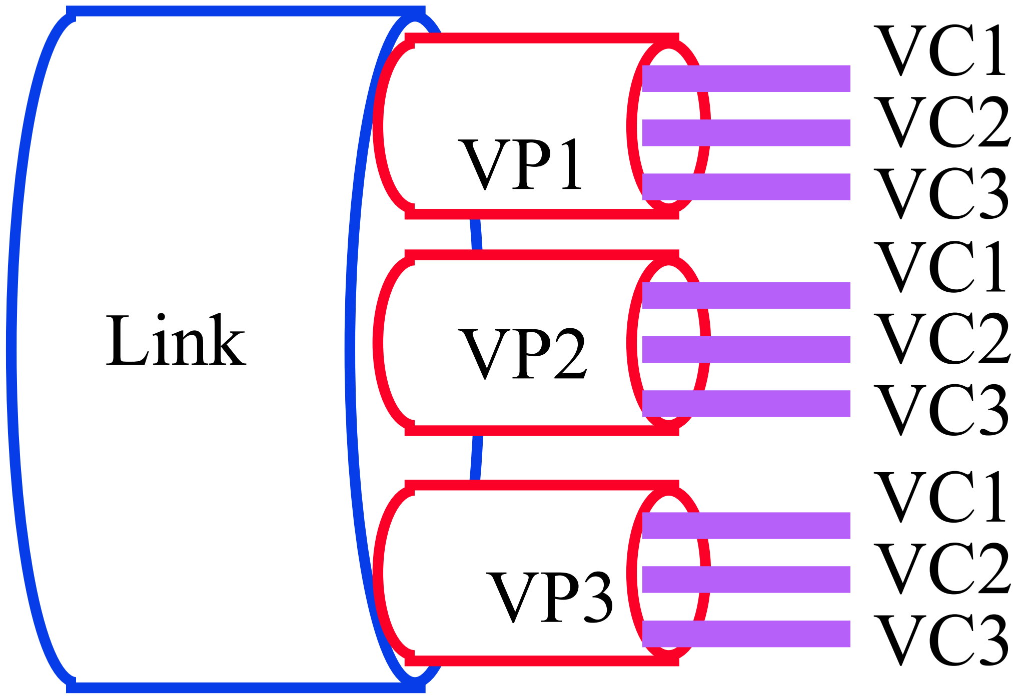


Classes of Service

- ABR (Available bit rate): Source follows feedback. Max throughput with minimum loss.
- UBR (Unspecified bit rate):
  User sends whenever it wants. No feedback. No guarantee. Cells may be dropped during congestion.
- CBR (Constant bit rate): User declares required rate. Throughput, delay and delay variation guaranteed.
- VBR (Variable bit rate): Declare avg and max rate.
  - rt-VBR (Real-time): Conferencing.
    Max delay guaranteed.
  - nrt-VBR (non-real time): Stored video.
- GFR (Guaranteed Frame Rate): Min Frame Rate

VC Aggregation

- 24/28-bit connection identifier
  First 8/12 bits: Virtual Path,
  Last 16 bits: Virtual Circuit
- VP service allows new VC's w/o orders to carriers



ATM QoS: Issues

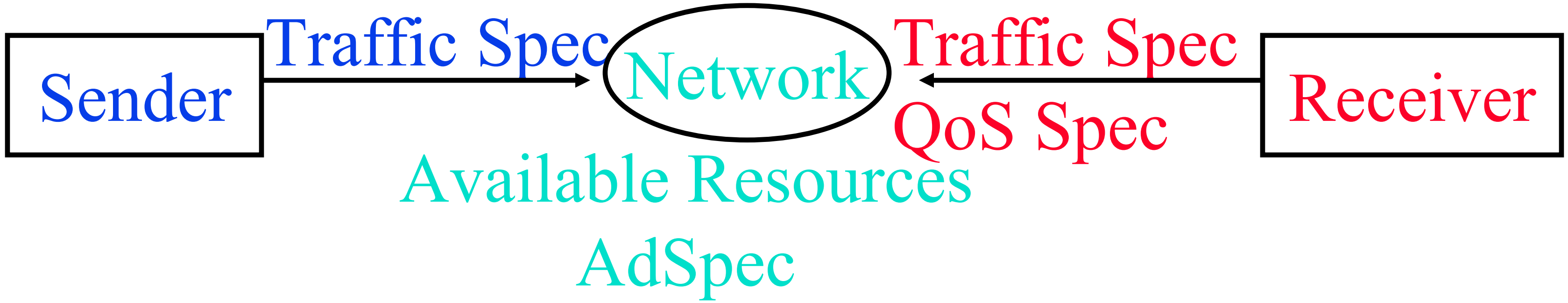
- Can't easily aggregate QoS: VP =?= Σ VCs
- Can't easily specify QoS: What is the cell delay variation (CDV) required for a movie?
- Signaling too complex ⇒ Need Lightweight Signaling
- Need Heterogeneous Point-to-Multipoint: Variegated VCs
- Need QoS Renegotiation
- Need Group Address
- Need priority or weight among VCs to map DiffServ and 802.1D

Integrated Services

- Best Effort Service: Like UBR.
- Controlled-Load Service: Performance as good as in an unloaded datagram network. No quantitative assurances. Like nrt-VBR or UBR w MCR
- Guaranteed Service: rt-VBR
  - Firm bound on data throughput and delay.
  - Delay jitter or average delay not guaranteed or minimized.
  - Every element along the path must provide delay bound.
  - Is not always implementable, e.g., Shared Ethernet.
  - Like CBR or rt-VBR

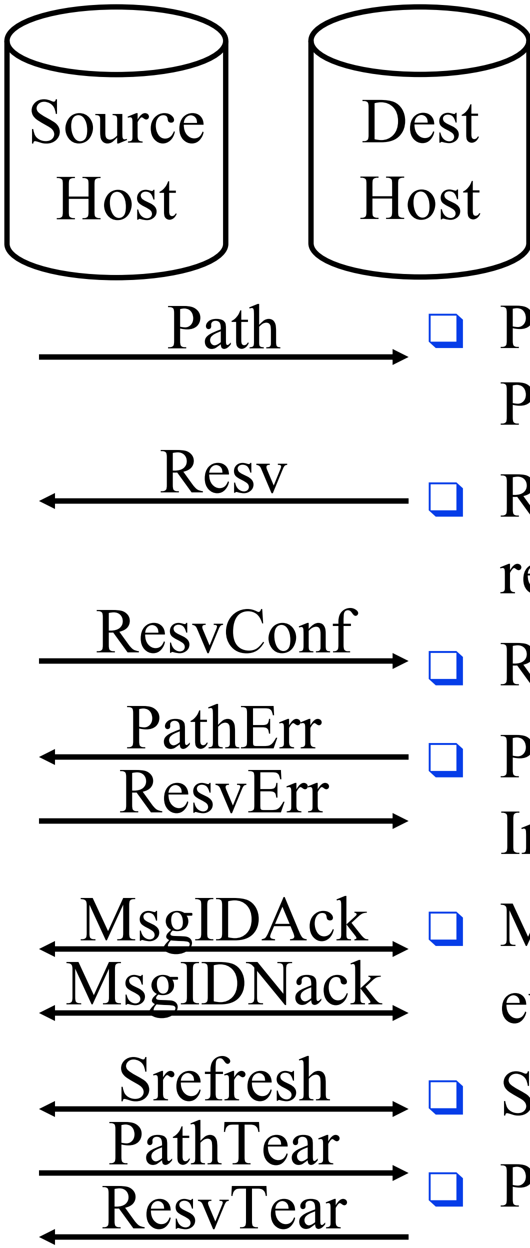
RSVP

- Resource ReSerVation Protocol
- Internet signaling protocol
- Carries resource reservation requests through the network including traffic specs, QoS specs, network resource availability
- Sets up reservations at each hop



Raj Jain

RSVP Messages



- Path: Sender's traffic spec
  Path state is created. Regularly referesheд.
- Resv: Resv state is created. Regularly referesheд
- ResvConf: Sent upon request to confirm Resv
- PathErr, ResvErr: Error in path or reservation Installation
- MsgIdAck/Nack: Trigger reporting of an event
- Srefresh: To refresh a group of path/resv states
- PathTear/ResvTear: Remove Path/Resv States

Problems with RSVP and Integrated Services

- Complexity in routers: multi-field packet classification, scheduling
- Per-flow signaling, packet handling, state.
  O(n) ⇒ Not scalable with # of flows.
  Number of flows in the backbone may be large.
  ⇒ Suitable for small private networks
- Need a concept of "Virtual Paths" or aggregated flow groups for the backbone
- Need policy controls: Who can make reservations?
  Support for accounting and security.
  ⇒ RSVP admission policy (rap) working group.

Problems (Cont)

- Receiver Based:
  Need sender control/notifications in some cases.
  Which receiver pays for shared part of the tree?
- Soft State: Need route/path pinning (stability).
  Limit number of changes during a session.
- RSVP does not have negotiation and backtracking
- Throughput and delay guarantees require support of lower layers. Shared Ethernet ⇒ IP can't do GS or CLS. Need switched full-duplex LANs.
- RSVP is being revived to for MPLS and DiffServ signaling. Also, policy, aggregation, security concepts are being developed

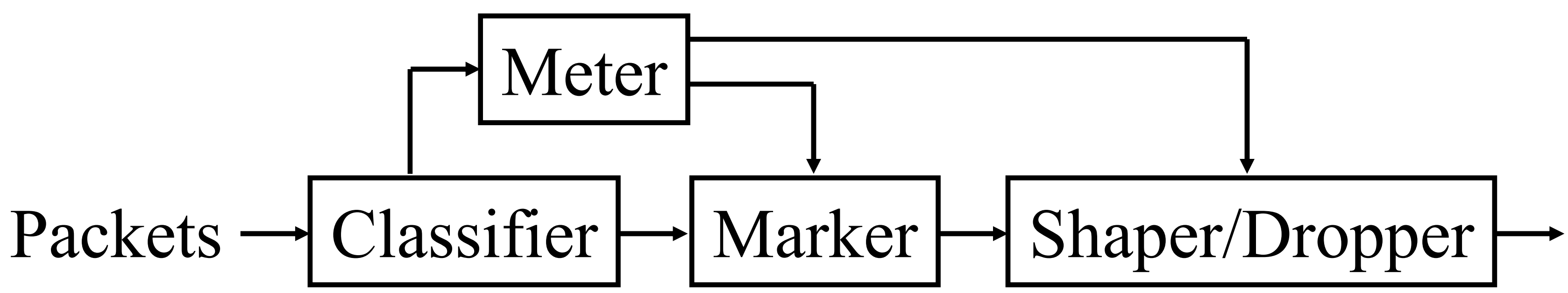
Differentiated Services

Ver	Hdr Len	Precedence	ToS	Unused	Tot Len
4b	4b	3b	4b	1b	16b

- IPv4: 3-bit precedence + 4-bit ToS
- OSPF and integrated IS-IS can compute paths for each ToS
- Many vendors use IP precedence bits but the service varies ⇒ Need a standard ⇒ Differentiated Services
- DS working group formed February 1998
- Charter: Define ds byte (IPv4 ToS field)
- Mail Archive: http://www-nrg.ee.lbl.gov/diff-serv-arch/

DiffServ Concepts

- Micro-flow = A single application-to-application flow
- Traffic Conditioners: Meters (token bucket), Markers (tag), Shapers (delay), Droppers (drop)
- Behavior Aggregate (BA) Classifier:
  Based on DS byte only
- Multi-field (MF) Classifiers:
  Based on IP addresses, ports, DS-byte, etc..

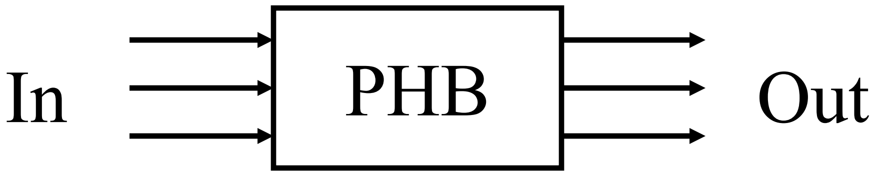


Raj Jain

Diff-Serv Concepts (Cont)

- Service: Offered by the protocol layer
  - Application: Mail, FTP, WWW, Video,...
  - Transport: Delivery, Express Delivery,...
    Best effort, controlled load, guaranteed service
  - DS group will not develop services
    They will standardize "Per-Hop Behaviors"

Per-hop Behaviors



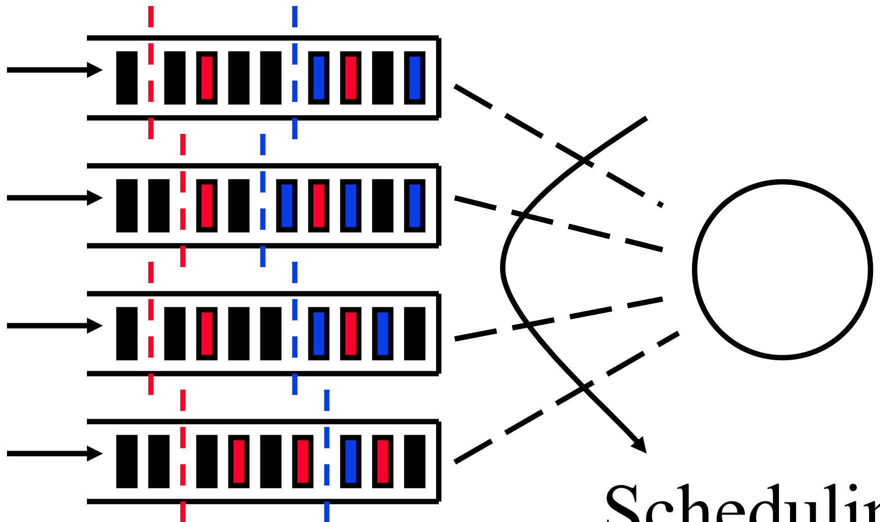
- Externally Observable Forwarding Behavior
- x% of link bandwidth
- Minimum x% and fair share of excess bandwidth
- Priority relative to other PHBs
- PHB Groups: Related PHBs. PHBs in the group share common constraints, e.g., loss priority, relative delay

Raj Jain

Expedited Forwarding

- Also known as "Premium Service"
- Virtual leased line
- Similar to CBR
- Guaranteed minimum service rate
- Policed: Arrival rate < Minimum Service Rate
- Not affected by other data PHBs
  ⇒ Highest data priority (if priority queueing)
- Code point: 101 110

Assured Forwarding



- PHB Group
- Four Classes: No particular ordering.
  $\Rightarrow$ Creates 4 distinct networks with specified QoS. Share unused capacity.
- Three drop preference per class

Raj Jain

Assured Forwarding (Cont)

- DS nodes SHOULD implement all 4 classes and MUST accept all 3 drop preferences. Can implement 2 drop preferences.
- Similar to nrt-VBR/ABR/GFR
- Code Points:

Drop Prec.	Class 1	Class 2	Class 3	Class 4
Low	001 010	010 010	011 010	100 010
Medium	001 100	010 100	011 100	100 100
High	001 110	010 110	011 110	100 110

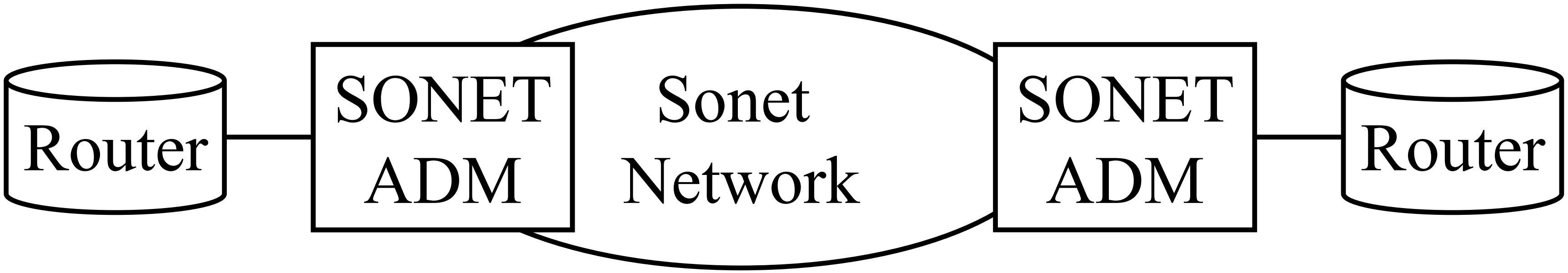
- Avoids xxx000 class selectors. Last bit 0 $\Rightarrow$ Standard

Per-Domain Behavior

- PDBs: Measurable edge to edge behavior across a cloud with same DS policies for all packets of a given PHB
- Existing PHBs have been extended to PDBs:
  - Virtual wire PDB: Based on EF
  - Assured Rate PDB: Based on AF. Min Rate. No delay or jitter guarantee
  - Bulk Handling PDB: Less than "Best Effort." Dropped if no resources. No need to police.

Problems with DiffServ

- per-hop $\Rightarrow$ Need at every hop
  One non-DiffServ hop can spoil all QoS
- End-to-end $\neq$ Σ per-Hop
  Designing end-to-end services with weighted guarantees at individual hops is difficult.
- How to ensure resource availability inside the network?

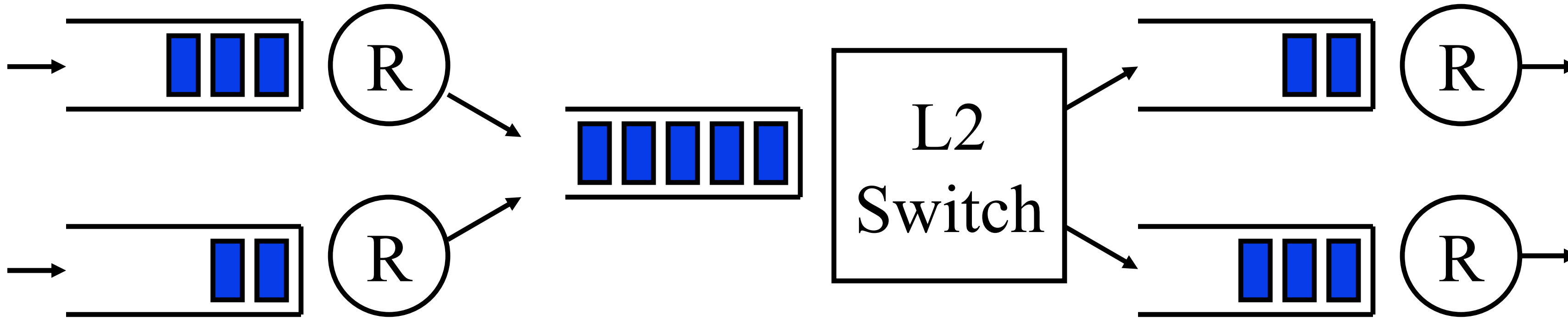


Raj Jain

DiffServ Problems (Cont)

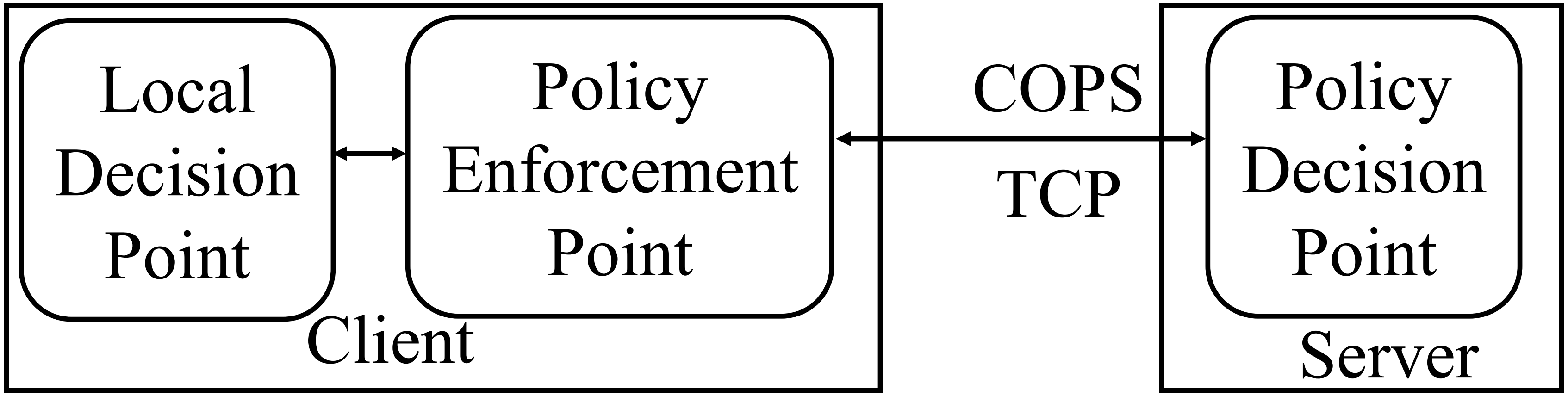
- QoS is for the aggregate not micro-flows.
  - Large number of short flows are better handled by aggregates.
  - High-bandwidth flows (1 Mbps video) need per-flow guarantees.
- Designed for static Service Level Agreements (SLAs) Both the network topology and traffic are highly dynamic.
- Need route pinning or connections.
- Not all DSCPs used by all vendors/providers. DSCPs rewritten at domain boundaries.

Subnet Bandwidth Manager (SBM)



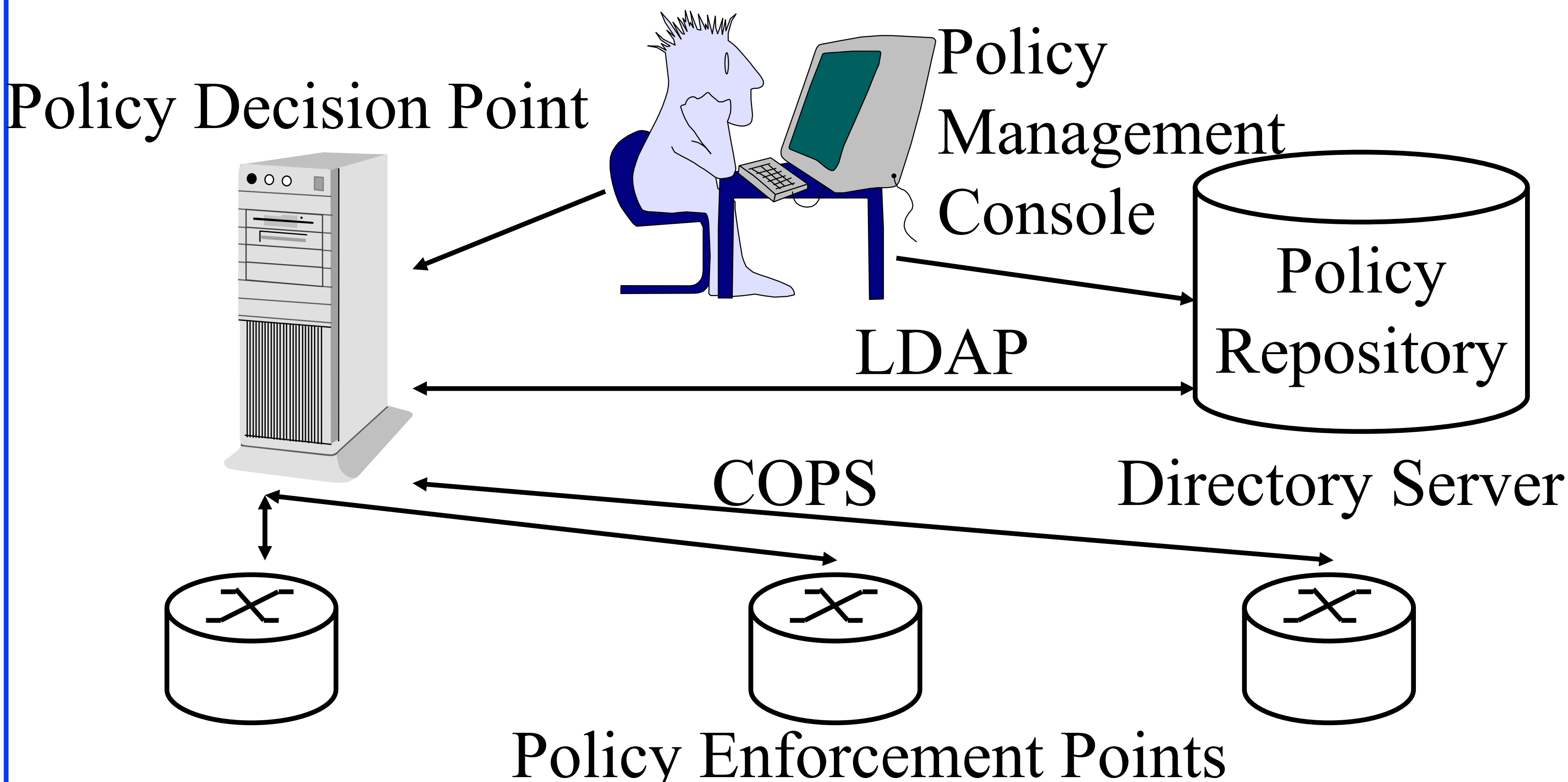
- Resources in a L2 switches may be a bottleneck
- SBM allows L2 switches to participate in RSVP admission control
- SBM capable switches and hosts elect a Designated SBM (DSBM)
- All RSVP messages are sent through DSBM

COPS Protocol



- Large and dynamic policy database ⇒ server
- Common Open Policy Service Protocol
- When the routers (clients) receive a RSVP message, they send the request the server and obtain authorization
- Will work with other (non-RSVP) signaling
- Routers can make local decisions when disconnected from PDP but should sync with PDP upon connection

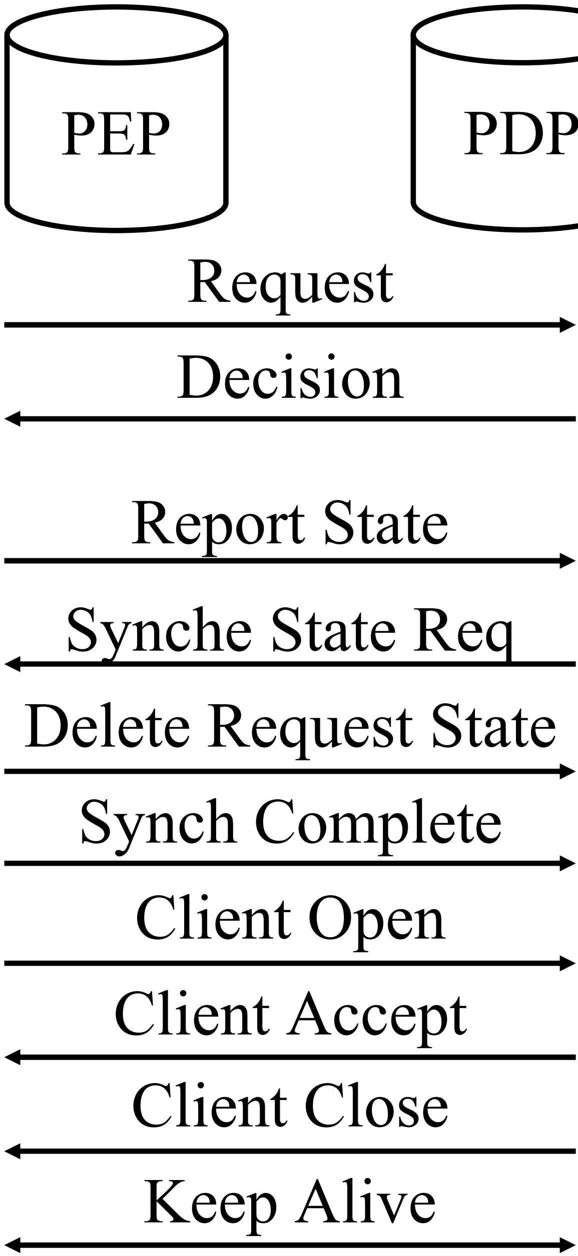
Policy Framework



- Policy based networking/Directory enabled networking (DEN)

Raj Jain

COPS Messages



- Request a decision regarding a client
- Decision
- Report success/failure of decision or accounting related changes
- Tell me about the state of this/all clients
- This client is no longer relevant
- Finished syncing all clients
- I can support these client types
- Client Ack. Here is the hello timer.
- Client Nack.
- Hello

Raj Jain

IEEE 802.1D Model

Dest Addr	Src Addr	Tag Prot ID	Pri	CFI	VLAN ID

802.1Q header (Tag Prot ID, Pri, CFI, VLAN ID)

Prot Type	Payload	FCS

CFI = Canonical Format Indicator (Source Routing)

- **Up to eight priorities:** Strict.

  1 Background

  2 Spare

  0 Best Effort

  3 Excellent Effort

  4 Control load

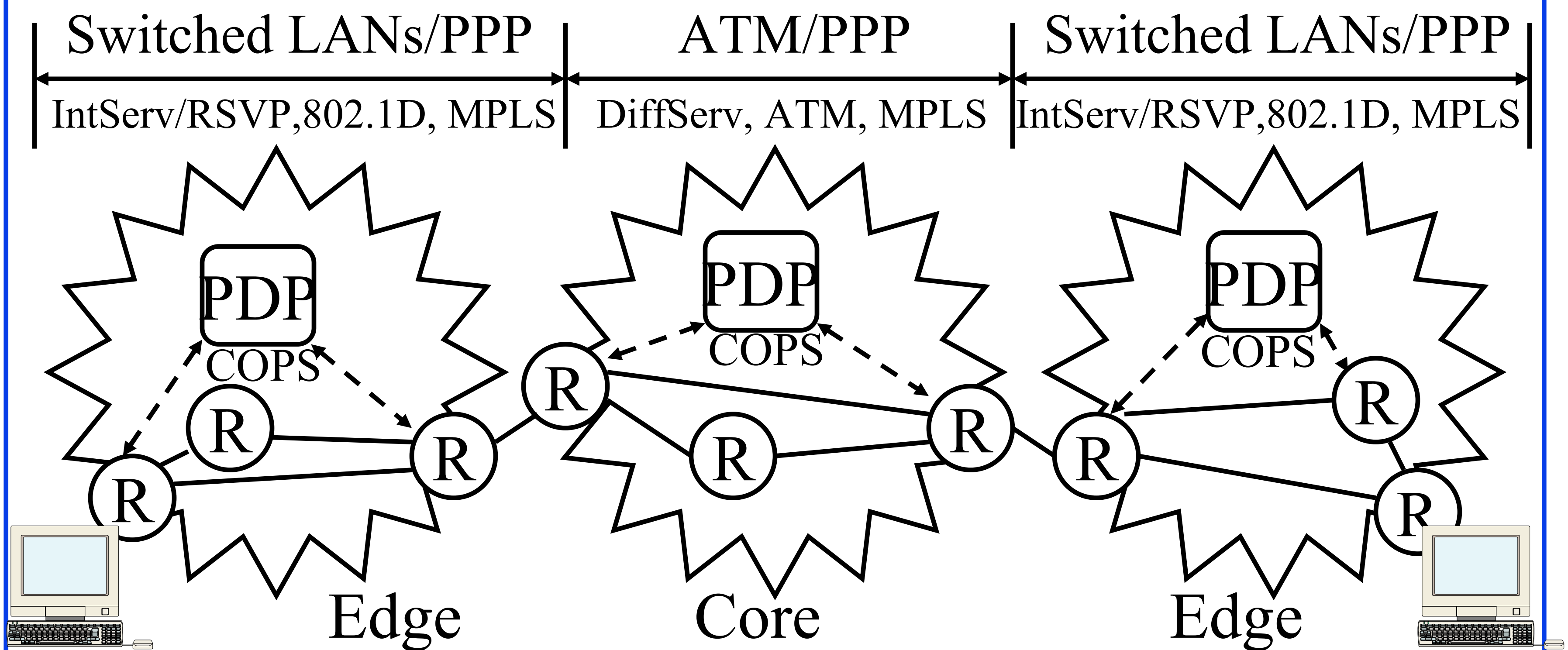
  5 Video (Less than 100 ms latency and jitter)

  6 Voice (Less than 10 ms latency and jitter)

  7 Network Control

End-to-end View

- ATM/PPP backbone, Switched LANs/PPP in Stub
- IntServ/RSVP, 802.1D, MPLS in Stub networks
- DiffServ, ATM, MPLS in the core



Raj Jain

QoS Implementation Example

- Windows 2000 has a QoS API
  - Uses RSVP to request bandwidth from network
  - Marks DSCP/802.1p
  - Netmeeting and Media Player can request QoS
- Most IP Phones do not directly support RSVP.
  - Need gatekeepers/proxies to request QoS
  - RSVP in near future to inter-operate with Netmeeting.

QoS Design Approaches

- Massive Bandwidth vs Managed Bandwidth
- Per-Flow vs Aggregate
- Source-Controlled vs Receiver Controlled
- Soft State vs Hard State
- Path based vs Access based
- Quantitative vs Qualitative
- Absolute vs Relative
- End-to-end vs Per-hop
- Static vs Feedback-based
- Homogeneous multicast vs heterogeneous multicast
- 1-to-n multicast vs n-to-1 multicast

Comparison of QoS Approaches

Issue	ATM	IntServ	DiffServ	MPLS	IEEE 802.1D
Massive Bandwidth vs Managed Bandwidth	Managed	Managed	Massive	Managed	Massive
Per-Flow vs Aggregate	Both	Per-flow	Aggregate	Both	Aggregate
Source-Controlled vs Receiver Controlled	Unicast Source, Multicast both	Receiver	Ingress	Both	Source
Soft State vs Hard State	Hard	Soft	None	Hard	Hard
Path based vs Access based	Path	Path	Access	Path	Access
Quantitative vs Qualitative	Quantitative	Quantitative +Qualitative	Mostly qualitative	Both	Qualitative
Absolute vs Relative	Absolute	Absolute	Mostly Relative	Absolute + relative	Relative

Raj Jain

Comparison (Cont)

Issue	ATM	IntServ	DiffServ	MPLS	IEEE 802.3D
End-to-end vs Per-hop	end-end	end-end	Per-hop	end-end	Per-hop
Static vs Feedback-based	Both	Static	Static	Static	Static
Homogeneous multicast vs heterogeneous multicast	Homo-geneous	Hetero-geneous	N/A	Homo-geneous	N/A
1-to-n vs n-to-1 multicast	1-to-n	1-to-n	N/A	Both	Both

Summary

1. ATM: CBR, VBR, ABR, UBR, GFR
2. Integrated Services: GS = rtVBR, CLS = nrt-VBR
3. Signaling protocol: RSVP
4. Differentiated Services will use the DS byte
5. 802.1D allows priority

Key References

- For a detailed list of references see:
  http://www.cis.ohio-state.edu/~jain/refs/ipqs_ref.htm
  Also reproduced in the back of this tutorial handout.
- QoS in Data Networks: Prototocols and Standards, http://www.cis.ohio-state.edu/~jain/cis788-99/qos_protocols/index.html
- Qos in Data Networks: Products, http://www.cis.ohio-state.edu/~jain/cis788-99/qos_products/index.html
- Integrated Services Overview, http://www.cis.ohio-state.edu/~jain/cis788-97/integrated_services/index.htm
- Multimedia over IP (RSVP, RTP, RTCP, RTSP), http://www.cis.ohio-state.edu/~jain/cis788-97/ip_multimedia/index.htm

Key References (Cont)

- QoS/Policy/Cinstraint Based Routing, http://www.cis.ohio-state.edu/~jain/cis788-99/qos_routing/
- QoS Forum, http://www.qosforum.com
- RSVP Project, http://www.isi.edu/div7/rsvp/rsvp.html

IETF Working groups:

- Diffserv, http://www.ietf.org/html.charters/diffserv-charter.html
- IntServ, http://www.ietf.org/html.charters/intserv-charter.html
- RSVP, http://www.ietf.org/html.charters/rsvp-charter.html
- Policy, http://www.ietf.org/html.charters/policy-charter.html
- ISSLL, http://www.ietf.org/html.charters/issll-charter.html

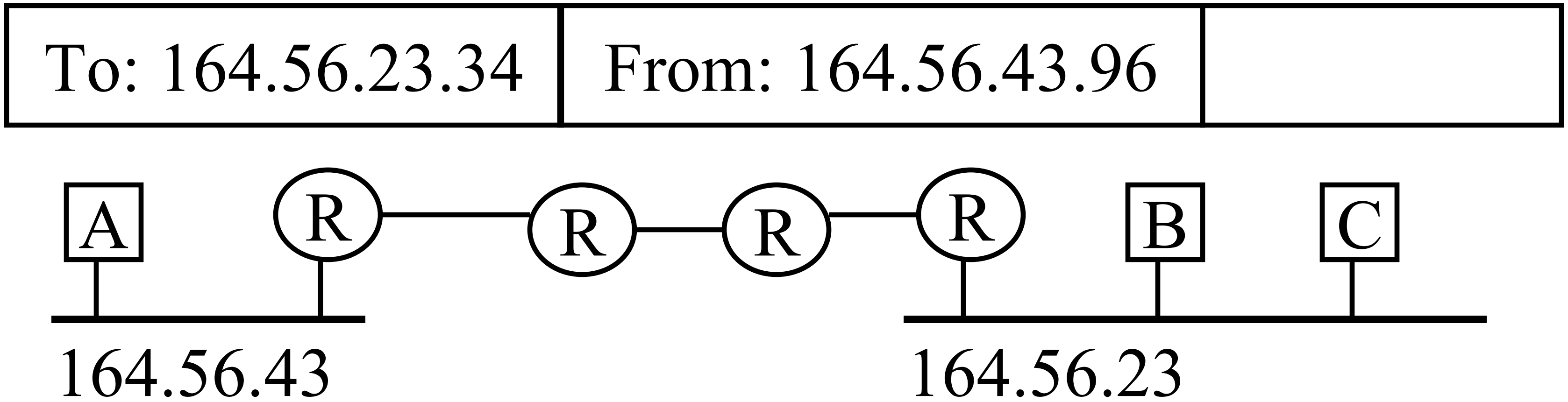
Multiprotocol Label Switching

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

- Routing vs Switching
- Tag Switching
- Multi-Protocol Label Switching
- Label Stacks
- Label Distribution Protocols: LDP, CR-LDP, RSVP-TE
- Traffic Engineering using MPLS

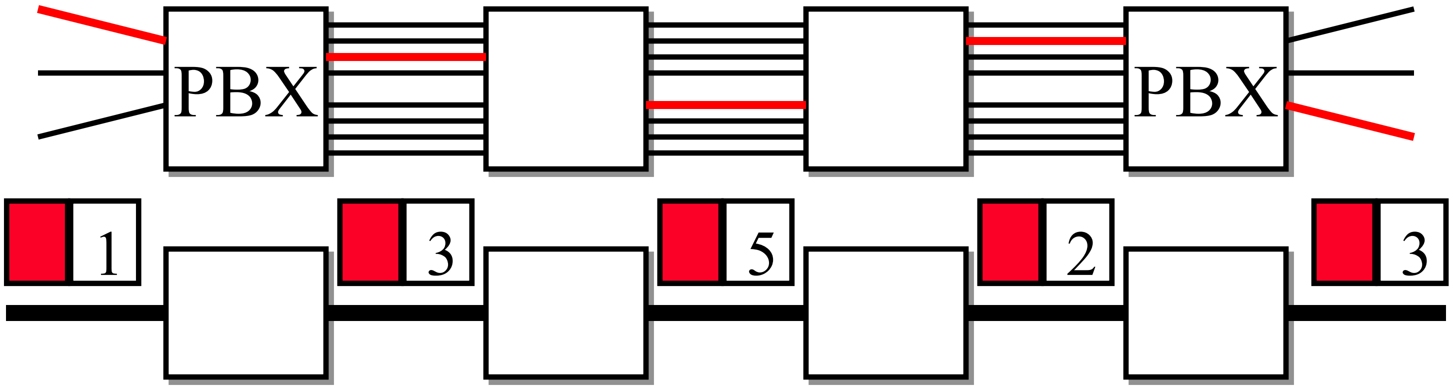
Raj Jain

IP Forwarding:Fundamentals



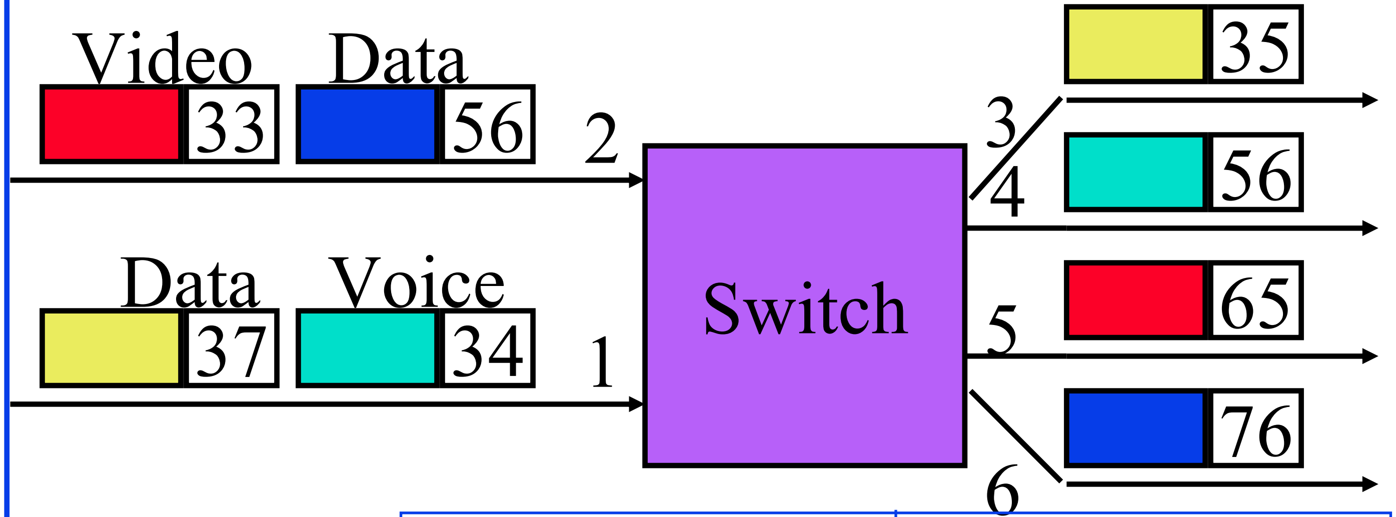
- IP = Internet protocol
- IP routers forward packets from source subnet (LAN) towards the destination subnet
- On the same subnet, routers are not required.
- Datagram or connectionless:
  Each packet has complete address.
  Each packet may potentially take a different path

Virtual Circuits



- Circuit = Line 1, 3, 5, 2, 3
- Virtual Circuit = VCI 1, 3, 5, 2, 3
- Used in X.25: Up to 4095 VCs on one X.25 interface
- Each packet or cell has a VC Id in the header
- A group of VCs is called "Virtual Path" (VP)

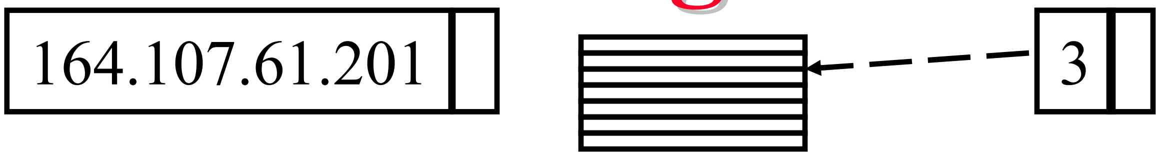
ATM VC Switching



In		Out	
Port	VCI	Port	VCI
1	37	3	35
1	34	4	56
2	33	5	65
2	56	6	76

Raj Jain

Routing vs Switching

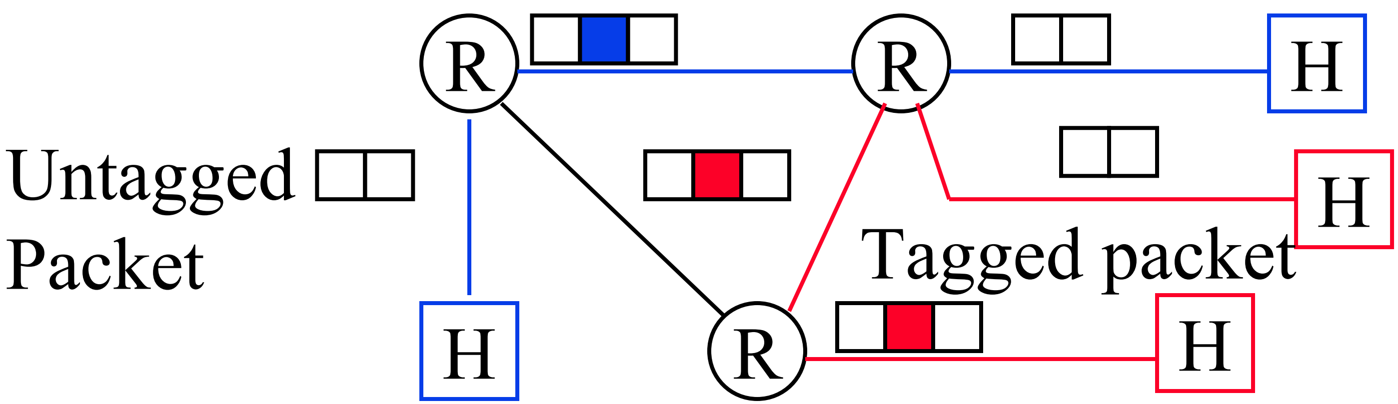


- Routing: Based on address lookup. Max prefix match.
  $\Rightarrow$ Search Operation
  $\Rightarrow$ Complexity $\approx O(\log_2 n)$
- Switching: Based on circuit numbers
  $\Rightarrow$ Indexing operation
  $\Rightarrow$ Complexity $O(1)$
  $\Rightarrow$ Fast and Scalable for large networks and large address spaces
- These distinctions apply on all datalinks: ATM, Ethernet, SONET

Tag Switching

- Proposed by CISCO
- Similar to ATM VC numbers
- Tags can be explicit or implicit L2 header

- Ingress router/host puts a tag. Exit router strips it off.

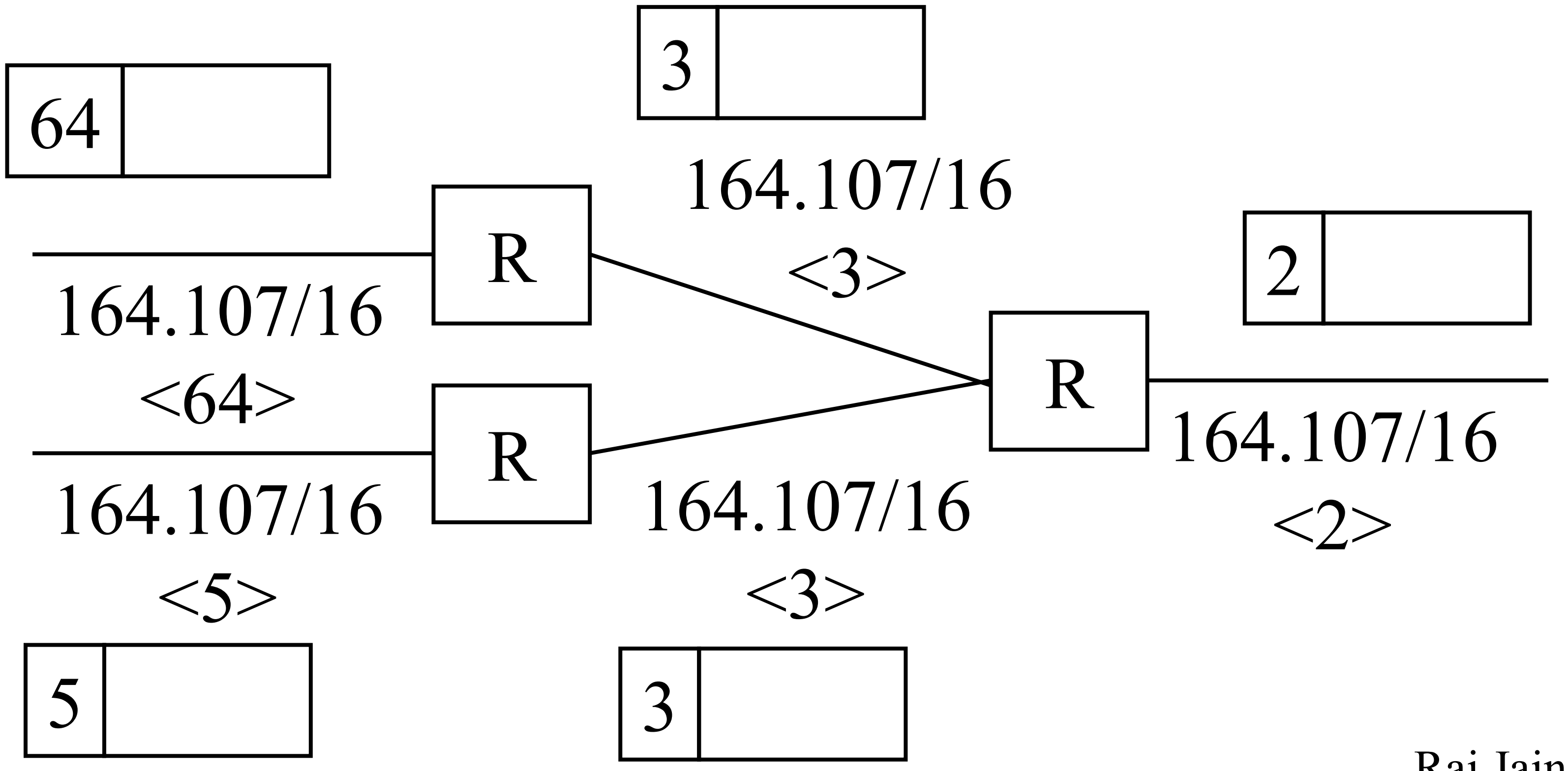


Tag Switching (Cont)

- Switches switch packets based on labels.
  Do not need to look inside ⇒ Fast.
- One memory reference compared to 4-16 in router
- Tags have local significance
  ⇒ Different tag at each hop (similar to VC #)

Tag Switching (Cont)

- One VC per routing table entry



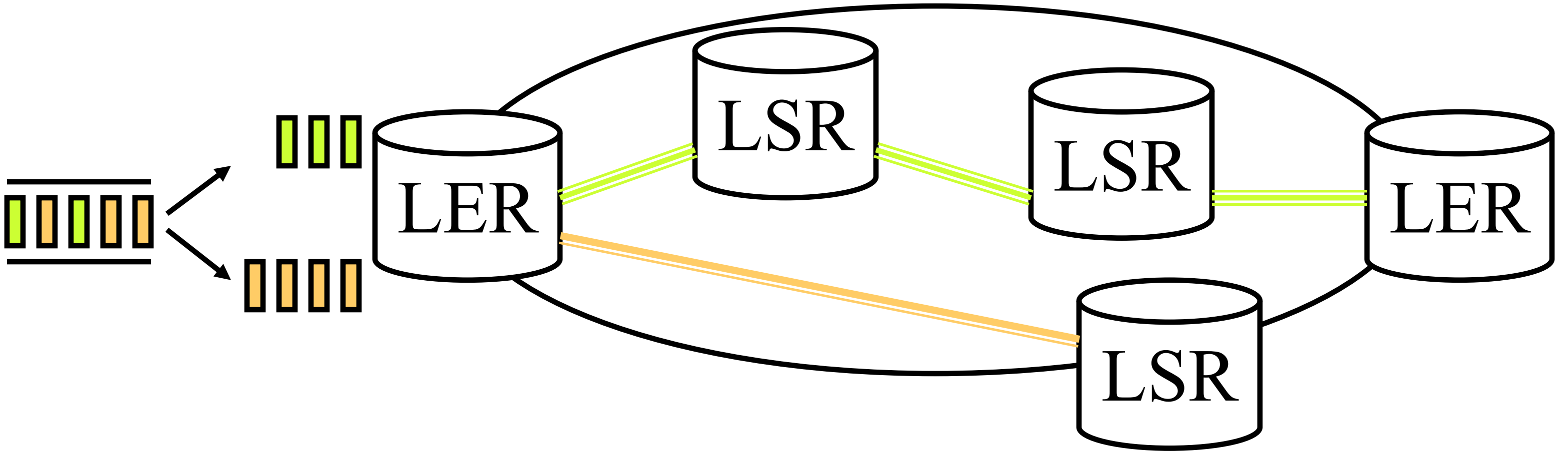
Raj Jain

MPLS

- Multiprotocol Label Switching
- IETF working group to develop switched IP forwarding
- Initially focused on IPv4 and IPv6. Technology extendible to other L3 protocols.
- Not specific to ATM. ATM or LAN.
- Not specific to a routing protocol (OSPF, RIP, ...)

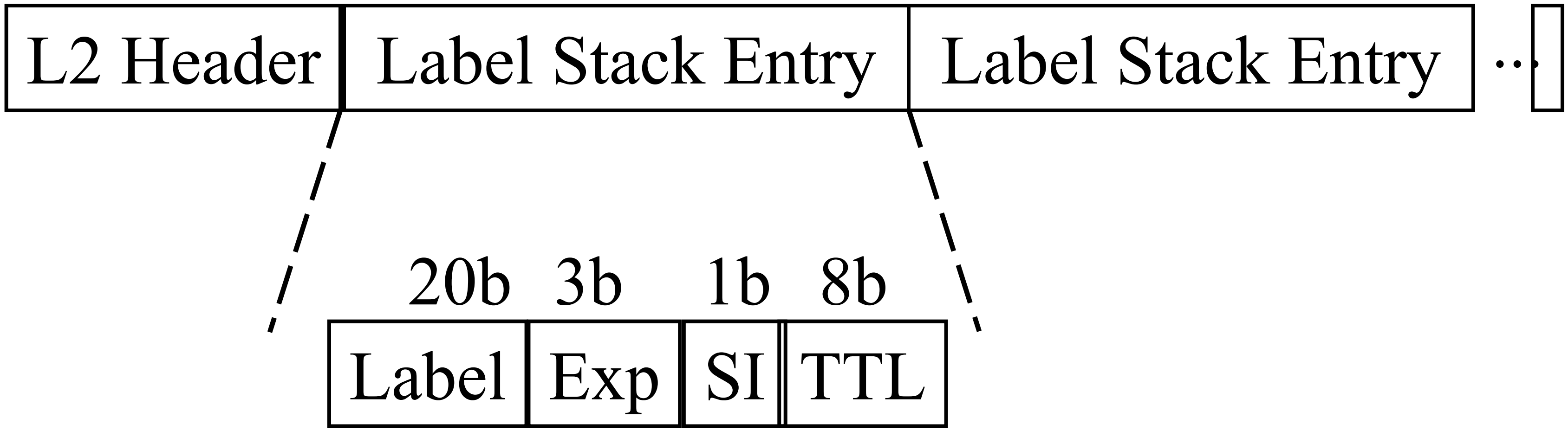
MPLS Terminology

- Label Edge Router (LER)
- Label Switching Router (LSR)
- Label Switched Path (LSP)
- Forwarding Equivalence Class (FEC)



Label Stack Entry Format

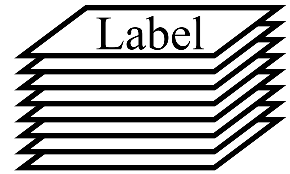
- Labels = Explicit or implicit L2 header
- TTL = Time to live
- Exp = Experimental
- SI = Stack indicator



Raj Jain

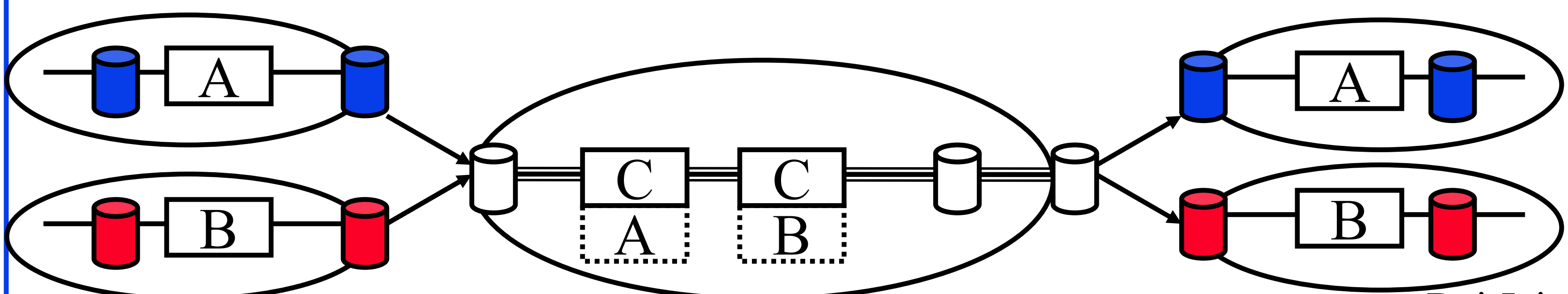
Label Stacks

- 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)

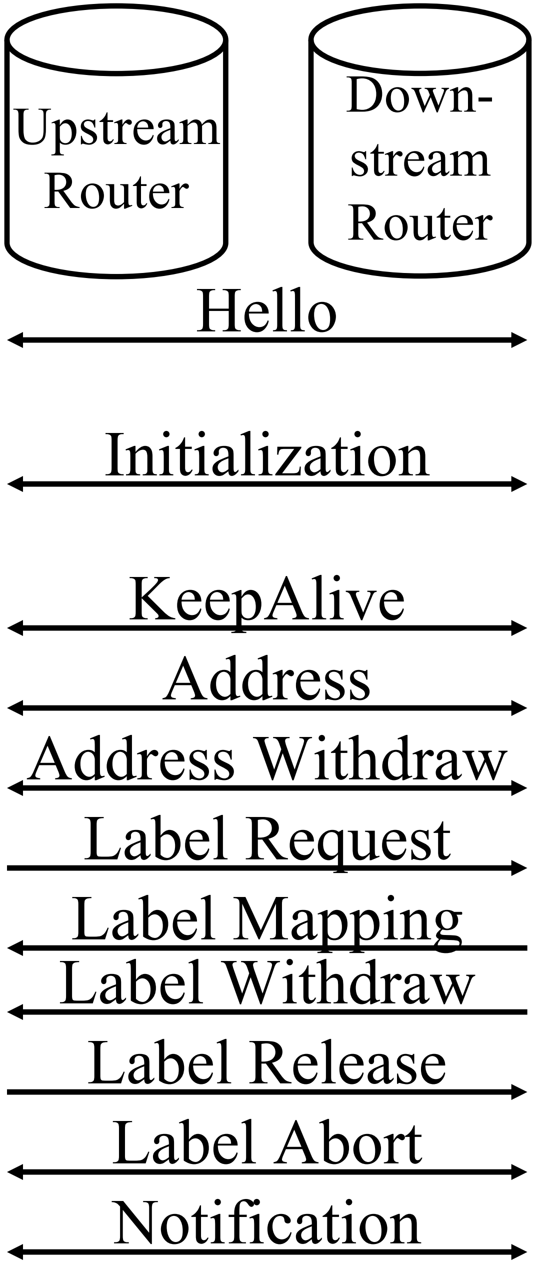
L2 Header	LSE 1	LSE 2	···	LSE n	



Label Assignment

- Unsolicited: Topology driven ⇒ 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 ⇒ latency
- A new Label Distribution Protocol called LDP is being defined.
- RSVP is being extended to allow label request and response

LDP Messages



- Hello: Sent periodically to "all routers on this subnet"
- Initialization: parameters and capability exchange
- KeepAlive: Keep the LDP session alive
- Address (advertise interface addresses)
- Address Withdraw
- Label Request
- Label Mapping (Label Response)
- Label Withdraw (by downstream)
- Label Release (by upstream)
- Label Abort Request
- Notification: Error/advisory information

CR-LDP

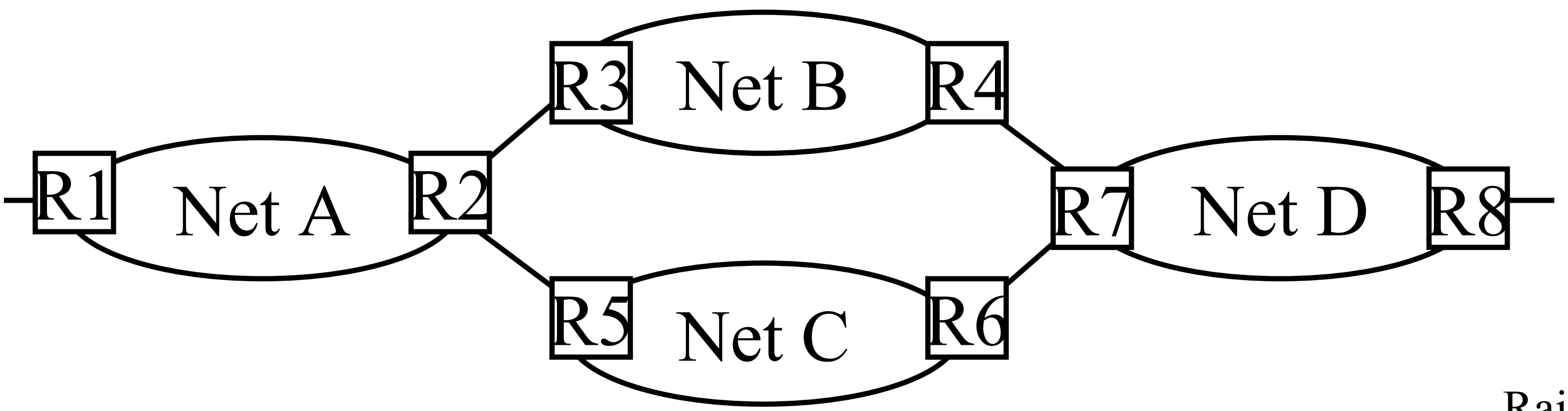
- Extension of LDP for constraint-based routing (CR)
- New Features:
  - Traffic parameters
  - Explicit Routing
  - Preemption of existing route. Based on holding priority of existing route and setup priority of new route
  - Route pinning: To prevent path changes
- No new messages
- Enhanced Messages: Label request, Label Mapping, Notification

RSVP Extensions

- Explicit Route Object (ERO): Path messages are forced to go along specified explicit route
- Record Route
- Message Bundling: Multiple messages in one packet
- Refresh Reduction: Srefresh refreshes all reservations related to a given message ID
- Node Failure Detection: Keep alive hello messages
- Aggregation: Resv messages include diffserv marking (DSCP code) or 802.1p tag for the upstream node
- Security: Flow = Dest IP + IPSec Protocol Type + Security Parameter Index (SPI) = Security Association

Explicit Route

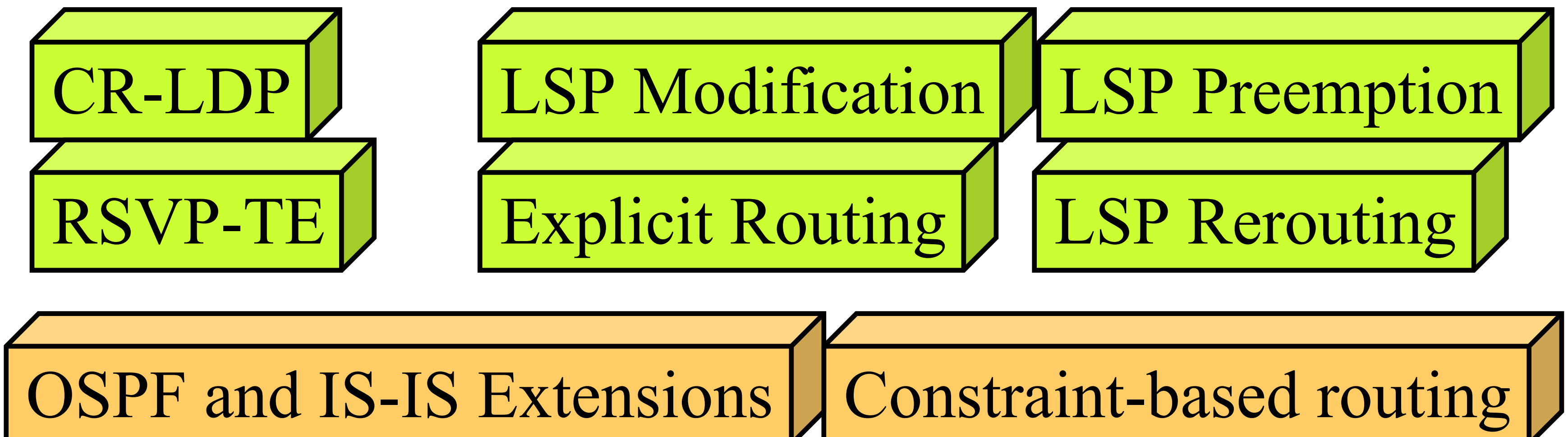
- Explicit route specified as a list of Explicit Route Hops (group of nodes)
- Hops can include IPv4 prefix, IPv6 prefix, MPLS tunnels or Autonomous systems
- Example: R1-R2-Net B-R7-R8
- Allows traffic engineering



Raj Jain

Traffic Engineering Building Blocks

- TE = Directing the traffic to where the capacity exists
- CR-LDP and RSVP-TE allow LSP explicit routing, rerouting, modification, preemption.
- OSPF and IS-IS are being modified to allow constraints



Summary

1. Switching = forwarding based on label indexing
2. Tag switching: Tags ≈ ATM's VC id
3. MPLS allows label stacks, TTL, QoS
4. MPLS signaling via RSVP, LDP, CR-LDP
5. Traffic engineering using explicit paths

Label Switching: Key References

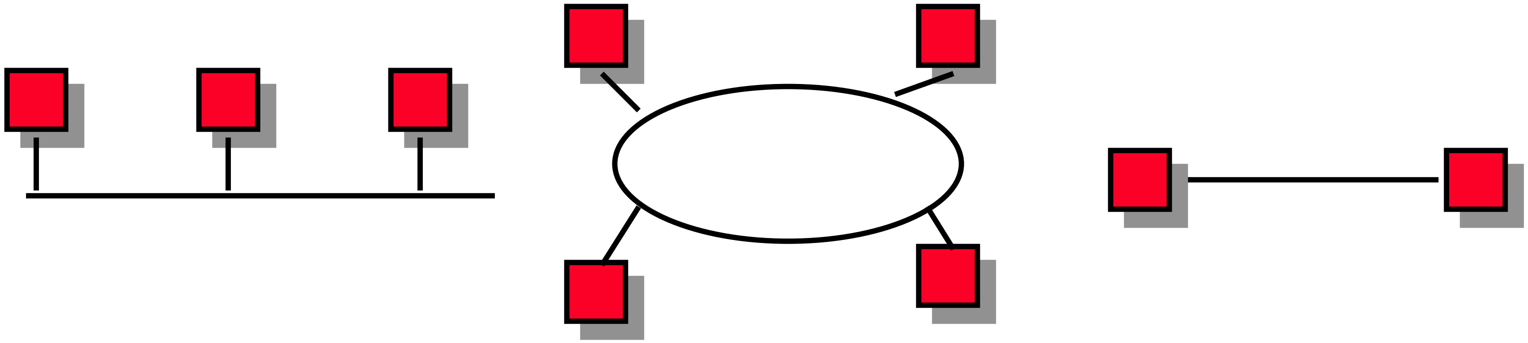
- See http://www.cis.ohio-state.edu/~jain/refs/ipsw_ref.htm
  Also reproduced at the end of this tutorial book.
- Multiprotocol Label Switching (mpls) working group at IETF. Email: mpls-request@cisco.com
- IP Switching, http://www.cis.ohio-state.edu/~jain/cis788-97/ip_switching/index.htm
- IP Switching and MPLS, http://www.cis.ohio-state.edu/~jain/ cis777-00/g_fipsw.htm
- MPLS Resource Center, http://www.mplsrc.com

Gigabit and 10 Gigabit Ethernet

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

- Distance-B/W Principle
- Gigabit MAC issues: Carrier Extension, Frame Bursting
- 10 GbE: Key Features, PMD Types
- 1G/10G Ethernet Switch Features
- Flow Control, Link Aggregation, Jumbo Frames
- Resilient Packet Rings
- Beyond 10 GbE

Distance-B/W Principle

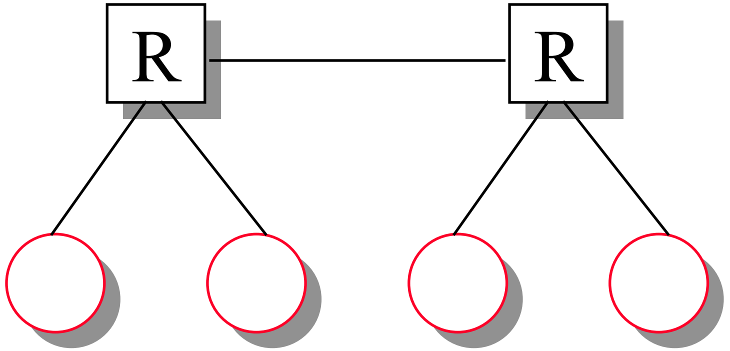


- Efficiency = Max throughput/Media bandwidth
- Efficiency is a non-increasing function of α
  α = Propagation delay /Transmission time
  = (Distance/Speed of light)/(Transmission size/Bits/sec)
  = Distance×Bits/sec/(Speed of light)(Transmission size)
- Bit rate-distance-transmission size tradeoff.
- 100 Mb/s ⇒ Change distance or frame size

Raj Jain

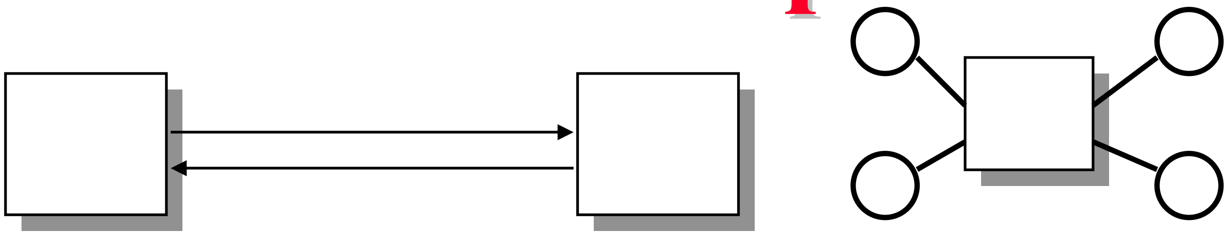
Ethernet vs Fast Ethernet

	Ethernet	Fast Ethernet
Speed	10 Mbps	100 Mbps
MAC	CSMA/CD	CSMA/CD
Network diameter	2.5 km	205 m
Topology	Bus, star	Star
Cable	Coax, UTP, Fiber	UTP, Fiber
Standard	802.3	802.3u
Cost	X	2X



Raj Jain

Full-Duplex LANs



- Uses point-to-point links between TWO nodes
- Full-duplex bi-directional transmission
  Transmit any time
- No collisions $\Rightarrow$ 50+ Km on fiber.
- Commonly used between servers and switches or between switches

How Much is a Gbps?

- 622,000,000 bps = OC-12
- 800,000,000 bps (100 MBps Fiber Channel)
- 1,000,000,000 bps
- 1,073,741,800 bps = 2^{30} bps (2^{10} = 1024 = 1k)
- 1,244,000,000 bps = OC-24
- 800 Mbps ⇒ Fiber Channel PHY
  ⇒ Shorter time to market
- Decision: 1,000,000,000 bps ⇒ 1.25 GBaud PHY
- Not multiple speed ⇒ Sub-gigabit Ethernet rejected
- 1000Base-X

Media Access Control Issues

- Carrier Extension
- Frame Bursting

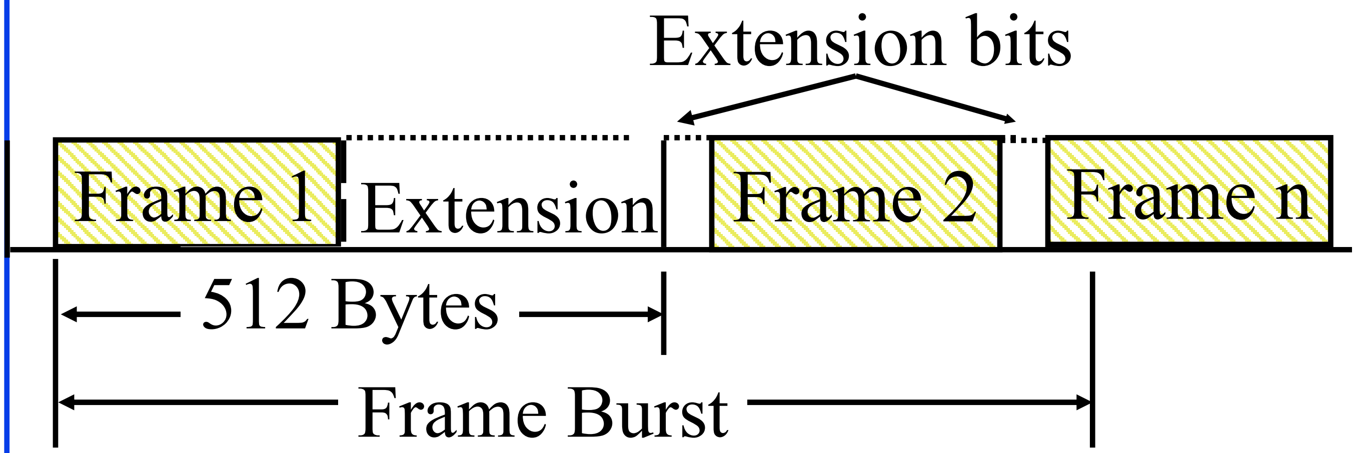
Raj Jain

Carrier Extension

- 10 Mbps at 2.5 km ⇒ Slot time = 64 bytes
- 1 Gbps at 200 m ⇒ Slot time = 512 bytes
- Continue transmitting control symbols. Collision window includes the control symbols
- Control symbols are discarded at the destination
- Net throughput for small frames is only marginally better than 100 Mbps

Raj Jain

Frame Bursting



- Don't give up the channel after every frame
- After the slot time, continue transmitting additional frames (with minimum inter-frame gap)
- Interframe gaps are filled with extension bits
- No no new frame transmissions after 8192 bytes
- Three times more throughput for small frames

Raj Jain

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

Maximum Distances for GbE

λ	Fiber	Core	Bandwidth	Attenu-ation	Dist.
nm		µm	MHz/km	dB/km	m
850	MMF	50	400	3.25	500
			500	3.43	550
		62.5	160	2.33	220
			200	2.53	275
1300	MMF	50	400/500	2.32	550
		62.5	500	2.32	550
	SMF	10	∞	4.5	5000

- All distances full duplex. Actual distances longer.

1000Base-T

- 100 m on 4-pair Cat-5 UTP
  ⇒ Network diameter of 200 m
- Applications
  - Server farms
  - High-performance workgroups
  - Network computers
- 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

1000BASE-T (Cont)

- Automatically detects and corrects pair-swapping, incorrect polarity, differential delay variations across pairs
- Autonegotiation $\Rightarrow$ Compatibility with 100Base-T
- 802.3ab-1999

How Much is 10 Gbps?

- 10,000,000,000 b/s
- 9,584,640,000 b/s (OC-192 payload rate)
- Both were accepted.
- LAN PHY at 10.000 Gbps
- WAN PHY at OC-192c payload rate 9.584640 Gbps
- Pacing Mechanism to adapt from LAN to WAN
  One extra byte in the inter-frame gap for every 13 bytes
- Both PHYs use the same MAC

10 GbE: Key Features

- P802.3ae ⇒ Update to 802.3
- Compatible with OC-192c Payload rate
- Compatible with 802.3 frame format, services, management
- LAN and WAN PHY families
- Cost = 3 × 1GbE
- Same min and max frame size as 10/100/1000 Mbps
- Full-duplex only ⇒ No CSMA/CD
- Star-wired point-to-point links
- 10.000 Gb/s at MAC interface

10GbE Physical Layer Proposals

Techno-logy	Wave-length	Cost	62.5 μm 160MHz	50 μm 400MHz	50 μm 2GHz	SMF	
Serial	1550 nm	5X	NA	NA	NA	40 km	√
Serial	1310 nm	1.8X	86 m	86 m	86 m	10 km	√
Serial	850 nm	X	28 m	65 m	300 m	NA	√
WWDM	1310 nm	3X	300 m	300 m	300 m	10 km	√
WWDM	850 nm	1.5X	100 m	300 m	550 m	NA	×

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

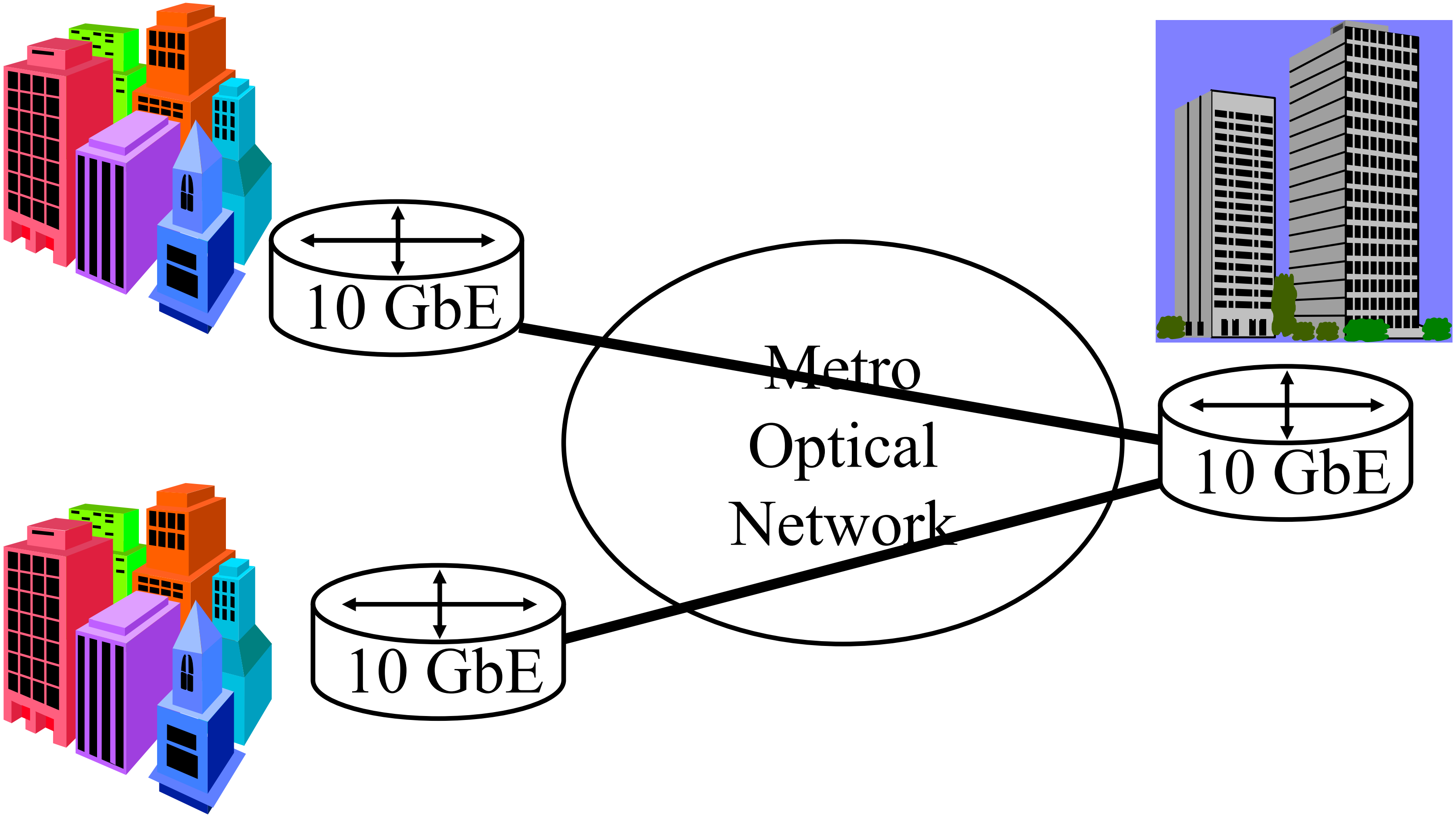
Raj Jain

10 GbE PMD Types

PMD	Description	MMF	SMF
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-LX4	1310nm WWDM LAN	300 m	10 km
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

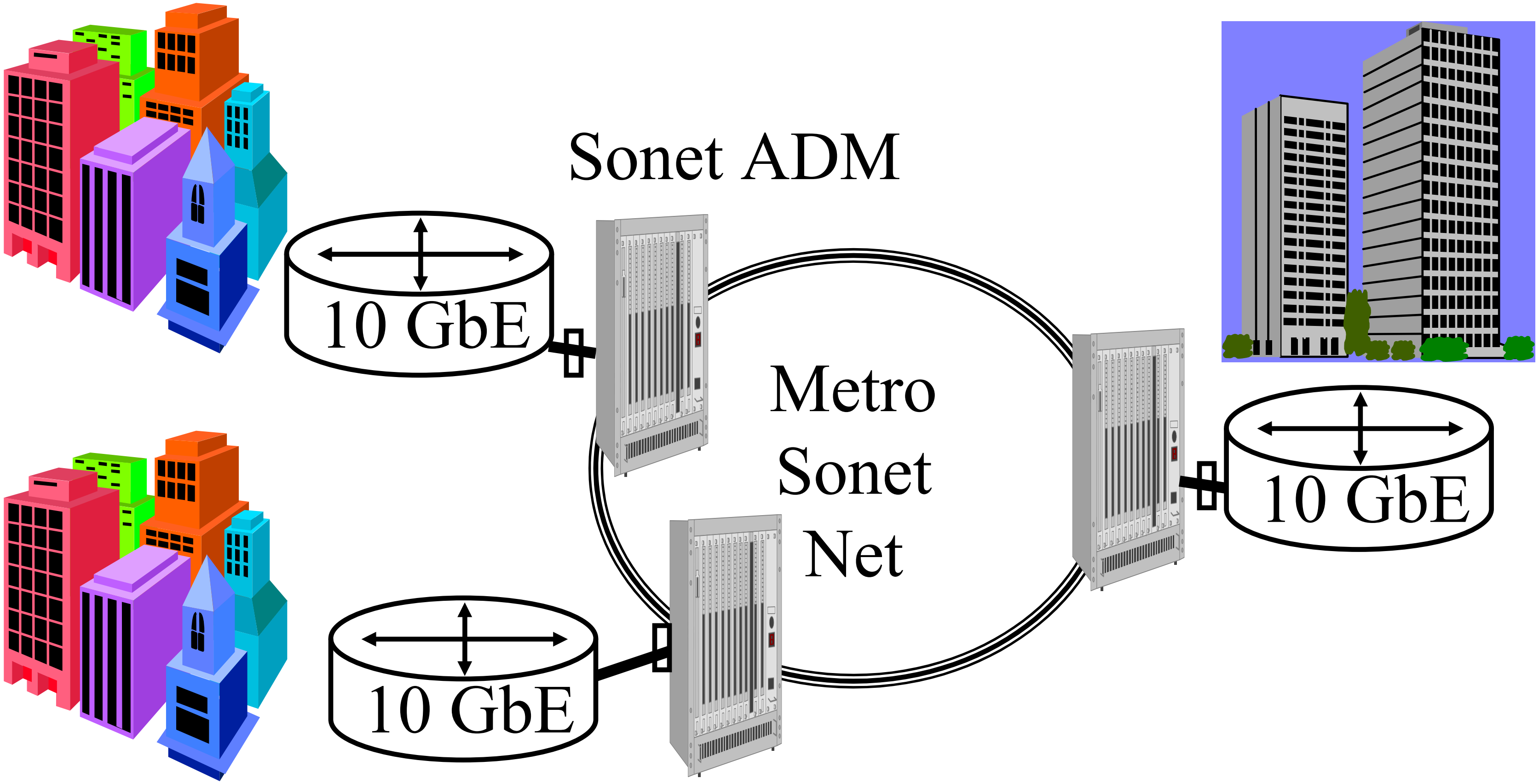
10 GbE over Dark Fiber



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

Raj Jain

10 GbE over Sonet/SDH



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

Raj Jain

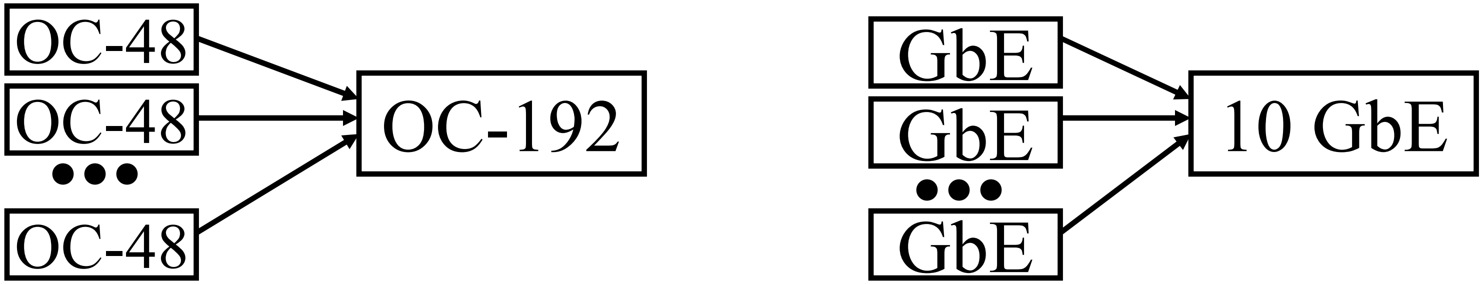
GbE vs 10 GbE

Feature	GbE	10 GbE
Media	UTP + Fiber	Fiber only
Distance	Up to 5 km	Up to 40km Longer distances via SONET
PMD	Updated Fiber Channel PMD	4 New PMDs (Being adapted by 10 G FC)
Coding	8b/10 coding	64b/66b coding for serial 8b/10b for WWDM
MAC	Half-Duplex + Full-Duplex	Full-Duplex only
MAC Extensions	Carrier Extension for CSMA/CD	Flow control for LAN/WAN speeds

10GbE Schedule

- Mar 1999: HSSG Formed
- Jan 2000: PAR Approved
- Sep 2000: First Draft
- Mar 2001: WG Ballot
- June 2001: Draft D3.1,
  Most discussion on Jitter and clock stability specs
  On Schedule
- Mar 2002: Standard

10 GbE Ethernet vs SONET OC-192



- SONET is isosynchronous
  10 GbE is a plesio-isosynchronous
- All transmitters of SONET have clocks synchronized to a common clock source. Bit muxed.
- All transmitters of Ethernet have independent clocks. Clock recovered from the received signal.Pkt Muxed. Interpacket gap is used to adjust clock differences.
- 10 GbE clock jitter is ±100 ppm (Draft 3.1).
  SONET requires 4.6 to 20 ppm.

Ethernet vs Sonet

Feature	SONET	Ethernet
Bit Rate (bps)	155 M, 622 M, 2.5 G, 10 G, 40 G, …	1M, 10 M, 100 M, 1 G, 10 G, …
Timing	Isochronous (Periodic 125μs)	Plesio-Isochronous
Multiplexing	Bit	Packet
Clocks	Common	Independent
Clock jitter	*4.6 to 20* ppm	100 ppm (May change)
Usage	Telecom	Enterprise
Volume	Millions	100's of Millions
Price (10 Gbps)	>10k	≈1k
Recovery	*50 ms*	Few Minutes
Topology	*Rings*	Mesh

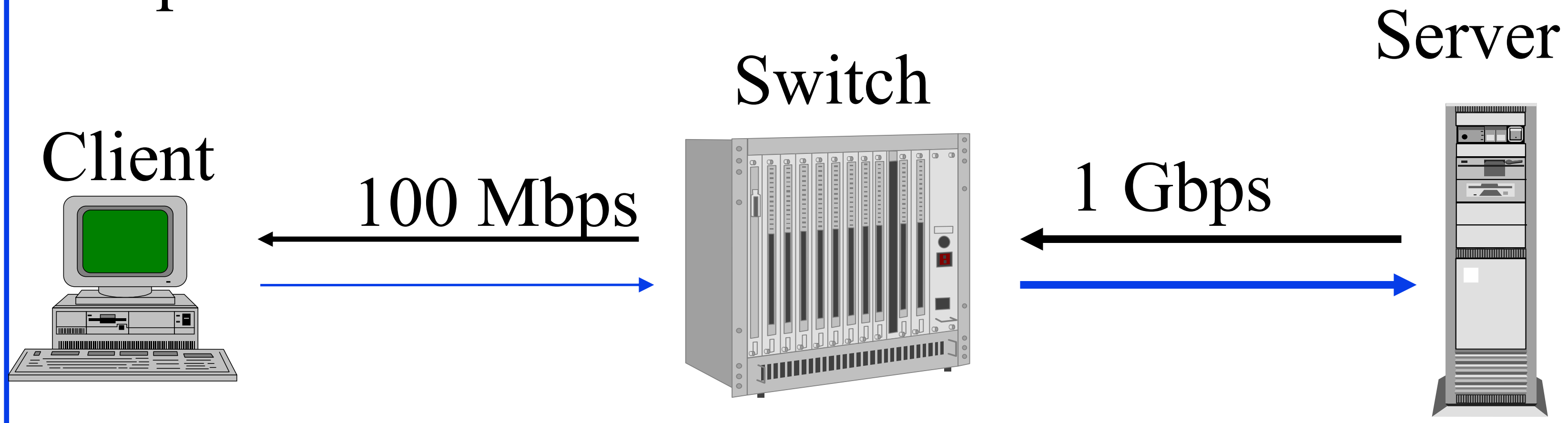
Raj Jain

1G/10G Ethernet Switch Features

- Stackable or Standalone
- Blocking or non-blocking
- Number of 10/100/1000/10G Ports
- Other LAN ports: ATM, FDDI
- Quality of Service: 802.1p+802.1Q, RSVP, WFQ
- Virtual LAN Support: 802.1Q, port, MAC, L3
- Layer 3 Switching: IP, IPX,AppleTalk
- Flow Control: 802.3x
- Link Aggregation
- Jumbo Frames

802.3x Full-Duplex Flow Control

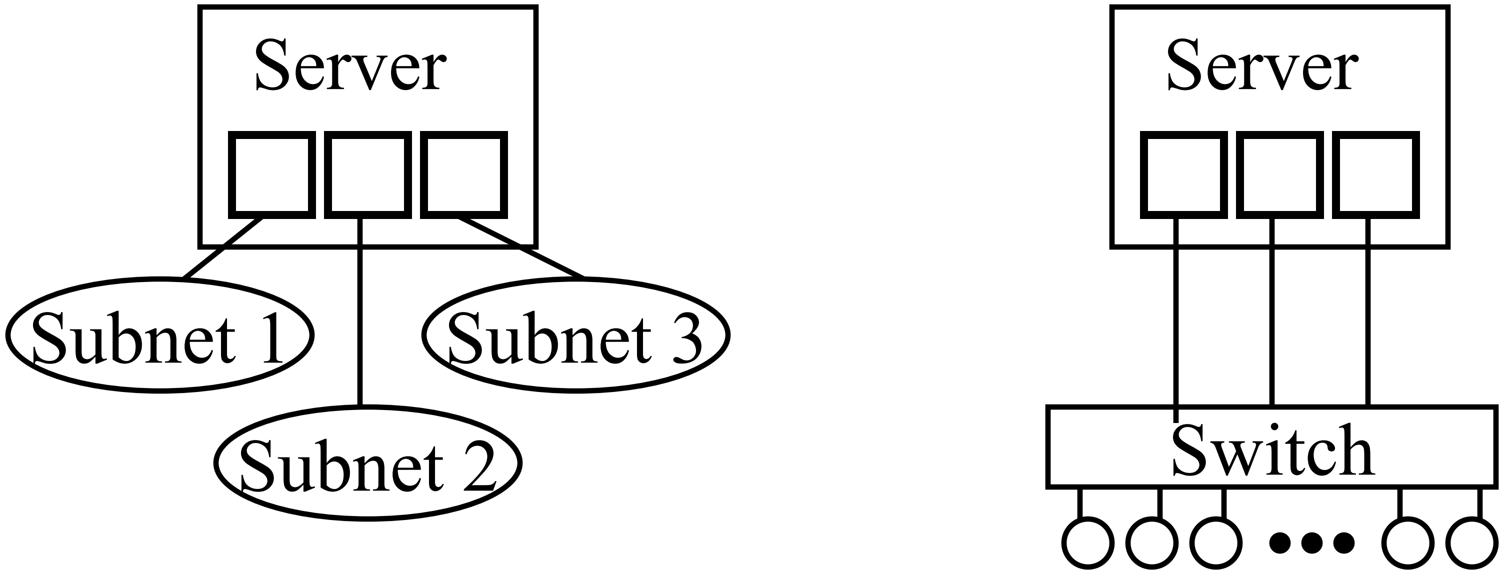
- Pause frame with pause time sent to multicast address 01-80-C2-00-00-01 not forwarded by bridges
- Autonegotiation updated to include a "flow-control capable" bit



1. Client disk full	2. Client sends Pause	3. Switch buffers fill up	4. Switch sends Pause	5. Server Stops

Raj Jain

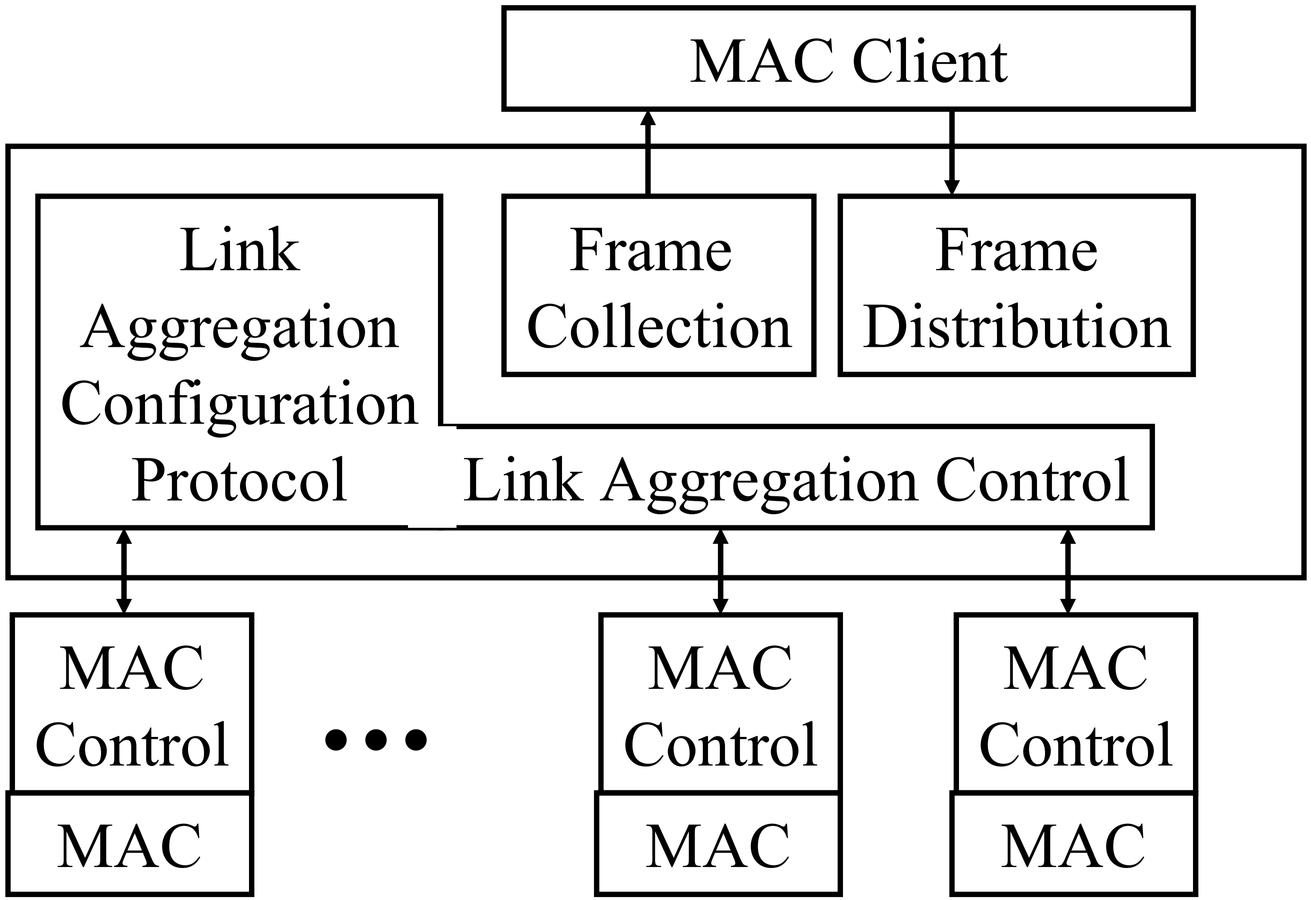
802.3ad Link Aggregation



- Allows n parallel links to act as one link
  ⇒ Server needs only one IP address.
- For redundancy and incremental bandwidth
- Cost < nX
- Ideal up to 4 links. Approved March 2000.

Raj Jain

Link Aggregation Sublayer



Raj Jain

Jumbo Frames

- Maximum Ethernet Frame Size = 1518 bytes or 1522 bytes (with VLAN Tags)
- Frame size too small at Gbps and higher speed
- 9kB implemented by Alteon WebSystems
- 9k-16kB being talked about in the industry
- Will not be an IEEE standard
- Ref: http://www.nwfusion.com/newsletters/lans/0614lan1.html

Future Possibilities

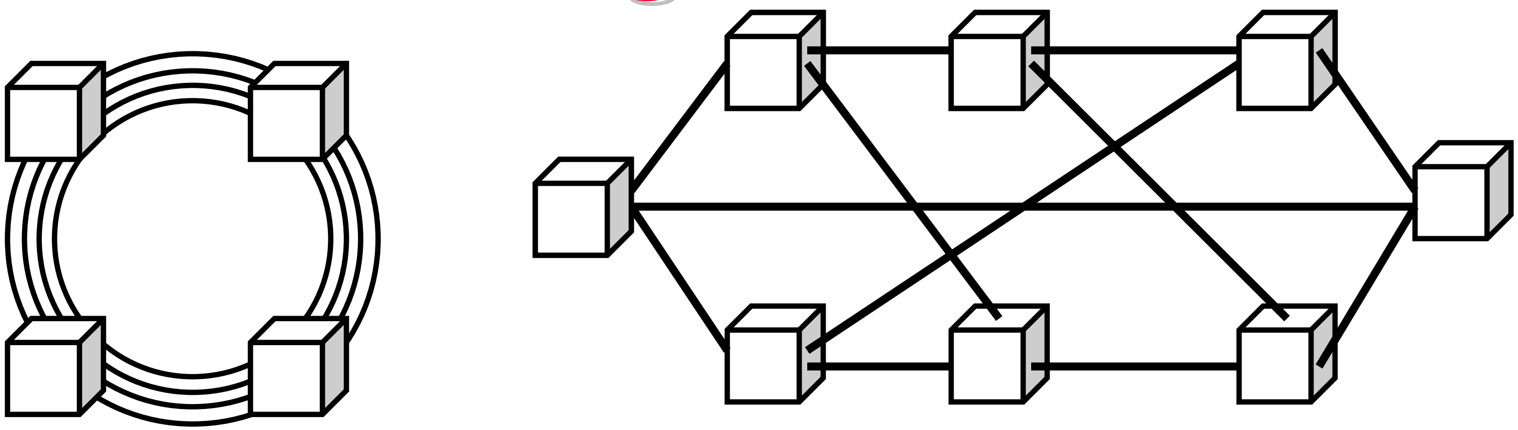
- 40 Gbps
- 100 Gbps:
  - 16λ×6.25 Gbps
  - 8λ × 12.5 Gbps
  - 4λ × 12.5 using PAM-5
- 160 Gbps
- 1 Tbps:
  - 12 fibers with 16λ × 6.25 Gbps
  - 12 fibers with 8λ × 12.5 Gbps
- 70% of 802.3ae members voted to start 40G in 2002

Networking and Religion

Both are based on a set of beliefs

Raj Jain

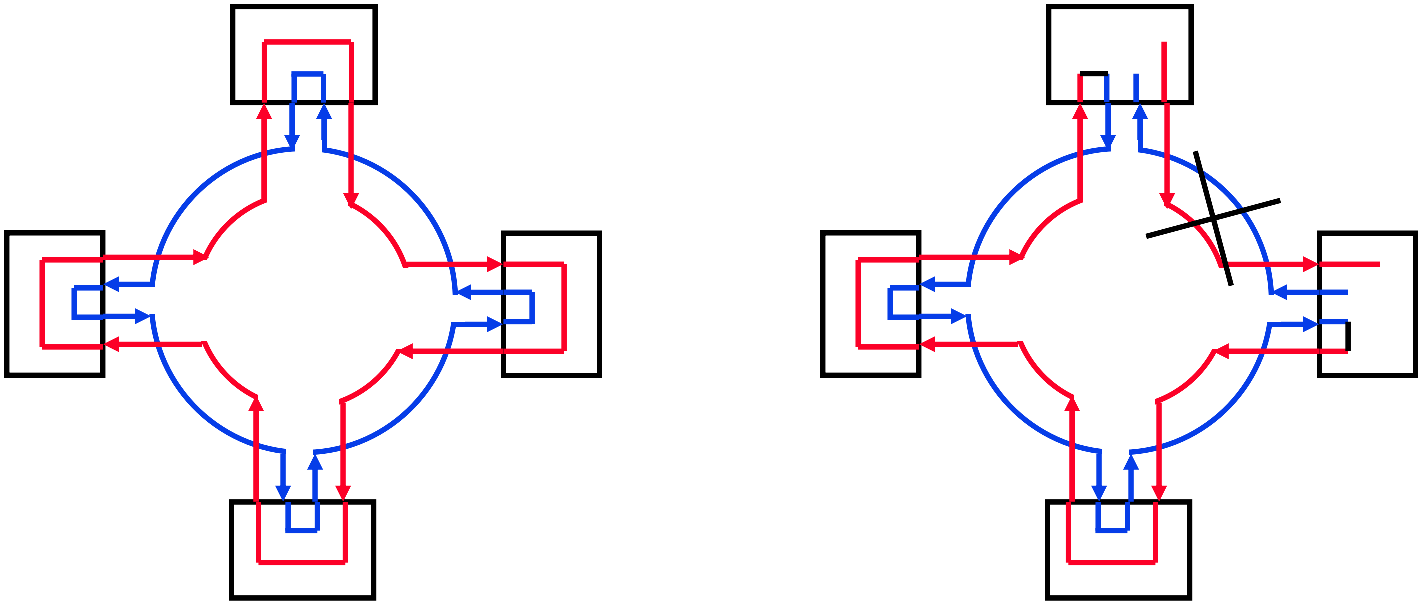
Ring vs Mesh



- On rings: All links same capacity ⇒ Not good for non-homogeneous or long-distance traffic
- Upgrade: All stations on the ring must be upgraded.
- Mesh typically requires 50% less restoration and 50% less working capacity than rings
- Mesh save more as degree of connectivity increases

Raj Jain

Resilient Packet Rings



- Dual Counter-rotating rings help protect against failure
- Used in SONET and FDDI
- Need to bring these concepts to Ethernet and IP

Raj Jain

RPR (Cont.)

- IEEE 802.17 Working Group. Started March 2001. First draft January 2002.
- Support packets of 64B to 1522+
- Destination stripping of unicast packets
  Source stripping of multicast-broadcast packets
- Support many Physical layers:
  SONET OC-3/12/48/192
  1000/10000 Ethernet LAN/WAN
- Fair and dynamic distribution of available bandwidth
- Distributed congestion management
- Data packets on both rings, corresponding control in the opposite direction.

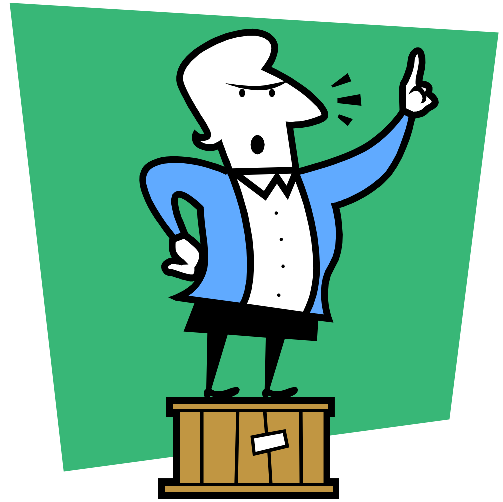
RPR (Cont)

- Support for guaranteed and committed data rates and delays
- Priority in the station. Priority on the ring
- Automatic topology discovery. No provisioning.
- Fast recovery. Sub 50ms restoration
- Coordination with layer 3
- Ref: http://www.ieee802.org/17/index.html
- IETF IPoRPR working group: L2-L3 interactions Started December 2000, http://www.ietf.org/html.charters/iporpr-charter.html

RPR (Cont)

- Products announcements from Mindspeed Technologies, and Connexant
- Rpralliance charter members: Cisco, Dynarc, Lantern Communications, Luminous, and Nortel, www.rpralliance.com

Summary



- Gigabit Ethernet runs at 1000 Mbps
- Standard allows both shared and full-duplex links
- 10 GbE for full duplex LAN and WAN links
- 1000 Mbps and 9,584.640 Mbps
- RPR will make it more suitable for Metro
- Higher speed are also coming...

GbE and 10 GbE: Key References

- For a detailed list of references, see http://www.cis.ohio-state.edu/~jain/refs/gbe_refs.htm Also reproduced at the end of this tutorial book.
- Gigabit Ethernet Overview, http://www.cis.ohio-state.edu/~jain/cis788-97/gigabit_ethernet/index.htm
- 10 Gigabit Ethernet, http://www.cis.ohio-state.edu/~jain/cis788-99/10gbe/index.html
- 10 Gigabit Ethernet Alliance, http://www.10gea.org
- 10 GbE Resource Site, http://www.10gigabit-ethernet.com

References (Cont)

- IEEE 802.3 Higher Speed Study Group, http://grouper.ieee.org/groups/802/3/10G_study/public/index.html
- Email Reflector, http://grouper.ieee.org/groups/802/3/10G_study/email/thrd1.html
- IEEE 802.3ae 10Gb/s Ethernet Task Force, http://grouper.ieee.org/groups/802/3/ae/index.html
- IEEE 802.3ae email list, send a message with "subscribe stds-802-3-hssg <email adr>" in body to majordomo@majordomo.ieee.org

IP over DWDM

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

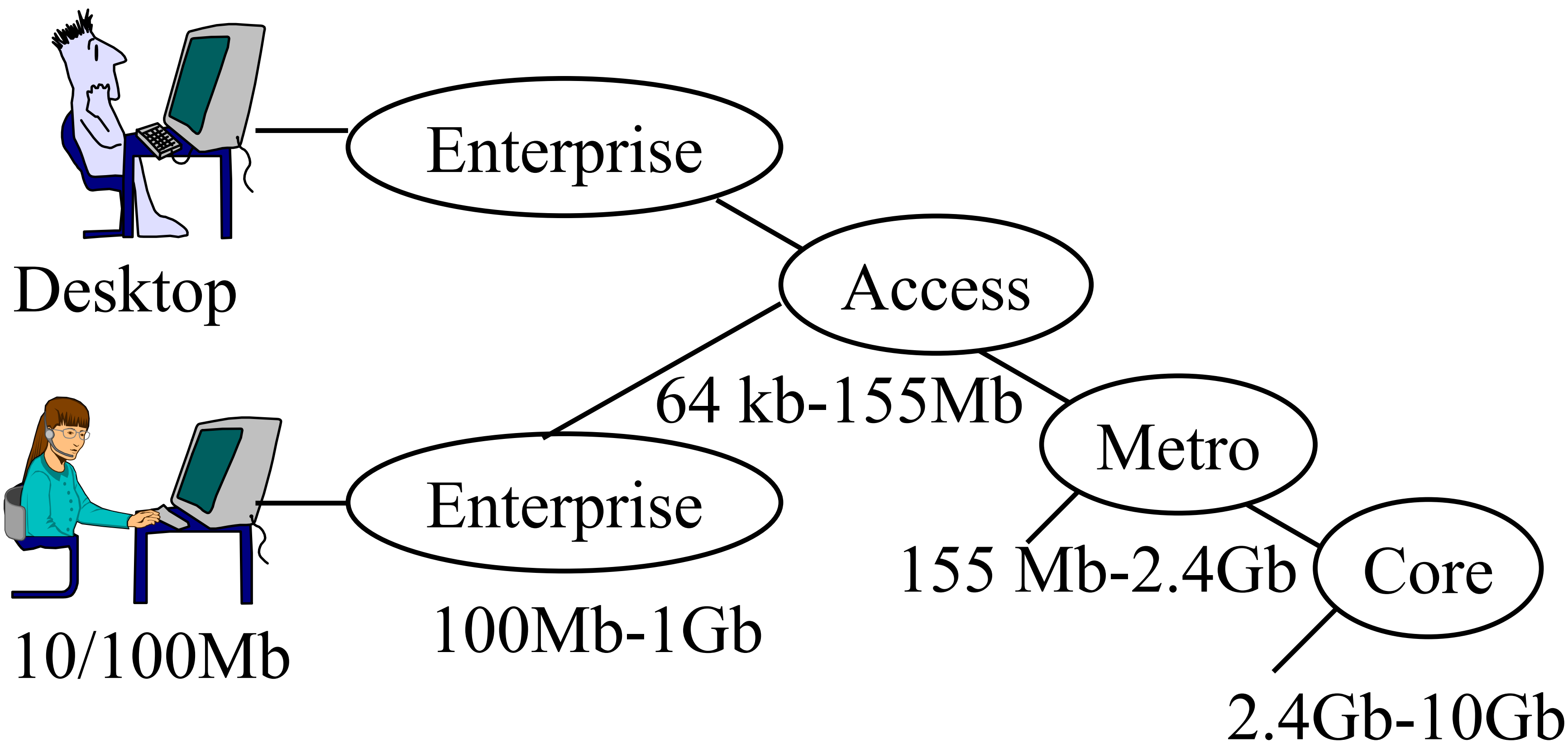
- Recent DWDM Records and Product Announcements
- Why IP over DWDM?
- How to IP over DWDM?
  - What changes are required in IP?
  - MPλS and GMPLS
  - UNI, LDP, RSVP, LMP
- Upcoming Optical Technologies

Past...

Who started optical networking?

Raj Jain

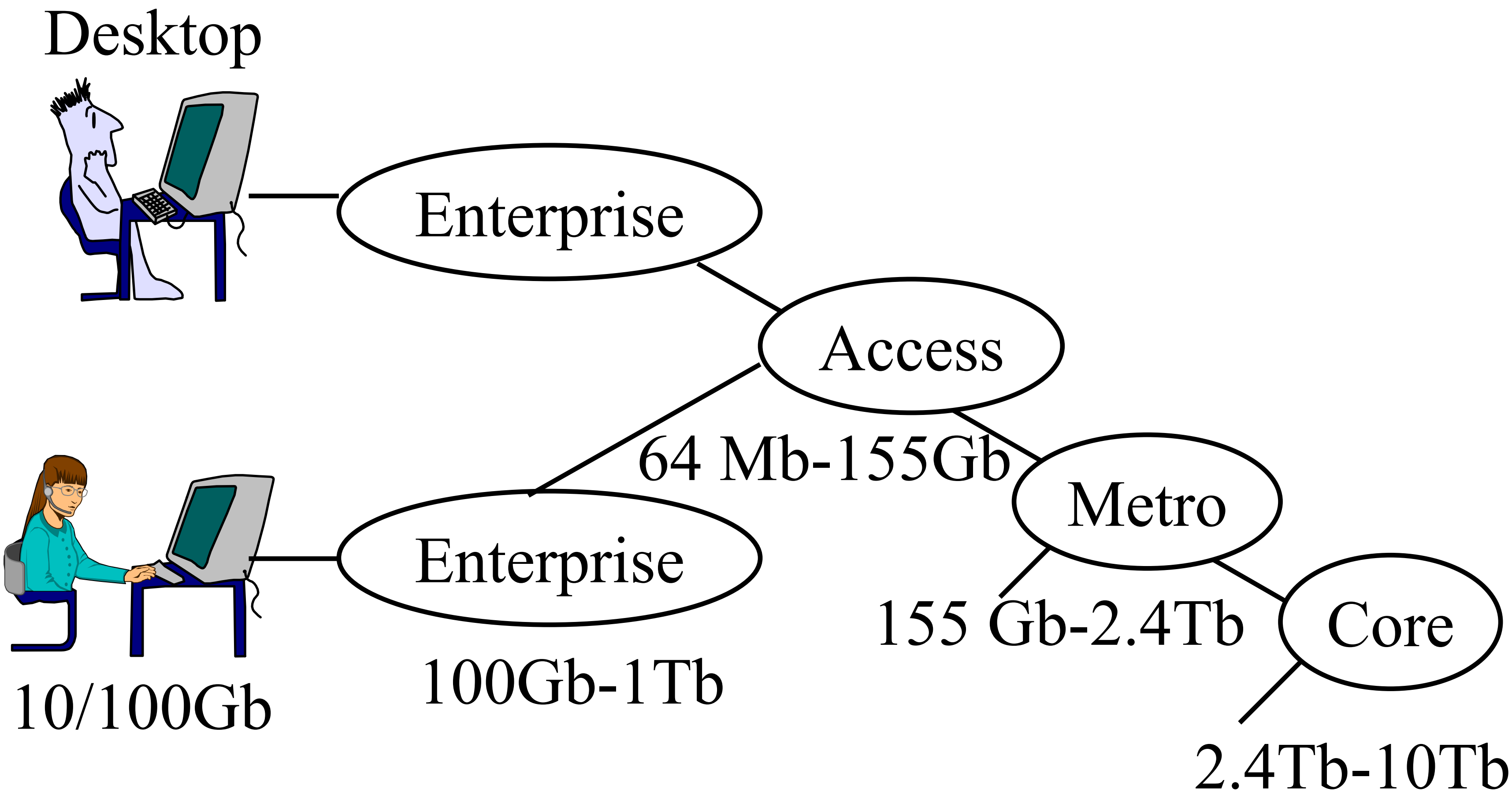
Issue: Access Bottleneck



- Access: Business bottleneck. Not technology.
  ⇒ New carriers for data services
  ⇒ ISP - Carrier convergence

Raj Jain

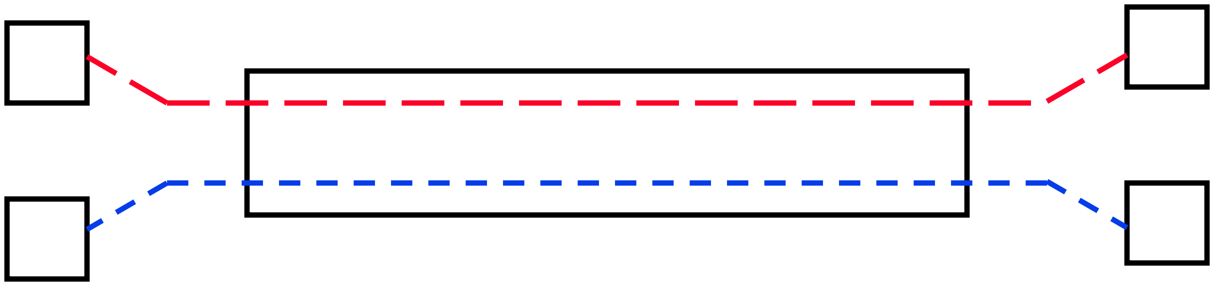
Bandwidth 2005



- Factor of 4/year = 1024 over 5 years

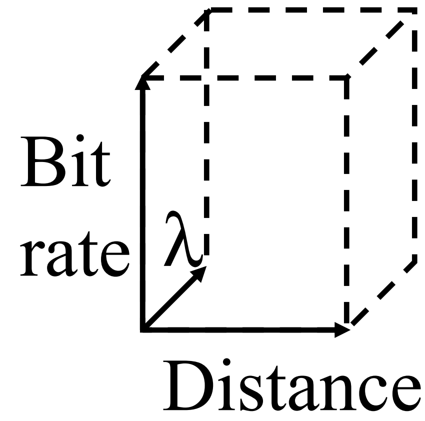
Raj Jain

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



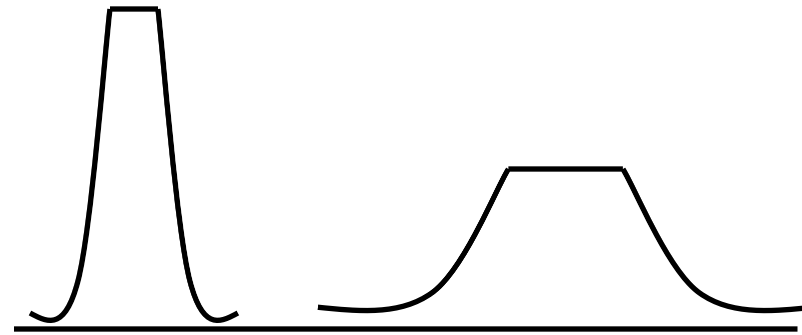
- 32λ× 5 Gbps to 9300 km (1998)
- 16λ× 10 Gbps to 6000 km (NTT'96)
- 160λ× 20 Gbps (NEC'00)
- 128λ× 40 Gbps to 300 km (Alcatel'00)
- 32λ× 40 Gbps to 2400 km (Alcatel'01)
- 19λ× 160 Gbps (NTT'99)
- 7λ× 200 Gbps (NTT'97)
- 1λ×1200 Gbps to 70 km using TDM (NTT'00)
- 1022 Wavelengths on one fiber (Lucent'99)

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

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

Raj Jain

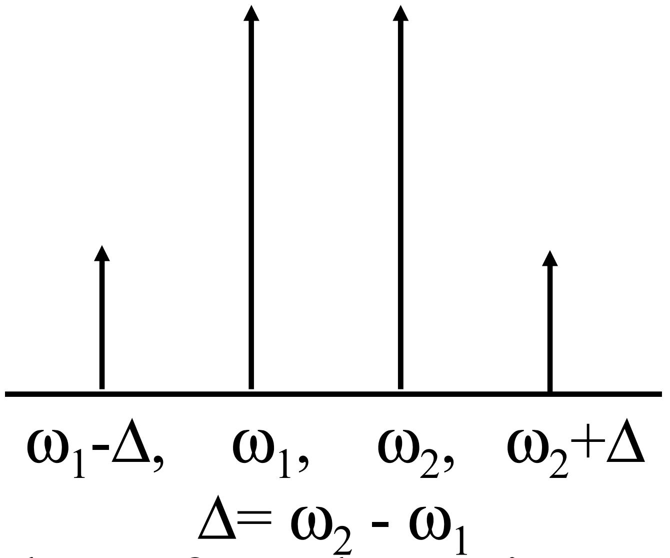
Attenuation and Dispersion



- Pulses become shorter and wider as they travel through the fiber

Raj Jain

Four-Wave Mixing



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

Raj Jain

New Developments

1. Higher Speed: 40 Gbps
2. More Wavelengths per fiber
3. Longer Distances
4. Larger Crossconnects
5. Newer places to install fibers

40 Gbps

Transmitter — Sources, Modulators, Wavelockers

Mux/Demux — Filters, Interleavers

Amplifier — Gain Equalizers, Performance Monitors

Switching — ADM

Receivers — Detectors

Dispersion compensators
PMD compensators

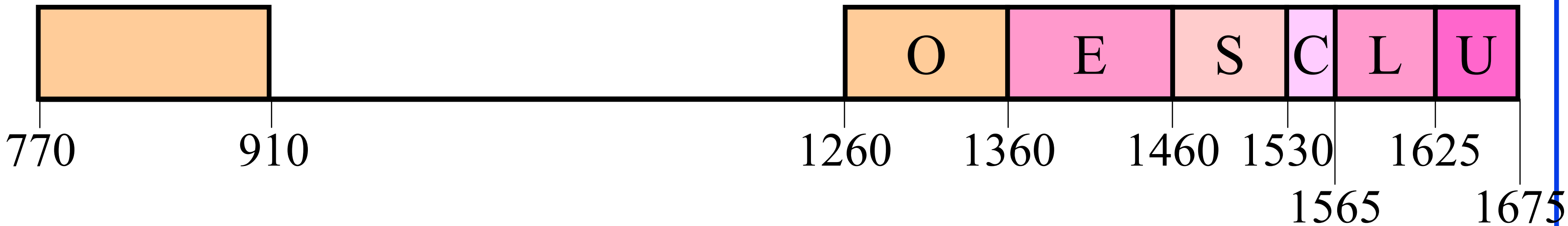
- Need all new optical and electronic components
- Non-linearity's reduced distance by square of rate.
- Deployment may be 2-3 years away
- Development is underway. To avoid 10 Gbps mistake.
- Cost goal: 2.5×10 Gbps

Raj Jain

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**



Raj Jain

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)

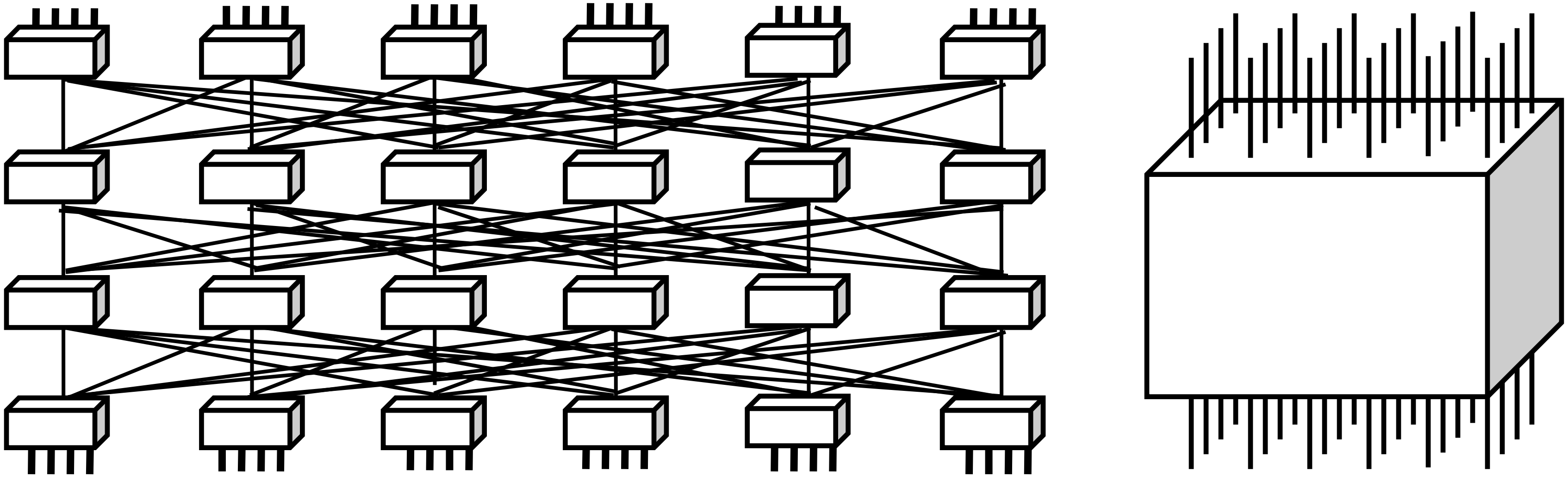
Ultra-Long Haul Transmission

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

Trend: Large Port Count

- Increasing traffic
  $\Rightarrow$ Increase number of ports or
  increase speed per port
- Increasing the port speed increases the number of muxing/demuxing (grooming) points
  Increases # of hops.
- Trend: Number of hops is decreasing (Avg 1.8)
  $\Rightarrow$ Larger number of ports per router
  E.g., Avici
- Also, larger # of wavelengths per fiber

Trend: Larger Crossconnects



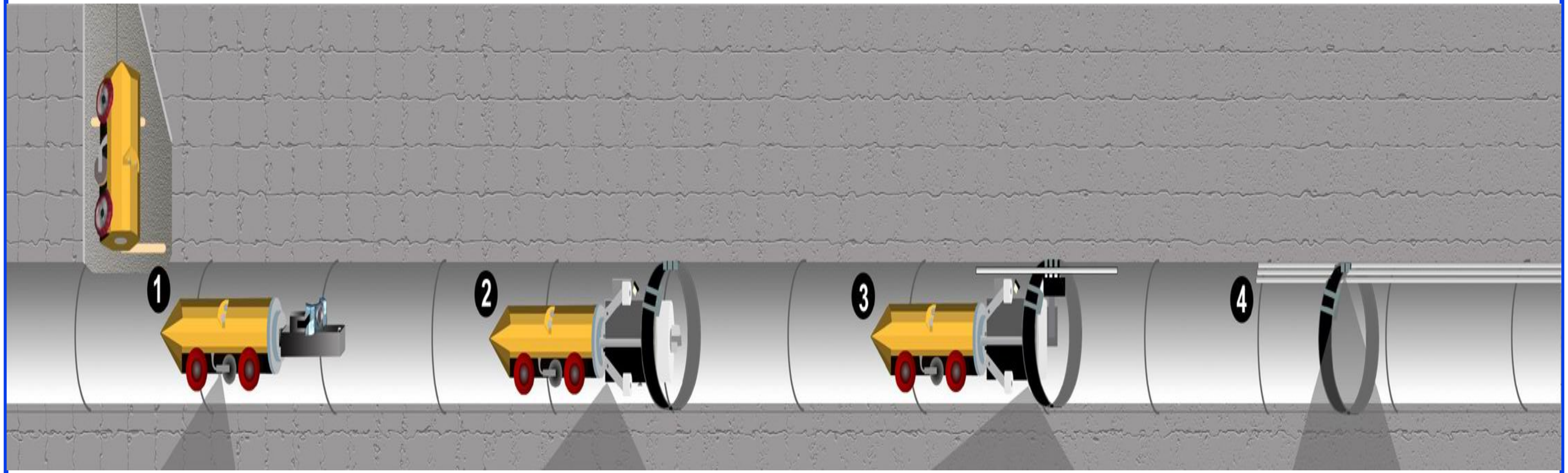
- Example: 24×24 using 4×4 switches ⇒ 24 switches ⇒ 48 External ports, 8×24=192 total ports ⇒ 25% port efficiency
- Crossconnect or routers with large number of ports are more cost effective

Raj Jain

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

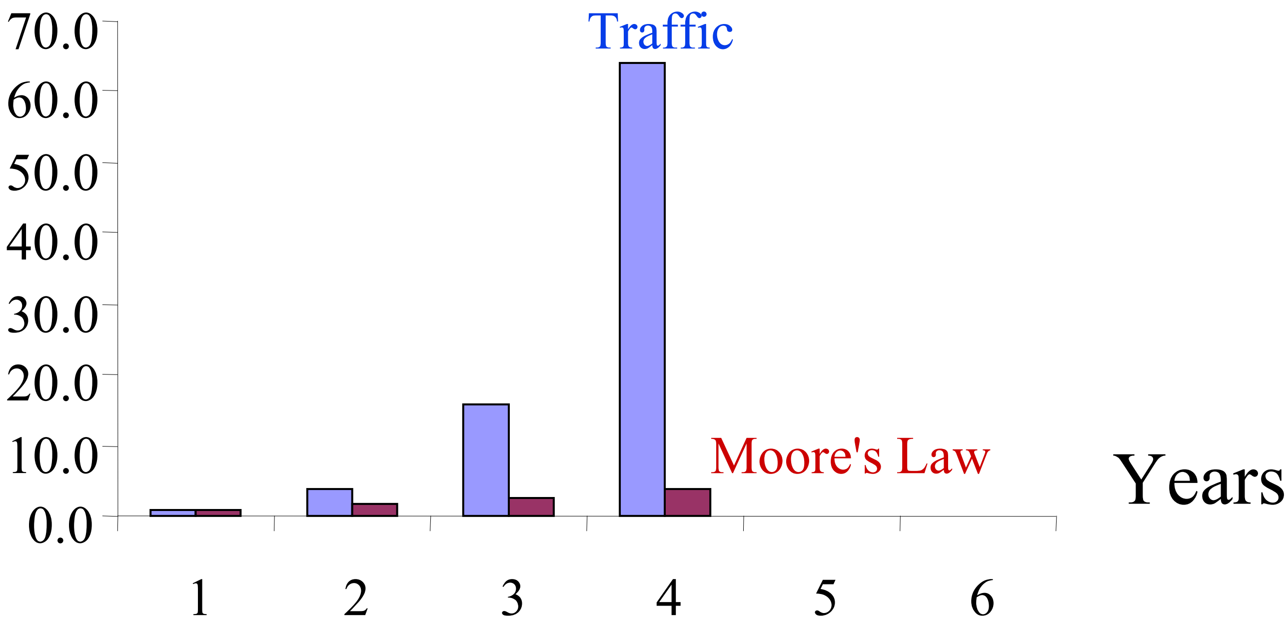
Raj Jain

Recent Products Announcements

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

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

Moore's Law is Too Slow



- Moore's Law: Factor of 2 every 1.5 years
  ⇒ 60%/year
- Internet Traffic: Factor of 4 per year
  ⇒ Need Optical Switching

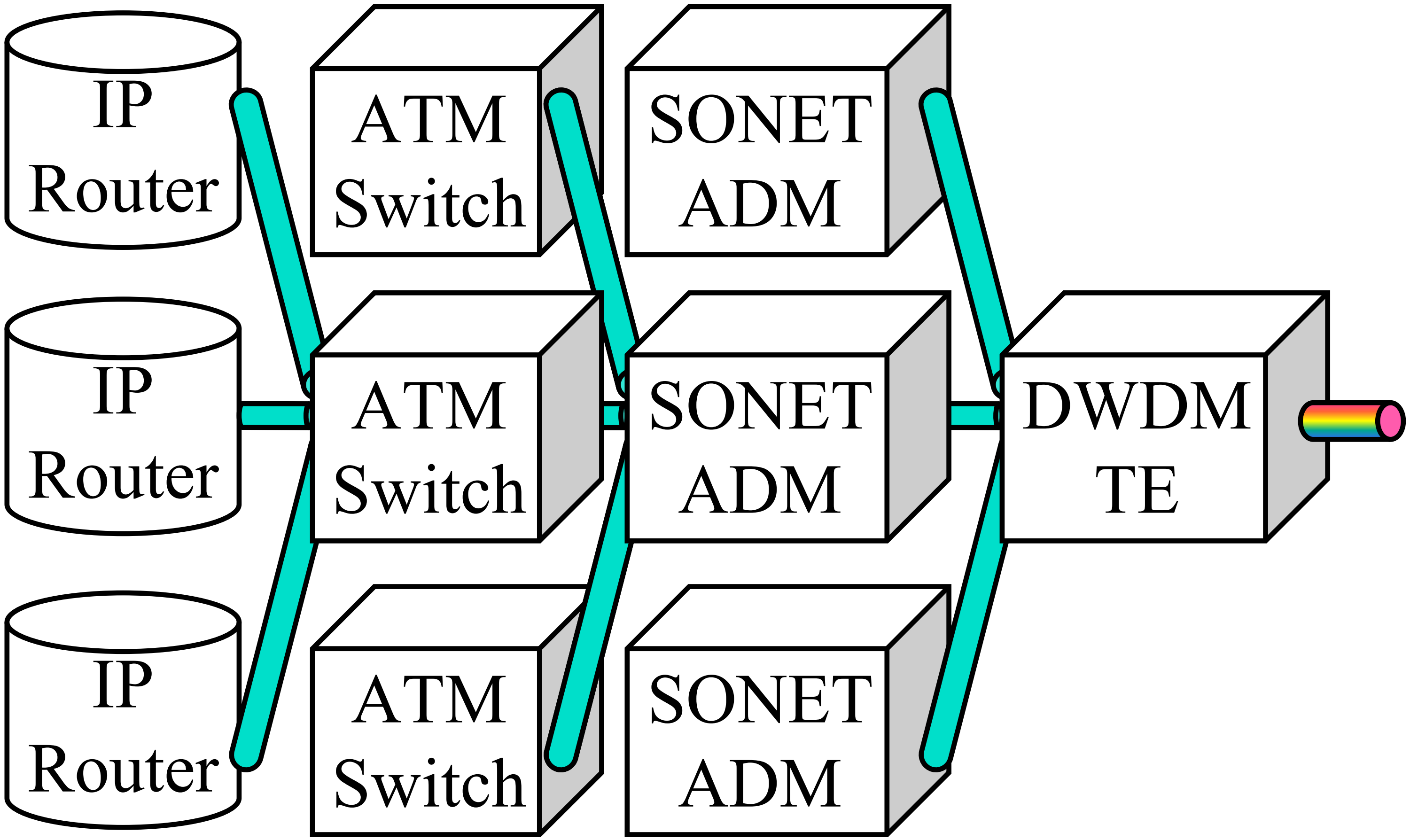
Raj Jain

OEO vs OOO

Feature	OEO	OOO
Data Format Independence	No	√ Yes
Cost/Space/Power independent of rate	No	√ Yes
Upgradeability to higher rate	No	√ Yes
Sub-Wavelength Switching	√ Yes	Future
Waveband Switching	No	√ Yes
Performance Monitoring	√ Bit error rate	Optical signal degradation
Wavelength Conversion	√ Built-in	1+ year away

Raj Jain

IP over DWDM (Past)



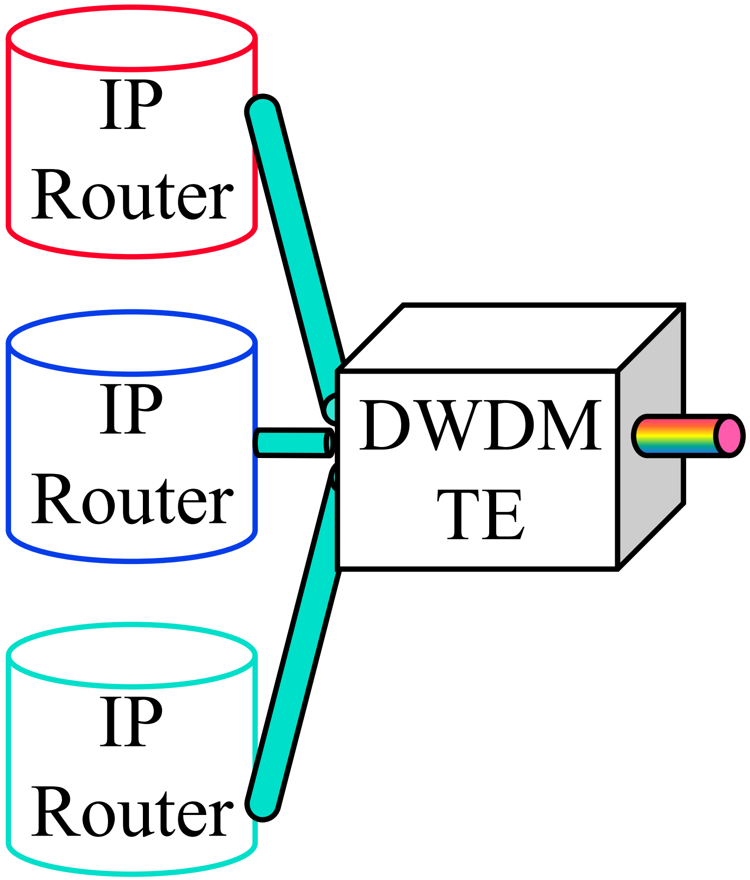
Raj Jain

IP over DWDM: Protocol Layers

1993	1996	1999	2001	2003
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

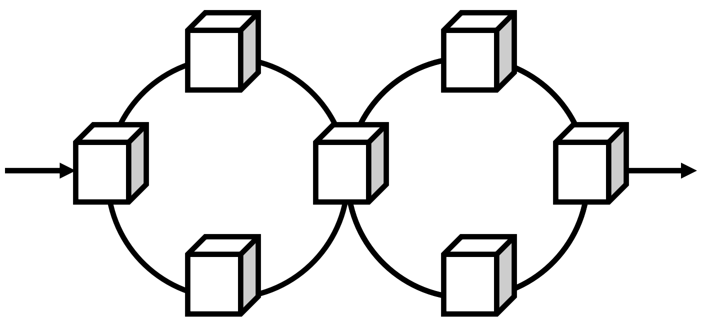
IP over DWDM (Future)

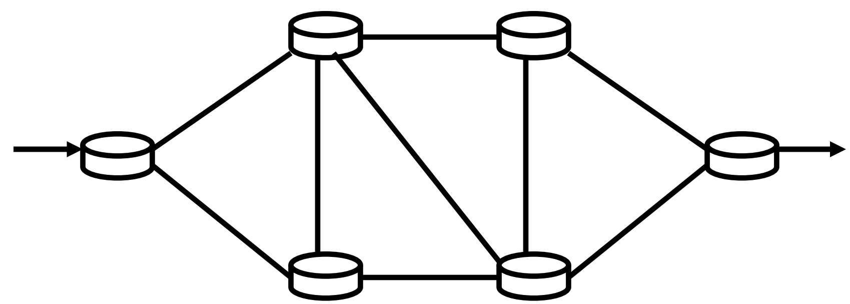


Raj Jain

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





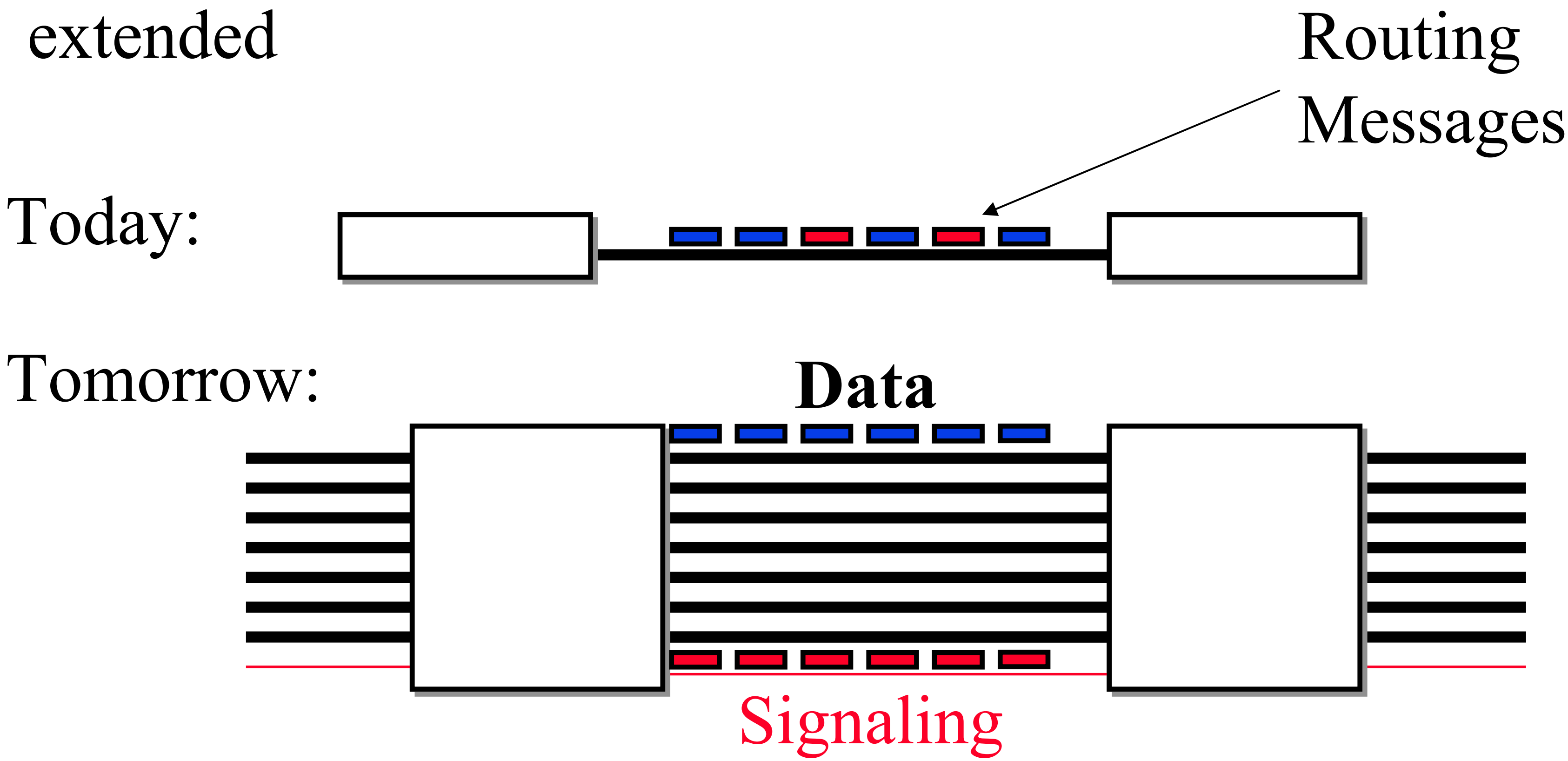
Raj Jain

IP over DWDM Issues

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

Issue: Control and Data Plane Separation

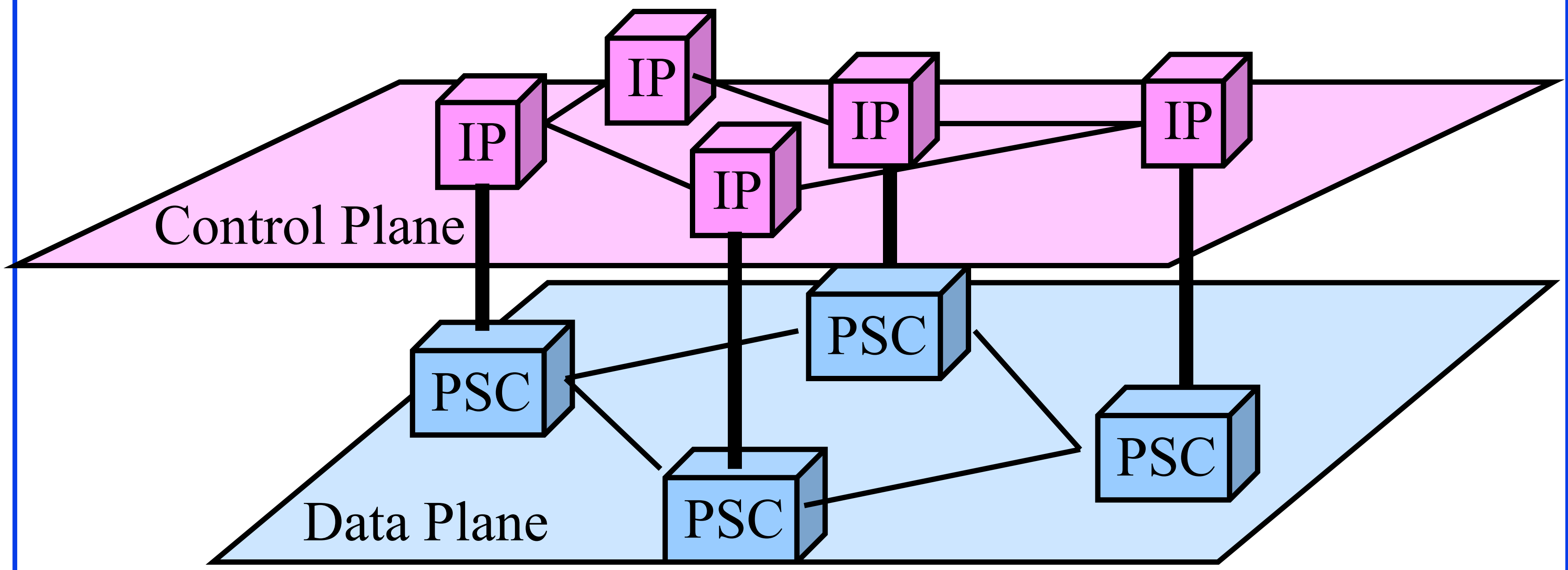
- Separate control and data channels
- IP routing protocols (OSPF and IS-IS) are being extended



Raj Jain

IP-Based Control Plane

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

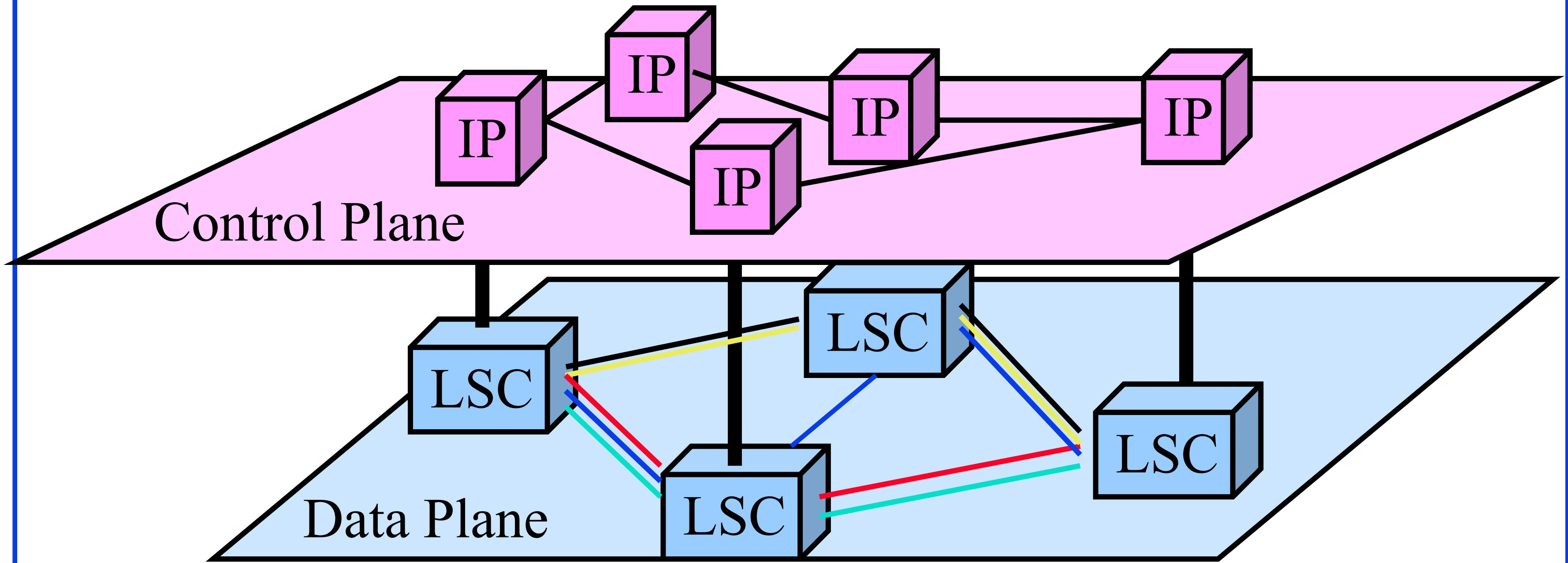


PSC = Packet Switch Capable Nodes

Raj Jain

MPλS

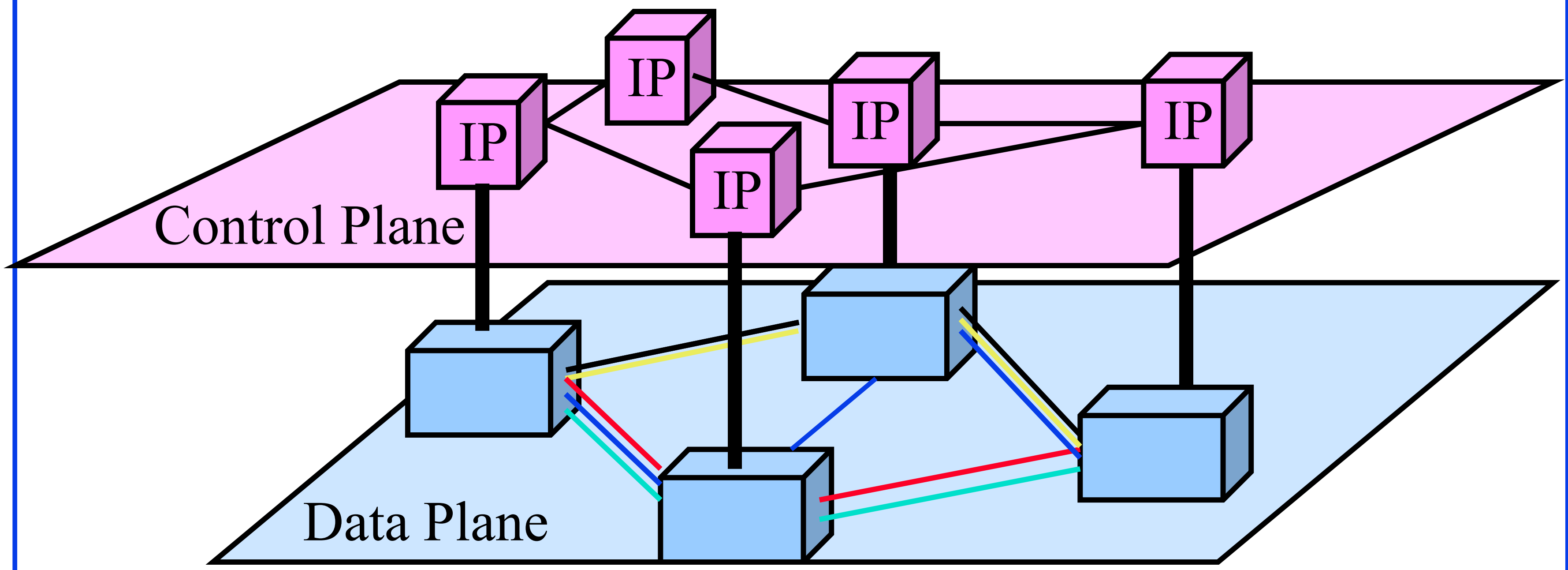
- Control is by IP packets (electronic).
  Data plane consists of wavelength circuits
  ⇒ Multiprotocol Lambda Switching (October 1999)



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

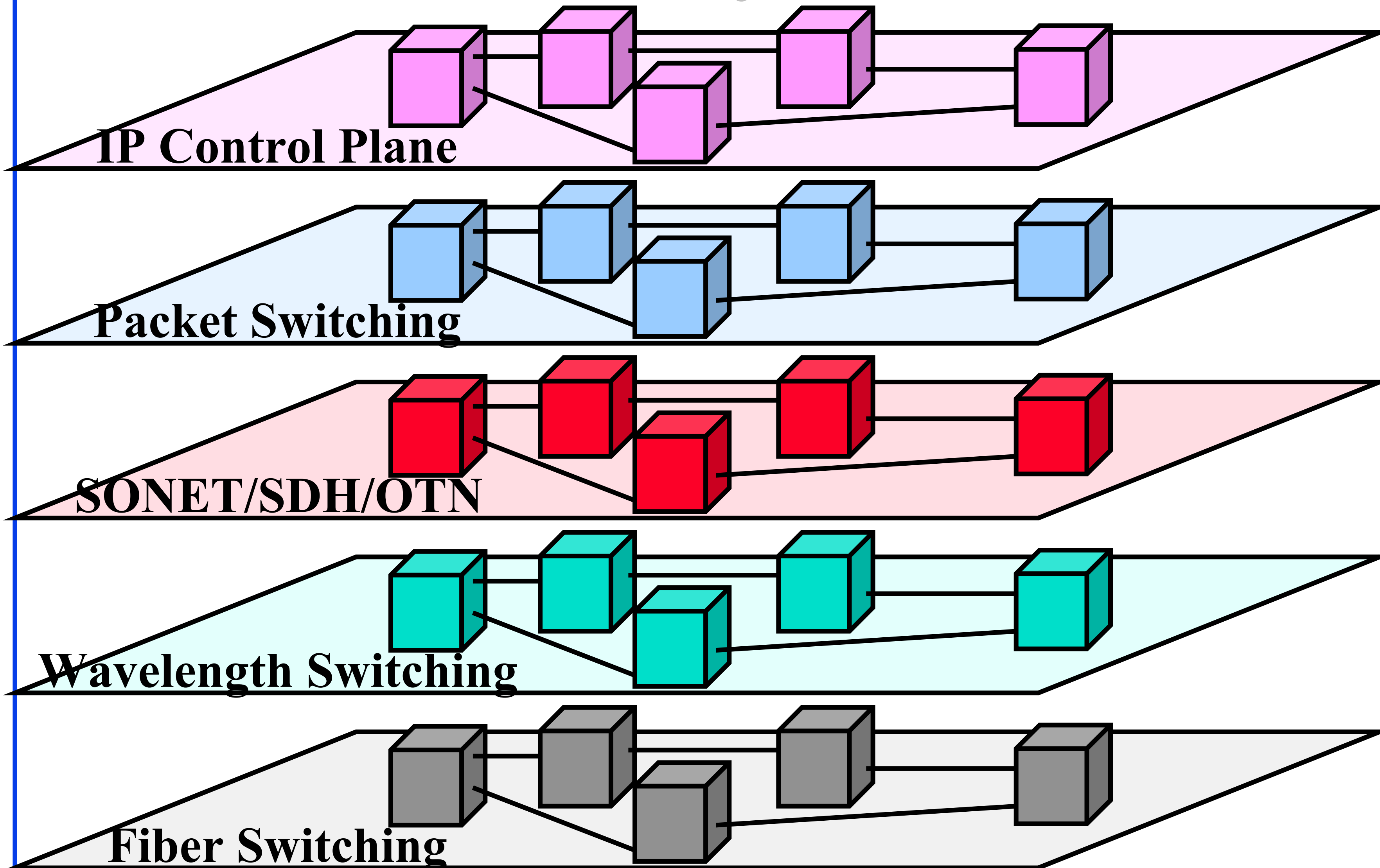
GMPLS

- Data Plane = Wavelengths, Fibers, SONET Frames, Packets (October 2000)



Raj Jain

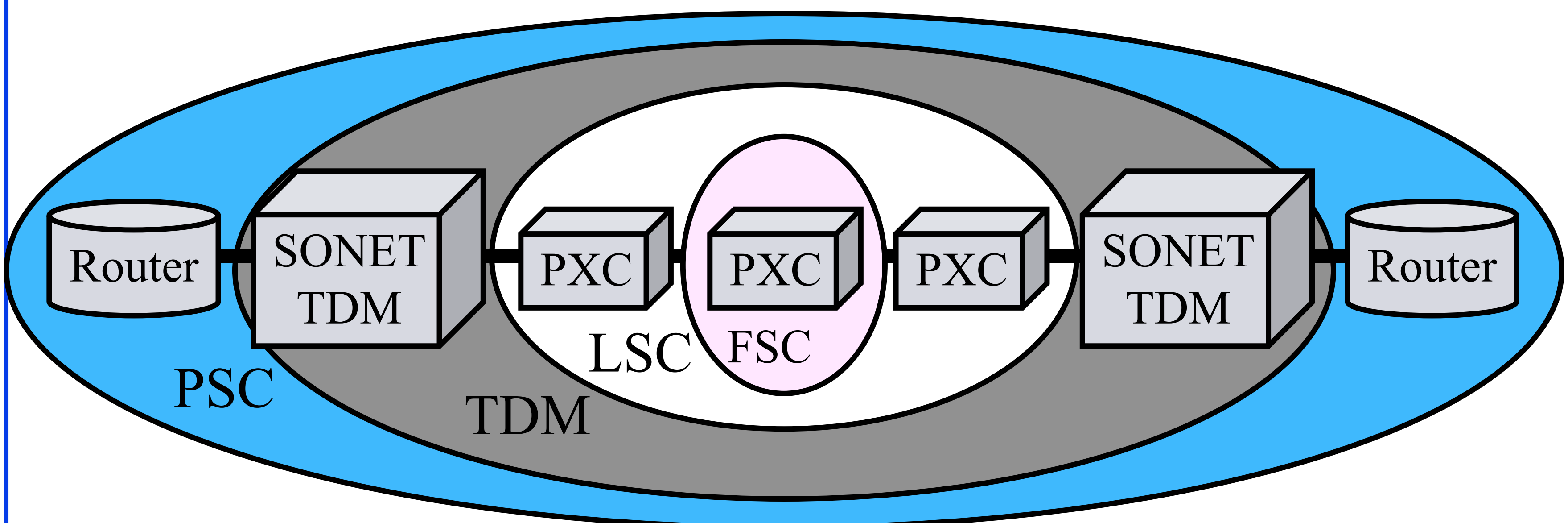
GMPLS: Layered View



Raj Jain

GMPLS: Hierarchical View

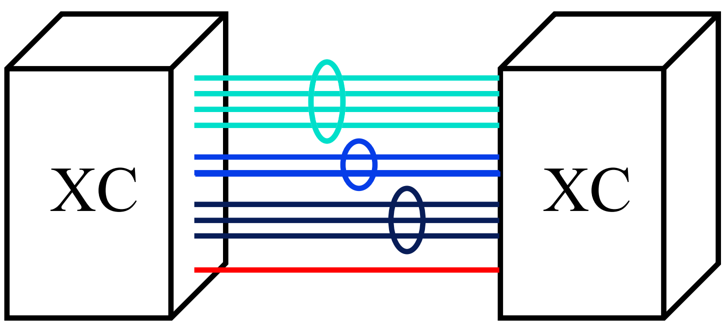
- Packets over SONET over Wavelengths over Fibers
- Packet switching regions, TDM regions, Wavelength switching regions, fiber switching regions



Raj Jain

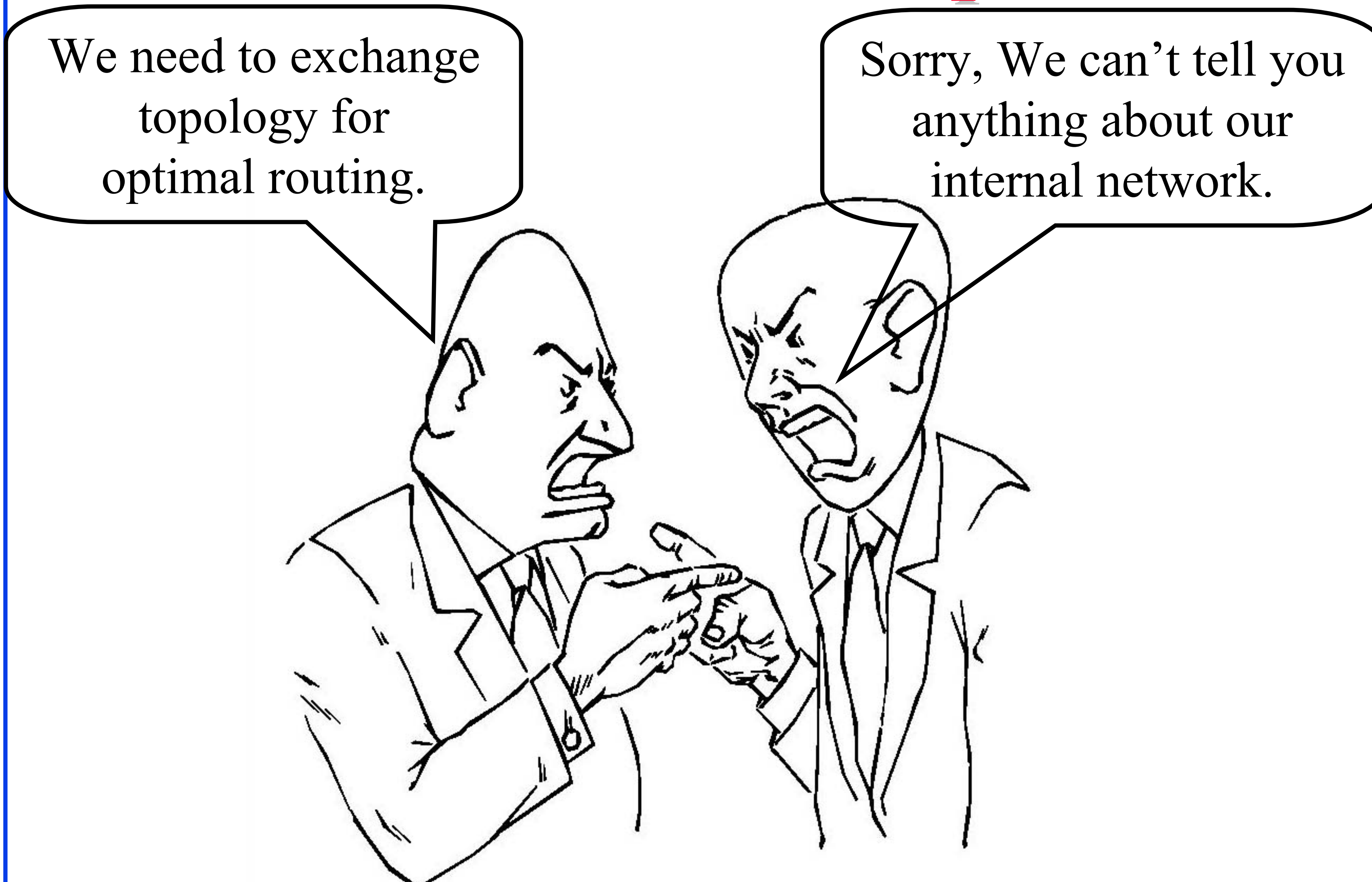
MPLS vs GMPLS

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	Unnumberred
Fault Detection	In-band	Out-of-band or In-Band



Raj Jain

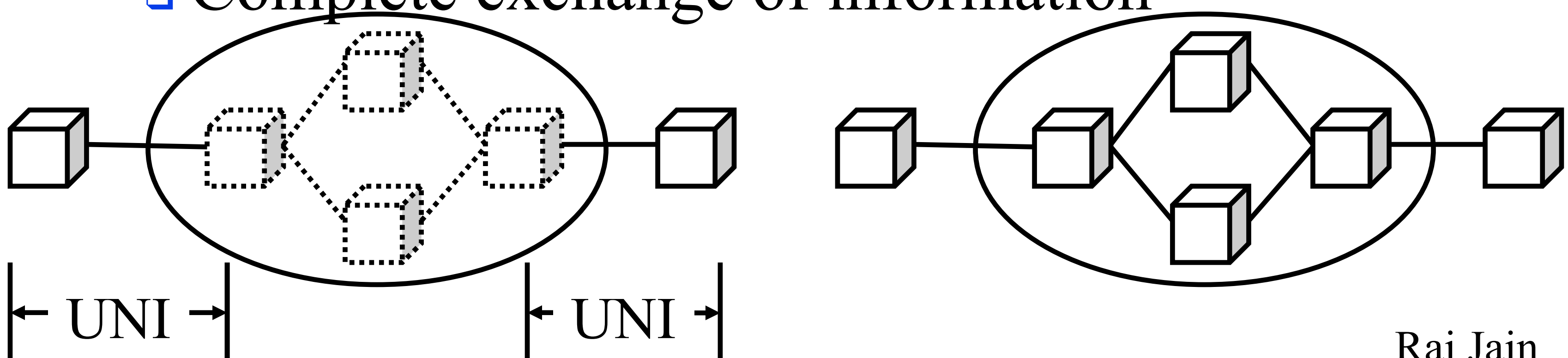
Carriers vs Enterprise



Raj Jain

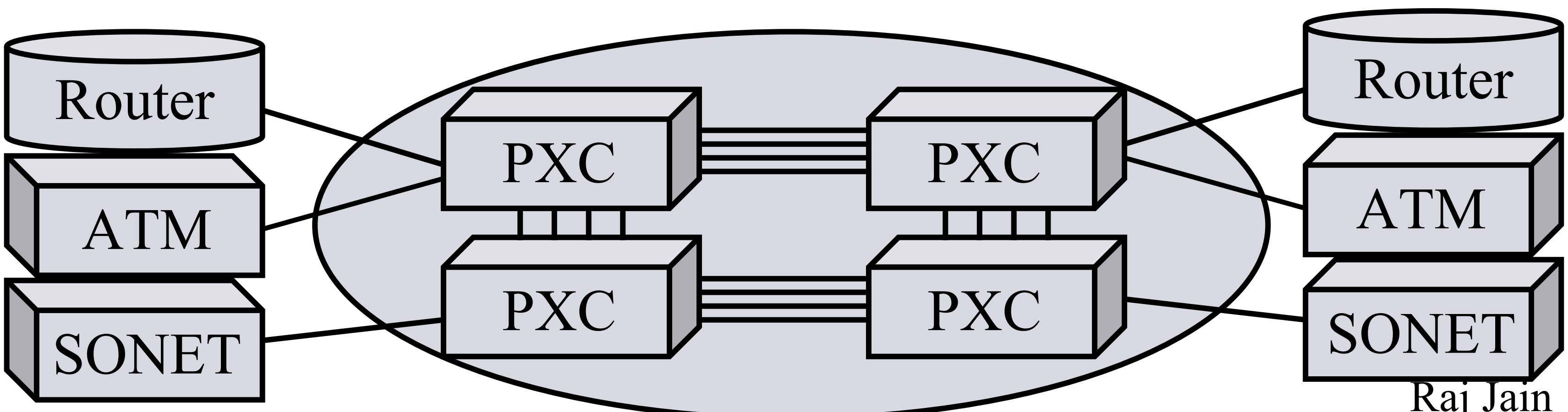
Issue: UNI vs Peer-to-Peer Signaling

- 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



Addressing

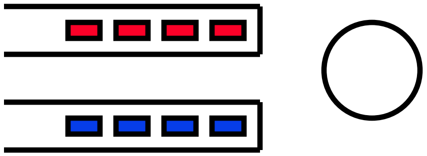
- Many different client types IP, ATM, SONET, …
  Each type has its own address: IPv4, IPv6, ATM, …
- Should a client be addressed by Switch and Port #?
- **Answer**: Optical Network Assigned Address (ONA)
  Globally Unique. Like Phone Number.
- Address Resolution Protocol to register and resolve name to ONA. Connect using ONA.

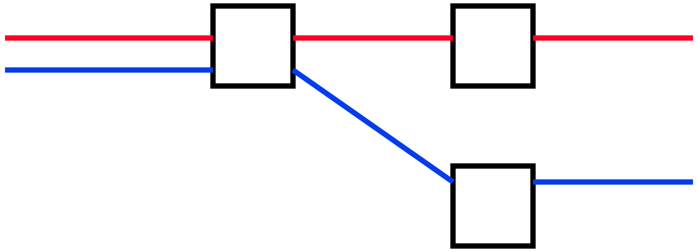


Raj Jain

Issue: Traffic Engineering

- Quickly create/destroy lightpaths on-demand
  - Bandwidth trading
  - Optical Dial Tone
- Dynamic topology for dynamic traffic
- Circuit-level priority for setup, holding, and restoration
- No packet-level queuing, marking, scheduling in the core





Raj Jain

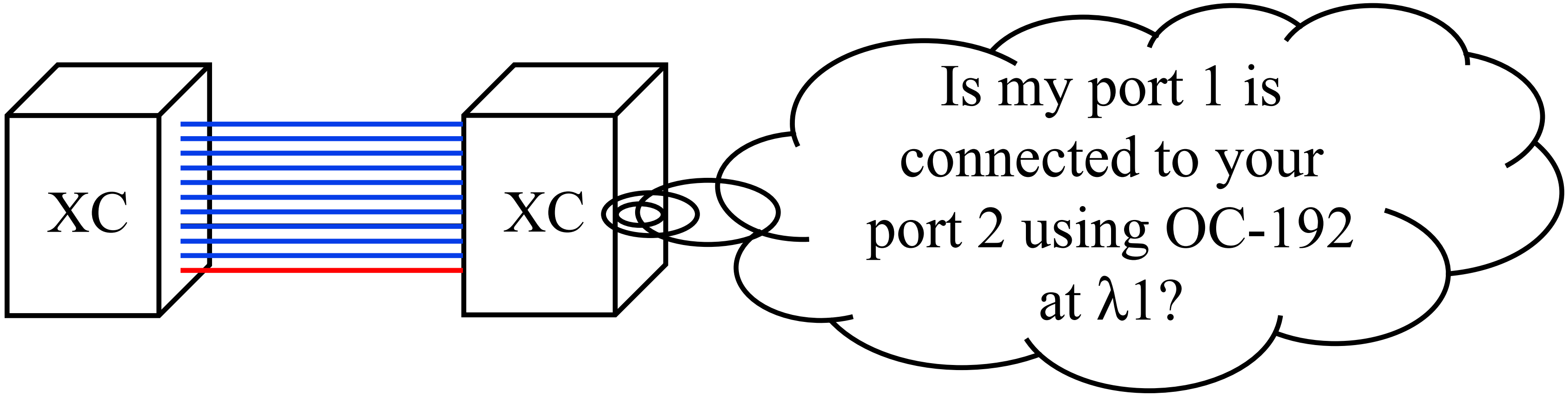
CR-LDP for UNI

- **Connection Management**
  - Connection create request ⇒ Label request
  - Connection create response ⇒ Label mapping
  - Connection delete request ⇒ Label release
  - Connection delete response ⇒ Label withdraw
- **Status**
  - Connection status query ⇒ Status query (new)
  - Connection status response ⇒ Status response (new)
- **Notification:**
  - Change Notification ⇒ CR-LDP notification
- **Address Resolution:**
  - Resolution query ⇒ Resolution query (new)
  - Resolution response ⇒ Resolution response (new)

RSVP for UNI

- **Connection Management**:
  - Connection create request ⇒ Path
  - Connection create response ⇒ Resv
  - Connection delete req ⇒ PathTear/ResvTear/PathErr
  - Connection delete resp ⇒ No state in next refresh
- **Status**:
  - Connection status query ⇒ Srefresh with MsgID
  - Connection status response ⇒ MsgID Ack/Nack
- **Notification**:
  - Change Notification ⇒ MsgID Ack/Nack
- **Address Resolution**:
  - Resolution query ⇒ Resolution query (new)
  - Resolution response ⇒ Resolution response (new)
- **Node Failure Detection**: Hello Message (new)

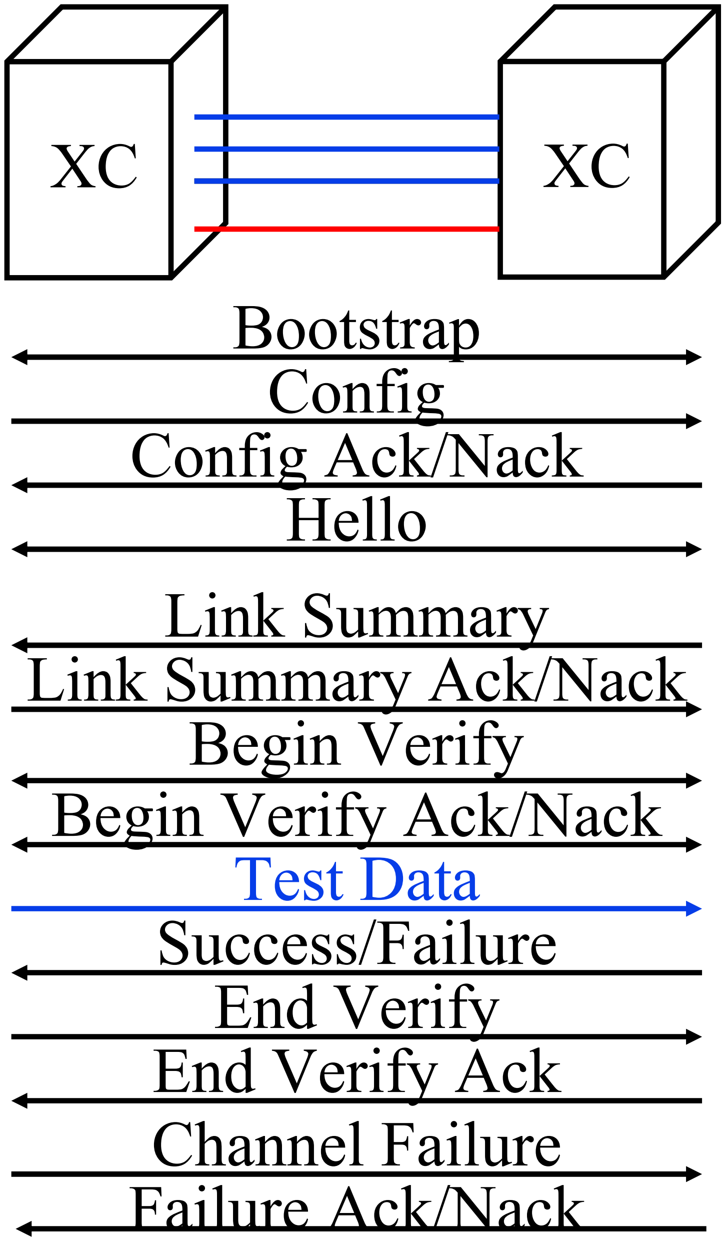
Link Management Protocol (LMP)



- Too many channels between crossconnects
- LMP allows connectivity verification, link parameter correlation, fault notification
- All communication takes place on control channel
- Only test messages on data channels to verify connectivity (optional)

Raj Jain

LMP Messages



- Are control channels connected?
- Parameter Negotiation: Hello interval, dead interval
- Keep Alive
- Link Property correlation
- Link Verification (optional)
- Fault reporting

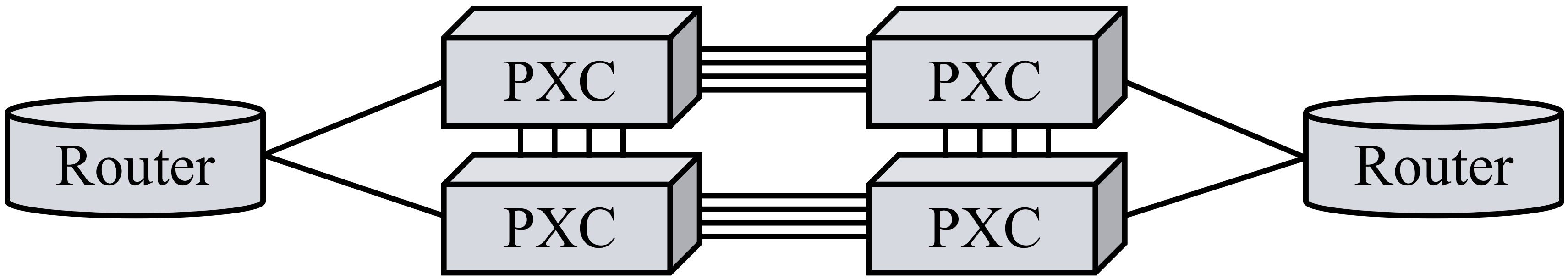
Raj Jain

Protection & Restoration

Raj Jain

Protection and Restoration

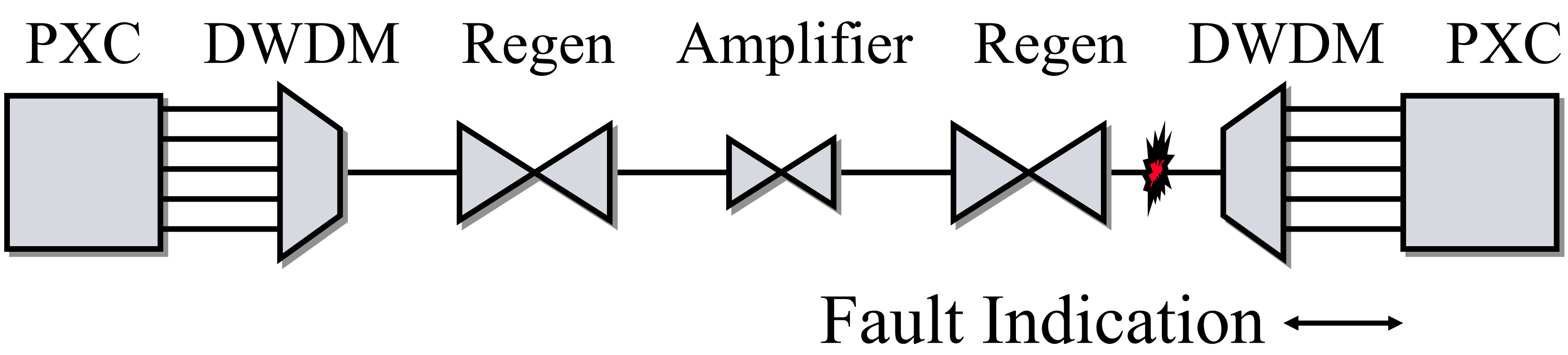
- Extent: SPAN vs PATH
- Topology: Ring vs Mesh
- Redundancy: 1+1, 1:1
- Finding Paths that do not share the same risk
  Each link has to be assigned a risk group
  Shared Risk Group (**SRG**) = All paths sharing a risk



Raj Jain

Fault Detection and Isolation

- SONET: Remote Defect Indicator, Alarm Indication Signal, Bit Interleaved Parity
- Photonic: Loss of signal, Optical degradation of signal
- Solution: A protocol for active devices to communicate fault information to Photonic switches Examples: LMP-DWDM, NTIP

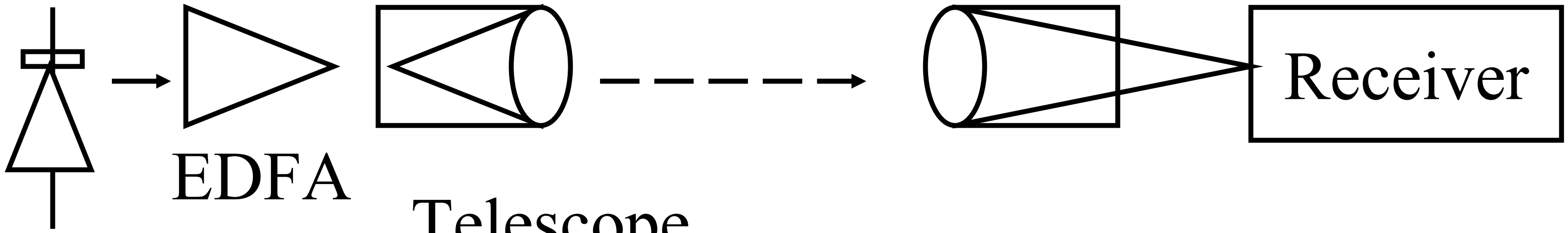


Raj Jain

Upcoming Technologies

- Optic Wireless
- Higher bit rate
  - Optical Time Domain Multiplexing (OTDM)
  - Optical Code Division Multilexing (OCDMA)
- Optical Packet Switching

Free Space Optical Comm



Laser
Source

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

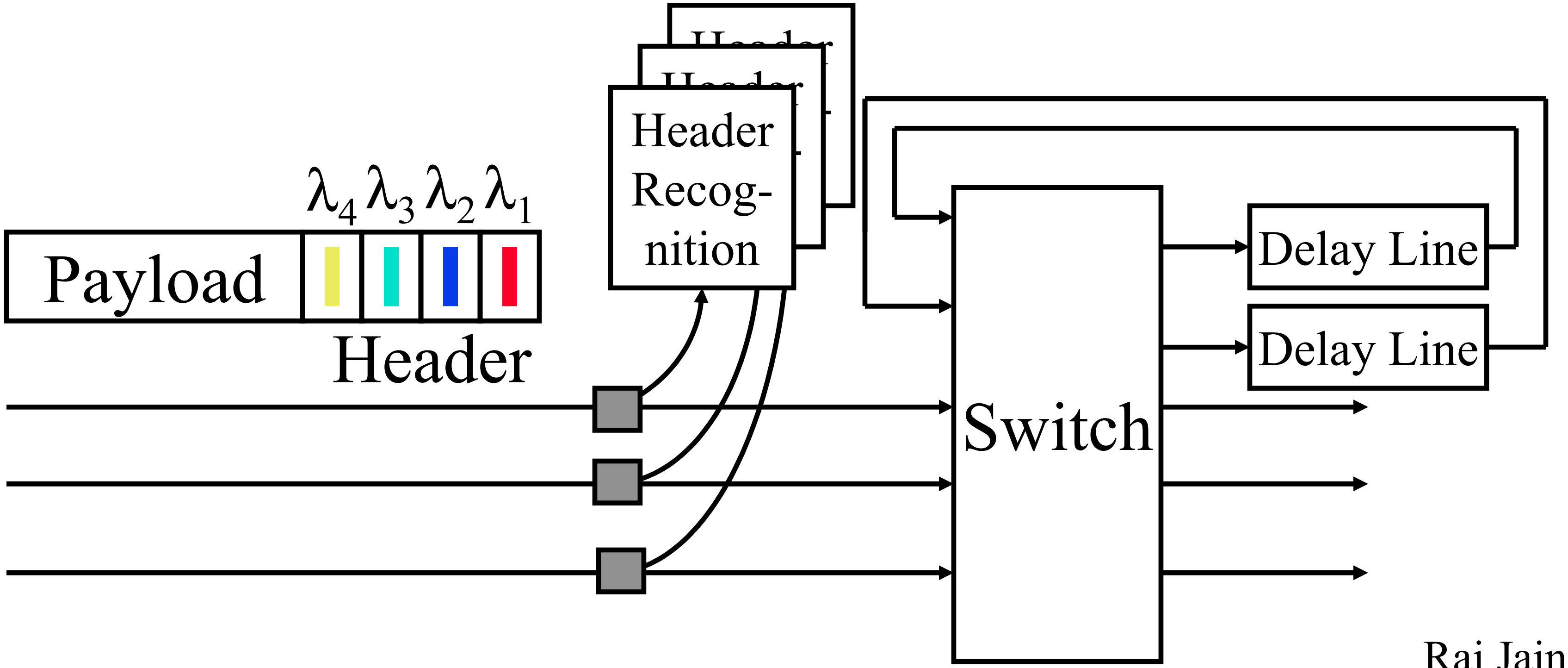
Raj Jain

Free Space Optical Comm

- No FCC Licensing required
- Immunity from interference
- Easy installation
  $\Rightarrow$ Unlimited bandwidth, Easy Upgrade
- Transportable upon service termination or move
- Affected by weather (fog, rain)
  $\Rightarrow$ 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

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

Raj Jain

IP over DWDM: Key References

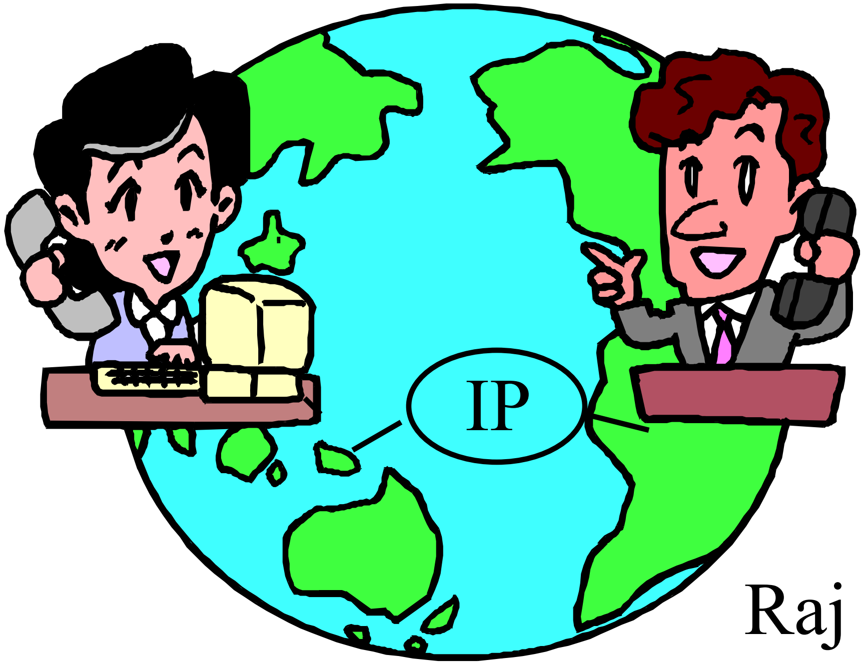
- Detailed references in http://www.cis.ohio-state.edu/~jain/refs/opt_refs.htm
  Also reproduced at the end of this tutorial book.
- Recommended books on optical networking, http://www.cis.ohio-state.edu/~jain/refs/opt_book.htm
- Optical Networking and DWDM, http://www.cis.ohio-state.edu/~jain/cis788-99/dwdm/index.html
- IP over Optical: A summary of issues, (internet draft) http://www.cis.ohio-state.edu/~jain/ietf/issues.html
- Lightreading, http://www.lightreading.com

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

Raj Jain

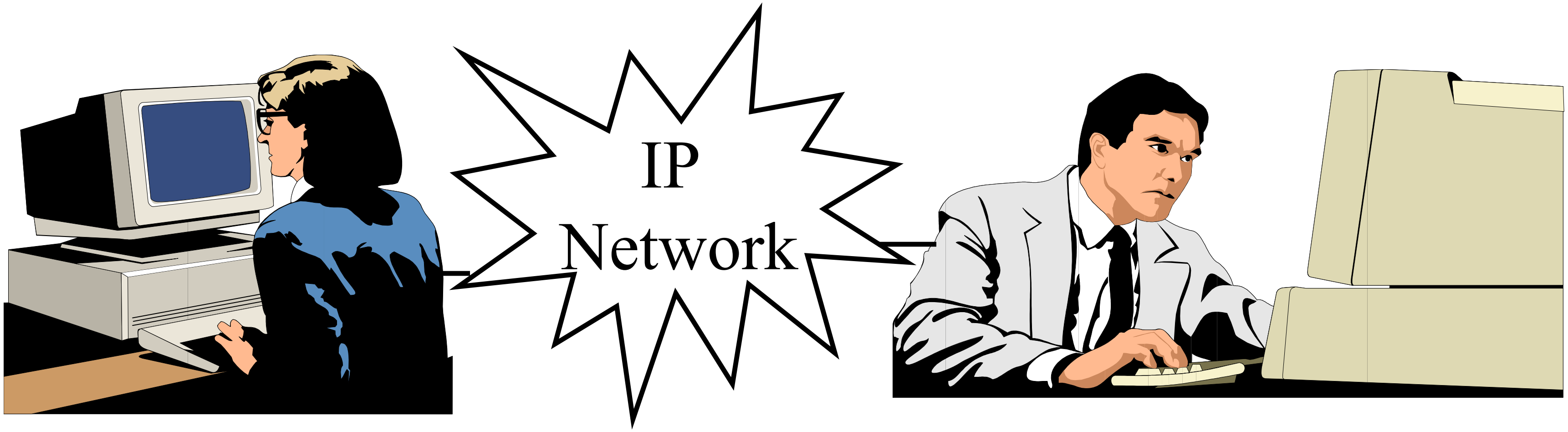
Voice over IP



Raj Jain
Co-Founder and CTO
Nayna Networks
481 Sycamore Dr, Milpitas, CA 95035
www.nayna.com and
http://www.cis.ohio-state.edu/~jain/

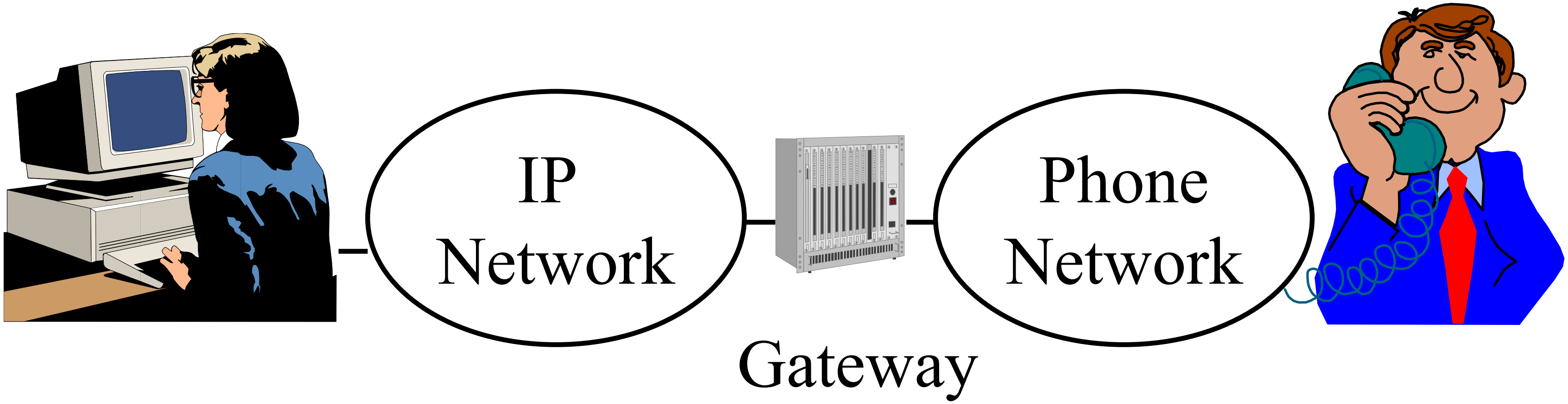
- Voice over IP: Why?
- Sample Products and Services
- 13 Technical Issues, 4 Other Issues
- H.323 Standard and Session Initiation Protocol (SIP)
- Media Gateway Control Protocol (MGCP) and Megaco
- Stream Control Transmission Protocol (SCTP)

Scenario 1: PC to PC



- Need a PC with sound card
- IP Telephony software: Cuseeme, Internet Phone, ...
- Video optional

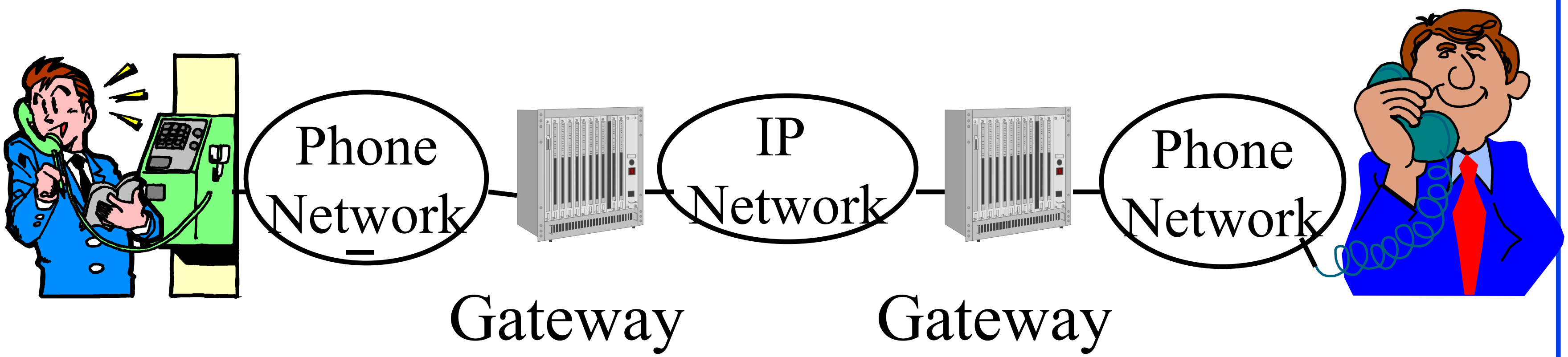
Scenario 2: PC to Phone



- Need a gateway that connects IP network to phone network (Router to PBX)

Raj Jain

Scenario 3: Phone to Phone



- Need more gateways that connect IP network to phone networks
- The IP network could be dedicated intra-net or the Internet.
- The phone networks could be intra-company PBXs or the carrier switches

Raj Jain

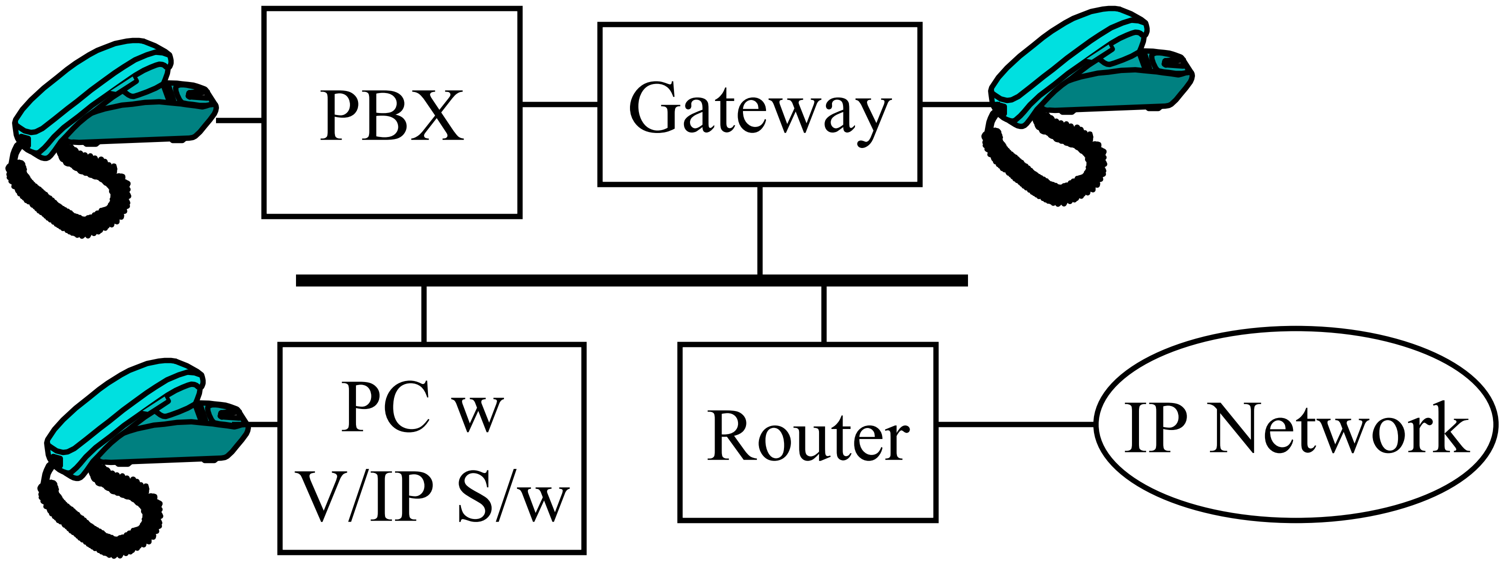
Applications

- Any voice communication where PC is already used:
  - Document conferencing
  - Helpdesk access
  - On-line order placement
- International callbacks
  (many operators use voice over frame relay)
- Intranet telephony
- Internet fax
- IP Control Plane

Sample Products

- VocalTec Internet Phone: PC to PC.
- Microsoft NetMeeting: PC to PC. Free.
- Internet PhoneJACK: ISA card to connect a standard phone to PC. Works with NetMeeting, InternetPhone etc. Provides compression.
- Internet LineJACK: Single-line gateway.
- Micom V/IP Family:
  - Analog and digital voice interface cards
  - PC and/or gateway

Products (Cont)



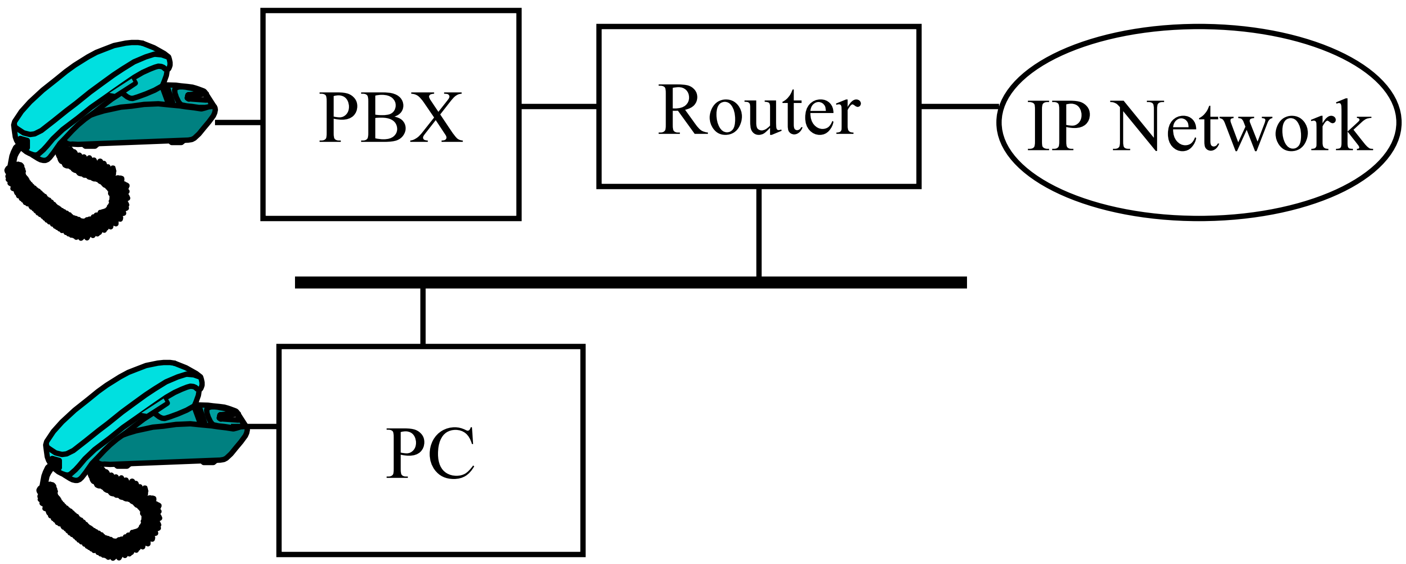
- Features:
  - Compression
  - Phone number to IP address translation.
  - Supports RSVP.
  - Limits number of calls.

Products (Cont)

- VocalTec Internet Telephony Gateway:
  - Similar to Micom V/IP
  - Interactive voice response system for problem reporting
  - Allows WWW plug in
  - Can monitor other gateways and use alternate routes including PSTN
  - Sold to Telecom Finland. New Zealand Telecom.
- Lucent's Internet Telephony Server: Gateway| Lucent PathStar Access Server

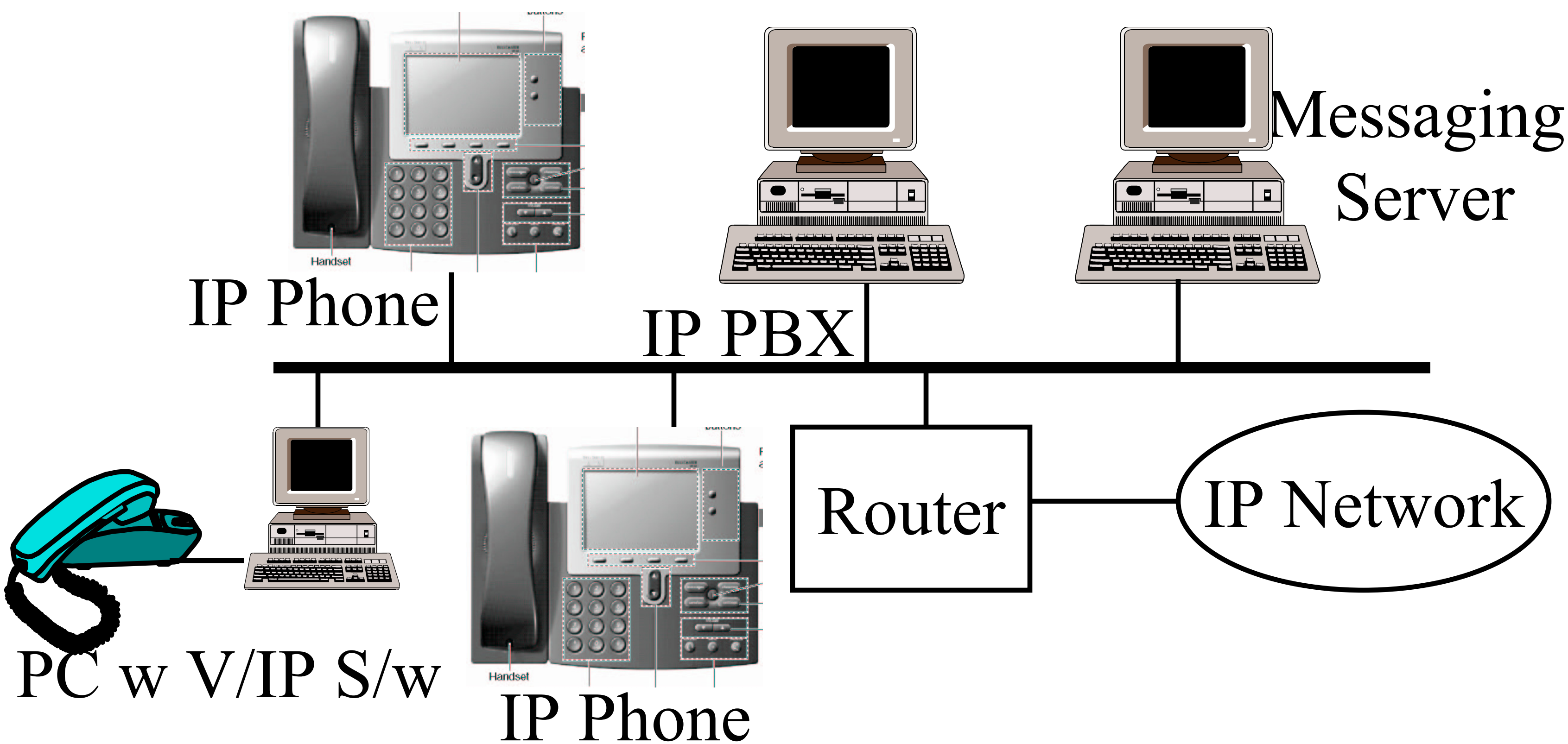
Products (Cont)

- CISCO 2600 Routers: Voice interface cards (VICs) Reduces one hop.
- Baynetworks, 3COM, and other router vendors have announced product plans



Products (Cont)

- IP Phone/IP PBX: Designed for enterprise market
- IP Voice Mail



Raj Jain

Sample Services

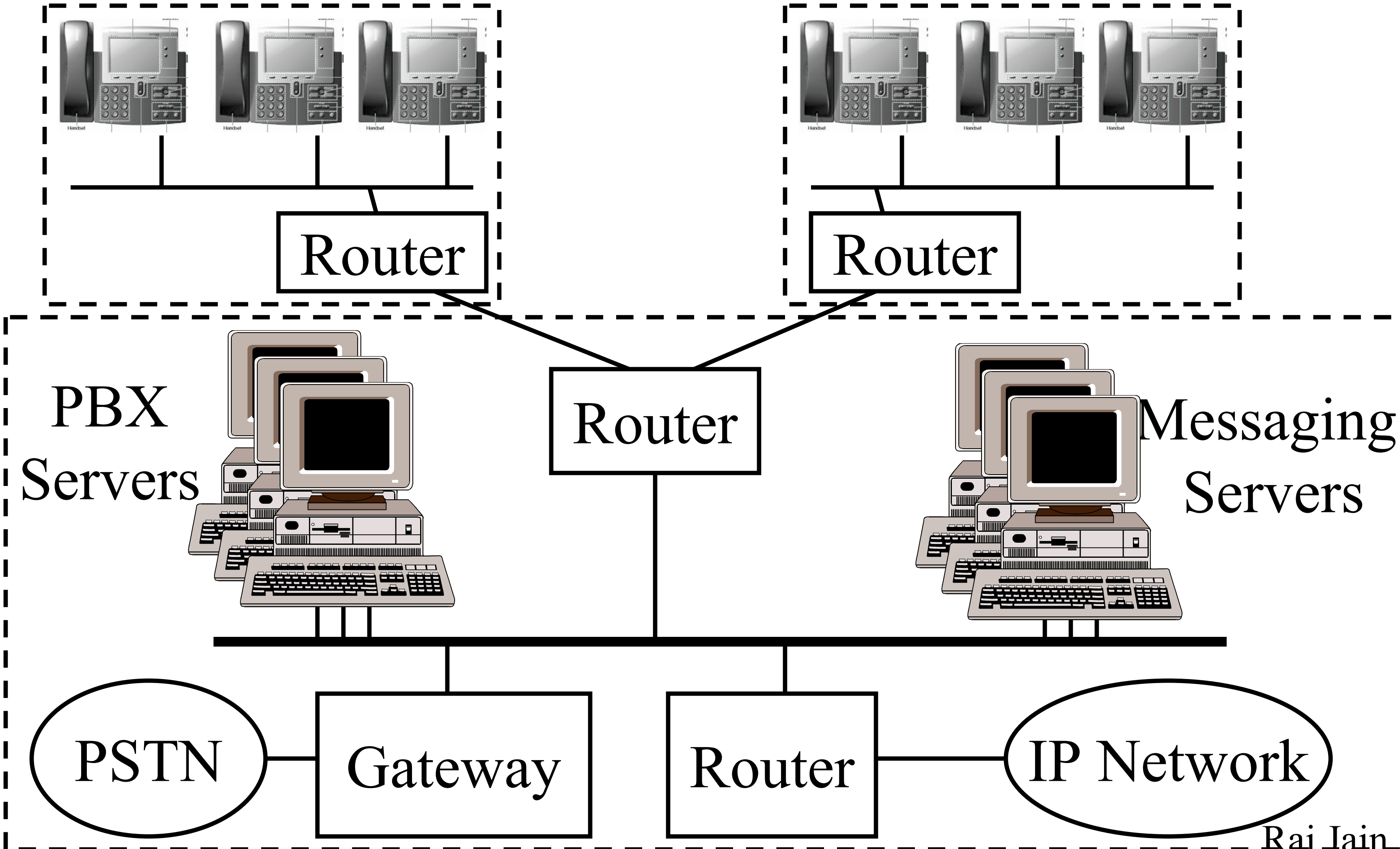
- IDT Corporation offers Net2Phone, Carrier2Phone, Phone2Phone services.
- Global Exchange Carrier offers international calls using VocalTec InternetPhone s/w and gateways
- Qwest offers 7.5¢/min VOIP Q.talk service in 16 cities.
- ITXC provides infrastructure and management to 'Internet Telephone Service Providers (ITSPs)'
- America On-line offers 9¢/min service.
- AT&T announced 7.5¢/min VOIP trials in 9 US cities.

Services (Cont)

- Other trials: USA Global link, Delta 3, WorldCom, MCI, U.S. West, Bell Atlantic, Sprint, AT&T/Japan, KDD/Japan, Dacom/Korea, Deutsche Telekom in Germany, France Telecom, Telecom Finland, and New Zealand Telecom.
- Level 3 is building a nation wide IP network for telephony.
- Bell Canada has formed 'Emergis' division.
- Bellcore has formed 'Soliant Internet Systems' unit
- Bell Labs has formed 'Elemedia' division

Services (Cont)

- Hosted PBX Service

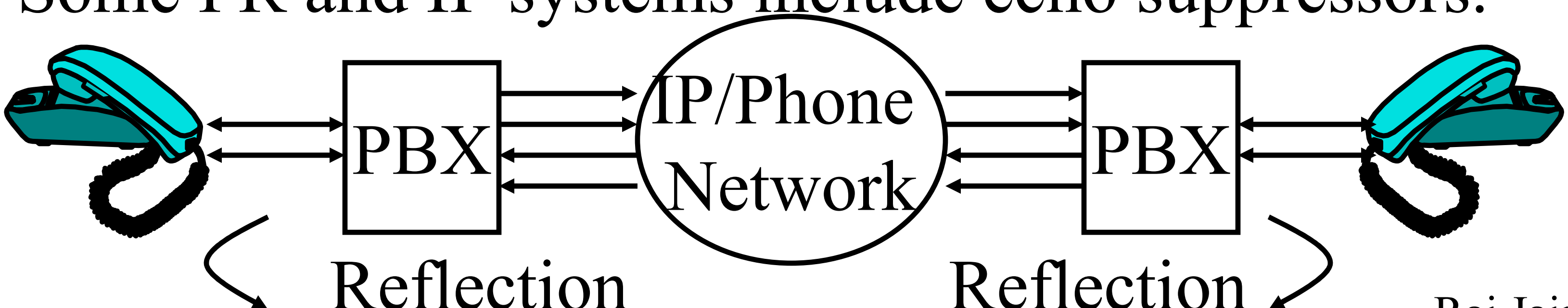


Technical Issues

1. **Large Delay**
   - Normal Phone: 10 ms/kmile ⇒ 30 ms coast-to-coast
   - G.729: 10 ms to serialize the frame + 5 ms look ahead + 10 ms computation = 25 ms one way algorithmic delay
   - G.723.1 = 100 ms one-way algorithmic delay
   - Jitter buffer = 40-60 ms
   - In one survey, 77% users found delay unacceptable.

Technical Issues (Cont)

2. Delay Jitter: Need priority for voice packets. Shorter packets? IP precedence (TOS) field.
3. Frame length: 9 kB at 64 kbps = 1.125 s
   Smaller MTU ⇒ Fragment large packets
4. Lost Packets: Replace lost packets by silence, extrapolate previous waveform
5. Echo cancellation: 2-wire to 4-wire.
   Some FR and IP systems include echo suppressors.



Raj Jain

Technical Issues (Cont)

6. Silence suppression
7. Address translation: Phone # to IP. Directory servers.
8. Telephony signaling: Different PBXs may use different signaling methods.
9. Bandwidth Reservations: Need RSVP.
10. Security: Firewalls may not allow incoming IP traffic
11. Insecurity of internet
12. Voice compression: Load reduction
13. Multiplexing: Subchannel multiplexing
    $\Rightarrow$ Multiple voice calls in one packet.

IP Tax

- RTP, UDP, IP headers increase the bandwidth required

Codec	10 ms Samples/ Frame	Rate	Delay	Bandwidth w/o header compression	Bandwidth with cRTP compression
G.711	1	64 kbps	10ms	112 kbps	81.6 kbps
G.711	2	64 kbps	20ms	96 kbps	80.8 kbps
G.729	1	8 kbps	15ms	40 kbps	9.6 kbps
G.729	2	8 kbps	25ms	24 kbps	11.2 kbps
G.729	4	8 kbps	45ms	16 kbps	8.4 kbps

Other Issues

1. Per-minute distance-sensitive charge vs
   flat time-insensitive distance-insensitive charge
2. Video requires a bulk of bits but costs little.
   Voice is expensive. On IP, bits are bits.
3. National regulations and government monopolies
   ⇒ Many countries forbid voice over IP
   In Hungary, Portugal, etc., it is illegal to access a web site with VOIP s/w. In USA, Association of Telecommunications Carriers (ACTA) petitioned FCC to levy universal access charges on ISPs
4. Modem traffic can't get more than 2400 bps.
5. No lossy compression on faxes. Fail on 1-2% pkt loss.

Compression Standards

- G.711: 64 kbps Pulse Code Modulation (PCM)
- G.721:
  - 32 kbps Adaptive Differential PCM (ADPCM).
  - Difference between actual and predicted sample.
  - Used on international circuits
- G.728: 16 kbps Code Excited Linear Prediction (CELP).
- G.729: 8 kbps Conjugate-Structure Algebraic Code Excited Linear Prediction (CS-ACELP).

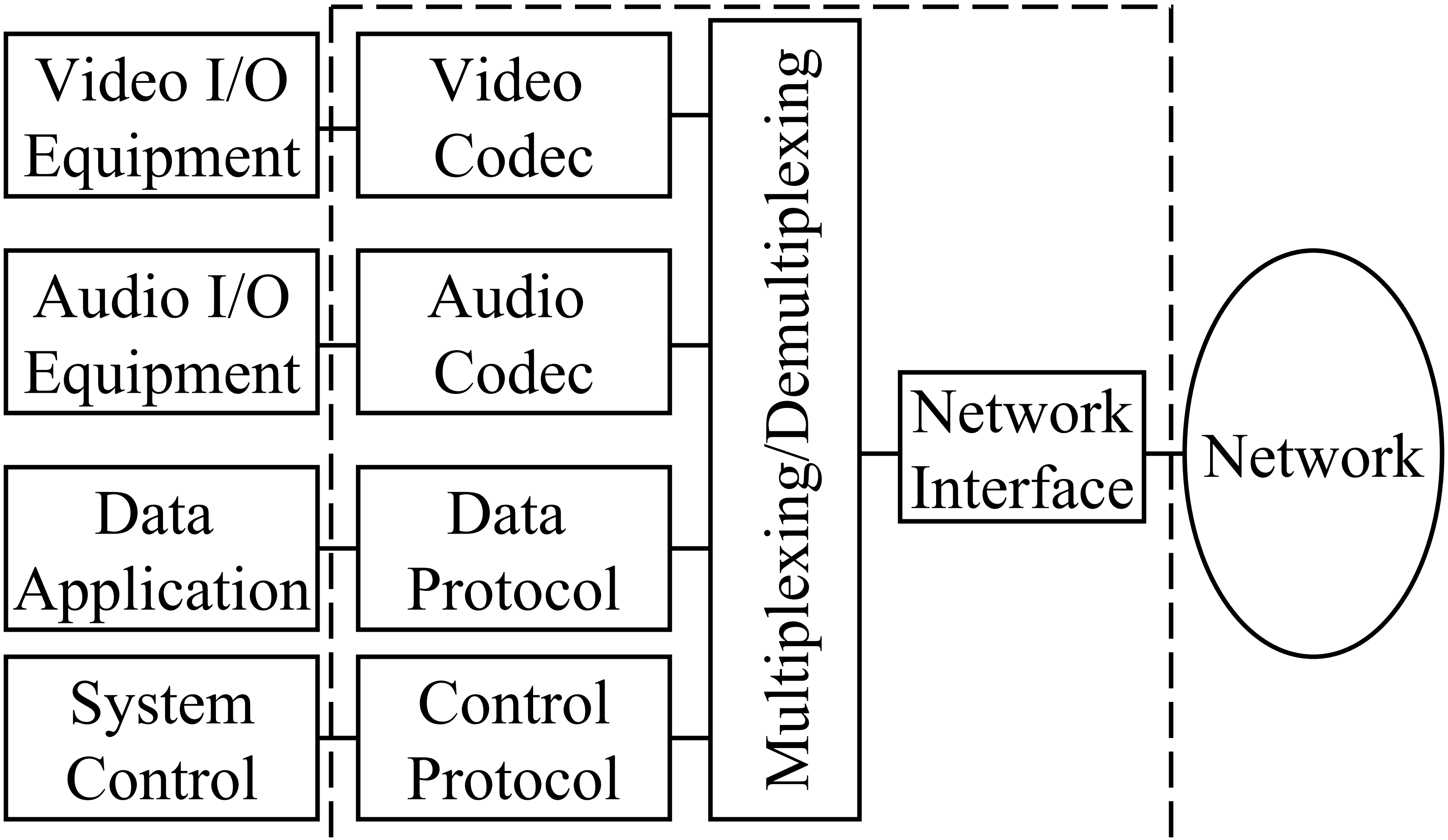
Compression (Cont)

- G.729A:
  - A reduced complexity version in Annex A of G.729.
  - Supported by AT&T, Lucent, NTT.
  - Used in simultaneous voice and data (SVD) modems.
  - Used in Voice over Frame Relay (VFRADs).
  - 4 kbps with proprietary silence suppression.

Compression (Cont)

- G.723.1: Dual rates (5.3 and 6.3 kbps).
  - Packet loss tolerant.
  - Silence suppression option.
  - Recommended by International Multimedia Teleconferencing Consortium (IMTC)'s VOIP forum as default for H.323.
  - Supported by Microsoft, Intel.
  - Mean opinion score (MOS) of 3.8.
    4.0 = Toll quality.

Telephony/Conferencing Systems



Raj Jain

Conferencing Standards

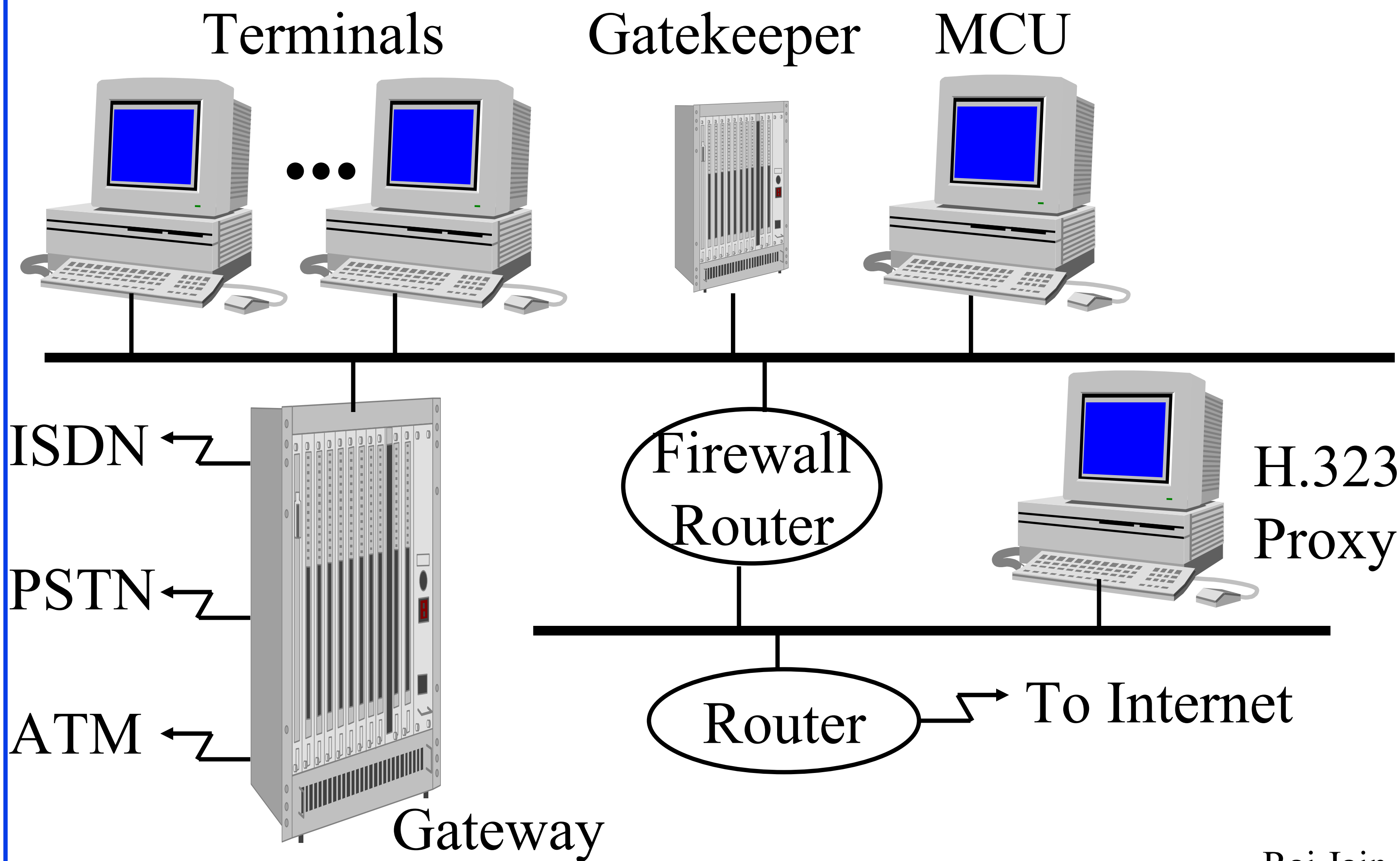
Network	ISDN	ATM	PSTN	LAN	POTs
Conf. Std.	H.320	H.321	H.322	H.323 V1/V2	H.324
Year Audio Codec	1990 G.711, G.722, G.728	1995 G.711, G.722, G.728	1995 G.711, G.722, G.728	1996/1998 G.711, G.722, G.723.1, G.728, G.729	1996 G.723.1, G.729
Audio Rates kbps Video Codec	64, 48-64 H.261	64, 48-64, 16 H.261, H.263	64, 48-64, 16 H.261, H.263	64, 48-64, 16, 8, 5.3/6.3 H.261 H.263	8, 5.3/6.3 H.261 H.263
Data Sharing	T.120	T.120	T.120	T.120	T.120
Control	H.230, H.242	H.242	H.242, H.230	H.245	H.245
Multiplexing	H.221	H.221	H.221	H.225.0	H.223
Signaling	Q.931	Q.931	Q.931	Q.931	-

H.323 Protocols

- Multimedia over LANs
- Provides component descriptions, signaling procedures, call control, system control, audio/video codecs, data protocols

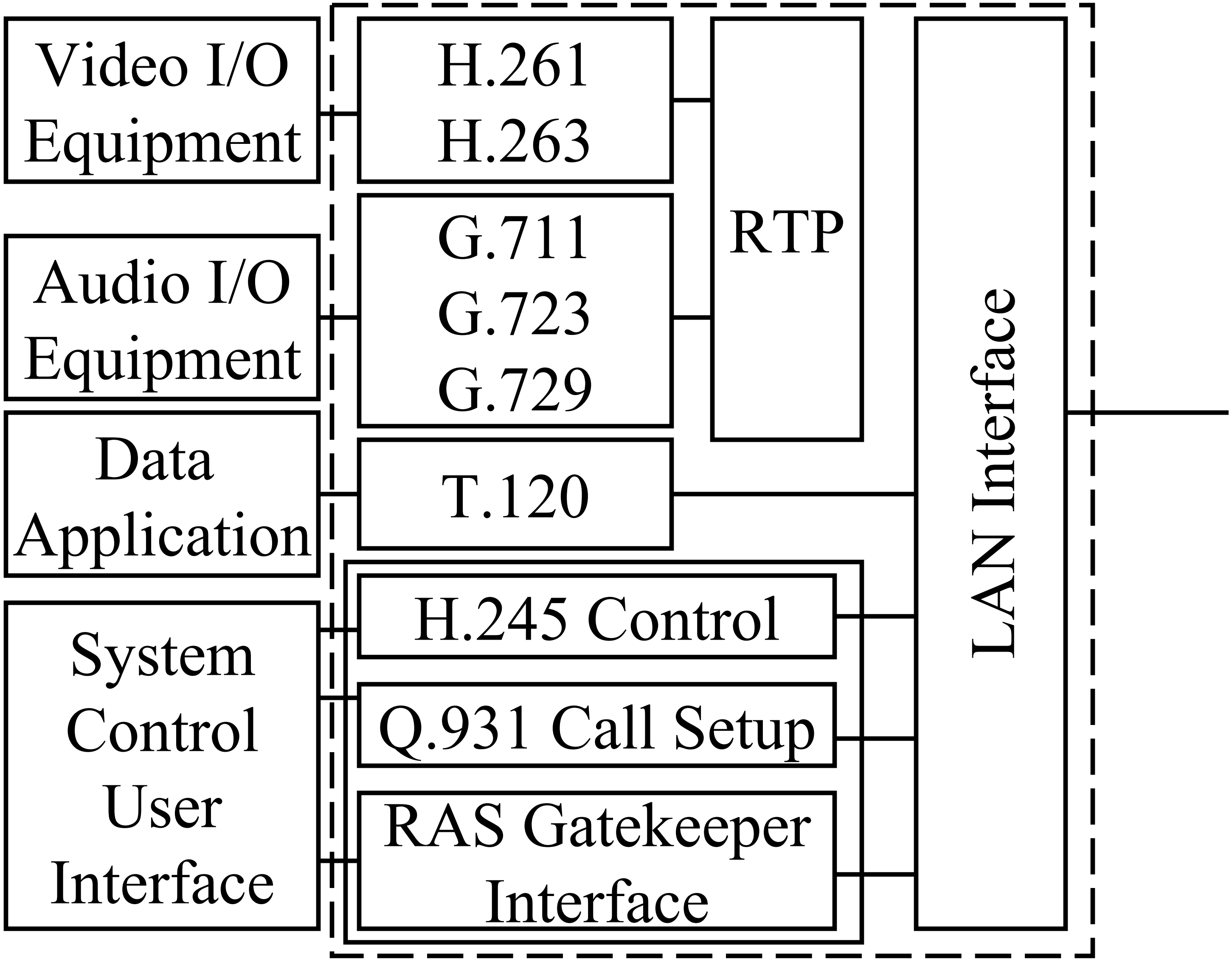
<table>
<tr><td>Video</td><td>Audio</td><td colspan="4">Control and Management</td><td>Data</td></tr>
<tr><td>H.261
H.263</td><td>G.711, G.722,
G.723.1, G.728,
G.729</td><td rowspan="2">RTCP</td><td>H.225.0
RAS</td><td>H.225.0
Signaling</td><td>H.245
Control</td><td>T.124</td></tr>
<tr><td colspan="2">RTP</td><td colspan="3">X.224 Class 0</td><td>T.125</td></tr>
<tr><td colspan="3">UDP</td><td colspan="3">TCP</td><td rowspan="2">T.123</td></tr>
<tr><td colspan="6">Network (IP)</td></tr>
<tr><td colspan="7">Datalink (IEEE 802.3)</td></tr>
</table>

H.323 Components



Raj Jain

H.323 Terminals



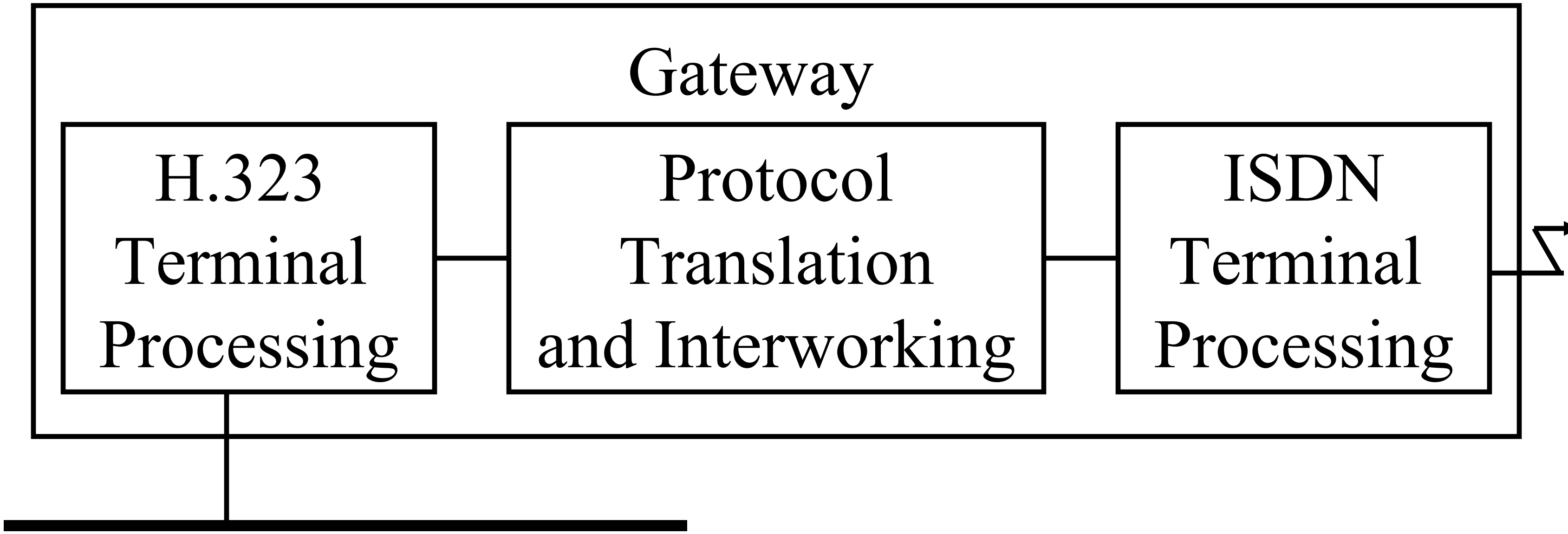
Raj Jain

H.323 Terminals

- Client end points. PCs.
- H.245 to negotiate channel usage and capabilities.
- Q.931 for call signaling and call setup.
- Registration/Admission/Status (RAS) protocol to communicate with gatekeepers.
- RTP/RTCP for sequencing audio and video packets.

H.323 Gateways

- Provide translation between H.323 and other terminal types (PSTN, ISDN, H.324)
- Not required for communication with H.323 terminals on the same LAN.



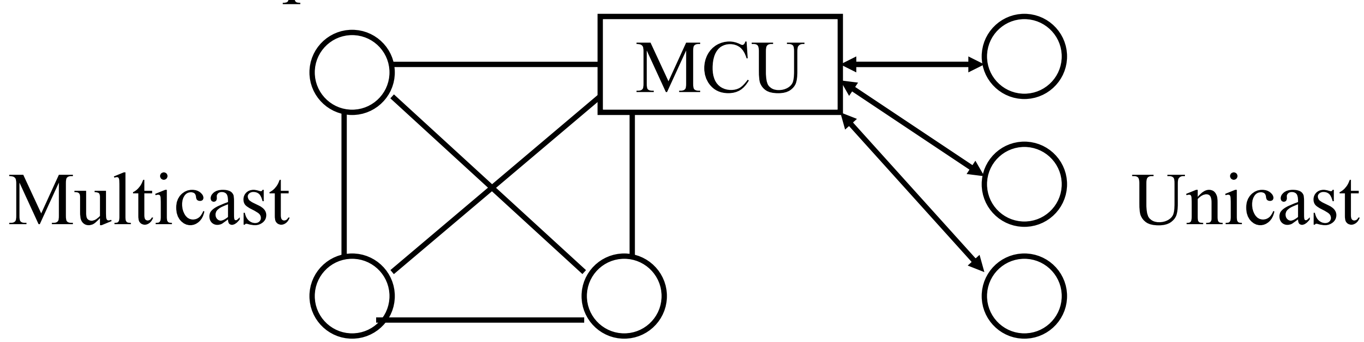
Raj Jain

H.323 Gatekeepers

- Provide call control services to registered end points.
- One gatekeeper can serve multiple LANs
- Address translation (LAN-IP)
- Admission Control: Authorization
- Bandwidth management
  (Limit number of calls on the LAN)
- Zone Management: Serve all registered users within its zone of control
- Forward unanswered calls
- May optionally handle Q.931 call control

H.323 MCUs

- Multipoint Control Units
- Support multipoint conferences
- Multipoint controller (MC) determines common capabilities.
- Multipoint processor (MP) mixes, switches, processes media streams.
- MP is optional. Terminals multicast if no MP.



Raj Jain

Session Initiation Protocol (SIP)

- Application level signaling protocol
- Allows creating, modifying, terminating sessions with one or more participants
- Carries session descriptions (media types) for user capabilities negotiation
- Supports user location, call setup, call transfers
- Supports mobility by proxying and redirection
- Allows multipoint control unit (MCU) or fully meshed interconnections
- Gateways can use SIP to setup calls between them

SIP (Cont)

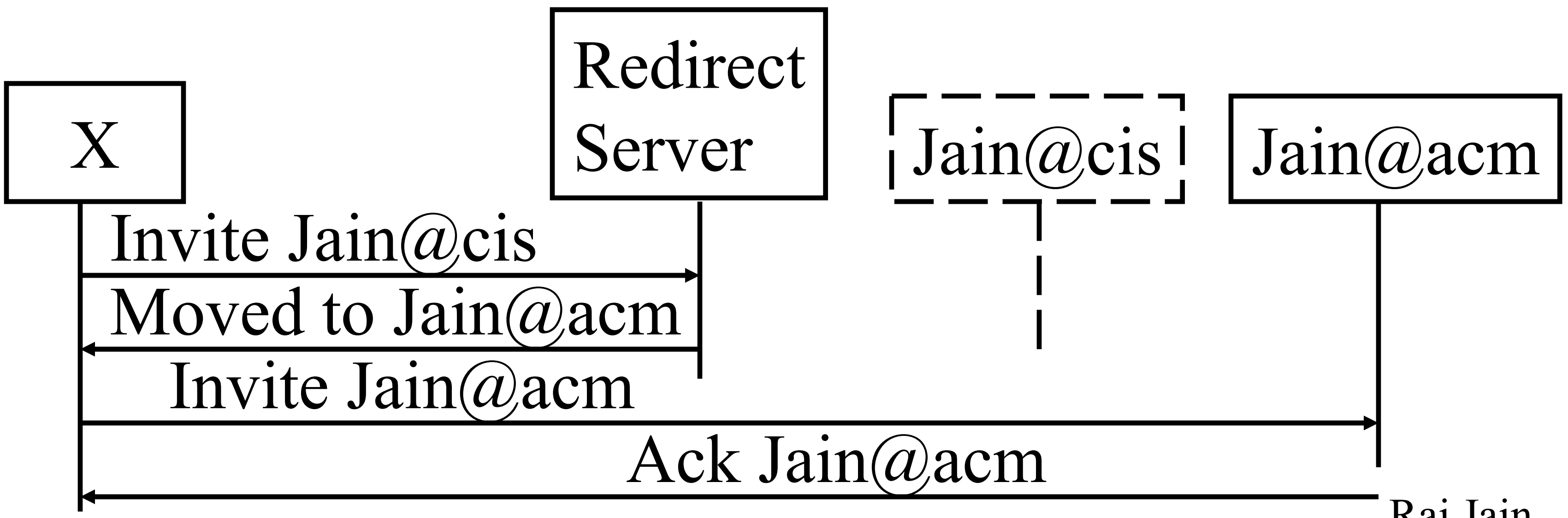
- SIP works in conjunction with other IP protocols for multimedia:
  - RSVP for reserving network resources
  - RTP/RTCP/RTSP for transporting real-time data
  - Session description protocol (SDP) for describing multimedia session
- Can also be used to determine whether party can be reached via H.323, find H.245 gateway/user address
- SIP is text based (similar to HTTP)
  ⇒ SIP messages can be easily generated by humans, CGI, Perl, or Java programs.

SIP (Cont)

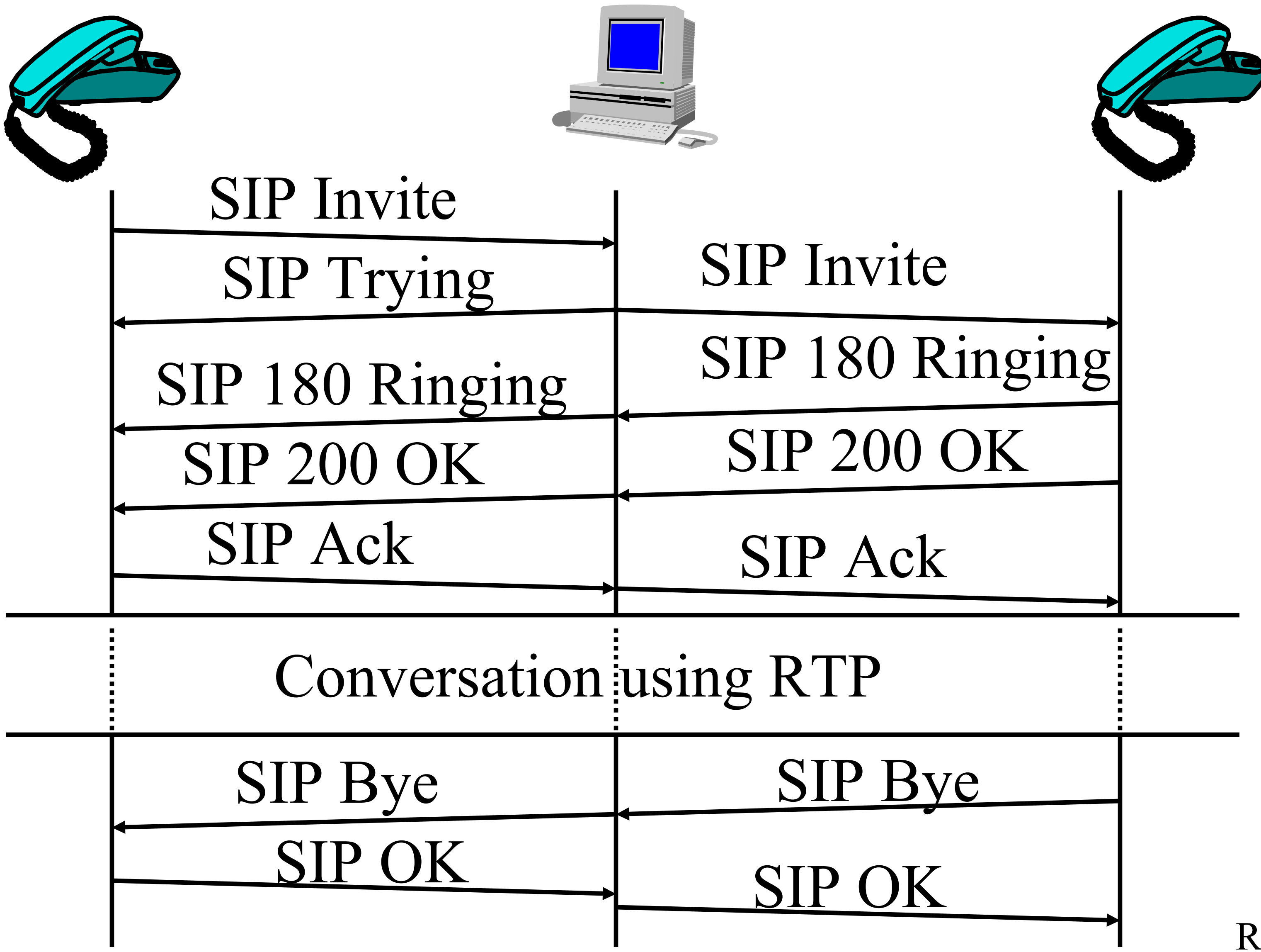
- SIP Uniform Resource Identfiers (URIs):
  Similar to email URLs
  sip:jain@cis.ohio-state.edu
  sip:+1-614-292-3989:123@osu.edu?subject=lecture
- SIP can use UDP or TCP
- SIP messages are sent to SIP servers:
  - Registrar: Clients register and tell their location to it
  - Location: Given name, returns possible addresses for a user. Like Directory service or DNS.
  - Redirect: Returns current address to requesters
  - Proxy: Intermediary. Acts like a server to internal client and like a client to external server

Locating using SIP

- Allows locating a callee at different locations
- Callee registers different locations with Registrar
- Servers can also use finger, rwhois, ldap to find a callee
- SIP Messages: Ack, Bye, Invite, Register, Redirection, ...



SIP Proxy

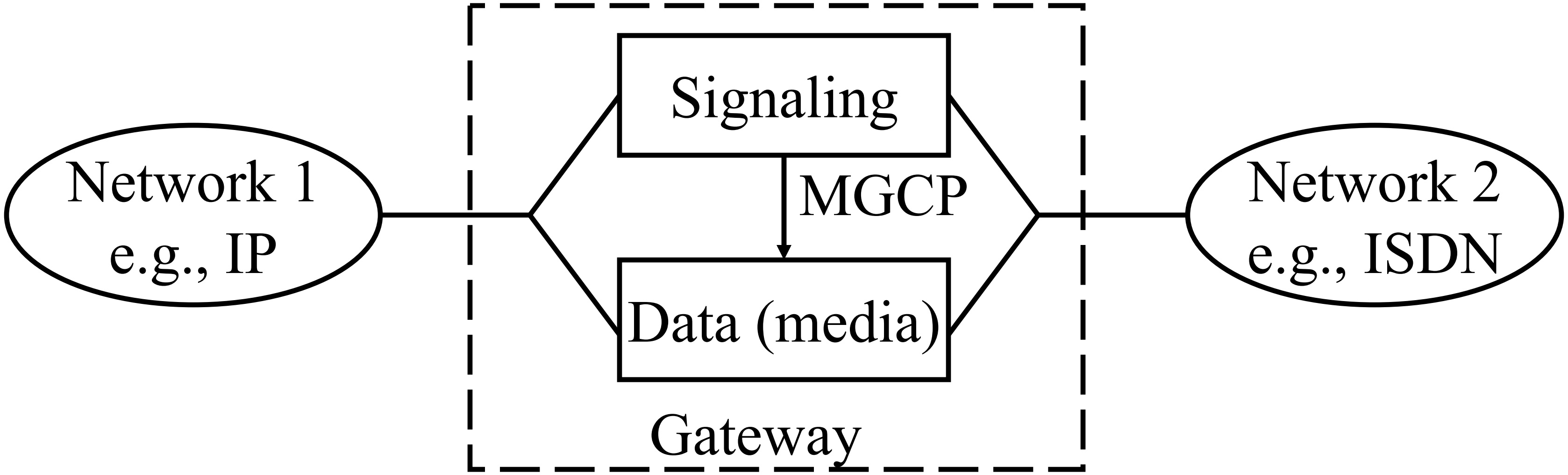


Raj Jain

Sample SIP Products and Services

- CISCO SIP Gateway, SIP Phone, SIP Proxy Server
- 3COM SIP Signaling Server software
- Level 3 uses SIP as part of all-IP carrier services
- Agilent Technologies and RADCOM make SIP network analyzers and test equipment
- Hughes Software Systems and Indigo Software sell SIP protocol stack for OEMs
- SIP support in **Windows XP**
- Ref: http://pulver.com/sip/products.html

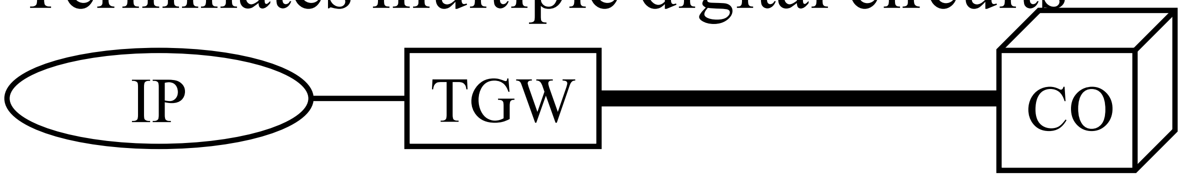
Media Gateway Control Protocol



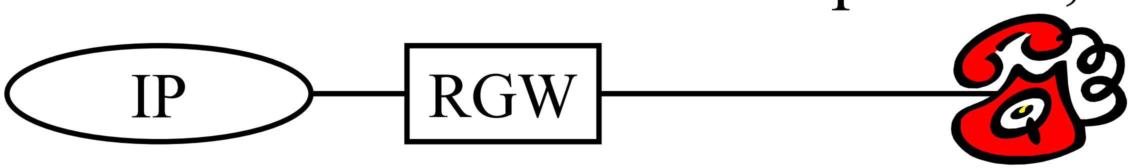
- Gateway = Signaling Fns + Media Transfer Fns
- Call Agents: Signaling functions ⇒ Intelligent
  ⇒ More complex ⇒ Fewer
  ⇒ Media Gateway Controller (MGC) or Softswitch
  ⇒ Control multiple media gateways ⇒ Need MGCP
- MGCP =Simple Gateway Control Protocol (SGCP)
  \+ Internet Protocol Device Control (IPDC)

Media Gateways: Examples

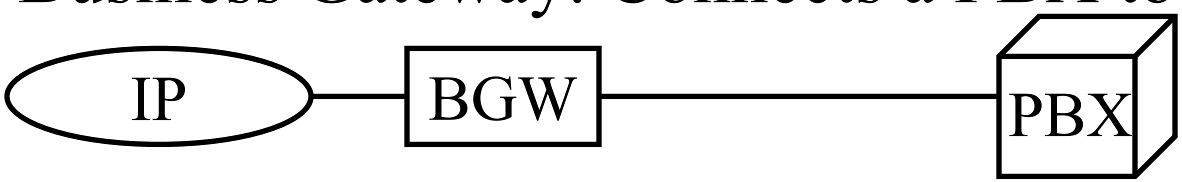
- Trunking Gateway: Connects a PSTN trunk to VOIP
  Terminates multiple digital circuits



- Residential Gateway: Connects a RJ11 to VOIP
  Will be used in cable set-top boxes, xDSL, ...



- Business Gateway: Connects a PBX to VOIP



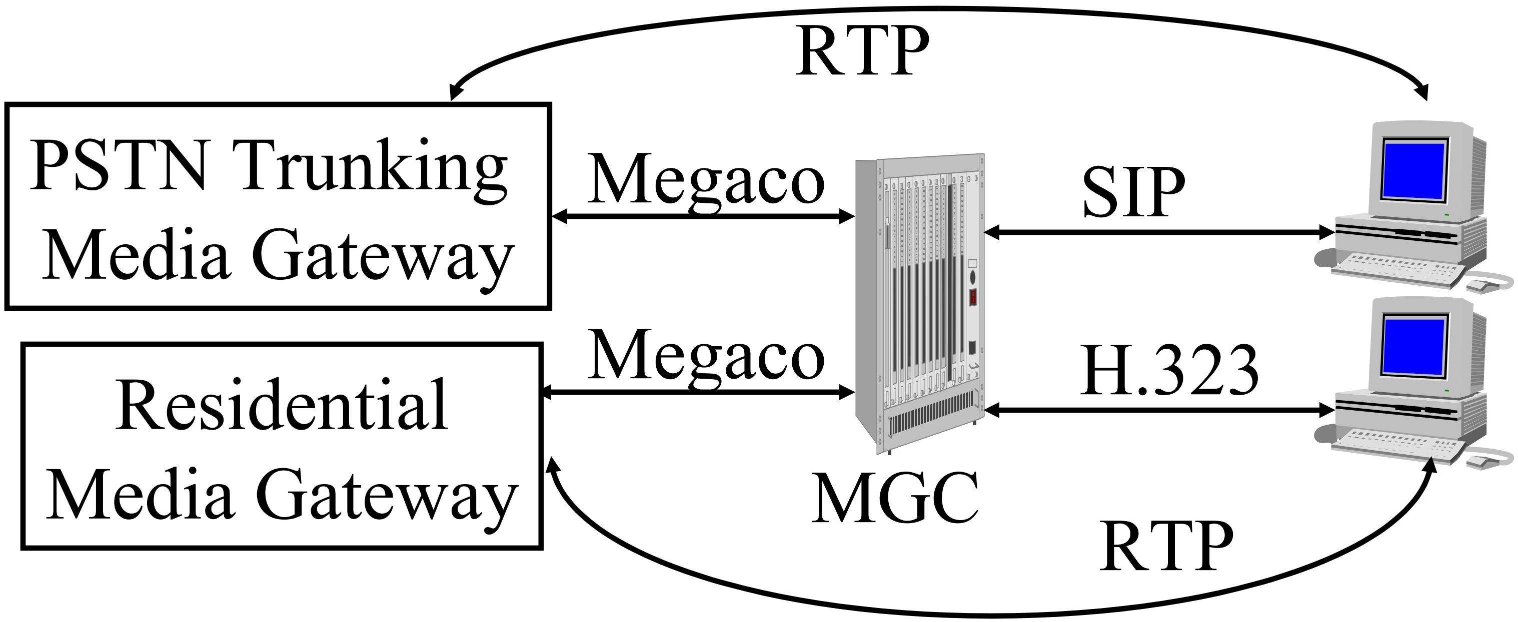
- Network Access Servers: Answer data + VOIP calls

Media Gateway Control (Megaco)

History:

- IP Device Control (IPDC)
  + Simple Gateway Control Protocol (SGCP)
  ⇒ Media Gateway Control Protocol (MGCP)
  Released as RFC 2705 Oct 99. Not fully accepted.
- May 99: ITU SG-16 initiated Gateway Control Protocol H.GCP starting from MDCP
- Summer 99: ITU SG16 and IETF Megaco agree to work together to create one standard
- November 2000: RFC 3015

Megaco Framework

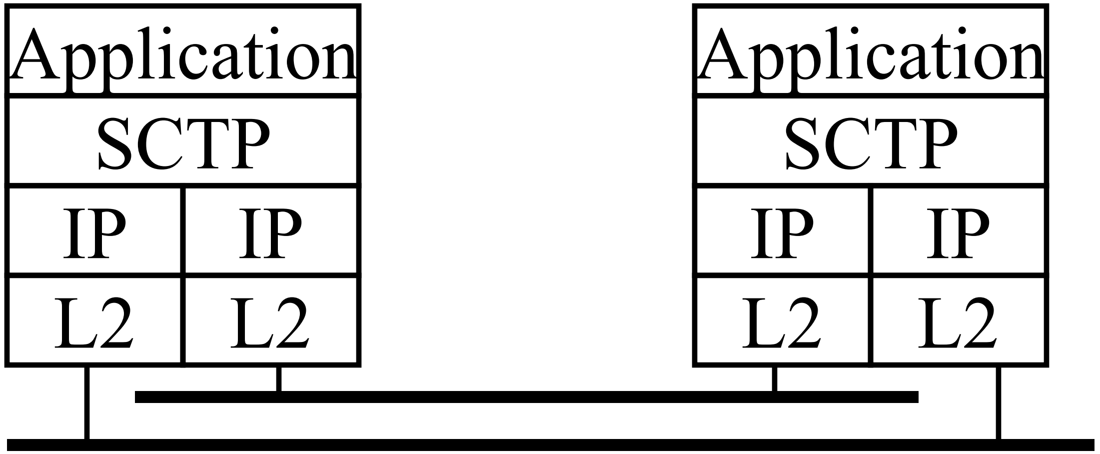


- MGC provides call processing, call routing
  Looks like H.323 gateway to H.323 devices and like a SIP server to SIP devices
- MG provides device control (ringing,…) and connection control (disconnect, connect).

Raj Jain

SIGTRAN and SCTP

- SIGTRAN: Signaling Transport Working Group at IETF
- SCTP:Stream Control Transmission Protocol [RFC2960]
  - Carrier-Grade Level 4 Protocol replacing TCP
  - Allows multiple redundant IP Addresses
  - Multiple parallel streams⇒No head of line blocking
  - Can be used between MGCs or MGs and MGCs



Raj Jain

Summary

- VOIP is ideal for computer-based communications
- IP needs QoS for acceptable quality
- H.323, SIP provide interoperability
- Megaco is the protocol between Media gateways and Media Gateway Controllers/Call agenets/Softswitches
- SCTP provides fault-tolerant transport for signaling

VOIP: Key References

- For a detailed list of references, see http://www.cis.ohio-state.edu/~jain/refs/ref_voip.htm Also reproduced at the end of this tutorial book.
- Voice and Telephony over ATM, http://www.cis.ohio-state.edu/~jain/cis788-99/vtoa/index.html
- VOIP Products,services, and issues, http://www.cis.ohio-state.edu/~jain/cis788-99/voip_products/index.html
- VOIP: Protocols and Standards, http://www.cis.ohio-state.edu/~jain/cis788-99/voip_protocols/index.html

Standards Organizations

- IETF, http://www.ietf.org
- International SoftSwitch Consortium, http://www.softswitch.org
- MultiService Switching Forum, http://msforum.org/
- International Multimedia Teleconferencing Consortium, http://www.imtc.org/
- Voice On The Net (VON) Coalition, http://www.von.org/
- MIT Internet Telephony Consortium, http://itel.mit.edu/
- Enterprise Computer Telephony Forum (ECTF), http://www.ectf.org/
- ITU, http://www.itu.ch

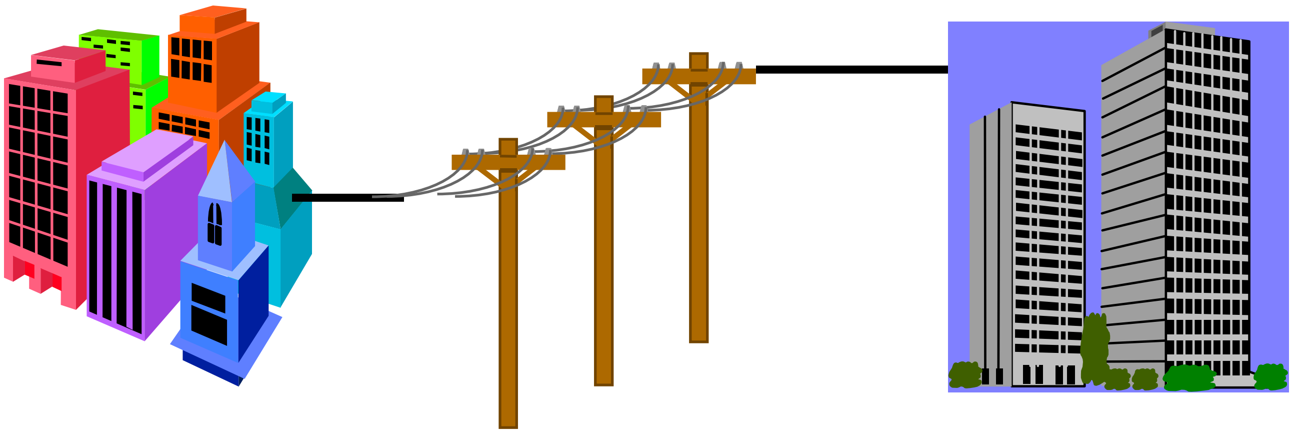
Virtual Private Networks

Raj Jain
Co-Founder and CTO
Nayna Networks
481 Sycamore Dr, Milpitas, CA 95035
www.nayna.com and
http://www.cis.ohio-state.edu/~jain/

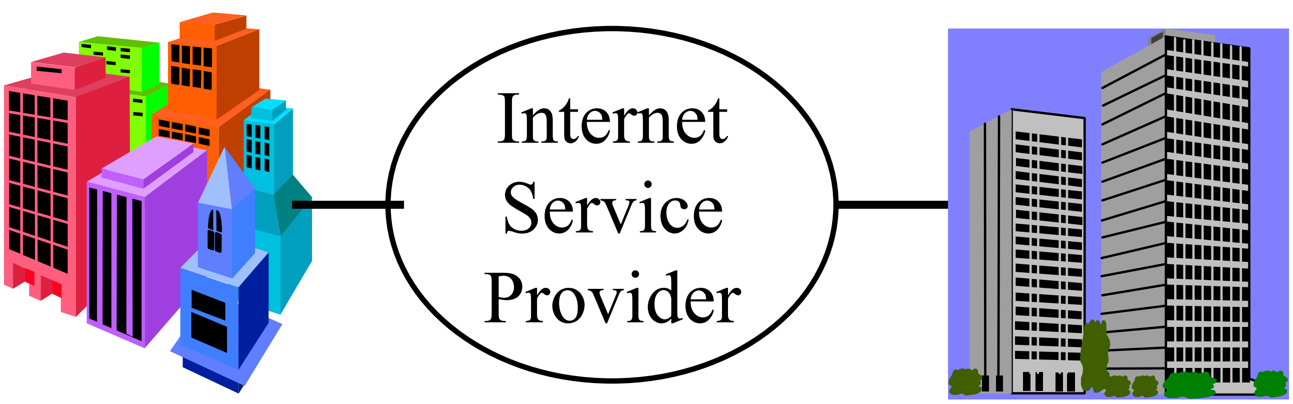
- Types of VPNs
- When and why VPN?
- VPN Design Issues
- Security Issues
- VPN Examples: PPTP, L2TP, IPSec
- Authentication Servers: RADIUS and DIAMETER
- VPNs using MPLS, Virtual Routers

What is a VPN?

- Private Network: Uses leased lines



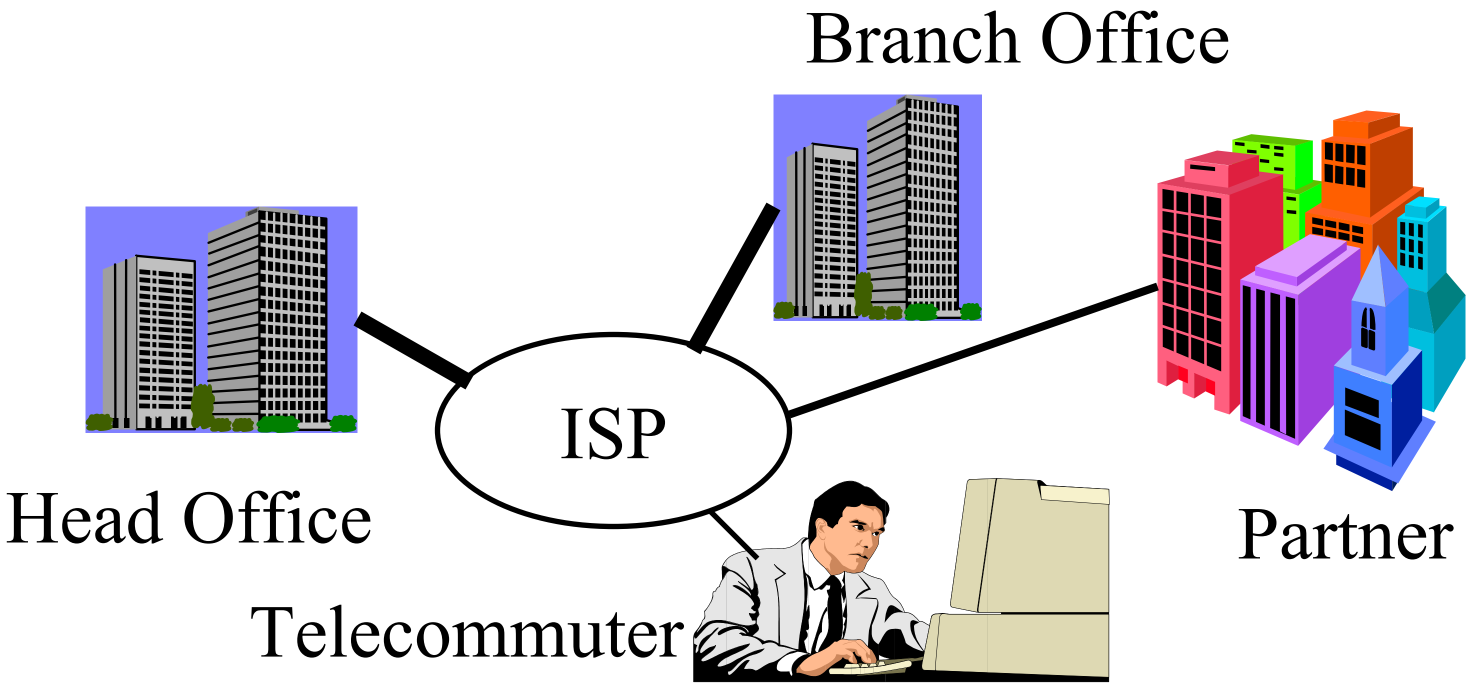
- *Virtual* Private Network: Uses public Internet



Raj Jain

Types of VPNs

- WAN VPN: Branch offices
- Access VPN: Roaming Users
- Extranet VPNs: Suppliers and Customers

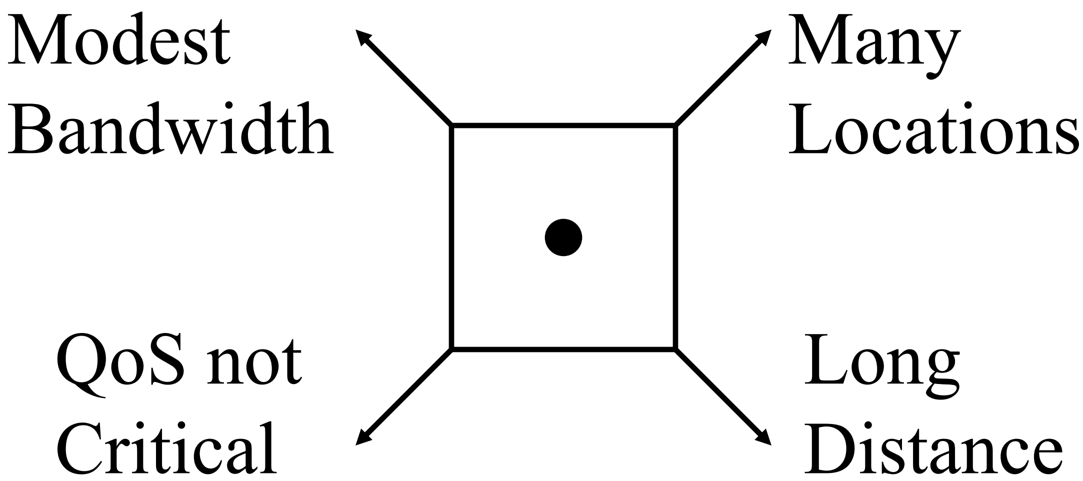


Raj Jain

Why VPN?

- Reduced telecommunication costs
- Less administration ⇒ 60% savings (Forester Res.)
- Less expense for client and more income for ISPs
- Long distance calls replaced by local calls
- Increasing mobility ⇒ More remote access
- Increasing collaborations
  ⇒ Need networking links with partners

When to VPN?



- More Locations, Longer Distances, Less Bandwidth/site, QoS less critical
  ⇒ VPN more justifiable
- Fewer Locations, Shorter Distances, More Bandwidth/site, QoS more critical
  ⇒ VPN less justifiable

VPN Design Issues

1. Security
2. Address Translation
3. Performance: Throughput, Load balancing (round-robin DNS), fragmentation, # of connections
4. Bandwidth Management: RSVP
5. Availability: Good performance at all times, automatic failover
6. Scalability: Number of locations/Users
7. Interoperability: Among vendors, ISPs, customers (for extranets) $\Rightarrow$ Standards Compatibility, With firewall

Design Issues (Cont)

8. Compression: Reduces bandwidth requirements
9. Manageability: SNMP, Browser based, Java based, centralized/distributed
10. Accounting, Auditing, and Alarming
11. Protocol Support: IP, non-IP (IPX)
12. Platform and O/S support: Windows, UNIX, MacOS, HP/Sun/Intel
13. Installation: Changes to desktop or backbone only
14. Legal: Exportability, Foreign Govt Restrictions, Key Management Infrastructure (KMI) initiative
    $\Rightarrow$ Need key recovery

Security Threat Statistics

- DoD networks were attacked 250000 times in 1995 (well before Internet popularity)
- Of 38000 friendly attacks, 65% succeded
- Only 4% of successful attacks were noticed by network administrators
- Only a small fraction of those noticed were reported to authorities
- FBI reports 163 organizations lost $123M in 1999
- Ref: M. Markow, "VPN for Dummies," IDG Books, 1999.

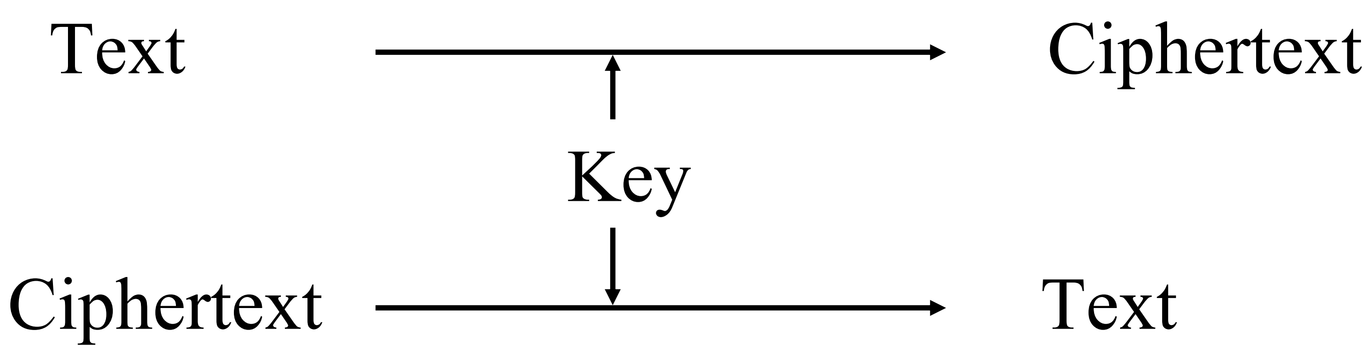
Security 101

- Integrity: Received = sent?
- Availability: Legal users should be able to use.
  Ping continuously ⇒ No useful work gets done.
- Confidentiality and Privacy:
  No snooping or wiretapping
- Authentication: You are who you say you are.
  A student at Dartmouth posing as a professor canceled the exam.
- Authorization = Access Control
  Only authorized users get to the data

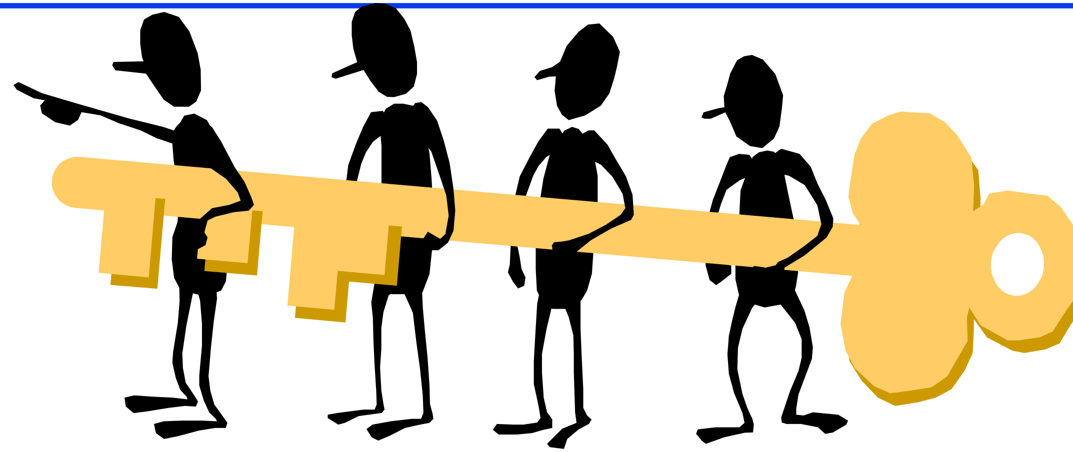
Raj Jain

Secret Key Encryption

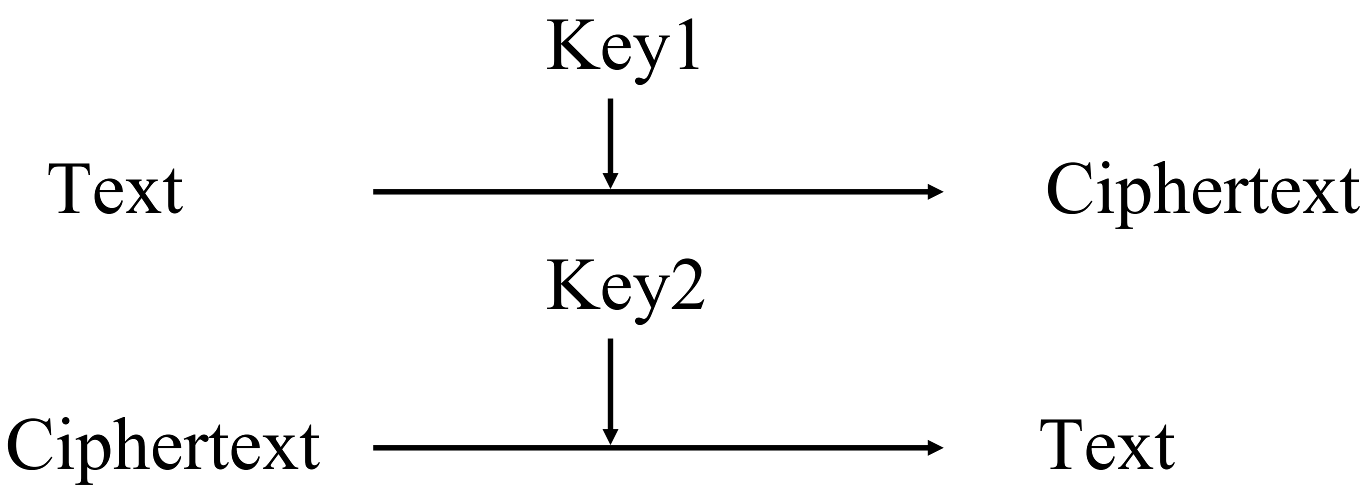
- Encrypted_Message = Encrypt(Key, Message)
- Message = Decrypt(Key, Encrypted_Message)
- Example: Encrypt = division
- 433 = 48 R 1 (using divisor of 9)



Public Key Encryption



- Invented in 1975 by Diffie and Hellman
- Encrypted_Message = Encrypt(Key1, Message)
- Message = Decrypt(Key2, Encrypted_Message)



Public Key Encryption

- RSA: Encrypted_Message = m^3 mod 187
- Message = Encrypted_Message^{107} mod 187
- Key1 = <3,187>, Key2 = <107,187>
- Message = 5
- Encrypted Message = 5^3 = 125
- Message = 125^{107} mod 187
  = $125^{(64+32+8+2+1)}$ mod 187
  = {(125^{64} mod 187)(125^{32} mod 187)...
  (125^2 mod 187)(125)} mod 187 = 5
- 125^4 mod 187 = (125^2 mod 187)^2 mod 187

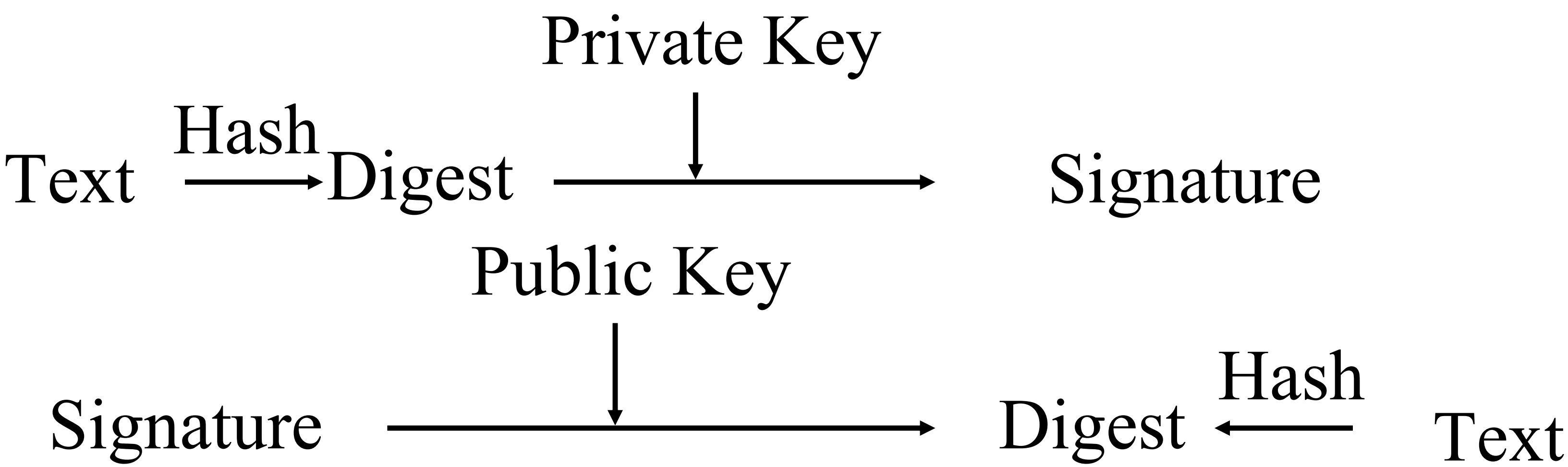
Raj Jain

Public Key (Cont)

- One key is private and the other is public
- Message = Decrypt(Public_Key,
  Encrypt(Private_Key, Message))
- Message = Decrypt(Private_Key,
  Encrypt(Public_Key, Message))

Digital Signature

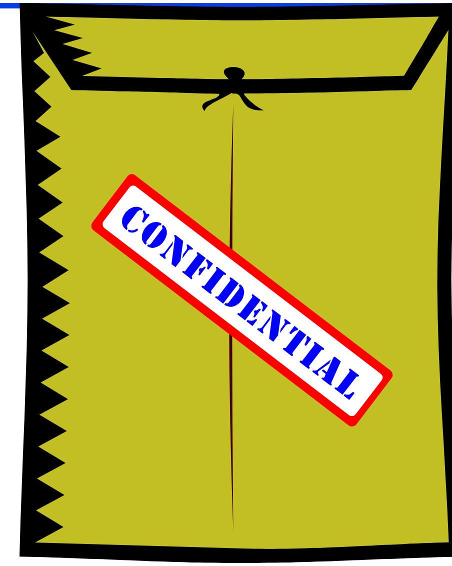
- Message Digest = Hash(Message)
- Signature = Encrypt(Private_Key, Hash)
- Hash(Message) = Decrypt(Public_Key, Signature)
  ⇒ Authentic



Digital Certificates

- Like driver license or passport
- Digitally signed by Certificate authority (CA) - a trusted organization
- Public keys are distributed with certificates
- CA uses its public key to sign the certificate
  $\Rightarrow$ Hierarchy of trusted authorities
- X.509 Certificate includes: Name, organization, effective date, expiration date, public key, issuer's CA name, Issuer's CA signature

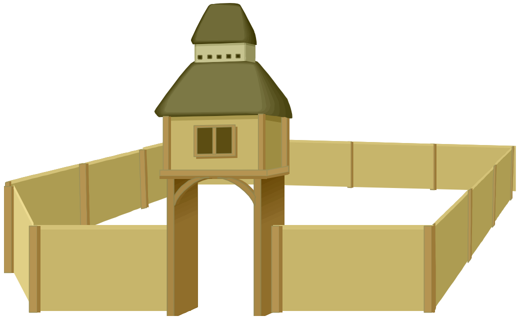
Confidentiality

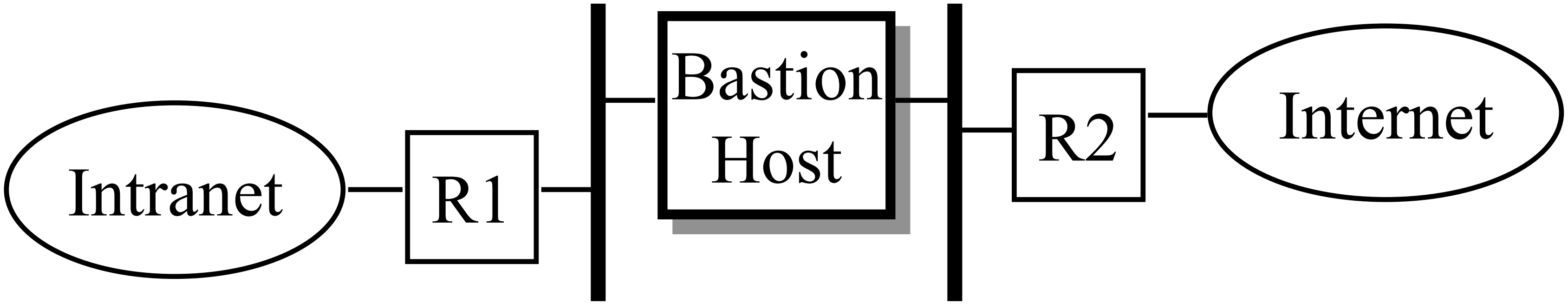


- User 1 to User 2:
- Encrypted_Message
  = Encrypt(Public_Key2,
  Encrypt(Private_Key1, Message))
- Message = Decrypt(Public_Key1,
  Decrypt(Private_Key2, Encrypted_Message)
  ⇒ Authentic and Private

Your Public Key | My Private Key | Message

Firewall: Bastion Host

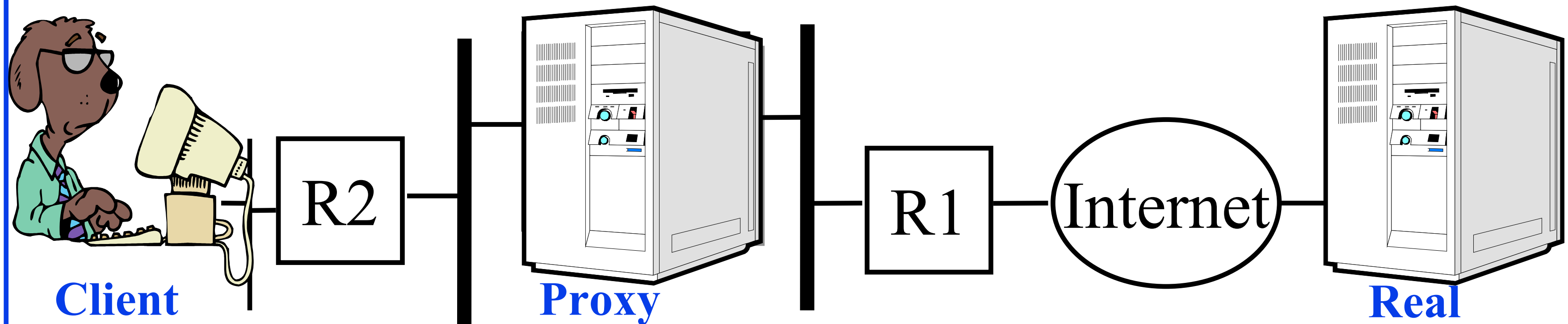




- Bastions overlook critical areas of defense, usually having stronger walls
- Inside users log on the Bastion Host and use outside services.
- Later they pull the results inside.
- One point of entry. Easier to manage security.

Raj Jain

Proxy Servers



- Specialized server programs on bastion host
- Take user's request and forward them to real servers
- Take server's responses and forward them to users
- Enforce site security policy ⇒ Refuse some requests.
- Also known as application-level gateways
- With special "Proxy client" programs, proxy servers are almost transparent

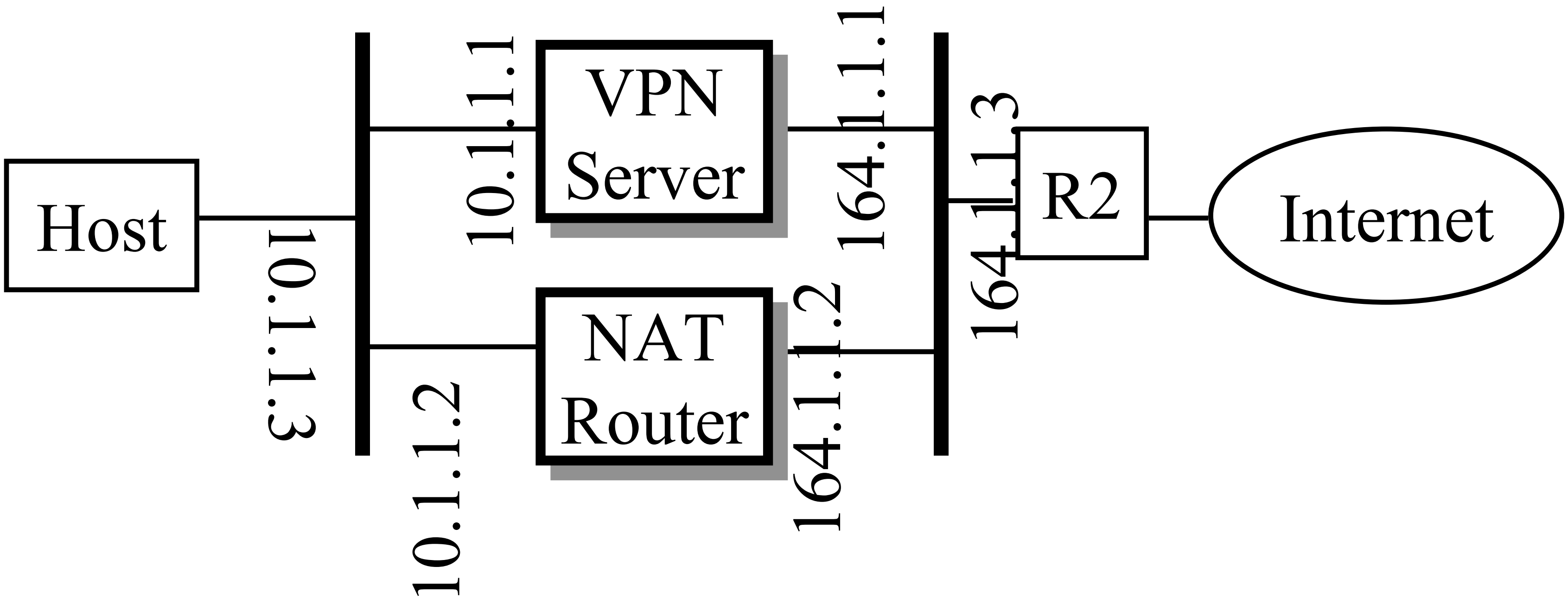
VPN Security Issues

- Authentication methods supported
- Encryption methods supported, Hardware encryption
- Key Management
- Data stream filtering for viruses, JAVA, active X
- Supported certificate authorities (X.509, Entrust, VeriSign)
- Encryption Layer: Datalink, network, session, application. Higher Layer $\Rightarrow$ More granular
- Granularity of Security: Departmental level, Application level, Role-based
- Key lengths: Twinkle optical computer can crack 512-bit RSA in 3 days. Need at least 1024-bit keys.

Private Addresses

- 32-bit Address ⇒ 4 Billion addresses max
- Subnetting ⇒ Limit is much lower
- Shortage of IP address ⇒ Private addresses
- Frequent ISP changes ⇒ Private address
- Private ⇒ Not usable on public Internet
- RFC 1918 lists such addresses for private use
- Prefix = 10/8, 172.16/12, 192.168/16
- Example: 10.207.37.234

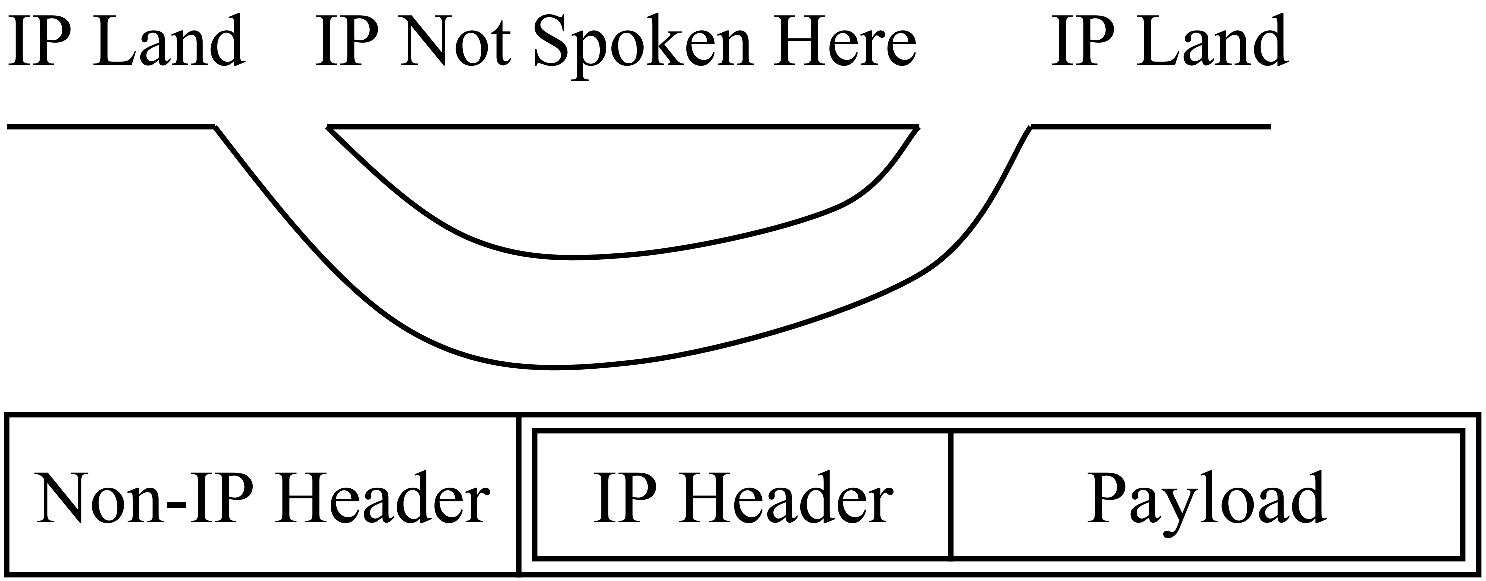
Address Translation



- NAT = Network Address Translation
  Like Dynamic Host Configuration Protocol (DHCP)
- IP Gateway: Like Firewall
- Tunneling: Encaptulation

Raj Jain

Tunnel



- Tunnel = Encaptulation
- Used whenever some feature is not supported in some part of the network, e.g., multicasting, mobile IP

VPN Tunneling Protocols

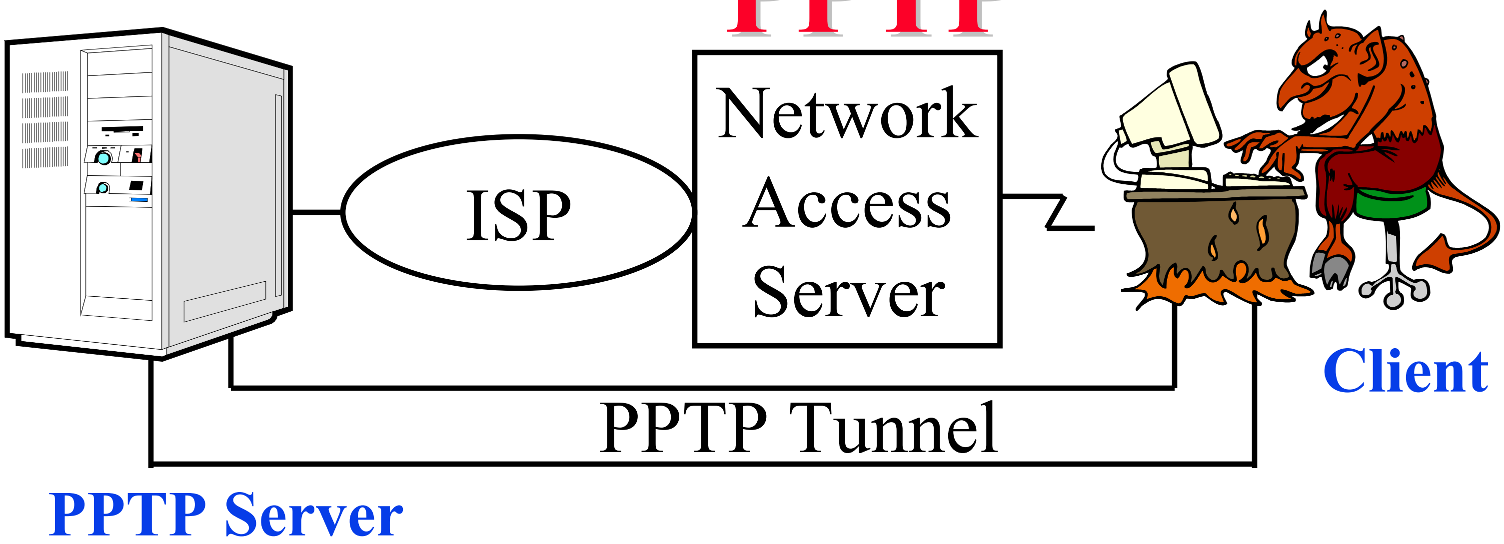
- GRE: Generic Routing Encaptulation (RFC 1701/2)
- PPTP: Point-to-point Tunneling Protocol
- L2F: Layer 2 forwarding
- L2TP: Layer 2 Tunneling protocol
- ATMP: Ascend Tunnel Management Protocol
- DLSW: Data Link Switching (SNA over IP)
- IPSec: Secure IP
- Mobile IP: For Mobile users

GRE

Delivery Header	GRE Header	Payload

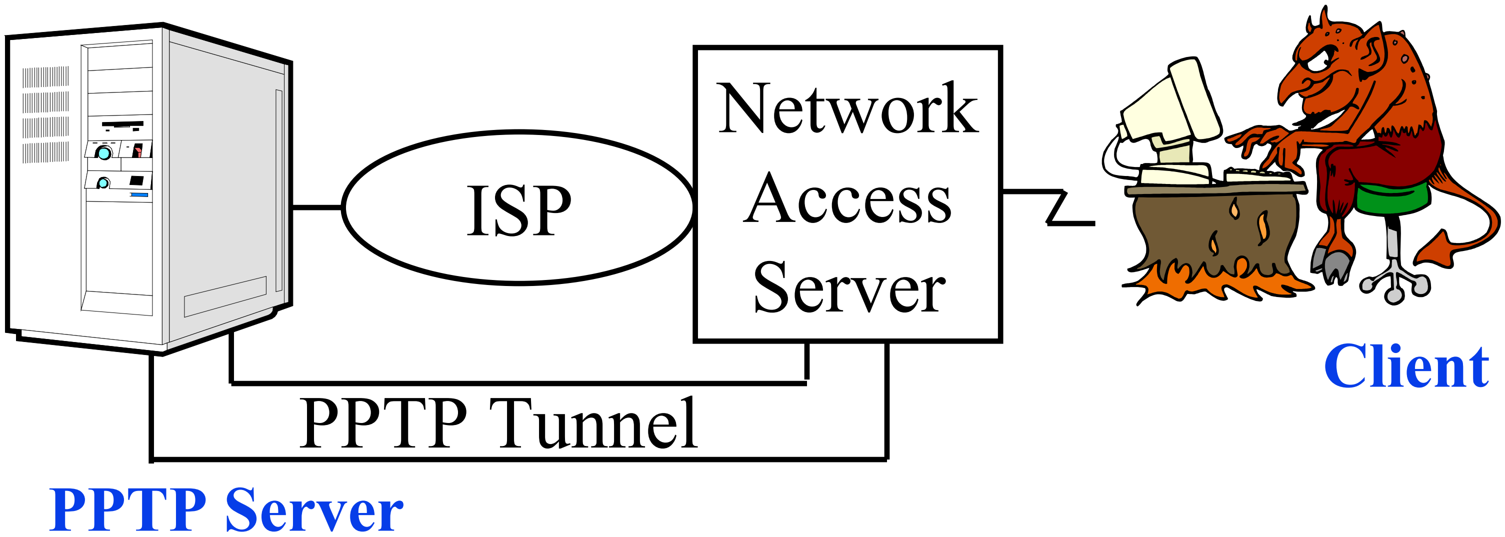
- Generic Routing Encaptulation (RFC 1701/1702)
- Generic ⇒ X over Y for any X or Y
- Optional Checksum, Loose/strict Source Routing, Key
- Key is used to authenticate the source
- Over IPv4, GRE packets use a protocol type of 47
- Allows router visibility into application-level header
- Restricted to a single provider network ⇒ end-to-end

PPTP



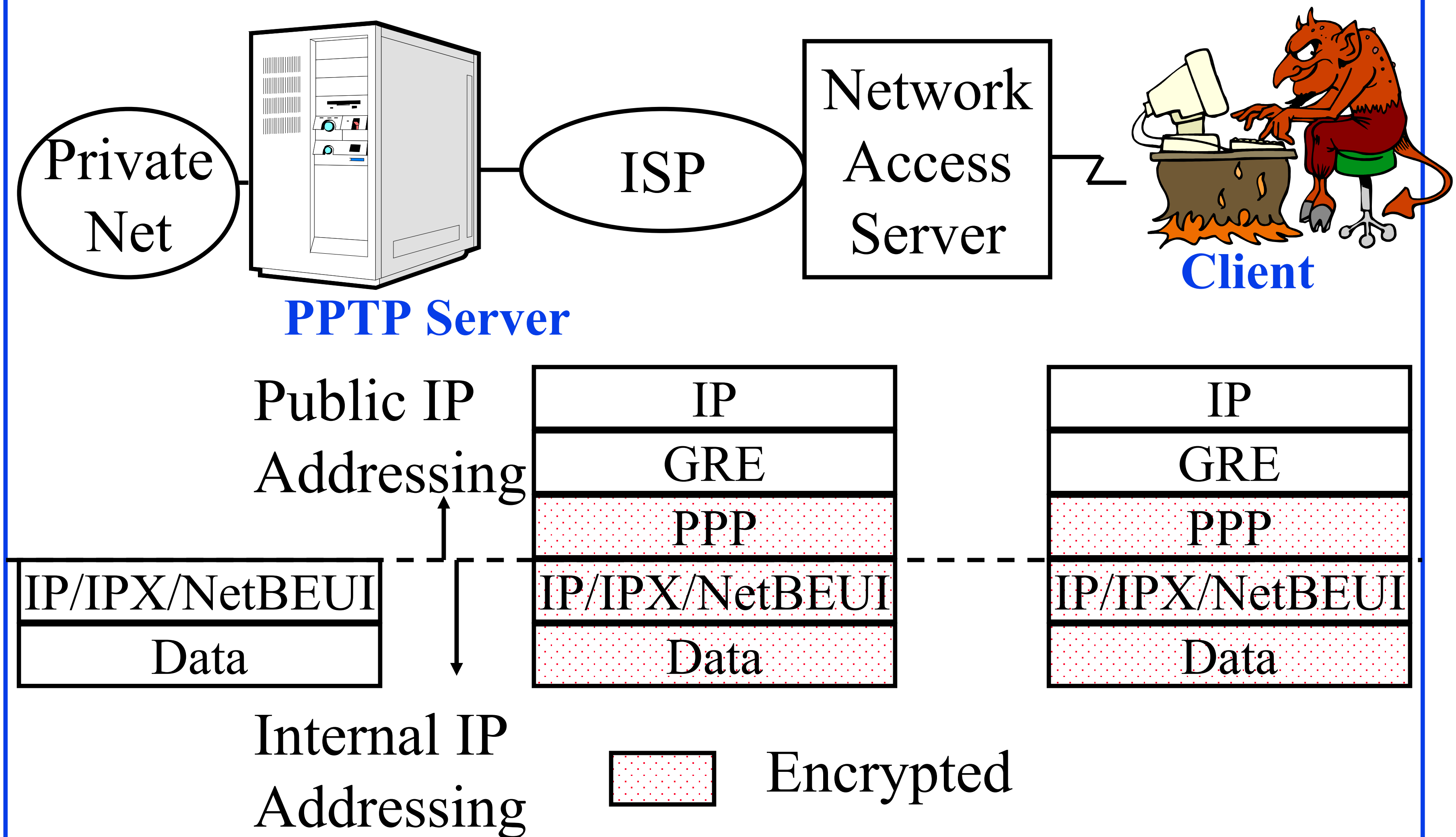
- PPTP = Point-to-point Tunneling Protocol
- Developed jointly by Microsoft, Ascend, USR, 3Com and ECI Telematics
- PPTP server for NT4 and clients for NT/95/98
- MAC, WFW, Win 3.1 clients from Network Telesystems (nts.com)

PPTP with ISP Support



- PPTP can be implemented at Client or at NAS
- With ISP Support: Also known as Compulsory Tunnel
- W/O ISP Support: Voluntary Tunnels

PPTP Packets

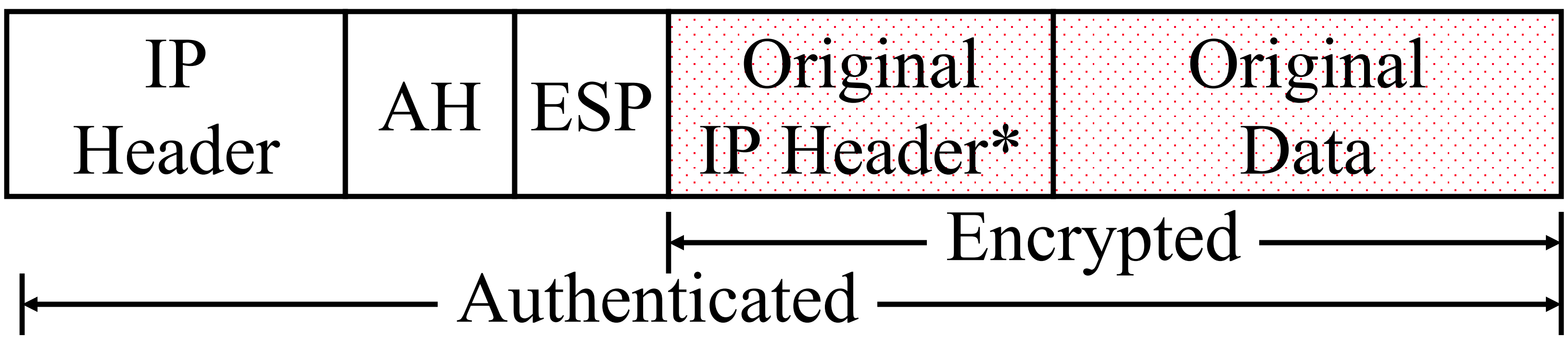


L2TP

- Layer 2 Tunneling Protocol
- L2F = Layer 2 Forwarding (From CISCO)
- L2TP = L2F + PPTP
  Combines the best features of L2F and PPTP
- Will be implemented in NT5
- Easy upgrade from L2F or PPTP
- Allows PPP frames to be sent over non-IP (Frame relay, ATM) networks also (PPTP works on IP only)
- Allows multiple (different QoS) tunnels between the same end-points. Better header compression. Supports flow control

IPSec

- Secure IP: A series of proposals from IETF
- Separate Authentication and privacy
- Authentication Header (AH) ensures data integrity and authenticity
- Encapsulating Security Protocol (ESP) ensures privacy and integrity



* Optional

IPSec (Cont)

- Two Modes: Tunnel mode, Transport mode
- Tunnel Mode ⇒ Original IP header encrypted
- Transport mode ⇒ Original IP header removed. Only transport data encrypted.
- Supports a variety of encryption algorithms
- Better suited for WAN VPNs (vs Access VPNs)
- A reference implementation (Cerberus) IPSec and interoperability tester are available from NIST

Cerberus = three headed dog guarding the underworld

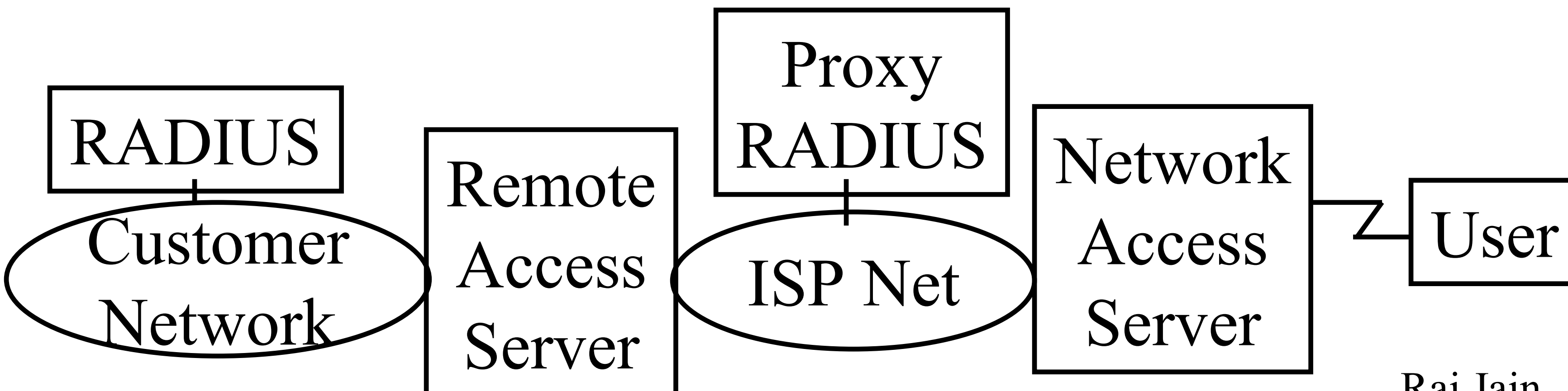
Application Level Security

- Secure HTTP
- Secure MIME
- Secure Electronic Transaction (SET)
- Private Communications Technology (PCT)

Raj Jain

RADIUS

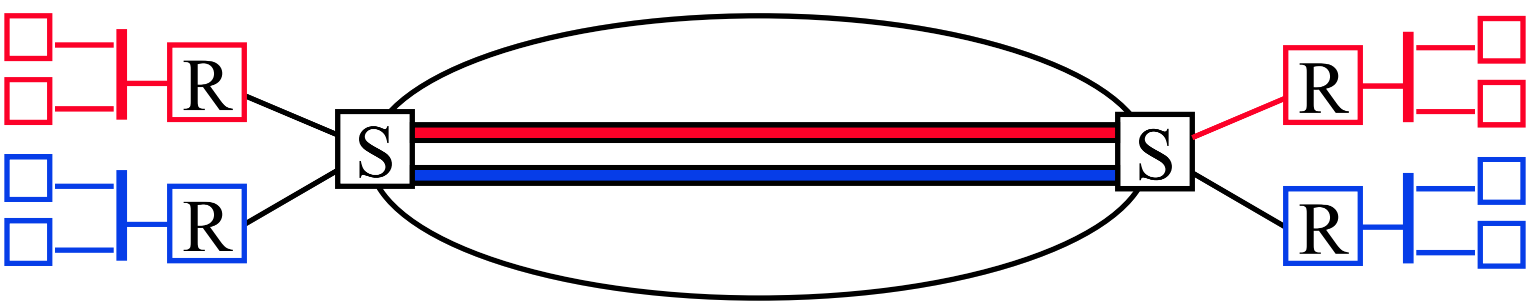
- Remote Authentication Dial-In User Service
- Central point for Authorization, Accounting, and Auditing data ⇒ AAA server
- Network Access servers get authentication info from RADIUS servers
- RADIUS Proxy Servers ⇒ ISP roaming alliances
- Diameter = Enhanced Radius



Raj Jain

VPN Support with MPLS

- Multiprotocol Label Switching
- Separate LSPs for customer traffic
- Issues:
  - Exchange of routing information between private and public routers
  - Class of Service: Included in Label encoding

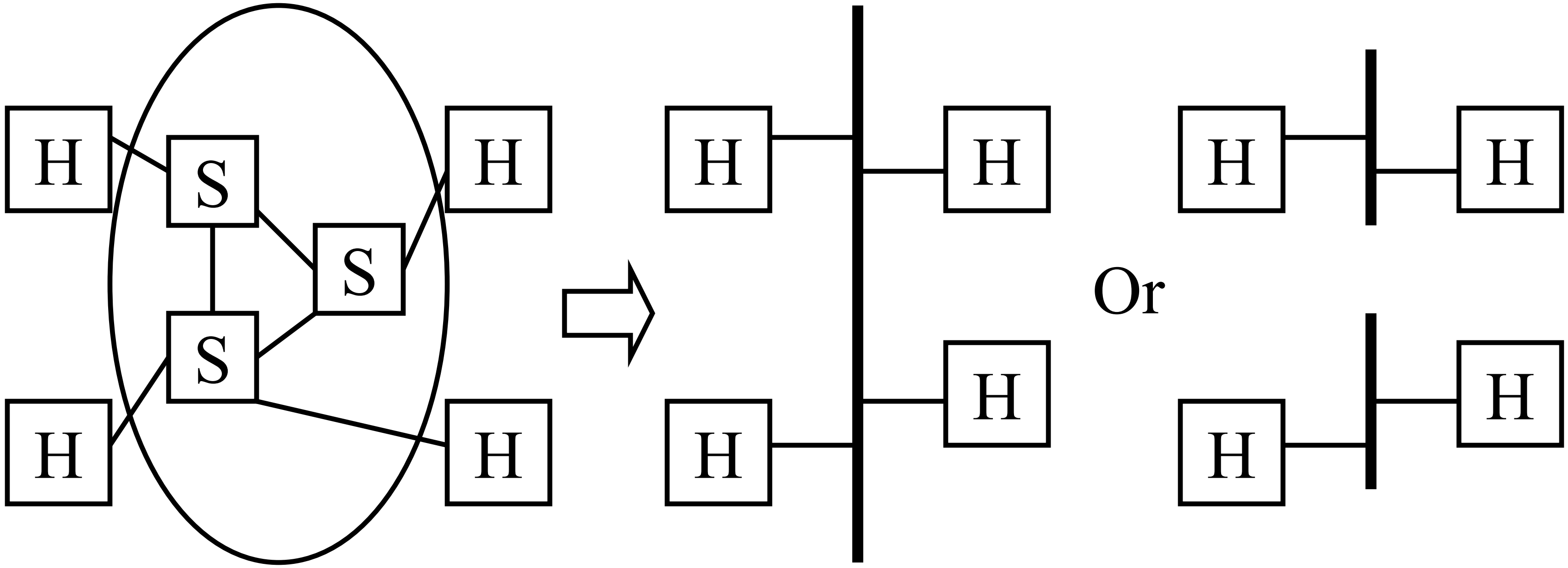


Raj Jain

CPE and Network Based VPNs

- CPE Based: CPE devices use real/virtual leased lines between them. Also known as overlay VPN.
- Network Based: Provider equipment takes the responsibility to provide secure communication for several customers on the same edge device
  - Cheaper since equipment shared
  - Issue: Exchange of routing information between customer and provider's routers
    1. Requires complex BGP routing [RFC2547]
    2. Alternative: Transparent LAN service

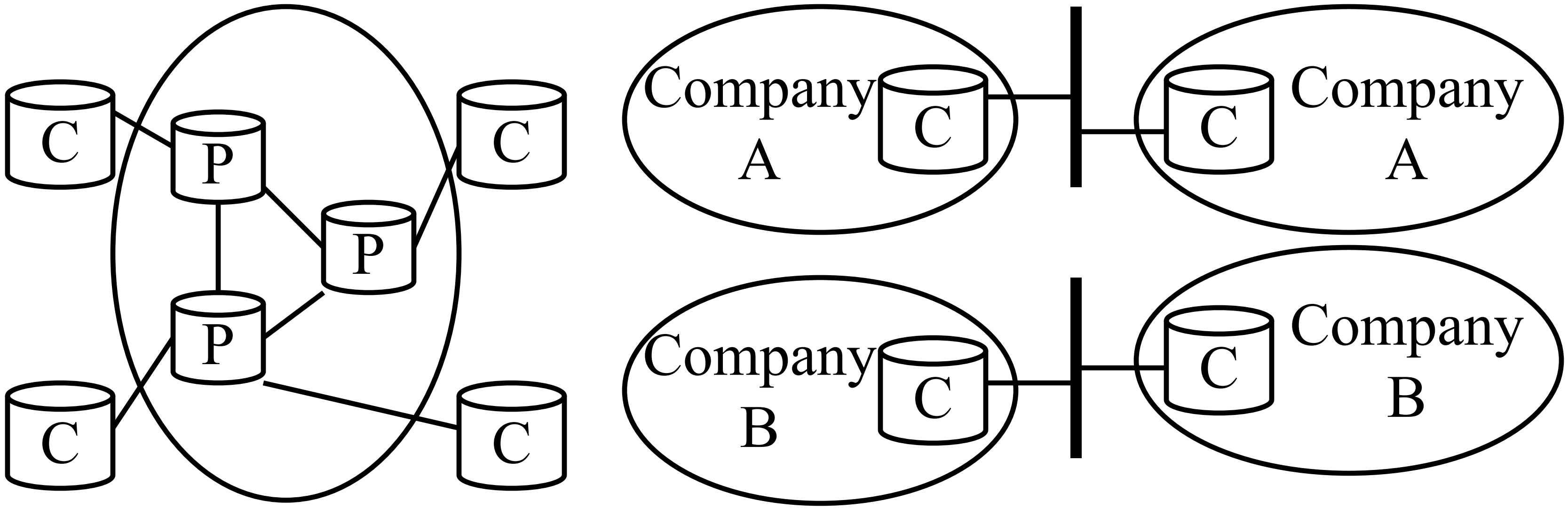
Transparent LAN Service



- ATM emulated LANs
- Hosts of ELAN driver and ATM card
- Destination hosts ATM address obtained by ARP
- Cells sent to the destination host and reassembled

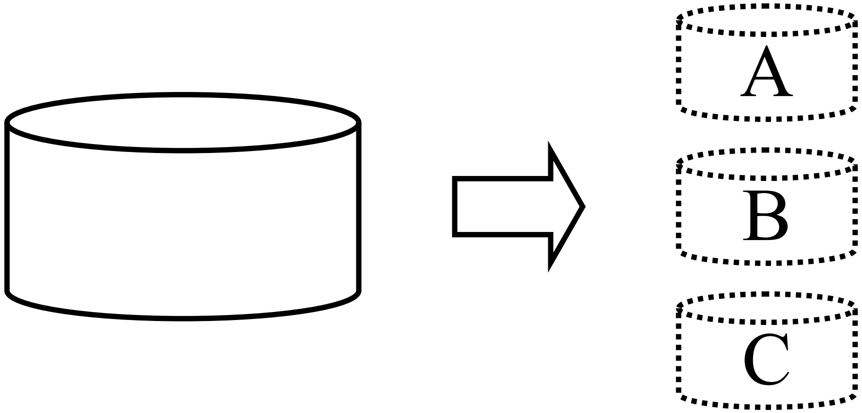
Raj Jain

Virtual Private LAN Segment



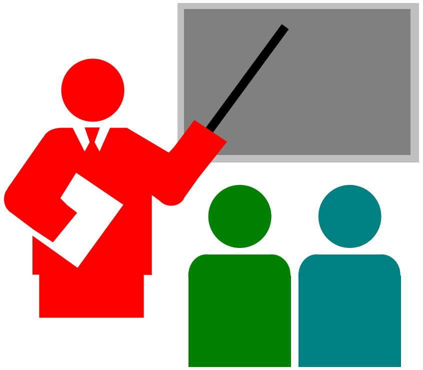
- Provider Edge Node provides link layer bridging
- IP multicast and broadcast is used to carry ARP
- Customer Switches/routers discover each other
- Scalable: many nodes per customer and many customers
- No routing info exchange between customer and provider
- Ref: RFC2764

Virtual Router



- Provider's edge device appears as several routers - each dedicated to a customer.
- Customers can configure, monitor, manage their VRs
- Many virtual routers may share physical resources but transparent to various customers. No leak of info.
- Provider specifies capacity, # of links, in each VR
- VRs use transparent LAN service to discover each other and setup MPLS tunnels [RFC 2917]

Summary



- VPN allows secure communication on the Internet
- Three types: WAN, Access, Extranet
- Key issues: address translation, security, performance
- Layer 2 (PPTP, L2TP), Layer 3 (IPSec), Layer 5 (SOCKS), Layer 7 (Application level) VPNs
- RADIUS allows centralized authentication server
- MPLS allows QoS and isolation

VPN Key: References

- For a detailed list of references, see http://www.cis.ohio-state.edu/~jain/refs/refs_vpn.htm
  Also reproduced at the end of this tutorial book
- VPN, http://www.cis.ohio-state.edu/~jain/cis788-99/h_7vpn.htm
- VPN Consortium, http://www.vpnc.org/
- VPN Insider, http://www.vpninsider.com
- VPDN.com, http://www.vpdn.com/
- Internet Week VPN Source Page, http://www.internetwk.com/VPN/default.html
- Network World VPN Resource page, http://www.nwfusion.com/netresources/vpn.html
- TeleChoice VPN Page, http://www.vpdn.com/home.asp

Final Review:25 Hot Facts

1. Networking is the key to productivity
2. Data traffic is exceeding voice traffic leading to carriers converting to data networks ⇒ voice over IP
3. Traffic growth is more than capacity leading to need for QoS and Traffic Engineering
4. Everything over Ethernet ⇒ LAN-WAN convergence
5. Integrated services allows three services: best effort, controlled load, and guaranteed service
6. RSVP allows signaling in IP networks

Final Review (Cont)

7. Both RSVP and Integrated services have overhead proportional to the number of flows. Not suitable for backbone.
8. Differentiated services will allow ISPs to monitor traffic and mark them with proper classes and drop precedences
9. Tag Switching allows packets to be switched based on tags (circuit numbers)
10. MPLS is an extension of tag switching and will allow interoperability, traffic engineering, and QoS.

Final Review (Cont)

11. Gigabit Ethernet comes in four varieties: SX, LX, and CX, T
12. Gigabit Ethernet supports both shared and full-duplex links. Most links are full-duplex
13. Ten-GbE will be not have a shared mode.
14. Ten-GbE will come in several varieties for various distances
15. Ten-GbE will run at two speeds: 10G and OC-192

Final Review (Cont)

16. High IP Routing speeds and volumes
    $\Rightarrow$ Need a full wavelength between routers
    $\Rightarrow$ Many ATM/SONET functions not needed
17. In Optical networks, data and control plane will be separated. Control Plane will be IP.
18. Overlay and Peer-to-peer modes for control. UNI for overlay mode.
19. Data will be circuit switched in the core
20. IP needs to be extended to provide addressing, signaling, routing, and protection for lightpaths

Final Review (Cont)

21. All-optical circuit/packet switching for Tbps and up
22. GMPLS allows SONET, Wavelength switching along with packet switching
23. H.323 is the conferencing standard designed for LANs and best effort networks.
24. Gatekeepers provide bandwidth management while Gateway provide protocol translation.
25. VPNs allow private networks over public Internet