Ethernet Basics

Rev. 02

PHOENIX CONTACT
INSPIRING INNOVATIONS

Contents

PHOENIX CONTACT
INSPIRING INNOVATIONS

Chapter 1

Introduction

1.1 The OSI model

In 1979, the International Organization for Standardization (ISO) had developed a model with the aim to structure and standardise the world of data communication and networks. The ISO is the committee that has developed the Open Systems Interconnection (OSI) reference model. The ISO's objective was to develop a reference model whereby mutual communication between two systems, e.g. two computers, could take place.

In accordance with the ISO
OSI model (also called the 7-layer model), system A can communicate with system B (2 systems from 2 different suppliers). Between these systems, different networks can be present; public as well as private networks.

A public network is a network that is accessible by everyone, provided that the conditions that apply to this network are complied with. A private network is mostly company-specific.

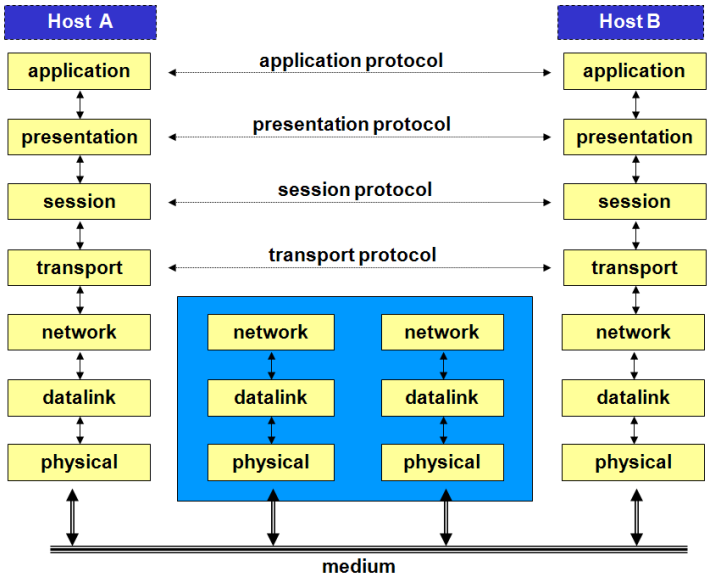


Figure 1.1: The OSI model

The OSI model consists of seven functional layers. Every layer contains a number of defined functions. A limited enumeration of the different layers is given below:

PHYSICAL LAYER (layer 1)
This layer ensures the connection with the medium via which the information is sent between two points in the network: this means that this layer provides the mechanical, electrical or optical entities that are required to realise, maintain and break off the physical connection.

DATA LINK LAYER (layer 2)
The protocols of layer 2 specify how the frames eventually have to be sent over the network. Layer 2 maintains an error detection- and correction mechanism in order to be sure that transmission errors are handled and that data are correctly received on the other side.

NETWERK LAYER (layer 3)
The addressing is configured on this level. This means that the network finds a route and avoids congestion within the network. The network layer ensures the transport of messages from one node to the other on the sender's route to the final receiver.

TRANSPORT LAYER(layer 4)
The transport layer is responsible for a reliable transmission of data. The transport layer ensures a logical connection between both end systems of the network (a logical point to point connection). This means that a faultless data transport can be realised whereby the data is received in correct order by the receiver.

SESSION LAYER (layer 5)
The control structure of the dialogue (session) between two applications over the network is provided for here, as well as the setting up and termination of such a session.

PRESENTATION LAYER (layer 6)
The protocols in layer 6 determine how data is represented: this is necessary as different computer systems represent numbers and characters in different ways. So, this layer ensures, amongst others, the conversion of character codes, e.g. from ASCII to EBCDIC.

APPLICATION LAYER (layer 7)
This layer provides service to applications that run for the benefit of network system users.

It has been agreed for the reference model that the message to be sent by the sender will run through these seven layers. Every layer of the model gives the message a header, starting from layer 7 and then descending until layer 1, see figure **??**. The header shows which data communication functions have to be carried out.
For the functioning of the communication protocols, every layer exchanges information with the corresponding layer on the other side of the connection, apart from the application data that the final users of the connection send to each other. In the OSI model, every layer adds a piece of information (header) to the user data on the sending side. The corresponding layer on the receiving side removes this information again. The data link layer not only places additional information in front of the transmitted data but often also behind it. This trailer contains a check code for the detection of possible transport errors. Only the physical layer does not add anything.

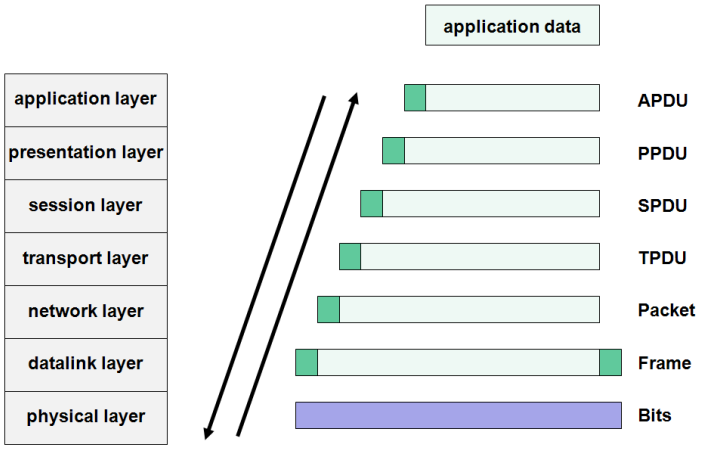


Figure 1.2: Protocol overhead in the OSI model

1.2 LAN

A local network (Local Area Network (LAN)) has been developed to ensure communication between computers, work stations and peripherals in an area of a very limited geographical size.
The connected stations in a LAN are autonomous, meaning that primary and secondary stations do not exist. Every station can set up, maintain and break off a connection with another station. With regard to public networks, the four bottom layers of the OSI model require a slightly different approach for a LAN.
The 802 committee of the Institute for Electrical and Electronic Engineers has established a number of standards for LANs.

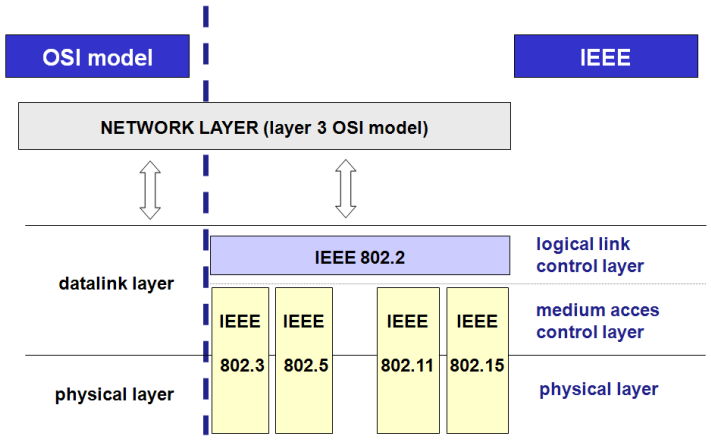


Figure 1.3: Location LAN within the OSI model

Figure **??** shows the filling in of layers 1 and 2 in the OSI model by the IEEE802 standard. Standard IEEE802.1 can be consulted for general concept on LANs.
Currently, the following work groups are active within the IEEE802 committee:

- IEEE802.1 Bridging (networking) and Network Management

- IEEE802.2 Logical Link Control
- IEEE802.3 CSMA
  CD (Ethernet)
- IEEE802.5 Token Ring
- IEEE802.11 Wireless LAN & Mesh (Wi-Fi certification)
- IEEE802.15 Wireless PAN
  - IEEE802.15.1 (Bluetooth certification)
  - IEEE802.15.4 (ZigBee certification)
- IEEE802.16 Broadband Wireless Access (WiMAX certification)
- IEEE802.16e (Mobile) Broadband Wireless Access
- IEEE802.16.1 Local Multipoint Distribution Service
- IEEE802.17 Resilient packet ring
- IEEE802.18 Radio Regulatory TAG
- IEEE802.19 Coexistence TAG
- IEEE802.20 Mobile Broadband Wireless Access
- IEEE802.21 Media Independent Handoff
- IEEE802.22 Wireless Regional Area Network

Chapter 2

Ethernet

2.1 Introduction

Ethernet is the basis of LAN networks. The current LAN market is characterised by an, up to now, unknown degree of standardisation on Ethernet. Due to its huge market share, Ethernet, despite some disadvantages, scores over all alternative technologies.

A short historical overview:

- 1980: Digital Equipment Corporation, Intel and Xerox released the first Ethernet specification, version 1.0, under the name *Ethernet Blue Book* or DIX standard. It defines *Thick Ethernet* in case of 10Mbps CSMA/CD. The first Ethernet controllers, based on the DIX standard, were available starting from 1982. The second and final version of the DIX standard, version 2.0, was released in November 1982: *Ethernet II*.
- 1983: The Institute of Electrical and Electronic Engineers (IEEE) launches the first IEEE standard for Ethernet technology. It was developed by the 802.3 group of the IEEE802 committee and this under the name *IEEE802.3 Carrier Sense Multiple Access with Collision Detection Access Method and Physical Layer Specifications*. IEEE reworked some parts of the DIX standard, especially with regard to the definition of the frame determination.
- 1985: IEEE802.3a; definition of thin Ethernet, cheapernet or 10Base2
- 1987: IEEE802.3d; Fiber Optic Inter Repeater Link (FOIRL). Use of two fibre optic cables to extend the distance between 10 Mbps repeaters up to 1000m.
- 1987: IEEE802.3e; 1Mbps over twisted pair
- 1990: IEEE802.3i; release of the popular 10Base-T; 10Mbps over UTP category 3
- 1993: IEEE802.3j; 10Base-F: distances greater than 2 km over fibre optic
- 1995: IEEE802.3u; 100Base-T and 100Base-F
- 1997: IEEE802.3x: full-duplex Ethernet
- 1997: IEEE802.3y; 100Base-T2
- 1998: IEEE802.3z; 1000Base-X standard; generally known by the name Gigabit Ethernet

- 1999: IEEE802.3ab; Gigabit Ethernet over twisted pair
- 1999: IEEE802.3ac; 802.1Q: definition of the Q tag with VLAN and priority information.
- 2003: IEEE802.3af; Power over Ethernet
- 2006: IEEE802.3an; 10GBase-T
- 2006: IEEE802.3aq; 10GBase-LRM, Ethernet over multimode fiber

Ethernet is only a specification of layers 1 and 2 in the OSI model. It is not a complete network protocol but a subnet on which other protocols such as the TCP/IP suite can work.

The most important functions of ETHERNET are:

- Filling in of the physical layer
  - sending and receiving the serial bit streams over the medium.
  - detecting collisions.
- Filling in of the data link layer
  - MAC sublayer:
    - access mechanism to the network (CSMA/CD).
    - building of the data frames.
  - LLC sublayer:
    - data reliability.
    - supply data channels for higher-level applications.

2.2 The physical implementations

The most important implementations over the years are:

- Thick Ethernet (10Base5)
- Thick Ethernet (10Base2)
- Broadband Ethernet (10Broad36)
- Ethernet over twisted pair (10Base-T)
- Ethernet over Fiber (10Base-F)
- Fast Ethernet (100Base-T / 100Base-F)
- Gigabit Ethernet (1000Base-T)
- Wireless Ethernet

2.2.1 Implementations based on coax

The original Ethernet was designed around the concept of a bus topology. The first implementations of Ethernet were based on a thick yellow coax cable - thick Ethernet - also named 10Base5.

Features of the original Ethernet:

- 10Mbps
- Baseband (basic band transmission)
- max. 5 x 100 = 500 meter
- max. 100 transceivers per segment

Thick Ethernet coax cables have a marking every 2.5 metres in order to ensure correct positioning of the 10Base5 transceivers (or MAUs). These transceivers are used to connect stations to the network. The transceivers can be positioned every 2.5 metres, this avoids reflections of the signals, resulting in a poor transmission quality.

This type of implementation has been superseded. The thick, rigid yellow coax was rapidly replaced by the black, more flexible coax which resulted in the implementation of thin Ethernet (10Base2). The connection of the different stations is realised by T-shaped BNC connectors whereby a maximum segment length of about 200 metres can be applied.

Important cabling detail that is required for most bus technologies: the terminating resistance (terminator) - a small, cheap device that has to be mounted on all endings of the coax cables that form an Ethernet. A terminating resistance consists of a resistance that connects the central core of the cable with the shielding: when an electrical signal reaches the terminating resistance, this is discarded. For the correct functioning of a network, the terminating resistance is indispensable as the end of the non-terminated cable reflects electrical signals just as a mirror reflects light. When a station tries to send a signal over a non-terminated cable, then this signal will be reflected by the cable end. When the reflection reaches the sending station, interference will occur.

2.2.2 Implementations based on twisted pair

The major problem with coax is that only half duplex communication can be applied. The applied bus structure is also not ideal if certain problems occur. In order to break through the bus topology, Ethernet has switched to a topology where twisted pair can also be used: all stations are connected with one or more central hubs. This way, a star topology can be worked out. The network can easily be extended and checked in this way and it facilitates error detection. The maximum segment length between a participant and a hub is 100 metres.

The variants on the basis of twisted pair have evolved from 10Base-T (10Mbps) to 100Base-T (100Mbps) to 1000Base-T (1000Mbps).

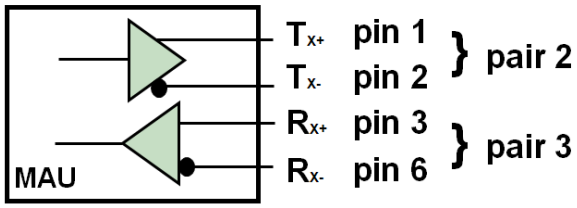


Figure 2.1: The MAU for 10/100Base-T

The MAU, developed for twisted pair, is equipped with 4 data pins: 2 for sending, 2 for receiving. This is the basis for full duplex Ethernet. In principle, any point to point communication is possible. Therefore, every host has to be connected directly with a structure element: a hub or a switch.

Fast Ethernet

The UTP cable, e.g. CAT5 (Class 5) UTP (Unshielded Twisted Pair), supports speeds up to 100Mbps. The cable consists of 8 wires, arranged in 4 pairs. The 4 pairs can be identified as 1 is always completely coloured and the other one has the same colour with white parts in between. Only 2 of the 4 pairs are used in 10/100Base-T (pair 2: orange/white and orange and pair 3: green/white and green).

The IEEE specification for Ethernet 10/100Base-T requires that the one used pair is connected to pin 1 and pin 2 of the connector while the second pair is connected to pin 3 and pin 6. The other two unused pairs will be connected to pin 4 and 5 and on pin 7 and 8.

Table 2.1: Pin configuration for Fast Ethernet

Pin	Colour	Function
1	green with white	+TD
2	green	-TD
3	orange with white	+RD
4	blue	unused
5	blue with white	unused
6	orange	-RD
7	brown with white	unused
8	brown	unused

Table 2.1 shows the pin configuration for 10/100Base-T. TD stands for Transmitted Data, RD stands for Received Data. The plus- and the minus signs indicate that the signal is sent mirrored via two data lines, also see figure 2.2.

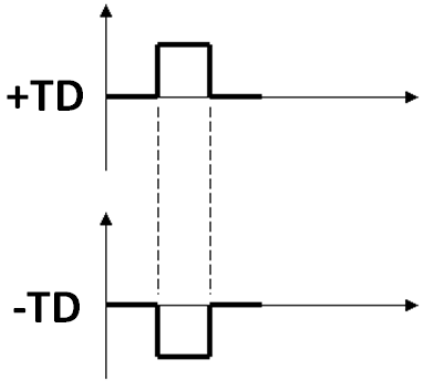


Figure 2.2: Transmission technology for 10/100Base-T

The straight-through cable, also called patch cable, is the cable that we get when we connect both sides of the cable pair 2 with pin 1 and pin 2, while pair 3 is connected with pin 3 and pin 6. This cable can be used for connections between the patch panel and the hub/switch, the PC and the hub/switch or the PC and the wall. This cable is generally used for the connection of a structure element and an end element.

A cross-over cable is required to set up the PC-PC connections (connection of two end elements) and to secure connections between hub/switch and another hub/switch (connection between two structure elements). In order to make a cross-over cable, we have to switch the used pairs. Along one side, pair 2 has to be connected with pin 3 and pin 6 while pair 3 has to be connected with pin 1 and pin 2.

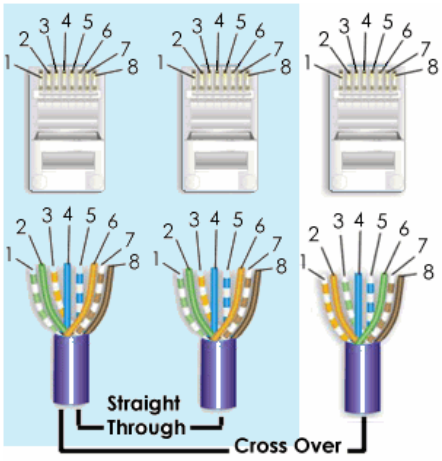


Figure 2.3: Twisted pair cabling, 10/100Base-T

Current Ethernet ports support autocrossing. This means that it can be detected automatically which cable is used and the crossing will be corrected internally if necessary.

The IEEE Fast Ethernet has defined 100Base-T as extension on the 10Base-T.

Fast Ethernet is characterised by:

- Data transmission at a speed of 100Mbps
- Full Duplex communication
- Wireless Ethernet

In Fast Ethernet, a mechanism is provided for auto negotiation: this makes it possible to built Ethernet interfaces that switch automatically between 10Mbps and 100Mbps.

For the 10Base-T standard, every data bit is coded in one physical bit. In other words, for a group of eight data bits, eight signals are generated in the cable. The 10Mbps data rate means a clock rate of 10MHz. For every clock pulse, one single bit is sent.

100Base-T uses the so-called 4B/5B scheme whereby each group of four bits is coded in a 5 bit signal. So, one single bit is not exactly converted into one single signal in the cable.

Data stream:	0111010000100000
4 bit pattern:	0111 0100 0010 0000
5 bit code:	01111 01010 10100 11110

The applied clock rate is 125MHz (5/4 x 100). Cat5 cables are certified for a transmission speed up to 125 MHz.

Gigabit Ethernet

Gigabit Ethernet targets a data rate of 1000Mbps. If the CAT5 Ethernet cables have to be used for this, for example, then this causes a problem as they only support a clock rate up to 125MHz. In order to realise this, the technology has to be adapted.

First, 1000Base-T codes two bits per clock signal (00, 01, 10 and 11) and uses four voltage levels for this.

Furthermore, 1000Base-T uses all four data pairs of an Ethernet cable. The four data pairs are applied here bi-directionally. Data are sent or received via all four data pairs.

Gigabit Ethernet therefore still uses the 100Base-T/Cat5 clock rate of 125MHz. A data rate of 1000Mbps is reached as 2 bits are being processed for every clock pulse and this is done via four data pairs. This modulation technology is called 4D-PAM5 and currently uses five different voltage levels. The fifth voltage level is used for the error mechanism. Table2.2 shows the Gigabit Ethernet pin configuration. BI stands for bi-directional while DA, DB, DC and DD stands for data A, data B, data C and data D.

Table 2.2: Pin configuration for Gigabit Ethernet

Pin	Colour	Function
1	green with white	+BI_DA
2	green	-BI_DA
3	orange with white	+BI_DB
4	blue	-BI_DB
5	blue with white	+BI_DC
6	orange	-BI_DC
7	brown with white	+BI_DD
8	brown	-BI_DD

2.2.3 Implementations based on fibre

In order to make longer segment distances possible, the glass fibre cable was integrated as a suitable interface. The first fibre glass variants are known by the name 10Base-F and 100Base-F. Separate fibres are used all the time for the sending and receiving of data.

Gigabit Ethernet over fibre has been developed for the full-duplex mode with a data rate of 1000Mbps. There are two different variants for Gigabit Ethernet. 1000Base-SX and 1000Base-LX.

1000Base-SX uses light pulses with short wavelength over multimode fibre. 1000Base-LX uses light pulses with long wavelength over multimode or single-mode fibre.
Recently, 10Gigabit Ethernet over fibre with different variants also has been added.

2.2.4 Wireless LAN

IEEE802.11

The IEEE defines different standards for wireless LAN in their IEEE802.11 description. The radio connections for a Wireless LAN take place in the 2.4 GHz frequency band, the so-called ISM band (Industrial, Scientific and Medical) or in the 5 GHz band. No licences are required for this. A Wireless LAN uses the so-called spread spectrum technology. This technology is specifically meant for fault-prone transmission channels. This is important as these frequency bands (especially the 2.4 Ghz) are also used by many other devices, e.g. Bluetooth.

A wireless network is in general much less fast than a fixed wired network. A major advantage is the flexibility.

With regard to physical implementation, the IEEE802.11 provides the infrastructure configuration or the Ad Hoc configuration.

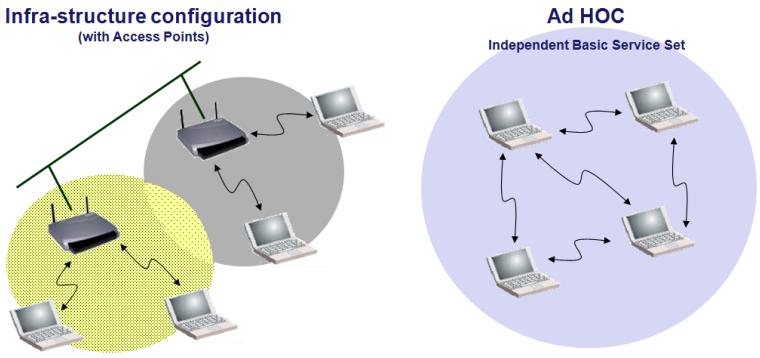


Figure 2.4: Physical implementation of WLAN

Infrastructure configuration is the configuration whereby a wireless access point is used to connect a wireless LAN with a cabled LAN. The wireless access point functions as central point for the routing of the all wireless data traffic. Wireless-enabled computers that are included in an infrastructure mode form a group that is called a Basic Service Set (BSS). At a certain moment, a maximum of 64 individual computers can be included in a BSS. This is because the capacity of the wireless access point is limited to 64 clients. The complete wireless network has a unique SSID (Service Set Identifier) and is also has a network name. This name only applies to the wireless network.

Ad hoc or peer-to-peer relates to a wireless configuration in which every participant communicates directly with the other. An actual organisation of the network is therefore not possible here. An ad hoc wireless LAN consists of a group of apparatuses each equipped with a wireless adaptor that is directly connected to each other and form an independent wireless LAN in this way.

WLAN standards

Different standards are defined within the IEEE802.11. These standards use different modulation technologies in order to obtain improved transmission speeds. Table 2.3 displays an overview of the different standards.

Table 2.3: WLAN standards within the IEEE802.11

Standard	Frequency band	Data transmission
IEEE802.11b	2.4GHz	11Mbps
IEEE802.11g	2.4GHz	54Mbps
IEEE802.11a	5GHz	54Mbps
IEEE802.11h	5GHz	54Mbps
IEEE802.11n	5GHz and/or 2.4GHz	600Mbps

IEEE802.11b/g

IEEE802.11b/g uses the 72 MHz band part of the 2.4 GHz band. 11 channels of 22MHz band are defined here, in accordance with the FCC rules. Theoretically this would mean that the bandwidth for these 11 channels is 242 Mbps (11x22 Mbps). In reality, this has to be reviewed as these channels overlap for a large part. Figure 2.5 shows that there are only three non-overlapping channels: channel 1, channel 6 and channel 11.

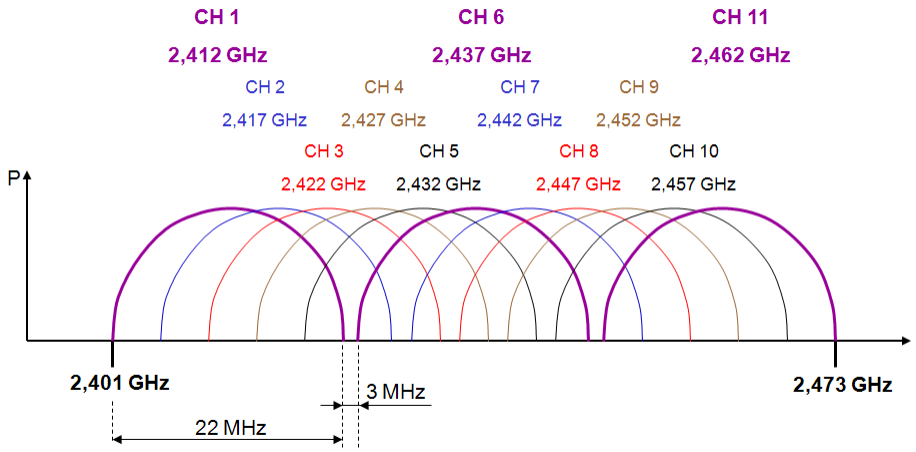


Figure 2.5: The 2.4GHz band for WLAN

The ETSI defines a slightly wider frequency band for Europe, including 13 channels of 22 MHz band. This means, in principle, that we can use 4 barely overlapping channels in Europe. These are channel 1,5,9 and 13.

The IEEE802.11b supports a maximum speed up to 11 Mbps. The IEEE802.11g supports a maximum speed up to 54 Mbps. The speed is decreased dynamically in case of a bad connection or great distance to the access point.

IEEE802.11a/h

IEEE802.11a uses the complete 5GHz band. With the application of OFDM (Orthogonal Frequency Division Multiplexing), the maximum (theoretical) speeds of up to 54Mbps are reached with IEEE802.11a. Figure 2.6 shows the different channels within the 5GHz band. Within Europe, this means that 8 non-overlapping channels of 20MHz wide can be used over the two lowest bands of the 5GHZ UNII band.

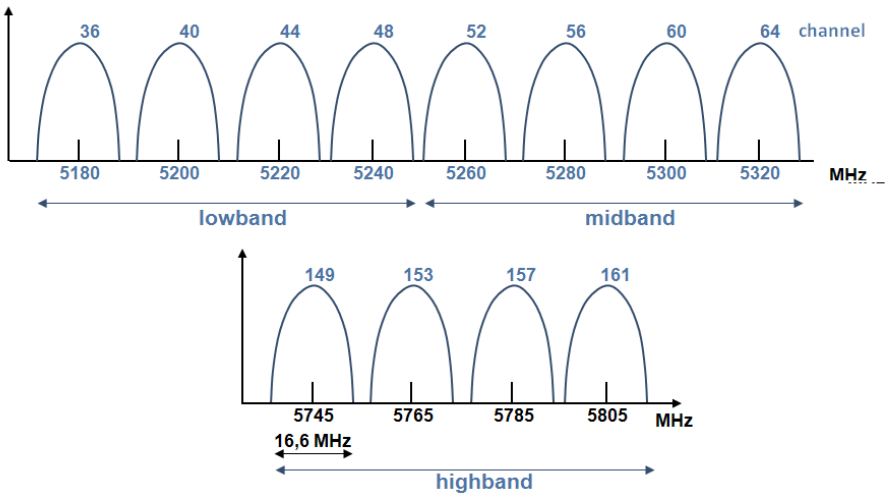


Figure 2.6: The 5GHz band for WLAN

As opposed to the USA, the use of the 5GHz band in Europe has quite a few restrictions. Therefore, the IEEE802.11a is converted into the IEEE802.11h. Two important protocols were added in order to eventually comply with the European regulations:

- DCS (Dynamic Channel Selection):the AP will automatically look for another channel if it appears that the channel is used by another application.
- TPC (Transmit Power Control): just the required capacity is transmitted, if two participants are in close vicinity, then the AP will adapt the capacity to the required level.

IEEE802.11n

This recent standard uses MIMO (multiple input - multiple output), a technique to transmit data wirelessly by means of several reception- and send antennas whereby a transmission speed of maximum 600Mbps is obtained if 4 channels of 40MHz each are used.

2.2.5 Bluetooth

The basic technology (two bottom layers of the OSI model) is standardised in the IEEE802.15.1. Moreover, the Bluetooth SIG (Special Interest Group) defines different application profiles, e.g. serial communication and transmission of Ethernet data frames.

Bluetooth uses the 2.4 GHz licence-free ISM band. As opposed to WLAN, the data to be sent are not spread out over a wider frequency band but FHSS (Frequency Hopping Spread Spectrum) is applied. The 2.4 GHz band is divided over 79 channels of 1 MHz. Figure 2.7 shows the functioning of FHSS. 1600 hops per second can be carried out. Each time, every data frame is sent on another frequency. This means that different logic channels can be active in parallel.

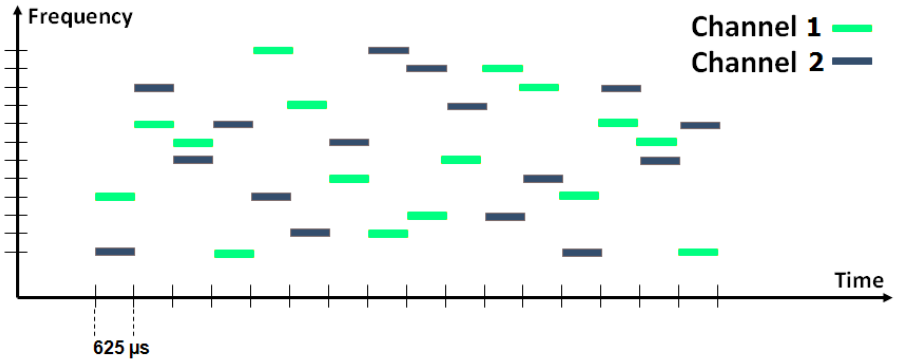


Figure 2.7: FHHS technology

A great advantage of the use of Bluetooth in the industry is the perfect co-existence with WLAN. If there is interference on a Bluetooth frequency as a WLAN channel is active on the same frequency, then Bluetooth can avoid this/these frequency (ies). As this is a frequently occurring issue, Bluetooth has integrated an automated co-existence mechanism: Adaptive Frequency Hopping (AFH).

This mechanism enables Bluetooth to suspend certain ’bad’ frequencies temporarily from the hopping list. Figure 2.8 shows how there is enough space in case of a full 2.4GHz band where three separate WLAN channels are active. The WLAN channel uses a statistic frequency band. Bluetooth can adapt and choose from adequate number of frequencies to avoid interference.

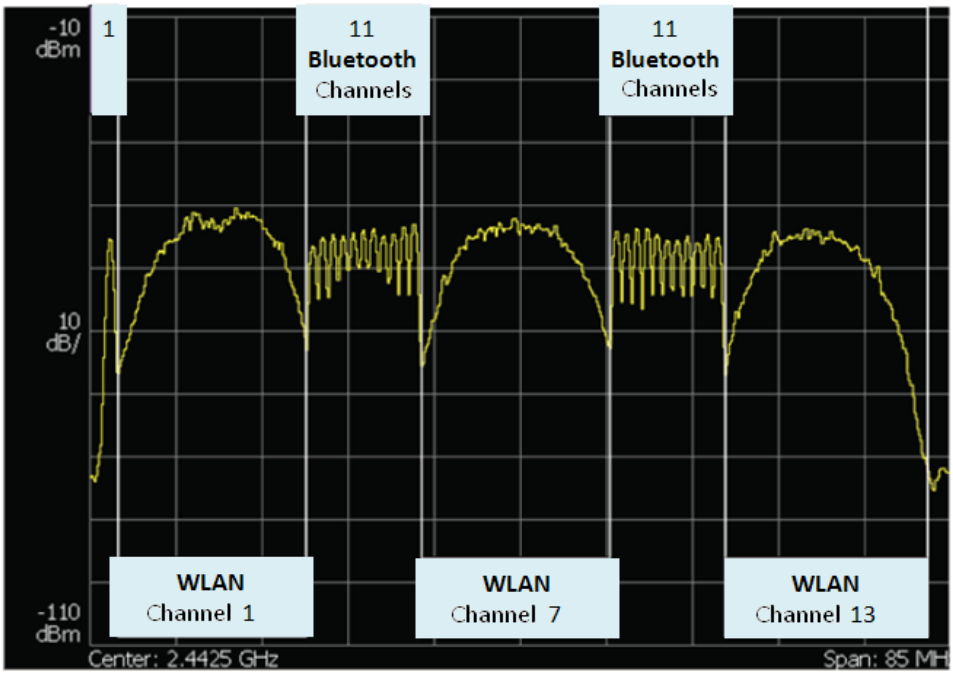


Figure 2.8: Co-existence of Bluetooth and WLAN

2.3 The data link layer

2.3.1 Introduction

Packet switching is used to send messages. Packet switching is mostly applied for computer to computer communication. In computer networks, a random quantity of data is not transported uninterruptedly. Instead, the network system divides the data into small blocks and packets that are sent separately. Computer networks are therefore also called packet switching networks.

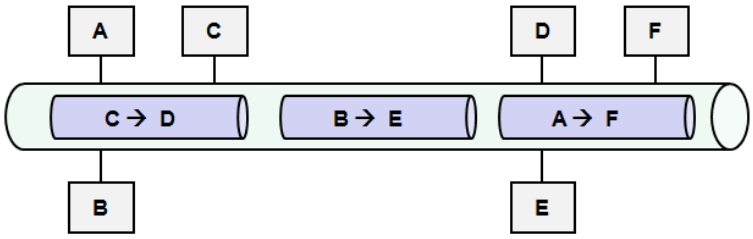


Figure 2.9: Packet switching

There are two reasons to choose usage of packets:

- Sender and receiver has to coordinate the transmission. In case of transmission errors, lot of data may be lost. If the data is divided into smaller blocks, then it is easier for the sender and receiver to determine which blocks are still intact on arrival and which aren't.
- Several computers make common use of underlying links and hardware. A network has to ensure that all computers have equal direct access to a shared communication facility. A computer cannot occupy a shared resource for longer than it takes to send one packet.

2.3.2 MAC address

On a common transmission medium of a LAN, every station has to have a unique address. Every participant has an Ethernet address, a physical address that is unique for the network card: the MAC address (Medium Access Control Address). Every manufacturer of network cards gives each card a unique address number that is stored in the ROM of the card.

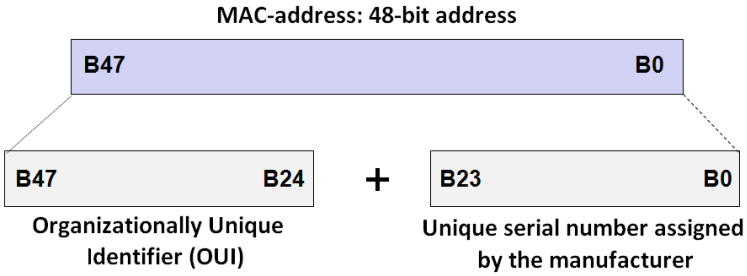


Figure 2.10: The MAC address

The MAC address consists of 48 bits (6 bytes) and is divided into two groups of three bytes. The highest 24 bits form a manufacturer number issued by XEROC. There are 4194302 possible manufacturer numbers. Phoenix Contact, for example, is assigned manufacturer number 00A045h.

The lowest 24 bits form a serial number. Every MAC address has to be unique.

2.3.3 The Ethernet dataframe

An Ethernet frame consists of at least 46 actual data bytes and a constant number of 26 protocol bytes (overhead). This minimum number of data bytes is necessary for the definition of the slot time.

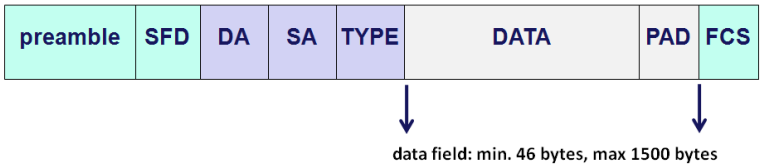


Figure 2.11: Building of an Ethernet data frame

The following fields are defined in an Ethernet data frame:

- Preamble: is a series of 56 bits alternating with 1 and 0. These bits are used for synchronisation and give each participant the time to observe the activity on the bus before the actual data arrives.
- SFD: the start of frame delimiter (10101011), the last byte of the preamble, indicates to the receiver that the actual data is on its way.
- DA: the destination address. The destination MAC address field identifies the station or the stations that have to receive the message. This field takes 6 bytes of space. The destination address can be an individual, a multicast or a broadcast address. The MAC broadcast address is FF FF FF FF FF FF.
- SA: the source address. The source MAC address field identifies the station from where the message originates. This field is 6 bytes long.

- TYPE: for the field type, a distinction is drawn between Ethernet II (DIX standard) and the IEEE802.3
  . For Ethernet II, the field type refers to the higher-level protocol that uses an Ethernet frame to send data. Xerox assigns a code of 2 bytes to every protocol that is developed for Ethernet. Some examples:

0600h	XNS
0800h	IP (Internet Protocol)
0806h	ARP protocol
0835h	Reverse ARP protocol
8100h	IEEE802 1.q tag frame (VLAN)

  The IEEE802.3 defines the field TYPE as LENGTH field in order to be able to send the number of actual data bytes.

  Xerox does not use type numbers below 1500 and as the maximum length of a data frame is 1500, no overlapping is possible and both definitions can be used.

- DATA: the data field contains the data to be sent. This data field is transparent- this means that the content of this field is completely free for Ethernet. Only the length has to be a minimum of 46 bytes and not more than 1500 bytes.
- PAD: the padding bits are random data bits that, if necessary, can be added to the data in order to reach the minimum required 46 bytes.
- FCS: the check sum is a 4-byte CRC value that the sender creates and sends. The receiver can check the integrity of the data with this code.

2.3.4 CSMA/CD

Ethernet uses the CSMA/CD (Carrier Sense Multiple Access / Collision Detect) protocol. With CSMA/CD, two or more stations can use a common transmission medium. In order to send a data frame, a station has to wait for an 'idle period'- when the bus is inactive and not a single participant is sending data. It will then send a message that is heard by all other participants. If a second participant is sending a message at the same time, then a collision will be detected. The participant that detects a collision first, sends an error frame.

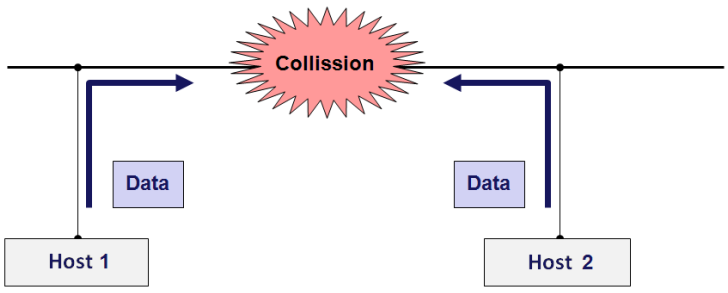


Figure 2.12: Collisions on an Ethernet segment

A collision domain is a multi-segment configuration in accordance with the CSMA/CD protocol whereby a collision will occur when 2 participants send a data frame at the same time on the segment.

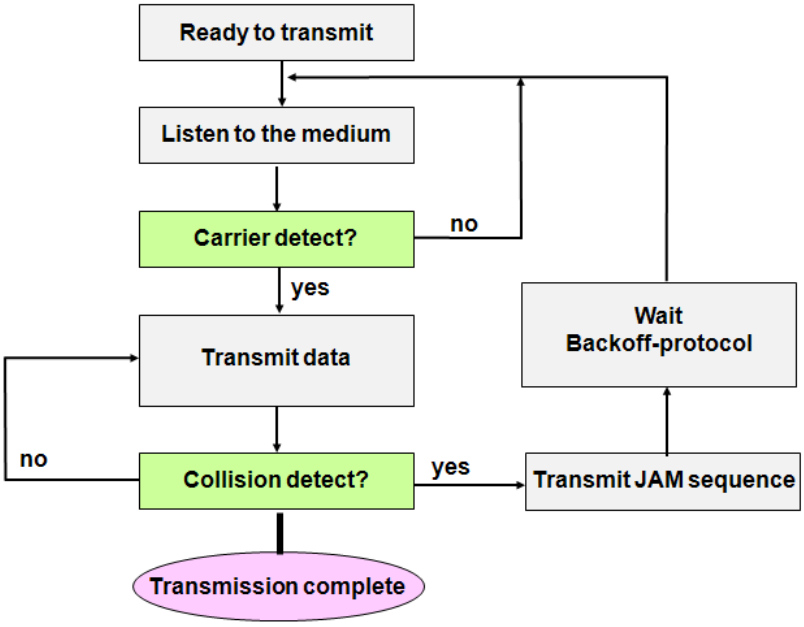


Figure 2.13: CSMA/CD flow

Figure 2.13 shows the CSMA/CD flow. A participant that wants to send data will first check the network on a *carrier*, or the presence of a station that is sending data. If an active carrier is detected, then the sending is delayed.

If no active carrier is detected for a period that is equal to or greater than the interframe gap, then this station can start sending the message. During the sending of the message, the participant will continue to check the medium on collisions. A network interface therefore has to send data and check the medium at the same time. If a collision occurs, then the participant stops the sending immediately and a 32-bit jam sequence is sent. If the collision is detected early, then the frame preamble will be sent before the jam sequence is sent. This jam sequence is necessary in order to make sure that the length of the collision is sufficiently long so that all participants can observe the collision. After sending the jam sequence, the participant will have to wait for a random period of time before making a new attempt: this process is called Backoff.

A few important additional definitions:

- Interframe gap: Ethernet participants have to plan a minimum period without activity (’idle period’) between the sending of two frames. The minimum interframe gap is 96 bit times (9.6μs for the 10Mbps version, 960ns for 100Mbps Ethernet and 96ns for Gigabit Ethernet.
- Slot time: this parameter is defined as 512 bit times for the 10Mbps and the 100Mbps versions, and 4096 bit times for Gigabit Ethernet. The minimum transmission time for

a complete data frame should be at least one slot time. The time required so that all participants can observe a collision, cannot be more than one slot time.

The slot time is an important Gigabit Ethernet parameter:

- it determines the minimum length of a data frame (64 bytes for 10Mbps and 100Mbps). Every frame shorter than 64 bytes is considered as a collision fragment.
- it determines the maximum length of a collision domain in order to avoid late collisions.
- it ensures that if a collision has to take place that it will happen within the 512 bit times of the frame transmission time.

2.3.5 CSMA/CA

The CSMA/CD technology of wired Ethernet cannot be applied to wireless Ethernet. The standard describes half-duplex radios, while sending the data it cannot be checked whether any collisions take place. In order to solve this, another technology is applied, namely CSMA/CA. Instead of detecting collisions, collisions will be avoided, CA: collision avoidance.

The chance of collisions is the greatest right after an occupied medium. That is why waiting times and a recovery phase are defined. Figure 2.14 shows some important parameters with regard to waiting times for the access to the medium. All parameters are related to the slot time (derived from the propagation time delay that the medium causes. These parameters are:

- SIFS (Short Interframe Spacing): shortest waiting time for medium access (thus highest priority). The access point uses this waiting time for the sending of ACK messages.
- PIFS (PCF Interframe Spacing): medium priority, this time is used for the polling actions of an access point.
- DIFS (DCF Interframe Spacing): lowest priority for medium access, applicable to normal participants on the wireless segment.

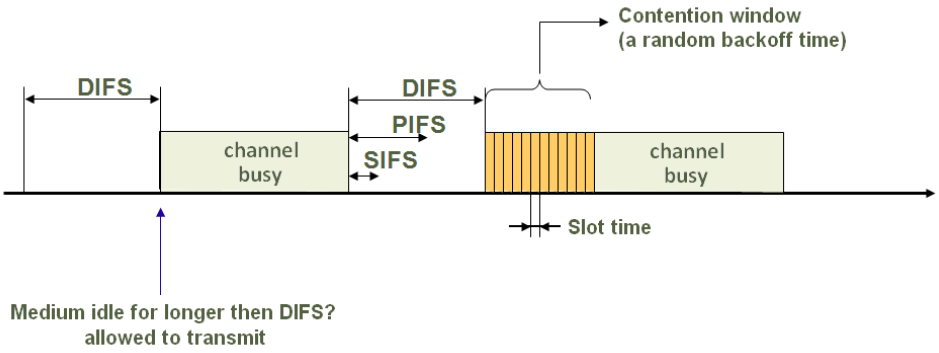


Figure 2.14: CSMA/CA

If a host wishes to send a message, then first the medium has to be listened to. If the medium is longer than the time DIFS is free, then this participant can take the initiative to send a message.

If it appears that the medium is occupied, then one has to wait until the sending participant has completed the sending. Then a DIFS time has to be waited. The access point has a higher priority and only has to wait for a SIFS time. If the medium is still free after the DIFS time, then the recovery phase starts whereby every host, that wants to send data, starts a random backoff timer. The participant that has completed the counting first, can take the initiative to use the medium and send the data.

2.4 Structure elements for Ethernet

2.4.1 The hub

The maximum segment length of a LAN is determined by the used medium and the applied access mechanism. In order to cancel the length restriction, methods are rapidly searched to link several segments one after another. The first and most simple method is to use a repeater. A repeater is a signal amplifier that transmits packets transparently, independent of the package content. A repeater is used to connect two or more Ethernet segments together. As can be seen on the slight , a repeater link takes place on the physical layer, in accordance with the ISO-OSI definitions.

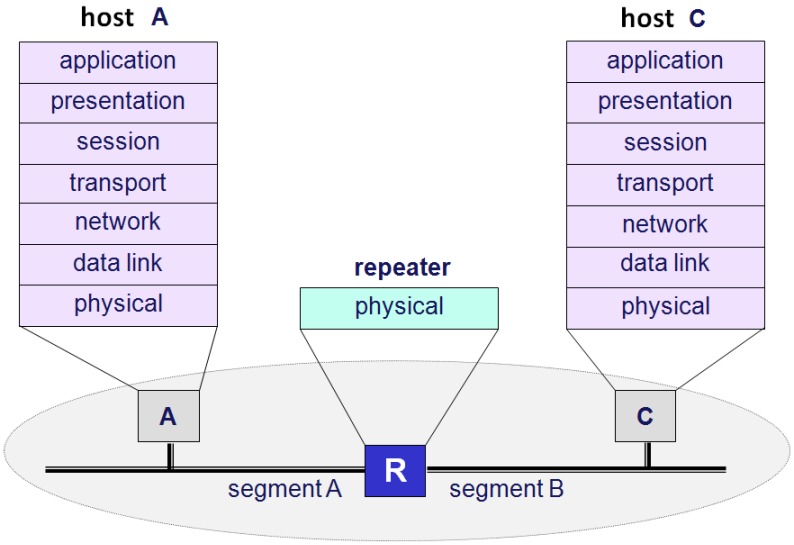


Figure 2.15: The repeater in accordance with the OSI model

Both segments can have a different medium. A 10Base-T based segment, for example, can be connected to a fibre glass segment by means of a repeater. Another important feature of a link on the basis of a repeater is that not only the data bits are transmitted but also any collisions and signal errors. Network segments that are connected mutually via a repeater are therefore prone to fault situations; a problem on one segment multiplies over all other segments. In modern local networks, based on Ethernet, repeaters are mainly used to connect segments of different media with each other. The backbone segments from fibre

glass cabling are always connected via optical repeaters to branch segments of twisted pair cabling.

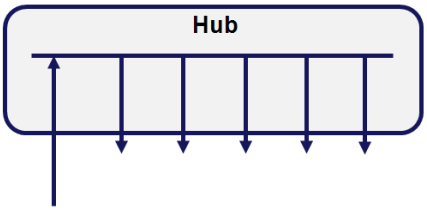


Figure 2.16: The hub

A hub is actually a multiport repeater: it regenerates incoming signals to all other ports as can be seen from figure 2.16. All segments that are connected with each other via a hub are a collision domain.

A hub is available in several different versions. These versions differ in the number of ports, the media types that are supported and the extensibility.

An important functionality of the modern hub is the option for network management. It is at least possible to switch off the ports and to detect whether failures have taken place. In order to make available this option, a modern hub is equipped with an SNMP agent that is controlled from a management station.

2.4.2 The switch

One of the options to interlink LAN segments with more intelligence is to use a bridge. A bridge is more than just a medium that transmits data like a repeater. Before a package is transmitted from one segment to the other segment via a bridge, a bridge checks the MAC address and on this basis the transport to the other segment takes place or not.

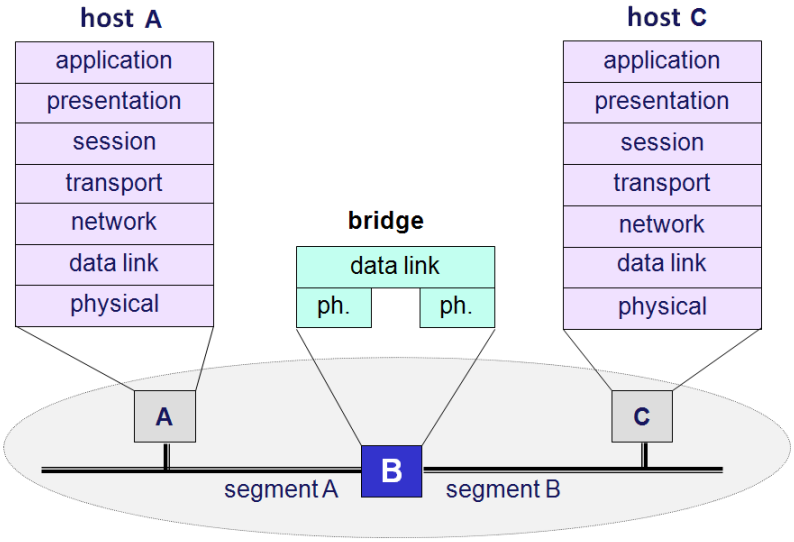


Figure 2.17: The bridge in accordance with the OSI model

A bridge can be equipped with more than two network ports. In that case, the term switch is used. A MAC address table is updated from a software point of view for every port. This table is filled by listening on the relevant segment of the network and by copying all MAC addresses that occur on that segment to the table. Every address is retained for a limited time and is deleted again as soon as a certain time (the hold time) has lapsed. This technique avoids that inactive stations are addressed or that stations are not recognised anymore.

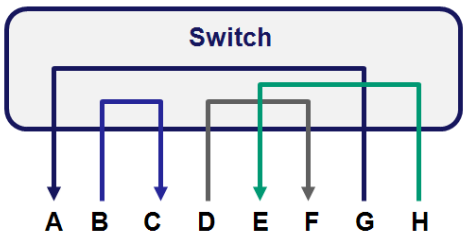


Figure 2.18: The switch

Linking the segments of a local network via a switch has a number of advantages over the link with a repeater or a hub. When using a switch, a segment is not loaded with the frames of the other segment that do not belong there from an addressing point of view. The load per segment is reduced by this bridge function. At the same time, fault situations are not transmitted as the switch also checks the correct building of the frame. Finally, the bridge also avoids that collisions between frames are transmitted from one segment to the other. Every port of a switch closes a collision domain. If every participant connects directly to the port of a switch, then many collision domains occur but each domain only contains one participant and no collisions can occur. The switch is elaborated upon in another part of the document.

2.5 IEEE802.1Q tagged frame

The IEEE802.1Q describes 4 extra bytes, divided into two extra fields in the Ethernet frame in order to use for new applications. One of these applications is VLAN (see also in this chapter).

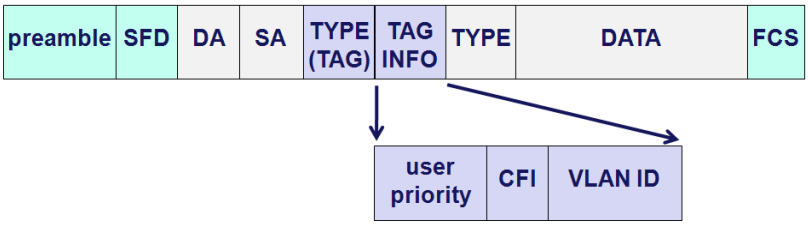


Figure 2.19: Building of a tagged frame

Description of the extra fields:

- TYPE(TAG), 2 bytes: has the value 8100h to specify that this frame is a tagged frame and therefore contains an extra information field
- VLAN TPID, 2 bytes: VLAN Tag Protocol Identifier
  - User priority, 3 bits: the priority of the frame is included, the priority code (a number between 0 and 7) is described in IEEE802.1p.
  - CFI: Canonical Format Indicator. The IEEE802.1Q is only developed for Ethernet or Token Ring. This bit is 0 for Ethernet and 1 for Token Ring.
  - VLAN ID: Identification of the VLAN, 4094 possibilities.

    FFFFh reserved
    0000h no VLAN, frames with priority (Profinet IO)

2.6 Power over Ethernet

The IEEE802.3af*Power over Ethernet* offers since June 2003 the possibility of common transmission of data and power over the same Ethernet cable.
PoE was developed for WLAN access points, Bluetooth access points, IP telephones (voice over IP), IP cameras, RFID reading units, touch panels, ... This had already been applied before via non-standardised systems, the unused lines of an Ethernet cable were used to transmit 24V or 48V.

The stream to the devices can be limited and checked via the IEEE802.3af standard. The use of PoE makes an extra power adapter superfluous. This is especially handy if the network device is used in a place where power supply via the electric point is difficult to realise.

The protocol defines two basic components. The PSE (Power Sourcing Equipment) and the PD (Powered Device)

2.6.1 PSE

The device that provides power to PoE is called a PSE (Power Sourcing Equipment). The voltage is 48V nominal (between 44V and 57V). Every port of a PSE should be able to supply 350mA for 44V (15.4 Watt).

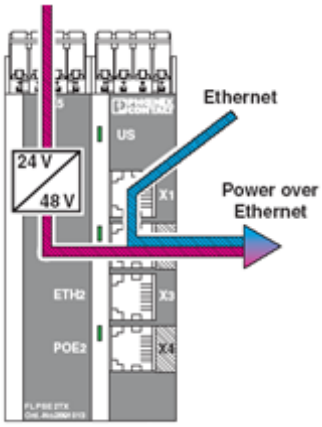


Figure 2.20: Building of a PSE

2 types are distinguished.

- End point PSE: the standard Ethernet switch is replaced with a PoE switch.
- Mid span PSE: this device is placed between the conventional switch and the network participant. Only functions on alternative B.

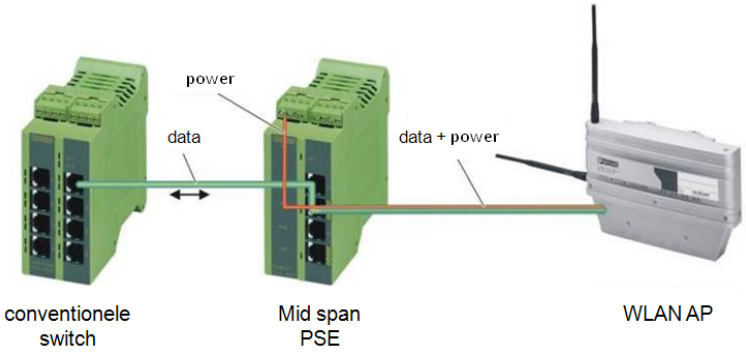


Figure 2.21: Use of a mid span PSE

Figure 2.21 shows the integration of a mid span PSE. An extra module is required every time to make PoE possible.

2.6.2 PD

The network participant that receives power over the Ethernet cable is called a PD (Powered Device). In order to avoid polarity related errors, an auto-polarity circuit is built in a PD. A PD has to support Alternative A and B, in accordance with the norm.

According to the norm, a PSE should be able to supply at least 15.4W and a PD cannot use more than 12.95W. This difference is used to cover the losses in the twisted pair cable. A 100 metres cable has a resistance that causes losses.
In order to protect devices against unexpected voltage, an identification process is carried out during the connection:

- If nothing is connected to the PSE, then the port will be dead
- A device reports a resistance of 25kΩ.
- The PSE applies a voltage of 10.1 V on the load and measures the power. If the required power is less than the minimum power, then no power is supplied.
- In order to determine the specific class from 0 to 3, the PSE applies a voltage of 20.5V on the load. After the determination of the class, the PSE applies a voltage of 48V on the load. The following protection classes are distinguished:

Class 0	0.44W to 12.95W
Class 1	0.44W to 3.84W
Class 2	3.84W to 6.49W
Class 3	6.49W to 12.95W

2.6.3 Alternative A

The power is transmitted via the data lines. The power is connected via transformers with center tap on pins 1-2 and 3-6 so that these are invisible for the data stream. Can be used for 10/100/1000Base-T.

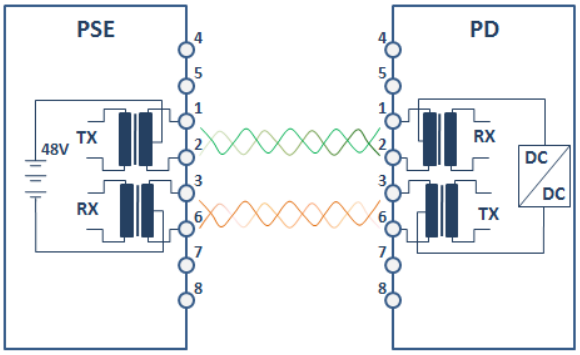


Figure 2.22: PoE, alternative A

2.6.4 Alternative B

The power is transmitted via the cores in a UTP cable that is not used for data. The pairs 4-5 and 7-8 are used in parallel so that more current can pass. The positive side of the 48V is connected to pin 4 and 5, the negative side is connected to pin 7 and 8.

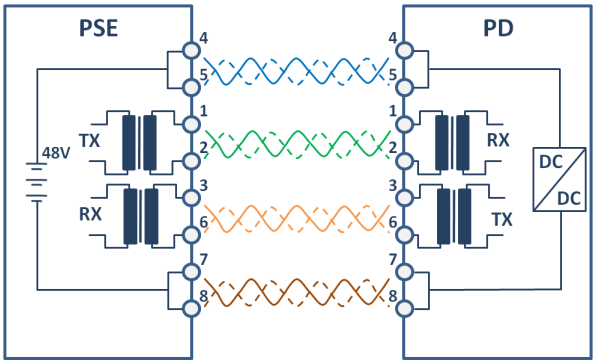


Figure 2.23: PoE, alternative B

Can only be used if pair 1 and pair 4 are available (certain industrial Ethernet cables only contain pair 2 and 3) or if pair 1 and 4 are not used (so 1000Base-T is not possible).

2.7 VLAN

A VLAN or Virtual Local Area Network is a group of participants in a large network that form a separate network in a logic manner. This means that several logic groups can be created on a large physical network. A VLAN has an own broadcast domain. Data packets are only transmitted within a VLAN. The participants may physically be far away from each other but have to be on one and the same physical network. Some examples of the organisation of a network:

- By department: one VLAN for Sales, another VLAN for Engineering and another VLAN for Automation.
- By hierarchy: one VLAN for management, another VLAN for managers and another VLAN for employees.
- By use: one VLAN for users that require e-mail and another for multimedia users

2.7.1 Advantages of VLANs

The greatest advantage of VLANs is the segmentation of the network. Other examples are additional security and restriction on network load.

- Moving devices around: it is easier to move devices around in the network In a traditional network, cabling has to be changed when a user moves from one subnet to another. The relocation from one VLAN to another does not require change in cabling. It only requires a setting on the switch. A station from Sales can, for example, be moved to a network connection that belongs to Engineering. The port has to be set up as a member of the VLAN Engineering but does not require new cabling.
- Additional security: devices of a VLAN can only communicate with devices in the same VLAN. If a device of VLAN Sales wants to communicate with the VLAN Automation, then this connection has to be set up in a router.
- Restriction on network traffic: for a traditional network, broadcasts can cause an overloaded network. Broadcast messages are often sent to devices that do not need these messages. VLANs limit this problem as a broadcast message from one VLAN is not sent to the other VLAN.

2.7.2 Trunking

Trunking is a method to send data from different VLANs between two switches. Only one port per device is required for this. There are different ways of trunking.

- ISL: InterSwitch Link, this is a widely used proprietary protocol of Cisco
- IEEE802.1Q: this is a standard that is supported by several switch manufacturers.

For trunking, a piece of code (tag) is added that states which VLAN the sent package comes from. Thanks to this system, the benefits of VLAN are retained. The VLANs remain separated, even if they are spread out over different switches. A router is needed to route data traffic between the different VLANs.

2.7.3 VLAN types

The different types of VLANs can be divided into two types: static and dynamic VLANs.

- Static VLANs are port-based. Depending on the port of a switch to which a user connects, this belongs to one or the other VLAN.

  Advantages:

  - Easy to configure
  - Everything is done using the switch. The user hardly notices anything

  Disadvantages:

  - If a user connects his PC to the wrong port, then the administrator has to do a reconfiguration.
  - If a second switch is connected to a port that belongs to a certain VLAN, then all computers that one connects to this switch will automatically belong to this VLAN.

- Dynamic VLANs: are not based on ports of a switch but on the address of the user or the used protocol.

  Advantage: everyone can connect his computer to any port and still be part of the correct VLAN.

  Disadvantage: the cost of this VLAN type is higher as it requires special hardware.

2.8 Network redundancy

2.8.1 Introduction

Network redundancy means the integration of hardware and software that ensures that the availability of the network remains optimal in case of a Single Point of Failure. The communication system - the network - is the core of every modern automation project. In order to handle network errors, different protocols can be integrated in the structure elements. Three important groups can be distinguished:

- STP/RSTP: (Rapid) Spanning Tree Protocol. Can be applied in mesh topologies (discussed elsewhere in this chapter).
- MRP: Media Redundancy Protocol, only for ring topologies.
- PRP: Parallel Redundancy Protocol

2.8.2 The Spanning Tree Protocol

The Spanning Tree Protocol (STP) is an open protocol that is described in the IEEE802.1d. It is an OSI layer-2 protocol that guarantees a closed loop free LAN. It is based on an algorithm developed by Radia Perlman (employee at Digital Equipment Corporation). Spanning tree makes it possible to extend a network whereby redundant links are integrated. This way, an automated back-up path is provided if an active link drops out for whatever reason, without creating closed loops in the network.

In order to apply this protocol, the used switches have to support the protocol. After the interruption of a segment, it can easily take 30-50 seconds before the alternative path becomes available. This delay is unacceptable for controls and 30 seconds is extremely long for monitor applications. An advantage of the STP is that it cannot be used for redundant ring structures.

2.8.3 The Rapid Spanning Tree Protocol

As a reply to the shortcomings of the spanning tree protocol, the IEEE formulated, in 2001, the rapid spanning tree protocol (RSTP). The protocol is described in the IEEE802.1w standard. Since 2004, the spanning tree protocol is described as superfluous in the IEEE802.1d and it is recommended to use the RSTP instead of the STP. The IEEE802.1w is therefore included in the 802.1d norm.

The recovery time of the RSTP is lower than of the STP (thus the name), namely 1 to 10 seconds instead of 30-50 seconds. Depending on the application, this recovery time may already be rather too quick.

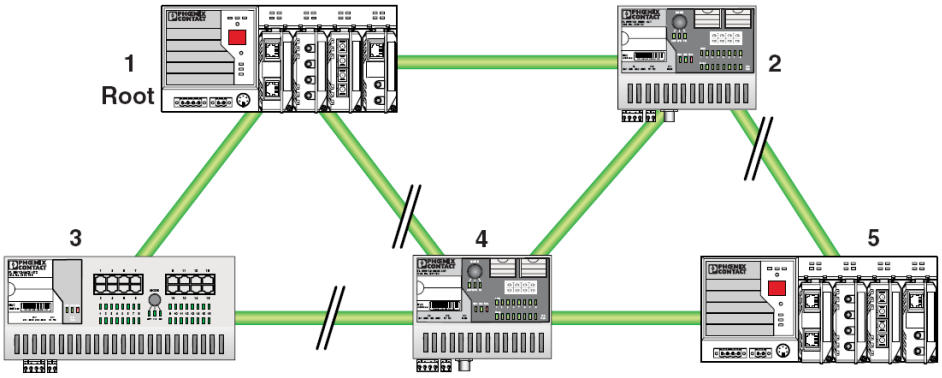


Figure 2.24: Possible tree topology by means of (R)STP

Figure 2.24 shows a network with five different structure elements. Different redundant connections are created. This results in the occurrence of unacceptable loops that will quickly congest the network. The RSTP protocol converts this topology into a tree structure by closing off a number of ports. One structure element is configured here as root. From this root, all other switches can be reached via one single path. If a network error occurs, then a new active path is created.

Extensions on RSTP

In order to meet the needs of the automation, many companies plan proprietary extensions on the RSTP protocol in order to attain recovery times of less than a second. This way, QoS is obtained for the redundant building of automation networks.

Fast Ring Detection is an extension of Phoenix Contact on the RSTP. When a network switch drops out, recovery times of 100... 500 ms are reached. Recovery times of not more than 500 ms are available for extensive automation networks with 1000 entered address tables in the switches. These times are shorter in case of less number of terminals in the network. This protocol, however, can only be used for 10 or 200Mbps.

2.8.4 Bridge Protocol Data Units (BPDUs)

The tree structure is calculated by means of a specific algorithm so that there is one switch configured as the root. Every switch must in fact have all the information required so as to be able to define the correct port lines. In order to ensure that every switch has sufficient, correct information, switches exchange information between them. Special frames are used for this, Bridge Protocol Data Units (BPDUs).

A bridge sends a BPDU, with it as the SA using the unique MAC of the port itself and, as the DA, the STP Multicast address 01:80:C2:00:00:00. There are various types of BPDU:

- Configuration BPDU (CBPDU) used for the calculation of the spanning tree
- Topology Change Notification BPDU (TCN), used to notify changes in the network
- Topology Change Notification Acknowledgement (TCA)

To create a network without loops, each port is allocated a specific status by a switch. The various statuses are:

- ROOT: port that forms the link to the root switch
- DESIGNATED: an active port that forms a link to an underlying switch in the tree structure.
- ALTERNATE: a port with a lower priority, an alternative link to the root

2.8.5 Multiple Spanning Tree Protocol (MSTP)

In an Ethernet environment where Virtual LANs are used, the Spanning Tree Protocol may also be used.

MSTP, originally defined in IEEE 802.1s and later adopted in IEEE 802.1q Version 2003, defines an extension of RSTP in combination with Virtual LANs. This combines the best of PVST (Per-VLAN Spanning Tree) in which each VLAN defines its own spanning tree and the original IEEE 802.1q with which only one spanning tree is created for the entire network.

With MSTP, various VLANs are split into logical instances (groups of VLANs with the same spanning tree topology).

MSTP bundles all the spanning tree information into one single BPDU in order to restrict the number of BPDUs and it is fully compatible with RSTP switches.

2.8.6 Media Redundancy Protocol

MRP is part of the PROFINET standard. In case of MRP, a ring manager blocks one port in order to obtain an active line structure. In case of a network error, the network splits up into two isolated lines that are linked together again when the blocked port is released. Recovery times are in the range of 100ms.

2.8.7 Parallel Redundancy Protocol

In contrast with the above technologies, PRP does not plan a change of the active topology in case of a network error. This protocol functions on two parallel networks. Every data frame is sent over the two networks. The receiving node processes the message that arrives first and rejects the copy message. PRP ensures the copying and rejection of the messages. PRP also makes the double network invisible for the higher layers in the communication stack.

2.9 Important additions

2.9.1 LLDP

The protocol IEEE802.1ab, link layer discovery protocol, is a standard that provides a solution for the configuration problems faced by extensive LAN structures. It defines a standard method for switches, routers, WLAN access points, ... in order to distribute information about itself to other network participants and to save information from neighbouring participants. LLDP is possible with all 802 media. It uses the data link layer.

A switch that support LLDP can carry out topology detection via other participants that also support LLDP.

Advantages:

- Improved detection of network errors
- Tool for replacing modules
- Better network configuration and network management

LDDP information is used within engineering tools to visualise a network topology in a graphic manner.

2.9.2 IEEE 802.1x

IEEE 802.1x is a security standard for authentication on each individual port of a switch. Authentication occurs even before the user is given access to the network. The recognition of an authorised user therefore occurs at Layer 2 of the OSI model. This can all be done, regardless of the hardware used, both wirelessly and with wires.

IEEE 802.1x uses a protocol to exchange information with a device/user that sends a request for access to a port. The messages contain a user name and a password. The switch does not perform the identification itself but forwards the request to a RADIUS authentication server on the network. The server processes the request and gives feedback to the switch, which then opens the port for the user.

There are three important players in the operation of the protocol:

- The user, the client, is known in the protocol as the supplicant.
- The access device, the switch or the access point is the authenticator.
- The controlling device, the RADIUS infrastructure, is an authentication server.

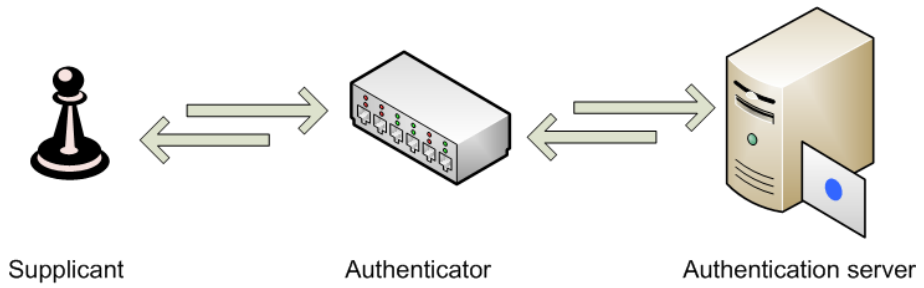


Figure 2.25: Three important players in the operation of the protocol

The authentication for 802.1X is performed with a flexible authentication mechanism called the Extensible Authentication Protocol (EAP), under which various forms of authentication are possible. This makes it possible to demand another form of authentication depending on the type of user: both strong and weak authentication. It is for instance possible to make a user name and a password compulsory for students and for staff to use a certificate. This aspect is dealt with in more detail on the section concerning security.

2.9.3 Link Aggregation with LACP to IEEE 802.3ad

Link Aggregation (also called trunking) is a way of physical network linking and is the English term for the aggregation of several network connections with the aim of achieving higher transfer speeds. Link Aggregation can also provide a redundant connection, which adds fault tolerance to critical business systems. The technique is applied both to switches and to network interface cards (NICs).

Link Aggregation is currently standardised in the IEEE 802.3ad standard. It offers the following advantages:

- Increased availability of the connections
- Capacity of a connection increased
- Higher performance with the available hardware

Today's LAN technologies provide data rates of 10 Mbps, 100 Mbps and 1,000 Mbps. Link Aggregation can create levels in between where necessary or if a data rate greater than 1,000 Mbps is required, a high-speed connection can be provided by the grouping of several 1,000 Mbps connections.

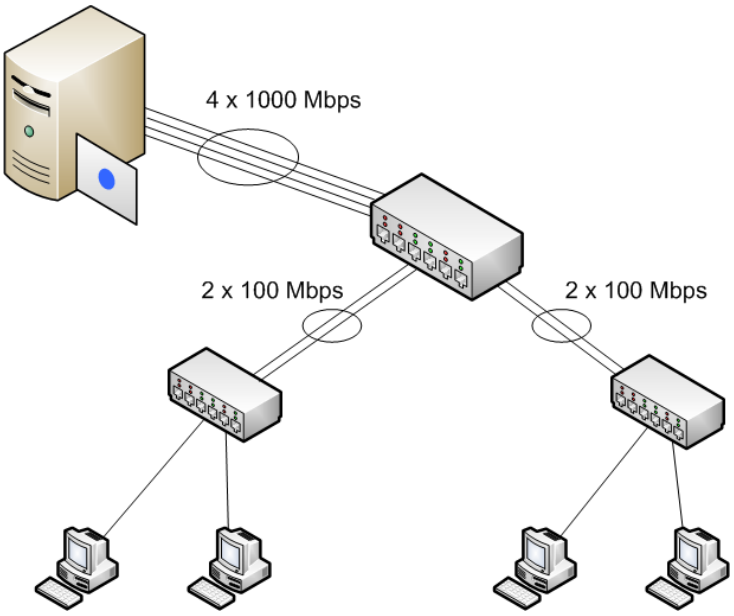


Figure 2.26: Link Aggregation

Link Aggregation may be used in several ways:

- Connection between two switches
- Connection between switch and end station
- Connection between two end stations

Diagram xxxx shows how switches are connected to each other via two 100 Mbps links. If one link is lost between the two switches, the other link in the link aggregation group will take over.

Link Aggregation is currently included in the IEEE 802.3ad standard: "Link Aggregation allows one or more links to be aggregated together to form a Link Aggregation Group, such that a MAC client can treat the Link Aggregation Group as if it were a single link' (IEEE Standard 802.3, 2000 Edition).

The IEEE 802.3ad standard also describes the use of the LACP (Link Aggregation Control Protocol) for easy exchange of configuration information between the various systems. This should make automatic configuration and also the monitoring of all link aggregation groups possible. This information exchange is by means of LACP frames as described in the standard.

2.10 Industrial Ethernet

In recent years, Ethernet is used more and more in an industrial environment. There are significant differences between the office environment and the industrial environment. Industrial Ethernet refers to the use of industrial products in order to meet the more specific requirements of the industrial world. The tables below show some important points of attention.

Table 2.4: Points of attention for installation of Ethernet

Office environment	**Industrial environment**
Fixed basic installation in the building	system-specific applications
Variable network connection for workstations	connection points with the network are seldom or never changed
Cables are placed in false floorings	
Prefabricated cables	connectors that can be assembled on the shop floor
Standard workstations on RJ45	RJ45 in the cupboards, M12 in the field
	regular use of fibre optics
	and cabling intended for use in moveable cable conductors
	carefully implemented earthing
230V AC power	24V DC power / Power over Ethernet (PoE)
Star topology	regular use of line topology or ring topology
	redundancy is often a requirement
19" switch cabinets (dimensions of standard office server cabinets)	
Service life of about 5 years	service life of about 10 years
	terminals suitable for assembly on a DIN rail
Devices with active cooling (ventilators)	passive cooling (fixed parts)
	alarm contact for error indication
	carefully implemented earthing

Table 2.5: Environmental effects

Office environment	**Industrial environment**
Moderate temperatures with low fluctuations	external temperatures with high fluctuations
Hardly any dust	a lot of dust
No humidity or water	humidity or water can be present
Hardly any shocks or vibrations	Vibrations or shocks are possible
Low level of EMC	High EMC level
Low mechanical load or danger	High mechanical load or danger
No chemical hazards	Chemical impact due to oily environments or aggressive atmosphere
No radiation hazards	High exposure to UV radiation in outdoor environments

Chapter 3

TCP/IP

3.1 Introduction

Transmission Control Protocol / Internet Protocol (TCP/IP) is a collection of industrial standard protocols designed for communication over large networks consisting of different network segments that are linked by routers.

TCP/IP is a protocol used on Internet, that is the collection of thousands of networks, spread out worldwide, that connects research centers, universities, libraries, companies, individuals, etc. with each other.

Internetworks is however, a very general concept. An internet is not limited in size: there are internets that consist of a couple of networks but also internets that consist of hundreds of networks. Internet with capital letter I relates to the worldwide Internet- also called the public Internet.

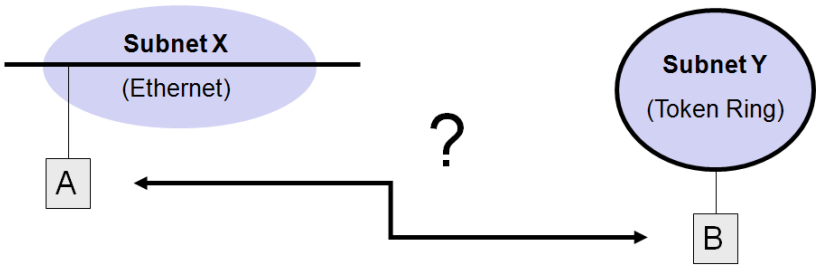


Figure 3.1: How to communicate over an internet?

The question that has to be asked is how two different hosts, connected to a different network, at a great distance, can communicate with each other? The answer to this question is twofold.

The first part of the answer is a hardware aspect: an internet consists of different networks that are connected to each other by routers. A router is a structure element with the special task to connect networks. Every router has a processor, a certain amount of memory and a separate interface for each network to which it is connected.

The second part of the answer is a software aspect: a universal communication service has to be active on every host. Although many software protocols are adapted for internet works, only one suite is really considered and that is used most for internet works. This suite is called TCP/IP suite.

The TCP/IP suite can be perfectly positioned in the OSI model. A simplified model is mostly used to represent the TCP/IP suite. A four layer model (the DoD (Department of Defence) model), the ARPANET Reference model or mostly just called the TCP/IP model.

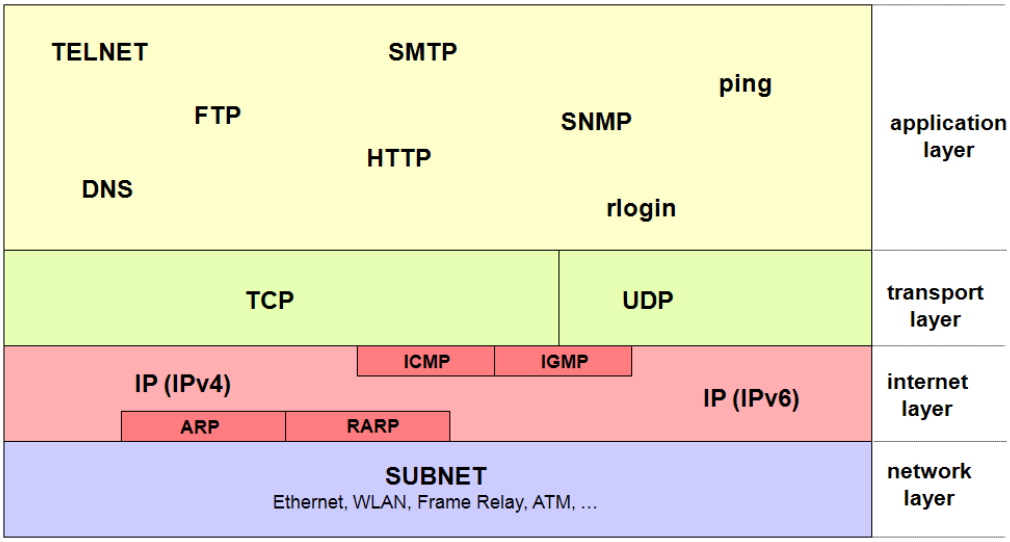


Figure 3.2: The TCP/IP suite

Central in this model is the internet layer and the transport later which is discussed in detail elsewhere in this chapter. The application layer collects and describes all protocols that use the TCP/IP protocol. The HTTP protocol belongs to this, for example. This is the protocol that makes it possible to surf to a certain web application. The TCP/IP protocol will then make a universal communication service possible so that the surf order is possible over the entire Internet. The network layer ensures the communication on the local network between the host and the router or between two routers mutually.

3.2 The Internet Protocol (IP)

3.2.1 Introduction

The most important features of the IP protocol are:

- Routing of a data packet over the Internet. Every host is identified by a 32-bit IP address.
- it is a connectionless protocol. Every packet can follow a different route to the same target host when sending different IP packets. No fixed physical connection is created.

- A universal data packet is built, consisting of a header and a data field. The header consists, amongst others, of the address of the sender and of the destination. The data packet is hardware independent and is encapsulated again on the local network before it can be transported.
- The IP protocol does not check if the data have been sent correctly and it also does not provide confirmation- or correction mechanisms: send-it and pray.
- The IP header is at least 20 bytes long. When using the options field, the header can be 60 bytes maximum. A header check sum is created.

The Internet Protocol(IP) is applicable to the network level (layer 3 of the OSI model). This layer is responsible for the presentation and transportation of information over different networks. Uniform addressing is required to realise this: the IP address.

This functionality is not required for as long as the information transfer takes place within the same network. The connection of different networks takes place by means of routers. When different networks are bundled into a bigger entity, each network should also be identifiable with an address. Each network will therefore get a unique network address. With this network address, every participant also gets a unique address number within this network address. Uniform addressing is based on this principle. This address is defined on the IP layer and is called IP address.

3.2.2 The IP address

General

An IP address consists of 32 bits, 4 bytes, represented as 4 decimals separated by a dot.

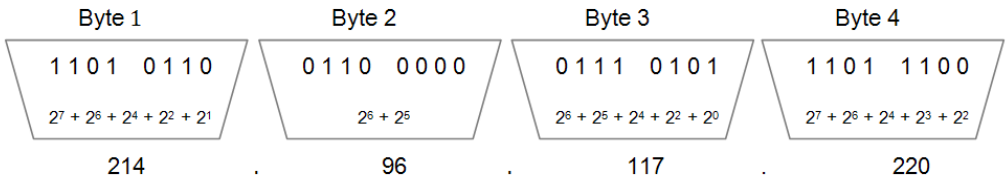


Figure 3.3: The IP address

Every network gets a name (Net ID) and every network participant gets a unique number within this network (Host ID). Net ID and Host ID together form the IP address. The network name is then the IP address whereby the Host ID is equal to zero.

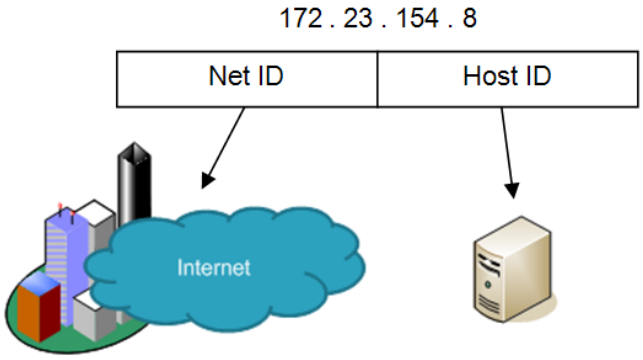


Figure 3.4: Building of the IP address

Classes of IP addresses

IP addresses are divided into different classes. Figure 3.5 shows an overview.

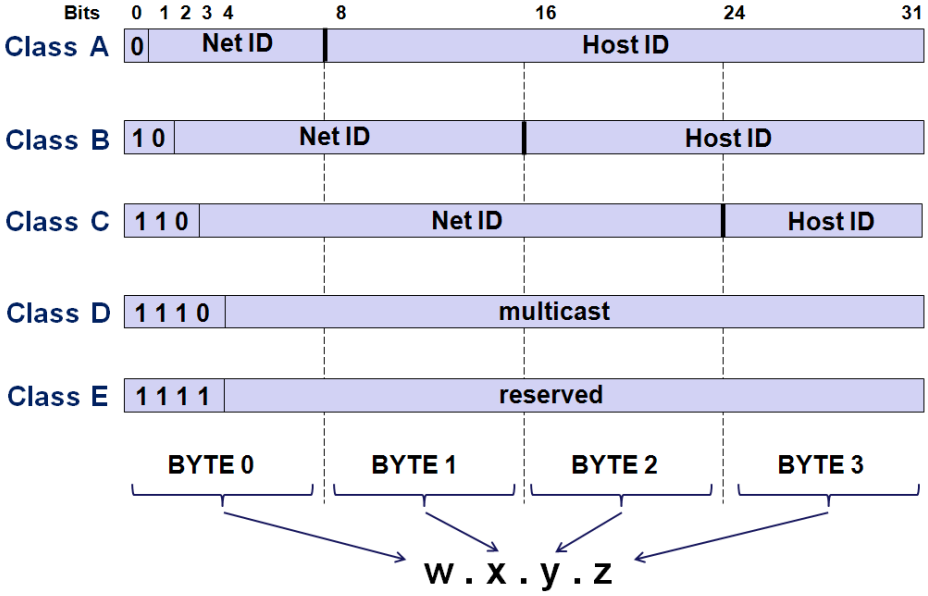


Figure 3.5: The different classes within IP addressing

Table 3.1 shows the features of class A, B and C. Class D is added in order to send multicast messages in a simple way. Class E currently has no function yet.

The distinction between the classes A, B and C is determined by the number of bytes that are part of the Net ID on the one hand and the number of bytes that are part of the Host ID on the other hand. The largest number of bits in the IP address determine to which class an IP address will belong. Table 3.1 sums up all features of the three different classes.

Table 3.1: Features of the different classes

Class A	Net ID	byte 1, first bit is always 0, (0 x x x x x x x) 126 possible network addresses
	Host ID	byte 2 + byte 3 + byte 4 16777214 possible hosts per network
	Range	1 . n . n . n → 126 . n . n . n
	Example	90.15.167.2 (network name 90.0.0.0)
Class B	Net ID	byte 1, first bits are always 1 0, (1 0 x x x x x x) + byte 2 16383 possible network addresses
	Host ID	byte 3 + byte 4 65534 possible hosts per network
	Range	128 . 0 . n . n → 191 . 255 . n . n
	Example	128.19.205.132 (network name 128.19.0.0)
Class C	Net ID	byte 1, first bits are always 1 1 0, (1 1 0 x x x x x) + byte 2 + byte 3 2097152 possible network addresses
	Host ID	byte 4 254 possible hosts per network
	Range	192 . 0 . 0 . n → 223 . 255 . 255 . n
	Example	192.147.25.112 (network name 192.147.25.0)

The use of IP addresses is controlled by the Internet Assigned Number Authority (IANA).

IP addresses for private networks

Distinction is drawn between public networks and private networks (corporate networks). On the Internet (the whole of public networks), every IP address has to be unique. Corporate networks are linked to the Internet via a router. In order to avoid conflicts between private networks and public networks, a series of IP addresses that are not used on the Internet are defined within each class. These are described in RFC 1597, Reserved Address Space. A corporate network preferably uses a value from this series as name, see Table 3.2 .

Table 3.2: IP addresses for private networks

Class A networks	10.0.0.0 → 10.255.255.255
Class B networks	172.16.0.0 → 172.31.255.255
Class C networks	192.168.0.0 → 192.168.255.255

Special IP addresses

Table 3.3 gives an overview of the special IP addresses

Table 3.3: Some important special IP addresses

Net ID	Host ID	Description
all zeroes	all zeroes	IP address of this computer is always used during the start
Net ID	all zeroes	The network address identifies a complete network
Net ID	all ones	Broadcast address on the network
127	random	IP address for the testing of network applications

An option has to be planned to send broadcast messages on a network. An IP broadcast address for a certain segment is obtained by setting all bits of the host ID to 1. The IP address 131.107.255.255 is a network broadcast address of the subnet with network address 131.107.0.0 .

3.2.3 Routers and subnet masking

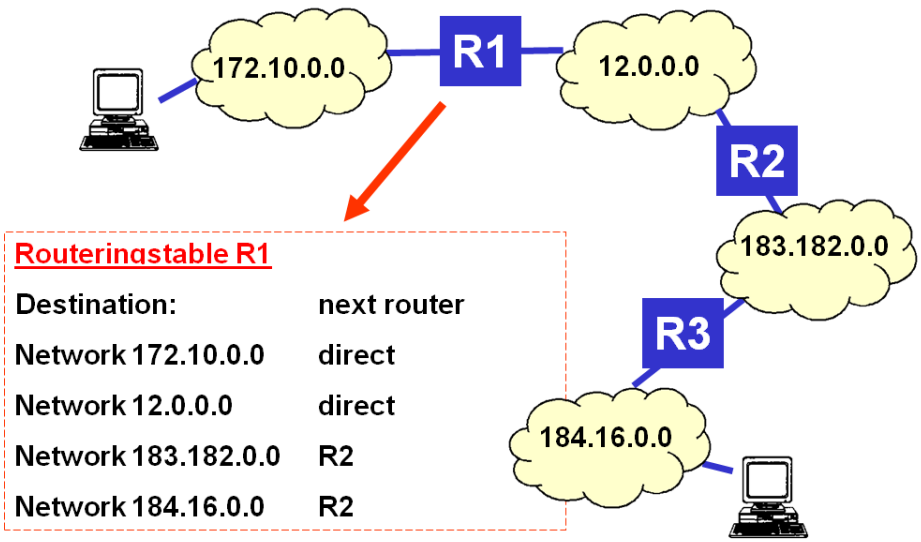


Figure 3.6: Functioning router

Despite the fact that the Internet is indicated in singular form, it consists of a large number of IP networks. Every ISP (Internet Service Provider) hooks up its network to at least one other network. As every network has its unique identification, the information can be sent from one to the other station. Routers ensure that information is routed correctly over the Internet. These routers maintain so-called routing tables in which is specified where certain IP addresses can be found. As soon as an IP packet arrives, the router compares the target address with its routing tables. If a correspondence is found, then the router knows to which port the relevant packet can be sent.

In order to facilitate the routing and to use the existing classes in a better way, the option was given, in 1985 with the RFC 950, to create groups of addresses within a class A, B or C. The prefix (NetID) is extended with a few bits (an extended network prefix) for the creation of a number of subnets within a class. The IP address does not change when using subnets. However, for the routers it is important to know which bits now form the NetID. The router uses a subnet mask for this purpose. The router filters the network part from the IP address by means of this mask.

How is the subnet mask set up?
The bits that represent the network part have the value 1
The bits that represent the host part have the value 0
Then the decimal conversion follows.

Example: a class C address is extended with four network bits, the subnet mask then becomes:

11111111 . 11111111 . 11111111 . 11110000
255 . 255 . 255 . 240

3.2.4 Subnetting

Subnetting is to generate several subnets from a given IP address.
Example: a company works with IP address 172.23.0.0 (class B).
Subnet mask is 255.255.0.0 or 1111 1111 1111 1111 0000 0000 0000 0000

The entire company has to be divided into 10 different subnets. All subnets can be linked to each other via a router.

With 4 bits, 16 different combinations can be created.
The required 10 subnets can be created by adding 4 bits to the Net ID.
Subnetmask 1111 1111 1111 1111 **1111** 0000 0000 0000 or 255 . 255 . 240 . 0

The following 10 subnets can be created in this way. Table 3.4 shows the different subnets.

Table 3.4: Subnetting and subnet masks

BYTE 3 (binary code)	BYTE 3 (decimal value)	Subnet	Subnetmask
0000 0000	0	172.23.0.0	255.255.240.0
0001 0000	16	172.23.16.0	255.255.240.0
0010 0000	32	172.23.32.0	255.255.240.0
0011 0000	48	172.23.48.0	255.255.240.0
0100 0000	64	172.23.64.0	255.255.240.0
0101 0000	80	172.23.80.0	255.255.240.0
0110 0000	96	172.23.96.0	255.255.240.0
0111 0000	112	172.23.112.0	255.255.240.0
1000 0000	128	172.23.128.0	255.255.240.0
1001 0000	144	172.23.144.0	255.255.240.0

3.2.5 Classless Inter-Domain Routing

The success of the Internet may lead to a shortage in IP addresses. The increasing number of networks also lead to a quickly increasing number of routes which will also cause a problem for global routing tables.

The solution to this problem consists of two steps:

- Restructuring of the IP addresses
- Hierarchical route structure in order to make the routing more efficient

CIDR, Classless Inter-Domain Routing is a new way of addressing for the Internet which would lead to a more efficient use of IP addresses in comparison to the classes A,B and C. This results directly from the subnetting concept.

The Net ID is not limited anymore to 8,16 or 24 bits. A CIDR address contains the 32-bit IP address and information about the number of bits that are part of the Net ID. In the address 206.13.01.48/25, the suffix '/25 means that the first 25 bits determine the network name and that the remaining bits are used to identify a certain participant on the network.

Table 3.5: Classless Inter-Domain Routing

CIDR code	subnet mask	binary	number of hosts
/28	255.255.255.240	11111111 11111111 11111111 11110000	16
/27	255.255.255.192	11111111 11111111 11111111 11100000	32
/26	255.255.255.192	11111111 11111111 11111111 11000000	64
/25	255.255.255.128	11111111 11111111 11111111 10000000	128
/24	255.255.255.0	11111111 11111111 11111111 00000000	256
/23	255.255.254.0	11111111 11111111 11111110 00000000	512
/22	255.255.252.0	11111111 11111111 11111100 00000000	1024
/21	255.255.248.0	11111111 11111111 11111000 00000000	2048
/20	255.255.240.0	11111111 11111111 11110000 00000000	4096
/19	255.255.224.0	11111111 11111111 11100000 00000000	8192
/18	255.255.192.0	11111111 11111111 11000000 00000000	16384
/17	255.255.128.0	11111111 11111111 10000000 00000000	32768
/16	255.255.0.0	11111111 11111111 00000000 00000000	65536
/15	255.254.0.0	11111111 11111110 00000000 00000000	131072
/14	255.252.0.0	11111111 11111100 00000000 00000000	262144
/13	255.248.0.0	11111111 11111000 00000000 00000000	524288

The CIDR addressing also makes ’route aggregation’ possible. One high-level route change can represent many lower-level routes in a global routing table. This way, a complete hierarchical structure can be elaborated which is comparable with the zonal division of telephone numbers.

3.2.6 Examples

- Show that the server with IP address 203.125.72.28/28 and the server with IP address 203.125.72.34/28 do not belong to the same network.
- The given IP address of a host is 192.168.100.102/27.
  - Show that this host belongs to the network: 192.168.100.96/27.
  - Show that the broadcast address of this network 192.168.100.127 is.
  - show that all participants on this network have an IP address which lies between 192.168.100.97 and 192.168.100.126.
- A corporate network is composed of different subnets.
  The participants with the following IP addresses belong to three different subnets: 172.23.136.45, 172.23.139.78 and 172.23.140.197.
  The participants with IP address 172.23.126.120 and 172.23.127.92 do belong to the same subnet.

  Show that the CIDR suffix /23 will be within the corporate network.

3.2.7 The IP packet

The data to be sent are transmitted by the transport layer to the internet layer. The internet layer packs the information in the data field and then adds an IP header. The whole is then transmitted to the network layer for further processing. The sending of information by means of the IP protocol takes place on the basis of IP packets.

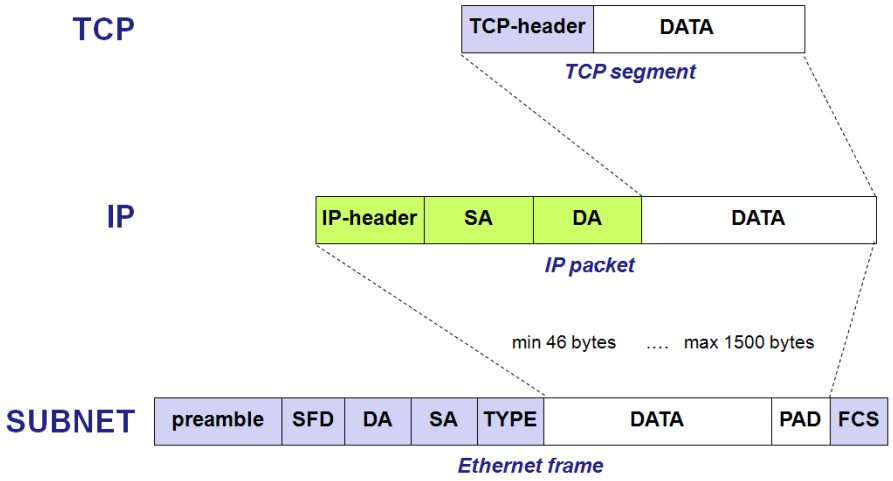


Figure 3.7: The IP packet

When a router receives an Ipv4 packet that is too big for the subnet on which the packet has to be transmitted, then IPv4 will divide this packet on the router into smaller packets that fit in the data frames of the relevant subnet. When these packets reach their final destination, IPv4 will put these packets back into the original order on the target host. When a packet has to be divided, the following takes place;

- Each packet has its own IP header
- All divided messages that belong to the same original message have the original identification field. The *more fragments flag* shows that other fragments follow. The *more fragments flag* is not placed in the last fragment.
- The *fragment offset field* specifies the position of this fragment in the total message.

In order to get a good idea of the functions of the IP protocol, the IP header is further explained. Figure 3.8 shows the different fields within the IP header. The header consists of at least 20 bytes.

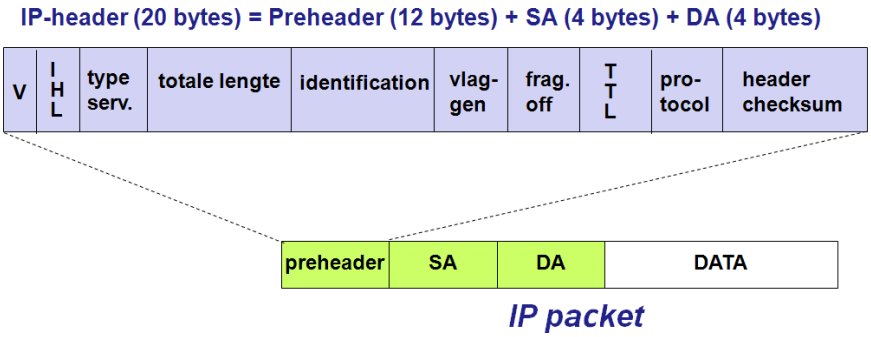


Figure 3.8: The IP header

- Version (V): field of 4 bits that represents the IP version.
- IHL: field of 4 bits that represents the length of the header (in bytes)
- Type of service: reserved/priority of the required service
- Total length: the total length in bytes of the complete IP packet
- Identification: if an IP packet has to be divided, then each packet gets a unique identification so that all packets can be merged back correctly on the receiving side.
- Flags: the flags are used to follow-up the fragmentation of the packets
- Fragment offset : when a data packet is divided then the position of the fragment in the entire packet is a 8-bit unit.
- Time to live (TTL): every time an IP packet passes a router, this value is reduced by 1. If this number is 0 then the relevant router will reject this message. This avoids that a message can exist forever.
- Protocol: the higher level protocol is represented here

  01h ICMP
  06h TCP
  11h UDP

- Header Checksum: a check value for the IP header. Every router will recalculate this header checksum.
- Source IP address: IP address of the sending participant
- Destination IP address: IP address of the receiving participant
- Options : other network information can be included in the IP header. If the options data do not end with a 32-bit word, then the rest is filled with padding zeroes.

3.2.8 IPv6

General

The most recent IP protocol that is discussed in this chapter has version number 4 (IPv4). The great success of the IP protocol results in the need for a new version. There is a pressing shortage of IP addresses and furthermore it is also important that new functionalities can be integrated in a simple way. Moreover, a new version of the IP protocol has to be able to guarantee a higher performance.

With the introduction of IPv6, there is also a practical problem: how will the publicly accessible Internet that functions on Ipv4 switch to Ipv6? The easiest way is the so-called dual-stack approach. This means the implementation of IPv6 and IPv4 on the nodes. These nodes can process IPv4 as well as IPv6 datagrams.

Currently, the automation world is not interested in the integration of IPv6.

Some features of the IPv6 are discussed below. However, as many IPv4 design features as possible are retained as these have made this version so successful.

IP address

IPv6 uses IP addresses of 128 bits. This creates extensive addressing options. Hexadecimal notation with double dots, addresses become 128 bits long: 8 groups of 4 hexadecimal digits

2000:0000:0000:0FED:CBA9:8765:4321
2000::FED:CBA9:8765:4321
IPv4 addresses ::192.32.20.46

The new addressing has to result in smaller routing tables.

IPv6 header

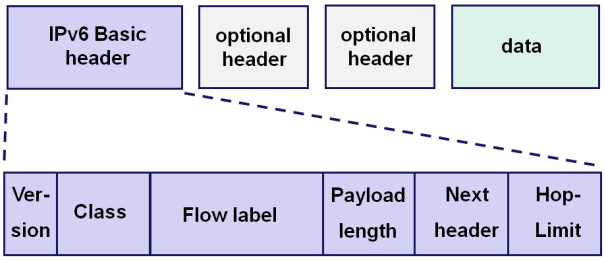


Figure 3.9: IPv6 header

The IP header is completely changed. A simpler basic header in combination with the option to integrate optional headers has to ensure that the header-processing time for the router is greatly reduced. A number of IPv4 fields are deleted or are only still available as option. The fields in the IPv6 header:

- Flow label: a 20-bit identification number to distinguish a packet in a data stream

- Hop limit: number of routers that can process a certain packet is limited
- Next header: defines the type of the first optional header
- Version field: This 4-bit field indicates the IP version number. For IPv6, this is value 6.
- Payload length field: this 16-bit number is an unsigned integer with which the number of bytes in the IPv6 datagram is indicated that follows after the standard header with a length of 40 bytes.

As the transport layer (TCP and UDP) and the data link layer (e.g. Ethernet) calculate protocols on the Internet checksums, the designers of IPv6 have decided that the calculation of checksums on the internet layer is no longer necessary.

3.3 Transmission Control Protocol (TCP)

3.3.1 Introduction

IP is a connectionless packet delivery service. TCP will have a difficult task. Using the unreliable packet service of IP, it has to provide a reliable data delivery service to different application programmes. For many applications it is essential that a transport system offers reliability: the system has to guarantee that data are not lost, cannot be duplicated or do not arrive in the right order.

3.3.2 End-to-end transport service

The TCP protocol is responsible for the correct sending of information over one or more networks. The exchange form of TCP is known as *connection oriented*: a logical connection is established, used and then stopped again. TCP is therefore a *end-to-end protocol*. Figure 3.10 shows why TCP is an end-to-end protocol. TCP sees IP as a mechanism with which TCP can exchange data on a certain host with TCP on a remote host.

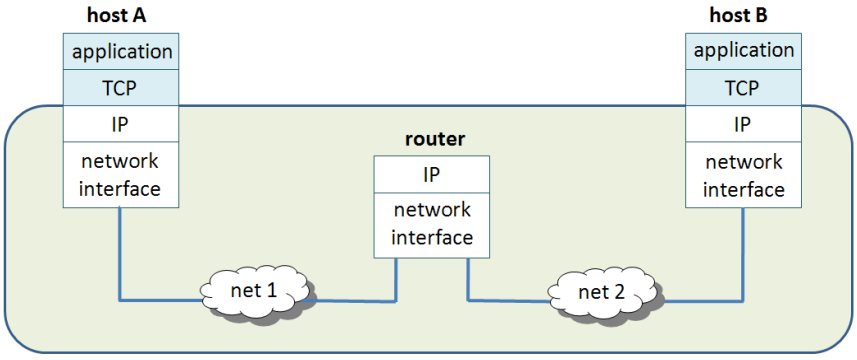


Figure 3.10: TCP if end-to-end transportprotocol

From the TCP point of view, the complete Internet is a communication system that can accept and deliver messages without changing or interpreting the content.

3.3.3 How reliability is achieved

TCP is a library with routines that applications can use when they want to start a reliable communication with another participant or host.
TCP uses different techniques to guarantee complete reliability.

Resending datagrams: when TCP receives data, it sends an acknowledgement to the sender. Every time that TCP sends data it starts a timer. If the timer ends before the confirmation was received, then the sender sends the data again, also see figure 3.11.

Window mechanism or organise the data stream. When a connection is made, each end of the connection reserves a buffer for the incoming and outgoing data and sends the size of the buffer to the other end. The available buffer space at a given moment is called window and the notification for the specification of the size is called window advertisement. A receiver sends a window advertisement for every acknowledgement. If the receiving application can read the data as fast as that they arrive, then the receiver sends a positive window advertisement with every confirmation. When the sending end however, works faster than the receiving end, then the incoming data will eventually fill the buffer of the receiver. This results at a certain moment in a zero window advertisement from the receiver. A sender that receives a zero window advertisement has to stop the sending until the receiver sends a positive window advertisement again.

Three-way handshake: In order to guarantee that connections are made and ended in a reliable way, TCP uses a three-way handshake in which three messages are exchanged. TCP uses the term synchronisation segment (SYN segment) for messages in a three-way handshake that is used for the setting up of the connection, and the term FIN segment for the description of message in a three-way handshake with which a connection is closed.

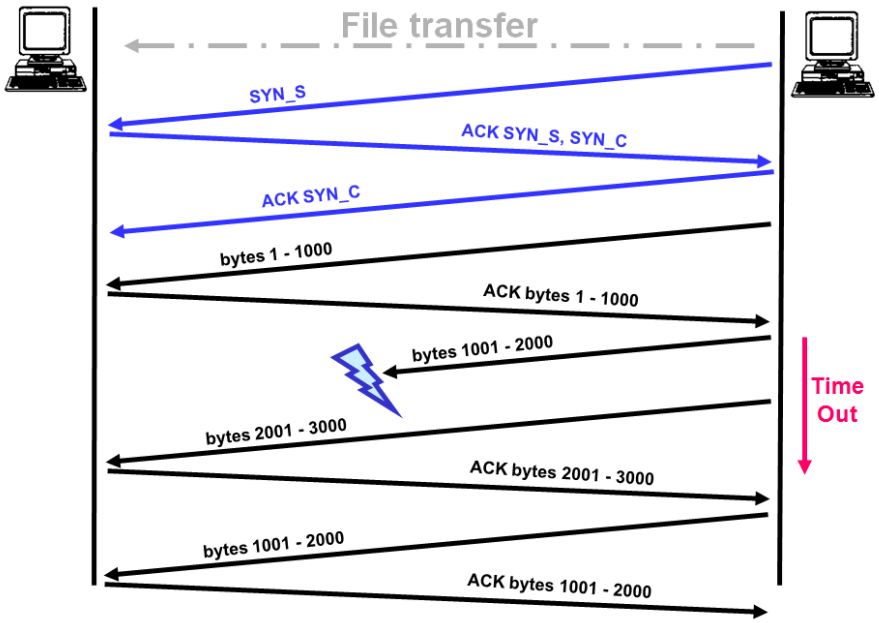


Figure 3.11: Three-way handshake

3.3.4 The TCP segment

The data to be sent are transmitted by the application layer to the transport layer. The transport layer packs the information in the data field and then adds a TCP header. The whole is then transmitted to the internet layer for further processing. The sending of information by means of the TCP protocol takes place on the basis of TCP segments.

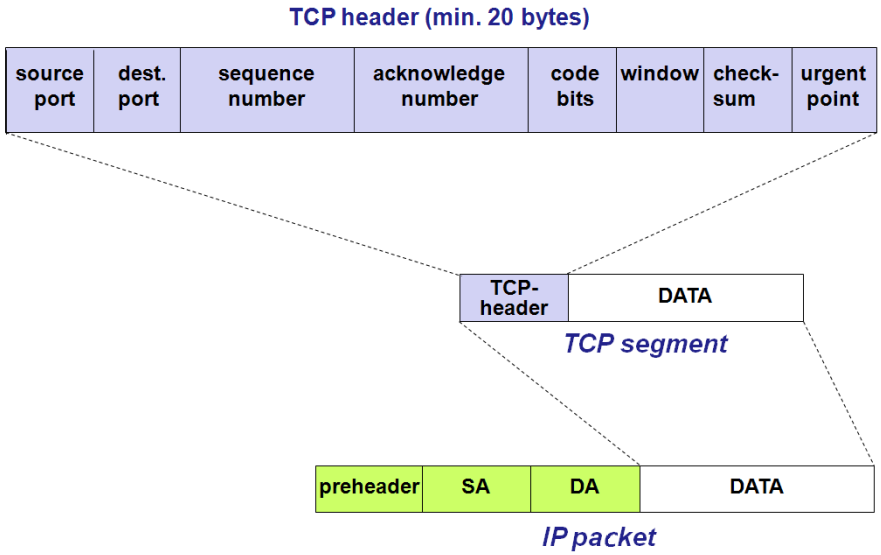


Figure 3.12: Building of a TCP segment

In order to get a good idea of the functions of the TCP protocol, the TCP header is further explained. Figure 3.12 shows the different fields within the TCP header. The header consists of 20 bytes.

- Source, Destination Port: for the different upper-layer applications, TCP is accessible via different port numbers. *Ports* is a unique 16-bit address. The combination of a port and an internet address is called a socket, in accordance with the original definition of a socket as defined by ARPA (1971). The use of port numbers is essential to build up a communication between the different applications. This is discussed in further detail elsewhere in the chapter. This course also includes table 3.6 that gives an overview of often used ports within the automation.
- Sequence number: Every byte of TCP has a number. The sequence number is the number of the first data byte in the TCP segment after the TCP header.
- Acknowledgement number: this field contains the next sequence number that is expected from the partner.
- Header Length: length of the TCP header in 32-bit words
- Code bits: different bits with which a number of statuses can be included.
  - the RST bit to initialise the communication again.
  - the SYN bit that is used to start a communication

  - the FIN bit that is used to indicate that a communication can be ended.

- Window: the window field indicates the maximum number of data bytes that can be sent before a confirmation is sent and received.
- Checksum: is a check value of the TCP packet
- Urgent Pointer: the value indicates where in the data field the urgent information starts. In order to include urgent information in a TCP packet, the URG code bit has to be set.

3.4 UDP

The protocol suite of the Internet also has a connectionless transport protocol, namely the UDP (User Data Protocol). With the UDP, applications can send IP packets without setting up a connection. Many Client Server applications that have one request and one answer use UDP instead of having to set up a connection and cancel this again later on. UDP is described in RFC 768.
UDP is more or less a zero protocol: the only services that it provides is a checksum for the data and the multiplexes of applications via port numbers. The UDP header is thus much simpler than the TCP header.

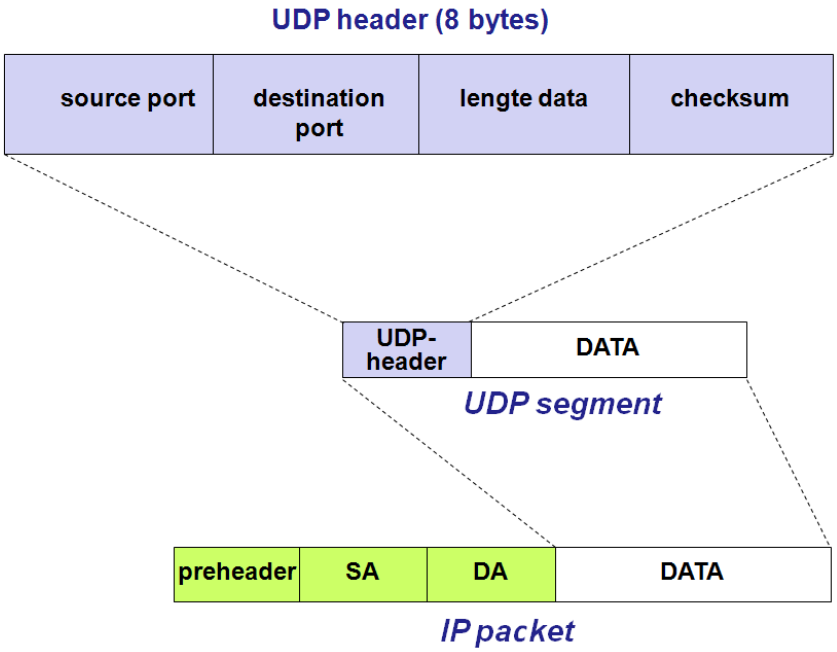


Figure 3.13: The UDP segment

A UDP segment consists of a header of 8 bytes followed by the data.
The header consists of:

- Source port (2 bytes): port number of the sender; this is equal to zero if no port is used.
- Destination port (2 bytes): port of the application to which this message is destined.
- Length (2 bytes): the length in bytes of the UDP header and the encapsulated data.
- Checksum (2 bytes)

A typical example of UDP is real time audio loosing data packets is a shame but has no influence on the further functioning of the application.

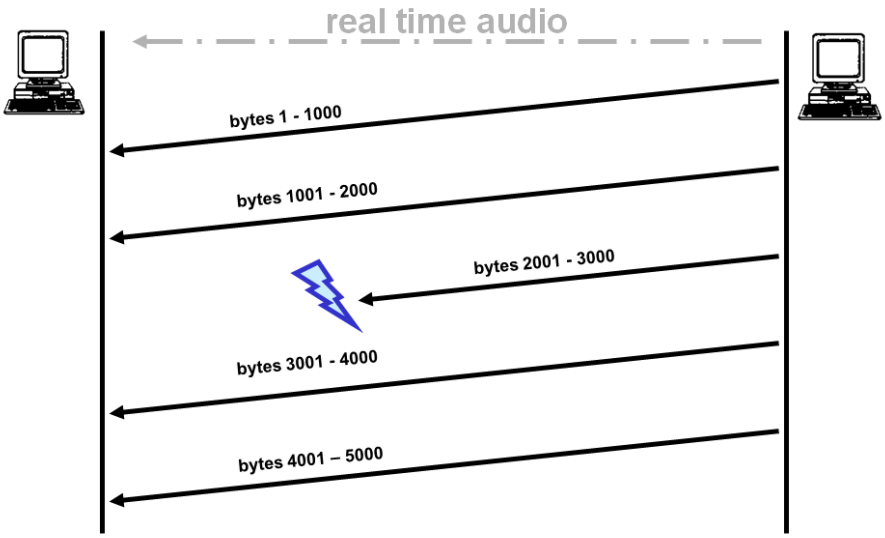


Figure 3.14: Real time audio as application of UDP

3.5 TCP and UDP ports within the automation.

This list gives an overview of some port numbers that are often used within the industrial automation.

Table 3.6: Frequently used TCP and UDP port numbers

Application	Port number / Protocol
FTP-data (File Transfer Protocol)	20 / TCP
FTP-control (File Transfer Protocol)	21 / TCP
SSH (Secure Shell)	22 / TCP,UDP
Telnet protocol	23 / TCP
BootP Server	67 / UDP
DHCP Server	67 / UDP
BootP Client	68 / UDP
DHCP Client	68 / UDP
TFTP (Trivial File Transfer Protocol)	69 / UDP
HTTP (Hypertext Transfer Protocol)	80 / TCP
NTP (Network Time Protocol)	123 / UDP
SNMP (Simple Network Management Protocol)	161 / TCP,UDP
SNMPTRAP (Simple Network Management Protocol Trap)	162 / TCP,UDP
HTTPS (Hypertext Transfer Protocol Secure)	443 / TCP
ISAKMP (Internet Security Association And Key Management Protocol	500 / UDP
MODBUS	502 / TCP; UDP
IPsec NAT traversal	4500 / UDP
EtherNet/IP	2222 / TCP; UDP
PROFINET such as connection establishment	0x8892 (34962) / UDP 0x8893 (34963) / UDP 0x8894 (34964) / UDP
IANA free ports reserved for dynamic and/or private port (profaned service)	0xC000 - 0xFFFF
DDI Device Driver Interface (special protocol used for diagnostic functionality)	1962 / TCP
SOCOMM Interface (Engineering Channel used for control communication)	20547 /TCP

3.6 Communication over TCP(UDP)/IP

3.6.1 Client Server model

An internet (TCP/IP) ensures a general communication infrastructure without specifying what services can be used. TCP/IP provides a basic communication service but this protocol software is not capable of making or accepting contact from a remote participant. Therefore, two application programmes have to be used for each communication. One application starts the communication and the other accepts this.

one significant problem: the protocol software cannot tell at all to an application programme that a request for communication has arrived. The communication between two participants will therefore be based on a model where one application is active (interaction requests) while the other is passive (listening and possibly accepting). Such a model is currently only applied to communication between two different hosts over TCP/IP and is called the *Client Server model*: a Server application waits passively on contact while the Client application starts the communication actively.

Features of Client software:

- Is an application programme that temporarily becomes a Client when remote access to a computer is required but that also carries out local calculations and operations.
- Is directly started by the user and is only carried out for one session.
- Runs locally on the user's PC
- Establishes active contact with a Server
- If necessary, can access several Servers but establishes active contact with one Server at a time
- Does not require special hardware or an advanced control system

Features of Server software:

- Is a specially designed application programme that supplies one specific service but can handle different Clients at the same time.
- Is automatically activated when a system starts and remains active for many sessions
- Waits passively until random, remote Clients look for contact
- Requires some powerful hardware and an advanced control system (depends on the application type)

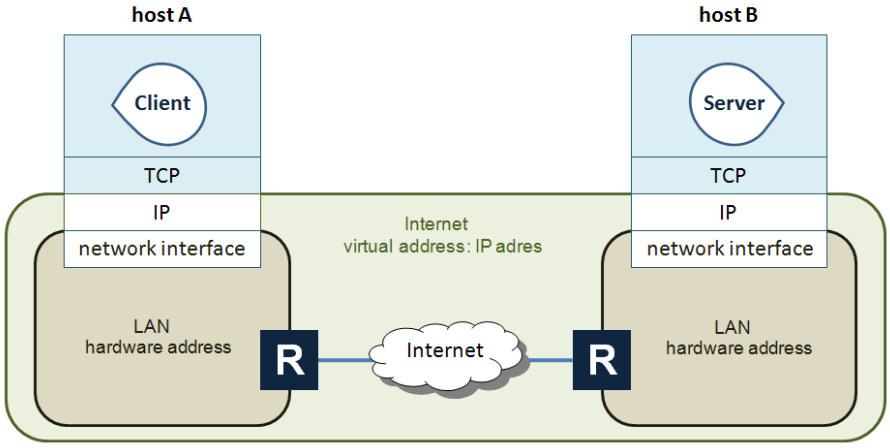


Figure 3.15: Client Server model over TCP/IP

3.6.2 Endpoint and Internetsocket

Figure 3.15 shows a Client Server communication over the TCP/IP stack. Several Clients and Servers can be active at the same time on computer systems. It is important here that every application is identified unambiguously although the computer, on which several applications run, only has one physical connection with the Internet.

For this reason, transport protocols give each communication service a unique name. TCP uses protocol port numbers. A specific protocol port number is assigned to each Server. The Server waits passively for communication via this port number. When sending a request, the Client mentions the port number of the required service. The TCP software on the Server's computer uses the destination port number in an incoming message to determine which Server has to handle the request.

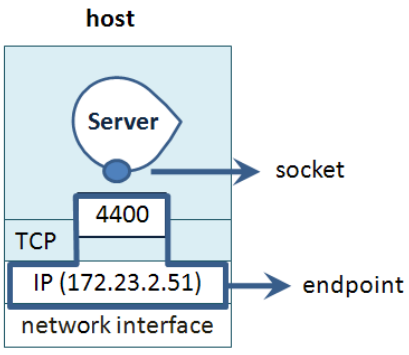


Figure 3.16: Endpoint and socket concepts

endpoint

The endpoint concept and socket concept can sometimes be confusing. The original definition of a socket according to ARPANET is the combination of the IP address and the port

number. This combination is now called an endpoint. An endpoint describes via which logical way an application can be accessed via an internet.

Internetsocket

The term socket is today only a software term. A socket organises the folders and the links of an application on the endpoint. This results in the Internet socket concept - also called network socket. An Internet socket, also called socket in short, is a bi-directional communication endpoint for a process to process connection and is defined by:

- The protocol
  - UDP protocol: datagramsockets of connectionless sockets
  - TCP protocol: streamsockets of connection-oriented sockets
  - raw IP packet (bv ICMP): rawsockets
- Local IP address
- Local protocol port number
- Remote IP address
- Remote protocol port number

3.6.3 Dynamic Servers

A computer system on which several application programmes can run at the same time is said to be a system that supports *concurrency*. A programme that has more than one control thread, or thread in short, process or task, is called a *concurrent* programme.

Concurrency is essential for the Client Server interaction model as a concurrent Server can serve several Clients at the same time, else this Client has to wait for other Clients to finish.

Most of concurrent Servers work dynamically. The server creates a new thread and a new process for every request that is received. In principle, a Server consists of two parts. A first part that accepts the requests and starts a new thread for this request. This first part is called the main thread. The second part consists of the code that can handle every individual request.

When a concurrent Server starts, only the main thread runs. When a request is received, the main thread produces a new thread that handles the request. Meanwhile, the main thread keeps the Server active and waits for a next incoming request.

3.6.4 Unambiguous communication

If several threads of a Server are active, then it is important that the incoming messages of a Client are linked to the correct Server thread.

TCP requires from every Client that it selects a local protocol number that is not allocated yet to a service. Every Client that sends a TCP segment will place this local port number in the

source port field. The protocol port number of the Server is then placed in the *destination port field*.

TCP uses the combination of the protocol port numbers and the IP addresses to identify a certain communication on the Server's computer. This way, messages from several Clients can be received on the same Server without causing any problems. In short, every (Internet) socket has to be unique.

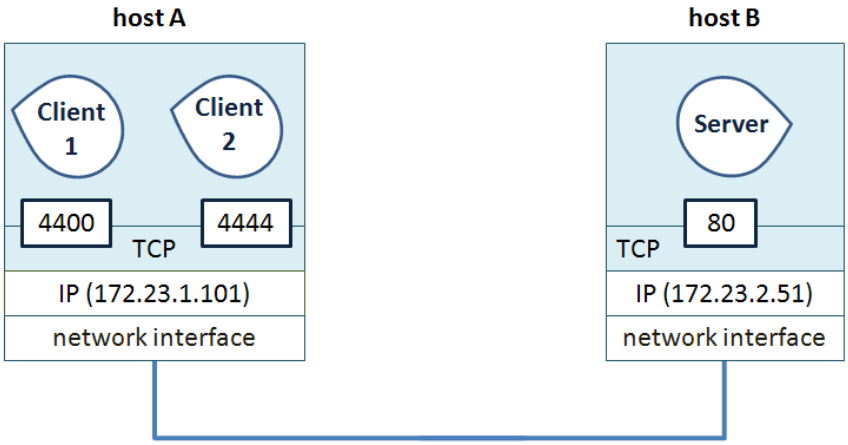


Figure 3.17: Client Server model over TCP/IP

Figure 3.17 shows how unambiguous communication is set up when two Clients set up a remote connection on the same PC with the same Server.

The communication between Client 1 and the Server is characterised by the following socket:

- Protocol: TCP
- SP: 4400
- DP: 80
- SA: 172.23.1.101
- DA: 172.23.2.51

The communication between Client 1 and the Server is characterised by the following socket:

- Protocol: TCP
- SP: 4444
- DP: 80
- SA: 172.23.1.101
- DA: 172.23.2.51

It is enough that one of the parameters is different in order to assign a unique identification to both connections.

3.6.5 Status of a socket

A TCP socket can be in the listening status. When a Server waits for a remote Client to request a communication, then the socket data are intended for the Server.

- Protocol: TCP
- SP: 80
- DP: 0
- SA: 172.23.1.51
- DA: 0.0.0.0

A TCP socket can have the following statuses

- listening
- established
- Syn-sent
- Syn-Recv
- Fin-wait1
- Fin-wait2
- Time-wait
- Close-wait
- Closed

A UDP socket cannot be in established status. A UDP Server cannot create new threads for every other Client. The main process processes incoming data packets in sequence, via the same local UDP socket.

3.6.6 Connection-oriented communication and connectionless communication

Transport protocols support two basic communication forms: Connection-oriented (TCP) and connectionless (UDP). Clients and Servers can use both basic forms for their communication.

If a Client uses TCP for a connection-oriented communication, then this Client has to ask TCP first to start a connection with another application. After the connection is established, the two applications can exchange data. TCP closes the connection after the two applications end the communication.

The alternative is a connectionless communication whereby an application can send a message at any moment to any destination. An application that uses UDP can send a series of messages whereby each message is sent to another destination.

Chapter 4

Extension protocols and network applications

4.1 ARP

4.1.1 Introduction

The IP address is a virtual address which is processed by software. No LAN hardware or WAN hardware can set a relationship between the NetID of an IP address and a network or between the HostID and an IP address and a host. In order to transport an IP packet, these data have to be wrapped in a frame that can be delivered by the local hardware to the right participant. This frame also has to contain the hardware addresses of the sender and the receiver.

4.1.2 Address Resolution Protocol (ARP)

If the IP protocol wants to send a message over Ethernet, then the MAC address of the destination also has to be known, besides the IP address of the destination. For this reason, the TCP/IP protocol suite contains an Address Resolution Protocol (ARP). The ARP defines two basic message types: a request and a reply. A request message contains an IP address and asks for the corresponding hardware address and the MAC address. The reply contains an IP address that was included in the request and the hardware address.

It is obviously hopelessly inefficient to first send an ARP request for every IP packet to be sent. For this purpose, the ARP protocol will temporarily store all information which is received in a table.

```
H:\PIH\personeel\henk.capoen>arp -a

Interface: 192.168.1.2 --- 0x3
  Internet-adres        Fysiek adres          Type
  192.168.1.1           00-12-bf-fa-0b-4e     dynamisch
  192.168.1.4           00-0c-41-62-8b-14     dynamisch

H:\PIH\personeel\henk.capoen>
```

Figure 4.1: The ARP cache

ARP manages this table like a cache: a small table with a limited number of bonds that are all the time overwritten or are deleted again after a period (a few minutes). Figure 4.1 shows how a current overview of the ARP cache is obtained with the DOS command arp -a.

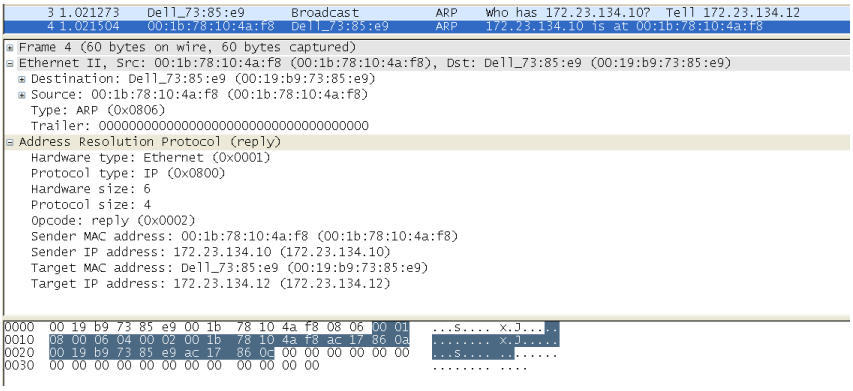


Figure 4.2: ARP reply in Wireshark

Figure4.2 shows the use of ARP in Wireshark (Wireshark is a packet sniffer and protocol analyser - a programme that is used to retrieve and analyse data on a computer network).
The RARP protocol does the reverse. It sends a request - a request with a hardware address. This results in a reply with the looked up IP address.

4.2 BootP and DHCP

4.2.1 Introduction

During the start of a host, a number of issues have to be configured before this host can actively participate in the network traffic. Every host has to obtain an IP address, the applied subnet mask, the IP address of the default gateway (this is the router that links the local network to other networks and to the Internet,...) and any data with regard to the DNS server (see also in this chapter). This data can be recorded statically in a host or can be assigned dynamically to a host. In this part is discussed how certain data can be configured automatically during the start up. The start-up process is also known by the name bootstrapping.

4.2.2 BootP

The Bootstap protocol is added to the TCP/IP suite to combine a number of dynamic configuration steps into one step. In order to obtain configuration information, the BootP protocol sends a request broadcast message. A BootP server recognises this message and returns a BootP reply message with all required information to the requesting participant. BootP uses an IP packet although the participant does not contain an IP address yet. A broadcast address is sent as target address only consisting of ’1’s and a source address only consisting of ’0’s. The BootP server can use the hardware address to send a reply.

BootP simplifies the configuration but the problem remains that a BootP server retrieves its information from a database that the administrator has to update manually.

4.2.3 DHCP

The IEFT has developed the Dynamic Host Configuration Protocol (DHCP) to further automate the configuration. DHCP is a protocol that can assist a host on a new network without manual intervention of an administrator. Figure 4.3 shows the different steps during the automated configuration of a host. It is a Client server protocol. The client is a new host that requests for the IP data. One or more DHCP servers will be present per network and these can assign these data.

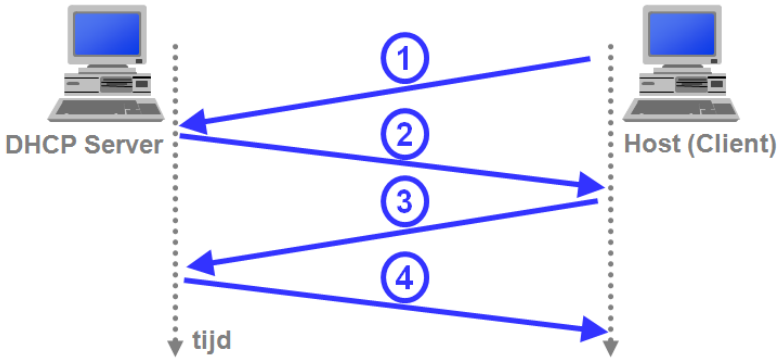


Figure 4.3: DHCP protocol

For a new host, the DHCP protocol consists of four steps:

- DHCP discover message: a Client sends an UDP message via port 67, packed in an IP packet in order to trace a DHCP server. A broadcast destination address (255.255.255.255) is used and 0.0.0.0 is assigned as source address.
- DHCP offer message: a DHCP server replies to the Client. This answer contains an IP address, a subnet mask and lease time for the IP address.
- DHCP request message: the host chooses from the various offers and replies to the chosen server with a request message that contains the configuration parameters.
- DHCP ACK message: the server replies with a confirmation.

4.2.4 DHCP Relay Agent - DHCP option 82

DHCP Relay Agent is a bootstrap protocol that can send DHCP messages between clients and servers for DHCP over various IP networks. In other words, a DHCP server can also serve a network to which it is not directly connected via a DHCP Relay Agent.

A DHCP Relay Agent looks out for broadcast messages from DHCP clients on the network via the familiar bootpc (67) client port. These messages are converted into unicast messages and they are then forwarded to the configured DHCP server. To do this, the DHCP Relay Agent fills in the "giaddr" field in these messages with its own IP address. The DHCP server can then send the reply as a unicast message to the Relay Agent. The Relay Agent then sends the reply either via a broadcast message or via a unicast message on the client’s network.

DHCP option 82 is a DHCP Relay Agent Information Option. This option has been developed so that a DHCP Relay Agent network can add specific information to a message that it forwards to a DHCP server. Here the option uses two associated pieces of data: Circuit ID and Remote ID. The DHSP server must obtain this information about the location of the host that is sending the request. This information is heavily dependent on the DHSP Relay Agent and for Ethernet-based networks this consists of the MAC addresses of the ports on the Relay Agent that form the link with the end host. This information may be used to indicate where an assigned IP address is physically located in the network. This information can also be used by the DHSP server when deciding on the assignment of a specific IP address.

4.3 ICMP

4.3.1 Introduction

Data packets can get lost in the IP communication service, the delivery of these can be greatly delayed or delivered in the wrong order. IP is not a reliable communication service but tries to avoid errors and reports any problems when they occur. A primary example of error detection is the header checksum. Every time a data packet is received, the checksum is checked to make sure that the header is not damaged. If a checksum error is observed, then this message is straight away discarded. No message is produced in this case as the source address is deleted together with the message. Other less serious problems may be reported though.

4.3.2 Internet Control Message Protocol

The TCP/IP protocol suite contains a protocol (Internet Control Message Protocol (ICMP)) to send error messages. This way it can be reported if a certain network feature is not available or if a certain host or router is not available. Sometimes, a computer user gets in touch directly with the ICMP protocol, especially when using the network diagnosis commands - ping and traceroute.

ICMP has five error messages and four informative messages. The five ICMP error messages are:

- Source quench: this is sent by a router if it is temporarily short of buffer space and therefore has to discard incoming IP packets. This message is sent to the host that has created the IP packet. The sending host will have to adapt the transmission speed.
- Time exceeded, this is sent by a router after the *Time to live* field has been reduced to zero.
- Destination unreachable: this is sent by a router if it notices that an IP packet cannot reach its destination. The error message distinguishes between a situation, whereby a complete network is temporarily delinked from the Internet (if a certain router is not functioning correctly) and the case where a certain host is temporarily offline.
- Redirect: this is sent by a router if it notices that the IP packet actually has to be sent to another router to reach the final destination.

- Fragmentation required: this is sent by a router if it notices that an IP packet is greater than the MTU (Maximum Transmission Unit) of the network over which the IP packet has to be sent.

ICMP also defines four informative messages:

- Echo request/reply. An echo request can be sent to any host. In return, an echo reply is sent. The answer contains the same data as the request.
- Address mask request/reply, a host sends an address mask request when it is started. A router replies with a message that contains the correct subnet mask that is used on the network.

4.3.3 ICMP message

The ICMP protocol is a supporting protocol for the IP protocol. It uses IP packets to send messages. Figure 4.4 Shows how an ICMP message is encapsulated in a data frame.
An ICMP error message is always produced as a reply to a certain IP packet. This is always returned to the source of the IP packet.

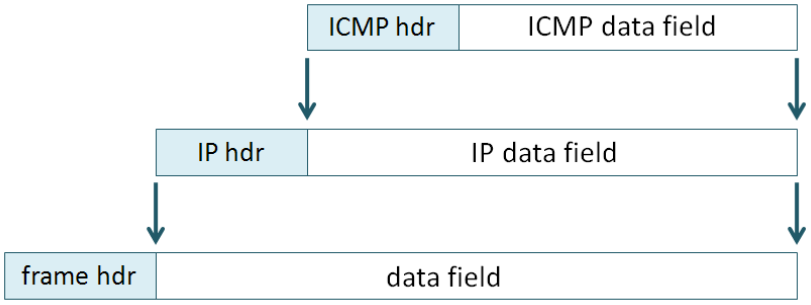


Figure 4.4: Encapsulation of an ICMP message

The different fields in the ICMP header are:

- TYPE:
- Code:
- Checksum:
- Identifier:
- Sequence number:

4.3.4 Check accessibility of a host

Many tools retrieve information over a network by sending test messages and then wait for the ICMP replies. One of the most important diagnosis tools is the ping programme. The ping programme (to be entered as DOS command) sends IP packets via ICMP to another participant to check if this host can be reached via the network. The target host has to send these small packets back immediately (as an echo). Moreover, a static summary of

the percentage of small packets for which no reply is given and the response time are also displayed. The IP address or the host name can be used.

ping www.google.be
ping 134.16.85.9

Quite a few suitable options are possible. An overview of all options is obtained if the ping command is entered as such.

```
H:\PIH\personeel\henk.capoen>ping 192.168.1.1

Pingen naar 192.168.1.1 met 32 byte gegevens:

Antwoord van 192.168.1.1: bytes=32 tijd=3 ms TTL=64
Antwoord van 192.168.1.1: bytes=32 tijd=4 ms TTL=64
Antwoord van 192.168.1.1: bytes=32 tijd=3 ms TTL=64
Antwoord van 192.168.1.1: bytes=32 tijd=4 ms TTL=64

Ping-statistieken voor 192.168.1.1:
    Pakketten: verzonden = 4, ontvangen = 4, verloren = 0
    (0% verlies).De gemiddelde tijd voor het uitvoeren van één bewerking in mil
iseconden:
    Minimum = 3ms, Maximum = 4ms, Gemiddelde = 3ms
```

Figure 4.5: Ping command

4.3.5 Trace a route

If the ping programme is only used to check if a certain host is accessible, then the tracert command will show the route to a certain host. Figure 4.6 shows in which way a tracert command displays all IP addresses of the routers that receive the test packet and returns them.

```
H:\PIH\personeel\henk.capoen>tracert www.google.be

Bezig met het traceren van de route naar www.l.google.com [74.125.79.104]
via maximaal 30 hops:

  1     1 ms     1 ms     1 ms  . [192.168.1.1]
  2     9 ms     7 ms     8 ms  86-39-8-1.customer.fulladsl.be [86.39.8.1]
  3     9 ms     9 ms     9 ms  83.217.77.30
  4     9 ms     9 ms     9 ms  ge-5-2-11.bb1.bru1.be.gbxs.net [193.27.64.161]
  5    17 ms    14 ms    14 ms  pos-6-0.bb1.bru2.be.gbxs.net [193.27.64.34]
  6    16 ms    13 ms    15 ms  so-7-0-0-0.bb1.ams3.nl.gbxs.net [83.143.243.18]

  7    16 ms    15 ms    14 ms  xe-3-1-0-26.bb1.ams1.nl.gbxs.net [193.27.64.178]

  8    17 ms    15 ms    15 ms  core2.ams.net.google.com [195.69.145.100]
  9    16 ms    15 ms    40 ms  209.85.254.90
 10    20 ms    18 ms    18 ms  209.85.248.79
 11    23 ms    19 ms    19 ms  209.85.255.20
 12    21 ms    23 ms    26 ms  209.85.255.122
 13    21 ms    19 ms    20 ms  ey-in-f104.google.com [74.125.79.104]

De trace is voltooid.
```

Figure 4.6: Tracert command

Tracert first sends a test packet with a time to live value of 1. The first router reduces this value to 0, discards the message and sends back an ICMP Time Exceeded message. This way, the IP address of the first router can be displayed. Next a test message is sent with a time to live value of 2. The first router reduces this value and forwards the message. The second router will set the time to live value to 0, discards the message and returns an ICMP

error message. This way, the IP address of the second router is known. This procedure is continued until the final host is reached.

4.4 IGMP

4.4.1 Introduction

IGMP (Internet Group Management Protocol) is the protocol for IP multicast applications on TCP/IP networks. This standard is defined in RFC 1112. Besides the definition of address- and host extensions for the support of multicasting by IP hosts, this RFC also contains a definition of version 1 of IGMP. IGMP version 2 is defined in RFC 2236. Both versions of IGMP offer a protocol with which information about the membership of a host of specific multicast groups can be exchanged and updated.

Multicast messages are sent to one address (multicast IP address) but are processed by several hosts. The collection of participants that listen to one specific multicast IP address is called a multicast group. Some important aspects of multicasting:

- The membership of a group is dynamic: hosts can enter or leave a group at any moment.
- Hosts can join multicast groups by sending IGMP messages
- There is no restriction with regard to the size of groups. The different participants can be spread out over several networks on condition that the intervening routers support IGMP.
- Hosts that do not belong to a certain group can send IP messages to that group.

4.4.2 IGMP messages

IGMP describes how the information about the membership status is exchanged between routers and the different participants of multicast groups. Some IGMP messages:

- Host membership report: is sent if a host becomes member of a multicast group and informs by way of this message all other members of the group. A router saves these reports and guarantees the maintenance of the multicast group in this way.
- Host membership query: is sent by routers to periodically inform the group members in a network. All members of a group reply once again with a membership report. Routers retain all information and ensure that no multicast messages are sent on networks where no members of the group are present.
- Leave group: this message is sent by a host that is the last one of a group within a certain network segment to leave a group.

The IGMP protocol is a supporting protocol for the IP protocol. It uses IP packets to send messages. Figure 4.7 Shows how an IGMP message is encapsulated in a data frame.

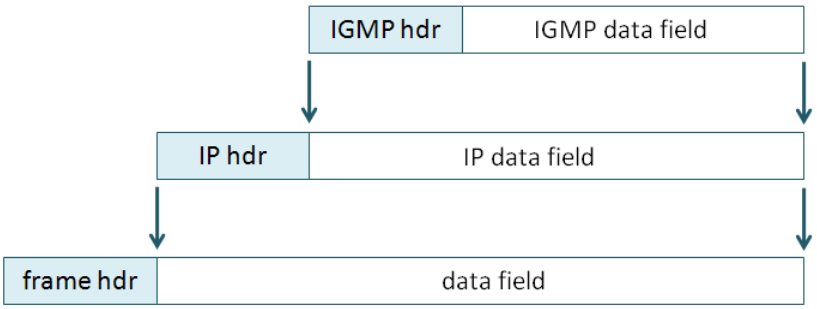


Figure 4.7: Encapsulation of an IGMP message

4.4.3 IGMP snooping

A switch that connects a member from a multicast group with a router can read IGMP messages and evaluate by means of IGMP snooping. IGMP snooping translates multicast IP addresses in multicast MAC addresses. A switch can store multicast MAC addresses this way in its multicast filter table. A switch can send multicast messages immediately to the correct ports in this way and thus ensures that multicast messages do not put a needless strain on a network. For switches, this is also called the dynamic use of multicasting. This is opposed to the static use of multicasting whereby the groups in all switches and for all ports have to be configured manually.

4.4.4 Multicast addresses

Multicast IP addresses are class D addresses that lie in the range 224.0.0.0 up to and including 239.255.255.255. For private networks it is recommended to use multicast IP addresses in the range 239.x.x.x.

The addresses within the range 224.0.0.1 up to and including 224.0.0.255 are reserved for multicast applications within one network. The time to live value of such IP packets is set to 1 so that such messages continue to be available on the network.

Multicast MAC addresses are also reserved. All addresses for which the first byte is equal to 01h are available for multicasting. Addresses that start with the value 01:00:5E:0 are multicast MAC addresses that are used for IP multicasting.

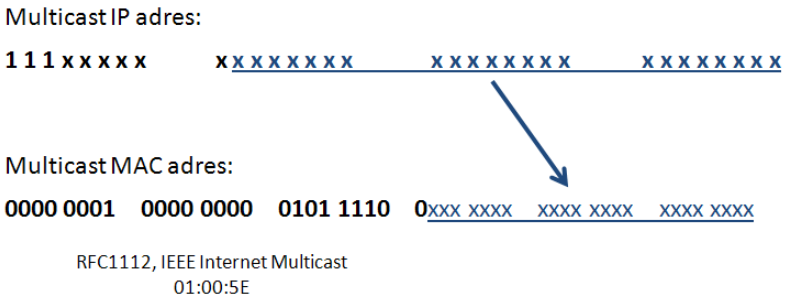


Figure 4.8: Conversion from multicast IP address to multicast MAC address

This conversion requires some attention. The most important bit of the second byte belongs to the recognition code of a multicast address and is therefore not mapped. Thus the multicast IP address 228.30.117.216 is converted to multicast MAC address 01:00:5E:1E:75:D8. However, the multicast IP address 228.158.117.216 is also converted to multicast MAC address 01:00:5E:1E:75:D8.

4.5 GMRP

4.5.1 IEEE 802.1p

Company networks are becoming ever larger and more complex. It is essential for growing network traffic to be managed efficiently. Here Quality of Service is an important tool for ensuring that the most critical data are delivered in the most predictable way. Thanks to the IEEE 802.1p protocol, switches can process network traffic in order of priority. Predictability and reliability of network traffic are thereby improved.

IEEE 802.1p defines a 3-bit field within a tagged Ethernet frame. This defines a priority level between 0 and 7 so as to be able to differentiate the network traffic.

The IEEE 802.1P standard also provides measures for filtering multicast messages so that they are not needlessly broadcast via Layer 2-based networks. These include GMRP (GARP Multicast Registration Protocol). GMRP and GARP are commercial protocols defined by IEEE 802.1P.

4.5.2 GMRP processing

GMRP processes multicast group addresses on Layer 2 (MAC layer). GMRP runs both on the switch and on the hosts. On the host, GMRP is used together with IGMP and Layer 2 data frames are created from the Layer 3 IGMP messages.

A switch receives both the Layer 2 GMRP messages and the Layer 3 IGMP ones. With the GMRP messages, the switch restricts the data traffic in the VLAN group in which the sending host is located. When the switch receives the GMRP join message, it will add the port on which it received this message to the relevant multicast group. The switch forwards the membership request message to all the other participants in the VLAN, which of course also

includes the multicast source. When the source sends a multicast message to the group, the switch forwards this message only to the members of this group.

The GMRP periodically sends queries. If a participant wishes to remain a member of a group, he sends a reply to the switch. A participant who no longer wishes to remain a member of the group can send a leave message or not reply at all. If the switch does not receive a reply or it receives a leave message from a specific host, it will delete this participant from the list.

4.6 DNS

4.6.1 Introduction

There are two important ways to indicate a host on the Internet. Besides the already mentioned IP address, it is also possible to give a participant a host name (a symbolic name) which is generally easier to use.
Host names, such as www.google.be (search engine), gaia.cs.umass.edu (computer networks research Group of the university of Massachusetts, Amherst) are easier to remember and therefore more user-friendly. A host name however, does not provide enough information about the location of that host within the Internet. A link has to be established between host names and IP addresses as users tend to use the host name whereas TCP/IP protocols are based on IP addressing. Domain Name System (DNS) provides a solution to this problem. Dr. Paul V. Mockapetris and Jon Postel are the inventors of the Domain Name System. In 1983, they set up a DNS architecture in RFC882 and RFC883.

In summary, DNS stands for:

- A distributed database which is implemented in a hierarchy of DNS servers
- An application layer protocol with which hosts and DNS servers can communicate during the translation session (conversion from an IP address to a host name and vice versa).

The DNS servers are often Unix machines on which the Berkeley Internet Name Domain (BIND) software or Microsoft DNS software is run. The DNS protocol works with UDP and uses port 53.

4.6.2 The structure of a host name

With regard to syntax, every host name consists of a series of alphanumerical segments that are separated by dots. Domain names have a hierarchical structure whereby the most significant part of the name is located to the right. The leftmost segment is the name of the individual host. Other segments in a domain name identify the group that is the owner of the name. DNS does not record how many segments a domain name has to contain. The DNS issues values for the most significant segment. Table 4.1 gives an overview of all names

Table 4.1: Names for the most significant part of a domain name

Domain name	Allocated to
com	commercial organisation
edu	educational institute
gov	government body
mil	military group
net	network supporting centre
org	other organisations
int	international organisation
country code	a country, be for Belgium

4.6.3 Functioning of the DNS protocol

Introduction

When an application (e.g. a web browser) on the host of a user has to convert a host name into an IP address, then this application will call the client component of the DNS with the host name that has to be translated. The DNS client component on the user host then takes over and sends a request message to the network. All DNS request- and reply messages are sent in UDP segments to port 53.
After a delay (which can vary from a few milliseconds to a few seconds), the DNS client component receives a DNS reply message with the requested reference on the user host. This reference is then transmitted to the application. This way, DNS is (from the perspective of the application on the user host) a blackbox that provides a simple, uncomplicated translation service.
In fact, the service that supplies the blackbox is very complex and consists of a great number of DNS servers, located all over the world, and an application layer protocol that determines how the DNS servers and the requesting hosts communicate with each other.
A simple design for DNS would consist of one DNS server that contains all references. In this centralised design, all clients would only have to send their requests to that one DNS server and then this server would process all requests. Although the simplicity of this design is attractive, this solution is unsuitable for the current Internet with the enormous (and fast increasing) number of hosts.

DNS uses a great number of DNS servers (located all over the world) in a hierarchical structure. There is no DNS server that contains all references for all hosts on the Internet. The references are divided over the DNS servers.
At first, there were three classes of DNS servers:

- Root DNS servers;
- Top level domain (TLD) DNS servers;
- Verifying DNS servers.

A distributed, hierarchical database

A DNS client first contacts one of the root servers that returns IP addresses for TLD servers for the Top level domain com. The client then contacts one of these TLD servers that returns

the IP address of a verifying server. Finally, the client contacts one of the verifying servers that will return the IP address of the host name.

Root DNS server
On the Internet, there are only 13 root DNS servers (indicated with the letters A-M) of which most are set up in North America. Although the 13 root DNS servers are indicated as one server, each *server* consists in fact of a cluster of replicated servers (for reasons of security and reliability).

Top level domain (TLD) server
These servers are responsible for top level domains such as com, org, net, edu and all countries top level domains such as be, nl, fr, jp,

Verifying DNS server
Every organisation which has hosts (on the Internet) that the public can access (such as webserver and Mailserver) is obliged to supply DNS data in which the names of these hosts are linked to IP addresses. These DNS data are stored on the verifying DNS server. An organisation can choose to implement a verifying DNS server with these data by itself but it can also ask a service provider (Telenet for example) to store these data, for a fee, on their verifying DNS server. Most universities and large companies implement and manage their own primary and secondary (as back-up) verifying DNS server.

The root, TLD- and verifying DNS servers all belong to the hierarchy of DNS servers. There is still another important type of DNS servers that are called local DNS servers. Strictly speaking, a local DNS server has no place in the hierarchy but is of vital importance to the DNS structure. Every ISP (Internet Server Provider) has a local DNS server (also called standard DNS server (default name server)). When a host wants to make a connection with an ISP, the ISP gives the host the IP addresses of one or more of its local DNS servers (mostly by means of DHCP).

4.7 SNMP

4.7.1 Introduction

SNMPv1: In 1990, the SNMP protocol was defined in RFC 1157. SNMP stands for Simple Network Management Protocol. This protocol describes a structure way of safeguarding and managing a certain network infrastructure. Rather quickly, this protocol was implemented on a large scale in commercial products and this protocol became the de facto standard for network management. SNMP is designed to be a simple protocol.

SNMPv2: With the experience gained, an improved version of SNMP was written in 1993 in RFC 1441 and RFC 1452 (co-existence between v1 and v2) to become the Internet standard in the end.

SNMPv3: The third version of the Internet Standard Management Framework (SNMPv3) is derived from and based on the previous versions, SNMPv1 and SNMPv2. SNMPv3 is therefore SNMPv2 supplemented by security and administration. The main features of SNMPv3 consist of:

- security

  - authentication and privacy
  - access control
- administration
  - user names and key management
  - designation of participants
  - policies

On a network, there are many interesting active participants that contain important status information for network management. Such participants can be hubs, switches, routers, printers or PCs. In order to be managed directly by SNMP, a node should be able to run an SNMP management process- an SNMP agent. All computers comply with this requirement, as well as many hubs, switches, routers and peripherals that have been designed for network use. Every agent keeps a local database with variables that represent its situation and history and influence its functioning.

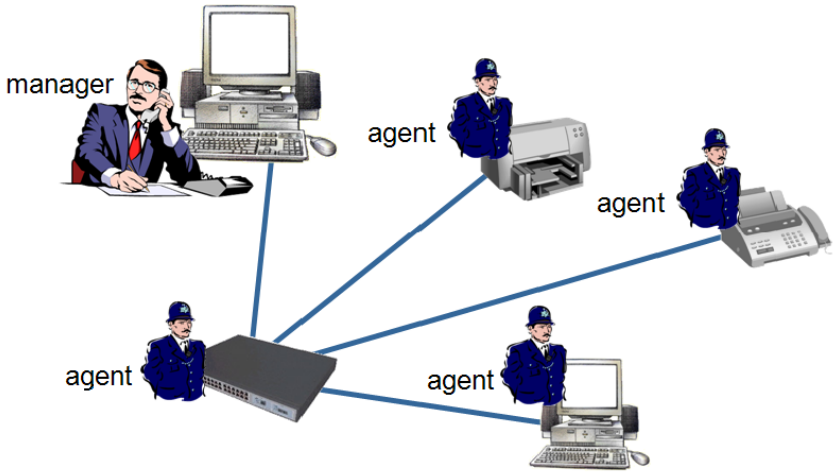


Figure 4.9: Managers and agents on a network

The network management is carried out from management stations: in fact with normal computers in which special management software is run. These stations contain one or more processes that communicate with agents via the network whereby they give orders and receive replies. In this set up, all intelligence lies in the management stations so that the agents are kept as simple as possible and to keep to a minimum the devices on which they run. Many management stations have a graphic user interface so that the network administrator can check the network status and take action accordingly.

4.7.2 SNMP structure

The SNMP framework consists of three essential components:

- the MIB (Management Information Base (RFC1213)) is the description of all variables that a certain network element contains.

- SMI (Structure of Management Information (RFC 1155)) is the structure for the storage of the network information.
- SNMP (RFC1157) is the communication protocol between the manager and a network participant.

Most existing networks are mixed products - with hosts of one or more manufacturers, hubs, switches and routers from other companies and printers from a different company. In order to make it possible that a management station (possibly from another supplier again) can talk with all these various components, the type of information, that is maintained by all these devices, has to be strictly specified. It is useless if the management station asks a router what its frequency of lost packets is when the router does not update this frequency. Therefore, SNMP describes the exact information that every type of agent has to maintain and the format in which the agent has to supply this information. The main part of the SNMP model is the definition of who has to maintain what and how this information has to be transmitted. In short, it comes down to that every device maintains one or more variables (objects) that describe the device status. The collection of all possible objects in a network can be found in the data structure called MIB (Management Information Base).
The SNMP protocol itself describes how the interaction between the management station and the agents is set up. Five different message types are defined here.

4.7.3 The MIB and SMI

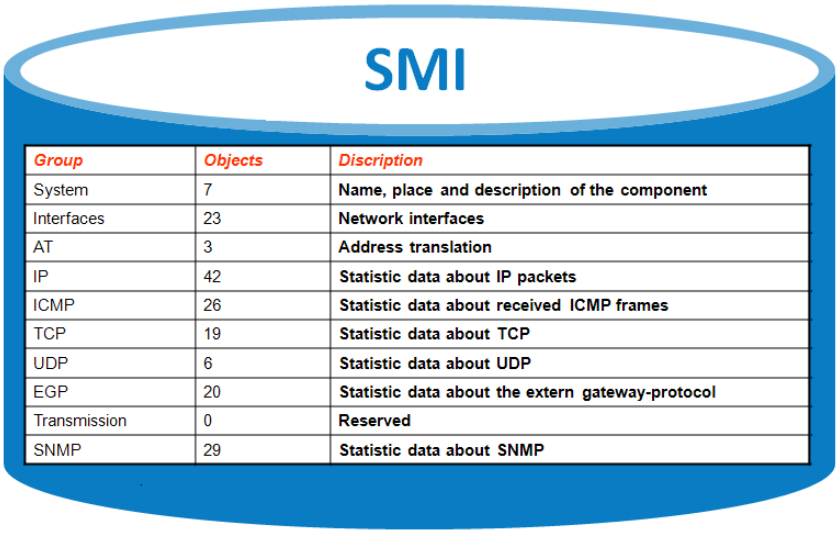


Figure 4.10: MIB is a database which contains all variables for network management

The collection of objects that is managed by SNMP is defined in the MIB and represented in figure 4.10. For sake of convenience, these objects are divided into different groups. These categories are meant to provide a basis for what a management station has to cover.

- The group System gives the manager the opportunity to figure out how the device is called, who has made it, which hardware and software it contains, where it is located and what it has to do. The timestamp of the last boot is also specified.

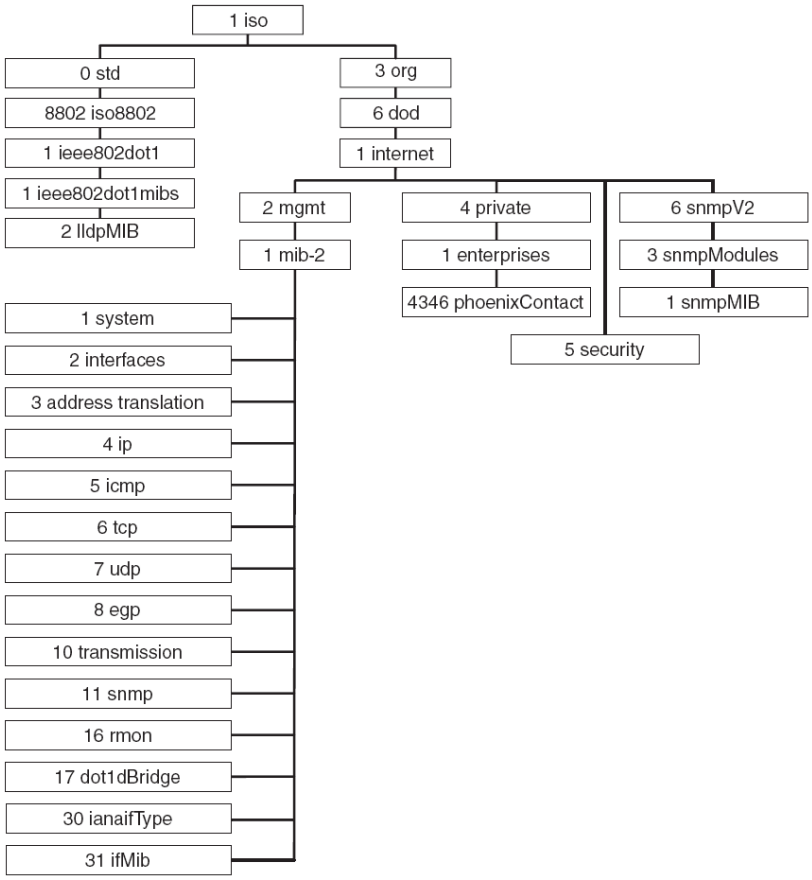


Figure 4.11: MIB tree structure

- The group Interfaces is responsible for network adapters. This updates how many packets and bytes are sent and received through the network, how many are discarded, how many broadcasts there are and how long is the current execution-queue.
- The group IP is responsible for IP traffic to and from the node. This has several counters on which are updated the number of packets which are discarded for various reasons. Statistics are also available about the fragmentation and setting the datagrams back again. All these items are very important for the management of the routers.
- The group ICMP relates to IP error messages. There is a counter for every ICMP message in which the number of each type observed are updated.
- The TCP group registers the actual number of opened connections, sent and received segments and various statistics about errors.
- The UDP group counts the number of sent and received UDP datagrams and also updates the number of received datagrams which were returned to sender due to an unknown port or for another reason.
- The last group is intended for the collection of statistics about the functioning of SNMP itself: the number of messages which were sent, the type of messages, etc.

Every variable, every object from the MIB is characterised by an object identifier (OID) and its type:

- The OID describes a path in the MIB tree structure. Figure 4.11 shows the structure of the MIB as used for SNMP. The object *sysObjectID*, which is part of the *system* group, can be accessed via OID 1.3.6.1.2.1.1.2.0
- Object types are built by means of fundamental types that are identified in the SMI.

Different MIBs are available. First of all, the global MIBs are described in RFCs. MIB2 is described in RFC1213, for example. Such MIBs have to be supported by all SNMP compatible devices. On the other hand, there are also manufacturer-specific MIB objects.

4.7.4 SNMP protocol

The normal use of SNMP is that the management station sends a request to an agent with a request for information or with the order to update the status in a certain way. In the ideal case, the agent only replies with the requested information or confirms that the status has been updated as requested. SNMP defines the different messages that can be sent.

Table 4.2: SNMP messages from manager to agent

Message	Description
Get request	Request the value of one or more variables
Get next request	Requests the variable after the current one
Get bulk request	Retrieves a large amount of information
Set request	Updates one or more variables
Inform request	Message between managers that describes the local MIB

In one specific case, the agent itself can take the initiative to send a message and this is done the moment an agent observes a certain critical event. Managed nodes can drop out and reboot, network segments can drop out and start up again, etc. Every relevant event is defined in an MIB module. When an agent observes that there has been a relevant event, then he immediately reports this event to all management stations in his configuration list. This message is called an SNMP trap. The message mostly only states that some sort of event has occurred. It is then the task of the management station to carry out requests to find out the details.

Table 4.3: SNMP messages from agent to manager

Message	Description
SNMP trap	Message about event from agent to manager

Figure 4.12 shows that SNMP messages use UDP protocol and the ports which are used for this purpose;

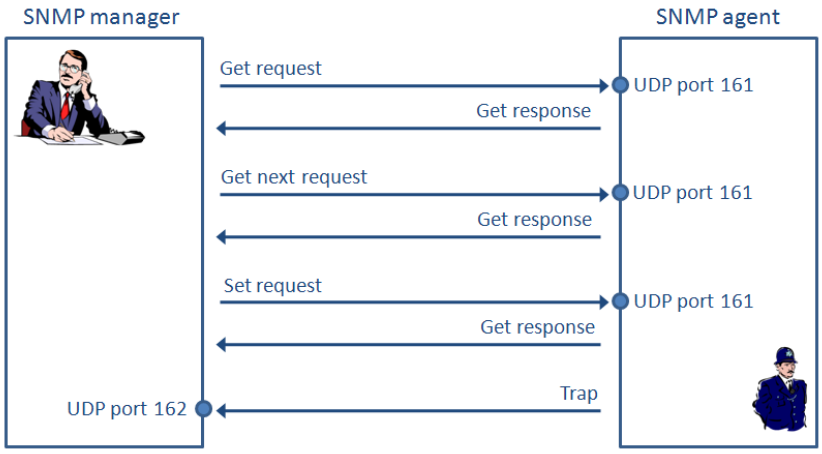


Figure 4.12: SNMP communication

4.8 HTTP and HTTPS

4.8.1 TLS/SSL

Transport Layer Security (TLS) , the successor to Secure Sockets Layer (SSL), is an encryption protocol for creating a secure data channel on an insecure network such as the internet. Both protocols run one layer below the application protocols such as http, SMYP, FTP, and above the TCP transport protocol. They form part of the TCP/IP protocol suite. One of its major goals is to make client/server applications secure. On the sending side the TLS Layer encrypts data from the application and forwards it to the correct TCP port. On the receiving side, TLS reads the data from the correct TCP port, decodes the data and sends it to the application. It is the task of the record layer to transfer the data.

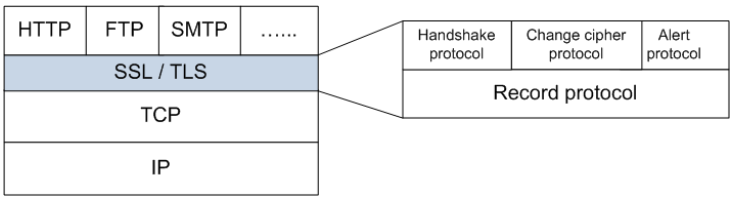


Figure 4.13: SSL TLS layer

TLS provides the following safeguards for client/server applications via TCP/IP:

- Authentication: an application permitted to verify the identity of another application with which it is communicating.
- Privacy: data that are transferred between applications cannot be misused or read.
- Integrity: applications detect whether data have been amended during transmission.

The techniques used are based on concepts such as public key and certificates (see Security section). If an application uses SSL/TLS, a handshake procedure is first initiated in which the encryption algorithm and the codes to be used are agreed and the client verifies the

server. Once this procedure has been completed, all the application data are encrypted. The execution and monitoring of this handshake procedure are carried out by the highest part layers of the protocol. See Diagram 4.13.

4.8.2 HTTP

HTTP (HyperText Transfer Protocol) is the protocol for the communication between a web client (a web browser) and a web server. This protocol is not only used a lot on the World Wide Web but also on local networks. The protocol defines the exact format of the requests, the requests of a web browser to the server and the format of the answers, and the responses returned by the web server. Each question contains a URL that refers to a web component or a static object such as a web page.

The http protocol uses port 80.

Every http URL begins with ’http://’.

HTTP is not secure and is vulnerable to man-in-the-middle attacks and eavesdropping.

4.8.3 HTTPS

HTTPS (Hypertext Transfer Protocol Secure) is an extension on the HTTP protocol with the aim to exchange data safely. When using HTTPS, the data are encrypted. This makes it difficult for an outsider to intercept the data. HTTPS is in principle HTTP with SSL/TLS being used to encrypt the data and to verify the server.

Every URL begins with ’https://’.

The protocol uses TCP port 443.

4.9 Overview of some other important applications

4.9.1 FTP

FTP (File Transfer Protocol) is a protocol that facilitates the exchange of files between different hosts. It allows the transmission of a random file and contains a mechanism with which restrictions can be put on files with regard to ownership and access rights. The protocol hides the details of an individual computer system for the user and is therefore suitable for heterogeneous situations. The protocol can transport a file between two totally random systems.

4.9.2 TFTP

TFTP (Trivial File Transfer Protocol) is a simplified version of FTP that is often used to provide firmware and configurations to devices like routers, switches, ...

4.9.3 NTP

NTP (Network Time Protocol) is a protocol with which computers can set their internal clock to the same time of another computer. NTP is based on the predictability of the network latency. The computer network is hierarchically organised whereby the computer with the most accurate time source is indicated as 'stratum 0'. The computer systems that derive their time from this computer, via NTP, are by definition 'stratum 1'. The protocol has some intelligent functions. An NTP Client can use several NTP servers whereby the NTP Client sorts out by itself server works the best. Based on a number of decision criteria, an NTP Client selects a server and synchronises with this server. Minor differences in time between server and client are updated by the client by having the time processing on the client's computer to function a bit faster or slower. This means that the difference can be eliminated without jumps in time.

4.9.4 SSH

Secure Shell is located in the application layer of the TCP/IP protocol. SSH replaces older protocols such as telnet and rlogin with a secure version of these. The protocol uses TCP port 22.

SSH permits secure login to another computer and the execution of remote commands on the other computer via a Shell. The encryption used makes it difficult for third parties to detect the original commands.

One major advantage of SSH is that authentication is also possible with the aid of a private/public key. As a result, SSH applications may be used automatically without the need for a password in the code. With the use of the private key, it is also possible to log into any system that recognises the public key.

4.9.5 CLI (Command Line Interface)

If an operating system provides a Command Line Interface, the user can make the system execute one or more tasks by means of a command line. Once the task has been executed, the user once more has an opportunity to input a subsequent command line. Commands are usually finished with the <enter> key. Well-known CLIs include command.com (DOS) and bash (UNIX).

In addition to operating systems, there are also other programs that can run with a CLI. FTP client and Microsoft's Telnet Client, for example, use a command line.

Most commercial switches are also configurable via a CLI.

Chapter 5

The switch

5.1 General

The switch is the basic component for the building of an Ethernet-based local network. The various participants of a LAN are linked to each other in an intelligent way by means of switches. A switch has several ports. A host (network participant) can be connected to each port of a switch or the connection can be made with another switch. This way, a network can be extended in a star-shaped manner. A network segment, a collision domain starts from every port.

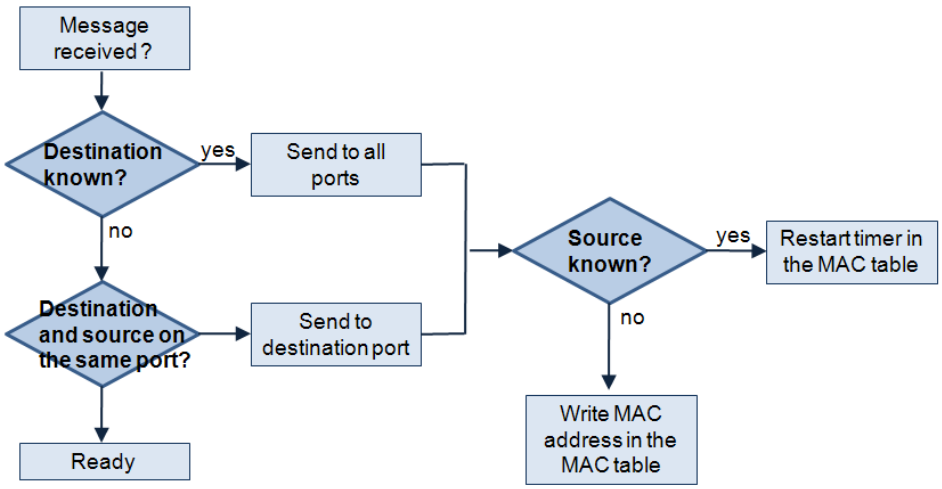


Figure 5.1: The switch is self-learning from the incoming messages

A MAC address table is updated from a software point of view for every port. The switch is self-learning, this MAC address table is filled by studying all incoming messages on the relevant segment. The source address of every incoming message on a certain port is copied to the table. Every address is retained for a limited time and is deleted again as soon as a certain time (the hold time) has lapsed. This technique avoids that inactive stations are addressed or that stations are not recognised anymore.

Before a data packet is transmitted from one segment to the other segment via a switch, a

switch checks the MAC address and on this basis the transport to the other segment takes place or not. Switches function in accordance with the store-and-forward principle.

A store-and-forward switch will first take the complete data frame in, check it for errors and then forwards it via the correct port. The latency depends on the message size.
The preamble precedes a message. Between the preamble and the Ethernet frame is an interval time- an inter frame gap. This time is equal to the time that is required to place 96 bits on the network. This time amounts to 100Mbps, thus 0.960μs.

The theoretical latency for a store-and-forward switch to forward a message with minimum length (64 bytes) at a transmission speed of 100 Mbps is determined with the formula:
TL = TIG (time for inter frame gap) + (64 $*$ 8 $*$ Bit time) [μs]
TL = 0,960 + (64 $*$ 8 $*$ 0,01) = 6,08 μs.
The maximum message size is 1518 bytes. This corresponds with the following latency:
TL = 0,960 + (1518 $*$ 8 $*$ 0,01) = 122,4 μs

5.2 Industrial switches

5.2.1 General

For industrial switches, the first distinction is made between two different categories:

- non-managed switches
- web-based managed switches

Nothing can be configured for the first group of switches. So nothing has to be configured for the general functioning of the switch.

The second group of switches can be configured via a web server. Such an approach is also interesting for diagnosis options of the network.

Figure 5.2: The FL SWITCH SFN 8GT

Figure 5.2 shows the industrial switch (FL SFN 8TX Gigabit switch) of Phoenix Contact. Some typical technical features of such switches are:

- 10/100/1000 TX, auto negotiation, auto crossing
- Non-managed, no software configuration

- mounted on DIN rail, alarm contact, redundant power supply,
- Temperature range: -25°C to +60°C

5.2.2 Technical description of an industrial switch

All potential features of a switch are discussed by means of the technical description of the FL SWITCH SMCS 8GT from the Factory Line of Phoenix Contact .

Figure 5.3: The FL SWITCH SMCS 8GT

SMCS stands for Smart Managed Compact Switch. This switch is built in conformity with the IEEE802.3 standard and is used to build up controllable automation networks on Ethernet. This switch is the type with eight RJ45 ports to which a twisted pair cable can be connected. All ports support 10/100/1000 Mbps. All ports support autonegotiation and autocrossing.

Besides the use as normal standard Ethernet switch, the switch is particularly suitable for Profinet RT and Ethernet/IP applications and supports the management functions that are required for this. The switch supports IGMP Snooping for Ethernet/IP.

Redundant network structures can be built in accordance with the (Rapid) Spanning Tree Protocol or the Media Redundancy Protocol. This guarantees an optimum functioning of the network- regardless of the used topology.

Within complete network systems, information can be retrieved from the switch via SNMP. The configuration and diagnosis can take place via a web server and V.24 (RS232) interface.

The SMCS SWITCH is of the store-and-forward type. All datagrams that reach the port switch are first stored in a buffer and their validity is checked. Corrupt data packets, i.e., packets that are larger than 1522 bytes, that are smaller than 64 bytes or packets with CRC errors, have to be deleted. Valid data packets are then immediately forwarded via the correct port. The transmission speed is determined per port by the linked network segment.

The switch dynamically learns all addresses of the different network participants by evaluating every incoming message on the source address. The switch should be able to store up to 8000 addresses in its MAC address table with an *aging time* of 40 seconds (default setting for delivery). This time can be set from 10 to 825 seconds via SNMP or Web-based management. This means that all addresses that are longer than this time are not used anymore and are automatically deleted from the MAC address table.

The switch is equipped with an alarm contact. The alarm contact is floating and closed when the switch functions correctly. This contact should open, in accordance with the functioning of the switch, in the situations described below. The switch will carry out hardware self-test in case of a restart. The alarm contact opens when an error occurs during this self-test. During the normal functioning, a watch dog will follow the cyclic execution of the software programme. If this watch dog is not triggered cyclically by the software, then the alarm contact will be opened.

Different status LEDs inform the user about the switch status. This enables a local diagnosis without having to use the matching tools.

The SMCS switch supports autocrossing. This means that no distinction is made any longer between a straight-through cable or a crossed Ethernet twisted-pair cable.

The SMCS supports autonegotiation. This must enable the switch to automatically detect the parameters of a certain subnet on every port and to apply these parameters on this RJ45 port. These parameters are the applied transmission speed (10, 100 or 1000 Mbps) and the transmission mode (half duplex or full duplex). This automated detection makes manual interventions by the user superfluous. The autonegotiation function can be activated or de-activated via web-based management.

When using twisted pair cables where a polarity is wrongly connected (more specifically, RD+ and RD- are reversed), then the switch will automatically switch the polarity. This feature is known as *auto polarity exchange*.

The switch checks, at regular intervals, the connected subnets on every port. The used link test pulses, as described in the IEEE 802.3 standard, check the connected TP/TX cables for short circuit or interruption.

The switch can obtain an IP address in two different ways. Via the BootP protocol or via the serial V.24 interface. On delivery, the allocation of the IP address is set on BootP. Configuration software is available to give the switch an IP address in a simple way. The mechanism to allocate an IP address can be manipulated via web-based management or V.24 interface.

The switch can be set to smart mode via a MODE button at the front of the module. In smart mode, the switch can be set to another user mode without the use of the management interface. Furthermore, it is possible to reset the switch in smart mode to default settings.
The switch can be configured as Profinet IO device. The operating mode can be set to default (normal Ethernet switch) via web-based management or in smart mode or on Profinet IO. If the switch is configured as Profinet IO device, then the switch can be included as Profinet IO device in the Profinet engineering software. This way a byte of diagnosis information is available per switch port in the engineering software.

The SMCS switch supports the LLDP protocol in conformity with the IEEE802.1ab. The switch sends and receives management- and connection information from neighbouring devices. This way, network architectures can be represented visually or followed up via available software tools. Profinet engineering software uses this to visually map out network diagnosis.

The switch support two queues for priority (traffic classes in accordance with IEEE802.1D). Received data packets are allocated to these queues, depending on their priority. The priority

is indicated in the VLAN tag of the Ethernet frame. This ensures that data with a high priority do not get delayed by large quantities of low priority data. In case of an overload, the data with a low priority are not read anymore. This principle is used by Profinet RT, amongst other, and is called Quality of Service.

A VLAN tag in conformity with the IEEE802.1Q can be processed by the switch. This tag consists of four bytes and can be found in the Ethernet frame between the source address and the field type. There are three bits in these four bytes that indicate the priority. Different VLANs per port can be set on the switch via the web-interface. This way, different VLANs can be built within a network architecture with such switches. This means that different logical networks can be created within one physical network.

The switch supports Spanning Tree Protocol (STP) and Rapid Spanning Tree protocol (RSTP). STP is described in the IEEE802.1d norm and allows the formation of ring structures or mesh structures in the network topology. Due to structuring of the mesh, several connection paths can exist between two devices. In order to prevent endless loops and broadcast storms, some connections are interrupter by the switch. In case of a break in the cable, the network will recover after a certain time (20..50s) by reconnecting the disconnected ports. Disconnected ports can still receive data but not send data anymore. Only connected ports can send data.

The RSTP is a updated version of the STP and can ensure recovery times from 1 to 10s. The RSTP also supports mesh- and ring structures. RSTP Fast Ring Detection function can be activated for the RSTP configuration. The function is only possible for 10 or 100 Mbps. Faster recovery times can be obtained with the Fast Ring Detection function.

The SMCS supports the Media Redundancy Protocol (MRP). For a ring topology, a recovery time of 200 to 500 ms is ensured in case of an error.

Via SNMP (Simple Network Management Protocol), the device can be monitored via the network. An SNMP management system has the option to read configuration data of the device and adapt and carry out a diagnosis. Versions 1 and 2c of SNMP are supported. The following MIBs are supported: RFC1213, RMON MIB, bridge MIB, If MIB, Etherlike MIB, Iana-address-family MIB, IANAifType MIB, SNMPv2 MIB, SNMP-FRAMEWORK MIB, P bridge MIB, Q bridge MIB and the own SNMP objects of Phoenix Contact (FL-SWITCH-M MIB).

A local connection with the switch can take place via an V.24 interface (RS232). The cable is connected to the COM port in case of a PC and to a mini-DIN socket in case of a switch. For this local connection, the communication takes place via a programme such as HyperTerminal. Via this interface, the IP address and the subnet mask can be set together with the standard gateway. BootP (used for the automated allocation of an IP address) can be switched on or off. The parameters can be stored via this interface and the device can be restarted. The resetting of the parameters to their standard default settings is also possible.

A following interface is the web interface. This interface allows diagnosis and configuration during the start, the use and in case of an error. The web interface also provides network- and device information. With web-based management complete device information can be requested using the popular Internet Explorer. Technical data, installation data, local diagnosis data and all commands for the serial interface can be requested. Furthermore, all configuration parameters can be checked and adapted. This means IP configuration, SNMP

configuration, software updates and passwords. The item 'Switch Station' enables the follow-up of all sorts of diagnosis information about the different ports and the alarm contact. Every port can be activated or de-activated individually. All transmission parameters can be adapted and statistics about the data themselves can be requested via web-based management. The function 'Port Mirroring' can also be activated. With this function, it is possible to send all data, that are sent via a certain port, via another port. This is important as errors can be detected this way via a network sniffer.

The SMCS switch is equipped with a memory plug, FL MEM PLUG.
Some general technical and mechanical data:

- The device is mounted on a DIN rail.
- The insulation class IP20 (protects against fixed objects greater than 12mm, no protection against water) DIN40050, IEC60529.
- The protection class is class 3 VDE 0106, IEC60536.
- Power supply 24V DC (18.5V - 30.5V) and can be effected on a section of maximum 2.5 mm^2.
- It is possible to provide the device with redundant power supply.
- The earthing is done via the DIN rail on which the device is mounted.
- The power consumption is 600mA (15W).
- The device is 128mm wide, 110mm high and 69 mm deep and weighs 650g.
- The operating temperature ranges from 0°C to 55°C and the storage temperature (idle condition) is between -40°C to 85°C.
- The switch should work and stored safely in areas with a humidity between 10% to 95% without condensation.
- The air pressure could be 80k Pa to 108 kPa at 2000m altitude when functioning and 70 kPa to 108 kPa at an altitude of 3000m above sea level when stored.

Chapter 6

The router

6.1 Introduction

A router is a device that connects two or more distinct computer networks with each other, for example internet and a corporate network. Figure 6.1 shows that a router can be considered as a switching device for data packets that are active in layer 3 of the OSI model.

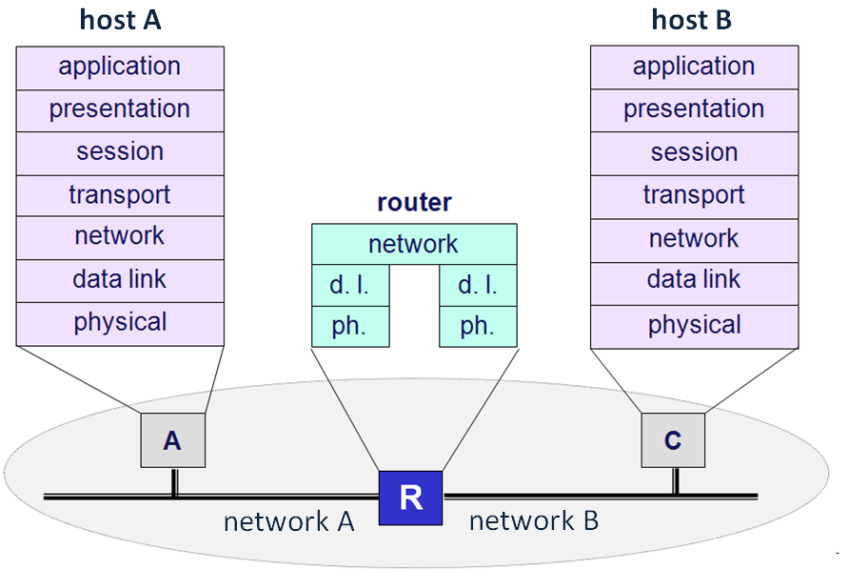


Figure 6.1: The router and the OSI model

6.2 Routing messages

A message that is sent from the one computer to the other computer over an internet has to be handled by several routers. A sender will send the IP packet first to a first router. The sender encapsulates the IP packet in a frame with a header in conformity with the physical network on which the sender and the router are linked.

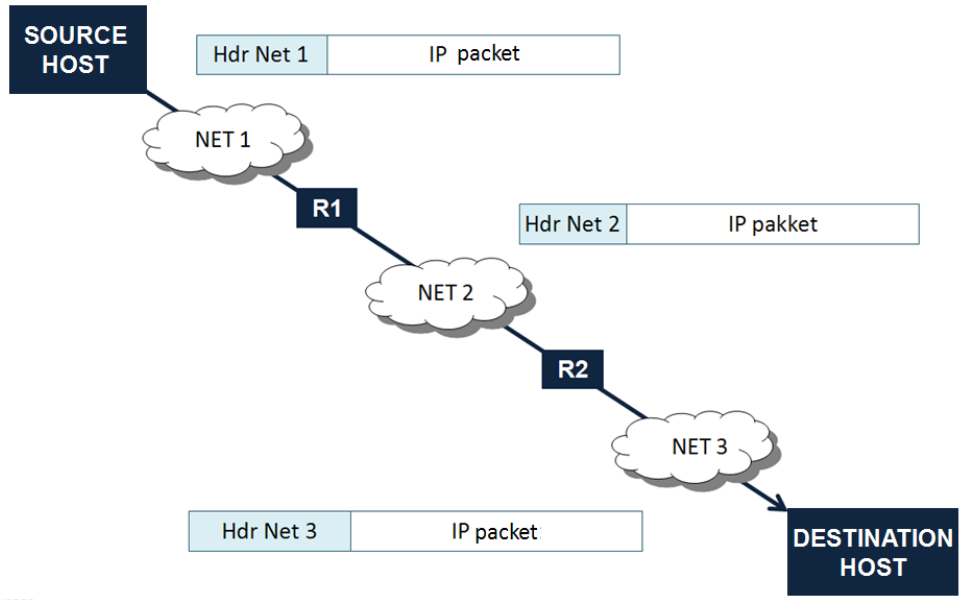


Figure 6.2: An IP packet at the different steps during the route over an internet

When the frame reaches the router, it unwraps the frame and checks the IP packet. The router must now know via which port it has to forward the message. In order to choose the correct outbound port, the router looks up the destination address of the packet to be routed on its routing table. For the TCP/IP protocol, a routing table consists of a table with IP addresses or grouped IP addresses (subnet) and the corresponding next node (next hop). The next node is usually another router that is linked via one of the router ports.

In case the destination address can be routed and is therefore present in the routing table, then the router will use the corresponding next node to determine the outbound port. The inbound IP packet is sent to the outbound port. The router will encapsulate the IP packet again with a specific header to the physical network to which both routers are linked. Figure 6.2 shows that an IP packet is encapsulated every time in a frame that corresponds with the physical network.

It is clear that a router has an IP address for every port that belongs to the range of the Net ID to which the router is linked. But every port also has a physical hardware address corresponding to the subnet protocol to which this router port is linked.

The router builds a routing table by exchanging route information with neighbour routers. This creates a complete picture of all routes in the IP network. The router will build a routing table based on the *shortest path algorithm* (Edsger Dijkstra), whereby the shortest path to the final destination is chosen. In other words: the node that is selected is part of the shortest path. There are different routing protocols for this.

A router is considered as an output device. A data packet can normally go through a certain number of routers before reaching the final destination, determined by the TTL value (Time to Live) of the packet.

6.3 Types of routers

There are many different types of routers. They can be distinguished by their shape, the router connections and all sorts of extra functions that are built in the router, such as a modem, firewall or a switch.

one can distinguish between software routers and hardware routers. By means of software, a simple PC, equipped with two network interfaces, can function as router. A hardware router is a separate device. The device is actually a small, simple computer that has been developed especially for the routing.

Commercial routers for home use are often combined with a switch, equipped with a modem and wireless AP, so that only one device is required to link a private network to the Internet.

Switches with router functionalities are available in the market. These devices are often called layer-3 switch.

Industrial routers are discussed elsewhere in this chapter. In its most simple form, such a router is equipped with a LAN and WAN interface. An industrial network can be linked with this to a corporate network or the Internet. The industrial routers also have all sorts of extra functionalities so that they can be deployed as a complete security module for a safe linking of industrial networks to corporate networks.

6.4 Layer 3 switch

As described, network switches operate at Layer 2 of the OSI model and network routers operate at Layer 3 of the OSI model. A Layer 3 switch is a high-performance device for network routing.

Layer 3 switches differ slightly from standard network routers. Both process incoming messages and, based on the addresses given in the messages, dynamic decisions are taken on how to forward (route) these messages. They were created in response to the need for routers that can easily be used in extended LAN networks such as company intranets.

The major difference between Layer 3 switches and standard routers is the hardware structure. In a Layer 3 switch, the hardware of a switch is combined with that of a router so as to be able to guarantee better performance in routing in larger LAN infrastructures. In its typical use for intranets, a Layer 3 switch has no WAN port and normally does not support typical WAN applications either.

6.5 Linking of a private network to the Internet

An industrial router will be used to link an automation network to a corporate network or to the Internet. For the automation network based on Ethernet, a Net ID will have to be selected. A Net ID preferably has to comply with the RFC 1597.

Figure 6.3 shows an example. The router will obtain an IP address on the LAN side that belongs to the range of the chosen Net ID. This is usually the first or the last free IP address

of the network. On the other hand, the network interface also has a MAC address on the LAN side. The router will function as default gateway on the network.

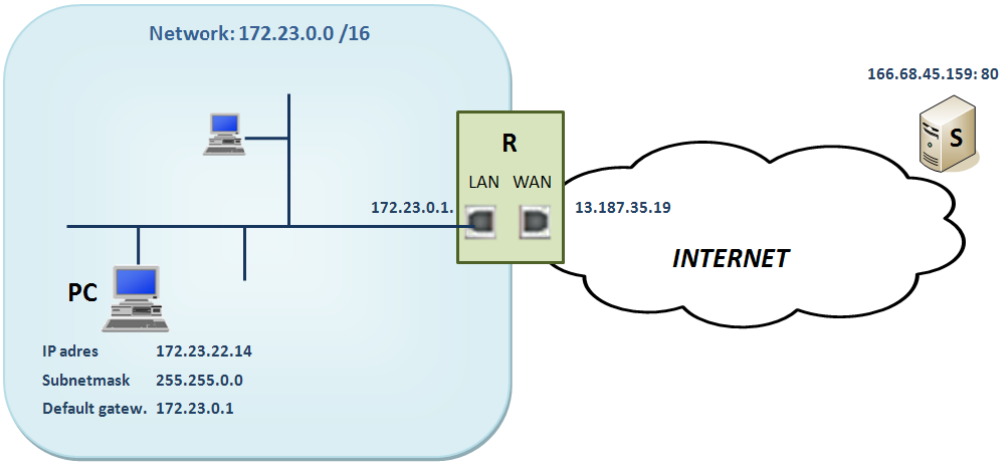


Figure 6.3: Linking of a private network to the Internet via a router

A network can be linked to the Internet via the WAN interface of the router. For this purpose, the router, usually via DHCP, gets a unique IP address on the Internet from the ISP (Internet Service Provider).

Every device on the network can now be configured as follows.

IP address	172.23.22.14
Subnetmask	255.255.0.0
Default gateway	172.23.0.1

Thus every participant gets an IP address where the Net ID for every participant is the same but the Host ID for every participant is unique.

If an application on a PC on the network wants to start a communication with a Server on the Internet, then the PC first has to build a first IP packet to request the connection. This IP packet is sent on the Internet via the default gateway. The PC encapsulates the IP packet in an Ethernet frame for this purpose. Figure 6.4 shows the data that are required to make the Ethernet frame. The MAC address of the router is obtained via the ARP protocol.

Once the message has arrived at the router, this will forward the IP packet via the WAN interface to another router on the Internet. As the private network is separated from the Internet, the router will replace the source IP address of the PC with its IP address on the WAN side. The private network can only be reached via this external router IP address, via the Internet.

The server can now send a reply and will address this reply to the external IP address of the router. A problem occurs now as the router will have to decide to which PC this message has to be sent. The Server's response does not contain any data anymore with regard to the original sender IP NAT was developed to solve this problem.

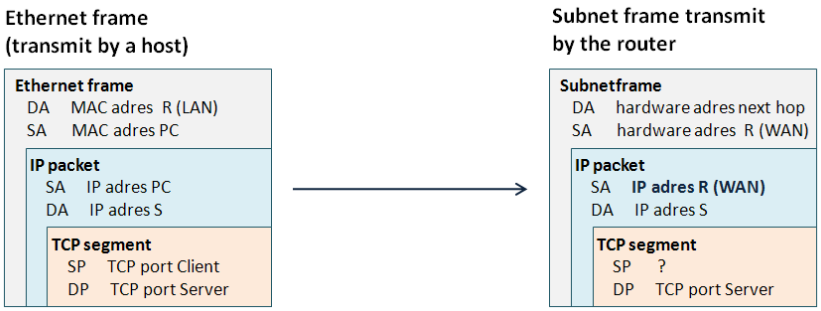


Figure 6.4: Adaptation of a data frame by a router

6.6 IP NAT

6.6.1 NAT: IP masquerading

Network Address Translation (NAT) is a protocol that allows to link a network to the Internet with unregistered IP addresses (a private network that complies with RFC 1597). As described in the previous paragraph, the router places its external IP address as source IP address in every message that is sent from the private network on Internet. All replies will now be addressed to the external IP address of the router.

The NAT protocol permits the router to change the TCP source port field. In a NAT table, all new port numbers are linked to an internal endpoint. Every reply from the Internet to a PC on the private network will be addressed to the external IP address of the router but will contain a port number from the NAT table of the router as TCP destination port. This way, the router knows for which endpoint this message is intended.

From a practical point of view, NAT is a protocol that translates an IP address on one network into an IP address known on another network. One network is called *the inside*, the other network the *the outside* Typically, a company will translate its local inside IP addresses into one or more global outside IP address and will translate the global IP addresses of inbound messages back into inside IP addresses. NAT allows a company to only use one global IP address for its communication with the outside world via the Internet. This contributes to the security concept as all outbound and inbound messages are subject to a translation of addresses.

Figure 6.5 shows the functioning of the NAT protocol. This is the dynamic use of the NAT protocol and therefore sometimes called the dynamic NAT.

6.6.2 Port Forwarding

Port forwarding is the static use of the NAT protocol. If Servers are present on the private network that have to be reached directly via the Internet, then the endpoints of these Servers can be linked in a static way to port numbers in the NAT table of the router. In order to reach this server on the Internet, the external IP address of the router has to be linked as endpoint to the port number from the NAT table. The router will translate the endpoint in the inbound message, intended for that specific Server, to the correct endpoint of the Server. This is

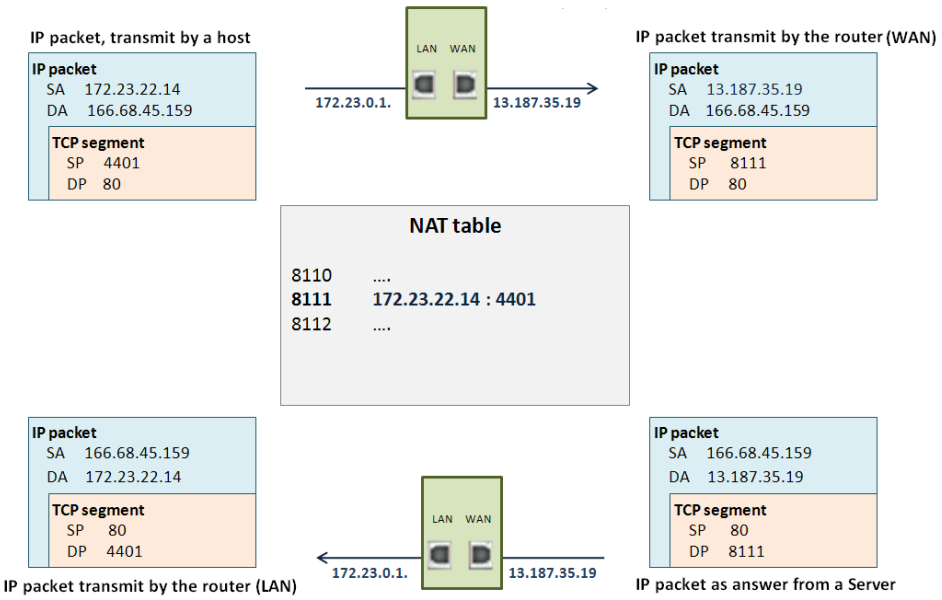


Figure 6.5: Functioning of the NAT protocol: on a PC with IP address 172.23.22.14 , the *http://166.68.45.159: 80* command is given

already an additional form of security. The exact IP data of the Server should not be made available so that hackers do not have any idea of the architecture of the network on which the Servers are located. Figure 6.6 shows which configuration has to take place for port forwarding or static NAT.

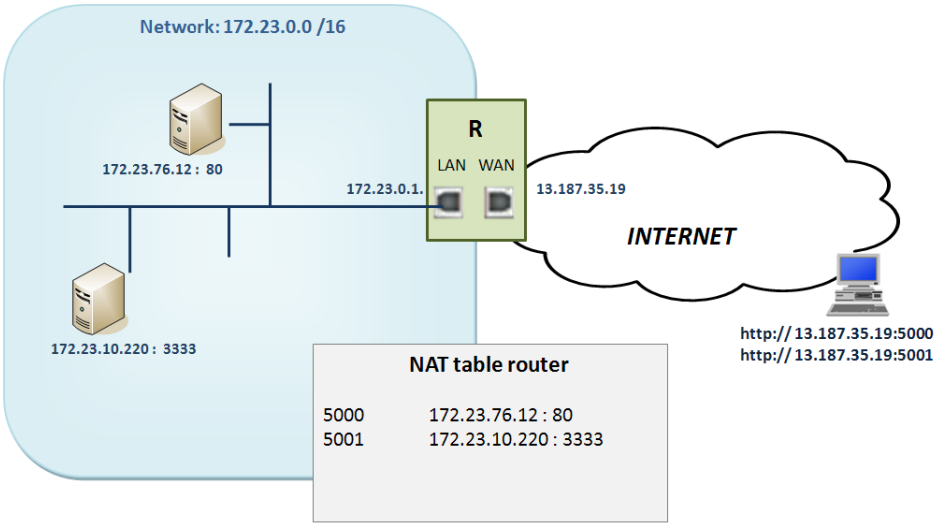


Figure 6.6: Port forwarding

6.7 1:1 NAT

1:1 NAT is a protocol whereby an IP address is translated into another IP address without changing the used TCP/UDP ports.

If a router is connected on the LAN side to the network 192.168.1.0/24 and via the WAN port with the network 10.1.0.0/16 and has 10.1.1.0/16 as external IP address, then the LAN participant with IP address 192.168.1.100 can be reached by means of the 1:1 NAT, via the WAN side, via the IP address 10.1.1.100.

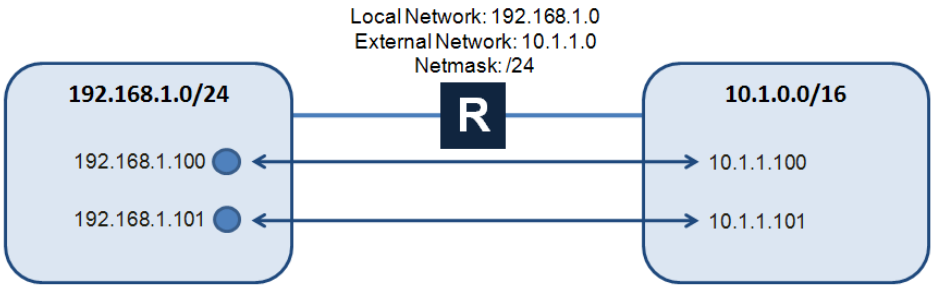


Figure 6.7: Mapping of the IP address with 1:1 NAT

1:1 NAT offers interesting options for the automation world:

- Different subnets can be linked to each other whereby a same IP addressing is used on all subnets.
- No additional routes have to be defined on the corporate network.
- An ARP demon on the mGuard processes the ARP requests from the external network.

- The IP mapping can immediately consult systems on subnets from the corporate network. The HOST ID is retained for this mapping and only the NET ID is adapted.

Figure 6.8 shows the functioning of 1:1 NAT.

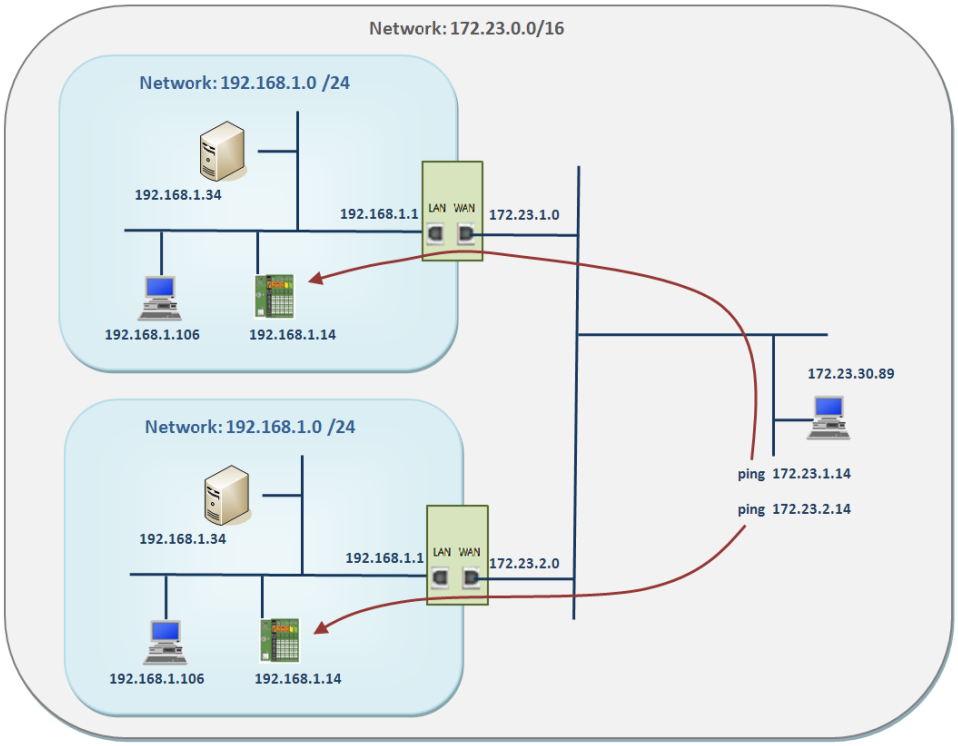


Figure 6.8: Principle of 1:1 NAT

Chapter 7

The firewall

7.1 Introduction

A firewall is an application that controls the access to the data on the network. A firewall is designed to refuse all traffic, except the traffic that is explicitly allowed to pass.

Sometimes, a router and a firewall are mistakenly confused even though there is a fundamental difference. A router is a network structure element meant for forwarding data traffic as fast and efficient as possible and is meant for blocking traffic.

The use of firewalls certainly does not have to be limited to Internet connections. It may also be interesting to deploy firewalls within internet networks to protect certain parts separately. Firewalls can usually be configured to also register all traffic in a log book and can carry out centralised management functions.

7.2 Types of firewalls

There are two types of firewalls. These are software firewalls and hardware firewalls. A software firewall is a firewall that is installed as a programme on a computer. A hardware firewall is a separate device such as a router with integrated firewall. Both types of firewalls (software- and hardware firewalls) function in a similar way. The used terminology is applicable to both.

Firewalls can also be differentiated based on method of functioning. The distinction is made based on the way in which a decision is made to let data pass or not.

- Packet filter: based on a number of rules, the firewall decides whether an IP packet can pass or is to be refused. Such rules are built by means of IP addresses, domain names, protocols (http, ftp, telnet,...) and port numbers. Such firewalls function in a simple and fast way. The packet filter actually functions as a gatekeeper. It roughly screens the passing messages. This way is checked whether the messages are incoming (inbound), outgoing (outbound) or just passing through (route). The specified, but easy to be falsified (spoofed) origin and the final destination (IP address and port number) are checked. The specified type and character of the message is checked in the transport layer. But what is actually in the message is not checked.

Packet filters are stateless. They check the origin and the destination but cannot evaluate suspicious pattern in a certain session. For example, it is not possible to detect when several data packets are suddenly exchanged between certain applications.

- Stateful inspection: beside the different rules in accordance with a packet filter, a firewall can retain intermediate information with regard to all connections that run over the firewall. Stateful Packet Inspection (SPI) means that each packet checks the context after the signing-on and the following handshaking between the communicating hosts. Stateful inspection will check during the complete session what is allowed in accordance with the connection request. First is checked, like with a stateless packet filter, whether the connection between the source and the target by itself is allowed. If not, then the synchronisation request is rejected. If the connection is allowed, then the information from the first datagram (that the session (SYN) builds during the session) is stored in a state-table database in the memory. If something strange occurs within the context of that connection (a host suddenly changes his IP address or his target port), then the session is aborted.

The latest firewalls are stateful inspection firewalls.

Chapter 8

VPN

8.1 Introduction

Data packets are completely unprotected when sent over the Internet. This means that there is no:

- data secrecy (encryption)
- identity guarantee of the sender (authentication)
- check whether data are corrupt or not (integrity)

A Virtual Private Network (VPN) is a private communication channel that functions as an umbrella above a public infrastructure- most of the time the Internet. The data to be sent via this service are secured in such a way that the integrity, authorisation and authentication of the data remains guaranteed. The final users will in principle not notice that a VPN is used. There are different protocols that provide this service.

8.2 Internet Protocol Security, IPsec

IPsec (Internet Protocol Security) is the most widely used VPN protocol. IPsec makes encryption of the data between two communication partners possible. IPsec can be implemented transparently in a network infrastructure. IPsec (an acronym for Internet Protocol Security) is a suite of protocols that ensure together that IP packets can be sent over an IP network. So, IP security offers internet users the option to send data in a secure way. For this purpose, the suite of protocols ensures that the following services are active during the sending of an IP data packet:

- Integrity: the protocol offers the guarantee that the sent packet is not modified by a third party during the transport.
- Authentication: the protocol records the identity of the communication parties. During a secured data transport it must be guaranteed that the party intended to receive the packet is the actual receiving party.
- Acknowledgement: the protocol shows that when a data transport takes place, the receiving party cannot deny this.

- Confidentiality: the protocol ensures the actual security of the data and guarantees the sender that only the receiving party can read the message.

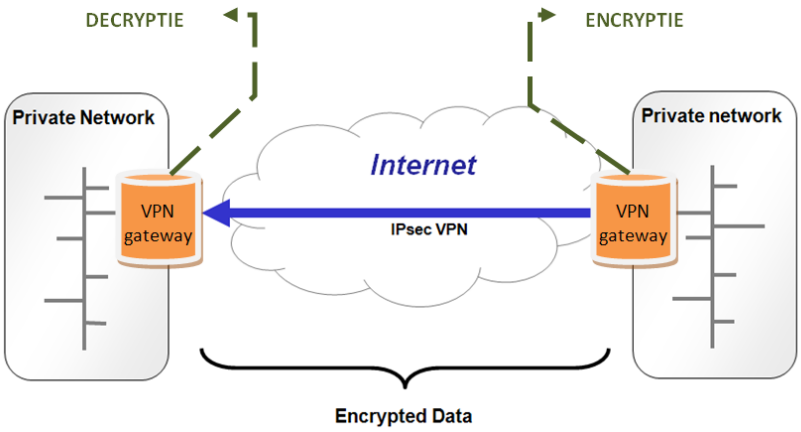


Figure 8.1: Internetworks

The protocol is mostly deployed for the sending of information via public connections and prevents so-called 'Man in the middle attacks' and 'Spoofing'. It uses the IKE protocol (Internet Key Exchange) for this with which the parties that want to set up a connection are identified. Next, a connection is set up and the data to be sent are secured by means of encryption .

Encryption is used for many protocols in order to realise data secrecy. For encryption, the data will be transformed into an illegible form, the so-called cyphertext. By means of a key, the receiver can carry out the reverse transformation (decryption) which makes the text legible again. The most used encryption techniques today are 3DES and AES. AES is always preferred for its stronger crypto properties compared to 3DES. 3DES is obsolete and is used only if it is not possible to use AES.

IPsec supports security as of the third layer of the OSI model (called the network layer). This means that TCP as well as UDP can use it but it also results in overhead with regard to e.g. SSL that functions on higher OSI levels (and cannot secure UDP). IETF established the standard in RFC's 2401-2412 - optional for IPv4 and mandatory for IPv6.

The protocol is built as follows:

- Authentication Header (AH): checksum check for the complete IP packet.
- Encapsulating Security Payload (ESP): safeguards against Man-in-the-middle attacks.
- IP payload compression (IPcomp): compression of the IP packet payload before the encryption takes place.
- Internet Key Exchange (IKE): assists in the setting up of the connection by safely transmitting keys/certificates.

IPsec exists in two variants:

- Transport: encrypts the content (payload) of the IP packet but not the header. No new IP packet is created in this mode but the headers (AH or ESP or both) are inserted into the IP packet. The source- and target addresses remain unchanged.
- Tunnel: encrypts the content as well as the header of the IP packet. In this mode, the complete IP packet is inserted into a completely new IP packet. The IP packet has the start and end point of the tunnel as source and target address.

8.3 VPN implementations

VPN can be implemented in three different ways.

1. Security gateway to security gateway

Figure 8.2: Security gateway to security gateway

2. Host to security gateway

Figure 8.3: Host to security gateway

3. Host to host gateway

Figure 8.4: Host to host gateway

Chapter 9

Automation networks & Security

9.1 Corporate network

A corporate network is the set of servers, computers and systems that enable the general functioning of the company on IT-level. Ethernet TCP/IP is already for years the standard when it comes to setting up IT networks in office and companies. A corporate network is, in its simplest form, linked to the Internet via a router and a firewall. Larger corporate networks also provide a DMZ - this is a part of the network containing public servers (mail server, web server, DNS server,...).

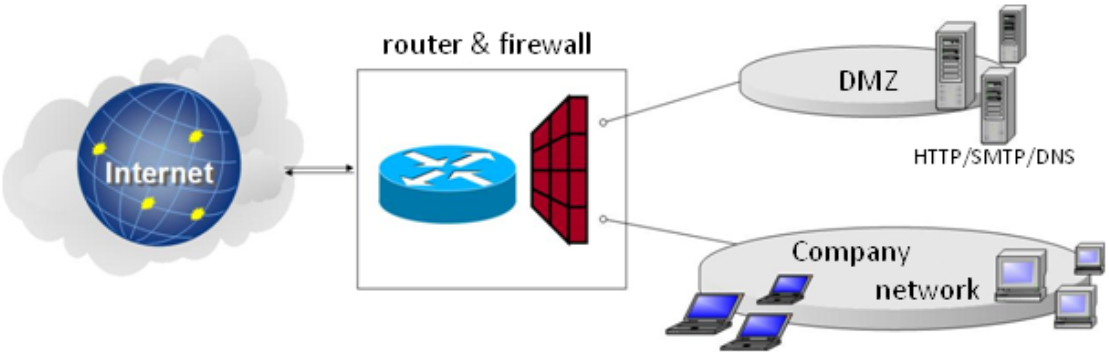


Figure 9.1: A corporate network

A router, in its simplest form, is a device that enables communication between two networks. Specifically, this is the corporate network (LAN) on one side and the Internet (WAN) on the other side. Firewalls are used to block unwanted communication whereby IP packets are filtered in accordance with rules that the user has established. Incoming as well as outgoing communication can be blocked. The filter criteria can be IP addresses, port numbers or certain protocols that can be blocked or released by choice.

9.2 Corporate network

9.2.1 Automation cell

An automation cell is the set of PCs, data servers, controllers, IO devices, sensors and actors that are required to execute the different functionalities of the automation concept.

An automation project is the set of

- Production lines and process installations
- PLC systems (Programmable Logic Controllers)
- ESD systems (Emergency Shut Down and Safety Controllers)
- DCS systems (Process and distributed Control Systems)
- SCADA systems (Supervisory Control and Data Acquisition)

Switches are the structure elements with which a complete automation cell is built further. The combination of different topologies and media makes a flexible, safe and controllable network (based on Ethernet TCP/IP) laid out on the industrial shop floor.

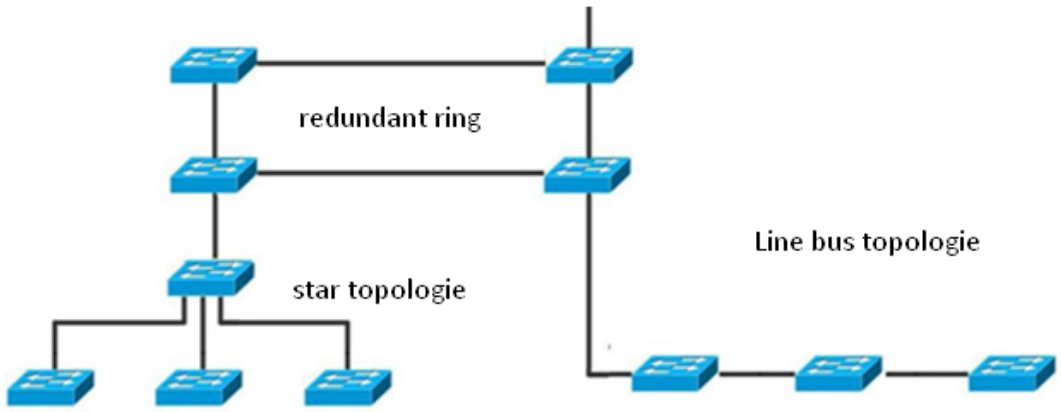


Figure 9.2: Different topologies in an automation cell

Important segments are connected to each other via switches in a redundant ring. In certain subsegments, network participants are connected in a star via switches (such as in normal IT networks). Where possible, a line structure is applied to connect participants with each other. In order to make a line structure possible, it is necessary that all IO devices are by default equipped with an integrated switch.

9.2.2 Automation network

An automation network consists of one or more automation cells. Every cell is separated here by a router.

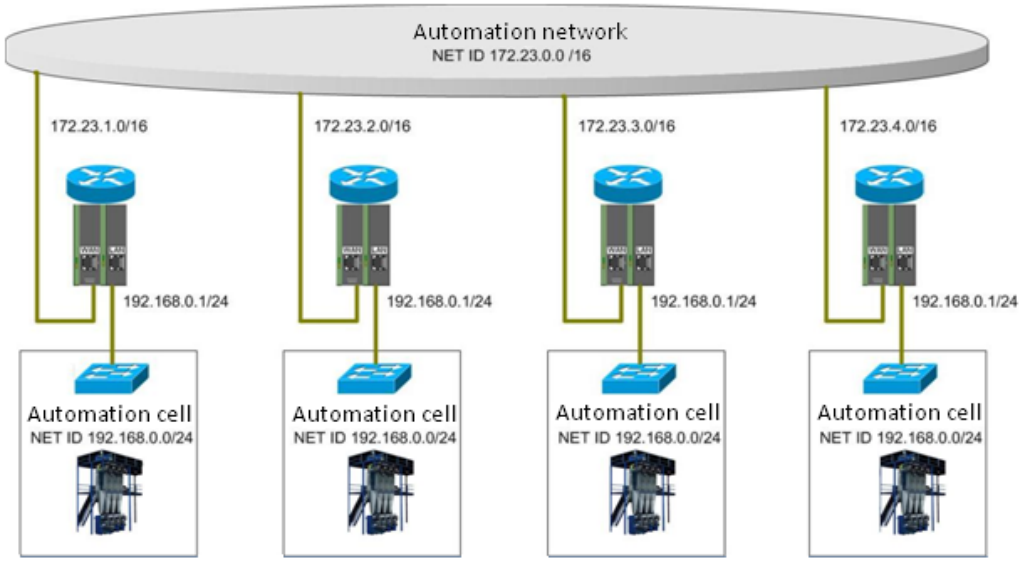


Figure 9.3: Automation network

Cabling and connectors in an automation network

The extension and the cabling as used in the office environment cannot be applied in the same manner to raw industrial environments.

Cables, connectors and infrastructure elements have to be aligned with the production environment where impacts like humidity, temperature variations, shocks or vibrations can occur. These elements, connectors and cables, have to comply with the quality requirements of the industry. This is the first step towards a reliable automation network.

It is important that all Ethernet cables can be configured in a simple way on the work floor. It is recommended to carry out the cabling first and then to implement the Ethernet connectors.

Use of switches

Only switches are used as structure elements. The building of the network is important in order to optimise the network load. This can never be higher than 60For the building of a redundant ring structure, the switches have to support RSTP. For the management and diagnosis of the network, the switches within the different automation cells have to support the following protocols:

- Web-based management - for a fast and clear configuration
- SNMP - for device diagnosis
- LLDP - for control and diagnosis of the network topology
- VLAN - for the structured division of the network

Switches should have the option to send SNMP trap messages for different events or to activate an alarm contact. A smart memory plug is used for the simple configuration of new devices. Another important step in a secure network is the use of VLAN.

9.2.3 Linking of an automation network to a corporate network

The linking of an automation network with the corporate network takes place by means of a router. This router ensures an ideal separation between the two networks that have completely different requirements. This router has to establish an open but highly secure communication structure between the corporate network and the automation network.

9.3 Necessity of security

9.3.1 Introduction

Automation networks are up to now mostly isolated networks with controllers and network protocols that are based on proprietary protocols. The production department itself is usually responsible for the industrial communication. Security is rarely a point of attention.

Modern automation projects are characterised by open systems and communication networks based on Ethernet TCP/IP. The IT department is therefore jointly responsible for the industrial communication. Security is an important point of attention.

Windows and Ethernet are sweeping through the production halls and that is an interesting development. But it is becoming increasingly clear that viruses and hackers also have a hold over machine parks and installations. It is therefore important to protect the automation world against the dangers that have been faced for years by the IT world.

9.3.2 Awareness

The realisation and the knowledge to secure the office network is well-established these days. It has become a standard practice to place a firewall between the office network and the Internet, supplemented with a number of additional security measures. This properly secures the office network.

The realisation and this knowledge is well-established on production level. The following questions on the work floor are therefore very obvious:

- Is the production IT so vulnerable that security is vital?
- Nothing can happen when the corporate network is properly secured, isn't it?
- Would any unauthorised person really hack the production equipment and bring the factory to a standstill?
- Moreover, certain other protocols, other than the known Microsoft protocols, run on the systems within the industrial IT. Doesn't this make the production networks less prone to attacks from the outside world?

This last statement was correct in earlier days but now the trend is to use open systems such as Windows-based software applications and protocols like HTTP, FTP or DCOM (used in OPC) till the PLC level. These open systems are virus-prone and can cause the blocking of the PLC.

On the one hand, the problems on the industrial work floor are not concentrated around the intentional hacking but more around the accidental errors within the production. For

example, cables that are pulled out or wrongly plugged in. Production standstill or something even worse can result from the use of an USB stick that it is infected with a virus and is plugged to a PC that is connected to a machine. Data traffic from the office network can cause delays on the production network.

On the other hand, there is a distinct possibility that the data that is hacked will be misused. The company can be blackmailed in such a case. Recent studies show an increasing trend in the area of industrial security incidents. The more accidental events are more and more supplemented with external incidents such as viruses, Trojan horses, system hacking, sabotage,... . Hackers have acquired more and more knowledge of control systems and SCADA applications. Hackers carry out their activity less and less for the fun of it and more and more with the intention to blackmail a certain company. This has become an organised crime.

Security is a *must*.

9.3.3 Objective of security

The main objectives of security are threefold:

- confidentiality: security that data do not end up with a third party.
- Data integrity: protection of the data against unwanted adaptations or against their destruction.
- availability: resources are available and function correctly at the time that they have to do so.

Security will therefore prevent an unauthorised person from entering the system, will make sure that the system functions normally at all times and that all data in the system can be handled in a confidential manner.

9.3.4 Security in the office world versus security in the automation world

Introduction

The integration of open systems can give the impression that the security problems within the production world can be solved by copying the approach in the office world. However, there are important differences between both domains. The office IT is not the same as production IT. It has to be checked what elements from the office IT are used and not used in the production IT. There is a standard under development (ANSI/ISA99) to completely describe the what, how and why of the security in the automation world.

Main objective of security

First of all, there is an important difference in the main objective of security.
In the office world, the main aim for security is always the confidential handling of data.
In the automation world, the main aim for security will always be the availability of the production system.

Network performances

Both domains have totally different performance requirements.
An overview:

Automation network	Office network
Real-time	Not real-time
Response is time critical	Response has to be reliable
Medium throughput acceptable	High throughput required
Significant delays are a problem	Significant delay and jitter is acceptable

Therefore, it is important to be able to correctly estimate the impact of security technologies on the performance of the system before this is implemented. A lot of encryption is applied, for instance, in the office world. Encryption however, does not stimulate the real-time functioning.

Reliability of a network

The requirements for reliability are also different for both domains.
An overview:

Automation network	Office network
Continuous operations	Planned operations
Power cuts are not acceptable	Power cuts are acceptable
Supposedly tested before the implementation	beta testing on location allowed
Formal certification is mandatory for applications	Little paperwork for applications

The installation of a new service pack is a good example for this. In the office world, this is a normal thing, whereas in the industrial world, the installation of a service pack is certainly not logical and in some industrial sectors it is not permitted.

Different risk outlook

Automation network	Office network
Human safety	Data integrity
Risk impact is loss of product or device	Risk impact is loss of data
Error tolerance is essential	Restart via reboot

Furthermore, there are critical response times on human interventions in the automation networks. Operating an emergency stop for example, cannot be obstructed by password securities.

Different security architecture

In the office IT, the central servers are the most critical devices to protect.
For the production IT, the end device, such as the PLC, is the most critical device and not the central data server which contains the historical process data.

Decision

The classic firewall protects against hackers, worm viruses and spyware. The software programmes protect the network or the specific computer against the outside world and allow only trusted messages to pass. One of the objectives is to block all suspect programmes that look for a connection with the Internet from the own location. In addition, there are programmes in the office IT such as anti-virus programmes, anti-spyware and anti- adware programmes that block every inbound file, check for threats hidden within the database and then declare the file clean or place it in quarantine. They also screen all called up files for the presence of viruses, spyware and adware that are included in the database. The principle of this IT protection is that the firewall views and analyses all inbound and outbound data traffic, data files, programmes,... Such firewalls slow down the system and can even prevent the programme from functioning, if necessary.

In order to realise a more industry-compatible firewall, techniques will have to be used that guarantee in the first place that there will be no delays and still guarantee security and reliability. One can use safe communication channels between, e.g., PLC and the master computer or data servers. In order to ensure real-time functioning, no delay should occur when checking the data that pass that firewall. Other technologies will have to be applied. An option is to check the applied protocols instead of the actual data.

9.3.5 Standardisation with regard to security in automation networks

Introduction

ANSI/ISA 99 gives directives for the carrying out of risk assessment, the setting up of a so-called cyber security management and carrying out that management. The standard is set up in collaboration with end users, system integrators and suppliers. The ANSI/ISA 99 norm is currently under development. Currently, 2 parts are available.

Figure 9.4: The ANSI/ISA 99 standard

Figure 9.4 shows an overview of the ISA 99 standard. Part 1 is meant to be a framework around all the other parts. At the moment, only part 1 and part 2.01 are available.

Part 1 (ANSI/ISA report TR99.00.01- 2007)

The title of the first part is: 'Security for Industrial Automation and Control Systems', the last version dates from 29 October 2007. It describes security technologies that are currently available for industrial production and control systems. In this part, technologies such as authentication and authorisation, firewalls, VPN,... are discussed.

Authentication is the process to be able to positively recognise users, devices, applications, resources. Authentication can take place by means of sometimes that has to be remembered (pin code, password,...) , something that is possessed (key, intelligent card, dongle,...) or something physical (finger print,...). A distinction must also be made between 2 different authentications: user authentication and network service authentication.

Part 2 (ANSI/ISA report TR99.00.02- 2004)

Originally, the second part had the title: *Integrating Electronic Security into the Manufacturing and Control Systems Environment*. The original part 2 and part 3 are now finally merged together as part 2. Part 2 will consist of two elements. The first part ISA 99 part 2.01 is ready and has the title 'Establishing an Industrial Automation and Control Systems Security Program'. It is meant to provide support to companies setting up a security management plan. The basis of such a plan is to map out all possible risks and then to formulate a number of solutions. ISA 99 part 2.02 will describe how such a security plan has to be executed.

Part 4

The requirements that are set for equipment and systems are described in part 4 so that these would comply with the ISA 99 standard.

9.3.6 A security programme

The elaboration of a security plan is more than just thinking of technical solutions such as firewalls and encryption of data. Various human factors can contribute to a successful implementation of a CSMS (Cyber Security Management System). Various points of attention that can result in a successful integration:

- Security management has to completely fit in the company policy
- The security programme has to fit in the corporate culture
- Support and commitment from the company management
- Clear budgeting of security management actions
- Separating functionalities: if a production in charge is also responsible for security, then security often comes second.
- Organise activities and training for all employees
- Distribute guidelines among all employees
-

9.4 Security in practice

Security actually means protecting the machine, the production or the process against certain human activities. Human activities can cause, intentionally or not, a production standstill.

In order to apply security, it is not just one rule but different concepts which have to be followed to ensure that human errors are minimised or that persons with malicious intentions are deterred from misusing the available data.

Security on the work floor can be integrated at different levels.

9.4.1 Layer 1 security

A first step for a well-secured network is the mechanical security of the network cables. Safe clips should be used to prevent easy removal of network cables from a network port. The access to free RJ45 ports of the different switches should be limited mechanically, in order to prevent access by unauthorised participants.

Figure 9.5 shows the option to block unused ports or fasten Ethernet connectors.

Figure 9.5: Layer 1 security

9.4.2 Layer 2 security

A second step for securing a network is to use available software to control switches.

Switches have to provide the option to set some important securities:

- Web-based management has to be protected by a password.
- Option should be available to allocate different rights (read-only or read-write protection) on the basis of IP addresses.
- Switches should have the option to set different securities per port. A list of authorised MAC addresses, for example, has to be set per port.

9.4.3 Layer 3 security

The most important step when securing an automation network is the separation of different segments by a security module. A security module is a router with the following options:

- NAT and 1:1 NAT: the application of NAT ensures the translation of IP addresses. This makes it more difficult for an outsider to retrieve the applied IP addressing on the network.

- Integrated stateful inspection firewall.
- User firewall: individual rules for different users.
- Support for VPN technology.

SCATTERGOOD & JOHNSON LTD
ELECTRICAL ENGINEERING & FLUID CONTROL DISTRIBUTORS
Est.1899

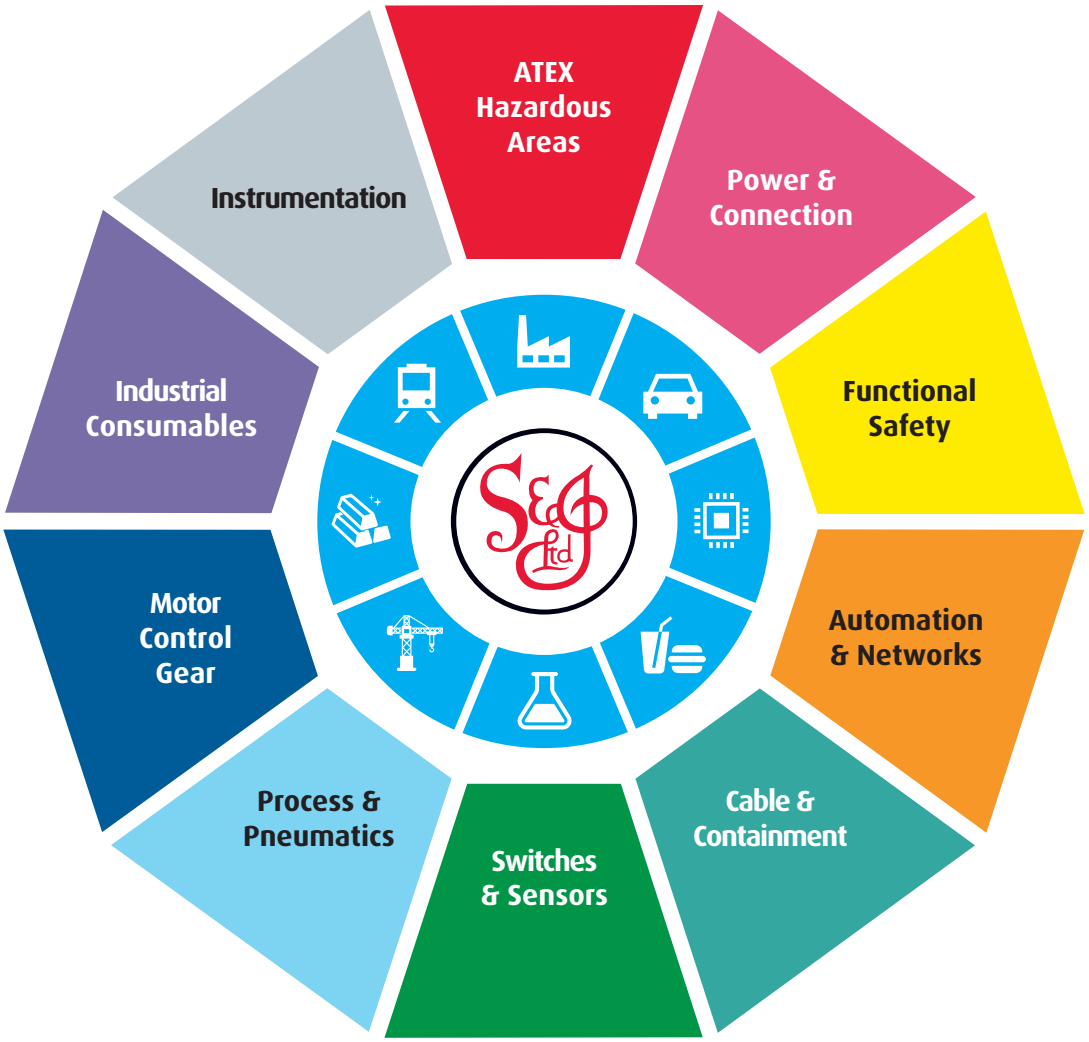
At Scattergood & Johnson Ltd, we pride ourselves on being a technical distributor to specialist industries.

Working with a range of quality product suppliers across a number of specialist markets, we are not your average 'box shifter' - we are your technical and supply chain partner.

We fully support every product we sell - for free! Our internal team and external sales engineers can answer any product or application question, no matter the complexity.

Backing up this technical ability is a range of 50,000+ products available from stock for nationwide next day delivery (same day if required!), or you can collect what you need from any of our trade counters around the UK.

Select your specialist interest below to learn more about how we can help.



Online, In Branch and On the Road - Scattergood & Johnson Ltd, there when you need us.

www.scatts.co.uk