The Link Layer

Michael Brodskiy
Professor: E. Bernal Mor

November 30, 2023

- Link Layer: Introduction
  - Terminology
    - Hosts, routers → nodes
    - Communication channels that connect adjacent nodes along communication path → links
      - Wired links
      - Wireless links
    - Over a given link, the transmitting node encapsulates the network-layer packet in a link-layer frame
  - Link layer has responsibility of transferring network-layer packets from one node to a physically adjacent node over a link
- Link Layer: Context
  - Packets transferred by different link protocols over different links
    - WiFi on first link
    - Ethernet on next link
    - Etcetera
  - Each link protocol provides different services
- Link Layer Services
  - Framing
    - Encapsulate packet into frame, adding header and maybe trailer
    - Addressing: "MAC" addresses used in frame headers to identify transmitter/receiver node → different from IP Address
  - Link access

        * Medium access control (MAC) protocol specifies the rules by which a frame is transmitted onto the link
    - Flow control
        * Pacing between adjacent sending and receiving nodes
    - Reliable delivery between adjacent nodes
        * We learned how to do this already (Transport layer)!
        * Seldom used on low error rate links, for example: fiber, some twisted pairs
        * Commonly used on high error rate links, like wireless ones
    - Error control
        * Errors caused by signal attenuation, noise
        * Error detection: receiver detects presence of errors
            · Ask sender for retransmission or drops frame
        * Error correction: receiver identifies and corrects bit error(s) without resorting to retransmission

- Half-Duplex and Full-Duplex Links
    - Unidirectional links (Simplex Links)
        * Communication occurs in one direction only
    - Bidirectional links
        * Half-Duplex Link — Communication occurs in both directions, but not at same time
        * Full-Duplex Link — Communication occurs in both directions at same time

- Where is the Link Layer Implemented?
    - For the most part, link layer is implemented on a chip called the network adapter, aka a Network Interface Card (NIC)
        * The NIC implements Link and Physical layers
        * *E.g.* Ethernet card, WiFi card or chip
    - NIC attaches into node's system buses
    - Link layer is implemented as a combination of hardware and software
        * Hardware: NIC implements most of the functions
        * Software: activating hardware controller, responds to controller interrupts, etc.

- Error Control
    - EDC → Error Detection/Correction bits (redundant bits)

  - D → Data protected by error control, may include header fields
  - Error control is not 100% reliable
    * Error control technique may miss some errors; we want to keep the probability of missing the errors small
    * Larger EDC field yields better detection and correction
    * Error correction needs more redundant bits than error detection for same number of errors
  - Parity Checking
    * Single bit parity: detect single bit errors
      · Even parity: set parity bit so there is an even number of 1's
      · Odd parity: set parity bit so there is an odd number of 1's
    * Two-dimensional bit parity: detect and correct single bit errors
      · Even parity: no errors
  - Cyclic Redundancy Check (CRC)
    * D: d data bits (given)
    * G: generator, bit pattern of $r+1$ bits where MSB must be 1 → transmitter and receiver agree on G (given)
    * R: r CRC bits, redundant bits
    * Transmitter: choose R, such that $\langle D, R\rangle$ is exactly divisible by G (modulo-2 arithmetic) → $D \cdot 2^r \text{XOR} R = nG$
    * Receiver: knows G and divides $\langle D, R\rangle$ by G → non-zero remainder: error detected!
      · Can detect all burst errors less than $r+1$ bits
    * More powerful error-detection technique: widely used in practice (Ethernet, WiFi)

- Types of Links
  - Point-to=point link: a single sender at one end on a link and a single receive at the other end of the link
    * PPP (Point-to-Point Protocol), switched Ethernet, etc.
  - Broadcast (shared medium) link: Multiple transmitting and receiving nodes all connected to the same, single shared broadcast link
    * Need to handle multiple access problem (classic Ethernet, 4G/5G, WiFi, etc.)

- Multiple Access Problem
  - Multiple access problem: how to coordinate the access of multiple transmitting and receiving nodes to a single, shared broadcast channel

  - Two or more simultaneous transmissions by nodes → interference
    - Collision: if a node receives two or more signals at the same time (collision happens in the receiver)
  - MAC (Medium Access Control) Protocol
    - Distributed algorithm that coordinates the frame transmissions of many nodes into the broadcast channel
    - Determines how nodes share channel and when nodes can transmit
    - Communication about channel sharing must use channel itself!
    - No out-of-band channel for coordination

- An Ideal MAC Protocol
  - Given: broadcast channel of rate R bps
  - Desirable Characteristics:
    1. When one node wants to transmit, it can send at rate R
    2. When M nodes want to transmit, each can send an average rate R/M
    3. Fully Decentralized
       - No special node to coordinate transmissions
       - No synchronizations of clocks, slots
    4. Simple

- MAC Protocols Taxonomy
  - Three Broad Classes:
    1. Channel Partitioning
       - Divide channel into smaller "pieces" (time slots, frequency, code)
       - Collision free: allocate piece of node for exclusive use to avoid collisions
    2. Random Access
       - Channel not divided, allow collisions
       - "Recover" from collisions
    3. Turn-Taking
       - Nodes take turns: tightly coordinate shared access to avoid collisions

- Channel Partitioning: FDMA
  - FDMA: Frequency Division Multiple Access
  - Frequency spectrum of the channel is divided into N frequency bands (each with transmission rate R/N bps)
  - Each node is assigned a fixed frequency band

  - Advantages: Avoids collisions, and divides the capacity link fairly
  - Drawback: Unused link capacity if frequency band goes idle

- Channel Partitioning: TDMA
  - TDMA: Time Division Multiple Access
  - Access to channel in "rounds" $\rightarrow$ time divided in N slots
  - Each node gets fixed time slot in each round (node average transmission rate R/N bps)
  - Advantages: it avoids collisions, and it divides the link capacity fairly
  - Drawback: unused link capacity if slots fo idle

- Random Access Protocols
  - A transmitting node always transmits at full channel rate, R bps
  - Two or more transmitting nodes create a collision
  - Random access MAC protcool specifics:
    * How to detect collisions
    * How to recover from collisions (like via delayed transmissions)

- Slotted ALOHA
  - Assumptions:
    * All frames same size
    * Time divided into equal size slots
      · Slot duration: time to transmit one frame, t_l
    * Nodes start to transmit only at the beginning of slot
    * Nodes are synchronized
    * If 2 or more nodes transmit in slot, all nodes detect collision
  - Operation:
    * When node obtains fresh frame, transmits in next slot
      · If no collision: node can send new frame in next slot
      · If collision: node retransmits frame in each subsequent slot with probability p until success

- Slotted ALOHA: Efficiency
  - Efficiency: long-run fraction of successful slots (many nodes, all with many frames to send)
  - Suppose N nodes with many frames to send, each transmits in slot with probability p

    * Probability that a given node has success in a slot: $p(1-p)^{N-1}$
    * Probability that any node has a success: $Np(1-p)^{N-1}$
    * Max efficiency: find p^* that maximizes $Np(1-p)^{N-1}$
    * For many nodes, take limit of $Np^*(1-p^*)^{N-1}$ as N goes to infinity gives:

$$\text{Max Efficiency: } \frac{1}{e} = .37$$

- Pure Aloha
  - Unslotted Aloha: simpler, no synchronization (no time slots)
  - Consider that t_f is the frame transmission time
  - When node obtains fresh frame, transmit immediately
  - Collision: retransmit with probability p immediately, repeat every t_f until the frame is transmitted
  - Collision probability increases:
    * Frame sent at t_0 collides with other frames sent in $[t_0 - t_f, t_0 + t_f]$
- Pure Aloha Efficiency
  - Efficiency at many nodes is .18, even worse than slotted aloha
- Carrier Sense Multiple Access (CSMA)
  - CSMA: listen before transmitting
    * If channel sensed idle: transmit entire frame
    * If channel sensed busy: defer transmission
  - CSMA/CD: CSMA with Collision Detection
    * Collisions detected within short time
    * Colliding transmissions aborted, reducing channel wastage
    * After collision, wait a random time before repeating the CSMA/CD cycle
- CSMA: Collisions
  - Collisions can still occur with carrier sensing:
    * Propagation delay: two nodes may not heard each other's just-started transmission
    * The longer the propagation delay from one node to another, the larger the probability that a node is not able to sense a transmission that has already begun at another node
    * Distance and propagation delay play a crucial role in determining collisoon probability

  - Collision: nodes continute to transmit their frames
    * Entire frame transmission time wasted

- CSMA/CD
  - Collision detection
    * Wired Links: while transmitting, node monitors for the presence of signal energy coming from others nodes
      · Signal energy from other nodes detected! Abort transmission
    * Wireless links: difficult $\rightarrow$ use CSMA/CA instead
  - CSMA/CD reduces the amount of time wasted in collisions
    * Transmission aborted on collision detection

- Ethernet CSMA/CD Algorithm
  - NIC receives packet from network layer, creates frame
  - NIC senses channel:
    * If idle: start frame transmission
    * If busy: wait until channel idle, then transmit
  - If NIC transmits entire frame without collision, NIC is done with frame!
  - If NIC detects another transmission while transmitting, abort and send jam signal
  - After aborting, NIC enters binary exponential backoff:
    * After n-th collision, NIC chooses K at random from $\{0, 1, 2, \ldots, 2^n - 1\}$
    * NIC waits $K \cdot 512$ bit duration times, returns to step 2
    * More collisions: longer backoff interval

- CSMA/CD Efficiency
  - t_{prop} is the maximum propagation delay between two nodes in broadcast link
  - t_{trans} is the time to transmit maximum-size frame

$$eff = \frac{1}{1 + 5t_{prop}/t_{trans}}$$

  - Efficiency goes to 1 when either propagation goes to zero or transmission goes to infinity
  - Better performance than ALOHA: and simple and cheap

- "Taking Turns" MAC Protocols
  - Master node "invites" slave nodes to transmit in turn

  - Typically used with "dumb" slave devices
  - Token Passing:
    * A small frame, called a token, is passed from one node to the next sequentially
    * The node that has the token can transmit
    * Concerns:
      · Token overhead
      · Latency
      · Single point of failure (token)

- Local Area Network (LAN)
  - A switch is a link-layer device
    * Switches operate at link layer
    * Switch frames
    * Do not recognize network-layer addresses

- MAC (or LAN or physical or Ethernet) Addresses
  - Link-layer address for interface
  - Function: used "locally" to get frame from one interface to another connected interface in the same subnet (in IP-addressing sense)
  - 48-bit MAC address (for most LANs) burned in NIC ROM, also sometimes software settable
  - For example: 1A:2F:BB:76:09:AD
  - Each interface on LAN
    * Has unique 48-bit MAC address
    * Has unique 32-bit/128-bit IP address
  - MAC address allocation administered by IEEE
  - Manufacturer buys portion of MAC address space (to assure uniqueness)
  - Analogy:
    * MAC address: like Social Security Number
    * IP address: like postal address
  - MAC flat address: portability
    * Can move interface from one LAN to another with the same MAC address
    * Recall IP address is hierarchical and not portable: depends on IP subnet to which node it attached
  - MAC broadcast address: FF:FF:FF:FF:FF:FF

    - Used by a sender that wants all the other interfaces on the LAN to receive and process a frame

- ARP: Address Resolution Protocol
  - ARP Protocol: to determine MAC address of interface knowing the IP address
  - ARP table: each IP node (host, router) on LAN has an ARP table
    - IP/MAC address mappings for some LAN nodes:

      $$< \text{IP address; MAC address; TTL} >$$

      - TTL (Time To Live): time after which address mapping will be forgotten (typically 20 min)

- IEEE 802.3: Ethernet
  - "Dominant" wired LAN technology:
    - First widely used LAN technology
    - Simple and cheap
    - Kept up with speed race: 10 Mbps — 400 Gbps
    - Single chip, multiple speeds

- Ethernet: Physical Topology
  - Bus: popular through mid 90s → classic Ethernet
    - All nodes in same collision domain (frames can collide with each other)
  - Hub-based: popular in late 90s → classic Ethernet
    - A hub is a physical-layer device: input signal sent out on all other interfaces
    - Hub in center logically operates as a bus: all nodes in same collision domain
  - Star with a switch: prevails today → switched Ethernet
    - Active link-layer switch in center
    - Each switch interface is in a different collision domain (frames in different interfaces do not collide with each other)
    - With full duplex linkes, MAC protocol unnecessary

- IEEE 802.3 Ethernet Standards: Link and Physical Layers
  - Many different Ethernet standards
    - Link layer and physical layer
    - Common MAC protocol and frame format
    - Different speeds: 2 Mbps, 10 Mbps, 100 Mbps, 1 Gbps, 10 Gbps, 40 Gbps, 100 Gbps, 400 Gbps, ...

        * Different physical layer and physical media (fiber, cable)

- Ethernet: Unreliable, Connectionless
    - Ethernet MAC protocol (if needed): unslotted CSMA/CD with binary backoff
    - Error detection: CRC
    - Connectionless: no handshaking between sending and receiving NICs
    - Unreliable: receiving NIC does not send acks or nacks to sending NIC
        * CSMA/CD (if needed): collisions are detected in the transmitter and colliding frames retransmitted
        * Error detection: when an error is detected, the frame is dropped
            · Data in dropped frames is recovered end-to-end only if initial sender uses higherlayer reliable data transfer (like TCP), otherwise dropped data is lost
        * It makes sense as the physical medium has low bit error rate and if collisions are possible, they can be detected (CSMA/CD) and the frame is recovered

- Ethernet Frame Structure
    - Preamble: 8 bytes
        * Used to synchronize receiver and sender clock rates (usually not considered part of the frame header)
        * 7 bytes with pattern 10101010 followed by one byte with pattern 10101011 (start of frame)
    - Addresses: 6-byte destination and source MAC addresses
        * If adapter receives frame with matching destination address, or with broadcast adress, it passes data in frame to higher-layer protocol
        * Otherwise, adapter discards frame
    - CRC: 4-byte cyclic redundancy check
        * Error detected: frame is dropped
    - Type: 2 bytes
        * If below 0x0600 (1536) → Length: indicates the length of the data (payload)
        * If above 0x0600 → Type: indicates the higher layer protocol (most common)
            · Mostly IP but others possible, like ARP
            · Used to demultiplex up at receiver

- Ethernet Switch
    - Switch is a link-layer device: takes an active role
        * Store and forward Ethernet frames

    - Examine incoming frame destination MAC address, selectively forward frame to one-or-more outgoing links or drop the frame
    - If needed, uses CSMA/CD to access link
        - Half-duplex link or when a hub is connected to a switch interface
  - Transparent: hosts and routers are unaware of presence of switches
    - A host/router addresses a frame to another host/router in LAN
  - Plug-and-play, self-learning
    - Switches do not need to be configured

- Switch: Self-learning
  - Switches are self-learning: switch table is built automatically, dynamically and autonomously (no intervention of admin)
  - Switch learns which MAC addresses can be reached through which interfaces
    - When frame received, switch checks source MAC address of frame
    - "Learns" location of sender: incoming LAN segment
    - Records sender/location pair in switch table

- Switch: Frame Filtering/Forwarding
  - When frame received at switch:
    1. Update switch table: record MAC address of sending host, incoming link
    2. Index switch table using MAC destination address
    3. If entry found for destination $\rightarrow$ if destination on link from which frame arrived $\rightarrow$ drop frame
       (a) Else $\rightarrow$ forward frame on interface indicated by entry
    4. Else $\rightarrow$ flood (forward on all interfaces except arriving interface)

- Switches vs. Routers
  - Both are store-and-forward:
    - Routers: network-layer devices (examine network-layer headers)
      - Use IP addresses to forward packet
    - Switches: link-layer devices (examine link-layer headers)
      - Use MAC addresses to forward frame
  - Both have forwarding table
    - Routers: compute tables using routing algorithms and IP addresses
    - Switches: learn forwarding tables using self-learning