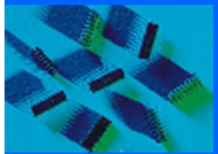


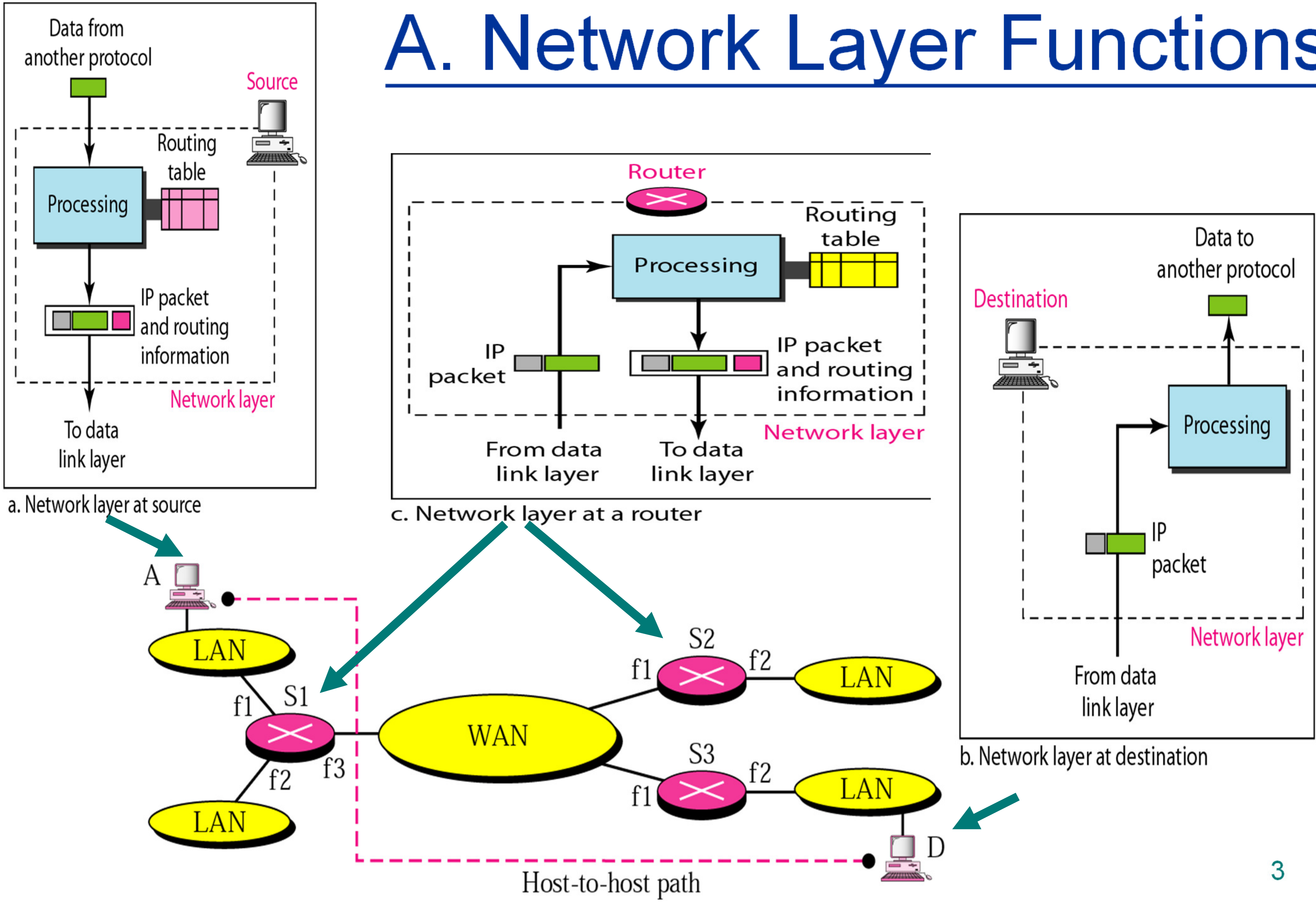
Lecture 8
Internet Protocol (IP) & Routing

Textbook: Ch.19, 20

Main Topics

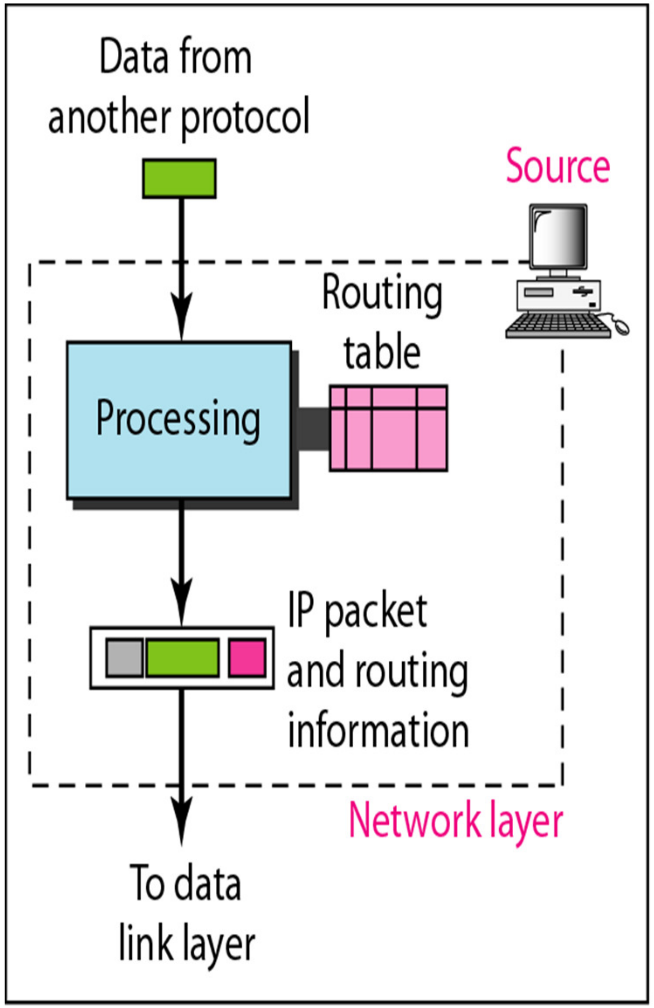
A. Network Layer Functions
B. Internet Protocol (IP)
   - TCP/IP
   - IP Datagram Format
   - IP Fragmentation
C. Routing
   - Routing Algorithms
   - Next-hop Routing Table
D. LAN Addressing
   - Address Resolution Protocol (ARP)

A. Network Layer Functions



Network Layer Functions

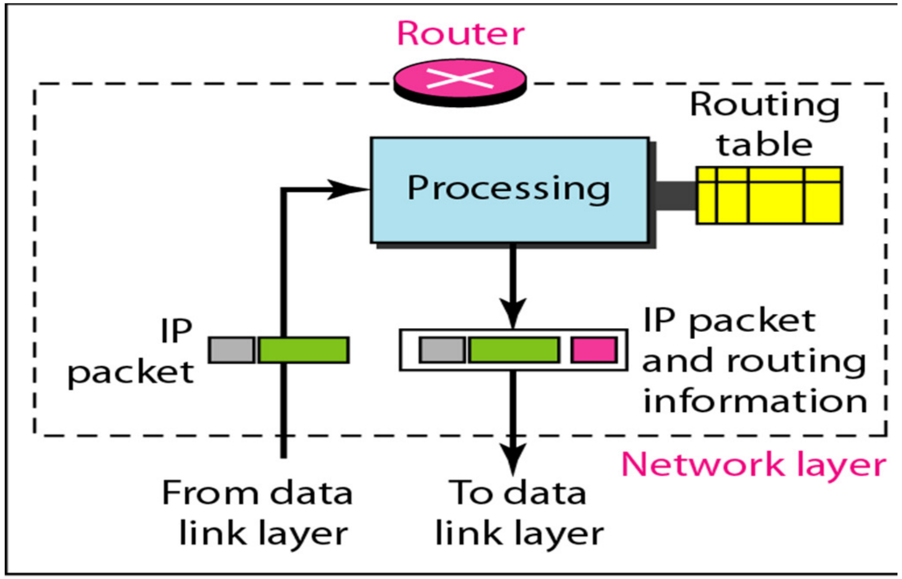
- **Source**
  - **Creating a packet** from the data coming from another protocol (usually the upper layer, TCP or UDP).
  - The **header** of packet contains the logical address (e.g. IP address) of the source and destination.
  - ***Identify*** the first **routing information**. (E.g. the address of next hop, such as default gateway.)
  - If the packet is too large, the packet is **fragmented**. (Similar to data size in the Ethernet)



a. Network layer at source

Network Layer Functions

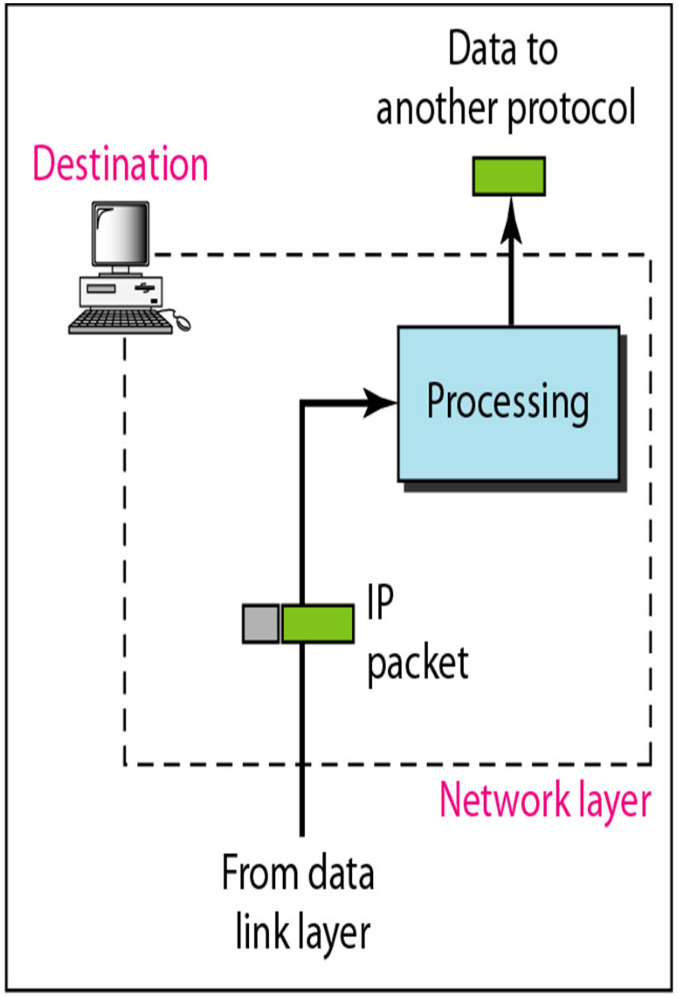
- **Router (or L3 Switch)**
  - **Routing** the packet.
  - It looks up the **routing table** find the interface that the packet should be sent.
  - The packet is sent to data link layer with the appropriate port.



c. Network layer at a router

Network Layer Functions

- **Destination**
  - Send the packet **to the upper layer**.
  - It performs **address verification** to ensure that the destination address on the packet is the same as the address of the host.
  - If the packet is a fragment, it waits until all fragments have arrived and **reassembles** them.

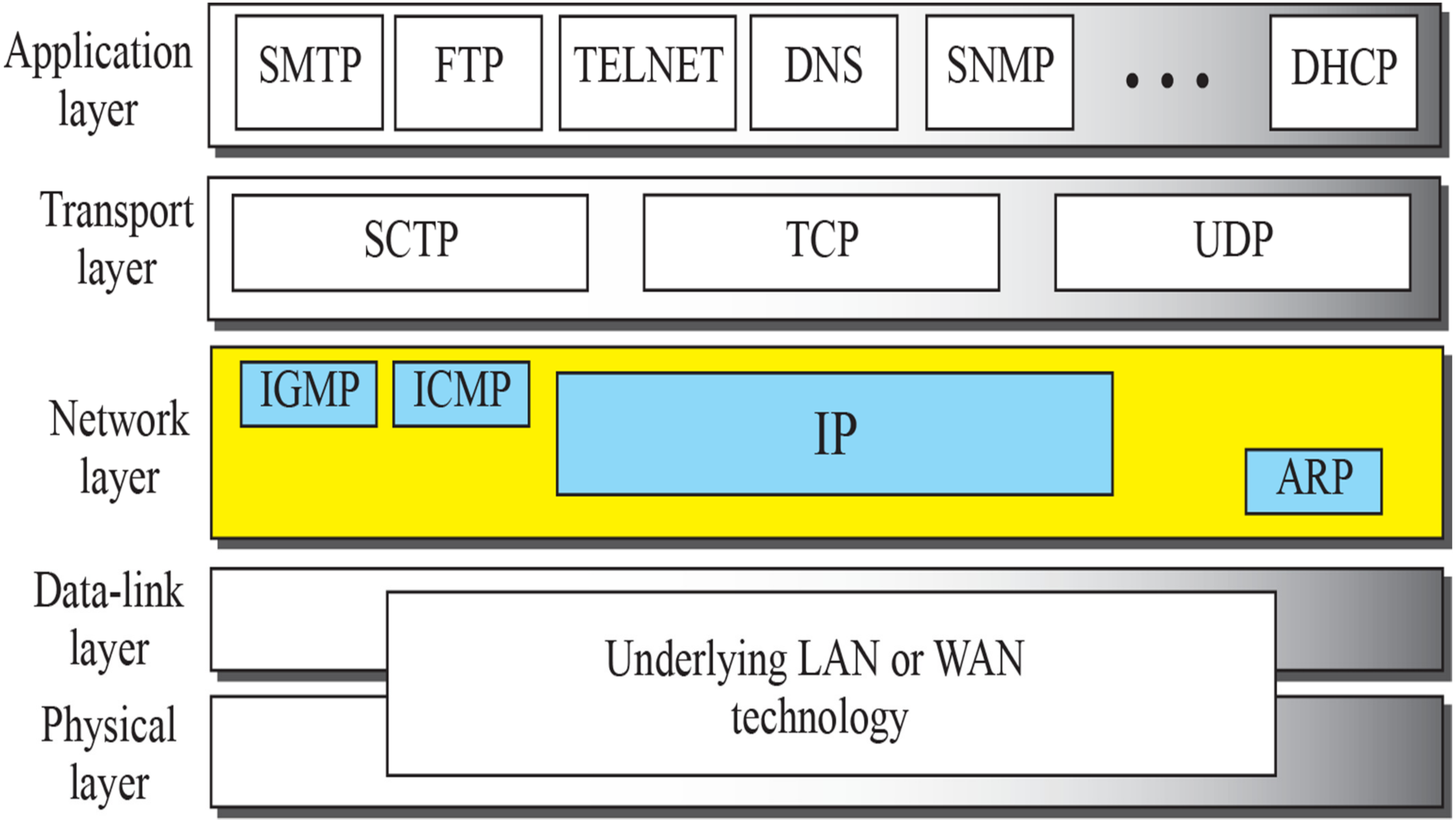


b. Network layer at destination

B. TCP/IP

- **A set of protocols**, or protocol suite, that defines how all transmissions are exchanged across the Internet
- TCP/IP is a **five-layer** protocol: physical, data link, network, transport and application
- ***Transport layer***: (2 protocols/services)
  - **Transmission Control Protocol** (**TCP**) - data unit is called TCP segment (VC service)
  - **User Datagram Protocol** (**UDP**) – data unit is called user datagram (Datagram service)
- ***Network layer***: **Internet Protocol** (**IP**)
- At least 6 protocols in the application layer

TCP/IP Protocol Suite



Internet Protocol (IP)

- The Internet Protocol version 4 (IPv4) is the network layer protocol used by TCP/IP
- Provides an unreliable, connectionless datagram best-effort delivery service
  - **Unreliable**

    packet may be lost, duplicated, delayed, out of order
  - **Connectionless datagram**

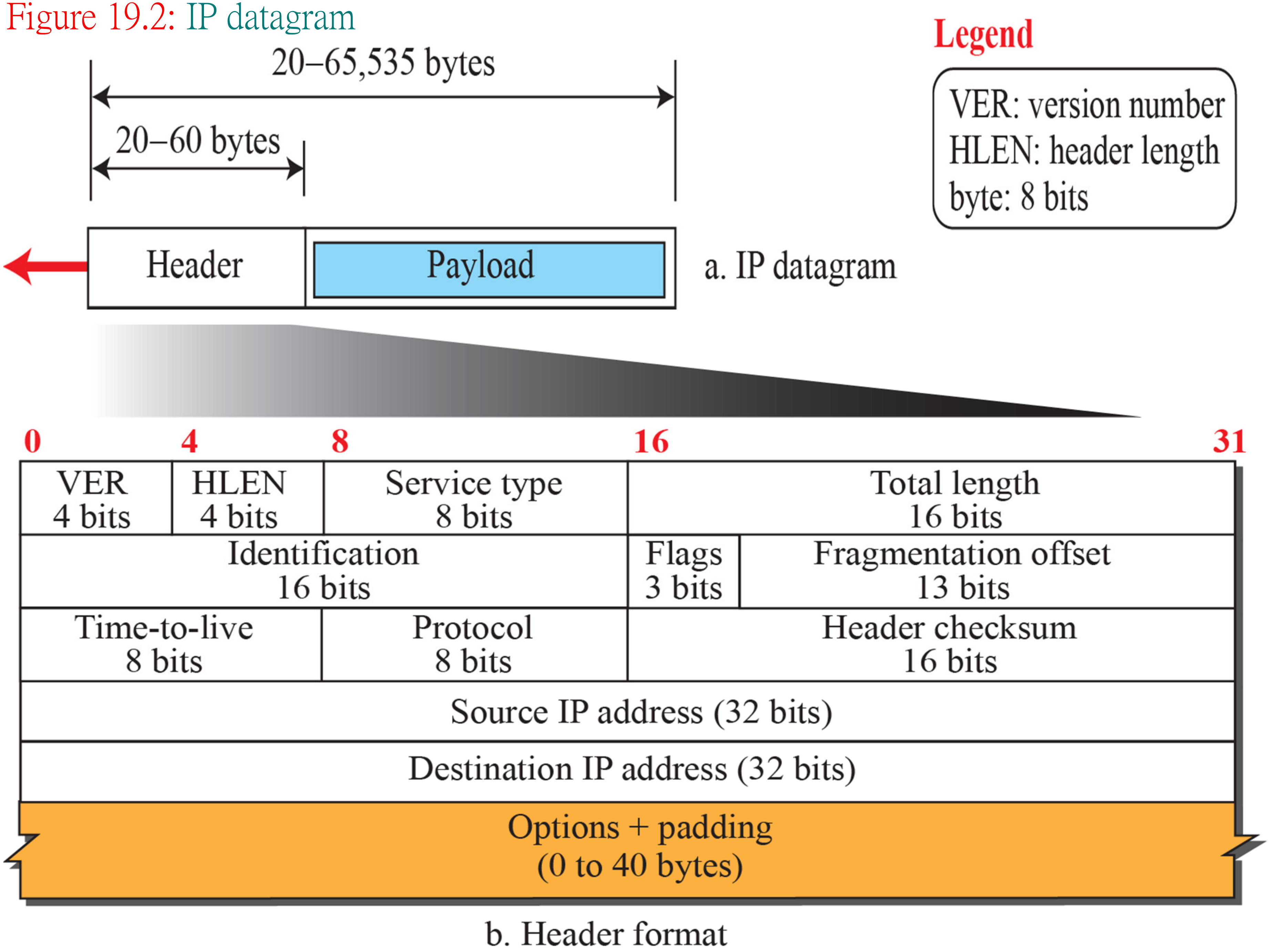
    sequence of packets (belong to the same data file) may travel along different paths
  - **Best-effort delivery**

    - Internet does not discard packets
    - unreliability arise only when resources are fully loaded or underlying networks fail

Internet Protocol (IP)

- Defines the format of the basic unit of data transfer
- The IP packet, called **IP datagram**, consists of variable-length header and data (*payload*) fields
- Setting IP address and routing tables
- Handle the routing decision and operation

Figure 19.2: IP datagram



IP Datagram Format

- **Version Number (VER)**
  - for IPv4, it is 4
- **Header Length (HLEN)**
  - min. 20 bytes to max. 60 bytes
  - (in 4-byte words) the header length in bytes is divided by 4 to get the value of this field
- **Total Length**
  - the total length (header plus data) in bytes
  - min. 20 bytes to max. 65535 bytes
- **Identification, Flags, and Fragment Offset**
  - for fragmentation (when the datagram is too long) and reassembly *(discuss later)*

IP Datagram Format

- **Time-to-live**
  - max. no. of remaining hops (or routers) allowed to visit
  - initially set to two times the path length
  - being decremented at each router
  - if it is zero, the datagram will be discarded
- **Protocol**
  - a number to indicate which upper layer protocol will get the data
- **Header Checksum**
  - just for detecting errors in the header

Example 19.1

An IPv4 packet has arrived with the first 8 bits as $(01000010)_2$ The receiver discards the packet. Why?

Solution

There is an error in this packet. The 4 leftmost bits $(0100)_2$ show the version, which is correct. The next 4 bits $(0010)_2$ show an invalid header length ($2 \times 4 = 8$). The minimum number of bytes in the header must be 20. The packet has been corrupted in transmission.

Example 19.2

In an IPv4 packet, the value of HLEN is $(1000)_2$. How many bytes of options are being carried by this packet?

Solution

The HLEN value is 8, which means the total number of bytes in the header is 8×4, or 32 bytes. The first 20 bytes are the base header, the next 12 bytes are the options.

Example 19.3 (modified)

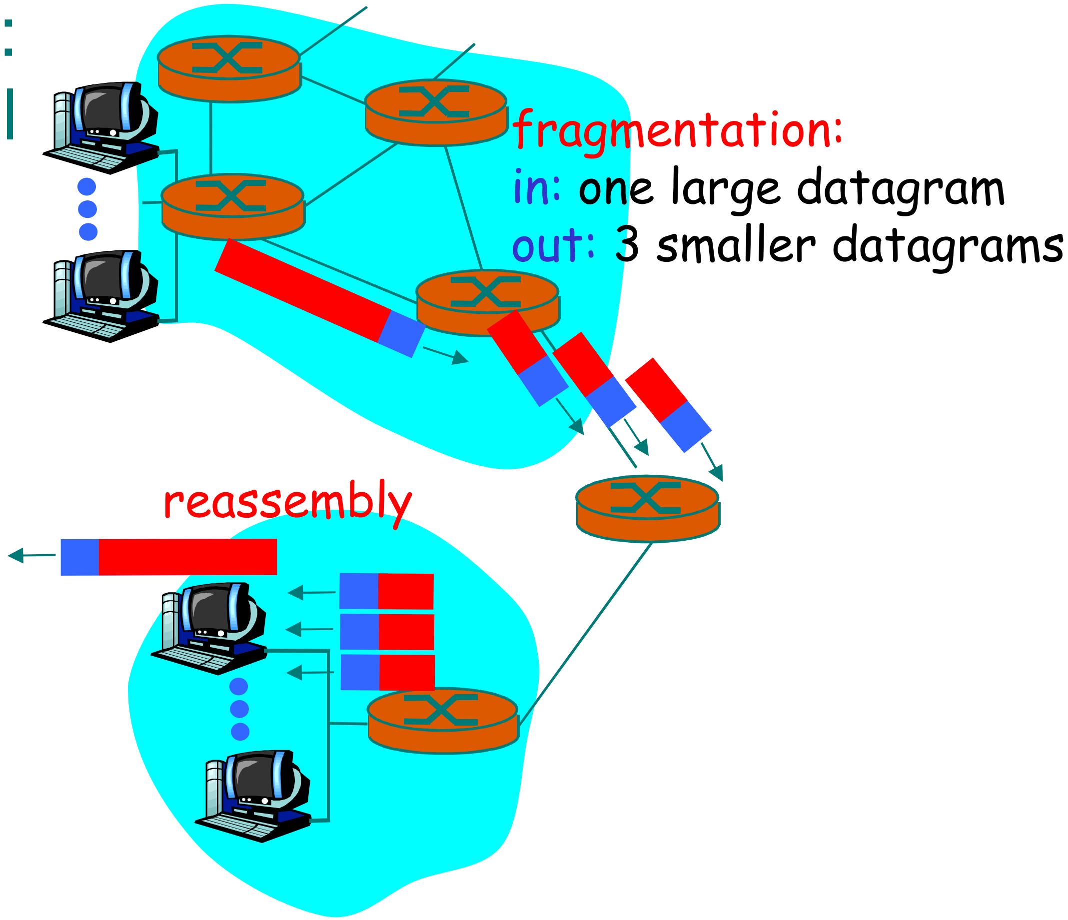
In an IPv4 packet, the value of HLEN is 5, and the value of the total length field is $(0038)_{16}$. How many bytes of data are being carried by this packet?

Solution

The HLEN value is 5, which means the total number of bytes in the header is 5×4, or 20 bytes (no options). The total length is $(0038)_{16}$ or 56 bytes, which means the packet is carrying 36 bytes of data (56 – 20).

IP Fragmentation & Reassembly

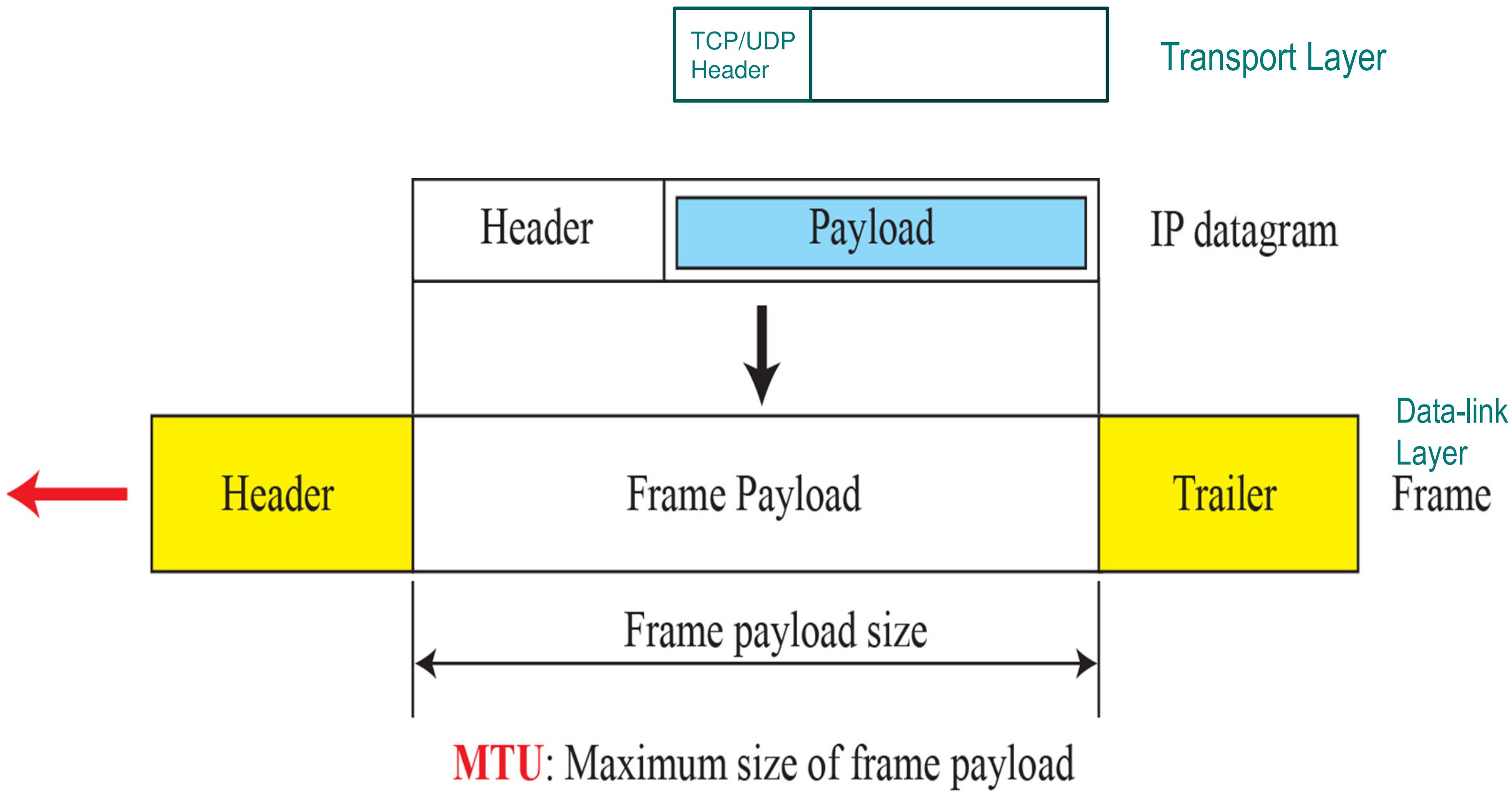
- Network links have **MTU** (maximum transfer unit) : largest possible link-level frame payload size
- Different link types, different sizes of MTUs
- The size of MTU includes both data and header of the IP datagram



IP Fragmentation & Reassembly (cont.)

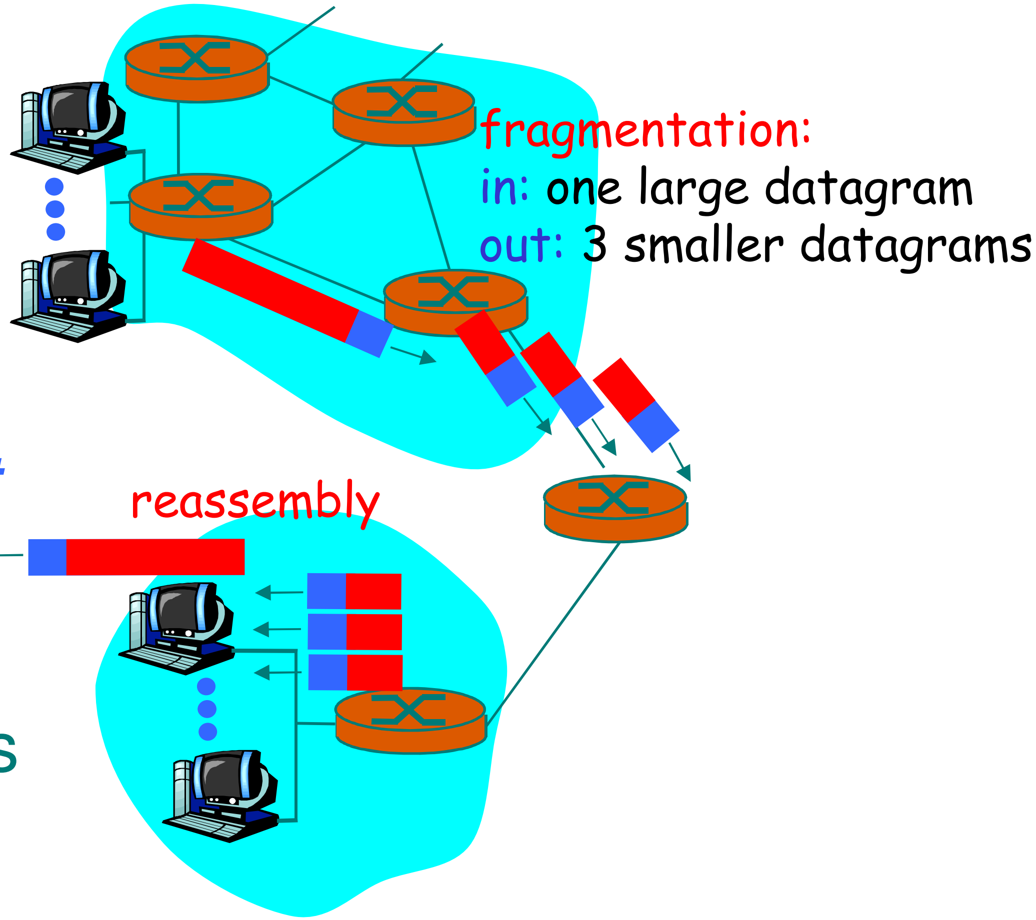
- A datagram can travel through different networks. Each router decapsulates the IP datagram from the frame it receives, processes it, and then encapsulates it in another frame.
- The format and size of the sent frame depend on the protocol used by the physical network through which the frame is going to travel.
- For example, if a router connects a LAN to a WAN, it receives a frame in the LAN format and sends a frame in the WAN format.

Figure 19.5: *Maximum transfer unit (MTU)*



IP Fragmentation & Reassembly (cont.)

- Large IP datagram is divided ("fragmented") within a network
- One datagram becomes several datagrams (fragments)
- ***"Reassembled" only at final destination***
- IP header bits used to identify related fragments and put them in order



IP Fragmentation Fields

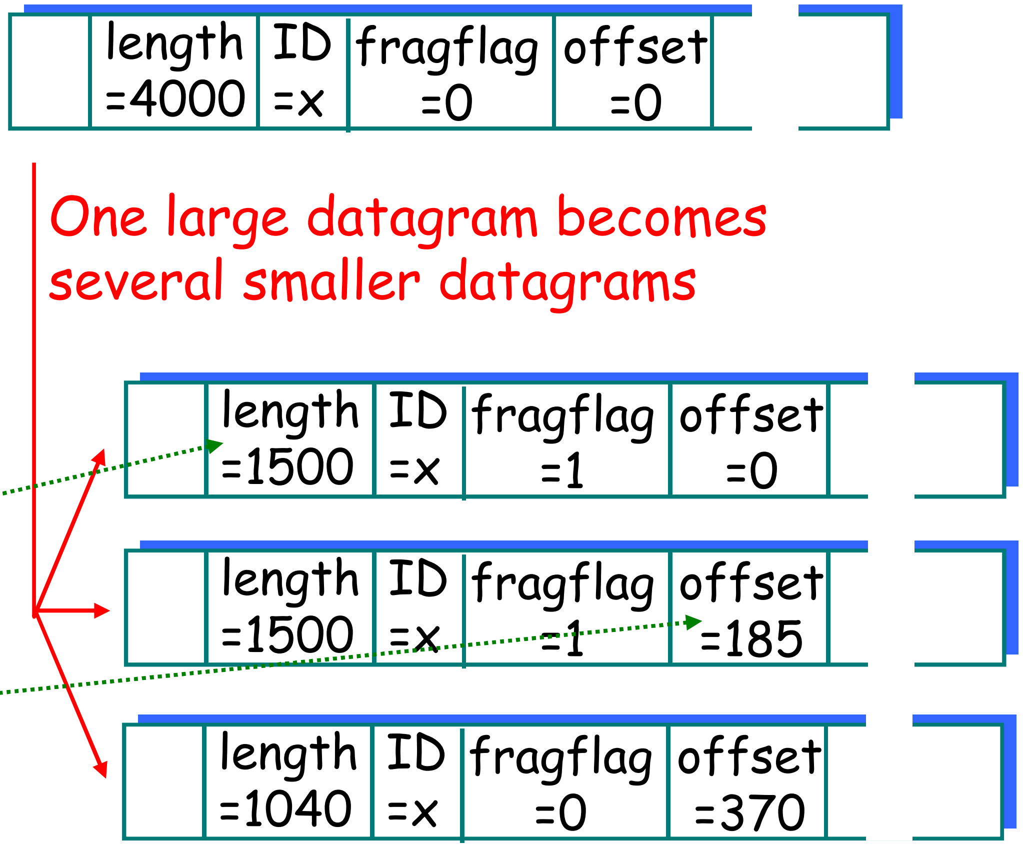
- The 16-bit **Identification** field identifies a datagram originating from the source
- When a datagram is fragmented, the ID is copied into all fragments
- The 3-bit **Flags** field defines three flags:
  - 1st bit – not used
  - 2nd bit (D-bit), "*do not fragment*" bit - if 1, means must not fragment; if 0, means can be fragmented if needed
  - 3rd bit (M-bit), "*more fragment*" bit – if 1, means more fragments after this one; if 0, means this is the last or only fragment
- The 13-bit **Fragmentation Offset** field shows the offset (position) of this fragment in the original datagram ***in units of 8 bytes***

IP Fragmentation

Example

- *Assume no optional fields in header*
- 4000 byte datagram
- MTU = 1500 bytes
- Data size =4000 - 20 =3980
- = 1480 + 1480 + 1020

The fragflag is the M-bit



1480 bytes in data field

offset = 1480/8

The offset value is specified in units of ***8 bytes***

IP Fragmentation

- (M-bit, offset)
- (0, 0) =>

  no fragmentation
- (1, 0) =>

  the first fragment
- (1, > 0) =>

  a middle fragment and more to be followed
- (0, > 0) =>

  the last fragment

The fragflag is the M-bit

length	ID	fragflag	offset
=4000	=x	=0	=0

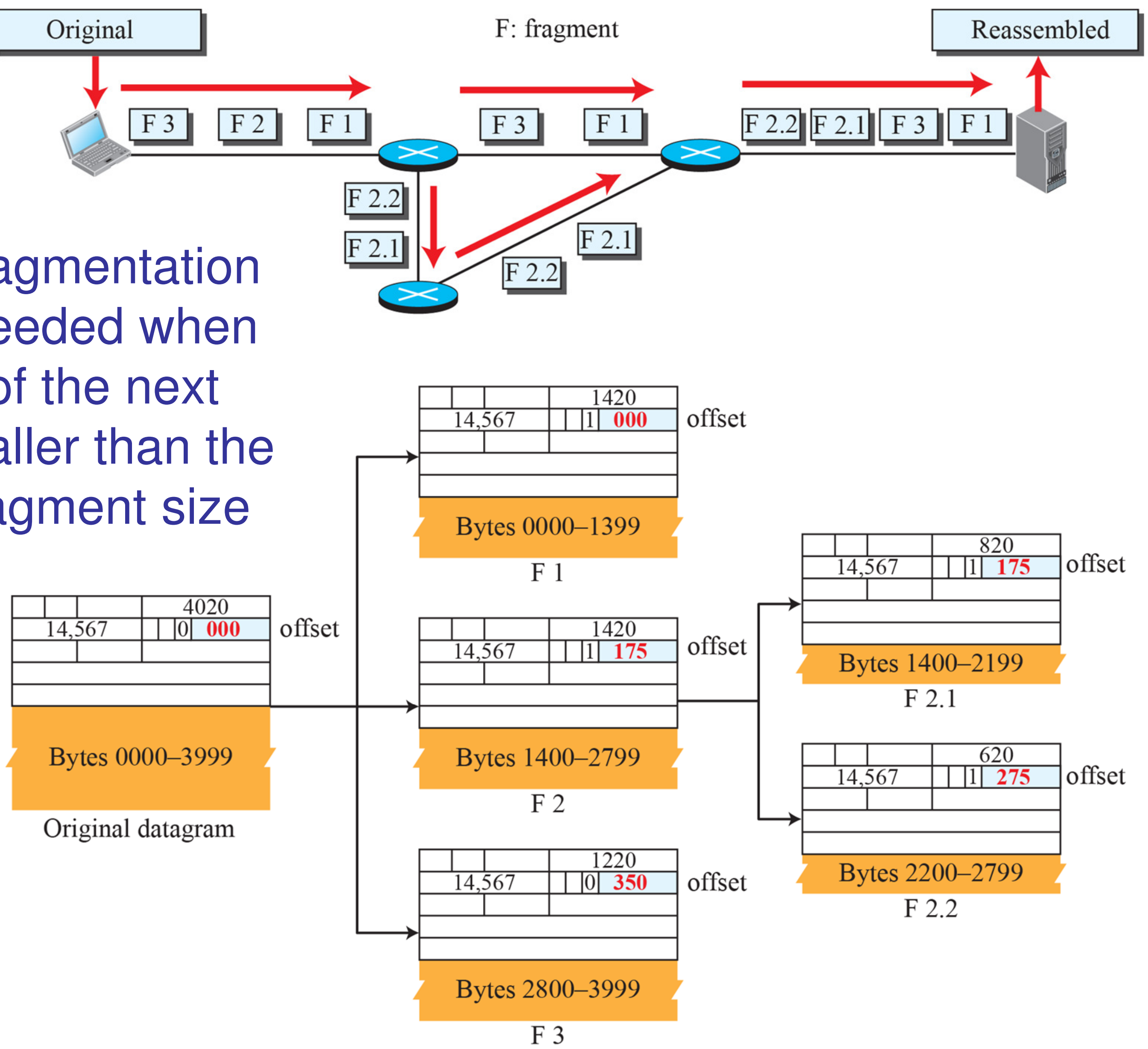
length	ID	fragflag	offset
=1500	=x	=1	=0

length	ID	fragflag	offset
=1500	=x	=1	=185

length	ID	fragflag	offset
=1040	=x	=0	=370

Figure 19.7: Detailed fragmentation example

Further fragmentation may be needed when the MTU of the next link is smaller than the current fragment size



Example 19.6

A packet has arrived with an M bit value of 0. Is this the first fragment, the last fragment, or a middle fragment? Do we know if the packet was fragmented?

Solution

If the M bit is 0, it means that there are no more fragments; the fragment is the last one.
However, we cannot say if the original packet was fragmented or not. A nonfragmented packet is considered the last fragment.

Example 19.7

A packet has arrived with an M bit value of 1. Is this the first fragment, the last fragment, or a middle fragment? Do we know if the packet was fragmented?

Solution

If the M bit is 1, it means that there is at least one more fragment.

This fragment can be the first one or a middle one, but not the last one. We don't know if it is the first one or a middle one; we need more information (the value of the fragmentation offset).

Example 19.9

A packet has arrived in which the offset value is 100. What is the number of the first byte? Do we know the number of the last byte?

Solution

To find the number of the first byte, we multiply the offset value by 8.

This means that the first byte number is 800. We cannot determine the number of the last byte unless we know the length of the data.

Example 19.10

A packet has arrived in which the offset value is 100, the value of HLEN is 5, and the value of the total length field is 100. What are the numbers of the first byte and the last byte?

Solution

The first byte number is $100 \times 8 = 800$.

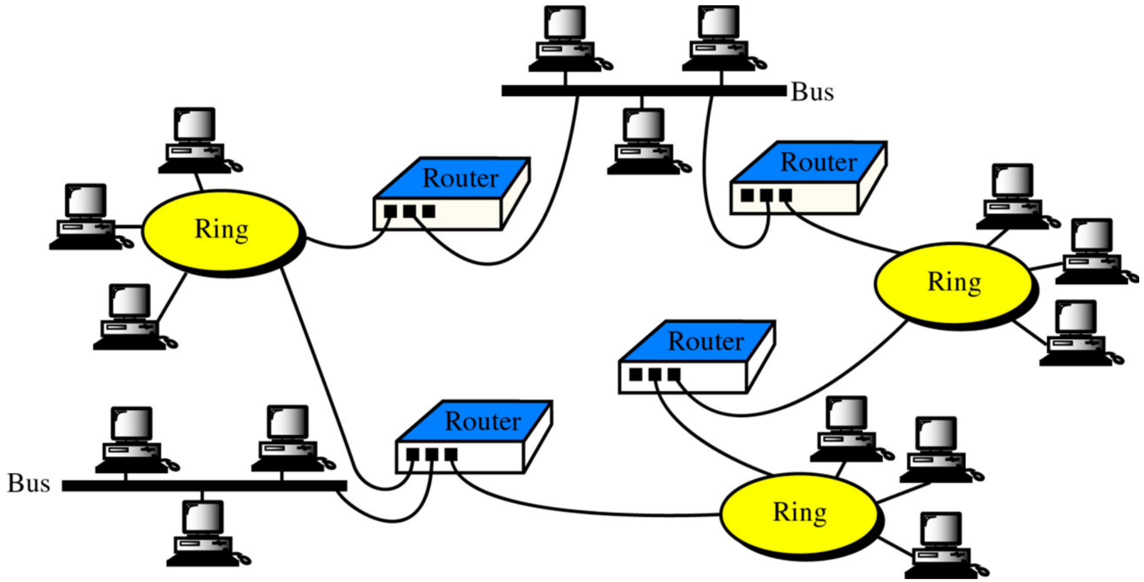
The total length is 100 bytes, and the header length is 20 bytes (5×4),

which means that there are 80 bytes in this datagram.

If the first byte number is 800, the last byte number must be 879.

Routers

- They connect networks and decide the path a packet should take
- The routing table are normally dynamic and are updated using routing protocols



C. Routing Algorithms

- Decide which **output link (path)** a packet should be transmitted in order to reach the destination
- There are many different routing algorithms:
  - It is used to exchange information among routers to build and maintain their ***routing tables***
  - Decision may base on some criteria: e.g. shortest path, least cost (Compare it with how you reach the campus from home)
  - In routing the term ***shortest*** can mean the combination of many factors including shortest, cheapest, fastest, most reliable and so on.

Routing Algorithm Classification

Global or decentralized?

Global:

- All routers have complete topology, link cost information
- e.g."link state" algorithms

Decentralized:

- Router knows neighbors' information, link costs to neighbors only
- Iterative process of computation, exchange of information with neighbors
- e.g."distance vector" algorithms

Static or dynamic?

Static:

- Routes (the routing tables) change slowly over time

Dynamic:

- Routes change more quickly
  - periodic update
  - in response to link cost changes

Next-Hop Routing

- Each host (source/destination) and router have their own (local) routing tables.
- Look at this table to find the route to the final destination (i.e. the next router to be visited)
- The table holds only the information that leads to the next hop (neighbor)
- The entries of the routing tables are generated by a selected algorithm and must be consistent with each other

Next-Hop Routing

Also see the example: SEHH2238_T8_IP_RoutingTableExample.pdf

Routing table for host S based on host-specific routing

Destination	Next Hop
A	R1
B	R1
C	R1
D	R1

Routing table for host S based on network-specific routing

Destination	Next Hop
N2	R1

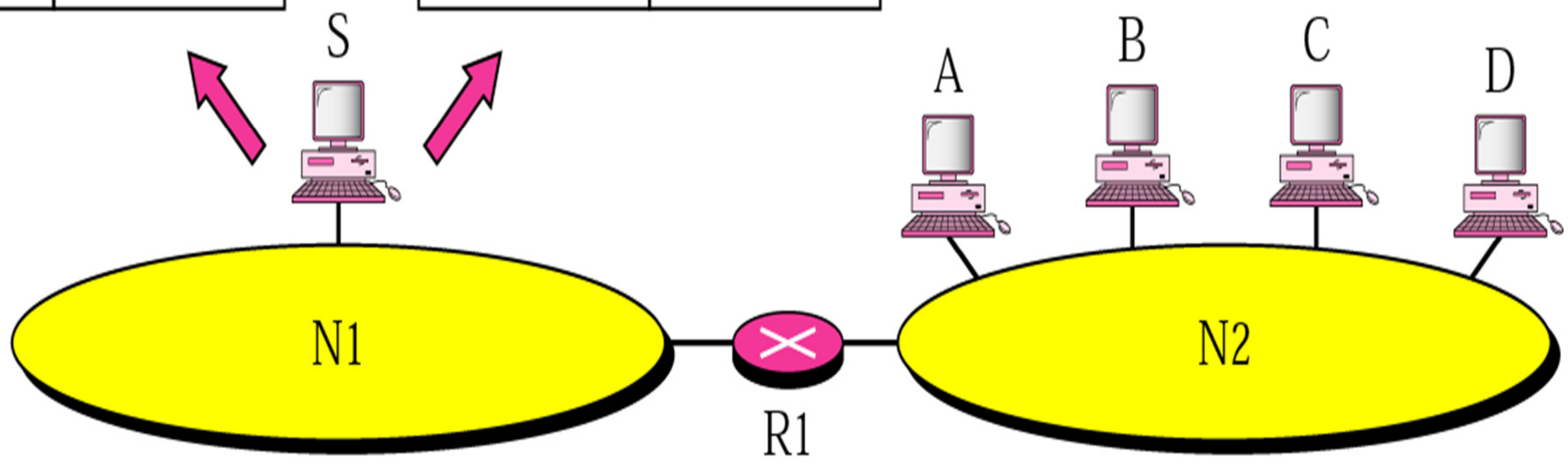
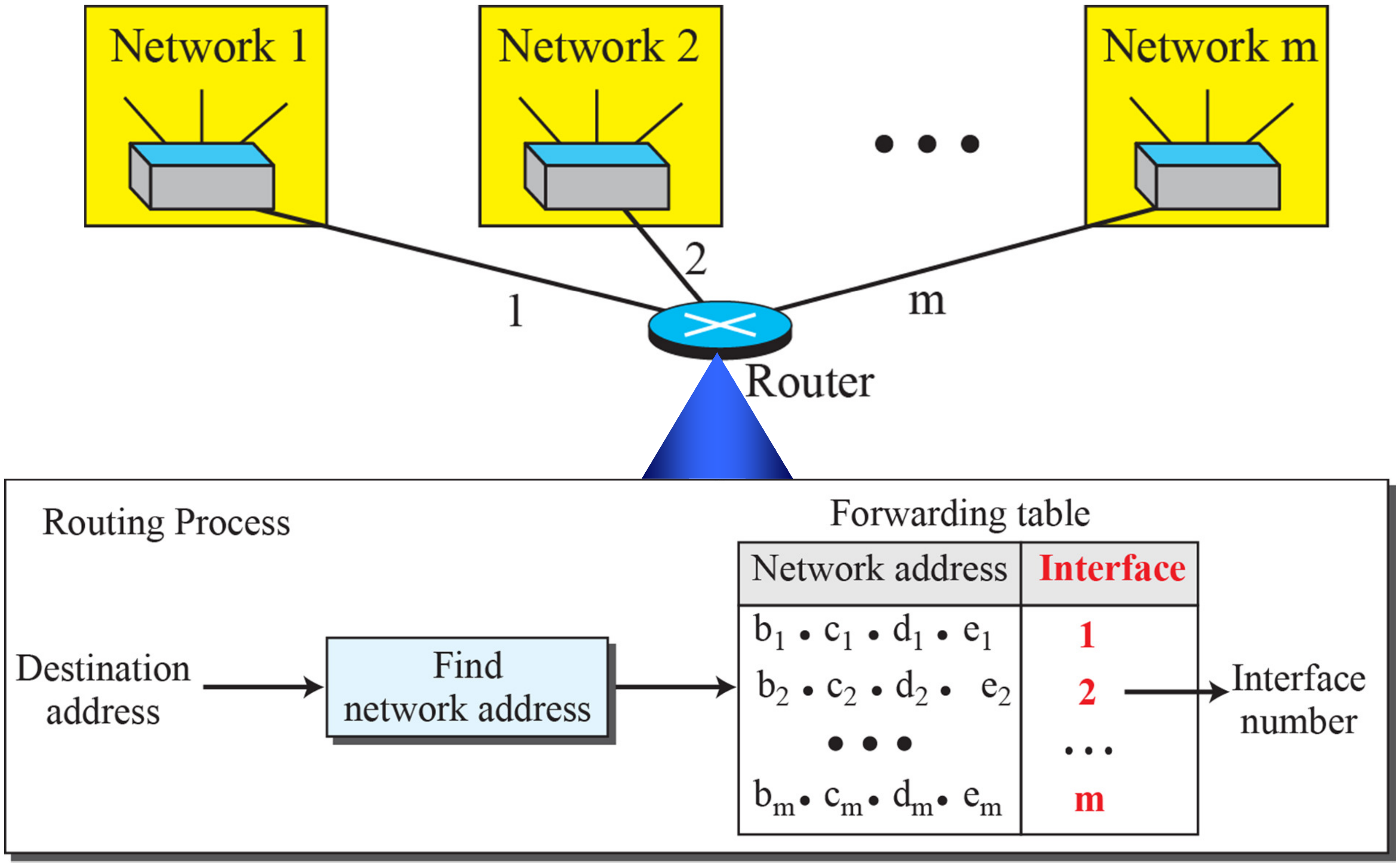


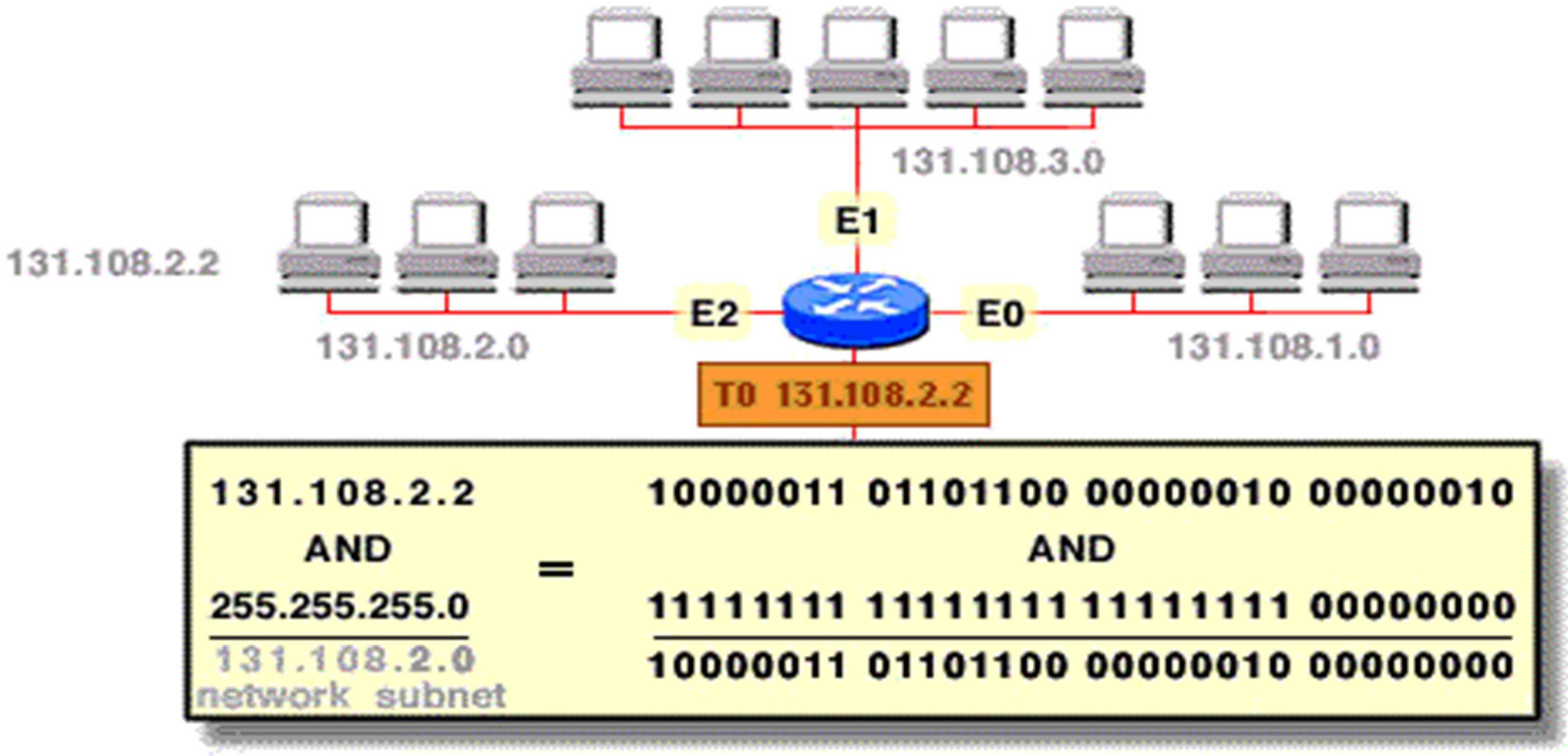
Figure 18.22: Network address

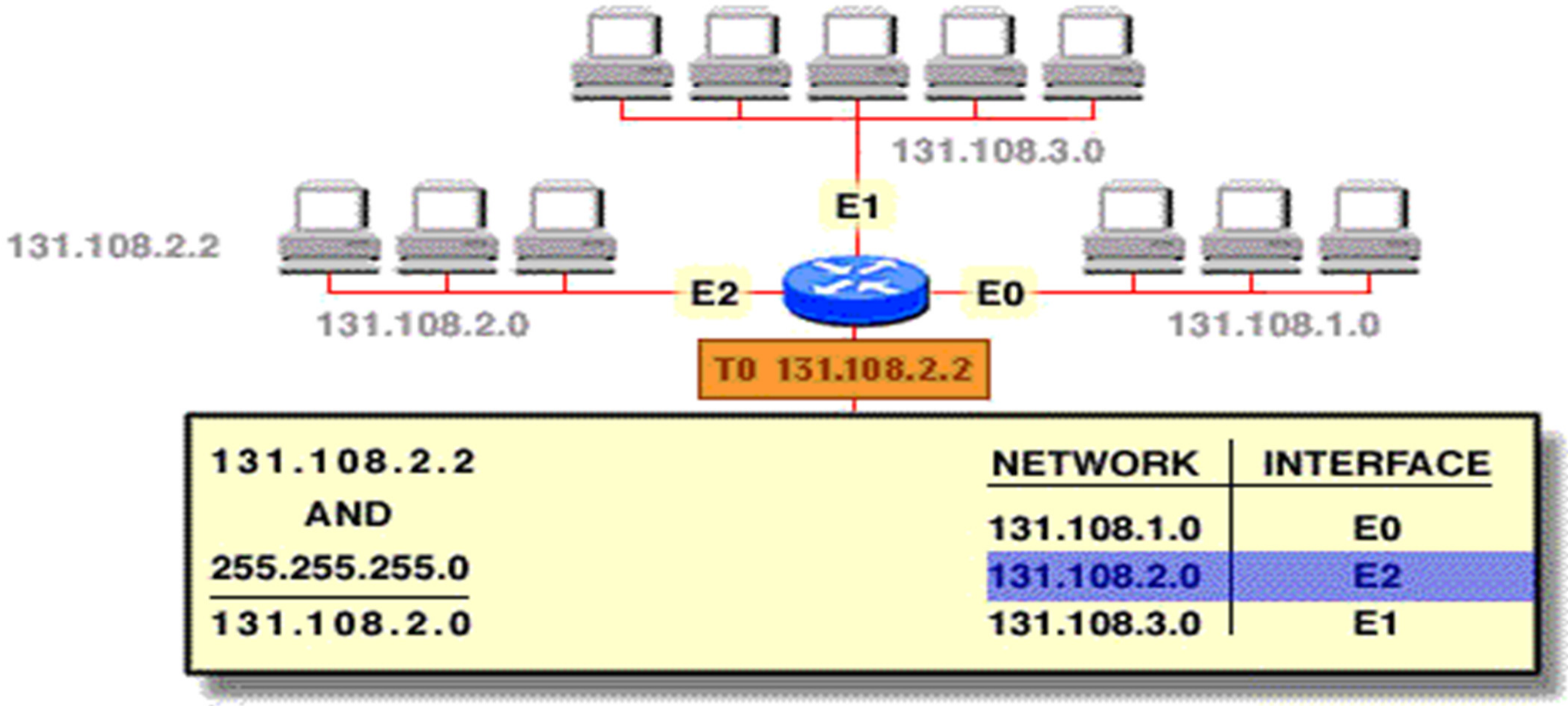


The network address is the identifier of the network and it is used by the forwarding (or routing) table in the Internet

Routing Example

- Assume that a station wants to send data with destination IP address of 131.108.2.2
- The data is sent out over the Internet until it reaches the router that is attached to the network
- The router in the destination network will determine which one of the subnets the data should be routed to
- The router knows that the subnet mask is 255.255.255.0


131.108.3.0
E1
131.108.2.2
E2
E0
131.108.2.0
131.108.1.0
TO 131.108.2.2
131.108.2.2
AND
255.255.255.0
131.108.2.0
network subnet
=
10000011 01101100 00000010 00000010
AND
11111111 11111111 11111111 00000000
10000011 01101100 00000010 00000000



Forwarding table
at the router

D. LAN Addresses and Address Resolution Protocol

32-bit IP address:

- *Network-layer* address
- Used to send datagram to destination IP network
- A logical address

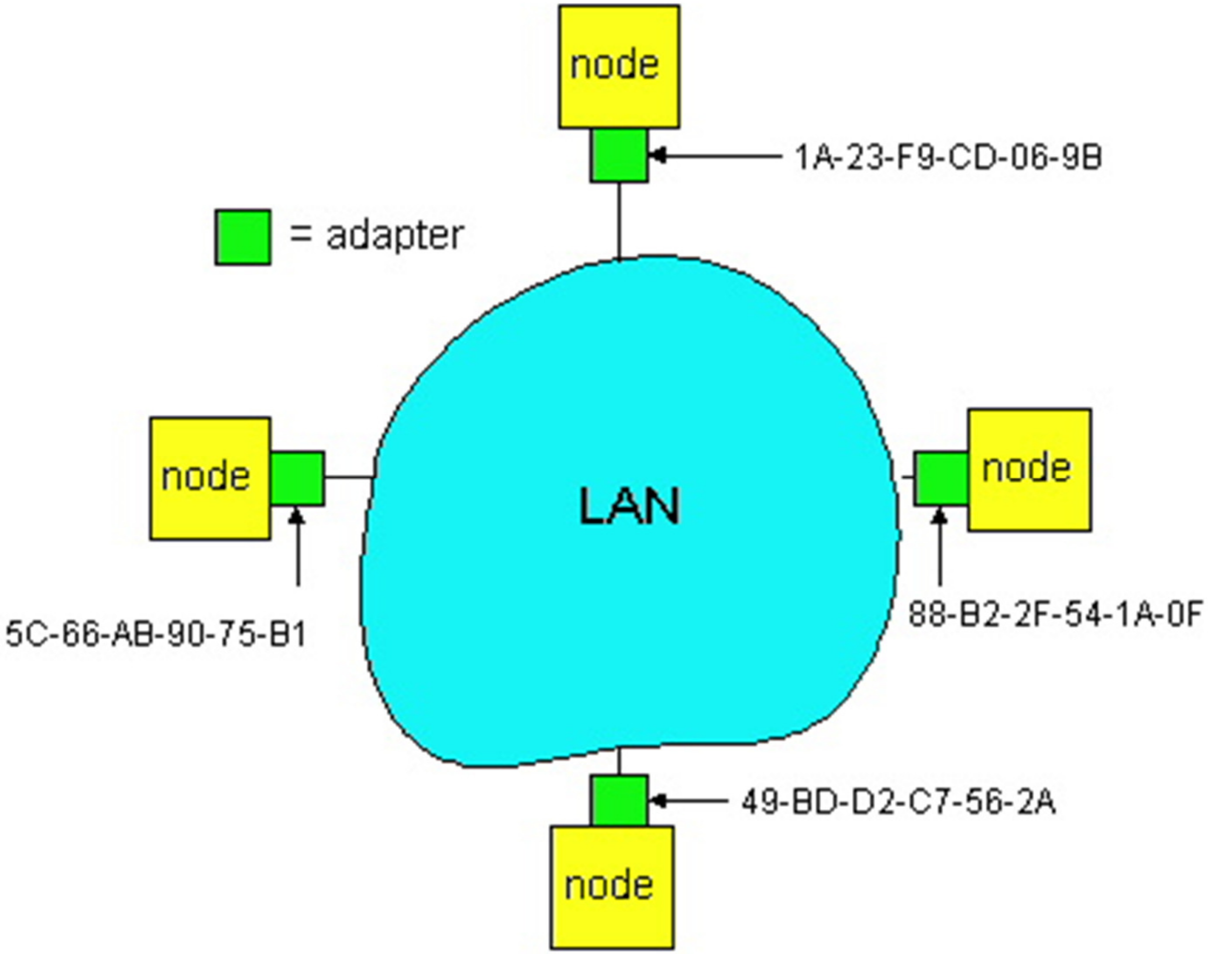
LAN (or MAC or Ethernet) address:

- A physical address
- Used to send datagram from one node to another physically-connected node (in the same network)
- **48 bit** MAC address (for most LANs) which is burned (hard-coded) in the adapter ROM

```
Physical Address. . . . . . . . . : 00-25-26-56-8D-30
```

LAN Addresses and ARP

Each adapter on LAN has unique LAN address



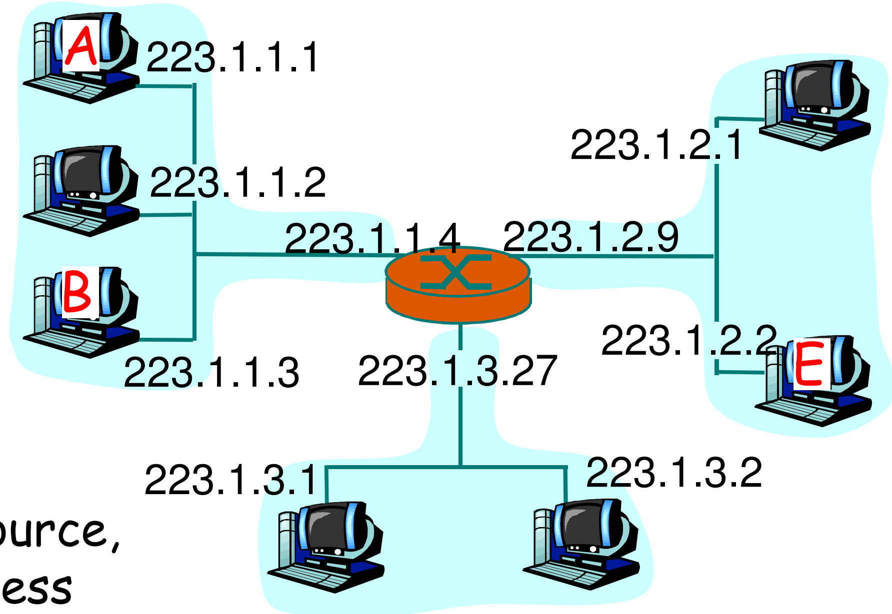
LAN Address (more)

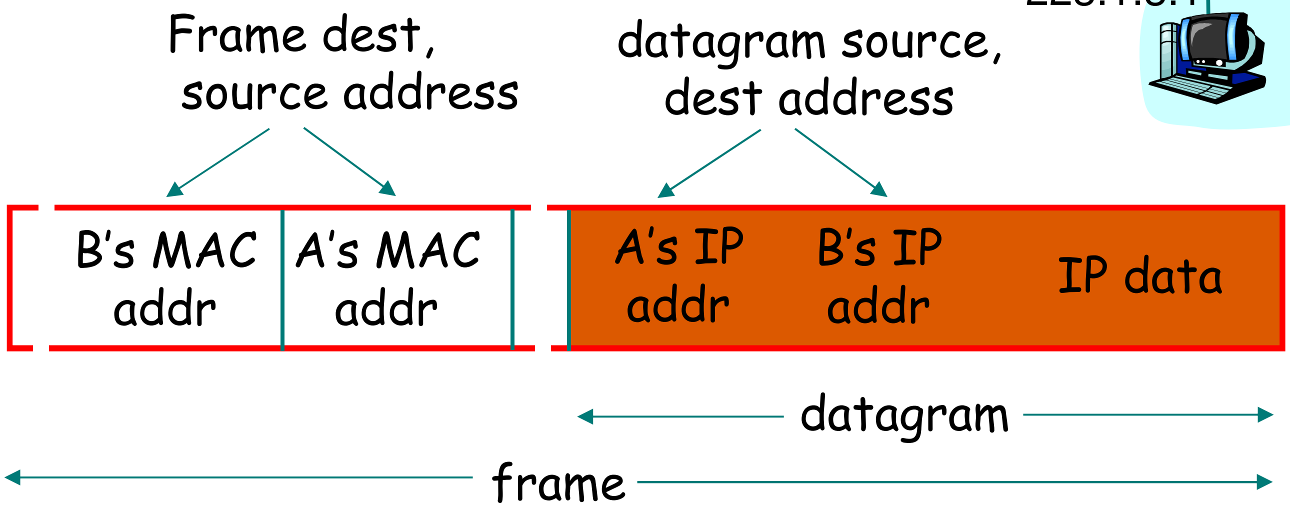
- MAC address allocation administered by IEEE
- Manufacturer buys MAC address space (to assure uniqueness)
- ***MAC address*** => **portable**
  - can move LAN card from one LAN to another LAN (the station (LAN card) still uses the same MAC address)
- ***IP address*** (structured) => **NOT portable**
  - depends on the IP of the attached network
  - when moving to another network, the station needs to change its IP address (without changing the MAC address if the same LAN card is used)

Recall earlier routing discussion

Starting at A, given IP datagram addressed to B:

- look up net. address of B, find B on same net. as A
- *link layer send datagram to B inside link-layer frame*

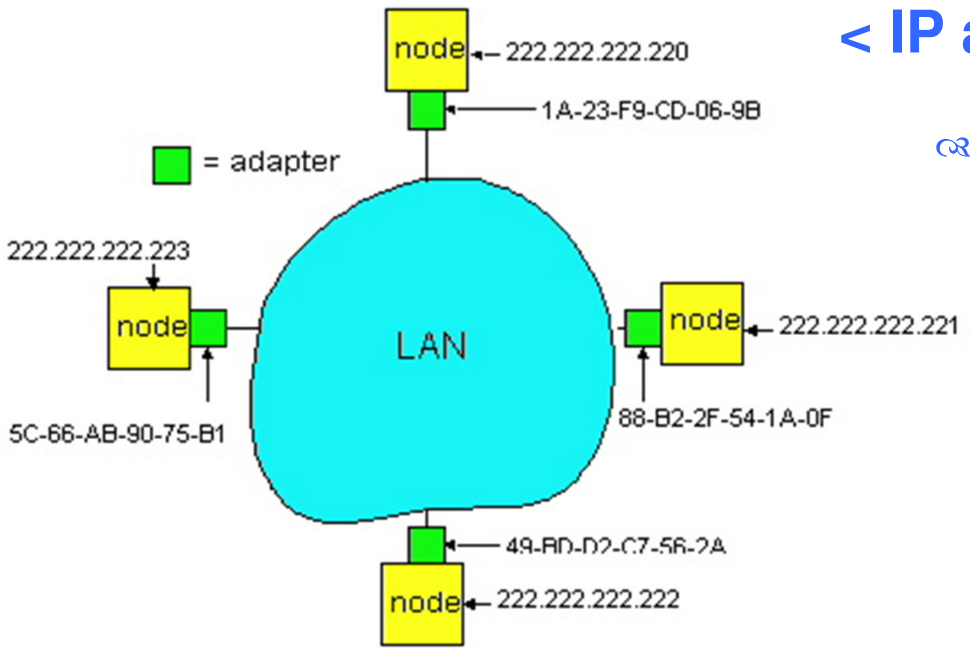




ARP: Address Resolution Protocol

How to determine MAC address of B knowing B's IP address?
- Done by ARP

- Each IP node (Host, Router) on LAN has an ARP table
- **ARP Table: IP/MAC address mappings** for some LAN nodes



< IP address; MAC address; TTL>

- TTL (Time To Live): time after which address mapping will be forgotten (typically 20 min)

ARP Protocol

- "A" wants to send datagram to "B", and A knows B's IP address.
- Suppose B's MAC address is not in A's ARP table.
- An ***ARP packet*** contains the sending and receiving IP and MAC addresses
- "A" **broadcasts** ***ARP query*** packet, containing B's IP address (as well as A's IP & MAC addresses)
  - all nodes on LAN receive ARP query

ARP Protocol (cont.)

- "B" receives ARP query packet, then replies to "A" with its (B's) MAC address
  - ***ARP reply*** packet sent to A's MAC address only

- After receiving the ARP reply packet, "A" ***caches*** (saves) IP-to-MAC address pair of "B" in its ARP table until information becomes old (times out- TTL expires)

Summary

- Network Layer Functions
- Internet Protocol (IP)
  - Connectionless, unreliable, best-effort
- Routing
  - Routing Algorithms, Next-hop Routing Table
- LAN Addressing
  - Address Resolution Protocol (ARP)
- Revision Quiz
  - http://highered.mheducation.com/sites/0073376221/student_view0/chapter19/quizzes.html