

Communications and Computer Networks

Summer Term 2023

Recap of last lecture (1/2)

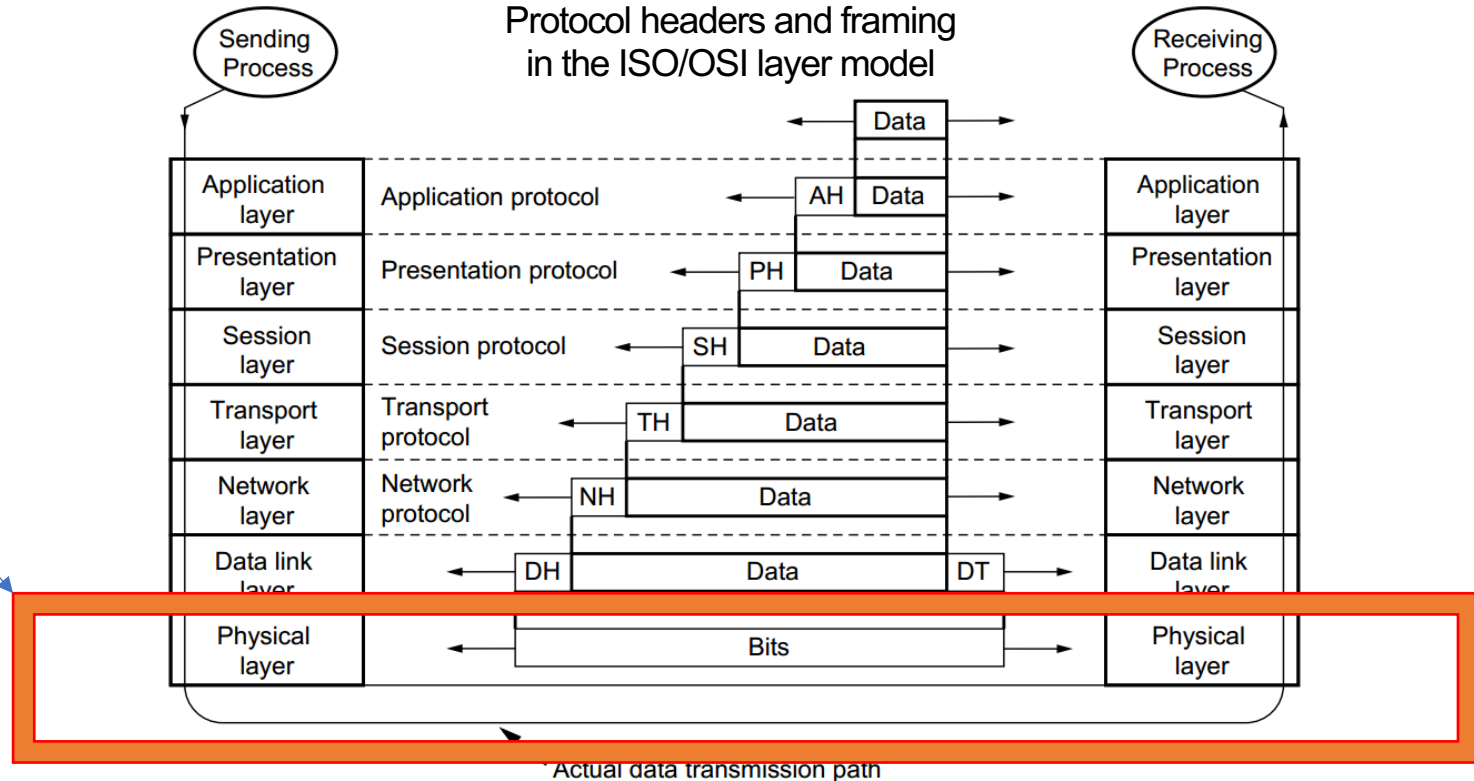
■ Physical layer and related topics

- You know the tasks of the physical layer
- You know fundamentals of data transmission
- You can explain what multiplexing techniques are needed for.
- You know the difference between baseband and broadband transmission
- You can explain the terms channel capacity, Nyquist theorem and Shannon-Hartley theorem
- You know the characteristics of the most important transmission media and hardware on layer 1

Layer	
7	Application
6	Presentation
5	Session
4	Transport
3	Network
2	Data Link
1	Physical

Recap of last lecture (2/2)

Discussion
till now:



Reasons for encoding

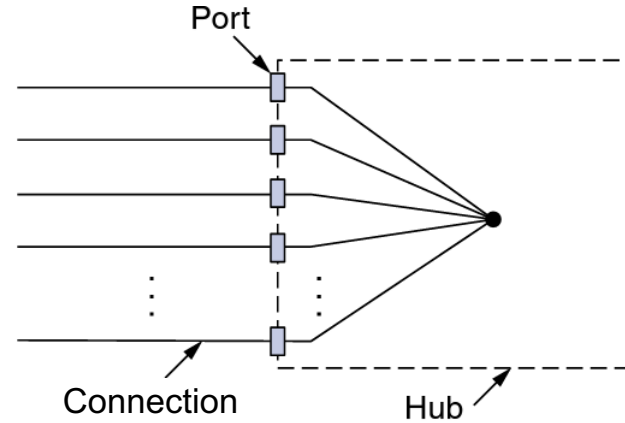
- Synchronization
- Clock recovery
- Minimize transmission hardware
- Convert signal to a specific medium
- Compression



- Optimized data transmission according to the medium



- Hubs are also devices of OSI Layer 1 and thus, completely transparent to the network.
- Hubs realize a star topology, with each port of the hub forming a single segment.
- Each device connected to a hub has the bandwidth of the LAN up to the hub port.
- Due to the star topology, the entire network is much safer against failures, since only the connected device/segment is affected in the event of cable problems and not the entire Ethernet network as with repeaters.
- Multiple hubs can be connected (cascaded) to each other, which makes it easy to increase the number of possible stations.



Layer 2 - Data Link Layer

- Fundamentals of the data link layer
- Ethernet
 - Short history
 - Carrier Sense Multiple Access/ Collision Detection (CSMA/CD)
 - Frame format
 - Ethernet and the physical layer
- Cyclic redundancy check (CRC)
- Hardware
- ARP
- VLAN

Learning Objectives

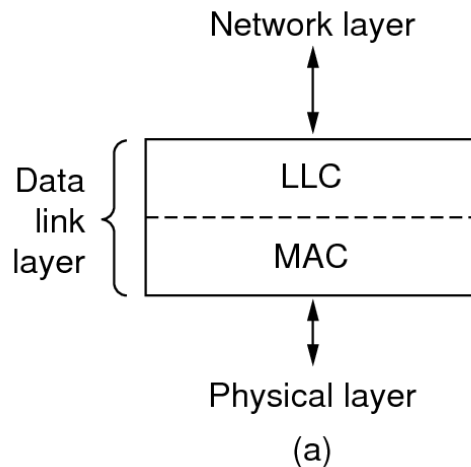
- You know fundamental tasks of the data link layer
- You know the basic principle of Ethernet and its technical development history.
- You can explain the CSMA/CD access procedure
- You know the structure of an Ethernet frame
- You know MAC-addresses and its usage in networks
- You can explain the task and procedure of the Address Resolution Protocol (ARP)
- You understand the technique of VLAN and the frame format according IEEE 802.1q

Layer 2 – Data link layer

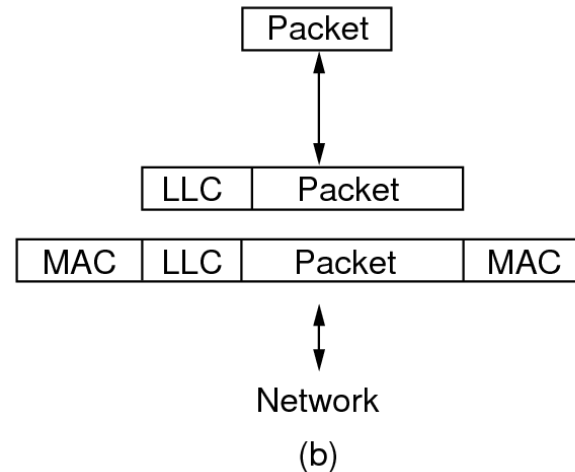
- The task of the data link layer is to **manage the incoming and outgoing data stream**. This is divided into transmission frames or packets, which are transmitted sequentially from the sender to the receiver.
- The **addressing scheme** on layer 2 provides a communication inside a restricted network area (**so called broadcast domain**)
- Furthermore, the link layer **has control mechanisms with which it can detect bit errors and „correct“ them by retransmission request (as far as possible)**.
- Relevant protocols:
 - Ethernet (IEEE 802.3)
 - Wi-Fi (IEEE 802.11)
 - *MPLS*
 - *PPP*
 - *ATM*

MAC and LLC

- Data Link Layer manages access to the medium (MAC – medium access control) and logical connection management for the upper layers (LLC – logical link control)



(a) LLC's position within the data link layer



(b) Package structure

Logical Link Control (LLC)

- LLC abstracts different MAC protocols according to IEEE 802.
- LLC is called layer 2b and together with the MAC layer (2a) forms the link layer.
- LLC distributes incoming data by forwarding it to the appropriate OSI Layer 3 instance protocols.
- It offers an unconfirmed, connectionless service (LLC 1), a connection-oriented service (LLC 2), as well as a confirmed connectionless service (LLC 3).
Nowadays, LLC 1 is used almost **exclusively**, as protocols on OSI layer 4 (e.B. TCP) take over the connection establishment and confirmation.

EDEC (Error Detection / Error Correction)

As discussed in CCN-1-PhysicalLayer, errors might occur when data is transferred

- Missing segments
- Lost packets
- Shifted bits
- Incomplete data

Detection:

- Checksum
- Parity bit
- CRC

Correction:

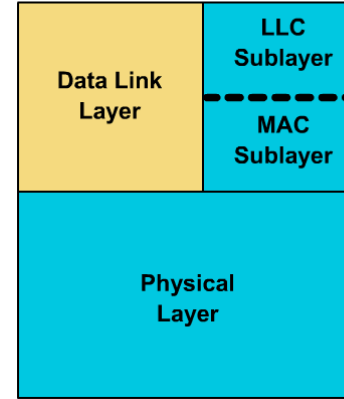
- Backward error correction
- Forward error correction

Ethernet

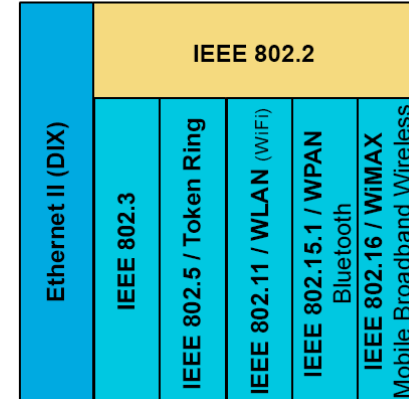
Ethernet

- The most relevant data link protocol nowadays
- Provides L2 **and** L1 specifications
- Multiple physical media support
- Intentionally a Baseband LAN technology
- Today basement for LAN, Highspeed Networks, data center, WAN
- Scalable and flexible, the same specification has been extended to:
 - Ethernet and IEEE 802.3—10 Mbps over coaxial cable, UTP and Fiber.
 - 100-Mbps Ethernet—Fast Ethernet, operating at 100 Mbps over UTP and Fiber cable.
 - 1000-Mbps Ethernet—Gigabit Ethernet, operating at 1000 Mbps (1 Gbps) over UTP and Fiber cable.
- Simplicity and low cost

OSI Layers



LAN Specification



Example Ethernet Standards

10Base5 (Thick-Ethernet, Yellow-Cable)

- Ethernet with 10Mb max. bandwidth
- Maximum cable length per segment is 500 meter.
- Both cable ends must be completed with termination resistors.
- 100 end devices can be connected per segment.
- Distance between two stubs must be a multiple of 2.5m.
- Stubs must not be longer than 50 meters.

100base2 (Thin-Ethernet, Cheapernet)

- Ethernet via RG58 50Ω coax cable.
- For 10Mb only.
- BNC connector and T-pieces for connections.
- 50Ω termination resistors on both ends.
- Maximum segment length of 185m.
- Minimum station distance: 0.5m.
- Maximum number of stations: 30
- No stubs are allowed.

10Base-T

- Transmission of 10Mb/s via Twisted Pair Cable
- Hub necessary.
- Maximum distance 100m.
- Two pairs of wires are needed.
- Cables from category 3.
- Full duplex mode available with switch.

100Base-T

- Transmission of 100Mb/s via Twisted Pair Cable
- Hub necessary.
- Maximum distance 100m.
- Two pairs of wires are needed.
- Cables from category 5.
- Full duplex mode available with switch.

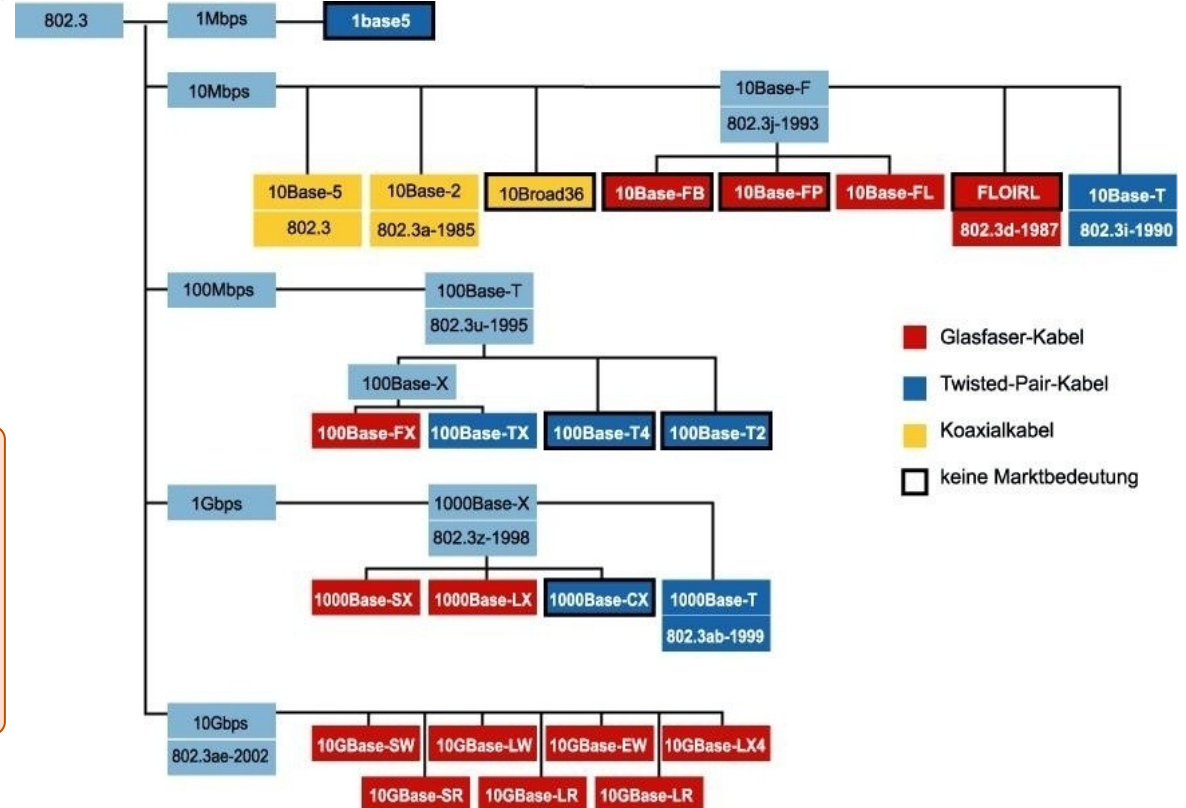
100BaseFX

- Transmission of 100Mb/s via optical fibre
- Maximum segment length of 400m.
- Max. distance between two 100BaseFX switches is 2000m

100BaseT4

- Transmission of 100Mb/s via Cat 3 cable
- All 4 pairs of wires are used.

Development of Ethernet standards



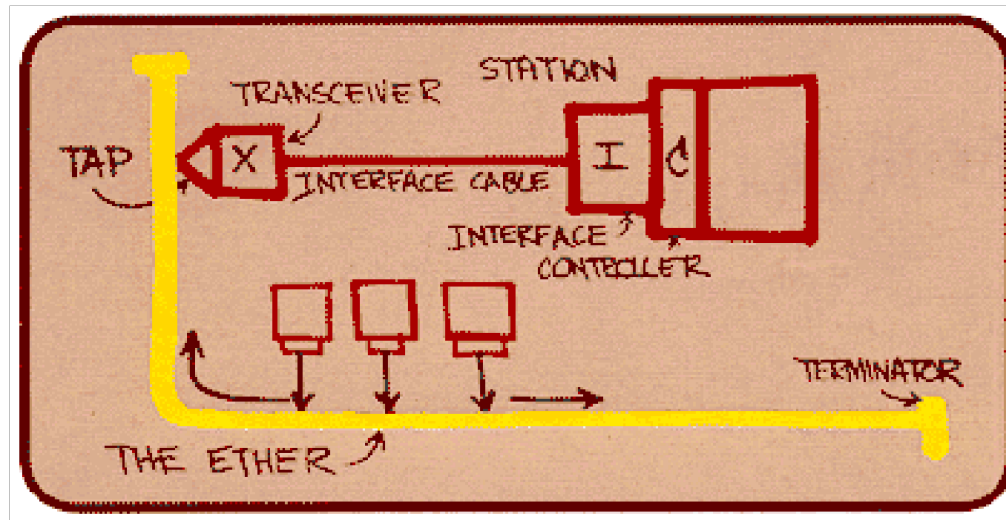
Current Ethernet standards:

- 802.3bz 2016-09:
2.5GBASE-T and 5GBASE-T via Cat-5/Cat-6 twisted pair
- 802.3bs 2017-12:
200GbE (200 Gbit/s) over single-mode fiber and 400GbE (400 Gbit/s) over optical physical media

Ethernet - History

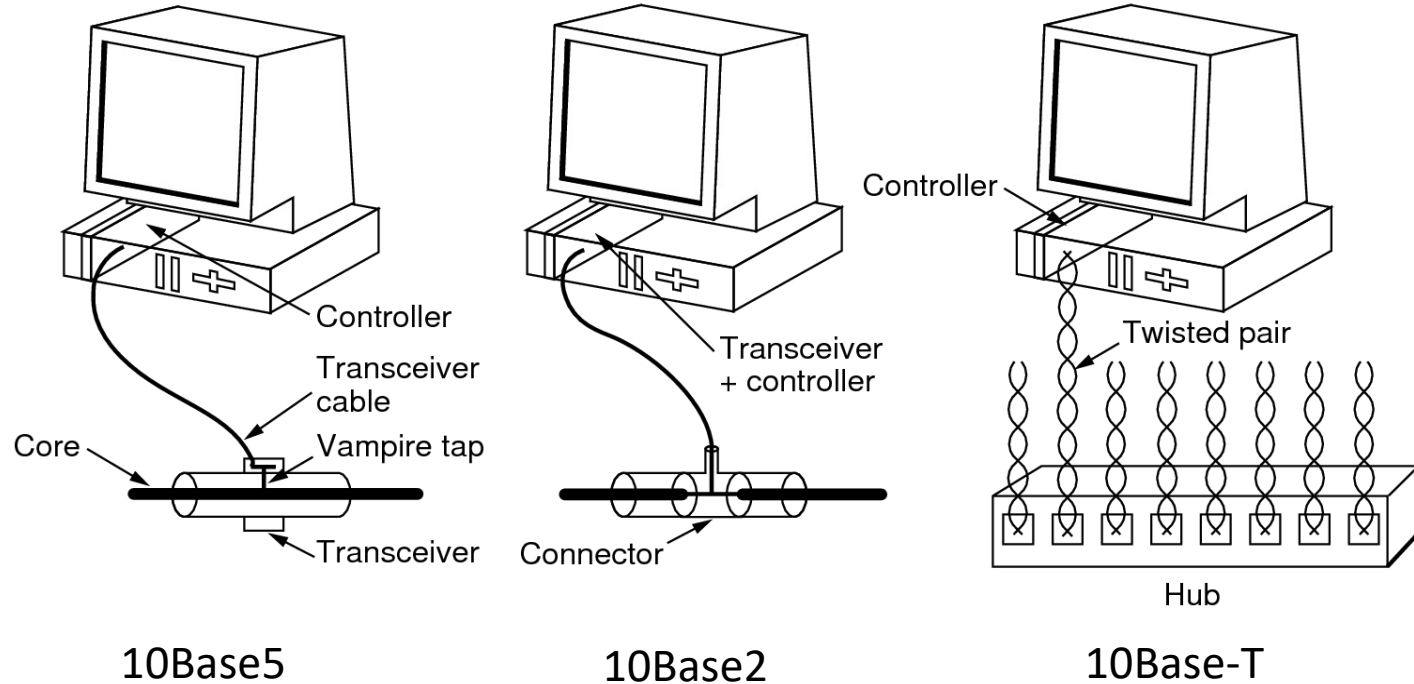
The Ethernet

Ethernet was developed in 1972 at the Xerox PARC (Palo Alto Research Center).

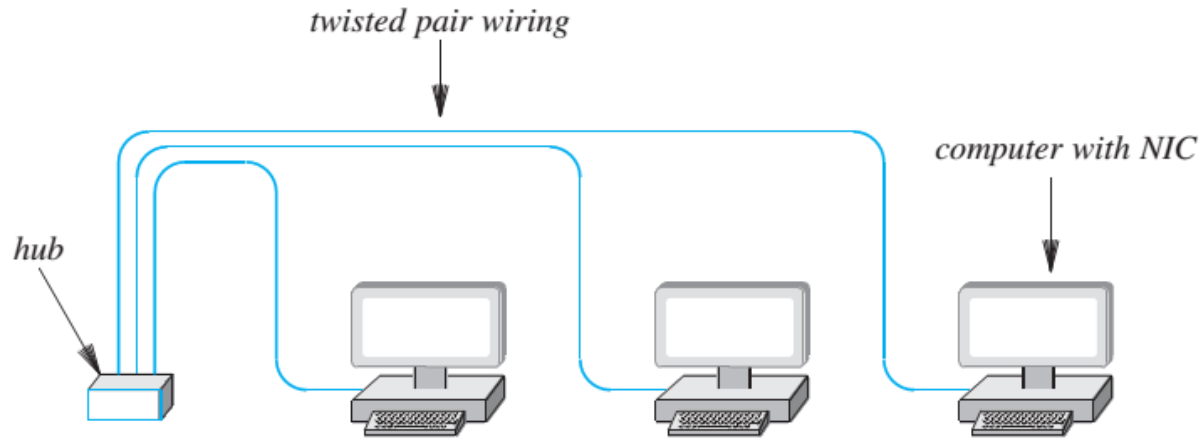


This drawing was made in June 1976 by Dr. Robert M. Metcalfe for presentation at the National Computer Conference and contains the original names.

Types of Ethernet Cabling (History)

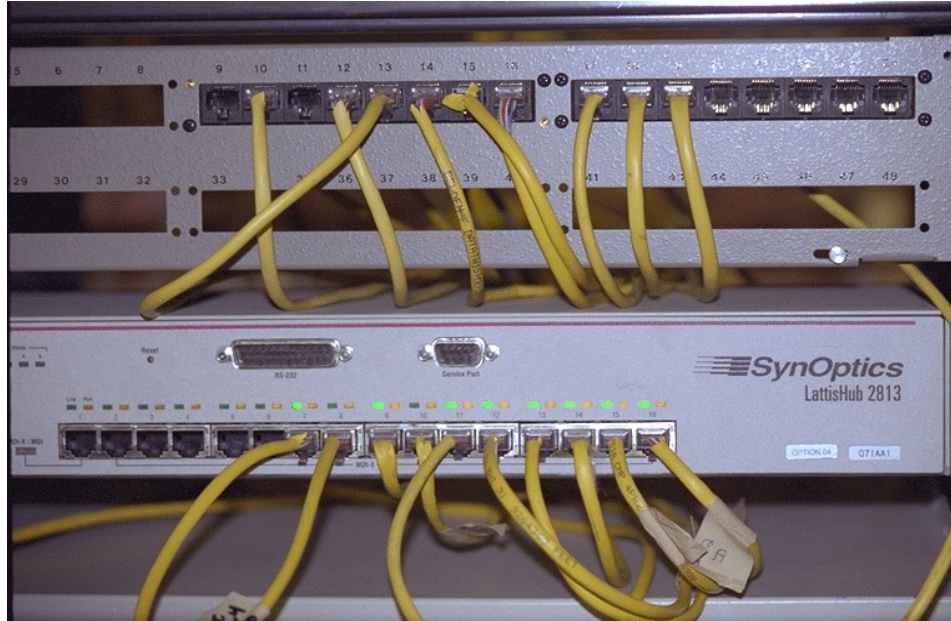


Twisted Pair Ethernet (10BaseT)



Twisted pair cable with RJ-45 connector, connected via hub

Twisted Pair Ethernet (10BaseT)



Ethernet – CSMA/CD

Network structure of classic Ethernet

- The basis of classic Ethernet is the Bus. This can be realized by different transmission media.
- In the classic Ethernet this was 10Base5, using a yellow coaxial cable with an impedance of 50Ω , an attenuation of 17 dB/km at 10 MHz and a signal propagation speed of $0.67 \cdot c$ (c = speed of light $\approx 3 \cdot 10^8$ m/s).
- If several subscriber stations are connected to one cable, a so-called network segment is created. An Ethernet can consist of several such network segments, which are then connected via repeaters.
- The maximum length of a network segment must not exceed a maximum length (500 m for 10Base5). The maximum network extension must not exceed 2.5 km and there must be no more than 4 repeaters between two stations.
- This restriction is specified by the MAC protocol (CSMA/CD).

Collision inside a network

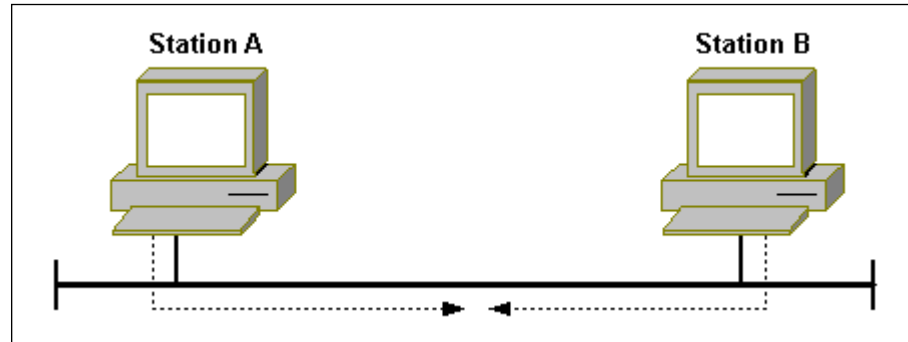
- The Ethernet is based on a bus structure both physically and logically, so that a packet reaches all connected stations and can be read by all (broadcast network).
- All stations **share** the transmission channel (characteristic of a broadcast network) and are thus in **competition for transmission capacity**.
- This competitive situation requires a protocol that ensures that **only one station transmits at a certain time**.
- A station of an Ethernet network can in principle access the transmission medium at any time in order to send its messages.
- But a station may only send if the medium is free and not already occupied by another station, since two or more simultaneously sent messages interfere with each other (collision) and thus cannot be received correctly by the respective recipient.

CSMA/CD

Carrier Sense Multiple Access/Collision Detection

The CSMA/CD (Carrier Sense Multiple Access/ Collision Detection) protocol provides the relevant aspects and restrictions to fulfil these requirements

- Carrier sense
- Multiple Access
- Collision detection



Characteristics of CSMA/CD

- A station willing to transmit must first listen to the transmission channel, which is also referred to as Listen Before Talking (LBT).
- A station willing to broadcast, which finds a free medium, immediately begins the transmission.
- During transmission, the **station continues to listen** to the medium in order to be able to detect collisions immediately. A collision is detected by the fact that the station hears something different on the medium than it has sent.
- If a collision is detected, the station immediately cancels the transmission and sends a so-called JAM signal (a 32-bit sequence of 1 and 0) so that all states on the bus notice that a collision has taken place.
- After a transmission has been aborted due to a collision, the station does not start a new transmission attempt until it has waited for a certain, random time.

Exponential backoff time

This random latency is determined by the backoff method and is called truncated binary exponential backoff. It is defined as follows:

$$W = i \cdot T$$

with

W waiting time

i Random number from the interval $0 \leq i \leq 2k$ with $k = \min(n, 10)$

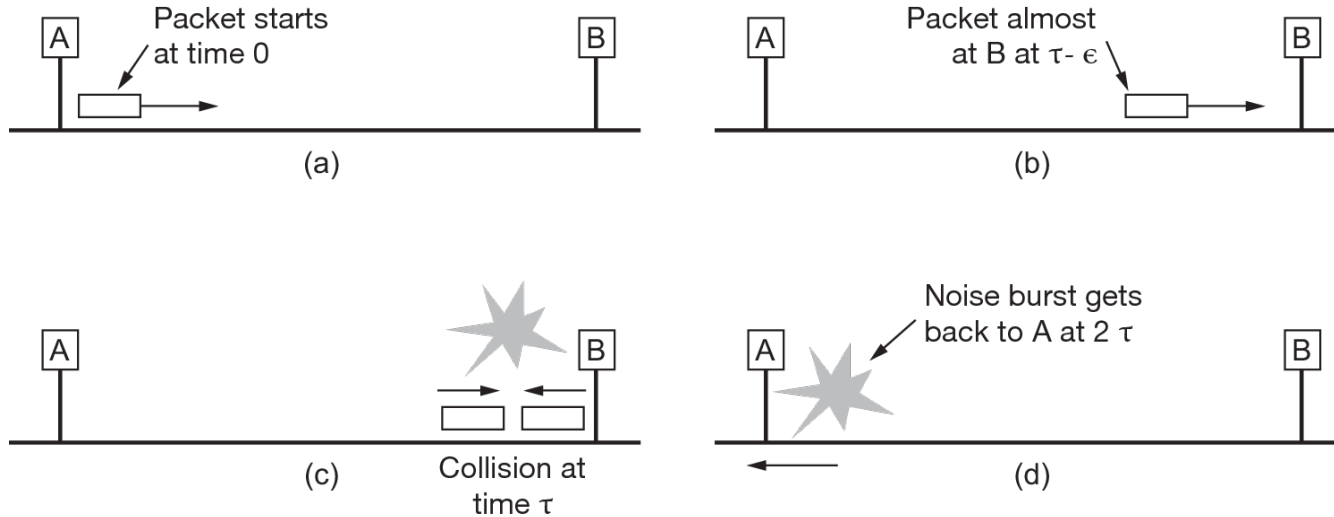
n Number of repetitions of the same block, i.e. collisions

T slot time, which corresponds to the round trip delay,
i.e. twice the maximum signal time of e.B. 51.2 μs .

After 10 unsuccessful transmission attempts, the waiting time does not increase any further. After 16 attempts, it aborts and generates an error message.

The probability of collisions increases with the number of stations and the length of the line.

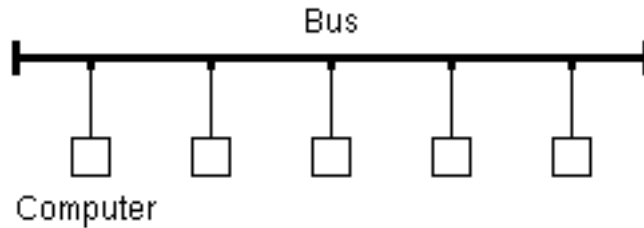
Timing consideration of collision detection (worst case)



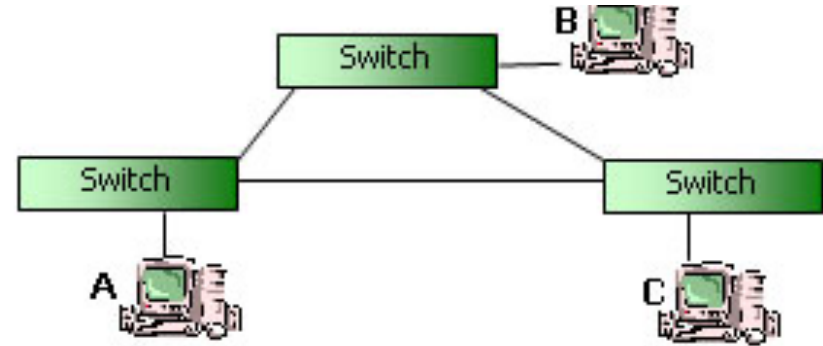
- (a) A sends packet ($t = 0$)
- (b) The package has almost reached B ($t \approx \tau$)
- (c) B sends packet because channel is still free \Rightarrow collision ($t = \tau$)
- (d) Destroyed packet reaches A after $t = 2\tau$

Ethernet – Current version

Comparison



old

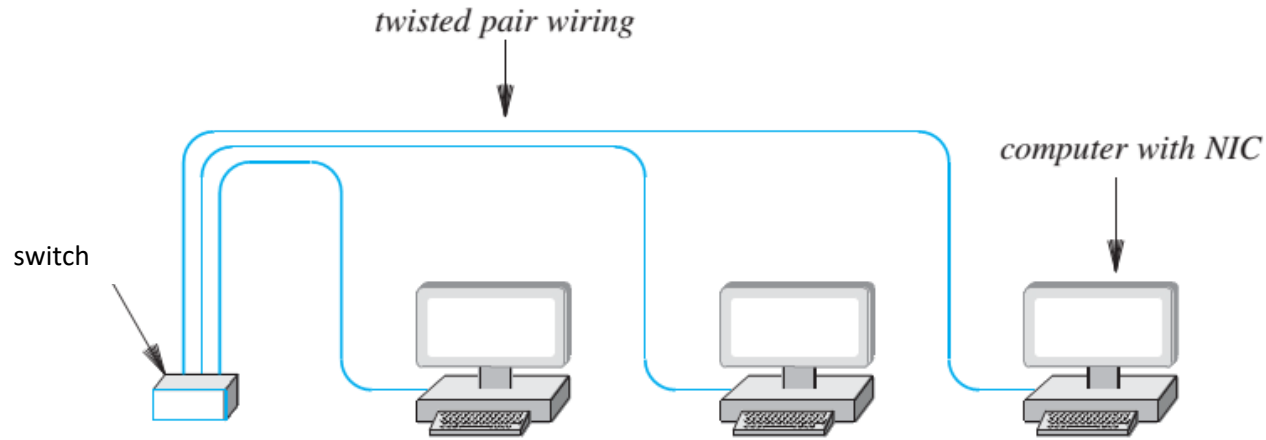


now

Network structure of modern Ethernet

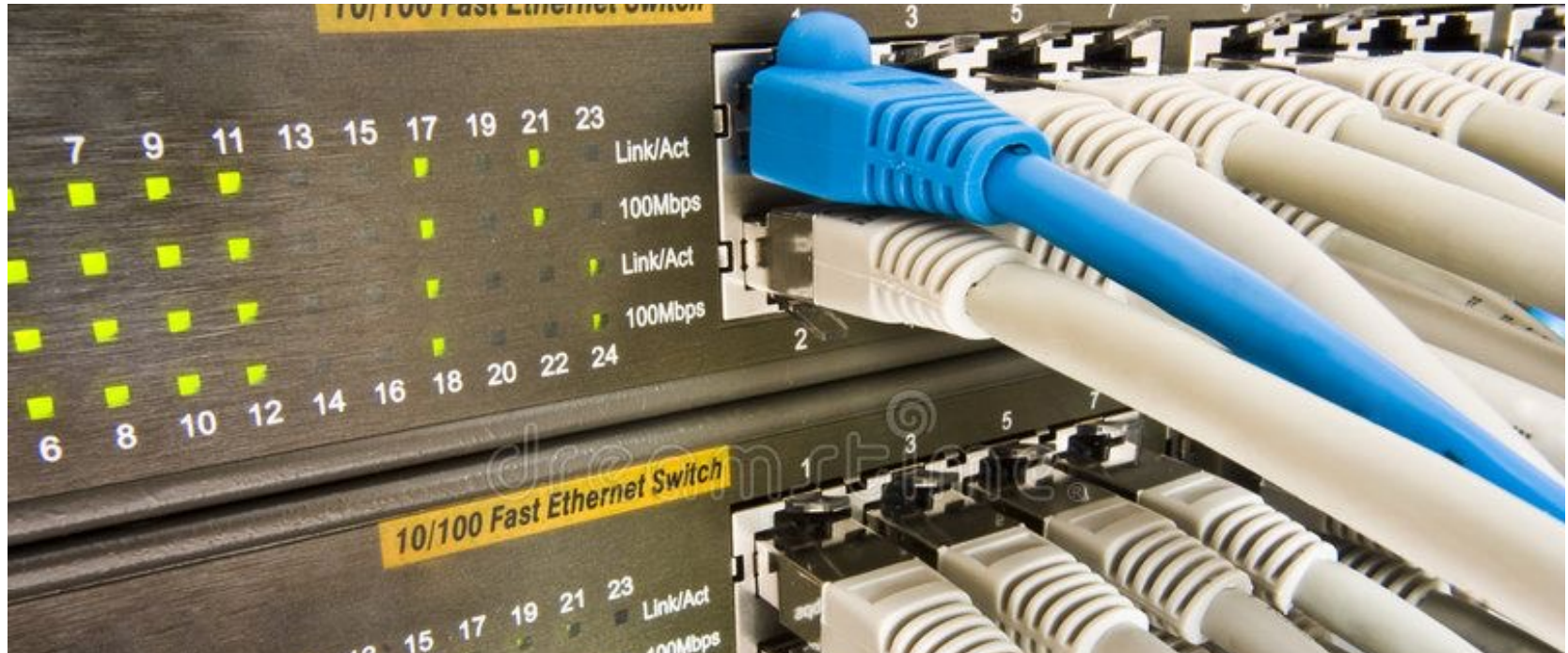
- Traditional Ethernet is based on a bus topology
- Modern Ethernet is based on the star topology, whereby the logical structure of a broadcast network is retained.
- The star point is a switch
- Multiple connections can be switched simultaneously via one switch.
- The cascading of several switches is possible.
- Various twisted-pair copper cables and fiber optic cables are used as transmission media.
- The transmission usually takes place full duplex, so that the MAC protocol CSMA/CD plays a **subordinate** role.
- Each connection between switch and a host is a single collision domain

Twisted Pair Fast Ethernet (100BaseT)

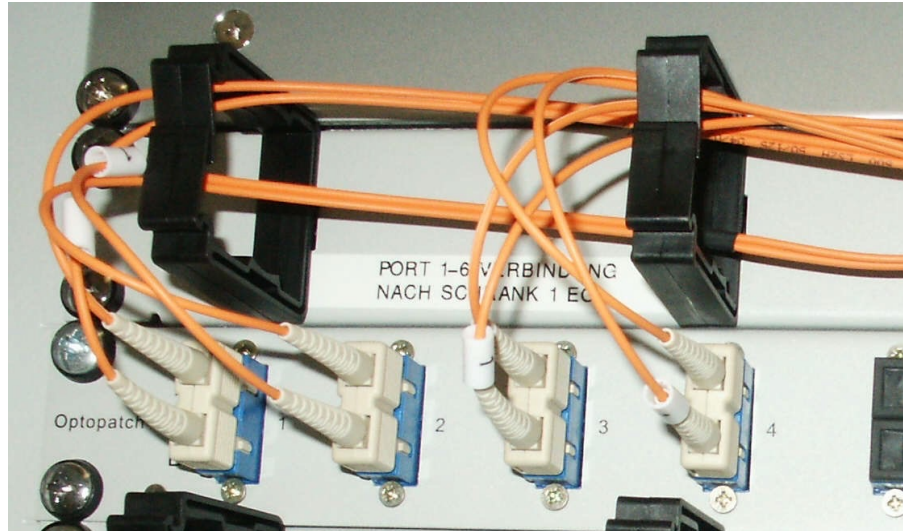


Twisted pair cable with RJ-45 connector, connected via Switch

Twisted Pair Fast Ethernet (100BaseT)



Fiber Patch Panel and 10GBase SR Ethernet Card

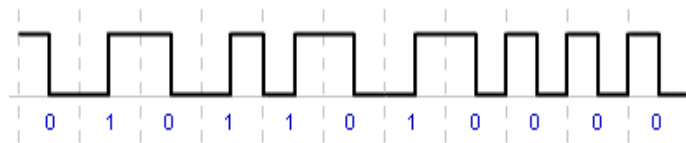


Ethernet – Physical Layer

Ethernet Encodings

- Ethernet has evolved from the original **10Mbps** copper based technology to **100Mbps** then **1Gbps** and **10Gbps** fiber based technology.
- In order to transport digital bits of data across any medium, various issues need to be considered; the bit transmission **speed**, the **frequency** of the carrier wave, **S/N** signal to noise ratio, **DC** balance and **clocking**.
- For each Ethernet technology, particular **encoding techniques** were developed such that data could be transmitted across the medium with the greatest efficiency.

Manchester encoding



- To send a logic '0' data bit, decrease the voltage down in the middle of the bit period.
- To send a logic '1' data bit, increase the voltage up in the middle of the bit period.
- Advantage: Always there is a voltage level transition in the middle of each bit-time. Thus, Manchester encoding is effectively **self-clocking**.
 - A separate clock signal is not required
 - There are no long strings of logic '0' or logic '1' levels to cause the clock to drift.
- Disadvantage: When a series of alternating 0's and 1's are sent, the frequency of the signal is equal to the bit rate, but, when a series of 0's or 1's are sent, an extra transition occurs between each bit. The frequency of the signal could be twice double the bit rate.
 - The **frequency** may reach **twice** that of NRZ
- Manchester encoding is used on **Legacy 10Mbps Ethernet** networks.

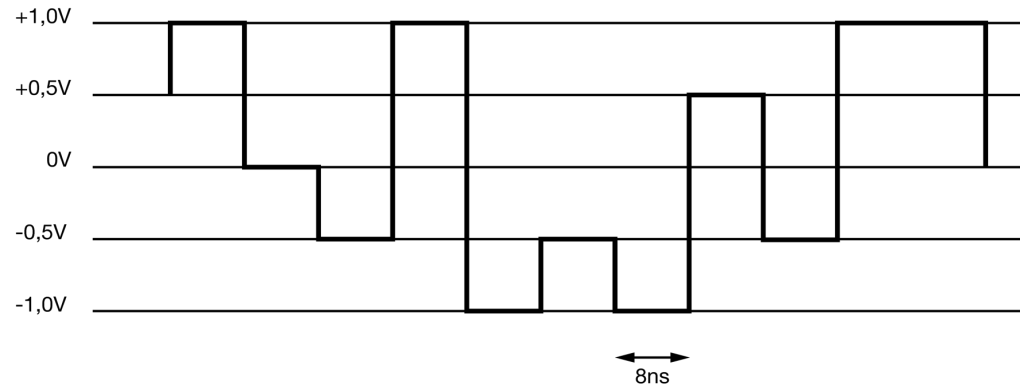
Pulse amplitude modulation

PAM-5 is the modulation method used for **Gigabit Ethernet** (1000Base-T).

It is a combination of pulse modulation and amplitude modulation (pulse amplitude modulation) with $n = 5$ stages per transmission step.

By using four double wires at the same time, a total of $5^4 = 625$ states per clock can be distinguished with the five stages, of which $2^8 = 256$ (one byte per clock) are used for data.

1 byte per 8 nanoseconds = 1 bit per nanosecond = 1000 megabits/s = 1 gigabit/s



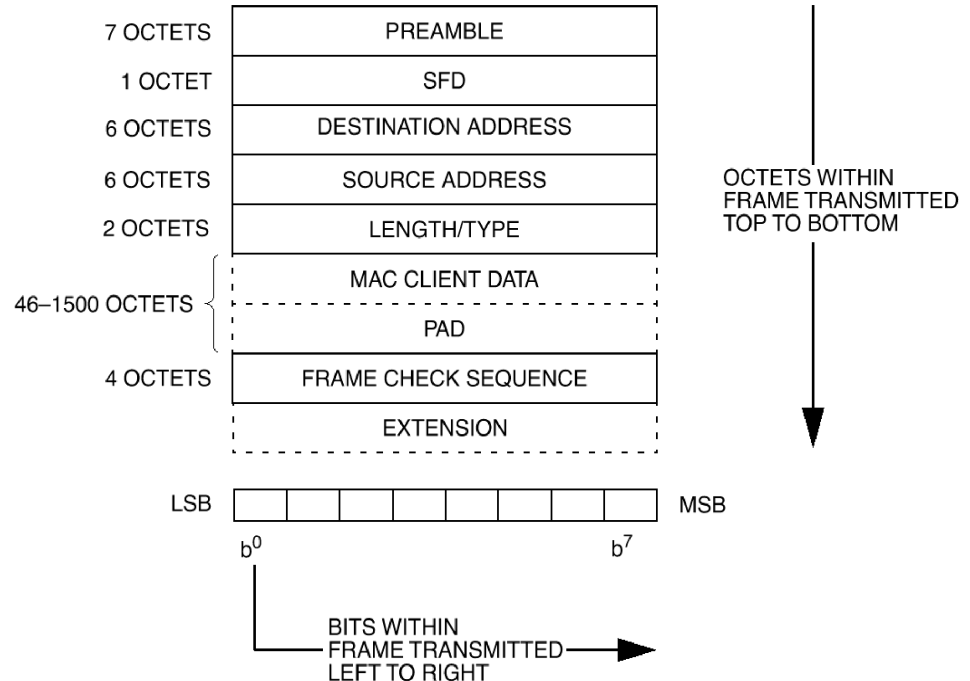
Ethernet Encodings

- Encoding on 10BASE-X Legacy Ethernet:
Manchester Encoding
- Encoding on 100Mbps Fast Ethernet:
100BASE-TX: 4B5B Encoding + MLT-3 Encoding
100BASE-FX: 4B5B Encoding + NRZ-I Encoding
- Encoding on 1000Mbps Gigabit Ethernet:
1000BASE-T: 8B1Q4 and 4D-PAM5 Encoding
1000BASE-X: 8B10B and NRZ Encoding
- Encoding on 10Gbps Ethernet:
10GBASE-X: 8B10B or 64B/66B (depending on the fiber) and NRZ Encoding

Transmission of octets and bits on the physical layer

Ethernet transmits the data **serially**, starting in each case with the lowest, lowest-order bit of an octet (b0).

The octets of the individual fields are transferred as **BigEndians**, i.e., the octet with the higher power first.

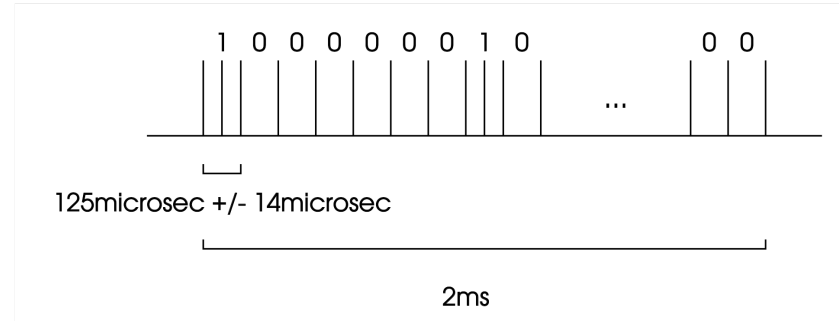


Auto-Negotiation

- Signaling mechanism between two systems to choose transmission parameters
 - Speed
 - Duplex
 - Flow control
- Defined in clause 28 of IEEE 802.3
- FE AN bases on Normal link pulse (NLP) of 10BaseT
- Uses *link integrity test* (LIT) pulses
 - positive-only electrical pulses
 - nominal duration of 100 ns,
 - maximum pulse width of 200 ns,
 - generated at a 16 ms time interval with a timing variation tolerance of 8 ms

- Change of LIT parameters led to Fast Link Pulse (FLP)
- The FLP burst consists of 17 NLP at a 125 μ s time interval with a tolerance of 14 μ s.
- If there is a additional positive pulse between a pair of NLP indicates a 1, an absence indicates 0
- As a result, every FLP transfers a 16bit word, named link code word (LCW)

Bits	Field
0-4	Selector Field
5-12	Technology field
13	Remote fault
14	Ack
15	Next page



Auto-Neg III

- Technology field (bit 5-12) defines the ability of the link

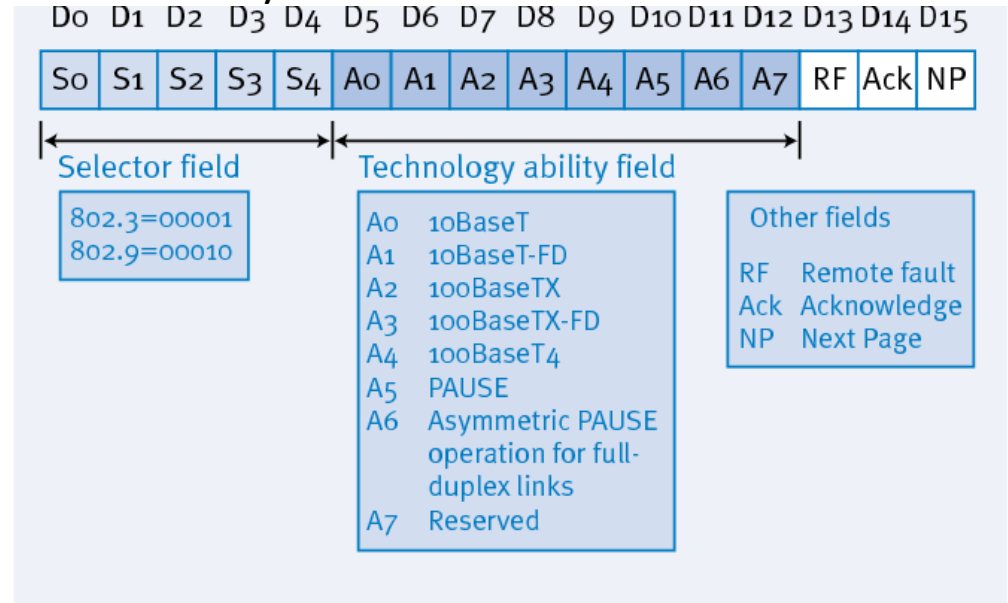
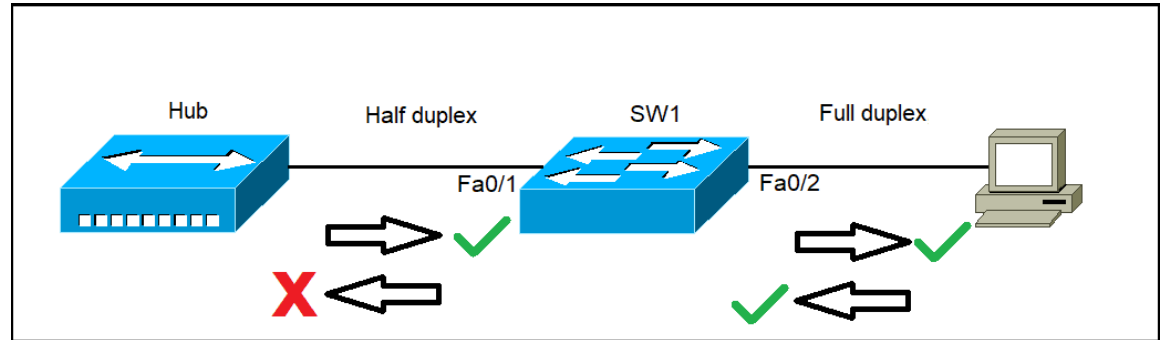


Figure 5. Base link code word definition

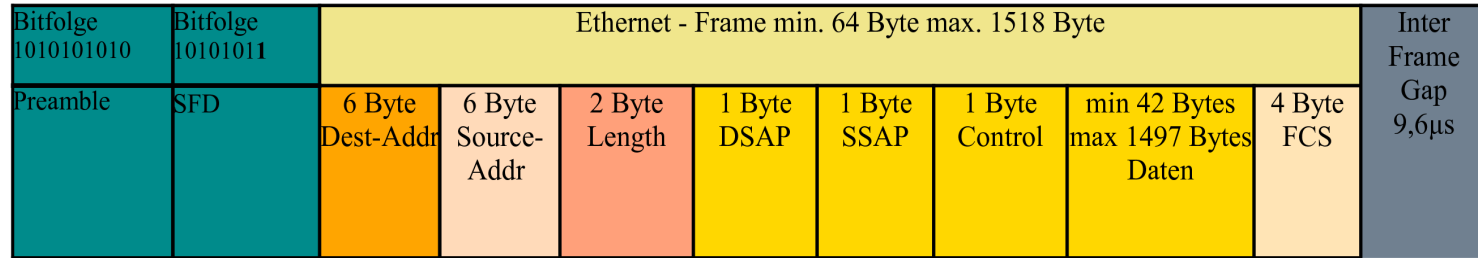
Duplex-mode

- Direction of communication
- Half-duplex (hdx): Only one partner can send at a time (like Walkie-Talkies)
- Full-duplex (fdx): Both partners can communicate simultaneously (like POTS)



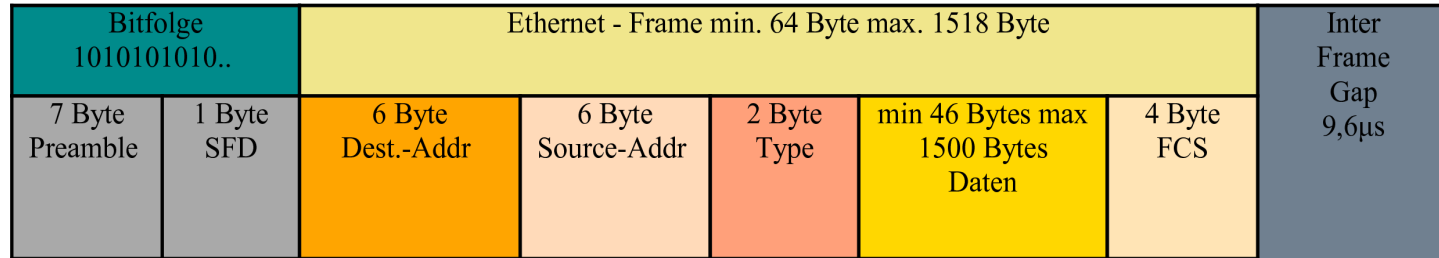
Ethernet – Protocol format

Overview of Ethernet frame acc. to IEEE 802.3



- IEEE 802.3 frames have inserted a 2-byte length field instead of the type field. It specifies the number of bytes in the data field including 802.2 LLC headers. If the value of the two bytes after the source address is less than or equal to the maximum possible 1500 bytes, it must be an IEEE 802.3 frame.
- IEEE 802.2 defines the Logical Link Control (LLC) protocol layer. Instead of the type field with the protocol ID, the Destination Service Access Point (DSAP) and the Source Service Access Point (SSAP) exist. The control field contains the type of LLC frame.
- In IEEE 802.3, the entire transmission unit, including Preamble and SFD, is called an Ethernet packet. The Ethernet frame starts after the SFD.

Overview of Ethernet frame acc. to Ethernet II



- The classic frame structure is Ethernet II.
- The preamble is used to synchronize the recipients.
Including the Start Frame Delimiter (SFD), it consists of a vibration of $6.4 \mu s$ length (sequence of 8 bytes with 1010...).
- The frame must be at least 64 bytes in size to achieve the minimum slot time to detect a collision. Otherwise, bits are added.
- Inter Frame Gap (also called inter frame spacing (IFS))) needed for recovery of clocking, power up after low-power state

Ethernet II – Protocol fields

Bitfolge 1010101010..		Ethernet - Frame min. 64 Byte max. 1518 Byte					Inter Frame Gap 9,6µs
7 Byte Preamble	1 Byte SFD	6 Byte Dest.-Addr	6 Byte Source-Addr	2 Byte Type	min 46 Bytes max 1500 Bytes Daten	4 Byte FCS	

Name	Size	Task
Dest-Addr	6 Byte	MAC-address of the destination
Source-Addr	6 Byte	MAC-address of the source
Type / Length	2 Byte	Ethertype, Values of 1500 and below mean that it is used to indicate the size of the payload in octets, while values of 1536 and above indicate that it is used as an EtherType, to indicate which protocol is encapsulated in the
Data	4 – 1500 Bytes	Payload of the Frame
FCS	4 Byte	Frame Check Sequence with a cyclic redundancy check CRC32 (Cyclic Redundancy Check with 32 bit).

MAC-Addresses I

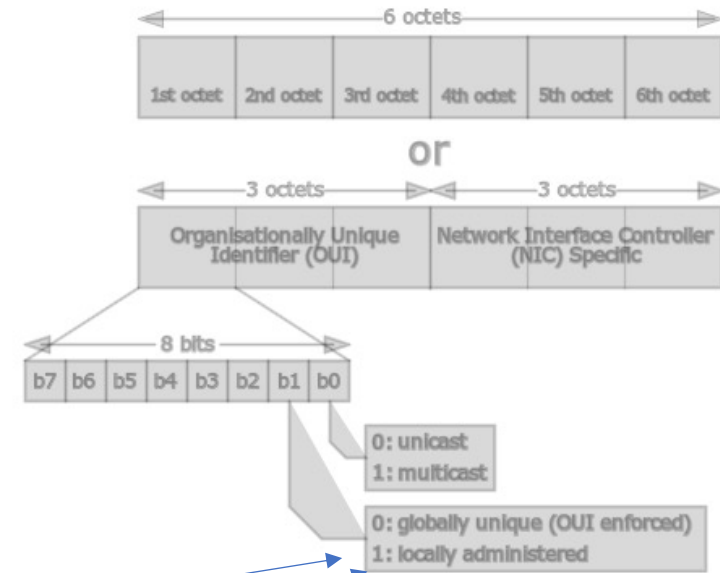
- Each MAC address has a length of six bytes or 48 bits.
- MAC addresses are traditionally represented as a sequence of six two-digit hex numbers separated by colons.
- Bit 0 of the first byte decides whether it is a unicast (0) or broadcast/multicast address (1).
- Bit 1 of the first byte decides whether the MAC address is administered globally (0) or locally (1). Network cards have a globally unique MAC address, which is managed globally by the IEEE and the manufacturer.
- The first 24 bits (bits 3 to 24) of a global address contain the manufacturer identifier (OUI – Organizationally Unique Identifier) assigned by the IEEE. The remaining 24 bits (bits 25 to 48) are individually defined by the respective manufacturer.
- Locally administered can be used for mac-address-randomization or VMs

MAC-Addresses II

- Different spellings (but always 48 Bit, 6 Byte):

- Cisco: 0123.4567.89ab
- Windows: 01-23-45-67-89-ab
- Linux: 01:23:45:67:89:ab
- Variations: 0123456789B, 01.23.45.67.89.AB

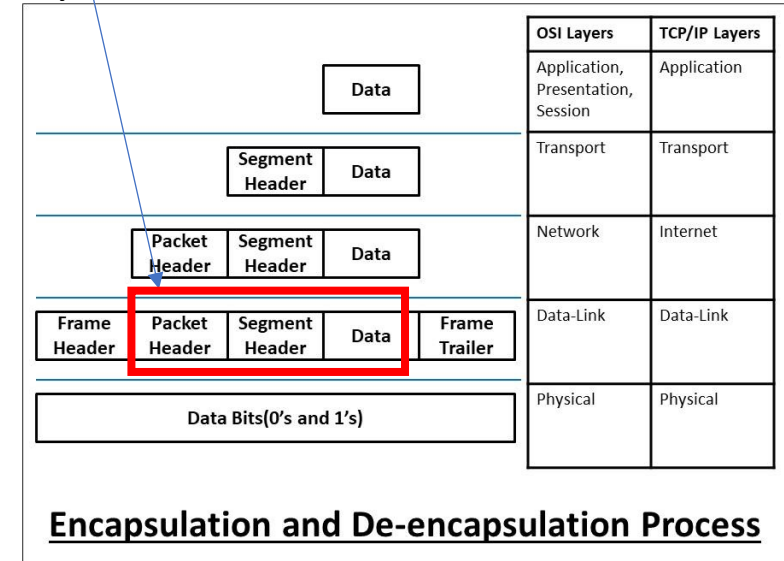
- Locally administered: 02:ab:ab:00:00:12 (First Byte: 00000010)
- Broadcast (all devices listen to this address): FF:FF:FF:FF:FF:FF
- Unique address (only one specific device listen to this): 5c:e9:1e:ae:7c:ef
- IPv6-Multicast (will be discussed later): 33:33:XX:XX:XX:XX



EtherType

- Indicates which protocol is encapsulated in the payload
- Values are assigned by IEEE Registration Authority

Value	Protocol
0x0800	IPv4
0x86DD	IPv6
0x0806	ARP
0x8100	VLAN 802.1q
0x22F3	TRILL
...	



Frame size

- Maximum frame size of an Ethernet frame is 1518 (with 802.1Q up to 1522) bytes
- Resulting max size of PDU is 1500 bytes, min size is 64 bytes
- Jumbo frame specification allows payload size of 9000 bytes
 - Improved transmission in specific situations (reduced CPU cycles)
 - Not part of the IEEE 802.3 standard
 - Primary use in some gigabit (GE) switches and NICs, partly in FE

Why must the IEEE 802.3 (Ethernet) frame be at least 64 bytes long?

- Frames must be at least 64 bytes long, not including the preamble, so, if the data field is shorter than 46 bytes, it must be compensated by the Pad field.
- Minimum length bases on the collision-detect mechanism. In CSMA/CD a station must never be allowed to believe it has transmitted a frame successfully if that frame has, in fact, experienced a collision.
- In the worst case it takes twice the maximum propagation delay across the network before a station can be sure that a transmission has been successful.

Why must the IEEE 802.3 (Ethernet) frame be at least 64 bytes long?

Calculations:

- LAN Length (L) = 500 m (per segment) x 5 segments = 2500 meters
- Velocity of propagation on the cable (V) = $2 * 10^8$ meters/sec
- Delay added by repeater (D) = $\sim 3 \mu\text{s} \times 2$ (Bi-Direction) x 4 Repeaters = $24 \mu\text{s}$
- Round Trip Delay (RTD) = (Total Distance/V) + Repeater Delays (D)
 - Total Distance/V = $(2 * 2500 / 2 * 10^8) = 25 * 10^{-6} \text{ sec} = 25 \mu\text{s}$
 - Hence RTD = $25 + 24 = 49 \mu\text{s}$
- Now, time to transmit 64 bytes = $512 \text{ bits} / 10 * 10^6 = 51.2 * 10^{-6} \text{ s}$ or $51.2 \mu\text{s}$ (referred to as slot time in the 802.3) which is greater than the RTD of $49 \mu\text{s}$.
- Hence the minimum frame size for the IEEE 802.3 (Ethernet) is 64 bytes.

OS commands I

Show own network configuration

- `ifconfig` (Linux, macOS)
- `ip link show` (Linux)
- `ipconfig /all` (Windows)

```
C:\WINDOWS\system32\cmd.exe
C:\Documents and Settings\Corey>ipconfig /all

Windows IP Configuration

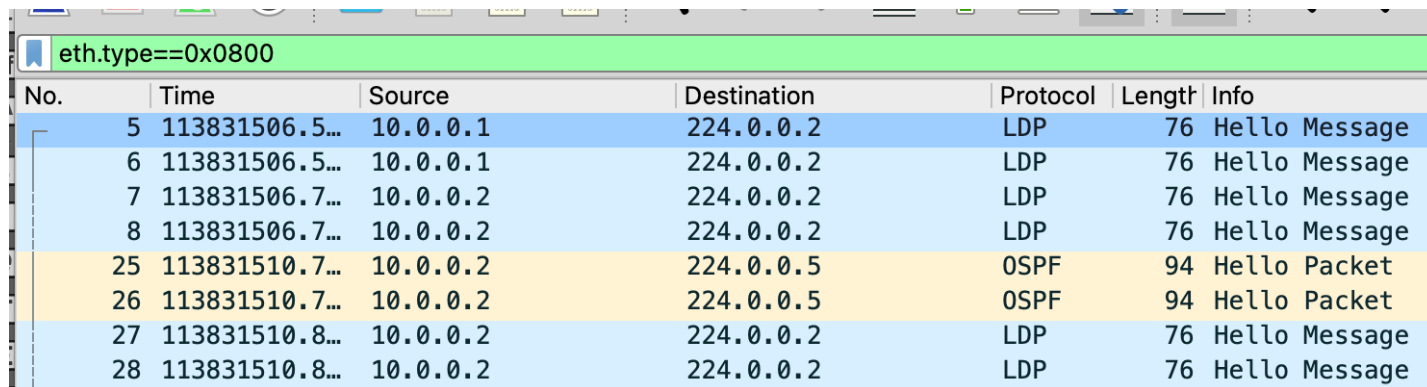
Host Name . . . . . : beatyou
Primary Dns Suffix . . . . . :
Node Type . . . . . : Unknown
IP Routing Enabled. . . . . : No
WINS Proxy Enabled. . . . . : No
DNS Suffix Search List. . . . . : na.dl.cox.net

Ethernet adapter Local Area Connection:

    Connection-specific DNS Suffix  . : na.dl.cox.net
    Description . . . . . : UIA Rhine II Fast Ethernet Adapter
    Physical Address. . . . . : 00-50-2C-A5-F5-73
    Dhcp Enabled. . . . . : Yes
    Autoconfiguration Enabled . . . . : Yes
    IP Address. . . . . : 192.168.1.30
    Subnet Mask . . . . . : 255.255.255.0
    Default Gateway . . . . . : 192.168.1.2
    DHCP Server . . . . . : 192.168.1.2
    DNS Servers . . . . . : 68.1.208.30
                           68.107.202.25
                           68.1.18.25
    Lease Obtained. . . . . : Monday, November 07, 2005 1:20:59 AM
```

Ethernet and Wireshark

- Displayfilter list at <https://www.wireshark.org/docs/dfref/e/eth.html>
- Most relevant:
 - eth.addr
 - eth.dst
 - eth.src
 - eth.type



No.	Time	Source	Destination	Protocol	Length	Info
5	113831506.5...	10.0.0.1	224.0.0.2	LDP	76	Hello Message
6	113831506.5...	10.0.0.1	224.0.0.2	LDP	76	Hello Message
7	113831506.7...	10.0.0.2	224.0.0.2	LDP	76	Hello Message
8	113831506.7...	10.0.0.2	224.0.0.2	LDP	76	Hello Message
25	113831510.7...	10.0.0.2	224.0.0.5	OSPF	94	Hello Packet
26	113831510.7...	10.0.0.2	224.0.0.5	OSPF	94	Hello Packet
27	113831510.8...	10.0.0.2	224.0.0.2	LDP	76	Hello Message
28	113831510.8...	10.0.0.2	224.0.0.2	LDP	76	Hello Message

Ethernet - Hardware

Switch



- A switch is an OSI Layer 2 device.
- Number of ports varies(4 - > 500, depending on the number of HE)
- For each frame that arrives at the switch, the destination address is examined
- Switches "learn" the MAC addresses of the connected hosts based on the packets (frames) sent to a port.
- Packets with an "unknown" destination (MAC) address are emitted on all ports (except the source port) (flooding). The same applies to the broadcast address.
- If you connect hosts via switches, the entire bandwidth of the network is available to them for communication.
- In contrast to normal shared LANs, switches thus realize dedicated connections between hubs or hosts and can thus greatly increase throughput in the network.

Switching techniques

Two technologies have established themselves for switches:

- Cut-through
examine only the first bytes of a packet (frames) to learn the source and destination addresses. They then forward the package directly without reading and checking the rest. As a result, invalid or defective packets can pass through the switch, but the delay time of this type is very short. A modification "Modified Cut Through" waits for the collision window (64 bytes) before a frame is forwarded.
- Store-and-forward
unlike the cut-through models, examine the entire package. To do this, they are cached in the switch and checked to see if they are invalid or defective. These are discarded, all others are forwarded to the corresponding port according to the destination address. These types of switches detect a large number of errors in network traffic and can thus reduce the network load. On the other hand, by using the method of caching and completely analyzing the packets, they introduce a delay that can cancel out this effect.

Switching internals

- Switches store the mapping between mac-address and port in the CAM (content addressable memory)
- Multiple addresses per port possible
- Number of entries in the CAM is vendor dependent
- Aging of entries (common def: 1800 sec) prevents overflow

```
switch1#show mac address-table
Mac Address Table
-----
```

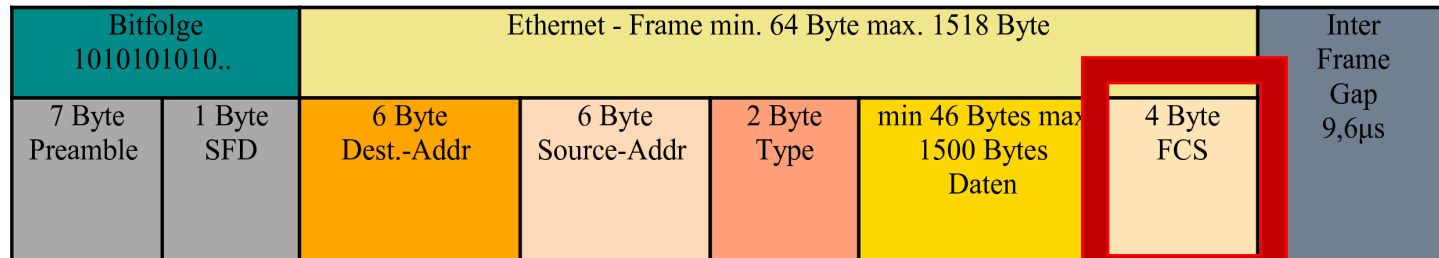
Vlan	Mac Address	Type	Ports
All	0011.5ccc.5c00	STATIC	CPU
All	0100.0ccc.cccc	STATIC	CPU
All	0100.0ccc.cccd	STATIC	CPU
All	0100.0cdd.dddd	STATIC	CPU
1	0009.5b44.9d2c	DYNAMIC	Fa0/1
1	000f.66e3.352b	DYNAMIC	Fa0/1
1	0012.8015.c940	DYNAMIC	Fa0/24
1	0012.8015.c941	DYNAMIC	Fa0/24
1	001a.adb3.bef7	DYNAMIC	Fa0/1
1	0025.2266.d104	DYNAMIC	Fa0/1
1	0026.b865.313e	DYNAMIC	Fa0/1
1	64a7.6973.8e4d	DYNAMIC	Fa0/1
1	6c71.d976.fce7	DYNAMIC	Fa0/1
1	74f6.12d4.1e1c	DYNAMIC	Fa0/1
1	a477.3344.98b6	DYNAMIC	Fa0/1

- Switches provide more techniques like VLAN, link aggregation (aka channel or bond), security, traffic steering
- Depends heavily on vendor and price

Ethernet - Error management

Error management

- Ethernet provides error detection, **no** error correction
- Detection of transmission errors by using redundancy.
- In practice, the cyclic redundancy check (CRC) has prevailed.
- FCS field is used to transfer the information
- Corrupted network frames are discarded
- No information to the upper layers



Cyclic Redundancy Check

- The CRC procedure is only designed for the detection of random errors, it is not suitable to ensure the integrity of the data.
 - The cyclic redundancy check is based on the representation of a bit chain as a polynomial with the coefficients 0 and 1.
 - Modulo 2 is used for the calculation with the polynomials. There is no carryover in addition and no borrowing in subtraction.
- CRC adds redundancy to the data (expanding the data without new information)
- Simplified process:
 - Sender adds the FCS to the frame (hence has to calculate it itself)
 - Receiver extracts FCS and calculates again
 - Match = packet ok
 - No Match = packet broken

Cyclic Redundancy Check in detail

- A cyclic binary code is a so-called linear code with m linearly independent **generator words** $g_1 \dots g_m$, which are derived from a **generator pattern** by position shifting.
- This generator pattern as well as the code word can be written as polynomials in u , whereby a multiplication by u corresponds to a position shift of the generator pattern (hardware realization as a **shift register**).
- For polynomial arithmetic, **Modulo 2** is used, that is, the overflow (carry) is not considered. Addition and subtraction are identical to the exclusive Or (*XOR*).

General representation of a generator polynomial:

$$G(u) = a_k * u^k \oplus \dots \oplus a_1 * u \oplus a_0$$

(\oplus is used as the operator character in the following. $*$ is used as the operator for multiplication.)

The degree of polynomial is k , the number of coefficients is $k + 1$, the coefficients are binary ($a_i = 0$ or 1)

Math behind Cyclic Redundancy Check

Code word (or check value) is represented as a **polynomial** as follows:

$$C(u) = c_1 * u^{n-1} \oplus \dots \oplus c_{n-1} * u^1 \oplus c_n * u^0$$

The degree of the polynomial here is $n - 1$, the number of coefficients is n , c_i are the bits of the code word.

The transferred code word polynomial is now formed in such a way that it contains the factor $G(u)$:

$$C(u) = G(u) * Q(u)$$

The **mod 2 calculation** is used. As a result, the following applies:

$$C(u)/G(u) = Q(u)$$

without rest. Therefore:

$$C(u) \bmod G(u) = 0$$

Cyclic Redundancy Check Example I

Short reminder: The degree of polynomial is k , the number of coefficients is $k + 1$, the coefficients are binary ($a_i = 0$ or 1)

$K = 3$, so number of coefficients is $3+1 = 4$, we choose 1011

The generator polynomial of **generator pattern 1011** is represented as follows:

$$G(u) = 1 * u^3 \oplus 0 * u^2 \oplus 1 * u^1 \oplus 1 * u^0 = u^3 \oplus u \oplus 1$$

Transmission data: **1 0 0 1** ($m = 4$ data bits)

Cyclic (7, 4) - Linear code: $m = 4$ (data bits), $k = 3$ (test bits),
total $n = m + k = 7$ bits, generator pattern = 1011 ($k + 1$ digits).

Cyclic Redundancy Check Example II

The divisor has 4 bits (because we choose a 3 degree polynomial), so append 3 (= degree) zero bits to the input pattern

Sender-side encoding:

1 0 0 1 0 0 0 / 1 0 1 1

1 0 1 1

0 0 1 0 0 0

1 0 1 1 0

1 1 0 → division remainder
→ CRC code

Code word to be sent:

1 0 0 1 1 1 0

(Transmission data + CRC code)

Receiver-side verification:

1 0 0 1 1 1 0 / 1 0 1 1

1 0 1 1

0 0 1 0 1 1

1 0 1 1 0

0 0 0 0 0 → division remainder = 0?

→ if division remainder = 0, then
transmission (probably) correct!

Standard polynomials for CRC

CRC-CCITT (CRC-4): $u^4 \oplus u \oplus 1$

Bluetooth: $u^5 \oplus u^4 \oplus u^2 \oplus 1$

CRC-12: $u^{12} \oplus u^{11} \oplus u^3 \oplus u^2 \oplus u \oplus 1$

CAN-CRC: $u^{15} \oplus u^{14} \oplus u^{10} \oplus u^8 \oplus u^7 \oplus u^4 \oplus u^3 \oplus 1$

IBM-CRC-16: $u^{16} \oplus u^{15} \oplus u^2 \oplus 1$

CRC-CCITT: $u^{16} \oplus u^{12} \oplus u^5 \oplus 1$

Ethernet-CRC-32: $u^{32} \oplus u^{26} \oplus u^{23} \oplus u^{22} \oplus u^{16} \oplus u^{12} \oplus u^{11} \oplus u^{10} \oplus u^8 \oplus u^7 \oplus u^5 \oplus u^4 \oplus u^2 \oplus u \oplus 1$

- The polynomials used are the result of extensive mathematical and empirical analyses.
- When choosing the generator pattern, highest and lowest bits must always be equal to 1.
- All burst errors of length $l \leq k$ are detected.

Ethernet - ARP

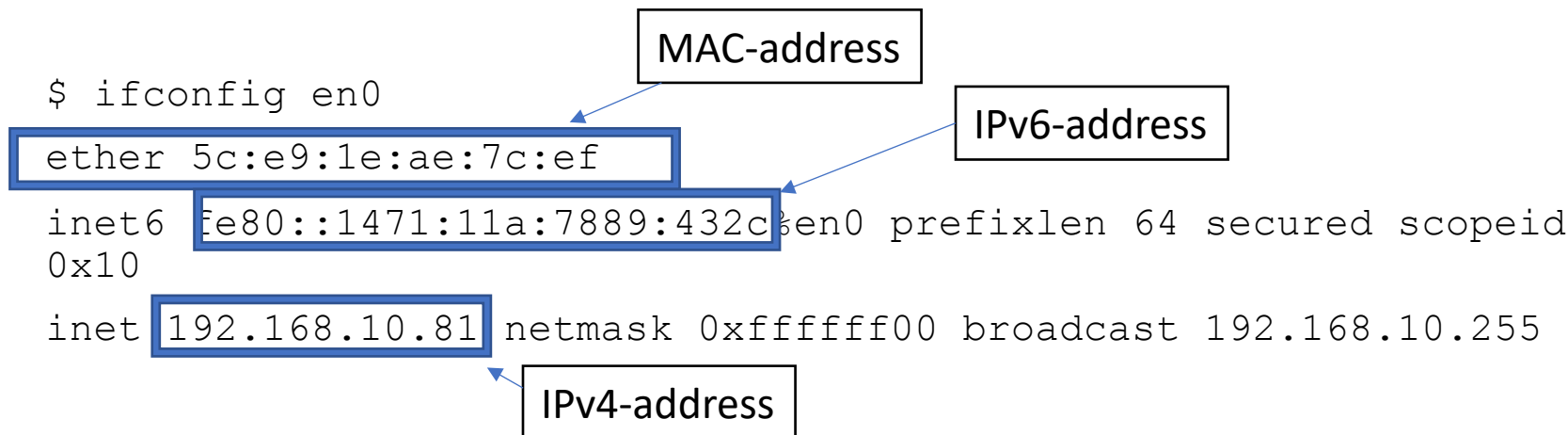
Addressing in networks

In a LAN, each network interface of each node has a **MAC Address** and (at least one) **IP address**

```
$ ifconfig en0
ether 5c:e9:1e:ae:7c:ef
inet6 fe80::1471:11a:7889:432c%en0 prefixlen 64 secured scopeid
0x10
inet 192.168.10.81 netmask 0xffffffff broadcast 192.168.10.255
```

Addressing in networks

In a LAN, each network interface of each node has a **MAC Address** and (at least one) **IP address**



We need these addresses to communicate with other nodes in the network

But how do we get the information of the other nodes?

- Dynamically discovers the mapping between a layer 3 (protocol) and a layer 2 (hardware) address
- Uses specialised request response protocol named ARP

- > Frame 26: 60 bytes on wire (480 bits), 60 bytes captured (480 bits) on interface en0, id 0
- ✓ Ethernet II, Src: Raspberr_a1:c6:18 (b8:27:eb:a1:c6:18), Dst: Broadcast (ff:ff:ff:ff:ff:ff)
 - > Destination: Broadcast (ff:ff:ff:ff:ff:ff)
 - > Source: Raspberr_a1:c6:18 (b8:27:eb:a1:c6:18)
 - Type: ARP (0x0806)
 - Padding: 00
- ✓ Address Resolution Protocol (request)
 - Hardware type: Ethernet (1)
 - Protocol type: IPv4 (0x0800)
 - Hardware size: 6
 - Protocol size: 4
 - Opcode: request (1)
 - Sender MAC address: Raspberr_a1:c6:18 (b8:27:eb:a1:c6:18)
 - Sender IP address: 192.168.10.76
 - Target MAC address: 00:00:00_00:00:00 (00:00:00:00:00:00)
 - Target IP address: 192.168.10.28

- Dynamically discovers the mapping between a layer 3 (protocol) and a layer 2 (hardware) address
- Uses specialised request response protocol named ARP
- If a computer A wants to send data to computer B on the same network, from which it initially only knows its IP address, it sends an **ARP broadcast**

- > Frame 26: 60 bytes on wire (480 bits), 60 bytes captured (480 bits) on interface en0, id 0
- ✓ Ethernet II, Src: Raspberr_a1:c6:18 (b8:27:eb:a1:c6:18), Dst: Broadcast (ff:ff:ff:ff:ff:ff)
 - > Destination: Broadcast (ff:ff:ff:ff:ff:ff)
 - > Source: Raspberr_a1:c6:18 (b8:27:eb:a1:c6:18)
 - Type: ARP (0x0806)
 - Padding: 00
- ✓ Address Resolution Protocol (request)
 - Hardware type: Ethernet (1)
 - Protocol type: IPv4 (0x0800)
 - Hardware size: 6
 - Protocol size: 4
 - Opcode: request (1)
 - Sender MAC address: Raspberr_a1:c6:18 (b8:27:eb:a1:c6:18)
 - Sender IP address: 192.168.10.76
 - Target MAC address: 00:00:00_00:00:00 (00:00:00:00:00:00)
 - Target IP address: 192.168.10.28

- Dynamically discovers the mapping between a layer 3 (protocol) and a layer 2 (hardware) address
- Uses specialised request response protocol named ARP
- If a computer A wants to send data to computer B on the same network, from which it initially only knows its IP address, it sends an ARP broadcast
- ARP packet contains the so-called target IP address of B.

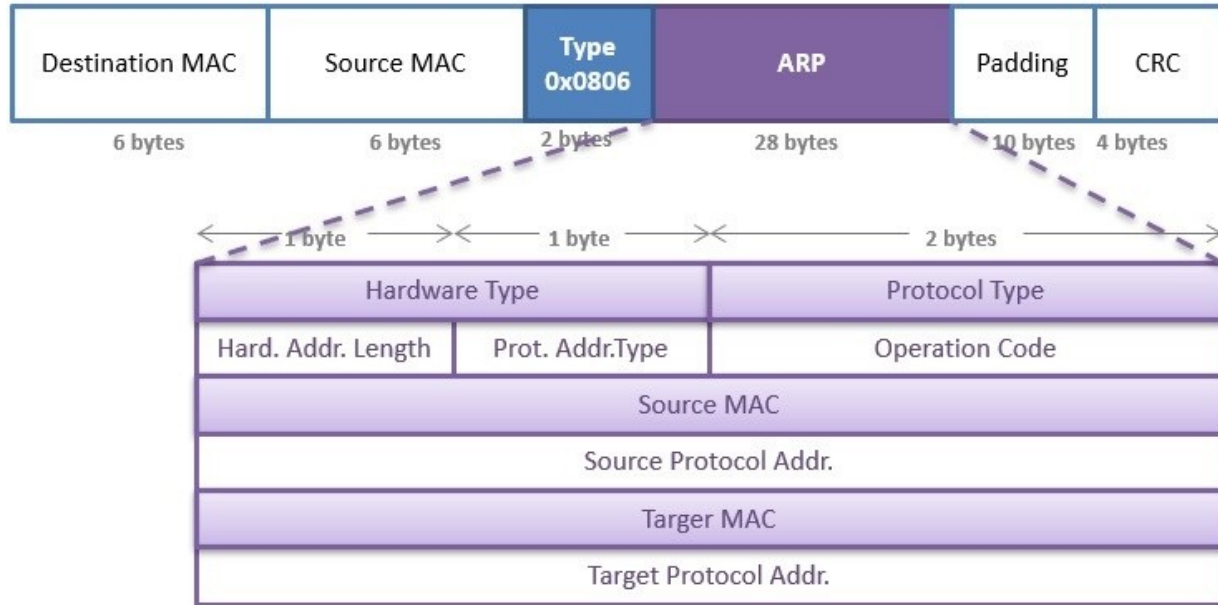
- > Frame 26: 60 bytes on wire (480 bits), 60 bytes captured (480 bits) on interface en0, id 0
- ✓ Ethernet II, Src: Raspberr_a1:c6:18 (b8:27:eb:a1:c6:18), Dst: Broadcast (ff:ff:ff:ff:ff:ff)
 - > Destination: Broadcast (ff:ff:ff:ff:ff:ff)
 - > Source: Raspberr_a1:c6:18 (b8:27:eb:a1:c6:18)
 - Type: ARP (0x0806)
 - Padding: 00
- ✓ Address Resolution Protocol (request)
 - Hardware type: Ethernet (1)
 - Protocol type: IPv4 (0x0800)
 - Hardware size: 6
 - Protocol size: 4
 - Opcode: request (1)
 - Sender MAC address: Raspberr_a1:c6:18 (b8:27:eb:a1:c6:18)
 - Sender IP address: 192.168.10.76
 - Target MAC address: 00:00:00:00:00:00 (00:00:00:00:00:00)
 - Target IP address: 192.168.10.28

- Dynamically discovers the mapping between a layer 3 (protocol) and a layer 2 (hardware) address
- Uses specialised protocol named ARP
- If a computer A wants to send data to computer B on the same network, from which it initially only knows its IP address, it sends an **ARP broadcast**
- **ARP packet contains the so-called target** IP address of B.
- All computers on the network receive this message and compare the Internet address with their own. Only computer B recognizes its address and sends a response with its physical address to the source of the **ARP request**.
- Computer A receives this response (**ARP reply**) and enters the pair of Internet and physical address into a so-called **ARP cache**. He now uses this information for further communication.

- > Frame 74: 42 bytes on wire (336 bits), 42 bytes captured (336 bits) on interface en0, id 0
- ✓ Ethernet II, Src: Apple_ae:7c:ef (5c:e9:1e:ae:7c:ef), Dst: be:53:a8:a0:01:82 (be:53:a8:a0:01:82)
 - > Destination: be:53:a8:a0:01:82 (be:53:a8:a0:01:82)
 - > Source: Apple_ae:7c:ef (5c:e9:1e:ae:7c:ef)
 - Type: ARP (0x0806)
- ✓ Address Resolution Protocol (reply)
 - Hardware type: Ethernet (1)
 - Protocol type: IPv4 (0x0800)
 - Hardware size: 6
 - Protocol size: 4
 - Opcode: reply (2)
 - Sender MAC address: Apple_ae:7c:ef (5c:e9:1e:ae:7c:ef)
 - Sender IP address: 192.168.10.81
 - Target MAC address: be:53:a8:a0:01:82 (be:53:a8:a0:01:82)
 - Target IP address: 192.168.10.75

Address Resolution Protocol

- Defined in RFC 826
- Transferred in the data section of the Ethernet frame



Address Resolution Protocol (reply)

Hardware type: Ethernet (1)

Protocol type: IPv4 (0x0800)

Hardware size: 6

Protocol size: 4

Opcode: reply (2)

Sender MAC address: Apple_ae:7c:ef (5c:e9:1e:ae:7c:ef)

Sender IP address: 192.168.10.81

Target MAC address: be:53:a8:a0:01:82 (be:53:a8:a0:01:82)

Target IP address: 192.168.10.75

RFC 826

- Defines details of ARP
- Every ARP packet is 28 byte long
 - Padding needed for min EFS
 - Padding might not be shown in WS

```
Frame Length: 42 bytes (336 bits)
Capture Length: 42 bytes (336 bits)
[Frame is marked: False]
[Frame is ignored: False]
[Protocols in frame: eth:ethertype:arp]
[Coloring Rule Name: ARP]
[Coloring Rule String: arp]
```

```
> Ethernet II, Src: AVMAudio_8e:23:a4 (5c:49:7
```

```
▼ Address Resolution Protocol (request)
```

```
Hardware type: Ethernet (1)
Protocol type: IPv4 (0x0800)
```

Let there exist machines X and Y that are on the same 10Mbit Ethernet cable. They have Ethernet address EA(X) and EA(Y) and DOD Internet addresses IPA(X) and IPA(Y). Let the Ethernet type of Internet be ET(IP). Machine X has just been started, and sooner or later wants to send an Internet packet to machine Y on the same cable. X knows that it wants to send to IPA(Y) and tells the hardware driver (here an Ethernet driver) IPA(Y). The driver consults the Address Resolution module to convert <ET(IP), IPA(Y)> into a 48.bit Ethernet address, but because X was just started, it does not have this information. It throws the Internet packet away and instead creates an ADDRESS RESOLUTION packet with

```
(ar$hrd) = ares_hrd$Ethernet
(ar$pro) = ET(IP)
(ar$hln) = length(EA(X))
(ar$pln) = length(IPA(X))
(ar$op) = ares_op$REQUEST
(ar$sha) = EA(X)
(ar$spa) = IPA(X)
(ar$tha) = don't care
(ar$tpa) = IPA(Y)
```

and broadcasts this packet to everybody on the cable.

Machine Y gets this packet, and determines that it understands the hardware type (Ethernet), that it speaks the indicated protocol (Internet) and that the packet is for it ((ar\$tpa)=IPA(Y)). It enters (probably replacing any existing entry) the information that <ET(IP), IPA(X)> maps to EA(X). It then notices that it is a request, so it swaps fields, putting EA(Y) in the new sender Ethernet address field (ar\$sha), sets the opcode to reply, and sends the packet directly (not broadcast) to EA(X). At this point Y knows how to send to X, but X still doesn't know how to send to Y.

Machine X gets the reply packet from Y, forms the map from <ET(IP), IPA(Y)> to EA(Y), notices the packet is a reply and throws it away. The next time X's Internet module tries to send a packet to Y on the Ethernet, the translation will succeed, and the packet will (hopefully) arrive. If Y's Internet module then wants to talk to X, this will also succeed since Y has remembered the information from X's request for Address Resolution.

ARP Protocol additions

- An ARP request is sent as an Ethernet broadcast (i.e. to ff:ff:ff:ff:ff:ff), since the "correct" Ethernet destination address is still unknown.
- The ARP response can be sent specifically to the requester, as their Ethernet address is known from the request.
- The Reverse Address Resolution Protocol (RARP) (RFCs 903, 2390) performs the opposite conversion of physical addresses into Internet addresses.
- ARP is stateless, if a system sends a response uncalled into the network, a receiving system adds this mapping in its own arp cache

ARP and Wireshark

- Displayfilter: arp
- Details in <https://gitlab.com/wireshark/wireshark/-/blob/master/epan/dissectors/packet-arp.c>

```
switch (ar_op) {
    case ARPOP_REQUEST:
        if (is_gratuitous)
            col_add_fstr(pinfo->cinfo, COL_INFO, "Gratuitous ARP for %s (Request)", tpa_str);
        else
            col_add_fstr(pinfo->cinfo, COL_INFO, "Who has %s? Tell %s", tpa_str, spa_str);
        break;
    case ARPOP_REPLY:
        if (is_gratuitous)
            col_add_fstr(pinfo->cinfo, COL_INFO, "Gratuitous ARP for %s (Reply)", spa_str);
        else
            col_add_fstr(pinfo->cinfo, COL_INFO, "%s is at %s",
                        spa_str,
                        tvb_arphrdaddr_to_str(pinfo->pool, tvb, sha_offset, ar_hln, ar_hrd));
        break;
```

OS commands II

Show (set or delete) learned mac-addresses of communication partners in the network

- arp

```
root@kali:~# arp -v
Address          HWtype  HWaddress      Flags Mask    Iface
machine1         ether    08:00:27:ad:87:b3  C             eth0
machine2         ether    08:00:27:27:d6:c7  C             eth0
10.0.2.3          ether    08:00:27:e5:fd:ed  C             eth0
_gateway         ether    52:54:00:12:35:00  C             eth0
Entries: 4      Skipped: 0      Found: 4
root@kali:~# arp -n
Address          HWtype  HWaddress      Flags Mask    Iface
10.0.2.4          ether    08:00:27:ad:87:b3  C             eth0
10.0.2.5          ether    08:00:27:27:d6:c7  C             eth0
10.0.2.3          ether    08:00:27:e5:fd:ed  C             eth0
10.0.2.1          ether    52:54:00:12:35:00  C             eth0
```

Ethernet - VLAN

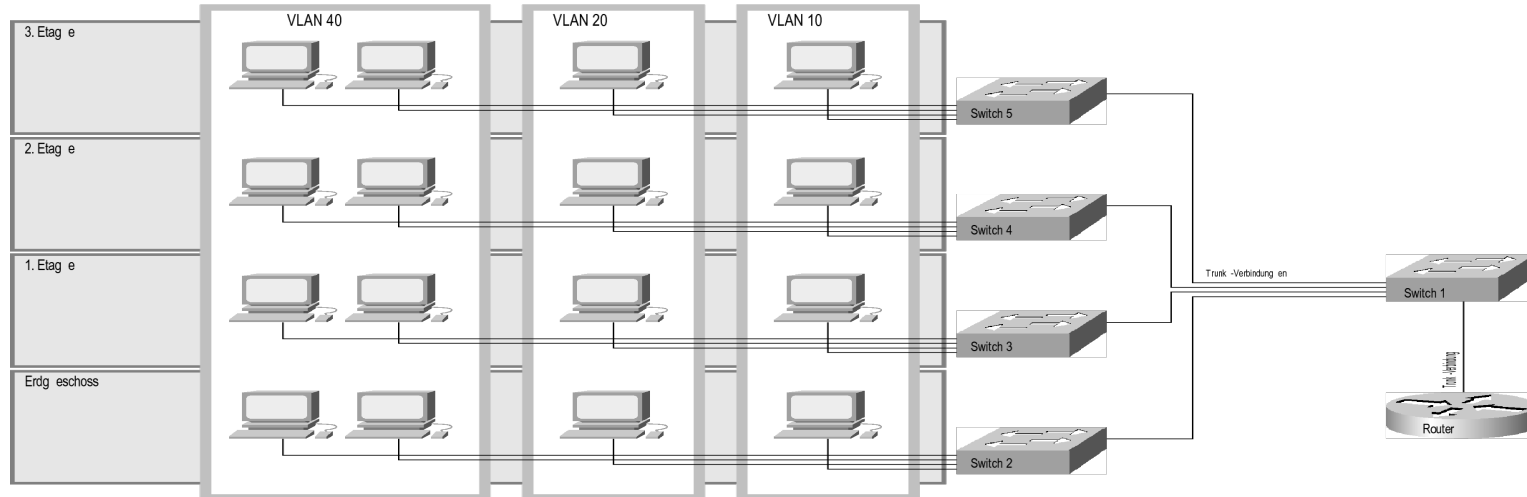
Does a switch solve all problems?

- A switch creates single collision domains between switch and host
- But a switch does not shrink a broadcast domain
- Huge number of hosts may result in a bunch of packet flooding
- No security between the hosts

- How can we separate domains inside a network?
 - E.g. in a company: Sales, development, CRM, management, HR, assembly
 - E.g. in a university: departments, HR, students, external, research

Separation

- Different subnets
- New cabling and new hardware
- Virtual separation based on VLANs



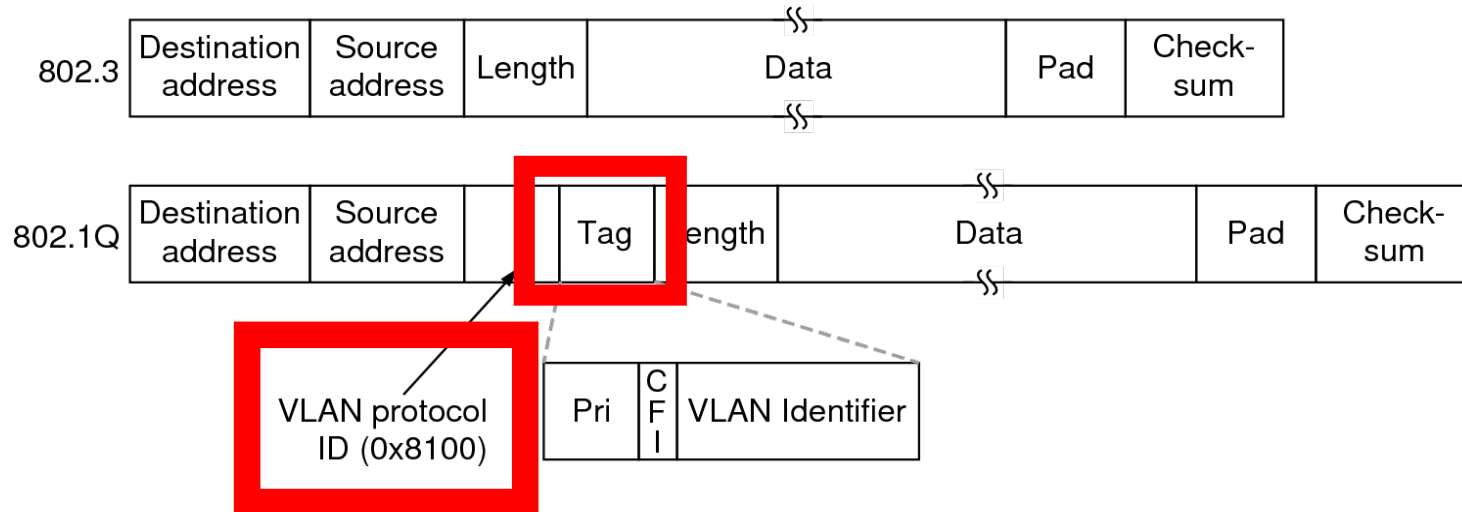
Benefits of VLANs

A virtual LAN (VLAN) is a **logical segment** in a **"switched" network**. VLANs allow logical segmentation of one or more LANs completely independent of the physical structure.

- **Broadcast Control:** By forming VLANs, the broadcast traffic in the individual segments can be specifically limited. Broadcast messages from computers from working groups that have nothing to do with each other no longer burden other segments.
- **Workgroup- and Network Security:** If a VLAN does not have a router, communication between users outside the VLAN and users inside is not possible. This extreme level of security may be desirable for certain areas.
- **Performance:** Dividing a LAN into several smaller segments can increase performance. The formation of several broadcast domains also reduces the number of broadcasts which the individual machines in the segment have to react to.
- **Network management:** VLAN formation simplifies the relocation of employees or computers, as only the computer at the new location has to be assigned to the corresponding VLAN.

VLAN based on IEEE 802.1Q

- To form the VLANs, the **LAN coupling elements** (switches) (not the end devices) must exchange information about the VLAN affiliation of the frames.
- **Frame tagging** has become the standard **according to IEEE 802.1Q**.



802.1Q Tagging Format

A VLAN tag consists of the following four fields:

- **Tag Protocol ID (16 bits)**: This field identifies the VLAN tag. Its value must be 0x8100. This value ensures that these two bytes are not interpreted as a type or length field.
- **Priority Code Point (3 bits)**: These three bits (usually abbreviated as PCP) enable quality of service on the link securing layer of the ISO-OSI model (layer 2). The value range is 0-7, with values 6 and 7 reserved for control data (e.B Resource ReSerVation Protocol (RSVP)). With a value of 0, a data packet is not given preferential treatment, a value of 5 gives the highest priority (for payloads).
- **DEI – Drop Eligible Indicator (1 bit)** formerly Canonical Frame Indicator (CFI): Can be used separately or in conjunction with PCP to indicate that frames can be discarded in the presence of congestion.
- **VLAN ID (12 bits)**: The value of this field specifies the affiliation of an Ethernet frame to a VLAN (packets with the same VLAN ID belong to the same VLAN). The value range is between 0 and 4095, whereby only values from 1 to 4094 can be assigned as IDs. A value of 0 does not indicate access to a VLAN, the value 4095 is reserved for experimental purposes and may not be used.

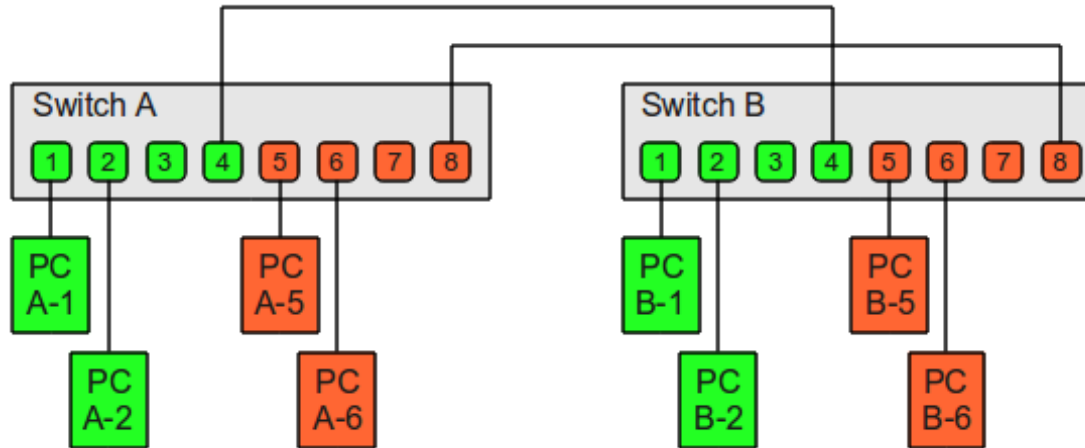
IEEE 802.1Q and Wireshark

■ Displayfilter: vlan

```
> Frame 1: 64 bytes on wire (512 bits), 64 bytes captured (512 bits) on interface 0  
> Ethernet II, Src: Cisco_ea:b8:c1 (00:19:06:ea:b8:c1), Dst: Broadcast  
✓ 802.1Q Virtual LAN, PRI: 0, DEI: 0, ID: 123  
    000. .... = Priority: Best Effort (default) (0)  
    ...0 .... = DEI: Ineligible  
    .... 0000 0111 1011 = ID: 123  
    Type: ARP (0x0806)  
    Padding: 00000000000000000000000000000000  
    Trailer: 00000000  
> Address Resolution Protocol (reply/gratuitous ARP)
```

Port-based VLAN (Layer 1 VLAN, Static VLAN)

- The port-based VLAN is the simplest form of virtual LAN formation
- The assignment of physical ports to the different VLANs is carried out by configuring the switches.

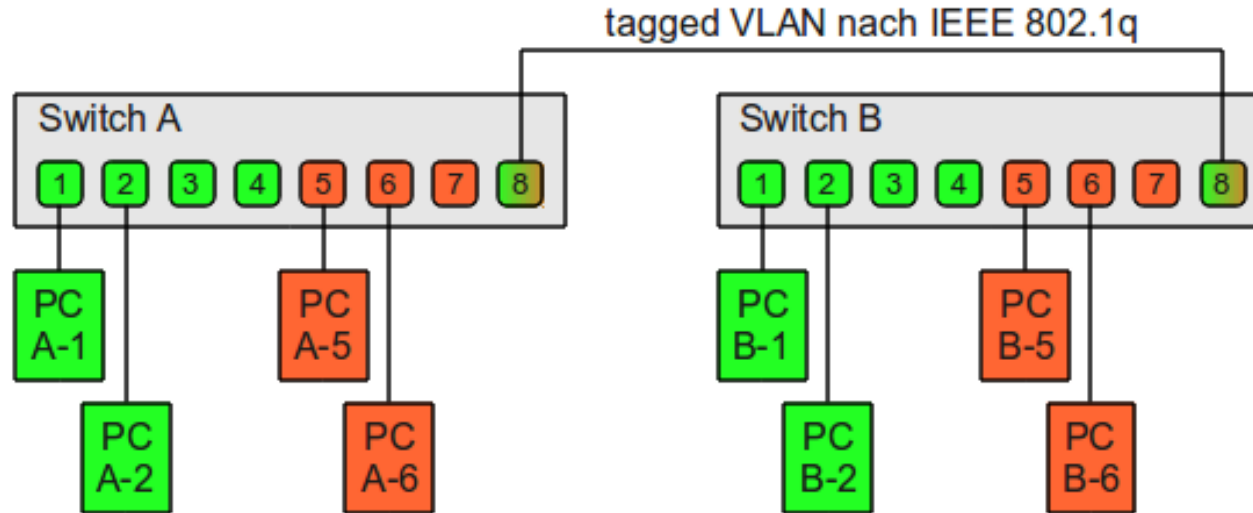


Access port: Switch port assigned to a specific VID, 802.1Q information are removed before sending the frame to the device

www.thomas-krenn.de

VLAN tagging

- Improvement installation with a trunk or tagged port



Trunk port (Tagged port):
Switch port, which does not
remove the 802.1Q
information before
forwarding the frame, used
for uplinks and switch-to-
switch-links

www.thomas-krenn.de

Further VLAN assignments

- Port-based vlan assignment is simple, but less powerful
- Different implementations provide more flexibility

MAC-based VLAN

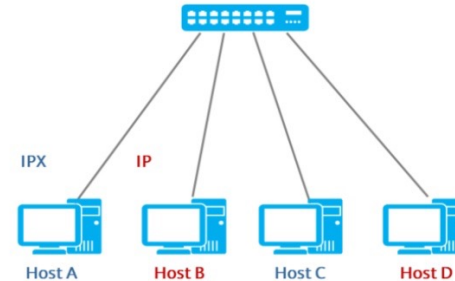
Protocol-based VLAN

- Save Configuration
- Configuration
- Monitor
- System
- Green Ethernet
- Ports
- DHCP
- Security
- LACP
- Loop Protection
- Spanning Tree
- MVR
- IPMC
- LLDP
- PoE
- MAC Table
- VLANs
- VCL
- sFlow

MAC Address Table

Start from VLAN and MAC address with

			Port Members								
Type	VLAN	MAC Address	CPU	1	2	3	4	5	6	7	8
Dynamic	1	00-E0-4C-36-01-44			✓						
Static	1	33-33-00-00-00-01	✓	✓	✓	✓	✓	✓	✓	✓	✓
Static	1	33-33-00-00-00-02	✓	✓	✓	✓	✓	✓	✓	✓	✓
Static	1	33-33-FF-74-59-17	✓	✓	✓	✓	✓	✓	✓	✓	✓
Dynamic	1	A8-2B-B5-4C-C1-1A						✓			
Static	1	A8-2B-B5-74-59-17	✓								
Static	1	FF-FF-FF-FF-FF-FF	✓	✓	✓	✓	✓	✓	✓	✓	✓
Dynamic	3	00-00-11-00-00-11			✓						



VLAN Table

Protocol	VLAN
IPX	VLAN5
IP	VLAN10
.....

- Multiple VLAN possible
- 802.1Q in 802.1Q
- Example usage:
Customer network through service network

```
> Frame 1: 64 bytes on wire (512 bits), 64 bytes captured (512 bits) on interface 0
> Ethernet II, Src: Cisco_ea:b8:c1 (00:19:06:ea:b8:c1), Dst: Broadcast (ff:ff:ff:ff:ff:ff)
✓ 802.1Q Virtual LAN, PRI: 0, DEI: 0, ID: 123
    000. .... .... = Priority: Best Effort (default) (0)
    ...0 .... .... = DEI: Ineligible
    .... 0000 0111 1011 = ID: 123
    Type: ARP (0x0806)
    Padding: 00000000000000000000000000000000
    Trailer: 00000000
> Address Resolution Protocol (reply/gratuitous ARP)
```

