Local Area Networks (LAN)

-

Introduction to Switching, IPv4 and Routing

Fundamentos de Redes

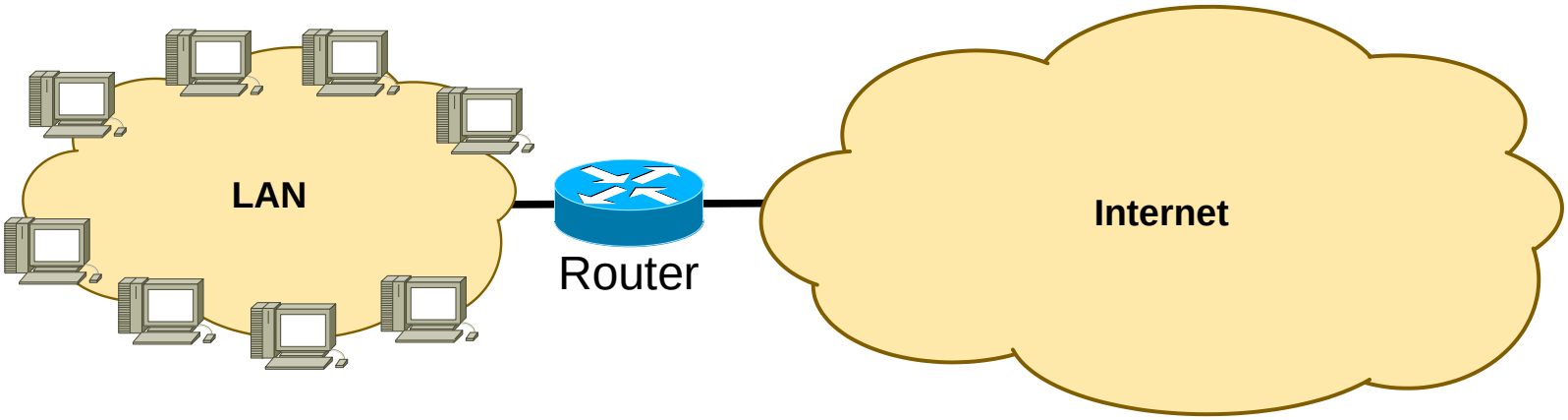
Mestrado Integrado em
Engenharia de Computadores e Telemática
DETI-UA

universidade de aveiro
deti.ua.pt

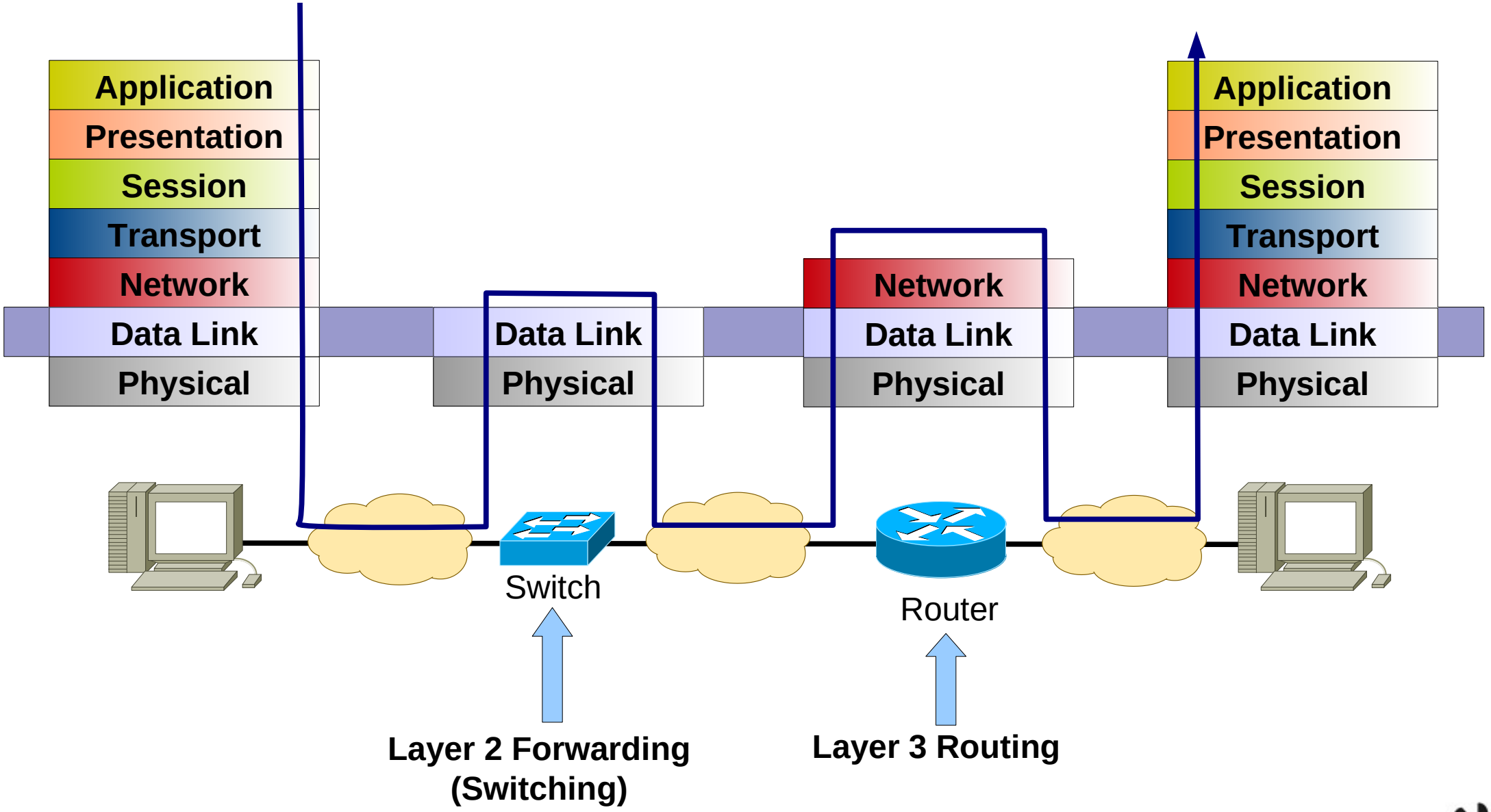
Local Area Network (LAN)

- Is a computer network within a small geographical area.
  - Home, school, room, office building or group of buildings.
- Is composed of inter-connected hosts capable of accessing and sharing data, network resources and Internet access.
  - Host refers generically to a PC, server, or any other terminal.
- Technologies
  - Current: Ethernet, 802.11 (Wi-Fi)

    Legacy: Token Ring, FDDI, …

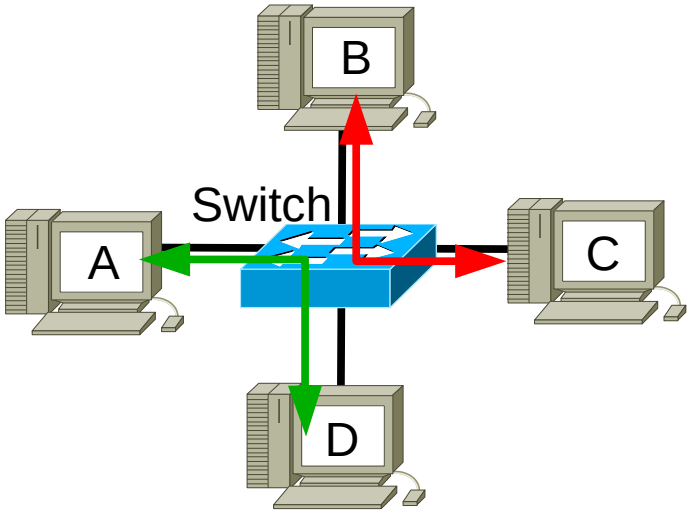


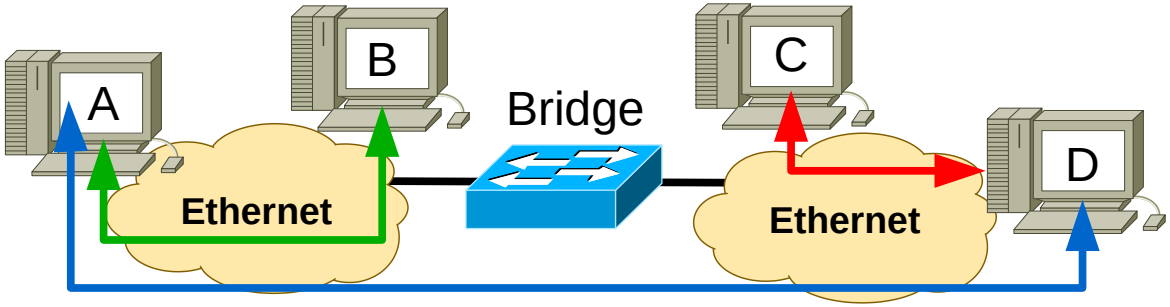
Local Area Network (LAN)



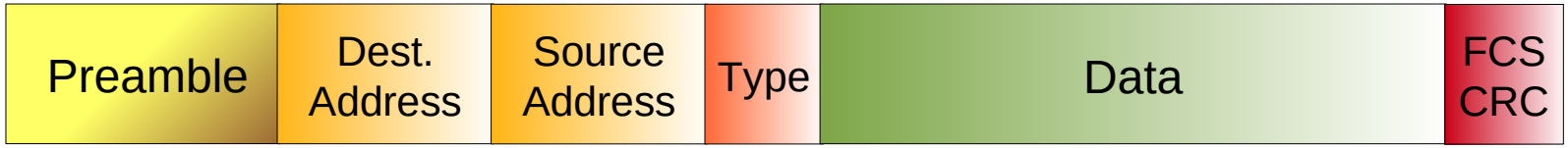
Switching

- With Switches/Bridges
  - Interconnection done at OSI Layer 2.
  - Hosts can transmit simultaneously.
  - A network of Switches is a **Broadcast Domain**
    - An Ethernet frame with destination FF:FF:FF:FF:FF:FF (Broadcast) will reach all connected switches and hosts.





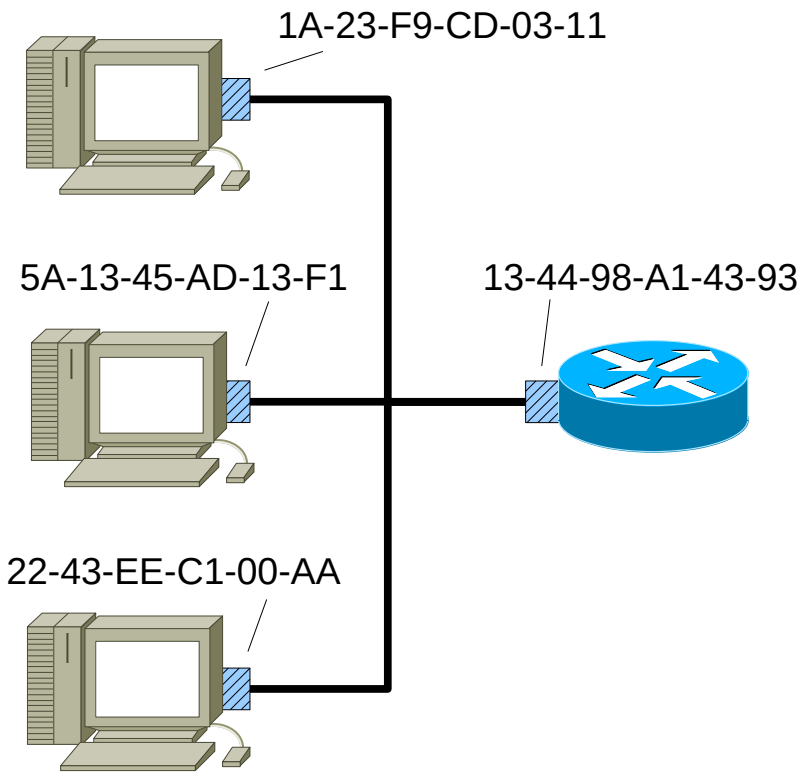
Ethernet Frame



- The sender's network card encapsulates an IP datagrama (or any other network protocol) in an Ethernet frame.
- Preamble:
  - 7 bytes with pattern 10101010 followed by one byte with pattern10101011.
  - Used to sincronize the sending and receiving clocks.
- Destination and Source addresses: 6 bytes Physical (MAC) address
  - If the network card receives a frame with destination equal to its own address or its the broadcast address, it will pass data to the network level process.
  - If not, drops the frame.
- Type defines which protocol is encapsulated in the frame (usually IPv4 or IPV6).
- The frame check sequence (FCS) is a four-octet cyclic redundancy check (CRC) that allows detection of corrupted data within the entire frame as received on the receiver side.

MAC Addresses

- MAC (Physical, Ethernet or LAN) Address:
  - Function: Allow the exchange of data between network interfaces connected using a Layer 2 network.
  - Have 6 bytes/48 bits.
  - Are unique.
  - Each network card has its own address.
  - Defined by manufacturer
    - Some hardware allows change.
    - First 24-, 28-, or 36-bits assign to manufacturer.
  - Hexadecimal notation
    - Broadcast: FF-FF-FF-FF-FF-FF



Ethernet Frame Minimum Size

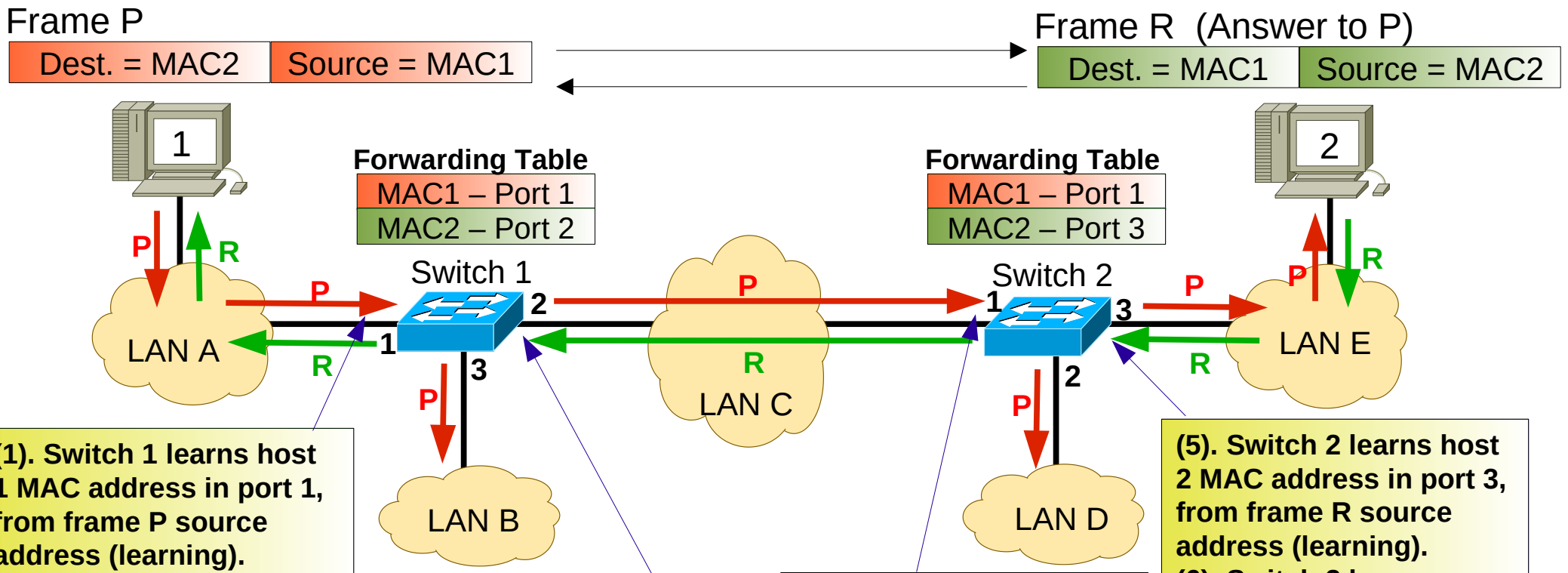
- Historically there were Ethernet technologies that allowed collisions and a collision detection mechanism had to be present (CSMA/CD).
- Depending on the technology and maximum cable size, the Ethernet frame had to be big enough to allow the collision detection mechanism to detect a frame being transmitted before the last frame byte leaving the source host.
- By legacy (it is possible to merge different Ethernet technologies) the **minimum frame size is 64 bytes**.
- If the frame's header plus data do not reach 64 bytes, a set of zeros must be added to the end of the frame to reach 64 bytes.
  - This is called **padding**.

Switches Basic Operations

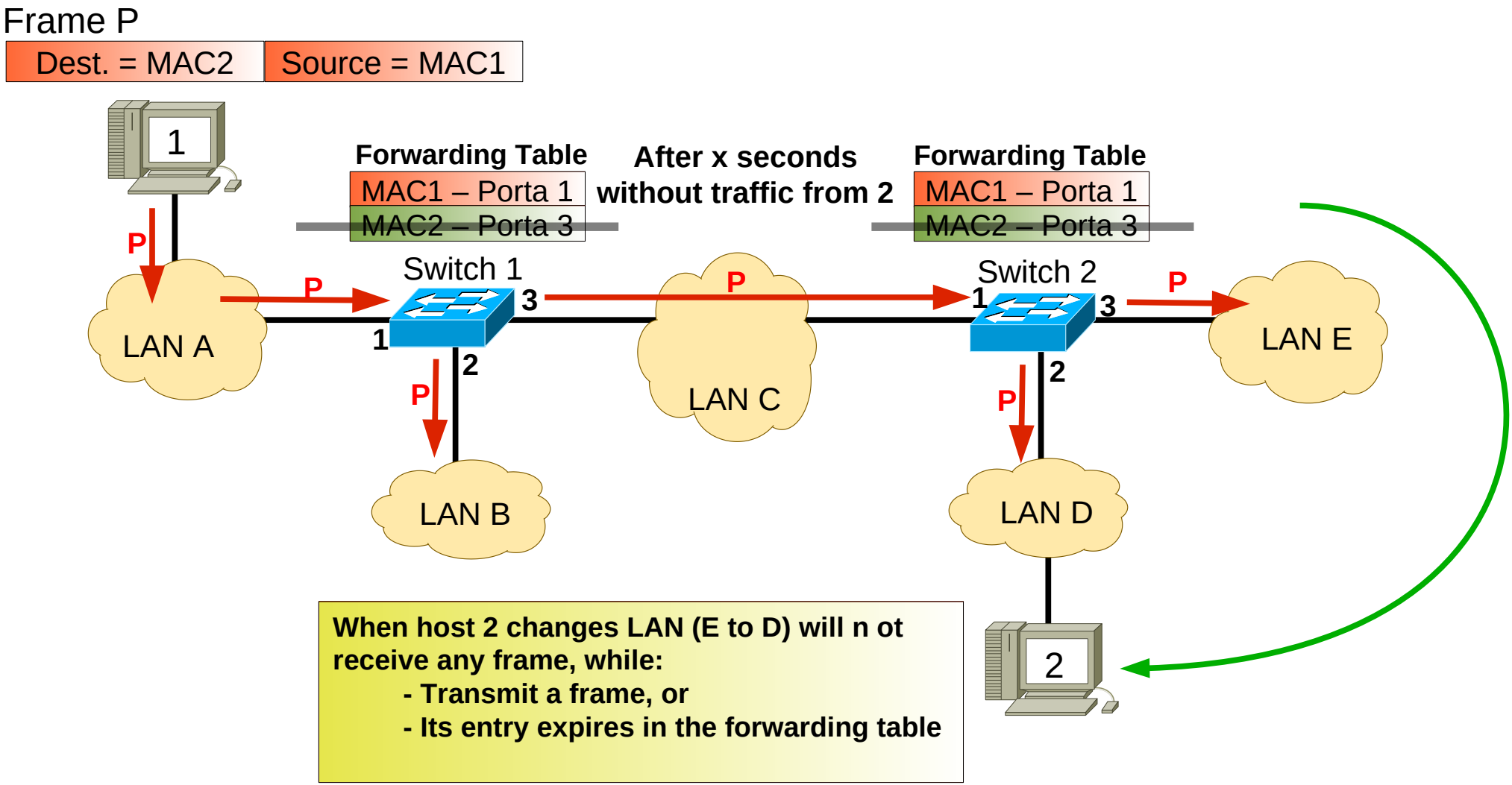
MAC	Porta
00:11:11:11:11:11	1
00:22:22:22:22:22	1
A1:33:33:33:33:33	2
44:44:44:44:44:44	3
55:55:55:00:00:55	3

- Switches have a **Forwarding Table**.
- When a switch receives an Ethernet frame:
  - Registers an entry at the Forwarding Table the frame's source MAC address and the port where the frame was received.
    - If no frames are received from that MAC address after some time (**aging time**) the entry is removed.
  - Searches the Forwarding Table for the frame's destination MAC address and forwards the packet according:
    - **Forwarding** mechanism:
      - If the frame's destination MAC address exists in the table, the switches forwards the frame through the port associated with that MAC address.
    - **Flooding** mechanism:
      - If the frame's destination MAC address DOES NOT exist in the table, the switches forwards the frame through all active ports (except the one where it was received).
        - » Note: Just within the same VLAN (more details later).

Learning, Flooding and Forwarding

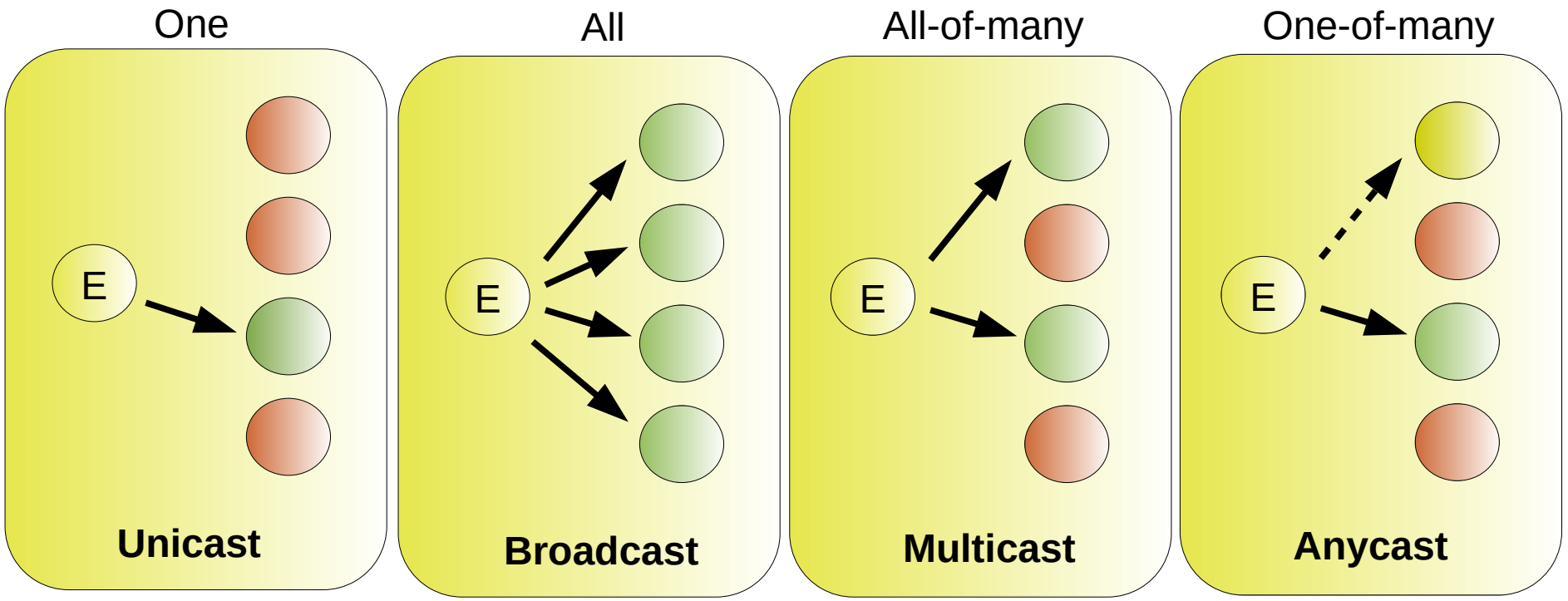


Forwarding Table Aging Time



Types of Addresses

- Unicast – Identify a single sender/receiver.
- Broadcast – All are receivers.
- Multicast – Identify all elements of a group as receivers (all-of-many)
- Anycast – Identifies any element of group as receiver (one-of-many)



IPv4 Addressing

- An IPv4 address is a unique address for a network interface
- Exceptions:
  - Dynamically assigned IPv4 addresses (DHCP)
  - IP addresses in private networks (NAT)

- An IPv4 address:
  - is a **32 bit long** identifier
  - encodes a network number (**network prefix**)
    and a **host identifier**

Network Prefix and Host Identifier

- The network prefix identifies a network and the host identifier identifies a specific host (actually, interface on the network).

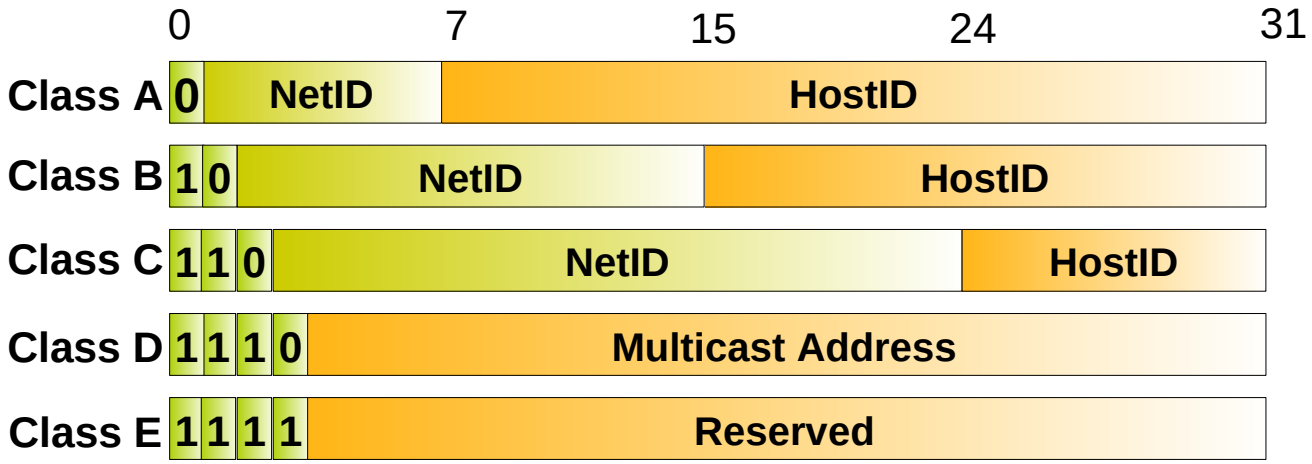
- How do we know how long the network prefix is?
  - **Before 1993:** The boundary between network prefix and host identifier is implicitly defined (**class-based/classful addressing**)

  or

  - **After 1993:** The boundary between network prefix and host identifier is indicated by a **netmask.**

IPv4 Classful Addressing

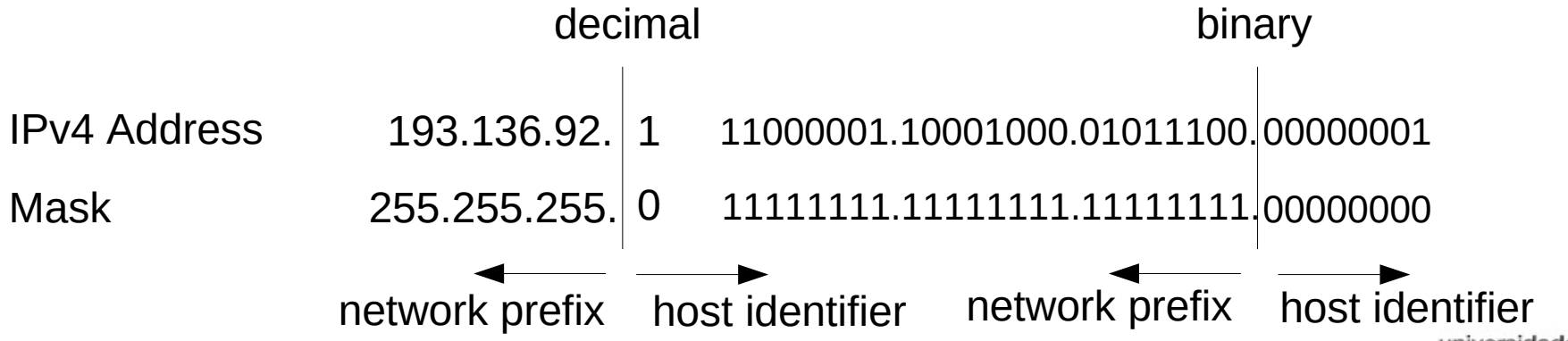
- Initially (until 1993) the boundary between the network prefix and host identifier was predefined by the value of the first byte (class).
- Resulted in a huge waste of addresses:
  - Classes A and B were to big,
  - Not enough class C networks.
- Routing Tables were becoming very long
  - It was not possible to merge (aggregate) networks to simplify routing tables.



Class	First Address	Last Address
A	1.0.0.0	126.0.0.0
B	128.0.0.0	191.255.0.0
C	192.0.0.0	223.255.255.0
D	224.0.0.0	239.255.255.255
E	240.0.0.0	255.255.255.254

Classless Inter-Domain Routing (CIDR)

- New interpretation of the IP addressing to increase efficiency and flexibility.
  - Network Masks were created to define the boundary between the IP network prefix and host identifier.
  - A bit of the mask equal to one indicate that that bit (in that position) of the address belongs to the network prefix.
    - A bit of the mask equal to zero indicate that that bit (in that position) of the address belongs to the host identifier.
  - Called VLSM (Variable Length Subnet Mask).
  - Must be provided with the IP address.
- Allowed the partition of a network in smaller networks or sub-networks (subnets).
- Allowed to merge several network under a single prefix (aggregation or summary process).



Mask Notations

- There are two notations for IPv4 masks:
  - Decimal: 4 bytes separated by dots.
  - CIDR: A slash (/) a a number with the number of bits of the network prefix.
- Both notations still exist today.
  - CIDR starts to become prevalent.
  - IPv6 only supports CIDR.

CIDR	Decimal
/21	255.255.248.0
/20	255.255.240.0
/19	255.255.224.0
/18	255.255.192.0
/17	255.255.128.0
/16	255.255.0.0
/15	255.248.0.0
/14	255.240.0.0
/13	255.224.0.0

CIDR	Decimal
/30	255.255.255.252
/29	255.255.255.248
/28	255.255.255.240
/27	255.255.255.224
/26	255.255.255.192
/25	255.255.255.128
/24	255.255.255.0
/23	255.255.254.0
/22	255.255.252.0

IPv4 Packet Format (1)

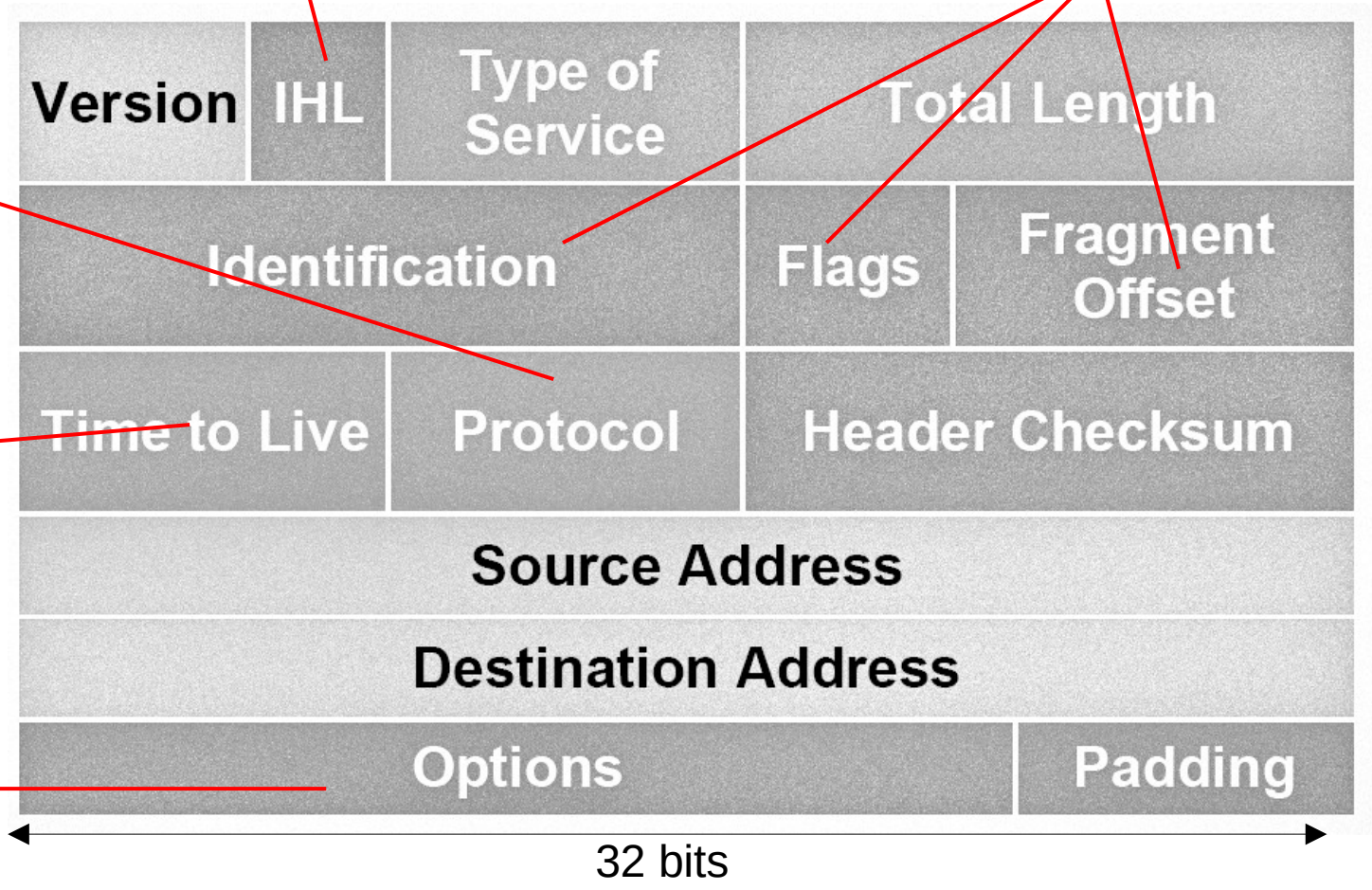
IP Header Length (bytes)

For use in fragmentation and reconstruction

Specifies the underlying protocol (ICMP (1), TCP (6), UDP (17), ...)

Maximum number of hops until the destination (decremented by one in each router on the path)

E.g.,Timestamps, Path record, Routing defined at the source (size must be a multiple of 32, zero padding may be needed)



IPv4 Packet Format (2)

- Version (4 bits) – Protocol version
- Header Length (4 bits) – Header size (number of blocks of of 4 bytes)
  - Without options, the header uses 5 blocks of 4 bytes (20 bytes) and the first byte of the header is 0x45 (version 4, 5 blocks of 4 bytes).
- Type od Service (1 byte) – To implement QoS
  - By default is 0x00.
- Total Length (2 bytes) – packet size in bytes including the header.
  - Maximum IPv4 packet size is 65 535 bytes.
  - Usually this value is limited by the local network Maximum Transport Unit (MTU).

IPv4 Packet Format (3)

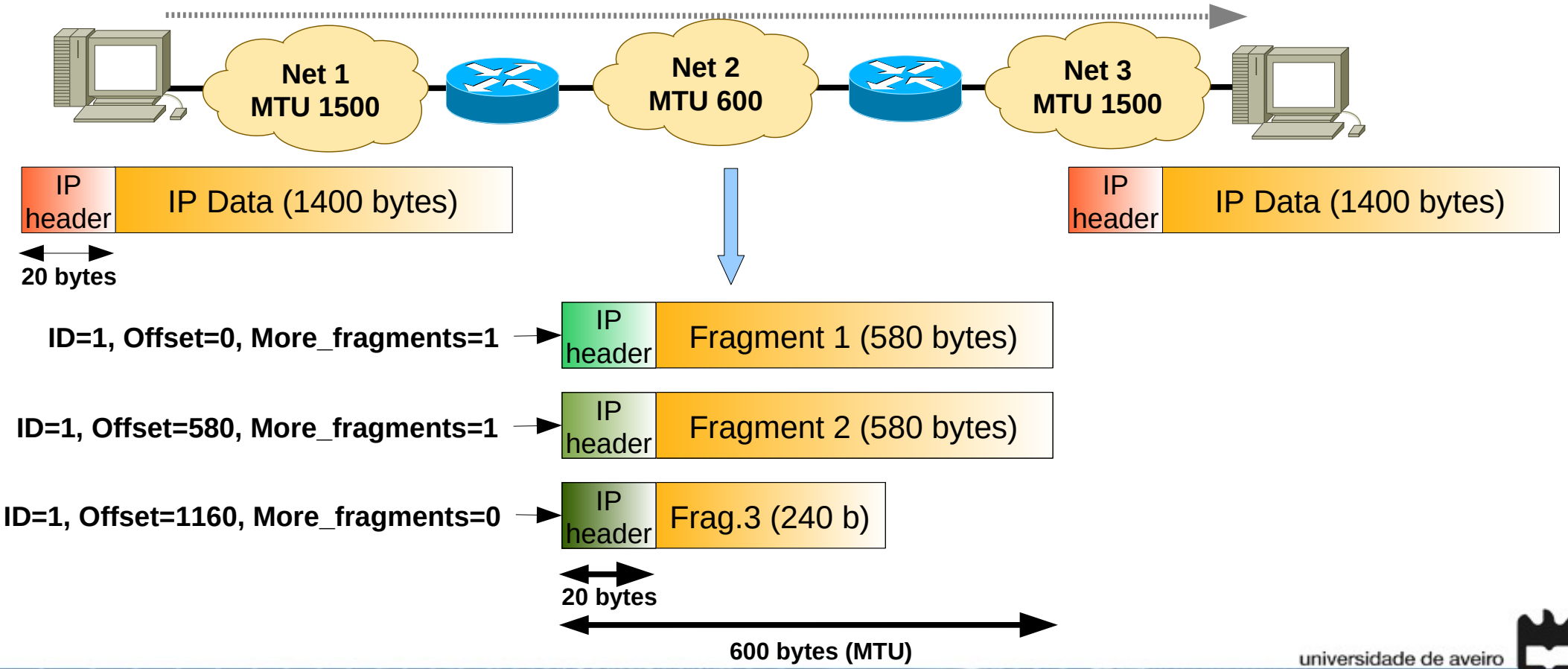
- Time to Live (1 byte) – maximum hops until destination
  - Each router on path reduces TTL by 1.
  - If TTL reaches 0 the packet is discarded and router may notify sender.
- Protocol (1 byte) – specifies the encapsulated protocol
- Header Checksum (2 bytes) – for header error detection
  - Each router on path must recalculate checksum.
    - Changes at least TTL.

IPv4 Packet Format (4)

- Identification (2 bytes) – identifies fragments of the same original IPv4 packet.
- Flags (3 bits)
  - First bit for future use (always 0).
  - Second bit is 0 if packet can be fragment, and 1 otherwise (do not fragment).
  - The third bit is 0 for the last fragment, and 1 otherwise (more fragments flag).
- Fragment Offset (13 bits) – position (in multiples of 8 bytes) of a fragment in the original IPv4 packet (for first fragment is 0x00).

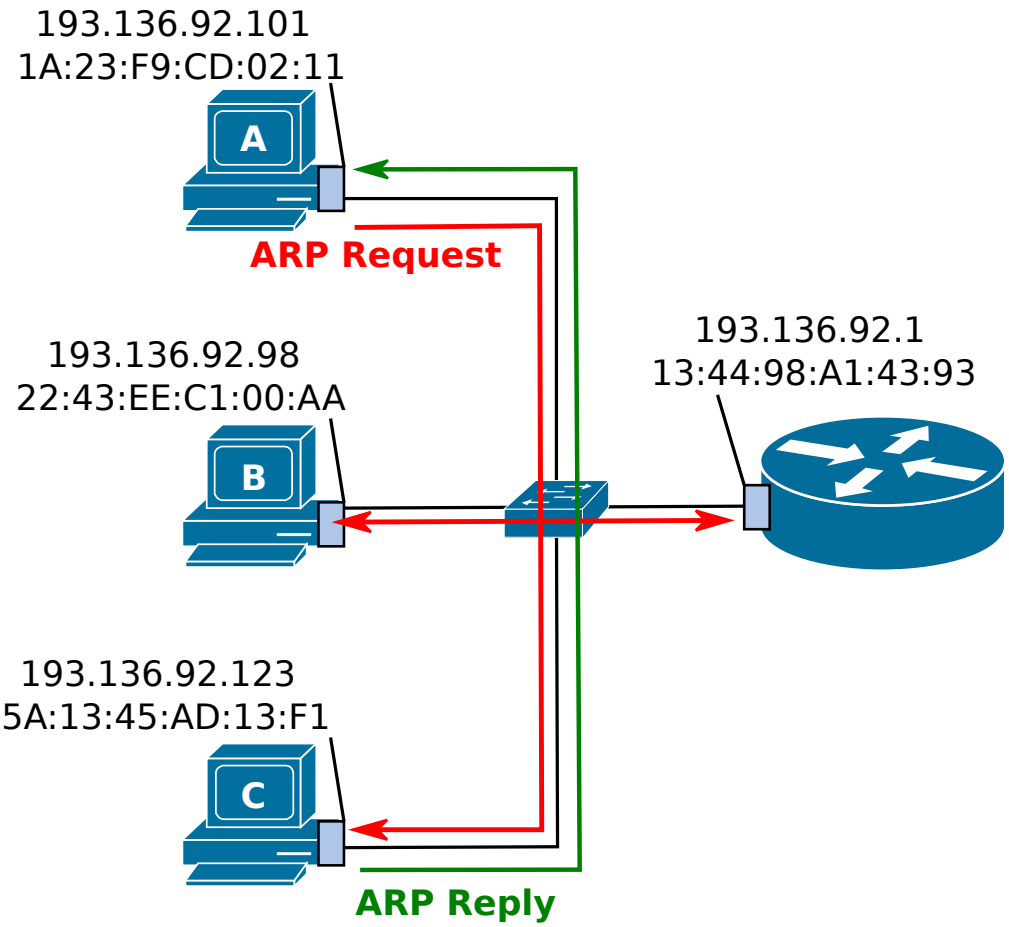
IPv4 Fragmentation and Reconstruction

- Each network defines the maximum packet that can be sent.
  - MTU - Maximum Transfer Unit
- For larger packets, the packet must be fragmented at entry and reconstructed after.
- Header fields used on the process:
  - Identification, fragment offset, flags: do not fragment e more fragments

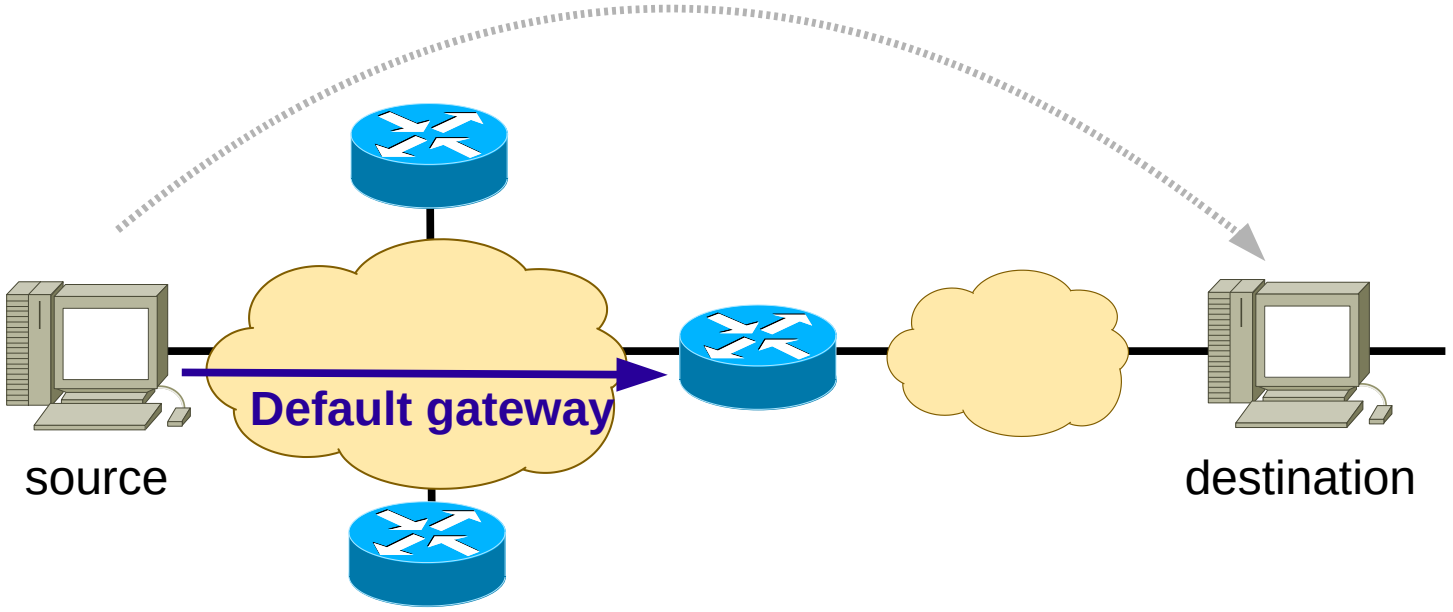


Address Resolution Protocol (ARP)

- IPv4: Address Resolution Protocol (ARP)
- Example:
  - When “A” wants to contact “C” by IPv4:
    - “A” requires “C” MAC address.
    - Only knows IPv4 address.
    - If “C” IPv4 address is not present in the ARP table, then:
      - “A” send an “ARP Request” in broadcast to the local network (destination MAC: FF:FF:FF:FF:FF:FF) with the IPv4 address of “C”,
      - All machines receive this packet,
      - “C” verifies that is IPv4 addess is on the the “ARP request”, responds directly to “A” with a “ARP reply” (destination MAC==MAC of “A”) with it’s on MAC address.
- MAC address resolution only happens in a the local network.
  - ARP packets to not pass through routers.



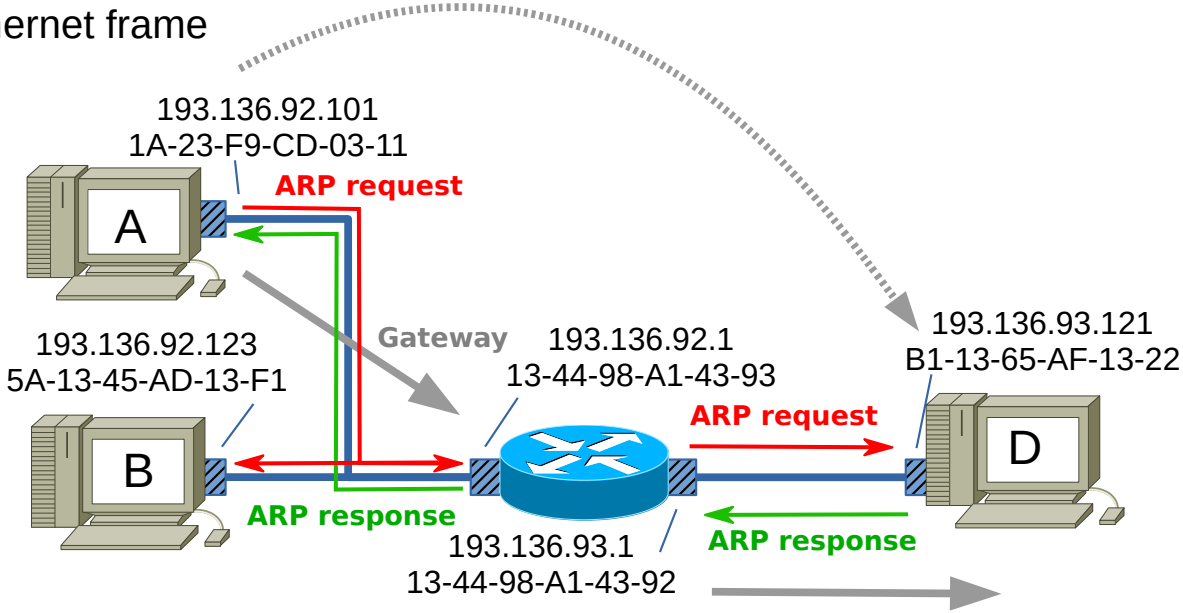
Routing to Another IP Network (1)



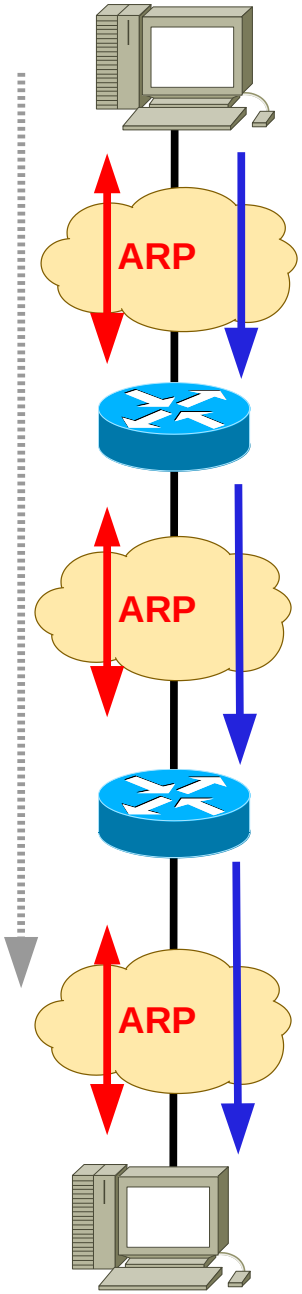
- When a host must send an IP packet to another IP, the packet must be sent to the **default gateway.**
- The **default gateway** must be provided at the same time than the IP address.
  - Manually or by self configuration.

Routing to Another IP Network (2)

- Sending an IP packet from host “A” to host “D”
  - “A” constructs the IP packet with the IPv4 address of “A” as source, and the IPv4 address of “D” as destination
  - “A” verifies that the address of “D” belongs to a different IPv4 network, “A” will send the packet to the configured gateway (router)
  - “A” determines the MAC address of the gateway (ARP)
  - “A” constructs Ethernet frame with the MAC address of “A” as source and the MAC address of the gateway as destination
  - “A” encapsulates the IP packet within the Ethernet frame
  - “A” send the Ethernet Frame
  - The router (GW) receives the Ethernet frame
  - The router removes the IP packet from the Ethernet frame, and verifies that the destination is “D”
  - The router determines the MAC address of “D” (ARP)
  - The router constructs a new Ethernet frame with the MAC address of the output interface as source and the MAC address of “D”as destination
  - The router encapsulates the received IP packet (changing just the TTL) within the Ethernet frame
  - The router sends the Ethernet Frame



Routing over Multiple IP networks



Datagrama IP

Ethernet header | Datagrama IP

MAC destino = MAC gateway

Datagrama IP

Ethernet header | Datagrama IP

MAC destino = MAC next hop

Datagrama IP

Ethernet header | Datagrama IP

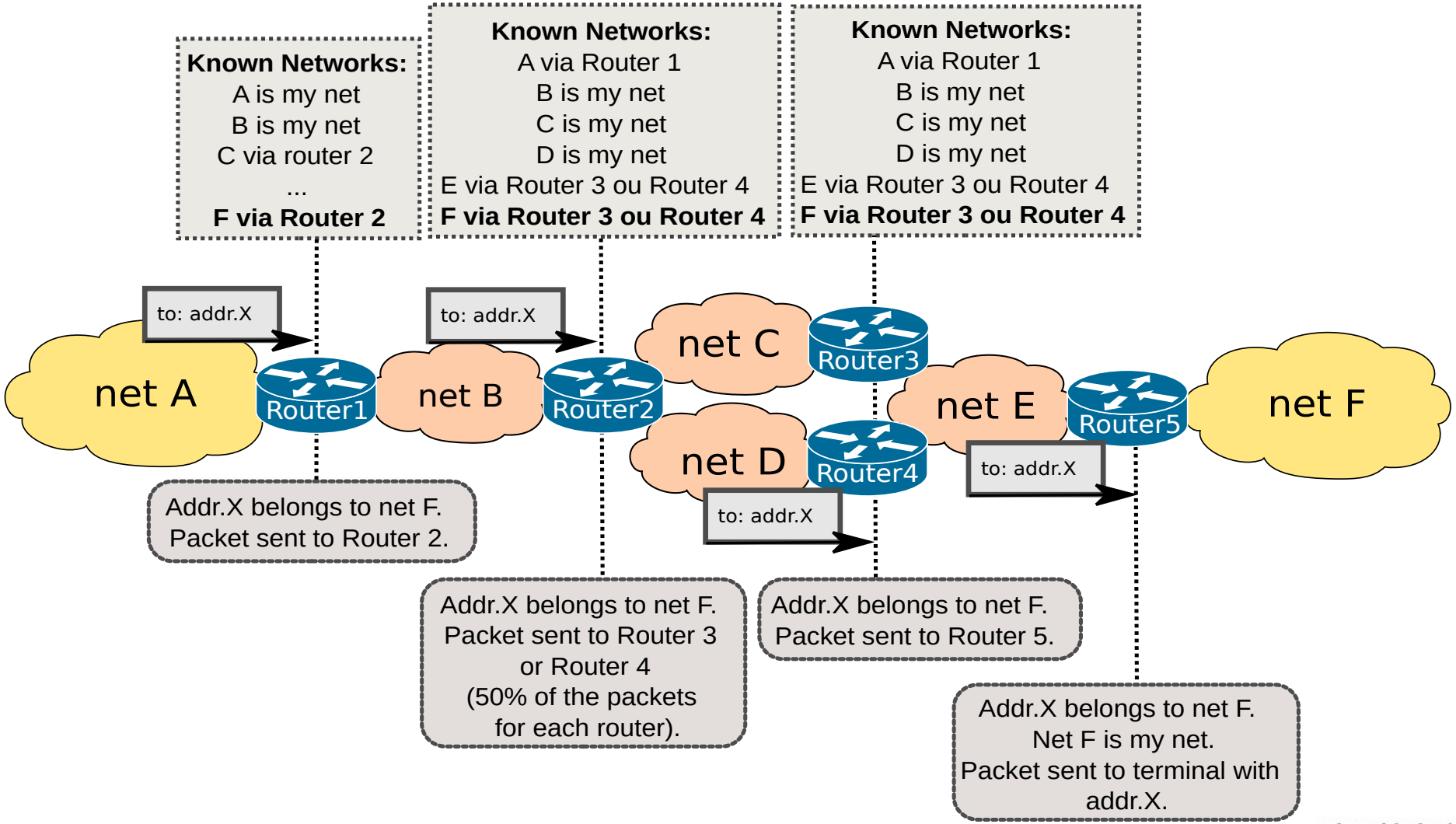
MAC destino = MAC PC destino

IP Routing Overview (1)

- Routers forward packets toward destination networks.
- Routers must be aware of destination networks to be able to forward packets to them.
- A router knows about the networks directly attached to its interfaces
- For networks not directly connected to one of its interfaces, however, the router must rely on outside information.
- A router can be made aware of remote networks by:
  - **Static routing**: An administrator manually configure the information.
  - **Dynamic routing**: Learns from other routers.
  - **Routing policies**: Manually routing rules that outweigh static/dynamic routing.

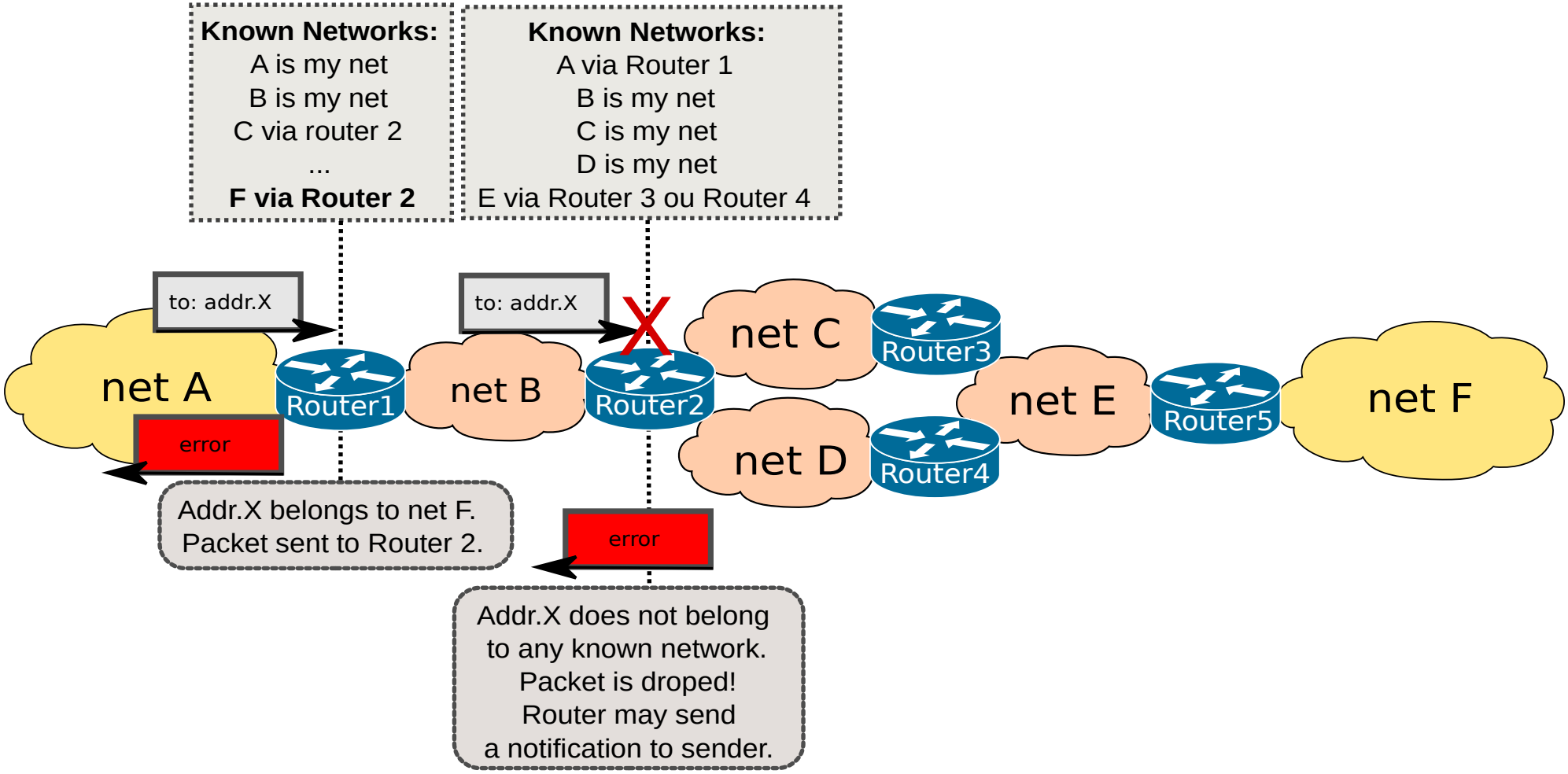
IP Routing Overview (2)

- Hop-by-hop decision:
  - Based on the packets' **IP Destination Address**.
  - Rules listed on the **IP Routing Table.**



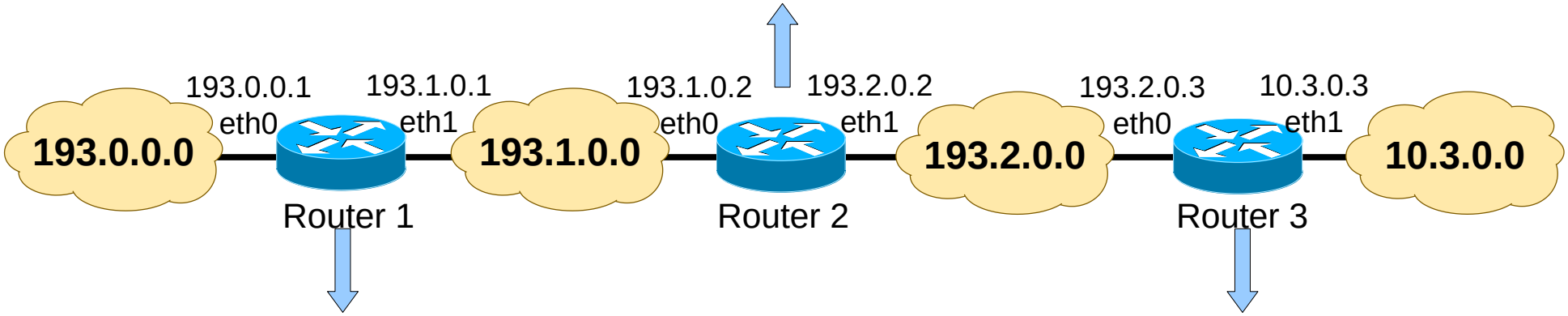
IP Routing Overview (3)

- Hop-by-hop decision:
  - If a packet for an unknown network reaches the router this will drop the packet, and MAY notify the sender about the routing error.



IP Routing Tables (1)

Dest.	Mask	Next hop	Interface	Metric
193.0.0.0	/24	193.1.01	eth0	1
193.1.0.0	/24	direct	eth0	-
193.2.0.0	/24	direct	eth1	-
10.3.0.0	/8	193.2.0.3	eth1	1



Dest.	Mask	Next hop	Interface	Metric
193.0.0.0	/24	direct	eth0	-
193.1.0.0	/24	direct	eth1	-
193.2.0.0	/24	193.1.0.2	eth1	1
10.3.0.0	/8	193.1.0.2	eth1	2

Dest.	Mask	Next hop	Interface	Metric
193.0.0.0	/24	193.2.0.2	eth0	2
193.1.0.0	/24	193.2.0.2	eth0	1
193.2.0.0	/24	direct	eth0	-
10.3.0.0	/8	direct	eth1	-

IP Routing Tables (2)

- Define how a remote network is reachable:
  - Next-hop (identified by its address), and
  - Local interface that provides connection.
- A network may be reachable using more than one path: (next-hop,local interface) pair.
- Mandatory elements
  - **Destination prefix**
  - **Destination mask**
  - **Metric**
    - Could be defined by key tags.
      - e.g., Directly Connected
  - One or both
    - **Next-hop address**
    - **Output interface**
- Optional elements
  - **Administrative distance**
  - **Protocol**
  - **Entry age (last time information received)**
  - **Scope**
  - **Flags**
  - **Source-specific**

Cisco IOS

```
R     200.1.1.0/24 [120/1] via 200.19.14.10, 00:00:16, FastEthernet0/1
      200.19.14.0/24 is variably subnetted, 2 subnets, 2 masks
C        200.19.14.0/24 is directly connected, FastEthernet0/1
L        200.19.14.4/32 is directly connected, FastEthernet0/1
R     200.38.0.0/24 [120/1] via 200.43.0.8, 00:00:03, FastEthernet1/1
      200.43.0.0/24 is variably subnetted, 2 subnets, 2 masks
C        200.43.0.0/24 is directly connected, FastEthernet1/1
L        200.43.0.1/32 is directly connected, FastEthernet1/1
```

Linux: route -n

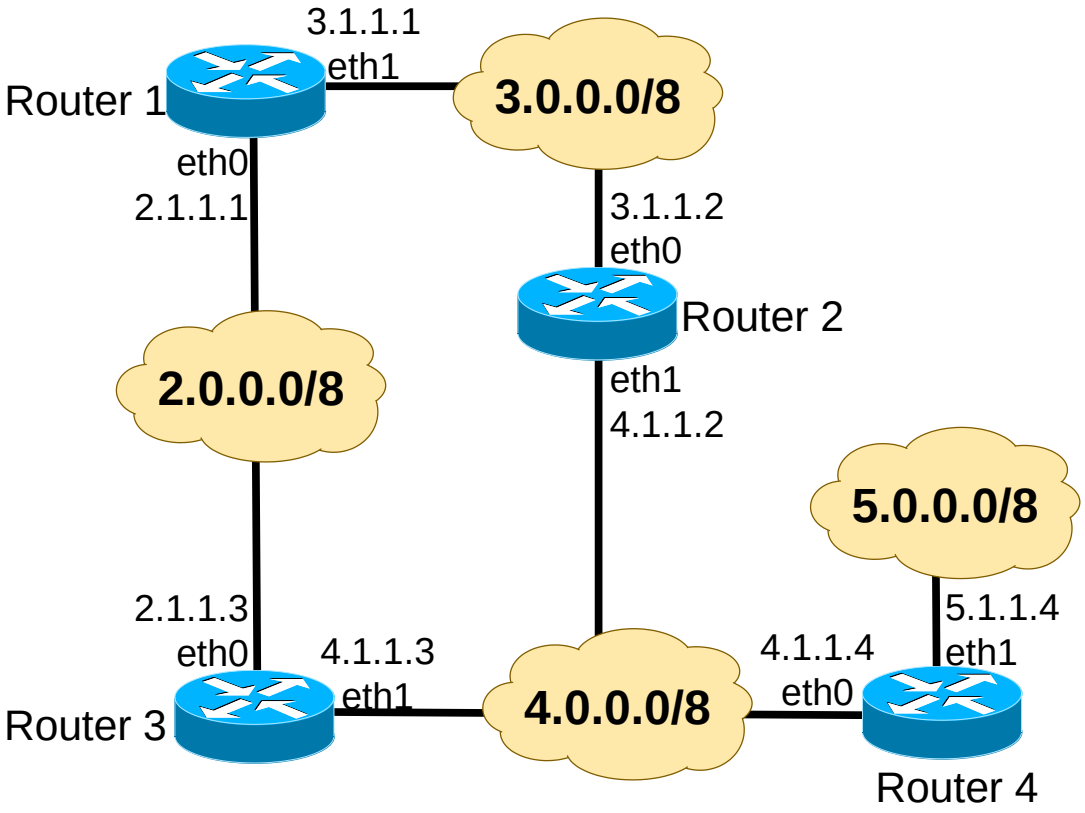
```
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
0.0.0.0         193.136.92.1    0.0.0.0         UG    100    0        0 enp5s0f1
169.254.0.0     0.0.0.0         255.255.0.0     U     1000   0        0 enp5s0f1
193.136.92.0    0.0.0.0         255.255.254.0   U     100    0        0 enp5s0f1
```

Linux: ip route

```
default via 193.136.92.1 dev enp5s0f1  proto static  metric 100
169.254.0.0/16 dev enp5s0f1  scope link  metric 1000
193.136.92.0/23 dev enp5s0f1  proto kernel  scope link  src 193.136.93.104  metric 100
```

- The next path hop (next hop address) may be found using more than one table entry (recursive resolution).
  - e.g., Network A is reachable through address from network B, Network B is reachable through address from network C, …
- The next-hop address may be obtained from external information (configurations or other mechanisms).
  - e.g., Tunnels, Point-to-point connections, etc…
- When an entry uses a next-hop address from an unknown network, that entry is removed.
- All entries obtain by dynamic methods may have an entry age (time since last update/confirmation).
  - After a timeout value without an update/confirmation the entry is removed.

IP Routing Example



C	2.0.0.0/8 is directly connected, Ethernet0
C	3.0.0.0/8 is directly connected, Ethernet1
R	4.0.0.0/8 [120/1] via 3.1.1.2, 00:00:16, Ethernet1
	[120/1] via 2.1.1.3, 00:00:12, Ethernet0
R	5.0.0.0/8 [120/2] via 3.1.1.2, 00:00:13, Ethernet1
	[120/2] via 2.1.1.3, 00:00:02, Ethernet0

Router 1

R	2.0.0.0/8 [120/1] via 4.1.1.3, 00:00:26, Ethernet1
	[120/1] via 3.1.1.1, 00:00:02, Ethernet0
C	3.0.0.0/8 is directly connected, Ethernet0
C	4.0.0.0/8 is directly connected, Ethernet1
R	5.0.0.0/8 [120/1] via 4.1.1.4, 00:00:23, Ethernet1

Router 2

C	2.0.0.0/8 is directly connected, Ethernet0
R	3.0.0.0/8 [120/1] via 4.1.1.2, 00:00:06, Ethernet1
	[120/1] via 2.1.1.1, 00:00:05, Ethernet0
C	4.0.0.0/8 is directly connected, Ethernet1
R	5.0.0.0/8 [120/1] via 4.1.1.4, 00:00:20, Ethernet1

Router 3

R	2.0.0.0/8 [120/1] via 4.1.1.3, 00:00:13, Ethernet0
R	3.0.0.0/8 [120/1] via 4.1.1.2, 00:00:08, Ethernet0
C	4.0.0.0/8 is directly connected, Ethernet0
C	5.0.0.0/8 is directly connected, Ethernet1

Router 4

Layer 2
Ethernet and Wi-Fi (802.11)

Fundamentos de Redes

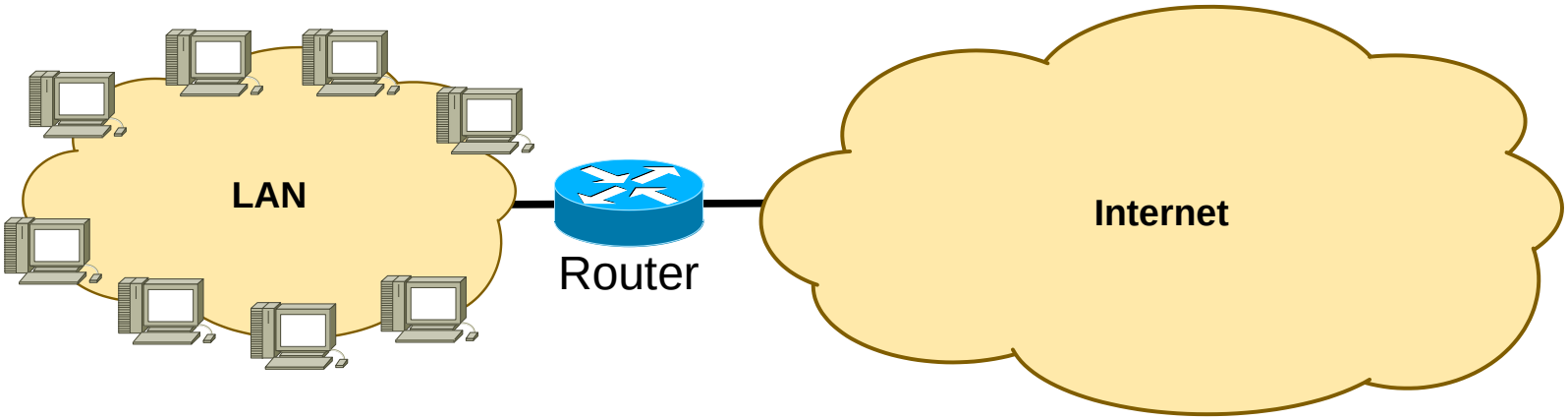
Mestrado Integrado em
Engenharia de Computadores e Telemática
DETI-UA

universidade de aveiro
deti.ua.pt

Local Area Network (LAN)

- Is a computer network within a small geographical area.
  - Home, school, room, office building or group of buildings.
- Is composed of inter-connected hosts capable of accessing and sharing data, network resources and Internet access.
  - Host refers generically to a PC, server, or any other terminal.
- Technologies
  - Current: Ethernet, 802.11 (Wi-Fi)

    Legacy: Token Ring, FDDI, …

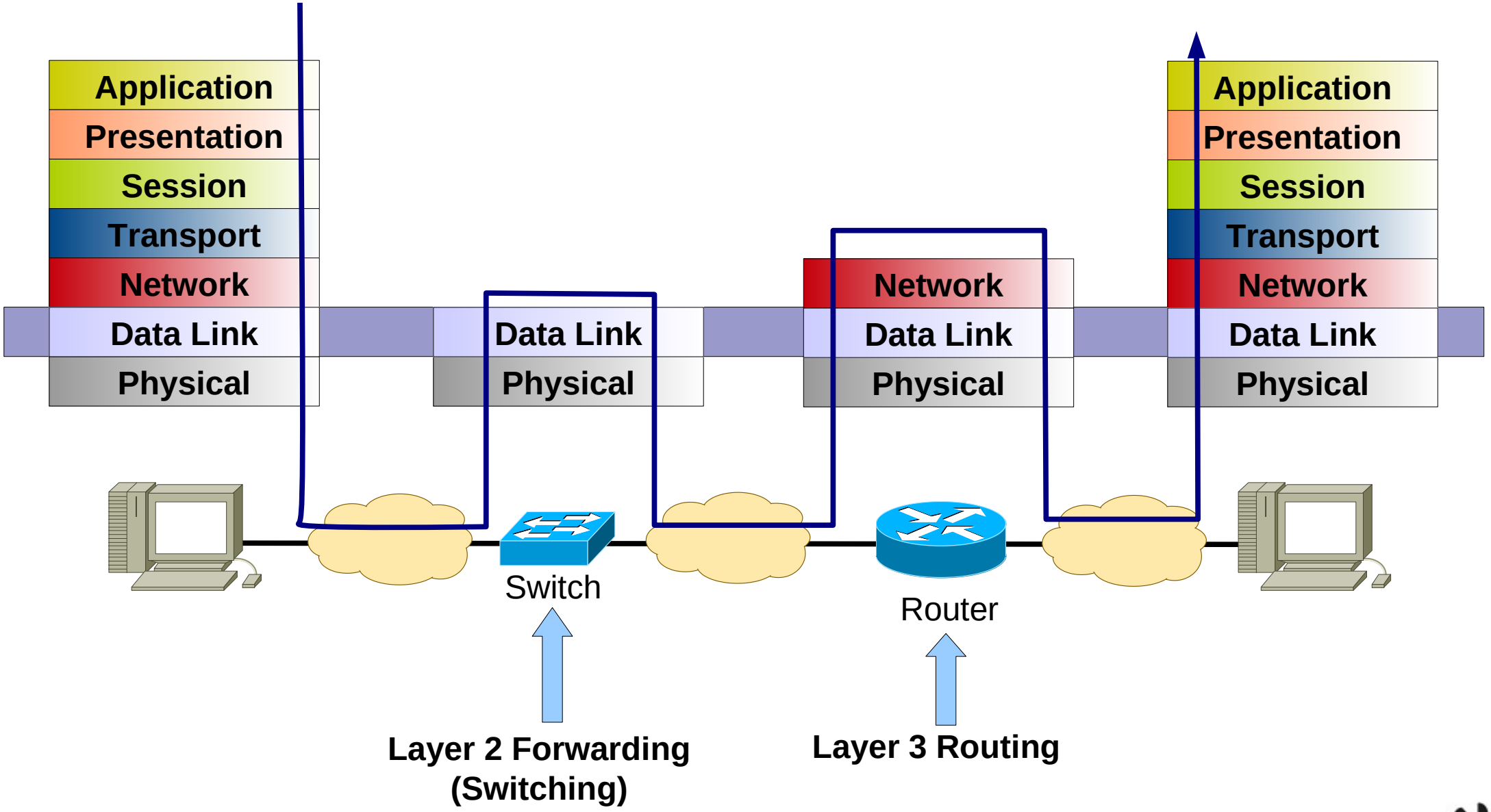


Ethernet

Ethernet (802.3)

- Most successful LAN technology.
- Invented at Xerox Palo Alto Research Center (PARC).
- Xerox, DEC and Intel defined in 1978 the standard for Ethernet 10Mbps.
- Uses “Carrier Sense/Multiple Access” with “Collision Detect” (CSMA/CD)
  - Carrier Sense: hosts can perceive is the communication channel is being used.
  - Multiple Access: multiple host can access simultaneously
  - Collision Detect: host “listen” the communication channel while transmitting to detect transmission collisions.
    - Collision: multiple physical signals overlapping and interfering with each other.

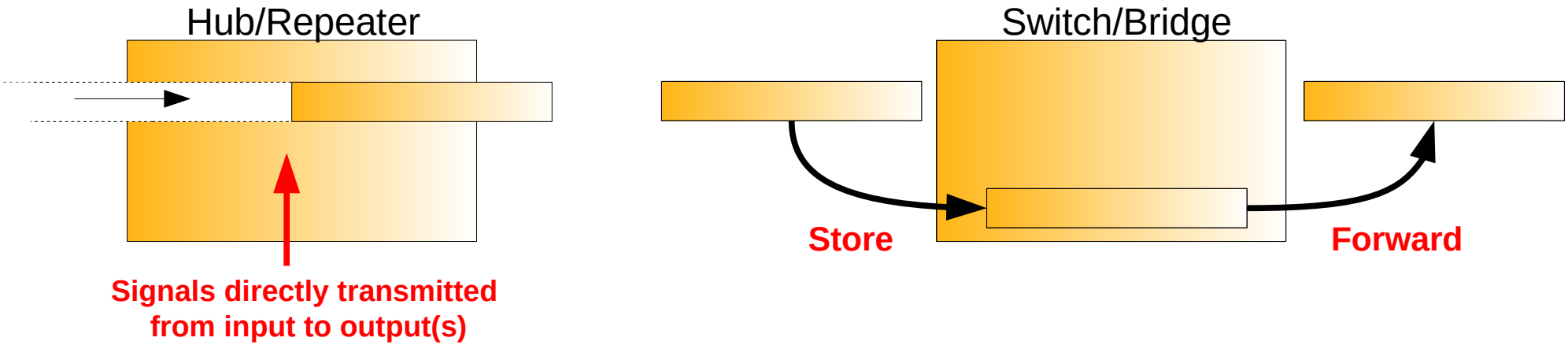
Ethernet based LAN



Ethernet Equipment

- Hub/Repeater:
  - Operates only at the physical level (OSI Layer 1).
  - Replicates and regenerates electrical signals.
  - Hub = repetidor com múltiplas portas.
  - **Não é usado nas redes locais actuais!**
- Switch/Bridge:
  - Store-and-forward operation.
  - Operates only at the data link level (OSI Layer 2).
  - Physically separates (and logically interconnects) different collision domains
    - Nowadays all Ethernet hosts are connected to a switch → There no Ethernet collision domains!
  - Forwards frames based on MAC addresses.
  - Switch = bridge with multiple ports.
- Router:
  - Store-and-forward operation.
  - Operates only at the network level (OSI Layer 3).
  - Routes packets based on network addresses (e.g., IPv4 and IPv6).

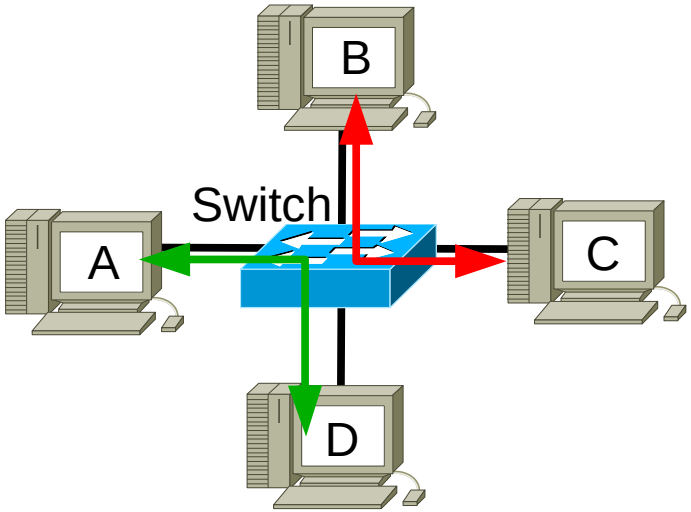
Switches/Bridges vs. Hubs/Repeaters

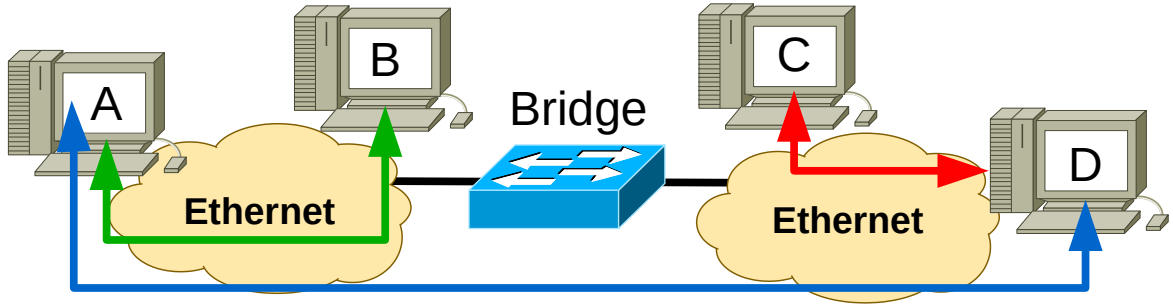


- Bridges/switches interconnect different local networks.
- Bridges/switches additional functions:
  - Store & Forward + Filtering
    - The Forwarding process decides to send a frame to a specid port based on the destination MAC address of the frame.
    - Ports may operate a different speeds.

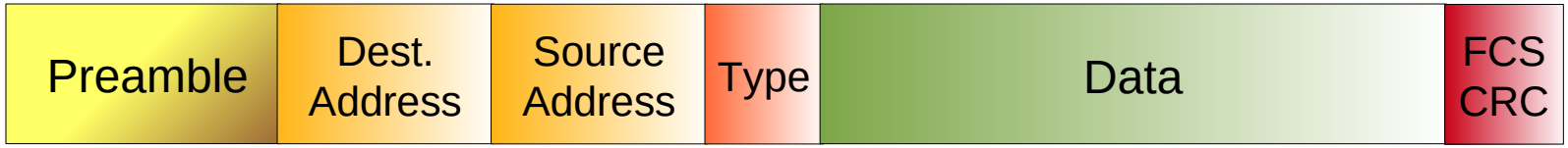
Switching

- With Switches/Bridges
  - Interconnection done at OSI Layer 2.
  - Hosts can transmit simultaneously.
  - A network of Switches is a **Broadcast Domain**
    - An Ethernet frame with destination FF:FF:FF:FF:FF:FF (Broadcast) will reach all connected switches and hosts.





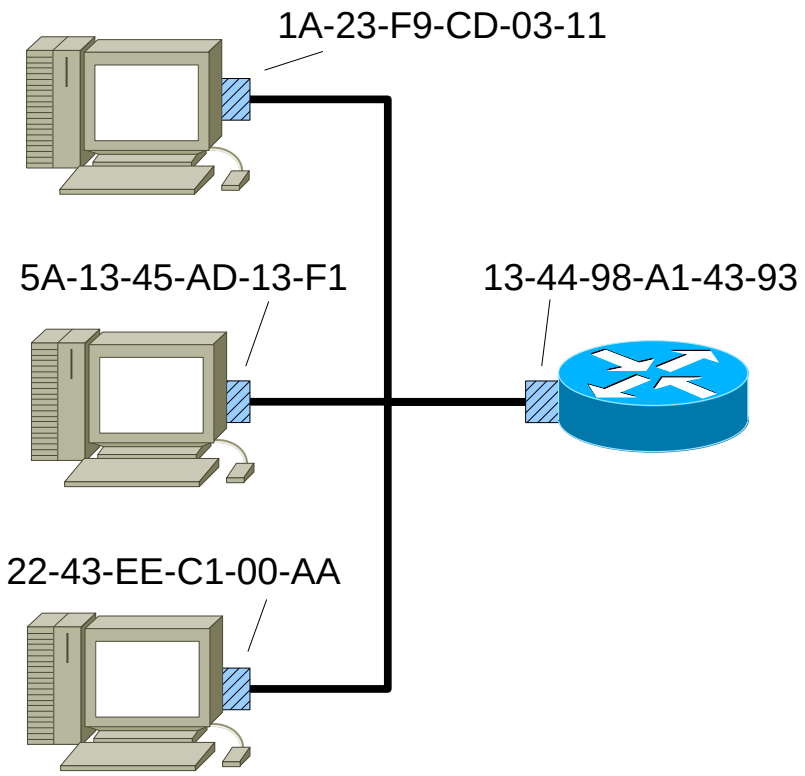
Ethernet Frame



- The sender's network card encapsulates an IP datagrama (or any other network protocol) in an Ethernet frame.
- Preamble:
  - 7 bytes with pattern 10101010 followed by one byte with pattern10101011.
  - Used to sincronize the sending and receiving clocks.
- Destination and Source addresses: 6 bytes Physical (MAC) address
  - If the network card receives a frame with destination equal to its own address or its the broadcast address, it will pass data to the network level process.
  - If not, drops the frame.
- Type defines which protocol is encapsulated in the frame (usually IPv4 or IPV6).
- The frame check sequence (FCS) is a four-octet cyclic redundancy check (CRC) that allows detection of corrupted data within the entire frame as received on the receiver side.

MAC Addresses

- MAC (Physical, Ethernet or LAN) Address:
  - Function: Allow the exchange of data between network interfaces connected using a Layer 2 network.
  - Have 6 bytes/48 bits.
  - Are unique.
  - Each network card has its own address.
  - Defined by manufacturer
    - Some hardware allows change.
    - First 24-, 28-, or 36-bits assign to manufacturer.
  - Hexadecimal notation
    - Broadcast: FF-FF-FF-FF-FF-FF



Ethernet Frame Minimum Size

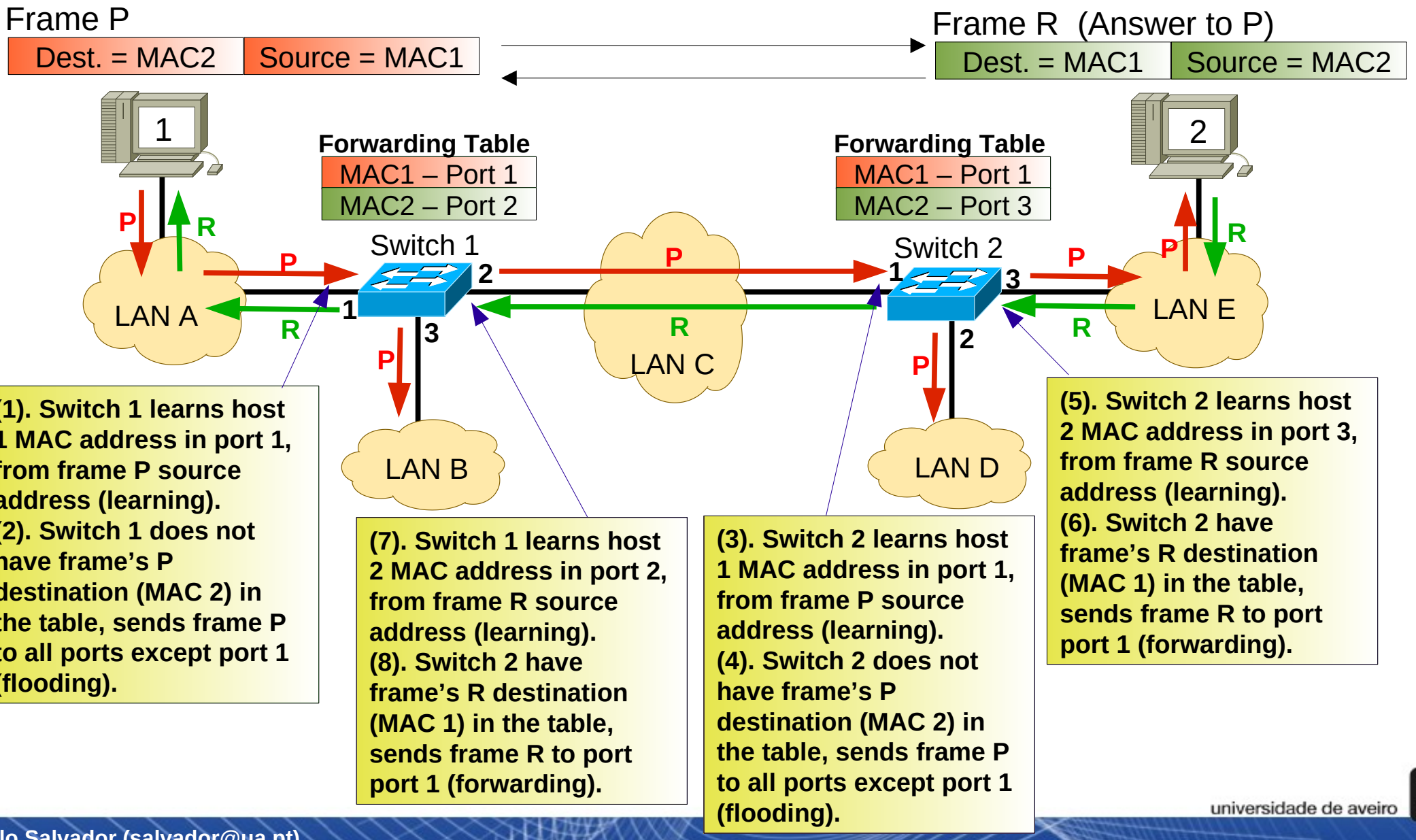
- Historically there were Ethernet technologies that allowed collisions and a collision detection mechanism had to be present (CSMA/CD).
- Depending on the technology and maximum cable size, the Ethernet frame had to be big enough to allow the collision detection mechanism to detect a frame being transmitted before the last frame byte leaving the source host.
- By legacy (it is possible to merge different Ethernet technologies) the **minimum frame size is 64 bytes**.
- If the frame's header plus data do not reach 64 bytes, a set of zeros must be added to the end of the frame to reach 64 bytes.
  - This is called **padding**.

Switches Basic Operations

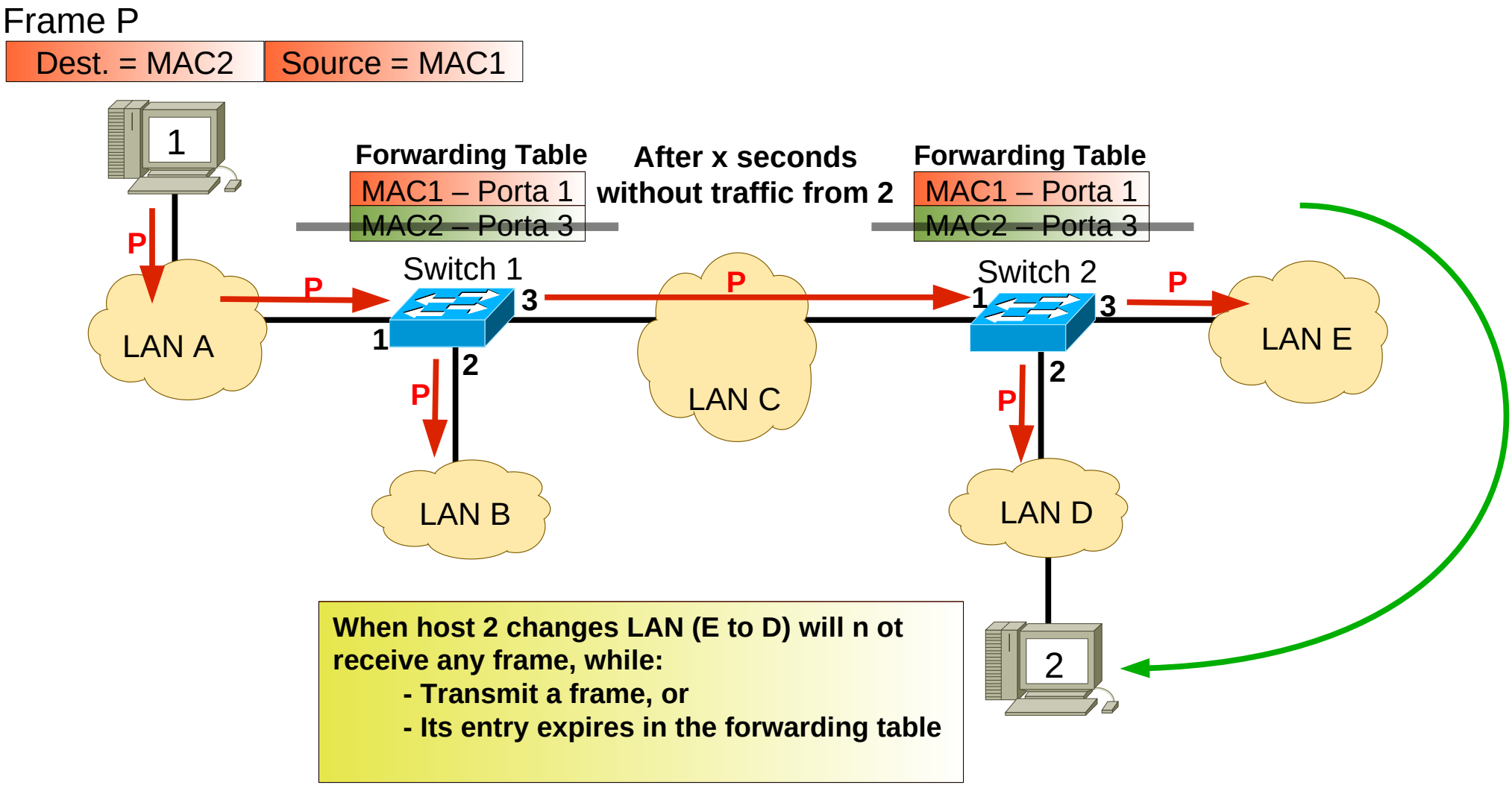
MAC	Porta
00:11:11:11:11:11	1
00:22:22:22:22:22	1
A1:33:33:33:33:33	2
44:44:44:44:44:44	3
55:55:55:00:00:55	3

- Switches have a **Forwarding Table**.
- When a switch receives an Ethernet frame:
  - Registers an entry at the Forwarding Table the frame's source MAC address and the port where the frame was received.
    - If no frames are received from that MAC address after some time (**aging time**) the entry is removed.
  - Searches the Forwarding Table for the frame's destination MAC address and forwards the packet according:
    - **Forwarding** mechanism:
      - If the frame's destination MAC address exists in the table, the switches forwards the frame through the port associated with that MAC address.
    - **Flooding** mechanism:
      - If the frame's destination MAC address DOES NOT exist in the table, the switches forwards the frame through all active ports (except the one where it was received).
        - » Note: Just within the same VLAN (more details later).

Learning, Flooding and Forwarding

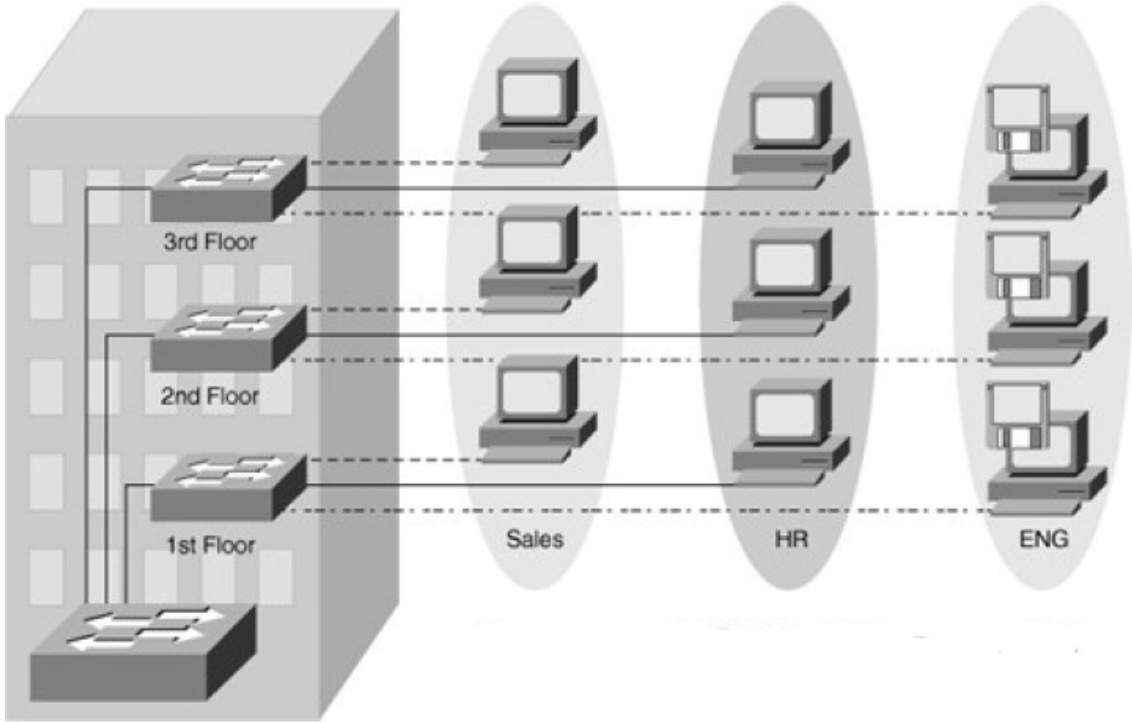


Forwarding Table Aging Time



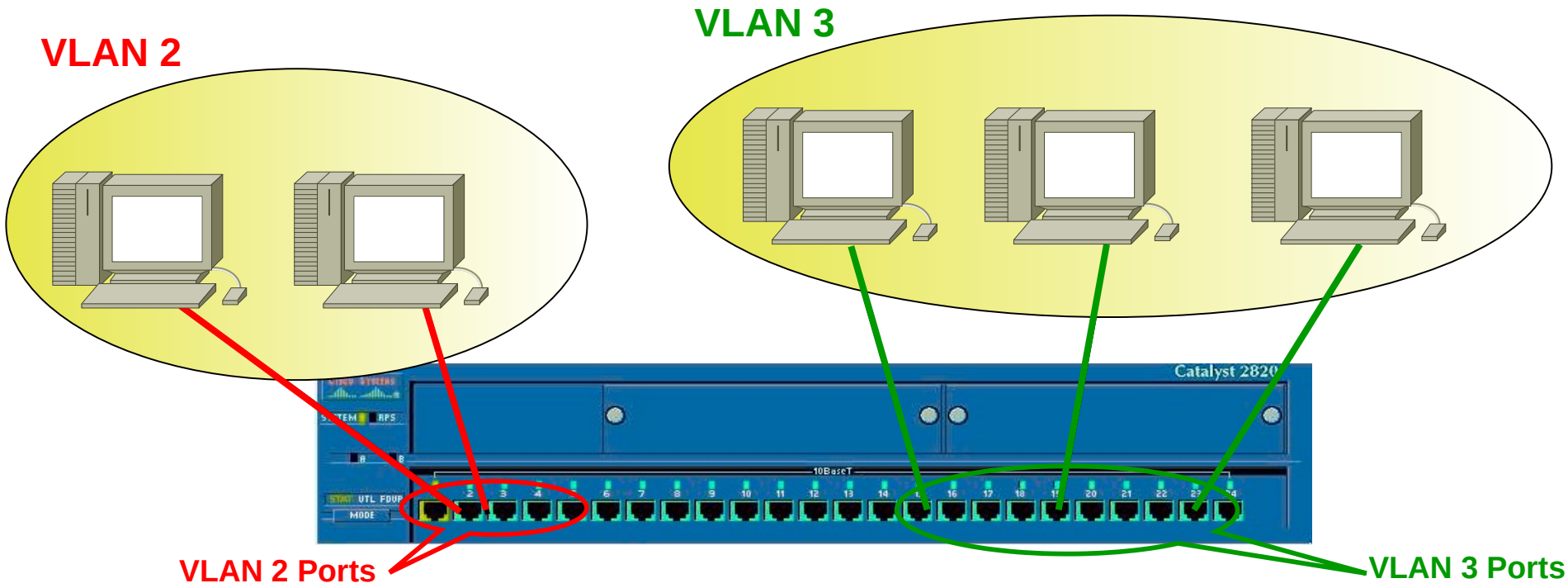
Virtual LAN (VLAN)

- A Virtual LAN (VLAN) is a group of hosts/users with a common set of requirements or characteristics in the same broadcast domain.
  - Independent of their physical location.
- Solves the scalability problems of large networks.
  - By breaking a single broadcast domain into several smaller broadcast domains.
  - Allows better/simpler network administration and security deployment.
- Hosts in different VLAN do not communicate by Layer 2.
  - Its communications are done at Layer 3 (with IP routing).



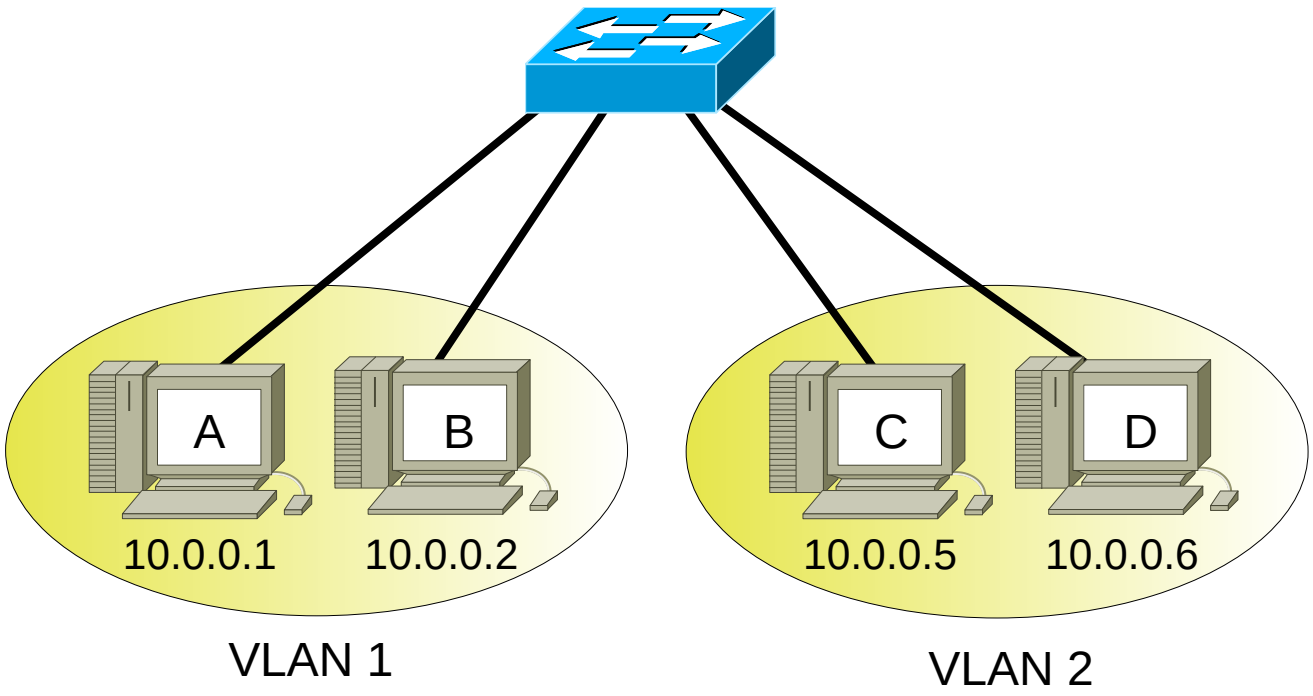
Defining Host VLAN

- The VLAN to which a host belongs depends only on the port of the switch.
  - Configured only in the switch.
  - Example: If port 1 is configure as VLAN 2, and port 20 is configured as VLAN 3:
    - If host is connected to port 1 it is on VLAN 2,
    - If host is connected to port 20 it is on VLAN 3.
- VLAN 1 is usually reserved to network administration.
  - Used to access configurations remotely via IP.



Example – VLAN

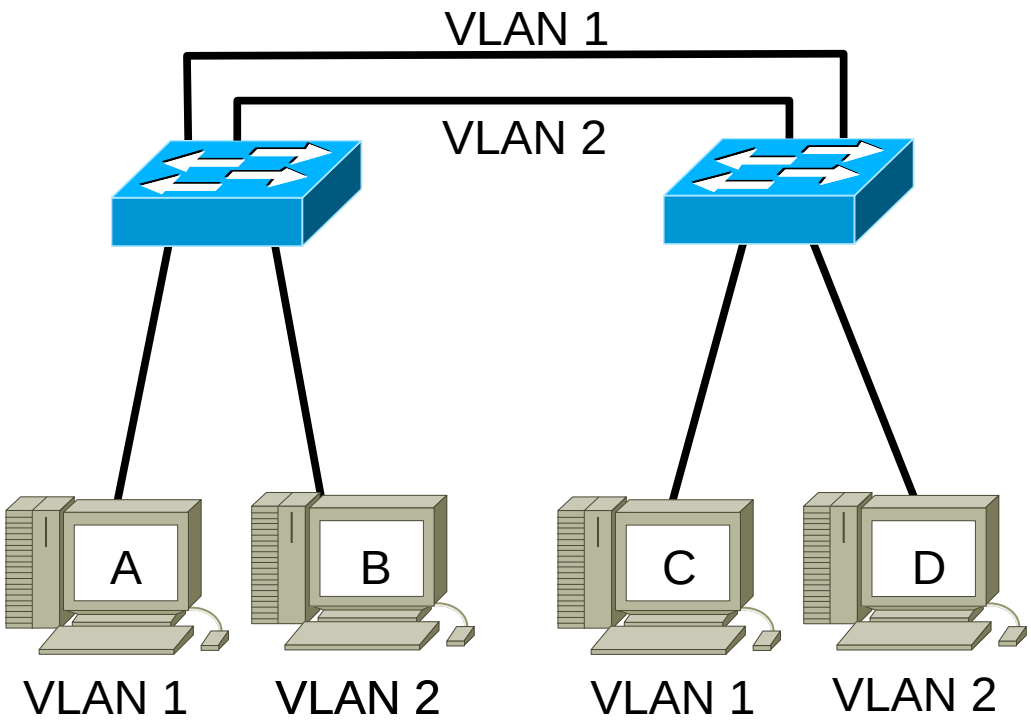
Pings sent by 10.0.0.1



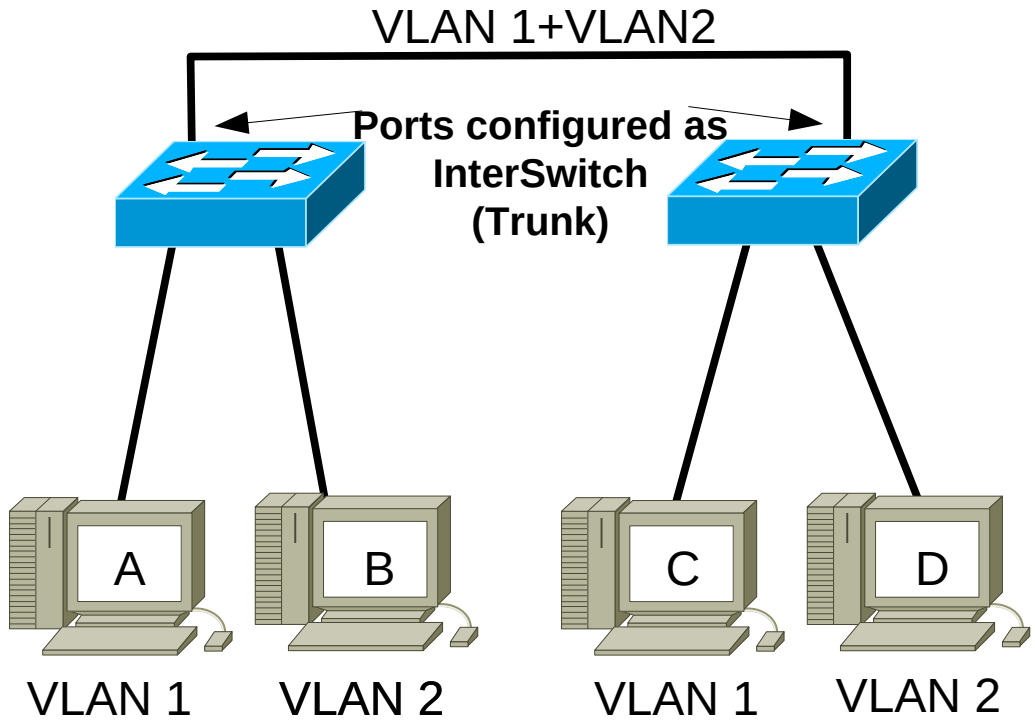

```
# ping 10.0.0.2

Pinging 10.0.0.2 with 32 bytes of data:

Reply from 10.0.0.2: bytes=32 time<10ms TTL=128
Reply from 10.0.0.2: bytes=32 time<10ms TTL=128
Reply from 10.0.0.2: bytes=32 time<10ms TTL=128
Reply from 10.0.0.2: bytes=32 time<10ms TTL=128

Ping statistics for 10.0.0.2:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 0ms, Maximum =  0ms, Average =  0ms

# ping 10.0.0.5

Pinging 10.0.0.5 with 32 bytes of data:

Request timed out.
Request timed out.
Request timed out.
Request timed out.

Ping statistics for 10.0.0.5:
    Packets: Sent = 4, Received = 0, Lost = 4 (100% loss),
Approximate round trip times in milli-seconds:
    Minimum = 0ms, Maximum =  0ms, Average =  0ms

# ping 10.0.0.6

Pinging 10.0.0.6 with 32 bytes of data:

Request timed out.
Request timed out.
Request timed out.
Request timed out.
```

Interconnection of Switches

- Physical link per VLAN

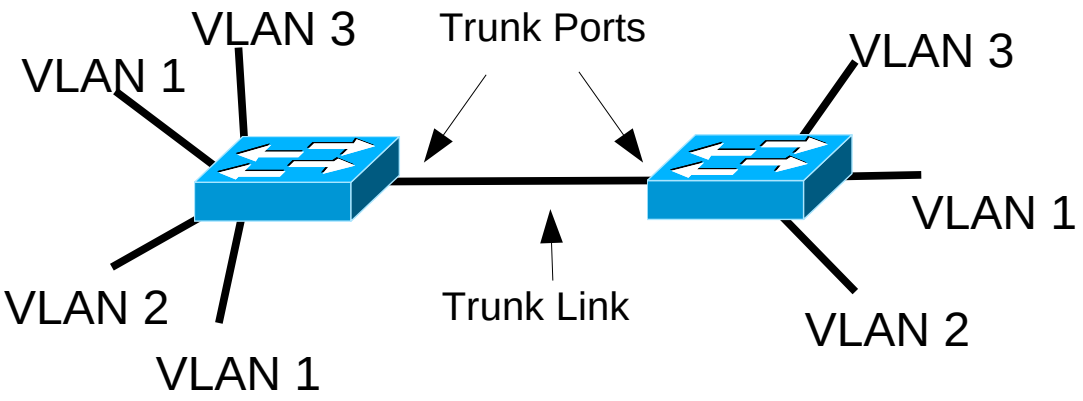


- With a single physical link.
- Using InterSwitch/Trunk port(s).

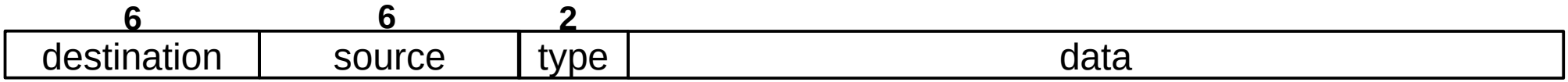


- Using a single physical link requires a mechanism to differentiate frames from different VLAN.
  - Frames must have a tagged
    - Added when forwarding to a trunk port.
    - Read and removed when receiving a frame from a trunk port

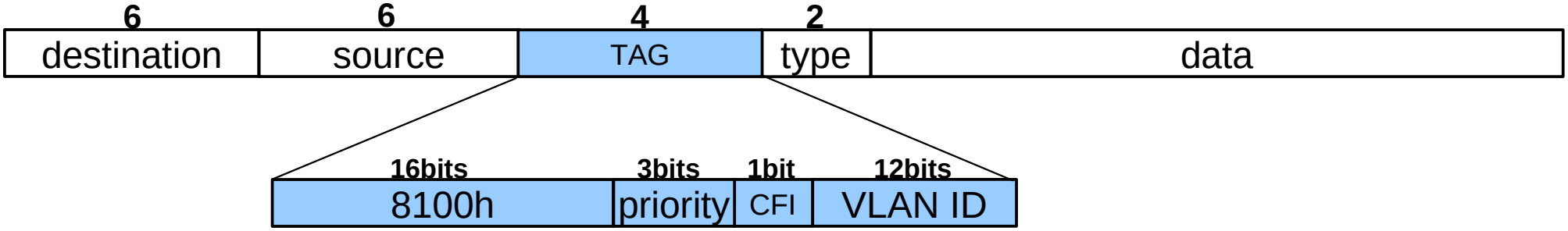
IEEE802.1Q Standard



Ethernet frame without a VLAN tag



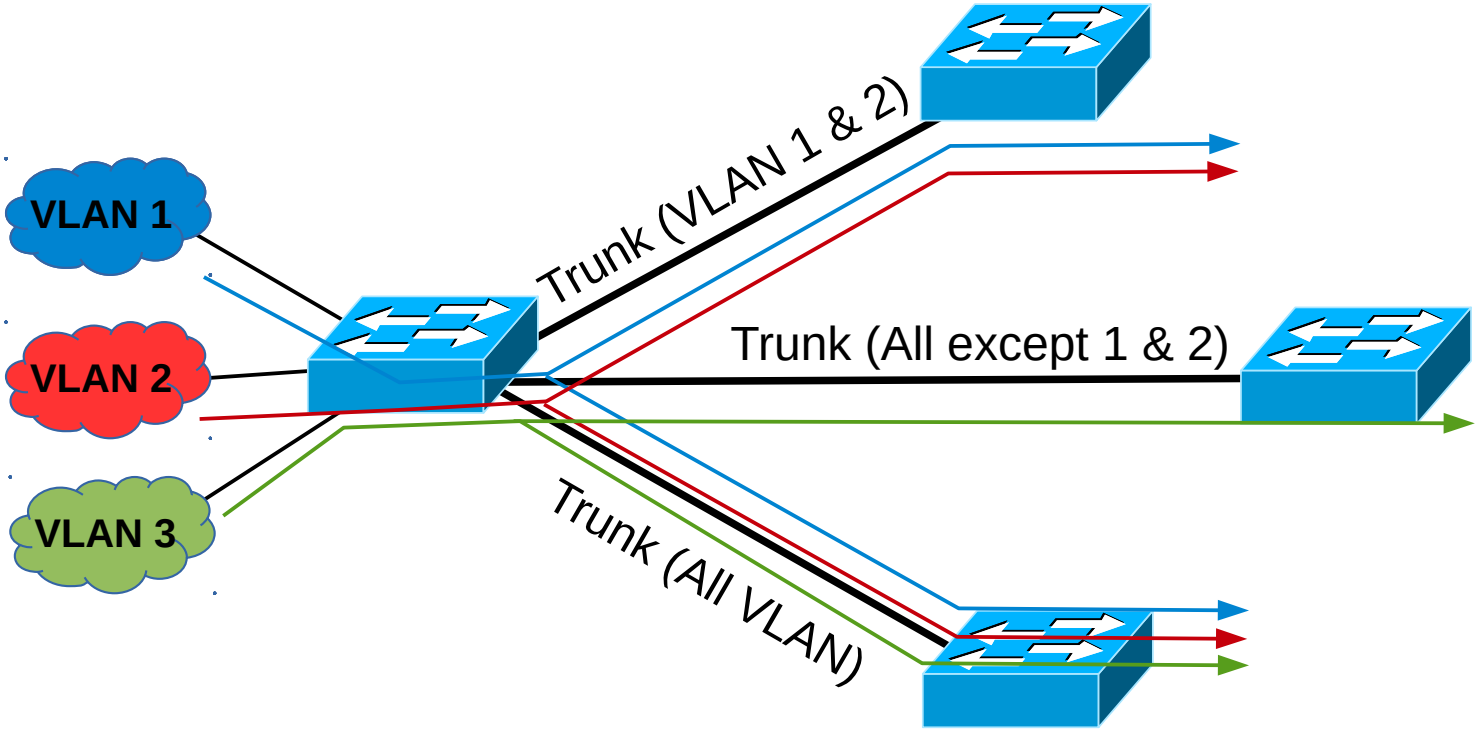
Ethernet frame with a VLAN tag



- Priority: Traffic relative priority according to standard 802.1q (0 to 7 values).
- CFI: Used to guarantee compatibility with older technologies (always zero in Ethernet).
- VLAN ID: VLAN identifier.

Trunk Links

- The physical link between two Trunk ports is called a Trunk link.
- A trunk carries traffic for multiple VLANs using IEEE 802.1Q.
  - Inter-Switch Link (ISL) encapsulation is an alternative but it getting obsolete.
- Trunks may transport all VLAN or only some!



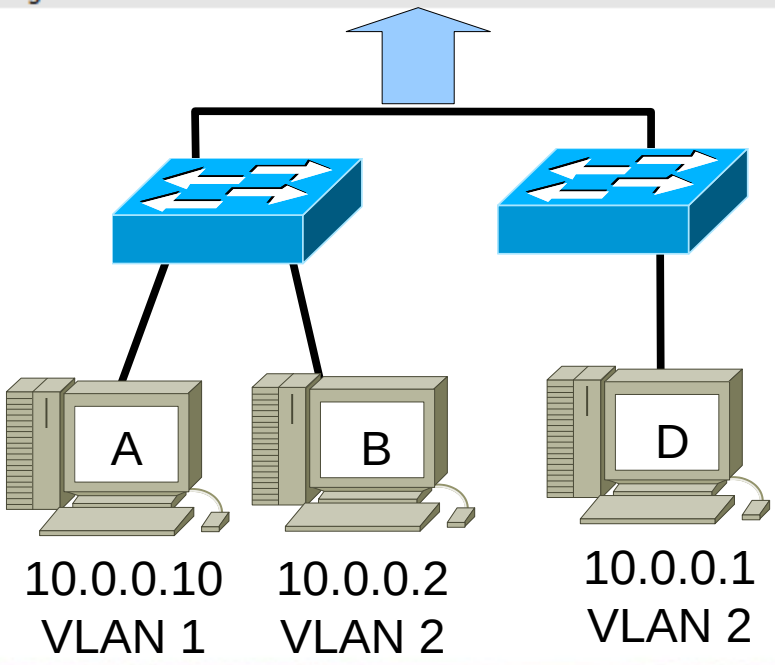
Example – InterSwitch/Trunk Ports

Filter: icmp | Expression... | Clear | Apply

No.	Time	Source	Destination	Protocol	Info
23	11.535990	10.0.0.2	10.0.0.1	ICMP	Echo (ping) request
24	11.536995	10.0.0.1	10.0.0.2	ICMP	Echo (ping) reply
27	12.538443	10.0.0.2	10.0.0.1	ICMP	Echo (ping) request
28	12.539186	10.0.0.1	10.0.0.2	ICMP	Echo (ping) reply

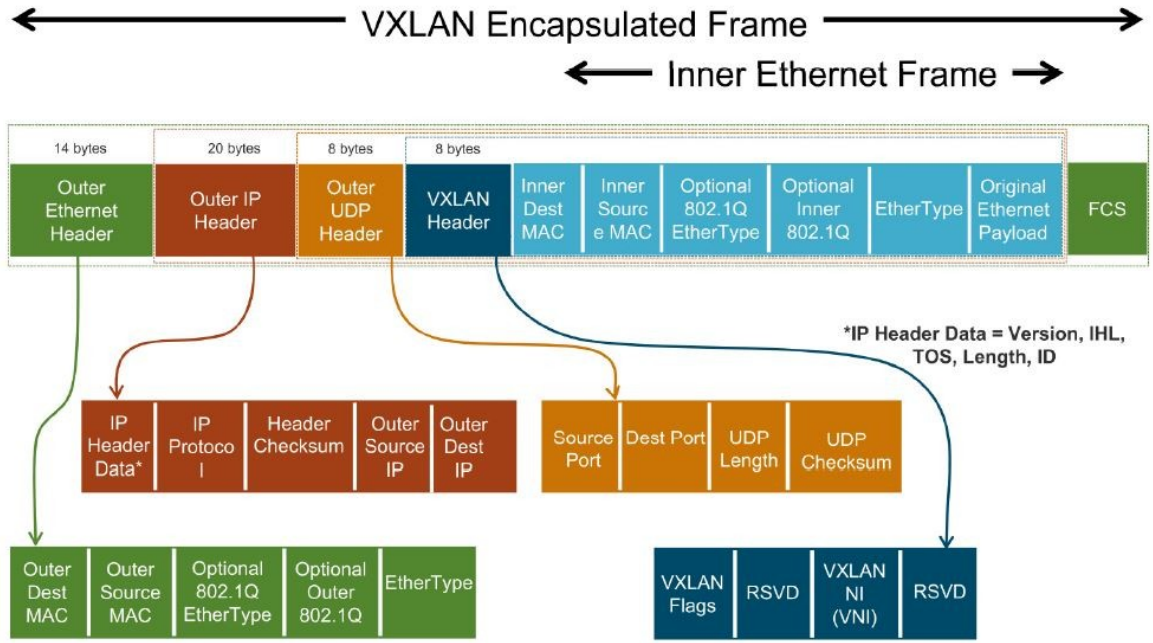
```
▷ Frame 23 (102 bytes on wire, 102 bytes captured)
▷ Ethernet II, Src: 00:aa:00:53:7c:00 (00:aa:00:53:7c:00), Dst: 00:aa:00:fa:67:00 (00:aa:00:fa:67:00)
▽ 802.1Q Virtual LAN, PRI: 0, CFI: 0, ID: 2
    000. .... .... .... = Priority: 0
    ...0 .... .... .... = CFI: 0
    .... 0000 0000 0010 = ID: 2
    Type: IP (0x0800)
▷ Internet Protocol, Src: 10.0.0.2 (10.0.0.2), Dst: 10.0.0.1 (10.0.0.1)
▷ Internet Control Message Protocol
```

ID:2 == VLAN 2



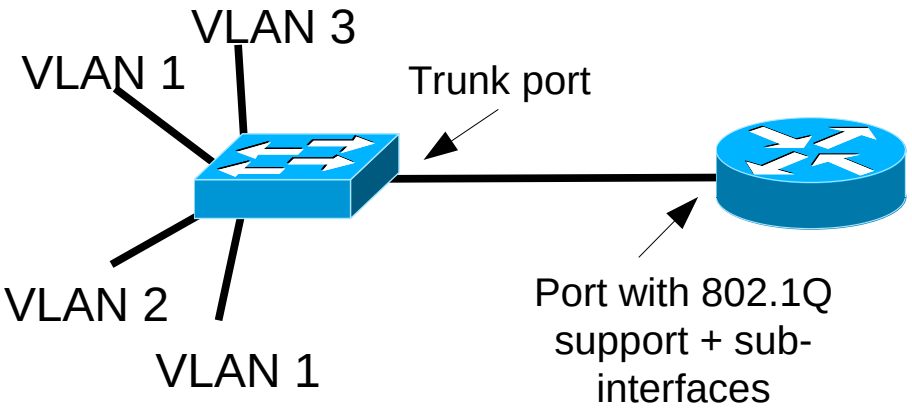
Virtual Extensible LAN (VXLAN)

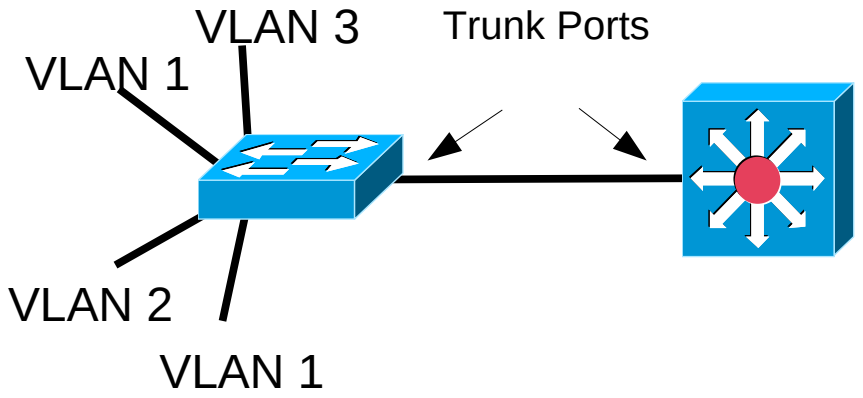
- Alternative/Complement to 802.1Q in Layer3 Switches.
- Encapsulates OSI Layer 2 Ethernet frames within Layer 4 UDP/IP datagrams .
  - Default port 4789.
- VLAN may be additionally identified by a VNI field with 24 bits.
  - 802.1Q tag only as 12 bits.
  - Allows for a very large number of VLAN.
- Usually used when connecting remote VLAN (connected only via IP) in Datacenter and Cloud scenarios.



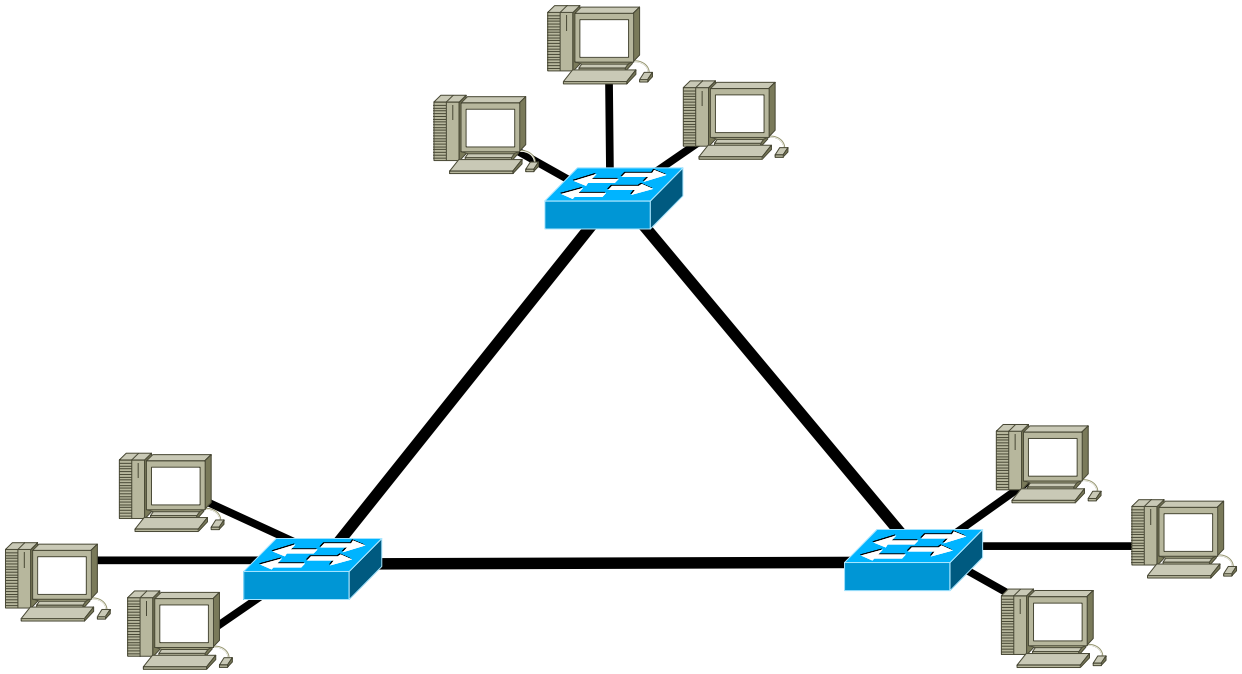
IP Connection between VLANs

- To communicate between different VLAN it is required to use Layer 3 (IP Routing).
- Common solutions:
  - A router with support to 802.1Q,
    - Connecting the physical router interface to a Trunk port.
    - The router's physical interface is sub-divided in sub-interfaces (one for each VLAN).
    - The IP gateway for a VLAN host is the IP address of the respective sub-interface in the Router.
  - A Layer 3 switch,
    - Connecting both switches (L3 and L2) using Trunk ports.
    - Each VLAN is mapped to a virtual Layer 3 interface.
    - The IP gateway for a VLAN host is the IP address of the respective virtual interface in the L3 switch.





Redundant Layer 2 Network

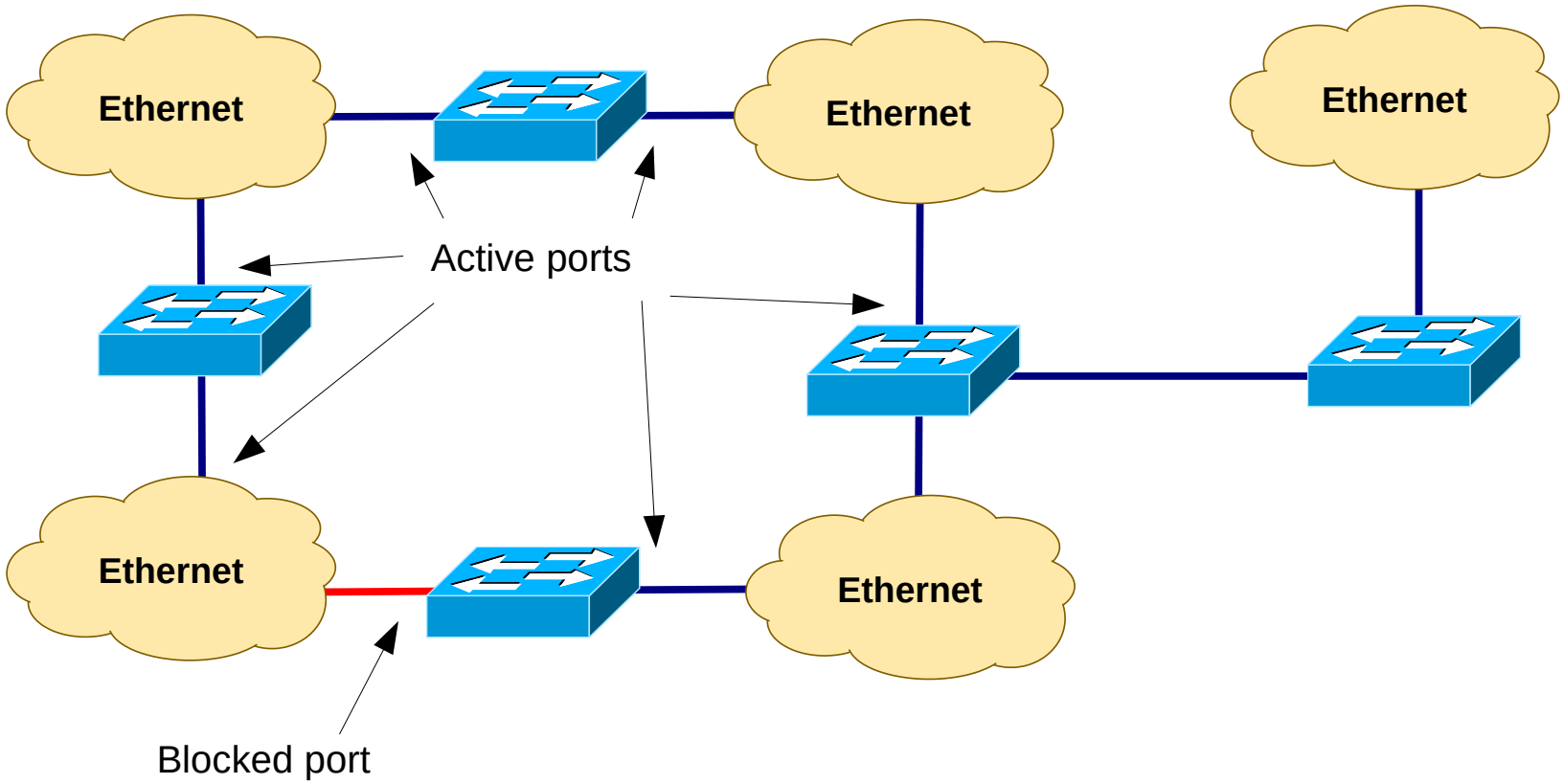


- Objective: Allow the network for dynamically recover from network failures.
- Problem: Link redundancy creates Layer 2 loops. Causes the collapse of communications when MAC frames with broadcast address are sent by any host due to infinite frame flooding.

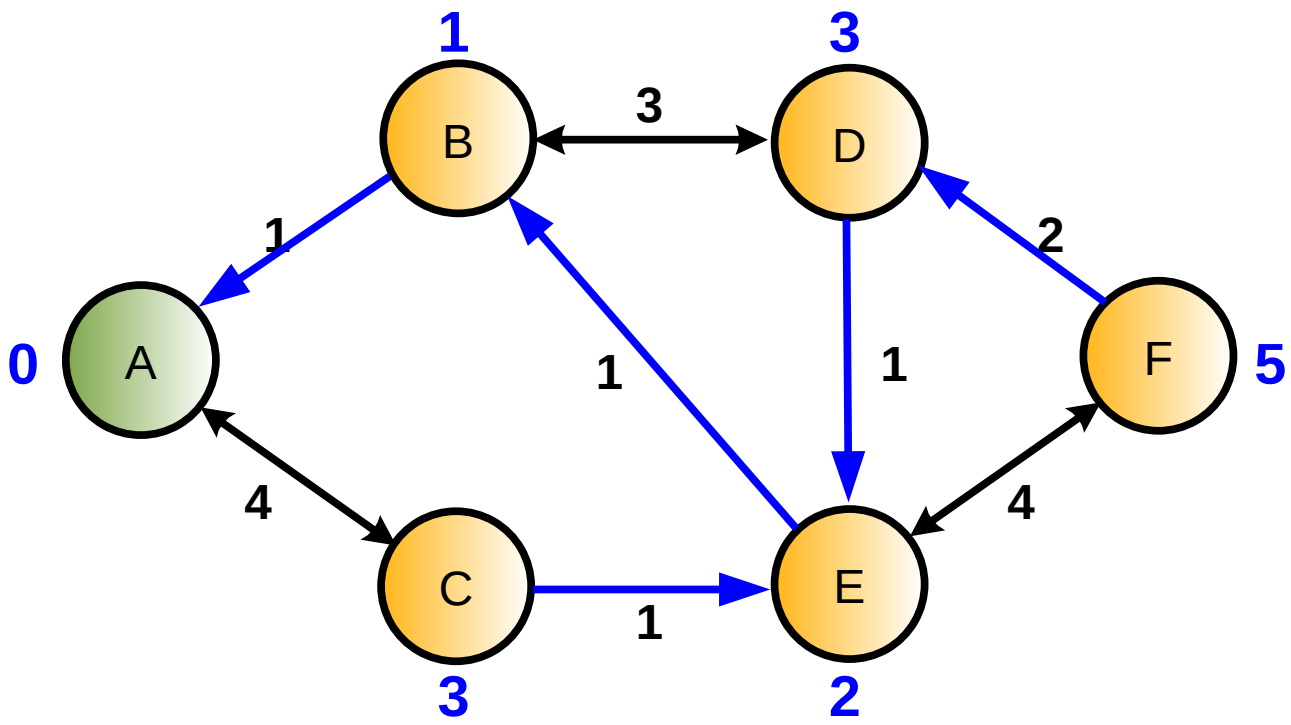
Spanning Tree Protocol (SPT)

- STP enables the network to deterministically block ports and provide a loop-free topology in a network with redundant links.
- There are several STP Standards and Features:
  - STP is the original IEEE 802.1D version (802.1D-1998) that provides a loop-free topology in a network with redundant links.
  - RSTP, or IEEE 802.1W, is an evolution of STP that provides faster convergence of STP.
  - Multiple Spanning Tree (MST) is an IEEE standard. MST maps multiple VLANs into the same spanning-tree instance.
  - Per VLAN Spanning Tree Plus (PVST+) is a Cisco enhancement of STP that provides a separate 802.1D spanning-tree instance for each VLAN configured in the network.
  - RPVST+ is a Cisco enhancement of RSTP that uses PVST+. It provides a separate instance of 802.1W per VLAN.

Spanning-Tree



Bellman Equations



- When link cost are not negative, then:

Shortest path from one node X to node A

=

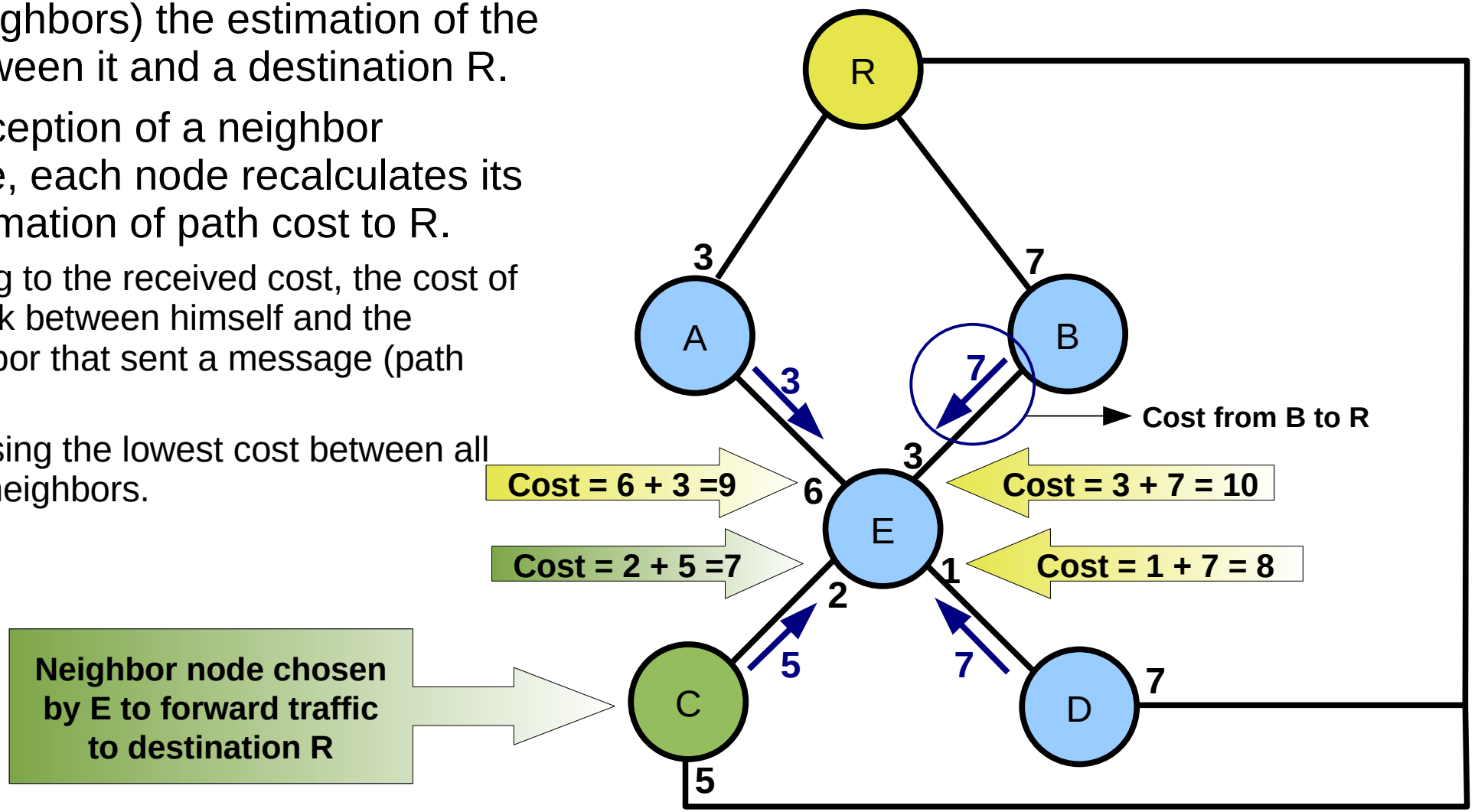
Cost of the link from that node X to the node that follows it in the shortest path to A

+

Shortest path from that node to node A

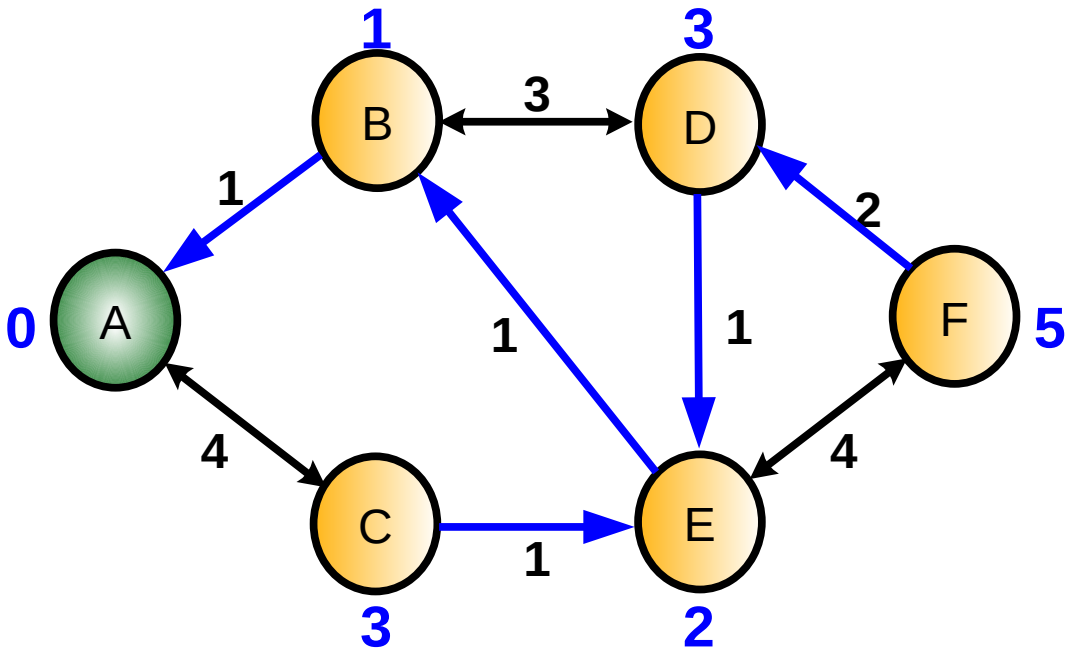
Bellman-Ford Distributed and Asynchronous Algorithm

- Each node transmits periodically (to all its neighbors) the estimation of the cost between it and a destination R.
- Upon reception of a neighbor message, each node recalculates its own estimation of path cost to R.
  - Adding to the received cost, the cost of the link between himself and the neighbor that sent a message (path cost).
  - Choosing the lowest cost between all links/neighbors.

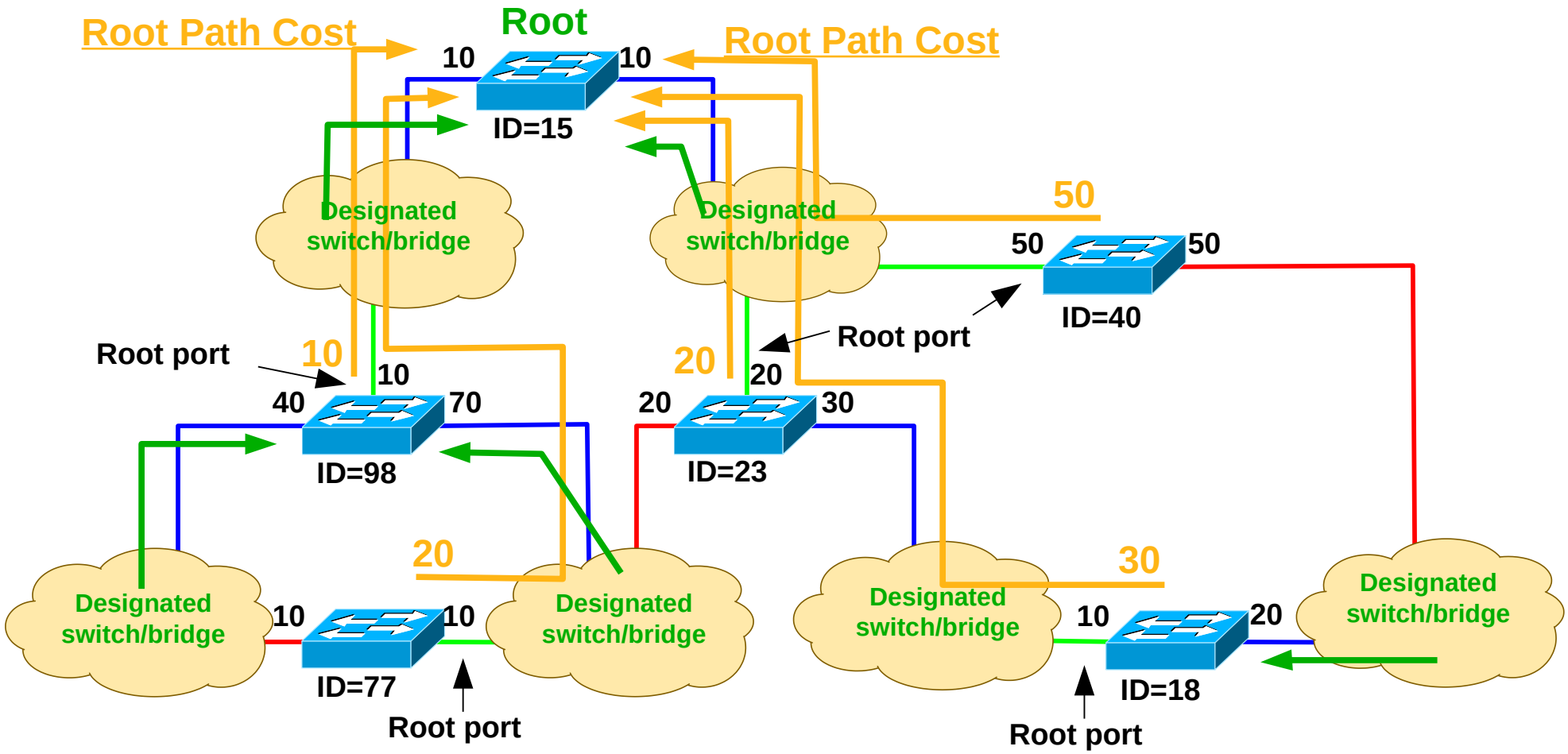


Routing based on Spanning Trees

- It is chosen an origin/root node.
- All nodes use the **Bellman-Ford Distributed and Asynchronous Algorithm** to calculate the neighbored node (and respective path cost) that provide the smallest cost to the origin/root node.
- The set of links used by all nodes to provide the shortest paths to the origin/root node is called the **Spanning Tree**.
- It is required a criteria to solve ties.

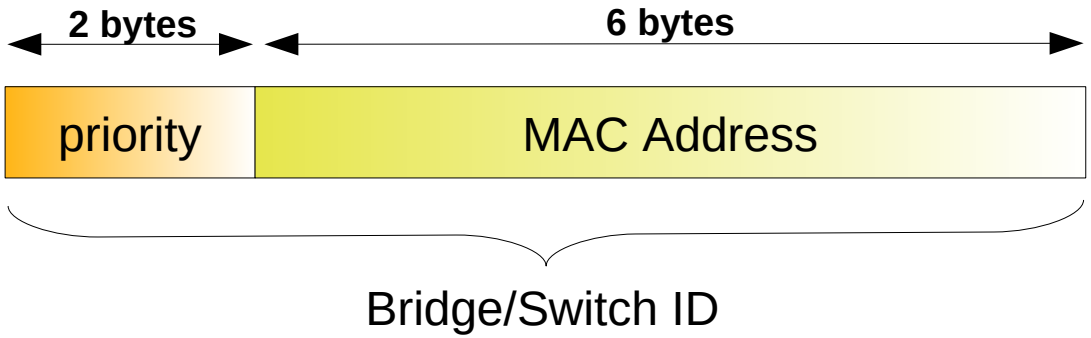


Spanning Tree Basic Concepts (1)



Spanning Tree Basic Concepts (2)

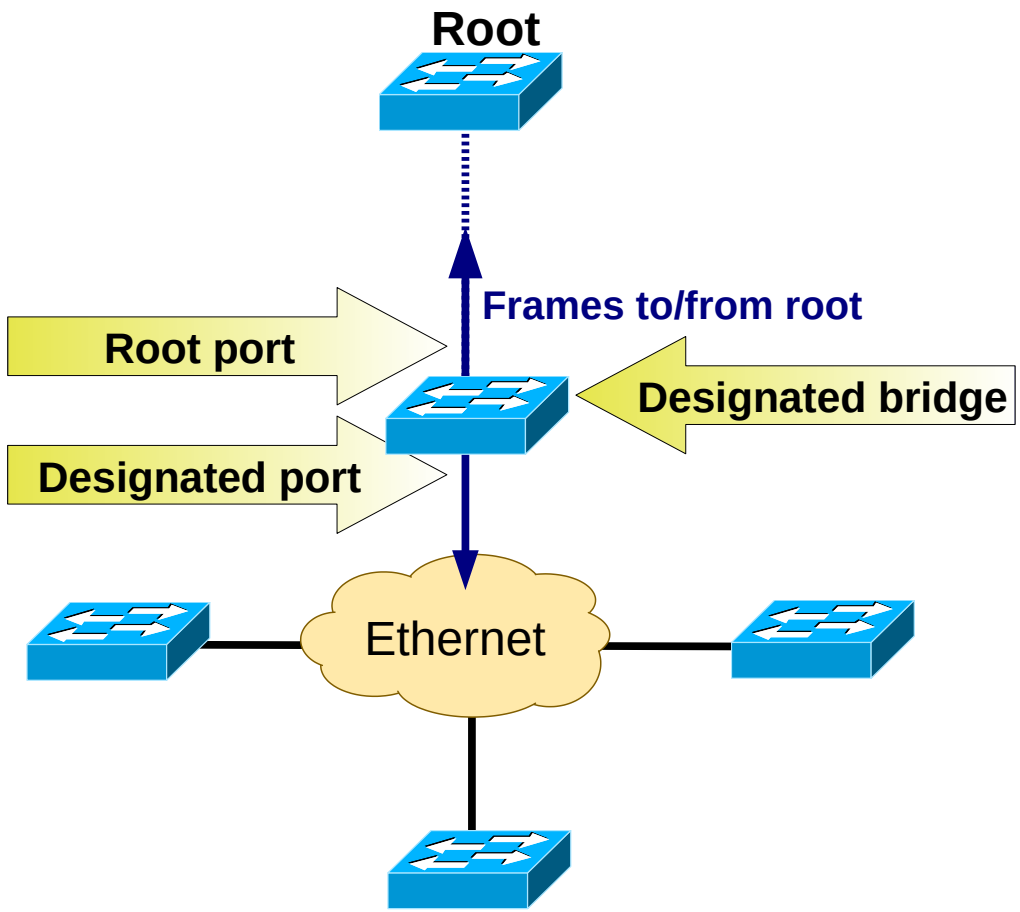
- Bridge/Switch ID – each switch is identified by an 8 bytes identifier based on:
  - 2 **Priority** bytes, defined by configuration.
  - 6 bytes (one of the **MAC Address** of the switch, or any other unique 48 bit sequence).
  - Priority has precedence over the 6 bytes sequence (usually MAC address).



- Root Switch/bridge – Switch chosen as origin/root of the spanning tree.
  - Switch com **lowest ID**.
- Path cost – Cost associated with each port.
  - Has a default value, but can be changed by configuration.

Spanning Tree Basic Concepts (3)

- Designated Bridge – Switch responsible to forward the packets from an Ethernet segment to and from the root.
  - The root bridge is the designated bridge to all Ethernet segments connected to it.
- Designated Port – Port of the designated bridge that connects an Ethernet segment (to which is designated).
- Root Port – Port of the designated bridge that provides the path to the root.



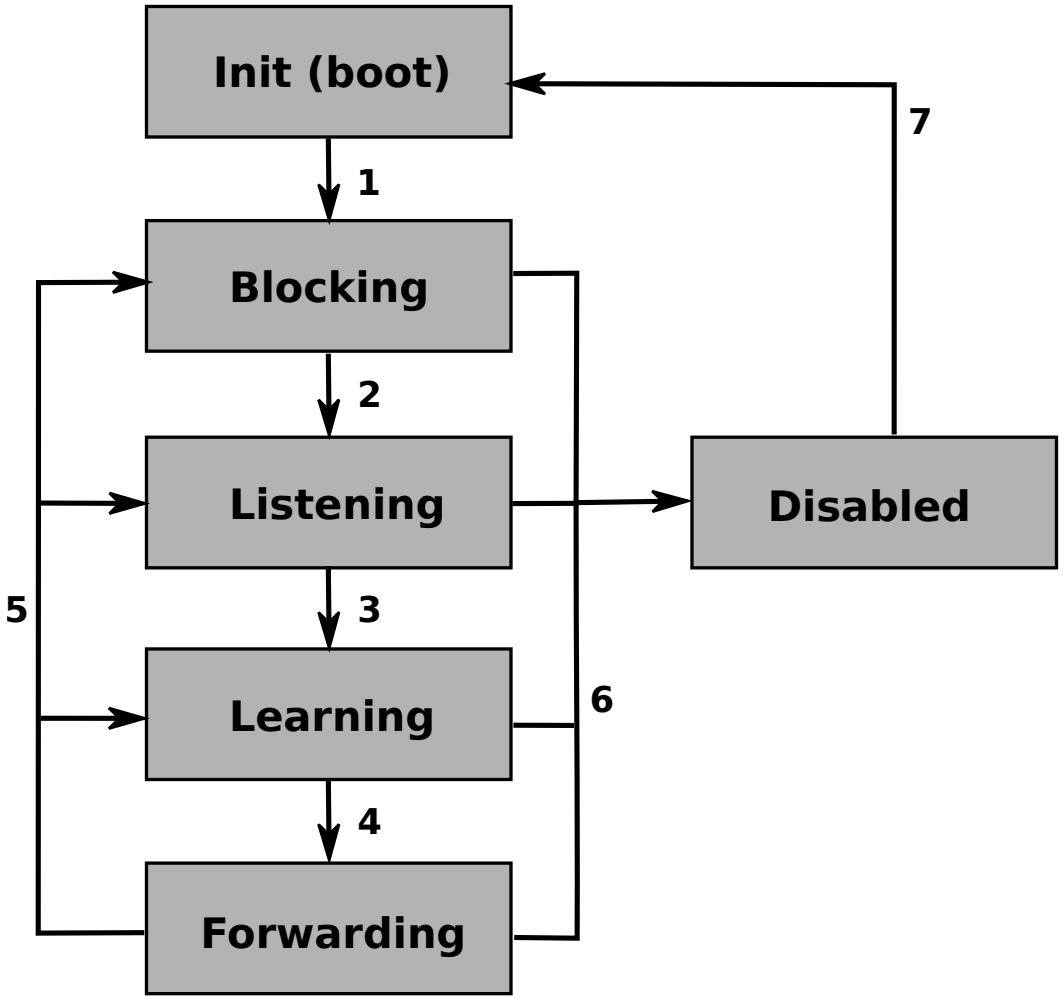
Spanning Tree Basic Concepts (4)

- Possible Port States
  - **Blocking** state:
    - MAC address learning and packet forwarding are disabled;
    - Receives and processes BPDU.
    - After *MaxAge* time without receiving BPDU, it transitions to Listening state.
  - **Listening** state:
    - MAC address learning and packet forwarding are disabled;
    - Receives and processes BPDU.
    - When *ForwardDelay* timer expires the port transitions to Learning state.
  - **Learning** state:
    - Learns MAC address;
    - Packet forwarding are disabled;
    - Receives and processes BPDU.
    - When *ForwardDelay* timer expires the port transitions to Forwarding state.
  - **Forwarding** state:
    - MAC address learning and packet forwarding are enabled;
    - Receives and processes BPDU.
  - **Disabled** state:
    - MAC address learning and packet forwarding are disabled;
    - Does not receive BPDU.

Spanning Tree Basic Concepts (5)

- Each switch has an associated cost of the shortest path to the root (Root Path Cost), given by the sum of the costs of all root ports along the path to the root.
- The Root Port, in each switch, is the port that provides the best path to the root (**lowest** Root Path Cost).
  - If more than one have the lowest cost, it is chosen the one with the neighbor with the lowest ID.
  - If more of one link is used to connect to the “best” neighbor it is used the one with the lowest (neighbor) port identifier.
- The Designated Bridge, from each Ethernet segment, is the switch with the **lowest** Root Path Cost from all connected to that segment.
  - If more than one have the lowest cost, it is chosen the one with the with the lowest ID.
- The Designated Port, from each Ethernet segment, is the port that connects it to its Designated Bridge.
- The root and designated ports will be in Forwarding state.
- All remaining ports will be in Blocking state.

Port States Diagram

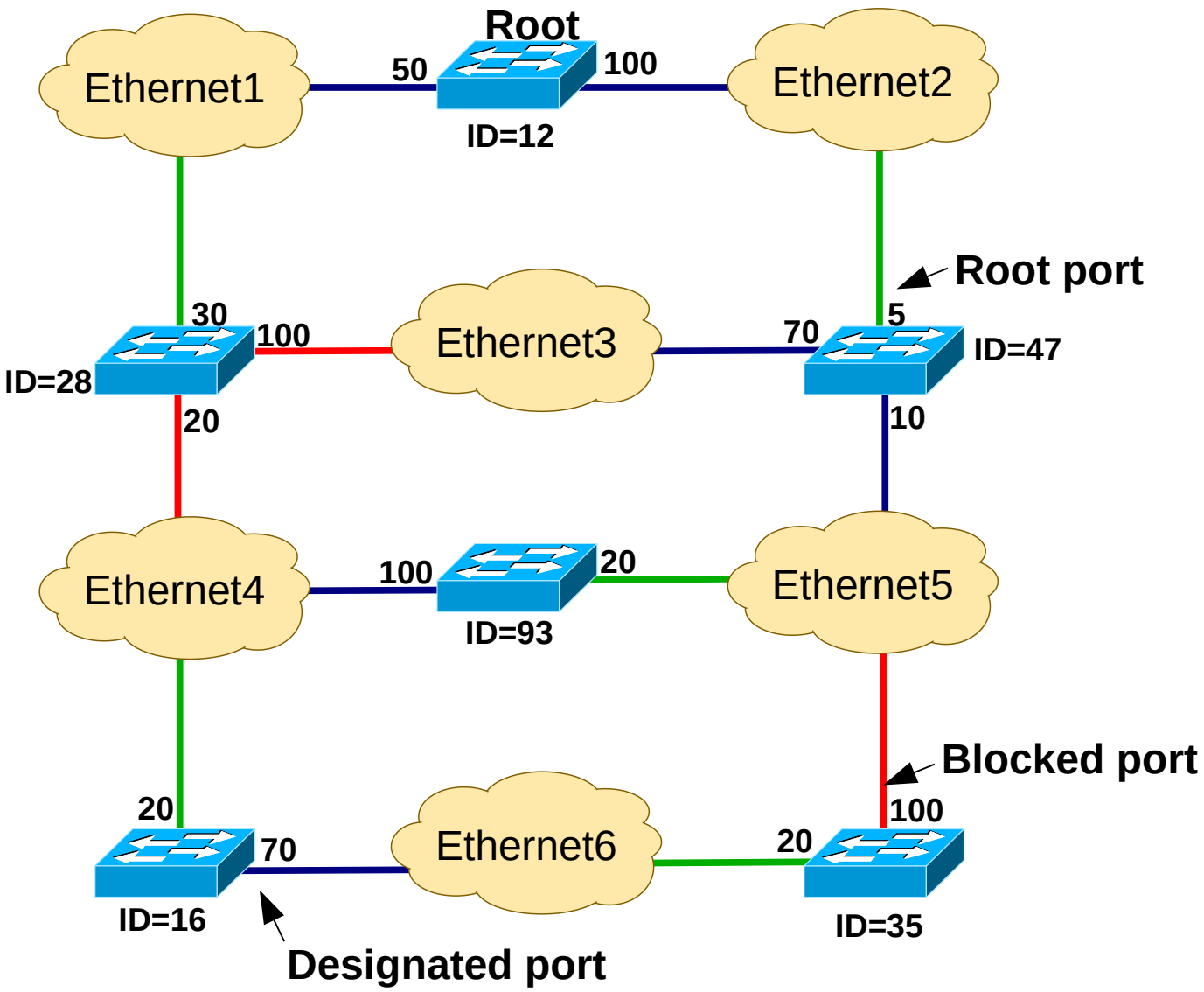


1) A port boots up and transitions to **Blocking** state.
2) When *MaxAge* timer expires the port transitions to **Listening** state.
3) When *ForwardDelay* timer expires the port transitions to **Learning** state.
4) When *ForwardDelay* timer expires the port transitions to **Forwarding** state.
5) After a topology change the port transitions immediately to **Blocking** state.
6) and 7) Administrative actions.

Example – Spanning Tree (1)

Designated bridges

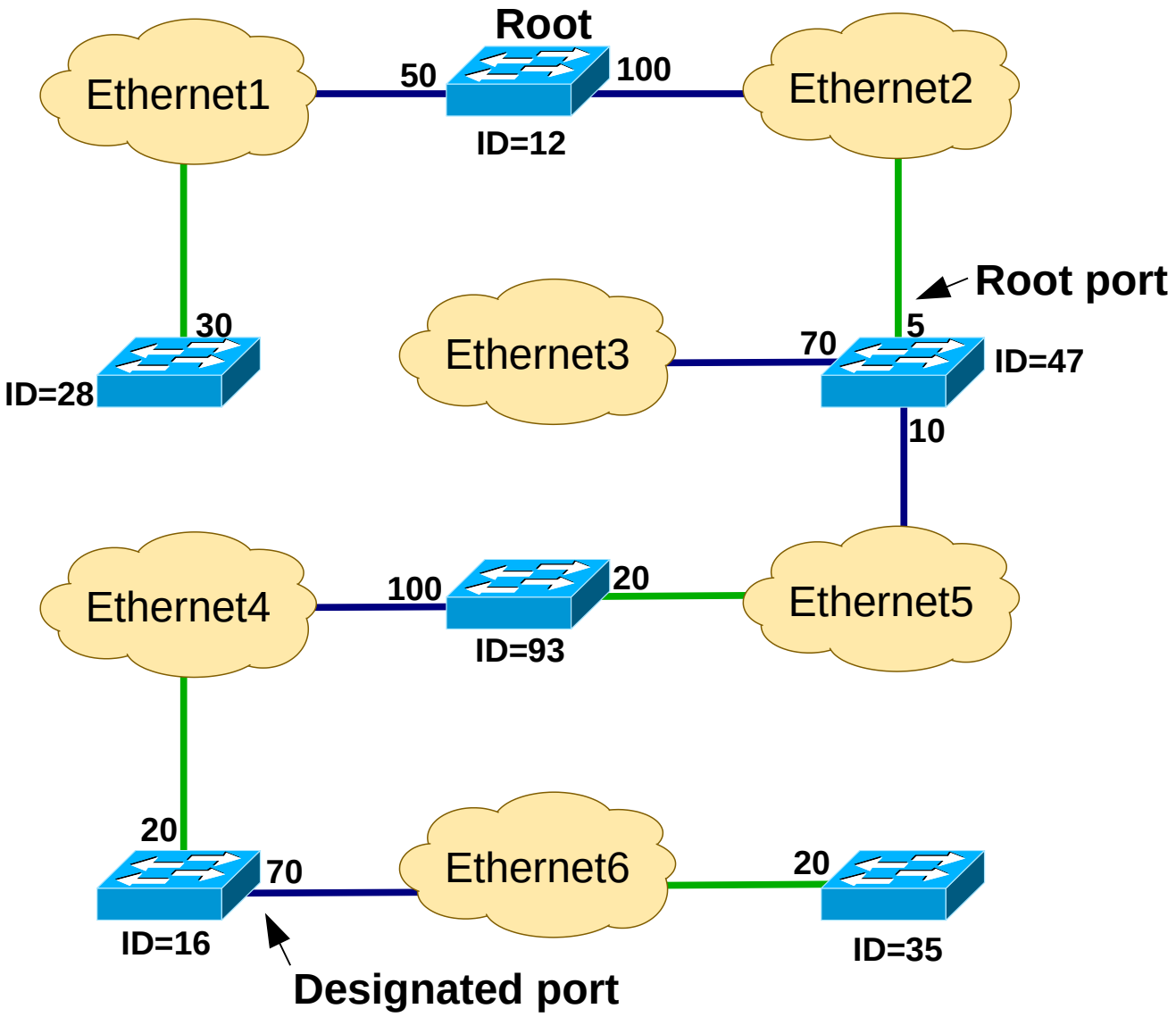
Eth1	12
Eth 2	12
Eth 3	47
Eth 4	93
Eth 5	47
Eth 6	16



Example – Spanning Tree (2)

Designated bridges

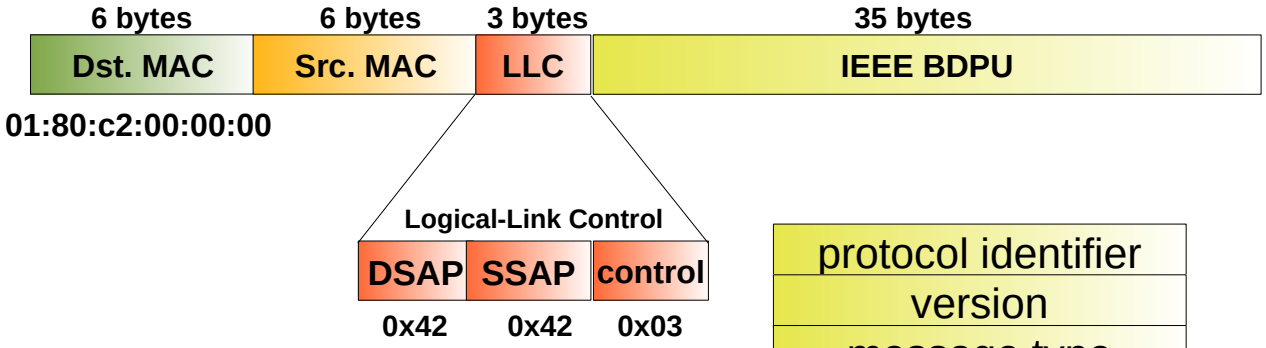
Eth1	12
Eth 2	12
Eth 3	47
Eth 4	93
Eth 5	47
Eth 6	16



Protocolo IEEE 802.1D
BPDUs (Bridge Protocol Data Units)

- To build the spanning tree, switches exchange special messages between them called Bridge Protocol Data Units (BPDU).
- There are two types: *Configuration* e *Topology Change Notification.*

IEEE 802.3 Ethernet
Destination: 01:80:c2:00:00:00 (01:80:c2:00:00:00)
Source: 00:16:e0:9a:c3:92 (00:16:e0:9a:c3:92)
Length: 39
Logical-Link Control
DSAP: Spanning Tree BPDU (0x42)
SSAP: Spanning Tree BPDU (0x42)
Control field: U, func=UI (0x03)
Spanning Tree Protocol
Protocol Identifier: Spanning Tree Protocol (0x0000)
Protocol Version Identifier: Spanning Tree (0)
BPDU Type: Configuration (0x00)
Root ID: 32768 / 00:05:1a:4e:fd:58
Root Path Cost: 200004
Bridge ID: 32768 / 00:16:e0:9a:c3:80
Port ID: 0x8012
Message Age: 1
Max Age: 20
Hello Time: 2
Forward Delay: 15



protocol identifier		
version		
message type		
TCA	reserved	TC
root ID		
root path cost		
bridge ID		
port ID		
message age		
max age		
hello time		
forward delay		

Configuration BPDU

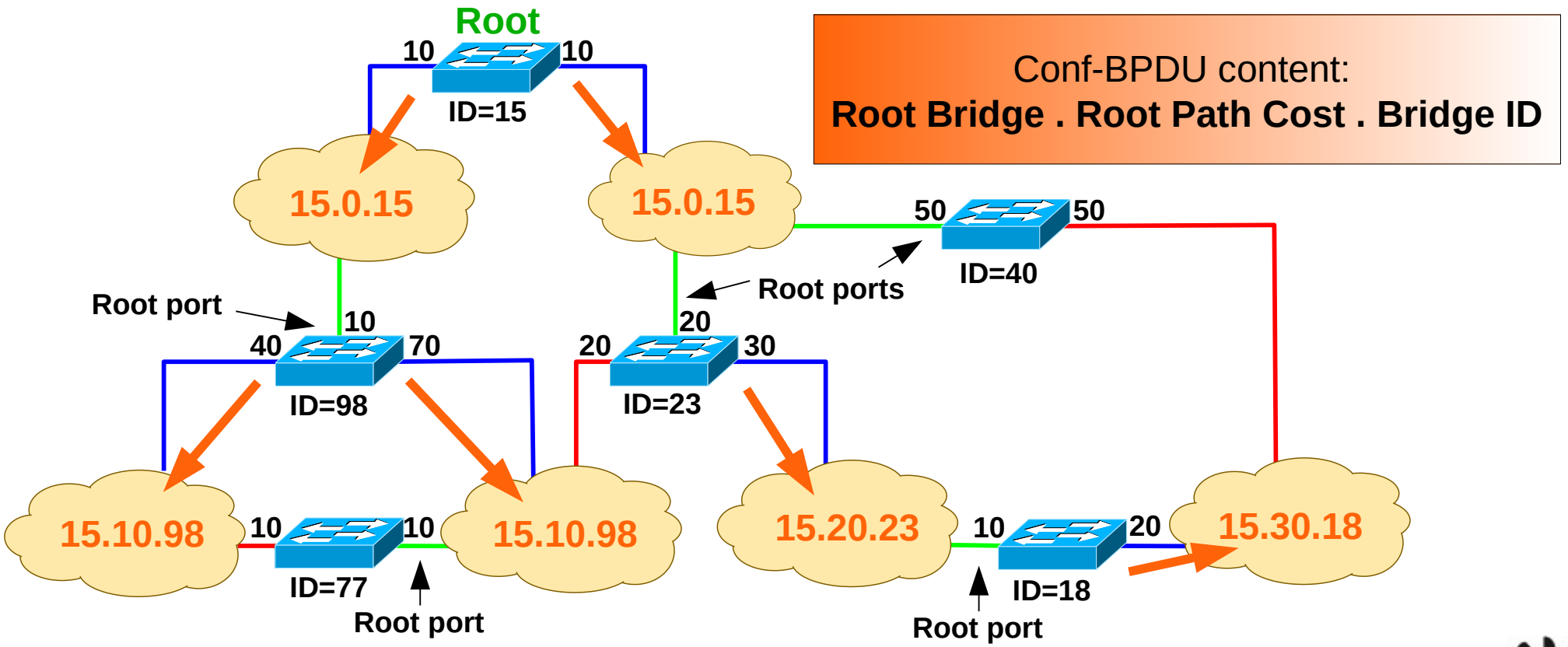
- The setup of the Spanning Tree id dune using Conf - BPDU (configuration messages).

IEEE 802.3 Ethernet
Destination: 01:80:c2:00:00:00 (01:80:c2:00:00:00)
Source: 00:16:e0:9a:c3:92 (00:16:e0:9a:c3:92)
Length: 39
Logical-Link Control
DSAP: Spanning Tree BPDU (0x42)
SSAP: Spanning Tree BPDU (0x42)
Control field: U, func=UI (0x03)
Spanning Tree Protocol
Protocol Identifier: Spanning Tree Protocol (0x0000)
Protocol Version Identifier: Spanning Tree (0)
BPDU Type: Configuration (0x00)
Root ID: 32768 / 00:05:1a:4e:fd:58
Root Path Cost: 200004
Bridge ID: 32768 / 00:16:e0:9a:c3:80
Port ID: 0x8012
Message Age: 1
Max Age: 20
Hello Time: 2
Forward Delay: 15

- More relevant fields:
  - Root ID: ID of the current root bridge.
  - Root Path Cost: estimation of the cost to the root.
  - Bridge ID: own bridge identifier.
  - Port ID: identifier of the port by which the BPDU was sent.
    - Port priority (1 byte) + Port number

Spanning Tree Maintenance

- Periodically switches sent Conf-BPDUs by its Designated Ports.
  - Periodicity of Conf-BPDU messages = hello time
    - Recommended Hello time: 2 seconds.
    - Defined at the root bridge.



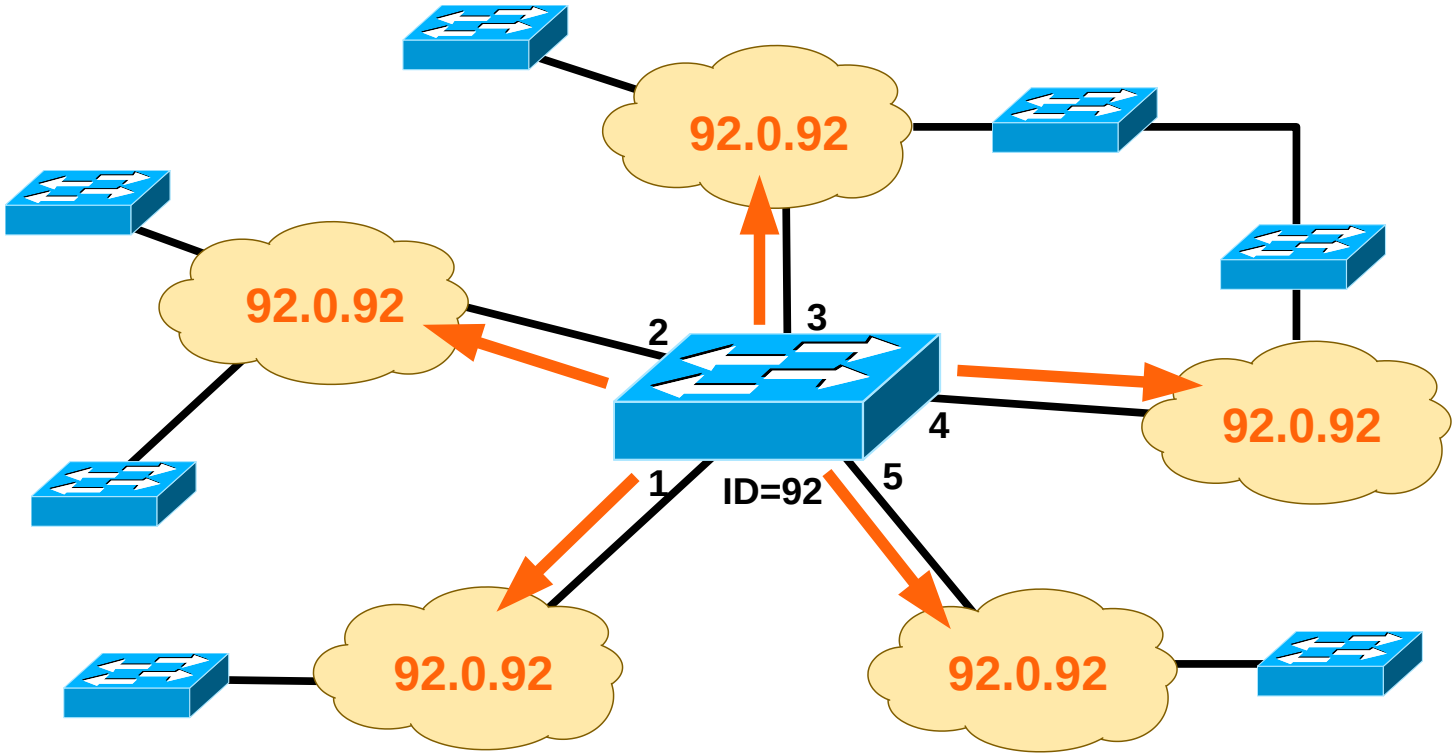
Sorting of Best BPDU

- A Conf-BPDU C1 is considered better than a Conf-BPDU C2 if:
  - The Root ID of C1 is lower than the one in C2,
  - With equal Root ID, if Root Path Cost of C1 is lower than the one in C2,
  - With equal Root ID and Root Path Cost, if the Bridge ID of C1 is lower than the one in C2,
  - With equal Root ID, Root Path Cost and Bridge ID, if the Port ID of C1 is lower than the one in C2.

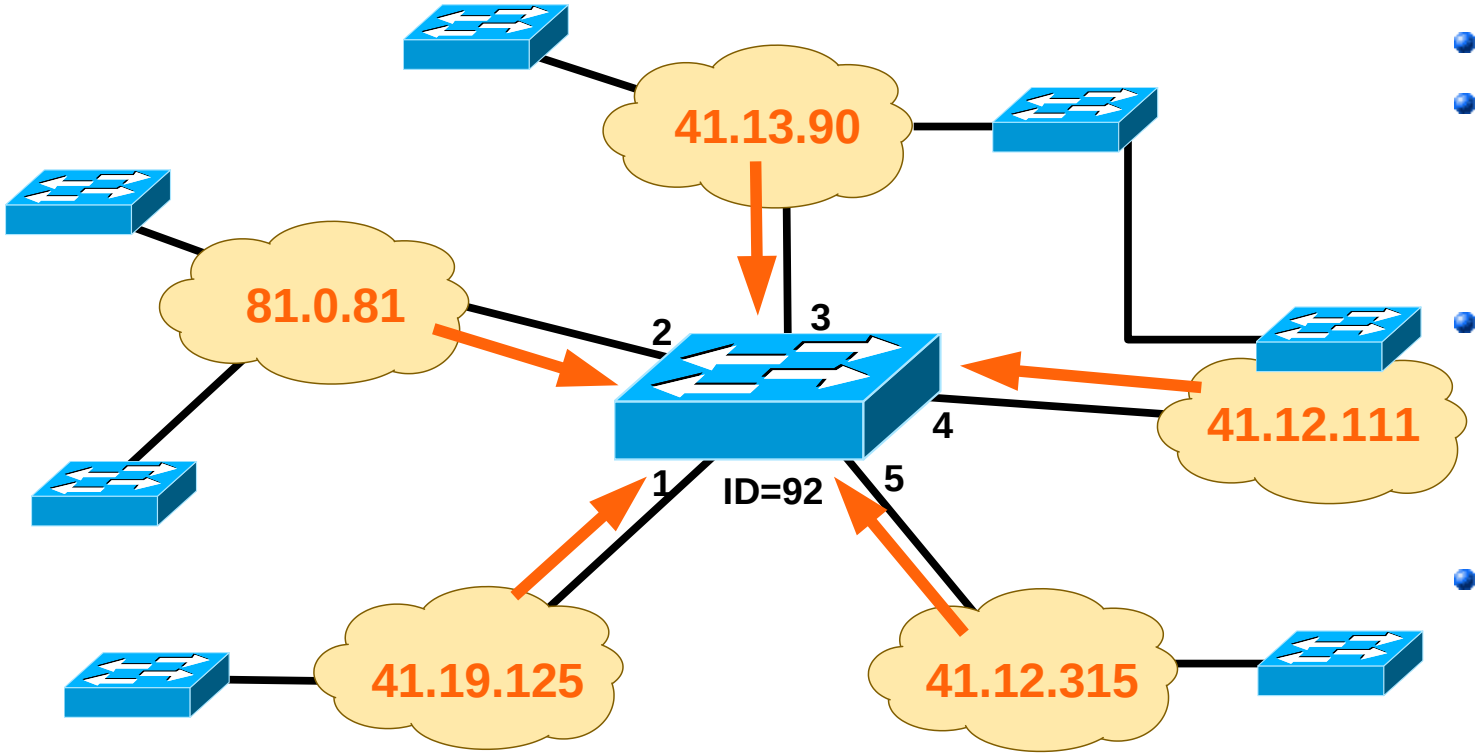
Root ID	Root Path Cost	Bridge ID	Port ID
18	27	32	2
18	27	32	4
18	27	43	1
18	35	23	3
23	31	45	2

Building the Spanning Tree (1)

- Each switch initially assumes to be the Root Bridge.
  - Assumes Root Path Cost = 0,
  - Sends Conf-BPDU to all its ports.



Building the Spanning Tree (2)



- Bridge92 is not root (BridgeID 92>41)
- Bridge 92 Root Port is 4.
  - Lowest RootID (41).
  - Lowest Root Path Cost (12+1=13).
  - Lowest Neighbor BridgeID (111<315)
- Bridge 92 is Designated Bridge via ports 1 and 2
  - Port 2, Lowest RootID (41).
  - Port 1, Same RootID (41) and Lowest Root Path Cost (13<19).
- Bridge 92 ports 3 and 5 are blocked.
  - Neighbors have the same RootID (41).
  - Via port 3, Neighbor has the same Root Path Cost (13), but lower BridgeID (90<92).
  - Via port 5, Neighbor has lower Root Path Cost (12).

Best Conf-BPDU received by Bridge 92 (until now)

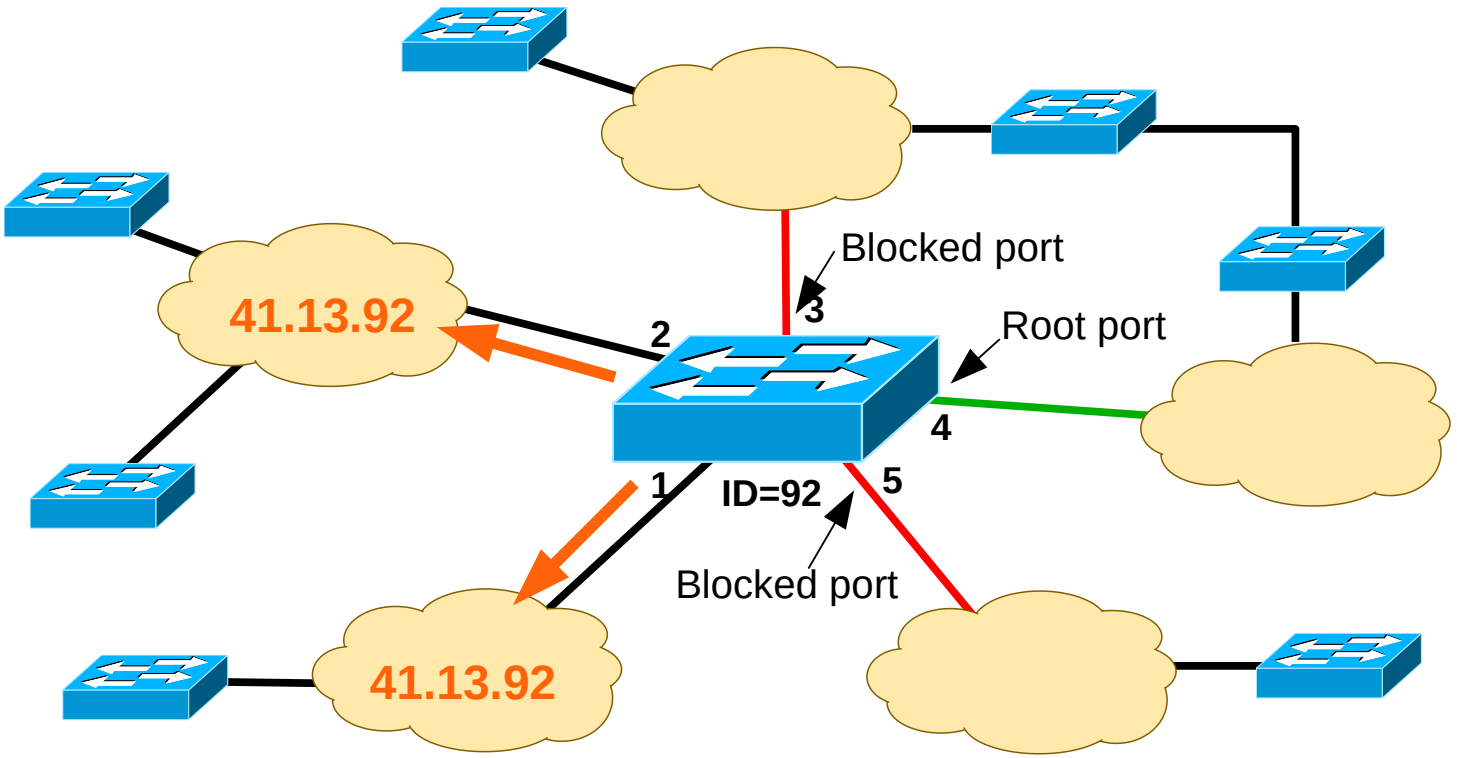
Estimations of Bridge 92 (assuming port costs equal to 1).

Root Bridge = 41

Root port = 4

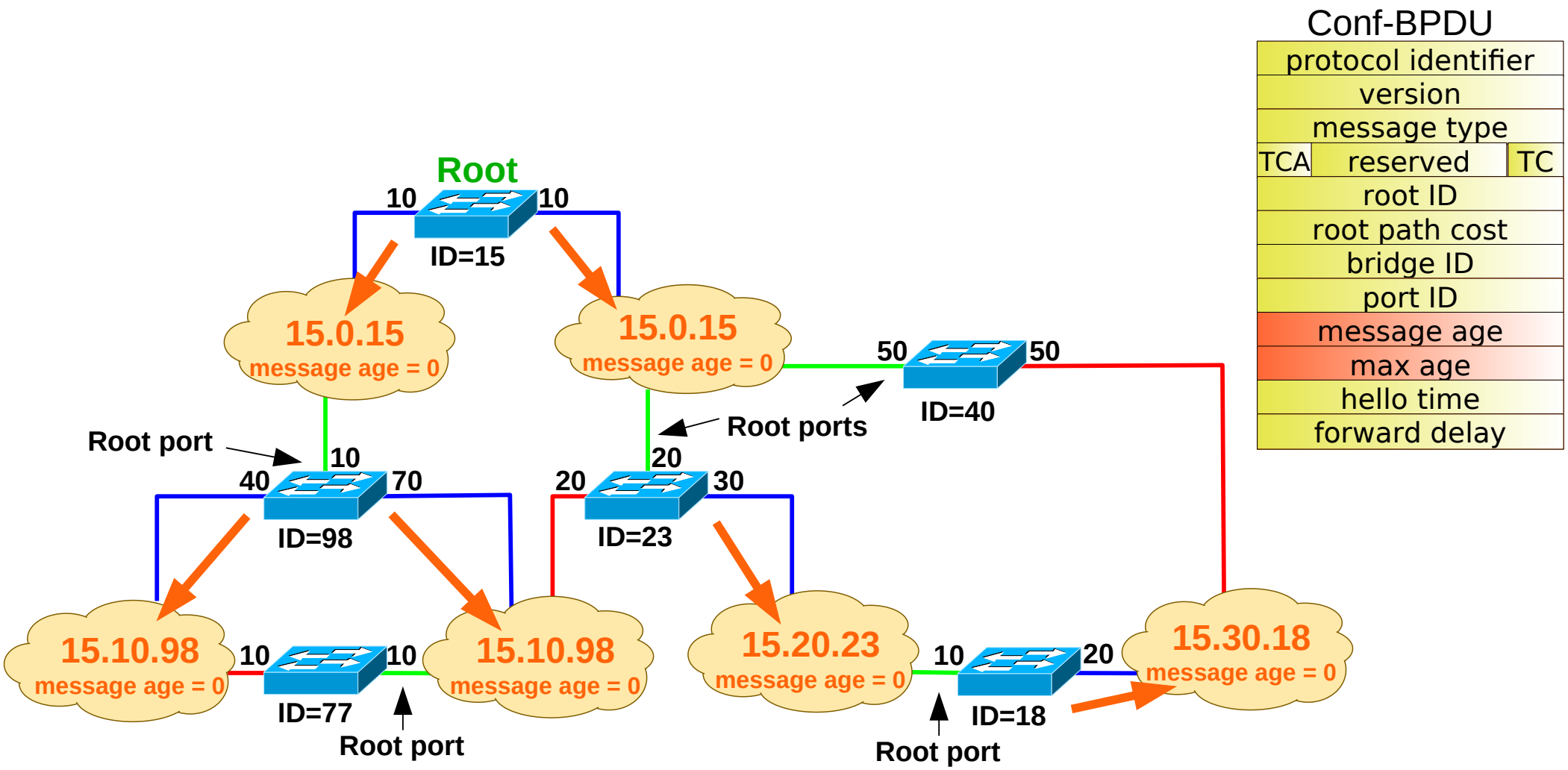
Root Path Cost = 12 + 1 **=13**

Building the Spanning Tree (3)



Conf-BPDU sent by Bridge 92 - 41.13.92

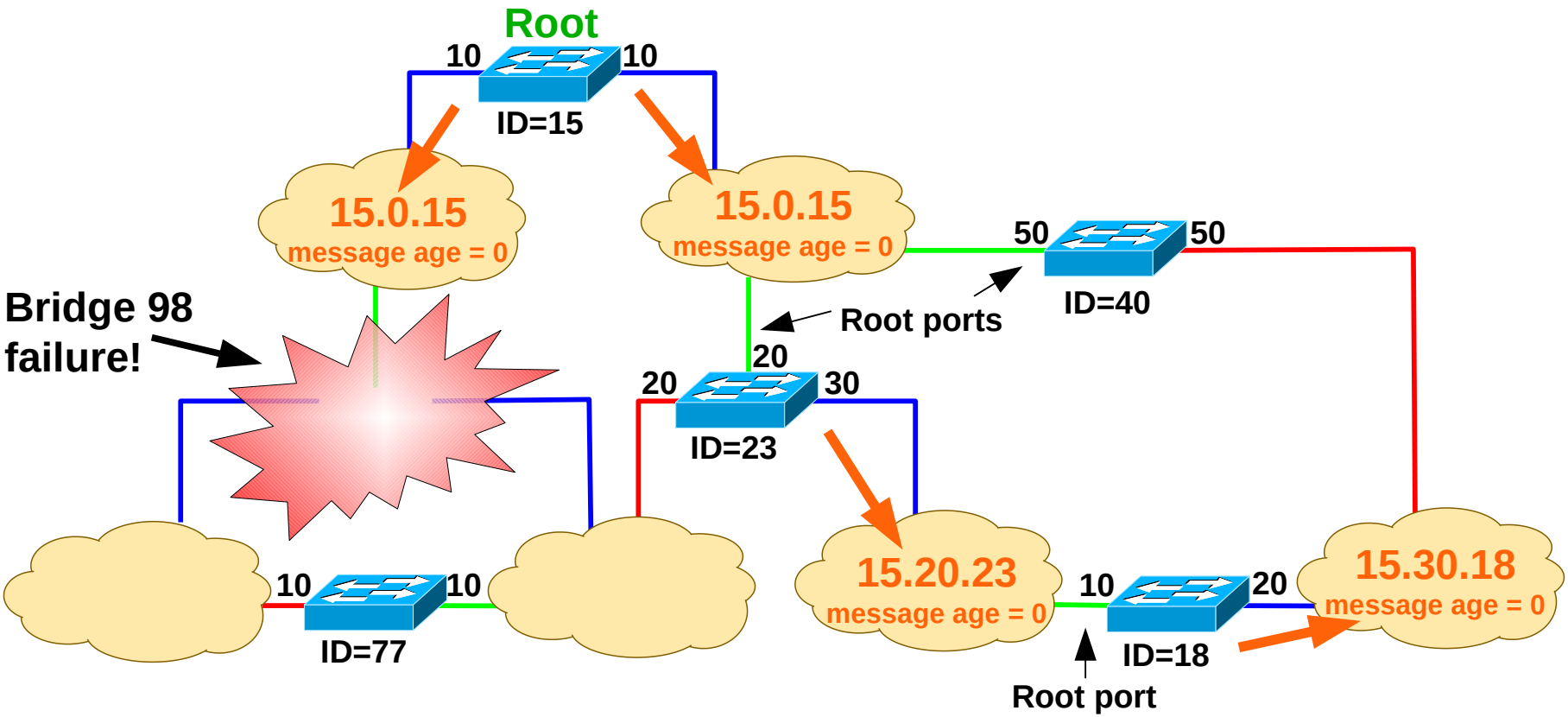
Network Failures (1)



Conf-BPDU

protocol identifier		
version		
message type		
TCA	reserved	TC
root ID		
root path cost		
bridge ID		
port ID		
message age		
max age		
hello time		
forward delay		

Network Failures (2)

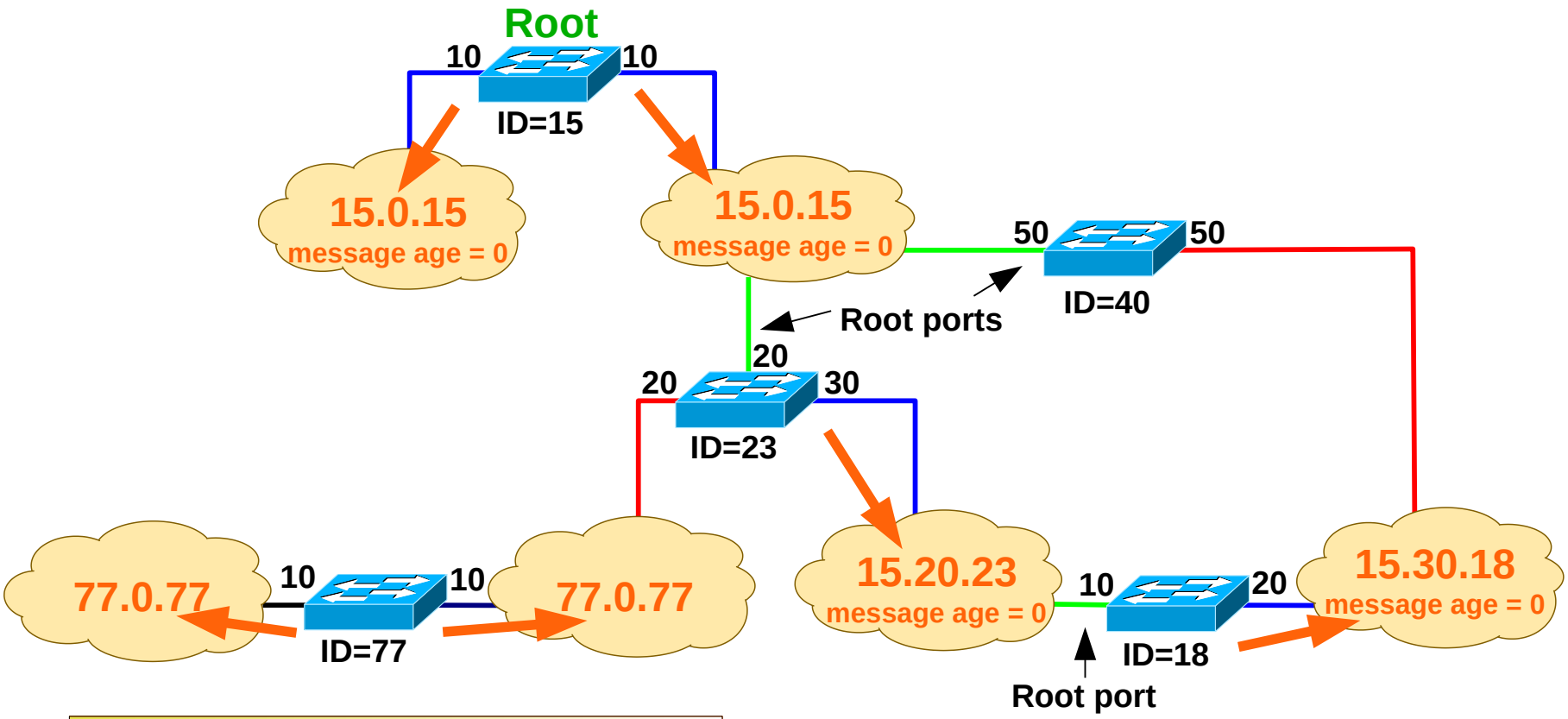


15.10.98 age = 0
15.10.98 age = 5
15.10.98 age = 10
.....
15.10.98 age = max age

15.10.98 age = 0
15.10.98 age = 5
15.10.98 age = 10
.....
15.10.98 age = max age

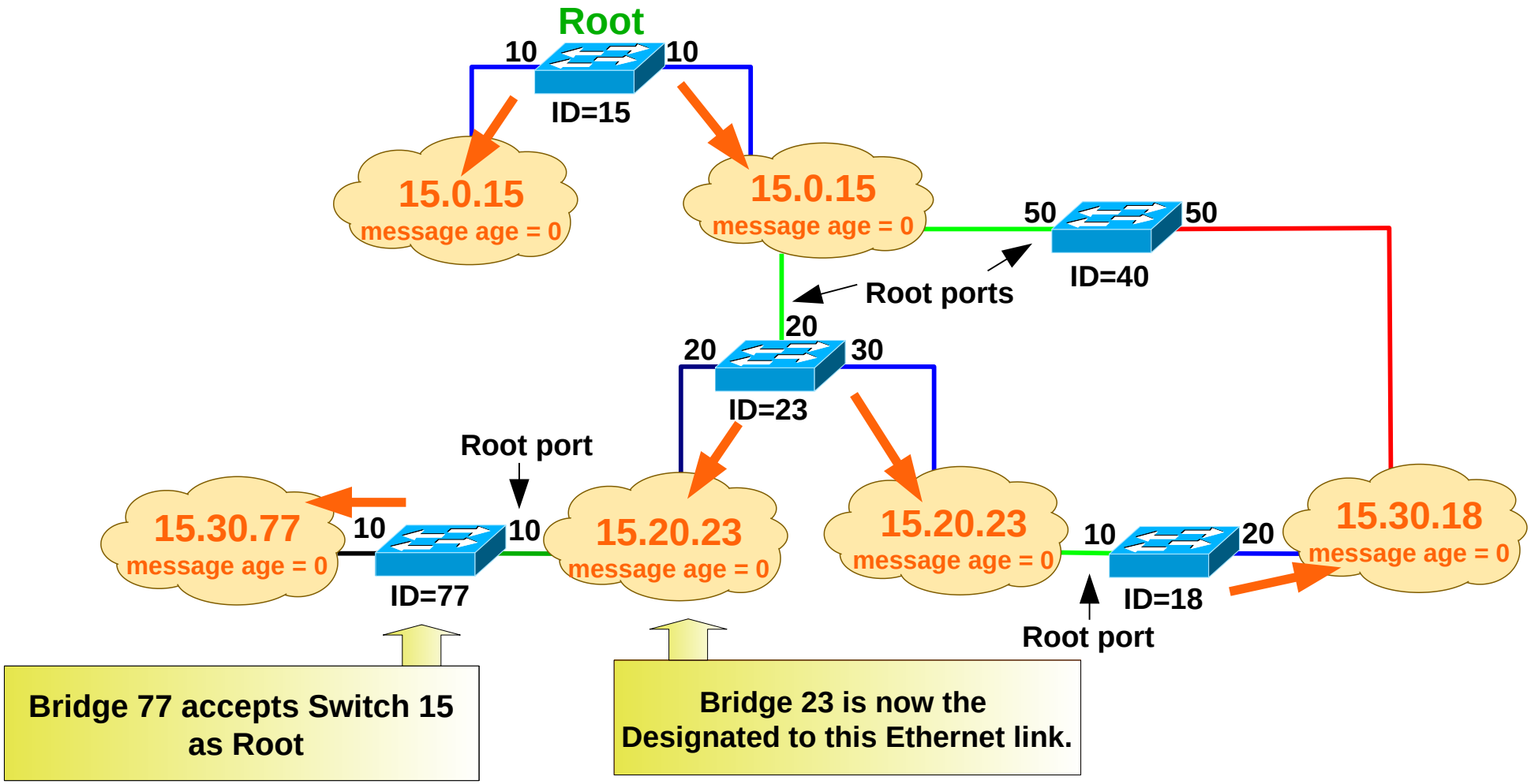
max age = 20 seconds

Network Failures (3)



Bridge 77 assumes to be Root

Network Failures (4)



Forwarding Tables Entries Lifetimes

- Forwarding Tables Long Lifetime – Many frames will be lost when network is changing topology.
- Forwarding Tables Short Lifetime – Creates to much traffic due to frequent flooding.

- There are two forwarding tables lifetimes:
  - **Long**: used by default (recommended value = 300 seconds)
  - **Short:** used when SPT is re-configuring (recommended value = 15 seconds)

Topology Change Notification

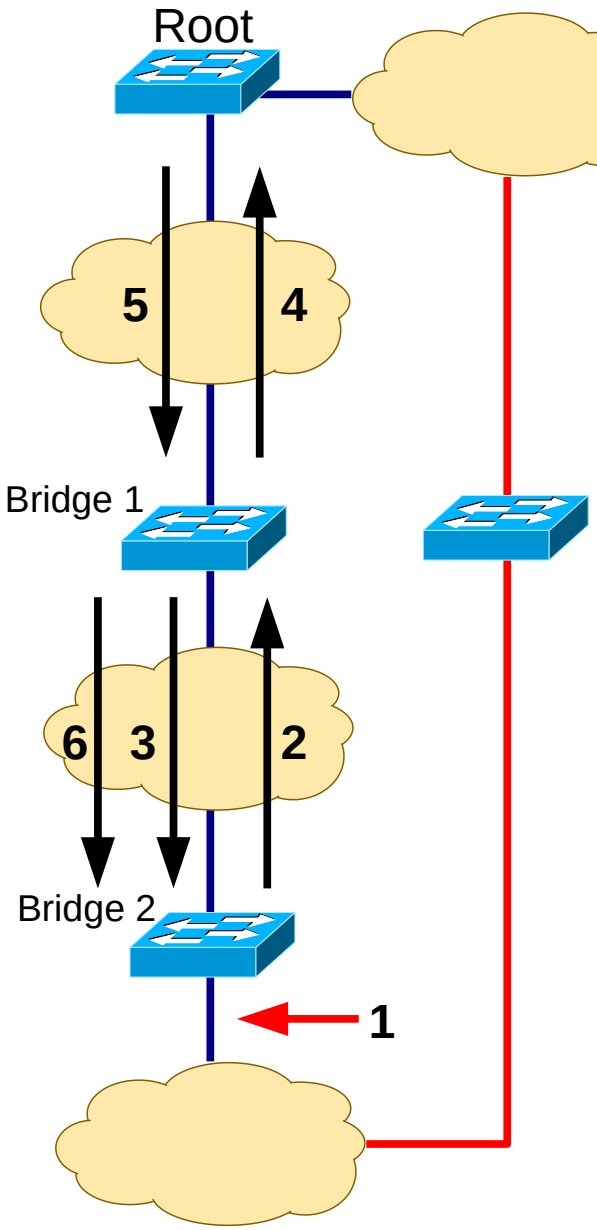
Conf (Configuration) BPDU

protocol identifier		
version		
message type = 0		
TCA	reserved	TC
root ID		
root path cost		
bridge ID		
port ID		
message age		
max age		
hello time		
forward delay		

TCA - flag Topology Change Acknowledgment
TC - flag Topology Change

TCN (Topology Change Notification) BPDU

protocol identifier
version
message type = 1



1. **Port changes state to disabled or blocking**
2. **Sends TCN-BPDU (periodicity = hello time)**
3. **Sends Conf-BPDU with TCA = 1 while receiving TCN-BPDU**
4. **Sends TCN-BPDU (periodicity = hello time)**
5. **Sends Conf-BPDU with TCA = 1 while receiving TCN-BPDU and with TC=1 for a period of time equal to *ForwardDelay* + *MaxAge***

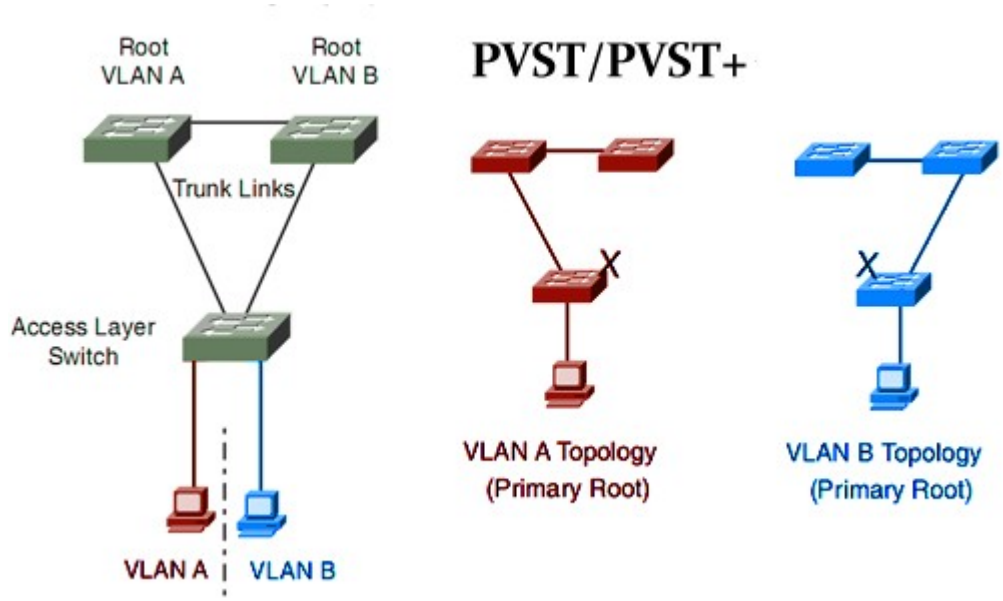
   Root bridge uses the forwarding table short lifetime during this period
6. **Sends Conf-BPDU with TC=1**

   Bridge 1 uses the forwarding table short lifetime while receiving Conf-BPDU with TC=1

   Bridge 2 uses the forwarding table short lifetime while receiving Conf-BPDU with TC=1

Other Protocols (1)

- Cisco's proprietary versions of SPT are:
  - Per-VLAN Spanning Tree (PVST).
  - Per-VLAN Spanning Tree Plus (PVST+).
- Create a different spanning tree for each VLAN.
  - Different roots, costs, blocked ports, etc...
  - In a complex switching network some switches may not have ports of all VLAN.




```
Ethernet II, Src: c2:00:05:7f:f1:01 (c2:00:05:7f:f1:01), Dst: PVST+ (01:00:0c:cc:cc:cd)
802.1Q Virtual LAN, PRI: 0, CFI: 0, ID: 1
   000. .... .... .... = Priority: 0
   ...0 .... .... .... = CFI: 0
   .... 0000 0000 0001 = ID: 1
   Length: 50
Logical-Link Control
Spanning Tree Protocol
   Protocol Identifier: Spanning Tree Protocol (0x0000)
   Protocol Version Identifier: Spanning Tree (0)
   BPDU Type: Configuration (0x00)
▷  BPDU flags: 0x00
▷  Root Identifier: 32768 / 0 / c2:00:05:7f:00:00
   Root Path Cost: 0
▷  Bridge Identifier: 32768 / 0 / c2:00:05:7f:00:00
   Port identifier: 0x802a
   Message Age: 0
   Max Age: 20
   Hello Time: 2
```

Identificador da VLAN

Other Protocols (2)

- IEEE 802.1p
  - Extension of IEEE 802.1Q.
  - Provides QoS based on relative priorities.
  - Defines the field *User Priority* (3 bits) that allows 8 levels of priority.
  - The standard recommends:
    - Priority 7 : Critical traffic,
    - Priorities 5–6 : Delay sensitive traffic (voice and live video),
    - Priorities 1–4 : Delay variation sensitive traffic (*streaming*),
    - Priority 0 : Other traffic.

Other Protocols (3)

- IEEE 802.1w Rapid Spanning Tree Protocol
  - Extension of IEEE 802.1D.
  - Speeds up the convergence time of the Spanning Tree in case of topology changes
    - There are only three port states in RSTP that correspond to the three possible operational states.
    - Adds two additional port roles to a port when in blocking state
      - Alternate port: possible alternative Root port.
      - Backup port: possible alternative Designated port.
  - Adds a negotiated mechanism between switches.
    - Uses the reserved bits in the Conf-BPDU.

Conf (Configuration) BPDU

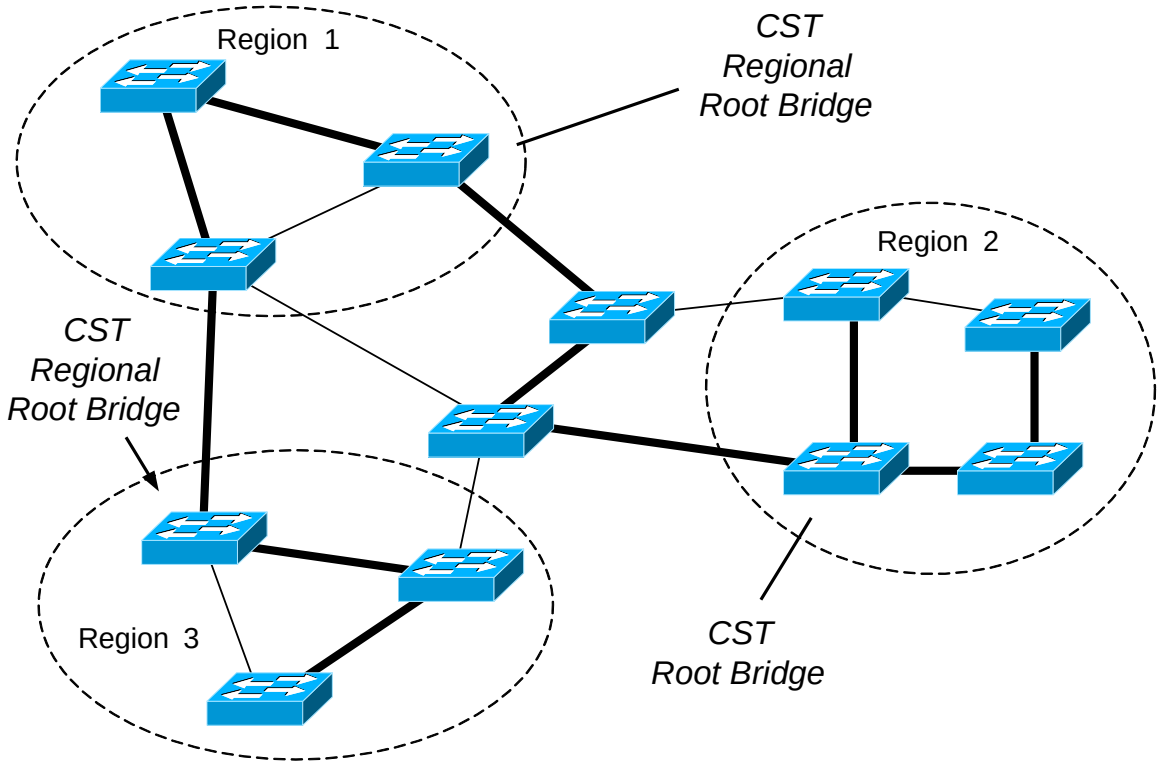
protocol identifier		
version		
message type = 0		
TCA	reserved	TC
root ID		
root path cost		
bridge ID		
port ID		
message age		
max age		
hello time		
forward delay		

STP (802.1D) Port State	RSTP (802.1w) Port State	Is Port Included in Active Topology?	Is Port Learning MAC Addresses?
Disabled	Discarding	No	No
Blocking	Discarding	No	No
Listening	Discarding	Yes	No
Learning	Learning	Yes	Yes
Forwarding	Forwarding	Yes	Yes

Other Protocols (4)

- IEEE 802.1s Multiple Spanning Tree Protocol
  - Creates multiple Spanning Trees.

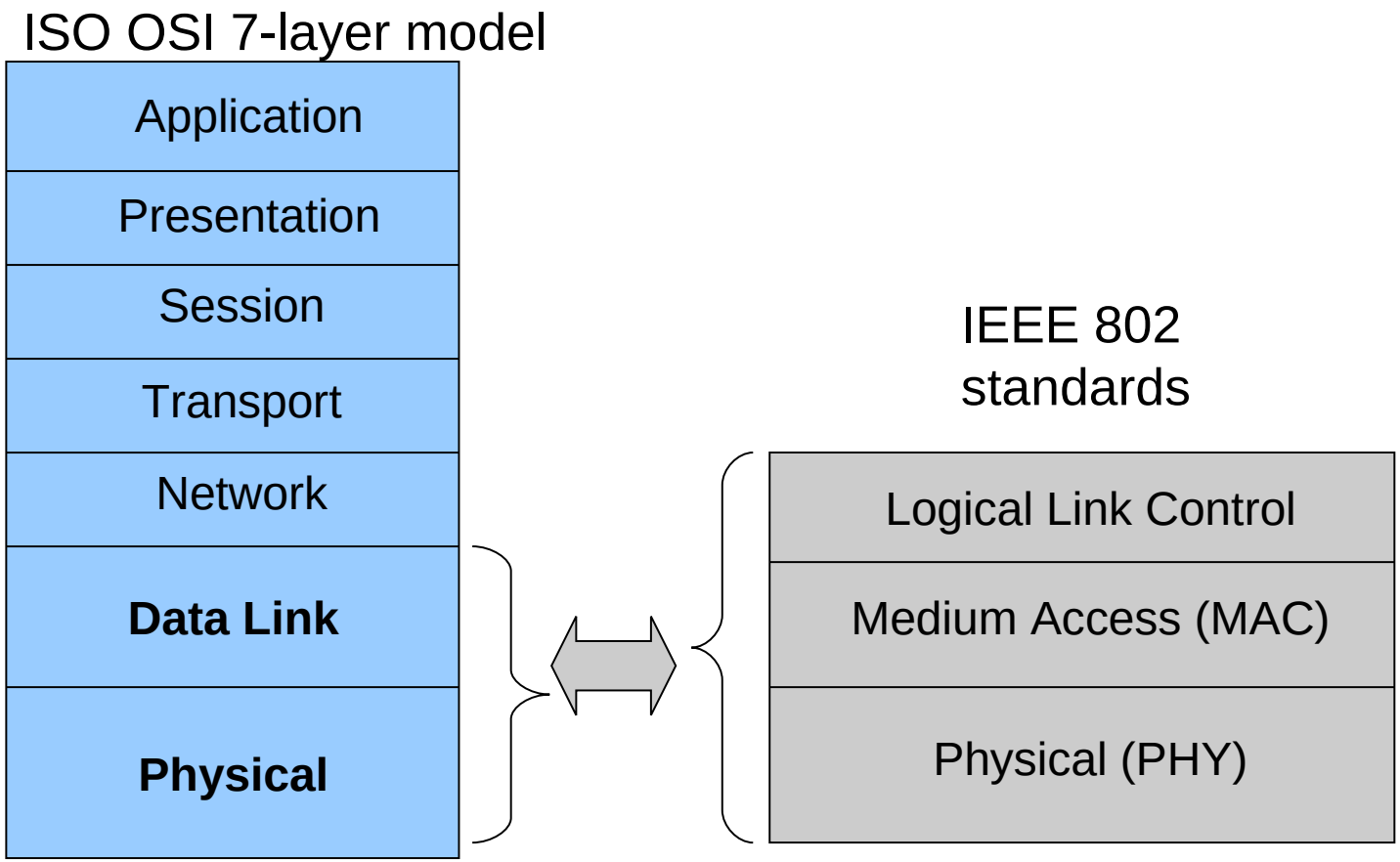
    Allows the assignment of a set of several VLAN to a specific Common Spanning Tree (CST).
    - CST are usually mapped to regions of the network.



Wi-Fi

Standardization of Wireless Networks

- Wireless networks are standardized by the IEEE under the 802 LAN MAN standards committee.



Wireless Networks

- Networks are designed according to the number of users and coverage area
- There are several scales on the number of users and coverage area
  - **Local: LANs → IEEE 802.11**
  - Personal: PANs → e.g. Bluetooth, ZigBee
  - Regional: WANs → GSM, UMTS, LTE, 5G, LoRa,...
  - Worldwide : Satellite → Iridium, SpaceX Starlink?

Wireless LAN: Oveview

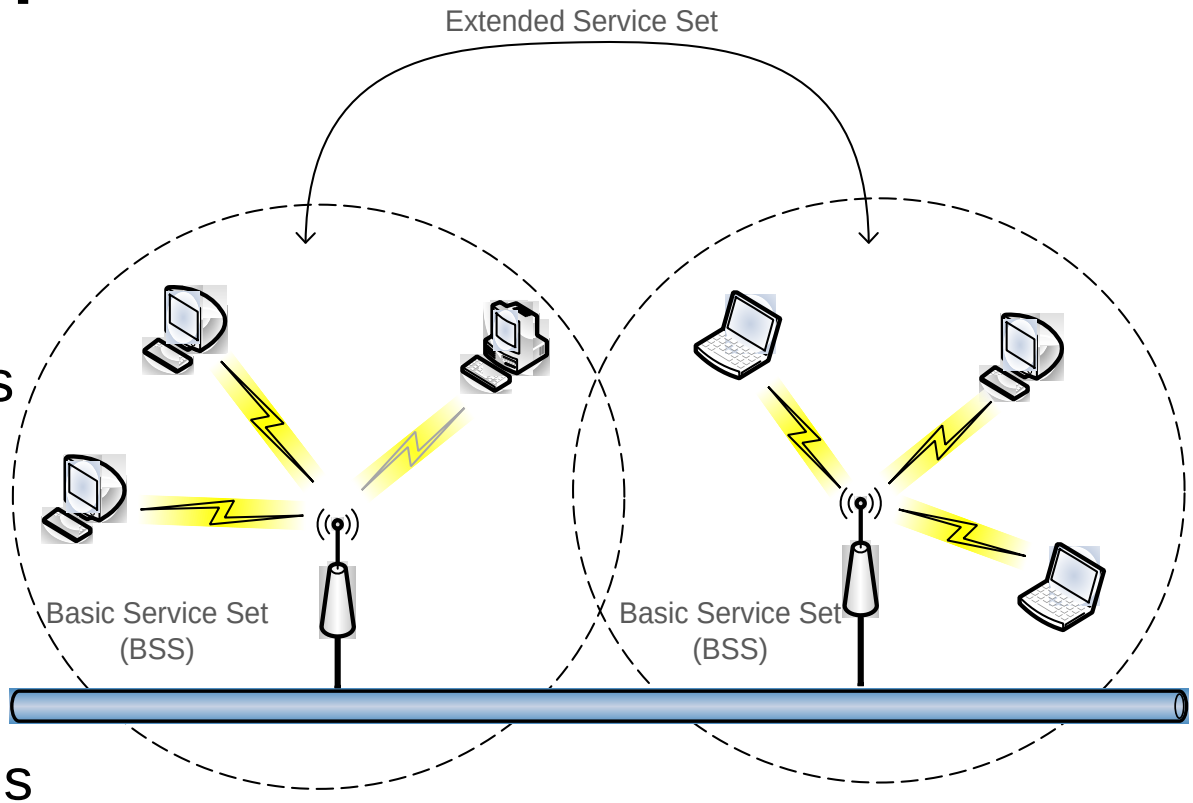
- Two Types
  - Infra-structured,
  - Ad-hoc.
- Advantages
  - Flexible installation (minimum cables).
  - More robust (no cable problems).
  - One-time installation (conferences, historic buildings).
- Problems
  - Many proprietary solutions.
  - Restrictions on the electromagnetic spectrum.
  - Subject to frame collision when accessing the transmission medium.
    - More on this later.
  - Lower bandwidths than cabled networks.

Evolution of WLAN standards

- WiFi 1 - 802.11b, 1999, 2.4 GHz band, 11 Mbps data rate
- WiFi 2 - 802.11a, 1999, 5 GHz band, 54 Mbps data rate
- WiFi 3 - 802.11g, 2003, 2.4 GHz band, 54 Mbps data rate
- WiFi 4 - 802.11n, 2009, 2.4 and 5 GHz bands, ~600 Mbps data rate
- WiFi 5 - 802.11ac, 2013, 5 GHz band, ~1.3 Gbps data rate
- WiFi 6 - 802.11ax, 2019, 1 to 7GHz bands, >11Gbps data rate

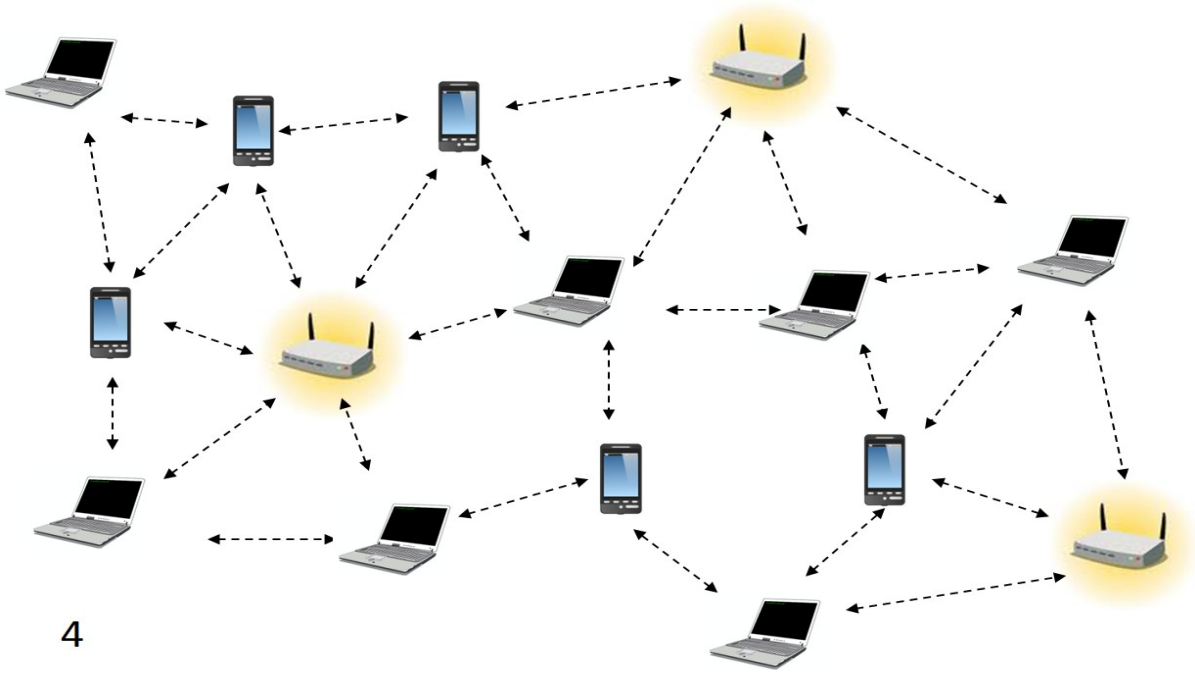
Components

- Station (STA)
  - Mobile terminal
- Access Point (AP)
  - STA connect to access points (infra-structured networks)
- Basic Service Set (BSS)
  - STA and AP with same coverage form a BSS
  - Group of IEEE 802.11 stations associated to an Access Point (AP)
  - Known through the SSID
- Extended Service Set (ESS)
  - Several BSSs interconnected by APs form a ESS



Ad-hoc Networks (IBSS)

- Temporary set of stations
- Forming an ad-hoc network – an independent BSS (IBSS), means that there is no connection to a wired network
- No AP
- No relay function (direct connection)
- Simple setup

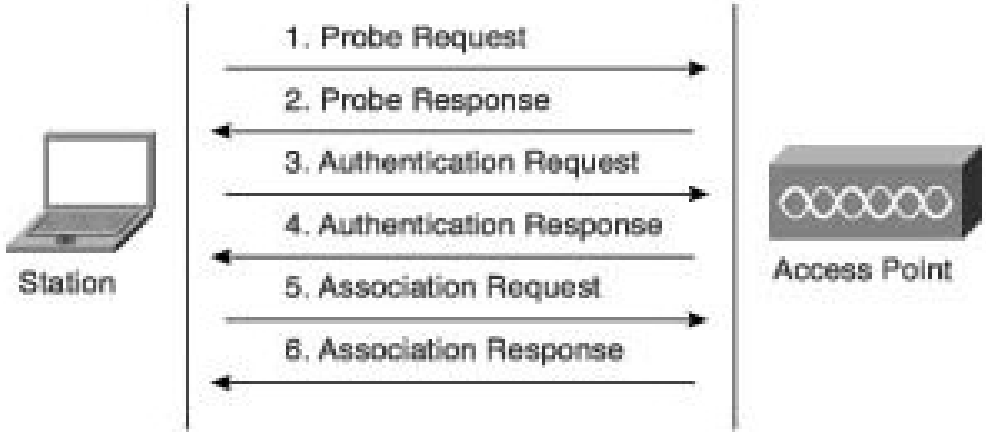


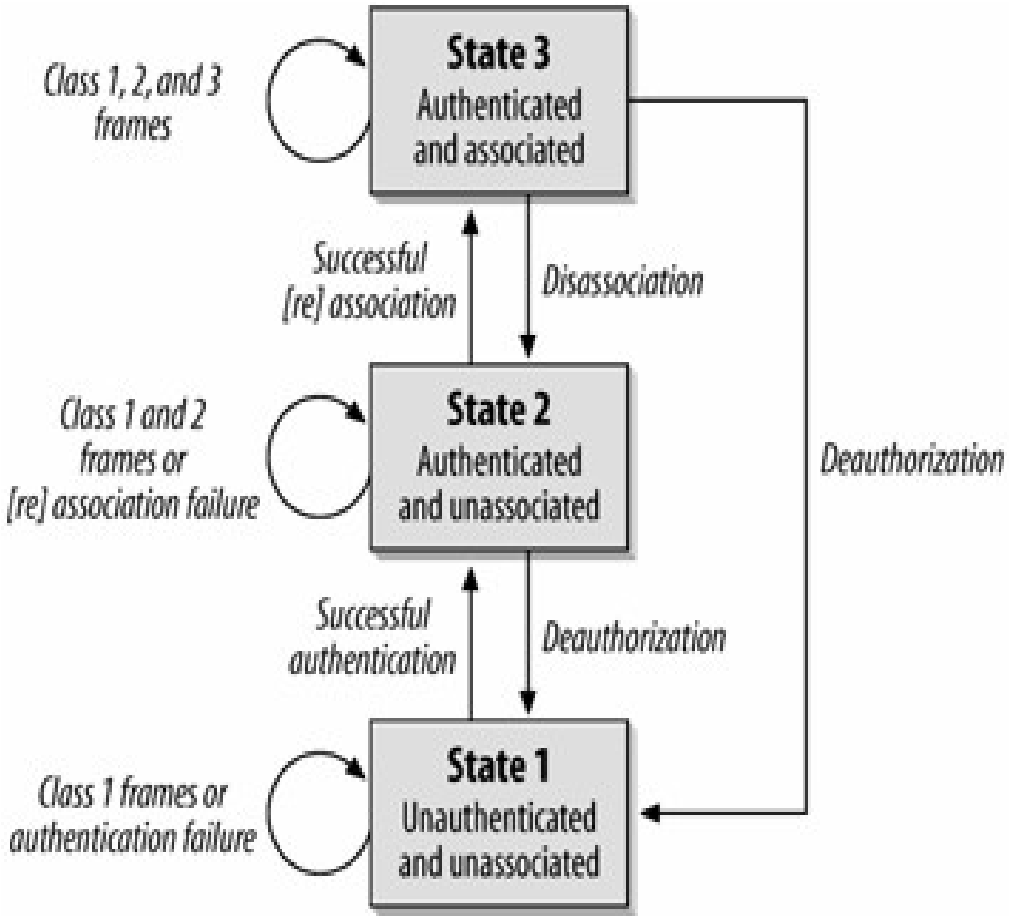
IEEE 802.11 services

- Station services (similar to wired network)
  - Authentication (login)
  - De-authentication (logout)
  - Privacy
  - Data delivery
- Distribution services
  - Association
    - Make logical connection between the AP and the station – the AP will not receive any data from a station before association
  - Re-association (similar to association)
    - Send repeatedly to the AP.
    - Help the AP to know if the station has moved from/to another BSS.
    - After Power Save
  - Disassociation
    - Manually disconnect (PC is shutdown or adapter is ejected)

Joining a BSS

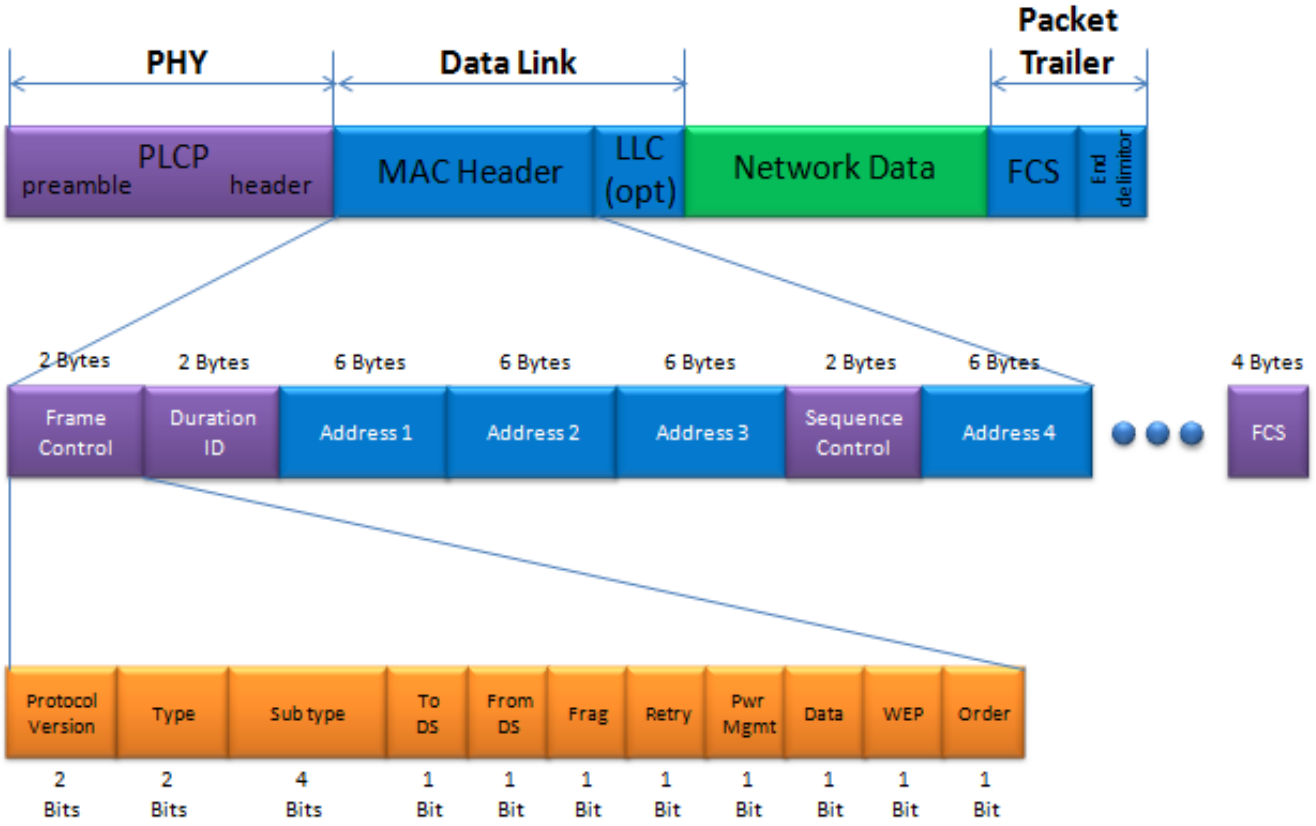
- Station finds BSS/AP by **Scanning/Probing**.
- BSS with AP: both **Authentication** and **Association** are necessary for joining a BSS.





WLAN Frames

- Three types of frames
  - Control: RTS, CTS, ACK
  - Management
  - Data
- Header is different for the different types of frames.



Joining BSS with AP: Scanning

- A station willing to join a BSS must get in contact with the AP. This can happen through:
- 1. Passive scanning
  - The station scans the channels for a Beacon frame that is sent periodically from an AP to announce its presence and provide the SSID, and other parameters for WNICs within range
- 2. Active scanning (the station tries to find an AP)
  - The station sends a Probe Request frame - Sent from a station when it requires information from another station
  - All AP's within reach reply with a Probe Response frame - Sent from an AP containing capability information, supported data rates, etc., after receiving a probe request frame

Beacon Frame

```
▾ IEEE 802.11 Beacon frame, Flags: ........C
    Type/Subtype: Beacon frame (0x0008)
  ▸ Frame Control Field: 0x8000
    .000 0000 0000 0000 = Duration: 0 microseconds
    Receiver address: Broadcast (ff:ff:ff:ff:ff:ff)
    Destination address: Broadcast (ff:ff:ff:ff:ff:ff)
    Transmitter address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    Source address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    BSS Id: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    .... .... .... 0000 = Fragment number: 0
    1001 1000 1010 .... = Sequence number: 2442
    Frame check sequence: 0x6f0b825c [unverified]
    [FCS Status: Unverified]
▾ IEEE 802.11 wireless LAN
  ▾ Fixed parameters (12 bytes)
      Timestamp: 660070796
      Beacon Interval: 0.102400 [Seconds]
    ▸ Capabilities Information: 0x0421
  ▾ Tagged parameters (123 bytes)
    ▸ Tag: SSID parameter set: LABCOM
    ▸ Tag: Supported Rates 1(B), 2(B), 5.5(B), 6, 9, 11(B), 12, 18, [Mbit/sec]
    ▸ Tag: DS Parameter set: Current Channel: 13
    ▸ Tag: Traffic Indication Map (TIM): DTIM 0 of 0 bitmap
    ▸ Tag: ERP Information
    ▸ Tag: Extended Supported Rates 24, 36, 48, 54, [Mbit/sec]
    ▸ Tag: Cisco CCX1 CKIP + Device Name
    ▸ Tag: Vendor Specific: Microsoft Corp.: WMM/WME: Parameter Element
    ▸ Tag: Vendor Specific: Cisco Systems, Inc.: Aironet Unknown (1) (1)
    ▸ Tag: Vendor Specific: Cisco Systems, Inc.: Aironet CCX version = 5
    ▸ Tag: Vendor Specific: Cisco Systems, Inc.: Aironet Unknown (11) (11)
    ▸ Tag: Vendor Specific: Cisco Systems, Inc.: Aironet Client MFP Disabled
```

Probe Request/Response Frames

```
▾ IEEE 802.11 Probe Request, Flags: ........C
    Type/Subtype: Probe Request (0x0004)
  ▸ Frame Control Field: 0x4000
    .000 0000 0000 0000 = Duration: 0 microseconds
    Receiver address: Broadcast (ff:ff:ff:ff:ff:ff)
    Destination address: Broadcast (ff:ff:ff:ff:ff:ff)
    Transmitter address: Microsof_0a:43:e3 (c0:33:5e:0a:43:e3)
    Source address: Microsof_0a:43:e3 (c0:33:5e:0a:43:e3)
    BSS Id: Broadcast (ff:ff:ff:ff:ff:ff)
    .... .... .... 0000 = Fragment number: 0
    1100 1011 0001 .... = Sequence number: 3249
    Frame check sequence: 0xc7056d0a [unverified]
    [FCS Status: Unverified]
▾ IEEE 802.11 wireless LAN
  ▾ Tagged parameters (62 bytes)
    ▸ Tag: SSID parameter set: TD_WIFI_GUEST
    ▸ Tag: Supported Rates 1, 2, 5.5, 6, 9, 11, 12, 18, [Mbit/sec]
    ▸ Tag: DS Parameter set: Current Channel: 13
    ▸ Tag: HT Capabilities (802.11n D1.10)
    ▸ Tag: Extended Supported Rates 24, 36, 48, 54, [Mbit/sec]
```

```
▾ IEEE 802.11 Probe Response, Flags: ........C
    Type/Subtype: Probe Response (0x0005)
  ▸ Frame Control Field: 0x5000
    .000 0001 0011 1010 = Duration: 314 microseconds
    Receiver address: IntelCor_d2:98:58 (28:b2:bd:d2:98:58)
    Destination address: IntelCor_d2:98:58 (28:b2:bd:d2:98:58)
    Transmitter address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    Source address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    BSS Id: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    .... .... .... 0000 = Fragment number: 0
    1010 0010 1001 .... = Sequence number: 2601
    Frame check sequence: 0x80831320 [unverified]
    [FCS Status: Unverified]
▾ IEEE 802.11 wireless LAN
  ▾ Fixed parameters (12 bytes)
      Timestamp: 664064263
      Beacon Interval: 0.102400 [Seconds]
    ▸ Capabilities Information: 0x0421
  ▾ Tagged parameters (117 bytes)
    ▸ Tag: SSID parameter set: LABCOM
    ▸ Tag: Supported Rates 1(B), 2(B), 5.5(B), 6, 9, 11(B), 12, 18, [Mbit/sec]
    ▸ Tag: DS Parameter set: Current Channel: 13
    ▸ Tag: ERP Information
    ▸ Tag: Extended Supported Rates 24, 36, 48, 54, [Mbit/sec]
    ▸ Tag: Cisco CCX1 CKIP + Device Name
    ▸ Tag: Vendor Specific: Microsoft Corp.: WMM/WME: Parameter Element
    ▸ Tag: Vendor Specific: Cisco Systems, Inc.: Aironet Unknown (1) (1)
    ▸ Tag: Vendor Specific: Cisco Systems, Inc.: Aironet CCX version = 5
    ▸ Tag: Vendor Specific: Cisco Systems, Inc.: Aironet Unknown (11) (11)
    ▸ Tag: Vendor Specific: Cisco Systems, Inc.: Aironet Client MFP Disabled
```

Joining BSS with AP: Authentication

- Once an AP is found/selected, a station goes through authentication
- Open system authentication (default, 2-step process)
  - Station sends authentication frame with its identity
  - AP sends frame as an Ack / NAck
- Shared key authentication
  - Stations receive shared secret key through secure channel independent of 802.11
  - After the WNIC sends its initial authentication request, it will receive an authentication frame from the AP containing a challenge text
  - The WNIC sends an authentication frame containing the encrypted version of the challenge text to the AP.
  - The AP ensures the text was encrypted with the correct key by decrypting it with its own key.
  - The result of this process determines the WNIC's authentication status.

Authentication Frames

- Nowadays, WPA* secure networks use “Open System”.
- Non-”Open System” authentication was used for WEP protected networks (unsecured and functionally deprecated).

```
IEEE 802.11 Authentication, Flags: ........
  Type/Subtype: Authentication (0x000b)
  Frame Control Field: 0xb000
  .000 0001 0011 1010 = Duration: 314 microseconds
  Receiver address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
  Destination address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
  Transmitter address: D-LinkIn_6a:cc:6e (84:c9:b2:6a:cc:6e)
  Source address: D-LinkIn_6a:cc:6e (84:c9:b2:6a:cc:6e)
  BSS Id: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
  .... .... .... 0000 = Fragment number: 0
  0001 0100 1011 .... = Sequence number: 331
IEEE 802.11 wireless LAN
  Fixed parameters (6 bytes)
    Authentication Algorithm: Open System (0)
    Authentication SEQ: 0x0001
    Status code: Successful (0x0000)
```

← From Station

From AP →

```
IEEE 802.11 Authentication, Flags: ........C
  Type/Subtype: Authentication (0x000b)
  Frame Control Field: 0xb000
  .000 0001 0011 1010 = Duration: 314 microseconds
  Receiver address: D-LinkIn_6a:cc:6e (84:c9:b2:6a:cc:6e)
  Destination address: D-LinkIn_6a:cc:6e (84:c9:b2:6a:cc:6e)
  Transmitter address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
  Source address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
  BSS Id: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
  .... .... .... 0000 = Fragment number: 0
  1010 1001 0000 .... = Sequence number: 2704
  Frame check sequence: 0x9f8350e1 [unverified]
  [FCS Status: Unverified]
IEEE 802.11 wireless LAN
  Fixed parameters (6 bytes)
    Authentication Algorithm: Open System (0)
    Authentication SEQ: 0x0002
    Status code: Successful (0x0000)
```

Joining BSS with AP: Association

- Once a station is authenticated, it starts the association process, i.e., information exchange about the AP/station capabilities and roaming
  - STA → AP: Associate Request frame
    - Enables the AP to allocate resources and synchronize. The frame carries information about the WNIC, including supported data rates and the SSID of the network the station wishes to associate with.
  - AP → STA: Association Response frame
    - Acceptance or rejection to an association request. If it is an acceptance, the frame will contain information such as association ID and supported data rates.
  - New AP informs old AP (if it is a handover).
- Only after association is completed, a station can transmit and receive data frames.

Association Request/Response Frames

← From Station

```
▾ IEEE 802.11 Association Request, Flags: ........
    Type/Subtype: Association Request (0x0000)
  ▸ Frame Control Field: 0x0000
    .000 0001 0011 1010 = Duration: 314 microseconds
    Receiver address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    Destination address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    Transmitter address: D-LinkIn_6a:cc:6e (84:c9:b2:6a:cc:6e)
    Source address: D-LinkIn_6a:cc:6e (84:c9:b2:6a:cc:6e)
    BSS Id: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    .... .... .... 0000 = Fragment number: 0
    0001 0100 1100 .... = Sequence number: 332
▾ IEEE 802.11 wireless LAN
  ▾ Fixed parameters (4 bytes)
    ▸ Capabilities Information: 0x0421
      Listen Interval: 0x000a
  ▾ Tagged parameters (43 bytes)
    ▸ Tag: SSID parameter set: LABCOM
    ▸ Tag: Supported Rates 1, 2, 5.5, 11, 6, 9, 12, 18, [Mbit/sec]
    ▸ Tag: Extended Supported Rates 24, 36, 48, 54, [Mbit/sec]
    ▸ Tag: Extended Capabilities (8 octets)
    ▸ Tag: Vendor Specific: Microsoft Corp.: WMM/WME: Information E
```

From AP →

```
▾ IEEE 802.11 Association Response, Flags: ........C
    Type/Subtype: Association Response (0x0001)
  ▸ Frame Control Field: 0x1000
    .000 0001 0011 1010 = Duration: 314 microseconds
    Receiver address: D-LinkIn_6a:cc:6e (84:c9:b2:6a:cc:6e)
    Destination address: D-LinkIn_6a:cc:6e (84:c9:b2:6a:cc:6e)
    Transmitter address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    Source address: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    BSS Id: Cisco_61:ee:d0 (00:1c:f6:61:ee:d0)
    .... .... .... 0000 = Fragment number: 0
    1010 1001 0001 .... = Sequence number: 2705
    Frame check sequence: 0xe7103b15 [unverified]
    [FCS Status: Unverified]
▾ IEEE 802.11 wireless LAN
  ▾ Fixed parameters (6 bytes)
    ▸ Capabilities Information: 0x0421
      Status code: Successful (0x0000)
      ..00 0000 0000 0001 = Association ID: 0x0001
  ▾ Tagged parameters (42 bytes)
    ▸ Tag: Supported Rates 1(B), 2(B), 5.5(B), 6, 9, 11(B), 12, 18, [Mbit/sec]
    ▸ Tag: Extended Supported Rates 24, 36, 48, 54, [Mbit/sec]
    ▸ Tag: Vendor Specific: Microsoft Corp.: WMM/WME: Parameter Element
```

Data Frame

```
▾ IEEE 802.11 QoS Data, Flags: .p.....TC
    Type/Subtype: QoS Data (0x0028)
  ▸ Frame Control Field: 0x8841
    .000 0001 0011 1010 = Duration: 314 microseconds
    Receiver address: Cisco_61:ee:d1 (00:1c:f6:61:ee:d1)
    Transmitter address: IntelCor_e8:14:53 (b8:8a:60:e8:14:53)
    Destination address: D-LinkIn_6a:cc:6e (84:c9:b2:6a:cc:6e)
    Source address: IntelCor_e8:14:53 (b8:8a:60:e8:14:53)
    BSS Id: Cisco_61:ee:d1 (00:1c:f6:61:ee:d1)
    STA address: IntelCor_e8:14:53 (b8:8a:60:e8:14:53)
    .... .... .... 0000 = Fragment number: 0
    0000 0000 0011 .... = Sequence number: 3
    Frame check sequence: 0xc72771e8 [unverified]
    [FCS Status: Unverified]
  ▸ Qos Control: 0x0000
  ▸ CCMP parameters
▾ Data (1244 bytes)
    Data: f8002648417037bc923106ead1717d4821fde0989beb08b1…
    [Length: 1244]
```

← Node that will receive frame (AP)
← Node that send frame
← Station to receive data
← Station who sent data

- Station “IntelCor*” sending data to station “D-LinkIn*” (via AP).
- Frame captured between station “IntelCor*” and AP (“Cisco*”).

Authentication and authorization mechanisms

- Changing according to the organization and the security level
  - Open network
  - Open network + MAC authentication
  - Open network + VPN-gateway
  - Open network + web-gateway
  - SSID
  - Shared key: WEP
  - Wi-Fi Protected Access (WPA)
  - IEEE 802.11i (WPA2)
  - IEEE 802.1X
  - Virtual Private Networks (VPNs)

Open Network(s)

- Open network
  - Network is open, providing IP addresses with DHCP
  - There is no authentication and access is free
  - Does not require specific software
  - Access control is complicated
  - It is possible to ‘see’ all traffic in the network (sniffing)
- Open network + MAC authentication
  - The control of the station MAC address is added
  - Larger management load
    - ... But MAC addresses can be falsified
    - ... Difficult to support guests
    - ... Impossible to use in public environments

Open Network + Gateways

- Open Network + VPN gateway.
  - Open network, with the client being authenticated in an IP VPN (L3) in order to be able to access its network from outside.
    - Requires VPN client software.
    - Difficult to use by guests.
    - Scalability is being enhanced.
    - VPN controllers can be expensive.
- Open network + web gateway.
  - Open network, with the client being authenticated in web server (L3), providing “credentials”.
    - Easy to use by guests.
    - Standardization is being enhanced.
    - Scalability is being enhanced.
    - A browser needs to be working during the session.

Service Set ID (SSID)

- **SSID – name of the network.**
- Identifies the BSS, emitted in the beacon.
- Networks can block beacon and force the AP to be directly specified by its name.
- This is not very efficient.
  - Operating systems are smarter.
  - The change of SSID requires a new advertisement to all stations.
  - With the increasing number of stations, security will decrease.
  - SSID is only useful to the self-organization of the stations, not to security.

WEP Protocol

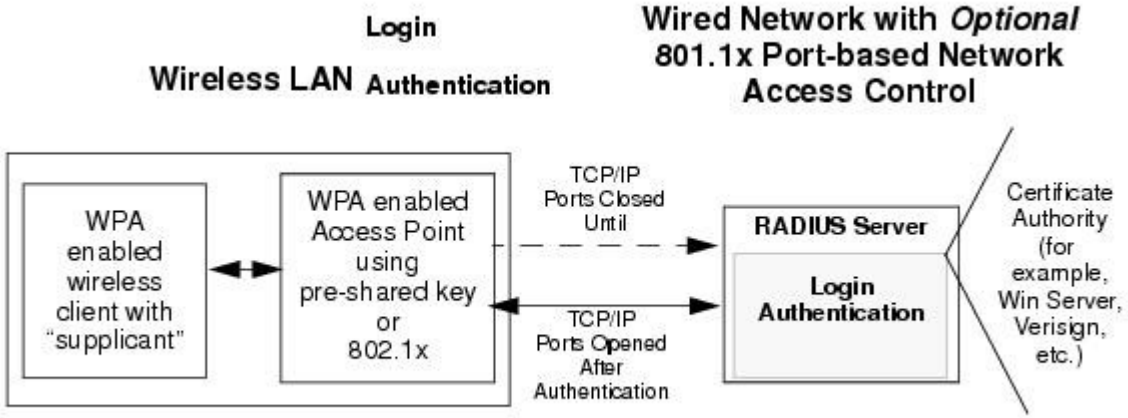
- Wired Equivalent Privacy → shared key scheme.
- Part of basic 802.11 standard.
- Security protocol at link layer (L2).
- Designed to be computationally efficient and self-synchronized.
- The station has to know the key (like a password) to access the AP.
- With passive monitoring, it can be broken (in seconds)
  - Header is not ciphered, all destinations and origins are visible.
  - Control frames are not ciphered, and then they can be changed.
  - AP is not authenticated and can be falsified.
  - **Should not be implemented!**

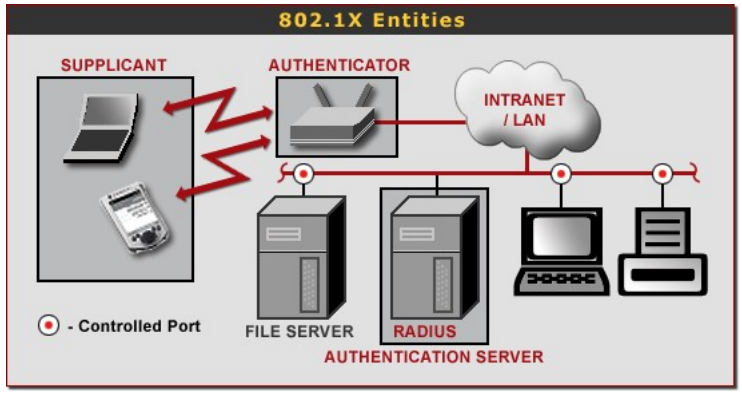
WPA and 802.11i (WPA2)

- **IEEE 802.11i - IEEE 802.11 task group “MAC enhancement for wireless security”.**
- **Wi-Fi Protected Access (WiFi Alliance), WPA, is a subset internal in 802.11i.**
  - Compatible with work developed in 802.11i.
  - Only supports BSS.
  - Defined to work in actual equipment.
    - Firmware update only.
  - Pass-phrase constant and shared, but keys are generated per session.
  - Used in the AP and station.
  - Uses “Open System” during authentication phase.
- WPA has two distinct components.
  - Authentication, based on 802.1X.
  - Ciphering based on TKIP (Temporal Key Integrity Protocol).

IEEE 802.1X

- Layer 2 solution between station and AP.
  - Available in many equipments (e.g. IEEE 802.xx).
  - Web systems frequently use 802.1X.
- Several authentication-mechanisms available (EAP-MD5, EAP-TLS, EAP-TTLS, PEAP)
- Multiple standard ciphering algorithms .
- Can cipher data with dynamic keys.
- Resorts to RADIUS servers.





WPA* Key Exchange

- Done during the Association process.
  - After Association Request/response frames.

```
205 595.669409767  IntelCor_e8:14:53   Cisco_61:ee:d1      802.11   110 Association Request, SN=38, FN=0, Flags=........, SSID=LABCOM_SEC
206 595.671214291  Cisco_61:ee:d1      IntelCor_e8:14:53   802.11   128 Association Response, SN=14, FN=0, Flags=........
207 595.673042781  Cisco_61:ee:d1      IntelCor_e8:14:53   EAPOL    211 Key (Message 1 of 4)
208 595.678333124  IntelCor_e8:14:53   Cisco_61:ee:d1      EAPOL    168 Key (Message 2 of 4)
209 595.681795313  Cisco_61:ee:d1      IntelCor_e8:14:53   EAPOL    269 Key (Message 3 of 4)
210 595.683690439  IntelCor_e8:14:53   Cisco_61:ee:d1      EAPOL    146 Key (Message 4 of 4)

▸ Frame 207: 211 bytes on wire (1688 bits), 211 bytes captured (1688 bits) on interface 0
▸ Radiotap Header v0, Length 56
▸ 802.11 radio information
▾ IEEE 802.11 QoS Data, Flags: ......F.
    Type/Subtype: QoS Data (0x0028)
  ▸ Frame Control Field: 0x8802
    .000 0001 0011 1010 = Duration: 314 microseconds
    Receiver address: IntelCor_e8:14:53 (b8:8a:60:e8:14:53)
    Transmitter address: Cisco_61:ee:d1 (00:1c:f6:61:ee:d1)
    Destination address: IntelCor_e8:14:53 (b8:8a:60:e8:14:53)
    Source address: Cisco_61:ee:d1 (00:1c:f6:61:ee:d1)
    BSS Id: Cisco_61:ee:d1 (00:1c:f6:61:ee:d1)
    STA address: IntelCor_e8:14:53 (b8:8a:60:e8:14:53)
    .... .... .... 0000 = Fragment number: 0
    0000 0001 1100 .... = Sequence number: 28
  ▸ Qos Control: 0x0007
▸ Logical-Link Control
▾ 802.1X Authentication
    Version: 802.1X-2004 (2)
    Type: Key (3)
    Length: 117
    Key Descriptor Type: EAPOL RSN Key (2)
    [Message number: 1]
  ▸ Key Information: 0x008a
    Key Length: 16
    Replay Counter: 1
    WPA Key Nonce: 4f65d0b4e9e77b88f2cbb135749eeb105a3aa1ef65de66a8…
    Key IV: 00000000000000000000000000000000
    WPA Key RSC: 0000000000000000
    WPA Key ID: 0000000000000000
    WPA Key MIC: 00000000000000000000000000000000
    WPA Key Data Length: 22
  ▸ WPA Key Data: dd14000fac046616ebb59b83e8cc1816ced0e542a935
```

Layer 3 - Addressing

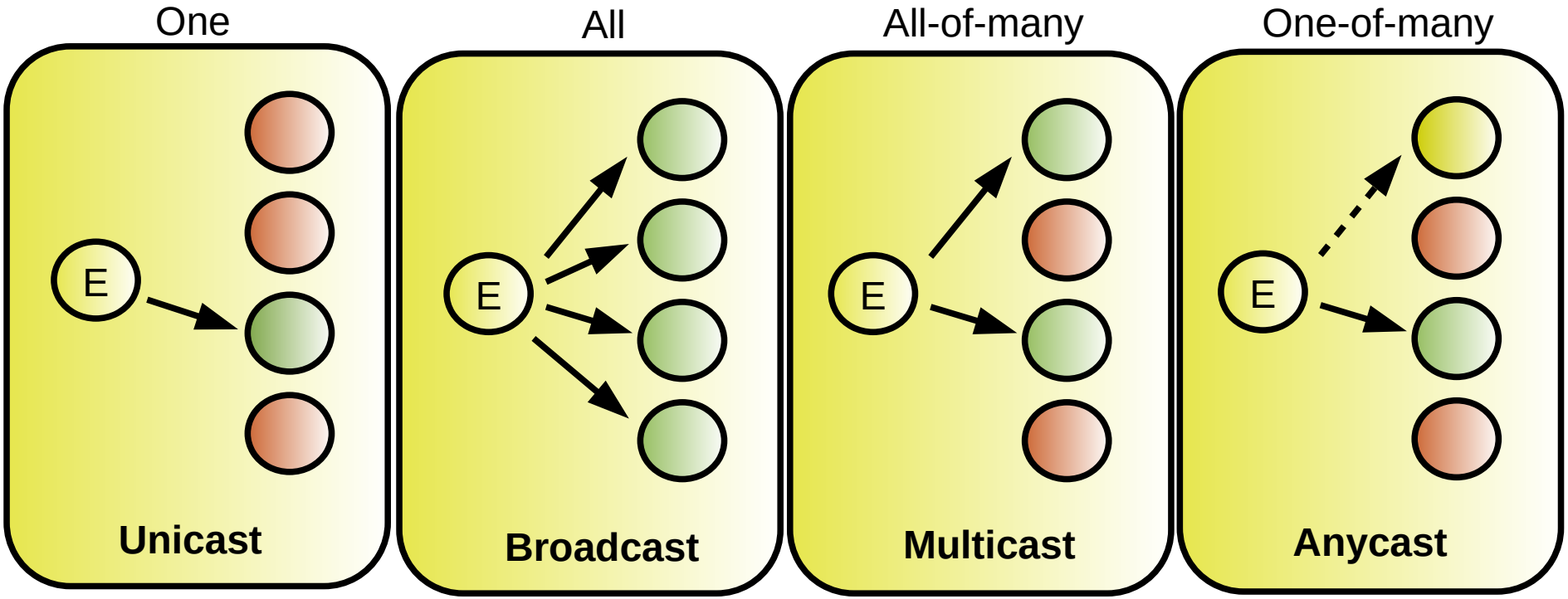
Fundamentos de Redes

Mestrado Integrado em
Engenharia de Computadores e Telemática
DETI-UA

universidade de aveiro
deti.ua.pt

Types of Addresses

- Unicast – Identify a single sender/receiver.
- Broadcast – All are receivers.
- Multicast – Identify all elements of a group as receivers (all-of-many)
- Anycast – Identifies any element of group as receiver (one-of-many)



IPv4 Addressing

- An IPv4 address is a unique address for a network interface
- Exceptions:
  - Dynamically assigned IPv4 addresses (DHCP)
  - IP addresses in private networks (NAT)

- An IPv4 address:
  - is a **32 bit long** identifier
  - encodes a network number (**network prefix**)
    and a **host identifier**

Network Prefix and Host Identifier

- The network prefix identifies a network and the host identifier identifies a specific host (actually, interface on the network).

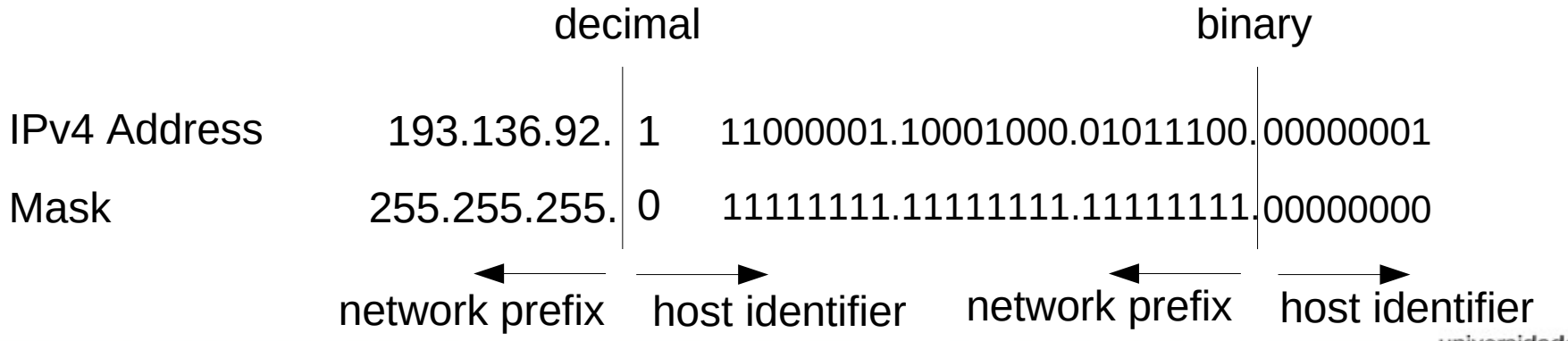
- How do we know how long the network prefix is?
  - **Before 1993:** The boundary between network prefix and host identifier is implicitly defined (**class-based/classful addressing**)

  or

  - **After 1993:** The boundary between network prefix and host identifier is indicated by a **netmask.**

Classless Inter-Domain Routing (CIDR)

- New interpretation of the IP addressing to increase efficiency and flexibility.
  - Network Masks were created to define the boundary between the IP network prefix and host identifier.
  - A bit of the mask equal to one indicate that that bit (in that position) of the address belongs to the network prefix.
    - A bit of the mask equal to zero indicate that that bit (in that position) of the address belongs to the host identifier.
  - Called VLSM (Variable Length Subnet Mask).
  - Must be provided with the IP address.
- Allowed the partition of a network in smaller networks or sub-networks (subnets).
- Allowed to merge several network under a single prefix (aggregation or summary process).



Mask Notations

- There are two notations for IPv4 masks:
  - Decimal: 4 bytes separated by dots.
  - CIDR: A slash (/) a a number with the number of bits of the network prefix.
- Both notations still exist today.
  - CIDR starts to become prevalent.
  - IPv6 only supports CIDR.

CIDR	Decimal
/21	255.255.248.0
/20	255.255.240.0
/19	255.255.224.0
/18	255.255.192.0
/17	255.255.128.0
/16	255.255.0.0
/15	255.248.0.0
/14	255.240.0.0
/13	255.224.0.0

CIDR	Decimal
/30	255.255.255.252
/29	255.255.255.248
/28	255.255.255.240
/27	255.255.255.224
/26	255.255.255.192
/25	255.255.255.128
/24	255.255.255.0
/23	255.255.254.0
/22	255.255.252.0

CIDR Address Blocks

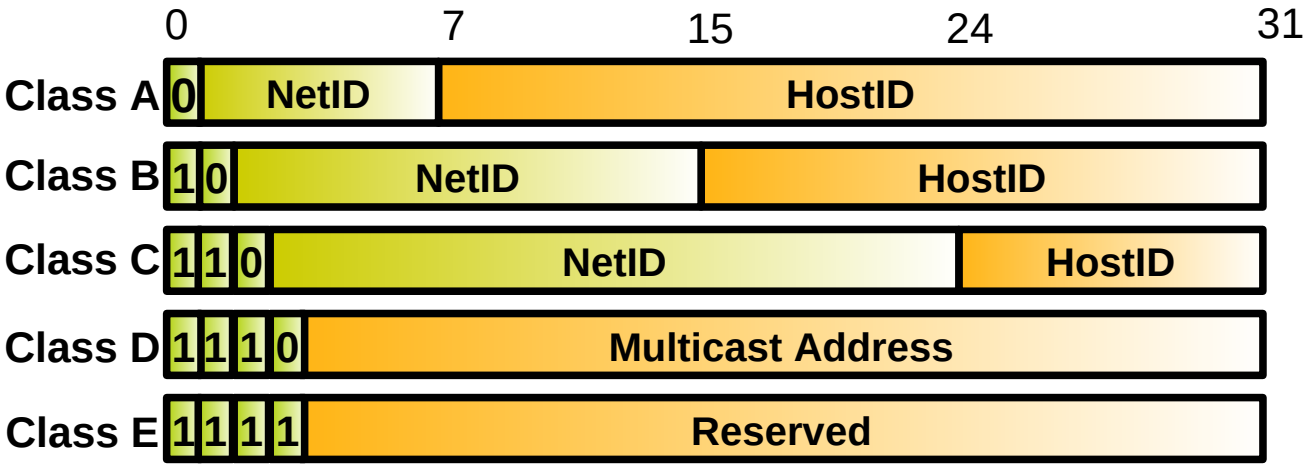
- CIDR defines a block of addresses.
- The addresses blocks are used to assign
- #Addresses= 2^(32-CIDR)
  - Example: \24→ 2^(32-24)=2^8=256, \28→ 2^(32-28)=2^4=16
- #Usable_Addresses = #Addresses – 2 addresses
  - Network prefix and broadcast address

CIDR	# of addresses	# usable addresses
21	2048	2046
20	4096	4094
19	8192	8190
18	16384	16382
17	32768	32766
16	65536	65534
15	131072	131070
14	262144	262142
13	524288	524286

CIDR	# of addresses	# usable addresses
30	4	2
29	8	6
28	16	14
27	32	30
26	64	62
25	128	126
24	256	254
23	512	510
22	1024	1022

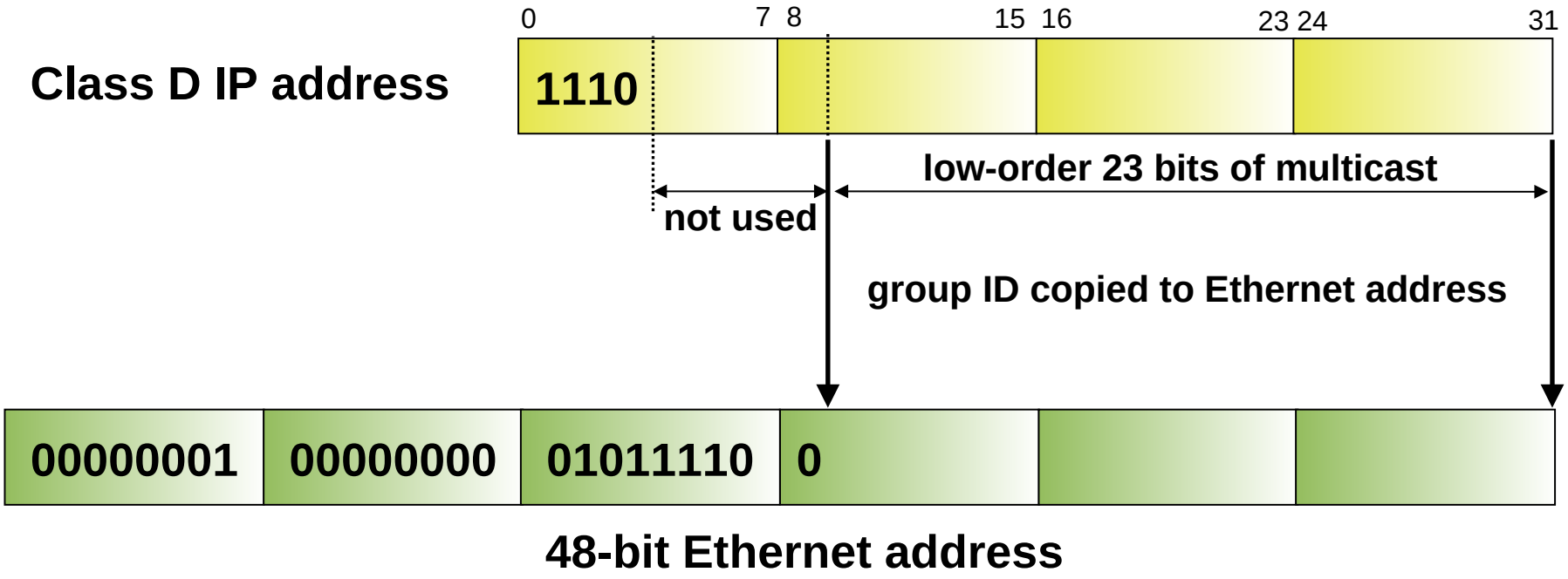
IPv4 Classful Addressing

- Initially (until 1993) the boundary between the network prefix and host identifier was predefined by the value of the first byte (class).
- Resulted in a huge waste of addresses:
  - Classes A and B were to big,
  - Not enough class C networks.
- Routing Tables were becoming very long
  - It was not possible to merge (aggregate) networks to simplify routing tables.



Class	First Address	Last Address
A	1.0.0.0	126.0.0.0
B	128.0.0.0	191.255.0.0
C	192.0.0.0	223.255.255.0
D	224.0.0.0	239.255.255.255
E	240.0.0.0	255.255.255.254

Conversion of Multicast IPv4 Address to Ethernet Address

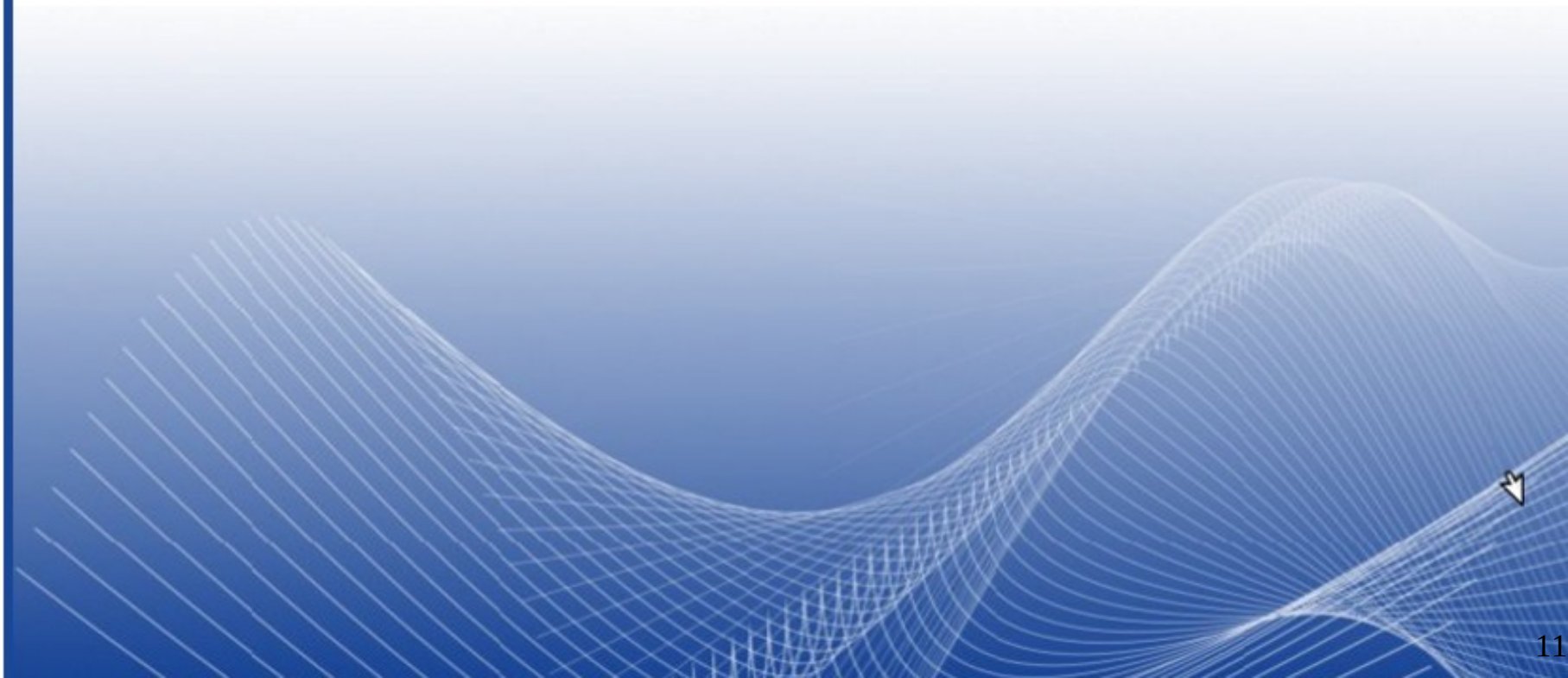


IPv4 Private Networks

Prefix	First Address	Last Address
10.0.0.0/8	10.0.0.0	10.255.255.255
172.16.0.0/12	172.16.0.0	172.31.255.255
192.168.0.0/16	192.168.0.0	192.168.255.255
169.254.0.0/16	169.254.0.0	169.254.255.255

- To be used within a local network.
- Packets with these addresses as destination are not routed to the Internet.
- Packets with these addresses as source should not be routed to the Internet.
  - Not default behavior!

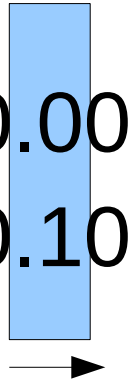
IPv4 Address Planning



IPv4 Network Sub-netting

- Made allowed by Variable Length Subnet Mask.
- Division of an IPv4 networks into smaller IPv4 networks.
- Allows to save IPv4 addresses.
  - Assign a large network to a small network will have many address not assigned.
  - A large network may divided into smaller networks and each one assign to different LAN.

193.136.92.0/24 → 193.136.92.0***/25*** + 193.136.92.128***/25***



.92.000=01011100.00000000$_2$

.92.128=01011100.10000000$_2$

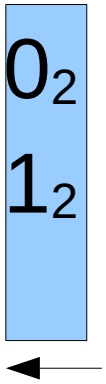
IPv4 Network Aggregation

- Inverse process to network sub-netting.
- Used to obtain a single network prefix to multiple networks.
  - Mainly used to simplify routing.
- Example:

  193.136.92.0/24 + 193.136.93.0/24 → 193.136.92.0**/23**

.92.=01011100$_2$

.93.=01011101$_2$



IPv4 Address Planning (1)

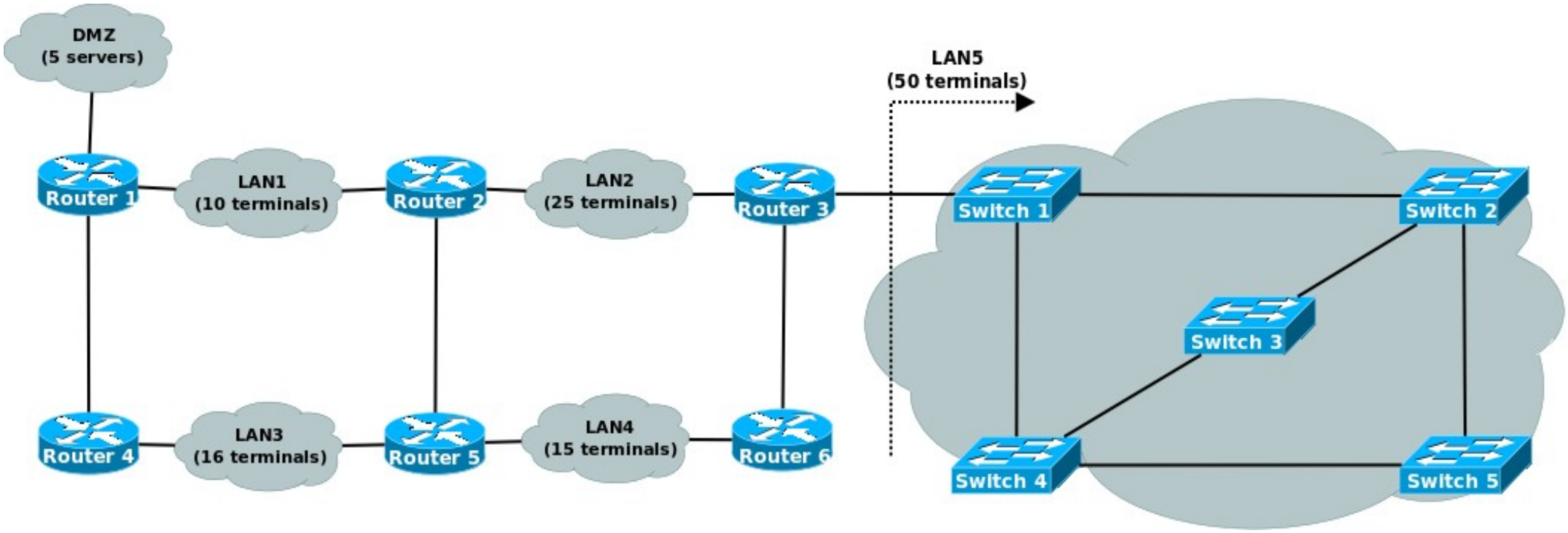
- Address planning is the assignment of an IP network to a (V)LAN.
  - To be assign address manually or dynamically (DHCP).
- Public addresses planning:
  - Limited number of available IPv4 addresses.
  - Planning ruled by the number of hosts in each LAN that require a public IPv4 address.
  - Not all LAN require IPv4 addresses.
  - Not all host in a LAN require IPv4 addresses.
  - Usually network managers receive /23, /24 or /25 networks.
- Private addresses planning:
  - Number of addresses is not an issue.
  - Number of hosts in a LAN is not so relevant.
    - Networks are usually divided in standard (/24), point-to-point (/30) and larger networks (may use /23, /22, /21, /20, etc…).

IPv4 Address Planning (2)

- Best practices:
  - Identify the available IPv4 network(s).
  - Identify the number of host in each (V)LAN.
    - Including terminals and routers (gateways).
  - Define each sub-network size.
    - Network prefix and broadcast addresses are not usable.
    - Define network mask.
  - Sort sub-networks from larger to smaller.
    - Smaller CIDR to higher CIDR.
  - Start from the available network.
    - Sub-divide in half.
    - If sub-network size is required → **Assigned it** → ITS SUB-NETWORKS ARE NOT USABLE IN OTHER LAN.
    - If sub-network size is larger that required → **Sub-divide it in half**.
    - Repeat until all LAN have an assigned IPv4 network.
      - The overall available network may not be enough to assigned sub-networks to all LAN. The solution is to reevaluate requirements and assigned smaller sub-networks.

Example – IPv4 Public Planning (1)

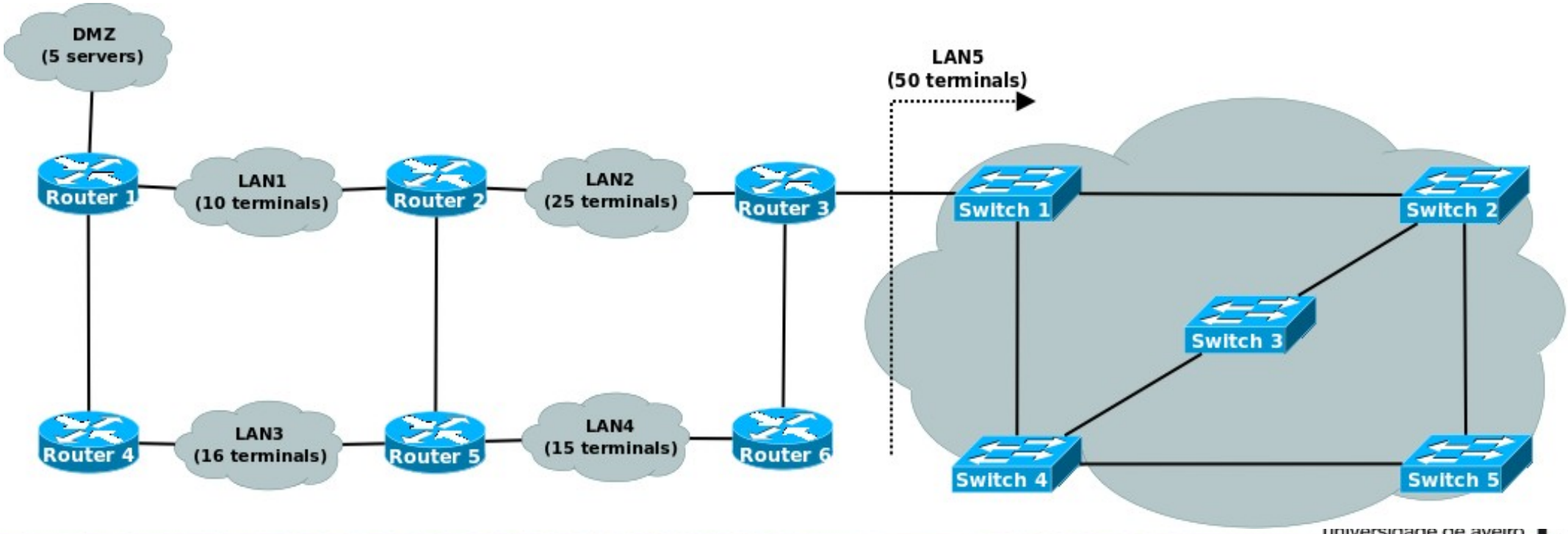
- Problem: Multiple (V)LAN require a small number of public IPv4 addresses. The public IPv4 network available is 193.1.1.0/24.
  - Note: All (V)LAN require IPv4 addresses, however may use private addresses (another IPv4 network).



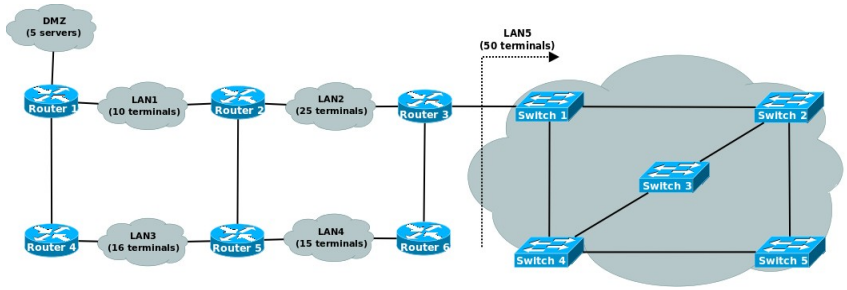
192.1.1.0/24

Example – IPv4 Public Planning (2)

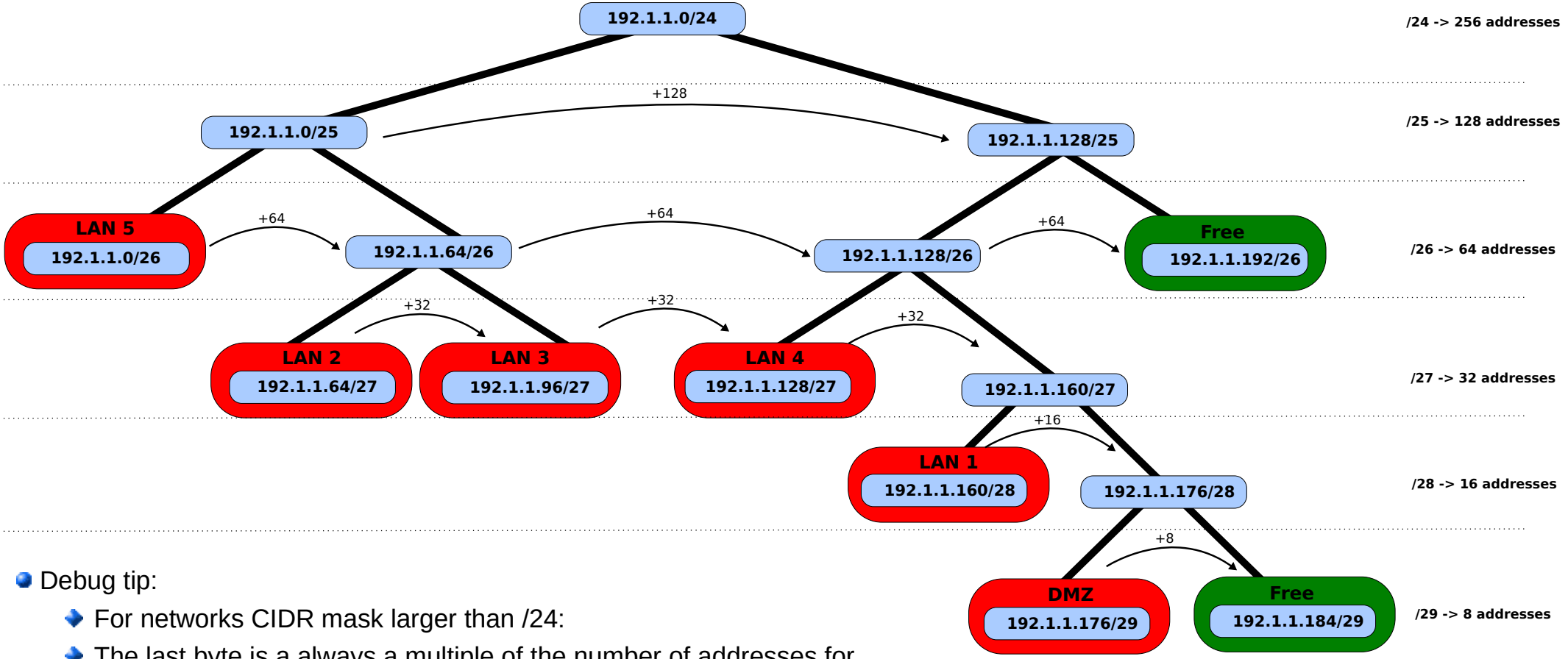
- LAN 1 → **10**+2 routers/gw+prefix+broadcast= 14 → 16 → /28 net
- LAN 2 → **25**+2 routers/gw+prefix+broadcast= 29 → 32 → /27 net
- LAN 3 → **16**+2 routers/gw+prefix+broadcast= 20 → 32 → /27 net
- LAN 4 → **15**+2 routers/gw+prefix+broadcast= 19 → 32 → /27 net
- LAN 5 → **50**+1 router/gw+prefix+broadcast= 53 → 64 → /26 net
- DMZ → **5**+1 router/gw+prefix+broadcast = 8 → 8 → /29 net



Example (3)



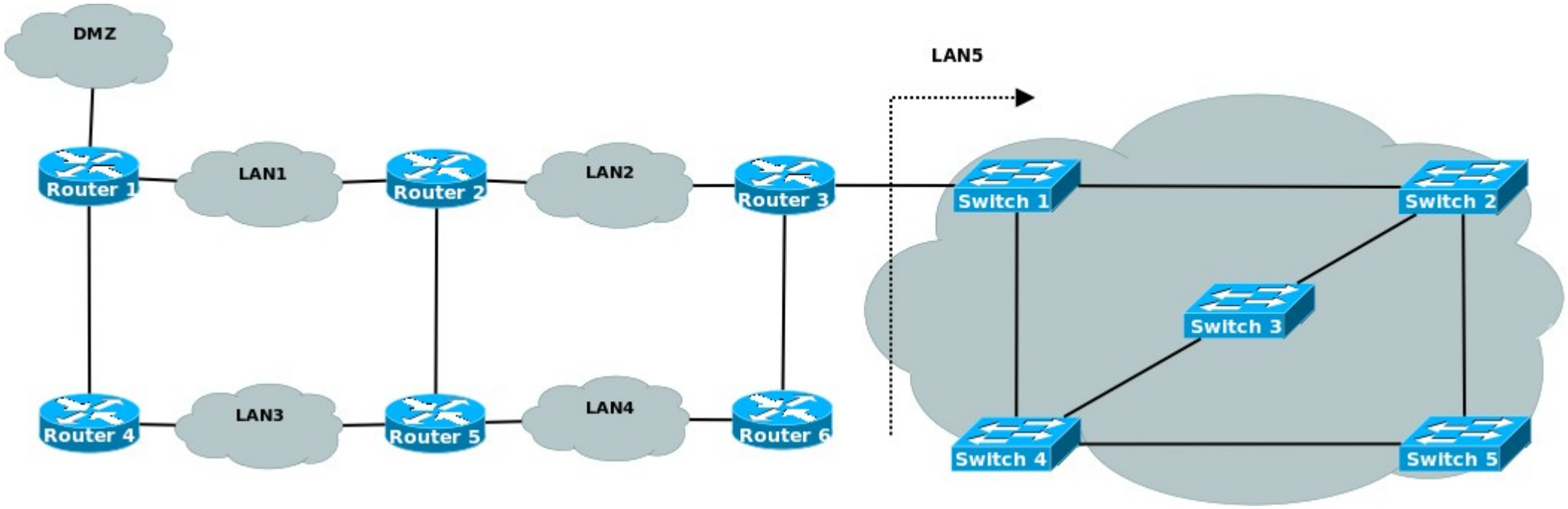
- LAN 1 → 10+2+2=14 → 16 → /28 net
- LAN 2 → 25+2+2=29 → 32 → /27 net
- LAN 3 → 16+2+2=20 → 32 → /27 net
- LAN 4 → 15+2+2=19 → 32 → /27 net
- LAN 5 → 50+1+2=53 → 64 → /26 net
- DMZ → 5+1+2=8 → 8 → /29 net



- Debug tip:
  - For networks CIDR mask larger than /24:
  - The last byte is a always a multiple of the number of addresses for that network size.
    - Example: 192 is multiple of 64, 176 is multiple of 16, and 184 is multiple of 8.

Example – IPv4 Private Planning (1)

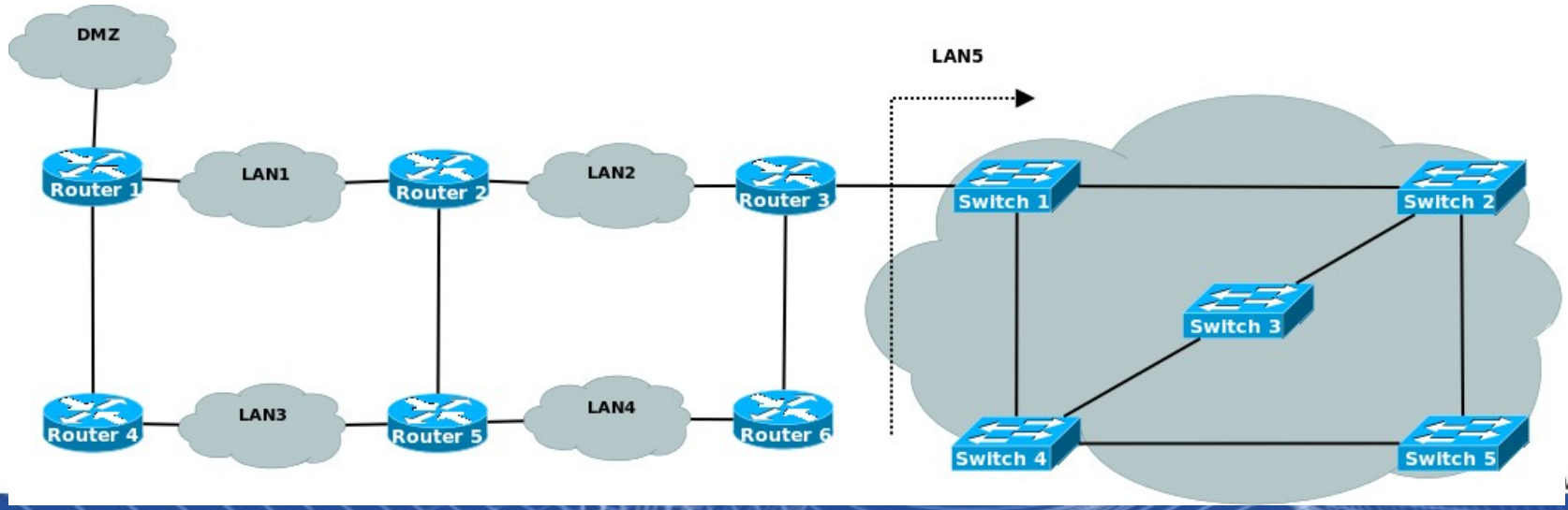
- Problem: All (V)LAN have a standard size, except LAN 5 that may have 1000 hosts.



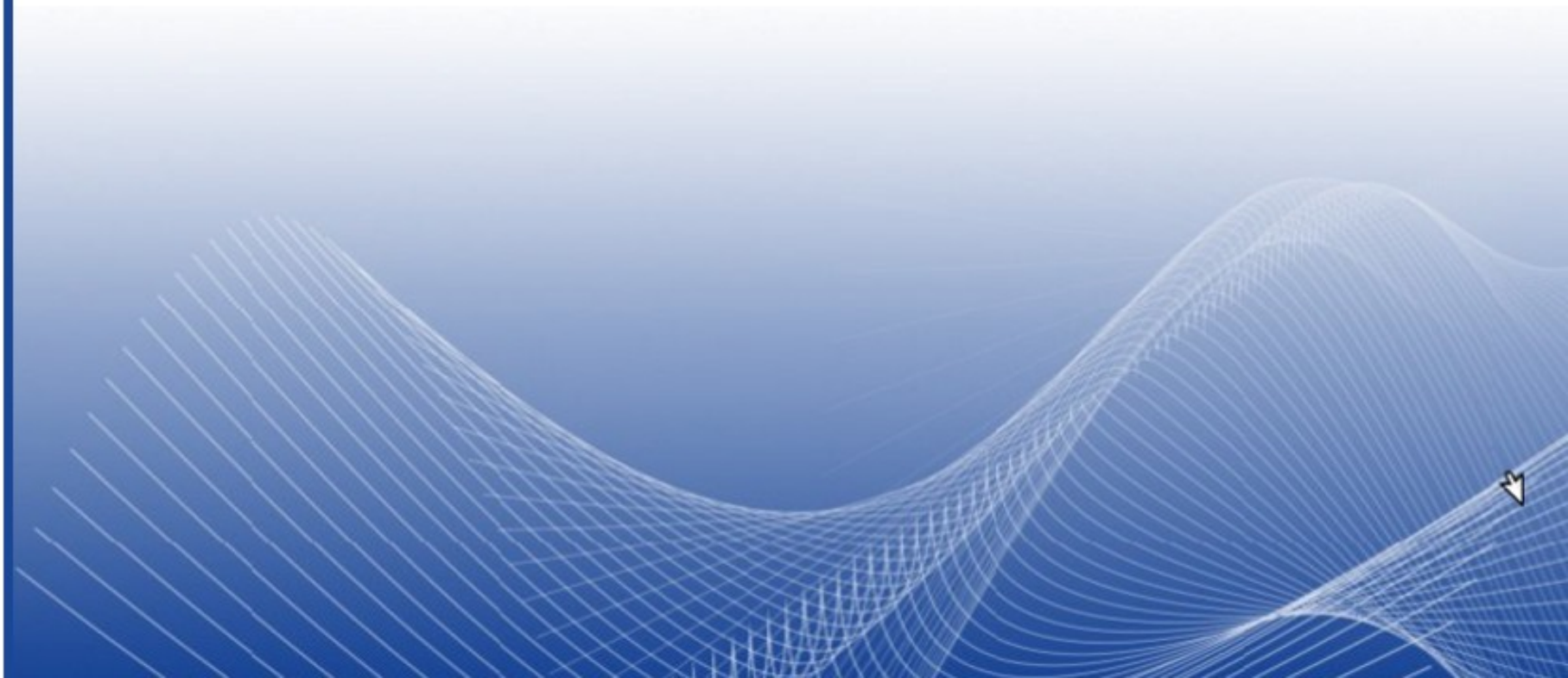
10.0.0.0/8

Example – IPv4 Private Planning (2)

- Easier approach is to start from /24 networks and perform sub-netting/aggregation as required.
- Start with larger networks.
- LAN5 with 1000 users (plus one router) will be a /22 network (2^(32-22)-2=1022 usable addresses).
  - Aggregation of networks 10.0.0.0/24, 10.0.1.0/24, 10.0.2.0/24 and 10.0.3.0/24.
  - Assigned: 10.0.0.0/22
- LAN1 to LAN4 and DMZ have a standard size and will be a /24 network.
  - Assigned: 10.0.4.0/24, 10.0.5.0/24, 10.0.6.0/24, 10.0.7.0/24, 10.0.8.0/24
- Point-to-point networks R1-R4, R2-R5 and R3-R6 will be /30 networks.
  - Network 10.0.9.0/24 will be used to perform the sub-netting.
  - Assigned: 10.0.9.0/30, 10.0.9.4/30, 10.0.9.8/30
  - Free: 10.0.9.12/30+10.0.9.16/28+10.0.9.32/27+10.0.9.64/26+10.0.9.128/25



DHCP



universidade de aveiro
deti.ua.pt

Dynamic Host Configuration Protocol (DHCP)

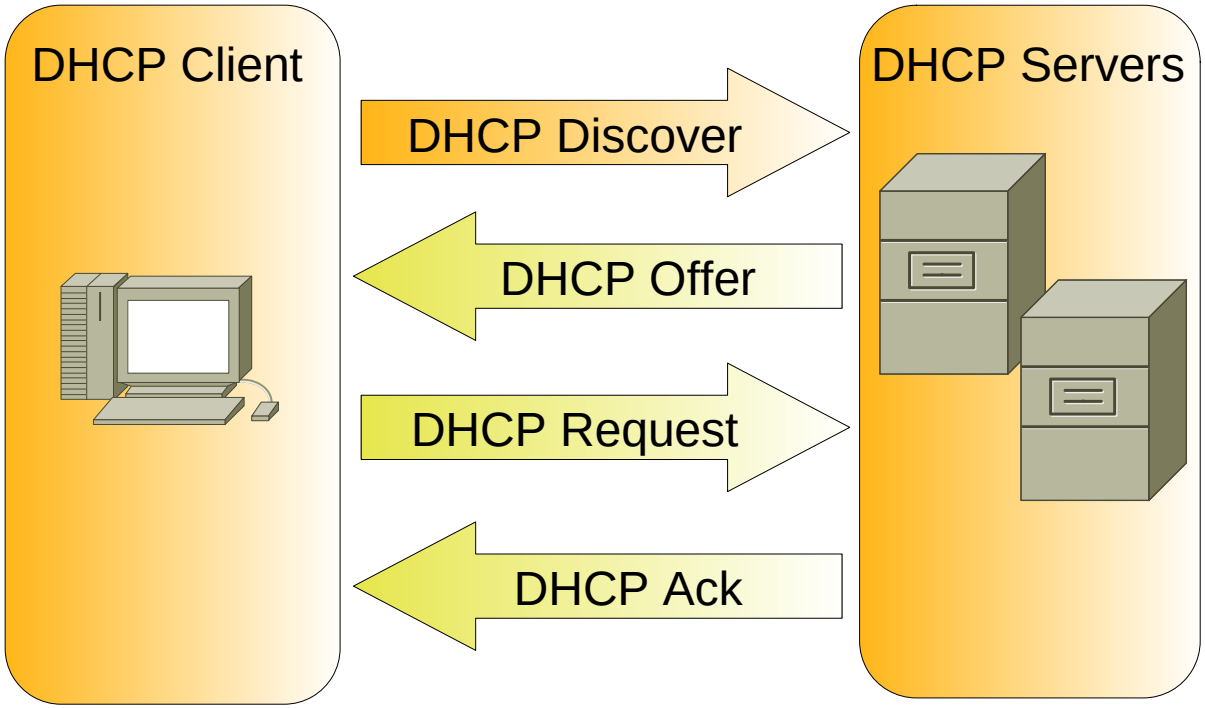
- Service for dynamic assignment of IP addresses.
  - Client-Server architecture.
- Extension of the Bootstrap Protocol, BOOTP, (RFC 1542)
  - Runs over UDP.
    - Server port 67 and client port 68.
- Address assignment follow a leasing paradigm.
- The assignment of address has four phases:
  - Discover
  - Offer
  - Request
  - Acknowledge
- DHCP servers provide:
  - Address, network mask and gateway.
  - May include additional information DNS server, Windows Domain Servers, etc…

DHCP Server

- Pool of (public or private) addresses
  - List of IPv4 public or private addresses to be assigned, usually defined as network or range of IPv4 addresses.
- Exclusion ranges
  - Set of IPv4 addresses that belong to a pool but must be assigned.
    - Usually manually assigned address to routers (gateways) and servers.
- Reserved addresses and static assignment
  - Based on the MAC address is possible to define a permanently assigned IPv4 address.
    - Usually used on servers, printers and other network devices.
    - Should not be used by routers.
- Lease time
  - Define for how long can a host use an assigned IPv4 address without a new interaction.
- To serve multiple IPv4 networks (LAN):
  - The server must have multiple pools of addresses,
  - The routers must have the BootP/DHCP Relay feature configured.
- A LAN may have multiple DHCP servers
  - For redundancy. Pools must be disjoint.

Phase One: *Discover*

- The *DHCP Discover* message id encapsulated into a *BootP Request* packet.
  - Source address is 0.0.0.0.
- It is used to discover the available DHCP server(s).
- The client may include the desired address.
  - Server is not obliged to obey.



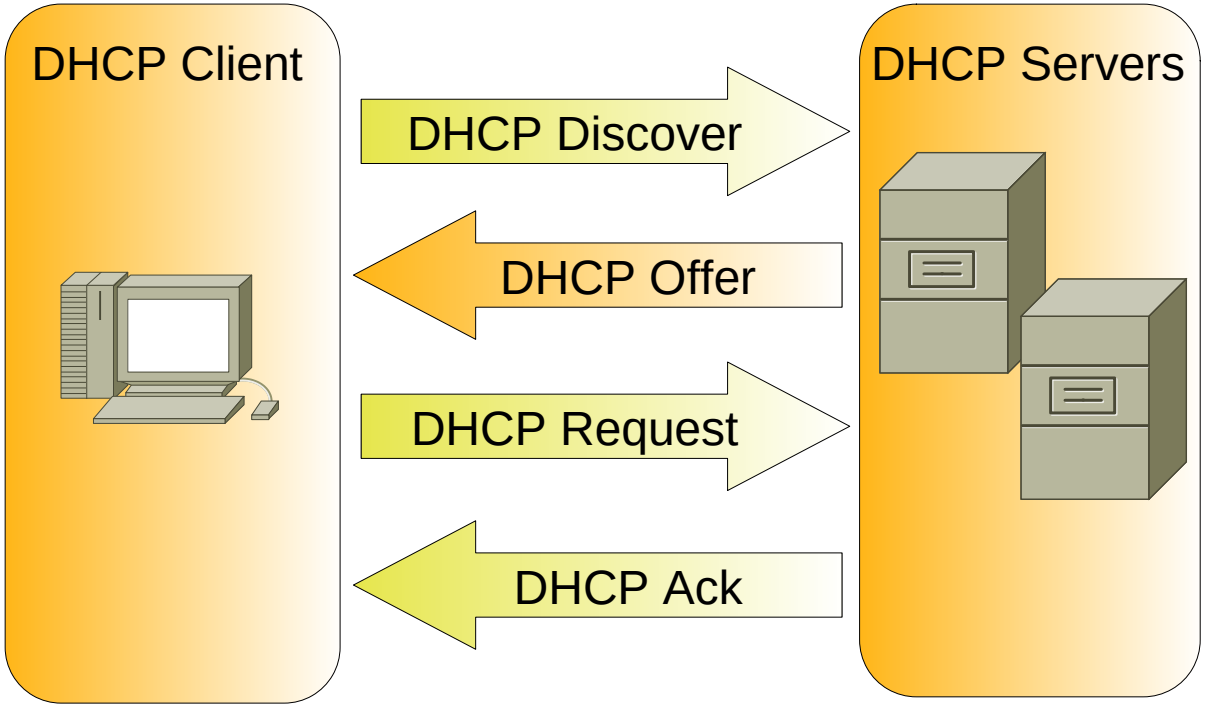
DHCP Discover

No.	Time	Source	Destination	Protocol	Info
1326	20.269579	0.0.0.0	255.255.255.255	DHCP	DHCP Discover
1337	20.561380	193.136.92.65	193.136.93.228	DHCP	DHCP Offer
1338	20.561592	0.0.0.0	255.255.255.255	DHCP	DHCP Request
1340	20.569560	193.136.92.65	193.136.93.228	DHCP	DHCP ACK

```
▷ Frame 1326 (342 bytes on wire, 342 bytes captured)
▷ Ethernet II, Src: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e), Dst: ff:ff:ff:ff:ff:ff (ff
▷ Internet Protocol, Src: 0.0.0.0 (0.0.0.0), Dst: 255.255.255.255 (255.255.255.255)
▷ User Datagram Protocol, Src Port: bootpc (68), Dst Port: bootps (67)
▽ Bootstrap Protocol
    Message type: Boot Request (1)
    Hardware type: Ethernet
    Hardware address length: 6
    Hops: 0
    Transaction ID: 0x42f5a54a
    Seconds elapsed: 0
  ▷ Bootp flags: 0x0000 (Unicast)
    Client IP address: 0.0.0.0 (0.0.0.0)
    Your (client) IP address: 0.0.0.0 (0.0.0.0)
    Next server IP address: 0.0.0.0 (0.0.0.0)
    Relay agent IP address: 0.0.0.0 (0.0.0.0)
    Client MAC address: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e)
    Client hardware address padding: 00000000000000000000
    Server host name not given
    Boot file name not given
    Magic cookie: (OK)
  ▷ Option: (t=53,l=1) DHCP Message Type = DHCP Discover
  ▷ Option: (t=50,l=4) Requested IP Address = 192.168.1.71
  ▷ Option: (t=12,l=15) Host Name = "salvador-laptop"
  ▷ Option: (t=55,l=13) Parameter Request List
    End Option
    Padding
```

Phase Two: *Offer*

- The *DHCP Offer* message is encapsulated into a *BootP Reply* packet.
- Each server proposes the lease of an IPv4 address to client.
  - If possible respect the client request (*Discovery*)



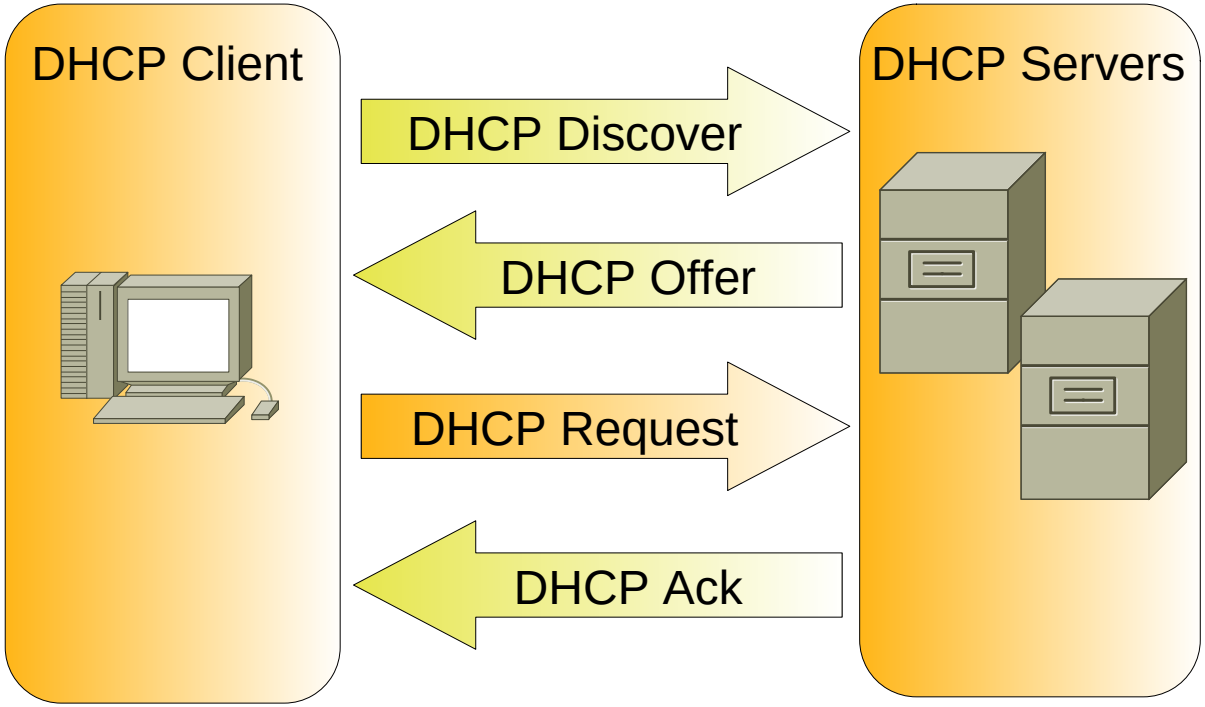
DHCP Offer

No.	Time	Source	Destination	Protocol	Info
1326	20.269579	0.0.0.0	255.255.255.255	DHCP	DHCP Discover
1337	20.561380	193.136.92.65	193.136.93.228	DHCP	DHCP Offer
1338	20.561592	0.0.0.0	255.255.255.255	DHCP	DHCP Request
1340	20.569560	193.136.92.65	193.136.93.228	DHCP	DHCP ACK

```
▷ Frame 1337 (342 bytes on wire, 342 bytes captured)
▷ Ethernet II, Src: 00:d0:b7:17:5b:6d (00:d0:b7:17:5b:6d), Dst: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e)
▷ Internet Protocol, Src: 193.136.92.65 (193.136.92.65), Dst: 193.136.93.228 (193.136.93.228)
▷ User Datagram Protocol, Src Port: bootps (67), Dst Port: bootpc (68)
▽ Bootstrap Protocol
    Message type: Boot Reply (2)
    Hardware type: Ethernet
    Hardware address length: 6
    Hops: 0
    Transaction ID: 0x42f5a54a
    Seconds elapsed: 0
  ▷ Bootp flags: 0x0000 (Unicast)
    Client IP address: 0.0.0.0 (0.0.0.0)
    Your (client) IP address: 193.136.93.228 (193.136.93.228)
    Next server IP address: 193.136.92.65 (193.136.92.65)
    Relay agent IP address: 0.0.0.0 (0.0.0.0)
    Client MAC address: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e)
    Client hardware address padding: 00000000000000000000
    Server host name not given
    Boot file name not given
    Magic cookie: (OK)
  ▷ Option: (t=53,l=1) DHCP Message Type = DHCP Offer
  ▷ Option: (t=54,l=4) DHCP Server Identifier = 193.136.92.65
  ▷ Option: (t=51,l=4) IP Address Lease Time = 10 minutes
  ▷ Option: (t=1,l=4) Subnet Mask = 255.255.254.0
  ▷ Option: (t=3,l=4) Router = 193.136.92.1
  ▷ Option: (t=15,l=8) Domain Name = "av.it.pt"
  ▷ Option: (t=6,l=4) Domain Name Server = 193.136.92.65
    End Option
    Padding
```

Phase 3: *Request*

- The *DHCP Request* message id encapsulated into a *BootP Request* packet.
- The client may chooses the offered IPv4 address (and DHCP server if more than on *offer* is received).



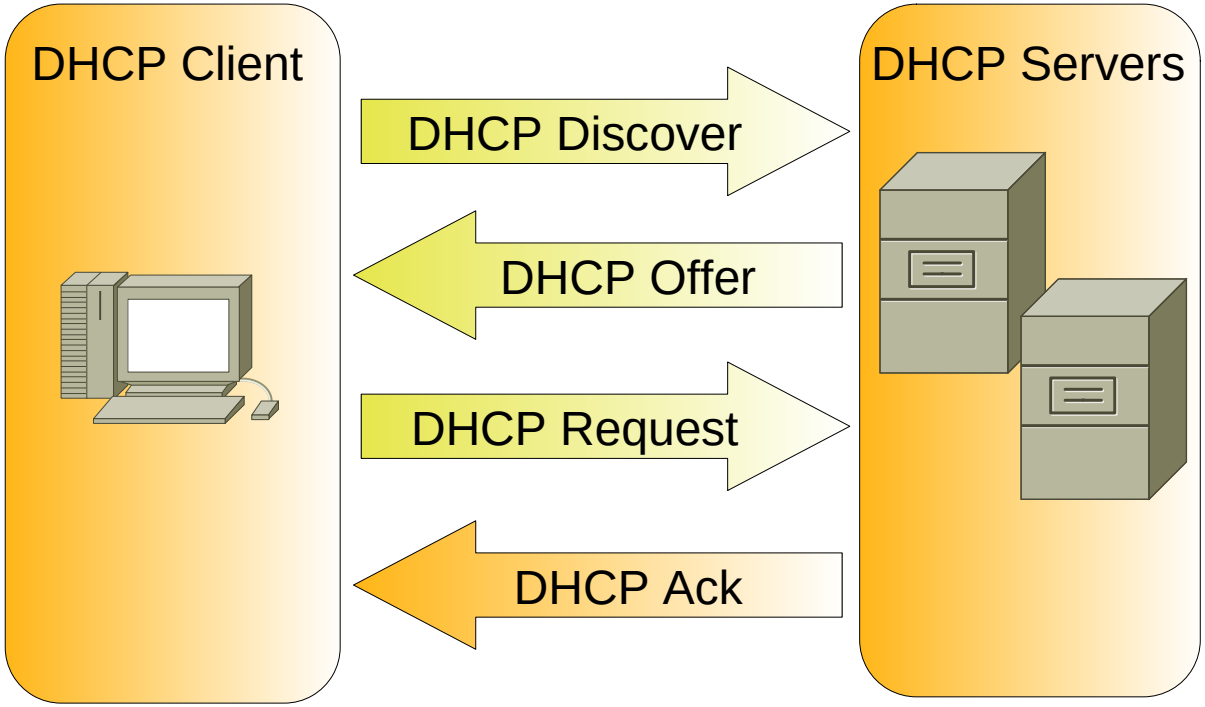
DHCP Request

No.	Time	Source	Destination	Protocol	Info
1326	20.269579	0.0.0.0	255.255.255.255	DHCP	
1337	20.561380	193.136.92.65	193.136.93.228	DHCP	
1338	20.561592	0.0.0.0	255.255.255.255	DHCP	
1340	20.569560	193.136.92.65	193.136.93.228	DHCP	

```
▷ Frame 1338 (342 bytes on wire, 342 bytes captured)
▷ Ethernet II, Src: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e), Dst: ff:ff:ff:ff:ff:ff (ff:ff:ff:ff:ff:ff)
▷ Internet Protocol, Src: 0.0.0.0 (0.0.0.0), Dst: 255.255.255.255 (255.255.255.255)
▷ User Datagram Protocol, Src Port: bootpc (68), Dst Port: bootps (67)
▽ Bootstrap Protocol
    Message type: Boot Request (1)
    Hardware type: Ethernet
    Hardware address length: 6
    Hops: 0
    Transaction ID: 0x42f5a54a
    Seconds elapsed: 0
  ▷ Bootp flags: 0x0000 (Unicast)
    Client IP address: 0.0.0.0 (0.0.0.0)
    Your (client) IP address: 0.0.0.0 (0.0.0.0)
    Next server IP address: 0.0.0.0 (0.0.0.0)
    Relay agent IP address: 0.0.0.0 (0.0.0.0)
    Client MAC address: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e)
    Client hardware address padding: 00000000000000000000
    Server host name not given
    Boot file name not given
    Magic cookie: (OK)
  ▷ Option: (t=53,l=1) DHCP Message Type = DHCP Request
  ▷ Option: (t=54,l=4) DHCP Server Identifier = 193.136.92.65
  ▷ Option: (t=50,l=4) Requested IP Address = 193.136.93.228
  ▷ Option: (t=12,l=15) Host Name = "salvador-laptop"
  ▷ Option: (t=55,l=13) Parameter Request List
    End Option
    Padding
```

Phase 4: *Acknowledge*

- The *DHCP Ack* message is encapsulated into a *BootP Reply* packet.
- The server confirms the IPv4 address lease and provides additional information:
  - Lease time, Gateway(s), DNS server, etc...



DHCP Ack

No.	Time	Source	Destination	Protocol	Info
1326	20.269579	0.0.0.0	255.255.255.255	DHCP	DHCP Discover
1337	20.561380	193.136.92.65	193.136.93.228	DHCP	DHCP Offer
1338	20.561592	0.0.0.0	255.255.255.255	DHCP	DHCP Request
1340	20.569560	193.136.92.65	193.136.93.228	DHCP	DHCP ACK

```
▷ Frame 1340 (342 bytes on wire, 342 bytes captured)
▷ Ethernet II, Src: 00:d0:b7:17:5b:6d (00:d0:b7:17:5b:6d), Dst: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e)
▷ Internet Protocol, Src: 193.136.92.65 (193.136.92.65), Dst: 193.136.93.228 (193.136.93.228)
▷ User Datagram Protocol, Src Port: bootps (67), Dst Port: bootpc (68)
▽ Bootstrap Protocol
    Message type: Boot Reply (2)
    Hardware type: Ethernet
    Hardware address length: 6
    Hops: 0
    Transaction ID: 0x42f5a54a
    Seconds elapsed: 0
  ▷ Bootp flags: 0x0000 (Unicast)
    Client IP address: 0.0.0.0 (0.0.0.0)
    Your (client) IP address: 193.136.93.228 (193.136.93.228)
    Next server IP address: 193.136.92.65 (193.136.92.65)
    Relay agent IP address: 0.0.0.0 (0.0.0.0)
    Client MAC address: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e)
    Client hardware address padding: 00000000000000000000
    Server host name not given
    Boot file name not given
    Magic cookie: (OK)
  ▷ Option: (t=53,l=1) DHCP Message Type = DHCP ACK
  ▷ Option: (t=54,l=4) DHCP Server Identifier = 193.136.92.65
  ▷ Option: (t=51,l=4) IP Address Lease Time = 10 minutes
  ▷ Option: (t=1,l=4) Subnet Mask = 255.255.254.0
  ▷ Option: (t=3,l=4) Router = 193.136.92.1
  ▷ Option: (t=15,l=8) Domain Name = "av.it.pt"
  ▷ Option: (t=6,l=4) Domain Name Server = 193.136.92.65
    End Option
    Padding
```

DHCP Operational Details

- Address Leasing Times
  - T1 Time (50% of Lease Time) – time after which the client must renew the address lease.

    T2 Time (85% of Lease Time) – time after which the client must renew the address lease if the first attempt failed.

    Lease Time – time after which the client can not used the leased address.
- DHCP allows multiple servers
  - Recommended for redundancy.
    - Requires
  - Advantage: resilience to operational failures.
  - Requirement: Disjointed pool of addresses in different servers.

DHCP Other Messages

- DHCP *Decline*:
  - Used by a client to reject the offer made by a server and musr restart the leasing process.
- DHCP *Nack*:
  - Used by a server informing that cannot satisfy the received request (DHCP *Request*).
- DHCP *Release*:
  - Used by a client informing the server that no longer requires an address. The lease is terminated.
- DHCP *Inform*:
  - Used by a client to request additional information after receiving an address.

DHCP Release

No.	Time	Source	Destination	Protocol	Info
1330	24.011686	193.136.93.228	193.136.92.65	DHCP	DHCP Release

```
▷ Frame 1330 (342 bytes on wire, 342 bytes captured)
▷ Ethernet II, Src: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e), Dst: 00:d0:b7:17:5b:6d (00:d0:b7:17:5b:6d)
▷ Internet Protocol, Src: 193.136.93.228 (193.136.93.228), Dst: 193.136.92.65 (193.136.92.65)
▷ User Datagram Protocol, Src Port: bootpc (68), Dst Port: bootps (67)
▽ Bootstrap Protocol
    Message type: Boot Request (1)
    Hardware type: Ethernet
    Hardware address length: 6
    Hops: 0
    Transaction ID: 0xc099a870
    Seconds elapsed: 0
  ▷ Bootp flags: 0x0000 (Unicast)
    Client IP address: 193.136.93.228 (193.136.93.228)
    Your (client) IP address: 0.0.0.0 (0.0.0.0)
    Next server IP address: 0.0.0.0 (0.0.0.0)
    Relay agent IP address: 0.0.0.0 (0.0.0.0)
    Client MAC address: 00:1d:ba:c0:a2:8e (00:1d:ba:c0:a2:8e)
    Client hardware address padding: 00000000000000000000
    Server host name not given
    Boot file name not given
    Magic cookie: (OK)
  ▷ Option: (t=53,l=1) DHCP Message Type = DHCP Release
  ▷ Option: (t=54,l=4) DHCP Server Identifier = 193.136.92.65
  ▷ Option: (t=12,l=15) Host Name = "salvador-laptop"
    End Option
    Padding
```

DHCP Inform

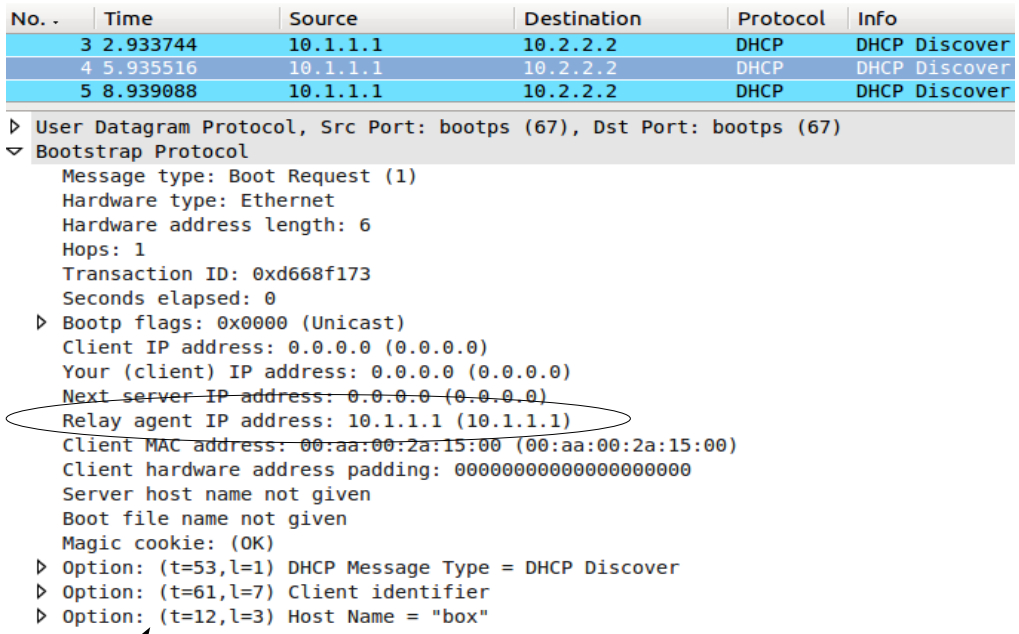
No.	Time	Source	Destination	Protocol	Info
4107	65.374546	193.136.93.173	255.255.255.255	DHCP	DHCP Inform
5446	86.143470	193.136.93.102	255.255.255.255	DHCP	DHCP Inform

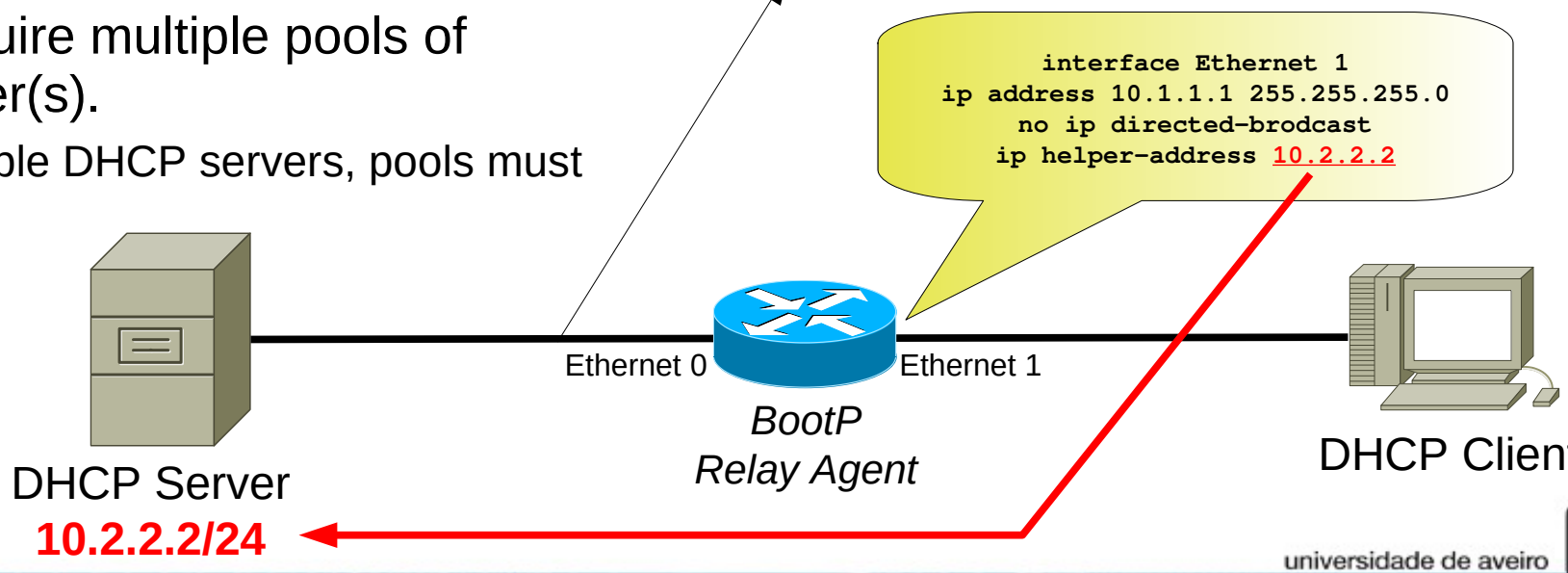
```
▷ Frame 4107 (342 bytes on wire, 342 bytes captured)
▷ Ethernet II, Src: d0:df:9a:cb:d1:3c (d0:df:9a:cb:d1:3c), Dst: ff:ff:ff:ff:ff:ff (ff:ff:ff:ff:ff:ff)
▷ Internet Protocol, Src: 193.136.93.173 (193.136.93.173), Dst: 255.255.255.255 (255.255.255.255)
▷ User Datagram Protocol, Src Port: bootpc (68), Dst Port: bootps (67)
▽ Bootstrap Protocol
    Message type: Boot Request (1)
    Hardware type: Ethernet
    Hardware address length: 6
    Hops: 0
    Transaction ID: 0xfb8eebf9
    Seconds elapsed: 0
  ▷ Bootp flags: 0x8000 (Broadcast)
    Client IP address: 193.136.93.173 (193.136.93.173)
    Your (client) IP address: 0.0.0.0 (0.0.0.0)
    Next server IP address: 0.0.0.0 (0.0.0.0)
    Relay agent IP address: 0.0.0.0 (0.0.0.0)
    Client MAC address: d0:df:9a:cb:d1:3c (d0:df:9a:cb:d1:3c)
    Client hardware address padding: 00000000000000000000
    Server host name not given
    Boot file name not given
    Magic cookie: (OK)
  ▷ Option: (t=53,l=1) DHCP Message Type = DHCP Inform
  ▷ Option: (t=61,l=7) Client identifier
  ▷ Option: (t=12,l=7) Host Name = "IT-TOSH"
  ▷ Option: (t=60,l=8) Vendor class identifier = "MSFT 5.0"
  ▷ Option: (t=55,l=13) Parameter Request List
    End Option
    Padding
```

```
▽ Option: (t=55,l=13) Parameter Request List
    Option: (55) Parameter Request List
    Length: 13
    Value: 010F03062C2E2F1F2179F92BFC
    1 = Subnet Mask
    15 = Domain Name
    3 = Router
    6 = Domain Name Server
    44 = NetBIOS over TCP/IP Name Server
    46 = NetBIOS over TCP/IP Node Type
    47 = NetBIOS over TCP/IP Scope
    31 = Perform Router Discover
    33 = Static Route
    121 = Classless Static Route
    249 = Private/Classless Static Route (Microsoft)
    43 = Vendor-Specific Information
    252 = Private/Proxy autodiscovery
```

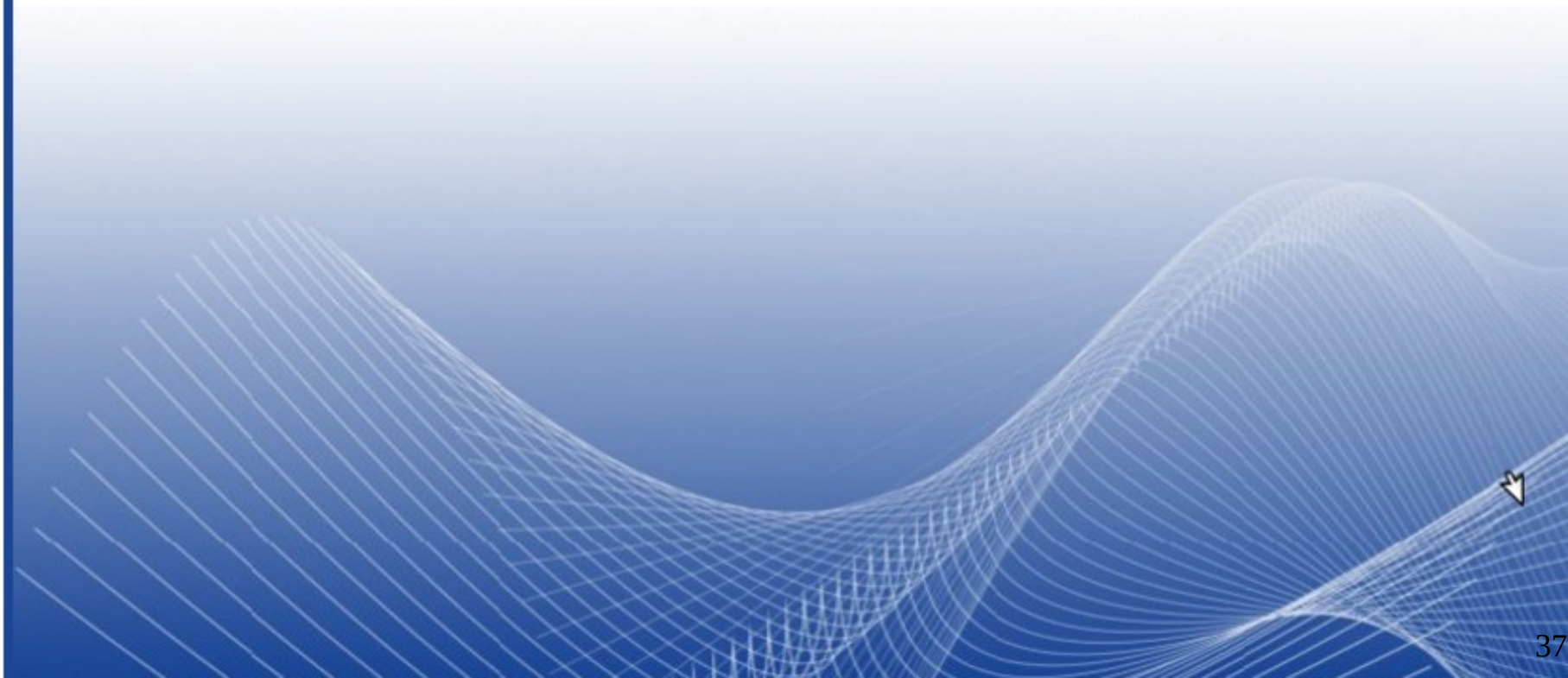
DHCP in Complex Environments

- In complex network environments where one (or more) DHCP server provide addresses to multiple (V)LAN.
  - Router must have a “BootP Relay Agent” configured and active.
  - Router redirects the client DHCP (broadcast) packets to DHCP server(s) using unicast,
    - Append information of the network/interface where it received the DHCP packet from client.
  - Router redirects server responses to the client.
  - From the client point of view, the Router behaves like a DHCP server.
- Multiple VLAN require multiple pools of addresses at server(s).
  - When using multiple DHCP servers, pools must be disjoint.





NAT and PAT

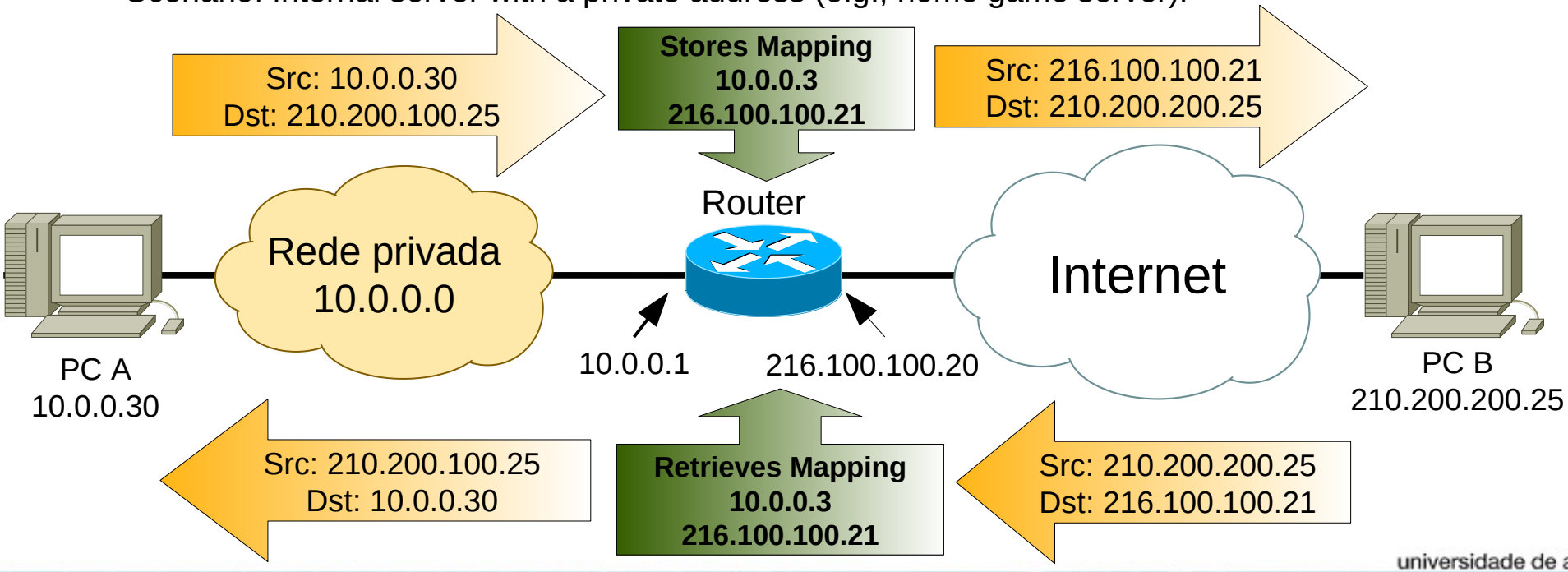


NAT (Network Address Translation) e PAT (Port Address Translation)

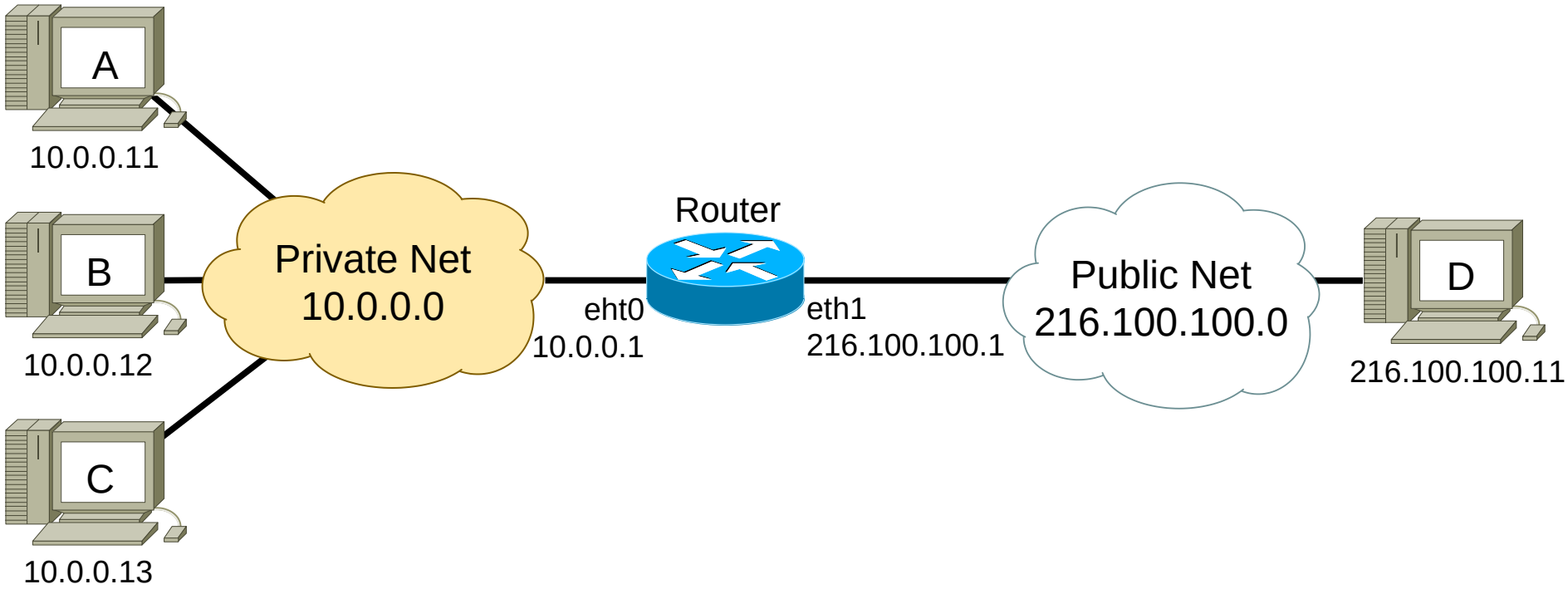
- NAT – Translates private address into public addresses.
- PAT – Translates address and also UDP/TCP ports.
  - ICMP does not have ports. ICMP identifier field is used instead.
  - Also called NAPT (Network Address and Port Translation)
- Mapping between a private and public address may be dynamic or static.
- Allows a LAN that has a limited number of IPv4 public address allow the connectivity of many internal host to the Internet.
  - The available IPv4 addresses are called the address pool.
  - A packet passing from a private network to a public network will have its IPV4 source address (and UDP/TCP port) changed to one of the available IPv4 public addresses (and ports).
  - That change will be store on the device on the boundary between the private and public network (Router, Firewall or Security Appliance).
    - Its called mapping or translation table.
  - The answer to that packet will have a reverse change.

NAT/PAT Mapping

- Dynamic Mapping:
  - The choice of public address (and port) and mapping to the private address (and port) is done automatically by the Router when it receives a packet from an inside host.
  - An external host cannot initiate a conversation with a inside host.
    - May respond to conversation initiated from an inside host.
- Static Mapping:
  - The choice of public address (and port) and mapping to the private address (and port) is done by configuration.
  - Allows an external host to initiate a conversation with an internal host with a private address.
    - External host contacts the public address/port statically mapped to the private address/port.
    - Scenario: Internal server with a private address (e.g., home game server).

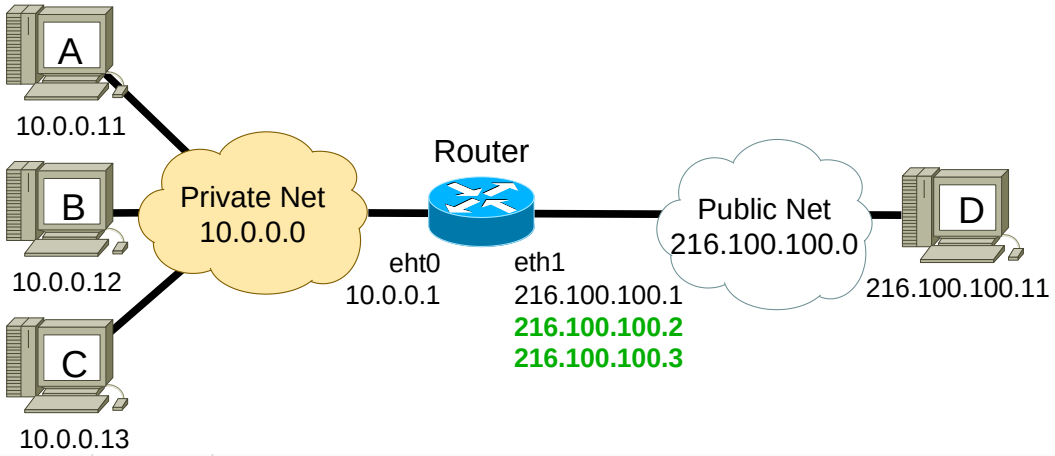


Example – NAT (1)



- Router configures with Dynamic NAT.
- Public IPv4 addresses:
  - 216.100.100.2 and 216.100.100.3 to NAT mappings,
  - 216.100.100.1 to be used by the interface.
    - The IPv4 on the interface may also be used for mapping.

Example – NAT (2)



Ping from 10.0.0.11 to 216.100.100.11:

No.	Time	Source	Destination	Protocol	Length	Info
6	15.892528	10.0.0.11	216.100.100.11	ICMP	98	Echo (ping) request id=0x6c3a, seq=1/256, ttl=64
7	15.911436	216.100.100.11	10.0.0.11	ICMP	98	Echo (ping) reply id=0x6c3a, seq=1/256, ttl=63
8	16.912087	10.0.0.11	216.100.100.11	ICMP	98	Echo (ping) request id=0x6d3a, seq=2/512, ttl=64
9	16.932449	216.100.100.11	10.0.0.11	ICMP	98	Echo (ping) reply id=0x6d3a, seq=2/512, ttl=63
10	17.933103	10.0.0.11	216.100.100.11	ICMP	98	Echo (ping) request id=0x6f3a, seq=3/768, ttl=64
11	17.952490	216.100.100.11	10.0.0.11	ICMP	98	Echo (ping) reply id=0x6f3a, seq=3/768, ttl=63
12	18.954005	10.0.0.11	216.100.100.11	ICMP	98	Echo (ping) request id=0x703a, seq=4/1024, ttl=64
13	18.974316	216.100.100.11	10.0.0.11	ICMP	98	Echo (ping) reply id=0x703a, seq=4/1024, ttl=63
14	19.975028	10.0.0.11	216.100.100.11	ICMP	98	Echo (ping) request id=0x713a, seq=5/1280, ttl=64
15	19.986293	216.100.100.11	10.0.0.11	ICMP	98	Echo (ping) reply id=0x713a, seq=5/1280, ttl=63

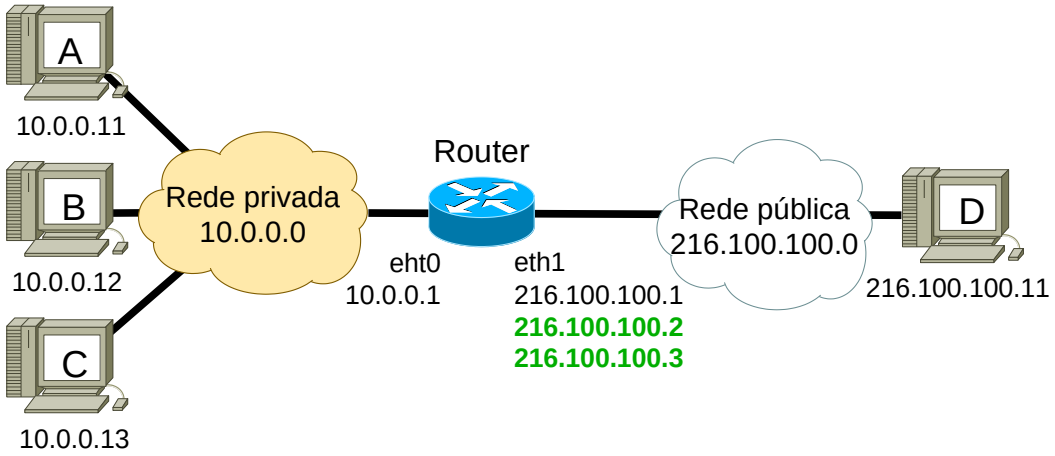
Private Network

No.	Time	Source	Destination	Protocol	Length	Info
2	3.913049	216.100.100.2	216.100.100.11	ICMP	98	Echo (ping) request id=0x6c3a, seq=1/256, ttl=63
3	3.913320	216.100.100.11	216.100.100.2	ICMP	98	Echo (ping) reply id=0x6c3a, seq=1/256, ttl=64
4	4.934041	216.100.100.2	216.100.100.11	ICMP	98	Echo (ping) request id=0x6d3a, seq=2/512, ttl=63
5	4.934405	216.100.100.11	216.100.100.2	ICMP	98	Echo (ping) reply id=0x6d3a, seq=2/512, ttl=64
6	5.954132	216.100.100.2	216.100.100.11	ICMP	98	Echo (ping) request id=0x6f3a, seq=3/768, ttl=63
7	5.954324	216.100.100.11	216.100.100.2	ICMP	98	Echo (ping) reply id=0x6f3a, seq=3/768, ttl=64
8	6.975911	216.100.100.2	216.100.100.11	ICMP	98	Echo (ping) request id=0x703a, seq=4/1024, ttl=63
9	6.976473	216.100.100.11	216.100.100.2	ICMP	98	Echo (ping) reply id=0x703a, seq=4/1024, ttl=64
10	7.987741	216.100.100.2	216.100.100.11	ICMP	98	Echo (ping) request id=0x713a, seq=5/1280, ttl=63
11	7.988265	216.100.100.11	216.100.100.2	ICMP	98	Echo (ping) reply id=0x713a, seq=5/1280, ttl=64

Public Network

```
Router#show ip nat translation
Pro Inside global      Inside local       Outside local      Outside global
--- 216.100.100.2      10.0.0.11          ---                ---
```

Example – NAT (3)



Ping from 10.0.0.12 to 216.100.100.11:

No.	Time	Source	Destination	Protocol	Length	Info
53	311.240021	10.0.0.12	216.100.100.11	ICMP	98	Echo (ping) request id=0x943b, seq=1/256, ttl=64
54	311.258670	216.100.100.11	10.0.0.12	ICMP	98	Echo (ping) reply id=0x943b, seq=1/256, ttl=63
56	312.259967	10.0.0.12	216.100.100.11	ICMP	98	Echo (ping) request id=0x953b, seq=2/512, ttl=64
57	312.280140	216.100.100.11	10.0.0.12	ICMP	98	Echo (ping) reply id=0x953b, seq=2/512, ttl=63
58	313.281645	10.0.0.12	216.100.100.11	ICMP	98	Echo (ping) request id=0x963b, seq=3/768, ttl=64
59	313.302003	216.100.100.11	10.0.0.12	ICMP	98	Echo (ping) reply id=0x963b, seq=3/768, ttl=63
60	314.303181	10.0.0.12	216.100.100.11	ICMP	98	Echo (ping) request id=0x973b, seq=4/1024, ttl=64
61	314.323635	216.100.100.11	10.0.0.12	ICMP	98	Echo (ping) reply id=0x973b, seq=4/1024, ttl=63
62	315.325157	10.0.0.12	216.100.100.11	ICMP	98	Echo (ping) request id=0x983b, seq=5/1280, ttl=64
63	315.345519	216.100.100.11	10.0.0.12	ICMP	98	Echo (ping) reply id=0x983b, seq=5/1280, ttl=63

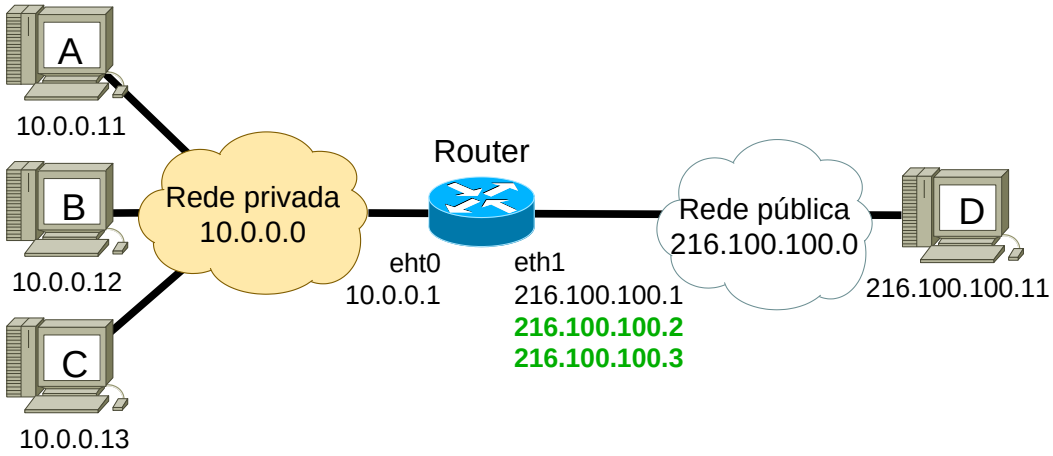
Private Network

No.	Time	Source	Destination	Protocol	Length	Info
47	299.260334	216.100.100.3	216.100.100.11	ICMP	98	Echo (ping) request id=0x943b, seq=1/256, ttl=63
48	299.260929	216.100.100.11	216.100.100.3	ICMP	98	Echo (ping) reply id=0x943b, seq=1/256, ttl=64
50	300.281677	216.100.100.3	216.100.100.11	ICMP	98	Echo (ping) request id=0x953b, seq=2/512, ttl=63
51	300.282286	216.100.100.11	216.100.100.3	ICMP	98	Echo (ping) reply id=0x953b, seq=2/512, ttl=64
52	301.303570	216.100.100.3	216.100.100.11	ICMP	98	Echo (ping) request id=0x963b, seq=3/768, ttl=63
53	301.304103	216.100.100.11	216.100.100.3	ICMP	98	Echo (ping) reply id=0x963b, seq=3/768, ttl=64
54	302.325227	216.100.100.3	216.100.100.11	ICMP	98	Echo (ping) request id=0x973b, seq=4/1024, ttl=63
55	302.325755	216.100.100.11	216.100.100.3	ICMP	98	Echo (ping) reply id=0x973b, seq=4/1024, ttl=64
56	303.347148	216.100.100.3	216.100.100.11	ICMP	98	Echo (ping) request id=0x983b, seq=5/1280, ttl=63
57	303.347704	216.100.100.11	216.100.100.3	ICMP	98	Echo (ping) reply id=0x983b, seq=5/1280, ttl=64

Public Network

```
Router#show ip nat translation
Pro Inside global      Inside local       Outside local      Outside global
--- 216.100.100.2      10.0.0.11          ---                ---
--- 216.100.100.3      10.0.0.12          ---                ---
```

Example – NAT (4)



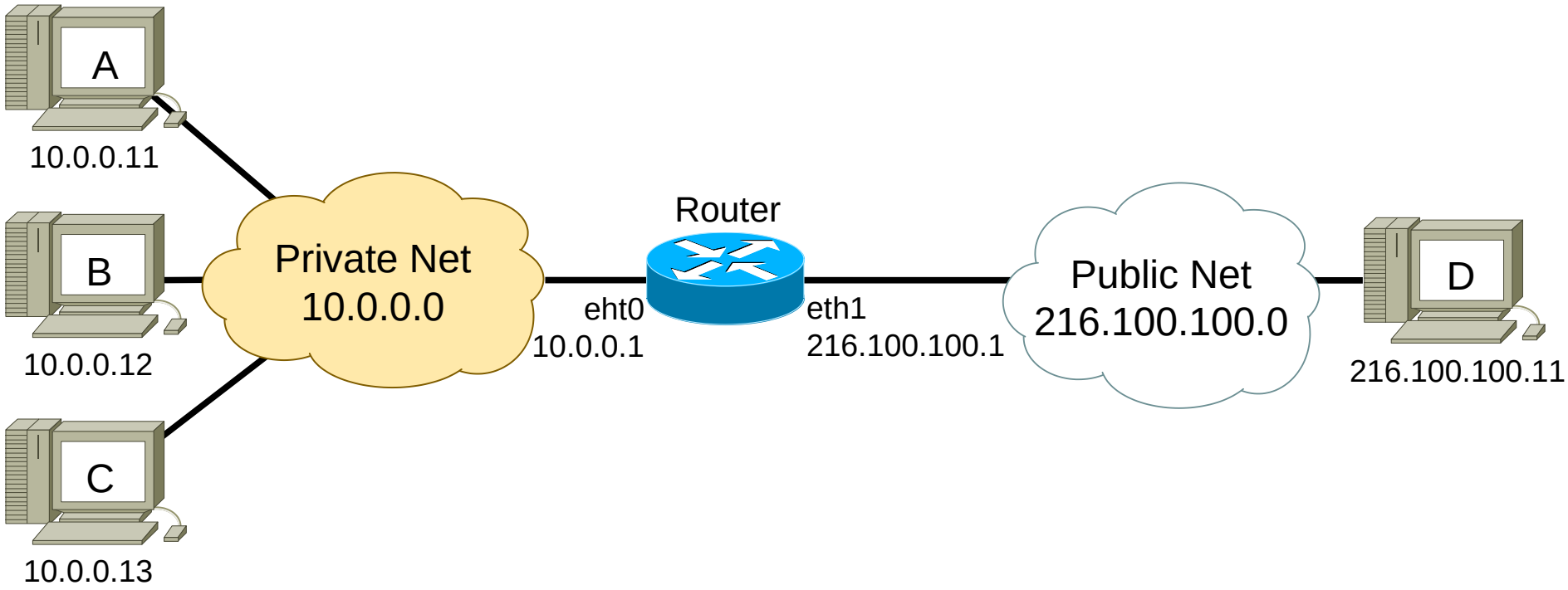
Ping from 10.0.0.13 to 216.100.100.11:

No.	Time	Source	Destination	Protocol	Length	Info
113	506.016226	10.0.0.13	216.100.100.11	ICMP	98	Echo (ping) request id=0x573c, seq=1/256, ttl=64
114	506.035020	10.0.0.1	10.0.0.13	ICMP	70	Destination unreachable (Host unreachable)
115	507.036014	10.0.0.13	216.100.100.11	ICMP	98	Echo (ping) request id=0x583c, seq=2/512, ttl=64
116	507.046188	10.0.0.1	10.0.0.13	ICMP	70	Destination unreachable (Host unreachable)
117	508.047177	10.0.0.13	216.100.100.11	ICMP	98	Echo (ping) request id=0x593c, seq=3/768, ttl=64
118	508.057193	10.0.0.1	10.0.0.13	ICMP	70	Destination unreachable (Host unreachable)
119	509.058553	10.0.0.13	216.100.100.11	ICMP	98	Echo (ping) request id=0x5a3c, seq=4/1024, ttl=64
120	509.068436	10.0.0.1	10.0.0.13	ICMP	70	Destination unreachable (Host unreachable)
121	510.069971	10.0.0.13	216.100.100.11	ICMP	98	Echo (ping) request id=0x5b3c, seq=5/1280, ttl=64
122	510.079907	10.0.0.1	10.0.0.13	ICMP	70	Destination unreachable (Host unreachable)

Private Network

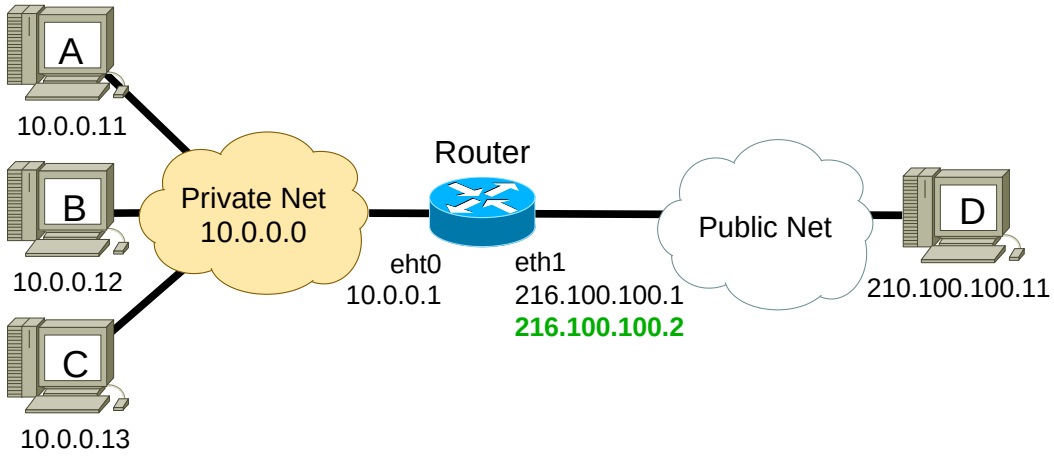
- Host C (10.0.0.13) cannot access the public network.
  - All IPv4 public address available on the Router have been mapped to Host A and Host B.
- All NAT mappings have a limited lifetime (*timeout).*
  - After some time without traffic to the public network the mappings will be deleted.

Example – NAT/PAT (1)



- Host D has a UDP server (ECHO) on port 5005.
- Public IPv4 addresses:
  - 216.100.100.2 and 216.100.100.3 to NAT mappings,
  - 216.100.100.1 to be used by the interface.

Example – NAT/PAT (2)



Hosts A, B and C access Host D (UDP Port 5005):

Source	Destination	Protocol	Length	Info
10.0.0.11	216.100.100.11	UDP	98	22147 → 5005
216.100.100.11	10.0.0.11	UDP	98	5005 → 22147
10.0.0.11	216.100.100.11	UDP	98	22147 → 5005
216.100.100.11	10.0.0.11	UDP	98	5005 → 22147

Source	Destination	Protocol	Length	Info
10.0.0.12	216.100.100.11	UDP	98	40521 → 5005
216.100.100.11	10.0.0.12	UDP	98	5005 → 40521
10.0.0.12	216.100.100.11	UDP	98	40521 → 5005
216.100.100.11	10.0.0.12	UDP	98	5005 → 40521

Source	Destination	Protocol	Length	Info
10.0.0.13	216.100.100.11	UDP	98	61252 → 5005
216.100.100.11	10.0.0.13	UDP	98	5005 → 61252
10.0.0.13	216.100.100.11	UDP	98	61252 → 5005
216.100.100.11	10.0.0.13	UDP	98	5005 → 61252

Private Network

Source	Destination	Protocol	Length	Info
216.100.100.2	216.100.100.11	UDP	98	1024 → 5005
216.100.100.11	216.100.100.2	UDP	98	5005 → 1024
216.100.100.2	216.100.100.11	UDP	98	1024 → 5005
216.100.100.11	216.100.100.2	UDP	98	5005 → 1024

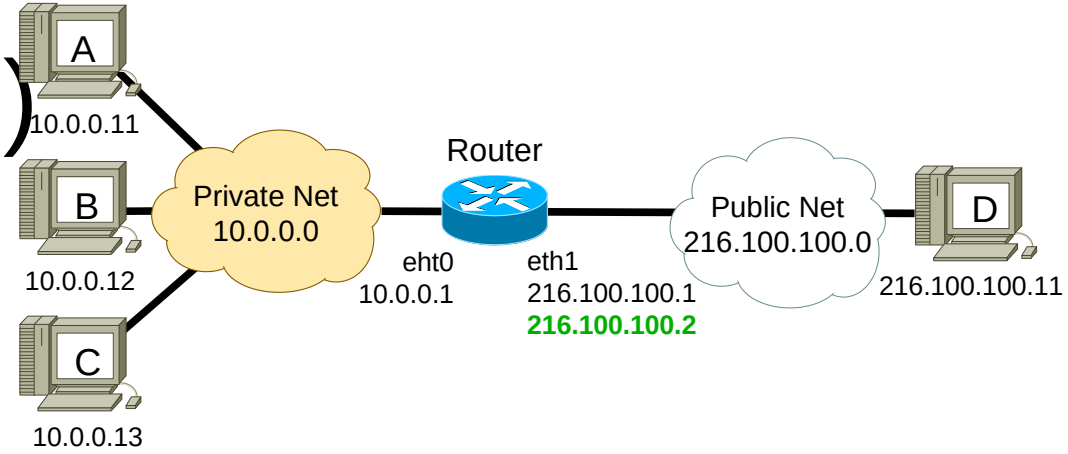
Source	Destination	Protocol	Length	Info
216.100.100.2	216.100.100.11	UDP	98	1025 → 5005
216.100.100.11	216.100.100.2	UDP	98	5005 → 1025
216.100.100.2	216.100.100.11	UDP	98	1025 → 5005
216.100.100.11	216.100.100.2	UDP	98	5005 → 1025

Source	Destination	Protocol	Length	Info
216.100.100.2	216.100.100.11	UDP	98	1026 → 5005
216.100.100.11	216.100.100.2	UDP	98	5005 → 1026
216.100.100.2	216.100.100.11	UDP	98	1026 → 5005
216.100.100.11	216.100.100.2	UDP	98	5005 → 1026

Public Network

- Mapping choices by the Router depends on local algorithm, is not defined by standards.
- All hosts were mapped to IPv4 216.100.100.2.
  - Host A used the UDP client port 22147, and was mapped to port 1024.
  - Host B used the UDP client port 40521, and was mapped to port 1025.
  - Host C used the UDP client port 61252, and was mapped to port 1026.

Example – NAT/PAT (3)



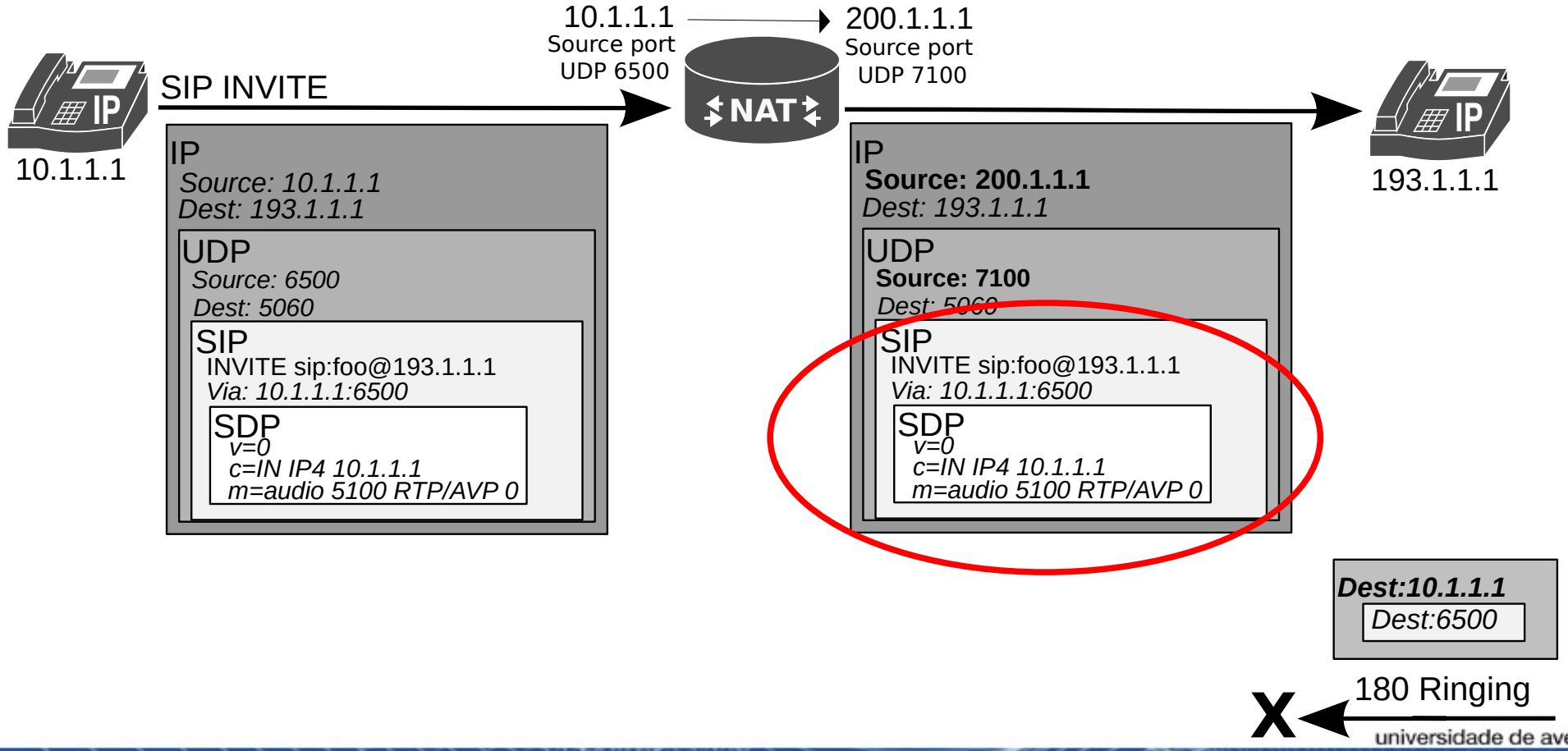
Hosts A, B and C access Host D (UDP Port 5005):

```
Router#show ip nat translation
Pro Inside global      Inside local       Outside local      Outside global
udp 216.100.100.2:1024 10.0.0.11:22147    216.100.100.11:5005 216.100.100.11:5005
udp 216.100.100.2:1025 10.0.0.12:40521    216.100.100.11:5005 216.100.100.11:5005
udp 216.100.100.2:1026 10.0.0.13:61252    216.100.100.11:5005 216.100.100.11:5005
```

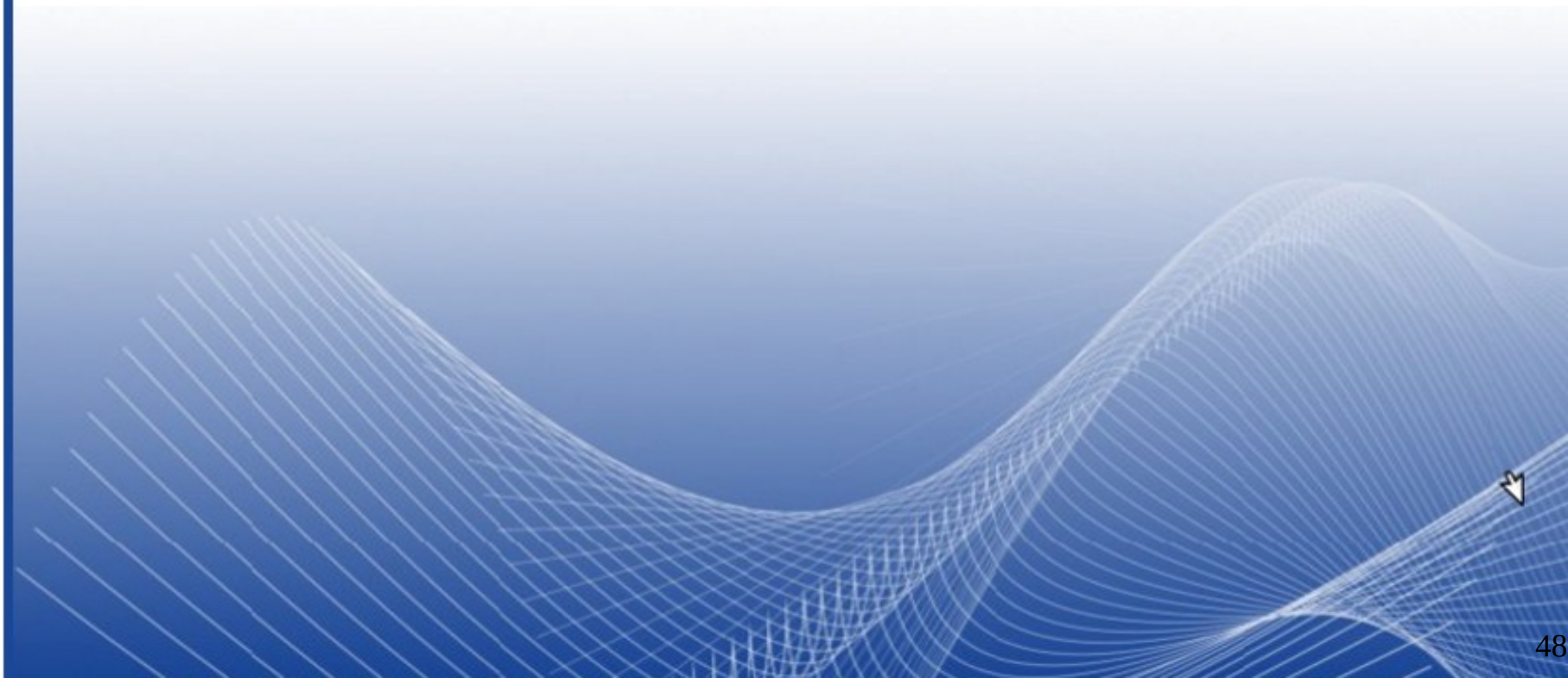
- All hosts were mapped to IPv4 address 216.100.100.2.
- Host A used the UDP client port 22147, and was mapped to port 1024.
- Host B used the UDP client port 40521, and was mapped to port 1025.
- Host C used the UDP client port 61252, and was mapped to port 1026.

Some Protocols Require Translation at the Application Level

- Some protocols (e.g., SIP) require the translation of addresses and ports also at the application protocol level.
  - Very computational demanding and not all devices allow it.



IPv6 Addressing



IPv6 Background

- ETF IPv6 WG began to work on a solution to solve addressing growth issues in early 1990s
- Reasons to late deployment
  - Classless Inter-Domain Routing (CIDR) and Network address translation (NAT) were developed
  - Investments on field equipments (not IPv6 aware) had to reach the predicted “return of investment”
  - Massive re-equipment price

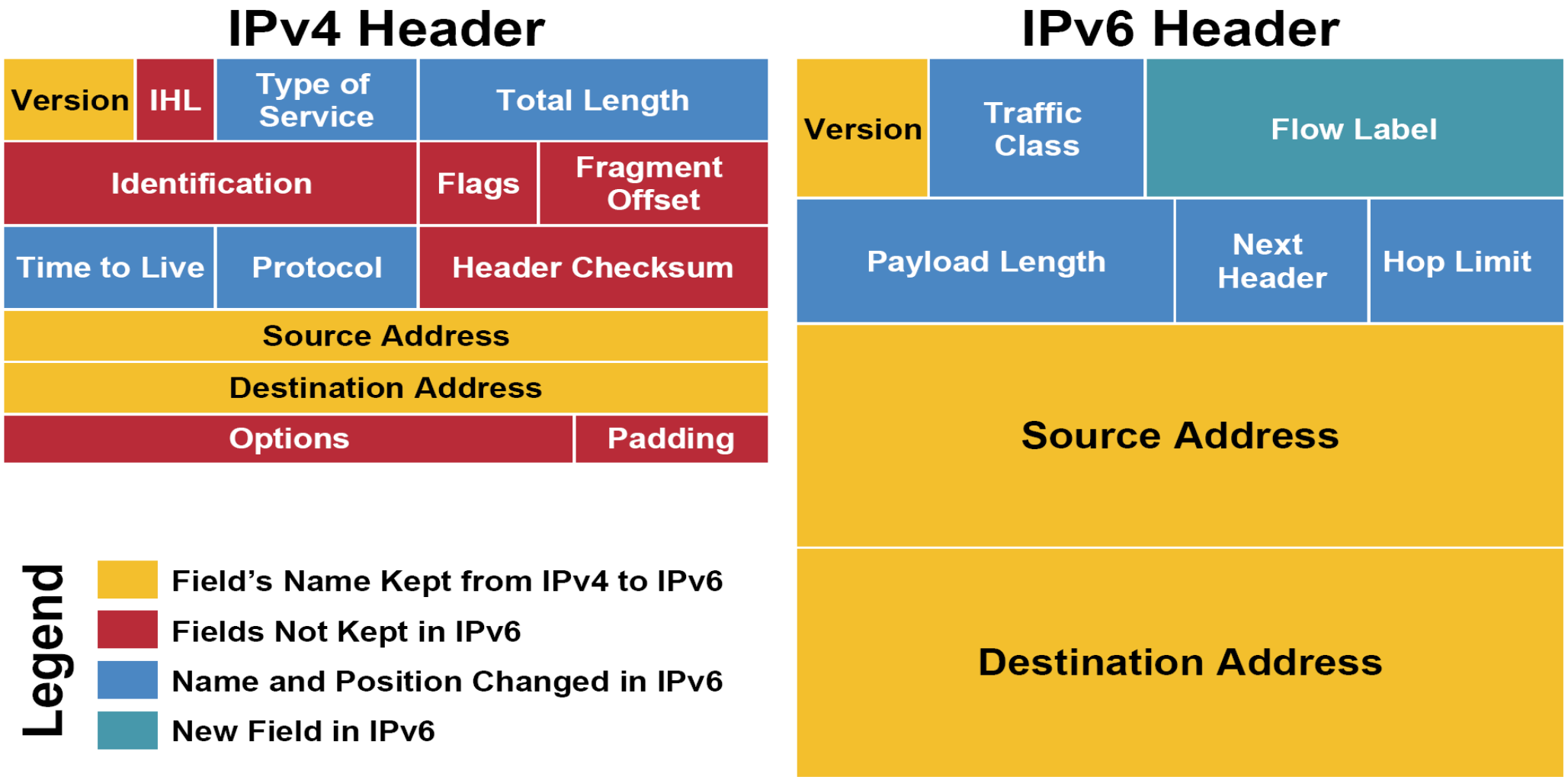
IPv6 Features

- Larger address space enabling:
  - Global reachability, flexibility, aggregation, multihoming, autoconfiguration, “plug and play” and renumbering
- Simpler header enabling:
- Routing efficiency, performance and forwarding rate scalability
- Improved option support

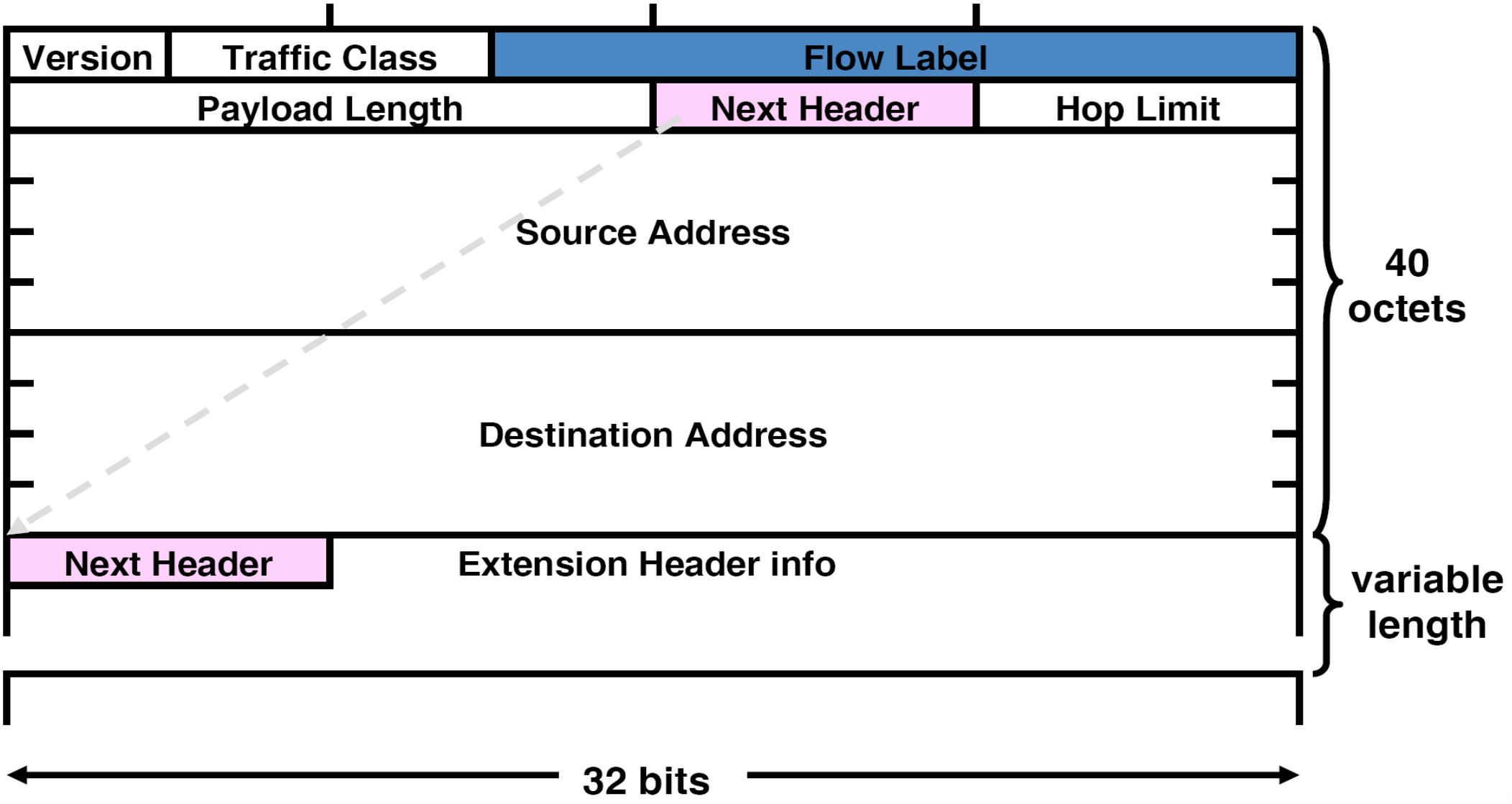
IPv6 Addressing

- IPv4: 4bytes/32 bits
  - ~ 4,294,967,296 possible addresses
- IPv6: 16bytes/128 bits
  - 340,282,366,920,938,463,463,374,607,431,768,211,456 possible addresses
- Representation
  - 16-bit hexadecimal numbers
    - Hex numbers are not case sensitive
  - Numbers are separated by (:)
    - Abbreviations are possible
      - Leading zeros in contiguous block could be represented by (::)
      - Example:
      - 2001:0db8:0000:130F:0000:0000:087C:140B = 2001:0db8:0:130F::87C:140B
      - Double colon only appears once in the address
  - Address's prefix is represented as: prefix/mask_number_of_bits

IPv4 vs. IPv6 Headers



IPv6 Header Format



IPv6 Addressing Model

- Interface have multiple addresses
- Addresses have scope:
  - Link Local
    - Valid within the same LAN or link
  - Unique Local
    - Valid within the same private domain
    - Can not be used in Internet
  - Global
- Addresses have lifetime
  - Valid and preferred lifetime

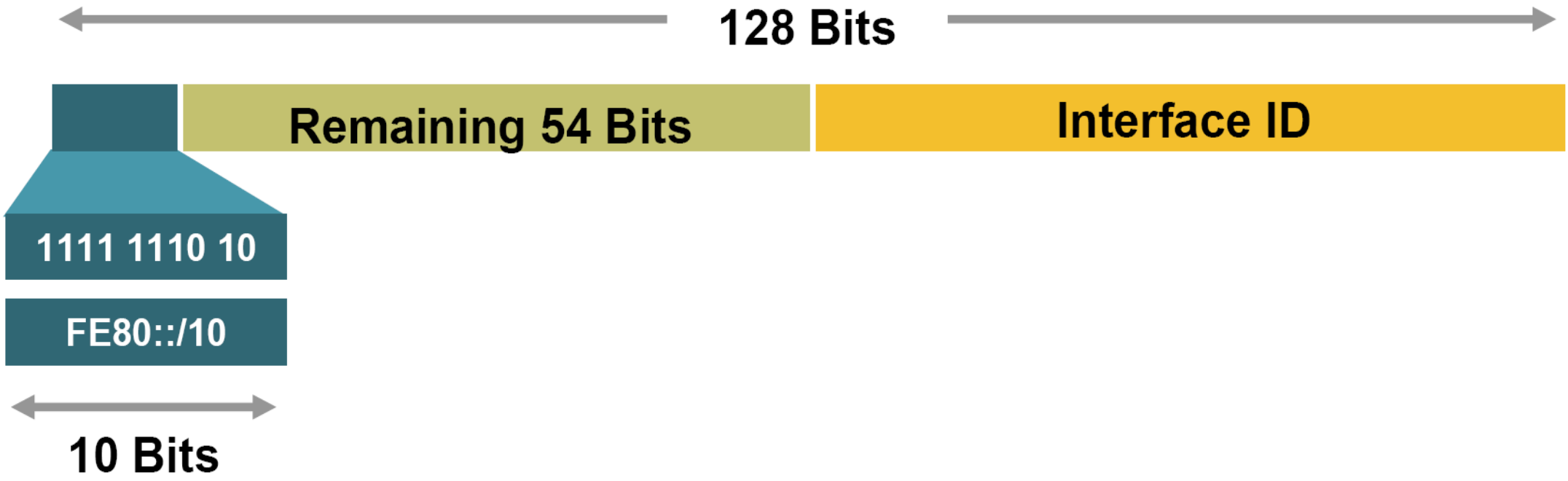
Types of IPv6 Addresses

- Unicast
  - Address of a single interface.
  - One-to-one delivery to single interface
- Multicast
  - Address of a set of interfaces.
  - One-to-many delivery to all interfaces in the set
- Anycast
  - Address of a set of interfaces.
  - One-to-one-of-many delivery to a single interface in the set that is closest
- No more broadcast addresses

IPv6 Addressing

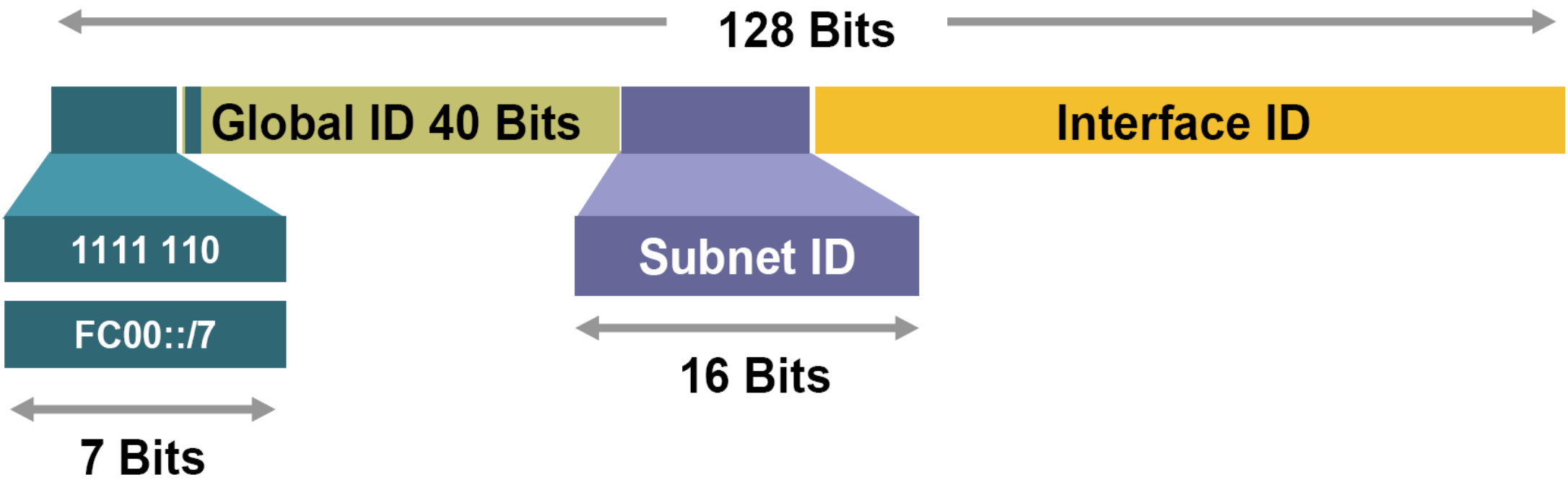
Type	Binary	Hexadecimal
Global Unicast Address	0010	2
Link-Local Unicast Address	1111 1110 10	FE80::/10
Unique-Local Unicast Address	1111 1100 1111 1101	FC00::/8 FD00::/8
Multicast Address	1111 1111	FF00::/16

Link-Local Address



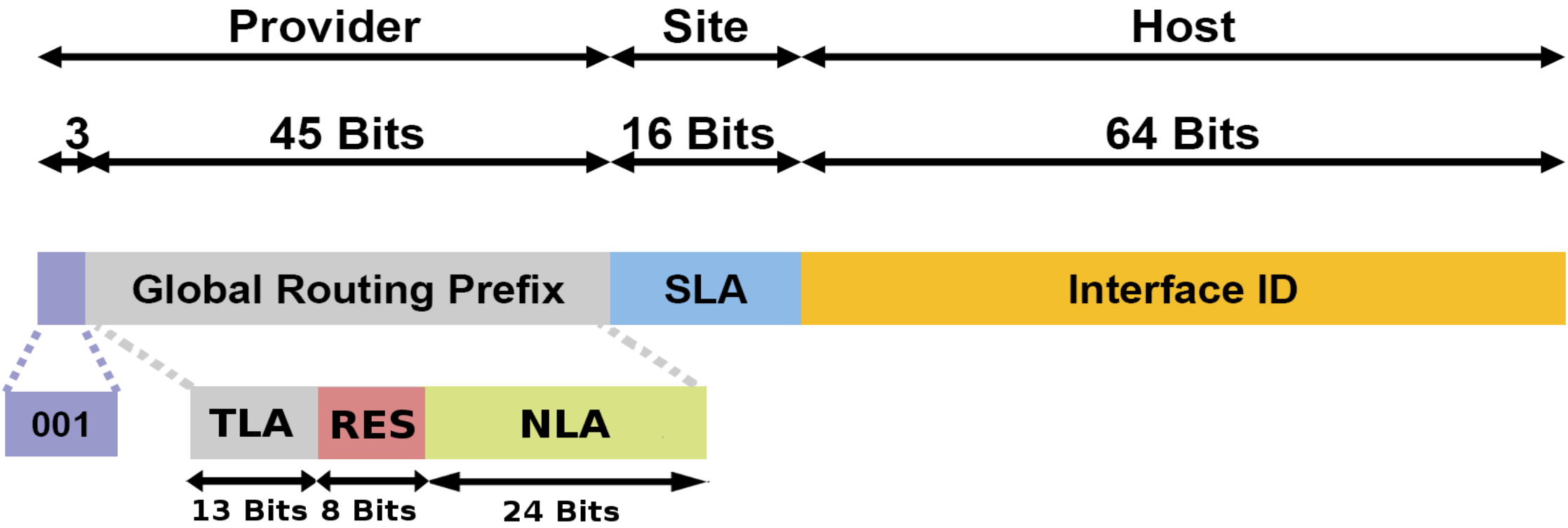
- Used For:
  - Mandatory address for local communication between two IPv6 devices
  - Next-Hop calculation in Routing Protocols
- Automatically assigned as soon as IPv6 is enabled
- Remaining 54 bits could be Zero or any manual configured value

Unique-Local Address



- Used For:
  - Local communications
  - Inter-site VPNs
- Can be routed only within the same Autonomous System
  - Can not be used on the Internet

Global Unicast Addresses



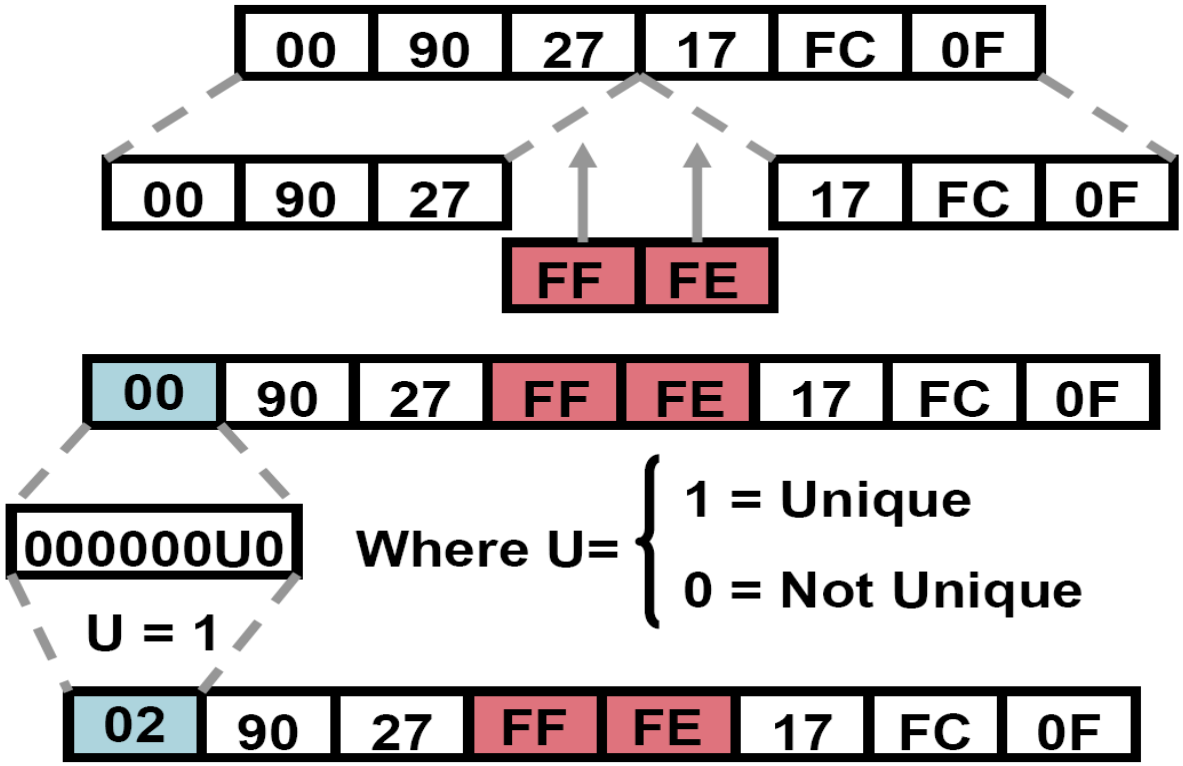
- LA, NLA and SLA used for hierarchical addressing
  - TLA - Top-Level Aggregation
  - RES – Reserved (must be zero)
  - NLA - Next-Level Aggregation Identifier
  - SLA - Site-Level Aggregation Identifier

IPv6 Interface Identifier

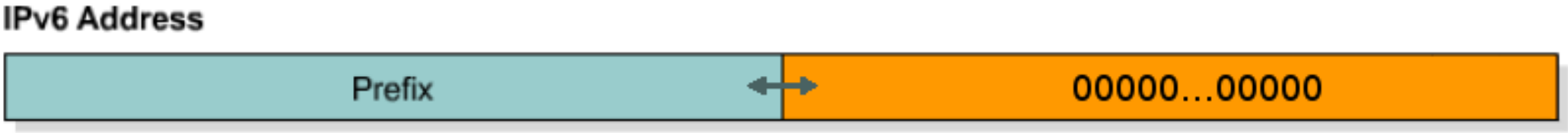
- Lowest-Order 64-Bit field of any address:
  - Auto-configured from a 64-bit EUI-64, or expanded from a 48-bit MAC address (e.g. Ethernet address)
  - Auto-generated pseudo-random number
  - Assigned via DHCP
  - Manually configured

MAC to Interface ID (EUI-64 format)

- Stateless auto-configuration
- Expands the 48 bit MAC address to 64 bits by inserting FFFE into the middle 16 bits
- To make sure that the chosen address is from a unique Ethernet MAC address
  - “u”bit is set to 1 for global scope
  - “u”bit is set to 0 for local scope



Anycast Address



- Address that is assigned to a set of interfaces
  - Typically belong to different nodes
- A packet sent to an Anycast address is delivered to the closest interface (determined by routing and timings)
- Anycast addresses can be used only by routers, not hosts
- Must not be used as the source address of an IPv6 packet
- Nodes to which the anycast address is assigned must be explicitly configured to recognize that the address is an Anycast address

Multicast Addresses

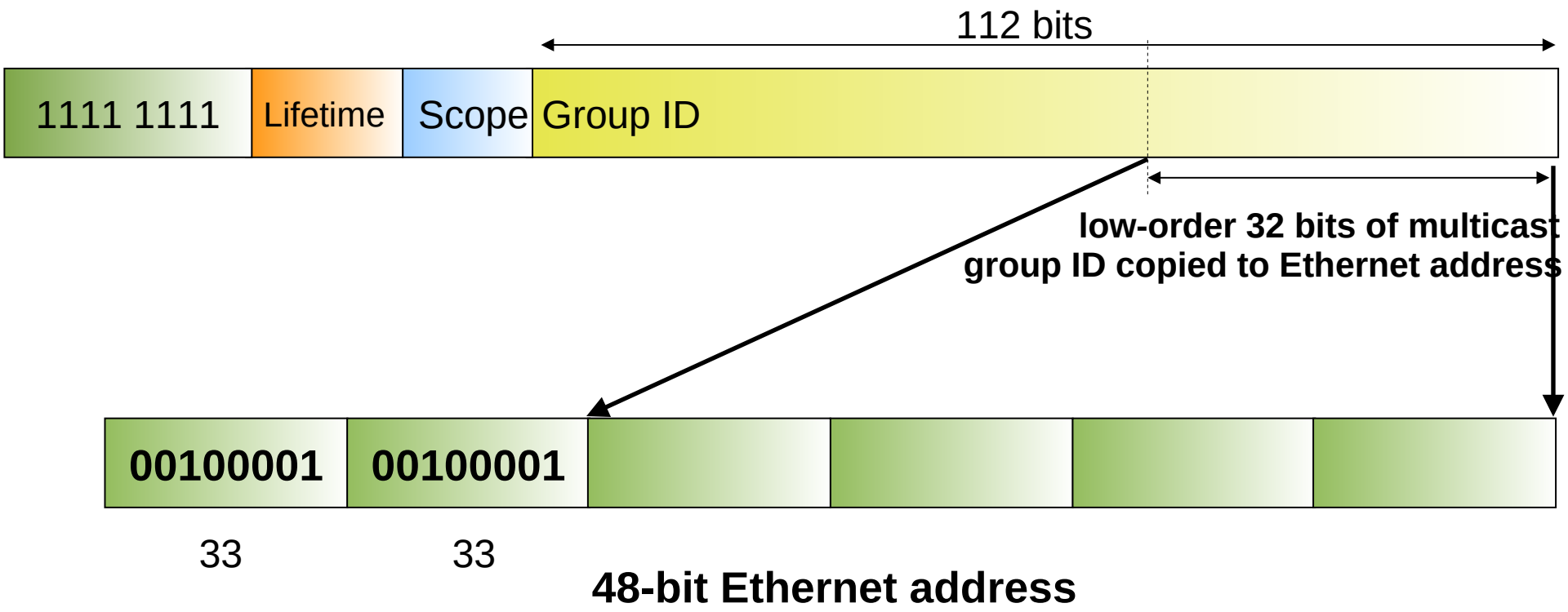
8-bit	4-bit	4-bit	112-bit
1111 1111	Lifetime	Scope	Group-ID

Lifetime	
0	If Permanent
1	If Temporary

Scope	
1	Node
2	Link
5	Site
8	Organization
E	Global

- Multicast addresses have a prefix FF00::/8
- The second byte defines the lifetime and scope of the multicast address.

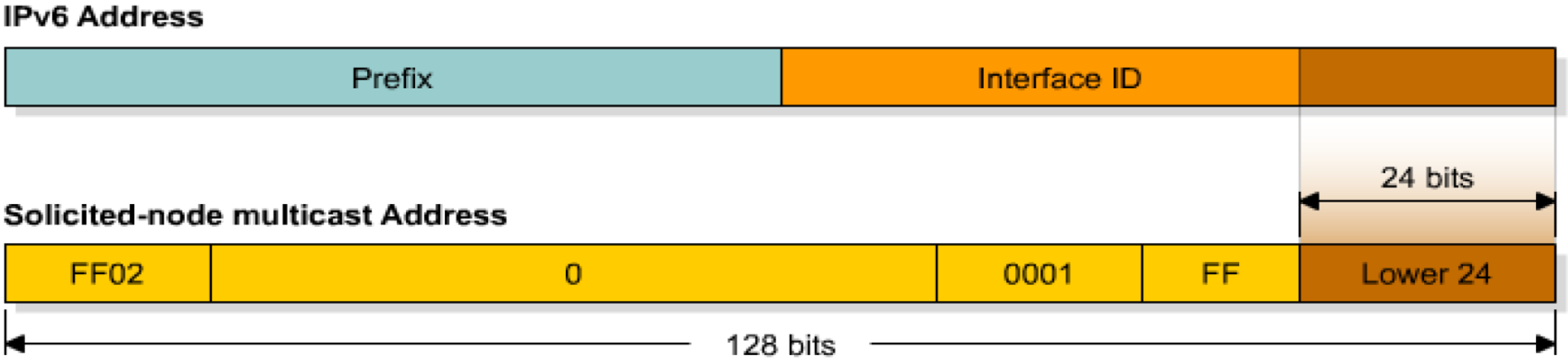
Mapping a IPv6 Multicast Address to Ethernet Address



Common Multicast Addresses

- Node Scope
  - FF01:::1 All Nodes Address (Node scope)
  - FF01:::2 All Routers Address (Node scope)
- Link Scope
  - FF02::1 All Nodes Address (Node scope)
  - FF02::2 All Routers Address
  - FF02::4 DVMRP Routers
  - FF02::5 OSPF IGP
  - FF02::6 OSPF IGP Designated Routers
  - FF02::9 RIP Routers
  - FF02::B Mobile-Agents
  - FF02::D All PIM Routers
  - FF02::E RSVP-ENCAPSULATION
  - FF02::16 All MLDv2-capable routers
  - FF02:::1:2 All DHCP agents

Solicited-Node Multicast Address



- For each unicast and anycast address configured there is a corresponding solicited-node multicast
- FF02::1:FF:<interface ID's lower 24 bits>
- This address has link local significance only
- Used in “Neighbour Solicitation Messages”
  - MAC/Physical addresses resolution
  - Duplicate Address Detection (DAD)
    - Random or assigned interface IDs may result in equal global/link addresses

Physical Addresses Resolution

- In IPv6 ARP does not exist anymore.
- ARP table is now called **NDP table**
  - NDP: Neighbor Discovery Protocol
  - Maintains a list of known neighbors (IPv6 addresses and MAC addresses).
- Uses ICMPv6 “Neighbor Solicitation” and “Neighbor Advertisement” messages.
  - To resolve an address a Neighbor Solicitation message is sent to the Solicited-Node multicast address of the target machine (IPv6 address).
  - Response is sent in unicast using a Neighbor Advertisement message.

ICMPv6

- Internet Control Message Protocol version 6 (ICMPv6) is the implementation ICMP for IPv6
  - RFC 4443
  - ICMPv6 is an integral part of IPv6.
- Have the same functionalities of ICMP, plus:
  - Replaces and enhances ARP,
    - ICMPv6 implements a Neighbor Discovery Protocol (NDP),
  - Hosts use it to discover routers and perform auto configuration of addresses,
  - Used to perform Duplicate Address Detection (DAD),
  - Used to test reachability of neighbors.

Neighbor Discovery

- Neighbor discovery uses ICMPv6 messages, originated from node on link local with hop limit of 255
- Consists of IPv6 header, ICMPv6 header, neighbor discovery header, and neighbor discovery options
- Five neighbor discovery messages
  - Router solicitation (ICMPv6 type 133)
  - Router advertisement (ICMPv6 type 134)
  - Neighbor solicitation (ICMPv6 type 135)
  - Neighbor advertisement (ICMPv6 type 136)
  - Redirect (ICMPV6 type 137)

Router Solicitation

- Host send to inquire about presence of a router on the link
- Send to all routers multicast address of FF02::2 (all routers multicast address)
- Source IP address is either link local address or unspecified IPv6 address

Router advertisement

- Sent out by routers periodically, or in response to a router solicitation
- Includes auto-configuration information
- Includes a "preference level" for each advertised router address
- Also includes a "lifetime" field

Neighbor Solicitation

- Send to discover link layer address of IPv6 node
- IPv6 header, source address is set to unicast address of sending node, or :: for DAD
- Destination address is set to
  - Unicast address for reachability
  - Solicited node multicast for address resolution and DAD

Neighbor Advertisement

- Response to neighbor solicitation message
- Also send to inform change of link layer address

Redirect

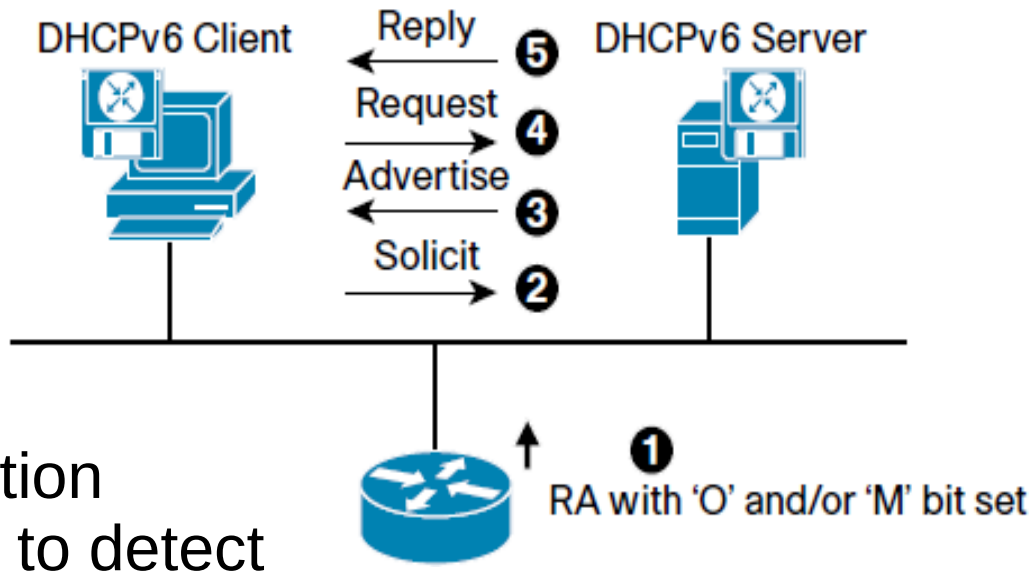
- Redirect is used by a router to signal the reroute of a packet to a better router

Auto-configuration

- Stateless
  - A node on the link can automatically configure global IPv6 addresses by appending its interface identifier (64 bits) to the prefixes (64 bits) included in the Router Advertisement messages
  - Additional/Other network information may be obtained
    - Additional fields in Router Advertisement messages,
    - Using a stateless DHCPv6 server.
- Stateful
  - Addresses are obtained using DHCPv6.

- The default gateway may send two configurable flags in Router Advertisements (RA)
  - Other flag bit: client can use DHCPv6 to retrieve other configuration parameters (e.g.: DNS server addresses)
  - Managed flag bit: client may use DHCPv6 to retrieve a Managed IPv6 address from a server

DHCPv6



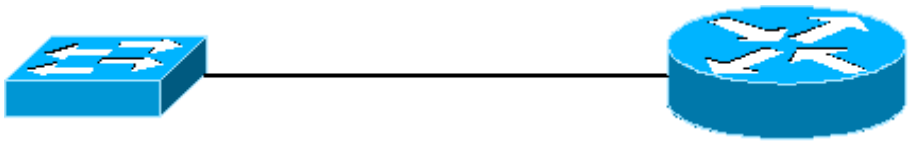
- Basic DHCPv6 concept is similar to DHCP for IPv4.
- If a client wishes to receive configuration parameters, it will send out a request to detect available DHCPv6 servers.
  - This done through the “Solicit” and “Advertise” messages.
  - Well known DHCPv6 Multicast addresses are used for this process.
- Next, the DHCPv6 client will ”Request” parameters from an available server which will respond with the requested information with a “Reply” message.
- DHCPv6 relaying works differently from DHCP for IPv4 relaying
  - Relay agent will encapsulate the received messages from the directly connected DHCPv6 client (RELAY-FORW message)
  - Forward these encapsulated DHCPv6 packets towards the DHCPv6 server.
  - In the opposite direction, the Relay Agent will decapsulate the packets received from the central DHCPv6 Server (RELAY-REPL message).

Multicast Listener Discovery (MLD)

- MLD permits the creation/management of multicast groups
- MLD is used by an IPv6 router to:
  - Discover the presence of multicast listeners on directly attached links
  - And to discover which multicast addresses are of interest to those neighboring nodes
  - Report interest in router specific multicast addresses
- Routers and hosts use MLD to report interest in respective Solicited-Node Multicast Addresses

- MLD will be studied later in detail.

IPv6 Start-up - Router



Multicast (all MLDv2-capable routers) ff02::16	**MLDv2 Report Message** (Multicast all routers)	**Null address ::**
Multicast (all MLDv2-capable routers) ff02::16	**MLDv2 Report Message** (Multicast solicited-node address)	**Null address ::**
Multicast solicited-node address ff02::1:ff+(address's last 24 bits)	**Neighbor Solicitation** (DAD link-local address)	**Null address ::**
Multicast (all hosts) ff02::1	**Neighbor Advertisement**	**Link-local address fe80::+(interface ID 64-bits)**
Multicast (all MLDv2-capable routers) ff02::16	**MLDv2 Report Message** (Multicast all routers)	**Link-local address fe80::+(interface ID 64-bits)**
Multicast (all MLDv2-capable routers) ff02::16	**MLDv2 Report Message** (Multicast solicited-node address)	**Link-local address fe80::+(interface ID 64-bits)**
Multicast solicited-node address ff02::1:ff+(address's last 24 bits)	**Neighbor Solicitation** (DAD global address)	**Null address ::**
Multicast (all hosts) ff02::1	**Router Advertisement**	**Link-local address fe80::+(interface ID 64-bits)**

Only if global address is configured

IPv6 Start-up – Terminal/Router Interaction

Null address **::**	**Neighbor Solicitation** (DAD link-local address) →	**Multicast solicited-node address** **ff02::1:ff+(address's last 24 bits)**
Link-local address **fe80::+(interface ID 64-bits)**	**Router Solicitation** →	**Multicast (all routers)** **ff02::2**
Null address **::**	**MLDv2 Report Message** (Multicast solicited-node address) →	**Multicast (all MLDv2-capable routers)** **ff02::16**
Multicast (all hosts) **ff02::1**	← **Router Advertisement**	**Link-local address** **fe80::+(interface ID 64-bits)**
Null address **::**	**Neighbor Solicitation** (DAD global address) →	**Multicast solicited-node address** **ff02::1:ff+(address's last 24 bits)**

Address Resolution and Ping6

Global address → **Neighbor Solicitation** (Target's IPv6 address) → **Multicast solicited-node address ff02::1:ff+(address's last 24 bits)**

Global address ← **Neighbor Advertisement** (MAC address) ← **Global address**

Global address → **ICMPv6 Echo Request** → **Global address**

Global address ← **ICMPv6 Echo Reply** ← **Global address**

Global/link-local address → **Neighbor Solicitation** (IPv6 address) → **Global/link-local address**

Global/link-local address ← **Global/link-local address**

To verify the reachability of a neighbor after physical address of a neighbor is identified

IPv6 Subnetting/Aggregation

- In IPv6 the same principles of IPv4 subnetting and aggregation are still valid.
  - Using the TLA, NLA and SLA bits of the IPv6 addresses.
  - Example: network 2001:A:A:/48 can be divided in 2^16 sub-networks with identifiers 2001:A:A:****:/64
- By standard, the maximum mask size is /64, however it is possible to subnet also the host part of the IPv6 address.
  - Usage of mask /120 to protect the network from NDP Table Exhaustion attacks.
    - With mask /120 the maximum size of the NDP table is limited to 2^8.
    - “Larger” masks also work.
  - Some tools/services may break.
  - Point-to-point links may use /126.
    - Some devices accept use /127, however in others may not work.
  - Requires manual, DHCPv6 address configuration or modified auto-configuration mechanisms.

IPv6 Addresses Planning

- Due to IPv6 nature, there are many networks and networks are large.
  - Number of hosts in LAN is not an issue!
  - Usually network managers receive /48 networks:
    - Allows for 2^16 /64 networks.
    - Standard LAN use /64
      - or /120 to protect against attacks, however breaks stateless assignment.
    - Point-to-point links use /126.
      - Usually a /64 network is sub-netted into multiple /126.