EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION,
YOUTH AND SPORTS

Communication Networks II
(Part 2)

Authors:
Libor Michalek
Petr Machník

Ostrava 2020

Content

1 Addressing in IPv4 and IPv6

1.1 IPv4 Address

The IPv4 is a data-oriented protocol that has been used in the networks with packet switching (e.g. Ethernet). It is a protocol that transports data without guarantee, i.e. it does not guarantee either delivery or retention of order of elimination of duplicates. The guarantee assurance is left on the higher layer that is represented by the TCP Protocol. Similarly, a data integrity control is left on the higher layer because the IPv4 datagram carries only the information about the checksum of the datagram header with service records.

The address is used to identify the computer interface. If the computer has multiple network cards, each card must have its own address. The IP addresses are located on the 3rd layer of the ISO/OSI model and their main goal is to establish a hierarchical network structure. The IPv4 provides limited address space – 2^{32} addresses (about 4×10^9 = 4 billion addresses) in theory. The example of an IPv4 address structure is shown in the Figure 1.1 and demonstrates that the address consists of two parts, i.e. the network identifier and station identifier. [1]

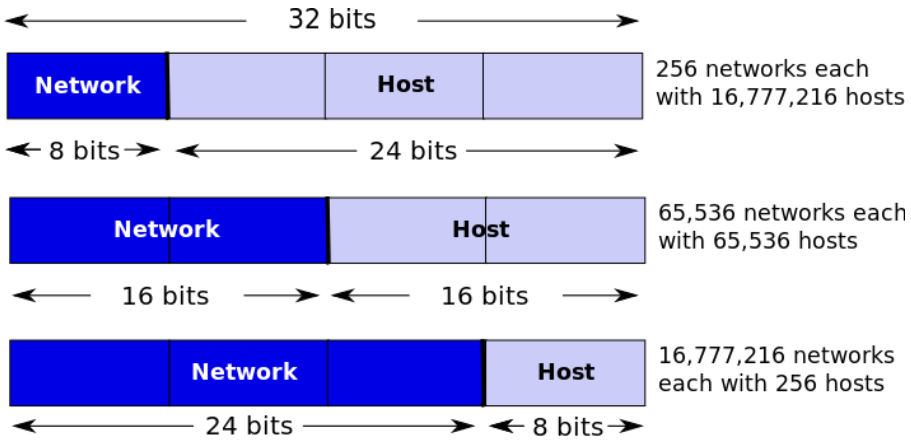

Fig 1.1 IPv4 address

1.2 The IPv4 Address Classes

The original design of the IPv4 assumed the division of the address into the network and fixed local part, first eight bits of the address determined the network. The rest was determined by the machine in the network. Therefore, there could be a maximum of 256 networks (however, each could have over 16 million stations). Along with the beginning of local networks, this system has become useless. The adopted solution inhered the introduction of address classes, class A for a small number of large networks, class B for a medium number of medium networks, and class C for a large number of small networks. Moreover, the class D was defined for multicasting and the class E remained as a reserve.

A successful applicant (which was an institution or company) for a network address was always assigned a network address of the required class so that he had the entire address

space on the subnet. This mechanism of address space management is referred to as classful.
The table 1.1. shows the division of the IP address classes.

Table 1.1 IP Address Classes

Class	Beginning (bin)	1st Byte	Standard Mask	Network Bits	Station Bits	Networks	Stations in Each Networks
A	0	0–127	255.0.0.0	8	24	2^7 = 128	2^{24}–2 = 16 777 214
B	10	128–191	255.255.0.0	16	16	2^{14} = 16384	2^{16}–2 = 65 534
C	110	192–223	255.255.255.0	24	8	2^{21} = 2 097 152	2^8–2 = 254
D	1110	224–239	*multicast*				
E	1111	240–255	*retained as a reserve*				

It is possible to determine the number of networks in each class.

$$N_{networks} = 2^n$$

where *n* is the number of bits of the network identifier.

In general, it is also possible to determine the number of hosts as

$$N_{hosts} = 2^n - 2$$

In case of the number of hosts, it is necessary to subtract two because the first address of each network range determines the network address and the last address of each range determines the so-called broadcast or broadcast address. It is not possible to use the network address and broadcast for addressing the station from the total number. If the broadcast is used as the destination address (it cannot be used as the source address), the packet is delivered to all stations in the network defined by the network mask.

1.3 Network Mask in IPv4

The network mask is a number that indicates how large a given network is. The numerical representation indicates how many bits from the left have a value of 1, for example, it is 11111111 for the mask 8 and the rest are all zeroes. In decimal notation, the mask 8 takes the form of 255.0.0.0. It is very easy to calculate how large the network will be. The mask determines which bits must not change for a given network. Those are marked with ones. The bits that carry a value of 0 can be changed within the network. If the mask has a value of

24, i.e. the mask contains 24 ones, then only 8 bits remain from the IP address, which can be changed. The number of addresses in a given network is 232-24 = 28 = 256.
If the network address is multiplied binary a bit by bit with the network mask, the network address is obtained, see the Table 1.2. If the result of a logical product is converted to a decimal form, the searched network address is obtained: 192.168.1.0.

Table 1.2 Calculation of the network address of the source station

Binary				
Source address	11000000	10101000	00000001	00111000
Network mask	11111111	11111111	11111111	00000000
Network address	11000000	10101000	00000001	00000000
Decimal				
Source address	192	168	1	135
Network mask	255	255	255	0
Network address	192	168	1	0

The routers process the received requests on forwarding packets to a specific destination address identically. If the router has the information about the destination network, to which the packet should be routed stored, in the routing table, it performs the binary multiplication of the destination address with the appropriate network mask, checks the network address and, based on this, selects the interface through which the packet will be forwarded to the destination address.

1.4 Private Networks

The IP Protocol is very popular and is thus deployed to networks which were not or should not be connected to the Internet. The private addresses are commonly used for home, office and corporate local networks (LAN), where the public addresses (i.e., globally routable on the Internet) are not desirable or available. The private addresses are referred to as private because they are not globally delegated, which means that they are not assigned to any specific organization and the IP packets addressed by them cannot be transmitted over the public Internet. Anyone can use these addresses without approval from the Regional Internet Registry. If such a private network needs an Internet connection it must use either Network Address Translation (NAT) or a proxy server.
Three blocks of so-called private addresses according to RFC 1918 were reserved for these networks, see Table 1.3

Table 1.3 Private IP address Ranges

RFC Designation 1918	**Range of IP Addresses**	**Number of Addresses**
24-bit block	10.0.0.0 – 10.255.255.255	16 777 216

20-bit block	172.16.0.0 – 172.31.255.255	1 048 576
16-bit block	192.168.0.0 – 192.168.255.255	65 536

1.5 Automatic Configuration in IPv4

The DHCP Protocol, through the DHCP server, allows to set automatically a set of parameters necessary for communication to the stations in the computer network by means of the IP protocol (i.e. to use the TCP/IP protocol family). The DHCP server typically assigns the IP address network mask, default gateway and DNS server address to the computers using the DHCP Protocol.

1.6 IPv6 Address

The IPv6 (Internet Protocol version 6) is in the name of the emerging protocol for communication in the current Internet (or computer networks that create the Internet). The IPv6 replaces the outgoing IPv4 protocol. In particular, it brings massive expansion of the address space (i.e. the ability to assign their own IPv6 address to all devices) and enhances the ability to transmit data at high speed.

The main change that IPv6 brings is much larger address space, allowing for more flexibility in the address allocation. The large IPv6 address space contains a total of 2^{128} (approximately $3{,}4\times10^{38}$) of addresses. It is impossible to use Network Address Translation (NAT) which was introduced for depletion of the IPv4 address space and for security.

The addresses IPv6 have 128 bits, i.e. 16 bytes. We record them in hexadecimal form unlike IPv6, always in pairs of bytes separated by colons, for example:

FEDC:02A5:0000:002A:00C0:7600:0C12:1C47

The insignificant (beginning) zeros of each quaternion can be omitted from the above-mentioned record and enter the address as **FEDC:2A5:0:2A:C0:7600:C12:1C47**. Since the longer part of the address is often zero, we can express more consecutive zero bits by the symbol "**::**".

For example, the address **FEDC:02A5:0000:0000:00C0:7600:0C12:1C47** can be shorten to the **FEDC:2A5::C0:7600:C12:1C47**. In addition, we have used the possibility to omit insignificant beginning zeros. The record "::" may also appear at the beginning or end of an address, e.g. **::FE12:CCD0 or DC80:FD87:A800::** For the reasons of unambiguity the record can appear only once in an address. For instance, the address **FD08:0000:0000:DAC8:0000:0000:0000:DACD** can be recorded as **FD08::DAC8:0000:0000:0000:DACD** or **FD08:0000:0000:DAC8::DACD**. If we used the (incorrect) record **FD08::DAC8::DACD** the address would not be unambiguous as it would not be clear from the record how many zero positions the first and second occurrence of symbols "::" represent

We often need to express an address prefix. We write the number of bits that form the prefix behind the slash, behind the prefix value, recorded according to the IPv6 address

convection. This is the same notation we are used to from the classless IPv4 addressing (CIDR). For example, the 60-bit prefix is written as **AAAA:BBBB:CCCC:DDD0::/60**.

1.7 Automatic Configuration in IPv6

The IPv6 offers two options – so-called stateful and stateless automatic configuration. The stateful configuration assumes the launching of the special server that allocates parameters of the connection on request. This is the same principle as in case of the well-known DHCP Protocol (Dynamic Host Configuration Protocol, RFC2131) widely used in the IPv4. The DHCPv6 Protocol, which is determined for assignment of the IPv6 parameters, has some extensions in comparison with the classic DHCP. However, the principle remains the same – initially, the station uses a group address to search for a DHCP server that offers it leasing of a certain network address and other parameters along with it, as the default gateway and DNS server address, for a certain period. The station chooses one of the offers and, along with the DHCP server that offered this address confirms the assignment request.

In case of the stateless configuration no DHCP servers are required, the stations will create their address from the MAC address and the local network prefix, which is periodically reported by the router on all interfaces of their LANs. An ICMPv6 Router Advertisement message of the protocol is used for this. The router will also serve as the default gateway for the station.

2 Subnetting

Classless routing CIDR (*Classless Interdomain Routing*) has developed as a solution to the lack of public IPv4 addresses along with massive spreading of the Internet around the world. Classless routing uses address space more efficiently. In case of common routing (classful Routing) a network mask is always assigned to a IP address class and cannot be changed. The IP address 10.10.10.1 where the network mask 255.0.0.0. is firmly fixed is an example. However, the fixed length of the net mask is not ideal and, above all, is not effective. In some cases, the address space may be unnecessarily large (wasting of IP addresses) and, on the contrary, in other cases may be insufficient (few IP addresses).

Therefore, the RFC recommendations from 1517 to 1520 were published in 1993.They radically modify the address space splitting strategy. The nets were no longer viewed in terms of 'classes' and the net mask started to be used exclusively. To avoid ambiguity, it is necessary to move away from the class view on the networks and always add appropriate network mask to the address. This method of subdividing a network into smaller parts is called subnetting. Therefore, the subnet mask has more ones than the standard mask for the given class. However, the ones can also be removed from the mask (replaced by zeros from the right) and thus creating larger networks, so-called supernetting. Then the subnet mask has fewer ones than the standard class mask.

As the decimal notation is very long, a simplified record of the netmask is called a prefix in the form of an integer number that is stated behind the network address after the slash. This number is equal to the number of ones in the netmask. For example, the netmask from the previous example 255.0.0.0 was written as **/8** and the whole record of the network address would look like **10.10.10.1/16**. [1, 2, 3]

2.1 VLSM (Variable Length Subnet Mask)

The VLSM or network mask of the variable length solves the need to divide the network into different sized parts. The practical usage is outlined in the following example:

Let us suppose that our task is to create the network addressing shown in the Figure 2.1. The network consists of five routers in different cities connected by interconnecting lines. In individual cities, the switches are connected, which ports carry stations of the local area network. The numbers listed for individual switches indicate how many stations connected to the network are counted at that location.

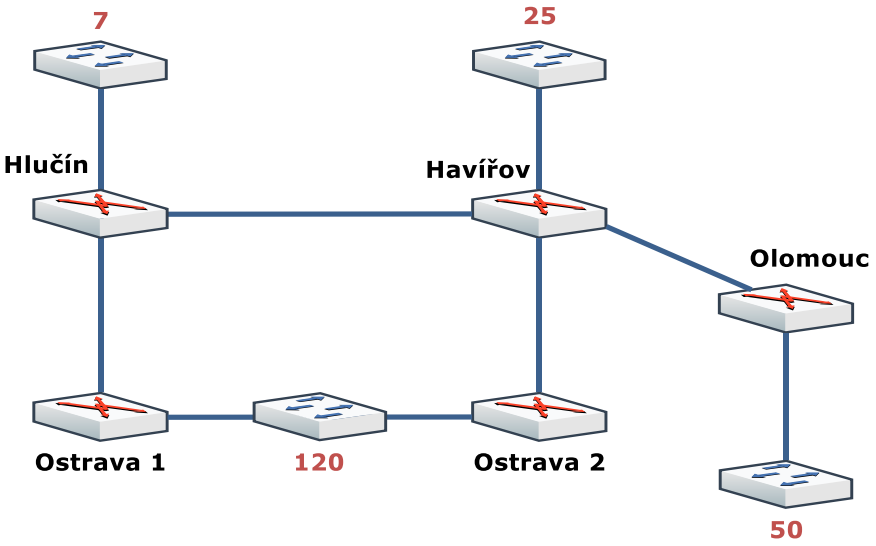

Fig 2.1 Network Architecture

To design a network addressing, first of all, we must determine the number of the IP subnets and specify the minimum number of bits that must be reserved for station addressing in each subnet. The areas forming individual subnets are shown in Fig 2.2.

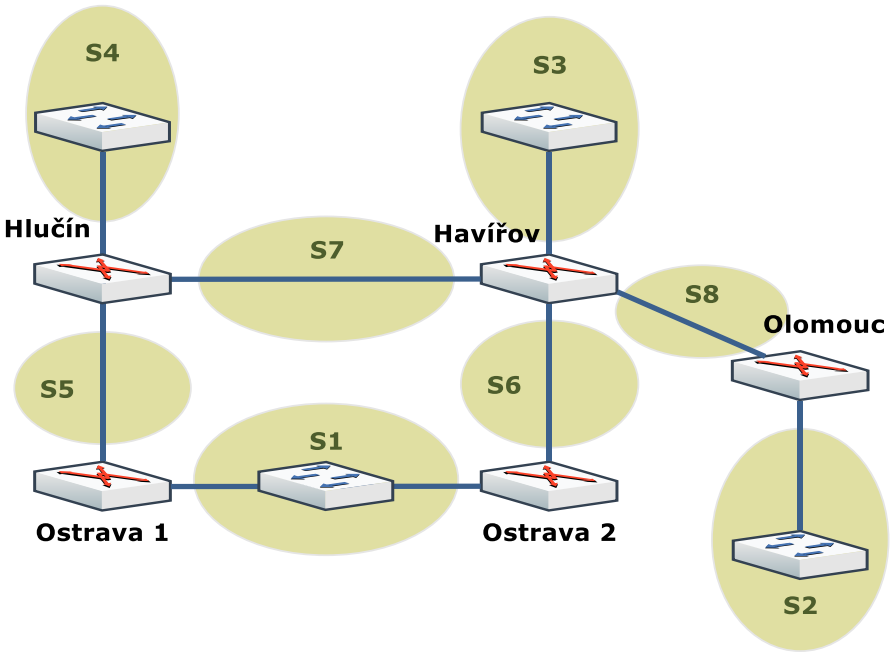

Fig 2.2 Marking individual subnets

The number of bits that will address all stations on a given subnet is determined as the smallest possible number of bits into which the number corresponding to the required number of stations can be encoded. Therefore, it is the nearest higher power of two greater than the required number of stations. Nevertheless, we cannot forget that the router interface to the subnet must also have its IP address, so we need to increase the number of

the required stations by one. In addition, we cannot forget that the bit combination containing all ones is reserved for the broadcast, a combination of all zeros to designate a subnet as such.

Therefore, three bits are not enough for a local network in Hlučín, as it might seem at first glance, but four bits are needed. In Fig 2.3 there are numbers of stations on individual segments adjusted to the nearest power of two including unused reserved addresses. The required number of addresses is also indicated for the connecting (point-to-point) lines. Note that two bits are needed to address stations of the connecting line, providing 4 combinations:

- Address of the first router,
- Address of the second router,
- Designation of the network as such
- Broadcast address

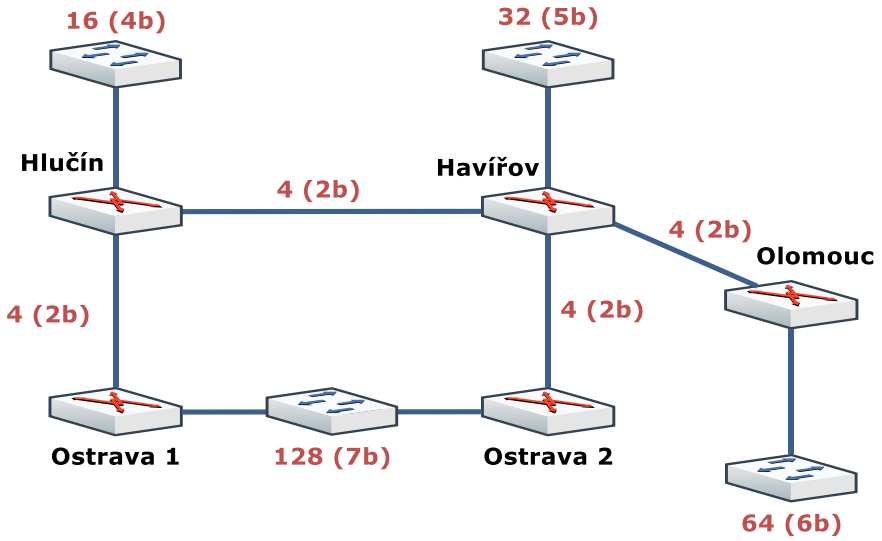

Fig 2.3 Designation of the number of required stations on each segment

The total number of addresses needed for a given network (adjusted for the router interfaces and unusable addresses and adjusted to the power of two in each subnet) is *128+64+32+16+4+4+4+4=256*. Therefore, the Internet provider assigned the prefix of the length of **24** bits (one C-class address) to our network, leaving 8 bits for subnetting. The assigned address is 13.1.20.0/24.

Let us try to address the network using a constant subnet mask. The subnet with the largest number of stations requires 7 bits. There is one bit remaining to determine the subnet, so we can only have two such large subnets. It is obvious that proposing addressing with a constant mask of the subnet will not be possible by means of using a 24-bit address prefix. Note that if the provider assigned multiple addresses (shorter prefix), addressing with a constant network mask would be very inefficient in our network: for example, the

connecting lines that require 4 addresses would be assigned 128 addresses, thus wasting 124 addresses

Therefore, we proceed to addressing with the variable VLSM subnet mask. We will no longer divide the allocated address space into blocks of the same size as with a constant subnet mask. Nevertheless, we will adapt the size of the blocks to the number of stations on individual subnets. The address blocks must not overlap. A unique IP address has to be assigned to each station. However, the number of bits, which unique combination determines the address blocks of individual subnets, will vary according to the block size.

In respect to address allocation of the VLSM we will start from the largest subnet – S1. We need 7 bits for it. There is one bit left for determining the subnet out of the bits usable for subnetting. We decide to assign addresses to the S1 network in the subnet, which subnet prefix is determined by the value1 in the bit 7, thus addresses *213.1.20.128* - *213.1.20.255* (the address 213.1.20.128 is the address of the network itself and 213.1.20.255 is the broadcast address). Therefore, the second half of the allocated range, i.e. the addresses 213.1.20.0 - 213.1.20.127, remains for all other networks.

For the sake of clarity, the prefixes of individual networks and their corresponding address ranges are listed in the Table 3.1. We can use the overview in the Figure 2.4.

Table 2.1 Address Ranges of Individual Subnets

location	name	network address	subnet mask	IP addresses range	broadcast
Ostrava	S1	213.1.20.128/25	255.255.255.128	213.1.20.129 - 213.1.20.254	213.1.20.255
Olomouc	S2	213.1.20.64/26	255.255.255.192	213.1.20.65 - 213.1.20.126	213.1.20.127
Havířov	S3	213.1.20.32/27	255.255.255.224	213.1.20.33 - 213.1.20.62	213.1.20.63
Hlučín	S4	213.1.20.16/28	255.255.255.240	213.1.20.17 - 213.1.20.30	213.1.20.31
link Hlučín-Ostrava1	S5	213.1.20.12/30	255.255.255.252	213.1.20.13 - 213.1.20.14	213.1.20.15
link Ostrava2-Havířov	S6	213.1.20.8/30	255.255.255.252	213.1.20.9 - 213.1.20.10	213.1.20.11
link Hlučín-Havířov	S7	213.1.20.4/30	255.255.255.252	213.1.20.5 - 213.1.20.6	213.1.20.7
link Havířov-Olomouc	S8	213.1.20.0/30	255.255.255.252	213.1.20.1 - 213.1.20.2	213.1.20.3

In the interests of clarification, we can also express the blocks of addresses that have been gradually allocated to individual subnets graphically. Let us imagine the address range of 256 addresses assigned by the provider as a 16 x 16 square. In its upper left corner there is the lowest address range, in the lower right corner there is the highest one, see Figure 2.4.

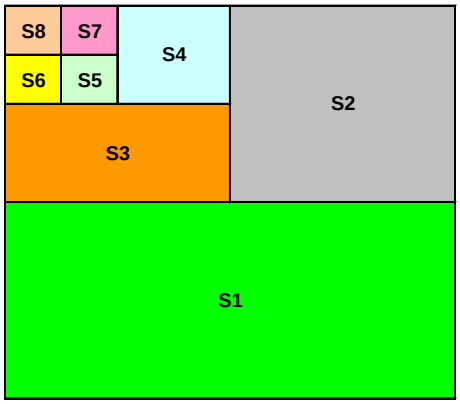

Fig. 2.4 Graphical representation of individual subnets

2.2 Support of Network Elements

Nevertheless, the occurrence of different subnet lengths, i.e. network masks within a single network, requires the implementation of a *Classless Routing Protocol*. It is supported by routing protocol RIPv2, EIGRP, OSPF, ISIS. Routing protocols RIP and IGRP do not support classless routing.

Class C Subnet Table	/24 .0 (00000000) 0 subnets 254 hosts	/25 .128 (10000000) 0 subnet 126 hosts	/26 .192 (11000000) 2 subnets 62 hosts	/27 .224 (11100000) 6 subnets 30 hosts	/28 .240 (11110000) 14 subnets14 hosts	/29 .248 (11111000) 30 subnets 6 hosts	/30 .252 (11111100) 62 subnets 2 hosts
.0	.0	.0	.0	.0	.0	.0	.0 (.1 -.2)
.4						(.1 - .6)	.4 (.5 - .6)
.8					(.1 - .14)	.8	.8 (.9 - .10)
.12				(.1 - .30)		(.9 - .14)	.12 (.13 - .14)
.16					.16	.16	.16 (.17 - .18)
.20						(.17 - .22)	.20 (.21 - .22)
.24					(.17 - .30)	.24	.24 (.25 - .26)
.28			(.1 - .62)			(.25 - .30)	.28 (.29 - .30)
.32				.32	.32	.32	.32 (.33 - .34)
.36						(.33 - .38)	.36 (.37 - .38)
.40					(.33 - .46)	.40	.40 (.41 - .42)
.44				(.33 - .62)		(.41 - .46)	.44 (.45 - .46)
.48					.48	.48	.48 (.49 - .50)
.52						(.49 - .54)	.52 (.53 - .54)
.56					(.49 - .62)	.56	.56 (.57 - .58)
.60		(.1 - .126)				(.57 - .62)	.60 (.61 - .62)
.64			.64	.64	.64	.64	.64 (.65 - .66)
.68						(.65 - .70)	.68 (.69 - .70)
.72					(.65 - .78)	.72	.72 (.73 - .74)
.76				(.65 - .94)		(.73 - .78)	.76 (.77 - .78)
.80					.80	.80	.80 (.81 - .82)
.84						(.81 - .86)	.84 (.85 - .86)
.88					(.81 - .94)	.88	.88 (.89 - .90)
.92			(.65 - .126)			(.89 - .94)	.92 (.93 - .94)
.96				.96	.96	.96	.96 (.97 - .98)
.100						(.97 - .102)	.100 (.101 - .102)
.104					(.97 - .108)	.104	.104 (.105 - .106)
.108				(.97 - .126)		(.105 - .108)	.108 (.107 - .108)
.112					.112	.112	.112 (.113 - .114)
.116						(.113 - .118)	.116 (.117 - .118)
.120					(.113 - .126)	.120	.120 (.121 - .122)
.124						(.121 - .126)	.124 (.125 - .126)
.128	(.1 - .254)	.128	.128	.128	.128	.128	.128 (.129 - .130)
.132						(.129 - .130)	.132 (.133 - .134)
.136					(.129 - .142)	.136	.136 (.137 - .138)
.140				(.129 - .158)		(.137 - .142)	.140 (.141 - .142)
.144					.144	.144	.144 (.145 - .146)
.148						(.145 - .150)	.148 (.149 - .150)
.152					(.145 - .158)	.152	.152 (.153 - .154)
.156			(.129 - .191)			(.153 - .158)	.156 (.157 - .158)
.160				.160	.160	.160	.160 (.161 - .162)
.164						(.161 - .166)	.164 (.165 - .166)
.168					(.161 - .174)	.168	.168 (.169 - .170)
.172				(.161 - .190)		(.169 - .174)	.172 (.173 - .174)
.176					.176	.176	.176 (.177 - .178)
.180						(.177 - .182)	.180 (.181 - .182)
.184					(.177 - .190)	.184	.184 (.185 - .186)
.188						(.185 - .190)	.188 (.189 - .190)
.192		(.129 - .254)	.192	.192	.192	.192	.192 (.193 - .194)
.196						(.193 - .198)	.196 (.197 - .198)
.200					(.193 - .206)	.200	.200 (.201 - .202)
.204				(.193 - .222)		(.201 - .206)	.204 (.205 - .206)
.208					.208	.208	.208 (.209 - .210)
.212						(.209 - .214)	.212 (.213 - .214)
.216					(.209 - .222)	.216	.216 (.217 - .218)
.220			(.191 - .254)			(.217 - .222)	.220 (.221 - .222)
.224				.224	.224	.224	.224 (.225 - .226)
.228						(.225 - .230)	.228 (.229 - .230)
.232					(.225 - .238)	.232	.232 (.233 - .234)
.236				(.225 - .254)		(.233 - .238)	.236 (.237 - .238)
.240					.240	.240	.240 (.241 - .242)
.244						(.241 - .246)	.244 (.244 - .246)
.248					(.241 - 254)	.248	.248 (.249 - .250)
.252						(.249 - .254)	.252 (.253 - .254)

Figure 2.4 Comprehensive Table for the VLSM subnetting

3 Ethernet

Ethernet is the term for the technology summary of the computer networks (LAN, MAN) largely standardized as IEEE 802.3 that use twisted-pair cables, optical cables (coaxial cables in older versions). The Ethernet implements the physical and link layers of the OSI reference model. Currently, Ethernet has been the most successful network technology. [1, 2, 3]

3.1 Ethernet Framework

Ethernet framework is the protocol data unit of the link layer in the Ethernet technology. The frameworks are subsequently used in the link layer to encapsulate packets that are transmitted from the network layer to be transmitted to a particular type of media.
The Ethernet frame has a variable length. The minimum length is 64 bytes, the maximum is 1518 bytes. The frame format is shown in Fig 3.1.

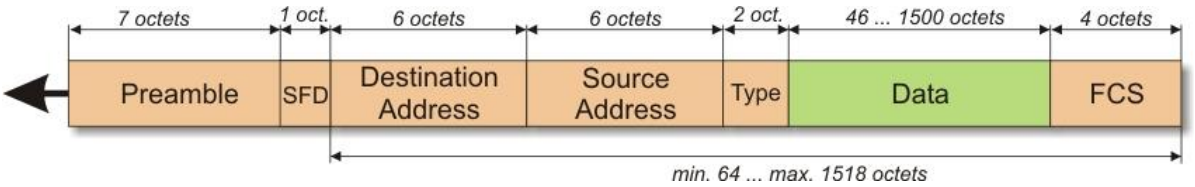

Fig 3.1. Ethernet frame

The meaning of individual fields is following:

- Preamble – initiation of framework, synchronization
- Destination Ethernet Address – address of the recipient (8 bytes)
- Source Ethernet Address – address of the sender (8 bytes)
- Framework type (Length or Type)
- Data - 46 bytes at minimum, 1500 bytes at maximum (maximum length transmitted is referred to as MTU – Maximum Transmission Unit)
- CRC (Cyclic Redundancy Check) – Frame Check Sequence

3.2 Ethernet Types

3.2.1 Fast Ethernet

An obvious change in comparison to the classic Ethernet is the end of support of coaxial cable and thus the bus network – they can no longer be used in the new Ethernet. A star has become a typical topology.
The 100Base-TX specification uses Unshielded Twisted Pair (UTP) by using two pairs. It enables to use Shielded Twisted Pair (STP) of the Type 1 and Type 2 (2 pairs). The maximum segment lengths can be 100m, the carrier frequency is 125 MHz and the data are encoded by using the 4B / 5B method.

3.2.2 Gigabit Ethernet (1GE)

The Gigabit Ethernet is a relatively significant progress in transfer rates. The basis of its physical layer was formed according to the Fibre Channel standard. It uses the CSMA/CD random access method in the MAC sublayer and the framework structure remains the same as with its predecessors.

The Gigabit Ethernet operates at the speed of 1 Gbps. In 1998 the IEEE 802.3z standard was adopted, including two specifications (1000Base-SX, 1000Base-LX) for fibre optic cables and one specification (1000Base-CX) for the metallic cabling (sometimes all of the specifications are referred to as 1000Base-X). Later, the second IEEE 802.3ab standard was adopted that covers the specification (1000Base-T), supporting the transmission through the UTP cables of the category 5.

The minimum frame size of 64B, which is given by the IEEE 802.3 standard, is maintained. This ensures the compatibility of older Ethernets. Although the Gigabit Ethernet uses the same 64 B minimum frame. An overview of individual versions of the Gigabit Ethernet is shown in the Table 3.1.

Table 3.1. Overview of individual versions of the Gigabit Ethernet

Term	Medium	Distance
1000BASE-CX	Twinaxial cable	25 meters
1000BASE-SX	Multi-mode optical fibre	from 220 to 550 meters depending on the fibre width and bandwidth
1000BASE-LX	Multi-mode optical fibre	500 meters
1000BASE-LX	Single-mode optical fibre	5 km
1000BASE-LX10	Single-mode optical fibre using wavelengths of 1.310 nm	10 km
1000BASE-ZX10	Single-mode optical fibre using a wavelength of 1.550nm	~ 70 km
1000BASE-BX10	Single-mode optical fibre in one direction with the wavelength of 1490 nm downstream and 1310 nm upstream	10 km
1000BASE-T	Twisted pair (cat-5, cat-5e, cat-6 or cat-7)	100 meters
1000BASE-TX	Twisted pair (cat-6, cat-7)	100 meters

3.2.3 10 Gigabit Ethernet (10 GE)

Another variant of the Ethernet technology – 10Gigabit Ethernet (10GE) according to the IEEE 802.3ae standard was approved in 2002. The standard IEEE 802.3ae for the 10Gbps Ethernet was designed only for full duplex operation. As a result, there is no restriction on the distance between the nodes, resulting from the transfer method principle. This distance is limited only by the physical properties of the transmission medium and the optical transmission elements. Both physical and logical buses are removed as the standard strictly uses only point-to-point connections.

The 10 Gigabit Ethernet is a widely used specification today. The individual specifications of the IEEE 802.3ae standard are listed in the Table 8.2. Unlike its predecessors it is not only used in local networks for element interconnection, but it is also used in the data centres, storage networks (SAN), metropolitan networks and WAN. It is combined with the standard of the 40 GBase or 100 GBase. In the LAN, it is used to join server clusters where there are 1000 Base-X or 1000 Base-T in a single link. It is also used to connect the servers, switches, or two switches within the data field.

Table 3.2. Overview of the 10 GE versions

Standard	Medium	Range [km]
10GBase-SR/SW (850 nm)	Multi-mode fibre	0,3
10GBase-LR/LW (1350 nm)	Single-mode fibre	10
10GBase-ER/EW (1550 nm)	Single-mode fibre	40
10GBase-LX4 (1550 nm)	Single-mode/multi-mode fibre	10/0,3

3.2.4 40 and 100 Gigabit Ethernet (40 / 100 GE)

The 40/100 GBase Ethernet technology defines the IEEE 802.3ba standard, see tab. 3.3. Likewise, the 10GBase Ethernet uses the full duplex. In contrast with other standards it has two speeds of 40 Gbps and 100 Gbps. The lower speed is used for the data centres interconnection and 100 Gb/s is mainly used for connecting switching elements.

Table 3.3. 40GbE and 100 GbE versions

Standard	Range [km]
40GBase-SR4	100/150 m
40GBase-FR	2 km
40GBase-LR4	10 km
40GBase-ER4	30 (40) km
40GBase-LM4	140/160 m
100GBase-SR10	100/150 m
100GBase-SR4	70/100 m
100GBase-LR4	10 km
100GBase-ER4	30 (40) km

3.2.5 200 and 400 Gigabit Ethernet. Terabit Ethernet

The Terabit Ethernet or the TbE is designation for the Ethernet with higher transfer rates than 100 Gbit /s. In 2017 the 400 Gigabit Ethernet (400G, 400GbE) and 200 Gigabit Ethernet (200G, 200GbE) standards that were developed by the IEEE P802.3bs group using approximately the same technology as the 100 Gigabit Ethernet, were approved. The Ethernet Alliance expects that Ethernet with the transfer rate of 800 Gbit/s and 1.6 Tbit/s will be standardized after 2020.

4 VLAN

The virtual LAN is used to logically divide the network independently of the physical layout. The main principle of VLAN is the segmentation into smaller networks within the physical structure of the original network. By using the VLAN we can create such two networks on one or more interconnected switches. [1, 2, 3]

The VLAN mechanism allows to divide the switch ports into groups – virtual networks, while traffic can only pass between the ports of the same group, see Fig 4.1.

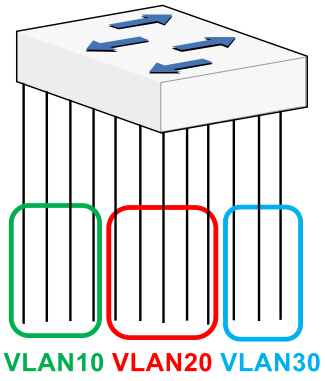

Fig 4.1 VLAN on the switch

For each virtual network, the switch creates and uses a separate switch table. The switch is logically divided into multiple independent logic switches, purely based on the switch configuration, without the need for any physical switching. We gain the possibility to group stations into mutually independent, so-called virtual, networks.

4.1 Standard IEEE 802.1q

In practice, it is necessary to interconnect virtual networks on ports of different switches. However, in this case, the connections between the switches must carry traffic from multiple virtual networks at the same time. That is why it is necessary to identify competence of individual frames to the virtual network, from which they were sent, on the connection between the switches. For this purpose, a special identification in the Ethernet frame header, so-called *tag,* is used. Connections between switches on which frames are marked with a tag are often called the *trunk*, see Fig 4.2.

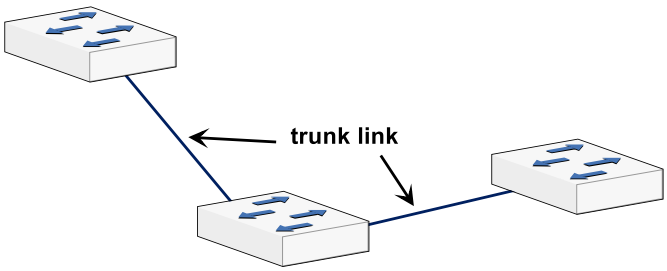

Fig 4.2 Trunk line between switches

The identification of the competent VLAN takes place in the format of the Ethernet frame. A header containing the VLAN number 0-4095 inserted in the *Tag Header* field is used. Apart from the VLAN number 802.1q, the header can also carry the frame priority. The original value of the Ether Type field follows the 802.1q. header, see Fig 4.3.

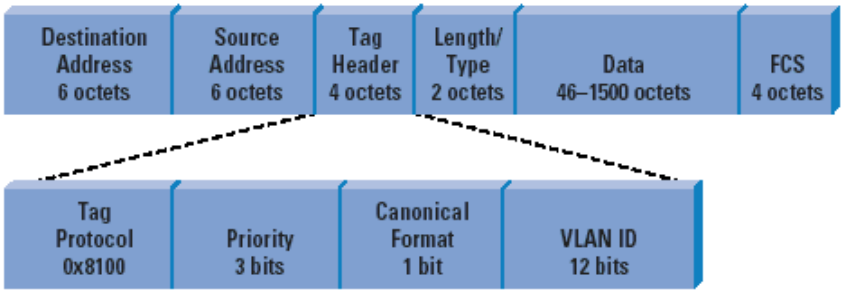

Fig 4.3 The Ethernet frame with 802.1q header with the VLAN identifier

The IEEE 802.1q header extends the frame by 2B. The network card hardware, respectively of the port switches, must be able to handle frames with a maximum length (MTU) of 2B higher than is defined by the classic Ethernet standard.

4.2 Protocol VTP

Most of the time we want the created VLANs to exist across the network on all switches. The VTP Protocol (VLAN Trunking Protocol) is used for the information transmitting about these VLANs. The VTP is the L2 protocol that is used to manage (add, delete, rename) the VLAN within the VTP domain. The VTP domain is formed by one or more network devices that have the same domain name (optionally a password) and are connected via a *trunk* link.

The switch can operate in the regime:

- Server – maintains a list of all VLANs, it has them stored in the NVRAM, it can create and delete the VLAN, receives and sends information about the VLAN via the trunk lines in the VTP domain

- Client – receives configuration from the server, maintains local copy of all the VLANs that cannot be changed and is not stored in the NVRAM, receives and sends the VLAN information
- Transparent – does not participate in the VTP, works independently, can create and delete VLANs, but changes are local, accepts advertisements and in version 2 it can forward them (but does not synchronize or publish their VLANs). It is the only mode where we can create the Extended and Private VLANs. The VTP and VLAN configurations are stored in the NVRAM.

4.2.1 VTP Protocol Trunking

The Ethernet frames that are sent to the broadcast address are sent to all trunk ports. However, there might be a situation when the frames can get to the target switch through the trunk lines even from those VLANs which have no port on the target switch, i.e. the VLAN is not on the target switch, see Fig 4.4.

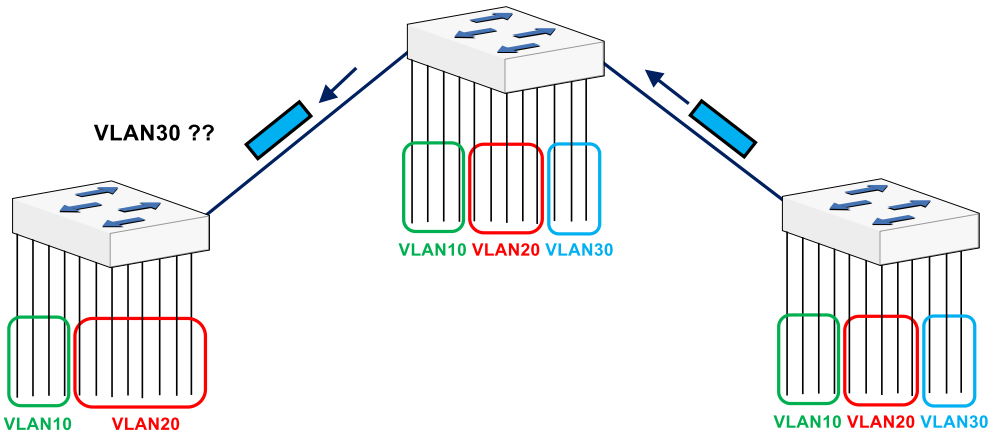

Fig 4.4 The Ethernet frame with VLAN30 sent to all switches

Therefore, the switches of some manufacturers implement the protocol for so-called **trunking** topology of individual VLANs, by means of which individually connected switches inform about the VLAN numbers that they have on each port. This is usually the Cisco VTP protocol, see Fig 4.5.

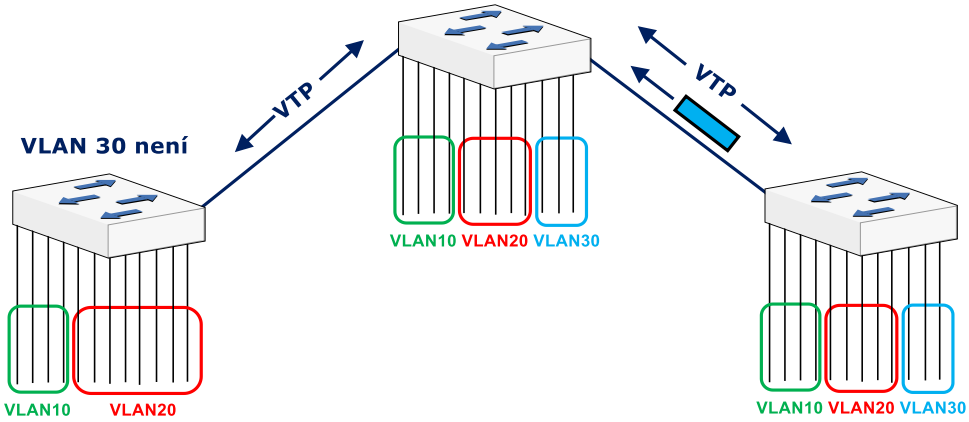

Fig 4.5 Protocol Application ATP

In case of a combination of manufacturers of equipment, and thus the inability to use a single trunking topology protocol, the static configurations are used to filter individual VLANs, i.e. the given switch is set to choose which VLAN to pass and which not on the appropriate trunk port

4.3 Routing between the VLANs

In some cases, it is useful to provide routing between two VLANs. We can do this in two ways:

- We will use the physical port of the router, see **Fig 4.6** for each VLAN
- or we enable routing by means of one physical port of the router, so-called R*outer on the stick*, Fig 4.7.

In the second case, one physical port of the router must have so much of the *subinterface* as the VLAN needs to route. We must not forget to set the encapsulation type on the L2 layer of the 802 1q and assign a separate IP address to each *subinterface*.

The advantage of this second approach is to save the number of router interfaces when the number of required router interfaces does not increase with the number of the VLANs between which the router is to route. The disadvantage is the lower throughput because one physical line is used for routing a packet between the VLANs.

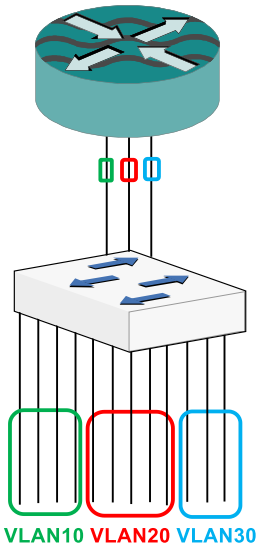

Fig 4.6 Routing between the VLANs using different router ports

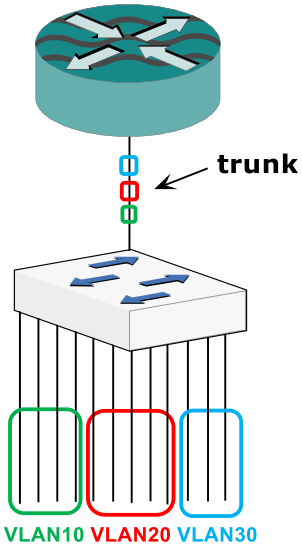

Fig 4.7 Routing between VLANs by means of the *Router on the stick*

4.4 Spanning Tree Protocol

Spanning Tree Protocol (STP) is an algorithm to create a logical network topology **without loops** in a switched network **with loops**. Let us consider the network in Fig 4.8. This network consists of two switches where data can be routed from the segment A to the segment B via two paths (via switch S1 or via switch S2). In this topology, this creates a **loop** in which, for example, the *broadcast* frames would circulate infinitely.

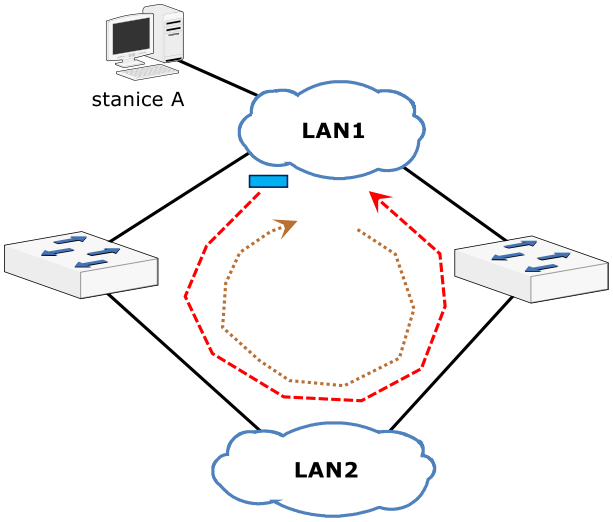

Fig 4.8 Frame circulation in the loop topology

In addition, the presence of a loop leads not only to circulation or generation of frame copies, but also completely breaks down the principle of automatic switch table of the switch. The circulating frames with the source address of the station **A** arrive to the switches from the switch LAN1 at one time and from the switch LAN2 at a different time. In this way, the switches continuously and in half of the cases incorrectly update information according to the port of the incoming frame, behind which port the station A is actually connected to.

It is clear from the above case that looping in the network topology is **undesirable.** However, it is desirable to have redundant links in case of connection backup. Therefore, we must have a mechanism that block some switch ports in the loop-topology so that the resulting topology is tree-like, for example, see **Fig 4.9**.

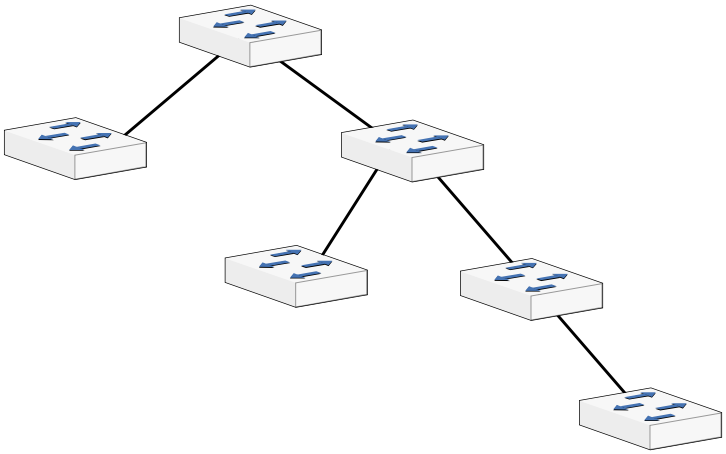

Fig. 4.9 Tree-like topology without created loops

This mechanism is provided by the **Spanning Tree** IEEE 802.1d Protocol. The task of this protocol is to **maintain constantly the tree structure** on a given switch topology. Important features of the Spanning Tree protocol are:

- the line blocking always takes place only from one side of the switch, see Fig 4.10.,
- the algorithm works constantly, in case of a line or switch port failure, the tree is recalculated automatically (one of the previously blocked ports is unblocked).

Creating a tree is done in two steps that, however, take place continuously and simultaneously:

1. choice of tree root, so-called *Root Bridge*
2. Creation of a tree of preferred (the lowest *cost*) routes from each switch to the *Root Bridge.*

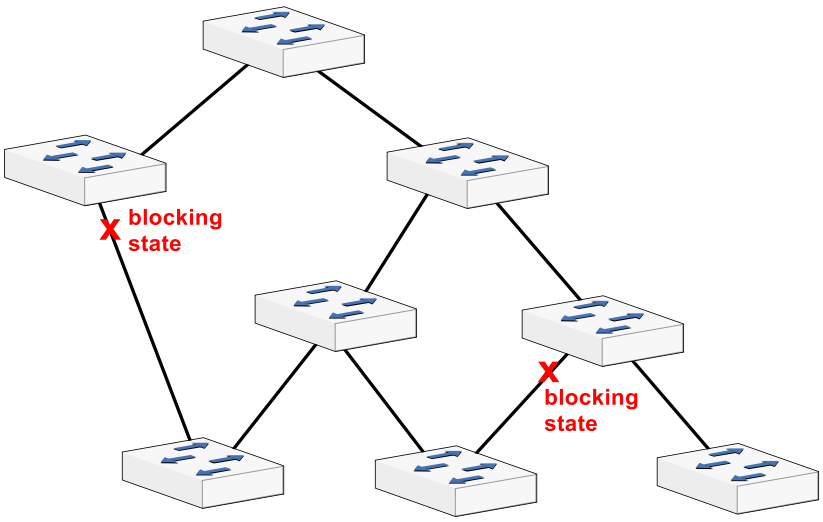

Fig 4.10 Port blocking when using Spanning Tree

The redundant connections that are not part of the shortest path of the tree are blocked. Data that arrive on the blocked connections are discarded. This creates a logical topology without loops. The **Spanning Tree** Protocol requires communication between devices to detect loops. The links that create loops are put into a *Blocking State*.

The switches are sending messages *Bridge Protocol Data Units* (BPDU) for obtaining information on logical topology without loops. The blocked ports receive the BPDUs and this ensures that when the active path or device fails, a new tree is calculated.

The BPDU contain information that enables the switches to:

- Choose one switch that will act as the *Root Bridge*,
- Calculate the shortest route to the *Root Bridge*,
- Choose one of your ports as the *Root Port* for each switch that is not the Root Port. The *Root Port* is the port with the best route to the Root Bridge.
- Set up ports that are the part of the spanning tree, so-called Designated Ports.

In order to prevent loops during the convergence (the *Root Bridge* option), the algorithm defines transient port states, the so-called *Learning* and *Listening*. The port in the state of *Listening* does not forward frames, but only monitors the surrounding traffic for 15

seconds to decide whether to switch to *Forwarding* (common operation) or *Blocking*. Moreover, before transition into the *Forwarding* state, the MAC address learning of the neighbouring stations takes place only for 15 seconds in the *Learning* state without forwarding frames. This limits the percentage of frames whose receiver is not in the switch Table and must therefore be sent to all ports.

The individual port phases are shown in Fig 4.11.

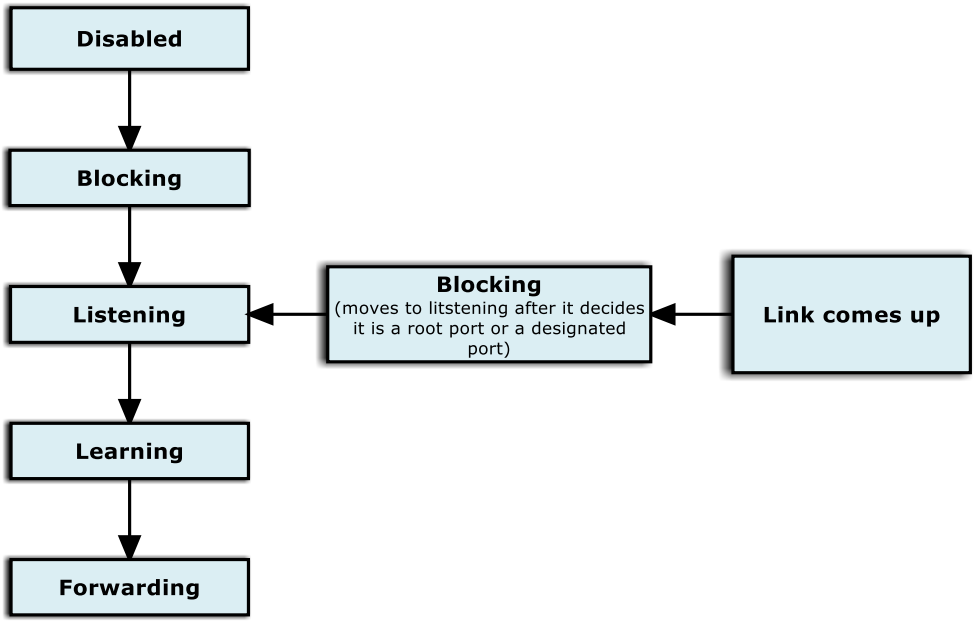

Fig 4.11 Port phases of the switch of the Spanning Tree Protocol

- **Blocking state** – ports only accept the BPDUs. Data are discarded and the switch does not learn the MAC addresses on this port. It lasts in this state for a maximum of 20 seconds.
- **Listening state** – the switch detects if there is another path to the root bridge. Paths that do not have the lowest *cost path* to the *root bridge* will return to the blocking state. The listening time is called the *forward delay* and lasts for 15 seconds. Data are not forwarded and the MAC addresses are not taught, but the BPDUs are still being processed.
- **Learning state** – the MAC address learning takes place, but data are still not forwarded. The BPDUs are still being processed. It takes 15 seconds.
- **Forwarding state** – Data are forwarded, the MAC address learning takes place and the BPDUs are being processed.
- The port may also be in the *Disabled State*. The port is in this state if the administrator turned it off or it has broken down.

5 Packet filters

A packet filter, or firewall, is a network device that serves for managing and securing network operation by blocking or allowing established communications based on predefined or dynamic rules and policies. Simply, its serves as a checkpoint that defines the rules for communication between the networks which it separates. Historically, these rules have always included the identification of the source and data destination (source and destination IP address) and the source and destination port. Today ‘s modern firewalls also check the status of the connection. They can intelligently filter the operation based on protocols of the application layer with capabilities similar to IDS (Intrusion Detection System).
Installing a firewall will increase the overall security of the IT infrastructure of the company or the household. If the firewall is well set up then it is the only entry point through which all the communication must pass. [1, 2, 3]

5.1 Basic principle of filtration

The simples and oldest form of the firewall function lies in the fact that the rules specify exactly from which address and port to which address and port a passing packet can be delivered. Therefore, the rules apply only to communication related to the network and transport layers of the OSI reference model.
The advantage of this solution is the high processing speed. That is why they are still used today in places where the accuracy or more thorough analysis of passing data is not required, just the basic functions.
Typical packet filter representatives include ACL (Access Control Lists) on Cisco routers, see Fig.5.1 or the tool iptables in Linux (Fig.5.2).

```
SWITCH(config)#access-list 5 deny host 10.5.1.10
SWITCH(config)#access-list 5 permit 10.5.1.10 0.0.0.255
SWITCH(config)#access-list 5 deny any
```

Fig. 5.1 Example ACL

```
iptables -A INPUT -p tcp --dport 80 -j DROP
iptables -A INPUT -p tcp --dport 90 -j DROP
iptables -A INPUT -p tcp --dport 100 -j DROP
```

Fig. 5.2 Example IP tables

5.2 State packet filters

State packet filters perform the control as well as simple packet filters. Moreover, they store the information about ongoing TCP protocol connections. This information can be used for decision making whether the packets belong to an already allowed connection and can be released or if they must go through the decision-making process.

Most of the stateless packet filters simply allow all ports above 1024 through the firewall because these ports are used for return sockets from the internal network. Nevertheless, this security is insufficient. For instance, nothing prevents the Trojans from waiting on any port higher than 1024 in the internal network. Stateless packet filters cannot prevent this type of penetration.

On the other hand, status packet filters do not let any services through Firewall, except for services for which they have permission set, and except for connections that they already have in their status tables.

Therefore, the most demanding control is performed at the time of the connection establishment during the TCP handshake process. After this process, all the packets (for a given relation) are being processed faster because it is easy to determine whether they belong to an existing session or not. When the connection is closed, the session is deleted from the table. The datagram can be found in the following states:

- NEW – datagram opens new communication,
- ESTABLISHED, RELATED – datagram belongs to an already established connection.
- INVALID – datagram does not belong to any connection or is unidentifiable.

The process handshake client is started by sending the SYN =1 flag in the packet header. Each packet that has SYN=1 is considered as an initiation packet for the new connection by the firewall. If the service requested by the client is available on the server, the server responds by sending a packet with the set command SYN=1 and ACK =1. Then the client responds with the packet in which only the bit ACK=1 is set and the connection is thus marked as ESTABLISHED. All the outgoing packets will pass through such a firewall, but those incoming packets will only pass if they are the part of the ESTABLISHED connection. This security prevents hackers from making unwanted connections. To prevent the constant table filling, there is an integrated system that after a certain time, when the connection is ‘silent’, deletes this connection from the table. Therefore, many applications send so-called ‘keepalive’ messages during the ‘silence’ so that the firewall does not choose to end the connection. It is worth mentioning that the most common attack on the Internet of the Denial type is the SYN-flood. This is an attack when the attacker sends many packets with the SYN flag to the target computer, but no longer responds. This results in an overflow of the status table and, among other things, a slowdown of the server, but also a system crash and the computer must then be restarted locally.

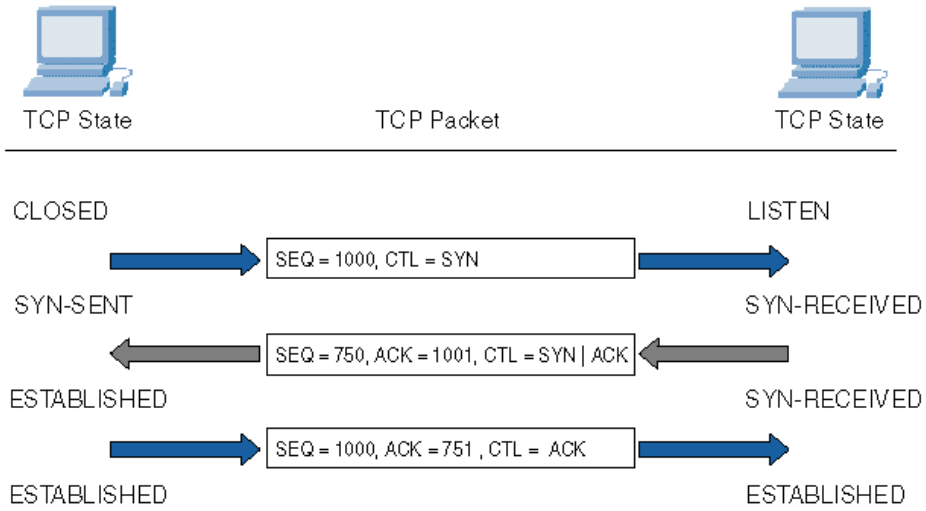

Fig. 5.3 Functionality of the statefull firewall

The biggest advantages of the status packet filters are their high speed, level of security and, in comparison with the above-mentioned application gateways and simple packet filters, many times easier configuration – and thanks to the simplification of configuration also lower probability of incorrect setting of rules by the operator.

6 NAT

The NAT (Network Address Translation) is the way of adjusting network traffic passing through the router by transcribing (so-called translation) of the source or destination IP address, or protocol headers of the higher layer, for instance, the port number for the TCP, UDP, etc. in the area of computer networks. The NAT can be implemented in software on the regular computer (e.g. in the Linux kernel using iptables /netfilter) or it can be implemented directly in the router‘s firmware or hardware. By means of translating the network addresses (NAT) the private IP addresses are translated in the private network into unique public IP addresses that can be used on the Internet. The routers have a table that consists of internal sockets assigned to external sockets for the translation of the network addresses.

The figure 6.1 shows IPTABLES scheme. The origin of the packet determines which chain it traverses initially. There are five predefined chains (mapping to the five available Netfilter hooks), though a table may not have all chains. Predefined chains have a policy, for example DROP, which is applied to the packet if it reaches the end of the chain. The system administrator can create as many other chains as desired. These chains have no policy; if a packet reaches the end of the chain it is returned to the chain which called it. A chain may be empty.

- PREROUTING - Packets will enter this chain before a routing decision is made.
- INPUT - Packet is going to be locally delivered. It does not have anything to do with processes having an opened socket; local delivery is controlled by the "local-delivery" routing table: ip route show table local.
- FORWARD - All packets that have been routed and were not for local delivery will traverse this chain.
- OUTPUT - Packets sent from the machine itself will be visiting this chain.
- POSTROUTING - Routing decision has been made. Packets enter this chain just before handing them off to the hardware.

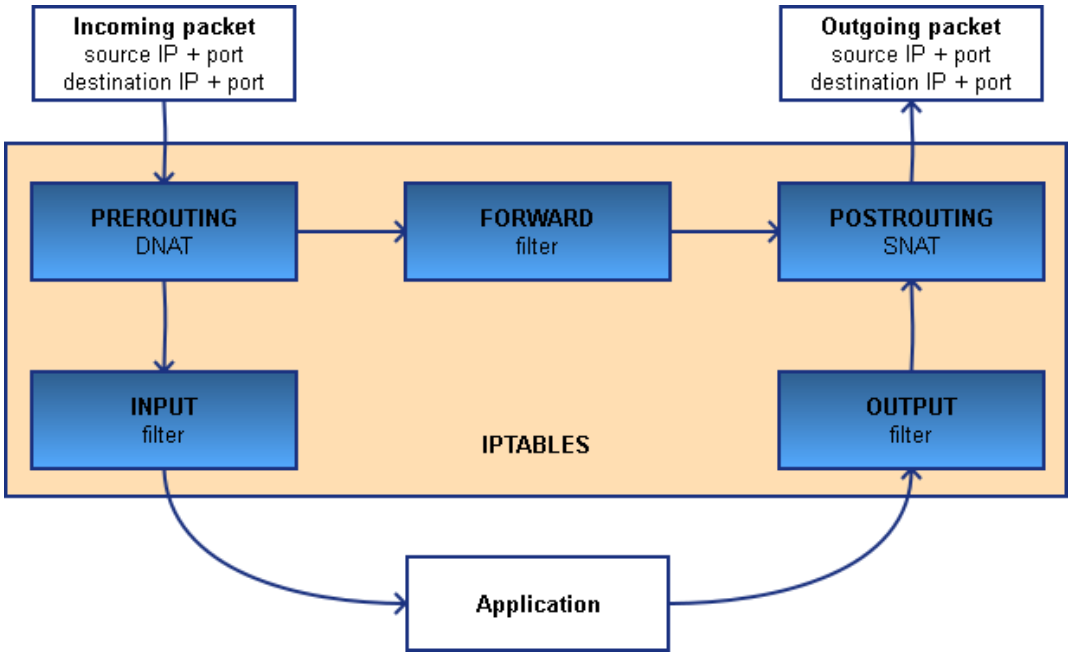

Fig. 6.1 IPTABLES scheme

The Figure 6.2 shows an example of the NAT functionality. Both the computers are connected to the Internet via a router where the address translation is performed. We distinguish between the DNAT, SNAT and the special case of SNAT - Masquerade. The SNAT or Source NAT is a technique in which the source IP address is changed and this happens after the process of postrouting. If we have a PC with the IP address 192.168.1.1 in the private network, then, after the packet passes through the router with the SNAT function, it changes to the address 12.1.1.1. Masquerade is a special and, at the same time, probably the most used case of the SNAT where there can be more devices in the private network. The individual IP addresses of these devices are translated into one external IP address. [4]

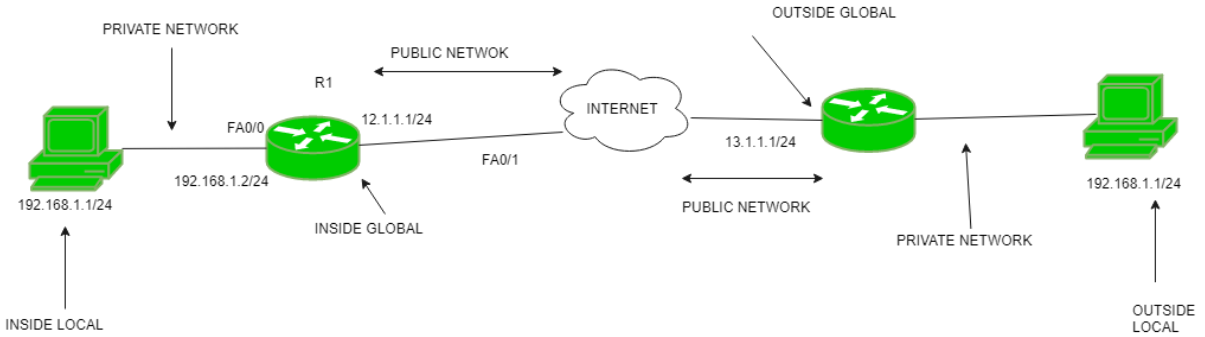

Fig. 6.2 NAT principle

On the other hand, the DNAT or Destination NAT is a technique in which the destination IP address of a device is being changed. For example, the port forwarding can be implemented. The DNAT is performed before prerouting.

References:

[1] TANENBAUM, Andrew S. *Computer networks*. 4th ed. New Jersey: Prentice-Hall, c2003. ISBN 9780130661029.

[2] ODOM, Wendell. *CCNA 200-301 Official Cert Guide Library*. 1. Cisco Press, 2020. ISBN 9780136755449.

[3] PYLES, James, Jeffrey L. CARRELL. *Guide to TCP/IP: IPv6 and IPv4*. Cengage Learning, 2016. ISBN 9781337020541.

[4] PETERSEN, Richard. *Ubuntu 20.04 LTS Server: Administration and Reference*. Surfing Turtle Press, 2020. ISBN 9781949857139.