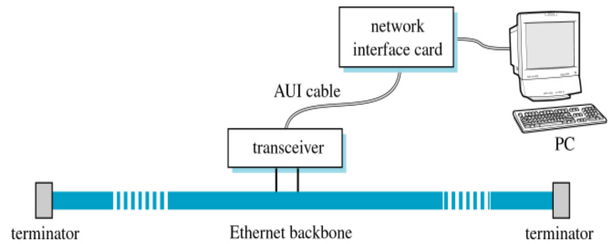
Lectures 15, 16, 17:
Ethernet - 802.3 and 802.11

Brian Carpenter

314 S2T 2009

Ethernet (Shay 9.3)

- IEEE 802.3: CSMA/CD on a shared “bus” cable
  - 802.3 is the number of an IEEE standards committee (under the main 802 committee)

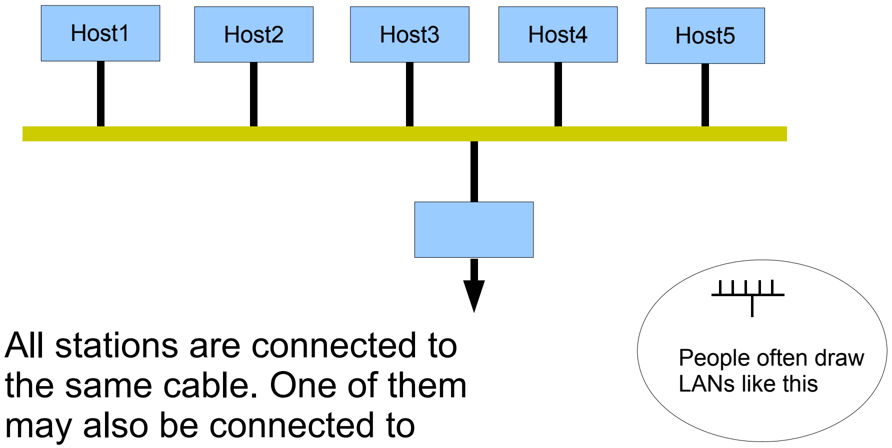


- Transceiver implements the MAC functions
- Originally 10 Mb/s on 50Ω coaxial cable with repeaters/bridges, later on UTP with hubs/switches

Original Ethernet cable and transceivers

Image from Wikimedia

Principle of original Ethernet cabling

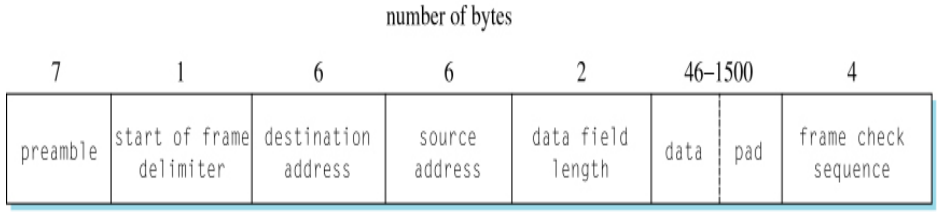


All stations are connected to the same cable. One of them may also be connected to the outside world.

Ethernet connection, step by step

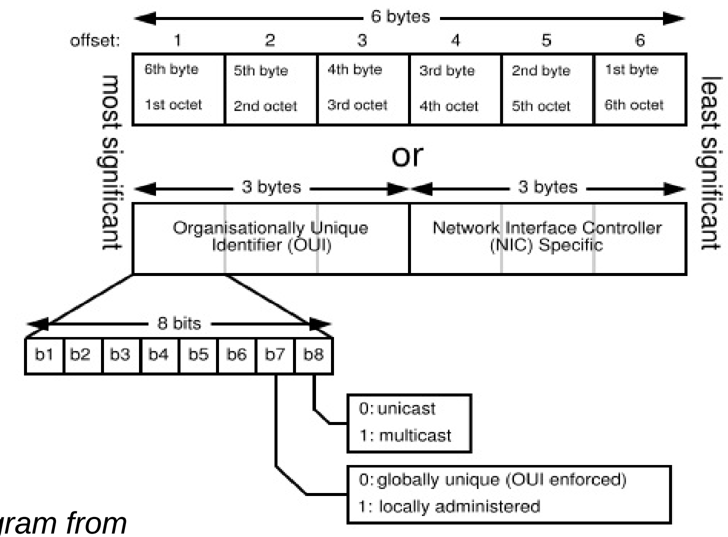
- Sending host builds a frame, sends it to Network Interface Card (NIC)
- NIC adds an Ethernet Header, waits for medium idle
- Sends packet, transceiver watches for collision. Tells NIC whether transmission succeeded or failed, NIC retries using *exponential backoff* algorithm
- Receiving host's transceiver sees packet, copies it to its NIC
- That NIC checks packet by computing CRC. If it was for this host (only, or as part of group), sends it to host via interrupt handler

Ethernet Frame, 802.2 encapsulation



- SFD and FCS are not counted as 'packet' bytes – they're not passed in to the host
  - which is why Wireshark can't see them
- Data starts with an 802.2 header (if used)
- Addresses (6-byte) are globally unique, 48 bits (MAC-48), see next slide
- Ethernet sends bytes in ascending order, bits in a byte low-order-bit-first

Ethernet Address Format (MAC-48)



- *Diagram from Wikipedia web page*

Looking at a real world address

Description: Broadcom NetXtreme 57xx Gigabit Controller

Physical Address: 00-1A-A0-4A-D6-80

OUI (manufacturer) **specific (single device)**

00 - 1A - A0

0000 0000 0001 1010 1010 0000

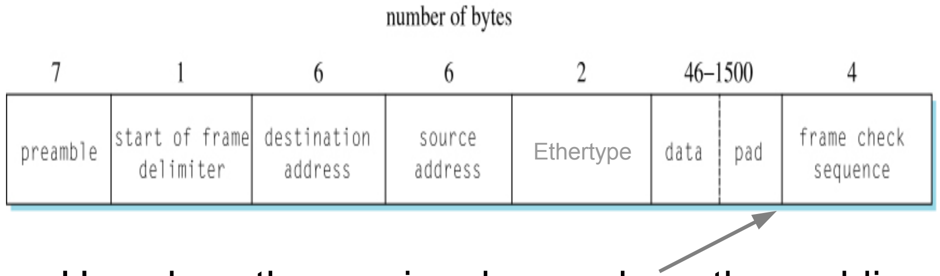
First bit on wire **9th bit on wire**

Ethernet Frame, 'native'

- One extra convention*:
  - Data Length field can carry an Ethertype instead, provided that the Ethertype value is > 1500_{10}, Ethernet's maximum packet size.
  - For example, Ethertype 0x0800 = 2048 (IP)
  - Length <= 1500 means that an 802.2 header follows
  - (In other words, this is a trick to avoid having to use an 802.2 header)

* This comes from the original industry standard that preceded the official IEEE standard. It saves bits, so is widely used.

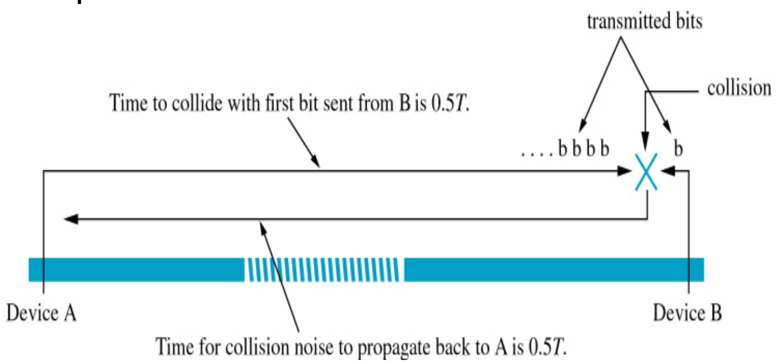
'Native' frame format



- How does the receiver know where the padding ends and the FCS starts?
  - there's no length field in the frame
  -

Detecting Collisions

- Max packet size stops a host from monopolising the medium
- Min packet size set for reliable collision detection



- Packet must take at least time T to transmit.
  T = 2 x (cable length) / (speed of signal)
  Packet size ≥ T x (bits per second)

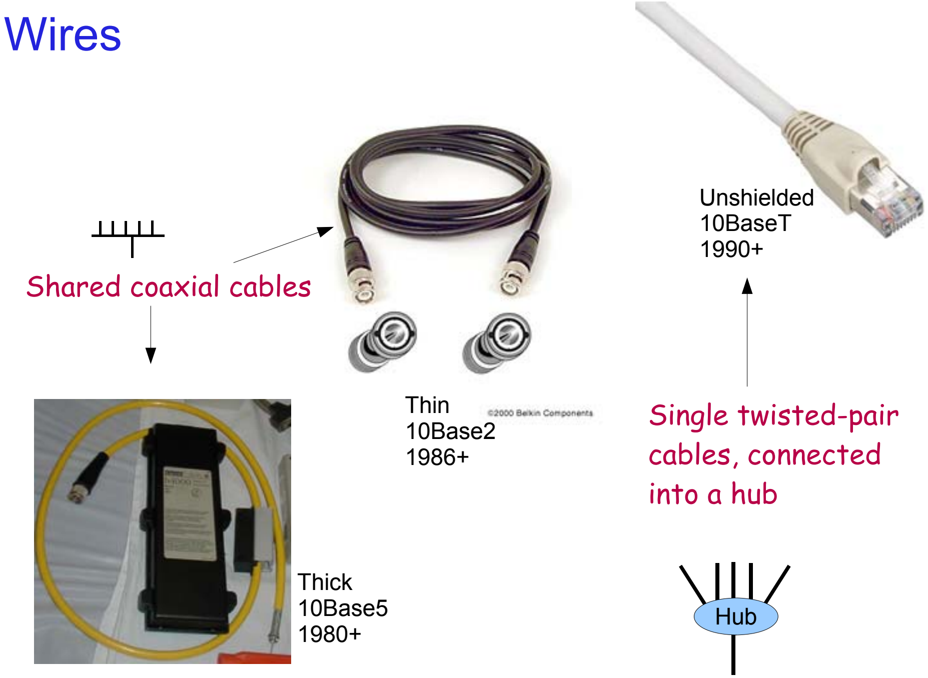
10Base5 Ethernet Specifications

- CSMA/CD occurs within a *collision domain*
  - Max *segment length* 500m
  - Max of *four repeaters* joining five segments
  - Collision domain = 2.5km
- 2 x 2500m/(2 x 10^8) m/s = 25 μs round-trip
  - Add 25 μs for (worst-case) repeater delay
- T = 50 μs at 10 Mb/s = 500 b, plus a few more
- 512 b = 64 B
- Min inter-packet gap is 12.5 μs (i.e. 2.5km of cable) for 10, 100 and 1000 Mb/s Ethernet

Physical Implementations

- 10Base5 = Thick Wire
  - thick coax, vampire taps, AUI on (50m) AUI cable
- 10Base2 = Thin Wire
  - thin coax, tee connectors, AUI built into NIC
- 10BaseT = UTP (unshielded twisted pair) wire
  - max UTP cable length 100 metres
  - UTP into hubs (multiport repeaters) or switches
  - no collisions in switches, allows full-duplex working
  - status pulse to verify link is connected (flashing *link light* on NIC) *[see Wikipedia for details]*

Wires

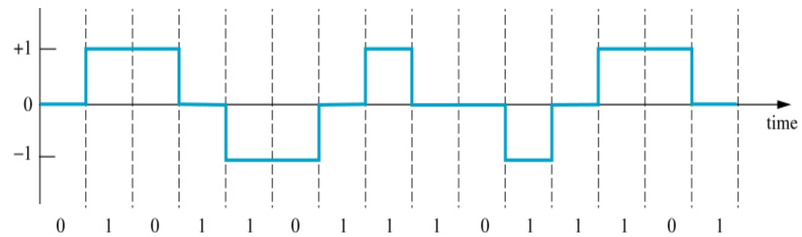


Fast (100 Mb/s) Ethernet (Shay 9.4)

- 100BaseTX standardised (802.3u) in 1995
- Changes to go from 10 to 100 Mb/s on UTP:
  - couldn't use NRZI encoding directly at 100 Mb/s, too much RF interference (noise)
  - 4B/5B block encoding for each *nibble*, so as to ensure short 'same-bit' runs (Shay Table 9.3)
  - e.g. 1010-0010-0000-0000-0000-0000 becomes 10110-10100-11110-11110-11110-11110
  - that reduced the noise, but not enough to allow use of NRZI
  - MLT-3 signaling ..

Fast (100 Mb/s) Ethernet (2)

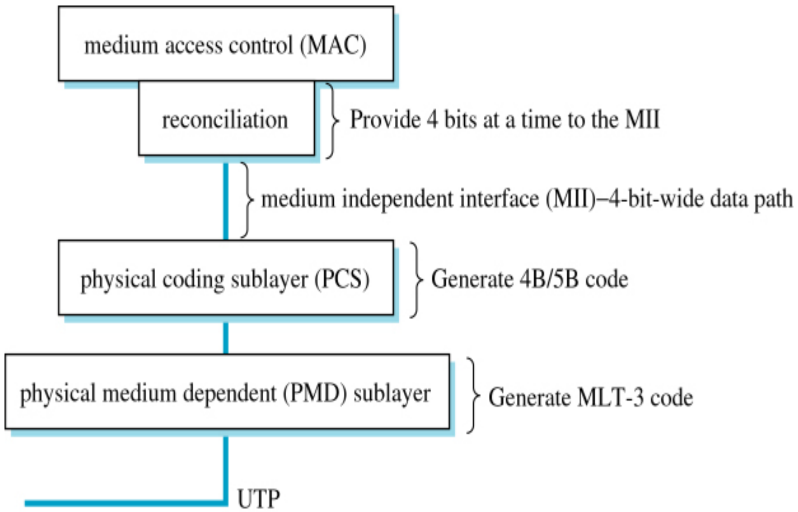
- MLT-3 signaling, Multilevel Line Transmission – Three signal Levels



- MLT-3 cycles through -1, 0, 1, 0, -1, ...
  - for a 1 bit, progress to next state
  - for a 0 bit, maintain same state
- Uses 25% max frequency compared to Manchester, works well over UTP

Fast (100 Mb/s) Ethernet (3)

- 100BaseTX physical layers



100BaseT4

- 100 Mb/s Ethernet on four Category 3 UTP cables
- Not widely used today

100BaseFX – 100 Mb/s on Fibre

- Multi-mode or single-mode fibre
- Segment length 412 metres if collisions can occur, 2 km in full duplex (i.e. using switches)
- Uses 4B/5B bock encoding, same as for UTP
- Uses NRZI signaling instead of MLT-3
- Normally use ST fibre connectors
  - ST connectors just push in
  - SC (an older type) is a bayonet-style connector

Collision Domain

- 10Mb/s Ethernet used a minimum frame size of 512 bits, (transmitted in 51.2 μs) for a maximum segment length of 2500m
- 100Mb/s Ethernet transmits a frame in 1/10 the time, so the max segment length decreases. For 100BaseTX it is only 100m
- 1GB/s Ethernet would require even less!

Gigabit Ethernet (Shay 9.5)

- Collision Domains again ..
  - 1000BaseX (fibre, 802.3z) and 1000BaseT (twisted pair, 802.3ab) allow collisions
  - when collisions are possible, need to use a longer minimum frame so as to keep 100BaseTX's maximum segment length of 100m
  - do that by using a min frame of 4096 bits, i.e. extra padding on short packets
    - can also send a group of packets back-to-back as a 'burst frame,' only the first packet needs to be 4096 bits long
  - collisions are not possible in full-duplex mode; that uses 512b minimum frames (same as earlier standards)

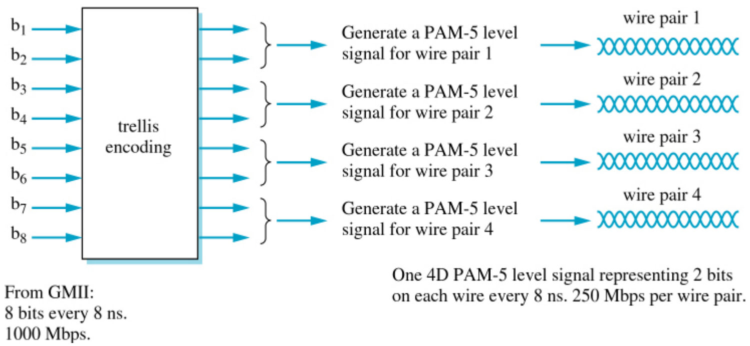
1000BaseX

- Gigabit Ethernet on fibre (or coax cable)
- Similar to 100Mb/s Ethernet, but uses GMII*
  - 8-bit-wide data path instead of 4
  - 1 bit of data (on all 8 lines) every 8 ns
- Uses 8B/10B block encoding instead of 4B/5B
  - code symbols are chosen so as to provide *DC balance*, i.e. equal numbers of 0s and 1 over the *long term*
  - has two encoder states and two alternate mappings for each symbol: 'more 0s' and 'more 1s'

*Gigabit Media Independent Interface

1000BaseT

- Gigabit Ethernet over Category 5 UTP
  - Note: 1000BaseTX is a different standard *[not widely used, see Wikipedia]*
- Much harder for UTP than fibre because of its high signal frequencies
- Uses all four twisted pairs in Cat5 cable to carry 250 Mb/s each

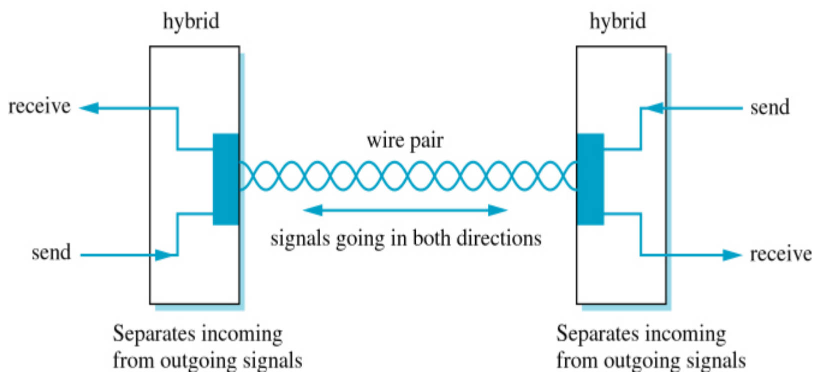


1000BaseT (2)

- 1000BaseT does *not* support half-duplex
- Each GMII octet is divided into four 2-bit groups
- 5-level signalling – PAM5 – is used to send the 2-bit groups. Having 5 levels provides support for some control functions
- Cat5 isn't quite able to carry this reliably, so the link needs error-correction codes to allow for possible errors
  - *trellis encoding* sends extra information, *Viterbi decoding* detects and corrects errors
  - we're not going into the details!

1000BaseT (3)

- All four Cat5 twisted pairs used for data
- Full-duplex carried over each pair at the same time using *hybrids* to combine/separate the signals

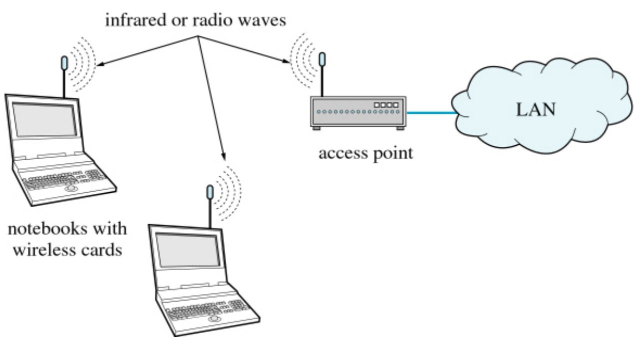


10 Gb/s Ethernet

- 802.3ae only works in full-duplex on fibre
- Standard specifies two physical layer types
  - LAN-PHY – for use in LANs
    - e.g. 10GBaseLX4, 300m
  - WAN-PHY – for linking LANs over a wide area
    - e.g. 10GBaseER, 40km
    - an alternative to SONET or ATM

Wireless Networks (Shay 9.7)

- 802.11 standard; link medium is radio or infrared
- Infrared can bounce of walls and ceiling, radio penetrates through walls
- Normally use one or more *access points* to provide connectivity to movable hosts



The wireless family

- 802.11 refers to a family of standards for wireless networks, 802.11a, b etc.
- Often called WLAN (Wireless LAN). Sometimes carelessly called “Wireless Ethernet”
  - Different from Ethernet, but the programming model viewed from Layer 3 is like Ethernet
- Marketed as “Wi-Fi”
- 802.16 refers to a future family of broadband wireless standards marketed as "WiMax" or "WiBro"
- There are other standards (Bluetooth, Zigbee)

Wireless basics

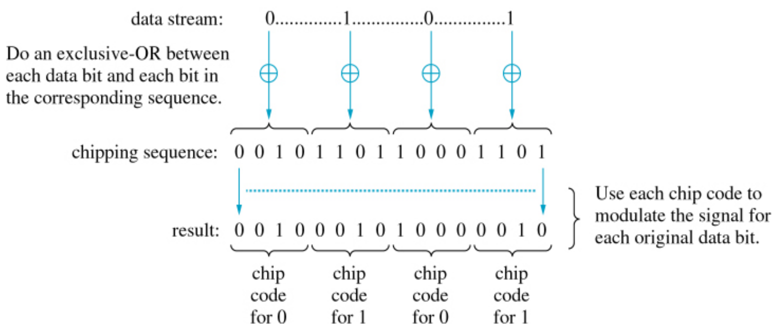
- Low power radio signals in 2.4 & 5 GHz bands
  - penetrate thin walls but bounce off concrete walls; effective range is tens of metres
    - 1 GigaHerz = 1000 MHz = one billion cycles/sec
- Infrared signals only work over a metre or so and any solid object blocks them
    - 802.11 over infrared is defined but really not very interesting...
- Bits can be modulated onto the radio wave using frequency modulation techniques
- The 2.4 GHz band is highly subject to interference (unregulated spectrum)
  - Many packets can be damaged in transit

Spread Spectrum Wireless

- Used by 802.11 to minimise interference and (maybe) provide (a little) security
- Two technologies: Frequency Hopping (FHSS) and Direct-Sequence (DSSS)
- FHSS:
  – use a set of frequencies (channels)
  – hop between them in an agreed pseudo-random sequence
  – 802.11 uses 79 channels and 22 hopping sequences

Spread Spectrum Wireless (2)

- DSSS (includes CDMA):
  – for each transmitted bit, send a *chip,* i.e. an n-bit pseudo-random sequence, as illustrated in this diagram



  – effect is to generate a high-bandwidth signal, that signal is modulated onto a 2.4 Ghz carrier
  – each station uses a different chipping sequence

Collision avoidance

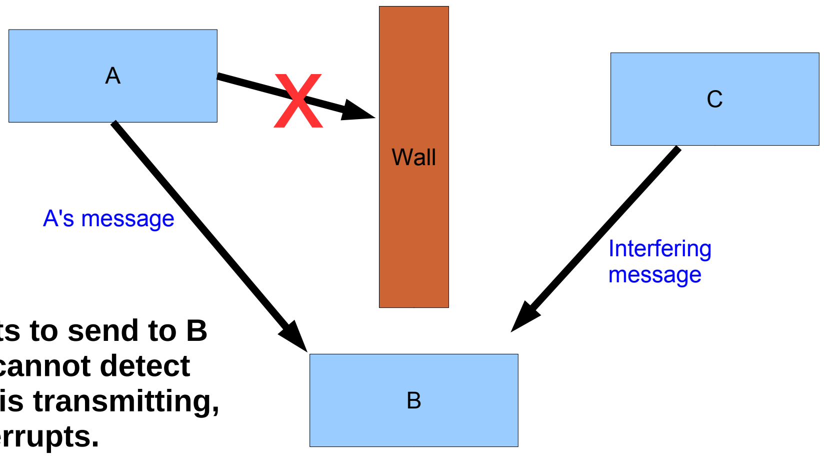
- Ethernet works by collision *detection* and retry
  - Drive out of the intersection, and get a new car if you crash
  - It's cheap to resend a packet
- 802.11 works by collision *avoidance*
  - Honk and listen before you drive out
  - Wireless transmission is expensive in battery-operated devices, so collision and retransmission is undesirable
  - A cheap radio can't detect collision anyway (its own signal drowns any incoming signal)
- CSMA/CA starts out like CSMA/CD
  - Wait until the channel is empty (no radio signal detected)
  - But then send a brief “I'm coming” signal and transmit if the channel stays empty

A better way to leave a stop sign

Contention, *Hidden Station* Problem

- Access Point (AP) can hear all stations, but they can't necessarily hear all of them
- That means they can't always detect a collision
- 802.11 has 'Distributed Coordination Function (DCF)' that implements CSMA/CA, i.e. Collision Avoidance even with hidden stations
- Next slide illustrates what happens when station A wants to send a message to station B ..

A hidden station interfering



A wants to send to B but C cannot detect that A is transmitting, so interrupts.

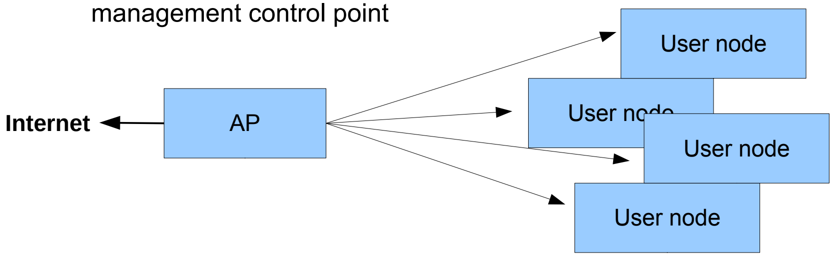
Solution: A and B exchange short "request to send" and "clear to send" packets. C missed the RTS but hears the CTS and keeps quiet. (Optional and not found on cheaper equipment.)

CTS/RTS Protocol

- All devices are contending for the medium. A waits until medium not busy, waits DIFS seconds and sends RTS to B
- B receives RTS and responds with CTS back to A; *however,* it waits for SIFS seconds (a little less than DIFS) before sending. Any other host wanting to send an RTS will wait for DIFS seconds
- If two hosts send RTS at same time the RTS messages will probably collide at B, so B will sense the collision and won't send CTS
- When A receives CTS it knows it has the medium and can send data. When B receives the data it replies with ACK
- Transmission from A to B is now complete, all hosts go back to contending again

Topology choices

- Access Point (AP) or “infrastructure” mode
  - AP also acts as management control point

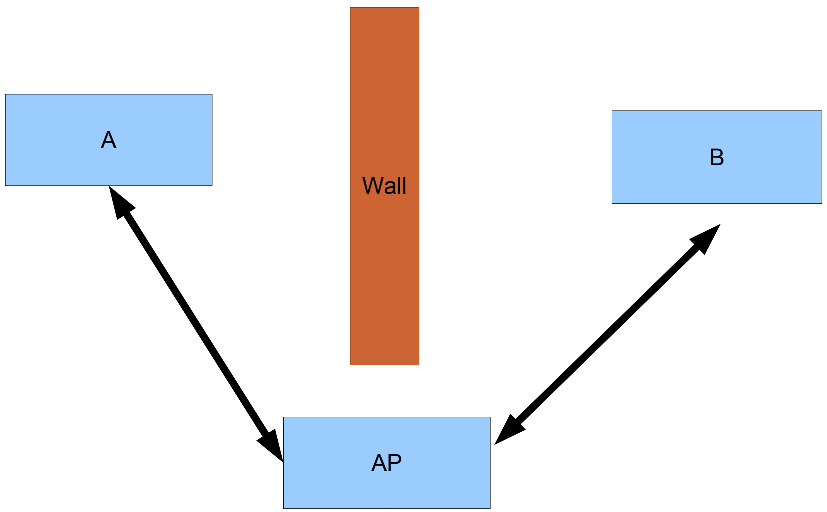


- Ad hoc mode (no AP)
  - To be avoided – operational nightmare

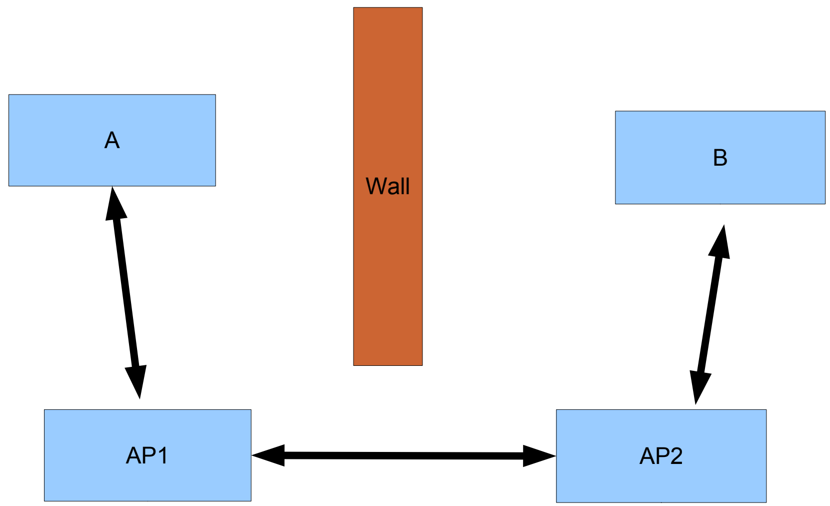
More topology

- If A can see the AP but can't see B, the AP must relay packets
- Multiple AP's may be connected to form a single network – then packets must be relayed from one AP to another
- Several wireless networks may overlap – they are distinguished by a network identifier (SSID or (Basic) Service Set Identifier)
  - When a new device joins the network, it will do so by requesting the SSID announced by the AP
  - When the AP accepts the request, the device is said to be “associated” with the AP

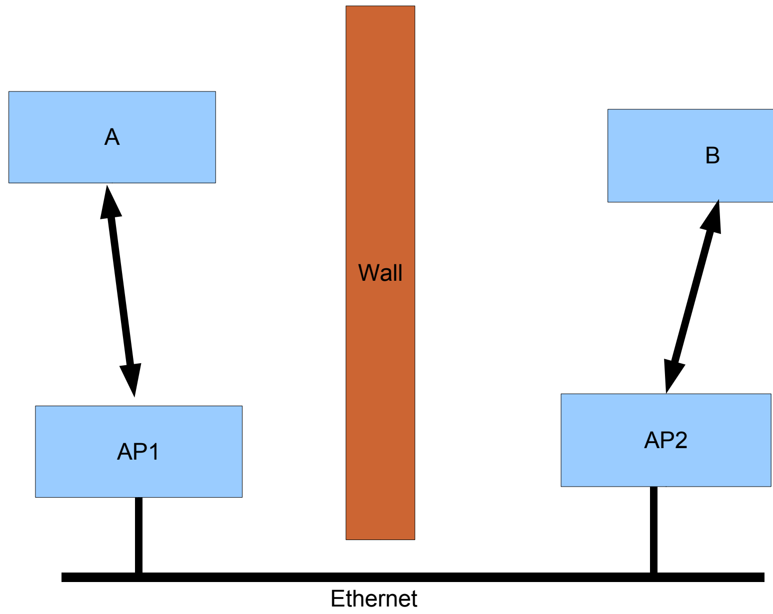
Access point as a relay



Indirect relay

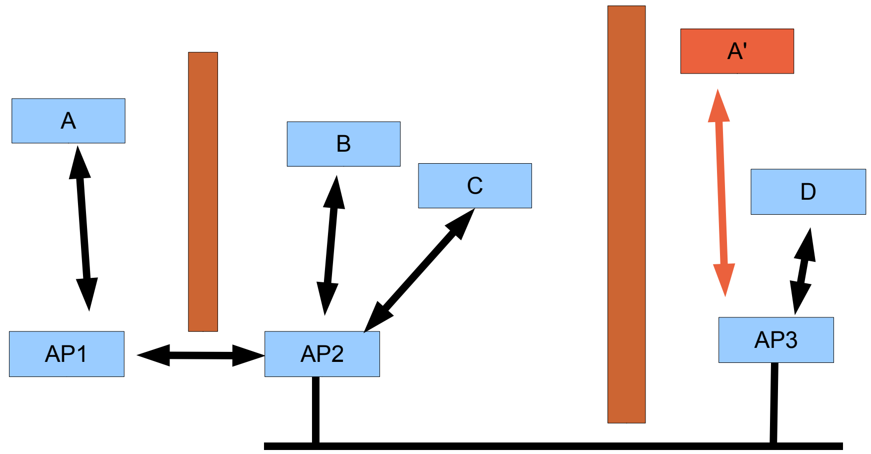


Indirect relay via cabled LAN



A quick look at the roaming problem

A is talking to B* but suddenly walks over to A'
- restore wireless connection
- restore network connection
- resume session with B



* B could also be elsewhere on the Internet; in fact that's more likely.

802.11 Addressing

- All the stations that communicate with a single AP form a Basic Service Set (BSS) identified by SSID
- BSSes may be connected via a (wired) Distribution System (DS)
- Externally, addressing looks like Ethernet: packets are sent from a source address to a destination address.
- But to allow for AP relaying, 802.11 frames have *four* address fields selected from
  - Destination Address (DA)
  - Source Address (SA)
  - Sending Wireless Access Point address
  - Receiving Wireless Access Point address
  - wireless network identifier (SSID)
- Usage listed in Shay Table 9.9.

802.11 Frame Format

bytes: 2	2	6	6	6	2	6	0–2312	4
Control	Duration	address1	address2	address3	sequence control	address4	data	CRC

- Duration: time message will require (for RTS/CTS frames)
- Control: includes ..
  - More Fragments bit. 802.11 may decrease max frame size, fragmenting and reassembling frames as needed. That's done to increase probability of error-free communication
  - To/From DS bit. Set for frames to/from the Distribution System
  - frame Type field. Distinguishes data / control / management frames. RTS, CTS and ACK are control frames

802.11 Management Frames

- Used for:
  - configuring a BSS; *Associate* Request/Response
  - find an AP; *Probe* Req/Resp
  - roaming; *Reassociate* Req/Resp
  - security; *Authenticate* frame, for exchanging security information [keys?]

802.11 Security/Privacy

- Obviously, a wireless network can be received by anyone in the area, so security is needed except for public-access networks.
- WEP – Wired Equivalent Privacy - specified in 802.11
  - WEP is a simple authentication/encryption scheme using RC4, a 40-bit secret key and a 24-bit initialisation vector. Each message uses a different initialisation vector.
  - Supposedly it makes 802.11 as safe as an Ethernet cable.
  - True; both can be tapped! WEP was “broken” in 2001
  - WEP can be cracked because the initialisation vector sequence may repeat often if traffic is heavy, and 40 bits is a rather short key.

Fixing the WEP weakness

- Quick fix in 2003 known as WPA (Wi-Fi Protected Access), also based on RC4
- 802.11i (= WPA2) is a better solution using AES.
  - Key exchange preceded by 802.1X authentication
- But any wireless network, including Wi-Fi, is a security headache - so we need security in higher level protcols, above layer 2

Some variants of 802.11

- 802.11b
  - choice of channels in the 2.4 Ghz band
    - legal channels vary between countries
  - max data rate 11 Mb/s
- 802.11g
  - As 11b but max data rate 54 Mb/s
- 802.11a
  - choice of channels in the 5 Ghz band
    - legal channels vary between countries
    - less interference than b/g but covers less distance
  - max data rate 54 Mb/s
- 802.11n is coming

Hints for setting up a WLAN

- You don't *need* to broadcast your SSID (WLAN identifier), but the standard says you should.
- It's simple to configure an AP to recognise only a small set of 802.11 MAC addresses
  - but then your packets are not encrypted, so "Eve" can monitor them.
- WEP is better than nothing.
- WPA is better.
- WPA2 is even better.
- If you do none of these things, your neighbours *will* borrow your bandwidth.