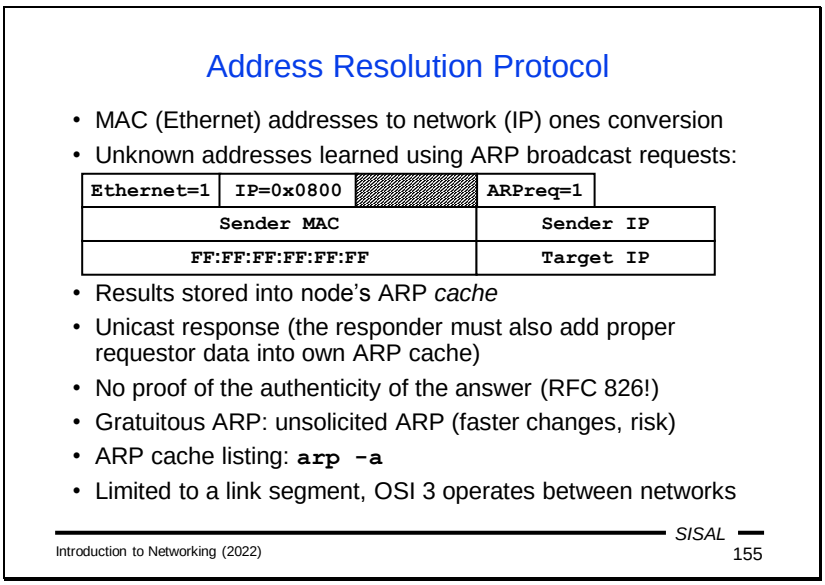



The Address Resolution Protocol is an ancillary technical protocol that represents a connection between the network and link layers. It allows nodes on a network to discover data link (MAC) addresses corresponding to specific network addresses. It's a generic protocol, it can be used for any network and data link addresses, but we'll be demonstrating its use on Ethernet and IP addresses.

We already know that the link layer gets a request from the network layer to deliver data by a directly connected segment of the network, either to the target machine or to the nearest router, i.e. a target node at the data-link layer. So it needs to correctly fill in a destination MAC address corresponding to a specific IP address in its data-link header. To find this MAC address, it sends a frame to the target node with an ARP query. However, it does not yet know the destination MAC address, so it will use the **broadcast** MAC address (FF:FF:FF:FF:FF:FF). This causes the frame to be delivered to all nodes on a given data-link network segment, but the frame will be ignored by everyone except the queried IP address holder.

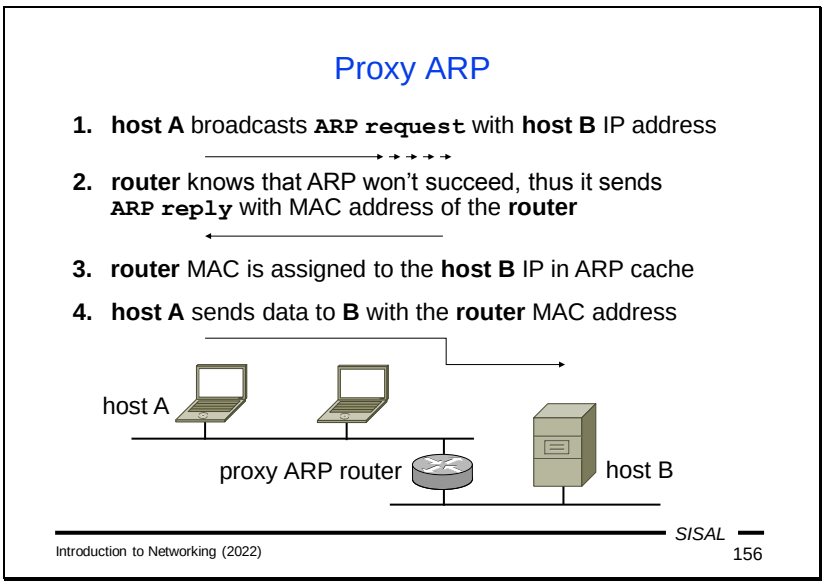
The node (i.e. an ARP server for this moment) responds to the query with a unicast ARP response containing the requested MAC address. The ARP client saves IP–to-MAC address assignments for further use in an **ARP cache**, in which a record lasts for a configurable time (of the order of minutes). The server will, however, perform the same caching operation. It is assumed that the communication that has just started will continue and the server would soon have to trace the client's address by itself. Therefore, the assignment of the client's IP and MAC address will be recorded in the server's cache, too. And that's why we can see machines in our ARP cache that we haven't communicated with – such hosts have simply sent us an ARP query. We can list the contents of the ARP cache with the **arp –a** command.

The ARP protocol (and thus the content of the ARP cache) is limited by the scope of the local data-link (OSI 2) network segment – we need to climb a layer up for communication outside this segment. Incidentally, this also means that network interface cards (NICs) with the same MAC address can be used on the same LAN if they cannot "see" each other, i.e. are separated by a router. Indeed, it is often the case that a router has only a single MAC address on all its network interfaces (although they have different IP addresses, of course).

A problem with ARP is its lack of security, i.e. that the broadcast query will reach everyone, so any node of the network can answer us. Worse, a potential attacker doesn't even have to wait for our request. An **unsolicited** ARP message (a so-called *Gratuitous ARP*) is possible, which is actually an answer without a prior query. Such messages are used e.g. for **cluster** solutions – important machines in the network can run redundantly, both sharing the same IP address. An active machine of the pair informs other machines on the network using gratuitous ARP about its MAC address as the one that they are currently expected to use for the shared IP address.

Some increase in security in certain types of networks can be achieved by **denying** certain important nodes (servers, routers) from using the ARP protocol and fixing their ARP cache content by a configuration.

Slide 156




In a complex LAN, network management can choose not to disclose details about subnets to all stations on the network, leaving the correct routing purely up to the routers. Let's imagine the simplified situation in the picture above: there are two subnets in the network, but the stations are not told this information and are sent a network mask that corresponds to the **entire** network. If host A wants to communicate with host B, according to the network mask, it believes they're on the same network, so it sends an ARP query. But as we know, a broadcast query only spreads across a data link segment, and therefore over the **subnet**, where computer A is connected, so it never reaches host B. For the entire network to work, we need to run an **ARP proxy** on the router that splits the network. The router with an ARP proxy intercepts the ARP query and, recognizing that the client would never receive a response, it sends a reply instead of the target host, returning the **router's MAC address** as the query response. The client will save this assignment to its ARP cache and use the router's MAC address for further communication with host B.

It may seem strange, but if we think about it more carefully, the client will actually behave in **exactly the same way** as if it had the right information about the netmask and knew about the router – if it did, it would also send packets for host B with the MAC address of the nearest router!

A user on the station can tell the difference by looking at the ARP cache. He will find **multiple IP** addresses with the **same MAC** address. If we have an ARP proxy in the network, this situation is expected. But if not, multiple occurrences of the same MAC address in the ARP cache may signal an attempt to attack our ARP cache using fake ARP messages.

## Data link layer (OSI 2)

- Separated into two sublayers:
  - Logical Link Control (LLC) multiplexes various network layer protocols approaching a media
  - Media Access Control (MAC) controls addressing and media access: who, when and how may send and receive data
- TCP/IP does not deal with this layer („network interface“)
- Network segment (physical network):
  - set of nodes sharing the same media
- Data link layer PDU: frame
  - format depends on the media used
  - in general: synchronization field, header (addresses, type, control data), data field and trailer (Frame Check Sequence - error detection)



For us, the ARP protocol represented a move from the network layer to the **data-link** layer, and thus a move away from TCP/IP as well.

The link layer represents a very important milestone; you could say that this is actually a step from software to hardware. Unlike some other layers of the OSI model, it performs a lot of work, and we are actually dividing it into two sub-layers:

- The upper sub-layer is called *Logical Link Control* (LLC) and is in charge of **multiplexing**. It is thus responsible for correctly storing data of various network protocols and identifying it so that the receiving data-link layer software is able to forward it to the corresponding network protocol software.
- The bottom sub-layer is called *Media Access Control* (MAC) and handles both the **addressing** and **access control** of individual nodes to physical media they share within the same line (physical) segment of the network.

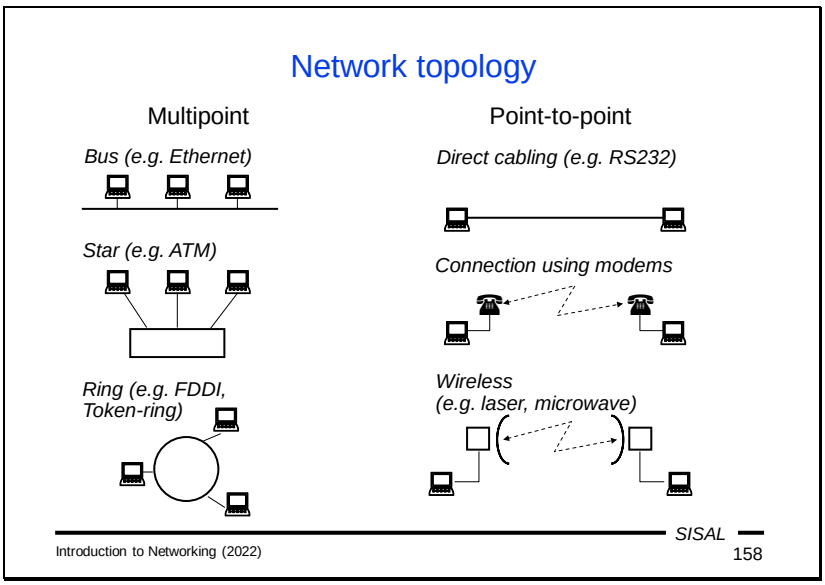
The MAC sub-layer is directly dependent on the underlying physical layer technology, and we will deal primarily with two of them, Ethernet (which now dominates cable networks) and wireless (WiFi) technology.

The data unit of a line protocol is called a **frame**, and its exact format may vary from one protocol to another, but it generally includes:

- A *Synchronization Field* – a special sequence of bits designed to "wake up" a target node, to distinguish subsequent data from “noise“. This field does not usually count as part of the actual contents of a frame.
- A *Header* – an initial part of the frame, which contains at least the MAC addresses of the recipient and the sender and LLC control information.
- *Data* (the payload) of a network protocol.

- A *footer* – a final part of the frame, which usually contains a so-called Frame Check Sequence (FCS), a value used to **check the accuracy** of delivery. In fact, the data-link layer is the last layer working with data above the physical layer, so it is desirable to check that the physical layer has transferred the data in order when it reaches the target node.

Slide 158




One of the criteria for classification of various technologies of the lower two OSI layers are their topology schemas (arrangement of nodes). A network topology can be studied from either a **physical** perspective, i.e. how the nodes are actually connected (e.g. by cable), or a **logical** one (how the nodes communicate with each other).

Technologies that can connect multiple nodes within a one network segment (**multipoint**) typically use one of the following topologies:

- A *bus* – all nodes are connected serially to the same media. They can all (more or less) simultaneously receive a signal that spreads across the medium (which is fine), but also they can all try to send a signal at the same time, which is more problematic. This situation is called a **collision**, and the protocol must address it. The physical bus topology of Ethernet networks built by coaxial cable was very popular in its day, with the great advantage that adding a new computer to the network could be done at virtually any time and anywhere, but its main disadvantage was that if the cable broke down anywhere, the **whole** network broke down.
- A star – one segment contains a central device and individual nodes are bound to that central element. This is the most common physical topology currently used by Ethernet networks. The term **structured cabling** is often used for it and stations are usually connected by a UTP cable. Logically, however, Ethernet's topology remains **bus-like**, but the bus is concentrated in that central device, where it is sufficiently physically protected, while damage-prone cabling links only the terminals, so that only the affected station is inoperable in the event of a problem.
- Circle - the individual nodes are connected to a circle. Token-ring and FDDI technologies, for example, work in this way.

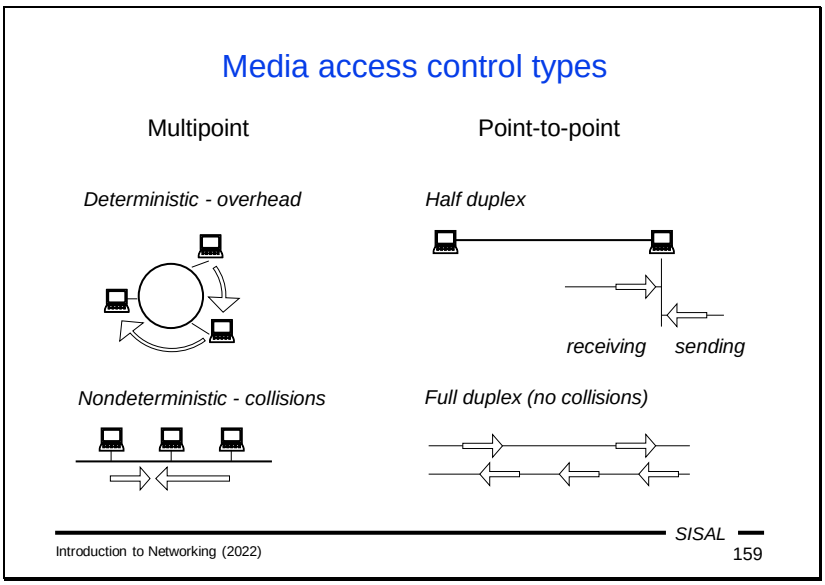
An alternative to multipoint technologies is technologies using point-to-point topology, such as RS 232. With them, only two nodes can connect to each other, so the address on the line protocol can be significantly simplified. But if we're talking about a physical point-to-point topology where we only connect two nodes together, for example, with an Ethernet cable, the same protocol as in multipoint wiring will still be used.

The point-to-point range of technology can be extended by replacing part of the line with another technology. A possible solution is to hook up a device (modem) to your computer that modulates data traffic so that it can be transmitted using a telephone connection.

The latest group of technologies using the point-to-point model are wireless, high-range links based on laser or radio waves. However, what one now typically envisions as a "wireless network" is another technology that actually works with a star topology.

*Note*: The central element of a star in a structured wiring tends to be a switch, and individual stations connect to it with a cable tucked into one socket. The drawer is commonly called a port. Be careful - don't confuse this term with the term port on the transport layer, it's a homonym! From now on, in my part of the lecture, we will use this word only in this new sense.

Slide 159



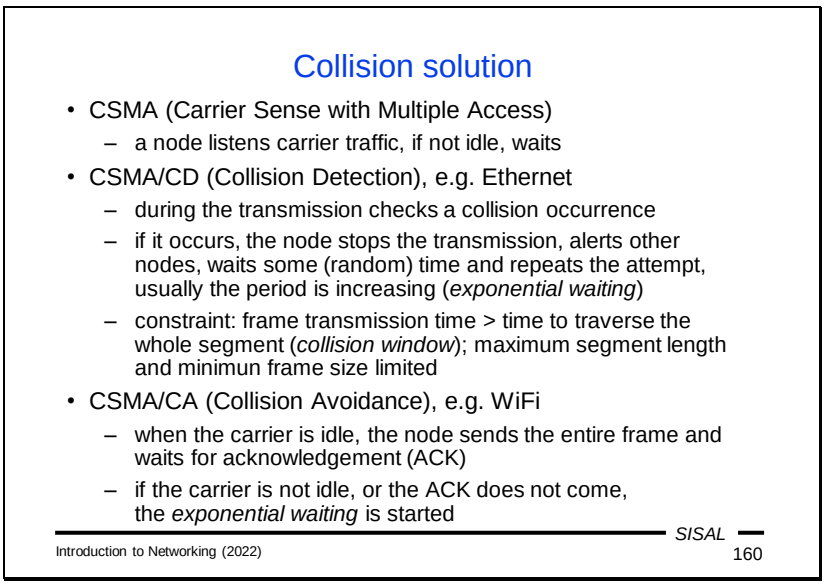

Network topology largely determines the capabilities of a given technology in terms of managing the access of nodes to a medium, so that the technology can cope with parallel transmission requirements.

Two basic approaches are used for multipoint topologies:

- In a *deterministic* means of control, there are no problems. Someone or something determines who is allowed to broadcast at any given moment. The downside is that if the "next" node has nothing to broadcast, its broadcast frequency will remain unused, increasing overhead and reducing network capacity.
    - In some networks (e.g. Token-ring), the controller role is played by a specific piece of data (a *token*) running from node to node. If a node wants to transmit, it waits for the token and sends its data over the network instead; the recipient deletes the data packet and sends a token back to the network.
    - In other technologies, there is a special control node in the network that sends a signal to other nodes when they have the right to transmit.
- In nondeterministic control (e.g. Ethernet), no one limits nodes in the broadcast, so **collisions** have to be dealt with subsequently (see the next slide).

With point-to-point topologies, we distinguish whether a node can simultaneously receive and transmit.

- If the node can't do this, it will work in the so-called *half-duplex* mode. If we plug in an Ethernet network this way, collisions will normally occur on it.
- If both nodes can handle both input and output simultaneously, we can set a *full-duplex* mode. For example, on Ethernet, this avoids collisions for a given segment, radically increasing throughput.




Networks that allow multiple access must somehow address when multiple nodes at once intend to transmit. The basic principle is that a node checks the "**carrier**," the transmission medium, for any current transmission. This step is called “carrier sense”. If the carrier is not free, the node is waiting. If the carrier is free, it can start transmitting. However, it is obvious that such a check does not guarantee that there will be no “multiple access” attempts. That's why there are different extensions that try to deal with **collisions**.

Ethernet uses the *Collision Detect* method. During broadcasting, a node checks the carrier at the same time, so it is able to **detect a collision**. If it occurs, the node will stop the broadcast, alert other stations on the network to the collision and start waiting for a new attempt. The wait time must be chosen at random from a certain interval to reduce the risk that the two nodes that detected the collision will wait the same amount of time and repeated transmissions will collide again. However, repeated attempts must not overwhelm the network, so the wait is **exponential**, i.e. the center of the interval from which the wait time is selected doubles with each subsequent attempt. However, a condition of the success of this method is that the frame broadcast time must be longer than the frame propagation time from one end of the segment to the other (the so-called *collision window*). If this condition is not met, it could be that a node completes broadcasting an entire frame before a conflicting frame arrives from the other end of a segment, the node does not detect a collision, does not do a repeated send attempt, and the frame is lost to most other nodes. The collision window determines both the maximum network segment length (segments that are too long must be split) and the minimum frame size (frames that are too short are padded).

WiFi uses the *Collision Avoidance* method. It uses a star topology – in the middle of a star, there is a so-called *access point* (AP), to which individual stations are connected. So any transmission at any given time is actually a point-to-point operation between the station and the AP, and confirmation of delivery can be implemented. A node detects a collision by not receiving a confirmation, in which case it begins exponential waiting.

## Ethernet

- History:
  - the first LAN attempts in Xerox company
  - standardization taken over by IEEE (Feb 1980 → IEEE 802)
  - two most common formats Ethernet II, IEEE 802.3
- Currently the top technology for LANs
  - can flexibly react to progressive HW evolution
  - can adapt to a wide range of transmission media
- Media access controlled by CSMA/CD
  - a „jam signal“ is sent by sender when a collision is detected
  - exponential waiting terminates after 16 attempts by an error
- Addresses:
  - 3 bytes prefix (producer, multicast...), 3 bytes node number
  - formerly „burned-in“ in NIC, nowadays programmable



Ethernet originated at the company Xerox, and it's a little surprising that a copier manufacturer was behind the most successful network technology. In February 1980, the IEEE (Institute of Electrical and Electronics Engineers) took over the standardization of the protocol, and it is said that this date is actually hidden behind the number 802 in the official standard identification (IEEE 802). This step caused a split in development, resulting, among other things, in a different format of the two most common versions of the protocol. In its early development phase, Ethernet had a number of competitors; IBM’s Token-ring technology was more successful for a time, but over the decades it became clear that Ethernet was best placed to keep up with hardware development.

The CSMA/CD method is used to deal with collisions on multipoint segments and half-duplex point-to-point segments, as described in the previous slide.

Ethernet addresses are 6 bytes long, of which the first three are a card manufacturer's prefix and the second three are the card's own number. The address is stored by the manufacturer into the card; formerly this was done in a fixed way, today it can be changed by software. But if no one has changed it, you can identify a manufacturer from an Ethernet address. Once upon a time, the rule existed that manufacturers guaranteed uniqueness of addresses, but that is no longer the case. It may well be that if you buy two cards from the same manufacturer, they will have the same address. Such cards can only be used in the same LAN simultaneously if they are separated by a router. Otherwise, at least one of the addresses needs to be changed. In addition to manufacturers' prefixes, there are other special prefixes for broadcasts and multicasts.

## IEEE 802.3 standards

| Standard | Year | Identification | Bandwidth | Media |
|---|---|---|---|---|
| 802.3 | 1983 | 10BASE5 | 10 Mbit/s | thick coaxial cable |
| 802.3a | 1985 | 10BASE2 | 10 Mbit/s | thin coaxial cable |
| 802.3i | 1990 | 10BASE-T | 10 Mbit/s | twisted pair (UTP) |
| 802.3j | 1993 | 10BASE-F | 10 Mbit/s | fiber optic |
| 802.3u | 1995 | 100BASE-TX,FX | 100 Mbit/s | UTP or fiber optic |
| 802.3z | 1998 | 1000BASE-X | 1 Gbit/s | fiber optic |
| 802.3ab | 1999 | 1000BASE-T | 1 Gbit/s | UTP |
| 802.3ae | 2003 | 10GBASE-SR,... | 10 Gbit/s | fiber optic |
| 802.3an | 2006 | 10GBASE-T | 10 Gbit/s | UTP |
| 802.3ba | 2010 | 100GBASE-SR | 100 Gbit/s | fiber optic |

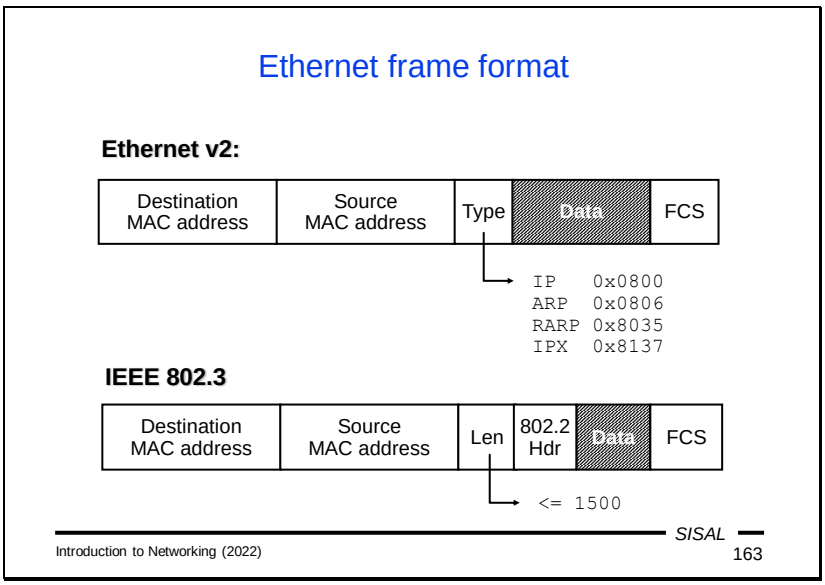
Unlike RFC, IEEE standards are licensed.



This table shows how Ethernet gradually changed as hardware improved over time. This information will not be a subject of our examination.

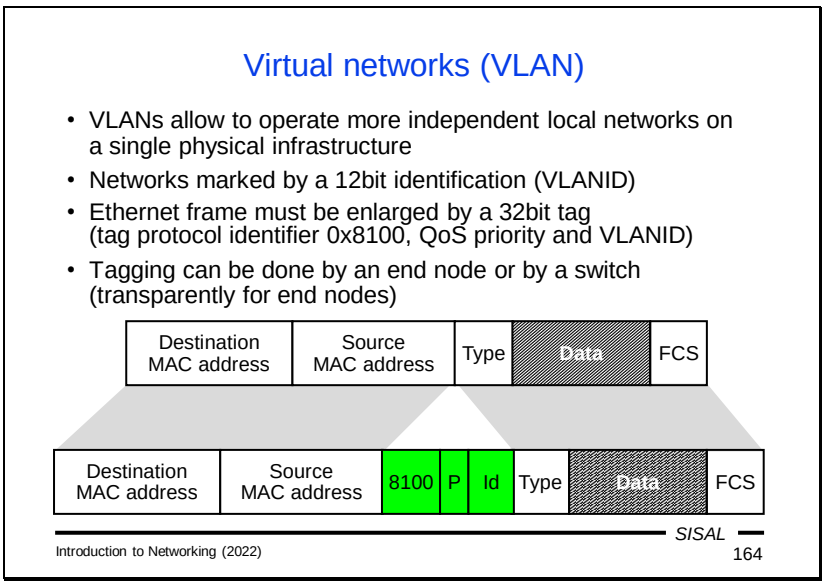
Important warning about protocol availability: IEEE standards, unlike RFCs, are **not public**. You have to buy them. Or borrow them from someone who has already bought them…

Slide 163




The long and turbulent evolution of Ethernet has also affected the structure of its frame format. While in the original concept the header contains a two-byte type of network protocol behind MAC addresses, in the IEEE version a frame length is in place and only behind it is the special added LLC header (IEEE 802.2).

Also note that, unlike in an IP header, the recipient's address precedes the sender's address here. This is so that hardware can detect as quickly as possible whether the incoming frame is intended for the node or not.

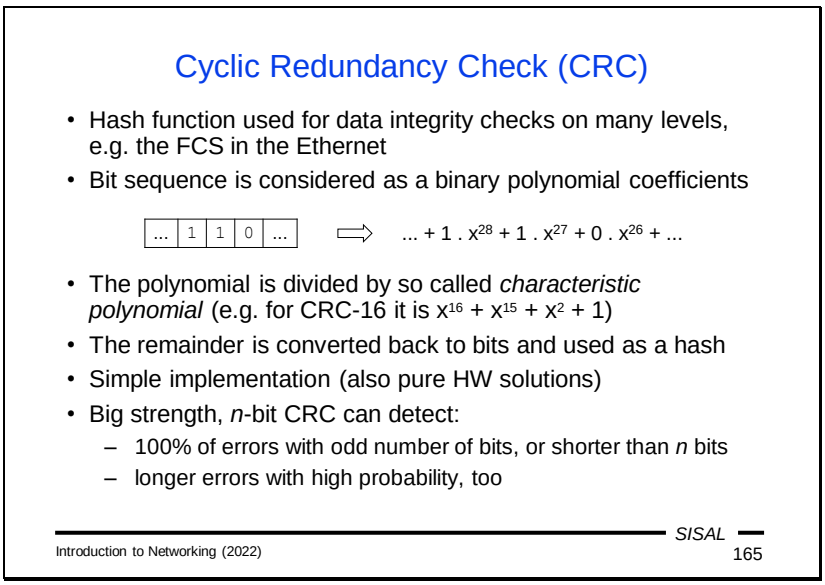



Ethernet version 2 frame format has enabled the deployment of an important network building tool, namely **virtual networks** (VLAN, Virtual LAN). It is used when one physical network needs to be split into multiple logical networks (e.g. when it is shared by multiple entities, or we want to separate segments with a different level of protection).

A principle of the method is to insert a four-byte portion (the so-called *VLAN tag*) into the frame, after MAC addresses, thus changing the frame **type** to a special “VLAN type”, and the inserted portion carries information about a virtual network number (VLANID). But an important feature is that this operation can happen **transparently**, without informing end stations about it. A station can be connected to a switch port that is configured as part of a network with a certain VLANID. In this case, from the station's point of view, the network will look like an ordinary network: the station sends normal frames, a switch inserts a VLAN tag with a correct VLANID into the frame and sends it on. Conversely, when the frame for that station arrives at the switch, the switch removes the VLAN tag and the station receives the frame as if it came from an ordinary network not using VLANs. The station does not have access to frames with a different VLANID that are transmitted over the network. This will logically separate traffic across different virtual networks. However, there are usually nodes in a network that need to have access to frames from all virtual networks (e.g. a central router). For them, the mount point is then configured as a so-called **trunk**, in which case the switch does not do any operations with the frame, and VLAN tag handling is left at the end node.

A slight complication is that each frame is lengthened by adding a tag, so that all the network devices that tagged frames pass through must be able to work with frames

longer than the maximum allowed. Another option is to inform all stations on the network that the maximum allowable frame is 4 bytes shorter than the standard.

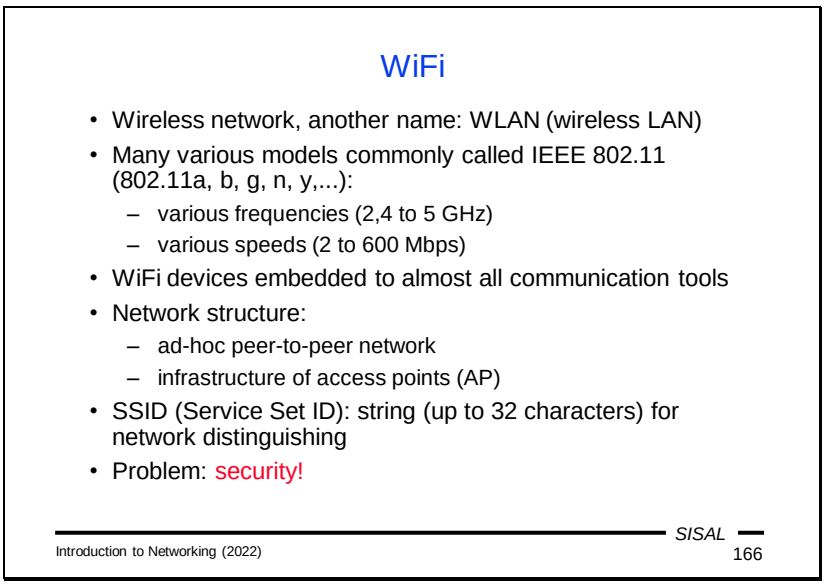



We have already encountered fields containing a content check several times when describing various PDUs. Examples include the IP header checksum, or the Frame Check Sequence in frame footer. Most of these control mechanisms use the **CRC** (Cyclic Redundancy Check) hashing function. This function is based on a division of polynomials, which is quite interesting, because it may seem like such a calculation would be quite complex, and yet it can be done very easily with hardware.

The basic idea is that we convert a sequence of bits into a polynomial with binary coefficients that correspond to individual bits and an order that corresponds to the number of bits. This polynomial is then divided by the so-called *characteristic polynomial*. This polynomial's degree must be equal to the number of bits in a control field. The polynomial that comes out as the remainder after the division is converted again into a sequence of bits, and this gives us the result of a fixed-size function corresponding to the order of the characteristic polynomial.

Despite its simple implementation, the method has surprisingly great power. It can detect all errors with an odd number of bits, all errors shorter than the number of bits of the control field, and even for errors of a different range it has a relatively high success rate.

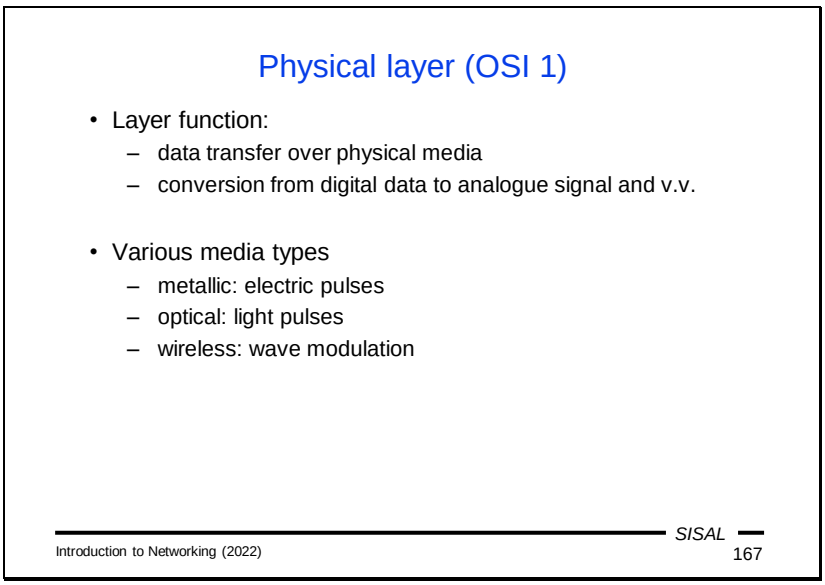
Slide 166




The term WiFi refers to a group of IEEE 802.11 protocols that are used for wireless communications in the unlicensed 2.4 and 5 GHz frequency bands. The name was originally meant to mean nothing, but over time it became a pun parodying Hi-Fi. Another name also used is Wireless LAN (WLAN). Unlike IEEE 802.3 protocols, however, the individual protocols of this group differ significantly from each other and do not form a coherent series.

Common features include the use of CSMA/CA as well as star network topology. Although a WiFi network can also be used in an ad-hoc peer-to-peer variant, it is usually used with a kind of infrastructure where there are *access points* (AP) at the center of each star, to which each individual terminal device connects within its reach. Planning the signal coverage of a campus or a building is complicated, because individual APs must either broadcast on non-overlapping frequency channels (of which there are only 13) or have a central control mechanism, which is relatively expensive.

A problem with WiFi networks is security. Since a potential attacker does not need physical access to the network (they can just stand outside the building), all possible security features must be taken into serious consideration on private networks. Each network is identified by a Service Set Identifier (SSID), but it is not used for security, only to distinguish different networks.

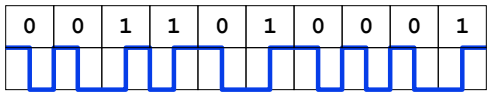



The last layer of the OSI model is the **physical** layer. Its task is to transmit a physical signal across a particular medium. For this, it is necessary to convert the digital information (bits) to analog (electrical pulses on metallic cables, light pulses on optical cables or various modulations of radio waves) first, and the opposite process is required when receiving.

Slide 168

## Data transfer modes

- Analogue vs. digital
  - in fact, everything is analogue (e.g. electric current)
  - digital: thresholds exist that decide whether a signal level belongs to proper interval (less impact of noise)
  - convertors: A→D (and back) *codec* (coder/decoder) D→A *modem* (modulator/demodulator),
- Baseband vs. broadband
  - baseband carries directly encoded signal itself, the Ethernet uses so called Manchester:




  - broadband carries the basic signal and modulates it (phase, amplitude, frequency)



Let's first explain some of the concepts about transmission variants, which can be a little confusing.

Analog vs. digital transmission.

- In fact, every transmission is analog, because the world around us is analog. Whatever medium is used, its physical nature is analogous. So what does **digital transmission** mean? Basically, the difference is that the receiving device is not trying to interpret any possible incoming signal value, but only values that lie within a certain range. One signal range is interpreted as the value 1, another range as the value 0, and a signal outside these intervals is ignored. The transmitting device, on the other hand, tries to minimize the time when the signal is outside the proper intervals. The consequence of these features is that digital transmission is more resistant to interference.

Baseband vs. Broadband transmission.

- The name **baseband** suggests that this form of transmission uses a signal value to encode digital information directly. An example would be the propagation of an electrical signal over a metallic line. We might expect that the easiest way to do this would be to simply encode ‘one’ as a high signal value and ‘zero’ as a low value. With this approach, however, the sender's clock might start to diverge with the recipient's, and the bits sent and received would cease to match. So a "tick clock" must also be transmitted, which is why Ethernet uses so-called "Manchester" coding, where in each "tick" the signal value must be changed, with ‘zero’ being the change from high to low and ‘one’ reversed. Of course, even baseband transmission is in its physical nature analog, so even here in reality there are **not as sharp edges** as in the picture.

- The name **broadband** is not exactly semantically identifiable with the transport it denotes. The name is based on the fact that it is used for broadband transmissions. But its essence is in that data is coded using some **modulation** of the base signal. Either phase shift, amplitude change (known from radio as AM) or frequency change (FM) are used.

Slide 169

## Unshielded Twisted Pair (UTP)

- The most common structured cabling media nowadays
- 4 pairs of copper conductors twisted around each other
  - twisting lowers both emission and reception of electromagnetic radiation (lower interference)
- 100Mb Ethernet uses only 2 pairs (cable can be “divided”)
- Connectors: RJ 45
- Cabling must reflect the nature of devices
  - nowadays usually the MDI/MDIX autodetection available

straight cable crossover

- Option: cable with metallic shielding (STP)



A commonly used connection method for stations on a local network is metallic cabling using the unshielded twisted pair (UTP) cable type. Again, the name is somewhat confusing, as there are **eight** wires in the cable. The essence of the name lies in the **way it is protected** from interference. It is based on the fact that if two conductors are twisted around each other at regular intervals, an electromagnetic field is created when an electric current passes through them, which forms a natural protection against a normal level of interference. In addition, the same signal is sent to both conductors, but with opposite polarity, so that any interference can be “subtracted” on the recipient's side. However, if we need to run the wiring in a highly “noisy” environment, an STP (Shielded Twisted Pair) cable can be used, which has additional metal **shielding** around the wiring.

A strange and interesting thing is why there are eight wires in the cable when Ethernet up to 100Mbps only uses four. However, this can be exploited to an advantage, using special forks to connect **two pairs** of computers with a single cable, each connection using a different four wires.

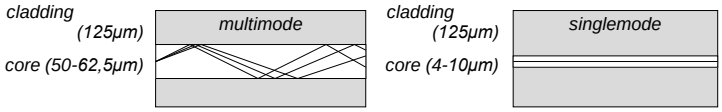
When connecting RJ 45 connectors to a UTP cable, it should be respected that the output pin on one side must be connected to the input pin on the other. So that we can use cables in which the same conductor is input on one side and output on the other (so-called **direct cables**), end stations (computers) use a different pin layout than switches. But this can also cause problems when we need to connect two identical devices (two computers or two switches). For such a connection, we should use a cable where the corresponding conductors are connected in reverse, i.e. a so-called **crossover cable**. Fortunately, when we connect a cable nowadays, network cards are able to negotiate with the counterparty when connecting a cable which wire will be used by which party for input and which for output, so in the normal case, if

both parties support this (so-called MDI/MDIX) auto-detection, there is no need to investigate the type of cable used. However, self-detection takes a certain amount of time, thus also increasing the time required to start communication, which may be unacceptable for some critical real-time applications.

Slide 170

## Optical fiber

- Signal is carried as visible light through a fiber of $SiO_2$
  - high frequency, large bandwidth
  - low attenuation, no interference
- Disadvantages:
  - higher price, demanding installation, don't look into cable
- Fiber types:
  - singlemode fiber - light source: laser => single wave, higher radius, bandwidth („speed", not speed), price
  - multimode fiber - light source: LED




SISAL



Where there is a physical limitation on the spread of an electrical signal over a metallic cable, optical cabling is used. Information is propagated by light pulses that are transmitted via a silicon fiber. The light beam does not have an interference problem, has a very low attenuation and a high bandwidth ("speed"). The downside is a higher price and more demanding handling (the optical fiber has a significantly greater **minimum bending radius**).

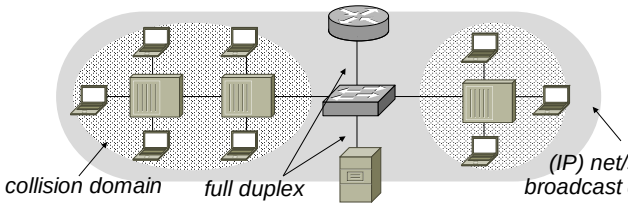
We recognize two kinds of optical fiber cables:

- *Single-mode* cables have a narrower silicon core and a laser is used as a light source. As a result, the refraction of light rays is significantly reduced, fundamentally increasing the range and transmission capacity. And the price. Therefore, these fibers are typically used for long-distance transmissions.
- In *multimode* cables, the silicon core is wider and LEDs can also be used as a light source. The beams are more refractory here, so the range and transmission capacity are smaller and are used most often for the LAN backhaul.

Slide 171

## Network segmentation

- Repeaters connect segments on the physical layer
  - solves: larger radius (eliminates cable attenuation)
  - does not solve: throughput (collision probability increases)
  - terminology of structured cabling: *hub*
- Bridges connect segments on the data link layer
  - solves: larger throughput (by splitting the collision domain)
  - terminology of structured cabling: *switch*






Now let's look at an example of connecting a small local network to explain some other concepts. For reasons of both greater extent and greater throughput, we will need to include several network devices.

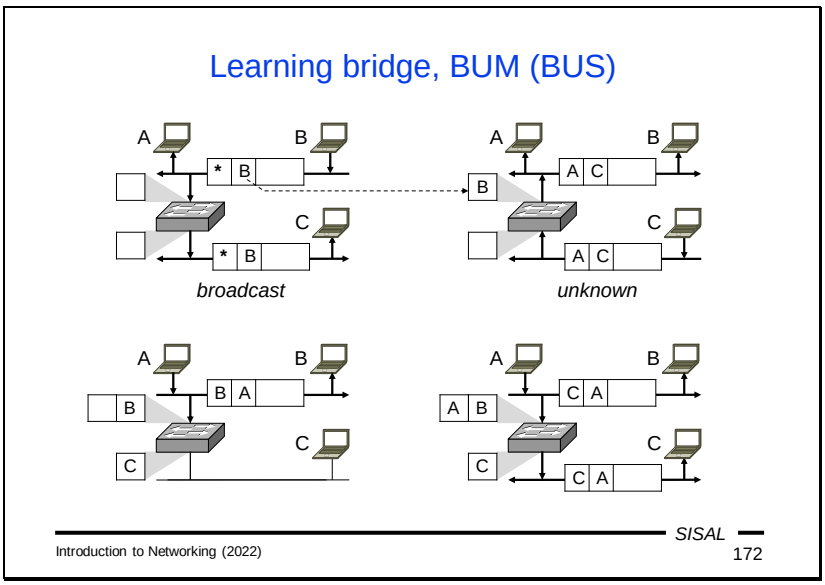
Some stations are far from the center and we can add *repeater* devices to the network to connect them. In structured wiring, the device is usually referred to as a “hub”. A repeater works at the **physical layer**, so its job is just to distribute the proper physical signal without understanding it in any way. This addresses the increase in signal range as attenuation is eliminated. But it does nothing to address the issue of network throughput or the number of collisions. The signal must be distributed from the source to all stations, so if there would be a collision without a repeater, it will occur with it as well. A repeater does not separate the so-called *collision domain*. Conversely, the likelihood of collisions with a repeater increases as a repeater takes time to transmit each frame from an input to an output interface, thereby extending the transmission time of a frame over a segment.

If we want to tackle the issue of **throughput**, we have to move up, to the **data-link** layer and use a *bridge*. In structured wiring, the term "switch" is usually used, which we've seen before. A bridge understands MAC addresses in transmitted frames (although its own MAC addresses don't feature here) and can send frames only where necessary. Each port of a bridge thus separates one collision domain, reducing the number of collisions rapidly. At the same time, it addresses throughput and, to some extent, safety, because each segment receives only traffic that belongs there according to its destination MAC address. Of course, such protection is not sufficient, because a potential attacker can fake an ARP transmission and convince our station to send its frames to him instead of the actual recipient.

To the central switch, two other nodes of exceptional status are connected in the image above – they are a central server and a central router. The two will most likely be connected separately, using a **full-duplex** line to give them maximum throughput.

The entire network, separated by a single router, represents one IP network (or subnet) as well as one *broadcast domain* (the area over which limited broadcasts are spread).

Slide 172




Switches typically operate in the so-called **learning bridge** mode. It means that for each port, they maintain a MAC address table of stations that are connected behind that port and maintain the table themselves by monitoring the traffic.

In the picture above, we see a switch that has just rebooted and whose tables are blank.

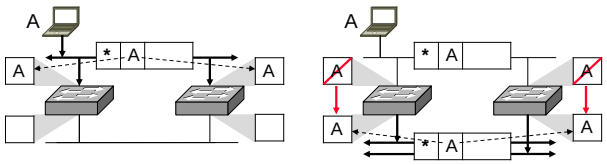
- Let's say the first frame to appear is a broadcast frame sent by station B from the upper segment. Since it is a **broadcast**, the switch must send it to all ports. At the same time, it looks at the sender's MAC address and adds the address of station B into the MAC address table of the upper segment.
- The second frame is sent by station C from the lower segment to station A on the upper segment. However, station A's location is still **unknown** at this time. Therefore, this frame must also be sent to all ports. At the same time, station C's MAC address will be added into the table for the lower port.
- If station A from the upper segment now sends a frame to station B, the switch can now detect that station B is on the same segment of the network (or switch port) and therefore there is **no need to send** the frame anywhere else! The station A entry will be added to the MAC address table, and from now on, the switch has the entire network map assigned to its ports.
- Therefore, when station A subsequently sends a frame to station C, the switch forwards it only to the port where it belongs. No other port or network segment will be burdened by this traffic anymore.

From step 3 onward, the switch will send all frames to the correct ports, with the exception of broadcasts, unknown unicasts and multicasts, which it will continue to send to all ports. That's why this behavior is sometimes called **BUS** (broadcast and unknown service) or **BUM** (broadcast, unknown and multicast).

Slide 173

## Spanning Tree Algorithm

- If two segments are connected by two switches, the network is flooded by forwarded frames, the learning bridge doesn't work




- Reason: the graph contains a loop
- Solution: to find an acyclic subset, spanning tree
- Switches must agree, which acts as backup one (forwarding no data, only monitoring status)
- Protocol (STP) needs timeouts, switch port start is slow
  - usually, the STA can be suspended („faststart"), use carefully



Important points or connections in a network are sometimes backed up by redundant switches due to robustness. But this raises one fundamental problem. If both switches worked, the network would be flooded with frame forwarding, and the learning bridge method would fail.

Let's look again at the example in the picture above. Let's say that station A sends out a broadcast. When it arrives at the switches, they both assign station A's MAC address to their correct port address table, and **both** send the frame to the lower segment. Note that switches, unlike routers, do **not interfere** with the frame content. Therefore, both frames will arrive at the other switch over the lower segment. Both switches will respond by **changing the MAC address assignment**. This will "move" station A to the lower segment and make it unavailable for a time. But at the same time, they re-transmit the frame to the upper segment, forming an endless loop.

If we think of the network as a graph, where individual segments are vertices and switches are edges, the problem is **the circle** that we created by adding the second switch. The problem of reducing the graph and getting rid of the circle is known as the problem of **finding a spanning tree** of a graph. A distributed version of the algorithm is used in computer networks using the Spanning Tree Protocol (STP). Dropping an edge is in practice implemented as converting (some ports of the) switch to a **blocking** mode, in which frames are not forwarded and the switch merely monitors for a failure of the second switch, which is in a **forwarding** mode.

However, STP needs some time for its work. If we plug a station into a normal switch port, STP must first take place and the station will be unavailable for a few seconds. This may sometimes matter, and in that case, certain switch ports can be configured in a "fast-start" mode and the use of the protocol suppressed on them. By doing so,

the administrator **takes responsibility** for not creating a cycle with additional connections in the network.

## Summary 8

- Describe the purpose and safety risks of ARP.

- Describe the purpose of the LLC and MAC layers.
- What is the difference between the physical and logical topology of a network?
- What is a collision, where can it occur and how is it dealt with?
- What is a VLAN for and how is it implemented?
- What is a CRC for and how does it work?

- Describe how data is transmitted depending on the underlying medium.
- Describe the various types of metallic and optical cables.
- What is a learning bridge and STP?

The End