Gigabit Ethernet: Architectural Design and Issues

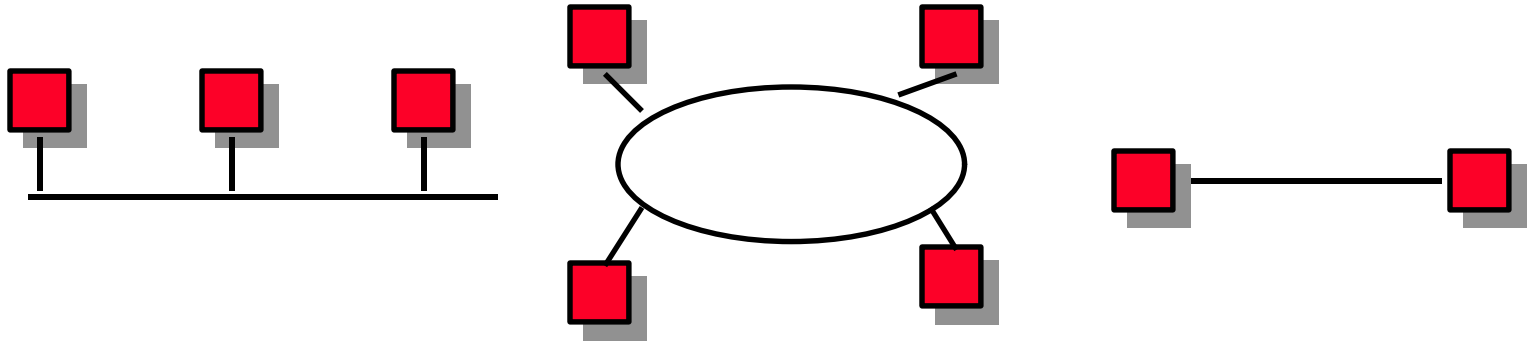
Raj Jain

Profe iences

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

- Distance-Bandwidth Principle
- 10 Mbps to 100 Mbps
- Gigabit PHY Issues
- Gigabit MAC Issues
- Status
- ATM vs Gigabit Ethernet

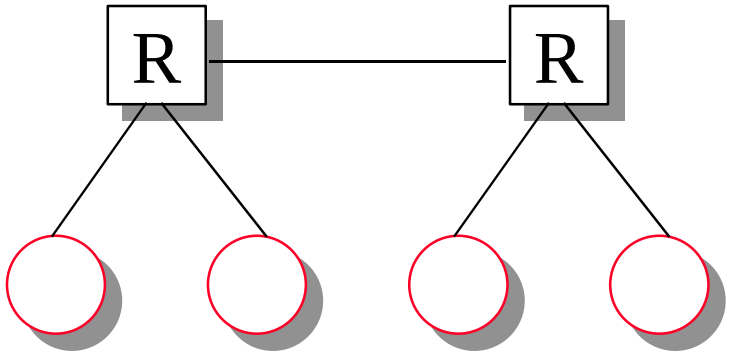
Distance-B/W Principle



- Efficiency = Max throughput/Media bandwidth
- Efficiency is a decreasing 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 $\Rightarrow$ Change distance or frame size

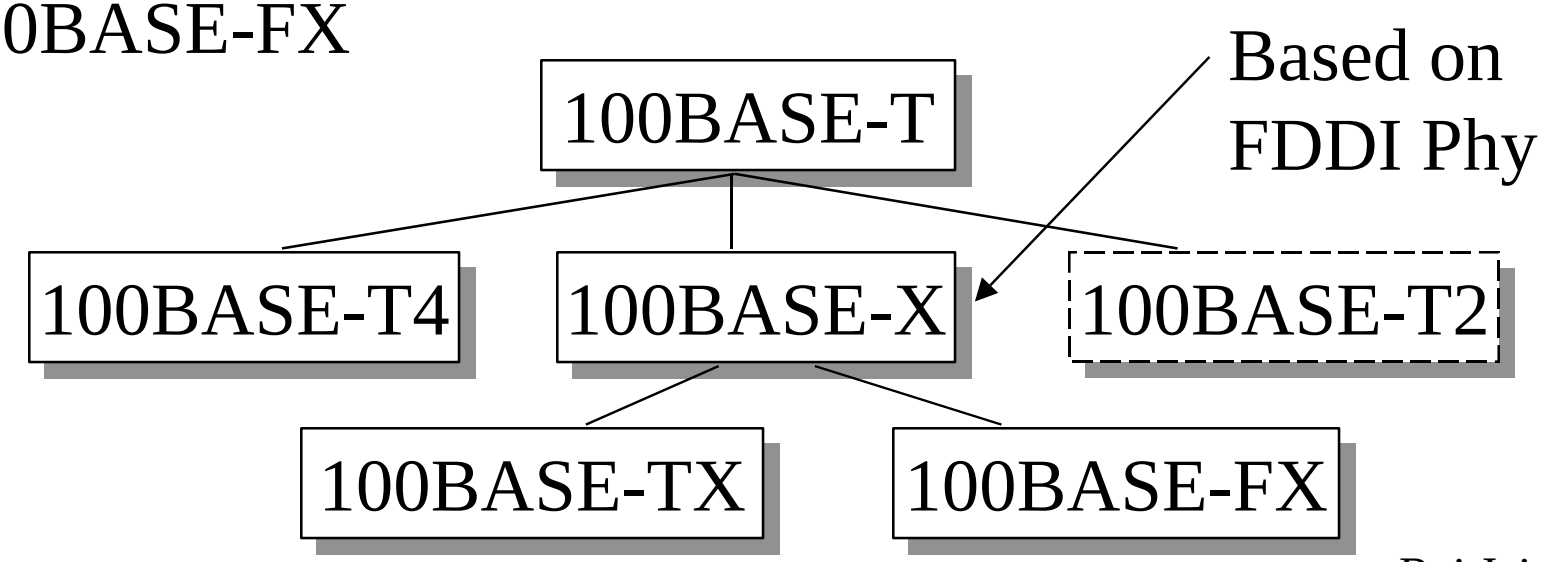
Ethernet vs Fast Ethernet

	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



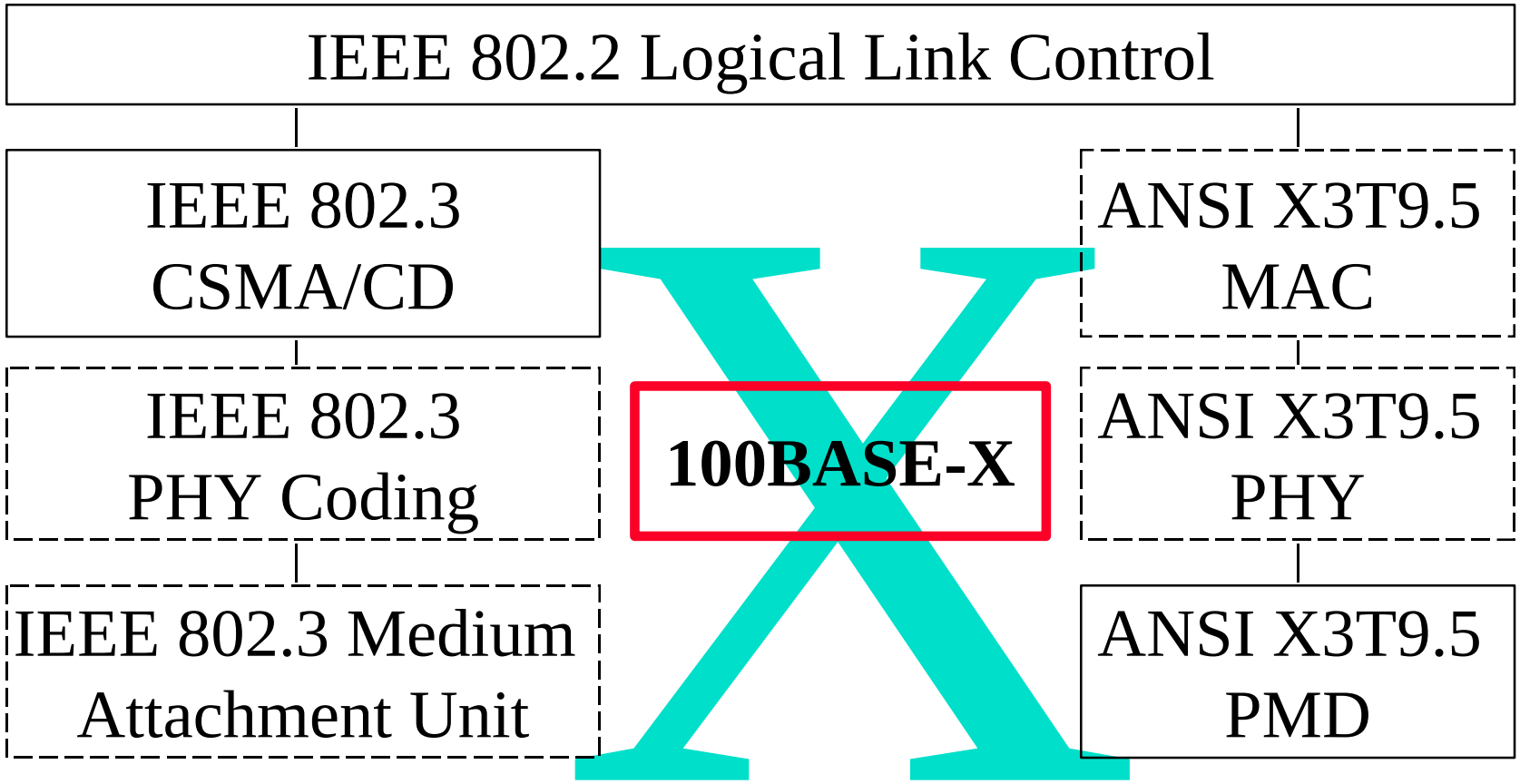
Fast Ethernet Standards

- **100BASE-T4:** 100 Mb/s over 4 pairs of CAT-3, 4, 5
- **100BASE-TX:** 100 Mb/s over 2 pairs of CAT-5, STP
- **100BASE-FX:** 100 Mbps CSMA/CD over 2 fibers
- **100BASE-X:** 100BASE-TX or 100BASE-FX
- **100BASE-T:** 100BASE-T4, 100BASE-TX, or 100BASE-FX

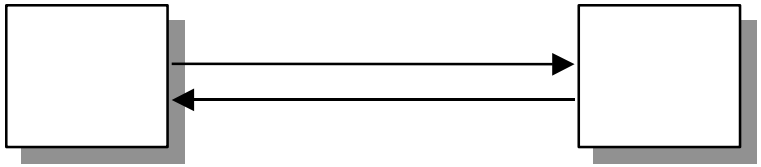


100 BASE-X

- X = Cross between IEEE 802.3 and ANSI X3T9.5



Full-Duplex Ethernet



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

Gigabit Ethernet

- Being standardized by 802.3z
- Project approved by IEEE in June 1996
- 802.3 meets every three months ⇒ Too slow
  ⇒ Gigabit Ethernet Alliance (GEA) formed.
  It meets every two weeks.
- Decisions made at GEA are formalized at 802.3 High-Speed Study Group (HSSG)
- Based on Fiber Channel PHY
- Shared (half-duplex) and full-duplex version
- Gigabit 802.12 and 802.3 to have the same PHY

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

Physical Media

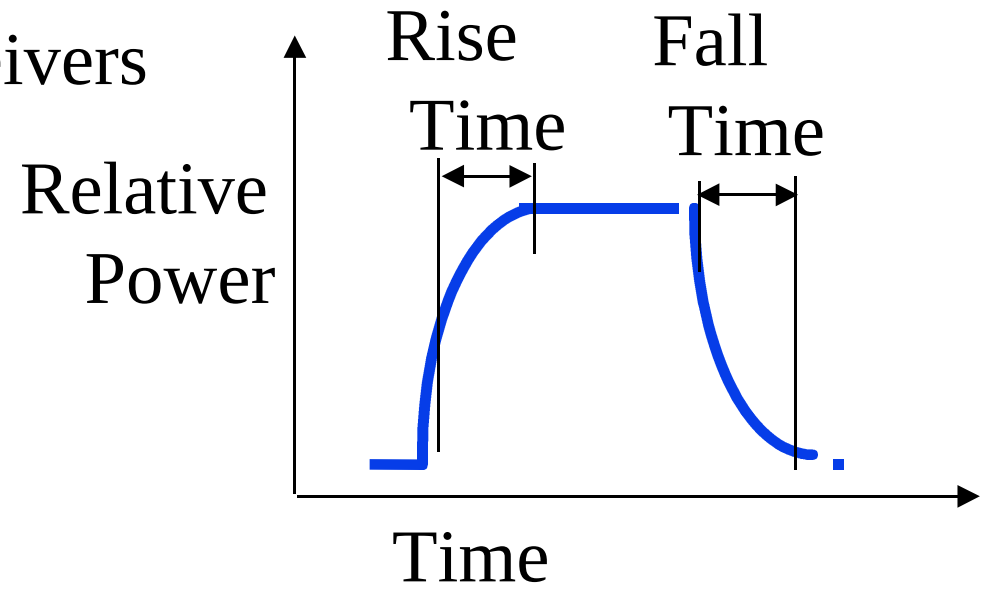
- Unshielded Twisted Pair (UTP-5): 4-pairs
- Shielded Twisted Pair (STP)
- Multimode Fiber: 50 μm and 62.5 μm
  - Use CD lasers
- Single-Mode Fiber
- Bit Error Rate better than 10^{-12}

How Far Should It Go?

- Full-Duplex:
  - Fiber Channel: 300 m on 62.5 μm
    at 800 Mbps ⇒ 230 m at 1000 Mbps
  - Decision: 500 m at 1000 Mbps
    ⇒ Minor changes to FC PHY
- Shared:
  - CSMA/CD without any changes
    ⇒ 20 m at 1 Gb/s (Too small)
  - Decision: 200 m shared
    ⇒ Minor changes to 802.3 MAC

PHY Issues

- Fiber Channel PHY:
  100 MBps = 800 Mbps
  $\Rightarrow$ 1.063 GBaud using 8b10b
- Changes to get 500 m on 62.5-μm multimode fiber
  - Modest decrease in rise and fall times of the transceivers



- Symbol Codes for Specific Signals: Jam, End-of-packet, beginning of packet
- PHY-based flow Control: No.
  Use the XON/XOFF flow control of 802.3x

850 nm vs 1300 nm lasers

- 850 nm used in 10Base-F
  - Cannot go full distance with 62.5-µm fiber
  - 500 m with 50-µm fiber
  - 250 m with 62.5-µm fiber
- 1300 nm used in FDDI but more expensive
  - Higher eye safety limits
  - Better Reliability
  - Start with 550 m on 62.5-µm fiber
  - Could be improved to 2 km on 62.5-µm fiber
    $\Rightarrow$ Needed for campus backbone

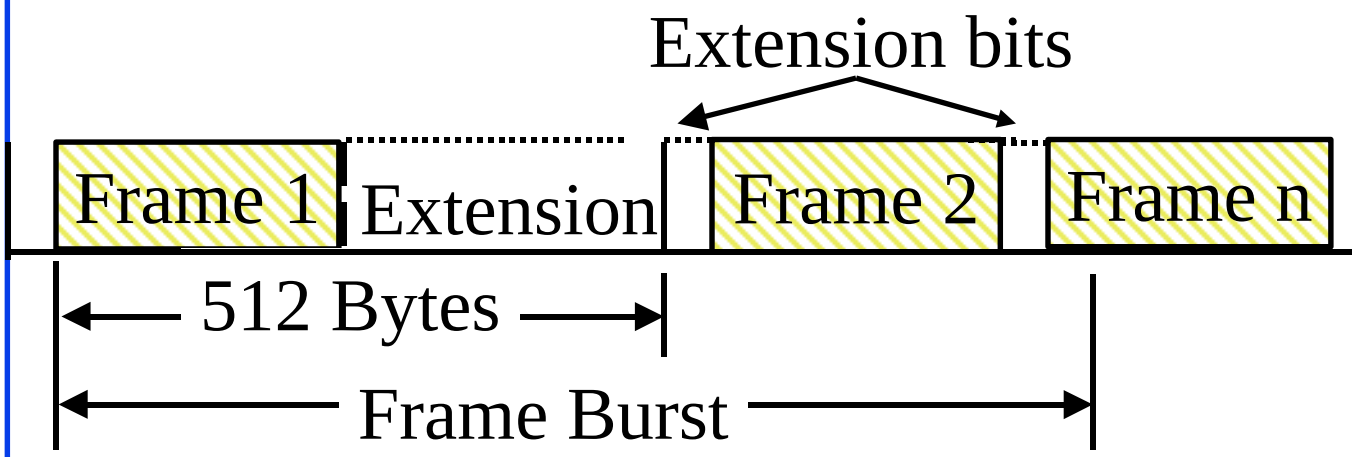
Media Access Control Issues

- Carrier Extension
- Frame Bursting
- Buffered Distributor

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

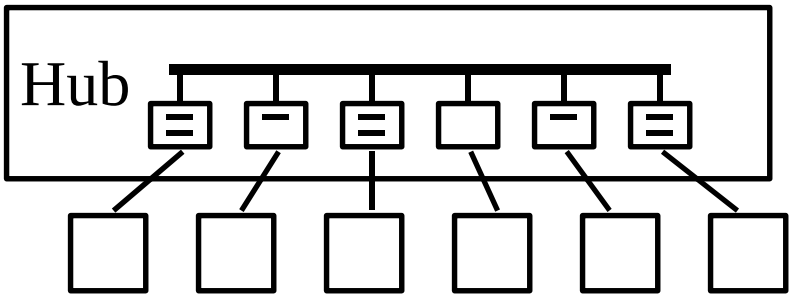


- Don't give up the channel after every frame
- After the slot time, continue transmitting additional frames (with minimum inter-frame gap)
- Interframe gaps are filled with extension bits
- No no new frame transmissions after 8192 bytes
- Three times more throughput for small frames

Frame Packing

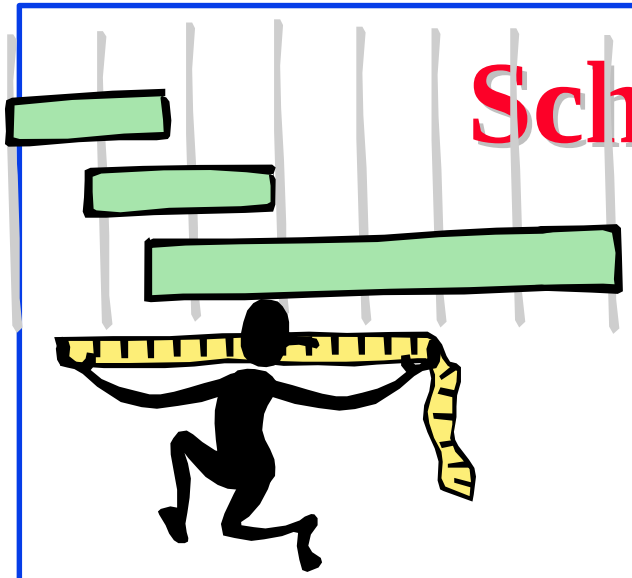
- Multiple small frames transmitted on a single carrier event
- Minimum inter-frame gap between frames
- Extend carrier if no more frames.
  No new frames after slot time.
- Sender retransmitts all frames on collision
- Receiver must discard all frames on collision
- Was considered but not accepted

Buffered Distributor



- All incoming frames are buffered in FIFOs
- CSMA/CD arbitration inside the box to transfer frames from an incoming FIFO to all outgoing FIFOs
- Previous slides were half-duplex. With buffered distributor all links are full-duplex with frame-based flow control
- Link length limited by physical considerations only

Schedule



- November 1996: Proposal cutoff
- January 1997: First draft
- March 1997: Second draft
- July 1997: Working Group Ballot
- March 1998: Approval

Status

- On Schedule
- First draft reviewed in January 97
- Fourth draft was issued in December'97
- 1000Base-X: Gigabit Ethernet based on Fiber Channel Phy
- Phy modified for 1000 Mbps operation
- Phy modified for ISO 11801 standard for premises cabling $\Rightarrow$ 550 m intra-building backbone runs
  $\Rightarrow$ 1300-nm lasers on 62.5-μm multimode fiber
  850-nm lasers on 62.5-μm fiber ok for 300 m

1000Base-X

- 1000Base-LX: 1300-nm laser transceivers
  - 2 to 550 m on 62.5-μm or 50-μm multimode, 2 to 3000 m on 10-μm single-mode
- 1000Base-SX: 850-nm laser transceivers
  - 2 to 300 m on 62.5-μm, 2 to 550 m on 50-μm. Both multimode.
- 1000Base-CX: Short-haul copper jumpers
  - 25 m 2-pair shielded twinax cable in a single room or rack.
    Uses 8b/10b coding ⇒ 1.25 Gbps line rate

1000Base-T

- 100 m on 4-pair Cat-5 UTP
  - Network diameter of 200 m
  - Requires new coding schemes
  - Under development.
  - New PAR approved in March 1997
  - 802.3ab task force

Design Parameter Summary

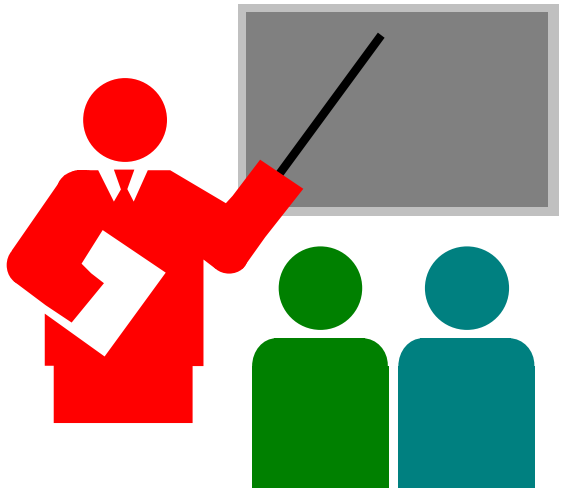
Parameter	10 Mbps	100 Mbps	1 Gbps
Slot time	512 bt	512 bt	4096 bt
Inter Frame Gap	9.6 μs	0.96 μs	0.096 μs
Jam Size	32 bits	32 bits	32 bits
Max Frame Size	1518 B	1518 B	1518 B
Min Frame Size	64 B	64 B	64 B
Burst Limit	N/A	N/A	8192 B

- bt = bit time

ATM vs Gb Ethernet

Issue	ATM	Gigabit Ethernet
Media	SM Fiber, MM Fiber, UTP5	Mostly fiber
Max Distance	Many miles using SONET	260-550 m
Data Applications	Need LANE, IPOA	No changes needed
Interoperability	Good	Limited
Ease of Mgmt	LANE	802.3Q VLANs
QoS	PNNI	802.1p (Priority)
Signaling	UNI	None/RSVP (?)
Traffic Mgmt	Sophisticated	802.3x Xon/Xoff

Summary



- Ethernet will run at 1000 Mbps
- Will compete with ATM for campus backbone and desktop
- Both shared and full-duplex links
- Fully compatible with current Ethernet

References

- For a detailed list of references, see http://www.cis.ohio-state.edu/~jain/refs/gbe_refs.htm
- "Media Access Control (MAC) Parameters, Physical Layer Repeater and Management Parameters for 1000 Mb/s Operation," IEEE Draft P802.3z/D4, December 22, 1997.

References (Cont)

- Email Reflector:
  stds-802-3-hssg@mail.ieee.org
  - To join send email to majordomo@mail.ieee.org
  - subscribe stds-802-3-hssg@mail.ieee.org <your email address>
- FTP Site:
  ftp://stdsbbs.ieee.org/pub/802_main/802.3/gigabit
- Gigabit Ethernet Consortium
  http://www.gigabit-ethernet.org