50005 - Networks and Communications - Lecture 8

Oliver Killane

18/04/22

Lecture Recording

Lecture recording is available here

Preamble

Device Terminology

Device	Layer	Description
Repeaters/Hubs	Physical	Boost signals by repeating all received.
Switches/Bridges	Data Link	Make interconnection decisions based on **MAC Addresses**.
Multi-Protocol Routers/Gateways	Data Link	Forwards (and possibly fragment) packets. Use **IP** addresses.

Transport Layer and **Application layer Gateways** also exist.

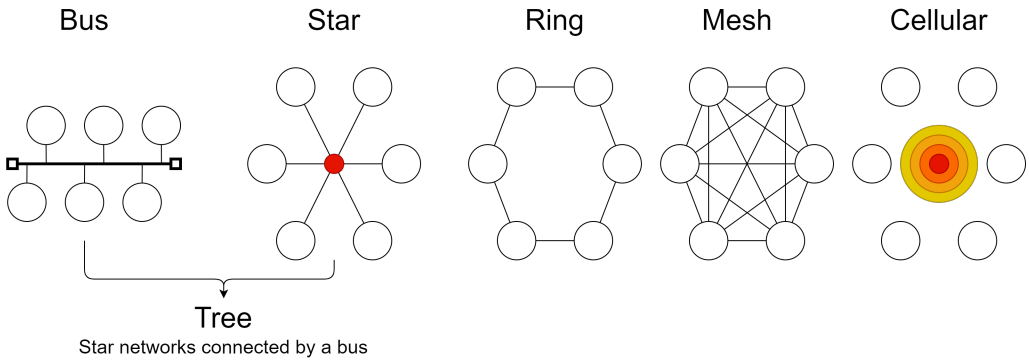
Routers act as **Gateways** to connect **IP**-based networks.

Network Types

Network Types ordered by size (small → large).

PAN	Personal Area Network
LAN	Local Area Network
MAN	Metropolitan Area Network
WAN	Wide Area Network

Network Topologies



Data Link Layer Protocols

802.3	Ethernet	LAN	1-persistent CSMA/CD	Star/Bus
802.4	Token Bus	LAN	Token Passing	Bus/Tree
802.5	Token Ring	LAN	Token Passing	Ring
802.6	DQDB	MAN	Distributed Queue	Bus
802.9	isoEthernet	LAN	Ethernet + ISDN	Star/Mesh
802.11	WiFi	LAN	CSMA/CA	Cellular
802.12	100BaseVG	LAN	Handshaking from hub	Star/Tree
802.15	Bluetooth	PAN	Adaptive FHSS	Cellular
802.16	WiMAX	MAN	Connection oriented	Cellular
802.17	Resilient Packet Ring	LAN to WAN	Distributed Queue	Ring

Ethernet

Definition: Ethernet

A **Data-Link Layer** protocol used for **LAN/MAN/WAN** communications.

- Specification created in 1980.
- Became IEEE Standard 802.3 in 1983.
- Originally coaxial cable (**10BASE5**), $\approx 2.94 Mbps$
- Currently fibre optic, twinaxial (two coaxial) cable $\approx 100 Gbps$.

Ethernet Cables

UTP	Unshielded Twisted Pair
STP	Shielded/Screened Twisted Pair
FTP	Foiled Twisted Pair
SFTP	Shielded & Foiled Twisted Pair

The most popular is **UTP** of which the most used version is **Cat5e**.

Cate6a,**Cat7a** exist and **Cat8** in development.

Cables use shielding to protect against **ElectroMagnetic Interference** (**EMI**) (e.g crosstalk between the wires) causing errors in data transmission.

Lantenna Attack

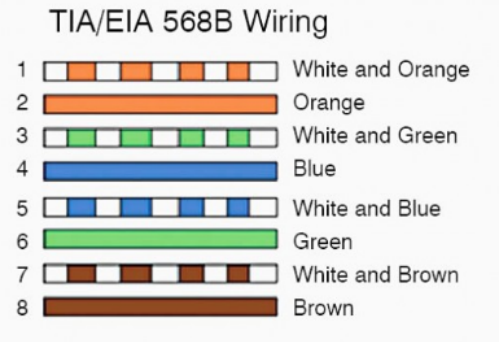
Sheilding can also help to reduce electromagnetic leakage from ethernet cables that can be sniffed and exploited.

A team from Ben-Gurion Univeristy have demonstrated this by attacking their own basic air-gapped networks. Their paper can be read here.

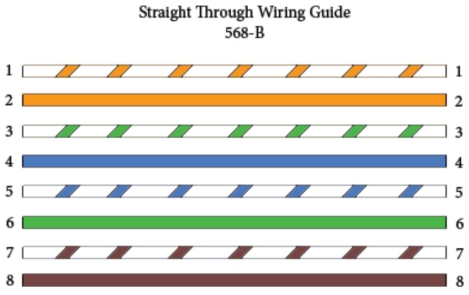
Cable Code	Name	Cable	Max Length	Topology
10Base5	Thick Ethernet	1/2 inch coaxial cable	500	Bus
10Base2	Thin Ethernet	75-Ohm coaxial cable	180	Bus
10BaseT	Twisted Pair Ethernet	Category 3 UTP	100	Star
100Base-TX	Fast Ethernet	Category 5 UTP	100	Star
100Base-FX	Fast Ethernet	Fiber Optic	185	Star
1000Base-T	Gigabit Ethernet	Category 6 UTP (4 pairs)	100	Star
10GBase-T	10 Gigabit Ethernet	Category 6a UTP (4 pairs)	100	Star

Ethernet Pinouts

There are two main pinout wirings, the only differences are the yellow and green cables are in swapped positions:



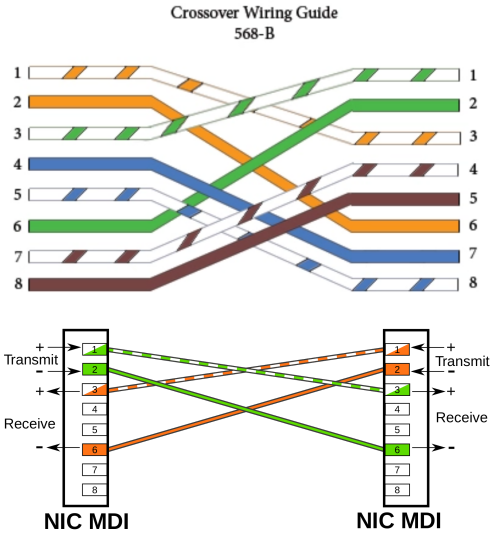
Straight Through



Communication between different OSI layers (e.g switch → router).

This is also called the **Media/Medium Dependent Interface** (**MDI**)

Crossover

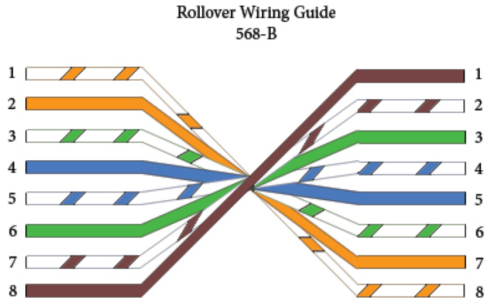


Communication between devices on the same OSI layer (e.g switch → switch).

This is also called the **Media/Medium Dependent Interface with Crossover** (**MDIX**)

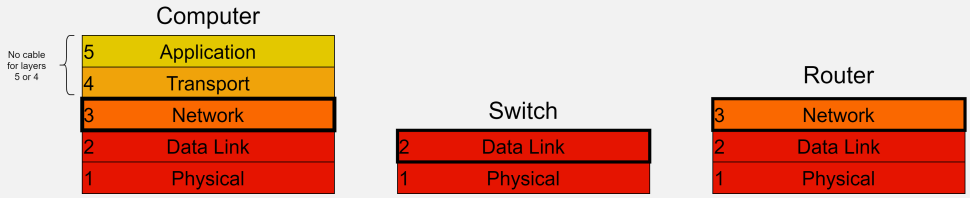
The crossover connects the transmit on one side, to the receive on the other side and vice-versa. As the devices are the same OSI layer, they can communicate back and forth through this.

Rollover



Used to directly tap into a network device (e.g a console to debug a router setup issue)

Example: Connect Computer to Router



When we connect to a router, we will be connecting to a switch (with a router attached).

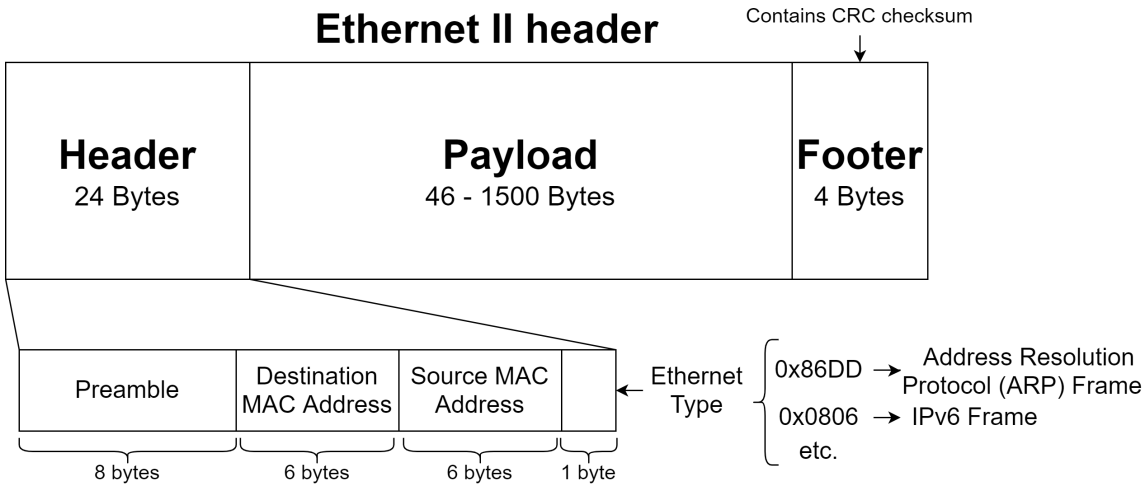
We consider the computer as **Network Layer** as that is the highest we can directly connect, and the switch as **Data Link Layer**. Hence a straight-through connection is used as they are different **OSI Layers**.

Ethernet Frame

Definition: Octet

A byte/8 bits. It is used as an unambiguous term as on older machines the definition of a byte was hardware dependent (e.g like the term word). The wikipedia page for byte contains many interesting sources with examples ranging from 1 bit to 48 bit bytes).

Nowadays a byte is always 8 bits, so can be used interchangably with **octet**

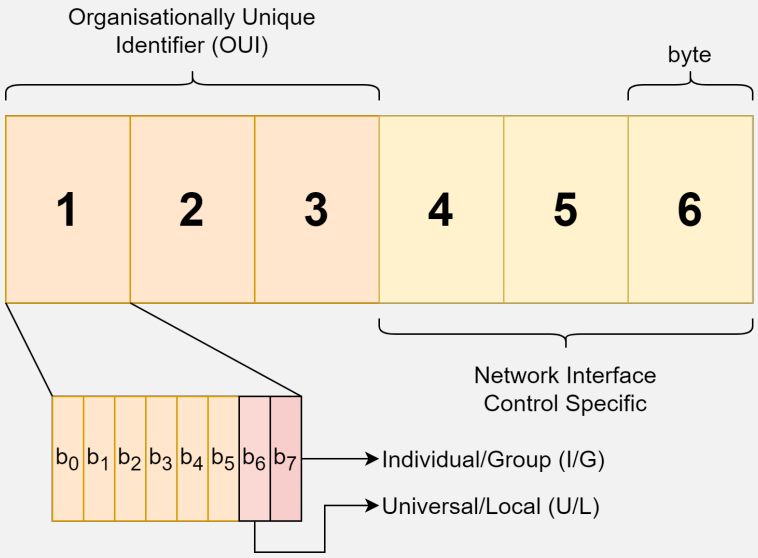


IEEE MAC Addressing

Definition: MAC Address

A 48 bit (6 byte) address etched into IEEE 802 conforming **NIC**s as a unique identifier an **NIC**.

$$< byte >:< byte >:< byte >:< byte >:< byte >:< byte >$$



I/G	0	Unicast, intended for one **NIC**.
	1	Multicast, **NIC** accepts based on other information (e.g list of accepted multicast addresses).
U/L	0	Globally Unique, enforced by the **OUI**.
	1	Locally Unique, e.g new MAC address burned in by network administrator.

A special address is the **Broadcast Address** $FF:FF:FF:FF:FF:FF$ (group Multicast, locally administered).

It is possible to filter sniffed traffic using mac addresses in applications such as **WireShark**.

eth	All Ethernet based traffic.
$eth.addr == 12.00.06.14.3a.fe$	All traffic to/from MAC Address $12:00:06:14:3a:fe$.
$eth.addr != < MAC\ Address >$	All traffic except $< MAC\ Address >$.
!(. . .)	Filter for all except frames traffic matching . . . .
$(eth.dst[0]\&1)$	Multicast only (bitwise and of 1 and the first byte).
$!(eth.dst[0]\&1)$	Unicast only.
$(eth.dst[0]\&2)$	Locally Unique addresses only.
$!(eth.dst[0]\&2)$	Globally Unique addresses only.

Switch

Definition: Switch

Allow many devices to be connected to the same subnet.

- Forwards messages to ports based on the MAC address of the device connected to each port.
- If the switch cannot determine which port to send to, it will send to all (flood).
- Uses a **Forwarding Information Base** (**FIB**) **MAC** table to remember addresses associated with ports.
- Difficult to network-sniff as packets are only directed to intended recipients.
- Can connect them to other **switches** or **hubs**, allowing networks to be connected together.
- Replaced network Bridges.

To allow messages to leave the subnet, a router and an **IP Address** provided by a **DHCP service** or set statically.

Store-and-forward Switching

Once a whole frame is received, check its integrity using the checksum. If it is correct, forward to the correct port based on the frame's destination **MAC Address**.

- Forwarding is slower, as the switch must wait to recieve the entire frame.
- Can check for errors at the switch, and drop the frame if invalid.
- Supported by **bridges** and **switches**.

Cut-through Switching

As soon as the enough information is received (e.g the destination address), start forwarding the packet.

- Faster forwarding.
- Does not error check, so final reciever must check footer checksum.
- Only supported by **switches**.

Wireless

Definition: Wireless Access Point (WAP)

Standardised by IEEE 802.11 for wireless communication.

- Uses $2.4Ghz$ or $5Ghz$ radio (open for unlicensed use).
- Acts as a hub (repeats all recieved)
- Can connect **WPAs** together to extend range.
- Can also act as a bridge to connect to a wired network.

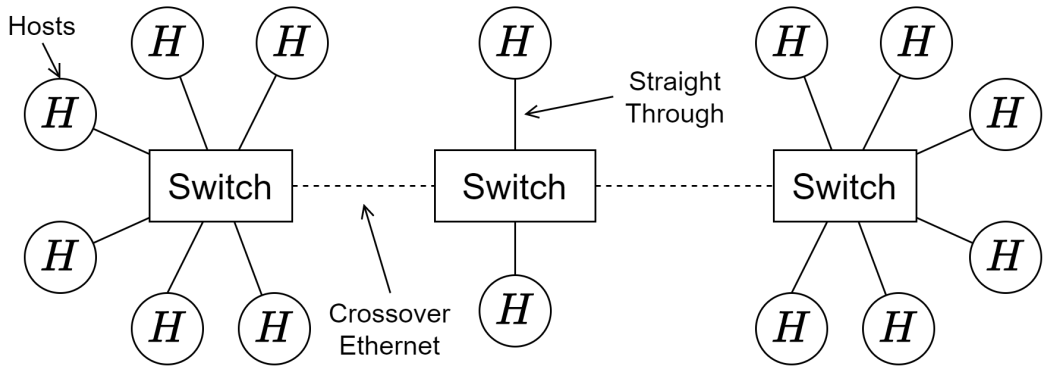
Note that it is very easy to network-sniff, as all devices within range of the **WPA** can receive frames.

Topologies

Lecture Recording

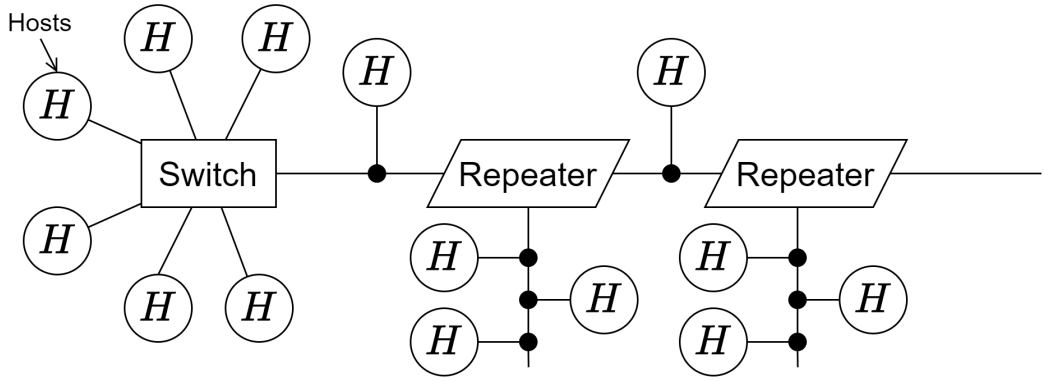
Lecture recording is available here

Switched Ethernet



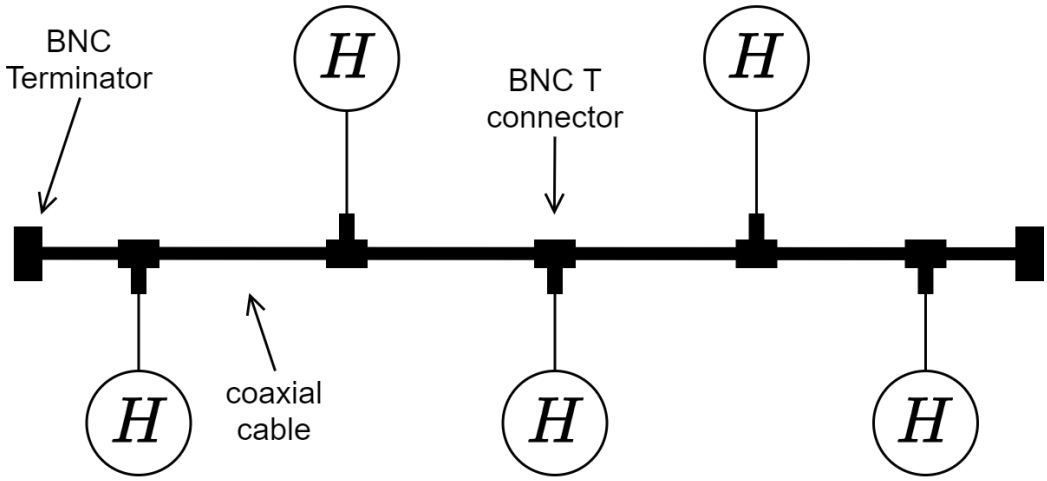
- Each switch port connected to another machine, or a host.
- Collisions avoided using small collision domains (the hosts on a single switch)
- Ideal, but expensive.

Internetworking Ethernet



- A combination of networks sharing a medium (e.g a cable).
- Repeaters boost signal to extend the range of the network (longer length of cables).
- Hubs are used (forward recieved frames out of every port) (use generally discouraged).

Bus

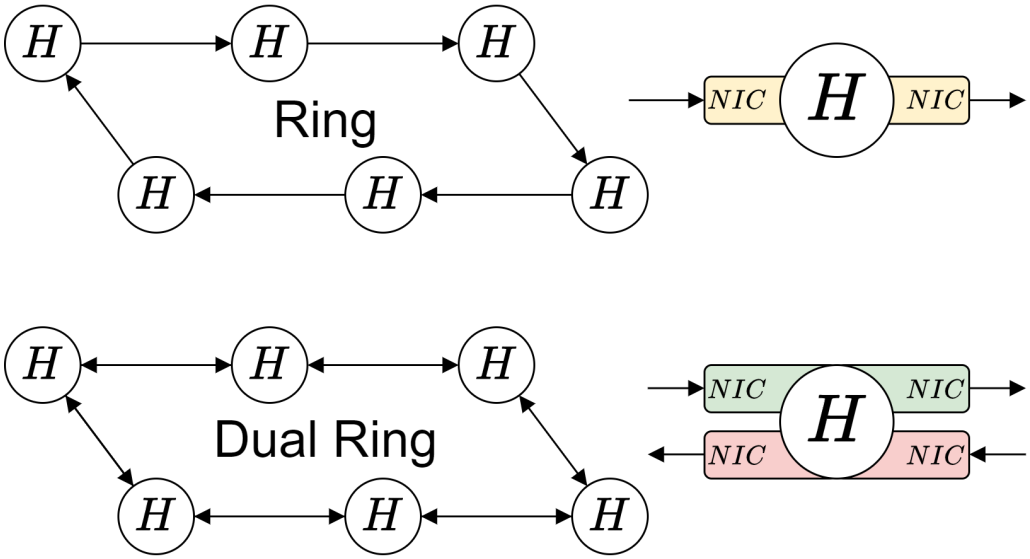


- Data travels up and down the **bus** coaxial cable.
- To add new hosts, the cable must be cut, and a **BNC T** connector used to connect the cable back together with the new host.
- Terminators at the end of the cable absorb signals, preventing them from being reflected back down the cable.

Cheapernet

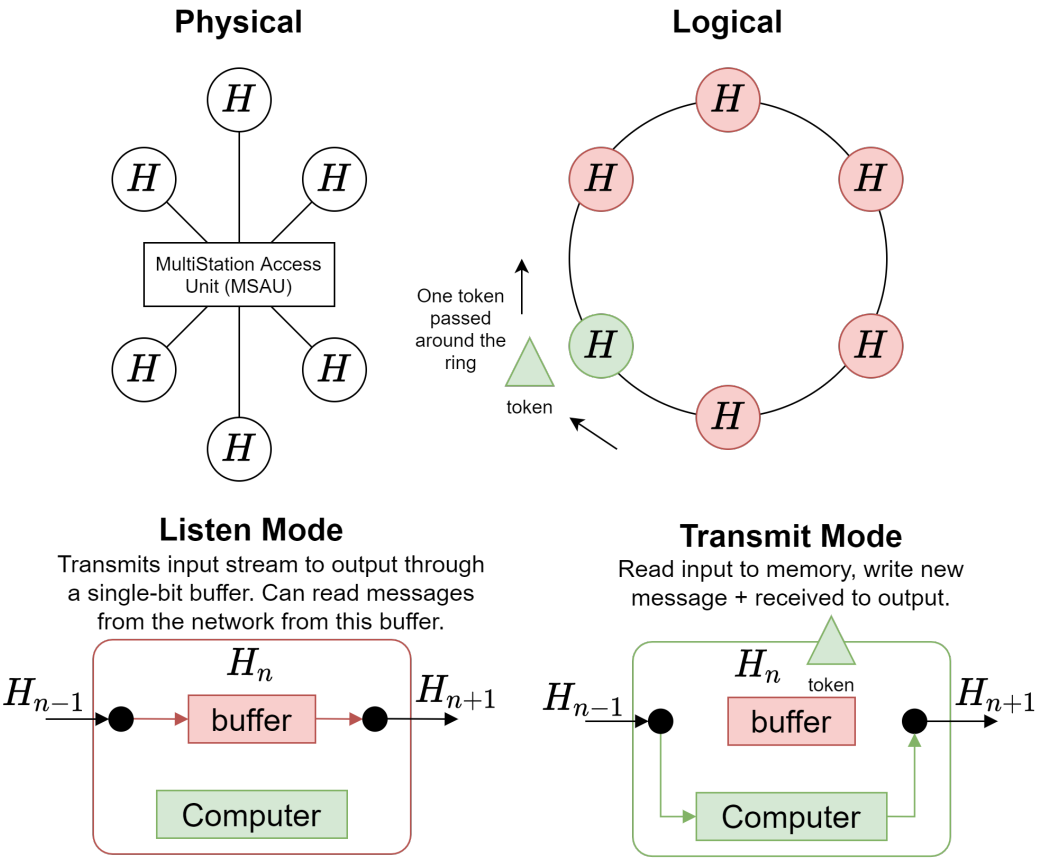
10BASE2 is the code for the coaxial cable used for ethernet on **LAN**s. It is rarely used. The wikipedia page covers the connector types and its replacement.

Ring



- Each host needs two **NIC**s for **Ring** and four for **Dual-Ring**.
- If a link is removed, the network fails (every host is a single point of failure) unless the network is designed to adjust (change flow, become a bus).
- For **single-ring** data flows one way, for **dual-ring** both ways.
- **Dual-Ring** allows for one ring being cut.

Token Ring (IEEE 802.5)



- Hosts connected to the **Multistation Access Unit** (**MSAU**), connected logically as a ring.
- One host has the token at a time. When a host has the token it is in **transmit mode** and can write to the network.
- No collisions as only one host can have the token & write at a time.
- All hosts can listen to the network (**listen mode**).
- Do not have to worry about frames not fitting inside the ring, as the host holding the token buffers with its own memory.

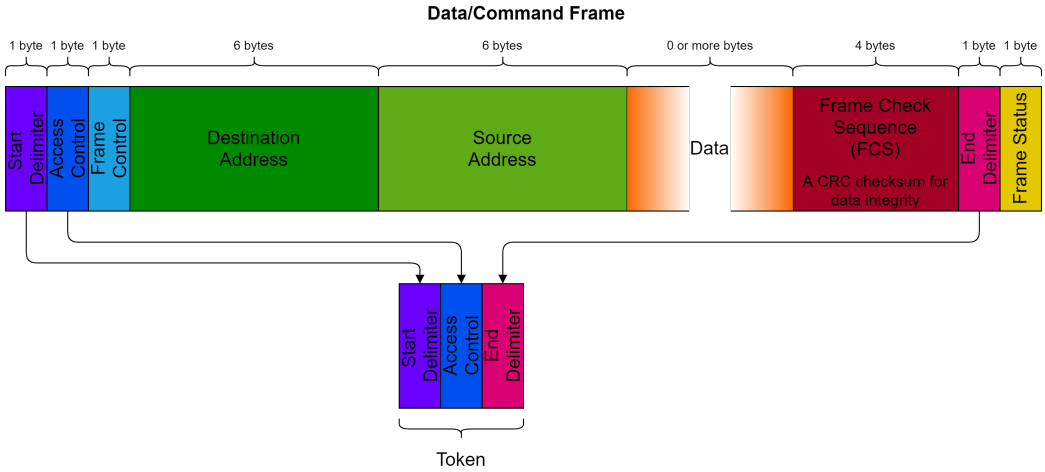
When a host has the token and wants to transmit new data:

1. Direct all recieved frames to memory.
2. Write new frame/s to the output.
3. Wait until the frame written is recieved (meaning it has traversed the entire ring)
4. Write all recieved (except frame written) to the output.
5. Pass token to next host.

Definition: Early Release

Hosts do not wait for a sent frame to traverse the ring before passing on the token. This increases the performance considerably.

Token Ring Frames



When there is no frame to be sent, the token is passed around the circle.

Frame Check Sequence	**FCS**	A **CRC checksum** used to ensure data integrity.
Frame Status	**FS**	Defines if the address was found & frame received/copied from the ring (set by **listening mode** destination host). When set it can be removed from the ring.
InterFrame Gap	**IFG**	Gaps between frames, smaller gap means less inefficiency, but need enough to recognise the start and end of frames. Defined by the protocol used.

Token Ring Priority & Reservation

A priority scheme can be implemented so hosts can only claim the token if the priority level of the the data they want to send is as high as the token's priority.

A host in **listen mode** may have high priority data to send. It can raise the reservation priority in the frame. When a token is created, it is created with the priority of the reservation bits in the frame.

- Can claim token if priority of data is as high as token priority.
- Low priority data may be delayed indefinitely.
- High priority data will be sent quickly (good for real-time applications).
- Used for **LAN**s, so it is assumed we can trust hosts to not abuse priority.

Token Ring Acknowledgement

The receiver can alter the **Frame Status**:

$A = 1$	Destination host is working.
$C = 1$	Destination host has correctly read the frame.

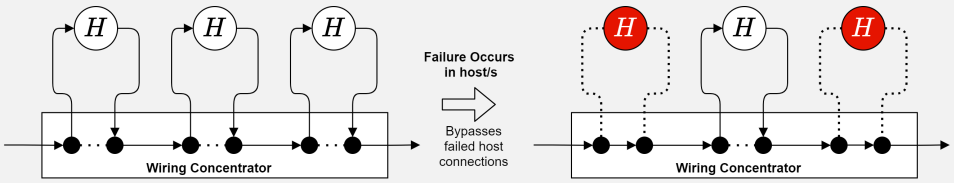
Ring Maintenance Complexity

- The **Frame Control** filed is used to create control frames.
- Frames may become orphaned (e.g never received, end up looping around indefinitely).
- One host is the **Active Monitor** and is responsible for generating tokens and removing orphaned frames.
- **Active Monitor** may fail, so any host must be able to become the **Active Monitor**.
- Contention rules/protocol needed to determine which host becomes **Active Monitor**.

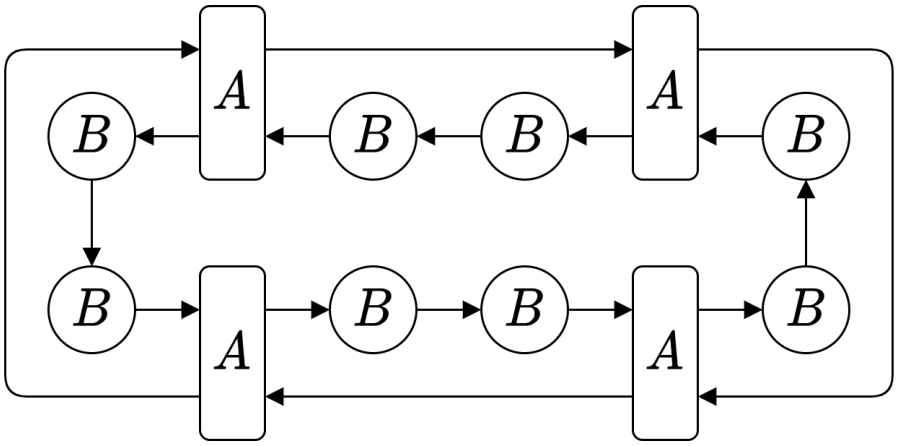
The above points add significant complexity to token passing and reduce reliability.

Definition: Wiring Concentrators

In order to improve reliability, when a host fails it can be bypassed.



Fibre Distributed Data Interface (FDDI)

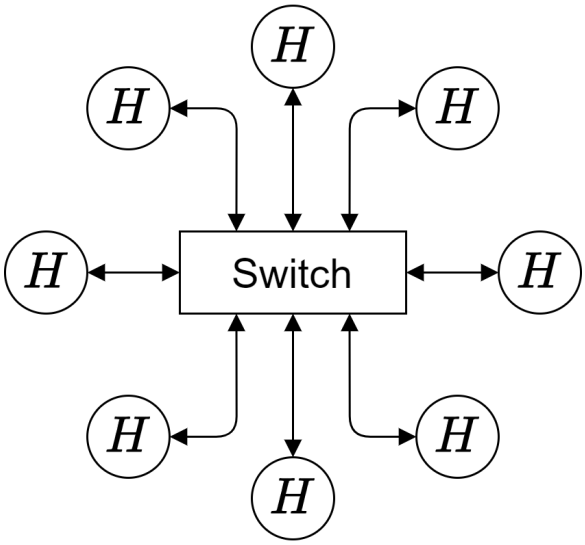


A ring-based token passing topology that was popular in the 1990s.

Hosts are divided into two classes, with one class (in our diagram class A) connected to both rings.

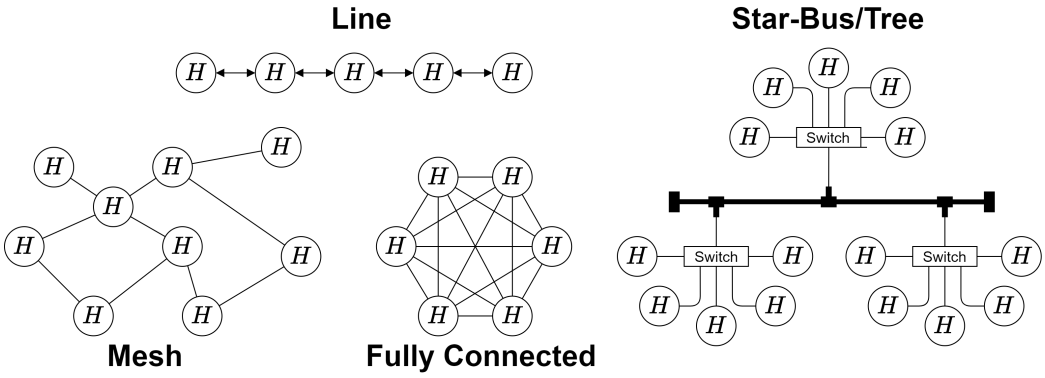
- Optical fibre cabling allows for networks to be geographically large.
- When a class B fails, data can be re-routed through class A hosts and the second ring.
- When a class A fails we can short-circuit two class As to create a single-ring (connecting two rings at disconnected ends).
- Rings can be up to $100km$ long, so FDDI must work with a length up to $200km$.
- No longer a popular.

Star



- All hosts connected directly to a **switch/multiport-bridge**.
- Any host can communicate with any other (provided they have a mechanism to prevent them talking over eachother).
- The central **switch** is a single point of failure (entire network fails it it fails).

Other Topologies

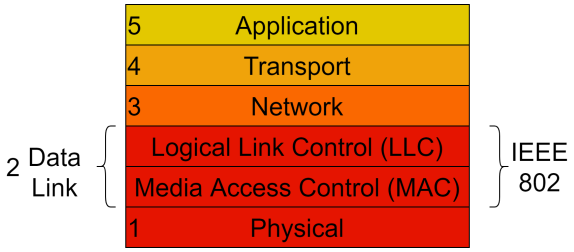


- Line is terrible (Split Ring).
- Star-Bus/Tree is a hybrid.
- Mesh is useful in some scenarios, but can be expensive.
- Mesh is ultra-fast (every host connected directly to every other host), but very expensive & difficult/nearly-impossible to effectively manage.

MAC

Lecture Recording

Lecture recording is available here



In **token ring** only one host can transmit at a time. However with other network types this is not the case.

For example a **broadcast** channel can have multiple receiving hosts & multiple concurrent transmissions can result in **frame** collisions.

We use **Medium Access Control** to coordinate channel access.

Stations

In these notes the term **station** is used to describe a host transmitting on the shared medium.

- In a wired network, frame collisions result in both frames needing to be re-transmitted.
- In a wireless network one transmitter may be stronger/a frame may be recieved even with a collision.

Medium Access Control Strategies

No Control

- When a frame is not recieved, the station retransmits as it pleases.
- Fine if channel ultilisation is low.
- Inefficient when contention is high (lots of transmitting stations → constant collisions & attempted re-transmissions).

Round Robin

- Stations take turns to transmit.
- Used in **token-based MAC** systems (only the station with the token can transmit).

Reservations

- Stations obtain *channel reservations* prior to transmitting.
- Stations can only transmit for the time interval they have reserved.
- Requires a system to manage reservations.
- Used in **slotted** systems.

Static Channel Allocations

Where each station is allocated a fixed schedule of times it is allowed to transmit.

For a channel shared between n different stations:

- **Time Division Multiplexing (TDM)**

  Stations waits for its time slot to transmit. each station's transmission rate limited to $\frac{R}{n}$ where R = maximum channel rate.

- **Frequency Division Multiplexing**

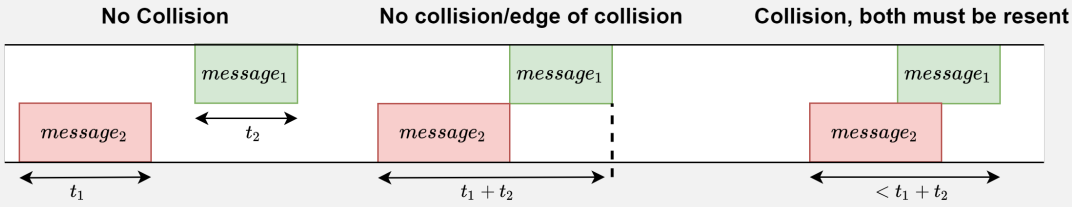
  Stations given a limited frequency band. Each station can use $\frac{B}{n}$ where B = total channel bandwidth.

  Bad for large n or traffic that is in bursts.

Dynamic Channel Allocation

Definition: ALOHA Protocol

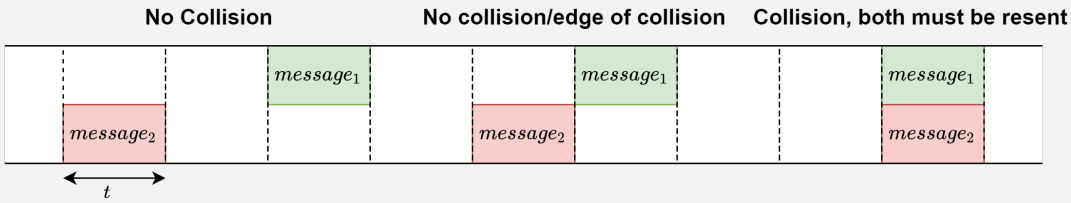
Stations transmit whenever they want to. If a collision occurs, stations wait a random period of time before attempting to re-transmit.



- The protocol suffers from low channel efficiency (worse with more contention) as there is a large **vulnerable period**.
- If a frame transmission is interrupted by another at any point, the both frames must be re-transmitted (new frames can destroy old frames).
- Maximum efficiency of 18% at 50% load.

Definition: Slotted ALOHA Protocol

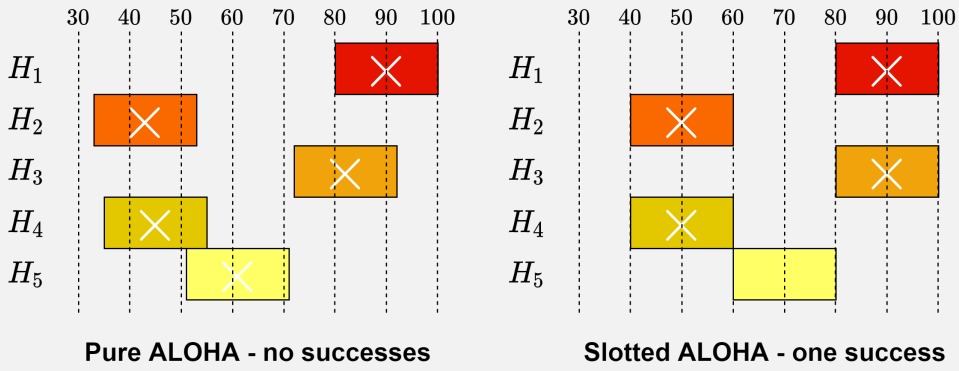
Like **ALOHA** but can only transmit on specific discrete time intervals (slots) (managed by a synchronous global clock).



- Reduces opportunities for a new frame to collide with an old one.
- Can only collide with exact overlap (contention for a slot)
- Maximum efficiency of 36% at 100% load.

Example: ALOHA Comparison

Given hosts start transmitting at times $H_1 : 80, H_2 : 33, H_3 : 72, H_4 : 35, H_5 : 51$ plot their transmissions and determine which transmissions collide given frames sent are $20s$ long.



Carrier Sense Multiple Access (CSMA)

Definition: Carrier Sensing

Listen before transmitting, transmission only occurs when the channel is idle/free.

- Reduces collisions over **ALOHA** as new frames are not sent during another's transmission.
- Collisions can still occur due to transmission delay (e.g two stations see idle channel, start transmitting, or one starts transmitting after another, but signal has yet to reach it).

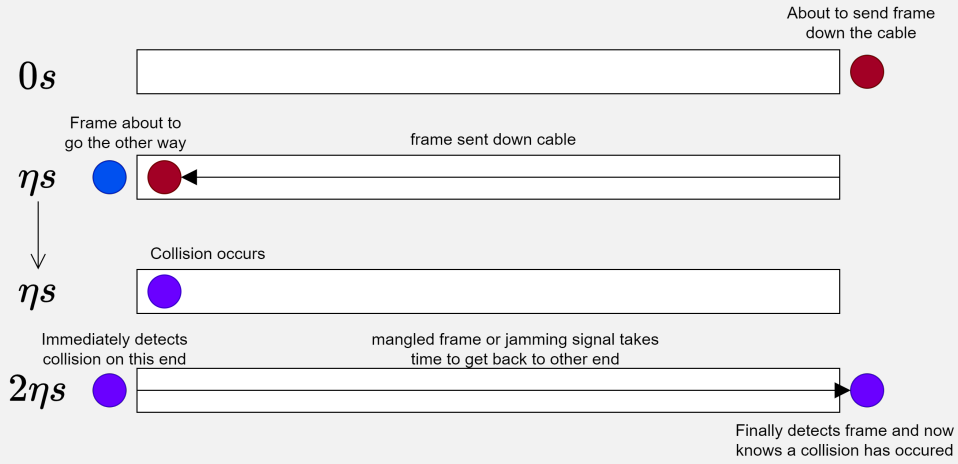
CSMA/CD	Collision Detection	Ethernet	IEEE 802.3
CSMA/CA	Collision Avoidance	WiFi	IEEE 802.11

Carrier Sense Multiple Access / Collision Detection (CSMA/CD)

Used for **ethernet** (wired networking), all collisions result in frames being destroyed.

- Station listens to channel during transmission to check for collisions.
- Transmission stop when collision detected, then sends a **jamming signal** (other transmitter will see the **jamming signal** and hence also know a collision has occurred).
- Host must transmit long enough to be able to tell the frame has not been collided. Hence minimum frame length is 2η where η = end-to-end transmission delay.

Required Delay



The minium size is required as we may not know the first bit (in the above example) has collided until $2\eta s$ after sending.

If the first bit does not collide (reaches end) then we have the line until we stop sending (as other side will not send while channel is not idle/free).

If the first bit does collide, then we should stop sending and attempt to recover (e.g resend using some policy/protocol).

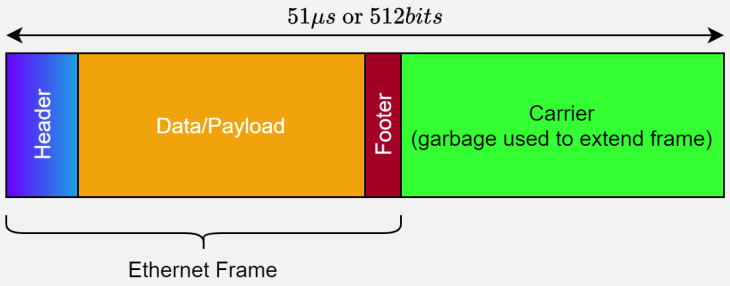
If the frame is not large enough, we will stop transmitting (and hence holding the channel) before we know if a collision may have occurred.

Collisions are inevitable as there is no central authority controlling transmission. Hence it is a best-effort service, in the worst case a frame may be indefinitely delayed.

- Suitable for most **LANs**
- Unacceptable for real-time systems (these require maximum wait time and minimum bandwidth assurances).

Definition: Carrier Extension

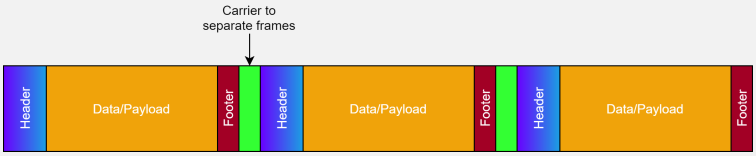
There is a minimum frame size requirement (to hold channel until bits reach destination), so for some frames an extension is required.



The wasted time transmitting the extension makes this inefficient.

Definition: Frame Bursting

Rather than padding with an extension, multiple frames are buffered, then packet together and sent at once.



Channel Back-Off

1-persistent	(Aggressive algorithm) Continually check channel. Transmit as soon as the channel is free. Used by **Ethernet**.
Non-persistent	(Non-Aggressive algorithm) Check channel, if idle translate immediately, else wait a random period of time before checking again.
P-persistent	Continually check channel, if it is free then transmit with probability p (between 1 and Non persistent).

Binary Exponential Back-Off

When network load is high (lots of contention for channels) **binary exponential back-off** is used.

- Slot length is the minimum frame length.
- If a collision occurs in transmission, wait 0 or 1 slots before attempting again.
- After c collisions, wait 0 to $2^c - 1$ slots (up to limit of 1023 slots / 10 collisions)

High contention → lots of collisions → **binary exponential back-off** → re-transmission attempts spread out → fewer collisions

Medium Access through Token Passing

There is a single token, stations can only transmit when they have the token.

- Token transfered with special token frame.
- If a station has the token but no frame to send, pass token on immediately.
- If a station has the token and a frame to send. It sets a timer and transmits until the timer expires or there is no more data to send. Then passes token.

Ethernet became much more popular, and hence has become the standard.

Avoiding Wired Collisions using Switches

A switch can remove the possibility of collisions by buffering frames and retransmitting when a channel becomes available.

- Each channel (ethernet cable) has only two stations (host and the switch)
- Hosts can transmit simultaneously, switch recieves and forwards frames.
- Maximum cable length determined by signal strength.
- **Switches** act as **repeaters**, refreshing the signal to pass it further.

Address Resolution Protocol (ARP)

Lecture Recording

Lecture recording is available here

In order to communicate outside of a network, an **IP Address** is required.

- Switches are in the **data-link layer** and do not use **IP Addresses** (in **network layer**)
- **IP Address** can be set statically (fixed **IP**, set manually) or dynamically (assigned to your **NIC**, e.g by **DHCP** service).
- **IP Addresses** specify hosts on the **internet**, it does not have to be translated when passing through a router (but can, e.g **NAT**).
- **MAC Addresses** specify hosts communicating on the same network/subnetwork. Typically do not change when passing through routers (as packets).

In order to translate **IP Addresses** to **MAC Addresses** and vice-versa we use **ARP**.

ARP Communication

1. **Router** Ask all hosts if they have a given **IP Address**
   Places **ARP Message** query in a Data-link frame and broadcasts.
2. **Host** Checks if it has the requested address, if so sends a reply with its **MAC Address**
3. **Router** Recieves **ARP Message** with **MAC Address** and uses it.
   Will forward **IP Datagrams** (encapsulated in a **Data-Link Frame**).
   Usually also cache the **IP** $\rightarrow$ **MAC** translation

Some optimisations include caching recent **ARP Message** replies, or having all hosts broadcast their **IP** and **MAC Address** on boot/connection (as a network policy).

Source			Destination			Message
Host	**IP**	**MAC**	**Host**	**IP**	**MAC**	
H_2	146.0.4.127	$0:1:2:a3:32:5$	All	146.0.4.98	$ff:ff:ff:ff:ff:ff$	ARP Req
H_1	146.0.4.98	$0:1:2:d5:6b:58$	H_2	146.0.4.127	$0:1:2:a3:32:5$	ARP Resp
H_2	146.0.4.127	$0:1:2:a3:32:5$	H_1	146.0.4.98	$0:1:2:d5:6b:58$	ICMP Echo Req
H_1	146.0.4.98	$0:1:2:d5:6b:58$	H_2	146.0.4.127	$0:1:2:a3:32:5$	ICMP Echo Resp
H_2	146.0.4.127	$0:1:2:a3:32:5$	H_1	146.0.4.98	$0:1:2:d5:6b:58$	ICMP Echo Req
H_1	146.0.4.98	$0:1:2:d5:6b:58$	H_2	146.0.4.127	$0:1:2:a3:32:5$	ICMP Echo Resp

```
1 arp who-has 146.169.4.98 tell 146.169.4.127 (0:90:27:a3:32:5)
2 arp reply 146.169.4.98 is-at 0:c0:4f:d5:6b:58 (0:90:27:a3:32:5)
3 146.169.4.127 > 146.169.4.98: icmp: echo request
4 arp who-has 146.169.4.127 tell 146.169.4.98 (0:c0:4f:d5:6b:58)
5 arp reply 146.169.4.127 is-at 0:90:27:a3:32:5 (0:c0:4f:d5:6b:58)
6 146.169.4.98 > 146.169.4.127: icmp: echo reply
7 146.169.4.127 > 146.169.4.98: icmp: echo request
8 146.169.4.98 > 146.169.4.127: icmp: echo reply
```

ARP cache poisoning

Malicious users can send spoof **ARP Messages** to attempt to associate their **MAC Address** with a victim's **IP Address** (thus receiving their **IP Datagrams**)

This is covered in detail on the wikipedia page.