Introduction to Internetworking

Introductory terms

Communications Network

Facility that provides data transfer services

An internet

Collection of communications networks interconnected by bridges and/or routers

The Internet - note upper case I

The *global collection* of thousands of individual machines and networks

Intranet

Corporate internet operating within the organization

Uses Internet (TCP/IP and http) technology to deliver documents and resources

End System (ES)

Device attached to one of the networks of an internet

Supports end-user applications or services

Intermediate System (IS)

Device used to connect two networks

Permits communication between end systems attached to different networks

Bridge /switch - which has more interfaces

IS used to connect two LANs using similar LAN protocols

Address filter passing on packets to the required network only

OSI layer 2 (Data Link)

Router

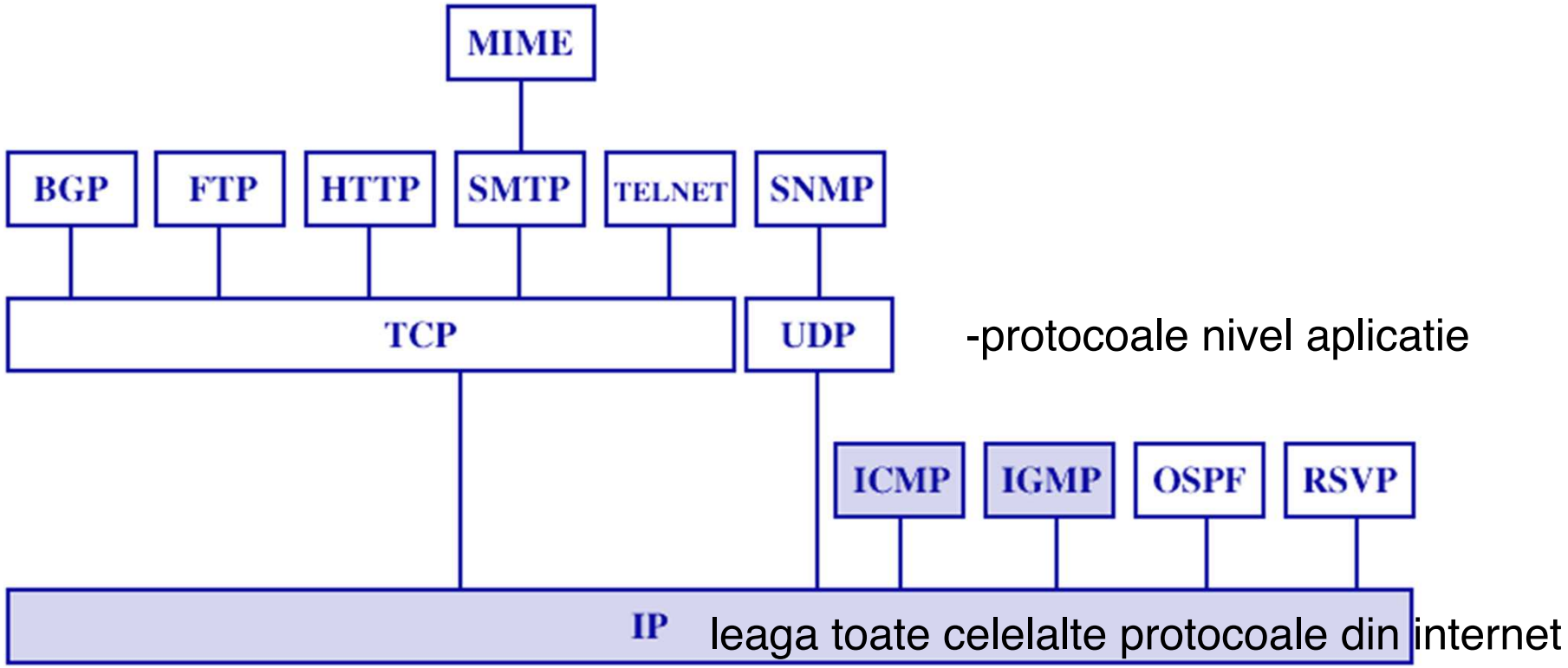
Connects two (possibly dissimilar) networks

Uses internet protocol present in each router and end system

OSI Layer 3 (Network)

Internetworking Protocols

TCP/IP stack and suite of internetworking protocols



Requirements of Internetworking

Link between networks

- Minimum physical and link layer

Routing and delivery of data between processes on different networks

Accounting services and status info

Independent of network architectures

Network Architecture Specific Features

Addressing

Packet size

Access mechanism

Timeouts

Error recovery

Status reporting

Routing

User access control

Architectural Approaches

Connection oriented

Connectionless

Connection Oriented

Assume that each network is connection oriented

IS connect two or more networks

- IS appear as DTE to each network
- Logical connection set up between DTEs
  - Concatenation of logical connections across networks
- Individual network virtual circuits joined by IS

May require enhancement of local network services

- 802, FDDI are datagram services

Connection Oriented IS Functions

Relaying

Routing

e.g. X.75 used to interconnect X.25 packet switched networks

Connection oriented not often used

Connectionless Operation

Corresponds to datagram mechanism in packet switched networks

Each NPDU treated separately

Network layer protocol common to all DTEs and routers

- Known generically as the internet protocol

Internet Protocol

- One such internet protocol developed for ARPANET
- RFC 791
- Lower layer protocol needed to access particular network

Connectionless Internetworking

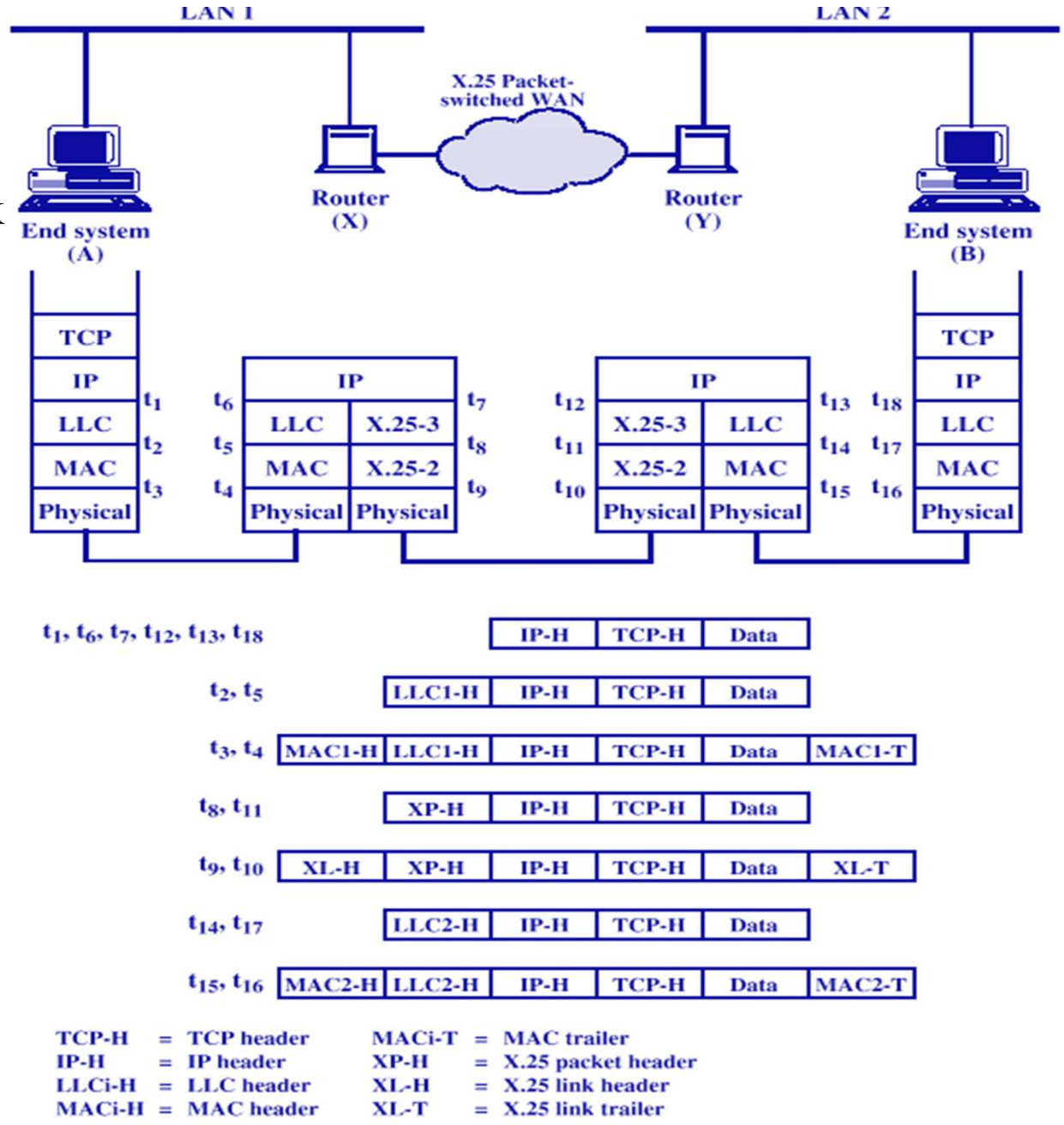
Advantages

- Flexibility
- Robust
- No unnecessary overhead

Main drawback: Unreliable

- Not guaranteed delivery
- Not guaranteed order of delivery
  - Packets can take different routes
- Reliability is responsibility of next layer up (e.g. TCP)

Example of an IP protocol operations, acting over a X.25 packet switched WAN network



Design Issues

- Routing
- Datagram lifetime
- Fragmentation and re-assembly
- Error control
- Flow control

Routing

End systems and routers maintain routing tables

- Indicate next router to which datagram should be sent
- Static
  - May contain alternative routes
- Dynamic
  - Flexible response to congestion and errors

Source routing

- Source specifies route as sequential list of routers to be followed
- Security
- Priority

Route recording

Datagram Lifetime

Datagrams could loop indefinitely

- Consumes resources
- Transport protocol may need upper bound on datagram life

Datagram marked with lifetime

- Time To Live field in IP
- Once lifetime expires, datagram discarded (not forwarded)
- Hop count
  - Decrement time to live on passing through a each router
- Time count
  - Need to know how long since last router

Fragmentation and Re-assembly

Different packet sizes

When to re-assemble

- At destination
  - Results in packets getting smaller as data traverses internet
- Intermediate re-assembly
  - Need large buffers at routers
  - Buffers may fill with fragments
  - All fragments must go through same router
    - Inhibits dynamic routing

IP re-assembles at destination only

Uses fields in header

Data Unit Identifier (ID)

- Identifies end system originated datagram
  - Source and destination address
  - Protocol layer generating data (e.g. TCP)
  - Identification supplied by that layer

Data length

- Length of user data in octets

Offset

- Position of fragment of user data in original datagram
- In multiples of 64 bits (8 octets)

More flag

- Indicates that this is not the last fragment

Dealing with Failure

Re-assembly may fail if some fragments get lost

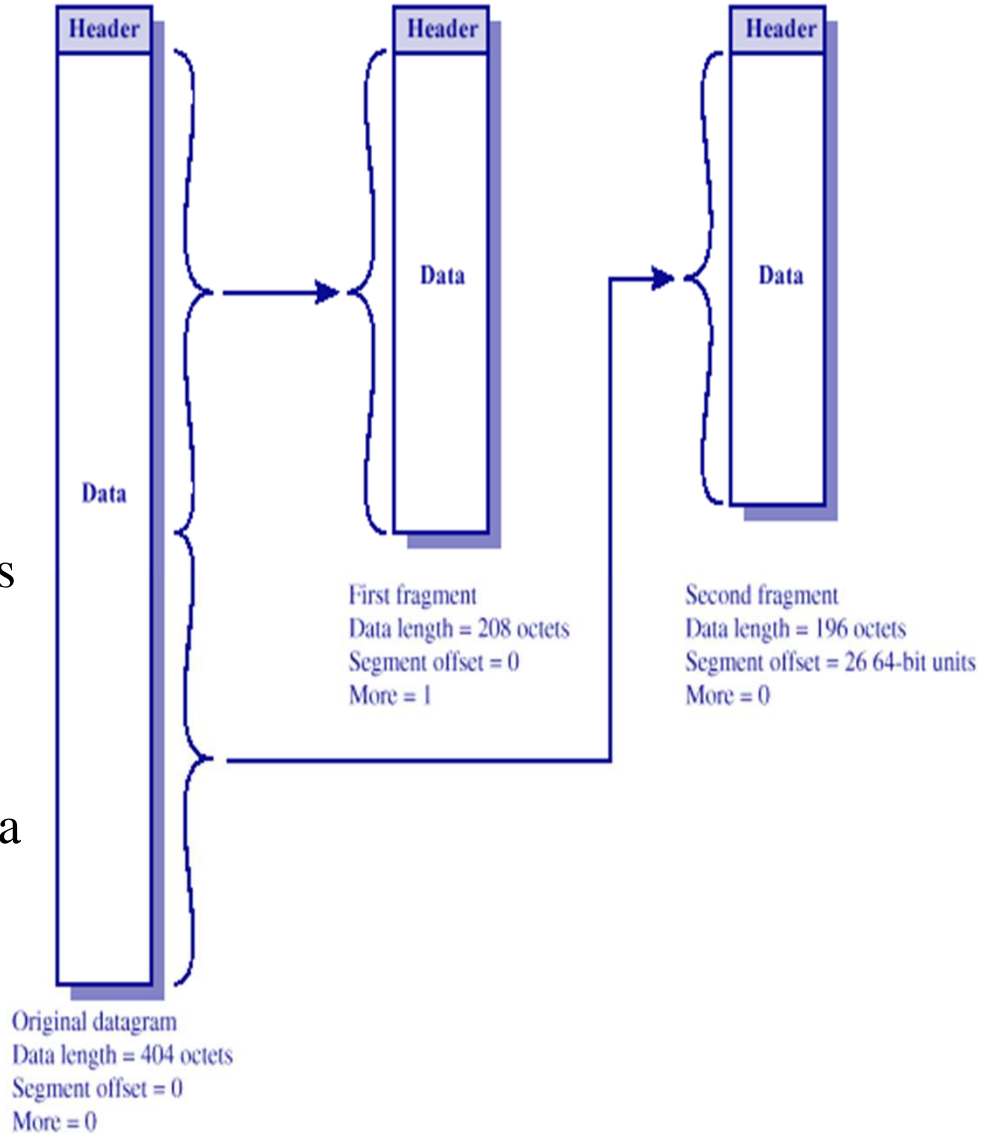
Need to detect failure

Re-assembly time out

- Assigned to first fragment to arrive
- If timeout expires before all fragments arrive, discard partial data

Use packet lifetime (time to live in IP)

- If time to live runs out, kill partial data



Error Control

Not guaranteed delivery

Router should attempt to inform source if packet discarded

e.g. for time to live expiring

Source may modify transmission strategy

May inform high layer protocol

Datagram identification needed

(Look up ICMP)

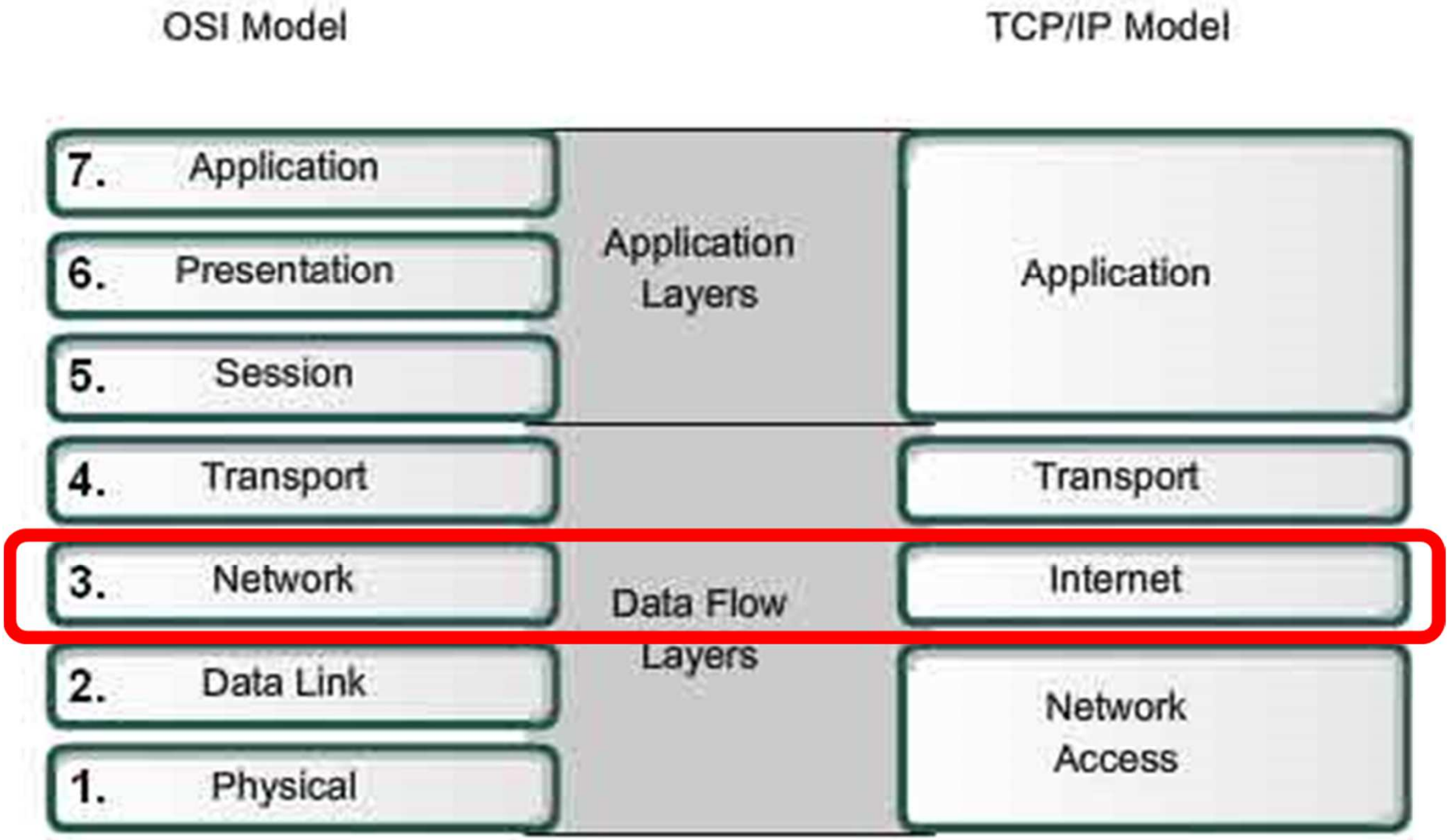
Flow Control

Allows routers and/or stations to limit rate of incoming data

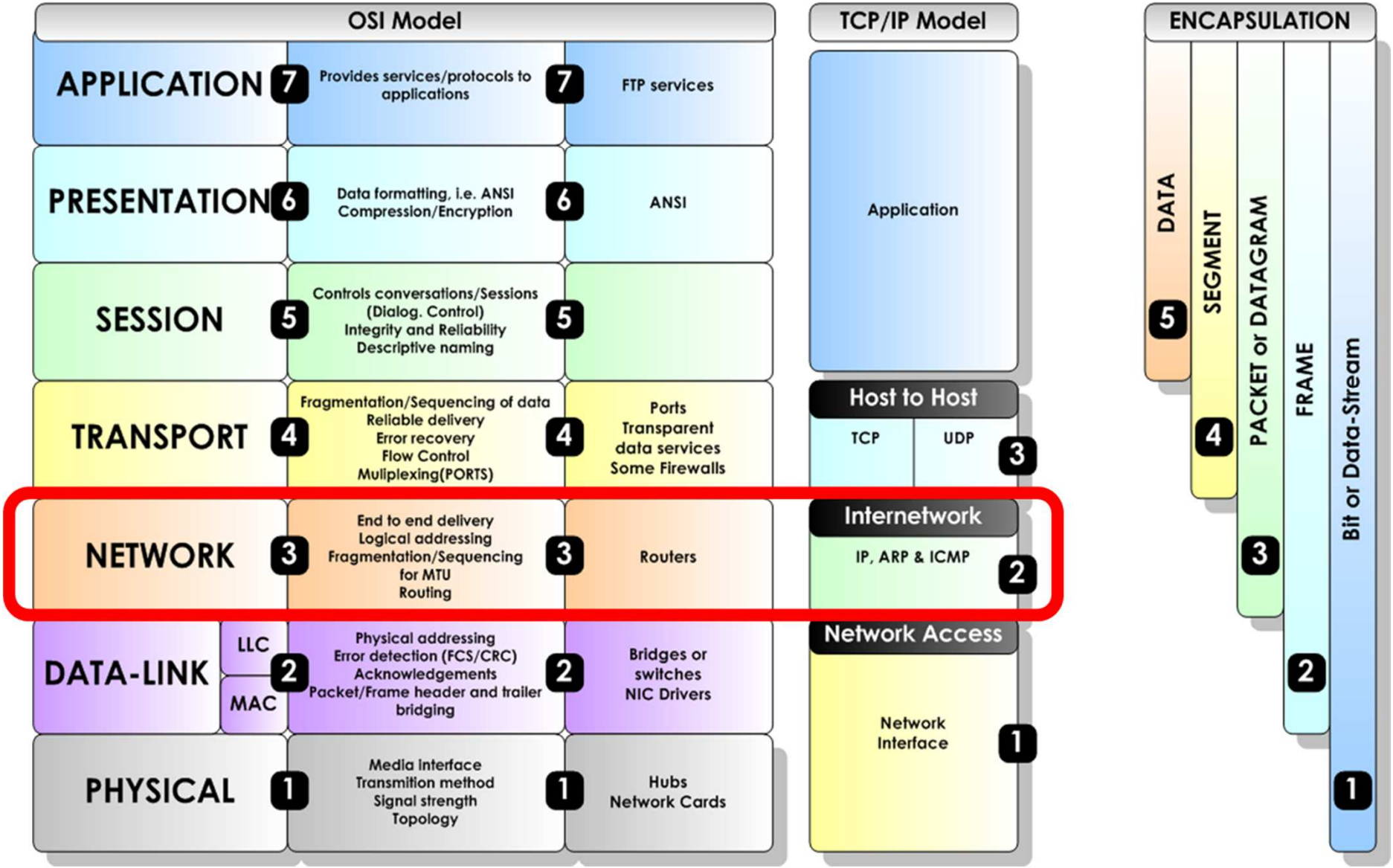
Limited in connectionless systems

Send flow control packets

Requesting reduced flow, e.g. ICMP


OSI Model
TCP/IP Model
7. Application
6. Presentation
5. Session
4. Transport
3. Network
2. Data Link
1. Physical
Application Layers
Data Flow Layers
Application
Transport
Internet
Network Access

The OSI Model (Open Systems Interconnection)



Internet Protocol (IP)

Part of TCP/IP, used by the Internet

Specifies interface with higher layer, e.g. TCP

Specifies protocol format and mechanisms

IP Services

Primitives

Functions to be performed

Form of primitive implementation dependent, e.g. subroutine call

Send

Request transmission of data unit

Deliver

Notify user of arrival of data unit

Parameters

Used to pass data and control info

Source address

Destination address

Protocol

Recipient, e.g. TCP

Type of Service

Specify treatment of data unit during transmission through networks

Identification of IP packet

Source, destination address and user protocol

Uniquely identifies PDU

Needed for re-assembly and error reporting

Send only

Don't fragment indicator

Can IP fragment data

If not, may not be possible to deliver

Send only

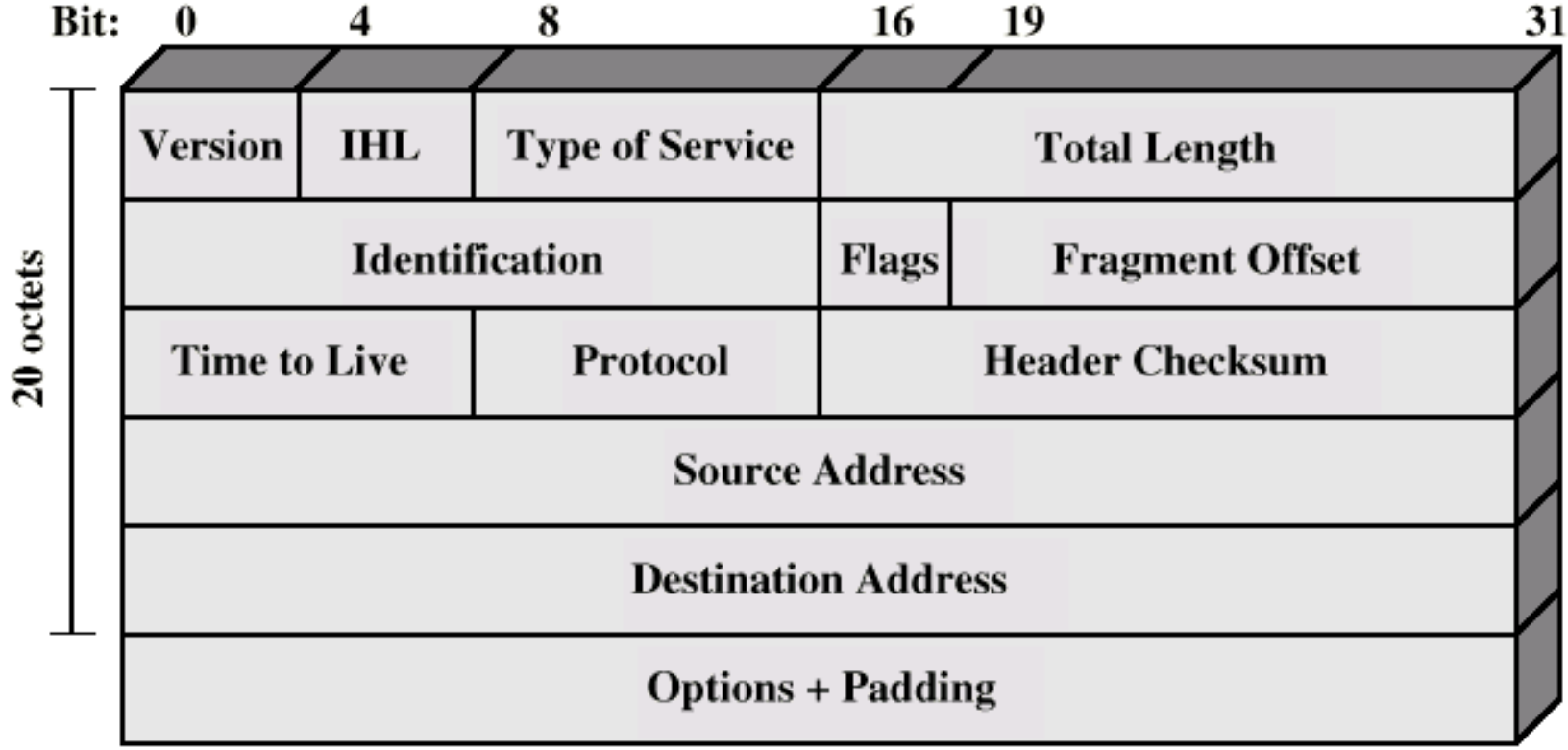
Time to live

Send only

Data length

Option data

User data



IP packet structure

Header Fields

Version

Currently 4

IP v6 - see later

Internet header length

In 32 bit words

Including options

Type of service

Total length

Of datagram, in octets

Identification

Sequence number

Used with addresses and user protocol to identify datagram uniquely

Flags

More bit

Don’t fragment

Fragmentation offset

Time to live -cand pachetul intra intr-o bucla sa pot arunca pachetele, deci ne zice pri cate hopuri poate sa treaca pachetul pana se atinge limita si e eliminat, daca nu ajunge destul de repede la destinatie

Protocol

Next higher layer to receive data field at destination /ce fel de pachet e encapsulat

Header checksum

Reverified and recomputed at each router

16 bit ones complement sum of all 16 bit words in header

Set to zero during calculation

Source address

Destination address

Options

Padding

To fill to multiple of 32 bits long

Data Field

Carries user data from next layer up

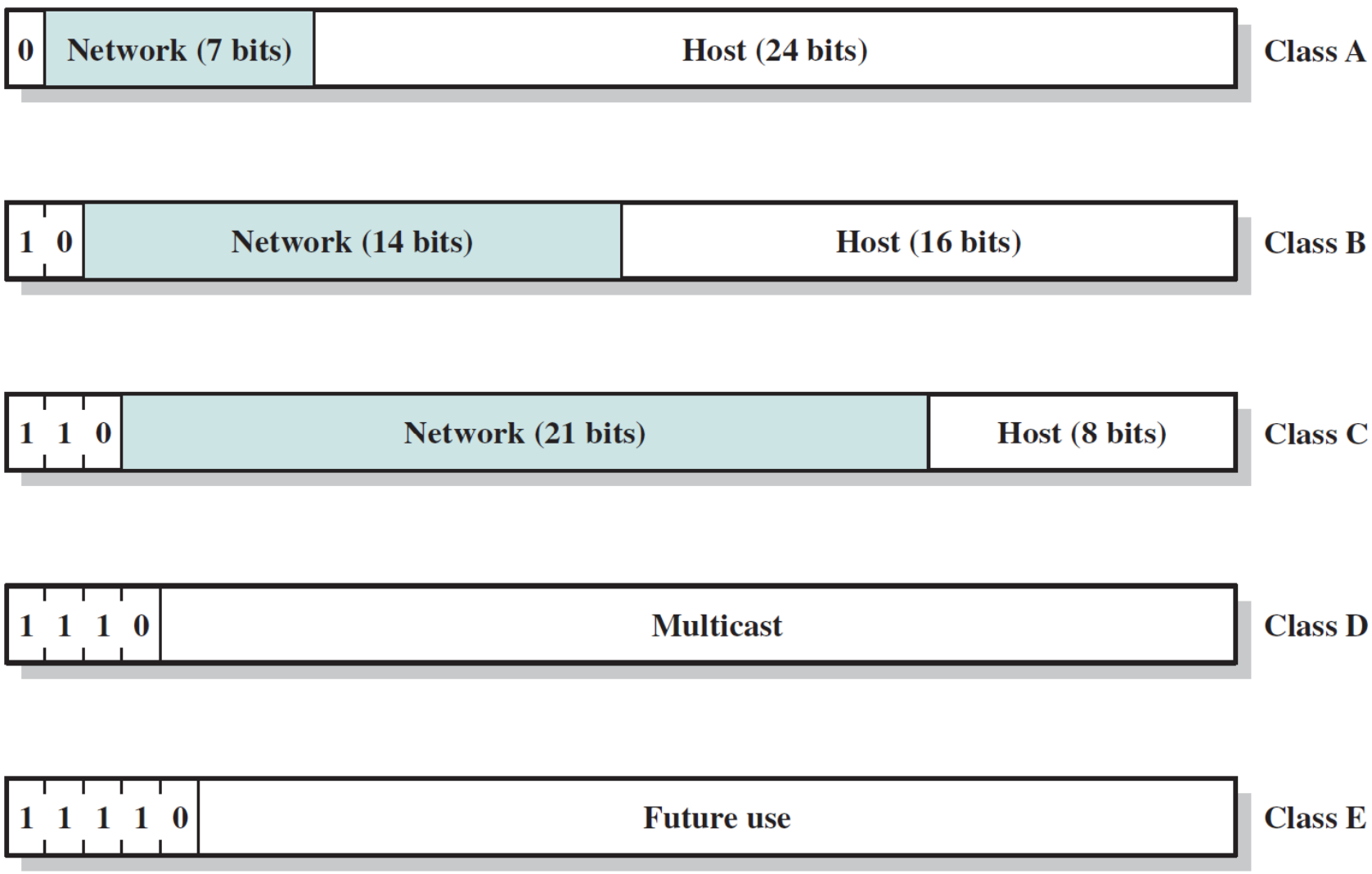
Integer multiple of 8 bits long (octet)

Max length of datagram (header plus data) 65,535 octets

IP Addresses

32 bit global internet address

Network part and host part


0
Network (7 bits)
Host (24 bits)
Class A
1 0
Network (14 bits)
Host (16 bits)
Class B
1 1 0
Network (21 bits)
Host (8 bits)
Class C
1 1 1 0
Multicast
Class D
1 1 1 1 0
Future use
Class E

Class A

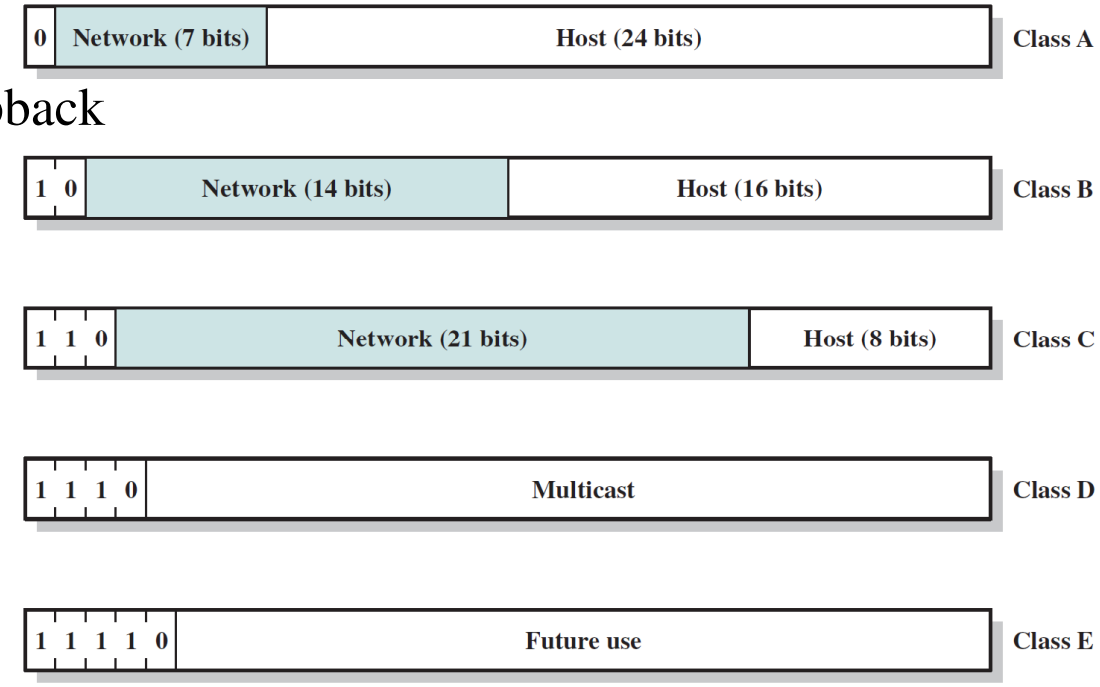
Start with binary 0

All 0 reserved

01111111 (127) reserved for loopback

Range 1.x.x.x to 126.x.x.x

All allocated



Class B

Start 10

Range 128.x.x.x to 191.x.x.x

Second Octet also included in network address

2^{14} = 16,384 class B addresses

All allocated

Class C

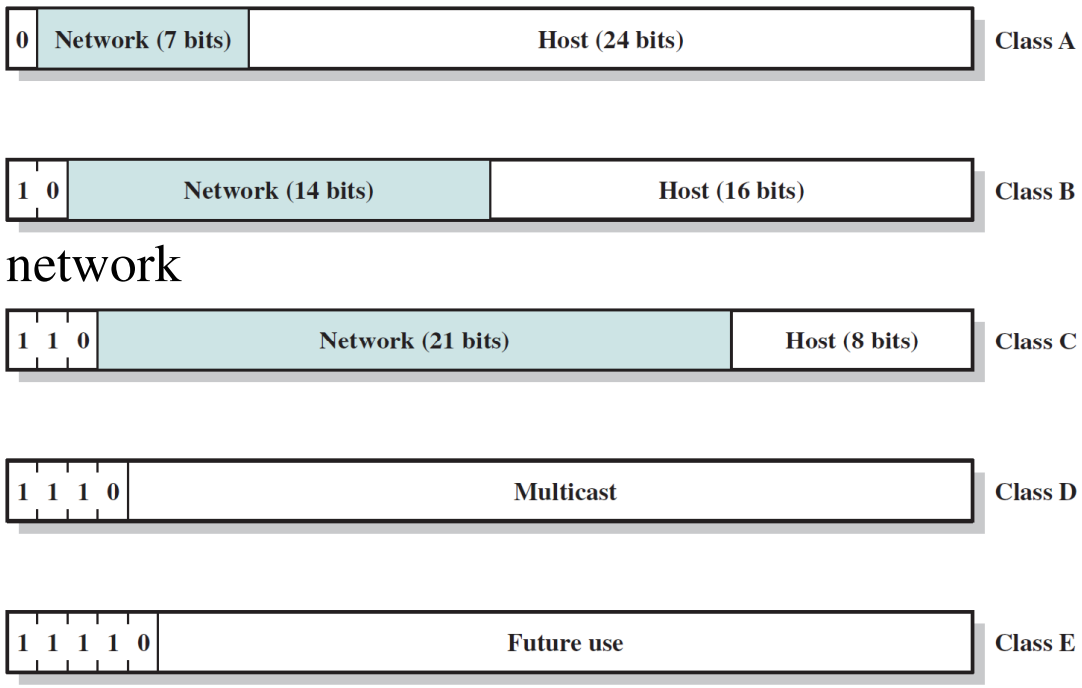
Start 110

Range 192.x.x.x to 223.x.x.x

Second and third octet also part of network address

2^{21} = 2,097,152 addresses

Nearly all allocated



Class D

Contain multicast addresses for group users

first decimal field is between 224 and 239

broadcast- 1 la toti
multicast: 1 la cativa

Class E

Rreserved for research and future developments

First decimal field between 240 and 255

Class	1st Octet Decimal Range	1st Octet High Order Bits	Network/Host ID (N=Network, H=Host)	Default Subnet Mask	
A	1 – 126*	0	N.H.H.H	255.0.0.0	/8
B	128 – 191	10	N.N.H.H	255.255.0.0	/16
C	192 – 223	110	N.N.N.H	255.255.255.0	/24
D	224 – 239	1110	Reserved for Multicasting		
E	240 – 254	1111	Experimental; used for research		

Note: Class A addresses 127.0.0.0 to 127.255.255.255 cannot be used and is reserved for loopback and diagnostic functions.

- *Private IP addresses* public - pt a trimite inafara retelei, in internet

Class A: 10.0.0.0 - 10.255.255.255 /8
Class B: 172.16.0.0 - 172.31.255.255 /16
Class C: 192.168.0.0 - 192.168.255.255 /24

	172	25	114	250
IP Address (B class)	10101100	00011001	01110010	11111010
Network Mask	11111111	11111111	00000000	00000000
	255	255	0	0
Network Address	10101100	00011001	00000000	00000000
	172	25	0	0
Broadcast Address	10101100	00011001	11111111	11111111
	172	25	255	255

AND

AND

0	0	0
0	1	0
1	0	0
1	1	1

- Network address: all host bits = 0
- Network broadcast address: all host bits = 1
- Total number of host bits: 2^{16} = 65,536
- Number of hosts: 2^{16} -2 = 65,536 – 2 = 65,534

	Binary Representation	Dotted Decimal
IP address	11000000.11100100.00010001.00111001	192.228.17.57
Subnet mask	11111111.11111111.11111111.11100000	255.255.255.224
Bitwise AND of address and mask (resultant network/subnet number)	11000000.11100100.00010001.00100000	192.228.17.32
Subnet number	11000000.11100100.00010001.001	1
Host number	00000000.00000000.00000000.00011001	25

(b) Default subnet masks

	Binary Representation	Dotted Decimal
Class A default mask	11111111.00000000.00000000.00000000	255.0.0.0
Example Class A mask	11111111.11000000.00000000.00000000	255.192.0.0
Class B default mask	11111111.11111111.00000000.00000000	255.255.0.0
Example Class B mask	11111111.11111111.11111000.00000000	255.255.248.0
Class C default mask	11111111.11111111.11111111.00000000	255. 255. 255.0
Example Class C mask	11111111.11111111.11111111.11111100	255. 255. 255.252

Subnets and Subnet Masks

Allow arbitrary complexity of internetworked LANs within organization

Insulate overall internet from growth of network numbers and routing complexity

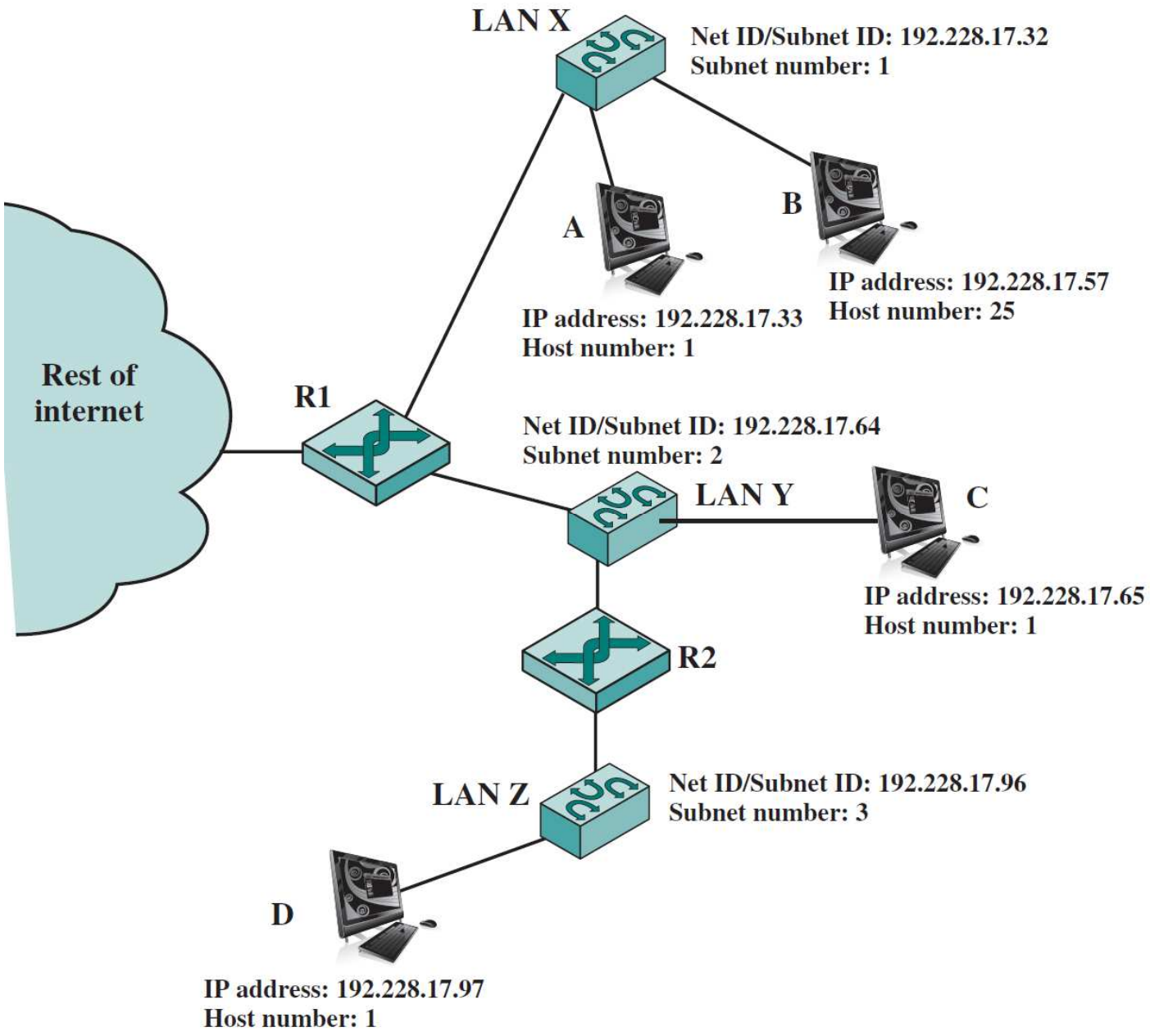
Site looks to rest of internet like single network

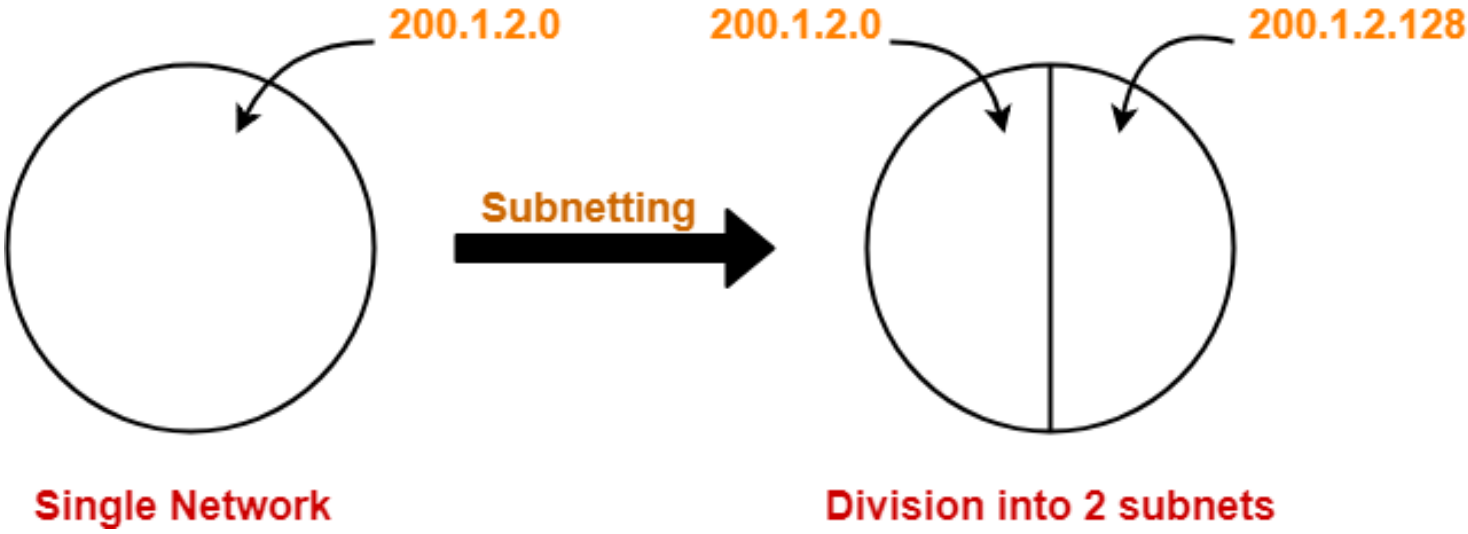
Each LAN assigned subnet number

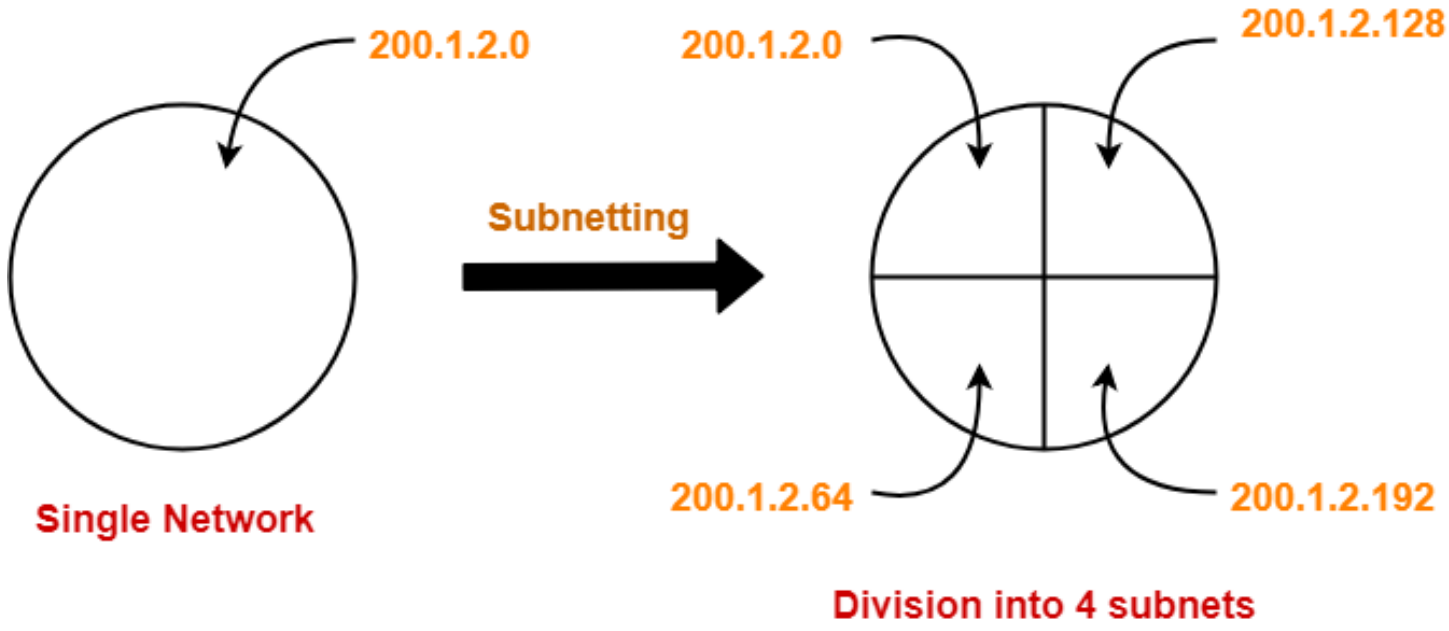
Host portion of address partitioned into subnet number and host number

Local routers route within subnetted network

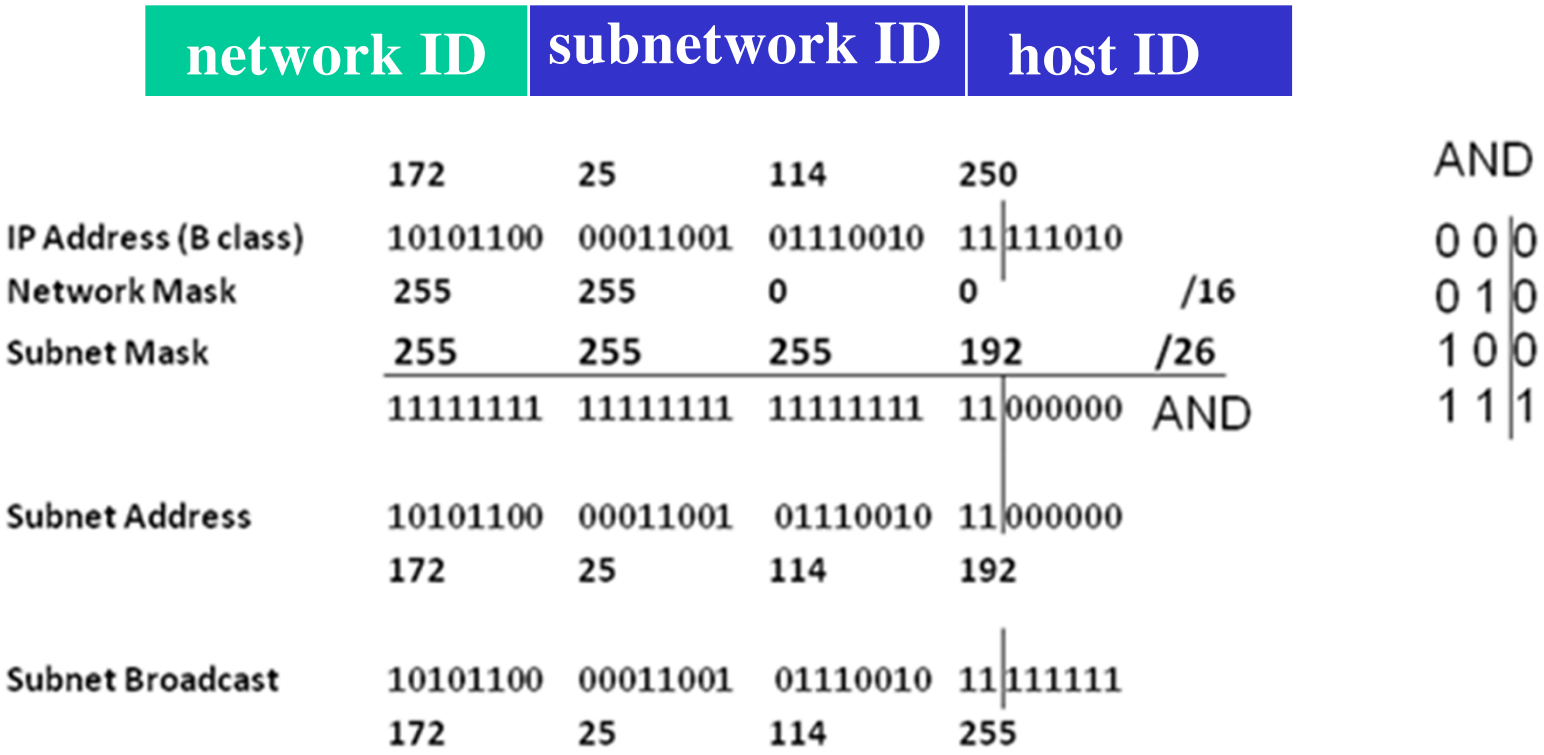
Subnet mask indicates which bits are subnet number (1s) and which are host number (0s)


LAN X
Net ID/Subnet ID: 192.228.17.32
Subnet number: 1
A
B
IP address: 192.228.17.33
Host number: 1
IP address: 192.228.17.57
Host number: 25
Rest of internet
R1
Net ID/Subnet ID: 192.228.17.64
Subnet number: 2
LAN Y
C
IP address: 192.228.17.65
Host number: 1
R2
LAN Z
Net ID/Subnet ID: 192.228.17.96
Subnet number: 3
D
IP address: 192.228.17.97
Host number: 1

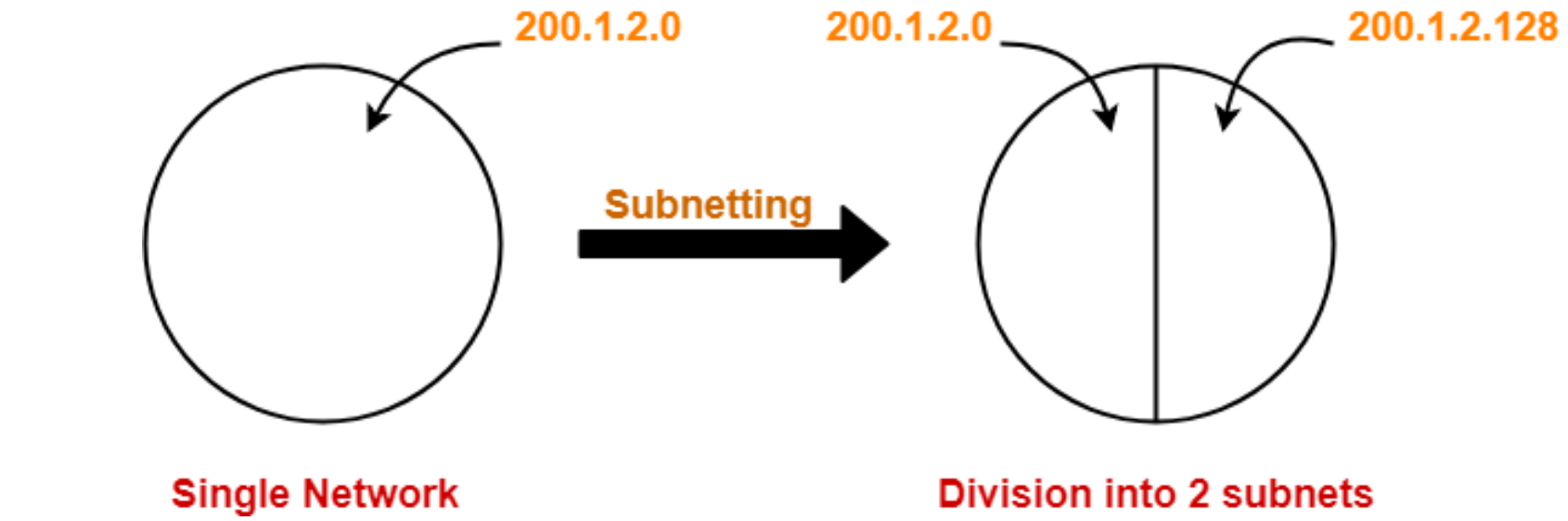

200.1.2.0
200.1.2.0
200.1.2.128
Subnetting
Single Network
Division into 2 subnets

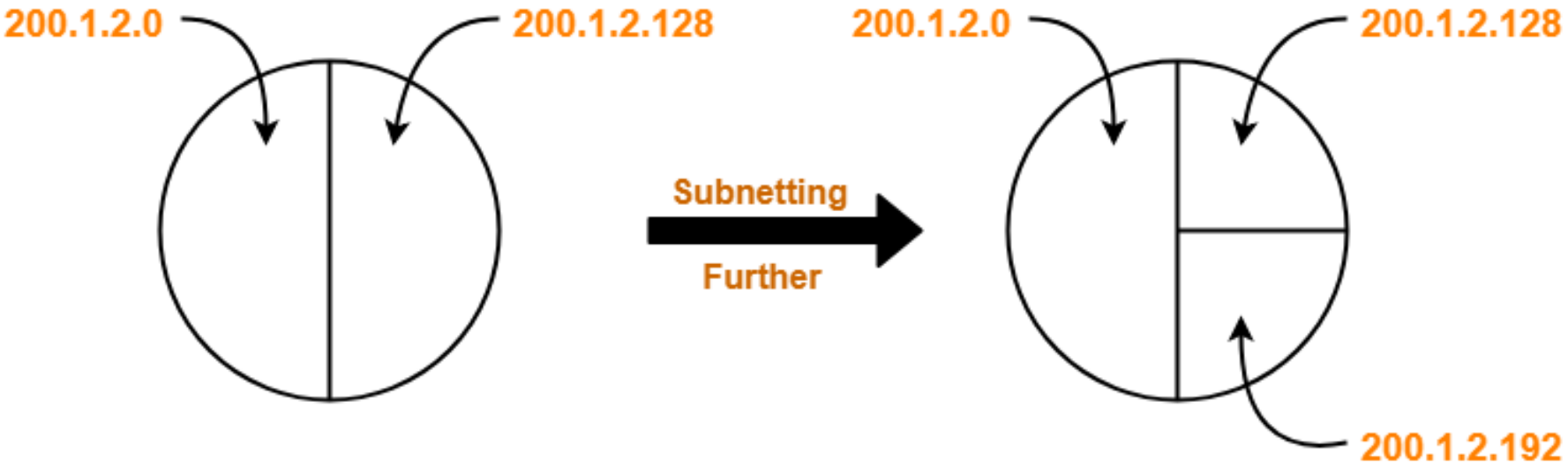

200.1.2.0
200.1.2.0
200.1.2.128
Subnetting
Single Network
200.1.2.64
200.1.2.192
Division into 4 subnets

- Creating subnets:
  - borrow bits from the host ID
  - create a new Network Mask to show the new structure of the IPv4



Total number of host bits: 2^6	
Number of hosts: 2^6 -2 = 64 – 2 = 62	First host IP on subnet: 172.25.114.193
Total number of subnet bits: 2^{10}	Last host IP on subnet: 172.25.114.254
Number of subnets: 2^{10} = 1024	


200.1.2.0
Subnetting
Single Network
200.1.2.0
200.1.2.128
Division into 2 subnets


200.1.2.0
200.1.2.128
Subnetting
Further
200.1.2.0
200.1.2.128
200.1.2.192

ICMP ex: ping

Need for appending to IP (used only for data transfer) of some **control protocols,** e.g. ICMP, ARP, RARP, BOOTP

Internet Control Message Protocol (ICMP) provides error-reporting mechanisms

RFC 792

Transfer of (control) messages from routers and hosts, to other hosts

Feedback about problems

e.g. time to live expired

ICMP packet encapsulated in IP datagram

Not a very reliable protocol, because (see next slide):

IP provides a *best-effort delivery* (not a secure one)
Internet layer can detect a small variety of errors:

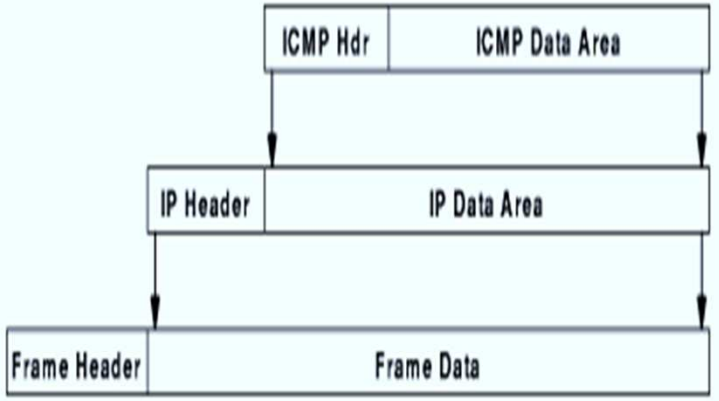
- •Checksum (header only!)
- •TTL expires
- •No route to destination network
- •Can't deliver to destination host (e.g., no ARP reply)

Internet layer discards datagrams with problems

Some - e.g., checksum error - can't trigger error messages (ICMP message)

Some errors can be reported

Remember:
ICMP encapsulated in IP (see drawing)
ICMP messages sent in response to incoming datagrams with problems
ICMP message **not** sent for ICMP message



Message type	Description
Destination unreachable	Packet could not be delivered
Time exceeded	Time to live field hit 0
Parameter problem	Invalid header field
Source quench	Choke packet
Redirect	Teach a router about geography
Echo request	Ask a machine if it is alive
Echo reply	Yes, I am alive
Timestamp request	Same as Echo request, but with timestamp
Timestamp reply	Same as Echo reply, but with timestamp

The principal ICMP message types.

ICMP and reachability

An internet host, *A*, is *reachable* from another host *B*, if datagrams can be delivered from *A* to *B*

TCP/IP *ping* program tests reachability - sends datagram from *B* to *A,* that *A echoes* back to B

Uses ICMP *echo request* and *echo reply* messages, e.g. Internet layer includes code to reply to incoming ICMP *echo request* messages

ICMP and internet routes

List of all routers on path from *A* to *B* is called the *route* from *A* to *B*

TCP/IP *traceroute* program uses UDP to non-existent port and TTL field to find route via *expanding ring* search

traceroute must accommodate varying network delays & dynamically changing routes

Sends ICMP echo messages with increasing TTL

- Router that decrements TTL to 0, sends ICMP *time exceeded* message, with router's address as source address
- First, with TTL 1, gets to first router, which discards and sends time exceeded message
- Next, with TTL 1, gets through first router to second router
- Continue until message from destination received

ICMP and path MTU (smallest accepted probe) discovery

Fragmentation should be avoided
How can source configure outgoing datagrams to avoid fragmentation?
Source determines *path MTU* - smallest network MTU (Minimum Transmission Unit) on path from source to destination
Source *probes* path using IP datagrams with *don't fragment* flag
Router responds with ICMP *fragmentation required* message
Source sends smaller probes until destination reached

ICMP and router discovery

Router can fail, causing "black-hole" or isolating host from internet

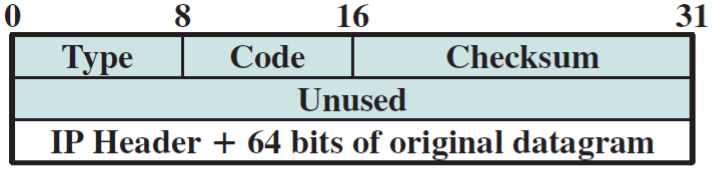
- •ICMP *router discovery* used to find new route
- •Host can broadcast request for router announcements to auto-configure default route
- •Host can broadcast request if router fails
- •Router can broadcast advertisement of existence when first connected

ICMP redirect

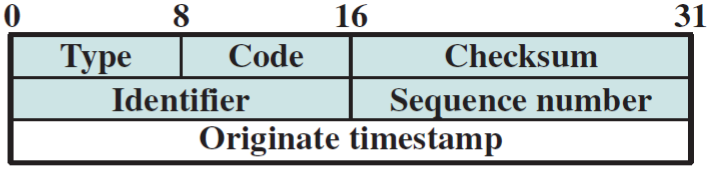
Default route may cause *extra hop*

- •Router that forwards datagram on same interface sends ICMP *redirect*
- •Host installs new route with correct router as next hop

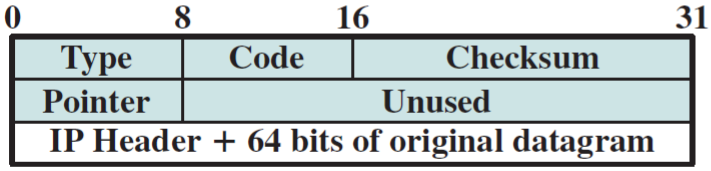
ICMP packet format



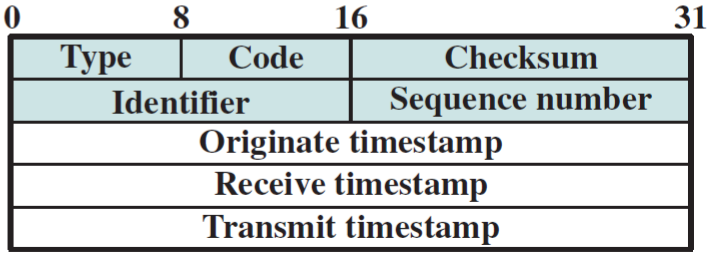
(a) Destination unreachable; time exceeded; source quench



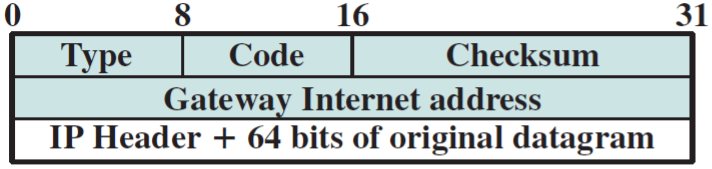
(e) Timestamp



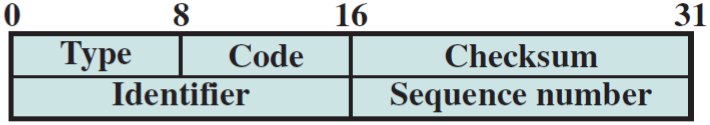
(b) Parameter problem



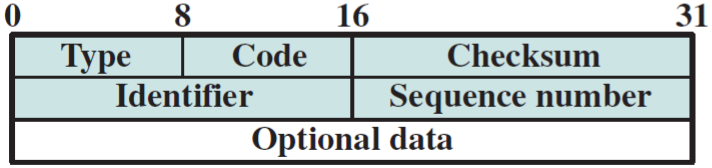
(f) Timestamp reply



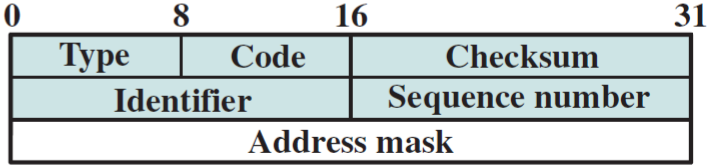
(c) Redirect



(g) Address mask request



(d) Echo, echo reply



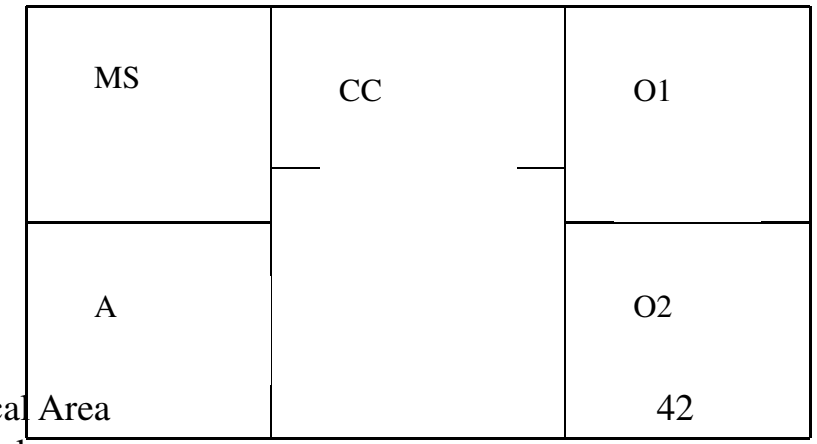
(h) Address mask reply

Proposed Problem

A company owns a building with one floor, having the following structure: **CC**-Communication Center, **O1**, **O2**-offices, **MS**-management, **A**- administration (see drawing). In room **MS** there are 2 computers, in room **A** are 5 computers, and the rooms **O1**, **O2** have 10 computers each.

The company gets for use the addresses 198.188.77.64 with the netmask 255.255.255.224.

Establish the address configuration of each subnetwork and computer located at this floor, knowing that in the both rooms **O** will be **one** subnetwork, the rooms **MS** and **A** belong to **one** other subnetwork , and **another** subnet will be setup for future applications in **CC**.



A *class C network* will be implemented.
Total number of *required subnetworks* will be *3*.

Bitwise AND of address and mask values (giving the resultant network/ subnet number) will be:

198.188.77.64

So the first subnet will have address: 198.188.77.64 and may serve rooms O1 and O2, with 20 stations together, filling address scheme from 198.188.77.65 to 198.188.77.84

The second subnet will have address: 198.188.77.128 and may serve rooms A+MS, with 7 computers, with addresses from 198.188.77.129 to 198.188.77.135

The third subnet with address 198.188.77.192 may serve the rest of the rooms.