Data Link Layer
- Media Access Control

Yan Wang
Two Types of Links

- **Point-to-point**
  - Point-to-Point Protocol (PPP) for dial-up access
  - point-to-point link between Ethernet switch and host
- **Broadcast (shared wire or medium)**
  - traditional Ethernet
  - 802.11 wireless LAN
Media Access Control (MAC)

• In our earlier discussion about DLL, we assumed that a link is associated with two nodes
• When a link/medium is shared among many nodes, we need to resolve the conflicts of multiple nodes transmitting at the same time.
• This issue is addressed by a sub-layer within the DLL: the media access control (MAC) sublayer.
Data Link Layer Model

Application
Presentation
Session
Transport
Network
Data Link Layer
Physical

Logical Link Control

Media Access Control
Discussion 6

• Imaging you are holding a big conference with colleagues in the same room, what rules will you make to ensure everyone has told others his/her idea?
Ideal MAC Protocols

Shared medium of capacity of R bps
1. When one node wants to transmit, it can send at rate $R$.
2. When $M$ nodes want to transmit, each can send at average rate $R/M$
3. Fully decentralized:
   ▫ No special node to coordinate transmissions
   ▫ No synchronization of clocks, slots
4. Simple
Three Approaches of MAC

• Fixed assignment -- multiplexing
  ▫ Divide channel into smaller “pieces” (time slots, frequency, code)
  ▫ Allocate piece to node for exclusive use

• Compete for the medium
  ▫ ALOHA
  ▫ Carrier Sense Multiple Access (CSMA)
  ▫ CSMA with Collision Detection (CSMA/CD)
  ▫ CSMA with Collision Avoidance (CSMA/CA)
  ▫ Wait for Your Turn

• Token Passing Ring
  ▫ Token Passing Bus
  ▫ FDDI
Time Division Multiple Access

- Access to channel in "rounds"
- Each station gets fixed length slot (length = frame transmission time) in each round
- Unused slots go idle
- Example: 6-station LAN, 1,3,4 have frames, slots 2,5,6 idle
Frequency Division Multiple Access

- Channel spectrum divided into frequency bands
- Each station assigned fixed frequency band
- Unused transmission time in frequency bands go idle
- Example: 6-station LAN, 1,3,4 have frames, frequency bands 2,5,6 idle
Code Division Multiple Access

- Unique “code” assigned to each user; i.e., code set partitioning
- Used mostly in wireless broadcast channels (cellular, satellite, etc.)
- All users share same frequency, but each user has own “chipping” sequence (i.e., code) to encode data
  - encoded signal = (original data) X (chipping sequence)
  - decoding: inner-product of encoded signal and chipping sequence
- Allows multiple users to “coexist” and transmit simultaneously with minimal interference (if codes are “orthogonal”)
Compete for Medium

- 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”
- MAC protocol specifies:
  - how to detect collisions
  - how to recover from collisions (e.g., via delayed retransmissions)
- Examples:
  - ALOHA, Slotted ALOHA
  - CSMA, CSMA/CD, CDMA/CA
ALOHA

- ALOHA
  - simple, no synchronization
  - when frame first arrives
    - transmit immediately
  - collision probability increases:
    - frame sent at $t$ collides with other frames sent before and after $t$

- Slotted ALOHA
  - Time is divided into slots
  - Only transmitted at the beginning of a slot
Performance of ALOHA

Slotted ALOHA: $S = Ge^{-G}$

Pure ALOHA: $S = Ge^{-2G}$
Carrier Sense Multiple Access (CSMA)

- Station listens to channel for ongoing transmissions.
- If so, the station waits until the channel is idle.
- When the channel is idle, the frame is transmitted.
- Collisions may still occur. How?

- If a collision occurs, the station waits random amount of time and retransmits.
- The longer the propagation delay, the longer the window of collisions and the worse the performance.
Collision Detection

- During a transmission, the station also listens to the channel and, if it detects collisions, it immediately aborts the transmission with a jamming signal.
  - This reduces the wasted time due to collisions.

- In the worst case, how long does a station take to detect a collision?
Worst Case

(a) A  B

(b) A  B

(c) A  B

(d) A  B
CSMA / Collision Detection (CD)

• Rules for CSMA/CD:
  ▫ If the channel is busy, then wait until idle and transmit immediately (called 1-persistent, a special case of p-persistent.)
  ▫ If the channel is quiet, then send the packet; continue listening to the transmission.
  ▫ If a collision occurs, then abort the transmission and send a short jamming signal.
  ▫ Wait a random number of slots before trying again.
  ▫ How to determine random number?
Binary Exponential Backoff

• To determine the number of time slots to wait before re-sending:
  ▫ On the first collision, wait either 0 or 1 slots.
  ▫ On the second, wait 0, 1, 2, or 3 slots.
  ▫ On the third, wait 0 to 7 slots.
• In general, after \( n \) collisions, wait
  ▫ anywhere from 0 to \( 2^n - 1 \) slots, if \( n \leq 10 \);
  ▫ or between 0 and 1023 slots, if \( n > 10 \).
• After 16 collisions, give up and report that packet could not be sent.
**Nonpersistent:***
- Transmit if idle
- If busy, wait random time and repeat process
- If collision, back off

**1-Persistent:***
- Transmit as soon as channel goes idle
- If collision, back off

**P-Persistent:***
- Transmit as soon as channel goes idle with probability P
- Otherwise, delay one time slot and repeat process
- If collision, back off

**Figure 16.1 CSMA Persistence and Backoff**
Ethernet Overview (IEEE 802.3 CSMA/CD)

- A broadcast-based LAN technology using CSMA/CD
- The official standard is IEEE 802.3.
- The Ethernet technology comprises two parts:
  - a DLL/MAC layer that defines frame format, error detection, CSMA/CD parameters, etc.
  - a family of physical layer standards
Ethernet Overview (2)

• Transceiver Hardware
  ▫ senses carrier: busy, idle, or collision
  ▫ sends/receives frames
  ▫ contains a hardwired address

• Ethernet protocol software/firmware
  ▫ formats header
  ▫ performs random backoff after collision
  ▫ implements binary exponential backoff algorithm

• No master station

• Speeds: 10 Mbps to 1 Gbps, 10 Gbps or higher standard
Typical Ethernet Hardware

- **Addresses**
  - unique, 48-bit MAC address assigned to each adapter
  - example: **80:00:e4:b1:ad:02**
  - broadcast: all 1s
  - multicast: first bit is 1
The “Classic” Ethernet

- 10 Mbps over coaxial cables
  - standard 10Base5 called “thick Ethernet”
  - standard 10Base2 called “thin Ethernet”

10 Base 5

Max. 10Mbps  Baseband signaling  Up to 500 m
10Base5

- Each cable segment up to 500 meters
- 4 repeaters can be used to cascade 5 segments, resulting in a diameter 2,500 meters
- Signal propagation delay in the cable is approximately $2 \times 10^8$ m/sec.
- Each repeater introduces a 0.5 μsec delay
- The slot time in the exponential backoff algorithm is 51.2 μsec, the transmission time of 64 bytes at speed 10 Mbps
- Manchester encoding
10Base5 (Continued)

• Why time slot is 51.2 μsec?
  ▫ because the time is enough for any other station to have realized that the first station has started transmitting
• Enforce a minimum frame length of 64 bytes. Why?
  
  • https://bharathisubramanian.wordpress.com/2009/08/05/minimum-ethernet-framepacket-size/
  • http://intronetworks.cs.luc.edu/current/html/ethernet.html
Frame Format

<table>
<thead>
<tr>
<th>Destination</th>
<th>Source</th>
<th>Length /Type</th>
<th>Frame Data</th>
<th>CRC Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>48 bits</td>
<td>48 bits</td>
<td>16 bits</td>
<td>368 to 12000 bytes</td>
<td>32 bits</td>
</tr>
</tbody>
</table>

- **Source Address:**
  - 48 bits each; first 24 bits vendor ID; second 24 bits assigned by the vendor
  - all Ethernet addresses are globally unique
Frame Format (2)

- **Destination Address:**
  - due to the broadcast nature of Ethernet, every frame will be seen by the physical layer modules at all stations
  - only the physical layer module at the destination station delivers the frame to its DLL layer
  - this filtering is performed by hardware
  - the destination address of all 1’s is called the **broadcast address**; frames destined to the address will be delivered to the DLL modules of all stations
  - Destination must be on the same LAN as the Source
Frame Format

- **Length/Type:**
  - If value < 1536 (0x0600), length of data
  - Otherwise, protocol type of data
    - 0x0800 for IP, 0x0806 for ARP, etc.
  - Frame Type describes the payload; each frame is self-identifying.

- **Frame Data:**
  - 46 to 1500 bytes
  - when necessary, this field is padded to 46 bytes, to ensure the 64-byte minimum frame length from destination address to checksum

- **CRC Checksum**
  - using the CRC-32 generator polynomial
Ethernet Algorithm Details

• If line is idle
  ▫ send immediately
  ▫ upper bound message size of 1500 bytes
  ▫ must wait 9.6us between back-to-back frames

• If line is busy
  ▫ wait until idle and transmit immediately
  ▫ called 1-persistent (special case of p-persistent)
Ethernet Algorithm Details (2)

- If collision
  - jam for 32 bits, then stop transmitting frame
  - minimum frame is 64 bytes (header + 46 bytes of data)
  - delay and try again
    - 1st time: 0 or 51.2us
    - 2nd time: 0, 51.2, or 102.4us
    - 3rd time 51.2, 102.4, or 153.6us
    - nth time: $k \times 51.2\text{us}$, for randomly selected $k=0..2^n - 1$
    - give up after several tries (usually 16)
    - exponential backoff
State Diagram for CSMA/CD

Packet?

Sense Carrier

Send

Detect Collision

Discard Packet

No

attempts < 16

Yes

attempts == 16

Jam channel
b=CalcBackoff();
wait(b);
attempts++;
LAN Address

- **32-bit IP address:**
  - network-layer address
  - used to get datagram to destination IP network
- **LAN (or MAC or physical or Ethernet) address:**
  - used to get datagram from one interface to another physically-connected interface (same network)
  - 48 bit MAC address (for most LANs) burned in the adapter ROM
Address Resolution Protocol

- Two addresses
  - How to know the MAC from IP?
- ARP
  - Each IP node (Host, Router) on LAN has ARP table
  - ARP Table: IP/MAC address mappings for some LAN nodes
    - (IP, MAC, TTL)
    - TTL (Time To Live): time after which address mapping will be forgotten (typically 20 min)
ARP Protocol

- A wants to send datagram to B, and A knows B's IP address.
- Suppose B's MAC address is not in A's ARP table.
- A broadcasts ARP query packet, containing B's IP address
  - all machines on LAN receive ARP query
- B receives ARP packet, replies to A with its (B's) MAC address
  - frame sent to A's MAC address (unicast)
- A caches (saves) IP-to-MAC address pair in its ARP table until information becomes old (times out)
  - soft state: information that times out (goes away) unless refreshed
- ARP is “plug-and-play”:
  - nodes create their ARP tables without intervention from net
  - administrator
Ethernet Over Twisted Pair

- 10BASET and 100BASET uses Ethernet software and protocols over twisted pair wiring
- Uses RJ-45 plug (similar to RJ-11 modular phone plug)
- Concentrator hubs provide immunity from Ethernet problems such as an open jack killing the network
  - Hubs are inexpensive because they are not true switches: a hub is basically a multiway repeater; it relays incoming signals to all ports
  - Hubs maintain the broadcast nature of Ethernet
  - 10BASET can be interfaced to 10BASE2
10Base-T: Twisted-pair Ethernet
Fast Ethernet

- 100 Mbps -- peak
- Use the original Ethernet MAC and format but operates at ten times the speed
- The network diameter is reduced by a factor of 10.
- Again, different physical layer standards support different transmission media:
  - 100Base-T and 100Base-T4 for twisted pairs
  - 100Base-F for fibers
- Uses star topology – similar to 10BaseT.
Discussion 7

- MAC addresses have limited numbers (48 bits = 281474976710655), what if it will be used up? What to do about it?
  - MAC address estimated to run out in about 100 years by 2014!
Global/Local bit: 1 -> Locally Administered, 0 -> Globally Unique
- If ‘1’ it means the Address is Locally Administered and may not be globally unique
- Duplicate MAC addresses are VERY bad for switches, thus these are rarely used

OUI: 24-bit Organizational Unique Identifier (purchased from IEEE)
ODI: 24-bit Organizational Defined Identifier (uniqueness by the Organization)
Possible Solutions

• Before IPv6, the IPv4 addressing problem was solved with routers – and this solution was so good it's still used
• You probably have a router in your home
  ▫ Its in the device used to connect your computers to the internet
  ▫ Your home uses only one unique IP address to connect the ‘outside’ world
  ▫ Each of your devices ‘inside’ your home get their IP addresses assigned
    ▪ either automatically or statically such that they are locally unique
  ▫ Your neighbor could be using the same IP addresses as your devices, but they don’t ‘see’ each other due to the router
  ▫ You can replace a computer and it automatically works!
  ▫ Many routers support security!
  ▫ These concepts port to MAC Addresses
Full Duplex Operation

• Traditional Ethernet half duplex
  ▫ Either transmit or receive but not both simultaneously
• With full-duplex, station can transmit and receive simultaneously
  ▫ 100-Mbps Ethernet in full-duplex mode, theoretical transfer rate 200 Mbps
• Attached stations must have full-duplex adapter cards
• Must use switching hub
  ▫ Each station constitutes separate collision domain
  ▫ In fact, no collisions
  ▫ CSMA/CD algorithm no longer needed
  ▫ 802.3 MAC frame format used
  ▫ Attached stations can continue CSMA/CD
Comparing MAC Protocols

- **Fixed assignment**
  - share channel efficiently and fairly at high load
  - inefficient at low load: delay in channel access, $1/N$ bandwidth allocated even if only 1 active node!
- **Compete for medium**
  - efficient at low load: single node can fully utilize channel
  - high load: collision overhead
- **Wait for your turn**
  - look for best of both worlds!
Wait for Your Turn

- **Polling**
  - master node “invites” slave nodes to transmit in turn
  - concerns:
    - polling overhead
    - latency
    - single point of failure (master)

- **Token passing**
  - control token passed from one node to next sequentially
  - token message
  - concerns:
    - token overhead
    - latency
    - single point of failure (token)
Wait for Your Turn: Token Passing

- A token circulates among all stations.
  - the token is a miniature, 3-byte frame (including start and end flags)
Token Passing

- When the token arrives, a station either **seizes the token** and sends a frame, or **passes the token** to the next station.
  - Assuming that station $i$ has a frame $f$ to send to station $j$.
  - Station $i$ waits for the arrival of the token and seizes the token.
  - Station $i$ sends $f$ to station $i + 1$, which in turns passes $f$ to station $i + 2$.
  - When frame $f$ arrives at station $j$, station $j$ picks up $f$ and simultaneously forwards $f$ to station $j + 1$.
  - Eventually, $f$ returns to station $i$, which passes the token, rather than $f$, to station $i + 1$. 
Token Ring

- Focus on Fiber Distributed Data Interface (FDDI)
  - 100 Mbps
  - Was (not is) a candidate to replace Ethernet
  - Used in some MAN backbones (LAN interconnects)
Example of Token Ring

- IBM: 4Mbps token ring
- IEEE 802.5: 16Mbps
FDDI with a Single Cable (Dual Ring)

- SAS: Single Attachment Station
- DAS: Dual Attachment Station
Basic Concepts of Token Ring

- Frames flow in one direction
  - Upstream to downstream
- Token
  - Special bit pattern rotates around ring
- Stations
  - Must capture token before transmitting
  - Must remove frame after it has cycled
  - Must release token after transmitting
- Service
  - Stations get round-robin service
- Immediate release
  - Used in FDDI
  - Token follows last frame immediately
- Delayed release
  - Used in IEEE 802.5
  - Token sent after last frame returns to sender
Timed Token Algorithm

- **Token Holding Time (THT)**
  - Upper limit on how long a station can hold the token
  - Each station is responsible for ensuring that the transmission time for its packet will not exceed THT

- **Token Rotation Time (TRT)**
  - How long it takes the token to traverse the ring
  - TRT <= ActiveNodes x THT + RingLatency
  - How to take the measurement?

- **Target Token Rotation Time (TTRT)**
  - Agreed-upon upper bound on TRT
Timing Algorithm: IEEE 802.5

- Each node measures TRT between successive tokens
- If measured-TRT > TTRT
  - Token is late
  - Don’t send
- If measured-TRT < TTRT
  - Token is early
  - OK to send
Token Maintenance: IEEE 802.5

• Monitoring for a Valid Token
  ▫ All stations should periodically see valid transmission (frame or token)
  ▫ Maximum gap
    • = ring latency + max frame < = 2.5ms
  ▫ Set timer at 2.5ms
    • send claim frame if timer expires
Two Classes of Traffic: FDDI

- Two classes of traffic
  - synchronous: can always send up to one TTRT per token rotation (delay sensitive)
  - asynchronous: can send only if token is early (throughput sensitive)
Timing Algorithm: FDDI

- Each station is allocated $S_i$ time units for synchronous traffic per TRT
- TTRT is negotiated
  - $S_1 + S_2 + \ldots + S_N + \text{RingLatency} \leq \text{TTRT}$
- Algorithm Goal
  - Keep actual rotation time less than TTRT
  - Allow station $i$ to send $S_i$ units of synchronous traffic per TRT
  - Fairly allocate remaining capacity to asynchronous traffic
  - Regenerate token if lost
Timing Algorithm: FDDI

- When a node gets the token
  - Set TRT = time since last token
  - If TRT > TTRT
    - Token is late
    - Send synchronous data
    - Don’t send asynchronous data
  - If TRT < TTRT
    - Token is early
    - OK to send any data
      - Send synchronous data
      - Set THT = TTRT – TRT
      - If THT > 0, send asynchronous data
Analysis

- **Worst case:** $2xTTRT$ between seeing token
  - Full asynchronous send followed by full synchronous send

- **Back-to-back** $2xTTRT$ rotations between seeing token not possible
  - After a cycle with $TRT = 2*TTRT$, no asynchronous traffic will be sent
Token Maintenance: FDDI

- **Lost Token**
  - No token when initializing ring
  - Bit error corrupts token pattern
  - Node holding token crashes

- **Monitoring for a valid token**
  - Should see valid transmission (frame or token) periodically—within $2 \times \text{TTRT}$
  - Maximum gap = RingLatency + MaxFrame $\leq 2.5\text{ms}$
Token Maintenance: FDDI (cont.)

- Generating a Token (and agreeing on TTRT)
  - Execute when joining ring or suspect a failure
  - Send a claim frame that includes the node’s TTRT bid
    - When receive claim frame, update the bid and forward
    - If your claim frame makes it all the way around the ring:
      - Your bid was the lowest
      - Everyone knows TTRT
      - You insert new token
Discussion

• This is a broadcast-based technology because all stations see every frame.
• The forwarding of both the token and data frames is performed by NICs in hardware.
• Three token-based LAN technologies
  ▫ 802.4 token bus
  ▫ 802.5 token ring
  ▫ Fiber Distributed Data Interface (FDDI)
Token-Based MAC vs. CSMA/CD

- When the medium is quite, a sending station
  - wastes no time in waiting with CSMA/CD
  - must wait for the arrival of the token with token-based approaches
  - thus, CSMA/CD outperforms token-based approaches in light traffic.

- Token-based MAC handles heavy loads better than CSMA/CD.
  - no waste of bandwidth due to collisions
Additionally ...

- Moreover, token-based MAC has advantages in handling real-time traffic
  - fairness: stations access the medium in a round-robin manner
  - bounded delays: a station can predict the next time the token returns to it (in CSMA/CD, a station waits for a “random period of time” before retransmission)
A Word about LAN and MAC

- The Broadcast/MAC technologies discussed in this lecture are inherently LAN technologies.
- However, the reverse is not always true; many LANs are based on switching.
  - ATM
  - switched Ethernet
Why did Ethernet Win?

• There are LOTS of LAN protocols
  ▫ **Price**
  ▫ Performance
  ▫ Availability
  ▫ Ease of use
  ▫ Scalability