Computer Networks and Distributed Systems

Data Link Layer

Dr Fidelis Perkonigg

February 22, 2018

Overview

- Overview of Data Link Layer
  - How do we transfer data over a non-ideal (noisy, data loss) physical channel?
  - How do we divide data into chunks for the physical layer?
  - How do we control access to a physical channel?
- Data Framing
- Medium Access Control
  - In wired networks
  - In wireless networks
- Repeaters, Hubs, Bridges/Switches

Data Link Layer

- Arranges data into bit stream for sending over physical link
  - Defines communication between two physically connected network nodes
  - Must cope with different physical layer technologies

Two sub-layers:

- Logical Link Control (LLC)
  - Low-level flow and error control for single hop
- Media Access Control (MAC)
  - Framing, addressing and channel access

Data Link Layer Services

- Unacknowledged connectionless service
  - Independent frames with no logical connection
  - No recovery from loss but fast
  - Common in LANs using reliable channels
- Acknowledged connectionless service
  - Each frame is acknowledged
  - Good for unreliable channels such as wireless
  - Out of order delivery possible
- Acknowledged connection-oriented service
  - Connection established before data is sent
  - Each frame numbered and guaranteed to be delivered exactly once and in order
  - Provides reliable bit stream
- Provide error detection and correction
  - Physical layer may introduce errors by adding, removing, or modifying bits

LLC - Error Detection and Correction

Detection

- Add 1 parity bit to data (e.g. 8 bits)
  - Makes total number of 1s odd (or even)
  - Detects all single (and odd numbered) bit errors, misses even bit errors
- Cyclic Redundancy Check (CRC)
  - Hash-based checksum (often implemented in H/W)

Correction

- Error correcting codes aka. Forward Error Correction (FEC)
  - Add more redundancy resulting in greater capacity to detect and correct bursts of errors
  - E.g. repetition of data or more complex functions
- Detection vs correction
  - fast or slow channels?
  - error-prone channels?

Data Framing

- Need to group bits into separate smaller messages
  - Costs less to retransmit if error detected
  - Smaller chance of collisions
  - This means more overhead
- Need to add meta data to control protocol
  - Addressing, length, frame type, CRC, ...

Framing Methods

- Insert gaps
  - But timing hard to guarantee
- Count bytes
  - Include length field to delimit data

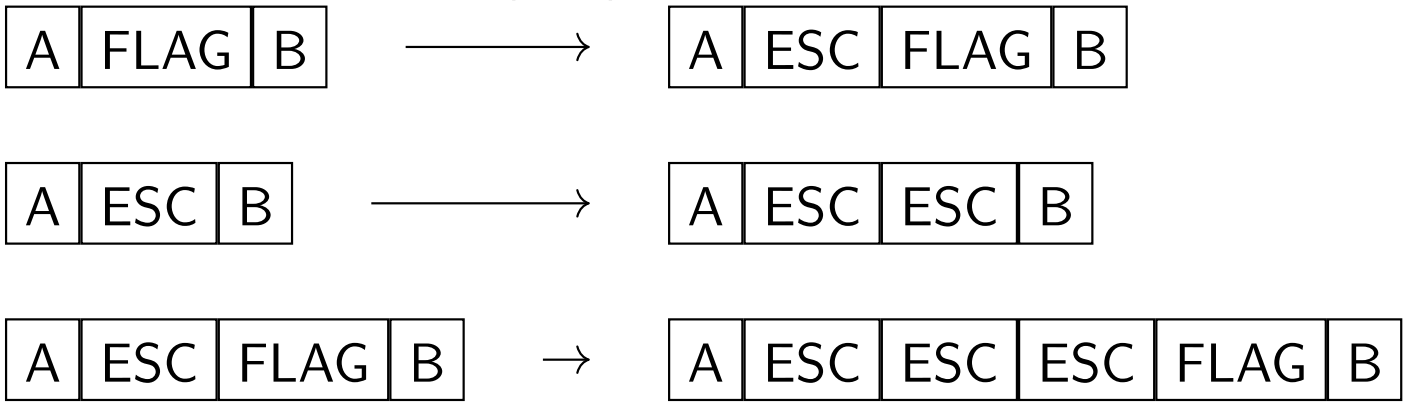
5	H	E	L	L	O	7	...

  - Any issues with this approach?
  - Synchronisation issues if the field length has been changed due to a transmission error

Framing Methods

- Use start and end flags
  - Special signal at start and end of frame (`FLAG`)
  - Search for flag if receiver loses track

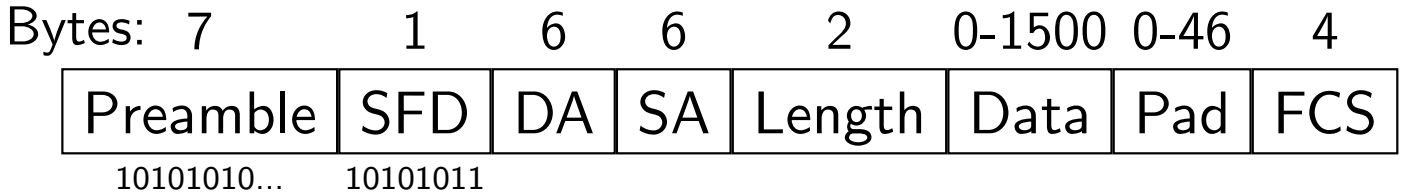
  - Any issues with this?
  - What if `FLAG` is part of the data?
- Uses byte stuffing
  - Identify data with same bit pattern as the flag
  - Use escape sequence (`ESC`) to identify if the next byte is data



Ethernet and IEEE 802.3

- Ethernet originally developed by Xerox
- Became open standard (IEEE 802.3)
- Two standards: IEEE 802.3 and Ethernet
  - Only small differences
  - Terms used interchangeably in common use
- Uses Manchester encoding
- Operates over various physical media
- Data link layer is separate to physical layer
- But physical layer affects parameters of Ethernet

IEEE 802.3 - Frame Format



- SFD: start of frame delimiter
- DA: destination address
- SA: source address
- FCS: Frame Check Sequence
- Ethernet standard slightly different
  - Does not have SFD field
  - Replaces Length with Type field

IEEE 802.3 - Frame Fields

- Preamble
  - 7-byte alternating 0s and 1s to establish synchronisation
  - Framing then by timing, spaces between frames plus counting from length field
- SFD (start of frame delimiter)
  - 10101011 indicates start of frame
  - Allows receiver to miss part of preamble without missing the start of the next field

IEEE 802.3 - Frame Fields

- Destination address
  - 16 or 48 bits (depending on implementation)
  - Host(s) intended to receive
    - Single host (unicast)
    - Group address (multicast)
    - Global address (broadcast)
- Source address
  - 16 or 48 bit address of sender

Ethernet Addresses

- Usually, Ethernet addresses are 48 bits
- Written as 6 pairs of hex digits (e.g. 00:11:85:7A:BC:E4)
- Medium Access Control (MAC) address
- Octets are received from left to right and least significant bit is received first for every octet (for IEEE 802.3)
- first 3 octets: vendor code (assigned by IEEE)
- last 3 octets: network card specific unique code set by vendor
- second bit of first octet: `0`: global address; `1`: local address (rest of vendor code is `0`)
- first bit of first octet: `0`: unicast addr; `1`: multicast

IEEE 802.3 - Frame Fields

- Type (Ethernet only)
  - Identifies higher level protocol
- Length (IEEE 802.3 only)
  - Bytes in this frame (optional)
- Data
  - Includes higher layer headers
  - Maximum transmission unit (MTU) of 1500 bytes
- Most hardware can find out if it is *type* or *length* (*type* values are $>$1500)

IEEE 802.3 - Frame Fields

- Pad
  - 0-46 bytes to ensure frame is long enough to enable collision detection
- FCS (Frame Check Sequence)
  - Enables error detection
  - CRC, based on all fields except preamble, SFD and FCS

Medium Access Control (MAC)

- How do we allocate communications channels?
  - Contention
  - Fairness
  - Access latency
- Physical channel supports multiplexing scheme
- Static allocation vs. dynamic allocation

Static Allocation

- Recall: TDM, FDM and CDMA
  - Static ways for stations to access fixed part of medium
- Properties
  - Guaranteed, allocated bandwidth
  - Bounded latency to transmit
- Challenges
  - In many computer networks most stations do not want to transmit at once
  - Do not want to waste bandwidth on silent stations
  - Alternative is to allocate dynamically

Dynamic Allocation

- Medium can be used on demand
- Single transmitter on medium is simpler electronically
- Challenges:
  - Need to ensure fair access to medium
  - Would like bounded delay to transmit
  - How to avoid collisions?

Propagation Delay

- Finite time for signal to go from one node to another: $delay = distance/speed$ (where $speed \approx 2 * 10^8 m/s$)
- Nodes will receive signal at different times
  - Depends on distance from sender
  - Need to keep this in mind when managing who transmits when

Medium Access Control: ALOHA

- Developed at University of Hawaii
- Send whenever data ready to go
- Central station broadcasts data back so that sender can check if a collision occurred
- Garbled data as a sign that collision occurred
  - If collision occurs, wait random time and try again
- Not very efficient: 18% theoretical maximum channel utilisation

Slotted ALOHA

- Divide time into slots
  - Start time of frames is synchronised
  - Probability of collisions is reduced
- Cannot assume synchronised clocks between stations
  - Master station sends short signal at start of each time frame
- Successful transmission 37% of the time

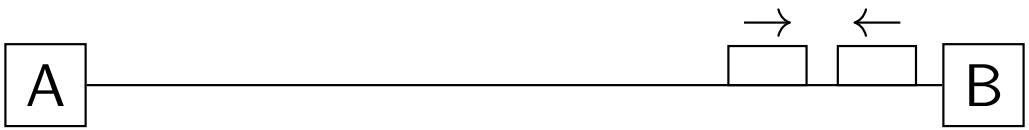
Improved Multiple Access

- ALOHA has a simple problem:
  - No-one listens before they start to send
  - Leads to lots of collisions
- Carrier Sense Multiple Access (CSMA)
  1. When ready to send, listen
  2. If channel busy, wait until idle
  3. When channel idle, send whole frame
  4. If collision, wait random time and start listening again

CSMA with Collision Detection

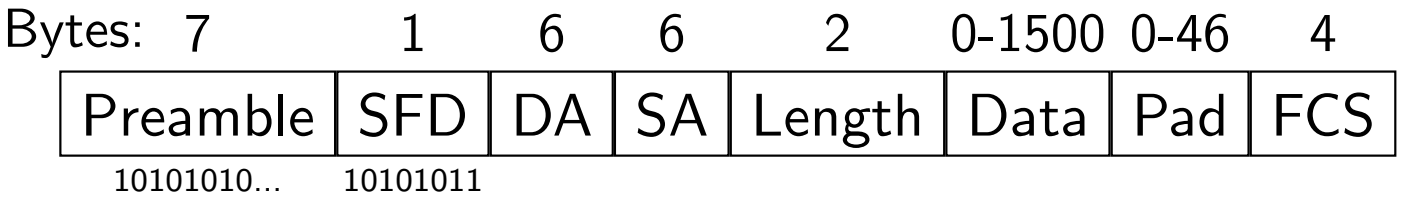
- What if 2 hosts transmit at the same time?
  - Two overlapping signals interfere; needs to be spotted as early as possible
- Collision Detection (CD)
  1. Listen to channel while sending
  2. If collision, abort signal immediately
  3. Wait random time and try again
- Properties
  - Does not waste channel sending broken frames
  - Gives unbounded time to access network
  - Designed for fair access

IEEE 802.3 - Collision Detection (CD)



- Ensure sender still sending when collision noise arrives
  - Must send for twice the propagation delay
- 802.3 allows for 2.5 km max LAN (with repeaters)
  - Specifies minimum frame length that takes at least $50\mu s$ to propagate (includes 4 repeaters)
  - Assuming 100 ns transmission delay for sending 1 bit (10 Mbps link), this means at least 500 bits, which is rounded up to 512 (safety), hence the pad

Data Frame Format

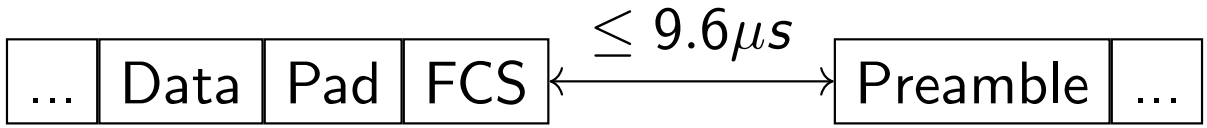


- $(7 + 1 + 2 + 2 + 2 + 46 + 4) * 8 = 512$ bits
- Takes $51.2\mu s$ on a 10 Mbps link
- Time to detect collision over longest network while still transmitting

IEEE 802.3 - Retry Timing

- Must avoid repeated collisions
  - At n^{th} retry, wait between 0 and 2^{n-1} slot times ($51.2\mu s$)
  - Do this up to a maximum of 1023 slot times
  - Give up on 16^{th} collision
- Properties
  - Low delay if frames of 2 hosts collide
  - Reasonable delay if frames of many hosts collide

IEEE 802.3 - Inter-frame Gap



- $9.6\mu s$ interval between successive frames of the same host
- Allows other hosts to use medium
- Initial frame can be transmitted immediately

CSMA/CD Summary

- Equal access and no priorities
- Unbounded access time; especially long times at heavy loads due to exponential back-off
- IEEE 802.3/Ethernet use this

Token Passing

- Arrange more orderly sharing of medium
  - Uses permission token
  - Access to medium signalled by passing token around
- Avoids collisions through strict control
  - But need to handle token control
  - Differentiate between tokens and data
  - Must handle token loss
- IEEE 802.5 Token Ring
  - Ring topology where data and token is passed around in one direction
  - When station needs to send data, it takes free token and sends data frame
  - Destination copies passing data
  - Sender removes frame on return and passes free token on
- Nice idea but complex in practice and rarely used

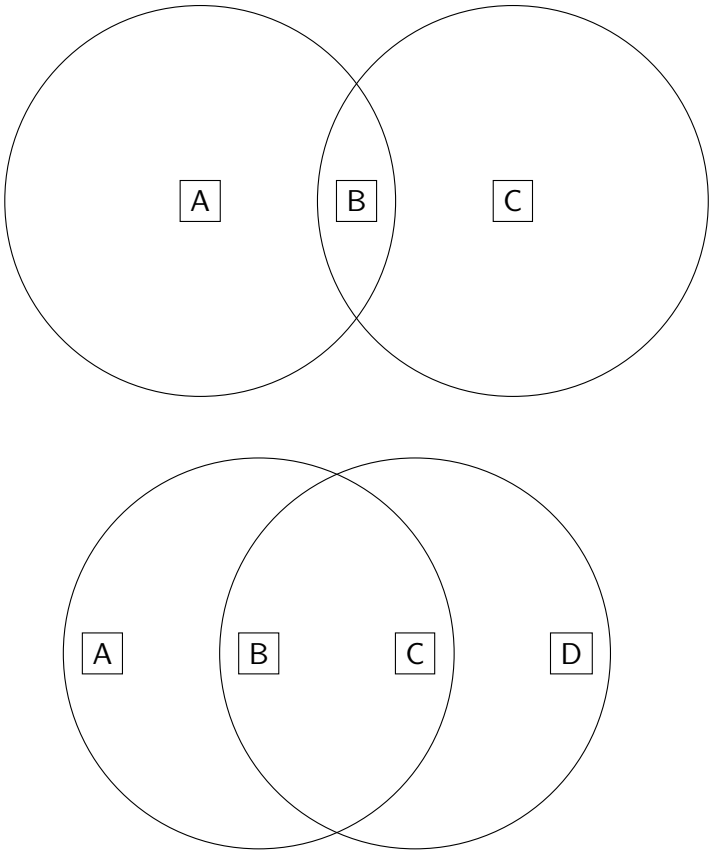
Summary: MAC in Wired LANs

- ALOHA
  - Contention based service
  - Low performance but simple, equal access
- CSMA/CD
  - Tries to avoid collisions, will detect collisions
  - Probabilistic/unbounded access time, equal access
- Token Passing
  - Avoids collisions
  - Bounded access time, access hierarchy but complex

MAC in Wireless LANs

- Centralised Medium Access Control
  - Good where data is time-sensitive or high priority
  - Suffers limits of centralisation
- Distributed Medium Access Control
  - Good for ad-hoc peers with bursty traffic
- Main challenges with distributed MAC
  - Collisions cannot be detected while sending data
  - No guarantee that all nodes can transmit frames to or receive frames from each other

Challenges

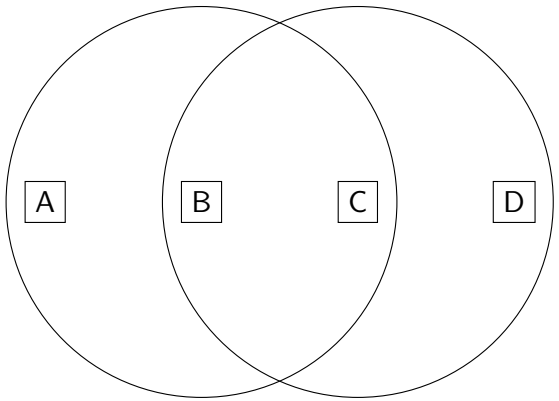


- Signals from A and C collide (at B) but A and C cannot hear each other to avoid the collision
- Hidden terminal problem

- B and C falsely conclude that they cannot send at the same time
- Exposed terminal problem

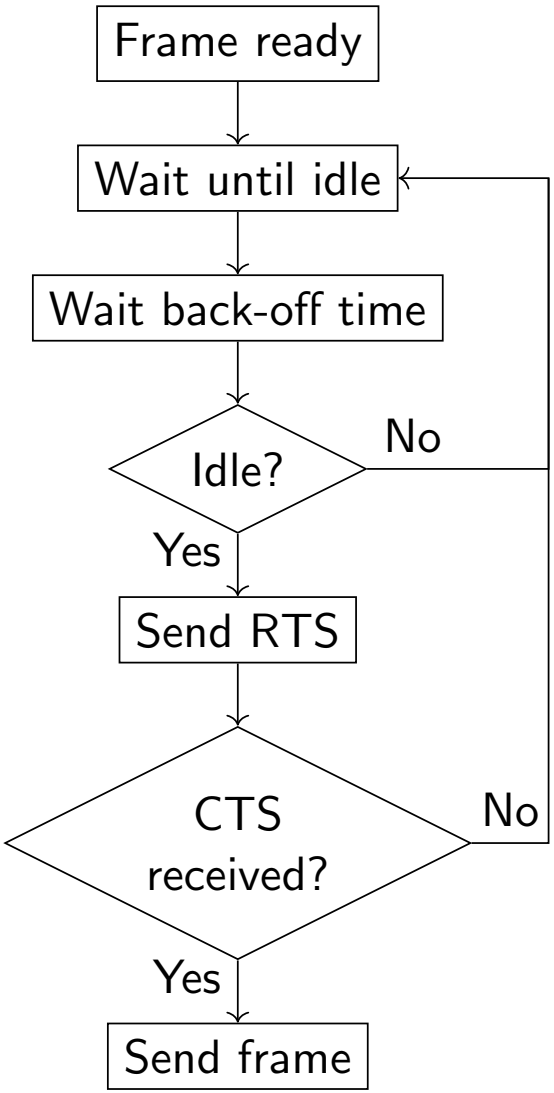
Collision Avoidance

- Idea: sender and receiver initially exchange short frames other stations can overhear
- RTS (Ready to Send) - request the channel
- CTS (Clear To Send) - response to RTS frame



- Other stations (A, D) hear exchange (RTS and/or CTS)
- Stations infer for how long the channel will be busy (RTS and CTS frames include the size)
- ACK (Acknowledgement) - sent on receipt of frame (provides efficient collision recovery)
- Repeated failures to transmit result in greater back-off time

CSMA with Collision Avoidance



- Station with frame to transmit senses medium and waits until it is idle
- When idle, wait for back-off time
- If still idle, transmit data
- If busy, wait until idle and then wait for back-off time
- Back-off time is adjusted exponentially and influenced by IFS (see next slide) and overheard RTS/CTS frames
- Distributed Coordination Function (DCF) as stations act independently
- Collisions are costly as the entire frame is transmitted

Inter Frame Spaces (IFS)

- Back-off times based on IFS
- Similar to IEEE 802.3 Inter-frame Gaps
- Help manage prioritising stations and data streams (e.g. VoIP)
- Urgent actions use shorter IFS (e.g. ACK, CTS)

IEEE 802.11 - WLAN

- Set of standards that governs wireless transmissions (physical layer, MAC layer)
- Basic Service Set (BSS)
  - Smallest building block with stations sharing medium using the same MAC protocol, aka cell
  - Access point and number of stations (laptops, mobile phones, etc.)
  - BSSes can overlap
- Extended Service Set (ESS)
  - Two or more BSSes, connected by distribution system
  - Appears as single logical LAN to higher levels

IEEE 802.11

- Operate in 2.4 and 5 GHz ISM frequency bands
- Industrial Scientific and Medical (ISM) unlicensed band
  - Regulates power usage to avoid interference (max. $1W$ and typical $50mW$)
  - Many other applications: Bluetooth, microwave ovens, garage door openers, surveillance systems, ...
- Rate adaptation (higher rate if signal is clear)

IEEE 802.11 - Standards

- 802.11b
  - Up to 11 Mbps; 2.4 GHz band
  - QPSK[1] and CDMA
- 802.11a
  - Up to 54 Mbps; 5 GHz band
  - Orthogonal Frequency Division Multiplexing (OFDM)
  - Reach is 7 times shorter compared to 2.4 GHz band
- 802.11g
  - Up to 54 Mbps; also uses OFDM but in the 2.4 GHz band
  - Better distance compared to 802.11a
- 802.11n
  - Up to 600 Mbps; operates on 2.4 GHz and 5 GHz band
  - Increase number of channels and reduce framing overhead (group frames together)
  - 4 antennas for 4 simultaneous streams
  - Uses Multiple Input Multiple Output (MIMO), to exploit multipath propagation

[1] quadrature phase-shift keying

IEEE 802.11 Summary

- Challenges of medium access control
- CSMA/CA but no CD
- Inter-frame spacing provides priority system using CSMA
- Various standards (802.11a/b/g/n)

Repeaters and Hubs

Repeaters:

- Makes two wires appear as one
- Amplifies electrical signal
- Improves signal propagation distance
- Ethernet (10 Mbps) up to 4 repeaters: 2.5 km max length

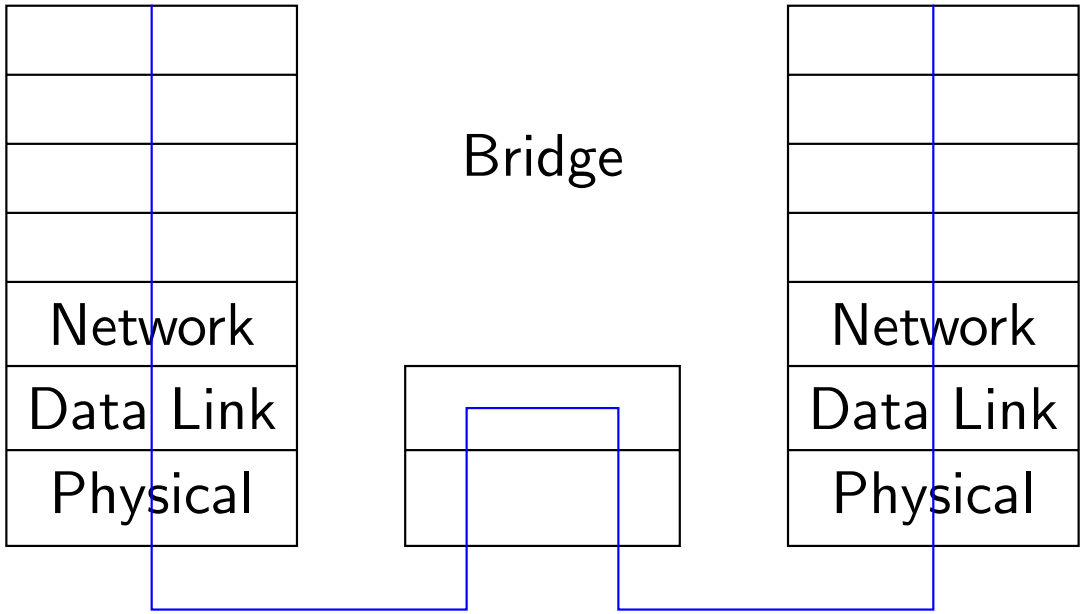
Hubs:

- Electronically join all input ports
- All ports at same speed
- Do not amplify electrical signal

Both:

- Operate at physical layer
- Transparent to higher layers
- No checking/generating of checksums
- Add to propagation delays (important to CSMA/CD)

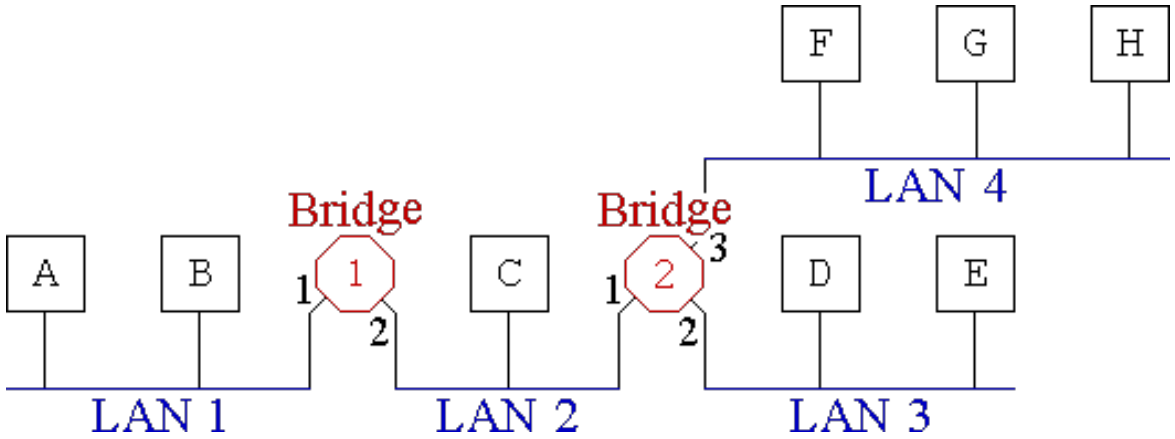
Bridge/Switch



- Operate on Data Link Layer (layer 2) and do not interpret higher layer information
- Conditional forwarding, which only forwards frames based on MAC addresses
- Bridges are used to join LANs together to make a larger LAN
- Traffic isolation

Separating Collision Domains

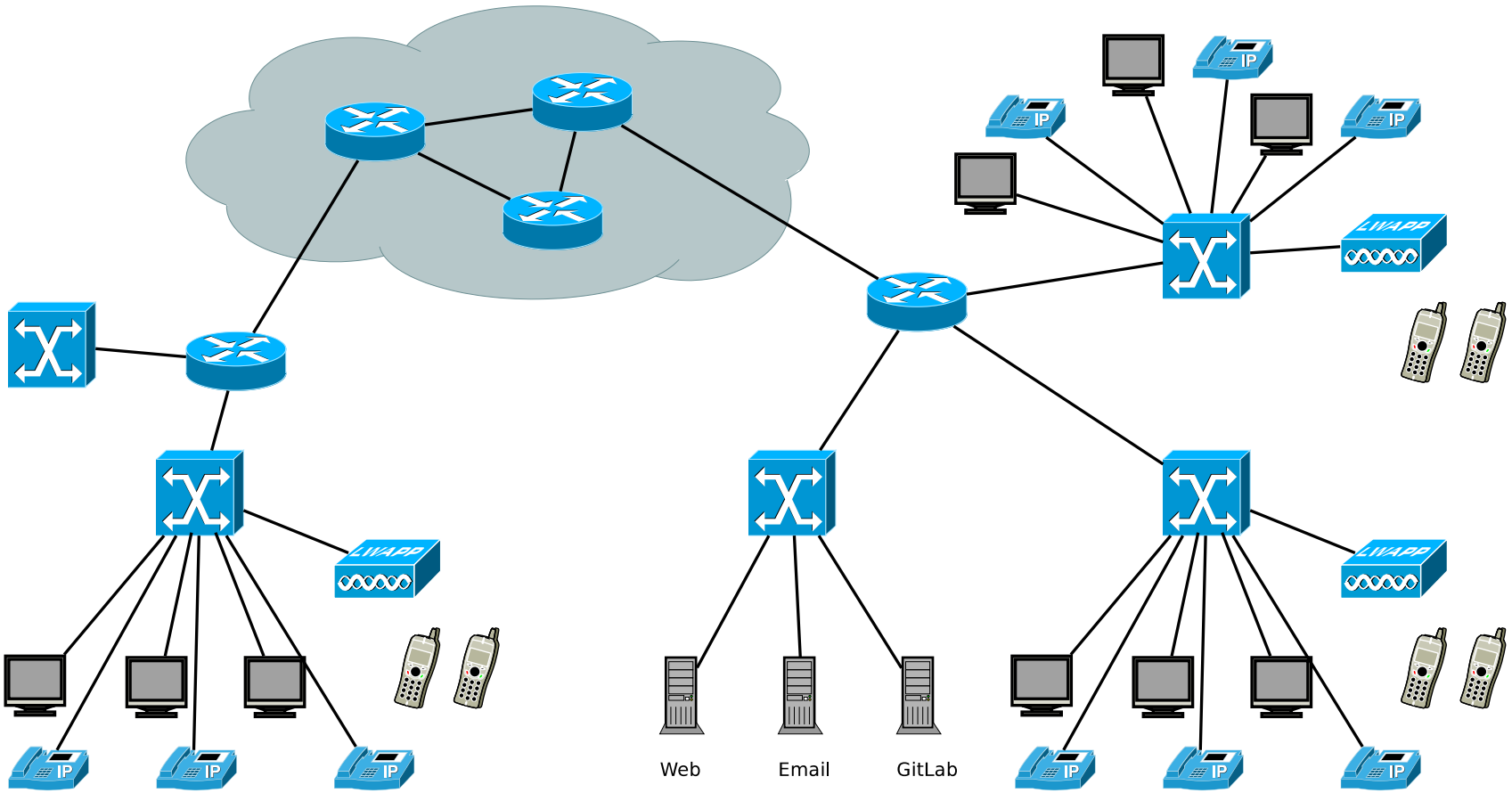
- Shared medium requires CSMA/CD to arbitrate
- Collisions can be problem on busy network
- Hosts in separate collision domains do not compete for media
- Switches form ends of collision domains (vs. hubs and repeaters)



Bridge/Switch
Cont.

- Extending physical limits
- Different speeds on different ports
- Interconnect devices that use different standards and protocols
- Separation of administration
- Idea of store and forward results in higher delay
- Ethernet calls it Ethernet Switch and these days they are also used to connect individual machines to LANs

Segmentation with Switches



- Transparent operation
  - Hosts and routers are oblivious to their presence in the network
- Keeps mapping of MAC addresses and ports in table (e.g. hash table)
- Relays frames based on table
  - If destination MAC on same port as source, do not forward
  - If port for destination MAC is known, only forward to that port
  - Otherwise use flooding (i.e. send to all ports except source port)
- Used in Ethernet

- Backwards learning
  - Listen to traffic and build address/port tables
  - Based on source and destination MAC addresses in the header
  - Purges entries that are a few minutes old to keep table up-to-date
- Loops in topology
  - Make determining location of source impossible
  - E.g. two switches that are connected by two links
- Use *Spanning Tree*
  - Switches agree on loop-free topology
- Network layer protocol often handles loops
  - Packets may have limited lifetime

- Adds transmission and processing delays
- Store-and-Forward
  - Store the entire frame and verify CRC before forwarding frame
- Cut-Through
  - Forward frame after reading destination MAC and without performing a CRC check

- Organisational changes happen regularly and require rewiring
- VLANs give a way to rewire in software
- 802.1Q: extends the header and adds a VLAN identifier
- Requires VLAN-aware switches (not backwards compatible)
- VLANs need to be configured on the switches (i.e. port/VLAN tables)