MODULE V

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

The Internet has several control protocols used in the network layer, including ICMP, ARP, RARP, BOOTP, and DHCP.

ICMP - The Internet Control Message Protocol

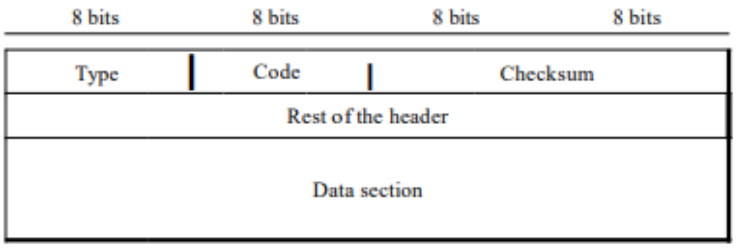
The operation of the Internet is monitored closely by the routers. When something unexpected occurs, the event is reported by the ICMP (Internet Control Message Protocol), which is also used to test the Internet. Each ICMP message type is encapsulated in an IP packet.

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 neighbors

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.

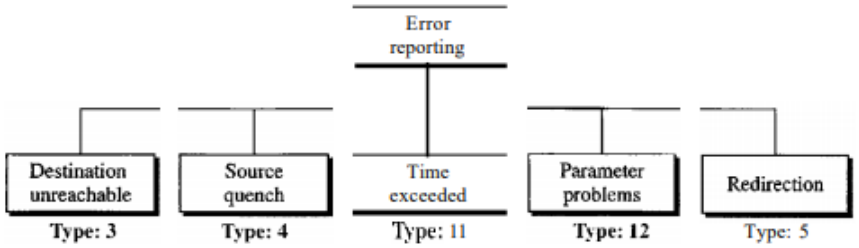
Figure 21.8 *Generalformat of[CM? messages*



ICMP type, defines the type of the message.
The **code field** specifies the reason for the particular message type.
The last common field is the **checksum field** for error handling.
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.

One of the main responsibilities of ICMP is to report errors. ICMP always reports error messages to the original source. Five types of errors are handled: destination unreachable, source quench, time exceeded, parameter problems, and redirection

Figure 21.9 *Error-reporting messages*



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 ICMP 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.

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, informa tion about the datagram itself. The 8 bytes of data are included because on UDP and TCP protocols, the first 8 bytes provide information about the port numbers (UDP and TCP) and sequence number (TCP). This information is needed so the source can inform the protocols (TCP or UDP) about the error.

The **DESTINATION UNREACHABLE** message is used

- when the subnet or a router cannot locate the destination or
- when a packet with the DF bit cannot be delivered because a "small-packet" network stands in the way.

The **TIME EXCEEDED** message is sent
The time-exceeded message is generated in two cases:

- routers use routing tables to find the next hop (next router) that must receive the packet. If there are errors in one or more routing tables, a packet can travel in a loop or a cycle, going from one router to the next or visiting a series of routers endlessly. Each datagram contains a field called time to live that controls this situation. When a datagram visits a router, the value of this field is decremented by 1. When the time-to-live value reaches 0, after decrementing, the router discards the datagram. How ever, when the datagram is discarded, a time-exceeded message must be sent by the router to the original source. Second, a time-exceeded message is also generated when not all fragments that make up a message arrive at the destination host within a certain time limit.
- when a packet is dropped because its counter has reached zero.
- This event is a symptom that packets are looping, that there is enormous congestion, or
- The timer values are being set too low.

- If there are errors in one or more routing tables, a packet can travel in a loop or a cycle, going from one router to the next or visiting a series of routers endlessly. each datagram contains a field called time to live that controls this situation. When a datagram visits a router, the value of this field is decremented by 1. When the time-to-live value reaches 0, after decrementing, the router discards the datagram. However, when the datagram is discarded, a time-exceeded message must be sent by the router to the original source. Second, a time-exceeded message is also generated when not all fragments that make up a message arrive at the destination host within a certain time limit.

The **PARAMETER PROBLEM** message indicates

- an illegal value has been detected in a header field.
- A bug in the sending host's IP software or possibly in the software of a router transited.
- If a router or the destination host discovers an ambiguous or missing value in any field of the datagram, it discards the datagram and sends a parameter-problem message back to the source.
- Any ambiguity in the header part of a datagram can Create serious problems as the data gram travels through the Internet. If a router or the destination host discovers an ambig uous or missing value in any field of the datagram, it discards the datagram and sends a parameter-problem message back to the source.

The **SOURCE QUENCH** message was used

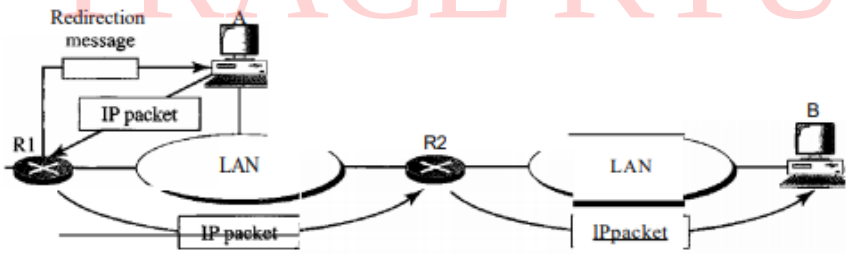
- to throttle hosts that were sending too many packets. When a host received this message, it was expected to slow down. It is rarely used any more because when congestion occurs, these packets tend to worsen it.
- The source-quench message in ICMP was designed to add a kind of flow control to the IP. When a router or host discards a datagram due to congestion, it sends a source-quench message to the sender of the datagram. This message has two purposes. First, it informs the source that the datagram has been discarded. Second, it warns the source that there is congestion somewhere in the path and that the source should slow down (quench) the sending process.
- The IP protocol is a connectionless protocol. There is no communication between the source host, which produces the datagram, the routers, which forward it, and the destination host, which processes it. One of the ramifications of this absence of communication is the lack of flow control. IP does not have a flow control mechanism embedded in the protocol. The lack of flow control can create a major problem in the operation of IP: congestion. The source host never knows if the routers or the destination host has been overwhelmed with datagrams. The source host never knows if it is producing datagrams faster than can be forwarded by routers or processed by the destination host. The lack of flow control can create congestion in routers or the destination host. A router or a host has a limited-size queue (buffer) for incoming datagrams waiting to be forwarded (in the case of a router) or to be processed (in the case of a host). If the datagrams are received much faster than they can be forwarded or processed, the queue may overflow. In this case, the router or the host has no choice but to discard some of the datagrams. The source-quench message in ICMP was designed to add a kind of flow control to the IP. When a router or host discards a datagram due to con gestion, it sends a source-quench message to the sender of the datagram. This

message has two purposes. First, it informs the source that the datagram has been discarded. Second, it warns the source that there is congestion somewhere in the path and that the source should slow down (quench) the sending process.

The **REDIRECT** message is used

- when a router notices that a packet seems to be routed wrong. It is used by the router to tell the sending host about the probable error.
- When a router needs to send a packet destined for another network, it must know the IP address of the next appropriate router. The same is true if the sender is a host. Both routers and hosts, then, must have a routing table to find the address of the router or the next router. Routing is dynamic.
- However, for efficiency, hosts do not take part in the routing update process because there are many more hosts in an internet than routerS. Updating the routing tables of hosts dynamically produces unacceptable traffic. The hosts usually use static routing. When a host comes up, its routing table has a limited number of entries. It usually knows the IP address of only one router, the default router. For this reason, the host may send a data gram, which is destined for another network, to the wrong router. In this case, the router that receives the datagram will forward the datagram to the correct router. However, to update the routing table of the host, it sends a redirection message to the host. This concept of redirection is shown in Figure 21.11. Host A wants to send a datagram to host B.

Figure **21.11** *Redirection concept*



- Router R2 is obviously the most efficient routing choice, but host A did not choose router R2. The datagram goes to R1 instead. Router R1, after consulting its table, finds that the packet should have gone to R2. It sends the packet to R2 and, at the same time, sends a redirection message to host A. Host A's routing table can now be updated.

Query

ICMP can diagnose some network problems. This is accomplished through the query messages, a group of four different pairs of messages, as shown in Figure 21.12. 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 tum is encapsulated in a data link layer frame.

Figure 21.12 *Query messages*

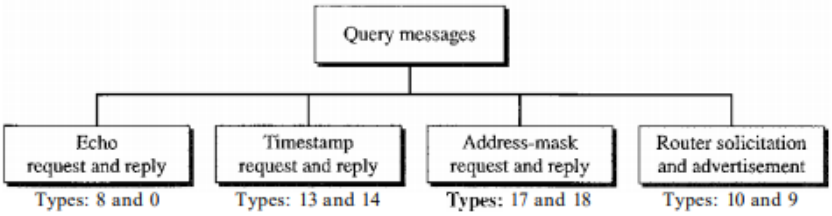
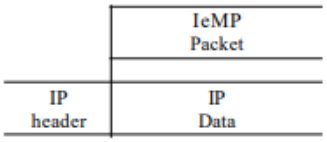


Figure 21.13 *Encapsulation ofICMP query messages*



The **ECHO and ECHO REPLY** messages are used to

- see if a given destination is reachable and alive.
- Upon receiving the ECHO message, the destination is expected to send an ECHO REPLY message back.
- 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

The **TIMESTAMP REQUEST and TIMESTAMP REPLY** messages are similar,

- except that the arrival time of the message and the departure time of the reply are recorded in the reply.
- This facility is used to measure network performance.
- 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

ROUTER SOLICITATION AND ADVERTISEMENT

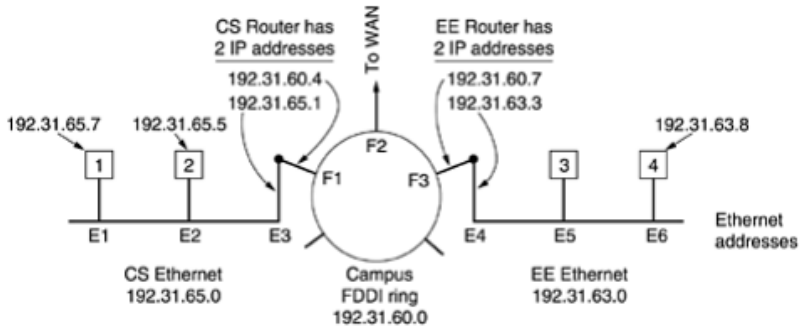
A host that wants to send data to a host on another network needs to know the address of routers connected to its own network. Also, the host must know if the routers are alive and functioning. The router-solicitation and router-advertisement messages can help in this situation. A host can broadcast (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 solic ited. 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.

ARP—The Address Resolution Protocol - ARP solves the problem of finding out which Ethernet address corresponds to a given IP address.

Every machine on the Internet has one (or more) IP addresses, these cannot actually be used for sending packets because the data link layer hardware does not understand Internet addresses. Most hosts at organizations are attached to a LAN by an interface board that only understands LAN addresses. For example, every Ethernet board ever manufactured comes equipped with a unique 48-bit Ethernet address. The boards send and receive frames based on 48-bit Ethernet addresses.

How do IP addresses get mapped onto data link layer addresses, such as Ethernet?

Figure 5-62. Three interconnected /24 networks: two Ethernets and an FDDI ring.



Small university with several class C (now called /24) networks is illustrated. We have two Ethernets, one in the Computer Science Dept., with IP address 192.31.65.0 and one in Electrical Engineering, with IP address 192.31.63.0. These are connected by a campus backbone ring (e.g., FDDI) with IP address 192.31.60.0. Each machine on an Ethernet has a unique Ethernet address, labeled E1 through E6, and each machine on the FDDI ring has an FDDI address, labeled F1 through F3.

- **a user on host 1 sends a packet to a user on host 2:**

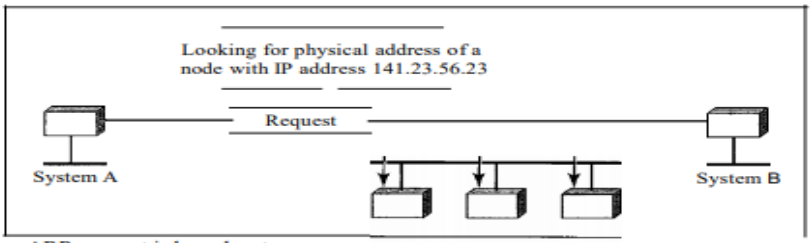
1. assume the sender knows the name of the intended receiver, eg:mary@eagle.cs.uni.edu
2. find the IP address for host 2, known as eagle.cs.uni.edu.

3. This lookup is performed by the Domain Name System, that DNS returns the IP address for host 2 (192.31.65.5).
4. The upper layer software on host 1 now builds a packet with 192.31.65.5 in the Destination address field and gives it to the IP software to transmit
5. The IP software can look at the address and see that the destination is on its own network, but it needs to find the destination's Ethernet address:

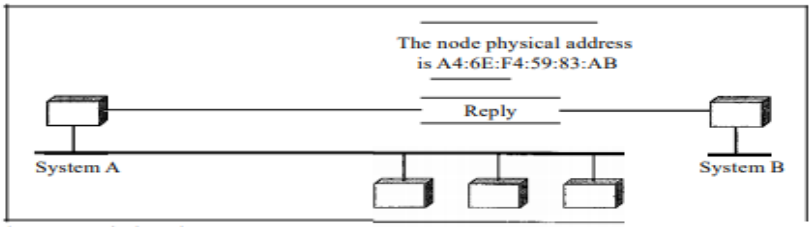
- One solution is to have a **configuration file/static mapping** somewhere in the system that maps IP addresses onto Ethernet addresses. While this solution is certainly possible, for organizations with thousands of machines, keeping all these files up to date is an error-prone, time-consuming job. Limitations of static mapping:
    - 1. A machine could change its NIC, resulting in a new physical address.
    - 2. In some LANs, such as LocalTalk, the physical address changes every time the computer is turned on.
    - 3. A mobile computer can move from one physical network to another, resulting in a change in its physical address.
    - To implement these changes, a static mapping table must be updated periodically. This overhead could affect network performance
- Host 1 to output a broadcast packet onto the Ethernet asking: Who owns IP address 192.31.65.5? The broadcast will arrive at every machine on Ethernet 192.31.65.0, and each one will check its IP address. Host 2 alone will respond with its Ethernet address (E2). The protocol used for asking this question and getting the reply is called **ARP (Address Resolution Protocol).** Almost every machine on the Internet runs it. ARP is defined in RFC 826.
- The advantage of using ARP over configuration files is the simplicity. The system manager does not have to do much except assign each machine an IP address and decide about subnet masks. ARP does the rest.

6. The IP software on host 1 builds an Ethernet frame addressed to E2, puts the IP packet (addressed to 192.31.65.5) in the payload field, and dumps it onto the Ethernet.
7. The Ethernet board of host 2 detects this frame, recognizes it as a frame for itself, scoops it up, and causes an interrupt.
8. The Ethernet driver extracts the IP packet from the payload and passes it to the IP software, which sees that it is correctly addressed and processes it.

Figure 21.1 *ARP operation*



a. ARP request is broadcast



b. ARP reply is unicast

Optimizations to make ARP work more efficiently:

- All machines on the Ethernet can enter this mapping into their ARP caches.

1. Once a machine has run ARP, it caches the result in case it needs to contact the same machine shortly. Next time it will find the mapping in its own cache, thus eliminating the need for a second broadcast.
2. Host 2 will need to send back a reply, forcing it, too, to run ARP to determine the sender's Ethernet address. This ARP broadcast can be avoided by having host 1 include its IP-to-Ethernet mapping in the ARP packet. When the ARP broadcast arrives at host 2, the pair (192.31.65.7, E1) is entered into host 2's ARP cache for future use.

- Every machine broadcast its mapping when it boots. This broadcast is generally done in the form of an ARP looking for its own IP address. There should not be a response, but a side effect of the broadcast is to make an entry in everyone's ARP cache. If a response does (unexpectedly) arrive, two machines have been assigned the same IP address. The new one should inform the system manager and not boot.
- every machine broadcast its mapping when it boots. This broadcast is generally done in the form of an ARP looking for its own IP address. There should not be a response, but a side effect of the broadcast is to make an entry in everyone's ARP cache. If a response does (unexpectedly) arrive, two machines have been assigned the same IP address. The new one should inform the system manager and not boot.

- **host 1 wants to send a packet to host 4 (192.31.63.8) / From host 1 to a distant network over a WAN :**

Using ARP will fail because host 4 will not see the broadcast (routers do not forward Ethernet-level broadcasts). There are two solutions:

1. First, the CS router could be configured to respond to ARP requests for network 192.31.63.0 (and possibly other local networks). In this case, host 1 will make an ARP cache entry of (192.31.63.8, E3) and send all traffic for host 4 to the local router. This solution is called **proxy ARP**.

2. The second solution is to have host 1 immediately see that the destination is on a remote network and just send all such traffic to a default Ethernet address that handles all remote traffic, in this case E3. This solution does not require having the CS router know which remote networks it is serving.

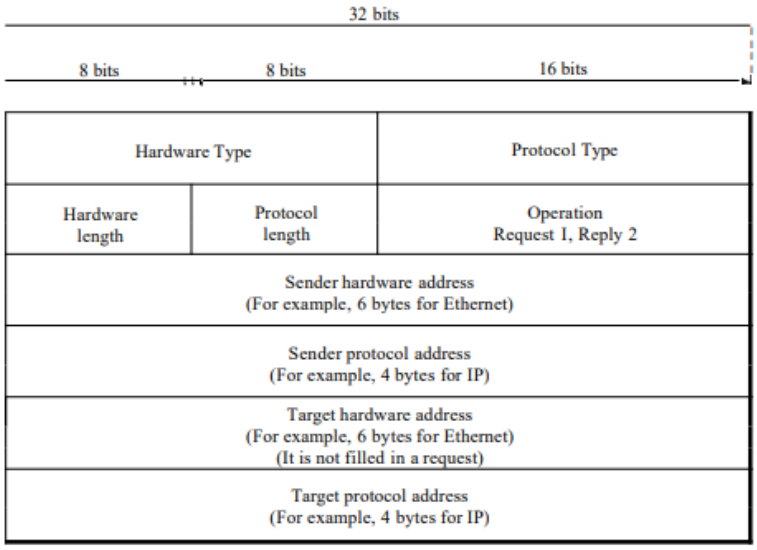
Either way, host 1 packs the IP packet into the payload field of an Ethernet frame addressed to E3. When the CS router gets the Ethernet frame, it removes the IP packet from the payload field and looks up the IP address in its routing tables. It discovers that packets for network 192.31.63.0 are supposed to go to router 192.31.60.7. If it does not already know the FDDI address of 192.31.60.7, it broadcasts an ARP packet onto the ring and learns that its ring address is F3. It then inserts the packet into the payload field of an FDDI frame addressed to F3 and puts it on the ring.

At the EE router, the FDDI driver removes the packet from the payload field and gives it to the IP software, which sees that it needs to send the packet to 192.31.63.8. If this IP address is not in its ARP cache, it broadcasts an ARP request on the EE Ethernet and learns that the destination address is E6,soit builds an Ethernet frame addressed to E6, puts the packet in the payload field, and sends it over the Ethernet. When the Ethernet frame arrives at host 4, the packet is extracted from the frame and passed to the IP software for processing.

From host 1 to a distant network over a WAN works essentially the same way, except that this time the CS router's tables tell it to use the WAN router whose FDDI address is F2.

ARP PACKET

Figure 21.2 *ARP packet*



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.
Protocol type. This is a 16-bit field defining the protocol. For example, the value of this field for the IPv4 protocol is 080016.

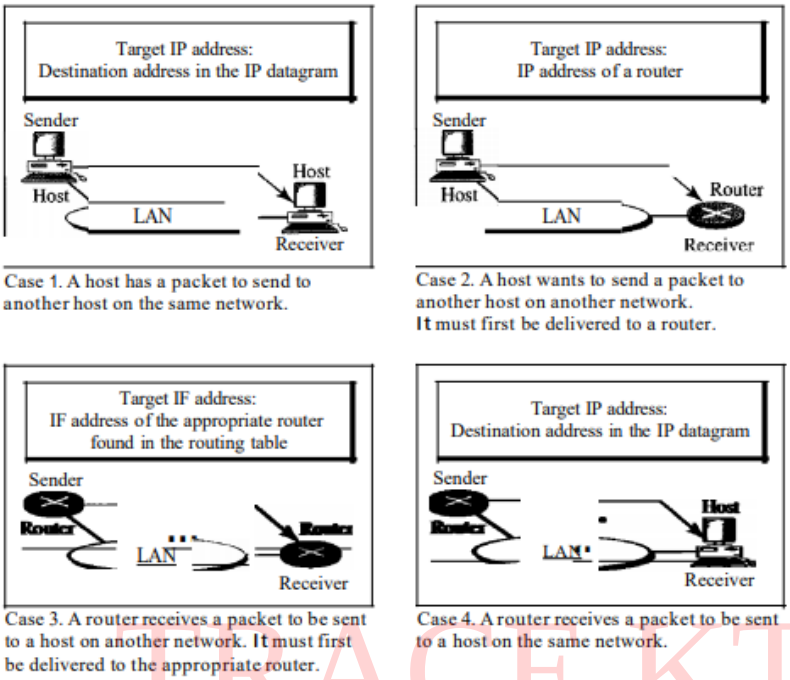
Hardware length. This is an 8-bit field defining the length of the physical address in bytes. For example, for Ethernet the value is 6.
Protocol length. This is an 8-bit field defining the length of the logical address in bytes. For example, for the IPv4 protocol the value is 4.
Operation. This is a 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 alI 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.

Operation
These are the steps involved in an ARP process:
1. The sender knows the IP address of the target.
2. IP asks ARP to create an ARP request message, filling in the sender physical address, the sender IP address, and the target IP address. The target physical address field is filled with Os.
3. The message is passed to the data link layer where it is encapsulated in a frame by using the physical address of the sender as the source address and the physical broadcast address as the destination address.
4. Every host or router receives the frame. Because the frame contains a broadcast destination address, all stations remove the message and pass it to ARP. All machines except the one targeted drop the packet. The target machine recognizes its IP address.
5. The target machine replies with an ARP reply message that contains its physical address. The message is unicast.
6. The sender receives the reply message. It now knows the physical address of the target machine.
7. The IP datagram, which carries data for the target machine, is now encapsulated in a frame and is unicast to the destination.

Four Different Cases

The following are four different cases in which the services of ARP can be used (see Figure 21.4).

Figure 21.4 *Four cases using ARP*



Case 1. A host has a packet to send to another host on the same network.

Case 2. A host wants to send a packet to another host on another network. It must first be delivered to a router.

Case 3. A router receives a packet to be sent to a host on another network. It must first be delivered to the appropriate router.

Case 4. A router receives a packet to be sent to a host on the same network.

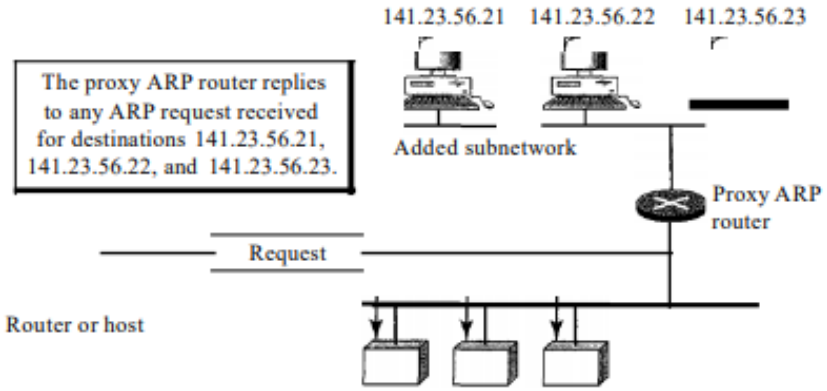
1. The sender is a host and wants to send a packet to another host on the same network. In this case, the logical address that must be mapped to a physical address is the destination IP address in the datagram header
2. The sender is a host and wants to send a packet to another host on another network. In this case, the host looks at its routing table and finds the IP address of the next hop (router) for this destination. Ifit does not have a routing table, it looks for the IP address of the default router. The IP address of the router becomes the logical address that must be mapped to a physical address.
3. The sender is a router that has received a datagram destined for a host on another network. It checks its routing table and finds the IP address of the next router. The IP address of the next router becomes the logical address that must be mapped to a physical address.
4. The sender is a router that has received a datagram destined for a host on the same network. The destination IP address of the datagram becomes the logical address that must be mapped to a physical address.

An ARP request is broadcast; an ARP reply is unicast.

ProxyARP

A technique called proxy ARP is used to create a subnetting effect. A proxy ARP is an ARP that acts on behalf of a set of hosts. Whenever a router running a proxy ARP receives an ARP request looking for the IP address of one of these hosts, the router sends an ARP reply announcing its own hardware (physical) address. After the router receives the actual IP packet, it sends the packet to the appropriate host or router.

Figure 21.6 *Proxy ARP*



In Figure 21.6 the ARP installed on the right-hand host will answer only to anARP request with a target IP address of 141.23.56.23.

The administrator may need to create a subnet without changing the whole system to recognize subnetted addresses. One solution is to add a router running a proxy ARP. In this case, the router acts on behalf of all the hosts installed on the subnet. When it receives an ARP request with a target IP address that matches the address of one of its proteges (141.23.56.21, 141.23.56.22, or 141.23.56.23), it sends an ARP reply and announces its hardware address as the target hardware address. When the router receives the IP packet, it sends the packet to the appropriate host.

Mapping Physical to Logical Address:

RARP, BOOTP, and DHCP There are occasions in which a host knows its physical address, but needs to know its logical address. This may happen in two cases:

1. A diskless station is just booted. The station can find its physical address by checking its interface, but it does not know its IP address.
2. An organization does not have enough IP addresses to assign to each station; it needs to assign IP addresses on demand. The station can send its physical address and ask for a short time lease

RARP - Reverse Address Resolution Protocol - (defined in RFC 903)

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

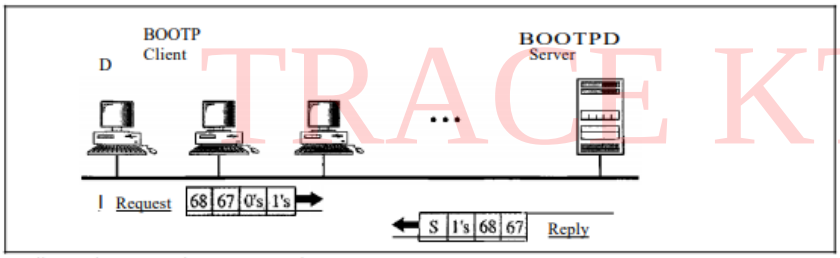
The machine can get its physical address (by reading its NIC, for example), which is unique locally. It can then use the physical address to get the logical address by using the RARP protocol. A RARP request is created and broadcast on the local network. Another machine on the local network that knows 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. Two protocols, BOOTP and DHCp, are replacing RARP.

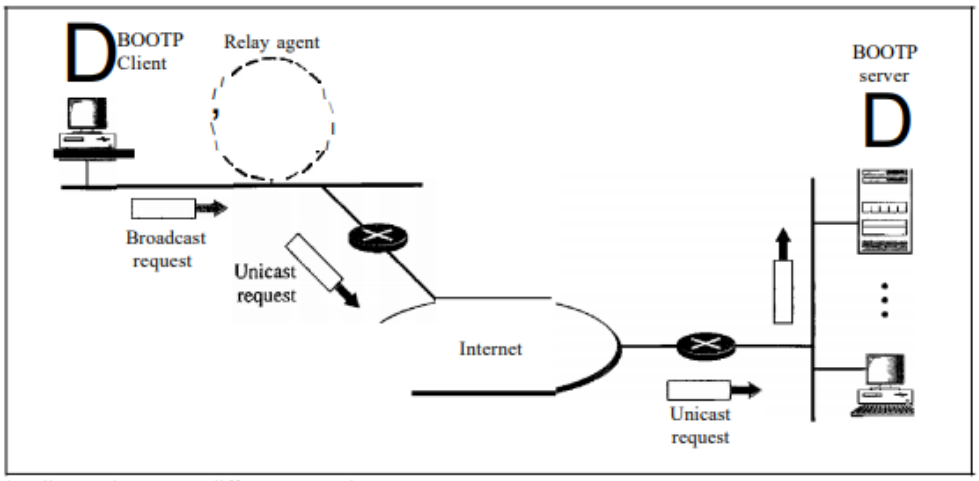
BOOTP

The Bootstrap Protocol (BOOTP) is a client/server protocol designed to provide physical address to logical address mapping. BOOTP is an application layer protocol. The administrator may put the client and the server on the same network or on different networks, as shown in Figure 21.7. BOOTP messages are encapsulated in a UDP packet, and the UDP packet itself is encapsulated in an IP packet.

Figure 21.7 *BOOTP client and server on the same and different network*



a. Client and server on the same network



b. Client and server on different networks

A client can send an IP datagram when it knows neither its own IP address (the source address) nor the server's IP address (the destination address) by using all 1 as as the source address and all 1s as the destination address.

One of the advantages of BOOTP over RARP is that the client and server are application-layer processes.

One problem that must be solved. The BOOTP request is broadcast because the client does not know the IP address of the server. A broadcast IP datagram cannot pass through any router. To solve the problem, there is a need for an intermediary. One of the hosts (or a router that can be configured to operate at the application layer) can be used as a relay. The host in this case is called **a relay agent**. The relay agent knows the unicast address of a BOOTP server. When it receives this type of packet, it encapsulates the message in a unicast datagram and sends the request to the BOOTP server. The packet, carrying a unicast destination address, is routed by any router and reaches the BOOTP server. The BOOTP server knows the message comes from a relay agent because one of the fields in the request message defines the IP address of the relay agent. The relay agent, after receiving the reply, sends it to the BOOTP client.

BOOTP is not a dynamic configuration protocol. When a client requests its IP address, the BOOTP server consults a table that matches the physical address of the client with its IP address. This implies that the binding between the physical address and the IP address of the client already exists. The binding is predetermined.

If a host moves from one physical network to another. If a host wants a temporary IP address. BOOTP cannot handle these situations because the binding between the physical and IP addresses is static and fixed in a table until changed by the administrator. BOOTP is a static configuration protocol

Internet Multicasting – IGMP:

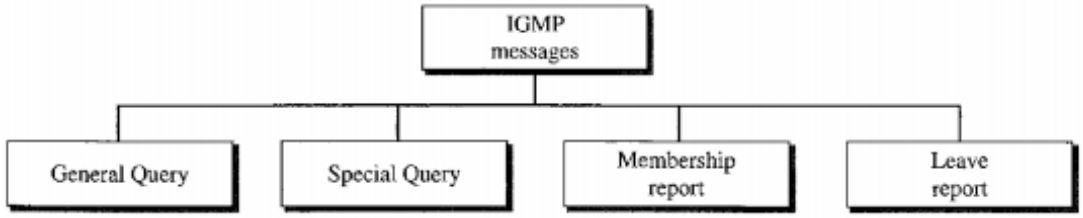
The IP protocol can be involved in two types of communication**: unicasting** and **multicasting**. **Unicasting** is the communication between one sender and one receiver. It is a one-to-one communication. Processes which send the same message to a large number of receivers simultaneously is called **multicasting**, which is a one-to-many communication. Examples are updating replicated, distributed databases, transmitting stock quotes to multiple brokers, and handling digital conference (i.e., multiparty) telephone calls.

The Internet Group Management Protocol (IGMP) is one of the necessary, but not sufficient protocols that is involved in multicasting. Group Management For multicasting in the Internet we need routers that are able to route multicast packets. The routing tables of these routers must be updated by using one of the multicasting routing protocols.

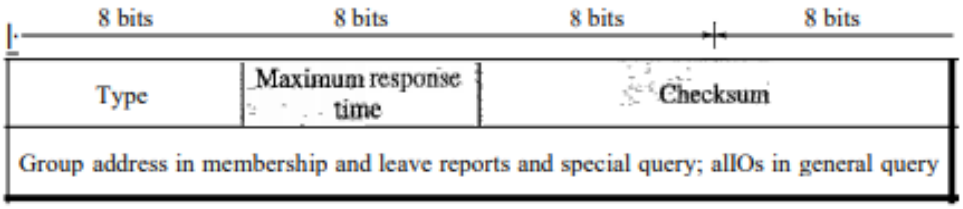
IGMP is not a multicasting routing protocol; it is a protocol that **manages group membership**. In any network, there are one or more multicast routers that distribute multicast packets to hosts or other routers. The IGMP protocol gives the multicast routers information about the membership status of hosts (routers) connected to the network. IGMP is a group management protocol. It helps a multicast router create and update a list of loyal members related to each router interface.

IGMP Messages:

IGMPv2 has three types of messages: the query, the membership report, and the leave report. There are two types of query messages: general and special.



Message Format:



Type. This 8-bit field defines the type of message, as shown in Table 21.1. The value of the type is shown in both hexadecimal and binary notation.

Table 21.1 *IGMP type field*

Type	*Value*
General or special query	0x11 or 00010001
Membership report	0x16 or 00010110
Leave report	0x17 or 00010111

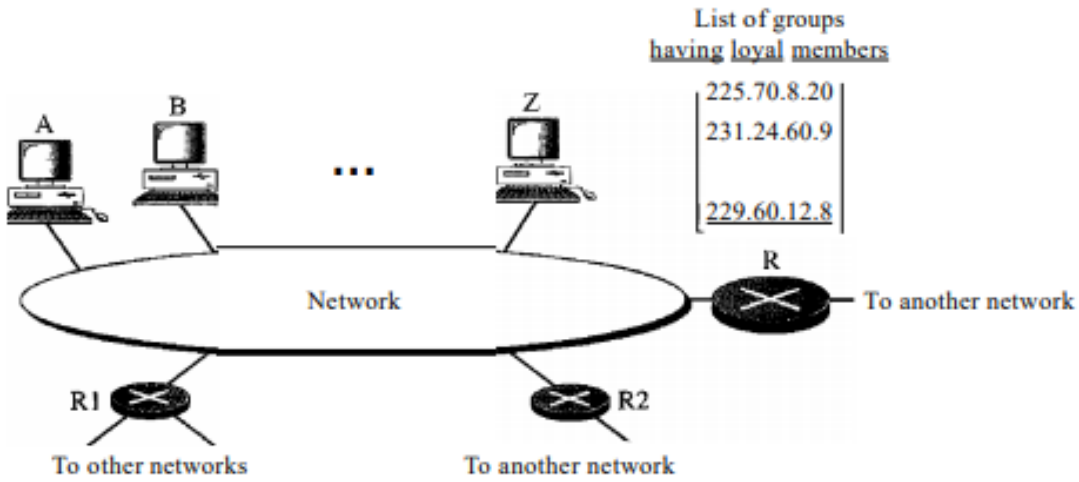
Maximum Response Time. This 8-bit field defines the amount of time in which a query must be answered. The value is in tenths of a second; for example, if the value is 100, it means 10 s. The value is nonzero in the query message; it is set to zero in the other two message types.

Checksum. This is a 16-bit field carrying the checksum. The checksum is calculated over the 8-byte message.

Group address. The value of this field is 0 for a general query message. The value defines the groupid (multicast address of the group) in the special query, the membership report, and the leave report messages.

IGMP Operation

IGMP operates locally. A multicast router connected to a network has a list of multicast addresses of the groups with at least one loyal member in that network.



For each group, there is one router that has the duty of distributing the multicast packets destined for that group. This means that if there are three multicast routers connected to a network, their lists of groupids are mutually exclusive. For example, in Figure only router R distributes packets with the multicast address of 225.70.8.20. A host or multicast router can have membership in a group. When a **host has membership**, it means that one of its processes (an application program) receives multicast packets from some group. When a **router has membership**, it means that a network connected to one of its other interfaces receives these multicast packets. We say that the host or the router has **an interest in the group**. In both cases, the host and the router keep a list of groupids and relay their interest to the distributing router.

In the Figure, router R is the distributing router. There are two other multicast routers (R1 and R2) that, depending on the group list maintained by router R, could be the recipients of router R in this network. Routers RI and R2 may be distributors for some of these groups in other networks, but not on this network.

1. **Joining a Group**

A host or a 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 the first entry for this particular group, the host sends a membership report message. If this is not the first entry, there is no need to send the membership report. The protocol requires that the membership **report be sent twice**, one after the other within a few moments. In this way, if the first one is lost or damaged, the second one replaces it.

2. **Leaving a Group**

When a host sees that no process is interested in a specific group, it sends a leave report. Similarly, when a router sees that none of the networks connected to its interfaces is interested in a specific group, it sends a leave report about that group. When a multicast router receives a leave report, 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.

3. Monitoring membership

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. 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. The general query message does not define a particular group.

The query message has a maximum response time of 10. 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**. The query message must be sent by only one router called **the query router**, also to prevent unnecessary traffic.

4. Delayed Response

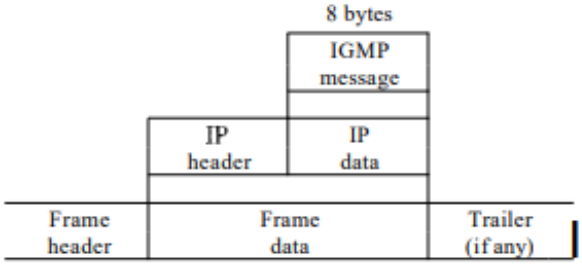
To prevent unnecessary traffic, IGMP uses a delayed response strategy. When a host or router receives a query message, it does not respond immediately; it delays the response. Each host or router uses a random number to create a timer, which expires between I and lOs. A timer is set for each group in the list. For example, the timer for the first group may expire in 2 s, 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, expires earlier, that host or router sends a membership report. Because, the report is broadcast, the waiting host or router receives the report and knows that there is no need to send a duplicate report for this group; thus, the waiting station cancels its corresponding timer.

5. Query Router

Query messages may create a lot of responses. To prevent unnecessary traffic, IGMP designates one router as the query router for each network. Only this designated router sends the query message, and the other routers are passive.

Encapsulation

The IGMP message is encapsulated in an IP datagram, which is itself encapsulated in a frame.



Encapsulation at Network Layer

The value of the protocol field is 2 for the IGMP protocol. Every IP packet carrying this value in its protocol field has data delivered to the IGMP protocol. When the message is encapsulated in the IP datagram, the value of TTL must be 1. This is required because the domain of IGMP is the LAN. No IGMP message must travel beyond the LAN. A TTL value of 1 guarantees that the

message does not leave the LAN since this value is decremented to 0 by the next router and, consequently, the packet is discarded.

Encapsulation at Data Link Layer
At the network layer, the IGMP message is encapsulated in an IP packet and is treated as an IP packet. However, because the IP packet has a multicast IP address, the ARP protocol cannot find the corresponding MAC (physical) address to forward the packet at the data link layer. What happens next depends on whether the underlying data link layer supports physical multicast addresses.

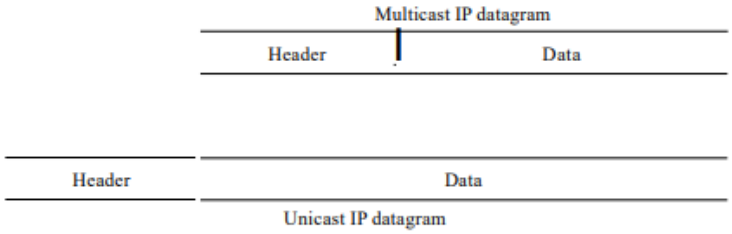
Physical Multicast Support Most LANs support physical multicast addressing. An Ethernet physical address (MAC address) is six octets (48 bits) long. If the first 25 bits in an Ethernet address identifies a physical multicast address for the TCP/IP protocol. The remaining 23 bits can be used to define a group. **To convert an IP multicast address into an Ethernet address**, the multicast router extracts the least significant 23 bits of a class D IP address and inserts them into a multicast Ethernet physical address.

An Ethernet multicast physical address is in the range 01 :00:5E:00:OO:OO to 01:00:5E:7F:FF:FF.

Change the multicast IP address 230.43.14.7 to an Ethernet multicast physical address.
Solution in two steps:
a. Write the rightmost 23 bits of the IP address in hexadecimal. This can be done by changing the rightmost 3 bytes to hexadecimal and then subtracting 8 from the leftmost digit if it is greater than or equal to 8. In our example, the result is 2B:OE:07.
b. We add the result of part a to the starting Ethernet multicast address, which is 01:00:5E:00:00:00. The result is 01:00:5E:2B:OE:07

Change the multicast IP address 238.212.24.9 to an Ethernet multicast address.
Solution
a. The rightmost 3 bytes in hexadecimal is D4: 18:09. We need to subtract 8 from the leftmost digit, resulting in 54:18:09.
b. We add the result of part a to the Ethernet multicast starting address. The result is 01:00:5E:54: 18:09

No Physical Multicast Support Most WANs do not support physical multicast addressing. To send a multicast packet through these networks, a process called **tunneling** is used. In tunneling, the multicast packet is encapsulated in a unicast packet and sent through the network, where it emerges from the other side as a multicast packet.



EXTERIOR ROUTING PROTOCOLS – BGP

Border Gateway Protocol (BGP) is an interdomain routing protocol using path vector routing. The Internet is divided into hierarchical domains called **autonomous systems.** Autonomous systems are divided into three categories: stub, multihomed, and transit.

Stub AS.
A stub AS has only one connection to another AS. The interdomain 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 stub AS is either a source or a sink. A good example of a stub AS is a small corporation or a small local ISP.

Multihomed AS.
A multihomed 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. It does not allow data coming from one AS and going to another AS to pass through. A good example of a multihomed AS is a large corporation that is connected to more than one regional or national AS that does not allow transient traffic.

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).

Attributes are divided into two broad categories: well known and optional. **A well-known** attribute is one that every BGP router must recognize. An **optional attribute** is one that needs not be recognized by every router.

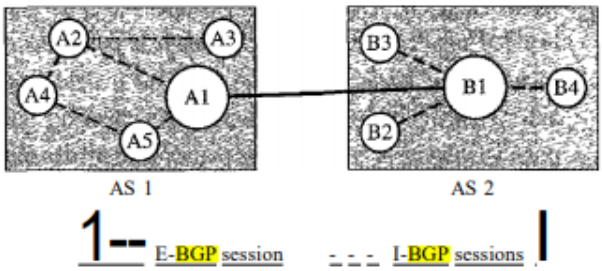
Well-known attributes are themselves divided into two categories: mandatory and discretionary. **A well-known mandatory attribute** is one that must appear in the description of a route. A well-known discretionary attribute is one that must be recognized by each router, but is not required to be included in every update message.
Examples:
ORIGIN. This defines the source of the routing information (RIP, OSPF, and so on).
AS_PATH. This defines the list of autonomous systems through which the destination can be reached.
NEXT-HOP, which defines the next router to which the data packet should be sent.

The optional attributes can also be subdivided into two categories: transitive and nontransitive. An **optional transitive attribute** is one that must be passed to the next router by the router that has not implemented this attribute.
An **optional nontransitive attribute** is one that must be discarded if the receiving router has not implemented it.

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. For this reason, BGP sessions are sometimes referred to as **semi-permanent connections**.

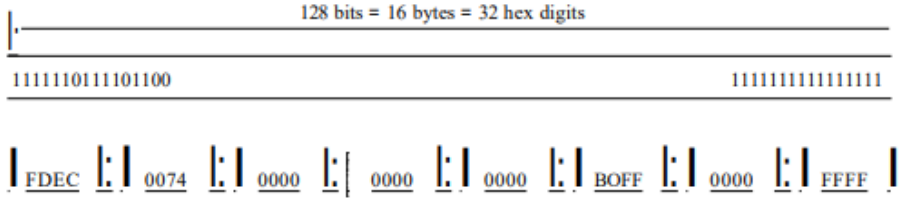
BGP can have two types of sessions: **external BGP (E-BGP)** and **internal BGP (I-BGP)** sessions.
The **E-BGP session** is used to exchange information between two speaker nodes belonging to two different autonomous systems.
The **I-BGP session**, on the other hand, is used to exchange routing information between two routers inside an autonomous systems.



IPV6 – ADDRESSING – ISSUES

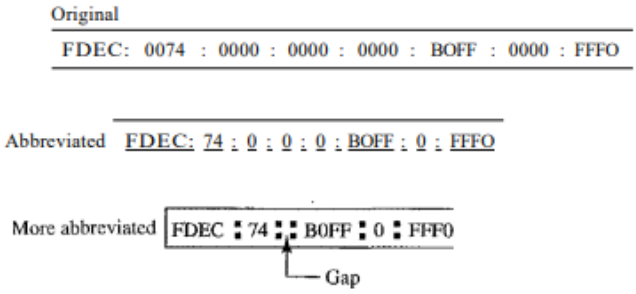
An IPv6 address consists of 16 bytes (octets); it is 128 bits long. IPv6 specifies hexadecimal colon notation. In this notation, 128 bits is divided into eight sections, each 2 bytes in length. Two bytes in hexadecimal notation requires four hexadecimal digits. Therefore, the address consists of 32 hexadecimal digits, with every four digits separated by a colon.

Figure 19.14 *IPv6 address in binary and hexadecimal colon notation*



Although the IP address, even in hexadecimal format, is very long, many of the digits are zer9s. In this case, we can abbreviate the address. The leading zeros of a section (four digits between two colons) can be omitted. Only the leading zeros can be dropped, not the trailing zeros.

Figure 19.15 *Abbreviated IPv6 addresses*



0074 can be written as 74, OOOF as F, and 0000 as O. We can remove the zeros altogether and replace them with a double semicolon. Note that this type of abbreviation is allowed only once per address. If there are two runs of zero sections, only one of them can be abbreviated.

Expand the address 0:15::1:12:1213 to its original.

```
xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx
 0:   15:                      1:   12:1213
```

This means that the original address is

0000:0015:0000:0000:0000:0001 :0012: 1213

Address Space

IPv6 has 2^{128} addresses available. IPv6 divided the address into several categories. A few leftmost bits, called **the type prefix**, in each address define its category. The type prefix is variable in length, but it is designed such that no code is identical to the first part of any other code. In this way, there is no ambiguity; when an address is given, the type prefix can easily be determined.

Table 19.5 *Type prefixes for IPv6 addresses*

Type Prefix	*Type*	*Fraction*
00000000	Reserved	1/256
00000001	Unassigned	1/256
0000001	ISO network addresses	1/128
0000010	IPX (Novell) network addresses	1/128
0000011	Unassigned	1/128
00001	Unassigned	1/32
0001	Reserved	*1/16*
001	Reserved	1/8
010	Provider-based unicast addresses	*1/8*

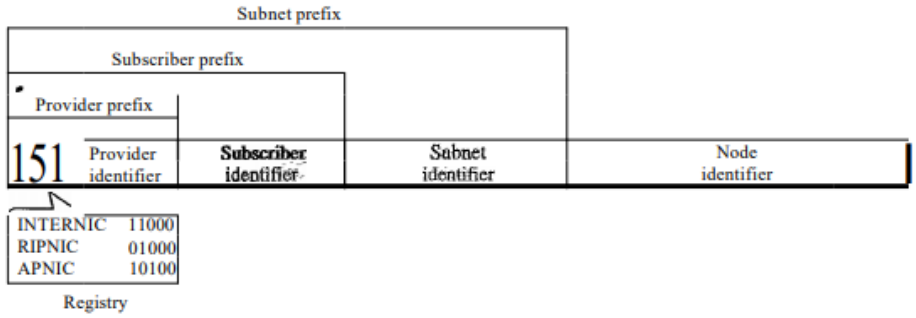
Table 19.5 *Type prefixes for IPv6 addresses (continued)*

Type Prefix	*Type*	*Fraction*
011	Unassigned	1/8
100	Geographic-based unicast addresses	1/8
101	Unassigned	1/8
110	Unassigned	1/8
1110	Unassigned	1116
11110	Unassigned	1132
1111 10	Unassigned	1/64
1111 110	Unassigned	1/128
11111110 a	Unassigned	1/512
1111 111010	Link local addresses	111024
1111 1110 11	Site local addresses	1/1024
11111111	Multicast addresses	1/256

Unicast Addresses

A unicast address defines a single computer. The packet sent to a unicast address must be delivered to that specific computer. IPv6 defines two types of unicast addresses: **geographically based** and **provider-based**. The provider-based address is generally used by a normal host as a unicast address.

Figure 19.16 *Prefixes for provider-based unicast address*

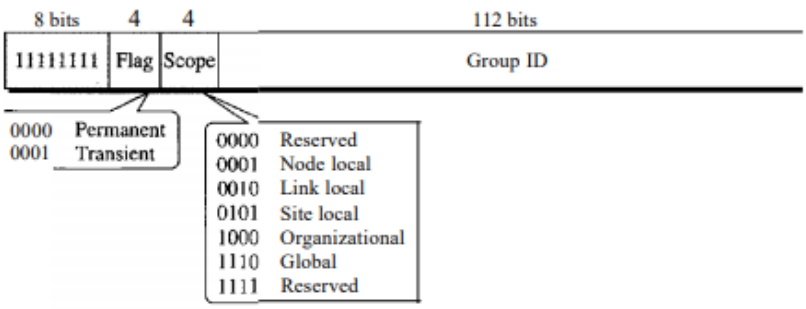


Type identifier. This 3-bit field defines the address as a provider-based address.
Registry identifier. This 5-bit field indicates the agency that has registered the address.
Provider identifier. This variable-length field identifies the provider for Internet access (such as an ISP). A 16-bit length is recommended for this field.
Subscriber identifier. When an organization subscribes to the Internet through a provider, it is assigned a subscriber identification. A 24-bit length is recommended for this field.
Subnet identifier. Each subscriber can have many different subnetworks, and each subnetwork can have an identifier. The subnet identifier defines a specific subnetwork under the territory of the subscriber. A 32-bit length is recommended for this field.
Node identifier. The last field defines the identity of the node connected to a subnet. A length of 48 bits is recommended for this field to make it compatible with the 48-bit link (physical) address used by Ethernet. In the future, this link address will probably be the same as the node physical address.

Multicast Addresses

Multicast addresses are used to define a group of hosts instead of just one. A packet sent to a multicast address must be delivered to each member of the group.

Figure 19.17 *Multicast address in IPv6*

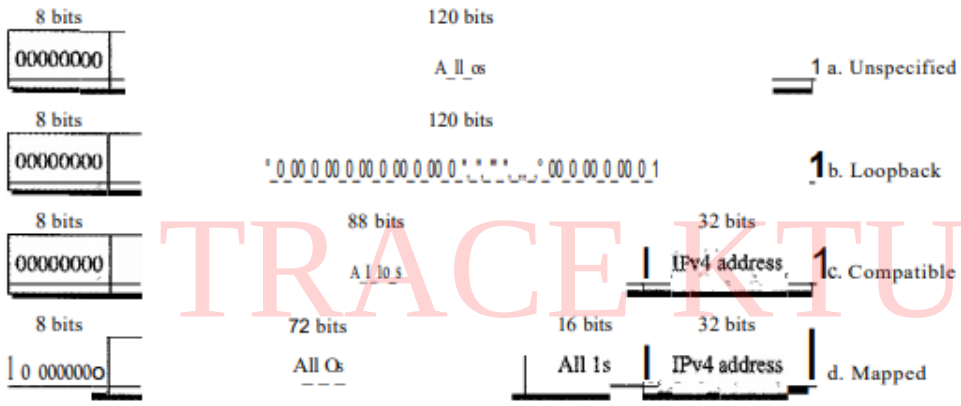


A flag that defines the group address as either permanent or transient. A **permanent group** address is defined by the Internet authorities and can be accessed at all times. A **transient group** address, on the other hand, is used only temporarily. Systems engaged in a teleconference, for example, can use a transient group address.
The third field defines **the scope** of the group address. Many different scopes are provided.

Anycast addresses.
An anycast address, like a multicast address, also defines a group of nodes. However, a packet destined for an anycast address is delivered to only one ofthe members ofthe anycast group, the nearest one (the one with the shortest route)

Reserved Addresses
These addresses start with eight Os (type prefix is 00000000). A few subcategories are defined in this category, as shown in Figure 19.18.

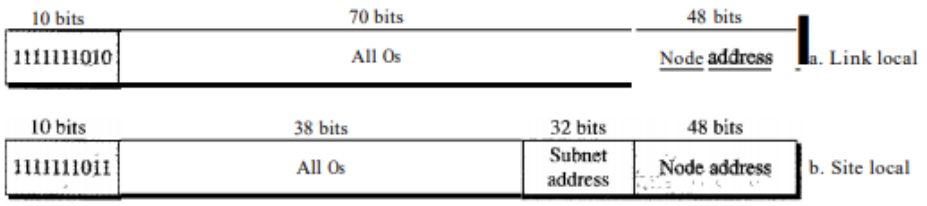
Figure 19.18 *Reserved addresses in IPv6*



An unspecified address is used when a host does not know its own address and sends an inquiry to find its address.
A loopback address is used by a host to test itself without going into the network.
A compatible address is used during the transition from IPv4 to IPv6. It is used when a computer using IPv6 wants to send a message to another computer using IPv6, but the message needs to pass through a part of the network that still operates in IPv4.
A mapped address is also used during transition. However, it is used when a computer that has migrated to IPv6 wants to send a packet to a computer still using IPv4.

Local Addresses
These addresses are used when an organization wants to use IPv6 protocol without being connected to the global Internet. In other words, they provide addressing for private networks. Nobody outside the organization can send a message to the nodes using these addresses. Two types of addresses are defined for this purpose:

Figure 19.19 *Local addresses in IPv6*



A link local address is used in an isolated subnet; a site local address is used in an isolated site with several subnets.

Advantages

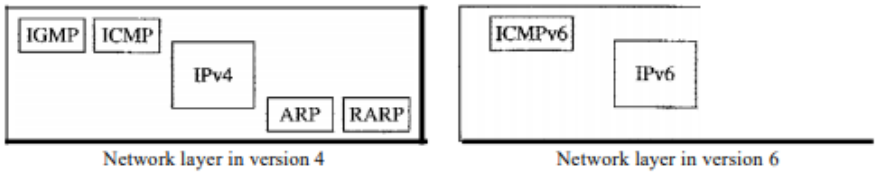
The next-generation IP, or IPv6, has some advantages over IPv4 that can be summarized as follows:

- Larger address space. An IPv6 address is 128 bits long, as we discussed in Chapter 19. Compared with the 32-bit address of IPv4, this is a huge (296) increase in the address space. O
- Better header format. IPv6 uses a new header format in which options are separated from the base header and inserted, when needed, between the base header and the upper-layer data. This simplifies and speeds up the routing process because most of the options do not need to be checked by routers.
- New options. IPv6 has new options to allow for additional functionalities.
- Allowance for extension. IPv6 is designed to allow the extension of the protocol if required by new technologies or applications.
- Support for resource allocation. In IPv6, the type-of-service field has been removed, but a mechanism (calledjlow label) has been added to enable the source to request special handling of the packet. This mechanism can be used to support traffic such as real-time audio and video.
- Support for more security. The encryption and authentication options in IPv6 provide confidentiality and integrity of the packet.

ICMPv6.

Comparison of the network layer of version 4 to version 6:

Figure 21.23 *Comparison ofnetwork layers in version* 4 *and version 6*



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.

Error Reporting

As we saw in our discussion of version 4, one of the main responsibilities of ICMP is to report errors. Five types of errors are handled: destination unreachable, packet too big, time exceeded, parameter problems, and redirection. ICMPv6 forms an error packet, which is then encapsulated in an IP datagram. This is delivered to the original source of the failed datagram.

Table 21.3 *Comparison of error-reporting messages in ICMPv4 and ICMPv6*

Type of Message	*Version 4*	*Version 6*
Destination unreachable	Yes	Yes
Source quench	Yes	No
Packet too big	No	Yes
Time exceeded	Yes	Yes
Parameter problem	Yes	Yes
Redirection	Yes	Yes

The source-quench message is eliminated in version 6 because the priority and the flow label fields allow the router to control congestion and discard the least important messages. In this version, there is no need to inform the sender to slow down. The packet-too-big message is added because fragmentation is the responsibility of the sender in IPv6. If the sender does not make the right packet size decision, the router has no choice but to drop the packet and send an error message to the sender.

- **Packet Too Big**

This is a new type of message added to version 6. If a router receives a datagram that is larger than the maximum transmission unit (MTU) size of the network through which the datagram should pass, two things happen. First, the router discards the datagram and then an ICMP error packet-a packet-too-big message-is sent to the source.

Query

ICMP can diagnose some network problems. This is accomplished through the query messages. Four different groups of messages have been defined: echo request and reply, router solicitation and advertisement, neighbor solicitation and advertisement, and group membership.

Table 21.4 *Comparison of query messages in ICMPv4 and ICMPv6*

Type of Message	*Version 4*	*Version 6*
Echo request and reply	Yes	Yes
Timestamp request and reply	Yes	No
Address-mask request and reply	Yes	No
Router solicitation and advertisement	Yes	Yes
Neighbor solicitation and advertisement	ARP	Yes
Group membership	IGMP	Yes

Two sets of query messages are eliminated from ICMPv6: time-stamp request and reply- and address-mask request and reply. The timestamp request and reply messages **are eliminated** because they are implemented in other protocols such as TCP and because they were rarely used in the past. The address-mask request and reply messages are eliminated in IPv6 because the subnet section of an address allows the subscriber to use up to 232 - 1 subnets. Therefore, subnet masking, as defined in IPv4, is not needed here.

Neighbor Solicitation and Advertisement

the network layer in version 4 contains an independent protocol called Address Resolution Protocol (ARP). In version 6, this protocol is eliminated, and its duties are included in ICMPv6. The idea is exactly the same, but the format of the message has changed.

Group Membership

The network layer in version 4 contains an independent protocol called IGMP. In version 6, this protocol is eliminated, and its duties are included in ICMPv6. The purpose is exactly the same.