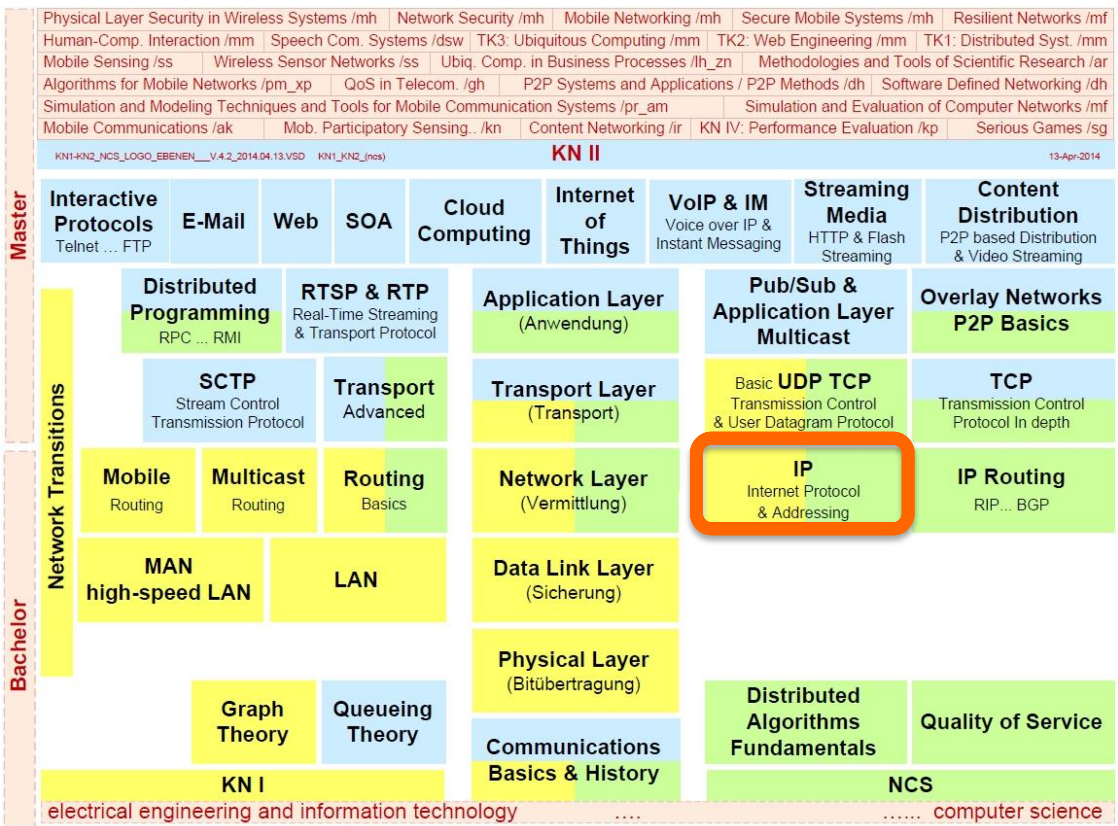
Communication Networks I

L3 - Internet Protocol:
IP Addressing – Interior and Exterior Gateway Protocols

TECHNISCHE UNIVERSITÄT DARMSTADT



Prof. Dr.-Ing. **Ralf Steinmetz**
KOM - Multimedia Communications Lab

Overview

1 History and Architecture

ARPANET

- Initiated and financed by ARPA
  - Advanced Research Projects Agency of the U.S. Department of Defense (DoD)

- Objective
  - Originally: network to survive nuclear war
  - Later: network to connect scientific and military institutions

- 1969
  - Experimental network with 4 nodes, followed by rapid growth, BBN first contractor

- Development of the INTERNET
  - Standardized protocols for comm. between networks: TCP/IP (1983)
  - Linking military networks (MILNET, MINET)
  - Linking satellite networks (SATNET, WIDEBAND)
  - Linking the LANs of the universities

- Fast spreading of TCP/IP technology as a part of UNIX
  → ARPANET growing rapidly
  - 1987: 15% per month
  - 1987: 20.000 computers, more than 100.000 users

1990: ARPANET replaced, MILNET still exists

- Services: E-mail, file transfer, remote login, later WWW. . .

The Internet and its Tasks

Internet (Internet Society)

- Mid-80s
  - a multiple of networks was designated as the "Internet"
- Jan. 1992:
  - founding of the (actual) Internet Society
  - objective: to spread the use of the Internet (protocols and services)

IAB: Internet Architecture Board

- Founded in 1983 to involve researchers in the ARPANET
- Today it is the supreme Internet board

IETF (INTERNET ENGINEERING TASKFORCE)

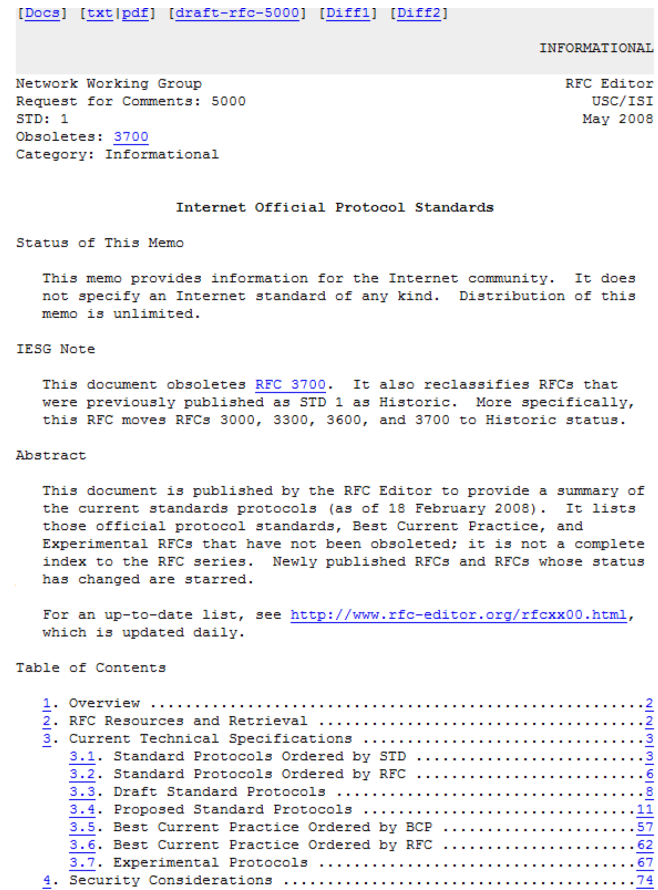
- Overseen by IAB oversees/nominates
- Divided into approx. 70 working groups (e. g. RSVP, ST-II)
- Actual governing board

IRTF (Internet Research Taskforce)

- Overseen by IAB oversees/nominates

RFC (REQUEST FOR COMMENTS) /www.ietf.org/rfc.html

- Recommendations, e.g. in May 2008 rfc 5000 appeared

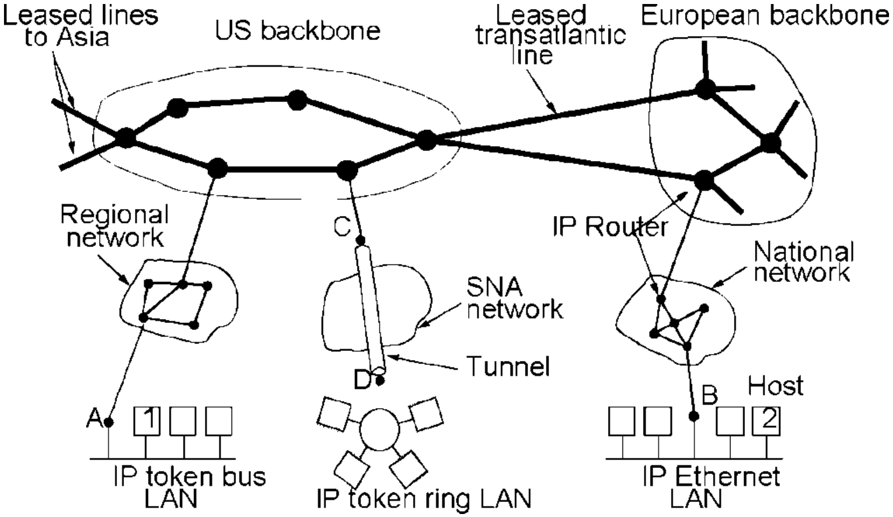


```
[Docs] [txt|pdf] [draft-rfc-5000] [Diff1] [Diff2]

                                                        INFORMATIONAL

Network Working Group                                      RFC Editor
Request for Comments: 5000                                    USC/ISI
STD: 1                                                       May 2008
Obsoletes: 3700
Category: Informational

                Internet Official Protocol Standards

Status of This Memo

   This memo provides information for the Internet community.  It does
   not specify an Internet standard of any kind.  Distribution of this
   memo is unlimited.

IESG Note

   This document obsoletes RFC 3700.  It also reclassifies RFCs that
   were previously published as STD 1 as Historic.  More specifically,
   this RFC moves RFCs 3000, 3300, 3600, and 3700 to Historic status.

Abstract

   This document is published by the RFC Editor to provide a summary of
   the current standards protocols (as of 18 February 2008).  It lists
   those official protocol standards, Best Current Practice, and
   Experimental RFCs that have not been obsoleted; it is not a complete
   index to the RFC series.  Newly published RFCs and RFCs whose status
   has changed are starred.

   For an up-to-date list, see http://www.rfc-editor.org/rfcxx00.html,
   which is updated daily.

Table of Contents

   1. Overview ......................................................2
   2. RFC Resources and Retrieval ...................................2
   3. Current Technical Specifications ..............................3
      3.1. Standard Protocols Ordered by STD ........................3
      3.2. Standard Protocols Ordered by RFC ........................6
      3.3. Draft Standard Protocols .................................8
      3.4. Proposed Standard Protocols .............................11
      3.5. Best Current Practice Ordered by BCP ....................57
      3.6. Best Current Practice Ordered by RFC ....................62
      3.7. Experimental Protocols ..................................67
   4. Security Considerations ......................................74
```

Tasks in the INTERNET

- To connect different networks over gateways
- Definition of
  - protocols that work on all subnetworks
  - standardized addressing pattern for a very large network
  - global routing architecture

Subnets in the INTERNET



Ethernet LANs
- Mainly large campus networks

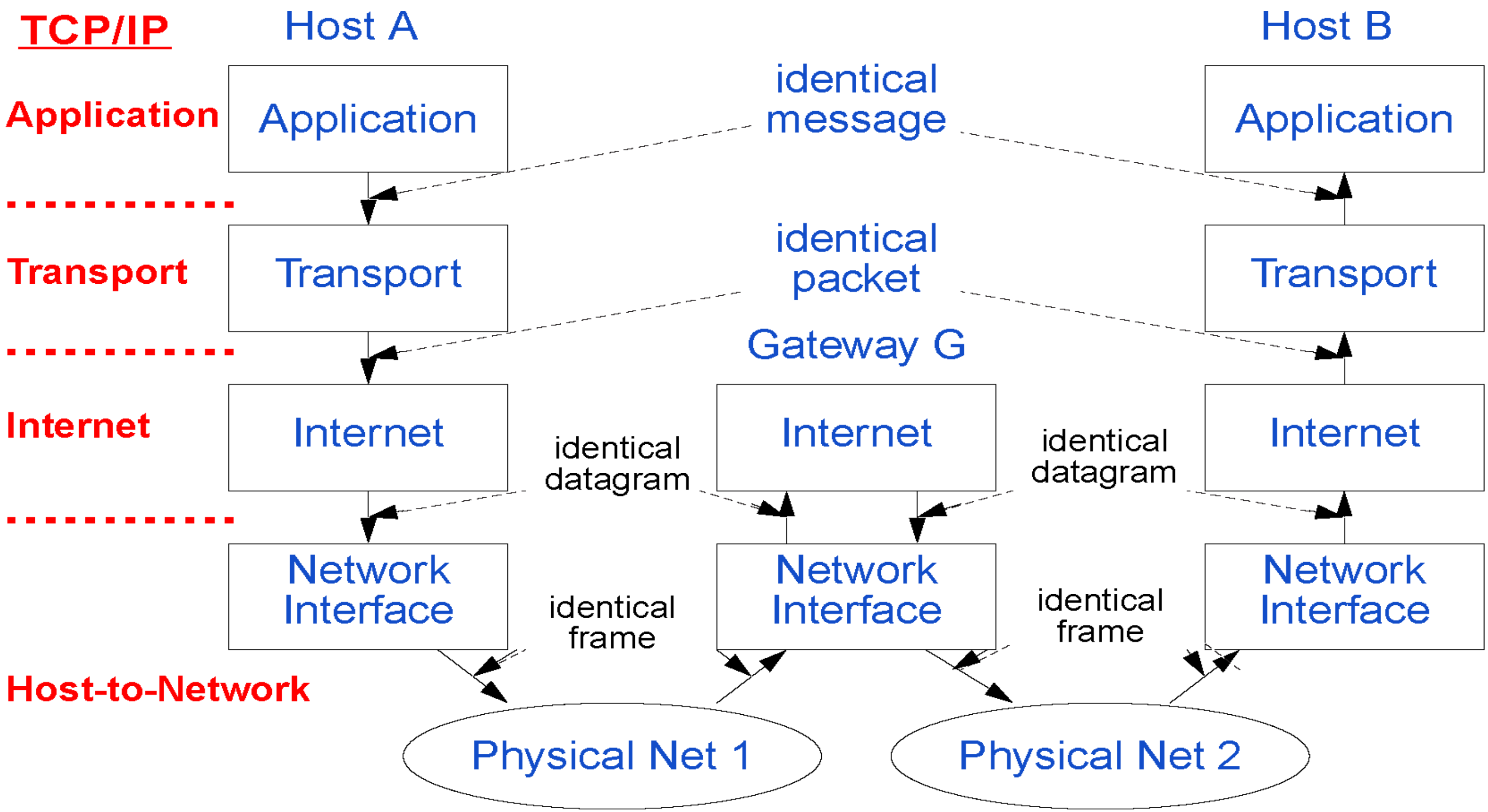
Other LANs
- Mainly smaller/experimental networks

Arpanet
- Network with specific protocols, partially connected over leased lines

NSF Net (National Science Foundation Network)
- Backbone consisting of leased high-speed lines
- Connecting the NSF supercomputers with each other and to regional networks and campus networks
- Later 1995 AOL, now a multitude of backbones in USA

CSNET (X.25 NET)
- Public packet relay network by X.25

Internet Architecture



i.e.

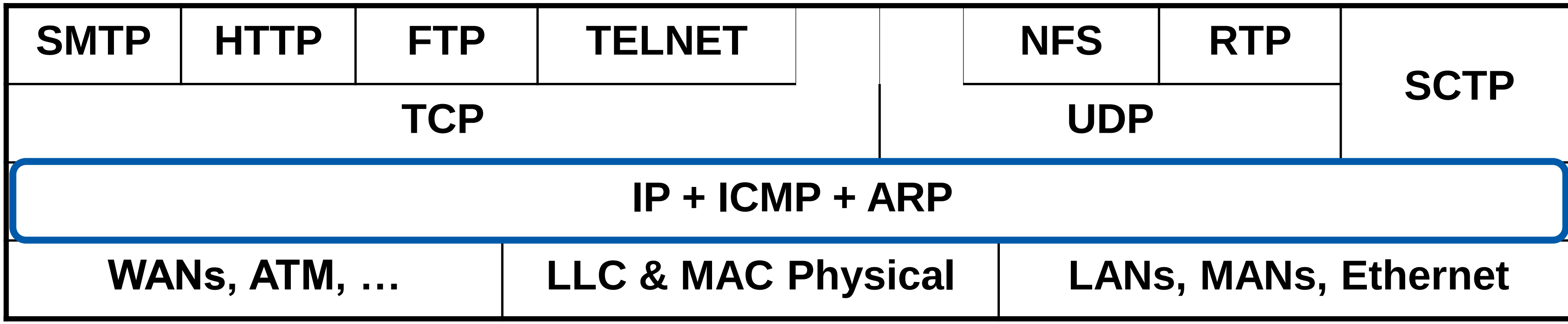
- ISO-OSI presentation and session layer not explicitly available
- Data link layer and physical layer combined

Well-Known Internet Protocols

<table>
<tr><td>SMTP</td><td>HTTP</td><td>FTP</td><td>TELNET</td><td></td><td></td><td>NFS</td><td>RTP</td><td rowspan="2">SCTP</td></tr>
<tr><td colspan="5">TCP</td><td colspan="3">UDP</td></tr>
<tr><td colspan="9">IP + ICMP + ARP</td></tr>
<tr><td colspan="3">WANs, ATM, …</td><td colspan="3">LLC & MAC Physical</td><td colspan="3">LANs, MANs, Ethernet</td></tr>
</table>

ARP = ADDRESS RESOLUTION PROTOCOL
FTP = File Transfer Protocol
HTTP = Hypertext Transfer Protocol
IP = INTERNET PROTOCOL
ICMP = INTERNET CONTROL MESSAGE PROTOCOL
LLC = Logical Link Control
MAC = Media Access Control
NFS = Network File System
SMTP = Simple Mail Transfer Protocol
TELNET = Remote Login Protocol
TCP = Transmission Control Protocol
UDP = User Datagram Protocol
RTP = Real-Time Transport Protocol

Well-Known Internet Protocols



INTERNET PROTOCOL IP basics

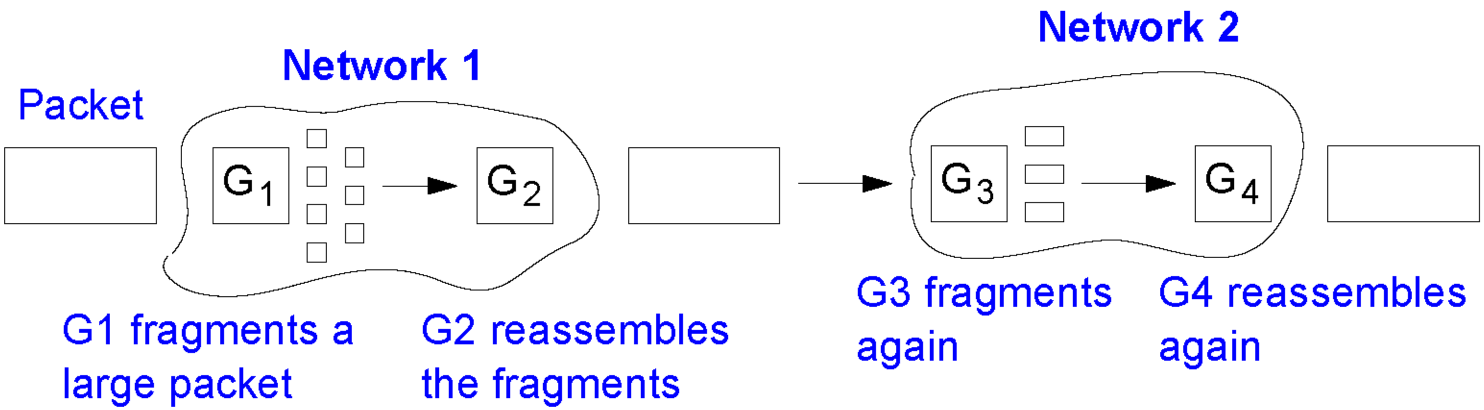
- Defined for the first time in 1981
  - J. Postel
  - RFC 791, September 1981
- Packet length
  - in theory: up to 64 kBytes
  - in real life: approx. 1500 Bytes

Connectionless service (datagram)

- Provide best-efforts (not guaranteed) way to transport datagrams
  - from source to destination
  - without regard whether
    - these machines are on the same network
    - there are other networks in between

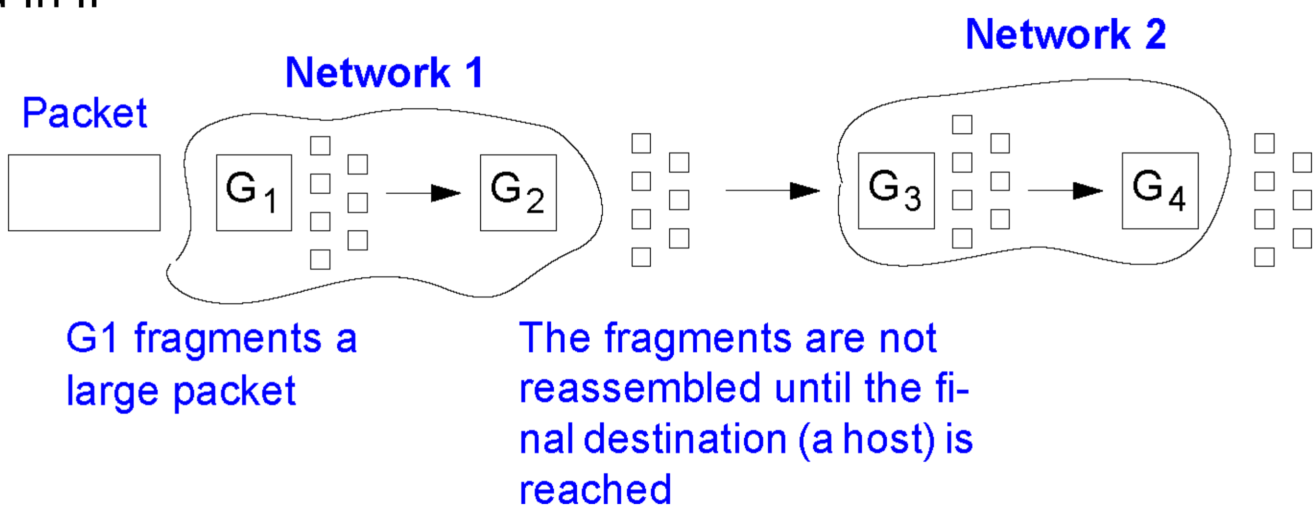
2.1 IP Segmentation/Reassembling

Transparent segmentation



Non-transparent segmentation

- Used In IP



2.2 IP Datagram Format (IPv4)

<table>
<tr><td colspan="2">0</td><td colspan="2">8</td><td colspan="2">16</td><td colspan="2">24 … 31</td></tr>
<tr><td>Version</td><td>IHL</td><td colspan="2">Type of Service</td><td colspan="4">Total Length</td></tr>
<tr><td colspan="4">Ident</td><td>Flags</td><td colspan="3">Fragment offset</td></tr>
<tr><td colspan="2">Time to live</td><td colspan="2">Protocol</td><td colspan="4">Header checksum</td></tr>
<tr><td colspan="8">Source Address</td></tr>
<tr><td colspan="8">Destination Address</td></tr>
<tr><td colspan="6">Options</td><td colspan="2">(Padding)</td></tr>
<tr><td colspan="8">Data</td></tr>
<tr><td colspan="8">...</td></tr>
</table>

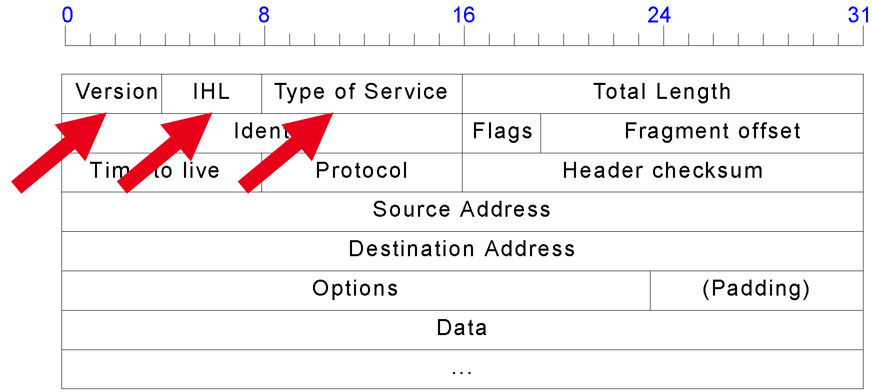
Comments

- Transfer in form of "Big Endian“
- From left to right, highest version bit first

- Big Endian
  - e.g. IBM PowerPC and SUN SPARC computers

- Little Endian
  - x86 architecture (e.g., Intel, AMD)
  - Conversion while receiving and sending

IP Datagram Format



Version

- IP v.4 actual protocol version
- IP v.5 (real time data transfer) ST-2
- IP v.6 successor to IP v.4

Header Length (IHL) (in 32 bit words)

- At least 5 words with 32 bit each = 20 bytes
- At most 15 words with 32 bit each = 60 bytes

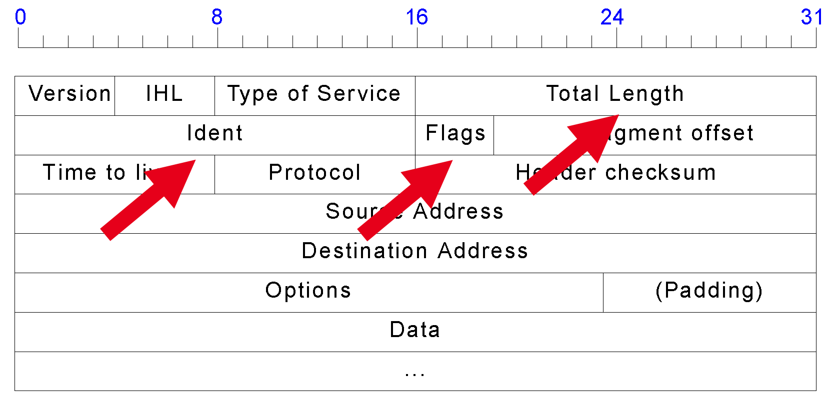
Type of Service

- Simple QoS: a combination of reliability and delay
  - precedence (3 bit):
    - priority 0 (normal) ...7 (network control)
    - influences the queuing scheme (and not routing)
  - D (1 bit): Delay, e.g. no satellite transmission
  - T (1 bit): Throughput, e.g. no telephone line
  - R (1 bit): Reliability, e.g. no radio channels
  - C (1 bit): low Cost, defined later on
  - 1 bit unused

    comment: C & D activated: e.g. invalid
- In practical use: ignored by routers
- Redefined for Differentiated Services (DiffServ)

IP Datagram Format

Total length



- Full length including the data
- Stated in bytes
- All hosts must be prepared to accept datagrams of up to 576 bytes
- Recommendation:
  - send larger datagrams only if assured that destination can handle these
- Max. 65.535 byte, often approximately 1500 byte sent

Identification

- Necessary for destination to determine datagram a fragment belongs to
- All fragments of a datagram contain same identification value

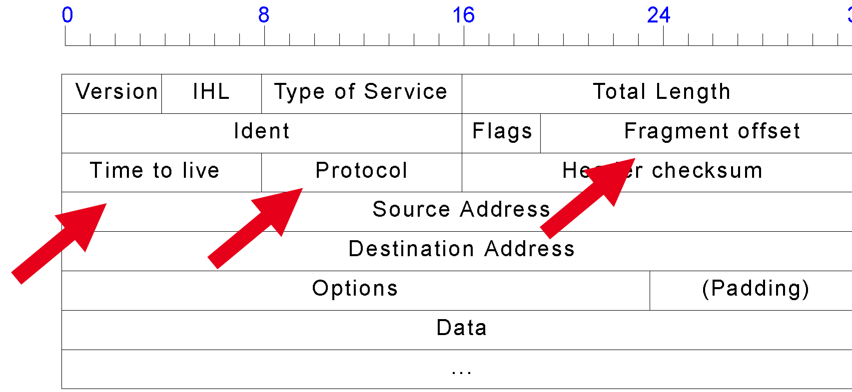
Flags

- 1 bit unused
- DF (1 bit): don’t fragment
  - packets may have a length of up to 576 byte
- MF (1 bit): more fragments
  - last fragment marked 0

IP Datagram Format

Fragment offset

- Offset of this fragment, i.e. the position within a datagram
- Stated in multiples of 8 bytes (elementary fragment unit)
- Length of 13 bits →
  - max. 8192 fragments / datagram
  - max. datagram len. 65536 bytes



Time To Live (TTL)

- Life cycle in seconds, max. 255 sec
- When 0: drop packet, feedback to sender
- Must be decremented per hop, in practical use: counts hops (not seconds)

Protocol: type of the higher level protocol for transmission

No.	Abbreviation	Protocol
0		**reserved**
1	**ICMP**	**Internet Control Message**
2	**IGMP**	**Internet Group Management**
3	**GGP**	**Gateway to Gateway**
4	**IP**	**IP in IP**
5	**ST**	**Stream**
6	**TCP**	**Transmission Control**
…	**…**	**…**

IP Datagram Format

Header Checksum

- Includes Source and Destination Addresses
- To detect errors generated
  by bad memory words inside an IS
- Observed each time
  when datagram is received
  - at IS and at ES
- If necessary datagram is dropped
- Certain summation of the header words
  - addition of all 16-bit halfwords in one’s complement arithmetic and use one’s complement of result (assume this field as zero upon arrival)
- Must be recomputed at each hop
  - due to change in Time-to-Live field

<table>
<tr><td colspan="8">0</td><td colspan="8">8</td><td colspan="8">16</td><td colspan="7">24</td><td colspan="1">31</td></tr>
<tr><td colspan="4">Version</td><td colspan="4">IHL</td><td colspan="8">Type of Service</td><td colspan="16">Total Length</td></tr>
<tr><td colspan="16">Ident</td><td colspan="3">Flags</td><td colspan="13">Fragment offset</td></tr>
<tr><td colspan="8">Time to live</td><td colspan="8">Protocol</td><td colspan="16">Header checksum</td></tr>
<tr><td colspan="32">Source Address</td></tr>
<tr><td colspan="32">Destination Address</td></tr>
<tr><td colspan="24">Options</td><td colspan="8">(Padding)</td></tr>
<tr><td colspan="32">Data</td></tr>
<tr><td colspan="32">...</td></tr>
</table>

Source Address

- Sender’s IP address

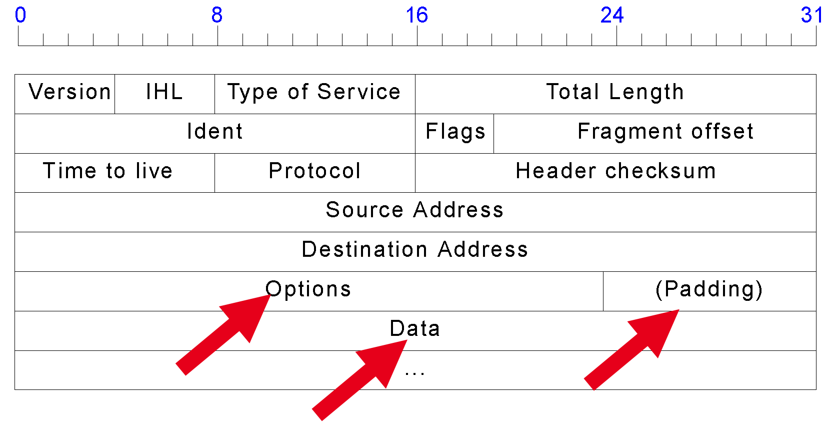
Destination Address

- Receiver’s IP address

IP Datagram Format

Options

- Options for routing, testing and debugging
- Conceptual design
  - as an enhancement for future versions
- Variable length
  - each begins with 1-byte identification code



E.g.

- Security
  - security degree
  - exclusion of routes, but ignored in practice
- Strict source routing
  - the exact route is specified
- Loose source routing
  - part of the route is given, i.e., list of routers to visit
- Record route
  - store IP addresses of routers,
  - but, nowadays headers are too small for this purpose
- Timestamp
  - like record route, but also timestamp added at router

Padding

- Fill up to the word limit

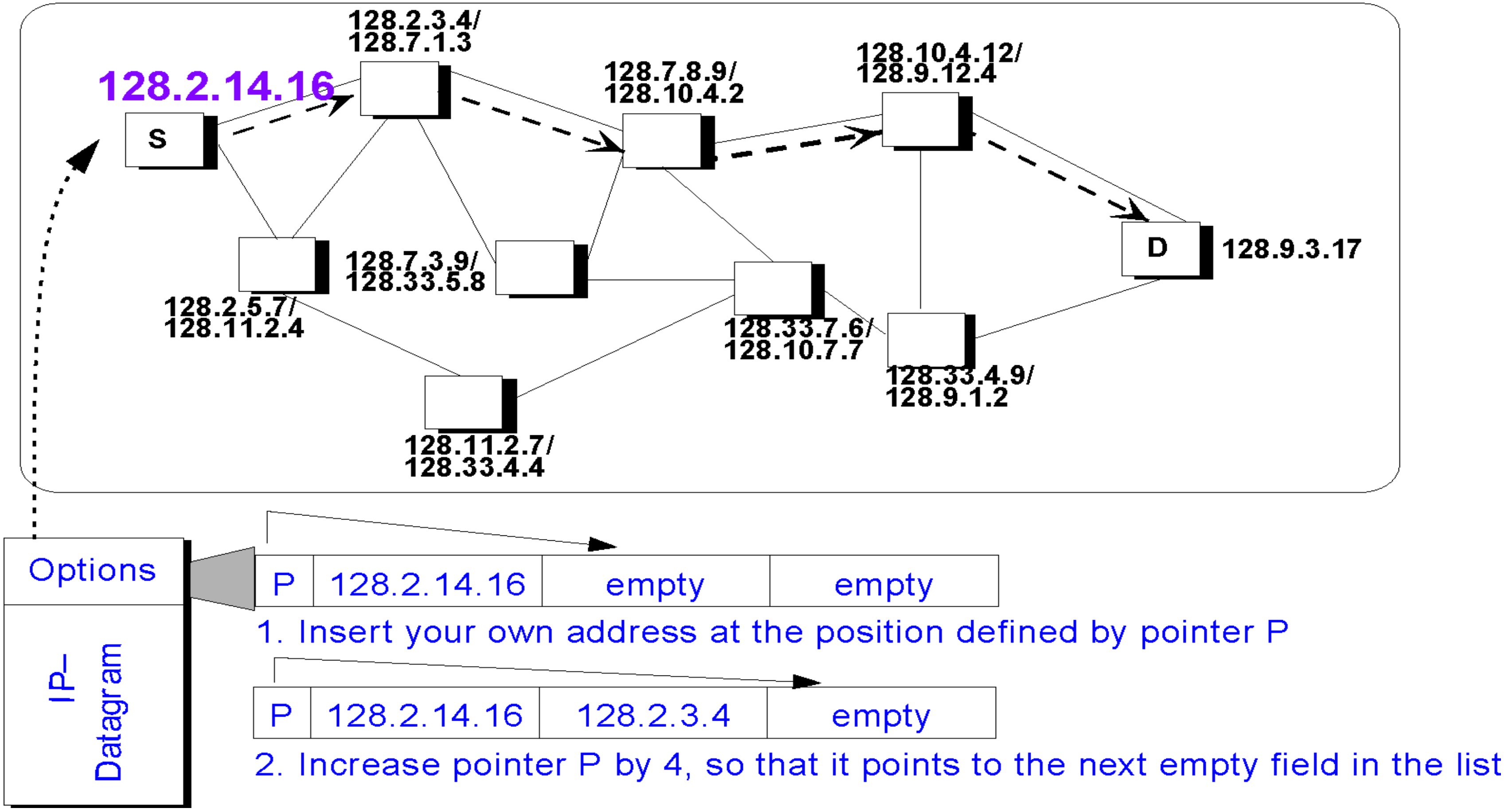
Data

- Field for user data

IP Datagram Format

TECHNISCHE UNIVERSITÄT DARMSTADT

Option: Record Route



3 Internet Control Message Protocol (ICMP)

<table>
<tr><td>SMTP</td><td>HTTP</td><td>FTP</td><td>TELNET</td><td></td><td></td><td>NFS</td><td>RTP</td><td rowspan="2">SCTP</td></tr>
<tr><td colspan="5">TCP</td><td colspan="3">UDP</td></tr>
<tr><td colspan="9">IP + ICMP + ARP</td></tr>
<tr><td colspan="3">WANs, ATM, …</td><td colspan="3">LLC & MAC Physical</td><td colspan="3">LANs, MANs, Ethernet</td></tr>
</table>

History:
J. Postel, RFC 792, Sept. 1981

Purpose
- To communicate network layer information
- Between hosts, routers (and gateways)
- Mostly for error reporting

Sent as IP packets
- I.e., the first 32 bits of the IP data field used as ICMP headers

ICMP origin, e.g.:
- A router was unable to find the given destination address
- Router sent back ICMP (Type 3) packet
- Sending host received the packet, returned error code to TCP
- TCP returned error code to application (e.g. ftp, telnet, http)

E.g., in ftp, telnet,
- http "destination network unreachable"

Internet Control Message Protocol (ICMP)

Header structure

32 bits

Type	Code	Checksum

Type: 16 types, e.g.

- Destination or port or protocol unreachable
- Fragmentation necessary but DF (don’t fragment) DF is set
- Source route failed, redirect (for routing)
- Used by ping program
  - echo request (e.g. for "ping" program)
  - echo reply (response in "ping" program)
- Used by traceroute program
- Source quench (previously congestion control: Choke packet)

Code

- States cause if type is "destination unreachable“ e.g.,
  - net, host, protocol, port unreachable
  - fragmentation needed,
  - source route failed

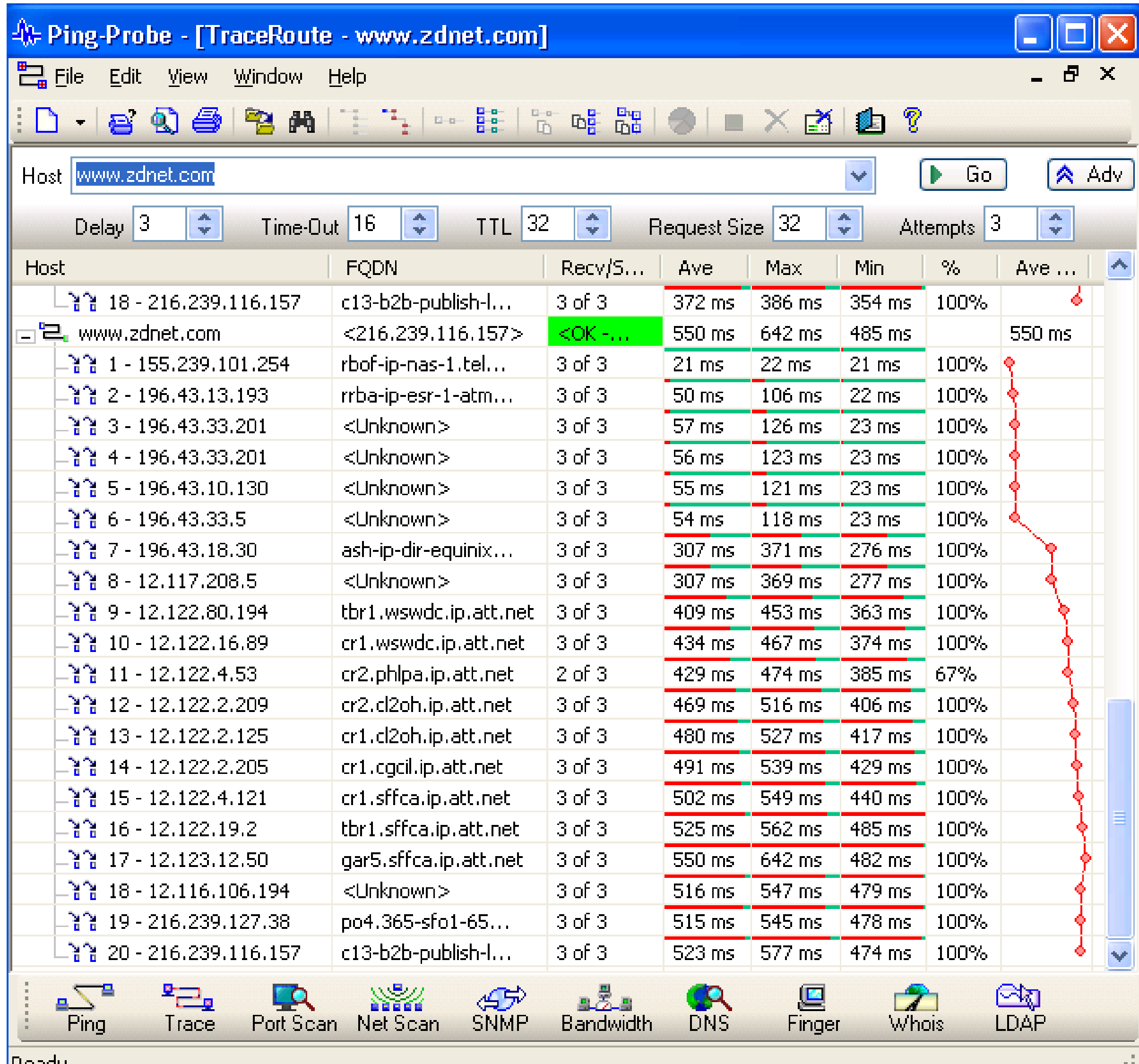
0 8 16 24 31

Version	IHL	Type of Service	Total Length
Ident		Flags	Fragment offset
Time to live	Protocol	Header checksum	
Source Address			
Destination Address			
Options			(Padding)
Data			
...			

Internet Control Message Protocol (ICMP)

....
- Used by ping program
  - echo request
  - echo reply

→ **Traceroute program**



Source: www.ping-probe.com

Motivation: Main issues

- Addressing (presently 32 bit) and
- Many other shortcomings in IP (QoS, mobility, ..)

Background & Status

- 1990: Call for Proposals
- 1992: 21 variants, with 7 possible candidates
- 1993: combination of 2 candidates: S. Deering and Francis (Xerox, Palo Alto)
- Result: RFC 1883-87 protocol, addressing, ICMP, RFC 1825-29, newer ones appeared later (RFC2460-2466)
- Since 2000:
  - On the way to become reality...

IPv6 Basics - Objectives

To support billions of end-systems

- Longer addresses

To reduce routing tables

To simplify protocol processing

- Simplified header

To increase security

- Security means integrated

To support real time data traffic (quality of service)

- Flow label, traffic class

To provide multicasting

To support mobility (roaming)

To be open for change (future)

- extension headers

To coexistence with existing protocols

4.1 Internet Addresses - IPv6

Prefix (binary)	Usage	Fraction
0000 0000	**Reserved (including IPv4)**	**1/256**
0000 0001	**Unassigned**	**1/256**
0000 001	**OSI NSAP addresses**	**1/128**
0000 010	**Novell Netware IPX addresses**	**1/128**
0000 011	**Unassigned**	**1/128**
0000 1	**Unassigned**	**1/32**
0001	**Unassigned**	**1/16**
001	**Unassigned**	**1/8**
010	**Provider-based addresses**	**1/8**
011	**Unassigned**	**1/8**
100	**Geographic-based addresses**	**1/8**
101	**Unassigned**	**1/8**
110	**Unassigned**	**1/8**
1110	**Unassigned**	**1/16**
1111 0	**Unassigned**	**1/32**
1111 10	**Unassigned**	**1/64**
1111 110	**Unassigned**	**1/128**
1111 11100	**Unassigned**	**1/512**
1111 111010	**Link local use addresses**	**1/1024**
1111 111011	**Site local use addresses**	**1/1024**
1111 1111	**Multicast**	**1/256**

IPv6 Addresses and Anycast

I. e.

- Provider based: approx. 16 mio. companies allocate addresses
- Geographically based: allocation as it is today
- Link, site-used: address has only local importance (security, Firewall concept)

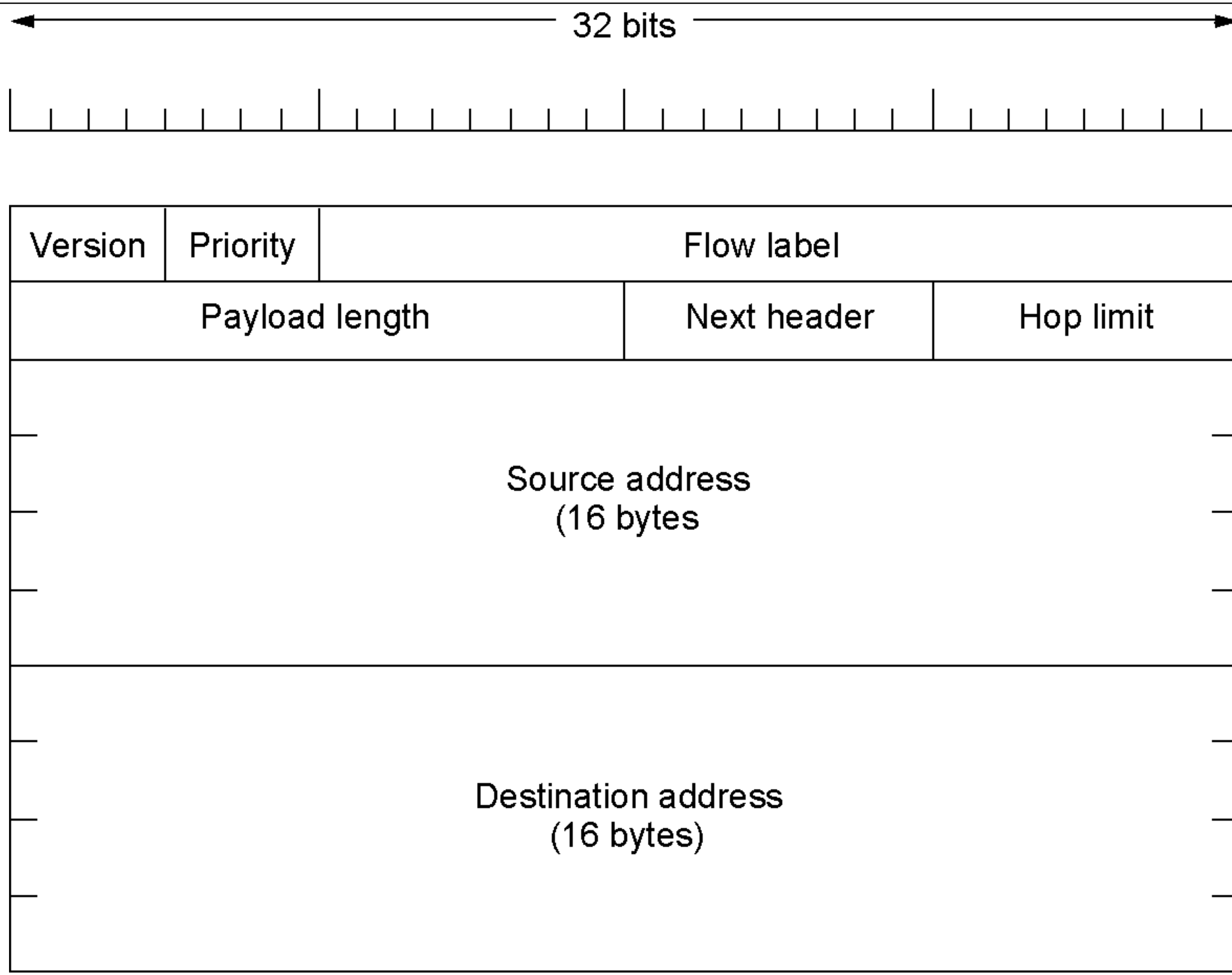
Anycast definition

- Previously
  - unicast, broadcast and multicast
- Now (new)
  - anycast
- Send data to one member of a group
  - for example to the member which is the nearest one geographically
  - i.e. a system within a pre-defined group is to be accessed

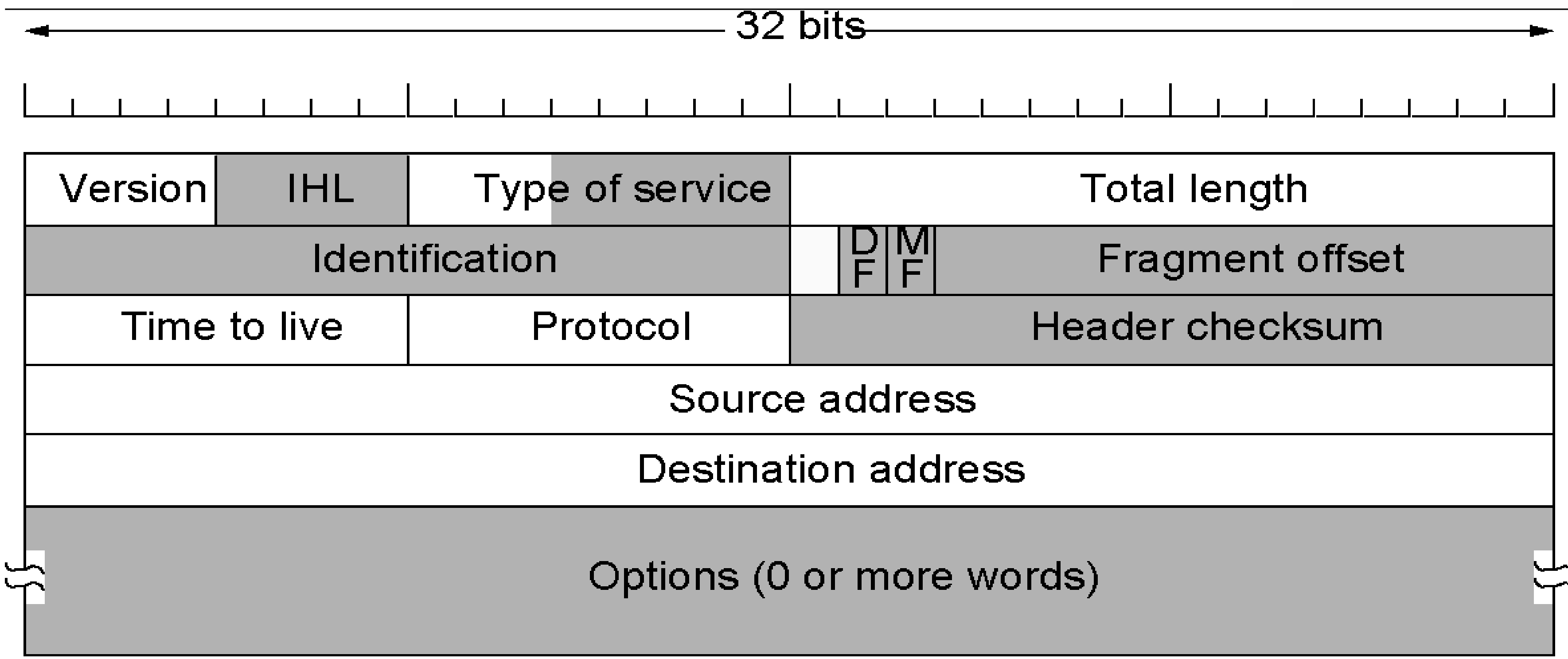
Anycast application

- To search for the nearest web-server
- To locate the nearest router of a multicast group
  - in order to participate in group communication

4.2 IPv6 Datagram Format



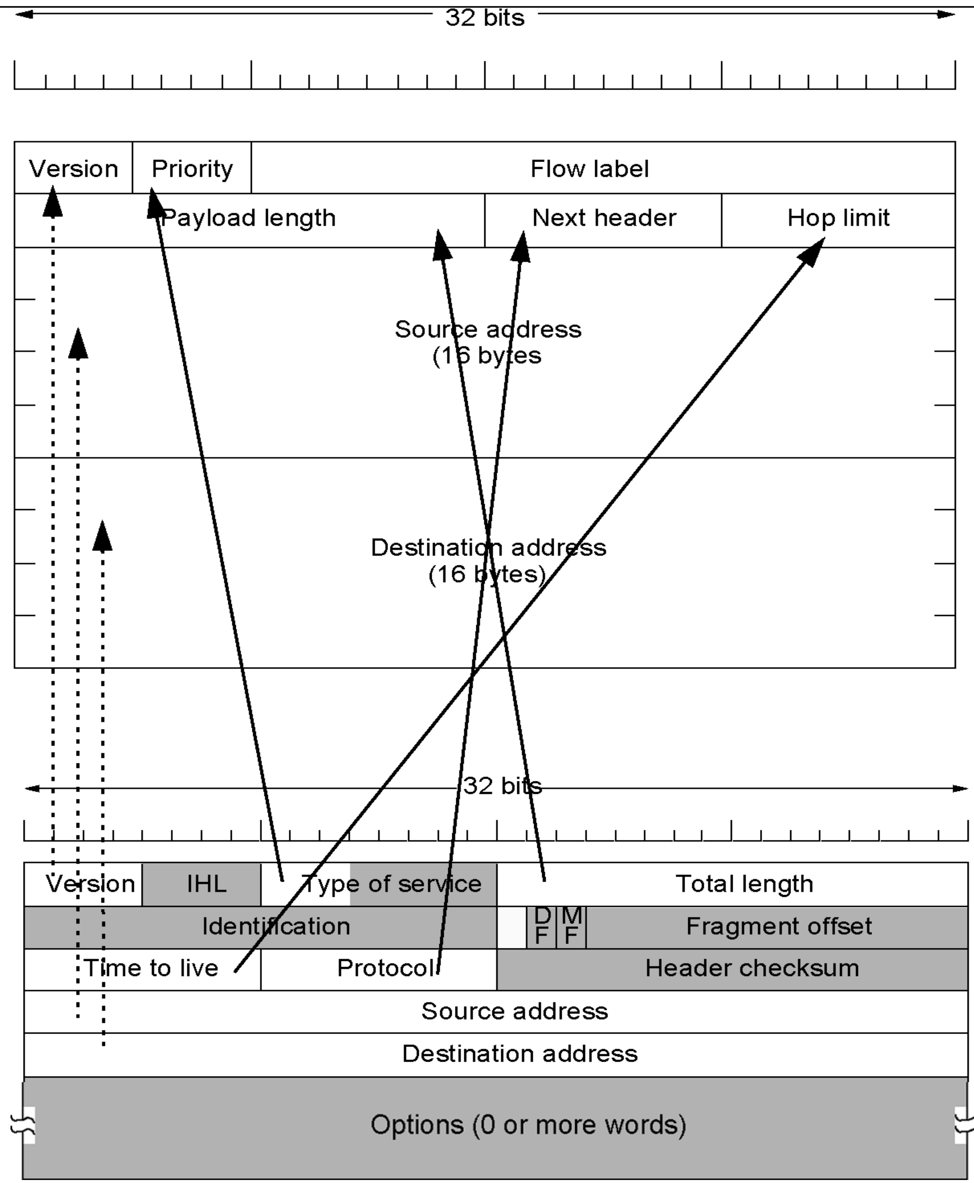
IPv4 Header



In the figure

- Hatching/shading means
  - significant differences to IPv4 or
  - does not exist in IPv6

IPv6 vs IPv4



32 bits
Version
Priority
Flow label
Payload length
Next header
Hop limit
Source address
(16 bytes)
Destination address
(16 bytes)
32 bits
Version
IHL
Type of service
Total length
Identification
D F
M F
Fragment offset
Time to live
Protocol
Header checksum
Source address
Destination address
Options (0 or more words)

IPv6 Header Fields

General comments

- Header with fixed length
- 64 bit alignment (IPv4: 32 bit)

Header length: (eliminated from v.4)

- Efficiency during processing

Version

Version	Description
0 -3	not in use
4	Internet Protocol, IP (presently used)
5	Stream Protocol, ST
6	**IPv6**
7	IPv77, TP/IX CATNIP
8	Pip
9	TUBA
10-15	not in use

IPv6 Header Fields

Priority (further development, precedence of IPv4)

- Differentiation of sources with/without flow control
- Within both groups:
  - low number means lower priority
  - 0..7: with flow control
- 8..15:
  - without flow control
  - continuous rate

Priority	Traffic Description	Example
0	not characterized	
1	"filler"	News
2	"unattended"	E-mail
3	reserved	
4	"attended bulk transfer"	FTP, NFS, HTTP
5	reserved	
6	interactive	Telnet, X
7	Internet Management	SNMP, routing packets

IPv6 Header Fields

TYPE OF Service:

- Precedence replaced by priority (see above)
- D T R C-Bits (QoS) eliminated/replaced by "Flow Label"

Flow Label

- Definition may still evolve
- Flow = Tupel (source ID, dest ID, No.)
- pre-defined
- handling defined by external auxiliary protocol

~~Total Length~~ → Payload Length

- Length including the data (but without the 40 byte header)
  - actually a maximum of 65.535 byte (plus 40 byte header)
- Possibly extension via "Jumbogram" (but then no fragmentation)
- Min. 576 byte
  - a maximum of 65.535 byte may not be enough for a major data transmission

IPv6 Header Fields

~~IDENT ification, FLAGS, FRAGMENT OFFSET~~

- minimum packet size of IPv6 increased
- if still too large packet is sent, then error message
  - L4 should then take over this task and
  - transfer the PDU with the appropriate size to L3

~~Protocol~~ → Next Header

- contains protocol identification
- options (presently):

Extension Header	Description
Hop-by-hop options	**Miscellaneous information for routers**
Routing	**Full or partial route to follow**
Fragmentation	**Management of datagram fragments**
Authentication	**Verification of the sender's identity**
Encrypted security payload	**Information about the encrypted contents**
Destination options	**Additional information for the destination**

~~TIME to live~~ = Hop limit

- life cycle in number of hops, max. 255
  this may not be sufficient, presently usually approx. 32 hops

IPv6 Header Fields

~~HEADER CHECKSUM~~

- L2 and L4 have sufficient mechanisms
- Communication channels better nowadays, at the expense of the performance

Source and Destination ADDRESS

- 32 bit → 128 bit

~~OPTIONS~~

Other changes compared to IPv4

- Checksum removed entirely
  - To speed up processing time at routers
  - Recall: IPv4 checksum has to be recalculated
    - At each router because
      - time to live (TTL) field changes
      - packet may have to be fragmented

- Fragmentation not allowed at routers
  - To speed up processing time at routers
  - Path MTU discovery algorithm used
    - To determine maximum packet size on the way to destination
  - Fragmentation option specified
    - But only for use at source

- ICMPv6 specifies new/adapted message types

Transition from IPv4 to IPv6

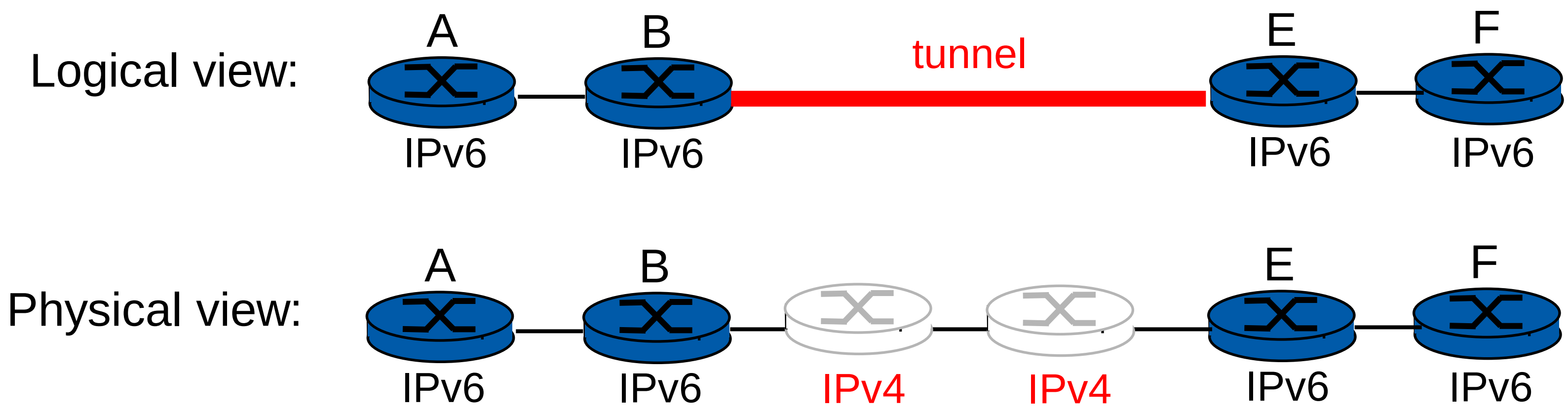
Not all routers are updated simultaneously

- No “flag day” as for switching ARPANET from NCP to IP

How to operate a network with a mix of IPv4 and IPv6 routers?

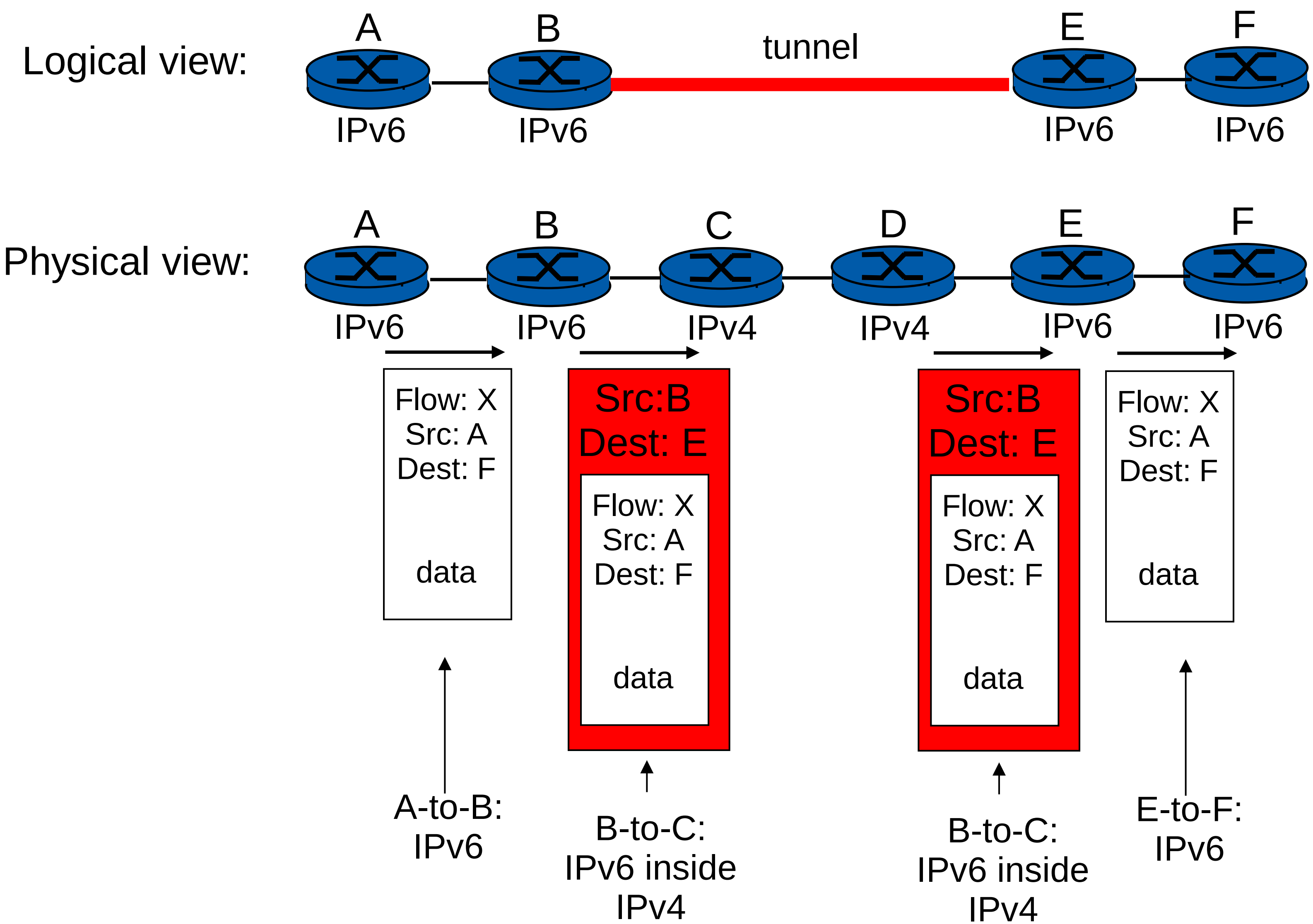
Solution: tunneling

- IPv6 packets carried as payload
  - In IPv4 packets
  - Between IPv4 routers
- IPv6 routers usually also able to handle IPv4
  - No IPv4 in IPv6 tunneling required



Slide source: Computer Networking: A Top Down Approach, 5th edition. J. Kurose, K. Ross

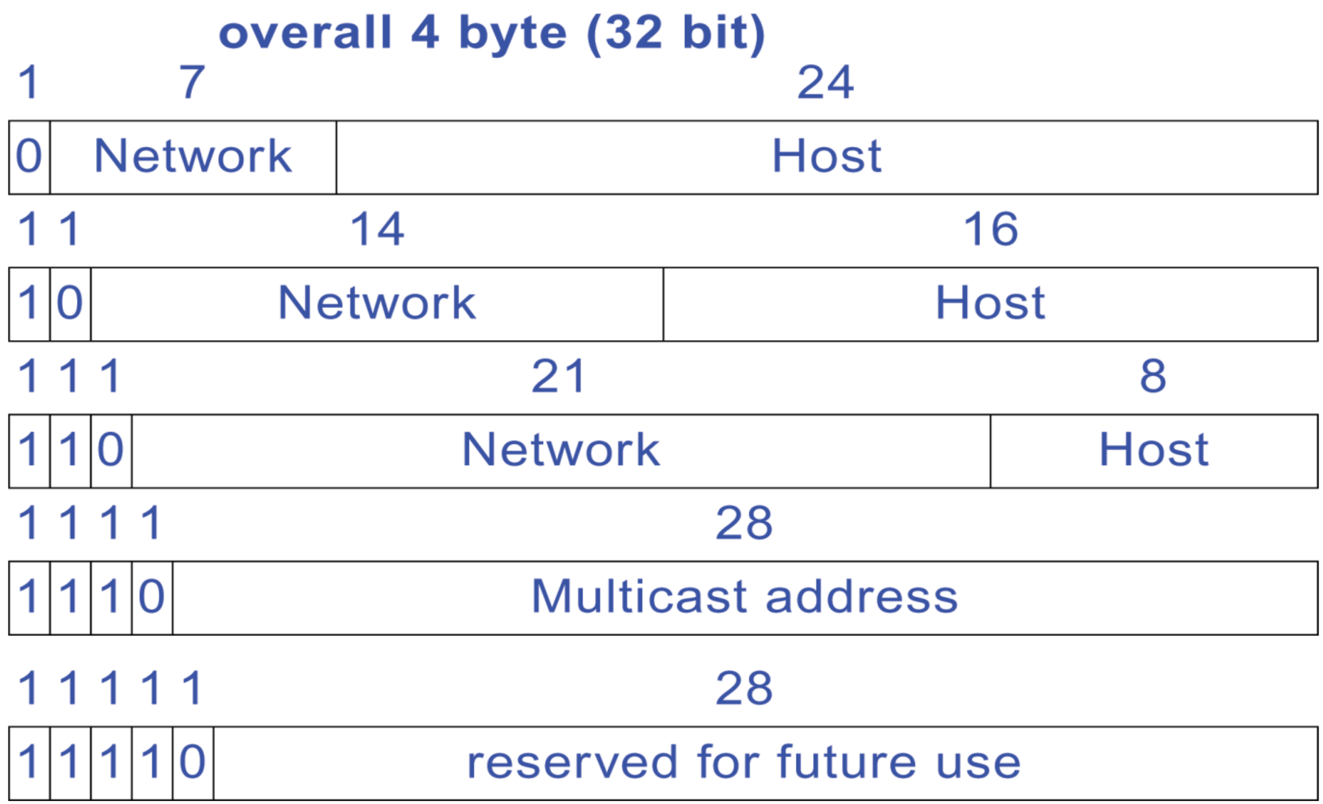
Transition from IPv4 to IPv6



Slide source: Computer Networking: A Top Down Approach, 5th edition. J. Kurose, K. Ross

Global addressing concept for ES (and IS) in the Internet

- 32 bit address (amount is limited!)
- Each address is unique worldwide
- Structure: Net-ID (Subnet-ID), ES-ID



Internet Addresses and Internet Subnetworks

over all 4 byte (32 bit)

1	7	24
0	Network	Host

1	1	14	16
1	0	Network	Host

1	1	1	21	8
1	1	0	Network	Host

1	1	1	1	28
1	1	1	0	Multicast address

1	1	1	1	1	28
1	1	1	1	0	reserved for future use

Global addressing concept for ES (and IS) in the Internet

- Unique 32 bit address with net-ID (subnetwork-Id), ES-Id
- I.e., each network interface (not ES) has its own unique address
- 5 classes

ICANN (Internet Corporation for Assigned Numbers and Names)

- Manages network numbers
- Delegates parts of the address space to regional authorities
  - NIC Network Information Center www.denic.de/

Network addresses typically written in dotted decimal notation

- E.g., 134.169.34.18 or at TUD e.g. 130.83.139.88
- Lowest 0.0.0.0 (0 means this host or network)
- Highest 255.255.255.255 (broadcast on local network)

Special Internet Addresses - Internet Subnetworks

Special IP addresses:

- Source 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 at 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 local network

Network	1 1 1 1 ... 1 1 1 1	Broadcast on distant network

127	(Anything)	Loopback

Structured networks growth

- Several networks instead of one preferable
- But getting several address areas is hard
  - since address space is limited
  - e.g.,
    - university may have started with class B address
    - but, doesn’t get second one

Problem (of classful IP addressing)

- Class A, B, C refer to
  - one network
  - not collection of LANs

Need

→ To allow a network to be split into several parts

- or internal use
- still look like single network to outside world

→ To provide for subnetworks

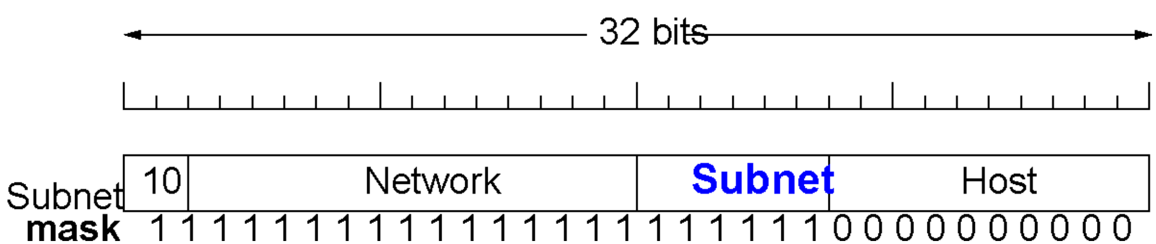
Internet Subnetworks

Subnets

- Ee.g., Ethernet-based LAN

Idea

- Local decision for subdividing host share
  into subnetwork portion and end system portion
- Example: class B address: max. 63 subnetworks



Use subnet mask to indicate split between network + subnet and host part routing with 3 levels of hierarchy

- Algorithm in router
  (by masking bits: i.e. AND between address and subnet mask):
  - packet to another network (yes, then to this router)
  - packet to local ES (yes, then deliver packet)
  - packet to other subnetwork (yes, then reroute to appropriate router)

Given constraints with classes

- IPs growth leads to lack of addresses
  - in principle many addresses due to 32-bit address space
  - but inefficient allocation due to class-based organization
    - class A network with 16 million addresses too big for most cases
    - class C network with 256 addresses is too small
    - most organizations are interested in class B network,
      - but there are only 16384
      - (in reality, class B too large for many organizations)
- Large number of networks leads to large routing tables

→ Introduction of CIDR (Classless InterDomain Routing) (RFC1519)

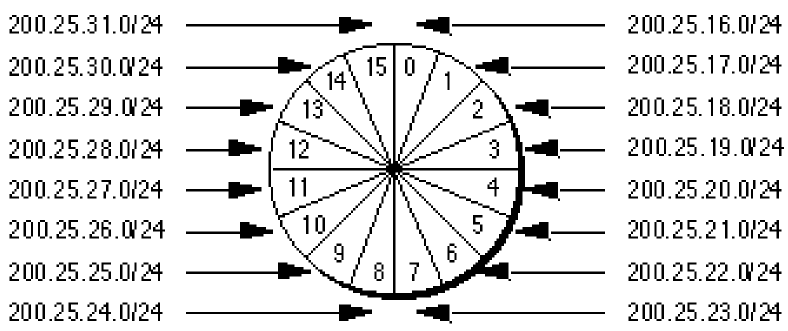
CIDR Principle

- To allocate **IP ADDRESSES** in **VARIABLE-SIZED** blocks
  - without regard to classes
- E.g., request for 2000 addresses would lead to
  - assignment of 2048 address block starting on 2048 byte boundary

But, dropping classes makes forwarding more complicated

CIDR: Classless InterDomain Routing

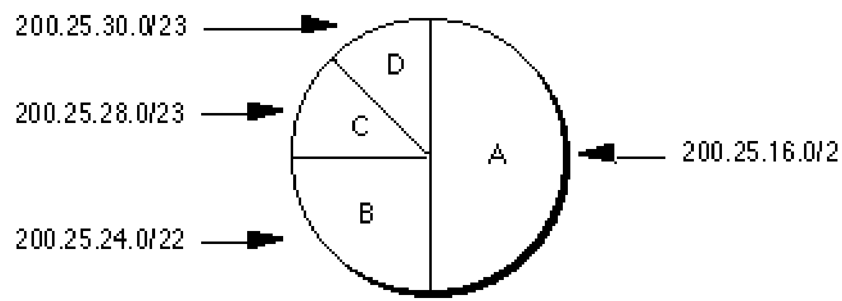
Classful IP Routing



→

In a classful environment, we are forced to use the 16 individual /24s.

Classless IP Routing



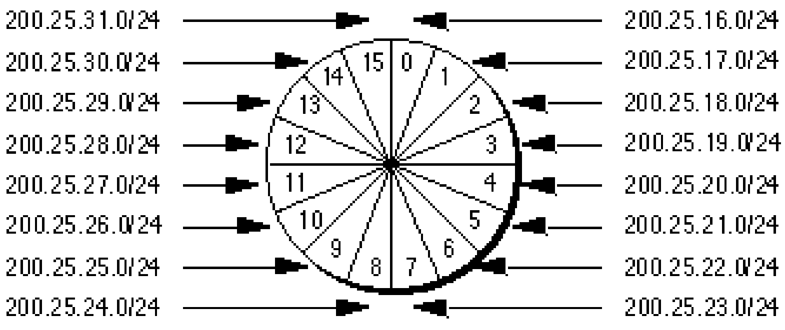
E.g.

- CIDR address 194.24.8.0 / 24,
  - the "/24" indicates
    - first 24 bits used
      to identify unique network
    - remaining bits
      to identify specific host

Source: http://elqui.dcsc.utfsm.cl/util/redes/3COM-Understanding-IP-addressing/nsc/501302.html

CIDR: Classless InterDomain Routing

Classful IP Routing



IP Address 1: 200.25.31.0/24
IP: 11001000.00011001.00011111.00000000
Mask: 11111111.11111111.11111111.00000000

IP Address 1: 200.25.30.0/24
IP: 11001000.00011001.00011110.00000000
Mask: 11111111.11111111.11111111.00000000

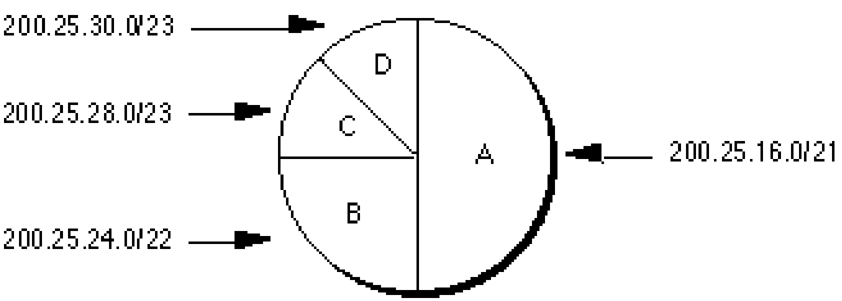
IP Address 1: 200.25.29.0/24
IP: 11001000.00011001.00011101.00000000
Mask: 11111111.11111111.11111111.00000000

IP Address 1: 200.25.28.0/24
IP: 11001000.00011001.00011100.00000000
Mask: 11111111.11111111.11111111.00000000

IP Address 1: 200.25.27.0/24
IP: 11001000.00011001.00011011.00000000
Mask: 11111111.11111111.11111111.00000000

…

Classless IP Routing



IP Address 1: 200.25.30.0/23
IP: 11001000.00011001.00011110.00000000
Mask: 11111111.11111111.11111110.00000000

IP Address 1: 200.25.28.0/23
IP: 11001000.00011001.00011100.00000000
Mask: 11111111.11111111.11111110.00000000

IP Address 1: 200.25.24.0/22
IP: 11001000.00011001.00011000.00000000
Mask: 11111111.11111111.11111100.00000000

IP Address 1: 200.25.16.0/21
IP: 11001000.00011001.00010000.00000000
Mask: 11111111.11111111.11111000.00000000

Source: http://elqui.dcsc.utfsm.cl/util/redes/3COM-Understanding-IP-addressing/nsc/501302.html

TECHNISCHE UNIVERSITÄT DARMSTADT

Additional example

Classfull IP Routing

IP Address 1: 192.168.10.100/24
IP: 11000000.10101000.00001010.01100100
Mask: 11111111.11111111.11111111.00000000
255 . 255 . 255 . 0

IP Address 1: 192.168.11.200/24
IP: 11000000.10101000.00001011.11001000
Mask: 11111111.11111111.11111111.00000000
255 . 255 . 255 . 0

→ hosts in two different networks

Classless IP Routing

IP Address 1: 192.168.10.100/23
IP: 11000000.10101000.00001010.01100100
Mask: 11111111.11111111.11111110.00000000
255 . 255 . 254 . 0

IP Address 1: 192.168.11.200/23
IP: 11000000.10101000.00001011.11001000
Mask: 11111111.11111111.11111110.00000000
255 . 255 . 254 . 0

→ hosts in same network

Source: http://elqui.dcsc.utfsm.cl/util/redes/3COM-Understanding-IP-addressing/nsc/501302.html

CIDR: Classless InterDomain Routing

TECHNISCHE UNIVERSITÄT DARMSTADT

CIDR basics

- Replacement for the old process of assigning Class A, B and C addresses
- With a generalized network "prefix"
  - Instead of being limited to network identifiers (or "prefixes") of 8, 16 or 24 bits
- Uses prefixes anywhere from 13 to 27 bits

  → blocks of addresses can be assigned to networks
  - as small as 32 hosts
  - until over 500.000 hosts

CIDR address

- Includes
  - the standard 32-bit IP address
  - information on how many bits are used for the network prefix
- E.g. CIDR address 194.24.8.0 / **22**,
  - the "**/22**" indicates
    - first 22 bits used to identify unique network
    - remaining bits to identify specific host

CIDR: Classless InterDomain Routing

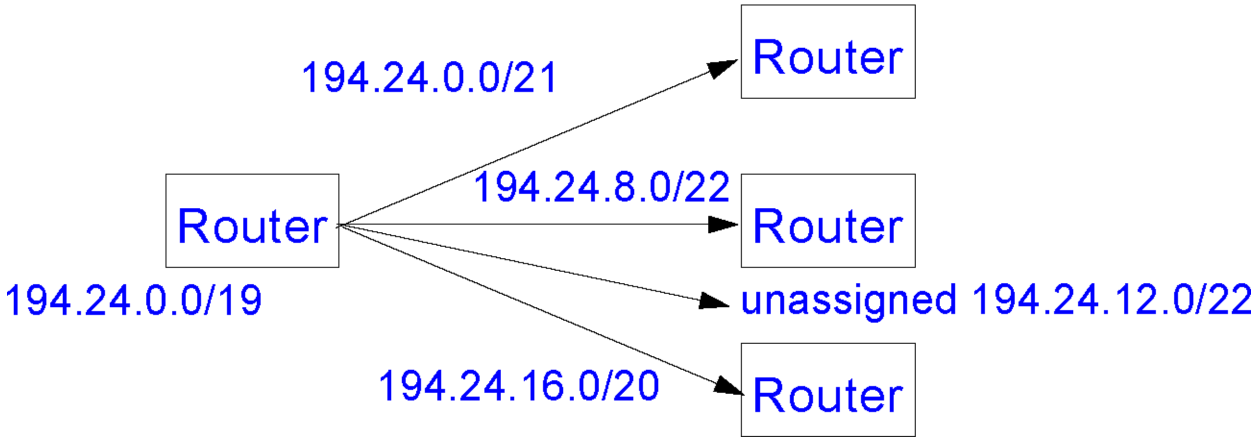
Search for longest matching prefix

- If several entries with different subnet mask length may match
  - then use the one with the longest mask
- I.e., AND operation for address & mask
  - to be performed for each table entry

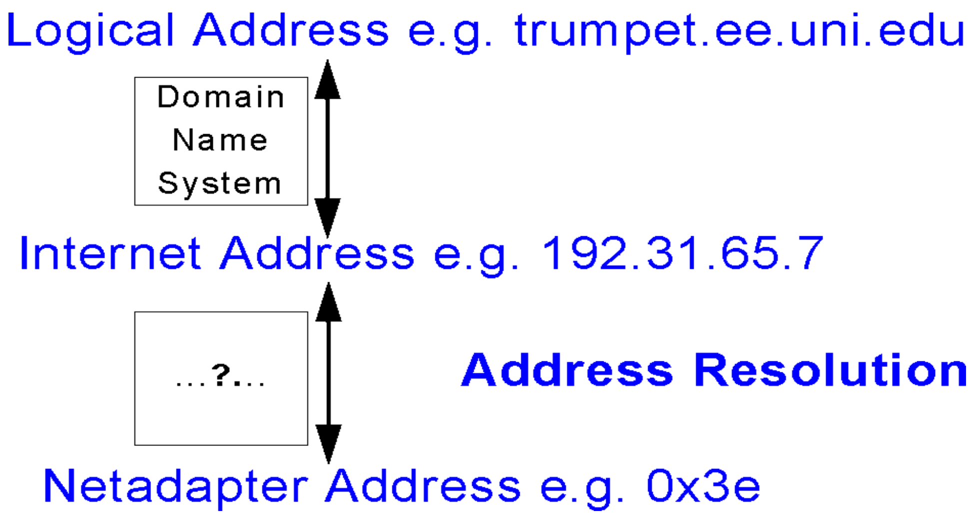
E.g.,

- Potentially several 'class C' networks can be characterized by one prefix

Entries may be aggregated to reduce routing tables



6 Address Resolution



Addressing levels

Host identification and routing specification within a subnetwork

- Based on the (local) physical network addresses of ES
  - e.g., station address of the adapter card

Problem:

- INTERNET address (32 bit) must be mapped onto the physical network address,
  - usually 48 bit (ADDRESS RESOLUTION)

Address Resolution: Methods

Address resolution in

- Source ES, if destination ES is local (direct routing)
- Gateway, if destination ES is not local

Solutions:

1. Direct HOMOGENEOUS ADDRESSING

- If the physical address can be dialed by the user, then the dial-up is:
  - physical address = Hostid of the INTERNET address

2. If the physical address is pre-defined or if it has to have a different format

- A mapping table from configuration data base (IPaddr → HWaddr),
  - e.g. in the Gateway,
  - may become maintenance nightmare
- The Address Resolution Protocol (ARP)
  - mainly applied in LANs with broadcasting facility

6.1 Address Resolution Protocol (ARP)

Process

1. Broadcast ARP request datagram on LAN

- including receiver’s INTERNET address (desired value)
- sender’s physical (HW) and INTERNET address (IP)

2. Every machine on LAN receives this request and checks address

3. Reply by sending ARP response datagram

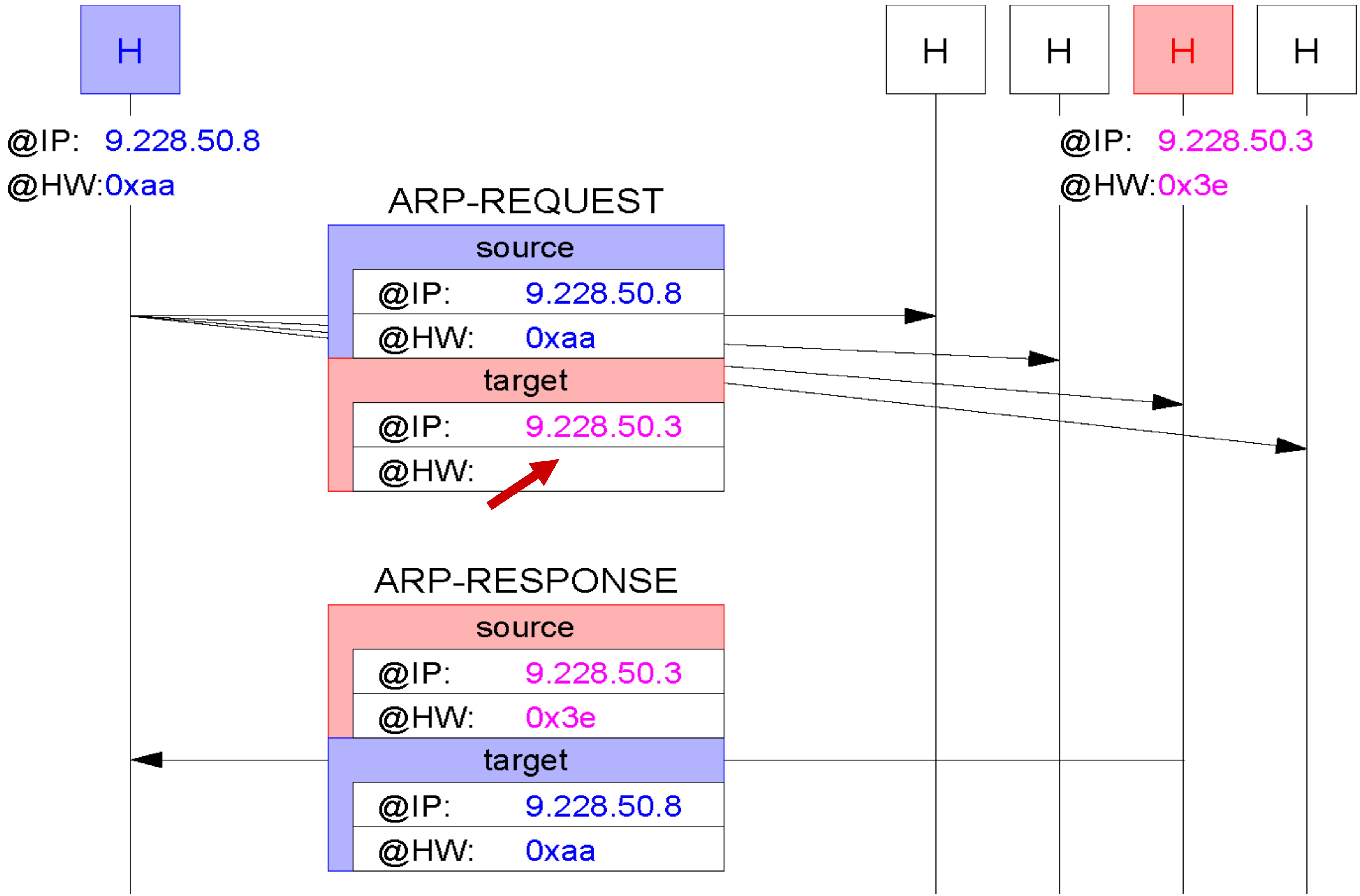
- machine which has requested address responses
- including the phyiscal address

4. Store pair (I,P) for future requests

Refinements

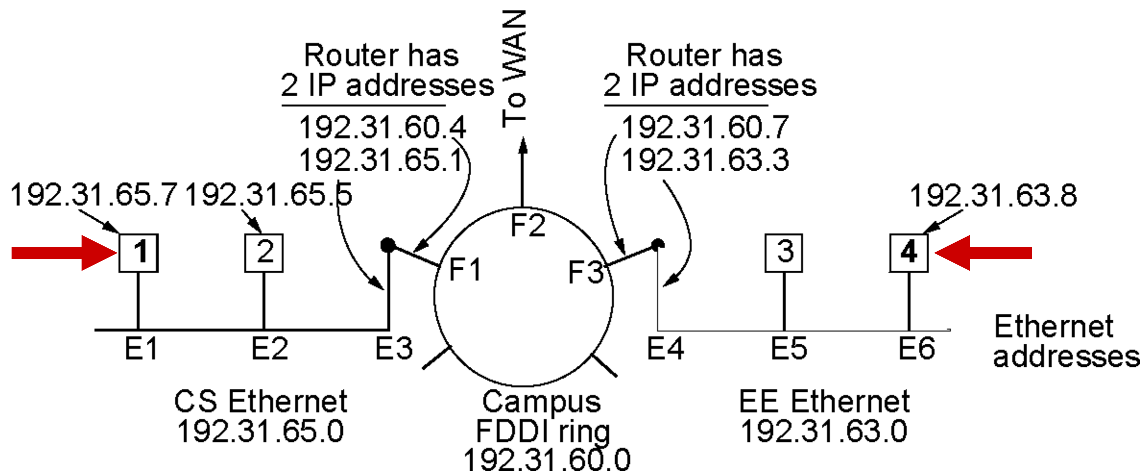
- Receiver of ARP request stores sender’s (I,P) pair in its cache
- Send own table during the boot process
  - (but may be too old)
- Entries in ARP cache should time out after some time
  - (few minutes)

Address Resolution Protocol (ARP)



Address Resolution Protocol (ARP)

End system not directly available by broadcast



Example: ES 1 to ES 4

- ARP would not receive a response
  - Ethernet Broadcast is not rerouted over a router

Solution: proxy ARP

- The local router knows all remote networks with their respective routers
  - responds to local ARP
- Local ES 1 sends data for ES 4 always to the local router,
  - his router forwards the data
    - (by interpreting the IP address contained in the data)

Solution: remote network address is known

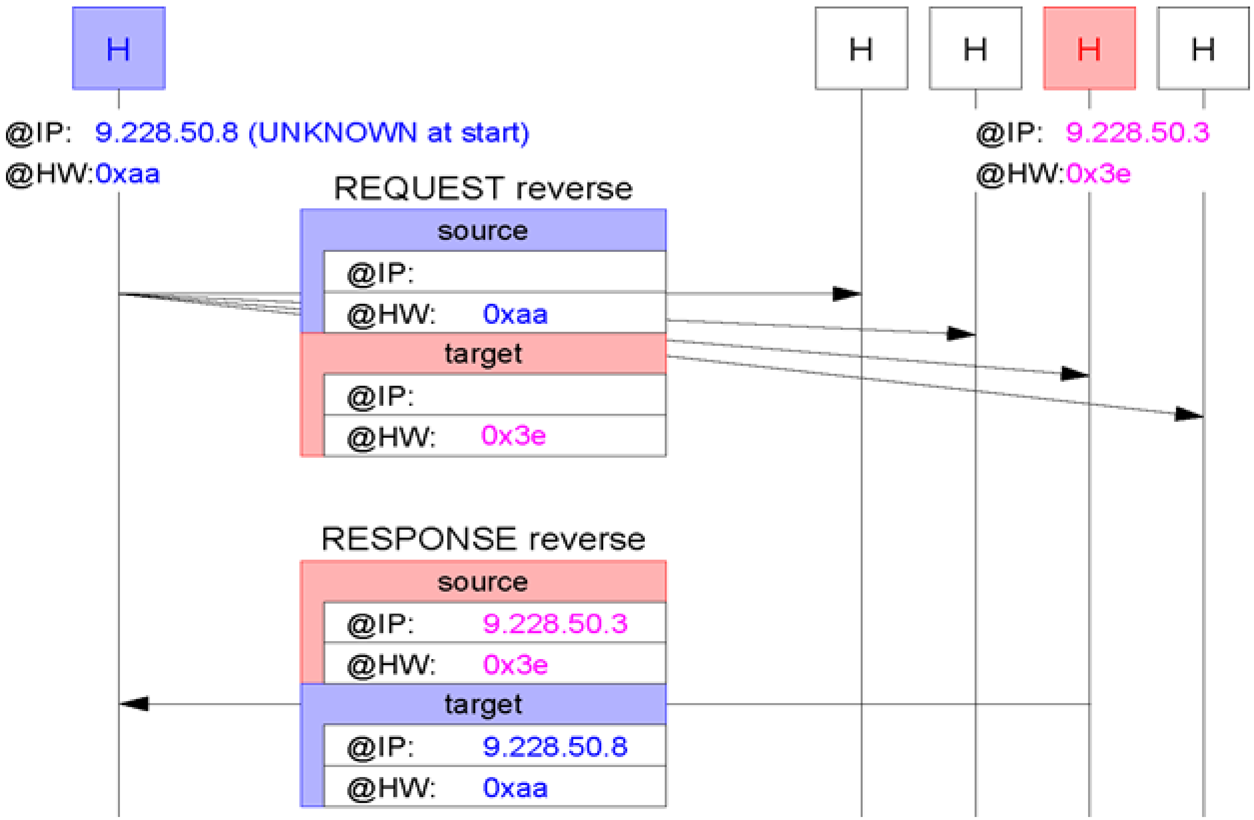
- Local ES 1 sends data to the appropriate remote router
- Local router forwards packets

6.2 Reverse Address Resolution Protocol (RARP)

RFC 903: Retrieve Internet address from knowledge of hardware address

- RARP server responds
- RARP server has to be available on the LAN

Application: diskless workstation boots over the network



6.3 DHCP: Dynamic Host Configuration Protocol

DHCP has largely replaced RARP (and BOOTP)

- Extends functionality

DHCP

- Simplifies installation and configuration of end systems
- Allows for manual and automatic IP address assignment
- May provide additional configuration information
  - (DNS server, netmask, default router, etc.)

DHCP server is used for assignment

- Request can be relayed by DHCP relay agent,
  - if server on other LAN

Client broadcasts DHCP DISCOVER packet

- Server answers

Address is assigned for limited time only

- Before the ’lease’ expires, client must renew it
- Allows to reclaim addresses of disappearing hosts

Overview: Routers are grouped into Autonomous Systems (AS)

Within an Autonomous System,
all routers run the same routing algorithm and
all know each other → Interior Gateway Protocols

- This routing algorithm is intra- Autonomous System routing algorithm
- Can be link state or distance vector (or something else)

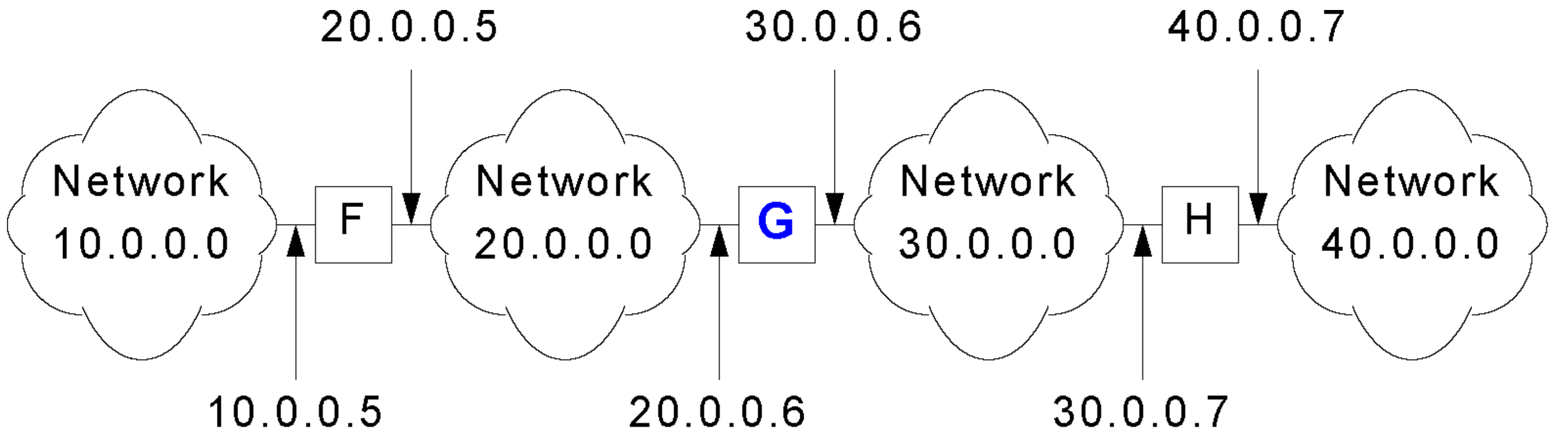
Some routers in an Autonomous System are gateway routers

- Gateway routers are also connected to other gateway routers in other AS

Gateway routers run inter- Autonomous System routing algorithm
→ Exterior Gateway Protocols

- Needs to be standardized

IP Routing



	TO REACH HOSTS ON NETWORK	ROUTE TO THIS ADDRESS
G:	20.0.0.0	DELIVER DIRECT
	30.0.0.0	DELIVER DIRECT
	10.0.0.0	20.0.0.5
	40.0.0.0	30.0.0.7

Routing tables of Gateways

Gateways may have incomplete information→ default paths

Autonomous Systems

Autonomous systems solve problems of scale and administrative authority

Router needs to know about the routers in its own Autonomous System and its own gateway routers

- Solves problem of scaling

Two-level routing

- Intra - Autonomous System: administrator is responsible for choice
  - Routing Information Protocol (RIP):
    - Distance Vector
  - Open Shortest Path First (OSPF):
    - Link State
  - Interior Gateway Routing Protocol (IGRP):
    - Distance Vector (Cisco proprietary)

- Inter - Autonomous System: unique standard
  - Border Gateway Protocol (BGP): Path Vector
    - (sort of distance vector, but with path information for loop avoidance)