

Computer Networks

Network Layer - Internet Protocol

Prof. Dr. Oliver Hahm

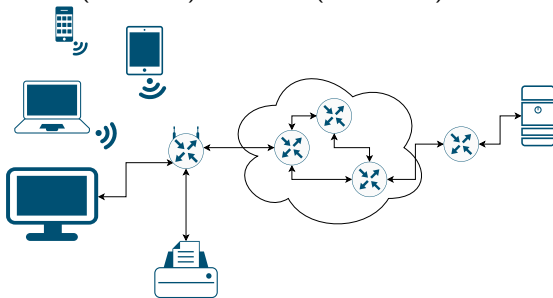
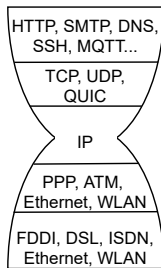
Frankfurt University of Applied Sciences
Faculty 2: Computer Science and Engineering
oliver.hahm@fb2.fra-uas.de
<https://teaching.dahahm.de>

December 07, 2021

The Narrow Waist of the Internet

Tasks of the Network Layer:

- Inter-Networking
- Providing logical addresses
- Forwarding packets
- Finding the best path → Routing
- Devices: Router
- Protocols: IPv4 (RFC 791) and IPv6 (RFC 2460)



Agenda

- Addressing
 - Purpose and Format
 - IPv4 Networks and Subnets
 - Private Networks and NAT
 - Fragmentation
 - IPv6 Networks

- Packet Structure
 - IPv4 Packet Structure
 - IPv6 Packet Structure

- ICMP

- Address Autoconfiguration

Addressing in the Network Layer

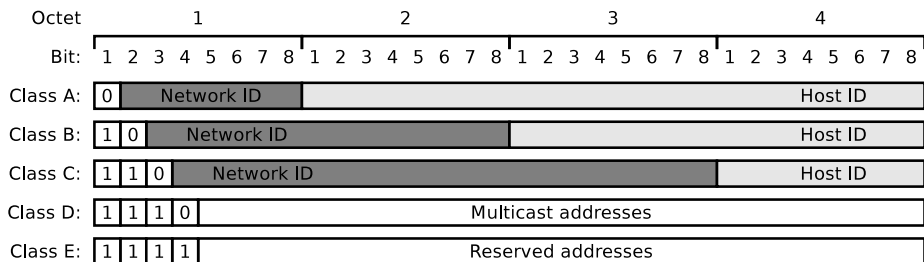
- Physical addresses (→ MAC addresses) are bound to a device
⇒ it is impossible to maintain a **logical hierarchy** or replace hosts in a **transparent** manner
- **Logical addresses** are required, which are independent from the specific hardware
 - Logical addressing separates the **logical position** within the network from a physical device

Address Assignment

For local networks manual address assignment is typically not desired, hence mechanisms for *address autoconfiguration* are required.

Network Identifier and Host Identifier

- Originally, IPv4 addresses were categorized into classes from A to C
 - Additionally, the classes D and E for special purposes existed
- The 32 bits of an IPv4 address are split into
 - Network identifier** (network ID)
 - Host identifier** (host ID)



- Class A: 7 bits for the network ID and 24 bits for the host ID
- Class B: 14 bits for the network ID and 16 bits for the host ID
- Class C: 21 bits for the network ID and 8 bits for the host ID

Address Classes

- The **prefixes** specify the address classes and their address ranges

Class	Prefix	Address range	Network ID	Host ID
A	0	0.0.0.0 - 127.255.255.255	7 bits	24 bits
B	10	128.0.0.0 - 191.255.255.255	14 bits	16 bits
C	110	192.0.0.0 - 223.255.255.255	21 bits	8 bits
D	1110	224.0.0.0 - 239.255.255.255	—	—
E	1111	240.0.0.0 - 255.255.255.255	—	—

- $2^7 = 128$ class A networks with a maximum of $2^{24} = 16,777,216$ host addresses each
- $2^{14} = 16,384$ class B networks with a maximum of $2^{16} = 65,536$ host addresses each
- $2^{21} = 2,097,152$ class C networks with a maximum of $2^8 = 256$ host addresses each
- Class D contains **multicast** addresses
- Class E is reserved for future purposes and experiments

Address Autoconfiguration

IPv4 has no builtin-mechanism for routable addresses. As a consequence, for local networks an additional protocol, like *DHCP* is required to assign the *host addresses*.

Drawback of Address Classes

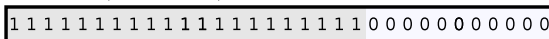
- The original intention was to identify physical networks in an unique way via the network ID
- Drawbacks of Address Classes:
 - It is impossible to dynamically adjust them
 - Many addresses are wasted
 - A class C network with 2 devices wastes 253 addresses
 - The address space of class C networks is quite small
 - A class B network with 256 devices wastes $> 64,000$ addresses
 - Only 128 class A networks exist
 - Migrating multiple devices to a different network class is complex task
- Solution: Logical networks are divided into **subnets**
 - 1993: Introduction of the **Classless Interdomain Routing** (CIDR)

Subnetting

Class B IP address



Subnet mask (255.255.248.0)



A part of the hosts IP address includes the subnet identifier



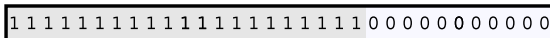
- For creating subnets, a (sub-)netmask is required
 - All hosts in a network have a subnet mask assigned
 - Length: 32 bits (4 bytes)
 - It is used to specify the number of subnets and hosts
- The subnet mask splits the host ID of an IP address into subnet ID and host ID
 - The network ID remains unchanged
 - The network mask adds another level of hierarchy into the IP address

The Structure of a Subnet Mask

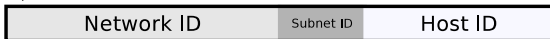
Class B IP address



Subnet mask (255.255.248.0)



A part of the hosts IP address includes the subnet identifier



- Structure of the **subnet mask**:
 - **1-bits** indicate, which part of the address space is used for **subnet IDs**
 - **0-bits** indicate, which part of the address space is used for **host IDs**
- Example: Splitting a class B network into 20 subnets requires 5 bits
 - Each subnet requires its own subnet ID and it must be represented in binary form
 - If 5 bits are used for the representation of the subnet IDs, 11 bits remain for host IDs

Syntax of the Classless Interdomain Routing (CIDR)

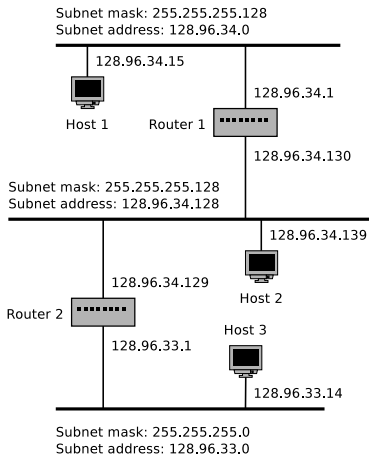
- According **CIDR** IP address ranges are represented by this notation:
First address/mask bits
 - The number of mask bits indicates the number of 1-bits (**prefix**) in the subnet mask
- The table shows the possible splits of a class C network into subnets

Mask bits (prefix)	/24	/25	/26	/27	/28	/29	/30	/31	/32
Subnet mask	0	128	192	224	240	248	252	254	255
Subnet bits	0	1	2	3	4	5	6	7	8
Subnets IDs	1	2	4	8	16	32	64	128	256
Host bits	8	7	6	5	4	3	2	1	0
Host IDs	256	128	64	32	16	8	4	2	—
Hosts (maximum)	254	126	62	30	14	6	2	0	—

2 Host IDs cannot be assigned to network devices, because each (sub-)network requires...

- an address for the network itself (all host ID bits are 0 bits)
- a broadcast address to address all devices in network (all bits of the host ID are 1 bits)

Subnets and Routing



- All hosts inside the **same subnet** have the **same subnet mask**
- If a host wants to **transmit** a packet, it performs a logical **AND** operation for its **own subnet mask** and the **destination IP address**
 - If the result is equal to the subnet address of the sender, the sender learns that the destination is inside the same subnet
 - If the result does not match the subnet address of the sender, the packet must be transmitted to a router, which forwards it to another subnet

Source: Computernetzwerke. Peterson and Davie. dpunkt (2000)

Agenda

- Addressing
 - Purpose and Format
 - IPv4 Networks and Subnets
 - Private Networks and NAT
 - Fragmentation
 - IPv6 Networks

- Packet Structure
 - IPv4 Packet Structure
 - IPv6 Packet Structure

- ICMP

- Address Autoconfiguration

Private IP Address Spaces

- In **private networks**, it is also required to assign IPs to network devices
 - These addresses are not allowed to interfere with global accessible internet services
- Several address spaces exist, containing private IP addresses
 - These address spaces are **not routed** in the internet

Address space: 10.0.0.0 to 10.255.255.255

CIDR notation: 10.0.0.0/8

Number of addresses: $2^{24} = 16,777,216$

Address class: Class A. 1 private network with 16,777,216 addresses

Address space: 172.16.0.0 to 172.31.255.255

CIDR notation: 172.16.0.0/12

Number of addresses: $2^{20} = 1,048,576$

Address class: Class B. 16 private networks with 65,536 addresses each

Address space: 192.168.0.0 to 192.168.255.255

CIDR notation: 192.168.0.0/16

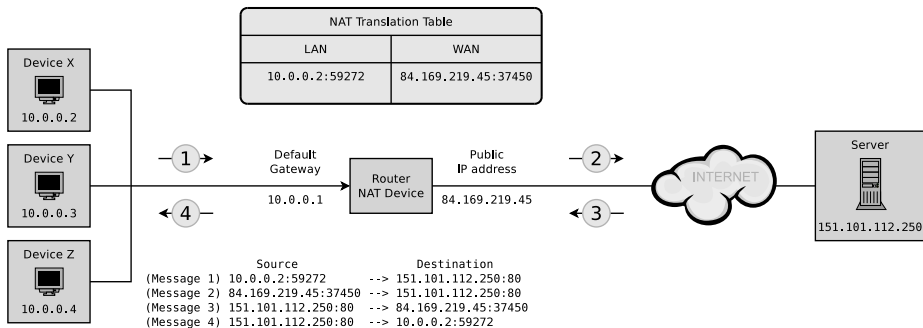
Number of addresses: $2^{16} = 65,536$

Address class: Class C. 256 private networks with 256 addresses each

Network Address Translation (NAT) (1/5)

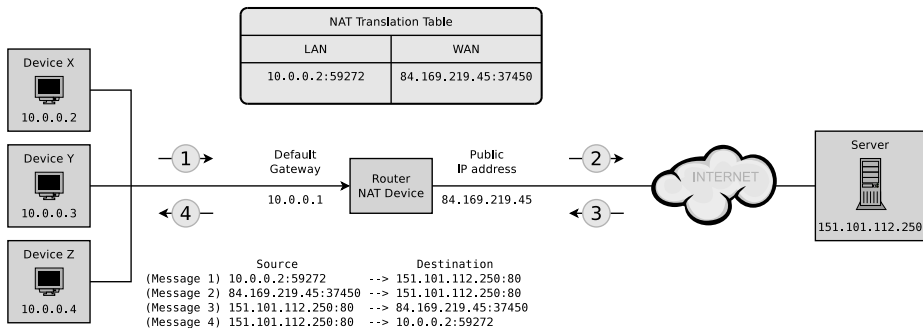
- **Problem:** Few households, businesses and educational/research institutions have enough public IPv4 addresses to equip all their network devices with **globally routable** IPs
 - Therefore, LANs usually use a **private IPv4 address space**
 - How can network devices in private networks communicate with network devices that have globally accessible addresses?
- **Solution:** **Network Address Translation (NAT)**
 - The local router presents itself as the source of those IP packets that it forwards from the directly connected private network to the Internet
 - In addition, it forwards incoming replies to the participants in the directly connected private network

Network Address Translation (2/5)



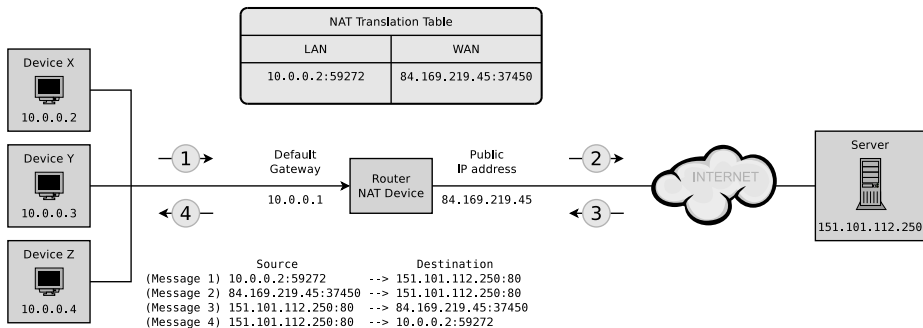
- Clients X, Y, and Z are inside a network with a **private IP address range**
- Only the **router** has a **globally** routable IP address
 - It does appear to the outside world as just a network device with a single public IP address and not as a router

Network Address Translation (3/5)



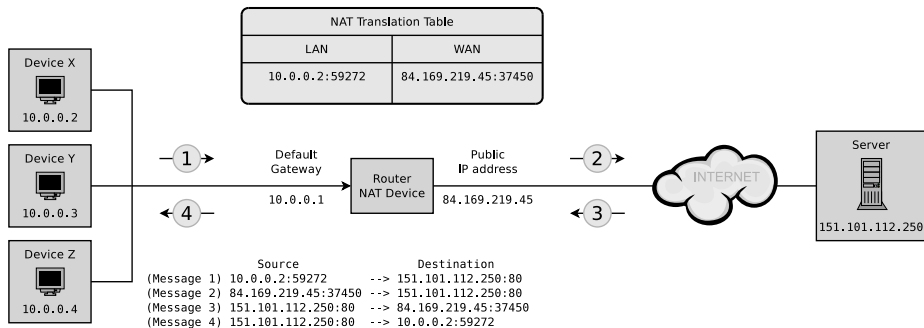
- Client X sends a request for a web page
 - The request (message 1) contains the IP address and port number of X as source addresses and the IP address and port number of the server as destination addresses
- The router replaces the IP and port number of the client with its own addresses inside the forwarded request (message 2)

Network Address Translation (4/5)



- The router stores the **mappings** between the router ports and the corresponding network devices inside its local **NAT table**
- The reply of the server (message 3) is targeted towards the IP of the router
 - The router replaces the address information according to the table and forwards the reply to X (message 4)

Network Address Translation (5/5)



- With IPv6, NAT is unnecessary because the address space is large enough to allocate globally accessible addresses to all network devices
 - However, NAT has advantages for network security because hosts, services, or the internal network structure are not exposed to the global Internet

Agenda

- Addressing
 - Purpose and Format
 - IPv4 Networks and Subnets
 - Private Networks and NAT
 - Fragmentation
 - IPv6 Networks

- Packet Structure
 - IPv4 Packet Structure
 - IPv6 Packet Structure

- ICMP

- Address Autoconfiguration

Packet Fragmentation (1/2)

- The split up (and reassembling) of IP packets into smaller packets (**fragments**) is called **Packet fragmentation**
- Either done by routers along the path or already at the sender
- Reason for packet fragmentation:
 - The maximum packet length depends on the network technology used
- The **Maximum Transmission Unit** (MTU) specifies the maximum payload of a frame (and thus the maximum size of an IP packet too)
 - MTU of Ethernet: usually 1,500 bytes
 - MTU of WLAN (IEEE 802.11): 2,312 bytes
 - MTU of PPPoE (e.g. DSL): $\leq 1,492$ bytes
 - MTU of ISDN: 576 bytes
 - MTU of FDDI: 4,352 bytes

Packet Fragmentation (2/2)

- IPv4 packets contain a flag which can be used to prohibit fragmentation
 - If a router needs to fragment a packet because it is too large to forward, but the fragmentation is prohibited in the packet, the router discards the packet because he cannot forward it
- If a network device does not receive all fragments of an IP packet within a certain period of time (a few seconds), the network device discards all received fragments
- Routers can split IP packets into smaller fragments, if the MTU makes this necessary and it is not prohibited in the packets
- But only the receiver can assemble fragments, none of the routers along the path

Agenda

- Addressing
 - Purpose and Format
 - IPv4 Networks and Subnets
 - Private Networks and NAT
 - Fragmentation
 - IPv6 Networks

- Packet Structure
 - IPv4 Packet Structure
 - IPv6 Packet Structure

- ICMP

- Address Autoconfiguration

A “new” Internet Protocol

Limitations of IPv4

- The IPv4 **packet format** has drawbacks
- **Newer hardware** obsoletes some of the design choices
- The **address space** is exhausted ³

A very short history of IPv6

- In 1992 the IETF working group *IPng* proposed seven ideas for a successor
- In 1995 IPv6 was specified as RFC 2460
- In 2011 all major OS provide a product-ready IPv6 implementation
- In 2018 only $\approx 25\%$ of all autonomous systems advertise IPv6 prefixes

³The IANA assigned the last free IPv4 address block to a Regional Internet Registry (RIR) in 2011.

IPv6 Improvements

■ Addressing

- $3.4 * 10^{38}$ addresses should suffice for the foreseeable future
- Simplifies address hierarchies
- More than one address per interface is common

■ Simplified administration

- Auto-configuration without additional protocols (like *DHCP* for IPv4)
- Renumbering of entire networks is much easier

■ Security

- The *IPsec* header extension enables authentication, integrity, and confidentiality

■ Simplified format

- Lean header with a fixed size plus optional next headers with a standardized format
- No checksum, no fragmentation

■ Improved Support for mobile applications

- Improved support for **multicast** and **anycast**
- Support for mobile devices

Representation of IPv6 Addresses

- Rules for simplification (RFC 5952):
 - **Leading zeros** within a block may be **omitted**
 - Successive blocks with value 0 (= 0000), may be omitted **exactly once within an IPv6 address**
 - If blocks are omitted, this is indicated by **two consecutive colons**
 - If several groups of null blocks exist, it is recommended to shorten the group with the most null blocks
- Example:
 - The IPv6 address of `j.root-servers.net` is:
2001:0503:0c27:0000:0000:0002:0030
⇒ 2001:503:c27::2:30

Notation of IPv6 addresses (URLs)

- IPv6 addresses are enclosed in square brackets
- Port numbers are appended outside the brackets
`http://[2001:500:1::803f:235]:8080/`
- This prevents the port number from being interpreted as part of the IPv6 address

Structure of IPv6 Addresses

- IPv6 addresses consist of two parts

64 Bits	64 Bits
Network Prefix	Interface Identifier
2001:638:208:ef34	:0:ff:fe00:65

1 Prefix (Network Prefix)

- Identifies the network

2 Interface identifier (Interface ID)

- Identifies a network device in a network
- Can be manually set, assigned via DHCPv6 or calculated from the MAC address of the network interface
- If the interface identifier is calculated from the MAC address, it is called **Extended Unique Identifier (EUI)**
 - When this is done, the MAC address (48 bits) is converted into a 64-bit address \implies **modified EUI-64 address format**

IPv6 Address Types

Described in RFC 4291.

■ Unicast

`fc00::/7` (1111 110) ⇒ Unique local address, may be routed only in private networks.⁴

`fe80::/10` (1111 1110 10) ⇒ Link local addresses, may not be routed.⁴

`::1/128` (0000..1) ⇒ Loopback address

`::/128` (0000..0) ⇒ Unspecified

Anything else ⇒ Global unicast address, e.g., `2000::/3` (2000... until `3fff...`)

■ Multicast

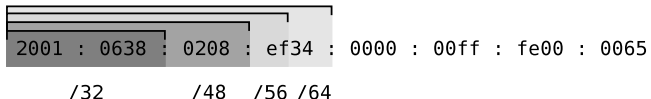
`ff00::/8` (1111 1111) ⇒ Multicast addresses. (No explicit broadcast addresses, but multicast groups for *all nodes* (`ff01::1` and `ff02::1`) and *all routers* (`ff01::2`, `ff02::2` and `ff05::2`)).

■ Anycast ⇒ from Unicast address range

⁴Only valid in the local network, not forwarded by routers in the Internet.

Structure of IPv6 Networks

- (Sub-)netmasks do not exist in IPv6
 - The subdivision of address ranges into subnets is done by specifying the **prefix length**
- IPv6 networks are specified in CIDR notation
 - The address of a single device sometimes has /128 attached
 - An example is the loopback address of IPv6: ::1/128
 - All bits – except the last one – have value 0
(For IPv4, the loopback address is: 127.0.0.1)
 - Internet Providers (ISPs) or operators of large networks get the first 32 or 48 bits assigned from a Regional Internet Registry (RIR)
 - The ISPs or network operators split this address space into subnets
 - **End users usually get a /64 or even a /56 network assigned**



Embed IPv4 Addresses into IPv6 (*IPv4 mapped*)

- A globally routed (unicast) IPv4 address can be represented as an IPv6 address and thus integrated into the IPv6 address space
 - In literature, this approach is called *IPv4 mapped*
- The IPv4 address gets a 96 bytes long prefix:
0:0:0:0:0:FFFF::/96

80 Bits					16 Bits	32 Bits
0000	0000	0000	0000	0000	FFFF	IPv4 address

- The IPv4 address may be represented in hexadecimal or decimal notation

- Example

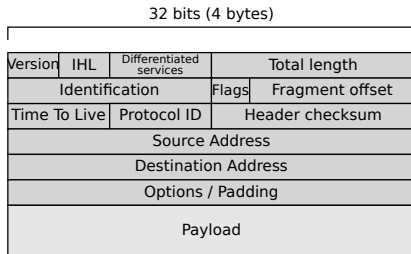
IPv4 address: 131.246.107.35
 IPv6 address: 0:0:0:0:0:FFFF:83F6:6B23
 Shorter notation: ::FFFF:83F6:6B23

Agenda

- Addressing
 - Purpose and Format
 - IPv4 Networks and Subnets
 - Private Networks and NAT
 - Fragmentation
 - IPv6 Networks
- Packet Structure
 - IPv4 Packet Structure
 - IPv6 Packet Structure
- ICMP
- Address Autoconfiguration

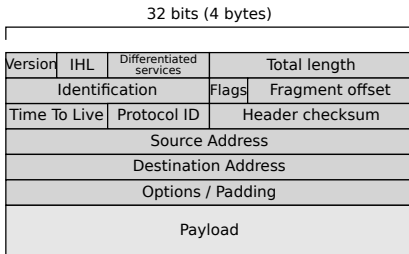
Structure of IPv4 Packets: Version, IHL, and DiffServ

- **Version** (4 bits)
 - Protocol version
 - Version = 4 \implies IPv4
 - Version = 6 \implies IPv6



- **IHL** = IP Header Length (4 bits)
 - Header length, represented as the number of 4 byte words
 - Example: IHL = 5 \implies 5 * 4 bytes = 20 bytes
 - Indicates where the payload begins
- **Differentiated services (DiffServ)** (8 bits)
 - Prioritization of IP packets is possible with this field (**Quality of Service (QoS)**)
 - The field slightly changed over the years (RFC 791, RFC 2474, RFC 3168)

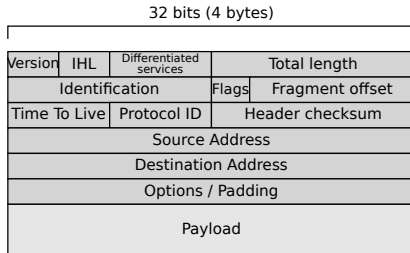
Structure of IPv4 Packets: Total Length



■ Total length (16 bits)

- This field defines the entire packet size (header and payload)
- This length of the field is 16 bits and therefore the maximum possible IPv4 packet length is 65,535 bytes

Structure of IPv4 Packets: TTL



■ Time To Live (8 bits)

- Specifies the maximum lifetime of an IP packet during transmission in seconds
- If the value is zero, the packet is discarded by the router
- Contains the maximum number of hops
 - Each router on the route to the destination decrements the value by one
- Prevents that undeliverable IP packets endlessly go in cycles on the network

The Role of ICMP

- The **Internet Control Message Protocol** (ICMP) is used for the exchange of...
 - **diagnostic**,
 - **control**, and
 - **error** messages
- ICMP is a component (*sub-protocol*) of IP
 - but it is treated as a separate protocol
- ICMPv4 is used for IPv4 networks, ICMPv6 is the corresponding protocol for IPv6 networks

Use Cases for ICMP

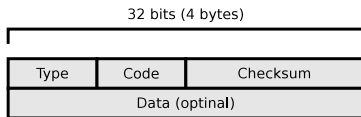
- All routers and terminal devices can handle ICMP
- Typical situations where ICMP is used:
 - A router **discards** an IP packet, because it does not know how to forward it
 - **Not all fragments** of an IP packet **arrives** at the destination
 - The destination of an IP packet cannot be reached, because the **Time To Live (TTL) has expired**
- ICMP specifies different sorts of messages, which can be send by a router as response to provide diagnostic information
- If an ICMP packet cannot be delivered, no further action is done

The most prominent example

The *ping* command uses ICMP messages.

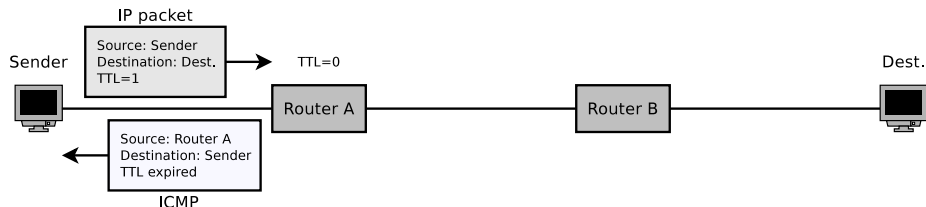
ICMP Message Structure

- The field **Type** in the ICMP header specifies its message type
- The field **Code** specifies the subtype of the message type
- The table contains some type-code combinations of ICMP messages



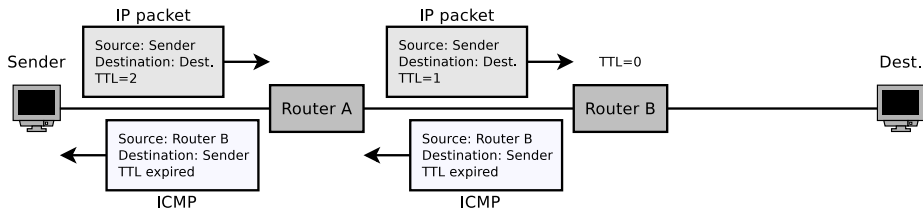
Type	Name of type	Code	Description
0	Echo reply	0	Echo reply (reply for ping)
3	Destination unreachable	0	Destination network unreachable
		1	Destination host unreachable
		2	Destination protocol unreachable
		3	Destination port unreachable
		4	Fragmentation required, but forbidden by the IP packet's flags
13		0	Firewall at destination site rejects the IP packet
		1	Firewall at destination site rejects the IP packet
4	Source Quench	0	Receive buffer is full, IP packet discarded (congestion control)
8	Echo Request	0	Echo request (ping)
11	Time Exceeded	0	TTL (Time To Live) expired
17	Address Mask Request	0	Request for the number of bits in the subnet mask
18	Address Mask Reply	0	Reply to message type 17
		1	Time limit exceeded during defragmentation
30	Traceroute	0	Determine the route to the destination
		1	Determine the route to the destination

Example of using ICMP: traceroute (1/3)



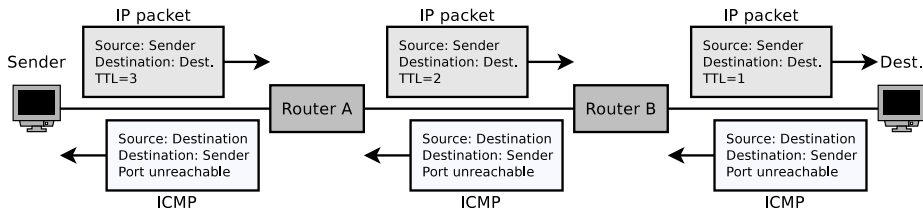
- Another application example of ICMP is the tool traceroute
- traceroute determines, which routers are used to forward packets to the destination site
- The sender transmits an IP packet to the destination with TTL=1
- Router A receives the IP packet, sets TTL=0, discards the IP packet and transmits an ICMP message of message type 11 and code 0 to the sender

Example of using ICMP: traceroute (2/3)



- Next, the sender transmits an IP packet to the destination with TTL=2
- The IP packet is forwarded by router A and thereby the value of TTL is decremented
- Router B receives the IP packet, sets TTL=0, discards the IP packet and transmits an ICMP message of message type 11 and code 0 to the sender

Example of using ICMP: traceroute (3/3)



- Once the value of TTL is big enough that the destination site can be reached, the receiver transmits an ICMP message of message type 3 and code 3 to the sender
- This way, the path from sender to receiver can be traced via ICMP

```
$ traceroute -q 1 wikipedia.de
traceroute to wikipedia.de (134.119.24.29), 30 hops max, 60 byte packets
 1 fritz.box (10.0.0.1)  1.834 ms
 2 p3e9bf6a1.dip0.t-ipconnect.de (62.155.246.161)  8.975 ms
 3 217.5.109.50 (217.5.109.50)  9.804 ms
 4 ae0.cr-polaris.fra1.bb.godaddy.com (80.157.204.146)  9.095 ms
 5 ae0.fra10-cr-antares.bb.gdinf.net (87.230.115.1)  11.711 ms
 6 ae2.cgn1-cr-nashira.bb.gdinf.net (87.230.114.4)  13.878 ms
 7 ae0.100.sr-jake.cgn1.dcnnet-emea.godaddy.com (87.230.114.222)  13.551 ms
 8 wikipedia.de (134.119.24.29)  15.150 ms
```

Agenda

- Addressing
 - Purpose and Format
 - IPv4 Networks and Subnets
 - Private Networks and NAT
 - Fragmentation
 - IPv6 Networks

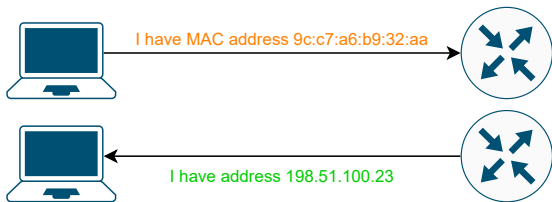
- Packet Structure
 - IPv4 Packet Structure
 - IPv6 Packet Structure

- ICMP

- Address Autoconfiguration

Reverse Address Resolution Protocol (RARP)

- Upon **booting** a network interface has no IP address assigned
- **Manual address** configuration is not desirable in many scenarios
- With the help of **Reverse ARP**, well-known hardware addresses are assigned to IP addresses, and recorded on a RARP server
- **Problem:** RARP requests are not passed on by routers, therefore a RARP server must be set up in each local network



RARP is obsolete. Replaced by DHCP (more modern and feature-rich).

