Fachhochschule
Dortmund
University of Applied Sciences and Arts

Communications and Computer Networks

Summer Term 2023

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

Recap of last lecture (1/7)

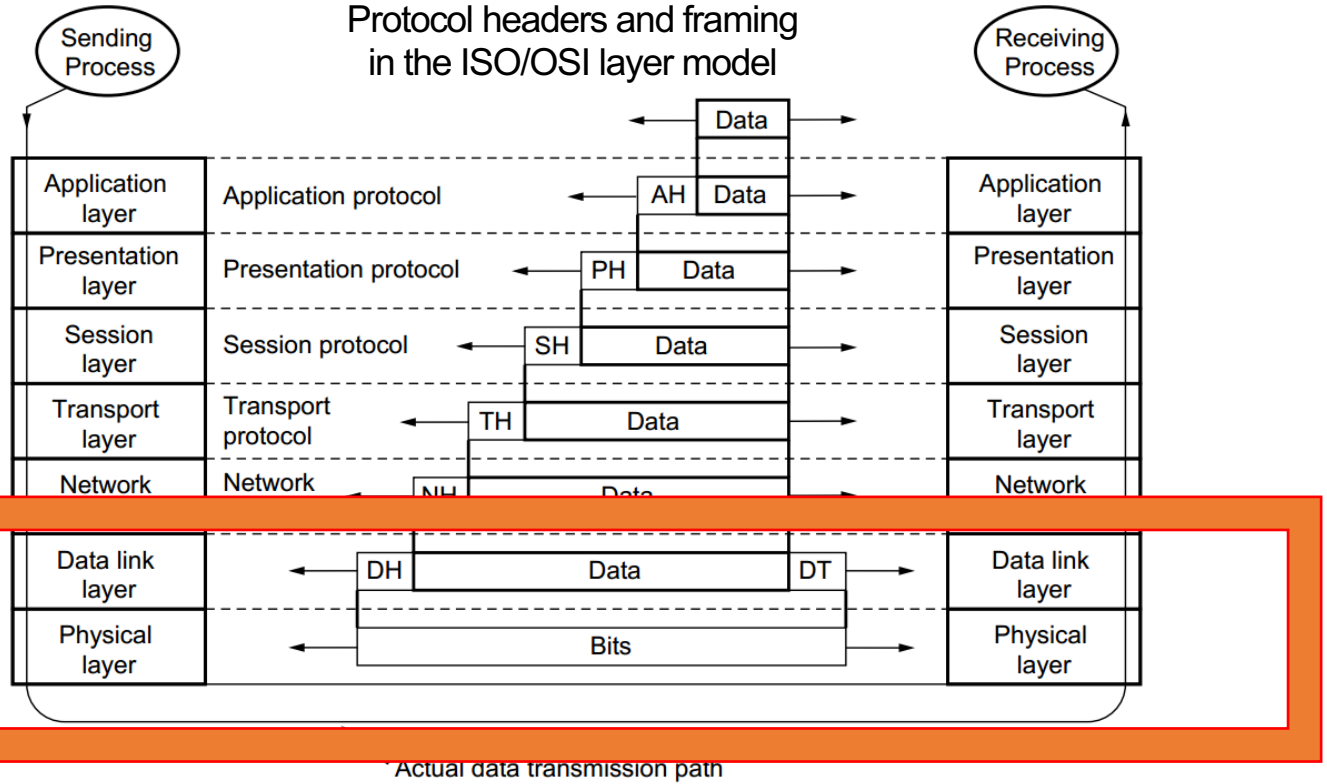
- Data link layer and related topics
  - Fundamental tasks of the data link layer
- Ethernet
  - You can explain the CSMA/CD access procedure
  - You know the structure of an Ethernet frame
  - You know MAC-addresses and its usage in networks
  - You can explain the task and procedure of the Address Resolution Protocol (ARP)
- VLAN based on IEEE 802.1q

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

Fachhochschule Dortmund
University of Applied Sciences and Arts

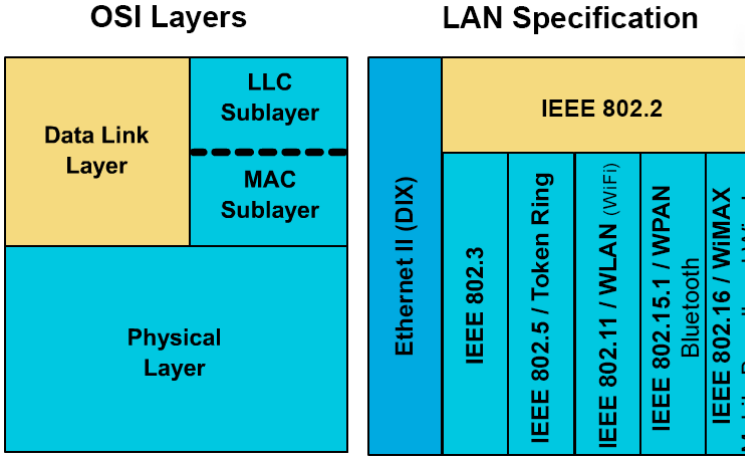
Recap of last lecture (2/7)

Discussion till now:



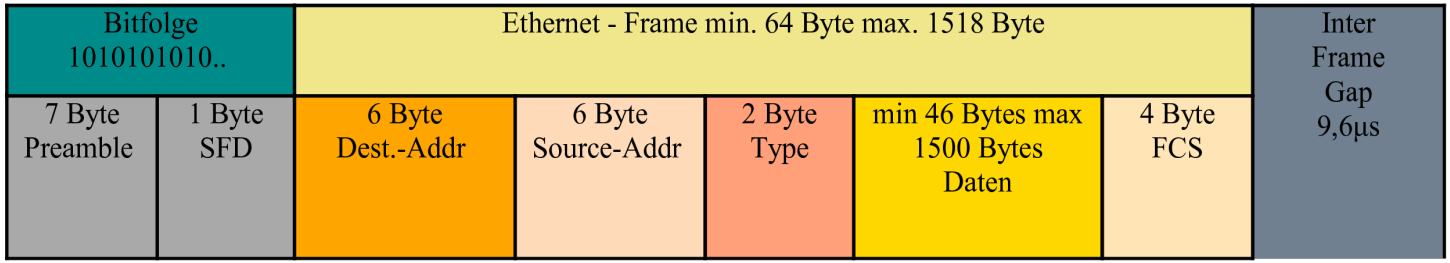
Fachhochschule Dortmund
University of Applied Sciences and Arts

Recap of last lecture (3/5)



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

Recap of last lecture (4/7)



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

Fachhochschule Dortmund
University of Applied Sciences and Arts

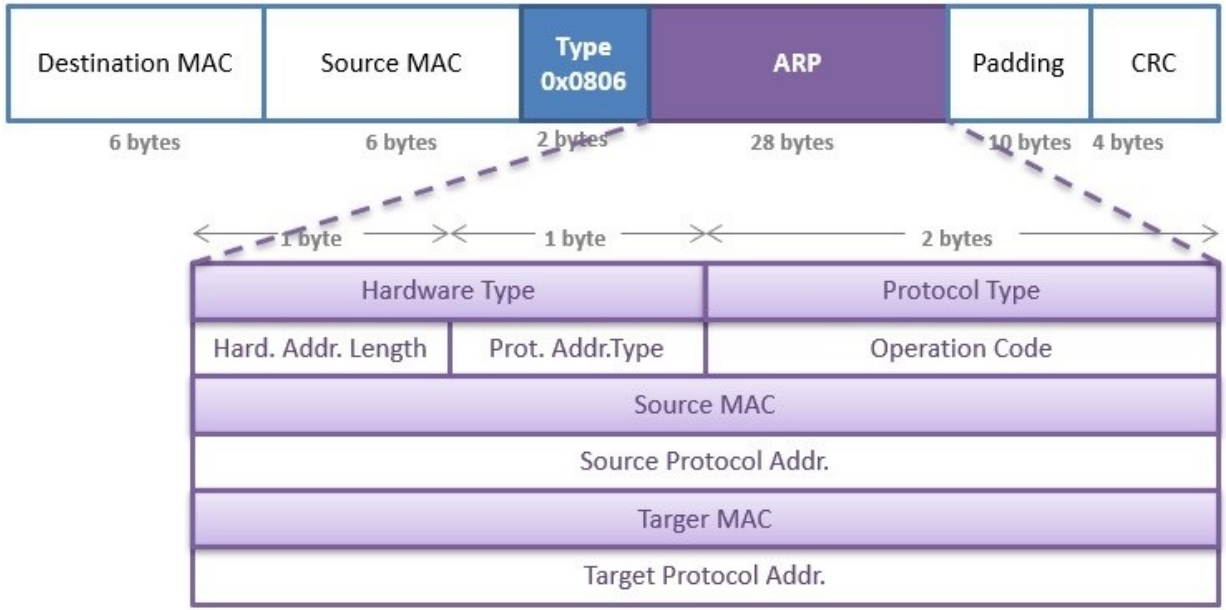
Recap of last lecture (5/7)

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

Fachhochschule Dortmund
University of Applied Sciences and Arts

Recap of last lecture (6/7)

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




```
Address Resolution Protocol (reply)
    Hardware type: Ethernet (1)
    Protocol type: IPv4 (0x0800)
    Hardware size: 6
    Protocol size: 4
    Opcode: reply (2)
    Sender MAC address: Apple_ae:7c:ef (5c:e9:1e:ae:7c:ef)
    Sender IP address: 192.168.10.81
    Target MAC address: be:53:a8:a0:01:82 (be:53:a8:a0:01:82)
    Target IP address: 192.168.10.75
```

Recap of last lecture (7/7)

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

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

Fachhochschule
Dortmund
University of Applied Sciences and Arts

Layer 3 - Network Layer

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

Fachhochschule Dortmund
University of Applied Sciences and Arts

Content Overview

Fundamentals of the network layer

IPv4

- Subnetting
- Fragmentation
- ICMP

IPv6

- Addressing
- ICMPv6

Routing / NAT

Fachhochschule Dortmund
University of Applied Sciences and Arts

Learning Objectives

- You know the structure of the IPv4 and IPv6 header and can explain the fields.
- You understand the structure of IP-addresses, subnets and its calculation
- You know reserved IP-addresses and its usage
- You understand IPv6-addresses and the use protocols like NDP or DAD
- You know the task of ICMP and can classify ICMP in the layer model.
- You can explain the process of fragmentation in IPv4

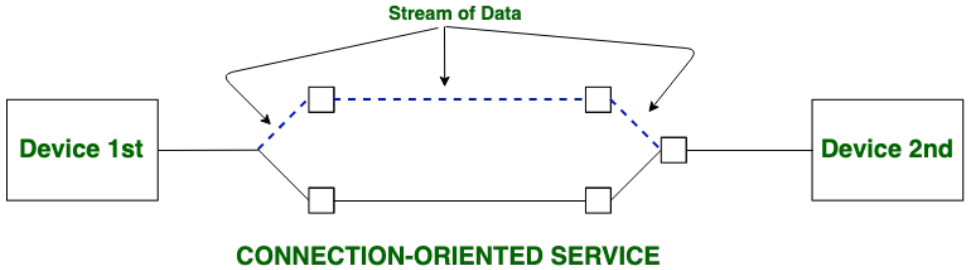
Fachhochschule Dortmund
University of Applied Sciences and Arts

Layer 3 – Network layer

- While the link layer is only responsible for the transmission between two successive transmission systems (broadcast domain), the network layer **carries out the transmission of the packets from the sender to the receiver.**
- Layer 3 provides addressing of endpoints with **IP-addresses**
- Layer 3 provides **the choice of route (routing)**. This includes the **possibility of deciding between several possible paths based on predetermined criteria (time, quality, costs).** A precise knowledge of the network structure is absolutely necessary for this.
- Layer 3 provides connection oriented (aka virtual circuits) or connectionless communication
- Relevant protocols:
  - IPv4
  - IPv6
  - ICMP

Connectionoriented Communication

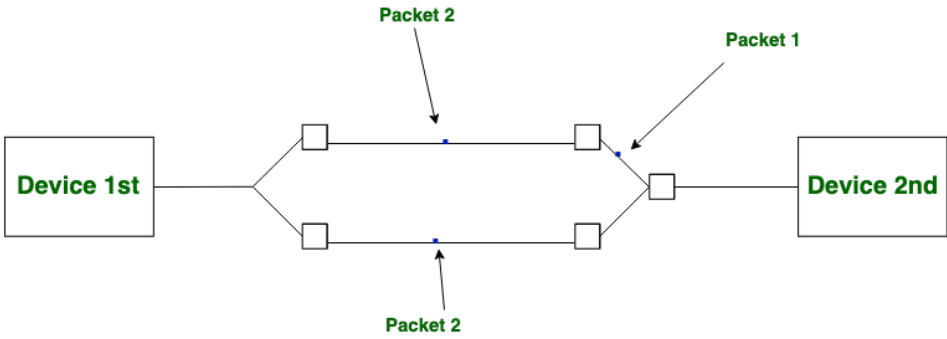
- Before data transmission is possible, a channel has to be established (virtual circuit, VC)
- Every packets is transferred inside the VC
- Out-of-sequence is uncommon
- Router failure is critical
- High overhead
- Good congestion control
- Easy QoS



https://www.geeksforgeeks.org/

Connectionless Communication

- Data transmission method used in packet switching networks in which each data unit is individually addressed and routed based on information carried in each unit
- A network packet can be routed on a different path than the prior or subsequent packet
- Out-of-sequence is possible
- CL has lower overhead (no special channel information needed)
- Router failure has low impact



CONNECTIONIESS SERVICE

https://www.geeksforgeeks.org/

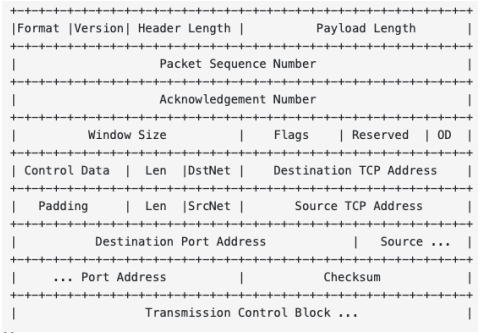
Fachhochschule
Dortmund
University of Applied Sciences and Arts

Network Layer – IPv4

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

Fachhochschule Dortmund
University of Applied Sciences and Arts

History



- May 1974: Vint Cerf and Bob Kahn published “A Protocol for Packet Network Intercommunication.”
  The paper describes the use of host-to-host AND process-to-process communication
- December 1974: RFC 675 specifies the paper
- 1977: TCP2 with an 4bit version field (later declared as IPv0)
- January 1978: Next revision led TCP3, RFC 755 later assigns IPv1 to this
- February 1978: IPv2 (IP separates from TCP)
- February 1978: Vint Cerf proposes a new header for IP (=IPv3) and TCP
- 1979 – 1995: Working on the IPv5 Internet Stream Protocol
- September 1981: Publishing RFC 791 with IPv4
- March 1982: US DoD defines IPv4 as standard for military comm.
- December 1998: IPv6 ratified

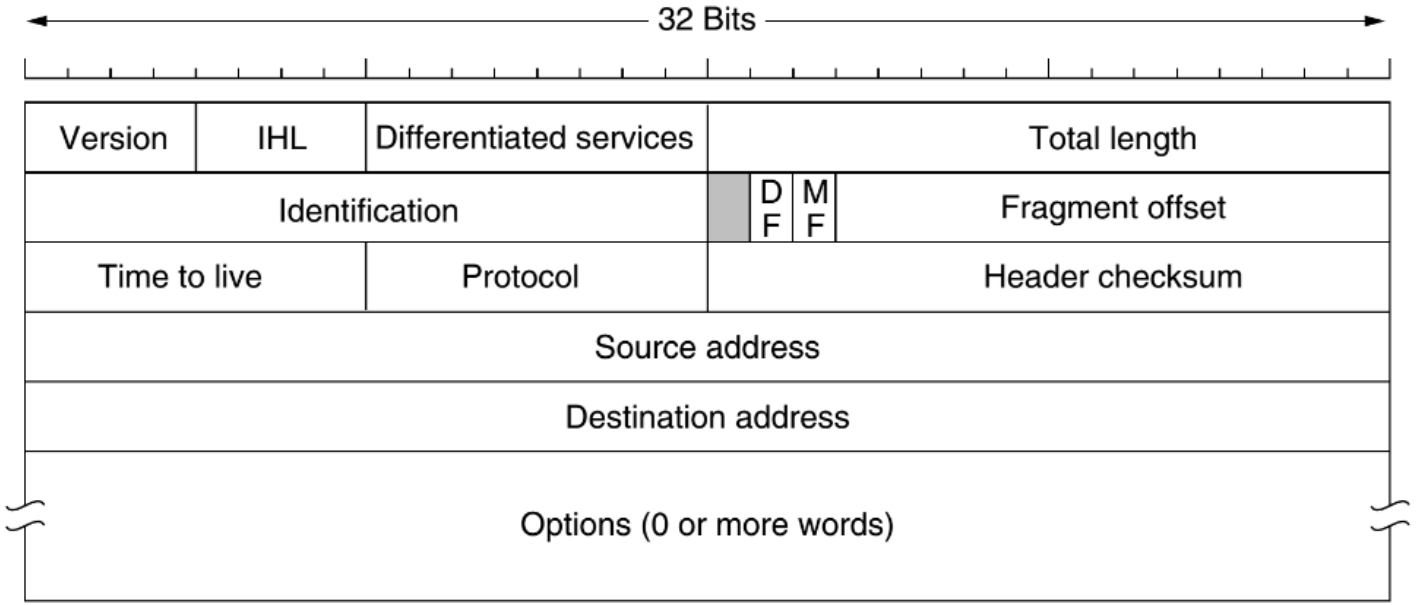
Internet Protocol IP (RFC 791)

- The Internet Protocol (IP) is responsible for transporting data across multiple networks.
- One of the core protocols of standards-based internetworking methods on the Internet and most packet-switched networks
- First version deployed 1982/1983
- Part of the so-called internet protocol suite TCP/IP
- IP takes data segments from layer 4 and packs them into packets called datagrams.
- IP provides an addressing mechanism that allows you to choose between networks. Each datagram consists of a piece of data and a header that contains both source and destination addresses
- IP is connectionless
- In order to support as many networks as possible with their different packet lengths, IP can divide datagrams on their way to the destination into smaller datagrams (Fragmentation)

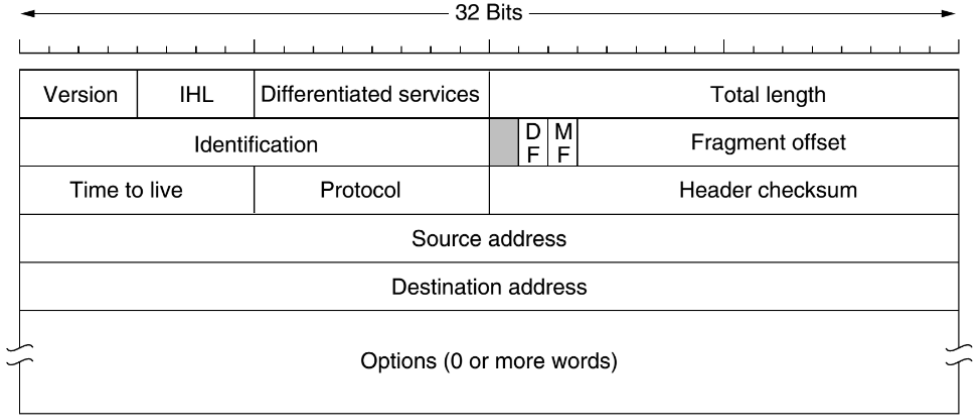
Fachhochschule Dortmund
University of Applied Sciences and Arts

Structure of an IPv4 packet

Typical header size is 20 Bytes
Minimum size is 21 Bytes
Maximum size is 65.535 Bytes

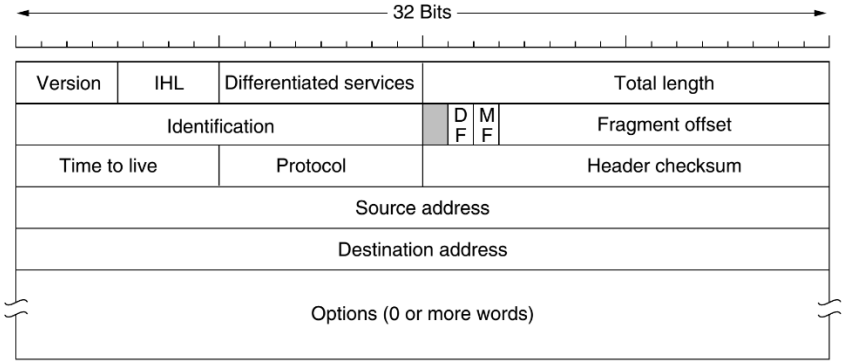


Explanation of the individual fields



- Version: Identifies the version of IP
- IHL Internet Header Length:
  Consists of 4 bits that specify how long the datagram header is (in 32bit words)
- Diff.Serv: The field can be set and evaluated for the prioritization of IP data packets (Quality of Service)
- Total length: Specifies the total length of an IP datagram (header and payload), can calculate the actual data length by subtracting IHL and total length
- Identification: IP packets are uniquely identified by a number
- Flags: 3 bit, first must be zero, second (DF) prevents fragmentation, third (MF) Indicates fragmentation of the datagram

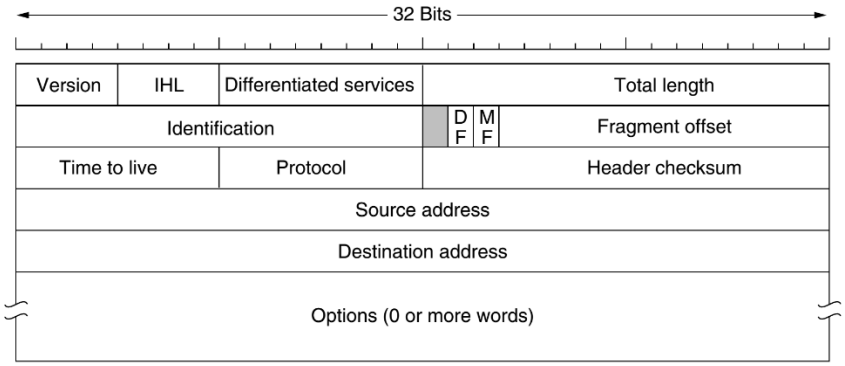
Explanation of the individual fields



- Fragment Offset: Specifies the relative position of the datagram fragment to the original
- Time to Live: Contains the number of routers to be crossed (so called hops) (decremented by 1 each time it is crossed).
- Protocol: Provides information of the PDU
- Header checksum: Detects changes in the header
- Source address: Address of the source node;
- Destination address: Address of the destination node
- Options (not necessarily present): Identifies additional services
- Padding: Ensure that the length of a datagram header is an integer multiple of 32 bits
- Data: User data

Fachhochschule Dortmund
University of Applied Sciences and Arts

Wireshark




```
Internet Protocol Version 4, Src: 10.0.0.2, Dst: 10.0.0.1
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
> Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
  Total Length: 100
  Identification: 0x0010 (16)
> 000. .... = Flags: 0x0
  ...0 0000 0000 0000 = Fragment Offset: 0
  Time to Live: 255
  Protocol: ICMP (1)
  Header Checksum: 0xa786 [validation disabled]
  [Header checksum status: Unverified]
  Source Address: 10.0.0.2
  Destination Address: 10.0.0.1
```

orks

Fachhochschule
Dortmund
University of Applied Sciences and Arts

IPv4 - Addressing

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

Fachhochschule Dortmund
University of Applied Sciences and Arts

IPv4 address

- In an IPv4 based network, each system is assigned a 32-bit Internet address (IP address), which is unique in this specific network
- The Internet Address (**IPv4 address**) is encoded in 32 bits (4 octets or bytes).

Byte3	Byte2	Byte1	Byte0

- Range of IPv4-addresses: 0.0.0.0 -> 255.255.255.255

Example:

01111010	.	11111111	.	11001111	.	01100000
122	.	255	.	207	.	96

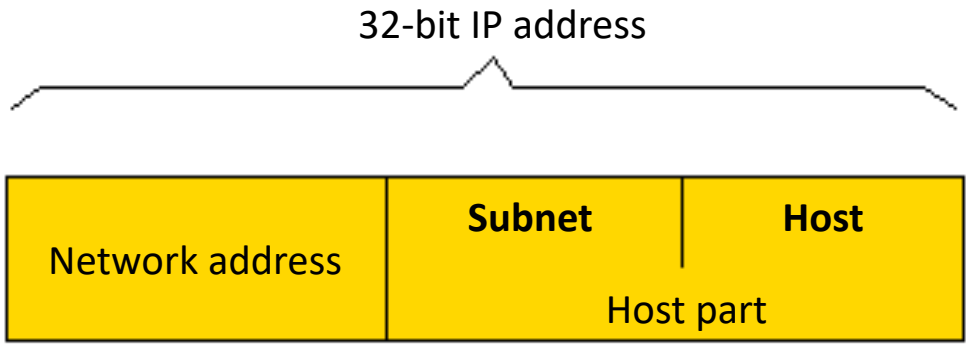
Structure and spelling of IP addresses

- An IPv4 address is a 32-bit number, the so-called Dotted Decimal Notation has prevailed as the notation.
  - Dotted Decimal Notation:
    - 211.15.24.143
    - 10.1.2.3
    - 192.168.10.99

- In principle, other spellings are also possible:
  193.25.22.84 = 0xC1191654 (Hex) = 030106213124 (Octal)= 3239646804 (Integer)
- The IP address C1:19:16:54 of a computer of the Faculty of Computer Science then has the form: 193.25.22.84.

IPv4 Addressing

- An IP-address consists of the network and the host part
  - The network field identifies the network on the Internet to which the computer is connected.
  - The host field identifies the computer within this network or within a smaller subnet.



- Separation is done with the subnet mask

Fachhochschule Dortmund
University of Applied Sciences and Arts

Subnet masks

- Subnet mask divides the address into the subnet and the host part.
- In order to identify the network to which it is connected, **every IP interface** of a host or a router must have a **netmask** assigned.
- The **netmask** is a **32-bit number**, which has **n** most significant bits at 1 and then (32-n) least significant bits at 0
- Like an IP address, a (sub-)netmask consists of 32-bit. The network and subnet fields are represented in the mask by set (=1) bits. The computer or host part has 0 bits.

11111111	.	11111111	.	11110000	.	00000000	
255	.	255	.	240	.	0	

- If one computer wants to communicate with another, it must first find out whether it is on the same network. To do this, it takes the destination address and hides the computer portion with the help of the subnet mask. If the network address obtained in this way is identical to your own, the target computer is in the same network. Otherwise, communication must take place via a gateway (=> Routing)

Fachhochschule Dortmund
University of Applied Sciences and Arts

IP subnetmask

- The n-bit **Netmask /n** allows the address space to be partitioned into a network or subnet of **2^(32-n)** addresses. The lowest address is used for the **Network Address** and the highest for the **Broadcast Address**. The rest can be used to identify the interfaces (hosts).

Netmask /n	Netmask ddn	2^(32-n)	# Hosts
/24	255.255.255.0	256	254
/25	255.255.255.128	128	126
/26	255.255.255.192	64	62
/27	255.255.255.224	32	30
/28	255.255.255.240	16	14
/29	255.255.255.248	8	6
/30	255.255.255.252	4	2

IP Addressing
Network and Broadcast Addresses

- Every IP interface of a system (host, router) must have an IP address and a netmask assigned.
- The Network address permit the identification of all the interfaces (hosts and routers) that are connected to the same network (or subnetwork).
- The Broadcast address allow datagrams to be broadcast to all hosts on the same network.

Network Address	Every host bits is 0	Address that identifies the network
Broadcast Address	Every host bits is 1	Datagram destinated to all hosts on the network

- The address of the network (32 bits) is obtained by making the AND between the IP Address and the mask of the interface.
- The Broadcast Address of the network (32 bits) is obtained by OR-ing the IP Address with the negated mask

Fachhochschule Dortmund
University of Applied Sciences and Arts

Determination of allocation

- IP-address + subnet mask define the allocation:
- 192.168.10.3 / 26

	IP	Subnet
Dot-decimal	192.168.10.3	255.255.255.192
binary	11000000.10101000.00001010.00000011	11111111.11111111.11111111.11000000

- IP: 11000000 .10101000 .00001010 .00000011
- Mask: 11111111 .11111111 .11111111 .11000000

Fachhochschule Dortmund
University of Applied Sciences and Arts

Determination of allocation

- IP-address + subnet mask define the allocation:
- 192.168.10.3 / 26

	IP	Subnet
Dot-decimal	192.168.10.3	255.255.255.192
binary	11000000.10101000.00001010.00000011	11111111.11111111.11111111.11000000

IP:	11000000	.10101000	.00001010	.00000011
Mask:	11111111	.11111111	.11111111	.11000000
Net:	192	.168	.10	.0
BC:	192	.168	.10	.63

Fachhochschule Dortmund
University of Applied Sciences and Arts

Determination of allocation

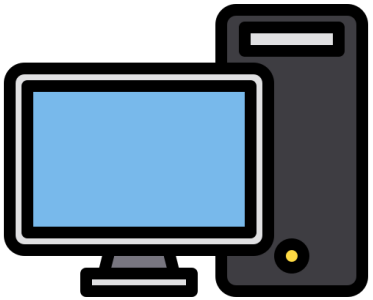
- IP-address + subnet mask define the allocation:
- 192.168.10.3 / 26

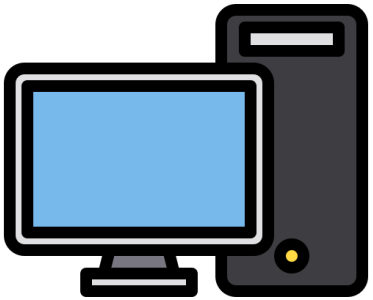
	IP	Subnet
Dot-decimal	192.168.10.3	255.255.255.192
binary	11000000.10101000.00001010.00000011	11111111.11111111.11111111.11000000

IP:	11000000	.10101000	.00001010	.00000011
Mask:	11111111	.11111111	.11111111	.11000000
Net:	192	.168	.10	.0
BC:	192	.168	.10	.63

6 Bit available = 2^6 = 64 addresses
All bits ”1” (= 111111) is BC-address = 63

Usage



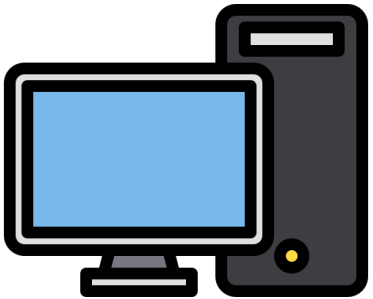


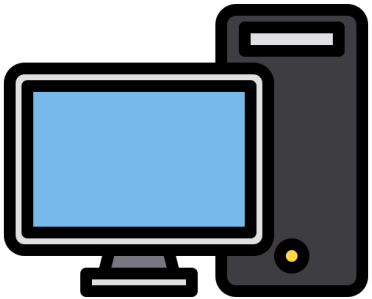
PC1: 106.93.22.7 / 14

PC2: 106.95.33.103 / 14

Are PC 1 and PC2 in the same network?

Usage





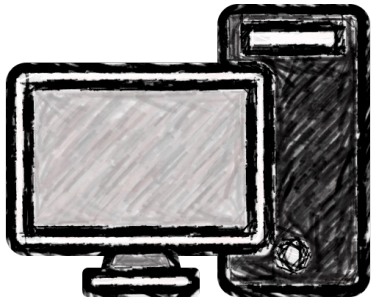
PC1: 106.93.22.7 / 14

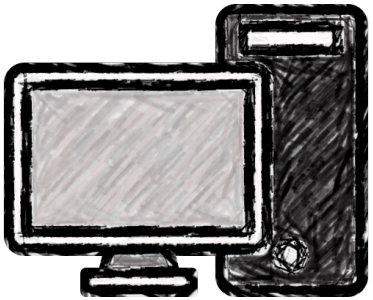
PC2: 106.95.33.103 / 14

106.93.22.7 255.240.0.0 => Net: 106.80.0.0, BC: 106.95.255.255

No Routing necessary, PC1 and PC2 are in the same subnet

Usage II



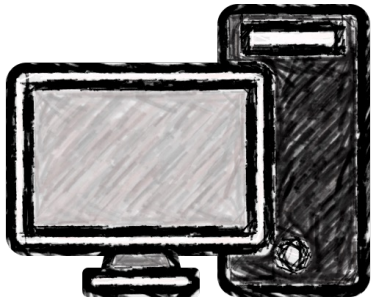


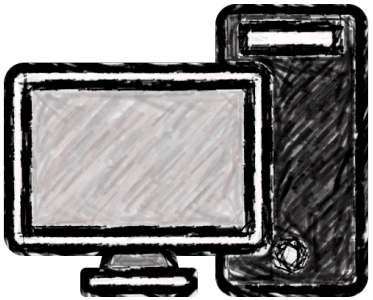
PC3: 88.12.33.125 / 26

PC4: 88.12.33.129 / 26

Are PC 3 and PC4 in the same network?

Usage II





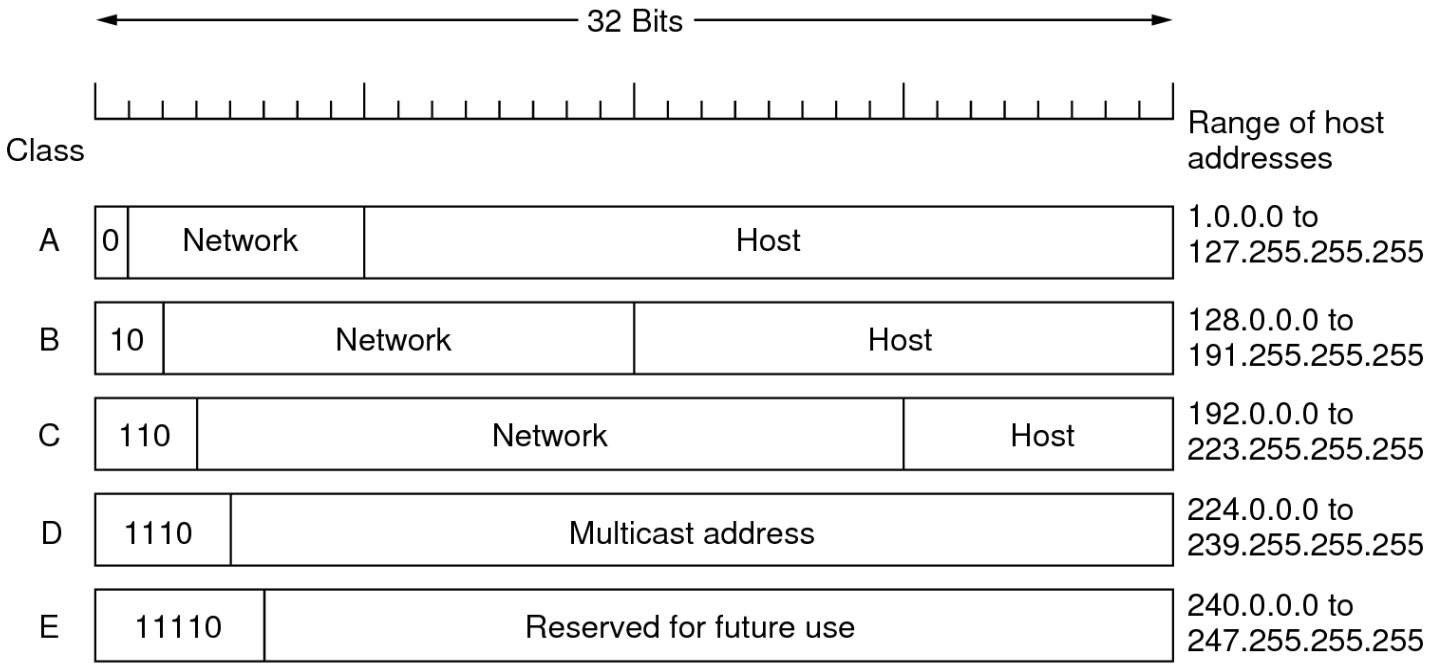
PC3: 88.12.33.125 / 26

PC4: 88.12.33.129 / 26

88.12.33.125 255.255.255.192 => Net: 88.12.33.64, BC: 88.12.33.127

PC3 and PC4 are NOT in the same subnet, routing is necessary (sending the packet to the default gatewas)

“Ancient” classification of IP addresses (default masks)



Fixed Length Subnet Mask

Subnet mask are used to separate networks

Using a FLSM, each network (e.g. block of IP addresses) is divided into multiple subnets of equal length, i.e. an equal number of IP addresses.

- Easy assignments and mappings
- Easy calculation of available blocks
- Easy routing tables

Available block: 10.0.0/8
256 networks with 10.0.x.0/24, each with 256 addresses
or 128 networks 10.x.y.z/23, each with 512 addresses

Variable Length Subnet Mask

When using different NetMasks (variable length) for the same NetNumber

partitioning of the address space of a network (identified with a Net#)

into subnets of different sizes.

- Prevents waste of addresses
- More complex calculation of addresses
- Maybe more complex routing tables

Available block: 10.0.0/8
10.x.0.0/9, 1 network with 8,388,608 addresses
AND 10.128.x.0/22, 5 networks with 1024 addresses
AND 10.128.y.0/25, 250 networks with 128 addresses
AND still addresses left

IP address assignment

- The network part of an IP address is assigned by central organizations.
- Internationally, the ICANN (Internet Corporation for Assigned Names and Numbers, www.icann.org for general coordination) and the IANA (Internet Assigned Numbers Authority, www.iana.org) are responsible for this.
- These include the Regional Inter-net Registries (RIR), for Europe e.B. the RIPE NCC (Réseaux IP Européens) as well as national registries (NIR) and, if applicable, local registries (LIR, ISP).
- Examples for german LIR:
  - Deutsche Telekom
  - Vodafone
  - de.government

Reserved IP-addresses

- IETF and IANA have reserved different ranges of IP-addresses for special purposes (RFC 6890)

Classification	Range	Definition
Localhost	127.0.0.1	RFC 1122
Private networks	10.0.0.0/8 172.16.0.0/12 192.168.0.0/16	RFC 1918
CGN	100.64.0.0/10	RFC 6598
APIPA	169.254.0.0/16	RFC 3297
Multicast	224.0.0.0/4	RFC 5771
Reserved	240.0.0.0/4	RFC 1700

Fachhochschule
Dortmund
University of Applied Sciences and Arts

Fragmentation

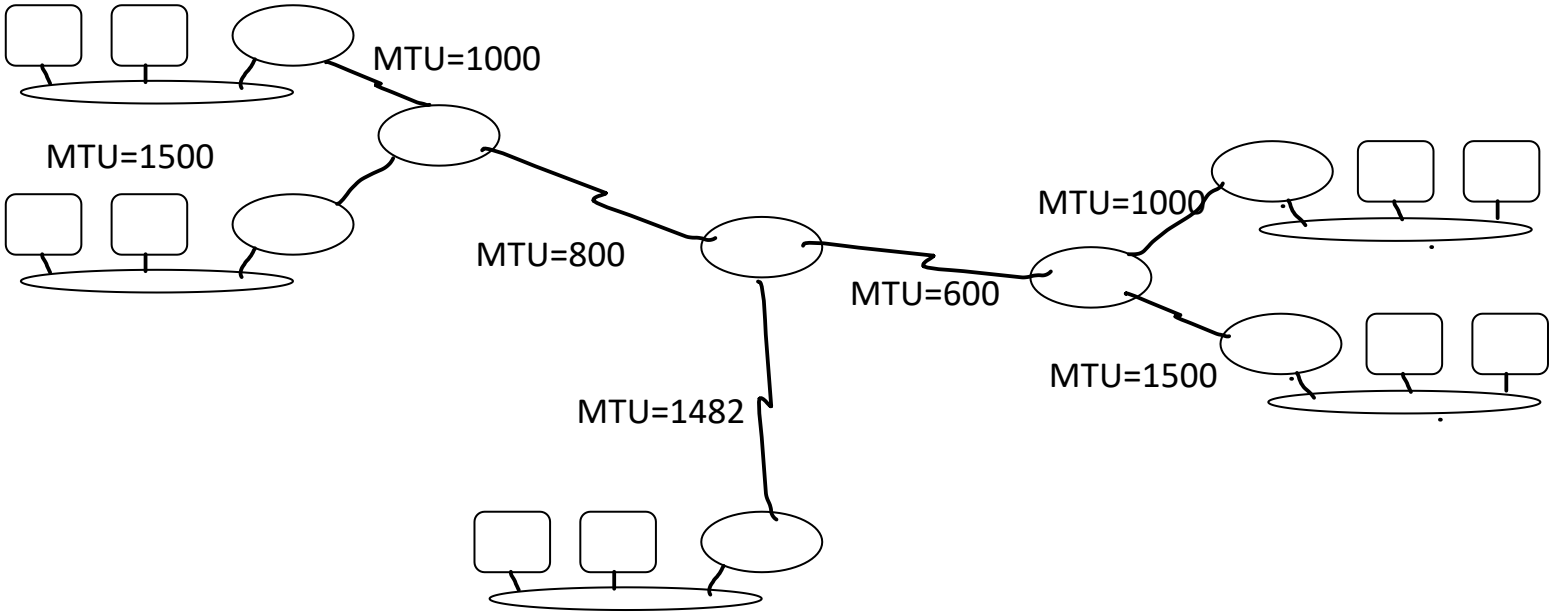
Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

Fragmentation

- The maximum packet size of a protocol of the switching layer (layer 3) of the OSI model, which fits into a frame of the link layer (layer 2) without fragmentation, is called the Maximum Transmission Unit (MTU).
- IP datagrams are transported via various technologies that can have different MTUs.
- In the case of a connection across several (different) networks, it can happen that networks are passed through that have a smaller MTU than the original network.
- In the IPv4 protocol, the datagrams in this case are divided by the routers into smaller packets (fragmented) and only in the target computer the parts are assembled (reassembled).

Fragmentation

- Why do we need to fragment?
  IP supports a maximum datagram size of 64 Kbytes
  IP is carried over different LAN&WAN network technologies **with different MTUs**



Fragmentation

- When is the IP datagram fragmented?
  It is fragmented when the datagram is **longer than the MTU** of the network through which the forwarding must be carried out.
- Who fragments?
  Fragmentation can occur either in the **source host** or intermediate **router**
- Who reassembles?
  Reassembly of a fragmented datagram is done only at the **destination host**.
  The destination host initializes a **reassembly timer**. If this timer expires before all fragments have been received, it will discard all fragments.

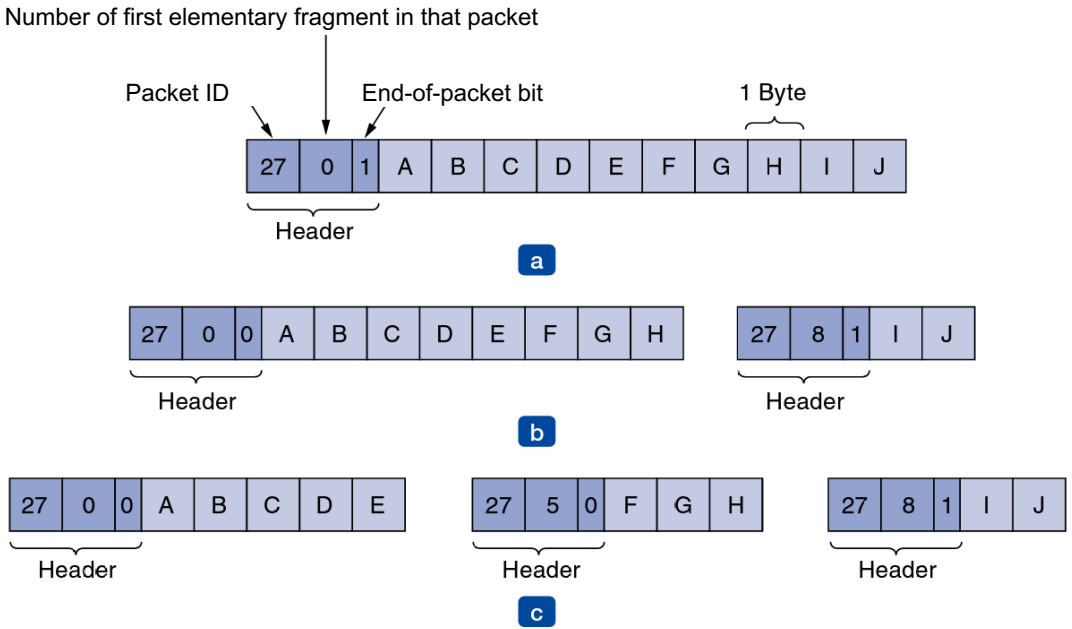
Fragmentation

- How is the datagram fragmented?
  The fragmentation process uses the "**offset**" field and the "**more fragments**" flag of the IP header. The "**Identifier**" field is used to differentiate fragments of originally different IP datagrams. This scheme allows fragmentation of fragments.

Fragment	Offset	MF Flag
1st	**0**	1
intermediate	> 0	1
last	> 0	**0**

Fachhochschule Dortmund
University of Applied Sciences and Arts

Example of fragmentation of IP datagrams in the router



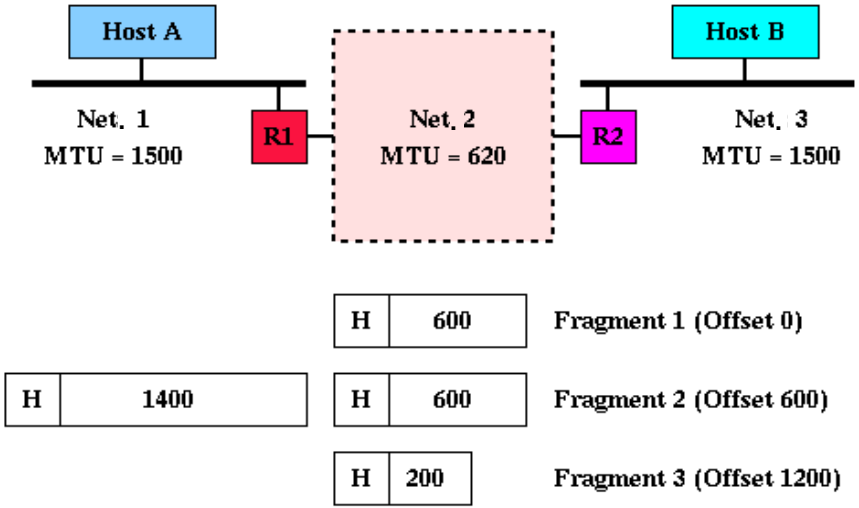
Original package with 10 data bytes

Fragments after limiting to 8 data bytes

Fragments after limiting to 5 data bytes

Fragmentation uses a flag that indicates whether the datagram represents the end of the package.

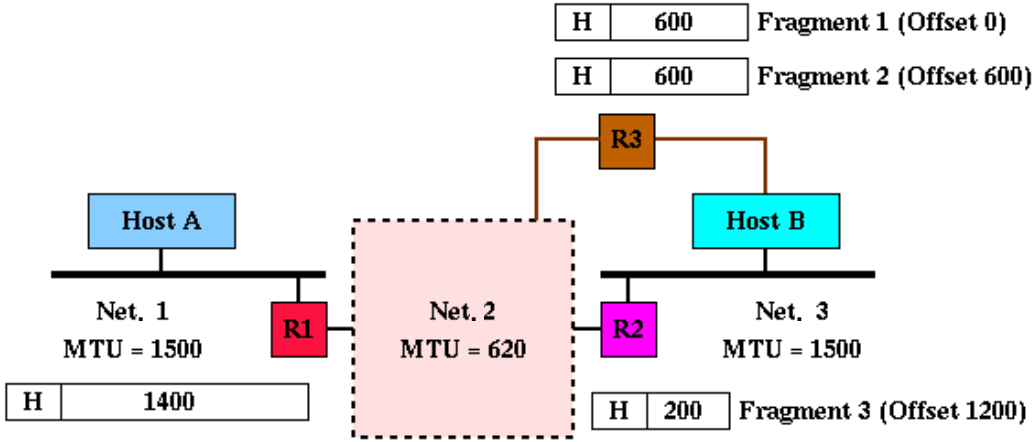
Fragmentation of IP datagrams in the router



Host A sends an IP datagram with a payload size of 1400 bytes. Router 1 breaks down the datagram into 3 fragments. Host B reassembles the 3 fragments like the.

Reassembling IP datagrams in the host

The **reassembly cannot be performed by the routers**, as different paths to the destination are possible.



Host B is connected to the network via two different paths. Now it can happen that e.B. the first two fragments come via R3, but the last one via R2. None of the routers involved "sees" all fragments, the reassembly is only possible at the destination,

2000byte Ping in Wireshark

```
Total Length: 1500
Identification: 0xa964 (43364)
001. .... = Flags: 0x1, More fragments
   0... .... = Reserved bit: Not set
   .0.. .... = Don't fragment: Not set
   ..1. .... = More fragments: Set
...0 0000 0000 0000 = Fragment Offset: 0
Time to Live: 64
Protocol: ICMP (1)
```

Fachhochschule Dortmund
University of Applied Sciences and Arts

2000byte Ping in Wireshark

- Total length = 2048?
- Length Packet 1:
  1480 PDU
  + 20 IP Header
  = 1500
- Length Packet 2:
  520 PDU (=PDU complete)
  + 20 IP Header
  = 540
- ?

```
Total Length: 1500
Identification: 0xa964 (43364)
001. .... = Flags: 0x1, More fragments
   0... .... = Reserved bit: Not set
   .0.. .... = Don't fragment: Not set
   ..1. .... = More fragments: Set
...0 0000 0000 0000 = Fragment Offset: 0
Time to Live: 64
Protocol: ICMP (1)
```

```
Total Length: 548
Identification: 0xa964 (43364)
000. .... = Flags: 0x0
   0... .... = Reserved bit: Not set
   .0.. .... = Don't fragment: Not set
   ..0. .... = More fragments: Not set
...0 0000 1011 1001 = Fragment Offset: 1480
Time to Live: 64
Protocol: ICMP (1)
```

Fachhochschule Dortmund
University of Applied Sciences and Arts

2000byte Ping in Wireshark

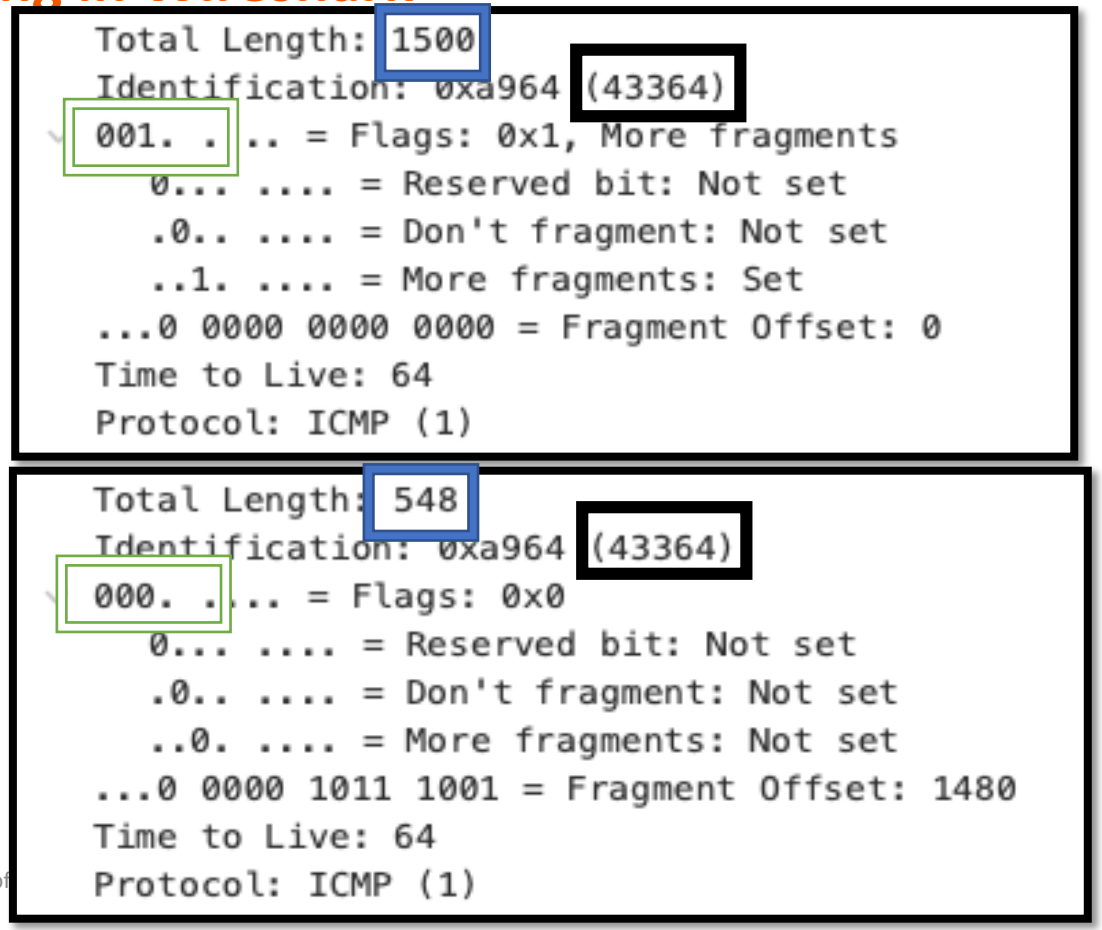
- Total length = 2048?
- Length Packet 1:
  1480 PDU
  + 20 IP Header
  = 1500
- Length Packet 2:
  520 PDU (=PDU complete)
  + 20 IP Header
  = 540
  + 8 ICMP-Header
  = 548

```
Total Length: 1500
Identification: 0xa964 (43364)
001. .... = Flags: 0x1, More fragments
  0... .... = Reserved bit: Not set
  .0.. .... = Don't fragment: Not set
  ..1. .... = More fragments: Set
...0 0000 0000 0000 = Fragment Offset: 0
Time to Live: 64
Protocol: ICMP (1)
```

```
Total Length: 548
Identification: 0xa964 (43364)
000. .... = Flags: 0x0
  0... .... = Reserved bit: Not set
  .0.. .... = Don't fragment: Not set
  ..0. .... = More fragments: Not set
...0 0000 1011 1001 = Fragment Offset: 1480
Time to Live: 64
Protocol: ICMP (1)
```

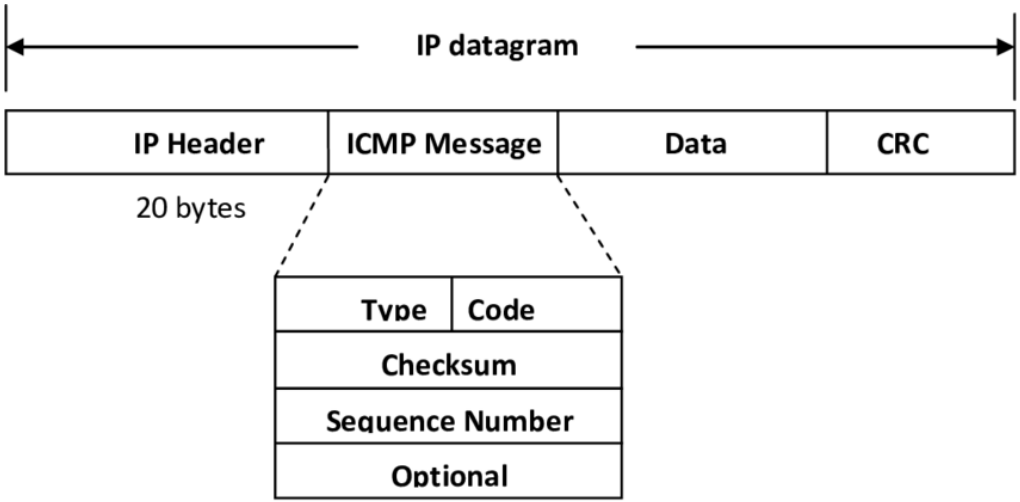
2000byte Ping in Wireshark

- Total Length
- Identification
- Flags



Fragmentation Length

- Most packets (with MF set) have a total length of 1500 Byte (1480 PDU)
- Last network packet (with MF not set) has different length
  - Last part of fragmented PDU
  - IPv4 header size (typ. 20 bytes)
  - Protocol depended header size



Disadvantages of fragmentation

- Fragmentation puts a strain on the routers involved, causing fragmentation to have a negative impact on data throughput.
- Although the possibility of fragmentation contributes to the robustness of the IP protocol in difficult network situations, because of the disadvantages it would be better if the sender immediately chooses the "optimal" packet length, i.e. the maximum packet length that is transported "just" without fragmentation.
- Current IPv4 implementations use the Path MTU Discovery (PMTUD) helper protocol to determine the optimal packet length.
- The IP protocol version 6 (IPv6) no longer provides for fragmentation in the router.

Fachhochschule
Dortmund
University of Applied Sciences and Arts

Network Layer - ICMP

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

Internet Control Message Protocol (RFC 792)

- ICMP is a supporting protocol of IP
- ICMP is used to manage networks and is mainly used to report errors. These error messages concern problems with the path selection, e.g., a network or a single computer is not reachable, or a computer uses incorrect routing tables.
- New ICMP messages are not generated from ICMP messages (except Echo request)
- ICMP only reports errors to the sending host of the datagram.
- The only part of the ICMP visible to the user is used by the ping program to test a connection. The associated messages of the ICMP are an echo request and an echo response.

ICMP

- The ICMP messages are transported in IP datagrams to their target computer or gateway and evaluated there by the IP software.

Overview ICMP Message Types

Message Type	Description
Destination unreachable	Packet could not be delivered
Time exceeded	The field Time to Live has value 0
Problem with parameters	Invalid header field
Source quench	Choke package
Redirecting	Tells the router about geography
Echo and echo reply	Checks if a router is still alive
Timestamp request/reply	As echo, but with a timestamp
Router advertisement/solicitation	Finds a nearby router

ICMP

The ICMP message is encapsulated in an IP datagram (Protocol: 0x01)

Ethernet Header	IP Header	ICMP Header	ICMP Data

Each message type has its own format, but they all have 3 fields in common:

- Type
- Code
- Checksum

Type	Code	ICMP header checksum
Data		

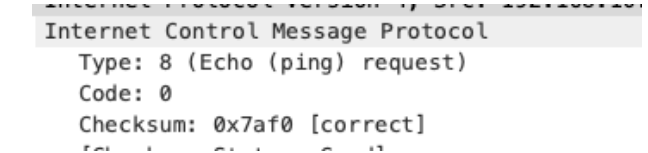
ICMP Type and Code

- The type field specifies the message.
- The code field interprets the message type more accurately.
- The data typically contains part of the original IP message.

Type	Type name	Code	Meaning
0	Echo Response	0	Echo Response
3	Destination uneachable	0	Network unreachable
		1	Host (target station) unreachable
		2	Protocol unreachable
		3	Port unreachable
		4	Fragmentation necessary, Don't fragment set
		5	Route not possible (the direction in IP header field option incorrectly specified)
		13	Communication administratively prohibited (packet is blocked by the recipient's firewall)
5	Redirect	0	Redirect datagram for the network
		1	Redirect datagram for the host
8	Echo Request	0	Echo request (better known as "ping")
11	Timeout exceeded	0	TTL (Time To Live) expired
		1	Timeout exceeded during reassembly

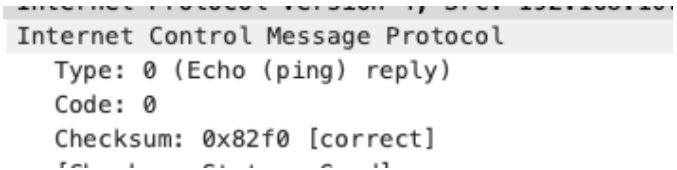
Fachhochschule Dortmund
University of Applied Sciences and Arts

ICMP - Ping



The "**ping**" command

- send an **ICMP_ECHO_REQUEST (Type 8, Code 0)**
- and expects to receive an **ICMP_ECHO_REPLY (Type 0, Code 0)** as a response



- before the timeout expires

The "ping" command lets you know the status of the network and the reachability of a destination host.

ICMP - Destination unreachable

- If an “**ICMP_DESTINATION_UNREACHABLE**” message is received
  it means that the IP datagram sent was **discarded** along the way by the IP entity (host or router) that appears in the IP Source address of the ICMP message.

  - If the Code indicates “**NETWORK_ UNREACHABLE**” (Type 3, Code 0)
    it means that it was **discarded for not matching any route**.

  - If the Code indicates “**HOST_ UNREACHABLE**” (Type 3, Code 1)
    it means that it was **discarded for not responding to the ARP_REQUEST**.

Fachhochschule Dortmund
University of Applied Sciences and Arts

ICMP - Traceroute

The traceroute command allows you to collect various information about the "path" to a destination host.

```
└─$ traceroute www.fh-dortmund.de                                    130 ↵
traceroute to www.fh-dortmund.de (193.25.16.26), 64 hops max, 52 byte packets
 1  fritz.box (192.168.178.1)  8.507 ms  3.660 ms  3.067 ms
 2  p3e9bf409.dip0.t-ipconnect.de (62.155.244.9)  11.784 ms  25.819 ms  7.364 ms
 3  d-ed6-i.d.de.net.dtag.de (62.159.98.130)  13.490 ms  12.416 ms  12.747 ms
 4  d-ed6-i.d.de.net.dtag.de (62.159.98.130)  11.705 ms  13.011 ms  12.680 ms
 5  193.159.165.115 (193.159.165.115)  19.081 ms  18.283 ms  18.363 ms
 6  cr-dui1-be17.x-win.dfn.de (188.1.144.189)  22.490 ms  23.611 ms  22.053 ms
 7  kr-fhdort11.x-win.dfn.de (188.1.233.242)  23.317 ms  23.619 ms  23.414 ms
 8  *^C
```

ICMP - Traceroute

The traceroute command allows you to collect various information about the "path" to a destination host.

The "**traceroute**" (or **tracert**) command

- send **ICMP_ECHO REQUEST** messages with TTL=1, 2, 3, …
- and expects to receive an **ICMP_TIME_EXCEEDED** response
- or an **ICMP_ECHO REPLY**

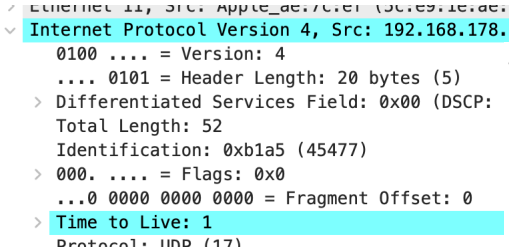
- By this, each router on the path provides some details and the path can be “reconstructed”

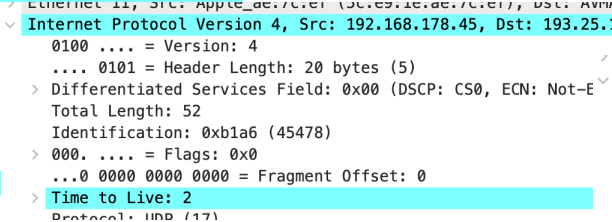
Fachhochschule Dortmund
University of Applied Sciences and Arts

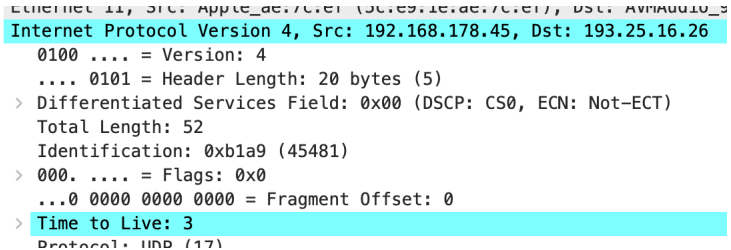
ICMP - Traceroute

The "**traceroute**" (or **tracert**) command

- send **ICMP_ECHO REQUEST** messages with TTL=1, 2, 3, …
- and expects to receive an **ICMP_TIME_EXCEEDED** response
- or an **ICMP_ECHO REPLY**







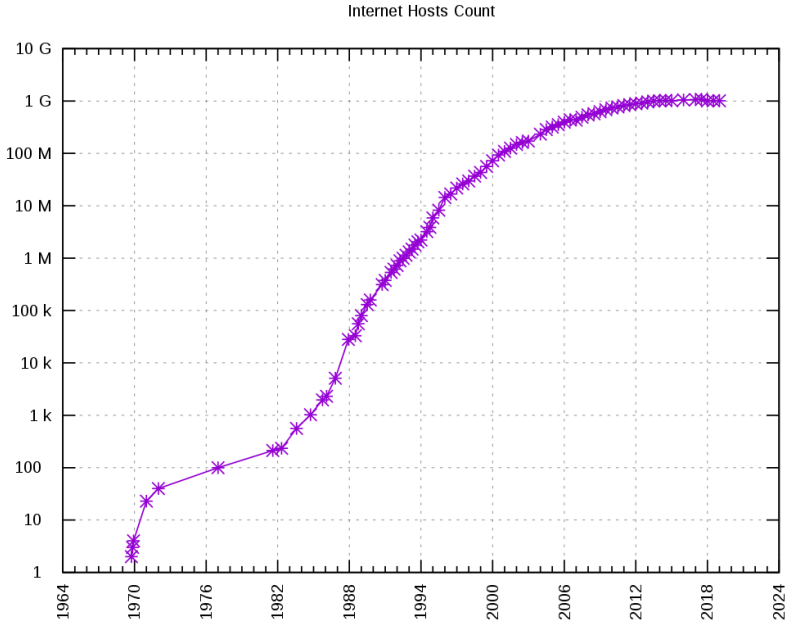
Fachhochschule
Dortmund
University of Applied Sciences and Arts

Network Layer – IPv6

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

Fachhochschule Dortmund
University of Applied Sciences and Arts

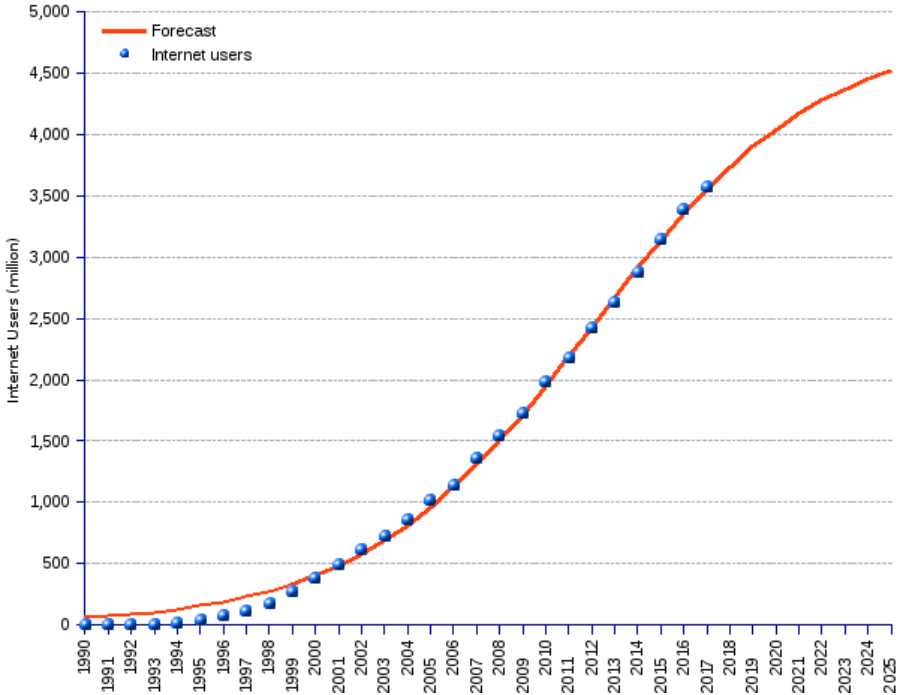
Why IPv6?



- **Address shortage**
  - IPv4 addresses 32-bit→ 4 billion addresses
  - Address space is poorly exploited by the network classes
  - Since the beginning of 2011, the IPv4 address space has been exhausted

- **Additional requirements** that led to IPv6:
  - **Security** (IPsec)
  - **Mobile** Internet (Mobile IP)
  - **QoS** Quality of Service
  - **Auto Configuration**, Plug & Play

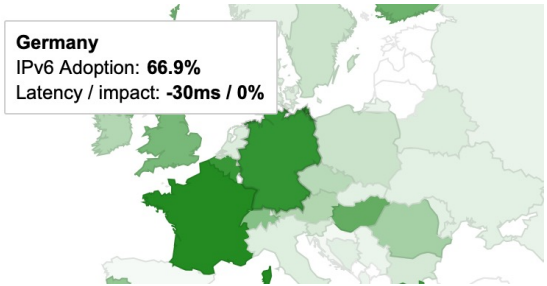
Growth of the Internet: number of hosts



Source: https://www.quora.com/What-was-the-year-the-Internet-became-essential-to-have

Internet Protocol version 6 (IPv6)

- Formerly also IP Next Generation (IPng)
- 1991 First considerations
- 1992 Establishment of working groups
- 1994 IPv6 selected as successor to IPv4
- 1998 IPv6 was published as **RFC 2460**
- 2002 the DFN Association starts IPv6 operation with the 6WiN
- 2004 Mobility Support in IPv6
- 2023 approx. 70% of german internet traffic is ipv6

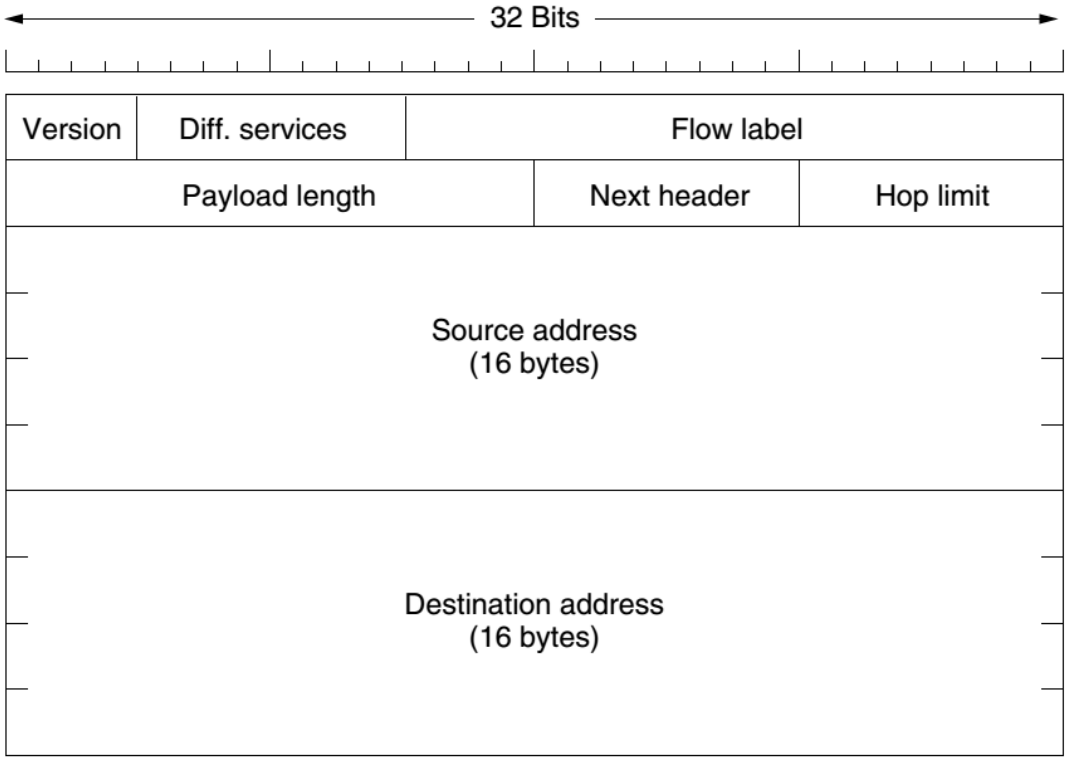


www.google.com

New Features

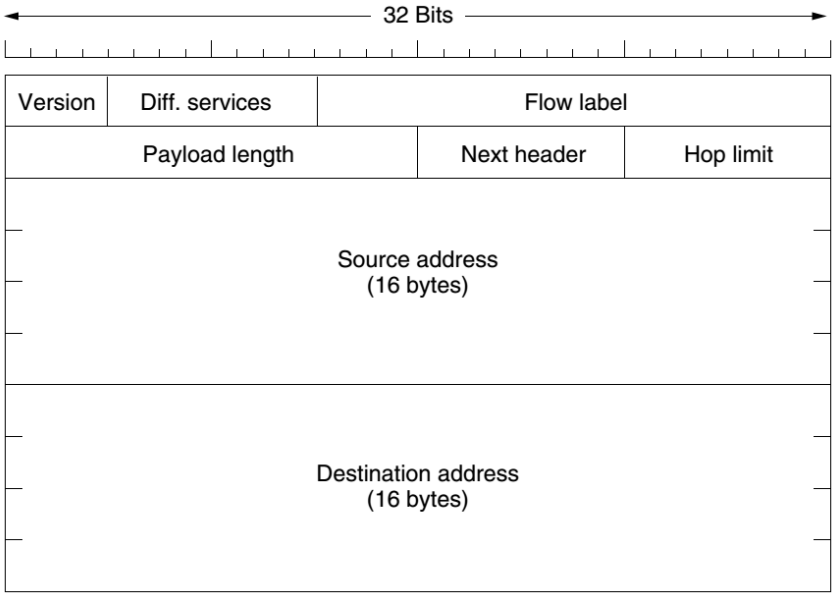
- Advanced **addressing** (**128-bit**) and routing
- Simplified **header** format
- Better support for options and extensions
- Support of:
  - **Security**: IPsec
  - **Mobility**: Mobile IP
  - Stateless Address **Autoconfiguration** (**SLAAC**) with ICMPv6 (Plug&Play)
  - **QoS** Quality Of Service (e.B. Video and Audio)

IPv6 Basic Header



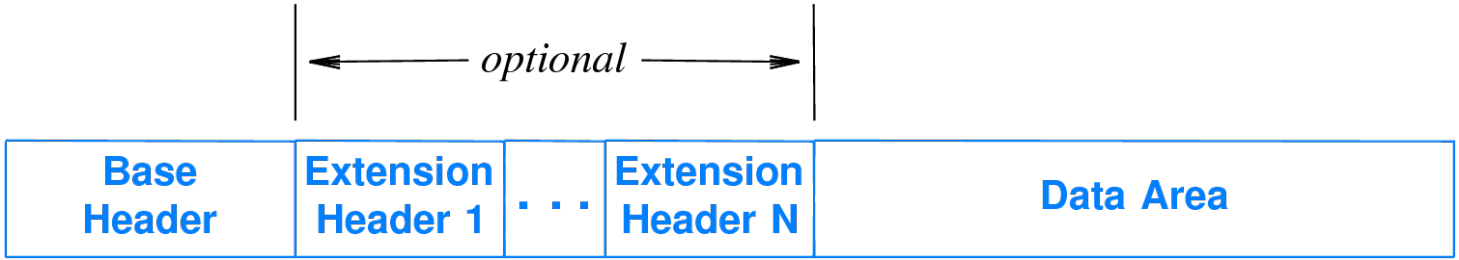
Meaning of the fields in the base header

- Version: Version 6
- Differential Services (before: Traffic Class): Quality of Service (Service Type for IPv4)
- Flow Label: identifies a route that meets the requirements of the Traffic Class. Packages bearing the same flow label, are treated equally.
- Payload Length: Length of user data
- Next Header: specifies the type of the next header or, if it does not exist, the type of data. (Protocol for IPv4)
- Hop limit like TTL for IPv4
- Source and destination address: Addressing of the communication partners



Fachhochschule Dortmund
University of Applied Sciences and Arts

IPv6 Extension Header



- Options are appended as a concatenated list if necessary
- Processing only at the target (except for hop-by-hop extensions)
- Eliminates IPv4 40 byte limit for options
  => easily expandable and future-proof

```
Internet Protocol Version 6, Src: 2001:470:e5bf:1001:8519:2d1f:c57d:fc4f, Dst: 2001:470:e5bf:dead:7db0:921:a2e9:1c21
    0110 .... = Version: 6
  > .... 0000 0000 .... .... .... .... .... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT)
    .... 0000 0000 0000 0000 0000 = Flow Label: 0x00000
    Payload Length: 8
    Next Header: Encap Security Payload (50)
    Hop Limit: 63
    Source Address: 2001:470:e5bf:1001:8519:2d1f:c57d:fc4f
    Destination Address: 2001:470:e5bf:dead:7db0:921:a2e9:1c21
> Encapsulating Security Payload
```

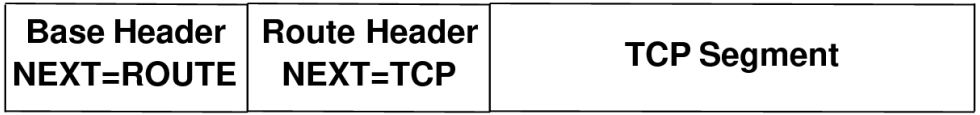
Currently defined extension headers

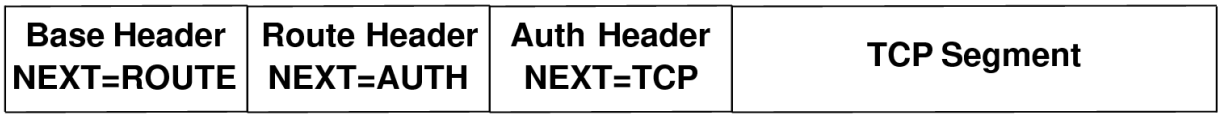
- Hop-By-Hop Options: Contains options that must be observed by all IPv6 devices that the packet passes through.
  RFC 2460, RFC 2675
- Destination Options: Contains options that only need to be considered by the target computer of the package.
  RFC 2460
- Routing: This header can influence the path of the packet through the network, it is used, among other things, for Mobile IPv6.
  RFC 2460, RFC 6275, RFC 5095
- Fragmentation: This header allows you to set the parameters of fragmentation.
  RFC 2460
- Authentication Header (AH): Ensure that packet comes from recipient. Contains additional data that can ensure the confidentiality of the package (IPsec).
  RFC 4302

Currently defined extension headers

- Encapsulating Security Payload (ESP): Contains data to encrypt the security (IPsec).
  RFC 4303
- Mobility: Contains data for Mobile IPv6.
  RFC 6275
- No Next Header: This type is just a placeholder to indicate the end of a header stack.
  RFC 2460

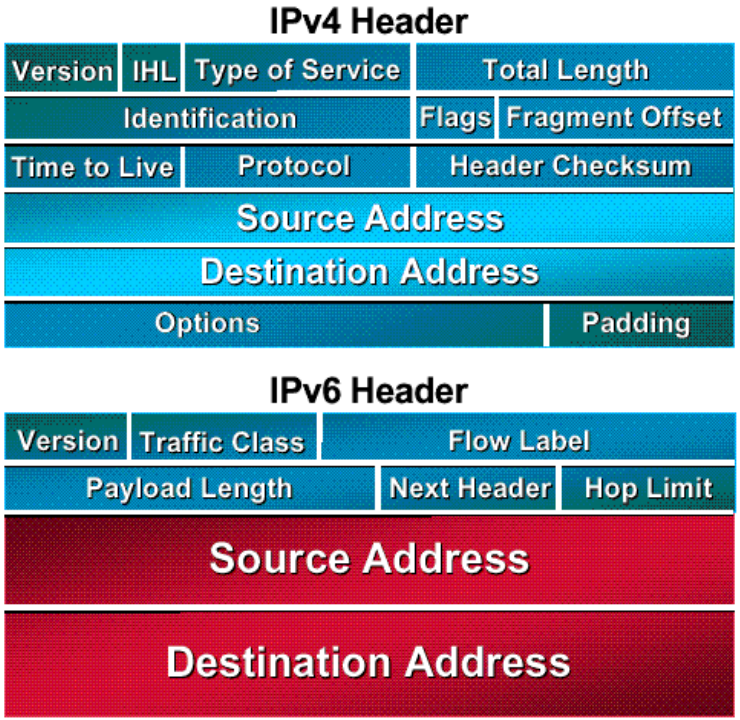
Examples extension headers





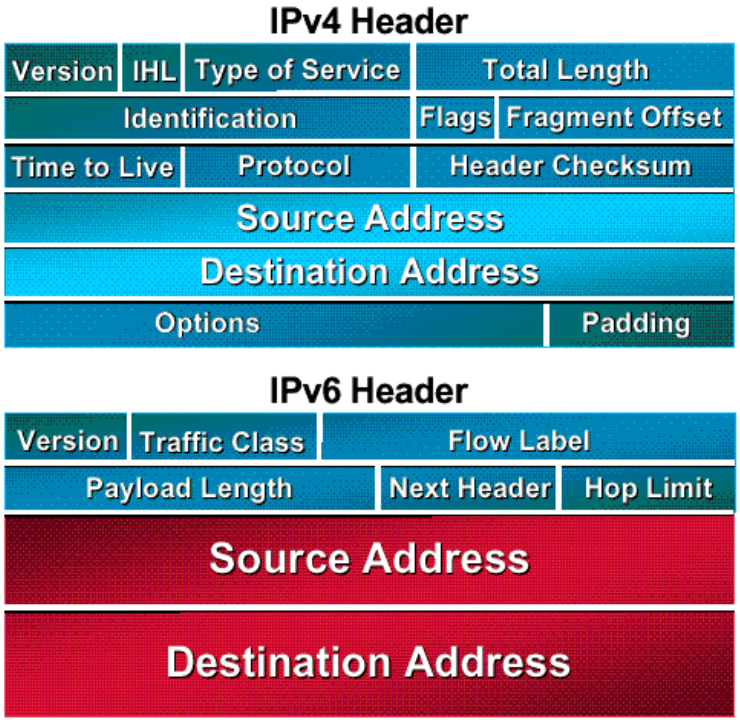
Differences between IPv4 and IPv6 defined in RFC 2460

- **Address length** quadrupled to **128 bits**
- **Fixed header length** (40 bytes) (Optional headers follow chained)
- No checksum (check is carried out in the Link layer)
- No hop-by-hop fragmentation (Path MTU discovery)
- Flow Label/Traffic Class (Integrated QoS Support)
- Concatenated extension headers



Differences between IPv4 and IPv6 defined in RFC 2460

- Version=6
- No IHL (Fixed Length Header)
- ToS -> Traffic Class
- Total Length -> Payload Length
- No ID, Flags, Frag Offset (No Frag on routers)
- TTL -> Hop Limit
- Protocol -> Next Header
- No Header Checksum (Check in Link Layer)
- Addresses 32 bits -> 128 bits
- No Options in Basic Header (Next Header)
- No Padding (Fixed Length Header)



128-bit IPv6 addresses – Basics

- So far: an IPv4 Internet address for about half of the people on earth
- Now IPv6: $2^{128} > 3{,}4 \cdot 10^{38}$
- I.e. 10^{24} addresses for every square meter of the earth
- Usually given in hexadecimal notation (colon hex), e. B. 68e6:8c64:ffff:ffff:0:1180:96a:ffff
- Multiple addresses per interface possible
- Different addresses predefined and ranges reserved for special purpose
- Typically, an Internet Service Provider (ISP) is assigned the first 32 bits (or less) as a network by a Regional Internet Registry (RIR).
  - This area is further divided into subnets by the provider. The length of the allocation to end customers is left to the ISP; the mini-number allocation of a /64 network is mandatory. Network part is typically with 48 or 56 bit long
  - A single network segment is usually assigned a 64-bit prefix.
  - Together with a 64-bit interface identifier, this forms the 128-bit-long address.
  - Separated in a network and a host part (Interface Identifier, IID)

128-bit IPv6 addresses – Format

Colon hexadecimal notation

0123 : 0002 : A234 : 0C00 : 0000 : 0000 : 00BB : 1003

- It is not necessary to put the zeros to the left

0123 : 2 : A234 : C00 : 0 : 0 : BB : 1003

- Zero strings can be abbreviated by using ::

0123 : 2 : A234 : C00 : : BB : 1003

FF01:0:0:0:0:0:0:101 (a multicast address)

FF01::101

0000 : 0000 : 0000 : 0000 : 0000 : 0000 : 0000 : 0001 (loopback address or local host)

::1

Fachhochschule Dortmund
University of Applied Sciences and Arts

Subnetting in IPv6

Similar to IPv4

Example:

- A123:B45::/6
- IP: A 1 2 3 B 4 5 6
- Binary: 1010 0001 0010 0011 1011 0100 0101 0110 …
- Mask (6bit): 1111 1100 0000 0000 0000 0000 0000 0000 …
- Starting: 1010 0000 0000 0000 0000 0000 0000 0000 …
- Ending: 1010 0011 1111 1111 1111 1111 1111 1111 …

Address range: A000:: - A3FF:FFFF:FFFF:FFFF …

Link Local Address

- Address range: fe80::/10
- Only valid inside a broadcast domain
- Basement for communication in IPv6
- Each interface with active IPv6-stack has a LLA

```
utun1: flags=8051<UP,POINTOPOINT,RUNNING,MULTICAST> mtu 2000
        inet6 fe80::bca6:ac81:b88a:2079%utun1 prefixlen 64 scopeid 0x13
        nd6 options=201<PERFORMNUD,DAD>
utun2: flags=8051<UP,POINTOPOINT,RUNNING,MULTICAST> mtu 1000
        inet6 fe80::ce81:b1c:bd2c:69e%utun2 prefixlen 64 scopeid 0x14
        nd6 options=201<PERFORMNUD,DAD>
```

- Protocols like NDP and DAD need LLA
- Localhost gets fe80::1
- Other interfaces calculate their LLA on their own

Globally Unicast Addresses (RFC 3587)

- Address range: Till now, only 2000::/3 is defined for GUA
  Range: 2000::/3 – 3fff::/3
- Globally unique
- Routable
- Assignment by Provider or RIR
- Reserved:
  - 2001:db8::/32 example and documentation
  - ::/96 for Transition
  - 64:ff9b::/96 for NAT64

Unique Local Address (RFC 6724)

- Address range: fc00::/7
- Routable **only** within a set of cooperating sites (similar to RFC 1918 private addresses)
- Introduced to replace the site-local addresses (now deprecated)
- Network address translateion needed for comm. with internet (so called NAT66)
- But:
  - ULA is less preferred (the precedence value is lower) than all IPv4
  - In case of Dual Stack, ULA is not used
  - So: ULA is broken

Then Tom Coffeen started putting a clearer picture on some real structural and architectural problems with ULA in his two part blog post titled "**3 Ways to Ruin Your Future Network with IPv6 Unique Local Addresses (Part 1 of 2)**" and "**3 Ways to Ruin Your Future Network with IPv6 Unique Local Addresses (Part 2 of 2)**", which covered:

Blogs.Infoblox.com

IPv6 Address Types

- Unicast (one-to-one)
  - link-local, **fe80::/10**, should not be forwarded by routers, are therefore only accessible in the same subnet.
  - unique-local address (ULA), **fc00::/7**, private addresses with unique site ID, described in RFC 4193. These addresses are routed only within a set of cooperating sites. These were introduced in the IPv6 to replace the site-local addresses. These addresses also provide a 40-bit pseudorandom number that reduces the risk of address conflicts.
- Anycast (one-to-nearest), addresses such as unicast
- Multicast (one-to-many), **ff00::/8**
  This prefix is offered by IPv6 to denote the multicast addresses. Any address carrying this prefix is automatically understood to be a multicast address.

Multicast

- Address range: ff00::/8
- Similar to multicast in IPv4, but much bigger relevance
- Used to send packet to receivers, whose LLA is unknown
- Uses scopes and groups
- Each host can join such a group
- No authentication or credentials necessary

Fachhochschule Dortmund
University of Applied Sciences and Arts

Multicast II

- ff0x
- ff0x::y

Scope	Name	Reference
0	Reserved	[RFC4291][RFC7346]
1	Interface-Local scope	[RFC4291][RFC7346]
2	Link-Local scope	[RFC4291][RFC7346]
3	Realm-Local scope	[RFC4291][RFC7346]
4	Admin-Local scope	[RFC4291][RFC7346]
5	Site-Local scope	[RFC4291][RFC7346]
6-7	Unassigned	
8	Organization-Local scope	[RFC4291][RFC7346]
9-D	Unassigned	
E	Global scope	[RFC4291][RFC7346]
F	Reserved	[RFC4291][RFC7346]

:1	All Nodes Address
:2	All Routers Address
:3	Unassigned
:4	DVMRP Routers
:5	OSPFIGP
:6	OSPFIGP Designated Routers
:7	ST Routers
:8	ST Hosts
:9	RIP Routers
:A	EIGRP Routers
:B	Mobile-Agents
:C	SSDP
:D	All PIM Routers
:E	RSVP-ENCAPSULATION
:F	UPnP
:10	All-BBF-Access-Nodes
:11	All-Homenet-Nodes
:12	VRRP

Example:

ff02::1 all nodes in network
ff02::2 all router in network
ff05::101 all NTP-server in site
ff02::1:2 all DHCP-agents in network

IPv6 Address Types

Address Type	Binary Prefix	IPv6 notation	Uses
Embedded IPv4 address	00…1111 1111 1111 1111 (96 bits)	::FFFF/96	Prefix for embedding IPv4 address in an IPv6 address
Loopback	00…1 (128 bits)	::1/128	Loopback address on every interface [RFC 2460]
Global unicast	001	2000::/3	Global unicast and anycast (allocated) [RFC 4291]
Global unicast	01 – 1111 1100 0	4000::/2 – FC00::/9	Global unicast and anycast (unallocated)
Teredo	0010 0000 0000 0001 0000 0000 0000 0000	2001:0000::/32	Teredo [RFC 4380]
Nonroutable	0010 0000 0000 0001 0000 1101 1011 1000	2001:DB8::/32	Nonroutable. Documentation purposes only [RFC 3849]
6to4	0010 0000 0000 0010	2002::/16	6to4 [RFC 3056]
6Bone	0011 1111 1111 1110	3FFE::/16	Deprecated. 6Bone testing assignment, 1996 through mid-2006 [RFC 3701]
Link-local unicast	1111 1110 10	FE80::/10	Link local unicast
Reserved	1111 1110 11	FEC0::/10	Deprecated. Formerly Site-local address space, unicast and anycast [RFC 3879]
Local IPv6 address	1111 110	FC00::/7	Unicast Unique local address space, unicast and anycast [RFC 4193]
Multicast	1111 1111	FF00::/8	Multicast address space [RFC 4291]

Address assignment

- Huge benefit is the implementation of (working) auto address assignment

- To perform address configuration on IPv6 there are a couple of familiar methods and a few additional methods, including:
  - static addressing,
  - static addressing with **DHCPv6** (stateless),
  - dynamic addressing via **DHCPv6** (RFC 3315) (Stateful),
  - Stateless Address Autoconfiguration (**SLAAC)** alone,
  - **SLAAC** with **DHCPv6** (Stateless).
- Especially SLAAC needs ICMPv6

Fachhochschule
Dortmund
University of Applied Sciences and Arts

Network Layer – ICMPv6

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

ICMPv6

- ICMPv6 is an integral part of IPv6.
- ICMPv6 performs **error reporting and diagnostic functions**.
- ICMPv6 messages: **error messages** and **information messages**.
- ICMPv6 messages are transported by IPv6 packets.
- IPv6 **Next Header** value for ICMPv6 is set to the value **58**.
- ICMPv6 has a framework for extensions to implement **new features**.
  - **Neighbor Discovery Protocol** (**NDP**) is a node discovery protocol based on ICMPv6 which replaces and enhances functions of ARP.
  - **Multicast Listener Discovery** (**MLD**) is used byIPv6 routers for discovering multicast listeners on a directly attached link, much like Internet Group Management Protocol (IGMP) is used in IPv4.
  - **Multicast Router Discovery** (**MRD**) allows the discovery of multicast routers.
- Blocking of ICMPv6 on a Firewall (common practice of mostly inept administrators with ICMPv4) is nonsense and dangerous
- IPv6 wont work without ICMPv6

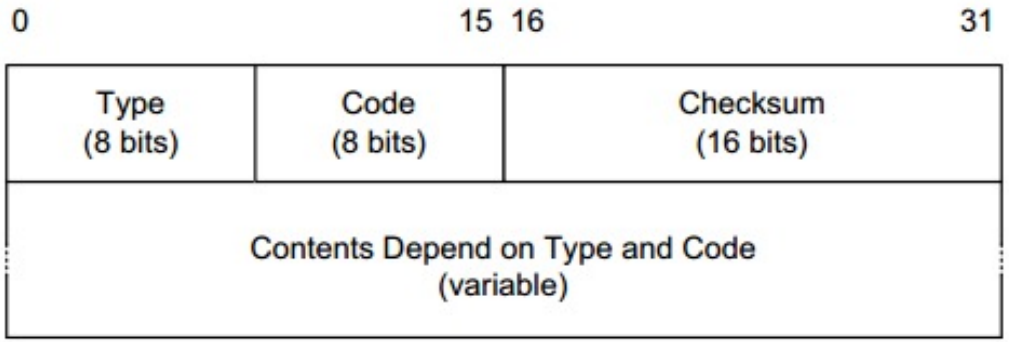
ICMPv6

ICMPv6 **header**: **TYPE** (8 bits), **CODE** (8 bits), and **Checksum** (16 bits).

TYPE: 0 to 127 (high-order bit is 0) indicate an **error message**,
TYPE: 128 to 255 (high-order bit is 1) indicate an **information message**.

CODE depends on the message type and provides additional information.

CHECKSUM provides a minimal level of integrity verification.



ICMPv6 – Error Code and type example

Type		Code	
Value	Meaning	Value	Meaning
ICMPv6 Error Messages			
1	Destination unreachable	0	no route to destination
		1	communication with destination administratively prohibited
		2	beyond scope of source address
		3	address unreachable
		4	port unreachable
		5	source address failed ingress/egress policy
		6	reject route to destination
		7	Error in Source Routing Header
2	Packet too big	0	
3	Time exceeded	0	hop limit exceeded in transit
		1	fragment reassembly time exceeded
4	Parameter problem	0	erroneous header field encountered
		1	unrecognized Next Header type encountered
		2	unrecognized IPv6 option encountered

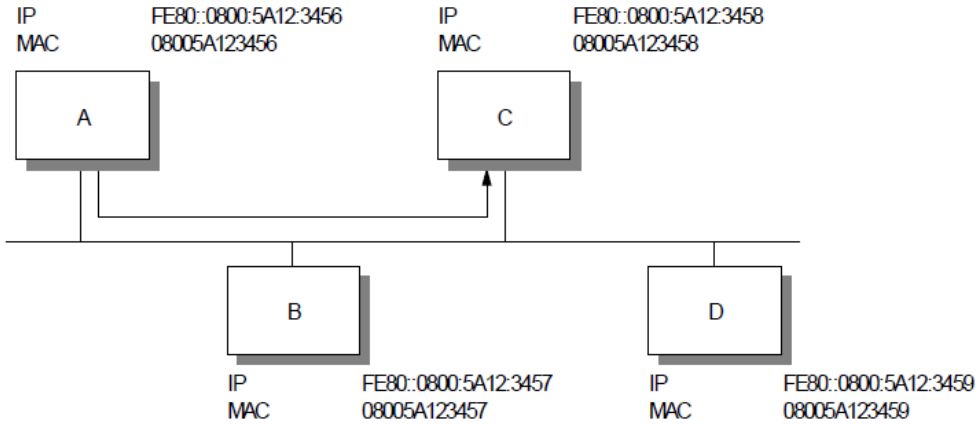
ICMPv6 - Informational code and type

Type		Code	
Value	Meaning	Value	Meaning
ICMPv6 Informational Messages			
128	Echo Request	0	
129	Echo Reply	0	
133	Router Solicitation (NDP)	0	
134	Router Advertisement (NDP)	0	
135	Neighbor Solicitation (NDP)	0	
136	Neighbor Advertisement (NDP)	0	
137	Redirect Message (NDP)	0	

ICMPv6 NDP Address Resolution

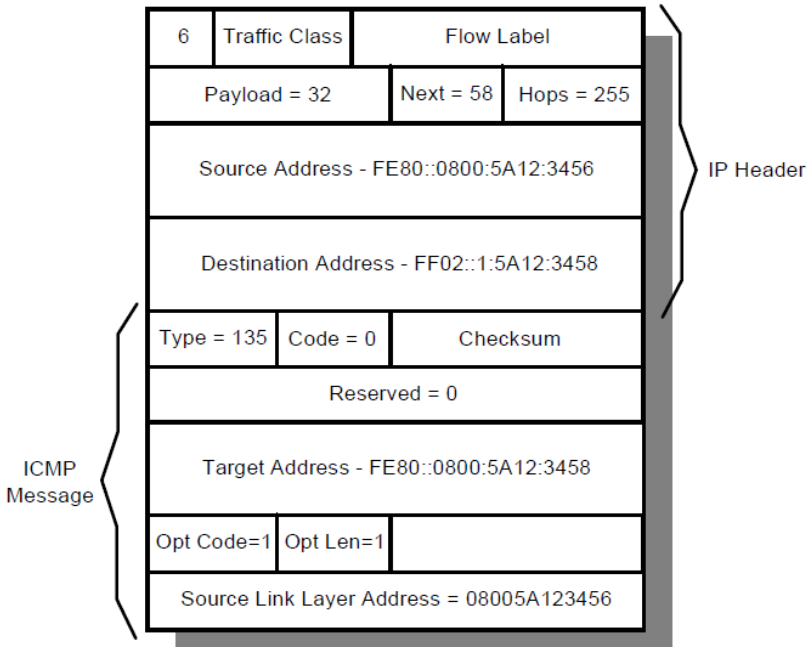
Workstation A needs to send data to workstation C. It knows the IPv6 address of workstation C, but it does not know how to send a packet, because it does not know its MAC address.

To discover this information, A sends a **Neighbor Solicitation** and C responds with a **Neighbor Advertisement**.



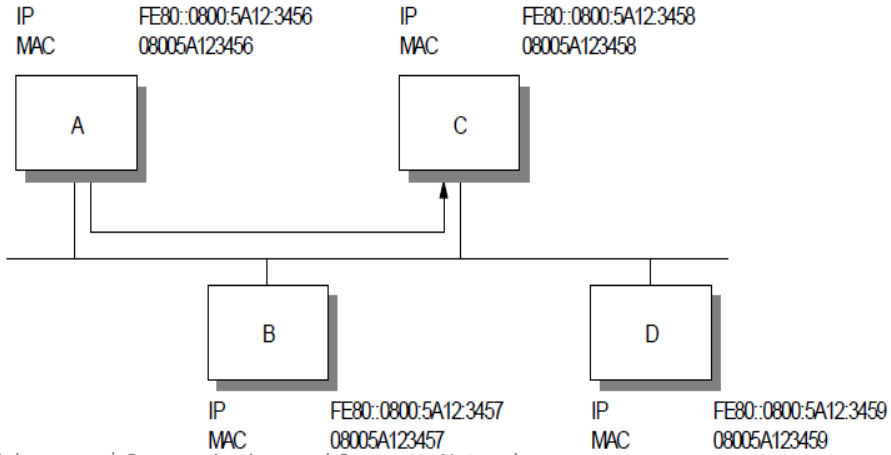
ICMPv6 NDP Address Resolution

Neighbor Solicitation



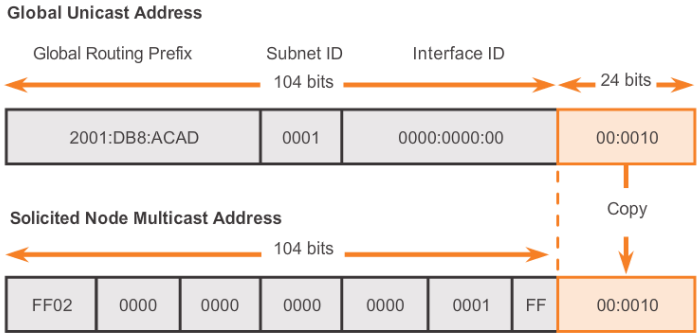
The IPv6 Destination Address is the **solicited node address** for the target workstation (a special type of **multicast**). Every workstation must respond to its own solicited node address but other workstations ignore it.

ICMP-Type is 135 (Neighbor Solicitation)



Solicited-Mulitcast Node Address - Layer 3

- Used for address lookup and duplicate address detection (DAD)
  - Every interface with IPv6 has a SMNA
  - Derived from Unicast or Anycast-address
  - Prefix: ff02::1:ff00:0/104
  - Last 24 Bit of Unicast or Anycast address were appended



IPv6 Global Unicast Address: 2001:DB8:ACAD:0001:0000:0000:0000:0010

IPv6 Solicited Node Multicast Address: FF02:0:0:0:0:1:FF00:0010

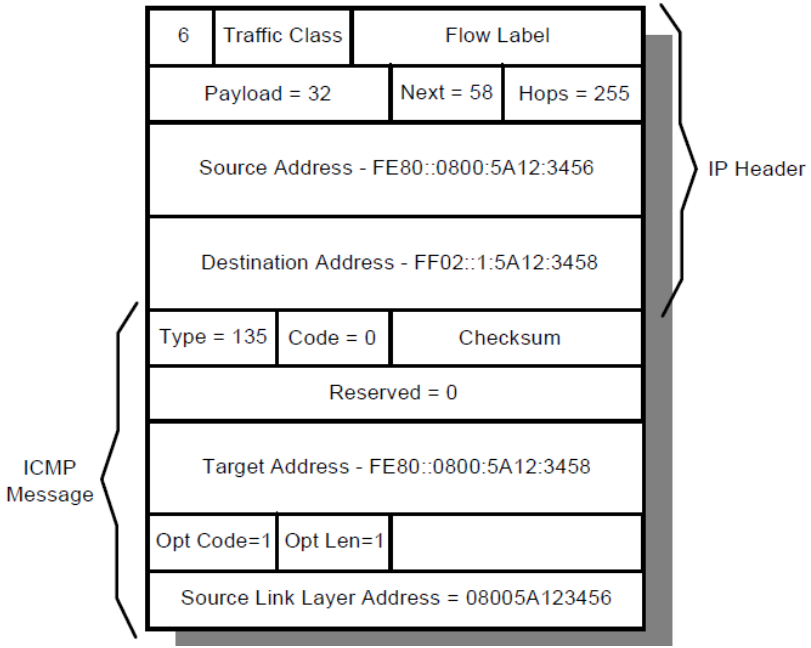
Solicited-Mulitcast Node Address - Layer 2

- Need for DAD and NDP
- Every interface listens for the MAC-address
- Creation:
  - Set first 16 Bit of the MAC to 33:33
  - Create last 32 Bit of Solicited-Mulitcast Node Address
- Example:
  1. Link Local Address: fe80::805:2bf8:fb25:31c9
  2. Append last 24 Bit at ff02::1:ff:0/104
  3. Now you have the SMNA: ff02::1:FF25:31c9
  4. Append these 32 Bit at MAC-address 33:33
  5. MAC-Address: 33:33:FF:25:31:c9

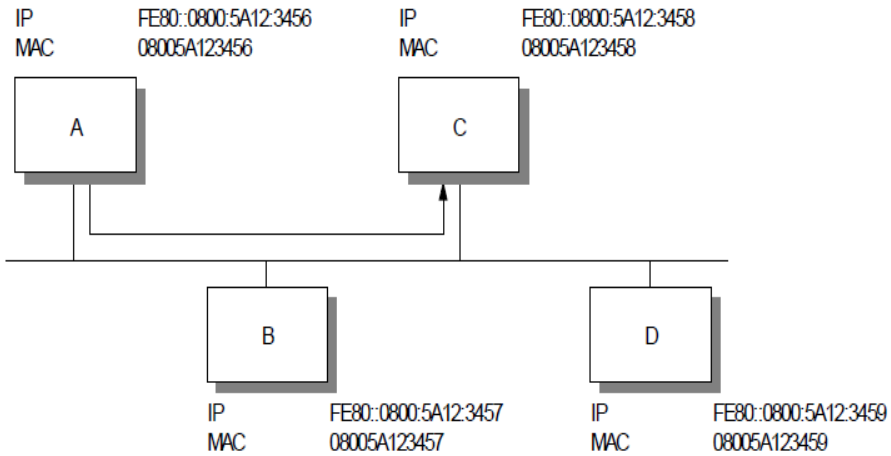
Fachhochschule Dortmund
University of Applied Sciences and Arts

ICMPv6 NDP Address Resolution

Neighbor Solicitation

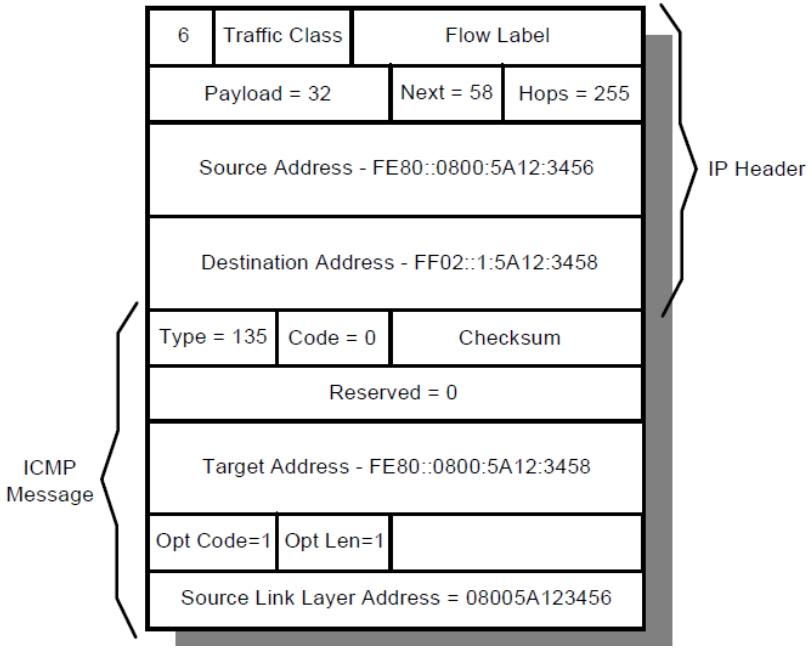


Target address is the known IP address of the target workstation.

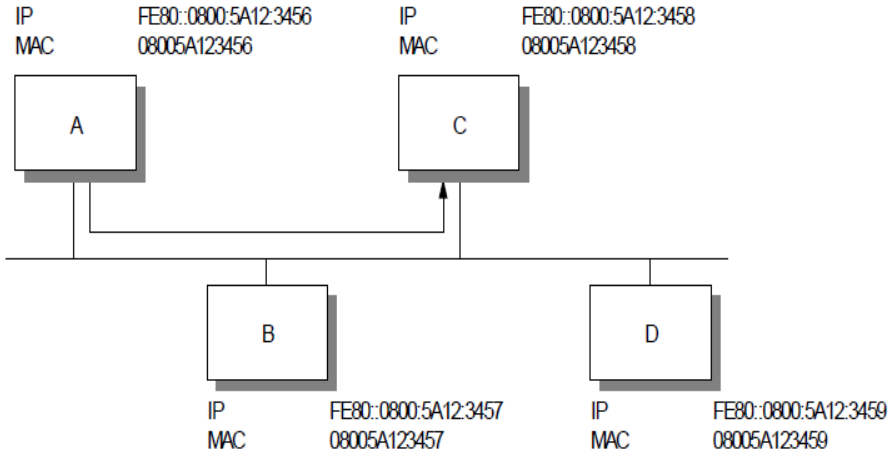


ICMPv6 NDP Address Resolution

Neighbor Solicitation



The Ethernet Source Address **(LLA)** is useful to the target workstation and saves it from having to initiate a neighbor discovery process of its own when it sends a packet back to the source workstation.

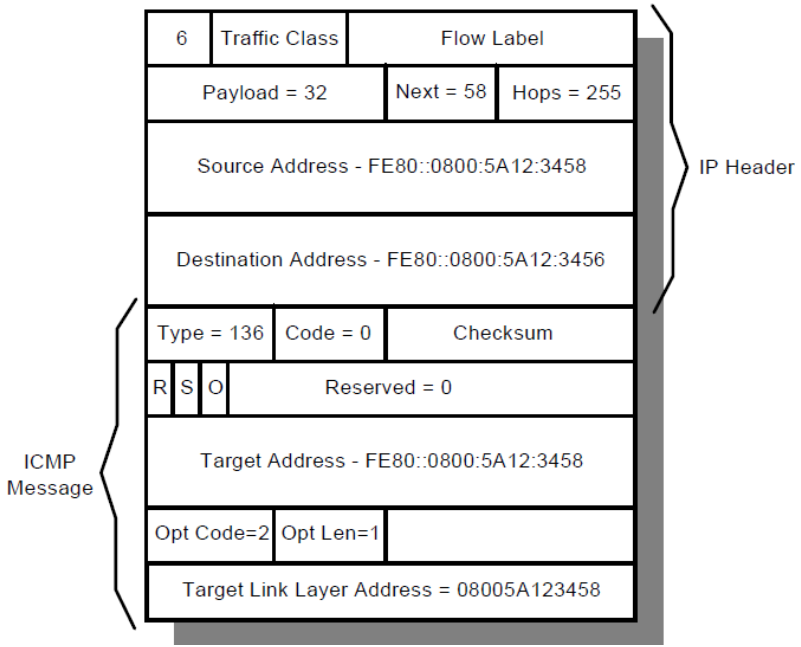


Solicited-Mulitcast Node Address - Layer 2 and 3

```
> Frame 13: 86 bytes on wire (688 bits), 86 bytes captured (688 bits)
v Ethernet II, Src: CeLink_df:00:d5 (a0:ce:c8:df:00:d5), Dst: IPv6mcast_ff:7f:f1:b5 (33:33:ff:7f:f1:b5)
  > Destination: IPv6mcast_ff:7f:f1:b5 (33:33:ff:7f:f1:b5)
  > Source: CeLink_df:00:d5 (a0:ce:c8:df:00:d5)
    Type: IPv6 (0x86dd)
v Internet Protocol Version 6, Src: ::, Dst: ff02::1:ff7f:f1b5
    0110 .... = Version: 6
  > .... 0000 0000 .... .... .... .... .... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT)
    .... 0000 0000 0000 0000 0000 = Flow Label: 0x00000
    Payload Length: 32
    Next Header: ICMPv6 (58)
    Hop Limit: 255
    Source Address: ::
    Destination Address: ff02::1:ff7f:f1b5
v Internet Control Message Protocol v6
    Type: Neighbor Solicitation (135)
    Code: 0
    Checksum: 0x0d27 [correct]
    [Checksum Status: Good]
    Reserved: 00000000
    Target Address: 2003:c5:e70d:c600:1cc4:50b8:6c7f:f1b5
  > ICMPv6 Option (Nonce)
```

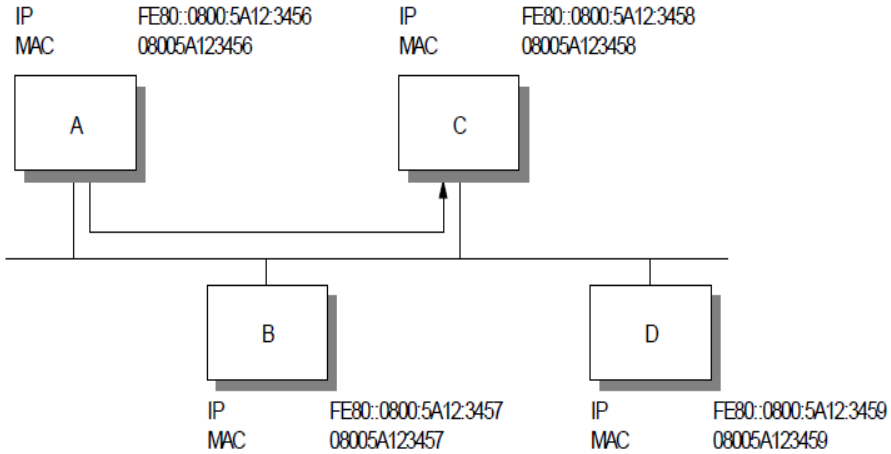
ICMPv6 NDP Address Resolution - Answer

Neighbor Advertisement



The neighbor advertisement is addressed directly back to workstation A. The ICMP message option contains the target IP address together with the **target's link layer (MAC) address**.

Answer has ICMP type 136: Neighbor Advertisement



ICMPv6 NDP Router and Prefix Discovery

A node needs to communicate not just with other nodes on the same link, but with nodes on other network segments that might be anywhere in the world.

In this case, there are two important pieces of information that a node needs to have:

- The address of a router that the node can use to reach the rest of the world.
- The prefix (or prefixes) that define the range of IP addresses on the same link as the node that can be reached without going through a router.

Routers use ICMP to convey this information to hosts with **Router Advertisements**.

ICMPv6 NDP Router and Prefix Discovery

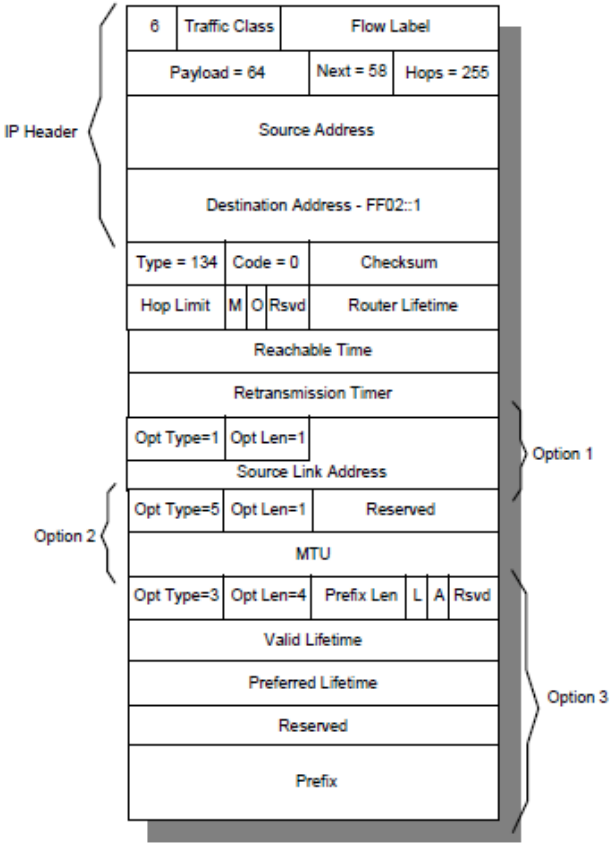
A router constantly sends unsolicited **Router Advertisements** at a frequency defined in the router configuration.

The Destination Address **FF02::1** is the special multicast address that defines all systems on the local link.

Router Lifetime: How long the node should consider this router to be available. If this time period is exceeded and the node does not receive another router advertisement message, the node should consider this router to be unavailable.

Reachable Time: This setting sets a parameter for all nodes on the local link. It is the time in milliseconds that the node should assume that a neighbor is still reachable after receiving a response to a neighbor solicitation.

Retransmission Timer: This field sets the time, in milliseconds, that nodes should allow between retransmitting neighbor solicitation messages if no initial response is received.



ICMPv6 NDP Router and Prefix Discovery

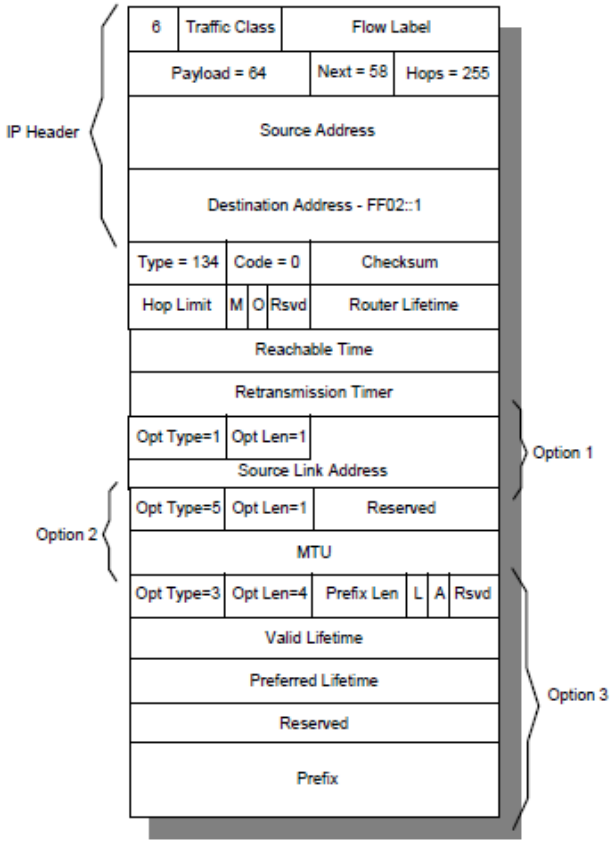
Router Advertisement generally has one or more attached options.

The three possible options in a router advertisement message are:

Option 1 (**Source Link Address**) Allows a receiving node to respond directly to the router without having to do a neighbor solicitation.

Option 5 (**MTU**) Specifies the maximum transmission unit size for the link. For some media, such as Ethernet, this value is fixed, so this option is not necessary.

Option 3 (**Prefix**) Defines the address prefix for the link. Nodes use this information to determine when they do, and do not, need to use a router. Prefix options used for this purpose have the L (link) bit set on. Prefix options are also used as part of address configuration, in which case the A bit is set on.



ICMPv6 NDP Router and Prefix Discovery

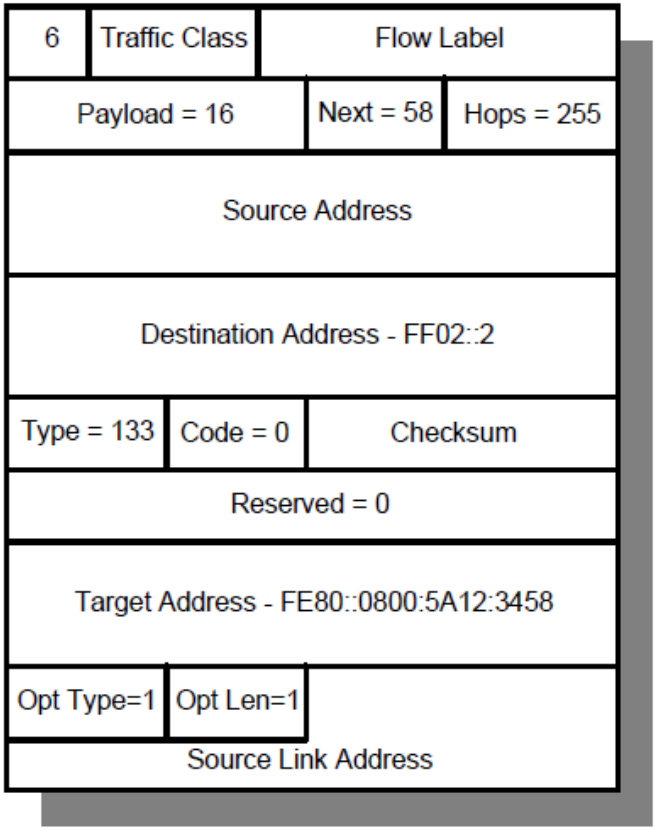
A node might want to obtain information about the nearest router without having to wait for the next scheduled advertisement (for example, a new workstation that has just attached to the network).

In this case, the node can send a **Router Solicitation** message.

The Destination Address **FF02::2** is the special multicast address that defines all routers on the local link.

Option 1 (**source link address**) Allows the receiving router to respond directly to the node without having to do a neighbor solicitation.

Each router that receives the solicitation message responds with a router advertisement sent **directly to the node** that sent the solicitation (not to the all systems link-local multicast address).



IPv6 SLAAC

SLAAC (**StateLess Address AutoConfiguration)** provides the ability to address a host based on a network prefix that is advertised from a local network router via **Router Advertisements** (RA).

RA messages are sent by default by most IPV6 routers; these messages are sent out periodically by the router and include information including: one or more IPv6 prefixes (Link-local scope), Prefix Lifetime information, Flag information, Default device information (Default router to use and its lifetime).

SLAAC is implemented on the **IPv6 client** by listening for the local RA's and then taking the prefix that is advertised to form a unique address that can be used on the network.

For this to work, the prefix that is advertised must advertise a **prefix length of 64 bits** (i.e., /64); SLAAC will then **dynamically form a host identifier** that is 64 bits long and will be **suffixed** to the end of the advertised prefix **to form an IPv6 address**.

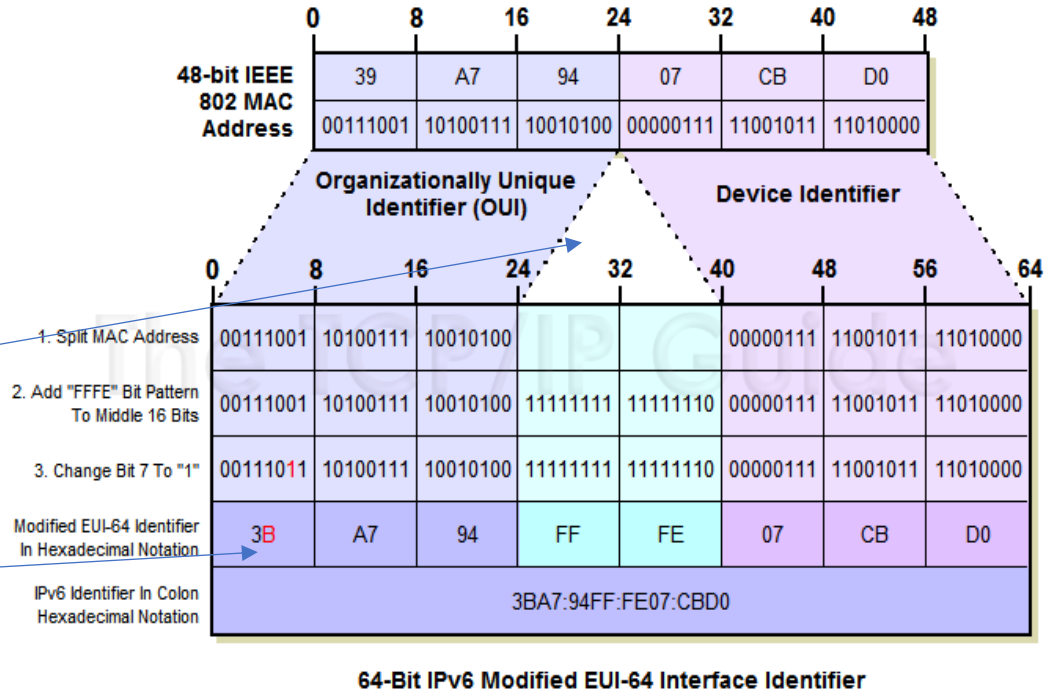
Originally, the host identifier was formed using the **EUI-64** rules (the same that are used to form link local addresses) and many devices still use this method.

Duplicate Address Detection

- If a device D has created an own ip-address (LLA with EUI-64 or GUA based on RA), D has to check for **duplicate addresses** in the network
- D sends an ICMPv6 Type 135 to the (its) SMNA, IPv6 src address is :: (unspecified)
- D waits a *few* seconds for an answer (RetransTimer defined in RFC 4861) Time depends on OS, TCP/IP-Stack, vendor and RFC implementation (i.e., RFC8415 and RFC4941), Optimistic DAD (RFC 4429), Enhanced DAD (RFC 7527)
- If D receives an answer (via ICMPv6 Type 136, NA), D calculates a new address, otherwise D keeps the address

Example: Determination of IEEE EUI-64 interface address

- EUI-64 (64-Bit Extended Unique Identifier) is an IEEE standardized MAC address format with a length of 64 bits
- MAC-48 is a 48 bit MAC address according to IEEE 802
- Insert FFFE between OUI and last 24 bits
- Changes the 7th bit to 1



64-Bit IPv6 Modified EUI-64 Interface Identifier

Privacy Extension (RFC 4941, RFC8981)

- EUI-64 links MAC-address and IP-address, which results in privacy issues
- The Privacy Extensions (RFC 4941, RFC 8981) allow a host to select random interface identifiers and change them over time
- PE prevents tracking of IPv6-addresses by providing a new algorithm for randomized IID creation
  - RID = F(Prefix, Net_Iface, Network_ID, Time, DAD_Counter, secret_key)
- Global prefix and randomized IID create temporary IPv6 address
- Use of PE is OS-dependet (Linux-Kernel):

```
use_tempaddr - INTEGER
Preference for Privacy Extensions (RFC3041).
<= 0 : disable Privacy Extensions
== 1 : enable Privacy Extensions, but prefer public addresses over temporary addresses.
> 1 : enable Privacy Extensions and prefer temporary addresses over public addresses.
```

OS commands

- IP-address configuration:
  - `ifconfig`
  - `ip addr show`
  - `ipconfig`

- Neighbor Discovery Table:
  - `ndp`
  - `ip -6 neighbor`
  - `netsh interface ipv6 show neighbors`

```
fe80::462:451c:9908:9b1b%en0          6e:98:e0:10:bf:3b     en0
fe80::106b:e4c1:ab04:d4e%en0          96:0:eb:9d:c0:5e      en0
```

Fachhochschule
Dortmund
University of Applied Sciences and Arts

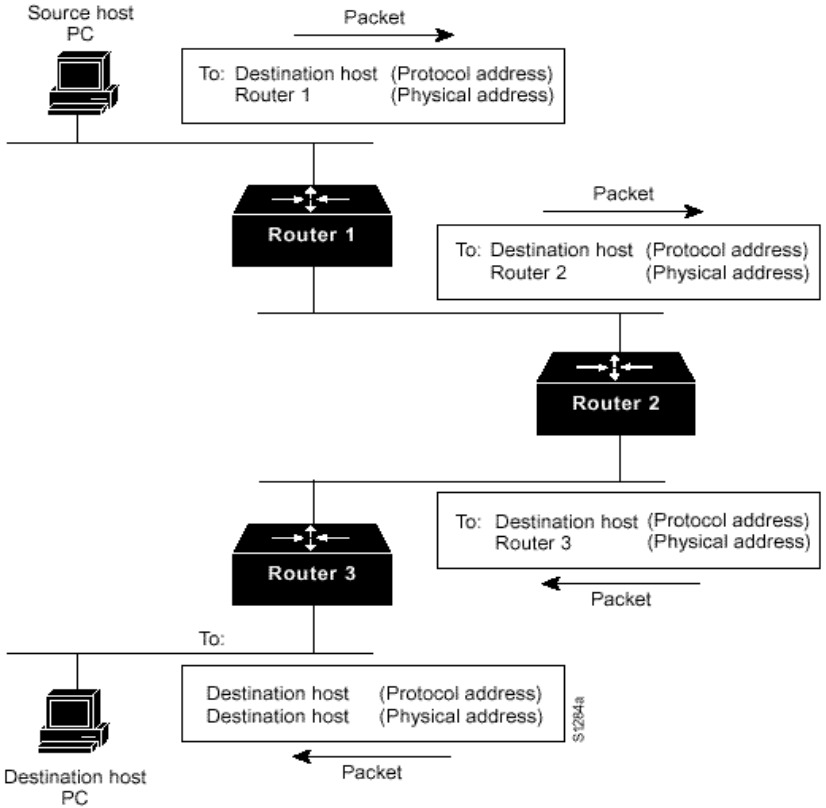
Network Layer – Hardware

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

Router and Gateway

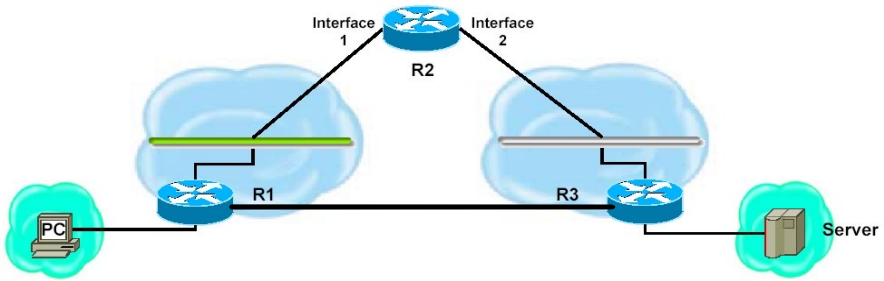
A **transit system**, which is called a **router** in TCP/IP parlance, is a special computer that is connected to two or more networks and transmits packets from one network to another.

A **gateway** is used to connect two or more dissimilar networks and translates between the different protocols



Router

- A router represents an OSI layer 3 device, so it can connect networks with different layer 1 and 2 topologies. However, all networks connected through a router must use the same addressing mechanisms.
- To be able to forward packets between the connected partial LANs, a router interprets the addresses on Layer 3. So, it does not work with the addresses of the MAC layer.
- In networks that are coupled via routers, the transmitting station does not need to know the MAC address of the target station in order to be able to address it – the Destination address from the protocol level (e.B. the IP address) is required.



Router internals

- This allows packets to be forwarded from one network segment to another, regardless of the topology of the connected networks. In addition, routers enable load balancing by using alternative routes to the destination address if a bottleneck occurs or a route fails completely.
- Routing tables are used for forwarding/distributing the packets.

DESTINATION		INTERFASE	NEXT HOP	

- Due to the need to read level 3 addresses from the packets and forward them based on the routing tables, the delay caused by routers is correspondingly greater than with bridges. In addition, routers must of course understand the corresponding network protocols in order to be able to interpret them. The number of supported protocols is therefore an important criterion.

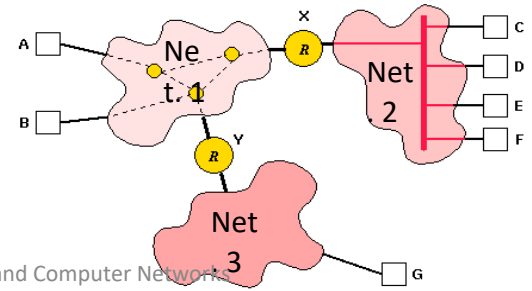
Fachhochschule
Dortmund
University of Applied Sciences and Arts

Network Layer – Routing / NAT

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science

Pathfinding in (connectionless) networks

1. Necessary when allocation check shows, that the communication partner is NOT in the same subnet.
2. Per default, these packets are sent to the default gateway
3. The default gw checks again, if the intended receiver is in its network, otherwise packet is sent to the next gateway (=router)
4. At least one router should have access to the intended receiver
5. This router delivers the network packet
6. If the receiver wants to send a response, a check for the allocation will be done
7. Process restarts (but from the other “side”)



Network Address Translation (NAT)

Faced with the "**scarcity**" of **IPv4 addresses**, starting in the 1990s, a workaround was chosen that consists of putting "private addresses" in corporate networks and SOHO (small office home office) networks, and in the perimeter with the Internet do **NAT** (Network Address Translation).

RFC 1918 reserves the following netblocks for use on **private networks**:

10.0.0.0	-	10.255.255.255/8
172.16.0.0	-	172.31.255.255/12
192.168.0.0	-	192.168.255.255/16

IP addresses from these ranges are not routed on the Internet and are intended for internal use in LANs.

NAT can be divided into two forms:

- Source Network Address Translation (SNAT):
  The source address in the header of the IP packet is exchanged.
- Destination Network Address Translation (DNAT):
  The destination address is exchanged in the header of the IP packet.

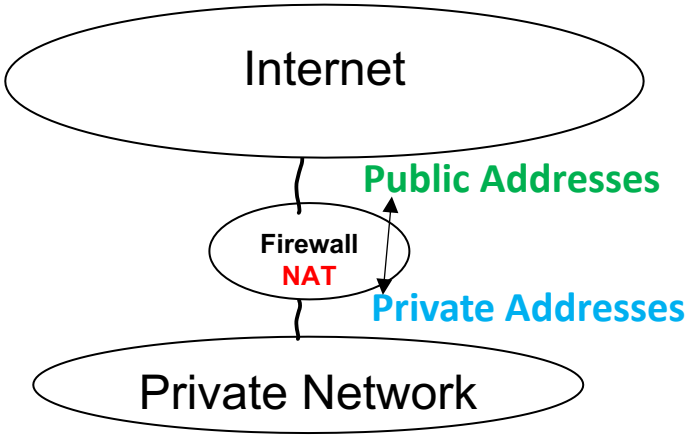
Network Address Translation (NAT)

The **perimeter control router** or **firewall** between the private network and the Internet
must implement the NAT (Network Address Translation) mechanism.

NAT **translates the Source Address** of the datagram **going out** to the Internet,
and **translates the Destination Address** of the datagram **going into** the private network.

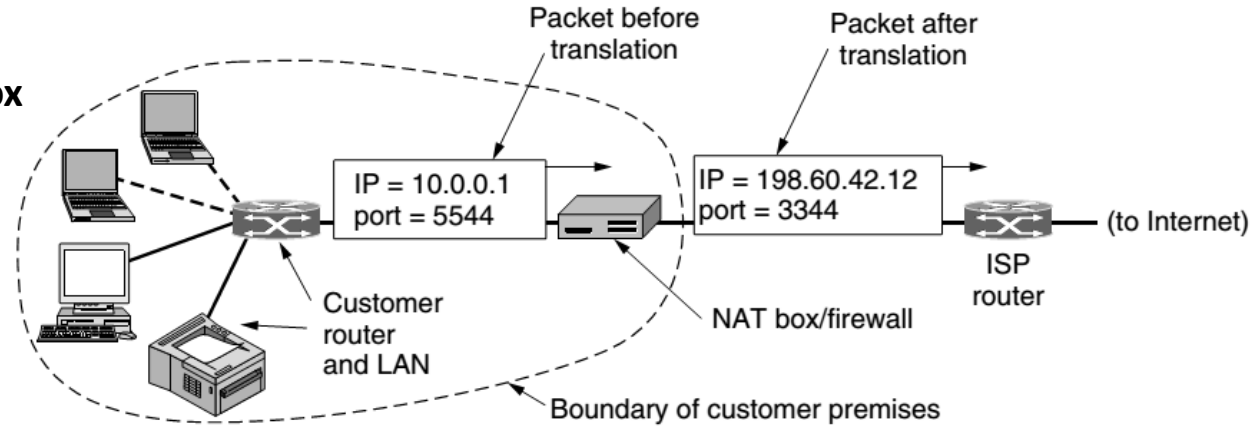
Types o NAT:

- One to one mapping (1 private to 1 public address)
- Many to one mapping (many private to 1 public address)
  - Using **PAT** (Port Address Translation)
  - Also called “masquerade” NAT



Network Address Translation (NAT)

Positioning and operation of a NAT box



Example ”PAT Translation Table" for two TCP connections to the same web server

Dir	Fields	Old value	New value
out	ip.src:tcp.srcport	10.0.0.1:30000	198.60.42.12:62000
out	ip.src:tcp.srcport	10.0.0.3:30001	198.60.42.12:48231
in	ip.dst:tcp.dstport	198.60.42.12:48231	10.0.0.3:30001
in	ip.dst:tcp.dstport	198.60.42.12:62000	10.0.0.1:30000

IP Routing

- It is the process of transferring a datagram from one host to another through intermediary nodes (routers or gateways).
- IP datagram forwarding is based on the **destination address** (DABR Destination Address Based Routing)
- The IP entity of every host or router uses an **IP Routing Table**
  - indicates by which **INTERFACE** each **DESTINATION** is reached
    - n a router it also indicates if it is reached **Directly** or **Indirectly**
  - in the case of an Indirect route, it indicates the **NEXT HOP**
  - If there is no route to the destination address, the IP entity drops (**DROP**) the datagram.

DESTINATION		INTERFASE	NEXT HOP	

Routing tables

- For the gateways to be able to forward packets correctly on their way to their destination, they need routing information that is stored in route tables.
- The route tables can be static or dynamic.
  - Static entries are entered manually.
  - Dynamic entries are generated using routing protocols.
- A route table contains pairwise entries from network addresses and associated gateways as well as the assigned interface and additional information such as the "costs" of the route (the so-called metric).
- To forward a packet, the *network part* of the destination address is looked up in the routing table and the packet is sent either directly or to the associated gateway via the corresponding interface.

Default gateway

- Since this type of routing is based on network addresses and not on computer addresses, the routing table can be kept relatively short.
- When evaluating the route table, "more concrete" entries with a longer netmask take precedence over "more general" ones at gateway (GW) via interface.
- Finding the destination subnet address is done in the following order:
  Find host destination address -> send to GW via interface
  Find network destination address -> send to GW via interface
  use default entry (0.0.0.0)-> send to default GW via interface
- Principle of the "Longest Match"
  A "match" exists if: (destination IP address & gene mask) == Destination
- The default entry always returns the "shortest match", as the following always applies:
  (destination IP address & 0.0.0.0) == 0.0.0.0

IP Classful & Classless Routing
FLSM & VLSM support

- **IP Classful Routing**
  uses ancient Network Classes for routing decisions (not used anymore)
- **CIDR Classless Inter Domain Routing** (routing tables with **routemask**)
  Supports FLSM, VLSM and route summarization/aggregation
  - The IP entity receives the datagram, and examine its Destination Address (DA)
  - If the DA is that of some local interface, it delivers the datagram to the upper layer indicated in Protocol
  - If the DA is NOT that of some local interface, it goes through the **Routing Table** in descending order of Route_Mask looking to "**match**" a **route**, for which it is set if **DestAddress & Route_Mask = Route_Destination**
  - If it matches more than 1 route, uses the route with the **highest mask** (the most specific)
  - If it matches more than 1 route with the same mask, use the route with the **lowest metric**
  - When the DA “matches” a route, it **forwards** the datagram along that route (directly or indirectly).
  - If it does not match any route, the IP entity does the **DROP** of the datagram

IP Direct Routing

Direct Routes (or connected routes)

- Any Direct Route allows reaching the destination hosts without going through any intermediary router, that is, with **METRIC 0**.
- The corresponding route is added to the routing table when an IP interface is configured with an IP address and mask.
- The destination address is obtained by making the AND between the IP address and the mask of the interface.

DESTINATION	MASK	INTERFASE	NEXT HOP	METRIC
Network_address	*Netmask*	*Interfase*	-	*0*

IP Indirect Routes

Indirect Routes (or remote routes)

- Any Indirect Route allows reaching the destination hosts through one or more intermediary routers, that is, with **METRIC >0**.
- The **NEXT HOP** indicates the IP address of the next router to reach the destination. The next hop should be reachable across a direct route.
  - **Static routing**: when indirect routes are added **manually** to the Routing Table
  - **Dynamic routing**: when Indirect Routes are added dynamically by a routing process that obtain routing information using some **routing protocol**.

DESTINATION	MASK	INTERFASE	NEXT HOP	METRIC
Network_address	*Netmask*	*Interfase*	*Next hop*	*>0*

IP CIDR & Route Aggregation/Summarization

- Routing tables might increase due to the number of connected subnets

DESTINATION	MASK	INTERFASE	NEXT HOP	METRIC
192.168.0.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*
192.168.1.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*
192.168.2.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*
192.168.33.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*
192.168.34.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*

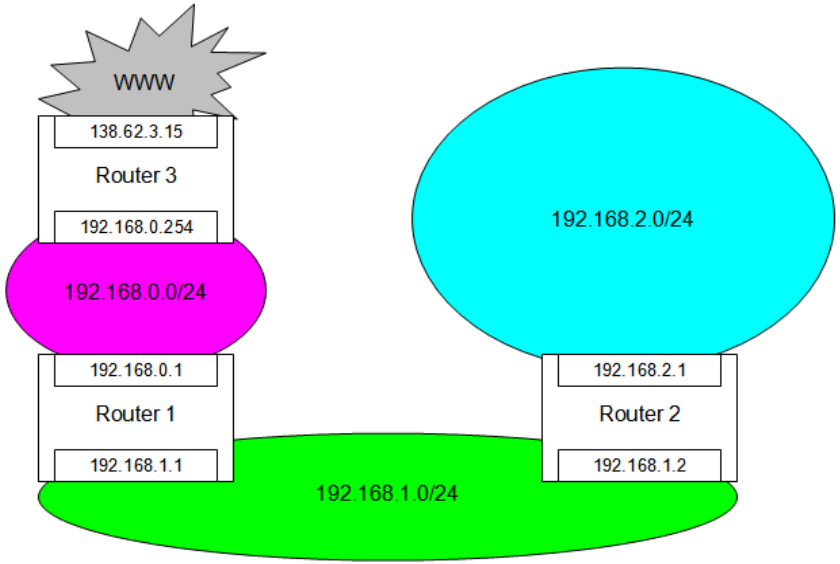
IP CIDR & Route Aggregation/Summarization

- Route Summerization with CIDR reduces the number of entries

DESTINATION	MASK	INTERFASE	NEXT HOP	METRIC
192.168.0.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*
192.168.1.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*
192.168.2.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*
192.168.33.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*
192.168.34.0	*255.255.255.0*	*eth0*	*10.0.0.3*	*>0*

DESTINATION	MASK	INTERFASE	NEXT HOP	METRIC
192.168.0.0	*255.255.254.0*	*eth0*	*10.0.0.3*	*>0*

Example Routing Table (Router 2)



Network address or IP address	Subnet or Netmask	Gateway or router	Interface	Metric (Number of hops)
192.168.0.0	255.255.255.0	192.168.1.1	192.168.1.2	2
192.168.1.0	255.255.255.0	192.168.1.2	192.168.1.2	1
192.168.2.0	255.255.255.0	192.168.2.1	192.168.2.1	1
0.0.0.0	0.0.0.0	192.168.1.1	192.168.1.2	3

Fachhochschule Dortmund
University of Applied Sciences and Arts

Routing table of a Linux server

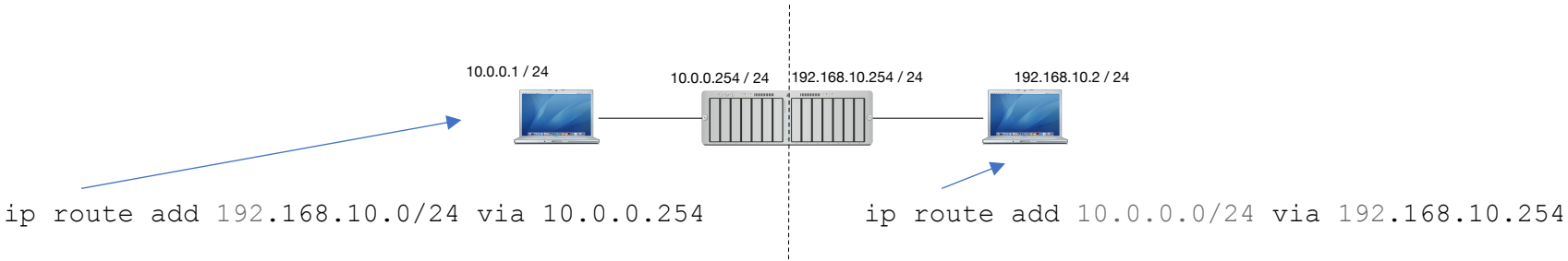
```
-bash-2.05b$ route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags   Metric  Iface
193.25.22.64    0.0.0.0         255.255.255.192 U       0       eth0
169.254.0.0     0.0.0.0         255.255.0.0     U 0     U 0     eth0
127.0.0.0       0.0.0.0         255.0.0.0       UG      0       lo
0.0.0.0         193.25.22.66    0.0.0.0                         eth0
```

Explanations of the flags:

U	(route is up)
H	(target is a host) G (use gateway)
R	(reinstate route for dynamic routing)
D	(dynamically installed by daemon or redirect)
M	(modified from routing daemon or redirect)
A	(installed by addrconf) C (cache entry)
!	(reject route)

Static routing

- Manually crafted routing tables are easy to manage
- Low CPU overhead
- Inflexible if changes in the network occur (manual reconfiguration needed)
- Configuration of a route on a linux system:



- Router needs ip_forwarding enabled:

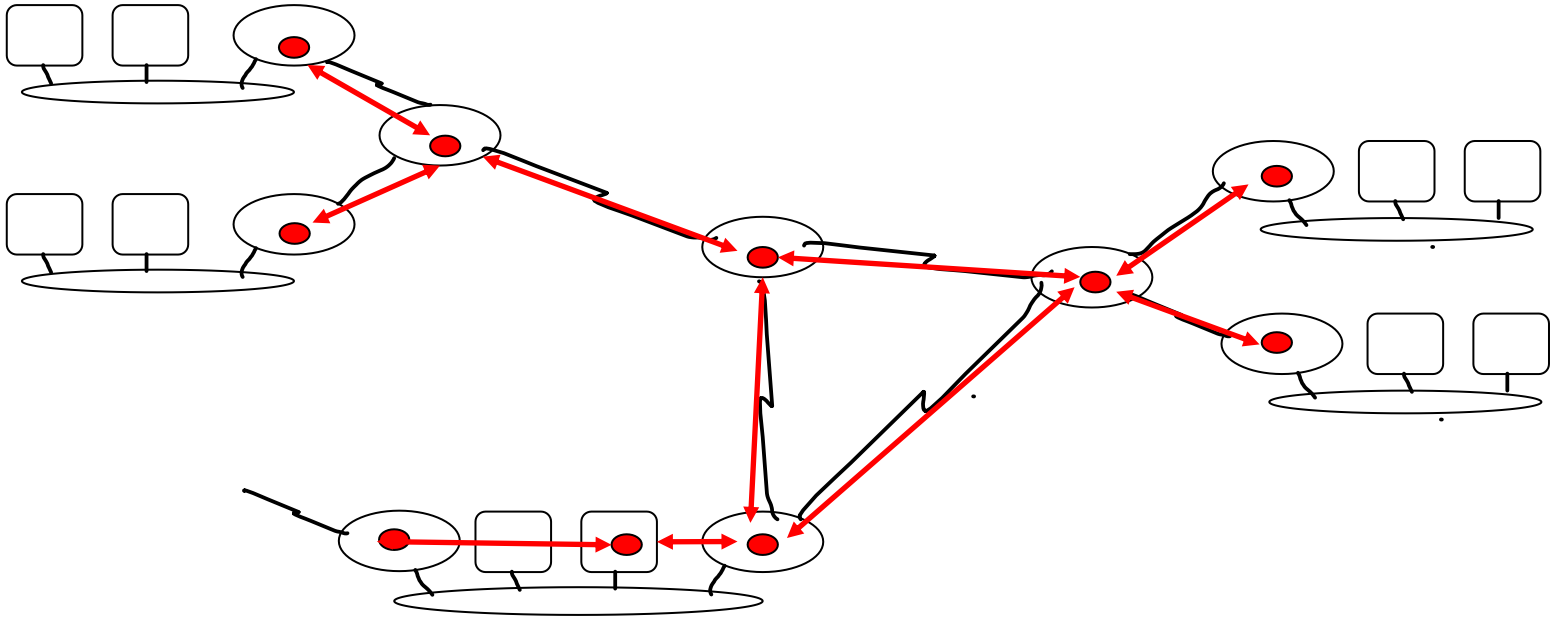
```
sysctl -w net.ipv4.ip_forward=1
```

Cisco Router Routing Table

```
ipv6gw#sh ip route
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
C       172.22.178.0/24 is directly connected, FastEthernet0/0.178
C       172.22.179.0/24 is directly connected, FastEthernet0/0.179
C       172.22.176.0/24 is directly connected, FastEthernet0/0.176
C       172.22.177.0/24 is directly connected, FastEthernet0/0.177
C       172.22.180.0/24 is directly connected, FastEthernet0/0.180
S       172.22.0.0/16 [1/0] via 192.168.255.217
     193.25.22.0/26 is subnetted, 1 subnets
C       193.25.22.64 is directly connected, FastEthernet0/1.8
     10.0.0.0/24 is subnetted, 1 subnets
C       10.11.10.0 is directly connected, FastEthernet0/0.10
     192.168.255.0/29 is subnetted, 1 subnets
C       192.168.255.216 is directly connected, FastEthernet0/1.112
S*   0.0.0.0/0 [1/0] via 193.25.22.66
```

Dynamic Routing with Routing Protocols

Routing processes dialog between routers to exchange routing information.



Routing Protocols

- In simple cases, you can set the entries in the routing tables of the end systems and routers manually.
- Larger networks are usually "living" structures, i.e. there are constant changes in the network topology. In the simplest case, only end systems are added or removed. However, failures of routers or lines are also to be expected, which requires a bypass of the "fault point".
- Routing protocols perform a (more or less) automatic setting of the routing tables. The procedure within an area of responsibility (intradomain) and between the areas of responsibility of different network operators (interdomain) differs.
- Routing protocols are used for communication between gateways. On the one hand, they ensure that gateways can be found at all, and on the other hand, the routing tables are constantly updated via the routing protocols.

Common routing protocols

- Intradomain routing protocols:
  - RIP – Routing Information Protocol: RFC 2453 (RIP 2)
  - OSPF – Open Shortest Path First: RFC 2328 (OSPF 2)

- Interdomain routing protocols:
  - BGP – Border Gateway Protocol: RFC 1771 (BGP4)
    BGP is a further development of the EGP, which in particular allows "policy" criteria to be included in the routing.

Basic Routing Algorithm Procedures

- Distance vector protocols
  - Tell your neighbors what their own route table looks like
  - Each node calculates the distance (cost) between itself and all the others within a network domain or AS and stores this information in a table.
  - Each node sends its table to all neighboring nodes.
  - When a node receives its neighbors distance tables, it calculates the shortest path to the other nodes and updates its table to reflect the changes.
- Link-State routing protocols
  - Tell everyone in a certain area who your neighbors are
  - Dijkstra's algorithm fixes a single node as the "source" node and finds shortest paths from the source to all other nodes in the topology graph, producing a shortest-path tree.

Routing Information Protocol

RIPv1

- RFC 1058 (1988)
- Distance vector routing protocol
- slow convergence
- Hop Metric
- Active & passive RIP systems
- Does NOT propagate the Netmasks of the routes
  - Does NOT support VLSM and CIDR

RIPv2

- RFC 1723 (1994)
- Propagate the Netmasks of the routes
  - Supports VLSM and CIDR
- Supports a simple authentication mechanism
- Hop Metric
- Triggered Updates

OSPF

Open Shortest Path First

- Proposed by the IETF (RFC 1247)
- Internal Gateway Protocol
- Link State (SPF) algorithm
- Supports VLSM and CIDR
- Authentication schemes
- Supports host-specific routes and network-specific routes.
- Allows multipath routing and load balancing

Fachhochschule
Dortmund
University of Applied Sciences and Arts

Additional slides

Prof. Dr. Daniel Spiekermann
Faculty of Computer Science