15-441 Computer Networking
Lecture 6
Data link Layer – Access Control

Peter Steenkiste

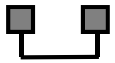
Fall 2010

www.cs.cmu.edu/~prs/15-441-F10

Datalink Functions

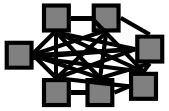
- **Framing: encapsulating a network layer datagram into a bit stream.**
  - **Add header, mark and detect frame boundaries, …**
- **Error control: error detection and correction to deal with bit errors.**
  - **May also include other reliability support, e.g. retransmission**
- **Flow control: avoid sender overrunning receiver.**
- **Media access: controlling which frame should be sent over the link next.**
  - **Easy for point-to-point links**
  - **Harder for multi-access links: who gets to send?**

So far …



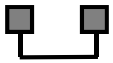
Can connect two nodes

- … But what if we want more nodes?



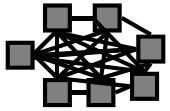
Wires for everybody!

So far …



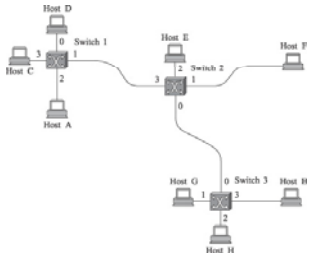
Can connect two nodes

- … But what if we want more nodes?



Wires for everybody!

Datalink Architectures



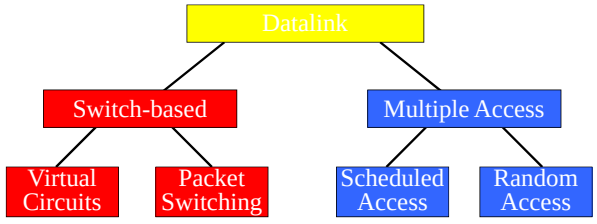
- Point-Point with switches

- Multiple access networks
- Media access control

Media Access Control

- How do we transfer packets between two hosts connected to the same network?
- Switches connected by point-to-point links -- store-and-forward.
  - » Used in WAN, LAN, and for home connections
  - » Conceptually similar to "routing"
    - – But at the datalink layer instead of the network layer
- Multiple access networks -- contention based.
  - » Multiple hosts are sharing the same transmission medium
  - » Used in LANs and wireless
  - » Need to control access to the medium

Datalink Classification

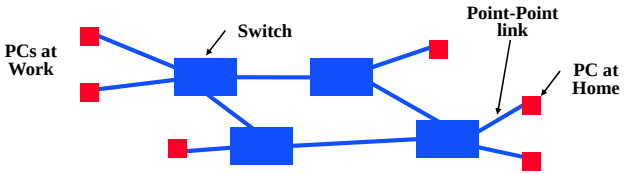


Switching

- Forward units of data based on address in header.
- Many data-link technologies use switching.
  - » Virtual circuits: Frame Relay, ATM, X.25, ..
  - » Packets: Ethernet, ...
- "Switching" also happens at the network layer.
  - » Layer 3: Internet protocol
  - » In this case, address is an IP address
  - » IP over SONET, IP over ATM, ...
  - » Otherwise, operation is very similar
- Switching is different from SONET mux/demux.
  - » SONET channels statically configured - no addresses

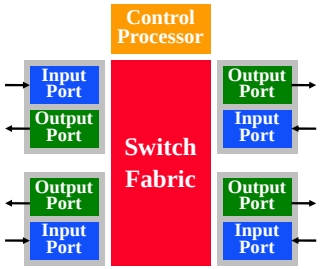
A Switch-based Network

- Switches are connected by point-point links.
- Packets are forwarded hop-by-hop by the switches towards the destination.
  - » Forwarding is based on the address
- How does a switch work?
- How do nodes exchange packets over a link?
- How is the destination addressed?



Switch Architecture

- Packets come in one interface, forwarded to output interface based on address.
  - » Same idea for bridges, switches, routers: address look up differs
- Control processor manages the switch and executes higher level protocols.
  - » E.g. routing, management, ...
- The switch fabric directs the traffic to the right output port.
- The input and output ports deal with transmission and reception of packets.

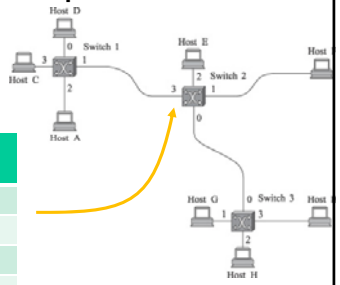


Connections or Not?

- Two basic approaches to packet forwarding
  - »Connectionless
  - »(virtual) Circuit switched
- When would you use?

Connectionless

- Host can send anytime anywhere
- No idea if resources are available to get to dest
- Forwarding is independent for each packet
- No setup time
- Fault tolerant
  - » More on this later

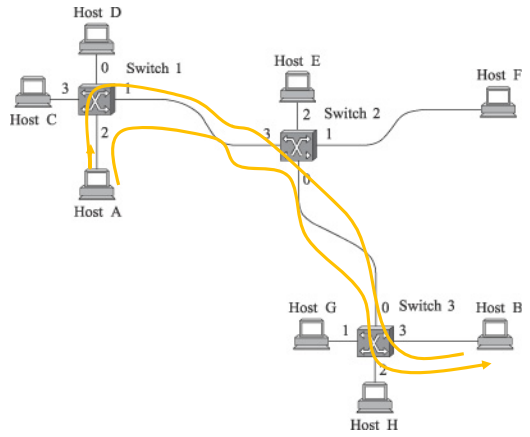


Destination	Port
A	3
B	0
C	
D	
E	

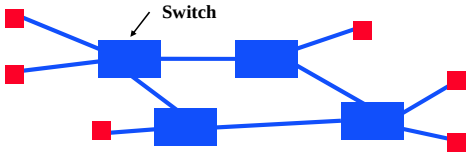
Virtual Circuit Switching

- Two stage process
  - » Setup connection (create VCIs)
  - » Send packets
- RTT introduced before any data is sent
- Per packet overhead can be smaller (VCI << adr)
- Switch failures are hard to deal with
- Reserves resources for connection

Setup, assign VCIs



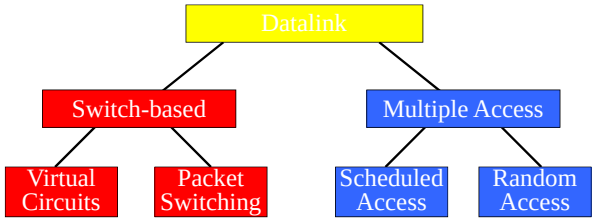
Packet Forwarding: Address Lookup



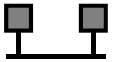
Address	Next Hop	Info
B3112381250	3	13
38913C3C213	3	-
A21023C9059	0	-
128.2.15.1	1	(2,34)

- Address from header.
  - » Absolute address (e.g. Ethernet)
  - » (IP address for routers)
  - » (VC identifier, e.g. ATM))
- Next hop: output port for packet.
- Info: priority, VC id, ..
- Table is filled in by protocol.

Datalink Classification

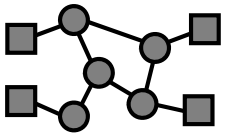


Problem: Sharing a Wire

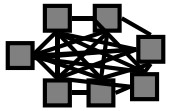


Learned how to connect hosts

- **… But what if we want more hosts?**

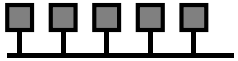


Switches

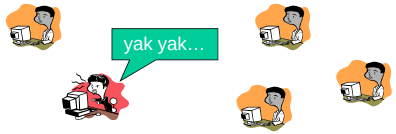


Wires for everybody!

- **Expensive! How can we share a wire?**



Listen and Talk



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

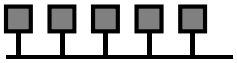
Listen and Talk

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

Listen and Talk

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

Problem: 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

Random Access Protocols

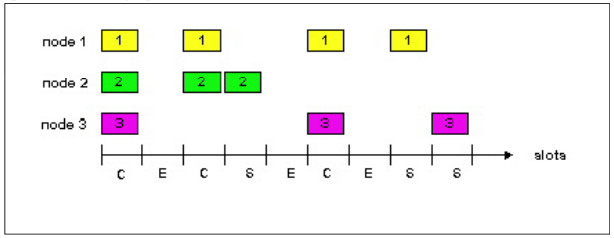
- When node has packet to send
  - Transmit at full channel data rate R
  - No *a priori* coordination among nodes
- Two or more transmitting nodes → "collision"
- Random access MAC protocol specifies:
  - How to detect collisions
  - How to recover from collisions (e.g., via delayed retransmissions)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

Aloha – Basic Technique

- First random MAC developed
  - For radio-based communication in Hawaii (1970)
- Basic idea:
  - When you are ready, transmit
  - Receivers 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

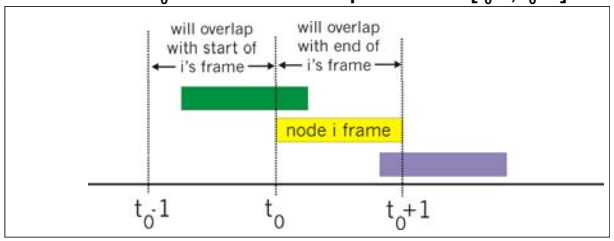
Slotted Aloha

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



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 $[t_0-1, t_0+1]$

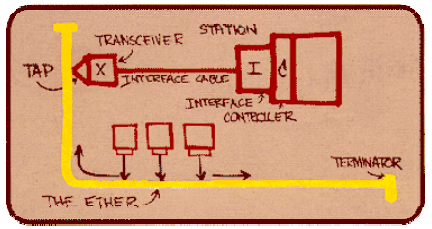


Outline

- Aloha
- Ethernet MAC
- Collisions
- Ethernet Frames

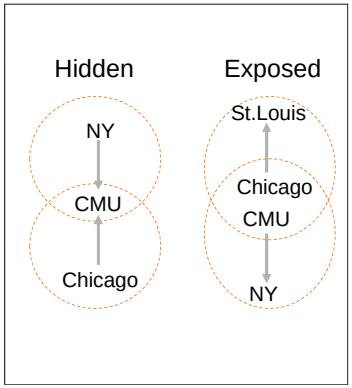
Ethernet

- 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



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**

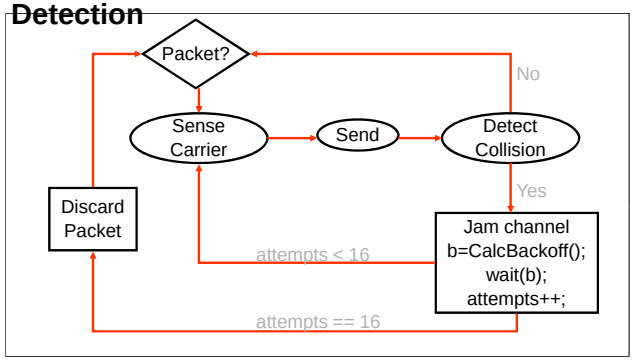


Ethernet MAC – Collision Detection

- **But: ALOHA has collision detection also?**
  - » **That was very slow and inefficient**
- **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 CSMA/CD: Making it word

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**

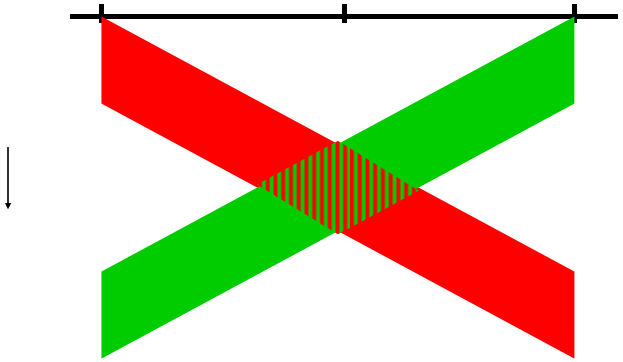
Ethernet Backoff Calculation

- 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

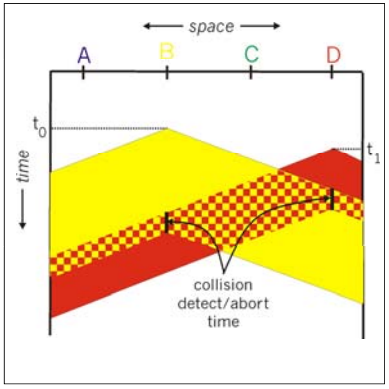
- Aloha
- Ethernet MAC
- Collisions
- Ethernet Frames

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

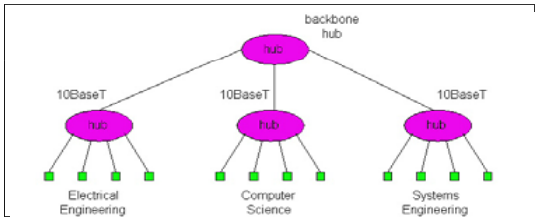
- c in cable = 60% * c in vacuum = 1.8 x 10^8 m/s
- Modern 10Mb Ethernet
  - 2.5km, 10Mbps
  - ~= 12.5us delay
  - +Introduced repeaters (max 5 segments)
  - Worst case – 51.2us round trip time!
- Slot time = 51.2us = 512bits in flight
  - After this amount, sender is guaranteed sole access to link
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

Packet Size

- What about scaling? 3Mbit, 100Mbit, 1Gbit...
  - Original 3Mbit Ethernet did not have minimum packet size
    - 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

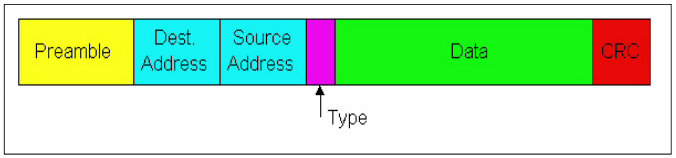
- Minimum packet size requirement
  - Make network smaller?
    - 512bits @ 1Gbps = 512ns
    - 512ns * 1.8 * 10^8 = 92meters = too small !!
  - Make min pkt size larger!
    - Gigabit Ethernet uses collision extension for small pkts and backward compatibility
- Maximum packet size requirement
  - 1500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

Outline

- Aloha
- Ethernet MAC
- Collisions
- Ethernet Frames

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 → 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 Address Assignment

- Each adapter is given a globally unique 6-byte 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

Why Did Ethernet Win?

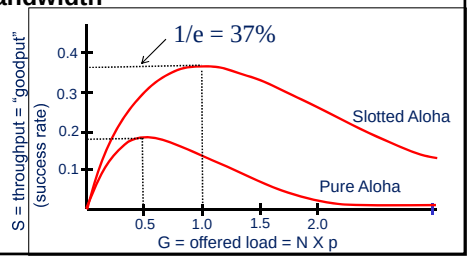
- Failure modes
  - » Token rings – network unusable
  - » Ethernet – node detached
- Good performance in common case
  - » Deals well with bursty traffic
  - » Usually used at low load
- Volume → lower cost → higher volume ….
- Adaptable
  - » To higher bandwidths (vs. FDDI)
  - » To switching (vs. ATM)
- Easy incremental deployment
- Cheap cabling, etc

And .. It is Easy to Manage

- You plug in the host and it basically works
  - » No configuration at the datalink layer
  - » Today: may need to deal with security
- Protocol is fully distributed
- 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

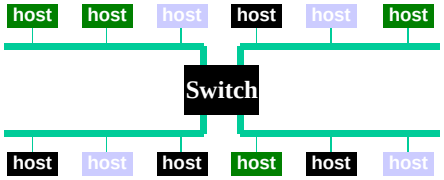
Ethernet Problems: Unstable at High Load

- Peak throughput worst with
  - » More hosts – more collisions to identify single sender
  - » Smaller packet sizes – more frequent arbitration
  - » Longer links – collisions take longer to observe, more wasted bandwidth
- But works well
  - » Can improve efficiency by avoiding above conditions



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



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?