CS 640: Introduction to Computer Networks

Aditya Akella

Lecture 6 -
Ethernet, Multiple Access and Bridging

The Road Ahead

- Multiple access protocols
  - Ethernet's CSMA/CD
- Bridging
- Spanning tree protocol

Multiple Access Protocols

- Prevent two or more nodes from transmitting at the same time over a *broadcast* channel.
  - If they do, we have a *collision*, and receivers will not be able to interpret the signal
- Several classes of multiple access protocols.
  - Partitioning the channel, e.g. frequency-division or time division multiplexing
  - Taking turns, e.g. token-based, reservation-based protocols, polling based
  - Contention based protocols, e.g. Aloha, Ethernet

Desirable MAC Properties

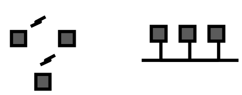
Broadcast channel of capacity R bps.

- 1 node ➔ throughput = R bps
- N nodes ➔ throughput = R/N bps, on average
- Decentralized
- Simple, inexpensive

Contention-Based Protocols

- Idea: access the channel in a "random" way - when collisions occur, recover.
  - Each node transmits at highest rate of R bps
  - Collision: two or more nodes transmitting at the same time
    - Each node retransmits until collided packet gets through
  - Key: don't retransmit right away
    - Wait a random interval of time first
- Examples
  - Aloha
  - Ethernet – focus today

Ethernet History



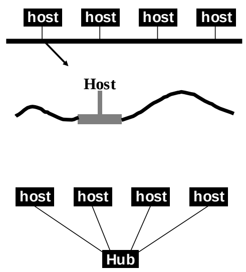
Aloha packet radio

Ethernet on coax
10base-2 (thinnet)
10base-5 (thicknet)

- 1978: 10-Mbps Ethernet standard defined
- Later adopted and generalized to the 802.3 IEEE standard
- 802.3 defined a much wider set of media
  - Also several recent extensions (covered later)
- We will focus on 10Mbps Ethernet, since it is commonly used for multi-access
  - Faster versions more for point to point links

Ethernet Physical Layer

- 10Base2 standard based on thin coax → 200m
  - Nodes are connected using thin coax cables and BNC "T" connectors in a bus topology
  - Thick coax no longer used
- 10BaseT uses twisted pair and hubs → 100m
  - Stations can be connected to each other or to hubs
  - Hub acts as a concentrator
    - Dumb device
- The two designs have the same protocol properties.
  - Key: electrical connectivity between all nodes
  - Deployment is different



Ethernet Frame Format

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

- Preamble marks the beginning of the frame.
  - Also provides synchronization
- Source and destination are 48 bit IEEE MAC addresses.
  - Flat address space
  - Hardwired into the network interface
- Type field is a demultiplexing field.
  - What network layer (layer 3) should receive this packet?
- Max frame size = 1500B; min = 46B
  - Need padding to meet min requirement
- CRC for error checking.

Ethernet host side

- Transceiver: detects when the medium is idle and transmits the signal when host wants to send
  - Connected to "Ethernet adaptor"
  - Sits on the host
- Any host signal broadcast to everybody
  - But transceiver accepts frames addressed to itself
  - Also frames sent to broadcast medium
  - All frames, if in promiscuous mode
- When transmitting, all hosts on the same segment, or connected to the same hub, compete for medium
  - Same collision domain
  - Bad for efficiency!

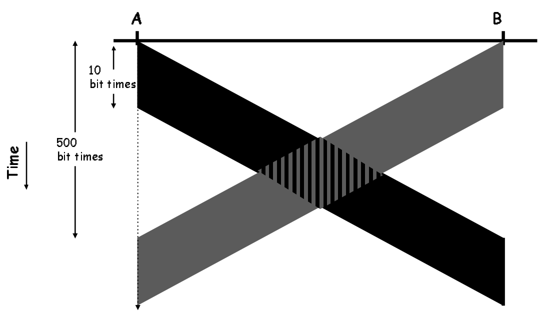
Sender-side: MAC Protocol

- Carrier-sense multiple access with collision detection (CSMA/CD).
  - MA = multiple access
  - CS = carrier sense
  - CD = collision detection

CSMA/CD Algorithm Overview

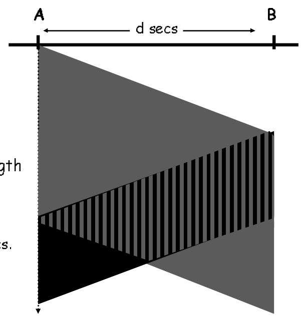
- Sense for carrier.
  - "Medium idle"?
- If medium busy, wait until idle.
  - Sending would force a collision and waste time
- Send packet and sense for collision.
- If no collision detected, consider packet delivered.
- Otherwise, abort immediately, perform *exponential back off* and send packet again.
  - Start to send after a random time picked from an interval
  - Length of the interval increases with every collision, retransmission attempt

Collision Detection



Collision Detection: Implications

- All nodes must be able to detect the collision.
  - Any node can be sender
- => Must either have short wires, long packets, or both
- If A starts at t, and wirelength is d secs,
  - In the worst case, A may detect collision at t+2d
  - ➔ Will have to send for 2d secs.
  - ➔ d depends on max length of ethernet cable



Minimum Packet Size

- Give a host enough time to detect a collision.
- In Ethernet, the minimum packet size is 64 bytes.
  - 18 bytes of header and 46 data bytes
  - If the host has less than 46 bytes to send, the adaptor pads bytes to increase the length to 46 bytes
- What is the relationship between the minimum packet size and the size of LAN?

*LAN size = (min frame size) * light speed / (2 * bandwidth)*

- How did they pick the minimum packet size?

CSMA/CD: Some Details

- When a sender detects a collision, it sends a "jam signal".
  - Make sure that all nodes are aware of the collision
  - Length of the jam signal is 32 bit times
  - Permits early abort - don't waste max transmission time
- Exponential backoff operates in multiples of 512 bit times.
  - RTT= 256bit times ➔ backoff time > Longer than a roundtrip time
  - Guarantees that nodes that back off will notice the earlier retransmission before starting to send
- Successive frames are separated by an "inter-frame" gap.
  - to allow devices to prepare for reception of the next frame
  - Set to 9.6 µsec or 96 bit times

Why Ethernet?

- Easy to manage.
  - You plug in the host and it basically works
  - No configuration at the datalink layer
- Cheap
  - No switches; only cables
- Broadcast-based.
  - In part explains the easy management
  - Some of the LAN protocols rely on broadcast
    - Resource discovery
    - Decide discovery (ARP)
    - Naturally fit with broadcast
  - Not having natural broadcast capabilities adds a lot of complexity to a LAN
- Drawbacks.
  - Broadcast-based: limits bandwidth since each packets consumes the bandwidth of the entire network
  - Works best under light loads
    - Limit on number of hosts
    - Distance

802.3u Fast Ethernet

- Apply original CSMA/CD medium access protocol at 100Mbps
- Must change either minimum frame or maximum diameter: change diameter
- No more "shared wire" connectivity.
  - Hubs and switches only

802.3z Gigabit Ethernet

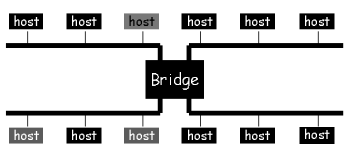
- Same frame format and size as Ethernet.
  - This is what makes it Ethernet
- Full duplex point-to-point links in the backbone are likely the most common use.
  - Added flow control to deal with congestion
- Alternative is half-duplex shared-medium access.
  - Cannot cut the diameter any more (set to 200m)
  - Raise the frame size to 512B
- Choice of a range of fiber and copper transmission media.
- Defining "jumbo frames" for higher efficiency.

LAN Properties

- Exploit physical proximity.
  - Often a limitation on the physical distance
  - E.g. to detect collisions in a contention based network
- Relies on single administrative control and some level of trust.
  - Broadcasting packets to everybody and hoping everybody (other than the receiver) will ignore the packet
- Broadcast: nodes can send messages that can be heard by all nodes on the network.
  - Almost essential for network administration
  - Can also be used for applications, e.g. video conferencing
- But broadcast fundamentally does not scale.

Building Larger LANs: Bridges

- Hubs are physical level devices
  - Don't isolate collision domains ➔ broadcast issues
- At layer 2, *bridges* connect multiple IEEE 802 LANs
  - Separate a single LAN into multiple smaller collision domains
    - Reduce collision domain size



Basic Bridge Functionality

- Bridges are full fledged packet switches
  - Saw bridge structure last class
- Frame comes in on an interface
  - Switch looks at destination LAN address
  - Determines port on which host connected
  - Only forward packets to the right port
  - Must run CSMA/CD with hosts connected to same LAN

"Transparent" Bridges

- Design features:
  - "Plug and play" capability
  - Self-configuring without hardware or software changes
  - Bridge do not impact the operation of the individual LANs
- Three components of transparent bridges:
  1) Forwarding of frames
  2) Learning of addresses
  3) Spanning tree algorithm

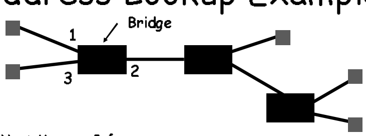
Frame Forwarding

- Each switch maintains a forwarding database:

```
<MAC address, port, age>
    MAC address: host or group address
    Port: port number on the bridge
    Age: age of the entry
```

- Meaning: A machine with MAC address lies in the direction of number port of the bridge
- For every packet, the bridge "looks up" the entry for the packet's destination MAC address and forwards the packet on that port.
  - No entry → packets are broadcasted

Address Lookup Example

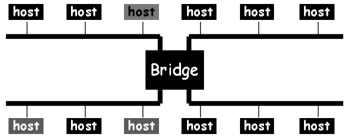


Address	Next Hop	Info
A21032C9A591	1	8:36
99A323C90B42	2	8:01
8711C9B900AA	2	8:15
301B2369011C	2	8:16
695519001190	3	8:11

- Address is a 48 bit IEEE MAC address.
- Next hop: output port for packet
- Timer is used to flush old entries
- Size of the table is equal to the number of hosts
- Flat address → no aggregation

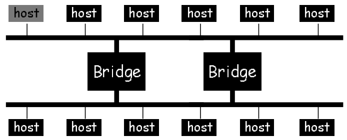
Learning Bridges

- Bridge tables can be filled in manually (flush out old entries etc)
  - Time consuming, error-prone
  - Self-configuring preferred
  - This is not done anyway; Instead bridges use "*learning*"
- Keep track of source address of packet (S) and the arriving interface (I).
  - Fill in the forwarding table based on this information
  - Packet with destination address S must be sent to interface I!



Spanning Tree Bridges

- More complex topologies can provide redundancy.
  - But can also create loops.
    - E.g. What happens when there is no table entry?
  - Multiple copies of data
  - → Could crash the network.



Spanning Tree Protocol Overview

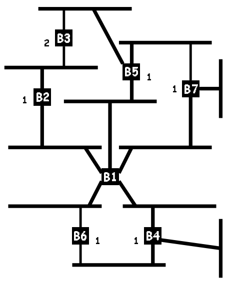
Embed a tree that provides a single unique path to each destination:

Bridges designated ports over which they will or will not forward frames

By removing ports, extended LAN is reduced to a tree

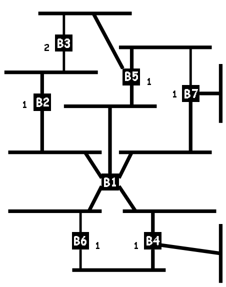
Spanning Tree Algorithm

- Root of the spanning tree is elected first → the bridge with the lowest identifier.
  - All ports are part of tree
- Each bridge finds shortest path to the root.
  - Remembers port that is on the shortest path
  - Used to forward packets
- Select for each LAN a designated bridge that will forward frames to root
  - Has the shortest path to the root.
  - Identifier as tie-breaker



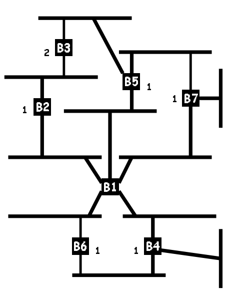
Spanning Tree Algorithm

- Each node sends configuration message to all neighbors.
  - Identifier of the sender
  - Id of the presumed root
  - Distance to the presumed root
- Initially each bridge thinks it is the root.
  - B5 sends (B5, B5, 0)
- When B receive a message, it decide whether the solution is better than their local solution.
  - A root with a lower identifier?
  - Same root but lower distance?
  - Same root, distance but sender has lower identifier?
- Message from bridge with smaller root ID
  - Not root; stop generating config messages, but can forward
- Message from bridge closer to root
  - Not designated bridge; stop sending any config messages on the port



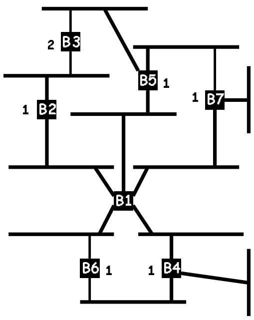
Spanning Tree Algorithm

- Each bridge B can now select which of its ports make up the spanning tree:
  - B's root port
  - All ports for which B is the designated bridge on the LAN
- States for ports on bridges
  - *Forward state* or *blocked state*, depending on whether the port is part of the spanning tree
- Root periodically sends configuration messages and bridges forward them over LANs they are responsible for



Spanning Tree Algorithm Example

- Node B2:
  - Sends (B2, B2, 0)
  - Receives (B1, B1, 0) from B1
  - Sends (B2, B1, 1) "up"
  - Continues the forwarding forever
- Node B1:
  - Will send notifications forever
- Node B7:
  - Sends (B7, B7, 0)
  - Receives (B1, B1, 0) from B1
  - Sends (B7, B1, 1) "up" and "right"
  - Receives (B5, B5, 0) - ignored
  - Receives (B5, B1, 1) – suboptimal
  - Continues forwarding the B1 messages forever to the "right"

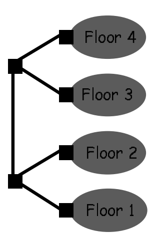


Ethernet Switches

- Bridges make it possible to increase LAN capacity.
  - Packets are no longer broadcasted - they are only forwarded on selected links
  - Adds a switching flavor to the broadcast LAN
  - Some packets still sent to entire tree (e.g., ARP)
- Ethernet switch is a special case of a bridge: each bridge port is connected to a single host.
  - Can make the link full duplex (really simple protocol!)
  - Simplifies the protocol and hardware used (only two stations on the link) – no longer full CSMA/CD
  - Can have different port speeds on the same switch
    - Unlike in a hub, packets can be stored

Example LAN Configuration

- 10 or 100 Mbit/second connectivity to the desk top using switch or hubs in wiring closets.
- 100 or 1000 Mbit/second switch fabric between wiring closets or floors.
- Management simplified by having wiring based on star topology with wiring closet in the center.
- Network manager can manage capacity in two ways:
  - speed of individual links
  - hub/bridge/switch tradeoff

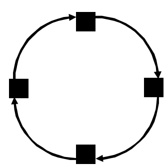


A Word about "Taking Turn" Protocols

- First option: Polling-based
  - Central entity polls stations, inviting them to transmit.
    - Simple design – no conflicts
    - Not very efficient – overhead of polling operation
    - Still better than TDM or FDM
    - Central point of failure
- Second (similar) option: Stations reserve a slot for transmission.
  - For example, break up the transmission time in contention-based and reservation based slots
    - Contention based slots can be used for short messages or to reserve time
    - Communication in reservation based slots only allowed after a reservation is made
  - Issues: fairness, efficiency

Token-Passing Protocols

- No master node
  - Fiber Distributed Data Interface (FDDI)
- One token holder may send, with a time limit.
  - known upper bound on delay.
- Token released at end of frame.
  - 100 Mbps, 100km
- Decentralized and very efficient
  - But problems with token holding node crashing or not releasing token



Next Lecture

- The IP layer lecture series begins..
  - Addressing
  - Forwarding in IP