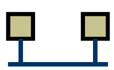
15-441 Computer Networking

Lecture 5 – Ethernet

Brainstorming

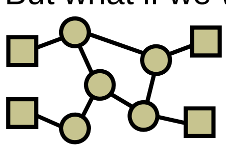
- One of class goals: learning how to design systems
- Take a look at problem and come up with your own solutions first → look at actual designs later
  - Learn to appreciate solution tradeoffs
  - Build confidence in your skills

Problem 1 – Sharing a Wire

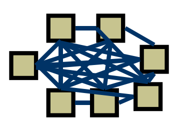


Learned how to connect hosts

- … But what if we want more hosts?

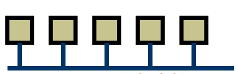


Switches

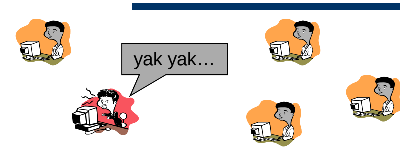


Wires for everybody!

- Expensive! How can we share a wire?

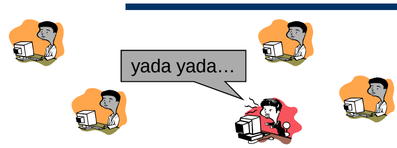


Problem 2 – Listen and Talk



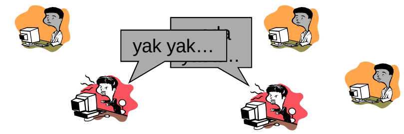
- Natural scheme – listen before you talk…
  - Works well in practice

Problem 2 – Listen and Talk



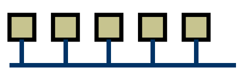
- Natural scheme – listen before you talk…
  - Works well in practice

Problem 2 – Listen and Talk



- Natural scheme – listen before you talk…
  - Works well in practice
- But sometimes breaks down
  - Why? How do we fix/prevent this?

Prob. 3 – Who is this packet for?



- Need to put an address on the packet
- What should it look like?
- How do you determine your own address?
- How do you know what address you want to send it to?

Outline

- Aloha
- Ethernet MAC
- Collisions
- Ethernet Frames
- "Taking Turns" MAC and Other LANs

MAC Protocols: A Taxonomy

Three broad classes:

- Channel partitioning
  - Divide channel into smaller "pieces" (time slots, frequency)
  - Allocate piece to node for exclusive use
- Random access
  - Allow collisions
  - "Recover" from collisions
- "Taking turns"
  - Tightly coordinate shared access to avoid collisions

Goal: efficient, fair, simple, decentralized

Random Access Protocols

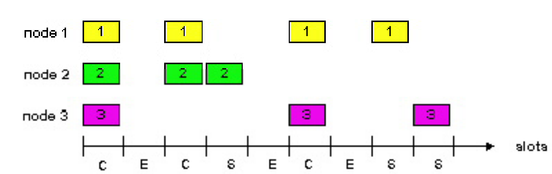
- When node has packet to send
  - Transmit at full channel data rate R
  - No *a priori* coordination among nodes
- Two or more transmitting nodes → "collision"
- Random access MAC protocol specifies:
  - How to detect collisions
  - How to recover from collisions (e.g., via delayed retransmissions)
- Examples of random access MAC protocols:
  - Slotted ALOHA
  - ALOHA
  - CSMA and CSMA/CD

Aloha – Basic Technique

- 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
    - Too long → underutilization

Slotted Aloha

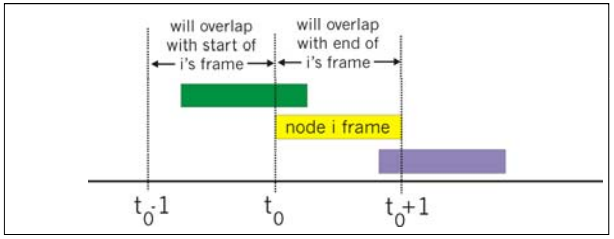
- Time is divided into equal size slots (= pkt trans. time)
- Node (w/ packet) transmits at beginning of next slot
- If collision: retransmit pkt in future slots with probability p, until successful



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

Pure (Unslotted) ALOHA

- Unslotted Aloha: simpler, no synchronization
- Pkt needs transmission:
  - Send without awaiting for beginning of slot
- Collision probability increases:
  - Pkt sent at t_0 collide with other pkts sent in $[t_0-1, t_0+1]$

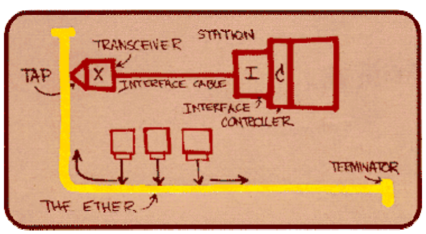


Outline

- Aloha
- Ethernet MAC
- Collisions
- Ethernet Frames
- “Taking Turns” MAC and Other LANs

Ethernet

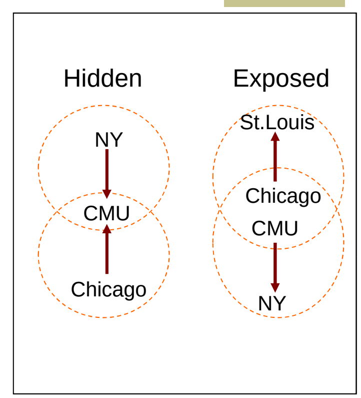
- First practical local area network, built at Xerox PARC in 70’s
- “Dominant” LAN technology:
  - Cheap
  - Kept up with speed race: 10, 100, 1000 Mbps



Metcalfe’s Ethernet sketch

Ethernet MAC – Carrier Sense

- Basic idea:
  - Listen to wire before transmission
  - Avoid collision with active transmission
- Why didn’t ALOHA have this?
  - In wireless, relevant contention at the receiver, not sender
    - Hidden terminal
    - Exposed terminal

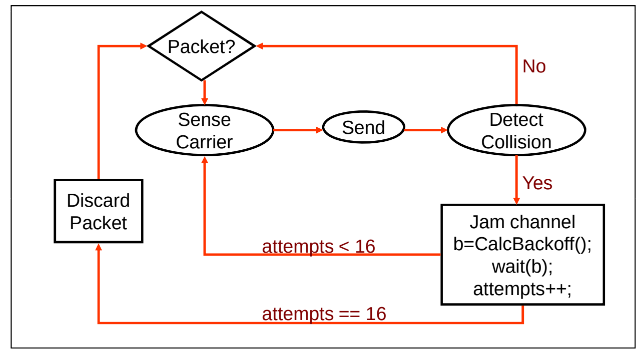


Ethernet MAC – Collision Detection

- Note: ALOHA has collision detection also, should really be called "Fast Collision Detection"
- Basic idea:
  - Listen while transmitting
  - If you notice interference → assume collision
- Why didn't ALOHA have this?
  - Very difficult for radios to listen and transmit
  - Signal strength is reduced by distance for radio
    - Much easier to hear "local, powerful" radio station than one in NY
    - You may not notice any "interference"

Ethernet MAC (CSMA/CD)

- Carrier Sense Multiple Access/Collision Detection



Ethernet's CSMA/CD (more)

Jam Signal: make sure all other transmitters are aware of collision; 48 bits;

Exponential Backoff:

- If deterministic delay after collision, collision will occur again in lockstep
- Why not random delay with fixed mean?
  - Few senders → needless waiting
  - Too many senders → too many collisions
- *Goal*: adapt retransmission attempts to estimated current load
  - heavy load: random wait will be longer

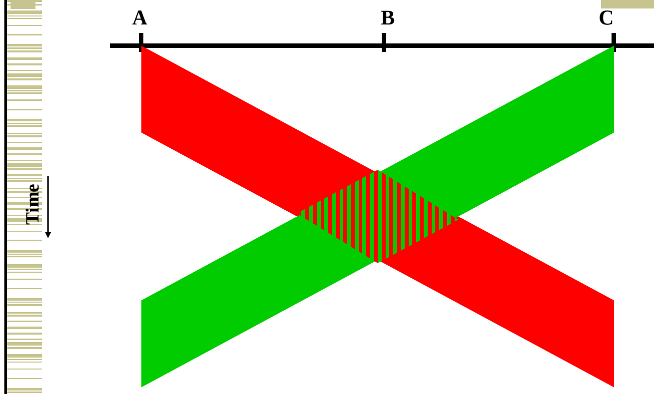
Ethernet Backoff Calculation

- Exponentially increasing random delay
  - Infer senders from # of collisions
  - More senders → increase wait time
- First collision: choose K from {0,1}; delay is K x 512 bit transmission times
- After second collision: choose K from {0,1,2,3}…
- After ten or more collisions, choose K from {0,1,2,3,4,…,1023}

Outline

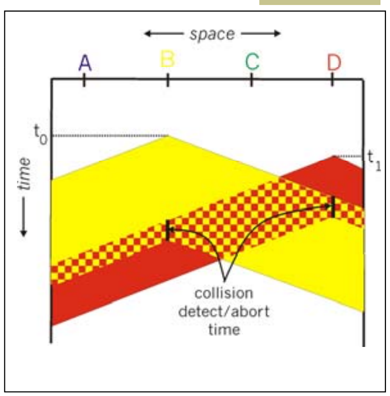
- Aloha
- Ethernet MAC
- Collisions
- Ethernet Frames
- “Taking Turns” MAC and Other LANs

Collisions



Minimum Packet Size

- What if two people sent really small packets
  - How do you find collision?



Ethernet Collision Detect

- Min packet length > 2x max prop delay
  - If A, B are at opposite sides of link, and B starts one link prop delay after A
- Jam network for 32-48 bits after collision, then stop sending
  - Ensures that everyone notices collision

End to End Delay

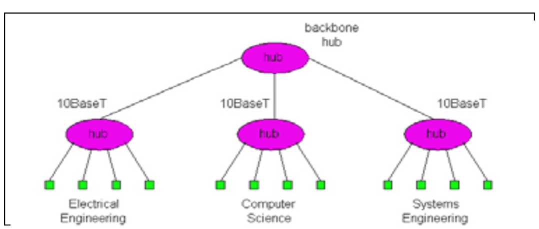
- c in cable = 60% * c in vacuum = 1.8 x 10^8 m/s
- Modern 10Mb Ethernet {
  - 2.5km, 10Mbps
  - ~= 12.5us delay
  - +Introduced repeaters (max 5 segments)
  - Worst case – 51.2us round trip time!
- Slot time = 51.2us = 512bits in flight
  - After this amount, sender is guaranteed sole access to link
  - 51.2us = slot time for backoff

Packet Size

- What about scaling? 3Mbit, 100Mbit, 1Gbit...
  - Original 3Mbit Ethernet did not have minimum packet size → bonus question!
    - Max length = 1Km and No repeaters
  - For higher speeds must make network smaller, minimum packet size larger or both
- What about a maximum packet size?
  - Needed to prevent node from hogging the network
  - 1500 bytes in Ethernet

10BaseT and 100BaseT

- 10/100 Mbps rate; latter called "fast ethernet"
- T stands for Twisted Pair (wiring)
- Minimum packet size requirement
  - Make network smaller → solution for 100BaseT



Gbit Ethernet

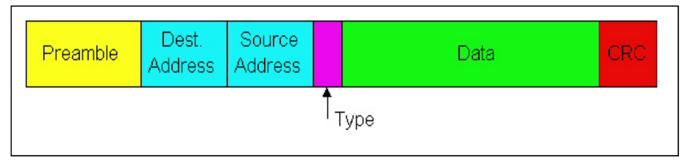
- Minimum packet size requirement
  - Make network smaller?
    - 512bits @ 1Gbps = 512ns
    - 512ns * 1.8 * 10^8 = 92meters = too small !!
  - Make min pkt size larger!
    - Gigabit Ethernet uses collision extension for small pkts and backward compatibility
- Maximum packet size requirement
  - 1500 bytes is not really "hogging" the network
  - Defines "jumbo frames" (9000 bytes) for higher efficiency

Outline

- Aloha
- Ethernet MAC
- Collisions
- Ethernet Frames
- “Taking Turns” MAC and Other LANs

Ethernet Frame Structure

- Sending adapter encapsulates IP datagram (or other network layer protocol packet) in Ethernet frame



Ethernet Frame Structure (cont.)

- Preamble: 8 bytes
  - 101010…1011
  - Used to synchronize receiver, sender clock rates
- CRC: 4 bytes
  - Checked at receiver, if error is detected, the frame is simply dropped

Ethernet Frame Structure (cont.)

- Each protocol layer needs to provide some hooks to upper layer protocols
  - Demultiplexing: identify which upper layer protocol packet belongs to
  - E.g., port numbers allow TCP/UDP to identify target application
  - Ethernet uses Type field
- Type: 2 bytes
  - Indicates the higher layer protocol, mostly IP but others may be supported such as Novell IPX and AppleTalk)

Addressing Alternatives

- Broadcast media → all nodes receive all packets
  - Addressing determines which packets are kept and which are packets are thrown away
  - Packets can be sent to:
    - Unicast – one destination
    - Multicast – group of nodes (e.g. "everyone playing Quake")
    - Broadcast – everybody on wire
- Dynamic addresses (e.g. Appletalk)
  - Pick an address at random
  - Broadcast "is anyone using address XX?"
  - If yes, repeat
- Static address (e.g. Ethernet)

Ethernet Frame Structure (cont.)

- Addresses: 6 bytes
  - Each adapter is given a globally unique address at manufacturing time
    - Address space is allocated to manufacturers
      - 24 bits identify manufacturer
      - E.g., 0:0:15:* → 3com adapter
    - Frame is received by all adapters on a LAN and dropped if address does not match
  - Special addresses
    - Broadcast – FF:FF:FF:FF:FF:FF is "everybody"
    - Range of addresses allocated to multicast
      - Adapter maintains list of multicast groups node is interested in

Outline

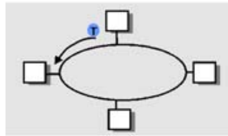
- Aloha
- Ethernet MAC
- Collisions
- Ethernet Frames
- "Taking Turns" MAC and Other LANs

"Taking Turns" MAC Protocols

- Channel partitioning MAC protocols:
  - Share channel efficiently at high load
  - Inefficient at low load: delay in channel access, 1/N bandwidth allocated even if only 1 active node!
- Random access MAC protocols
  - Efficient at low load: single node can fully utilize channel
  - High load: collision overhead
- "Taking turns" protocols
  - Look for best of both worlds!

Token Rings

- Control token passed from one node to next sequentially
  - Fair, real-time bandwidth allocation
  - Every host holds token for limited time
- Concerns:
  - Token overhead
  - Latency
  - Single point of failure (token)



- Fiber Distributed Data Interface (FDDI)
  - Optical version of 802.5 token ring
  - 100 Mbps, 100km
  - Optional dual ring for fault tolerance.

Why Did Ethernet Win?

- Failure modes
  - Token rings – network unusable
  - Ethernet – node detached
- Good performance in common case
- Volume → lower cost → higher volume ….
- Adaptable
  - To higher bandwidths (vs. FDDI)
  - To switching (vs. ATM)
- Completely distributed, easy to maintain/administer
- Easy incremental deployment
- Cheap cabling, etc

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 natural broadcast capabilities adds complexity to a LAN
    - Example: ATM
- Drawbacks.
  - Broadcast-based: limits bandwidth since each packets consumes the bandwidth of the entire network
  - Distance

Summary

- CSMA/CD → carrier sense multiple access with collision detection
  - Why do we need exponential backoff?
  - Why does collision happen?
  - Why do we need a minimum packet size?
    - How does this scale with speed?
- Ethernet
  - What is the purpose of different header fields?
  - What do Ethernet addresses look like?
- What are some alternatives to Ethernet design?

EXTRA SLIDES

Outline

- Random Access Analysis

Slotted Aloha Efficiency

Q: What is max fraction slots successful?

A: Suppose N stations have packets to send

- Each transmits in slot with probability *p*
- Prob. successful transmission *S* is:

by single node: $S = p\,(1-p)^{(N-1)}$

by any of N nodes

S = Prob (only one transmits)

$= N\,p\,(1-p)^{(N-1)}$

… choosing optimum p as N -> infty …

… p = 1/N

$= 1/e = .37$ *as N -> infty*

At best: channel use for useful transmissions 37% of time!

Pure Aloha (cont.)

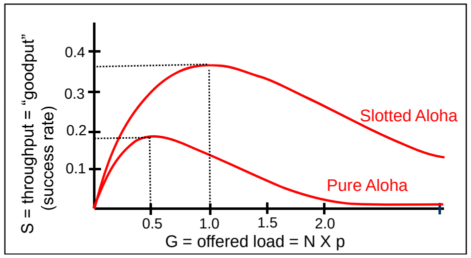
P(success by given node) = P(node transmits) X P(no other node transmits in $[p_0-1,p_0]$ X P(no other node transmits in $[p_0-1,p_0]$

$= p \times (1-p)^{(N-1)} \times (1-p)^{(N-1)}$

P(success by any of N nodes) $= N\,p \times (1-p)^{(N-1)} \times (1-p)^{(N-1)} = 1/(2e) = .18$

… choosing optimum p as N → infty → p = 1/2N …

protocol constrains effective channel throughput!



Simple Analysis of Efficiency

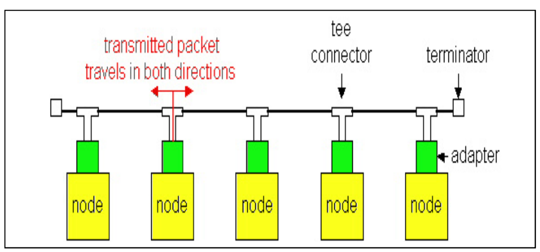
- Key assumptions
  - All packets are same, small size
    - Packet size = size of contention slot
  - All nodes always have pkt to send
  - p is chosen carefully to be related to N
    - p is actually chosen by exponential backoff
  - Takes full slot to detect collision (I.e. no “fast collision detection”)

Ethernet Problems

- Key concern: Ethernet (like Aloha) is unstable at high loads
  - Peak utilization approx. = 1/e = 37%
- Peak throughput worst with
  - More hosts – more collisions needed to identify single sender
  - Smaller packet sizes – more frequent arbitration
  - Longer links – collisions take longer to observe, more wasted bandwidth
  - Can improve efficiency by avoiding these conditions
    - Works well in practice

Ethernet Technologies: 10Base2

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



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

Gbit Ethernet

- Use standard Ethernet frame format
- Allows for point-to-point links and shared broadcast channels
- In shared mode, CSMA/CD is used; short distances between nodes to be efficient
- Uses hubs, called here “Buffered Distributors”
- Full-Duplex at 1 Gbps for point-to-point links

Virtual LANs

- Single physical LAN infrastructure that carries multiple "virtual" LANs simultaneously.
- Each virtual LAN has a LAN identifier in the packet.
  - Switch keeps track of what nodes are on each segment and what their virtual LAN id is
- Can bridge and route appropriately.
- Broadcast packets stay within the virtual LAN.
  - Limits the collision domain for the packet