15-441 Computer Networks

Lecture 6

Link-Layer (2)

Dave Eckhardt

Roadmap

- **Last time**
  - What's a link layer?
- **Today**
  - Ethernet
  - Some things which aren't Ethernet
- **Next time**
  - Ethernet empires
  - What's a switch?

Outline

- **Ethernet**
  - Conceptual history
  - Carrier sense, Collision detection
  - Ethernet history, operation (CSMA/CD)
  - Packet size
  - Ethernet evolution
- **Some things which aren't Ethernet**
  - Token Bus, Token Ring, FDDI, Frame Relay
  - 802.11

Approach

- **A word on approach**
  - We will discuss many “obsolete” technologies
  - This can be a good way to grasp the underlying ideas
    - ...which keep turning up in different contexts
    - A good arrangement of ideas is an easier advance than a genuinely new thing

Reminder: Medium Access Control (MAC)

- **Share a communication medium among multiple**
- **Arbitrate between connected hosts**
- **Goals:**
  - High resource utilization
  - Avoid starvation
  - Simplicity (non-decentralized algorithms)
- **Approaches**
  - Taking turns, random access, really-random access (SS)
  - Random access = allow collisions
    - – Manage & recover from them

Ethernet in Context

- **ALOHA**
  - When you're ready, transmit
  - Detect collisions by waiting (a long time)
  - Recover from collision by trying again
    - ...after a random delay...
      - » Too short, entire network collapses
      - » Too long, every user gets bored
- **Things to try**
  - Slotted ALOHA – reduce collisions (some, not enough)
  - Listen before transmit
  - True collision detection

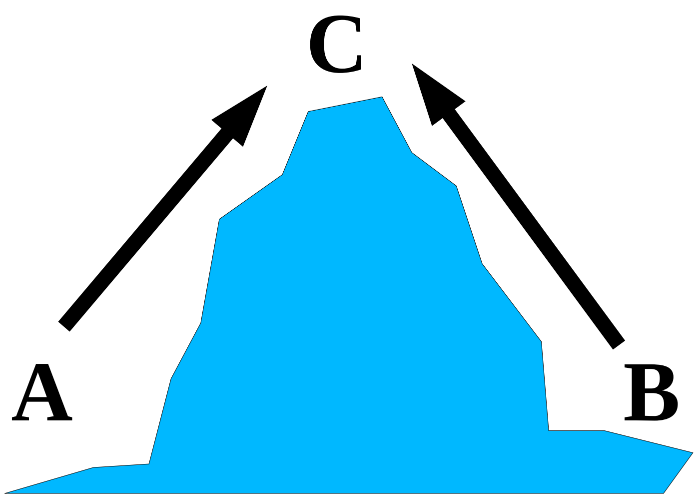
Hui Zhang, Dave Eckhardt

Listen Before Transmit

- **Basic idea**
  - Detect, avoid collisions – *before they happen*
  - Listen before transmit (officical name: "Carrier Sense")
    - Don't start while anybody else is already going
- **Great idea! Why didn't ALOHA do it?**
  - “Hidden terminal problem”

Hidden Terminal Problem

- **A and B are deaf to each other**
  - Can't sense each other's carrier
  - Carrier sense “needs help” in this kind of environment
- **But CS can work really well in an enclosed environment (wire)**



Collision Detection

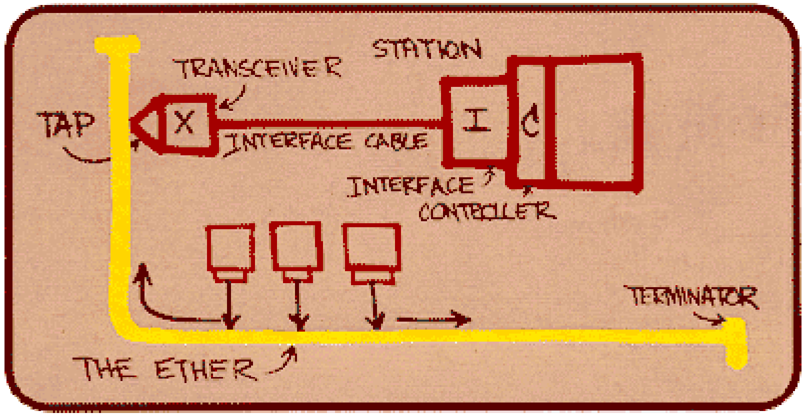
- **Is Carrier sense enough?**
  - Sometimes there is a “race condition”
    - – Two stations listen at the same time
    - – Both hear nothing, start to transmit
    - – Result: collision
      - » Could last “for a while”
      - » Can we detect it while it's happening?
- **Collision Detection**
  - Listen while you transmit
  - If your signal is “messed up”, assume it's due to a collision
- **Great idea! Why didn't ALOHA do it?**

Hui Zhang, Dave Eckhardt

Collision Detection

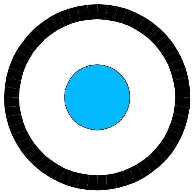
- **Collision detection difficult for radios**
  - “Inverse-square law” relates power to distance
    - – At A, A's transmission drowns out B's
    - – At B, B's transmission drowns out A's
    - – Neither can hear each other, C hears mixture (collision)
  - Many radios disable receiver while transmitting
    - – Huge power of local transmitter may damage receiver
- **Collision detection *can* be done inside a wire**

Original Xerox PARC Ethernet Design



www.ethermanage.com/ethernet

Coaxial cable



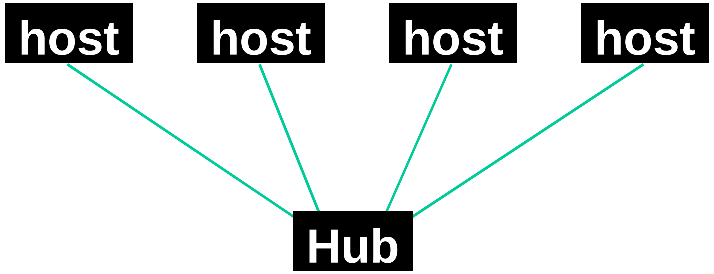
- **Medium – one long cable snaked through your building**
- **Transceiver – fancy radio with collision detection**
- **“Vampire tap”**
  - Drill hole into cable (carefully!)
  - Insert pin to touch center connector (carefully!!)

Original Xerox PARC Ethernet Design

- **Carrier-sense multiple access with collision detection (CSMA/CD).**
  - MA = multiple access
  - CS = carrier sense
  - CD = collision detection
- **PARC Ethernet parameters**
  - 3 Mb/s (to match Xerox Alto workstation RAM throughput)
  - 256 stations (1-byte destination, source addresses)
  - 1 kilometer of cable

802.3 Ethernet

Broadcast technology



- **DEC/Intel/Xerox ("DIX") Ethernet standardized by IEEE**
  - Throughput 3 Mb/s ⇒ 10 Mb/s
  - Station addresses 1 byte ⇒ 6 bytes
- **Growth over the years**
  - Hubs, bridges, switches
  - 100Mbps, 1Gbps, 10Gbps
  - Thin coax, twisted pair, fiber, wireless

CSMA/CD Algorithm

- **Listen for carrier**
- **If carrier sensed, wait until carrier ends.**
  - Sending would force a collision and waste time
- **Send packet while listening for collision.**
- **If no collision detected by end, consider packet delivered.**
- **Otherwise**
  - Abort immediately
    - Transmit "jam signal" (32 bits) to fill cable with errors
  - Perform “exponential back-off” to try packet again.

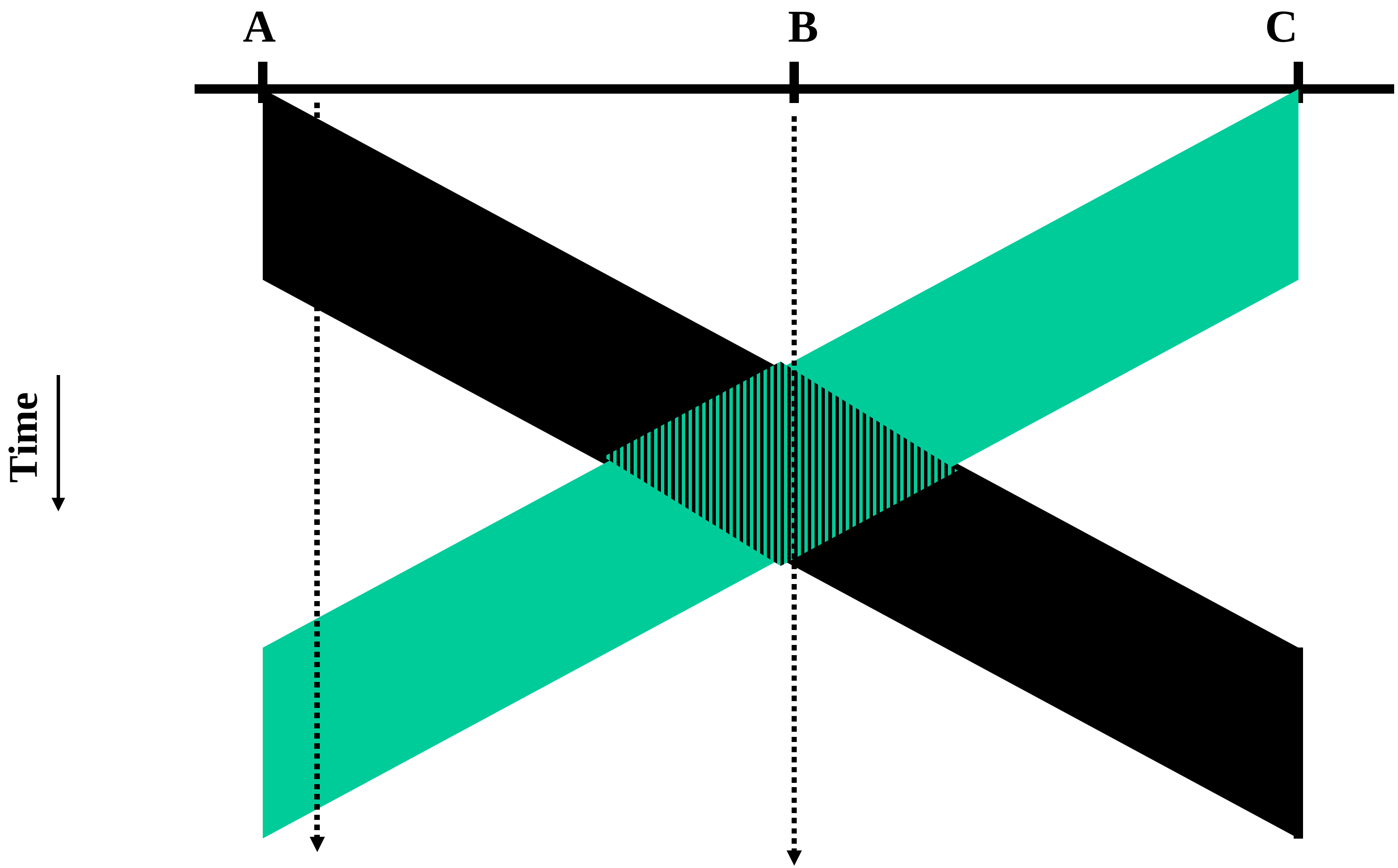
Exponential Back-off

- **Basic idea**
  - Two or more stations want the medium, want to take turns
    - – Idea: each station rolls an N-sided die
      - » Wait that many packet times, try transmitting then
    - – Issues: what is N?
      - » How many stations want the medium?

Exponential Back-off

- **Exponential Back-off**
  - First collision: delay 0 or 1 periods (512 bits)
    - Each station chooses with 50/50% probability
    - Result is 50% probability of resolving the conflict
    - Appropriate if two stations contending for medium
  - Second collision: delay 0...3 periods
    - Will work well if “roughly 4” stations contending
  - Third collision: delay 0...7 times
  - Ten collisions?
    - Give up, tell device driver “transmission failed”

Collision Detection



Hui Zhang, Dave Eckhardt

Collision Detection: Implications

Goal: every node detects collision as it's happening

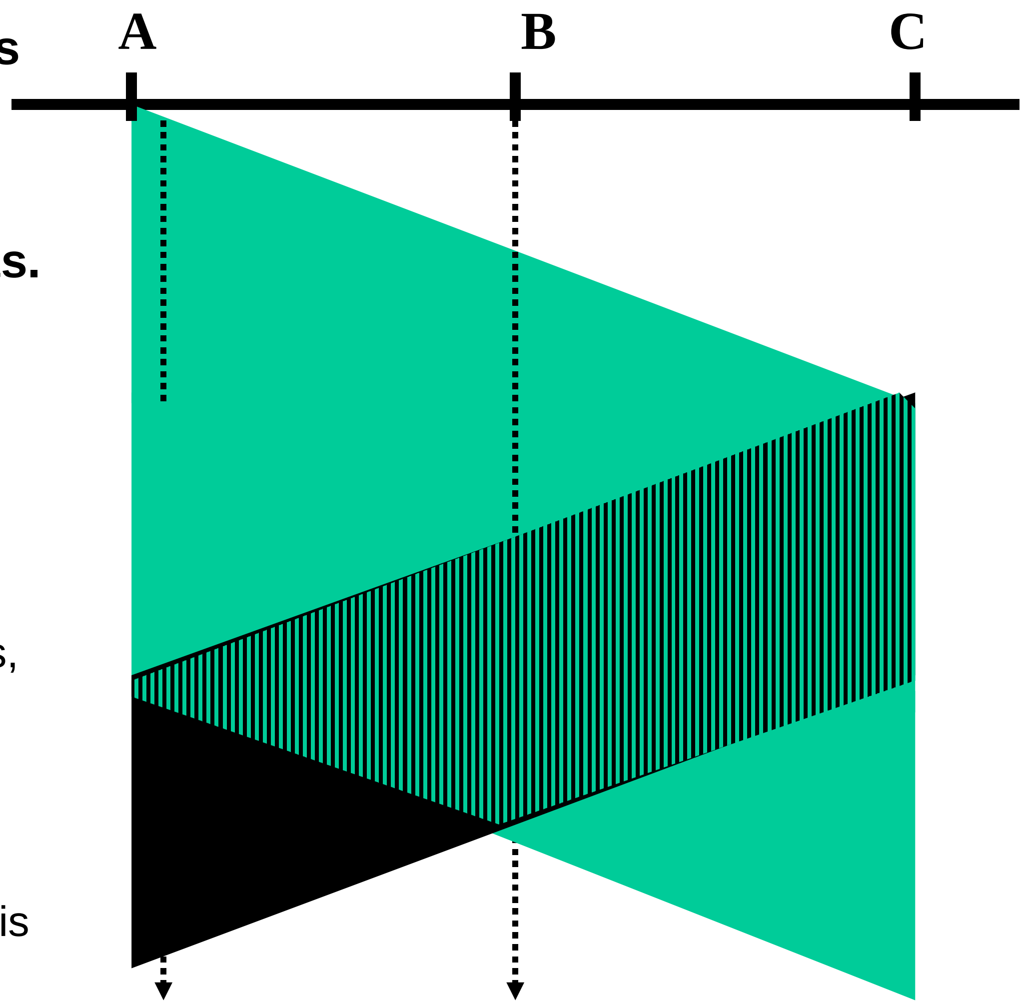
- Any node can be sender

So: need short wires, or long packets.

- Or a combination of both

Can calculate length/distance based on transmission rate and propagation speed.

- Messy: propagation speed is medium-dependent, low-level protocol details, ..
- Minimum packet size is 64 bytes
  - Cable length ~256 bit times
- Example: maximum coax cable length is 2.5 km



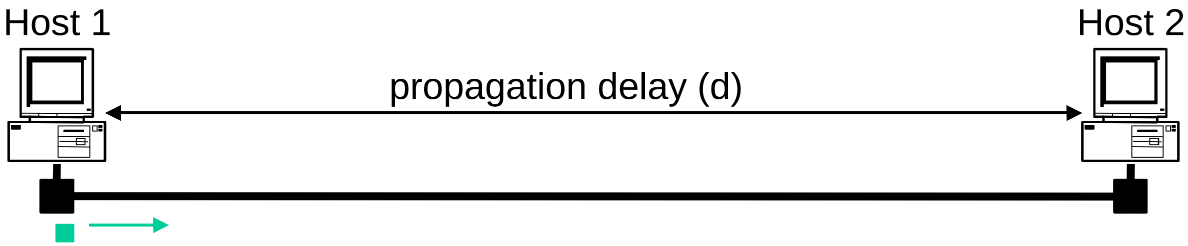
Hui Zhang, Dave Eckhardt

Minimum Packet Size

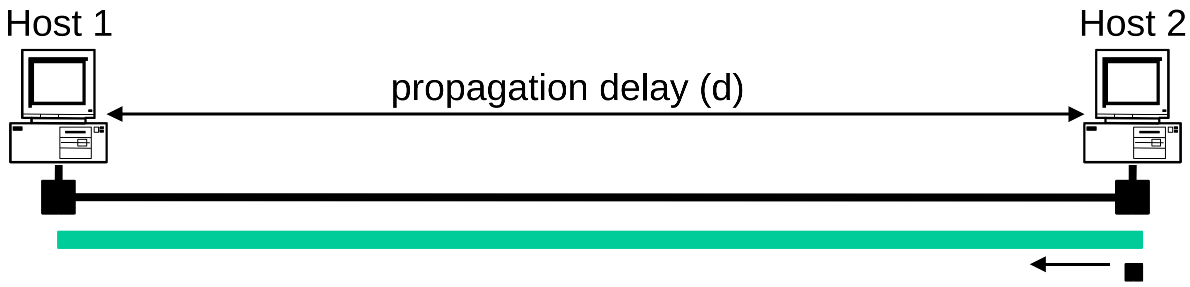
- Why put a minimum packet size?
- Give a host enough time to detect collisions
- In Ethernet, minimum packet size = 64 bytes (two 6-byte addresses, 2-byte type, 4-byte CRC, and 46 bytes of data)
- If host has less than 46 bytes to send, the adaptor pads (adds) bytes to make it 46 bytes
- What is the relationship between minimum packet size and the length of the LAN?

Minimum Packet Size (more)

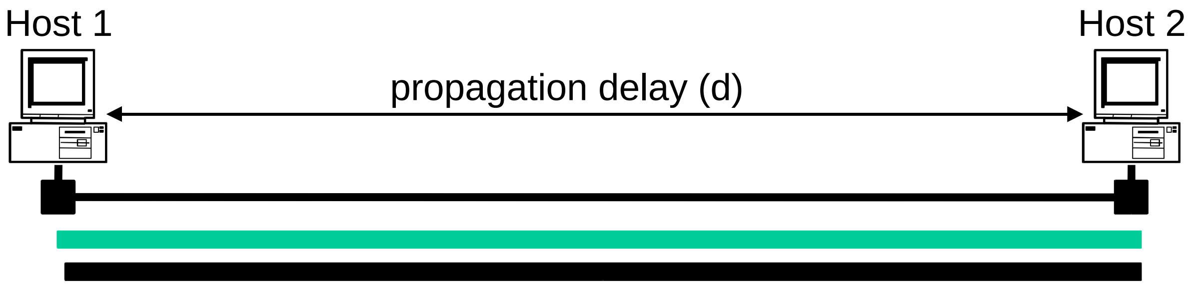
a) Time = t; Host 1 starts to send frame



b) Time = t + d; Host 2 starts to send a frame just before it hears from host 1's frame



c) Time = t + 2*d; Host 1 hears Host 2's frame: detects collision



LAN length = (min_frame_size)*(light_speed)/(2*bandwidth) =

= $(8*64b)*(2*10^8 mps)/(2*10^7\ bps)$ = 5.12 km

Issues: Old (10Base5) Ethernet had 500m segments, repeaters

Hui Zhang, Dave Eckhardt

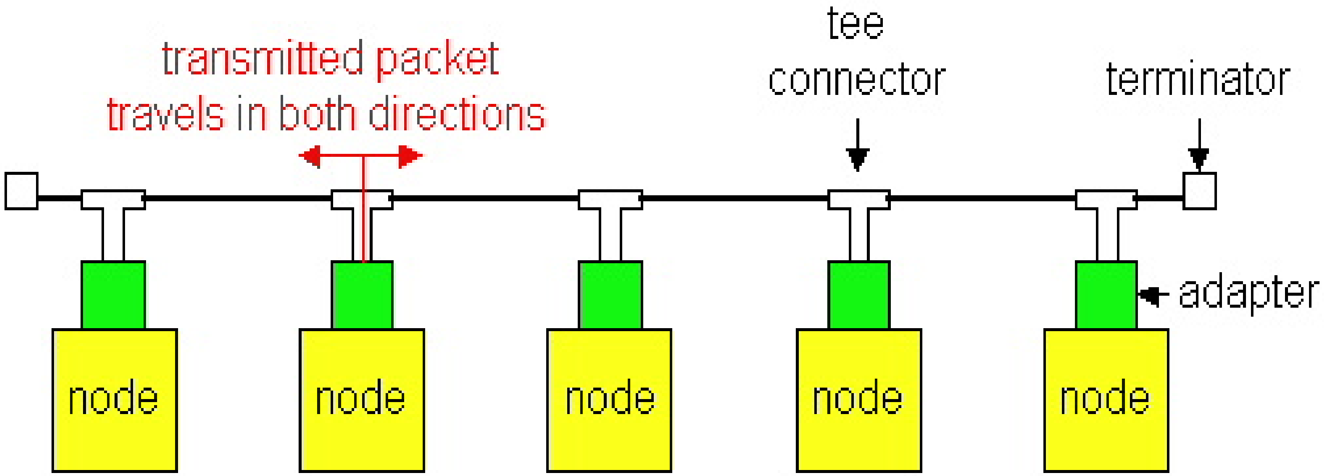
Ethernet Frame Format

8	6	6	2			4
Preamble	**Dest**	**Source**	**Type**	**Data**	**Pad**	**CRC**

- **Preamble marks the beginning of the frame.**
  - Also provides clock synchronization
- **Source and destination are 48 bit IEEE MAC addresses.**
  - Flat address space
  - Hardwired into the network interface
- **Type field (DIX Ethernet) is a demultiplexing field.**
  - Which network (layer 3) protocol should receive this packet?
  - 802.3 uses field as length instead
- **CRC for error checking.**

Ethernet Technologies: 10Base2

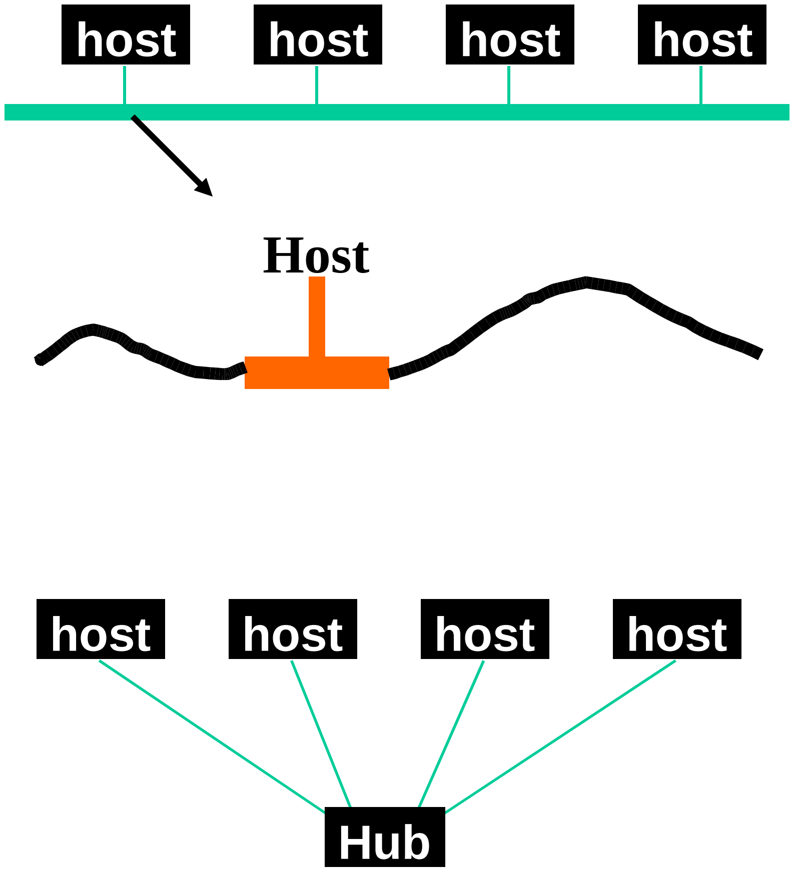
- **10: 10Mbps; 2: under 200 meters max cable length (185 m)**
- **Thin coaxial cable in a bus topology**



- **Repeaters used to connect multiple segments**
- **Repeater repeats bits it hears on one interface to its other interfaces: physical layer device only!**

Compatible Physical Layers

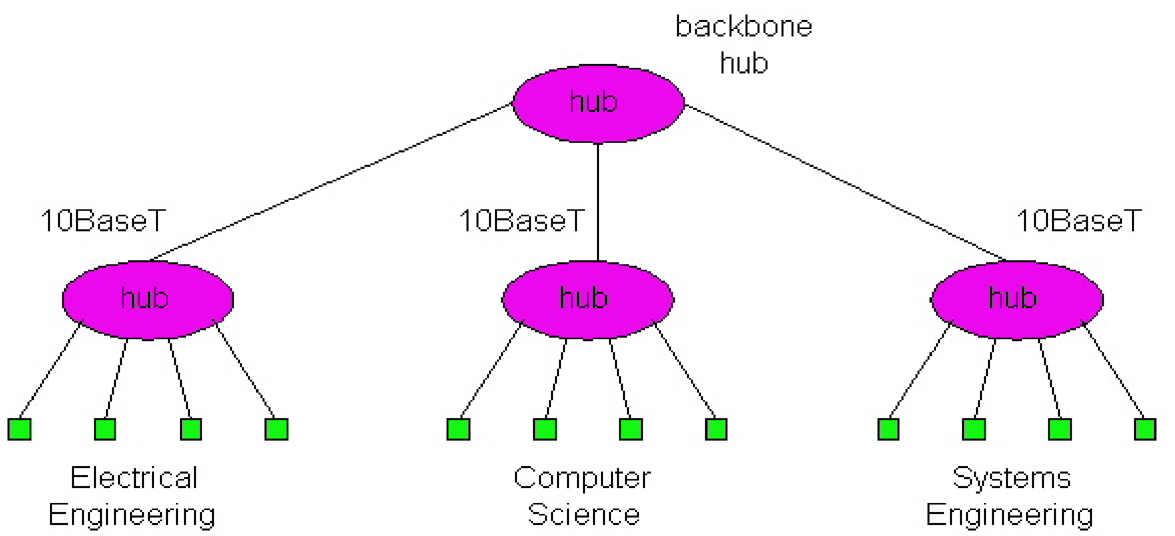
- **10Base2 standard**
  - Thin coax, point-to-point "T" connectors
  - Bus topology
- **10-BaseT: twisted pair**
  - Hub acts as a concentrator
- **3 layers, same protocol!**
  - Key: electrical connectivity between all nodes
  - Deployment is different



Hui Zhang, Dave Eckhardt

10BaseT and 100BaseT

- **10/100 Mbps rate; later called “Fast Ethernet”**
- **T stands for Twisted Pair**
- **Hub to which nodes are connected by twisted pair, thus “star topology”**



Hui Zhang, Dave Eckhardt

10BaseT and 100BaseT (more)

- **Max distance from node to Hub is 100 meters**
- **Hub can disconnect "jabbering" adapter**
- **Hub can gather monitoring information, statistics for display to LAN administrators**

- **Hubs still preserve one collision domain**
  - Every packet is forwarded to all hosts
- **Use *bridges* to address this problem**
  - Bridges forward a packet only to the port leading to the destination

Hui Zhang, Dave Eckhardt

802.3u Fast Ethernet

- **Apply original CSMA/CD medium access protocol at 100Mbps**
- **Must change either minimum frame or maximum diameter: change diameter**
- **Requires**
  - 2 UTP5 pairs (4B5B) or
  - 4 UTP3 pairs (8B6T) or
  - 1 fiber pair
- **No more "shared wire" connectivity.**
  - Hubs and switches only
- **4B/5B encoding**

802.3z Gigabit Ethernet

- **Same frame format and size as Ethernet.**
  - This is what makes it Ethernet
- **Two cases**
  - Shared (broadcast) CSMA/CD – not frequently used
  - Point-to-point links – two stations
    - "Full-duplex": "both sides transmit simultaneously"
    - Added flow control to deal with congestion
- **Choice of a range of fiber and copper transmission media.**
- **“Jumbo frames” (larger than 1500 bytes) allow higher efficiency**

Traditional IEEE 802 Networks: MAC in the LAN and MAN

- **“Ethernet” often considered same as IEEE 802.3.**
  - Not quite identical
- **The IEEE 802.* set of standards defines a common framing and addressing format for LAN protocols.**
  - Simplifies interoperability
  - Addresses are 48 bit strings, with no structure
- **802.3 (Ethernet)**
- **802.5 (Token ring)**
- **802.X (Token bus)**
- **802.6 (Distributed queue dual bus)**
- **802.11 (Wireless)**

LAN Properties

- **Exploit physical proximity**
  - Typically there is a limitation on the physical distance between the nodes
  - E.g., to collect collisions in a contention based network
  - E.g., to limit the overhead introduced by token passing or slot reservations
- **Rely on single administrative control and some level of trust**
  - Broadcasting packets to everybody and hoping everybody (other than the receiver) will ignore the packet
  - Token passing protocols assume everybody plays by the rules

Why Ethernet?

- **Easy to manage**
  - You plug in the host and it basically works
  - No configuration at the datalink layer
- **Broadcast-based**
  - In part explains the easy management
  - Some of the LAN protocols (e.g. ARP) rely on broadcast
    - Networking would be harder without ARP
  - Not having “free” broadcast adds complexity to a LAN
    - Example: ATM
- **Drawbacks.**
  - Broadcast-based: limits throughput since each packet consumes the throughput of the entire network
  - Distance

Outline

- **Ethernet**
  - Conceptual history
  - Carrier sense, Collision detection
  - Ethernet history, operation (CSMA/CD)
  - Packet size
  - Ethernet evolution
- ☞ **Some things which aren't Ethernet**
  - Token Ring, FDDI, Token Bus
  - Wireless, including 802.11

Token Ring

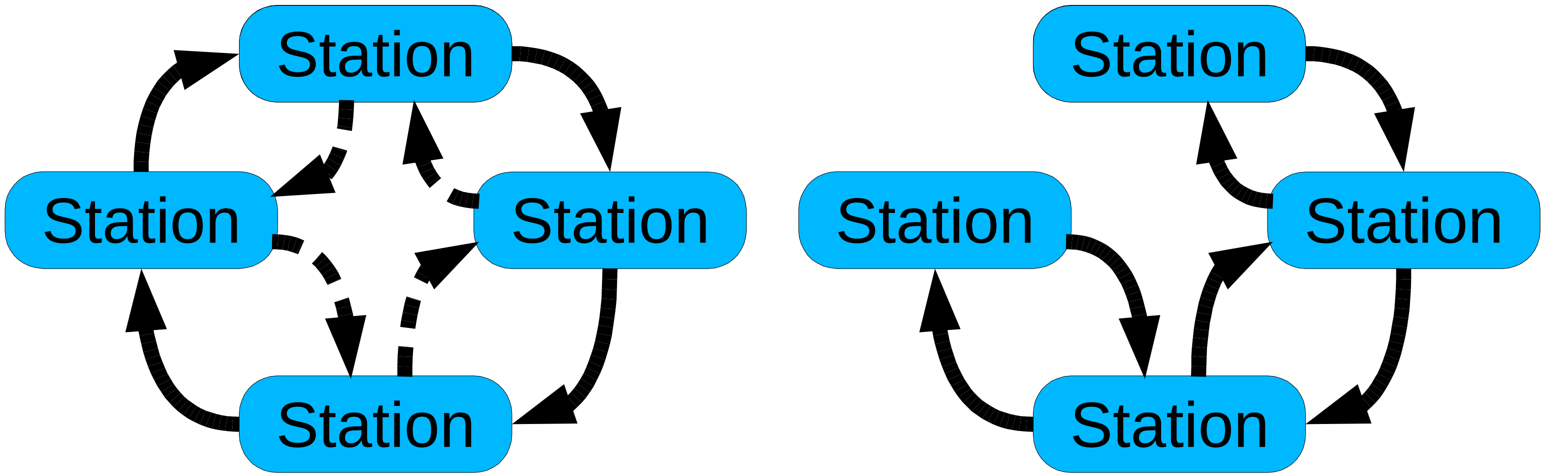
- **IBM Token Ring**
  - Competed with 10 megabit Ethernet
  - 4 megabit transfer rate
    - Lots of heated arguments about how 4 > 10
  - Logical ring
    - Each station has left neighbor, right neighbor
    - Distributed elections, token passing, whee!
  - Physical star
    - Each station actually connected to a box in a closet
    - If you stopped making sense you were cut out of the ring

Hui Zhang, Dave Eckhardt

FDDI

- **Fiber Distributed Data Interface**
  - "Token ring grown up"
  - 100 Mbit/s
  - Nodes connected by fiber
    - – Multi-mode fiber driven by LED
    - – Single-mode fiber driven by laser (long distance)
  - Up to 500 nodes in ring, total fiber length 200 km
  - Organized as *dual ring*

FDDI – Fault Recovery



Token Bus

- **Basic idea**
  - Ethernet is cool
    - ...run one cable throughout building
    - ...popular technology, commodity, cheap
  - Factory automation people worry about frame delay
    - ...must bound delay from sensor to controller to robot
  - Token ring is cool - firm bound on transmission delay
- **Virtual network**
  - Run token-ring protocol on Ethernet frames
    - No collisions, gain delay bound (though generally worse)
  - May be a nested lie: bus atop bridge atop star!

NCR WaveLAN

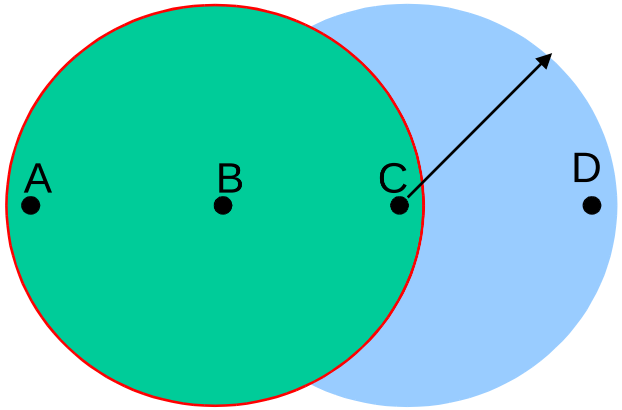
- **Basic idea**
  - Ethernet is cool
    - ..."wireless Ethernet" would be cooler
    - ...re-use addresses, bridging protocols, ...
- **Recall: radio collision detection is hard**
  - Undetected collisions waste a lot of time
  - Hack: collision <u>*inference*</u>
    - Is medium busy when you want to transmit?
      - » Assume true of other stations too
      - » Assume you will all collide when the medium is free
      - » “Back off” pro-actively

Wireless (802.11)

- **Designed for use in limited geographical area (i.e., couple of hundreds of meters)**
- **Designed for three physical media (run at either 1Mbps or 2 Mbps)**
  - Two based on spread spectrum radio
  - One based on diffused infrared

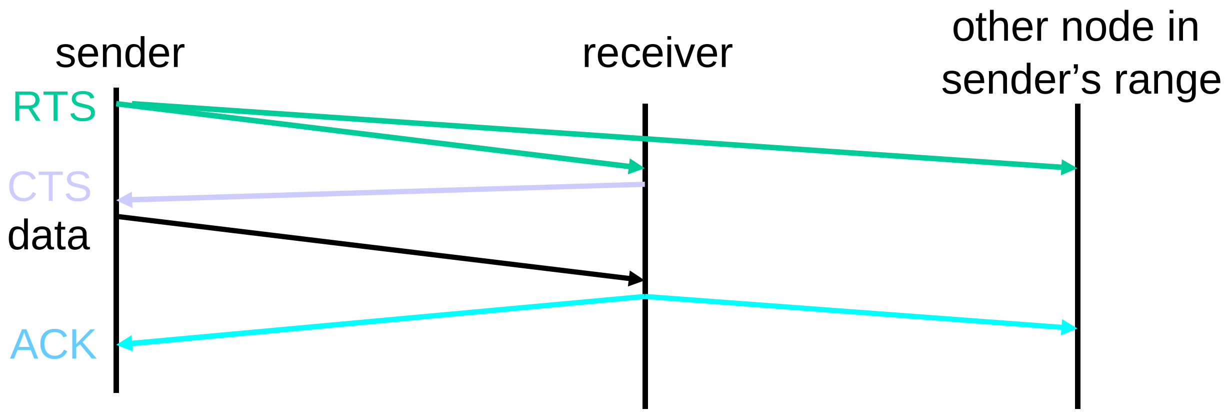
Collision Avoidance: The Problems

- **Reachability is not transitive: if A can reach B, and B can reach C, it doesn't necessary mean that A can reach C**



- **Hidden nodes: A and C send a packet to B; neither A nor C will detect the collision!**
- **Exposed node: B sends a packet to A; C hears this and decides not to send a packet to D (despite the fact that this will not cause interference)!**

Multiple Access with Collision Avoidance (MACA)



- **Before every data transmission**
  - Sender sends a Request to Send (RTS) frame containing the length of the transmission
  - Receiver respond with a Clear to Send (CTS) frame
  - Sender sends data
  - Receiver sends an ACK; now another sender can send data
- **When sender doesn't get a CTS back, it assumes collision**

Summary

- **MAC Problem: arbitrate between multiple hosts sharing a common communication media**
- **Wired solution: Ethernet (use CSMA/CD protocol)**
  - Detect collisions
  - Backoff exponentially on collision
- **Alternate wired "solution"**
  - If you have have enough wires, "full duplex" link (2 links)
- **Wireless solution: 802.11**
  - Use MACA protocol
  - Cannot detect collisions; try to avoid them
  - Distribution system & frame format in discussion sections