GPUR

NPTEL ONLINE
CERTIFICATION COURSES

NPTEL

COMPUTER NETWORKS AND INTERNET PROTOCOLS

Data Link Layer - Ethernet

IIT KHARAGPUR
SANDIP CHAKRABORTY
COMPUTER SCIENCE AND
ENGINEERING IIT KHARAGPUR

SOUMYA K GHOSH
COMPUTER SCIENCE AND ENGINEERING

Shared Access Networks

- Shared Access

Networks assume multiple nodes on the same physical link

– Bus, ring and wireless structures

  – Transmission sent by one node is received by all others
  – No intermediate switches
- Methods for moderating access (MAC protocols)
  – Fairness
  – Performance

Random Access MAC Protocols

• When node has packet to send

- Transmit at full channel data rate R
- No *a priori* coordination among nodes

• Two or more transmitting nodes “collision” • **Random access MAC protocol** specifies:

- How to detect collisions
- How to recover from collisions (e.g., via delayed retransmissions)

• Examples of random access MAC protocols: – ALOHA

– Slotted ALOHA
– CSMA and CSMA/CD

Aloha – Basic Approach

• First random MAC developed

– For radio-based communication in Hawaii (1970)

- Basic idea:
  - – When you’re ready, transmit
  - – Receiver’s send ACK for data
  - – Detect collisions by timing out for ACK
  - – Recover from collision by trying after random delay •

Too short large number of collisions

Aloha Networks - Overview

- Developed by Norm Abramson at Univ. of Hawaii for use with packet radio systems

- Any station can send data at any time
- Receiver sends an ACK for data
- Timeout for ACK signals that there was a collision
  - What happens if timeout is poorly timed?
- If there is a collision, sender will resend data after a random backoff

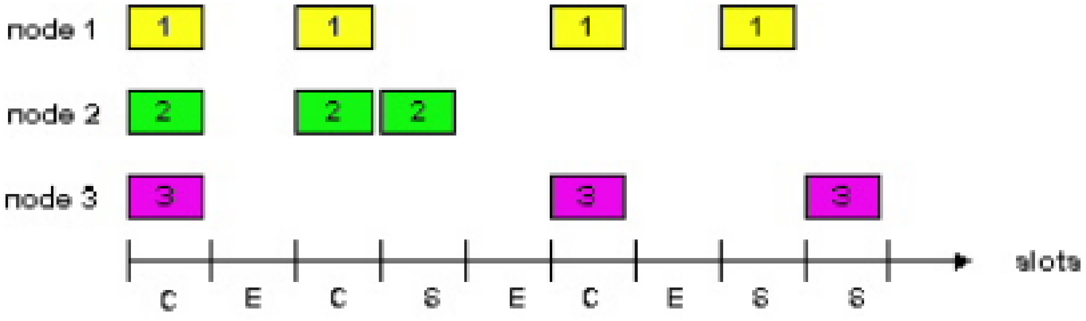
- Utilization (fraction of transmitted frames avoiding collision for N nodes) was pretty bad
  - Max utilization = 18%
- Slotted Aloha (dividing transmit time into windows) helped
  - Max utilization increased to 36%

Slotted Aloha

- Time is divided into equal size

slots (i.e. packet transmission time) • Node (w/ packet) transmits at beginning of next slot

- If collision: retransmit packet in future slots with probability p, until successful



Success (S), Collision (C), Empty (E) slots

Pure (Unslotted) ALOHA

- Unslotted Aloha: simpler, no synchronization
- Packet needs transmission:
  – Send without awaiting for beginning of slot
- Collision probability increases:

1]

Ethernet

- First practical local area network, built at Xerox PARC in 70's
- "Dominant" LAN technology:
  – Cheap

Ethernet MAC – Carrier Sense

10, 100, 1000 Mbps

Metcalfe’s Ethernet sketch

- Basic idea:
  - Listen to wire before transmission
  - Avoid collision with active transmission

Hidden Exposed

- Why didn't ALOHA have this?
  - In wireless, relevant contention at the *receiver*, not sender
    - Hidden terminal
    - Exposed terminal

NY

Chicago CMU

CMU

NY

Chicago

St.Louis

Multiple Access Methods

- Fixed assignment

– Partition channel so each node gets a slice of the bandwidth

– Essentially circuit switching – thus inefficient

– Examples:

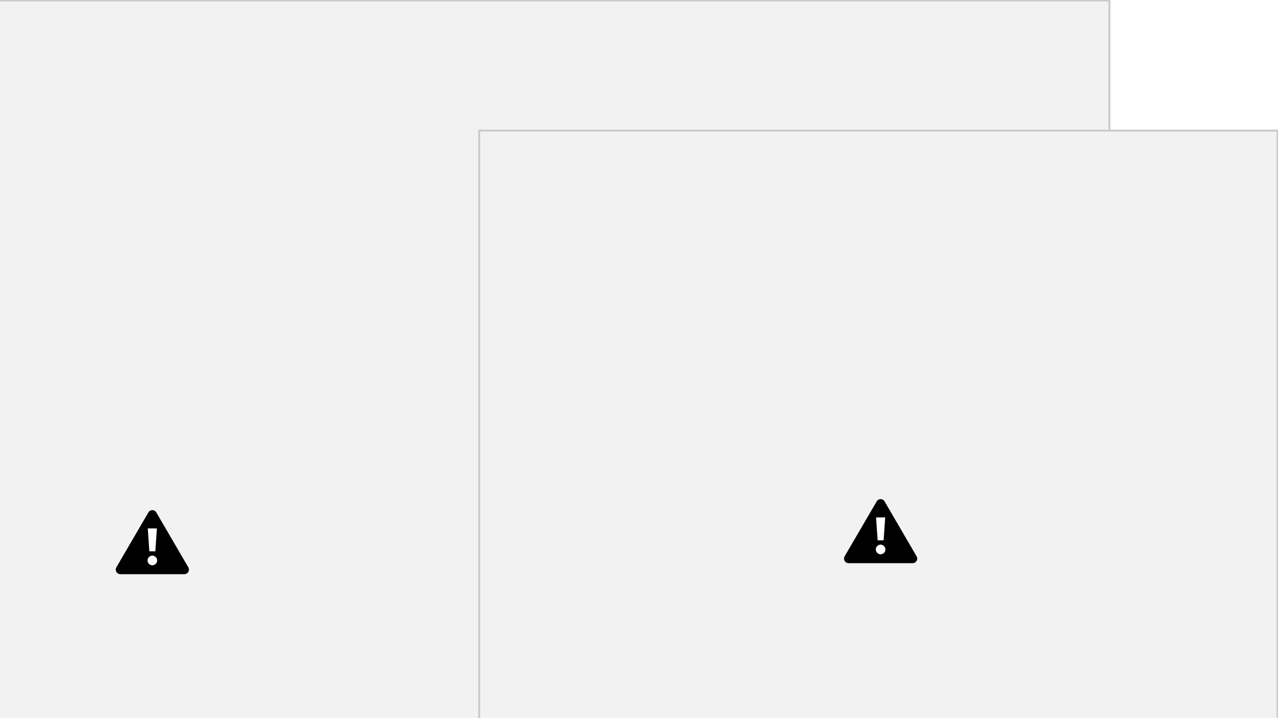
TDMA, FDMA,

CDMA (all used in wireless/cellular environments) •

Contention-based

- Nodes contends equally for bandwidth and recover from collisions
- Examples: Aloha, Ethernet

- Token-based or reservation-based
  - Take turns using the channel
  - Examples: Token ring



Ethernet

- Background
  -

Developed by Bob Metcalfe and others at Xerox PARC in mid-1970s –
Roots in Aloha packet-radio network

– Standardized by Xerox, DEC, and Intel in 1978

– LAN standards define MAC and physical layer connectivity •

IEEE 802.3 (CSMA/CD - Ethernet) standard – originally 2Mbps •
IEEE 802.3u standard for 100Mbps Ethernet

- IEEE 802.3z standard for 1,000Mbps Ethernet

- CSMA/CD: Ethernet's Media Access Control (MAC) policy –

CS = Carrier Sense

- Send only if medium is idle

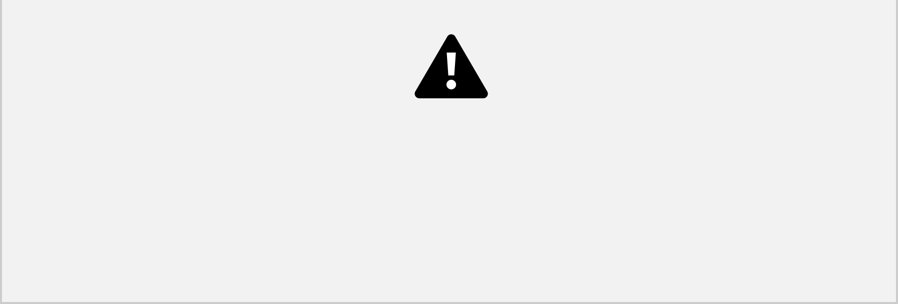
– MA = Multiple Access

– CD = collision detection

- Stop sending immediately if collision is detected

Ethernet Standard

- 802.3 standard defines both MAC *and* physical layer details



Metcalfe’s original Ethernet sketch

Ethernet Technologies: 10Base2

- ***10:*** 10Mbps; ***2:*** under 185 (~200) meters cable length
- Thin coaxial cable in a bus topology
- Repeaters used to connect multiple segments

– Repeater repeats bits it hears on one interface to its other interfaces: physical layer device only!

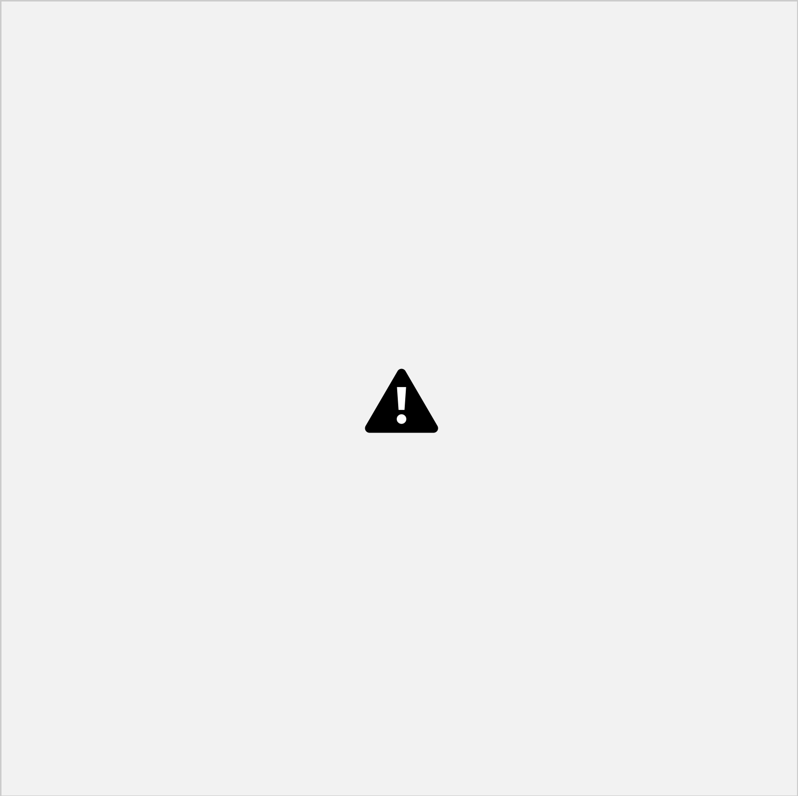
10BaseT and 100BaseT

- 10/100 Mbps rate
- ***T*** stands for Twisted Pair
- Hub(s) connected by twisted pair facilitate “star topology”

– Distance of any node to hub must be < 100M

configurations for 802.3

• Physical



layer

configurations are specified in three parts • Data rate (10, 100, 1000)

– 10, 100, 1000Mbps

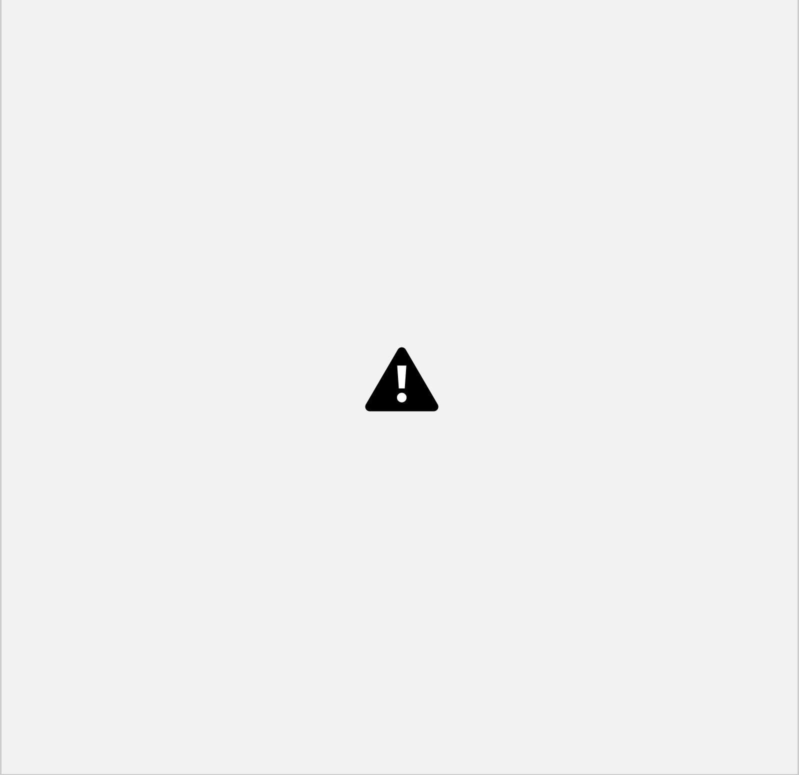
Ethernet Overview

- Signaling method (base, broad)
  - Baseband
    - Digital signaling
  - Broadband
    - Analog signaling
- Cabling (2, 5, T, F, S, L)
  - 5 - Thick coax (original Ethernet cabling)
  - F – Optical fiber
  - S – Short wave laser over multimode fiber
  - L – Long wave laser over single mode fiber

- Most popular packet-switched LAN technology
- Bandwidths: 10Mbps, 100Mbps, 1Gbps
- Max bus length: 2500m
  - 500m segments with 4 repeaters
- Bus and Star topologies are used to connect hosts
  - Hosts attach to network via Ethernet transceiver or hub or switch
    - Detects line state and sends/receives signals
  - Hubs are used to facilitate shared connections
  - All hosts on an Ethernet are competing for access to the medium
    - Switches break this model

- Challenge: Distributed algorithm that provides fair access

- Ethernet by definition is a broadcast protocol
  - Any signal can be received by all hosts
  - Switching enables individual hosts to communicate
- Network layer packets are transmitted over an Ethernet by encapsulating
- Frame Format



64 48 32

48 16

Preamble CRC Src

Dest addr Body

addr Type

Switched Ethernet

• Switches forward and filter frames based on LAN addresses – It’s not a bus or a router (although simple forwarding tables are maintained) • Very scalable

– Options for many interfaces

– Full duplex operation (send/receive frames simultaneously)

• Connect two or more “segments” by

copying data frames between them – Switches only copy data when needed

- key difference from repeaters

- Higher link bandwidth
  - Collisions are completely avoided
- Much greater aggregate bandwidth
  - Separate segments can send at once

Ethernet Frames

- Preamble is a sequence of 7 bytes, each set to

"10101010" – Used to synchronize receiver before actual data is sent

- Addresses
  - unique, 48-bit unicast address assigned to each adapter
    - example: 3**8:10:e4:b1:29:07**
    - Each manufacturer gets their own address range
  - broadcast: all **1**s
  - multicast: first bit is **1**
- *Type* field is a demultiplexing key used to determine which higher level protocol the frame should be delivered to
- Body can contain up to 1500 bytes of data

Ethernet's MAC Algorithm

- In Aloha, decisions to transmit are made without paying attention to what other nodes might be doing

- Ethernet uses CSMA/CD – listens to line before/during sending •
If line is idle (no carrier sensed)
  - send packet immediately
  - upper bound message size of 1500 bytes
  - must wait 9.6us between back-to-back frames
- If line is busy (carrier sensed)
  - wait until idle and transmit packet immediately
    - called *1-persistent* sending
- If collision detected
  - Stop sending and jam signal
  - Try again later

State Diagram for CSMA/CD

Packet?

No

Send Detect Collision

Sense
Carrier

Yes

Jam channel

Discard
Packet

16

b=CalcBackoff();
wait(b);
attempts++;

attempts < 16

attempts ==

COMPUTER NETWORKS AND INTERNET PROTOCOLS

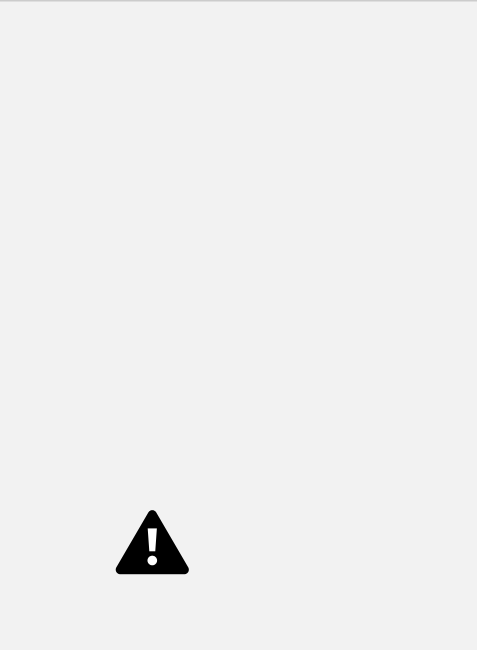
Data Link Layer – Ethernet (contd.)

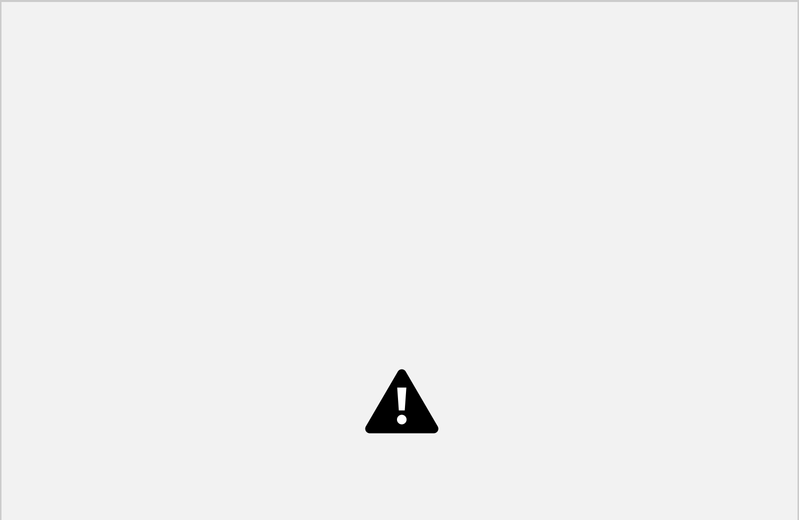
SOUMYA K GHOSH
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

SANDIP CHAKRABORTY
COMPUTER SCIENCE AND ENGINEERING IIT KHARAGPUR

Ethernet

- Background
  –





Developed by Bob Metcalfe and others at Xerox PARC in mid-1970s –
Roots in Aloha packet-radio network

– Standardized by Xerox, DEC, and Intel in 1978

– LAN standards define MAC and physical layer connectivity •

IEEE 802.3 (CSMA/CD - Ethernet) standard – originally 2Mbps •
IEEE 802.3u standard for 100Mbps Ethernet

• IEEE 802.3z standard for 1,000Mbps Ethernet

• CSMA/CD: Ethernet's Media Access Control (MAC) policy –

CS = Carrier Sense

• Send only if medium is idle

– MA = Multiple Access

– CD = collision detection

cted

Collisions

Collisions are
caused when two

adaptors transmit at the same time (adaptors sense collision based on voltage differences) • Both found line to be idle

• Both had been waiting to for a busy line to become idle

A B

A starts at
time 0

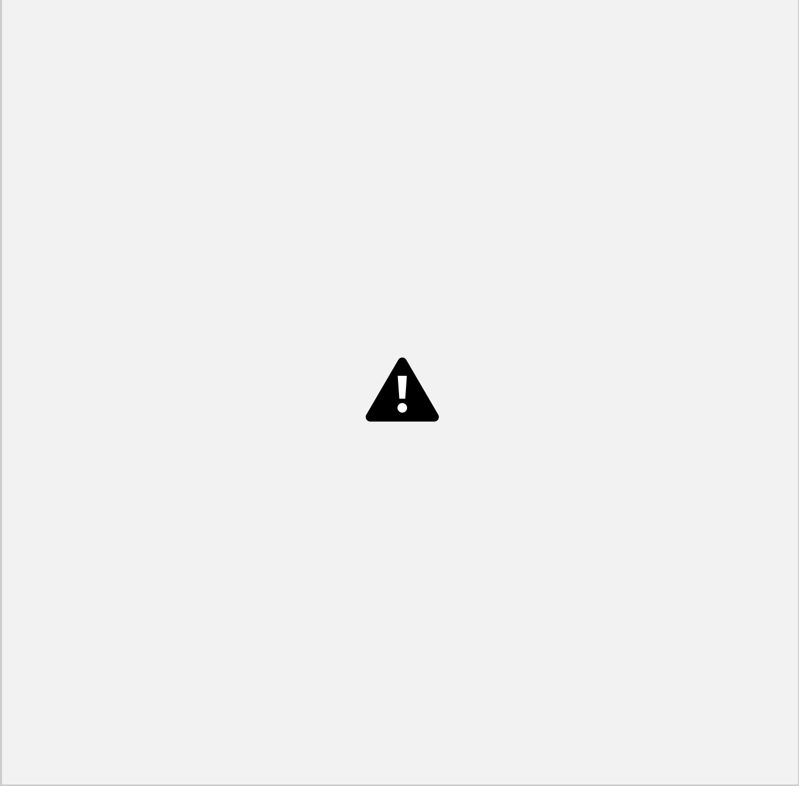
Message almost

there at time T when B
starts – collision!

A B

How can we be sure A knows about the collision?

- How can A know that a collision has taken place?
  - There must be a mechanism to insure retransmission on collision – A's message reaches B at time T

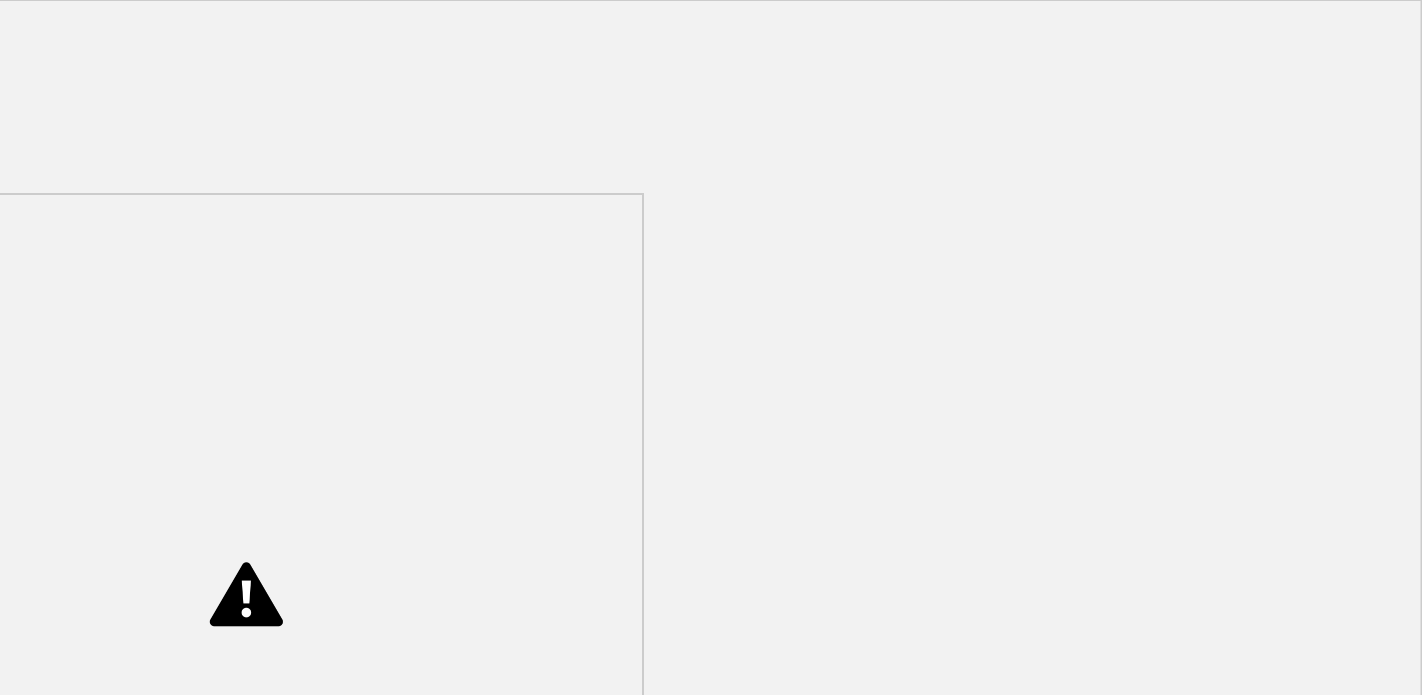


  - B's message reaches A at time 2T
  - So, A must still be transmitting at 2T
- IEEE 802.3 specifies max value of 2T to be 51.2us

Collision Detection (contd…)

- This relates to maximum distance of 2500m between hosts
- At 10Mbps it takes 0.1us to transmit one bit so 512 bits (64B) take 51.2us to send
- So, Ethernet frames must be at least 64B long
  - 14B header, 46B data, 4B CRC
  - Padding is used if data is less than 46B

- Send jamming signal after collision is detected to insure all hosts see collision – 48 bit signal



A B

time = 0

A B

time = T

A B

time = 2T

Exponential Backoff

- If a collision is detected, delay and try again
- Delay time is selected using binary exponential backoff
  – 1st time: choose K from {0,1} then delay = K * 51.2us
  – 2nd time: choose K from

{0,1,2,3} then delay = K * 51.2us

– *nth* time: delay = K x 51.2us, for K=0..2^n – 1

  - Note max value for k = 1023

– give up after several tries (usually 16)

  - Report transmit error to host

- If delay is not random, then there is a chance that sources would retransmit in lock step
- Why not just choose from small set for K

  – This works fine for a small number of hosts

MAC Algorithm from the Receiver Side

Ethernet Frame (as per 802.3

- Senders handle all access control
- Receivers simply read frames with acceptable address –
  - Address to host
  - Address to broadcast
  - Address to multicast to which host belongs
  - All frames if host is in promiscuous mode

standard)

• **Preamble –** informs the receiving system that a frame is starting and enables synchronization. • **SFD (Start Frame Delimiter)** – signifies that the Destination MAC Address field begins with the next byte. • **Destination MAC** – identifies the receiving system.

- **Source MAC** – identifies the sending system.
- **Type** – defines the type of protocol inside the frame, for example IPv4 or IPv6.
- **Data** and **Pad** – contains the payload data. Padding data is added to meet the minimum length requirement for this field (46 bytes).
- **FCS (Frame Check Sequence)** – contains a 32-bit **Cyclic Redundancy Check (CRC)** which allows detection of corrupted data.

Ethernet Address?

Fast and Gigabit Ethernet

- Fast Ethernet (100Mbps) has technology very similar to 10Mbps Ethernet
  - Uses different physical layer encoding (4B5B)
  - Many NIC’s are 10/100

capable

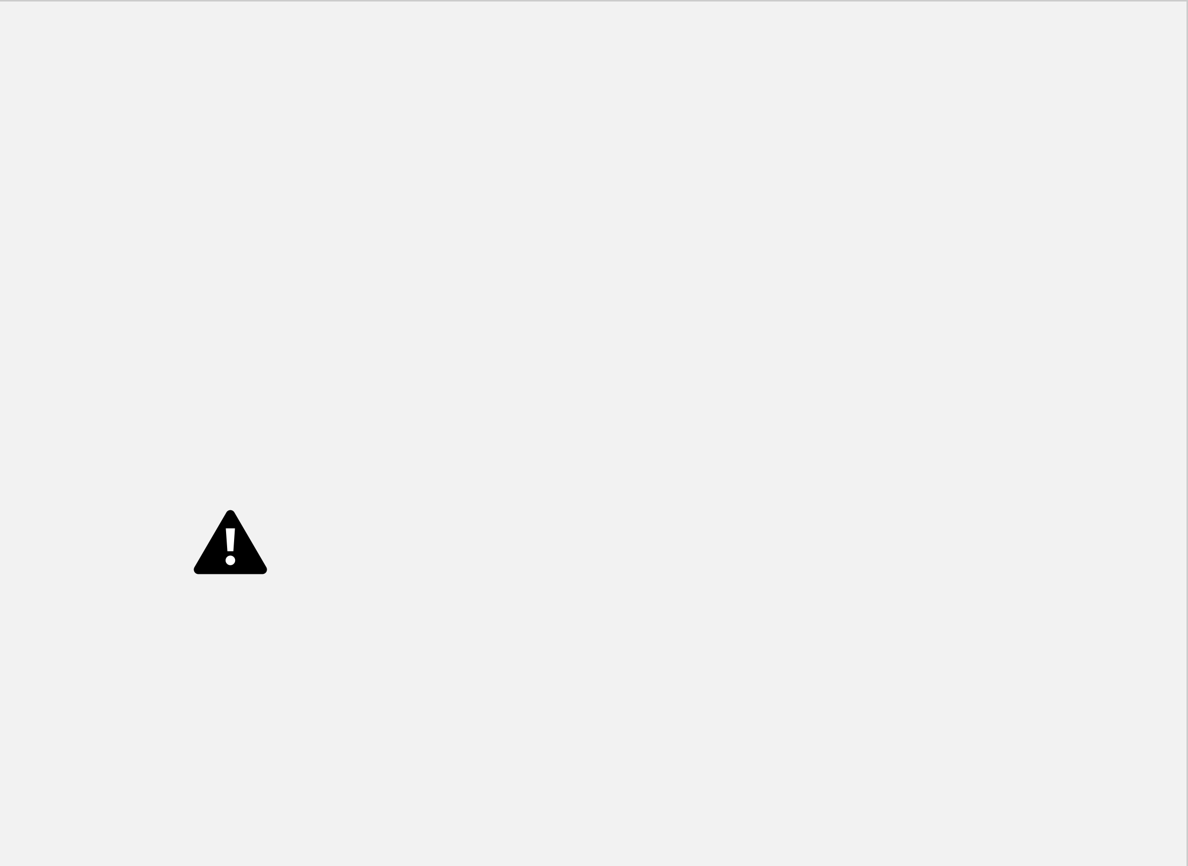
        - Can be used at either speed

- Gigabit Ethernet (1,000Mbps)
    - – Compatible with lower speeds
    - – Uses standard framing and CSMA/CD algorithm
    - – Distances are severely limited
    - – Typically used for backbones and inter-router connectivity
    - – Becoming cost competitive
    - – How much of this bandwidth is realizable?

Ethernet – Practical facts

- Ethernets work best under light loads
  - Utilization over 30% is considered heavy
    - Network capacity is wasted by collisions
- Most networks are limited to about 200 hosts
  - Specification allows for up to 1024
- Most networks are much shorter
  - 5 to 10 microsecond RTT

- Transport level flow control helps reduce load (number of back to back packets)
- Ethernet is inexpensive, fast and easy to administer!



Ethernet Issues

- Ethernet’s peak utilization is pretty low
- Peak throughput worst with
  - More hosts
    - More collisions needed to identify single sender
  - Smaller packet sizes
    - More frequent arbitration
  - Longer links

llisions take longer to observe, more wasted bandwidth –
ciency cane be improved by avoiding these conditions