Computer Networks

Internet Protocol

Chapter

1. Introduction
2. Protocols
3. Application layer
4. Web services
5. Distributed hash tables
6. Time synchronization
7. Error control
8. Transport layer
9. Network layer
10. **Internet protocol**
    - **IPv4**
    - **ICMP**
    - **NAT**
    - **IPv6**
11. Data link layer
12. WLAN

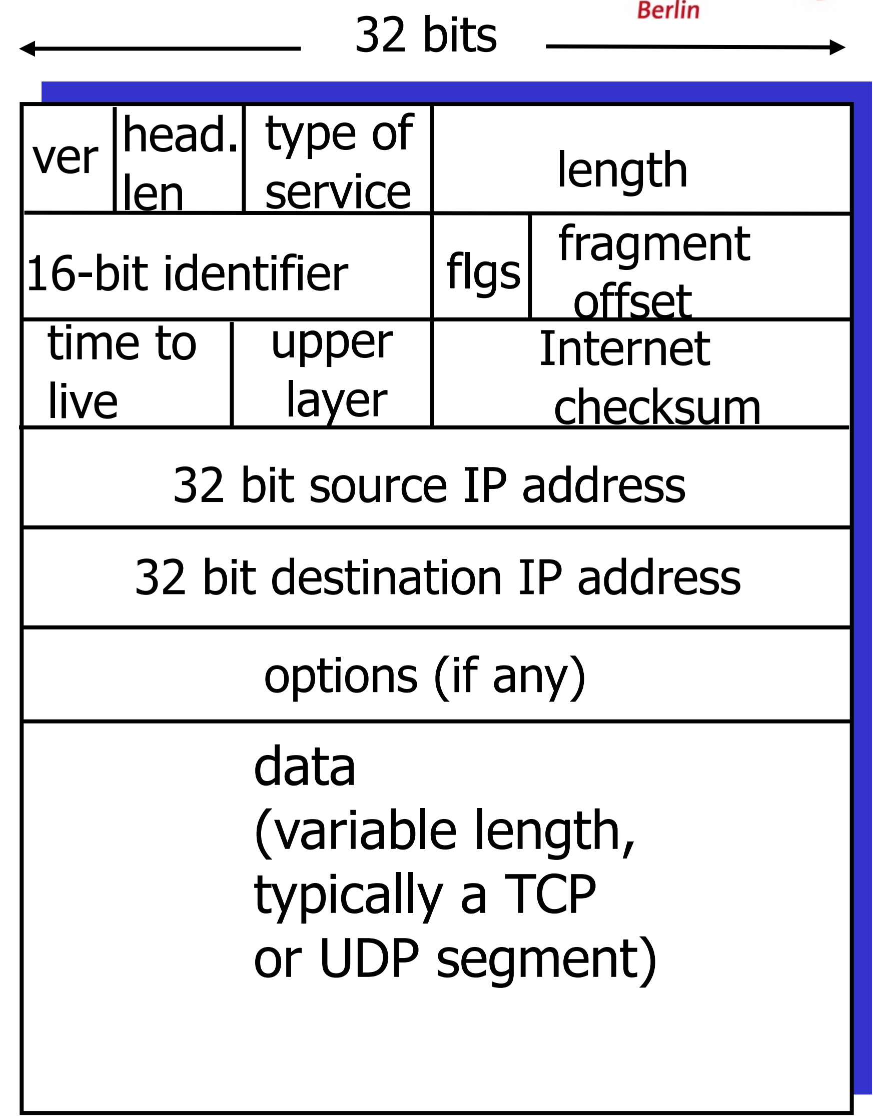
Top-Down-Approach

- Application Layer
- Presentation Layer
- Session Layer
- Transport Layer
- Network Layer
- Data link Layer
- Physical Layer

Internet Protocol (IPv4)

Packet Format

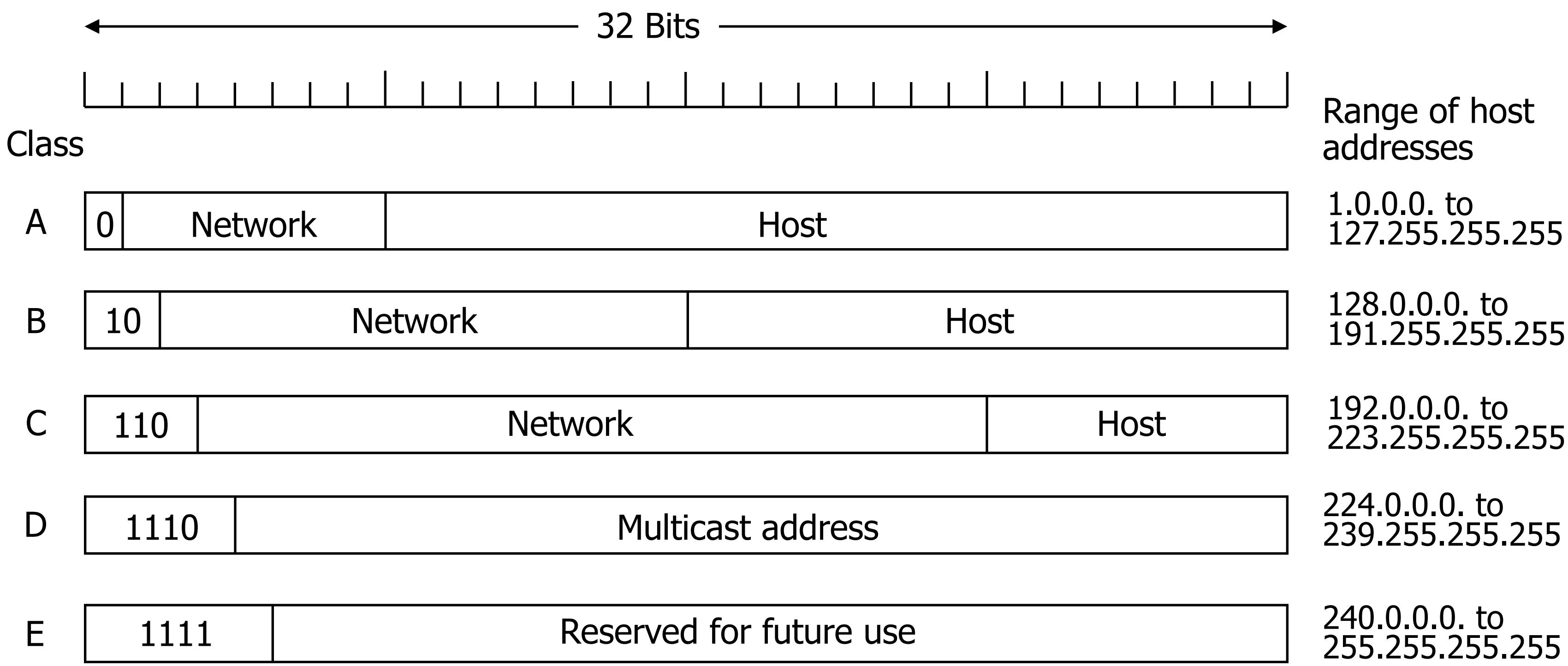
- ver: 4 for IPv4
- header length: 4 bit, length in 32-bit-words
- type of service: 2 bit for quality of service
- length: total length in byte (max. 2^{16} = 65.535 byte)
- identifier, flgs, fragment offset: used for fragmentation
- time to live: max. number of hops, every router on path decrements
- upper layer: higher layer protocol
- options: e.g., time stamps, path through network
- total length without options 20 byte



IP Address

- Identifies the **interface** of a host or routter
  - Hosts with multiple interfaces (multi-homed) and routers need multiple IP addresses
- 32 bit / 4 byte, separated in **network** and **host** identifier
- Centralized address coordination
  - Internet Corporation for Assigned Names and Numbers (ICANN)
- Authorized address registries release IP networks to ISPs, these to their customers
  - American Registry for Internet Numbers (ARIN),
    Reseaux IP Europeens (RIPE), …
- **Dotted decimal** representation
  - $d_1.d_2.d_3.d_4$ with d_j = decimal representation of j-th byte
  - Ex: $10000000\ 10000111\ 01000100\ 00000101_2$ is written as 128.135.68.5

Class-based Addressing



Special IP Addresses

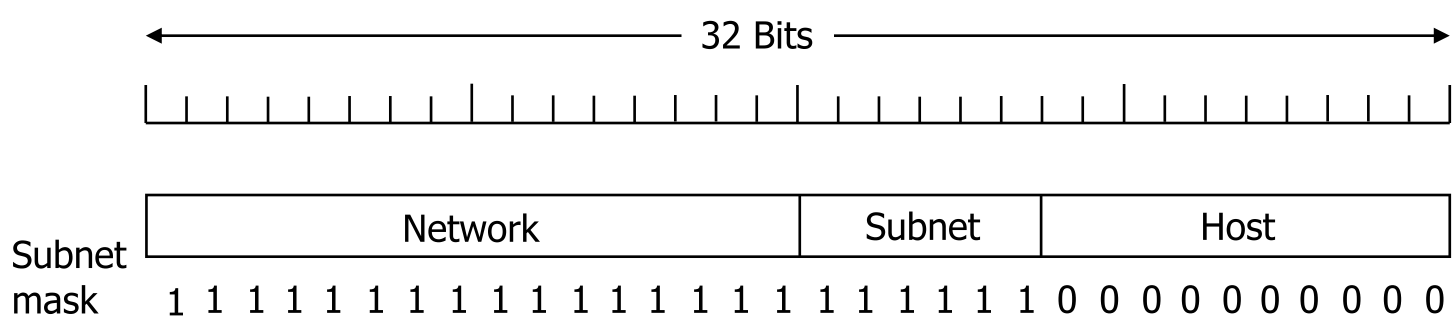
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	This host
0 0 · · · 0 0 | Host	A host on this network
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Broadcast on the local network
Network | 1 1 1 1 · · · 1 1 1 1	Broadcast on a distant network
127 | (Anything)	Loopback

Class-based Addressing

- Advantages
  - **Self-identifying address**: the first bits decide about the class
  - Forwarding tables are rather short
- Disadvantages
  - Inflexible and usually overly-large address space (what if too few / too many nodes are in a network)
  - Since the early 1990ies, it was clear that the address space is insufficient for further growth of the internet

Classless Addressing

- Subnetworks
  - IP address is split into network and host part
  - Network mask: leading 1-bits AND IP address results in network part
  - Notation: IP address/length of mask, e.g., 150.100.12.176/22

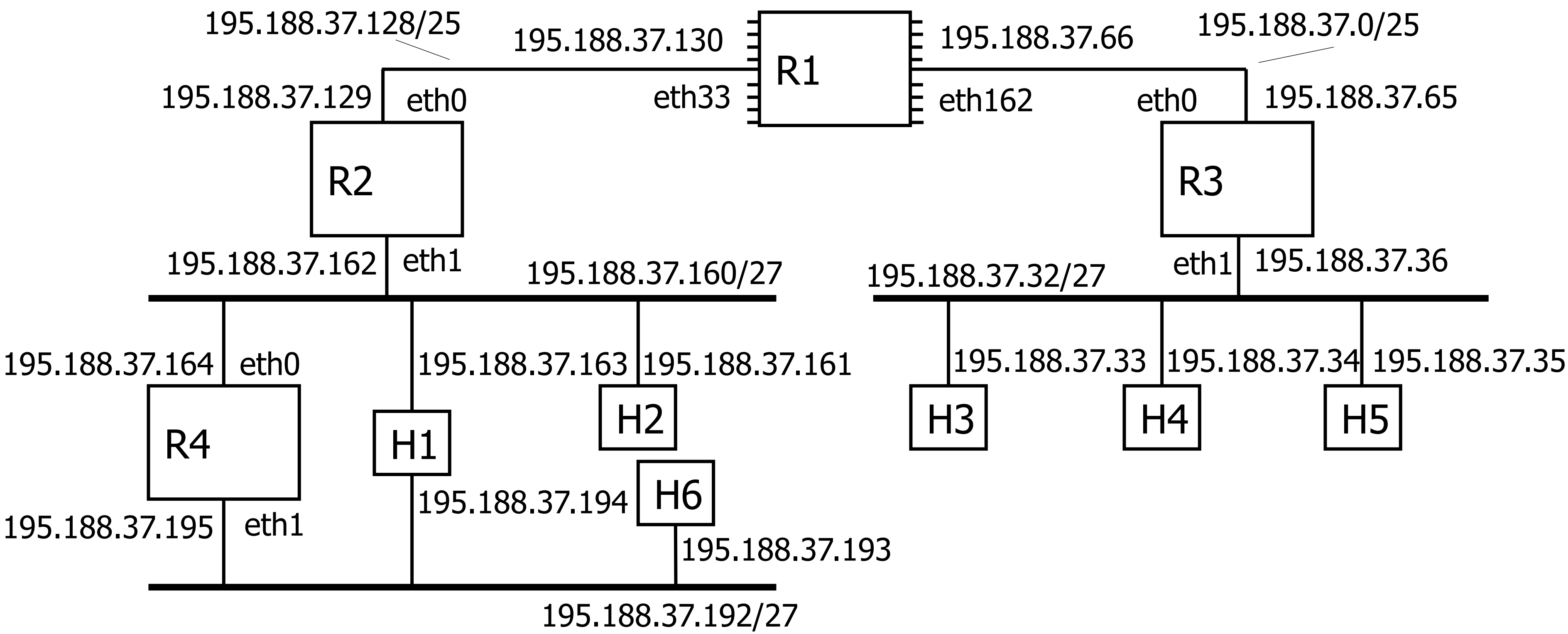


Classless Addressing: Example

- Organization has class-B-network (i.e., 16 bit for network part): 150.100.0.0/16
- Subnetworks are required for about 100 hosts each
  - 7 bit are sufficient (= 126 host addresses, **why 126?**)

- Example: IP address 150.100.12.176/25
  - Binary = `10010110 01100100 00001100 10110000`
  - network mask = `11111111 11111111 11111111 10000000`
  - AND = `10010110 01100100 00001100 10000000`
  - Subnetwork address = 150.100.12.128/25

Classless Addressing: Example Network



Classless Inter-Domain Routing (CIDR)

- Fixed-size routing tables in class-based routing is replaced by subnetwork-based tables
  - Forwarding now based on variable length of network part
  - Address and mask are used in routing tables
  - Results in much larger routing tables

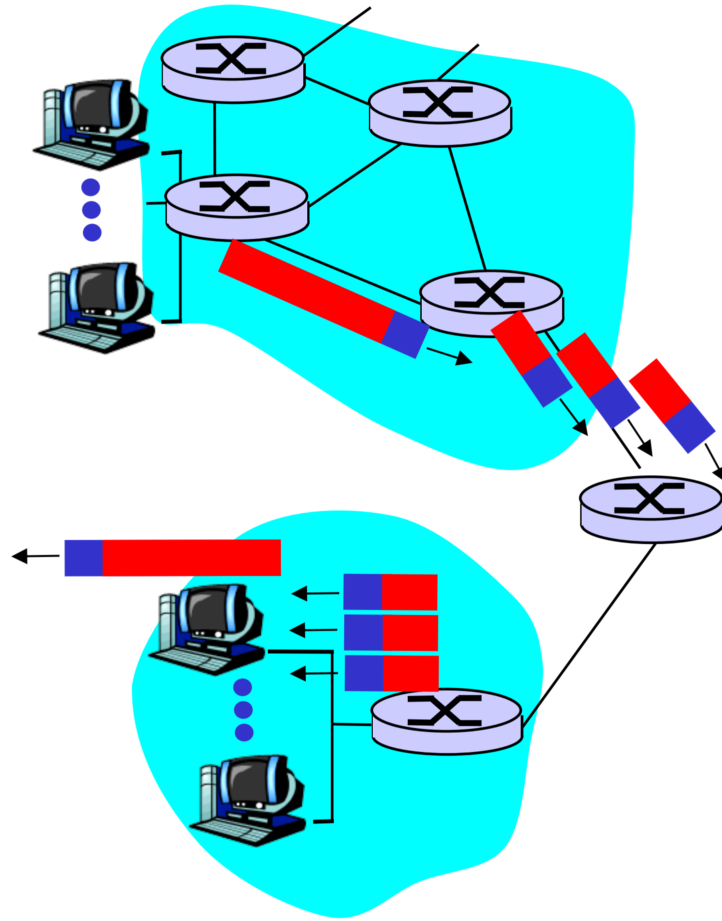
- **Subnetting** or **Supernetting**: Combining networks into larger subnetwork
  - This is today's standard
  - Routers use **Longest-Prefix-Match** to select the outgoing interface
  - This is quite expensive, thus, special data structures (variants of binary trees) are used for improved efficiency

Fragmantation

Fragmentation

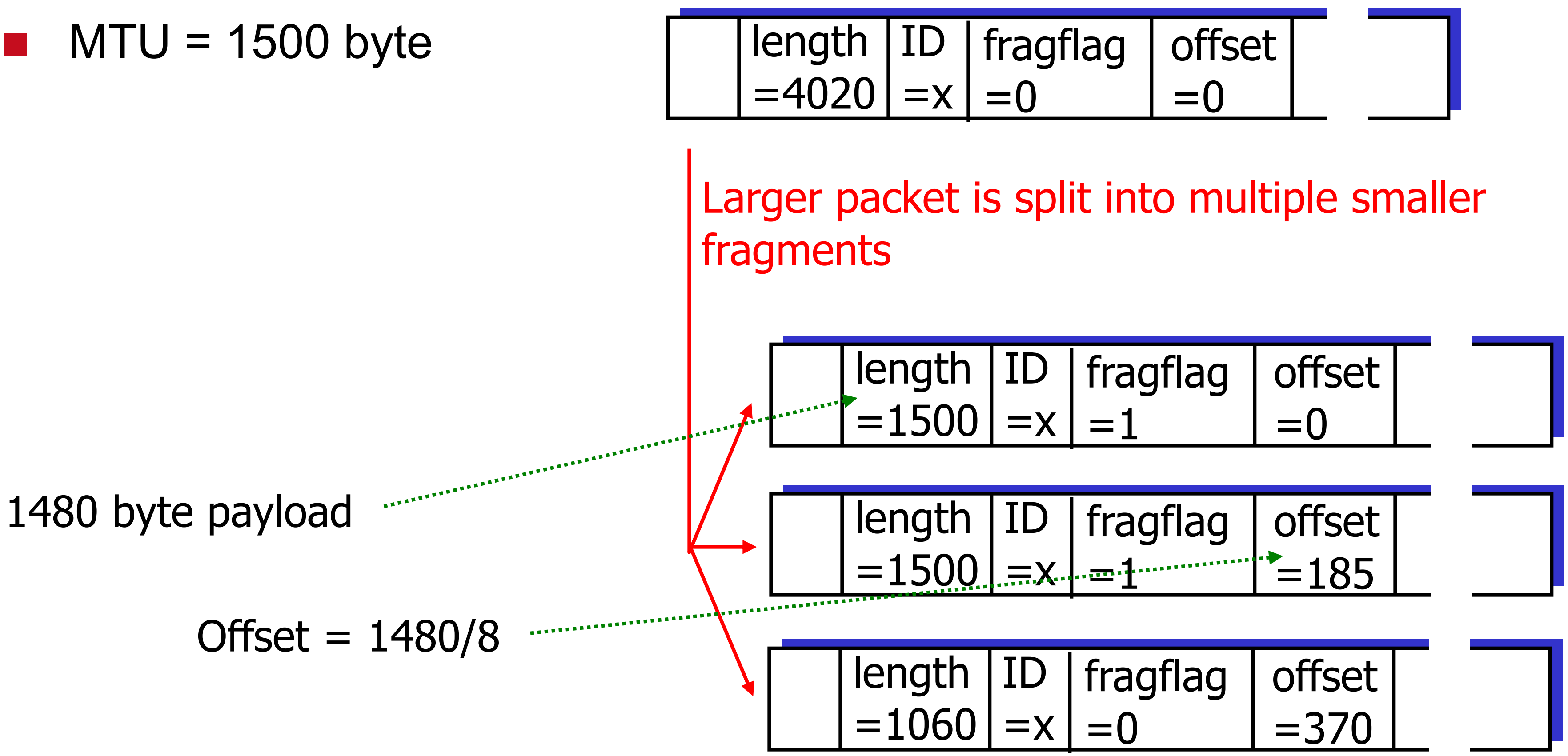
- Fragmentation
  - Every connection has a maximum transmission unit (MTU) defined by lower layer protocols
  - If a connection has a smaller MTU than the size of the IP packet, IP packet is split into smaller fragments

    (this can repeat multiple times)
  - IP header support
    - identifier: ID of fragments belonging to the same packet
    - flag: another fragment follows
    - offset: position within the payload data (using offset*8)
- Reassembly
  - Only at the final destination (**why?**)



Fragmentation: Example

- 4000 byte packet
- MTU = 1500 byte



Dynamic Host Configuration Protocol

Dynamic Host Configuration Protocol (DHCP)

- If I get a new laptop, it has
  - Ethernet interface with its specific MAC address
  - WiFi interface with its specific MAC address
  - Probably more…

- If I connect to a new network, I need an IP address per interface
  - Campus: assigned by system administrator?
  - At home: I am responsible myself?
  - Coffee shop: it's getting tedious…

- Obviously, the IP address is depending on the internet service provider (ISP)

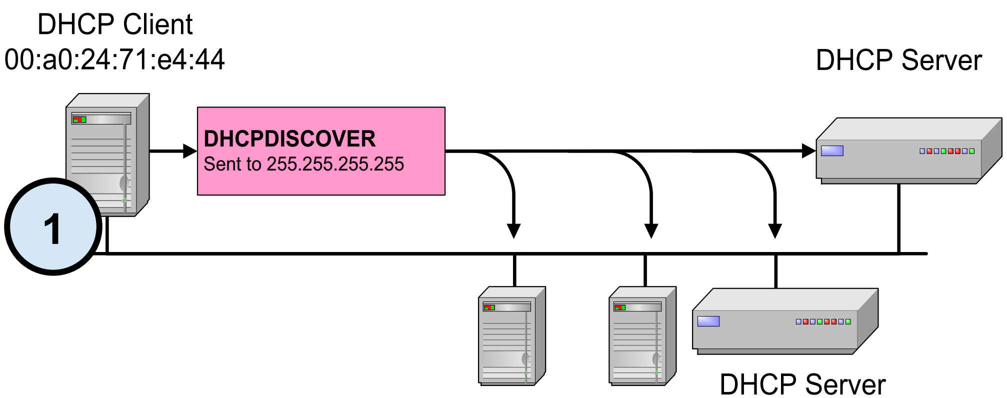
Dynamic IP Address Assignment

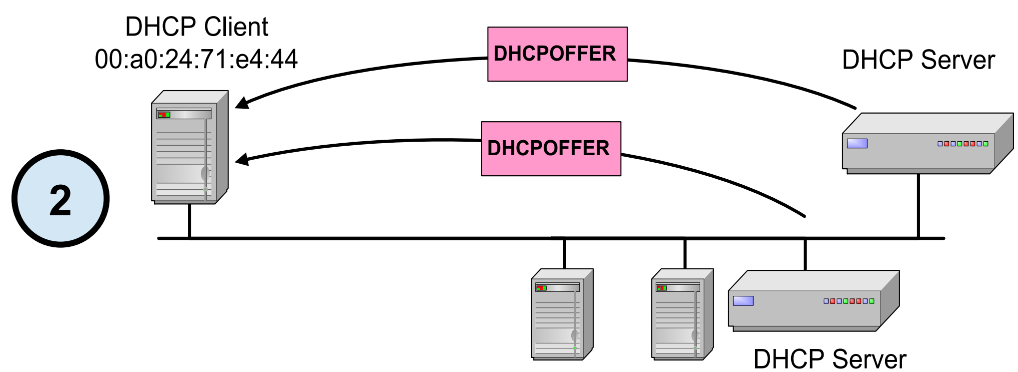
- Dynamic IP address assignment is helpful
  - Manual configuration is tedious, error prone, …
  - Mobility makes things worse

- Protocols for IP address assignment
  - RARP (until 1985, no longer used)
  - BOOTP (1985-1993)
  - **Dynamic Host Configuration Protocol (DHCP)** (since 1993)

- Today: only **DHCP** is relevant

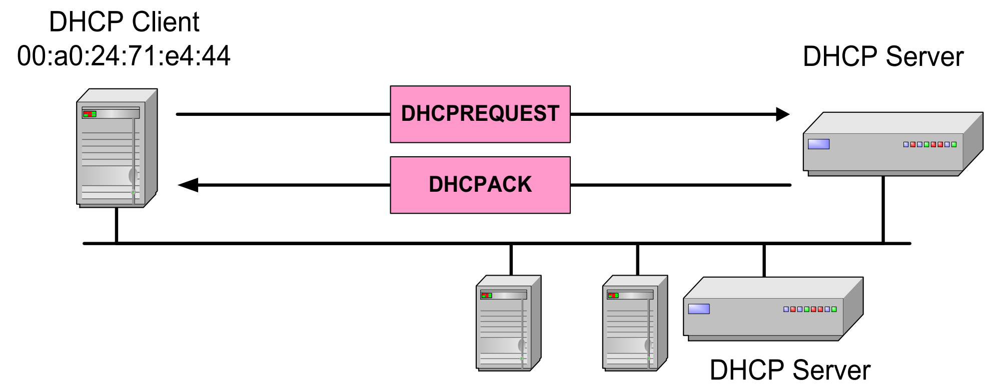
DHCP Protocol

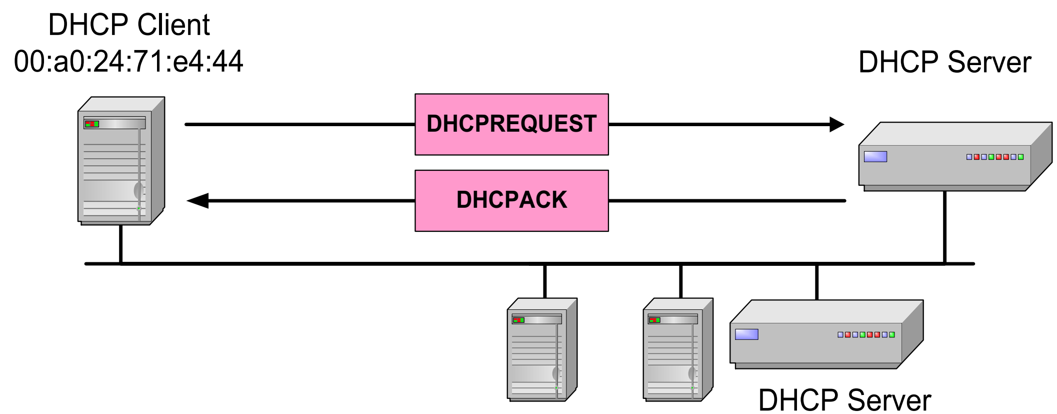




DHCP Protocol

- Based on DHCPOFFER, client choses an IP address; this one is requested; can be used after DHCPACK
- "Lease" needs to be renewed after 50% of expiry time
- DHCP server can explicitly release an IP address using DHCPNACK
- **Soft state concept!**



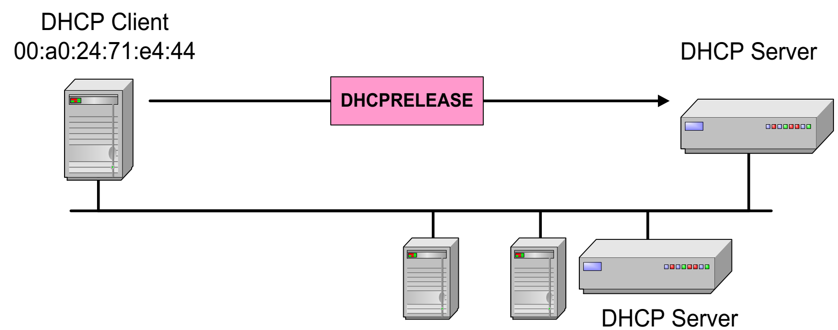


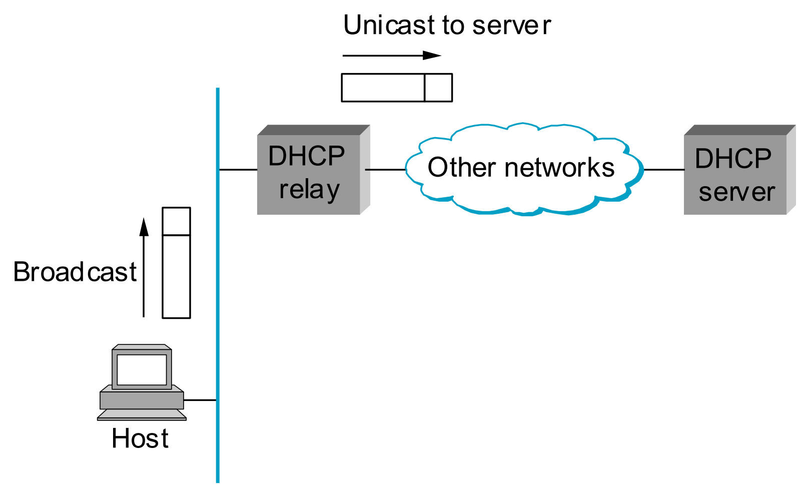
Telecommunication Networks Group

DHCP Protocol

- Client should explicitly release address when disconnecting → in reality implicit if address is not renewed (**why?**)

- **Note**: DHCP relay can be used if DHCP server is not connected to local network (**useful?**)





DHCP Requirements

- Every address must be assigned to at most one host
- Address assignment must survive restart of server and/or client
  - Server needs to persist current assignments
- Server should support fixed IP to MAC assignment (you always get the same IP address)

- Most modern DHCP servers not only assign IP addresses but also
  - Default router
  - DNS server
  - Proxy server
  - …

Internet Control Message Protocol

Internet Control Message Protocol (ICMP)

- Control messages between hosts and routers
  - E.g., address unreachable, max. number of hops reached, echo request/echo reply
- Format:
  - Type
  - Code
  - Checksum
  - Data
- ICMP messages are transported as payload of IP packets

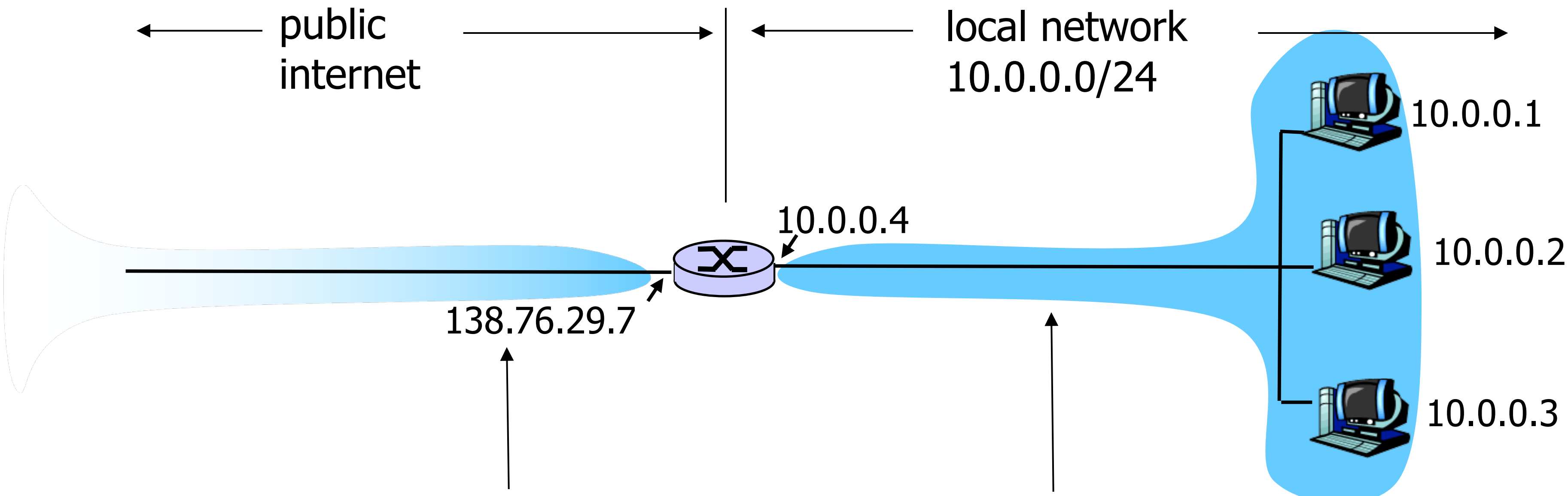
Type	Code	Description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest. host unreachable
3	2	dest. protocol unreachable
3	3	dest. port unreachable
3	6	dest. network unknown
3	7	dest. host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

Network Address Translation

Network Address Translation (NAT)

- IP(v4) address space is very limited (and already exhausted)
- Idea: use internally other (and more) addresses than externally known
  - Use of private IP addresses (not routed in the internet)
  - 10.0.0.0/8, 172.16.0.0/12, 192.168.0.0/16
- Externally, only a few (in most cases exactly one) global IP addresses are used
- External connections are now represented using a trick:
  - Instead of using IP addresses only (as we learned is correct for protocol stacks), in addition transport layer port numbers are used
  - Translation table at NAT gateway to overwrite IP addresses and UDP/TCP ports for every incoming/outgoing packet
  - Number of connections limited by number of ports
- Advantage: requires changes only at NAT gateway (usually the router to the internet)
- Disadvantage: strict layering is violated, end-to-end-connections are manipulated

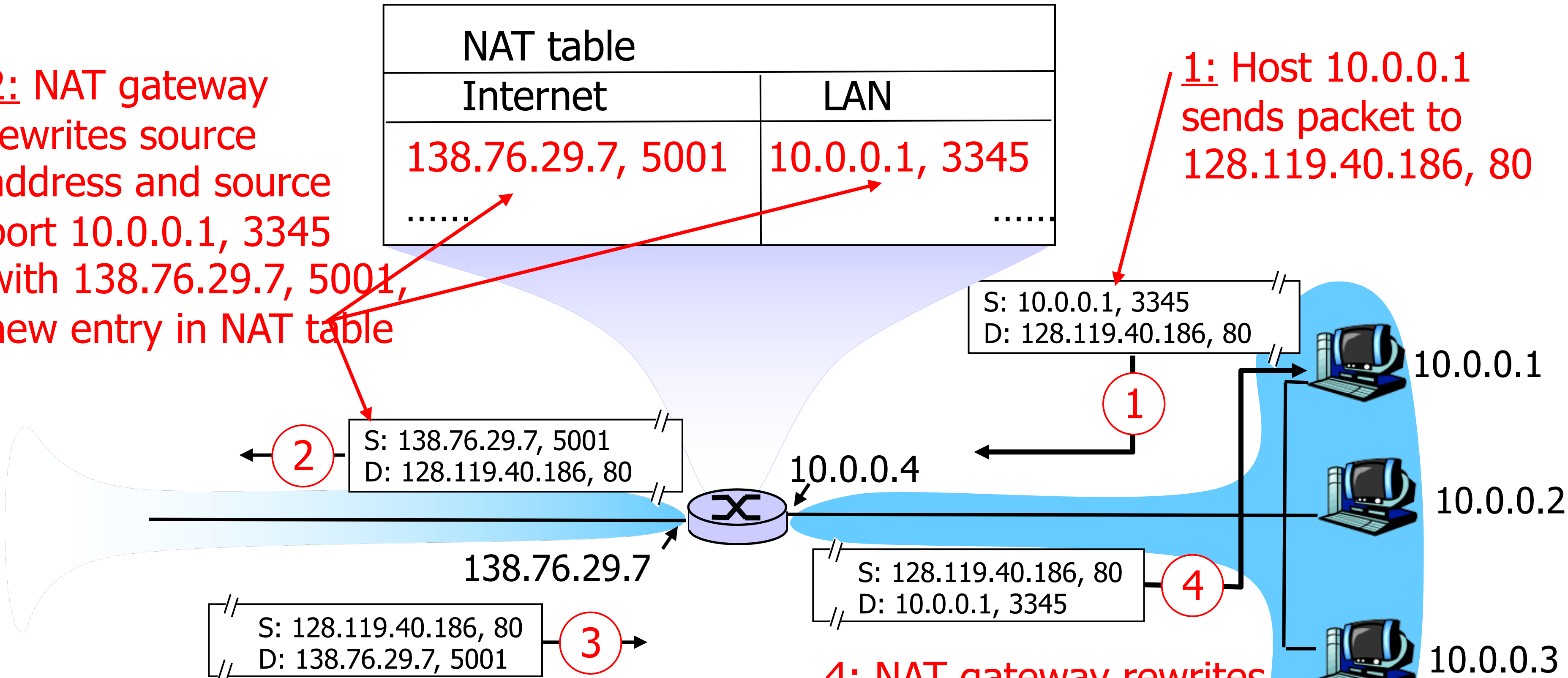
NAT: Example



assigned globally valid IP address, translation to local addresses using port numbers

10.0.0.0/24 subnetwork, assigned addresses not routed in the internet

NAT: Example



1: Host 10.0.0.1 sends packet to 128.119.40.186, 80

2: NAT gateway rewrites source address and source port 10.0.0.1, 3345 with 138.76.29.7, 5001, new entry in NAT table

3: response with destination address and port: 138.76.29.7, 5001

4: NAT gateway rewrites destination address and port 138.76.29.7, 5001 with 10.0.0.1, 3345

IPv6

IPv6

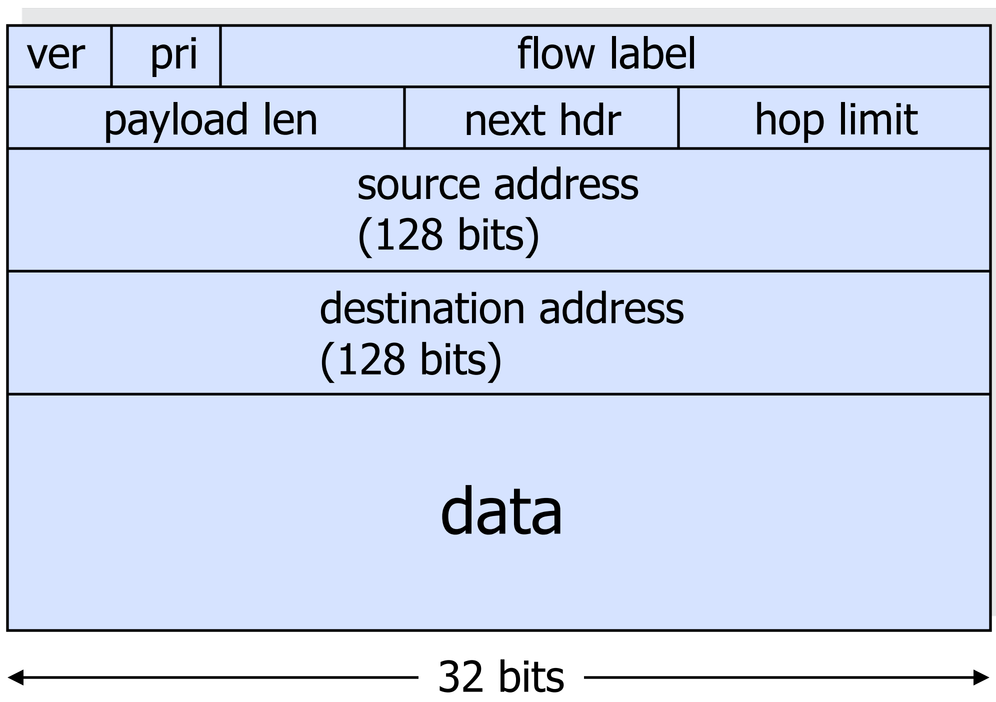
- IPv6 (IP version 6)
  - Initially called IP Next Generation (IPng)
  - IETF standardization was mainly triggered by limited IP address space, but IPv6 was designed to solve many other issues
    - Header with fixed length (**why?**)
    - No fragmentation needed (initial MTU path discovery)
    - No checksums (error control at higher layers)
    - Additional options in form of chained headers
    - Auto configuration (meanwhile all DHCP)
    - Quality of service (realized by IntServ and DiffServ)
  - However, still problems making IPv6 the only internet protocol (IPv4 no “can” do many things itself)

IPv6

- Address categories
  - Unicast: to one specific destination
  - Multicast: to all nodes (in a network)
  - Anycast: to a node out of a group
- Addresses
  - 128 bit, grouped into blocks of 16 bit, written as 8 hexadecimal numbers connected by a colon
  - Example: 4BF5:AA12:0216:FEBC:BA5F:039A:BE9A:2176
  - For easier use, short representation possible:
    - 4BF5:0000:0000:0000:BA5F:039A:000A:2176
    - Compact nulls: 4BF5:0:0:0:BA5F:39A:A:2176
    - Left out nulls: 4BF5::BA5F:39A:A:2176
  - Mixed notation for IPv4-IPv6 translation:
    - Last 32 bit are used for IPv4 address, e.g., ::FFFF:128.155.12.198

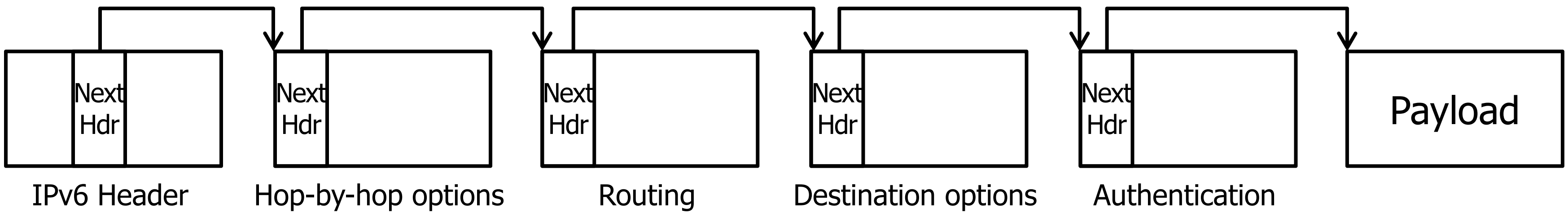
IPv6 Packet Format

- Address space of 128 bit should be sufficient “for a while”



IPv6 Header Concept

- For optional tasks, additional ”extension header” of fixed size
- Very fast processing, every system only looks at known / required headers



- Examples
  - TCP (6)
  - UDP (17)
  - ESP (50) / AH (51)
  - ICMPv6 (58)
  - SCTP (132)