UNIT 1 NETWORK LAYER

Structure **Page No.**

1.0 Introduction 5
1.1 Objectives 5
1.2 Switching 6
1.3 Routing Algorithm 7
1.3.1 Classification of Routing Algorithms
1.3.2 Non-Adaptive Routing Algorithm (Static Routing)
1.3.3 Dynamic Routing Algorithm (Adaptive)
1.3.4 Comparison Link State Versus Distance Vector Routing
1.4 Congestion Control 15
1.4.1 Algorithm For Congestion Control
1.5 Network Addressing 20
1.5.1 Classful Addressing
1.5.2 NetID and HostID
1.6 Fragmentation 26
1.7 Error Messaging Services 28
1.7.1 ICMP (Internet Control Message Protocol)
1.7.2 IGMP(Internet Group Message Protocol)
1.8 Summary 32
1.9 Further Reading 32
1.10 Solution/Answers 33

1.0 INTRODUCTION

As you know, the network layer is one of the important layers of OSI model. It is responsible for different tasks of networking, but mainly its role is to determine addresses and finding a route between a source and destination node or between two intermediate devices. It establishes and maintains a logical connection between these two nodes, either a connectionless or a connection oriented communication. The basic purpose of the network layer is to provide a network to network communication capability in contrast to machine to machine common provided by data line layer. The network layer controls the operation of the subnet. A key design issue is determining how packets are routed from source to destination. Routes can be determined based on static tables that are "wired into" the network and rarely changed. If too many packets are present in the subnet at the same time, they will get in one another's way forming bottlenecks. The controlling such congestion also belongs to the network layer. The quality of service also depends on network layer issue.

In this unit, we will study the fundamental Issues of network layer. These issues are designing interface between the host and the network, the routing methods, congestion control methods and Internetworking issues. In this unit we will study how routing is done at network layer using adaptive and non adaptive algorithm. We will also discuss the Network addressing. Further, some Error reporting protocols ICMP and IGMP on network layer will be discussed.

1.1 OBJECTIVES

After going through this unit, you should be able to:

- Know the basic issues of network layer
- Understand the different switching methods used at network layer
- Know the routing mechanisms

- Understand the congestion control methods
- Differentiate between adaptive and non adaptive algorithm
- Know process of Error reporting protocols at network layer

1.2 SWITCHING

As you have studied earlier in block 1, unit 3, that Switching is used to determine the path to be used for forwarding the information to the receiver. You also know that the Switching methods are mainly divided into Circuit, Message and Packet switching. In this section, we will explore the other switching mechanism like virtual circuit and datagram.

Virtual circuit is a connection oriented communication service that is delivered by means of packet mode communication. After a connection or virtual circuit is established between two nodes or application processes, a bit stream may be delivered between the nodes. It is similar to the circuit switching only in virtual circuit permanent/physical connection are not established. If router fails all virtual circuits that pass through, the failed router are terminated.

Datagram is opposite of Virtual circuit, it is connection less service. A datagram or packet needs to be self-contained without any dependency on earlier data-transfer because there is no connection of fixed duration between the two communicating nodes as shown in figure 1. Following Table 1 show the difference between circuit switching, virtual circuit and datagram:

Table 1: Difference Between Circuit Switching, Virtual Circuit and Datagram

S.No.	**Circuit Switching**	**Virtual Circuit**	**Datagram**
1.	Dedicated path between sender and receiver	Non-dedicated between sender and receiver.	Non-dedicated.
2.	Connection oriented Highly Reliable	Connection oriented	Connection less
3.	Data transfer in continuous form.	Data transfer in packets	Data transfer in packets.
4.	Bandwidth is fixed.	Not fined dynamic.	Dynamic
5.	E.g. telephone services	Subnet	Internet
6.	Data transfer in voice form.	Data transfer usually in text from.	Text form
7.	Call setup delay is maximum.	Delay Moderate	Call set up negligible.
8.	Transmission delay minimum	Moderate	Maximum transmission delay
9.	Unused bandwidth is just wasted	Moderate	No wasted.

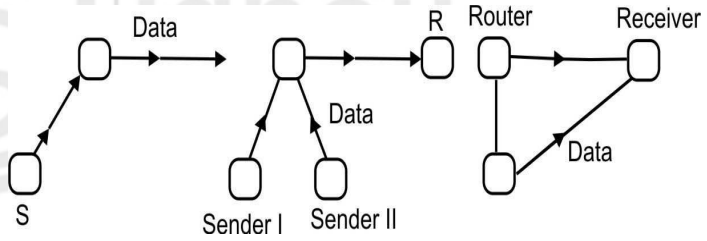


Figure 1: Circuit switching, Virtual circuit and Datagram (left to right)

1.3 ROUTING ALGORITHM

The main function of the network layer is routing packets from the source machine to the destination machine. So, the algorithm that choose the routes and the data structure that they use are a major area of network layer design.

It is that part of the network layer responsible for deciding which output line an incoming packet should be transmitted on.

Desired Properties of a Routing Algorithm

1. **Correction:** The routing should be done properly and correctly so that the packets may reach their proper destination.

2. **Simplicity:** The routing should be done in a simple manner so that the over head is as low as possible.

3. **Robustness:** Once a major network becomes operative, it may be expected to run continuously for years without any failure.

4. **Stability:** The routing algorithm should be stable under all possible circumstances.

5. **Fairness:** Every node connected to the network gets a fair change of transmitting their packets. This is generally done on a first come first serve basis.

6. **Optimality:** The routing algorithms should be optimal in terms of throughput and minimizing mean packet delays. Here there is a trade off and one has to choose depending on his suitability.

1.3.1 Classification of Routing Algorithms

Routing algorithm may be classified as follows:

1. Adaptive Algorithm

2. Non-adaptive algorithms

Adaptive algorithms use such dynamic information as current topology, load delay etc to select routes.

Non adaptive algorithms, routes never changes once initial routes have been selected. Also, called static routing.

Adaptive Routing Algorithm (Dynamic Routing)

It changes their routing decision to reflect changes in the topology and in traffic as well. These get their routing information from adjacent routers or from all routers. Routing decision may be changed when network topology and/or traffic load changes. The optimization parameters are the distance, number of hops and estimated transit time.

Adaptive routing algorithms can be further classified as follows:

1. **Isolated:** Each router makes its routing decisions using only the local information it has on hand. Specifically, routers do not even exchange information with their neighbors.

2. **Centralized:** A centralized node makes all routing decision specifically the centralized node has access to global information.

3. **Distributed:** Algorithm that uses a combination of local and global information.

Isolated: In this method, the node decides the routing without seeking information from other node. The disadvantage is that the packet may be sent through a congested route resulting in a delay.
Some of the examples of this type of algorithm for routing are :

- **Hot Potato Routing:** Form of routing in which the nodes of a network have no buffer to store packets in before they are moved on to their final predetermined destination.
- In normal routing situation, when multiple packets contend for a single outgoing channel, packets that are not buffered are dropped to avoid congestion.
- Backward Learning: In this method the routing tables at each node gets modified by information from the incoming packets. Backward learning routing algorithm used for routing traffic that makes decisions by assume that a can optimally reach B through C.

Centralized Routing

Advantage: Only one node is required to keep the information.

Disadvantage: If the central node goes down the entire network is down, i.e. single point of failure.

Distributed: It receives information from its neighboring nodes and then takes the decision about which way to send the packet.

Disadvantages: If in between the interval it receives information and sends the packet, something changes then packet may be delayed.

Optimality Principle

Optimality principle is a general statement about optimal routes regardless of network topology or traffic.

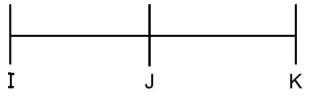


Figure 2: An example of Optimality Principle

With reference to Figure 2 above, Optimality principle states that if router I is on the optimal path from router 'I' to router 'K' then the optimal path from 'J' to 'K' also falls along the same route.

To prove the above statement we can say, if there was a better way from J to K, then you could use that with the path from I to J for a better path from I to K, so your starting point (the path from I to K was optimal) is contradicted.

1.3.2 Non-Adaptive Routing Algorithm (Static Routing)

These algorithms do not take their routing decisions on measurements and estimates of the current traffic and topology. Instead the route to be taken from one node to the other is computed in advance. This is also known as static routing.

1. **Shortest Path Routing:** According to this algorithm build a graph of the subnet, with each node of graph representing a router and each arc of the graph representing communication line. To choose a route between a pair of routers, just finds the shortest path between them on the graph.

   Two ways of measuring distance

   a) Distance in terms of link delay.

   b) Measuring path length in number of hops.

      In the most general case, the labels on the arc could be computed as a function of the bandwidth, average traffic communication cost, mean queue length measured delay, and other factors.

   c) **Flooding:** According to this algorithm every incoming packet is sent out on every outgoing line except the cone it arrived on.

      Flooding generates vast number of duplicate packets unless some measures are taken to damp the process. One such measure is to have a hop counter contained in the header of each packet, which is decremented at each hop with the packet being discarded when the counter reaches zero. The hop counter should be initialized to the length of path from source to destination. If the sender does not know how long the path is, it can be initialize the counter to the worst case, namely, the full diameter of the subnet. An alternative technique for damming the flood is to keep track of which packets have been flooded, to avoid sending them out a second time.

   d) **Selective Flooding**

      In this algorithm the routers do not send every incoming packet out on every line, only on those line that are going approximately in right direction.

      Though flooding is not practical in most application, but (yes) is does have some uses.

      i) **In military applications:** Tremendous robustness of flooding is highly desirable.

      ii) **In distributed data base application:** It is sometimes necessary to update all the data bases concurrently, in which case flooding can be useful.

Disadvantages:

1. Duplicacy

2. Infinite looping.

3. **Flow-Based Routing:** This algorithm considers two strategies in account to decide the route.

a) Topology.
b) Load for routing

Previous static algorithm only considers topology in account not the load for routing. The basic idea behind the analysis is that for a given line, if capacity and average flow is known, it is possible to compute the mean packet delay on that line from queuing theory. The routing problem then reduces to finding the routing algorithm that produces the minimum average delay for the subnet.

For this technique, certain information must be known in advance.

c) Topology

d) Traffic Matrix Fij

e) Line capacity matrix Cij

Now $T = \frac{1}{\mu_c - \lambda}$

Where T = Mean delay

$\frac{1}{\mu}$ = Mean packet size

λ = Mean flow in packet/sec. (No of arrival frame on particular line).

C = Capacity

Two common methods are used to calculate the shortest path between two routers.

1. Distance vector routing (Bellman ford routing algorithm and the Ford Fulkerson algorithm).
2. Link state routing (based on Dikastra's algorithm).

1.3.3 Dynamic Routing Algorithm (Adaptive)

Routing algorithms can be classified based on inter-domain and intra-domain as shown in figure 3 given below.

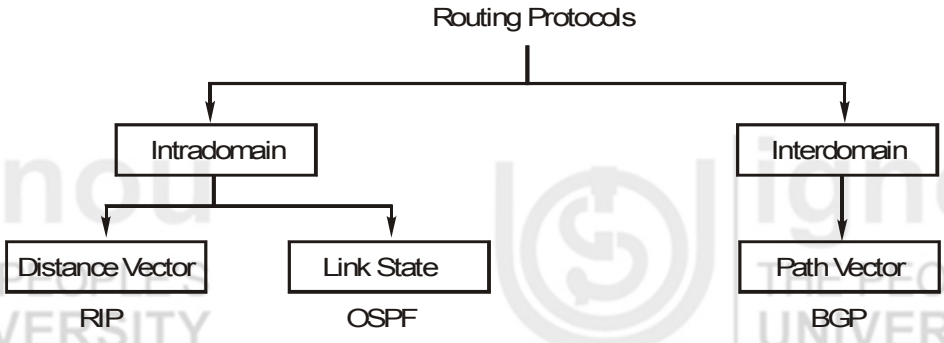


Figure 3: Classification of routing algorithms

RIP □ Routing information protocol

OSPF □ Open shortest path first

BGP □ Border gateway protocol.

1. **Distance vector routing**

Distance Vector Routing: According to this algorithm, each router maintains a table (vector) giving the best known distance to each destination and which line to use to get there. These tables are updated by exchanging information with the neighbors. This algorithm is also called **Bellman-Ford** or the **Ford-Fulkerson Algorithm**.

In distance vector routing, each router maintains a routing table indexed by, and containing one entry for each router in the subnet. This entry contains two parts

a) The preferred outgoing line to use for that destination.

b) An estimate of the time as distance to that destination.

The router is assumed to know the distance to each of its neighbors.
For example, consider a subnet as given below in figure 4.

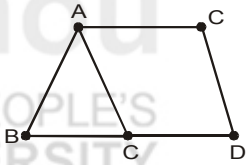


Figure 4: Subnet Diagram

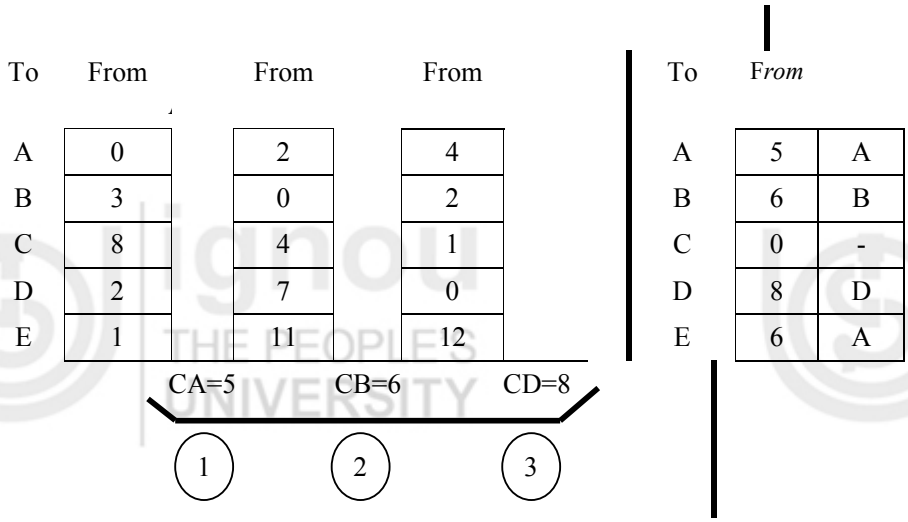


Figure 5: An example of Distance Vector Routing

Part (a) shows a subnet.

Part (b) the first 3 column shows the delay received from neighbor of router 'C' is A and B and D.

For example : as shown in the figure 5, 'A' claims to have 3 msec delay to B 8 msec delay to 'C' and so on. Similarly 'B' claims to have 2 msec delay to A, 4 msec delay to 'C' and so on.

'C' has estimated his delay to neighbour A, B, D as 5, 6, 8 respectively (CA = 5, CB = 6, CD = 8).

Now (4) column shows how router 'C' decides his new route to router 'E'. There are three ways

a) If 'C' follows line 'A' then delay is CE = CA ® AE = 5 + 1 = 6 msec.

c) If ‘C’ follows line ‘B’ then delay is CE = CB ® BE = 6 + 11 = 17 msec.

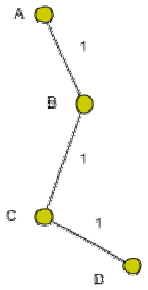
d) If ‘C’ follows line ‘D’ then delay is CD ® DE ® 8 + 12 = 20 msec.

Min delay time is via neighbor route ‘A’ so from C to E line is chosen ‘A’ in column (4)

The same calculations is performed for all destination, with the new routing table as (4)

Problem of Distance Vector Routing

1.**Count to Infinity Problem**-Distance vector routing has a serious drawback in its receptivity. In particular, it reacts rapidly to good news, but slowly to bad news. Following illustration shows an imagined network and denotes the distances from router A to every other router in the figure 6. Until now everything works fine.



B	C	D
1	2	3

Figure 6: Count to Infinity Problem

The illustration shows that link (A, B) is broken. Router B observed it, but in its routing table he sees, that router C has a route to A with 2 hops.

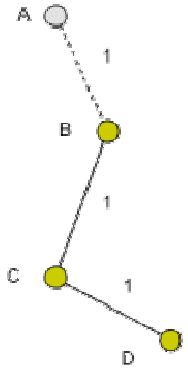
The problem is, that router B doesn't know that C has router B as successor in his routing table on the route to A.

That occurs followed count-to-infinity problem. B actualizes his routing table and takes the route to A over router C.

In the next figure 7; we can see the new distances to A. In C's routing the route to A contains router B as next hop router, so if B has increase his costs to A, C is forced to do so. Router C increases his cost to A about B + 1 = 4.

Now we see the consequence of the distributed Bellman-Ford protocol: Because router B takes the path over C to A, it updates its routing table and so on!
At the end, this problem is going to immobilize the whole network.

1.



B	C	D
3	2	3

B	C	D
3	4	3

B	C	D
5	4	5

Figure 7: Count to Infinity Problem illustration

2. **Hierarchical Routing:** As networks grow in size, the router routing tables grow proportionally. Not only the router memory consumed but also the more CPU time is needed to scan them and more bandwidth is needed.

The problem can be solved to some extent by using Hierarchical routing. In this routers are divided into regions as depicted in figure 8, with each router knowing all the details how to route packets to destination within it own region, but nothing about the internal structure of other regions.

For huge network, a two level hierarchy may be insufficient, it may be necessary to group the regions into clusters, the clusters into zones, the zones into groups and so on.

For example : When different networks are connected together, it is natural to regard each one as a separate region in order to free the routers in one network from having to know the topological structure of other ones.

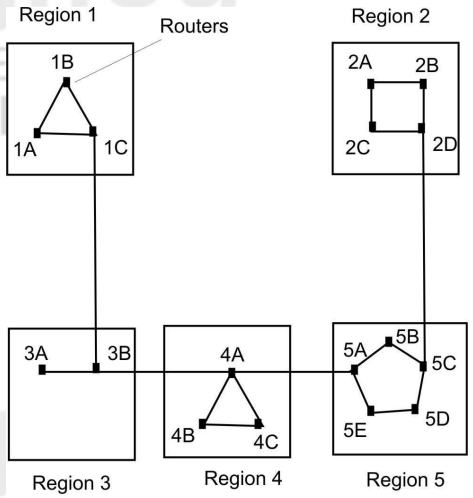


Figure 8: Hierarchical Routing

Full Table for 1A (Dest=Destination)

Dest	Line	Hops
1A		
1B	1B	1
1C	1C	1
2A	1B	2
2B	1B	3
2C	1B	3
2D	1B	4
3A	1C	3
3B	1C	2
4A	1C	2
4B	1C	4
4C	1C	4
5A	1C	4
5B	1C	5
5C	1B	5
5D	1C	6
5E	1C	5

Hierarchical Table for 1A

Dest	Line	Hops
1A		
1B	1B	1
1C	1C	1
2	1B	2
3	1C	2
4	1C	3
5	16	4

The full routing table as shown above for 1A has 17 entries but when routing is done hierarchically there are only 7 entries (e entries for local routers 4 entries for regions which are considered as single router. All the traffic for region 2 goes by 1B-2A line but rest of traffic goes by 1C-3B line.

The basic concept of link-state routing is that every node constructs a map of the connectivity to the network, in the form of a graph, showing which nodes are connected to which other nodes. Each node then independently calculates the next best logical path from it to every possible destination in the network. The collection of best paths will then form the node's routing table.

Disadvantages:

1. There is a penalty to be paid in the form of increased path length. For example the best route from 1A to 5C is via region 2, but with hierarchical routing all traffic to region 5 goes via region 3, because that is better for most destinations in region 5.

2. If single network become very large then multilevel hierarchy can be used.

   The presence of congestion means that the load is (temporarily) greater than the resources can handle.

3. Link state Routing

   Link State routing protocols (an adaptive routing algorithm) do not view networks in terms of adjacent routers and hop counts, but they build a comprehensive view of the overall network which fully describes the all possible routes along with their costs. Using the SPF (Shortest Path First) algorithm, the router creates a "topological database" which is a hierarchy reflecting the network routers it knows about. It then puts it's self on the top of this hierarchy, and has a complete picture from it's own perspective.

The complete working of algorithm can be divided into Five Steps:

1. Discover your neighbors and learn their addresses.

   In this process send “Hello”, packet on each point-to-point line. After receiving the hello packet Destination, node replies with its address.

2. Measure the cost (delay) to each neighbor.

   Send an “ECHO” packet over the line. Destination is required to respond to “ECHO” packet immediately. Measure the time required for this operation.

3. Construct a packet containing all this information

   The information tables are creating having all details of neighboring nodes.

4. Send this packet to all other routers.

   Use selective flooding. Sequence numbers prevent duplicate packets from being propagated. Lower sequence numbers are rejected as obsolete

5. Compute the shortest path to every other router.

   Dijkstra’s Shortest Path algorithm is used to determine the shortest path to each destination.

   When a router using a Link State protocol, such an OSPF (Open Shortest Path First) knows about a change on the network, it will broadcast this change instantly, there for flooding the network with this information. The information routers require to build their databases is provided in the form of Link State advertisement packets (LSAP). Routers do not advertise their entire routing tables; instead each router advertises only its information regarding immediately adjacent routers.

1.3.4 Comparison Link State Versus Distance Vector Routing

- Link state has big memory requirements
- In link state shortest path computations require many CPU circles
- Link state, If network is stable little bandwidth is used; react quickly to topology changes
- In link state announcements cannot be “filtered”. All items in the database must be sent to neighbors
- In link state all neighbors must be trusted
- In link state authentication mechanisms can be used to avoid undesired adjacencies
- In link state no split horizon techniques are possible

Even though Link State protocols work more efficiently, problem can arise. Usually problems occur cause of changes in the network topology (links go up-down), and all routers don't get updated immediately cause they might be on different line speeds, there for, routers connected via a fast link will receive these changes faster than the others on a slower link.

Different techniques have been developed to deal with these problem and these are:

1. Dampen update frequency
2. Target link-state updates to multicast
3. Use link-state area hierarchy for topology
4. Exchange route summaries at area borders
5. Use Time-stamps Update numbering & counters
6. Manage partitions using a area hierarchy

1.4 CONGESTION CONTROL

Congestion: When too many packets are in a subnet or a part of subnet, performance degrades as depicted in figure 9. This situation is called congestion.

Factors Causing the Congestion

1. Many input lines demanding the same output lines.
2. Slow receiver fast sender.
3. Low bandwidth lines can also cause congestion.
4. Congestion itself (duplicacy).
5. Traffic is bursty.

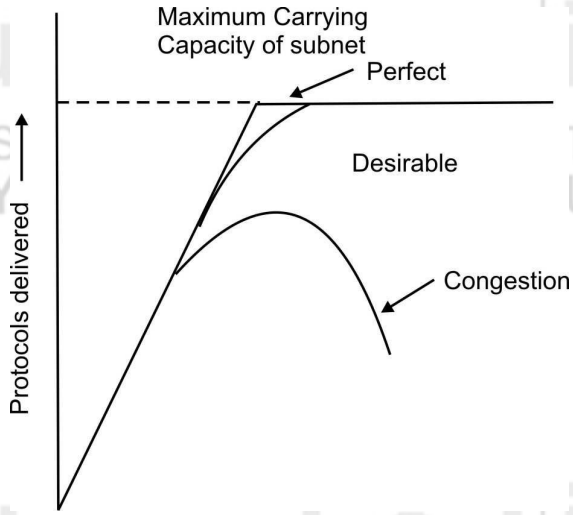


Figure 9: capacity of subnet and congestion

Congestion control principles are divided into two categories:

a) Open loop: In open loop solution the good designs are being developed to solve the problem so that congestion does not occur at first place once the system is setup and running, no mid pores connection is made. In open loop control, tools are included to decide when to accept new traffic, when to discard packets and which ones. And making scheduling decisions at various points in the network. The decisions are offline decision is not based on current state of network close loop solution.

   The concept of feedback loop is used in closed loop solution. This approach has three parts, when apply to the congestion control.

   i) Monitor the system to detect when and where congestion occurs.

   ii) Pass the information to places where the action can be taken.

   iii) Adjust system operation to correct the problem.

1.4.1 Algorithm for Congestion Control

Two major criterion of congestion control are

1. To decrease load

2. To increase capacity

Traffic Shaping (Congestion Control Policy in ATM)

One of the main causes of congestion is that traffic is often bursty. If hosts could be made to transmit a uniform rate, congestion would be less common. Another open loop method to help manage congestion is forcing the packet to be transmitted at a more predictable rate. This approach to congestion management is widely used in ATM networks and is called traffic shaping.

Traffic shaping is about regulating the average rate (and burstiness) of data transmission.

Leaky Bucket Algorithm

Imagine a bucket with a small hole in the bottom as depicted in figure 10. No matter at what rate enters the bucket, the outflow is at a constant rate, when there is any

water is bucket and zero when the bucket is empty. Also once the bucket is full, any additional water entering it spills over the sides and is lost.

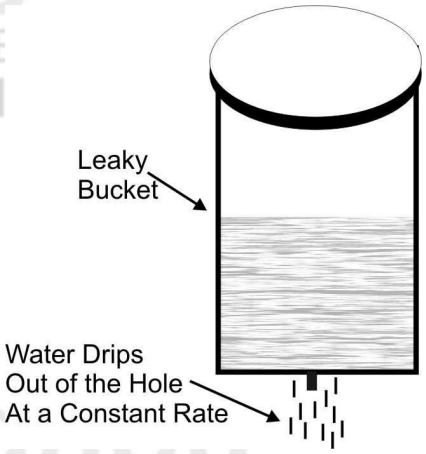


Figure 10: Normal Leaky bucket

The same idea can be applied to packets conceptually; each host is connected to the network by an interface containing a leaky bucket, i.e. a finite internal queue. If a packet arrives at the queue when it is full, the packet is discarded. This arrangement can be built into the hardware interface. It was first proposed by turner and is called leaky bucket algorithm as given figure 11.

Leaky bucket algorithm can be understood as "The leaky bucket consists of finite queue when a packet arrives, if there is room on the queue it is appended to the queue, otherwise it is discarded. At every clock tick, one packet is transmitted".

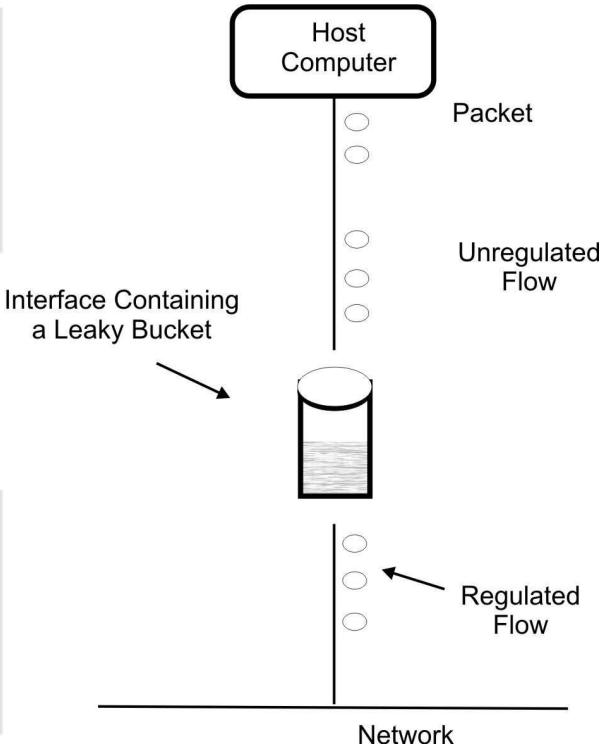


Figure 11: Leaky bucket algorithm

Advantage

This algorithm smoothens the bursts and greatly reduces the chances of congestion.

Disadvantages

1. When the queue is full, packets are discarded.
2. Sometimes it is necessary to speed up the output which is not possible in leaky bucket algorithm.

Token Bucket Algorithm: The leaky bucket algorithm has a rigid output pattern at the average rate, no matter how bursty the traffic is.

In many applications, it is better to allow the output to speed up somewhat when large bursts arrive so a more flexible algorithm is needed, preferably one that never losses data one such algorithm is token bus algorithm.

In token bucket algorithm, the leaky bucket holds 'tokens' generated by a clock at the rate of one token every DT sec.

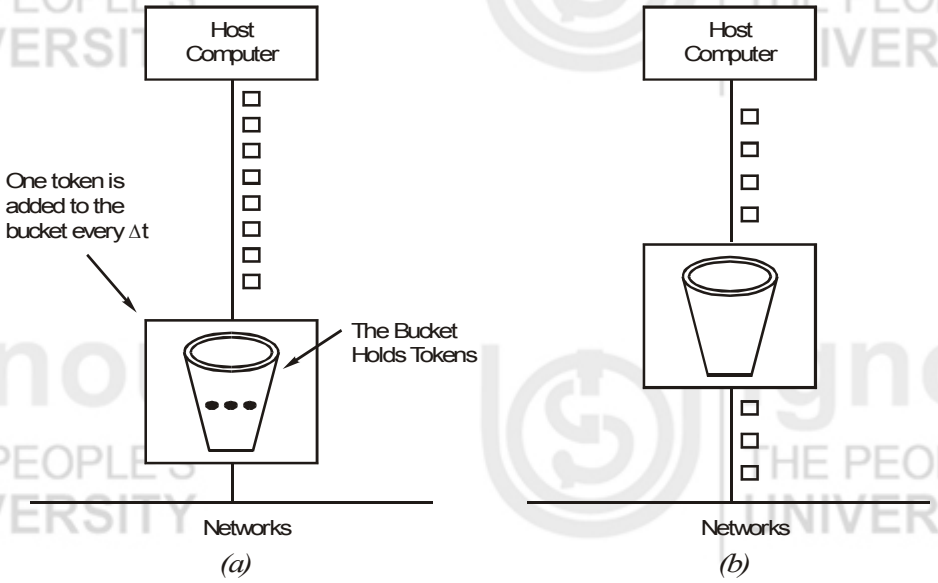


Figure 12: Token Bucket Algorithm

In above figure 12, we see a bucket holding three tokens, with five packets waiting to be transmitted. For a packet to be transmitted, it must capture and destroy one token. Three of fine packets have gone through, but other two are waiting for tokens to be generated.

The token bucket algorithm provides a different kind of traffic shaping than the leaky bucket algorithm. Bust of up to n. packets can be sent at once, allowing some burstiness in the output stream and giving faster response to sodden bursts of input.

A token bucket algorithm throws away tokens when the bucket fills up but never discards packets.

Table 1: Token Bucket V/S Leaky Bucket

S.No.	Token Bucket	Leaky Bucket
1.	The algorithm shaping is quite different.	In this there is a trade off between memory and bandwidth and packet life time.
2.	It allows saving up to maximum size of 'n' i.e. burst can be send of size 'n' at once.	It has constant traffic depending on the Leakage.
3.	Token bucket discarded token when bucket fills up.	This discards the packets when bucket fills

☞ Check Your Progress 1

1. The shortest path in routing can refer to……..

   a) The least expensive path
   b) The least distant path
   c) The path with the smallest number of hops
   d) Any or a combination of above

   ……

   ……

2. In distance vector routing, each router receives vector from

   a) Every router in the network
   b) Every router less than two units away
   c) A table stored by the software
   d) Its neighbor only

   ……

   ……

3. In link state routing, flooding allows changes to be recorded by

   a) All router
   b) Neighbor router only
   c) Some routers
   d) All networks

   ……

   ……

4. In which type of switching, do all the datagrams of a message follow the same channel of a path

   a) Circuit switching
   b) Datagram packet switching
   c) Virtual circuit packet switching
   d) Message switching

   ……

   ……

5. Which type of switching uses the entire capacity of a dedicated link

a) Circuit switching
b) Datagram packet switching
c) Virtual circuit packet switching
d) Message switching

……

……

1.5 NETWORK ADDRESSING

IP address versions

IP became the official protocol for the internet in 1983.As the internet has evolved, so has the IP. There have been six versions since its inception. Three versions are main.

1. Version 4(IPv4)
2. Version 5(Ipv5)
3. Version 6(Ipv6)

IP addressing

The identifier used in the IP layer of the TCP/IP protocol suite to identify each device connected to the Internet is called the Internet address or IP address.

OR

An IP address (Ipv4) is a 32-bit address that uniquely and universally defines the connection of a host or a router to the Internet.

There are **three** common ways in which IP addresses can be represented.

1. There is the binary notation which uses the base two number system to represent numbers.
2. There is the decimal notation which uses the base ten number system to represent numbers.
3. There is the hexadecimal notation which uses the base sixteen number system to represent numbers.

Do you know?

An IP address is a 32-bit(4-bytes) address.

Example 1

Assume, IGNOU's IP address is 142.190.23.180. This IP address consists of four bytes. The first byte has the value of 142. The second byte has the value of 190. The third byte has the value of 23, and the fourth byte has the value of 18.

Do you know?

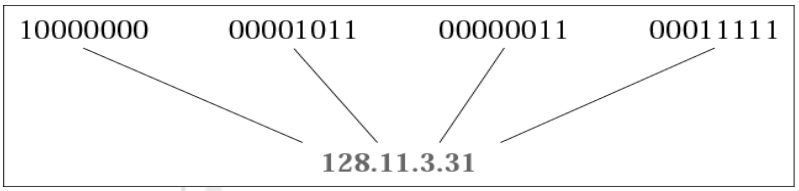
IP addresses are unique.

IP addresses are unique in the sense that each address defines one, and only one, connection to the Internet. Two devices on the Internet can never have the same address.

Do you know?

The address space of IPv4 is 2^{32} or 4,294,967,296.

Notations of IP addresses



Example 1

Change the following IP addresses from binary notation to dotted-decimal notation.

a) 10000001 00001011 00001011 11101111
b. 11000001 10000011 00011011 11111111
c. 11100111 11011011 10001011 01101111
d. 11111001 10011011 11111011 00001111

Solution

We replace each group of 8 bits with its equivalent decimal number and add dots for separation:

a) 129.11.11.239
b) 193.131.27.255
c) 231.219.139.111
d) 249.155.251.15

Example 2

Change the following IP addresses from dotted-decimal notation to binary notation.

a) 111.56.45.78
b) 221.34.7.82
c) 241.8.56.12
d) 75.45.34.78

Solution

We replace each decimal number with its binary equivalent:

a) 01101111 00111000 00101101 01001110
b) 11011101 00100010 00000111 01010010
c) 11110001 00001000 00111000 00001100
d) 01001011 00101101 00100010 01001110

Example 3

Find the error, if any, in the following IP addresses:

a) 111.56.045.78
b) 221.34.7.8.20
c) 75.45.301.14
d) 11100010.23.14.67

Solution

a) There are no leading zeroes in dotted-decimal notation (045).
b) We may not have more than four numbers in an IP address.
c) In dotted-decimal notation, each number is less than or equal to 255; 301 is outside this range.
d) A mixture of binary notation and dotted-decimal notation is not allowed.

Example 4

Change the following IP addresses from binary notation to hexadecimal notation.

a) 10000001 00001011 00001011 11101111
b) 11000001 10000011 00011011 11111111

Solution

We replace each group of 4 bits with its hexadecimal equivalent (see Appendix B). Note that hexadecimal notation normally has no added spaces or dots; however, 0X (or 0x) is added at the beginning or the subscript 16 at the end to show that the number is in hexadecimal.

a) 0X810B0BEF or 810B0BEF16

b) 0XC1831BFF or C1831BFF16

Classless Inter-Domain Routing allocates address space to Internet service providers and end users on any address bit boundary, instead of on 8-bit segments. CIDR notation is a syntax of specifying IP addresses and their associated routing prefix. It appends to the address a slash character and the decimal number of leading bits of the routing prefix, e.g., 192.0.2.0/24 for IPv4

1.5.1 Classful Addressing

IP addresses, when started a few decades ago, used the concept of classes. This architecture is called classful addressing. In the mid-1990s, a new architecture, called classless addressing, was introduced and will eventually supersede the original architecture. However, part of the Internet is still using classful addressing, but the migration is very fast.

Occupation of the address space

Class	Number of Addresses	Percentage
A	2^{31} = 2,147,483,648	50%
B	2^{30} = 1,073, 741, 824	25%
C	2^{29} = 536,870,912	12.5%
D	2^{28} = 268,435,456	6.25%
E	2^{28} = 268,435,456	6.25%

Table 13: Addresses per Class

When IP addresses were first started, they used the concept of classes. The range of IP addresses were divided into five classes: As shown in the figure 13 and 14, A, B, C, D, and E. Class A used up 50% of the address space, class B used up 25%, class C used up 12.5%, class D used up 6.25%, and class E also used up 6.25%.

Figure 14: Address Spaces of IPv4 classes

The way you recognize which class an IP address belongs to is by analyzing the first byte. If the number in the first byte is between 0-127, then the IP address is in the Class A range as shown in figure 15. If it is between 128-191 it is in Class B. If it is between 192-223 it is in the Class C range. If it is between 224-239 it is in the Class D range, and if it is between 240-255, then it belongs to Class E.

Class in decimal notations

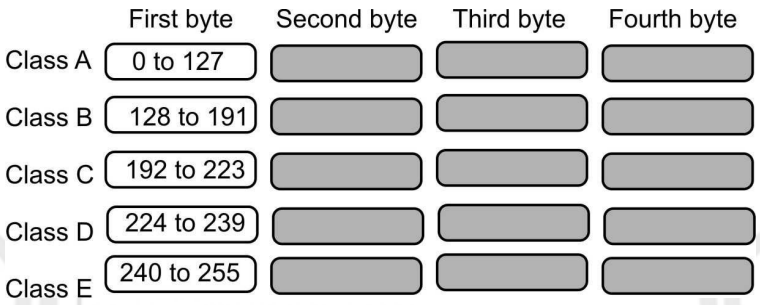


Figure 15: Classes of IPv4in decimal notation

Example

If IGNOU's IP address is 140.192.23.180. Looking at this address we can see that the first byte is 140. Since 140 is between the numbers 128-191, we know that it is in the Class B range.

Class in binary notation

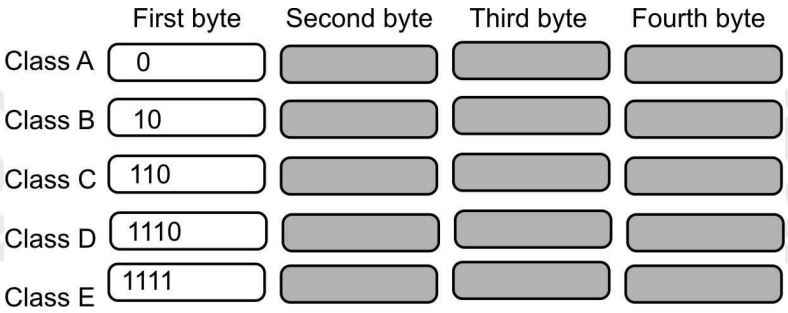


Figure 16: Classes of IPv4in binary notation

According to the figure 16, we can device a mechanism as given figure 17 for finding the address class in binary notation like:

If first left most bit is 0 then it is class A
If first bit is 1 and second bit is 0 then it is class B
If first two bits are 1 and third bit is 0 then it is class C
If first three bits are 1 and fourth bit is 0 then it is class D
If all the four bits are 1 then it is class E

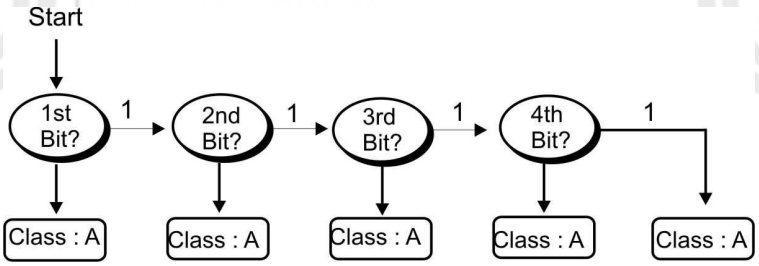


Figure 17: Finding the address class in binary notation

Example 5

How can we prove that we have 2,147,483,648 addresses in class A?

Solution

In class A, only 1 bit defines the class. The remaining 31 bits are available for the address. With 31 bits, we can have 2^{31} or 2,147,483,648 addresses.

Example 6

Find the class of each address:

a) 00000001 00001011 00001011 11101111
b) 11000001 10000011 00011011 11111111
c) 10100111 11011011 10001011 01101111
d) 11110011 10011011 11111011 00001111

Solution

a) The first bit is 0. This is a class A address.
b) The first 2 bits are 1; the third bit is 0. This is a class C address.
c) The first bit is 1; the second bit is 0. This is a class B address.
d) The first 4 bits are 1s. This is a class E address

Example 7

Find the class of each address:

a) 227.12.14.87
b) 193.14.56.22
c) 14.23.120.8
d) 252.5.15.111
e) 134.11.78.56

Solution

a) The first byte is 227 (between 224 and 239); the class is D.
b) The first byte is 193 (between 192 and 223); the class is C.
c) The first byte is 14 (between 0 and 127); the class is A.
d) The first byte is 252 (between 240 and 255); the class is E.
e) The first byte is 134 (between 128 and 191); the class is B.

Example 8

In Example 5 we showed that class A has 2^{31} (2,147,483,648) addresses. How can we prove this same fact using dotted-decimal notation?

Solution

The addresses in class A range from 0.0.0.0 to 127.255.255.255. We need to show that the difference between these two numbers is 2,147,483,648. This is a good exercise because it shows us how to define the range of addresses between two addresses. We notice that we are dealing with base 256 numbers here. Each byte in the notation has a weight. The weights are as follows

$$256^3, 256^2, 256^1, 256^0$$

Now to find the integer value of each number, we multiply each byte by its weight:

Last address: $127 \times 256^3 + 255 \times 256^2 +$

$$255 \times 256^1 + 255 \times 256^0 = 2,147,483,647$$

First address: = 0

If we subtract the first from the last and add 1 to the result (remember we always add 1 to get the range), we get 2,147,483,648 or 2^{31}.

1.5.2 NetID and HostID

An IP address is divided into a network ID (netid) and a host ID (hostid) as depicted in figure 18. The lengths of the netid vary depending on the class the IP address belongs to. In class A the netid occupies the first byte and the hostid occupies the remaining three bytes. In class B the netid occupies the first two bytes and the hostid occupies the remaining two bytes. In class C the first three bytes define the netid and the last remaining byte defines the hostid. Class D and E are not divided into netid and hostid. The following figure 18 shows how the netid and hostid are divided in Classes A, B, and C.

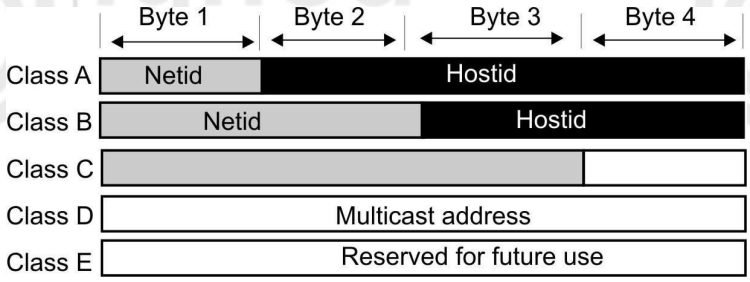


Figure 18: Network ID and a host ID of IPv4 classes

Do you know?

Class D addresses are used for multicasting; there is only one block in this class and Class E addresses are reserved for future purposes; most of the block is wasted.

In classful addressing the netid and hostid are easily distinguishable by looking at the IP address. First you have to determine which class the IP address belongs to and from there you can tell which part is the netid and which part is the hostid. If it is in Class A, then the first byte represents the netid and the last three represent the hostid, and so on.

Disadvantages of class full addressing

It wastes a lot of IP addresses and since the Internet keeps growing larger, we can't afford to throw away IP addresses.

That is why a new addressing scheme was devised. It is called classless addressing because it doesn't use the classes which were used in classful addressing.

Do you know?

Millions of class A and class B addresses and are wasted in class full addressing

Do you know?

The number of addresses in class C is smaller than the needs of most organizations.

Example 9

Given the network address 17.0.0.0, find the class, the block, and the range of the addresses.

Solution

The class is A because the first byte is between 0 and 127. The block has a netid of 17. The addresses range from 17.0.0.0 to 17.255.255.255.

Example 10

Given the network address 132.21.0.0, find the class, the block, and the range of the addresses.

Solution

The class is B because the first byte is between 128 and 191. The block has a netid of 132.21. The addresses range from 132.21.0.0 to 132.21.255.255.

Example 11

Given the network address 220.34.76.0, find the class, the block, and the range of the addresses.

Solution

class is C because the first byte is between 192 and 223. The block has a netid of 220.34.76. The addresses range from 220.34.76.0 to 220.34.76.255.

☞ Check Your Progress 2

1. Which IP address class has few hosts per network

   a) Class A
   b) Class B
   c) Class C
   d) Class D

   ……
   ……

2. Which of the following is true about IP addresses

   a) It is divided into exactly two classes
   b) It contains a fixed length host-id
   c) It was established as a user friendly interface
   d) It is 32 bits long

   ……
   ……

3. Which of the following is class C host address

   a) 230.0.0.0
   b) 130.4.4.6
   c) 200.1.2.3
   d) 30.4.5.6

   ……
   ……

1.6 FRAGMENTATION

Each network imposes some maximum size on its packets. A problem appears when a large packet wants to travel through a network whose maximum packet size is too small. One solution is to make sure the problem does not occur in the first place. In other words, the internet should use a routing algorithm that avoids sending packets through networks that cannot handle them. However, this solution is no solution at all. What happens if the original source packet is too large to be handled by the destination network? The routing algorithm can hardly bypass the destination.

Basically, the only solution to the problem is to allow gateways to break up packets into **fragments,** sending each fragment as a separate internet packet. However, converting a large object into small fragments is considerably easier than the reverse process.

Two opposing strategies exist for recombining the fragments back into the original packet.

1. Transparent Fragmentation
2. Non Transparent Fragmentation

Transparent Fragmentation

The first strategy is to make fragmentation caused by a "small-packet" network transparent to any subsequent networks through which the packet must pass on its way to the ultimate destination. This option is shown in Figure 19 (a). In this approach, the small-packet network has gateways that interface to other networks. When an oversized packet arrives at a gateway, the gateway breaks it up into fragments. Each fragment is addressed to the same exit gateway, where the pieces are recombined. In this way, passage through the small-packet network has been made transparent. Subsequent networks are not even aware that fragmentation has occurred.

Do you know?

ATM networks have special hardware to provide transparent fragmentation of packets into cells and then reassembly of cells into packets. In the ATM world, fragmentation is called segmentation

Transparent fragmentation is straightforward as shown in figure 19 a.

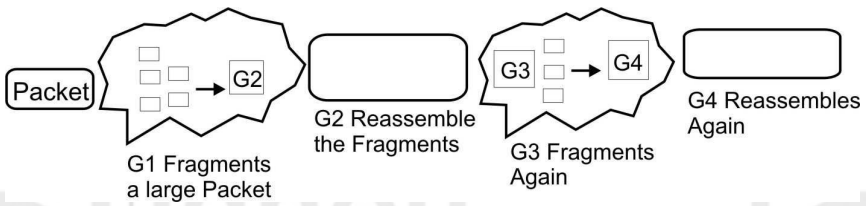


Figure 19 a: Transparent fragmentation

Drawback of transparent fragmentation

1. The exit gateway must know when it has received all the pieces, so either a count field or an "end of packet" bit must be provided.

2. All packets must exit via the same gateway. By not allowing some fragments to follow one route to the ultimate destination and other fragments a disjoint route. some performance may be lost.

3. A last problem is the overhead required to repeatedly reassemble and then re-fragment a large packet passing through a series of small packet networks.

Tips

ATM requires transparent fragmentation.

Non Transparent Fragmentation

The nontransparent fragmentation strategy refrains from recombining fragments at any intermediate gateways. Once a packet has been fragmented, each fragment is treated as though it were an original packet. All fragments are passed through the exit gateway (or gateways), as shown in Figure 19 (b). Recombination occurs only at destination host. IP works this way.

Non Transparent fragmentation also has some problems. For example, it requires every host to be able to do reassembly. Yet another problem is that when the large packet is fragmented the total overhead increases, because each fragment must have a header.

An advantage of this method is that multiple exit gateways can now be used and higher performance can be achieved

When a packet is fragmented, the fragments must be numbered in such a way that the original data stream can be reconstructed. One way of numbering the fragments is to use a tree. If packet 0 must be split up, the pieces are called 0.0, 0.1, 0.2 etc. If these fragments themselves must be fragmented later on, the pieces are numbered 0.0.0, 0.0.1, 0.0.2, ….0.1.2 etc. If enough fields have been reserved in the header for the worst case and no duplicates generated anywhere, this scheme is sufficient to ensure that all the pieces can be correctly reassembled at the destination, no matter what order they arrive in.
However, if even one network loses or discards packets, end-to-end retransmissions are needed, with unfortunate effects for the numbering system. Suppose that a 1024-bit packet is initially fragmented into four equal-sized fragments, 0.0, 0.1, 0.2 and 0.3. Fragment 0.1 is lost, but the other parts arrive at the destination. Eventually, the source times out and retransmits the original packet again. Only this time the route taken passes through a network with a 512-bit limit, so two fragments are generated. When the new fragment 0.1 arrives at the destination, the receiver will think that all four pieces are now accounted for a reconstruct the packet incorrectly.

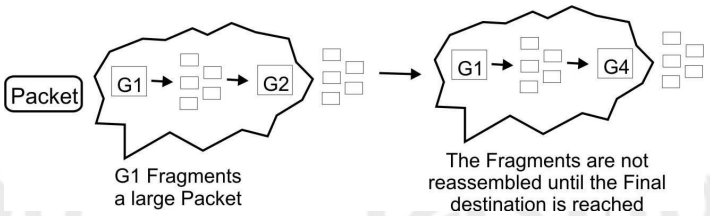


Figure 19 b: Nontransparent fragmentation

1.7 ERROR MESSAGING SERVICES

1.7.1 ICMP (Internet Control Message Protocol)

The Internet Control Message Protocol (ICMP) is a helper protocol that supports IP with facility for Error reporting Simple queries.

ICMP messages are sent in following situations

- when a datagram cannot reach its destination,
- when the gateway does not have the buffering capacity to forward a datagram,
- When the gateway can direct the host to send traffic on a shorter route.

Do you know?
ICMP is considered an integral part of IP as shown in figure 20.

The Internet Protocol (IP) is not designed to be absolutely reliable. The purpose of these control messages is to provide feedback about problems in the communication environment, not to make IP reliable. There are still no guarantees that a datagram will be delivered or a control message will be returned. Some datagrams may still be undelivered without any report of their loss. The higher level protocols that use IP must implement their own reliability procedures if reliable communication is required. The ICMP messages typically report errors in the processing of datagrams. To avoid the infinite regress of messages about messages etc., no ICMP messages are sent about ICMP messages.

TIPS
ICMP provides error reporting, flow control and first-hop gateway redirection.

Position of ICMP in the network layer

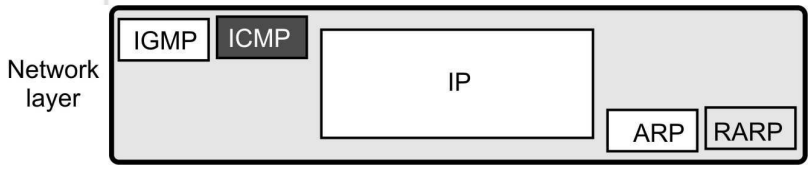


Figure 20: Protocols of Internet layer (TCP/IP)

ICMP header format

It consists of following fields

- Type
- Code
- ICMP header checksum
- Data

00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Type								Code								ICMP header checksum															
Data																															

The fields can be described as follows

1. **Type.** It is of 8 bits. It specifies the format of the ICMP message.
2. **Code.** It is of 8 bits. It further qualifies the ICMP message.
3. **ICMP Header Checksum.** It is of 16 bits. It is checksum that covers the ICMP message. This is the 16-bit one's complement of the one's complement sum of the ICMP message starting with the Type field. The checksum field should be cleared to zero before generating the checksum.
4. **Data.** It is of variable length. It contains the data specific to the message type indicated by the Type and Code fields

Types of ICMP messages

Each ICMP message contains three fields that define its purpose and provide a checksum. They are

- TYPE,
- CODE, and
- CHECKSUM fields (described above).

The TYPE field identifies the ICMP message, the CODE field provides further information about the associated TYPE field, and the CHECKSUM provides a method for determining the integrity of the message.

Tips

ICMP message are sent as packet so these are also called ICMP packet

ICMP messages are divided into two categories

- Error-reporting messages.
- Query messages.

Tips

ICMP messages are identified by "type" numbers

The **error-reporting** messages report problems that a router or a host (destination) may encounter. The **query messages** get specific information from a router or another host. For example, this can be used by the hosts to discover the routers present in their network. The host would send a ICMP query asking for routers to respond. The outers present in the network will respond with an ICMP reply message. The host would get information about the router from this reply.

Examples of error reporting messages

- Destination unreachable
- Source quench
- Time exceeded
- Parameter problem
- Redirection

Example of query messages

- Echo request and reply
- Timestamp request and reply
- Address mask request and reply
- Router solicitation and advertisement

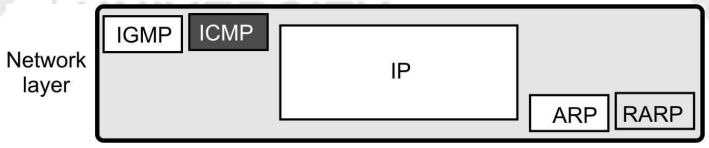
1.7.2 IGMP(Internet Group Message Protocol)

IGMP is a protocol that manages group membership. The IGMP protocol gives the multicast routers information about the membership status of hosts (routers) connected to the network.

Tips

Internet Group Management Protocol (IGMP) is the protocol used to support multicasting.

Position of IGMP in the network layer



Types of messages in IGMP

IGMP has three types of messages:

- the query,
- the membership report,
- and the leave report.

There are two types of query messages as shown in figure 21, general and special

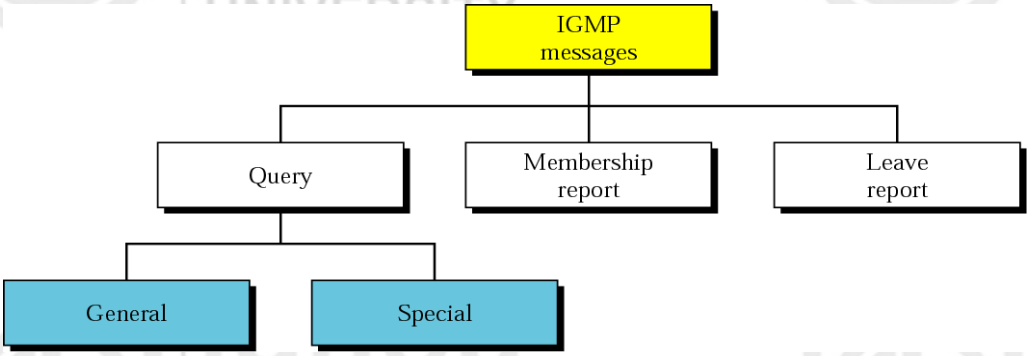


Figure 21: Types of query messages in IGMP

IGMP frame format

It consists of following fields

- Type
- Maximum response time
- Checksum
- Group addresses

00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Type								Code								IGMP Checksum															
Identifier																															
Group Address																															
Access Key																															

The fields can be described as follows

1. **Type.** It is of 8 bits. Following types of IGMP messages are possible

2. **Code.** It is of 8 bits. In a Create Group Request message, this field indicates if the new host group is to be public or private. In all other Request messages, this field is set to zero.

   In a Reply message, the *Code* field specifies the outcome of the request.

3. **IGMP Checksum.** It is of 16 bits. The checksum is the 16-bit one's complement of the one's complement sum of the IGMP message starting with the IGMP Type. For computing the checksum, the checksum field should first be cleared to 0. When the data packet is transmitted, the checksum is computed and inserted into this field. When the data packet is received, the checksum is again computed and verified against the checksum field. If the two checksums do not match then an error has occurred.

4. **Identifier.** It is of 32 bits. In a confirm Group Request message, the identifier field contains zero. In all other Request messages, the identifier field contains a value to distinguish the request from other requests by the same host. In a Reply message, the identifier field contains the same value as in the corresponding Request message.

5. **Group Address.** It is of 32 bits. In a Create Group Request message, the group address field contains zero. In all other Request messages, the group address field contains a host group address. In a Create Group Reply message, the group address field contains either a newly allocated host group address (if the request is granted) or zero (if denied). In all other Reply messages, the group address field contains the same host group address as in the corresponding Request message.

6. **Access Key.** This field is of 64 bits. In a Create Group Request message, the access key field contains zero. In all other Request messages, the access key field contains the access key assigned to the host group identified in the Group Address field (zero for public groups). In a Create Group Reply message, the access key field contains either a non-zero 64-bit number (if the request for a private group is granted) or zero. In all other Reply messages, the access key field contains the same access key as in the corresponding Request.

Do you know?

IGMP is defined in RFC 1112.

1.8 SUMMARY

In this unit, we studied various design issues of network layer. Network layer provides best route from source to destination using adaptive routing algorithm like distance vector routing and link state routing. A serous drawback of distance vector routing is count to infinity problem. It is also responsible for congestion control using leaky bucket and token leaky bucket algorithm. The four main protocols that operates on network layer are ARP, RARP, ICMP, IGMP. Network layer mainly works on IP address. IP addresses are 32bits.IP addresses have been divided into five classes namely A,B,C,D,E.ICMP and .ICMP and IGMP are error reporting protocols.ARP and RARP are used for address translation.

1.9 REFERENCES/FURTHER READING

1. Introduction to Data Communication & Networking, 3rd Edition, Behrouz Forouzan, Tata McGraw Hill.

2. Computer Networks, A. S. Tanenbaum 4th Edition, Practice Hall of India, New Delhi. 2003.
3. Douglas E. Comer, Internetworking with TCP/IP Vol.1: Principles, Protocols, and Architecture (4th Edition).
4. James F. Kurose, Computer Networking: A Top-Down Approach Featuring the Internet (3rd Edition).
5. Larry L. Peterson, Computer Networks: A Systems Approach, 3rd Edition (The Morgan Kaufmann Series in Networking).
6. www. wikipedia.org
7. W. Richard Stevens, The Protocols (TCP/IP Illustrated, Volume 1).
8. William Stallings, Data and Computer Communications, Seventh Edition.

1.10 SOLUTION/ANSWERS

☞ Check Your Progress 1

1. D
2. D
3. A
4. C
5. A

☞ Check Your Progress 2

1. C
2. D
3. C

UNIT 2 TRANSPORT LAYER

Structure **Page No.**

2.0	Introduction	34
2.1	Objective	34
2.2	Addressing	35
2.3	Reliable delivery	35
2.4	Flow control	38
2.5	Connection Management	38
2.6	Multiplexing	40
2.7	Congestion Control	40
2.8	Quality of Cervices (QoS)	42
2.9	TCP window Management	43
2.10	Ports	44
2.11	Summary	46
2.12	References/Further Reading	46
2.13	Solution /Answers	47

2.0 INTRODUCTION

The transport layer supports two protocols in TCP/IP protocol suite. One is Transmission Control Protocol (TCP). TCP is connection oriented that provides reliable end-to-end transmission. Another protocol is User Datagram Protocol (UDP). UDP is simple and provides well sequenced transport function when reliability and serving are less important than size and speed. Transport layer services are implemented by transport protocols used between two transport entities. Transport layer services are similar to the data link services. Data link layer is designed to provide its services within a single network, while the transport layer provides services across an inter network made up of many networks. There are seven categories of services provided by the transport layer. These services are End to end delivery, Addressing, Reliable delivery, Flow control, Connection management, Multiplexing and Congestion Control.

2.1 OBJECTIVE

After going through this unit, you should be able to:

- Know the Functions and Services of transport layer
- Understand the Working of transport layer
- Understand the TCP Window management
- Know the different transport layer design issues

End to End Delivery

As shown in figure 1, network layer answers the end to end delivery of individual packets from a machine to another machine in different network, but does not see any relationship between those packets. It treats each as an independent entity. Further, the packet needs to be delivered to the last participating entity, i.e. a process engaged in data exchanged. But the transport layer makes sure that the entire message (not a single packets receives) is delivered to a process that is the end (last) entity participating in message exchange. So it provides process-to-process or end-to-end delivery of an entire message.

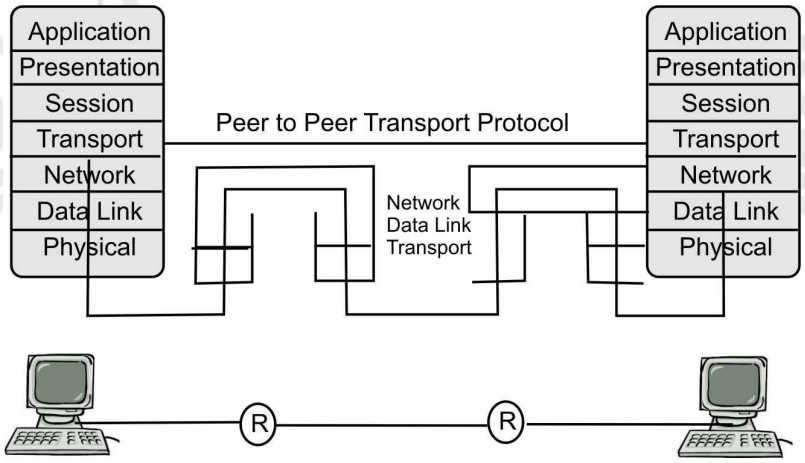


Figure 1: End to End delivery of packets

2.2 ADDRESSING

Transport layer interact with the functions of session layer. Many protocols combine session, presentation and application level protocols into a single package called an application. In these cases delivery to the session layer functions is, in effect delivery to the application. So communication occurs not just from end machine to end machine but from end application to end application. Data generated by an application on one machine must be received not just by other machines but by the correct application on that machine.

In most cases, we end up with the communication between many to many entities, called service access points as shown in figure 2 given below. To ensure accurate delivery from service access point to service access point we used another level of addressing in addition to the network and data link level.

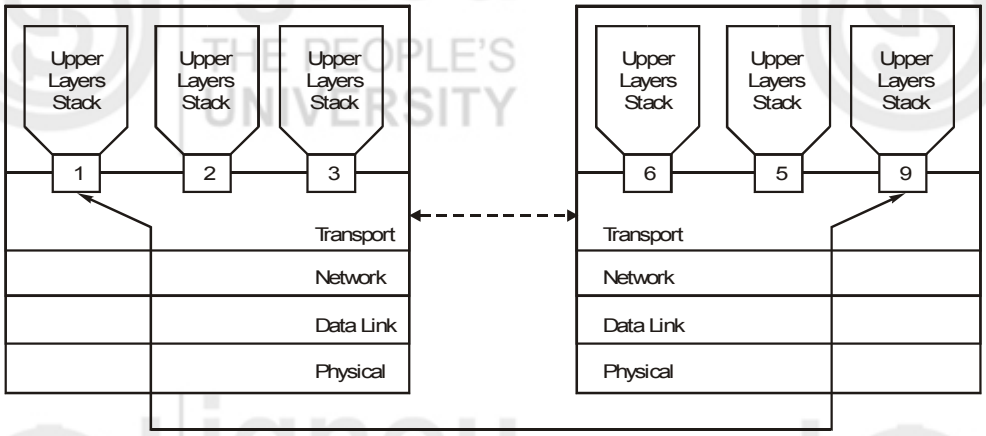


Figure 2: Service Access Points

2.3 RELIABLE DELIVERY

It is responsible for reliable delivery of data by providing the following methods also shown in figure 3:

- Error control
- Sequence control
- Loss control

- Duplication control

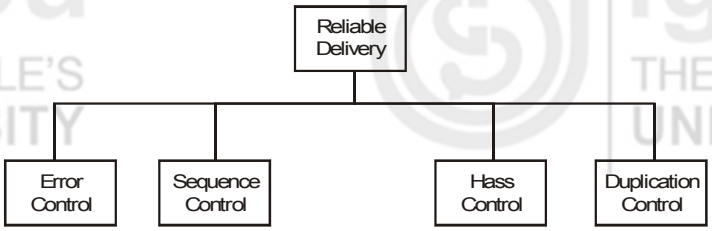


Figure 3: Methods of reliable delivery of packets

a) **Error Control:** When transferring data the primary goal of reliability if occur control. Data must be delivered to their destination. Exactly as they originated from the source. The reality of physical data transport are that while 100 per cent error free delivery is probably impossible, transport layer protocols are designed to come as close as possible.

Mechanisms full errors handling at this layer are based on error detection and retransmission. With the error handling, performed using algorithms implemented in software such as checksum "error detection and correction".

b) **Sequence Control:** Second aspect of reliability implemented at the transport layer is sequence control. On the sending end, the transport layer is responsible for ensuring that data with received from the upper layers are usable by the lower layers. ON the receiving end it is responsible for ensuring that the various pieces of a transmission are correctly reassembled.

Segmentation: When the size of the data units received from the upper layer is too long for the network layer datagram and data link layer frame to handle, the transport layer divides it into smaller usable blocks. This dividing process is called segmentation.

Concatenation: When the sizes of the data units belonging to a single session are so small that several can fit together into a single data queue are frame, the transport protocol combines them into a single data unit. This combining process is called concatenation.

Sequence Number: Most transport layer services add sequence number at the end of each segment.

If a longer data unit has been segmented the sequence number indicate the reassembly.

If several shorter units have been concatenated the numbers indicate the end of each subunit and allow them to be separated accurately at the destination.

c) **Loss Control:** The third aspect of reliability covered by the transport layer is loss control as depicted in figure 4. The transport layer ensures all pieces of the transmission arrive at the destination, not just some of them. When data have been segmented for delivery, some segments may be lost in transmit. Sequence number allows the receiver's transport layer protocol to identify any missing segment and request in delivery.

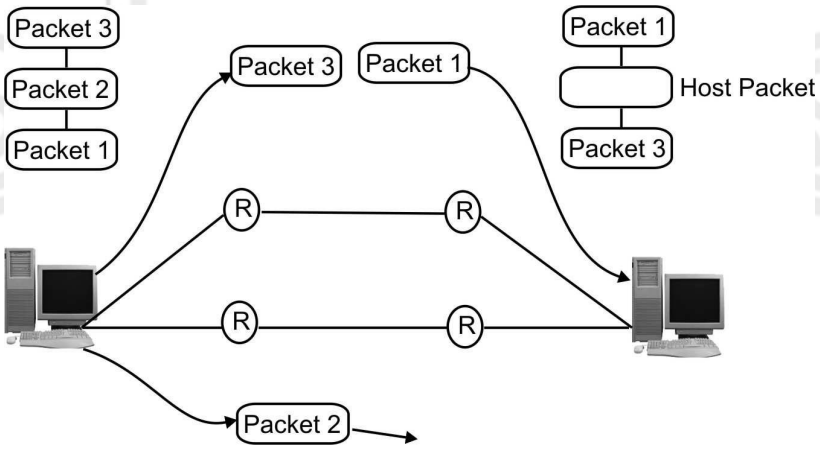


Figure 4: Loss Control

d) **Duplication Controls:** The fourth aspect of reliability by the transport layer is duplication control as shown in figure 5. Transport layer functions must guarantee that no places of data arrive at the receiving system duplicated. As they allow identification of last packets, sequence no. allows the receiver to identify and discard duplicate segments.

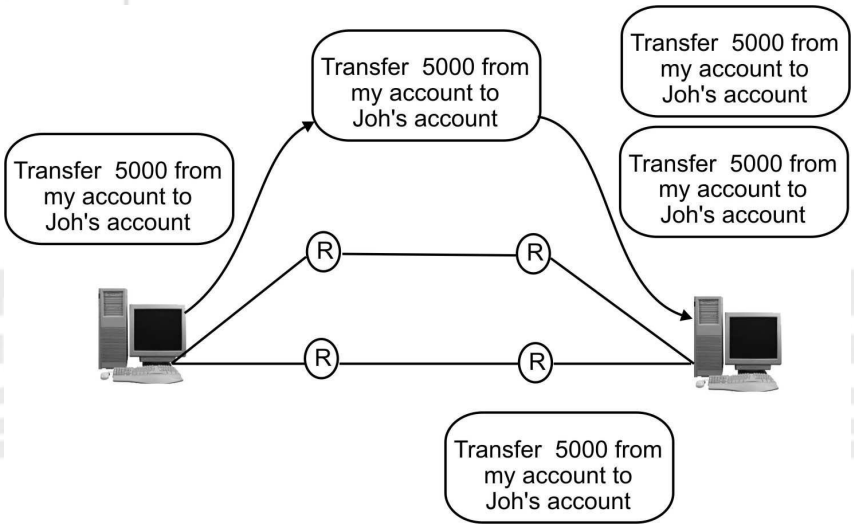


Figure 5: Duplication Control

☛ Check Your Progress 1

1. Which layer ensures the process-to-process or end-to-end delivery of an entire message? Explain.

……

……

……

2. List the methods or mechanism provided by Transport layer for reliable delivery of data.

……

……

2.4 FLOW CONTROL

Like the data link layer transport layer is responsible for flow control. Flow control is performed end to end rather than across a single link. Transport layer flow control uses a sliding window protocol. The window at the transport layer can vary in size to accommodate buffer occupancy as depicted in figure 6 given below.

Sliding window is used to make data transmission more efficient as well as to control the flow of data so that the receiver does not become overwhelmed. Sliding window used at the transport layer are usually byte oriented rather than frame oriented.

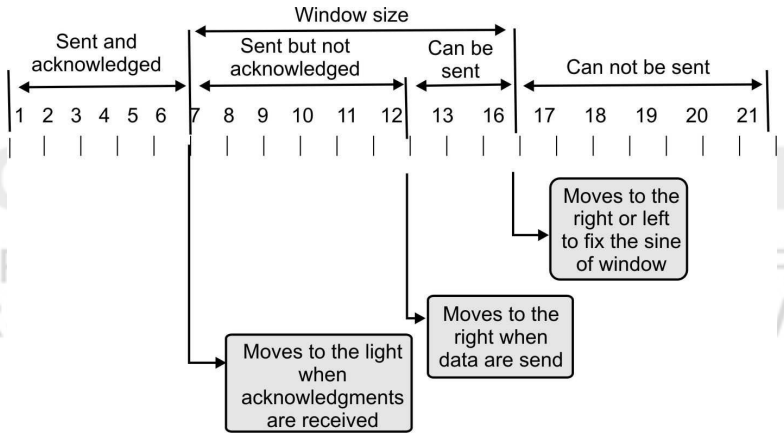


Figure 6: Sliding Window for Flow Control

Some points about sliding window at the transport layer are as follows:

1. Sender does not have to send a full window's worth of data.
2. An acknowledgement can expand the size of the window based on the sequence number of the acknowledged data segment.
3. The size of the window can be increased as decreased by the receiver.
4. The receiver can send acknowledgement at anytime.

2.5 CONNECTION MANAGEMENT

End to end delivery can be accomplished in two ways connection oriented and connectionless. The connection oriented mode is most commonly used from both two modes. A connection oriented protocol establishes a virtual circuited or pathway through the internal between sender and receiver. All of the packets belonging to a message are then sent over this same path. Using a single pathway for the entire message facilitates the acknowledgement process and retransmission of damaged and lost frames connection oriented services is generally considered reliable.

Connection Oriented transmission has three stages:

1. Connection establishment.
2. Data transfer
3. Connection termination.

Connection Establishment: Before communicating device can send data to the other, the initializing device must first determine the availability of the other to exchange data and a pathway must be found through the network by which the data can be sent.

This step is called connection establishment. Connection establishment requires three actions called three way handshake as shown in figure 7 given below.

- The computer requesting the connection sends a connection request packet to the intended receiver.
- The responding computer returns a confirmation packet to the requesting computer.
- The requesting computing returns a packet acknowledging the confirmation.

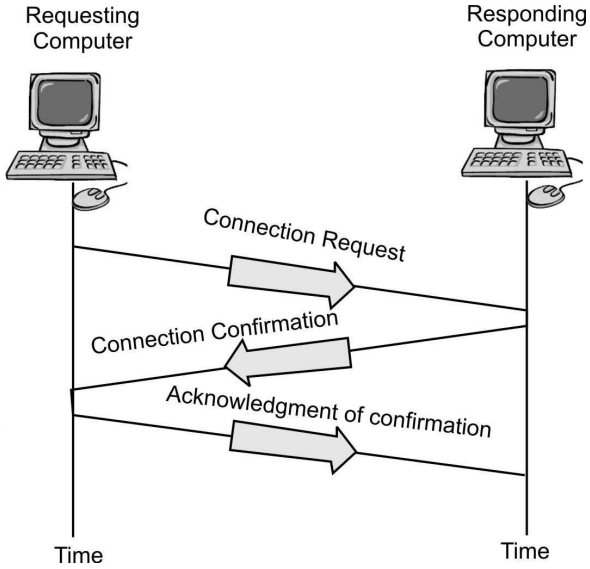


Figure 7: Connection Establishment

Connection Termination: Once all of the data have been transferred, the connection must be terminated.

Connection termination also requires three way handshake as shown in figure 8:

- Requesting computer sends a disconnection request packet.
- Responding computers confirms the disconnection request.
- The requesting computer acknowledges the confirmation.

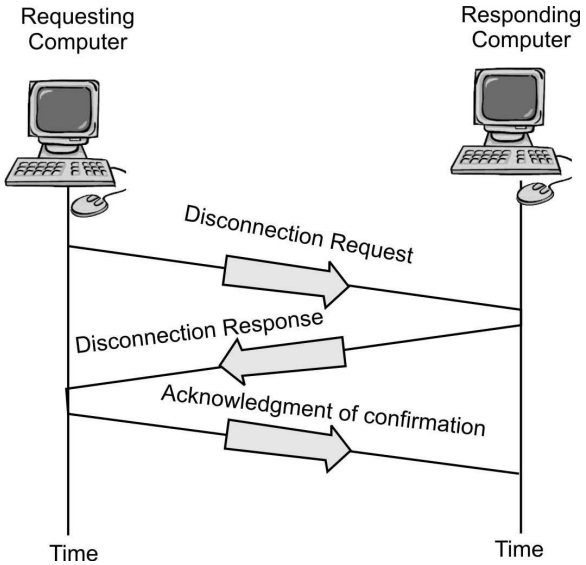


Figure 8: Connection Termination

2.6 MULTIPLEXING

To improve transmission efficiency, the transport layer has the option of multiplexing. Multiplexing at this layer occurs in two ways as shown in figure 9:

1. **Upward** -meaning that multiple transport layer connections use the same network connection

2. **Downward**- (meaning that one transport layer connection uses multiple network connections.

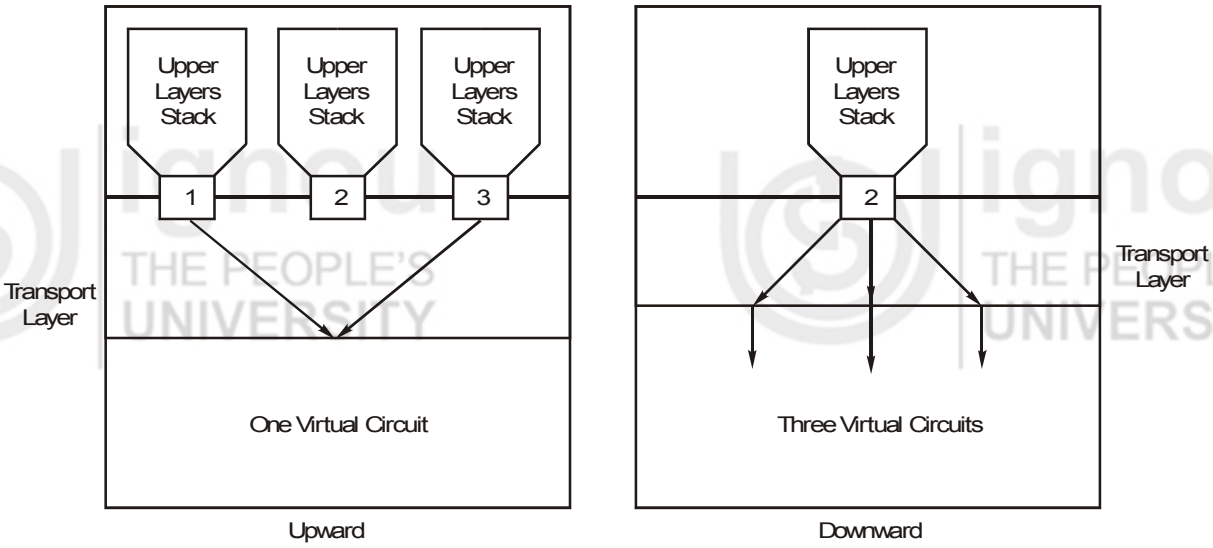


Figure 9: Upward and downward Multiplexing

Upward Multiplexing: Transport layer uses virtual circuits based on the services of the lower three layers. The underlying network charge for each virtual circuit connection. To make well cost-effective use of an established circuit, the transport layer sends several transmissions based for the same destination along the same path by upward multiplexing.

Downward Multiplexing: It allows the transport layer to split a single connection among several different paths to improve throughput (speed of delivery) as shown in figure 9. This option is useful when the underlying networks have low or slow capacity. For example, some network layer protocols have restriction on the sequence number that can be handled .X.25 uses a three bit numbering code, so sequence number are restricted to the range of 0 to 7. In this case throughput can be unacceptably low. To counteract this problem the transport layer can use more than one virtual circuit at the network layer to improve throughput by sending several data segment at once delivery is faster.

2.7 CONGESTION CONTROL

Congestion control refers to techniques and mechanisms that can either prevent congestion, before it happens, or remove congestion, after it has happened. Congestion control mechanism is divided into two broad categories:

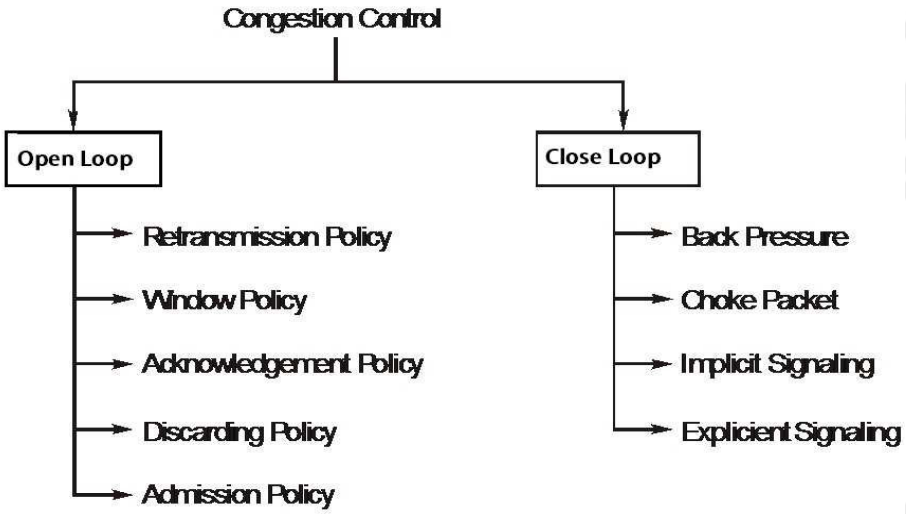


Figure 10:

Open Loop: Polices are applied to prevent congestion before it happens.

Closed Loop: Mechanism try to alleviate congestions after it happens.

Congestion Control Policy

TCP's general policy for handling congestion is based on 3 phases:

- Slow start,
- Congestion avoidance, and
- Congestion detection

The sender starts with a very slow rate of transmission but increases the rate rapidly to reach a threshold. When threshold is reached, the data rate is reduced to avoid congestion.

Congestion Avoidance: Additive Increase

In the congestion avoidance algorithm, the size of the congestion window increases additively until congestion is detected.

Congestion Detection: Multiplicative Decrease

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

1. If detection is by time-out, a new slow-start phase starts.
2. If detection is by three (acknowledgements) Acks, a new congestion Avoidance phase starts.

Slow Start: Exponential Increase

At the beginning of the connection, set the congestion window size to the maximum segment size. For each segment that is acknowledged, increase the size of the congestion window by one maximum size until you reach a threshold of half the allowable window size. This is some timer called slow start which is totally misleading because the process is not slow at all.

The size of the congestion window increase exponentially. The sender sends one segment receives one acknowledgement, increases the size to two segments, sends two segments, receives ack for two segments; increase the size to four and so on.

In other words, after receiving the 3rd ACK, the size of the window has been increased to eight segments (i.e. $2^3 = 8$). To avoid congestion before it happen one must slow down this exponential growth. After the size reaches the threshold, the size is increased one segment for each acknowledgement even if as ACK is for several segment.

Multiplicative Decrease

If the congestion occurs the congestion windows size must be decreased. The only way the sender causes that connection has occur through a lost segment. If the sender does not receive an acknowledgement for a segment before it transmission timer matured, it assumes that there is congestion.

The strategy says if a time out occurs, the threshold must be set to half of the congestion window size and the congestion or size should start from one again.

2.8 QUALITY OF SERVICES (QOS)

A stream of packet, from a source to a destination is called a flow. In connection oriented network, all the packets belonging to a flow the same route. But in connection less network they may follow different routes. The need of each flow can be characterized by four primary parameters:

1. Reliability: Reliability is the ability of a system to perform and maintain its functions in normal conditions as well as under unexpected conditions.

2. Delay is defined as the time interval elapsed between the departures of data from the source to its arrival at the destination.

3. Jitter**:** Jitter refers to the variation in time between packets arriving at the destination.

4. Bandwidth refers to the data rate supported by a network connection or interface.

Techniques to Improve QOS

1. **Over Provisioning:** An easy solution is to provide so much router capacity, buffer space and bandwidth that the packets just fly through costly.

2. **Buffering:** Flows can be buffered on the receiving side before being delivered. Buffering those does not affected the reliability of bandwidth and increases the delay but it smoothes out the jitter.

3. **Scheduling:** Packets from different flows arrive at a switch or router for processing. A good scheduling technique treats the different flows in a fair and appropriate manner.

   Here we discuss some scheduling techniques to improve the quality of service such as:

   i) FIFO Queuing

   ii) Priority Queuing

   iii) Weighted Fair Queuing

FIFO Queuing

In FIFO Queuing packets wait in a buffer (Queue) until the node is ready to process them. If the average rate is higher than the average processing rate, the queue will fill up and new packets will be discarded.

Priority Queuing

In this packets are first assigned to **priority class**. Each priority class has its own Queue. The packets in the highest priority Queue are processed first. But if there is a continuous flow in high-priority Queue, the packets in the low priority Queues will never have a chance to be processes.

Weighted Fair Queuing

In this method, the packets are still assigned to different classes and admitted to different queues. The queues, however, are weighted based on the priority of the queues; higher priority means a higher weight.

The system processes packets in each queue in a round-robin fashion with the number of packets selected from each queue based on the corresponding weight.

For example: If the weights are 3, 2 and 1, three packets are processed from the first queue, two from the second queue and one from the third queue.

If the system does not impose priority on the classes, all weights can be equal. In thus way, we have fair queuing with priority.

4. Traffic Shaping
   a) Leaky bucket
   b) Token bucket

2.9 TCP WINDOW MANAGEMENT

TCP uses two buffers and one window, to control the flow of data coming from the sending application program. The application program creates data and writes it to the buffer. The sender imposes a window on this buffer and sends the segments as long as the size of the window is not zero. The TCP receiver has buffer also. It receives data, checks them, and stores them in buffer to be consumed by the receiving application program.

A sliding window is used to make transmission more efficient as well as to control the flow of data so that the destination does not become over whelmed with data. TCP's sliding window is byte oriented

TCP's Sliding Window

The source does not have to send a full window's worth of data. The size of the window can be increased or decreased by the destination. The destination can send an acknowledgement at any time.

Silly Window Syndrome

Silly Window Syndrome is another problem that can degrade the TCP performance. This problem occurs when the sender transmits data in large blocks, but an interactive application on the receiver side reads data 1 byte at a time as shown in figure 11.

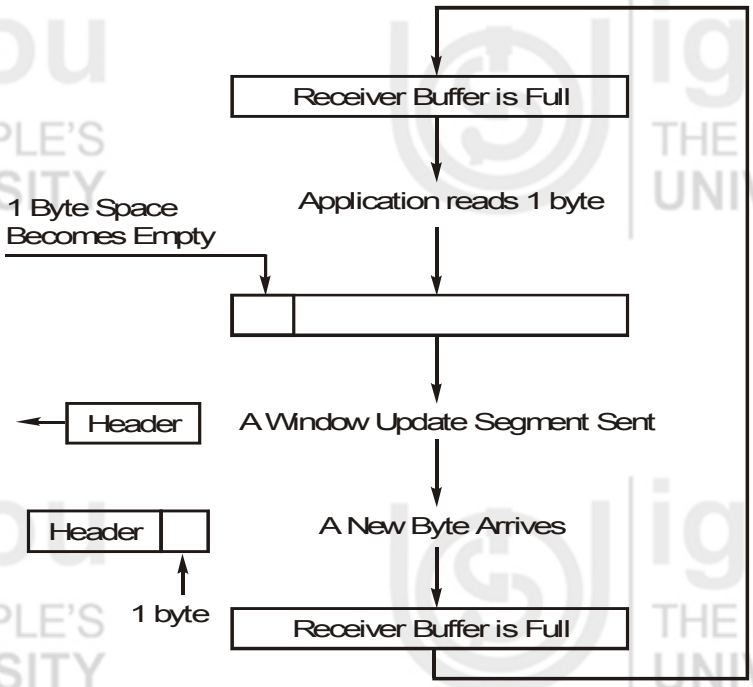


Figure 11: Silly Window Syndrome

1. Initially the receiver's buffer is full so it sends a window size 0 to block the sender.

2. But the interactive application reads one byte from the buffer, so one byte space becomes empty.

3. The receiving TCP sends a window update to the sender informing that it can send 1 byte.

4. The sender sends 1-new byte.

5. The buffer is full again and the window size is 0. The behavior can continue forever's.

Solution to Silly Window Syndrome

A solution to silly window syndrome was suggested as follows:
It was suggested that the **receiver should not send** a window update for 1 byte. Instead it must wait until it has a substantial amount of buffer space available and then sends the window update. To be specific, the receiver should not send a window update until it can handle the maximum window size, it has advertised at the time of establishing a connection or its buffer is half empty, whichever is smaller.

The sender can also help to improve the situation. It should not send tiny segments instead it must wait and send a full segment or at least one containing half of the receivers buffers size.

2.10 PORTS

In computer networking of connection-based communication port is like a medium through which, an application establish a connection with another application by binding a socket by a port number. Addressing the information and the port number, accompanied the data transfer over the network.

The Ports are used by TCP and UDP to deliver the data to the right application, are identified by a 16-bit number present in the header of a data packet. Ports are typically used to map data to a particular process running on a client. If we consider a letter (data packet) sent to a particular apartment (IP) with house no. (Port no), at this time the port no. is the most important part for the delivery of the letter. In order for the delivery to work, the sender needs to include a house number along with the address to ensure the letter gets to the right destination.

Do you know?

TCP and UDP ports are 16 bit number

The TCP and UDP protocols use ports to map incoming data to a particular process running on a computer.

- Port is represented by a positive (16-bit) integer value
- Some ports have been reserved to support common/well known services:
  - FTP 21/TCP
  - TELNET 23/TCP
  - SMTP 25/TCP
  - LOGIN 513/TCP
- User level process/services generally use port number value >= 1024

Types of Port

1. Well known port (0 to 1023)-They are controlled by IANA
2. Registered Port (1024-49159)
3. Dynamic port (49152-65535)

If we consider the client-server architecture, a server application binds a socket to a specific port number in connection-based communication. It registered the server with the system where all the data destined for that port.

Do you know?

Port number permits unique identification of several simultaneous processes using TCP/UDP

Now we are aware of the importance of the port number. In the same order there are some ports which are predefine and called reserved ports. Some of them are given in Table 1given below:

Table 1: Reserved Port Numbers.

Service	Port no.
ECHO	7
DAYTIME	13
FTP	21
TELNET	23
SMTP	25
FINGER	79
HTTP	80
POP3	110

Do you know?

If we consider the range of the port numbers, there are 0 to 65,535 ports available.

Tips

The port numbers ranging from 0 - 1023 are reserved ports or we can say that are restricted ports. All the 0 to 1023 ports are reserved for use by well-known services such as FTP, telnet and http and other system services. These ports are called well-known ports.

☞ Check Your Progress 2

1. List three stages of Connection Oriented transmission.

   ……

   ……

   ……

   ……

2. Compare and contrast between Upward and Downward Multiplexing.

   ……

   ……

   ……

   ……

2.11 SUMMARY

Transport layer is mainly responsible for end to end reliable delivery, segmentation and concatenation. Two main protocols that operate on transport layer are TCP and UDP.TCP provides reliable connection oriented service while UDP provides unreliable connectionless service. The data link and transport layer perform many of the same duties .The data link layer function in a single network, while the transport layer operates across an internet. Flow control at the transport layer is handles by three walled sliding window. Multiplexing can be downward of upward in transport layer. Connection establishment and termination can be done by using three way handshakes. Transport layer works on port address. To know more about the Transport layer and its protocols (TCP and UDP), student may please refer to the course material of BCS-61TCP/IP Programming or BCS-54 Network Programming.

2.12 REFERENCES/FURTHER READING

1. Computer Networks, A. S. Tanenbaum 4th Edition, Practice Hall of India, New Delhi. 2003.

2. Introduction to Data Communication & Networking, 3rd Edition, Behrouz Forouzan, Tata McGraw Hill.

3. Douglas E. Comer, *Internetworking with TCP/IP Vol.1: Principles, Protocols, and Architecture* (4th Edition).

4. James F. Kurose, *Computer Networking*: A Top-Down Approach Featuring the Internet (3rd Edition).

5. Larry L. Peterson, *Computer Networks*: A Systems Approach, 3rd Edition (The Morgan Kaufmann Series in Networking).

6. W. Richard Stevens, *The Protocols (TCP/IP Illustrated, Volume 1)*.

7. www.wikipedia.org

8. William Stallings, *Data and Computer Communications*, Seventh Edition.

2.13 SOLUTION /ANSWERS

☞ Check Your Progress 1

1. Transport layer makes sure that the entire message (not a single packets receives) is delivered to a process that is the end (last) entity participating in message exchange. So it provides process-to-process or end-to-end delivery of an entire message.

2. Transport Layer is responsible for reliable delivery of data, it provide following methods to provide reliable delivery of data.
   - Error control
   - Sequence control
   - Loss control
   - Duplication control

☞ Check Your Progress 2

1. Connection Oriented transmission has following three stages:
   i) Connection establishment.
   ii) Data transfer
   iii) Connection termination.

2. **Upward Multiplexing:** Transport layer uses virtual circuits based on the services of the lower three layers. The underlying network charge for each virtual circuit connection. To make well cost-effective use of an established circuit, the transport layer sends several transmissions based for the same destination along the same path by upward multiplexing. **Downward Multiplexing:** Allows the transport layer to split a single connection among several different path to improve throughput (speed of delivery). This option is useful when the underlying networks have low or slow capacity. For example, some network layer protocols have restriction on the sequence number that can be handled .X.25 uses a three bit numbering code, so sequence number are restricted to the range of 0 to 7. In this case throughput can be unacceptably low. To counteract this problem, the transport layer can use more than one virtual circuit at the network layer to improve throughput by sending several data segment at once delivery is faster.

UNIT 3 APPLICATION LAYER

Structure		Page No.
3.0	Introduction	48
3.1	Objectives	49
3.2	Client Server Architecture	49
3.3	Domain Name Server (DNS)	49
	3.3.1 The DNS name space	
	3.3.2 Resource records	
	3.3.3 Name servers	
	3.3.4 Remote login (telnet)	
	3.3.5 The telnet application	
	3.3.6 The telnet protocol	
3.4	Remote Login (Telnet)	54
	3.4.1 The FTP application	
	3.4.2 The FTP protocol	
3.5	File Transfer Protocol (FTP)	55
3.6	Network Management	56
	3.6.1 Configuration management	
	3.6.2 Reconfiguration	
	3.6.3 Documentation	
	3.6.4 Fault management	
	3.6.5 Reactive fault management	
	3.6.6 Proactive fault management	
	3.6.7 Performance management	
	3.6.8 Security management	
	3.6.9 Accounting management	
	3.6.10 SNMP protocol	
3.7	Word Wide Web and Client Server Applications	61
	3.7.1 Architectural overview (www)	
3.8	Electronic Mail	70
	3.8.1 Architecture and services	
	3.8.2 The user agent	
3.9	Summary	73
3.10	References/Further Readings	73
3.11	Solutions / Answers	73

3.0 INTRODUCTION

The application layer contains a variety of protocols that are commonly needed by users. One widely-used application protocol is HTTP (Hyper Text Transfer Protocol), which is the basis for the World Wide Web. When a browser wants a Web page, it sends the name of the page it wants to the server using HTTP. The server then sends the page back. Other application protocols are used for file transfer, electronic mail, and network news. The application layer enables the user, whether human or software, to access the network. It provides user interfaces and support for services such as electronic mail, remote file access and transfer, shared database management, and other types of distributed information services. Application-layer functions typically include identifying communication partners, determining resource availability, and synchronizing communication. When identifying communication partners, the application layer determines the identity and availability of communication partners for an application that has data to transmit. When determining resource availability, the application layer must decide whether sufficient network or the requested communication exists. This unit covers the details of all important functions and

services of Application layer including different protocols available in Application layer.

3.1 OBJECTIVES

After going through this unit, you will be able to:

- define and know the functions of Application layer
- Understand the working of Domain Name Server
- Know the features and services of Telnet and FTP
- Understand the Network Management issues
- Understanding of the WWW Client/Server architecture
- Know the Email Architecture and Services

3.2 CLIENT SERVER ARCHITECTURE

Client/server architecture means one or more computers is/are acting as client one computer is behaving as serve which preferably of high configuration. In simple words client is defined as requester of services and server is defined as provider of services. Client and Servers are nothing but programs, which generally run on different machines/computers.

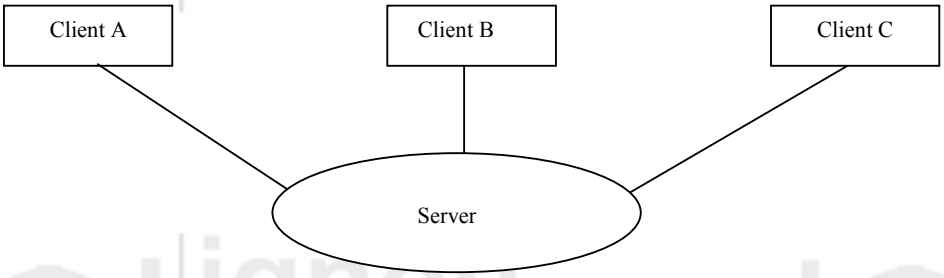


Figure 1: Client/Server Architecture

When client wants some information or wants to perform some task by server it has to contact the server as shown in figure1. Therefore, the client initially needs the address of the server, and tries to contact the server using that address. Server is a program which is always running and waiting to serve clients. Server welcomes the client (if free) as the client contacts it. After contacting the server, the server welcomes that client (if free), the client informs about its own address to servers so that server can reply the client also.

When both clients and servers exchange important information to each other connection establishes. Once with connection or link established, both client and server could exchange any information with each other. We will read more details of client/server architecture further in this unit during discussion about Word Wide Web.

3.3 DOMAIN NAME SERVER (DNS)

Programs theoretically could refer to hosts, mailboxes, and other resources by their network (e.g., IP) addresses, these addresses are hard for people to remember. Also, sending e-mail to *ravi@128.111.45.46* means that if Ravi's ISP or organization moves the mail server to a different machine with a different IP address, his e-mail address has to change. Consequently, ASCII names were introduced to decouple machine names from machine addresses. In this way, Ravi's address might be something like ravi@art.ucsb.edu. Nevertheless, the network itself understands only numerical addresses, so some mechanism is required to convert the ASCII strings to network addresses.

Way back in the ARPANET, there was simply a file, hosts.txt that listed all the hosts and their IP addresses. Every night, all the hosts would fetch it from the site at which it was maintained. For a network of a few hundred large timesharing machines, this approach worked reasonably well.

However, when thousands of minicomputers and PCs were connected to the net, everyone realized that this approach could not continue to work forever. For one thing, the size of the file would become too large. However, even more important, host name conflicts would occur constantly unless names were centrally managed, something unthinkable in a huge international network due to the load and latency. To solve these problems, **DNS** (the **Domain Name System**) was invented.

The essence of DNS is the invention of a hierarchical, domain-based naming scheme and a distributed database system for implementing this naming scheme. It is primarily used for mapping host names and e-mail destinations to IP addresses but can also be used for other purposes. Very briefly, the way DNS is used as follows. To map a name onto an IP address, an application program calls a library procedure called the **resolver**, passing it the name as a parameter. The resolver sends a UDP packet to a local DNS server, which then looks up the name and returns the IP address to the resolver, which then returns it to the caller. Armed with the IP address, the program can then establish a TCP connection with the destination or send it UDP packets.

3.3.1 The DNS Name Space

Managing a large and constantly changing set of names is a nontrivial problem. In the postal system, name management is done by requiring letters to specify (implicitly or explicitly) the country, state or province, city, and street address of the addressee. Conceptually, the Internet is divided into over 200 top-level **domains**, where each domain covers many hosts. Each domain is partitioned into sub domains, and these are further partitioned, and so on. All these domains can be represented by a tree, as shown in Figure 2. The leaves of the tree represent domains that have no sub domains (but do contain machines, of course). A leaf domain may contain a single host, or it may represent a company and contain thousands of hosts.

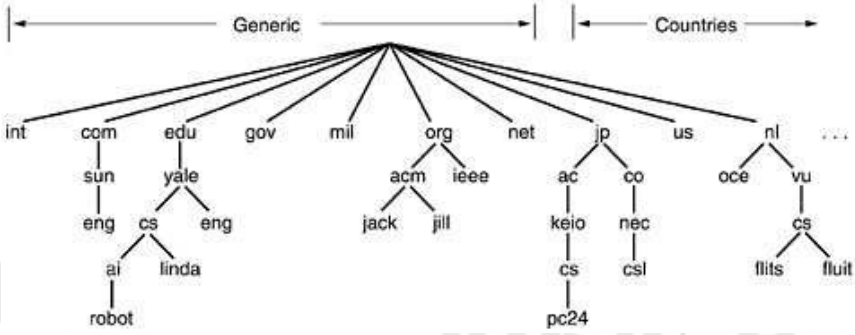


Figure 2: A portion of the Internet domain name space

The top-level domains come in two flavors: **generic** and **countries**. The original generic domains were com (commercial), edu (educational institutions), gov (the Government), int (certain international organizations), mil (the U.S. armed forces), net (network providers), and org (nonprofit organizations). The country domains include one entry for every country, as defined in ISO 3166.

In November 2000, ICANN approved four new, general-purpose, top-level domains, namely, biz (businesses), info (information), name (people's names), and pro (professions, such as doctors and lawyers). In addition, three more specialized top-level domains were introduced at the request of certain industries. These are aero (aerospace industry), coop (co-operatives), and museum (museums). Other top-level domains will be added in the future.

As an aside, as the Internet becomes more commercial, it also becomes more contentious. Take pro, for example. It was intended for certified professionals.

In general, getting a second-level domain, such as name-of-company.com, is easy. It merely requires going to a registrar for the corresponding top-level domain (com in this case) to check if the desired name is available and not somebody else's trademark. If there are no problems, the requester pays a small annual fee and gets the name. By now, virtually every common (English) word has been taken in the com domain. Try household articles, animals, plants, body parts, etc. Nearly all are taken.

Each domain is named by the path upward from it to the (unnamed) root. The components are separated by periods (pronounced "dot"). Domain names can be either absolute or relative. An absolute domain name always ends with a period (e.g., eng.sun.com.), whereas a relative one does not. Relative names have to be interpreted in some context to uniquely determine their true meaning. In both cases, a named domain refers to a specific node in the tree and all the nodes under it.

Domain names are case insensitive, so edu, Edu, and EDU mean the same thing. Component names can be up to 63 characters long, and full path names must not exceed 255 characters.

In principle, domains can be inserted into the tree in two different ways. For example, cs.yale.edu could equally well be listed under the US country domain as cs.yale.ct.us. In practice, however, most organizations in the United States are under a generic domain, and most outside the United States are under the domain of their country. There is no rule against registering under two top-level domains, but few organizations except multinationals do it (e.g., sony.com and sony.nl).

Each domain controls how it allocates the domains under it. For example, Japan has domains ac.jp and co.jp that mirror edu and com. The Netherlands does not make this distinction and puts all organizations directly under nl. Thus, all three of the following are university computer science departments:

1. cs.yale.edu(Yale University, in the United States)

2. cs.vu.nl(Vrije Universiteit, in The Netherlands)

3. cs.keio.ac.jp(Keio University, in Japan)

To create a new domain, permission is required of the domain in which it will be included. For example, if a VLSI group is started at Yale and wants to be known as vlsi.cs.yale.edu, it has to get permission from whoever manages cs.yale.edu.

Similarly, if a new university is chartered, say, the University of Northern South Dakota, it must ask the manager of the edu domain to assign it unsd.edu. In this way, name conflicts are avoided and each domain can keep track of all its subdomains. Once a new domain has been created and registered, it can create subdomains, such as cs.unsd.edu, without getting permission from anybody higher up the tree.

Naming follows organizational boundaries, not physical networks. For example, if the computer science and electrical engineering departments are located in the same building and share the same LAN, they can nevertheless have distinct domains.

3.3.2 Resource Records

Every domain, whether it is a single host or a top-level domain, can have a set of **resource records** associated with it. For a single host, the most common resource record is just its IP address, but many other kinds of resource records also exist. When a resolver gives a domain name to DNS, what it gets back are the resource records associated with that name. Thus, the primary function of DNS is to map domain names onto resource records.

A resource record is a *five-tuple*. Although they are encoded in binary for efficiency, in most expositions, resource records are presented as ASCII text, one line per resource record. The format we will use is as follows:

- **Domain_name, Time_to_live, Class Type, Value**

The Domain_name tells the domain to which this record applies. Normally, many records exist for each domain and each copy of the database holds information about multiple domains. This field is thus the primary search key used to satisfy queries. The order of the records in the database is not significant.

The Time_to_live field gives an indication of how stable the record is. Information that is highly stable is assigned a large value, such as 86400 (the number of seconds in 1 day). Information that is highly volatile is assigned a small value, such as 60 (1 minute).

The third field of every resource record is the Class. For Internet information, it is always IN. For non-Internet information, other codes can be used, but in practice, these are rarely seen.

The Type field tells what kind of record this is. The most important types are listed in Table. 1.2.2.

Table 1: The principal DNS Resource Record

Type	Meaning	Value
SOA	Start of Authority	Parameters for this zone
A	IP address of a host	32-Bit integer
MX	Mail exchange	Priority, domain willing to accept e-mail
NS	Name Server	Name of a server for this domain
CNAME	Canonical name	Domain name
PTR	Pointer	Alias for an IP address
HINFO	Host description	CPU and OS in ASCII
TXT	Text	Uninterpreted ASCII text

3.3.3 Name Servers

In theory at least, a single name server could contain the entire DNS database and respond to all queries about it. In practice, this server would be so overloaded as to be useless. Furthermore, if it ever went down, the entire Internet would be crippled.

To avoid the problems associated with having only a single source of information, the DNS name space is divided into non-overlapping **zones**. One possible way to divide

the name space is shown in figure 3. Each zone contains some part of the tree and also contains name servers holding the information about that zone.

Normally, a zone will have one primary name server, which gets its information from a file on its disk, and one or more secondary name servers, which get their information from the primary name server. To improve reliability, some servers for a zone can be located outside the zone.

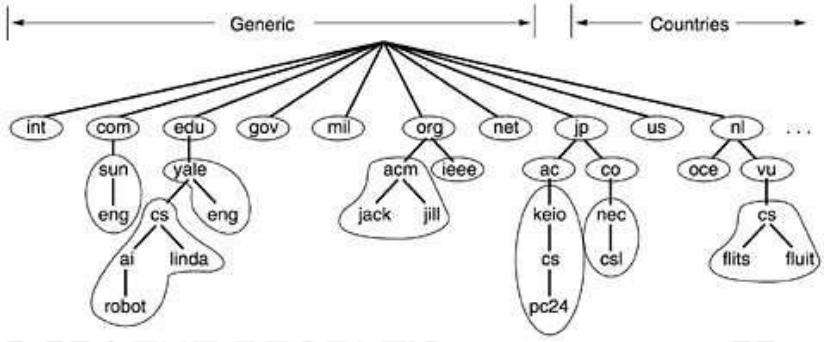


Figure 3: Part of the DNS name space showing the division into zones.

Where the zone boundaries are placed within a zone is up to that zone's administrator. This decision is made in large part based on how many name servers are desired, and where.

☞ Check Your Progress 1

1. Select the right choice.

   a) Which Layer is not present in TCP/IP model?
   (A) Application Layer (B) Internet Layer
   (C) Transport Layer (D) Presentation Layer

   ……

   ……

   b) Let most segment of a name inn DNS represents
   (A) Individual Network. (B) Individual computer.
   (C) Domain name (D) Network type.

   ……

   ……

   c) Address 192.5.48.3 belongs to
   (A) Class A. (B) Class B.
   (C) Class C. (D) Class D.

   ……

   ……

2. Discuss the DNS Name Space?

   ……

   ……

3. Write a note on Name Server?

3.4 REMOTE LOGIN (TELNET)

Telnet permits a user to connect to an account on a remote machine. A client program running on the user's machine communicates using the Telnet protocol with a server program running on the remote machine.

3.4.1 The Telnet Application

The user (say Ravi) has an account on both the local and remote machines. For example, 'Ravi' on farkas.mcmaster.ca types telnet optlab.cas.mcmaster.ca at his user prompt. Here, farkas.mcmaster.ca is the client and optlab.cas.mcmaster.ca is the server. The Telnet client would perform a methotrexate() call to determine the IP address of optlab.cas.mcmaster.ca. Then the client would create a socket to communicate with the Telnet server. The server prompts 'Ravi' for a login identifier - the name of the user's account on the remote server followed by a password. The Telnet users interact with the remote machine in the same way they would interact with their local machine. The client relays Ravi's keystrokes to the remote server, and the remote server displays them on its pseudo terminal (which is actually the display screen on the client machine).

3.4.2 The Telnet Protocol

The Telnet client program performs two important functions - **interacting** with the user terminal on the local host and **exchanging messages** with the Telnet server. The client connects to port 23 on the remote machine, which is the port number reserved for Telnet servers. The TCP connection persists for the duration of the login session. The client and the server maintain the connection, even when the user interrupts the transfer of data, for example by hitting *cntl-C*. Since Telnet is designed to work over two hosts on different platforms, the protocol assumes that the two hosts run a Network Virtual Terminal (NVT).The TCP connection is set up across these two NVT terminals. The NVT is a very simple character device with a keyboard and a printer - data typed by the user on the keyboard is translated by the client software into NVT format and sent via its NVT terminal to the server, and data received in NVT format from the server is translated by the client into the local machine format and output to the printer.

The NVT terminals on the two hosts exchange data in the 7-bit U.S. variant of the ASCII format, with each character sent as an octet with the first bit set to19-21-10. Some control information, such as end-of-line indication, is transmitted as the character sequence CR (carriage return) followed by an LF (linefeed). Each Telnet control message starts with the special octet (Interpret as Command (IAC)) octet of all 1s) to ensure that the recipient interprets the subsequent octets as a command. Otherwise, each octet is interpreted as data (e.g., a user keystroke). Sending control messages on the same connection as the data is referred to as *inband signaling*. The initial control messages between the client and the server are used to exchange information about their capabilities (Telnet option negotiation). For example, the client may indicate the type and speed of its terminal, and whether data is to be sent one character or one line at a time. After the capabilities exchange, the server instructs the client to send a login identifier and password. Once the authentication completes, the user interacts directly with the remote machine. The client application relays user keystrokes to the server, and the server relays the output back to the client, using inband signaling, with the interpretation that commands follow the IAC octet of all ones. Telnet cannot rely on the conventional data stream alone to carry such control sequences between client and server.

3.5 FILE TRANSFER PROTOCAL (FTP)

FTP allows a user to copy files to and from a remote machine. The client program also sends commands to the server program to coordinate the copying of files between the two machines on behalf of the user.

3.5.1 The FTP Application

The FTP client connects to the remote machine which prompts the user to enter a login identifier and a password. However, some users may not have their own accounts on the remote machine. To grant access to a broad set of users, many FTP servers have a special account (e.g. *anonymous*) that does not require password information. Instead, the user logs in using *guest* or his email address as password. The FTP server coordinates access to a collection of files in various directories. In case of anonymous FTP, the server typically has a special directory, with one or more subdirectories, that can be accessed by the client. The user logged into the FTP server can traverse through the directories of files on the remote machine, and send or receive files. This is typically done via the command-line interface. The interface may also allow the client to send or receive multiple files with a single command. Recent FTP client applications provide a menu-based graphical user interface. For instance, a Web browser allows users to specify the desired file as an URL (e.g.,ftp://ftp.optlab.cas.mcmaster.ca/midterm.pdf). In this case, the web browser connects to the FTP server as an anonymous user and sends a sequence of FTP commands to fetch the requested file.

3.5.2 The FTP protocol

FTP differs from other applications such as Telnet since it uses separate TCP connections for control and data. Recall that in Telnet both control information and data are sent over the same TCP connection using in-band signaling. The two TCP connections in FTP are:

1. **The control connection** is established in the normal client-server fashion. In this case, the server does a *passive open* (is listening) on port 21 for FTP, and waits for the client connection. The client does an *active open* (the2nd handshake in a TCP connection) to establish the control connection. The client uses an ephemeral port number (above 1023) for the control connection. This control connection stays up for the entire time that the client communicates with this server. This connection is used for commands from the client to the server and for the server's replies. The IP type of service for the control connection should be to minimize delay in passing these commands over the TCP connection.

2. **A data connection** is created each time a file is transferred between the client and the server. The IP type of service for the data connection should be to maximize throughput since this connection is file transfer, and we want to send this entire file over a high bandwidth line. The specification of FTP includes more than 30 different commands, which are transmitted over the control connection in NVT ASCII format. The commands are not case-sensitive and may have arguments; each command ends with a two character sequence of a carriage return (CR) followed by a line feed (LF). It must be emphasized here that these commands are different from the commands typed by the user at the interface provided by the client. Transferring a single file for instance requires only a single user-level command (e.g., *put* or *get*), but this single command triggers the client to send a set of FTP commands to the server. The FTP server responds to each command with a three-digit reply code (for the FTP client) and an optional text message (for the user).

The control connection persists over a sequence of FTP commands, as the client and the server continue their dialogue. The typical interaction starts with a command that identifies the user on the server machine followed by another command to send the user password. The arguments for these commands are gleaned from the user's input (his account name and password). The server uses this information to verify whether the user has an authorized account on the remote machine, and in the case of anonymous FTP decides on the set of19-21-3directories to which the anonymous guest has access. The next set of commands depends on the user request to send, receive, or view the files in a present directory.

The actual file (data) transfer uses a separate TCP connection established by the host sending the data. For instance, if the user wants to retrieve the file *midterm.pdf* from the remote server, the server initiates the creation of the TCP data connection. In case, the user wants to put a file into the remote machine, it is the client who initiates the creation of the TCP connection. The data connection is usually established on port 20 on the server machine. In the former case (when the file is to be retrieved from the server), the server does not know the destination port for the FTP client. So before sending the command to retrieve the file, the client instructs its operating system to allocate a port number (above 1023) for such a transaction. This information is given to the server via the control connection. The data connection is created (using the usual TCP 3 way handshake), and the server writes the contents of the file, and closes the connection. The client reads the bytes from its socket upto the end of file (EOF) character. Also, unlike Telnet, FTP does not require the data transfer to 7 bit ASCII characters (NVT format); it actually permits a wide range of data types including binary files. The client requests the form of data transfer using the control connection. In practice, each data transfer requires a separate TCP connection. In contrast, the control connection can persist across multiple data transfers.

3.6 NETWORK MANAGEMENT

We can define **network management** as monitoring, testing, configuring, and troubleshooting network components to meet a set of requirements defined by an organization. These requirements include the smooth, efficient operation of the network that provides the predefined quality of service for users. To accomplish this task, a network management system uses hardware, software, and humans. In this chapter, first we briefly discuss the functions of a network management system. Then we concentrate on the most common management system, the Simple Network Management Protocol (SNMP).

We can say that the functions performed by a network management system can be divided into **five broad categories**: configuration management, fault management, performance management, security management, and accounting management.

3.6.1 Configuration Management

A large network is usually made up of hundreds of entities that are physically or logically connected to one another. These entities have an initial configuration when the network is set up, but can change with time. Desktop computers may be replaced by others; application software may be updated to a newer version; and users may move from one group to another. The **configuration management** system must know, at any time, the status of each entity and its relation to other entities. Configuration management can be divided into two subsystems: reconfiguration and documentation.

3.6.2 Reconfiguration

Reconfiguration, which means adjusting the network components and features, can be a daily occurrence in a large network. There are three types of reconfiguration: hardware reconfiguration, software reconfiguration, and user-account reconfiguration. Hardware reconfiguration covers all changes to the hardware. For example, a desktop computer may need to be replaced. A router may need to be moved to another part of the network. A sub network may be added or removed from the network. All these need the time and attention of network management. In a large network, there must be specialized personnel trained for quick and efficient hardware reconfiguration. Unfortunately, this type of reconfiguration cannot be automated and must be manually handled case by case.

Software reconfiguration covers all changes to the software. For example, new software may need to be installed on servers or clients. An operating system may need updating. Fortunately, most software reconfiguration can be automated. For example, updating an application on some or all clients can be electronically downloaded from the server.

User-account reconfiguration is not simply adding or deleting users on a system. You must also consider the user privileges, both as an individual and as a member of a group. For example, a user may have read and write permission with regard to some files, but only read permission with regard to other files. User-account reconfiguration can be, to some extent, automated. For example, in a college or university, at the beginning of each quarter or semester, new students are added to the system. The students are normally grouped according to the courses they take or the majors they pursue.

3.6.3 Documentation

The original network configuration and each subsequent change must be recorded meticulously. This means that there must be documentation for hardware, software, and user accounts. **Hardware documentation** normally involves two sets of documents: maps and specifications. Maps track each piece of hardware and its connection to the network. There can be one general map that shows the logical relationship between each sub-network. There can also be a second general map that shows the physical location of each sub-network. For each sub network, then, there are one or more maps that show all pieces of equipment. The maps use some kind of standardization to be easily read and understood by current and future personnel. Each piece of hardware also needs to be documented. There must be a set of specifications for each piece of hardware connected to the network. These specifications must include information such as hardware type, serial number, vendor (address and phone number), time of purchase, and warranty information. All software must also be documented. Software documentation includes information such as the software type, the version, the time installed, and the license agreement. Most operating systems have a utility that allows the documentation of user accounts and their privileges. The management must make sure that the files with this information are updated and secured. Some operating systems record access privileges in two documents one shows all files and access types for each user; the other shows the list of users that have a particular access to a file.

3.6.4 Fault Management

Complex networks today are made up of hundreds and sometimes thousands of components. Proper operation of the network depends on the proper operation of each component individually and in relation to each other. **Fault management** is the area of network management that handles this issue. An effective fault management system has two subsystems: reactive fault management and proactive fault management.

3.6.5 Reactive Fault Management

A reactive fault management system is responsible for detecting, isolating, correcting, and recording faults. It handles short-term solutions to faults. The first step taken by a reactive fault management system is to detect the exact location of the fault. A fault is defined as an abnormal condition in the system. When a fault occurs, either the system stops working properly or the system creates excessive errors. A good example of a fault is a damaged communication medium. This fault may interrupt communication or produce excessive errors.

The next step taken by a reactive fault management system is to isolate the fault. A fault, if isolated, usually affects only a few users. After isolation, the affected users are immediately notified and given an estimated time of correction. The third step is to correct the fault. This may involve replacing or repairing the faulty component(s).After the fault is corrected, it must be documented. The record should show the exact location of the fault, the possible cause, the action or actions taken to correct the fault, the cost, and time it took for each step. Documentation is extremely important for several reasons:

- The problem may recur. Documentation can help the present or future administrator or technician solve a similar problem.
- The frequency of the same kind of failure is an indication of a major problem in the system. If a fault happens frequently in one component, it should be replaced with a similar one, or the whole system should be changed to avoid the use of that type of component.
- The statistic is helpful to another part of network management, performance management.

3.6.6 Proactive Fault Management

Proactive fault management tries to prevent faults from occurring. Although this is not always possible, some types of failures can be predicted and prevented. For example, if a manufacturer specifies a lifetime for a component or a part of a component, it is a good strategy to replace it before that time. As another example, if a fault happens frequently at one particular point of a network, it is wise to carefully reconfigure the network to prevent the fault from happening again.

3.6.7 Performance Management

Performance management, which is closely related to fault management, tries to monitor and control the network to ensure that it is running as efficiently as possible. Performance management tries to quantify performance by using some measurable quantity such as capacity, traffic, throughput, or response time.

Capacity

One factor that must be monitored by a performance management system is the capacity of the network. Every network has a limited capacity, and the performance management system must ensure that it is not used above this capacity. For example, if a LAN is designed for 100 stations at an average data rate of2 Mbps, it will not operate properly if 200 stations are connected to the network. The data rate will decrease and blocking may occur.

Traffic

Traffic can be measured in two ways: internally and externally. Internal traffic is measured by the number of packets (or bytes) traveling inside the network for a given period. External traffic is measured by the exchange of packets (or bytes) outside the network. During peak hours, when the system is heavily used, blocking may occur if there is excessive traffic.

Throughput

We can measure the throughput of an individual device (such as a router) or a part of the network. Performance management monitors the throughput to make sure that it is not reduced to unacceptable levels.

Response Time

Response time is normally measured from the time a user requests a service to the time the service is granted. Other factors such as capacity and traffic can affect the response time. Performance management monitors the average response time and the peak-hour response time. Any increase in response time is a very serious condition as it is an indication that the network is working above its capacity.

3.6.8 Security Management

Security management is responsible for controlling access to the network based on the predefined policy.

3.6.9 Accounting Management

Accounting management is the control of users' access to network resources through charges. Under accounting management, individual users, departments, divisions, or even projects are charged for the services they receive from the network. Charging does not necessarily mean cash transfer; it may mean debiting the departments or divisions for budgeting purposes. Today, organizations use an accounting management system for the following reasons:

- It prevents users from monopolizing limited network resources.
- It prevents users from using the system inefficiently.
- Network managers can do short- and long-term planning based on the demand for
- Network use.

3.6.10 SNMP Protocol

Since it was developed in 1988, the Simple Network Management Protocol has become the de facto standard for internetwork management. Since it is a simple solution, requiring little code to implement, vendors can easily build SNMP agents to their products. SNMP is extensible, allowing vendors to easily add network management functions to their existing products. SNMP also separates the management architecture from the architecture of the hardware devices, which broadens the base of multivendor support. Perhaps most important, unlike other so-called standards, SNMP is not a mere paper specification, but an implementation that is widely available today. To know more about SNMP you may refer to fifth semester course BCS-052 Network programming and Administration.

SNMP is based on the manager/agent model consisting of a manager, an agent, a database of management information, managed objects and the network protocol. The manager provides the interface between the human network manager and the management system. The agent provides the interface between the manager and the physical device(s) being managed (see the illustration above).

A typical agent usually

- Implements full SNMP protocol.
- Stores and retrieves management data as defined by the Management Information Base.
- Can asynchronously signal an event to the manager
- Can be a proxy for some non-SNMP manageable network node.

A typical manager usually

- Implemented as a Network Management Station (the NMS)
- Implements full SNMP Protocol Able to
- Query agents
- Get responses from agents
- Set's variables in agents
- Acknowledge's asynchronous events from agents.

☞ Check Your Progress 2

1. Select the right choice.

   a) FTP does not use

   (A) Two transfer mode.
   (B) Control connection to remote computer before file can be transferred.
   (C) User Datagram Protocol.
   (D) Authorization of a user through login and password verification.

   ……
   ……

   b) Protocol used to monitor and control network devices operates at:

   (A) Application layer (B) Transport layer
   (C) Network layer (D) Data Link layer

   ……
   ……

2. Discuss TCP connections in FTP.

   ……
   ……

3. What is configuration management in computer network management?

   ……
   ……

3.7 WORD WIDE WEB AND CLIENT SERVER APPLICATIONS

The World Wide Web is an architectural framework for accessing linked documents spread out over millions of machines all over the Internet. In the last decade on so it went from being a way to distribute high-energy physics data to the application that millions of people think of as being "The Internet." Its enormous popularity stems from the fact that it has a colorful graphical interface that is easy for beginners to use, and it provides an enormous wealth of information on almost every conceivable subject.

The Web (also known as **WWW**) began in 1989 at CERN, the European center for nuclear research. CERN has several accelerators at which large teams of scientists from the participating European countries carry out research in particle physics. These teams often have members from half a dozen or more countries. Most experiments are highly complex and require years of advance planning and equipment construction. The Web grew out of the need to have these large teams of internationally dispersed researchers collaborate using a constantly changing collection of reports, blueprints, drawings, photos, and other documents.

The initial proposal for a web of linked documents came from CERN physicist Tim Berners-Lee in March 1989. The first (text-based) prototype was operational 18 months later. In December 1991, a public demonstration was given at the Hypertext '91 conference in San Antonio, Texas.

This demonstration and its attendant publicity caught the attention of other researchers, which led Marc Andreessen at the University of Illinois to start developing the first graphical browser, Mosaic. It was released in February 1993. Mosaic was so popular that a year later, Andreessen left to form a company, Netscape Communications Corp., whose goal was to develop clients, servers, and other Web software. When Netscape went public in 1995, investors, apparently thinking this was the next Microsoft, paid $1.5 billion for the stock. This record was all the more surprising because the company had only one product, was operating deeply in the red, and had announced in its prospectus that it did not expect to make a profit for the foreseeable future. For the next three years, Netscape Navigator and Microsoft's Internet Explorer engaged in a "browser war," each one trying frantically to add more features (and thus more bugs) than the other one. In 1998, America Online bought Netscape Communications Corp. for $4.2 billion, thus ending Netscape's brief life as an independent company.

In 1994, CERN and M.I.T. signed an agreement setting up the **World Wide Web Consortium** (sometimes abbreviated as **W3C**), an organization devoted to further developing the Web, standardizing protocols, and encouraging interoperability between sites. Berners-Lee became the director. Since then, several hundred universities and companies have joined the consortium. Although there are now more books about the Web than you can shake a stick at, the best place to get up-to-date information about the Web is (naturally) on the Web itself.

3.7.1 Architectural Overview (WWW)

From the users' point of view, the Web consists of a vast, worldwide collection of documents or **Web pages**, often just called **pages** for short. Each page may contain links to other pages anywhere in the world. Users can follow a link by clicking on it, which then takes them to the page pointed to. This process can be repeated indefinitely. The idea of having one page point to another, now called **hypertext**, was

invented by a visionary M.I.T. professor of electrical engineering, Vannevar Bush, in 1945, long before the Internet was invented.
Pages are viewed with a program called a **browser**, of which Internet Explorer and Netscape Navigator are two popular ones. The browser fetches the page requested, interprets the text and formatting commands on it, and displays the page, properly formatted, on the screen. Web pages, starts with a title, contain some information, and ends with the e-mail address of the page's maintainer. Strings of text that are links to other pages, called **hyperlinks**, are often highlighted, by underlining, displaying them in a special color, or both. To follow a link, the user places the mouse cursor on the highlighted area, which causes the cursor to change, and clicks on it. Although non graphical browsers, such as Lynx, exist, they are not as popular as graphical browsers. Voice-based browsers are also being developed.

The basic model of how the Web works is shown in Figure 4, Here the browser is displaying a Web page on the client machine. When the user clicks on a line of text that is linked to a page on the *abcd.com* server, the browser follows the hyperlink by sending a message to the *abcd.com* server asking it for the page. When the page arrives, it is displayed. If this page contains a hyperlink to a page on the *xyz.com* server that is clicked on, the browser then sends a request to that machine for the page, and so on indefinitely.

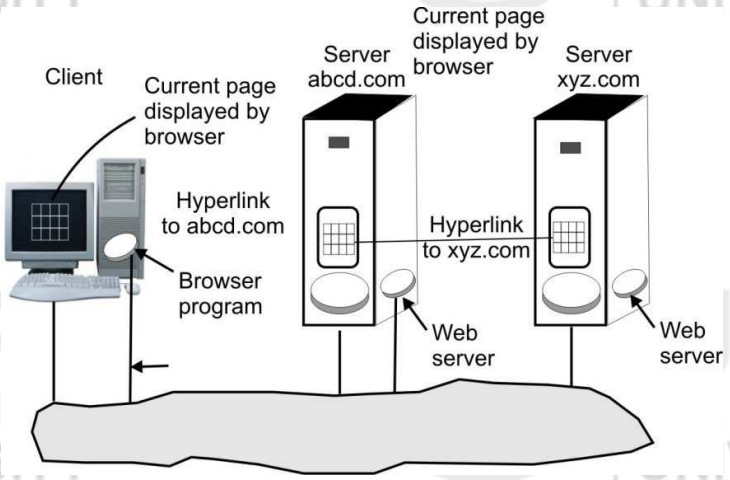


Figure 4: The parts of the Web model.

The Client Side

In essence, a browser is a program that can display a Web page and catch mouse clicks to items on the displayed page. When an item is selected, the browser follows the hyperlink and fetches the page selected. Therefore, the embedded hyperlink needs a way to name any other page on the Web. Pages are named using **URLs** (**Uniform Resource Locators**). A typical URL is http://www.abcd.com/products.html

It is sufficient to know that a URL has three parts: the name of the protocol (*http*), the DNS name of the machine where the page is located (*www.abcd.com*), and (usually) the name of the file containing the page (*products.html*).

When a user clicks on a hyperlink, the browser carries out a series of steps in order to fetch the page pointed to. Suppose that a user is browsing the Web and finds a link on Internet telephony that point to ITU's home page, which is *http://www.itu.org/home/index.html*. Let us trace the steps that occur when this link is selected.

1. The browser determines the URL (by seeing what was selected).

2. The browser asks DNS for the IP address of www.itu.org.
3. DNS replies with 156.106.192.32.
4. The browser makes a TCP connection to port 80 on 156.106.192.32.
5. It then sends over a request asking for file /home/index.html.
6. The www.itu.org server sends the file /home/index.html.
7. The browser displays all the text in /home/index.html.
8. The browser fetches and displays all images in this file.

Many browsers display which step they are currently executing in a status line at the bottom of the screen. In this way, when the performance is poor, the user can see if it is due to DNS not responding, the server not responding, or simply network congestion during page transmission.

Although a browser is basically an HTML interpreter, most browsers have numerous buttons and features to make it easier to navigate the Web. Most have a button for going back to the previous page, a button for going forward to the next page (only operative after the user has gone back from it), and a button for going straight to the user's own start page. Most browsers have a button or menu item to set a bookmark on a given page and another one to display the list of bookmarks, making it possible to revisit any of them with only a few mouse clicks. Pages can also be saved to disk or printed. Numerous options are generally available for controlling the screen layout and setting various user preferences.

In addition to having ordinary text (not underlined) and hypertext (underlined), Web pages can also contain icons, line drawings, maps, and photographs. Each of these can (optionally) be linked to another page. Clicking on one of these elements causes the browser to fetch the linked page and display it on the screen, the same as clicking on text. With images such as photos and maps, which page is fetched next may depend on what part of the image was clicked on.

Not all pages contain HTML. A page may consist of a formatted document in PDF format, an icon in GIF format, a photograph in JPEG format, a song in MP3 format, a video in MPEG format, or any one of hundreds of other file types. Since standard HTML pages may link to any of these, the browser has a problem when it encounters a page it cannot interpret.

Rather than making the browsers larger and larger by building in interpreters for a rapidly growing collection of file types, most browsers have chosen a more general solution. When a server returns a page, it also returns some additional information about the page. This information includes the MIME type of the page. Pages of type *text/html* are just displayed directly, as are pages in a few other built-in types. If the MIME type is not one of the built-in ones, the browser consults its table of MIME types to tell it how to display the page.

There are two possibilities: plug-ins and helper applications. A **plug-in** is a code module that the browser fetches from a special directory on the disk and installs as an extension to itself, as illustrated in Figure 5(a) Because plug-ins run inside the browser, they have access to the current page and can modify its appearance. After the plug-in has done its job (usually after the user has moved to a different Web page), the plug-in is removed from the browser's memory.

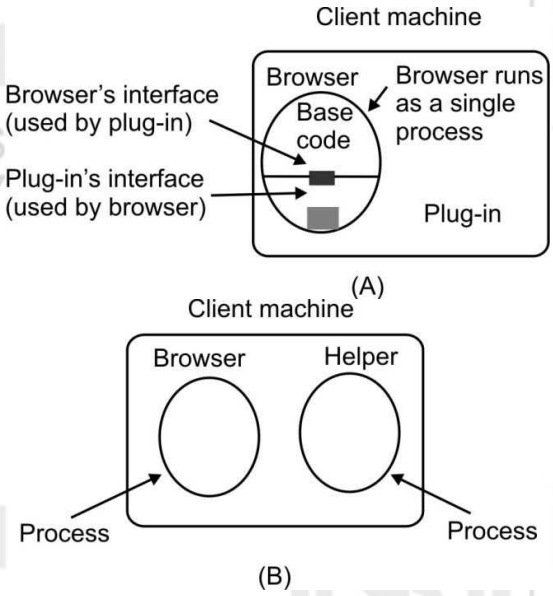


Figure 5: (a) A browser plug-in. (b) A helper application.

Each browser has a set of procedures that all plug-ins must implement so the browser can call the plug-in. For example, there is typically a procedure the browser's base code calls to supply the plug-in with data to display. This set of procedures is the plug-in's interface and is browser specific.

In addition, the browser makes a set of its own procedures available to the plug-in, to provide services to plug-ins. Typical procedures in the browser interface are for allocating and freeing memory, displaying a message on the browser's status line, and querying the browser about parameters.

Before a plug-in can be used, it must be installed. The usual installation procedure is for the user to go to the plug-in's Web site and download an installation file. On Windows, this is typically a self-extracting zip file with extension *.exe*. When the zip file is double clicked, a little program attached to the front of the zip file is executed. This program unzips the plug-in and copies it to the browser's plug-in directory. Then it makes the appropriate calls to register the plug-in's MIME type and to associate the plug-in with it. On UNIX, the installer is often a shell script that handles the copying and registration.

The other way to extend a browser is to use a **helper application**. This is a complete program, running as a separate process. It is illustrated in Figure 5(b). Since the helper is a separate program, it offers no interface to the browser and makes no use of browser services. Instead, it usually just accepts the name of a scratch file where the content file has been stored, opens the file, and displays the contents. Typically, helpers are large programs that exist independently of the browser, such as Adobe's Acrobat Reader for displaying PDF files or Microsoft Word. Some programs (such as Acrobat) have a plug-in that invokes the helper itself.

Many helper applications use the MIME type *application*. A considerable number of subtypes have been defined, for example, *application/pdf* for PDF files and *application/msword* for Word files. In this way, a URL can point directly to a PDF or Word file, and when the user clicks on it, Acrobat or Word is automatically started and handed the name of a scratch file containing the content to be displayed. Consequently, browsers can be configured to handle a virtually unlimited number of document types with no changes to the browser. Modern Web servers are often

configured with hundreds of type/subtype combinations and new ones are often added every time a new program is installed.

Helper applications are not restricted to using the *application* MIME type. Adobe Photoshop uses *image/x-photoshop* and Real One Player is capable of handling *audio/mp3*, for example.

On Windows, when a program is installed on the computer, it registers the MIME types it wants to handle. This mechanism leads to conflict when multiple viewers are available for some subtype, such as *video/mpg*. What happens is that the last program to register overwrites existing (MIME type, helper application) associations, capturing the type for itself. As a consequence, installing a new program may change the way a browser handles existing types.

On UNIX, this registration process is generally not automatic. The user must manually update certain configuration files. This approach leads to more work but fewer surprises.

Browsers can also open local files, rather than fetching them from remote Web servers. Since local files do not have MIME types, the browser needs some way to determine which plug-in or helper to use for types other than its built-in types such as *text/html* and *image/jpeg*. To handle local files, helpers can be associated with a file extension as well as with a MIME type. With the standard configuration, opening *foo.pdf* will open it in Acrobat and opening *bar.doc* will open it in Word. Some browsers use the MIME type, the file extension, and even information taken from the file itself to guess the MIME type. In particular, Internet Explorer relies more heavily on the file extension than on the MIME type when it can.

Here, too, conflicts can arise since many programs are willing, in fact, eager, to handle, say, *.mpg*. During installation, programs intended for professionals often display checkboxes for the MIME types and extensions they are prepared to handle to allow the user to select the appropriate ones and thus not overwrite existing associations by accident. Programs aimed at the consumer market assume that the user does not have a clue what a MIME type is and simply grab everything they can without regard to what previously installed programs have done.

The ability to extend the browser with a large number of new types is convenient but can also lead to trouble. When Internet Explorer fetches a file with extension *exe*, it realizes that this file is an executable program and therefore has no helper. The obvious action is to run the program. However, this could be an enormous security hole. All a malicious Web site has to do is produce a Web page with pictures of, say, movie stars or sports heroes, all of which are linked to a virus. A single click on a picture then causes an unknown and potentially hostile executable program to be fetched and run on the user's machine. To prevent unwanted guests like this, Internet Explorer can be configured to be selective about running unknown programs automatically, but not all users understand how to manage the configuration.

The Server Side

Now let us take a look at the server side. As we saw above, when the user types in a URL or clicks on a line of hypertext, the browser parses the URL and interprets the part between *http://* and the next slash as a DNS name to look up. Armed with the IP address of the server, the browser establishes a TCP connection to port 80 on that server. Then it sends over a command containing the rest of the URL, which is the name of a file on that server. The server then returns the file for the browser to display.

The steps that the server performs in its main loop are as follows:

1. Accept a TCP connection from a client (a browser).
2. Get the name of the file requested.
3. Get the file (from disk).
4. Return the file to the client.
5. Release the TCP connection.

Modern Web servers have more features, but in essence, this is what a Web server does.

A problem with this design is that every request requires making a disk access to get the file. The result is that the Web server cannot serve more requests per second than it can make disk accesses.

One obvious improvement (used by all Web servers) is to maintain a cache in memory of the *n* most recently used files. Before going to disk to get a file, the server checks the cache. If the file is there, it can be served directly from memory, thus eliminating the disk access. Although effective caching requires a large amount of main memory and some extra processing time to check the cache and manage its contents, the savings in time are nearly always worth the overhead and expense.

The next step for building a faster server is to make the server multithreaded. In one design, the server consists of a front-end module that accepts all incoming requests and *k* processing modules, as shown in Figure 6. The $k + 1$ threads all belong to the same process so the processing modules all have access to the cache within the process' address space. When a request comes in, the front end accepts it and builds a short record describing it. It then hands the record to one of the processing modules. In another possible design, the front end is eliminated and each processing module tries to acquire its own requests, but a locking protocol is then required to prevent conflicts.

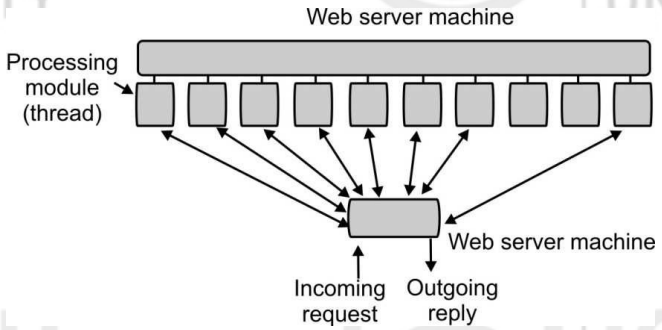


Figure 6: A multithreaded Web server with a front end and processing modules.

The processing module first checks the cache to see if the file needed is there. If so, it updates the record to include a pointer to the file in the record. If it is not there, the processing module starts a disk operation to read it into the cache (possibly discarding some other cached files to make room for it). When the file comes in from the disk, it is put in the cache and also sent back to the client.

The advantage of this scheme is that while one or more processing modules are blocked waiting for a disk operation to complete (and thus consuming no CPU time), other modules can be actively working on other requests. Of course, to get any real

improvement over the single-threaded model, it is necessary to have multiple disks, so more than one disk can be busy at the same time. With *k* processing modules and *k* disks, the throughput can be as much as *k* times higher than with a single-threaded server and one disk.

In theory, a single-threaded server and *k* disks could also gain a factor of *k*, but the code and administration are far more complicated since normal blocking READ system calls cannot be used to access the disk. With a multithreaded server, they can be used since then a READ blocks only the thread that made the call, not the entire process.

Modern Web servers do more than just accept file names and return files. In fact, the actual processing of each request can get quite complicated. For this reason, in many servers each processing module performs a series of steps. The front end passes each incoming request to the first available module, which then carries it out using some subset of the following steps, depending on which ones are needed for that particular request.

1. Resolve the name of the Web page requested.
2. Authenticate the client.
3. Perform access control on the client.
4. Perform access control on the Web page.
5. Check the cache.
6. Fetch the requested page from disk.
7. Determine the MIME type to include in the response.
8. Take care of miscellaneous odds and ends.
9. Return the reply to the client.
10. Make an entry in the server log.

Step 1 is needed because the incoming request may not contain the actual name of the file as a literal string. For example, consider the URL *http://www.cs.vu.nl*, which has an empty file name. It has to be expanded to some default file name. Also, modern browsers can specify the user's default language (e.g., Italian or English), which makes it possible for the server to select a Web page in that language, if available. In general, name expansion is not quite so trivial as it might at first appear, due to a variety of conventions about file naming.

Step 2 consists of verifying the client's identity. This step is needed for pages that are not available to the general public. We will discuss one way of doing this later in this chapter.

Step 3 checks to see if there are restrictions on whether the request may be satisfied given the client's identity and location. Step 4 checks to see if there are any access restrictions associated with the page itself. If a certain file (e.g., *.htaccess*) is present in the directory where the desired page is located, it may restrict access to the file to particular domains, for example, only users from inside the company.
Steps 5 and 6 involve getting the page. Step 6 needs to be able to handle multiple disk reads at the same time.

Step 7 is about determining the MIME type from the file extension, first few words of the file, a configuration file, and possibly other sources.

Step 8 is for a variety of miscellaneous tasks, such as building a user profile or gathering certain statistics.

Step 9 is where the result is sent back and step 10 makes an entry in the system log for administrative purposes. Such logs can later be mined for valuable information about user behavior, for example, the order in which people access the pages.

If too many requests come in each second, the CPU will not be able to handle the processing load, no matter how many disks are used in parallel. The solution is to add more nodes (computers), possibly with replicated disks to avoid having the disks become the next bottleneck. This leads to the **server farm** model of Figure 7. A front end still accepts incoming requests but sprays them over multiple CPUs rather than multiple threads to reduce the load on each computer. The individual machines may themselves be multithreaded and pipelined as before.

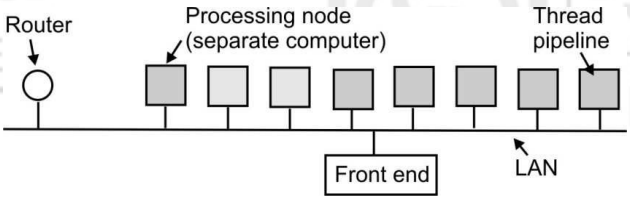


Figure 7: A server farm.

One problem with server farms is that there is no longer a shared cache because each processing node has its own memory unless an expensive shared-memory multiprocessor is used. One way to counter this performance loss is to have a front end keep track of where it sends each request and send subsequent requests for the same page to the same node. Doing this makes each node a specialist in certain pages so that cache space is not wasted by having every file in every cache. Another problem with server farms is that the client's TCP connection terminates at the front end, so the reply must go through the front end. This situation is depicted in Figure 8(a), where the incoming request (1) and outgoing reply (4) both pass through the front end.

Sometimes a trick, called **TCP handoff**, is used to get around this problem. With this trick, the TCP end point is passed to the processing node so it can reply directly to the client, shown in Figure 8(b). This handoff is done in a way that is transparent to the client.

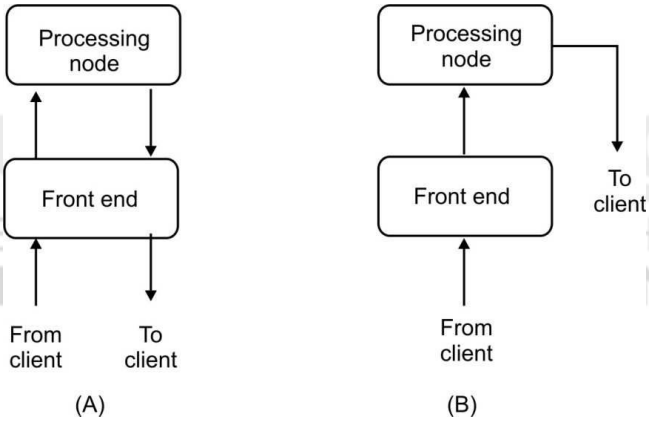


Figure 8: (a) Normal request-reply message sequence. (b) Sequence when TCP handoff is used.

HTTP

Hyper Text Transfer Protocol (HTTP) is used mainly to transfer data on World Wide Web. The commands from the client are embedded in a request message .The contents of the request message are embedded in a response message. HTTP uses the services of TCP at port 80.

HTTP is a stateless protocol since each transaction is independent of the previous transaction. The TCP connection between the client and the server is established for every page. It does not remember anything about the previous request. Keeping HTTP stateless was aimed at making the Web simple.

Sample HTTP request and response transaction is shown below:

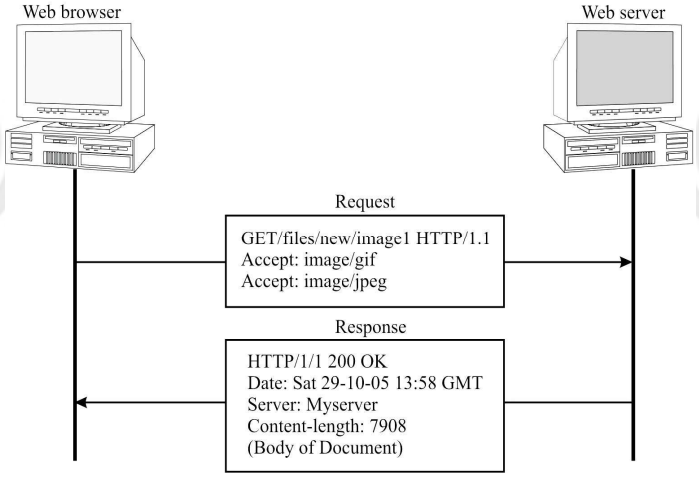


Figure 9: HTTP request and response

As shown in figure 9 above, the GET command requests the web server at www.myserver.com for an image file image1 .The HTTP/1.0 indicates that the browser uses the 1.0 version of the HTTP protocol.

The first line of response indicates that the server is also using HTTP version 1.0 .the return code of 200 indicates that the server processed the request successfully. Request message can be of following types:

Method	**Action**
GET	Requests a document from the server
HEAD	Request information about the document but not the document itself; i.e. head of the HTML page
POST	Sends some information from client to server. It appends the data to the existing document.
PUT	Sends a document from to server. It replaces the existing document.
TRACE	Echoes the incoming request
CONNECT	Reserved
OPTION	Enquire about available options

3.8 ELECTRONIC MAIL

Electronic mail, or **e-mail**, as it is known to its many fans, has been around for over two decades. Before 1990, it was mostly used in academia. During the 1990s, it became known to the public at large and grew exponentially to the point where the number of e-mails sent per day now is vastly more than the number of **snail mail** (i.e., paper) letters.

E-mail, like most other forms of communication, has its own conventions and styles. In particular, it is very informal and has a low threshold of use. People who would never dream of calling up or even writing a letter to a Very Important Person do not hesitate for a second to send a sloppily-written e-mail.

E-mail is full of jargon such as BTW (By The Way), ROTFL (Rolling On The Floor Laughing), and IMHO (In My Humble Opinion). Many people also use little ASCII symbols called **smileys** or **emoticons** in their e-mail.

The first e-mail systems simply consisted of file transfer protocols, with the convention that the first line of each message (i.e., file) contained the recipient's address. As time went on, the limitations of this approach became more obvious. Some of the complaints were as follows:

1. Sending a message to a group of people was inconvenient. Managers often need this facility to send memos to all their subordinates.

2. Messages had no internal structure, making computer processing difficult. For example, if a forwarded message was included in the body of another message, extracting the forwarded part from the received message was difficult.

3. The originator (sender) never knew if a message arrived or not.

4. If someone was planning to be away on business for several weeks and wanted all incoming e-mail to be handled by his secretary, this was not easy to arrange.

5. The user interface was poorly integrated with the transmission system requiring users first to edit a file, then leave the editor and invoke the file transfer program.

6. It was not possible to create and send messages containing a mixture of text, drawings, facsimile, and voice.

3.8.1 Architecture and Services

In this section we will provide an overview of what e-mail systems can do and how they are organized. They normally consist of two subsystems: the **user agents**, which allow people to read and send e-mail, and the **message transfer agents**, which move the messages from the source to the destination. The user agents are local programs that provide a command-based, menu-based, or graphical method for interacting with the e-mail system. The message transfer agents are typically system **daemons**, that is, processes that run in the background. Their job is to move e-mail through the system. Typically, e-mail systems support **five basic functions**.

Composition refers to the process of creating messages and answers. Although any text editor can be used for the body of the message, the system itself can provide assistance with addressing and the numerous header fields attached to each message. For example, when answering a message, the e-mail system can extract the

originator's address from the incoming e-mail and automatically insert it into the proper place in the reply.

Transfer refers to moving messages from the originator to the recipient. In large part, this requires establishing a connection to the destination or some intermediate machine, outputting the message, and releasing the connection. The e-mail system should do this automatically, without bothering the user.

Reporting has to do with telling the originator what happened to the message. Was it delivered? Was it rejected? Was it lost? Numerous applications exist in which confirmation of delivery is important and may even have legal significance ("Well, Your Honor, my e-mail system is not very reliable, so I guess the electronic subpoena just got lost somewhere").

Displaying incoming messages is needed so people can read their e-mail. Sometimes conversion is required or a special viewer must be invoked, for example, if the message is a PostScript file or digitized voice. Simple conversions and formatting are sometimes attempted as well.

Disposition is the final step and concerns what the recipient does with the message after receiving it. Possibilities include throwing it away before reading, throwing it away after reading, saving it, and so on. It should also be possible to retrieve and reread saved messages, forward them, or process them in other ways.

In addition to these basic services, some e-mail systems, especially internal corporate ones, provide a variety of advanced features. Let us just briefly mention a few of these. When people move or when they are away for some period of time, they may want their e-mail forwarded, so the system should be able to do this automatically. Most systems allow users to create **mailboxes** to store incoming e-mail. Commands are needed to create and destroy mailboxes, inspect the contents of mailboxes, insert and delete messages from mailboxes, and so on.

Corporate managers often need to send a message to each of their subordinates, customers, or suppliers. This gives rise to the idea of a **mailing list**, which is a list of e-mail addresses. When a message is sent to the mailing list, identical copies are delivered to everyone on the list.

3.8.2 The User Agent

E-mail systems have two basic parts, as we have seen: **the user agents** and **the message transfer agents**. In this section, we will look at the user agents. A user agent is normally a program (sometimes called a mail reader) that accepts a variety of commands for composing, receiving, and replying to messages, as well as for manipulating mailboxes. Some user agents have a fancy menu- or icon-driven interface that requires a mouse, whereas others expect 1-character commands from the keyboard. Functionally, these are the same. Some systems are menu-or icon-driven but also have keyboard shortcuts.

Sending E-mail

To send an e-mail message, a user must provide the message, the destination address, and possibly some other parameters.

Reading E-mail

Typically, when a user agent is started up, it looks at the user's mailbox for incoming e-mail before displaying anything on the screen. Then it may announce the number of

messages in the mailbox or display a one-line summary of each one and wait for a command.

SMTP server (Simple Mail Transfer Protocol)

For email messaging, every domain maintains an email server. The server runs protocols software that enable email communication. There are two main emails protocols: POP and SMTP. Because both the email protocol software programs run on server computers, the server computers are themselves called POP server and SMTP server. A single server can host both the POP and SMTP server programs.
SMTP is the Internet protocol used to transfer electronic mail between computers. The second generation of SMTP is called ESMTP (for Extended SMTP), but the differences are not important for this introduction.

It actually transfers the email message from the SMTP server of the sender to the SMTP server of the recipient. Its main job is to carry the message between the sender and the receiver. It uses TCP/IP underneath. That is, it runs on top of TCP/IP.
At the sender's site, an SMTP server takes the message sent by a user's computer. The SMTP server at the sender's end then transfers the message to the SMTP server of the recipient.

The SMTP server at the recipient's end takes the message and stores it in the appropriate user's mailbox.

☞ Check Your Progress 3

1. Select the right choice

   a) Internet's initial development was supported by:
   (A) ARPANET
   (B) Bill Rogers
   (C) Bill Gates
   (D) Microsoft

   ……
   ……

   b) What are the uses of the Internet?

   (A) Communication
   (B) Information Retrieval
   (C) Presentation of Information
   (D) All of the above

   ……
   ……

   c) net domain is used for

   (A) Educational institution
   (B) Internet infrastructure and service providers
   (C) International organizations
   (D) None of the above

   ……
   ……

2. Differentiate between http and ftp.

……

……

3.9 SUMMARY

This completes our discussion on the application layer. Include identifying communication partners, determining resource availability, and synchronizing communication. When identifying communication partners, the application layer determines the identity and availability of communication partners for an application with data to transmit. When determining resource availability, the application layer must decide whether sufficient network or the requested communication exists. The unit very well defines the concept of DNS and various internet and communication related issues like www, emailing system, FTP, Telnet etc. To know further about different application layer protocol students may refer to the course material of BCS-052: Network Programming and Administration

3.10 REFERENCES/FURTHER READING

1. Introduction to Data Communication & Networking, 3rd Edition, Behrouz Forouzan, Tata McGraw Hill.
2. Computer Networks, A. S. Tanenbaum 4th Edition, Practice Hall of India, New Delhi. 2003.
3. Douglas E. Comer, Internetworking with TCP/IP Vol.1: Principles, Protocols, and Architecture (4th Edition).
4. James F. Kurose, Computer Networking: A Top-Down Approach Featuring the Internet (3rd Edition).
5. Larry L. Peterson, Computer Networks: A Systems Approach, 3rd Edition (The Morgan Kaufmann Series in Networking).
6. www. wikipedia.org
7. W. Richard Stevens, The Protocols (TCP/IP Illustrated, Volume 1).
8. William Stallings, Data and Computer Communications, Seventh Edition.

3.11 SOLUTIONS / ANSWERS

☞ **Check Your Progress 1**

1. a) D
   b) B
   c) C
2. Managing a large and constantly changing set of names is a nontrivial problem. In the postal system, name management is done by requiring letters to specify (implicitly or explicitly) the country, state or province, city, and street address of the addressee. Conceptually, the Internet is divided into over 200 top-level **domains**, where each domain covers many hosts. Each domain is partitioned into sub domains, and these are further partitioned, and so on. All these domains can be represented by a tree. The top-level domains come in two flavors: **generic** and **countries**.

3. A single name server could contain the entire DNS database and respond to all queries about it. In practice, this server would be so overloaded as to be useless. Furthermore, if it ever went down, the entire Internet would be crippled. To avoid the problems associated with having only a single source of information, the DNS name space is divided into non-overlapping **zones**. A zone will have one primary name server, which gets its information from a file on its disk, and one or more secondary name servers, which get their information from the primary name server. To improve reliability, some servers for a zone can be located outside the zone.

☞ Check Your Progress 2

1. a) (C)

   b) (A)

2. Two TCP connections in FTP are:

   i) **The control connection** is established in the normal client-server fashion. In this case, the server does a *passive open* (is listening) on port 21 for FTP, and waits for the client connection. The client does an *active open* (the2nd handshake in a TCP connection) to establish the control connection. The client uses an ephemeral port number (above 1023) for the control connection. This control connection stays up for the entire time that the client communicates with this server. This connection is used for commands from the client to the server and for the server's replies. The IP type of service for the control connection should be to minimize delay in passing these commands over the TCP connection.

   ii) **A data connection** is created each time a file is transferred between the client and the server. The IP type of service for the data connection should be to maximize throughput since this connection is file transfer, and we want to send this entire file over a high bandwidth line. The specification of FTP includes more than 30 different commands, which are transmitted over the control connection in NVT ASCII format. The commands are not case-sensitive and may have arguments; each command ends with a two character sequence of a carriage return (CR) followed by a line feed (LF). It must be emphasized here that these commands are different from the commands typed by the user at the interface provided by the client. Transferring a single file for instance requires only a single user-level command (e.g., *put* or *get*), but this single command triggers the client to send a set of FTP commands to the server. The FTP server responds to each command with a three-digit reply code (for the FTP client) and an optional text message (for the user).

3. **Configuration management** system must know, at any time, the status of each entity and its relation to other entities. Configuration management can be divided into two subsystems: reconfiguration and documentation.

   Reconfiguration, which means adjusting the network components and features, can be a daily occurrence in a large network. There are three types of reconfiguration: hardware reconfiguration, software reconfiguration, and user-account reconfiguration. Hardware reconfiguration covers all changes to the hardware. For example, a desktop computer may need to be replaced. A router may need to be moved to another part of the network. A sub network may be added or removed from the network.

   The original network configuration and each subsequent change must be recorded meticulously. This means that there must be documentation for hardware, software, and user accounts. **Hardware documentation** normally

involves two sets of documents: maps and specifications. Maps track each piece of hardware and its connection to the network. There can be one general map that shows the logical relationship between each sub-network. There can also be a second general map that shows the physical location of each sub-network. For each sub network, then, there are one or more maps that show all pieces of equipment. The maps use some kind of standardization to be easily read and understood by current and future personnel. Each piece of hardware also needs to be documented. There must be a set of specifications for each piece of hardware connected to the network

☛ Check Your Progress 3

1. a) (A)
   b) (D)
   c) (C)

2. FTP and HTTP were developed to make Internet transmission better. FTP is used to exchange files between computer accounts, to transfer files between an account and a desktop computer (upload), or to access software archives on the Internet. It's also commonly used to download programs and other files to your computer from other servers. It transfers files in two different formats ASCII for text files and Binary format for binary files. This allows a user to perform basic file and directory management operations such as deleting, copying, or renaming. HTTP is used primarily in today's society as a set of rules for exchanging files (text, graphic images, sound, video, and other multimedia files) on the World Wide Web. It also provides access to other protocols like FTP, SMTP, NNTP, WAIS, Gopher, Telnet, and TN3270. Essential concepts that are part of HTTP include (as its name implies) the idea that files can contain references to other files whose selection will elicit additional transfer requests. Any web server machine contains, in addition to the HTML and other files it can serve, an HTTP daemon, a program that is designed to wait for HTTP requests and handle them when they arrive.

UNIT 4 NETWORK APPLICATIONS

Structure **Page Nos.**

4.0 Introduction 76
4.1 Objectives 76
4.2 Internet Applications 76
4.2.1 Email
4.2.2 Chatting
4.3 Social Networking 79
4.3.1 Blogs
4.3.2 Online multiplayer gaming
4.3.3 Facebook
4.3.4 Emerging Trends
4.3.5 Characteristics of social Networking
4.4 Railway Reservation System 85
4.5 Information Sharing 91
4.6 Electronic Governance 92
4.7 Online Processing and Collaborations 95
4.8 Mobile Applications 98
4.9 Summary 100
4.10 References/Further Readings 100
4.11 Solutions/Answers 101

4.0 INTRODUCTION

In this unit we are concentrating on the kind of applications that are used on the **Internet**. It is the part of network protocol (in the sense that they exchange messages with their peers on other machines) and part of traditional application program (in the sense that they interact with the windowing system, the file system, and ultimately, the user). It includes some of the most popular network applications available today Like: The World Wide Web and Email etc. We have also discussed some of the real-time applications like social networking, chatting, Railway Reservation system, Mobile Applications etc.

4.1 OBJECTIVES

After going through this unit you will be able to:

- define the logical structure of the Internet Applications;
- define the structure and working of network applications;
- define the concept of information sharing;
- discuss the basic features of E-Governance; and
- define the various components of mobile Applications;

4.2 INTERNET APPLICATIONS

The Internet is a global system of interconnected computer networks that uses the standard Internet protocol suite (often called TCP/IP, although not all applications use TCP) to serve billions of users worldwide. It is a network of networks that IS MODE of millions of private, public, academic, business, and government networks, of local to global scope, that are linked by a broad array of electronic, wireless and optical networking technologies. The Internet carries an extensive range of information

resources and services, such as the inter-linked hypertext documents of the World Wide Web (WWW) and the infrastructure to support email.

Most traditional communications media including telephone, music, film, and television are reshaped or redefined by the Internet, giving birth to new services such as Voice over Internet Protocol (VoIP) and Internet Protocol Television (IPTV). Newspaper, book and other print publishing are adapting to Web site technology, or are reshaped into blogging and web feeds. The Internet has enabled and accelerated new forms of human interactions through instant messaging, Internet forums, and social networking. Online shopping has boomed both for major retail outlets and small artisans and traders. Business-to-business and financial services on the Internet affect supply chains across entire industries.

The origins of the Internet reach back to research of the 1960s, commissioned by the United States government in collaboration with private commercial interests to build robust, fault-tolerant, and distributed computer networks. The funding of a new U.S. backbone by the National Science Foundation in the 1980s, as well as private funding for other commercial backbones, led to worldwide participation in the development of new networking technologies, and the merger of many networks. The commercialization of what was by the 1990s an international network resulted in its popularization and incorporation into virtually every aspect of modern human life. As of 2011, more than 2.2 billion people – nearly a third of Earth's population — use the services of the Internet.

The Internet has no centralized governance in either technological implementation or policies for access and usage; each constituent network sets its own standards. Only the overreaching definitions of the two principal name spaces in the Internet, the Internet Protocol address space and the Domain Name System, are directed by a maintainer organization, the Internet Corporation for Assigned Names and Numbers (ICANN). The technical underpinning and standardization of the core protocols (IPv4 and IPv6) is an activity of the Internet Engineering Task Force (IETF), a non-profit organization of loosely affiliated international participants that anyone may associate with by contributing technical expertise.

Internet is a short form of the technical term internetwork, the result of interconnecting computer networks with special gateways or routers. The Internet is also often referred to as the Net.

The term the Internet, when referring to the entire global system of IP networks, has been treated as a proper noun and written with an initial capital letter. In the media and popular culture, a trend has also developed to regard it as a generic term or common noun and thus write it as "the internet", without capitalization. Some guides specify that the word should be capitalized as a noun but not capitalized as an adjective.

The terms Internet and World Wide Web are often used in everyday speech without much distinction. However, the Internet and the World Wide Web are not one and the same. The Internet establishes a global data communications system between computers. In contrast, the Web is one of the services communicated via the Internet. It is a collection of interconnected documents (web pages) and other resources, linked by hyperlinks and URLs. In addition to the Web, the Internet also powers a multitude of other services, including (among others) email, file transfer, newsgroups, and online games. Web services can exist apart from the internet, such as on a private intranet.

The Internet allows greater flexibility in working hours and location, especially with the spread of unmetered high-speed connections. The Internet can be accessed almost anywhere by numerous means, including through mobile Internet devices. Mobile phones, data cards, handheld game consoles and cellular routers allow users to connect to the Internet wirelessly. Within the limitations imposed by small screens and other limited facilities of such pocket-sized devices, the services of the Internet, including email and the web, may be available. Service providers may restrict the services offered and mobile data charges may be significantly higher than other access methods.

Educational material at all levels from pre-school to post-doctoral is available from websites. Examples range from CBeebies, through school and high-school revision guides, virtual universities, to access to top-end scholarly literature through the likes of Google Scholar. For distance education, help with homework and other assignments, self-guided learning, whiling away spare time, or just looking up more detail on an interesting fact, it has never been easier for people to access educational information at any level from anywhere. The Internet in general and the World Wide Web in particular are important enablers of both formal and informal education.

The low cost and nearly instantaneous sharing of ideas, knowledge, and skills has made collaborative work dramatically easier, with the help of collaborative software. Not only can a group cheaply communicate and share ideas but the wide reach of the Internet allows such groups more easily to form. An example of this is the free software movement, which has produced, among other things, Linux, Mozilla Firefox, and OpenOffice.org. Internet chat, whether in the form of an IRC chat room or channel, via an instant messaging system, or a social networking website, allows colleagues to stay in touch in a very convenient way when working at their computers during the day. Messages can be exchanged even more quickly and conveniently than via email. These systems may allow files to be exchanged, drawings and images to be shared, or voice and video contact between team members.

4.2.1 Email

Electronic mail is one of the most popular tools made available through the Internet. It is an efficient and effective means of network communication. You can call it as an electronic postal system. One of the most valuable features of communicating via email is that it is asynchronous, meaning the recipient need not be at a computer to receive the message you send. The message will be stored and available to be read when the recipient is ready to read it. In order to send and receive email, you must have access to an Email account.

4.2.2 Chatting

Chatting may refer to any kind of communication over the Internet that offers a real time direct transmission of text-based messages from sender to receiver, hence the delay for visual access to the sent message shall not hamper the flow of communications in any of the directions. Online chat may address point-to-point communications as well as multicast communications from one sender to many receivers and voice and video chat or may be a feature of a Web conferencing service.

Online chat in a lesser stringent definition may be primarily any direct text-based or video-based (webcams), one-on-one chat or one-to-many group chat (formally also known asynchronous conferencing), using tools such as instant messengers, Internet Relay Chat, talkers and possibly MUDs. The expression online chat comes from the word chat which means "informal conversation". Online chat includes web-based applications that allow communication - often directly addressed, but anonymous - between users in a multi-user environment. Web conferencing is a more specific

online service, which is often sold as a service, hosted on a web server controlled by the vendor.

"Real-time communication between two users via computer. Once a chat has been initiated, either user can enter text by typing on the keyboard and the entered text will appear on the other user's monitor. Most networks and online services offer a chat feature."

Issues with Chatting

There has been much criticism about what online chatting has done in today's society. Many people are accusing it of replacing proper English with short hand with an almost completely new hybrid language.

Writing is changing as it takes on some of the functions and features of speech. Internet chat-rooms and rapid real-time conferencing allow users to interact with whoever happens to coexist in cyberspace. These virtual interactions involve us in `talking' more freely and more widely than ever before (Merchant, 2001). With chat-rooms replacing many face-to-face conversations it is necessary to be able to have quick conversation as if the person were present; so many people learn to type as quickly as they would normally speak. Critics are wary that this casual form of speech is being used so much that it will slowly take over common grammar; however, such a change has yet to be seen. With the increasing population of online chat-rooms there has been a massive growth of new words created or slang words, many of them documented on the website Urban Dictionary online chatting can be defined as:

"as new electronic modes of communication provoke similar anxieties amongst critics who express concern that young people are at risk, endangered by a rising tide of information over which the traditional controls of print media and the guardians of knowledge have no control on it".

4.3 SOCIAL NETWORKING

Many people use the World Wide Web to access news, weather and sports reports, to plan and book vacations and to find out more about their interests. People use chat, messaging and email to make and stay in touch with friends worldwide, sometimes in the same way as some previously had pen pals. The Internet has seen a growing number of Web desktops, where users can access their files and settings via the Internet.

Social networking websites such as Facebook, Twitter, and MySpace have created new ways to socialize and interact, some of these social networking groups are depicted in the Figure 1. Users of these sites are able to add a wide variety of information to pages, to pursue common interests, and to connect with others. It is also possible to find existing acquaintances, to allow communication among existing groups of people. Sites like LinkedIn foster commercial and business connections. YouTube and Flickr specialize in users' videos and photographs.

Figure 1: Some of the Social networking groups

The Internet has been a major outlet for leisure activity since its inception, with entertaining social experiments such as MUDs and MOOs being conducted on university servers, and humor-related Usenet groups receiving much traffic. Today, many Internet forums have sections devoted to games and funny videos; short cartoons in the form of Flash movies are also popular. Over 6 million people use blogs or message boards as a means of communication and for the sharing of ideas. The pornography and gambling industries have taken advantage of the World Wide Web, and often provide a significant source of advertising revenue for other websites. Although many governments have attempted to restrict both industries' use of the Internet, in general this has failed to stop their widespread popularity.

4.3.1 Blogs

Blog is a website where entries are written as commentary or news on a particular subject such as food, politics, or local news; some function as more personal online diaries. A typical blog combines text, images, and links to other blogs, web pages, and other media related to its topic. The ability for readers to leave comments in an interactive format is an important part of many blogs. Most blogs are primarily textual, although some focus on art, photographs, videos and music and are part of a wider network of social media. One of the Blogging service named Blogger is depicted in the Figure 2.

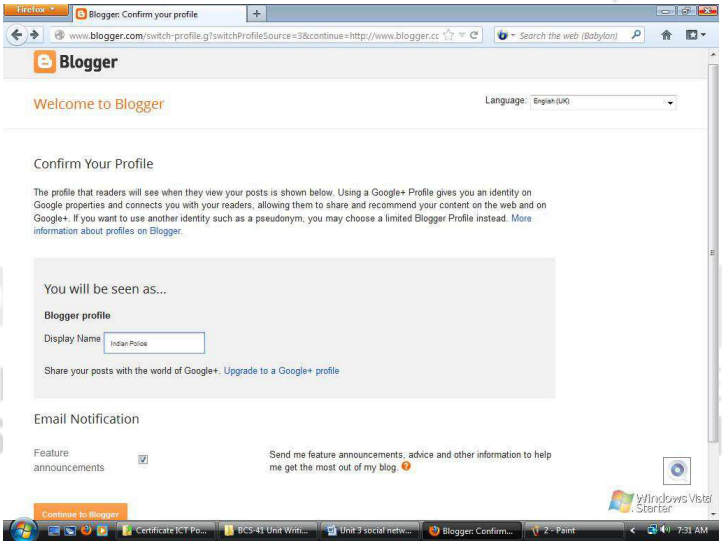


Figure 2: One of the Blogging service named Blogger

Blogging is gaining popularity in education, as it removes the technical barriers of writing and publishing online, which encourages students to keep a record of their ideas and thinking over time. Blogging also facilitate readers to give critical feedback on any topic, readers can add comments, where readers can be teachers, other students or a wider viewers. Teachers should investigate the potential of blogs, media-sharing services and other social software, which can be used create new learning opportunities. Students can also use the blogs as blog can provide a personal space online, to ask questions, comment on other questions, publish work, and link to other web sources. However a blog needn't be restricted to a single author, it can merge different kinds of ideas, including fellow students, teachers, and subject specialists. An example: http://edu.blogs.com/.

Blogging platform alternatives are following:

- WordPress.org is not the only blogging platform out there. There are several popular alternatives. First of all there is WordPress.com. Learn here about WordPress.org vs. WordPress.com.
- Google's Blogspot Blogger is another popular option. See here a bit on the difference between WordPress or Blogger.
- Tumblr is another popular choice. If you're wondering which on to pick, check Tumblr vs WordPress article.
- Google offer Blogger, which is easy way to collaborate, discuss, or share your thoughts with others. In the following section we will discuss how to create a blog, and discuss the main features of Blogger.

4.3.2 Online multiplayer gaming

Another area of leisure activity on the Internet is multiplayer gaming. This form of recreation creates communities, where people of all ages and origins enjoy the fast-paced world of multiplayer games. These range from MMORPG to first-person shooters, from role-playing video games to online gambling. While online gaming has been around since the 1970s, modern modes of online gaming began with subscription services such as GameSpy and MPlayer. Non-subscribers were limited to certain types of game play or certain games. Many people use the Internet to access and download music, movies and other works for their enjoyment and relaxation. Free and fee-based services exist for all of these activities, using centralized servers and distributed peer-to-peer technologies. Some of these sources exercise more care with respect to the original artists' copyrights than others.

Internet usage has been correlated to users' loneliness. Lonely people tend to use the Internet as an outlet for their feelings and to share their stories with others, such as in the "I am lonely will anyone speak to me" thread.

Cyber-sectarianism is a new organizational form which involves: "highly dispersed small groups of practitioners that may remain largely anonymous within the larger social context and operate in relative secrecy, while still linked remotely to a larger network of believers who share a set of practices and texts, and often a common devotion to a particular leader. Overseas supporters provide funding and support; domestic practitioners distribute tracts, participate in acts of resistance, and share information on the internal situation with outsiders. Collectively, members and practitioners of such sects construct viable virtual communities of faith, exchanging personal testimonies and engaging in collective study via email, on-line chat rooms and web-based message boards."

A **social networking service** is an online service, platform, or site that focuses on facilitating the building of social networks or social relations among people who, for

example, share interests, activities, backgrounds, or real-life connections. A social network service consists of a representation of each user (often a profile), his/her social links, and a variety of additional services. Most social network services are web-based and provide means for users to interact over the Internet, such as e-mail and instant messaging. Online community services are sometimes considered as a social network service, though in a broader sense, social network service usually means an individual-centered service where as online community services are group-centered. Social networking sites allow users to share ideas, activities, events, and interests within their individual networks.

The main types of social networking services are those that contain category places (such as former school year or classmates), means to connect with friends (usually with self-description pages), and a recommendation system linked to trust. Popular methods now combine many of these, with Facebook, Twitter and Google+ widely used worldwide.

Some social networks have additional features, such as the ability to create groups that share common interests or affiliations, upload or stream live videos, and hold discussions in forums. Geosocial networking co-opts Internet mapping services to organize user participation around geographic features and their attributes.

There is a trend towards more interoperability between social networks led by technologies such as OpenID and OpenSocial. In most mobile communities, mobile phone users can now create their own profiles, make friends, participate in chat rooms, create chat rooms, hold private conversations, share photos and videos, and share blogs by using their mobile phone. Some companies provide wireless services that allow their customers to build their own mobile community and brand it; one of the most popular wireless services for social networking in North America is Facebook Mobile.

4.3.3 Facebook

At present, Facebook is one of the popular Social Networking sites used by millions of people around the world, especially young people including your learners to connect to each other; Figure 3 shows the Login/Signup page for Facebook. This site is a free and effective way of communication on-line with your learners. You can, for example, send the messages, assignments and on-line resources for your subject. They in turn can communicate with you by posting questions etc.

Figure 3: Login/Signup page for Facebook

4.3.4 Emerging Trends

As the increase in popularity of social networking is on a constant rise, new uses for the technology are constantly being observed.

At the forefront of emerging trends in social networking sites is the concept of "real-time web" and "location-based." Real-time allows users to contribute content, which is then broadcast as it is being uploaded - the concept is analogous to live radio and television broadcasts. Twitter set the trend for "real-time" services, wherein users can broadcast to the world what they are doing, or what is on their minds within a 140-character limit. Facebook followed suit with their "Live Feed" where users' activities are streamed as soon as it happens. While Twitter focuses on words, Clixtr, another real-time service, focuses on group photo sharing wherein users can update their photo streams with photos while at an event. Facebook, however, remains the largest photo sharing site - Facebook application and photo aggregator Pixable estimates that Facebook will have more 100 billion photos by mid of 2011. In April, 2012, the image-based social media network *Pinterest* had become the third largest social network in the United States.

Companies have begun to merge business technologies and solutions, such as cloud computing, with social networking concepts. Instead of connecting individuals based on social interest, companies are developing interactive communities that connect individuals based on shared business needs or experiences. Many provide specialized networking tools and applications that can be accessed via their websites, such as LinkedIn. Others companies, such as Monster.com, have been steadily developing a more "socialized" feel to their career center sites to harness some of the power of social networking sites. These more business related sites have their own nomenclature for the most part but the most common naming conventions are "Vocational Networking Sites" or "Vocational Media Networks", with the former more closely tied to individual networking relationships based on social networking principles.

Foursquare gained popularity as it allowed for users to "check-in" to places that they are frequenting at that moment. *Gowalla* is another such service that functions in much the same way that Foursquare does, leveraging the GPS in phones to create a location-based user experience. Clixtr, though in the real-time space, is also a location-based social networking site, since events created by users are automatically geotagged, and users can view events occurring nearby through the Clixtr iPhone app. Recently, Yelp announced its entrance into the location-based social networking space through check-ins with their mobile app; whether or not this becomes detrimental to Foursquare or Gowalla is yet to be seen, as it is still considered a new space in the Internet technology industry.

One popular use for this new technology is social networking between businesses. Companies have found that social networking sites such as Facebook and Twitter are great ways to build their brand image.

There are five major uses for businesses and social media: to create brand awareness, as an online reputation management tool, for recruiting, to learn about new technologies and competitors, and as a lead generation tool to intercept potential prospects. These companies are able to drive traffic to their own online sites while encouraging their consumers and clients to have discussions on how to improve or change products or services.

4.3.5 Characteristics of social Networking

Social networks and science

By sharing information and knowledge with one another, people are able to "increase both their learning and their flexibility in ways that would not be possible within a self-contained hierarchical organization." Social networking is allowing scientific groups to expand their knowledge base and share ideas, and without these new means of communicating their theories might become "isolated and irrelevant".

Social networks and education

Social networks are also being used by teachers and students as a communication tool. Because many students are already using a wide range of social networking sites, teachers have begun to familiarize themselves with this trend in order to leverage student interest in relation to curriculum content. Some of this includes creating chat-room forums and groups to extend classroom discussion to posting assignments, tests and quizzes, through to assisting with homework outside of the classroom setting. Social network services are also being used to foster teacher-parent communication. These services make it possible and more convenient for parents to ask questions and voice concerns without having to meet face-to-face with their children's teachers. The advent of social networking platforms may also be impacting the way(s) in which learners engage with technology in general. The use of online social networks by school libraries is also increasingly prevalent and they are being used to communicate with potential library users, as well as extending the services provided by individual school libraries.

Social networks and their educational uses are of interest to many researchers. “Social networking sites, like much else on the internet, represent a moving target for researchers and policy makers.” Recent trends indicate that 47% of American adults use a social network. A national survey in 2009 found that 73% of online teenagers use SNS, which is an increase from 55% three years earlier. Recent studies have shown that social network services provide opportunities within professional education, curriculum education, and learning. However, there are constraints in this area.

Learning uses within education

Educators and advocates of new digital literacy are confident that social networking encourages the development of transferable, technical, and social skills of value in formal and informal learning. In a formal learning environment, goals or objectives are determined by an outside department or agency. Tweeting, instant messaging, or blogging enhances student involvement. Students who would not normally participate in class are more apt to partake through social network services.

Networking allows participants the opportunity for just-in-time learning and higher levels of engagement. The use of SNSs allow educators to enhance the prescribed curriculum. When learning experiences are infused into a website, students utilize everyday for fun, students realize that learning can and should be a part of everyday life. It does not have to be separate and unattached. Informal learning consists of the learner setting the goals and objectives. It has been claimed that media no longer just influence our culture. They are our culture. With such a high number of users between the ages of 13-18, a number of skills are developed. Participants hone technical skills in choosing to navigate through social networking services. This includes elementary items such as sending an instant message or updating a status. The developments of new media skills are paramount in helping youth navigate the digital world with confidence. Social networking services foster learning through "Participatory Culture." A participatory culture consists of a space that allows engagement, sharing, mentoring, and an opportunity for social interaction. Participants of social network services avail of this opportunity. Informal learning, in the forms of participatory and social learning online, is an excellent tool for teachers to sneak in material and ideas that students will identify with and therefore, in a secondary manner, students will

learn skills that would normally be taught in a formal setting in the more interesting and engaging environment of social learning.

Sites like Twitter provide students with the opportunity to converse and collaborate with others in real time. Social networking services provide a virtual "space" for learners.

Social Interaction

Social networking is a way for one person to meet up with other people on the net. People use social networking sites for meeting new friends, finding old friends, or locating people who have the same problems or interests they have, called niche networking.

More and more relationships and friendships are being formed online and then carried to an offline setting. The relationships which start online are much more likely to succeed.

Being able to meet some-one as a "friend" and see what common interests you share and how you have built up your friend base and "likes" you can truly see a fuller picture of the person you are talking with. Most sites are free instead of being paid based which allows younger people with stricter budgets to enjoy some of the same features as those of adults who are more likely to be able to afford pay based sites. While not the intended or original use for these social sites, a large area of their current function has stemmed from people wanting to meet other people in person and with the extremely busy schedules of most people, it is a fast, reliable and easy way in which to do so that costs you little time and money (if any). Users do not necessarily share with others the content which is of most interest to them, but rather that which projects a good impression of themselves.

4.4 RAILWAY RESERVATION SYSTEM

Journey by rail has its own charm and glitz. And, railway reservation in India is no more a hassle. You can go by online train reservation services or any outlet for that matter. Despite the coming up of cheap fairs in domestic airlines market, a substantial number of passengers and visitors yet journey by train. However, a train travel is both safe & comfortable and cheap. Indian people like to travel by train. A journey by train takes you to unearth the otherwise unexplored sites and mysteries of Mother India. As far as the railway reservation in India is concerned, there are myriad options at ones disposal. You can go by online train reservation system or any railway reservation booking outlet scattered everywhere.

However, Indian railway ticket reservation is no more a tedious job. Just lay your hands on any railway reservation booking outlet around you and make your way to the differing journey. With the onset of online railway reservation system things got much simpler for the passengers to book railway tickets online.

Indian railway is working incessantly to endow simply the best services to the passengers in India. Anyone with a system can have rail reservation instantly with no hassle. However, there are also systems of making railway reservation enquiry from virtually any place with your computer. This is how the whole system of booking railway tickets got easier in terms of accessibility and affordability. You are no longer required to sweat and fret over train ticket reservation in India. For any inquiry or for that matter any info you require regarding railway reservations just log onto the official site of Indian railways and you will have it.

Sitting at your home in front of a computer can give you all the relevant information on booking tickets in Indian railways. Booking any train on Indian Railways computerized passenger reservation system (PRS) network from any originating station or train passing through system station to any destination is that much easy nowadays. Be it about booking tickets, reservation enquiry, internet tickets (i-tickets), electronic tickets (e-tickets) or cancellation of tickets, things are just in place for the convenience of passengers.

Indian Railways-System and Network

Indian Railways is the world's second-largest railway, with 6,853 stations, 63,028 kilometers of track, 37,840 passenger coaches and 222,147 freight cars. Annually it carries some 4.83 billion passengers and 492 million tons of freight. Of the 11 million passengers who climb aboard one of 8,520 trains each day, about 550,000 have reserved accommodations. Their journeys can start in any part of India and end in any other part, with travel times as long as 48 hours and distances up to several thousand kilometers. The challenge is to provide a reservation system that can support such a huge scale of operations — regardless of whether it's measured by kilometers, passenger numbers, routing complexity, or simply the sheer scale of India.

The main challenges in front of the Indian railways are:

- Provide a reservation system that efficiently serves more than half a million people each day
- Ensure maximum uptime so reservation/ticketing/inquiry application is available 24 x 7
- Create a Web site that can accommodate more than one million hits per day
- Traveling on High Technology Indian Railways is one of the most advanced ministries in India, with an innovative and extensive IT environment — and a leading-edge reservation system powered by HP AlphaServer systems running the HP OpenVMS™ operating system and HP Reliable Transaction Router (RTR) middleware. Consider the scope of the operation.
- Good Technology means good service in 1986, the Ministry of Railways established the Centre for Railway Information Systems (CRIS) as an umbrella for all computer activities on Indian Railways. CRIS is responsible for designing, developing and implementing all major computer systems for the Railways. With its own R&D effort, CRIS has become a frontrunner in its field. One of CRIS's key technical achievements is a sophisticated reservation and ticketing application called Country-Wide Network for Enhanced Reservation and Ticketing (CONCERT), which runs on the OpenVMS AlphaServer platform. Centre for Railway Information Systems at Indian Railways, "OpenVMS is an extremely rugged and reliable operating system. Its built-in auditing feature provides us with excellent security." The primary challenge for CRIS is to provide an efficient passenger service by ensuring maximum uptime for its reservation/ticketing and inquiry application. The Railway must prepare charts that map passengers with their seats, and must post these charts outside each coach. CONCERT software enables the preparation of skeleton charts in advance for each train for the next three journey days. Indian Railway's current CONCERT application represents a steady progression of using the latest technologies available. In the mid-1980s, Indian Railways first computerized its reserved ticketing operation on VAX systems running VMS. This was done from five regional passenger reservation centres, each of which was a stand-alone site with its own local database. During the mid- to late 1990s, CRIS introduced CONCERT, which linked the five passenger reservation centres so

that reserved tickets from any station of Indian Railways could be issued to any other station from a single window. "CONCERT from CRIS has been able to improve the services to the passenger by offering single-window service to the passengers. RTR gives the user location transparency for the distributed database system. Thus, the reservation from one station to any other station can be given from a single window covering the round trip, which means passengers only have to stand in one queue. Indian railways perform various services to the passengers by using information technology, these are as follows:

i) Passenger Reservation System Solution

ii) Unreserved Ticketing System for Railways

iii) Mobile Ticketing

iv) Web Ticketing

v) Kiosk-based Ticketing

vi) Centralized (Hybrid) Ticketing System

vii) Time Table and Scheduling System

viii) Traffic Management Systems

ix) Passenger Information Display System

As more and more people turned to the Web to find information about various services, Indian Railways decided to provide information related to passenger reservations to the public over the Internet. In 2000, CRIS designed and implemented Indian Railways' own Web site, which receives a staggering 1.2 million hits per day. The site is hosted by CRIS and runs on the OpenVMS AlphaServer platform.

In 1985, CMC Software company (a subsidiary company of TATA Consultancy Services) developed IMPRESS, the railway reservation system based online transaction processing (OLTP), for the Indian Railways, which has been successfully operating it since 1987. Since then, however, the system has undergone a major change for networking all nodes in the railway network. The current software is CONCERT, implemented by the Centre for Railway Information Systems (CRIS).

The impact of IMPRESS / CONCERT on the system's users as well as on the Indian Railways has been tremendous. The benefits include substantial savings in transportation costs and in reservation time, telescopic fare benefits for cluster journeys, reduced malpractice and, above all, a modern, efficient and convenient system.

For the Railways, there is substantial reduction in cost per ticket issued, manpower savings (a 40 per cent increase in transactions handled per day), savings in space required, less strenuous work, higher productivity and fewer errors in fare computation, concession calculations, etc.

IMPRESS is being enhanced proactively, using state-of-the-art relational database management systems on open systems. The enhanced IMPRESS is built around an RDBMS core and supports full client-server architecture. It can also work on character-based terminals (used in the reservation and charting modules) in a host-based environment.

The application has been designed as an open distribution system, so that the data and transaction volume can be segregated between multiple host sites. Networking is an inherent feature of the application.

- The IMPRESS software can support both graphic user interface (GUI) and character-based terminals, which act as front-ends installed at the booking counters to cater to passenger requests.
- This software conforms to open standards. Hence, it can be interfaced to other applications like airline reservation systems, hotel reservation systems, etc., which are also based on open standards.
- The IMPRESS software is 'parametric' in terms of data and business rules, for fare computation, refund rules, cancellation, break journey rules, etc. Here, business rules are also kept as data items in the back-end repository instead of being part of the application logic. Therefore, the system can absorb changes in business rules immediately, without having to regenerate the object code.
- The application is secured against intrusion by two-level user authentications as the topmost guard. Below it, the data is secured from external access through multiple-level privileges. A data encryption facility is available across the WAN to prevent hacking.
- Ticketing for Indian Railways

Almost 14 million of the 15 million people whom the Railways transports every day travel on unreserved tickets. Handling them has been a huge problem. The Indian Railways plans to cover 943 more stations in 2006-07, and ensure that a total of 6,000 stations have UTS as of March 31, 2009. Unreserved tickets were earlier offered only two hours before the scheduled departure of trains. This not only caused inconvenience to passengers (as they wait in long queues to purchase their tickets) but it also affected the IR adversely in terms of loss in revenues, cumbersome reporting and in poor demand analysis. While some of the trains ran overcrowded, the others went partially vacant.

The implementation of UTS eliminated the earlier lapses in ticketing and helped the IR to substantially control overcrowding. The system comprised a network of terminals wherefrom the passengers could buy unreserved tickets for any journey 30 days in advance.

The Unreserved Ticketing System allowed advance planning and rational analysis of passenger demands for unreserved coaches. It also helped the IR to effectively monitor sale of tickets on various trains and regulate the train capacities to the fluctuating demands of passengers. With an aggressive use of leading hardware, data management and network technology, IR could successfully address the needs of the passengers of unreserved trains.

A network covering 63, 140 route kms (as on March 31, 2002), the Indian Railways traverse the length and breadth of the country. Even though the railways have been divided into zones for better management and functionality, the railway reservations or the process for booking train tickets is centrally computerised. Operating 14,444 trains daily, the IRCTC or Indian Railway Catering and Tourism Corporation Limited, ensures that train schedule and train timings remain prompt. In effect, Indian Railways is the largest railway system in the world to be functioning under a single management.

The most cost efficient mode of transport, Indian Railways enjoy preference over other public transport systems. Used extensively for passenger and freight transfer, Indian Railways proves itself the forerunner in the transport sector as the most affordable, convenient and well connected network. With thousands of railway stations across the country, superior safety standards, lower environmental hazard and relatively low train fare, Indian Railways is the first choice for transfer goods and commodities .

While the Indian Railway booking system has always been well organised, one immensely crucial step has been the launch of the IRCTC website (IRCTC online rail ticket reservation system is depicted in the following Figure 4), making all processes related to Indian Railways a breeze. Besides allowing online booking of rail ticket (s), it offers everything from railway map to railway ticket fare to timetable of train (s). A complete railway enquiry system is in place, with a click of the mouse allowing one to access / check the Indian Railways timetable, railway ticket availability and booking for any sector besides complete online train ticket reservations. Visitors to the website can use it not only for railway ticket reservation / booking but also to find out ticket availability and ticket confirmation or PNR status for any ticket on any train within the Indian Railways network

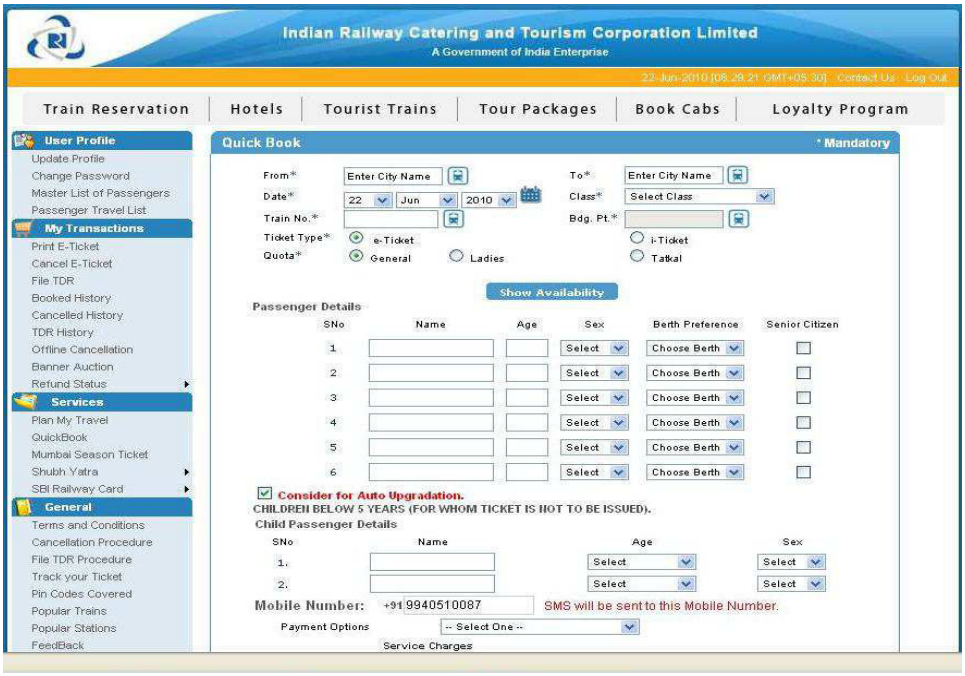


Figure 4: IRCTC online rail ticket reservation system

With various seating arrangements like AC 1 tier, AC 2 tier, AC 3 tier, AC Chair Car, Sleeper and General in most long distance trains and another range of seat availability in the metro trains within urban areas, the Indian Railways offers a plethora of choices. Travellers can take their pick and enjoy the convenience and affordability of this mode of transport.

The list of **major Indian Railway Zones** is as follows:

Central Indian Railways: This is the oldest of Indian railway zones and one of the largest of the 16 zones formed by Indian railways.

Eastern Indian Railways: The Eastern Railway (ER) zone is one of the important Indian Railway zones. With its headquarters in Kolkata, the Eastern Zone is divided further into four divisions namely Mald, Howrah, Asansol and Sealdah for better working.

Northern Indian Railways: It is one of the nine older zones of Indian Railways. New Delhi, the national capital of India serves as headquarters of this Indian Railway Zone

Southern Indian Railways: This is the first zone formed after India got liberated

from British Rule. Southern Indian Railway Zone was established on April 14, 1951 by combination of always of three states.

Western Indian Railways: This Indian railways zone is amongst the most hustling and lively rail networks of the country. The headquarters of Western Indian Railway is situated in Mumbai city.

☛ Check Your Progress 1

1. Select the right choice.
   a) Which of the following identifies a specific web page and its computer on the Web?
      A) Web site
      B) Web site address
      C) URL
      D) Domain Name

   ……

   ……

   b) What type of telecommunications hardware allows you to access the web?
      A) Browser
      B) Modem
      C) FTP protocol
      D) IRC

   ……

   ……

   c) The mail server as defined in the text uses the _____ protocol.
      A) HTTP
      B) FTP
      C) POP
      D) SMTP

   ……

   ……

2. Discuss standards of email protocol?

   ……

   ……

   ……

   ……

3. Write a note on social networking service?

   ……

   ……

   ……

   ……

4.5 INFORMATION SHARING

Information sharing describes "the exchange of data between various organizations, people and technologies".

There are several types of information sharing:

- Information shared by individuals (such as a video shared on Facebook or YouTube).
- Information shared by organizations (such as the RSS feed of an online weather report).
- Information shared between firmware/software (Such as the IP addresses of available network nodes or the availability of disk space).

The advent of wide distributed networks, intranets, cross-platform compatibility, application porting, and standardization of IP protocols have all facilitated the huge growth in global information sharing.

When it comes to personal information however, no matter how easy it is to port the actual data, there are laws in most countries prohibiting the sharing of personal data without explicit permission being granted. In the US and Europe it is a criminal offense to share any personal data about anyone without such explicit permission.

There is plenty of other information sharing that does not fall under the law and information sharing is increasing as more networks and organizations connect and information becomes easier to share across the internet.

Data was formerly frequently kept in silos and often not shared among other entities due to its proprietary, non-portable format or the inability to import/export data. Even simple items such as dates were stored in a whole range of different formats making the sharing of such a simple field a potential nightmare. The same applied to a whole range of data, and even if it was compatible it was often not possible to physically transfer the data from one platform to another.

Today these problems have all been coded out and information sharing is common between computer networks; information sharing has become especially prevalent due to social networking. These 21st century network models actively encourage the sharing of information across social networks.

Facebook has 750 million accounts, YouTube has over 400 million and the other social networking sites and applications have established between them a sharing network of over a billion people. In terms of information sharing this is a global proportion with almost 10% of the world's population sharing information across common networks regularly.

After the terrorist attacks of September 11th, information sharing became one of the United States government's goals in developing their resources to try to avert such atrocities. It was mandated among government agencies and departments that personnel create a methodology for regularly sharing relevant information. The US needed information sharing improvements to respond to various threats more effectively. The lesson was learned that when information is hoarded instead of shared, those needing it may not be able to react in a timely manner Using information sharing intelligently has been shown to be a more effective way to manage any organization; a government or a business.

Information sharing is crucial to many businesses, helping to promptly meet customer and client needs through customer relationship systems which share information about products and services and improve access to their customers.

Information sharing has also allowed easy availability of credit history details which helps consumers access more services. Consumers can have access to banking, financial and credit products from across the nation and even internationally where appropriate.

Hospitals sharing medical records (under stringent conditions) about people so that their medical personnel can make better decisions, is a good example of how organizations can share information for productive purposes rather than for social entertainment as with Facebook.

Overall, when used intelligently, information sharing is a useful way of lowering costs, improving overall accuracy of public data and allowing organizations and individuals alike to have access to information that they might need and entertainment that they want to experience.

4.6 ELECTRONIC GOVERNANCE

It is the use of a range of modern Information and Communication Technologies such as Internet, Local Area Networks, mobiles etc. by Government to improve the effectiveness, efficiency, service delivery and to promote democracy. E-Government can transform citizen service, provide access to information to empower citizens, enable their participation in government and enhance citizen economic and social opportunities, so that they can make better lives, for themselves and for the next generation. Governments are specialized institutions that contribute to governance. Representative governments seek and receive citizen support, but they also need the active cooperation of their public servants. Governance is the outcome of politics, policies, and programs.

The primary delivery models of e-Government can be divided into:

- Government-to-Citizen or Government-to-Consumer (G2C)
- Government-to-Business (G2B)
- Government-to-Government (G2G)
- Government-to-Employees (G2E)

Within each of these interaction domains, five kinds of activities take place:

1. Informing the citizen
2. Representing the citizen
3. Encouraging the citizen to vote
4. Consulting the citizen
5. Involving the citizen

The ultimate goal of the E-Government is to be able to offer an increased portfolio of public services to citizens in an efficient and cost effective manner. E-government allows for government transparency. Government transparency is important because it allows the public to be informed about what the government is working on as well as the policies they are trying to implement. Simple tasks may be easier to perform through electronic government access. Many changes, such as marital status or address changes can be a long process and take a lot of paper work for citizens. E-government allows these tasks to be performed efficiently with more convenience to

individuals. E-government is an easy way for the public to be more involved in political campaigns. It could increase voter awareness, which could lead to an increase in citizen participation in elections. It is convenient and cost-effective for businesses, and the public benefits by getting easy access to the most current information available without having to spend time, energy and money to get it.

E-government helps simplify processes and makes access to government information more easily accessible for public sector agencies and citizens. In addition to its simplicity, e-democracy services can reduce costs. The anticipated benefits of e-government include efficiency, improved services, better accessibility of public services, and more transparency and accountability. One goal of e-government will be greater citizen participation. Through the internet, people from all over the country can interact with politicians or public servants and make their voices heard. Blogging and interactive surveys will allow politicians or public servants to see the views of the people they represent on any given issue. Chat rooms can place citizens in real-time contact with elected officials, their offices or provide them with the means to replace them by interacting directly with public servants, allowing voters to have a direct impact and influence in their government. These technologies can create a more transparent government, allowing voters to immediately see how and why their representation in the capital is voting the way they are. This helps voters better decide who to vote for in the future or how to help the public servants become more productive. A government could theoretically move more towards a true democracy with the proper application of e-government. Government transparency will give insight to the public on how decisions are made and hold elected officials or public servants accountable for their actions. The public could become a direct and prominent influence in government legislature to some degree.

Characteristics of E-Governance

E-government allows citizens to interact with computers to achieve objectives at any time and any location, and eliminates the necessity for physical travel to government agents sitting behind desks and windows. Improved accounting and record keeping can be noted through computerization, and information and forms can be easily accessed, equaling quicker processing time. A network architecture of E-governance is shown in the Figure 5 given blow. It shows the different parties and agencies of governance can be connected using different networks devices and Internet.

Figure 5: Network Architecture of E-Governance. Source: http://www.mapit.gov.in

On the administrative side, access to help find or retrieve files and linked information can now be stored in databases versus hardcopies stored in various locations. Individuals with disabilities or conditions no longer have to be mobile to be active in government and can be in the comfort of their own homes.

The primary delivery models of e-Government are classified depending on who benefits. In the development of public sector or private sector portals and platforms, a system is created that benefits all constituents. Citizens needing to renew their vehicle registration have a convenient way to accomplish it while already engaged in meeting the regulatory inspection requirement. On behalf of a government partner, business provides what has traditionally, and solely, managed by government and can use this service to generate profit or attract new customers. Government agencies are relieved of the cost and complexity of having to process the transactions.
To develop these public sector portals or platforms, governments have the choice to internally develop and manage, outsource, or sign a self-funding contract. The self-funding model creates portals that pay for themselves through convenience fees for certain e-government transactions, known as self-funding portals.

Social networking is an emerging area for e-democracy. The social networking entry point is within the citizens' environment and the engagement is on the citizens' terms. Proponents of e-government perceive government use of social networks as a medium to help government act more like the public it serves. Examples can be found at almost every state government portal through Facebook, Twitter, and YouTube widgets. Government and its agents also have the opportunity to follow citizens to monitor satisfaction with services they receive. Through ListServs, RSS feeds, mobile messaging, micro-blogging services and blogs, government and its agencies can share information to citizens who share common interests and concerns. Government is also beginning to Twitter.

E-Government Forum

Since electronic government is new to everyone throughout the world, why couldn't an international matching system be developed so that anyone around the world can search by government function (e.g., Chief Information Officer); role (e.g., state or federal Department of Transportation); city, state, or country; type of technology used in development of the e-government system; sharing of best practices; licensing or buying existing software or platforms from the agency concerned; collaborating in the integrated development of a vertical (statewide) or horizontal (national or international) systems; or exchanging ideas. Theoretically, ALL government employees in the United States or throughout the world could be included. This could be easily achieved using modern technology and foster greater peace and understanding. Almost like an online Peace Corps.

☞ Check Your Progress 2

1. Select the right choice.
   a) Why do we store information? To refer to it:
      (A) When a decision is to be made or to confirm a fact.
      (B) If the original data is lost or corrupted.
      (C) Analyze to make some interpretation, explanation or resolution with it.
      (D) Distribute for further use
      (E) All of the above.

   ……

   ……

   b) Most of the electronic payment systems on internet use _____ to ensure confidentiality and security of the payment information.
      (A) Quantum Computing
      (B) Cryptography
      (C) Both (A) & (B)
      (D) None of the above

   ……

   ……

   c) An electronic check book device is a combination of
      (A) Hardware and a digital signature
      (B) Software and information about user
      (C) Secure hardware and software
      (D) None of the above

   ……

   ……

2. Discuss the Goal's of e-government?

   ……

   ……

   ……

   ……

3. Write a note on social networking service?

   ……

4.7 ONLINE PROCESSING AND COLLABORATIONS

The terms "**online**" and "**offline**" (also stylized as "**on-line**" and "**off-line**") have specific meanings in regard to computer technology and telecommunications. In general, "online" indicates a state of connectivity, while "offline" indicates a disconnected state. In common usage, "online" often refers to the Internet or the World Wide Web.

The concepts have however been extended from their computing and telecommunication meanings into the area of human interaction and conversation, such that even offline can be used in contrast to the common usage of online. For example, discussions taking place during a business meeting are "online", while issues that do not concern all participants of the meeting should be "taken offline" — continued outside of the meeting.

To be considered online, one of the following must apply to a device:

- Under the direct control of another device
- Under the direct control of the `system` with which it is associated
- Available for immediate use on demand by the system without human intervention
- Connected to a system, and is in operation
- Functional and ready for service

In contrast, a device that is offline meets none of these criteria (e.g., its main power source is disconnected or turned off, or it is off-power).

An online system is differs from an offline system in that templates are employed whenever the user inserts content into a web page for publication. These templates area said to be ‘on demand’ in that they are applied as per the user’s requirements.

Content is stored in a database which is a mandatory requirement, especially if the user wishes to view, amend or upgrade that content.
Template processing takes place as and when requested compared to offline systems which pre-process all of their content and apply their templates beforehand.

One advantage of an online processing is its extensibility which means greater functionality in the form of add-on’s such as image galleries, forum, wikis and blogs. This scalability is an important feature of these types of CMS.
Other advantages include:

- The user views content in real-time (published)
- The user is only shown relevant content, e.g. content which is out of date is withdrawn once it has reached it’s ‘sell by date’.
- The content can be accessed by multiple authors/editors in a variety of locations.
- Content can be maintained from any location which has internet access.

The downsides to this type of system include a slowing down of performance due to the fact that every time a user views a page, that page is retrieved from the database which slows down the processing speed exponentially.

This system requires a fast connection and an up to date browser especially in regard to content creation. Both of these are improving all the time but performance speeds are quicker on a local computer as used for an offline system.

Collaboration is working together to achieve a goal. It is a recursive process where two or more people or organizations work together to realize shared goals, (this is more than the intersection of common goals seen in co-operative ventures, but a deep, collective, determination to reach an identical objective. Most collaboration requires leadership, although the form of leadership can be social within a decentralized and egalitarian group. In particular, teams that work collaboratively can obtain greater resources, recognition and reward when facing competition for finite resources. Collaboration is also present in opposing goals exhibiting the notion of adversarial collaboration, though this is not a common case for using the word. Structured methods of collaboration encourage introspection of behavior and communication. These methods specifically aim to increase the success of teams as they engage in collaborative problem solving. Forms, rubrics, charts and graphs are useful in these situations to objectively document personal traits with the goal of improving performance in current and future projects.

Types of Collaborations:

Project management

Project Management developed from different fields of application including construction, engineering, and defense.

Learning community

A learning community is a group of people who share common emotions, values or beliefs, are actively engaged in learning together from each other, and by habituation. Such communities have become the template for a cohort-based, interdisciplinary approach to higher education. This may be based on an advanced kind of educational or 'pedagogical' design. the participants of learning community must feel some sense of loyalty and belonging to the group (membership) that drive their desire to keep working and helping others, also the things that the participant do in must affect what happened in the community, that means, an active and not just a reactive performance (influence). Besides a learning community must give the chance to the participants to meet particular needs (fulfillment) by expressing personal opinions, asking for help or specific information and share stories of events with particular issue included (emotional connections) emotional experiences.

Business

Collaboration in business can be found both inter- and intra-organization and ranges from the simplicity of a partnership and crowd funding to the complexity of a multinational corporation. Collaboration between team members allows for better communication within the organization and throughout the supply chains. It is a way of coordinating different ideas from numerous people to generate a wide variety of knowledge. Collaboration with a selected few firms as opposed to collaboration with a large number of different firms have been shown to positively impact firm performance and innovation outcomes. The recent improvement in technology has provided the world with high speed internet, wireless connection, and web-based collaboration tools like blogs, and wikis, and has as such created a "mass collaboration." People from all over the world are efficiently able to communicate and

share ideas through the internet, or even conferences, without any geographical barriers. The power of social networks it beginning to permeate into business culture where many collaborative use are being found including file sharing and knowledge transfer.

Education

Generally defined, an Educational Collaborative Partnership is ongoing involvement between schools and business/industry, unions, governments and community organiza tions. Educational Collaborative Partnerships are established by mutual agreement between two or more parties to work together on projects and activities that will enhance the quality of education for students while improving skills critical to success in the workplace.

Collaboration in Education- two or more co-equal individual voluntarily brings their knowledge and experience together by interacting toward a common goal in the best interest of students for the betterment of their education success. Students achieve team building and communication skills meeting many curricular standards. Students have the ability to practice real-world communication experiences. Students gain leadership through collaboration and empower peer to peer learning.

Technology

Due to the complexity of today's business environment, collaboration in technology encompasses a broad range of tools that enable groups of people to work together including social networking, instant messaging, team spaces, web sharing, audio conferencing, video, and telephony. Broadly defined, any technology that facilitates linking of two or more humans to work together can be considered a collaborative tool. Wikipedia, Blogs, even Twitter are collaborative tools. Many large companies are developing enterprise collaboration strategies and standardizing on a collaboration platform to allow their employees, customers and partners to intelligently connect and interact.

4.8 MOBILE APPLICATIONS

A **mobile application** (or **mobile app**) is a software application designed to run on smart phones, tablet computers and other mobile devices. They are available through application distribution platforms, which are typically operated by the owner of the mobile operating system, such as the Apple App Store, Google Play, Windows Phone Marketplace and BlackBerry App World. Some apps are free, while others have a price. Usually, they are downloaded from the platform to a target device, such as an iPhone, BlackBerry, Android phone or Windows Phone 7, but sometimes they can be downloaded to less mobile computers, such as laptops or desktops. For apps with a price, generally a percentage, 20-30%, goes to the distribution provider (such as iTunes), and the rest goes to the producer of the app.

The popularity of mobile applications has continued to rise, as their usage has become increasingly prevalent across mobile phone users.

For mobile applications, the fixed telephone system is not suitable. Mobile phones are currently in widespread use for voice and will soon be in widespread use for data. The first generation was analog, dominated by AMPS. The second generation was digital, with D-AMPS, GSM, and CDMA the major options. The third generation will be digital and based on broadband CDMA.

An alternative system for network access is the cable television system, which has gradually evolved from a community antenna to hybrid fiber coax. Potentially, it

offers very high bandwidth, but the actual bandwidth available in practice depends heavily on the number of other users currently active and what they are doing.

Mobile apps were originally offered for general productivity and information retrieval, including email, calendar, contacts, and stock market and weather information. However, public demand and the availability of developer tools drove rapid expansion into other categories, such as mobile games, factory automation, GPS and location-based services, banking, order-tracking, and ticket purchases. The explosion in number and variety of apps made discovery a challenge, which in turn led to the creation of a wide range of review, recommendation, and curation sources, including blogs, magazines, and dedicated online app-discovery services.

Mobile Apps are apps or services that can be pushed to a mobile device or downloaded and installed locally.

Classification of Mobil Apps:

- **Browser-based**: apps/services developed in a markup language
- **Native**: compiled applications (device has a runtime environment). Interactive apps such as downloadable games.
- **Hybrid**: the best of both worlds (a browser is needed for discovery)

Mobile application development is the process by which application software is developed for low-power handheld devices, such as personal digital assistants, enterprise digital assistants or mobile phones. These applications can be pre-installed on phones during manufacture, downloaded by customers from various mobile software distribution platforms, or delivered as web applications using server-side or client-side processing (e.g. JavaScript) to provide an "application-like" experience within a Web browser.

☛ Check Your Progress 3

1. Select the right choice.
   a) What do the letters GSM currently mean?
      (A) Global Special Mobile
      (B) Greater System's Mobile
      (C) Global Systems for Mobile Communications
      (D) None of the above!

   ……

   ……

   b) How is this problem solved in analogue cellular network?
      (I) Base Station continuously transmits the Mobile Identification Number (MIN) and received by the mobile phone.
      (II) Mobile phone continuously transmits the Signal Audio Tone (SAT) and received by the base station.
      (III) Overall size of the cluster is increased
      (IV) Increasing the bandwidth allocated to each user.

   ……

   ……

   c) Why does GSM use TDMA, as opposed to CDMA?

(I) When GSM was planned, CDMA was not approved as a multiple access system.
(II) TDMA is better than CDMA.
(III) CDMA is not really needed in Europe.
(IV) CDMA is too expensive to implement.

(A) I and III
(B) I only
(C) II and III
(D) I and IV

……

……

2. What is collaboration? Discuss in detail.

……

……

……

……

……

3. Discuss the advantages of Mobile applications?

……

……

……

……

……

4.9 SUMMARY

This completes our discussion on the introductory concepts of Network Applications. The internet architecture discussed in the unit, we have also discussed WWW and email formats in detail. The information given on various topics such as Information Sharing, Railway Reservation System, E-Governance, Social Networking, Online Processing etc. We have also includes the discussions based on latest trends and technology used for the internet applications (Mobile Applications) which helps the reads to keep your knowledge up to date. In addition to further readings and test their skills question answer sessions are included at the end of each sections. In the next block of this course, you will learn fundamental details for setting up a small local area network including wired and wireless setup. Next block will also cover the foundational details of network security protocols and wireless networking.

4.10 REFERENCES/FURTHER READING

1. Introduction to Data Communication & Networking, 3rd Edition, Behrouz Forouzan, Tata McGraw Hill.
2. Computer Networks, A. S. Tanenbaum 4th Edition, Practice Hall of India, New Delhi. 2003.

3. Douglas E. Comer, Internetworking with TCP/IP Vol.1: Principles, Protocols, and Architecture (4th Edition).

4. James F. Kurose, Computer Networking: A Top-Down Approach Featuring the Internet (3rd Edition).

5. Larry L. Peterson, Computer Networks: A Systems Approach, 3rd Edition (The Morgan Kaufmann Series in Networking).

6. www.wikipedia.org

7. W. Richard Stevens, The Protocols (TCP/IP Illustrated, Volume 1).

8. William Stallings, Data and Computer Communications, Seventh Edition.

4.11 SOLUTIONS/ANSWERS

☞ Check Your Progress 1

1. a) (D)
   b) (B)
   c) (C)

2. The main standards that relate to the protocols of email transmission and reception are:

   i) **S**imple Mail Transfer Protocol (SMTP) - which is used with the TCP/IP protocol suite? It has traditionally been limited to the text based electronic messages.

   ii) **M**ultipurpose Internet Mail Extension (MIME) - Which allows the transmission and reception of mail that contains various types of data, such as speech, images, and motion video? It is a newer standard than STMP and uses much of its basic protocol.

   iii) **S**/MIME (Secure MIME). RSA Data security created S/MIME which supports encrypted e-mail transfer and digitally signed electronic mail.

3. A **social networking service** is an online service, platform, or site that focuses on facilitating the building of social networks or social relations among people who, for example, share interests, activities, backgrounds, or real-life connections. A social network service consists of a representation of each user (often a profile), his/her social links, and a variety of additional services. Most social network services are web-based and provide means for users to interact over the Internet, such as e-mail and instant messaging. Online community services are sometimes considered as a social network service, though in a broader sense, social network service usually means an individual-centered service where as online community services are group-centered. Social networking sites allow users to share ideas, activities, events, and interests within their individual networks.

☞ Check Your Progress 2

1. a) (D)
   b) (B)
   c) (B)

2. **Goal's of e-government**

One goal of e-government will be greater citizen participation. Through the internet, people from all over the country can interact with politicians or public servants and make their voices heard. Blogging and interactive surveys will allow politicians or public servants to see the views of the people they represent on any given issue. Chat rooms can place citizens in real-time contact with elected officials, their offices or provide them with the means to replace them by interacting directly with public servants, allowing voters to have a direct impact and influence in their government. These technologies can create a more transparent government, allowing voters to immediately see how and why their representation in the capital is voting the way they are. This helps voters better decide who to vote for in the future or how to help the public servants become more productive. A government could theoretically move more towards a true democracy with the proper application of e-government.

3. A **social networking service** is an online service, platform, or site that focuses on facilitating the building of social networks or social relations among people who, for example, share interests, activities, backgrounds, or real-life connections. A social network service consists of a representation of each user (often a profile), his/her social links, and a variety of additional services. Most social network services are web-based and provide means for users to interact over the Internet, such as e-mail and instant messaging. Online community services are sometimes considered as a social network service, though in a broader sense, social network service usually means an individual-centered service where as online community services are group-centered. Social networking sites allow users to share ideas, activities, events, and interests within their individual networks.

☞ Check Your Progress 3

1. a) (C)
   b) (D)
   c) (A)

2. **Collaboration** is working together to achieve a goal. It is a recursive process where two or more people or organizations work together to realize shared goals, (this is more than the intersection of common goals seen in co-operative ventures, but a deep, collective, determination to reach an identical objective. Most collaboration requires leadership, although the form of leadership can be social within a decentralized and egalitarian group. In particular, teams that work collaboratively can obtain greater resources, recognition and reward when facing competition for finite resources. Collaboration is also present in opposing goals exhibiting the notion of adversarial collaboration, though this is not a common case for using the word. Structured methods of collaboration encourage introspection of behavior and communication. These methods specifically aim to increase the success of teams as they engage in collaborative problem solving. Forms, rubrics, charts and graphs are useful in these situations to objectively document personal traits with the goal of improving performance in current and future projects.

3. Mobile Apps are internet applications designed to run on smartphones and other mobile devices. mobile applications help users by connecting them to Internet services more commonly accessed on desktop or notebook computers. While opportunities abound, we have identified three advantages of using mobile apps for your business: speed, volume of information, and advertising.