SREELA CHANDRAN, AP

CST 303 COMPUTER NETWORKS

NETWORK LAYER IN INTERNET

MODULE 4

IP protocol, IP addresses, Internet Control Message Protocol (ICMP), Address Resolution Protocol (ARP), Reverse Address Resolution Protocol (RARP), Bootstrap Protocol (BOOTP), Dynamic Host Configuration Protocol (DHCP). Open Shortest Path First(OSPF) Protocol, Border Gateway Protocol (BGP), Internet multicasting, IPv6, ICMPv6.

NETWORK LAYER IN INTERNET

- At the network layer, the Internet can be viewed as a collection of subnetworks or autonomous systems that are connected together
- The glue that holds the Internet together is the network layer protocol, IP (Internet Protocol)
- Its job is to provide a best - efforts way to transport datagrams from source to destination, without regard to whether or not these machines are on the same network, or whether or not there are other networks in between them.
- Communication in the Internet works as follows. The transport layer takes data streams and breaks them up into datagrams. In theory, datagrams can be up to 64 Kbytes each, but in practice they are usually around 1500 bytes. Each datagram is transmitted through the Internet, possibly being fragmented into smaller units as it goes. When all the pieces finally get to the destination machine, they are reassembled by the network layer into the original datagram. This datagram is then handed to the transport layer, which inserts it into the receiving process input stream.

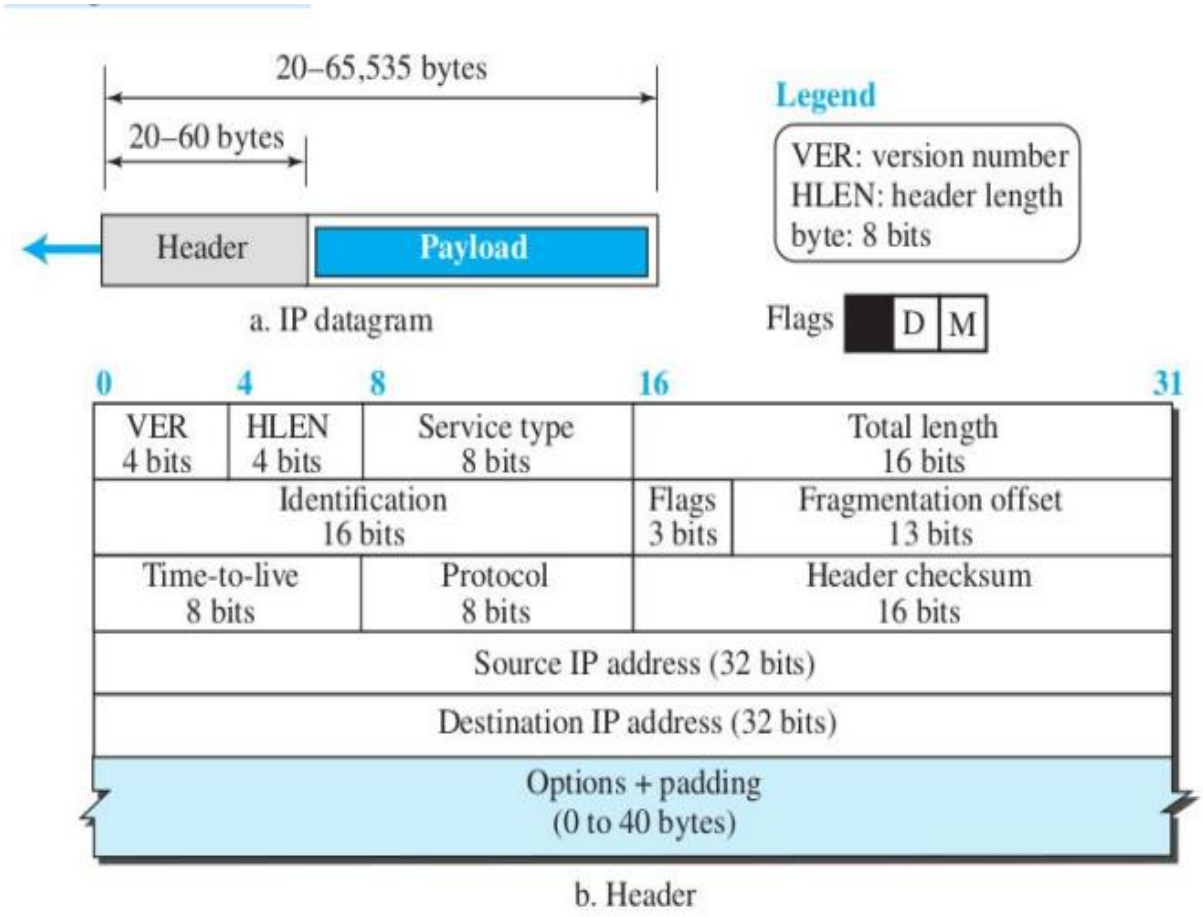
The IP Protocol

- A connectionless protocol.
- An unreliable datagram protocol to provide best-effort delivery service.
- IPv4 packets can be corrupted, be lost, arrive out of order, or be delayed, and may create congestion for the network.

- If reliability is important, IPv4 must be paired with a reliable transport-layer protocol such as TCP.
- Supported by many auxiliary protocols, such as;
    - The Internet Group Management Protocol (IGMP) helps IPv4 in multicasting.
    - Internet Control Message Protocol (ICMPv4) helps IPv4 to handle someerrors that may occur in the network-layer delivery.
    - Address Resolution Protocol (ARP) is used to glue the network anddata-link layers in mapping network-layer addresses to link-layeraddresses.

IPV4 Datagram Format

- An IP datagram consists of a header part and a payload part.
- The header has a 20-byte fixed part and a variable length optional part.



Version Number:

- 4-bit
- Defines the version of the IPv4 protocol, which has the value of 4 (01002

Header Length:

- 4-bit
- Defines the total length of the datagram header in 4-byte words.
- Varies from 5 to 15 (means 20 bytes to 60 bytes)
- In binary, 0101 to 1111
- Eg: If the value in the header length field is 8, it means the header has 8x4=32 bytes, which also means 20 bytes basic header and 12 bytes options.

Service Type:

- 8-bits
- Specifies differentiated services (DiffServ) for different types of protocols.

Total Length:

- 16 bit field.
- Defines the total length (header plus data) of the IP datagram in bytes.
- Most useful when zero-padding occurs due to small datagram size (Standard Ethernet frame needs minimum 46 bytes payload).

Identification, Flags, and Fragmentation Offset:

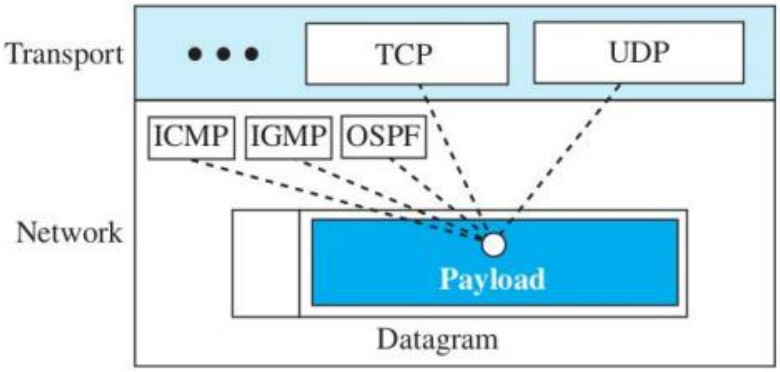
- Total 32 bits
- Related to the fragmentation of the IP datagram when the size of the datagram is larger than the underlying network can carry.

Time-to-live (TTL)

- 8 bits
- To control the maximum number of hops (routers) visited by the
- datagram.
- When a source host sends the datagram, it stores a number in this field (which may be double the maximum number routers up to the destination).
- Each router decrements this value when the datagram reaches the router.
- When the TTL becomes zero, the packet is discarded.

Protocol

- 8 bits.
- Specifies the protocol of the payload.



Some protocol values

ICMP	01
IGMP	02
TCP	06
UDP	17
OSPF	89

Header checksum:

- 16 bits.
- Set at the source to protect the header from errors.
- Also recalculated at the routers as values of some fields such as TTL changes at routers.
- Checksum is the complement of the sum of other fields calculated using 1s complement arithmetic.

Source and Destination Addresses:

- 4 bytes each.

Options:

- Single-byte and multiple-byte options.
- Used for testing and debugging

FRAGMENTATION

- Occurs if the datagram size is greater than the Maximum Transfer Unit(MTU) of the data link layer.
- Fragmentation is the process of dividing the datagram into smallfragments so as to limit the size within the MTU.
- MTU: the maximum size of the payload that can be encapsulated in aframe; it is the restriction imposed by the hardware and software.

  MTU of Standard Ethernet is 1500 bytes.
- Occurs at the source or at the intermediate routers.

- Reassembly occurs at the destination only, as the fragments may go through different routes.
- Fields Related to Fragmentation
    - Identification - 16 bits
    - Flags - 3 bits
    - Fragmentation offset - 13 bits

Identification field: 16 bits

- Identifies a datagram originating from the source host.
- All fragments of a datagram will have the same identification number.
- Helps the destination in reassembling the datagram.
- The combination of the identification and source IP address must uniquely define a datagram as it leaves the source host.
- The source uses a counter for this field.

Flags field: 3 bits

1. Reserved (not used)
2. D bit (Do not fragment bit):

    Value 0: Can be fragmented, if needed.

    Value 1:The machine must not fragment the datagram. Discards the datagram and sends an ICMP error message to the source host
3. M bit (More fragments bit):

    Value 0: This is the last/only fragment of the datagram.

    Value 1: There are more fragments after it.

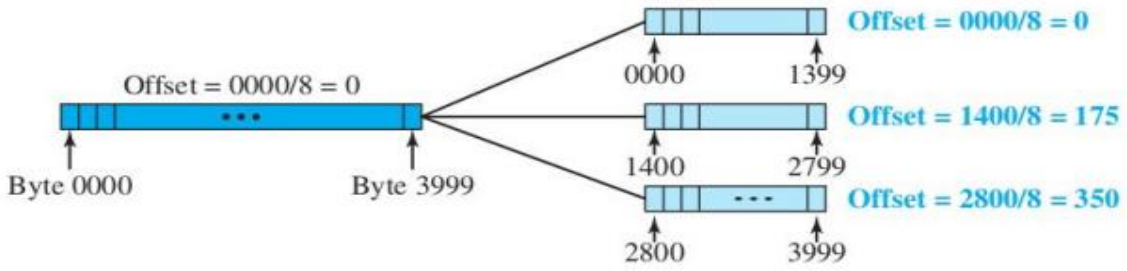
Fragmentation offset field: 13 bits

Shows the relative position of this fragment with respect to the whole datagram.

It is the offset of the data in the original datagram measured in units of 8 bytes.

Calculation Example:

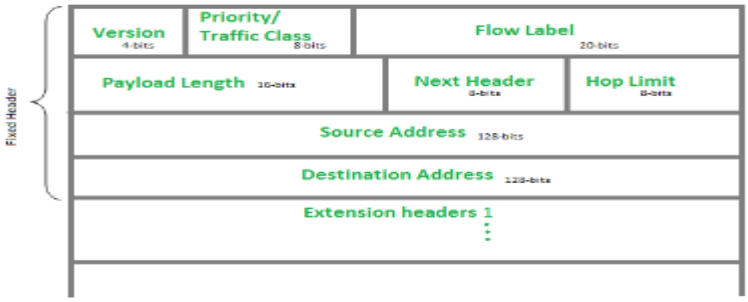
Consider a datagram with 4000 bytes data; divided with a maximum of 1400 bytes. (data bytes numbered 0 to 3999)



IPv6 Header

The IPv6 header is a part of the data packet structure used in Internet Protocol version 6 (IPv6), which is the latest version of the Internet Protocol. IPv6 is designed to replace IPv4, offering a much larger address space and improved features. The header in IPv6 contains important information needed for routing and delivering packets across networks.

he IPv6 header representation is a structured layout of fields in an IPv6 packet, including source and destination addresses, traffic class, flow label, payload length, next header, and hop limit. It ensures proper routing and delivery of data across networks.



Version (4-bits)

The size of this field is 4-bit. Indicates the version of the Internet Protocol, which is always 6 for IPv6, so the bit sequence is 0110.

Traffic Class(8-bit)

These 8 bits are divided into two parts. The most significant 6 bits are used for Type of Service to let the Router Known what services should be provided to this packet. The least significant 2 bits are used for Explicit Congestion Notification (ECN).

Flow Label (20-bits)

This label is used to maintain the sequential flow of the packets belonging to a communication. The source labels the sequence to help the router identify that a particular packet belongs to a specific flow of information. This field helps avoid re-ordering of data packets.

Payload Length (16-bits)

This field is used to tell the routers how much information a particular packet contains in its payload. Payload is composed of Extension Headers and Upper Layer data. With 16 bits, up to 65535 bytes can be indicated

Next Header(8 bits)

This field is used to indicate either the type of Extension Header, or if the Extension Header is not present then it indicates the Upper Layer PDU. The values for the type of Upper Layer PDU are same as IPv4's.

Hop Limit (8-bits)

This field is used to stop packet to loop in the network infinitely. This is same as TTL in IPv4. The value of Hop Limit field is decremented by 1 as it passes a link (router/hop). When the field reaches 0 the packet is discarded.

Source Address (128-bits)

Source Address is the 128-bit IPv6 address of the original source of the packet.

Destination Address (128-bits)

The destination Address field indicates the IPv6 address of the final destination(in most cases). All the intermediate nodes can use this information in order to correctly route the packet.

Extension Headers

In IPv6, the Fixed Header contains only that much information which is necessary, avoiding those information which is either not required or is rarely used. All such information is put between the Fixed Header and the Upper layer header in the form of Extension Headers. Each Extension Header is identified by a distinct value.

When Extension Headers are used, IPv6 Fixed Header's Next Header field points to the first Extension Header. If there is one more Extension Header, then the first Extension Header's 'Next-Header' field points to the second one, and so on. If the Next Header field contains the value 59, it indicates that there are no headers after this header, not even Upper Layer Header.

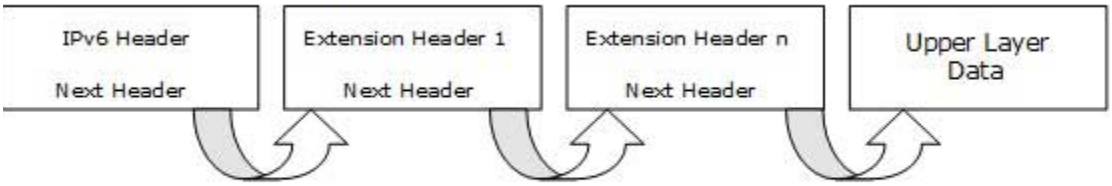
The following Extension Headers must be supported.

Extension Header	Next Header Value
Hop-by-Hop Options header	0
Routing header	43
Fragment header	44
Destination Options header	60
Authentication header	51
Encapsulating Security Payload header	50

The sequence of Extension Headers should be:

IPv6 header
Hop-by-Hop Options header
Destination Options header[1]
Routing header
Fragment header
Authentication header
Encapsulating Security Payload header
Destination Options header[2]
Upper-layer header

Extension Headers are arranged one after another in a linked list manner, as depicted in the following diagram:



[Image: Extension Headers Connected Format]

IP ADDRESSING

The identifier used in the internet layer (network layer) of the TCP/IP protocol suite to identify the connection of each device to the Internet is called the Internet address or Internet Protocol address (IP address).

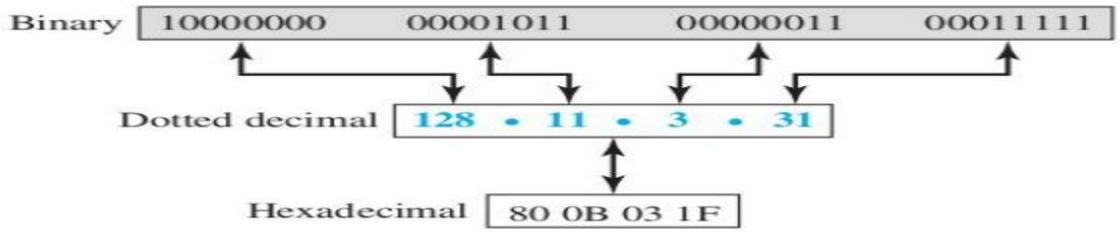
There are two versions of IP addresses: IPv4 (32 bit) and IPv6 (128 bit).

IPv4 address is a 32-bit address that uniquely and universally defines the connection of a host or a router to the Internet.The IP address is the address of the connection, not the host or the router, because if the device is moved to another network, the IP address may be changed.If a device has two connections to the Internet, via two networks, it has two IPv4 addresses.

IPV4 ADDRESS

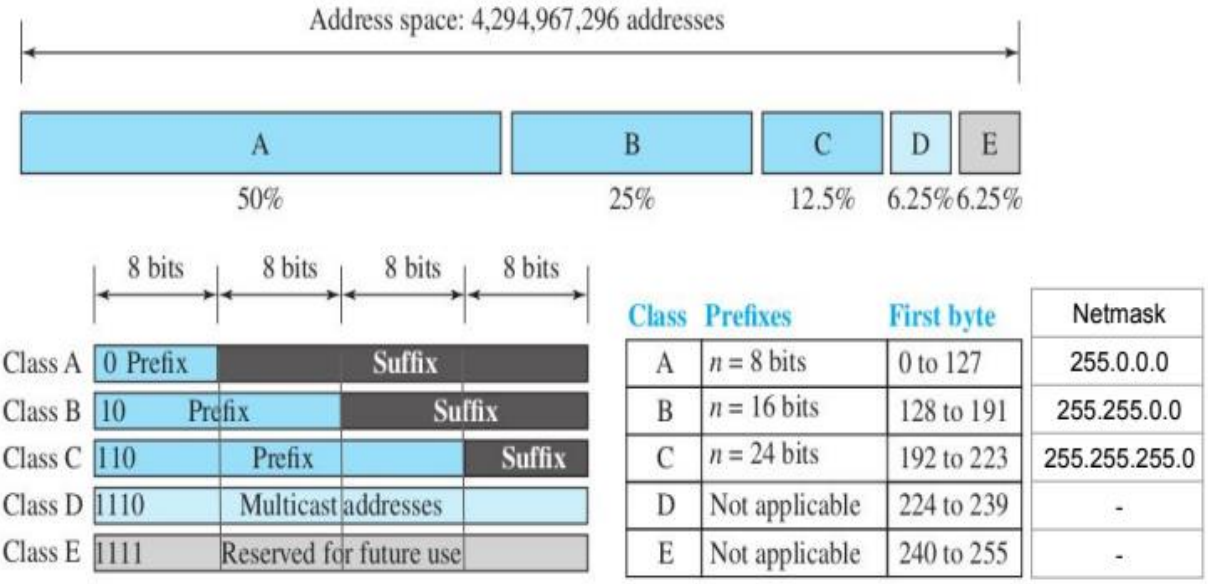
Address Space: An address space is the total number of addresses used by the protocol.IPv4 uses 32-bit addresses, which means that the address space is 232or 4,294,967,296 (more than four billion).

Notation



Classful Addressing

The whole address space was divided into 5 classes: Class A, B, C, D & E. Each address is divided into prefix and suffix.To accommodate both small and large networks, three fixed-length prefixeswere designed (n = 8, n = 16, and n = 24).



Class	Prefixes	First byte	Netmask
A	$n = 8$ bits	0 to 127	255.0.0.0
B	$n = 16$ bits	128 to 191	255.255.0.0
C	$n = 24$ bits	192 to 223	255.255.255.0
D	Not applicable	224 to 239	-
E	Not applicable	240 to 255	-

Class	Leading bits	Size of *network number* bit field	Size of *rest* bit field	Number of networks	Addresses per network	Start address	End address
Class A	0	8	24	128 (2^7)	16,777,216 (2^{24})	0.0.0.0	127.255.255.255
Class B	10	16	16	16,384 (2^{14})	65,536 (2^{16})	128.0.0.0	191.255.255.255
Class C	110	24	8	2,097,152 (2^{21})	256 (2^8)	192.0.0.0	223.255.255.255
Class D (multicast)	1110	not defined	not defined	not defined	not defined	224.0.0.0	239.255.255.255
Class E (reserved)	1111	not defined	not defined	not defined	not defined	240.0.0.0	255.255.255.255

Extracting Information from a Classful Address

Eg: 192.168.1.5

- Class C (ie, 24 bits network part, 8 bits host part)
- Total addresses in this network: 256
- Network: 192.168.1.0
- First host address: 192.168.1.1
- Last host address: 192.168.1.254
- Broadcast address in this network: 192.168.1.255
- Max. number of hosts in this network: 254
- Netmask (or commonly subnet mask): 255.255.255.0

Address depletion in Classful Addressing

- Due to fixed length network fields, there were no more addressesavailable for new organizations and individuals that needed to beconnected to the Internet.
- Eg: Class A can be assigned to only 128 organizations in the world, buteach organization needs to have a single network (seen by the rest ofthe world) with 16,777,216 nodes (computers in this single network).
- Too few organizations with that number of computers.
    - ○ Many Class B addresses remained unused, but couldn't be given toothers.
    - ○ Class C networks are more, but each have very few addresses (256).

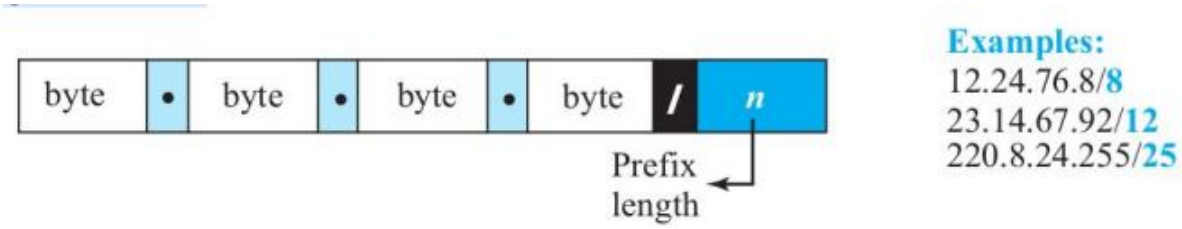
Subnetting and Supernetting in Classful Addressing

- In subnetting, a class A or class B block is divided into several subnets.
- Eg: If a network in class A is divided into four subnets, each subnethas a prefix of nsub = 10 bits.
- Most organizations were unhappy in dividing their network.

- Supernetting was devised to combine several class C blocks into alarger block to be attractive to organizations that need more than the256 addresses available in a class C block.
- Did not work as it made the routing of packets more difficult.

Classless Addressing

- Variable number of prefix bits based on requirement.
- Divides the network into blocks of variable length (ie, number of nodes),which is a power of 2.
- An address in Class A can be thought of as a classless address withprefix length 8, and so on.
- An address with prefix length n is represented with a slash notation asgiven below;



- This notation is officially called as classless interdomain routing or CIDRnotation.

Subnetting in Classless network

It is a method of dividing a single physical network into numerous smaller logical sub-networks. These subnetworks are referred to as subnets. An IP address is formed by combining a network and host segments. A subnet is created by accepting bits from the IP address host part and is used to split the original network into smaller subnetworks.

Supernetting is the inverse of subnetting, in which many networks are combined into a single network. During supernetting, the mask bits are moved to the left of the default mask. It is sometimes referred to as router summarization and aggregation. It leads to the production of more addresses at the cost of network addresses, where network bits are essentially turned into host bits.

COMPARISON OF SUBNETTING AND SUPERNETTING

Features	Subnetting	Supernetting
Definition	It is a method of dividing a single physical network into numerous smaller logical sub-networks.	It is the inverse of subnetting, in which many networks are integrated into a single network.
Purpose	It is utilized to decrease address depletion	It is utilized to simplify and speeds up the routing process.
Procedure	It transforms host bits into network bits and helps to increase the number of network bits.	It converts network bits to host bits and helps to increase the number of host bits.
Mask bits	Mask bits are relocated to the right of the default mask during subnetting.	Supernetting shifts the mask bits to the left of the normal mask.
Implementation	It is implemented via VLSM and FL techniques.	It is implemented via the CIDR technique.

Dividing a network into smaller subnets by increasing the number of bits in the prefix (network part).

eg: We have a network 192.168.2.0/24, ie 254 hosts in this network.

It can be divided into 4 subnets by changing 2 bits from suffix toprefix. It will become four /26 networks.

192.168.2.*00 000000* – 192.168.2.0 Network 1
192.168.2.*00 000001* – 192.168.2.1 First host
192.168.2.*00 111110* – 192.168.2.62 Last host
192.168.2.*00 111111* – 192.168.2.63 Broadcast 1
192.168.2.*01 000000* – 192.168.2.64 Network 2
192.168.2.*01 000001* – 192.168.2.65 First host
192.168.2.*01 111110* – 192.168.2.126 Last host
192.168.2.*01 111111* – 192.168.2.127 Broadcast 2
192.168.2.*10 000000* – 192.168.2.128 Network 3
192.168.2.*10 000001* – 192.168.2.129 First host
192.168.2.*10 111110* – 192.168.2.190 Last host
192.168.2.*10 111111* – 192.168.2.191 Broadcast_3
192.168.2.*11 000000* – 192.168.2.192 Network 4
192.168.2.*11 000001* – 192.168.2.193 First host
192.168.2.*11 111110* – 192.168.2.254 Last host
192.168.2.*11 111111* – 192.168.2.255 Broadcast 4

Subnet mask
255.255.255.*11 000000*
ie, 255.255.255.192

12

Extracting Information from a Classless Address

For an address with prefix length n,

Total number of addresses in the block, N = 232 − n

First address (ie, network address): the address with 32-n rightmostbits set to 0.

Last address (ie, broadcast address): the address with 32-nrightmost bits set to 1.

Eg: 167.199.170.82/ 27

Address: 10100111 11000111 10101010 01010010

First Address: 10100111 11000111 10101010 01000000 167.199.170.64/ 27

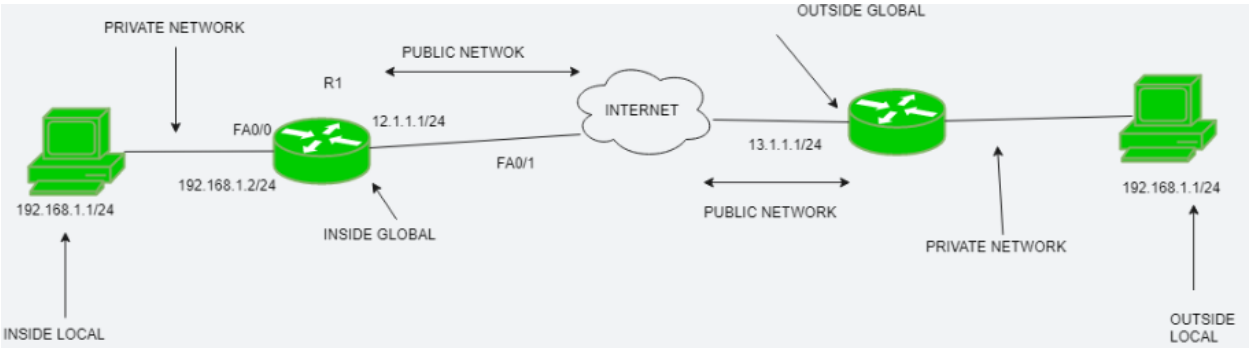
Last Address: 10100111 11000111 10101010 01011111 167.199.170.95/ 27

NETWORK ADDRESS TRANSLATION

One public IP address is needed to access the Internet, but we can use a private IP address in our private network. The idea of NAT is to allow multiple devices to access the Internet through a single public address. To achieve this, a private IP address must be translated into a public IP address.

Network Address Translation (NAT) is a process in which one or more local IP addresses are translated into one or more Global IP addresses and vice versa to provide Internet access to the local hosts. It also does the translation of port numbers, i.e., masks the port number of the host with another port number in the packet that will be routed to the destination. It then makes the corresponding entries of IP address and port number in the NAT table.

Generally, the border router is configured for NAT i.e. the router which has one interface in the local (inside) network and one interface in the global (outside) network. When a packet traverse outside the local (inside) network, then NAT converts that local (private) IP address to a global (public)IP address. When a packet enters the local network, the global (public) IP address is converted to a local (private) IP address.



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.

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 is known as ICMP. The protocol is at the network layer. It is mostly utilized on network equipment like routers and is utilized for error handling at the network layer. Since there are various kinds of network layer faults, ICMP can be utilized to report and troubleshoot these errors.

- **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**

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.

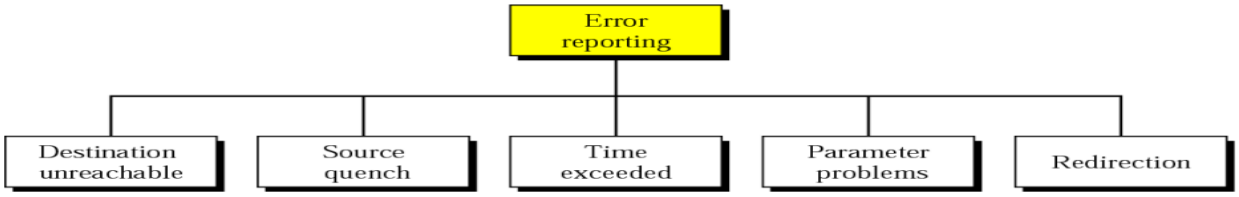
- E.g.: Query messages are used, if a node need **redirect** it message.

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

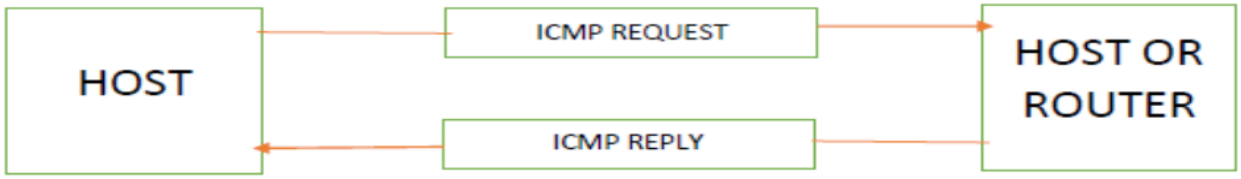
Types of Error messages



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

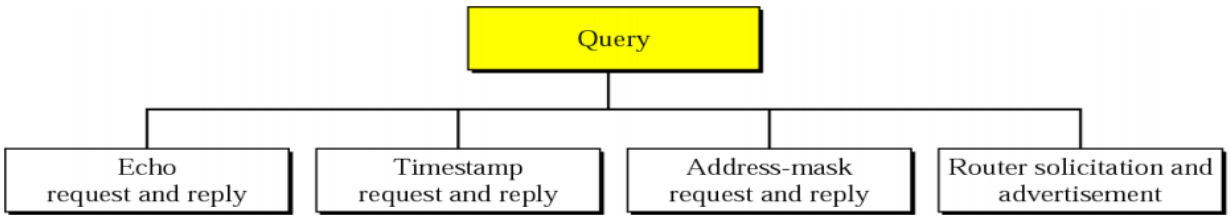
2) Query Message

- 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)

Types of query messages



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.

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.

- 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)

ARP PROTOCOL

ARP stands for “Address Resolution Protocol”. It is a network protocol used to determine the MAC address (hardware address) from any IP address.ARP is used to mapping the IP Address into MAC Address. When one device wants to communicate with another device in a LAN (local area network) network, the ARP protocol is used.IP address is used to communicate with any device at the application layer. But to communicate with a device at the data link layer or to send data to it, a MAC address is required.When data is sent to a local host, the data travels between networks via IP address. But to reach that host in LAN, it needs the MAC address of that host. In this situation the address resolution protocol plays an important role.

- 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

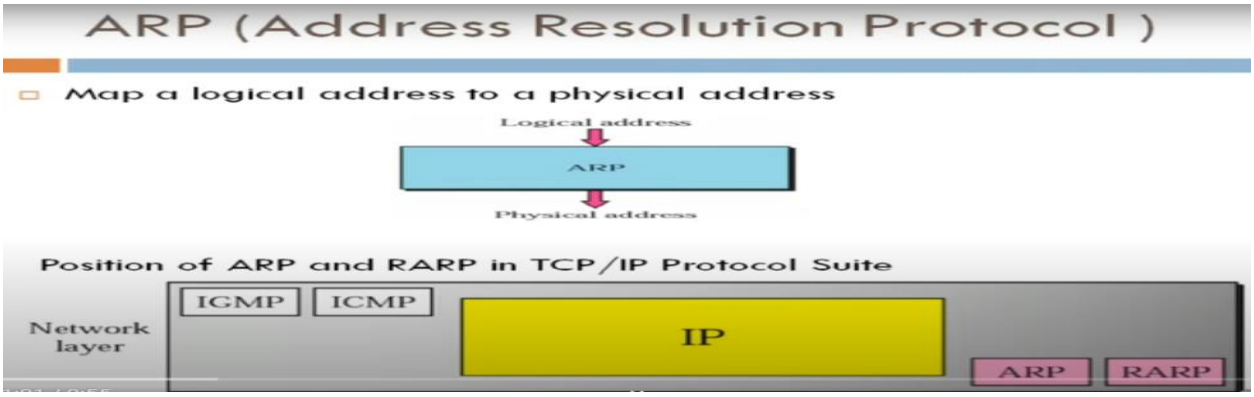


Fig: ARP request is broadcast

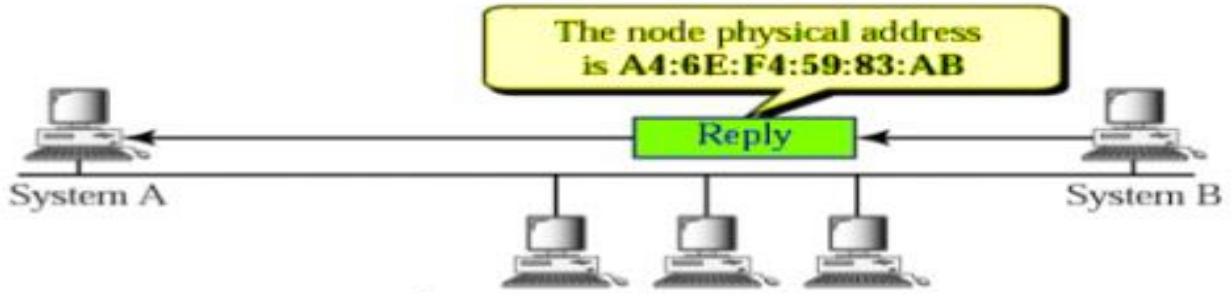
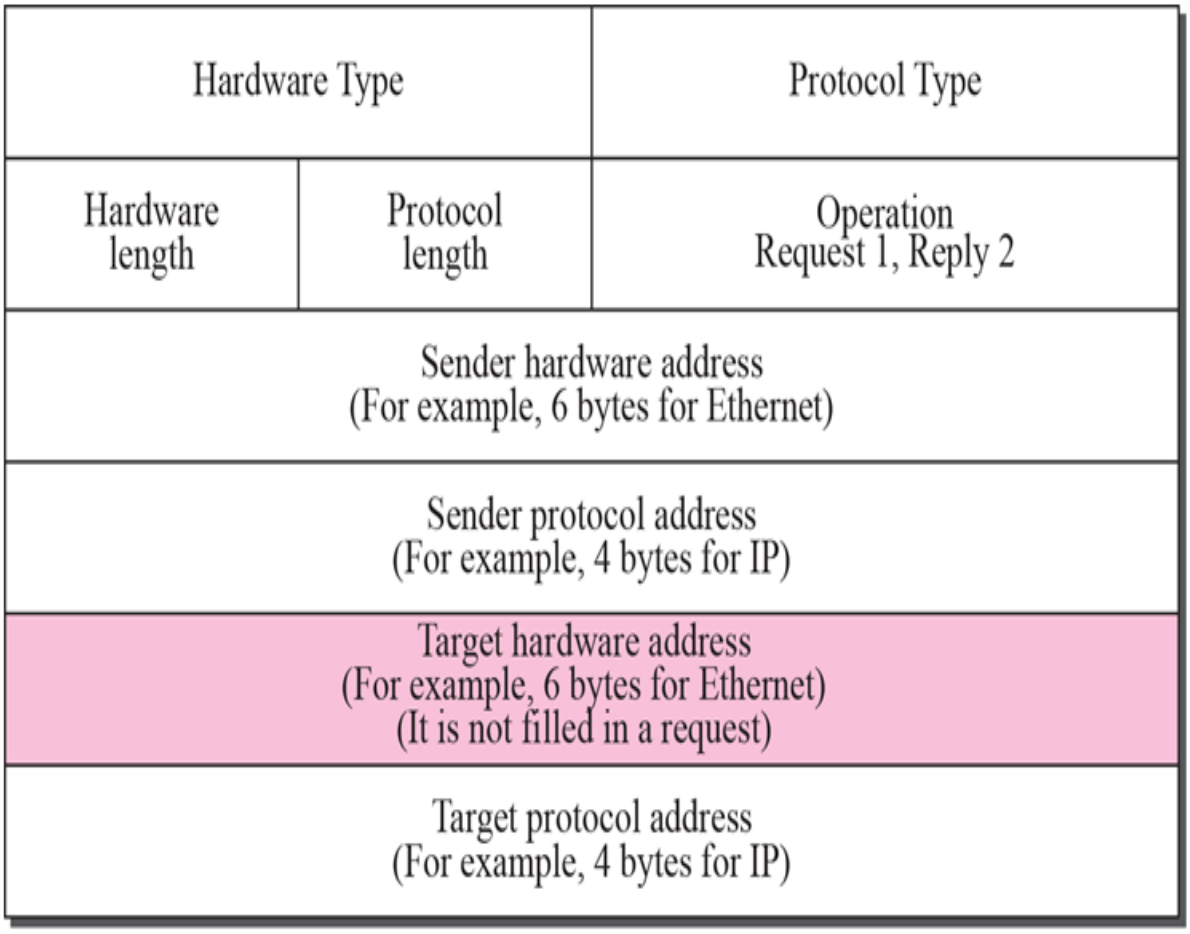


Fig: ARP reply is unicast

- **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 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(TYPE1)**

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(IPV4 080016)

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)

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:

Define the logical address of the target(IP 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.

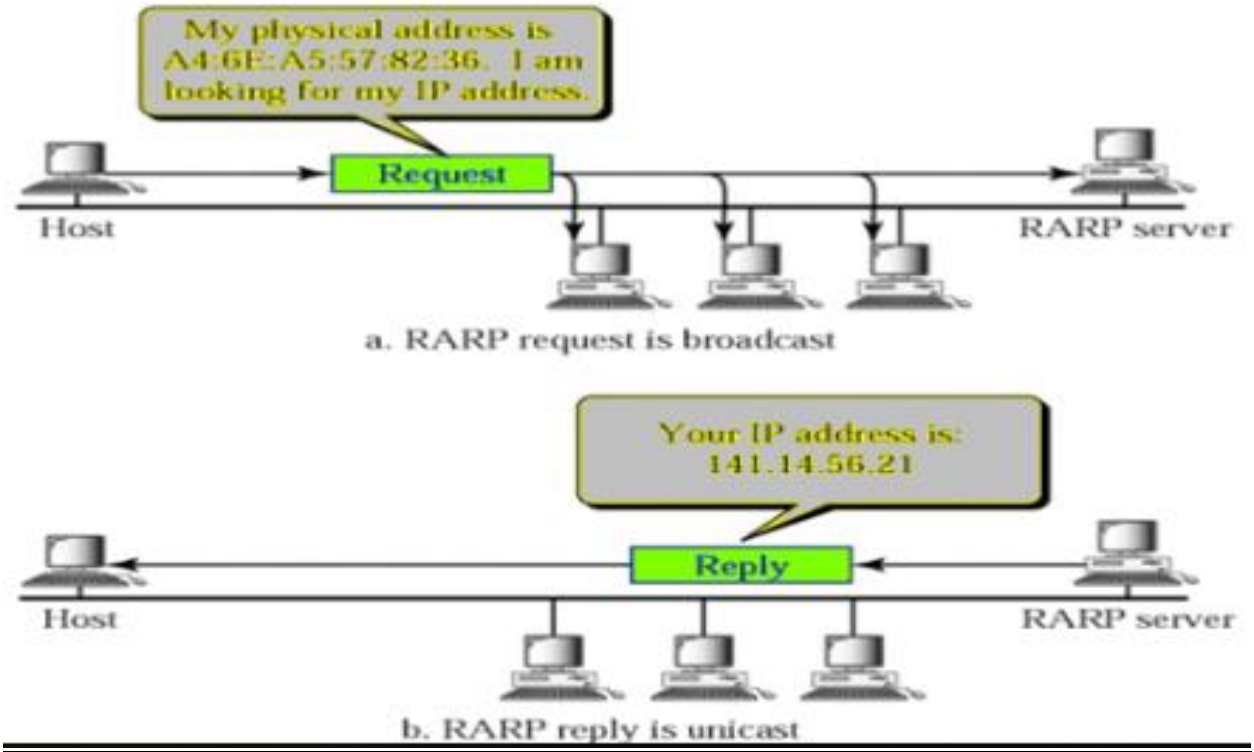
2) An **organization** doesn't have enough IP address to assign to **each station**

- It needs to assign IP address on demand.

The station can send its physical addresses and ask for a short time lease

- 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 operation



- 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

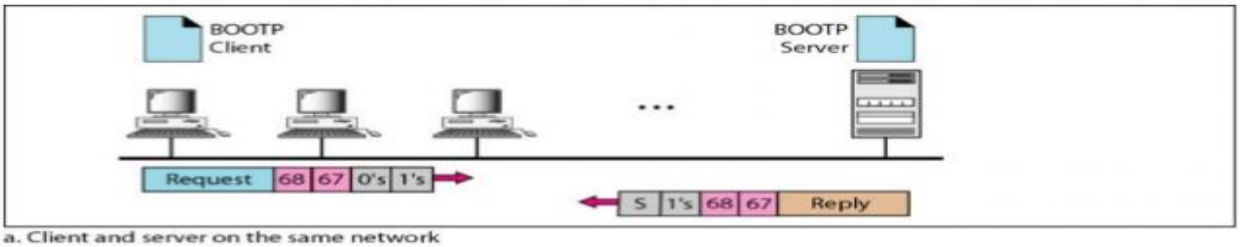
2) DHCP

BOOTP (Bootstrap Protocol)

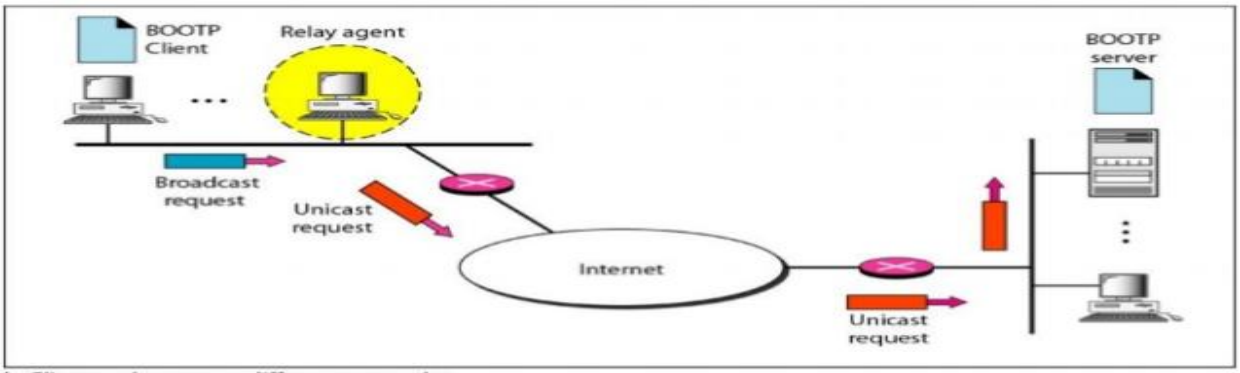
BOOTP is a **client/server protocol** designed to provide physical address to logical address mapping.

- 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.
- 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.



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

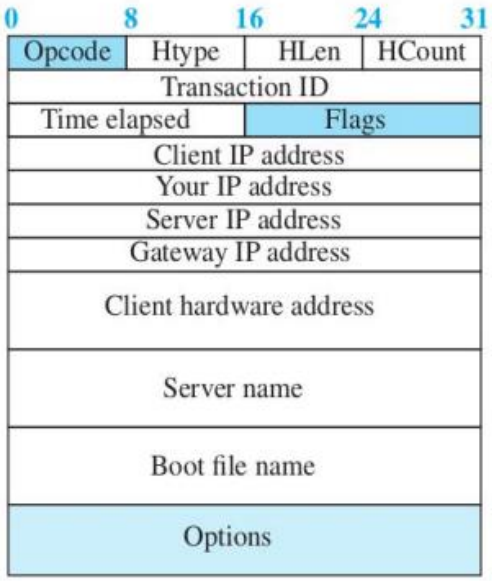
DYNAMIC HOST CONFIGURATION PROTOCOL

The Dynamic Host Configuration Protocol (DHCP) provides computers essential information when connecting to an IP network. This is necessary because a computer (for example, a mobile laptop) does not have an IP address to use.

The computer needing an IP address will first send a broadcast request for an IP address. A DHCP server will reply with the IP address, netmask, and gateway router information the computer should use. The address provided comes from a pool of available IP addresses, which is managed by the DHCP server. Therefore the DHCP provides a method of sharing IP addresses among a group of hosts that will change over time.

- An Application layer protocol that helps the network layer to assign IPaddress to hosts.
- Assign IP addresses dynamically/automatically to each host in thenetwork.
- Also provide other network parameters such as subnet mask, DNS serveraddress, gateway address etc.
- Uses client-server paradigm.

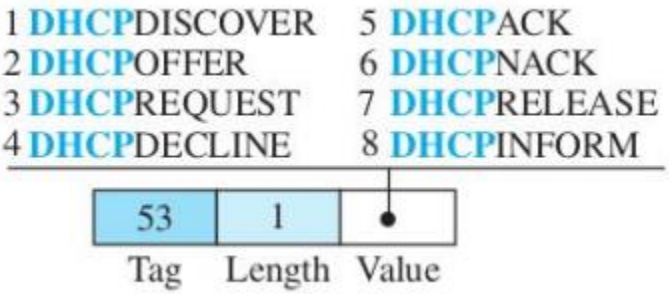
- Can assign IP parameters permanently / temporarily (for some time) tohosts.
- Uses a request-reply mechanism.



Fields:
Opcode: Operation code, request (1) or reply (2)
Htype: Hardware type (Ethernet, ...)
HLen: Length of hardware address
HCount: Maximum number of hops the packet can travel
Transaction ID: An integer set by the client and repeated by the server
Time elapsed: The number of seconds since the client started to boot
Flags: First bit defines unicast (0) or multicast (1); other 15 bits not used
Client IP address: Set to 0 if the client does not know it
Your IP address: The client IP address sent by the server
Server IP address: A broadcast IP address if client does not know it
Gateway IP address: The address of default router
Server name: A 64-byte domain name of the server
Boot file name: A 128-byte file name holding extra information
Options: A 64-byte field with dual purpose described in text

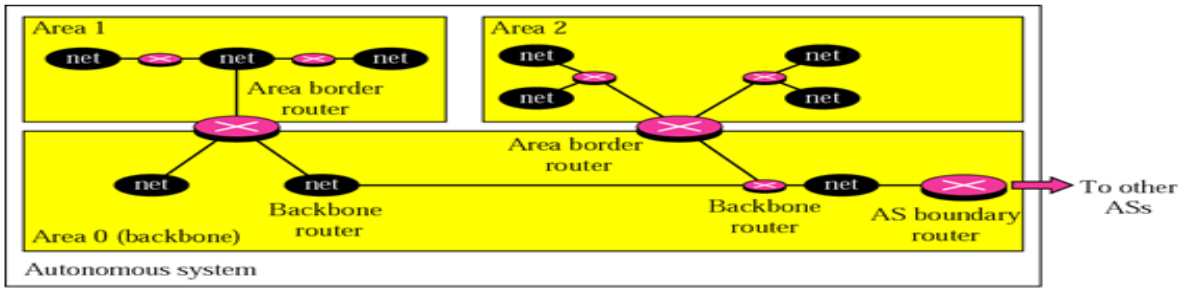
Message Format

- Options: 64 bytes
- Carry either additional information or some specific vendorinformation.
- If the option field is present, it will start with a magic cookie (a fourbyte number - 99.130.83.99), the remaining 60 bytes are options.
- An option is composed of three fields: a 1-byte tag field, a 1-bytelength field, and a variable-length value field.
- If the tag field is 53, the value field defines one of the 8 messagetypes.



OSPF (Open Shortest Path First) protocol

- The OSPF (Open Shortest Path First) protocol is one of a family of IP Routing protocols and is an Interior Gateway Protocol (IGP) for the Internet, used to distribute IP routing information throughout a single Autonomous System (AS) in an IP network.
- The Open Shortest Path First (OSPF) protocol is an intra-domain routing protocol based on link state routing.
- Divides an AS into areas
- Special routers (autonomous system boundary routers) or backbone routers responsible to dissipate(disappear) information about other AS into the current system.
- Metric based on type of service : Minimum delay , maximum throughput, reliability, etc.



Features of OSPF

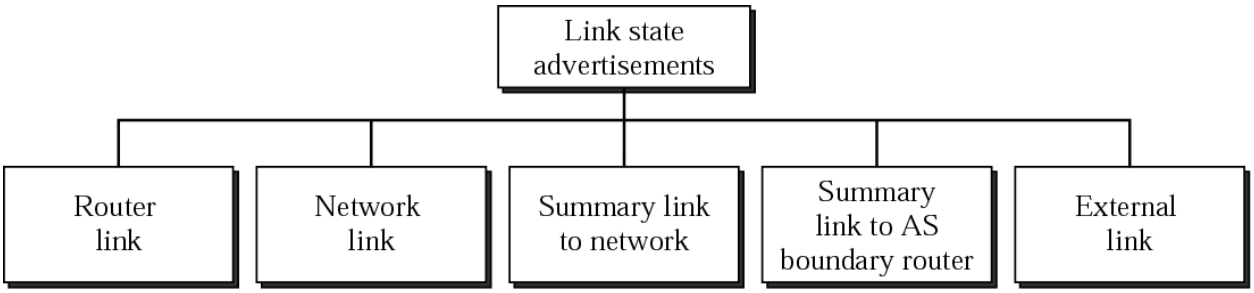
- Provides authentication of **routing messages**
- Enables load balancing by allowing **traffic** to be split evenly across routes with equal cost
- **Type-of-Service** routing allows to setup different routes
- Supports **subnetting** (is the practice of dividing a network into two or smaller networks)
- Supports **multicasting**
- Allows **hierarchical routing**

✓ A multi access network is one that can have multiple routers on it, each of which can directly communicate with all the others

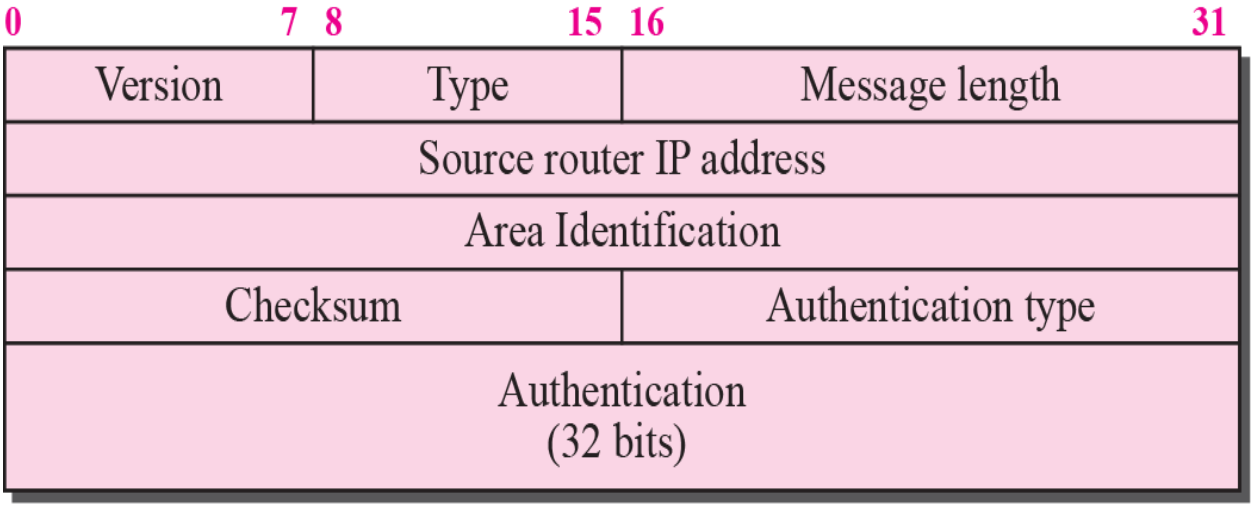
✓ OSPF operates by abstracting the collection of actual networks, routers, and lines into a **directed graph** in which each **arc** is assigned a cost (distance, delay, etc.).

✓ It then computes the shortest path based on the weights on the arcs.

✓ A **serial connection** between two routers is represented by a pair of arcs, one in each direction



OSPF Packet format :



The five types of OSPF messages

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

- **Version**: version of the OSPF
- **Type** : It specifies the type of the packets, there are 4 types of packets

  * hello packet

  * link state request

  * link state update

  * Link state acknowledgement

- **Source address** :It defines the address of the node to be send.
- **Area id** : It specifies ID of different areas
- **Check sum** :It deals with error detection and correction
- **Authentication type** :It has two value either zero or other numbers except zero
- **Authentication data**: used by authentication procedure

RIP	OSPF
It is a distance vector protocol	It is a link state protocol
The metrics used in RIP is hop count	The metrics used in OSPF are bandwidth and delay
RIP uses distance vector algorithm to calculate the best path	OSPF uses the SPF algorithm to calculate the best path
In RIP protocol, networks are not divided in areas or tables	In OSPF, routing is carried out in autonomous system, into areas, sub areas as well as backbone areas
Maximum hop count is 15	No hop count

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.

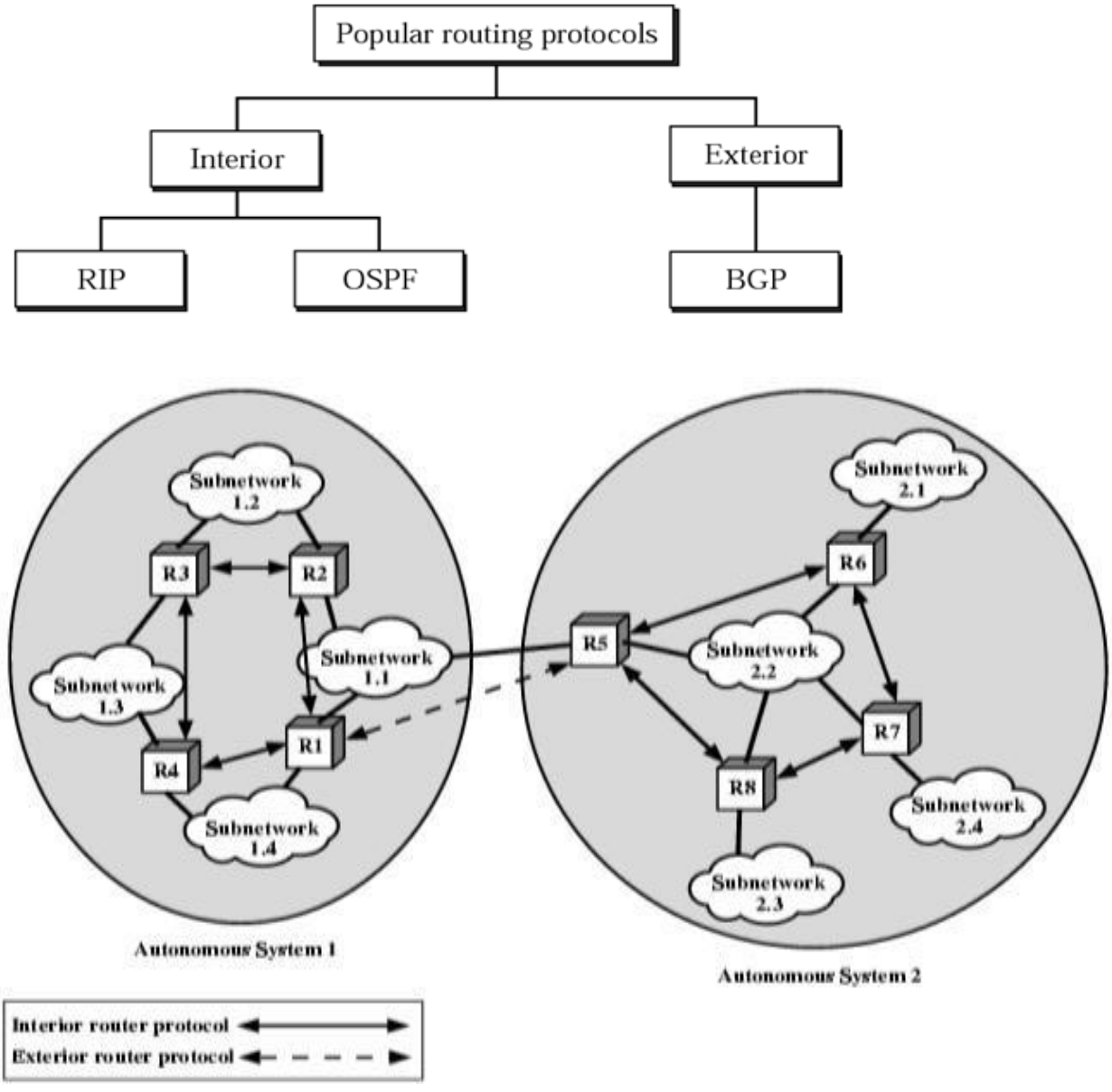
Allows a **subnet** to advertise its existence to rest of the Internet: "I am here"

- 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
- 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



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 **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.

Transit AS

- 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 coming from one AS and going to another AS to pass through.

A good example is a large corporation that is connected to more than one regional or national AS that does not allow transient traffic

- 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.
- 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.**