Chapter 6 – Ethernet Fundamentals

Reiner Nitsch

FB Informatik

Office: F16/15c

Phone: 9255

Email: r.nitsch@fbi.fh-darmstadt.de

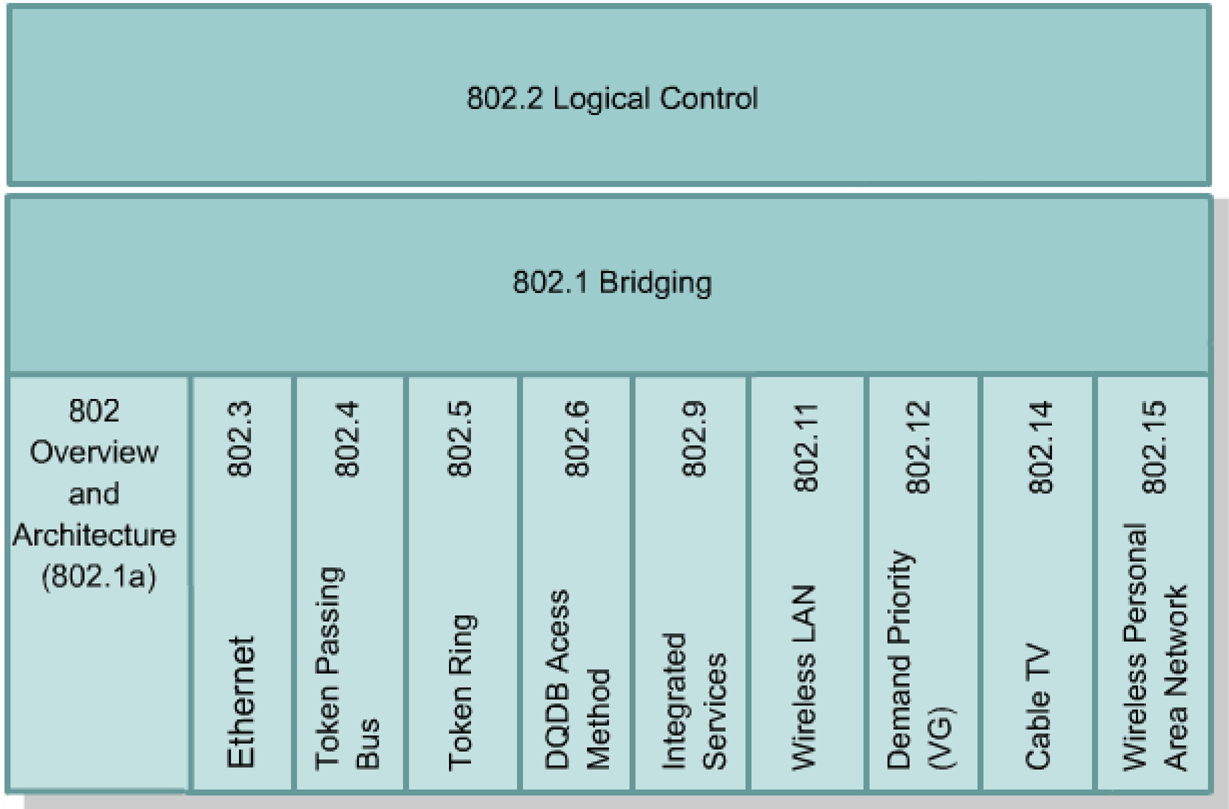
Introduction to Ethernet

- Ethernet developed in the 1970s
- success of Ethernet is due to the following factors:
  - Simplicity and ease of maintenance
  - Ability to incorporate new technologies
  - Reliability
  - Low cost of installation and upgrade

History

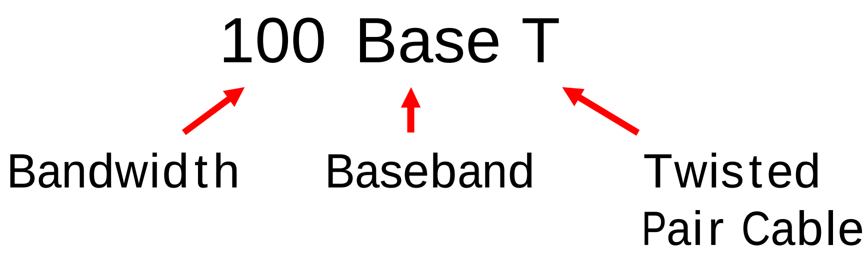
- Early 1970s: At the University of Hawaii a system called Alohanet was developed to control access of various stations to the shared radio frequency band in the atmosphere. This work formed the basis for the Ethernet access method known as CSMA/CD.
- 1980: First Ethernet standard; published by Digital Equipment Company, Intel, and Xerox (DIX); open standard; up to 10 Mbps; Thicknet; <= 2000 m
- 1985: **802.3 Ethernet** standard is published **by IEEE**; complies to ISO/OSI model; only small modifications to original (DIX)-Ethernet.
- Any Ethernet network interface card (NIC) can transmit and receive both Ethernet and 802.3 frames.
- 1995: IEEE announced a standard for a 100-Mbps Ethernet (100BASE-T, 100m).
- 1998 and 1999: IEEE standards for Gigabit Ethernet (1000BASE-T, 100m).
- All the standards are essentially compatible with the original Ethernet standard.
- Many Ethernet standard supplements were added in order to use different transmission media and higher transmission rates.

IEEE 802.X Standards

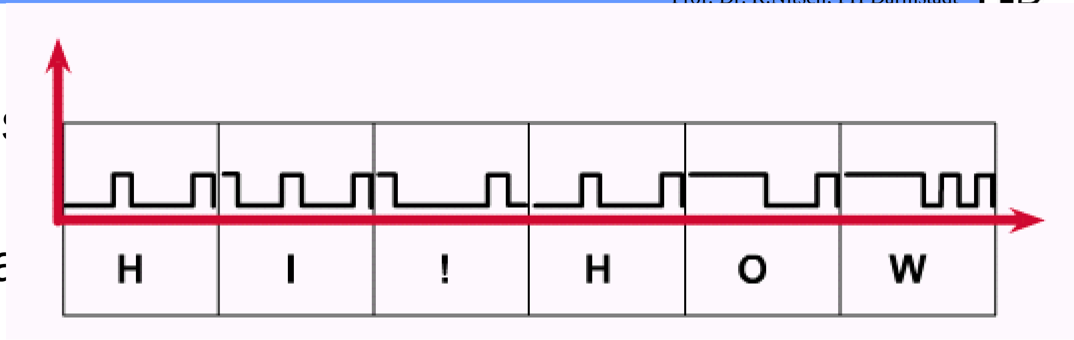


IEEE Ethernet naming rules

- Ethernet is a family of networking technologies:
  - Legacy Ethernet,
  - Fast Ethernet, and
  - Gigabit Ethernet.
- Ethernet speeds can be 10, 100, 1000, or 10,000 Mbps.
- Basic frame format remain consistent across all forms of Ethernet.
- Ethernet family relies on baseband signaling
- The abbreviated technology description consists of:
  - A number indicating the number of Mbps transmitted.
  - The word base, indicating that baseband signaling is used.
  - One or more letters of the alphabet indicating the type of medium used (F= fiber optical cable, T = copper unshielded twisted pair).



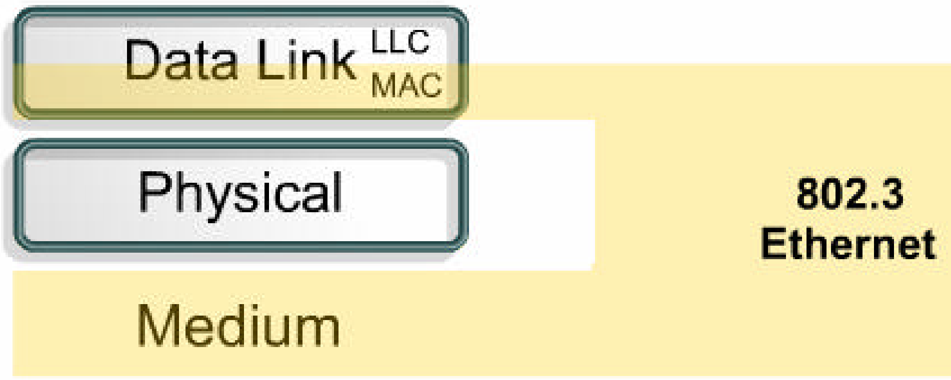
- Layer 1 involves
  - media, signals, bit stream that travel on media,
  - components that put signa on media, and
  - various topologies.

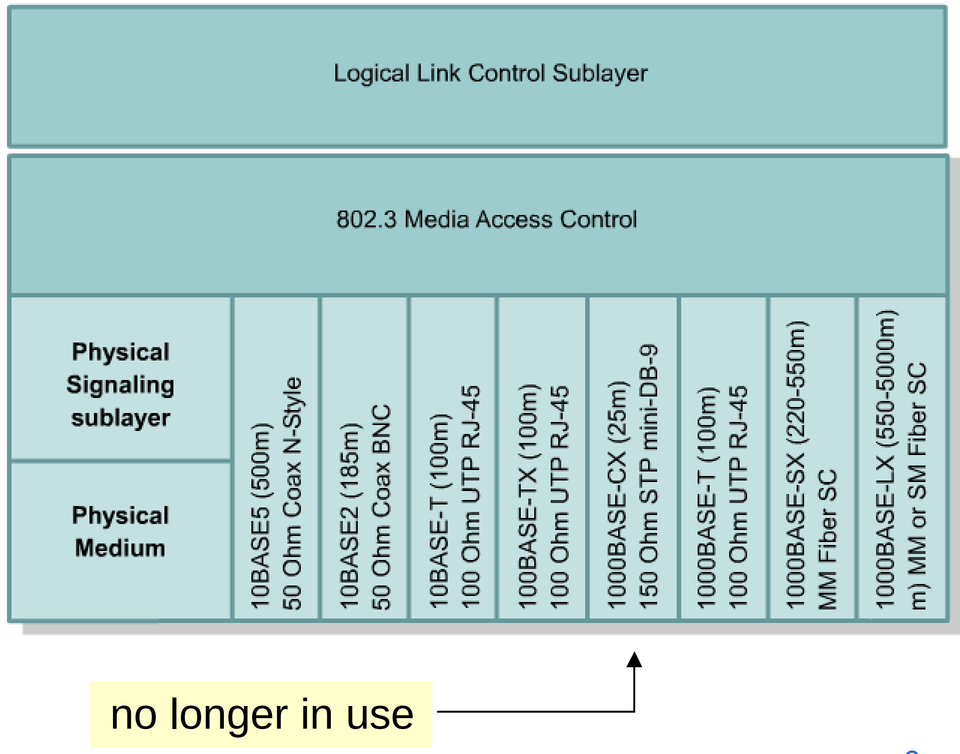


- Layer 2 tasks are
  - communicate with the upper-level layers using ***Logical Link Control (LLC)*** Sublayer
  - identify computers using a Layer 2 ***addressing*** (or naming) process.
  - organize or group the bits using a Layer 2 ***frame format*** to.
  - decide which computer will transmit binary data from a group that are all trying to transmit at the same time using a system called ***Media Access Control (MAC).***

Layer-2 (Data Link) Sublayers

- The *IEEE* divides the OSI data link layer into two separate sublayers.
  - Media Access Control (MAC) (transitions down to media)
  - Logical Link Control (LLC) (transitions up to the network layer)
- The MAC sublayer is concerned with the physical components that will be used to communicate the information.
- The LLC sublayer remains relatively independent of the physical equipment
- The LLC, as a sublayer, participates in the encapsulation process.





10Base5 (Thick-Ethernet, Yellow-Cable)

•Ethernet mit einer Bandbreite von 10 Mb über Yellow-Cable.
•Die maximale Kabellänge eines Segments beträgt 500 Meter.
•Die beiden Kabelenden müssen mit Abschlußwiderständen abgeschlossen werden.
•Pro Segment dürfen 100 Endgeräte angeschlossen werden.
•Der Abstand zwischen zwei Stichleitungen muß ein Vielfaches von 2,5m betragen.
•Die Stichleitungen dürfen nicht länger als 50 Meter sein.

10Base2 (Thin-Ethernet, Cheapernet)

•Ethernet über RG58 50O Koax-Kabel.
•Nur für 10Mb.
•BNC-Stecker und T-Stücke zur Verbindung.
•An beiden Enden ein 50O Abschlußwiderstand.
•Maximale Segmentlänge: 185m.
•Mindest Stationsabstand: 0,5m.
•Maximale Anzahl Stationen: 30
•Keine Stichleitungen erlaubt.

10Base-T

•Übertragung von 10Mb/s über Twisted Pair Kabel
•Hub notwendig.
•Maximale Entfernung 100m.
•Es werden zwei Adernpaare benötigt.
•Kabel ab Kategorie drei.
•Mit Switch Full-Duplex-Betrieb möglich.
•Verbindung von 2 Stationen ohne Hub mit Cros-Over-Kabel möglich.

100Base-T

•Übertragung von 100Mb/s über Twisted Pair Kabel
•Hub notwendig.
•Maximale Entfernung 100m.
•Es werden zwei Adernpaare benötigt.
•Kabel ab Kategorie fünf.
•Mit Switch Full-Duplex-Betrieb möglich.
•Verbindung von 2 Stationen ohne Hub mit Cros-Over-Kabel möglich.

10BaseFB

- 10Base-FB (B für Backbone)
  - Übertragung von 10Mb/s über LWL
  - Maximal 15 Repeater
  - Maximale Segmentlänge 2000m
  - Signalisierung ist synchron mit Fehlererkennung

10BaseFL

- 10Base-FL
  - Übertragung von 10Mb/s über LWL
  - Maximal 5 Repeater
  - Maximale Segmentlänge 2000m.

100BaseFX

- Übertragung von 100Mb/s über LWL
- Maximale Segmentlänge von 400m.
- Max. Distanz zwischen zwei 100BaseFX Switches 2000m

100BaseT4

- Übertragung von 100Mb/s über Cat 3 Kabel
- Es werden alle 4 Adernpaare benutzt.

1000BaseLX

- Übertragung von 1000Mb/s über LWL
- L für Long Wavelength von 1300nm
- Max. Distanz mit Multimodefasern 550m
- Max. Distanz mit Monomodefaser 3000m
- Full-Duplex

1000BaseSX

- Übertragung von 1000Mb/s über LWL
- S für Short Wavelength 850nm
- Max. Distanz mit 62,5µm Multimodefasern 270m
- Max. Distanz mit 50µm Multimodefasern 550m
- Full-Duplex

1000BaseCX

- C für Copper.
- Übertragung von 1000Mb/s über STP-Kabel 150O Wellenwiderstand
- Max. Distanz 25m.
- Full-Duplex

10GBaseLX4

- Übertragung von 10000Mb/s über LWL
- L für Long Wavelength
- X für WWDW (Wide Wave Division Multiplexing) 8B/10B Encoding
- 4 für die vier Bereiche:
    - 1269,0nm - 1282,4nm
    - 1293,5nm - 1306,9nm
    - 1318,0nm - 1331,4nm
    - 1342,5nm - 1355,9nm
- Max. Distanz mit 10µm Monomodefasern 10km
- Max. Distanz mit 50µm Multimodefasern 300m
- nur Full-Duplex

1000BaseT

- Übertragung von 1000Mb/s über UTP CAT 5 Kabel
- T für Twisted Pair.
- Max. Distanz 100m.
- Alle vier Paare sind notwendig.
- Full-Duplex

10GBaseSX4

- Übertragung von 10000Mb/s über LWL
- S für Short Wavelength
- X für CWDW (Coarse Wave Division Multiplexing)
- 4 für die vier Bereiche:
    - 773,5nm - 786,5nm
    - 789,5nm - 811,5nm
    - 823,5nm - 836,5nm
    - 848,5nm -861,5nm
- Max. Distanz mit 62µm Multimodefasern 100m
- Max. Distanz mit 50µm Multimodefasern 300m
- nur Full-Duplex, nur Multimodefasern kein Standard

10GBaseSR

- Übertragung von 10000Mb/s über LWL
- S für Short Wavelength von 850nm
- R für Serial 64B/66B Encoding
- Max. Distanz mit 50µm Multimodefasern 65m
- nur Full-Duplex, nur Fiber.

10GBaseSW

- Übertragung von 10000Mb/s über LWL
- S für Short Wavelength von 850nm
- W für Serial WIS (WAN Interface Sublayer) Encoding Ethernet in SONET STS192c
- Max. Distanz mit 50µm Multimodefasern 65m
- nur Full-Duplex, nur Fiber.

10GBaseLW

- Übertragung von 10000Mb/s über LWL
- L für Long Wavelength von 1310nm
- W für Serial WIS (WAN Interface Sublayer) Encoding Ethernet in SONET STS192c
- Max. Distanz mit 10µm Monomodefasern 10km
- nur Full-Duplex, nur Fiber.

10GBaseLR

- Übertragung von 10000Mb/s über LWL
- L für Long Wavelength von 1310nm
- R für Serial 64B/66B Encoding
- Max. Distanz mit 10µm Monomodefasern 10km
- nur Full-Duplex, nur Fiber.

10GBaseER

- Übertragung von 10000Mb/s über LWL
- E für Long Wavelength von 1550nm
- R für Serial 64B/66B Encoding
- Max. Distanz mit 10µm Monomodefasern 40km
- nur Full-Duplex, nur Fiber.

10GBaseEW

- Übertragung von 10000Mb/s über LWL
- E für Long Wavelength von 1550nm
- W für Serial WIS (WAN Interface Sublayer) Encoding Ethernet in SONET STS192c
- Max. Distanz mit 10µm Monomodefasern 40km
- nur Full-Duplex, nur Fiber.

10GBaseSR

•Übertragung von 10000Mb/s über LWL
•S für Short Wavelength von 850nm
•R für Serial 64B/66B Encoding
•Max. Distanz mit 50µm Multimodefasern 65m
•nur Full-Duplex, nur Fiber.

10GBaseSW

•Übertragung von 10000Mb/s über LWL
•S für Short Wavelength von 850nm
•W für Serial WIS (WAN Interface Sublayer) Encoding Ethernet in SONET STS192c
•Max. Distanz mit 50µm Multimodefasern 65m
•nur Full-Duplex, nur Fiber.

10GBaseLW

•Übertragung von 10000Mb/s über LWL
•L für Long Wavelength von 1310nm
•W für Serial WIS (WAN Interface Sublayer) Encoding Ethernet in SONET STS192c
•Max. Distanz mit 10µm Monomodefasern 10km
•nur Full-Duplex, nur Fiber.

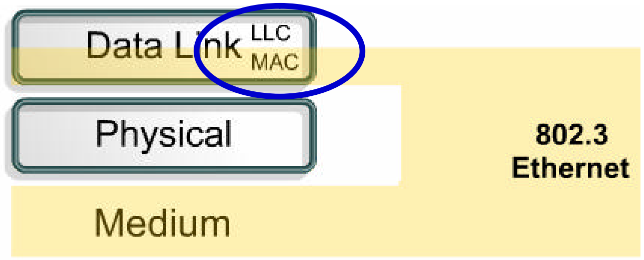
10GBaseLR

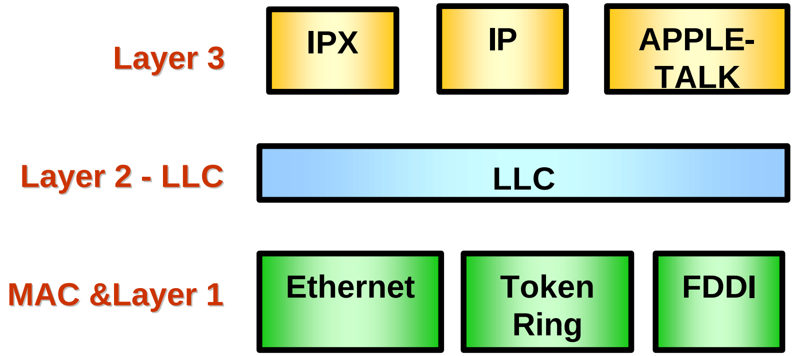
•Übertragung von 10000Mb/s über LWL
•L für Long Wavelength von 1310nm
•R für Serial 64B/66B Encoding
•Max. Distanz mit 10µm Monomodefasern 10km
•nur Full-Duplex, nur Fiber.

MAC – Media Access Control Sublayer

The Media Access Control (MAC) sublayer

- deals with the protocols that a host follows in order to access the physical media.
- is responsible for the actual framing
  - builds the 1s and 0s to hand off to the physical layer.
- is responsible for media access: (later)
  - Contention (ge: Auseinandersetzung)
  - Token Passing
  - Polling (Master sequentially asks all slaves wether they have to transmit data or not)





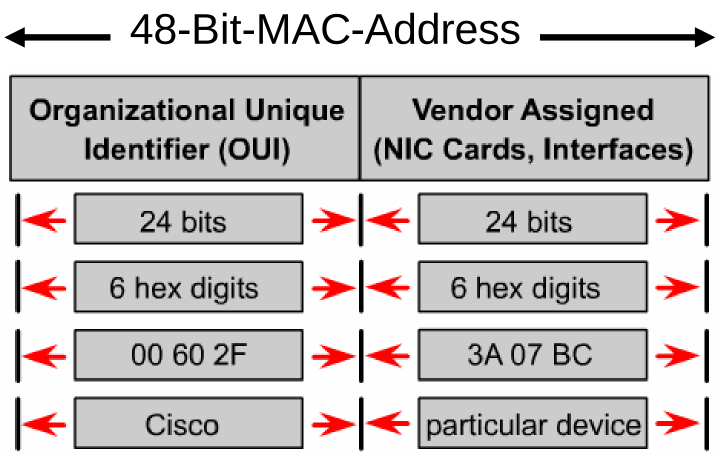
- is concerned with physical naming (addressing); defines MAC addressing;
- The NIC uses the MAC address to assess whether the message is destined for that host and therefore should be passed onto the upper layers of the OSI model.
- The NIC makes this assessment without using CPU processing time.

BTW: Ethernet vs IEEE 802.3

- Most of the time, the term “Ethernet” is used to mean IEEE 802.3
- For the most part, **Ethernet** and **IEEE 802.3** are used interchangeably, even though they are not really the same thing.
- We will discuss this more later.

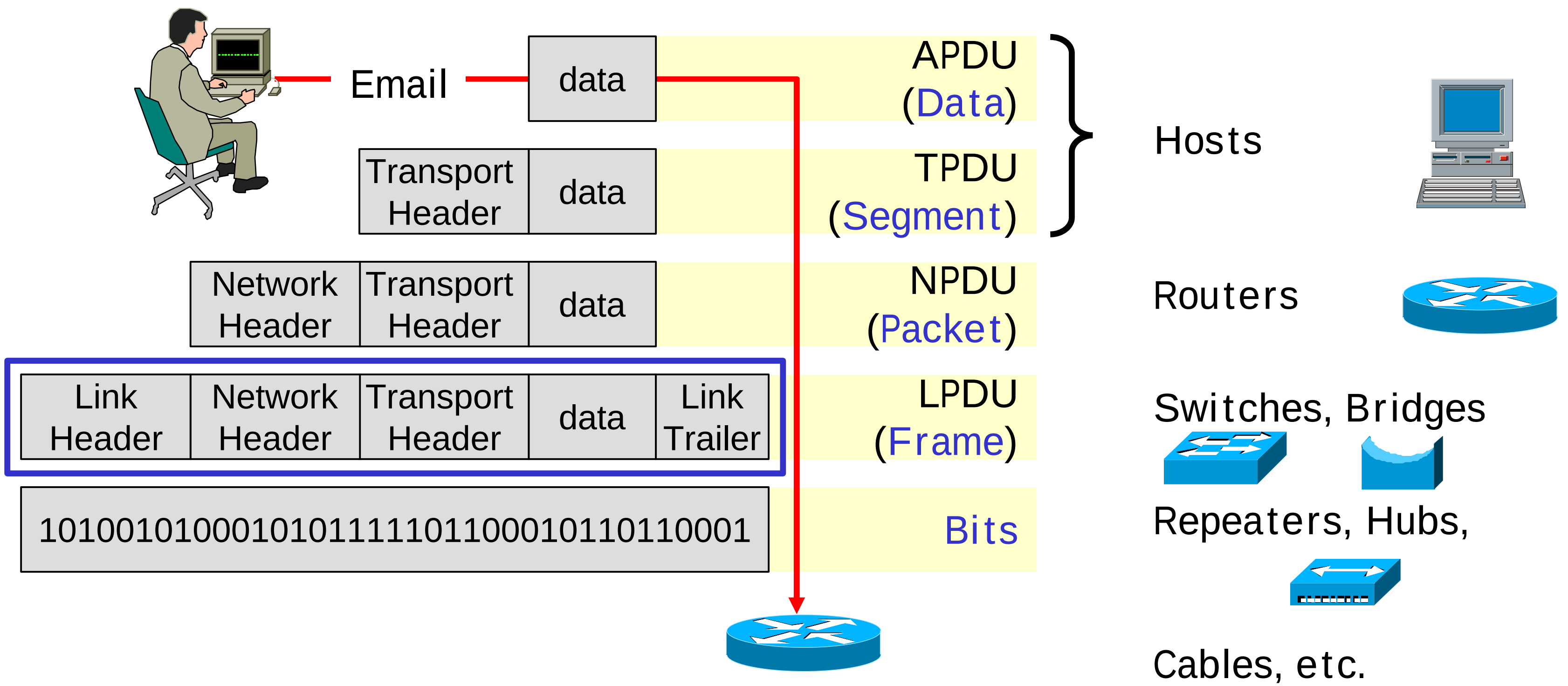
Naming Computers with MAC Addresses

- MAC addresses provide a way for computers to identify themselves within LANs
- MAC addresses are:
  - 48 bits in length
  - Expressed as twelve hexadecimal digits.
  - The first six hexadecimal digits, which are administered by the IEEE "universally administered address", identify the manufacturer or vendor and thus comprise the *Organizational Unique Identifier (OUI)*.
  - The remaining six hexadecimal digits comprise the *interface serial number*, or another value administered by the specific vendor.
- MAC addresses are sometimes referred to as *burned-in addresses (BIAs)* because they are burned into read-only memory (ROM) and are copied into random-access memory (RAM) when the NIC initializes
- The PC software (in PROTOCOL.INI or NET.CFG) can be configured to substitute a different address number. When this option is used, it is called a "locally administered address."



- IEEE OUI FAQs: http://standards.ieee.org/faqs/OUI.html

Review: Encapsulation Example



Let us focus on the Layer 2, Data Link, Ethernet Frame for now.

Generic Data Link Frame

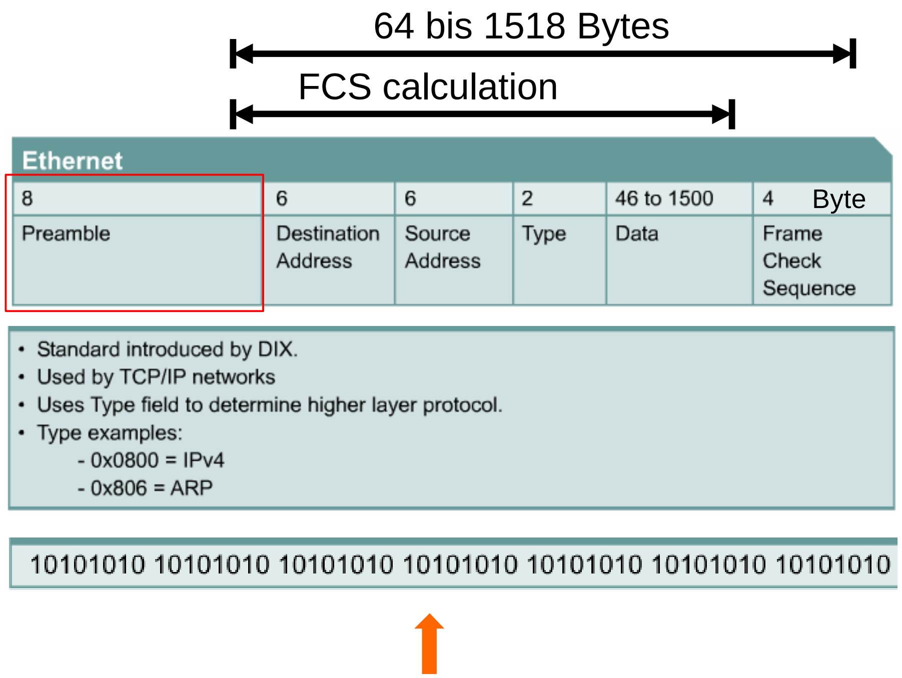
- Framing is the Layer 2 encapsulation process.
- A frame is the Layer 2 protocol data unit (2-PDU).
- Framing provides order, or structure, to the bitstream.
- There are many different types of frames described by various standards.
- A single generic frame has sections called fields
- Eeach field is composed of bytes.
- The names of the fields are as follows:
  - Start frame field
  - Address field
  - Length / type field
  - Data field
  - Frame check sequence field (FCS)

Field Names				
A	B	C	D	E
Start Frame Field	Address Field	Type/Length Field	Data Field	FCS Field

- Logical link control (LLC) bytes are also included with the data field in the IEEE standard frames.
- There are three common conventions for the format of the remainder of the frame:
  - IEEE 802.3 and 802.2
  - SNAP
  - Ethernet II or DIX

Ethernet II or DIX

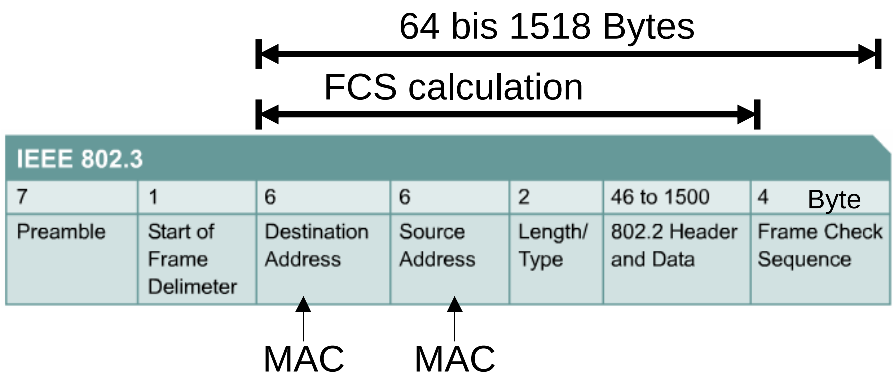
- Before the development of international standards, Xerox administered the Ethernet conventions.
- A two byte Type code was assigned by Xerox to identify higher-layer protocols developed by different vendors.
- The given out code values are
  - 0x0600 for XNS (the Xerox own protocol),
  - 0x6003 for DECNET,
  - 0x0800 for IP, and
  - 0x8137 for Novell IPX.
- **Problem**: To allow collision detection, the 10 megabit Ethernet requires a minimum packet size of 64 bytes. Any shorter message must be padded with zeros. Therefore each of these higher level protocols needs to have either a larger minimum message size (>=46 Bytes) or an internal length field that can be used to distinguish data from padding.



Preamble: 64 bits of timing synchronization information

IEEE 802.3 and 802.2

- The DIX standard did not need a length field because the vendor protocols that used it (XNS, DECNET, IPX, IP) all had their own length fields. However, the 802 committee needed a standard that did not depend on the good behavior of other programs. The 802.3 standard therefore replaced the two byte type field with a two byte length field.



- Maximum size of Ethernet frame payload (data) is 1500 and Xerox did not assign type values below 1500. This allows DIX and 802 standards to overlap conflictlessly.
- The 802.2 header follows the 802.3 header (and also follows the comparable fields in a Token Ring, FDDI, or other types of LAN).

Interpretation of Length/Type Field:

- Value >= 0x600 (hex): Type field (Ethernet II or DIX, Ethernet 802.3)
- Value < 0x600 (hex): Length field (Ethernet 802.3)

IEEE 802.2

- The 802.2 header is three bytes long for control packets or the kind of connectionless data sent by all the old DIX protocols (IP, IPX, DECNET, XNS).
- A four byte header is defined for connection oriented data, which refers primarily to SNA and NETBEUI.
- The first two bytes identify the SAP.
- In current use, the two SAP fields are set to 0x0404 for SNA and 0xF0F0 for NETBEUI.

6	6	2	1	1	1	N	4	Byte
DA	SA	Length/Type	DSAP	SSAP	Control	3-SDU	FCS	

DSAP: Destination Service Access Point
SSAP: Source Service Access Point
SDU: Service Data Unit = Higher Layer PDU

802.3 802.2

Subnetwork Access Protocol (SNAP)

- The IEEE left all the other protocols in a confusing situation: A one byte SAP could not substitute for the two byte type field.
- The compromise was to create a special version of the 802.2 header that conformed to the standard but actually repackaged the old DIX conventions.
- Under SNAP, the 802.2 header appears to be a datagram message (control field 0x03) between SAP ID 0xAA. The first five bytes of what 802.2 considers data are actually a subheader ending in the two byte DIX type value. Any of the old DIX protocols can convert their existing logic to legal 802 SNAP by simply moving the DIX type field back eight bytes from its original location.

6	6	2	1	1	1	3	2	N	4	Byte
DA	SA	Length/Type	DSAP=AA	SSAP=AA	Control=03	000000	Type	3-SDU	FCS	
Data Link Header			LLC Header			SNAP Header				

DSAP: Destination Service Access Point
SSAP: Source Service Access Point
SDU: Service Data Unit = Higher Layer PDU
LLC: Link Layer Control

802.2 802.3

Pause: Info

- Let's pause here for a moment and figure all of this out!

- Let's bring the following together:
  - Ethernet Frames and MAC Addresses
  - Sending and receiving Ethernet frames on a bus
  - CSMA/CD
  - Sending and receiving Ethernet frames via a hub
  - Sending and receiving Ethernet frames via a switch
  - 5-4-3 rule

Ethernet frame structure and MAC Addresses

- The frame structure is nearly identical for all speeds of Ethernet from 10 Mbps to 10,000 Mbps.
- Interpretation of Length/Type Field:
  - Value >= 0x600 (hex): Type field (Ethernet II or DIX, Ethernet 802.3)
  - Value < 0x600 (hex): Length field (Ethernet 802.3)
  - Ethernet requires that the frame be not less than 46 octets or more than 1500 octets.
- Padding bytes may be added so frames have a minimum length of 64 Bytes for timing purposes.
- 10 Mbps Ethernet is asynchronous: Each receiving station will use the preamble to synchronize the receive circuit to the incoming data, and then discard it.
- 100 Mbps and higher speed implementations of Ethernet are synchronous: The timing information is not required, however for compatibility reasons the Preamble and SFD are present.

Preamble: 64 bits of timing synchronization information

64 bis 1518 Bytes

FCS calculation

IEEE 802.3

7	1	6	6	2	46 to 1500	4 Byte
Preamble	Start of Frame Delimeter	Destination Address	Source Address	Length/ Type	802.2 Header and Data	Frame Check Sequence

MAC MAC

EEE 802.3 Ethernet frame fields

Octets	Description
• 7	Preamble
• 1	Start Frame Delimeter(SFD)
• 6	Destination MAC Address
• 6	Source MAC Address
• 2	Lenght/Type Field (Length if less than 0600 in hexadecimal,otherwise protocol Type)
• 46 to 1500	Data* (If less than 46 octets, then a pad must be added to the end)
• 4	Frame Check Sequence (CRC Checksum)

Length/Type field: usage as type field

Ethernet

8	6	6	2	46 to 1500	4 Byte
Preamble	Destination Address	Source Address	Type	Data	Frame Check Sequence

- Standard introduced by DIX.
- Used by TCP/IP networks
- Uses Type field to determine higher layer protocol.
- Type examples:
  - 0x0800 = IPv4
  - 0x806 = ARP

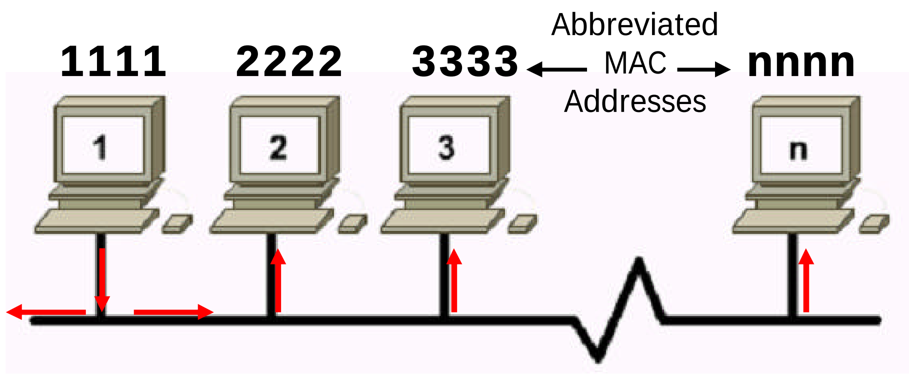
10101010 10101010 10101010 10101010 10101010 10101010 10101010

Sending and receiving Ethernet frames on a bus

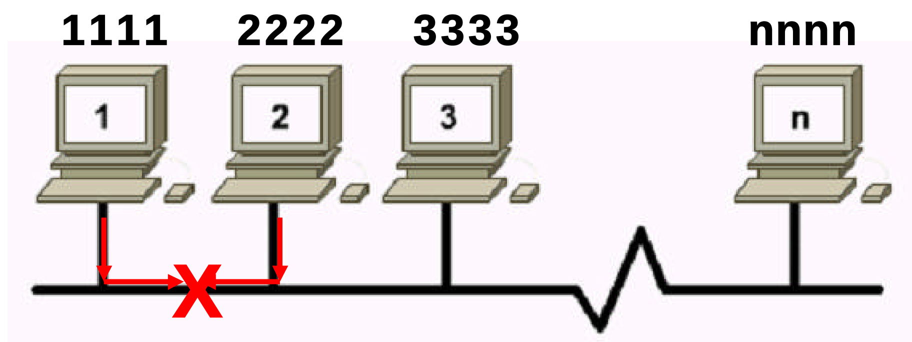
- When an Ethernet frame is sent out on the “bus” all devices on the bus receive it.

What do they do with it?

- Each NIC card compares its own MAC address with the Destination MAC Address.
- If it matches, it copies in the rest of the frame.
- If it does NOT match, it ignores (filters) the rest of the frame …
  - … unless you are running a Sniffer program
- So, what happens when multiple computers try to transmit at the same time?
- They produce a collision. Collisions are the most common error condition on an Ethernet.



Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	1111				



Collision!

Media Access Control (MAC)

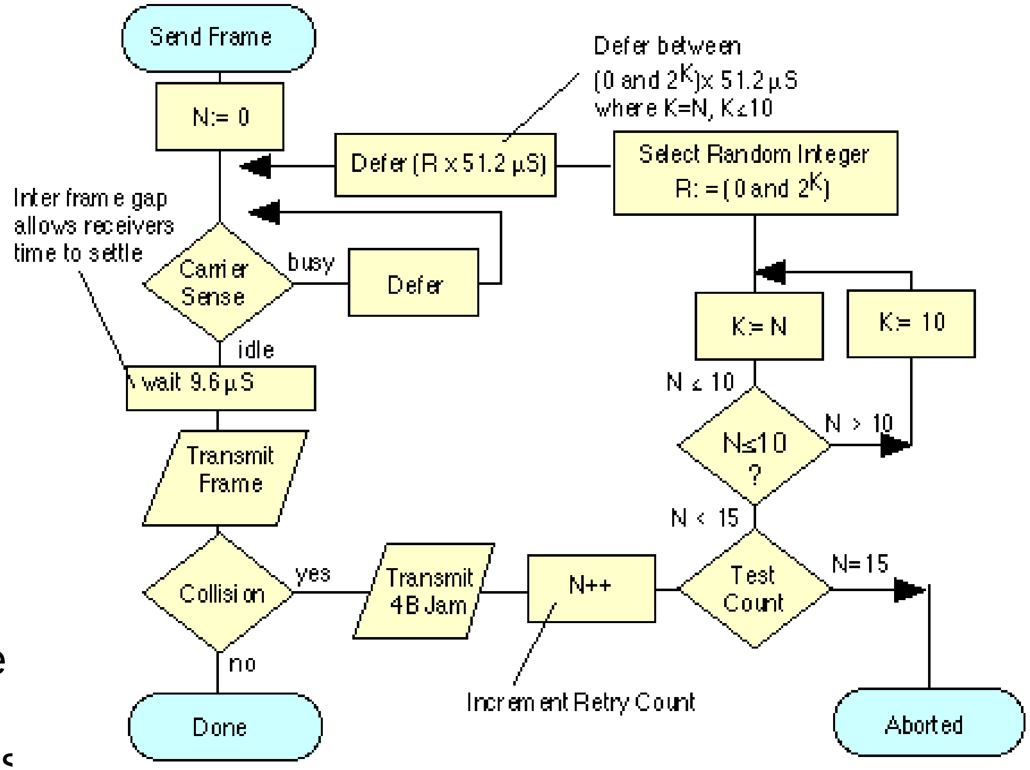
Two common types of access methods for LANs include

- **Non-Deterministic**: Contention (ge: Auseinandersetzung, Kampf) methods (Ethernet, IEEE 802.3)
  - use a first-come, first-served approach
  - CSMA/CD used by Ethernet is a simple system
  - Collisions are a normal occurrence on an Ethernet/802.3 LAN

- **Deterministic**: Token Passing (Token Ring, FDDI)
  - more later

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

- Ethernet is a shared-media broadcast technology.
- The access method CSMA/CD used in Ethernet performs three functions:
  - Transmitting and receiving data packets.
  - Decoding data packets and checking them for valid addresses before passing them to the upper layers of the OSI model.
  - Detecting errors within data packets or on the network.
- Collision detection method: Monitor the amplitude of the signal during transmission. If the amplitude increases on the networking media a collision has occurred.

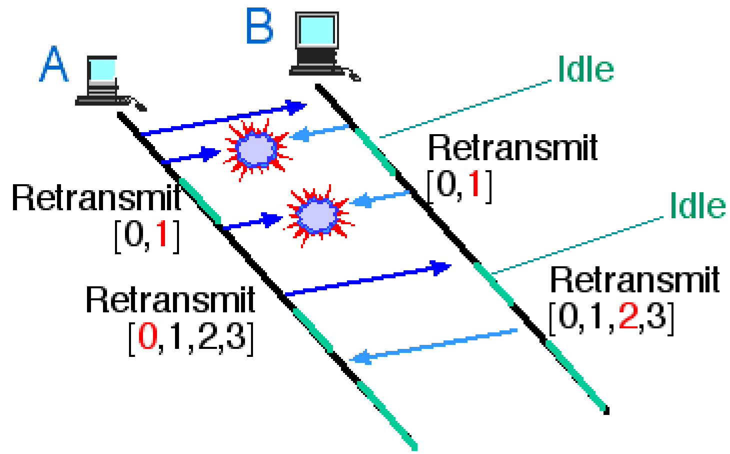


- In case of collision the nodes stop transmitting for a random period of time (backoff time), which is (hopefully) different for each device.
- When a device's backoff delay period expires, that device can attempt to gain access to the networking media.

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

- Listens to the network's shared media to see if any other users is "on the line" by trying to sense a neutral electrical signal or carrier.
- If no transmission is sensed, then *multiple access* allows anyone onto the media without any further permission required.
- If two NICs detect a neutral signal and access the shared media at the exact same time, a *collision* occurs and is *detected*.
- The PCs sense the collision by wether excess voltage is on the line additional to their own transmission voltage.
- When a collision occurs, a 32-bit jamming signal is sent out by the first NIC that detects the collision in order to enforce the collision .
- A random back-off scheme, is used to prevent colliding retransmissions.
- If collisions continue to occur, the NICs random interval is doubled, lessening the chances of a collision.
- **Late Collisions:** In a proper functioning Ethernet network, a NIC may experience collision within the first slot time (minimum frame period, 51.2 µS) after it starts transmission. This is the reason why an Ethernet NIC monitors the CD signal during this time and use CSMA/CD. A faulty CD circuit, or misbehaving NIC or transceiver may lead to a late collision (i.e. after one slot time).

Exponential Back-off Algorithm



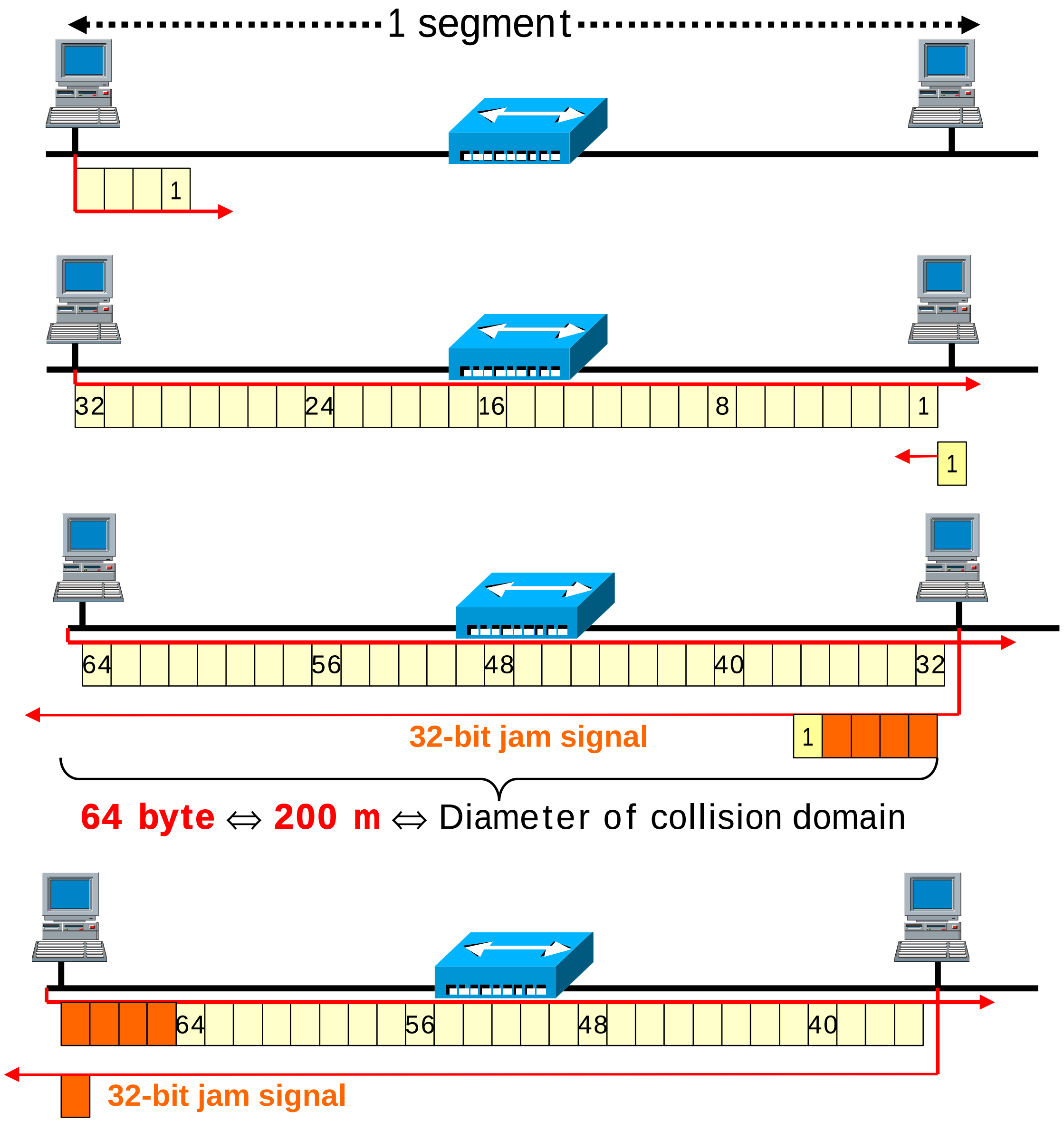
- If all NICs attempted to retransmit immediately following a collision, then this would certainly result in another collision.
- Ethernet uses a random back-off period to minimize the propability of this event.
- Each node selects a random number, multiplies this by the slot time (minimum frame period, 51.2 µS) and waits for this random period before attempting retransmission. The small Inter-Frame Gap (IFG) (e.g., 9.6 microseconds) is also added.
- On a busy network, a retransmission may still collide with another retransmission (or possibly new frames being sent for the first time by another NIC). The protocol therefore counts the number of retransmission attempts (using a variable N in the above figure) and attempts to retransmit the same frame up to 15 times.
- For each retransmission, the transmitter constructs a set of numbers:

  {0, 1, 2, 3, 4, 5, ... L} where L is ([2 to the power (K)]-1) and where K=N; K<= 10;
- A random value R is picked from this set, and the transmitter waits for a period

  R x (slot time) i.e. R x 51.2 µs
- For example, after two collisions, N=2, therefore K=2, and the set is {0, 1, 2, 3} giving a one in four chance of collision. This corresponds to a wait selected from {0, 51.2, 102.4, 153.6} micro seconds.

What are collision domains?

- Host A senses no voltage on the circuit. It begins transmitting a frame. The signal voltage travels through the wire with about light velocity. Repeaters and hubs flood the incoming signals out any other outputs
- Host B also wants to transmit a frame. It senses a free line up to the time when host A's first frame byte arrives at host B. Immediately after Host B starts sending it detects the excessive voltage from host A's frame. It stops sending and starts sending a 32-bit jam signal
- Host A cannot detect the collision until the collision fragments of Host B arrives. It **MUST** continue sending his frame until that time.
- When the collision fragments finally reaches Host A, it also truncates the current transmission and substitutes the jam signal.
- A hub or series of hubs is a single **collision domain**.



Slot Times required

- In order to have CSMA/CD work over a distance of 200 m the Ethernet standard requires a minimum frame transmission time called **slot time**.
- Slot time is calculated assuming maximum cable lengths on the largest legal network architecture. All hardware propagation delay times are at the legal maximum and the **32-bit jam signal** is used when collisions are detected.
  - Slot time for 10-Mbps Ethernet is 512 bit-times (=51.2µs), or 64 octets.
  - Slot time for 100-Mbps Ethernet is 512 bit-times (5.12µs), or 64 octets.
  - Slot time for 1000-Mbps Ethernet is 4096 bit-times (4.096µs), or 512 octets.
- To allow 1000-Mbps Ethernet to operate in **half duplex** the **extension field** was added to expand the frame length to 512 bytes for small data payloads.
- This field is present only on 1000-Mbps, half-duplex links and allows minimum-sized frames to be long enough to meet slot time requirements. Extension bits are discarded by the receiving station.
- Propagation velocity is about 20.3 cm (8 in) per nanosecond in a UTP cable.
- It takes just under 1 µs for the signal (10 bit-times for a 10BASE-T, 100 bit-times for 100BaseT, and 1000 bit-times for 1000BaseT) to travel 200 m length of UTP cable.
- Therefore half duplex is not permitted in 10-Gigabit Ethernet.

Interframe spacing and backoff

Interframe spacing

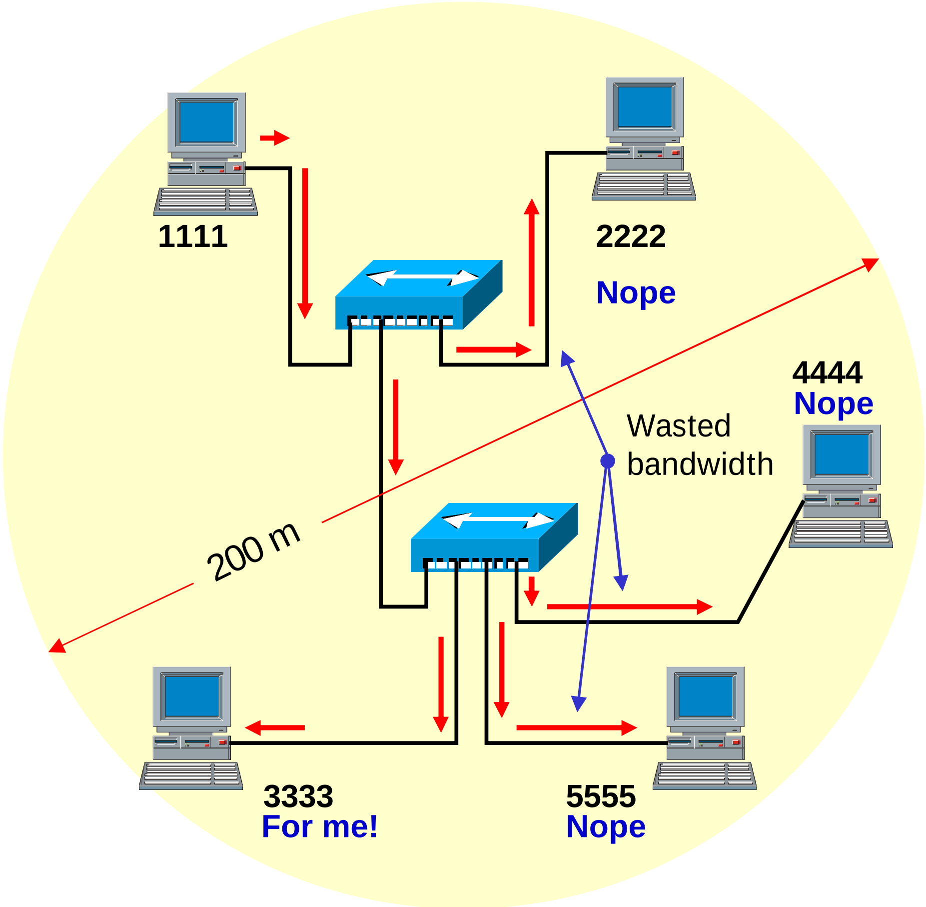
- After a frame has been sent, all stations on a 10-Mbps Ethernet are required to wait a minimum of 96 bit-times (9.6 microseconds) before any station may legally transmit the next frame.
- This minimum spacing between two non-colliding frames is also called the **interframe gap**. It is measured from the last bit of the FCS field of the first frame to the first bit of the preamble of the second frame.
- This gap is intended to allow slow stations time to process the previous frame and prepare for the next frame.

Backoff Time

- After a collision occurs and all stations wait the full interframe spacing
- The stations that collided must wait an additional backoff time before retransmitting the collided frames.
- The waiting period is intentionally random so that two stations do not delay for the same amount of time before retransmitting which would result in more collisions.
- The waiting period is measured in increments of the parameter "slot time".
- If the MAC layer is unable to send the frame after sixteen attempts, it gives up and generates an error to the network layer.

Sending and receiving Ethernet frames via a hub

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	1111				



- So, what does a hub do when it receives a frame?
- Remember, a hub is nothing more than a multiport repeater acting as a layer 1 device.
- The hub will **flood** it out all ports except for the incoming port.
- A hub does NOT look at layer 2 addresses, so it is fast in transmitting data.
- Disadvantage with hubs: A hub or series of hubs is a single **collision domain**.
- A collision will occur if any two or more devices transmit at the same time within the collision domain.
- More on this later.
- Another disadvantage with hubs is that it takes up unnecessary bandwidth on other links.

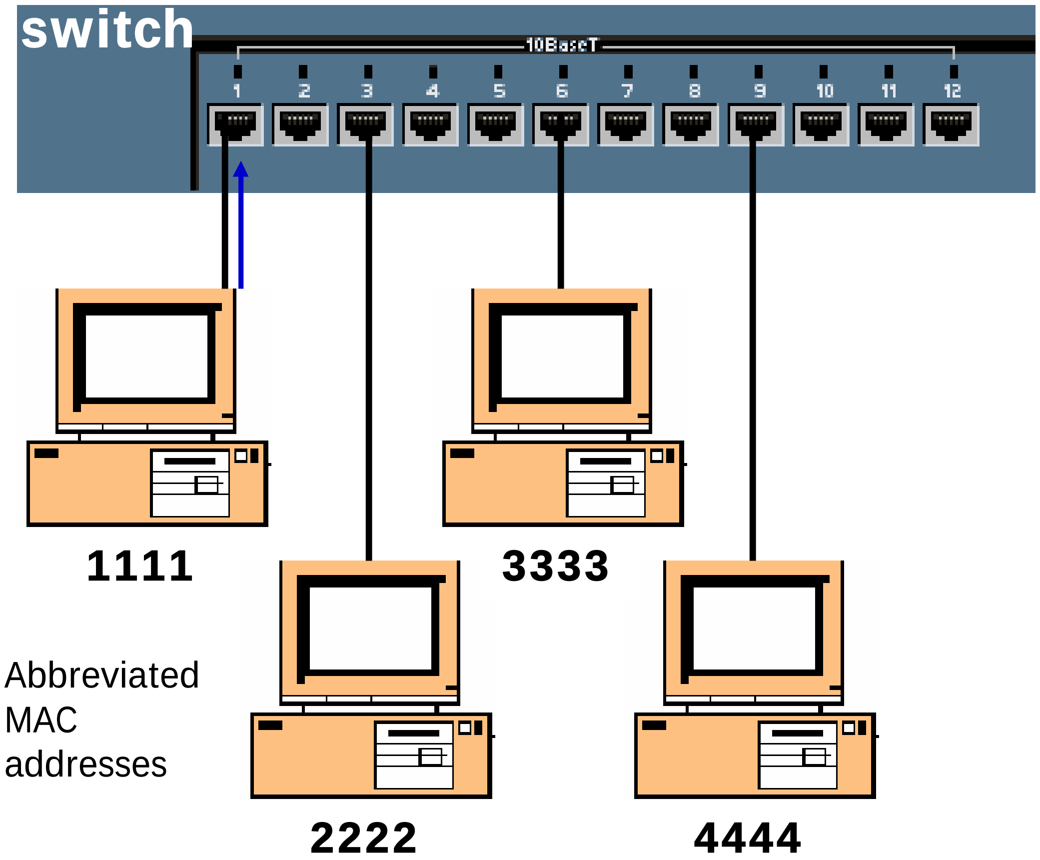
Nope: no operation

Sending and receiving Ethernet frames via a switch

Source Address Table

Port	Source MAC Add.	Port	Source MAC Add.

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	**1111**				



Abbreviated MAC addresses

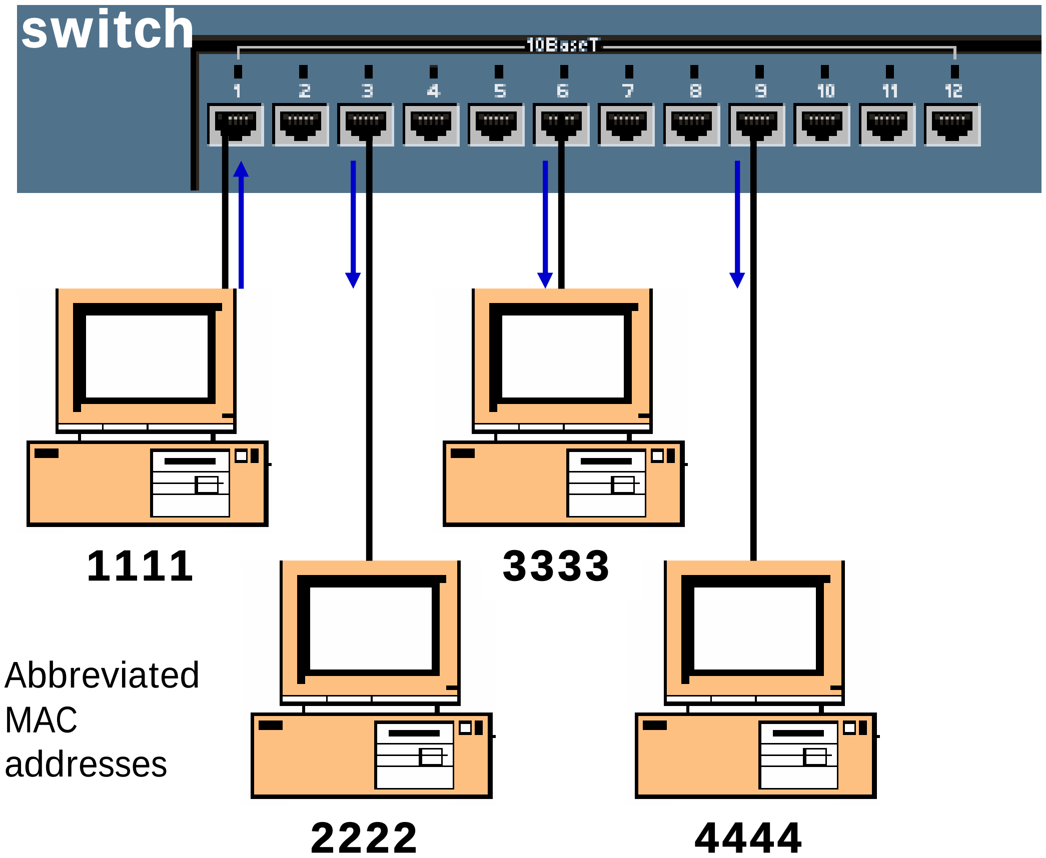
- Switches are also known as **learning bridges** or **learning switches**.
- A switch has a source address table in cache (RAM) where it stores source MAC address after it learns about them.
- A switch receives an Ethernet frame it searches the source address table for the Destination MAC address.
- If it finds a match, it **filters** the frame by only sending it out that port.
- If there is not a match if **floods** it out all ports.

No Destination Address in table, Flood

Source Address Table

Port	Source MAC Add.	Port	Source MAC Add.
1	**1111**		

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	**1111**				



- How does it learn source MAC addresses?
- First, the switch will see if the SA (1111) is in it's table.
- If it is, it resets the timer (more in a moment).
- If it is NOT in the table it adds it, with the port number.

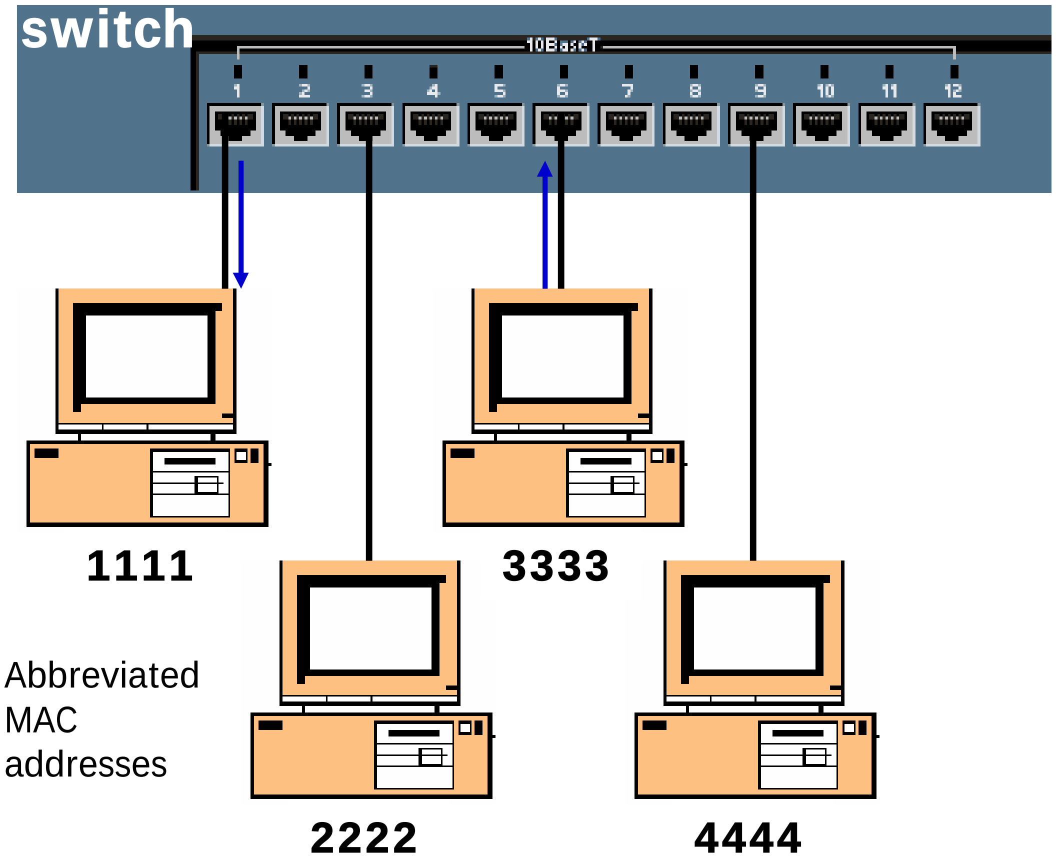
- Next, in our scenario, the switch will **flood** the frame out all other ports, because the DA is not in the source address table.

Destination Address in table, Filter

Source Address Table

Port	Source MAC Add.	Port	Source MAC Add.
1	**1111**	**6**	**3333**

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	1111	**3333**				



- Now 3333 sends data back to 1111.
- The switch sees if it has the SA stored.
- It does NOT so it adds it. (This will help next time 1111 sends to 3333.)
- Next, it checks the DA and in our case it can **filter** the frame, by sending it only out port 1.

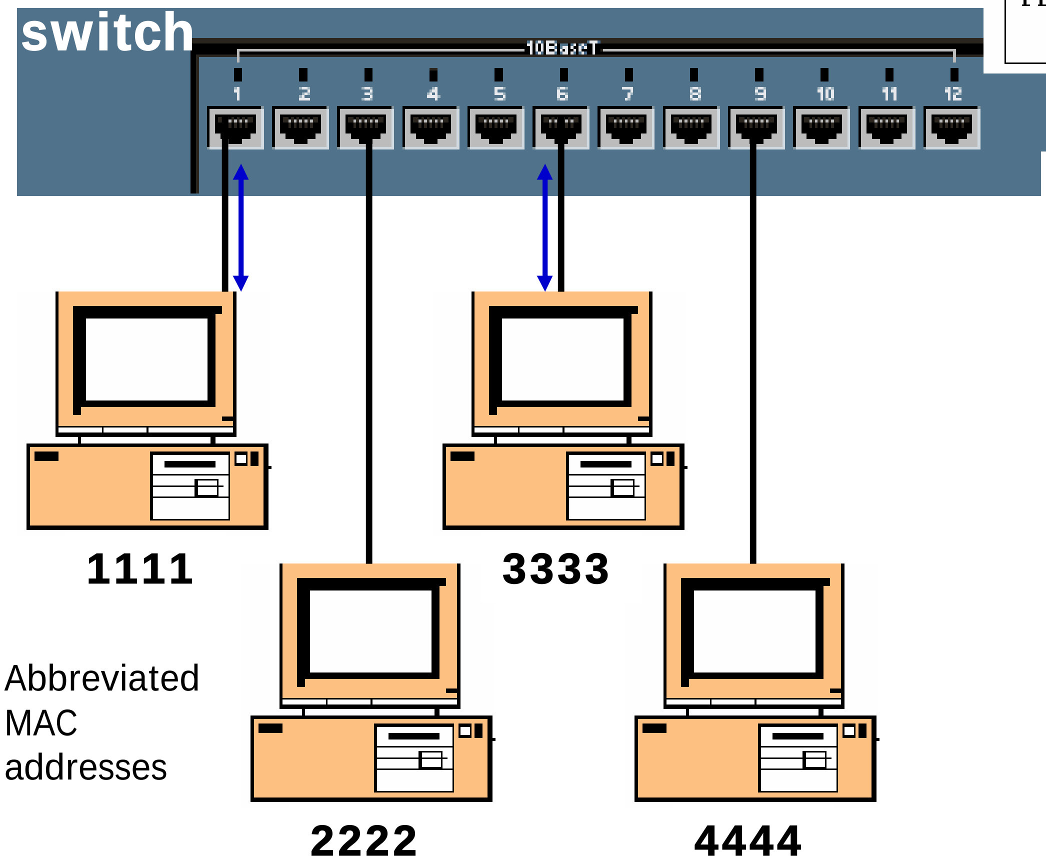
Destination Address in table, Filter

Source Address Table

Port	Source MAC Add.	Port	Source MAC Add.
1	**1111**	**6**	**3333**

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	**1111**				

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	1111	**3333**				



Abbreviated MAC addresses

- Now, because both MAC addresses are in the switch's table, any information exchanged between 1111 and 3333 can be sent (filtered) out the appropriate port.

- **What happens when two devices send to same destination?**
- **What if this was a hub?**
- **Where is (are) the collision domain(s) in this example?**

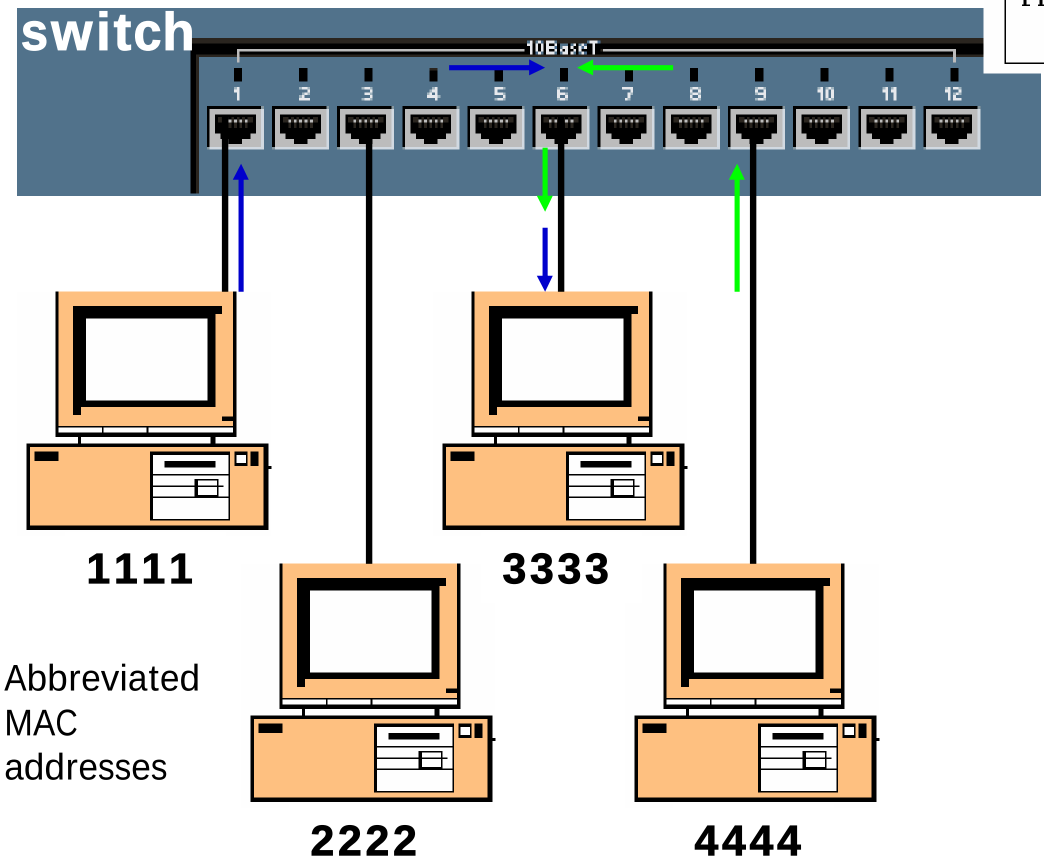
No Collisions in Switch, Buffering

Source Address Table

Port	Source MAC Add.	Port	Source MAC Add.
1	**1111**	**6**	**3333**
9	**4444**		

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	**1111**				

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	**4444**				



- Unlike a hub, a collision does NOT occur within these segments, which would cause the two PCs to have to retransmit the frames.
- Instead the switch buffers the frames and sends them out port #6 one at a time.
- The sending PCs have no idea that their was another PC wanting to send to the same destination.

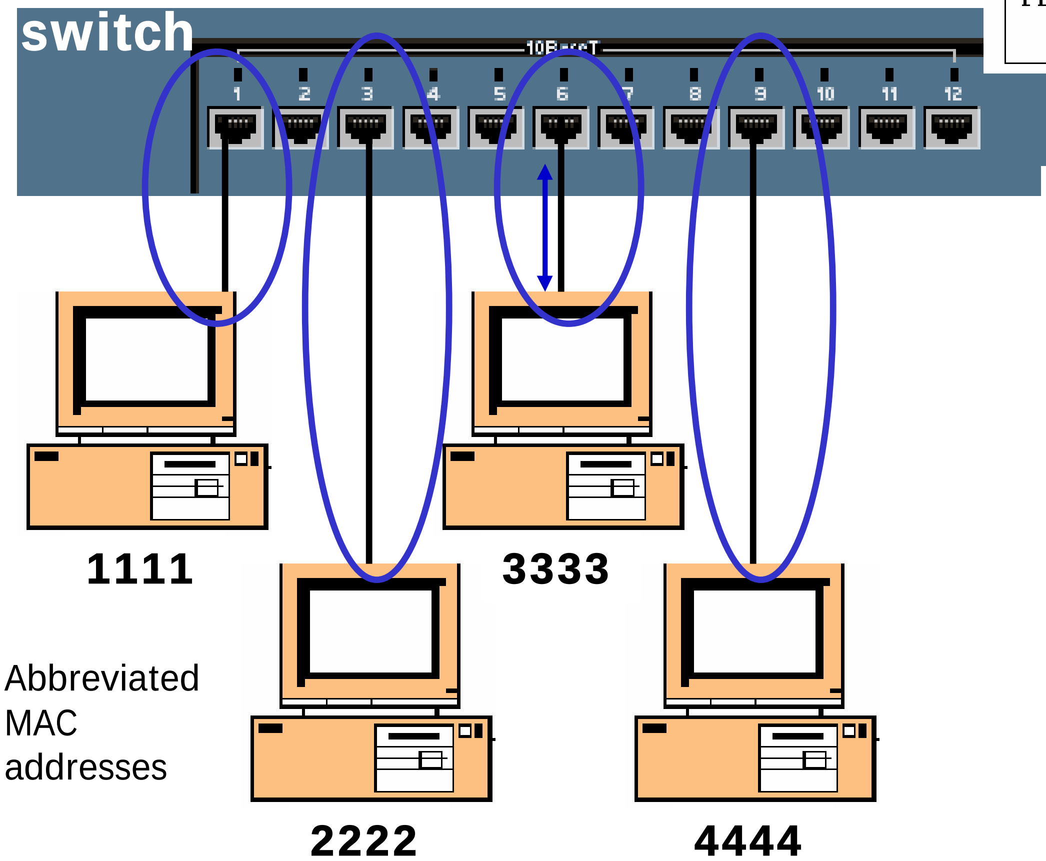
Source Address Table

Port	Source MAC Add.	Port	Source MAC Add.
1	1111	6	3333
9	4444		

Collision Domains

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	1111				

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	4444				



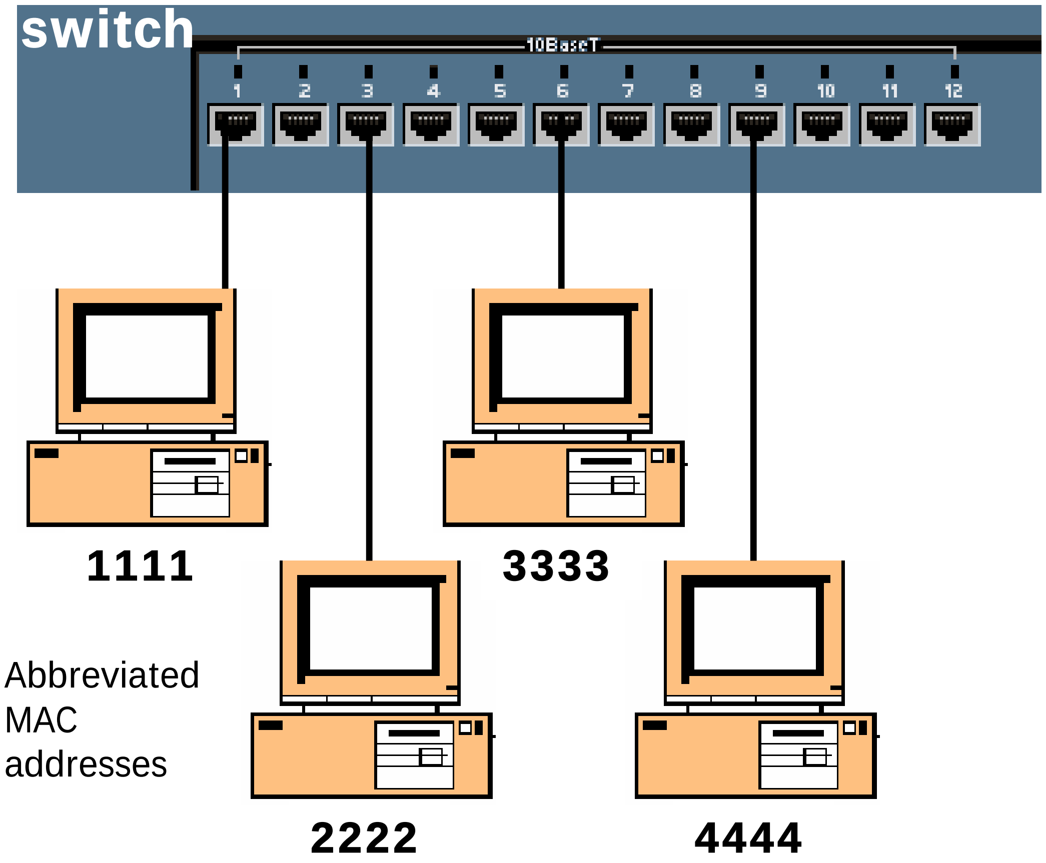
Abbreviated MAC addresses

- When there is only one device on a switch port, the collision domain is only between the PC and the switch. (Cisco curriculum is inaccurate on this point.)
- With a **full-duplex** PC and switch port, there will be no collision, since the devices and the medium can send and receive at the same time.
- Full-duplex operation allows for larger network architecture designs since the timing restriction for collision detection is removed.

Other Information

Source Address Table

Port	Source MAC Add.	Port	Source MAC Add.
1	**1111**	**6**	**3333**
9	**4444**		



Preamble	Destination Address	Source Address	Type	Data	Pad	CRC

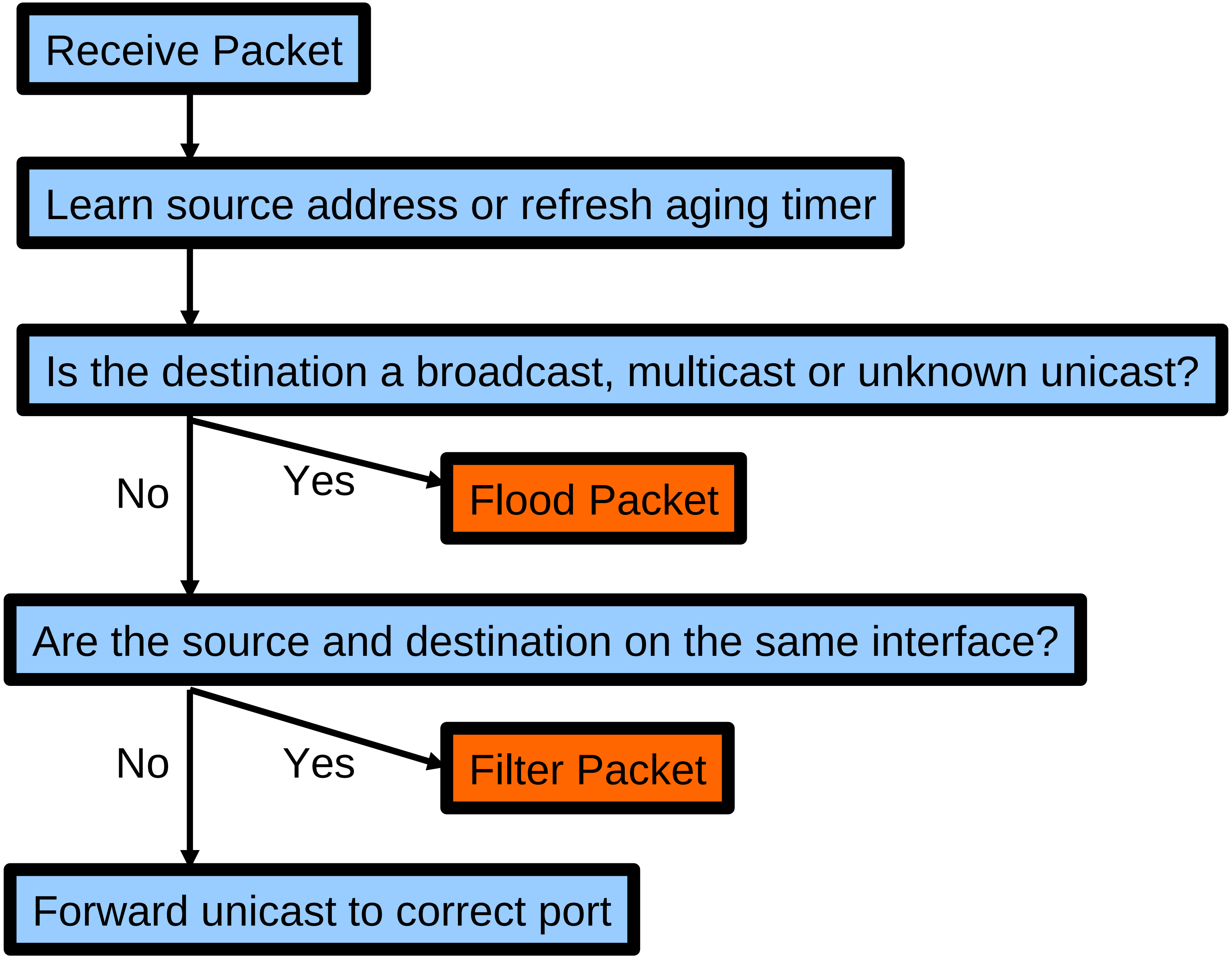
- How long are addresses kept in the Source Address Table?
  - 5 minutes is common on most vendor switches.
- How do computers know the Destination MAC address?
  - ARP Caches and ARP Requests (later)
- How many addresses can be kept in the table?
  - Depends on the size of the cache, but 1,024 addresses is common.
- What about Layer 2 broadcasts?
  - Layer 2 broadcasts (DA = all 1's) is flooded out all ports.

Side Note - Transparent Bridging

- Transparent bridging (normal switching process) is defined in IEEE 802.1d describing the five bridging processes of:
  - learning
  - flooding filtering
  - forwarding
  - aging

- These will be discussed further in STP (Spanning Tree Protocol)

Transparent Bridge Process - Jeff Doyle

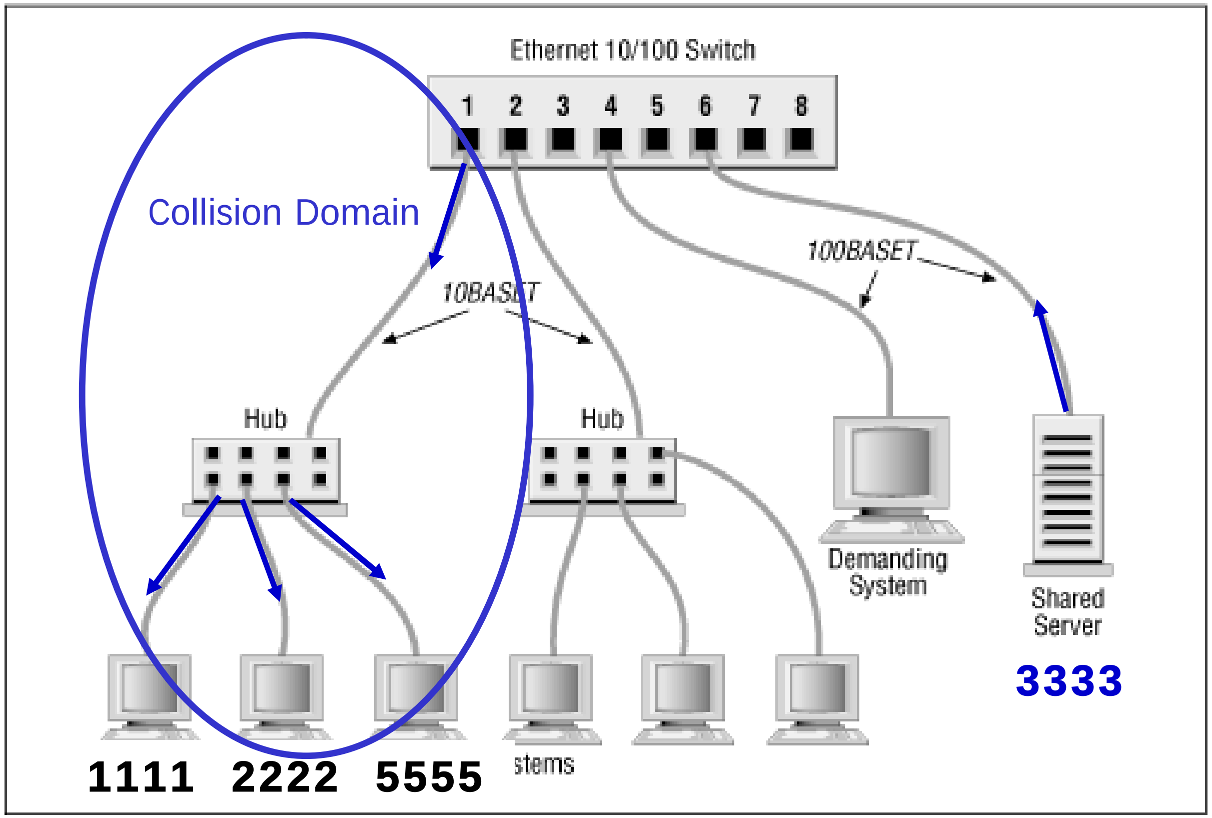


What happens here?

Source Address Table

Port	Source MAC Add.	Port	Source MAC Add.
1	**1111**	6	3333
1	**2222**	**1**	**5555**

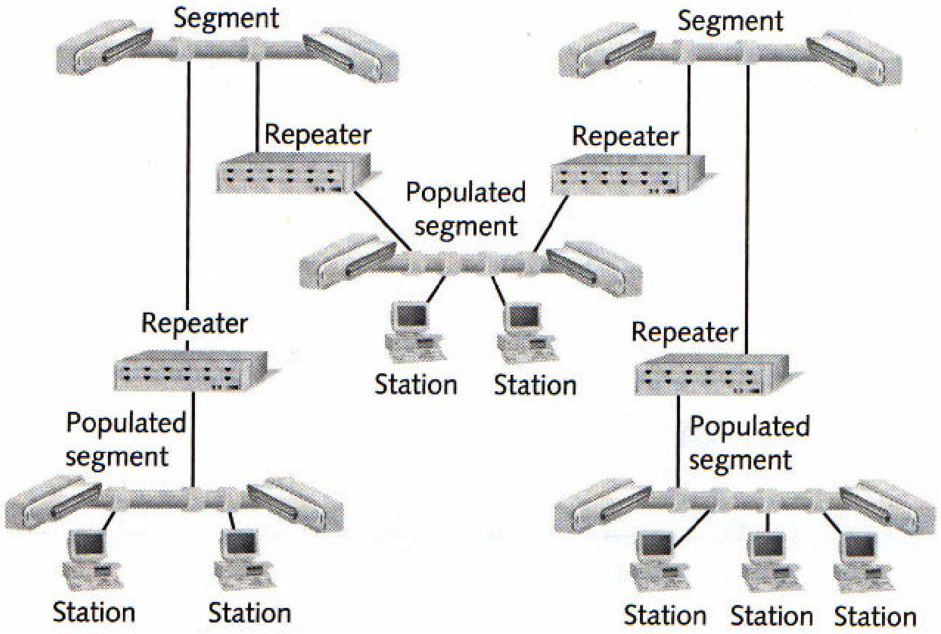
Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	1111	**3333**				



- Notice the Source Address Table has multiple entries for port #1.
- Having these small collision domains with a switch is called "**micro segmentation**"
- The switch filters the frame out port #1.
- But the hub is only a layer 1 device, so it floods it out all ports.
- Where is the collision domain?

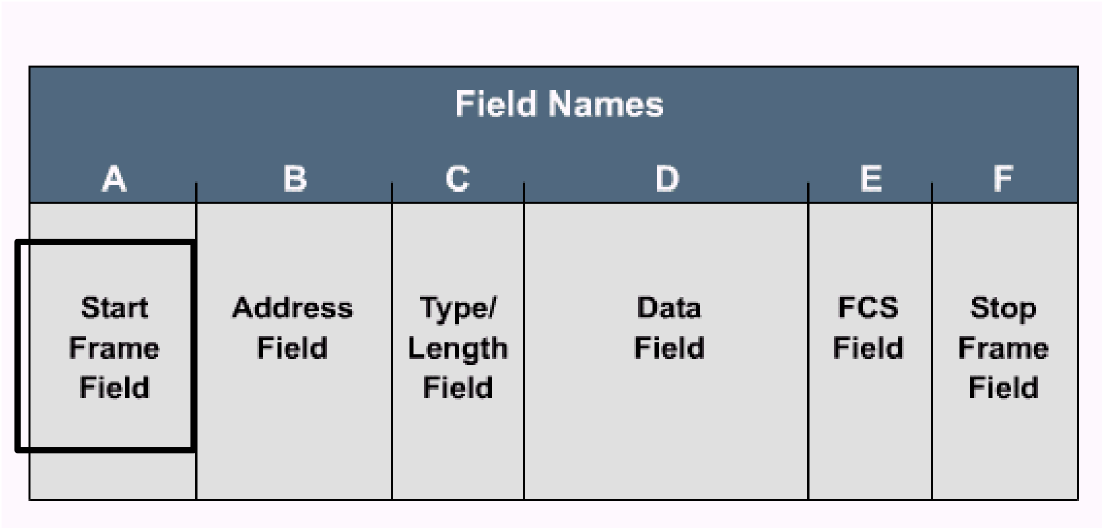
5-4-3 rule

- Ethernet and IEEE 802.3 implement a rule, known as the *5-4-3 rule*, for the number of repeaters and segments on shared access Ethernet backbones in a tree topology. The 5-4-3 rule divides the network into two types of physical segments: populated (user) segments, and unpopulated (link) segments. User segments have users' systems connected to them. Link segments are used to connect the network's repeaters together. The rule mandates that between any two nodes on the network, there can only be a maximum of **five** segments, connected through **four** repeaters, or *concentrators*, and only **three** of the five segments may contain user connections.



- The Ethernet protocol requires that a signal sent out over the LAN reach every part of the network within a specified length of time. The 5-4-3 rule ensures this. Each repeater that a signal goes through adds a small amount of time to the process, so the rule is designed to minimize transmission times of the signals.
- Note: This is really no longer an issue with switched networks. The 5-4-3 rule -- which was created when Ethernet, 10Base5, and 10Base2 were the only types of Ethernet network available -- only applies to shared-access Ethernet backbones. A switched Ethernet network is freed from the 5-4-3 rule because each switch has a buffer to temporarily store data and all nodes can access a switched Ethernet LAN simultaneously.

Generic Data Link Frame Format



Start Field

- When computers are connected to a physical medium, there must be a way they can grab the attention of other computers to broadcast the message, "Here comes a frame!"
- Various technologies have different ways of doing this process, but all frames, regardless of technology, have a beginning signaling sequence of bytes.

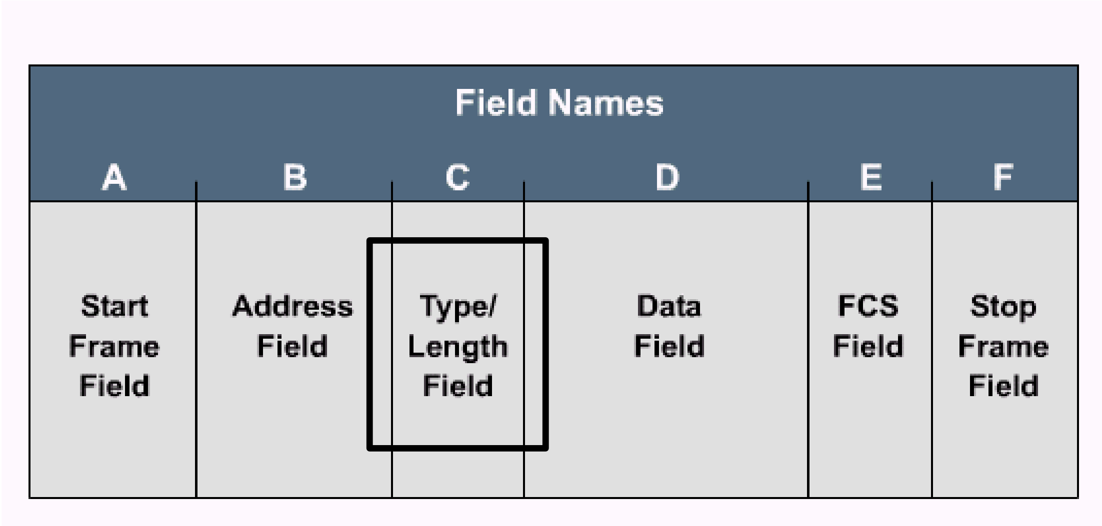
Generic Data Link Frame Format

Field Names					
A	B	C	D	E	F
Start Frame Field	Address Field	Type/ Length Field	Data Field	FCS Field	Stop Frame Field

Address Field

- We saw how IEEE 802.3 uses Destination and Source Addresses.
- BTW: Any idea how a serial data link frame is addressed?
  - – Dedicated Links - Broadcast
  - – Non-broadcast Multiple Access (NBMA), Frame Relay - DLCIs

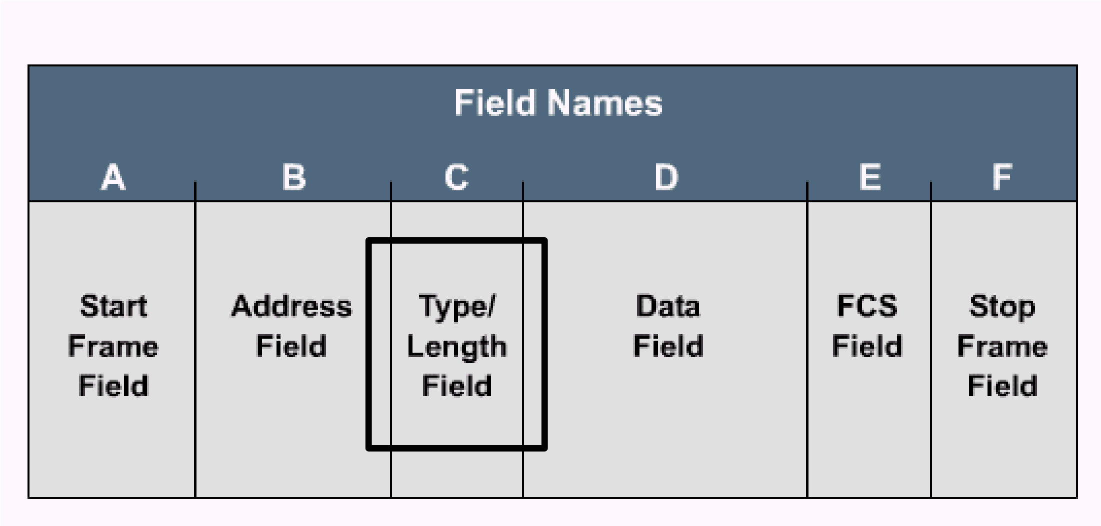
Generic Data Link Frame Format



Type Field

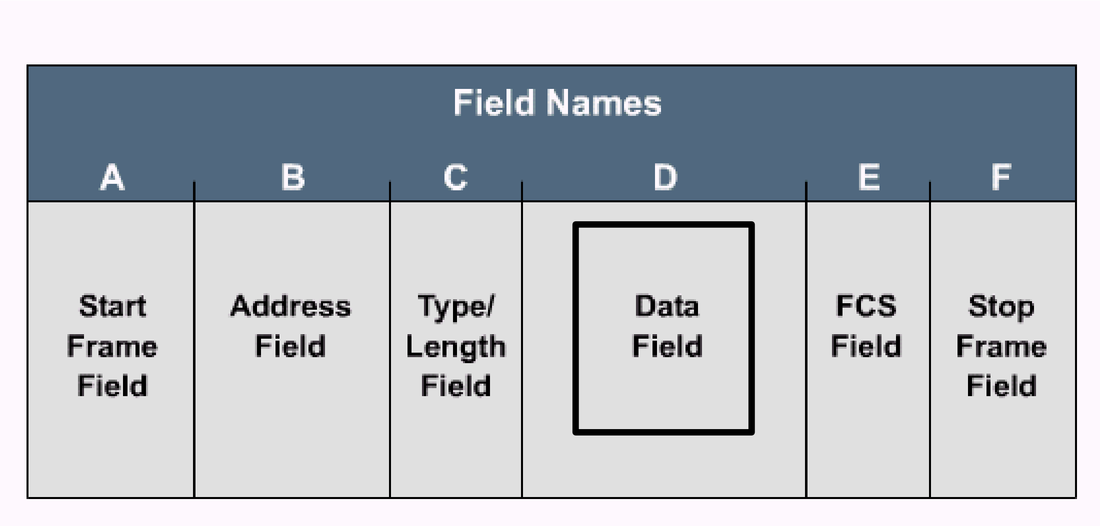
- Usually information indicating the layer 3 protocols in the data field, I.e. IP Packet.
- Type field values of particular note for IEEE 802.3 frames include:
  - 0x0600 XNS (Xerox)
  - 0x0800 IP (the Internet protocol)
  - 0x8137 Novell NetWare packet formatted for Ethernet II
  - 0x6003 DECNET

Generic Data Link Frame Format



Length Field

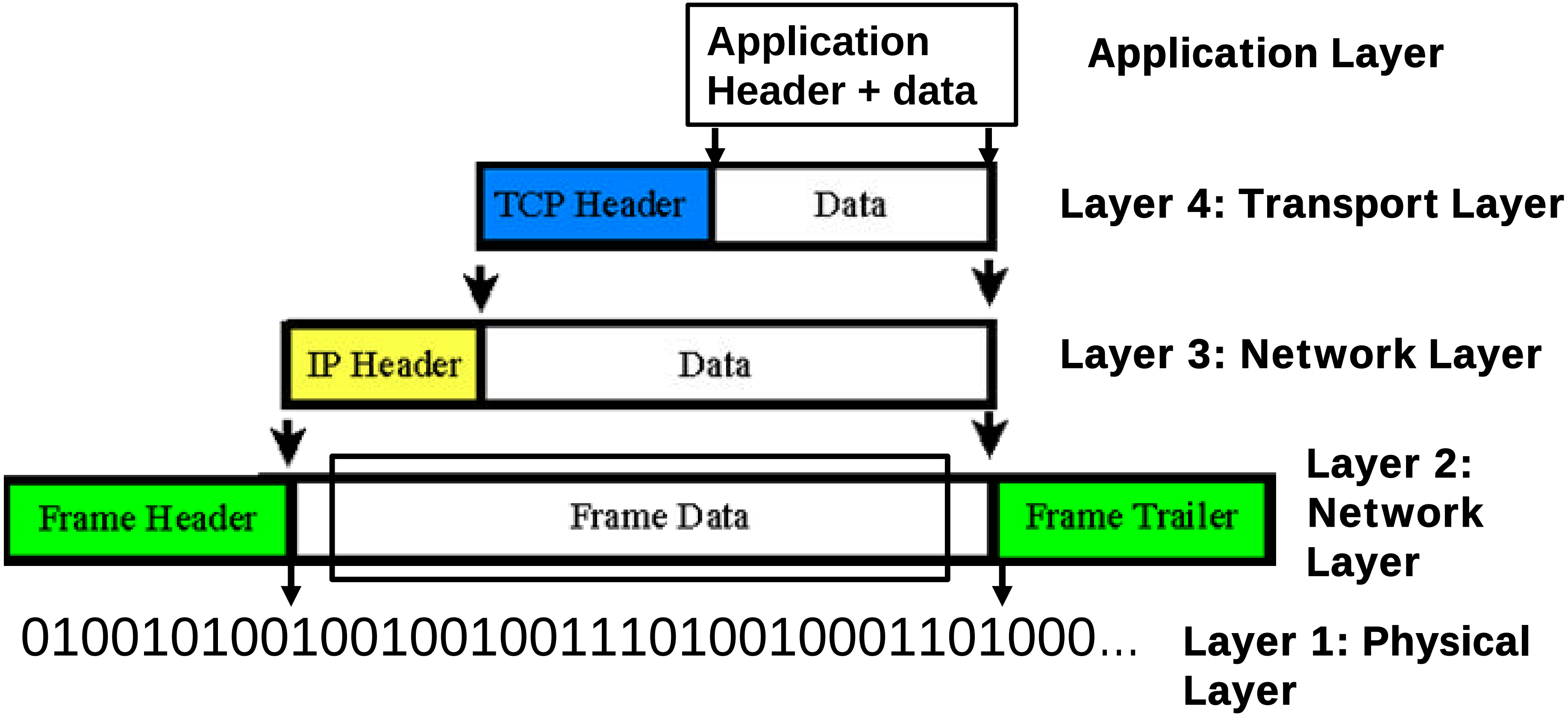
- In some technologies, a length field specifies the exact length of a frame.



Data Field

- Included along with this data, you must also send a few other bytes.
- They are called *padding bytes*, and are sometimes added so that the frames have a minimum length for timing purposes.
- LLC bytes are also included with the data field in the IEEE standard frames. (later)

Data Encapsulation Example



Generic Data Link Frame Format

Field Names					
A	B	C	D	E	F
Start Frame Field	Address Field	Type/ Length Field	Data Field	FCS Field	Stop Frame Field

FCS

- Used to insure that the data has arrived without corruption.
- Stations receiving the frame recalculate the FCS to determine if the incoming message is valid and then pass valid messages to the next higher layer in the protocol stack.

Three types of FCS

- Cyclic redundancy check (CRC)
  - – performs polynomial calculations on the data
- Two-dimensional parity
  - – adds an 8th bit that makes an 8-bit sequence have an odd or even number of binary 1s
- Internet checksum
  - – adds the numbers to determine a number

Generic Data Link Frame Format

Field Names					
A	B	C	D	E	F
Start Frame Field	Address Field	Type/ Length Field	Data Field	FCS Field	Stop Frame Field

Stop Field

- The computer that transmits data must get the attention of other devices, in order to start a frame, and then claim it again, to end the frame.
- The length field implies the end, and the frame is considered ended after the FCS.
- Sometimes there is a formal byte sequence referred to as an end-frame delimiter.