Imperial College London

Computer Networks and Distributed Systems
Part 3 – Data Link Layer

Course 527 – Spring Term 2015-2016

Emil C Lupu and Daniele Sgandurra

e.c.lupu@imperial.ac.uk, d.sgandurra@imperial.ac.uk

Part 3 – Contents

Overview of Data Link Layer
- How do we divide data into chunks for the physical layer?
- How do we control access to a physical channel?

Data Framing
- Gaps, counting, delimiters
- Ethernet/IEEE 802.3 formats

Medium Access Control
- In wired networks
  - ALOHA, Ethernet (CSMA/CD), Token Ring
- In wireless networks
  - IEEE 802.11 (CSMA/CA)

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

☛ *Not really covered in this course*

Media Access Control (MAC)
- Framing, addressing and channel access

Data Link Layer Services

1. Unacknowledged connectionless service
   - Independent frames with no logical connection
   - No recovery from loss but fast
   - Common in LANs at data link layer e.g. Ethernet

2. Acknowledged connectionless service
   - Each frame in acknowledged
   - Out of order delivery possible
   - Good for unreliable channels such as wireless e.g. 802.11

3. 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

LLC: Error Detection and Correction

Detection

Serial connections use 8 data + 1 parity bit
- 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

Forward Error Correction (FEC)
- Add more redundancy ➔ greater capacity to detect/correct bursts of errors
- e.g. use 5 bit codewords to encode 2 data bits

Data Framing

Why are frames necessary? Min size? Max size?

Need to group bits into separate messages
- Large **frames** have less overhead, but:
  - Have greater chance of collision
  - Cost more to retransmit if error detected

Need to add meta data to control protocol
- Addressing, length, frame type, CRC, …

Need to provide error detection/correction
- Physical layer may introduce errors by adding, removing, or modifying bits

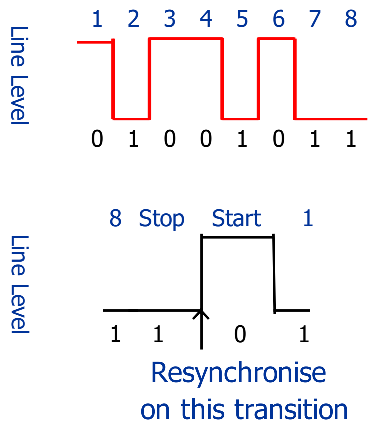
How to break bit streams into frames?

Insert gaps
- But timing hard to guarantee. How big a gap?

Count characters
- Include length field to delimit data
- Often used with another framing method

Count: 5	H	E	L	L	O	Count: 8	…

Serial Line Framing



Counting for framing
- Data transmitted at agreed rate
- But clocks may not be accurate

Use **start/stop** bits
- At least 1 transition per byte
- Bit asynchronous and byte synchronous

Framing: Flags

Start and end flags
- Special signal at start & end of frame: "FLAG"
- Search for flag if receiver loses track

FLAG	Header	Payload data from network layer	Trailer	**FLAG**	**FLAG**	Header

Uses **byte stuffing**
- Identify data with same bit pattern as the flag
- Use escape sequence to identify "next byte is data"

Stuffing

A	**ESC**	FLAG	B

C	**ESC**	**ESC**	D

IEEE 802.3 and Ethernet

Originally developed by Xerox
- Became open standard
- These days it's almost a marketing term…

Uses Manchester encoding for line transitions

Operates over various physical media
- Data link layer is separate to physical layer
- But physical layer affects parameters of Ethernet

Two standards: IEEE 802.3 and Ethernet

IEEE 802.3: Frame Format

Bytes	7	1	2 or 6	2 or 6	2	0 to 1500	0 to 46	4
	Preamble	SFD	DA	SA	Length	Data	Pad	FCS
	10101010	10101011				46 to 1500 Bytes Data and Pad		

Ethernet standard slightly different
- Doesn't have SFD field
- Replaces Length with Type field

Preamble

- 7-byte alternating 0s and 1s to establish synchronisation
- Framing then by timing, spaces between frames (96 bytes) plus counting from length field

SFD (start frame delimiter)

- 10101011 indicates start of frame
- Allows receiver to miss start of preamble and still synchronise
- Compatibility with 802.4 and 802.5 (Token Ring)

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

Type (Ethernet only)

- Identifies higher level protocol

Length (IEEE802.3 only)

- Bytes in this frame (optional)

Data

- Speaks for itself
- Includes higher layer headers

Pad

- 0-46 bytes to ensure frame long enough to enable collision detection (*a few slides away*)

FCS (Frame Check Sequence)

- CRC, based on all fields except preamble, SD and FCS
- Enables error detection

Ethernet Addresses

Usually, Ethernet addresses are 48 bits:

Bit 47: 0 = ordinary addr 1 = group addr

Bit 46: 0 = global addr (fixed in hardware), 1 = local addr (assigned by admin)

Bits 23-45: Vendor code (IEEE assigned)

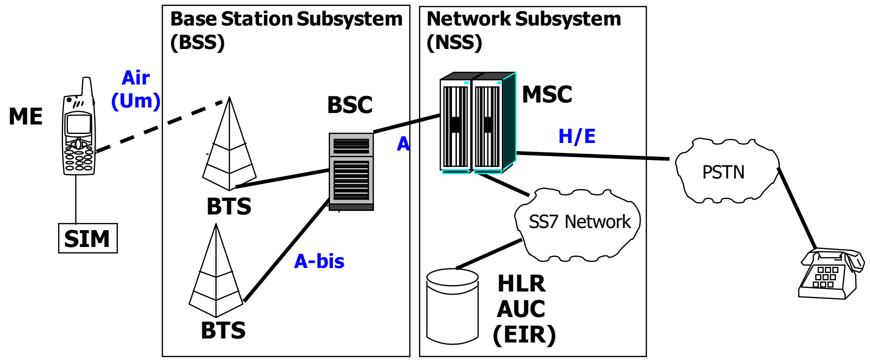
Bits 0-23: Unique code, set by vendor

$2^{46} \Rightarrow 7 * 10^{13}$ possible global addresses

Written as 6 pairs of hex digits

– e.g. 00:11:85:7A:BC:E4

Global System for Mobile Communication (GSM)



SIM: Subscriber Identity Module
ME: Mobile Equipment
BTS: Base Transceiver Station

BSC: Base Station Controller
MSC: Mobile Switching Center
HLR/AUC/EIR: various databases

More info: https://styx.uwaterloo.ca/~jscouria/GSM/gsmreport.html

GSM Station Types

Mobile Equipment (ME) (terminal)

- Fixed (e.g. in cars; max power 20W)
- Portable (max power 8W)
- Handheld (max power 2W)
  - Power down to 0.8W as technology evolves

Base Transceiver Station (BTS)

- Defines cell
- Many to deploy; need to be rugged, reliable

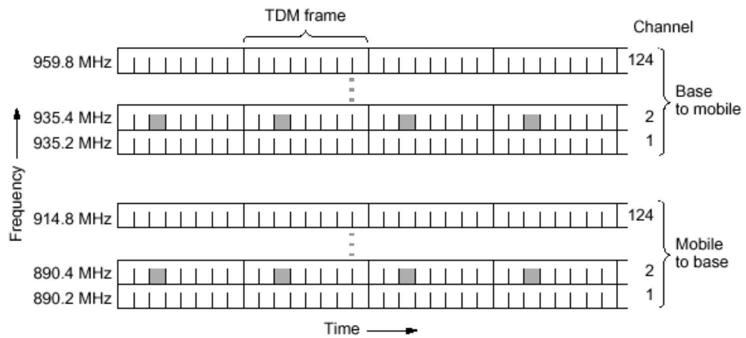
Base Station Controller (BSC)

- Manages radio resources for 1+ base stations
  - Channels, frequency hopping, handovers

Mobile Service Switching Centre (MSC)

- Acts like PSTN switch
- Mobile subscription handling
  - Registration, authentication, location updating, handovers and call routing

Multiple Access Control (FDM/TDM)



Single mobile phone connection

TDM extends number of users each frequency band can accommodate with FDM

GSM: Channels and Frames

25 MHz frequency band in each direction

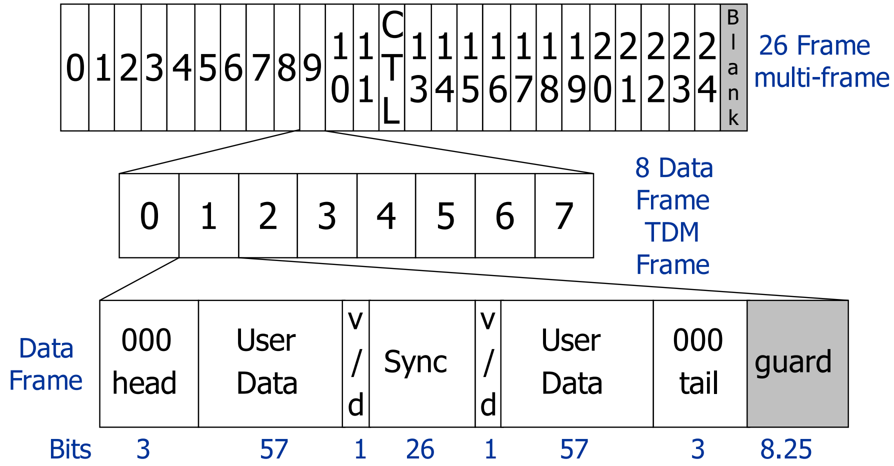
124 frequency channels (FDM) in band

- 200 kHz wide + guard channel to avoid interference

GSM Framing

- 8 data frames in 4.615ms TDM frame
- 26 TDM frames in 120ms multi-frame
  - 24 TDM frames (slots 0-11, 13-24)
  - 1 control data (slot 12)
  - 1 reserved (slot 25)

GSM Framing



Data Frame Structure

Data frame is 54μs

- Two 3-bit **head/tail** fields during power ramping
- Two 57-bit user **data** fields + 1 **data/voice** bit
- 26-bit **sync** field to synchronise frame boundaries and manage multi-path fading
- 8.25bit (30μs) **guard** to separate signals of MEs while signals ramp up

ME can send one data frame every 4.615 ms

- Downlink and uplink separated by 3 frames
- MEs need not transmit & receive at same time

GSM Error Detection/Correction

Some bits of voice data more important than others

Voice data very delay sensitive

3 levels of error correction/detection codes

- Include none on least important bits
- Add forward error correction (FEC) to others
- Replace lost important bits with previous sample

Summary: GSM

Wireless (digitised) voice network
- Also supports data circuits

Communication always via base station controller

FDM/TDM allocation of medium
- Stations assigned bandwidth by controller

Different levels of FEC for voice bits

☛ *There's a lot of GSM which we've not covered here!*

IEEE 802.11 vs. GSM

IEEE 802.11	GSM
Connectionless	Connection oriented
Packet based – Allows contention for medium	Frequency/time slots – Fixed bandwidth for data connection
Distributed and centralised MAC	MAC done by base station
Error/ACK scheme designed for data	Error correction designed for voice

☛ Very different beasts!

Medium Access Control

Physical channel supports multiplexing scheme

But how do we allocate communications channels?
- Contention
- Fairness
- Access latency

<u>Static</u> allocation vs. <u>dynamic</u> allocation

Static Allocation

In Part 2, we looked at TDM, FDM and CDMA
- Static ways for stations to access fixed part of medium

Properties
- Connection-oriented service
- Guaranteed, allocated bandwidth
- Bounded latency to transmit

Dynamic Allocation

But in many computer networks the following applies:

- Most stations do not want to transmit at once
  - Don't waste bandwidth on silent stations
- Need to ensure fair access to medium
  - Would like bounded delay to transmit
- Single transmitter on medium is simpler electronically

Use dynamic allocation

- Allocate time to use medium on demand
- Connection-less service

Use **statistical multiplexing** for TDM

Propagation Delay

Finite time for signal to go from one node to another:

delay = distance / speed (where speed=2 * 10^8 m/s)

Finite time to send each signal:

time = 1 / baud rate

Nodes will receive signal at different times

- Depends on distance from sender
- Different nodes will perceive medium to be busy/quiet at different times

Need to keep this in mind when managing who transmits when!

Collision Detection: ALOHA

Original contention network

- Developed at U. Hawaii

Send whenever data ready to go

Two stations whose signals overlap get garbled data

- By listening can detect collision
- Wait random time and try again

Not very efficient

- 18% theoretical maximum channel utilisation

Slotted ALOHA

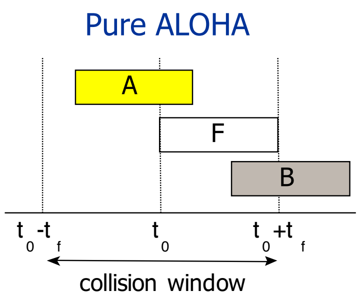
Divide time into slots, each corresponding to one frame

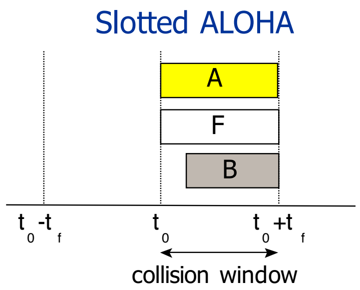
- Start time of frames is synchronised
- Probability of collisions is reduced

Cannot assume synchronised clocks between stations

- Master station sends short signal at start of each time frame

Successful transmission 37% of the time





Assume fixed frames size t_f

- Otherwise even worse performance

Collision window is $2t_f$

Max frame size = slot size t_f

Collision destroys one slot

- Remember propagation delay

Collision window is t_f

CSMA (or being polite)

ALOHA has simple problem:

- No-one listens before they start to send
- Leads to lots of collisions

Carrier Sense Multiple Access (CSMA)

- When ready to send, listen
- If channel busy, wait until idle
- When channel idle, send
- If collision, wait random time and start listening again

CSMA with Collision Detection

What if two hosts want to transmit?

- Two overlapping signals interfere, needs to be spotted

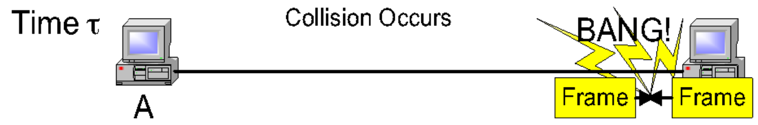
Collision Detection (CD)

- Listen to channel while sending
- When collision, abort signal with noise burst
- Wait random time and try again

Properties

- Designed for fair access
- Gives unbounded time to access network
- Doesn't waste channel sending broken frames

IEEE 802.3: Collision Detection



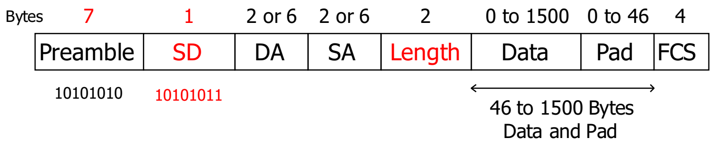
802.3 allows for 2.5 km max LAN (with repeaters)

- Specifies minimum frame length must be 51.2μs

Ensure sender still sending when collision noise arrives

- Must send for twice the propagation delay
- Assuming 100 ns transitions for sending 1 bit, this means at least 512 bits, hence the pad

Data Frame Format



7+1+2+2+2+46+4 bytes * 8 bits = 512 bits = 51.2 μs

- Time to detect collision over longest network while still transmitting
- When finished sending, cannot guarantee that still listening to hear collision

IEEE 802.3: Retry Timing

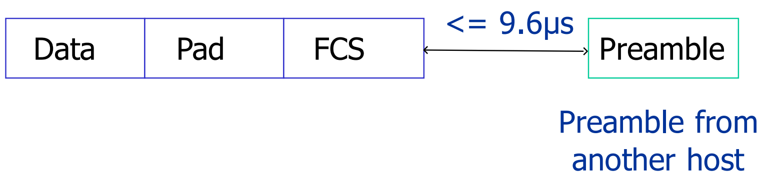
Must avoid repeated collisions

- At n^{th} retry, wait between 0 and 2^{n-1} slot times (51.2 msec)
- Do this up to a maximum of 1023 slot times
- Give up on 16^{th} collision

Properties

- Low delay for two hosts' frames colliding
- Reasonable delay for many hosts' frames colliding

IEEE 802.3: Inter-frame Gap



9.6 μs interval between successive frames from host

- Allows other hosts to use medium
- Initial frame can be transmitted immediately

CSMA/CD Summary

Fairness

- Equal access to all stations
- No priorities

Probabilistic

- Unbounded access time
- Bad at heavy loads due to exponential back-off

IEEE 802.3/Ethernet use this

Token Passing ("Collision Free")

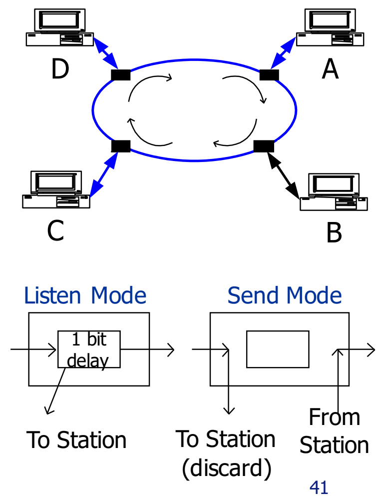
Arrange more orderly sharing of medium
- Uses permission **token**
- Access to medium signalled by passing token around

Avoids collisions through strict control
- But need to handle token control
- Differentiate between tokens and data

IEEE 802.5 Token Ring

Developed by IBM
- Token Ring and Bus

Token frame inserted by Active Ring Monitor (ARM)
1. Any station takes token and sends data frame
2. Destination copies passing data
3. Sender removes frame on return



Token Ring: Frame Format

Token (Bytes)

1	1	1
SD	AC	ED

Data (Bytes)

1	1	1	2 or 6	2 or 6	no limit	4	1	1
SD	AC	FC	DA	SA	Data	FCS	ED	FS

Starting Delimiter (SD): `JK0JK000`
- J = high-high, K=low-low transitions
- J & K are invalid in diff. Manchester encoding
  - This is another type of framing

Ending Delimiter (ED): `JK1JK1IE`
- I=1 ⇒ intermediate, I=0 ⇒ last frame
- E=0 from source, set to 1 if error detected (checksum)

Access Control (AC): PPPTMRRR
- Priority P – token priority (different access levels for hosts)
- Token T – flag indicating token or data
- Monitor M – handle failure of source to remove frame
- Reservation R – used to request priority level

Frame Control: FFZZZZZZ
- FF - indicates data / control frame
  - If FF ⇒ data, Z interpreted by destination
  - If FF ⇒ control, all hosts act on control bits, Z

Source/destination addresses
- Similar to IEEE 802.3

Data

- Variable length within token holding time

Frame Check Sequence (FCS)

- 32 bit CRC from SA, DA, Data

Frame Status (FS): ACxxACxx

- A = address recognised, C = frame copied
- Form acknowledgement for each frame
- Repeated for error robustness

Token Passing Summary

Bounded access delay for fair use

- Control passes round all nodes
- But no instant on demand access

Supports giving some stations priority over others

- Not just equal/fair access of contention like CSMA
- Good for real-time control systems

Nice idea but complex in practice ➔ rarely used

- All stations must cooperate
- Must handle token loss

IEEE 802.4/802.5 Token Bus/Ring and FDDI use this

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

<u>Centralised</u> Medium Access Control

- Good where data is time-sensitive or high priority
- Suffers limits of centralisation

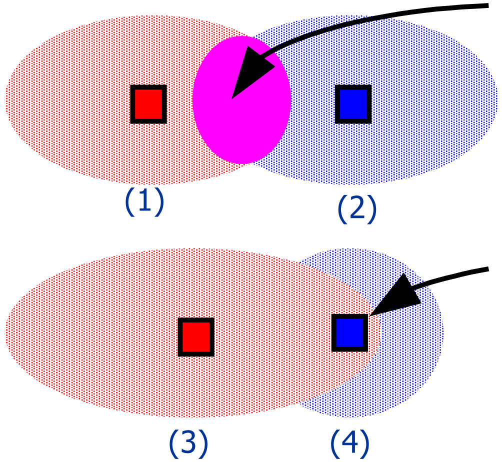
<u>Distributed</u> Medium Access Control

- Good for ad-hoc peers with bursty traffic

No guarantee that all nodes can hear each other

- Makes collision detection harder
- Collision Avoidance (CA) rather than detection (CD)

Undetectable Collisions



Signals from (1) and (2) collide

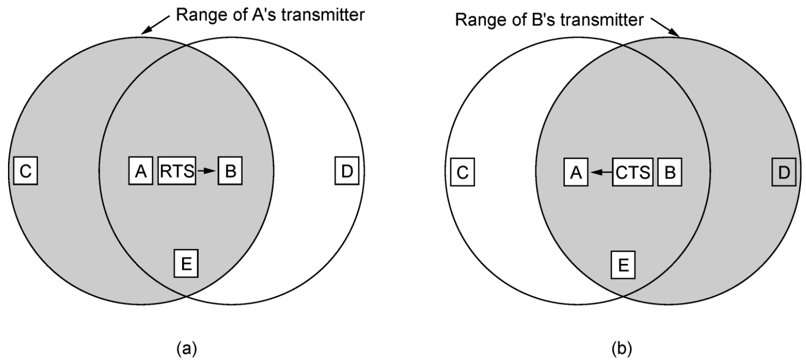
– But (1) cannot hear (2), and vice versa

Hidden receiver problem

– (4) can hear (3)

– But (3) cannot hear (4)

Basic Idea: MACA (Multiple Access with Collision Avoidance)



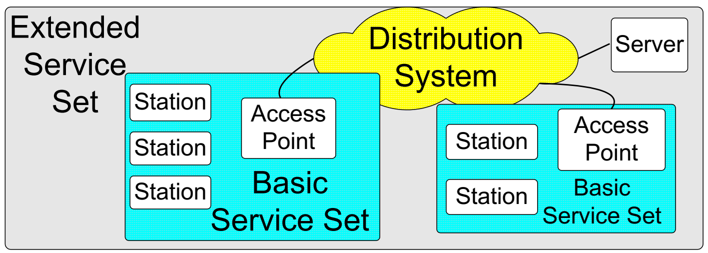
A sends RTS (Request to Send)

– Anybody who can hear A will hear RTS

B replies with CTS (Clear to Send)

– Anybody who can hear B will hear CTS

IEEE 802.11 - WLAN



Basic Service Set (BSS)

– Smallest building block with stations sharing medium using the same MAC protocol, aka **cell**

– BSSes can overlap

Extended Service Set (ESS)

– Two or more BBSes, connected by distribution system

– Appears as single logical LAN to higher levels

IEEE 802.11 Station Types

Station types based on mobility:

No Transition

– Stationary/only moves within range of one BSS

BSS Transition

– Moves between BSSes in one ESS

– Addressing must recognise new location and deliver via appropriate BSS

ESS Transition

– Moves between BSSes in different ESSes

– Does not guarantee connection to upper layers

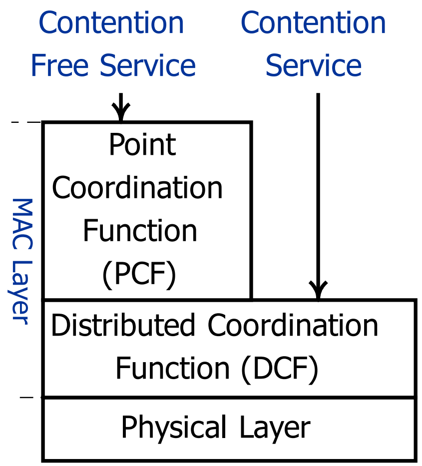
IEEE 802.11 Media Types

Frequency-hopping spread spectrum
- 2.4GHz ISM band with 20 x 1MHz hopping channels
- 1 or 2Mb/s with different FSK encodings
- Low bandwidth but good interference resistance

Direct-sequence spread spectrum (similar to CDMA)
- 2.4GHz ISM band at up to 11Mb/s (802.11b)
- Very good range and variable speed

Orthogonal FDM (similar to ADSL)
- 5Ghz ISM band with 52 narrow bands (802.11a)
- 2.4Ghz ISM band (802.11g)
- Up to 54Mb/s but lower range

Distributed Foundation Wireless MAC



Distributed access control
- With optional centralised control by base station

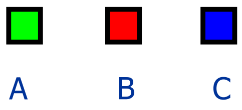
DCF uses **CSMA/CA**
- Collision Avoidance but no detection (not practical)
- Inter Frame Spaces (**IFS**) give fair access with priorities

Collision Avoidance

RTS (Ready to Send) – request the channel

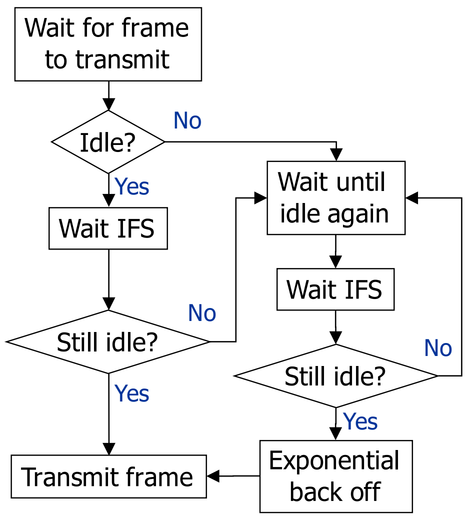
CTS (Clear To Send) – response to RTS frame

ACK (Acknowledgement) – sent on receipt of frame
- MAC-level ACK provides efficient collision recovery



Other stations hear exchange and "sense" channel
- Stations infer how long the channel will be busy
- Repeated failures to transmit ➔ greater back-off time

CSMA with Collision Avoidance



Station with frame to transmit senses medium

If idle and remains idle for IFS period ➔ transmit immediately

If busy ➔ wait until transmission ends + another IFS

If still idle ➔ back off random amount of time (exponential algorithm) + then transmit, otherwise wait until idle again

Priority and Timing

Short IFS (SIFS)

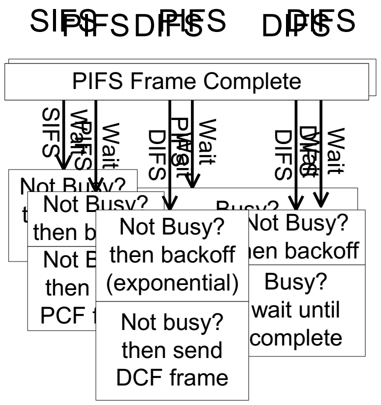
- Immediate response actions (e.g. ACK, CTS)
- Short IFS gets medium first

PCF IFS (PIFS)

- Medium length
- Polls from central controller

DCF IFS (DIFS)

- Ordinary & management data
- 1st MAC Protocol Data Units (MPDU) of series



Contention example

Multi-Frame Transmissions

Data unit broken into multiple frames

- Individual ACKs good for noisy channels

Once medium acquired send data without interruption

- Following frames sent on receipt of ACK

Use DCF IFS to initiate connection (1st MPDU)

Use Short IFS for later frames of MPDU

IEEE 802.11 Summary

Wireless LAN network

Distributed communication

- Not always via centralised controller

CSMA/CA but no CD

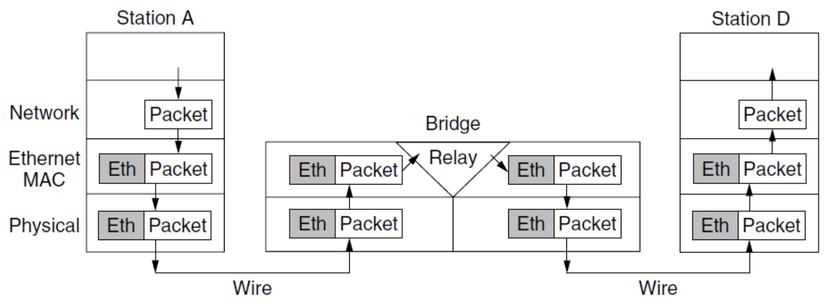
Inter-frame spacing

- Provide priority system using CSMA

Data Link Layer Switching

Join LANs together to make a larger LAN

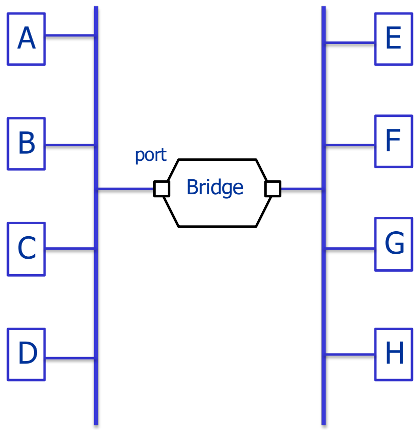
- Layer 2 level **Bridges**
- a bridge for Ethernet it called an **Ethernet Switch**
- the bridge does not interpret higher layer information



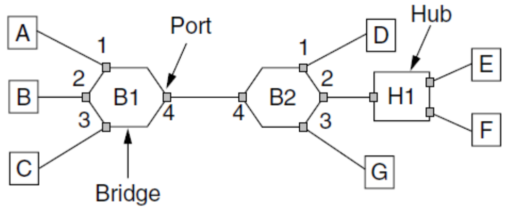
Backward Learning to only relay relevant frames.

- Keep table address / port mappings
- Start by relaying everything
- Learn which addresses are on which ports.

Spanning Tree Algorithm used to eliminate cycles in complex topologies

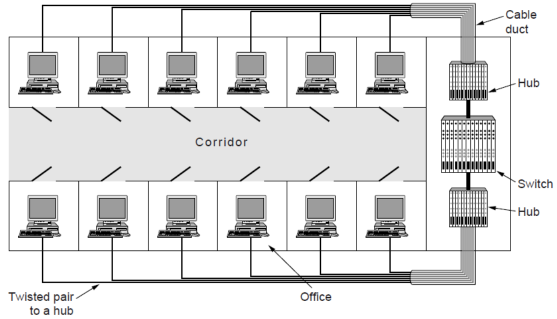


More Complex Topologies



Modern Ethernet switches can connect individual stations

Typical Office Layout



VLANs

Require configuration and changes to the Ethernet headers to use VLAN identifiers.

- Individual stations do not need to be VLAN-aware
- Remember VLANs are layer 2. To connect VLANs need layer 3 routing.