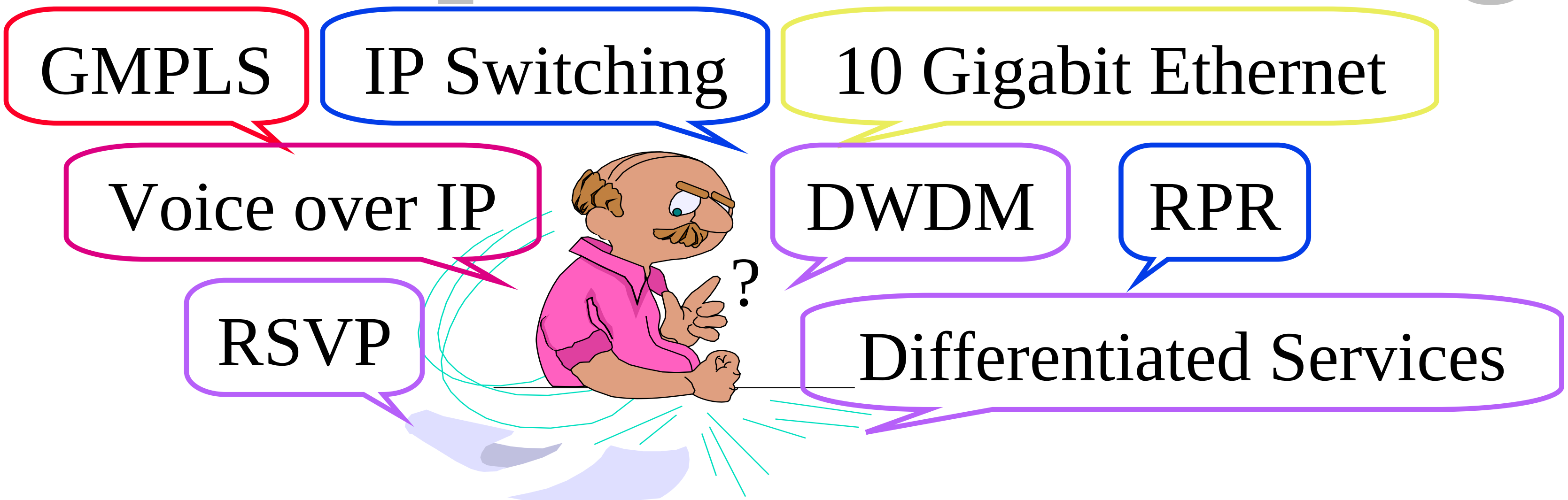
Hot Topics in Networking



Raj Jain

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

www.n ... lu/~jain/

1. Networking trends
2. QoS over data networks
3. Label switching
4. Gigabit, 10 Gb Ethernet, RPR
5. Storage area networks
6. IP over DWDM
7. Wireless
8. Voice over IP

1. Networking Trends

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

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
- Multi-Protocol Label Switching
- Label Stacks
- Label Distribution Protocols: LDP, CR-LDP, RSVP-TE
- Traffic Engineering using MPLS
- Draft-Martini

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. Storage Area Networks

- Five Trends in Storage
- What is SAN?
- SAN vs NAS
- Fibre Channel, ESCON
- SAN Devices
- IP Storage: iSCSI, iFCP, FCIP, iSNS

6. 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

7. Wireless Data Networks

- Spread Spectrum, Frequency Hopping, Direct-Sequence, OFDM
- IEEE 802.11, 11b, 11a, 11g LANs
- HiperLAN, HiperLAN2
- PANs: IrDA, Bluetooth, HomeRF
- WAP, WML
- Note: wireless phone services and standards not covered.

8. 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)

Day 1 (Tentative)

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

Day 2

8:30 - 9:30	Storage Area Networks
9:30 -10:15	IP Over DWDM I
10:15 -10:30	*Coffee Break*
10:30 -11:30	IP Over DWDM II
11:30 -12:00	Wireless Data Networks I
12:00 - 1:00	*Lunch Break*
1:00 - 2:30	Wireless Data Networks II
2:30 - 2:45	*Coffee Break*
2:45 - 4:30	Voice over IP

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
- Guaranteed quality and controlled load services
- Integrated vs Differentiated Services
- Expedited forwarding vs Assured Forwarding
- Min packet sizes on 10Base-T and 1000Base-T
- Token Ring and Resilient Packet Ring
- Ring wrapping vs Steering
- iSCSI and iSNS
- FCIP and iFCP

Pre-Test (Cont)

- ❑ MPλS and GMPLS
- ❑ 802.11a and 802.11b
- ❑ OFDM and CDMA
- ❑ Bluetooth and HomeRF
- ❑ H.323 and Session Initiation Protocol
- ❑ Gatekeeper and Gateway
- ❑ Media Gateway and Signaling Gateway

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

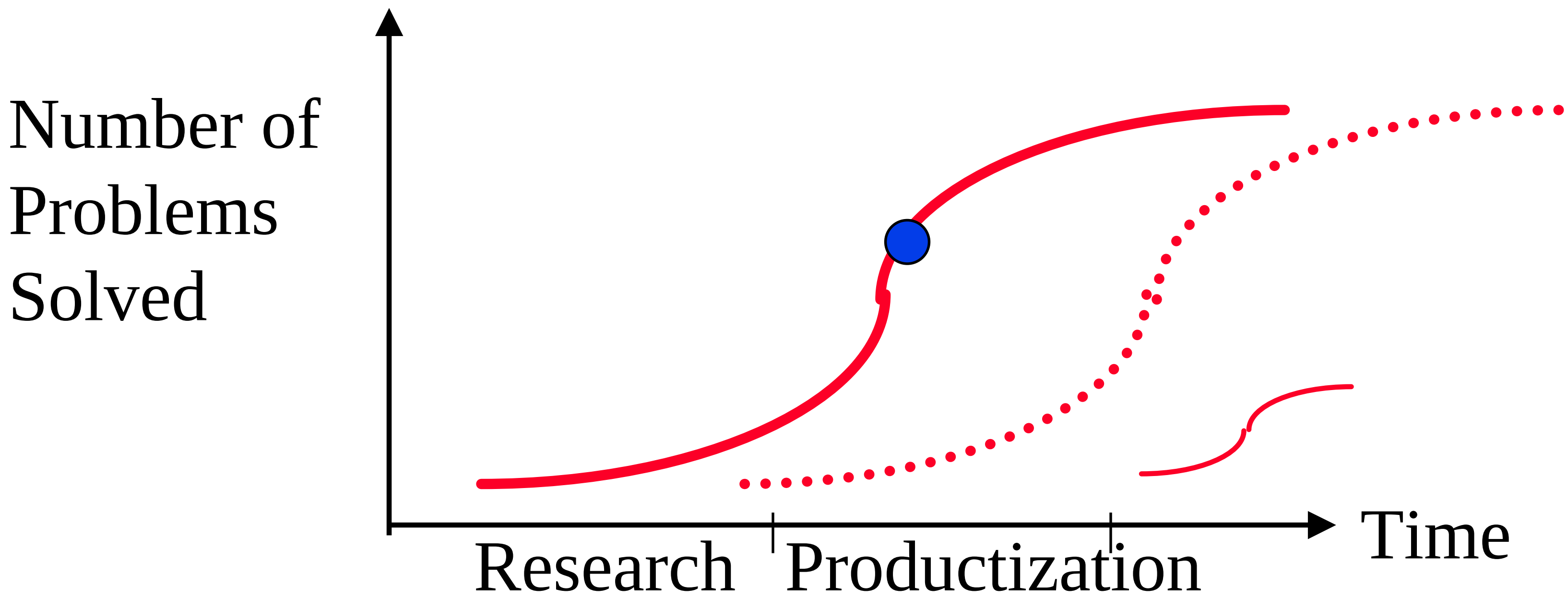
- The technologies are currently evolving.
  $\Rightarrow$ Many statements are subject to change.
- Features not in a technology may be implemented later in that technology.
- Problems claimed to be in a technology may later not be a problem.

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/

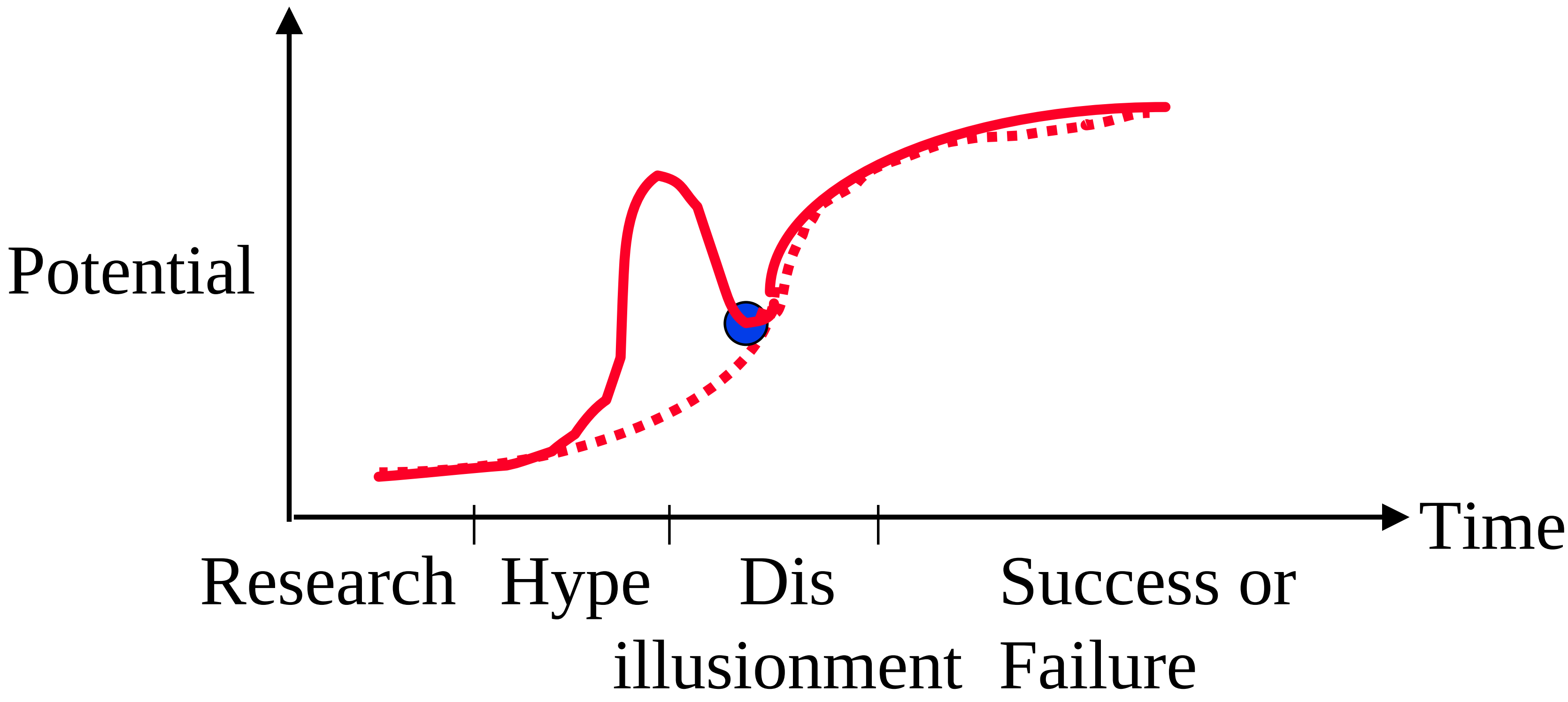
- Life Cycles of Technologies
- Traffic vs Capacity Growth
- Trend: Ethernet Everywhere
- Technology Failures vs Successes
- Trend: LAN - WAN Convergence
- Ethernet vs SONET
- Trend: Everything over IP

Life Cycles of Technologies



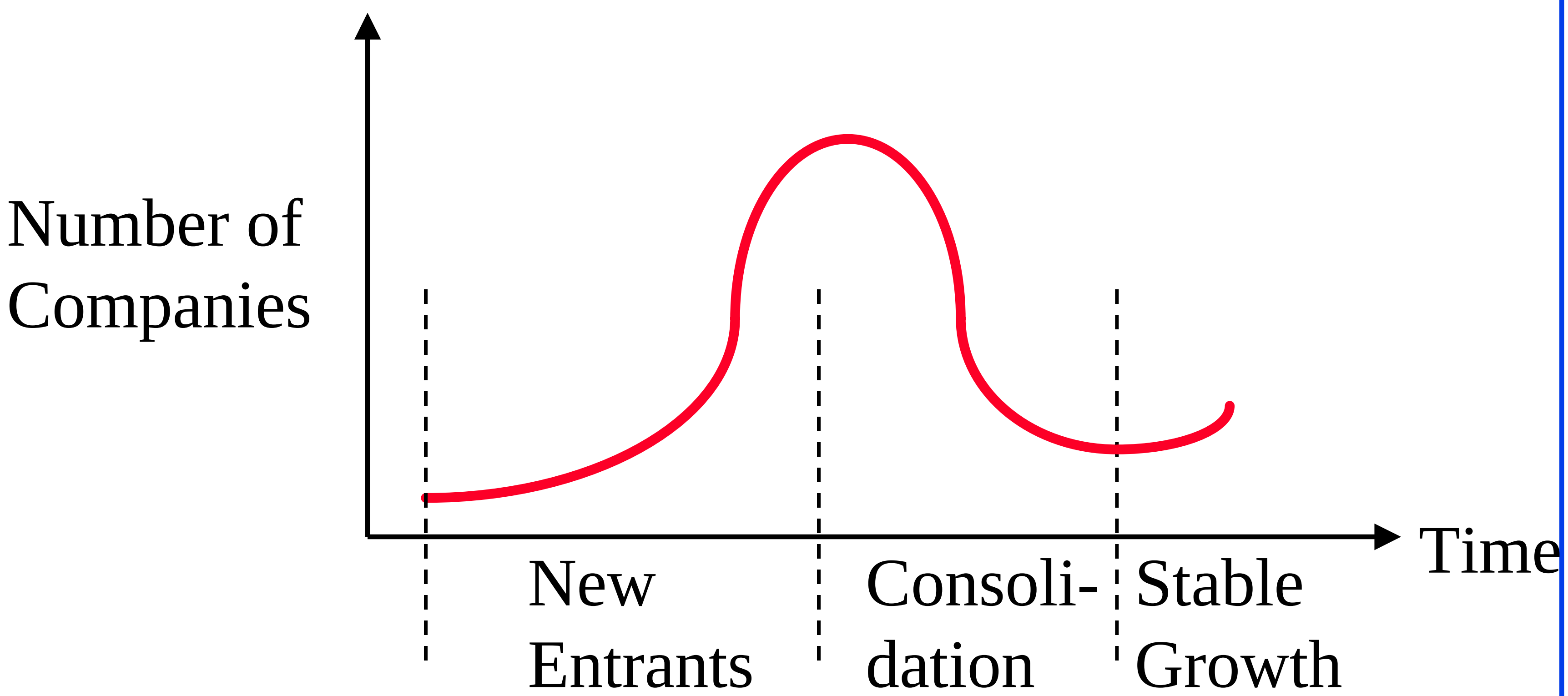
Raj Jain

Hype Cycles of Technologies



Raj Jain

Industry Growth



Raj Jain

Traffic vs Capacity Growth

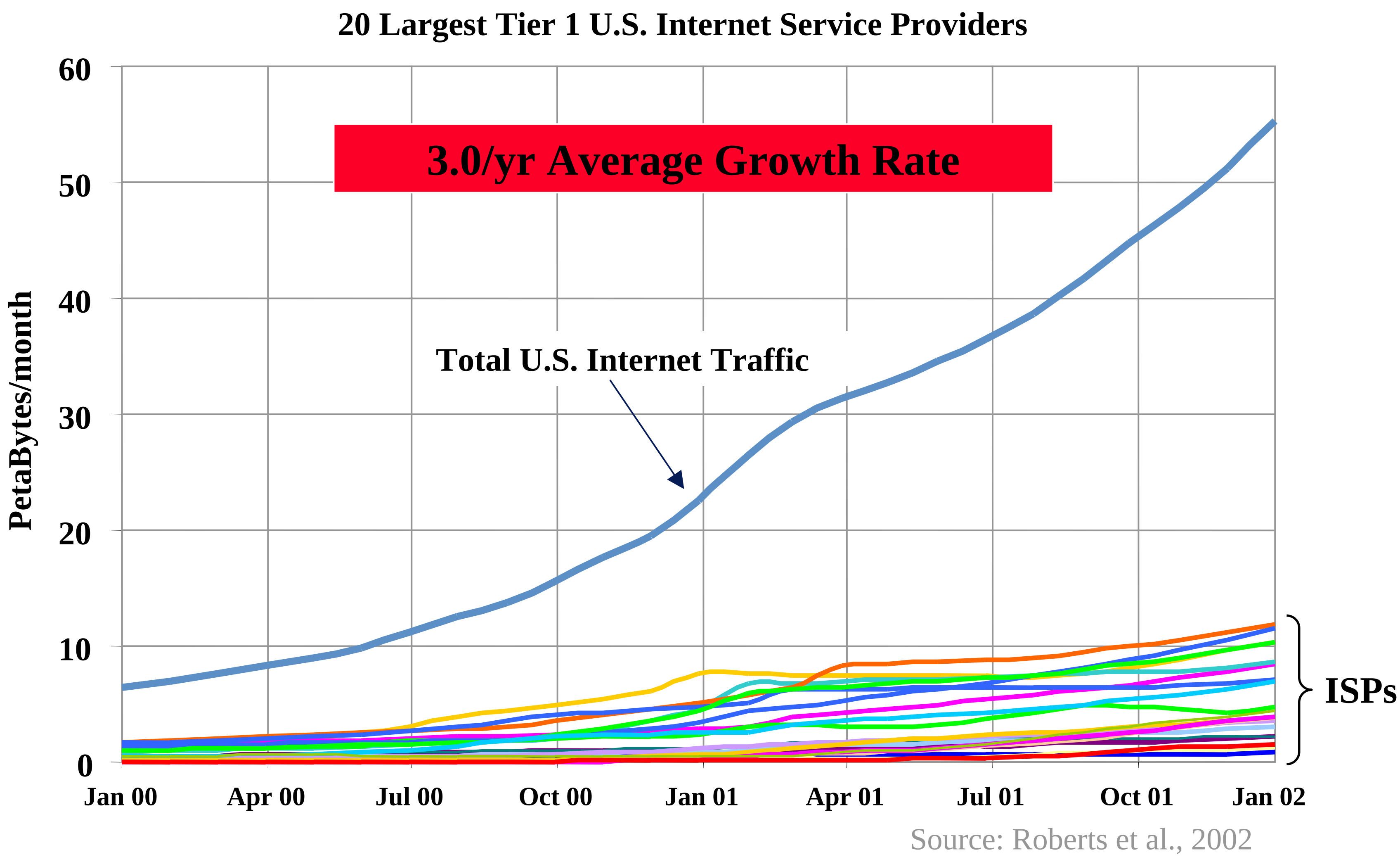
Expensive Bandwidth	Cheap Bandwidth
❑ Sharing	❑ No sharing
❑ Multicast	❑ Unicast
❑ Virtual Private Networks	❑ Private Networks
❑ Need QoS	❑ QoS less of an issue
❑ Likely in WANs	❑ Possible in LANs

Raj Jain

Is Internet Traffic 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
- Carriers are using only *avg* 2.7% of their total *lit* fiber capacity - Michael Ching, Marril Lynch & Co. in Wall Street Journal
- 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

Total U.S. Internet Traffic



Raj Jain

Trend: Ethernet Everywhere

- Ethernet in Enterprise Backbone
  - Ethernet vs ATM (Past)
- Ethernet in Metro: Ethernet vs SONET
  - 10 G Ethernet
  - Survivability, Restoration $\Rightarrow$ Ring Topology
- Ethernet in Access: EFM
- Ethernet in homes: Power over Ethernet

Networking: Failures vs Successes

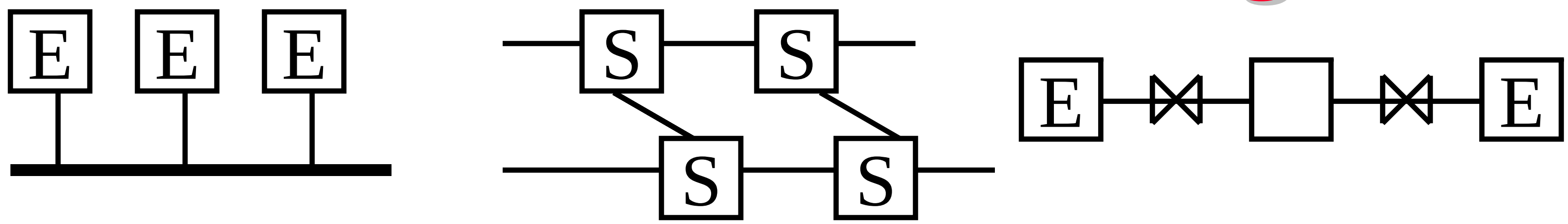
- 1980: Broadband (vs baseband)
- 1984: ISDN (vs Modems)
- 1986: MAP/TOP (vs Ethernet)
- 1988: OSI (vs TCP/IP)
- 1991: DQDB
- 1994: CMIP (vs SNMP)
- 1995: FDDI (vs Ethernet)
- 1996: 100BASE-VG or AnyLan (vs Ethernet)
- 1997: ATM to Desktop (vs Ethernet)
- 1998: Integrated Services (vs MPLS)
- 1999: Token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

Requirements for Success

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

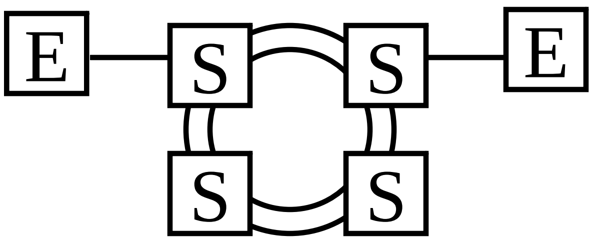
Raj Jain

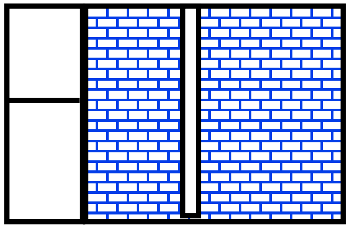
Trend: LAN - WAN Convergence



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

SONET Functions





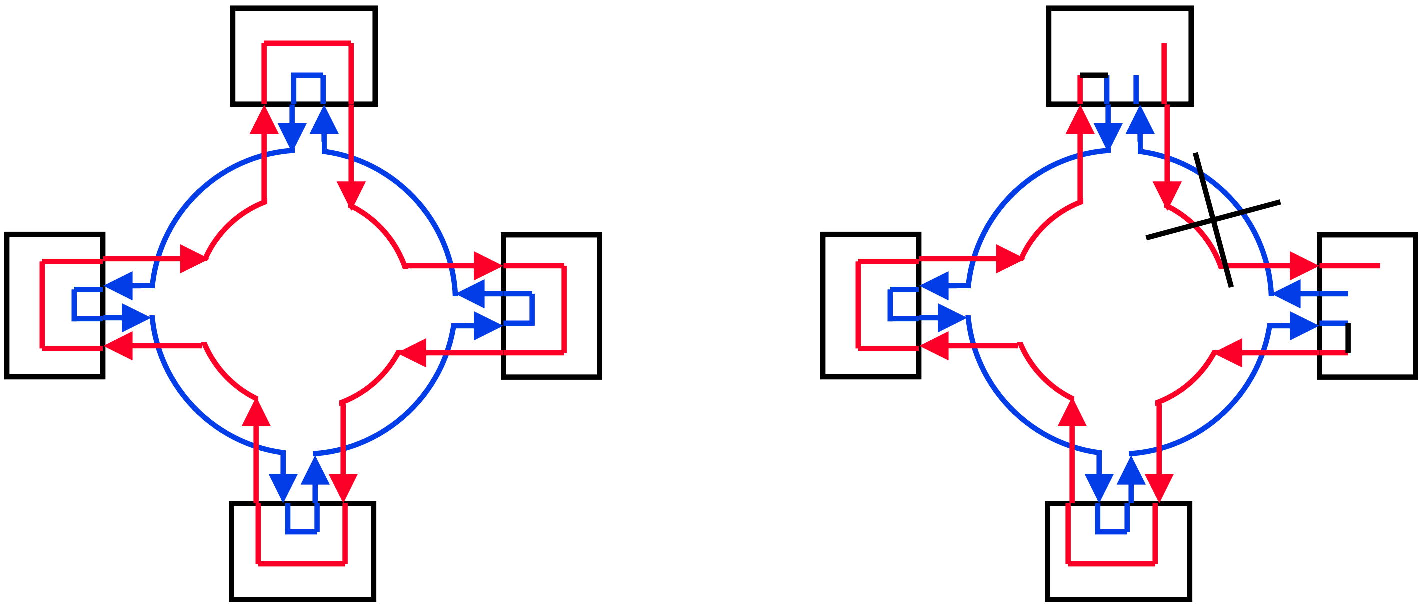
- Protection: Allows redundant Line or paths
- Fast Restoration: 50ms using rings
- Sophisticated OAM&P
- Ideal for Voice: No queues. Guaranteed delay
- Fixed Payload Rates: 51M, 155M, 622M, 2.4G, 9.5G
  Rates do not match data rates of 10M, 100M, 1G, 10G
- Static rates not suitable for bursty traffic
- One Payload per Stream
- High Cost

Raj Jain

SONET: 2001 Developments

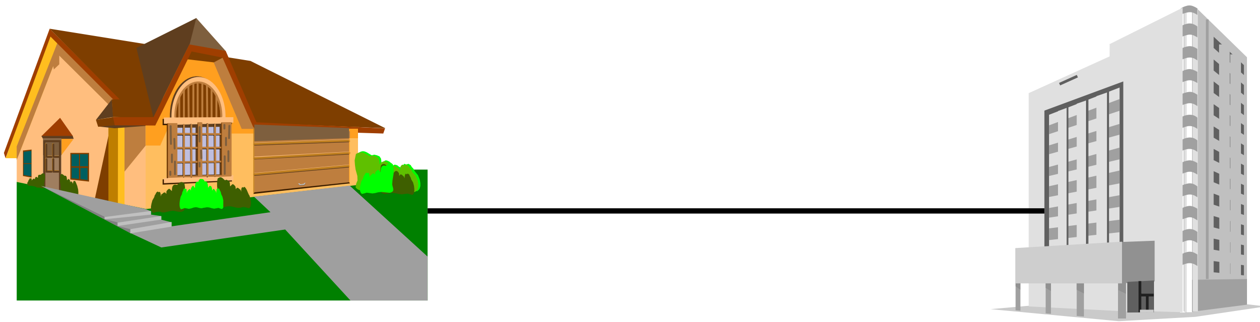
- Fixed Payload Rates: 51M, 155M, 622M, 2.4G, 9.5G
  Virtual concatenation allows any multiple of T1/STS1
  10M = 7 T1, 100M=2 STS-1, 1G=7 STS-3c's
- Static rates not suitable for bursty traffic
  Link Capacity Adjustment Scheme (LCAS) allows dynamic adjustment of number of T1's or STS's
- One Payload per Stream
  Generic Framing Protocol (GFP) allows multiple payloads per stream
- High Cost
  ASICs are being developed to reduce cost

Resilient Packet Rings



- Dual Counter-rotating rings help protect against failure
- Allows TDM traffic like T1, T3, SONET over RPR
- Will Ethernet with RPR be cheaper than SONET?

Ethernet in the First Mile



- IEEE 802.3 Study Group started November 2000
- Originally called Ethernet in the Last Mile
- Current Technologies: ISDN, xDSL, Cable Modem, Satellite, Wireless, PON
- EFM Goals: Media: Phone wire, Fiber, Air
  - Speed: 125 kbps to 1 Gbps
  - Distance: 1500 ft, 18000 ft, 1 km - 40 km
- Ref: http://www.ieee802.org/3/efm/public/index.htm

Raj Jain

Power over Ethernet

- IEEE 802.3af group approved 30 January 2000 Power over 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
- Standard Expected: November 2002
- Ref: http://grouper.ieee.org/groups/802/3/power_study/public/nov99/802.3af_PAR.pdf

Recent Networking Trends

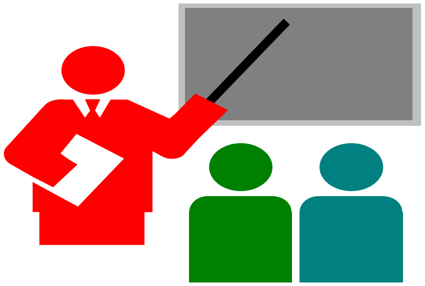
- Hottest Technologies: Storage, IP, Ethernet, Wireless, Optical
- Hottest Applications: Peer-to-peer (no money to be made by carriers), Storage, VOIP
- Changing Traffic Mix: 80/20 to 20/80 LAN:WAN Ratio (because of IP addressing and distance independent billing)
- Enterprise Market > Access > Metro > Core
- Financial Markets: No CLECs
- Advances in Optical technologies: 40G, Long Haul, More wavelengths

Networking Trends (Cont)

- Glut of Fiber in long haul but shortage in Metro/Access
- Emergence of Ethernet Metro
- Bandwidth prices are dropping (in the long haul) 2c/min
- Emphasis on Security
- Emphasis on Mobility

Trend: Everything over IP

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



Summary

- Traffic > Capacity
  $\Rightarrow$ Need QoS, traffic engineering in WANs
- Ethernet everywhere
  $\Rightarrow$ Rings, many rates, longer distances, Power
- SONET is also adapting to data traffic
  $\Rightarrow$ SONET will stay longer than expected.
- Convergence at L3 $\Rightarrow$ Everything over IP
  $\Rightarrow$ IP needs circuits, traffic engineering, data and control plane separation

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.

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/

Overview

- 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

ATM Classes of Service

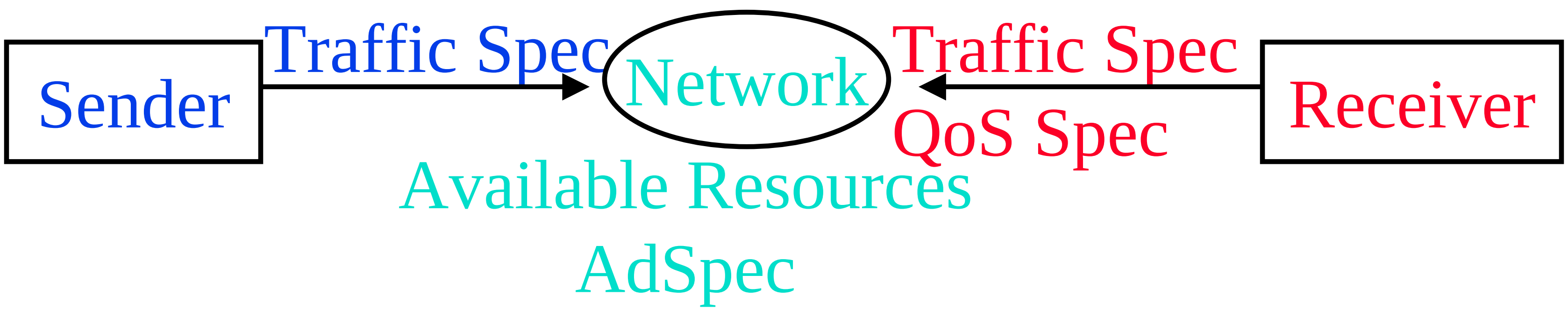
- ABR (Available bit rate): Source follows feedback. Max throughput with minimum loss.
- UBR (Unspecified bit rate):
  User sends whenever it wants. No feedback. No guarantee. Cells may be dropped during congestion.
- CBR (Constant bit rate): User declares required rate. Throughput, delay and delay variation 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

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

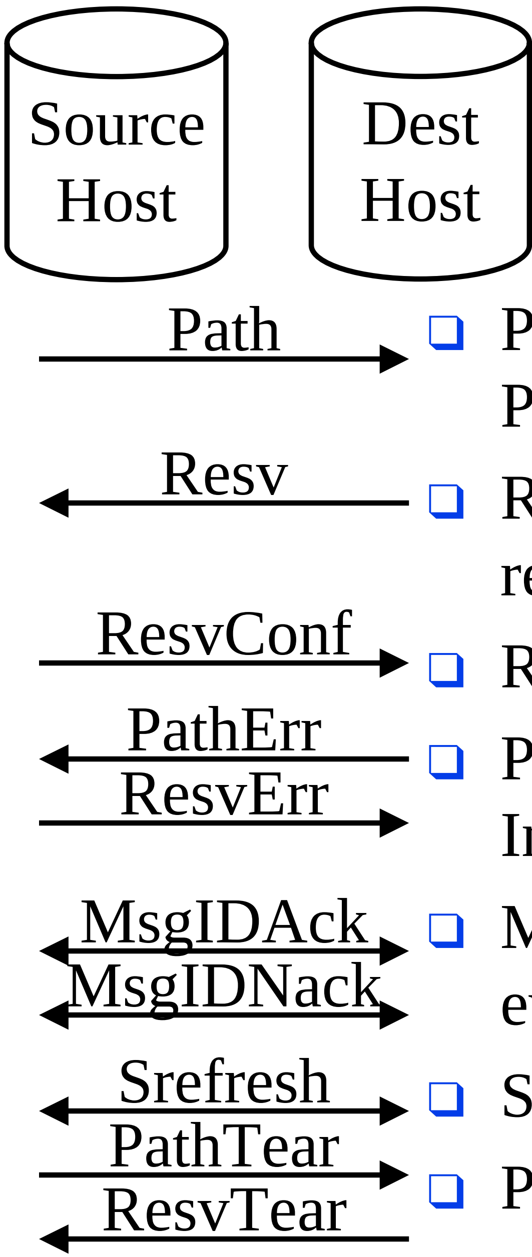
RSVP

- 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 refereshed.
- Resv: Resv state is created. Regularly refereshed
- 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) $\Rightarrow$ Not scalable with # of flows.
  Number of flows in the backbone may be large.
  $\Rightarrow$ 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.
  $\Rightarrow$ 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

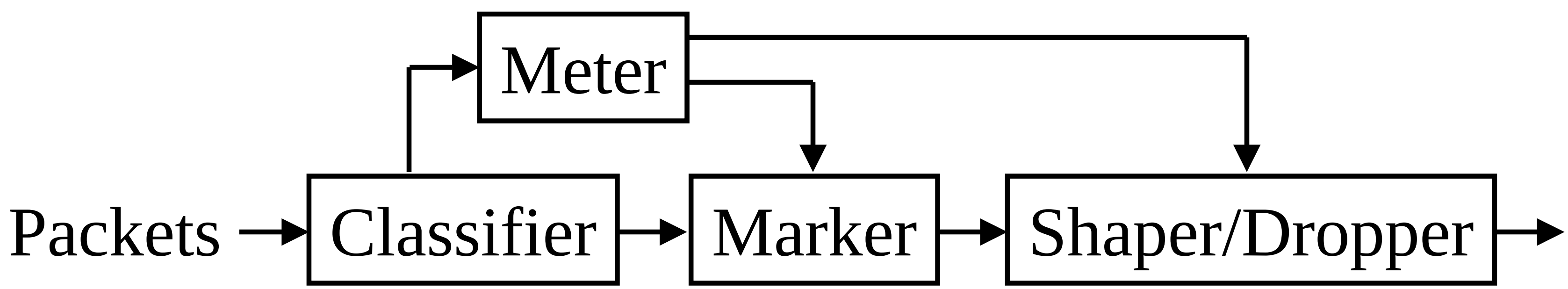
Differentiated Services

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

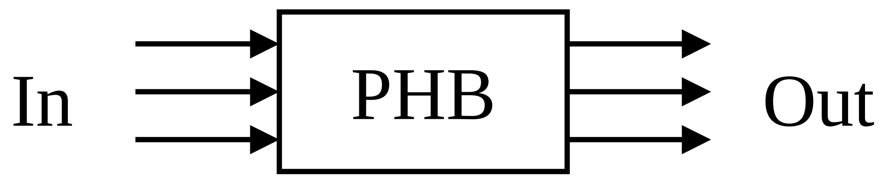
- Micro-flow = A single application-to-application flow
- Traffic Conditioners: Meters (token bucket), Markers (tag), Shapers (delay), Droppers (drop)
- Behavior Aggregate (BA) Classifier:
  Based on DS byte only
- Multi-field (MF) Classifiers:
  Based on IP addresses, ports, DS-byte, etc..



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

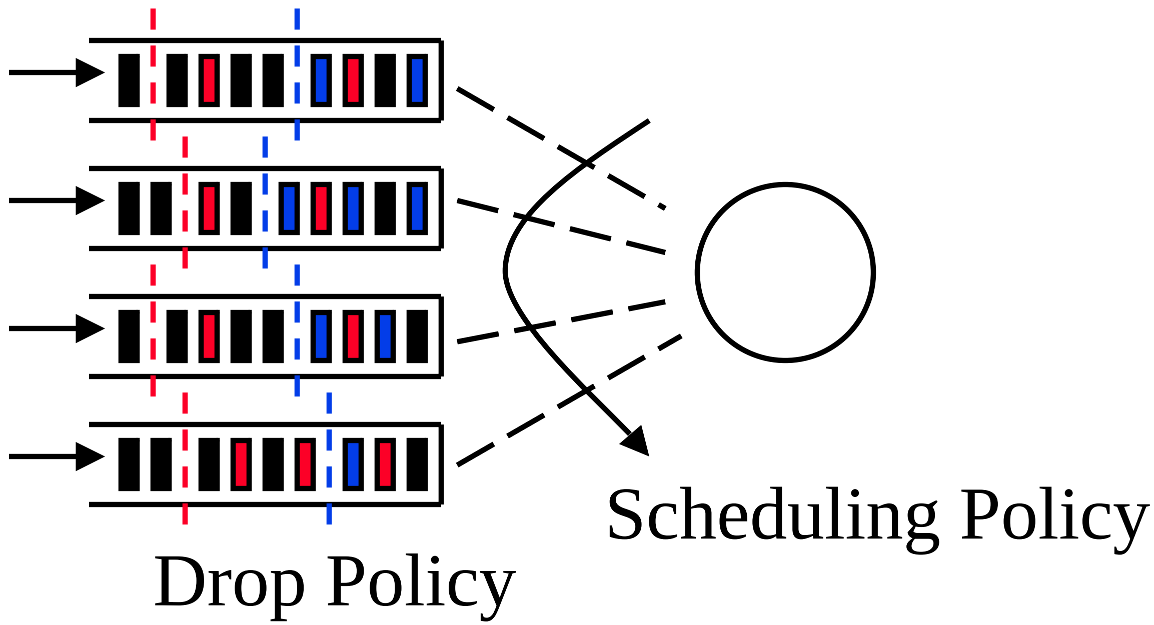


- Externally Observable Forwarding Behavior
- x% of link bandwidth
- Minimum x% and fair share of excess bandwidth
- Priority relative to other PHBs
- PHB Groups: Related PHBs. PHBs in the group share common constraints, e.g., loss priority, relative delay

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
  $\Rightarrow$ Highest data priority (if priority queueing)
- Code point: 101 110

Assured Forwarding



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

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

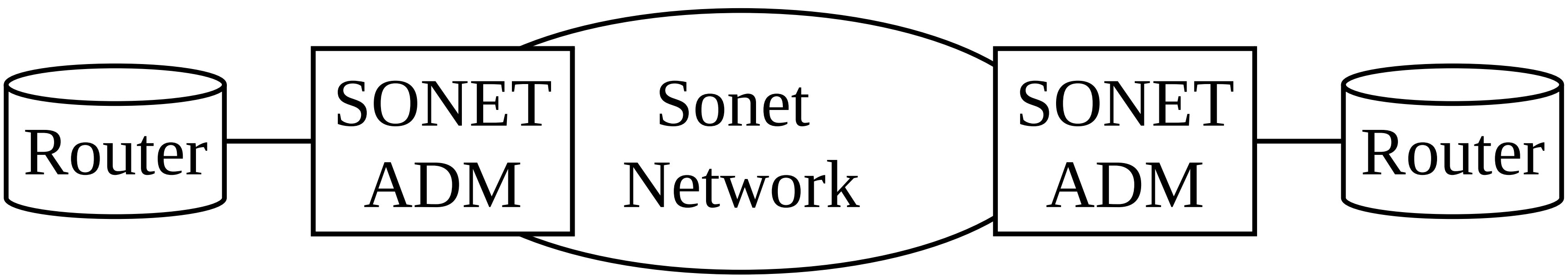
- Avoids xxx000 class selectors. Last bit 0 ⇒ Standard

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?

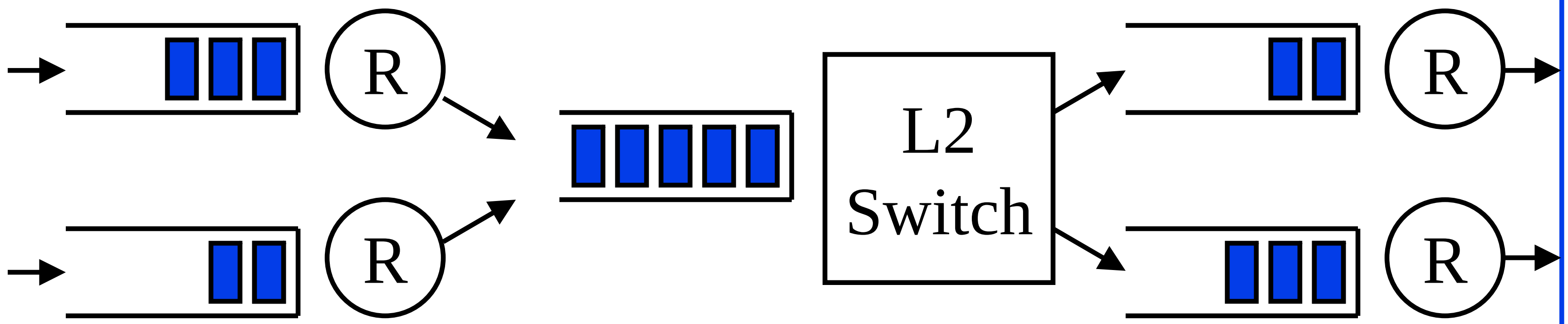


Raj Jain

DiffServ Problems (Cont)

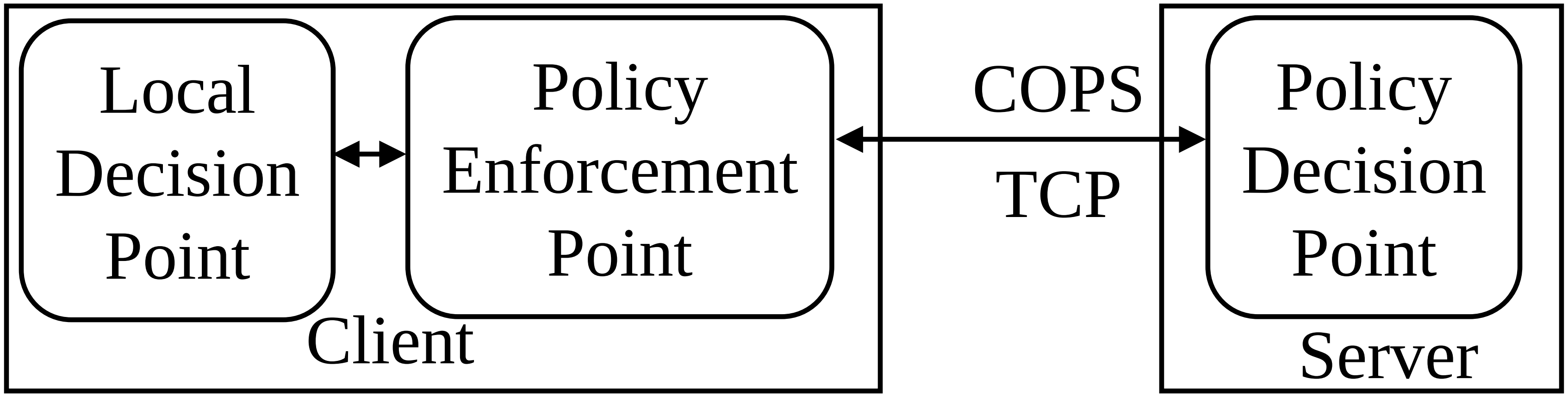
- QoS is for the aggregate not micro-flows.
  - Large number of short flows are better handled by aggregates.
  - High-bandwidth flows (1 Mbps video) need per-flow guarantees.
- Designed for static Service Level Agreements (SLAs) Both the network topology and traffic are highly dynamic.
- Need route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

Subnet Bandwidth Manager (SBM)



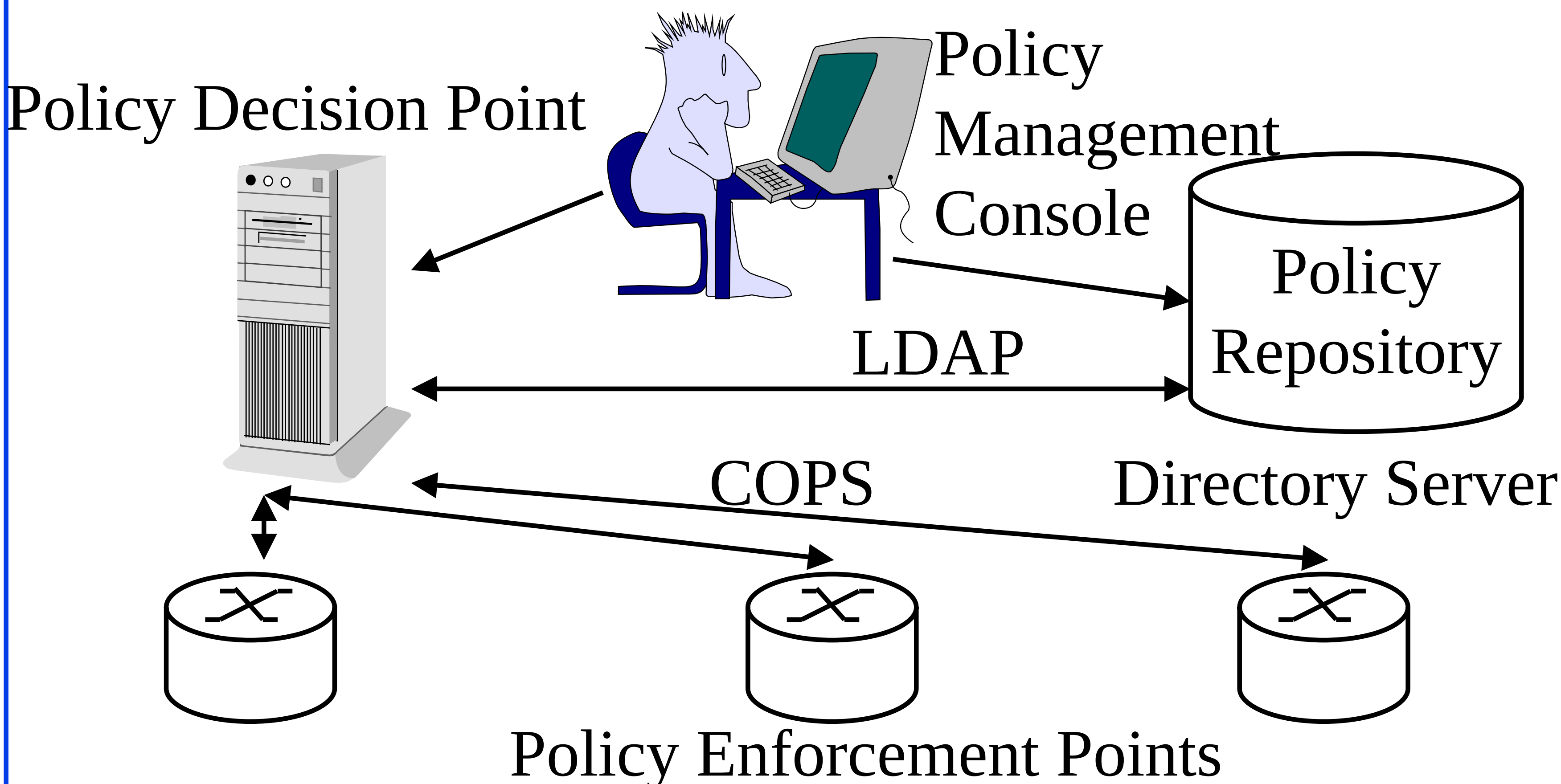
- Resources in a L2 switches may be a bottleneck
- SBM allows L2 switches to participate in RSVP admission control
- SBM capable switches and hosts elect a Designated SBM (DSBM)
- All RSVP messages are sent through DSBM
- Ref: RFC 2814, May 2000.

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

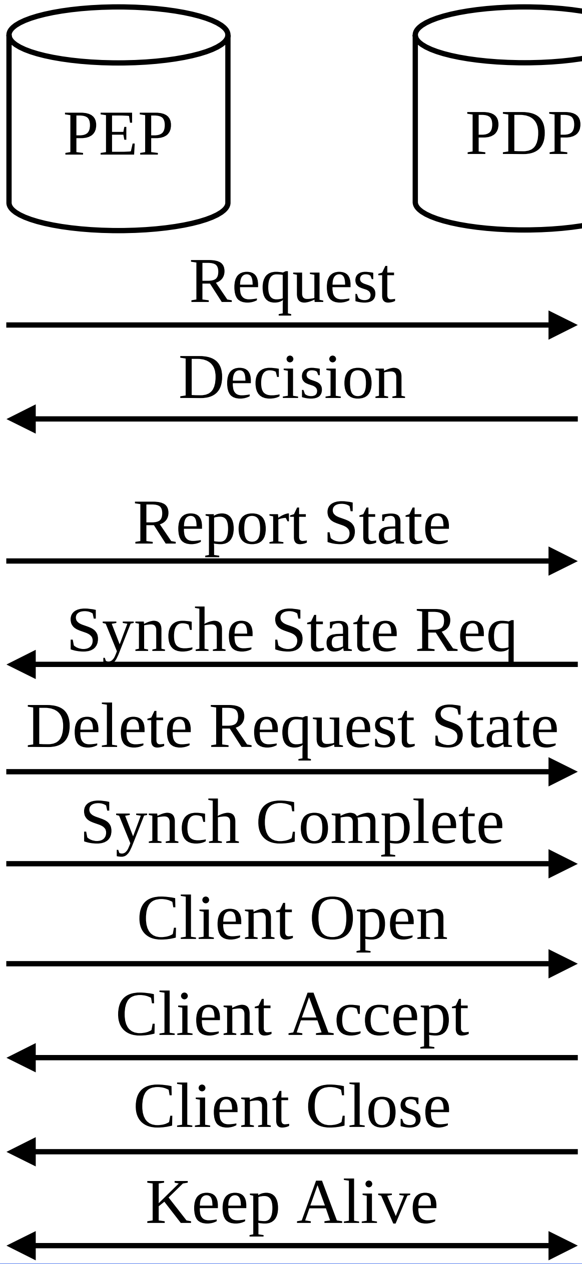
Policy Framework



- 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

IEEE 802.1D Model

Dest Addr	Src Addr	Tag Prot ID	Pri	CFI	VLAN ID

|←——— 802.1Q header ———→|

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

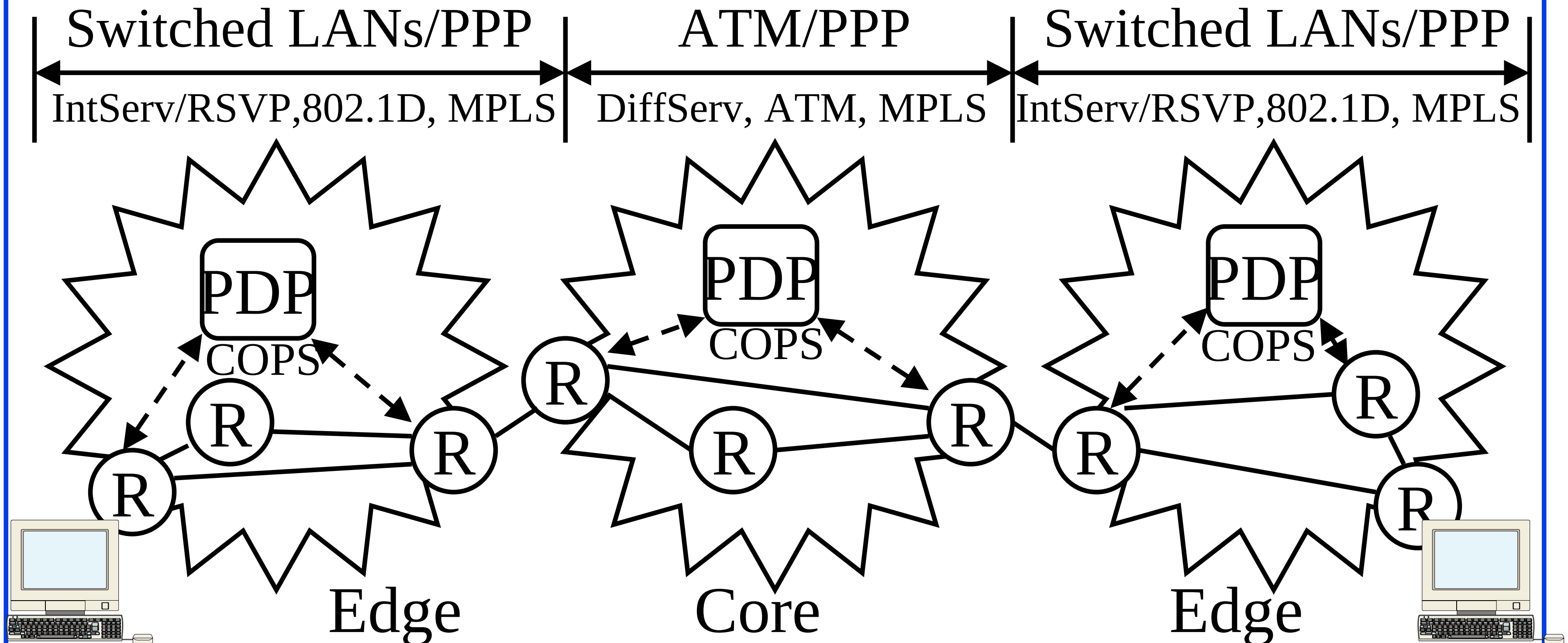
  5 Video (Less than 100 ms latency and jitter)

  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



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.

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

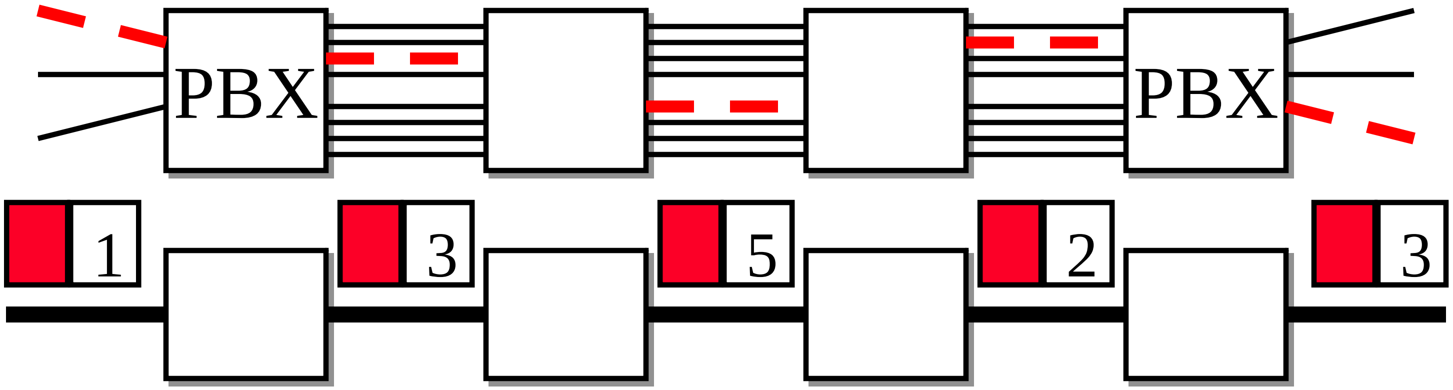
Multiprotocol Label Switching

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
- Multi-Protocol Label Switching
- Label Stacks
- Label Distribution Protocols: LDP, CR-LDP, RSVP-TE
- Traffic Engineering using MPLS
- Draft-Martini

Raj Jain

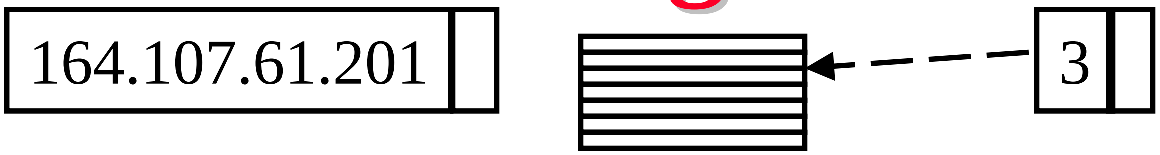
Multiprotocol Label Switching (MPLS)



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

Raj Jain

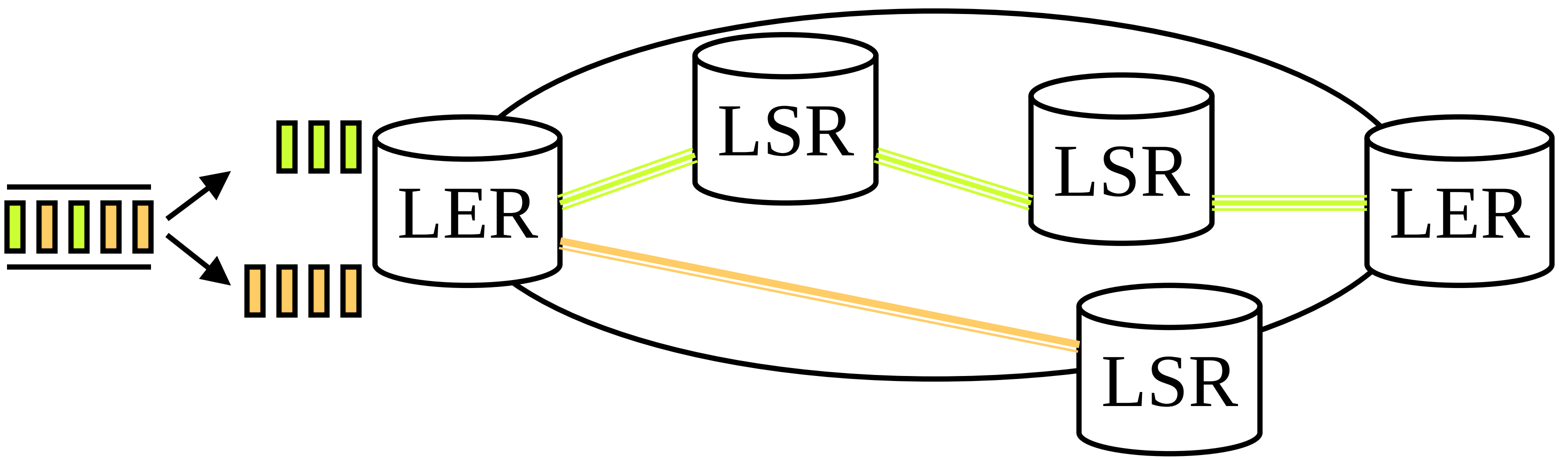
Routing vs Switching



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

MPLS Terminology

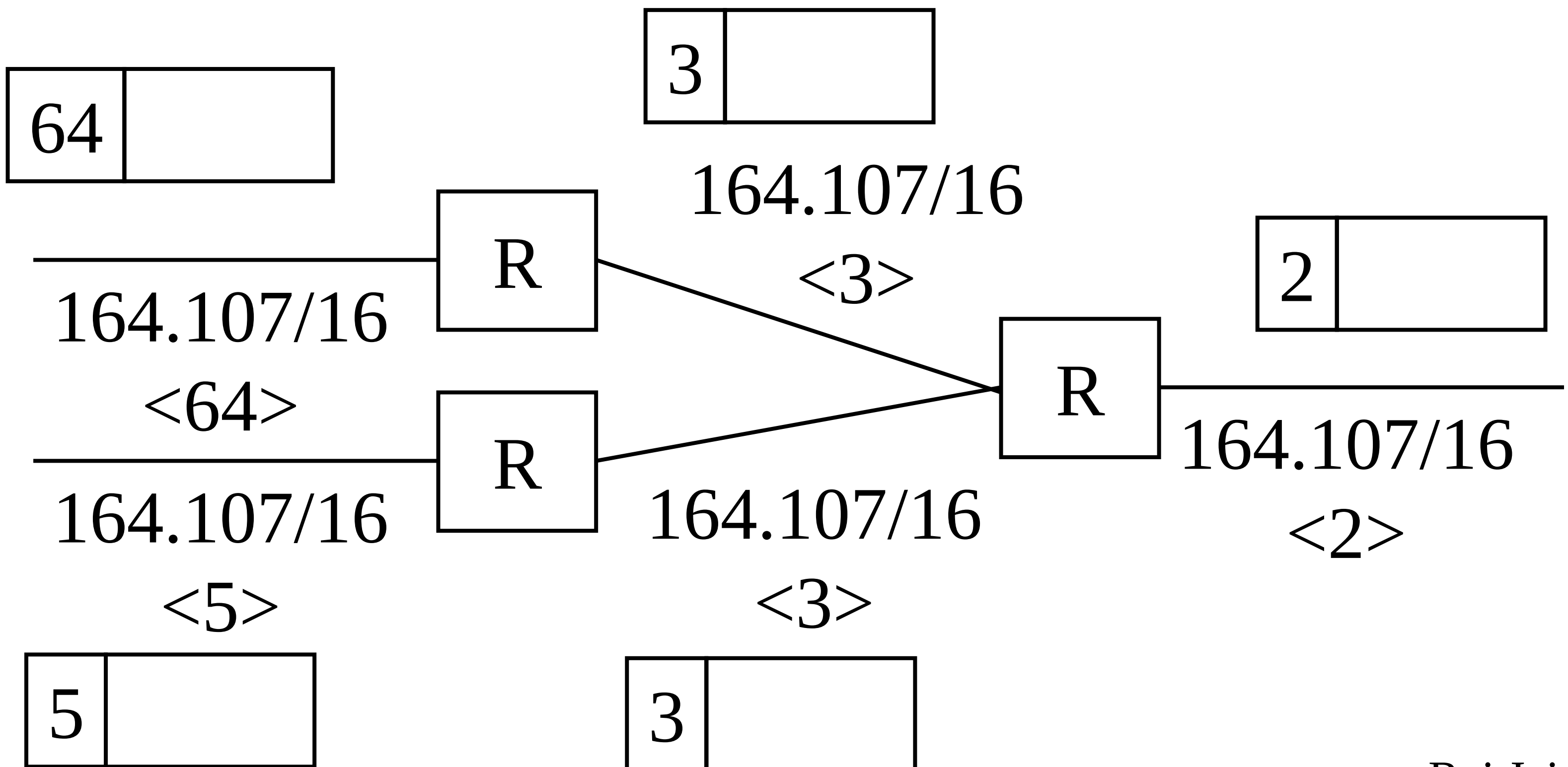
- Label Edge Router (LER)
- Label Switching Router (LSR)
- Label Switched Path (LSP)
- Forwarding Equivalence Class (FEC)



Raj Jain

Label Switching Example

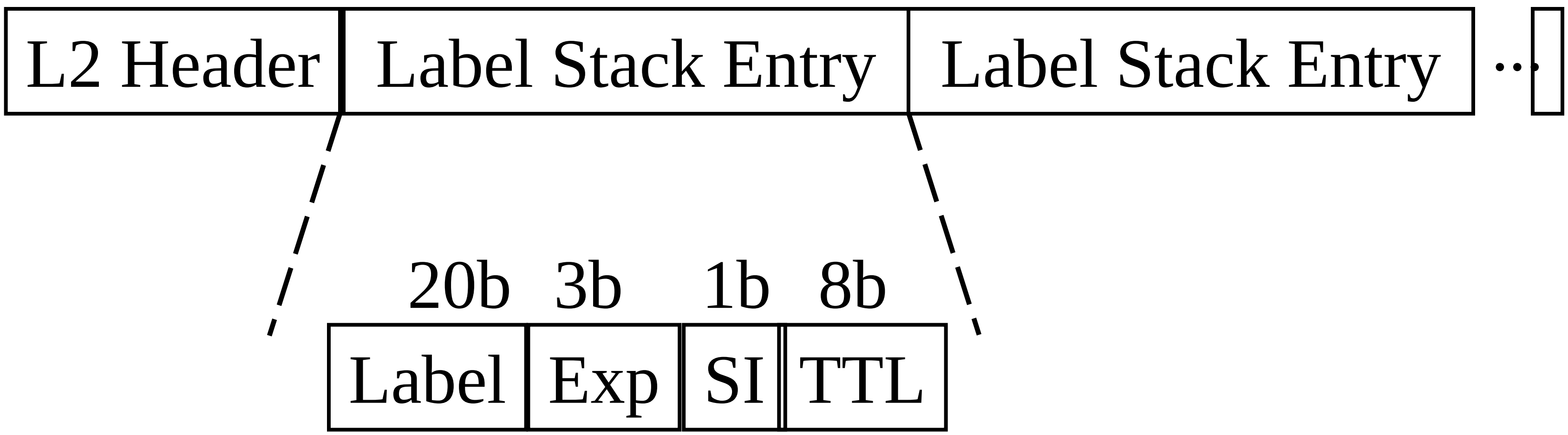
- One VC per routing table entry



Raj Jain

Label Stack Entry Format

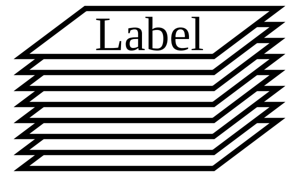
- Labels = Explicit or implicit L2 header
- TTL = Time to live
- Exp = Experimental
- SI = Stack indicator



Raj Jain

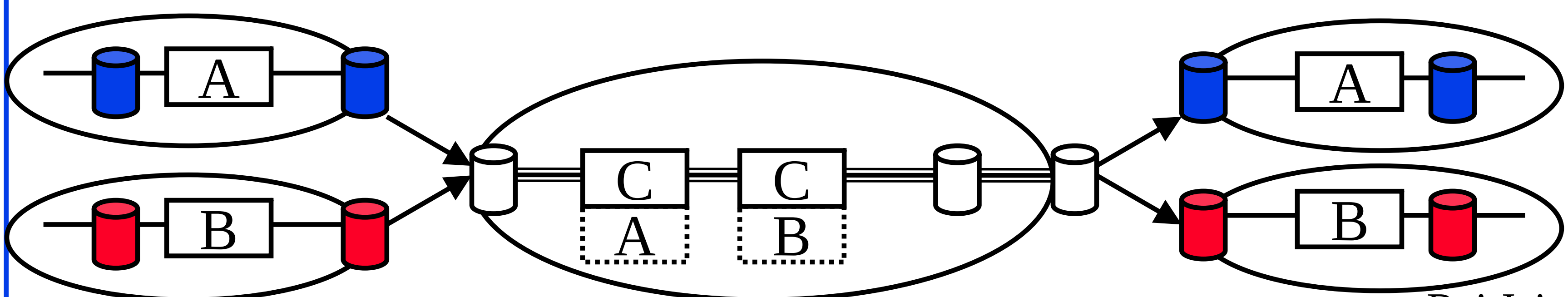
Label Stacks

- 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)

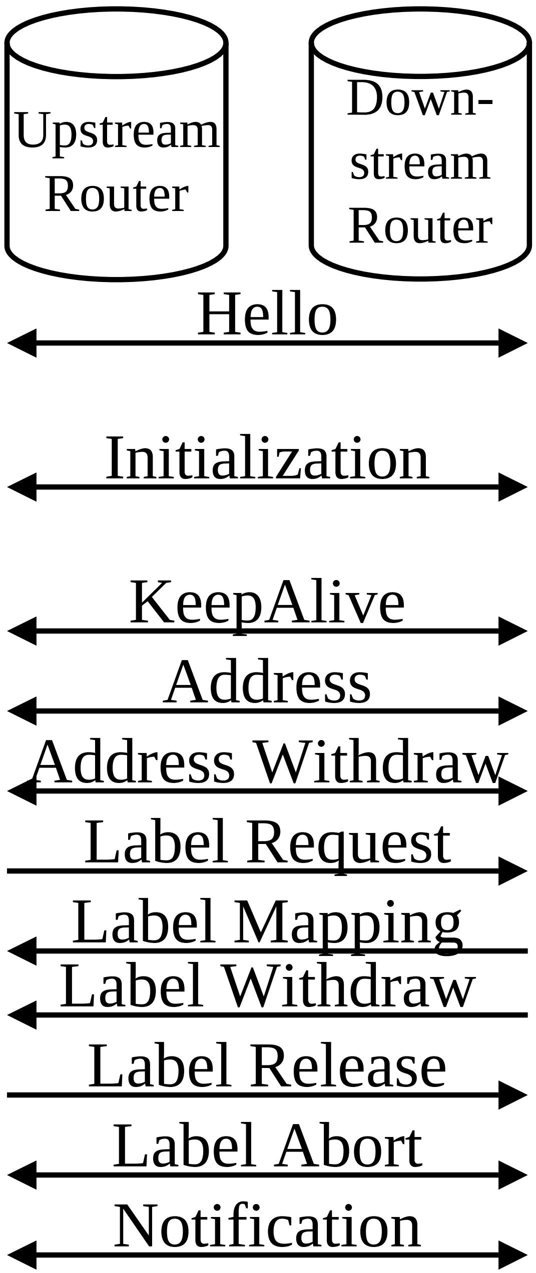
L2 Header	LSE 1	LSE 2	...	LSE n	



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
- Label Distribution Protocol called LDP
- RSVP has been 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

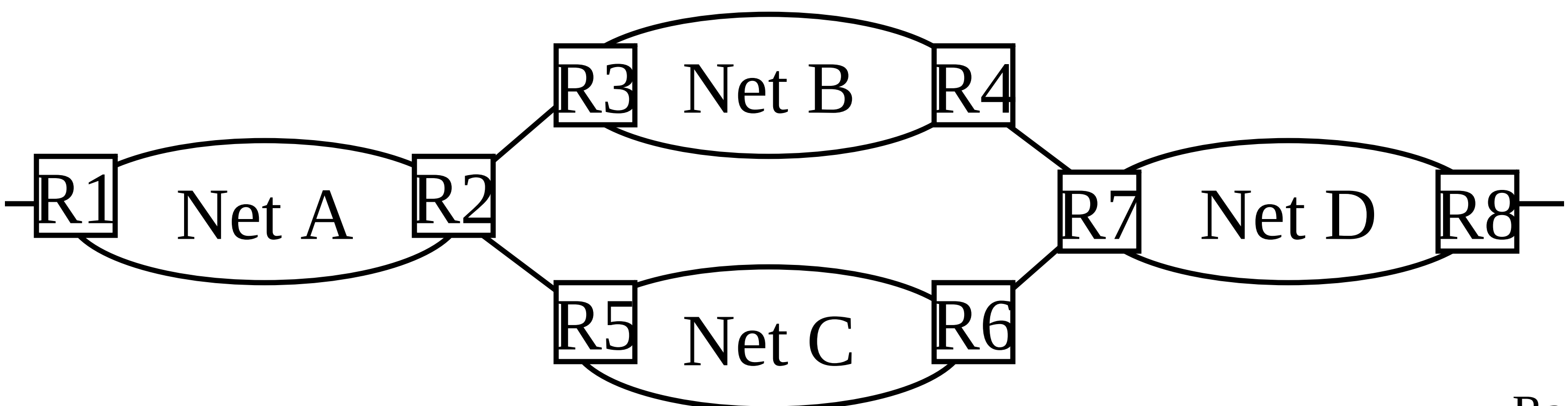
- Extension of LDP for constraint-based routing (CR)
- New Features:
  - Traffic parameters
  - Explicit Routing with Egress Label
  - Preemption of existing route. Based on holding priorities and setup priorities
  - Route pinning: To prevent path changes
  - Label Set: Allows label constraints (wavelengths)
- 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
- Quick Fault Notify: Notify msg direct to initiator (and terminator if bidirectional). Multi failures in one msg.
- 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

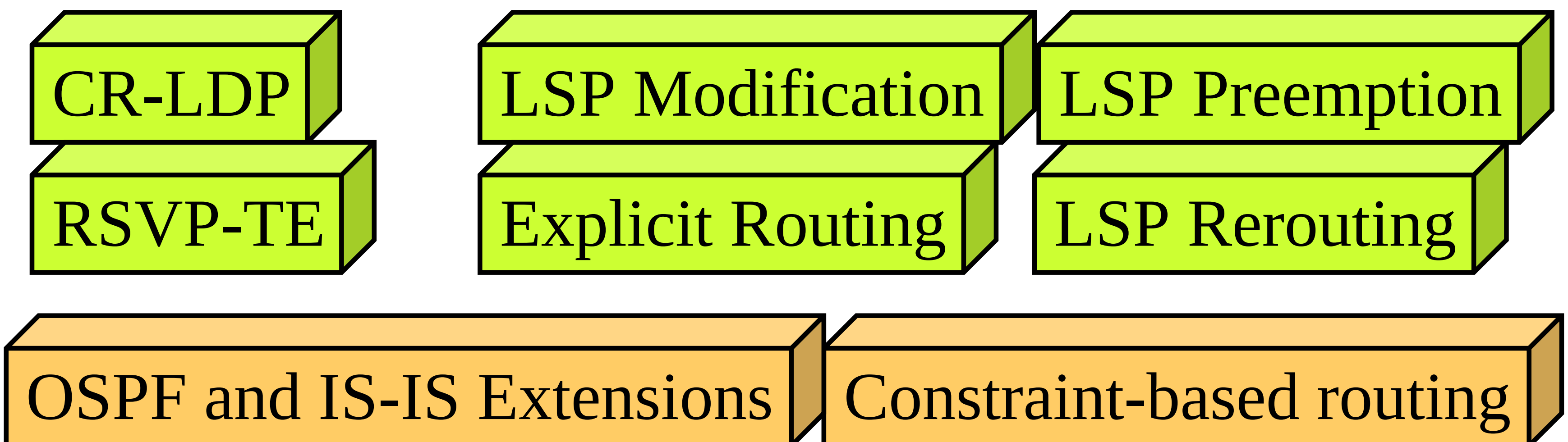


Hop-by-Hop vs Explicit Routing

Issue	Hop-by-hop	Explicit
Topology Awareness	Everywhere	Edge only
Circuit Management	None	LSP setup/ teardown
Signaling	Not required	Requires LDP or RSVP-TE
Recovery Time	Convergence time of routing Protocol	Path switch time
Routing	Fixed	QoS, Policy, or arbitrary
Traffic Engineering	Difficult	Easy

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



Draft Martini

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

IP		
Ethernet	ATM	PPP

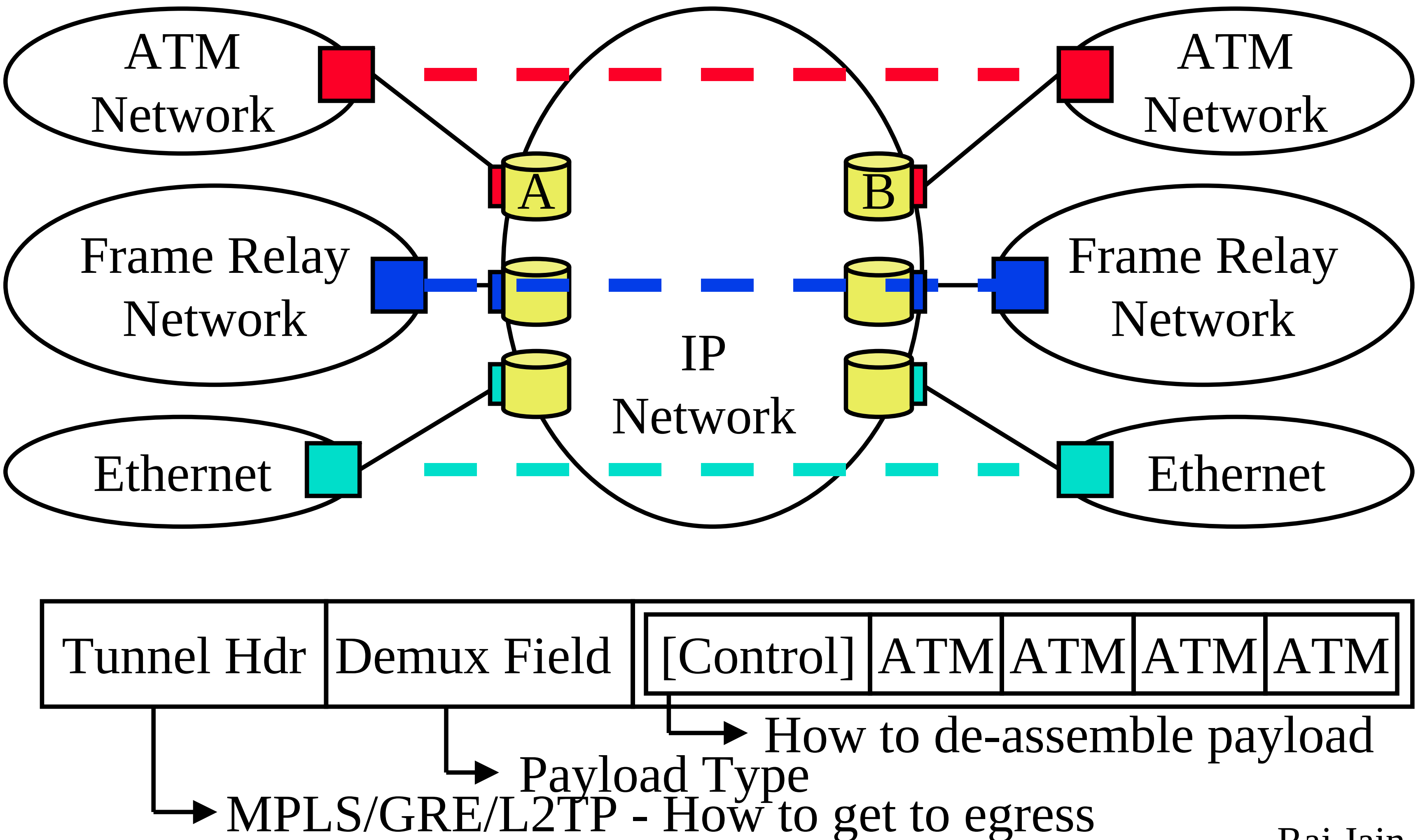
- 2000+: ATM over IP
  Ethernet over IP
  SONET over IP

Ethernet	ATM	PPP
IP		

- Ref: draft-martini-*.txt

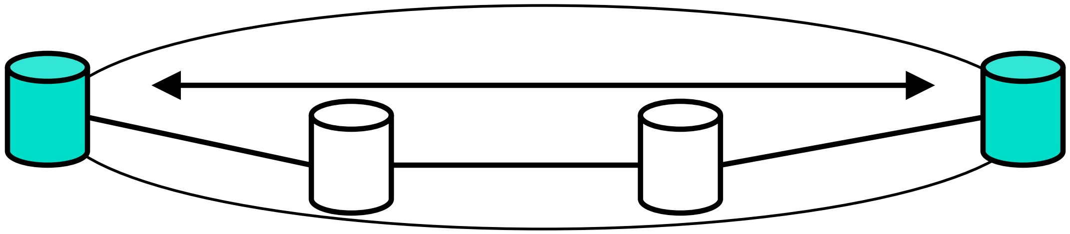
Raj Jain

L2 Circuits over IP



Raj Jain

VC Label



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

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

Raj Jain

ATM over MPLS

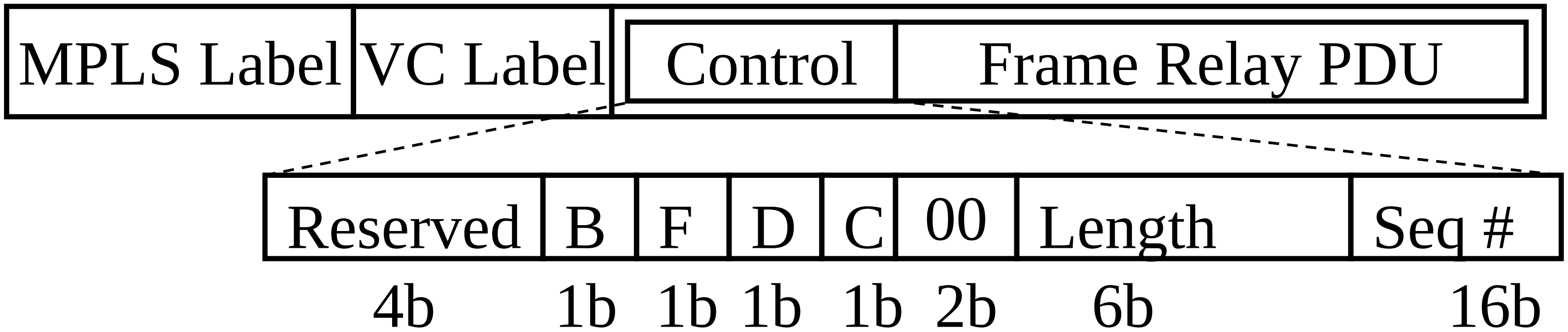
MPLS Label	VC Label	[Control]	ATM	ATM	ATM	ATM

MPLS Label	VC Label	Control	ATM AAL5 SDU

Reserved	T	E	C	C/R	00	Length	Seq #
4b	1b	1b	1b	1b	2b	6b	16b

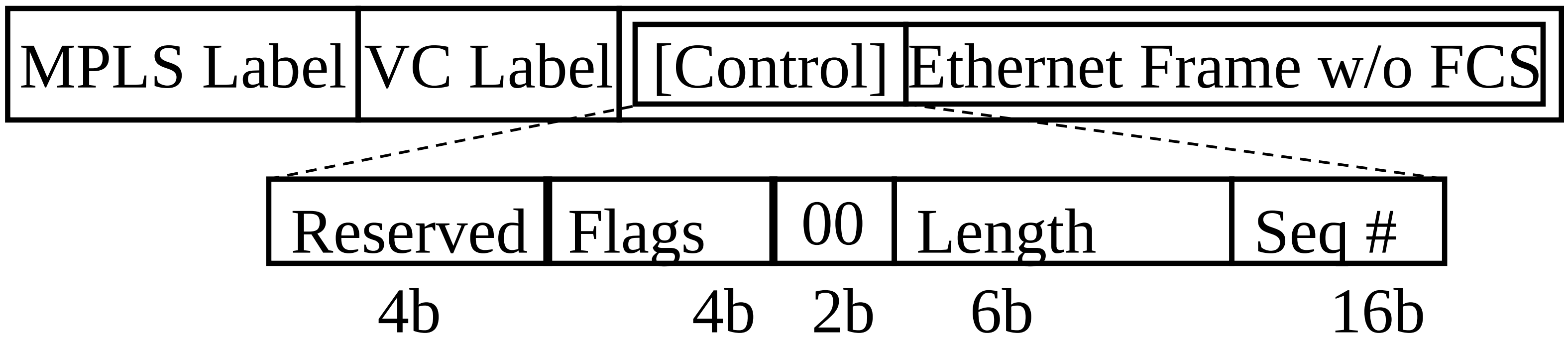
- T = Transport Type: 0=> Cells, 1=> SDU
- E = EFCI
- C = CLP
- C/R = Command/Response
- Length of payload + Control word
  0 => Greater than or equal to 64 bytes
- Ref: draft-martini-atm-encap-mpls-00.txt

Frame Relay over MPLS



- B = BECN = Backward Explicit Cong Notification
- F= FECN = Forward Explicit Cong Notification
- D = DE = Discard Eligible
- C = C/R = Command/Response
- B/F/D/C are copied from incoming frame to control at ingress and from control to outgoing frame at egress
- Length of payload + Control word if less than 64
- Ref: draft-martini-frame-encap-mpls-00.txt

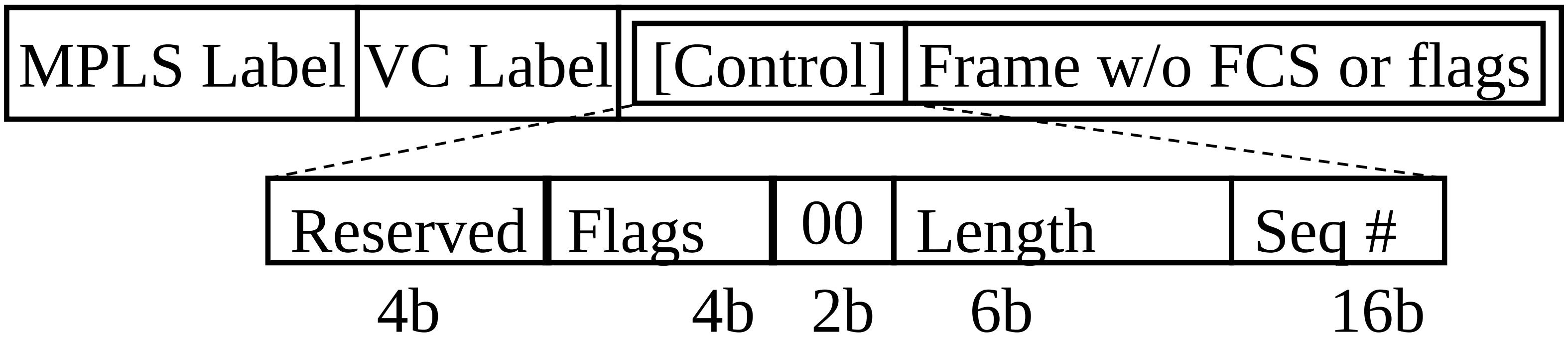
Ethernet over MPLS



- Control word is optional
- Flags are not used
- May put 802.1p priority in exp field of MPLS label
- Ref: draft-martini-ethernet-encap-mpls-00.txt

Raj Jain

PPP/HDLC over MPLS



- Control word is optional
- Flags are not used
- Bit/byte stuffing is undone
- Ref: draft-martini-ppp-hdlc-encap-mpls-00.txt

Raj Jain

Summary

1. Switching = forwarding based on label indexing
2. Labels ≈ ATM's VC id
3. MPLS allows label stacks, TTL, QoS
4. MPLS signaling via RSVP, LDP, CR-LDP, RSVP-TE
5. Traffic engineering using explicit paths
6. Draft-martini allows ATM, FR, Ethernet, PPP over MPLS

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

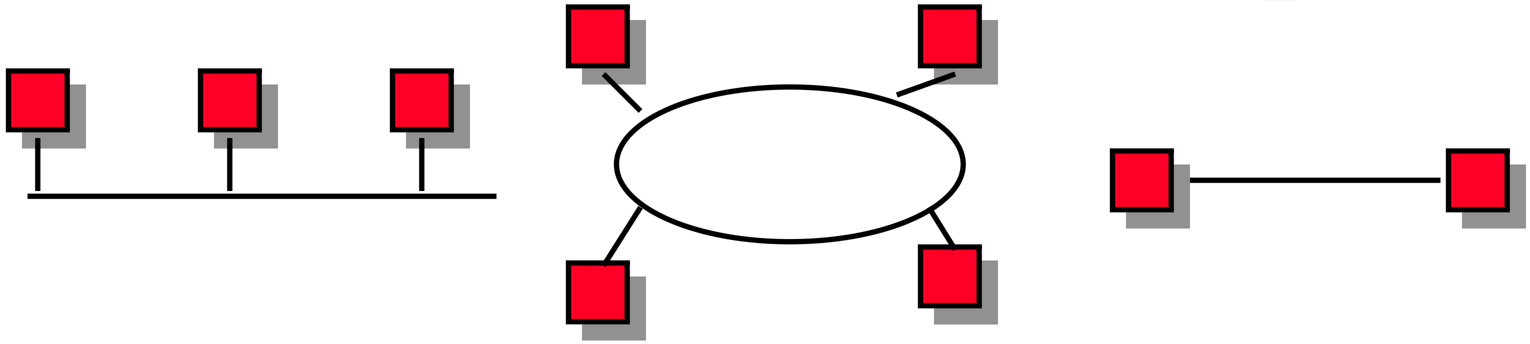
Gigabit Ethernet, 10 Gigabit Ethernet, and RPR

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

Raj Jain

Distance-B/W Principle

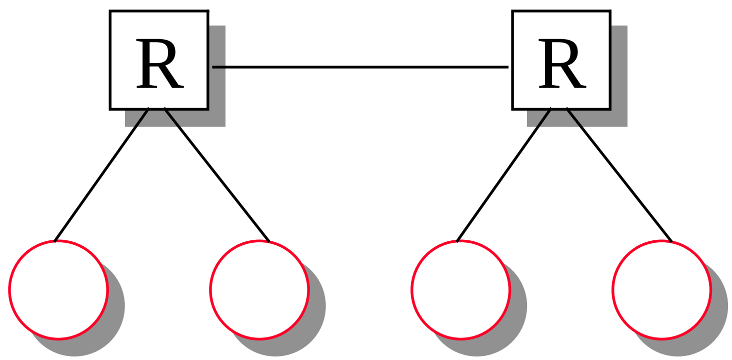


- Efficiency = Max throughput/Media bandwidth
- Efficiency is a non-increasing function of α
  α = Propagation delay /Transmission time
  = (Distance/Speed of light)/(Transmission size/Bits/sec)
  = Distance×Bits/sec/(Speed of light)(Transmission size)
- Bit rate-distance-transmission size tradeoff.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

Raj Jain

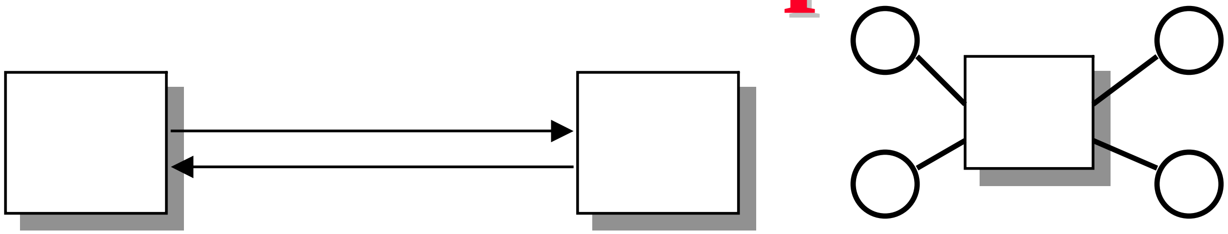
Ethernet vs Fast Ethernet

	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

Raj Jain

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

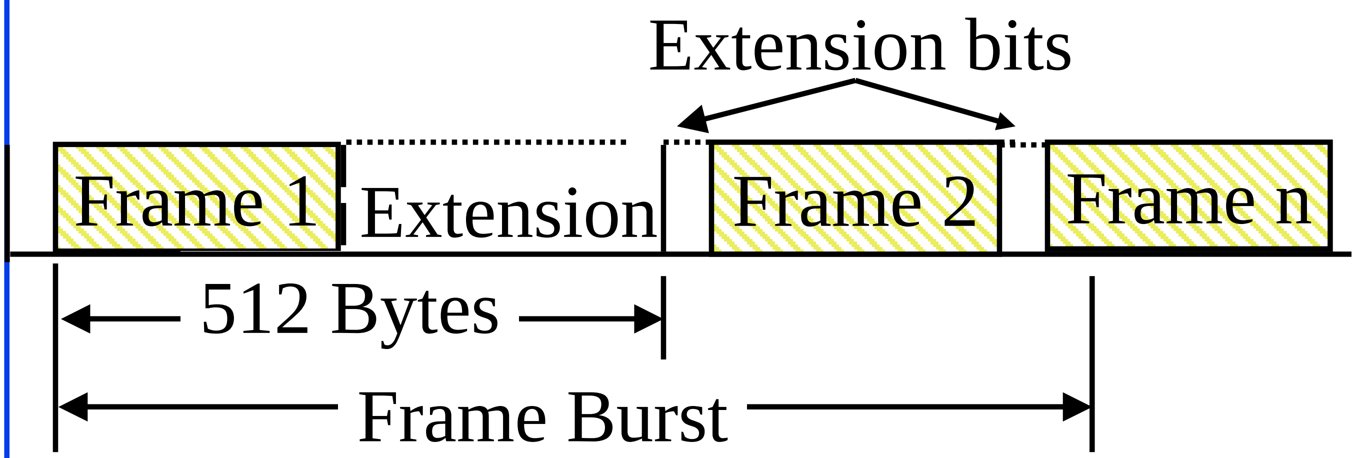
Media Access Control Issues

- Carrier Extension
- Frame Bursting

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

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

1000Base-X

- 1000Base-LX: 1300-nm <u>laser</u> 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 <u>laser</u> 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 <u>shielded</u> twinax cable in a single room or rack.

    Uses 8b/10b coding $\Rightarrow$ 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.

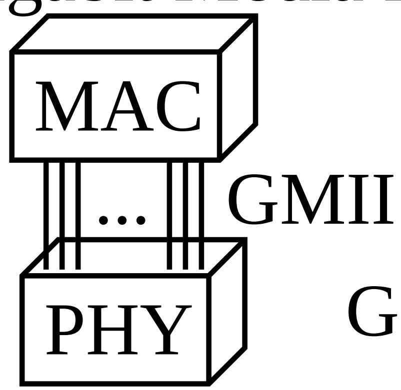
Raj Jain

1000Base-T

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

1000BASE-T (Cont)

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



GBICs??

Raj Jain

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

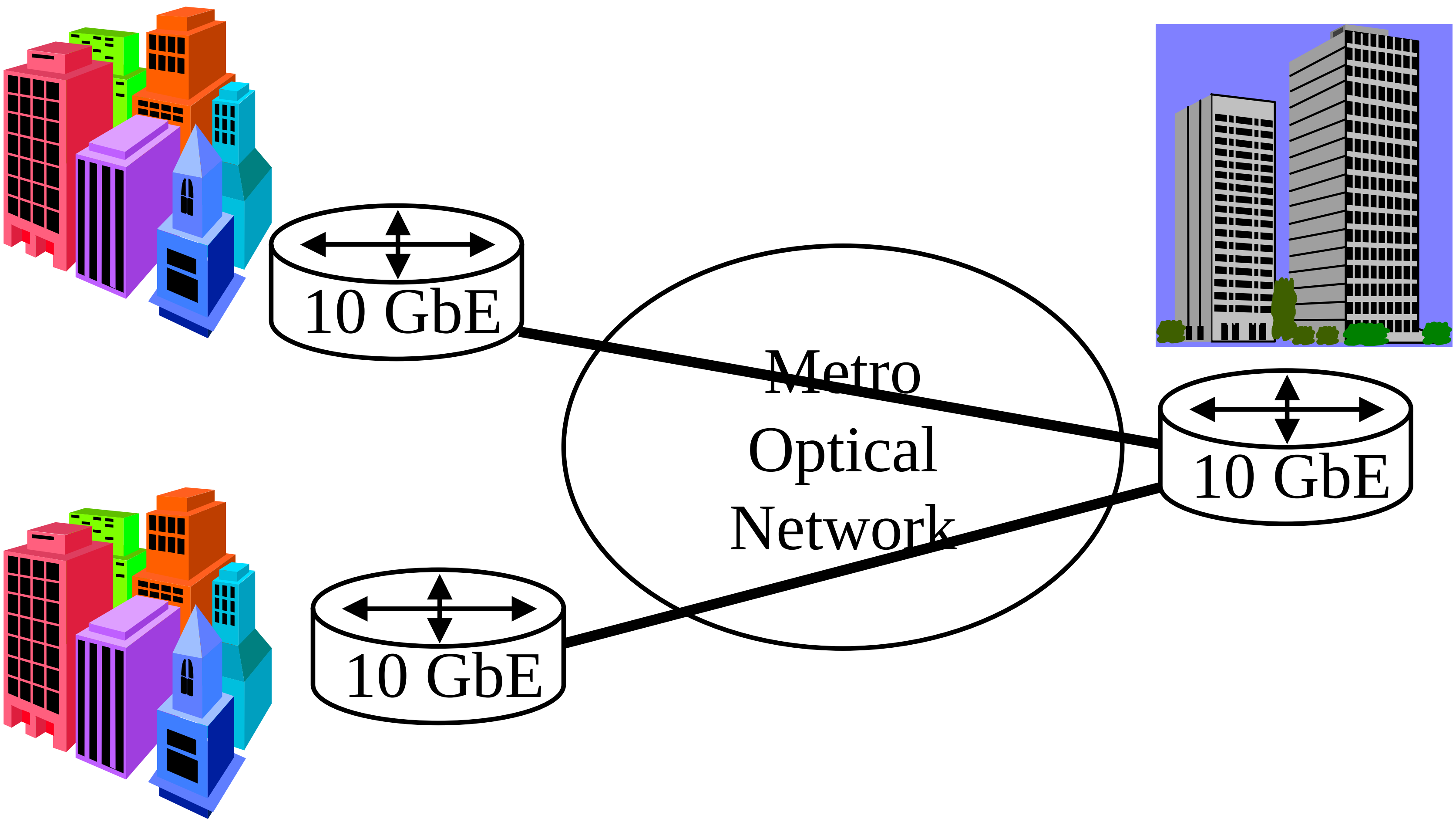
- P802.3ae ⇒ Update to 802.3
- Compatible with OC-192c Payload rate
- Compatible with 802.3 frame format, services, management
- LAN and WAN PHY families
- Cost = 3 × 1GbE
- Same min and max frame size as 10/100/1000 Mbps
- Full-duplex only ⇒ No CSMA/CD
- Star-wired point-to-point links
- 10.000 Gb/s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

10 GbE PMD Types

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

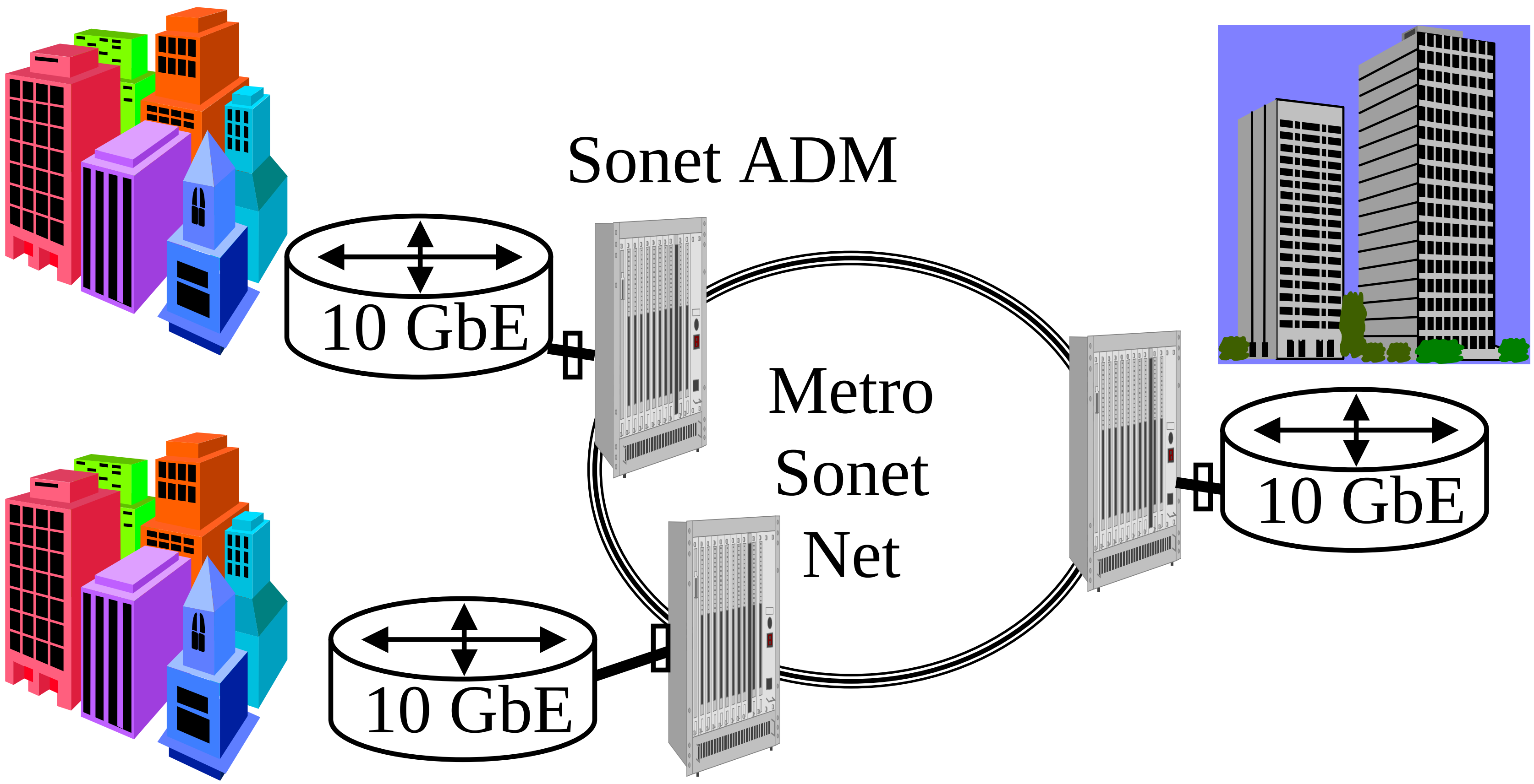
- S = Short Wave, L=Long Wave, E=Extra Long Wave
- R = Regular reach (64b/66b), W=WAN (64b/66b + SONET Encapsulation), X = 8b/10b
- 4 = 4 λ's

10 GbE over Dark Fiber



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

10 GbE over Sonet/SDH



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

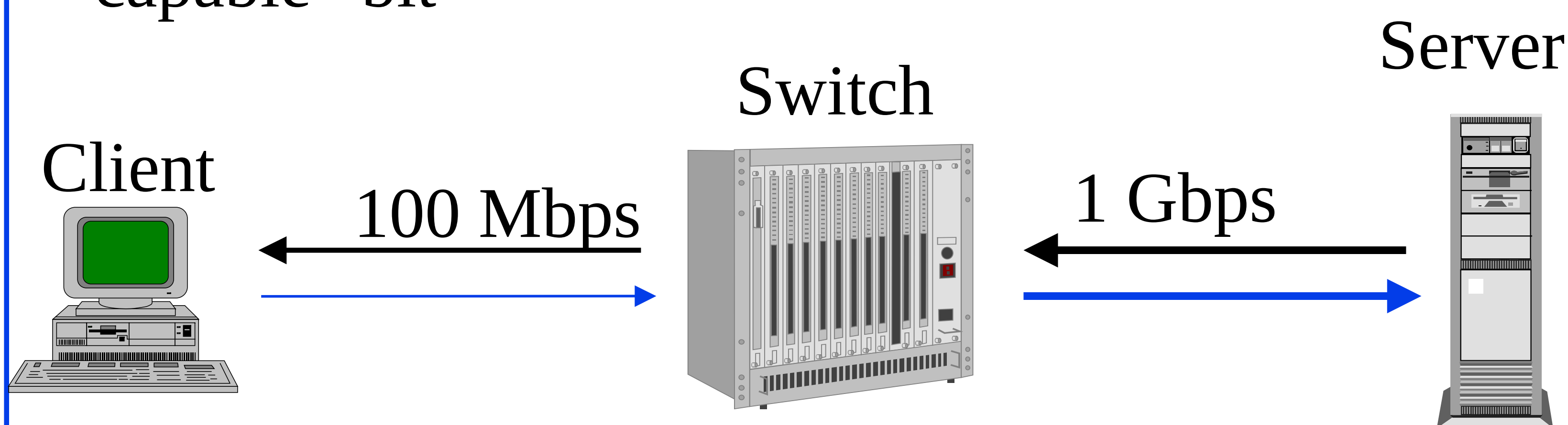
Feature	SONET	Ethernet	Remedy
Payload Rates	51M, 155M, 622M, 2.4G, 9.5G	10M, 100M, 1G, 10G	10GE at 9.5G
Payload Rate Granularity	Fixed	√Any	Virtual Concatenation
Bursty Payload	No	√Yes	Link Capacity Adjustment Scheme
Payload Count	One	√Multiple	Packet GFP
Protection	√Ring	Mesh	Resilient Packet Ring (RPR)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

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

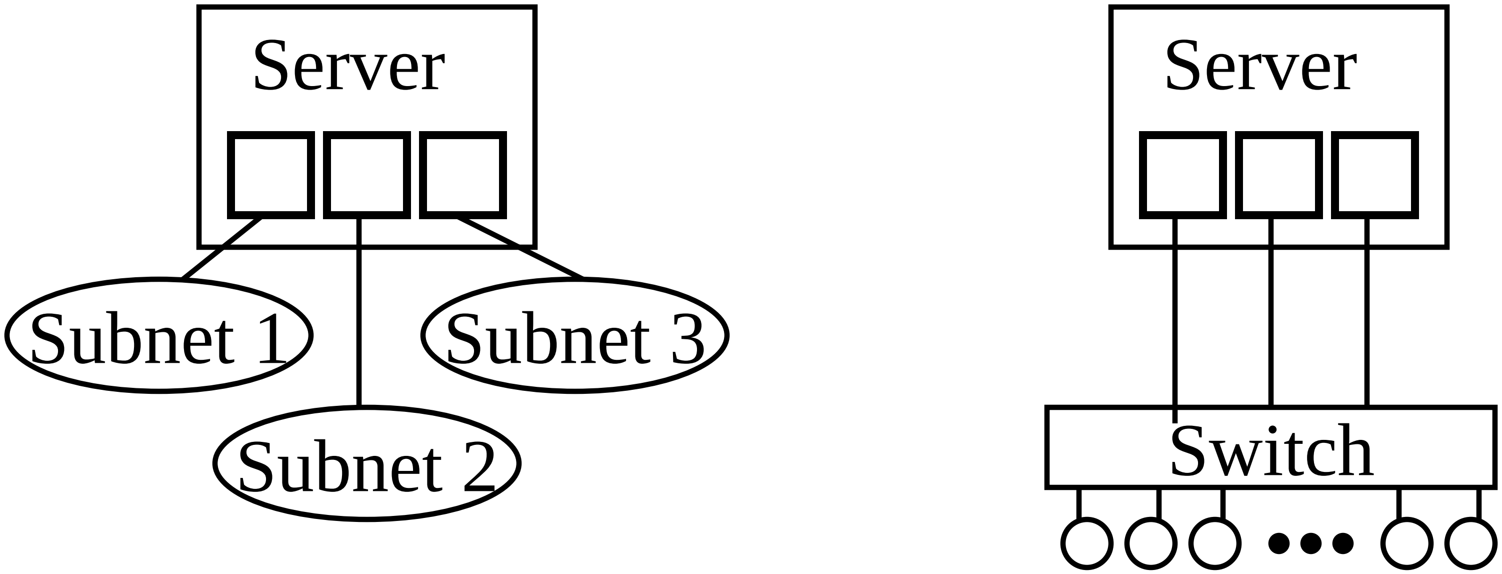
- Pause frame with pause time sent to multicast address 01-80-C2-00-00-01 not forwarded by bridges
- Autonegotiation updated to include a "flow-control capable" bit



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

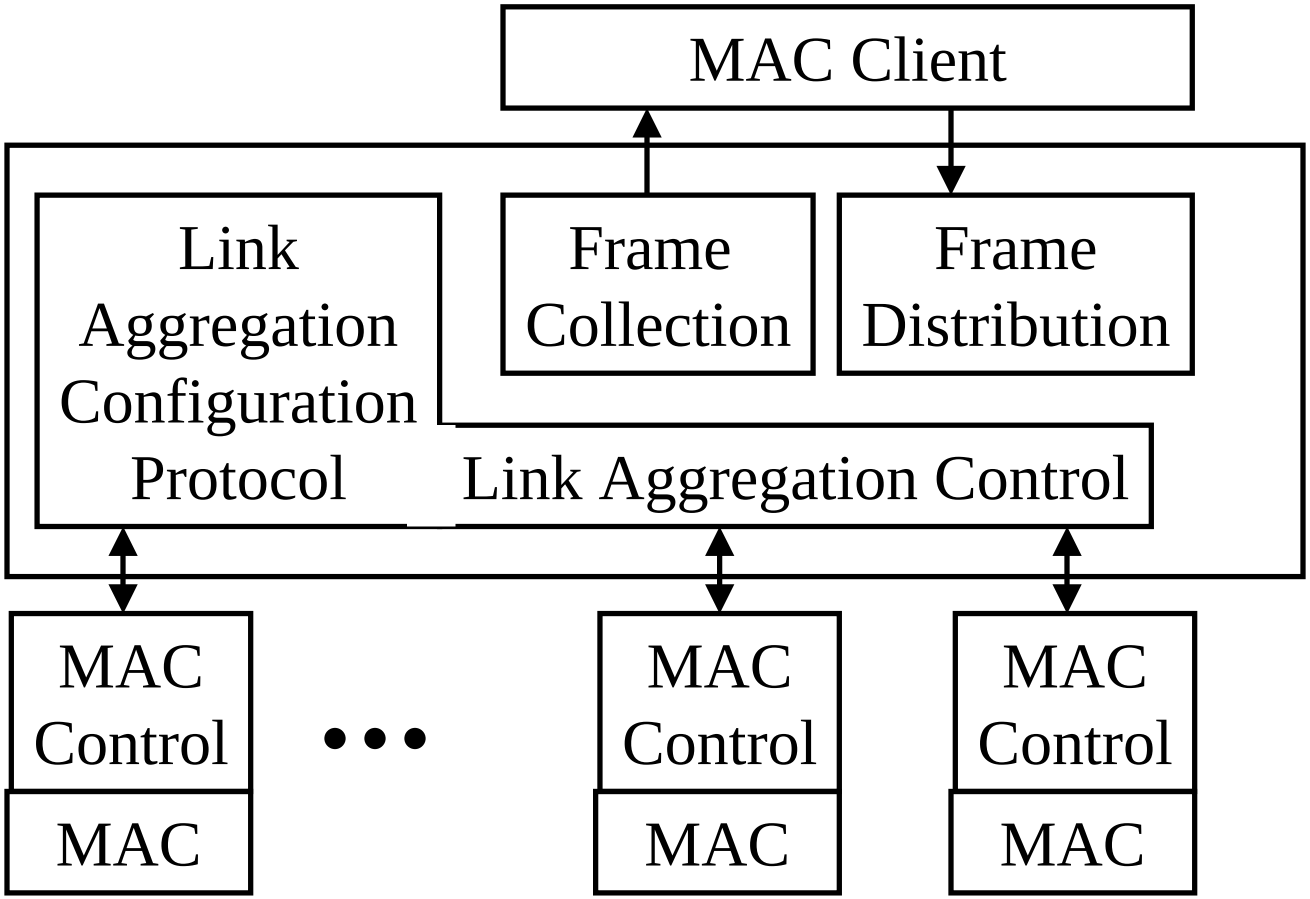
Raj Jain

802.3ad Link Aggregation



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

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
- Is not an IEEE standard
- Ref: http://www.nwfusion.com/newsletters/lans/0614lan1.html

Rapid Spanning Tree Protocol (RSTP)

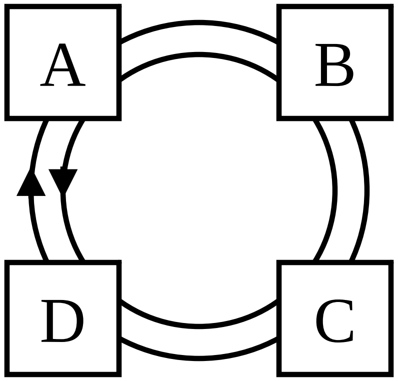
1. Builds upon the known topology rather than starting fresh.
2. Topology change is sent along designated ports (to sub-tree). Not all ports.
3. If the root port becomes disabled, alternate port becomes root port.
4. The learned address database (stations towards the root) is not flushed but transferred.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

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

Multiple Spanning Tree

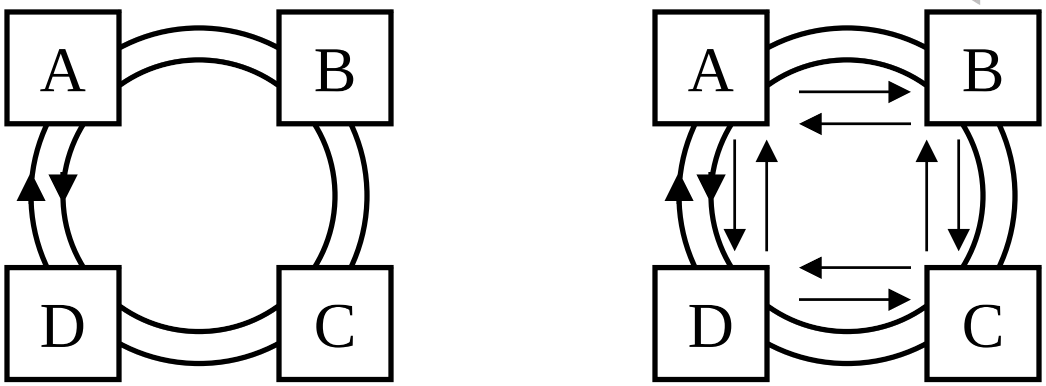
- 802.1s for VLANs
- Old Bridge ID = 16-bit priority + 48-bit MAC Address
- New Bridge ID = 4-bit priority + 12-bit VLAN ID + 48-bit MAC Address

RPR: Key Features



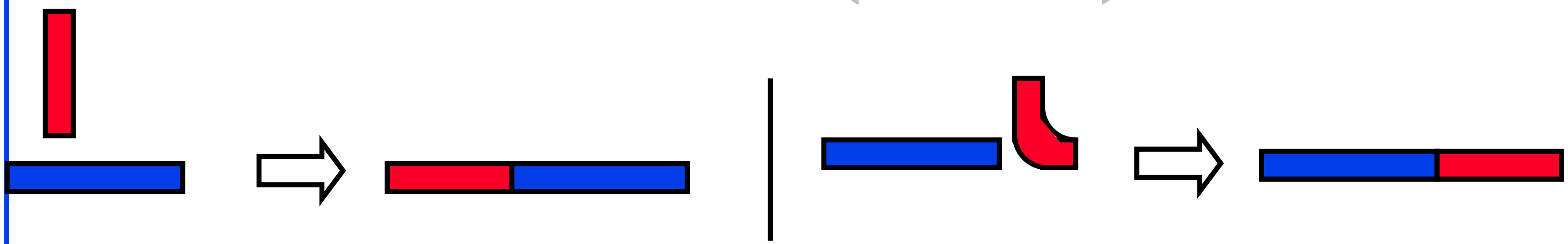
- Dual Ring topology (like FDDI)
- Supports broadcast and multicast
- Packet based ⇒ Continuous bandwidth granularity
- Max 256 nodes per ring
- MAN distances: Several hundred kilometers.
- Gbps speeds: Up to 10 Gbps

RPR Features (Cont)



- Both rings are used (unlike SONET)
- Normal transmission on the shortest path
- Destination stripping ⇒ Spatial reuse
  Multicast packets are source stripped
- Five Classes of traffic: Reserved, High-Priority, Medium Priority, Low Priority, Control

RPR (Cont)

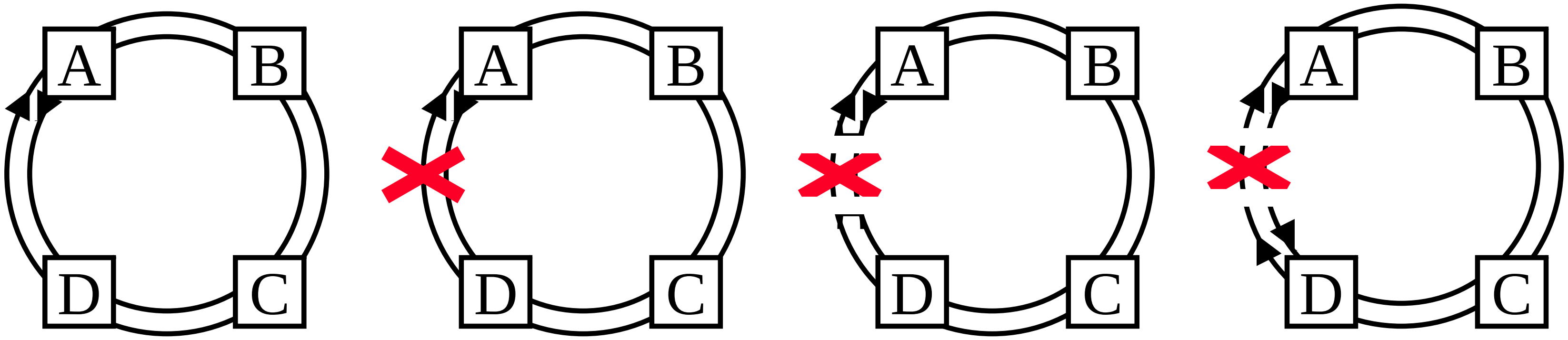


- Buffer Insertion Ring: Absolute but non-preemptive priority to pass-through traffic
- Cut-through of transit packets optional.
- Bandwidth management: Unused bandwidth is advertised so that others can use it
- Fairness Algorithm for fair and efficient bandwidth use
- Physical Layer Independent: GbE/10GE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

RPR Traffic Classes

1. **Reserved**: Pre-allocated. Not used even if idle. TDM.
2. **High Priority**: Bounded Delay and jitter.
   Shaped at ingress to Committed Information Rate (CIR), Excess Burst Size (Be), and Committed Rate Measurement Interval (Tc). Out-of-profile pkts dropped.
3. **Medium Priority**: Guaranteed throughput.
   No delay sensitivity. Shaped to CIR, Be, and Tc.
   Out-of-profile packets are tagged.
   Total throughput = CIR + Fair allocation.
4. **Low Priority**: Best effort traffic class.
   No throughput or delay guarantee.
5. **Control**: Protection/Fairness/Topology messages

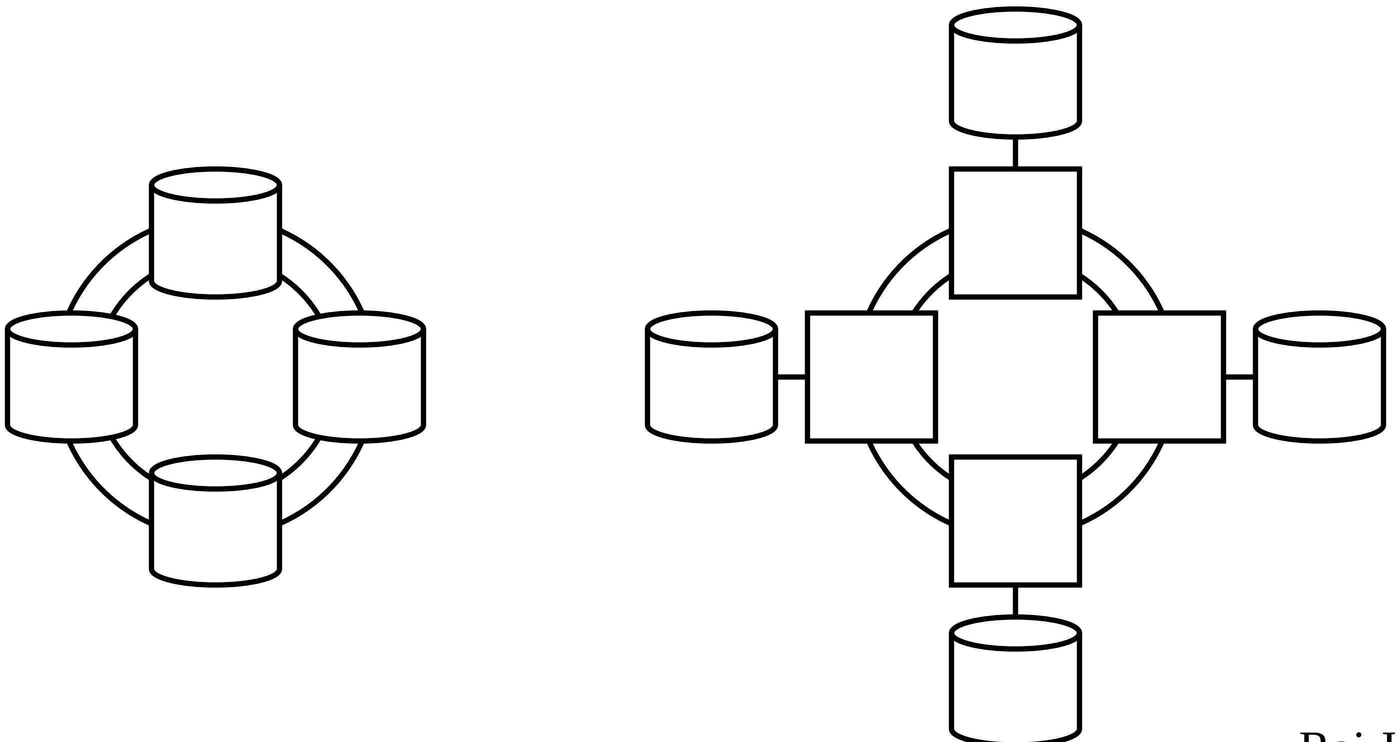
RPR Protection Mechanisms



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

RPR Issues

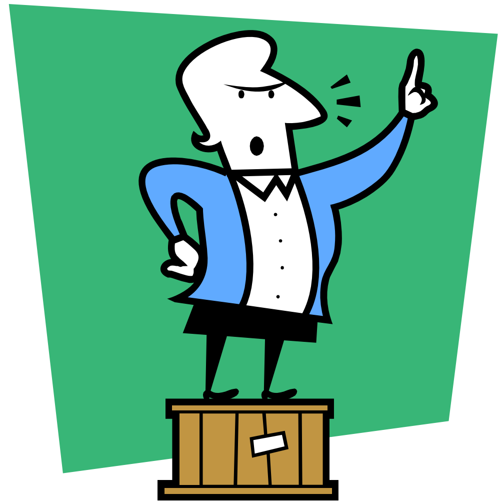
- Ring vs Mesh (Atrica)
- Router Feature vs Dedicated RPR Node (Cisco, Redback, Riverstone vs Luminous)



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

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, 10 GbE, RPR: 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
- RPR Alliance, http://www.rpralliance.org/

Raj Jain

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

Storage Area 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/

- Five Trends in Storage
- What is SAN?
- SAN vs NAS
- Fibre Channel, ESCON
- SAN Devices
- IP Storage: iSCSI, iFCP, FCIP, iSNS

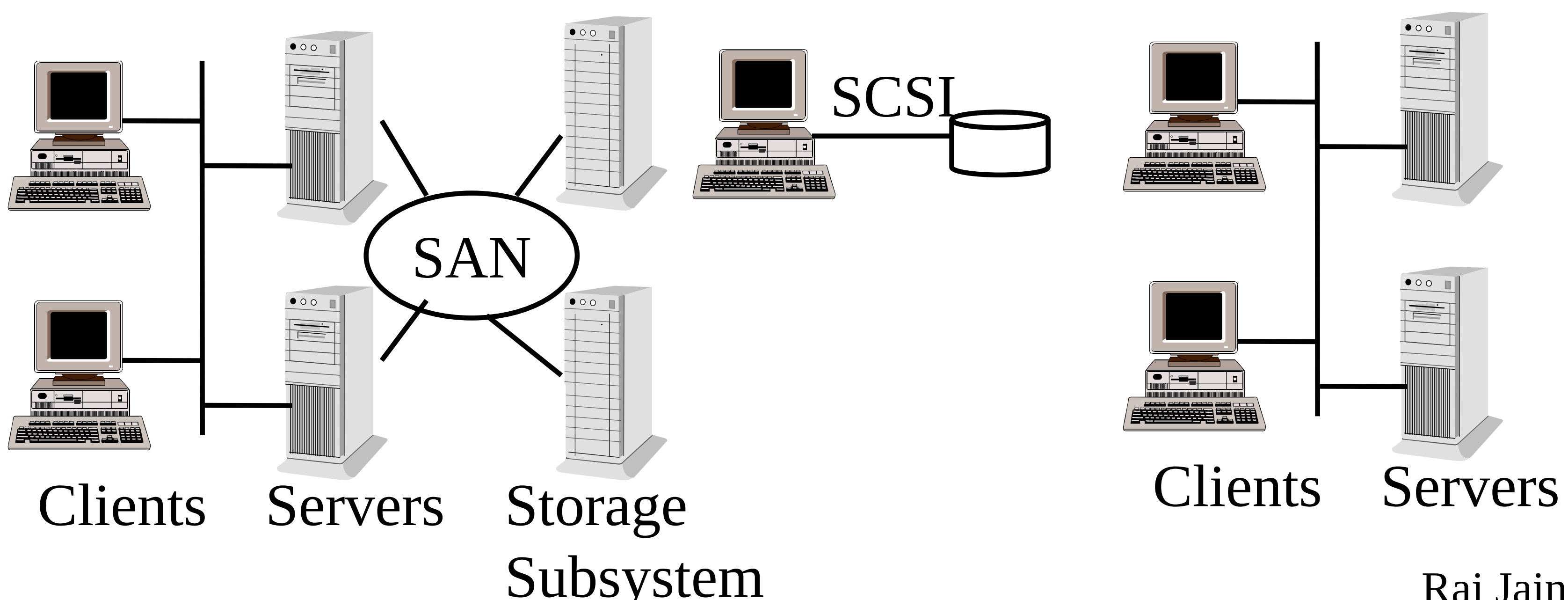
Five Trends in Storage

1. New applications are fueling storage demands. Streaming video, digital photography, MP3 Audio...
2. High-speed networking ⇒ increasing flow of info.
3. Storage prices are declining 43% per year - Solomon Smith Barney ⇒ Demand is doubling every year
4. Data is critical for businesses. Financial Brokerage Houses can loose $6M/hour ⇒ Data needs to be replicated for high-availability.
5. Centralized storage: Reduced OpEx. Easy replication and disaster recovery. Higher storage utilization.

All of these trends are leading to storage access over networks.

What is SAN?

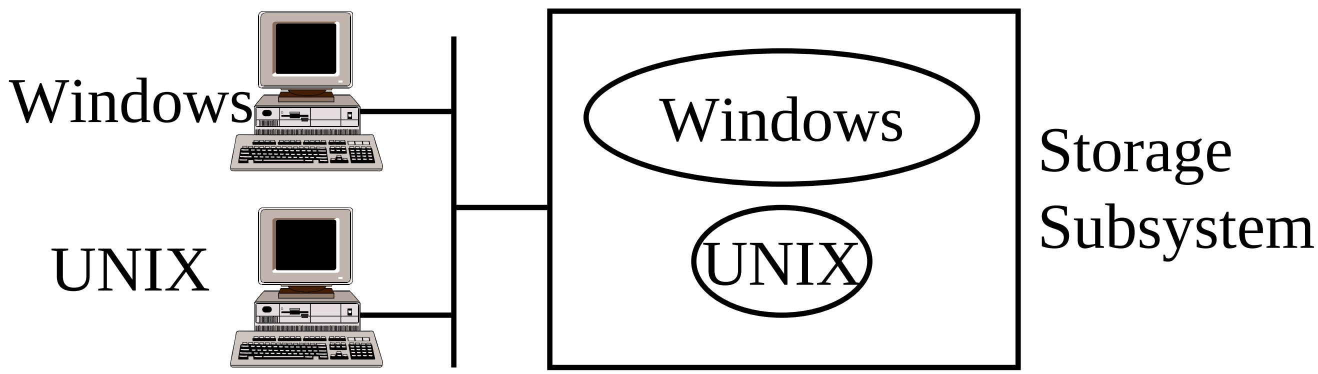
- SAN = A network exclusively for storage
- Direct Attached Storage (DAS)
- Network Attached Storage (NAS): Share existing networks for data and storage



Raj Jain

SAN vs NAS

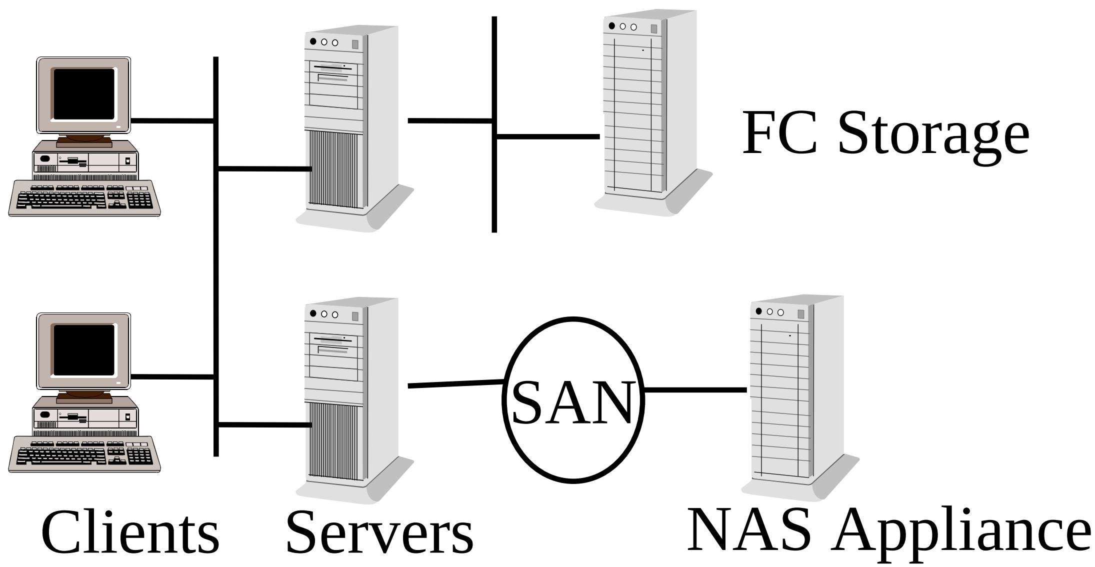
	SAN	NAS
Protocol	Fibre Channel	NFS, TCP/IP, Ethernet
Network	Fibre Channel	Ethernet
Transfer	Block	File
OS Independent	Yes	Some



- Block level operations like disk mirroring and striping are easier with SAN than NAS

Trend: SAN and NAS Convergence

- Debates have subsided
- SAN protocols over shared LANs (iFCP or FCIP)
  $\Rightarrow$ SAN appliances over LANs and vice versa
- LANs improved to meet storage QoS



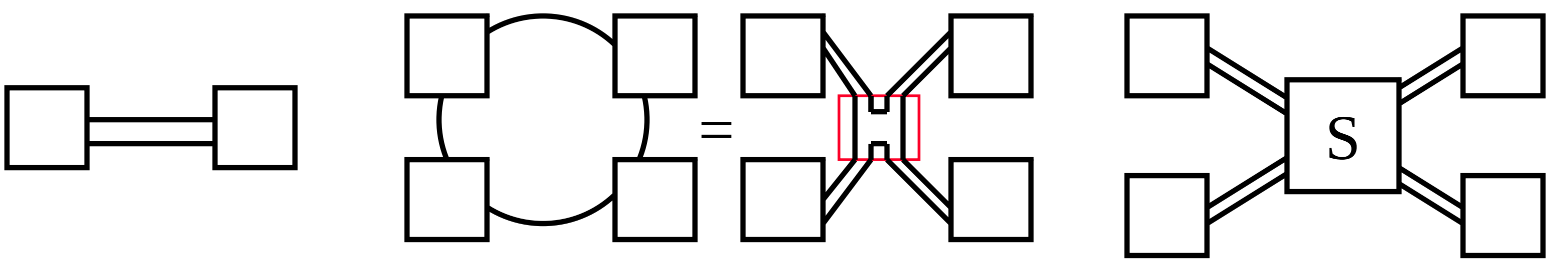
Raj Jain

Fibre Channel

- Developed in 1980's as a LAN
- French spelling to emphasize copper medium too.
- Designed to be a datalink similar to Ethernet to carry IP, FICON, SCSI traffic.
- Revitalized by EMC and Brocade as Enhanced SCSI in mid-1990s $\Rightarrow$ Took off in 1999
- 126 Devices up to 10 km initially
- Serial $\Rightarrow$ Longer distances
- 1 Gbps FC runs at 100 MBps
  2 Gbps FC runs at 200 MBps

Fibre Channel (Cont)

- Full-duplex ⇒ Simultaneous transmit/receive
- Supports point-to-point, arbitrated loop, and switched topologies
- Switched FC allows 16 million devices



FC Protocol Layers

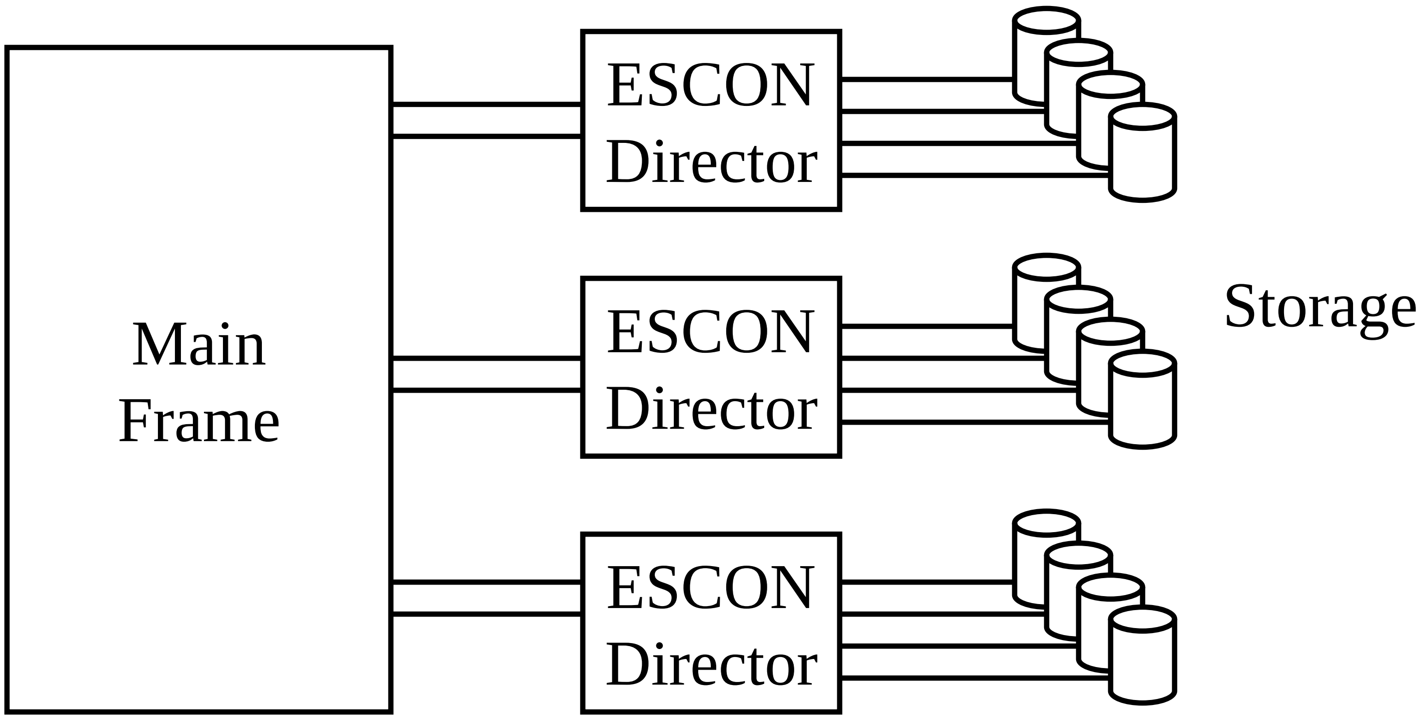
FC-4	Upper Layer Protocol Mapping. SCSI, FICON, 802.2, HIPPI, and IP
FC-3	Common Services - striping and mirroring (multicast)
FC-2	Data Delivery - framing, CoS, flow control. Topology support.
FC-1	Byte Encoding - 8b/10b for serial transmission.
FC-0	Physical Layer - Defines transmission rates, Cables, connectors, transmitters, receivers. 26 different media (historical) at quarter, half, full speeds

ESCON

- Introduced by IBM in 1991
- Enterprise Systems Connectivity (ESCON) = Mainframe SCSI
- 17 MBps half-duplex to 3 km on Fiber
- Fibre Connection Channel (FICON) is a FC-4 layer protocol for using Fibre Channel
- 100 MBps full-duplex up to 100 km being developed at T11

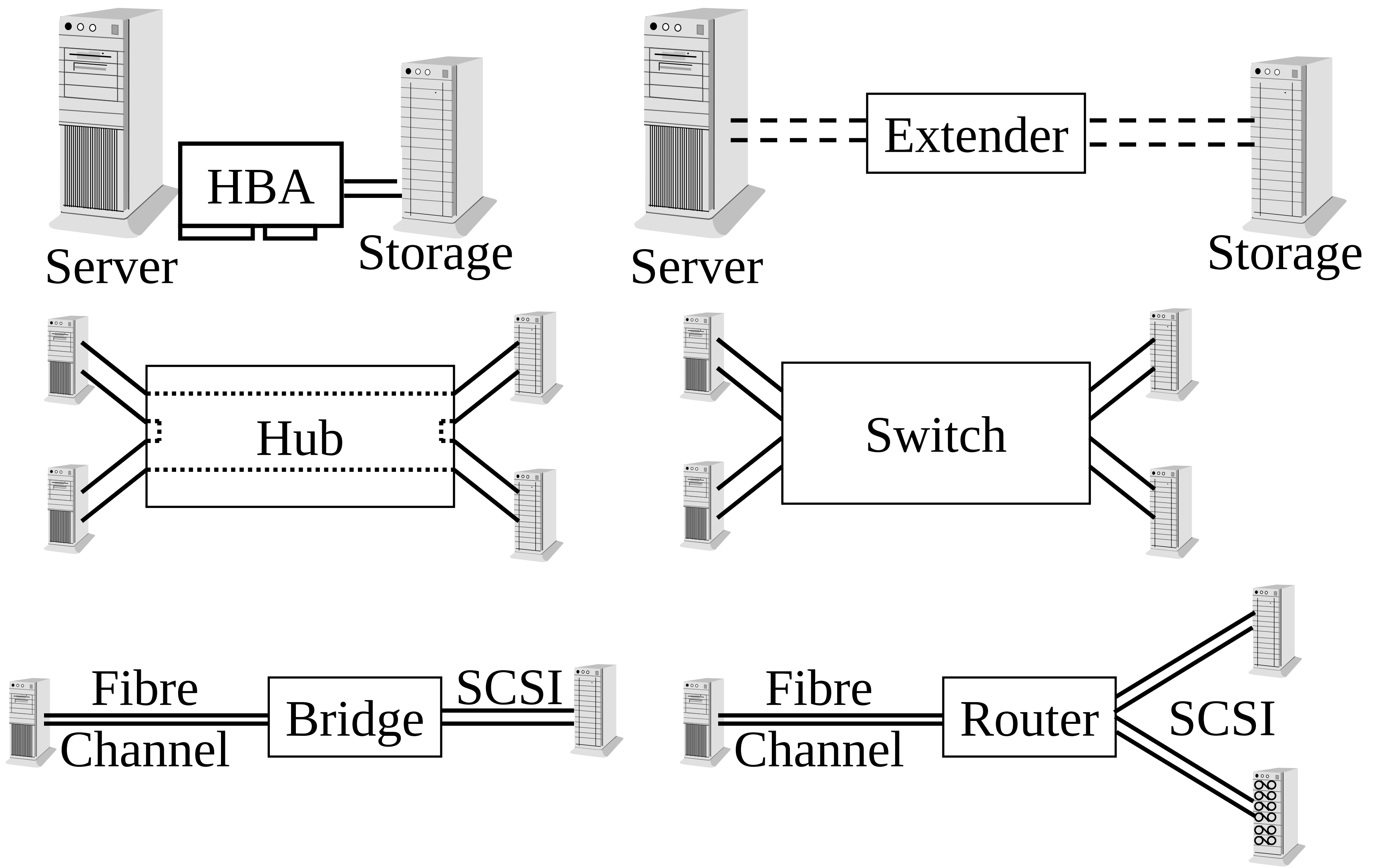
ESCON (Cont)

- Problem: S/390 has max 256 ESCON channels
- Multiple systems need access to same data
- Solution: ESCON Directors



Raj Jain

SAN Devices



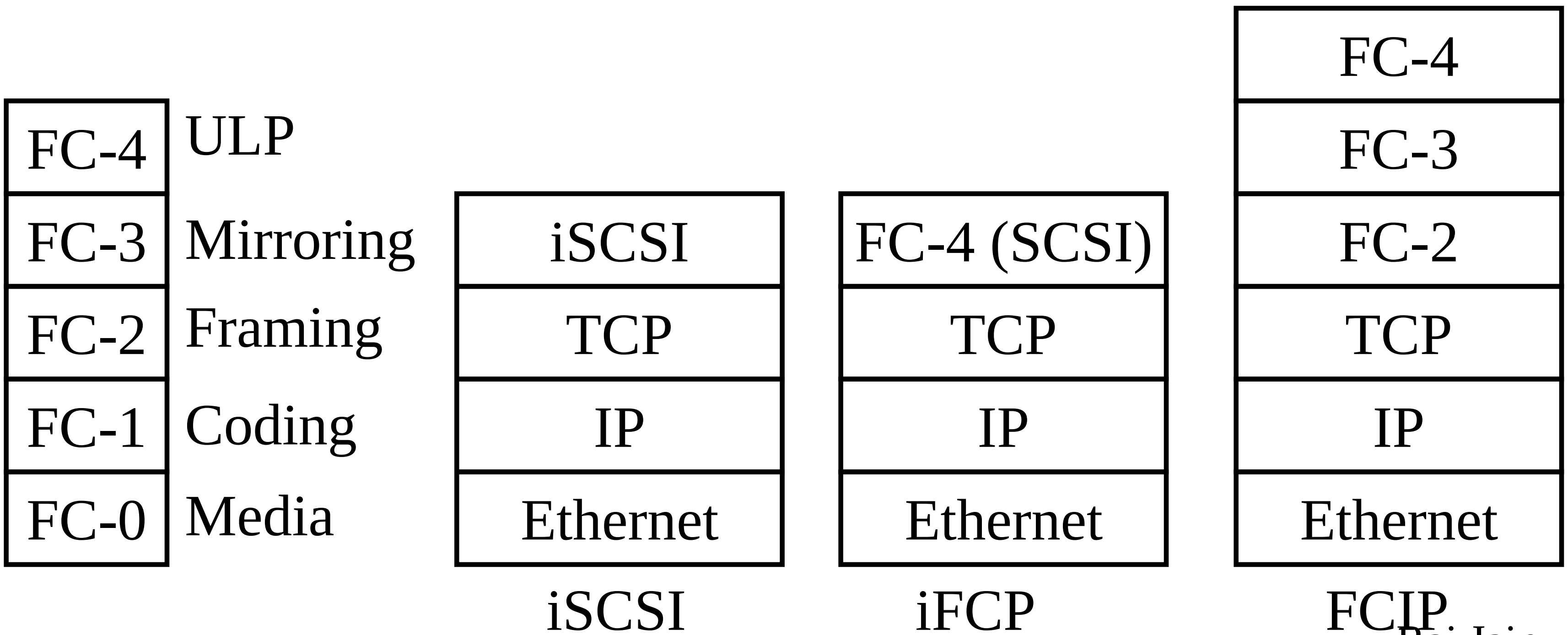
Raj Jain

SAN Devices

- Servers and Storage
- HBA: Host Bus Adapters, e.g., Adaptec A5158A-FC
- Extenders
- Hub: Based on loop $\Rightarrow$ One at a time connectivity. Low end, e.g., HP Hub S10
- Switch: Provides simultaneous any to any connectivity. Up to 16 ports, e.g., HP Switch F16
- Bridge: Connects two dissimilar SANs (SCSI and FC), e.g., HP Bridge FC 4/2
- Router: Connects multiple dissimilar SANs/LANs, e.g., FC to SCSI or Ethernet to FC
- Multiservice box: FC, ESCON, LAN, TDM

IP Storage

- iSCSI: Send SCSI commands as data over TCP/IP.
- iFCP: Send SCSI commands using FC-4 over TCP/IP. Allows FC software to be reused
- FCIP: Tunnel FC over IP. Allows FC hardware on IP.



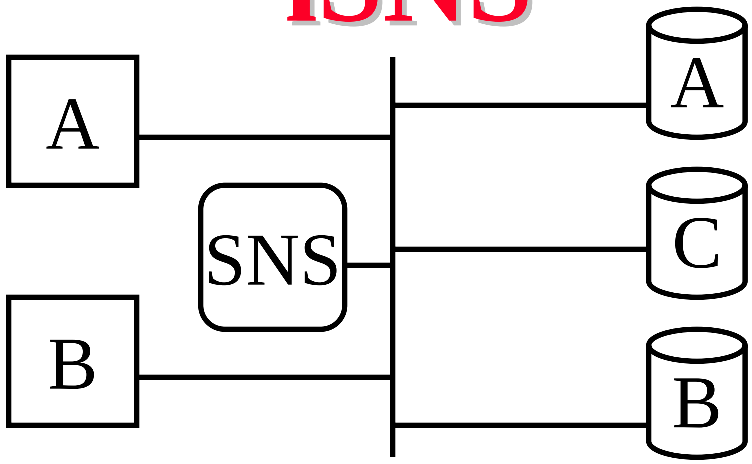
iSCSI Products

- HBAs combine SCSI HBA+Ethernet NIC
- Software iSCSI HBA allow SCSI transfers over Ethernet NICs
- Adaptec AEA-7110C iSCSI HBA
- Emulex GN9000/SI iSCSI HBA
- IBM IP Storage 200i NAS appliance
- Brocade Silkworm Fibre Channel Switches
- Cisco SN 5420 Storage Router

iSCSI Operation

- All nodes (initiators and targets) have names, e.g., iqn.1998-03.com.disk-vendor.diskarrays.sn.45678
- All nodes also have addresses IP:port, e.g., 192.64.107.61:4260
- Default port 3260 assigned by IANA
- Targets listen to well-known TCP port
- Initiators send login requests to setup a connection
- Login = Allows mutual authentication, Parameter exchange
- Ref: draft-ietf-ips-iSCSI-10.txt, 20-Jan-2002

iSNS

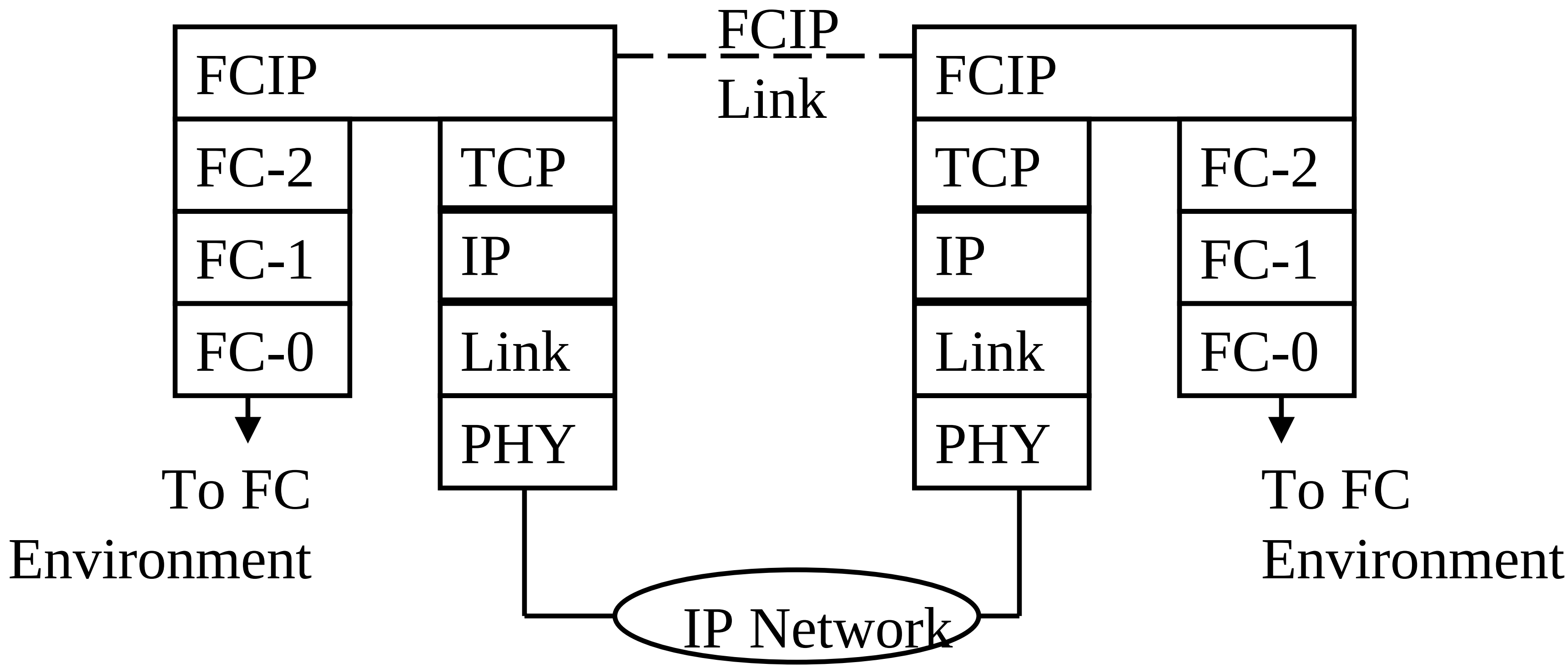


- Internet Storage Name Service (iSNS)
- Allows discovery and management of iSCSI and FC devices in IP networks
- Both targets and initiators can register
- Extension of Fibre Channel Generic Services Name Server FC-GS-3
- Discovery domains like VLANs

iSNS (Cont)

- Each node can be member of multiple domains
- Login Control: Targets obtain the list of allowed initiators and their authentication info from iSNS
- State change notification service: provided by iSNS
- Allows proxy services for iSCSI devices accessed by FC and vices versa
- iSNS and DNS may be in one box and may use a common database
- Ref: draft-ietf-ips-isns-08.txt, February 2002

FCIP Protocol Stack



- Ref: draft-ietf-ips-fcovertcpip-09.txt, January 2002.

Raj Jain

Summary

- Storage is growing fast and centralized storage is easier to manage.
- Storage area network is a network designed specifically for storage.
- Fibre Channel is the primary SAN protocol with FC Hubs, switches, bridges, routers, and directors.
- SAN vs NAS debate is now over. SAN protocols can be tunneled through networks.
- Storage will be managed over WAN distances using IP transport iSCSI, iFCP, FCIP

SANs: Key References

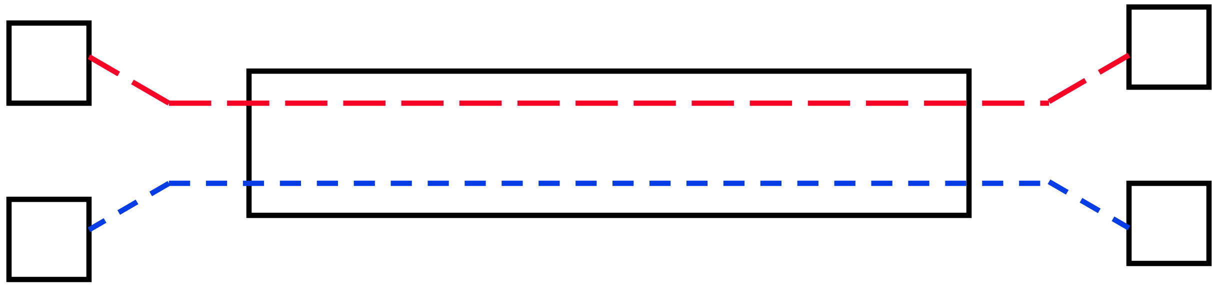
- See detailed references in http://www.cis.ohio-state.edu/~jain/refs/san_refs.htm
- IP Storage (ips), http://www.ietf.org/html.charters/ips-charter.html
- Fibre Channel Association (FCA), www.fibrechannel.com

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/

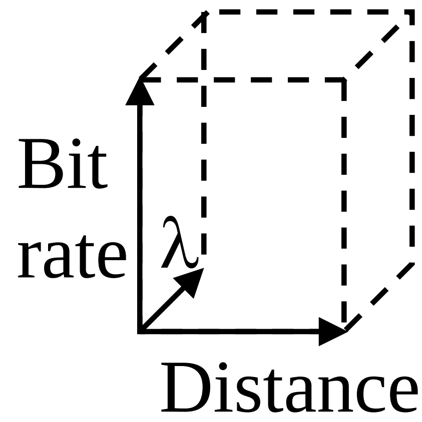
- Recent DWDM Records and Product Announcements
- Why IP over DWDM?
- How to IP over DWDM?
  - What changes are required in IP?
  - MPλS and GMPLS
  - UNI, LDP, RSVP, LMP
- Upcoming Optical Technologies

Sparse and Dense WDM



- 10Mbps Ethernet (10Base-F) uses 850 nm
- 100 Mbps Ethernet (100Base-FX) + FDDI use 1310 nm
- Some telecommunication lines use 1550 nm
- WDM: 850nm + 1310nm or 1310nm + 1550nm
- Dense ⇒ Closely spaced ≈ 0.1 - 2 nm separation
- Coarse = 2 to 25 nm = 4 to 12 λ's
- Wide = Different Wavebands

Recent DWDM Records



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

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

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

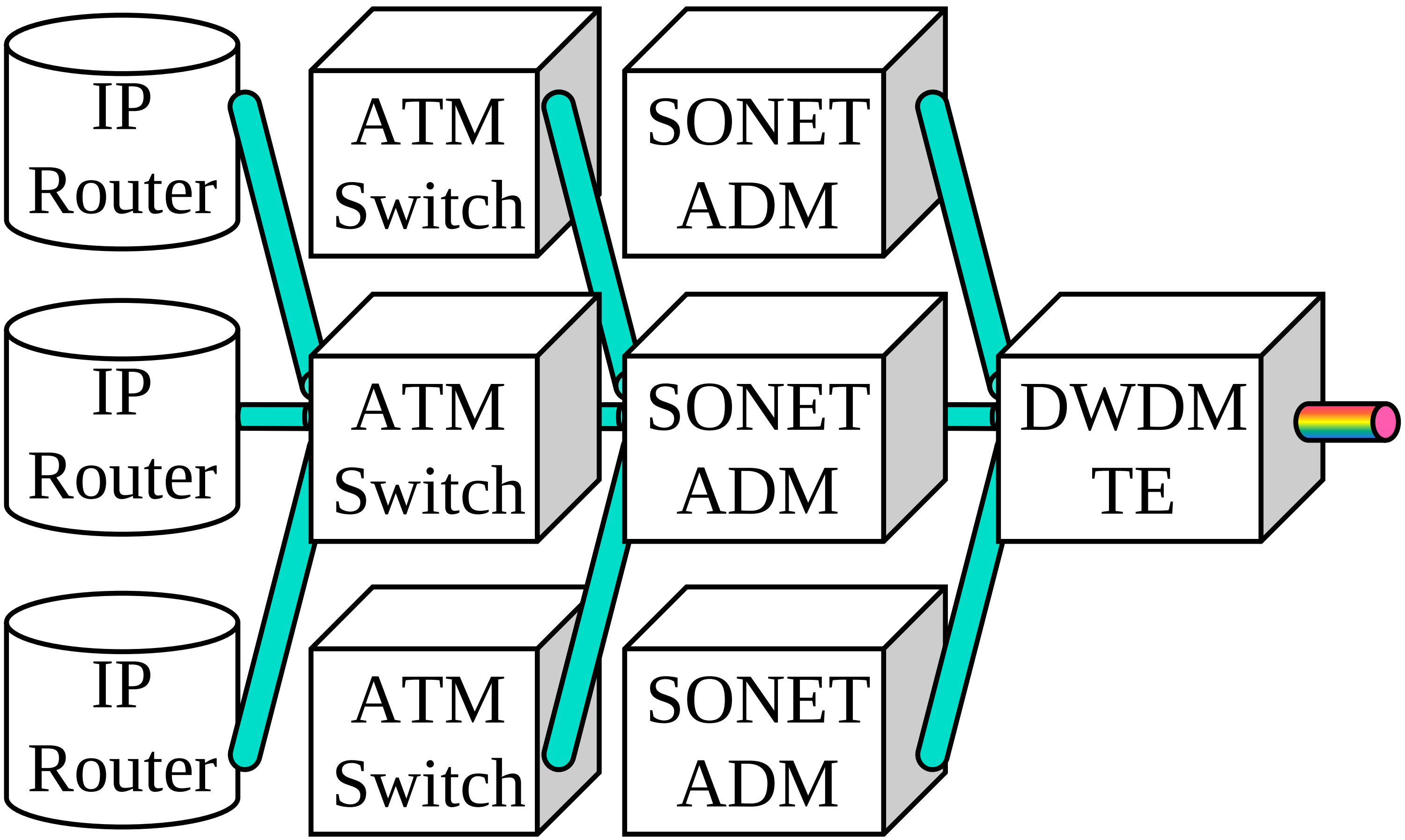
Recent Products Announcements

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
Cisco ONS 15800	160	10	2000	2002

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

Raj Jain

IP over DWDM (Past)



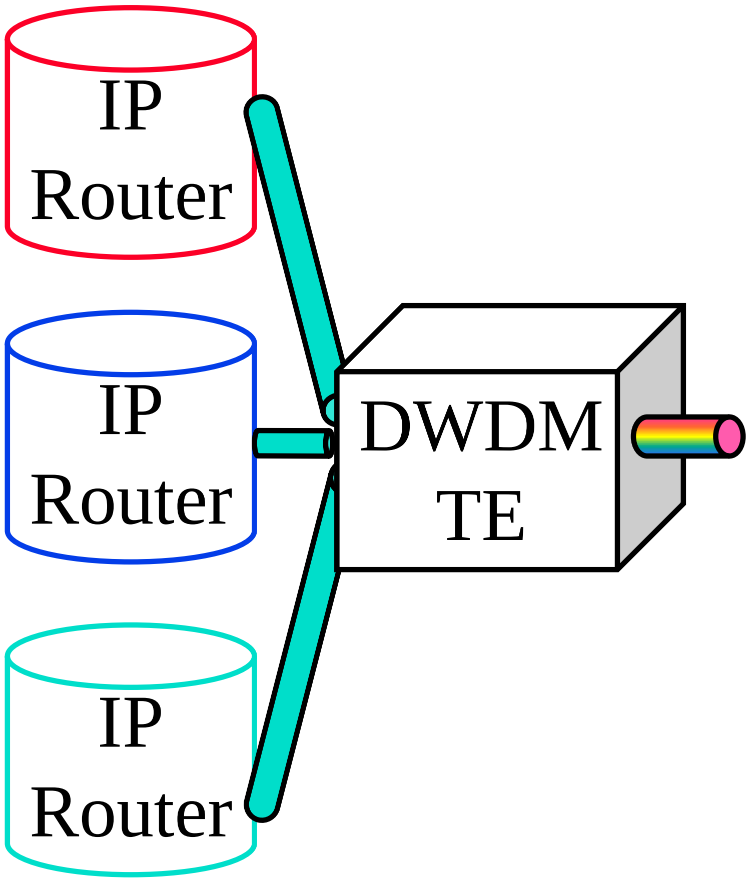
Raj Jain

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
  $\Rightarrow$ Intersection of features and union of problems

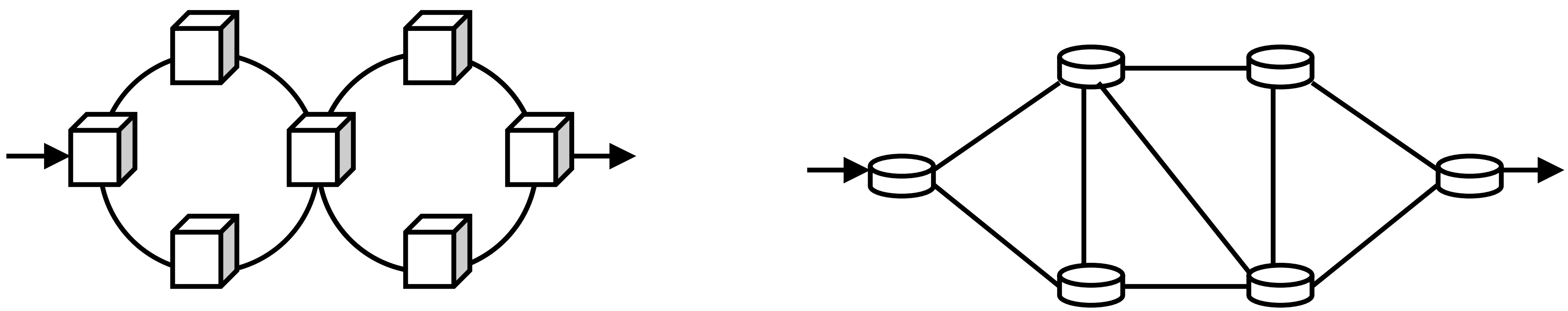
IP over DWDM (Future)



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



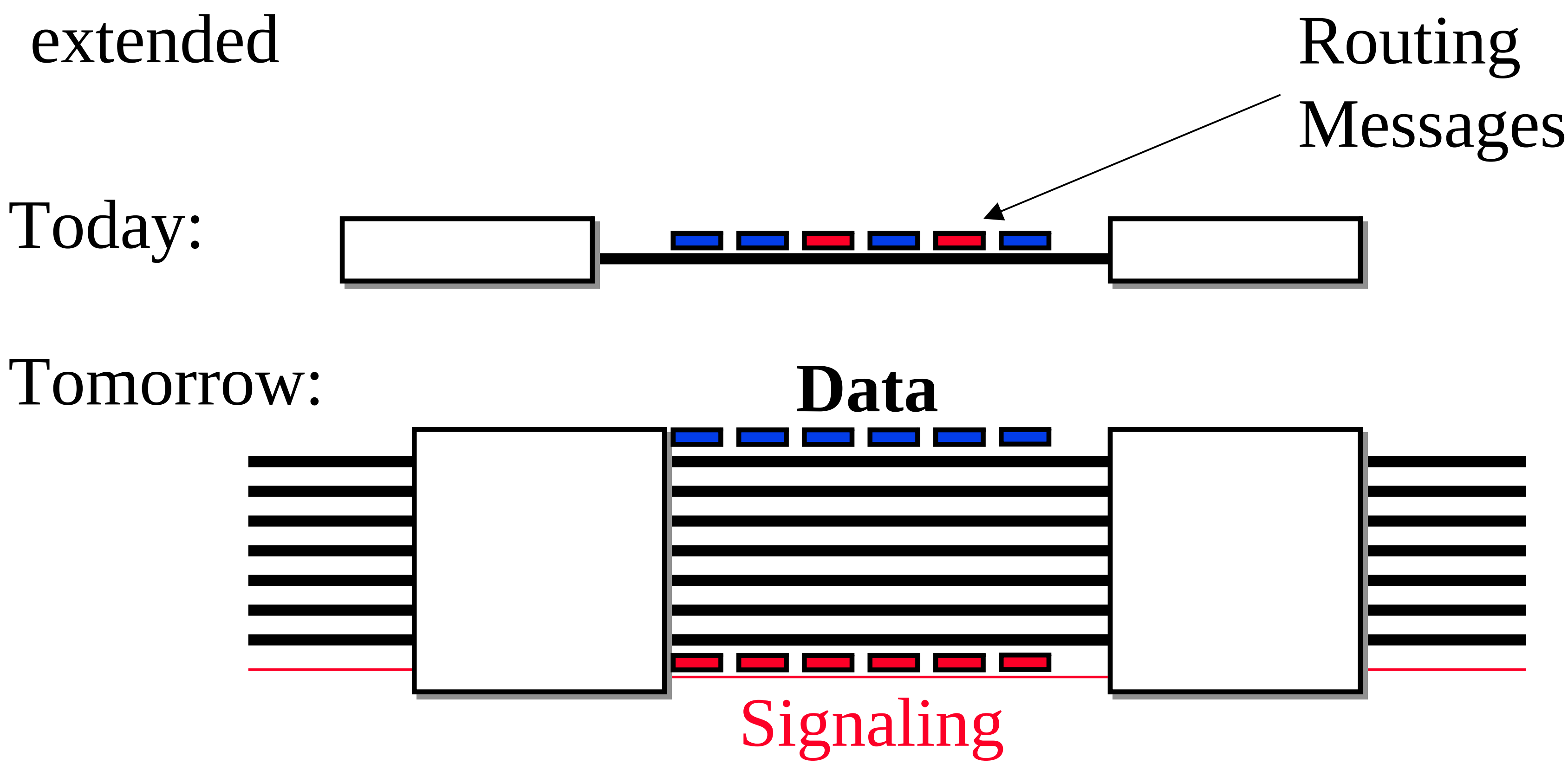
Raj Jain

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

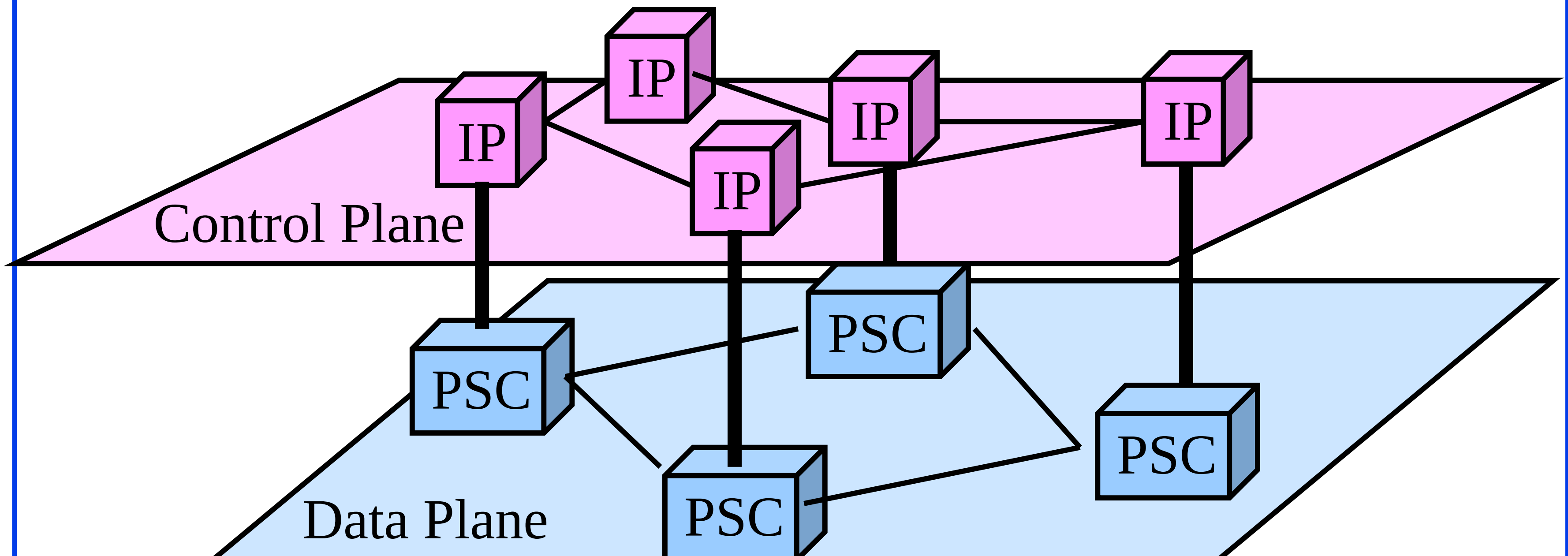
- Separate control and data channels
- IP routing protocols (OSPF and IS-IS) are being extended



Raj Jain

IP-Based Control Plane

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

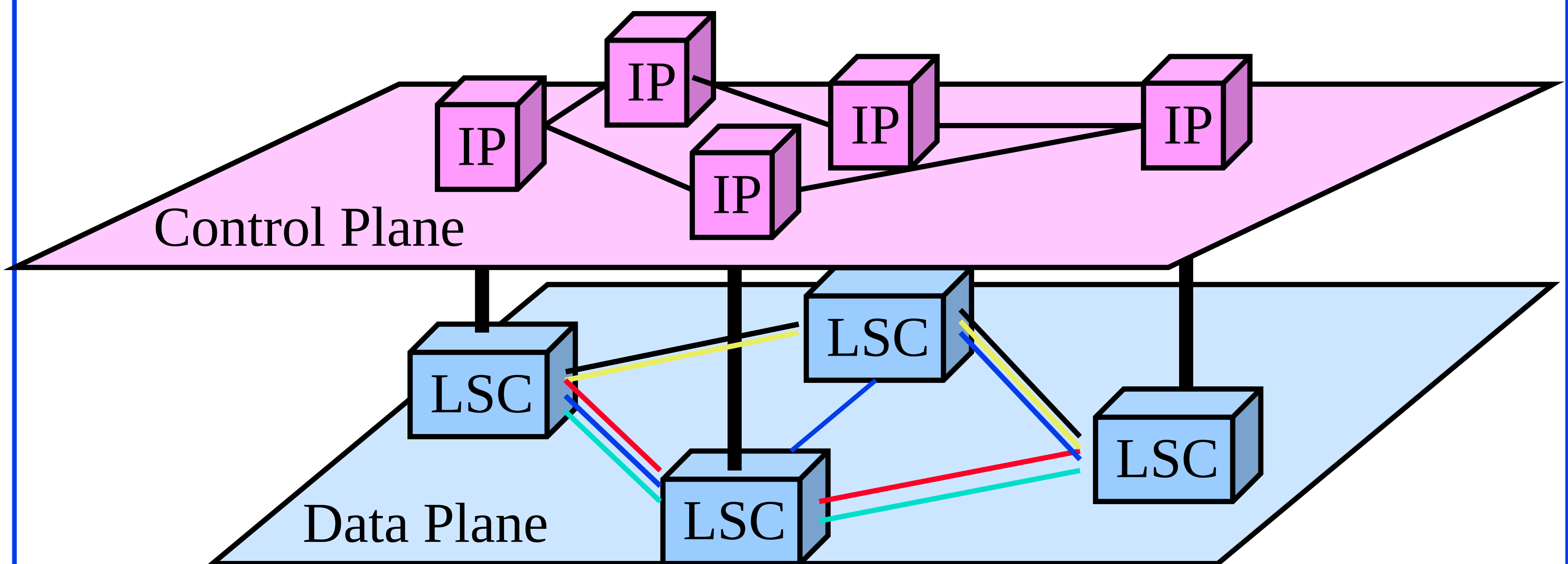


PSC = Packet Switch Capable Nodes

Raj Jain

MPλS

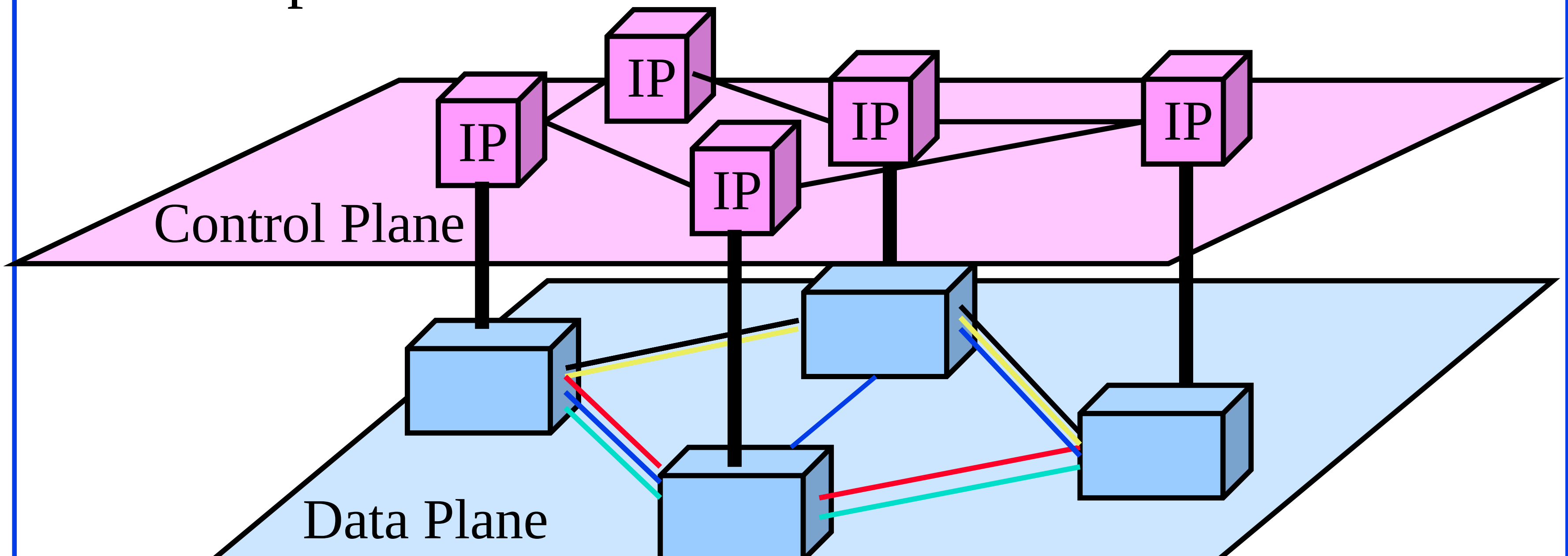
- Control is by IP packets (electronic).
  Data plane consists of wavelength circuits
  ⇒ Multiprotocol Lambda Switching (October 1999)



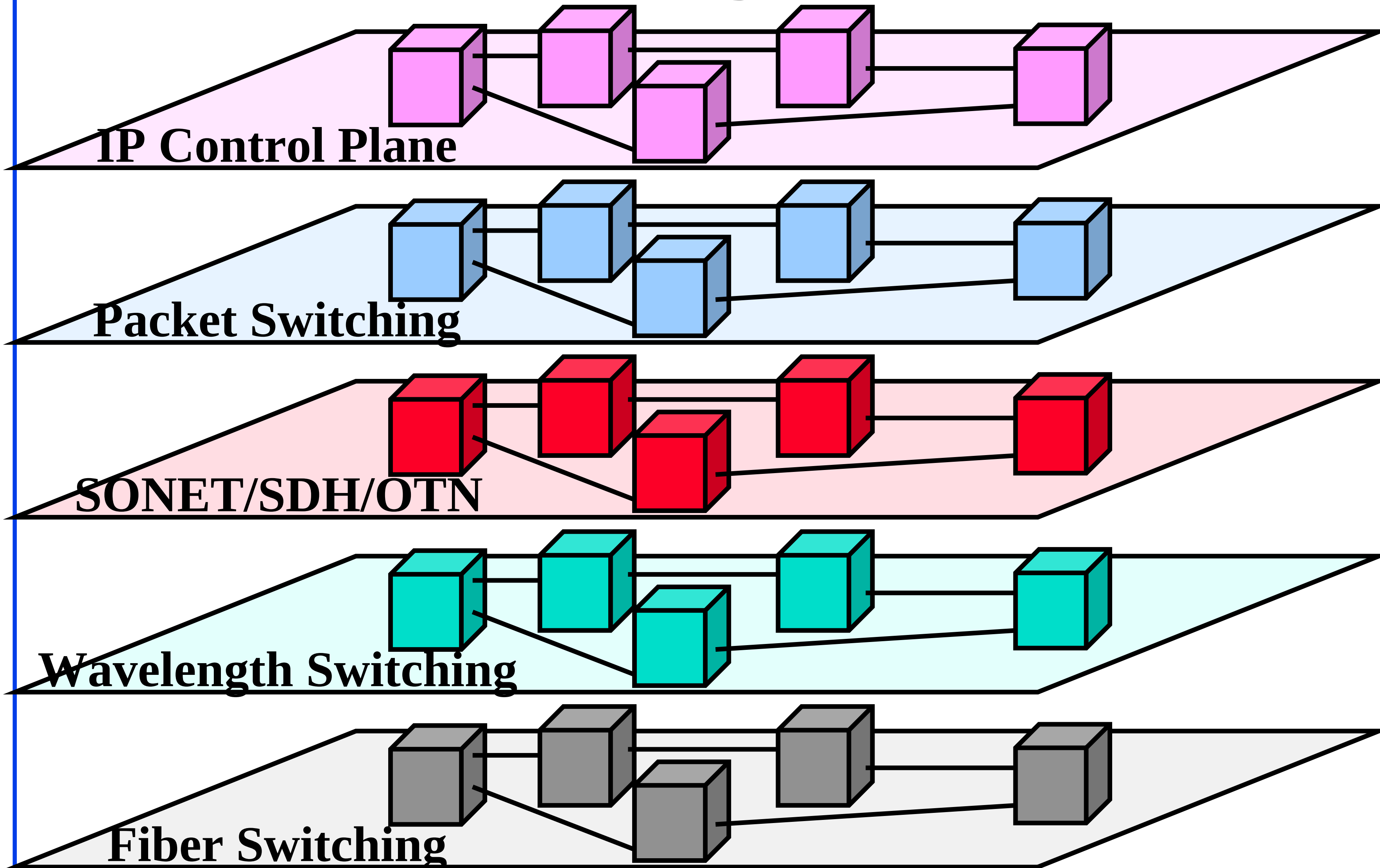
LSC = Lambda Switch Capable Nodes
= Optical Cross Connects = OXC

GMPLS

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



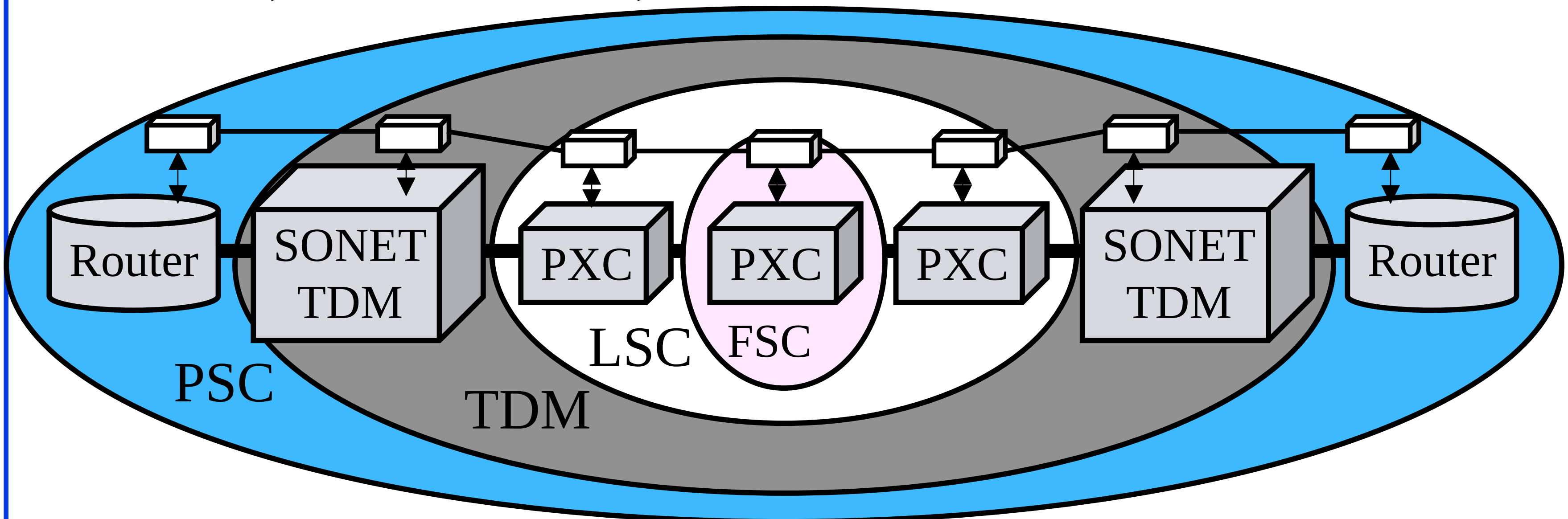
GMPLS: Layered View



Raj Jain

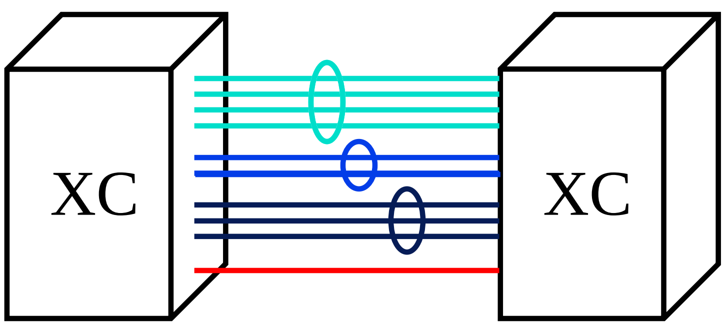
GMPLS: Hierarchical View

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



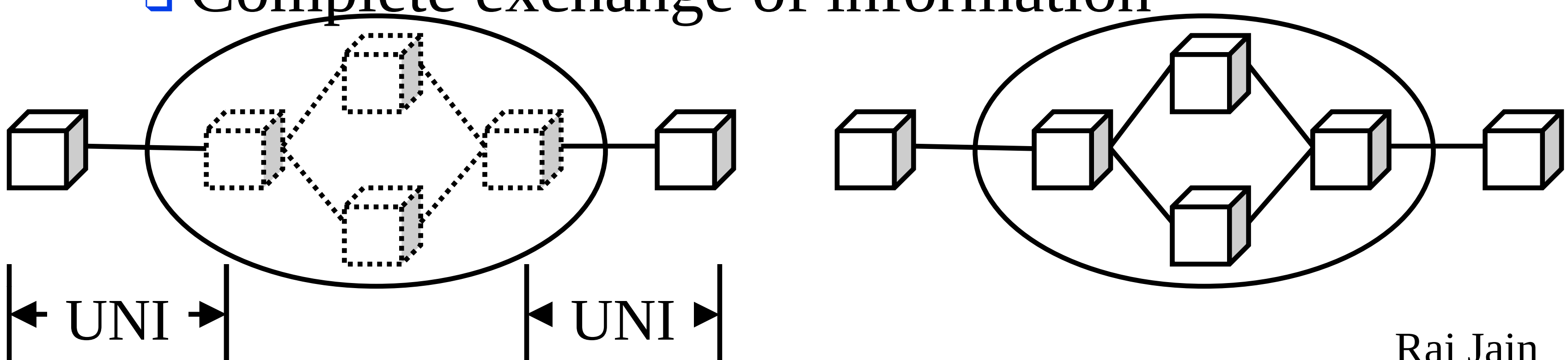
MPLS vs GMPLS

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



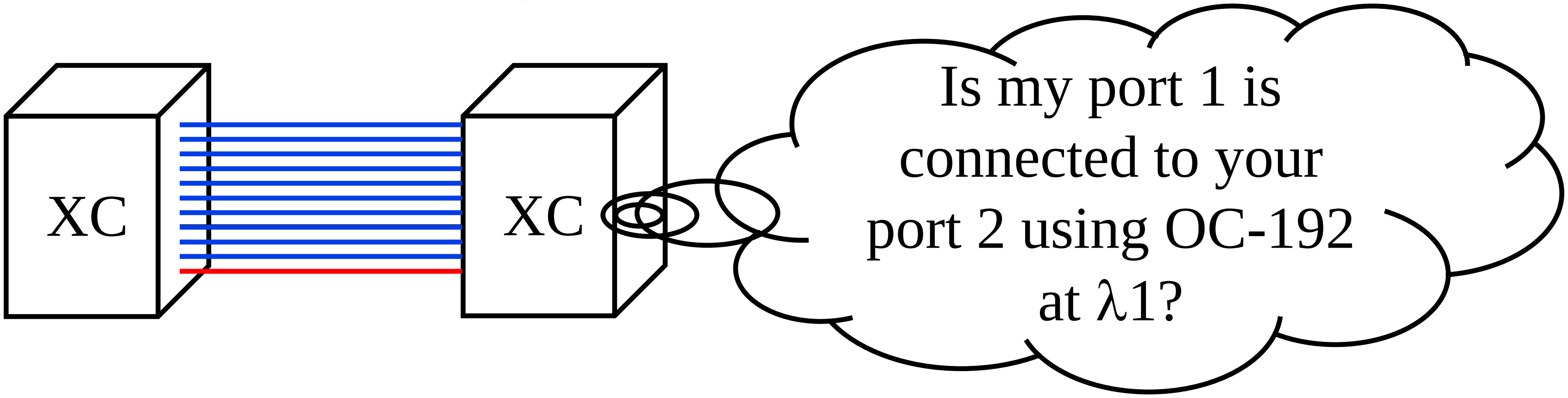
Issue: UNI vs Peer-to-Peer Signaling

- 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



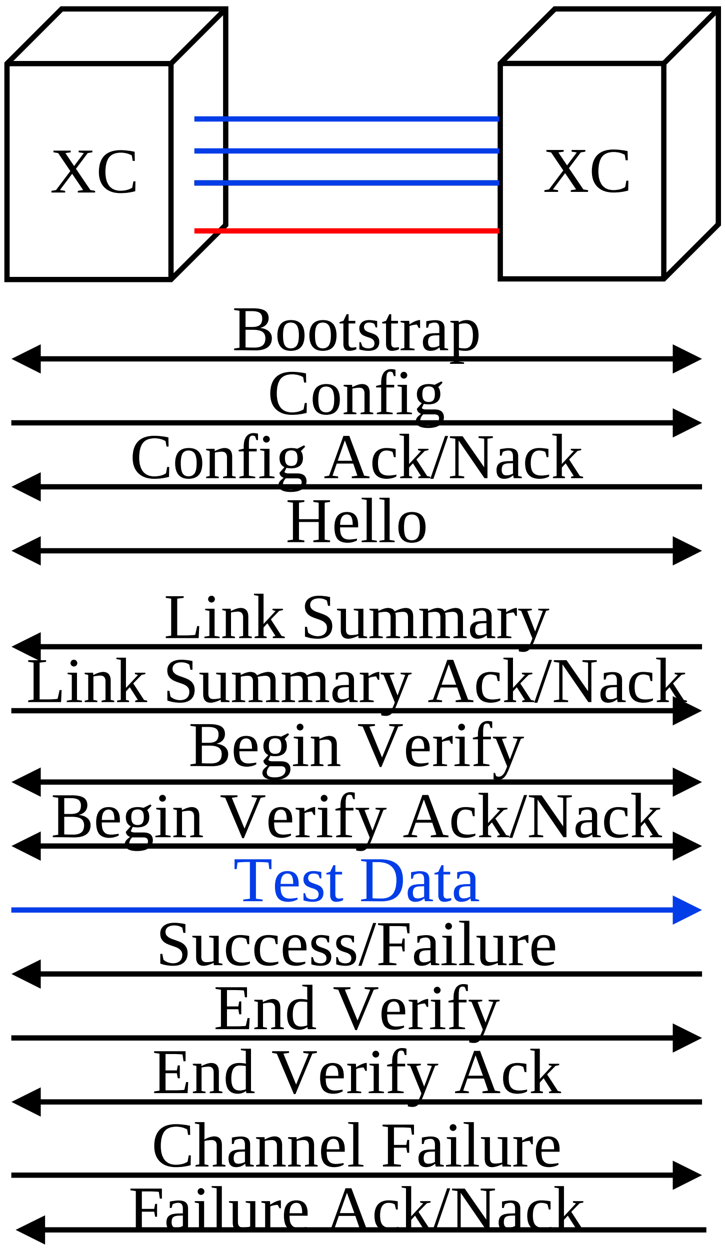
Raj Jain

Link Management Protocol (LMP)



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

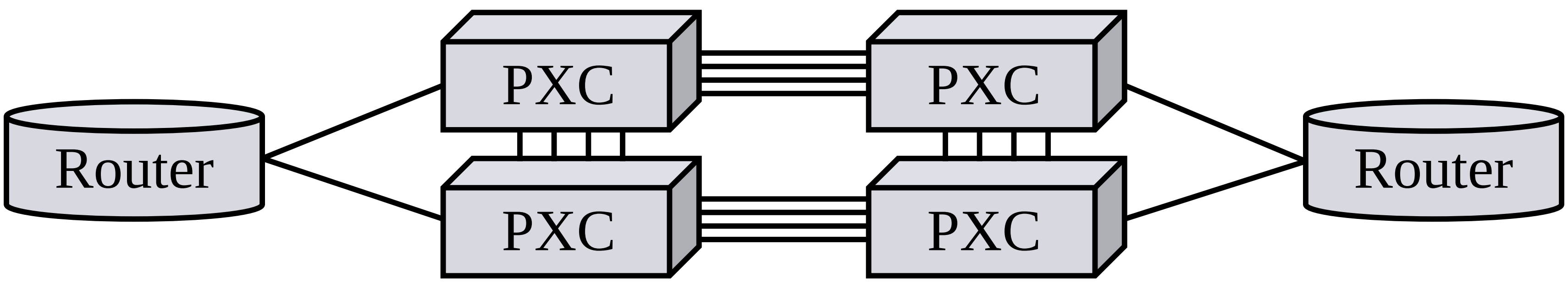
LMP Messages



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

Protection and Restoration

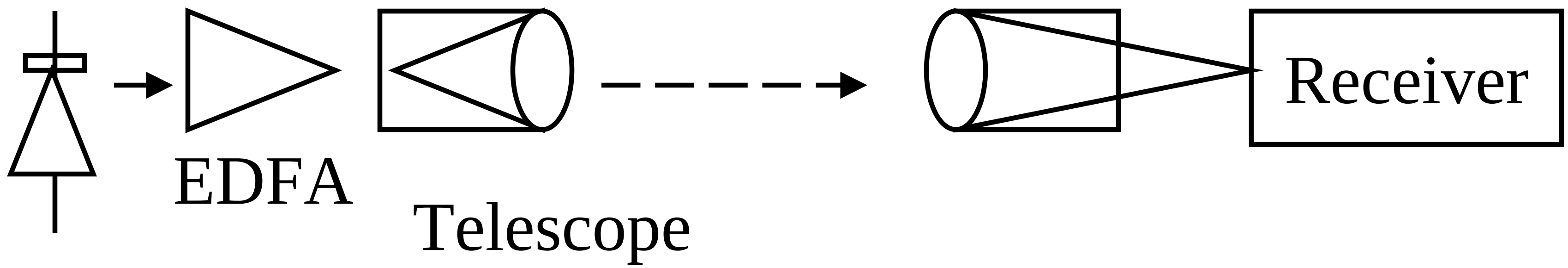
- Extent: Span vs Path
- Topology: Ring vs Mesh
- Redundancy: 1+1, 1:1
- Finding Paths that do not share the same risk
  Each link has to be assigned a risk group
  Shared Risk Group (**SRG**) = All paths sharing a risk



Upcoming Technologies

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

Free Space Optical Comm



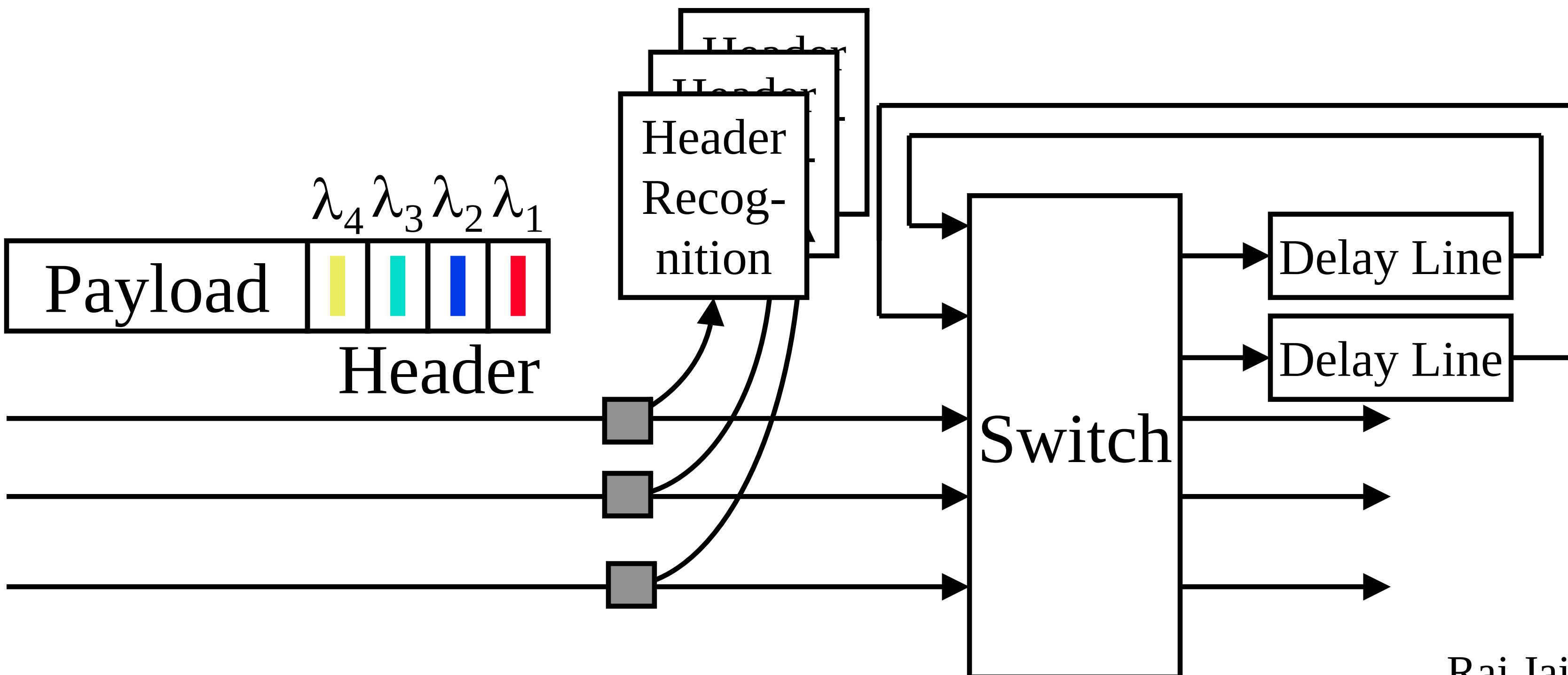
- Uses WDM in open air
- Sample Product:
  Lucent WaveStar OpticAir: 4×2.5Gbps to 5 km
  Available March'00.
- EDFA = Erbium Doped Fiber Amplifier

Free Space Optical Comm

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

Optical Packet Switching

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

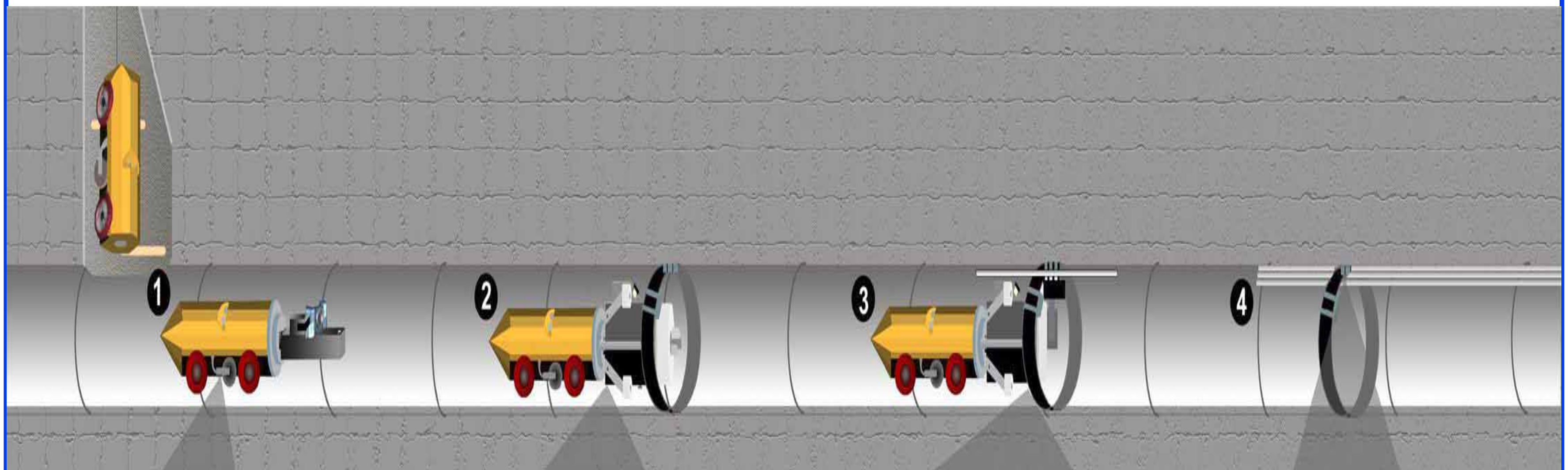


Raj Jain

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

Summary

1. Large number of wavelengths per fiber
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

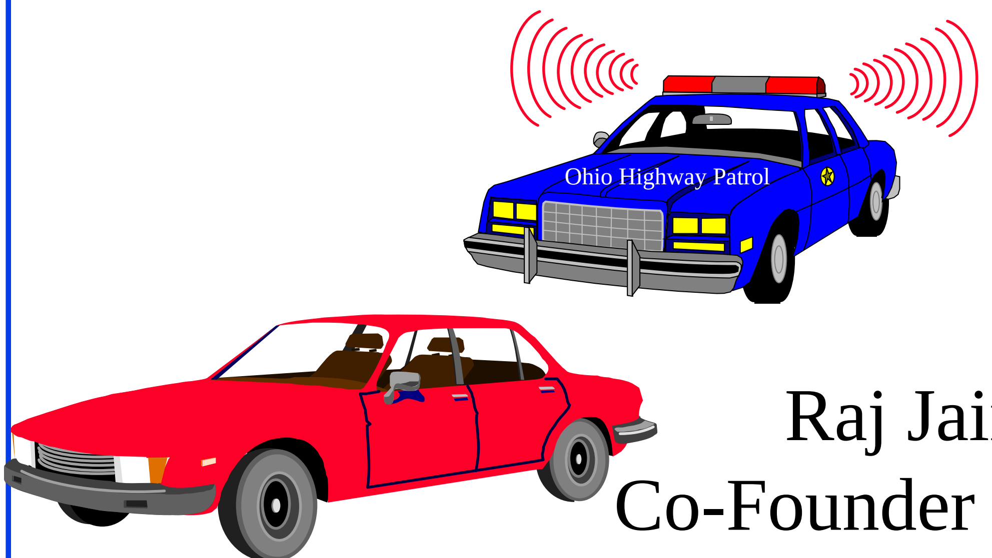
IP over DWDM: Key References

- 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

Wireless Data Networks

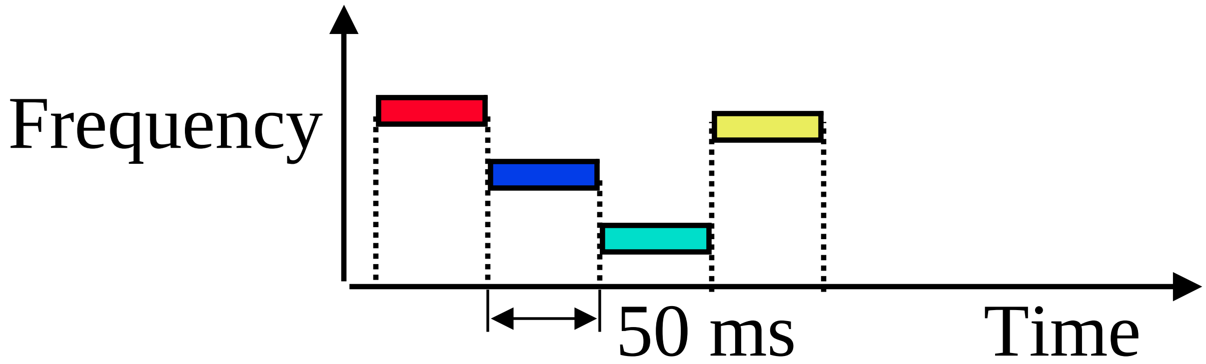


Raj Jain
Co-Founder and CTO
Nayna Networks
481 Sycamore Dr, Milpitas, CA 95035
www.nayna.com and
http://www.cis.ohio-state.edu/~jain/

- Spread Spectrum, Frequency Hopping, Direct-Sequence, OFDM
- IEEE 802.11, 11b, 11a, 11g LANs
- HiperLAN, HiperLAN2
- PANs: IrDA, Bluetooth, HomeRF
- WAP, WML

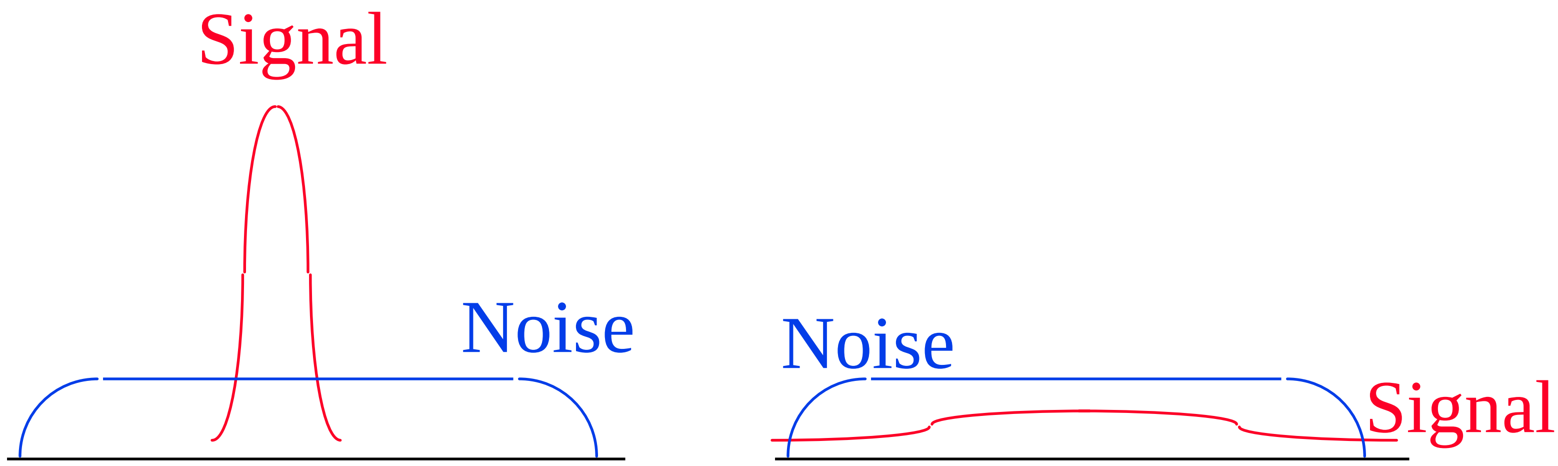
Note: wireless phone services and standards not covered.

Frequency Hopping Spread Spectrum



- Pseudo-random frequency hopping
- Spreads the power over a wide spectrum
  $\Rightarrow$ Spread Spectrum
- Developed initially for military
- Patented by actress Hedy Lamarr
- Narrowband interference can't jam

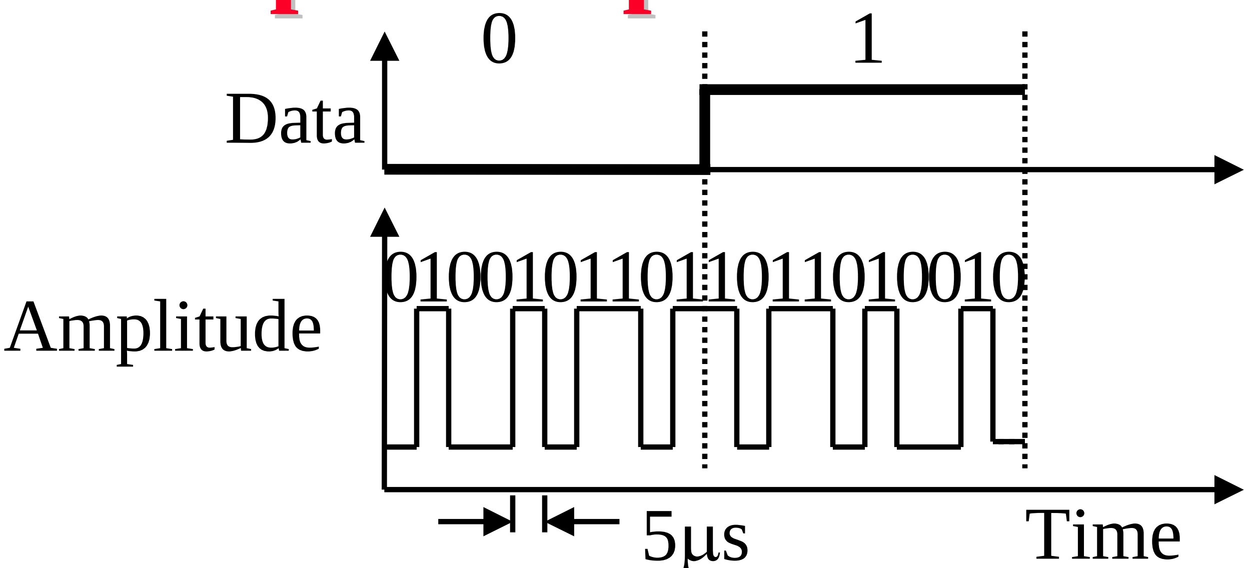
Spectrum



(a) Normal (b) Frequency Hopping

Raj Jain

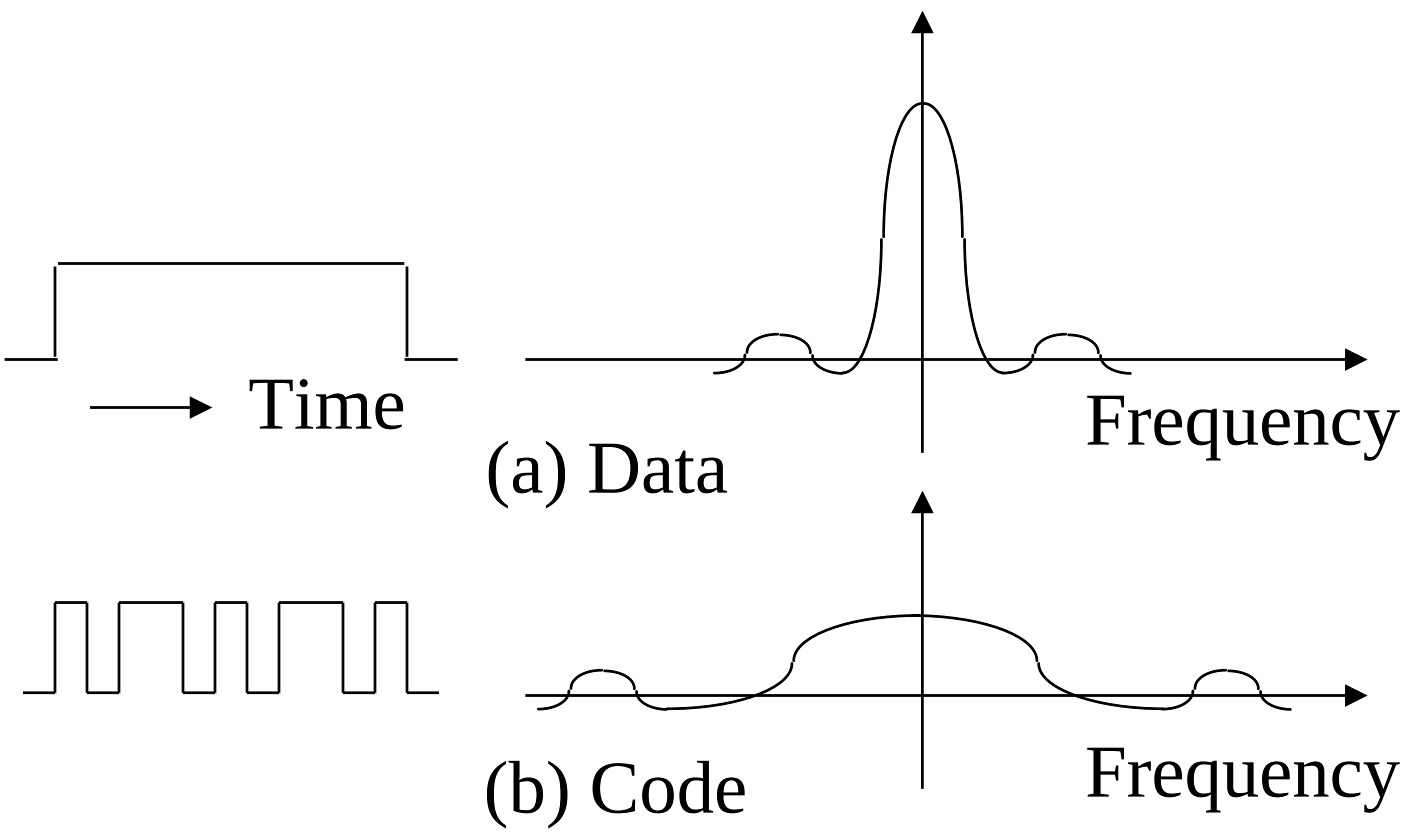
Direct-Sequence Spread Spectrum



- Spreading factor = Code bits/data bit, 10-100 commercial (Min 10 by FCC), 10,000 for military
- Signal bandwidth >10 × data bandwidth
- Code sequence synchronization
- Correlation between codes ⇒Interference⇒ Orthogonal

DS Spectrum

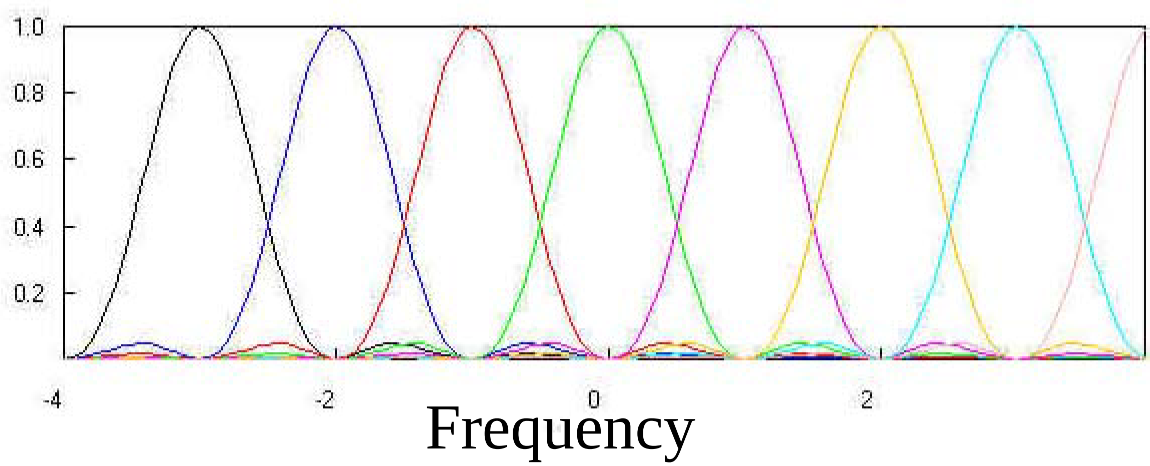
Time Domain   Frequency Domain



Raj Jain

OFDM

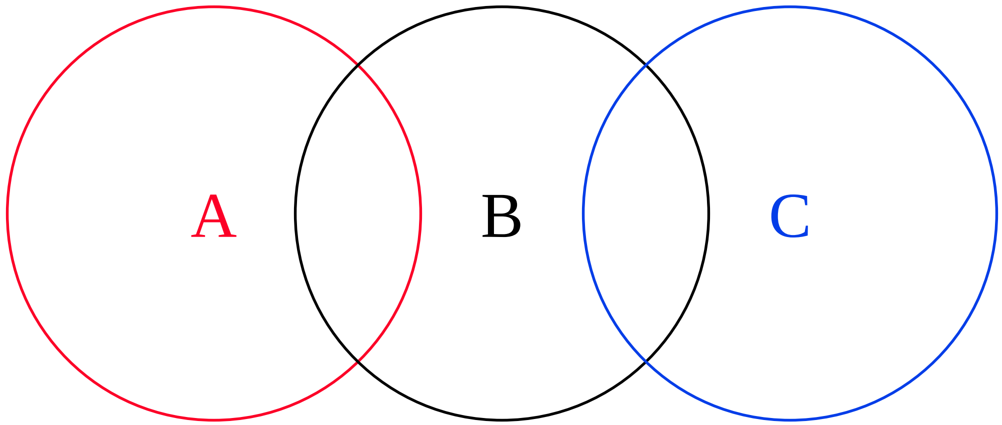
- Orthogonal Frequency Division Multiplexing (OFDM)
- Multi-Carrier Transmission
- Frequencies are selected such that all other signals are zero at the peak



IEEE 802.11 Features

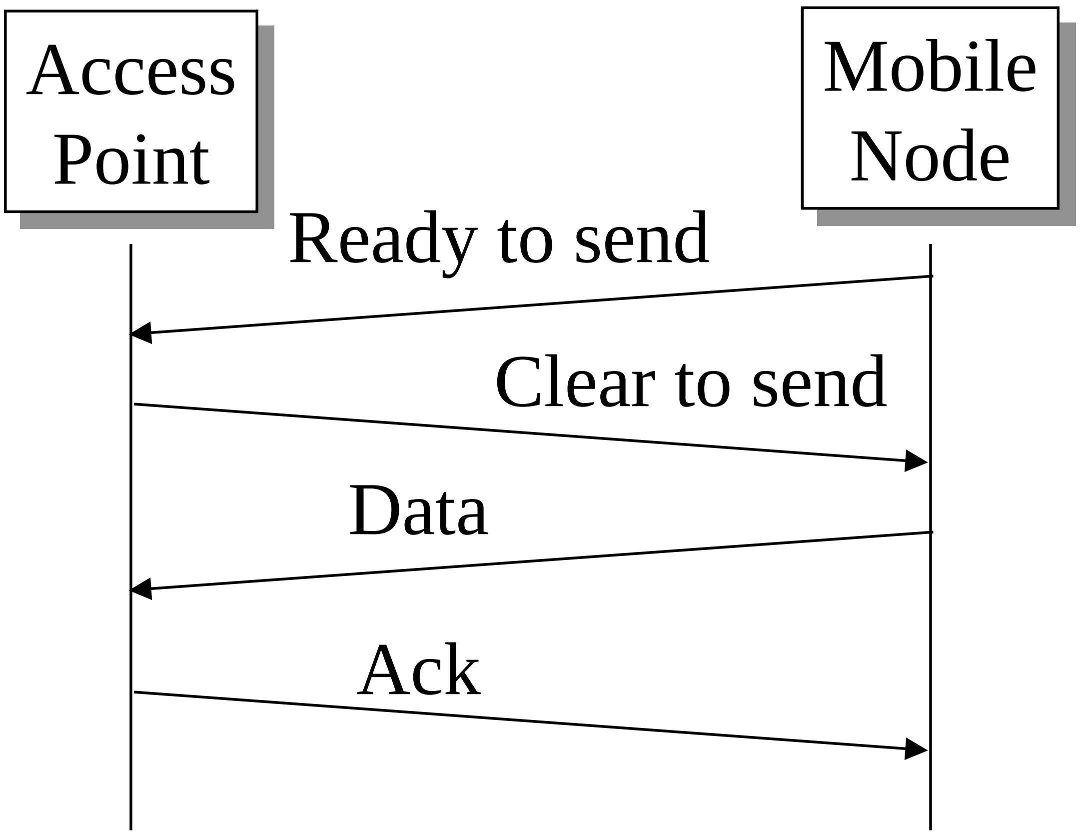
- 1 and 2 Mbps
- Supports both Ad-hoc and base-stations
- Spread Spectrum ⇒ No licensing required.
  Three Phys: Direct Sequence, Frequency Hopping, 915-MHz, **2.4 GHz** (Worldwide ISM), 5.2 GHz, and Diffused Infrared (850-900 nm) bands.
- Supports multiple priorities
- Supports time-critical and data traffic
- Power management allows a node to doze off

Hidden Node Problem



- C cannot hear A.
  It may start transmitting while A is also transmitting
  $\Rightarrow$ A and C can't detect collision.
- Only the receiver can help avoid collisions

4-Way Handshake



Raj Jain

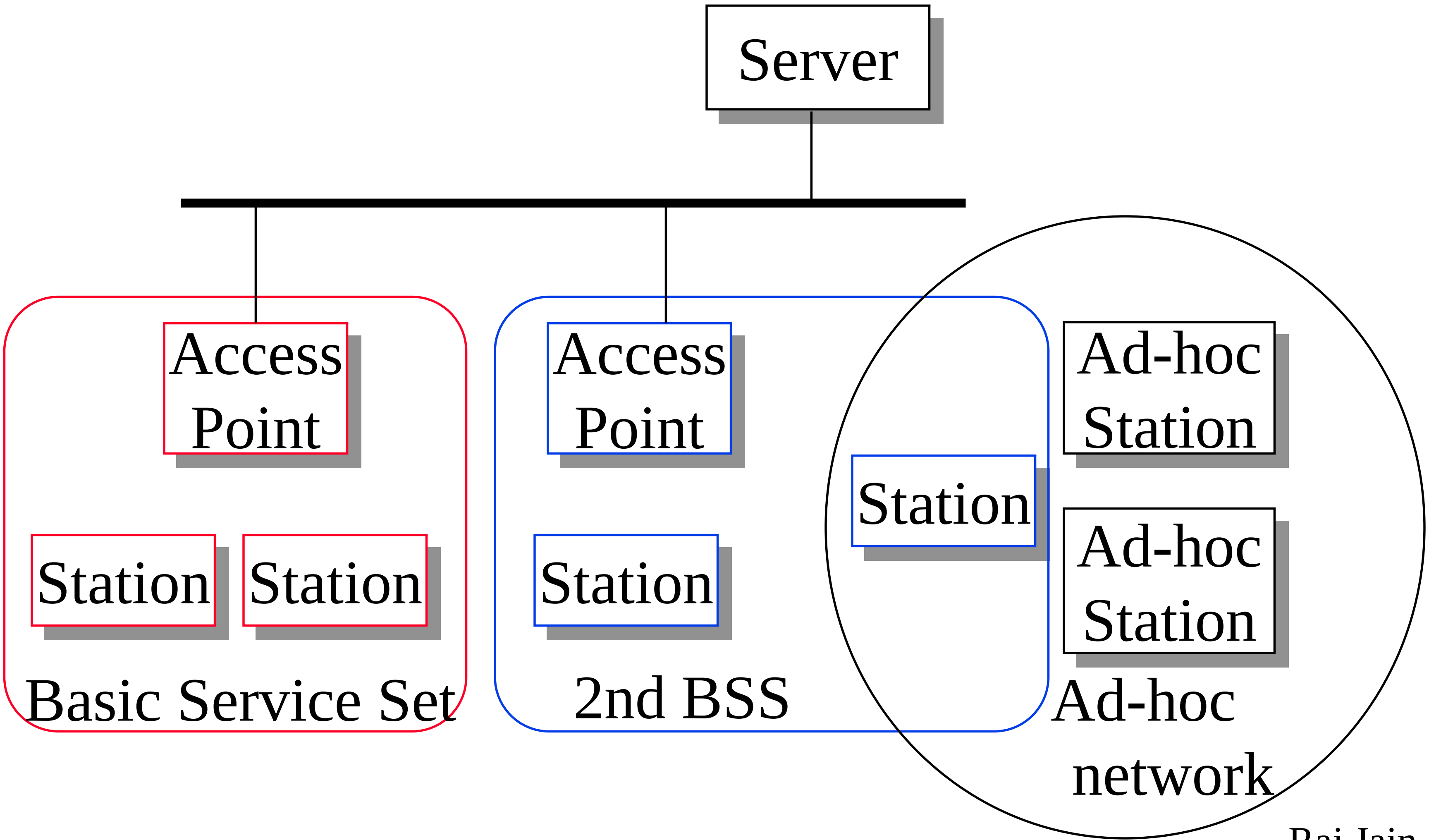
IEEE 802.11 MAC

- Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA)
- Listen before you talk. If the medium is busy, the transmitter backs off for a random period.
- Avoids collision by sending a short message:
  Ready to send (RTS)
  RTS contains dest. address and duration of message.
  Tells everyone to backoff for the duration.
- Destination sends: Clear to send (CTS)
- Can not detect collision $\Rightarrow$ Each packet is acked.
- MAC level retransmission if not acked.

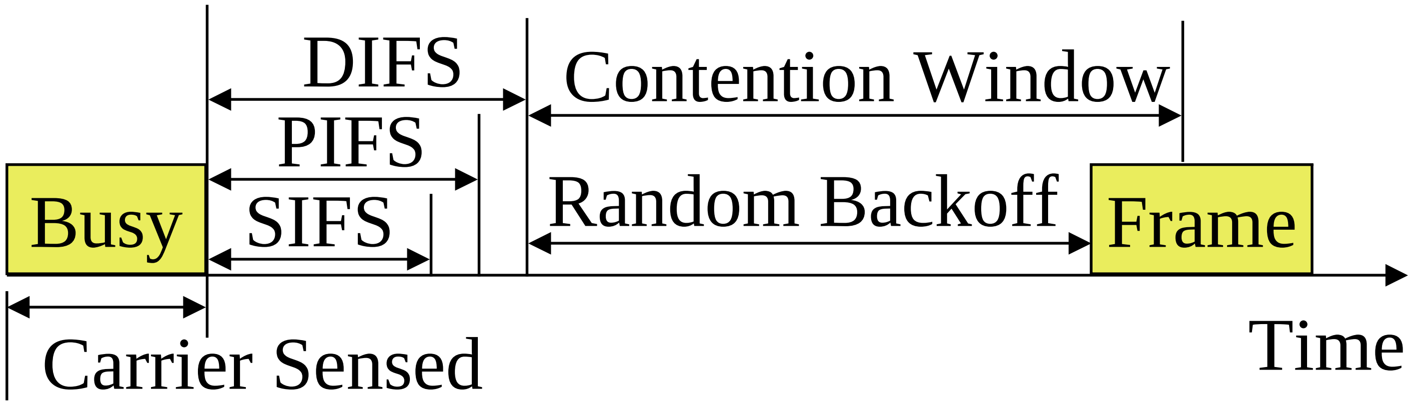
Peer-to-Peer or Base Stations?

- Ad-hoc (Autonomous) Group:
  - Two stations can communicate
  - All stations have the same logic
  - No infrastructure, Suitable for small area
- Infrastructure Based: Access points (base units)
  - Stations can be simpler than bases.
  - Base provide connection for off-network traffic
  - Base provides location tracking, directory, authentication ⇒ Scalable to large networks
- IEEE 802.11 provides both.

IEEE 802.11 Architecture

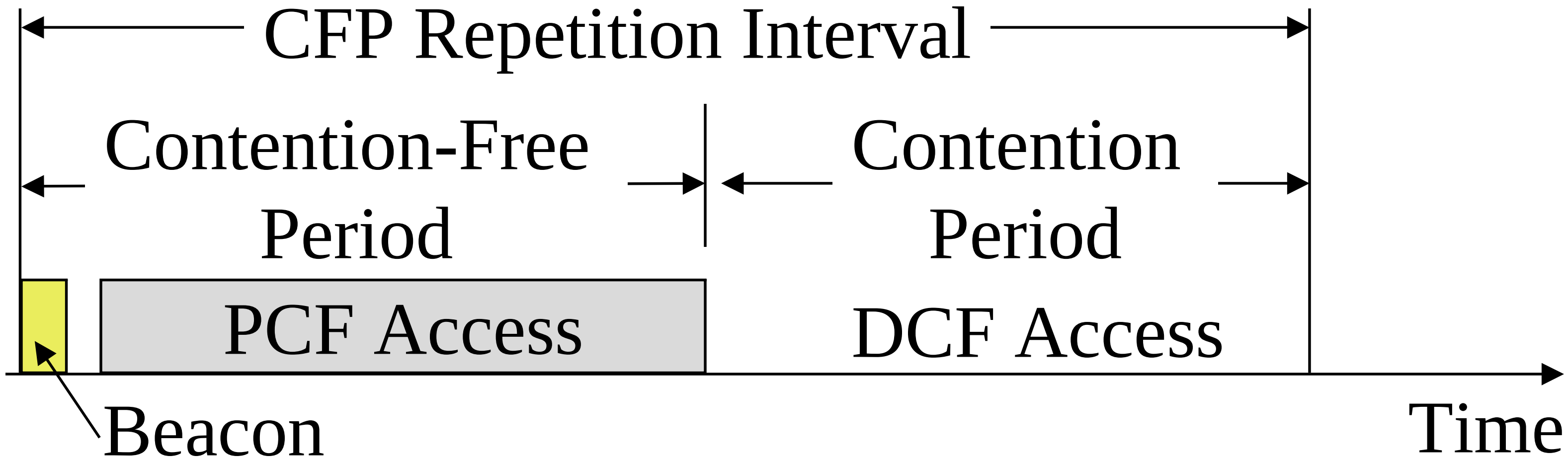


IEEE 802.11 Priorities



- Initial interframe space (IFS)
- Highest priority frames, e.g., Acks, use short IFS (SIFS)
- Medium priority time-critical frames use "Point Coordination Function IFS" (PIFS)
- Asynchronous data frames use "Distributed coordination function IFS" (DIFS)

Time Critical Services



- Timer critical services use Point Coordination Function
- The point coordinator allows only one station to access
- Coordinator sends a beacon frame to all stations. Then uses a polling frame to allow a particular station to have contention-free access
- Contention Free Period (CFP) varies with the load.

Power Management

- A station can be in one of three states:
  - Transmitter on
  - Receiver only on
  - Dozing: Both transmitter and receivers off.
- Access point (AP) buffers traffic for dozing stations.
- AP announces which stations have frames buffered. Traffic indication map included in each beacon. All multicasts/broadcasts are buffered.
- Dozing stations wake up to listen to the beacon. If there is data waiting for it, the station sends a poll frame to get the data.

Status and Future

- 802.11 including both MAC and PHY approved June 1997.
- More bandwidth in future by:
  1. Better encoding: Multilevel modulation $\Rightarrow$ 8 Mbps
  2. Fewer channels with more bandwidth $\Rightarrow$ 4 MHz channels. Or Entire ISM band for one channel.
  3. Find another band. May get 150 MHz band in 5-GHz band. Fifteen 10-MHz channels with 15-20 Mb/s.

802.11 Versions

Version	Frequency Band	Data Rate	Coding
802.11	2.4 GHz	1 and 2 Mbps	DPBSK, DQPSK
802.11b (Wi-Fi)	2.4 GHz	1, 2, 5.5, 11 Mbps	CCK with DPBSK, DQPSK
802.11a (Coming soon)	5 GHz	6, 9, 12, 18, 24, 36, 48, 54 Mbps	OFDM w BPSK/QPSK/ QAM-64
802.11g (2H02)	2.4 GHz	54 Mbps	CCK+OFDM+ PBCC

- Wi-Fi ⇒ 802.11b products certified to interoperate (see www.wi-fi.com)

IEEE 802.11 Coding/Modulation

- Differential Binary Phase Shift Keying (DBPSK)
- Differential Quadrature Phase Shift Keying (DQPSK)
- Orthogonal Frequency Division Multiplexing (OFDM)
- Complementary Code Keying (CCK) modulation
- Binary Phase Shift Keying (BPSK)
- Quadrature Phase Shift Keying (QPSK)
- Quadrature Amplitude Modulation (QAM)
- Packet Binary Convolutional coding (PBCC)

HiperLAN

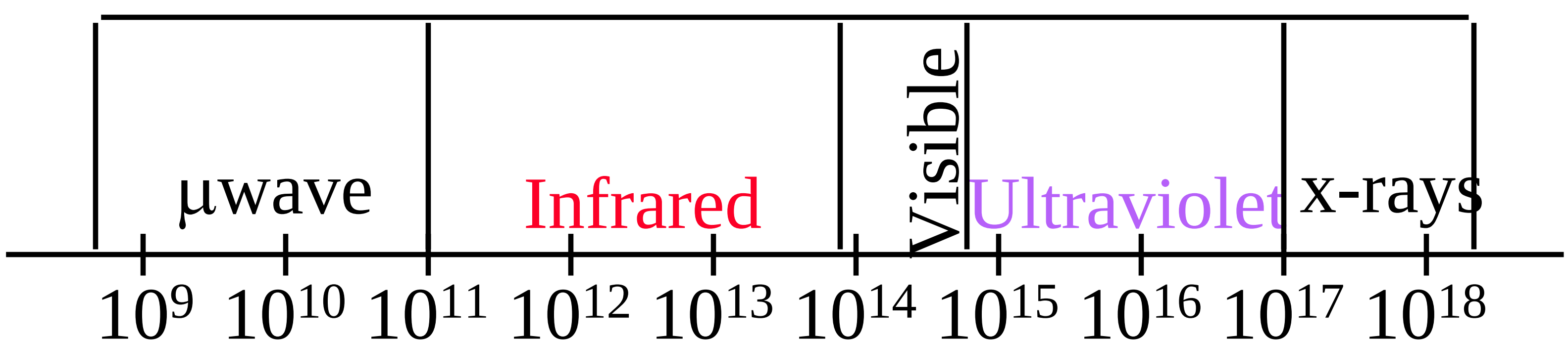
- High PErformance Radio Local Area Network
- European Technical Standards Inst (ETSI) Standard
- Short range, very high data rate
- 5.15 to 5.25 GHz, 5 Channels, 23.5 Mbps in air
- Output power 10mW, 100mW, 1W
- Both ad-hoc and infrastructure.
- Data and Real-time video
- Power management: Nodes announce their wakeup cycle. Other nodes send according to the cycle. A low-bit rate header allows nodes to keep most ckts off.

HiperLAN2

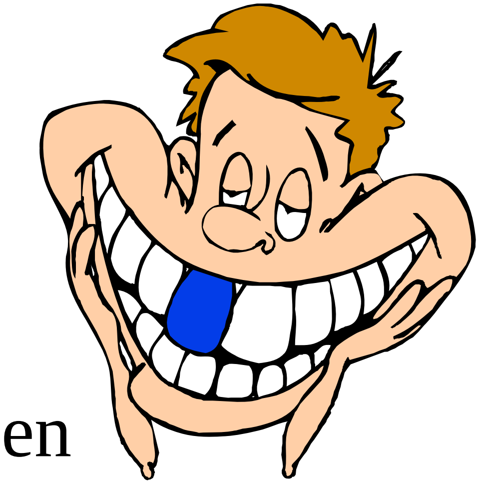
- Similar to 802.11a
- 54 Mbps in 5 GHz band
- Dual standard devices possible: HiperLAN2 and 802.11a
- Panasonic demonstrated HiperLAN2 at CeBIT2002

Infrared Data Association (IrDA)

- IR ⇒ Line of sight, short range, indoors
- Up to 1-2m
- Bi-directional
- 9.6 kbps to to 4 Mbps in steps of 115 kbps
- CRC-16 up to 1.152 Mbps, CRC-32 at higher speeds



Bluetooth



- Named after Harald Bluetooth
  - King of Denmark around 980AD.
- Initially designed to avoid cables between neighboring devices
- Electronic payments: Toll, Movie Tickets
- Local information exchange in public areas
- Personal area network or piconet
- 2.4 GHz ISM Band, ad-hoc
- Omni-directional, non-line of sight through walls
- IEEE 802.15.1-2002 licensed Bluetooth V1.1 from Bluetooth SIG, Inc.⇒ IEEE 802 LLC compatibility

Bluetooth (Cont)

- ISM Band is also used by baby monitors, garage-door openers, and some cordless phones
- 1 Mbps total. 20% overhead
  $\Rightarrow$ 800 kbps in two directions
  $\Rightarrow$ 721 kbps+57.6 kbps or 432.6 kbps full-duplex
- Frequency hopping spread spectrum.
  1600 times per sec.
- 1 mW Power $\Rightarrow$ 10m range.
  With external amplifier to 100m.
- All electronics in one IC
- 128-bit encryption and authentication

HomeRF

- Home Radio Frequency Working Group (HomeRF)
- HomeRF V1.0 at 1.2 Mbps to 40m
- HomeRF V2.0 at 10 Mbps
- 2.45 GHz ISM band
- 50 hops/s FHSS
- Data and Voice
- Shared Wireless Access Protocol (SWAP)
  = CSMA/CA similar to 802.11
- 127 nodes Infrastructure or Ad-Hoc
- 100mW
- Blowfish Encryption

WAP

- Wireless Application Protocol (WAP)
- Developed jointly by Ericsson, Nokia, Motorola, and Phone.com
- For Web access through small displays in cell phones
- WAP gateways can reduce normal web pages by looking for WML markups
- Uses Header compressing and UDP to reduce bandwidth

WML

Dept of CIS
[Faculty]
[Research]
[Admission]

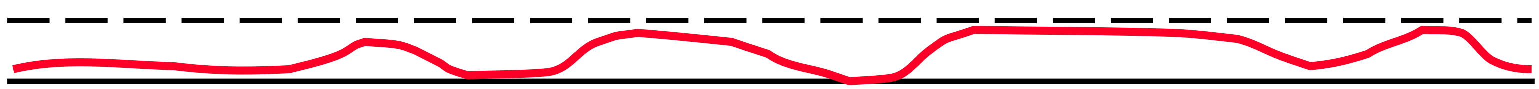
- Wireless Markup Language (WML)
- Deck of Cards

```
<card id="OSU" title="Raj Jain" ontimer="#C">
<Timer value="15 />
<P><img src="http://www.cis.ohio-
   state.edu/wap/OSUlogo.wbsp" alt="OSU Logo">
</card>
<card id="C" title="Dept of CIS">
<P>
<a title="Faculty" href="#F">Faculty</a><BR/>
<a title="Research" href="#R">Research</a><BR/>
<a title="Admission" href="#A">Admission</a><BR/>
</card>
```

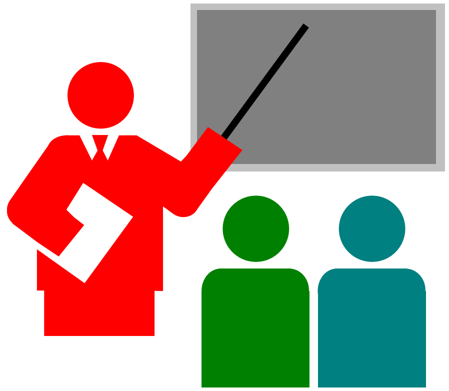
- See WAP emulator at www.pyweb.com

Future

- Millimeter Wave Transmission:
  - 1 to 10 mm = 30 to 300 GHz
  - FCC has allocated 7 GHz in 60 GHz Band
    ⇒ Wireless GbE
  - Affected by rain and fog
- Ultra wide band transmission:
  Wideband = Spread spectrum over a few MHz
  Ultra wide band = Spread spectrum over a GHz
  ⇒ Power level ≤ FCC Noise limit



Summary

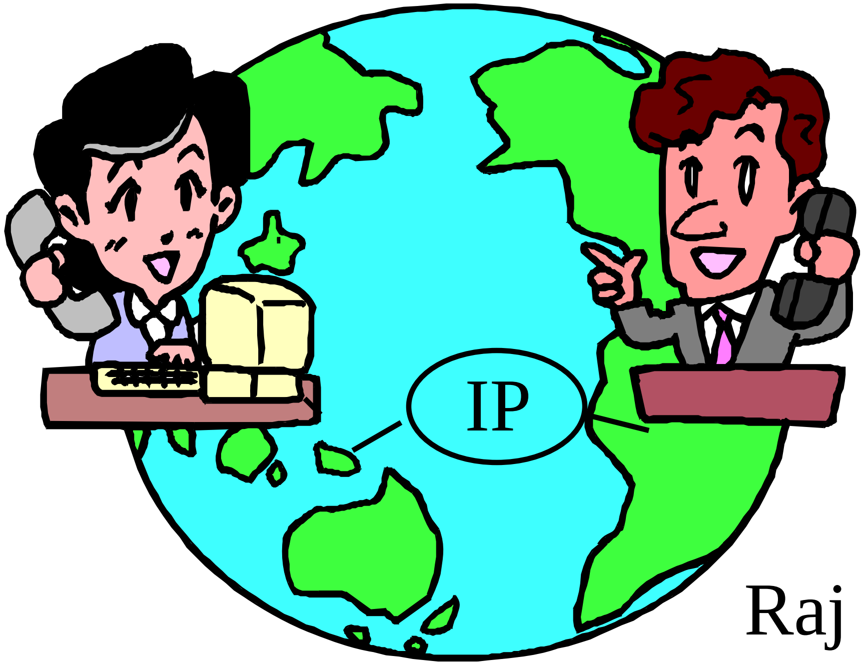


- Spread spectrum allows simultaneous use of spectrum by several users ⇒ Unlicensed operation
- 802.11b (Wi-Fi) LANs run at 11 Mbps.
  11a at 54 Mbps is coming soon.
- HomeRF V2 runs at 10 Mbps to 40m
- Bluetooth allows 800 kbps upto 10 m
- WAP enables web pages on cell phone screens.

Wireless: Key References

- For a detailed list of references see: http://www.cis.ohio-state.edu/~jain/refs/wir_refs.htm
  Also reproduced in the back of this tutorial handout.
- "Wireless local area networks," Aug 97, http://www.cis.ohio-state.edu/~jain/cis788-97/wireless_lans/index.htm
- "In-building wireless LANs," http://www.cis.ohio-state.edu/~jain/cis788-99/wireless_lans/index.Html

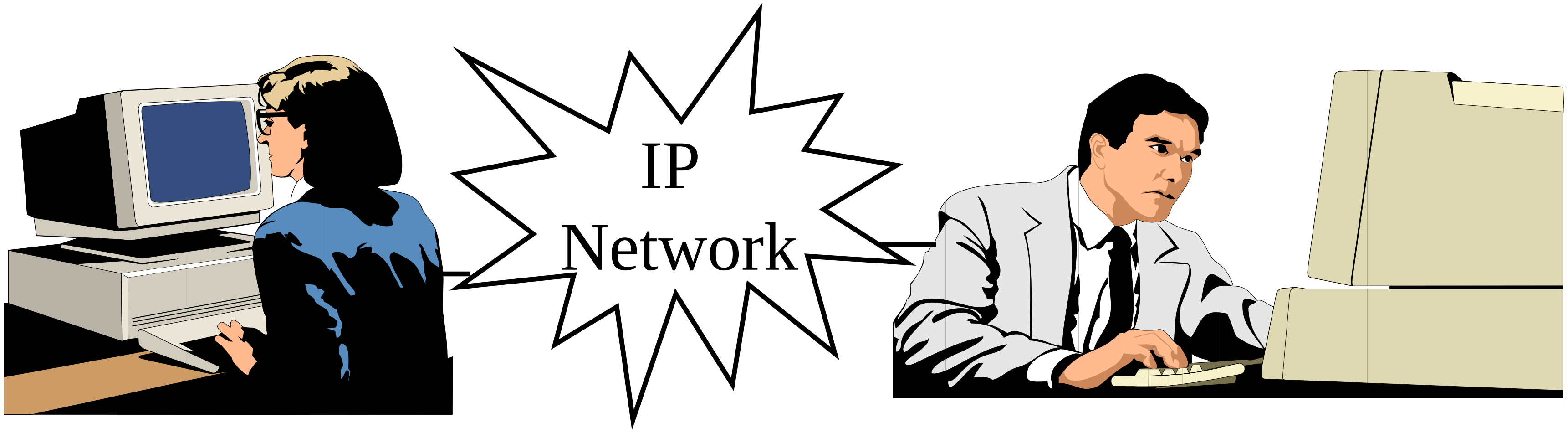
Voice over IP



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/

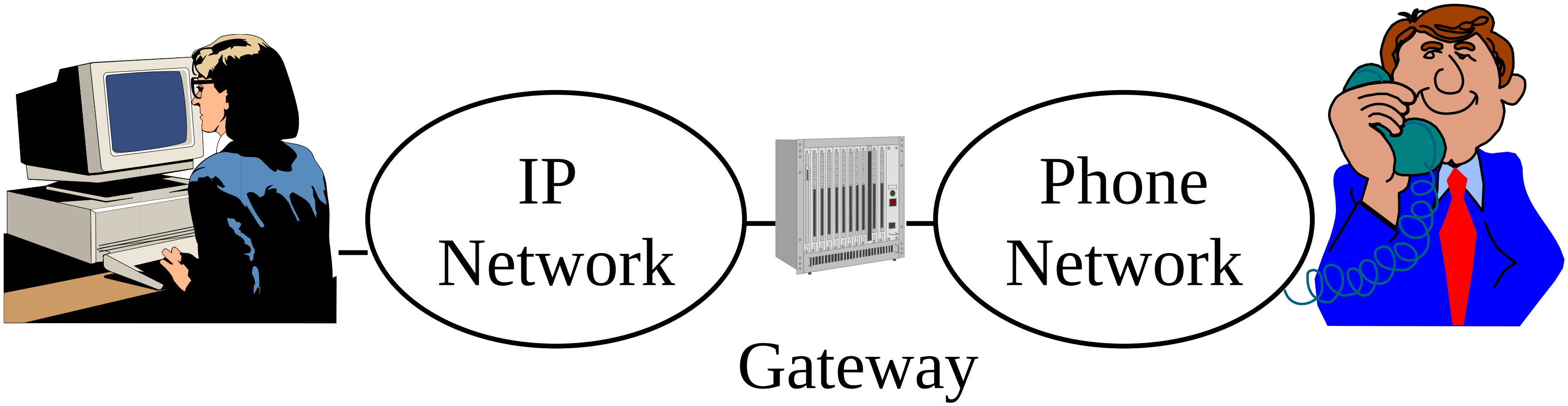
- Voice over IP: Why?
- Sample Products and Services
- 13 Technical Issues, 4 Other Issues
- H.323 Standard and Session Initiation Protocol (SIP)
- Media Gateway Control Protocol (MGCP) and Megaco
- Stream Control Transmission Protocol (SCTP)

Scenario 1: PC to PC



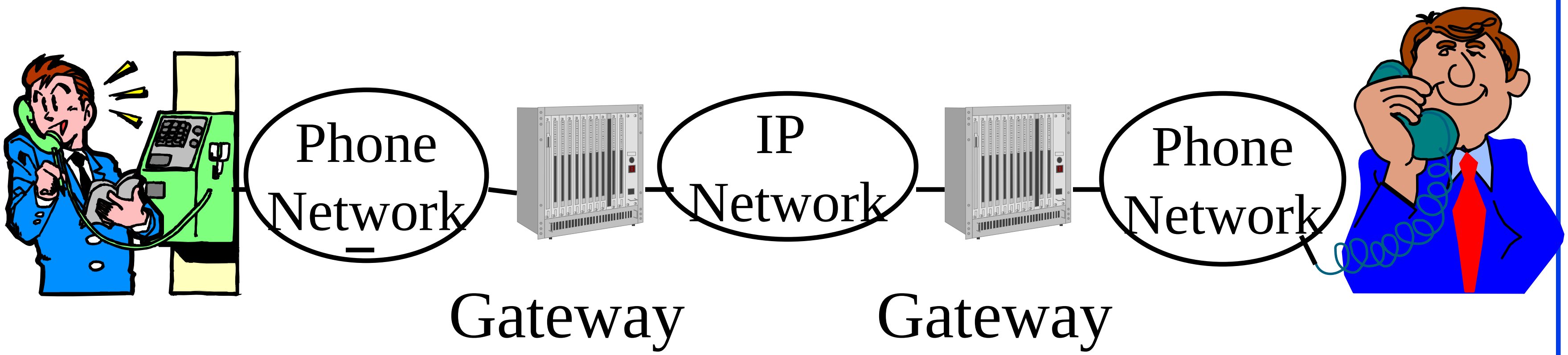
- Need a PC with sound card
- IP Telephony software: Cuseeme, Internet Phone, ...
- Video optional

Scenario 2: PC to Phone



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

Scenario 3: Phone to Phone



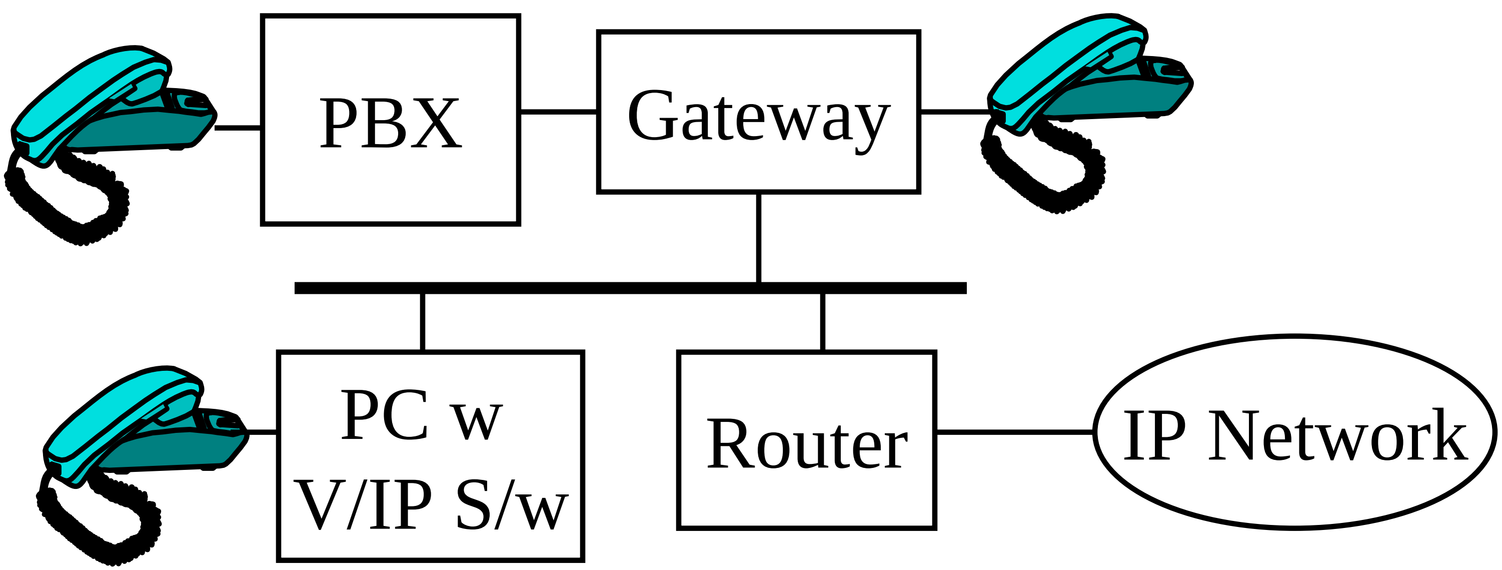
- Need more gateways that connect IP network to phone networks
- The IP network could be dedicated intra-net or the Internet.
- The phone networks could be intra-company PBXs or the carrier switches

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

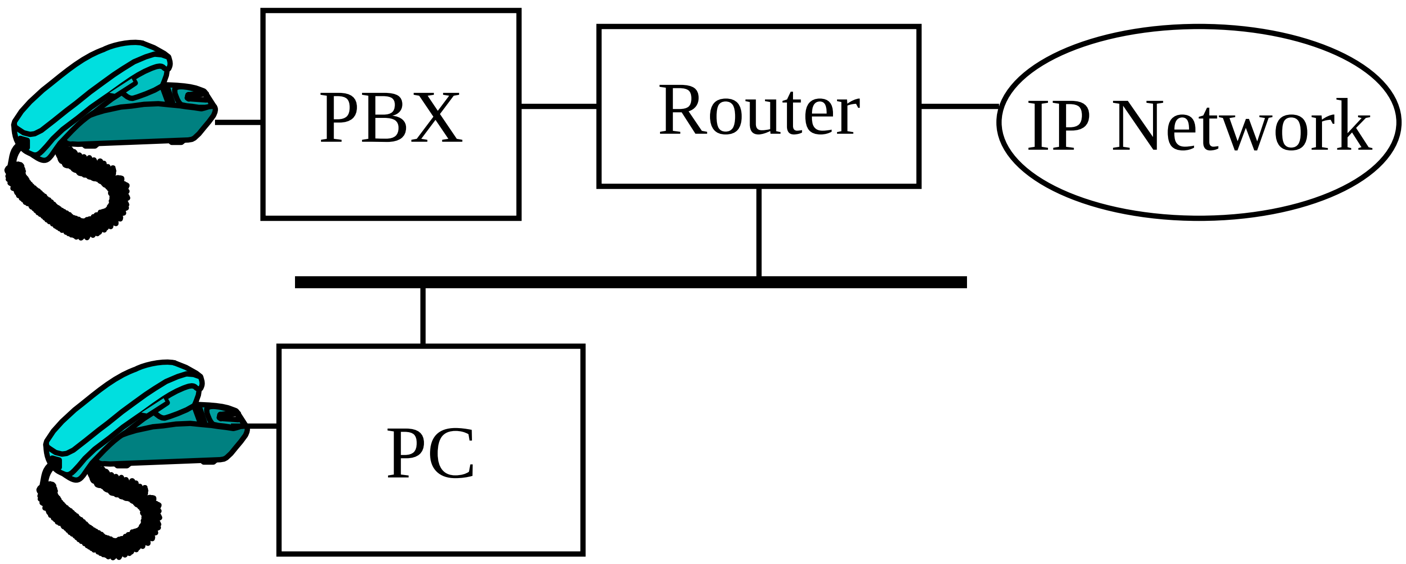
Sample Products

- Microsoft NetMeeting: PC to PC. Free.
- Micom/VocalTec/Lucent:
  - Analog and digital voice interface cards
  - PC and/or gateway. RSVP. Limits # of calls.



Products (Cont)

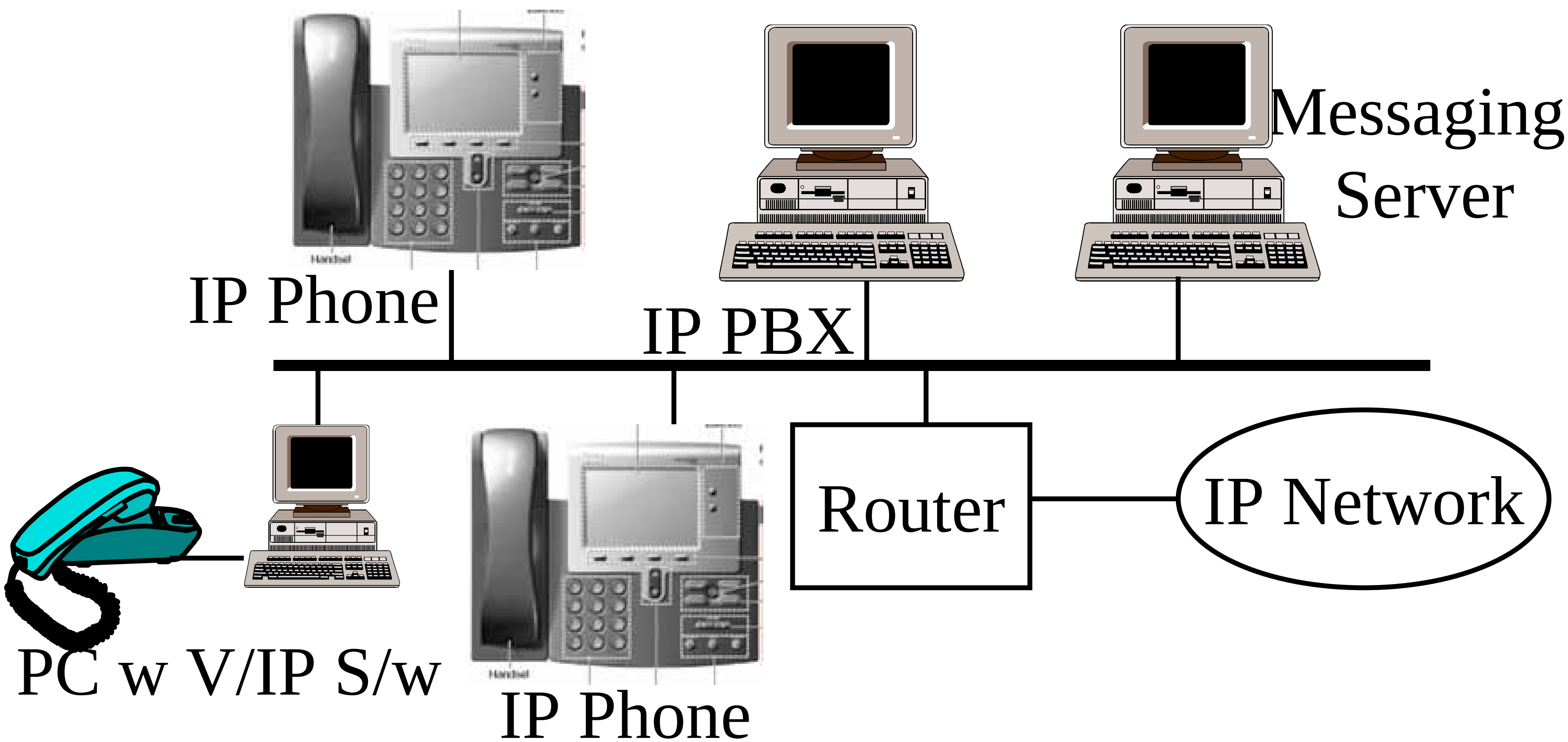
- CISCO, Baynetworks, 3COM, and others: Voice enable routers



- Amisystech's Klik2Talk enables online users to speak instantly with a company's customer service with a single mouse click.

Products (Cont)

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



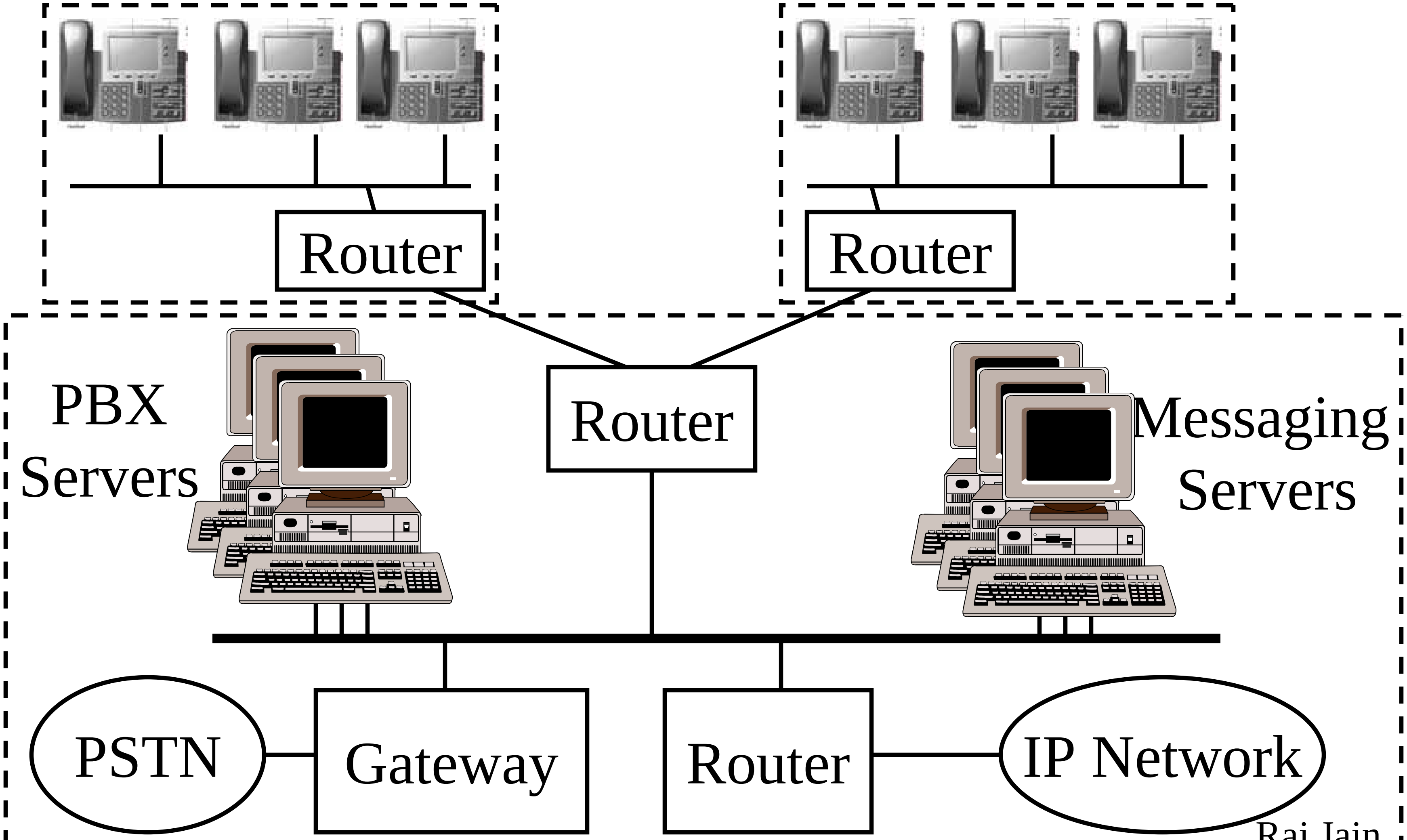
Raj Jain

Sample Services

- IDT Corporation offers Net2Phone, Carrier2Phone, Phone2Phone services.
- ITXC provides infrastructure and management to 'Internet Telephone Service Providers (ITSPs)'

Services (Cont)

- Hosted PBX Service



Raj Jain

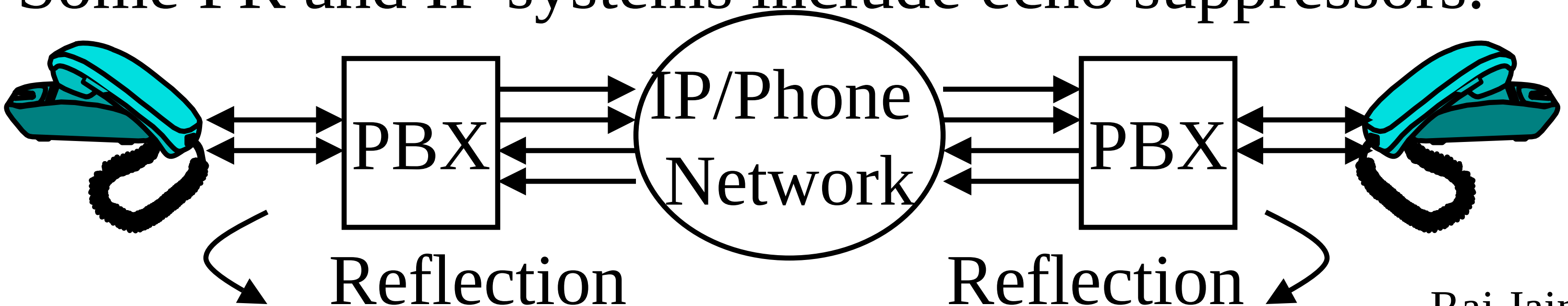
Technical Issues

1. Large Delay

- Normal Phone: 10 ms/kmile $\Rightarrow$ 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 $\Rightarrow$ 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.



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

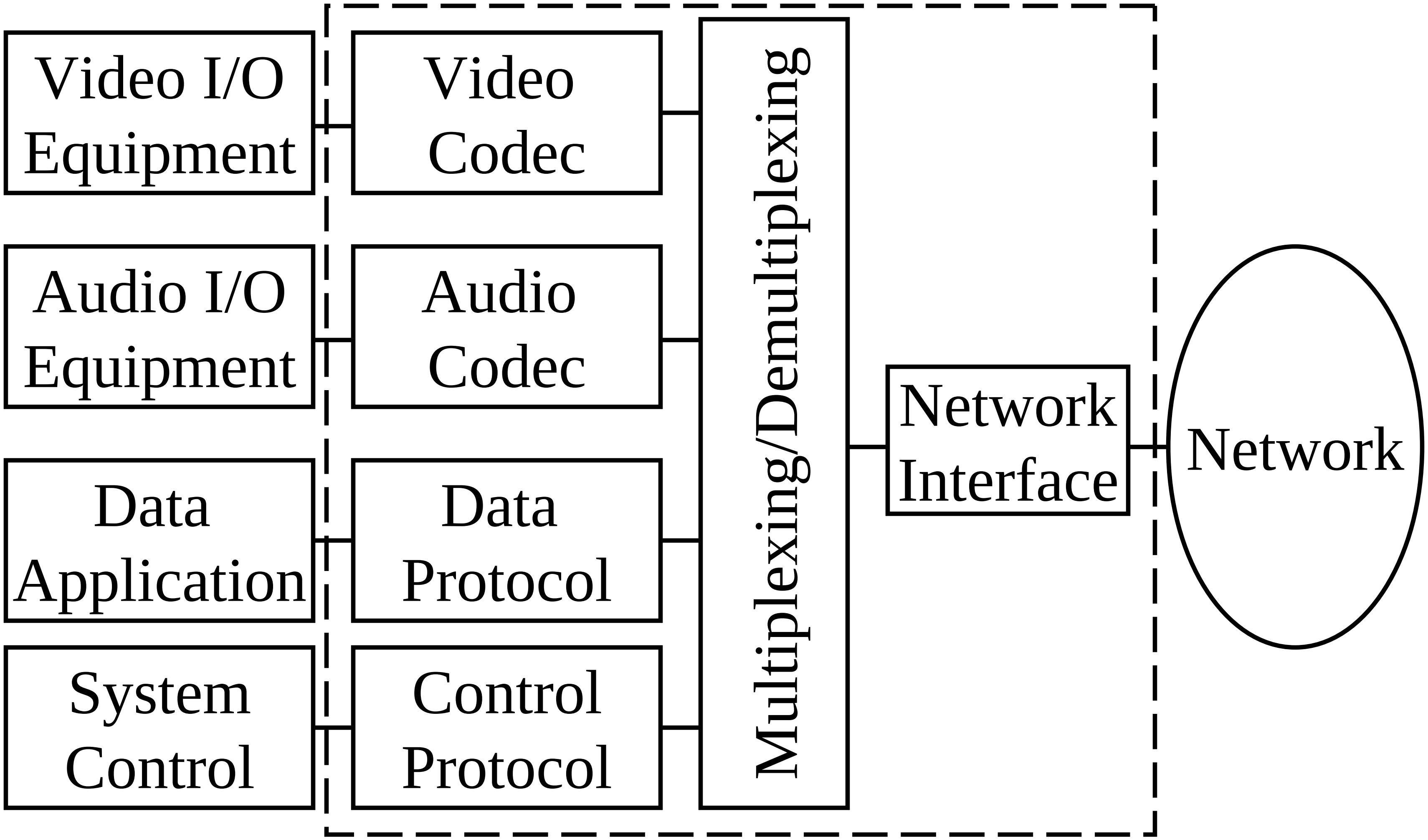
- RTP, UDP, IP headers increase the bandwidth required

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.

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	-

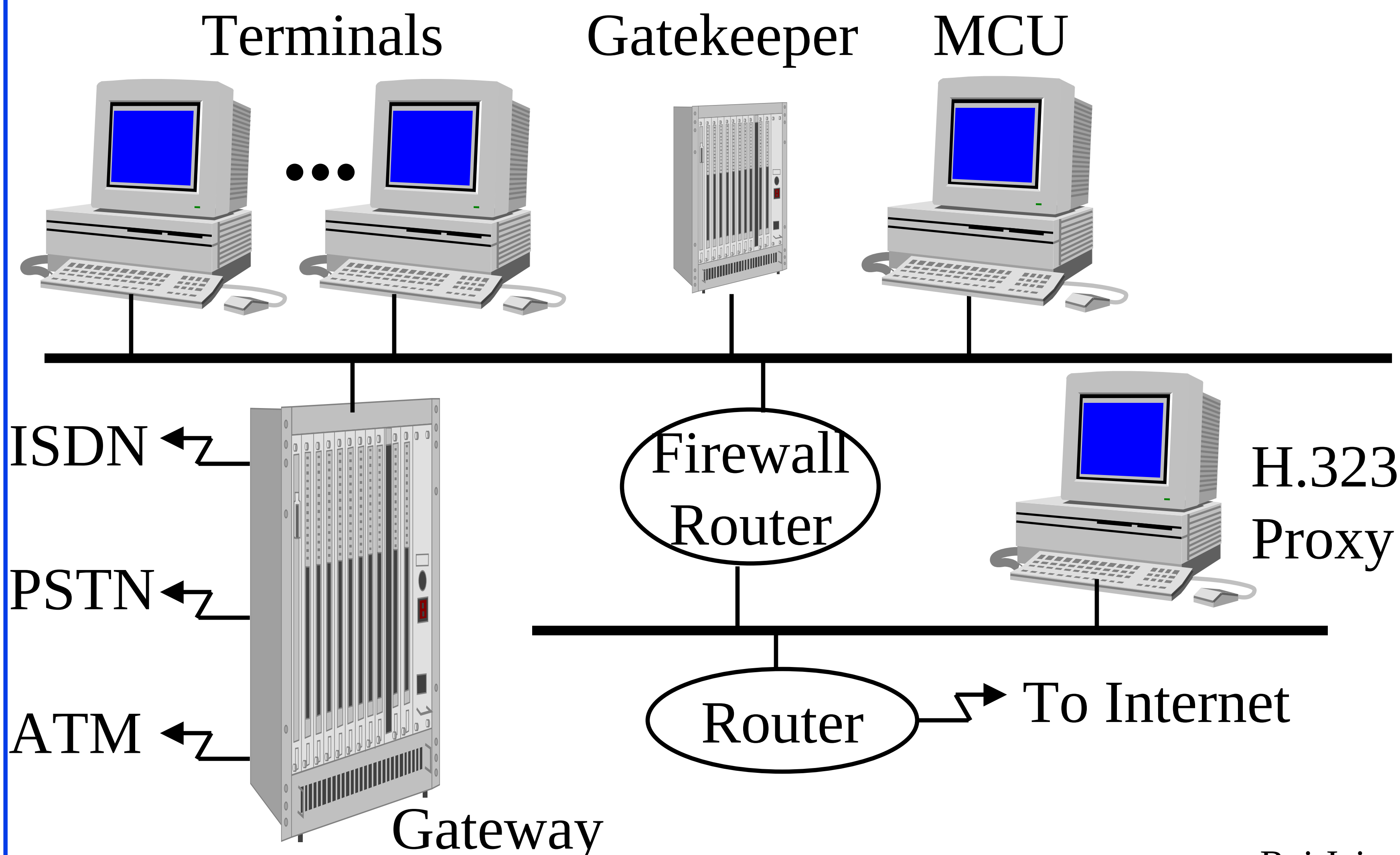
Raj Jain

H.323 Protocols

- Multimedia over LANs, V1 (June 96), V2(Feb 98)
- Provides component descriptions, signaling procedures, call control, system control, audio/video codecs, data protocols

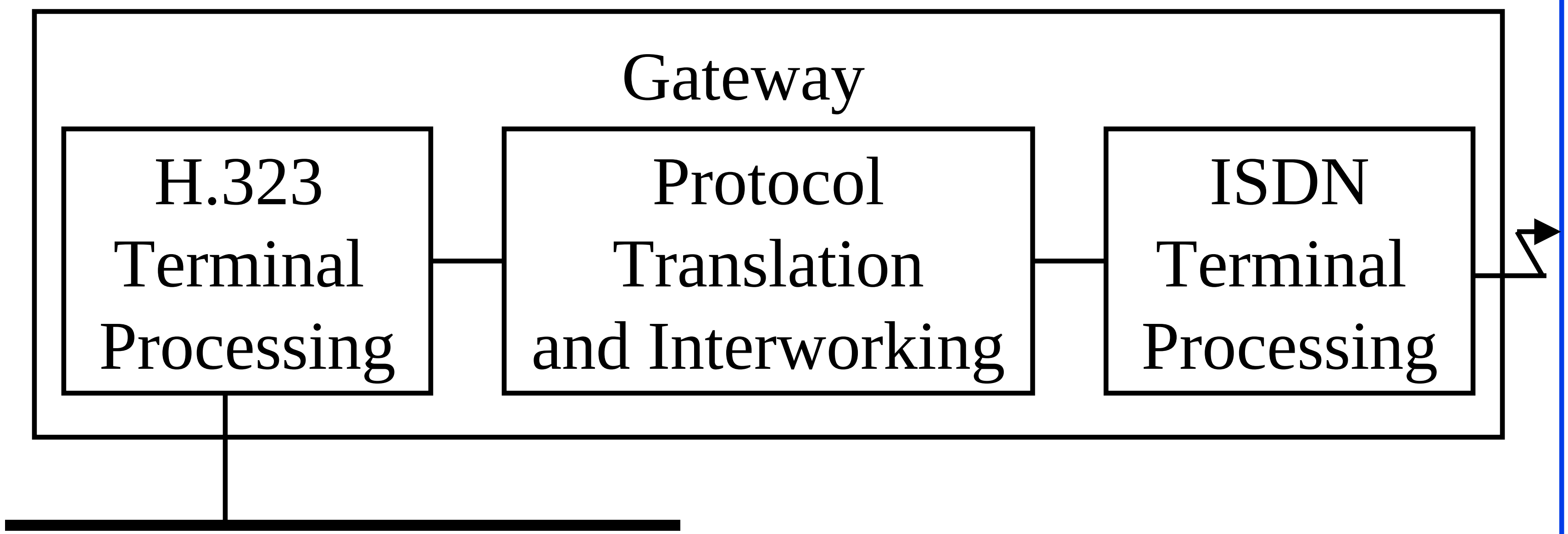
Video	Audio	Control and Management				Data
H.261 H.263	G.711, G.722, G.723.1, G.728, G.729	RTCP	H.225.0 RAS	H.225.0 Signaling	H.245 Control	T.124
RTP			X.224 Class 0			T.125
UDP			TCP			T.123
Network (IP)						
Datalink (IEEE 802.3)						

H.323 Components



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.



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

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)

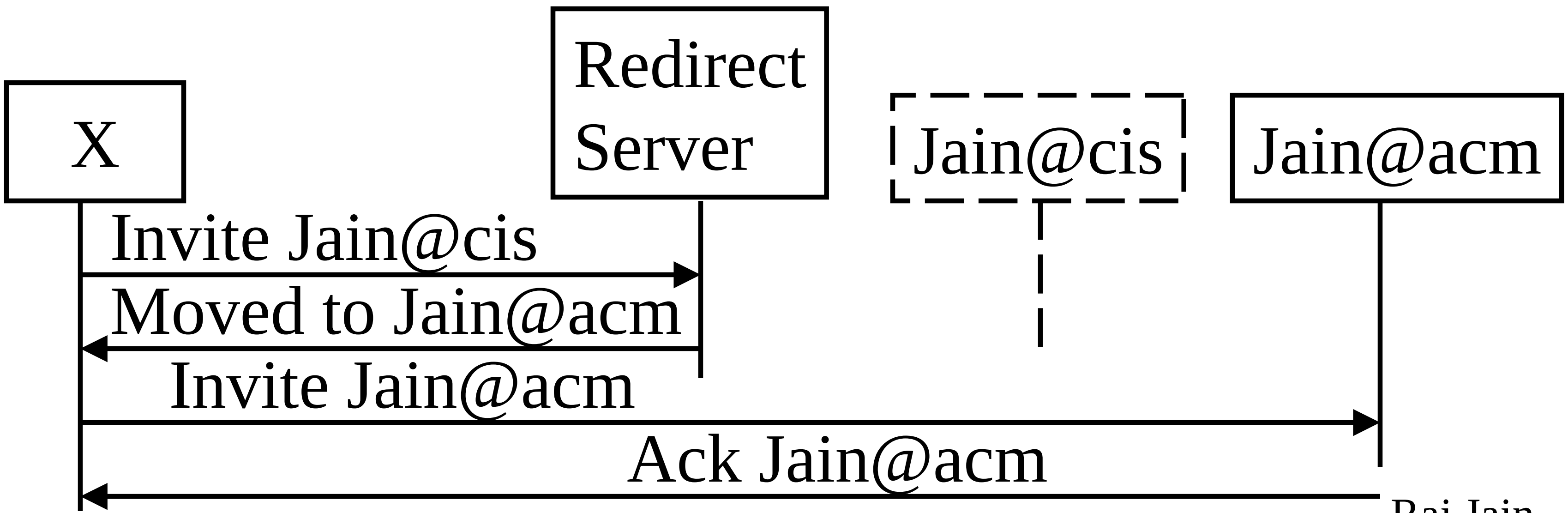
- SIP works in conjunction with other IP protocols for multimedia:
  - RSVP for reserving network resources
  - RTP/RTCP/RTSP for transporting real-time data
  - Session description protocol (SDP) for describing multimedia session
- Can also be used to determine whether party can be reached via H.323,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

SIP (Cont)

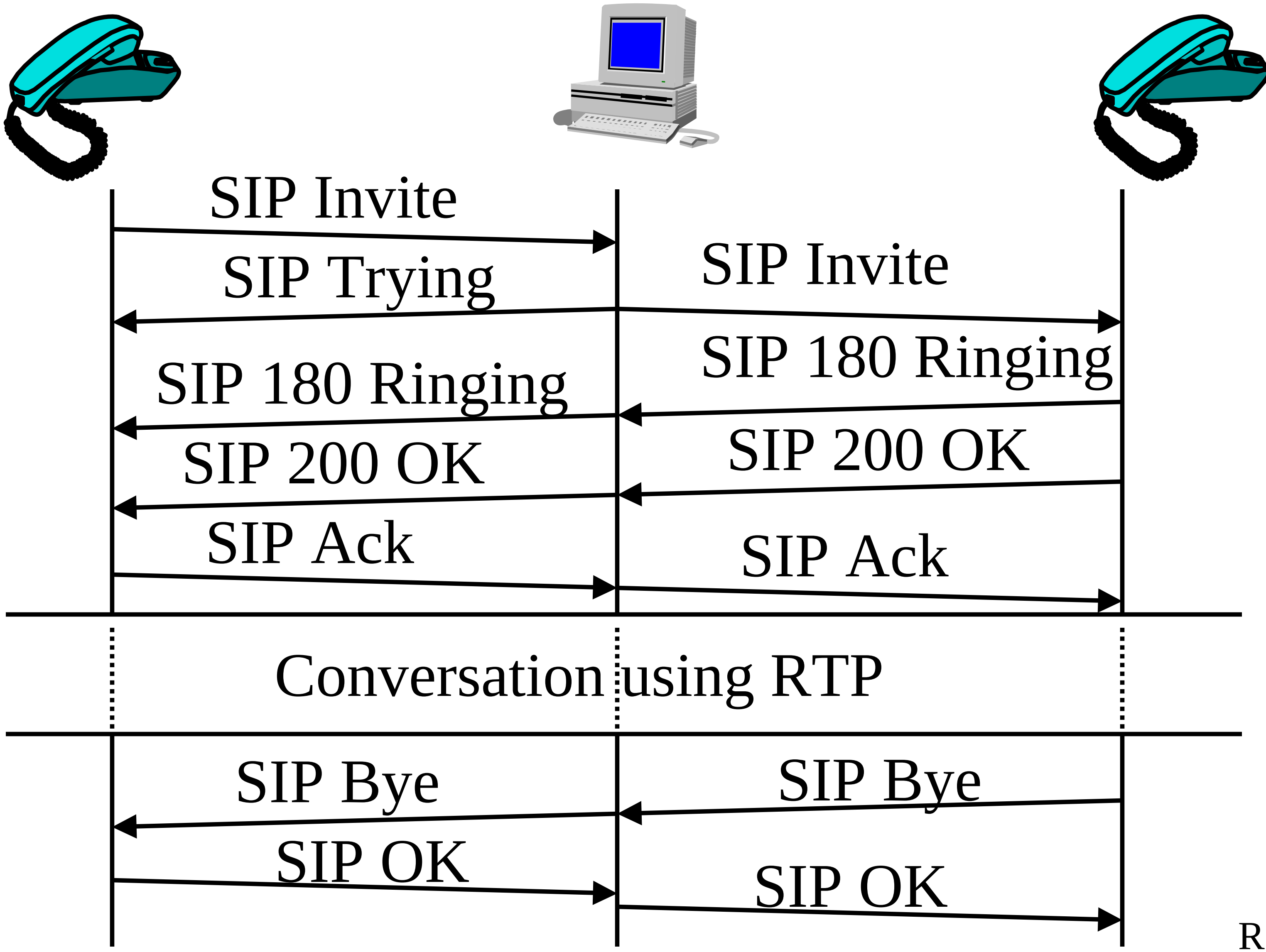
- SIP Uniform Resource Identfiers (URIs):
  Similar to email URLs
  sip:jain@cis.ohio-state.edu
  sip:+1-614-292-3989:123@osu.edu?subject=lecture
- SIP can use UDP or TCP
- SIP messages are sent to SIP servers:
  - Registrar: Clients register and tell their location to it
  - Location: Given name, returns 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

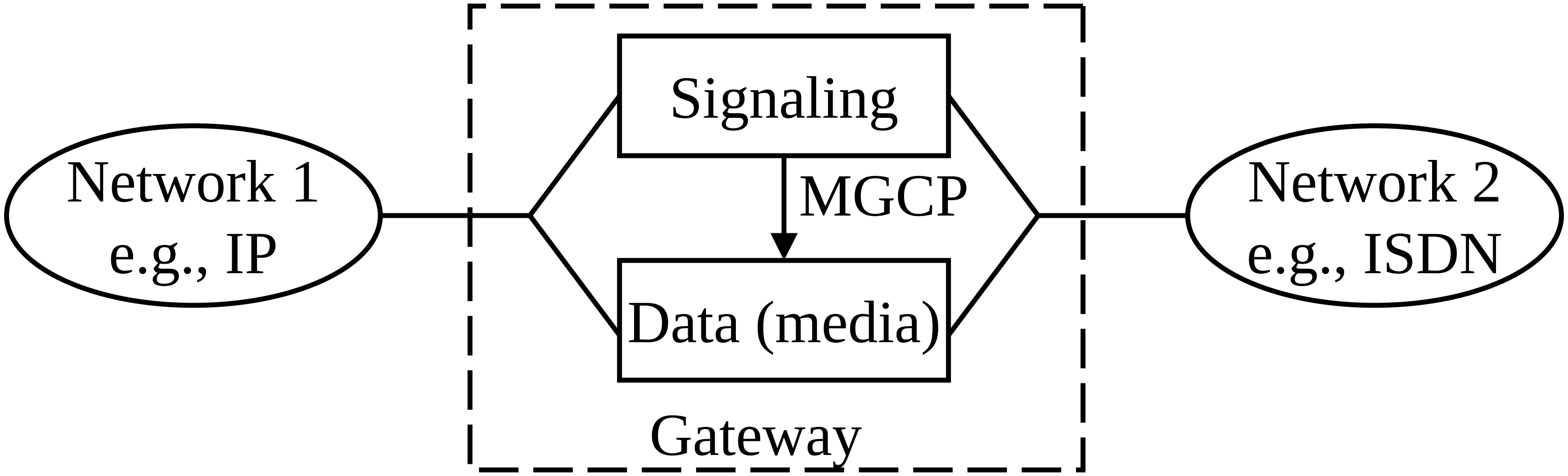


Raj Jain

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

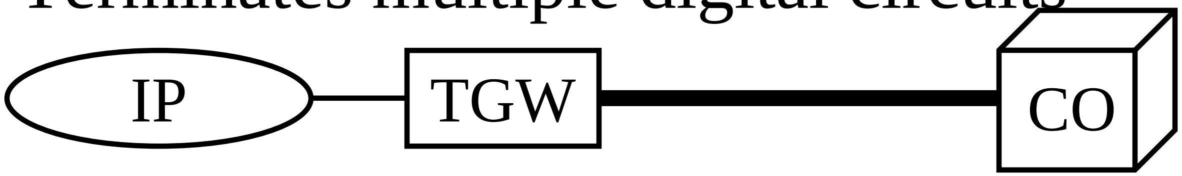
Media Gateway Control Protocol



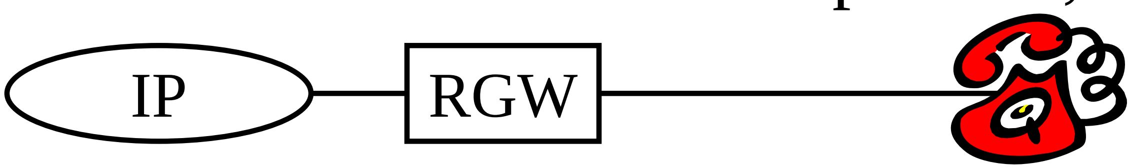
- Gateway = Signaling Fns + Media Transfer Fns
- Call Agents: Signaling functions ⇒ Intelligent
  ⇒ More complex ⇒ Fewer
  ⇒ Media Gateway Controller (MGC) or Softswitch
  ⇒ Control multiple media gateways ⇒ Need MGCP
- MGCP =Simple Gateway Control Protocol (SGCP)
  + Internet Protocol Device Control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

Media Gateways: Examples

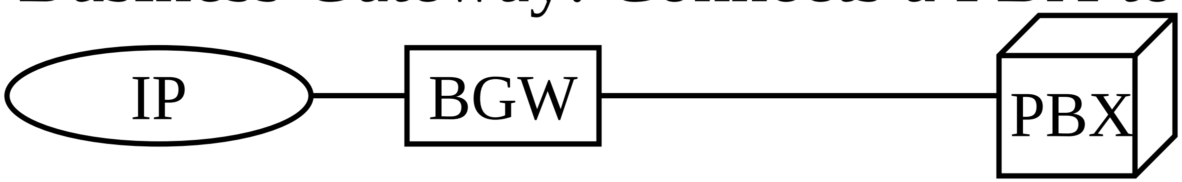
- Trunking Gateway: Connects a PSTN trunk to VOIP
  Terminates multiple digital circuits



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



- Business Gateway: Connects a PBX to VOIP



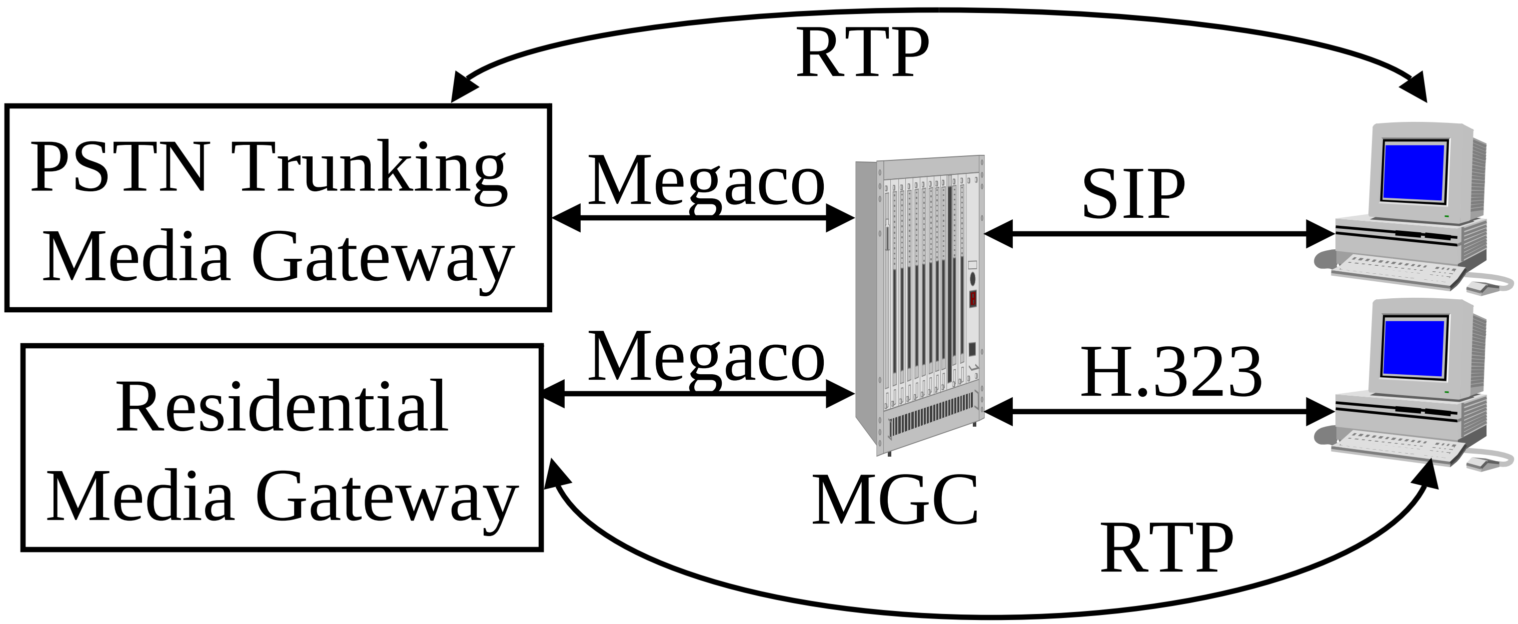
- Network Access Servers: Answer data + VOIP calls

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

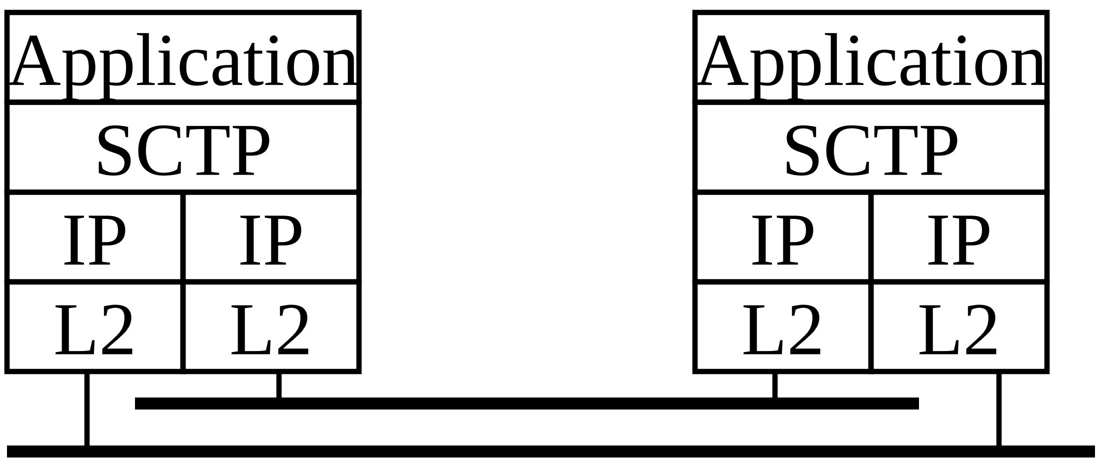
Megaco Framework



- 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).

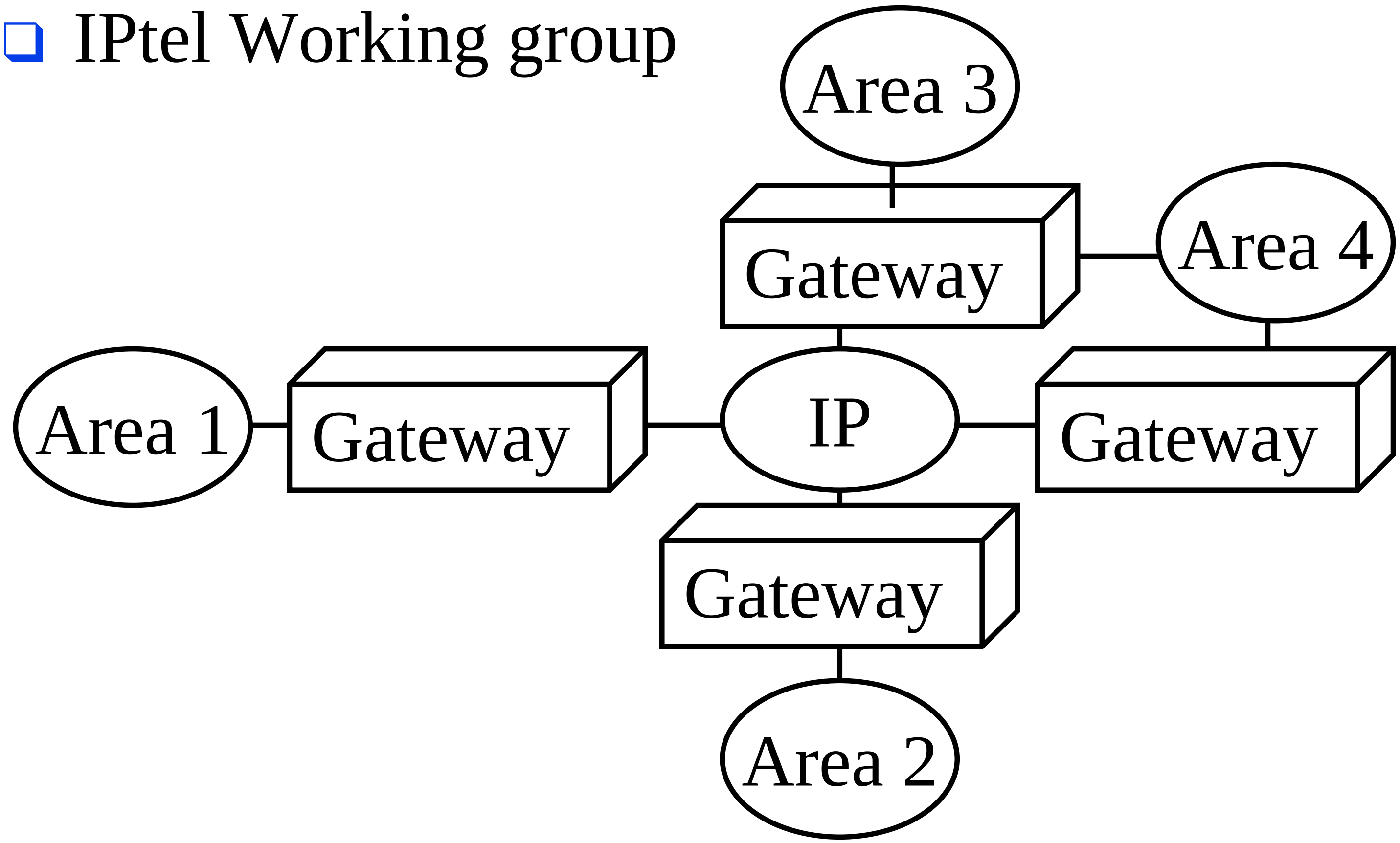
SIGTRAN and SCTP

- 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

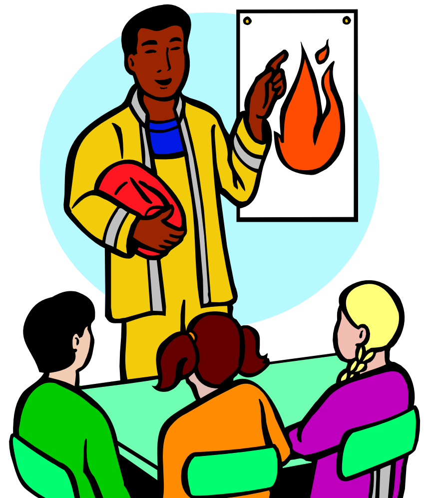


TRIP

- Telephony Routing over IP (TRIP)
- Allows locating gateways and selection of call routes
- IPtel Working group



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