MODULE-5

Internet Control Protocols – ICMP, ARP, RARP, BOOTP. Internet Multicasting – IGMP, Exterior Routing Protocols – BGP. IPv6 – Addressing – Issues, ICMPv6.

Internet control protocol

- IP protocol has some **deficiencies** :

  (1) IP protocol has no error-reporting or error correcting mechanism and

  (2) IP protocol also lacks a mechanism for host and management queries.

- Internet protocol (IP)
  - ➢ provides **unreliable** and **connectionless datagram delivery**
    - ✓ which means it has **no** error reporting or error correcting mechanism.
    - ✓ When **error happens**, router must discard the datagram.

- Therefore, in addition to IP, **internet control protocol** are used in **network layer** for flow control & error control.

Internet Control Protocols

- In addition to IP, **Internet Control Protocol** is used for data transfer.
- The Internet has several control protocols used in the **network layer**.
- They are

  1) ICMP (Internet control message protocol)

  2) ARP (Address resolution protocol)

  3) RARP (Reverse address resolution protocol)

  4) BOOTP (Bootstrap protocol)

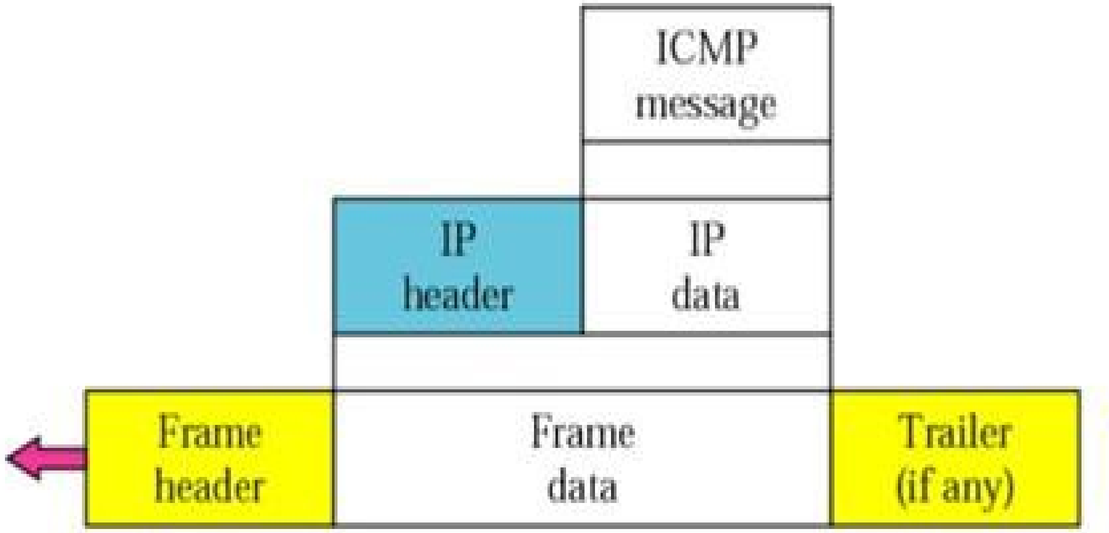
  5) DHCP (Dynamic host control protocol)

ICMP(Internet control message protocol)

(1) no error-reporting or error correcting mechanism and

(2) lacks a mechanism for host and management queries.

- **ICMP** has been designed to compensate the deficiencies of IP.
- **ICMP messages** are encapsulated within IP datagrams
- ICMP provides **error reporting, congestion reporting, and first-hop router redirection.**



Internet Control Message Protocol (ICMP)

- ICMP is **error-reporting protocol** network devices like routers use to generate error messages to the source IP address when the network problems prevent delivery of IP packets.
- Any **IP network device** has the capability to send, receive or process ICMP messages.
- When something unexpected occur, the **event** is reported by ICMP, which is also used to test the internet.
  - ➢ Therefore, ICMP is a companion to the IP protocol.

Internet Control Message Protocol (ICMP)

- ICMP messages are divided into two broad categories.

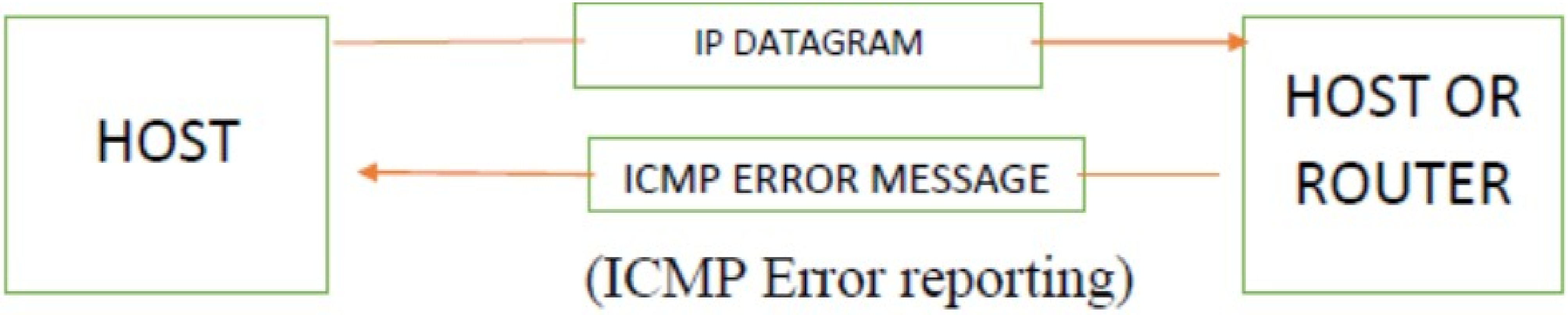
  1) Error-reporting message

  2) Query Message

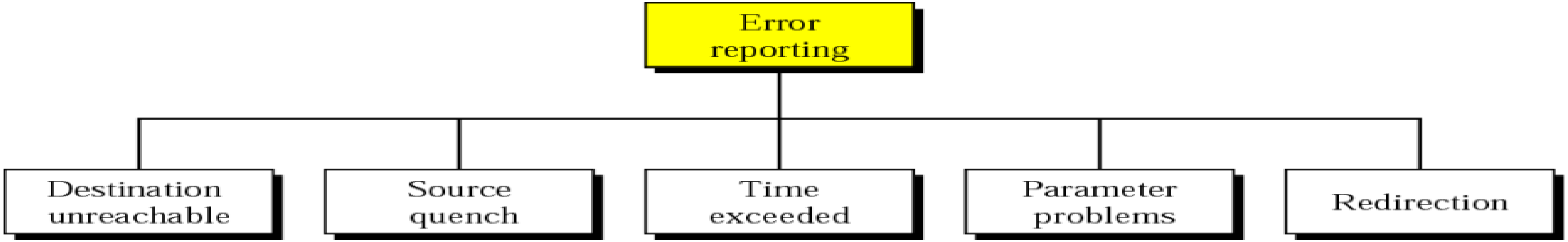
- **Error-reporting message :** report **problems** that a **router or a host(Destination)** may meet unexpected when it processes an IP packet.
- The **query messages**: help a **host or a network manager** to get specific information from a router or another host.
- E.g.: Query messages are used, if a node need **redirect** it message.

1) Error-reporting

- The main responsibility of ICMP is to report errors.
- **ICMP** doesn’t correct error-it simply report them (Error correction can be done by high-level protocol)
- ICMP always reports **error messages** to the original source.



(ICMP Error reporting)



Error reporting

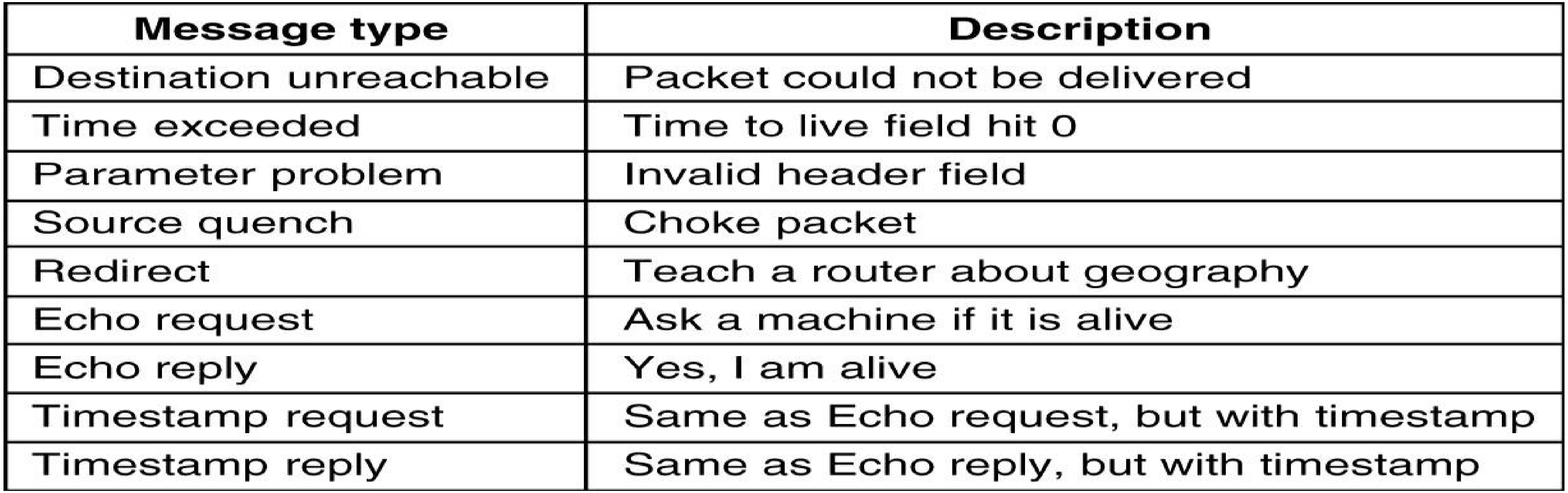
Destination unreachable

Source quench

Time exceeded

Parameter problems

Redirection

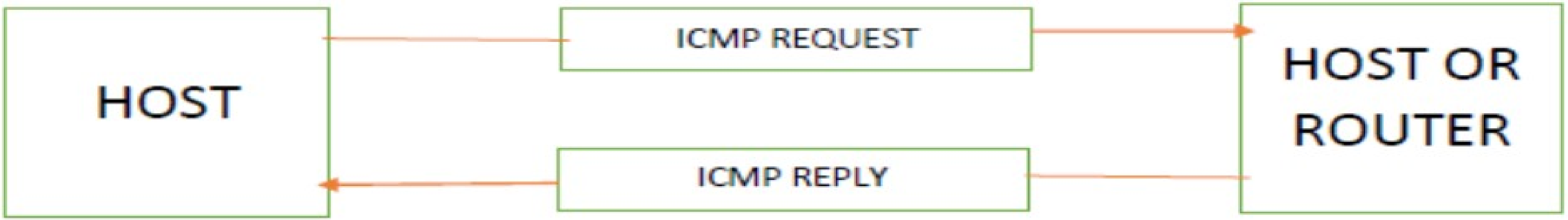


Message type	Description
Destination unreachable	Packet could not be delivered
Time exceeded	Time to live field hit 0
Parameter problem	Invalid header field
Source quench	Choke packet
Redirect	Teach a router about geography
Echo request	Ask a machine if it is alive
Echo reply	Yes, I am alive
Timestamp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

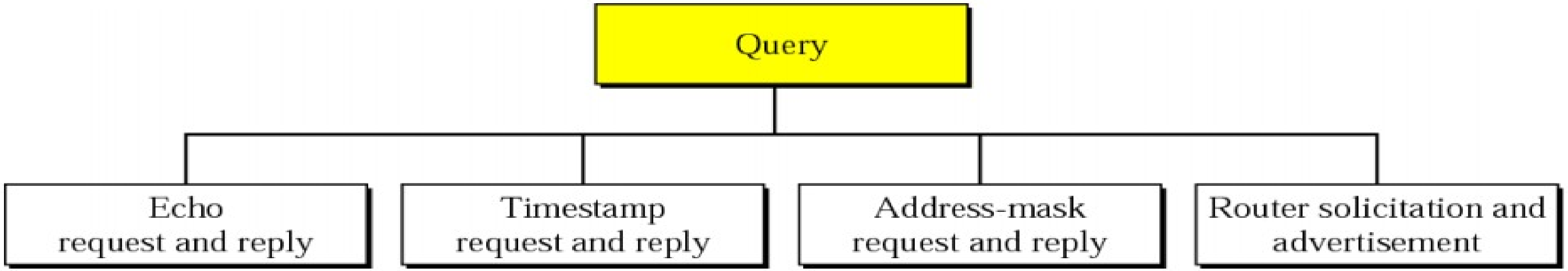
Figure 3-31. The principal ICMP message types.

2) Query Message

- ICMP can diagnose some network problems. This is accomplished through the **query messages**.
- In **query message**,
  - a **node** sends a message to **destination** and
  - an answer message in a specific format by **destination** to **source**.
- It is encapsulated in an **IP packet** (for transmission)



(ICMP query message)


Query

Echo request and reply

Timestamp request and reply

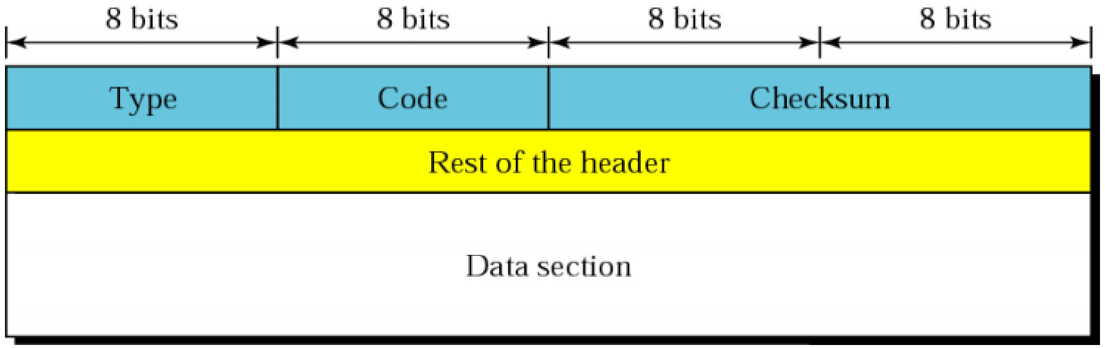
Address-mask request and reply

Router solicitation and advertisement

Message type	Description
Echo request	Ask a machine if it is alive
Echo reply	Yes, I am alive
Time stamp request	Same as echo request, but with time stamp
Time stamp reply	Same as echo reply, but with time stamp
Address-mask request and reply	To obtain the mask of IP address and reply provide the necessary mask for the host
Router solicitation	To know the address or routing information of router connected to its own network, by broadcasting router solicitation message
Router advertisement	Reply for router solicitation message broadcast routing information using this message.

ICMP Message Format

- ICMP messages are send in **IP datagram**.
- The **IP header** will always have a protocol number of 1, indicating **ICMP** and a type of service of zero.
- The **IP data field** will contain the actual ICMP message in the format shown in the figure below.



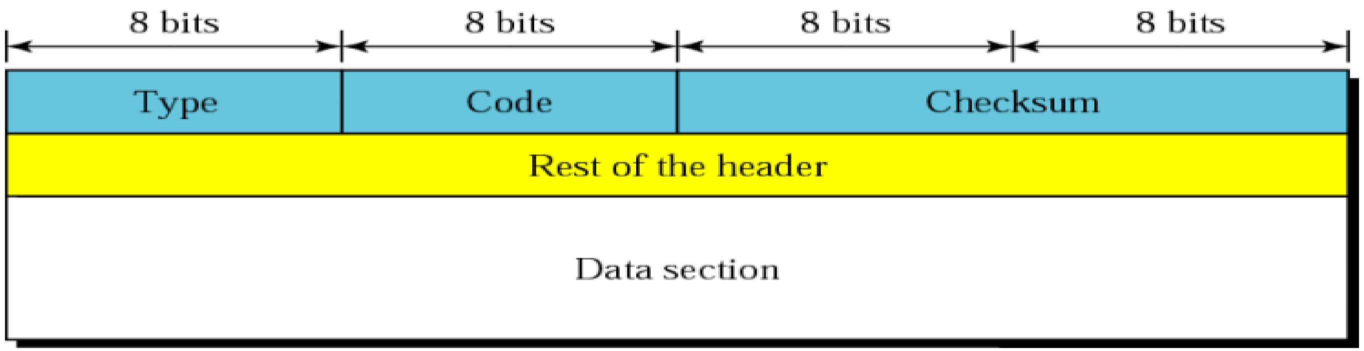
Type: This is an 8bit filed defining the type of ICMP message.

Code: The code field specifies the reason for the particular message type.

Checksum: Error calculation, in ICMP the checksum is calculated over the entire message (Header and Data)

Rest of the header: Depends on different messages (specific for each message type)

The data section: the data section in error message carries information for finding the original packet that had the error. In query messages the data section carries extra information based on type of query.



8 bits | 8 bits | 8 bits | 8 bits

Type | Code | Checksum

Rest of the header

Data section

Category	*Type*	*Message*
Error-reporting messages	3	Destination unreachable
	4	Source quench
	11	Time exceeded
	12	Parameter problem
	5	Redirection
Query messages	8 or 0	Echo request or reply
	13 or 14	Timestamp request or reply
	17 or 18	Address mask request or reply
	10 or 9	Router solicitation or advertisement

Address

- Two types of address for computer Network:
  1. IP address (logical address)
  2. MAC address (physical address)
- At the **physical level**,
  - ➢ the IP address is not useful
  - ➢ the hosts and routers are recognized by their MAC address.
- A **MAC** address is a local address.
- The IP and MAC address are two different identities and both of them are need.
- An example of **physical address** is the 48-bit MAC address.
  - ➢ In the ETHERNET protocol, which is imprinted on the NIC in the host or router.

Address

- **Data link layer** protocol use physical address
- **Network layer** use logical address.
- Eg : Ethernet or LAN have two **different protocols**
    - ➢ at network layer IP
    - ➢ data link layer Ethernet protocol.
- This means that delivery of a packet to a host or router require two level of addressing, logical and physical addressing.
- We need to able to map a logical address to its corresponding physical address and vice-versa.

Address Mapping

1. Static mapping

2. Dynamic mapping

1. Static mapping: A **table** is created and stored in each machine.

- This table is associates an **IP address** with a **MAC address**.
- If a machine knows IP address of another machine, then it can search for corresponding **MAC address** in its table.
- The limitation of statically mapping is that the MAC address can change.
- To implement statically mapping, the static mapping table need to be updated periodically.

2. Dynamic Mapping: A **protocol** is used for finding the other address.

- There are two protocols designed to perform the dynamic mapping.

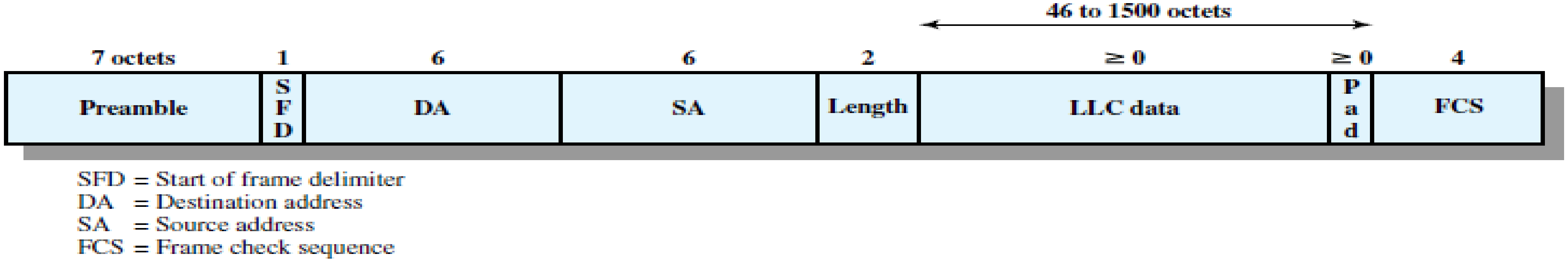
1. **ARP** (Address resolution protocol)

2. **RARP** (Reverse address resolution protocol)

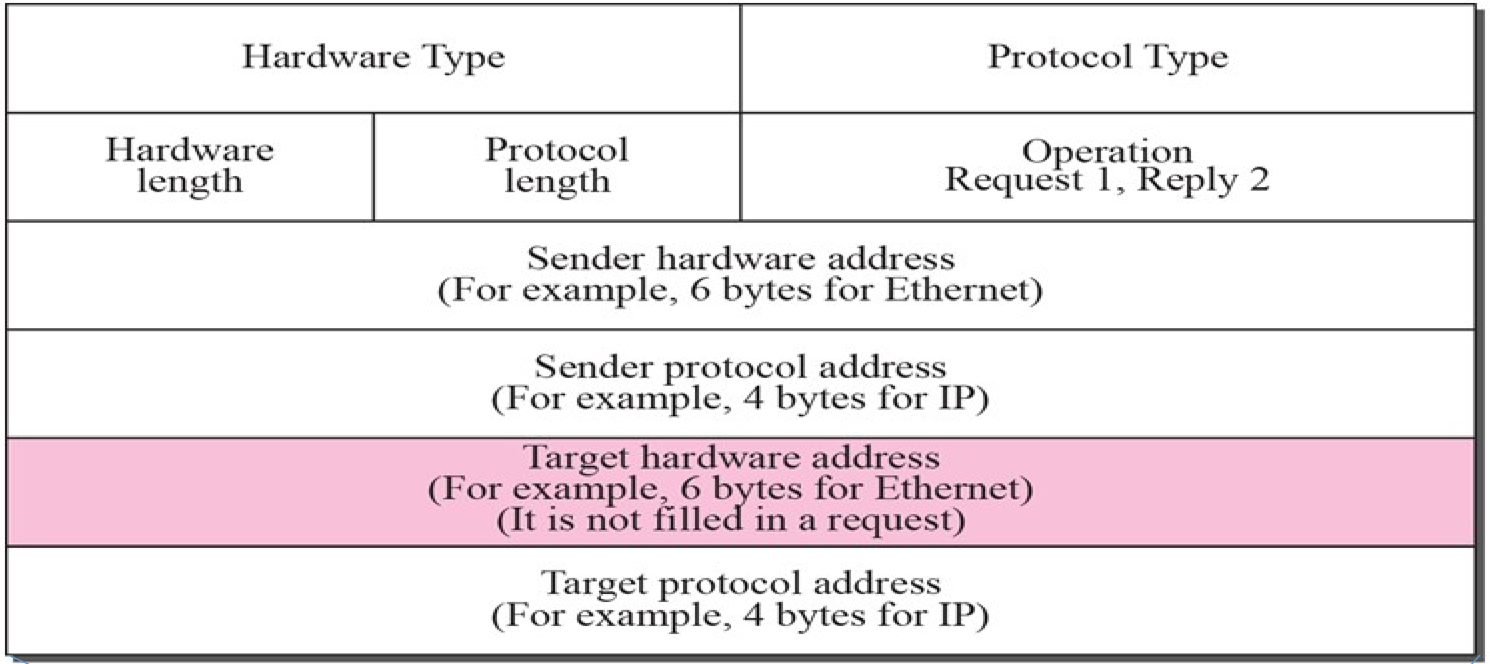
Ethernet -802.3 Frame Format

- **PREAMBLE:** alternating 0s and 1s that alerts the receiving system to coming frame and enables it to **synchronize** its input timing.
- **START FRAME DELIMITER(SFD)**: (1byte:10101011)signals the **beginning** of the frame .
  - The last 2 bits is11 and alerts the receive that the next field is the destination address.
- **DESTINATION ADDRESS(DA):** physical address of the **destination** station
- **SOURCE ADDRESS** : physical address of the **sen**

- **LENGTH/TYPE** : Ethernet used this field to define the **upper –layer protocol** using the MAC frame.
- **DATA** : This field carries **data encapsulated** from the upper–layer protocols.
- **FCS : Frame Checksum sequence** The last filed contains **error detection** information.

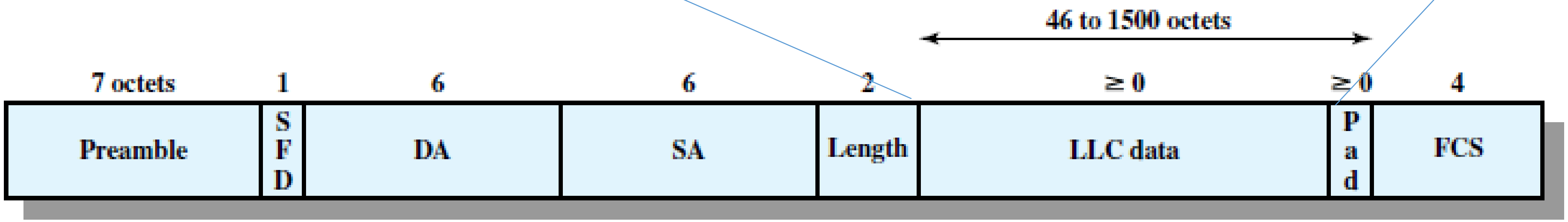


IEEE 802.3 Frame Format



Hardware Type		Protocol Type
Hardware length	Protocol length	Operation Request 1, Reply 2
Sender hardware address (For example, 6 bytes for Ethernet)		
Sender protocol address (For example, 4 bytes for IP)		
Target hardware address (For example, 6 bytes for Ethernet) (It is not filled in a request)		
Target protocol address (For example, 4 bytes for IP)		

Fig: ARP packet format



7 octets	1	6	6	2	≥ 0	≥ 0	4
Preamble	SFD	DA	SA	Length	LLC data	Pad	FCS

SFD = Start of frame delimiter
DA = Destination address
SA = Source address
FCS = Frame check sequence

Ethernet/IEEE 802.3 Frame Format

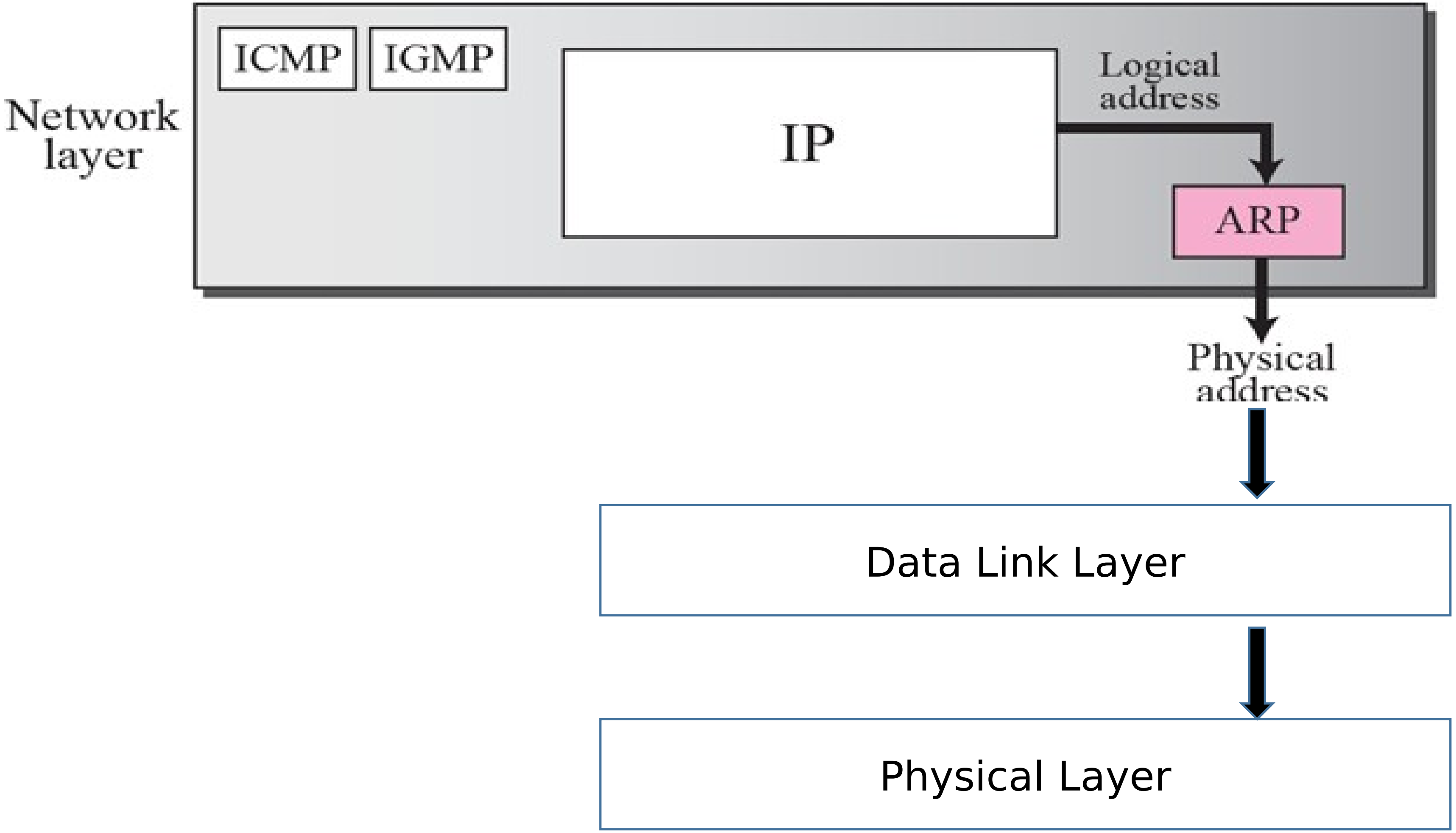
ARP– The Address Resolution Protocol

- Anytime a **host or a router** has an IP datagram to send to another host or router, it has the **logical (IP) address** of the receiver.
- But the **IP datagram** must be encapsulated in a **frame** to be able to pass through the **physical network**.
- This means that the sender needs the **physical(MAC) address** of the receiver.
- A mapping corresponds a logical address to a physical address.
- **ARP** accepts a logical address from the **IP protocol** maps the address to the corresponding physical address and passes it to the data link layer.

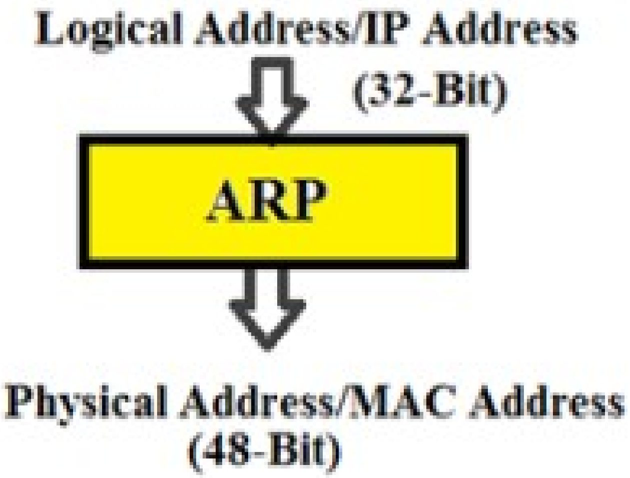
Accept logical address

- ARP □ Physical address □ passes to data link layer

ARP □ Accept logical (IP) address Physical(MAC) address □ passes to data link layer



ARP Operation:



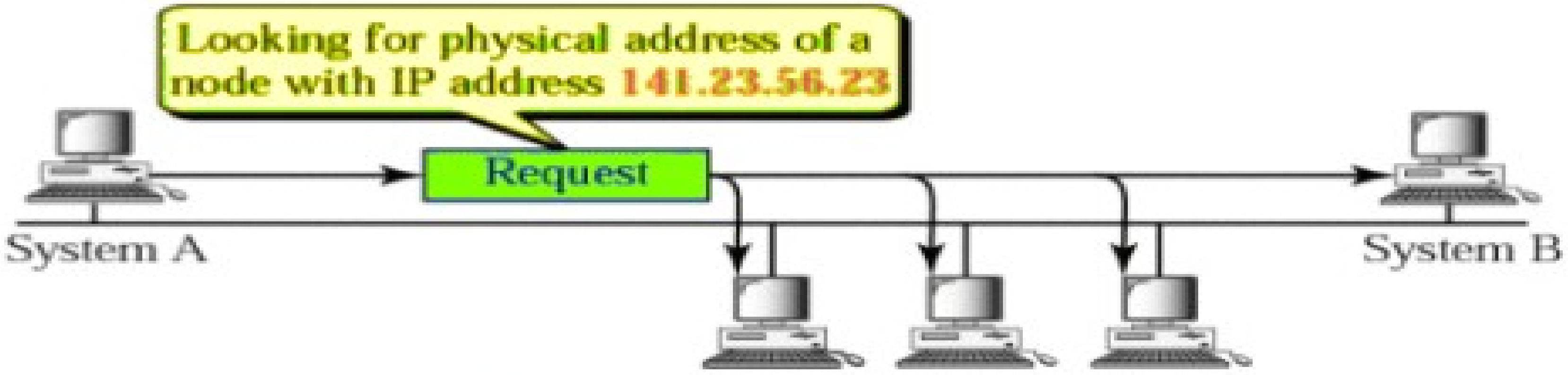

Looking for physical address of a node with IP address 141.23.56.23
Request
System A
System B

Fig: ARP request is broadcast

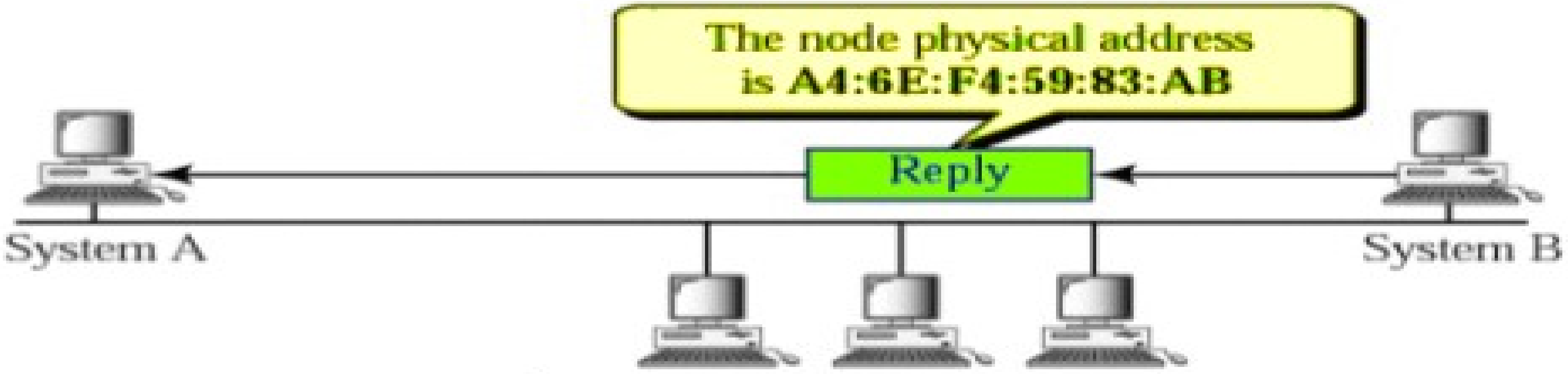
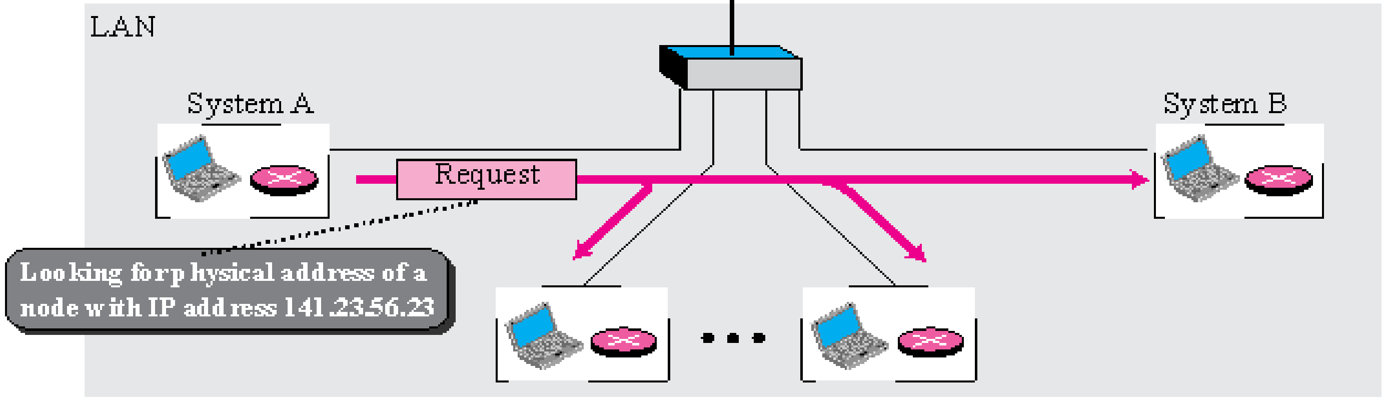


Fig: ARP reply is unicast

Mapping logical to physical address: ARP (Address resolution protocol)

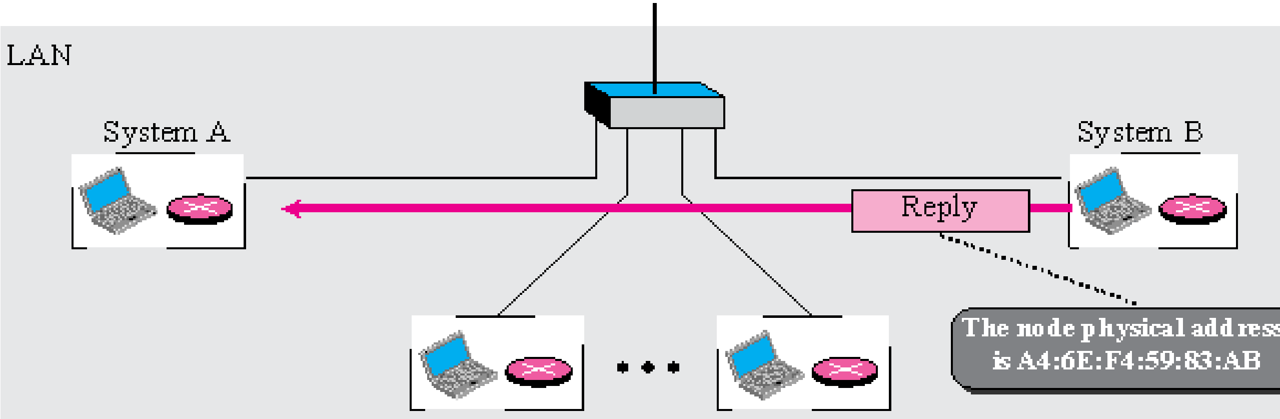
- **ARP** is used to mapping logical to physical address mapping.
- The **router or host**, who wants to find the MAC address of some other router, sends an ARP request packet.
- **ARP request packet** consist of IP and MAC address of sender and IP address of receiver/destination.
- The **request packet** is broadcasted over the network.
- Every host and router on the network receives and processes the ARP request packet.
- But only the intended receiver recognizes its IP address in the request packet and send back an ARP response packet.
- **ARP response packet** contains the IP Physical address of the **receiver**.
- ARP response packet is delivered only to **sender(unicast)**using A’s physical address in the ARP request packet.

ARP :
(The Address Resolution


LAN
System A
System B
Request
Looking for physical address of a node with IP address 141.23.56.23

a. ARP request is broadcast

- A **system A** has a packet that needs to be delivered to another **system B** with **IP address** 141.23.56.23.
- **System A** needs to pass the packet to its data link layer for the actual delivery.
- But it does not know the physical address of the receipt.
- It uses the **service ARP** by asking the **ARP protocol** to send a broadcast ARP request to ask for the PA of the system with an IP address of 141…..
- This request is received by all the system but the original system will answer it**.**



b. ARP reply is unicast

ARP Packet Format

1.Hardware Type (16-bit field)

- Defining the type of the network on which ARP run.
- ARP can run on any physical network use Ethernet frame.
- H/w type: **Ethernet**

2. Protocol type

- Defining the Protocol using ARP.
- ARP can be used with any higher level protocol.
  - ➢ IP address used in higher level as protocol type

3. Hardware length (8-bit field)

- Used to define the length of physical address in bytes.

E.g.: Length is 6 bytes MAC address for **Ethernet.**

4. Protocol length

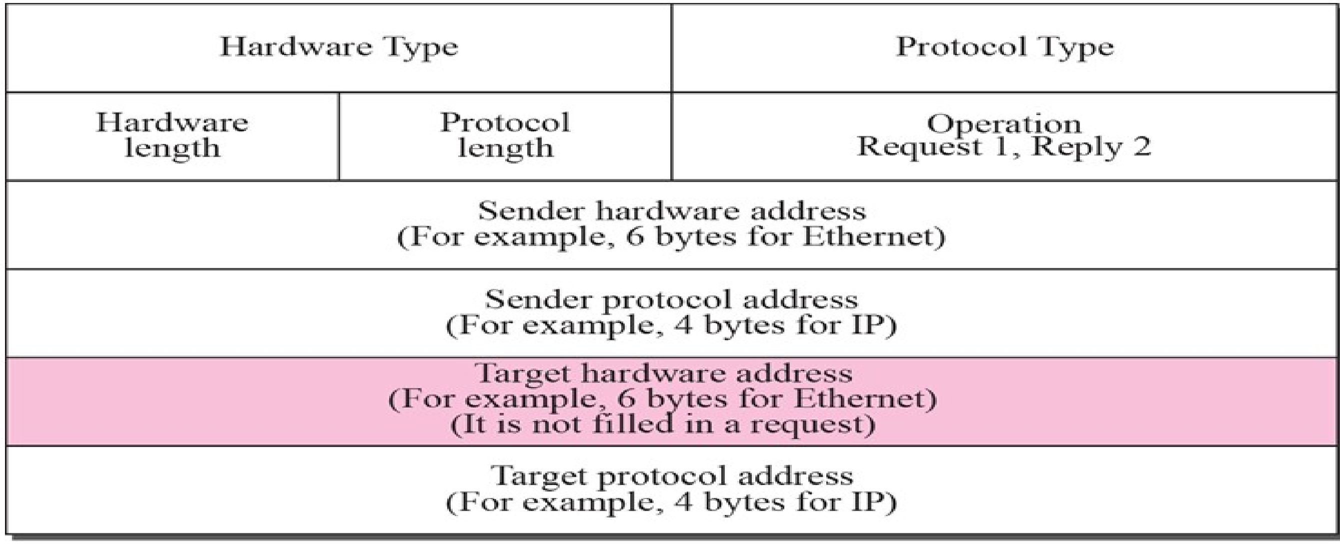
- Define the length of the IP address in bytes

E.g:IPV4-4 bytes(32 bits)

5. Operation

- Define the type of packet
- The possible type of packets are

1. ARP Request (field value-1)
2. ARP Reply (field value-2)



Hardware Type		Protocol Type
Hardware length	Protocol length	Operation Request 1, Reply 2
Sender hardware address (For example, 6 bytes for Ethernet)		
Sender protocol address (For example, 4 bytes for IP)		
Target hardware address (For example, 6 bytes for Ethernet) (It is not filled in a request)		
Target protocol address (For example, 4 bytes for IP)		

6. Sender hardware address

- Defining the physical address of the sender.
- MAC address of source

7. Sender protocol address

- Defining the logical address of sender.
- IP address of source

8 . Target Hardware address

- Define the physical/MAC address of the target.
- For **ARP request packet**,
  - ➢ the field contains all zeros .
  - ➢ Because the sender doesn’t know the receivers physical address or MAC address.

9. Target protocol address:

RARP (Reverse Address Resolution Protocol)

- Mapping Physical to logical address:
  1. **RARP- Reverse address resolution protocol**
  2. **BOOTP - Bootstrap Protocol**
  3. **DHCP - Dynamic Host Control Protocol**
- There are occasions in which a **host** knows its physical address and unknowns its logical address.
- This may happen in two case.

1) A **diskless station** is just booted.

   - The situation can find its physical address by checking its **interface**, but it does not know its IP address.

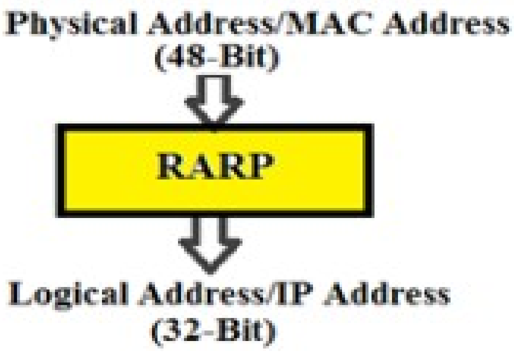
2) An **organization** doesn't have enough IP address to assign to **each station**

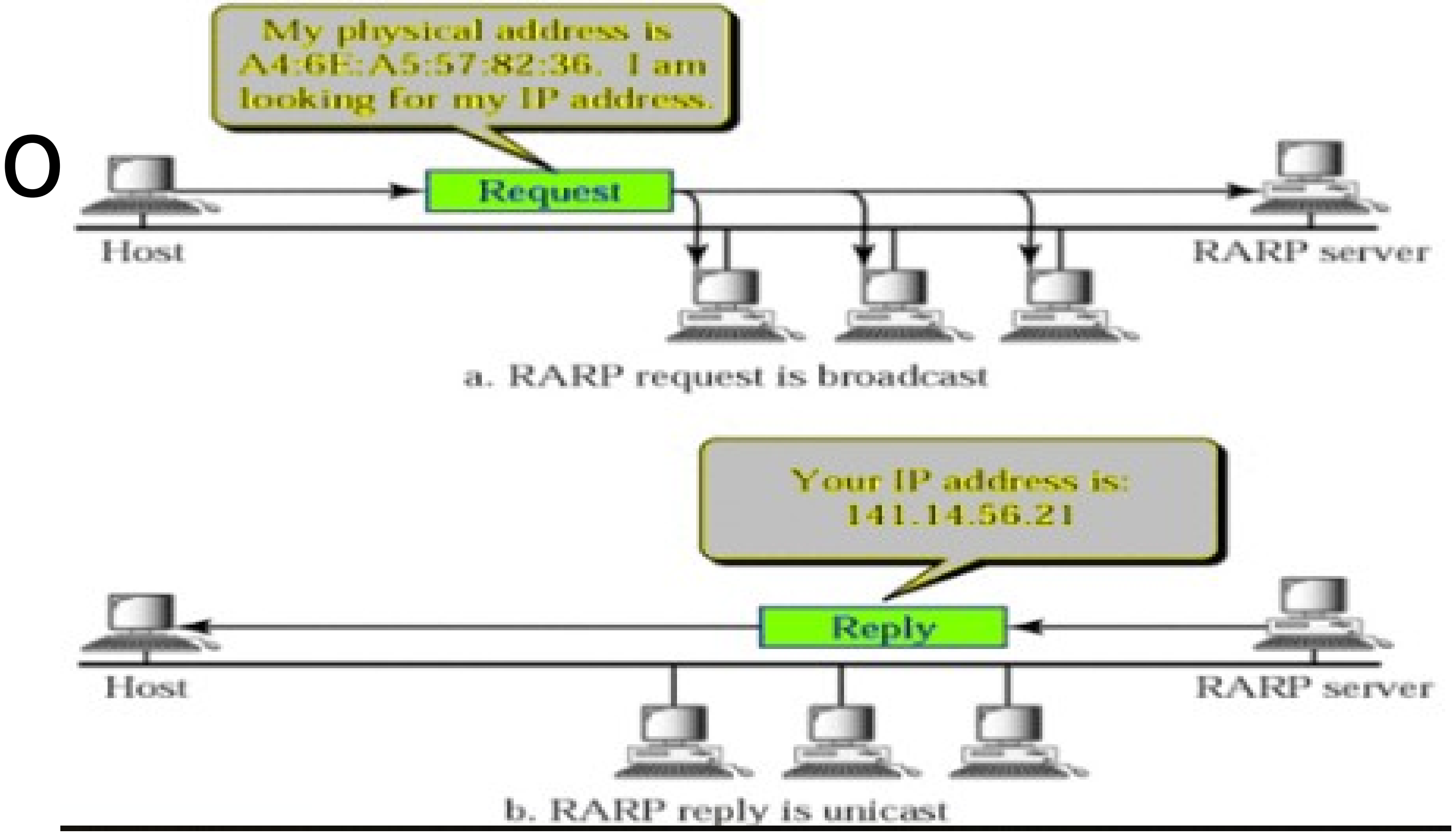
   - It needs to assign IP address on demand.
   - The station can send its physical addresses and ask for a short time lease.

RARP

- The **IP address** of a machine is usually reach from its configuration file stored on a **disk file**.
- A **diskless machine**
  - ➢ is usually booted from ROM, which has minimum booting information.
  - ➢ The **ROM** is installed by the manufacture.
  - ➢ It can’t include the IP address because IP address on a network are assigned by the **network administrator**.
  - ➢ The **machine** can get its physical address, which is unique locally (by reading its NIC).
  - ➢ It can then use the physical address to get the logical address by using the RARP protocol.

RARP operatio




My physical address is A4:6E:A5:57:82:36. I am looking for my IP address.
Request
Host
RARP server
a. RARP request is broadcast
Your IP address is: 141.14.56.21
Reply
Host
RARP server
b. RARP reply is unicast

- 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 and the **responding machine** must be running a RARP server program

Problem of RARP

- Broadcasting is done at the **data link layer**.
- The **physical broadcast address** (all 1’s in the case of ETHERNET) doesn’t pass the boundaries of network.
- This means that if an **administrator** has several networks or several subnets it need to assign a **RARP server** for each network or subnet.
  - ➢ This is the reason that RARP is almost outdated.
- Two protocols are commonly used for replacing RARP

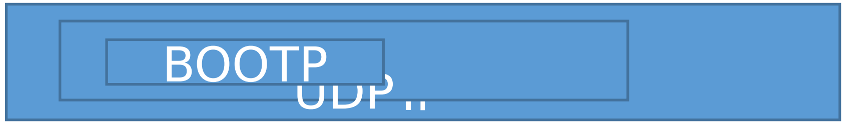
1) BOOTP

2) DHCP

BOOTP (Bootstrap Protocol)

BOOTP is a **client/server protocol** designed to provide physical address to logical address mapping.

- BOOTP is an **application layer protocol**, administrator may put the client and server on the same network or on different network.
- **BOOTP message** are encapsulated in a **UDP packet**, and the UDP packet itself encapsulated in an **IP packet**.

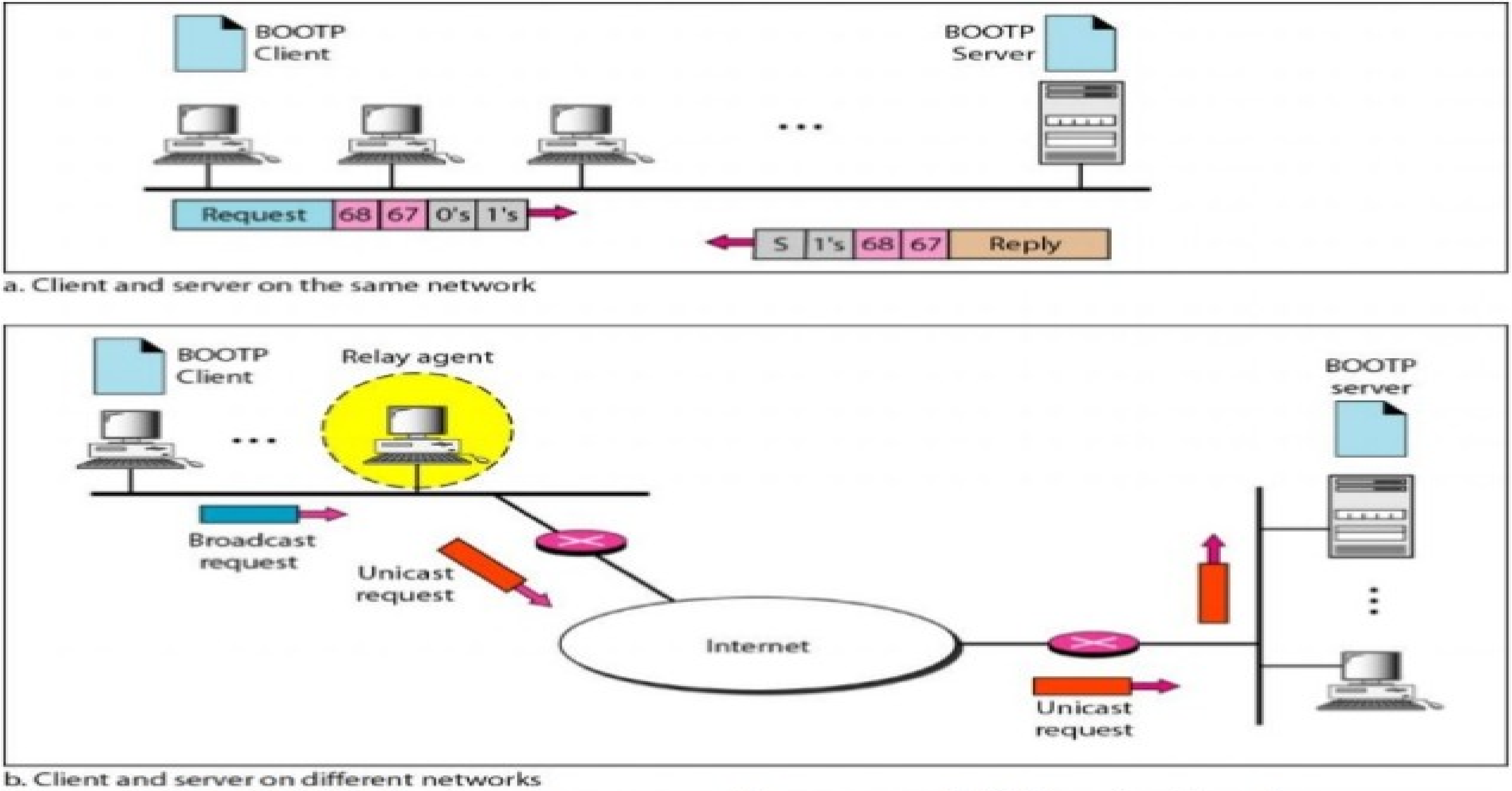


- The **client** may unknown about IP address, but it need to send IP datagram.
- The **client** simply uses all 0's as the **source address** and all 1's as the **destination address**.
- One of the advantage of BOOTP over RARP is that the client and server are application layer processes.

In client and server on different network:

- The **BOOTP request** is broadcast because the client doesn’t know the IP address of server.
- A broadcast **IP datagram** cannot pass through any router.
- So there is a need for an intermediary.
- One of the **host** can be used as a **relay (Relay agent)**
- The **relay agent** know the unicast address of **BOOTP server**.
- When relay agent receives BOOTP request packet, it encapsulates the message in a **unicast datagram** and send the request to the **BOOTP server**.
- **BOOTP server** know the message comes from a relay agent because one of the **field** in the request message define the IP address of relay agent.
- The relay agent, **after receiving reply**, send it to BOOTP client.

BOOTP (Bootstrap Protocol)



a. Client and server on the same network

b. Client and server on different networks

Figure 3.34 BOOTP client and server on the same and different networks

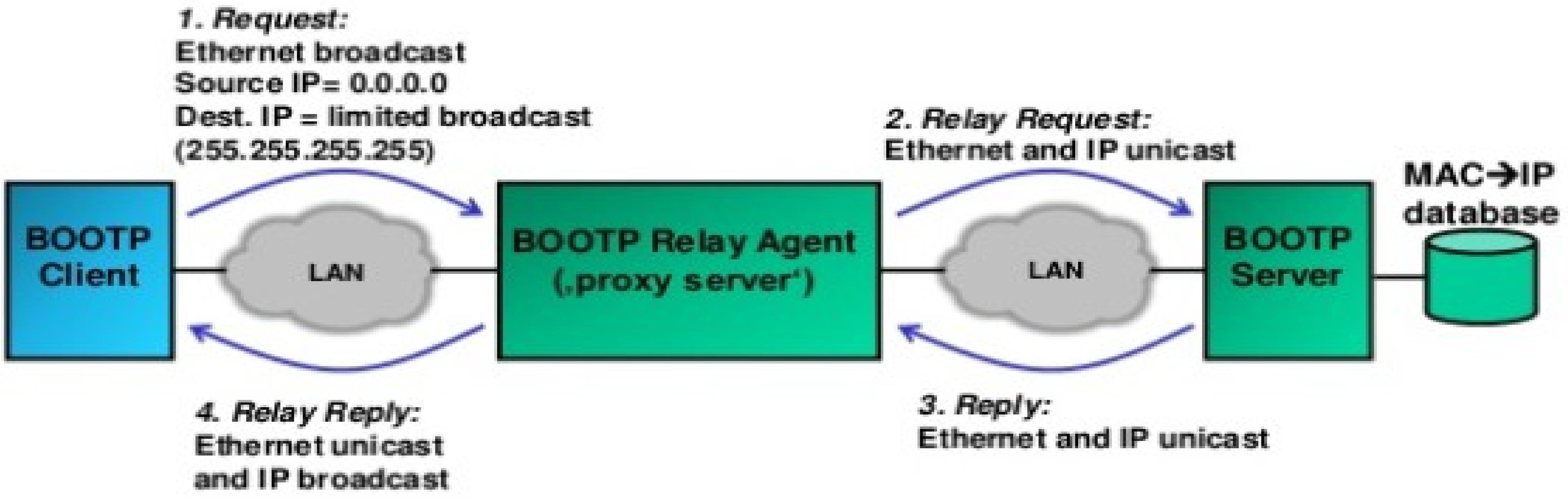
BOOTP BOOTstrap Protocol RFC951

BOOTP relay:

With BOOTP relay a single BOOTP server services BOOTP requests from multiple LAN segments. This may be used to ease administration in large networks (operate only one single BOOTP server).

The BOOTP relay agent forwards (unicast) BOOTP requests to a pre-configured BOOTP server.

To reduce broadcasts in the networks, the BOOTP reply is sent back as an Ethernet unicast addressed packet.



DHCP Dynamic Host Configuration Protocol RFC2131

Differences between BOOTP and DHCP:

DHCP is an evolution of BOOTP. DHCP was defined to fix the limitations of BOOTP.

1. IP address lease (time):

BOOTP assigns IP addresses without a lease time.
In order to free an IP address assigned by BOOTP, a host needs to be rebooted.
In DHCP an IP address is assigned to a client only for a limited (but extensible) time. If the IP address is not renewed it goes back to the pool of free IP addresses.
Typically a client renews a lease when half of the lease time has elapsed.

2. Vendor options:

In DHCP vendor specific options are no longer restricted to 64 bytes.
This gives more possibilities to carry vendor specific information such as configuration files from DHCP server to client.

DHCP Dynamic Host Configuration Protocol RFC2131

BOOTP and DHCP interoperability:

BOOTP and DHCP are interoperable because DHCP is basically an extension of BOOTP.

a. Same ports:

Both BOOTP and DHCP use UDP ports 68 (client) and 67 (server). BOOTP and DHCP use 2 different ports for client and server so that clients do not receive messages to servers and vice versa (DHCP and BOOTP use UDP thus source IP and port number can not be used to distinguish a session or connection).

b. Same packet format:

BOOTP and DHCP use the same packet format. Thus a DHCP server can receive BOOTP packets and vice versa.

c. DHCP transaction on top of BOOTP request-reply:

DHCP uses a 4-way transaction for assigning IP addresses, BOOTP only knows Request and Reply. Therefore the DHCP request packets (client to server) are mapped to BOOTPREQUEST and DHCP server responses are mapped to BOOTPREPLY (every DHCP message is either a BOOTPREQUEST or BOOTPREPLY message).

- DHCP server may not be reachable by broadcasting, a DHCP relay agent is needed on each LAN, as shown in Fig.

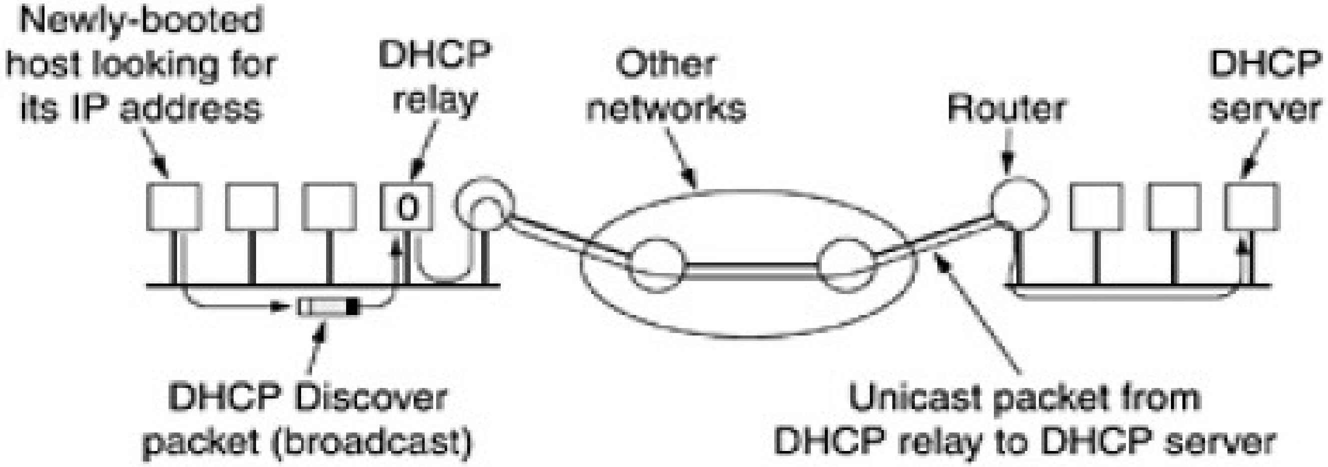
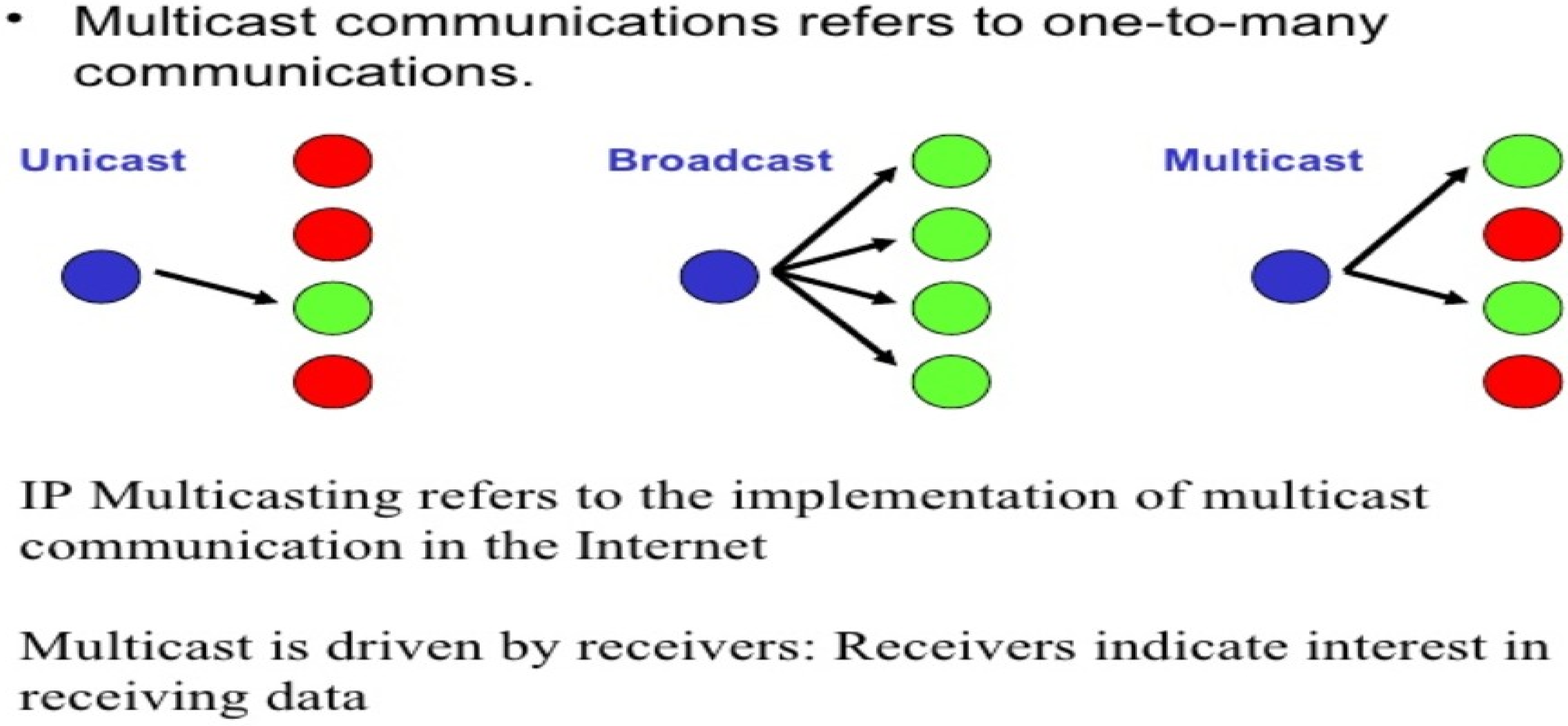

Newly-booted host looking for its IP address
DHCP relay
0
Other networks
Router
DHCP server
DHCP Discover packet (broadcast)
Unicast packet from DHCP relay to DHCP server

Figure. Operation of DHCP.

Internet Multicasting

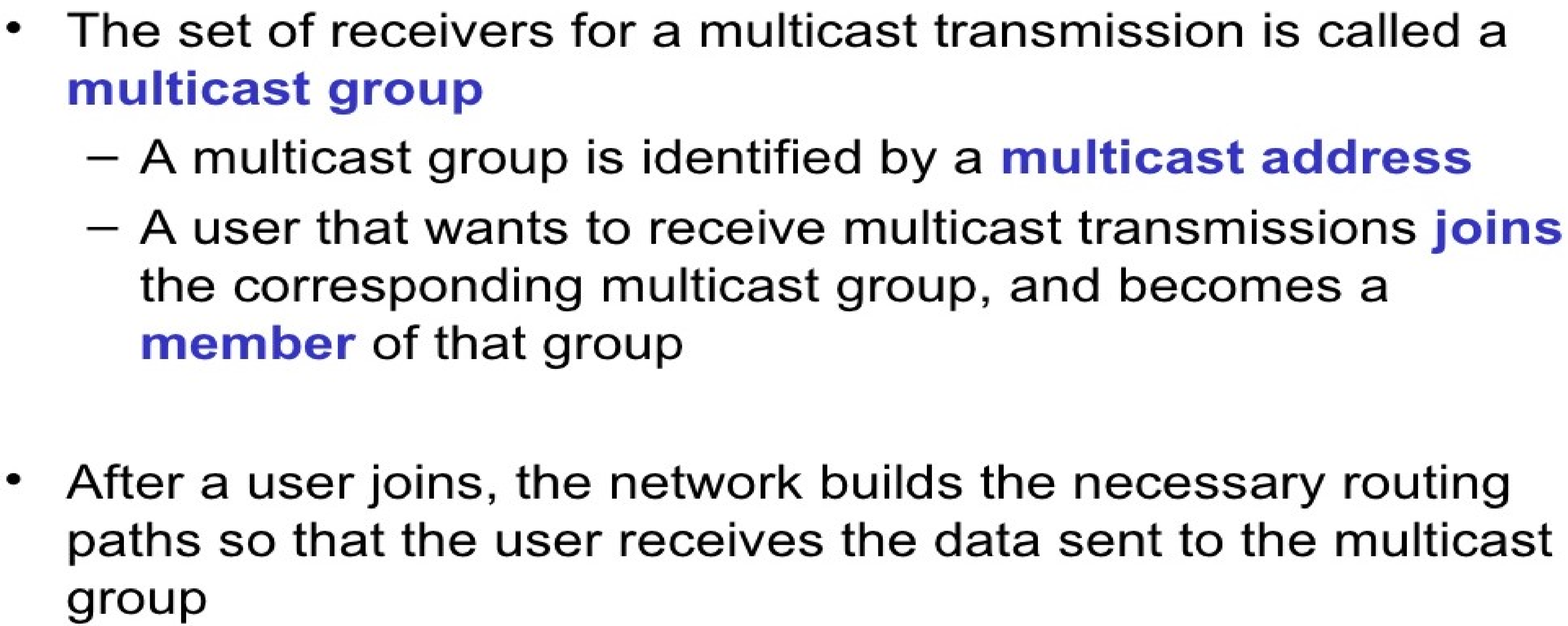


- Multicast communications refers to one-to-many communications.

IP Multicasting refers to the implementation of multicast communication in the Internet

Multicast is driven by receivers: Receivers indicate interest in receiving data

Multicast group



- The set of receivers for a multicast transmission is called a **multicast group**
  - A multicast group is identified by a **multicast address**
  - A user that wants to receive multicast transmissions **joins** the corresponding multicast group, and becomes a **member** of that group

- After a user joins, the network builds the necessary routing paths so that the user receives the data sent to the multicast group

Internet Multicasting protocols- IGMP

- In **multicasting communication** there is **one source** and **more than one destinations**
- **Communication protocol** used by host and adjacent routers on **IPv4** networks to establish **multicast group membership**.
- IGMP is defined in RFC 1112.
- **IGMP** operates on a physical network (e.g., single Ethernet Segment.)
- The **IGMP protocol** gives the **multicast routers** information about the membership status of hosts (routers) connected to the network.
- Support for:
  - **Joining a multicast group**
  - **Query membership**
  - **Send membership reports**

IGMP- Internet Group Management Protocol

- A **host** sends an IGMP report when it joins a multicast group.
- Note: multiple processes on a **host** can join. A **report** is sent only for the first process.
- No report is sent when a **process leaves a group**.
- A **multicast router** regularly multicasts an **IGMP query** to all **hosts** (group address is set to zero).

- A **host** responds to an IGMP query w
- **N**
  tl
  jo
- T
  tl

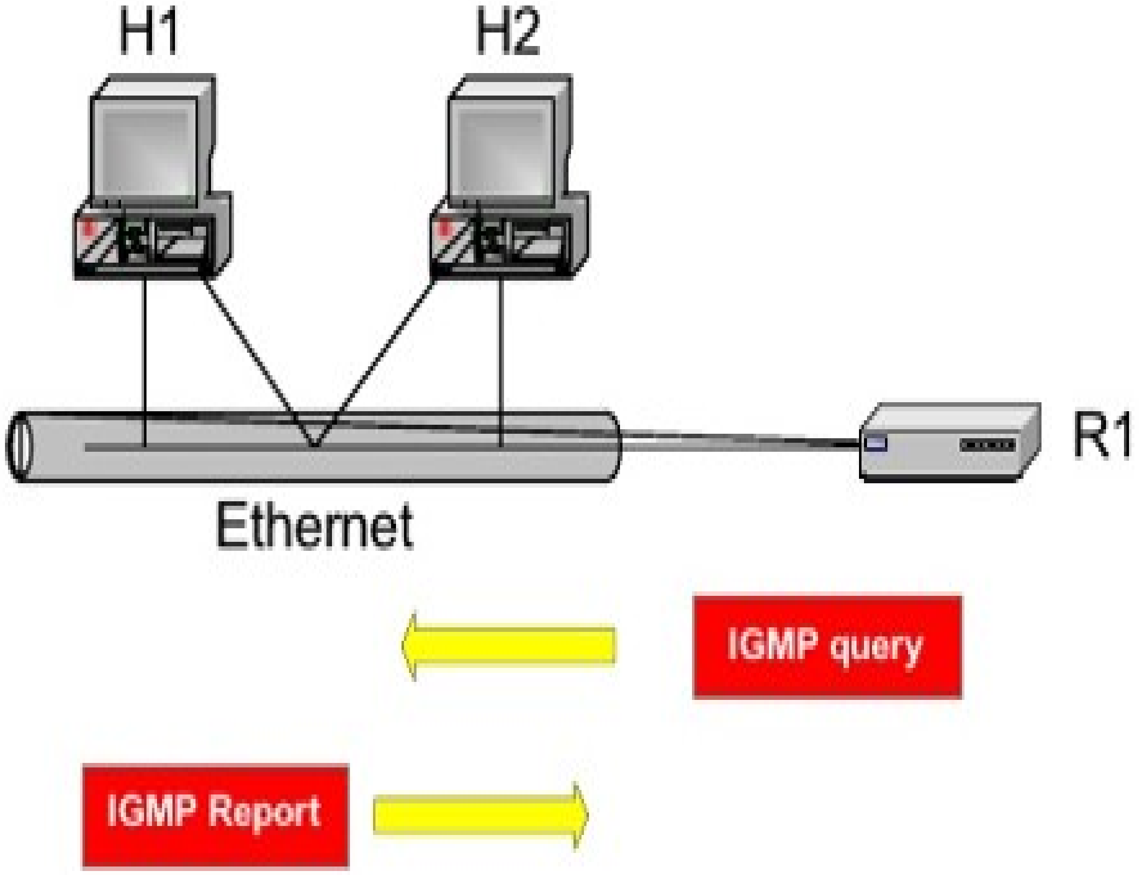
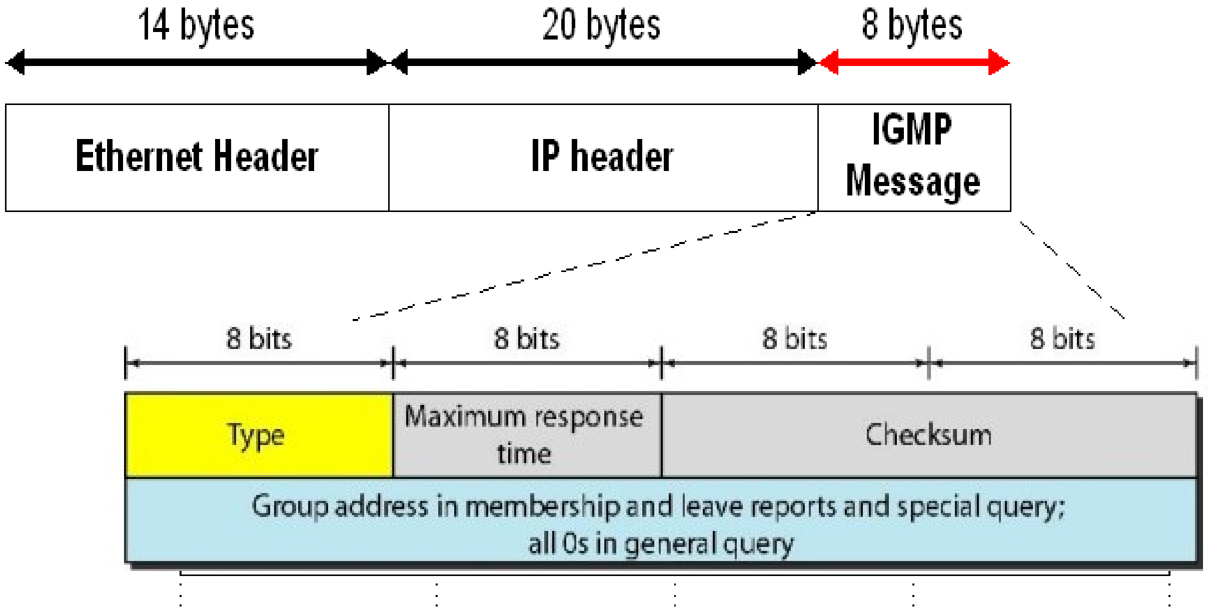


Fig: IGMP Protocol

IGMP Packet Format



•Type: 1 = sent by router, 2 = sent by host

Figure IGMP message format

- **Type:** Type of **IGMP message**.
- There are three types:
    - Membership Query,
    - Membership Report
    - Leave Group.
- **Maximum Response Time.** This 8-bit field defines the amount of time in which a query must be answered.
- **Checksum**: This is the **one's compliment** of the **one's complement sum** of the entire IGMP message, the IGMP datagram is encapsulated within the IP datagram .
- **Group Address:**
    - Behavior of this field varies by the type of message.
    - The value defines the **group id** (multicast address of the group) in the **special query**, the **membership report**, and the **leave report** messages.

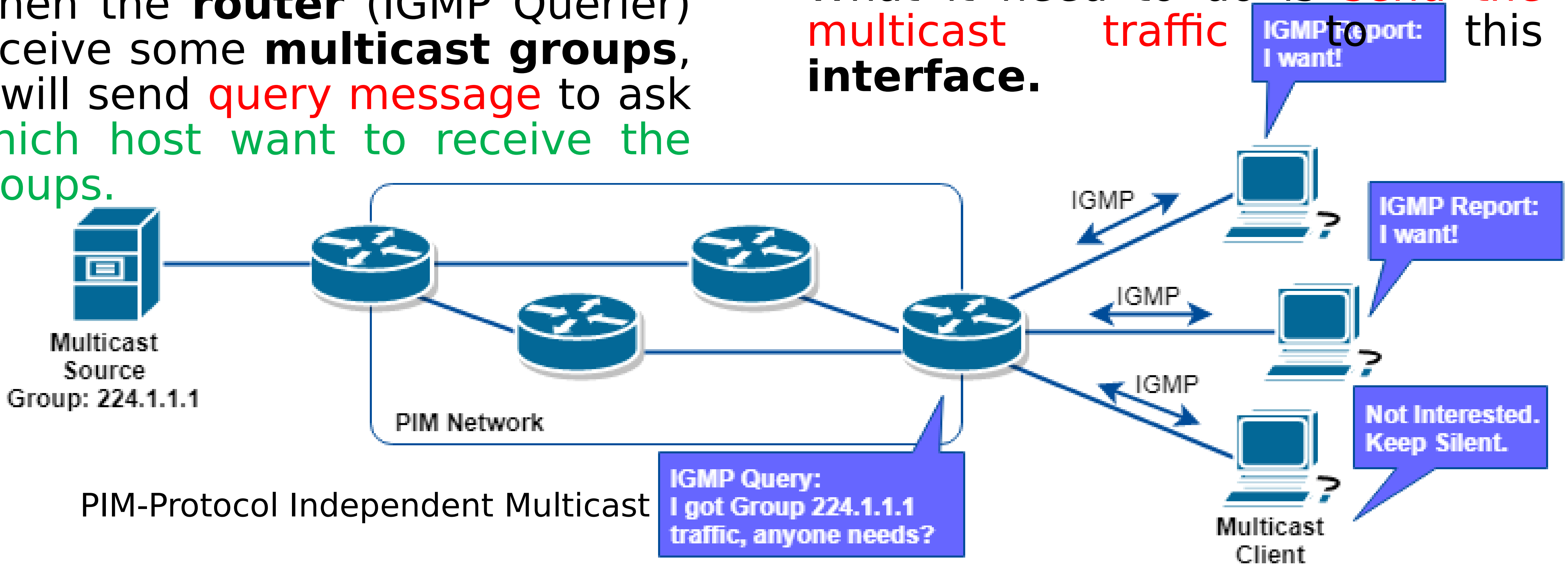
IGMP

- IGMP is the **language** between host and router to tell the router that which host is requesting which multicast group.
- **Mechanism:**

➢ When the **router** (IGMP Querier) receive some **multicast groups**, it will send query message to ask which host want to receive the groups.

➢ If a **host** (receiver) needs, it will reply an IGMP membership report message and the **router** knows **some hosts** on this interface need the traffic.

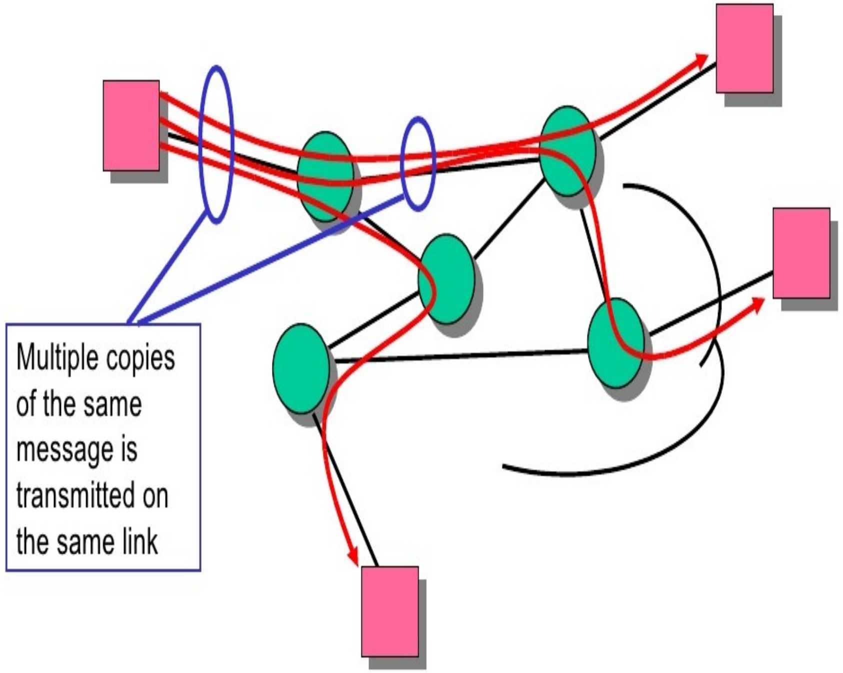
➢ What it need to do is send the multicast traffic to this **interface.**



PIM-Protocol Independent Multicast

IGMP

- If a router has no knowledge about the **membership status** of the hosts, it must **broadcast** all these packets
- This creates a lot of traffic and consumes b
- A better **solution** is
  - ➢ to keep a list of groups in the network
  - ➢ there is at least one **loyal member**.
- **IGMP** helps
  - ➢the multicast router create and updat
- IGMP operates locally.



Multiple copies of the same message is transmitted on the same link

- For each group, there is one **router** that has the duty of distributing the multicast packets destined for that group.
- In fig **Router R** distributes packets with the **multicast address** of 255.70.8.20.
- A **host** or **multicast router** can have **membership in a group**.

✓ When a **hos**
means that c
or applicati
packets from

✓ When a **router** has membership, it means that a **network** connected to one of its **other interfaces** receives this multicast packets.

- Here **R** is the distributing router.
- R1 and R2 are two multicast routers that depending on the group list maintained by router R.

uters for
n other

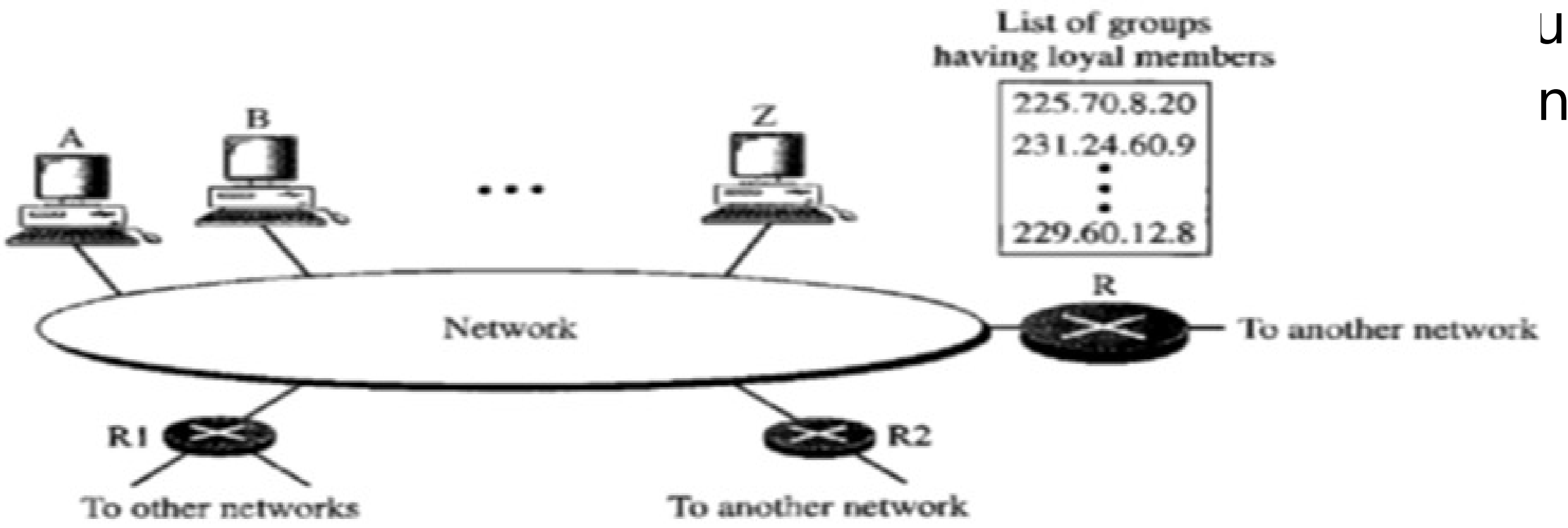


Fig. IGMP operation

Joining a Group

- A **host** or **router** can join a group
- A **host** maintains a list of processes that have membership in a group.
- When a process wants to join a new group,
  - it sends its request to the **host**.
  - The **host** adds the name of the process and the name of the requested group to its list .
- If this is **not the first entry**,
  - there is no need to send the membership report since the host is already a member of the group;
  - it already receives multicast packets for this group.

Leaving a group

- When a **host** sees that no process is interested in a specific group, it sends a leave report
- When a r**outer** sees that none of the networks connected to its interfaces in a specific group ,it sends a leave report about that group.
- When a **multicast router** receives a leave report, it cannot immediately purge that group from its list because the report comes from just one host or router; there may be other hosts or routers that are still interested in that group.
- To make sure, the **router** sends a **special query message** and inserts the groupid or multicast address, related to the group.
- The **router** allows a specified time for any host or router to respond.
- If during this time, no interest(membership report)is received ,the router assumes that there are **no loyal members** in the network and purges the group from its list.

Monitoring Membership

- A **host or router** can **join** a group by sending a membership report message.
- It can leave a group by sending a leave report message.
- However, sending these two types of reports is not enough.
- Consider the situation in which there is only one host interested in a group, but the host is shut down or removed from the system.
- The **multicast router** will never receive a leave report. How is this handled?
- The multicast router is responsible for **monitoring** all the hosts or routers in a LAN to see if they want to continue their membership in a group.
- The router periodically (by default, every 125 s)sends a **general query message**.
- In this message, the group address field is set to 0.0.0.0.
- This means the **query** for membership continuation is for all groups in which a host is involved, not just one.

Delayed Response

- The **router** expects an answer for **each group** in its group list; even new groups may respond.
- When a host or router receives the general query message, it responds with a membership report if it is interested in a group.
- However, if there is a common interest (two hosts, for example, are interested in the same group),only one response is sent for that group to prevent unnecessary traffic.
- This is called a **delayed response**.

Delayed Response

- To prevent unnecessary traffic, IGMP uses a delayed response strategy.
- When a host or router receives a query message it does not respond immediately; it delays the response.
- Each host or router uses a random number to create a timer, which expires between 1 and 10 s.
- A timer is set for each group in the list.
- For example, the timer for the **first group** may expire in 2s, but the timer for the **third group** may expire in 5 s.
- Each **host or router** waits until its timer has expired before sending a membership report message.
- During this **waiting time** ,if the timer of another host or router , for the same group ,expire earlier, that host or router sends a membership report.

Query Router

- **Query message** may create a lot of responses.
- To prevent unnecessary traffic, IGMP designates one router as **the query router** for each network.
- Only this **designed router** sends the query message and the **other routers** are passive(they receive responses and update their lists).

Internet inter-AS routing: BGP(Border Gateway Protocol)

- BGP is an inter domain routing protocol using **path vector routing**.
- It first appeared in 1989 and has gone through four versions.
- BGP provides each AS a means to:

  1. Obtain **subnet** reachability information from **neighboring ASs**.

  2. **Propagate** the reachability information to all routers internal to the AS.

  3. Determine **“good”** routes to subnets based on reachability information and policy.

- Allows a **subnet** to advertise its existence to rest of the Internet: “I am here”

BGP history

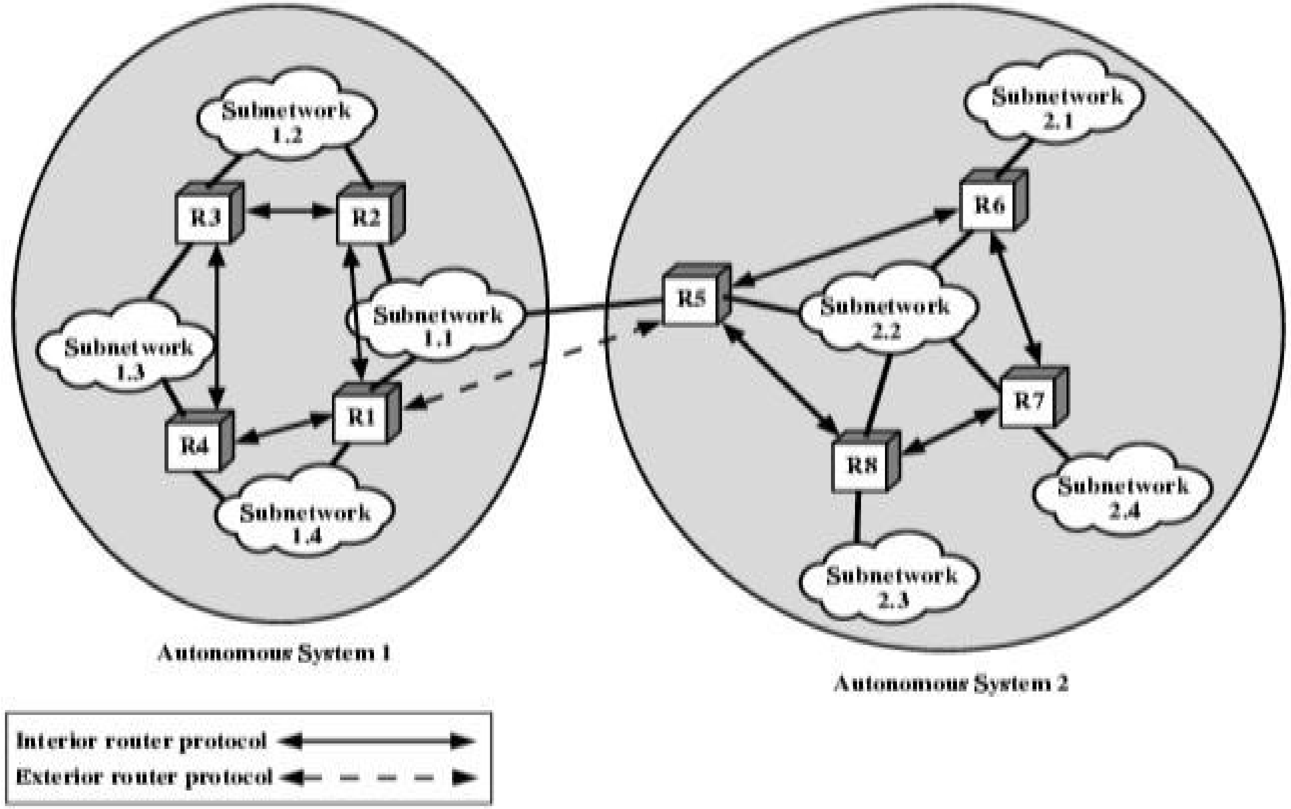
- 1989: BGP-1 [RFC 1105]
  - Replacement for EGP (1984, RFC 904)
- 1990: BGP-2 [RFC 1163]
- 1991: BGP-3 [RFC 1267]
- 1995: BGP-4 [RFC 1771]
  - Support for CIDR

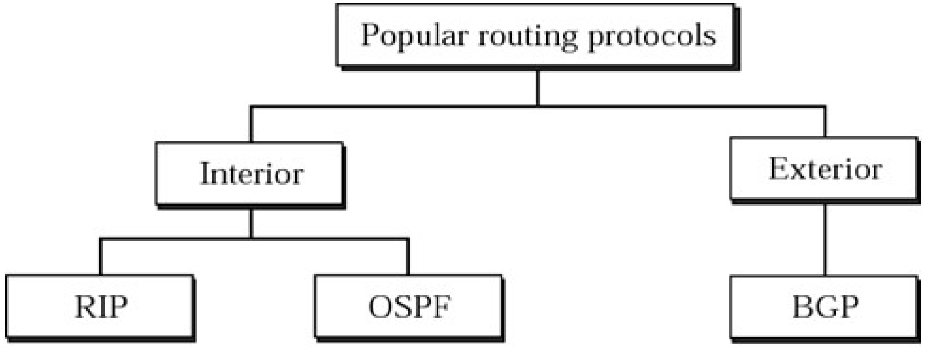
Policies with BGP

- BGP provides capabilities for enforcing various **policies.**
- Policies are not part of BGP!
- **Policies** are used to configure BGP.
- BGP enforces **policies** by choosing paths from multiple alternatives and controlling advertisements to other AS’s

Routing Protocols

- Interior Routing Protocols
- Exterior Routing Protocols





Popular routing protocols

Interior

Exterior

RIP

OSPF

BGP

Interior Routing Protocols

- Passes routing information between routers within AS.
- Does not need to be implemented outside of the AS.

Exterior Routing Protocols

- Protocol used to pass routing information between routers in different Ass.
- If a datagram is to be transferred from a host in one AS to a host in another AS.
  - ✓ **Router** in first system determines route to target AS.
  - ✓ **Routers** in target AS then co-operate to deliver datagram.

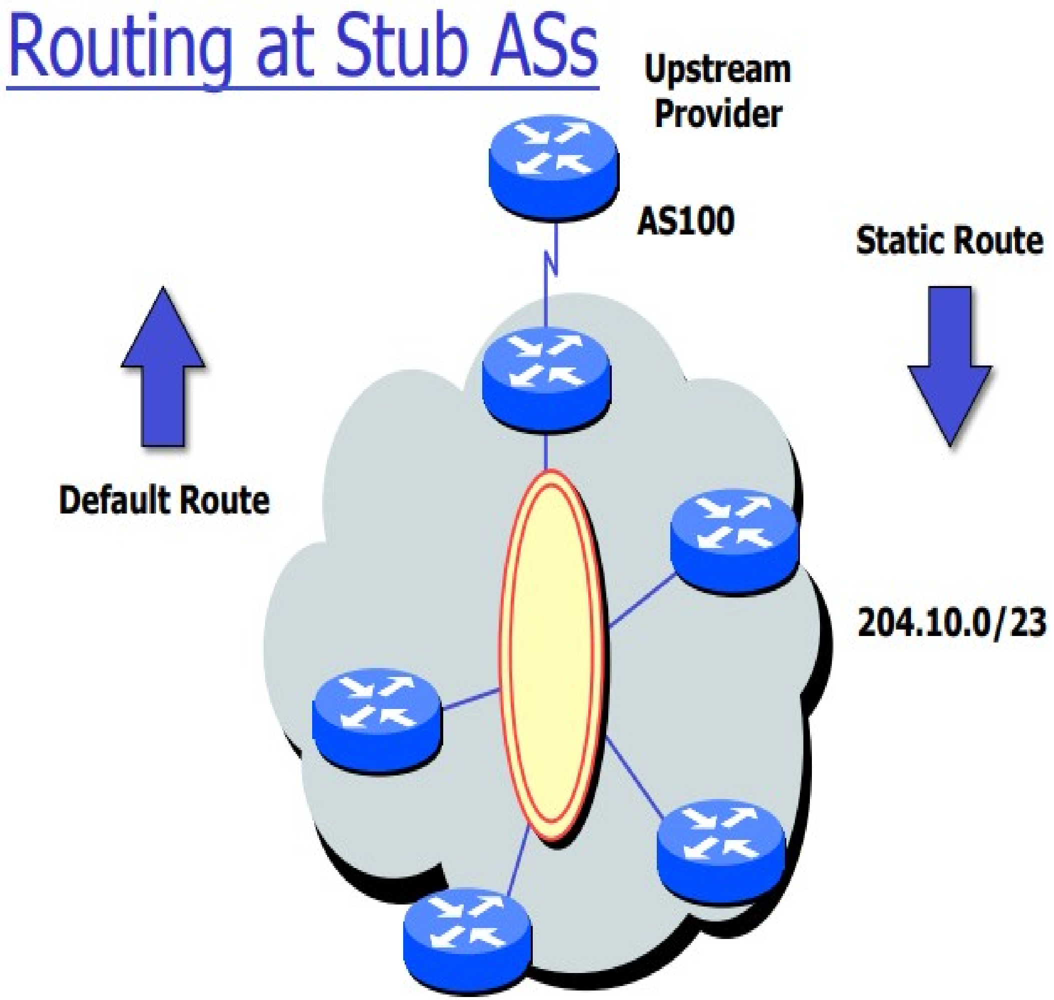
BGP – The Exterior Gateway Routing Protocol

- The **Internet** is divided into hierarchical domains called **autonomous systems**.
- For example,
  - A **large corporation** that manages its own network and has full control over it is an autonomous system.
  - A **local ISP** that provides services to local customers is an **autonomous system**.
- We can divide **autonomous systems** into three categories:
  - Stub,
  - Transit
  - Multi homed

Stub AS

- A **stub AS** has only one connection to another AS.
- The inter domain data traffic in a stub AS can be either **created** or **terminated** in the AS.
- The **hosts** in the AS can send data traffic to other ASs.
- The **hosts** in the AS can receive data coming from hosts in other ASs.

- **Data traffic**, however, cannot pass through a stub AS.
- A or a
- A AS is a sr



Transit AS

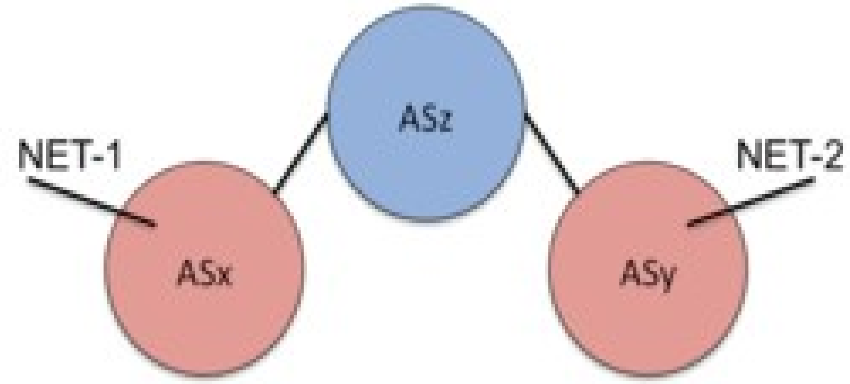


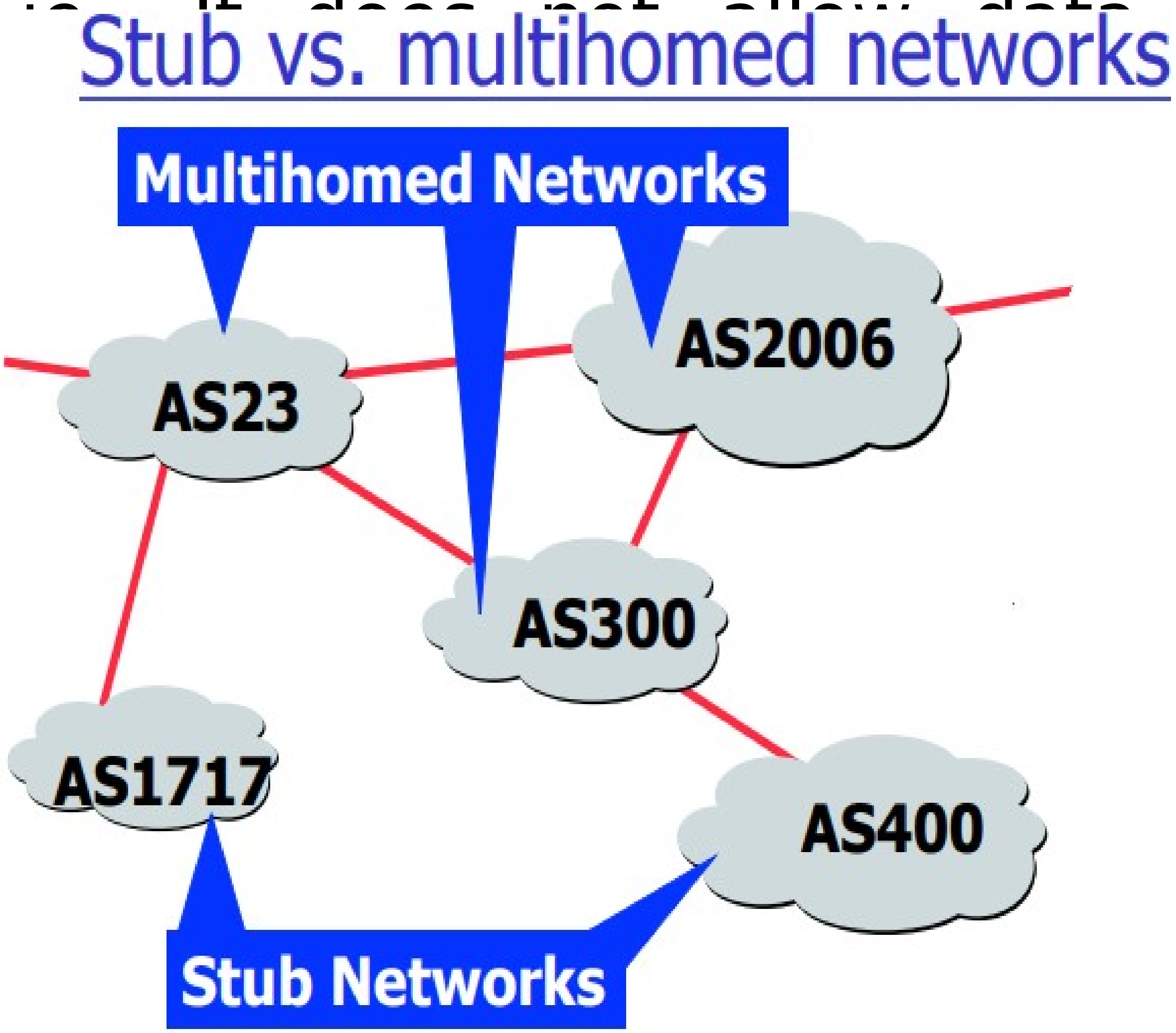
Figure Transit AS: Az handling traffic between Ax and Ay

- A **transit autonomous system** is one that offers the ability to route data from one AS to another AS.
- It allows **traffic** with neither source nor destination within AS to flow across the network.
- For example, if ASx can route date to ASy by going through ASz, ASz is a transit AS.
- A transit AS is a multihomed AS that also allows transient traffic.
- Good examples of transit ASs are **national and international ISPs** (Internet backbones).

Multi homed AS

- A multi homed AS has more than one connection to other ASs, but it is still only a source or sink for data traffic.
- It can receive data traffic from more than one AS.
- It can send data traffic to more than one AS, but there is no transient traffic.

- ie. It does not allow data c t
- A **c** t n t



BGP

- BGP is a **distance-vector protocol** used to communicate between different ASes.
- Instead of maintaining just the **cost** to each destination, each BGP router keeps track of the exact path used.
- Every **BGP router** contains a **module** that examines **routes to a given destination** and **scores them returning a number** for destination to each route.
- Functional procedures
  - Neighbor acquisition (open message, acceptance through Keepalive message)
  - Neighbor reachability (periodic Keepalive messages)
  - Network reachability (broadcast an update message)
    - Each routers maintains a database of networks that can be reached
    - preferred route to this network.

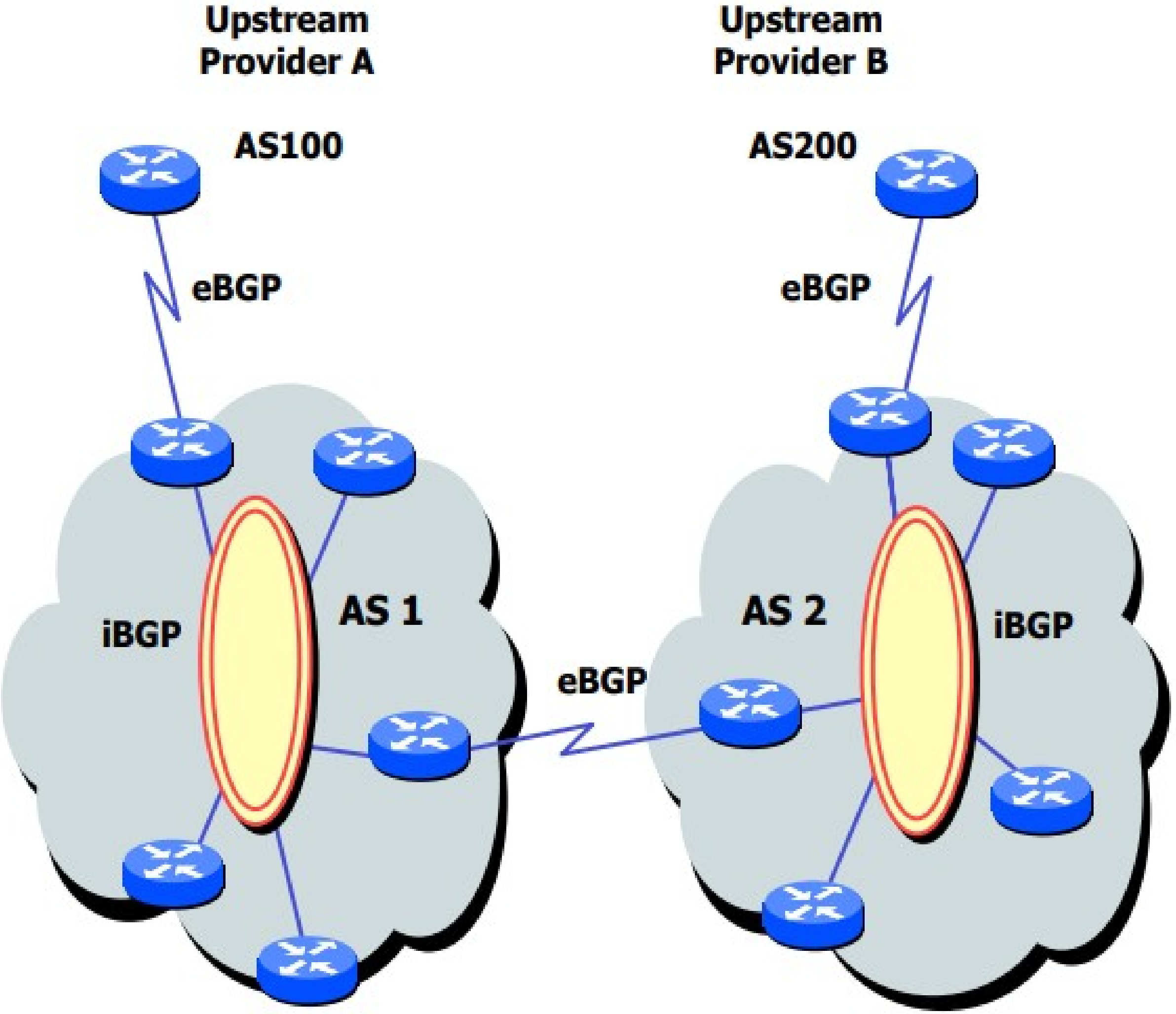
BGP Sessions:

- The exchange of **routing information** between two routers using BGP takes place in a **session**.
- A **session** is a connection that is established between two BGP routers only for the sake of exchanging routing information.
- To create a reliable environment, BGP uses the **services of TCP**.
- There is a difference between a connection in TCP made for BGP and other application programs.
- When a TCP connection is created for BGP, it can last for a long time, until something unusual happens.
- For this reason, BGP sessions are sometimes referred to as **semi-permanent connections.**

External and Internal BGP

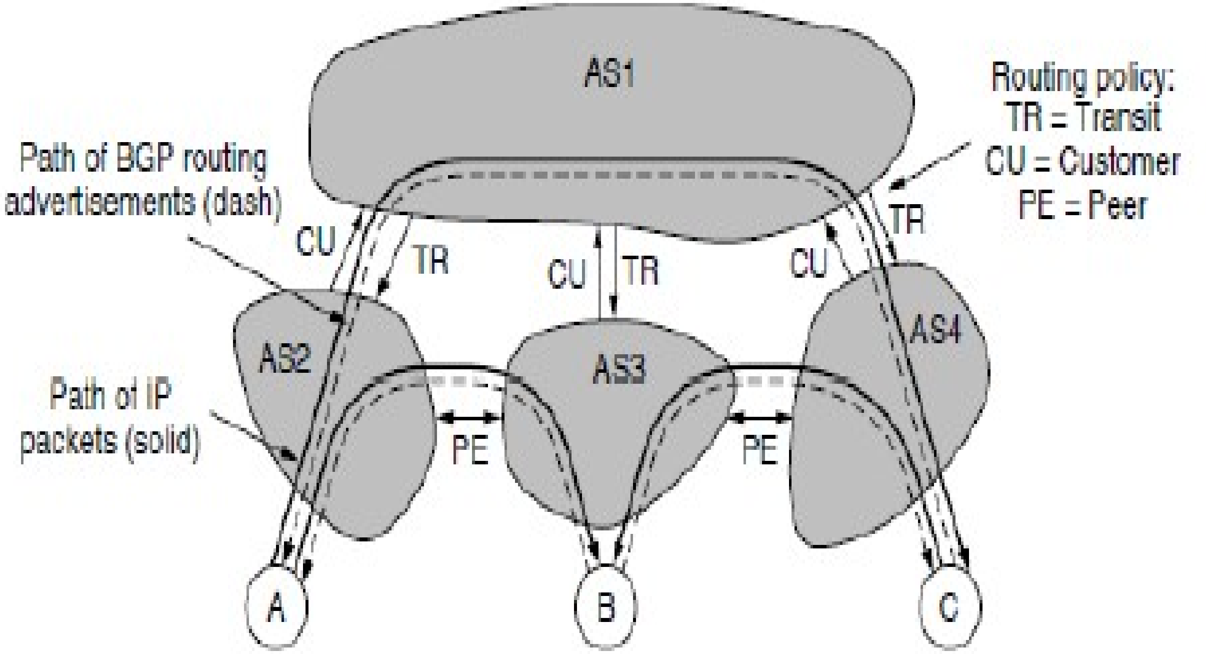
- BGP can have two types of **sessions**:
  - external BGP (E-BGP)
  - internal BGP (I-BGP) sessions

- The E-BGP session is used to
-



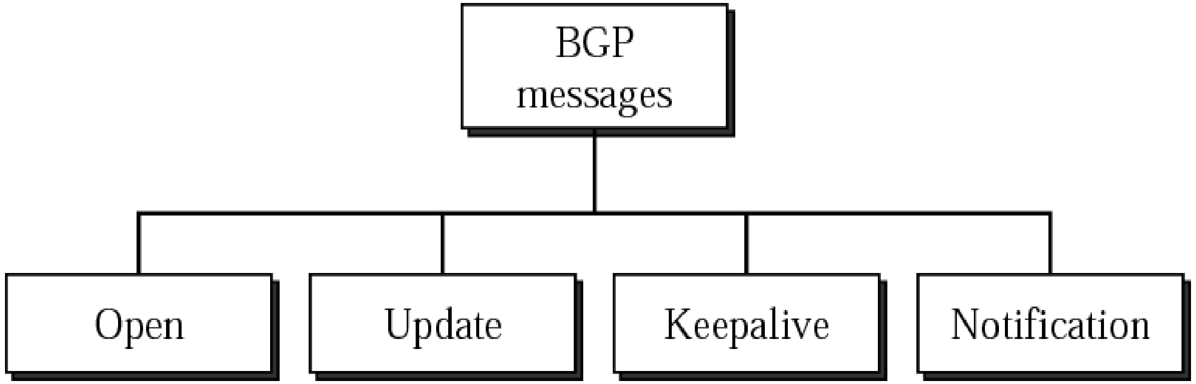
- There are four ASes that are connected.
- The connection is often made with a link at IXPs (Internet eXchange Points), connecting with other ISPs.
- **AS2, AS3, and AS4** are customers of **AS1**.
- They buy **transit service** from it.
- Thus, when **source A** sends to **destination C**, the packets travel from AS2 to AS1 and finally to AS4.
- The **routing advertisements** travel in the opposite direction to the packets.
- **AS4** advertises C as a destination to its **transit provider, AS1**, to let sources reach C via AS1.
- Later, **AS1** advertises a route to C to its other customers, including **AS2**, to let the customers know that they can send traffic

- All of the other ASes buy transit service from AS1.
  - This provides them with connectivity so they can interact with any host on the Internet.
  - Suppose that AS2 and AS3 exchange a lot of traffic.
  - Given that their networks are connected already, they can send traffic directly to



BGP messages

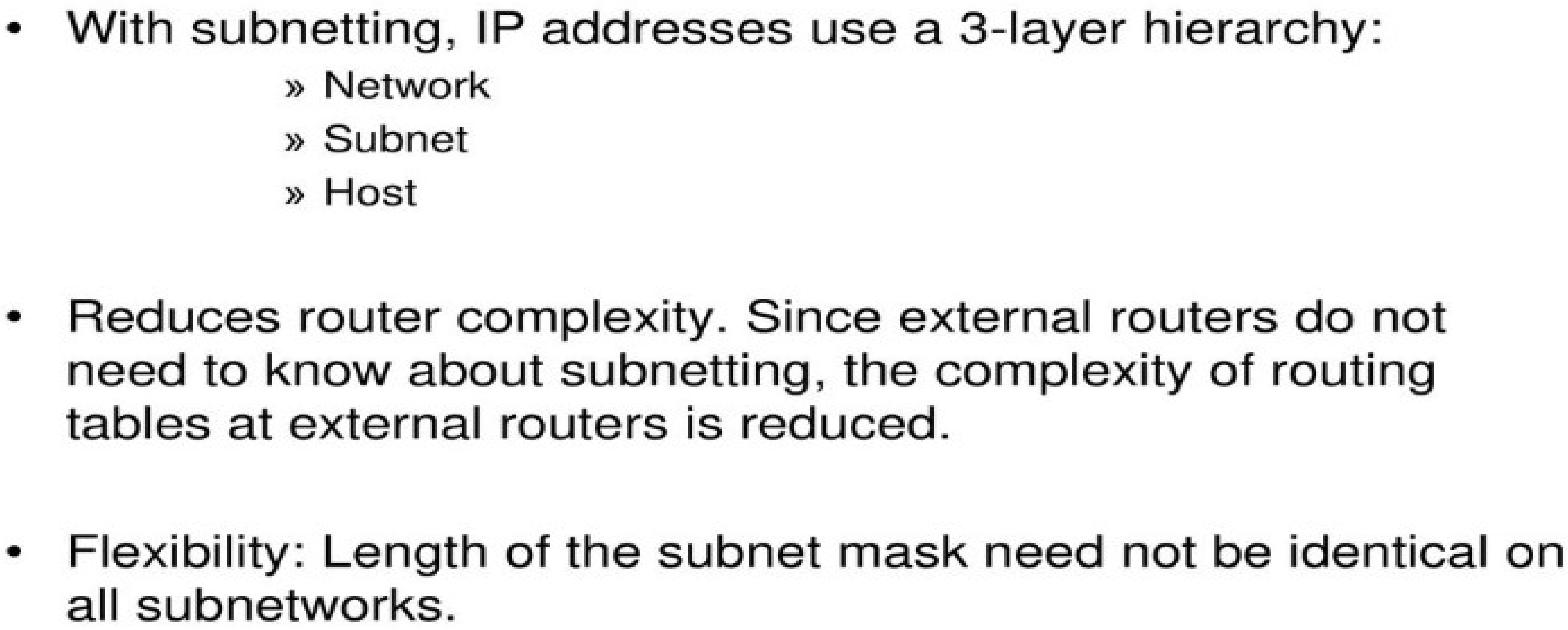
- Peers exchange BGP messages using **TCP**
- BGP messages:
  - ○ OPEN:
    - Opens TCP conn. to peer
    - Authenticates sender
  - ○ UPDATE:
    - Advertises new path (or withdraws old)
  - ○ KEEPALIVE:
    - Keeps conn alive in absence of UPDATES
    - Serves as ACK to an OPEN request
  - ○ NOTIFICATION:
    - Reports errors in previous msg;
    - Closes a connection



IPv4 Address Exhaustion

- Though the 32-bit address space of IPv4 supports about 4 billion IP devices, the **IPv4 addressing** scheme is not optimal because of recent exponential growth of the Internet.
- Current allocation trends predict exhaustion (insufficient capacity in the design of the original **Internet infrastructure**) of IPv4 space by 2008.
- **solution**
  - ➢ to switch over some scheme that could overcome this exponential growth by expanding size of IP addresses.

IPv4 - Advantages of Subnetting



- With subnetting, IP addresses use a 3-layer hierarchy:
  - » Network
  - » Subnet
  - » Host

- Reduces router complexity. Since external routers do not need to know about subnetting, the complexity of routing tables at external routers is reduced.

- Flexibility: Length of the subnet mask need not be identical on all subnetworks.

Technique to reduce address shortage in IPv4

- **Sub netting** - efficiently used IP addresses
- **CIDR** - way of designing IP addresses
- **NAT** – can allow more than 60,000 private IP addresses with a single public IP addresses.
- **Dynamic IPv4 address assignment (DHCP).**

1) A **diskless station** is just booted.

➢The situation can find its physical address by checking its **interface**, but it does not know its IP address.

2) An **organization** doesn't have enough IP address to assign to **each station**

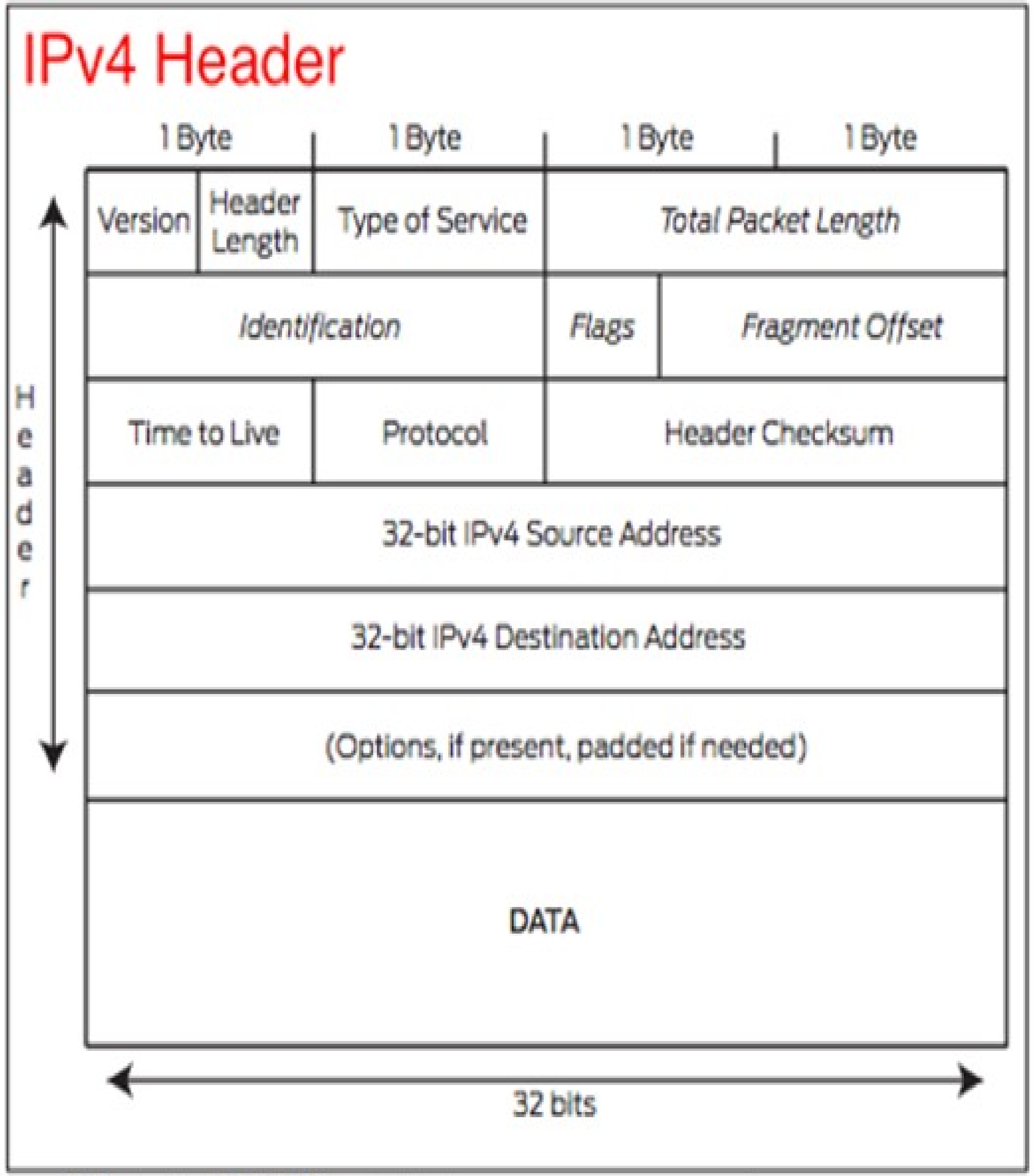
➢ It needs to assign IP address on demand.

➢The station can send its physical addresses

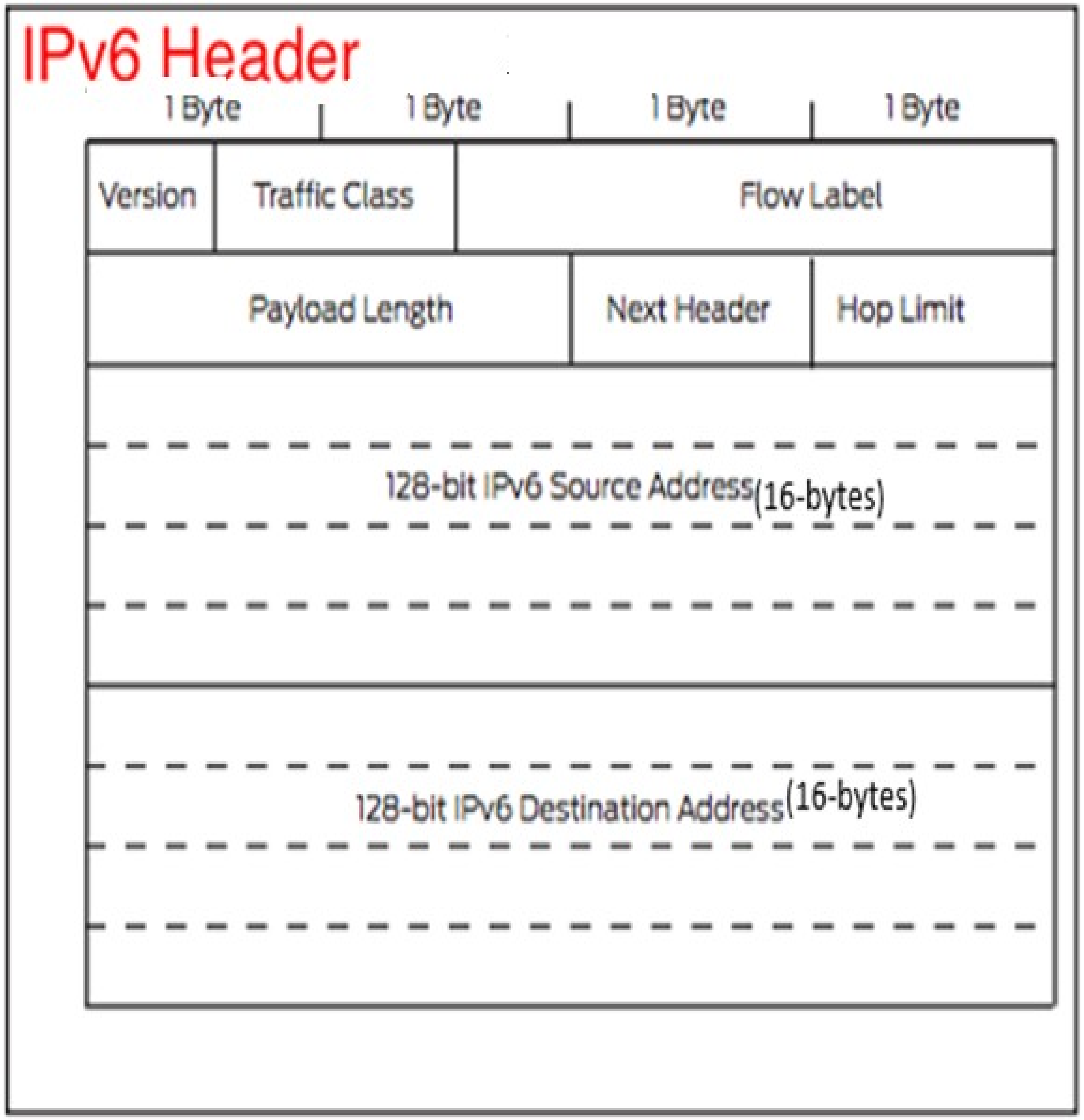
How Do IPv4 and IPv6 Work?

- The 128-bits in the **IPv6** address are **eight 16-bit hexadecimal blocks** separated by colons. For example, 2dfc:0:0:0:0217:cbff:fe8c:0.
- The 32-bits in the **IPv4** addresses are divided into “classes” with Class A networks for a few huge networks, Class C networks for thousands of small networks, and Class B networks that are in between.(**decimal format**)
- **IPv4** uses class-type address space for multicast use (224.0.0.0/4). **IPv6** uses an integrated address space for multicast, at FF00::/8.
- **IPv4** uses “broadcast” addresses that forced each device to stop and look at packets. **IPv6** uses multicast groups.
- **IPv4** uses **0.0.0.0** as an unspecified address, and **class-type address (127.0.0.1**) for loopback. **IPv6** uses **:: and ::1** as unspecified and loopback address respectively.
- **IPv4** uses globally unique public addresses for traffic and “private” addresses. **IPv6** uses globally unique unicast addresses and local addresses (FD00::/8).

IPv4 Header



IPv6 Header



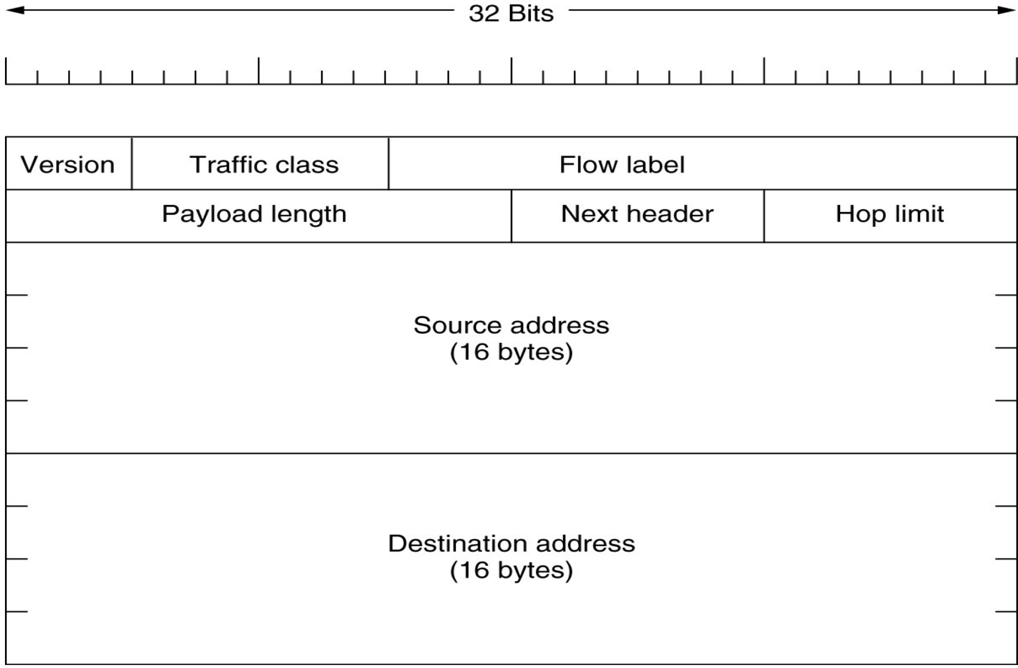
IPv6 Header

1. Version (4-bits) :

- ➢ Indicates version of Internet Protocol which contains bit sequence 0110.
- ➢ It is always 6 for IPv6

2. Traffic Class (8-bits) :

- ➢ The Traffic Class field indicates class or priority of IPv6 packet which is similar to ***Service Field*** in IPv4 packet.
- ➢ It helps **routers** to handle the traffic based on priority of the packet.
- ➢ If congestion occurs on router then packets with least priority will be discarded.



32 Bits

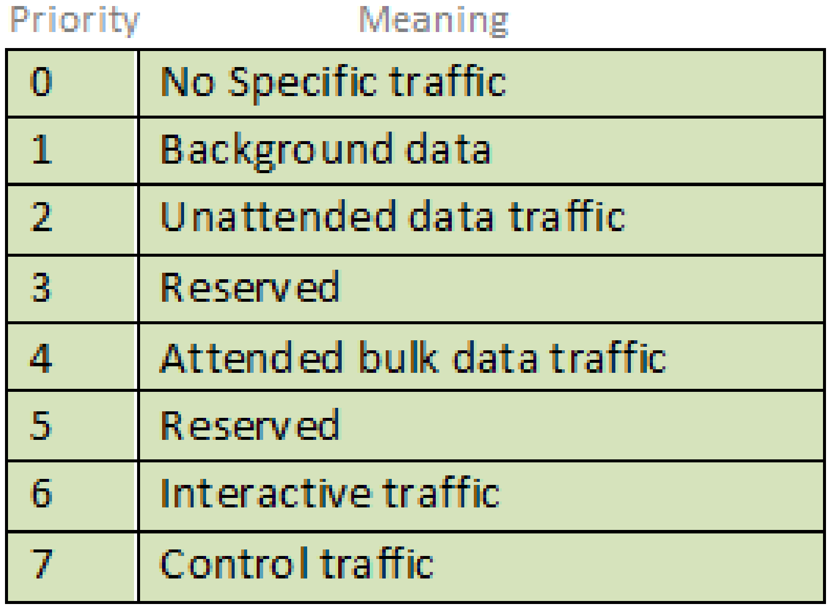
Version | Traffic class | Flow label

Payload length | Next header | Hop limit

Source address (16 bytes)

Destination address (16 bytes)

➢ **Priority assignment** of Congestion controlled traffic :



Priority	Meaning
0	No Specific traffic
1	Background data
2	Unattended data traffic
3	Reserved
4	Attended bulk data traffic
5	Reserved
6	Interactive traffic
7	Control traffic

IPv6 Header

3. Flow Label (20-bits) :

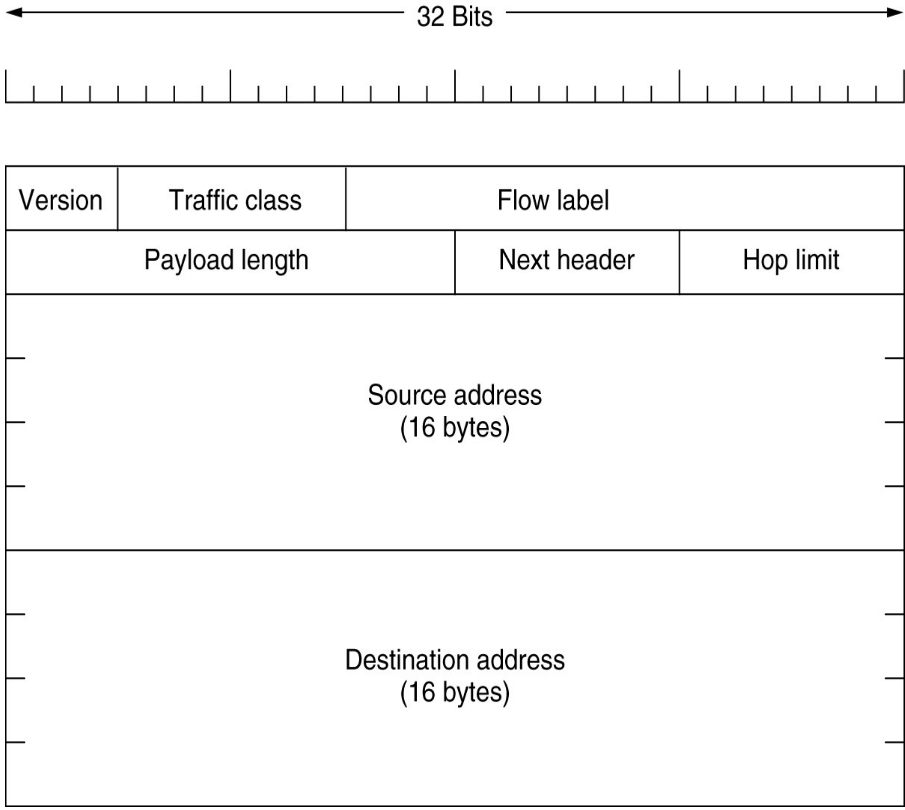
- **Routers** use the value in the flow label field to route the datagram.
- In order to distinguish the flow, intermediate **router** can use source address, destination address and flow label of the packets.

4. Payload Length (16-bits) :

- It is a 16-bit (unsigned integer) field, indicates **total size** of the payload which tells routers about amount of information a particular packet contains in its payload.
- Payload Length field includes extension headers(if any) and upper layer packet.

5. Next Header (8-bits) :

- Indicates type of extension header(if present) immediately following the IPv6 header.
- Whereas In some cases it indicates the **protocols** contained within upper-layer packet, such as TCP, UDP.



IPv6 Header

6. Hop Limit (8-bits) :

- This field is same as TTL in **IPv4** packets.
- It indicates the maximum number of intermediate nodes **IPv6 packet** is allowed to travel.
- Its value gets decremented by one, by each node that forwards the packet and packet is discarded if **value** decrements to 0.

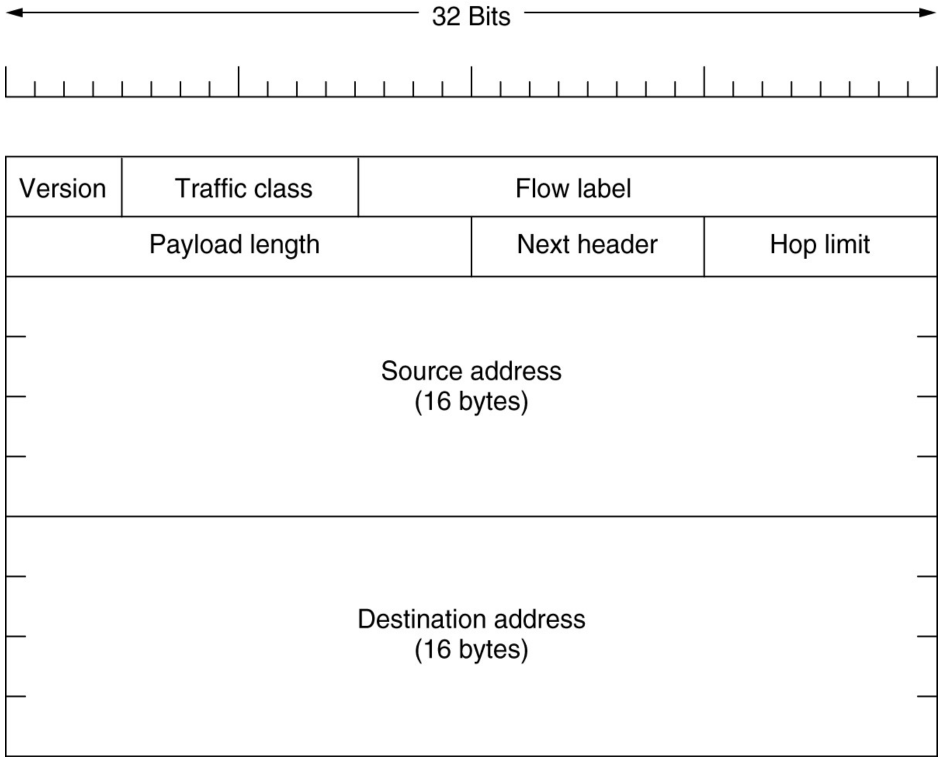
7. Source Address (128-bits) :

- 128-bit IPv6 address of the original source of the packet.

8. Destination Address (128-bits) :

- 128-bit IPv6 address of the final destination(in most cases).
- All the **intermediate nodes** can use this information in order to correctly route the packet.
- 16 bytes address are written as **eight groups** of four hexadecimal digits with colons between the groups:

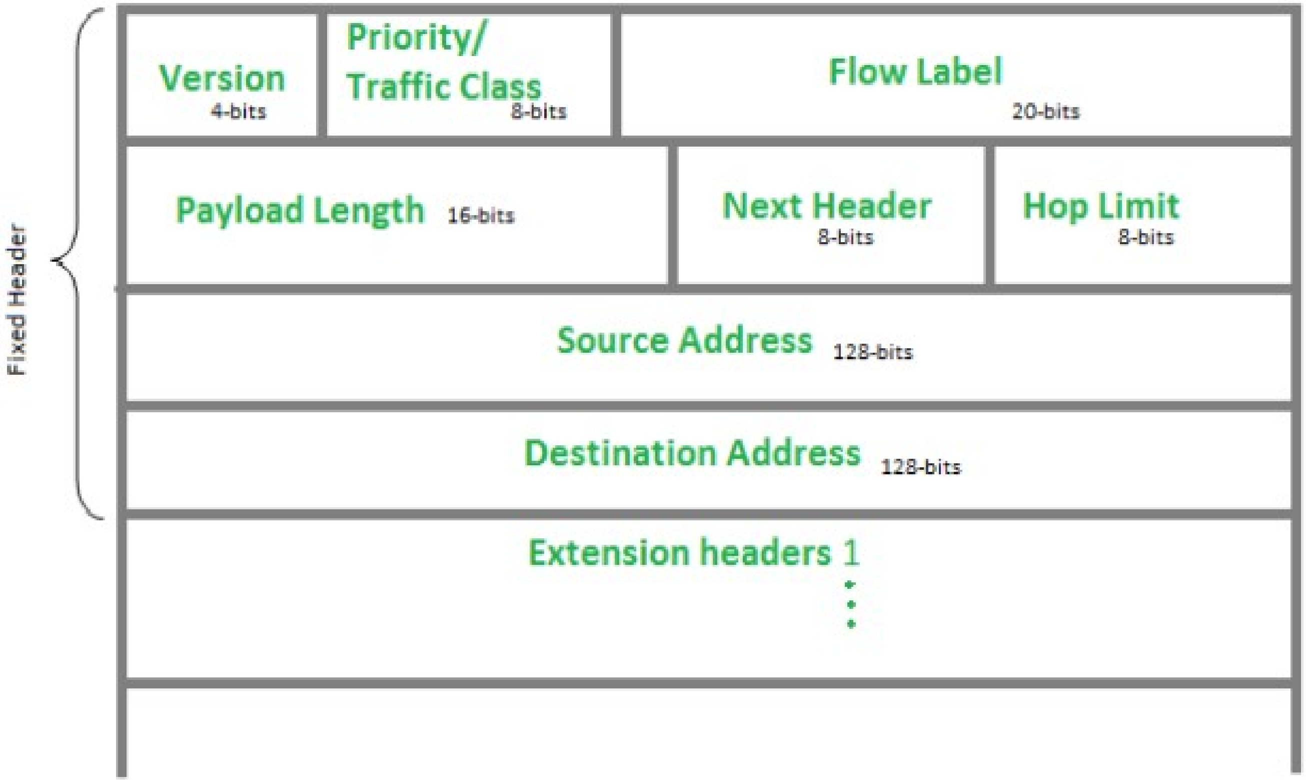
➢ 8000:0000:0000:0000:0123:4567:89AB:CDEF

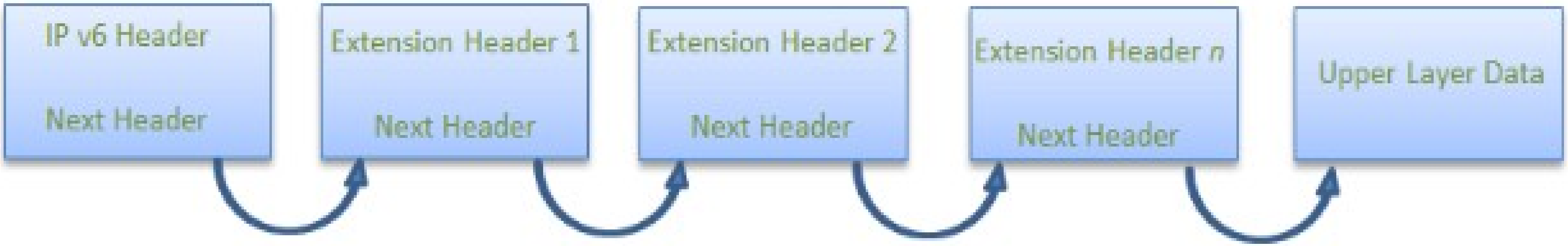


IPv6 Header

9. Extension Headers :

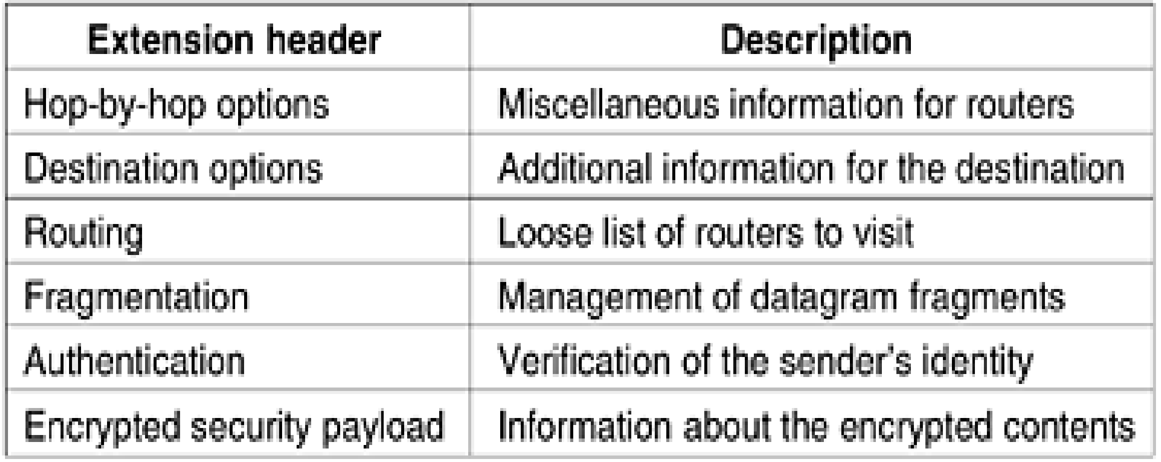
- In order to rectify the limitations of *IPv4 Option Field*, Extension Headers are introduced in IPv6.
- **Next Header** field of IPv6 fixed header points to the first Extension Header and this **first extension header** points to the second extension header and so on.



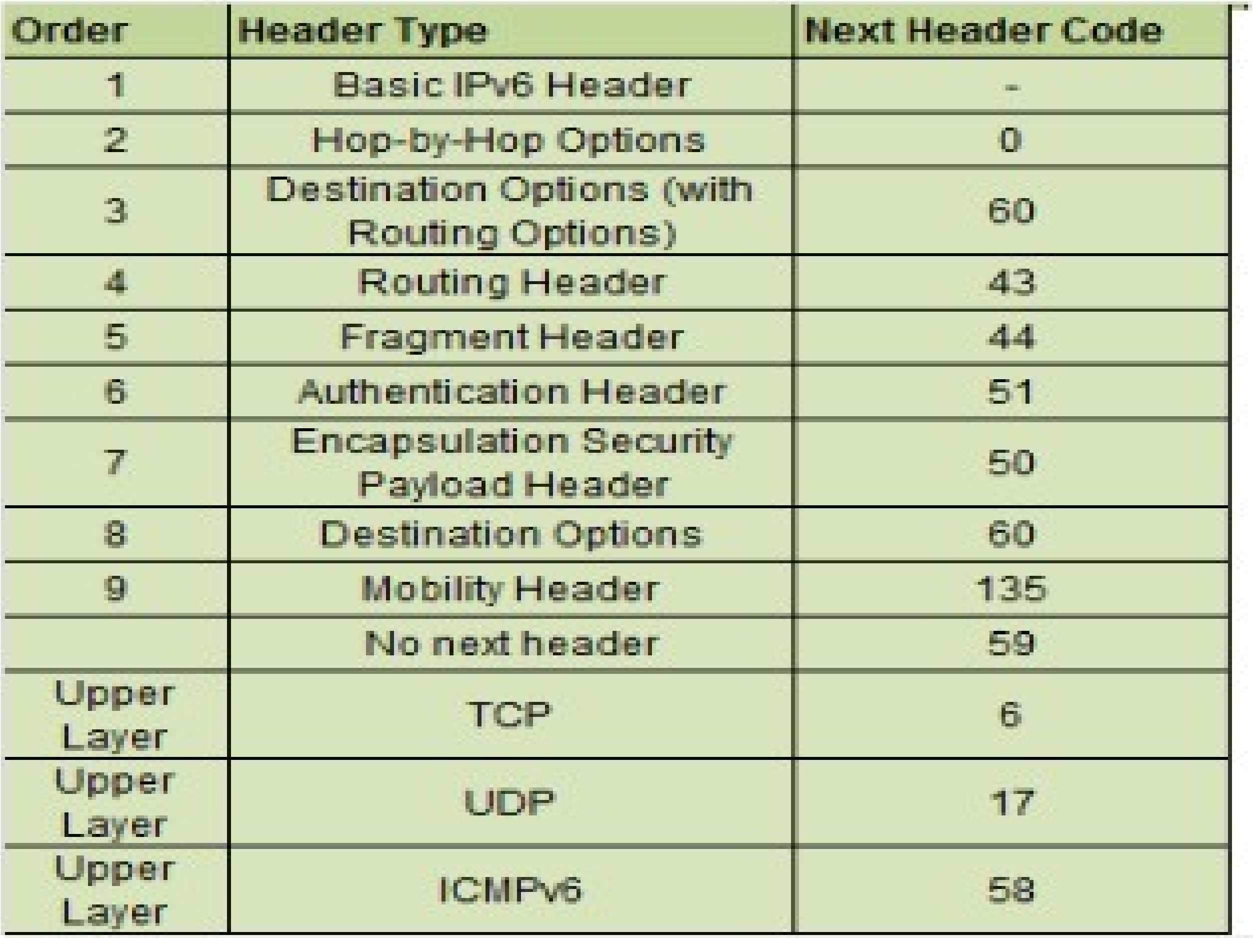


IPv6 Extension Headers

- Six kinds of **extension headers** are defined at present
- IPv6 packet may contain zero, one or more extension headers but these should be present in their recommended order:

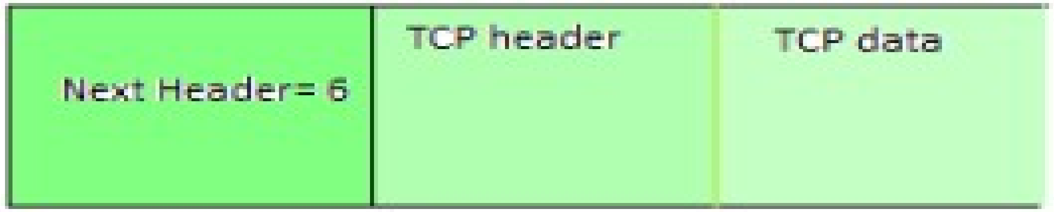


Extension header	Description
Hop-by-hop options	Miscellaneous information for routers
Destination options	Additional information for the destination
Routing	Loose list of routers to visit
Fragmentation	Management of datagram fragments
Authentication	Verification of the sender's identity
Encrypted security payload	Information about the encrypted contents

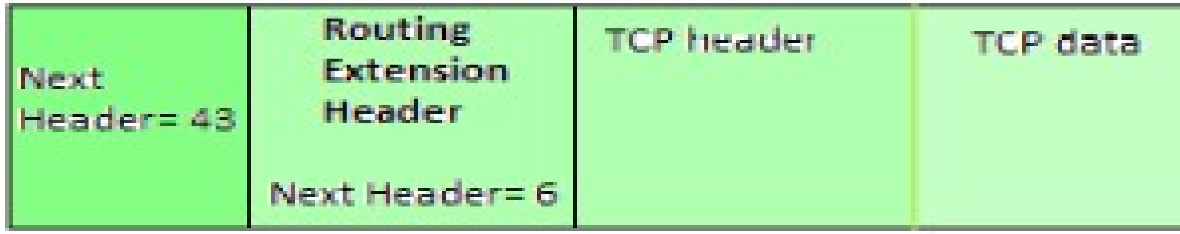


Order	Header Type	Next Header Code
1	Basic IPv6 Header	-
2	Hop-by-Hop Options	0
3	Destination Options (with Routing Options)	60
4	Routing Header	43
5	Fragment Header	44
6	Authentication Header	51
7	Encapsulation Security Payload Header	50
8	Destination Options	60
9	Mobility Header	135
	No next header	59
Upper Layer	TCP	6
Upper Layer	UDP	17
Upper Layer	ICMPv6	58

Example: *TCP is used in IPv6 packet*



Example2:



IPv6 features

1. Larger Address Space

- IPv6 has 128-bit (16-byte) source and destination IP addresses.
- In contrast to IPv4, IPv6 uses 4 times more bits to address a device on the Internet.
- According to an estimate, 1564 addresses can be allocated to every square meter of this earth.

2. Simplified Header

- IPv6's header has been simplified by moving all unnecessary information and options (which are present in IPv4 header) to the end of the IPv6 header.

3. Efficient and Hierarchical Addressing and Routing Infrastructure

- IPv6 global addresses that are used on the IPv6 portion of the Internet are designed to create an efficient, hierarchical, and summarizable routing infrastructure.

IPv6 features

4. Stateless and Stateful Address Configuration

- To simplify host configuration,
  - ➢ IPv6 supports both stateful address configuration (as in the presence of a DHCP server)
  - ➢ and stateless address configuration (as in the absence of a DHCP server).
    - ✓ With **stateless** address configuration, hosts on a link automatically configure themselves with IPv6 addresses for the link (called link-local addresses)

5. Built-in Security

- The IPv6 protocol suite support for IPSec.
- This requirement provides a standards-based solution for network security needs and promotes interoperability between different IPv6 implementations.

IPv6 features

6. Better Support for QoS

- New fields in the IPv6 header define how traffic is handled and identified.
- Because the IPv6 header identifies the traffic, **QoS** can be supported even when the packet payload is encrypted through IPSec.

7. New Protocol for Neighboring Node Interaction

- The Neighbor Discovery protocol for IPv6 is a series of **Internet Control Message Protocol for IPv6 (ICMPv6)**
- **messages** that manage the interaction of nodes on the same link (known as neighboring nodes).
- **Neighbor Discovery** replaces the broadcast-based Address Resolution Protocol (ARP), ICMPv4 Router Discovery, and ICMPv4 Redirect messages with efficient multicast and unicast Neighbor Discovery messages.

IPv6 features

8. Extensibility

- IPv6 can easily be extended by adding **extension headers** after the IPv6 header.

9. End-to-end Connectivity

- Every system now has unique IP address and can traverse through the Internet without using NAT or other translating components.

IPv6 features

10. Communication Types for IPv6

- There are three types of communication for IPv6

1. Unicast :

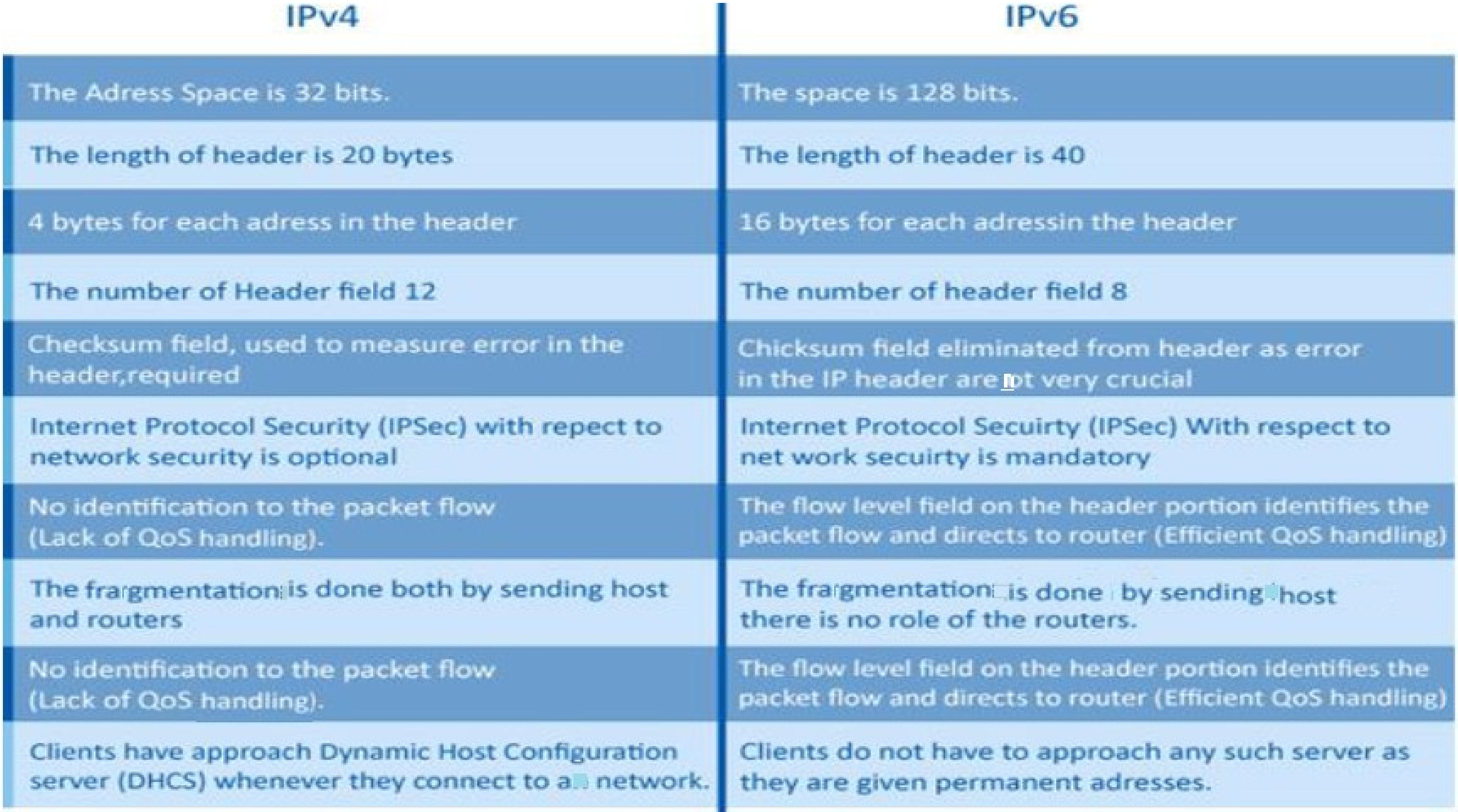
- used for one-to-one communication.
- A unicast address identifies a **single network interface.**
- The **protocol** delivers packets sent to a unicast address to that specific interface.
- There are 3 types of unicast addresses namely **global, unique-local and link-local**

2. Multicast :

- used for one-to-many communication.
- A packet that is sent to a **multicast address** is delivered to all interfaces identified by that address.
- Multicast addresses are easily identifiable because the **value** of a IPv6 multicast address begins with "FF"

3. Anycast :

- used for one-to-one-of-many communication
- An anycast address is assigned to a **group of interfaces**, usually belonging to different nodes.
- A packet sent to an **anycast address** is delivered to just one of the member interfaces, typically the **"nearest"** according to the routing protocol's choice



IPv4	IPv6
The Adress Space is 32 bits.	The space is 128 bits.
The length of header is 20 bytes	The length of header is 40
4 bytes for each adress in the header	16 bytes for each adressin the header
The number of Header field 12	The number of header field 8
Checksum field, used to measure error in the header,required	Chicksum field eliminated from header as error in the IP header are not very crucial
Internet Protocol Security (IPSec) with repect to network security is optional	Internet Protocol Secuirty (IPSec) With respect to net work secuirty is mandatory
No identification to the packet flow (Lack of QoS handling).	The flow level field on the header portion identifies the packet flow and directs to router (Efficient QoS handling)
The frargmentation is done both by sending host and routers	The frargmentation is done by sending host there is no role of the routers.
No identification to the packet flow (Lack of QoS handling).	The flow level field on the header portion identifies the packet flow and directs to router (Efficient QoS handling)
Clients have approach Dynamic Host Configuration server (DHCS) whenever they connect to a network.	Clients do not have to approach any such server as they are given permanent adresses.

Comparison Chart

BASIS OF COMPARISON	IPV4	IPV6
Address Configuration	Supports Manual and DHCP configuration.	Supports Auto-configuration and renumbering
End-to-end connection integrity	Unachievable	Achievable
Address Space	It can generate 4.29 x 10^9 addresses.	It can produce quite a large number of addresses, i.e., 3.4 x 10^38.
Security features	Security is dependent on application	IPSEC is inbuilt in the IPv6 protocol
Address length	32 bits (4 bytes)	128 bits (16 bytes)
Address Representation	In decimal	In hexadecimal
Fragmentation performed by	Sender and forwarding routers	Only by the sender
Packet flow identification	Not available	Available and uses flow label field in the header
Checksum Field	Available	Not available
Message Transmission Scheme	Broadcasting	Multicasting and Anycasting
Encryption and Authentication	Not Provided	Provided

ICMPv6

- Another protocol that has been modified in version 6 of the TCP/IP protocol suite is ICMP (ICMPv6).
- This new version follows the same strategy and purposes of version 4.
- ICMPv4 has been modified to make it more suitable for IPv6.
- In addition, some protocols that were independent in version 4 are now part of Internetworking Control Message Protocol (ICMPv6).

Internet Control Message Protocol (ICMP)

- ICMP messages are divided into two broad categories.

  1) Error-reporting message

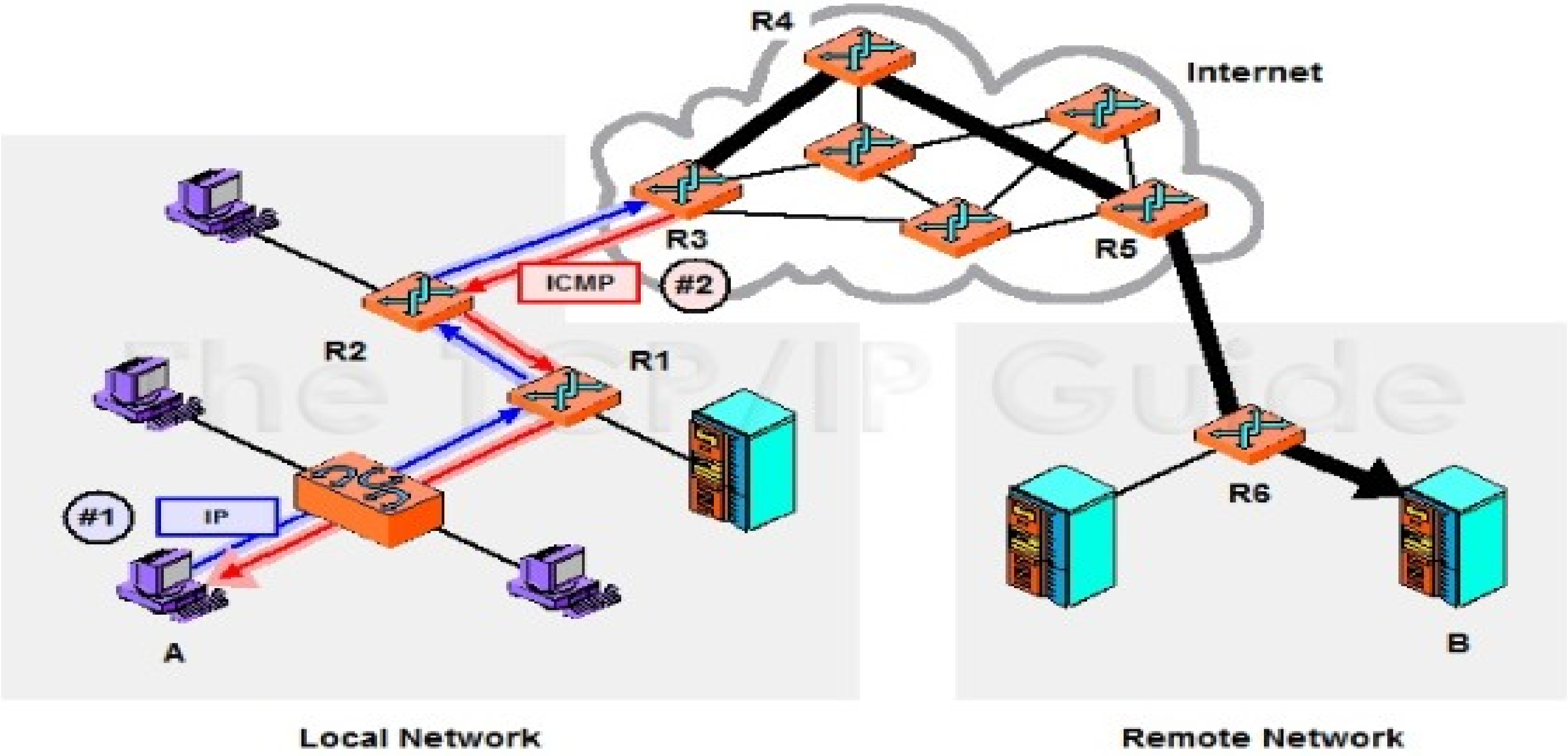
  2) Query Message

- **Error-reporting message :** report **problems** that a **router or a host(Destination)** may meet unexpected when it processes an IP packet.

- The **query messages**: help a **host or a network manager** to get specific information from a router or another host.

- E.g.: Query messages are used, if a node need **redirect** it message.

ICMP General operation



ICMPv6

- The ARP and IGMP protocols in version 4 are **combined in ICMPv6**.
- The RARP protocol is dropped from the suite because it was rarely used and BOOTP has the same functionality.
- Just as in ICMPv4, we divide the ICMP messages into two categories.
- es than

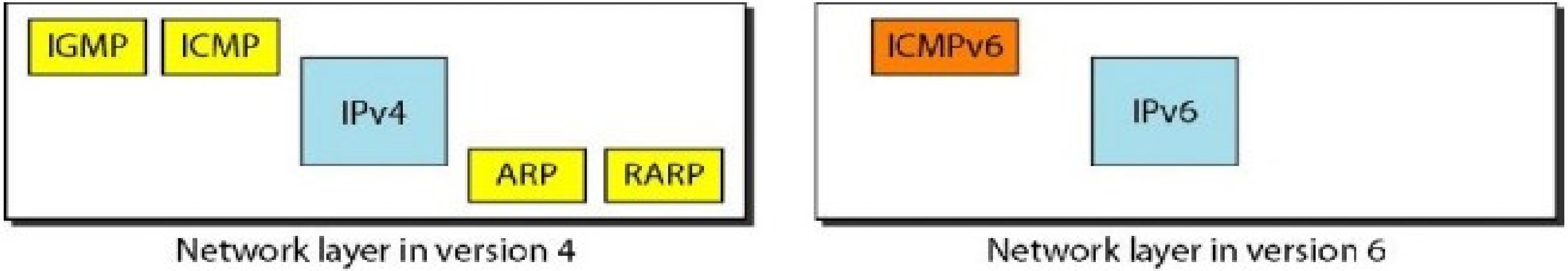
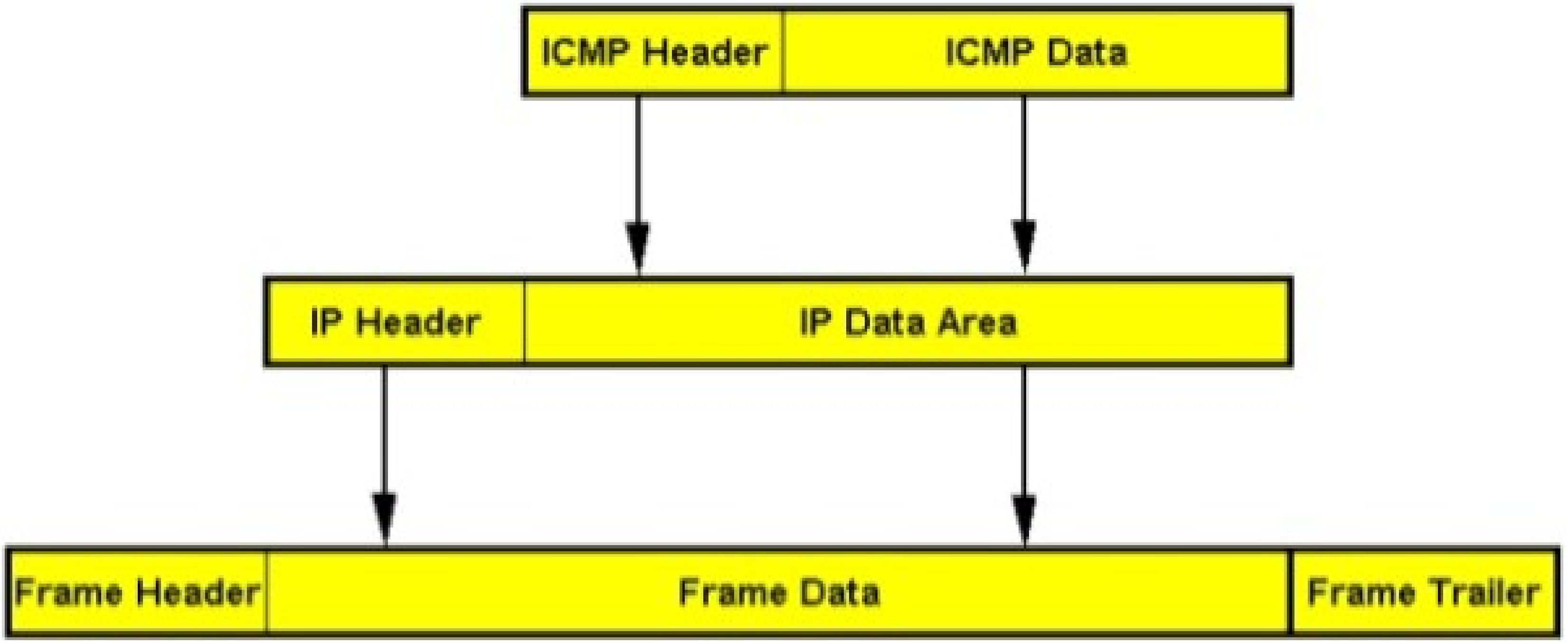


Figure 3.38 Comparison of network layers in version 4 and version 6

Internet Control Message Protocol version 6 (ICMPv6)

- ICMPv6 is defined in RFC 4443.
- ICMPv6 is an integral part of IPv6 and performs **error reporting and diagnostic functions (e.g., ping),** and has a framework for extensions to implement future changes.
- Neighbor Discovery Protocol (NDP) is a node discovery protocol in IPv6 which replaces and **enhances functions of ARP**.
- Secure Neighbor Discovery (SEND) is an extension of NDP with **extra security**.
- Multicast Listener Discovery (MLD) is used by IPv6 routers for discovering multicast listeners on a directly attached link, much **like Internet Group Management Protocol (IGMP) is used in IPv4.**
- Multicast Router Discovery (MRD) allows **discovery of multicast routers**.

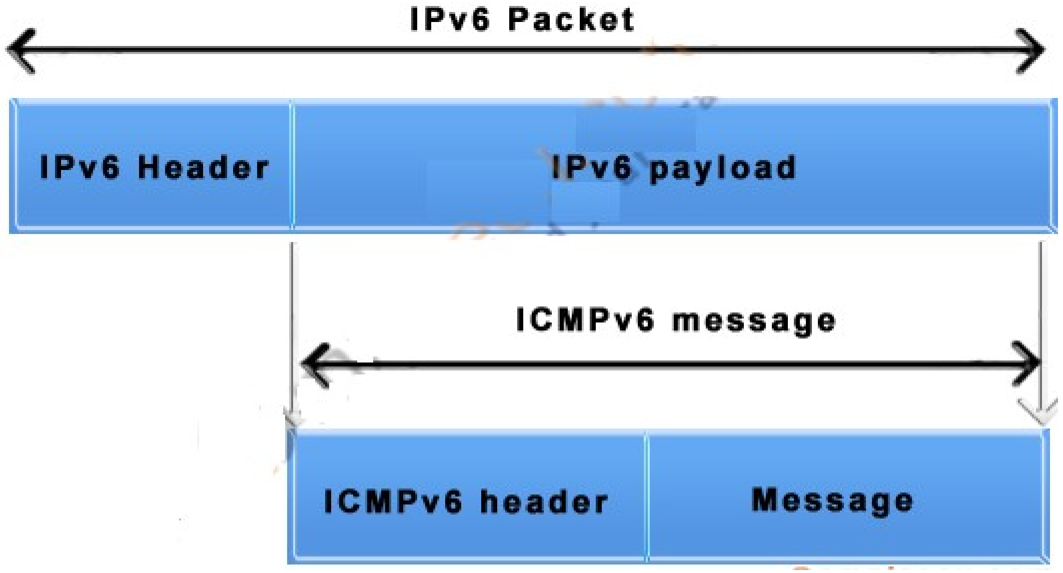


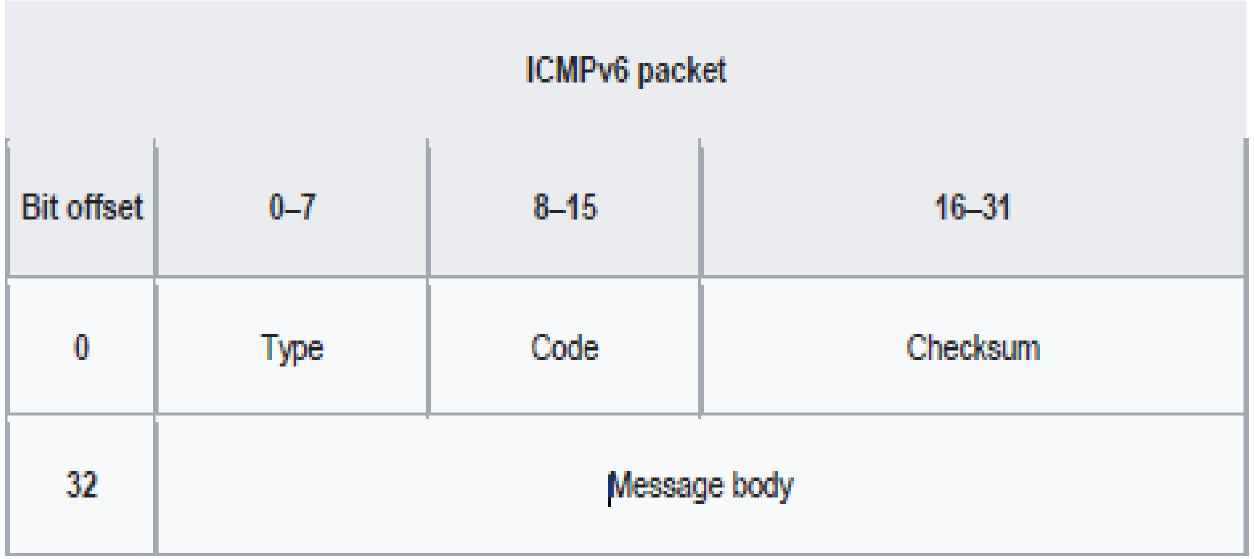
ICMP Header | ICMP Data

IP Header | IP Data Area

Frame Header | Frame Data | Frame Trailer

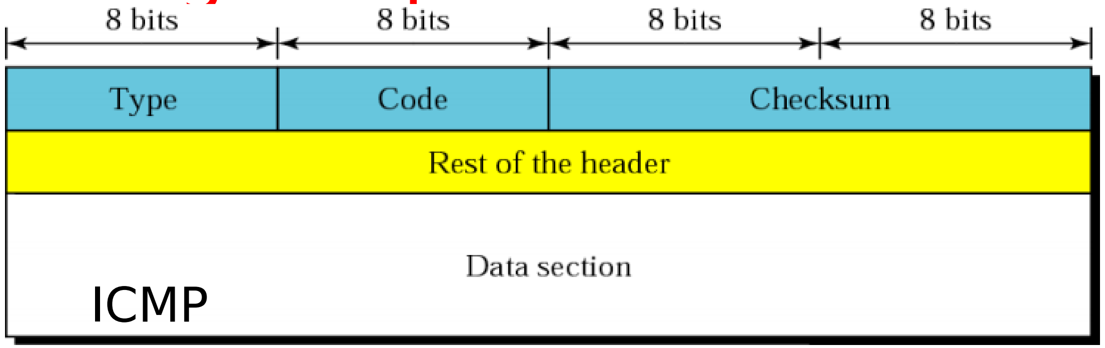
ICMPv6





ICMPv6 packet			
Bit offset	0–7	8–15	16–31
0	Type	Code	Checksum
32	Message body		

- ICMPv6 messages may be classified as **error messages** and **information messages**.
- ICMPv6 messages are transported by IPv6 packets.
- **ICMPv6 messages** are encapsulated with IPv6 datagram packets as shown



ICMP

Message types and formats

- The ICMPv6 message consists of a header and the protocol payload.
- The header contains only three fields:
  - ➢ type (8 bits),
  - ➢ code (8 bits),
  - ➢ checksum (16 bits).

1. Type

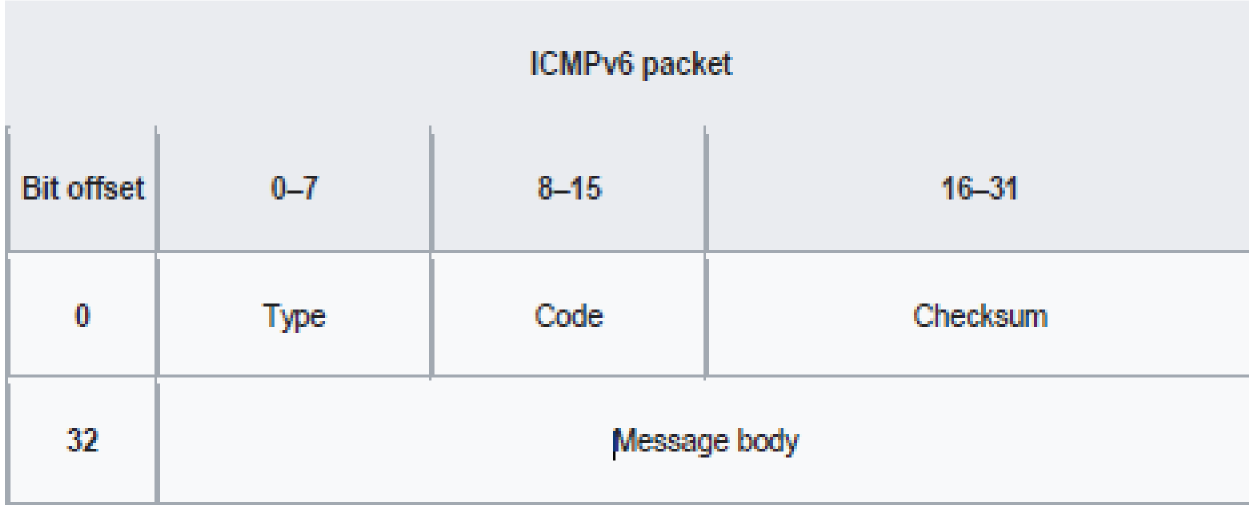
- ✓ specifies the type of the message.
- ✓ **Values** in the range from 0 to 127 (high-order bit is 0) indicate an **error message**, while values in the range from 128 to 255 (high-order bit is 1) indicate an **information message.**

2. code

- ✓ field value depends on the message type
- ✓ provides an additional level of message granularity.

3. checksum

- ✓ ICMPv6 provides a minimal level of



ICMPv6 packet

Bit offset	0–7	8–15	16–31
0	Type	Code	Checksum
32	Message body		

ICMPv6 Error Messages

- **ICMPv6 error messages** are used to report errors in the forwarding or delivery of IPv6 packets.
- The ICMPv6 **"Type field"** values for the error message are between **0 and 127**.
- ICMPv6 error messages : **Destination Unreachable, Packet Too Big, Time Exceeded, and Parameter Problem**.

1. "Destination Unreachable" :

- ➢ message is generated by the **source host** or a **router** when an IPv6 datagram packet cannot be delivered for any reason other than congestion.

2. "Packet Too Big" :

- ➢ messages are generated by the **router** when a packet cannot be forwarded to the next hop link because the size of the IPv6 datagram is larger than the MTU (Maximum Transmission Unit) of the link.
- ➢ message includes the **MTU** (Maximum Transmission Unit) of the next link also.
- ➢ MTU (Maximum Transmission Unit) is the size of the largest protocol data unit that is supported over the link.

ICMPv6 Error Messages

3. "Time Exceeded" :

- Similar to the Time-to-Live field value in IPv4 datagram header, IPv6 header includes a **Hop Limit field**.
- The **Hop Limit field** value in IPv6 header is used to prevent routing loops.
- Hop Limit field in IPv6 datagram header is decremented by each router that forwards the packet.
- When the Hop Limit field value in IPv6 header reaches zero, the router discards the IPv6 datagram packet and returns a "Time Exceeded" ICMPv6 error message to the **source host**.

4. "Parameter Problem" :

- message is typically related with the problems and mistakes related with IPv6 header itself.
- When a problem or mistake with an IPv6 header make a router
  - ✓ cannot process the packet,
  - ✓ the router stops processing the IPv6 datagram packet,
  - ✓ discards the packet and returns a "Parameter Problem" ICMPv6 error message to the source host.

ICMPv6 informational messages

- ICMPv6 informational messages are used for **network diagnostic functions and additional critical network functions** : **Neighbor Discovery, Router Solicitation & Router Advertisements, Multicast Memberships**.
- **Echo Request and Echo Reply** (used by many commands and utilities like "ping" for network diagnostics and communication trouble shooting) are also ICMPv6 informational messages.
- The ICMPv6 informational messages have values for the Type field (8 bit binary number) between 128 and 255.

1. Diagnostic Messages:

- ➢ ICMPv6 Echo request and Echo reply are the Diagnostic messages.
- ➢ Every IPv6 host must return an ICMPv6 Echo reply when it receives an ICMPv6 Echo request.
- ➢ Echo request and Echo reply messages are used by the ping command to check the **network connectivity between two IPv6 hosts**.

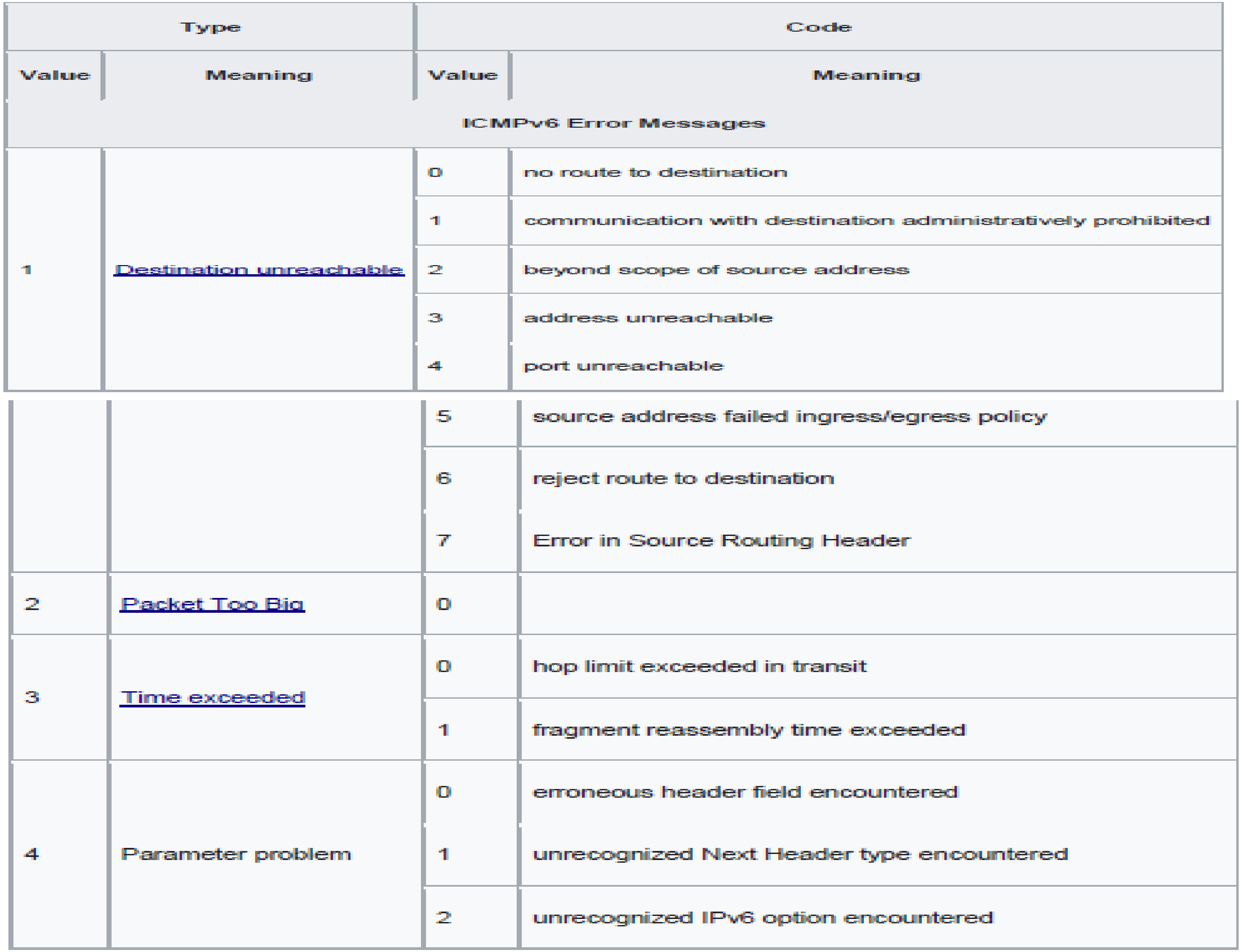
ICMPv6 informational messages

2. MLD (Multicast Listener Discovery) Messages:

- ICMPv6 MLD (Multicast Listener Discovery) Messages are used by an IPv6 enabled router to discover **hosts** who are interested in **multicast packets**, and the multicast addresses they are interested.
- MLD (Multicast Listener Discovery) messages are used by MLD (Multicast Listener Discovery) Protocol.
- MLD (Multicast Listener Discovery) Protocol is the IPv6 equivalent of IGMP (Internet Group Management) Protocol in IPv4.

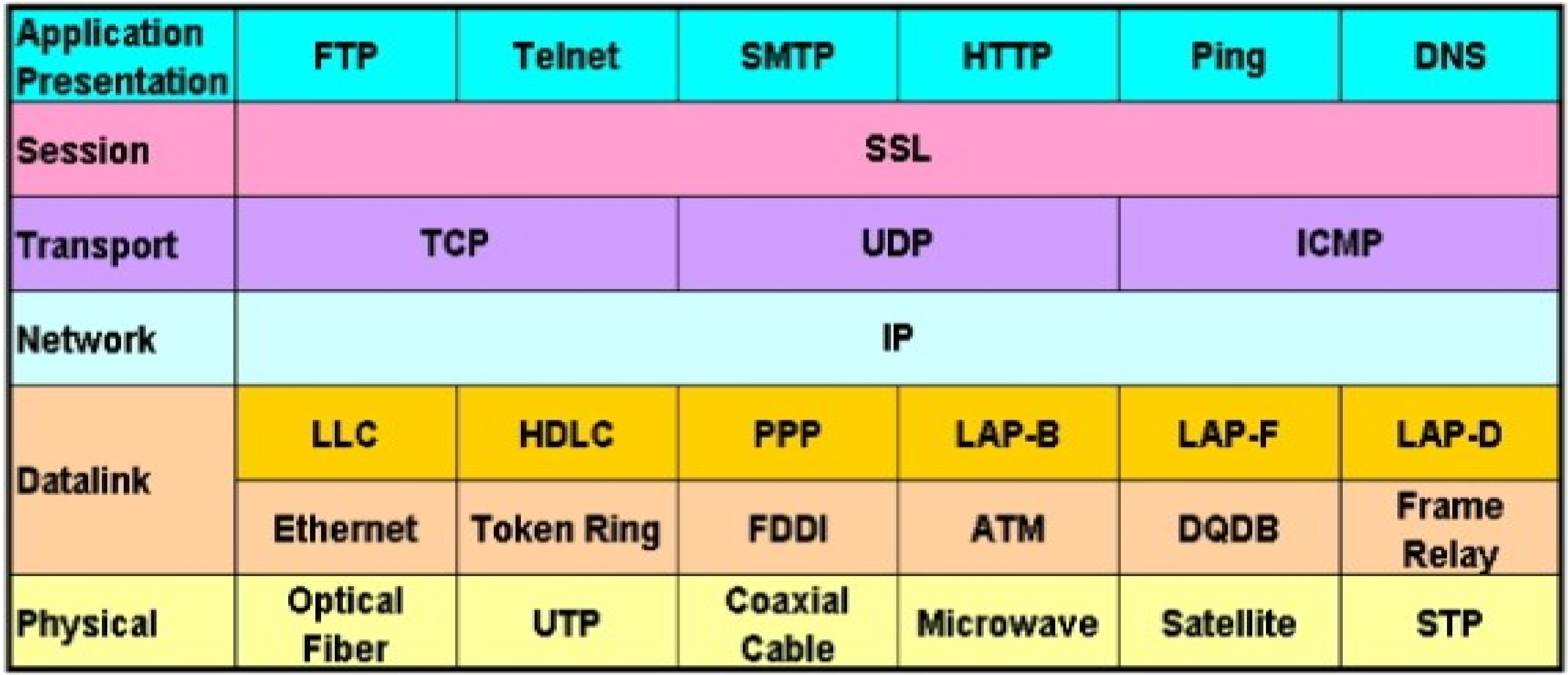
3. ND (Neighbor Discovery) Messages:

- ICMPv6 ND (Neighbor Discovery) Messages are used for the Neighbor Discovery Protocol (NDP).
- ND (Neighbor Discovery) Messages includes **Router Solicitation & Router Advertisement, Neighbor Solicitation and Neighbor Advertisement**.



Type		Code	
Value	Meaning	Value	Meaning
ICMPv6 Error Messages			
1	Destination unreachable	0	no route to destination
		1	communication with destination administratively prohibited
		2	beyond scope of source address
		3	address unreachable
		4	port unreachable
		5	source address failed ingress/egress policy
		6	reject route to destination
		7	Error in Source Routing Header
2	Packet Too Big	0	
3	Time exceeded	0	hop limit exceeded in transit
		1	fragment reassembly time exceeded
4	Parameter problem	0	erroneous header field encountered
		1	unrecognized Next Header type encountered
		2	unrecognized IPv6 option encountered

IP and ICMP



Application Presentation	FTP	Telnet	SMTP	HTTP	Ping	DNS
Session	SSL					
Transport	TCP		UDP		ICMP	
Network	IP					
Datalink	LLC	HDLC	PPP	LAP-B	LAP-F	LAP-D
	Ethernet	Token Ring	FDDI	ATM	DQDB	Frame Relay
Physical	Optical Fiber	UTP	Coaxial Cable	Microwave	Satellite	STP