UNIT – 4 /Lecture-01/ Lecture-02

logical addressing

Logical Addressing[RGPV/Jun 2014]

A person's name usually does not change. A person's address on the other hand, relates to where they live and can change. On a host, the MAC address does not change; it is physically assigned to the host NIC and is known as the physical address. The physical address remains the same regardless of where the host is placed on the network.

The IP address is similar to the address of a person. It is known as a logical address because it is assigned logically based on where the host is located. The IP address, or network address, is assigned to each host by a network administrator based on the local network.

IP addresses contain two parts. One part identifies the local network. The network portion of the IP address will be the same for all hosts connected to the same local network. The second part of the IP address identifies the individual host. Within the same local network, the host portion of the IP address is unique to each host.

Both the physical MAC and logical IP addresses are required for a computer to communicate on a hierarchical network, just like both the name and address of a person are required to send a letter.

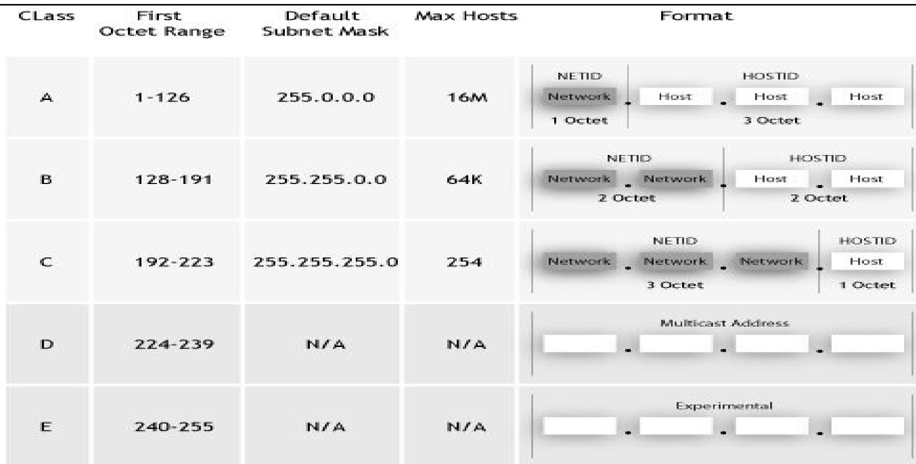
IPv4

Internet Protocol version 4 (**IPv4**) is the fourth version in the development of the Internet Protocol (IP) Internet, and routes most traffic on the Internet

IPv4 Address Classes

The IPv4 address space can be subdivided into 5 **classes** –

- Class A
- Class B
- Class C
- Class D
- Class E



CLass	First Octet Range	Default Subnet Mask	Max Hosts	Format
A	1-126	255.0.0.0	16M	NETID: Network (1 Octet) . HOSTID: Host . Host . Host (3 Octet)
B	128-191	255.255.0.0	64K	NETID: Network . Network (2 Octet) . HOSTID: Host . Host (2 Octet)
C	192-223	255.255.255.0	254	NETID: Network . Network . Network (3 Octet) . HOSTID: Host (1 Octet)
D	224-239	N/A	N/A	Multicast Address
E	240-255	N/A	N/A	Experimental

FIGURE : IP ADDRESSING

Each class consists of a contiguous subset of the overall IPv4 address range. With a few special exceptions explained further below, the values of the leftmost four bits of an IPv4 address determine its class as follows:

Class	**Leftmost bits**	**Start address**	**Finish address**
A	0xxx	0.0.0.0	127.255.255.255
B	10xx	128.0.0.0	191.255.255.255
C	110x	192.0.0.0	223.255.255.255
D	1110	224.0.0.0	239.255.255.255
E	1111	240.0.0.0	255.255.255.255

All Class C addresses, for example, have the leftmost three bits set to '110', but each of the remaining 29 bits may be set to either '0' or '1' independently (as represented by an x in these bit positions):

110xxxxx xxxxxxxx xxxxxxxx xxxxxxxx

Converting the above to dotted decimal notation, it follows that all Class C addresses fall in the range from 192.0.0.0 through 223.255.255.255.

IP Address Class E and Limited Broadcast

The IPv4 networking standard defines Class E addresses as reserved, meaning that they should not be used on IP networks. Some research organizations use Class E addresses for experimental purposes. However, nodes that try to use these addresses on the Internet will be unable to communicate properly.

A special type of IP address is the limited broadcast address 255.255.255.255. A broadcast involves delivering a message from one sender to many recipients. Senders direct an IP broadcast to 255.255.255.255 to indicate all other nodes on the local network (LAN) should pick up that message. This broadcast is 'limited' in that it does not reach every node on the Internet, only nodes on the LAN.

Technically, IP reserves the entire range of addresses from 255.0.0.0 through 255.255.255.255 for broadcast, and this range should not be considered part of the normal Class E range.

IP Address Class D and Multicast

The IPv4 networking standard defines Class D addresses as reserved for multicast. Multicast is a mechanism for defining groups of nodes and sending IP messages to that group rather than to every node on the LAN (broadcast) or just one other node (unicast).

Multicast is mainly used on research networks. As with Class E, Class D addresses should not be used by ordinary nodes on the Internet.

IP Address Class A, Class B, and Class C

Class A, Class B, and Class C are the three classes of addresses used on IP networks in common practice, with three exceptions as explained next.

IP Loopback Address

127.0.0.1 is the loopback address in IP. Loopback is a test mechanism of network adapters. Messages sent to 127.0.0.1 do not get delivered to the network. Instead, the adapter intercepts all loopback messages and returns them to the sending application. IP applications often use this feature to test the behavior of their network interface.

As with broadcast, IP officially reserves the entire range from 127.0.0.0 through 127.255.255.255 for loopback purposes. Nodes should not use this range on the Internet, and it should not be considered part of the normal Class A range.

Zero Addresses

As with the loopback range, the address range from 0.0.0.0 through 0.255.255.255 should not be considered part of the normal Class A range. 0.x.x.x addresses serve no particular function in IP, but nodes attempting to use them will be unable to communicate properly on the Internet.

Private Addresses

The IP standard defines specific address ranges within Class A, Class B, and Class C reserved for use by private networks (intranets). The table below lists these reserved ranges of the IP address space.

Class	Private start address	Private finish address
A	10.0.0.0	10.255.255.255
B	172.16.0.0	172.31.255.255
C	192.168.0.0	192.168.255.255

Nodes are effectively free to use addresses in the private ranges if they are not connected to the Internet, or if they reside behind firewalls or other gateways that use Network Address Translation (NAT).

Classful network[RGPV/Dec 2009]

A **classful network** is a network addressing architecture used in the Internet from 1981 until the introduction of Classless Inter-Domain Routing in 1993. The method divides the address space for Internet Protocol Version 4 (IPv4) into five address classes. Each class, coded in the first four bits of the address, defines either a different network size

---

Special-use addresses

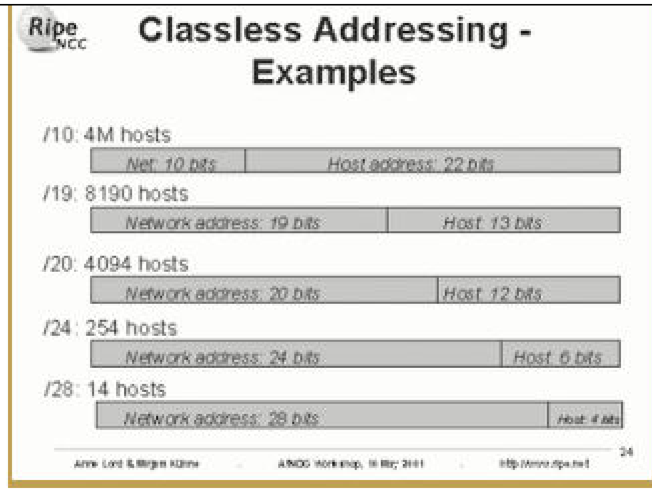
Reserved address blocks

Range	**Description**
0.0.0.0/8	Current network (only valid as source address)
10.0.0.0/8	Private network

100.64.0.0/10	Shared Address Space
127.0.0.0/8	Loopback
169.254.0.0/16	Link-local
172.16.0.0/12	Private network
192.0.0.0/24	IETF Protocol Assignments
192.0.2.0/24	TEST-NET-1, documentation and examples
192.88.99.0/24	IPv6 to IPv4 relay
192.168.0.0/16	Private network
198.18.0.0/15	Network benchmark tests
198.51.100.0/24	TEST-NET-2, documentation and examples
203.0.113.0/24	TEST-NET-3, documentation and examples
224.0.0.0/4	IP multicast (former Class D network)
240.0.0.0/4	Reserved (former Class E network)
255.255.255.255	Broadcast

Classless Addressing :[RGPV/Dec 2009]

Classless addressing uses a variable number of bits for the network and host portions of the address.



Classless addressing treats the IP address as a 32 bit stream of ones and zeroes, where the boundary between network and host portions can fall anywhere between bit 0 and bit 31.Classless addressing system is also known as CIDR (Classless Inter-Domain Routing).Classless addressing is a way to allocate and specify the Internet addresses used in inter-domain routing more flexibly than with the original system of Internet Protocol (IP) address classes. CIDR (Classless Internet Domain Routing) defines arbitrarily-sized subnets solely by base address and number of significant bits in the address. A CIDR address of 192.168.0.0/24 defines a block of addresses in the range 192.168.0.0 through 192.168.0.255, while 192.168.0.0/20 would define a network 16 times as large - from 192.168.0.0 through 192.168.15.255.

Decimal	192	160	20	48	
Binary	11000000	10100000	00010100	0011	0000
	<-------- 28 bits Network ------->				4 bits host

What is the difference between classless and classful IP address?

Your default class addresses are Class A 0-127, Class B - 128-191, Class C - 192-223 for the 1st octet values .Classful IP addresses are IP addresses that follow this standard subnet ranges for class A, B, C so a classful router protocol like ripv1 will always assume that the address 172.16.1.2 has a subnet mask of 255.255.0.0 even if you want it to have a subnet of 255.255.255.0 so on a classful router protocol 172.16.1.2 will always have the range 172.16.0.0 - 172.16.255.255 (because the value 172 in the 1st octet falls in the Class B range of 128-191 and class B addresses have the subnet mask set to 255.255.0.0)

Classless IP addresses mean that the address range is determined by the subnet mask and hence the same address 172.16.1.2 255.255.255.0 will now be looked at as having its range as 172.16.1.0 - 255 because 255.255.255.0 corresponds to that range.

Public IP Address and Private IP Address

Public Host

Any computer accessing a public network like internet must have a unique ip address.Such a host is termed as public host. IP address of the public host is termed as public IP address.

Private Host

The total IP addresses available are very limited. So it is not possible to assign unique ip address to all computers in the world. Here comes the importance of private IP address. The following ranges of IP addresses are reserved for private IP.

- 10.0.0.0 to 10.255.255.255
- 172.16.0.0 to 172.31.255.255
- 192.168.0.0 to 192.168.255.255

Inside a LAN or a private network computers can use these ip addresses. Two different private network may use the same set of private IP addresses. These private hosts can access internet or a public network through a public host.That is the private host's identity will not be visible in the public network, but it shows only the IP address of the public host through which it is connected to the internet. That is Private host shows the IP address of the public host in the Internet environment,but in the local network it shows the private IP address.

A private network is typically a network that uses private IP address space.Private IP addresses were originally created due to the shortage of publicly registered IP addresses created by the IPv4 standard.

Merging two Private Networks

Internal networks of two different organizations may use the same private IP addresses.Problem occurs when trying to merge two such networks. We may apply the following solutions.

- One network must renumber
- A NAT router must be placed between the networks.

CIDR - Classless Inter-Domain Routing

IPv4 TCP/IP Subnet Table

While subnetting might be easy enough to grasp as a concept, it can be a bit involved, and even mind-boggling in part due to the required manipulations of binary numbers. Many people understand the ideas behind subnetting, but find it hard to follow the actual steps required to subnet a network. The table below is intended as a quick reference and a fairly complete example of IPv4 subnetting.

Introduction to Subnet Masks

Subnet masks are one of the most interesting aspects of TCP/IP. Subnet masks point out to IP which bits of the 32-bit IP address refer to the network. A good network administrator understands how to determine and use subnet masks.

What Is a Subnet Mask?

A subnet mask is a number that looks like an IP address. It shows TCP/IP how many bits are used for the network portion of the IP address by covering up, or "masking," the IP addresses network portion. As you learned in Chapter 6, an IP address is made up of two parts: the network portion and the host portion. For every outgoing packet, IP has to determine whether the destination host is on the same local network or on a remote network. If the destination is local, then IP uses an ARP broadcast to find out the hardware address of the destination host. If the destination host is not on the local network, then ARP broadcasts request for the hardware address of the router. Therefore, IP sends packets that are bound for a remote network directly to the router, which is also known as the default gateway. The router then sends the packet to the next network on its journey to the correct destination network. Just as the telephone system uses an area code to determine whether a number is local or long distance, TCP/IP uses the subnet mask to determine whether the destination of a packet is a host on the local network or a host on a remote network. In the same way that every U.S. telephone number must have an area code, every IP address must have a subnet mask. If, for example, your telephone number is (619) 555-1212, and you call someone whose telephone number is (619) 345-1111, it is a local call. You know that because you can look at the numbers between the parentheses and see that they have the same value. If, on the other hand, your number is (619) 555-1212, and you call someone whose number is (213) 888-8146, it's a long distance call. You know that because the numbers inside of the parentheses are different. You can think of the subnet mask as the area code in the parentheses of a telephone number. Just as an area code determines a phone call's destination, a subnet mask tells IP how many bits to look at when determining if the destination IP address is local or remote. The following graphic shows Harry calling Amber. Since Amber has a different area code, the phone call will have to go through the router. When Harry calls Sally, however, it is a local call and does not need to go through the router. When determining if the packet is bound for the local network or a remote network, IP compares the network portion of the sender's IP address with the same number of bits from the destination's IP address. If the bit values are exactly the same, the packet's destination is determined to be local. If there are any differences in the bit values, the packet's destination is determined to be remote.

To know how many bits to compare, IP evaluates the subnet mask of the sending host. In the subnet mask, there is a series of 1s, and then the rest of the bits are set to 0. When IP evaluates the subnet mask, it is looking specifically for the answer to the question, "How many bits are set to 1?" Once IP determines how many bits are set to 1, it knows how many bits of the source host's IP address and the destination host's IP address will be compared.You can think of the number of bits that are set to 1 in the subnet mask as the number of digits inside the parentheses in a telephone number—if that number could change (in other words, if it's variable). If, for example, a telephone number has 10 digits, imagine if the parentheses include 4, 5, or 6 digits.

You would then evaluate the number to be local or long distance based on the digits that are in the arentheses. If there are 8 bits set to 1 in the subnet mask, IP will compare the first 8 bits of the host with the first 8 bits of the destination. If there are 16 bits in the subnet mask that are set to 1, IP will compare the first 16 bits of host and destination. A subnet mask is a required element of every IP address. When you want to type in the IP address for a host, the only two required elements are the IP address itself and the subnet mask. Likewise, when you want to call someone, it is required that you know the correct area code for the phone number.

You then compare the first three characters of your phone number (your area code) with the first three characters of their phone number (their area code). If the area codes are the same, you don't need to dial the area code, nor do you have to pay for a long distance call, because it is a local call. If the area code is not the same, however, you'll have to dial their area code so that the telephone system can route your call to their city. You'll see over the next several pages that IP looks at everything in binary. Subnet masks and routing will become clearer if you think about the IP addresses and subnet masks in binary, so begin now to think of IP addresses and subnet masks as 32 bits. When thinking in binary, do not pay attention to the periods that we use in the decimal representation.

Subnet Mask

For each class of address, there is a standard, or default, subnet mask. Each is discussed in the following sections.

Class A Addresses

The standard subnet mask for a Class A address is 255.0.0.0. This tells IP that the first 8 bits are used for the network portion of the IP address, and the remaining 24 bits are used for the host portion. IP looks at the 32 bits and uses the subnet mask to mask out the network portion of the address: NNNN NNNN.HHHH HHHH.HHHH HHHH.HHHH HHHH

Because 24 bits are left for the host portion of the address, there are almost 17 million unique host IP addresses for each Class A network address.

Class B Address

A Class B address has a standard subnet mask of 255.255.0.0. This mask tells IP that the first 16 bits are used for the network portion of the address, and the remaining 16 bits are used for the host portion: NNNN NNNN.NNNN NNNN.HHHH HHHH.HHHH HHHH The 16 bits that are used for the host portion of the address can uniquely address more than 16,000 hosts on each Class B network.

Default Mask

When a router receives a packet with a destination address, it needs to route the packet. A router outside the organization route the packet based on the network address and a router inside the organization route the packet based on the subnetwork address.

- Network Address
- Sub Network Address
- Default Mask
- Subnet Mask

The router outside the organization has a routing table with one coulmn based on the network address. The router inside the organization has a routing table based on the subnetwork address.

IP Default Subnet Masks For Address Classes A, B and C

Subnetting is the process of dividing a Class A, B or C network into subnets, as we've seen in the preceding topics. In order to better understand how this "division of the whole" is accomplished, it's worth starting with a look at how the "whole" class A, B and C networks are represented in a subnetted environment. This is also of value because there are situations where you may need to define an unsubnetted network using subnetting notation.

Just as is always the case, the subnet mask for a default, unsubnetted class A, B or C network has ones for each bit that is used for network ID or subnet ID, and zeroes for the host ID bits. Of course, we just said we aren't subnetting, so there **are** no subnet ID bits! Thus, the subnet mask for this default case has 1s for the network ID portion and 0s for the host ID portion. This is called the default subnet mask for each of the IP address classes.

Since classes A, B and C divide the network ID from the host ID on octet boundaries, the subnet mask will always have all ones or all zeroes in an octet. Therefore, the default subnet masks will always have 255s or 0s when expressed in decimal notation

So, the three default subnet masks are 255.0.0.0 for Class A, 255.255.0.0 for class B, and 255.255.255.0 for Class C. Note that while all default subnet masks use only "255" and "0", not all subnet masks with "255" and "0" are defaults. There are a small number of custom subnets that divide on octet boundaries as well. These are:

- **255.255.0.0:**,This is the default mask for Class B, but can also be the custom subnet mask for dividing a Class A network using 8 bits for the subnet ID (leaving 16 bits for the host ID).
- **255.255.255.0:** This is the default subnet mask for Class C, but can be a custom Class A with 16 bits for the subnet ID or a Class B with 8 bits for the subnet ID.

S.NO	RGPV QUESTIONS	Year	Marks
Q.1	What do you mean by classless and classful addressing? How many maximum network and host ids are there in class A, B and C networks?	Dec2009	7
Q.2	For the given data IP address 172.16.0.0 Subnet mask 255.255.248.0 Find the- (i) No of subnets (ii) No of host (iii) Subnet ip (iv) Range of ip address.	Jun 2010	7
Q.3	A host in an organization has an IP address 150.37.64.34 and a subnet mask 255.255.240.0. What is the address of this subnet? What is the range of IP address that a host can have on this subnet?	Dec 2010	7

Unit-04/Lecture-03

IPV4

IPV4 HEADER FORMAT[RGPV/Jun 2010, Jun 2011, Jun 2013,Dec 2012, Dec 2013]

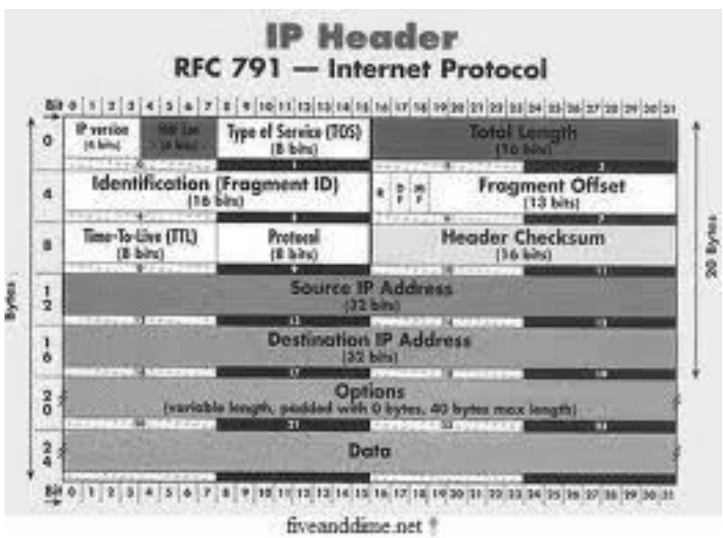


FIGURE IP HEADER

Version It defines the version of the IPv4 protocol. Currently the version is 4. However, version 6 (or IPng) may totally replace version 4 in the future. This field tells the IPv4 software running in the processing machine that the datagram has the format of version 4. All fields must be interpreted as specified in the fourth version of the protocol. If the machine is using some other version of IPv4, the datagram is discarded rather than interpreted incorrectly.

Header length (HLEN) This 4-bit field defines the total length of the datagram header in 4-byte words. This field is needed because the length of the header is variable (between 20 and 60 bytes). When there are no options, the header length is 20 bytes, and the value of this field is 5 (5 x 4 = 20). When the option field is at its maximum size, the value of this field is 15 (15 x 4 = 60).

Services 8-bit field, this field, previously called service type, is now called differentiated services.

Service type or differentiated services

- D: Minimize delay
- R: Maximize reliability
- T: Maximize throughput
- C: Minimize cost

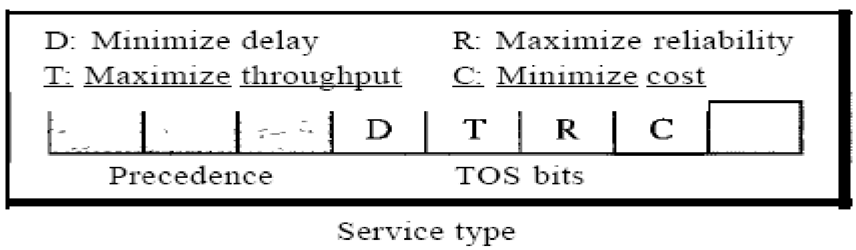


FIGURE : IP HEADER SERVICE TYPE

In this interpretation, the first 3 bits are called precedence bits. The next 4 bits are called type of service (TOS) bits, and the last bit is not used. a. Precedence is a 3-bit subfield ranging from 0 (000 in binary) to 7 (111 in binary). The precedence defines the priority of the datagram in issues such as congestion.

If a router is congested and needs to discard some datagrams, those datagrams with lowest precedence are discarded first. Some datagrams in the Internet are more important than others. For example, a datagram used for network management is much more urgent and important than a datagram containing optional information for a group.

The precedence subfield was part of version 4, but never used.

TOS bits is a 4-bit subfield with each bit having a special meaning. Although a bit can be either 0 or 1, one and only one of the bits can have the value of 1 in each datagram.

Types of service

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

FIGURE : IP HEADER TOS BITS DESCRIPTION

Total length This is a In-bit field that defines the total length (header plus data) of the IPv4 datagram in bytes. To find the length of the data coming from the upper layer, subtract the header length from the total length. The header length can be found by multiplying the value in the HLEN field by 4.

Length of data =total length - header length

Since the field length is 16 bits, the total length of the IPv4 datagram is limited to 65,535 (216 - 1) bytes, of which 20 to 60 bytes are the header and the rest is data from the upper layer. The total length field defines the total length of the datagram including the header. Though a size of 65,535 bytes might seem large, the size of the IPv4 datagram may increase in the near future as the underlying technologies allow even more throughput (greater bandwidth).

Fragmentation The datagram must be fragmented to be able to pass through those networks.

Identification This field is used in fragmentation

Flags This field is used in fragmentation

Fragmentation offset This field is used in fragmentation

Time to live A datagram has a limited lifetime in its travel through an internet. This field was originally designed to hold a timestamp, which was decremented by each visited router. The datagram was discarded when the value became zero. However, for this scheme, all the machines must have synchronized clocks and must know how long it takes for a datagram to go from one machine to another. Today, this field is used mostly 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. This value is approximately 2 times the maximum number of routes between any two hosts. Each router that processes the datagram decrements this number by 1. If this value, after being decremented, is zero, the router discards the datagram.

This field is needed because routing tables in the Internet can become corrupted. A datagram may travel between two or more routers for a long time without ever getting delivered to the destination host. This field limits the lifetime of a datagram. Another use of this field is to intentionally limit the journey of the packet. For example, if the source wants to confine the packet to the local network, it can store 1 in this field. When the packet arrives at the first router, this value is decremented to 0, and the datagram is discarded.

Protocol. This 8-bit field defines the higher-level protocol that uses the services of the IPv4 layer. An IPv4 datagram can encapsulate data from several higher-level protocols such as TCP, UDP, ICMP, and IGMP. This field specifies the final destination protocol to which the IPv4 datagram is delivered. In other words, since the IPv4 protocol carries data from different other protocols, the value of this field helps the receiving network layer know to which protocol the data belong

Protocol values

Value	*Protocol*
1	ICMP
2	IGMP
6	TCP
17	UDP
89	OSPF

FIGURE : IP HEADER PROTOCOL

Source address This 32-bit field defines the IPv4 address of the source. This field must remain unchanged during the time the IPv4 datagram travels from the source host to the destination host.

o Destination address. This 32-bit field defines the IPv4 address of the destination. This field must remain unchanged during the time the IPv4 datagram travels from the source host to the destination host.

Fragmentation The fields that are related to fragmentation and reassembly of an IPv4 datagram are the identification, flags, and fragmentation offset fields.

Identification This 16-bit field identifies a datagram originating from the source host. The combination of the identification and source IPv4 address must uniquely define a datagram as it leaves the source host. To guarantee uniqueness, the IPv4 protocol uses a counter to label the datagrams. The counter is initialized to a positive number. When the IPv4 protocol sends a datagram, it copies the current value of the counter to the identification field and increments the counter by 1. As long as the counter is kept in the main memory, uniqueness is guaranteed. When a datagram is fragmented, the value in the identification field is copied to all fragments. In other words, all fragments have the same identification number, the same as the original datagram. The identification number helps the destination in reassembling the datagram. It knows that all fragments having the same identification value must be assembled into one datagram.

Flags This is a 3-bit field. The first bit is reserved. The second bit is called the do not fragment bit. If its value is 1, the machine must not fragment the datagram. If it cannot pass the datagram through any available physical network, it discards the datagram and sends an ICMP error message to the source host. If its value is 0, the datagram can be fragmented if necessary. The third bit is called the more fragment bit. If its value is 1, it means the datagram is not the last fragment; there are more fragments after this one. If its value is 0, it means this is the last or only fragment

Flags used in fragmentation

- D: Don’t fragment
- M: More fragments

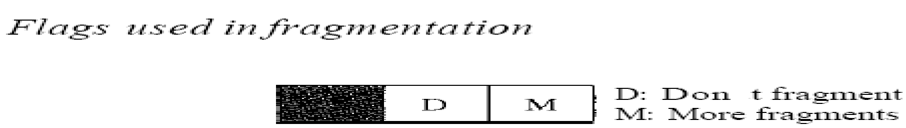


FIGURE : IP HEADER FLAGS

Fragmentation offset This 13-bit field shows the relative position of this fragment with respect to the whole datagram. It is the offset of the data in the original datagram measured in units of 8 bytes. Figure 20.11 shows a datagram with a data size of 4000 bytes fragmented into three fragments.

The bytes in the original datagram are numbered 0 to 3999. The first fragment carries bytes 0 to

1399. The offset for this datagram is 0/8 =O. The second fragment carries bytes 1400 to 2799; the offset value for this fragment is 1400/8 = 175. Finally, the third fragment carries bytes 2800 to 3999. The offset value for this fragment is 2800/8 =350.

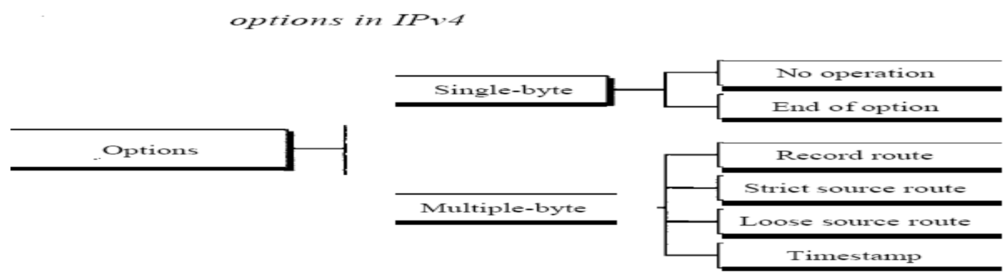


FIGURE : IP HEADER OPTIONS

S.NO	RGPV QUESTIONS	Year	Marks
Q.1	Write short note on IPv4.	Jun 2011	7
Q.2	Explain the frame format of IPv4.	Jun 2010	7
Q.3	Draw the frame format of IPv4	Dec 2013	7
Q.4	Discuss in detail the various aspects of IPV4?	Jun 2013	7
Q.5	Explain the function of 3 flags in the IPv4 header	Dec 2012	5
Q.6	How is the IPV4 header checksum calculated?	Dec 2012	5

Unit-04/Lecture-04

IPV6

Introduction to IPv6[RGPV/Dec 2009, Jun 2013]

There are legitimate reasons for designing and developing the new Internet Protocol IPv6:

- The recent exponential growth of the Internet and the impending exhaustion of the IPv4 address space. IPv4 addresses have become relatively scarce, forcing some organizations to use a network address translator (NAT) to map multiple private addresses to a single public IP address. While NATs promote reuse of the private address space, they do not support standards-based network layer security or the correct mapping of all higher layer protocols and can create problems when connecting two organizations that use the private address space. Additionally, the rising prominence of Internet-connected devices and appliances assures that the public IPv4 address space will eventually be depleted.
- The growth of the Internet and the ability of Internet backbone routers to maintain large routing tables. Because of the way in which IPv4 network IDs have been and are currently allocated, there are routinely over 70,000 routes in the routing tables of

Internet backbone routers. The current IPv4 Internet routing infrastructure is a combination of both flat and hierarchical routing.

- The need for simpler configuration. Most current IPv4 implementations must be configured either manually or through a statefull address configuration protocol such as Dynamic Host Configuration Protocol (DHCP). With more computers and devices using IP, there is a need for a simpler and more automatic configuration of addresses and other configuration settings that do not rely on the administration of a DHCP infrastructure.
- The requirement for security at the IP level. Private communication over a public medium like the Internet requires encryption services that protect the data sent from being viewed or modified in transit. Although a standard now exists for providing security for IPv4 packets (known as Internet Protocol security or IPSec), this standard is optional and proprietary solutions are prevalent.
- The need for better support for real-time delivery of data (also known as quality of service). While standards for quality of service (QoS) exist for IPv4, real-time traffic support relies on the IPv4 Type of Service (TOS) field and the identification of the payload, typically using a UDP or TCP port. Unfortunately, the IPv4 TOS field has limited functionality and has different interpretations.

IPv6 features

The following are the features of the IPv6 protocol:

- new header format;
- large address space;
- efficient and hierarchical addressing and routing infrastructure;
- stateless and state full address configuration;
- built-in security;
- better support for quality of service (QoS);
- new protocol for neighboring node interaction;
- Extensibility.

The IPv6 header has a new format that is designed to minimize header overhead. This is achieved by moving both nonessential fields and option fields to extension headers that are placed after the IPv6 header. The streamlined IPv6 header provides more efficient processing at intermediate routers.

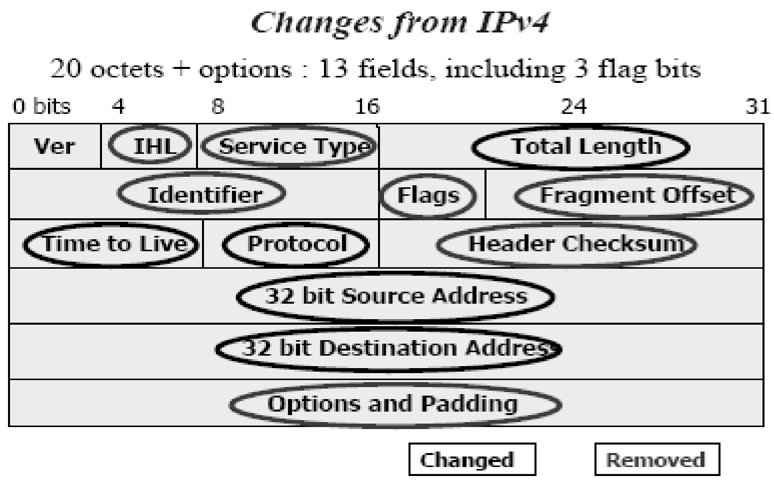


FIGURE: IPV4 HEADER

FIGURE: IPV6 HEADER

IPv6 has 128-bit (16-byte) source and destination addresses. Although 128 bits can provide over 3.4×10^{38} possible combinations, the large address space of IPv6 has been designed to allow for multiple levels of subnetting and address allocation from the Internet backbone to the individual subnets within an organization.

IPv6 global addresses used on the IPv6 portion of the Internet are designed to create an efficient and hierarchical routing infrastructure that addresses the common occurrence of multiple levels of Internet service providers. On the IPv6 Internet, backbone routers have much smaller routing tables.

IPv6 can be extended for new features by adding extension headers after the IPv6 header. Unlike the IPv4 header, which can only support 40 bytes of options, the size of IPv6 extension headers is only constrained by the size of the IPv6 packet.

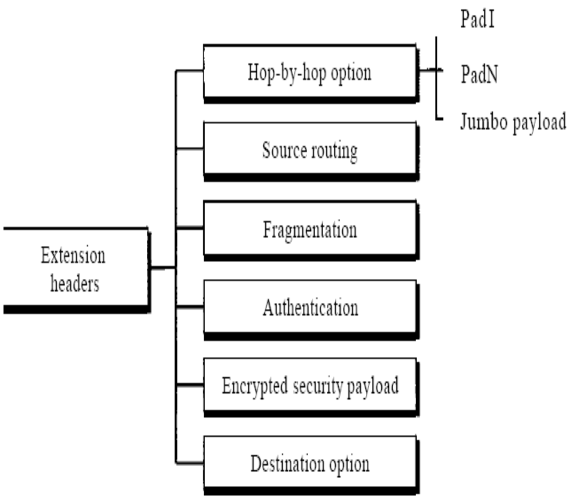
Security features in IPv6

The IPv6 protocol incorporates Internet Protocol security (IPSec), which provides protection of IPv6 data as it is sent over the network. IPSec is a set of Internet standards that uses cryptographic security services to provide the following:

- Confidentiality: IPSec traffic is encrypted. Captured IPSec traffic cannot be deciphered without the encryption key.
- Authentication: IPSec traffic is digitally signed with the shared encryption key so that the receiver can verify that the IPSec peer sent it.
- Data integrity: IPSec traffic contains a cryptographic checksum that incorporates the encryption key. The receiver can verify that the packet was not modified in transit.

The IPv6 protocol for Windows XP also provides support for anonymous addresses. Anonymous addresses provide a level of anonymity when accessing Internet resources.

Extension header types



Next header codes for IPv6

Code	*Next Header*
0	Hop-by-hop option
2	ICMP
6	TCP
17	UDP
43	Source routing
44	Fragmentation
50	Encrypted security payload
51	Authentication
59	Null (no next header)
60	Destination option

IPv4	**IPv6**
Addresses are 32 bits (4 bytes) in length.	Addresses are 128 bits (16 bytes) in length
Address (A) resource records in DNS to map host names to IPv4 addresses.	Address (AAAA) resource records in DNS to map host names to IPv6 addresses.
Pointer (PTR) resource records in the IN-ADDR.ARPA DNS domain to map IPv4 addresses to host names.	Pointer (PTR) resource records in the IP6.ARPA DNS domain to map IPv6 addresses to host names.
IPSec is optional and should be supported externally	IPSec support is not optional
Header does not identify packet flow for QoS handling by routers	Header contains Flow Label field, which Identifies packet flow for QoS handling by router.
Both routers and the sending host fragment packets.	Routers do not support packet fragmentation. Sending host fragments packets

Header includes options.	Optional data is supported as extension headers.
ARP uses broadcast ARP request to resolve IP to MAC/Hardware address.	Multicast Neighbor Solicitation messages resolve IP addresses to MAC addresses.
Internet Group Management Protocol (IGMP) manages membership in local subnet groups.	Multicast Listener Discovery (MLD) messages manage membership in local subnet groups.
Broadcast addresses are used to send traffic to all nodes on a subnet.	IPv6 uses a link-local scope all-nodes multicast address.
Configured either manually or through DHCP.	Does not require manual configuration or DHCP.
Must support a 576-byte packet size (possibly fragmented).	Must support a 1280-byte packet size (without fragmentation).

S.NO	RGPV QUESTIONS	Year	Marks
Q.1	Differentiate IPv4 & IPv6.	Dec 2009	7
Q.2	Define the type of the following destination addresses: i. 4A:30:10:21:10:1A ii. 47:20:1B:2E:08:EE iii. FF:FF:FF:FF:FF:FF	Jun 2013	7

Unit-04/Lecture-05

CONGESTION CONTROL

CONGESTION

An important issue in a packet-switched network is **congestion.** Congestion in a network may occur if the **load** on the network-the number of packets sent to the network-is greater than the *capacity* of the network-the number of packets a network can handle.

Congestion control refers to the mechanisms and techniques to control the congestion and keep the load below the capacity.

Congestion in a network or internetwork occurs because routers and switches have queues-buffers that hold the packets before and after processing. A router, for example, has an input queue and an output queue for each interface.

Delay Versus Load

Note that when the load is much less than the capacity of the network, the delay is at a minimum. This minimum delay is composed of propagation delay and processing delay, both of which are negligible. However, when the load reaches the network capacity, the delay increases sharply because we now need to add the waiting time in the queues (for all routers in the path) to the total delay. Note that the delay becomes infinite when the load is greater than the capacity. If this is not obvious, consider the size of the queues when almost no packet reaches the destination, or reaches the destination with infinite delay; the queues become longer and longer. Delay has a negative effect on the load and consequently the congestion. When a packet is delayed, the source, not receiving the acknowledgment, retransmits the packet, which makes the delay, and the congestion, worse.

Throughput Versus Load

We defined throughput as the number of bits passing through a point in a
second. We can extend that definition from bits to packets and from a point to a network. We can define throughput in a network as the number of packets passing through the network in a unit of time. Notice that when the load is below the capacity of the network, the throughput increases proportionally with the *load.* We expect the throughput to remain constant after the load reaches the capacity, but instead the throughput declines sharply. The reason is the discarding of packets by the routers. When the load exceeds the capacity, the queues become full and the routers have to discard some packets. Discarding packet does not reduce the number of packets in the network because the sources retransmit the packets, using time-out mechanisms, when the packets do not reach the destinations.

CONGESTION CONTROL[RGPV/Dec 2003, Dec 2010,Jun 2013, Dec 2013]

Congestion control refers to techniques and mechanisms that can either prevent congestion, before it happens, or remove congestion, after it has happened. In general, we can divide congestion control mechanisms into two broad categories:

- **open-loop congestion control (prevention)**
- **closed-loop congestion control (removal)**

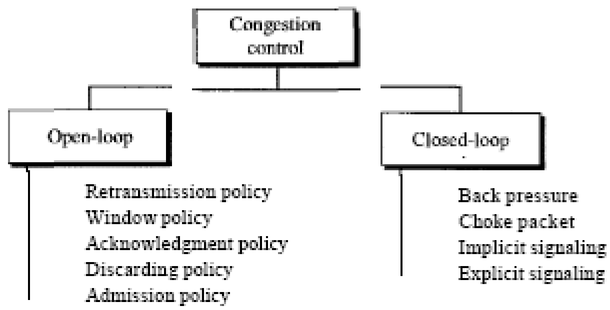


FIGURE : CONGESTION CONTROL

Open-Loop Congestion Control

In open-loop congestion control, policies are applied to prevent congestion before it happens. In these mechanisms, congestion control is handled by either the source or the destination. We give a brief list of policies that can prevent congestion.

Retransmission Policy

Retransmission is sometimes unavoidable. If the sender feels that a sent packet is lost or corrupted, the packet needs to be retransmitted. Retransmission in general may increase congestion in the network. However, a good retransmission policy can prevent congestion. The retransmission policy and the retransmission timers must be designed to optimize efficiency and at the same time prevent congestion. For example, the retransmission policy used by TCP (explained later) is designed to prevent or alleviate congestion.

Window Policy

The type of window at the sender may also affect congestion. The Selective Repeat window is better than the Go-Back-N window for congestion control. In the *Go-Back-N* window, when the timer for a packet times out, several packets may be resent, although some may have arrived safe and sound at the receiver. This duplication may make the congestion worse. The Selective Repeat window, on the other hand, tries to send the specific packets that have been lost or corrupted.

Acknowledgment Policy

The acknowledgment policy imposed by the receiver may also affect congestion. If the receiver does not acknowledge every packet it receives, it may slow down the sender and help prevent congestion. Several approaches are used in this case. A receiver may send an acknowledgment only if it has a packet to be sent or a special timer expires. A receiver may decide to acknowledge only *N* packets at a time. We need to know that the acknowledgments are also part of the load in a network. Sending fewer acknowledgments means imposing less load on the network.

Discarding Policy

A good discarding policy by the routers may prevent congestion and at the same time may not harm the integrity of the transmission. For example, in audio transmission, if the policy is to discard less sensitive packets when congestion is likely to happen, the quality of sound is still preserved and congestion is prevented or alleviated.

Admission Policy

An admission policy, which is a quality-of-service mechanism, can also prevent congestion in virtual-circuit networks. Switches in a flow first check the resource requirement of a flow before admitting it to the network. A router can deny establishing a virtual circuit connection if there is congestion in the network or if there is a possibility of future congestion.

Closed-Loop Congestion Control

Closed-loop congestion control mechanisms try to alleviate congestion after it happens. Several mechanisms have been used by different protocols. We describe a few of them here.

Backpressure

The technique of *backpressure* refers to a congestion control mechanism in which a congested node stops receiving data from the immediate upstream node or nodes. This may cause the upstream node or nodes to become congested, and they, in turn, reject data from their upstream nodes or nodes. And so on. Backpressure is a node-to-node congestion control that starts with a node and propagates, in the opposite direction of data flow, to the source. The backpressure technique can be applied only to virtual circuit

networks, in which each node knows the upstream node from which a flow of data is corning.

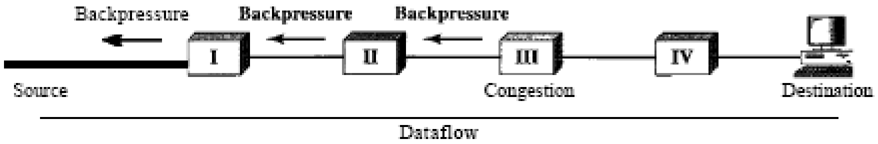


FIGURE : BACKPRESSURE

Node III in the figure has more input data than it can handle. It drops some packets in its input buffer and informs node II to slow down. Node II, in turn, may be congested because it is slowing down the output flow of data. If node II is congested, it informs node I to slow down, which in turn may create congestion. If so, node I inform the source of data to slow down. This, in time, alleviates the congestion. Note that the *pressure* on node III is moved backward to the source to remove the congestion. None of the virtual-circuit networks we studied in this book use backpressure. It was, however, implemented in the first virtual-circuit network, X.25. The technique cannot be implemented in a datagram network because in this type of network, a node (router) does not have the slightest knowledge of the upstream router.

Choke Packet[RGPV/Jun 2003,Dec 2010, Dec 2006, Jun 2013]

A choke packet is a packet sent by a node to the source to inform it of congestion. Note the difference between the backpressure and choke packet methods. In backpressure, the warning is from one node to its upstream node, although the warning may eventually reach the source station. In the choke packet method, the warning is from the router, which has encountered congestion, to the source station directly. The intermediate nodes through which the packet has traveled are not warned. When a router in the Internet is overwhelmed with IP datagram's, it may discard some of them; but it informs the source host, using a source quench ICMP message. The warning message goes directly to the source station; the intermediate routers, and does not take any action. Figure shows the idea of a choke packet.

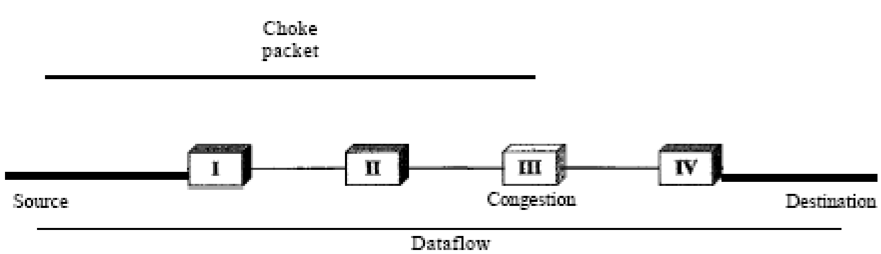


FIGURE 4.1: CHOKE PACKET

Implicit Signaling

In implicit signaling, there is no communication between the congested node or nodes and the source. The source guesses that there is a congestion somewhere in the network from other symptoms. For example, when a source sends several packets and there is no acknowledgment for a while, one assumption is that the network is congested. The delay in receiving an acknowledgment is interpreted as congestion in the network; the source should slow down.

Explicit Signaling

The node that experiences congestion can explicitly send a signal to the source or destination. The explicit signaling method, however, is different from the choke packet method. In the choke packet method, a separate packet is used for this purpose; in the explicit signaling method, the signal is included in the packets that carry data. Explicit signaling can occur in either the forward or the backward direction.

Backward Signaling

A bit can be set in a packet moving in the direction opposite to the congestion. This bit can warn the source that there is congestion and that it needs to slow down to avoid the discarding of packets.

Forward Signaling

A bit can be set in a packet moving in the direction of the congestion. This bit can warn the destination that there is congestion. The receiver in this case can use policies, such as slowing down the acknowledgments, to alleviate the congestion.

Congestion Control in TCP[RGPV/Dec 2012/ Jun 2014]

How TCP uses congestion control to avoid congestion or alleviate congestion in the network.

Congestion Window

The sender window size is determined by the available buffer space in the receiver *(rwnd).* In other words, we assumed that it is only the receiver that can dictate to the sender the size of the sender's window. We totally ignored

another entity here-the network. If the network cannot deliver the data as fast as they are created by the sender, it must tell the sender to slow down. In other words, in addition to the receiver, the network is a second entity that determines the size of the sender's window. Today, the sender's window size is determined not only by the receiver but also by congestion in the network. The sender has two pieces of information: the receiver-advertised window size and the congestion window size. The actual size of the window is the minimum of these two.

Congestion Policy

TCP's general policy for handling congestion is based on three phases:

- slow start
- congestion avoidance
- congestion detection

In the slow-start phase, the sender starts with a very slow rate of ransmission, but increases the rate rapidly to reach a threshold. When the threshold is reached, the data rate is reduced to avoid congestion. Finally if congestion is detected, the sender goes back to the slow-start or congestion avoidance phase based on how the congestion is detected.

Slow Start: Exponential Increase

One of the algorithms used in TCP congestion control is called slow start. This algorithm is based on the idea that the size of the congestion window *(cwnd)* starts with one maximum segment size (MSS). The MSS is determined during connection establishment by using an option of the same name. The size of the window increases one MSS each time an acknowledgment is received. As the name implies, the window starts slowly, but grows exponentially. We have used segment numbers instead of byte numbers (as though each segment contains only 1 byte). We have assumed that *rwnd* is much higher than *cwnd,* so that the sender window size always equals *cwnd.* We have assumed that each segment is acknowledged individually. The sender starts with *cwnd* =1 MSS. This means that the sender can send only one segment. After receipt of the acknowledgment for segment 1, the size of the congestion window is increased by 1, which means that *cwnd* is now 2. Now two more segments can be sent. When each acknowledgment is received, the size of the window is increased by 1 MSS. When all seven segments are acknowledged, *cwnd* = 8.

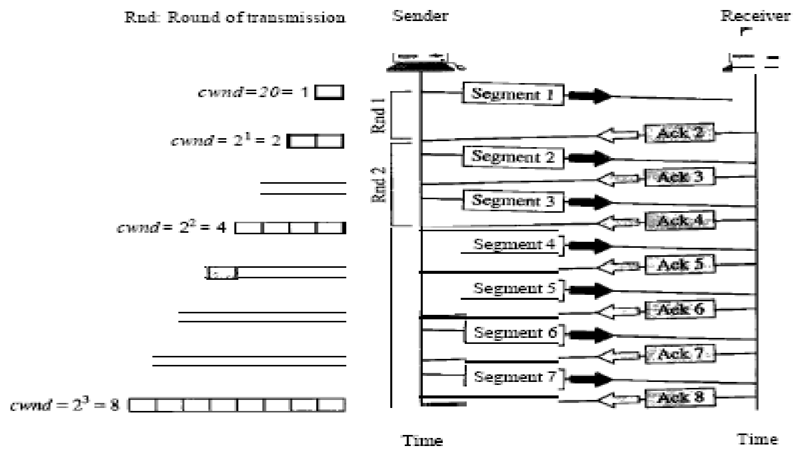


FIGURE : CONGESTION CONTROL,SLOW START

Congestion Avoidance: Additive Increase

If we start with the slow-start algorithm, the size of the congestion window increases exponentially. To avoid congestion before it happens, one must slow down this exponential growth. TCP defines another algorithm called congestion avoidance, which undergoes an additive increase instead of an exponential one. When the size of the congestion window reaches the slow-start threshold, the slow-start phase stops and the additive phase begins. In this algorithm, each time the whole window of segments is acknowledged (one round), the size of the congestion window is increased by 1. To show the idea, we apply this algorithm to the same scenario as slow start, although we will see that the congestion avoidance algorithm usually starts when the size of the window is much greater than 1.

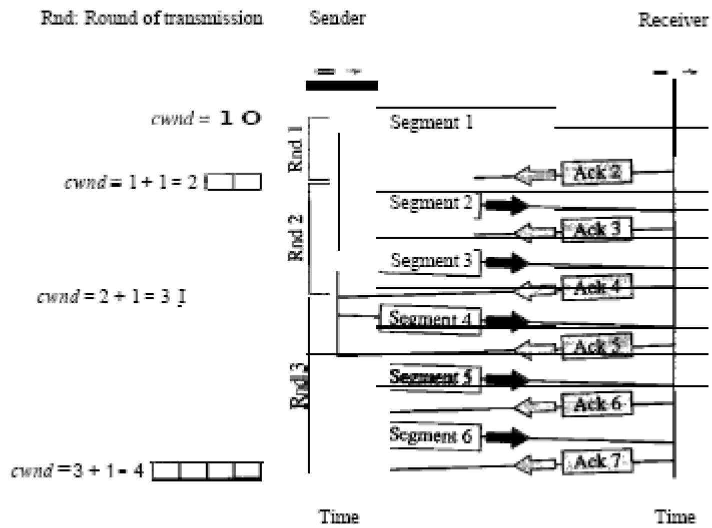


FIGURE : CONGESTION CONTROL ADDITIVE INCREASE

Congestion Detection: Multiplicative Decrease

If congestion occurs, the congestion window size must be decreased. The only way the sender can guess that congestion has occurred is by the need to retransmit a segment. However, retransmission can occur in one of two cases: when a timer times out or when three ACKs are received. In both

cases, the size of the threshold is dropped to one-half, a multiplicative decrease. Most TCP implementations have two reactions:

I. If a time-out occurs, there is a stronger possibility of congestion; a segment has probably been dropped in the network, and there is no news about the sent segments.

In this case TCP reacts strongly:

a. It sets the value of the threshold to one-half of the current window size.

b. It sets *cwnd* to the size of one segment.

c. It starts the slow-start phase again.

2. If three ACKs are received, there is a weaker possibility of congestion; a segment may have been dropped, but some segments after that may have arrived safely since three ACKs are received. This is called fast transmission and fast recovery. In this case, TCP has a weaker reaction:

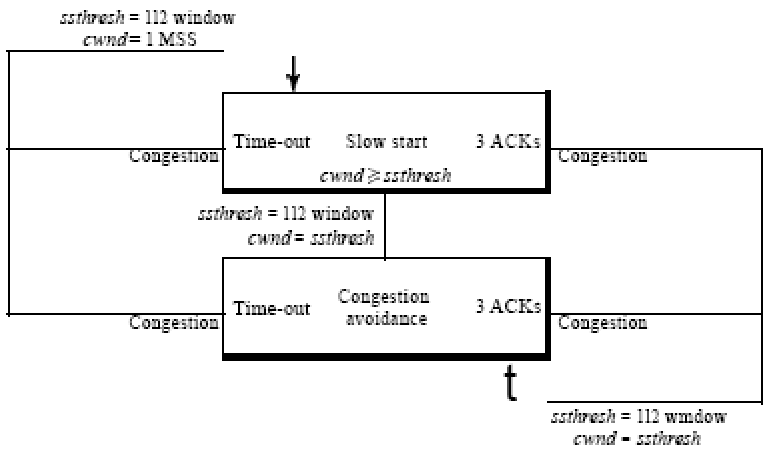
a. It sets the value of the threshold to one-half of the current window size.

b. It sets *cwnd* to the value of the threshold (some implementations add three segment sizes to the threshold).

c. It starts the congestion avoidance phase.

An implementations reacts to congestion detection in one of the following ways:

- If detection is by time-out, a new *slow-start* phase starts.
- If detection is by three ACKs, a new *congestion avoidance* phase starts.

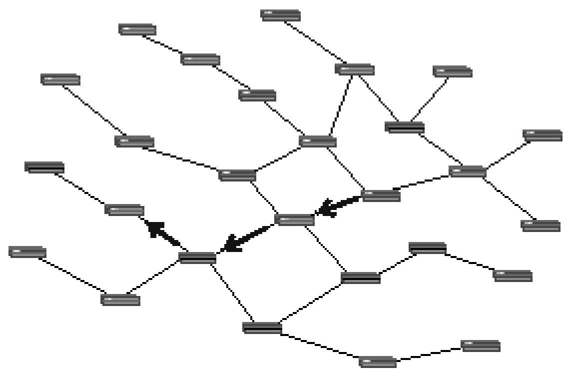


S.NO	RGPV QUESTIONS	Year	Marks
Q.1	Explain various congestion control techniques.	Dec 2003	7
Q.2	Explain congestion control by choke packet method.	Jun 2003	7
Q.3	Explain leaky bucket algorithms in congestion control.	Dec 2003	7
Q.4	What are the policies and algorithm used for prevention of congestion? What are the policies that affect congestion?	Dec 2002	7
Q.5	Give an argument why the leaky bucket algorithm should allow just one packet per tick, independent of how large the packet is?	Dec 2010	7
Q.6	Explain congestion control in virtual circuit.	Dec 2013	7
Q.7	What is congestion control and how it is implemented in network layer?What is the role of choke packet in managing congestion?	Jun 2013	7
Q.8	How can TCP used to deal with network or internet congestion?	Dec 2012	7
Q.9	Enumerate various measures taken by TCP to avoid congestion on networks.	Jun 2014	7

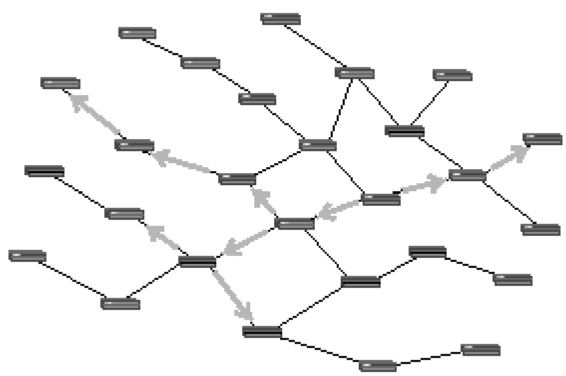
Unit-04/Lecture-06/ Lecture-07/ Lecture-08/ Lecture-09/ Lecture-10/ Lecture-11
ROUTING PROTOCOLS

Packet forwarding

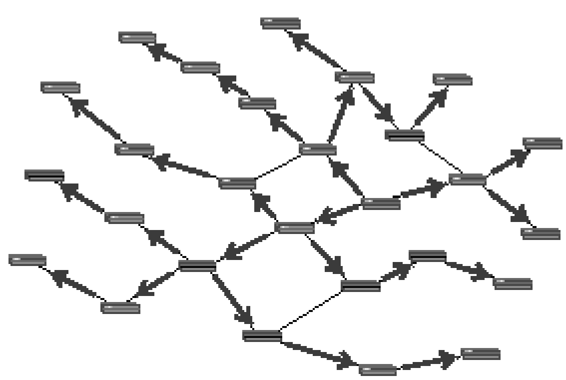
Packet forwarding is the relaying of packets from one network segment to another by nodes in a computer network.



A unicast forwarding pattern, typical of many networking technologies including the overwhelming majority of Internet traffic



A multicast forwarding pattern, typical of PIM



A broadcast forwarding pattern, typical of bridged Ethernet

The Network Layer of the OSI Layer is responsible for Packet Forwarding. The simplest forwarding model — unicasting — involves a packet being relayed from link to link along a chain leading from the packet's source to its destination. However, other forwarding strategies are commonly used. Broadcasting requires a packet to be duplicated and copies sent on multiple links with the goal of delivering a copy to every device on the network. In practice, broadcast packets are not forwarded everywhere on a network, but only to devices within a broadcast domain, making *broadcast* a relative term. Less common than broadcasting, but perhaps of greater utility and theoretical significance, is multicasting, where a packet is selectively duplicated and copies delivered to each of a set of recipients.

Direct Versus Indirect Delivery

The delivery of a packet to its final destination is accomplished by using two different methods of delivery

- direct
- indirect

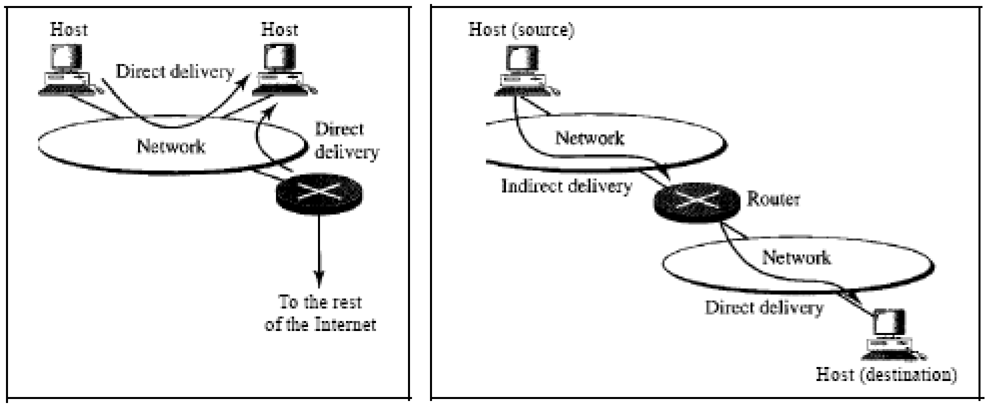
Direct Delivery

In a direct delivery, the final destination of the packet is a host connected to the same physical network as the deliverer. Direct delivery occurs when the source and destination of the packet are located on the same physical network or when the delivery is between the last router and the destination host.

The sender can easily determine if the delivery is direct. It can extract the network address of the destination (using the mask) and compare this address with the addresses of the networks to which it is connected. If a match is found, the delivery is direct.

Indirect Delivery

If the destination host is not on the same network as the deliverer, the packet is delivered indirectly. In an indirect delivery, the packet goes from router to router until it reaches the one connected to the same physical network as its final destination. Note that a delivery always involves one direct delivery but zero or more indirect deliveries.



a. Direct delivery b. Indirect and direct delivery

FIGURE : DIRECT DELIVERY

FORWARDING

Forwarding means to place the packet in its route to its destination. Forwarding requires a host or a router to have a routing table. When a host has a packet to send or when a router has received a packet to be forwarded, it looks at this table to find the route to the final destination. However, this simple solution is impossible today in an internetwork such as the Internet because the number of entries needed in the routing table would make table lookups inefficient.

Forwarding Techniques

Several techniques can make the size of the routing table manageable and also handle issues such as security. We briefly discuss these methods here.

Next-Hop Method Versus Route Method

One technique to reduce the contents of a routing table is called the next-hop method. In this technique, the routing table holds only the address of the next hop instead of information about the complete route (route method). The entries of a routing table must be consistent with one another.

Network-Specific Method Versus Host-Specific Method

A second technique to reduce the routing table and simplify the searching process is called the network-specific method. Here, instead of having an entry for every destination host connected to the same physical network (host-specific method), we have only one entry that defines the address of the destination network itself. In other words, we treat all hosts connected to the same network as one single entity. For example, if 1000 hosts are attached to the same network, only one entry exists in the routing table instead of 1000.

Host-specific routing is used for purposes such as checking the route or providing security measures.

Default Method

Another technique to simplify routing is called the default method. Host A is connected to a network with two routers. Router Rl routes the packets to hosts connected to network N2. However, for the rest of the Internet, router R2 is used. So instead of listing all networks in the entire Internet, host A can just have one entry called the *default* (normally defined as network address 0.0.0.0).

Routing protocol [RGPV/ Dec 2012, Dec 2013, Jun 2014]

A **routing protocol** specifies how routers communicate with each other, disseminating information that enables them to select routes between any two nodes on a computer network. Routing algorithms determine the specific choice of route. Each router has a priori knowledge only of networks attached to it directly. A routing protocol shares this information first among immediate neighbors, and then throughout the network. This way, routers gain knowledge of the topology of the network.

Routing table

In computer networking a **routing table**, or **routing information base (RIB)**, is a data table stored in a router or a networked computer that lists the routes to particular network destinations, and in some cases, metrics (distances) associated with those routes. The routing table contains information about the topology of the network immediately around it. The construction of routing tables is the primary goal of routing protocols. Static routes are entries made in a routing table by non-automatic means and which are fixed rather than being the result of some network topology "discovery" procedure.

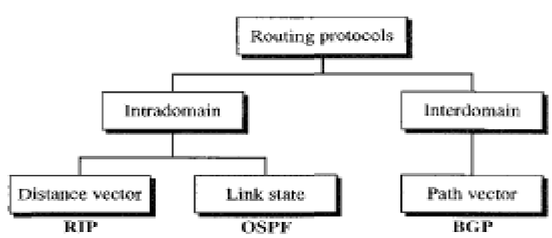


FIGURE : ROUTING PROTOCOL

Contents of routing tables

The routing table consists of at least three information fields:

1. the network id: i.e. the destination subnet
2. cost/metric: i.e. the cost or metric of the path through which the packet is to be sent
3. next hop: The next hop, or gateway, is the address of the next station to which the packet is to be sent on the way to its final destination

Depending on the application and implementation, it can also contain additional values that refine path selection:

1. quality of service associated with the route. For example, the U flag indicates that an IP route is up.
2. links to filtering criteria/access lists associated with the route
3. interface: such as eth0 for the first Ethernet card, eth1 for the second Ethernet card, etc.

Routing tables are also a key aspect of certain security operations, such as unicast reverse path forwarding (uRPF) In this technique, which has several variants, the router also looks up, in the routing table, the source address of the packet. If there exists no route back to the source address, the packet is assumed to be malformed or involved in a network attack, and is dropped.

Mask	Network address	Next-hop address	Interlace		Reference count	Use

FIGURE : ROUTING TABLE

Static Routing Table

A **static routing table** contains information entered manually. The administrator enters the route for each destination into the table. When a table is created, it cannot update automatically when there is a change in the Internet. The table must be manually altered by the administrator.

A static routing table can be used in a small internet that does not change very often, or in an experimental internet for troubleshooting. It is poor strategy to use a static routing table in a big internet such as the Internet.

Dynamic Routing Table

A dynamic routing table is updated periodically by using one of the dynamic routing protocols such as RIP, OSPF, or BGP. Whenever there is a change in the Internet, such as a shutdown of a router or breaking of a link, the dynamic routing protocols update all the tables in the routers (and eventually in the host) automatically. The routers in a big internet such as the Internet need to be updated dynamically for efficient delivery of the IP packets.

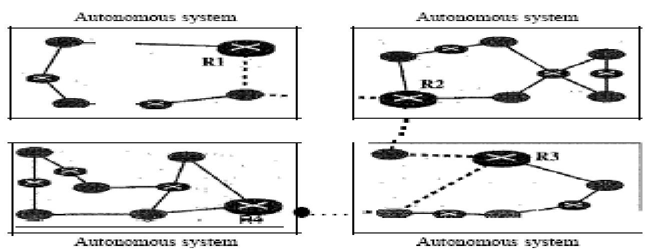


FIGURE : AUTONOMOUS SYSTEM

Autonomous System

Within the Internet, an **autonomous system** (**AS**) is a collection of connected Internet Protocol (IP) routing prefixes under the control of one or more network operators that presents a common, clearly defined routing policy to the Internet.

Originally the definition required control by a single entity, typically an Internet service provider or a very large organization with independent connections to multiple networks, that adhere to a single and clearly defined routing policy, as originally defined in RFC 1771. The newer definition in RFC 1930 came into use because multiple organizations can run BGP using private AS numbers to an ISP that connects all those organizations to the Internet. Even though there may be multiple autonomous systems supported by the ISP, the Internet only sees the routing policy of the ISP. That ISP must have an officially registered **autonomous system number** (**ASN**).

Intra- and Interdomain Routing

Today, an internet can be so large that one routing protocol cannot handle the task of updating the routing tables of all routers. For this reason, an internet is divided into autonomous systems. An autonomous system (AS) is a group of networks and routers under the authority of a single administration. Routing inside an autonomous system is referred to as intradomain routing. Routing between autonomous systems is referred to as interdomain routing. Each autonomous system can choose one or more intradomain routing protocols to handle routing inside the autonomous system. However, only one interdomain routing protocol handles routing between autonomous systems.

Distance-vector routing protocol [RGPV/Dec 2010]

In computer communication theory relating to packet-switched networks, a **distance-vector routing protocol** is one of the two major classes of routing protocols, the other major class being the link-state protocol. Distance-vector routing protocols use the Bellman–Ford algorithm to calculate paths.

A distance-vector routing protocol requires that a router informs its neighbors of topology changes periodically. Compared to link-state protocols, which require a router to inform all the nodes in a network of topology changes, distance-vector routing protocols have less computational complexity and message overhead.

The term distance vector refers to the fact that the protocol manipulates vectors (arrays) of distances to other nodes in the network. The vector distance algorithm was the original ARPANET routing algorithm and was also used in the internet under the name of RIP (routing internet protocol).Examples of distance-vector routing protocols include RIPv1 and RIPv2 and IGRP.

Method

Routers using distance-vector protocol do not have knowledge of the entire path to a destination. Instead they use two methods:

1. Direction in which router or exit interface a packet should be forwarded.
2. Distance from its destination

Distance-vector protocols are based on calculating the direction and distance to any link in a network. "Direction" usually means the next hop address and the exit interface. "Distance" is a measure of the cost to reach a certain node. The least cost route between any two nodes is the route with minimum distance. Each node maintains a vector (table) of minimum distance to every node. The cost of reaching a destination is calculated using various route metrics. RIP uses the hop count of the destination whereas IGRP takes into account other information such as node delay and available bandwidth.

Updates are performed periodically in a distance-vector protocol where all or part of a router's routing table is sent to all its neighbors that are configured to use the same distance-vector routing protocol. RIP supports cross-platform distance vector routing whereas IGRP is a Cisco Systems proprietary distance vector routing protocol. Once a router has this information it is able to amend its own routing table to reflect the changes and then inform its neighbors of the changes. This process has been described as ‘routing by rumor’ because routers are relying on the information they receive from other routers and cannot determine if the information is actually valid and true. There are a number of features which can be used to help with instability and inaccurate routing information.

EGP and BGP are not pure distance-vector routing protocols because a distance-vector protocol calculates routes based only on link costs whereas in BGP, for example, the local route preference value takes priority over the link cost.

Count-to-infinity problem

The Bellman–Ford algorithm does not prevent routing loops from happening and suffers from the **count-to-infinity problem**. The core of the count-to-infinity problem is that if A tells B that it has a path somewhere, there is no way for B to know if the path has B as a part of it. To see the problem clearly, imagine a subnet connected like A–B–C–D–E–F, and let the metric between the routers be "number of jumps". Now suppose that A is taken offline. In the vector-update-process B notices that the route to A, which was distance 1, is down – B does not receive the vector update from A. The problem is, B also gets an update from C, and C is still not aware of the fact that A is down – so it tells B that A is only two jumps from C (C to B to A), which is false. This slowly propagates through the network until it reaches infinity (in which case the algorithm corrects itself, due to the relaxation property of Bellman–Ford).

Workarounds and solutions

RIP uses the split horizon with poison reverse technique to reduce the chance of forming loops and uses a maximum number of hops to counter the 'count-to-infinity' problem. These measures avoid the formation of routing loops in some, but not all, cases. The addition of a *hold time* (refusing route updates for a few minutes after a route retraction) avoids loop formation in virtually all cases, but causes a significant increase in convergence times.

More recently, a number of loop-free distance vector protocols have been developed notable examples are EIGRP, DSDV and Babel. These avoid loop formation in all cases, but suffer from increased complexity, and their deployment has been slowed down by the success of link-state routing protocols such as OSPF.

BELLMAN FORD ALGORITHM [RGPV/Jun 2006 / Dec 2008/Jun 2011/ Dec 2012/ Dec 2013]

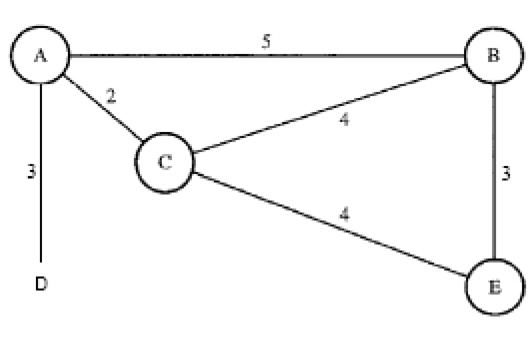
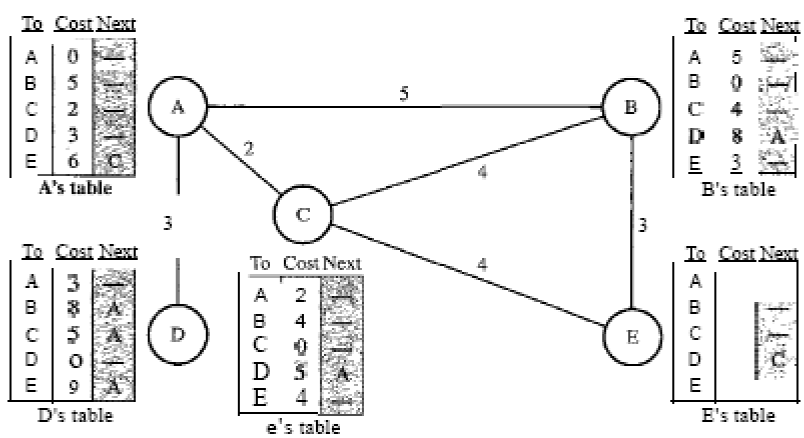


FIGURE : NETWORK



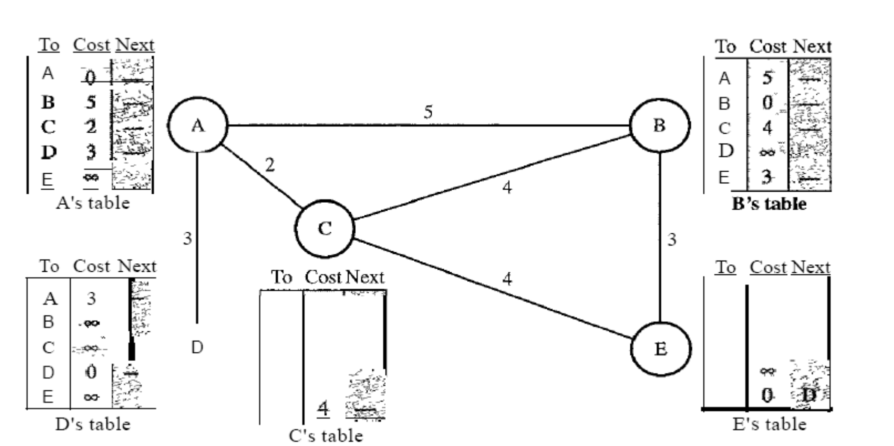
The table for node A shows how we can reach any node from this node. For example, our least cost to reach node E is 6. The route passes through C.

Initialization

The tables in Figure are stable; each node knows how to reach any other node and the cost. At the beginning, however, this is not the case. Each node can know only the distance between itself and its immediate neighbors, those directly connected to it.

So for the moment, we assume that each node can send a message to the immediate neighbors and find the distance between itself and these neighbors. Figure shows the initial tables for each node. The distance for any entry that is not a neighbor is marked as infinite (unreachable).

Initialization of tables in distance vector routing



Sharing

The whole idea of distance vector routing is the sharing of information between neighbors. Although node A does not know about node E, node C does. So if node C shares its routing table with A, node A can also know how to reach node E. On the other hand, node C does not know how to reach node D, but node A does. If node A shares its routing table with node C, node C also knows how to reach node D. In other words, nodes A and C, as immediate neighbors, can improve their routing tables if they help each other. There is only one problem. How much of the table must be shared with each neighbor? A node is not aware of a neighbor's table. The best solution for each node is to send its entire table to the neighbor and let the neighbor decide what part to use and what part to discard. However, the third column of a table (next stop) is not useful for the neighbor. When the neighbor receives a table, this column needs to be replaced with the sender's name. If any of the rows can be used, the next node is the sender of the table. A node therefore can send only the first two columns of its table to any neighbor. In other words, sharing here means sharing only the first two columns.

Updating

When a node receives a two-column table from a neighbor, it needs to update its routing table.

Updating takes three steps:

1. The receiving node needs to add the cost between itself and the sending node to each value in the second column. The logic is clear. If node C claims that its distance to a destination is *x* mi, and the distance between A and C is *y* mi, then the distance between A and that destination, via C, is *x* + *y* mi.

2. The receiving node needs to add the name of the sending node to each row as the third column if the receiving node uses information from any row. The sending node is the next node in the route.

3. The receiving node needs to compare each row of its old table with the corresponding row of the modified version of the received table.

a. If the next-node entry is different, the receiving node chooses the row with the smaller cost. If there is a tie, the old one is kept.

b. If the next-node entry is the same, the receiving node chooses the new row. For example, suppose node C has previously advertised a route to node X with distance 3. Suppose that now there is no path between C and X; node C now advertises this route with a distance of infinity. Node A must not ignore this value even though its old entry is smaller. The old route does not exist any more. The new route has a distance of infinity.

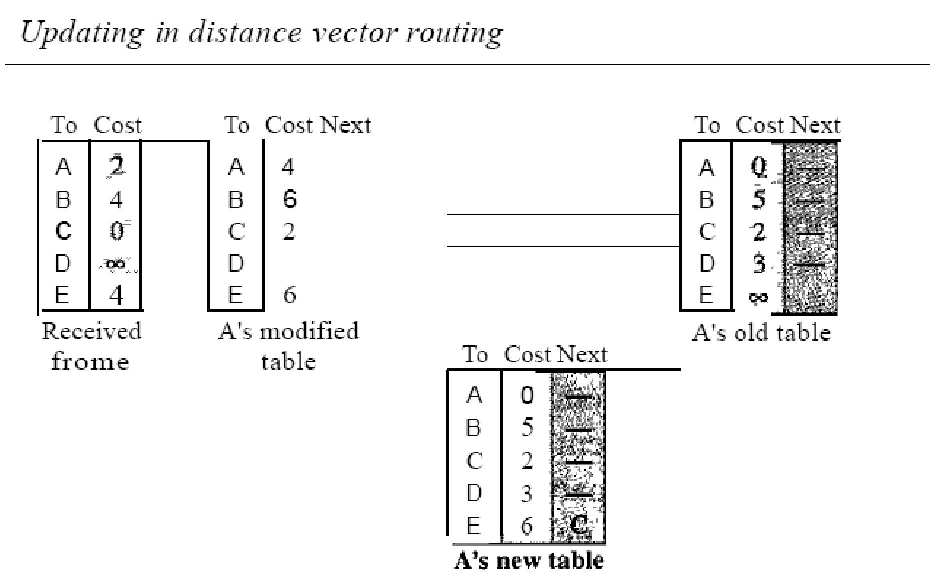


FIGURE : BELLMAN FORD ALGORITHM

There are several points we need to emphasize here. First, as we know from mathematics, when we add any number to infinity, the result is still infinity. Second, the modified table shows how to reach A from A via C. If A needs to reach itself via C, it needs to go to C and come back, a distance of 4. Third, the only benefit from this updating of node A is the last entry, how to reach E. Previously, node A did not know how to reach E (distance of infinity); now it knows that the cost is 6 via C.

Each node can update its table by using the tables received from other nodes. In a short time, if there is no change in the network itself, such as a failure in a link, each node reaches a stable condition in which the contents of its table remains the same.

When to Share

The question now is, When does a node send its partial routing table (only two columns) to all its immediate neighbors? The table is sent both periodically and when there is a change in the table.

Periodic Update

A node sends its routing table, normally every 30 s, in a periodic update. The period depends on the protocol that is using distance vector routing.

Triggered Update

A node sends its two-column routing table to its neighbors anytime there is a change in its routing table. This is called a triggered update. The change can result from the following.

1. A node receives a table from a neighbor, resulting in changes in its own table after updating.

2. A node detects some failure in the neighboring links which results in a distance change to infinity.

RIP

The Routing Information Protocol (RIP) is an intradomain routing protocol used inside an autonomous system. It is a very simple protocol based on distance vector routing. RIP implements distance vector routing directly with some considerations:

1. In an autonomous system, we are dealing with routers and networks (links). The routers have routing tables; networks do not.

2. The destination in a routing table is a network, which means the first column defines a network address.

3. The metric used by RIP is very simple; the distance is defined as the number of links (networks) to reach the destination. For this reason, the metric in RIP is called a hop count.

4. Infinity is defined as 16, which means that any route in an autonomous system using RIP cannot have more than 15 hops.

5. The next-node column defines the address of the router to which the packet is to be sent to reach its destination.

The table of each router is also shown. Let us look at the routing table for Rl. The table has seven entries to show how to reach each network in the autonomous system. Router Rl is directly connected to networks 130.10.0.0 and 130.11.0.0, which means that there are no next-hop entries for these two networks. To send a packet to one of the three networks at the far left, router Rl needs to deliver the packet to R2. The next-node entry for these three networks is the interface of router R2 with IP address 130.10.0.1. To send a packet to the two networks at the far right, router Rl needs to send the packet to the interface of router R4 with IP address 130.11.0.1.

Link State Routing [RGPV/Dec 2009/ Jun 2010]

Link state routing has a different philosophy from that of distance vector routing. In link state routing, if each node in the domain has the entire topology of the domain the list of nodes and links, how they are connected including the type, cost (metric), and condition of the links (up or down)-the node can use Dijkstra's algorithm to build a routing table.

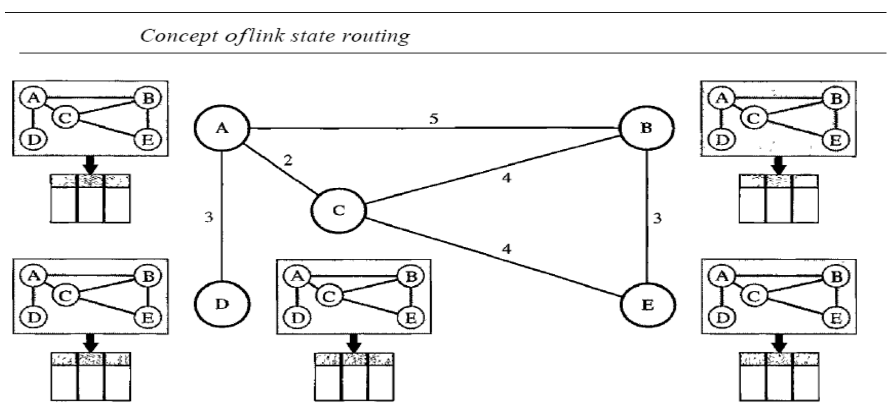


FIGURE: LINK STATE ROUTING

The figure shows a simple domain with five nodes. Each node uses the same topology to create a routing table, but the routing table for each node is unique because the calculations are based on different interpretations of the topology. This is analogous to a city map. While each person may have the same map, each needs to take a different route to reach her specific destination.

The topology must be dynamic, representing the latest state of each node and each link. If there are changes in any point in the network (a link is down, for example), the topology must be updated for each node. How can a common topology be dynamic and stored in each node? No node can know the topology at the beginning or after a change somewhere in the network. Link state routing is based on the assumption that, although the global knowledge about the topology is not clear, each node has partial knowledge: it knows the state (type, condition, and cost) of its links. In other words, the whole topology can be compiled from the partial knowledge of each node. Figure shows the same domain as in Figure, indicating the part of the knowledge belonging to each node.

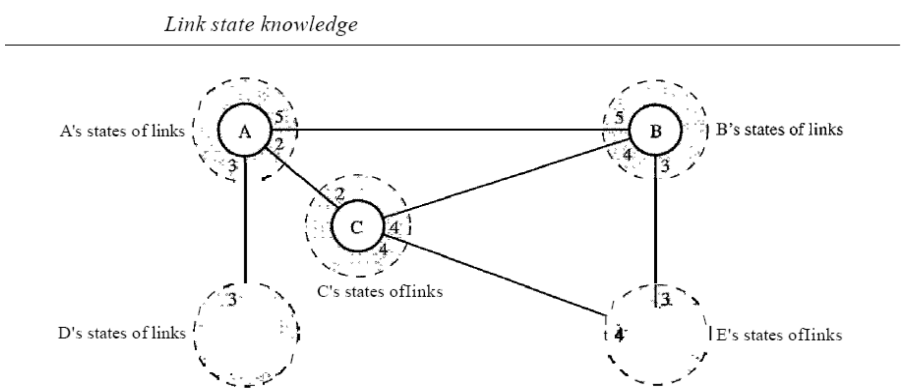


FIGURE: LINK STATE KNOWLEDGE

Node A knows that it is connected to node B with metric 5, to node C with metric 2, and to node D with metric 3. Node C knows that it is connected to node A with metric 2, to node B with metric 4, and to node E with metric 4. Node D knows that it is connected only to node A with metric 3. And so on. Although there is an overlap in the knowledge, the overlap guarantees the creation of a common topology-a picture of the whole domain for each node.

Dijkstra Algorithm[RGPV/Dec 2007, Dec 2010,Dec 2012]

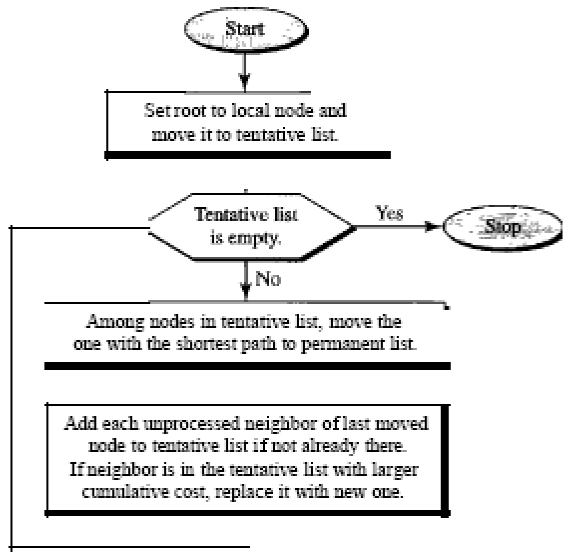


FIGURE : DIJKSTRA ALGORITHM FLOW CHART

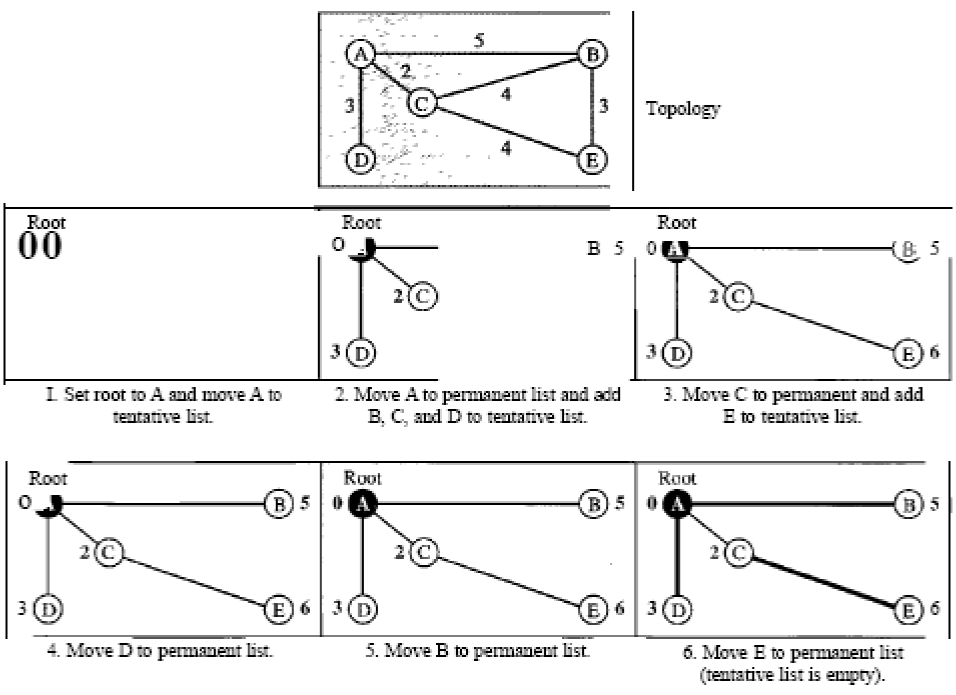


FIGURE : DIJKSTRA ALGORITHM

1. We make node A the root of the tree and move it to the tentative list. Our two lists are Permanent list: empty Tentative list: A(O)

2. Node A has the shortest cumulative cost from all nodes in the tentative list. We move A to the permanent list and add all neighbors of A to the tentative list. Our new lists are

Permanent list: A(O) Tentative list: B(5), C(2), D(3)

3. Node C has the shortest cumulative cost from all nodes in the tentative list. We move C to the permanent list. Node C has three neighbors, but node A is already processed, which makes the unprocessed neighbors just B and E. However, B is already in the tentative list with a cumulative cost of 5. Node A could also reach node B through C with a cumulative cost of 6. Since 5 is less than 6, we keep node B with a cumulative cost of 5 in the tentative list and do not replace it. Our new lists are

Permanent list: A(O), e(2) Tentative list: B(5), 0(3), E(6)

4. Node D has the shortest cumulative cost of all the nodes in the tentative list. We move D to the permanent list. Node D has no unprocessed neighbor to be added to the tentative list. Our new lists are

Permanent list: A(O), C(2), 0(3) Tentative list: B(5), E(6)

5. Node B has the shortest cumulative cost of all the nodes in the tentative list. We move B to the permanent list. We need to add all unprocessed neighbors of B to the tentative list (this is just node E). However, E(6) is already in the list with a smaller cumulative cost. The cumulative cost to node E, as the neighbor of B, is 8. We keep node E(6) in the tentative list. Our new lists are

Permanent list: A(O), B(5), C(2), 0(3) Tentative list: E(6)

6. Node E has the shortest cumulative cost from all nodes in the tentative list. We move E to the permanent list. Node E has no neighbor. Now the tentative list is empty. We stop; our shortest path tree is ready. The final lists are Permanent list: A(O), B(5), C(2), D(3), E(6) Tentative list: empty

OSPF

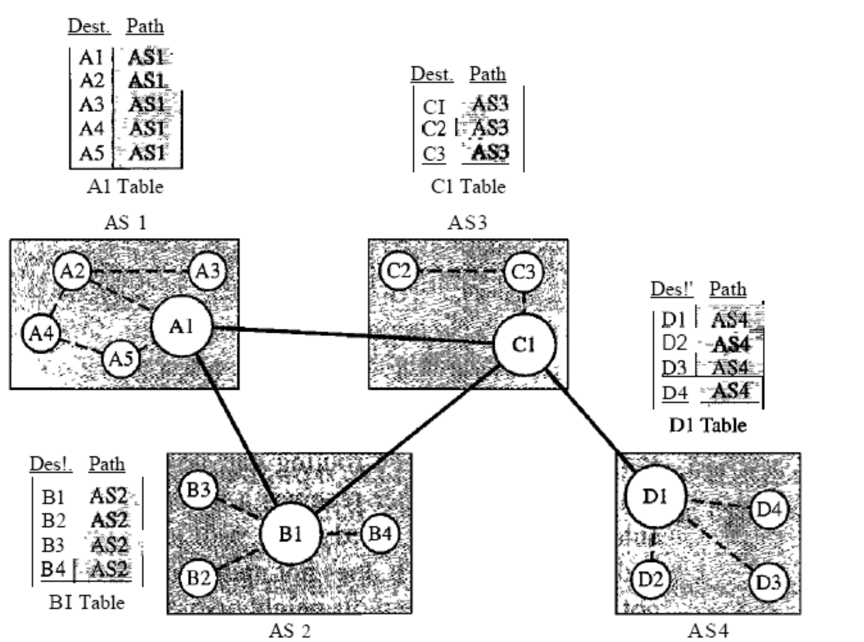
The Open Shortest Path First or OSPF protocol is an intradomain routing protocol based on link state routing. Its domain is also an autonomous system.

Areas To handle routing efficiently and in a timely manner, OSPF divide an autonomous system into areas. An area is a collection of networks, hosts, and routers all contained within an autonomous system. An autonomous system can be divided into many different areas. All networks inside an area must be connected. Routers inside an area flood the area with routing information. At the border of an area, special routers called area border routers summarize the information about the area and send it to other areas. Among the areas inside an autonomous system is a special area called the *backbone;* all the areas inside an autonomous system must be connected to the backbone. In other words, the backbone serves as a primary area and the other areas as secondary areas. This does not mean that the routers within areas cannot be connected to each other, however. The routers inside the backbone are called the backbone routers. Note that a backbone router can also be an area border router. If, because of some problem, the connectivity between a backbone and an area is broken, a virtual link between routers must be created by an administrator to allow continuity of the functions of the backbone as the primary area. Each area has area identification. The area identification of the backbone is zero.

Path Vector Routing

Distance vector and link state routing are both intradomain routing protocols. They can be used inside an autonomous system, but not between autonomous systems. These two protocols are not suitable for interdomain routing mostly because of scalability. Both of these routing protocols become intractable when the domain of operation becomes large. Distance vector routing is subject to instability if there are more than a few hops in the domain of operation. Link state routing needs a huge amount of resources to calculate routing tables. It also creates heavy traffic because of flooding. There is a need for a third routing protocol which we call path vector routing. Path vector routing proved to be useful for interdomain routing. The principle of path vector routing is similar to that of distance vector routing. In path vector routing, we assume that there is one node (there can be more, but one is enough for our conceptual discussion) in each autonomous system that acts on behalf of the entire autonomous system. Let us call it the speaker node. The speaker node in an AS creates a routing table and advertises it to speaker nodes in the neighboring ASs. The idea is the same as for distance vector routing except that only speaker nodes in each AS can communicate with each other. However, what is advertised is different. A speaker node advertises the path, not the metric of the nodes, in its autonomous system or other autonomous systems

Initial routing tables in path vector routing



BGP

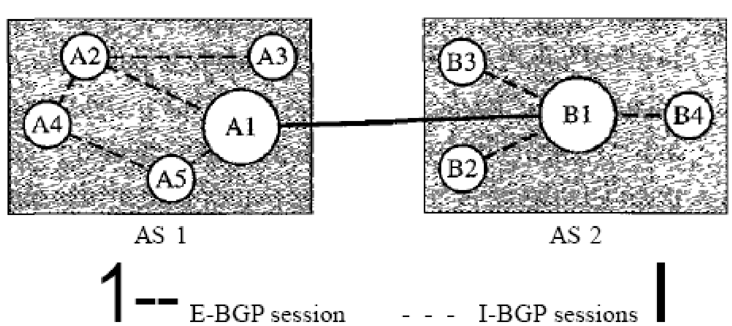
Border Gateway Protocol (BGP) is an interdomain routing protocol using path vector routing. It first appeared in 1989 and has gone through four versions. Types **of Autonomous** Systems As we said before, the Internet is divided into hierarchical domains called autonomous systems. For example, a large 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, 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).

Internal and external BGP sessions



S.NO	RGPV QUESTIONS	Year	Marks
Q.1	Explain distance vector routing with suitable example.	Dec 2012	7
Q.2	Write a short note on bellman-ford routing algorithm. Explain the drawback of count to infinity in bellman-ford algorithm.	Jun 2006 Dec 2008 Jun 2011	7
Q.3	Explain the two classes of routing algorithm- (i) Adaptive algorithm (ii) Non-adaptive algorithm Discuss about multipath routing.	Dec 2011	7
Q.4	compare the following- (i) Adaptive vs. non-adaptive routing	Jun 2011	7

	(ii) Centralized, isolated and distributed routing.		
Q.5	Explain least cost routing algorithm with example.	Jun 2010	7
Q.6	What is optimality principle? Why it is used in routing? Explain the shortest path routing algorithm.	Dec 2008 Dec 2009	7
Q.7	Apply dijkstra routing algorithm to calculate shortest path with source vertex	Dec 2007 Dec 2010	
Q.8	Describe about unicast routing protocol and multicast routing protocol.	Dec 2013	
Q.9	Explain bellman ford algorithm with example.	Dec 2012 Dec 2013	
Q.10	Describe the different adaptive routing strategies. What are the advantages and disadvantages of adaptive routing strategies?	Dec 2012	
Q.11	State and Describe the dijkstra algorithm with example.	Dec 2012	
Q.12	Name different unicast routing protocols any explain any of them in details	Jun 2014	
Q.13	Name different multicast routing protocols any explain any of them in details	Jun 2014	

UNIT 4/LECTURE 12

Internetworking Device

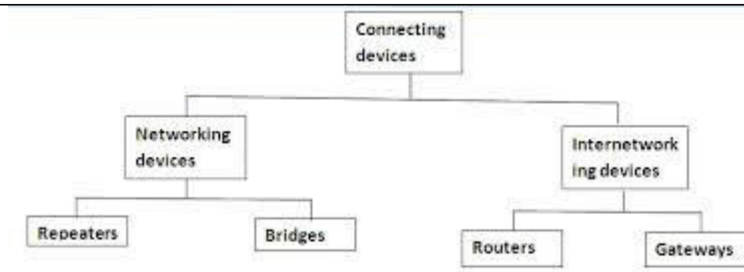


FIGURE: INTERNETWORKING DEVICE

Internetworking Device [RGPV/Dec 2011, Jun 2013]

An internetworking device is a widely-used term for any hardware within networks that connect different network resources. Key devices that comprise a network are

- routers
- bridges
- repeaters
- gateways

Routers

Routers are highly intelligent network devices that are primarily used for large networks and provide the best data path for effective communication. Routers have memory chips which store large quantities of network addresses.

A router is a device that analyzes the contents of data packets transmitted within a network or to another network. Routers determine whether the source and destination are on the same network or whether data must be transferred from one network type to another, which requires encapsulating the data packet with routing protocol header information for the new network type. When several routers are used in a collection of interconnected networks, they exchange and analyze information, and then build a table of the preferred routes and the rules for determining routes and destinations for that data. As a network interface, routers convert computer signals from one standard protocol to another that's more appropriate for the destination network. Large routers determine interconnectivity within an enterprise, between enterprises and the Internet, and between different internet service providers (ISPs); small routers determine interconnectivity for office or home networks. ISPs and major enterprises exchange routing information using border gateway protocol (BGP).

Bridges

Bridges are used to connect two large networks by providing different network services.

A bridge is a type of computer network device that provides interconnection with other bridge networks that use the same protocol.

Bridge devices work at the data link layer of the Open System Interconnect (OSI) model, connecting two different networks together and providing communication between them. Bridges are similar to repeaters and hubs in that they broadcast data to every node. However, bridges maintain the media access control (MAC) address table as soon as they discover new segments, so subsequent transmissions are sent to only to the desired recipient.

Bridges are also known as Layer 2 switches.

A network bridge device is primarily used in local area networks because they can potentially flood and clog a large network thanks to their ability to broadcast data to all the nodes if they don’t know the destination node's MAC address.

A bridge uses a database to ascertain where to pass, transmit or discard the data frame.

1. If the frame received by the bridge is meant for a segment that resides on the same host network, it will pass the frame to that node and the receiving bridge will then discard it.
2. If the bridge receives a frame whose node MAC address is of the connected network, it will forward the frame toward it.

Repeaters

Repeaters are used for signal and data regeneration and are primarily responsible for data amplification.

The term "repeater" originated with telegraphy in the 19th century, and referred to an electromechanical device used to regenerate telegraph signals. Use of the term has continued in telephony and data communications.

In telecommunication, the term repeater has the following standardized meanings:

1. An analog device that amplifies an input signal regardless of its nature (analog or digital).
2. A digital device that amplifies, reshapes, retimes, or performs a combination of any of these functions on a digital input signal for retransmission. A repeater that includes the retiming function is also known as a regenerator.

In computer networking, because repeaters work with the actual physical signal, and do not attempt to interpret the data being transmitted, they operate on the physical layer, the first layer of the OSI model.

Gateways

Gateways are internetworking devices used to convert formats and are the backbone of any network architecture, the term gateway has the following meaning:

- Gateway is a router or a proxy server that routes between networks
- Gateway Rule - Gateway should belong to same subnet to which your PC belongs
- In a communications network, a network node equipped for interfacing with another network that uses different protocols.
    - A gateway may contain devices such as protocol translators, impedance matching devices, rate converters, fault isolators, or signal translators as necessary to provide system interoperability. It also requires the establishment of mutually acceptable administrative procedures between both networks.
    - A protocol translation/mapping gateway interconnects networks with different network protocol technologies by performing the required protocol conversions.
- Loosely, a computer or computer program configured to perform the tasks of a gateway. For a specific case, see default gateway.

Gateways, also called protocol converters, can operate at any network layer. The activities of a gateway are more complex than that of the router or switch as it communicates using more than one protocol.

Both the computers of Internet users and the computers that serve pages to users are host nodes, while the nodes that connect the networks in between are gateways. For example, the computers that control traffic between company networks or the computers used by internet service providers (ISPs) to connect users to the internet are gateway nodes.

A gateway is often associated with both a router, which knows where to direct a given packet of data that arrives at the gateway, and a switch, which furnishes the actual path in and out of the gateway for a given packet.

Switch

A switch, in the context of networking is a high-speed device that receives incoming data packets and redirects them to their destination on a local area network (LAN). A LAN switch operates at the data link layer (Layer 2) or the network layer of the OSI Model and, as such it can support all types of packet protocols.

A switch in an Ethernet-based LAN reads incoming TCP/IP data packets/frames containing destination information as they pass into one or more input ports. The destination information in the packets is used to determine which output ports will be used to send the data on to its intended destination.

Switches are similar to hubs, only smarter. A hub simply connects all the nodes on the network communication are essentially in a haphazard manner with any device trying to communicate at any time, resulting in many collisions. A switch, on the other hand, creates an electronic tunnel between source and destination ports for a split second that no other traffic can enter. This results in communication without collisions.

Switches are similar to routers as well, but a router has the additional ability to forward packets between different networks, whereas a switch is limited to node-to-node communication on the same network.

Hub

A hub is the connection point in a computer device where data from many directions converge and are then sent out in many directions to respective devices. A hub may also act as a switch by preventing specific data packets from proceeding to a destination.

In addition to receiving and transmitting communication data, a hub may also serve as a switch. For example, an airport acts much like a hub in the sense that passengers converge there and head out in many different directions. Suppose that an airline passenger arrives at the airport hub and is then called back home unexpectedly, or receives instructions to change his or her destination. The same may occur with a computing hub when it acts as a switch by preventing specific data packets from proceeding to a destination, while sending other data packets on a specific route. Where packets are sent depends on attributes (MAC addresses) within the data packets. A switch may also act as a hub.

S.NO	RGPV QUESTIONS	Year	Marks
Q.1	Write short notes on the following networking devices- (a)Switches (b)Bridges (c)Hubs (d)gateway	Dec 2011	7
Q.2	Explain the following: (i) Bridges (ii)Routers(iii)Gateways	Jun 2013	7