Network layer

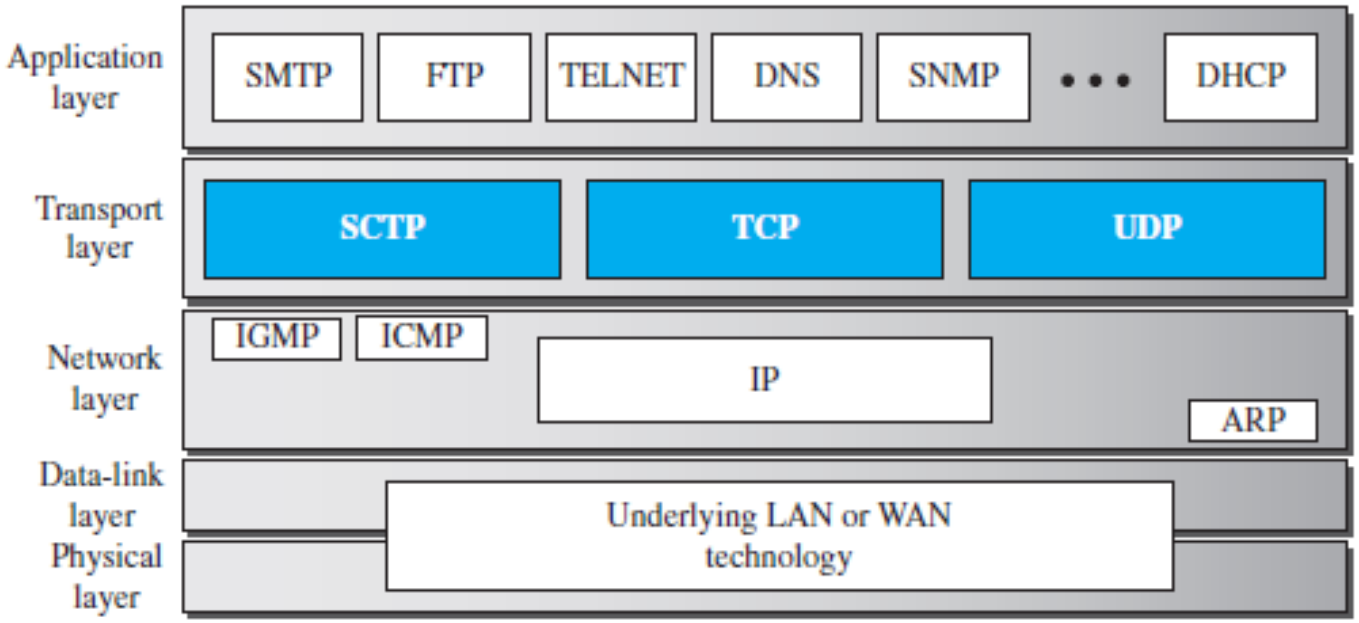
- The network layer is responsible for the source-to-destination delivery of a packet, possibly across multiple networks (links) i.e. through internet.

NETWORK-LAYER SERVICES

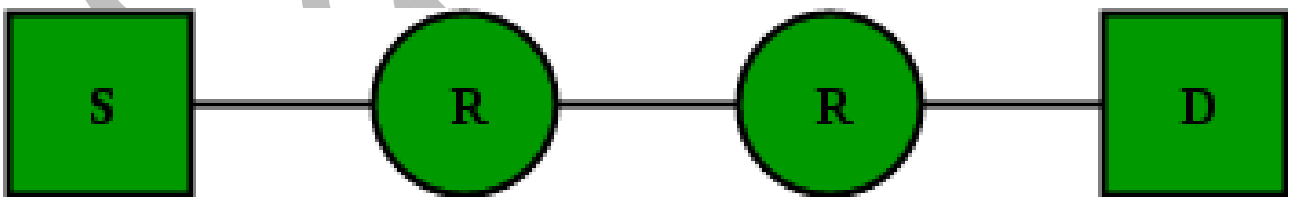
- **Logical addressing**: If a packet passes the network boundary, we need another addressing system to help distinguish the source and destination systems, the logical addresses of the sender and receiver.
-
- **Routing**: When independent networks or links are connected to create internetworks (network of networks) or a large network, the connecting devices (called routers or switches) route or switch the packets to their final destination. One of the functions of the network layer is to provide this mechanism. Network layer is responsible for routing the packet from its source to the destination.
- There is more than one route from the source to the destination.
- The network layer is responsible for finding the best one routes using routing protocols.

- **Packetizing:** Encapsulating the payload (data received from upper layer) in a network-layer packet at the source and decapsulating the payload from the network-layer packet at the destination.
- Adds a header that contains the source and destination addresses and some other information that is required by the network-layer protocol and delivers the packet to the data-link layer.
- The source is not allowed to change the content of the payload unless it is too large for delivery and needs to be fragmented.
- The routers in the path are not allowed to decapsulate the packets they received unless the packets need to be fragmented.

- **Error Control:** Error control is not directly provided in the Network layer, but checksum is added in datagram to control any corruption in header, but not in whole datagram. Although we use a protocol ICMP which provides some level of error control.
  - **Flow Control:** Network Layer does not directly provide any flow control, the job of the network layer at the receiver is so simple that it may rarely be overwhelmed.
  - **Congestion Control:** Congestion in the network layer is a situation in which too many datagrams are present in an area of the Internet. Congestion may occur if the number of datagrams sent by source computers is beyond the capacity of the network or routers. Leaky bucket, Token bucket can be used.



Q Assume that source S and destination D are connected through two intermediate routers labelled R. Determine how many times each packet has to visit the network layer and the data link layer during a transmission from S to D. **(GATE-2013) (1 Marks)**



(A) Network layer – 4 times and Data link layer – 4 times
(B) Network layer – 4 times and Data link layer – 3 times
(C) Network layer – 4 times and Data link layer – 6 times
(D) Network layer – 2 times and Data link layer – 6 times
Answer: (C)

Q Which one of the following statements is FALSE? **(GATE-2004) (1 Marks)**

(A) Packet switching leads to better utilization of bandwidth resources than circuit switching.

(B) Packet switching results in less variation in delay than circuit switching.

(C) Packet switching requires more per packet processing than circuit switching

(D) Packet switching can lead to reordering unlike in circuit switching

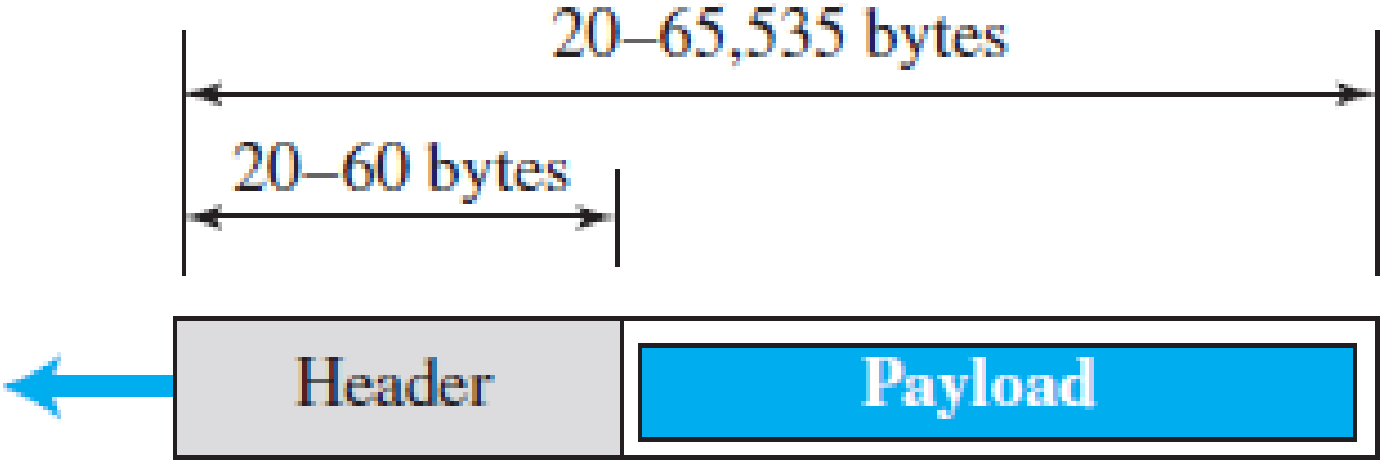
Answer: (B)

IPv4

- IPv4 is an ***unreliable connectionless datagram protocol***—a best-effort delivery service.
- The term *best-effort* means that IPv4 packets can be corrupted, maybe lost, arrive out of order, or be delayed, and may create congestion for the network.
- ***datagram*** approach means Each datagram (Packet) is handled independently, and each datagram can follow a different route to the destination.
- If reliability is important, IPv4 must be paired with a reliable protocol such as TCP, so the delivery mechanism used is TCP/IP protocols.

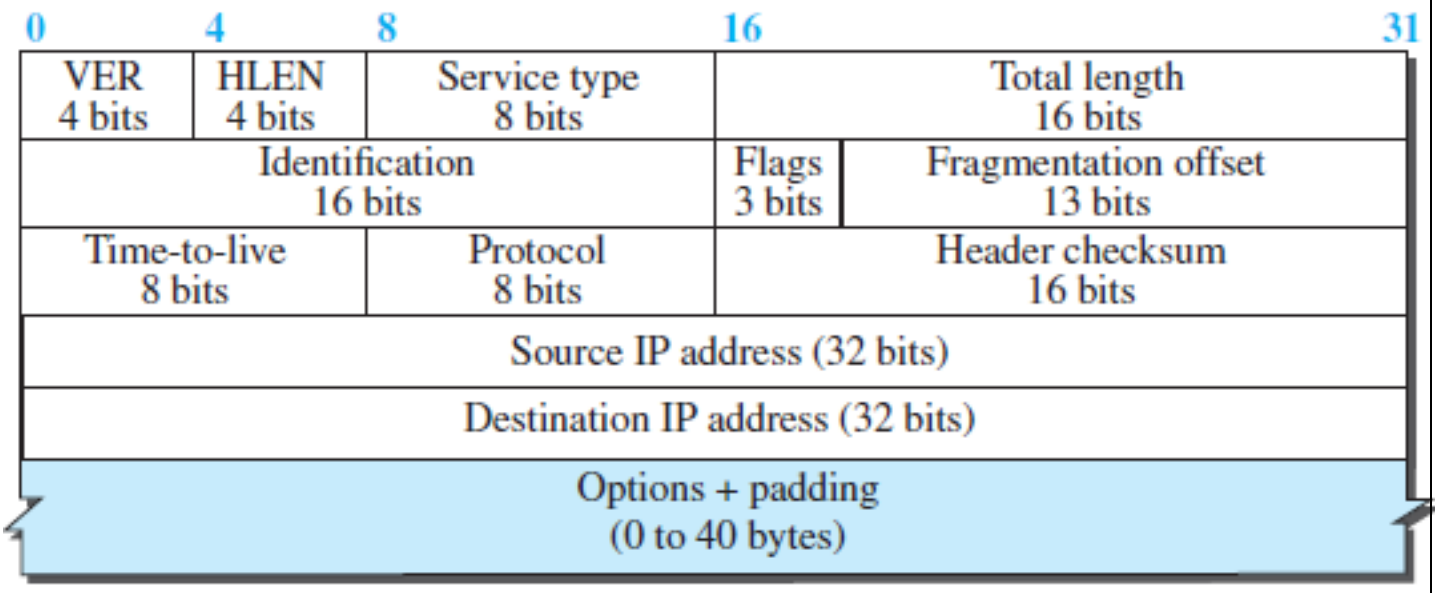
Datagram Format

- Packets used by the IP are called ***datagrams*.**
- A datagram is a variable-length packet consisting of two parts: header and payload (data).
- The header is 20 to 60 bytes in length and contains information essential to routing and delivery.



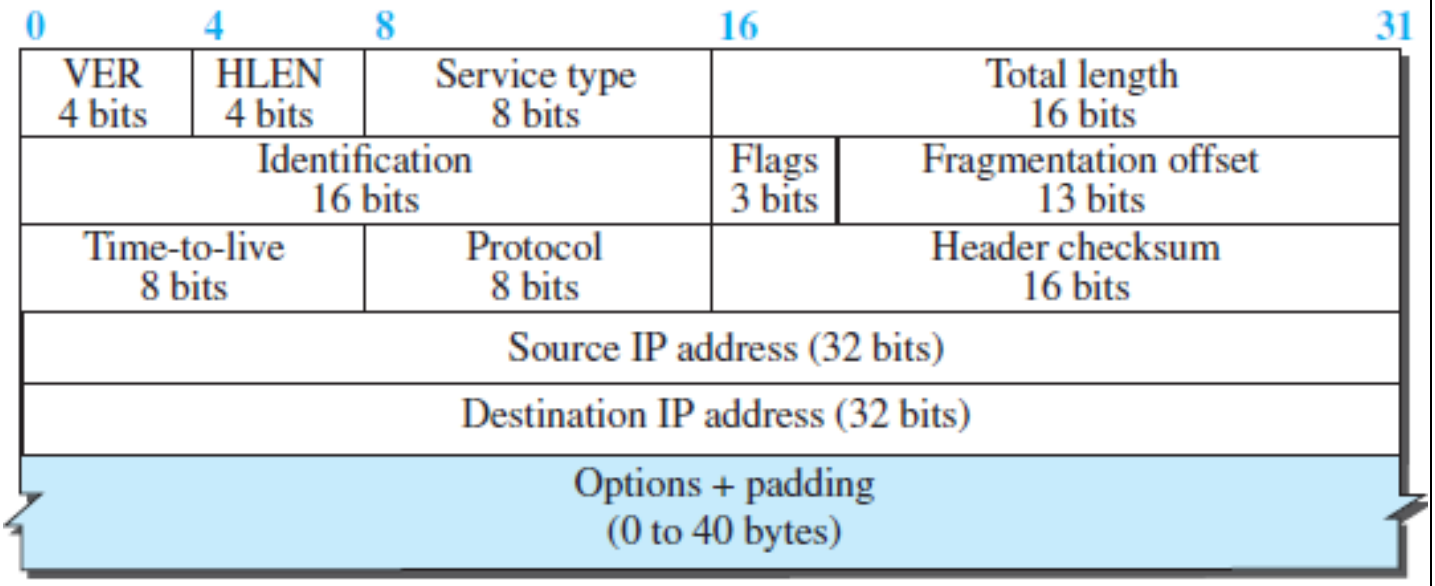
IP Datagram

Header



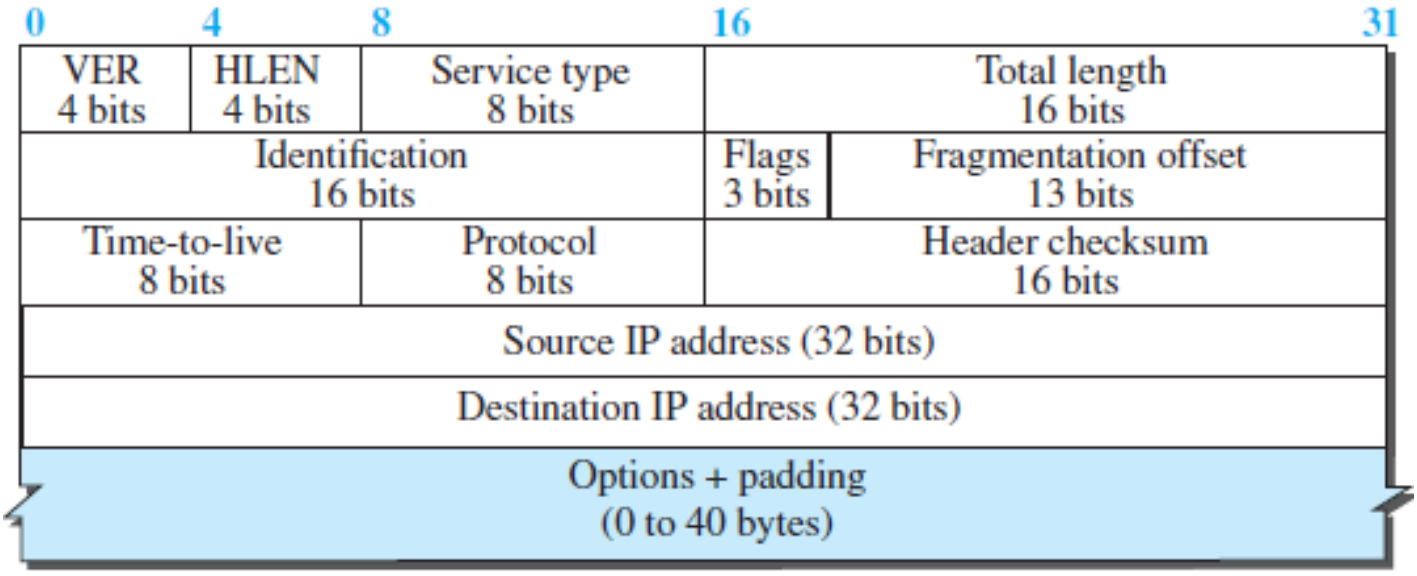
- ***Version Number.*** The 4-bit version number (VER) field defines the version of the IPv4 protocol, which, has the value of 4.

Header Length

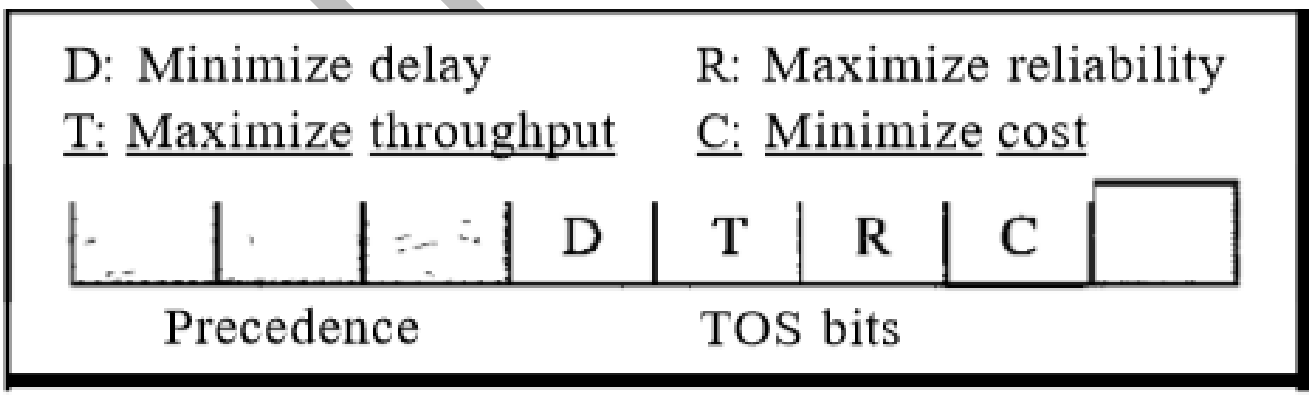


- **Header Length.** The 4-bit header length (HLEN) field defines the total length of the datagram header. The IPv4 datagram has a variable-length header.
- **Scaling Factor:**
  - To make the value of the header length (number of bytes) fit in a 4-bit header length, the total length of the header is calculated as 4-byte words.
  - The total length is divided by 4 and the value is inserted in the field.
  - The receiver needs to multiply the value of this field by 4 to find the total length.
  - Example: If header length field contains decimal value 5 (represented as 0101), then - Header length = 5 x 4 = 20 bytes
- **Point to Note**
  - The length of IP header always lies in the range of [20 bytes, 60 bytes]
  - The initial 5 rows of the IP header are always used. So, ***minimum length of IP header*** = 5 x 4 bytes = 20 bytes.
  - The size of Options field can go up to 40 bytes. So, ***maximum length of IP header*** = 20 bytes + 40 bytes = 60 bytes.
  - The range of header length field value is always [5, 15] as [20/4 = 5, 60/4 = 15]
  - The range of header length is always [20, 60].

Service Type



- ***Service Type*.** It defines how the datagram should be handled. Service type is an 8-bit field that is used for Quality of Service (QoS).
- IETF has changed the interpretation and name of this 8-bit field. This field, previously called service type, is now called differentiated services.
- Precedence is a 3-bit subfield ranging from 0 (000 in binary) to 7 (111 in binary). The precedence defines the priority of the datagram in issues such as congestion. If a router is congested and needs to discard some datagrams, those datagrams with lowest precedence are discarded first.
- TOS bits is a 4-bit subfield with each bit having a special meaning. Although a bit can be either 0 or 1, one and only one of the bits can have the value of 1 in each datagram

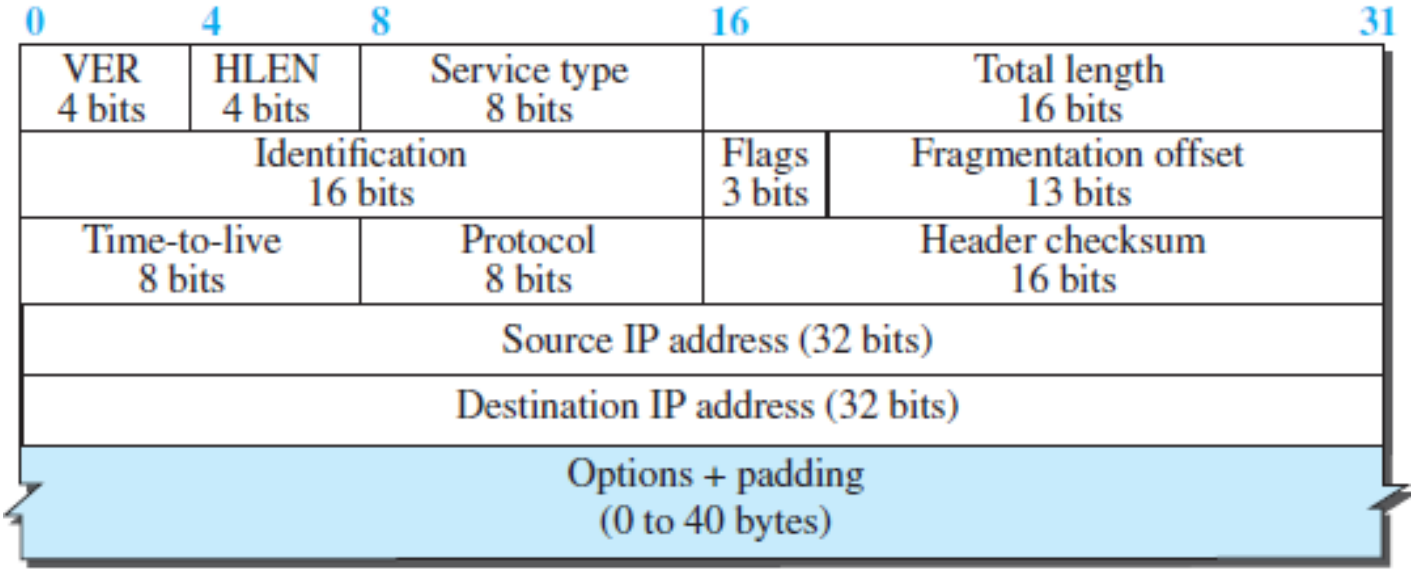


Service type

TOS Bits	*Description*
0000	Normal (default)
0001	Minimize cost
0010	Maximize reliability
0100	Maximize throughput
1000	Minimize delay

Protocol	*TOS Bits*	*Description*
ICMP	0000	Normal
BOOTP	0000	Normal
NNTP	0001	Minimize cost
IGP	0010	Maximize reliability
SNMP	0010	Maximize reliability
TELNET	1000	Minimize delay
FTP (data)	0100	Maximize throughput
FTP (control)	1000	Minimize delay
TFTP	1000	Minimize delay
SMTP (command)	1000	Minimize delay
SMTP (data)	0100	Maximize throughput
DNS (UDP query)	1000	Minimize delay
DNS (TCP query)	0000	Normal
DNS (zone)	0100	Maximize throughput

Total Length

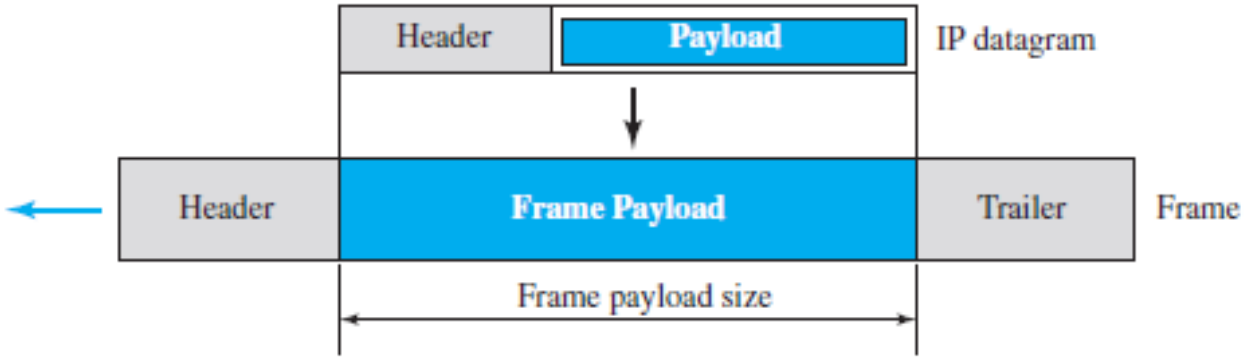


- ***Total Length.*** It defines the total length (header plus data) of the IP datagram in bytes. This field helps the receiving device to know when the packet has completely arrived.
- **Minimum total length of datagram** = 20 bytes (20 bytes header + 0 bytes data)
- **Maximum total length of datagram** = Maximum value of 16-bit word = 65535 bytes
- To find the length of the data coming from the upper layer, subtract the header length from the total length.
- ***Length of data = total length – (HLEN) × 4***

Maximum Transfer Unit (MTU)

- Each link-layer protocol has its own frame format.
- One of the features of each format is the maximum size of the payload that can be encapsulated.
- In other words, when a datagram is encapsulated in a frame, the total size of the datagram must be less than this maximum size.

Protocol	*MTU*
Hyperchannel	65,535
Token Ring (16 Mbps)	17,914
Token Ring (4 Mbps)	4,464
FDDI	4,352
Ethernet	1,500
X.25	576
PPP	296



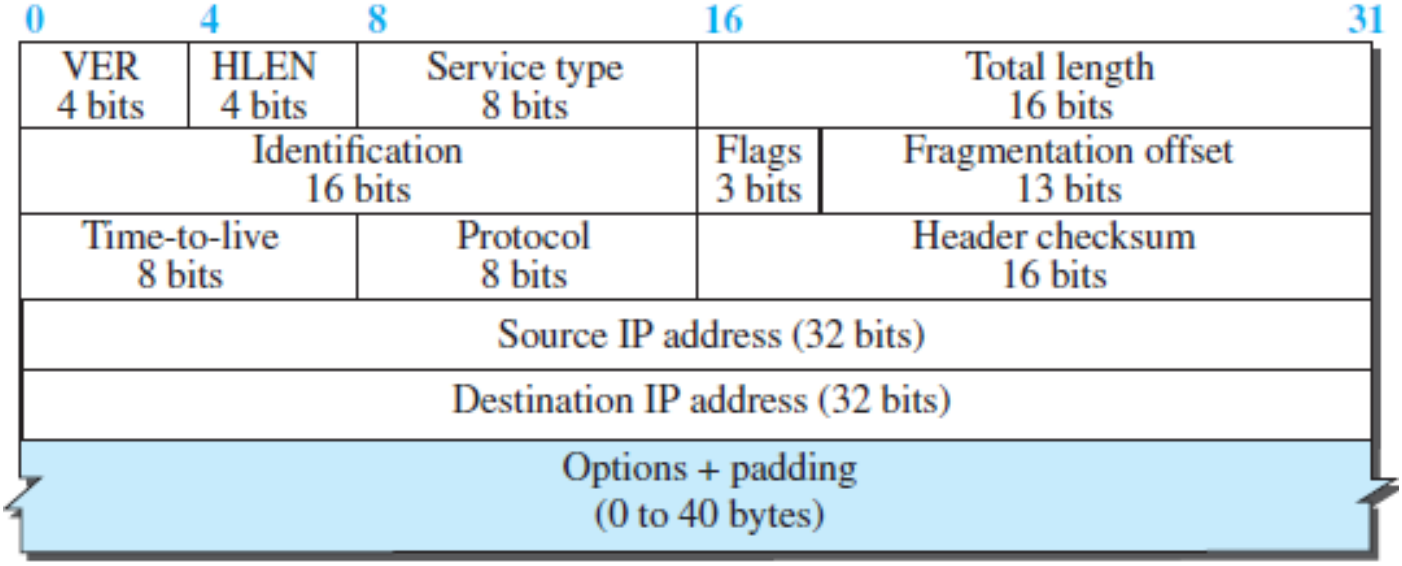
MTU: Maximum size of frame payload

- ***The value of the MTU differs from one physical network protocol to another.*** For example, the value for a LAN is normally 1500 bytes, but for a WAN it can be larger or smaller.
- When a datagram is fragmented it means that the payload of the IP datagram is fragmented and each fragment has its own header with most of the fields repeated, but some have been changed such as flags, fragmentation offset, and total length and checksum is recalculated at each point.
- A fragmented datagram may itself be fragmented if it encounters a network with an even smaller MTU. Thus, ***datagram may be fragmented several times before it reaches the final destination.***

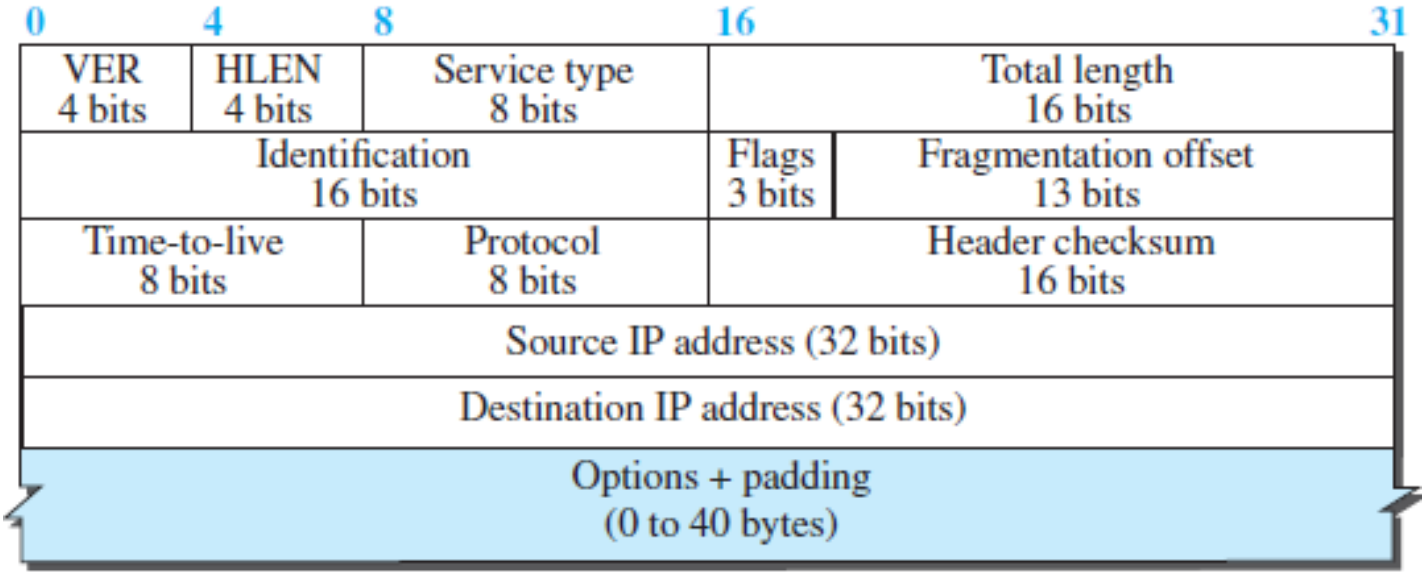
Fragmentation

- Fragmentation is a process of dividing the datagram into fragments during its transmission.
- Datagram can be fragmented by the source host or any router in the path.
- The reassembly of the datagram, is done only by the destination host, because each fragment becomes an independent datagram.
- The fragmented datagram can travel through different routes

Fields Related to Fragmentation

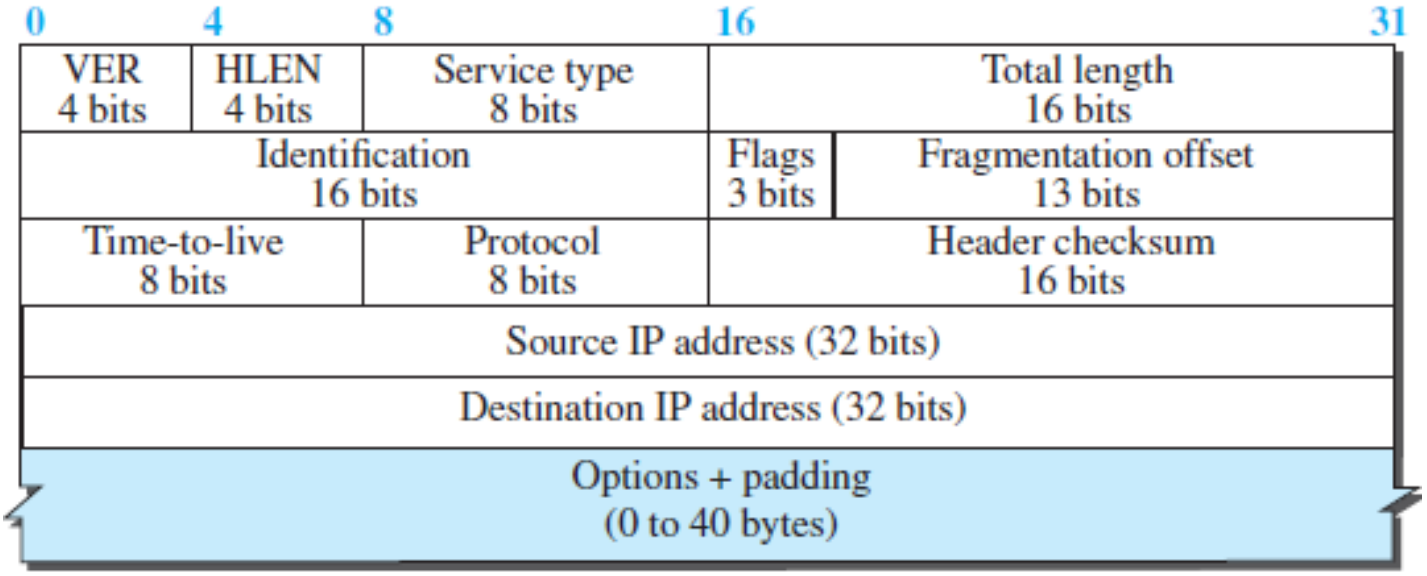


- **Identification**: 16-bit *identification field* identifies a datagram originating from the source host. To guarantee uniqueness, IP protocol uses a counter to label the datagrams.
- The counter is initialized to a positive number. When the IP protocol sends a datagram, it copies the current value of the counter to the identification field and increments the counter by one.
- When a datagram is fragmented, the value in the identification field is copied into all fragments so used for the identification of the fragments of an original IP datagram.
- The identification number helps the destination in reassembling the datagram.



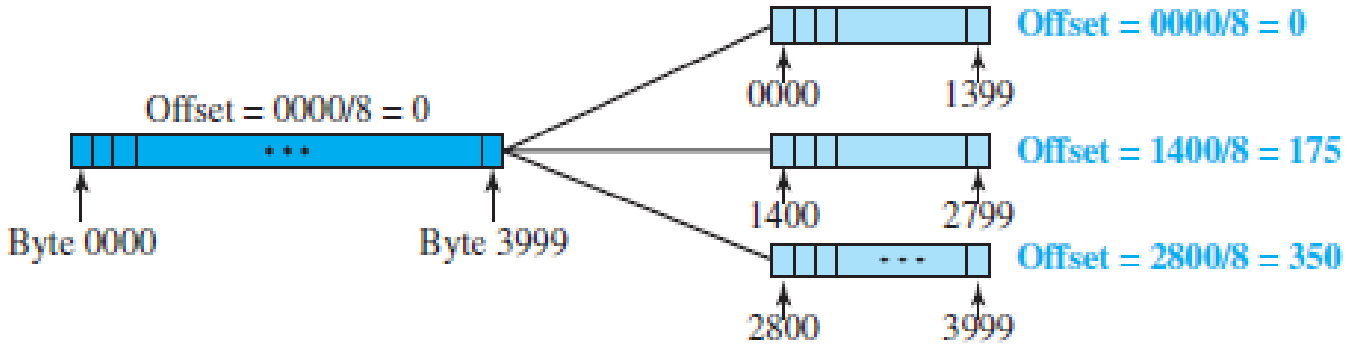
D: Donat fragment
M: More fragments

- **Flag Field:** The 3-bit *flags field* defines three flags.
  - The leftmost bit is reserved (not used).
  - The second bit (D bit) is called the *do not fragment* bit.
    - If its value is 1, the machine must not fragment the datagram.
    - If its value is 0, the datagram can be fragmented if necessary.
  - The third bit (M bit) is called the *more fragment bit*.
    - If its value is 1, it means the datagram is not the last fragment; there are more fragments after this one.
    - If its value is 0, it means this is the last or only fragment.



- **<u>Fragmentation Offset</u>**: The 13-bit *fragmentation offset field* shows the relative position of this fragment with respect to the whole datagram.
- It is the offset of the data in the original datagram measured in units of 8 bytes.

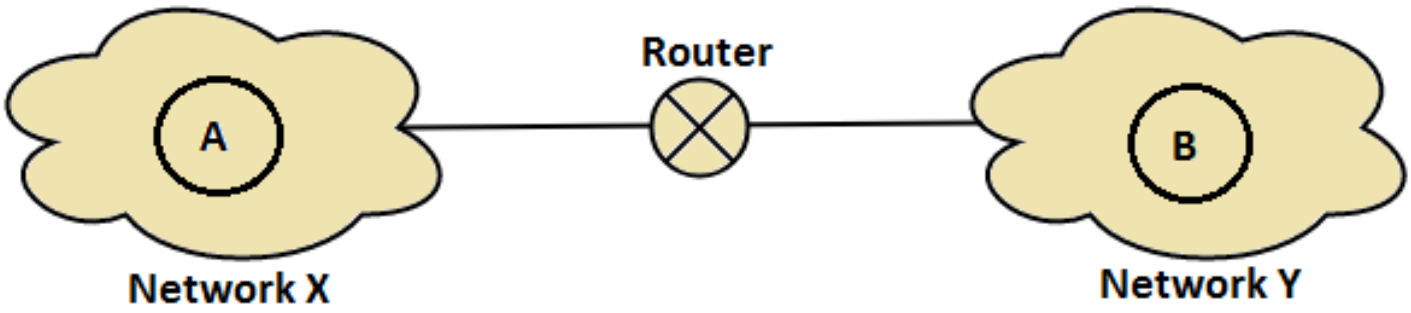
Example:



A datagram with a data size of 4000 bytes fragmented into three fragments.

- The bytes in the original datagram are numbered 0 to 3999.
- The first fragment carries bytes 0 to 1399. The offset value => 0/8 = 0.
- The second fragment carries bytes 1400 to 2799; the offset value => 1400/8 = 175.
- The third fragment carries bytes 2800 to 3999. The offset value => 2800/8 = 350.

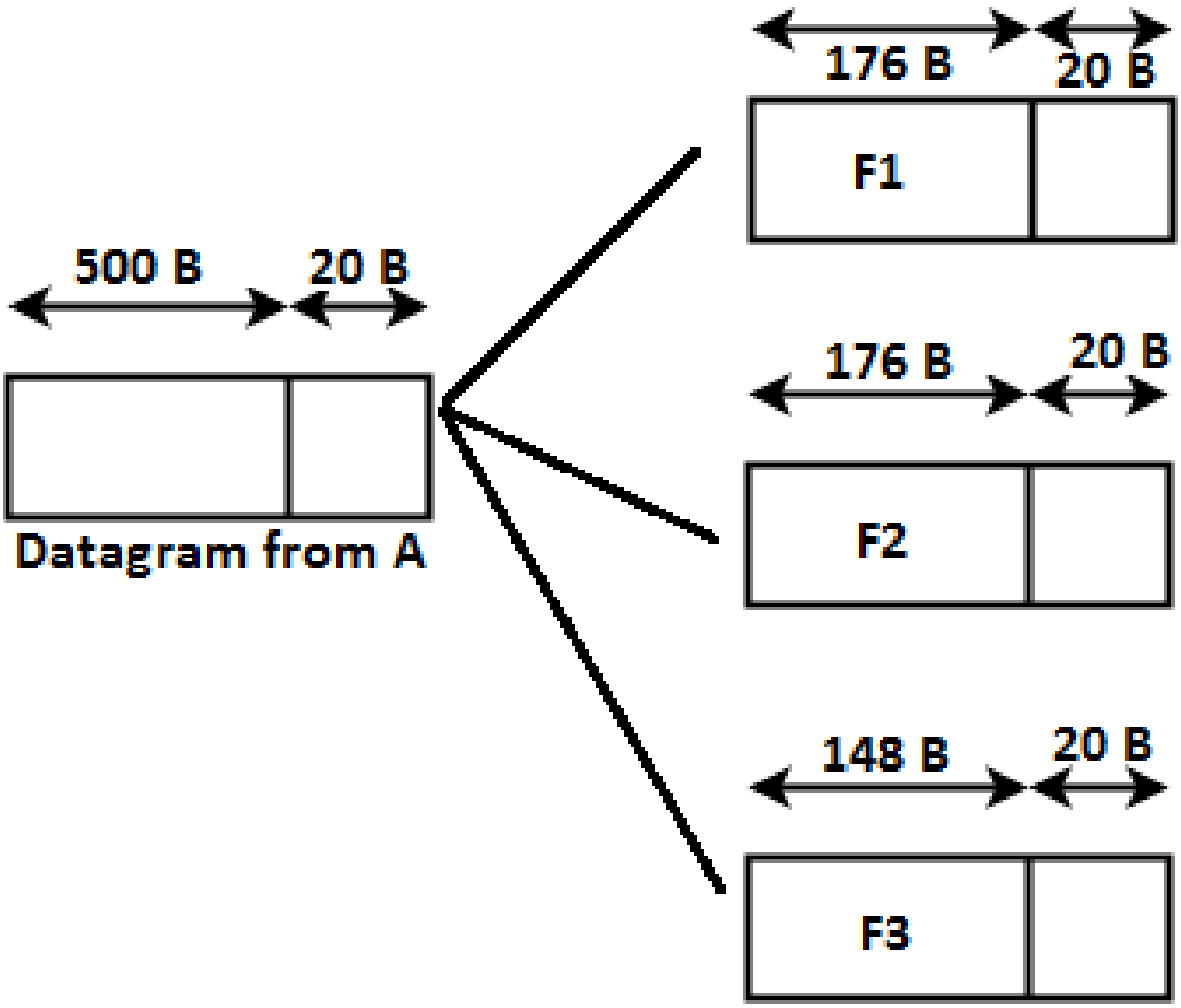
Example: Consider host A is present in network X having MTU = 520 bytes. There is another host B present in network Y having MTU = 200 bytes. Now, host A wants to send a message to host B.



- Consider that the router has received the datagram from A which have: IP Header of 20 bytes and payload length of 500 bytes and D bit is set to be 0.
- Now, router examines the MTU's of both the network and D bit and knows that it has to fragment the packet.
- The amount of data that can be sent is 200 Bytes, out of which 20 Bytes will be taken by header so the maximum data that can be sent is 180 bytes.

Now certain rules that needs to be followed during fragmentation are:

- The amount of data to be sent in one fragment is chosen in such a way that:
- It is as large as possible but should be less than or equal to MTU.
- It is a multiple of 8 so that pure decimal value can be obtained for the fragment offset field.
- It is not mandatory for the last fragment to contain the amount of data that is a multiple of 8, because it does not have to decide the fragment offset value for any other fragment.
- Using the above rules, the router is able to send maximum 176 bytes of data in one fragment as it is the greatest value that is a multiple of 8 and less than MTU.
- Now, the router will create 3 fragments of the original datagram such that:
    - **F1 = 176 Bytes**
    - **F2 = 176 Bytes**
    - **F3 = 148 Bytes**


500 B
20 B
Datagram from A
176 B
20 B
F1
176 B
20 B
F2
148 B
20 B
F3

Now, the information contained in each fragment's header will be:

- ***F1 Header***
  - Header length field value = 20 / 4 = 5
  - Total length field value = 176 + 20 = 196
  - M bit = 1
  - Fragment offset field value = 0
  - Header checksum is recalculated.
  - Identification number is same as that of original datagram.
- ***F2 Header***
  - Header length field value = 20 / 4 = 5
  - Total length field value = 176 + 20 = 196
  - M bit = 1
  - Fragment offset field value = 176 / 8 = 22
  - Header checksum is recalculated.
  - Identification number is same as that of original datagram.
- ***F3 Header***
  - Header length field value = 20 / 4 = 5
  - Total length field value = 148 + 20 = 168
  - M bit = 0
  - Fragment offset field value = (176 + 176) / 8 = 44
  - Header checksum is recalculated.
  - Identification number is same as that of original datagram.
- **At destination, the receiver is receiving 3 fragmented frames, the final destination host can reassemble the original datagram from the fragments received (if none of them is lost) using the following strategy:**
  A. The first fragment has an offset field value of zero.
  B. Divide the length of the first fragment by 8. The second fragment has an offset value equal to that result.
  C. Divide the total length of the first and second fragment by 8. The third fragment has an offset value equal to that result.
  D. Continue the process. The last fragment has its M bit set to 0.
  E. Continue the process. The last fragment has a *more* bit value of 0.

Fragmentation Overhead

- Fragmentation increases the overhead as with each fragment we have to append the header.
    - **Total Overhead = (Total number of fragmented datagrams – 1) x size of IP header**
    - **Efficiency = Data without header / data with header**
    - **Throughput = Efficiency x Bandwidth**

- **Few Important Points to Note:**
    - Source does not require fragmentation due to wise segmentation by transport layer.
    - If a datagram goes through a path where different intermediary paths are having different bandwidths. Then, while calculating the throughput, we consider the minimum bandwidth since it acts as a bottleneck.

Example: 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?

Since, M bit is set to 0 it is definitely the last fragment. We cannot know whether the packet was fragmented or not as a non-fragmented packet is considered the last fragment.

Example: A packet has arrived with an M bit value of 1 and a fragmentation offset value of 0. Is this the first fragment, the last fragment, or a middle fragment?

Since the M bit value is 1 it is definitely not the last fragment, and since the fragmentation offset is 0 we can conclude that it is the first fragment.

Example: 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?

To get the number of first byte we multiply the offset value with 8, 100 x 8 = 800 but we cannot get the last value.

Example: 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?

Now, the first byte is 100 x 8 = 800.

Since the HLEN value is 5, total length of header = 5 x 4 = 20 BytesTotal Length given is 100 Bytes out of which 20 bytes is header, which means 80 bytes are present in the datagram. If 800 is the first byte then 879 will be the last one.

Q Consider an IP packet with a length of 4,500 bytes that includes a 20-byte IPv4 header and a 40-byte TCP header. The packet is forwarded to an IPv4 router that supports a Maximum Transmission Unit (MTU) of 600 bytes. Assume that the length of the IP header in all the outgoing fragments of this packet is 20 bytes. Assume that the fragmentation offset value stored in the first fragment is 0. The fragmentation offset value stored in the third fragment is _____. **(Gate-2018) (2 Marks)**
Ans: 144

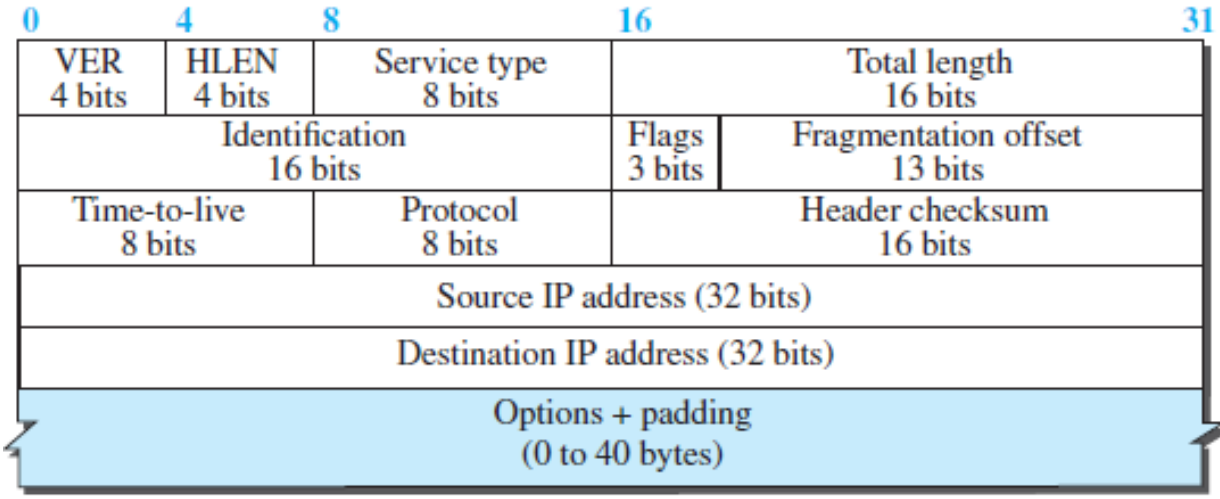
Q An IP datagram of size 1000 bytes arrives at a router. The router has to forward this packet on a link whose MTU (maximum transmission unit) is 100 bytes. Assume that the size of the IP header is 20 bytes. The number of fragments that the IP datagram will be divided into for transmission is _____. **(Gate-2016) (2 Marks)**
ANSWER 13

Q Host A sends a UDP datagram containing 8880 bytes of user data to host B over an Ethernet LAN. Ethernet frames may carry data up to 1500 bytes (i.e. MTU = 1500 bytes). Size of UDP header is 8 bytes and size of IP header is 20 bytes. There is no option field in IP header. How may total number of IP fragments will be transmitted and what will be the contents of offset field in the last fragment? **(Gate-2015) (2 Marks)**
(A) 6 and 925 **(B)** 6 and 7400
(C) 7 and 1110 **(D)** 7 and 8880
Answer: (C)

Q An IP router with a Maximum Transmission Unit (MTU) of 1500 bytes has received an IP packet of size 4404 bytes with an IP header of length 20 bytes. The values of the relevant fields in the header of the third IP fragment generated by the router for this packet are **(Gate-2014) (2 Marks)**
(A) MF bit: 0, Datagram Length: 1444; Offset: 370
(B) MF bit: 1, Datagram Length: 1424; Offset: 185
(C) MF bit: 1, Datagram Length: 1500; Offset: 37
(D) MF bit: 0, Datagram Length: 1424; Offset: 2960
Answer: (A)

Q In an IPv4 datagram, the M bit is 0, the value of HLEN is 10, the value of total length is 400 and the fragment offset value is 300. The position of the datagram, the sequence numbers of the first and the last bytes of the payload, respectively are **(Gate-2013) (2 Marks)**
(A) Last fragment, 2400 and 2789 **(B)** First fragment, 2400 and 2759
(C) Last fragment, 2400 and 2759 **(D)** Middle fragment, 300 and 689
Answer: (C)

Time-to-live

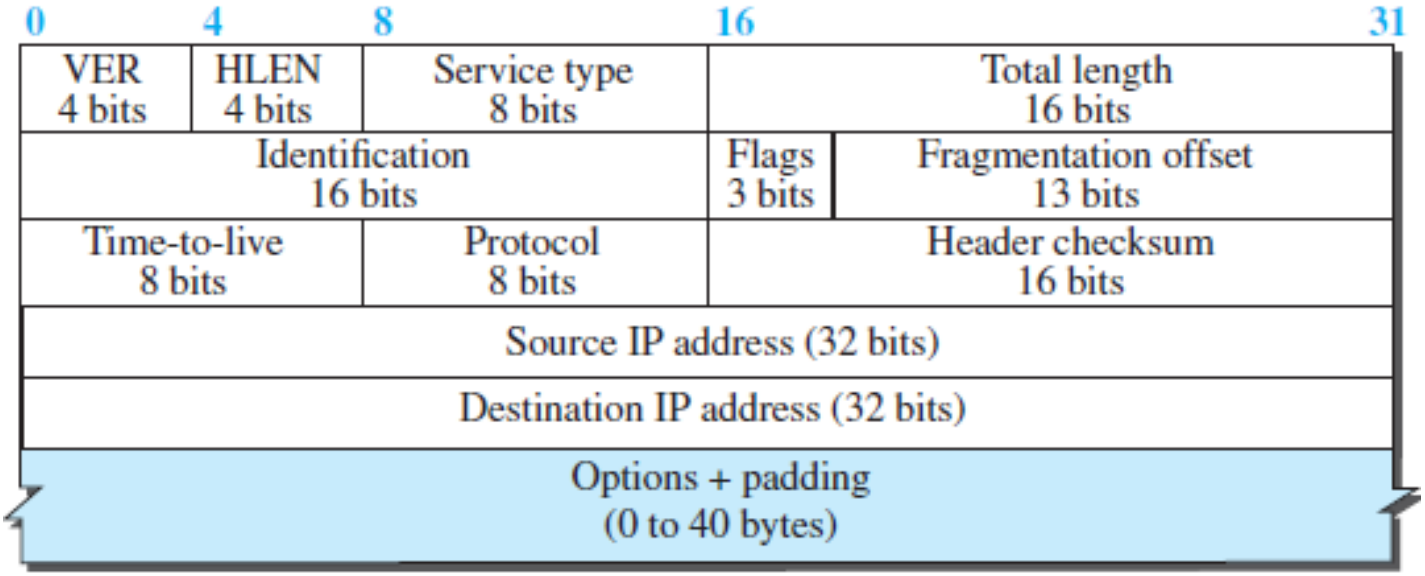


- **Time-to-live.** The time-to-live (TTL) field is used to control the maximum number of hops (routers) visited by the datagram.
- This value is approximately two times the maximum number of routers between any two hosts.
- Each router that processes the datagram decrements this number by one.
- If this value, after being decremented, is zero, the router discards the datagram.
- This field is needed because routing tables in the Internet can become corrupted. A datagram may travel between two or more routers for a long time without ever getting delivered to the destination host. This field limits the lifetime of a datagram.
- Another use of this field is to intentionally limit the journey of the packet. For example, if the source wants to confine the packet to the local network, it can store 1 in this field. When the packet arrives at the first router, this value is decremented to 0, and the datagram is discarded.

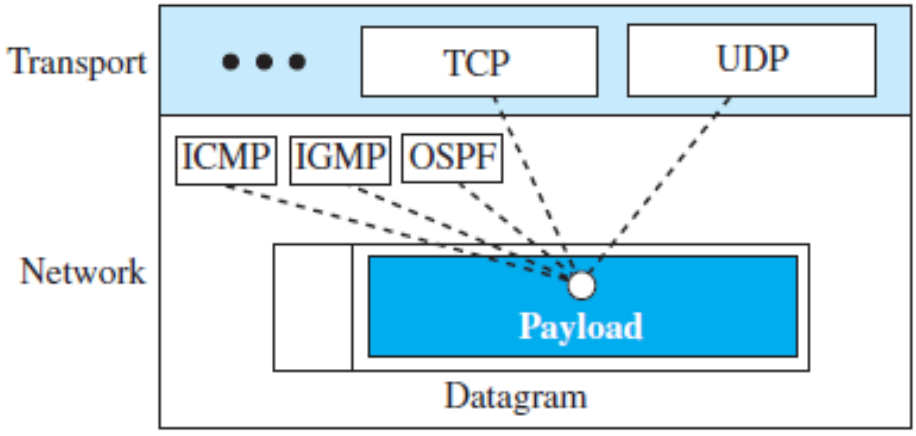
Q One of the header fields in an IP datagram is the Time to Live (TTL) field. Which of the following statements best explains the need for this field? **(Gate-2010) (1 Marks)**
(A) It can be used to prioritize packets
(B) It can be used to reduce delays
(C) It can be used to optimize throughput
(D) It can be used to prevent packet looping
Answer (D)

Q For which one of the following reasons does Internet Protocol (IP) use the time-to-live (TTL) field in the IP datagram header **(Gate-2006) (1 Marks)**
(A) Ensure packets reach destination within that time
(B) Discard packets that reach later than that time
(C) Prevent packets from looping indefinitely
(D) Limit the time for which a packet gets queued in intermediate routers.
Answer: (C)

Protocol



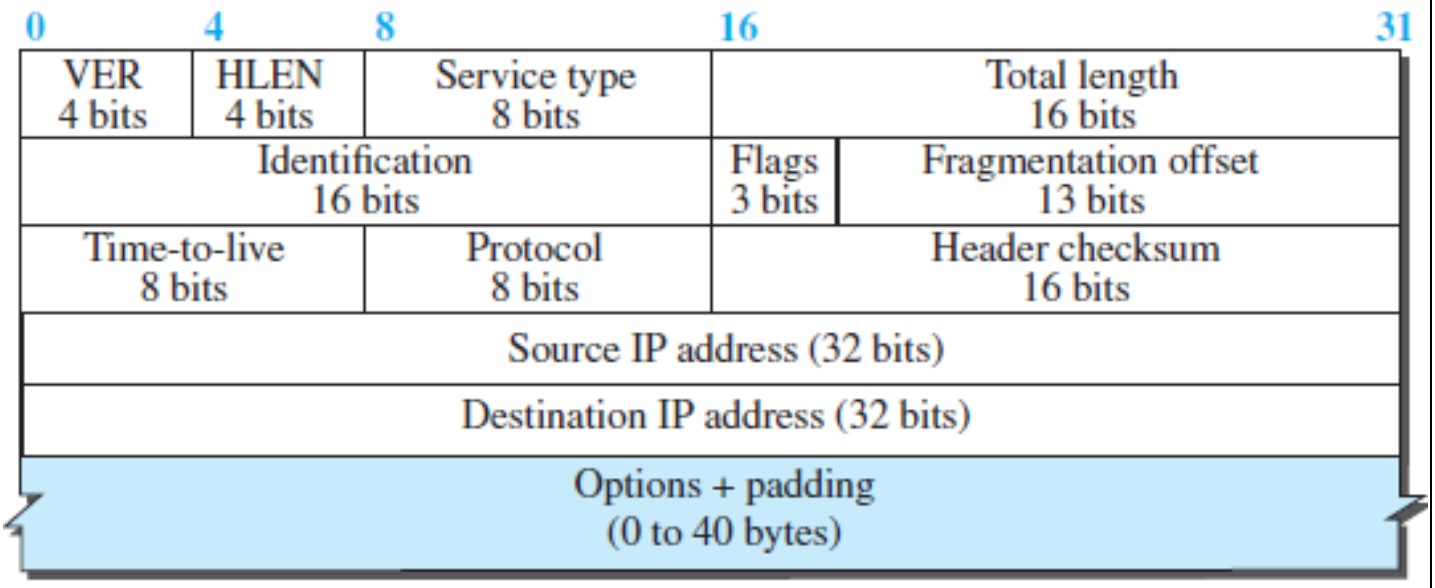
- **Protocol**. In TCP/IP, the data section of a packet, called the payload, carries the whole packet from another protocol. A datagram, for example, can carry a packet belonging to any transport-layer protocol such as UDP or TCP.
- When the datagram arrives at the destination, the value of this field helps to define to which protocol the payload should be delivered.



Some protocol values

ICMP	01
IGMP	02
TCP	06
UDP	17
OSPF	89

Header checksum



- **Header checksum.** IP adds a header checksum field to check the header, but not the payload.
    - IP is not a reliable protocol; it does not check whether the payload carried by a datagram is corrupted during the transmission.
    - The datagram header, is added by IP, and its error-checking is the responsibility of IP.
    - Since the value of some fields, such as TTL, may change from router to router, the checksum needs to be recalculated at each router.
    - First, all higher-level protocols that encapsulate data in the IPv4 datagram have a checksum field that covers the whole packet. Therefore, the checksum for the IPv4 datagram does not have to check the encapsulated data.
    - Second, the header of the IPv4 packet changes with each visited router, but the data do not. So, the checksum includes only the part that has changed. If the data were included, each router must recalculate the checksum for the whole packet, which means an increase in processing time.

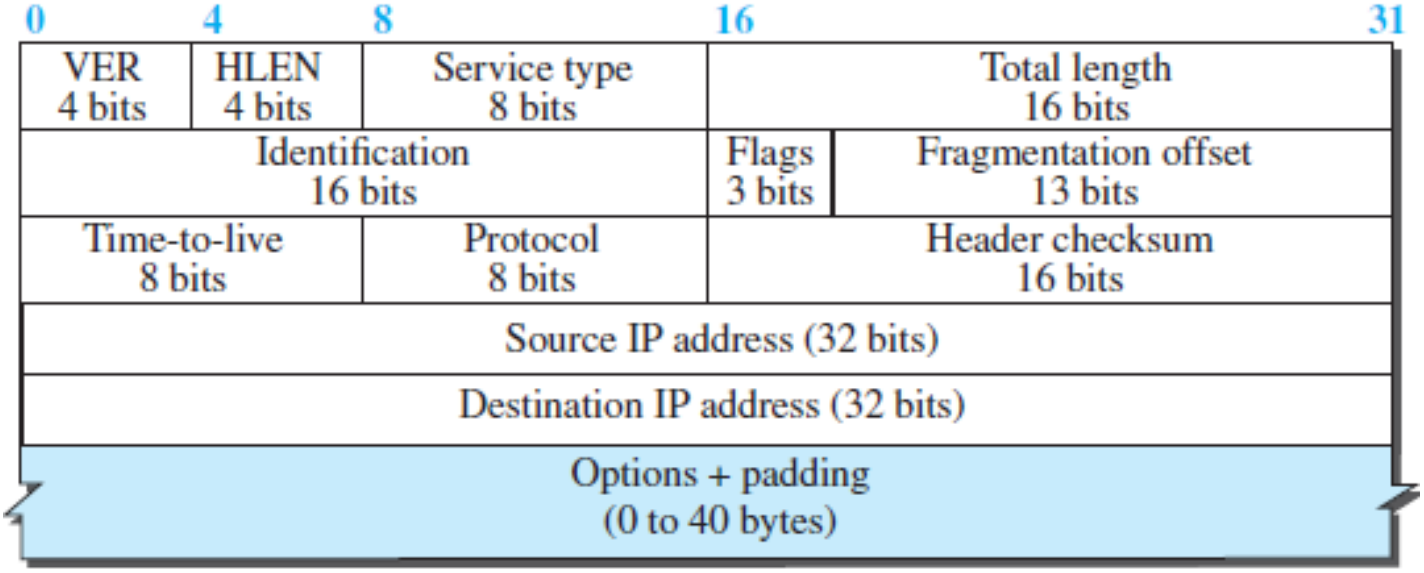
Q Host A (on TCP/IP v4 network A) sends an IP datagram D to host B (also on TCP/IP v4 network B). Assume that no error occurred during the transmission of D. When D reaches B, which of the following IP header field(s) may be different from that of the original datagram D? **(Gate-2014) (1 Marks)**

(i) TTL **(ii)** Checksum **(iii)** Fragment Offset

(A) (i) only
(B) (i) and (ii) only
(C) (ii) and (iii) only
(D) (i), (ii) and (iii)
Answer: (D)

Q Which of the following statements is TRUE? **(Gate-2006) (1 Marks)**
(A) Both Ethernet frame and IP packet include checksum fields
(B) Ethernet frame includes a checksum field and IP packet includes a CRC field
(C) Ethernet frame includes a CRC field and IP packet includes a checksum field
(D) Both Ethernet frame and IP packet include CRC fields
Answer: (C)

Source and Destination Addresses



- **Source and Destination Addresses.** These 32-bit source and destination address fields define the IP address of the source and destination respectively.
- **Options.** A datagram header can have up to 40 bytes of options. Options can be used for network testing and debugging.
    - They are not a required part of the IP header
- **Payload.** Payload, or data, is the main reason for creating a datagram. Payload is the packet coming from other protocols that use the service of IP.
    - Payload is the content of the package; the header is only the information written on the package.

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

- 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: In an IPv4 packet, the value of HLEN is 5, and the value of the total length field is $(0028)_{16}$. How many bytes of data are being carried by this packet?

The HLEN value is 5, thus the total number of bytes in the header is $5 \times 4 = 20$ bytes (no options).

The total length is $(0028)_{16} = 16^1 \times 2 + 16^0 \times 8 = 40$ bytes, which means the packet is carrying 20 bytes of data (40 − 20).

Example: An IPv4 packet has arrived with the first few hexadecimal digits as shown.

$$(45000028000100000102\ldots)_{16}$$

How many hops can this packet travel before being dropped?

- Each hexadecimal digit i.e. Base 16 is equivalent to 4 binary digits i.e Base 2

So, each digit in hexadecimal notation defines 4 binary digits. TTL field is after 64 binary digits.

So, 64 / 4 = 16 hexadigits

Skipping 16 hexadigits we get, 01 as our answer (as TTL field is of 8 bits, thus will require two hexadecimal digits)

$$(4500002800010000\mathbf{01}02\ldots)_{16}$$

$(01)_{16} = (1)_{10}$, so after 1 hop it will get discarded.

Variable part

- The header of the IPv4 datagram is made of two parts: a fixed part and a variable part.
- The fixed part is 20 bytes long and was discussed in the previous section.
- The variable part comprises the options that can be a maximum of 40 bytes. Options, as the name implies, are not required for a datagram.
- They can be used for network testing and debugging. Although options are not a required part of the IPv4 header, option processing is required of the IPv4 software.
- This means that all implementations must be able to handle options if they are present in the header.

- **End of Option**
  - An end-of-option option is a 1-byte option used for padding at the end of the option field. It, however, can only be used as the last option.
- **Record Route**
  - A record route option is used to record the Internet routers that handle the datagram. It can list up to nine router addresses. It can be used for debugging and management purposes.
- **Strict Source Route**
  - A strict source route option is used by the source to predetermine a route for the datagram as it travels through the Internet. Dictation of a route by the source can be useful for several purposes.
  - The sender can choose a route with a specific type of service, such as minimum delay or maximum throughput.
  - Alternatively, it may choose a route that is safer or more reliable for the sender's purpose. For example, a sender can choose a route so that its datagram does not travel through a competitor's network.
  - If a datagram specifies a strict source route, all the routers defined in the option must be visited by the datagram. A router must not be visited if its IPv4 address is not listed in the datagram. If the datagram visits a router that is not on the list, the datagram is discarded and an error message is issued.
  - If the datagram arrives at the destination and some of the entries were not visited, it will also be discarded and an error message issued.
- **Loose Source Route**
  - A loose source route option is similar to the strict source route, but it is less rigid. Each router in the list must be visited, but the datagram can visit other routers as well.

- **Timestamp**
  - A timestamp option is used to record the time of datagram processing by a router. The time is expressed in milliseconds from midnight, Universal time or Greenwich mean time.
  - Knowing the time, a datagram is processed can help users and managers track the behaviour of the routers in the Internet. We can estimate the time it takes for a datagram to go from one router to another. We say estimate because, although all routers may use Universal time, their local clocks may not be synchronized.

Q The maximum number of IPv4 router addresses that can be listed in the record route (RR) option field of an IPv4 header is _____ **(Gate-2017) (1 Marks)**
Ans: 9

Q Which one of the following fields of an IP header is NOT modified by a typical IP router? **(Gate-2015) (1 Marks)**
(A) Checksum
(B) Source address
(C) Time to Live (TTL)
(D) Length
Answer: (B)

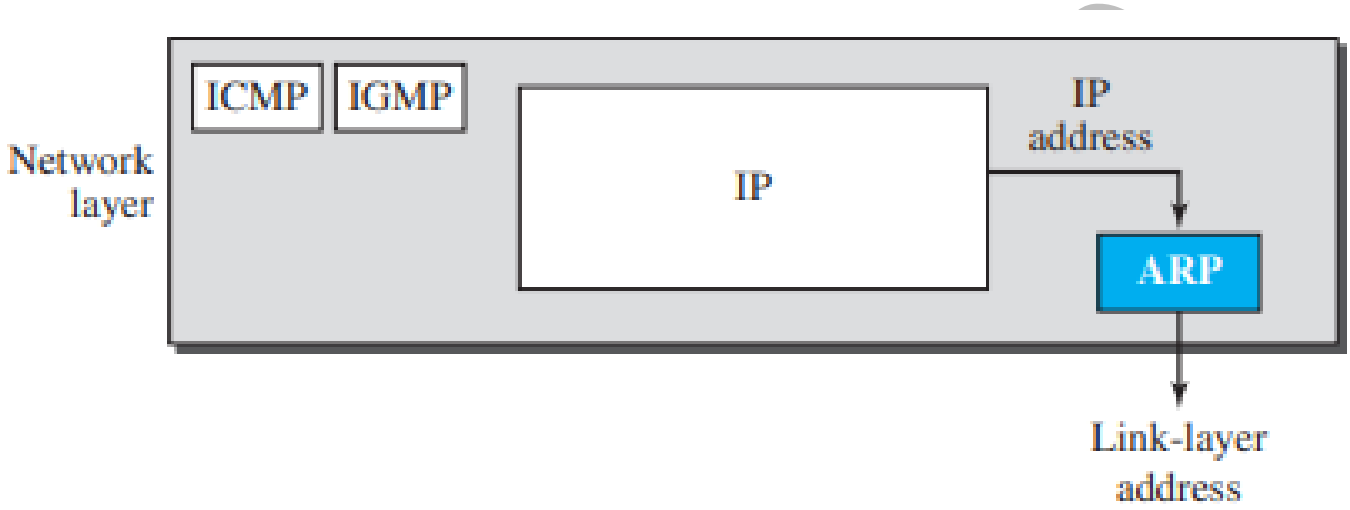
Q Which of the following assertions is FALSE about the Internet Protocol (IP)? **(Gate-2003) (1 Marks)**
(A) It is possible for a computer to have multiple IP addresses
(B) IP packets from the same source to the same destination can take different routes in the network
(C) IP ensures that a packet is discarded if it is unable to reach its destination within a given number of hops
(D) The packet source cannot set the route of an outgoing packets; the route is determined only by the routing tables in the routers on the way
Answer: (D)

Additional protocols

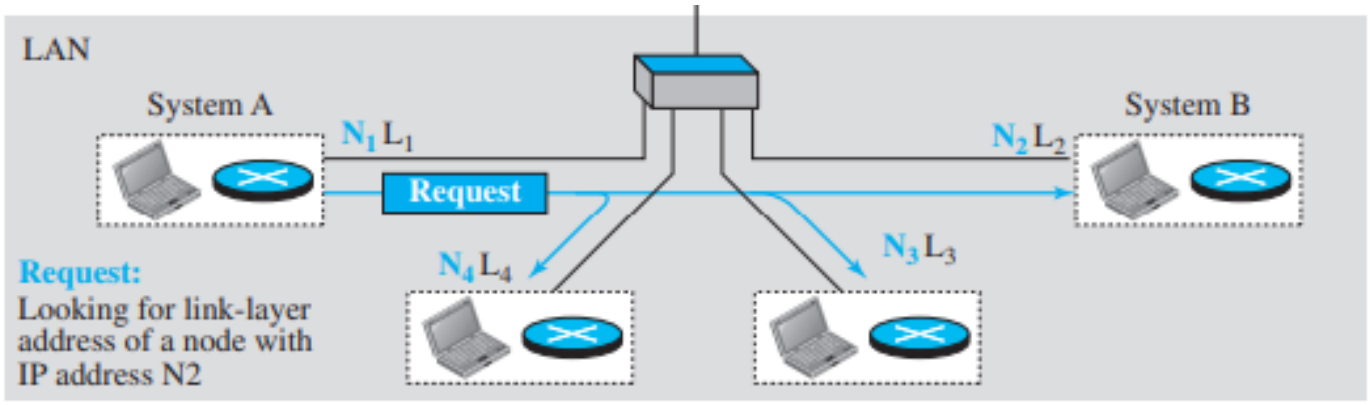
- IP packets, however, need to be encapsulated in a frame, which needs physical addresses (node-to-node). We will see that a protocol called ARP, the Address Resolution Protocol.
- We sometimes need reverse mapping-mapping a physical address to a logical address. For example, when booting a diskless network or leasing an IP address to a host, RARP is used.
- Lack of flow and error control in the Internet Protocol has resulted in another protocol, ICMP, that provides alerts. It reports congestion and some types of errors in the network or destination host
- IP was originally designed for unicast delivery, one source to one destination. As the Internet has evolved, the need for multicast delivery, one source to many destinations, has increased tremendously. IGMP gives IP a multicast capability.

Address Resolution Protocol (ARP)

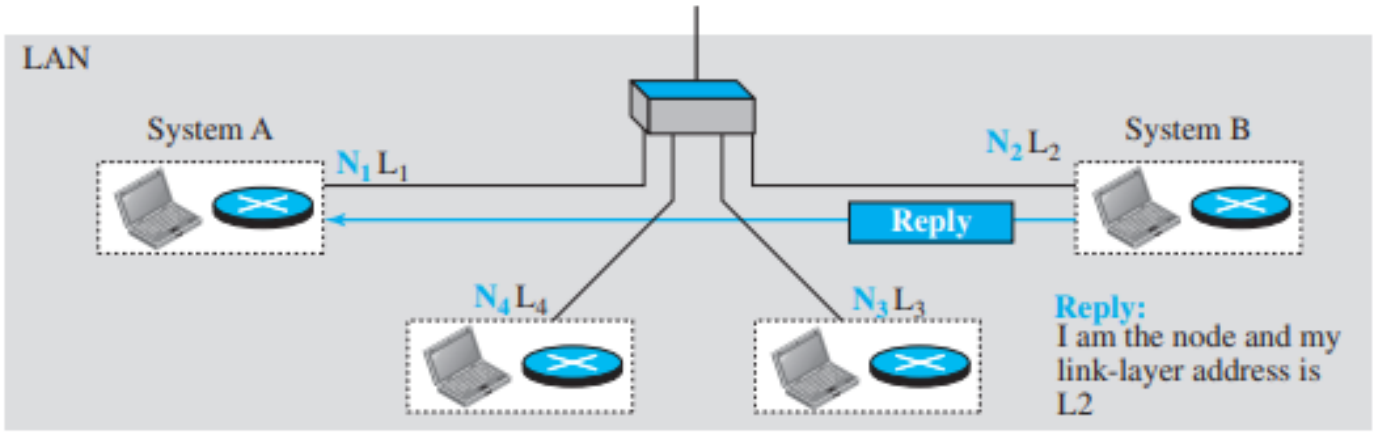
- The IP address of the next node alone is not helpful in moving a frame through a link; we need the link-layer address of the next node so that data link layer can work.
- ARP maps an IP address to a logical-link address.
- ARP accepts an IP address from the IP protocol, maps the address to the corresponding link-layer address, and passes it to the data-link layer.
- The ARP protocol is one of the auxiliary protocols defined in the **network layer.**



- Anytime a host or a router needs to find the link-layer address of another host or router in its network, it sends an ARP request packet.
- The packet includes the ***link-layer and IP addresses of the sender and the IP address of the receiver.***
- Because the sender does not know the link-layer address of the receiver, the query is broadcast over the link.
- Every host or router on the network receives and processes the ARP request packet, but only the intended recipient recognizes its IP address and sends back an ARP response packet.
- The response packet contains the recipient's IP and link-layer addresses.
- ***The packet is unicast directly to the node that sent the request packet.***



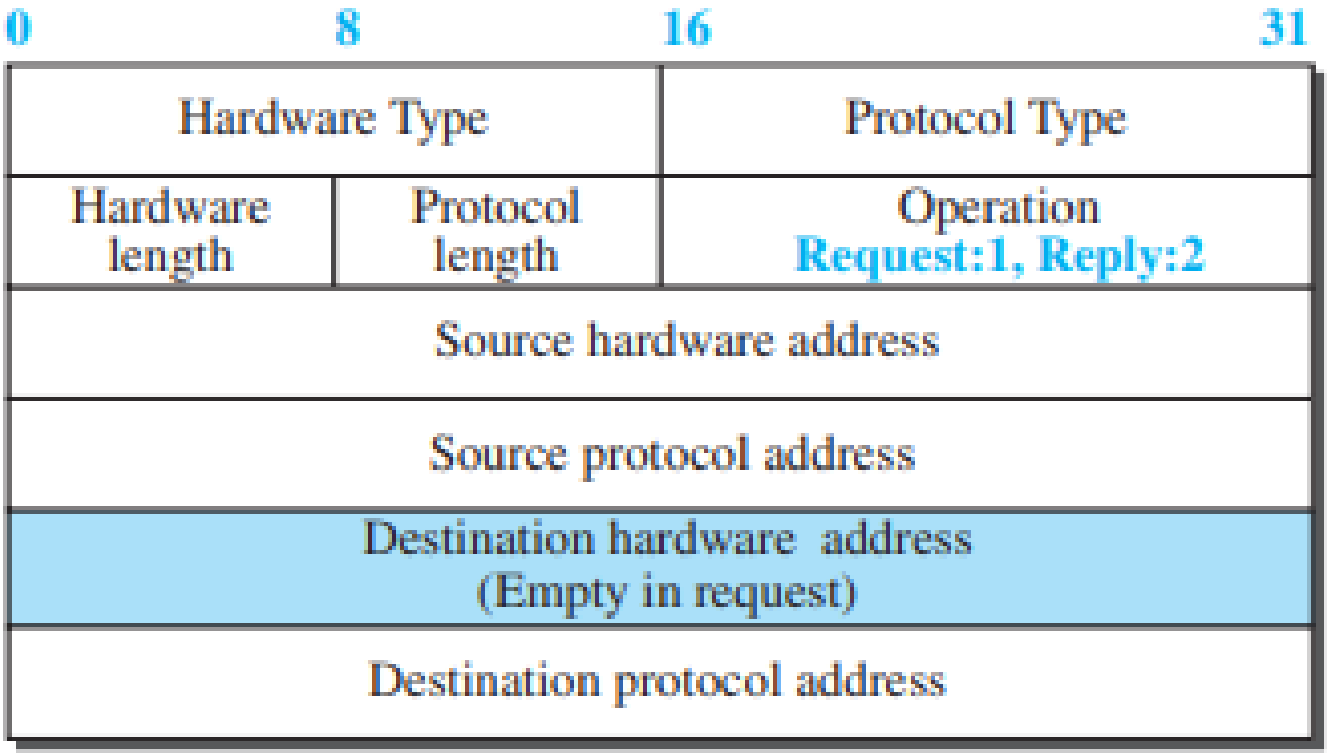
a. ARP request is broadcast



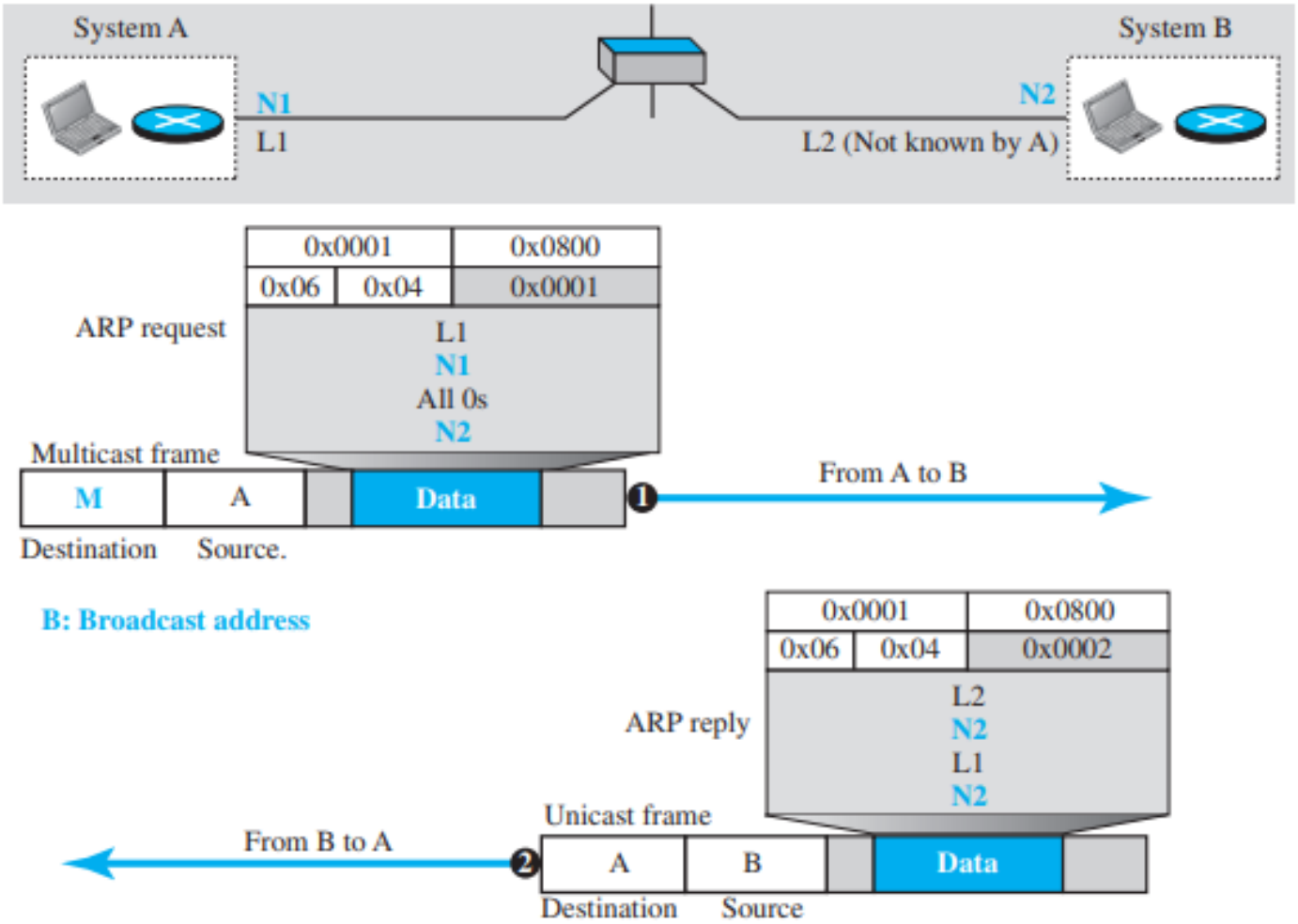
b. ARP reply is unicast

ARP Packet Format

- **Hardware type:** This is a 16-bit field defining the type of the network on which ARP is running. Each LAN has been assigned an integer based on its type. For example, Ethernet is given type 1. ARP can be used on any physical network.
- **Protocol type:** This is a 16-bit field defining the protocol. For example, the value of this field for the IPv4 protocol is 080016, ARP can be used with any higher-level protocol.
- Hardware length. This is an 8-bit field defining the length of the physical address in bytes. For example, for Ethernet the value is 6.
- Protocol length. This is an 8-bit field defining the length of the logical address in bytes. For example, for the IPv4 protocol the value is 4.
- Operation. This is a 16-bit field defining the type of packet. Two packet types are defined: ARP request (1) and ARP reply (2).
- Sender hardware address. This is a variable-length field defining the physical address of the sender. For example, for Ethernet this field is 6 bytes long.
- Sender protocol address. This is a variable-length field defining the logical (for example, IP) address of the sender. For the IP protocol, this field is 4 bytes long.
- Target hardware address. This is a variable-length field defining the physical address of the target. For example, for Ethernet this field is 6 bytes long. For an ARP request message, this field is all Os because the sender does not know the physical address of the target.
- Target protocol address. This is a variable-length field defining the logical (for example, IP) address of the target. For the IPv4 protocol, this field is 4 bytes long

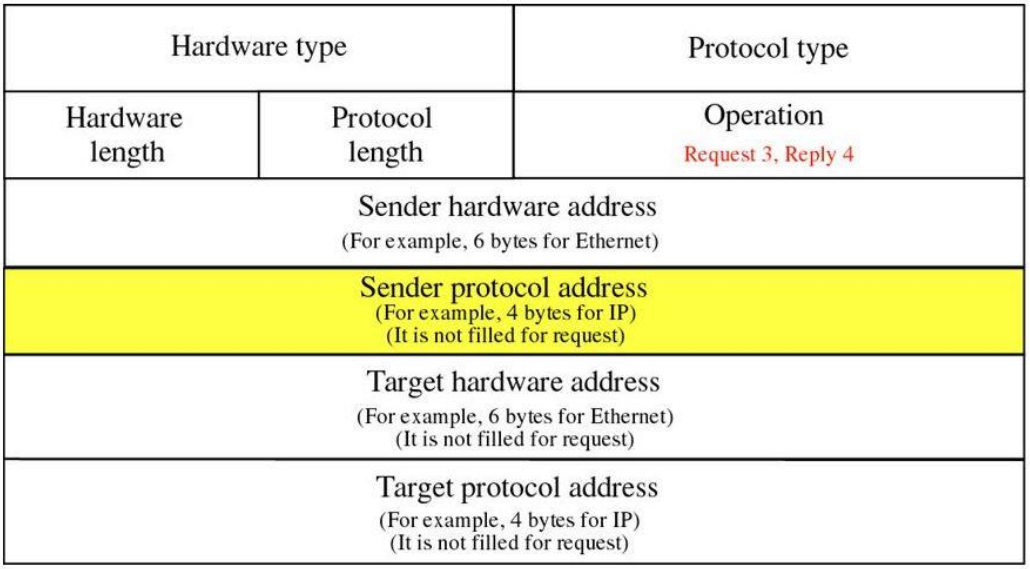

0
8
16
31
Hardware Type
Protocol Type
Hardware length
Protocol length
Operation
Request:1, Reply:2
Source hardware address
Source protocol address
Destination hardware address
(Empty in request)
Destination protocol address

Example: A host with IP address N1 and MAC address L1 has a packet to send to another host with IP address N2 and physical address L2 (which is unknown to the first host). The two hosts are on the same network. Figure below shows the ARP request and response messages.



RARP

- Reverse Address Resolution Protocol (RARP) finds the logical address for a machine that knows only its physical address.
- Each host or router is assigned one or more logical (IP) addresses, which are unique and independent of the physical (hardware) address of the machine. To create an IP datagram, a host or a router needs to know its own IP address or addresses.
- The IP address of a machine is usually read from its configuration file stored on a disk file. However, a diskless machine is usually booted from ROM, which has minimum booting information. The ROM is installed by the manufacturer. It cannot include the IP address because the IP addresses on a network are assigned by the network administrator.



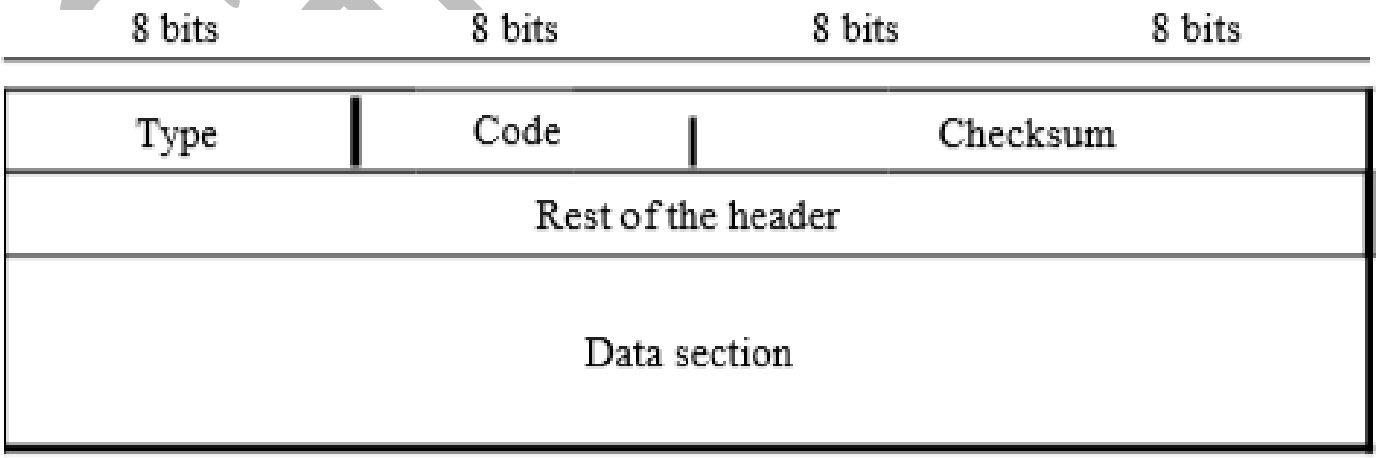
- The machine can get its physical address (by reading its NIC, for example), which is unique locally. It can then use the physical address to get the logical address by using the RARP protocol.
- A RARP request is created and broadcast on the local network. Another machine on the local network that knows all the IP addresses will respond with a RARP reply.
- The requesting machine must be running a RARP client program; the responding machine must be running a RARP server program. There is a serious problem with RARP: Broadcasting is done at the data link layer.
- The physical broadcast address, allis in the case of Ethernet, does not pass the boundaries of a network. This means that if an administrator has several networks or several subnets, it needs to assign a RARP server for each network or subnet. This is the reason that RARP is almost obsolete.
- BOOTP and DHCp, are replacing RARP.

ICMP

- IP has two deficiencies: lack of error control and lack of assistance mechanisms. The IP protocol has no error-reporting or error-correcting mechanism. What happens if something goes wrong? What happens if a router must discard a datagram because it cannot find a router to the final destination, or because the time-to-live field has a zero value? What happens if the final destination host must discard all fragments of a datagram because it has not received all fragments within a predetermined time limit?
- These are examples of situations where an error has occurred and the IP protocol has no built-in mechanism to notify the original host. The IP protocol also lacks a mechanism for host and management queries. A host sometimes needs to determine if a router or another host is alive. And sometimes a network administrator needs information from another host or router.
- The Internet Control Message Protocol (ICMP) has been designed to compensate for the above two deficiencies. It is a companion to the IP protocol.

Message Format

- An ICMP message has an 8-byte header and a variable-size data section. Although the general format of the header is different for each message type, the first 4 bytes are common to all message.
- The code field specifies the reason for the particular message type. The last common field is the checksum field. The rest of the header is specific for each message type.
- The data section in error messages carries information for finding the original packet that had the error. In query messages, the data section carries extra information based on the type of the query.



Types of Messages

- ICMP messages are divided into two broad categories: error-reporting messages and query messages.
- The error-reporting messages report problems that a router or a host (destination) may encounter when it processes an IP packet.
- The query messages, which occur in pairs, help a host or a network manager get specific information from a router or another host. For example, nodes can discover their neighbours. Also, hosts can discover and learn about routers on their network, and routers can help a node redirect its message.

Error Reporting

- One of the main responsibilities of ICMP is to report errors. Although technology has produced increasingly reliable transmission media, errors still exist and must be handled. IP, is an unreliable protocol.
- This means that error checking and error control are not a concern of IP.
- ICMP was designed, in part, to compensate for this shortcoming. However, ICMP does not correct errors-it simply reports them.
- Error correction is left to the higher-level protocols. Error messages are always sent to the original source because the only information available in the datagram about the route is the source and destination IP addresses.
- ICMP uses the source IP address to send the error message to the source (originator) ofthe datagram.

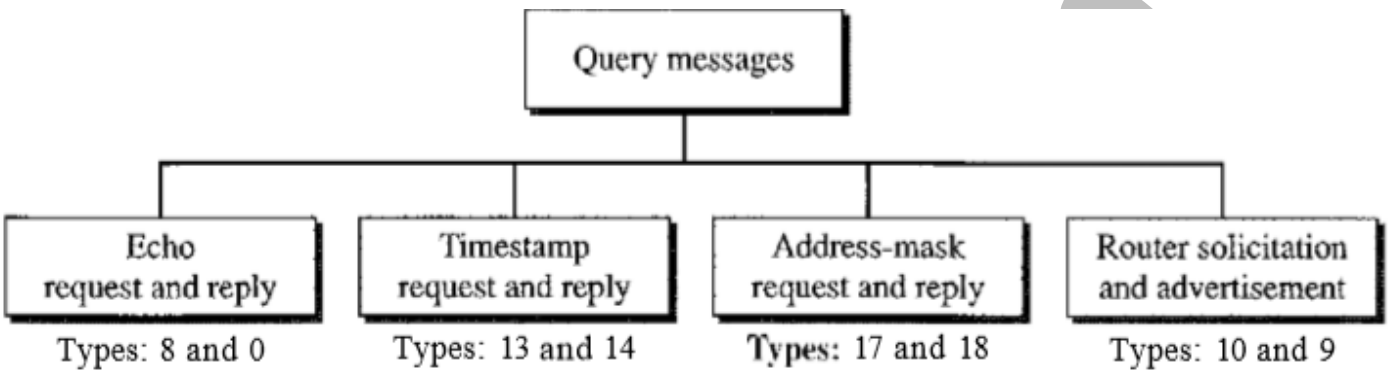
The following are important points about ICMP error messages:

- No ICMP error message will be generated in response to a datagram carrying an ICMP error message.
- No ICMP error message will be generated for a fragmented datagram that is not the first fragment.
- No IeMP error message will be generated for a datagram having a multicast address.
- No ICMP error message will be generated for a datagram having a special address such as 127.0.0.0 or 0.0.0.0.

- Note that all error messages contain a data section that includes the IP header of the original datagram plus the first 8 bytes of data in that datagram.
- The original datagram header is added to give the original source, which receives the error message, information about the datagram itself.

Query

- In addition to error reporting, ICMP can diagnose some network problems. This is accomplished through the query messages, a group of four different pairs of messages.
- In this type of ICMP message, a node sends a message that is answered in a specific format by the destination node. A query message is encapsulated in an IP packet, which in turn is encapsulated in a data link layer frame. However, in this case, no bytes of the original IP are included in the message



Echo Request and Reply

- The echo-request and echo-reply messages are designed for diagnostic purposes.
- Network managers and users utilize this pair of messages to identify network problems.
- The combination of echo-request and echo-reply messages determines whether two systems (hosts or routers) can communicate with each other.
- The echo-request and echo-reply messages can be used to determine if there is communication at the IP level. Because ICMP messages are encapsulated in IP datagrams, the receipt of an echo-reply message by the machine that sent the echo request is proof that the IP protocols in the sender and receiver are communicating with each other using the IP datagram.
- Also, it is proof that the intermediate routers are receiving, processing, and forwarding IP datagrams. Today, most systems provide a version of the ping command that can create a series (instead of just one) of echo-request and echo-reply messages, providing statistical information.

Timestamp Request and Reply

- Two machines (hosts or routers) can use the timestamp request and timestamp reply messages to determine the round-trip time needed for an IP datagram to travel between them. It can also be used to synchronize the clocks in two machines

Address-Mask Request and Reply

- A host may know its IP address, but it may not know the corresponding mask. For example, a host may know its IP address as 159.31.17.24, but it may not know that the corresponding mask is /24.
- To obtain its mask, a host sends an address-mask-request message to a router on the LAN. If the host knows the address of the router, it sends the request directly to the router. If it does not know, it broadcasts the message.
- The router receiving the address-mask-request message responds with an address-mask-reply message, providing the necessary mask for the host. This can be applied to its full IP address to get its subnet address.

Router Solicitation and Advertisement

- As we discussed in the redirection message section, a host that wants to send data to a host on another network needs to know the address of routers connected to its own network.
- Also, the host must know if the routers are alive and functioning. The router-solicitation and router-advertisement messages can help in this situation.
- A host can broadcast (or multicast) a router-solicitation message. The router or routers that receive the solicitation message broadcast their routing information using the router-advertisement message.
- A router can also periodically send router-advertisement messages even if no host has solicited. Note that when a router sends out an advertisement, it announces not only its own presence but also the presence of all routers on the network of which it is aware.

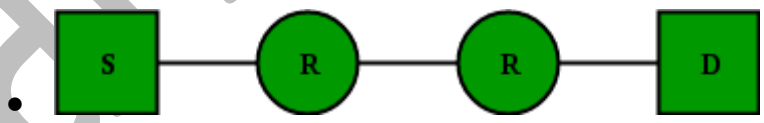
IGMP

- The IP protocol can be involved in two types of communication: unicasting and multicasting. Unicasting is the communication between one sender and one receiver. It is a one-to-one communication.
- However, some processes sometimes need to send the same message to a large number of receivers simultaneously. This is called multicasting, which is a one-to-many communication.
- Multicasting has many applications. For example, multiple stockbrokers can simultaneously be informed of changes in a stock price, or travel agents can be informed of a plane cancellation. Some other applications include distance learning and video-on-demand. The Internet Group Management Protocol (IGMP) is one of the necessary, but not sufficient (as we will see Internet Protocol version 4 (IPv4), is responsible for packetizing, forwarding, and delivery of a packet at the network layer.

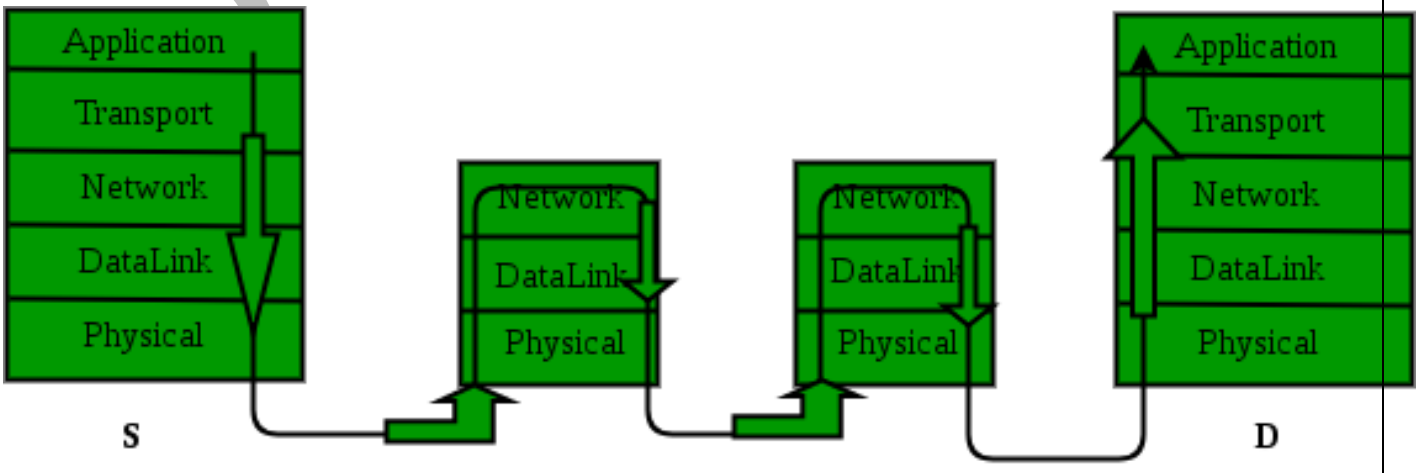
- The Internet Control Message Protocol version 4 (ICMPv4) helps IPv4 to handle some errors that may occur in the network-layer delivery.

- The Internet Group Management Protocol (IGMP) is used to help IPv4 in multicasting.

- The Address Resolution Protocol (ARP) is used to glue the network and data-link layers in mapping network-layer addresses to link-layer addresses.
- **Example:** Assume that source S and destination D are connected through two intermediate routers labelled R. Determine how many times each packet has to visit the network layer and the data link layer during a transmission from S to D.



- Ans. 4 times Network layer and 6 times data link layer.



NETWORK-LAYER PERFORMANCE

The performance of a network can be measured in terms of *delay, throughput,* and *packet loss.*

Delay: Total delay is similar to the delays in data link layers.

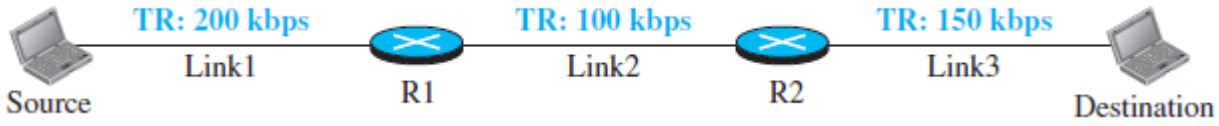
$$\textbf{\textit{Total Delay} = } T_t + T_p + T_{que} + T_{proc}$$

Throughput

- Throughput at any point in a network is defined as the number of bits passing through the point in a second.

- In a path from source to destination, a packet may pass through several links (networks), each with a different transmission rate.

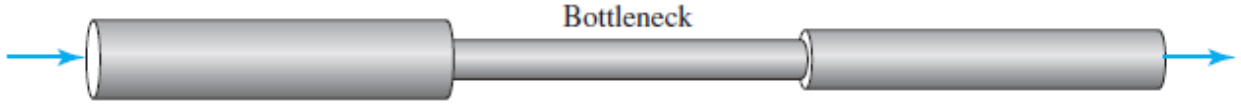
Example: To identify the throughput consider the example, where we have three links, each with a different transmission rate:



- The data can flow at the rate of 200 kbps in Link1.

- When the data arrives at router R1, it cannot pass at this rate. Data needs to be queued at the router and sent at 100 kbps.

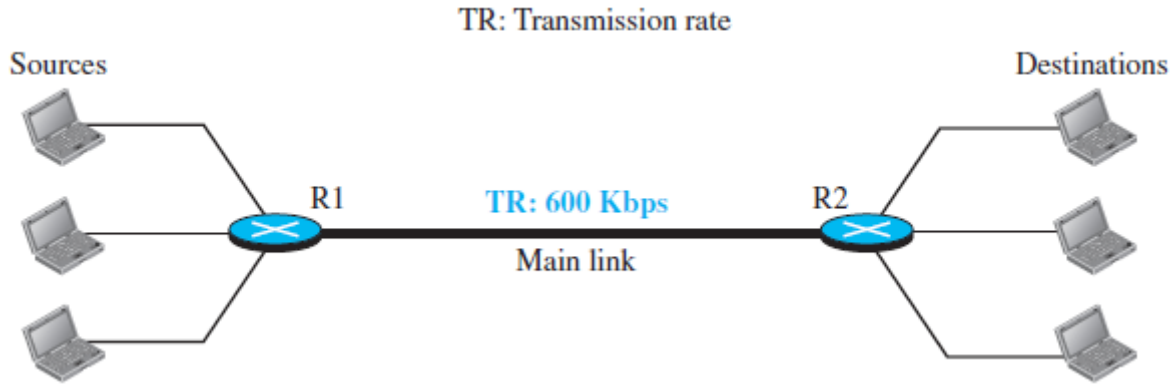
- When data arrives at router R2, it could be sent at the rate of 150 kbps, but there is not enough data to be sent, the average rate of the data flow in Link3 is also going to be 100 kbps, due to ***bottlenecking***.



- So, the Throughput is: ***min {TR_1, TR_2, TR3 …TRn},*** where TR is transmission rate of different links.

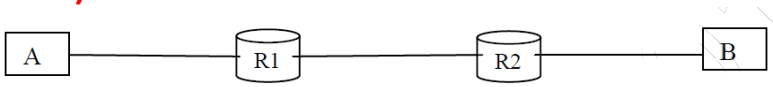
<u>Effective Throughput in shared links</u>

Consider the figure below:



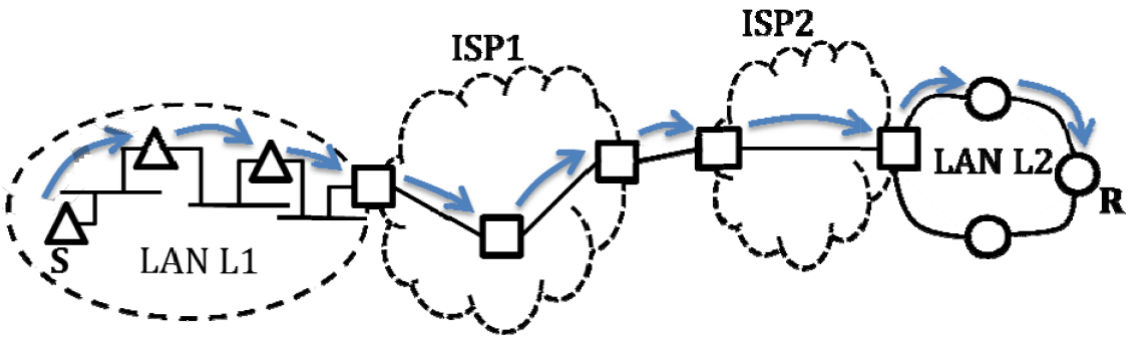
- In the figure above the transmission rate of the main link in the calculation of the throughput is only 200 kbps because the link is shared between three paths.

Q Consider the store and forward packet switched network given below. Assume that the bandwidth of each link is 10^6 bytes / sec. A user on host A sends a file of size 10^3 bytes to host B through routers R_1 and R_2 in three different ways. In the first case a single packet containing the complete file is transmitted from A to B. In the second case, the file is split into 10 equal parts, and these packets are transmitted from A to B. In the third case, the file is split into 20 equal parts and these packets are sent from A to B. Each packet contains 100 bytes of header information along with the user data. Consider only transmission time and ignore processing, queuing and propagation delays. Also assume that there are no errors during transmission. Let T_1, T_2 and T_3 be the times taken to transmit the file in the first, second and third case respectively. Which one of the following is CORRECT? **(Gate-2014) (2 Marks)**



(A) $T_1 < T_2 < T_3$
(B) $T_1 > T_2 > T_3$
(C) $T_2 = T_3, T_3 < T_1$
(D) $T_1 = T_3, T_3 > T_2$
Answer: (D)

Q In the diagram shown below L_1 is an Ethernet LAN and L_2 is a Token-Ring LAN. An IP packet originates from sender S and traverses to R, as shown. The link within each ISP, and across two ISPs, are all point to point optical links. The initial value of TTL is 32. The maximum possible value of TTL field when R receives the datagram is **(Gate-2014) (1 Marks)**



(A) 25 **(B)** 24 **(C)** 26 **(D)** 28
Answer: (C)

Q A link of capacity 100 Mbps is carrying traffic from a number of sources. Each source generates an on-off traffic stream; when the source is on, the rate of traffic is 10 Mbps, and when the source is off, the rate of traffic is zero. The duty cycle, which is the ratio of on-time to off-time, is 1 : 2. When there is no buffer at the link, the minimum number of sources that can be multiplexed on the link so that link capacity is not wasted and no data loss occurs is S_1. Assuming that all sources are synchronized and that the link is provided with a large buffer, the maximum number of sources that can be multiplexed so that no data loss occurs is S_2. The values of S_1 and S_2 are, respectively, **(Gate-2006) (2 Marks)**
(A) 10 and 30 **(B)** 12 and 25 **(C)** 5 and 33 **(D)** 15 and 22
Answer: (A)

Q In a communication network, a packet of length L bits takes link L_1 with a probability of p_1 or link L_2 with a probability of p_2. Link L_1 and L_2 have bit error probability of b_1 and $b_2$respectively. The probability that the packet will be received without error via either L_1 or L_2 is **(Gate-2005) (2 Marks)**
(A) $(1 – b_1)^L p1 + (1 – b_2)^L p_2$ **(B)** $[1 – (b_1 + b_2)^L]p_1p_2$
(C) $(1 – b_1)^L (1 – b_2)^L p_1p_2$ **(D)** $1 – (b_1^L p_1 + b_2^L p_2)$
Answer: (A)

Q In a packet switching network, packets are routed from source to destination along a single path having two intermediate nodes. If the message size is 24 bytes and each packet contain a header of 3 bytes, then the optimum packet size is: **(Gate-2005) (2 Marks)**
(a) 4 **(b)** 6 **(c)** 7 **(d)** 9
Answer (d)

Q The address resolution protocol (ARP) is used for **(Gate-2005) (1 Marks)**
(A) Finding the IP address from the DNS
(B) Finding the IP address of the default gateway

(C) Finding the IP address that corresponds to a MAC address
(D) Finding the MAC address that corresponds to an IP address
Answer: (D)

Q Consider three IP networks A, B and C. Host HA in network A sends messages each containing 180 bytes of application data to a host HC in network C. The TCP layer prefixes a 20-byte header to the message. This passes through an intermediate network B. The maximum packet size, including 20 byte IP header, in each network is
A : 1000 bytes
B : 100 bytes
C : 1000 bytes
The network A and B are connected through a 1 Mbps link, while B and C are connected by a 512 Kbps link (bps = bits per second).

Assuming that the packets are correctly delivered, how many bytes, including headers, are delivered to the IP layer at the destination for one application message, in the best case ? Consider only data packets. **(Gate-2004) (2 Marks)**
(A) 200 **(B)** 220 **(C)** 240 **(D)** 260
Answer: (D)

Q What is the rate at which application data is transferred to host HC? Ignore errors, acknowledgements, and other overheads. **(Gate-2004) (2 Marks)**
(A) 325.5 Kbps **(B)** 354.5 Kbps **(C)** 409.6 Kbps **(D)** 512.0 Kbps
Answer: (B)

Q Traceroute reports a possible route that is taken by packets moving from some host A to some other host B. Which of the following options represents the technique used by traceroute to identify these hosts **(GATE-2005) (1 Marks)**
(A) By progressively querying routers about the next router on the path to B using ICMP packets, starting with the first router
(B) By requiring each router to append the address to the ICMP packet as it is forwarded to B. The list of all routers en-route to B is returned by B in an ICMP reply packet
(C) By ensuring that an ICMP reply packet is returned to A by each router en-route to B, in the ascending order of their hop distance from A

(D) By locally computing the shortest path from A to B
Answer: (A)

Q Which of the following is NOT true with respect to a transparent bridge and a router? **(Gate-2004) (1 Marks)**
(A) Both bridge and router selectively forward data packets
(B) A bridge uses IP addresses while a router uses MAC addresses
(C) A bridge builds up its routing table by inspecting incoming packets
(D) A router can connect between a LAN and a WAN
Answer: (B)

Q Every host in an IPv4 network has a 1-second resolution real-time clock with battery backup. Each host needs to generate up to 1000 unique identifiers per second. Assume that each host has a globally unique IPv4 address. Design a 50-bit globally unique ID for this purpose. After what period (in seconds) will the identifiers generated by a host wrap around? **(Gate-2014) (1 Marks)**

Q For a host machine that uses the token bucket algorithm for congestion control, the token bucket has a capacity of 1 megabyte and the maximum output rate is 20 megabytes per second. Tokens arrive at a rate to sustain output at a rate of 10 megabytes per second. The token bucket is currently full and the machine needs to send 12 megabytes of data. The minimum time required to transmit the data is _____ seconds. **(Gate-2017) (2 Marks)**

Q A computer on a 10Mbps network is regulated by a token bucket. The token bucket is filled at a rate of 2Mbps. It is initially filled to capacity with 16Megabits. What is the maximum duration for which the computer can transmit at the full 10Mbps? **(GATE-2008) (2 Marks)**
(A) 1.6 seconds **(B)** 2 seconds **(C)** 5 seconds **(D)** 8 seconds
Answer: (B)